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

LUANA CECÍLIA MEIRELES DA SILVA

DIAGNÓSTICO EM MODELOS DE REGRESSÃO SIMPLEX

Recife

2019

LUANA CECÍLIA MEIRELES DA SILVA

DIAGNÓSTICO EM MODELOS DE REGRESSÃO SIMPLEX

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

Área de Concentração: Estatística Aplicada.

Orientadora: Prof.ª Dr.ª Patrícia Leone Espinheira Ospina
Coorientador: Prof. Dr. Francisco Cribari Neto

Recife
2019

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

S586d Silva, Luana Cecília Meireles da
Diagnóstico em modelos de regressão simplex / Luana Cecília Meireles da Silva. – 2019.
120 f.: il., fig., tab.

Orientadora: Patrícia Leone Espinheira Ospina.
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN, Estatística, Recife, 2019.
Inclui referências.

1. Estatística. 2. Modelos de regressão. I. Ospina, Patrícia Leone Espinheira (orientadora). II. Título.

310 CDD (23. ed.) UFPE- MEI 2019-030

LUANA CECÍLIA MEIRELES DA SILVA

DIAGNÓSTICO EM MODELOS DE REGRESSÃO SIMPLEX

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

Aprovada em: 22 de fevereiro de 2019.

BANCA EXAMINADORA

Prof.(ª) Patrícia Leone Espinheira Ospina (Orientadora)
Universidade Federal de Pernambuco

Prof.(ª) Fernanda De Bastiani (Examinador Interno)
Universidade Federal de Pernambuco

Prof.(º) Roberto Ferreira Manghi (Examinador Interno)
Universidade Federal de Pernambuco

Prof.(ª) Tarciana Liberal Pereira (Examinador Externo)
Universidade Federal da Paraíba

Prof.(ª) Michelli Karinne Barros da Silva (Examinador Externo)
Universidade Federal de Campina Grande

Este trabalho é carinhosamente dedicado

aos meus pais, João e Luzimar,

e ao meu marido, Rafael.

AGRADECIMENTOS

À Deus, por sua infinita bondade e por todas as bençãos concedidas.

Aos meus pais, João e Luzimar, pelo amor incondicional, pelo apoio, incentivo e pelos valiosos conselhos. Vocês são os principais responsáveis por essa conquista.

Ao meu esposo, Rafael, pelo amor, carinho, cuidado, incentivo e paciência durante esse período. Muito obrigada por tudo.

À professora Patrícia Espinheira, por acima de tudo ser uma grande amiga. Obrigada por tornar mais leve essa caminhada, pela excelente orientação, pelo carinho, incentivo, dedicação e confiança.

Ao professor Cribari, pela dedicação, confiança e pelos ensinamentos compartilhados.

A toda minha família. Em especial, ao meu irmão, Jonas, à minha cunhada, Carina, e meus sobrinhos, por todo amor e incentivo.

À família de Rafael, que se tornou minha família, por estarem sempre presentes em minha vida.

Aos professores da Universidade Federal de Pernambuco que contribuíram para a minha formação, em especial aos professores Patrícia Leone Espinheira, Francisco Cribari-Neto, Raydonal Ospina, Klaus Vasconcellos, Gauss Cordeiro, Audrey Cysneiros, Francisco Cysneiros. Obrigada por todo ensinamento compartilhado.

Aos professores do Departamento de Estatística da Universidade Federal da Paraíba, pelo incentivo e pela torcida.

À minha turma de doutorado, Wênia, Neto, Marley, Thiago, Renata, Fernanda, Fábio, Eberson, Fernando, que se tornaram grandes amigos. Obrigada pelos momentos de estudo e pelos momentos de descontração.

Aos amigos da "República Paraibana": Jodavid, Ramon, Camila, Pedro e Andreza, pelo maravilhoso convívio e pelos momentos divertidos.

Aos amigos Wanessa, Alisson e Antônio, por estarem sempre torcendo por mim, em especial, agradeço à Wanessa pelos momentos compartilhados, pelo carinho, incentivo e pela torcida.

Aos demais amigos da Pós-graduação em Estatística da Universidade Federal de Per-

nambuco.

Aos funcionários do Departamento de Estatística da Estatística da Universidade Federal de Pernambuco, em especial à Valéria Bittencourt, pelo carinho, cuidado e dedicação.

Aos participantes da banca examinadora, desde já agradeço pelas valiosas sugestões.

À CAPES, pelo apoio financeiro.

“A persistência é o caminho do êxito.”

(CHAPLIN, 1997, p.118)

RESUMO

Em muitas situações práticas existe a necessidade de modelar dados no intervalo $(0, 1)$. Esses dados podem ser interpretados como taxas ou proporções e, em geral, apresentam assimetria e heteroscedasticidade, não satisfazendo as suposições do modelo de regressão linear clássico. Diversos modelos de regressão estão sendo estudados com esse objetivo. Por exemplo, o modelo de regressão beta (FERRARI & CRIBARI-NETO, 2004), o modelo de regressão Kumaraswamy (MITNIK & BAEK, 2013), o modelo Johnson S_b (LEMONTE & BAZAN, 2016), o modelo gama unitário (MOUSA et al., 2013), o modelo de regressão simplex (BARNDORFF-NIELSEN & JøRGENSEN, 1991), entre outros. O modelo de regressão simplex, em especial, faz parte dos modelos de dispersão (JøRGENSEN, 1997) que estendem os modelos lineares generalizados (MCCULLAGH & NELDER, 1989). Uma fase muito importante para a escolha de um modelo de regressão é a análise de diagnóstico, visto que todo o desempenho inferencial é baseado no modelo selecionado. Nessa fase, os resíduos desempenham um papel crucial para a verificação da adequação do modelo. A estatística $PRESS$ pode ser utilizada como uma indicação do poder preditivo do modelo e o coeficiente de predição, P^2, para selecionar modelos com a perspectiva de predição. Nesta tese propomos um novo resíduo para a classe de modelos de regressão simplex não linear. Propomos a estatística $PRESS$ e o coeficiente de predição P^2 baseados no resíduo ponderado e no novo resíduo. Além disso, avaliamos algumas medidas de qualidade de ajuste (BAYER & CRIBARI-NETO, 2017). Apresentamos resultados de simulações de Monte Carlo para o novo resíduo e paras as estatísticas de predição e de qualidade de ajuste sob diversos cenários. Por fim, apresentamos e discutimos várias aplicações à dados reais.

Palavras-chave: Coeficientes de predição. Distribuição simplex. Modelo de regressão simplex. $PRESS$. Resíduo combinado.

ABSTRACT

In many practical situations there is a need to model data in the interval $(0, 1)$. These data can be interpreted as rates or proportions and, in general, have asymmetry and heteroscedasticity, not satisfying the assumptions of the classical linear regression model. Several regression models are being studied for this purpose. For example, the beta regression model (FERRARI & CRIBARI-NETO, 2004), Kumaraswamy regression model (MITNIK & BAEK, 2013), Johnson S_b regression model (LEMONTE & BAZAN, 2016), unit gamma model (MOUSA et al., 2013), simplex regression model (BARNDORFF-NIELSEN & JøRGENSEN, 1991), among others. The simplex regression model, in particular, is part of the dispersion models (JøRGENSEN, 1997) that extend generalized linear models (MCCULLAGH & NELDER, 1989). A very important phase for choosing a regression model is the diagnostic analysis, since all inferential performance is based on the selected model. At this stage, the residuals plays a crucial role in verifying the adequacy of the model. The $PRESS$ statistic can be used as an indication of the predictive power of the model and the prediction coeficiente, P^2 , to select models from the prediction perspective. In this thesis we propose a new residual for the class of nonlinear simplex regression models. We propose the $PRESS$ statistic and the prediction coefficient P^2 based on the weighted residual and the new residual. In addition, we evaluated some measures of goodness of fit (BAYER & CRIBARI-NETO, 2017). We present results of Monte Carlo simulations for the new residual and for the prediction and fit quality statistics under different scenarios. Finally, we present and discuss various applications to real data.

Keywords: Prediction coefficients. Simplex distribution. Simplex regression model. $PRESS$. Combined residual.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

SUMÁRIO

1 INTRODUÇÃO

Em muitas situações práticas existe a necessidade de modelar dados no intervalo (0,1). Esses dados podem ser interpretados como taxas ou proporções e, em geral, apresentam assimetria e heteroscedasticidade, não satisfazendo as suposições do modelo linear clássico. Portanto, é de grande importância levar em consideração o comportamento dos dados para manipulá-los adequadamente. Diversos autores propuseram modelos de regressão considerando distribuições mais adequadas para variáveis respostas com tais características. Um dos modelos mais estudados atualmente é o modelo de regressão beta proposto por diversos autores como por exemplo, PAOLINO (2001), KIESCHNICK & MCCULLOUGH (2003), FERRARI & CRIBARI-NETO (2004), VASCONCELLOS & CRIBARI-NETO (2005), SMITHSON & VERKULIEN (2006), SIMAS et al. (2010) e ROCHA & SIMAS (2011) e entre outros. Em FERRARI & CRIBARI-NETO (2004), os autores fazem uma reparametrização de modo a permitir a modelagem da média da variável resposta envolvendo um parâmetro de dispersão. Além disso, esse modelo é muito utilizado pois está implementado no *software* estatístico `R` (`http://www.r-project.org`) por meio do pacote `betareg`. SIMAS et al. (2010) estende esse modelo ao introduzir uma estrutura de regressão não linear para a média e para o parâmetro de dispersão.

No entanto, vários modelos alternativos têm sido estudados tais como, modelo de regressão Johnson S_b (LEMONTE & BAZAN, 2016), modelo de regressão gama unitária (MOUSA et al., 2013), modelo de regressão Kumarashuamyn (MITNIK & BAEK, 2013), modelo de regressão simplex)(BARNDORFF-NIELSEN & JøRGENSEN, 1991), entre outros.

O modelo de regressão simplex faz parte dos modelos de dispersão (JøRGENSEN, 1997) que estendem os modelos lineares generalizados (MCCULLAGH & NELDER, 1989). SONG & TAN (2000) desenvolvem o modelo marginal para dados de proporções longitudinais com dispersão constante utilizando equações de estimação generalizadas (EEG) e SONG et al. (2004) generalizam esse modelo introduzindo uma estrutura de regressão para o parâmetro de dispersão e correlação, além da média populacional. SONG (2009) apresenta uma revisão da teoria de análise de regressão baseado nos modelos de disper-

são e utilizam o método de máxima verossimilhança para a estimação dos parâmetros. MIYASHIRO (2008) propôs o modelo de regressão simplex com dispersão constante similar ao modelo de regressão beta (FERRARI & CRIBARI-NETO, 2004) e define o resíduo baseado no processo iterativo escore de Fisher para a estimação de β. Mais recentemente, ESPINHEIRA & SILVA (2018) propôs a classe de modelos de regressão simplex não linear. Os autores utilizam o método de máxima verossimilhança para estimar os parâmetros do modelo e derivam as quantidades de influência local para o modelo.

A análise de diagnóstico é uma fase muito importante para a escolha de um modelo de regressão, visto que todas as conclusões sobre a relação entre as variáveis se baseiam em um modelo postulado. Dentre as várias formas de avaliarmos a qualidade do ajuste de um modelo ajustado aos dados reais temos a análise de resíduos, que visa identificar discrepâncias entre os valores ajustados a partir do modelo e os valores observados. Na classe de modelos para dados que se distribuem de forma contínua no intervalo $(0, 1)$ muitos resultados de diagnóstico foram desenvolvidos em especial para a classe de modelos de regressão beta (FERRARI & CRIBARI-NETO, 2004). Em ESPINHEIRA et al. (2008b) são desenvolvidos resíduos e medidas de influência local para a classe de modelos de regressão beta linear. Além disso, ESPINHEIRA et al. (2017) apresentam um novo resíduo para a classe de modelos de regressão beta não linear e LEMONTE & BAZAN (2016) também propuseram resíduos e derivaram as quantidades de influência local para a classe de modelos Johnson S_b. Recentemente, ESPINHEIRA & SILVA (2018) desenvolveram um resíduo baseado no processo iterativo escore de Fisher para β e medidas de influência local considerando diversos esquemas de perturbação para os modelos de regressão simplex em que tanto a média quanto a dispersão podem ser explicados com base em preditores não lineares.

Além disso, outras medidas são de grande importância para a seleção de modelos tais como, seleção *stepwise forward* ou eliminação *backward* (DRAPER & SMITH, 1981), seleção *best subset* (GARSIDE, 1965), Cp de Mallows (MALLOWS, 1973), Validação cruzada (STONE, 1974), critério de informação de Akaike (AIC) (AKAIKE, 1973), critério bayesiano de Schwarz (SBC) (SCHWARZ, 1978), soma de quadrados de resíduos (SQR), e várias funções do SQR tais como R^2 e o R^2 ajustado (BAYER & CRIBARI-NETO, 2017). No entanto, tais medidas não fornecem informação sobre o poder preditivo do modelo.

Com esse objetivo, ALLEN (1971) porpôs o critério $PRESS$ (Predictive Residual Sum of Squares) que pode ser utilizado como uma indicação do poder preditivo do modelo. O cálculo da estatística $PRESS$ consiste no ajuste do modelo, repetidamente, deixando de fora uma observação de cada vez. Em cada repetição o modelo é utilizado para predizer a observação que ficou de fora. Similarmente a abordagem do R^2, WOLD (1982) propôs o P^2, um coeficiente de predição baseado na $PRESS$. A estatística P^2 pode ser utilizada para selecionar modelos com a perspectiva de predição.

Esta tese tem como objetivo propor um novo resíduo para a classe de modelos de regressão simplex. O novo resíduo é facilmente calculado e é baseado nos processos iterativos escore de Fisher para a estimação dos parâmetros que modelam a média e a dispersão, tal como em ESPINHEIRA et al. (2017) para o modelo de regressão beta. Além disso, também temos por objetivo propor versões das estatísticas $PRESS$ e dos coeficientes de predição P^2 para o modelo de regressão simplex linear e não linear e avaliar o comportamento de medidas de qualidade de ajuste através de simulações de Monte Carlo e aplicações a dados reais.

1.1 Organização da Tese

O presente trabalho é constituído por cinco capítulos. No Capítulo 2, apresentamos o modelo de regressão simplex linear e não linear, assim como a função escore, matriz de informação de Fisher e estimação dos parâmetros. No Capítulo 3 definimos o resíduo combinado para o modelo de regressão simplex não linear, apresentamos simulações para avaliar sua distribuição empírica e apresentamos aplicações a dados reais. No Capítulo 4 definimos as estatísticas $PRESS$ e P^2 para o modelo simplex e definimos a distância de Cook. Além disso, avaliamos o desempenho destas medidas através de simulações de Monte Carlo considerando diversos cenários, assim como algumas medidas de qualidade de ajuste e apresentamos algumas aplicações a dados reais. Por fim, no Capítulo 5, as principais conclusões desse trabalho são discutidas.

1.2 Suporte Computacional

As avaliações numéricas apresentadas nesta tese foram realizadas utilizando a linguagem de programação Ox para o sistema operacional Windows (DOORNIK, 2001). Ox é uma linguagem matricial de programação com sintaxe semelhante a C e C++, possuindo uma ampla biblioteca numérica. Foi criada e desenvolvida por Doornik em 1994 na Universidade de Oxford (Inglaterra) e é distribuída gratuitamente para uso acadêmico no site <http://www.doornik.com>. Mais informações sobre essa linguagem podem ser encontrados em DOORNIK (2001), DOORNIK (2006), DOORNIK & OOMS (2006) e DOORNIK (2013).

Para a análise gráfica utilizamos o ambiente computacional R em sua versão 3.4.2 para o sistema operacional Windows. R é uma ferramenta de programação, análise de dados e geração de gráficos que recebe contribuições de pessoas em todo o mundo e encontra-se disponível gratuitamente em <http://www.r-project.org>. Para mais detalhes ver VANABLES & RIPLEY (2002), DALGAARD (2002) e VENABLES et al. (2018).

Essa tese foi digitada através do sistema tipográfico LaTeX, desenvolvido por Leslie Lamport na década de 1980. Consiste em uma série de macros ou rotinas do sistema TeX, criado por Donald Knuth na Universidade de Stanford, que facilitam o desenvolvimento da edição de textos (KNUTH, 1986). Mais detalhes podem ser encontrados em LAMPORT (1994) ou através do site <http://www.tex.ac.uk/CTAN/latex>.

2 MODELO DE REGRESSÃO SIMPLEX

2.1 Introdução

A distribuição simplex faz parte da classe de modelos de dispersão e é bastante útil para modelar dados contínuos no intervalo unitário padrão. Nesta seção, introduziremos o modelo de regressão simplex linear e não linear. Apresentaremos a função escore, matriz de informação de Fisher e a estimação dos parâmetros, assim como algumas propriedades da distribuição Simplex.

2.2 Distribuição Simplex

A distribuição simplex é utilizada para modelar dados restritos ao intervalo $(0, 1)$, que podem ser interpretados como taxas ou proporções, por exemplo. A distribuição foi proposta por BARNDORFF-NIELSEN & JøRGENSEN (1991) e faz parte da classe dos modelos de dispersão introduzida por JøRGENSEN (1997), que estende os modelos lineares generalizados (NELDER & WEDDERBURN, 1972). Por definição, um modelo de dispersão $\mathbf{DM}(\mu; \sigma^2)$ com parâmetro de locação μ e parâmetro de dispersão σ^2 é uma família de distribuições que possui função densidade de probabilidade com a seguinte forma

$$p(y; \mu, \sigma^2) = a(y; \sigma^2)\exp\left\{-\frac{1}{2\sigma^2}d(y; \mu)\right\}, \quad y \in C \tag{2.1}$$

em que C é o suporte da distribuição, $\mathbb{E}(Y) = \mu \in \Omega$, Ω é o espaço paramétrico de μ, $\sigma^2 > 0$ e $a \geq 0$ é um termo de normalização adequado que é independente de μ. A função bivariada $d(\cdot; \cdot)$ é chamada desvio unitário definido em $(y, \mu) \in C \times \Omega$ e satisfaz duas propriedades:

- quando o valor observado y é igual ao valor esperado μ, o desvio unitário é igual a zero,

$$d(y; y) = 0, \quad \forall y \in \Omega;$$

- quando o valor observado y é diferente do valor esperado μ, o desvio unitário é positivo,

$$d(y;\mu) > 0, \quad \forall y \neq \mu.$$

Além disso, o desvio unitário é chamado regular se a função $d(\cdot;\cdot)$ for duas vezes continuamente diferenciável com relação a (y,μ) em $\Omega \times \Omega$ e satisfazer

$$\frac{\partial^2 d}{\partial \mu^2}(y;\mu)\bigg|_{y=\mu} > 0, \forall \mu \in \Omega.$$

Para o desvio unitário regular, a função de variância é uma função $V : \Omega \to (0,\infty)$, definida por

$$V(\mu) = \frac{2}{\frac{\partial^2 d}{\partial \mu^2}(y;\mu)\big|_{y=\mu}}.$$

Diversas distribuições, discretas e contínuas, pertencem aos modelos de dispersão, a exemplo da Normal, Poisson, Binomial, Binomial Negativa, Gamma, Normal Inversa, von Mises, Simplex, entre outras.

Em especial, quando uma variável aleatória y segue distribuição simplex com parâmetros $\mu \in (0,1)$ e $\sigma^2 > 0$, a função densidade em (2.1) possui a seguinte forma

$$p(y;\mu,\sigma^2) = \{2\pi\sigma^2[y(1-y)]^3\}^{-1/2}\exp\left\{-\frac{1}{2\sigma^2}d(y;\mu)\right\}, \ 0 < y < 1, \tag{2.2}$$

em que $d(y;\mu)$ é o desvio unitário dado por

$$d(y;\mu) = \frac{(y-\mu)^2}{y(1-y)\mu^2(1-\mu)^2}, \tag{2.3}$$

que é um desvio regular e consequentemente $V(\mu) = \mu^3(1-\mu)^3$. Utilizando alguns resultados encontrados em JøRGENSEN (1997) temos que $\mathbb{E}(Y) = \mu$ e

$$\mathrm{Var}(Y) = \mu(1-\mu) - \sqrt{\frac{1}{2\sigma^2}}\exp\left\{\frac{1}{\sigma^2\mu^2(1-\mu)^2}\right\}\Gamma\left\{\frac{1}{2}, \frac{1}{2\sigma^2\mu^2(1-\mu)^2}\right\}.$$

em que $\Gamma(a,b)$ é a função gama incompleta definida por $\Gamma(a,b) = \int_b^\infty x^{a-1}\mathrm{e}^{-x}dx$.

A Figura 1 apresenta algumas densidades da distribuição simplex para diferentes valores de (μ,σ^2). Pode-se observar que a densidade pode apresentar diversas formas simétricas e assimétricas, entre elas forma de banheira e bimodal.

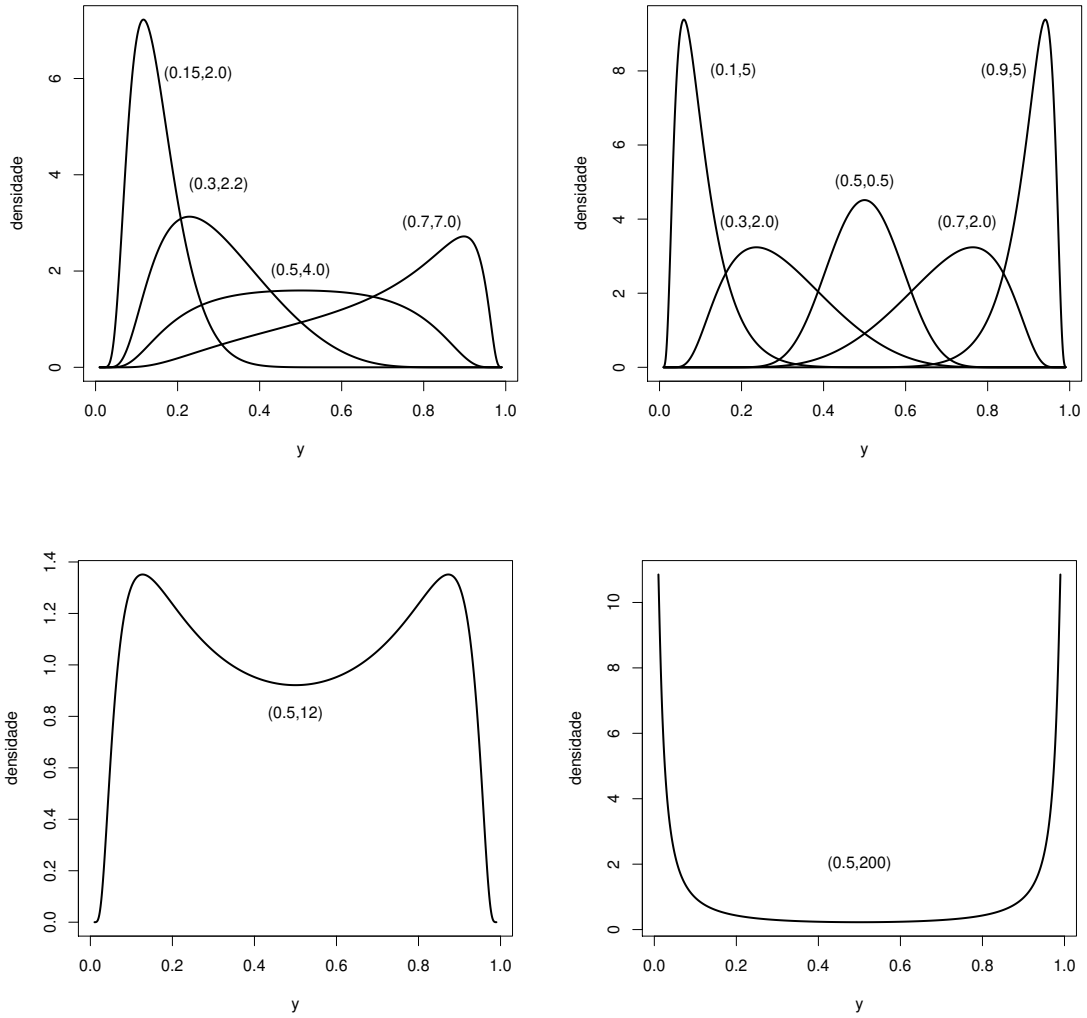

Figura 1 – Densidade da distribuição simplex para diferentes valores de (μ, σ^2).

2.2.1 Propriedades da distribuição simplex

Seja y uma variável aleatória que segue distribuição simplex com média μ e parâmetro de dispersão σ^2. Então, são válidas as seguintes propriedades

(a) $\mathbb{E}[d(y;\mu)] = \sigma^2$;

(b) $\mathbb{E}[(y-\mu)d'(y;\mu)] = -2\sigma^2$;

(c) $\mathbb{E}[(y-\mu)d(y;\mu)] = 0$;

(d) $\mathbb{E}[(y-\mu)d^2(y;\mu)] = 0$;

(e) $\frac{1}{2}\mathbb{E}[d''(y;\mu)] = \frac{3\sigma^2}{\mu(1-\mu)} + \frac{1}{\mu^3(1-\mu)^3}$;

(f) $\text{Var}[d(y;\mu)] = 2(\sigma^2)^2$;

(g) $\mathbb{E}[d'(y;\mu)] = 0$;

em que $d'(y;\mu) = \partial d(y;\mu)/\partial\mu$ e $d''(y;\mu) = \partial^2 d(y;\mu)/\partial\mu^2$. Para mais detalhes sobre as propriedades da distribuição simplex, ver SONG & TAN (2000), SONG et al. (2004) e CLOTILDE (2016).

2.3 Modelo de Regressão Simplex Linear

Nesta seção, apresentaremos o modelo de regressão simplex na qual a média da variável resposta e o parâmetro de dispersão σ^2 estão relacionados às covariáveis através de preditores lineares.

Sejam $y_1, \ldots, y_n$ variáveis aleatórias independentes, em que cada y_t, $t = 1, \ldots, n$, segue distribuição simplex cuja densidade é dada em (2.2), com média μ_t e parâmetro de dispersão σ_t^2. O modelo de regressão simplex linear assume que a média e o parâmetro de dispersão satisfazem as seguintes relações funcionais

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

em que $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)^\top$ e $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_q)^\top$ são vetores de parâmetros de regressão desconhecidos, $\boldsymbol{\beta} \in \mathbb{R}^k$ e $\boldsymbol{\gamma} \in \mathbb{R}^q$, $k + q < n$, η_{1t} e η_{2t} são preditores lineares, e $x_{t1}, \ldots, x_{tk}, z_{t1}, \ldots, z_{tq}$ são observações em k e q covariáveis conhecidas. Ambas, $g : (0,1) \to \mathbb{R}$ e $h : (0,\infty) \to \mathbb{R}$ são funções de ligação estritamente monótonas e duas vezes diferenciáveis. Diferentes funções de ligação podem ser utilizadas em g e h. Para μ, temos a função Logit $g(\mu) = \log(\mu/1-\mu)$, Probit $g(\mu) = \Phi^{-1}$, em que $\Phi(\cdot)$ é a função acumulada da distribuição normal, C-log-log $g(\mu) = \log(-\log(1-\mu))$, Log-log $g(\mu) = -\log(-\log(\mu))$, entre outras. Em particular, quando a função de ligação Logit é utilizada, os parâmetros de regressão podem ser interpretados em termos da razão de chances (*odds ratio*). Para σ^2 podemos utilizar função logarítmica $h(\sigma^2) = \log(\sigma^2)$, a função identidade $h(\sigma^2) = \sigma^2$ e a função raiz quadrada $h(\sigma^2) = \sqrt{\sigma^2}$. Para mais detalhes, ver MCCULLAGH & NELDER (1989) e ATKINSON (1985).

2.3.1 Função Escore, Matriz de Informação de Fisher e Estimação dos Parâmetros

Baseado em (2.2), o logaritmo da função da função de verossimilhança é dado por

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{t=1}^{n} \ell_t(\mu_t, \sigma_t^2),$$

em que

$$\ell_t(\mu_t, \sigma_t^2) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{3}{2}\log[y_t(1-y_t)] - \frac{1}{2\sigma_t^2} d(y_t; \mu_t).$$

Os componentes do vetor escore $U_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})$, são obtidos através da diferenciação do logaritmo da função verossimilhança com relação a β_i, $i = 1, \ldots, k$, e são dados por

$$\frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \beta_i} = \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i},$$

em que $d\mu_t/d\eta_{1t} = 1/g'(\mu_t)$, $\partial \eta_{1t}/\partial \beta_i = x_{ti}$ e

$$\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t} = -\frac{1}{2\sigma_t^2} d'(y_t; \mu_t), \tag{2.4}$$

com

$$d'(y_t; \mu_t) = -\frac{2(y_t - \mu_t)}{\mu_t(1-\mu_t)} \left[d(y_t; \mu_t) + \frac{1}{\mu_t^2(1-\mu_t)^2} \right]. \tag{2.5}$$

Matricialmente, a função escore para o vetor $\boldsymbol{\beta}$ pode ser expressa da seguinte forma:

$$U_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma}) = X^\top \Sigma T U (\boldsymbol{y} - \boldsymbol{\mu}),$$

em que X é uma matriz $n \times k$ cuja t-ésima linha é $x_t = (x_{t1}, x_{t2}, \ldots, x_{tk})$, $\Sigma = \text{diag}\{1/\sigma_1^2, \ldots, 1/\sigma_n^2\}$, e T e U são dadas, respectivamente, por

$$T = \text{diag}\{1/g'(\mu_1), \ldots, 1/g'(\mu_n)\} \quad \text{e} \quad U = \text{diag}\{u_1, \ldots, u_n\}, \tag{2.6}$$

em que

$$u_t = \frac{1}{\mu_t(1-\mu_t)} \left[d(y_t; \mu_t) + \frac{1}{\mu_t^2(1-\mu_t)^2} \right]. \tag{2.7}$$

Analogamente, os componentes do vetor escore $U_\gamma(\boldsymbol{\beta},\boldsymbol{\gamma})$ são obtidos através da diferenciação do logaritmo da função verossimilhança com relação a γ_j, $j=1,\ldots,q$, que são dados por

$$\frac{\partial \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial \gamma_j}=\sum_{t=1}^{n}\frac{\partial l_t(\mu_t;\sigma_t^2)}{\partial \gamma_j}=\sum_{t=1}^{n}\frac{\partial \ell_t(\mu_t,\sigma_t^2)}{\partial \sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\frac{\partial \eta_{2t}}{\partial \gamma_j},$$

em que $d\sigma_t^2/d\eta_{2t}=1/h'(\sigma_t^2)$, $\partial\eta_{2t}/\partial\gamma_j=z_{tj}$ e

$$\frac{\partial \ell_t(\mu_t,\sigma_t^2)}{\partial \sigma_t^2}=-\frac{1}{2\sigma_t^2}+\frac{d(y_t;\mu_t)}{2(\sigma_t^2)^2}. \tag{2.8}$$

A função escore para $\boldsymbol{\gamma}$ pode ser expressa matricialmente por

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

em que Z é uma matriz $n\times q$ cuja t-ésima linha é $\boldsymbol{z}_t=(z_{t1},z_{t2},\ldots,z_{tq})$, $H=\text{diag}\{1/h'(\sigma_1^2),\ldots,1/h'(\sigma_n^2)\}$ e $\boldsymbol{a}=(a_1,\ldots,a_n)^\top$, com

$$a_t=-\frac{1}{2\sigma_t^2}+\frac{d(y_t;\mu_t)}{2(\sigma_t^2)^2}, \tag{2.9}$$

em que $d(y_t;\mu_t)$ está definido em (2.3).

A matriz de informação de Fisher para os vetores de parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ é obtida através dos cálculos das segundas derivadas do logaritmo natural da função de verossimilhança com respeito a $\boldsymbol{\beta}$ e a $\boldsymbol{\gamma}$, respectivamente. Dessa forma, para $i,r=1,\ldots,k$, temos

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i\partial\beta_r} &= \frac{\partial}{\partial\beta_r}\left[\sum_{t=1}^{n}\frac{\partial \ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t}\frac{d\mu_t}{d\eta_{1t}}\frac{\partial\eta_{1t}}{\partial\beta_i}\right]\\ &= \sum_{t=1}^{n}\frac{\partial}{\partial\beta_r}\left[\frac{\partial \ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t}\frac{d\mu_t}{d\eta_{1t}}\right]x_{ti}\\ &= \sum_{t=1}^{n}\left(\frac{\partial^2 \ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t^2}\frac{d\mu_t}{d\eta_{1t}}+\frac{\partial \ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t}\frac{\partial}{\partial\mu_t}\frac{d\mu_t}{d\eta_{1t}}\right)\frac{d\mu_t}{d\eta_{1t}}x_{ti}x_{tr}.\end{aligned} \tag{2.10}$$

É possível mostrar que $\mathbb{E}(\partial \ell_t(\mu_t,\sigma_t^2)/\partial\mu_t)=0$. Portanto,

$$\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i\partial\beta_r}\right)=\sum_{t=1}^{n}\mathbb{E}\left(\frac{\partial^2 \ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t^2}\right)\left(\frac{d\mu_t}{d\eta_{1t}}\right)^2 x_{ti}x_{tr}.$$

De (2.4) temos que

$$\frac{\partial^2 \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t^2} = -\frac{1}{2\sigma_t^2} d^{''}(y_t, \mu_t)$$

e pela Proposição (e) na seção 2.2.1 segue que

$$\frac{1}{2}\mathbb{E}[d^{''}(y_t, \mu_t)] = \frac{3\sigma_t^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}.$$

Logo, a esperança de (2.10) fica dada por

$$\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \beta_i \partial \beta_r}\right) = -\sum_{t=1}^{n} \frac{1}{\sigma_t^2} w_t x_{ti} x_{tr},$$

em que

$$w_t = \left(\frac{3\sigma_t^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}\right) \frac{1}{[g^{'}(\mu_t)]^2}. \tag{2.11}$$

Matricialmente, temos que a informação de Fisher para $\boldsymbol{\beta}$ é dada por

$$K_{\beta\beta} = -\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \beta_i \partial \beta_r}\right) = X^\top \Sigma W X, \tag{2.12}$$

em que $W = \text{diag}(w_1, \ldots, w_n)$ e $\Sigma = \text{diag}(1/\sigma_1^2, \ldots, 1/\sigma_n^2)$.

Temos ainda que a derivada de segunda ordem do logaritmo da função de verossimilhança com relação a β_i e γ_j, $i = 1, \ldots, k$ e $j = 1, \ldots, q$, é dada por

$$\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_j} = \sum_{t=1}^{n} \left(\frac{d^{'}(y_t; \mu_t)}{2(\sigma_t^2)^2}\right) \frac{1}{g^{'}(\mu_t)} \frac{1}{h^{'}(\sigma_t^2)} \frac{\partial \eta_{1t}}{\partial \beta_i} \frac{\partial \eta_{2t}}{\partial \gamma_j}. \tag{2.13}$$

Pela Proposição (g) na Seção 2.2.1, temos que $\mathbb{E}(d^{'}(y_t; \mu_t)) = 0$. Portanto,

$$\mathbb{E}\left(\frac{\partial^2 \ell(\beta, \gamma)}{\partial \beta_i \partial \gamma_j}\right) = 0,$$

e, consequentemente, $K_{\beta\gamma} = K_{\gamma\beta}^\top = 0$, o que significa que os parâmetros que indexam a distribuição simplex são ortogonais.

Por fim, a derivada de segunda ordem de $\ell(\boldsymbol{\beta}, \boldsymbol{\gamma})$ com relação a γ_j e γ_s, para $j, s = 1, \ldots, q$ é dada por

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s} &= \frac{\partial}{\partial\gamma_s}\left[\sum_{t=1}^{n}\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\frac{\partial\eta_{2t}}{\partial\gamma_j}\right] \\ &= \sum_{t=1}^{n}\frac{\partial}{\partial\gamma_s}\left[\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\right]z_{tj} \\ &= \sum_{t=1}^{n}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}\frac{d\sigma_t^2}{d\eta_{2t}}+\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{\partial}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\right)\frac{d\sigma_t^2}{d\eta_{2t}}z_{tj}z_{ts}.\end{aligned} \tag{2.14}$$

Como $\mathbb{E}(\partial\ell_t(\mu_t,\sigma_t^2)/\partial\sigma_t^2)=0$, temos que

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s}\right)=\sum_{t=1}^{n}\mathbb{E}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}\right)\left(\frac{d\sigma_t^2}{d\eta_{2t}}\right)^2 z_{tj}z_{ts}.$$

De (2.8), temos que

$$\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}=\frac{1}{2(\sigma_t^2)^2}-\frac{d(y_t;\mu_t)}{(\sigma_t^2)^3}.$$

Logo,

$$\mathbb{E}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}\right)=\frac{1}{2(\sigma_t^2)^2}-\frac{1}{(\sigma_t^2)^3}\mathbb{E}[d(y_t;\mu_t)].$$

Pela Proposição (a) na Seção 2.2.1, segue que

$$\mathbb{E}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}\right)=-\frac{1}{2(\sigma_t^2)^2},$$

e, portanto,

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s}\right)=-\sum_{t=1}^{n}v_t z_{tj}z_{ts},$$

em que

$$v_t=\frac{1}{2(\sigma_t^2)^2}\frac{1}{[h'(\sigma_t^2)]^2}. \tag{2.15}$$

Matricialmente, temos que a informação de Fisher para $\boldsymbol{\gamma}$ pode ser escrita como

$$K_{\gamma\gamma}=-\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s}\right)=Z^\top V Z \tag{2.16}$$

em que $V=\mathrm{diag}(v_1,\ldots,v_n)$.

Assim, a matriz de informação de Fisher para o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ é dada por

$$K = K(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{pmatrix} K_{\beta\beta} & 0 \\ 0 & K_{\gamma\gamma} \end{pmatrix},$$

em que $K_{\beta\beta}$ e $K_{\gamma\gamma}$ são definidas em (2.12) e (2.16), respectivamente.

Sob condições gerais de regularidade (SEN & SINGER, 1993) e para grandes amostras, a distribuição aproximada dos estimadores de máxima verossimilhança é dada por

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} \sim N_{k+q}\left(\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}, K^{-1}\right),$$

em que $\widehat{\boldsymbol{\beta}}$ e $\widehat{\boldsymbol{\gamma}}$ são os estimadores de máxima verossimilhança de β e γ, respectivamente, e $K^{-1} = K^{-1}(\boldsymbol{\beta}, \boldsymbol{\gamma})$ é a inversa da matriz de informação de Fisher definida por

$$K^{-1} = K^{-1}(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \begin{pmatrix} K^{\beta\beta} & 0 \\ 0 & K^{\gamma\gamma} \end{pmatrix},$$

em que $K^{\beta\beta} = (X^\top \Sigma W X)^{-1}$ e $K^{\gamma\gamma} = (Z^\top V Z)^{-1}$.

Os estimadores de máxima verossimilhança de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são obtidos através da solução do sistema de equações $U_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{0}$ e $U_\gamma(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \mathbf{0}$. No entanto, sabe-se que tais estimadores não possuem expressões em formas fechadas, sendo necessário a utilização de algoritmos numéricos tais como Newton-Raphson, Escore de Fisher, BHHH, BFGS. O chute inicial para $\boldsymbol{\beta}$ pode ser o mesmo utilizado para o modelo de regressão beta, ou seja, $\beta^{(0)} = (X^\top X)^{-1} X^\top g(y_t)$. Já para $\boldsymbol{\gamma}$ pode ser dado por $\gamma^{(0)} = (Z^\top Z)^{-1} Z^\top d(y_t; \breve{\mu}_t)$, em que $\breve{\mu}_t = g^{-1}(X(X^\top X)^{-1} X^\top g(y_t))$.

2.4 Modelo de Regressão Simplex Não Linear

Apresentaremos, nesta seção, o modelo de regressão simplex não linear, no qual consideramos uma estrutura não linear para o sub-modelo da média e para o sub-modelo da dispersão, simultaneamente.

Sejam $y_1, \ldots, y_n$ variáveis aleatórias independentes, em que cada y_t, $t = 1, \ldots, n$, segue distribuição simplex cuja densidade é dada em (2.2), com média μ_t e parâmetro de

dispersão σ_t^2. O modelo de regressão simplex não linear assume que a média e o parâmetro de dispersão satisfazem as seguintes relações funcionais:

$$g(\mu_t) = f_1(x_t^\top; \boldsymbol{\beta}) = \eta_{1t} \quad \text{e} \quad h(\sigma_t^2) = f_2(z_t^\top; \boldsymbol{\gamma}) = \eta_{2t}, \tag{2.17}$$

em que $\boldsymbol{\beta} = (\beta_1, ..., \beta_k)^\top$ e $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_q)^\top$ são vetores de parâmetros de regressão desconhecidos, $\boldsymbol{\beta} \in \mathbb{R}^k$ e $\boldsymbol{\gamma} \in \mathbb{R}^q$, $k + q < n$, η_{1t} e η_{2t} são preditores não lineares, e $\boldsymbol{x}_t = (x_{t1}, ..., x_{tk_1})^\top$ e $\boldsymbol{z}_t = (z_{t1}, ..., z_{tq_1})^\top$ são observações em k_1 e q_1 covariáveis conhecidas, $k_1 \leq k$ e $q_1 \leq q$. Ambas, $g : (0, 1) \to \mathbb{R}$ e $h : (0, \infty) \to \mathbb{R}$ são funções de ligação estritamente monótonas e duas vezes diferenciáveis. Diversas funções de ligação foram discutidas na seção anterior. O modelo acima foi proposto por ESPINHEIRA & SILVA (2018).

2.4.1 Função Escore, Matriz de Informação de Fisher e Estimação dos Parâmetros

Baseado em (2.2), o logaritmo da função da função de verossimilhança é dado por

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \sum_{t=1}^{n} \ell_t(\mu_t, \sigma_t^2),$$

em que

$$\ell_t(\mu_t, \sigma_t^2) = -\frac{1}{2}\log(2\pi) - \frac{1}{2}\log(\sigma_t^2) - \frac{3}{2}\log[y_t(1 - y_t)] - \frac{1}{2\sigma_t^2} d(y_t; \mu_t).$$

Os componentes do vetor escore $U_\beta(\boldsymbol{\beta}, \boldsymbol{\gamma})$ são obtidos através da diferenciação do logaritmo da função verossimilhança com relação a β_i, $i = 1, \ldots, k$, e são dados por

$$\frac{\partial \ell(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \beta_i} = \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t} \frac{d\mu_t}{d\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i},$$

em que $d\mu_t/d\eta_{1t} = 1/g'(\mu_t)$ e

$$\frac{\partial \ell_t(\mu_t, \sigma_t^2)}{\partial \mu_t} = -\frac{1}{2\sigma_t^2} d'(y_t; \mu_t), \tag{2.18}$$

com $d'(y_t; \mu_t)$ definido em (2.5).

Reescrevendo a função escore para o vetor $\boldsymbol{\beta}$ em forma matricial temos

$$U_\beta(\boldsymbol{\beta},\boldsymbol{\gamma}) = J_1^\top \Sigma T U(\boldsymbol{y}-\boldsymbol{\mu}),$$

em que $J_1 = \partial\eta_1/\partial\beta$ é uma matriz de derivadas $n \times k$, $\Sigma = \text{diag}\{1/\sigma_1^2,\ldots,1/\sigma_n^2\}$, e T e U são dadas em (2.6).

Analogamente, os componentes do vetor escore $U_\gamma(\boldsymbol{\beta},\boldsymbol{\gamma})$ são obtidos através da diferenciação do logaritmo da função verossimilhança com relação a γ_j, $j = 1,\ldots,q$, que são dados por

$$\frac{\partial\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j} = \sum_{t=1}^{n}\frac{\partial l_t(\mu_t;\sigma_t^2)}{\partial\gamma_j} = \sum_{t=1}^{n}\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\frac{\partial\eta_{2t}}{\partial\gamma_j},$$

em que $d\sigma_t^2/d\eta_{2t} = 1/h'(\sigma_t^2)$ e

$$\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2} = -\frac{1}{2\sigma_t^2} + \frac{d(y_t;\mu_t)}{2(\sigma_t^2)^2}. \tag{2.19}$$

Matricialmente, temos que

$$U_\gamma(\boldsymbol{\beta},\boldsymbol{\gamma}) = J_2^\top H\boldsymbol{a},$$

em que $J_2 = \partial\eta_2/\partial\beta$ é uma matriz de derivadas $n \times q$, $H = \text{diag}\{1/h'(\sigma_1^2),\ldots,1/h'(\sigma_n^2)\}$ e $\boldsymbol{a} = (a_1,\ldots,a_n)^\top$, com

$$a_t = -\frac{1}{2\sigma_t^2} + \frac{d(y_t;\mu_t)}{2(\sigma_t^2)^2}.$$

De forma similar ao modelo linear, a matriz de informação de Fisher para os vetores de parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ é uma matriz bloco diagonal e é obtida através dos cálculos das derivadas de segunda ordem de $\ell(\boldsymbol{\beta},\boldsymbol{\gamma})$. Logo, para $i,r = 1,\ldots,k$, temos

$$\begin{aligned}
\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i\partial\beta_r} &= \frac{\partial}{\partial\beta_r}\left[\sum_{t=1}^{n}\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t}\frac{d\mu_t}{d\eta_{1t}}\frac{\partial\eta_{1t}}{\partial\beta_i}\right] \\
&= \sum_{t=1}^{n}\frac{\partial}{\partial\beta_r}\left[\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t}\frac{d\mu_t}{d\eta_{1t}}\right]\frac{\partial\eta_{1t}}{\partial\beta_i} \\
&= \sum_{t=1}^{n}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t^2}\frac{d\mu_t}{d\eta_{1t}} + \frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t}\frac{\partial}{\partial\mu_t}\frac{d\mu_t}{d\eta_{1t}}\right)\frac{d\mu_t}{d\eta_{1t}}\frac{\partial\eta_{1t}}{\partial\beta_r}\frac{\partial\eta_{1t}}{\partial\beta_i}.
\end{aligned} \tag{2.20}$$

É possível mostrar que $\mathbb{E}(\partial\ell_t(\mu_t,\sigma_t^2)/\partial\mu_t)=0$. Portanto,

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i\partial\beta_r}\right)=\sum_{t=1}^{n}\mathbb{E}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial\mu_t^2}\right)\left(\frac{d\mu_t}{d\eta_{1t}}\right)^2\frac{\partial\eta_{1t}}{\partial\beta_i}\frac{\partial\eta_{1t}}{\partial\beta_r}. \tag{2.21}$$

De (2.18) e pela proposição (e) na Seção 2.2.1

$$\frac{1}{2}\mathbb{E}[d^{''}(y_t,\mu_t)]=\frac{3\sigma_t^2}{\mu_t(1-\mu_t)}+\frac{1}{\mu_t^3(1-\mu_t)^3},$$

a expressão em (2.21) pode ser reescrita por

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i\partial\beta_r}\right)=-\sum_{t=1}^{n}\frac{1}{\sigma_t^2}w_t j_{1ti}j_{1tr},$$

em que j_{1ti} é o vetor da matriz J_1 referente a covariada i e

$$w_t=\left(\frac{3\sigma_t^2}{\mu_t(1-\mu_t)}+\frac{1}{\mu_t^3(1-\mu_t)^3}\right)\frac{1}{[g'(\mu_t)]^2}.$$

Matricialmente, temos que a informação de Fisher para β é dada por

$$K_{\beta\beta}=-\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i\partial\beta_r}\right)=J_1^{\top}\Sigma W J_1. \tag{2.22}$$

Para obtermos a informação de Fisher de γ, temos que a derivada de segunda ordem de $\ell(\boldsymbol{\beta},\boldsymbol{\gamma})$ com relação a γ_j e γ_s, para $j,s=1,\ldots,q$ é dada por

$$\begin{aligned}\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s} &= \frac{\partial}{\partial\gamma_s}\left[\sum_{t=1}^{n}\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\frac{\partial\eta_{2t}}{\partial\gamma_j}\right]\\ &= \sum_{t=1}^{n}\frac{\partial}{\partial\gamma_s}\left[\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\right]\frac{\partial\eta_{2t}}{\partial\gamma_j}\\ &= \sum_{t=1}^{n}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}\frac{d\sigma_t^2}{d\eta_{2t}}+\frac{\partial\ell_t(\mu_t,\sigma_t^2)}{\partial\sigma_t^2}\frac{\partial}{\partial\sigma_t^2}\frac{d\sigma_t^2}{d\eta_{2t}}\right)\frac{d\sigma_t^2}{d\eta_{2t}}\frac{\partial\eta_{2t}}{\partial\gamma_j}\frac{\partial\eta_{2t}}{\partial\gamma_s}.\end{aligned} \tag{2.23}$$

Como $\mathbb{E}(\partial\ell_t(\mu_t,\sigma_t^2)/\partial\sigma_t^2)=0$, temos que

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s}\right)=\sum_{t=1}^{n}\mathbb{E}\left(\frac{\partial^2\ell_t(\mu_t,\sigma_t^2)}{\partial(\sigma_t^2)^2}\right)\left(\frac{d\sigma_t^2}{d\eta_{2t}}\right)^2\frac{\partial\eta_{2t}}{\partial\gamma_j}\frac{\partial\eta_{2t}}{\partial\gamma_s}.$$

De (2.19), temos que

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j\partial\gamma_s}\right)=-\sum_{t=1}^{n}v_t j_{2tj}j_{2ts},$$

em que j_{2tj} é o vetor da matriz J_2 referente a covariada j e v_t é dado em (2.15).

Matricialmente, temos que a informação de Fisher para γ é dada por

$$K_{\gamma\gamma} = -\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\gamma_j \partial\gamma_s}\right) = J_2^\top V J_2$$

em que $V = \text{diag}\{v_1, \ldots, v_t\}$.

Finalmente, baseando-se em (2.13) e na Proposição (g) (Seção 2.2.1), é fácil ver que

$$\mathbb{E}\left(\frac{\partial^2 \ell(\boldsymbol{\beta},\boldsymbol{\gamma})}{\partial\beta_i \partial\gamma_j}\right) = 0,$$

ou seja, $K_{\beta\gamma} = K_{\gamma\beta}^\top = 0$.

Assim, a matriz de informação de Fisher para o vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ para o modelo de regressão simplex não linear é dada por

$$K = K(\boldsymbol{\beta},\boldsymbol{\gamma}) = \begin{pmatrix} K_{\beta\beta} & 0 \\ 0 & K_{\gamma\gamma} \end{pmatrix}, \tag{2.24}$$

em que $K_{\beta\beta} = J_1^\top \Sigma W J_1$ e $K_{\gamma\gamma} = J_2^\top V J_2$.

Para amostras grandes e sob condições gerais de regularidade, a distribuição aproximada dos estimadores de máxima verossimilhança é dada por

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} \sim N_{k+q}\left(\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}, K^{-1}\right),$$

em que $\widehat{\boldsymbol{\beta}}$ e $\widehat{\boldsymbol{\gamma}}$ são os estimadores de máxima verossimilhança de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, respectivamente, e $K^{-1} = K^{-1}(\boldsymbol{\beta},\boldsymbol{\gamma})$ é a inversa da matriz de informação de Fisher definida em (2.24).

Os estimadores de máxima verossimilhança de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são obtidos através da solução do sistema de equações $U_\beta(\boldsymbol{\beta},\boldsymbol{\gamma}) = \mathbf{0}$ e $U_\gamma(\boldsymbol{\beta},\boldsymbol{\gamma}) = \mathbf{0}$. No entanto, sabe-se que tais estimadores não possuem expressões em formas fechadas, sendo necessário a utilização de algoritmos numéricos tais como Newton-Raphson, Escore de Fisher, BHHH, BFGS.

O chute inicial para $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ no modelo de regressão simplex não linear é baseado em mínimos quadrados não linear e uma aproximação não linear (BATES & WATTS, 1988). Suponha que $k_1 = k$ e $q_1 = q$. A expansão de Taylor de primeira ordem de $f_1(x_t;\boldsymbol{\beta})$ em $\boldsymbol{\beta}^{(0)}$ é

$$f_1(x_t;\boldsymbol{\beta}) \approx f_1(x_t;\boldsymbol{\beta}^{(0)}) + \sum_{i=1}^{k}\left[\frac{\partial f_1(x_t;\boldsymbol{\beta})}{\partial\beta_i}\right]_{\boldsymbol{\beta}=\boldsymbol{\beta}^{(0)}} (\beta_i - \beta_i^{(0)}),$$

em que $\boldsymbol{\beta}^{(0)} = (\beta_1^{(0)}, \ldots, \beta_k^{(0)})$ é um valor inicial. Assim,

$$f_1(x_t; \boldsymbol{\beta}) = f_1(x_t; \boldsymbol{\beta}^{(0)}) + \sum_{i=1}^{k} j_{ti}^{(0)}(\beta_i - \beta_i^{(0)}).$$

Similarmente ao que acontece com os modelos normais não lineares, em que $f_1(x_t; \boldsymbol{\beta}) = y_t$, aqui $f_1(x_t; \boldsymbol{\beta}) = g(y_t)$. Além disso, $\theta_i^{(0)} = (\beta_i - \beta_i^{(0)})$, segue que $g(y_t) - f_1(x_t; \boldsymbol{\beta}) = \sum_{i=1}^{k} j_{ti}^{(0)}\theta_i^{(0)}$, que pode ser visto como um modelo linear para o qual o estimador de mínimos quadrados de $\theta^{(0)}$ é dado por $\widehat{\theta^{(0)}} = (J_1^{(0)\top} J_1^{(0)})^{-1} J_1^{(0)\top}(g(\boldsymbol{y}) - f_1(x; \boldsymbol{\beta}^{(0)}))$, com $J_1^{(0)} = [\partial \boldsymbol{\eta_1}/\partial \boldsymbol{\beta}]_{\beta=\beta^{(0)}}$ e $\theta_i^{(0)} = (\widehat{\beta_i}^{(1)} - \beta_i^{(0)})$. Consequentemente, $\widehat{\beta_i}^{(1)} = \theta_i^{(0)} + \beta_i^{(0)}$. Logo,

$$\boldsymbol{\beta}_{NL}^{(0)} = (J_1^{(0)\top} J_1^{(0)})^{-1} J_1^{(0)\top}(g(\boldsymbol{y}) - f_1(x; \boldsymbol{\beta}_L^{(0)})),$$

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

Para o submodelo da dispersão consideramos $h(\sigma_t^2) = f_2(z_t; \boldsymbol{\gamma})$. Assim,

$$\boldsymbol{\gamma}_{NL}^{(0)} = (J_2^{(0)\top} J_2^{(0)})^{-1} J_2^{(0)\top}(h(\sigma_{NL}^2)) - f_2(z; \boldsymbol{\gamma}_L^{(0)})),$$

em que $\boldsymbol{\gamma}_L^{(0)} = (Z^\top Z)^{-1} Z^\top h(\sigma_L^2)$, $J_2^{(0)} = [\partial \boldsymbol{\eta_2}/\partial \boldsymbol{\gamma}]_{\gamma=\gamma^{(0)}}$ e $\sigma_{tL}^2 = d(y_t; \breve{\mu}_{tL})$, com $\breve{\mu}_{tL} = g^{-1}(\widehat{\eta}_{1L}) = g^{-1}(x_t^\top \boldsymbol{\beta}_L^{(0)})$. Finalmente,

$$\sigma_{NL}^{2(0)} = d(y_t; \breve{\mu}_{tNL})$$

com $\breve{\mu}_{tNL} = g^{-1}(\widehat{\eta}_{1NL}) = g^{-1}(x_t^\top \boldsymbol{\beta}_{NL}^{(0)})$.

Muitas vezes $k_1 < k$ e/ou $q_1 < q$, isto é, em um ou em ambos os submodelos há mais parâmetros do que covariadas na função não linear. Quando esse é o caso, consideramos dois subconjuntos de $\boldsymbol{\beta}_L^{(0)}$ e obtemos seus valores separadamente. Primeiro, definimos os parâmetros de valores arbitrários, de modo que ficamos com um preditor formado por covariadas que não envolvem parâmetros desconhecidos. Construímos agora uma matriz X usando tais regressores e computamos $(X^\top X)^{-1} X^\top g(\boldsymbol{y})$. Usando esse procedimento em dois passos podemos obter $\boldsymbol{\beta}_L^{(0)}$ quando há mais parâmetros que variáveis independentes. Esse chute inicial foi proposto por ESPINHEIRA & SILVA (2018).

3 RESÍDUO COMBINADO PARA O MODELO DE REGRESSÃO SIMPLEX NÃO LINEAR

3.1 Introdução

Diversos modelos de regressão para dados restritos ao intervalo $(0, 1)$ estão sendo estudados na literatura, tais como o modelo de regressão beta (FERRARI & CRIBARI-NETO, 2004), o modelo de regressão Kumaraswamy (MITNIK & BAEK, 2013), o modelo Johnson S_b (LEMONTE & BAZAN, 2016), o modelo gama unitário (MOUSA et al., 2013), o modelo de regressão simplex (BARNDORFF-NIELSEN & JøRGENSEN, 1991), entre outros. Em especial, diversos autores vêm se dedicando ao modelo de regressão simplex, como por exemplo, SONG & TAN (2000), SONG et al. (2004), QUI et al. (2008) e SONG (2009). Mais recentemente, ZHANG (2016) implementaram o pacote simplexreg no software R que está disponível no Comprehensive R Archive Network (CRAN) em `https://CRAN.R-project.org/package=simplexreg`, e ESPINHEIRA & SILVA (2018) propõem a classe de modelos de regressão simplex não linear. Os autores apresentam as expressões fechadas para a função escore e a matriz de informação de Fisher, estimação dos parâmetros através do método de máxima verossimilhança e algumas medidas de diagnóstico, tais como resíduo e medidas de influência local.

Após a estimação dos parâmetros e do ajuste do modelo, é preciso verificar a adequação do mesmo aos dados. Na análise de diagnóstico, verificamos possíveis problemas nas suposições feitas para o modelo, como por exemplo, problemas na escolha da função de ligação, problemas na escolha das covariadas, problemas de não-linearidade e até mesmo problemas com observações discrepantes que interferem nas estimativas dos parâmetros. Uma das técnicas de diagnóstico mais utilizada é a análise de resíduos. Os resíduos desempenham um papel importante na verificação da adequação do modelo e na identificação de "outliers" entre os valores ajustados a partir do modelo e os valores observados. A análise de diagnóstico foi amplamente estudada em alguns modelos de regressão e com isso vários resíduos foram propostos. ATKINSON (1985), CORDEIRO & DEMETRIO (2008), MC-

CULLAGH & NELDER (1989), PAULA (2013), entre outros propuseram resíduos para os modelos lineares generalizados. Na classe de modelos para dados que se distribuem de forma contínua no intervalo $(0, 1)$ muitos resultados de diagnóstico foram desenvolvidos em especial para a classe de modelos de regressão beta (FERRARI & CRIBARI-NETO, 2004). Em ESPINHEIRA et al. (2008b) são desenvolvidos resíduos e medidas de influência local para a classe de modelos de regressão beta linear. Além disso, ESPINHEIRA et al. (2017) apresentam um novo resíduo para a classe de modelos de regressão beta não linear e LEMONTE & BAZAN (2016) também propuseram resíduos e medidas de influência local para a classe de modelos Johnson S_b. Recentemente, ESPINHEIRA & SILVA (2018) desenvolveram um resíduo baseado no processo iterativo escore de Fisher para β e medidas de influência local considerando diversos esquemas de perturbação para os modelos de regressão simplex em que tanto a média quanto a dispersão podem ser explicados com base em preditores não lineares.

Neste capítulo nós propomos um novo resíduo para a classe de modelos de regressão simplex não linear. O novo resíduo é facilmente calculado e é baseado nos processos iterativos escore de Fisher para a estimação dos parâmetros do modelo da média e do modelo da dispersão tal como em ESPINHEIRA et al. (2017) para o modelo de regressão beta. O capítulo é organizado da seguinte maneira. Na Seção 3.2, introduzimos o novo resíduo para modelo de regressão simplex denominado resíduo combinado. Na Seção 3.3 apresentamos resultados de simulações de Monte Carlo sob diversos cenários. Aplicações utilizando dados simulados e dados reais são apresentados na Seção 3.4. Finalmente, algumas observações finais são encontradas na Seção 3.5.

3.2 Resíduo Combinado

Nossa proposta é definir o resíduo combinado para a classe de modelos de regressão simplex não linear. Esse resíduo baseia-se no algoritmo iterativo escore de Fisher para estimar $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ (ESPINHEIRA et al., 2017).

Seja o modelo de regressão simplex não linear definido em (2.17), o processo iterativo escore de Fisher para a estimação de $\boldsymbol{\beta}$ é dado por

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + (J_1^\top \Sigma^{(m)} W^{(m)} J_1)^{-1} J_1^\top \Sigma^{(m)} T^{(m)} U^{(m)} (\boldsymbol{y} - \boldsymbol{\mu}^{(m)}), \tag{3.1}$$

em que as matrizes T e U estão definidas em (2.6) e W é uma matriz diagonal com elementos definidos em (2.11). Além disso, $J_1 = \partial\boldsymbol{\eta}_1/\partial\boldsymbol{\beta}$ é uma matriz de derivadas n x k e $\Sigma = \text{diag}\{1/\sigma_1^2, \ldots, 1/\sigma_n^2\}$. Similarmente, o m-ésimo passo do algoritmo escore de Fisher para estimar $\boldsymbol{\gamma}$ é dado por

$$\boldsymbol{\gamma}^{(m+1)} = \boldsymbol{\gamma}^{(m)} + (J_2^\top V^{(m)} J_2)^{-1} J_2^\top H^{(m)} \boldsymbol{a}^{(m)}, \tag{3.2}$$

em que V está definida em (2.15), $H = \text{diag}\{1/h'(\sigma_1^2), \ldots, 1/h'(\sigma_n^2)\}$, $J_2 = \partial\boldsymbol{\eta}_2/\partial\boldsymbol{\gamma}$ é uma matriz de derivadas $n \times p$ e o t-ésimo elemento de $\boldsymbol{a}$ é dado por

$$a_t = \left\{ \frac{d(y_t; \mu_t)}{2(\sigma_t^2)^2} - \frac{1}{2\sigma_t^2} \right\}.$$

É possível reescrever (3.1) e (3.2) em termos dos estimadores de mínimos quadrados ponderados, como

$$\boldsymbol{\beta}^{(m+1)} = (J_1^\top \Sigma^{(m)} W^{(m)} J_1)^{-1} J_1^\top \Sigma^{(m)} W^{(m)} \boldsymbol{u}_1^{(m)} \quad \text{e} \quad \boldsymbol{\gamma}^{(m+1)} = (J_2^\top V^{(m)} J_2)^{-1} J_2^\top V^{(m)} \boldsymbol{u}_2^{(m)}$$

,

respectivamente. Temos que $\boldsymbol{u}_1^{(m)} = J_1\boldsymbol{\beta}^{(m)} + W^{-1(m)} U^{(m)} T^{(m)} (\boldsymbol{y} - \boldsymbol{\mu}^{(m)})$, $\boldsymbol{u}_2^{(m)} = J_2\boldsymbol{\gamma}^{(m)} + V^{-1(m)} H^{(m)} \boldsymbol{a}^{(m)}$.

Após convergência, obtemos

$$\begin{gathered} \widehat{\boldsymbol{\beta}} = (J_1^\top \widehat{\Sigma}\widehat{W} J_1)^{-1} J_1^\top \widehat{\Sigma}\widehat{W}\boldsymbol{u}_1 \quad \text{e} \quad \widehat{\boldsymbol{\gamma}} = (J_2^\top \widehat{V} J_2)^{-1} J_2^\top \widehat{V}\boldsymbol{u}_2, \quad \text{com} \\ \boldsymbol{u}_1 = \widehat{\boldsymbol{\eta}_1} + \widehat{W}^{-1}\widehat{U}\widehat{T}(\boldsymbol{y} - \widehat{\boldsymbol{\mu}}) \quad \text{e} \quad \boldsymbol{u}_2 = \widehat{\boldsymbol{\eta}_2} + \widehat{V}^{-1}\widehat{H}\widehat{\boldsymbol{a}}. \end{gathered} \tag{3.3}$$

em que $\widehat{W}$, $\widehat{T}$, $\widehat{H}$ e $\widehat{V}$ são as matrizes W, T, H e V. avaliadas nos estimadores de máxima verossimilhança, respectivamente. Temos ainda que $\widehat{\boldsymbol{\beta}}$ e $\widehat{\boldsymbol{\gamma}}$ em (3.3) podem ser vistos como estimadores de mínimos quadrados de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ através das regressões $\widehat{\Sigma}^{1/2}\widehat{W}^{1/2}\boldsymbol{u}_1$ com $\widehat{\Sigma}^{1/2}\widehat{W}^{1/2}J_1$ e $\widehat{V}^{1/2}\boldsymbol{u}_2$ com $\widehat{V}^{1/2}J_2$. Os resíduos ordinários obtidos através do processo iterativo de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são dados por $r^\beta = \widehat{\Sigma}^{1/2}\widehat{W}^{1/2}(\boldsymbol{u}_1 - J_1\widehat{\boldsymbol{\beta}}) = \widehat{\Sigma}^{1/2}\widehat{W}^{-1/2}\widehat{T}\widehat{U}(\boldsymbol{y} - \widehat{\boldsymbol{\mu}})$ e $r^\gamma = \widehat{V}^{1/2}(\boldsymbol{u}_2 - J_2\widehat{\boldsymbol{\gamma}}) = \widehat{V}^{-1/2}\widehat{H}\widehat{\boldsymbol{a}}$, e podem ser reescritos por

$$r_t^\beta = \frac{\widehat{u}_t(y_t - \widehat{\mu}_t)}{\sqrt{\widehat{b}_t}} \quad \text{e} \quad r_t^\gamma = \frac{\widehat{a}_t}{\sqrt{[2(\widehat{\sigma^2}_t)^2]^{-1}}}, \tag{3.4}$$

em que u_t e a_t são dados em (2.7) e (2.9), respectivamente, e b_t é dado por

$$b_t = \sigma_t^2 \left\{ \frac{3\sigma_t^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3} \right\}.$$

Aqui r_t^β é o resíduo ponderado proposto por ESPINHEIRA & SILVA (2018). Nós podemos ainda padronizar r^β utilizando a variância de u_1. Note que $\mathrm{Cov}(\widehat{\boldsymbol{\beta}}) \approx (J_1^\top \widehat{\Sigma}\widehat{W} J_1)^{-1}$, logo, $\mathrm{Cov}(\boldsymbol{u}_1) \approx (\widehat{\Sigma}\widehat{W})^{-1}$. Assim, temos que $\widehat{\mathrm{Cov}}(r^\beta) = (1 - H^*)$ em que

$$H^* = (\widehat{W}\widehat{\Sigma})^{1/2} J_1 (J_1 \widehat{\Sigma}\widehat{W} J_1)^{-1} J_1^\top (\widehat{\Sigma}\widehat{W})^{1/2}. \tag{3.5}$$

Portanto, o resíduo ponderado padronizado (ESPINHEIRA & SILVA, 2018) é definido por

$$r_t^\beta = \frac{\widehat{u}_t(y_t - \widehat{\mu}_t)}{\sqrt{\widehat{b}_t(1-h_{tt})}}, \tag{3.6}$$

em que h_{tt} é o t-ésimo elemento da diagonal principal da matriz H^* dada em (3.5).

Finalmente, o resíduo combinado para o modelo de regressão simplex é obtido através da combinação de r^β e r^γ e é dado por $r^{\beta\gamma} = \widehat{u}_t(y_t - \widehat{\mu}_t) + \widehat{a}_t$. Temos que os vetores de parâmetros $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$ são ortogonais (COX & REID, 1987) e consequentemente, $\mathrm{cov}(\widehat{u}_t(y_t - \widehat{\mu}_t)\widehat{a}_t) = 0$. Dessa forma, o resíduo combinado padronizado é dado por

$$r_{p.t}^{\beta\gamma} = \frac{\widehat{u}_t(y_t - \widehat{\mu}_t) + \widehat{a}_t}{\sqrt{\widehat{b}_t + [2(\widehat{\sigma^2}_t)^2]^{-1}}}. \tag{3.7}$$

Nosso objetivo é avaliar as características desse novo resíduo para a classe de modelos de regressão simplex não linear. Para isso, compararemos com o resíduo ponderado e o resíduo ponderado padronizado (ESPINHEIRA & SILVA, 2018).

3.3 Avaliação Numérica

Com o objetivo de avaliar a distribuição empírica do resíduo combinado definido em (3.7), realizamos simulações de Monte Carlo com 10000 réplicas sob diferente cenários. Juntamente com o resíduo combinado, utilizamos nas simulações os resíduos ponderado e ponderado padronizado definidos em (3.4) e (3.6), respectivamente. 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 simplex linear com dispersão variável dado por

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4} + \beta_5\, x_{t5}, \\ \log(\sigma_t^2) &= \gamma_1 + \gamma_2\, z_{t2} + \gamma_3\, z_{t3} + \gamma_4\, z_{t4} + \gamma_5\, z_{t5}, \quad t = 1, \ldots, n. \end{aligned} \tag{3.8}$$

As realizações das covariadas foram geradas através da distribuição uniforme, $x_i \sim U(0,1)$ e $z_i \sim U(-0.5, 0.5)$, $i = 2, \ldots, 5$ e mantidas fixas para cada réplica. Consideramos três intervalos para a média da variável resposta, a saber: $\mu \in (0.02, 0.15)(\beta = (-3.5, 1.2, 0.7, -1.3, 1.0))$, $\mu \in (0.32, 0.71)(\beta = (-1.9, 1.2, 1.0, 1.1, 1.3))$ e $\mu \in (0.80, 0.98)$ $(\beta = (2.0, 1.4, 0.8, -1.3, 1.0))$. Além disso, consideramos três valores para o grau de dispersão não constante: $\lambda = 20$ $(\gamma = (-1.2, 2.2, 1.3, 1.0, 1.0))$, $\lambda = 50$ $(\gamma = (-2.2, -1.9, 1.4, 2.2, 1.9))$ e $\lambda = 100$ $(\gamma = (-1.2, -2.2, 1.3, 3.18, 2.0))$, em que $\lambda = \max(\sigma_t^2)/\min(\sigma_t^2)$, $t = 1, \ldots, n$.

As Tabelas 1, 2 e 3 mostram as médias, erros-padrão, assimetrias e curtoses dos resíduos r^β, r_p^β e $r^{\beta\gamma}$ considerando os três cenários da média da variável resposta, $\lambda = 20$ e $n = 20$. É possível notar que as médias dos três resíduos são próximas de zero nos três cenários. Temos ainda que os erros-padrão estão bem próximos do valor um. Quanto à assimetria, observou-se que o resíduo combinado possui, em geral, assimetria positiva, com exceção das observações 4, 5, 6 e 13 da Tabela 3. Em contrapartida, os resíduos ponderado e ponderado padronizado apresentam assimetria positiva quando $\mu \approx 0$ e assimetria negativa nos demais cenários. As curtoses dos resíduos r^β e r_p^β estão mais próximas do valor três que a curtose do resíduo combinado, que por sua vez apresenta alguns valores superiores, como por exemplo, 4.093, 5.769 e 4.183 na Tabela 1.

Nas Tabelas 4, 5 e 6 contém os valores das estatísticas referentes aos três resíduos quando $\lambda = 50$. Nota-se o mesmo comportamento dos resíduos quando $\lambda = 20$. A média e o erro-padrão das 20 réplicas são próximas de zero e um, respectivamente, para os três resíduos. Observa-se que a assimetria do resíduo combinado é positiva e maior que a dos resíduos ponderado e ponderado padronizado. O resíduo combinado continua apresentando curtose acima de três e valores discrepantes para algumas observações. Na Tabela 4 verifica-se que as observações 12 e 17 apresentam valores de curtose iguais a

22.358 e 10.345, respectivamente. O caso 12 é a observação que apresenta o menor valor da variável resposta. Além disso, está associado com um valor extremo da covariada x_{t2}, 0.080934, o que gera um ponto de alta alavancagem, interferindo assim na curtose do resíduo combinado. O mesmo acontece nas Tabelas 5 e 6.

Por fim, as Tabelas 7, 8 e 9 mostram o comportamento dos resíduos ponderado, ponderado padronizado e combinado para os três cenários da média da variável resposta com $\lambda = 100$. O comportamento das estatísticas descritas das réplicas dos vinte resíduos mostra-se bastante semelhante aos demais valores de graus de dispersão não constante considerados. Além disso, para os três valores de λ, a assimetria e curtose do resíduo combinado são maiores quando a média da variável resposta assume valores centrais do intervalo $(0, 1)$ quando comparados aos valores de μ próximo dos extremos. No entanto, o inverso acontece para os resíduos ponderado e ponderado padronizado, ou seja, a assimetria e a curtose são menores nesse intervalo.

As Figuras 2 - 6 apresentam os gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado considerando o modelo da média da variável resposta dado em (3.8) e o parâmetro de dispersão constante. Na Figura 2 consideramos o valor $\sigma^2 = 3.5$ para o parâmetro de dispersão e o tamanho da amostra igual a $n = 20$. Observamos através dos QQ plots que as distribuições dos resíduos são bastante semelhantes e que as médias e erros-padrão se aproximam da distribuição normal padrão, no entanto, a assimetria dos três resíduos, em geral, são diferentes de zero e a curtose em alguns cenários é maior que três. Na Figura 3 apresentamos os gráficos dos quantis empíricos dos resíduos contra os quantis da normal padrão considerando $n = 40$, $\sigma^2 = 6.0$ e os mesmos cenários para a média da variável resposta. Mesmo com o aumento do tamanho amostral os três resíduos continuam sendo um pouco assimétricos, principalmente quando μ está próxima dos extremos do intervalo $(0, 1)$. As Figuras 4 e 5 apresentam os mesmos gráficos para $n = 60$ e $\sigma^2 = 0.4$ e $\sigma^2 = 3.5$, respectivamente. Observa-se que com exceção do resíduo combinado com $\mu \in (0.20; 0.88)$ na Figura 4, os três resíduos se comportam de forma similar. Na Figura 5 por exemplo, os três resíduos apresentam assimetria positiva quando $\mu \approx 0$ e assimetria negativa quando $\mu \approx 1$. O mesmo pode ser observado na Figura 6 que refere-se aos gráficos dos quantis empíricos dos resíduos contra os quantis da normal padrão considerando $n = 120$, $\sigma^2 = 6.0$.

Fizemos simulações considerando diversos outros cenários para o modelo de regressão simplex linear, no entanto, os resultados são similares. Observamos que a assimetria e a curtose dos resíduos não se aproximam com as 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, 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 (ESPINHEIRA et al., 2017).

Tabela 1 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 20$ e $\mu \in (0.02, 0.15)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	0.017	0.019	0.204	1.184	1.316	1.159	0.161	0.136	1.128	2.270	2.094	4.093
2	0.029	0.067	−0.206	0.750	1.199	0.741	0.149	0.025	1.028	2.829	2.117	3.640
3	0.010	0.008	0.109	1.112	1.222	1.096	0.278	0.243	1.458	2.651	2.519	5.769
4	−0.013	−0.014	0.002	0.977	1.028	0.950	0.357	0.318	0.562	2.489	2.506	2.793
5	0.017	0.014	0.001	1.054	1.144	1.039	0.329	0.252	0.608	2.230	2.132	2.524
6	0.008	0.008	−0.015	0.925	1.046	0.917	0.417	0.342	0.636	2.873	2.781	3.103
7	0.002	0.002	−0.023	0.985	1.109	0.942	0.335	0.240	0.940	2.433	2.285	3.313
8	0.020	0.020	0.080	1.060	1.163	1.001	0.264	0.203	0.966	2.305	2.215	3.318
9	0.020	0.018	0.172	1.236	1.360	1.179	0.143	0.117	0.644	1.851	1.763	2.682
10	0.030	0.034	0.017	0.998	1.153	0.985	0.183	0.142	0.735	2.667	2.531	3.373
11	0.036	0.068	−0.087	0.902	1.301	0.862	0.098	0.024	0.781	2.222	1.917	2.823
12	0.011	−0.002	−0.163	0.536	1.262	0.577	0.209	−0.010	0.746	4.021	2.642	3.639
13	0.007	0.007	−0.017	0.903	0.979	0.898	0.380	0.322	0.575	2.692	2.635	2.875
14	0.008	0.004	0.004	0.993	1.113	0.981	0.295	0.239	0.610	2.716	2.576	3.111
15	−0.001	0.000	0.062	1.079	1.175	1.001	0.195	0.171	0.864	2.163	2.151	3.235
16	−0.011	−0.019	0.376	1.289	1.522	1.191	0.175	0.107	0.947	1.964	1.981	3.165
17	0.025	0.030	−0.041	0.958	1.158	0.935	0.211	0.165	0.788	2.491	2.369	3.275
18	−0.038	−0.068	−0.024	1.008	1.370	0.893	0.101	0.070	0.728	1.906	1.788	2.643
19	0.013	0.014	−0.046	0.937	1.069	0.925	0.224	0.159	1.040	2.739	2.483	4.183
20	−0.001	−0.006	−0.109	0.842	1.020	0.823	0.243	0.158	0.867	2.699	2.520	3.354

Para avaliar o comportamento da distribuição dos resíduos sob a modelagem do parâmetro de dispersão consideramos o modelo de regressão simplex não linear no submodelo da média e no submodelo da dispersão dado por

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}, \\ \log(\sigma_t^2) &= \gamma_1 + z_{t2}^{\gamma_2}, \end{aligned} \tag{3.9}$$

Tabela 2 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 20$ e $\mu \in (0.23, 0.85)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.037	−0.040	0.225	1.158	1.322	1.135	0.064	0.075	1.395	2.420	2.192	4.746
2	0.005	0.001	−0.180	0.862	1.228	0.695	−0.066	−0.035	1.444	2.420	2.029	5.470
3	0.008	0.010	0.103	1.064	1.217	1.113	0.028	0.019	2.000	2.942	2.570	8.323
4	−0.005	−0.003	0.075	1.020	1.078	0.951	−0.216	−0.187	0.325	2.407	2.392	3.241
5	−0.001	0.002	0.073	1.066	1.172	1.002	−0.182	−0.139	0.456	2.284	2.139	2.983
6	−0.010	−0.011	−0.061	0.915	1.076	0.900	−0.320	−0.216	0.402	3.170	2.899	3.887
7	0.009	0.013	0.007	0.960	1.143	0.909	−0.115	−0.070	0.917	2.525	2.295	4.174
8	−0.010	−0.001	−0.049	0.929	1.115	0.894	−0.202	−0.117	0.934	2.859	2.489	4.027
9	0.000	0.001	0.329	1.187	1.302	0.990	−0.013	−0.013	0.752	1.908	1.844	2.978
10	−0.024	−0.028	0.133	1.090	1.229	1.050	0.016	0.030	1.395	2.573	2.395	5.307
11	−0.022	−0.029	−0.116	0.913	1.321	0.749	−0.088	−0.021	0.940	2.271	1.945	4.320
12	0.006	−0.002	−0.267	0.736	1.209	0.656	0.259	0.064	1.907	3.175	2.649	7.094
13	0.033	0.040	−0.001	0.924	1.024	0.898	0.311	0.253	0.869	2.834	2.678	3.565
14	−0.010	−0.015	0.051	1.044	1.160	1.009	0.162	0.143	1.055	2.677	2.529	4.176
15	−0.020	−0.019	0.097	1.064	1.181	0.949	−0.061	−0.034	0.902	2.170	2.182	3.587
16	−0.019	−0.019	0.121	1.094	1.352	0.959	−0.115	−0.067	1.219	2.253	2.203	4.312
17	0.004	0.001	−0.094	0.922	1.142	0.862	0.036	−0.001	1.265	2.686	2.420	4.985
18	0.014	0.009	0.219	1.109	1.424	0.763	−0.044	−0.025	0.498	1.715	1.661	2.630
19	0.030	0.040	−0.004	0.932	1.088	0.901	−0.211	−0.152	1.377	2.938	2.589	5.640
20	−0.001	0.008	−0.105	0.864	1.060	0.797	−0.066	−0.029	1.247	2.838	2.684	4.700

Tabela 3 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 20$ e $\mu \in (0.80, 0.98)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.003	−0.005	−0.066	0.999	1.215	1.006	−0.119	−0.064	0.463	2.363	2.184	2.849
2	−0.028	−0.040	−0.122	0.923	1.221	0.782	−0.077	−0.016	1.029	2.149	2.005	4.029
3	−0.031	−0.036	−0.099	0.919	1.187	0.927	−0.113	−0.039	0.463	2.982	2.451	3.827
4	0.007	0.005	0.060	1.109	1.158	1.088	−0.311	−0.293	−0.139	2.219	2.212	2.518
5	0.001	0.008	−0.029	1.023	1.133	1.018	−0.318	−0.233	−0.004	2.400	2.277	2.551
6	−0.007	−0.006	−0.032	0.976	1.124	0.972	−0.355	−0.267	−0.127	2.688	2.470	2.754
7	0.010	0.020	−0.078	0.911	1.138	0.912	−0.186	−0.109	0.231	2.521	2.303	2.599
8	0.019	0.021	−0.109	0.747	1.015	0.784	−0.196	−0.106	0.313	3.517	2.894	3.121
9	−0.036	−0.040	0.121	1.108	1.228	0.996	−0.109	−0.079	0.317	1.910	1.937	2.656
10	−0.005	−0.007	0.091	1.157	1.275	1.140	−0.186	−0.157	0.493	2.258	2.172	3.196
11	−0.026	−0.051	−0.054	0.971	1.338	0.858	−0.096	−0.028	0.568	2.017	1.855	3.027
12	0.012	0.020	−0.080	0.910	1.151	0.804	−0.220	−0.140	1.030	2.447	2.349	4.328
13	0.023	0.025	−0.003	1.005	1.091	1.004	−0.435	−0.376	−0.224	2.479	2.401	2.479
14	−0.014	−0.014	−0.004	1.107	1.215	1.099	−0.302	−0.250	0.024	2.383	2.293	2.556
15	0.016	0.016	0.021	1.061	1.192	1.042	−0.146	−0.115	0.174	2.124	2.089	2.352
16	−0.015	−0.022	−0.062	0.955	1.221	0.943	−0.110	−0.077	0.512	2.559	2.302	3.321
17	−0.029	−0.024	−0.119	0.865	1.092	0.860	−0.239	−0.119	0.200	2.672	2.542	2.813
18	0.019	0.032	0.321	1.217	1.506	0.901	−0.102	−0.074	0.245	1.560	1.544	2.521
19	0.008	0.013	−0.046	0.994	1.163	0.988	−0.218	−0.140	0.428	2.547	2.319	3.186
20	−0.008	−0.016	−0.079	0.917	1.108	0.905	−0.233	−0.166	0.278	2.572	2.408	2.791

Tabela 4 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 50$ e $\mu \in (0.02, 0.15)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.005	−0.009	0.210	1.114	1.203	0.944	0.121	0.095	0.898	1.972	1.956	3.461
2	0.020	0.028	−0.079	0.870	1.023	0.825	0.087	0.083	0.760	2.216	2.235	2.958
3	0.033	0.036	0.211	1.106	1.162	1.110	0.138	0.135	1.706	2.568	2.524	6.323
4	−0.001	−0.001	0.060	0.988	1.024	0.930	0.180	0.177	0.589	2.285	2.318	2.876
5	0.022	0.029	0.123	1.112	1.257	1.079	0.152	0.100	0.969	2.289	2.106	3.484
6	0.001	0.000	0.010	0.941	0.988	0.919	0.236	0.223	0.661	2.573	2.599	3.167
7	−0.022	−0.015	0.012	0.988	1.277	0.791	0.094	0.081	1.378	2.267	2.038	4.673
8	−0.015	−0.017	0.022	0.902	0.920	0.832	0.282	0.266	1.019	2.590	2.571	4.025
9	0.000	−0.010	0.299	1.193	1.524	0.930	0.042	0.034	0.804	1.825	1.661	2.885
10	−0.007	−0.012	0.178	1.125	1.284	1.103	0.101	0.072	1.619	2.475	2.276	6.002
11	0.007	0.011	−0.021	0.875	0.999	0.807	0.087	0.083	0.572	2.260	2.319	2.883
12	−0.002	−0.010	−0.584	0.347	1.417	0.275	0.043	0.011	3.749	5.734	2.157	22.358
13	0.005	0.005	0.008	0.917	0.965	0.896	0.217	0.191	0.631	2.452	2.500	3.018
14	0.017	0.016	0.025	0.965	1.118	0.900	0.104	0.078	1.177	2.493	2.422	4.439
15	−0.013	−0.014	0.321	1.199	1.439	1.015	0.066	0.060	0.989	1.973	1.847	3.179
16	−0.005	−0.007	0.596	1.355	1.454	1.099	0.045	0.034	0.717	1.712	1.702	2.708
17	0.017	0.042	−0.229	0.820	1.294	0.709	0.054	0.012	2.433	3.089	2.343	10.345
18	0.001	−0.002	0.025	0.998	1.153	0.883	0.086	0.068	0.581	1.897	1.944	2.597
19	0.005	0.002	−0.001	0.958	1.181	0.933	0.064	0.030	1.904	2.972	2.361	7.003
20	0.029	0.041	−0.220	0.818	1.144	0.695	0.067	−0.008	1.853	2.749	2.313	6.402

Tabela 5 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 50$ e $\mu \in (0.23, 0.85)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	0.003	−0.001	0.180	1.021	1.117	0.815	0.016	0.006	1.277	2.329	2.283	4.839
2	−0.015	−0.020	−0.081	0.904	1.001	0.680	−0.046	−0.049	1.264	2.180	2.236	4.399
3	−0.019	−0.019	0.190	1.106	1.178	1.066	0.059	0.055	1.838	2.542	2.466	7.026
4	−0.009	−0.009	0.090	0.968	1.005	0.849	−0.092	−0.083	0.894	2.458	2.494	3.919
5	−0.009	−0.008	0.426	1.241	1.382	1.168	−0.065	−0.059	0.844	2.137	1.924	2.801
6	−0.006	−0.003	0.019	0.912	0.968	0.858	−0.174	−0.148	0.993	2.844	2.837	4.634
7	−0.018	−0.022	−0.064	0.939	1.315	0.753	−0.011	0.006	1.595	2.458	2.022	5.478
8	0.010	0.013	0.008	0.829	0.860	0.753	−0.157	−0.130	0.866	2.892	2.878	4.210
9	0.018	0.027	0.323	1.203	1.586	0.956	−0.014	−0.013	0.639	1.882	1.645	2.371
10	0.035	0.041	0.429	1.252	1.362	1.289	−0.008	−0.016	1.485	2.382	2.213	4.750
11	−0.007	−0.011	−0.043	0.833	0.942	0.705	−0.133	−0.123	0.857	2.566	2.599	4.103
12	−0.006	−0.027	−0.514	0.523	1.442	0.349	0.084	−0.015	4.277	4.169	2.079	30.716
13	−0.002	0.001	0.009	0.890	0.945	0.805	0.137	0.156	1.123	2.650	2.694	4.406
14	0.005	0.005	0.078	0.992	1.113	0.860	0.054	0.048	1.584	2.568	2.513	5.740
15	0.000	0.003	0.057	1.034	1.382	0.923	0.008	0.028	1.420	2.504	2.106	4.307
16	0.014	0.017	0.497	1.301	1.435	1.046	−0.054	−0.047	0.785	1.774	1.706	2.763
17	−0.025	−0.050	−0.344	0.710	1.338	0.579	0.005	0.033	2.732	3.603	2.430	12.302
18	−0.014	−0.011	0.113	1.059	1.162	0.755	−0.039	−0.020	0.783	1.819	1.859	2.967
19	−0.030	−0.035	0.094	1.046	1.259	1.042	−0.074	−0.016	1.829	2.870	2.251	6.074
20	−0.008	−0.012	−0.196	0.848	1.123	0.662	0.001	0.021	1.799	2.678	2.263	6.228

Tabela 6 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 50$ e $\mu \in (0.80, 0.98)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta,\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	0.002	0.002	0.063	1.024	1.149	0.954	−0.113	−0.097	0.500	2.151	2.146	2.923
2	0.004	0.005	−0.005	0.941	1.007	0.807	−0.111	−0.096	0.797	2.102	2.142	3.179
3	−0.015	−0.015	0.011	1.047	1.176	1.047	−0.072	−0.063	0.771	2.468	2.350	4.258
4	0.007	0.009	0.102	1.030	1.069	0.975	−0.212	−0.192	0.270	2.321	2.351	3.041
5	0.006	0.003	0.141	1.163	1.291	1.172	−0.094	−0.073	0.717	2.215	2.000	3.098
6	0.024	0.024	0.045	0.963	1.024	0.952	−0.209	−0.202	0.224	2.649	2.675	2.995
7	0.003	0.009	−0.143	0.898	1.295	0.802	−0.067	−0.020	1.248	2.454	2.010	4.451
8	−0.011	−0.012	−0.040	0.856	0.908	0.855	−0.128	−0.126	0.087	2.609	2.546	2.628
9	−0.014	−0.023	0.391	1.238	1.598	0.874	−0.009	0.008	0.525	1.658	1.540	2.518
10	−0.008	−0.011	0.406	1.259	1.352	1.250	−0.092	−0.083	1.329	2.268	2.156	4.593
11	0.001	0.001	−0.022	0.868	0.953	0.801	−0.180	−0.164	0.341	2.395	2.432	2.868
12	0.024	0.038	−0.224	0.823	1.469	0.542	−0.046	−0.046	1.690	2.253	1.846	6.737
13	0.008	0.007	0.011	0.957	1.014	0.932	−0.224	−0.201	0.188	2.421	2.450	2.709
14	0.001	0.001	0.051	1.032	1.144	0.987	−0.105	−0.084	1.108	2.464	2.372	4.249
15	0.008	0.009	−0.101	0.922	1.296	0.832	−0.038	−0.002	1.485	2.593	2.213	5.278
16	0.013	0.015	0.216	1.188	1.374	1.086	−0.036	−0.025	0.981	1.974	1.821	3.711
17	0.011	0.019	−0.370	0.675	1.256	0.578	−0.025	−0.039	2.666	3.477	2.495	12.492
18	−0.028	−0.030	0.106	1.058	1.138	0.852	−0.080	−0.066	0.552	1.836	1.876	2.731
19	0.017	0.022	0.024	0.995	1.253	1.000	−0.062	−0.062	1.853	2.883	2.312	6.788
20	−0.020	−0.020	−0.212	0.838	1.131	0.719	−0.021	−0.001	1.678	2.597	2.258	5.916

Tabela 7 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 100$ e $\mu \in (0.02, 0.15)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.005	−0.002	0.138	1.088	1.193	0.974	0.165	0.133	0.845	2.066	2.036	3.302
2	−0.001	−0.001	−0.050	0.882	0.989	0.861	0.150	0.115	0.444	2.229	2.212	2.452
3	0.010	0.010	0.166	1.140	1.182	1.142	0.252	0.245	1.393	2.529	2.480	5.333
4	0.004	0.005	0.022	0.930	0.956	0.909	0.349	0.344	0.581	2.622	2.646	2.974
5	−0.022	−0.021	0.038	1.195	1.286	1.190	0.316	0.269	0.783	2.313	2.125	2.977
6	0.002	0.002	0.012	0.923	0.954	0.918	0.419	0.399	0.595	2.904	2.912	3.160
7	−0.014	−0.017	0.021	1.004	1.286	0.839	0.087	0.060	1.382	2.243	1.943	4.677
8	0.013	0.014	0.019	0.855	0.865	0.831	0.463	0.451	0.777	3.087	3.086	3.589
9	−0.003	−0.006	0.198	1.140	1.539	0.963	0.029	0.018	1.002	2.016	1.711	3.270
10	0.004	−0.001	0.170	1.187	1.361	1.184	0.149	0.107	1.364	2.455	2.195	4.876
11	−0.005	−0.003	−0.016	0.832	0.922	0.808	0.252	0.206	0.482	2.574	2.657	2.825
12	0.002	0.007	−0.564	0.317	1.463	0.292	−0.090	−0.039	2.975	6.009	2.065	15.054
13	0.020	0.020	0.016	0.891	0.921	0.885	0.388	0.370	0.555	2.702	2.736	2.941
14	−0.007	−0.006	−0.002	0.958	1.127	0.903	0.130	0.112	0.843	2.555	2.494	3.625
15	−0.008	0.001	0.269	1.187	1.485	1.030	0.081	0.072	1.034	2.020	1.808	3.328
16	0.000	−0.001	0.592	1.371	1.457	1.143	0.072	0.062	0.731	1.719	1.696	2.671
17	0.006	−0.006	−0.258	0.787	1.292	0.733	0.131	−0.017	2.336	3.521	2.449	9.321
18	0.010	0.011	0.023	1.017	1.129	0.964	0.111	0.084	0.378	1.898	1.936	2.222
19	0.011	0.020	0.036	1.013	1.279	1.058	0.103	0.051	1.878	3.058	2.237	6.532
20	0.005	0.008	−0.268	0.778	1.167	0.685	0.087	0.011	1.969	3.019	2.292	7.021

Tabela 8 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 100$ e $\mu \in (0.23, 0.85)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	0.010	0.010	0.179	1.009	1.118	0.868	0.001	−0.005	1.245	2.478	2.396	4.807
2	−0.007	−0.005	−0.038	0.898	0.975	0.772	−0.075	−0.066	0.829	2.200	2.266	3.360
3	0.005	0.005	0.245	1.154	1.211	1.146	0.005	0.006	1.520	2.500	2.439	5.713
4	0.011	0.013	0.074	0.892	0.921	0.845	−0.233	−0.209	0.427	2.983	3.024	3.994
5	0.003	−0.005	0.456	1.327	1.417	1.314	−0.126	−0.108	0.601	2.089	1.920	2.646
6	−0.018	−0.021	−0.005	0.861	0.898	0.848	−0.335	−0.331	0.254	3.375	3.354	4.113
7	0.013	0.009	−0.034	0.963	1.323	0.788	−0.072	−0.054	1.548	2.424	1.982	5.328
8	−0.006	−0.006	0.004	0.814	0.832	0.779	−0.343	−0.330	0.179	3.274	3.233	3.722
9	0.016	0.011	0.238	1.152	1.603	0.960	0.003	−0.017	0.785	2.055	1.717	2.604
10	0.002	0.002	0.509	1.309	1.426	1.375	0.018	0.011	1.406	2.364	2.160	4.415
11	−0.008	−0.005	−0.042	0.767	0.851	0.713	−0.303	−0.207	0.181	3.101	3.119	3.598
12	−0.003	−0.012	−0.538	0.494	1.501	0.318	0.055	0.000	4.306	4.514	2.020	31.406
13	0.001	0.001	0.003	0.840	0.875	0.813	0.298	0.286	0.811	3.062	3.104	3.888
14	−0.004	−0.006	0.074	0.965	1.088	0.874	0.092	0.070	1.517	2.762	2.693	5.631
15	0.002	0.016	0.006	1.001	1.417	0.918	−0.012	−0.015	1.555	2.675	2.070	4.923
16	−0.023	−0.028	0.545	1.333	1.462	1.086	−0.026	−0.012	0.773	1.748	1.681	2.733
17	−0.003	−0.013	−0.355	0.696	1.371	0.603	0.122	0.008	3.008	3.993	2.466	15.251
18	−0.011	−0.016	0.103	1.058	1.131	0.881	−0.066	−0.051	0.560	1.873	1.873	2.711
19	0.022	0.025	0.173	1.096	1.346	1.167	−0.049	−0.048	1.657	2.928	2.180	5.093
20	0.012	0.007	−0.229	0.813	1.152	0.664	0.033	0.008	1.972	2.859	2.319	7.126

Tabela 9 – Médias, erros-padrão, assimetrias e curtoses empíricas dos resíduos ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). 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 $\log(\sigma_t^2) = \gamma_1+\gamma_2 z_{t2}+\gamma_3 z_{t3}+\gamma_4 z_{t4}+\gamma_5 z_{t5}$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, 20$, $\lambda = 100$ e $\mu \in (0.80, 0.98)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	0.021	0.025	0.024	1.020	1.165	0.988	−0.179	−0.120	0.273	2.198	2.198	2.592
2	−0.009	−0.007	−0.010	0.932	0.983	0.862	−0.234	−0.200	0.234	2.200	2.233	2.558
3	−0.008	−0.008	−0.005	1.065	1.171	1.061	−0.151	−0.139	0.168	2.453	2.376	2.781
4	−0.001	−0.002	0.036	0.976	1.004	0.955	−0.325	−0.306	−0.111	2.582	2.617	2.937
5	0.011	0.008	0.103	1.214	1.302	1.224	−0.228	−0.203	0.236	2.180	2.028	2.639
6	0.003	0.004	0.008	0.960	0.996	0.956	−0.445	−0.421	−0.307	2.917	2.908	2.979
7	0.005	0.007	−0.159	0.901	1.302	0.860	−0.064	−0.032	1.006	2.458	2.042	3.838
8	0.001	0.003	−0.014	0.867	0.897	0.868	−0.291	−0.265	−0.214	2.685	2.615	2.632
9	−0.018	−0.020	0.379	1.237	1.621	0.889	−0.027	−0.001	0.452	1.663	1.543	2.401
10	0.000	0.000	0.369	1.312	1.405	1.349	−0.118	−0.109	1.164	2.211	2.085	4.186
11	−0.017	−0.020	−0.023	0.838	0.904	0.811	−0.329	−0.305	−0.100	2.646	2.692	2.825
12	0.002	0.005	−0.250	0.805	1.502	0.546	−0.055	−0.037	1.804	2.374	1.848	7.062
13	0.008	0.009	0.006	0.931	0.969	0.923	−0.424	−0.385	−0.270	2.653	2.690	2.734
14	0.024	0.025	0.062	1.043	1.166	1.015	−0.207	−0.177	0.654	2.529	2.444	3.330
15	−0.012	−0.017	−0.184	0.874	1.298	0.825	−0.019	0.002	1.274	2.780	2.262	4.771
16	−0.007	−0.005	0.165	1.219	1.392	1.147	−0.057	−0.048	0.602	1.931	1.786	2.892
17	0.000	−0.004	−0.355	0.663	1.269	0.615	−0.071	−0.053	2.031	3.669	2.655	8.493
18	0.014	0.014	0.086	1.054	1.113	0.966	−0.203	−0.184	0.167	1.920	1.934	2.382
19	−0.002	0.013	−0.043	1.010	1.336	1.027	−0.102	−0.019	1.834	3.019	2.271	6.772
20	−0.029	−0.037	−0.265	0.796	1.156	0.702	−0.086	−0.015	1.535	2.784	2.313	5.594

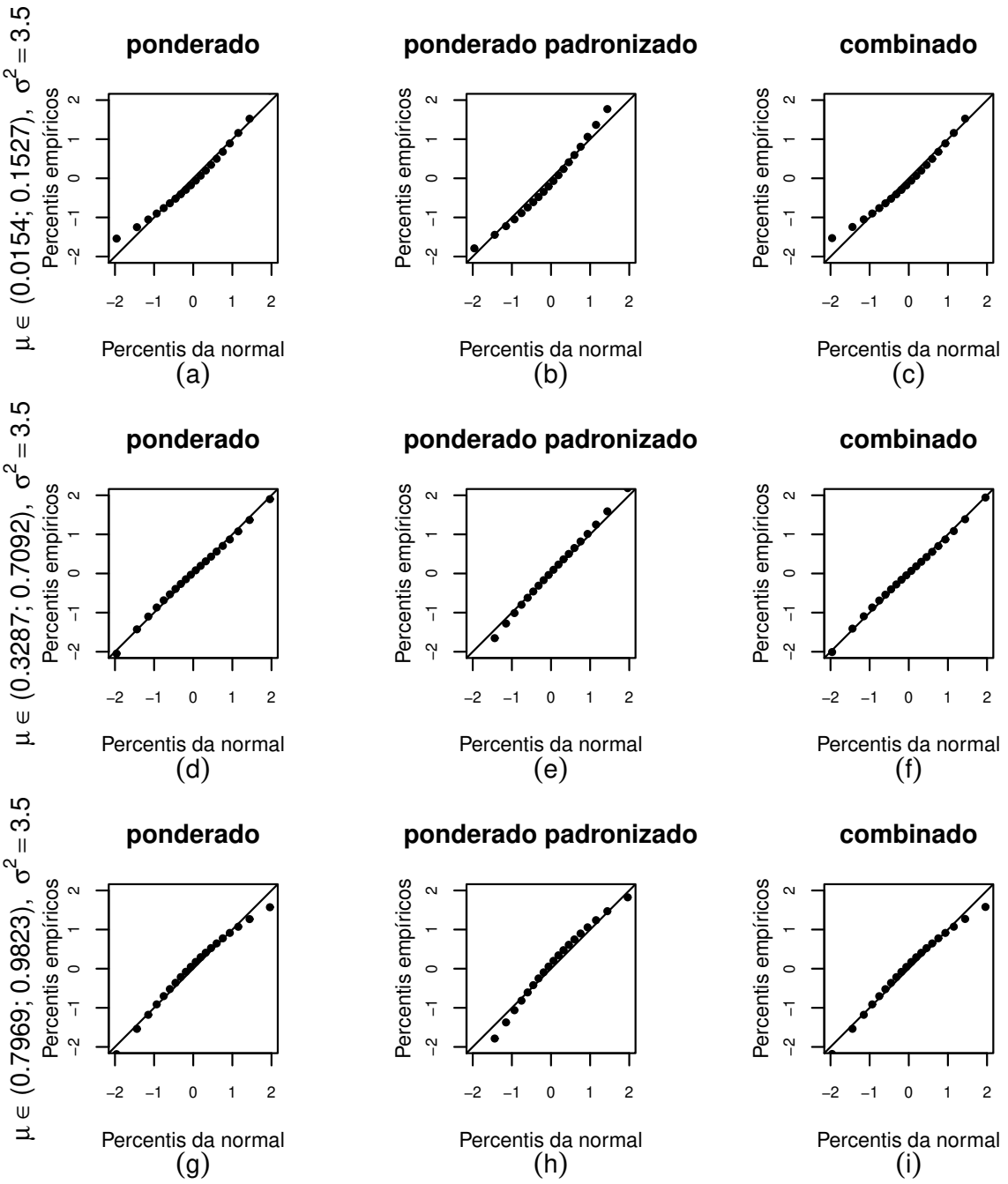

Figura 2 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}+\beta_3 x_{t3}+\beta_4 x_{t4}+\beta_5 x_{t5}$, $t = 1, \ldots, 20$, $\sigma^2 = 3.5$.

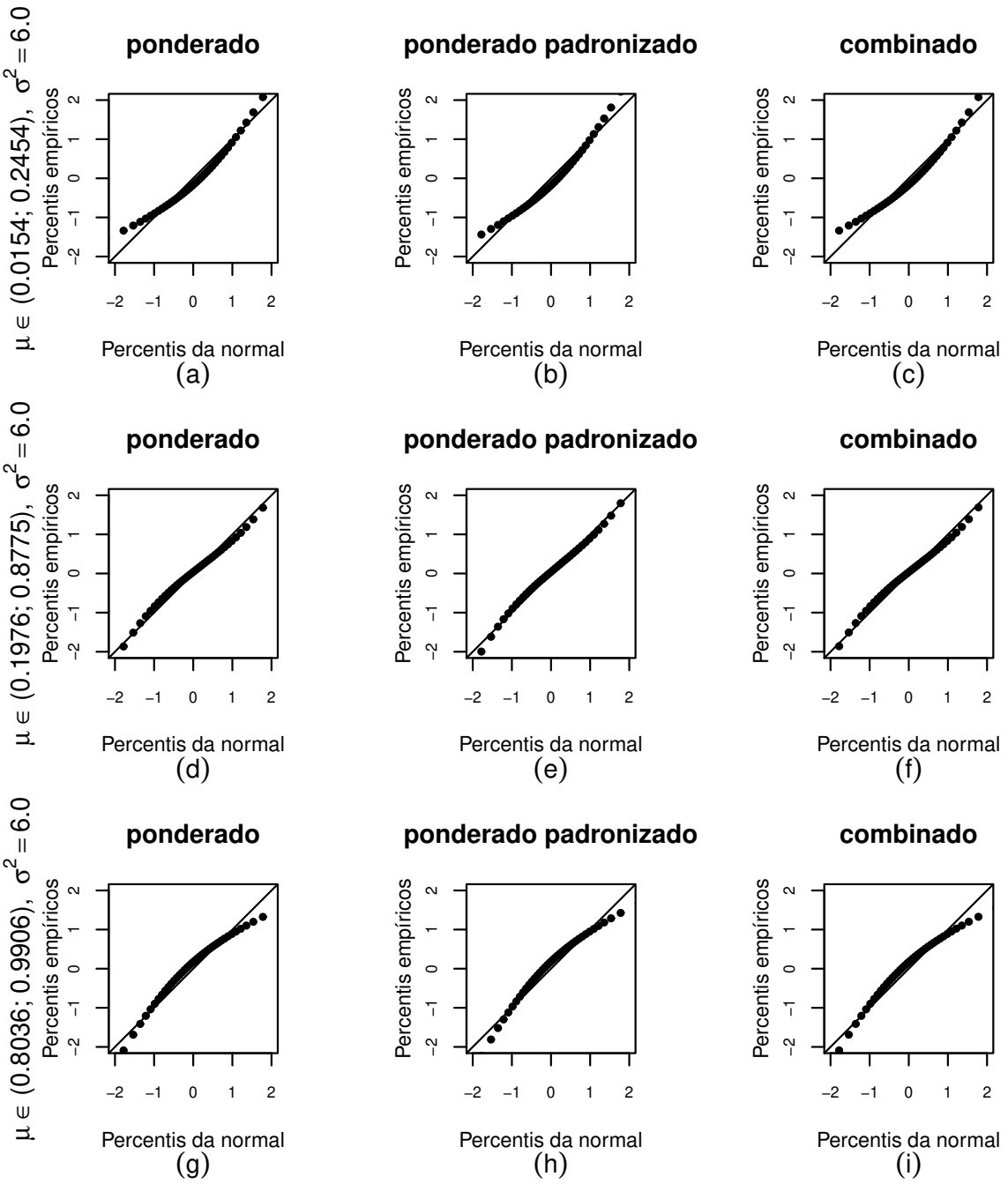

Figura 3 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}+\beta_3 x_{t3}+\beta_4 x_{t4}+\beta_5 x_{t5}$, $t = 1, \ldots, 40$, $\sigma^2 = 6.0$.

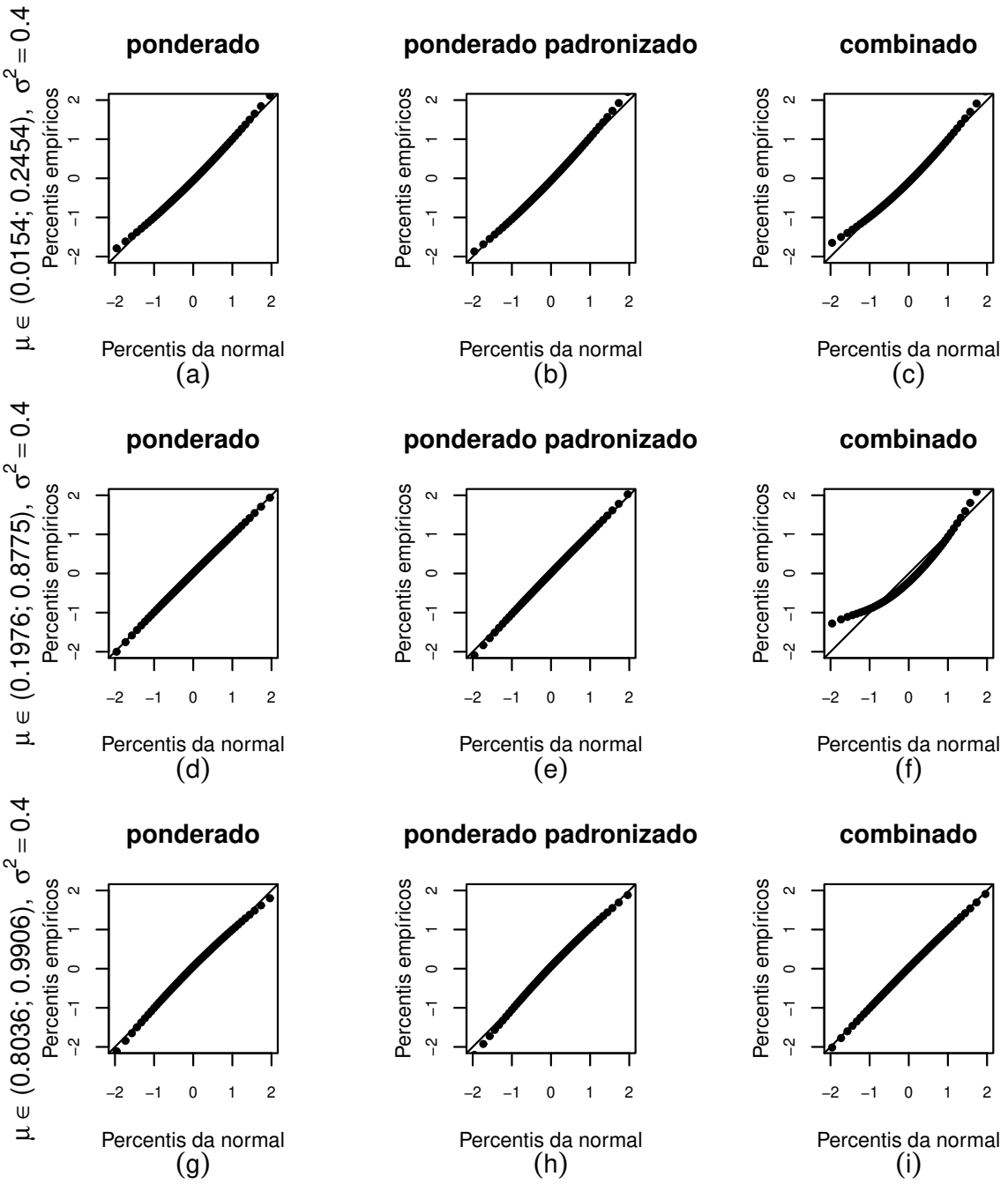

Figura 4 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}+\beta_3 x_{t3}+\beta_4 x_{t4}+\beta_5 x_{t5}$, $t = 1, \ldots, 60$, $\sigma^2 = 0.4$.

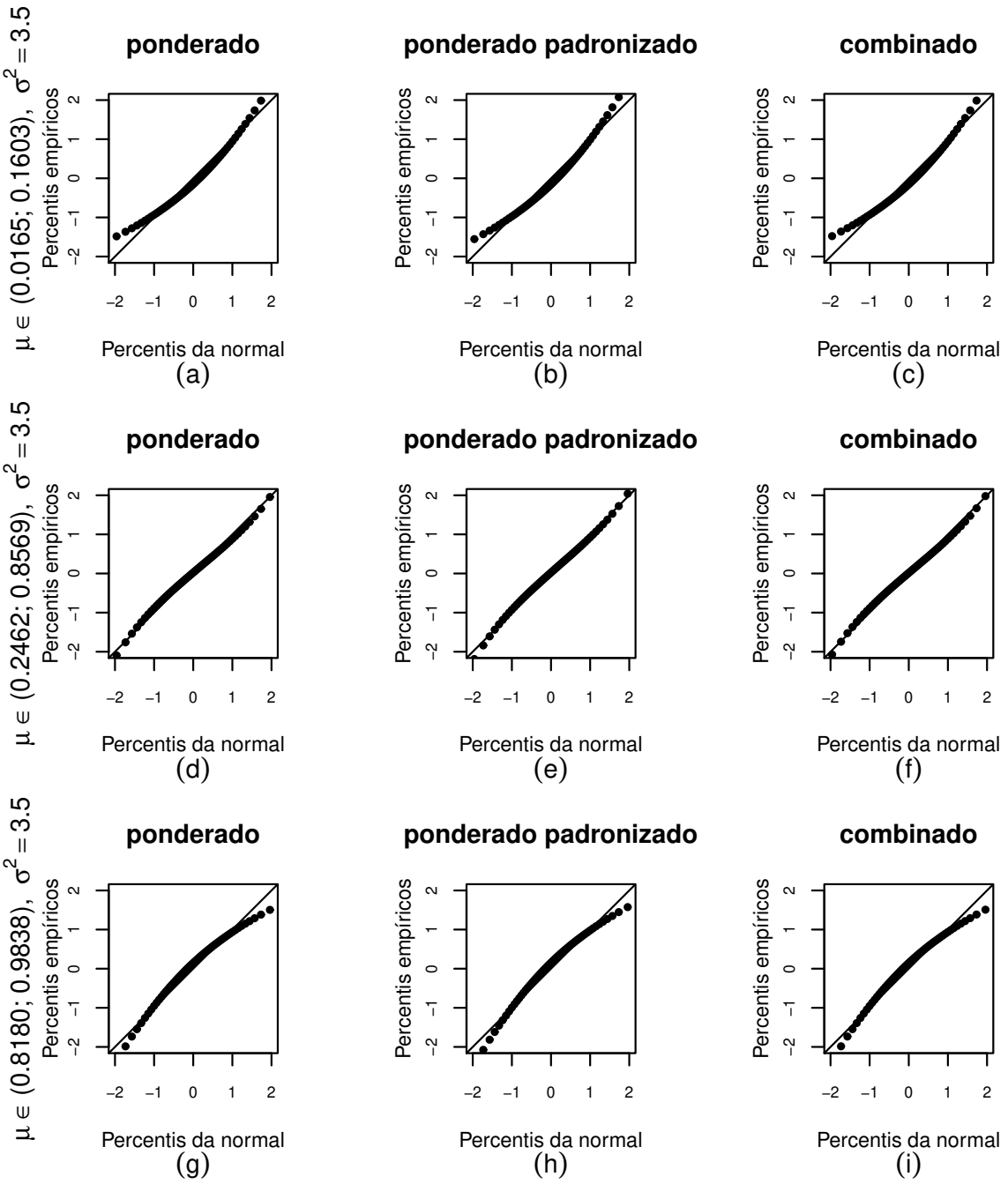

Figura 5 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}+\beta_3 x_{t3}+\beta_4 x_{t4}+\beta_5 x_{t5}$, $t = 1, \ldots, 60$, $\sigma^2 = 3.5$.

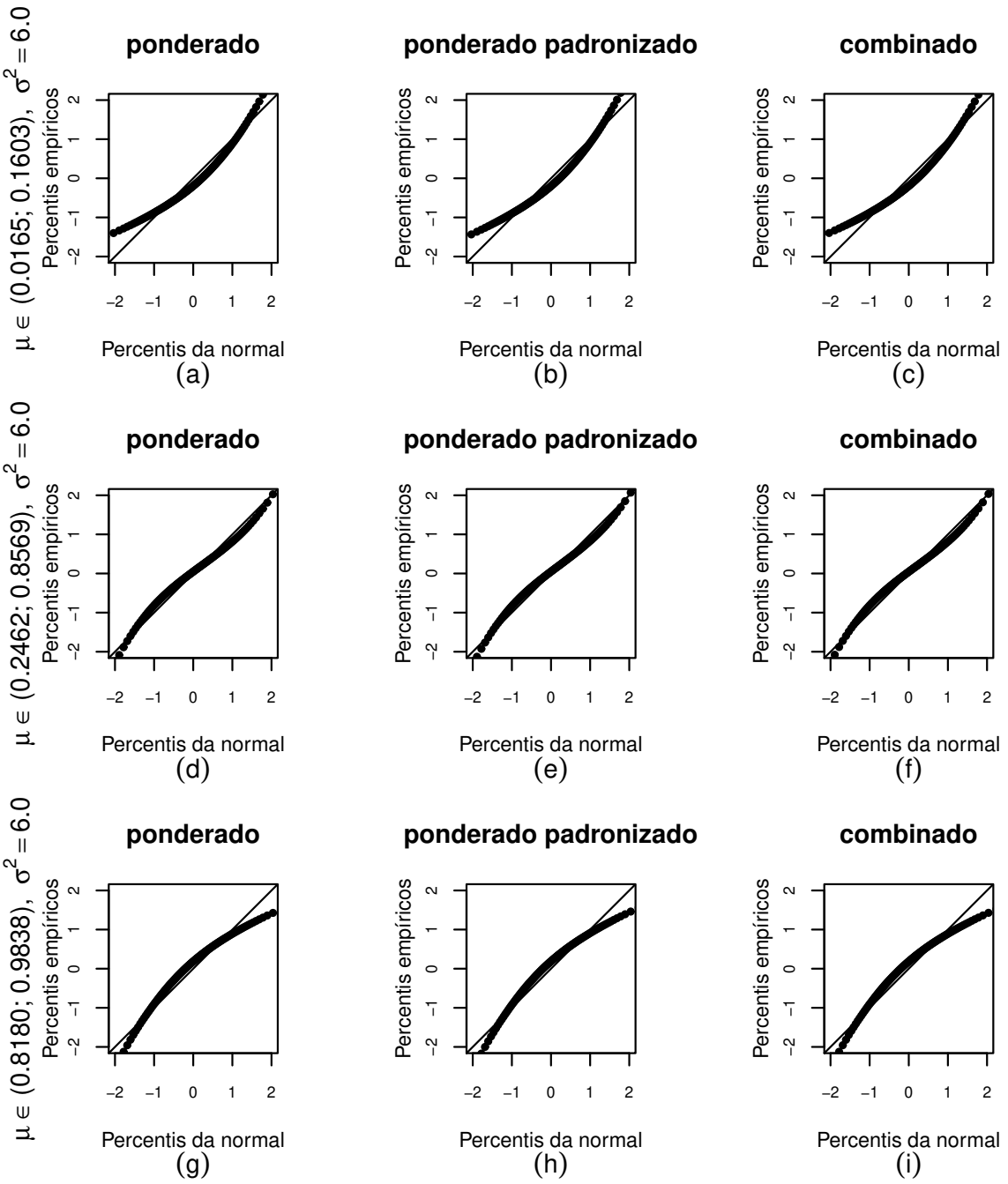

Figura 6 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}+\beta_3 x_{t3}+\beta_4 x_{t4}+\beta_5 x_{t5}$, $t = 1, \ldots, 120$, $\sigma^2 = 6.0$.

em que $t = 1, \ldots, n$. Aqui consideramos 10000 réplicas de Monte Carlo. Nós medimos o grau de dispersão não constante por $\lambda = \sigma^2_{t,max}/\sigma^2_{t,min}$, $t = 1 \ldots, n$. Consideramos os três cenários para a média da variável resposta como definidos anteriormente, ou seja, $\mu \in (0.02, 0.33)(\beta = (-2.4, 1.4, -1.5, -1.7))$. $\mu \in (0.20, 0.87)(\beta = (-1.7, -1.8, 1.2, -1.3))$ e $\mu \in (0.79, 0.99)(\beta = (2.1, -1.5, -1.6, -1.2))$. As realizações das covariadas foram geradas através das seguintes distribuições: $x_{t2} \sim U(-0.5, 0.5)$, $x_{t3} \sim U(0, 1)$, $x_{t4} \sim U(-0.5, 0.5)$ e $z_{t2} \sim U(-0.5, 0.5)$, e foram mantidos fixas para cada réplica.

As Tabelas 10, 11 e 12 mostram as médias das estatísticas descritivas referentes aos resíduos r^β, r^β_p e $r^{\beta\gamma}$ considerando os três cenários para a média da variável resposta e $\lambda = 20$ ($\gamma = (-1.3, -1.82)$). Percebemos que as médias dos três resíduos estão próximos de zero, os erros-padrão do valor um e a similaridade também ocorre com a assimetria e curtose, que apresentam valores um pouco maiores que 0 e 3, respectivamente, valores referentes a assimetria e a curtose da distribuição normal padrão. O mesmo acontece quando aumenta o grau de dispersão não constante como pode ser observado nas Tabelas 13 a 18, em que é considerado os valores $\lambda = 50$ ($\gamma = (-1.3, 3.75)$) e $\lambda = 150$ ($\gamma = (-1.3, -2.44)$). No entanto, observa-se que para $\lambda = 150$ algumas observações mostram um valor relativamente alto para a assimetria e curtose, como por exemplo, na Tabela 16 as observações 9 e 17.

Para comparar os quantis empíricos dos resíduos com os quantis teóricos da distribuição normal padrão construímos os gráficos normais de probabilidade utilizando o modelo dado em 3.9 (Figuras 7 e 8). Na Figura 7 o tamanho amostral utilizado foi $n = 40$ e $\lambda = 20$. Mais uma vez, verificou-se que os três resíduos são bastante semelhantes e a assimetria é diferente de zero principalmente quando a média da variável resposta está próxima de zero e próxima de um. Além disso, vê-se que a assimetria é positiva quando $\mu \approx 0$ e negativa quando $\mu \approx 1$. O mesmo acontece na Figura 8, em que $n = 80$ e $\lambda = 50$. Nas Figuras 9 - 11 consideramos o submodelo da média da variável dado em 3.9 com parâmetro de dispersão constante $\sigma^2 = 0.4$, 3.5 e 6.0, respectivamente, e $n = 120$. Observamos novamente a semelhança entre os três resíduos que confirmam uma leve assimetria em suas distribuições. Como comentado anteriormente, deve-se utilizar como limites de detecção de pontos aberrantes os quantis empíricos dos resíduos gerados com base em suas distribuições estimadas para a construção das bandas do envelope dos gráficos normais

de probabilidade.

Tabela 10 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 20$ e $\mu \in (0.02, 0.15)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.010	−0.013	−0.026	0.981	1.192	0.934	0.222	0.206	0.951	2.416	2.386	3.449
2	−0.016	−0.023	−0.106	0.734	1.167	0.706	0.097	0.062	0.565	2.527	2.395	2.838
3	−0.001	−0.001	0.021	1.073	1.193	1.053	0.191	0.192	0.808	2.618	2.621	3.420
4	0.012	0.013	0.022	1.145	1.195	1.142	0.587	0.586	0.733	3.064	3.057	3.249
5	−0.005	−0.005	0.002	1.116	1.161	1.112	0.616	0.615	0.776	3.546	3.516	3.779
6	−0.024	−0.052	−0.028	0.527	1.146	0.527	−0.096	−0.093	−0.081	2.746	2.645	2.738
7	0.009	0.009	0.010	1.039	1.058	1.039	0.981	0.973	1.030	4.481	4.435	4.569
8	−0.011	−0.012	−0.009	1.021	1.196	0.973	0.156	0.155	1.005	2.393	2.382	3.622
9	−0.009	−0.009	−0.009	0.859	0.864	0.859	1.298	1.295	1.309	5.081	5.081	5.114
10	0.034	0.036	0.036	1.105	1.166	1.104	0.561	0.560	0.591	3.103	3.093	3.134
11	−0.018	−0.028	−0.177	0.719	1.173	0.636	0.042	0.029	0.938	2.393	2.357	3.428
12	−0.002	−0.002	−0.002	0.991	1.030	0.991	0.940	0.914	0.950	3.903	3.798	3.919
13	−0.015	−0.016	−0.009	1.110	1.156	1.108	0.609	0.604	0.749	3.595	3.569	3.784
14	0.009	0.009	0.018	1.143	1.179	1.144	0.617	0.617	0.819	3.646	3.643	3.949
15	0.003	0.003	0.002	0.813	0.825	0.813	1.069	1.058	1.093	4.759	4.722	4.792
16	0.028	0.030	0.056	1.118	1.225	1.100	0.304	0.298	0.869	2.437	2.422	3.274
17	0.005	0.005	0.005	0.887	0.908	0.887	1.206	1.184	1.214	4.741	4.679	4.755
18	0.030	0.036	0.018	0.978	1.187	0.951	0.121	0.116	0.678	2.258	2.260	2.850
19	0.012	0.012	0.013	1.082	1.106	1.082	0.863	0.859	0.887	3.753	3.734	3.797
20	−0.018	−0.019	−0.017	1.068	1.111	1.067	0.897	0.888	0.944	3.983	3.935	4.075

Tabela 11 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 20$ e $\mu \in (0.23, 0.85)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.005	−0.005	−0.003	1.011	1.176	0.970	−0.093	−0.090	0.984	2.411	2.392	3.842
2	−0.010	−0.013	−0.221	0.763	1.111	0.709	0.115	0.143	1.570	2.851	2.676	6.023
3	0.001	0.001	0.020	1.053	1.174	1.039	0.023	0.023	0.683	2.622	2.621	3.308
4	−0.007	−0.008	0.009	1.137	1.185	1.134	0.465	0.465	0.758	3.087	3.088	3.503
5	0.003	0.004	0.012	1.138	1.164	1.136	0.662	0.663	0.828	3.614	3.603	3.873
6	0.008	0.014	−0.020	0.662	1.119	0.668	−0.062	−0.054	0.119	2.677	2.634	2.760
7	0.014	0.014	0.015	1.056	1.077	1.057	−0.010	−0.011	0.080	3.343	3.327	3.361
8	−0.009	−0.009	0.017	1.065	1.181	1.030	0.142	0.140	1.053	2.439	2.429	3.868
9	−0.011	−0.011	−0.011	0.830	0.838	0.830	0.260	0.260	0.312	3.319	3.326	3.377
10	0.005	0.005	0.012	1.085	1.153	1.087	−0.507	−0.506	−0.320	3.231	3.222	3.126
11	−0.001	0.000	−0.155	0.707	1.146	0.634	0.002	−0.008	0.825	2.360	2.348	3.113
12	0.001	0.001	−0.002	0.919	1.064	0.920	−0.450	−0.447	−0.397	3.803	3.537	3.776
13	−0.004	−0.004	0.002	1.093	1.157	1.092	0.149	0.145	0.336	3.296	3.273	3.404
14	0.011	0.012	0.018	1.098	1.165	1.097	0.484	0.482	0.687	3.376	3.364	3.613
15	−0.004	−0.004	−0.004	0.827	0.839	0.826	0.288	0.285	0.340	3.384	3.387	3.437
16	0.011	0.012	0.045	1.081	1.191	1.064	−0.035	−0.034	0.891	2.424	2.414	3.724
17	−0.013	−0.013	−0.014	0.866	0.904	0.867	−0.728	−0.705	−0.694	3.960	3.894	3.935
18	0.008	0.009	0.012	0.995	1.153	0.918	0.140	0.137	1.257	2.344	2.346	4.269
19	−0.021	−0.022	−0.018	1.074	1.115	1.072	0.020	0.022	0.154	3.382	3.353	3.429
20	−0.002	−0.003	−0.002	1.002	1.105	1.004	−0.574	−0.565	−0.475	3.439	3.367	3.368

Tabela 12 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 20$ e $\mu \in (0.80, 0.98)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.014	−0.017	−0.022	0.959	1.164	0.958	−0.179	−0.167	−0.010	2.390	2.362	2.404
2	−0.008	−0.009	−0.042	0.702	1.097	0.706	−0.014	0.089	0.140	2.807	2.577	2.840
3	0.013	0.015	0.015	1.055	1.173	1.056	−0.356	−0.355	−0.257	2.636	2.634	2.596
4	−0.001	−0.001	0.007	1.147	1.186	1.149	−0.675	−0.676	−0.551	3.243	3.243	3.110
5	0.017	0.018	0.021	1.139	1.163	1.140	−0.849	−0.845	−0.790	3.854	3.839	3.755
6	0.023	0.033	0.015	0.788	1.136	0.791	−0.231	−0.213	−0.144	2.717	2.685	2.704
7	0.009	0.010	0.010	1.057	1.076	1.057	−0.833	−0.827	−0.825	3.514	3.495	3.503
8	−0.011	−0.011	−0.006	1.074	1.173	1.076	−0.299	−0.294	−0.059	2.463	2.453	2.472
9	−0.009	−0.009	−0.009	0.904	0.911	0.903	−1.120	−1.115	−1.112	4.212	4.199	4.199
10	0.021	0.023	0.022	1.062	1.150	1.063	−0.572	−0.571	−0.541	3.047	3.042	3.017
11	0.006	0.009	−0.064	0.838	1.133	0.826	−0.107	−0.095	0.577	2.361	2.370	2.909
12	−0.004	−0.002	−0.005	0.890	1.062	0.891	−0.495	−0.416	−0.494	3.105	2.888	3.104
13	0.012	0.012	0.012	1.035	1.138	1.036	−0.496	−0.495	−0.485	2.967	2.946	2.958
14	0.025	0.027	0.025	1.060	1.151	1.061	−0.506	−0.503	−0.486	3.013	3.004	2.994
15	−0.007	−0.007	−0.007	0.907	0.915	0.907	−1.166	−1.160	−1.157	4.339	4.321	4.327
16	−0.004	−0.005	0.001	1.081	1.187	1.082	−0.298	−0.294	−0.088	2.401	2.394	2.423
17	−0.016	−0.016	−0.016	0.921	0.947	0.921	−1.129	−1.109	−1.117	4.335	4.274	4.318
18	−0.015	−0.019	−0.032	0.952	1.149	0.942	−0.145	−0.140	0.413	2.351	2.356	2.803
19	−0.024	−0.024	−0.022	1.101	1.133	1.100	−0.935	−0.929	−0.894	3.955	3.925	3.885
20	−0.004	−0.006	−0.004	0.972	1.116	0.972	−0.361	−0.353	−0.357	2.850	2.792	2.848

Tabela 13 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 50$ e $\mu \in (0.02, 0.15)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.010	−0.013	−0.035	0.962	1.197	0.910	0.198	0.185	0.957	2.410	2.381	3.455
2	−0.019	−0.031	−0.117	0.704	1.180	0.675	0.071	0.041	0.553	2.528	2.385	2.817
3	0.000	0.000	0.021	1.073	1.191	1.052	0.186	0.187	0.750	2.592	2.596	3.274
4	0.012	0.012	0.019	1.155	1.197	1.153	0.661	0.661	0.770	3.201	3.197	3.350
5	−0.004	−0.003	0.001	1.131	1.166	1.129	0.723	0.721	0.828	3.830	3.809	3.997
6	−0.025	−0.057	−0.028	0.503	1.147	0.503	−0.097	−0.092	−0.086	2.736	2.654	2.730
7	0.002	0.002	0.003	1.038	1.051	1.038	1.212	1.205	1.233	5.313	5.273	5.359
8	−0.008	−0.009	−0.010	1.012	1.198	0.963	0.150	0.150	1.013	2.383	2.377	3.614
9	−0.006	−0.006	−0.006	0.825	0.827	0.825	1.902	1.899	1.905	7.879	7.864	7.887
10	0.033	0.035	0.034	1.113	1.166	1.112	0.665	0.664	0.685	3.284	3.275	3.309
11	−0.020	−0.033	−0.191	0.699	1.188	0.610	0.009	0.006	0.940	2.384	2.357	3.389
12	0.006	0.007	0.006	1.013	1.037	1.013	1.250	1.228	1.255	4.946	4.857	4.955
13	−0.017	−0.018	−0.013	1.119	1.157	1.118	0.728	0.725	0.816	4.027	4.008	4.160
14	0.008	0.009	0.015	1.157	1.187	1.157	0.703	0.703	0.836	3.949	3.946	4.163
15	−0.002	−0.002	−0.002	0.765	0.772	0.765	1.446	1.439	1.453	6.409	6.372	6.421
16	0.029	0.032	0.055	1.108	1.223	1.088	0.285	0.281	0.856	2.418	2.407	3.243
17	0.003	0.003	0.003	0.862	0.874	0.862	1.723	1.708	1.725	6.979	6.900	6.984
18	0.034	0.042	0.020	0.975	1.194	0.947	0.108	0.105	0.699	2.282	2.279	2.909
19	0.013	0.013	0.014	1.092	1.111	1.092	1.083	1.080	1.096	4.445	4.427	4.474
20	−0.018	−0.018	−0.017	1.084	1.115	1.083	1.106	1.100	1.133	4.698	4.653	4.763

Tabela 14 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 50$ e $\mu \in (0.23, 0.85)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.004	−0.005	−0.008	0.999	1.184	0.955	−0.081	−0.079	1.053	2.402	2.391	3.948
2	−0.016	−0.022	−0.273	0.701	1.113	0.645	0.131	0.193	1.725	3.092	2.794	6.972
3	−0.001	−0.001	0.017	1.052	1.172	1.040	0.026	0.026	0.637	2.607	2.607	3.232
4	−0.001	−0.001	0.012	1.146	1.188	1.145	0.522	0.522	0.743	3.192	3.195	3.503
5	0.003	0.003	0.008	1.148	1.168	1.146	0.753	0.752	0.860	3.863	3.855	4.031
6	0.005	0.010	−0.019	0.642	1.114	0.646	−0.049	−0.045	0.094	2.673	2.631	2.744
7	0.012	0.012	0.013	1.057	1.073	1.057	−0.006	−0.007	0.033	3.699	3.683	3.701
8	−0.008	−0.009	0.016	1.059	1.185	1.024	0.138	0.137	1.087	2.430	2.426	3.986
9	−0.010	−0.011	−0.011	0.777	0.782	0.777	0.336	0.336	0.352	3.819	3.822	3.834
10	0.007	0.007	0.012	1.093	1.156	1.095	−0.576	−0.575	−0.451	3.434	3.425	3.339
11	0.001	0.002	−0.166	0.695	1.160	0.616	−0.029	−0.033	0.833	2.346	2.337	3.089
12	0.010	0.012	0.009	0.943	1.055	0.943	−0.607	−0.601	−0.583	4.341	4.099	4.323
13	−0.001	−0.001	0.003	1.104	1.160	1.104	0.160	0.156	0.275	3.517	3.500	3.572
14	0.013	0.014	0.018	1.113	1.172	1.113	0.564	0.562	0.699	3.623	3.613	3.789
15	0.000	0.000	−0.001	0.773	0.782	0.773	0.373	0.371	0.388	3.954	3.954	3.965
16	0.008	0.009	0.042	1.074	1.194	1.055	−0.035	−0.035	0.913	2.423	2.415	3.783
17	−0.011	−0.011	−0.011	0.833	0.860	0.833	−0.973	−0.955	−0.963	4.925	4.850	4.915
18	0.008	0.010	0.011	0.997	1.166	0.918	0.135	0.133	1.333	2.362	2.361	4.555
19	−0.021	−0.021	−0.019	1.086	1.121	1.086	0.032	0.035	0.104	3.761	3.730	3.778
20	0.001	0.001	0.002	1.020	1.105	1.022	−0.710	−0.703	−0.653	3.826	3.758	3.765

Tabela 15 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 50$ e $\mu \in (0.80, 0.98)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.015	−0.018	−0.025	0.941	1.178	0.940	−0.145	−0.136	0.035	2.351	2.334	2.376
2	−0.005	−0.005	−0.046	0.621	1.104	0.625	0.057	0.159	0.228	3.071	2.678	3.115
3	0.014	0.015	0.016	1.055	1.175	1.056	−0.357	−0.356	−0.265	2.610	2.609	2.574
4	−0.001	−0.001	0.005	1.156	1.190	1.157	−0.744	−0.745	−0.650	3.379	3.383	3.262
5	0.020	0.021	0.022	1.147	1.165	1.147	−0.991	−0.989	−0.952	4.262	4.255	4.184
6	0.021	0.032	0.014	0.760	1.134	0.763	−0.230	−0.215	−0.159	2.736	2.704	2.726
7	0.010	0.010	0.010	1.056	1.069	1.056	−1.090	−1.085	−1.086	4.275	4.256	4.268
8	−0.011	−0.012	−0.006	1.068	1.179	1.070	−0.279	−0.275	−0.026	2.432	2.426	2.454
9	−0.008	−0.008	−0.008	0.866	0.870	0.866	−1.649	−1.645	−1.647	6.359	6.341	6.353
10	0.018	0.019	0.019	1.066	1.146	1.067	−0.661	−0.660	−0.640	3.190	3.183	3.165
11	0.008	0.010	−0.070	0.828	1.154	0.815	−0.088	−0.082	0.637	2.361	2.358	2.977
12	−0.006	−0.005	−0.007	0.938	1.056	0.938	−0.675	−0.610	−0.675	3.295	3.152	3.294
13	0.015	0.016	0.015	1.047	1.132	1.047	−0.604	−0.604	−0.598	3.132	3.119	3.125
14	0.025	0.027	0.026	1.068	1.147	1.069	−0.616	−0.615	−0.602	3.229	3.226	3.212
15	−0.001	−0.001	−0.001	0.862	0.866	0.862	−1.686	−1.681	−1.684	6.485	6.461	6.479
16	−0.008	−0.009	−0.002	1.079	1.197	1.080	−0.278	−0.276	−0.060	2.371	2.368	2.407
17	−0.006	−0.006	−0.006	0.883	0.898	0.883	−1.627	−1.609	−1.622	6.354	6.286	6.346
18	−0.016	−0.020	−0.035	0.948	1.166	0.940	−0.111	−0.109	0.486	2.340	2.342	2.854
19	−0.018	−0.019	−0.017	1.106	1.131	1.106	−1.156	−1.151	−1.135	4.702	4.668	4.656
20	−0.005	−0.006	−0.005	1.004	1.118	1.004	−0.493	−0.483	−0.491	3.017	2.965	3.015

Tabela 16 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 150$ e $\mu \in (0.02, 0.15)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.007	−0.009	−0.040	0.943	1.201	0.887	0.169	0.159	0.950	2.386	2.361	3.423
2	−0.020	−0.034	−0.125	0.676	1.187	0.645	0.025	0.017	0.506	2.433	2.371	2.665
3	−0.001	−0.001	0.018	1.072	1.186	1.053	0.186	0.187	0.703	2.586	2.591	3.179
4	0.013	0.013	0.018	1.166	1.202	1.164	0.760	0.760	0.838	3.422	3.422	3.541
5	−0.007	−0.007	−0.003	1.144	1.171	1.143	0.854	0.853	0.919	4.238	4.224	4.348
6	−0.026	−0.061	−0.029	0.486	1.153	0.486	−0.124	−0.106	−0.115	2.733	2.643	2.727
7	−0.002	−0.002	−0.002	1.042	1.051	1.042	1.570	1.565	1.578	7.069	7.030	7.091
8	−0.007	−0.007	−0.013	1.002	1.198	0.952	0.140	0.140	1.020	2.377	2.374	3.630
9	−0.011	−0.011	−0.011	0.758	0.759	0.758	2.776	2.774	2.777	13.499	13.484	13.501
10	0.027	0.028	0.027	1.121	1.166	1.121	0.821	0.820	0.834	3.648	3.640	3.667
11	−0.020	−0.035	−0.203	0.679	1.192	0.588	−0.007	−0.006	0.930	2.367	2.349	3.294
12	0.008	0.008	0.008	1.025	1.038	1.025	1.706	1.692	1.708	6.949	6.873	6.954
13	−0.021	−0.021	−0.019	1.121	1.151	1.120	0.826	0.824	0.877	4.456	4.441	4.539
14	0.002	0.003	0.006	1.165	1.191	1.165	0.812	0.812	0.895	4.364	4.364	4.508
15	0.000	0.000	0.000	0.716	0.720	0.716	1.899	1.894	1.900	8.809	8.776	8.812
16	0.030	0.034	0.056	1.100	1.223	1.080	0.271	0.268	0.849	2.408	2.401	3.239
17	0.006	0.006	0.006	0.838	0.844	0.838	2.689	2.678	2.690	13.312	13.216	13.314
18	0.038	0.047	0.022	0.968	1.197	0.940	0.099	0.096	0.719	2.309	2.304	2.972
19	0.012	0.012	0.013	1.102	1.116	1.102	1.383	1.381	1.389	5.542	5.528	5.559
20	−0.014	−0.014	−0.014	1.099	1.122	1.099	1.386	1.382	1.400	5.847	5.811	5.890

Tabela 17 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 150$ e $\mu \in (0.23, 0.85)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.003	−0.004	−0.012	0.988	1.192	0.941	−0.075	−0.075	1.097	2.385	2.379	3.993
2	−0.018	−0.027	−0.318	0.633	1.112	0.579	0.196	0.263	1.941	3.330	2.951	8.252
3	0.006	0.007	0.021	1.053	1.170	1.042	0.021	0.020	0.584	2.591	2.591	3.125
4	−0.005	−0.005	0.005	1.149	1.186	1.148	0.581	0.581	0.744	3.333	3.336	3.563
5	−0.002	−0.002	0.002	1.158	1.174	1.157	0.896	0.896	0.962	4.382	4.377	4.499
6	0.007	0.014	−0.012	0.628	1.114	0.632	−0.070	−0.055	0.042	2.642	2.601	2.688
7	0.016	0.016	0.017	1.062	1.074	1.061	0.002	0.001	0.017	4.263	4.247	4.262
8	−0.006	−0.006	0.015	1.052	1.188	1.017	0.135	0.135	1.111	2.431	2.429	4.034
9	−0.013	−0.013	−0.013	0.724	0.728	0.724	0.401	0.402	0.408	4.435	4.440	4.432
10	0.000	0.000	0.003	1.104	1.161	1.106	−0.659	−0.657	−0.581	3.693	3.682	3.615
11	0.005	0.008	−0.173	0.684	1.167	0.603	−0.048	−0.053	0.844	2.363	2.345	3.111
12	0.016	0.018	0.016	0.964	1.048	0.964	−0.808	−0.802	−0.799	5.313	5.070	5.304
13	−0.004	−0.005	−0.002	1.114	1.162	1.114	0.179	0.175	0.244	3.824	3.812	3.855
14	0.008	0.008	0.011	1.122	1.173	1.122	0.667	0.664	0.751	3.994	3.983	4.114
15	−0.001	−0.001	−0.001	0.723	0.729	0.723	0.472	0.471	0.474	4.531	4.528	4.533
16	0.006	0.007	0.039	1.068	1.196	1.047	−0.040	−0.040	0.919	2.412	2.409	3.751
17	−0.008	−0.008	−0.009	0.789	0.810	0.789	−1.226	−1.211	−1.224	6.198	6.122	6.195
18	0.009	0.010	0.007	0.994	1.172	0.912	0.126	0.124	1.384	2.374	2.371	4.737
19	−0.026	−0.026	−0.025	1.101	1.131	1.100	0.047	0.050	0.081	4.297	4.263	4.306
20	0.005	0.005	0.005	1.040	1.107	1.040	−0.898	−0.893	−0.868	4.523	4.459	4.476

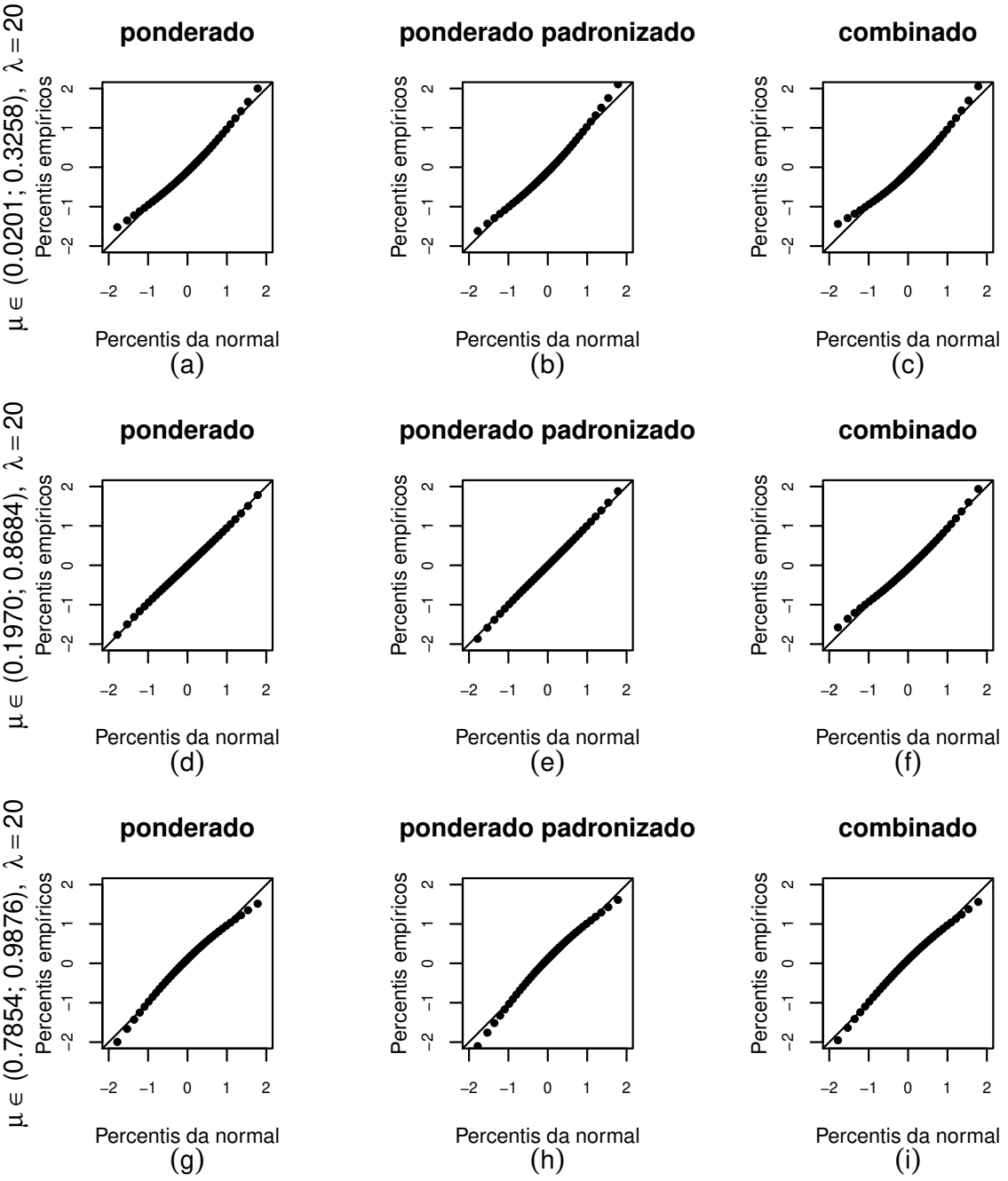

Figura 7 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 40$, $\lambda = 20$.

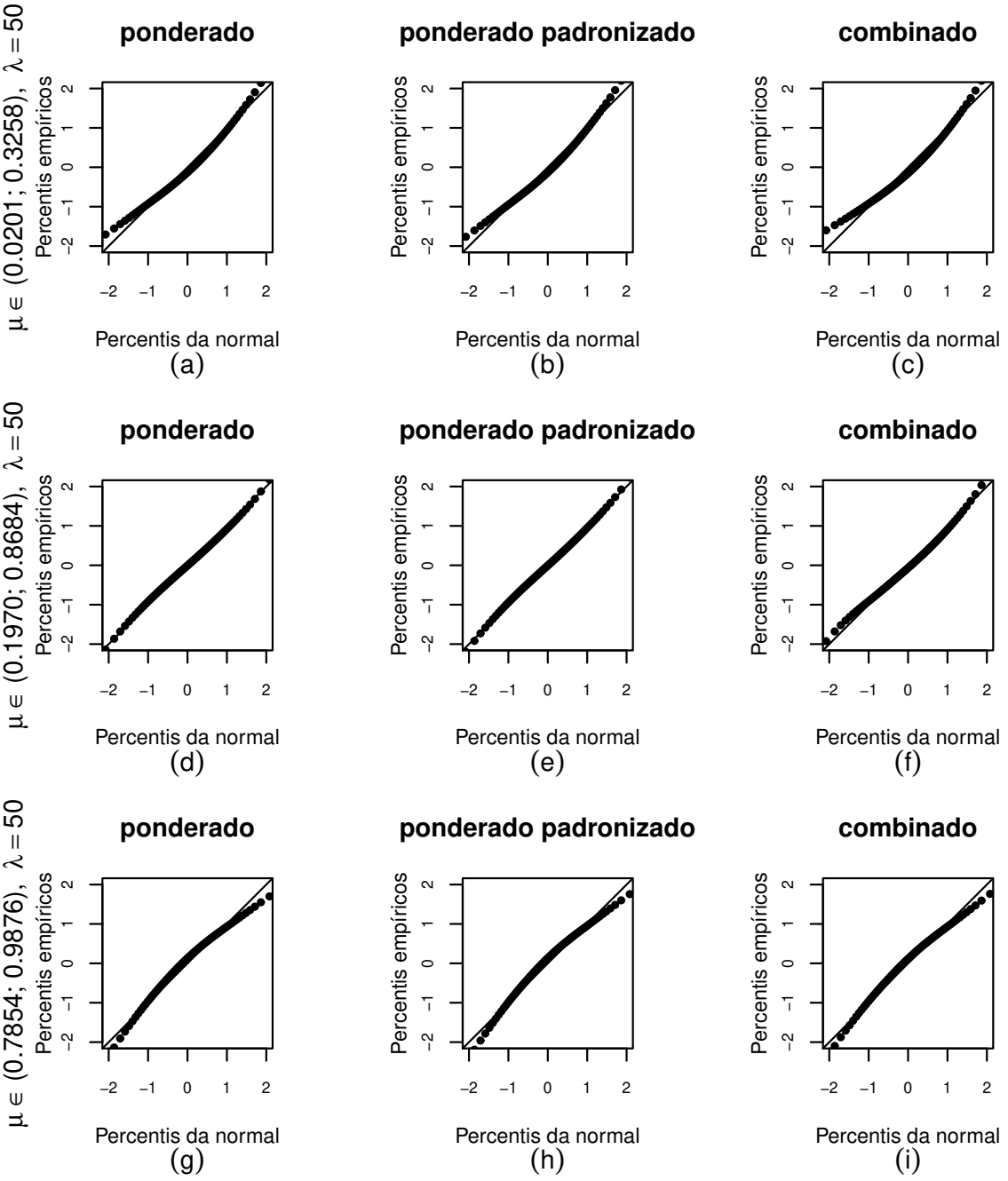

Figura 8 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 80$, $\lambda = 50$.

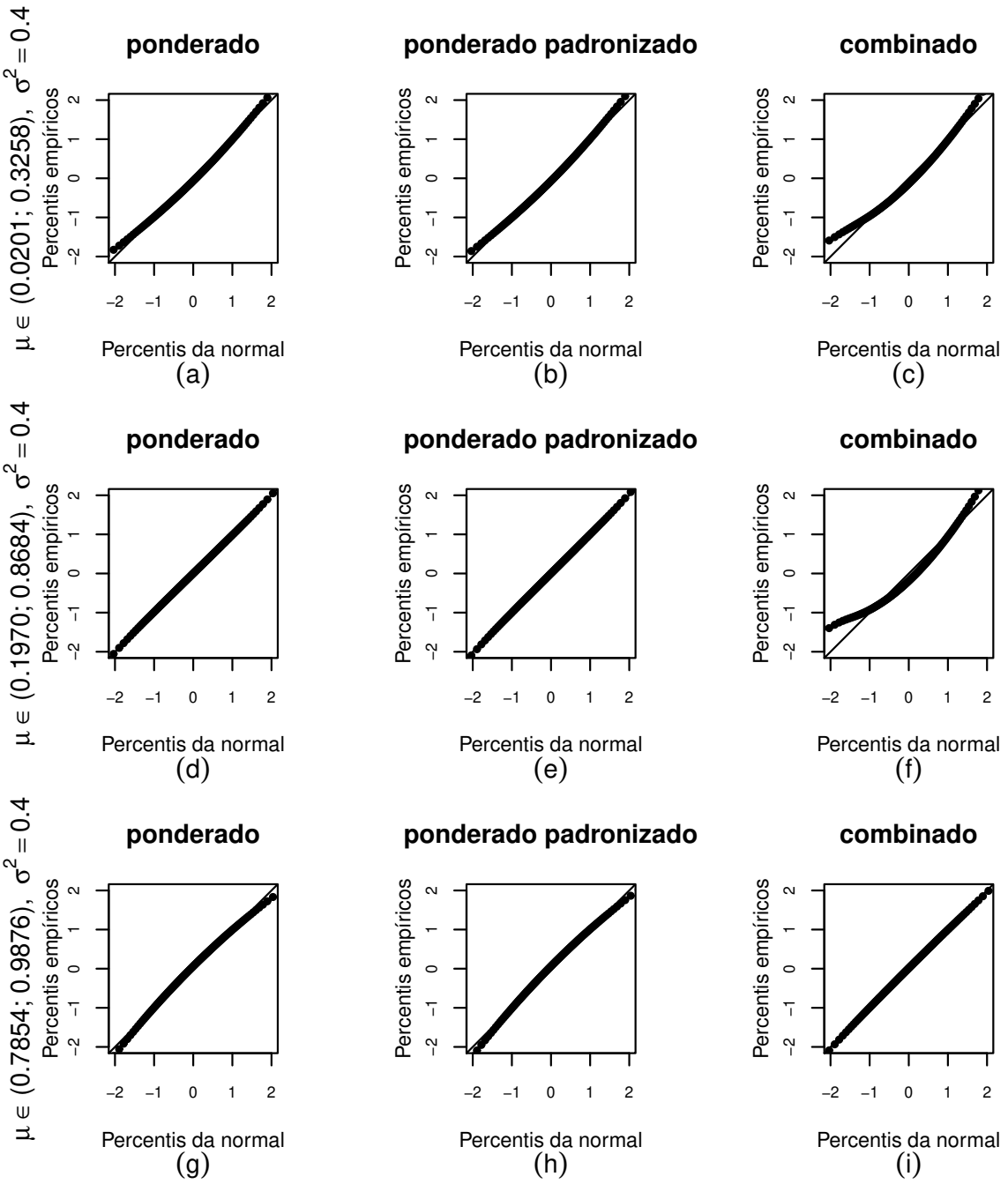

Figura 9 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$, $t = 1, \ldots, 120$, $\sigma^2 = 0.4$.

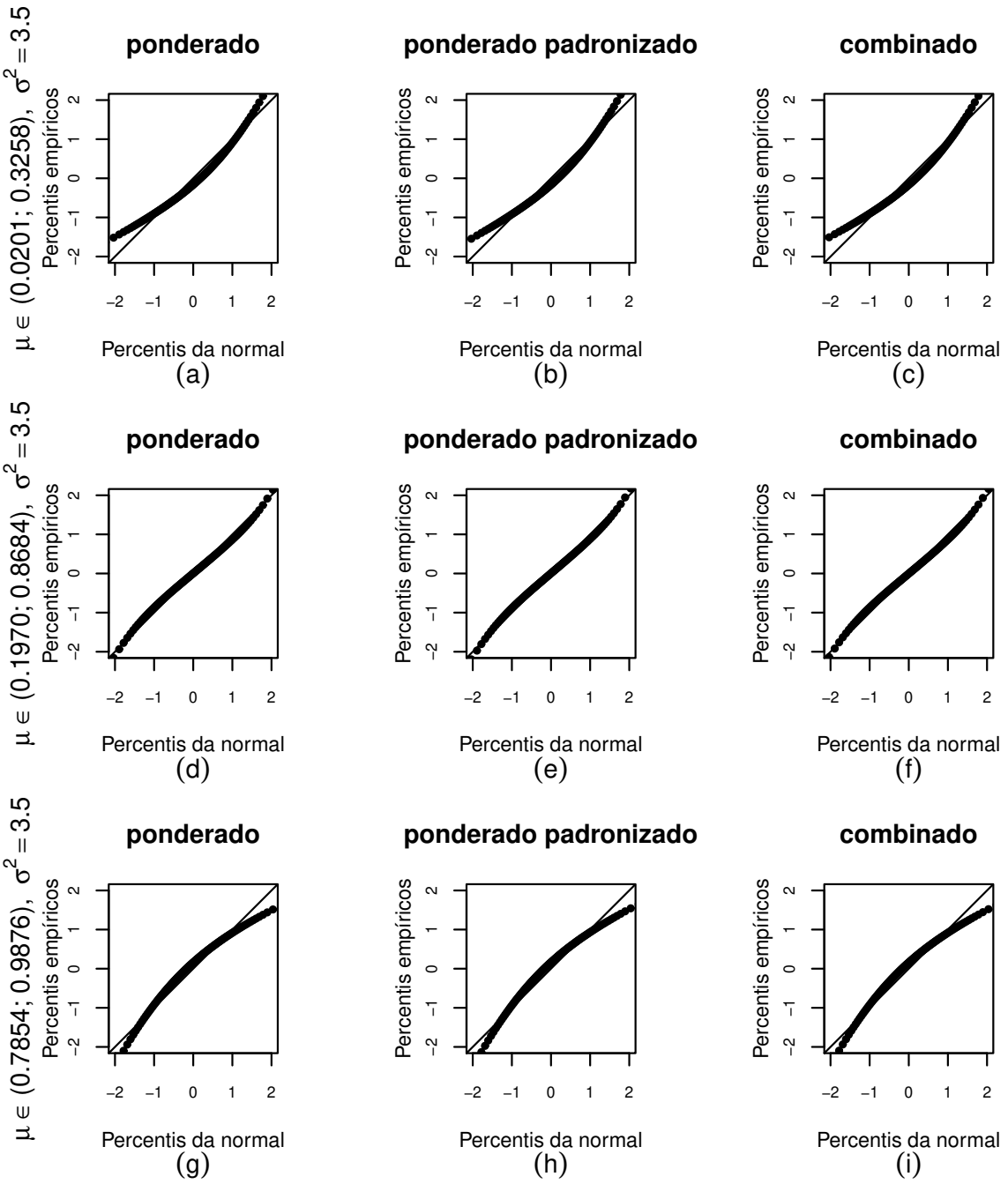

Figura 10 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$, $t = 1, \ldots, 120$, $\sigma^2 = 3.5$.

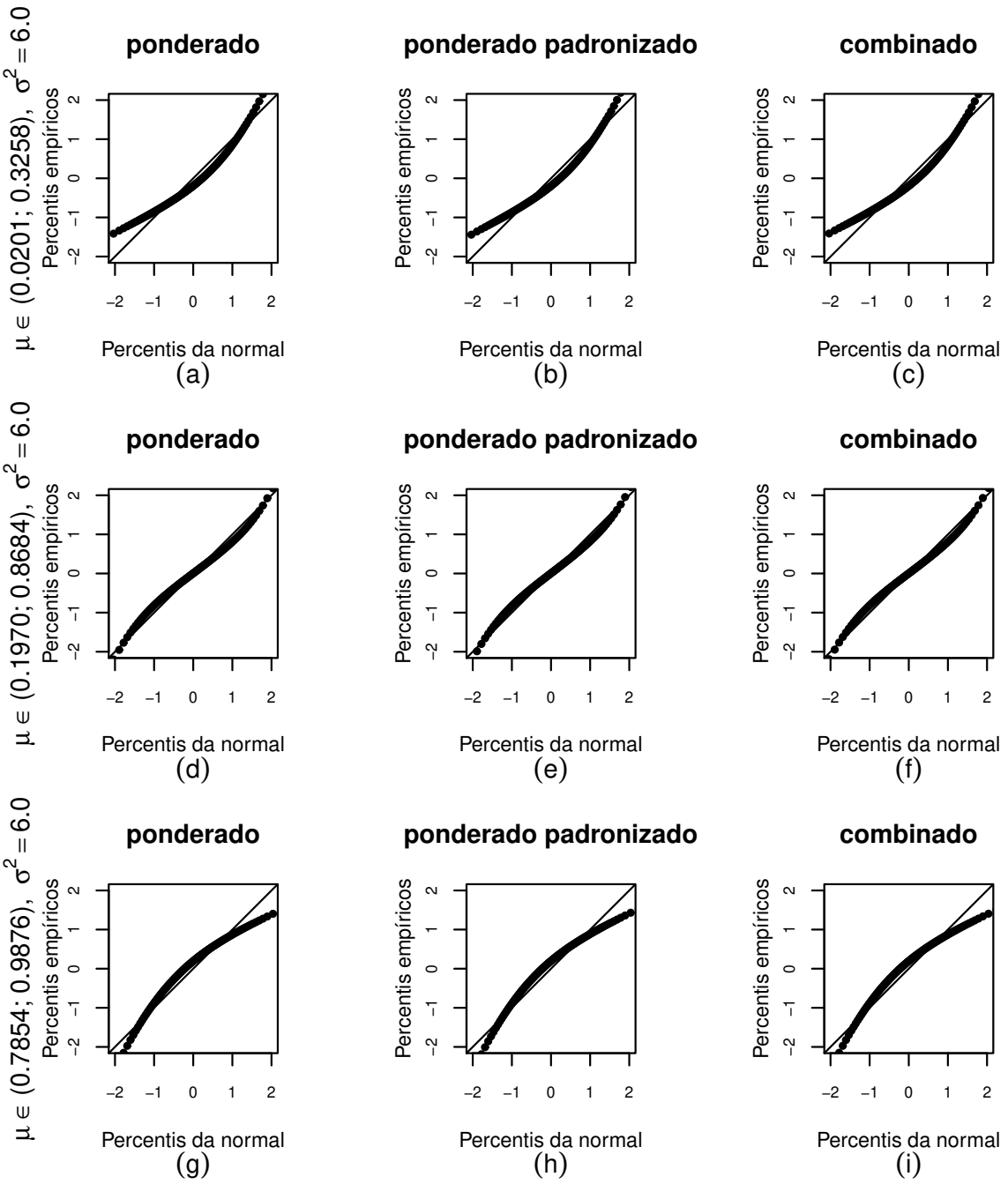

Figura 11 – Gráficos normais de probabilidade dos resíduos ponderado, ponderado padronizado e combinado. Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$, $t = 1, \ldots, 120$, $\sigma^2 = 6.0$.

Tabela 18 – Médias, erros-padrão, assimetrias e curtoses dos resíduos: ponderado (r^{β}), ponderado padronizado (r_p^{β}) e combinado ($r^{\beta\gamma}$). Modelo: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$, $t = 1, \ldots, 20$, $\lambda = 150$ e $\mu \in (0.80, 0.98)$.

	Média			Erro-padrão			Assimetria			Curtose		
t	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$	r^{β}	r_p^{β}	$r^{\beta\gamma}$
1	−0.016	−0.020	−0.028	0.922	1.190	0.922	−0.114	−0.107	0.078	2.335	2.321	2.377
2	−0.006	−0.006	−0.054	0.528	1.112	0.533	0.149	0.268	0.332	3.428	2.835	3.492
3	0.015	0.016	0.017	1.057	1.178	1.058	−0.369	−0.369	−0.285	2.607	2.607	2.571
4	−0.006	−0.006	−0.002	1.162	1.192	1.163	−0.811	−0.812	−0.744	3.522	3.525	3.428
5	0.020	0.020	0.021	1.155	1.169	1.155	−1.175	−1.175	−1.151	4.879	4.878	4.819
6	0.025	0.039	0.019	0.736	1.135	0.738	−0.231	−0.219	−0.174	2.720	2.686	2.714
7	0.012	0.012	0.012	1.053	1.061	1.053	−1.501	−1.498	−1.500	5.948	5.929	5.944
8	−0.007	−0.007	−0.002	1.062	1.186	1.065	−0.264	−0.262	0.001	2.422	2.418	2.453
9	−0.004	−0.004	−0.004	0.811	0.813	0.811	−2.644	−2.641	−2.644	12.375	12.349	12.374
10	0.016	0.017	0.016	1.074	1.143	1.074	−0.782	−0.780	−0.768	3.444	3.437	3.424
11	0.012	0.016	−0.075	0.814	1.167	0.800	−0.070	−0.069	0.680	2.358	2.347	3.028
12	−0.005	−0.003	−0.005	0.979	1.051	0.979	−0.953	−0.909	−0.953	3.862	3.744	3.862
13	0.014	0.015	0.014	1.063	1.131	1.063	−0.753	−0.752	−0.749	3.425	3.418	3.420
14	0.029	0.031	0.029	1.074	1.141	1.074	−0.750	−0.750	−0.742	3.508	3.506	3.496
15	0.000	0.000	0.000	0.803	0.805	0.803	−2.672	−2.668	−2.672	12.520	12.494	12.518
16	−0.009	−0.010	−0.002	1.077	1.206	1.079	−0.268	−0.266	−0.041	2.359	2.356	2.401
17	−0.001	−0.001	−0.001	0.835	0.843	0.835	−2.445	−2.434	−2.445	10.851	10.780	10.848
18	−0.015	−0.019	−0.037	0.939	1.176	0.932	−0.085	−0.084	0.542	2.344	2.343	2.901
19	−0.022	−0.023	−0.022	1.112	1.131	1.111	−1.450	−1.446	−1.440	5.881	5.847	5.853
20	−0.005	−0.006	−0.005	1.036	1.120	1.036	−0.664	−0.651	−0.662	3.326	3.278	3.324

3.4 Aplicações

Nesta seção apresentamos três aplicações. A primeira utilizamos dados simulados e as demais são dados reais relacionados ao processo de oxidação de amônia e ao processo de craqueamento catalítico fluido (FCC).

3.4.1 Dados simulados

Nosso objetivo aqui é avaliar o comportamento dos resíduos ponderado padronizado e combinado com a presença de pontos aberrantes. Para isso consideramos o seguinte modelo com dispersão não constante

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}, \quad \text{e} \quad \log(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3},$$

$t = 1, \ldots, 40$. Os valores das covariadas foram obtidos de forma independente através das seguintes distribuições $x_{1t} \sim U(0,1)$, $x_{2t} \sim t_3$, $x_{3t} \sim U(0,1)$, $x_{4t} \sim t_3$, $z_{1t} \sim U(0.3, 1.3)$ e $z_{2t} \sim U(-0.5, 0.5)$. Os valores verdadeiros dos parâmetros são $\beta = (1.0, 1.4, -1.3, 1.2, 1.2)$ e $\gamma = (-1.3, 1.3, 2.1)$, de modo que $\mu \in (0.059, 0.999)$ e $\lambda = 15.93$. Observe que os valores

de duas covariáveis foram obtidos a partir da distribuição t_3, que tem cauda pesada. Portanto, é provável que alguns pontos dos dados sejam atípicos.

Na Figura 12 apresentamos os gráficos dos resíduos ponderado padronizado e combinado versus os índices das observações (Figura 12 (a) e (b)) e os gráficos normais de probabilidade com envelopes simulados (Figura 12 (c) e (d)). Na Figura 12 (a) e (b) é possível ver que o comportamento dos dois resíduos são bem similares, ambos destacam as observações 21 e 37 como pontos aberrantes. Além disso, o resíduo ponderado padronizado também destaca a observação 30 como aberrante. Nos gráficos normais de probabilidade com envelopes simulados, o resíduo combinado destaca os pontos {1,2,10,32} como possivelmente influentes enquanto que o resíduo ponderado padronizado destaca os pontos {10,14,16}. Para investigar melhor o comportamento dessas observações nós retiramos esses casos e reestimamos o modelo para verificar o impacto delas nas estimativas dos parâmetros. A Tabela 19 apresenta as mudanças relativas nas estimativas dos parâmetros (%), nas estimativas dos erros-padrão (%) e os p-valores dos testes em que os parâmetros relevantes foram iguais a zero. Também contém as estimativas de parâmetros, erros padrão e p-valores obtidos usando os dados completos.

Para o modelo simplex considerado, todos os parâmetros são significativos ao nível de 5% com exceção de β_5 (p-valor = 0.101). Esse fato ocorre pois os testes de hipóteses e processo de estimação dos parâmetros são afetados com a presença de pontos aberrantes e influentes nos dados. Com a retirada das observações 21 e 30, individualmente, não observou-se nenhuma mudança expressiva nos p-valores e nas estimativas dos parâmetros. No entanto, ao retirar a observação 37, o p-valor referente a covariada β_5 passa de 0.101 para 0.213. Ao retirar apenas a observação 1 vimos uma mudança na direção oposta do p-valor relacionado a covariada x_5, que diminui para 0.083 e passa a ser significativa a 10%. Observe que apenas o resíduo combinado identificou o caso 1 como atípico (ver Figura 12). A hipótese nula sob avaliação também é rejeitada ao nível de significância de 10% quando {1,2}, {10,14} e {10,16} não estão nos dados. Vale ressaltar que na ausência dos casos {1,2,10} a hipótese nula é rejeitada ao nível nominal de 5% e conclui-se que a covariada x_5 é relevante para explicar o comportamento da variável dependente. Vale salientar também que tais pontos foram destacados pelo resíduo combinado que se mostrou muito eficaz em revelar a presença de pontos influentes a partir do gráfico normal

de probabilidades com envelopes simulados.

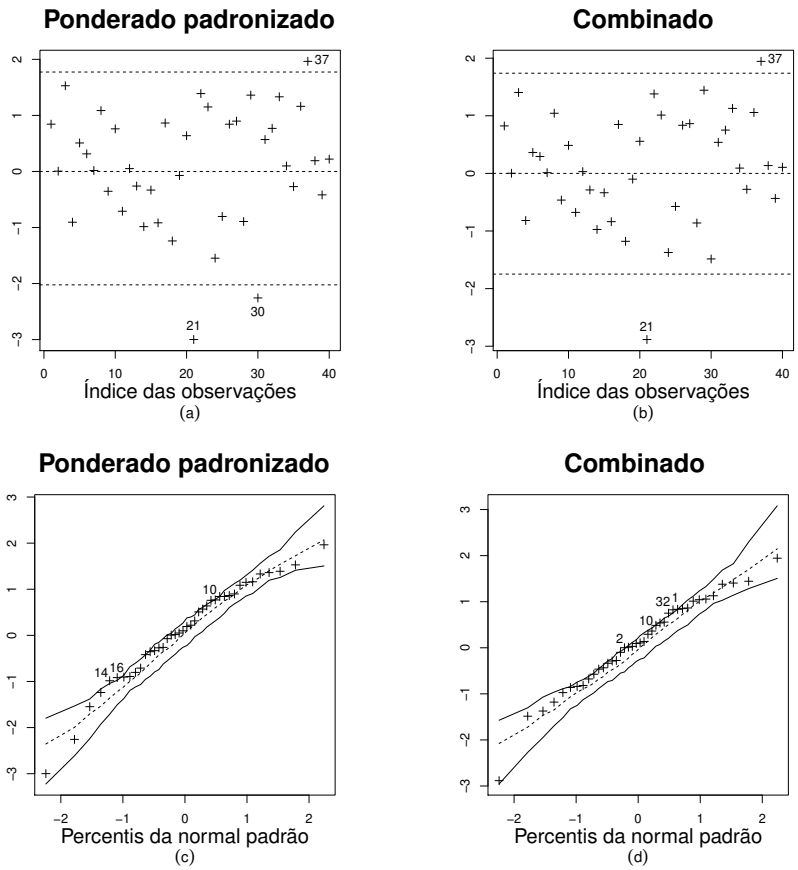

Figura 12 – Gráficos dos resíduos. Dados simulados.

3.4.2 Dados de amônia

Nesta aplicação consideramos os dados apresentados por BROWNLEE (1965, p. 454) que foram obtidos em 21 dias de processos de oxidação de amônia como um estágio para a produção do ácido nítrico em uma planta industrial. A variável resposta é a proporção de amônia que não foi convertida em ácido nítrico (y), isto é, uma medida inversa da eficiência total da planta industrial. As covariáveis são corrente de ar (x_2); temperatura da água utilizada para o resfriamento da reação (x_3) e concentração de ácido (x_4), medida com 10×(concentração de ácido - 50). Avaliamos diferentes formas para o preditor linear e consideramos o seguinte modelo de regressão simplex linear

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 (x_{t2} \times x_{t3}), \quad \text{e} \\ \log(\sigma_t^2) &= \gamma_1 x_{t3} + \gamma_2 (x_{t2} \times x_{t3}), \end{aligned}$$

Tabela 19 – Estimativas dos parâmetros, erros-padrão (e.p.), mudanças relativas das estimativas e dos erros-padrão para casos excluídos e respectivos p-valores. Dados simulados.

	Parâmetro	β_1	β_2	β_3	β_4	β_5	γ_1	γ_2	γ_3
Dados	est.	0.875	1.458	−1.303	1.331	0.044	−2.581	2.586	−1.977
completos	e.p.	0.044	0.069	0.013	0.065	0.027	0.724	0.917	0.798
	p-v	0.000	0.000	0.000	0.000	0.101	0.000	0.005	0.013
obs. 21	mud. est.	4.089	−12.767	1.110	9.186	−21.634	−0.074	−15.780	−124.582
Deletado	mud. e.p.	24.731	−0.947	−12.642	2.115	−11.113	1.484	0.763	2.410
	p-v	0.000	0.000	0.000	0.000	0.148	0.000	0.019	0.552
obs. 30	mud. est.	7.954	−2.309	0.023	−4.036	−14.274	18.312	18.283	32.322
Deletado	mud. e.p.	0.098	−6.776	−7.917	−10.850	−10.877	6.647	5.533	3.635
	p-v	0.000	0.000	0.000	0.000	0.114	0.000	0.002	0.002
obs. 37	mud. est.	2.669	−2.166	0.105	−0.411	−20.481	−38216	−62.049	−3.493
Deletado	mud. e.p.	16.170	−3.124	−7.927	−3.882	4.739	3.168	4.996	0.058
	p-v	0.000	0.000	0.000	0.000	0.213	0.033	0.308	0.017
obs. 1	mud. est.	−0.161	0.747	0.028	−0.866	5.381	1.600	2.409	3.033
Deletado	mud. e.p.	−1.532	0.038	0.115	0.790	−0.164	1.484	0.763	2.410
	p-v	0.000	0.000	0.000	0.000	0.083	0.000	0.004	0.013
obs. {1,2}	mud. est.	−0.354	1.278	−0.008	−1.196	8.399	4.761	8.038	7.572
Deletado	mud. e.p.	−3.539	0.656	1.106	1.234	−0.468	2.726	3.185	2.799
	p-v	0.000	0.000	0.000	0.000	0.074	0.000	0.003	0.010
obs. {10,14}	mud. est.	0.328	1.460	−0.146	−1.454	8.476	3.655	6.379	6.116
Deletado	mud. e.p.	−2.339	1.155	1.404	2.244	−0.127	0.672	1.250	3.674
	p-v	0.000	0.000	0.000	0.000	0.075	0.000	0.003	0.011
obs. {10,16}	mud. est.	−3.421	1.421	−0.270	2.420	9.539	3.783	5.639	5.204
Deletado	mud. e.p.	25.623	10.197	2.054	8.131	−0.271	12.176	9.864	3.991
	p-v	0.000	0.000	0.000	0.000	0.071	0.001	0.007	0.012
obs. {1,2,10}	mud. est.	−5.072	5.337	−0.332	−0.987	25.490	18.238	25.999	25.673
Deletado	mud. e.p.	10.824	4.146	0.739	−1.281	−4.861	9.646	9.043	6.382
	p-v	0.000	0.000	0.000	0.000	0.030	0.000	0.001	0.003

$t = 1, \ldots, 21$. O grau de dispersão não constante estimado para o modelo simplex é $\widehat{\lambda} = 16.214$, com $\widehat{\sigma^2}_{min} = 0.536$ e $\widehat{\sigma^2}_{max} = 8.686$. Para avaliar o ajuste do modelo construímos os gráficos dos resíduos ponderado padronizado e combinado versus o índice das observações, a covariada corrente de ar e os valores preditos para o modelo acima (Figura 13). Nesses gráficos vimos que os resíduos estão distribuídos aleatoriamente em torno do zero, destacando os pontos $\{4, 20, 21\}$ como aberrantes. Na Figura 14 apresentamos os gráficos normais de probabilidade com envelopes simulados. Os pontos estão dentro das bandas de confiança com exceção da observação 3 que encontra-se em cima da linha. Além disso, quando o gráfico normal de probabilidade com envelope simulados é construído utilizando o resíduo combinado, a observação 10 está muito próxima da banda de confiança do envelope. Estas duas observações podem ser influentes e merecem uma análise mais aprofundada.

Para verificar se esses pontos são influentes retiramos esses casos individualmente e reestimamos o modelo. A Tabela 20 apresenta as mudanças relativas nas estimativas dos parâmetros (%), nas estimativas dos erros-padrão (%) e os p-valores dos testes z com a exclusão das observações para esse conjunto de dados. As observações 3, 20 e 21 mudam

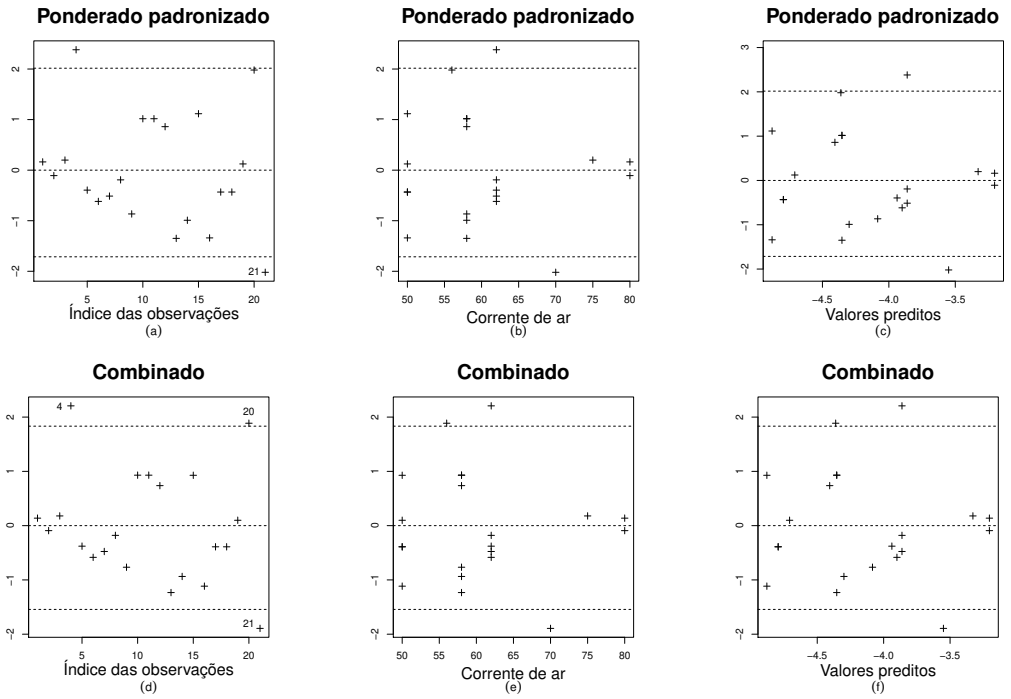

Figura 13 – Gráficos de resíduos. Modelo simplex: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4(x_{t2} \times x_{t3})$ e $\log(\sigma_t^2) = \gamma_1 x_{t3} + \gamma_2(x_{t2} \times x_{t3})$, $t = 1, \ldots, 21$. Dados de amônia.

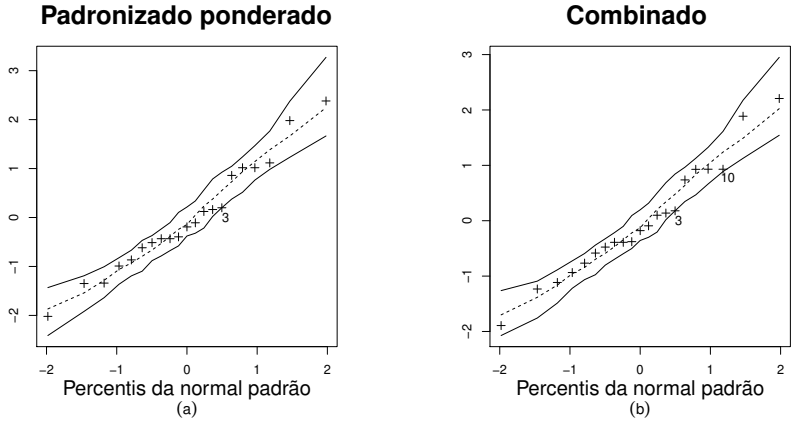

Figura 14 – Gráficos normais de probabilidade com envelopes simulados. Modelo simplex: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4(x_{t2} \times x_{t3})$ e $\log(\sigma_t^2) = \gamma_1 x_{t3} + \gamma_2(x_{t2} \times x_{t3})$, $t = 1, \ldots, 21$. Dados de amônia.

de forma expressiva as estimativas dos parâmetros e os p-valores, como por exemplo, ao retirar a observação 3 o p-valor referente a covariada temperatura passa de 0.019 para 0.035. Além disso, com a retirada da observação 3 a interação das covariadas corrente de ar e temperatura deixa de ser significativa a 5%. A remoção individual das observações 20 e 21 também tem um grande impacto inferencial, especialmente a retirada da observação 21. Com a exclusão conjunta das observações {3,10}, os p-valores associados a x_3 e $x_2 \times x_3$ passa de 0.019 e 0.05 para 0.041 e 0.094, respectivamente. Aqui, os resíduos ponderado e combinado se comportam de maneira semelhante. No entanto, o resíduo combinado sugere que as observações 10 e 20 são atípicas. Como vimos, tais observações são consideradas influentes. Está claro, portanto, que a análise de diagnóstico baseada no resíduo combinado supera o resíduo ponderado.

Tabela 20 – Estimativas dos parâmetros, erros-padrão (e.p.), mudanças relativas das estimativas e dos erros-padrão para casos excluídos e respectivos p-valores. Dados de amônia.

Modelo : $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4(x_{t2} \times x_{t3})$ e $\log(\sigma_t^2) = \gamma_1 x_{t3} + \gamma_2(x_{t2} \times x_{t3})$							
	Parâmetro	β_1	β_2	β_3	β_4	γ_1	γ_2
Dados	est.	−13.272	0.137	0.283	−0.004	−0.217	0.004
completos	e.p.	2.401	0.041	0.121	0.002	0.094	0.002
	p-v	0.000	0.001	0.019	0.050	0.021	0.013
obs. 3	mud. est.	3.592	6.346	7.699	9.987	31.856	31.633
Deletado	mud. e.p.	17.978	19.019	19.751	21.470	4.755	6.268
	p-v	0.000	0.003	0.035	0.075	0.004	0.002
obs. 4	mud. est.	−3.986	−4.357	−8.702	−6.297	−3.019	−7.720
Deletado	mud. e.p.	−15.320	−15.523	−15.926	−16.348	0.142	0.011
	p-v	0.000	0.000	0.011	0.028	0.026	0.021
obs. 10	mud. est.	−3.417	−7.176	−5.642	−9.173	−15.822	−15.061
Deletado	mud. e.p.	−3.684	−3.526	−4.536	−4.778	0.671	0.388
	p-v	0.000	0.001	0.021	0.061	0.054	0.035
obs. 20	mud. est.	−7.021	−11.289	−16.463	−19.140	22.204	18.330
Deletado	mud. e.p.	0.845	1.288	1.305	1.986	1.487	1.008
	p-v	0.000	0.003	0.053	0.119	0.006	0.004
obs. 21	mud. est.	3.691	6.981	5.759	8.117	−157.858	−161.432
Deletado	mud. e.p.	−16.143	−41.913	−50.771	−52.207	0.809	1.304
	p-v	0.000	0.000	0.000	0.000	0.187	0.130
obs. {3,10}	mud. est.	0.362	−0.344	2.127	1.170	18.800	19.189
Deletado	mud. e.p.	15.966	17.151	17.085	18.640	5.325	6.572
	p-v	0.000	0.004	0.041	0.094	0.009	0.005
obs. {3,20}	mud. est.	−5, 268	−5, 268	−12, 872	−14, 352	51.324	47.281
Deletado	mud. e.p.	16.316	17.740	18.124	20.330	6.135	7.178
	p-v	0.000	0.009	0.084	0.162	0.001	0.001
obs. {4,20}	mud. est.	−9.099	−12.484	−20.385	−19.547	25.929	14.015
Deletado	mud. e.p.	−20.613	−20.306	−20.759	−20.432	1.683	1.033
	p-v	0.000	0.000	0.019	0.047	0.004	0.005
obs. {3,4,20}	mud. est.	−3.244	−2.605	−6.543	−2.378	50.188	38.277
Deletado	mud. e.p.	−8.911	−7.857	−7.944	−6.430	6.219	7.178
	p-v	0.000	0.000	0.017	0.040	0.001	0.001

Ajustamos ainda a esse conjunto de dados o modelo de regressão beta. Inicialmente testamos o mesmo modelo proposto para o modelo simplex. No entanto, o p-valor para o teste em que a hipótese nula $\beta_4 = 0$ é igual a 0.140. Adicionalmente, quando retiramos os casos possivelmente influentes como as observações 10 e 20 individualmente o p-valor aumenta para 0.153 e 0.234, respectivamente. Assim, vimos que o impacto inferencial das duas observações influentes é mais nítido no modelo beta do que no modelo simplex.

Consideramos um grande número de preditores lineares com base na distribuição beta. No entanto apenas o modelo apresentou regressores estatisticamente significantes ao nível de significância de 5%

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}, \quad \text{e} \quad \log(\phi_t) = \gamma_1 + \gamma_2 x_{t2},$$

$t = 1, \ldots, 21$.

Para o modelo beta acima temos que $\widehat{\lambda} = 670.47$, em que $\widehat{\phi}_{min} = 41.088$ e $\widehat{\phi}_{max} = 27548$. As Figuras 15 e 16 mostram os gráficos dos resíduos ponderado padronizado e combinado versus elementos do modelo e os gráficos normais de probabilidade com envelopes simulados, respectivamente.

As observações apontadas como atípicas são as mesmas que no modelo simplex, a saber: observações $\{4, 20, 21\}$. No entanto, os gráficos normais de probabilidade mostram que o ajuste do modelo é ruim, pois muitos pontos estão fora dos envelopes simulados. Vale ressaltar que todos os valores da variável resposta estão próximos ao limite inferior do intervalo unitário padrão $\min(y_t) = 0.007$ e $\max(y_t) = 0.042$. Adicionalmente, $\widehat{\sigma^2}_{min} = 0.536$ e $\widehat{\sigma^2}_{max} = 8.686$ que são valores baixos de dispersão para o modelo simplex. Neste cenário, tipicamente o processo de estimação do modelo simplex tem se mostrado menos sensível que o processo de estimação do modelo beta, como mencionado por ESPINHEIRA & SILVA (2018). Também notamos que $\widehat{\lambda}_{simplex} = 16.214$ e $\widehat{\lambda}_{beta} = 670.47$, ou seja, o modelo simplex apresenta substancialmente menos heterogeneidade do que o modelo beta.

Quando todos os valores da variável resposta estão próximos de um dos limites do intervalo unitário padrão, e há pontos influentes, o modelo simplex normalmente fornece estimativas de parâmetros que são mais estáveis (ou seja, menos influenciadas por observações atípicas) do que o modelo beta.

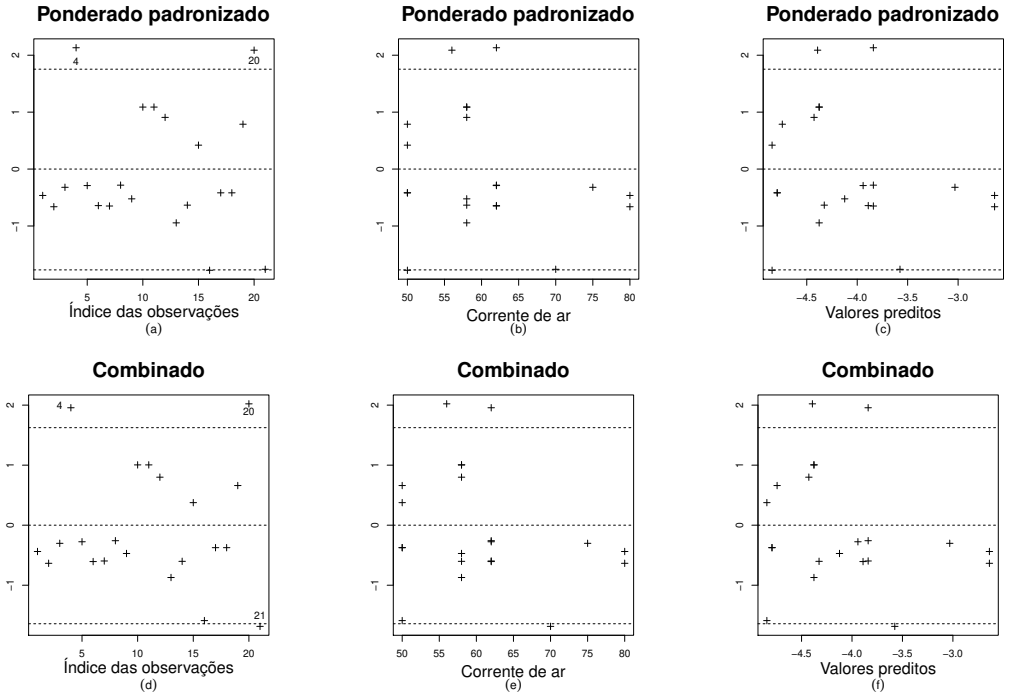

Figura 15 – Gráficos de resíduos. Modelo beta: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 x_{t2}$, $t = 1, \ldots, 21$. Dados de amônia.

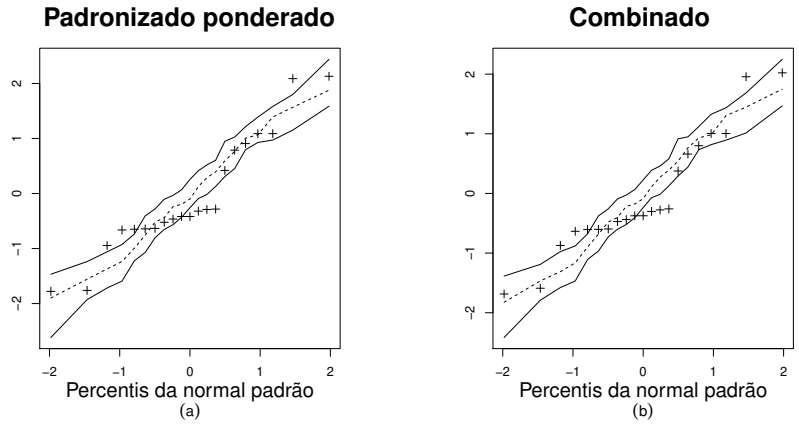

Figura 16 – Gráficos normais de probabilidade com envelopes simulados. Modelo beta: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}$ e $\log(\phi_t) = \gamma_1 + \gamma_2 x_{t2}$, $t = 1, \ldots, 21$. Dados de amônia.

3.4.3 Dados de craqueamento catalítico fluido (FCC)

O processo FCC (Fluid Catalytic Cracking) ou ruptura catalítica é usado para converter hidrocarbonetos de alto peso molecular em pequenas moléculas de maior valor comercial, através do contato destes com um catalisador. Ele é usado para converter frações pesadas de petróleo em gasolina, olefinas C3 e C4, GLP e frações que permitem a formulação de combustível diesel. O processo da FCC é muitas vezes considerado o coração de uma refinaria, uma vez que permite adaptar a produção aos produtos com maior demanda e/ou alta rentabilidade (SALAZAR, 2005). O catalisador do processo é formado por partículas finas de 10 a 150 microns, facilmente fluidizável tendo como componente principal o zeólito Y incorporado numa matriz amorfa de aluminossilicato e argila (SALAZAR, 2005).

Sabe-se que cada 1000 ppm de vanádio no catalisador a produção de gasolina diminui em cerca de 2.3%. Além disso, este componente químico é conhecido por participar da destruição do catalisador, reduzindo a superfície ativa, a seletividade e a cristalinidade do zeólito Y especialmente na presença de vapor. O vanádio é depositado na superfície externa das partículas do catalisador no reator da unidade FCC, esses complexos sofrem decomposição parcial e são transferidos para o regenerador onde eles são queimados com a coca e o vanádio é oxidado. Esta reação depende da temperatura do regenerador que deve ser próximo a 720 graus Celsius (SALAZAR, 2005).

O objetivo aqui é modelar a variável resposta porcentagem de cristalinidade do zeólito Y (y) através das covariadas vapor d'água (x_2), temperatura do processo (x_3) e concentração de vanádio (x_4). Inicialmente, consideramos um modelo linear de dispersão variável. Por uma questão de brevidade, não apresentaremos a análise de diagnóstico para esse modelo. Devemos destacar apenas que os gráficos dos resíduos sugeriram que existe tendência não-linear.

Em seguida, procuramos por um modelo não linear que se ajustasse bem aos dados. Para esse fim, usamos como linha de base o modelo utilizado por ESPINHEIRA & SILVA (2018) que utilizaram o seguinte modelo não linear:

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + \beta_2 x_{t2}/(x_{t2}+\beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}} \quad \text{e} \\ \log(\sigma_t^2) &= \gamma_1 + \gamma_2 x_{t4}^2, \end{aligned}$$

$t = 1, \ldots, 28$.

Para a construção do chute inicial temos que a t-ésima linha da matriz $J_1^{(0)} = [\partial\eta/\partial\beta]_{\beta=\beta^{(0)}}$ é definida como $(1, x_{t2}/(x_{t2}+\beta_3^{(0)}), -\beta_2^{(0)}x_{t2}/(x_{t2}+\beta_3^{(0)})^{-2}, x_{t3}, \sqrt{x_{t4}})$, $t = 1, \ldots, 28$. Como estamos no caso em que temos mais parâmetros que covariadas, obtemos $(\beta_1^{(0)}, \beta_2^{(0)}, \beta_4^{(0)}, \beta_5^{(0)})^\top$ através de $(X^\top X)^{-1}X^\top g(y)$, em que a t-ésima linha de X é dada por $x_t = (1, x_{t2}/(x_{t2}+\beta_3^{(0)}), x_{t3}, \sqrt{x_{t4}})$, $t = 1, \ldots, 28$, e atribuímos um valor numérico plausível para $\beta_3^{(0)}$. Temos que os valores que a covariada vapor d'água pode assumir estão entre $(-20, 80)$. Considere adicionalmente as seguintes relações baseadas na variável vapor d'agua: $(x_{t2} + \beta_3^{(0)}) = -20$ e $(x_{t2} + \beta_3^{(0)}) = 80$. Se x_{t2} assume o mínimo, então $\beta_3^{(0)} \in (-20, 80)$, se x_{t2} assume o máximo, $\beta_3^{(0)} \in (-75.8.24.2)$. Após a escolha de $\beta_3^{(0)}$, obtemos $\beta_{NL}^{(0)} = (J_1^{(0)\top}J_1^{(0)})^{-1}J_1^{(0)\top}(g(y) - f(x, \beta_L^{(0)}))$, em que $f(x, \beta_L^{(0)}) = \beta_1^{(0)} + \beta_2^{(0)}x_{t2}/(x_{t2} + \beta_3^{(0)}) + \beta_4^{(0)}x_{t3} + \beta_5^{(0)}\sqrt{x_{t4}}$. Testamos vários valores para $\beta_3^{(0)}$ e entre esses valores $\beta_3^{(0)} = -19$ resultou na convergência do processo iterativo e os resultados abaixo.

Após convergência nós construímos os gráficos dos resíduos contra elementos do modelo (índice das observações, covariada concentração de vanádio, valores preditos) para o modelo não linear proposto (Figura 17). Observamos que o resíduo ponderado padronizado e combinado se distribuem de forma aleatória em torno do zero, em especial nos gráficos dos resíduos contra os índices das observações e valores preditos. Vimos ainda que o resíduo ponderado padronizado destaca as observações 10 e 24 como aberrantes e o resíduo combinado destaca mais enfaticamente a observação 10. Ao retirar essas observações verificamos que a inferência do modelo não foi alterada o que significa que não são observações influentes. Adicionalmente, na Figura 18 apresentamos os gráficos dos resíduos ponderado padronizado e combinado com envelopes simulados. Podemos notar que o gráfico do resíduo combinado revela a existência de quatro pontos fora do envelope que são pontos potencialmente influentes, a saber: as observações 8, 13, 14 e 27.

Dessa forma, excluímos essas observações individualmente e reestimamos o modelo e computamos as mudanças relativas (%) nas estimativas. Notamos que nenhuma delas

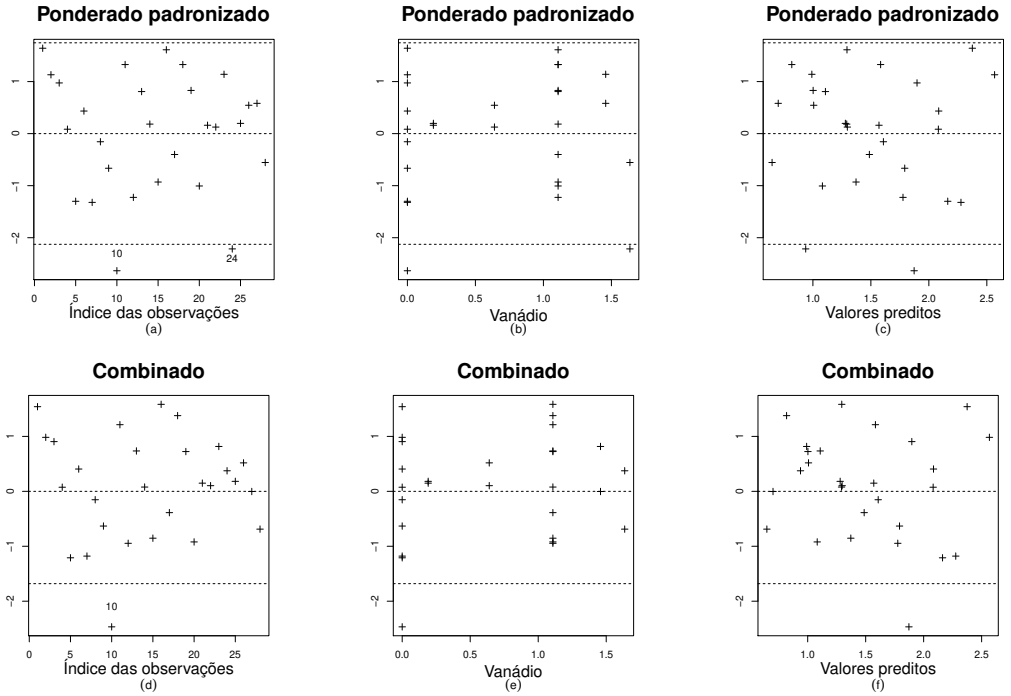

Figura 17 – Gráficos dos resíduos. Modelo simplex: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}/(x_{t2} + \beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t4}^2$, $t = 1, \ldots, 28$. Dados FCC.

mudou de forma expressiva os valores das estimativas ou dos p-valores o que significa que essas observações não são individualmente influentes. Decidimos ainda avaliar o impacto com a retirada das observações conjuntamente. Para isso removemos sequencialmente todos os subconjuntos dos quatro casos, estimamos os parâmetros do modelo, e calculamos as mudanças relativas nas estimativas e erros- padrão (%) e os correspondentes p-valores. Os resultados encontram-se na Tabela 21. Quando os pontos $\{10, 13, 14, 27\}$ são excluídos percebe-se o forte impacto no processo de inferência no modelo. Vimos, por exemplo, o p-valor associado a covariada vapor d'água que passa de 0.0116 para 0.0338 e o p-valor associado a covariada concentração de vanádio que passa de 0.0066 para 0.0214.

Com isso podemos perceber que o resíduo combinado se mostrou mais sensível para detectar a presença de pontos influentes que o resíduo ponderado. No entanto, como os pontos influentes não alteram as conclusões inferenciais temos que o modelo de regressão simplex mostra-se uma boa alternativa para a modelagem do percentual de cristalinidade do zeólito Y, fator importante na produção da gasolina. Ressaltamos, que ESPINHEIRA & SILVA (2018) realizou análise de influência local para esse mesmo modelo simplex e para um modelo competidor baseado na regressão beta não linear. Os autores con-

cluíram que o processo de estimação do modelo simplex mostrou-se menos sensível aos pontos influentes. Para esse conjunto de dados observamos que os valores da variável resposta estão concentrados no limite superior do intervalo unitário com $\min(y_t) = 0.643$ e $\max(y_t) = 0.963$. Temos ainda que $\widehat{\sigma^2}_{min} = 0.088$ e $\widehat{\sigma^2}_{max} = 2.281$. Quando o valor de σ^2 está nessa faixa de valores, a distribuição simplex normalmente oferece ajustes muito bons.

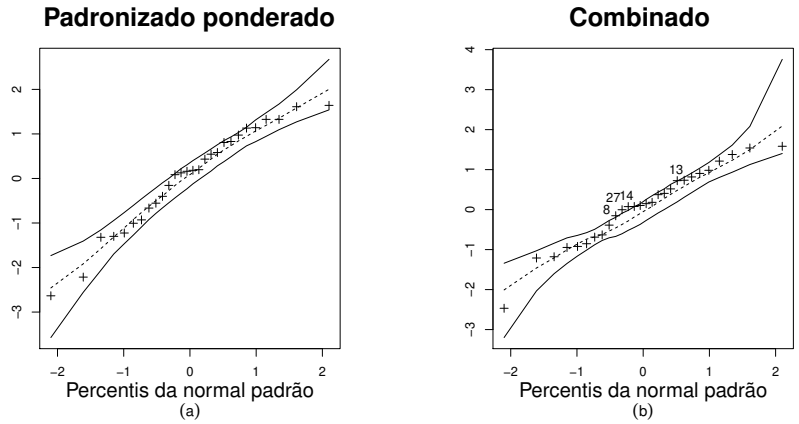

Figura 18 – Gráficos normais de probabilidade com envelopes simulado. Modelo simplex: $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}/(x_{t2}+\beta_3)+\beta_4 x_{t3}+\beta_5\sqrt{x_{t4}}$ e $\log(\sigma_t^2) = \gamma_1+\gamma_2 x_{t4}^2$, $t = 1, \ldots, 28$. Dados FCC.

3.5 Conclusão

Neste capítulo nós propomos um novo resíduo, denominado resíduo combinado, para a classe de modelos de regressão simplex não linear. Esse resíduo é baseado no processo iterativo escore de Fisher para a estimação de $\boldsymbol{\beta}$ e $\boldsymbol{\gamma}$, os vetores de parâmetros que indexam os submodelos da média e do parâmetro de dispersão.

Apresentamos resultados de simulações de Monte Carlo para avaliar o resíduo combinado considerando diversos cenários, diferentes tamanhos amostrais, valores para o parâmetro de dispersão e graus de dispersão não constante. Verificamos que a distribuição empírica do resíduo possui uma leve assimetria e por isso os limites -2 e 2 para detectar pontos aberrantes podem não ser adequados. Sugerimos a utilização dos quantis empíricos dos resíduos obtidos a partir do envelope simulado da distribuição normal padrão.

Apresentamos três aplicações. Uma com dados simulados na qual verificamos a eficiência do resíduo combinado em detectar pontos influentes nos dados, e duas aplicações a

Tabela 21 – Estimativas dos parâmetros, erros-padrão (e.p.), mudanças relativas das estimativas e dos erros-padrão para casos excluídos e respectivos p-valores. Dados FCC.

Modelo simplex : $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}/(x_{t2}+\beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t4}^2$								
	Parâmetro	β_1	β_2	β_3	β_4	β_5	γ_1	γ_2
Dados	est.	2.375	−0.107	−27.803	−0.290	−0.751	0.825	−1.217
completos	e.p.	0.149	0.042	4.021	0.107	0.142	0.370	0.314
	p-v	0.000	0.012	0.000	0.007	0.000	0.026	0.000
obs. {8,14}	mud.est.	−0.013	0.128	−0.396	−3.504	1.485	12.420	3.062
Deletado	mud. e.p.	7.260	4.757	10.137	4.552	10.240	3.914	2.139
	p-v	0.000	0.016	0.000	0.012	0.000	0.016	0.000
obs. {8,27}	mud.est.	0.882	1.261	0.270	7.209	1.749	7.137	−0.193
Deletado	mud. e.p.	5.122	3.908	2.238	9.658	8.008	3.872	6.491
	p-v	0.000	0.014	0.000	0.008	0.000	0.021	0.000
obs. {14,27}	mud.est.	0.638	2.680	−0.495	10.955	0.689	−1.218	−7.497
Deletado	mud. e.p.	5.219	4.068	7.810	14.229	2.143	0.484	5.362
	p-v	0.000	0.013	0.000	0.008	0.000	0.028	0.001
obs. {13,14,27}	mud.est.	0.163	4.123	0.295	3.494	0.205	3.991	−4.090
Deletado	mud. e.p.	6.295	7.861	5.076	14.839	3.719	0.754	6.063
	p-v	0.000	0.015	0.000	0.014	0.000	0.021	0.001
obs. {10,13,27}	mud.est.	3805	−6831	3258	−5983	10781	−17550	−6400
Deletado	mud. e.p.	−4766	6955	−4940	5886	−4365	4144	7078
	p-v	0000	0028	0000	0016	0000	0078	0001
obs. {10,14,27}	mud.est.	4.002	−7.065	2.051	0.543	10.945	−20.032	−11.680
Deletado	mud. e.p.	−2.562	8.803	6.420	10.115	−3.597	4.144	7.078
	p-v	0.000	0.031	0.000	0.013	0.000	0.087	0.001
obs. {10,13,14,27}	mud.est.	3.662	−5.959	2.780	−5.961	10.818	−15.409	−8.776
Deletado	mud. e.p.	−1.542	11.840	3.504	11.033	−2.286	4.445	7.726
	p-v	0.000	0.034	0.000	0.021	0.000	0.071	0.001

dados reais, uma linear e outra não linear. Na aplicação linear comparamos o modelo de regressão simplex com o modelo de regressão beta e vimos que o modelo simplex se destaca quando a média da variável resposta está próxima de zero. Na aplicação não linear vimos mais uma vez a habilidade do resíduo combinado em destacar pontos aberrantes e influentes.

Por fim, concluímos que quando os dados estão concentrados nos extremos do intervalo unitário padrão, ou seja, próximos de zero ou de um, o processo de estimação por máxima verossimilhança do modelo simplex é mais robusto que o do modelo de regressão beta, como foi visto nas aplicações.

4 ESTATÍSTICA DE PREDIÇÃO PARA O MODELO DE REGRESSÃO SIMPLEX NÃO LINEAR

4.1 Introdução

É comum investigarmos uma variável de interesse com base em um conjunto de variáveis que possam descrevê-la. Diversos modelos de regressão são estudados na literatura e são úteis para esse tipo de problema. Em especial, existem modelos adequados para modelar dados no intervalo (0,1). Como citado anteriormente, alguns modelos de regressão existentes para modelar taxas e proporções são: modelo de regressão simplex (BARNDORFF-NIELSEN & JøRGENSEN, 1991), modelo de regressão beta (FERRARI & CRIBARI-NETO, 2004), modelo de regressão Johnson S_b (LEMONTE & BAZAN, 2016), modelo de regressão gama unitária (MOUSA et al., 2013), modelo de regressão Kumarashuamyn (MITNIK & BAEK, 2013), entre outros.

O modelo de regressão simplex vem sido amplamente estudado recentemente por fazer parte do modelos de dispersão, mais detalhes ver JøRGENSEN (1997). Diversos autores tem estudado esse modelo como por exemplo, SONG & TAN (2000), SONG et al. (2004), MIYASHIRO (2008) e mais recentemente ESPINHEIRA & SILVA (2018) propôs a classe de modelos de regressão simplex não linear. Os autores utilizam o método de máxima verossimilhança para estimar o parâmetros e derivam as quantidades de influência local para o modelo.

Um passo importante na modelagem é a escolha do modelo mais adequado para os dados. Além da análise residual, existem diversas medidas e técnicas na literatura que são utilizadas para a seleção de modelos. Quanto a qualidade de ajuste as mais conhecidas são: o critério de informação de Akaike (AIC) (AKAIKE, 1973), critério bayesiano de Schwarz (SBC) (SCHWARZ, 1978), seleção *stepwise forward* ou eliminação *backward* (DRAPER & SMITH, 1981), funções da soma de quadrados de resíduos como R^2 e suas versões, entre outras. Quanto a qualidade de predição do modelo existe a estatística $PRESS$ (Predictive Residual Sum of Squares) proposta por ALLEN (1971) para o modelo de

regressão normal linear. A estatística $PRESS$ é calculada ajustando-se repetidamente o modelo, deixando de fora uma observação por vez. Em cada repetição o modelo é utilizado para predizer a observação que ficou de fora, sendo assim independente da qualidade do ajuste do modelo. Similarmente a abordagem do R^2, WOLD (1982) propôs a P^2, um coeficiente de predição baseado na $PRESS$. Seguindo essa abordagem, ESPINHEIRA et al. (2019) desenvolveu os cálculos da medidas de predição para o modelo de regressão beta não linear. Além disso, BRITO (2018) propôs essas mesmas medidas para o modelo de regressão beta com erros nas variáveis. Nosso objetivo aqui é propor as estatísticas $PRESS$ e P^2 para o modelo de regressão simplex não linear e avaliar o comportamento das medidas de qualidade de ajuste R^2_{FC} e R^2_{LR}.

Neste capítulo nós propomos a estatística $PRESS$ para o modelo de regressão simplex não linear assim como as respectivas medidas P^2. A organização desse capítulo é dada da seguinte forma. Na Seção 4.2, introduzimos as estatísticas $PRESS$, as medidas P^2 e definimos algumas medidas de qualidade de ajuste para o modelo simplex. A distância de Cook para o modelo simplex é apresentada na Seção 4.3. Na Seção 4.4 apresentamos resultados de simulações de Monte Carlo considerando diversos cenários. Aplicações a dados reais são apresentados na Seção 4.5. Por fim, algumas observações finais são encontradas na Seção 4.6.

4.2 Estatística PRESS

O coeficiente de determinação, R^2, é uma medida de diagnóstico muito utilizada para seleção de variáveis e seleção de modelos. No entanto, nem o R^2, nem estatísticas usuais, por exemplo, teste t, teste F, fornecem qualquer conhecimento sobre a qualidade dos valores preditos ou do impacto de pontos influentes nos valores preditos (MEDIAVILLA at al., 2008). A estatística $PRESS$ (Prediction Sum of Squares), por sua vez, é utilizada como um indicativo do poder preditivo do modelo, ou seja, é um critério que mede quão bem o uso dos valores ajustados por um modelo postulado pode prever as respostas observadas (NETER et al., 1996).

A $PRESS$ é calculada ajustando o modelo, repetidamente, deixando de fora uma observação de cada vez. Em cada repetição o modelo é usado para predizer a observação que ficou de fora. Modelos com valores da $PRESS$ pequenos são considerados bons

modelos candidatos pois terão pequenos erros de predição.

Inicialmente, ALLEN (1971) propôs a estatística $PRESS$ para o modelo linear clássico, definida por $PRESS = \sum_{t=1}^{n} e_{(t)}^2 = \sum_{t=1}^{n}(y_t - \widehat{y}_{(t)})^2$, em que $e_{(t)} = y_t - \widehat{y}_{(t)}$ é chamado erro predito e $\widehat{y}_{(t)} = x_t^\top \widehat{\beta}_{(t)}$ é o valor predito sem a t-ésima observação. A estatística $PRESS$ pode ainda ser reescrita em função da matriz chapéu $H = X(X^\top X)^{-1}X^\top$, ou seja, $PRESS = \sum_{t=1}^{n}(y_t - \widehat{y}_t)^2/(1-h_{tt})^2$, em que h_{tt} é o t-ésimo elemento da diagonal principal da matriz H.

No caso do modelo de regressão simplex temos que $\widehat{\beta}$ em (3.3) pode ser visto como a solução de mínimos quadrados ordinários da regressão

$$\check{\boldsymbol{y}} = \widehat{\Sigma}^{1/2}\widehat{W}^{1/2}\boldsymbol{u}_1 \text{ contra } \check{J}_1 = \widehat{\Sigma}^{1/2}\widehat{W}^{1/2}J_1,$$

em que $\Sigma = \text{diag}\{1/\sigma_1^2, \ldots, 1/\sigma_n^2\}$, $\boldsymbol{u}_1$ é definido em (3.3) e W é uma matriz diagonal cujos elementos estão em (2.11). Portanto, o erro predito é definido por $\check{y}_t - \widehat{\check{y}}_{(t)} = (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}u_{1,t} - (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}j_{1t}^\top\widehat{\beta}_{(t)}$, em que $j_{1t}^\top$ t-ésima linha da matriz J_1. Utilizando um resultado de PREGIBON (1981), temos que

$$\widehat{\boldsymbol{\beta}}_{(t)} = \widehat{\boldsymbol{\beta}} - \{(J_1^\top\widehat{\Sigma}\widehat{W}J_1)^{-1}j_{1t}(1/\widehat{\sigma_t^2})^{1/2}\widehat{w}_t^{1/2}r_t^\beta\}/(1-h_{tt}^*),$$

em que h_{tt}^* é o t-ésimo elemento da diagonal principal da matriz

$$H^* = (\widehat{W}\widehat{\Sigma})^{1/2}J_1(J_1\widehat{\Sigma}\widehat{W}J_1)^{-1}J_1^\top(\widehat{\Sigma}\widehat{W})^{1/2}. \tag{4.1}$$

Aqui, r_t^β é o resíduo ponderado (ESPINHEIRA & SILVA, 2018), definido por

$$r_t^\beta = \frac{\widehat{u}_t(y_t - \widehat{\mu}_t)}{\sqrt{\widehat{b}_t}},$$

em que u_t é dado em (2.7) e

$$b_t = \sigma_t^2\left\{\frac{3\sigma_t^2}{\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}\right\}.$$

Com base nessas quantidades podemos reescrever o erro predito para o modelo de regressão simplex não linear

$$
\begin{aligned}
\check{y}_t \;-\; \widehat{\check{y}}_{(t)} &= (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}u_{1,t} - (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}j_{1t}^{\top}\widehat{\beta}_{(t)} \\
&= (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}u_{1,t} - (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}j_{1t}^{\top}\left\{\hat{\beta} - \frac{(J_1^{\top}\widehat{\Sigma}\widehat{W}J_1)^{-1}j_{1t}(1/\widehat{\sigma}_t^2)^{1/2}\hat{w}_t^{1/2}r_t^{\beta}}{1-h_{tt}^*}\right\} \\
&= (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}u_{1,t} - (1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}j_{1t}^{\top}\hat{\beta} \\
&+ \frac{(1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}j_{1t}^{\top}(J_1^{\top}\widehat{\Sigma}\widehat{W}J_1)^{-1}j_{1t}(1/\widehat{\sigma}_t^2)^{1/2}\widehat{w}_t^{1/2}r_t^{\beta}}{1-h_{tt}^*} \\
&= (1/\widehat{\sigma}_t^2)_t^{1/2}\widehat{w}_t^{1/2}(u_{1,t}-\widehat{\eta}_t) + \frac{h_{tt}^*r_t^{\beta}}{1-h_{tt}^*} = \frac{r_t^{\beta}}{1-h_{tt}^*}.
\end{aligned}
$$

Finalmente, a estatística $PRESS$ para o modelo de regressão simplex não linear fica dada por

$$PRESS^{\circledast} = \sum_{t=1}^{n}(\check{y}_t - \widehat{\check{y}}_{(t)})^2 = \sum_{t=1}^{n}\left(\frac{r_t^{\beta}}{1-h_{tt}^*}\right)^2. \tag{4.2}$$

Segundo LIU el al.(1999), a estatística $PRESS$ identifica o modelo com a melhor habilidade em predizer. Através da $PRESS$, WOLD (1982) propôs um coeficiente de predição para o modelo normal linear, denominado estatística P^2, que é calculada de forma similar ao coeficiente de determinação R^2. Portanto, a estatística P^2 para o modelo de regressão simplex é dada por

$$P^{2\circledast} = 1 - \frac{PRESS^{\circledast}}{SST_{(t)}^*}, \tag{4.3}$$

em que $SST_{(t)}^* = \sum_{t=1}^{n}(\check{y}_t - \bar{\check{y}}_{(t)})^2$ e $\bar{\check{y}}_{(t)}$ é a média aritmética dos $n-1$ valores do vetor $\check{\boldsymbol{y}} = \widehat{\Sigma}^{1/2}\widehat{W}^{1/2}\boldsymbol{u}_1$ excluindo a t-ésima observação. Pode-se mostrar que $SST_{(t)}^* = [n/(n-p)]^2SST^*$, em que p é o número de parâmetros do modelo com $p = k+q$ e SST^* é a Soma de Quadrados Totais com os dados completos.

COOK & WEISBERG (1982) sugerem outras versões da estatística $PRESS$ baseada em diferentes resíduos. Assim, consideramos a estatísticas $PRESS$ e P^2 baseadas no resíduo combinado (Capítulo 3), ou seja,

$$PRESS_{\beta\gamma}^{\circledast} = \sum_{t=1}^{n}\left(\frac{r_{p,t}^{\beta\gamma}}{1-h_{tt}^*}\right)^2 \quad \text{e} \quad P_{\beta\gamma}^{2\circledast} = 1 - \frac{PRESS_{\beta\gamma}^{\circledast}}{SST_{(t)}^*}, \tag{4.4}$$

em que

$$r_{p,t}^{\beta\gamma} = \frac{\widehat{u}_t(y_t - \widehat{\mu}_t) + \widehat{a}_t}{\sqrt{\widehat{b}_t + [2(\widehat{\sigma^2}_t)^2]^{-1}}} \quad \text{e} \quad a_t = -\frac{1}{2\sigma_t^2} + \frac{d(y_t;\mu_t)}{2(\sigma_t^2)^2}.$$

Visto que a $PRESS$ é positiva, as estatísticas $P^{2\circledast}$ e $P_{\beta\gamma}^{2\circledast}$ definidas em (4.3) e (4.4), respectivamente, assumem valores no intervalo $(-\infty; 1]$. Quanto maior for o valor da medida, maior o indicativo do poder preditivo do modelo.

Para avaliar a qualidade de ajuste do modelo consideramos o R_{FC}^2 proposto por BAYER & CRIBARI-NETO (2017) para o modelo de regressão beta, que é definido como o quadrado da correlação entre $g(\boldsymbol{y})$ e η_1 e o R_{LR}^2 baseado na razão de verossimilhanças definido por $R_{LR}^2 = 1 - (L_{null}/L_{fit})^{2/n}$, 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. Além disso, consideramos as versões corrigidas sugeridas por BAYER & CRIBARI-NETO (2017):

$$R_{FC_c}^2 = 1 - (1 - R_{FC}^2)(n-1)/(n-(k_1+q_1)),$$

$$R_{LR_c}^2 = 1 - (1 - R_{LR}^2)\left(\frac{n-1}{n-(1+\alpha)k_1-(1-\alpha)q_1}\right)^{\delta},$$

em que k_1 e q_1 são os números de covariadas no submodelo da média e no submodelo da dispersão, respectivamente, $\alpha \in [0,1]$ e $\delta > 0$. Os autores sugerem $\alpha = 0.4$ e $\delta = 1$. Similarmente a $R_{FC_c}^2$ definimos as versões corrigidas de P^2 e $P_{\beta\gamma}^2$ dadas, respectivamente, por $P_c^2 = 1-(1-P^2)(n-1)/(n-(k_1+q_1))$ e $P_{\beta\gamma_c}^2 = 1-(1-P_{\beta\gamma}^2)(n-1)/(n-(k_1+q_1))$.

4.3 Distância de Cook

As técnicas de diagnósticos de modelos de regressão são de extrema importância, uma vez que, aspectos importantes de um modelo podem ser confundidos por causa de apenas uma observação. Vários métodos foram propostos para detectar observações influentes como por exemplo, a distância de Cook (COOK, 1977), as matrizes de alavanca e as medidas de influência local.

A distância de Cook (COOK, 1977) mede o impacto de uma dada observação nas estimativas dos coeficientes de regressão a partir de sua exclusão do conjunto de dados.

Baseada em aproximações para a versão da medida LD (Likelihood Displacement), a distância de Cook tem sido proposta para diferentes classes de modelos de regressão. Usaremos o deslocamento de verossimilhança (COOK & WEISBERG, 1982) que resulta da remoção da t-ésima observação dos dados e é definida como $LD_t = 2\{\ell_t(\widehat{\boldsymbol{\beta}}) - \ell_t(\widehat{\boldsymbol{\beta}}_{(t)})\}$, em que $\ell_t(\widehat{\boldsymbol{\beta}})$ e $\ell_t(\widehat{\boldsymbol{\beta}}_{(t)})$ são a função log-verossimilhança avaliada na estimativa de máxima verossimilhança de $\boldsymbol{\beta}$ para os dados completos e a função log-verossimilhança avaliada na estimativa de máxima verossimilhança de $\boldsymbol{\beta}$ sem a t-ésima observação, respectivamente.

Nem sempre é possível obter uma expressão fechada para o deslocamento de verossimilhanças sem a t-ésima observação LD_t, $t = 1, \ldots, n$. Assim, é comum utilizar a aproximação

$$LD_t \approx (\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(t)})^\top K_{\beta\beta}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(t)}), \tag{4.5}$$

em que $K_{\beta\beta}$ é dado em (2.22). Para obter $\boldsymbol{\beta}_{(t)}$, PREGIBON (1981) utiliza o esquema de perturbação no qual $\ell_\delta(\beta) = \sum_{t=1}^n \delta_t \ell_t(\mu_t, \sigma_t^2)$, com $\delta_t = 0$ ou $\delta_t = 1$. Quando $\delta_t = 1$, $\forall t$, não existe perturbação, por outro lado, $\delta_t = 0$ implica que a t-ésima observação foi excluída dos dados. O estimador de máxima verossimilhança de $\boldsymbol{\beta}$ para o modelo perturbado pode ser obtido utilizando o método escore de Fisher no qual

$$\boldsymbol{\beta}^{(m+1)} = \boldsymbol{\beta}^{(m)} + (J_1^\top \Sigma^{(m)} \Lambda W^{(m)} J_1)^{-1} J_1^\top \Sigma^{(m)} T^{(m)} \Lambda U^{(m)} (\boldsymbol{y} - \boldsymbol{\mu}^{(m)}),$$

em que $\Lambda = \text{diag}(\delta_1, \ldots, \delta_n)$. PREGIBON (1981) sugere começar o processo iterativo acima com o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ e terminar após um passo. Essa aproximação de um passo gera $\widehat{\boldsymbol{\beta}}_\delta$ que é dado por

$$\widehat{\boldsymbol{\beta}}_\delta = \widehat{\boldsymbol{\beta}} + (J_1^\top \widehat{\Sigma} \Lambda \widehat{W} J_1)^{-1} J_1^\top \widehat{\Sigma} \widehat{T} \Lambda \widehat{U} (\boldsymbol{y} - \widehat{\boldsymbol{\mu}}).$$

que pode ser reescrito por

$$\widehat{\boldsymbol{\beta}}_\delta = (J_1^\top \widehat{\Sigma} \Lambda \widehat{W} J_1)^{-1} J_1^\top \widehat{\Sigma} \Lambda \widehat{W} \boldsymbol{u}_1,$$

em que

$$\boldsymbol{u}_1 = \widehat{\boldsymbol{\eta}}_1 + \widehat{W}^{-1} (\boldsymbol{y} - \widehat{\boldsymbol{\mu}}),$$

com $\widehat{\boldsymbol{\eta}}_1 = J_1\widehat{\boldsymbol{\beta}}$. Podemos interpretar $\widehat{\boldsymbol{\beta}}_\delta$ como a solução de mínimos quadrados da regressão linear de $\widehat{\Sigma}^{1/2}\widehat{W}^{1/2}\boldsymbol{u}_1$ contra $\widehat{\Sigma}^{1/2}\widehat{W}^{1/2}J_1$ com pesos Λ. Seja $\delta_t = 0$, $\delta_l = 1$, $\forall l \neq t$, ou seja, a t-ésima observação excluída dos dados. Segue que

$$\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(t)} \approx \frac{(J_1^\top \widehat{\Sigma}\widehat{W}J_1)^{-1} j_{1t}(/1/\sigma_t^2)^{1/2} w_t^{1/2} r_t^\beta}{1 - h_{tt}^*}, \tag{4.6}$$

em que r_t^β é dado em (3.4) e h_{tt}^* é o t-ésimo elemento da diagonal principal da matriz H^* dada em 4.1.

Utilizando as equações (4.5), (2.22) e (4.6) nós obtemos a distância de Cook para o modelo de regressão simplex não linear que é dado por

$$LD_t \approx \frac{h_{tt}^*(r_t^\beta)^2}{(1 - h_{tt}^*)(1 - h_{tt}^*)}.$$

A estatística $PRESS$ possui uma relação com medidas de influência. Para o modelo simplex a distância de Cook mede o impacto de uma dada observação na estimação dos parâmetros do submodelo da média removendo-a dos dados. Observe a relação

$$LD_t \approx \frac{h_{tt}^*(r_t^\beta)^2}{(1 - h_{tt}^*)^2} \Rightarrow \frac{(r_t^\beta)^2}{(1 - h_{tt}^*)^2} \approx \frac{LD_t}{h_{tt}^*}. \tag{4.7}$$

Utilizando (4.2) e (4.7) nós obtemos

$$PRESS^{\circledast} \approx \sum_{t=1}^{n} \frac{LD_t}{h_{tt}^*},$$

em que LD_t é a distância de Cook para o modelo de regressão simplex. Além disso, LESAFFRE & VERBEKE (1998) mostraram que $LD_t \approx C_t$ em que C_t é a influência local total da observação t, definida por

$$C_t = 2|\Delta_t^\top \ddot{\ell}^{-1}\Delta_t|,$$

em que Δ_t é a t-ésima coluna de $\Delta = \partial^2\ell_\delta(\theta)/\partial\theta\partial\delta^\top$, $\ell_\delta(\theta)$ é a log verossimilhança do modelo perturbado para um dado δ e $\ddot{\ell} = \partial^2\ell(\widehat{\theta})/\partial\theta\partial\theta^\top$. Assim, $PRESS^{\circledast} \approx \sum_{t=1}^{n} \frac{C_t}{h_{tt}^*}$ mostra a relação desta estatística com as medidas de influência local. LESAFFRE & VERBEKE (1998) também sugerem que observações tais que $C_t > 2\sum_{t=1}^{n} C_t/n$ podem ser tomadas como individualmente influentes. Quando os preditores em (2.17) são funções lineares dos parâmetros, a expressão em (4.2) representa a estatística $PRESS$ para

classe de modelos de regressão simplex linear com $p = k + q$ parâmetros de regressão desconhecidos.

4.4 Avaliação Numérica

Nesta seção utilizamos simulação de Monte Carlo para avaliar o comportamento das distribuições das estatísticas de predição para o modelo de regressão simplex linear e não linear. Todas as simulações foram realizadas com 10000 réplicas através do programa de linguagem matricial Ox, para mais detalhes ver `http://www.doornik.com`. Nosso objetivo também é avaliar as medidas de qualidade de ajuste R^2_{FC} e R^2_{LR}, assim como suas versões corrigidas, para o modelo simplex.

A Tabela 22 mostra os valores médios das estatísticas obtidas das simulações utilizando o modelo de regressão simplex com dispersão constante dado por

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

As realizações das covariadas foram geradas de forma independente através da distribuição $x_t \sim U(0,1)$, $t = 1, \ldots, n$, e mantiveram-se fixas em todas as réplicas. Foram considerados os seguintes valores para o parâmetro de dispersão $\sigma^2 = (0.4, 3.5, 6.0)$ e o tamanho da amostra igual a 80. Além disso, consideramos que a média da variável resposta pertence aos intervalos $\mu \in (0.02, 0.25)$, $\mu \in (0.20, 0.88)$ e $\mu \in (0.80, 0.99)$. Para verificar o comportamento das estatísticas sob omissão de covariadas consideramos os cenários 1, 2 e 3, nos quais são omitidas 3, 2 e 1 covariadas, respectivamente. No cenário 4, o modelo está corretamente especificado, ou seja, é gerado e estimado com quatro covariadas. Ressaltamos que na Tabela 22 são apresentados os valores médios das 10000 réplicas de Monte Carlo das estatísticas.

Inicialmente verificamos que os valores das estatísticas P^2, $P^2_{\beta\gamma}$ e suas versões corrigidas aumentam à medida que as covariadas são inseridas no modelo. O mesmo acontece quando o parâmetro de dispersão é menor, por exemplo, no cenário 1 para $\mu \in (0.20; 0.88)$, $n = 40$ e $\sigma^2 = (0.4, 3.5, 6.0)$, os valores da $P^2_{\beta\gamma}$ corrigida foram 0.245, 0.053 e 0.012, respectivamente. Isso é observado para todos os tamanhos amostrais considerados. É possível notar ainda que os valores dos coeficientes de predição são maiores quando a

média da variável resposta se encontra próxima dos extremos do intervalo unitário padrão, comparados com os mesmos cenários quando $\mu \approx (0.20; 0.88)$, indicando maior dificuldade de predição quando μ assume valores centrais do intervalo (0,1). Como vimos no capítulo anterior, o processo de estimação por máxima verossimilhança do modelo simplex é menos sensível que o do modelo de regressão beta, nesse cenário.

Os valores das estatísticas R^2_{LR} e $R^2_{LR_c}$ também aumentam à medida que covariadas importantes são incorporadas ao modelo. No entanto, quando μ está próximo dos extremos do intervalo unitário, os valores destas estatísticas são muito altas mesmo sob omissão de três covariadas (Cenário 1). Neste cenário, enquanto os valores das medidas de predição estão próximas de 0.70 para $\sigma^2 = 0.4$, por exemplo, os valores do R^2_{LR} estão próximos de 0.90.

É muito interessante notar como tanto o poder preditivo do modelo quanto a qualidade do ajuste são afetados negativamente quando a média da variável resposta está concentrada em torno de 0.5, valores centrais do intervalo unitário, em especial quando a dispersão do modelo aumenta. Esse fato fica evidente no cenário 4 em que o modelo está corretamente especificado. Para $\sigma^2 = 0.4$ e $n = 120$ os valores de todas as estatísticas estão próximas de 0.90. Quando $\sigma^2 = 3.5$ e $\sigma^2 = 6.0$ os valores médios das estatísticas caem para aproximadamente 0.55 e 0.45, respectivamente. Já para $\mu \approx 0$ e $\mu \approx 1$ essas mesmas estatísticas apresentam valores médios aproximadamente iguais a 0.99 ($\sigma^2 = 0.4$), 0.96 ($\sigma^2 = 3.5$) e 0.94 ($\sigma^2 = 6.0$).

Na Tabela 23 apresentamos os valores médios da P^2, R^2_{FC}, R^2_{LR} e suas versões corrigidas considerando o modelo com dispersão variável dado por

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4} + \beta_5\, x_{t5}, \\ \log(\sigma_t^2) &= \gamma_1 + \gamma_2 z_{t2} + \gamma_3 z_{t3} + \gamma_4 z_{t4} + \gamma_5 z_{t5}, \end{aligned} \tag{4.8}$$

$t = 1, \ldots, n$.

Como anteriormente, foram considerados os tamanhos amostrais $n = (40, 80, 120)$ e os intervalos para a média da variável resposta $\mu \in (0.02, 0.25)$, $\mu \in (0.20, 0.88)$ e $\mu \in (0.80, 0.99)$. Além disso, consideramos $\lambda = (20, 50, 100)$.

Para verificar o comportamento das distribuições das estatísticas consideramos quatro cenários corretamente especificados. No primeiro, foi gerada uma covariada para o

Tabela 22 – Valores médios das estatísticas. Modelo verdadeiro: $g(\mu_t) = \log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4} + \beta_5\, x_{t5},\ x_{ti} \sim U(0,1)$, $i = 2,3,4,5,\ t = 1,\ldots,n$ e σ^2 constante. Modelo mal especificado: omissão de covariadas (Cenários 1, 2 e 3).

	Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
	Modelo estimado	$g(\mu_t) = \beta_1 + \beta_2\, x_{t2}$			$g(\mu_t) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3}$			$g(\mu_t) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4}$			$g(\mu_t) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4} + \beta_5\, x_{t5}$		
μ	$\mu \in (0.20; 0.88)$; $\beta = (-1.9, 1.2, 1.0, 1.1, 1.3)^\top$.												
n	σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
40	P^2	0.278	0.102	0.063	0.472	0.256	0.206	0.663	0.411	0.351	0.910	0.622	0.547
	P^2_c	0.239	0.053	0.012	0.428	0.194	0.140	0.625	0.343	0.277	0.897	0.567	0.481
	$P^2_{\beta\gamma}$	0.284	0.102	0.063	0.470	0.256	0.206	0.660	0.410	0.351	0.911	0.622	0.547
	$P^2_{\beta\gamma_c}$	0.245	0.053	0.012	0.426	0.193	0.140	0.621	0.343	0.276	0.898	0.566	0.481
	R^2_{LR}	0.474	0.295	0.251	0.576	0.367	0.315	0.698	0.451	0.389	0.894	0.581	0.503
	$R^2_{LR_c}$	0.446	0.257	0.211	0.541	0.314	0.258	0.664	0.388	0.319	0.879	0.520	0.430
80	P^2	0.336	0.151	0.110	0.496	0.260	0.206	0.665	0.382	0.315	0.901	0.579	0.492
	P^2_c	0.318	0.129	0.087	0.476	0.231	0.175	0.647	0.349	0.278	0.894	0.551	0.457
	$P^2_{\beta\gamma}$	0.341	0.151	0.110	0.495	0.260	0.206	0.662	0.382	0.315	0.901	0.579	0.492
	$P^2_{\beta\gamma_c}$	0.324	0.129	0.087	0.475	0.231	0.175	0.644	0.349	0.278	0.895	0.550	0.457
	R^2_{LR}	0.471	0.284	0.238	0.571	0.350	0.295	0.692	0.428	0.362	0.887	0.555	0.472
	$R^2_{LR_c}$	0.457	0.265	0.218	0.554	0.324	0.267	0.675	0.397	0.328	0.880	0.525	0.437
120	P^2	0.352	0.163	0.122	0.503	0.260	0.204	0.665	0.372	0.302	0.898	0.564	0.472
	P^2_c	0.341	0.149	0.107	0.490	0.241	0.184	0.653	0.351	0.278	0.894	0.545	0.449
	$P^2_{\beta\gamma}$	0.357	0.164	0.122	0.501	0.260	0.204	0.662	0.372	0.302	0.898	0.564	0.472
	$P^2_{\beta\gamma_c}$	0.346	0.149	0.107	0.489	0.240	0.184	0.650	0.350	0.278	0.894	0.545	0.449
	R^2_{LR}	0.470	0.280	0.233	0.570	0.344	0.288	0.690	0.420	0.353	0.885	0.547	0.462
	$R^2_{LR_c}$	0.461	0.268	0.220	0.559	0.327	0.269	0.679	0.400	0.331	0.880	0.527	0.439
μ	$\mu \in (0.80.0.99)$; $\beta = (2.0, 1.4, 0.8, -1.3, 1.0)^\top$.												
n	σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
40	P^2	0.751	0.628	0.554	0.853	0.781	0.732	0.973	0.934	0.905	0.996	0.967	0.946
	P^2_c	0.737	0.607	0.530	0.841	0.763	0.710	0.970	0.926	0.894	0.996	0.962	0.938
	$P^2_{\beta\gamma}$	0.751	0.628	0.554	0.853	0.781	0.732	0.973	0.934	0.905	0.996	0.967	0.946
	$P^2_{\beta\gamma_c}$	0.737	0.607	0.530	0.841	0.763	0.710	0.970	0.926	0.894	0.996	0.962	0.938
	R^2_{LR}	0.909	0.886	0.870	0.929	0.905	0.889	0.977	0.949	0.930	0.996	0.966	0.946
	$R^2_{LR_c}$	0.904	0.880	0.863	0.923	0.897	0.880	0.975	0.943	0.922	0.995	0.961	0.939
80	P^2	0.775	0.649	0.572	0.869	0.785	0.728	0.974	0.931	0.898	0.996	0.964	0.939
	P^2_c	0.769	0.640	0.560	0.864	0.777	0.717	0.973	0.927	0.893	0.995	0.961	0.935
	$P^2_{\beta\gamma}$	0.775	0.649	0.572	0.869	0.785	0.728	0.974	0.931	0.898	0.996	0.964	0.939
	$P^2_{\beta\gamma_c}$	0.769	0.640	0.561	0.864	0.777	0.717	0.973	0.927	0.893	0.995	0.961	0.935
	R^2_{LR}	0.909	0.884	0.868	0.929	0.903	0.885	0.977	0.947	0.926	0.995	0.964	0.943
	$R^2_{LR_c}$	0.906	0.881	0.864	0.926	0.899	0.881	0.976	0.944	0.922	0.995	0.961	0.939
120	P^2	0.782	0.654	0.576	0.872	0.785	0.725	0.974	0.930	0.896	0.996	0.962	0.936
	P^2_c	0.778	0.649	0.568	0.869	0.780	0.718	0.973	0.927	0.892	0.995	0.961	0.934
	$P^2_{\beta\gamma}$	0.782	0.654	0.576	0.872	0.785	0.725	0.974	0.930	0.896	0.996	0.962	0.936
	$P^2_{\beta\gamma_c}$	0.778	0.649	0.568	0.869	0.780	0.718	0.973	0.927	0.892	0.995	0.961	0.934
	R^2_{LR}	0.909	0.884	0.867	0.929	0.902	0.884	0.977	0.946	0.925	0.995	0.963	0.941
	$R^2_{LR_c}$	0.907	0.882	0.865	0.927	0.899	0.881	0.976	0.944	0.923	0.995	0.961	0.939
μ	$\mu \in (0.02; 0.25)$; $\beta = (-3.5, 1.2, 0.7, -1.0, 1.0)^\top$.												
n	σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
40	P^2	0.650	0.506	0.428	0.796	0.686	0.622	0.944	0.885	0.844	0.994	0.954	0.925
	P^2_c	0.631	0.479	0.397	0.779	0.660	0.590	0.938	0.872	0.827	0.994	0.947	0.913
	$P^2_{\beta\gamma}$	0.650	0.506	0.428	0.796	0.686	0.621	0.944	0.885	0.844	0.995	0.954	0.925
	$P^2_{\beta\gamma_c}$	0.631	0.479	0.397	0.779	0.660	0.590	0.938	0.872	0.826	0.994	0.947	0.913
	R^2_{LR}	0.915	0.892	0.876	0.931	0.907	0.891	0.973	0.946	0.928	0.996	0.967	0.948
	$R^2_{LR_c}$	0.911	0.886	0.869	0.926	0.900	0.882	0.970	0.940	0.919	0.995	0.962	0.940
80	P^2	0.673	0.526	0.446	0.802	0.682	0.610	0.947	0.881	0.835	0.994	0.948	0.914
	P^2_c	0.664	0.514	0.431	0.794	0.670	0.595	0.945	0.875	0.826	0.994	0.945	0.908
	$P^2_{\beta\gamma}$	0.673	0.526	0.446	0.802	0.682	0.610	0.947	0.881	0.835	0.994	0.948	0.914
	$P^2_{\beta\gamma_c}$	0.664	0.514	0.431	0.794	0.669	0.595	0.945	0.875	0.826	0.994	0.945	0.908
	R^2_{LR}	0.915	0.890	0.873	0.931	0.905	0.888	0.973	0.944	0.924	0.996	0.965	0.944
	$R^2_{LR_c}$	0.913	0.887	0.870	0.928	0.902	0.884	0.971	0.941	0.920	0.995	0.962	0.940
120	P^2	0.679	0.527	0.449	0.803	0.679	0.603	0.948	0.879	0.831	0.994	0.946	0.910
	P^2_c	0.674	0.519	0.439	0.798	0.671	0.593	0.946	0.875	0.825	0.993	0.944	0.906
	$P^2_{\beta\gamma}$	0.679	0.527	0.448	0.803	0.679	0.603	0.948	0.879	0.830	0.994	0.946	0.910
	$P^2_{\beta\gamma_c}$	0.674	0.519	0.439	0.798	0.671	0.593	0.946	0.875	0.825	0.993	0.944	0.906
	R^2_{LR}	0.915	0.889	0.872	0.931	0.905	0.887	0.972	0.943	0.923	0.995	0.964	0.943
	$R^2_{LR_c}$	0.914	0.887	0.870	0.929	0.902	0.884	0.972	0.941	0.920	0.995	0.962	0.940

submodelo da média e uma covariada para o submodelo da dispersão, e os modelos foram estimados de forma correta, ambos com uma covariada. No segundo, terceiro e quarto cenários foram gerados 2, 3 e 4 covariadas para o submodelo da média e da dispersão, respectivamente, e estimados da mesma forma. As realizações das covariadas foram geradas de forma independente através da distribuição $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $t = 1, \ldots, n$, e mantiveram-se fixas em todas as réplicas.

Obtemos os seguintes valores para o máximo e o mínimo de σ^2 com $n = 80$, Cenário 1: (8.4167,0.4071) para $\lambda = 20$, (12.3038,0.2465) para $\lambda = 50$ e (14.7883,0.1475) para $\lambda = 100$; Cenário 2: (11.4773,0.5783) para $\lambda = 20$, (16.8088,0.3357) para $\lambda = 50$ e (22.3774,0.2233) para $\lambda = 100$; Cenário 3: (13.9517,0.6845) para $\lambda = 20$, (17.8377,0.3559) para $\lambda = 50$ e (33.0118,0.3302) para $\lambda = 100$; e por fim Cenário 4: (10.1452,0.5035) para $\lambda = 20$, (20.4167,0.4029) para $\lambda = 50$ e (23.0288,0.2289) para $\lambda = 100$.

Mais uma vez vimos que as estatísticas se comportam melhor quando os dados estão concentrados próximo de 0 ou 1. Como já sabemos é uma característica do modelo simplex. Os valores iguais a um das estatísticas R^2_{LR} e $R^2_{LR_c}$ se dá ao fato da função de verossimilhança maximizada do modelo ajustado não apresentar um valor válido, ou seja, o quociente entre a função de verossimilhança maximizada do modelo sem regressores com a função de verossimilhança maximizada do modelo ajustado tende à zero e consequentemente R^2_{LR} assume o valor um.

Na Figura 19 apresentamos resultados com especificação incorreta do modelo, a saber: omissão de três covariadas importantes. A distribuição das estatísticas de predição apresentam maior dispersão à medida que os valores de σ^2 aumentam. No entanto, em todos os casos a média aritmética é praticamente igual à mediana. Os valores mais distantes de um indicam que o poder preditivo está sendo afetado por alguma quebra de suposições do modelo, neste caso a suposição que o modelo verdadeiro é formado de fato por quatro covariadas ao invés de apenas uma. Neste cenário as estatísticas de predição podem apresentar valores muito reduzidos até próximos de zero. A distribuição da estatística R^2_{FC} apresenta menos dispersão e os valores estão mais concentrados e abaixo de 0.4.

Na Figura 20 apresentamos os boxplots das 10000 réplicas de Monte Carlo da estatística R^2_{LR} e sua versão corrigida. Consideramos o Cenário 4 do modelo com dispersão constante na qual possui quatro covariadas importantes e o modelo é estimado correta-

Tabela 23 – Valores das médias das estatísticas. Modelo corretamente especificado. $g(\mu_t) = \log(\mu_t/(1-\mu_t))$ e $h(\sigma_t^2) = \log(\sigma_t^2)$, $x_t \sim U(0,1)$ e $z_t \sim U(-0.5, 0.5)$, $i = 2, 3, 4, 5,\ t = 1, \ldots, n$.

	Cenários	Cenário 1			Cenário 2			Cenário 3			Cenário 4		
	Modelo da média	$g(\mu_t) = \beta_1 + \beta_2\, x_{t2}$			$g(\mu_t) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3}$			$g(\mu_t) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4}$			$g(\mu_t) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4} + \beta_5\, x_{t5}$		
	Modelo da dispersão	$h(\sigma_t^2) = \gamma_1 + \gamma_2\, z_{t2}$			$h(\sigma_t^2) = \gamma_1 + \gamma_2\, z_{t2} + \gamma_3\, z_{t3}$			$h(\sigma_t^2) = \gamma_1 + \gamma_2\, z_{t2} + \gamma_3\, z_{t3} + \gamma_4\, z_{t4}$			$h(\sigma_t^2) = \gamma_1 + \gamma_2\, z_{t2} + \gamma_3\, z_{t3} + \gamma_4\, z_{t4} + \gamma_5\, z_{t5}$		
μ		$\mu \in (0.20; 0.88)$.											
n	λ	20	50	100	20	50	100	20	50	100	20	50	100
40	P^2	0.877	0.902	0.928	0.737	0.745	0.715	0.724	0.801	0.788	0.794	0.839	0.870
	$P^2_{\mathcal{L}}$	0.867	0.894	0.923	0.698	0.708	0.673	0.663	0.757	0.741	0.733	0.791	0.831
	$P^2_{\beta\gamma}$	0.877	0.903	0.929	0.737	0.746	0.715	0.725	0.801	0.788	0.797	0.842	0.874
	$P^2_{\beta\gamma_c}$	0.867	0.895	0.924	0.699	0.708	0.673	0.665	0.758	0.742	0.737	0.795	0.836
	R^2_{LR}	0.999	0.999	0.999	0.976	0.987	0.997	0.988	0.986	0.995	0.989	0.993	0.994
	$R^2_{LR_c}$	0.999	0.999	0.999	0.971	0.984	0.996	0.984	0.982	0.993	0.984	0.990	0.991
80	P^2	0.861	0.890	0.919	0.687	0.696	0.662	0.644	0.737	0.720	0.706	0.767	0.809
	$P^2_{\mathcal{L}}$	0.856	0.885	0.916	0.666	0.676	0.639	0.610	0.712	0.693	0.668	0.737	0.785
	$P^2_{\beta\gamma}$	0.862	0.890	0.920	0.687	0.696	0.662	0.645	0.738	0.720	0.707	0.768	0.811
	$P^2_{\beta\gamma_c}$	0.856	0.886	0.917	0.666	0.676	0.639	0.610	0.712	0.693	0.669	0.738	0.786
	R^2_{LR}	0.999	1.000	1.000	0.981	0.991	0.999	0.992	0.990	0.998	0.991	0.995	0.996
	$R^2_{LR_c}$	0.999	1.000	1.000	0.979	0.990	0.999	0.990	0.989	0.997	0.989	0.994	0.995
120	P^2	0.856	0.885	0.916	0.669	0.678	0.642	0.616	0.715	0.697	0.676	0.742	0.788
	$P^2_{\mathcal{L}}$	0.852	0.882	0.914	0.654	0.664	0.626	0.592	0.698	0.678	0.649	0.720	0.771
	$P^2_{\beta\gamma}$	0.856	0.885	0.916	0.669	0.678	0.642	0.617	0.715	0.697	0.676	0.742	0.789
	$P^2_{\beta\gamma_c}$	0.852	0.883	0.914	0.654	0.664	0.626	0.593	0.698	0.678	0.650	0.721	0.772
	R^2_{LR}	0.999	1.000	1.000	0.983	0.993	0.999	0.993	0.992	0.998	0.992	0.996	0.997
	$R^2_{LR_c}$	0.999	1.000	1.000	0.982	0.992	0.999	0.992	0.991	0.998	0.991	0.995	0.996
μ		$\mu \in (0.80.0.99)$.											
n	λ	20	50	100	20	50	100	20	50	100	20	50	100
40	P^2	0.992	0.994	0.996	0.982	0.988	0.988	0.963	0.983	0.983	0.995	0.996	0.997
	$P^2_{\mathcal{L}}$	0.991	0.994	0.996	0.980	0.986	0.986	0.955	0.980	0.979	0.994	0.995	0.996
	$P^2_{\beta\gamma}$	0.992	0.994	0.996	0.982	0.988	0.988	0.963	0.983	0.983	0.995	0.996	0.997
	$P^2_{\beta\gamma_c}$	0.991	0.994	0.996	0.980	0.986	0.986	0.955	0.980	0.979	0.994	0.995	0.996
	R^2_{LR}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$R^2_{LR_c}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
80	P^2	0.991	0.994	0.996	0.979	0.986	0.986	0.949	0.976	0.976	0.993	0.994	0.995
	$P^2_{\mathcal{L}}$	0.990	0.993	0.996	0.977	0.985	0.985	0.943	0.974	0.973	0.992	0.993	0.994
	$P^2_{\beta\gamma}$	0.991	0.994	0.996	0.979	0.986	0.986	0.948	0.976	0.976	0.993	0.994	0.995
	$P^2_{\beta\gamma_c}$	0.990	0.993	0.996	0.977	0.985	0.985	0.943	0.974	0.973	0.992	0.993	0.994
	R^2_{LR}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$R^2_{LR_c}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
120	P^2	0.990	0.993	0.996	0.978	0.985	0.985	0.943	0.974	0.973	0.992	0.993	0.994
	$P^2_{\mathcal{L}}$	0.990	0.993	0.996	0.977	0.984	0.984	0.940	0.972	0.972	0.991	0.992	0.993
	$P^2_{\beta\gamma}$	0.990	0.993	0.996	0.978	0.985	0.985	0.943	0.974	0.973	0.992	0.993	0.994
	$P^2_{\beta\gamma_c}$	0.990	0.993	0.996	0.977	0.984	0.984	0.940	0.972	0.972	0.991	0.992	0.993
	R^2_{LR}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$R^2_{LR_c}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
μ		$\mu \in (0.02; 0.25)$.											
n	λ	20	50	100	20	50	100	20	50	100	20	50	100
40	P^2	0.844	0.904	0.943	0.974	0.980	0.979	0.993	0.986	0.986	0.993	0.993	0.994
	$P^2_{\mathcal{L}}$	0.830	0.897	0.938	0.970	0.977	0.976	0.991	0.983	0.983	0.991	0.990	0.992
	$P^2_{\beta\gamma}$	0.844	0.905	0.943	0.974	0.980	0.979	0.993	0.986	0.986	0.993	0.993	0.994
	$P^2_{\beta\gamma_c}$	0.831	0.897	0.939	0.970	0.977	0.976	0.991	0.983	0.983	0.991	0.990	0.992
	R^2_{LR}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$R^2_{LR_c}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
80	P^2	0.827	0.894	0.936	0.967	0.976	0.975	0.990	0.980	0.979	0.989	0.987	0.989
	$P^2_{\mathcal{L}}$	0.820	0.890	0.934	0.965	0.974	0.973	0.989	0.978	0.977	0.987	0.985	0.988
	$P^2_{\beta\gamma}$	0.827	0.895	0.937	0.967	0.976	0.975	0.990	0.980	0.979	0.989	0.987	0.989
	$P^2_{\beta\gamma_c}$	0.820	0.890	0.934	0.965	0.974	0.973	0.989	0.978	0.977	0.987	0.985	0.988
	R^2_{LR}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$R^2_{LR_c}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
120	P^2	0.821	0.890	0.934	0.965	0.974	0.973	0.989	0.977	0.977	0.987	0.985	0.988
	$P^2_{\mathcal{L}}$	0.816	0.887	0.932	0.964	0.973	0.972	0.988	0.976	0.975	0.986	0.984	0.987
	$P^2_{\beta\gamma}$	0.821	0.890	0.934	0.965	0.974	0.973	0.989	0.977	0.977	0.987	0.985	0.988
	$P^2_{\beta\gamma_c}$	0.816	0.888	0.932	0.964	0.973	0.972	0.988	0.976	0.976	0.986	0.984	0.987
	R^2_{LR}	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	$R^2_{LR_c}$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

mente para $n = 40$ observações. Através dessa figura podemos notar que as estatísticas R^2_{LR} e $R^2_{LR_c}$ permanecem próximas de um quando a média da variável resposta está próxima dos extremos do intervalo (0,1), já para μ central tais estatísticas diminuem conforme σ^2 aumenta.

As Figuras 21 e 22 apresentam os boxplots das 10000 réplicas de Monte Carlo das estatísticas de predição e de qualidade de ajuste sob especificação incorreta. Geramos o modelo dado em 4.8 com apenas uma covariada para ambos os submodelos e estimamos com o modelo simplex com dispersão constante também com uma covariada. Na Figura 21 consideramos o grau de dispersão não constante $\lambda = 20$ e na Figura 22 $\lambda = 100$, com $n = 120$. É possível perceber que o coeficiente de determinação R^2_{FC} (FERRARI & CRIBARI-NETO, 2004) possui mediana baixa quando $\mu \in (0.80, 0.99)$ e uma variação bastante elevada quando comparada com as demais medidas, que permanecem sempre próximas de um. Quando $\mu \in (0.20, 0.88)$ e $\mu \approx 1$, as estatísticas de predição são similares aos cenários em que não há erros de especificação, o que significa que as medidas não conseguem identificar esse tipo de erro. No entanto, quando $\mu \approx 0$, diminuem de forma significativa. Não existe muita diferença quando o grau de dispersão não constante aumenta de 20 para 100.

A seguir apresentamos o comportamento das distribuições das estatísticas de predição e de qualidade de ajuste para o modelo de regressão simplex não linear. Inicialmente consideramos o seguinte modelo com dispersão constante

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}, \tag{4.9}$$

em que $t = 1, \ldots, n$. As realizações das covariadas foram geradas através das distribuições $x_{t2} \sim U(-0.5, 0.5)$, $x_{t3} \sim U(0, 1)$ e $x_{t4} \sim U(-0.5, 0.5)$. Consideramos ainda três intervalos para a média da variável resposta $\mu \in (0.0314; 0.3654)(\beta = (-2.4, 1.4, -1.5, -1.7))$, $\mu \in (0.1970; 0.8684)(\beta = (-1.7, -1.8, 1.2, -1.3))$ e $\mu \in (0.7854; 0.9876)(\beta = (2.1, -1.5, -1.6, -1.2))$, e os seguintes valores para o parâmetro de dispersão $\sigma^2 = (0.4, 3.5, 6.0)$.

A Tabela 24 mostra a média das 10000 réplicas das medidas P^2, $P^2_{\beta\gamma}$, R^2_{LR} e suas versões corrigidas com omissão de covariadas, ou seja, consideramos os cenários 1 e 2, na qual são omitidas 2 e 1 covariada, respectivamente, e o cenário 3, em que o modelo está corretamente especificado, pois é gerado e estimado com três covariadas. Vemos que à

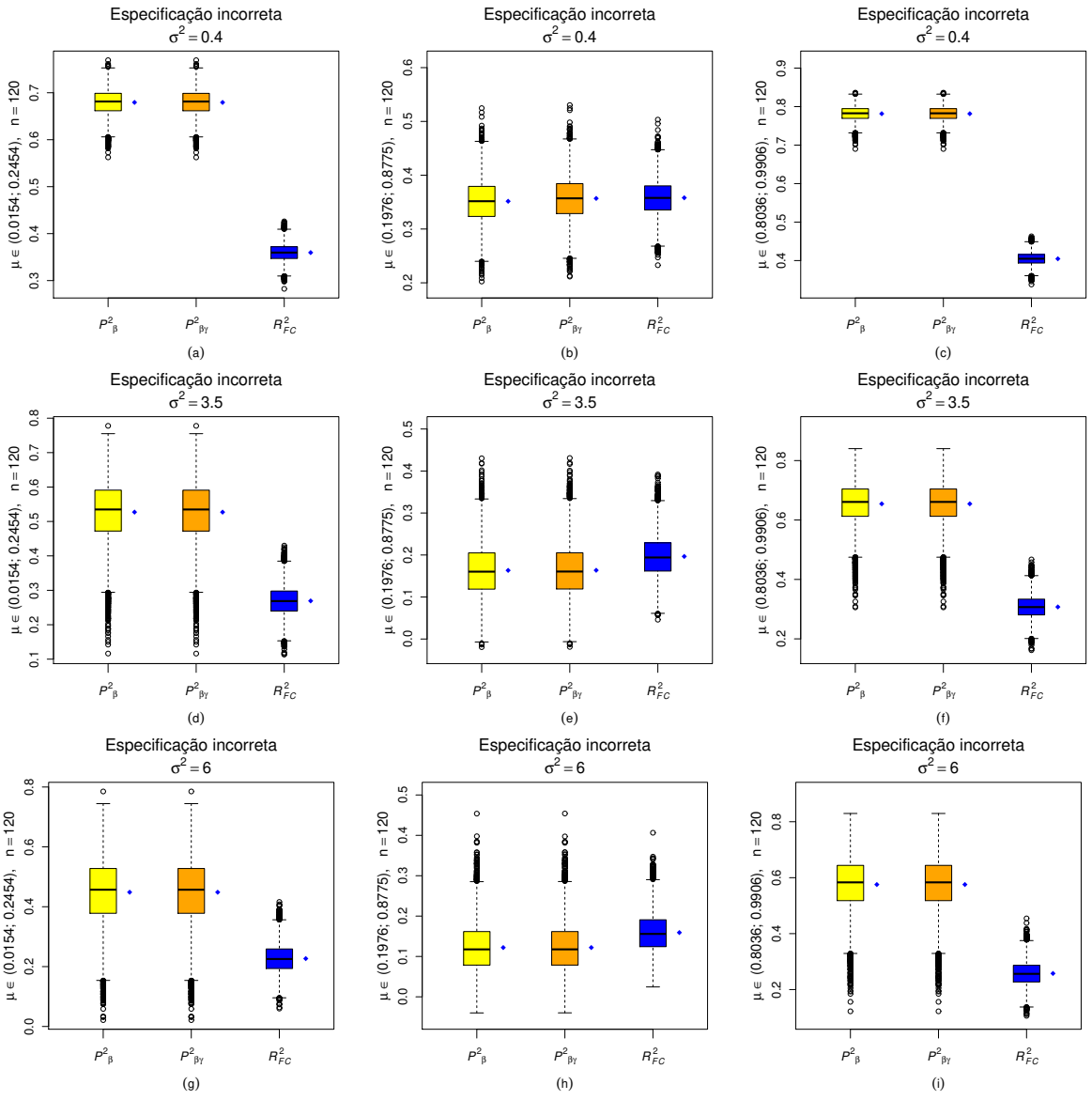

Figura 19 – Boxplots das 10000 réplicas de Monte Carlo das estatísticas P^2_β, $P^2_{\beta\gamma}$ e R^2_{FC}. Modelo verdadeiro: $g(\mu_t) = \log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2\, x_{t2} + \beta_3\, x_{t3} + \beta_4\, x_{t4} + \beta_5\, x_{t5}$. Modelo estimado: $g(\mu_t) = \log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2\, x_{t2}$, $n = 120$.

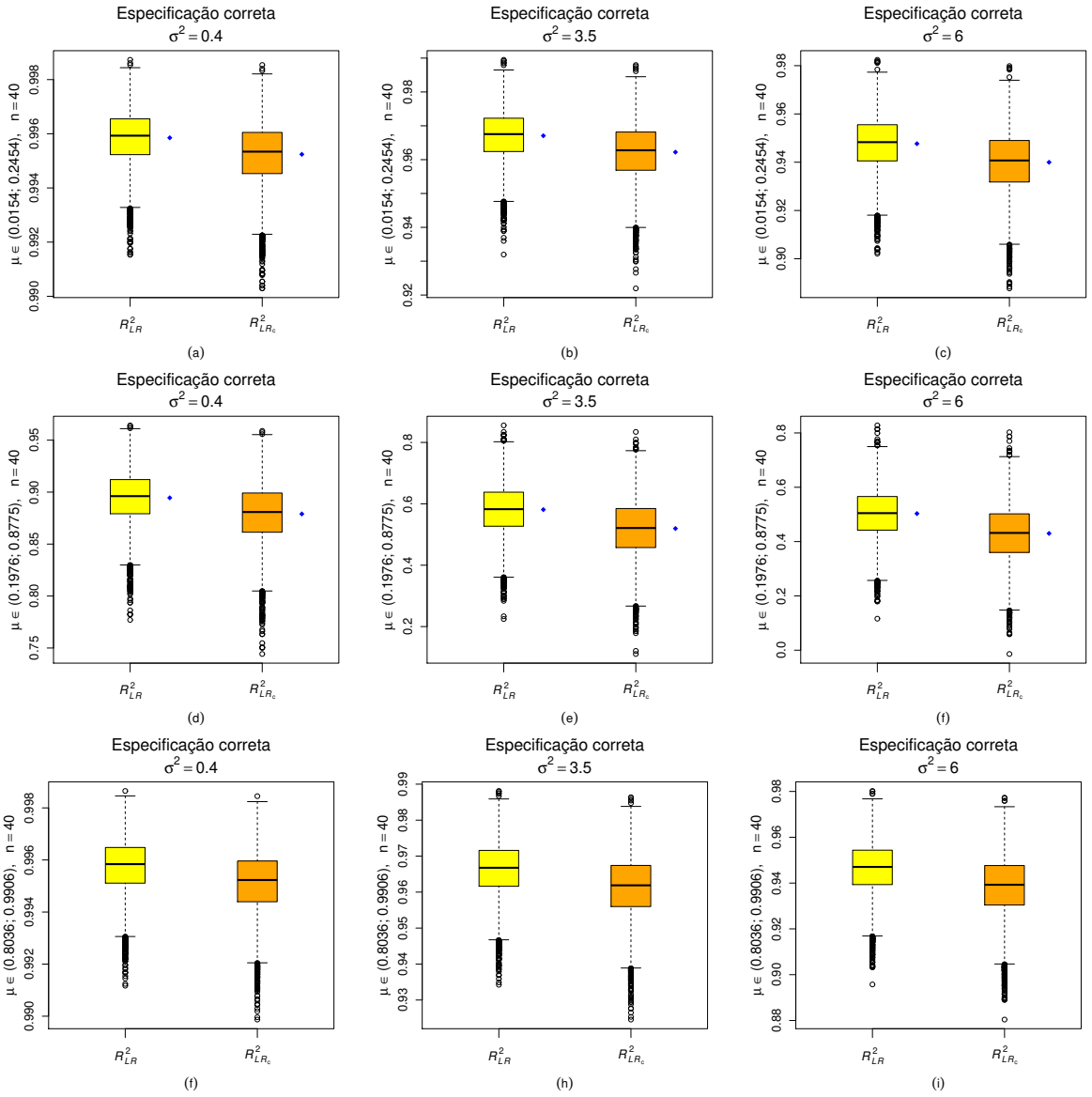

Figura 20 – Boxplots das 10000 réplicas de Monte Carlo das estatísticas R^2_{LR} e R^2_{LRc}. Modelo verdadeiro: $g(\mu_t) = \log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$. Modelo estimado: $g(\mu_t) = \log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \beta_4 x_{t4} + \beta_5 x_{t5}$, $n = 40$.

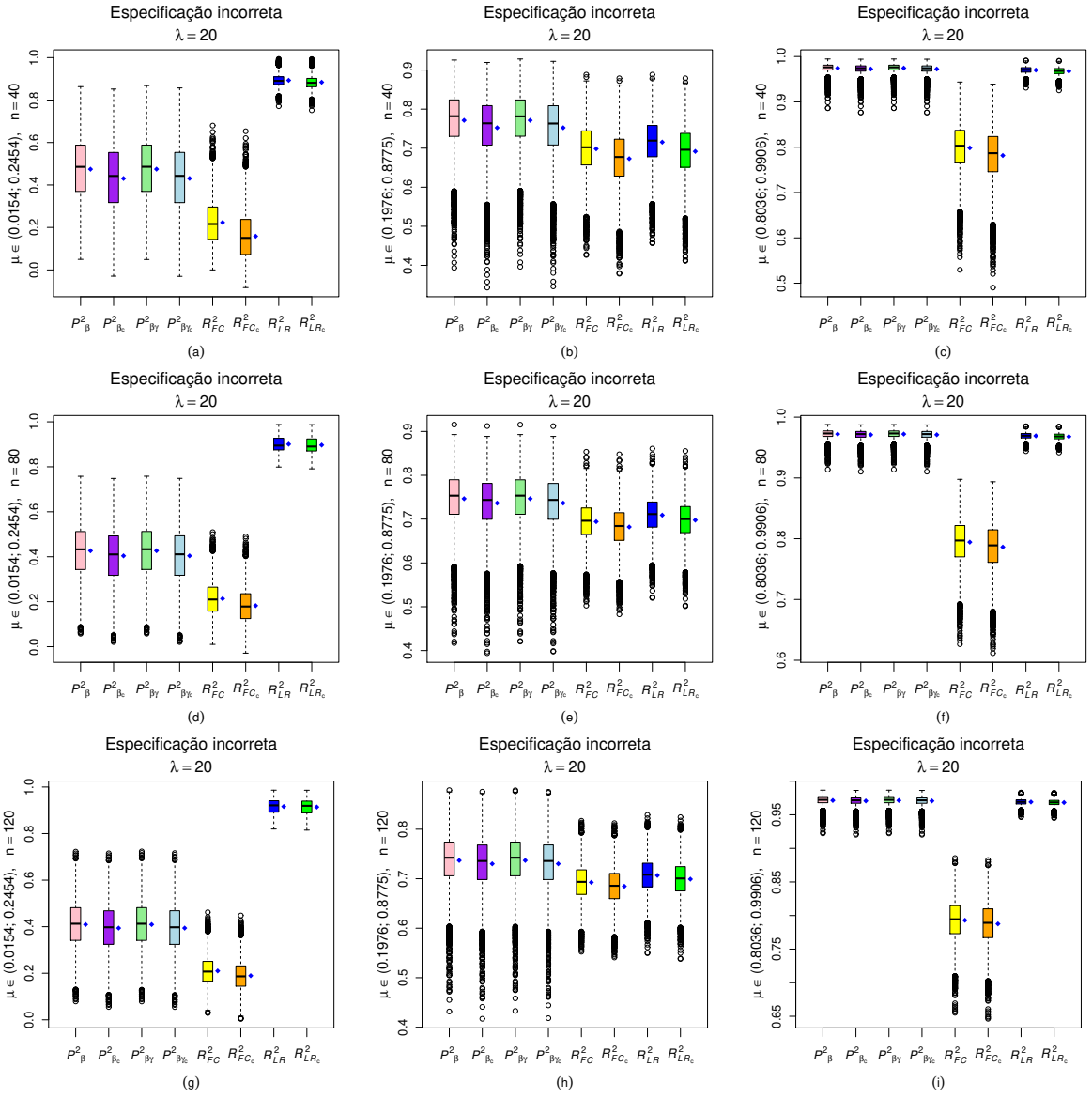

Figura 21 – Boxplots das 10000 réplicas de Monte Carlo das estatísticas P^2_β, $P^2_{\beta\gamma}$, R^2_{FC}, R^2_{LR} e suas versões corrigidas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}$ e $\log(\sigma^2_t) = \gamma_1 + \gamma_2 z_{t2}$. Modelo estimado: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}$, $\lambda = 20$.

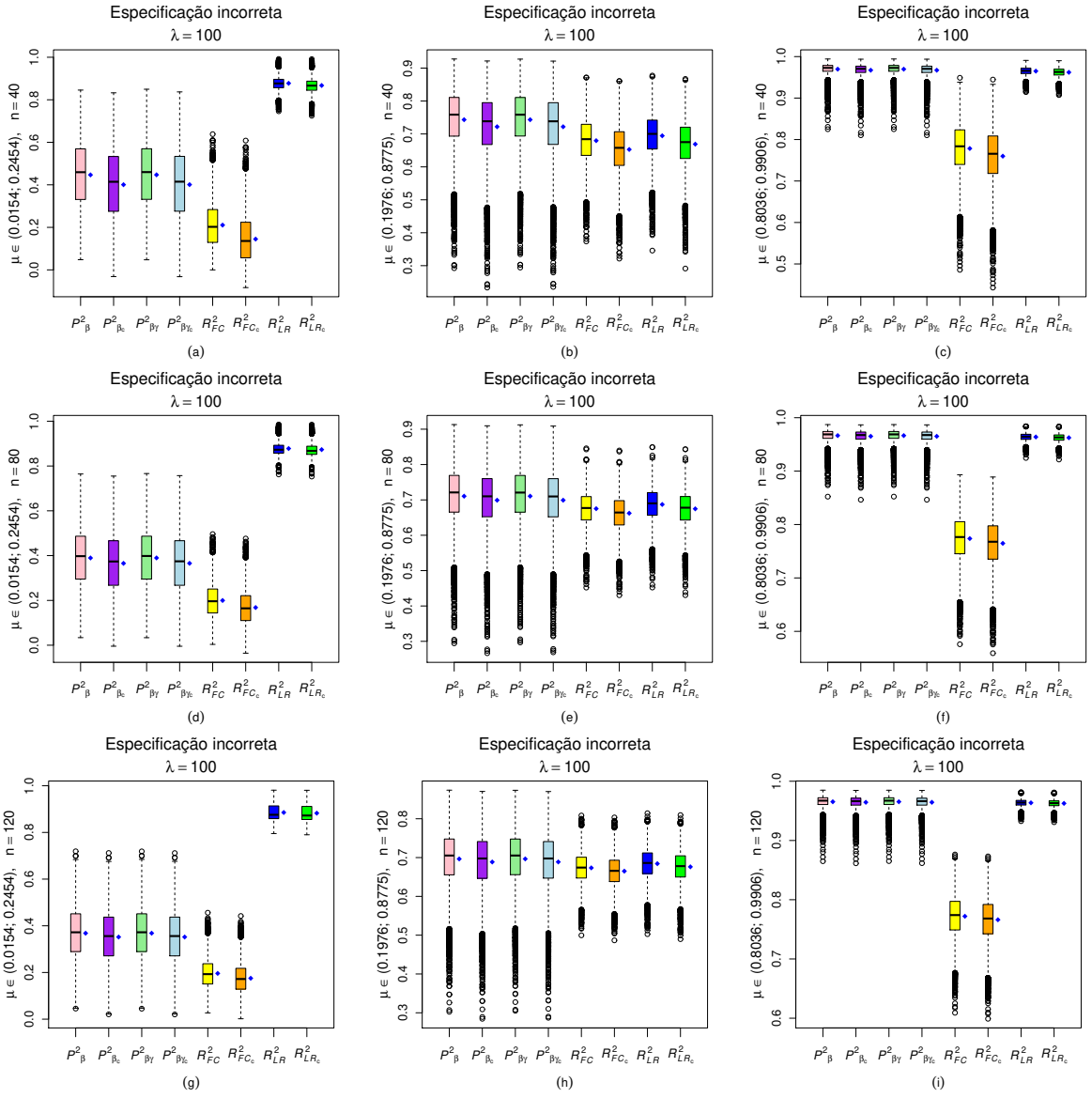

Figura 22 – Boxplots das 10000 réplicas de Monte Carlo das estatísticas P^2_{β}, $P^2_{\beta\gamma}$, R^2_{FC}, R^2_{LR} e suas versões corrigidas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}$ e $\log(\sigma^2_t) = \gamma_1 + \gamma_2 z_{t2}$. Modelo estimado: $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2}$, $\lambda = 100$.

Tabela 24 – Valores médios das estatísticas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1+x_{t2}^{\beta_2}+\beta_3 x_{t3}+\beta_4 x_{t4}$,$x_{ti} \sim U(0,1)$, $i = 2,3,4,5$, $t = 1,\ldots,n$ e σ^2 constante. Especificação incorreta: omissão de covariadas (Cenários 1 e 2).

	Cenários	Cenário 1			Cenário 2			Cenário 3		
	Modelo estimado	$g(\mu_t)=\beta_1+x_{t2}^{\beta_2}$			$g(\mu_t)=\beta_1+x_{t2}^{\beta_2}$ $+\beta_3 x_{t3}$			$g(\mu_t)=\beta_1+x_{t2}^{\beta_2}$ $+\beta_3 x_{t3}+\beta_4 x_{t4}$		
μ		$\mu \in (0.02, 0.15)$; $\beta = (-2.4, 1.4, -1.5, -1.7)^\top$.								
n	σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
40	P^2	0.541	0.385	0.323	0.752	0.600	0.526	0.988	0.909	0.854
	P^2_c	0.516	0.352	0.287	0.731	0.566	0.487	0.987	0.898	0.838
	$P^2_{\beta\gamma}$	0.540	0.385	0.323	0.751	0.600	0.526	0.988	0.909	0.854
	$P^2_{\beta\gamma_c}$	0.516	0.352	0.287	0.731	0.566	0.487	0.987	0.898	0.838
	R^2_{LR}	0.278	0.207	0.176	0.454	0.342	0.293	0.945	0.686	0.576
	$R^2_{LR_c}$	0.239	0.164	0.132	0.408	0.287	0.234	0.939	0.651	0.527
80	P^2	0.522	0.344	0.275	0.742	0.563	0.474	0.987	0.900	0.839
	P^2_c	0.510	0.327	0.256	0.731	0.545	0.453	0.987	0.895	0.830
	$P^2_{\beta\gamma}$	0.522	0.344	0.275	0.741	0.562	0.474	0.987	0.900	0.839
	$P^2_{\beta\gamma_c}$	0.509	0.327	0.256	0.731	0.545	0.453	0.987	0.895	0.830
	R^2_{LR}	0.278	0.203	0.172	0.452	0.333	0.281	0.942	0.675	0.562
	$R^2_{LR_c}$	0.259	0.183	0.150	0.430	0.306	0.253	0.939	0.657	0.538
120	P^2	0.516	0.328	0.254	0.738	0.548	0.452	0.987	0.897	0.832
	P^2_c	0.507	0.317	0.241	0.731	0.536	0.438	0.986	0.893	0.827
	$P^2_{\beta\gamma}$	0.515	0.328	0.254	0.738	0.547	0.452	0.987	0.897	0.832
	$P^2_{\beta\gamma_c}$	0.507	0.317	0.241	0.731	0.536	0.438	0.987	0.893	0.827
	R^2_{LR}	0.278	0.202	0.170	0.451	0.329	0.277	0.941	0.671	0.557
	$R^2_{LR_c}$	0.265	0.188	0.155	0.437	0.312	0.259	0.939	0.659	0.542
μ		$\mu \in (0.20; 0.88)$; $\beta = (-1.7, -1.8, 1.2, -1.3)^\top$.								
n	σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
40	P^2	0.640	0.456	0.404	0.785	0.569	0.507	0.931	0.694	0.623
	P^2_c	0.621	0.427	0.372	0.767	0.533	0.466	0.923	0.659	0.580
	$P^2_{\beta\gamma}$	0.638	0.456	0.404	0.784	0.569	0.507	0.931	0.694	0.623
	$P^2_{\beta\gamma_c}$	0.619	0.427	0.372	0.766	0.533	0.466	0.923	0.659	0.580
	R^2_{LR}	0.570	0.381	0.328	0.726	0.492	0.427	0.909	0.622	0.542
	$R^2_{LR_c}$	0.547	0.348	0.292	0.703	0.450	0.379	0.898	0.579	0.490
80	P^2	0.629	0.422	0.363	0.773	0.533	0.462	0.925	0.663	0.581
	P^2_c	0.619	0.407	0.346	0.764	0.514	0.441	0.921	0.645	0.559
	$P^2_{\beta\gamma}$	0.626	0.422	0.362	0.773	0.533	0.462	0.925	0.663	0.581
	$P^2_{\beta\gamma_c}$	0.617	0.407	0.346	0.764	0.514	0.441	0.921	0.645	0.559
	R^2_{LR}	0.568	0.375	0.322	0.722	0.481	0.414	0.904	0.607	0.524
	$R^2_{LR_c}$	0.556	0.359	0.304	0.711	0.460	0.391	0.899	0.586	0.499
120	P^2	0.624	0.408	0.346	0.769	0.520	0.445	0.923	0.651	0.566
	P^2_c	0.617	0.398	0.335	0.763	0.507	0.431	0.921	0.639	0.551
	$P^2_{\beta\gamma}$	0.621	0.408	0.346	0.769	0.520	0.445	0.923	0.651	0.566
	$P^2_{\beta\gamma_c}$	0.615	0.398	0.335	0.763	0.507	0.431	0.921	0.639	0.551
	R^2_{LR}	0.567	0.373	0.319	0.721	0.477	0.410	0.902	0.601	0.518
	$R^2_{LR_c}$	0.559	0.362	0.307	0.714	0.464	0.395	0.899	0.587	0.501
μ		$\mu \in (0.80.0.99)$; $\beta = (2.1, -1.5, -1.6, -1.2)^\top$.								
n	σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
40	P^2	0.843	0.751	0.694	0.952	0.890	0.846	0.994	0.954	0.924
	P^2_c	0.834	0.737	0.677	0.948	0.881	0.833	0.994	0.948	0.915
	$P^2_{\beta\gamma}$	0.843	0.751	0.694	0.952	0.890	0.846	0.994	0.954	0.924
	$P^2_{\beta\gamma_c}$	0.834	0.737	0.677	0.948	0.881	0.833	0.994	0.948	0.915
	R^2_{LR}	0.468	0.364	0.311	0.719	0.560	0.479	0.960	0.746	0.638
	$R^2_{LR_c}$	0.439	0.330	0.274	0.695	0.524	0.436	0.956	0.717	0.597
80	P^2	0.841	0.732	0.665	0.950	0.881	0.832	0.994	0.950	0.916
	P^2_c	0.837	0.725	0.657	0.948	0.877	0.825	0.994	0.947	0.911
	$P^2_{\beta\gamma}$	0.841	0.732	0.665	0.950	0.881	0.832	0.994	0.950	0.916
	$P^2_{\beta\gamma_c}$	0.837	0.725	0.657	0.948	0.877	0.825	0.994	0.947	0.911
	R^2_{LR}	0.467	0.358	0.303	0.717	0.552	0.469	0.958	0.736	0.625
	$R^2_{LR_c}$	0.453	0.341	0.285	0.706	0.534	0.448	0.956	0.722	0.605
120	P^2	0.839	0.724	0.652	0.949	0.878	0.826	0.994	0.948	0.913
	P^2_c	0.837	0.719	0.646	0.948	0.875	0.822	0.993	0.946	0.910
	$P^2_{\beta\gamma}$	0.839	0.724	0.652	0.949	0.878	0.826	0.994	0.948	0.913
	$P^2_{\beta\gamma_c}$	0.837	0.720	0.646	0.948	0.875	0.822	0.993	0.946	0.910
	R^2_{LR}	0.466	0.356	0.301	0.717	0.550	0.466	0.957	0.732	0.620
	$R^2_{LR_c}$	0.457	0.345	0.289	0.709	0.538	0.452	0.956	0.723	0.607

medida que covariadas são incluídas no modelo estimado os valores das estatísticas P^2 e $P^2_{\beta\gamma}$, assim como de suas versões corrigidas vão aumentando. Esse fato é mais perceptível quando a média da variável resposta encontra-se perto de zero ou de valores mais centrais do intervalo (0,1). É possível perceber também que as medidas de predição e de qualidade de ajuste aumentam quando o parâmetro de dispersão diminui. Da mesma forma, R^2_{LR} e $R^2_{LR_c}$ apontam para o problema de qualidade de ajuste quando as covariadas são omitidas em todos os intervalos de μ.

As Figuras 23 e 24 apresentam os boxplots das 10000 réplicas de P^2_β, $P^2_{\beta\gamma}$, R^2_{FC}, R^2_{LR} e suas versões corrigidas para os Cenários 1 e 3, respectivamente, descritos na Tabela 24, ou seja, a Figura 23 apresenta resultados para o modelo estimado incorretamente (omissão de duas covariadas importantes) enquanto que a Figura 24 apresenta resultados para o modelo estimado corretamente, considerando todas as covariadas que formam o modelo verdadeiro. A comparação entre essas duas figuras revela de forma nítida o aumento dos valores das estatísticas do cenário em que falta duas covariadas importantes para o cenário com todas as covariadas, principalmente quando $\mu \approx 0$ e $\mu \approx 1$. Observamos também que os valores das medidas diminuem quando a dispersão do modelo aumenta. Na Figura 24 observamos que as estatísticas R^2_{FC} e R^2_{LR}, assim como suas versões corrigidas identificam problemas no ajuste do modelo nos três cenários da média. A estatística P^2 pode ser vista como uma medida do viés do modelo enquanto que a estatística R^2 é uma medida da variância do modelo, ou seja, seleciona o modelo que é capaz de explicar melhor a variabilidade da resposta. Portanto, é plausível que estas medidas apresentem valores inferiores quando as distorções entre as variâncias reais e estimadas da variável resposta são grandes. Vale ressaltar que nós também consideramos a estatística P^2 para selecionar o modelo com o melhor ajuste aos dados. De fato, o melhor modelo ajustado deve mostrar valores altos e próximos para as três medidas, assim como suas versões penalizadas.

A Tabela 25 mostra a média das 10000 réplicas das medidas de predição e de qualidade de ajuste sob o cenário de especificação correta. Para isso utilizamos o modelo de regressão simplex com dispersão variável não linear dado por

$$\begin{aligned} \log\left(\frac{\mu_t}{1-\mu_t}\right) &= \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}, \\ \log(\sigma_t^2) &= \gamma_1 + z_{t2}^{\gamma_2}, \end{aligned} \tag{4.10}$$

em que $t = 1, \ldots, n$. Consideramos três tamanhos amostrais $n = (40, 80, 120)$ e $\lambda = (20, 50, 150)$. Como anteriormente, consideramos os seguintes cenários para a média de y: $\mu \in (0.0314; 0.3654)$, $\mu \in (0.1970; 0.8684)$ e $\mu \in (0.7854; 0.9876)$. As realizações das covariadas foram geradas através das seguintes distribuições: $x_{t2} \sim U(-0.5, 0.5)$, $x_{t3} \sim U(0, 1)$, $x_{t4} \sim U(-0.5, 0.5)$ e $z_{t2} \sim U(-0.5, 0.5)$ e foram mantidos fixas para cada réplica. Novamente, percebemos que as medidas de predição indicam uma maior facilidade para predizer quando os dados estão perto dos extremos, ou seja, perto de 0 e 1. O mesmo acontece com as medidas de qualidade. Vimos também um leve aumento das estatísticas à medida que o grau de dispersão não constante vai aumentando, e ainda, não se vê muita diferença nos valores das medidas quando comparamos os tamanhos amostrais considerados.

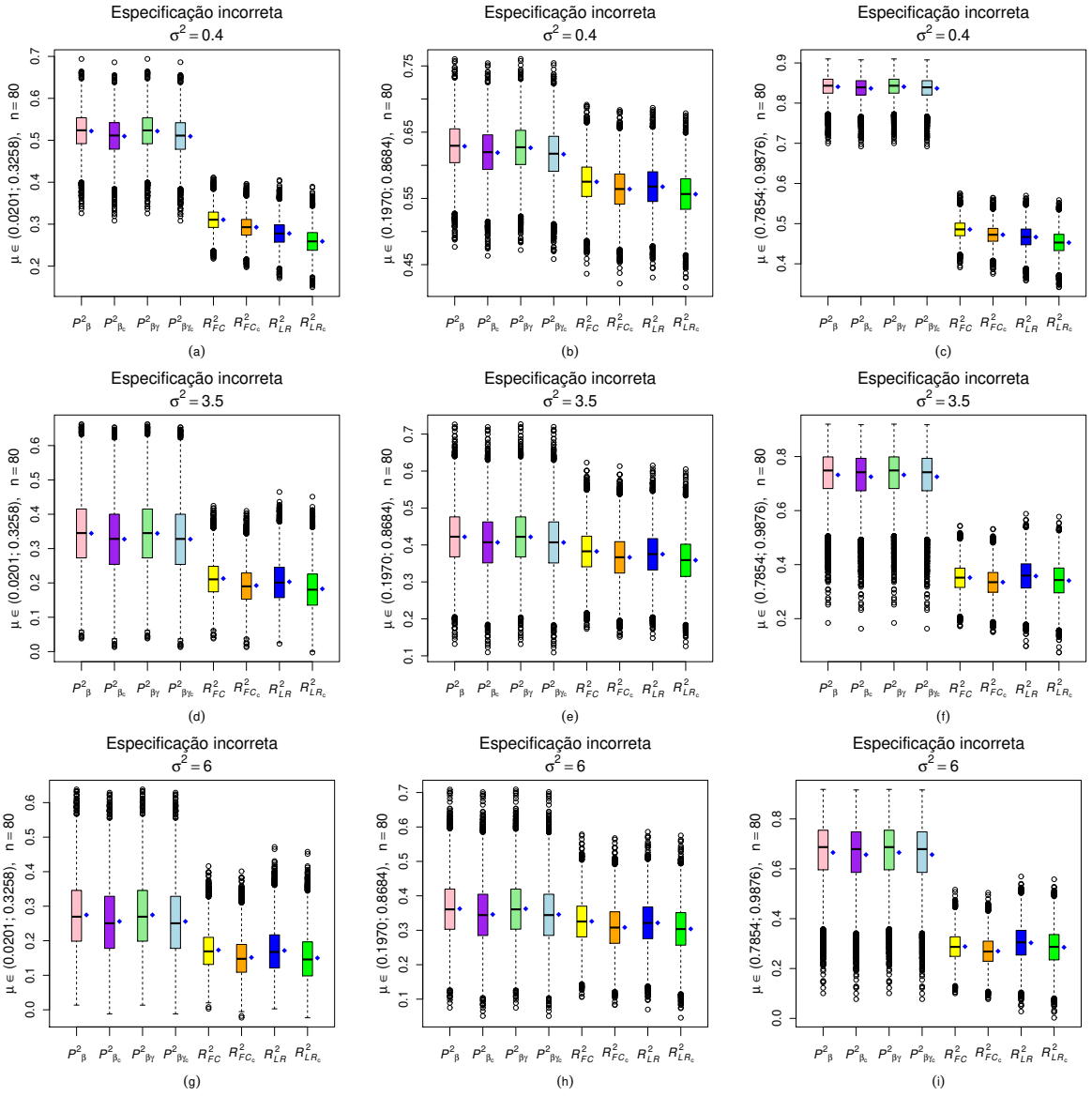

Figura 23 – Boxplots das 10000 réplicas de Monte Carlo das estatísticas P^2_β, $P^2_{\beta\gamma}$, R^2_{FC}, R^2_{LR} e suas versões corrigidas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$. Modelo estimado: $\log(\mu_t/[1-\mu_t]) = \beta_1 + x_{t2}^{\beta_2}$, $n = 80$.

Também avaliamos a distribuição das estatísticas de predição e de qualidade de ajuste considerando diferentes funções de ligação para o submodelo da média da variável resposta. Na Tabela 26 apresentamos a média das 10000 réplicas de Monte Carlo considerando o modelo dado em (4.10) para as seguintes funções de ligação: Logit($g(\mu) = \log(\mu/(1-\mu))$), C-log-log ($g(\mu) = -\log(-\text{-}\log(\mu))$), Log-log ($g(\mu) = \log(-\log(1-\mu))$), Cauchy ($g(\mu) = \tan\{\pi(\mu - 0.5)\}$) e Probit ($g(\mu) = \Phi^{-1}(\mu)$). Consideramos $\lambda = (20, 50, 150)$ e os três cenários para a μ: $\mu \in (0.01; 0.25)$, $\mu \in (0.20; 0.85)$ e $\mu \in (0.85; 0.99)$. O objetivo aqui é avaliar o desempenho das funções de ligação em cada intervalo da média

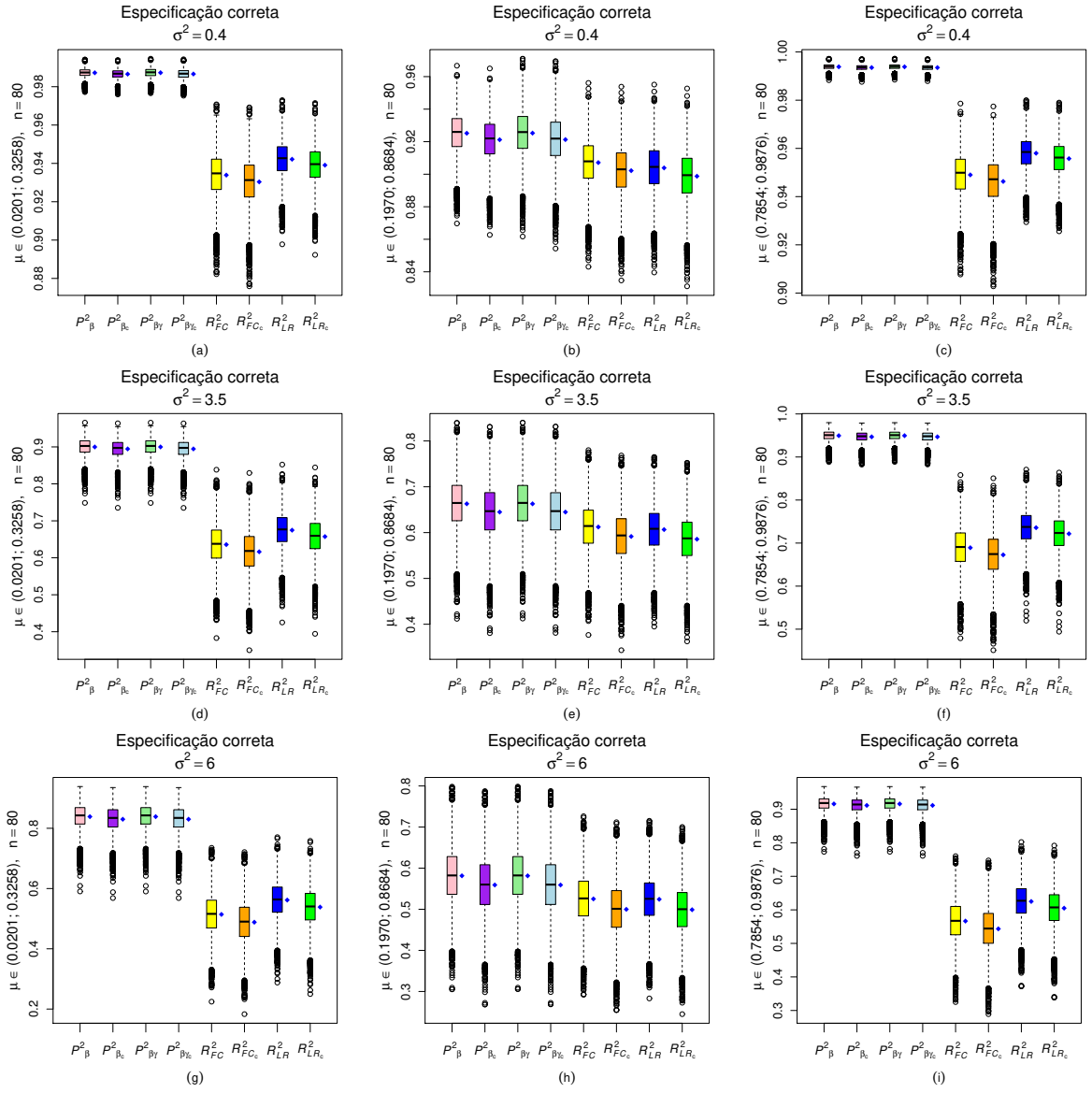

Figura 24 – Boxplots das 10000 réplicas de Monte Carlo das estatísticas P^2_β, $P^2_{\beta\gamma}$, R^2_{FC}, R^2_{LR} e suas versões corrigidas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$. Modelo estimado: $\log(\mu_t/[1-\mu_t]) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$, $n = 80$.

Tabela 25 – Valores médios das estatísticas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\phi_t) = \gamma_1 + z_{t2}^{\gamma_2}$. Modelo corretamente especificado.

	Cenários	$\mu \in (0.02, 0.15)$ $\beta = (-2.4, 1.4, -1.5, -1.7)^\top$			$\mu \in (0.20; 0.88)$ $\beta = (-1.7, -1.8, 1.2, -1.3)^\top$			$\mu \in (0.80.0.99)$ $\beta = (2.1, -1.5, -1.6, -1.2)^\top$		
n	λ	20	50	150	20	50	150	20	50	150
40	P^2	0.9824	0.9833	0.9843	0.8620	0.8576	0.8545	0.9875	0.9878	0.9882
	P_c^2	0.9798	0.9808	0.9820	0.8417	0.8367	0.8331	0.9857	0.9860	0.9865
	$P_{\beta\gamma}^2$	0.9824	0.9833	0.9844	0.8622	0.8578	0.8547	0.9875	0.9878	0.9882
	$P_{\beta\gamma_c}^2$	0.9798	0.9809	0.9821	0.8419	0.8369	0.8333	0.9857	0.9860	0.9865
	R_{LR}^2	0.8651	0.8607	0.8694	0.8171	0.8251	0.8517	0.8950	0.8873	0.8878
	$R_{LR_c}^2$	0.8453	0.8403	0.8502	0.7902	0.7993	0.8299	0.8796	0.8708	0.8713
80	P^2	0.9802	0.9813	0.9824	0.8425	0.8379	0.8346	0.9857	0.9861	0.9867
	P_c^2	0.9788	0.9800	0.9812	0.8319	0.8270	0.8234	0.9848	0.9852	0.9858
	$P_{\beta\gamma}^2$	0.9802	0.9813	0.9824	0.8424	0.8378	0.8345	0.9857	0.9861	0.9867
	$P_{\beta\gamma_c}^2$	0.9789	0.9800	0.9812	0.8318	0.8269	0.8233	0.9848	0.9852	0.9858
	R_{LR}^2	0.8588	0.8564	0.8690	0.8083	0.8198	0.8524	0.8884	0.8808	0.8829
	$R_{LR_c}^2$	0.8493	0.8467	0.8601	0.7954	0.8077	0.8425	0.8808	0.8728	0.8750
120	P^2	0.9794	0.9805	0.9817	0.8354	0.8306	0.8271	0.9851	0.9855	0.9861
	P_c^2	0.9784	0.9796	0.9809	0.8282	0.8232	0.8195	0.9845	0.9849	0.9855
	$P_{\beta\gamma}^2$	0.9794	0.9805	0.9817	0.8353	0.8305	0.8270	0.9851	0.9855	0.9861
	$P_{\beta\gamma_c}^2$	0.9785	0.9796	0.9809	0.8281	0.8231	0.8194	0.9845	0.9849	0.9855
	R_{LR}^2	0.8567	0.8550	0.8696	0.8054	0.8183	0.8532	0.8862	0.8787	0.8815
	$R_{LR_c}^2$	0.8504	0.8486	0.8639	0.7969	0.8103	0.8468	0.8812	0.8734	0.8763

de y, principalmente quando $\mu \in (0.20; 0.85)$, uma vez que a distribuição simplex tem mais dificuldade em modelar dados nesse intervalo. Os resultados da Tabela 26 mostram que quando a média da variável resposta encontra-se próxima de zero o uso das funções de ligações Log-log e Probit levam a modelos com melhor poder preditivo. No entanto, a Probit também se destaca com relação as medidas de qualidade de ajuste. Quando $\mu \in (0.20; 0.85)$, a função Probit leva a resultados melhores considerando todas as medidas. Por fim, quando a média de y está próxima de um, as funções C-Log-log e Probit apresentam melhores resultados. Concluímos que uma função de ligação adequada pode melhorar a robustez da estimação do parâmetros no modelo. As mesmas conclusões tiramos da Tabela 27 que apresenta a distribuição das estatísticas considerando as cinco funções de ligação sob modelo com dispersão constante (4.9).

Tabela 26 – Valores médios das estatísticas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$ e $\log(\sigma_t^2) = \gamma_1 + z_{t2}^{\gamma_2}$. Modelo corretamente especificado.

μ	$\mu \in (0.01; 0.25)$			$\mu \in (0.20; 0.85)$			$\mu \in (0.85; 0.99)$		
λ	20	50	150	20	50	150	20	50	150
Estatísticas	Logit								
P^2	0.980	0.981	0.982	0.843	0.838	0.835	0.986	0.986	0.987
P_c^2	0.979	0.980	0.981	0.832	0.827	0.823	0.985	0.985	0.986
R_{FC}^2	0.762	0.687	0.580	0.725	0.662	0.577	0.813	0.751	0.653
$R_{FC_c}^2$	0.745	0.666	0.551	0.706	0.639	0.548	0.800	0.734	0.630
R_{LR}^2	0.859	0.856	0.869	0.808	0.820	0.852	0.888	0.881	0.883
$R_{LR_c}^2$	0.849	0.847	0.860	0.795	0.808	0.842	0.881	0.873	0.875
	C-Log-Log								
P^2	0.980	0.981	0.982	0.792	0.797	0.802	0.999	0.999	0.999
P_c^2	0.978	0.980	0.981	0.778	0.784	0.789	0.999	0.999	0.999
R_{FC}^2	0.790	0.722	0.619	0.595	0.513	0.409	0.900	0.867	0.809
$R_{FC_c}^2$	0.776	0.704	0.593	0.567	0.480	0.369	0.893	0.858	0.796
R_{LR}^2	0.869	0.866	0.875	0.735	0.755	0.803	0.996	0.995	0.994
$R_{LR_c}^2$	0.860	0.857	0.866	0.717	0.739	0.790	0.996	0.995	0.994
	Log-Log								
P^2	0.999	0.999	0.999	0.853	0.847	0.841	0.985	0.986	0.987
P_c^2	0.999	0.998	0.998	0.843	0.836	0.830	0.984	0.985	0.986
R_{FC}^2	0.845	0.791	0.709	0.757	0.696	0.611	0.816	0.755	0.655
$R_{FC_c}^2$	0.835	0.777	0.689	0.740	0.676	0.585	0.804	0.738	0.632
R_{LR}^2	0.967	0.962	0.958	0.829	0.835	0.858	0.889	0.882	0.883
$R_{LR_c}^2$	0.965	0.960	0.956	0.818	0.824	0.848	0.881	0.874	0.875
	Cauchy								
P^2	0.570	0.624	0.666	0.769	0.764	0.760	0.603	0.660	0.703
P_c^2	0.541	0.598	0.643	0.753	0.748	0.744	0.576	0.637	0.683
R_{FC}^2	0.294	0.178	0.078	0.612	0.466	0.284	0.261	0.163	0.085
$R_{FC_c}^2$	0.247	0.122	0.016	0.585	0.429	0.235	0.211	0.107	0.023
R_{LR}^2	0.622	0.674	0.758	0.823	0.832	0.861	0.570	0.633	0.732
$R_{LR_c}^2$	0.597	0.652	0.742	0.811	0.821	0.852	0.541	0.608	0.714
	Probit								
P^2	0.999	0.999	0.999	0.996	0.996	0.996	0.999	0.999	0.999
P_c^2	0.999	0.999	0.999	0.996	0.996	0.996	0.999	0.999	0.999
R_{FC}^2	0.946	0.924	0.886	0.892	0.852	0.792	0.938	0.917	0.884
$R_{FC_c}^2$	0.943	0.919	0.878	0.884	0.842	0.778	0.934	0.912	0.876
R_{LR}^2	0.995	0.994	0.993	0.957	0.953	0.950	0.993	0.992	0.990
$R_{LR_c}^2$	0.994	0.993	0.992	0.954	0.950	0.947	0.992	0.991	0.989

Tabela 27 – Valores médios das estatísticas. Modelo verdadeiro: $\log(\mu_t/(1-\mu_t)) = \beta_1 + x_{t2}^{\beta_2} + \beta_3 x_{t3} + \beta_4 x_{t4}$,$x_{ti} \sim U(0,1),\ \ i = 2,3,4,5,\ t = 1,\ldots,n$ e σ^2 constante. Modelo corretamente especificado.

μ	$\mu \in (0.01; 0.25)$			$\mu \in (0.20; 0.85)$			$\mu \in (0.85; 0.99)$		
σ^2	0.4	3.5	6.0	0.4	3.5	6.0	0.4	3.5	6.0
Estatísticas	Logit								
P^2	0.987	0.900	0.839	0.925	0.663	0.581	0.994	0.950	0.916
P_c^2	0.987	0.895	0.830	0.921	0.645	0.559	0.994	0.947	0.911
R_{FC}^2	0.934	0.636	0.514	0.907	0.612	0.525	0.949	0.689	0.567
$R_{FC_c}^2$	0.930	0.616	0.488	0.902	0.592	0.500	0.946	0.673	0.543
R_{LR}^2	0.942	0.675	0.562	0.904	0.607	0.524	0.958	0.736	0.625
$R_{LR_c}^2$	0.939	0.657	0.538	0.899	0.586	0.499	0.956	0.722	0.605
	C-Log-Log								
P^2	0.988	0.897	0.833	0.869	0.493	0.399	0.999	0.999	0.999
P_c^2	0.987	0.892	0.824	0.862	0.466	0.366	0.999	0.999	0.999
R_{FC}^2	0.944	0.667	0.548	0.855	0.467	0.370	0.970	0.813	0.735
$R_{FC_c}^2$	0.941	0.649	0.523	0.848	0.438	0.337	0.969	0.803	0.721
R_{LR}^2	0.949	0.697	0.589	0.861	0.489	0.401	0.999	0.989	0.982
$R_{LR_c}^2$	0.946	0.680	0.567	0.853	0.462	0.369	0.999	0.989	0.981
	Log-Log								
P^2	0.999	0.997	0.995	0.938	0.692	0.610	0.994	0.947	0.912
P_c^2	0.999	0.997	0.995	0.935	0.676	0.589	0.994	0.945	0.907
R_{FC}^2	0.962	0.767	0.675	0.920	0.637	0.548	0.954	0.699	0.577
$R_{FC_c}^2$	0.960	0.754	0.658	0.915	0.618	0.524	0.951	0.683	0.555
R_{LR}^2	0.989	0.915	0.865	0.922	0.646	0.563	0.960	0.739	0.631
$R_{LR_c}^2$	0.989	0.910	0.858	0.918	0.627	0.539	0.958	0.726	0.611
	Cauchy								
P^2	0.438	0.096	0.067	0.885	0.544	0.457	0.347	0.074	0.054
P_c^2	0.408	0.048	0.017	0.879	0.519	0.428	0.313	0.025	0.003
R_{FC}^2	0.726	0.217	0.139	0.901	0.557	0.459	0.687	0.185	0.116
$R_{FC_c}^2$	0.711	0.175	0.093	0.895	0.533	0.430	0.670	0.142	0.069
R_{LR}^2	0.765	0.310	0.231	0.916	0.632	0.552	0.709	0.250	0.182
$R_{LR_c}^2$	0.753	0.273	0.190	0.911	0.613	0.528	0.694	0.210	0.138
	Probit								
P^2	0.999	0.999	0.999	0.998	0.985	0.975	0.999	0.999	0.999
P_c^2	0.999	0.999	0.999	0.998	0.984	0.973	0.999	0.999	0.999
R_{FC}^2	0.987	0.907	0.856	0.975	0.845	0.781	0.984	0.882	0.822
$R_{FC_c}^2$	0.987	0.902	0.848	0.974	0.837	0.770	0.983	0.876	0.813
R_{LR}^2	0.998	0.985	0.974	0.985	0.893	0.840	0.998	0.981	0.968
$R_{LR_c}^2$	0.998	0.984	0.973	0.984	0.887	0.831	0.998	0.980	0.966

4.5 Aplicações

4.5.1 Dados de craqueamento catalítico fluido (FCC)

O processo FCC (Fluid Catalytic Cracking) ou ruptura catalítica é usado para converter hidrocarbonetos de alto peso molecular em pequenas moléculas de maior valor comercial, através do contato destes com um catalisador. Ele é usado para converter frações pesadas de petróleo em gasolina, olefinas C3 e C4, GLP e frações que permitem a formulação de combustível diesel. O processo da FCC é muitas vezes considerado o coração de uma refinaria, uma vez que permite adaptar a produção aos produtos com maior demanda e/ou alta rentabilidade (SALAZAR, 2005). O catalisador do processo é formado por partículas finas de 10 a 150 microns, facilmente fluidizável tendo como componente principal o zeólito Y incorporado numa matriz amorfa de aluminossilicato e argila (SALAZAR, 2005).

Sabe-se que cada 1000 ppm de vanádio no catalisador a produção de gasolina diminui em cerca de 2.3%. Além disso, este componente químico é conhecido por participar da destruição do catalisador, reduzindo a superfície ativa, a seletividade e a cristalinidade do zeólito Y especialmente na presença de vapor. O vanádio é depositado na superfície externa das partículas do catalisador no reator da unidade FCC, esses complexos sofrem decomposição parcial e são transferidos para o regenerador onde eles são queimados com a coca e o vanádio é oxidado. Esta reação depende da temperatura do regenerador que deve ser próximo a 720 graus Celsius (SALAZAR, 2005).

O objetivo aqui é modelar a variável resposta porcentagem de cristalinidade do zeólito Y (y) através das covariadas vapor d'água (x_1), temperatura do processo (x_2) e concentração de vanádio (x_3). Inicialmente realizaremos uma análise descritiva da variável resposta e das covariadas. Na Figura 25 apresentamos os boxplots com algumas estatísticas descritivas da variável resposta e das covariadas vapor d'água e concentração de vanádio. Observamos que a variável resposta está concentrada no extremo superior do intervalo unitário padrão, sendo o mínimo igual a 0.6430, um ponto aberrante. Temos ainda uma variável levemente assimétrica com média e mediana bem próximas. Quanto as covariadas, observamos que vapor d'água possui alta dispersão com mínimo e máximo dados por 0 e 55.8, respectivamente. No caso da covariada vanádio vemos que o mínimo é

igual ao primeiro quantil assumindo valor zero, ou seja, um quarto dos valores assumidos pela covariada é igual a zero. A covariada temperatura assume dois valores, 700º ou 760º, e cada um desses valores representa 50% dos dados.

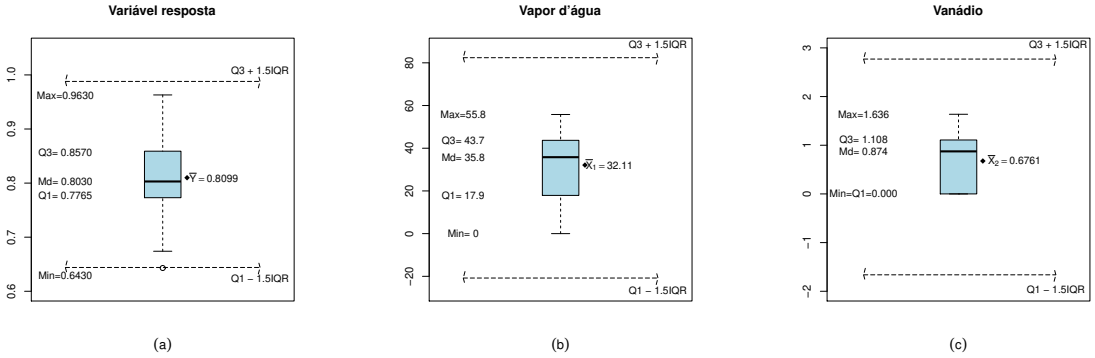

Figura 25 – Boxplots da variável resposta e das variáveis vapor d'água e vanádio.

ESPINHEIRA & SILVA (2018) analisaram esses dados e propuseram um modelo de regressão beta não linear utilizando a função de ligação Logit. Aqui consideramos a mesma função dos preditores para o modelo de regressão simplex, ou seja,

$$g(\mu_t) = \beta_1 + \beta_2 x_{t2}/(x_{t2} + \beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}} \quad \text{e} \quad \log(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t4}^2 \tag{4.11}$$

$t = 1, \ldots, 28$. No entanto, consideramos três funções de ligação para o submodelo da média, a saber: Logit, Probit e C-log-log. As funções de ligação Log-log e Cauchy não apresentaram todos os parâmetros significativos utilizando o modelo dado em 4.11. Inicialmente, faremos uma análise residual e de influência para o modelo 4.11 considerando as funções Logit, Probit e C-log-log. Para tanto, construímos os gráficos da distância de Cook versus o índice das observações e os gráficos normais de probabilidade com envelopes simulados considerando os resíduos ponderado (ESPINHEIRA & SILVA, 2018) e combinado. Os gráficos da distância de Cook versus o índice das observações identificaram os observações {10,12,24} como possivelmente influentes para as três funções de ligação. No entanto, observamos que a função Logit destaca a observação 24 com mais ênfase, uma vez que a distância de Cook para essa observação apresenta valor próximo de quatro e a complemento Log-log, por exemplo, apresenta valor abaixo de três. Para investigar melhor o comportamento dessas observações, retiramos esses pontos e reestimamos o modelo para verificar o impacto delas nas estimativas dos parâmetros. A Tabela 28 apresenta as

mudanças relativas nas estimativas dos parâmetros, nas estimativas dos erros-padrão e os correspondentes p-valores com a exclusão das observações para as funções de ligação Logit, Probit e C-log-log. Podemos observar que com a exclusão individual e conjunta das observações {10,12,24} não houve nenhuma mudança expressiva nos p-valores, no entanto, houve mudanças significativas nas estimativas dos parâmetros. Mesmo com a retirada desses pontos o processo de estimação por máxima verossimilhança do modelo simplex continua robusto, uma vez que mesmo ocorrendo mudanças nas estimativas dos parâmetros o modelo continua sendo estatisticamente significativo. Observamos ainda que ao comparar as três funções de ligação, o processo de estimação da C-log-log mostrou ser menos sensível com a presença de pontos influentes que o das demais. Os gráficos normais de probabilidade com envelopes simulados dos resíduos ponderado e combinado (Figura 26) apontam para a função C-log-log como a mais adequada para esses dados.

Na Tabela 29 apresentamos os valores dos critérios para seleção de modelo proposto nesta tese. Tais medidas apontam mais uma vez para a função de ligação C-log-log como a mais adequada para os esses dados, tanto as medidas de predição como as de qualidade de ajuste foram superiores às demais. No entanto, é preciso ressaltar que são as medidas de predição que apontam mais enfaticamente este modelo como o mais adequado. Com base na Tabela 29 percebe-se como as medidas de predição associadas ao ajuste Logit são inferiores ao do modelo C-log-log, fato este relacionado a influência do caso 24. Já as quantidades R^2 estão muito próximas. As medidas R^2 estão relacionadas com a variabilidade do modelo, enquanto as medidas P^2 podem ser consideradas uma medida de viés do modelo. Assim, é interessante a escolha de um modelo em que tanto o viés quanto a variabilidade estejam bem controlados. Como foi visto nas simulações de Monte Carlo, quando a média da variável resposta está próxima do extremo superior do intervalo unitário padrão, as funções de ligação que apresentaram melhores resultados foram a C-log-log e a Probit. Essa aplicação evidencia esses resultados e mostra como os critérios P^2, R^2_{FC} e R^2_{LR} são úteis para selecionar modelos simplex.

Tabela 28 – Estimativas dos parâmetros, erros-padrão (e.p.), mudanças relativas das estimativas e dos erros-padrão para casos excluídos e respectivos p-valores. Modelo: $g(\mu_t) = \beta_1 + \beta_2 x_{t2}/(x_{t2} + \beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t4}^2$. Dados FCC.

Logit						
	Parâmetro	β_1	β_2	β_3	β_4	β_5
Dados	est.	2.375	-0.107	-27.803	-0.290	-0.751
completos	e.p.	0.149	0.042	4.021	0.107	0.142
	p-v	0.000	0.012	0.000	0.007	0.000
obs.{10}	mud. est.	3.605	-9.286	2.414	-7.951	10.820
Deletado	mud. e.p.	-7.202	2.803	-2.200	-3.193	-6.001
	p-v	0.000	0.026	0.000	0.010	0.000
obs.{24}	mud. est.	0.917	-12.883	-1.833	38.854	-4.631
Deletado	mud. e.p.	-3.029	-15.563	7.781	-5.424	-1.136
	p-v	0.000	0.009	0.000	0.000	0.000
obs.{12}	mud. est.	2.054	26.043	-2.981	19.149	-5.650
Deletado	mud. e.p.	-0.455	6.515	-11.996	5.832	-0.606
	p-v	0.000	0.003	0.000	0.002	0.000
obs.{10,24}	mud. est.	4.047	-16.913	-0.197	28.087	5.459
Deletado	mud. e.p.	-8.225	-12.752	6.189	-5.873	-5.822
	p-v	0.000	0.016	0.000	0.000	0.000
obs.{10,12}	mud. est.	5.030	11.921	-0.198	6.823	5.420
Deletado	mud. e.p.	-7.027	12.334	-9.670	2.137	-4.746
	p-v	0.000	0.012	0.000	0.005	0.000
obs.{10,12,24}	mud. est.	5.774	13.589	-4.417	47.165	-2.356
Deletado	mud. e.p.	-8.526	-10.994	-13.162	-4.859	-5.123
	p-v	0.000	0.001	0.000	0.000	0.000
Probit						
Dados	est.	1.399	-0.061	-27.843	-0.182	-0.420
completos	e.p.	0.077	0.023	3.879	0.063	0.075
	p-v	0.000	0.007	0.000	0.004	0.000
obs.{10}	mud. est.	2.715	-9.555	2.525	-9.234	9.001
Deletado	mud. e.p.	-8.381	3.183	-2.520	-3.579	-6.273
	p-v	0.000	0.019	0.000	0.006	0.000
obs.{12}	mud. est.	1.961	25.811	-2.680	19.825	-6.963
Deletado	mud. e.p.	-1.568	3.895	-14.474	4.567	0.355
	p-v	0.000	0.001	0.000	0.001	0.000
obs.{10,24}	mud. est.	3.045	-16.512	-0.175	19.521	3.885
Deletado	mud. e.p.	-9.263	-11.848	6.454	-6.708	-7.159
	p-v	0.000	0.011	0.000	0.000	0.000
obs.{10,12}	mud. est.	4.025	11.017	0.154	5.840	2.636
Deletado	mud. e.p.	-8.786	10.278	-11.753	1.000	-3.743
	p-v	0.000	0.007	0.000	0.002	0.000
obs.{10,12,24}	mud. est.	4.582	13.307	-3.988	38.584	-5.157
Deletado	mud. e.p.	-9.734	-10.916	-14.478	-6.201	-5.300
	p-v	0.000	0.001	0.000	0.000	0.000
C-log-log						
Dados	est.	0.952	-0.053	-27.909	-0.180	-0.355
completos	e.p.	0.058	0.018	3.680	0.057	0.059
	p-v	0.000	0.003	0.000	0.002	0.000
obs.{10}	mud. est.	2.229	-9.580	2.541	-11.181	6.653
Deletado	mud. e.p.	-9.081	3.700	-2.630	-4.004	-5.687
	p-v	0.000	0.011	0.000	0.003	0.000
obs.{10,24}	mud. est.	2.357	-15.808	-0.202	8.701	1.610
Deletado	mud. e.p.	-9.394	-10.407	7.171	-7.278	-8.010
	p-v	0.000	0.006	0.000	0.000	0.000
obs.{10,12}	mud. est.	3.602	9.273	0.727	2.410	-1.099
Deletado	mud. e.p.	-10.084	8.680	-13.877	-0.850	-0.794
	p-v	0.000	0.003	0.000	0.001	0.000
obs.{10,12,24}	mud. est.	3.924	11.707	-3.127	26.362	-9.103
Deletado	mud. e.p.	-10.010	-9.172	-15.046	-7.255	-3.930
	p-v	0.000	0.000	0.000	0.000	0.000

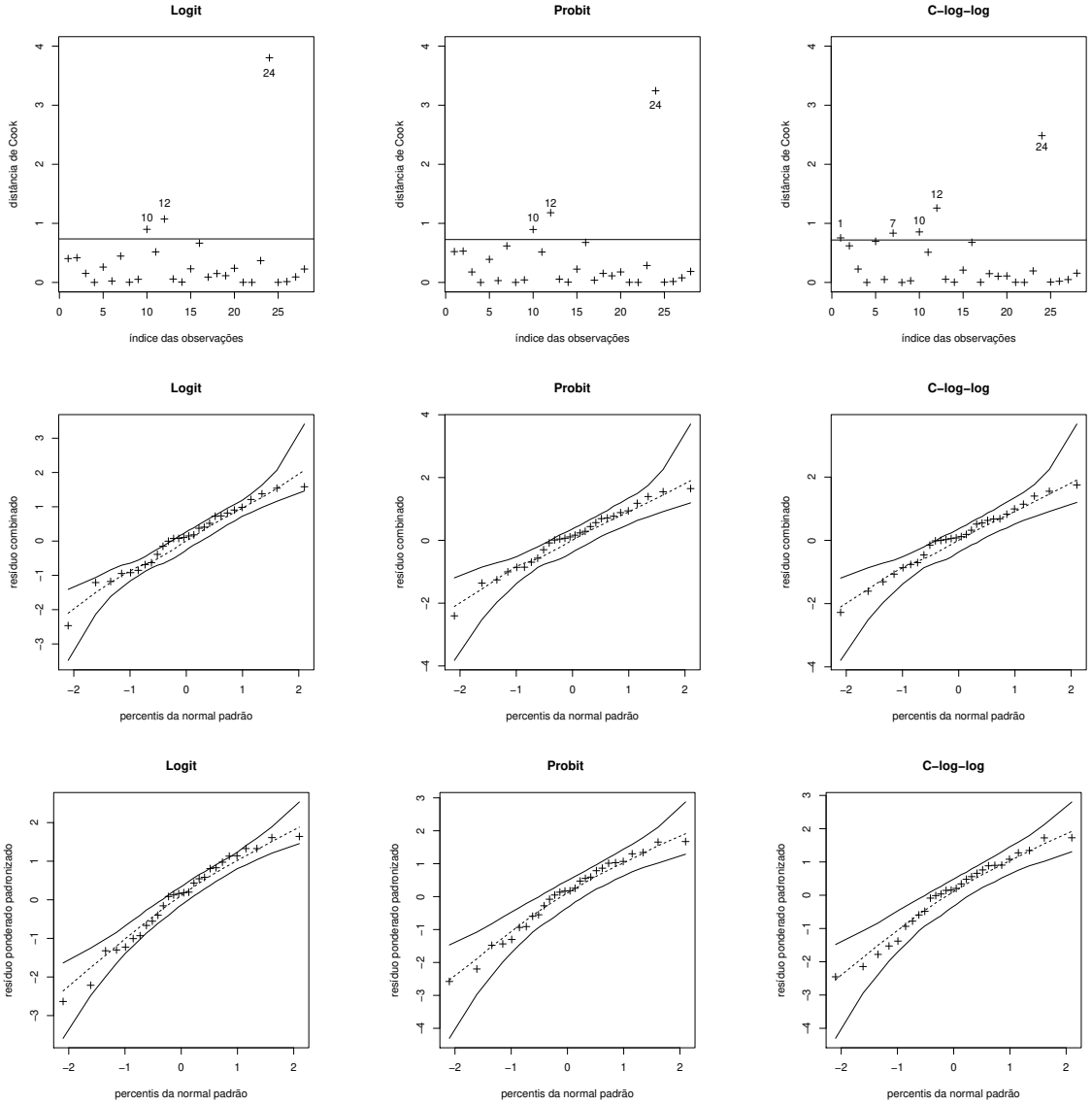

Figura 26 – Gráficos das distâncias de Cook e dos resíduos. Modelo: $\mathrm{g}(\mu_t) = \beta_1 + \beta_2 x_{t2}/(x_{t2} + \beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t4}^2$. Dados FCC.

Tabela 29 – Valores dos critérios de predição e de qualidade de ajuste. Modelo: $\mathrm{g}(\mu_t) = \beta_1 + \beta_2 x_{t2}/(x_{t2} + \beta_3) + \beta_4 x_{t3} + \beta_5\sqrt{x_{t4}}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 z_{t4}^2$. Dados FCC.

Estatísticas	Logit	Probit	C-log-log
P^2_{β}	0.62757	0.72742	0.82339
$P^2_{\beta_c}$	0.52116	0.64954	0.77293
$P^2_{\beta\gamma}$	0.72371	0.78602	0.84704
$P^2_{\beta\gamma_c}$	0.64477	0.72489	0.80333
R^2_{FC}	0.66028	0.67335	0.68125
$R^2_{FC_c}$	0.56322	0.58002	0.59018
R^2_{LR}	0.71885	0.72276	0.72879
$R^2_{LR_c}$	0.63853	0.64355	0.65130

4.5.2 Dados dos Transplantes Autólogos de Células Tronco do Sangue Periférico

Os dados dessa aplicação foram utilizados por ALLAN et al. (2002) e YANG et a. (2005) e trata-se de um estudo sobre transplantes autólogos de células tronco do sangue periférico (CTSP). Os CTSP têm sido bastante utilizados para a recuperação hematológica após a terapia mieloablativa para diversas doenças hematológicas malignas. Esse estudo foi desenvolvido por Edmonton Hematopoietic Stem Cell Lab, no Cross Cancer Institute, em Alberta, Canadá. Um grupo de 239 pacientes realizaram o transplante autólogo de CTSP após receber doses de quimioterapia mieloablativos entre os anos de 2003 e 2008. Através dos dados obtidos modelamos a taxa de recuperação das células CD34+ (y) na qual consideramos as seguintes covariadas:

- idade ajustada (x_2) - idades menores que 40 anos são definidas como idades da linha de base, idades acima de 40 anos são ajustadas subtraindo o valor 40 da idade dos pacientes;
- quimio (x_3) - corresponde a uma variável dummy, indicando se o paciente recebe uma quimioterapia no protocolo de 1 dia (quimio=0) ou de 3 dias (quimio=1);
- idade dos pacientes (x_4).

Com o objetivo de conhecer melhor a variável resposta e as covariáveis, realizamos uma análise descritiva dos dados. A Figura 27 apresenta o boxplot das observações da variável resposta taxa de recuperação das células CD34+ e o histograma das observações da variável da resposta com a curva estimada supondo distribuição simplex. É possível notar que os dados estão concentrados no extremo superior do intervalo unitário padrão e que a variável possui muitos outliers inferiores a aproximadamente 0.5. Já o histograma evidencia que a massa de probabilidade da variável resposta encontra-se no intervalo (0.7,0.9) e que a curva estimada com base na distribuição simplex se ajusta de forma razoável à variável resposta.

Além disso, construímos os boxplots das observações das covariáveis candidatas idade ajustada e idade (Figura 28). Notamos uma dispersão menor para a covariada idade

ajustada quando comparada com idade e a covariada quimio apresenta 54% dos dados iguais a 0 e 46% iguais a 1.

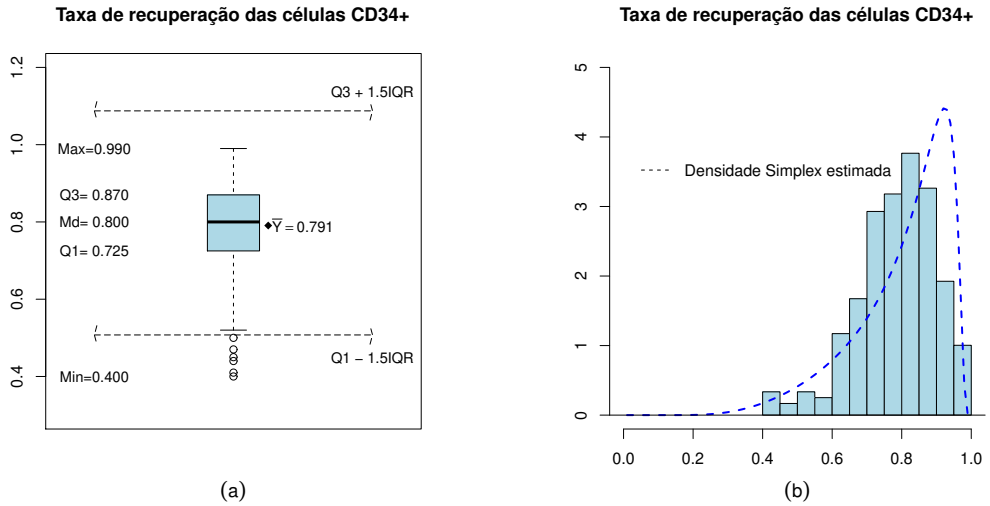

Figura 27 – Boxplot e histograma das observações da variável resposta.

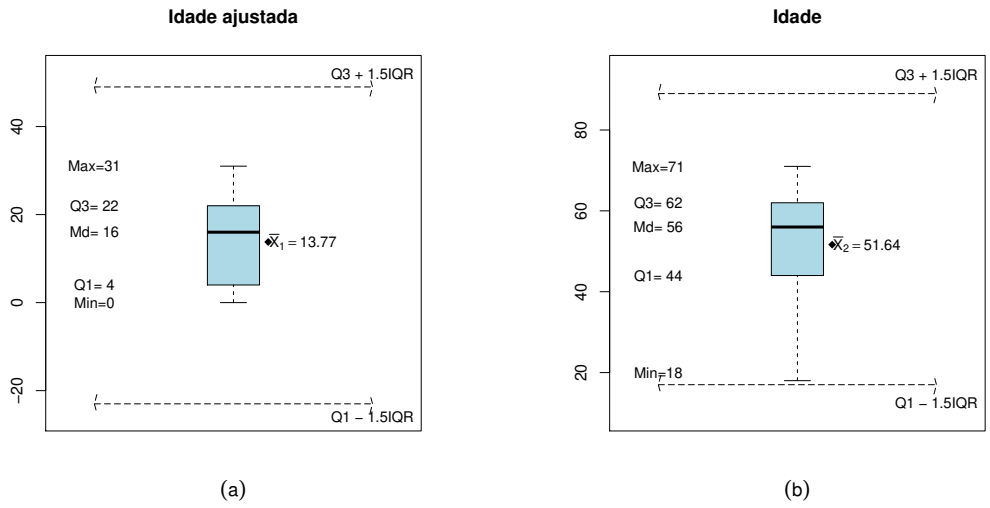

Figura 28 – Boxplots das observações das covariadas candidatas ao modelo.

Inicialmente consideramos a modelagem com dispersão constante. Após avaliar diferentes preditores lineares e diferentes funções de ligação escolhemos o seguinte modelo com a função de ligação Logit:

$$\text{M1} : \log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}, \quad t = 1, \ldots, 239.$$

As Figuras 29 e 30 mostram os gráficos dos resíduos ponderado e combinado contra diferentes elementos do modelo e os gráficos normais com probabilidades com envelopes simulados do modelo M1, respectivamente. Os gráficos dos resíduos evidenciam a ocorrência de muitos pontos aberrantes e os gráficos de envelopes indicam para uma má qualidade do ajuste do modelo aos dados.

A Tabela 30 mostra as estimativas dos parâmetros, erros-padrões e p-valores do modelo M1. Estamos em um cenário em que a estimativa de σ^2 não é tão alta, ou seja, $\hat{\sigma^2} = 6.4$, e em que os valores da variável resposta estão mais próximas do limite superior do intervalo (0,1), tanto que sua mediana é igual a 0.8. Nestes cenários nas simulações tanto as medidas de predição quanto as de qualidade de ajuste são altas. No entanto, a Tabela 31 contendo as P^2's e R^2's para o modelo M1 mostra que são medidas extremamente pequenas, o que só verificamos nos cenários com omissão de covariáveis importantes. Assim, a princípio temos indícios que estão faltando covariáveis importantes para a explicação da média da variável resposta.

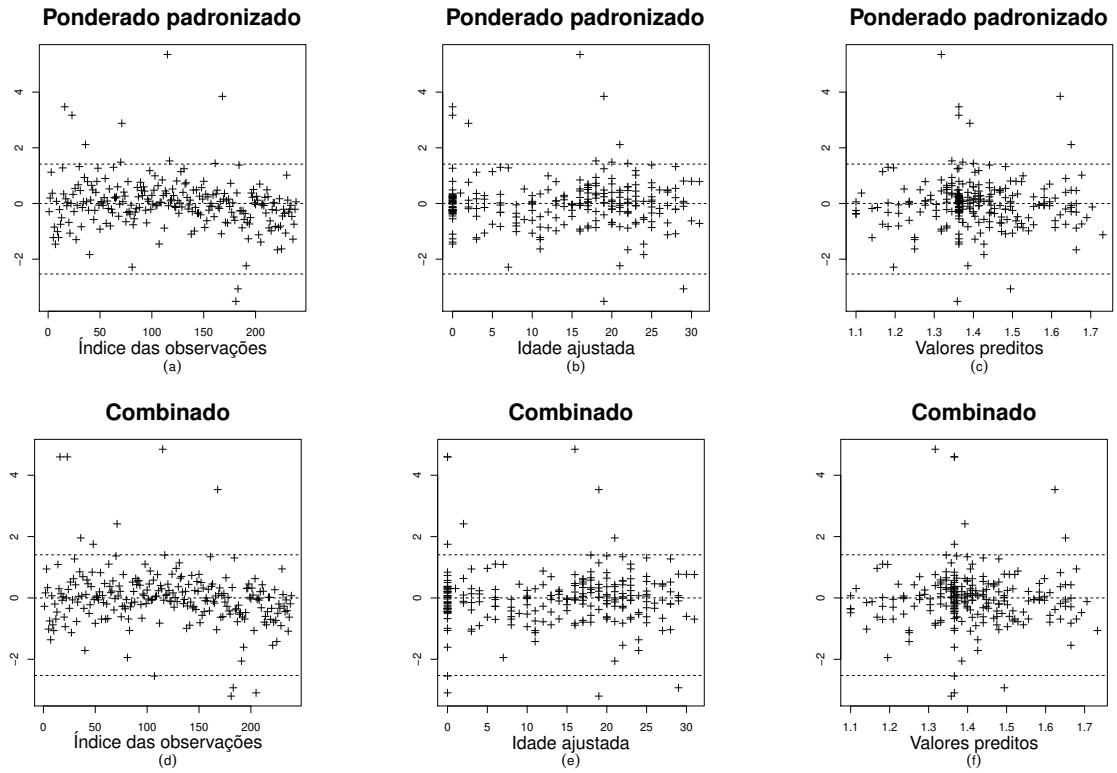

Figura 29 – Gráficos de resíduos do modelo $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}$, $t = 1, \ldots, 239$, σ^2 constante. Dados dos transplantes.

Consideramos ainda a modelagem da dispersão. Com base no modelo M1 propomos

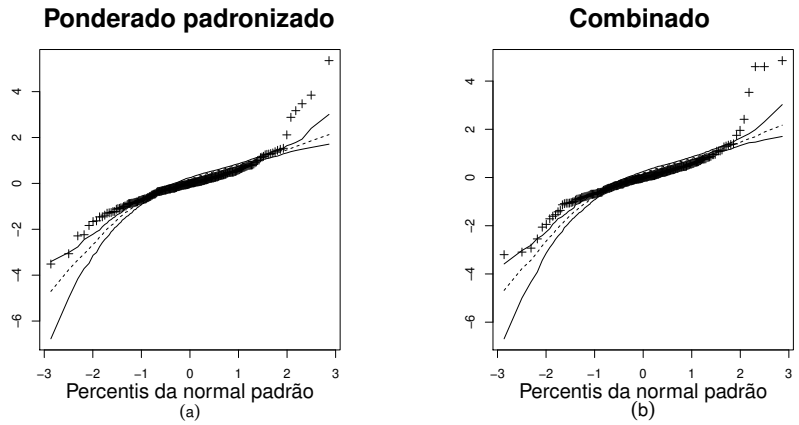

Figura 30 – Gráficos normais de probabilidades com envelopes simulados para o modelo $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}$, $t = 1, \ldots, 239$, σ^2 constante. Dados dos transplantes.

Tabela 30 – Estimativas dos parâmetros, erros-padrões e p-valores do modelo $\log(\mu_t/(1-\mu_t)) = \beta_1+\beta_2 x_{t2}+\beta_3 x_{t2}$, $t = 1, \ldots, 239$, σ^2 constante. Dados dos transplantes.

	Estimativa	Erro-padrão	p-valor
β_1	1.1002	0.1401	0.0000
β_2	0.0136	0.0065	0.0365
β_3	0.2661	0.1245	0.0325
σ^2	6.3966	0.0915	0.0000

Tabela 31 – Medidas de predição e qualidade de ajuste para o modelo $\log(\mu_t/(1-\mu_t)) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}$, $t = 1, \ldots, 239$, σ^2 constante. Dados dos transplantes.

Medidas	Valores
P^2_{β}	0.044340
$P^2_{\beta_c}$	0.032140
$P^2_{\beta\gamma}$	0.044134
$P^2_{\beta\gamma_c}$	0.031931
R^2_{FC}	0.026124
$R^2_{FC_c}$	0.013691
R^2_{RV}	0.024712
$R^2_{RV_c}$	0.012262

o seguinte modelo com dispersão variável

$$\text{M2} : g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}, \quad \text{e} \quad \log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t4}, \quad t = 1, \ldots, 239.$$

Para $g(\mu_t)$ consideramos diversas funções de ligações e avaliamos as estatísticas de predição e qualidade do ajuste. Na Tabela 32 apresentamos as estatísticas P^2's e R^2's para o modelo M2 com diferentes funções de ligação. De maneira geral observamos que a função de Probit se mostra melhor analisando como um todo. Assim, apresentaremos os gráficos dos resíduos para o seguinte modelo

$$\text{M3} : \Phi^{-1}(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}, \quad \text{e} \quad \log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t4}, \quad t = 1, \ldots, 239.$$

Tabela 32 – Medidas de predição e qualidade de ajuste para o modelo $g(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t4}$, $t = 1, \ldots, 239$. Dados dos transplantes. Diferentes funções de ligação.

Estatísticas	Cauchy	C-log-log	Log-log	Logit	Probit
P^2_{β}	0.036640	0.089213	0.081162	0.081410	0.096697
$P^2_{\beta_c}$	0.015967	0.069668	0.061445	0.061698	0.077313
$P^2_{\beta\gamma}$	0.036741	0.089662	0.081539	0.081789	0.097168
$P^2_{\beta\gamma_c}$	0.016070	0.070126	0.061830	0.062085	0.077794
R^2_{FC}	0.004317	0.035101	0.023835	0.026293	0.030207
$R^2_{FC_c}$	−0.017050	0.014395	0.0028868	0.0053983	0.0093955
R^2_{LR}	0.063576	0.061745	0.062406	0.062241	0.062042
$R^2_{LR_c}$	0.043481	0.041611	0.042286	0.042117	0.041914

As Figuras 31 e 32 apresentam os gráficos dos resíduos contra vários elementos do modelo M3 e os gráficos dos resíduos com envelopes simulados. Mesmo com a modelagem do parâmetro da dispersão, os valores das estatísticas de predição e de qualidade do ajuste continuam bastante baixos e os gráficos de resíduos evidenciam que de fato o modelo postulado está muito distante do verdadeiro modelo que ajusta bem esses dados. Como discutido anteriormente, pode ser falta de alguma covariada importante para o modelo. Além disso, as estatísticas de predição são afetadas com a presença de pontos influentes como visto anteriormente.

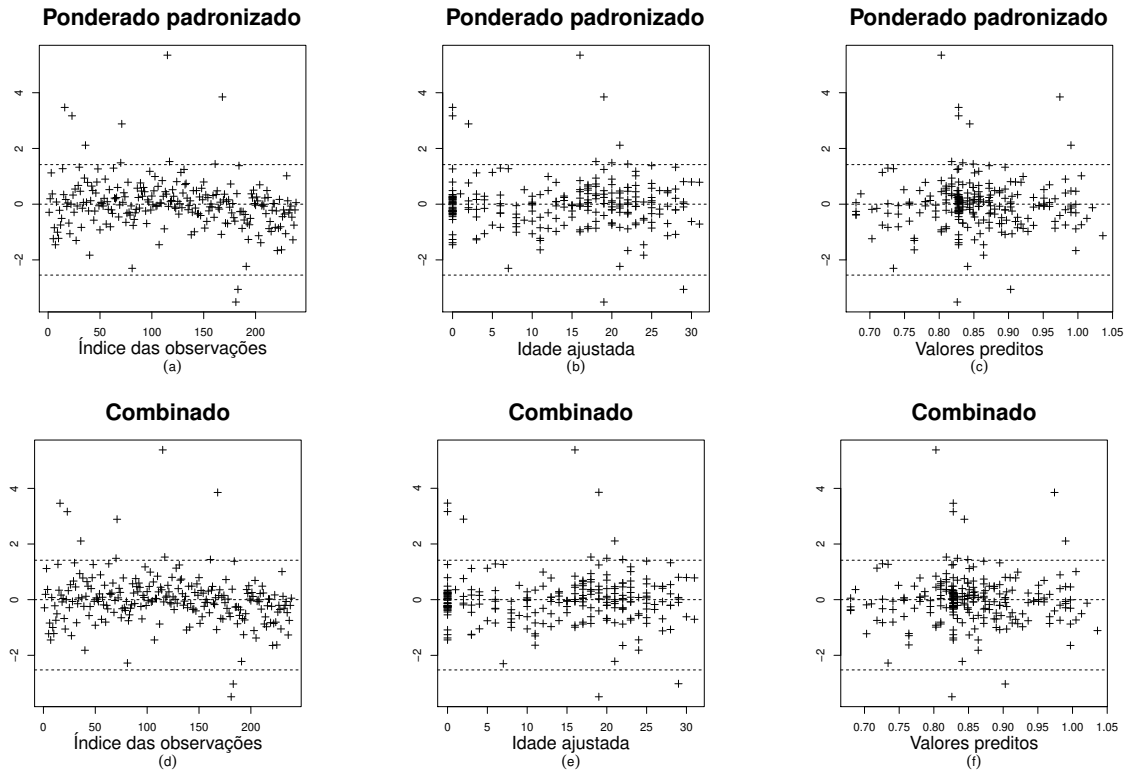

Figura 31 – Gráficos de resíduos do modelo $\Phi^{-1}(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t4}$, $t = 1, \ldots, 239$. Dados dos transplantes. Diferentes funções de ligação.

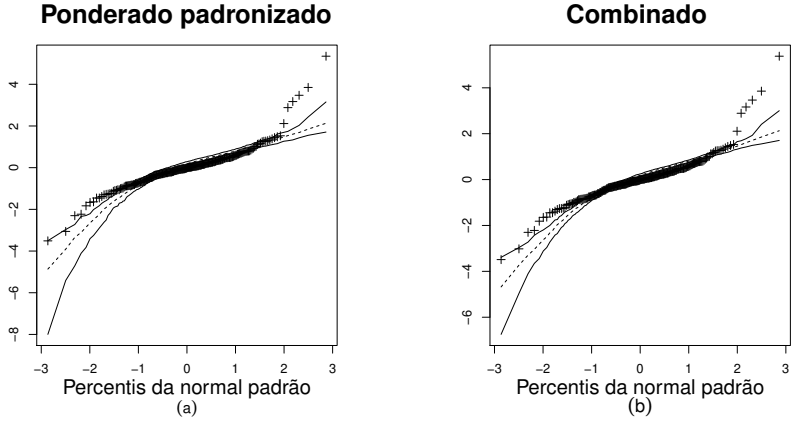

Figura 32 – Gráficos normais de probabilidades com envelopes simulados para o modelo $\Phi^{-1}(\mu_t) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t2}$ e $\log(\sigma_t^2) = \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t4}$, $t = 1, \ldots, 239$. Dados dos transplantes. Diferentes funções de ligação.

4.6 Conclusão

Neste capítulo propomos as estatísticas $PRESS$ e P^2 como critérios de seleção de modelos para a classe de modelos de regressão simplex não linear. Apresentamos uma versão baseada no resíduo ponderado (ESPINHEIRA & SILVA, 2018) e outra no resíduo combinado. Além disso, construímos a distância de Cook para o modelo simplex não linear e mostramos sua relação com a estatística $PRESS$ (COOK & WEISBERG, 1982).

Apresentamos resultados de simulação de Monte Carlo para avaliar a distribuição dos critérios de predição P^2 e $P^2_{\beta\gamma}$ considerando diversos cenários incluindo especificação correta e incorreta do modelo para diferentes tamanhos amostrais, valores para o parâmetro de dispersão e grau de dispersão não constante. Apresentamos também resultados de simulação dos critérios de qualidade de ajuste R^2_{FC} e R^2_{LR}. Em geral, vimos que os valores das estatísticas são maiores quando a média da variável resposta encontra-se perto dos extremos do intervalo unitário padrão comparados com os mesmos cenários em que μ assume valores centrais. Como já foi visto no capítulo anterior, o processo de estimação por máxima verossimilhança do modelo simplex em alguns casos é menos sensível que o do modelo de regressão beta, por exemplo, quando os dados estão próximos de zero ou de um. Avaliamos ainda os critérios de predição e qualidade de ajuste considerando funções de ligação diferentes para o submodelo da média da variável resposta.

Por fim, mostramos duas aplicações a dados reais, as quais mostraram a eficiência dos critérios propostos.

5 CONSIDERAÇÕES FINAIS

Nesta tese apresentamos algumas ferramentas úteis para análise de diagnóstico na classe de modelos de regressão simplex não linear. Propomos um novo resíduo para essa classe de modelos, aqui denotado por resíduo combinado. Esse resíduo é baseado no processo iterativo escore de Fisher para a estimação de β e γ, os vetores de parâmetros que indexam os submodelos da média e do parâmetro de dispersão. Além disso, introduzimos a estatística $PRESS$ e o coeficiente de predição P^2 baseado no resíduo ponderado e no resíduo combinado.

Apresentamos resultados de simulações de Monte Carlo para avaliar o novo resíduo e a performance das medidas de predição. Além disso avaliamos o comportamento das distribuições de algumas medidas de qualidade ajuste tais como R^2_{LR}, R^2_{FC} e suas versões corrigidas. Consideramos diversos cenários incluindo especificação correta e incorreta do modelo para diferentes tamanhos amostrais, valores para o parâmetro de dispersão e valores para o grau de dispersão não contante.

Verificamos que a distribuição empírica do resíduo combinado possui uma leve assimetria e por isso os limites -2 e 2 para detectar pontos aberrantes podem não ser adequados. Sugerimos a utilização dos 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 (BAYER & CRIBARI-NETO, 2017).

Vimos ainda através das medidas de predição que seus valores são maiores quando a média da variável resposta encontra-se perto dos extremos do intervalo unitário padrão comparados com os mesmos cenários em que μ assume valores centrais do intervalo $(0, 1)$. Observamos que em alguns casos o processo de estimação por máxima verossimilhança do modelo simplex é menos sensível que o do modelo de regressão beta, por exemplo, quando os dados estão próximos de zero ou de um. Avaliamos ainda os critérios de predição e qualidade de ajuste considerando funções de ligação diferentes para o submodelo da média da variável resposta.

Por fim, as aplicações a dados reais mostraram a eficiência do resíduo combinado

em detectar pontos aberrantes e influentes nos dados. Em muitos exemplos tal resíduo detectou pontos que os demais não conseguiram detectar. Além disso, vimos como as medidas de predição e de qualidade de ajuste auxiliam na escolha de modelos.

5.1 Trabalhos futuros

Serão objetivos de nossas próximas pesquisas:

- Desenvolver métodos de diagnóstico para o modelo simplex longitudinal.
- Desenvolver intervalos de predição bootstrap para o modelo de regressão simplex.

REFERÊNCIAS

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory*, p. 267–281, 1973.

ALLAN, D. S. et al. Number of viable cd34+ cells reinfused predicts engraftment in autologous hematopoietic stem cell transplantation. *Bone Marrow Transplantation*, v. 29, p. 967–972, 2002.

ALLEN, D. M. The prediction sum of squares as a criterion for selecting predictor variables. *University of Kentucky, Department of Statistics Technical Report 23*, 1971.

ALLEN, D. M. The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, v. 16, p. 125–127, 1974.

ATKINSON, A. C. Plots, transformations and regression: An introduction to graphical methods of diagnostic regression analysis. *New York: Oxford University Press*, 1985.

BARNDORFF-NIELSEN, O. E.; JøRGENSEN, B. Some parametric models on the simplex. *Journal of Multivariate Analysis*, v. 39, p. 106–116, 1991.

BATES, D. M.; WATTS, D. G. Nonlinear regression analysis and its applications. *New York, John Wiley*, 1988.

BAYER, F. M.; CRIBARI-NETO, F. Model selection criteria in beta regression with varying dispersion. *Communications in Statistics, Simulation and Computation*, v. 46, p. 720–746, 2017.

BRITO, D. *Modelo de regressão beta não linear com erros nas variáveis.* Dissertação (Msc Thesis) — Universidade Federal de Pernambuco, 2018. Disponível em: <http://www3.ufpe.br/ppge/images/teses/tese_35.pdf>.

BROWNLEE, K. *Statistical Theory and Methodology in Science and Engineering*. [S.l.]: Wiley New York, 1965. ISBN 978-0898747485.

CHAPLIN, C. *Chaplin - Vida e Pensamentos*. [S.l.]: Martin Claret, 1997.

CLOTILDE, F. *Teste de Diagnóstico Baseado em Influência Local Aplicado ao Modelo de Regressão Simplex*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2016.

COOK, R. D. Detection of influential observation in linear regression. *Technometrics*, v. 19(1), p. 15–18, 1977.

COOK, R. D.; WEISBERG, S. Residuals and influence in regression. *Chapman and Hall*, 1982.

CORDEIRO, G. M.; DEMETRIO, C. G. B. *Modelos Lineares Generalizados e Extensões*. [S.l.]: Departamento de Estatística e Informática, UFRPE e ESALQ/USP, 2008.

COX, D. R.; REID, N. Parameter orthogonality and aproximate conditional inference. *Jornal of the Royal Statistical Society Ser. B*, v. 49, p. 1–39, 1987.

DALGAARD, P. *Introductory Statistics with R*. [S.l.]: Springer, 2002. ISBN 978-0-387-79053-4.

DOORNIK, J. A. *Ox: An object-oriented matrix programming language* . [S.l.]: Timberlake Consulants LTD, 2001.

DOORNIK, J. A. *An Introduction to OxMetrics 4* . [S.l.]: Timberlake Consulants Press, 2006. ISBN 0-9552127-0-7.

DOORNIK, J. A. *Object-Oriented Matrix Programming using Ox*. [S.l.]: Timberlake Consulants Press, 2013. ISBN 978-0-9571708-1-0.

DOORNIK, J. A.; OOMS, M. *Introduction to Ox 4*. [S.l.]: Timberlake Consulants Press, 2006. ISBN 0-9552127-0-7.

DRAPER, N. R.; SMITH, H. *Applied Regression Analysis*. [S.l.]: Springer Verlag, 1981. ISBN 9780471170822.

ESPINHEIRA, P. L.; FERRARI, S. L. P.; CRIBARI-NETO, F. Influence diagnostics in beta regression. *Computaional Statistics and Data Analysis*, v. 52, p. 4417–4431, 2008.

ESPINHEIRA, P. L.; FERRARI, S. L. P.; CRIBARI-NETO, F. On beta regression residuals. *Journal of Applied Statistics*, v. 35, n. 4, p. 407–419, 2008.

ESPINHEIRA, P. L.; SANTOS, E. G.; CRIBARI-NETO, F. On nonlinear beta regression residuals. *Biometrical Journal*, v. 59, 2017.

ESPINHEIRA, P. L.; SILVA, A. O. Nonlinear simplex regression models. *arXiv:1805.10843 [math.ST]*, 2018.

ESPINHEIRA, P. L. et al. Model selection criteria on beta regression for machine learning. *Machine Learning and Knowledge Extraction*, v. 1(1), p. 427–449, 2019.

FERRARI, S. L. P.; CRIBARI-NETO, F. Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, v. 31, n. 7, p. 799–815, 2004.

GARSIDE, M. J. The best subset in multiple regression analysis. *Journal of the Royal Statistical Society C*, v. 14, p. 196–200, 1965.

JøRGENSEN, B. *The theory of dispersion models*. [S.l.]: Chapman and Hall, 1997. ISBN 9780412997112.

KIESCHNICK, R.; MCCULLOUGH, B. Regression analysis of variates observed on (0,1): percentages, proportions, and fractions. *Statist. Model*, v. 3, p. 193–213, 2003.

KNUTH, D. E. *The TEXbook*. [S.l.]: Addisson-Wesley, 1986.

LAMPORT, L. *LATEX A Document Preparation System*. [S.l.]: Addison-Wesley, 1994. ISBN 0-201-52983-1 978-0387954578.

LEMONTE, A. J.; BAZAN, J. L. New class of johnson sb distributions and its associated regression model for rates and proportions. *Biometrical Journal*, v. 58, p. 727–746, 2016.

LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. *Biometrics*, v. 54(2), p. 570–582, 1998.

LIU, H. et al. Press model selection in repeated measures data. *Computational Statistics & Data Analysis*, v. 30, p. 169–184, 1999.

MALLOWS, C. L. Some comments on cp. *Technometrics*, v. 15, p. 661–675–213, 1973.

MCCULLAGH, P. *Tensor Methods in Statistics*. [S.l.]: Chapman and Hall, 1987. ISBN 9780412274800.

MCCULLAGH, P.; NELDER, J. A. *Generalized Linear Models*. 2. ed. London: Chapman and Hall, 1989.

MEDIAVILLA, F. A. M.; LANDRUM, F.; SHAH, V. A. A comparison of the coefficient of predictive power, the coefficient of determination and aic for linear regression. In: . [S.l.: s.n.], 2008. p. 1261–1266.

MITNIK, P. A.; BAEK, S. The kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. *Statistical Papers*, v. 54(1), p. 177–192, 2013.

MIYASHIRO, E. S. *Modelos de regressão beta e simplex para ánalise de proporções*. Dissertação (Msc Thesis) — University of Sao Paulo, 2008. Disponível em: <{http://www.teses.usp.br/teses/disponiveis/45/45133/tde-06112009-224039/pt-br.ph}.>

MOUSA, A. M.; EL-SHEIKH, A. A.; ABDEL-FATTAH, M. A. A gamma regression for bounded continuous variables. *Advances and Applications in Statistics*, v. 49(4), p. 305–326, 2016.

NELDER, J. A.; WEDDERBURN, R. W. M. Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, v. 135, p. 370–384, 1972.

NETER, J. et al. *Applied linear statistical models*. [S.l.]: Chicago: Irwin., 1996. ISBN 0-07-238688-6.

PAOLINO, P. Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Anal*, v. 9, p. 325–346, 2001.

PAULA, G. A. Influence and residuals in restricted generalized linear models. *Journal of Statistical Computation and Simulation*, v. 51, p. 315–352, 1995.

PAULA, G. A. *Modelos de Regressão com apoio computacional*. [S.l.]: Instituto de Matemática e Estatística-USP, 2013.

PREGIBON, D. Logistic regression diagnostics. *Ann. Statist.*, The Institute of Mathematical Statistics, v. 9, n. 4, p. 705–724, 07 1981.

QIU, Z.; SONG, P. X. K.; TAN, M. Simplex mixed-effects models for longitudinal proportional data. *Scandinavian Journal of Statistics*, v. 35, p. 577–596, 2008.

ROCHA, A. V.; SIMAS, A. B. Influence diagnostics in a general class of beta regression models. *Test*, v. 20, n. 1, p. 95–119, 2011.

SALAZAR, S. M. G. *Contribuición al estudio de la reacción de decomposición de la Zeolita Y em presencia de vapor de agua y vanadio*. Dissertação (Mestrado) — Universidad Nacional de Colombia, 2005.

SCHWARZ, G. Estimating the dimension of a model. *Ann. Statist.*, v. 6, p. 461–464, 1978.

SEN, P. K.; SINGER, J. M. Large sample methods in statistics. *Chapman and Hall*, 1993.

SIMAS, A. B.; BARRETO-SOUZA, W.; ROCHA, A. V. Improved estimators for a general class of beta regression models. *Computational Statistics & Data Analysis*, v. 54, n. 2, p. 348–366, 2010.

SMITHSON, M.; VERKULIEN, J. A better lemon squeezer? Maximum-likelihood regression with beta-disbruted dependent variables. *Psichological Methods*, v. 11, n. 1, p. 54–71, 2006.

SONG, P. X. K. Dispersion models in regression analysis. *Pak. J. Statist.*, v. 25, p. 529–551, 2009.

SONG, P. X. K.; QIU, Z.; TAN, M. Modelling heterogeneous dispersion in marginal models for longitudinal proportional data. *Biometrical Journal*, v. 46, p. 540–553, 2004.

SONG, P. X. K.; TAN, M. Marginal models for longitudinal continuous proportional data. *Biometrics*, v. 56, p. 496–502, 2000.

SOUZA, F. A. M. de; PAULA, G. A. Deviance residuals for an angular response. *Australian and New Zealand Journal of Statistics*, v. 44, p. 345–356, 2002.

STONE, M. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society Ser. B*, v. 36, p. 111–147, 1974.

VANABLES, W. B.; RIPLEY, B. D. *Modern applied statistics with S.* [S.l.]: New York: Springer, 2002. ISBN 978-0387954578.

VASCONCELLOS, K. L. P.; CRIBARI-NETO, F. Improved maximum likelihood estimation in a new class of beta regression models. *Brazilian J. Probab. Statist*, v. 19, p. 13–31, 2005.

VENABLES, W. N.; SMITH, D. M.; TEAM, R. C. *An Introduction to R. Notes on R: Programming Environment for Data Analysis and Graphics. Version 3.5.2.* [S.l.: s.n.], 2018.

WOLD, H. Soft modelling: the basic design and some extensions. *Communications in Statistics, Simulation and Computation Systems under indirect observation (Part II)*, p. 1–53, 1982.

YANG, H. et al. Association of post-thaw viable cd34+ cells and cfu-gm with time to hematopoietic engraftment. *Bone Marrow Transplantation*, v. 35(9), p. 1–7, 2005.

ZERBINATTI, L. F. M. *Predicao de ator de simutaneidade atarves de modelos de reggressao para proporcoes continuas.* Dissertação (Mestrado) — University of Sao Paulo, 2008.

ZHANG, P.; QIU, Z.; SHI, C. simplexreg: An r package for regression analysis of proportional data using the simplex distribution. *Journal of Statistical Software, Articles*, v. 71, n. 11, p. 1–21, 2016.