UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA

PÓS-GRADUAÇÃO EM ESTATÍSTICA

REGRESSÃO SIMPLEX NÃO LINEAR: INFERÊNCIA E DIAGNÓSTICO

ALISSON DE OLIVEIRA SILVA

Brasil
2015

UNIVERSIDADE FEDERAL DE PERNAMBUCO

CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA

PÓS-GRADUAÇÃO EM ESTATÍSTICA

REGRESSÃO SIMPLEX NÃO LINEAR: INFERÊNCIA E DIAGNÓSTICO

ALISSON DE OLIVEIRA SILVA

Orientadora: Profª. Drª. PATRÍCIA LEONE ESPINHEIRA OSPINA

Área de Concentração: Estatística Aplicada

Dissertação submetida como requerimento parcial para obtenção do grau de Mestre em Estatística pela Universidade Federal de Pernambuco

Brasil
2015

Catalogação na fonte
Bibliotecária Jane Souto Maior, CRB4-571

S586r Silva, Alisson de Oliveira
Regressão simplex não linear: inferência e diagnóstico / Alisson de Oliveira Silva. – Recife: O Autor, 2015.
85 f.: il., fig., tab.

Orientador: Patrícia Leone Espinheira Ospina.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Estatística, 2015.
Inclui referências e apêndices.

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

310 CDD (23. ed.) UFPE- MEI 2015-37

ALISSON DE OLIVEIRA SILVA

REGRESSÃO SIMPLEX NÃO LINEAR: INFERÊNCIA E DIAGNÓSTICO

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

Aprovada em: 25 de fevereiro de 2015.

BANCA EXAMINADORA

Prof.ª Doutora Patrícia Leone Espinheira Ospina - Presidente
Universidade Federal de Pernambuco

Prof. Doutor Raydonal Ospina Martínez (Examinador Interno)
Universidade Federal de Pernambuco

Prof.ª Doutora Tatiene Correia de Souza (Examinadora Externa)
Universidade Federal da Paraíba

Este trabalho é carinhosamente dedicado
aos meus pais, Maria e Antonio.

AGRADECIMENTOS

Agradeço primeiramente a DEUS, pelo dom da vida e por me abençoar em todos os momentos.

Aos meus pais, Maria e Antonio, pelo amor incondicional, carinho, dedicação e por tudo que me ensinaram ao longo desses anos. Sem o apoio deles seria impossível à concretização de mais uma etapa da minha vida.

À minha irmã, Annanda, pelo carinho, incentivo e pelos momentos de descontração.

À minha namorada, Camila, pelo carinho, amor, dedicação, companheirismo, paciência inesgotável e por está sempre comigo ao longo da minha caminhada.

À Professora Patrícia, pela orientação excepcional, confiança, paciência e dedicação.

Aos Professores do Departamento de Estatística da UFPB, pelo incentivo.

Aos amigos do mestrado William, Dora, Raquel, Luana, Wanessa e Jeniffer por fazerem parte desta conquista.

Aos meus amigos Antonio, Aldair e Frederico, pelo incentivo e pelos momentos de descontração.

Aos Professores Raydonal e Tatiene, antecipadamente, pelas contribuições para melhoria deste trabalho.

A todos os Professores do Programa de Pós-graduação em Estatística da UFPE, por contribuírem para minha formação acadêmica.

À Valéria Bittencourt, pela paciência, dedicação e carinho por todos os alunos da Pós-graduação em Estatística.

Aos demais colegas do Programa de Pós-graduação em Estatística da UFPE.

Ao CNPq, pelo apoio financeiro.

É muito melhor lançar-se em busca de conquistas grandiosas,
mesmo expondo-se ao fracasso,
do que alinhar-se com os pobres de espírito,
que nem gozam muito nem sofrem muito,
porque vivem numa penumbra cinzenta,
onde não conhecem nem vitória, nem derrota.

Theodore Roosevelt.

RESUMO

Em diversas situações práticas, sejam experimentais ou observacionais, há o interesse em investigar como um conjunto de variáveis se relaciona com percentagens, taxas ou razões. Dados restritos ao intervalo contínuo (0,1), em geral, exibem assimetria e possuem um padrão específico de heteroscedasticidade, tornando o modelo normal linear inadequado. Nesse sentido, uma classe de modelos de regressão beta foi proposta por Ferrari e Cribari–Neto (2004), em que a média da variável resposta está relacionada com um preditor linear, através de uma função de ligação, e o preditor linear envolve covariáveis e parâmetros desconhecidos. Uma alternativa competitiva à distribuição beta é o modelo simplex proposto por Barndorff–Nielsen e Jorgensen (1991). A distribuição simplex faz parte dos modelos de dispersão definidos por Jorgensen (1997) que estendem os modelos lineares generalizados. Nesta dissertação, propomos uma extensão do modelo de regressão simplex (Miyashiro, 2008), em que tanto a média da variável resposta quanto o parâmetro de precisão estão relacionados às covariáveis por meio de preditores não lineares. Apresentamos expressões em forma fechada para o vetor escore, matriz de informação de Fisher e sua inversa. Desenvolvemos técnicas de diagnósticos para o modelo de regressão simplex não linear baseadas no método de influência local (Cook, 1986), sob cinco esquemas de perturbação. Além disso, propomos um resíduo para o modelo através do processo iterativo escore de Fisher, e obtemos uma expressão matricial para a alavanca generalizada com base na definição geral apresentada por Wei et al. (1998). Aplicações a dados reais e dados simulados são apresentadas para ilustrar a teoria desenvolvida.

Palavras-chaves: Alavanca generalizada. Distribuição simplex. Influência local. Modelo de regressão simplex não linear. Resíduo.

ABSTRACT

In many practical situations, whether experimental or observational, there is interest in investigating how a set of variables relates to percentages, rates or fractions. Restricted data to continuous interval (0.1), in general, exhibit asymmetry and have a specific pattern of heteroscedasticity, making the normal linear model inappropriate. In this sense, a class of beta regression models was proposed by Ferrari and Cribari–Neto (2004), in which the mean response is related to a linear predictor through a link function, and the linear predictor includes regressors and regression parameters. A useful alternative to the beta distribution is the simplex model proposed by Barndorff–Nielsen and Jorgensen (1991). Simplex distribution is part of the dispersion models defined by Jorgensen (1997) that extend generalized linear models. In this paper, we propose an extension of the simplex regression model (Miyashiro, 2008), in which both the mean response as the precision are related to covariates via non-linear predictors. We provide closed-form expressions for the score function, for Fisher's information matrix and its inverse. Some diagnostic measures are introduced. We propose a new residual obtained using Fisher's scoring iterative scheme for the estimation of the parameters that index the regression non-linear predictor to the mean response and numerically evaluate its behaviour. We also derive the appropriate matrices for assessing local influence on the parameter estimates under diferent perturbation schemes and provide closed-form to generalized leverage matrix proposed by Wei, Hu and Fung (1998). Finally, applications using real and simulated data are presented and discussed.

Key-words: Leverage generalized. Simplex distribution. Local influence. Nonlinear simplex regression model. Residual.

LISTA DE ILUSTRAÇÕES

LISTA DE TABELAS

SUMÁRIO

CAPÍTULO 1

INTRODUÇÃO

Em diversas situações práticas, sejam experimentais ou observacionais, há o interesse em investigar a relação entre uma variável aleatória de interesse (y), denominada de variável resposta, e um conjunto de p variavéis explicativas ($x_1, x_2, \ldots, x_p$), através de um modelo de regressão. Durante muito tempo, os modelos normais lineares foram utilizados na tentativa de descrever a maioria dos fenômenos aleatórios. Mesmo quando a suposição de normalidade não era razoável, algum tipo de transformação era sugerida a fim de obter a normalidade desejada (Paula, 2013). No entanto, o modelo normal linear torna-se inadequado quando o objetivo recai em investigar como um conjunto de variáveis se relaciona com percentagens, taxas ou razões. Dados restritos ao intervalo (0,1), em geral, exibem assimetria e possuem um padrão específico de heterocedasticidade, violando as suposições básicas do modelo normal linear. Kieschnick e McCullough (2003) identificam sete modelos de regressão que têm sido utilizados para modelar a distribuição de dados de proporções no intervalo (0,1). Dentre eles destacam-se os modelos baseados nas distribuições beta e simplex e o modelo de quasi-verossimilhança.

O modelo de regressão beta proposto por Ferrari e Cribari–Neto (2004) assume que a variável resposta y segue distribuição beta e que a média de y está relacionada a um preditor linear por meio de uma função de ligação. O preditor linear envolve covariáveis e parâmetros de regressão desconhecidos. Além disso, a reparametrização utilizada pelos autores considera um parâmetro de precisão ϕ supostamente constante.

Diversos autores têm-se dedicado a diferentes aspectos de inferência e diagnóstico no modelo de regressão beta. Por exemplo, Smithson e Verkuilen (2006) propuseram um modelo de regressão beta em que tanto a média da variável resposta quanto o parâmetro de precisão são modelados através de covariáveis. Espinheira et al. (2008a, b) apresentam resíduos e medidas de influência local para análise de diagnóstico no modelo de regressão beta. Ospina et al. (2006) obtêm melhoramentos em estimação pontual e intervalar para o modelo de regressão beta com dispersão constante. Simas et al. (2010) apresentam uma generalização do modelo de

regressão beta proposto por Ferrari e Cribari–Neto (2004), considerando a modelagem conjunta dos parâmetros da distribuição e a inclusão de funções não lineares. Os autores apresentam ainda correções analíticas de viés para os estimadores de máxima verossimilhança, generalizando os resultados apresentados em Ospina et al. (2006). Rocha e Simas (2011) apresentam medidas de diagnóstico baseadas no método de influência local para o modelo de regressão beta introduzido em Simas et al. (2010). Bayer e Cribari–Neto (2014) avaliou o desempenho de diferentes critérios de seleção de modelos em amostras de tamanho finito no modelo de regressão beta. Recentemente, Espinheira et al. (2014) propuseram intervalos de predição para o modelo de regressão beta baseados em reamostragem bootstrap e avaliaram seu desempenho em amostras de tamanho finito.

Uma alternativa competitiva à distribuição beta é o modelo simplex proposto por Barndorff–Nielsen e Jorgersen (1991). A distribuição simplex faz parte dos modelos de dispersão definidos por Jorgersen (1997) que estendem os modelos lineares generalizados. A distribuição simplex tem sido utilizada por diversos autores para modelar dados restritos ao intervalo (0,1). Song e Tan (2000), por exemplo, propuseram um modelo de regressão simplex com dispersão constante para modelar dados contínuos de proporção longitudinais, sob o enfoque de equações de estimação generalizadas (EEG). Essa abordagem foi modificada por Song et al. (2004), assumindo que o parâmetro de dispersão varia ao longo das observações. Venezuela (2007) apresenta medidas de diagnóstico baseadas no método de influência local (Cook, 1986) para o modelo de regressão simplex também sob a perspectiva de EEG. Seguindo Ferrari e Cribari–Neto (2004), Miyashiro (2008) propôs o modelo de regressão simplex com dispersão constante, apresentando expressões analíticas para o vetor escore e para a matriz de informação de Fisher correspondente ao vetor de parâmetros, além de medidas de diagnóstico. Mais recentemente, López (2013) avaliou através de simulações de Monte Carlo o comportamento de estimadores bayesianos para os parâmetros do modelo de regressão simplex com dispersão variável.

Até o momento, nenhum trabalho tem abordado um modelo de regressão simplex geral, em que tanto a média da variável resposta quanto o parâmetro de precisão estão relacionados às covariáveis por meio de preditores não lineares. Desse modo, a presente proposta de dissertação de mestrado tem por objetivo propor um modelo de regressão simplex não linear geral, bem como desenvolver vários aspectos de inferência e diagnóstico para essa classe de modelos. No que se refere a análise de diagnóstico, propomos resíduo e medidas de influência local baseadas no método proposto por Cook (1986). Além disso, ilustraremos a utilidade do modelo proposto através de aplicações a dados reais e dados simulados.

1.1 Organização da dissertação

Esta dissertação encontra-se dividida em cinco capítulos. No segundo capítulo definimos a distribuição simplex para modelar dados contínuos no intervalo unitário e propomos uma

extensão do modelo de regressão simplex constante (Miyashiro, 2008), em que a média da variável resposta e o parâmetro de precisão são modelados através de funções não lineares. Além disso, apresentamos expressões em forma fechada para o vetor escore, matriz de informação de Fisher e sua inversa, e discutimos a estimação dos parâmetros do modelo pelo método de máxima verossimilhança.

No terceiro capítulo, desenvolvemos técnicas de diagnóstico para o modelo simplex não linear. Nesse sentido, propomos um resíduo para o modelo baseado no processo iterativo escore de Fisher e avaliamos o comportamento de sua distribuição empírica através de simulações de Monte Carlo. Além disso, obtemos medidas de influência local, baseadas no método proposto por Cook (1986), sob diferentes esquemas de perturbação, e derivamos uma expressão matricial para a alavanca generalizada através da definição geral para o método de máxima verossimilhança apresentada por Wei et al. (1998).

No quarto capítulo, apresentamos aplicações da teoria desenvolvida a três conjuntos de dados. Na primeira, geramos uma única amostra com a presença de erro na entrada dos dados, simulando uma observação alavanca e influente. Na segunda aplicação, consideramos os dados de oxidação de amônia, analisados por Miyashiro (2008) e na terceira, os dados de proporção de gafanhotos mortos, apresentados em McCullagh e Nelder (1989, pg.384).

1.2 Suporte computacional

Todos os resultados gráficos e numéricos apresentados nesta dissertação foram obtidos utilizando o ambiente de programação, análise de dados e gráficos `R`, em sua versão 3.0.2 para sistema operacional Microsoft Windows, que se encontra disponível gratuitamente através do site `http://www.R-project.org`. O `R` foi criado por Ross Ihaka e Robert Gentleman na Universidade de Auckland com o objetivo de produzir um ambiente de programação parecido com S, uma linguagem desenvolvida no AT&T Bell Laboratories, cuja versão comercial é o S-Plus, tendo as vantagens de ser de livre distribuição e possuir código fonte aberto. Maiores detalhes sobre o `R` podem ser encontrados em Cribari–Neto e Zarkos (1999).

Esta dissertação foi digitada utilizando o sistema de tipografia LaTeX desenvolvido por Leslie Lamport em 1985, que 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 do texto. Detalhes sobre o sistema de tipografia LaTeX podem ser encontrados em Lamport (1994) ou através do site `http://www.tex.ac.uk/CTAN/latex`.

CAPÍTULO 2

MODELO DE REGRESSÃO SIMPLEX NÃO LINEAR

2.1 Distribuição simplex

A distribuição simplex, proposta por Barndorff–Nielsen e Jorgensen (1991) para modelagem de dados restritos ao intervalo contínuo (0,1), pertence aos modelos de dispersão introduzidos por Jorgensen (1997), que estendem os modelos lineares generalizados (Nelder e Wedderburn, 1972). De modo geral, dizemos que uma variável aleatória y possui distribuição pertencente aos modelos de dispersão se sua função densidade de probabilidade pode ser escrita na 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, $\mu \in \Omega$ é o parâmetro de locação, $\sigma^2 > 0$ é o parâmetro de dispersão e $a \geq 0$ é uma constante normalizadora, independente de μ. A função $d(y;\mu)$ é conhecida como componente do desvio e é definida em $(y,\mu) \in (C,\Omega)$ e deve satisfazer as seguintes propriedades

$$d(y;y) = 0, \forall\, y = \mu \in \Omega \quad \text{e} \quad d(y;\mu) > 0, \forall\, y \neq \mu.$$

A função de variância para os modelos de dispersão é definida como

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

desde que $d(y;\mu)$ seja contínua e duas vezes diferenciável com respeito μ e satisfaça

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

Em particular, se uma variável aleatória y segue a distribuição simplex, denotada por $\mathcal{S}^{-1}(\mu, \lambda)$ com parâmetros $\mu \in (0,1)$ e $\lambda > 0$, em que $\lambda = 1/\sigma^2$, a expressão (2.1) toma a forma

$$p(y;\mu,\lambda) = \left\{\frac{\lambda}{2\pi\{y(1-y)\}^3}\right\}^{\frac{1}{2}} \exp\left\{-\frac{\lambda}{2}d(y;\mu)\right\}, y \in (0,1), \tag{2.2}$$

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

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

A função de variância para a distribuição simplex é dada por $V(\mu) = \mu^3(1-\mu)^3$. Usando resultados apresentados em Jorgensen (1997) segue ainda que $\mathrm{E}(Y) = \mu$ e

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

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

A Figura 2.1.1 apresenta diferentes densidades simplex correspondentes a alguns valores dos parâmetros (μ,λ). Em particular, para $\mu = 1/2$ a densidade simplex é simétrica, enquanto que para $\mu \neq 1/2$ apresenta forma assimétrica. Adicionalmente, a densidade pode apresentar forma de J, U, J invertido e bimodal.

2.2 Definição e estimação do modelo

Nesta seção, apresentamos uma extensão do modelo de regressão simplex proposto por Miyashiro (2008), em que tanto a média da variável resposta μ, quanto o parâmetro de precisão λ estão relacionados às covariáveis por meio de preditores não lineares.

Sejam $y_1, \ldots, y_n$ variáveis aleatórias independentes, em que cada y_t, $t = 1, \ldots, n$, segue a densidade (2.2), isto é, $y_t \sim \mathcal{S}^{-1}(\mu_t, \lambda_t)$. O modelo de regressão simplex não linear é definido por (2.2) e pelas componentes sistemáticas

$$g(\mu_t) = f_1(x_t^\top;\beta) = \eta_t \quad \text{e} \quad h(\lambda_t) = f_2(z_t^\top;\gamma) = \zeta_t, \tag{2.4}$$

em que $\beta = (\beta_1, \ldots, \beta_k)^\top$ e $\gamma = (\gamma_1, \ldots, \gamma_q)^\top$ são vetores de parâmetros de regressão desconhecidos tais que $\beta \in \mathbb{R}^k$ e $\gamma \in \mathbb{R}^q$, $k+q < n$, η_t e ζ_t são preditores não lineares, e $x_{t1}, \ldots, x_{tk}$, $z_{t1}\ldots, z_{tq}$ são observações de covariáveis conhecidas, que podem coincidir total ou parcialmente. Assumimos que as funções de ligação $g : (0,1) \to \mathbb{R}$ e $h : (0,\infty) \to \mathbb{R}$ são estritamente monótonas e duas vezes diferenciáveis. Diferentes funções de ligação podem ser escolhidas para g e h. Por exemplo, para μ podemos usar a especificação logito $g(\mu) = \log\{\mu/(1-\mu)\}$, a função probito $g(\mu) = \Phi^{-1}(\mu)$, em que $\Phi(\cdot)$ denota a função de distribuição normal padrão, a função log-log complementar $g(\mu) = \log\{-\log(1-\mu)\}$ e a função log-log $g(\mu) = \log\{-\log(\mu)\}$,

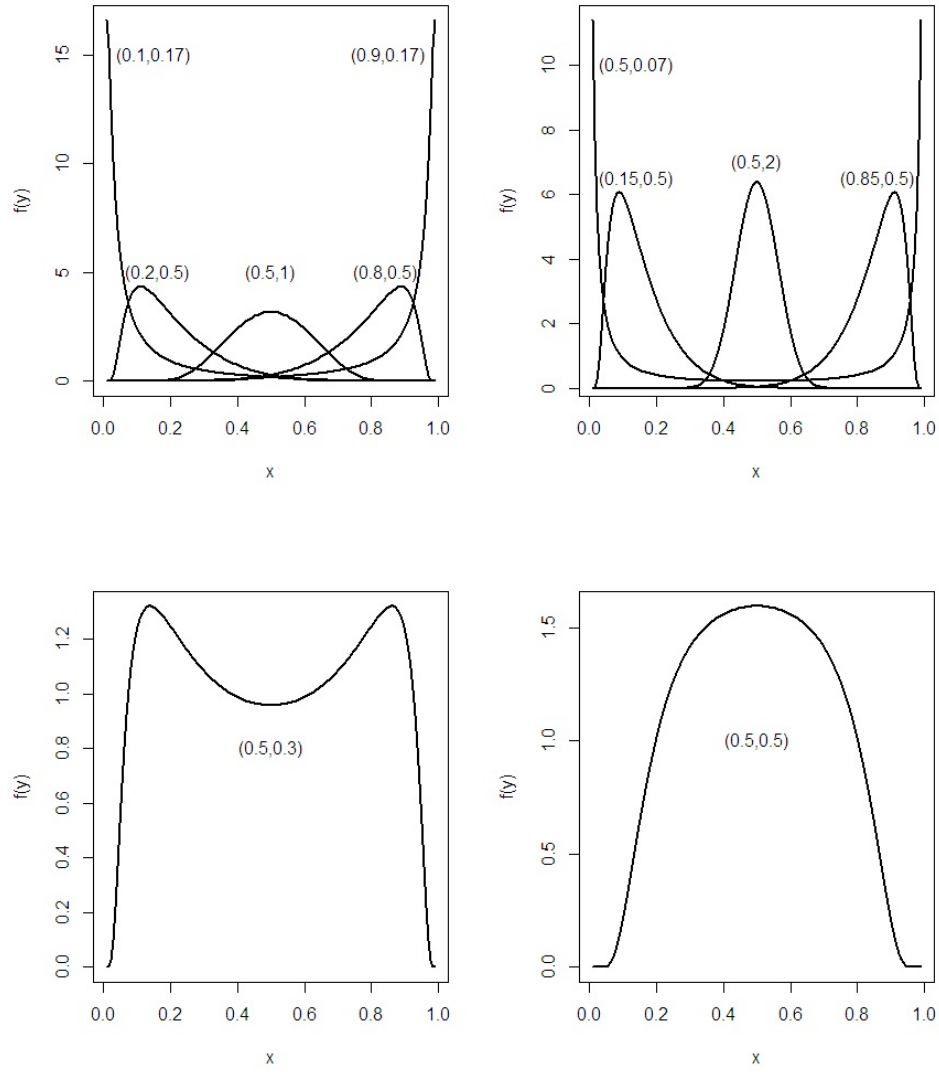

Figura 2.1.1 – Densidades simplex para diferentes valores de (μ,λ).

entre outras. Como $\lambda > 0$, podemos usar a função logarítmica $h(\lambda) = \log(\lambda)$, a função raiz quadrada $h(\lambda) = \sqrt{\lambda}$ e a função identidade $h(\lambda) = \lambda$, com atenção especial para a positividade das estimativas. Uma discussão detalhada destas funções de ligação pode ser encontrada em McCullagh and Nelder (1989); ver também Atkinson (1985).

Baseado em (2.2) seque que o logaritmo da função de verossimilhança tem a forma

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

em que

$$\ell_t(\mu_t, \lambda_t) = \frac{1}{2}\log \lambda_t - \frac{1}{2}\log 2\pi - \frac{3}{2}\log\{y_t(1-y_t)\} - \frac{\lambda_t}{2} d(y_t; \mu_t). \tag{2.5}$$

Os componentes do vetor escore $U_\beta(\beta, \gamma)$, obtidos por diferenciação do logaritmo da

função verossimilhança com respeito a β_i, são dados, para $i = 1, \ldots, k$ por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t, \lambda_t)}{\partial \mu_t} \frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t} \frac{\partial \eta_t}{\partial \beta_i}, \tag{2.6}$$

em que $\mathrm{d}\mu_t/\mathrm{d}\eta_t = 1/g'(\mu_t)$ e

$$\frac{\partial \ell_t(\mu_t, \lambda_t)}{\partial \mu_t} = -\frac{\lambda_t}{2} \frac{\partial d(y_t; \mu_t)}{\partial \mu_t}. \tag{2.7}$$

De (A.1) (ver Apêndice A) temos que $-(1/2)\partial d(y_t; \mu_t)/\partial \mu_t = (y_t - \mu_t)u_t$. Assim,

$$\frac{\partial \ell(\beta, \gamma)}{\partial \beta_i} = \sum_{t=1}^{n} \lambda_t (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} \frac{\partial \eta_t}{\partial \beta_i}, \tag{2.8}$$

em que

$$u_t = \frac{1}{\mu_t(1-\mu_t)} \left\{ d(y_t; \mu_t) + \frac{1}{\mu_t^2(1-\mu_t)^2} \right\}. \tag{2.9}$$

Matricialmente, a função escore para o vetor $\beta = (\beta_1, \ldots, \beta_k)^\top$ pode ser expressa por

$$U_\beta(\beta, \gamma) = \tilde{X}^\top \Lambda U T (y - \mu), \tag{2.10}$$

em que $\tilde{X} = \partial\eta/\partial\beta$ é uma matriz de derivadas de dimensão $n \times k$, $U = \mathrm{diag}\{u_1, \ldots, u_n\}$,

$$\Lambda = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\} \tag{2.11}$$

e

$$T = \mathrm{diag}\{1/g'(\mu_1), \ldots, 1/g'(\mu_n)\}. \tag{2.12}$$

Analogamente, diferenciando o logaritmo da função de verossimilhança com respeito a γ_j, $j = 1, \ldots, q$, temos que

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t, \lambda_t)}{\partial \lambda_t} \frac{\mathrm{d}\lambda_t}{\mathrm{d}\zeta_t} \frac{\partial \zeta_t}{\partial \gamma_j}, \tag{2.13}$$

com $\mathrm{d}\lambda_t/\mathrm{d}\zeta_t = 1/h'(\lambda_t)$. Além disso,

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

Assim, o j-ésimo componente do vetor $U_\gamma(\beta, \gamma)$ é dado por

$$\frac{\partial \ell(\beta, \gamma)}{\partial \gamma_j} = \sum_{t=1}^{n} \left\{ \frac{1}{2\lambda_t} - \frac{d(y_t; \mu_t)}{2} \right\} \frac{1}{h'(\lambda_t)} \frac{\partial \zeta_t}{\partial \gamma_j}.$$

A função escore para $\gamma = (\gamma_1, \ldots, \gamma_q)^\top$ pode ser expressa em forma matricial como

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

sendo $\tilde{Z} = \partial\zeta/\partial\gamma$ uma matriz de derivadas de dimensão $n \times q$, $a = (a_1, \ldots, a_n)^\top$, com

$$a_t = \frac{1}{2\lambda_t} - \frac{d(y_t;\mu_t)}{2} \tag{2.14}$$

e

$$H = \text{diag}\{1/h'(\lambda_1), \ldots, 1/h'(\lambda_n)\}. \tag{2.15}$$

Para obtenção da matriz de informação de Fisher conjunta dos vetores β e γ, torna-se necessário o cálculo das derivadas de segunda ordem do logaritmo da função de verossimilhança. De (2.6) temos que para $i, p = 1, \ldots, k$

$$\begin{aligned}
\frac{\partial^2\ell(\beta,\gamma)}{\partial\beta_i\partial\beta_p} &= \sum_{t=1}^{n}\frac{\partial}{\partial\mu_t}\left(\frac{\partial\ell_t(\mu_t,\lambda_t)}{\partial\mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\right)\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial\eta_t}{\partial\beta_i}\frac{\partial\eta_t}{\partial\beta_p} \\
&+ \sum_{t=1}^{n}\frac{\partial\ell_t(\mu_t,\lambda_t)}{\partial\mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial^2\eta_t}{\partial\beta_i\partial\beta_p} \\
&= \sum_{t=1}^{n}\left(\frac{\partial^2\ell_t(\mu_t,\lambda_t)}{\partial\mu_t^2}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t} + \frac{\partial\ell_t(\mu_t,\lambda_t)}{\partial\mu_t}\frac{\partial}{\partial\mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\right)\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial\eta_t}{\partial\beta_i}\frac{\partial\eta_t}{\partial\beta_p} \\
&+ \sum_{t=1}^{n}\frac{\partial\ell_t(\mu_t,\lambda_t)}{\partial\mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\frac{\partial^2\eta_t}{\partial\beta_i\partial\beta_p}.
\end{aligned}$$

Como $\mathrm{E}(\partial\ell_t(\mu_t,\lambda_t)/\partial\mu_t) = 0$, temos que

$$\mathrm{E}\left(\frac{\partial^2\ell(\beta,\gamma)}{\partial\beta_i\partial\beta_p}\right) = \sum_{t=1}^{n}\mathrm{E}\left(\frac{\partial^2\ell_t(\mu_t,\lambda_t)}{\partial\mu_t^2}\right)\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t}\right)^2\frac{\partial\eta_t}{\partial\beta_i}\frac{\partial\eta_t}{\partial\beta_p}.$$

De (2.7) seque que

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

Assim, pela Proposição A.1 (e) (Apêndice A)

$$\frac{1}{2}\mathrm{E}\left(\frac{\partial^2 d(y_t;\mu_t)}{\partial\mu_t^2}\right) = \frac{3}{\lambda_t\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}$$

e definindo

$$w_t = \left\{\frac{3}{\lambda_t\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}\right\}\frac{1}{g'(\mu_t)^2}, \tag{2.16}$$

temos que

$$\mathrm{E}\left(\frac{\partial^2\ell(\beta,\gamma)}{\partial\beta_i\partial\beta_p}\right) = -\sum_{t=1}^{n}\lambda_t w_t\frac{\partial\eta_t}{\partial\beta_i}\frac{\partial\eta_t}{\partial\beta_p}. \tag{2.17}$$

A expressão em (2.17) pode ser escrita matricialmente como

$$\mathrm{E}\left(\frac{\partial^2\ell(\beta,\gamma)}{\partial\beta\partial\beta^\top}\right) = -\tilde{X}^\top\Lambda W\tilde{X},$$

em que $W = \text{diag}\{w_1, \ldots, w_n\}$.

Adicionalmente, segue de (2.8) que as derivadas de segunda ordem de $\ell(\beta, \gamma)$ com respeito a β_i e γ_j, para $i = 1, \ldots, k$ e $j = 1, \ldots, q$, são dadas por

$$\frac{\partial^2 \ell(\beta,\gamma)}{\partial\beta_i \partial\gamma_j} = \sum_{t=1}^{n} (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} \frac{1}{h'(\lambda_t)} \frac{\partial \eta_t}{\partial \beta_i} \frac{\partial \zeta_t}{\partial \gamma_j}.$$

De (A.2) segue que $\text{E}\{(y_t - \mu_t)u_t\} = 0$ (Ver Apêndice A). Assim, obtemos

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

Agora, de (2.13) temos que as segundas derivadas de $\ell(\beta, \gamma)$ com respeito a γ_j e γ_l, para $j, l = 1, \ldots, q$, são dadas por

$$\begin{aligned}
\frac{\partial^2 \ell(\beta,\lambda)}{\partial\gamma_j \partial\gamma_l} &= \sum_{t=1}^{n} \frac{\partial}{\partial \lambda_t}\left(\frac{\partial \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t}\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\right)\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\frac{\partial \zeta_t}{\partial \gamma_j}\frac{\partial \zeta_t}{\partial \gamma_l} \\
&+ \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t}\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\frac{\partial^2 \zeta_t}{\partial \gamma_j \partial \gamma_l} \\
&= \sum_{t=1}^{n}\left(\frac{\partial^2 \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t^2}\frac{\text{d}\lambda_t}{\text{d}\zeta_t} + \frac{\partial \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t}\frac{\partial}{\partial \lambda_t}\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\right)\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\frac{\partial \zeta_t}{\partial \gamma_j}\frac{\partial \eta_t}{\partial \gamma_l} \\
&+ \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t}\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\frac{\partial^2 \zeta_t}{\partial \gamma_j \partial \gamma_l}.
\end{aligned}$$

Dado que $\text{E}(\partial \ell_t(\mu_t, \lambda_t)/\partial \lambda_t) = 0$, temos

$$\text{E}\left(\frac{\partial^2 \ell(\beta,\gamma)}{\partial\gamma_j \partial\gamma_l}\right) = \sum_{t=1}^{n} \text{E}\left(\frac{\partial^2 \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t^2}\right)\left(\frac{\text{d}\lambda_t}{\text{d}\zeta_t}\right)^2 \frac{\partial \zeta_t}{\partial \gamma_j}\frac{\partial \zeta_t}{\partial \gamma_l}.$$

Além disso, de (2.14) obtemos

$$\frac{\partial^2 \ell_t(\mu_t,\lambda_t)}{\partial \lambda_t^2} = -\frac{1}{2\lambda_t^2}.$$

Desse modo,

$$\text{E}\left(\frac{\partial^2 \ell(\beta,\gamma)}{\partial\gamma_j \partial\gamma_l}\right) = -\sum_{t=1}^{n} d_t \frac{\partial \zeta_t}{\partial \gamma_j}\frac{\partial \zeta_t}{\partial \gamma_l},$$

que pode ser expressa matricialmente por

$$\text{E}\left(\frac{\partial^2 \ell(\beta,\gamma)}{\partial\gamma \partial\gamma^\top}\right) = -\tilde{Z}^\top D \tilde{Z},$$

em que $D = \text{diag}\{d_1, \ldots, d_n\}$, com

$$d_t = \left\{\frac{1}{2\lambda_t^2}\right\}\frac{1}{h'(\lambda_t)^2}. \tag{2.18}$$

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

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

em que

$$K_{\beta\beta} = \tilde{X}^\top \Lambda W \tilde{X} \tag{2.19}$$

e

$$K_{\gamma\gamma} = \tilde{Z}^\top D \tilde{Z}.$$

Como K é uma matriz bloco diagonal, os vetores de parâmetros β e γ são ortogonais, de modo que seus estimadores de máxima verossimilhança $\widehat{\beta}$ e $\widehat{\gamma}$, respectivamente, são assintoticamente independentes.

Para grandes amostras e sob condições de regularidade, a distribuição aproximada conjunta de $\widehat{\beta}$ e $\widehat{\gamma}$ é normal $(k+q)$ multivariada, de forma que

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

em que

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

com $K^{\beta\beta} = (\tilde{X}^\top \Lambda W \tilde{X})^{-1}$ e $K^{\gamma\gamma} = (\tilde{Z}^\top D \tilde{Z})^{-1}$.

Os estimadores de máxima verossimilhança de β e γ são obtidos como solução do sistema $U_\beta(\beta, \gamma) = 0$ e $U_\gamma(\beta, \gamma) = 0$. No entanto, destacamos que os estimadores de máxima verossimilhança em modelos não lineares, em geral, não apresentam expressões analíticas em forma fechada, tornando-se necessário a utilização de métodos iterativos, tal como os algoritmos quasi-Newton (por exemplo, BFGS); ver Nocedal e Wright (1999) e Press et al. (2007). Os algoritmos de otimização requerem a especificação de um valor $\theta = (\beta^\top, \gamma^\top)^\top$ para inicializar o processo iterativo. Para o modelo de regressão beta não linear, Simas et al.(2010) sugerem utilizar como valor inicial para o vetor θ as estimativas obtidas no ajuste do modelo

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

assumindo que a variável y_t segue distribuição normal com média μ_t e variância σ_t^2, ou seja, $y_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$, $t = 1, \ldots, n$. Para o modelo de regressão simplex não linear, no entanto, sugerimos como estimativa inicial para o vetor $\theta = (\beta^\top, \gamma^\top)^\top$, as estimativas obtidas através do método de mínimos quadrados não lineares. Para isso, consideramos o pacote `nls2` implementado no software `R`.

CAPÍTULO 3

TÉCNICAS DE DIAGNÓSTICO

3.1 Introdução

Após a estimação de um modelo de regressão é necessário verificar possíveis afastamentos das suposições iniciais feitas para o modelo, uma vez que este representa uma aproximação para o verdadeiro processo gerador dos dados. Nesse sentido é importante detectar a presença de observações extremas no conjunto de dados e avaliar o impacto dessas observações nos resultados inferenciais. Por outro lado, avaliar a adequação da distribuição de probabilidades aos dados é fundamental para a validação do modelo estimado. Para tanto, a análise de diagnóstico, que teve início com a análise de resíduos, configura-se como uma etapa imprescindível, no sentido de analisar a estabilidade e robustez do processo de estimação do modelo postulado e avaliar a adequação da distribuição aos dados.

A análise de resíduos pode ser baseada nos resíduos ordinários, em versões padronizadas, em resíduos construídos a partir dos componentes da função desvio (McCullagh e Nelder, 1989) ou em resíduos generalizados (Cox e Snell, 1968). Diversos trabalhos na literatura têm abordado diferentes aspectos dos resíduos, tais como definição, comportamento distribucional e correção (Ospina, 2008). Nesta direção, destacam-se os trabalhos de Pregibon (1981), McCullagh (1987), Copas (1988) Williams (1984, 1987), Davison e Gigli (1989), Pierce e Schafer (1986) , Farhrmeir e Tutz (1994), Paula (1995), de Souza e Paula (2002), Dunn e Smyth (1996), Ferrari e Cribari–Neto (2004) e Espinheira et al. (2008a), entre outros. As técnicas gráficas utilizando resíduos são frequentemente adotadas para a análise de diagnóstico. O uso de envelopes simulados, por exemplo, conforme proposto por Atkinson (1981) inicialmente para o modelo de regressão normal, permite avaliar a adequação do modelo postulado.

Outro aspecto importante na análise de diagnóstico é a detecção de observações que exercem efeitos desproporcionais sobre o ajuste do modelo. Nesse contexto, Cook (1986) propôs o método de influência local, que consiste em avaliar a influência das observações a partir de

pequenas perturbações introduzidas no modelo ou nos dados. Se pequenas modificações na formulação original do modelo causam efeitos desproporcionais sobre os resultados inferenciais, existem fortes evidências de falta de qualidade no ajuste ou violação das suposições assumidas. Diversos trabalhos na literatura tem abordado o método de influência local em modelos de regressão. Por exemplo, Tsai e Wu (1992) investigam medidas de influência local em modelos auto-regressivos de primeira ordem e modelos heterocedásticos, e Paula (1996) em modelos próprios de dispersão. Galea et al. (1997), Liu (2000) e Galea et al. (2003) apresentam medidas de influência local para modelos elípticos lineares e Liu (2002) para modelos elípticos multivariados. Ortega et al. (2003) estuda o método de influência local para os modelos log-gama generalizados com dados censurados. Para os modelos de regressão beta, destacam-se os trabalhos de Espinheira et al. (2008b) e Rocha e Simas (2011).

Este capítulo encontra-se organizado da seguinte forma. Na Seção 3.2, propomos o resíduo ponderado padronizado para o modelo de regressão simplex não linear baseado no trabalho de Espinheira et al. (2008a). Na Seção 3.3 apresentamos medidas de influência local para o modelo de regressão simplex não linear sob diferentes esquemas de perturbação. Por fim, na Seção 3.5, apresentamos uma medida de alavancagem baseada na definição geral apresentada por Wei et al. (1998).

3.2 Resíduo ponderado padronizado

Podemos definir como resíduo uma medida que objetiva identificar discrepâncias entre o modelo estimado e os dados. Desse modo, é compreensível que a maioria das definições de resíduos estejam baseadas na distância $y_t - \widehat{\mathrm{E}(y_t)}$. No entanto, respeitado o formato da distribuição de probabilidades da variável resposta, é mais interessante pensar no resíduo como uma função $r(y_t, \widehat{\mathrm{E}(y_t)})$, definição geral de resíduos proposta por Cox e Snell (1968). Sob essa perspectiva, Espinheira et al. (2008a) sugerem utilizar resíduos padronizados obtidos da convergência do processo iterativo escore de Fisher para estimação do vetor de parâmetros de regressão. Neste sentido, estendemos tal ideia para o modelo de regressão simplex não linear definido no Capítulo 2.

Considerando o modelo simplex não linear em (2.4), temos que o processo iterativo escore de Fisher de β é dado por

$$\beta^{(m+1)} = \beta^{(m)} + (K_{\beta\beta}^{(m)})^{-1} U_{\beta}^{(m)}, \tag{3.1}$$

em que $m = 0, 1, 2, \ldots$ denota os passos do processo iterativo que é repetido até que a distância entre $\beta^{(m+1)}$ e $\beta^{(m)}$ seja menor que um valor de tolerância especificado, ocorrendo assim a convergência do processo. Considerando a função escore e a matriz de informação de Fisher para o vetor β definidas em (2.10) e (2.19), temos que o processo iterativo em (3.1) após a

convergência é dado por

$$\widehat{\beta} = (\widehat{\tilde{X}}^{\top}\widehat{\Lambda}\widehat{W}\widehat{\tilde{X}})^{-1}\widehat{\tilde{X}}^{\top}\widehat{\Lambda}\widehat{W}\omega, \tag{3.2}$$

com

$$\omega = \widehat{\tilde{X}}\widehat{\beta} + \widehat{W}^{-1}\widehat{U}\widehat{T}(y - \widehat{\mu}), \tag{3.3}$$

que pode ser interpretado como a solução de mínimos quadrados da regressão linear de ω contra $\widehat{\tilde{X}}$ com pesos $\widehat{\Lambda}\widehat{W}$. Assim, o resíduo ordinário é dado por $r^* = (\widehat{\Lambda}\widehat{W})^{1/2}(\omega - \widehat{\tilde{X}}\widehat{\beta}) = \widehat{\Lambda}^{1/2}\widehat{W}^{-1/2}\widehat{U}\widehat{T}(y - \widehat{\mu})$. Usando as definições de Λ , W, U e T dadas em (2.11), (2.16), (2.9) e (2.12), respectivamente, propomos o resíduo ponderado para o modelo simplex não linear, definido por

$$r_t^* = \frac{\widehat{u}_t(y - \widehat{\mu_t})}{\sqrt{\widehat{v_t}}}$$

em que

$$v_t = \frac{1}{\lambda_t}\left\{\frac{3}{\lambda_t\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}\right\}.$$

Uma possível padronização para o resíduo ponderado baseia-se na variância de ω. Reescrevendo (3.2) como $(\widehat{\tilde{X}}^{\top}\widehat{\Lambda}\widehat{W}\widehat{\tilde{X}}) = \widehat{\tilde{X}}^{\top}\widehat{\Lambda}\widehat{W}\omega$ e usando o fato de que $\mathrm{cov}(\widehat{\beta}) \approx (\widehat{\tilde{X}}^{\top}\widehat{\Lambda}\widehat{W}\widehat{\tilde{X}})^{-1}$ segue que $\mathrm{cov}(\omega) \approx (\widehat{\Lambda}\widehat{W})^{-1}$. Assim, a partir de (3.3) obtemos

$$\mathrm{cov}(r^*) = \mathcal{I}_n - H^*,$$

em que

$$H^* = (\widehat{\Lambda}\widehat{W})^{1/2}\widehat{\tilde{X}}(\widehat{\tilde{X}}^{\top}\widehat{\Lambda}\widehat{W}\widehat{\tilde{X}})^{-1}\widehat{\tilde{X}}^{\top}(\widehat{\Lambda}\widehat{W})^{1/2}$$

é a matriz de projeção e $\mathcal{I}_n$ é a matriz identidade de ordem n. Logo, o resíduo ponderado padronizado é definido por

$$r_t^{pp} = \frac{r_t^*}{\sqrt{(1-h_{tt}^*)}} = \frac{\widehat{u}_t(y - \widehat{\mu_t})}{\sqrt{\widehat{v}_t(1-h_{tt}^*)}}, \tag{3.4}$$

sendo h_{tt}^* o t-ésimo elemento da diagonal principal da matriz H^*. Gráficos de r_t^{pp} versus os índices das observações, contra os valores das covariáveis ou contra os valores preditos, podem ser utilizados para avaliar a adequação da componente sistemática do modelo. Além disso, sugere-se a utilização do gráfico normal de probabilidades com envelopes para investigar possíveis afastamentos das suposições assumidas quanto a parte aleatória do modelo.

3.2.1 Avaliação numérica: modelo simplex linear constante

Para investigar o comportamento da distribuição empírica do resíduo ponderado padronizado em (3.4), realizamos simulações de Monte Carlo com 5000 réplicas. Inicialmente, consideramos o modelo de regressão simplex com

$$g(\mu_t) = \beta_1 + \beta_2 x_t, \quad t = 1, \ldots, 20, \tag{3.5}$$

em que $g(\cdot)$ é a função de ligação logito e os valores da variável explicativa x_t são conhecidos para cada observação. Vale destacar que os valores da covariável x_t foram obtidos de realizações independentes da distribuição de probabilidades $\mathcal{U}(0,1)$ e mantidos fixos para cada réplica de Monte Carlo. Os valores do parâmetro de precisão λ foram 0,5; 1,25 e 3,5. Além disso, consideramos três cenários diferentes. No primeiro, os valores da média da variável resposta encontram-se no extremo inferior do intervalo (0,1), mas especificamente, $\mu \in (0,03; 0,18)$ e neste caso, $\beta_1 = -1,4$ e $\beta_2 = -2,5$. No segundo cenário, fixamos $\beta_1 = 1,2$ e $\beta_2 = -2,1$, o que conduz a valores de $\mu \in (0,35; 0,76)$. Finalmente, consideramos o caso em que os valores da média da variável resposta encontram-se próximos ao extremo superior do intervalo unitário, isto é, $\mu \in (0,81; 0,95)$, sendo neste cenário $\beta_1 = 1,4$ e $\beta_2 = 1,7$.

Nas Tabelas 3.2.1–3.2.9, apresentamos medidas descritivas para as 5000 réplicas do t-ésimo resíduo r_t^{pp}, $t = 1, \ldots, 20$. Nota-se, de modo geral, que a média do resíduo ponderado padronizado é aproximadamente zero em todos cenários investigados e para todos os valores de λ. Além disso, observa-se que o desvio-padrão de r_t^{pp} é ligeiramente maior que o desvio-padrão da distribuição normal padrão para $\lambda = 0,5$ (Tabelas 3.2.1, 3.2.4 e 3.2.7). No entanto, a medida que aumentamos a precisão, $\lambda = 1,25$ e $\lambda = 3,5$, observa-se para todos os cenários que o desvio-padrão do resíduo apresenta valores inferiores a 1. Quanto à forma da distribuição do resíduo, verifica-se, para o cenário em que $\mu \in (0,03; 0,18)$, que esta apresenta assimetria positiva e achatamento superior ao da distribuição $\mathcal{N}(0,1)$ para o caso em que $\lambda = 0,5$. Para valores maiores de λ, no entanto, o resíduo tende a apresentar menores valores de assimetria e curtose próxima de 3. No cenário em que a média da variável resposta concentra-se no extremo superior do intervalo (0,1), isto é, $\mu \in (0,81; 0,95)$ (Tabelas 3.2.7–3.2.9), nota-se que o resíduo apresenta assimetria negativa, diferentemente do que ocorre no cenário em que $\mu \in (0,03; 0,18)$ (Tabelas 3.2.1–3.2.3). Além disso, a curtose da distribuição de r_t^{pp} tende a acompanhar a curtose da distribuição normal padrão a medida que aumentamos a precisão dos dados. Já para o caso que $\mu \in (0,35; 0,76)$, observa-se que a distribuição do resíduo ponderado padronizado apresenta assimetria ligeiramente menor quando comparada aos demais cenários considerados para a média da variável resposta, independente do valor de λ.

Tabela 3.2.1 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,03; 0,18)$, $\lambda = 0,5$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0598$	1,3733	0,5741	3,2488
2	0,0322	1,4408	0,6596	3,2362
3	0,0352	1,4474	0,9770	4,3322
4	0,0030	1,4448	0,8591	3,8545
5	$-0,0413$	1,3835	0,5029	3,0576
6	0,0222	1,4592	1,0761	4,4781
7	$-0,0424$	1,3922	0,8776	4,1629
8	$-0,0096$	1,4220	1,0042	4,3481
9	$-0,0207$	1,4213	0,8756	4,2129
10	$-0,0046$	1,4108	0,8439	3,8077
11	0,0476	1,4543	0,8846	3,9057
12	$-0,0013$	1,4378	1,0310	4,5947
13	0,0358	1,4432	1,0084	4,2669
14	0,0347	1,4608	0,9887	4,3475
15	$-0,0216$	1,4190	1,0740	4,5331
16	$-0,0109$	1,4242	1,0128	4,3700
17	0,0130	1,4219	1,0520	4,5833
18	$-0,0072$	1,4634	1,0707	4,5357
19	$-0,0031$	1,4097	0,7475	3,5446
20	$-0,0150$	1,4197	0,8550	3,8124

Tabela 3.2.2 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,03; 0,18)$, $\lambda = 1,25$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0152$	0,9108	0,2420	2,7752
2	0,0200	0,8857	0,2447	2,8660
3	0,0204	0,9077	0,4085	3,0830
4	$-0,0022$	0,9220	0,3506	2,9010
5	$-0,0010$	0,9125	0,2601	2,8376
6	0,0044	0,8913	0,4171	3,0014
7	$-0,0105$	0,9102	0,4169	3,0489
8	0,0140	0,9143	0,4226	3,1117
9	0,0166	0,9108	0,3783	3,0815
10	$-0,0046$	0,9023	0,2984	2,8672
11	0,0097	0,8948	0,4209	3,1628
12	0,0442	0,9186	0,4320	3,0339
13	0,0012	0,9053	0,4519	3,0443
14	-0,0169	0,9049	0,3528	2,9168
15	$-0,0222$	0,8930	0,4215	3,1143
16	$-0,0091$	0,8954	0,5043	3,3770
17	0,0012	0,9058	0,4791	3,1979
18	$-0,0071$	0,9133	0,4063	2,9882
19	$-0,0204$	0,9045	0,2927	2,8545
20	$-0,0254$	0,8976	0,3852	3,0111

Tabela 3.2.3 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,03; 0,18)$, $\lambda = 3,5$.

t	média	desvio-padrão	assimetria	curtose
1	0,0020	0,5525	0,0647	2,7397
2	0,0041	0,5540	0,0757	2,7946
3	0,0127	0,5531	0,1065	2,6762
4	0,0020	0,5544	0,1457	2,8014
5	$-0,0155$	0,5397	0,1122	2,7173
6	$-0,0069$	0,5358	0,1817	2,8180
7	$-0,0035$	0,5438	0,1562	2,7961
8	$-0,0072$	0,5443	0,1714	2,8370
9	0,0042	0,5392	0,2070	2,8662
10	$-0,0001$	0,5421	0,0835	2,7976
11	0,0131	0,5471	0,1863	2,8457
12	$-0,0033$	0,5377	0,1144	2,7603
13	0,0009	0,5529	0,1908	2,8434
14	$-0,0008$	0,5382	0,1371	2,8298
15	0,0011	0,5408	0,0947	2,8658
16	$-0,0078$	0,5380	0,1378	2,7778
17	$-0,0139$	0,5385	0,1282	2,8402
18	0,0068	0,5502	0,1866	2,7795
19	0,0042	0,5444	0,1017	2,7226
20	0,0062	0,5411	0,1542	2,8432

Tabela 3.2.4 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,35; 0,76)$, $\lambda = 0,5$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0319$	1,4921	0,2910	3,6625
2	0,0254	1,5432	0,4267	4,0952
3	0,0511	1,5624	0,0072	4,3961
4	$-0,0076$	1,5870	0,3452	4,2665
5	$-0,0269$	1,4890	0,4057	3,8151
6	$-0,0045$	1,5628	$-0,4123$	4,3452
7	$-0,0030$	1,4852	$-0,5416$	3,8511
8	$-0,0008$	1,5726	$-0,3342$	4,3939
9	0,0174	1,4339	$-0,5405$	3,5719
10	0,0336	1,5722	0,3746	3,9909
11	0,0389	1,5985	0,1407	3,9877
12	$-0,0540$	1,5713	$-0,1205$	4,2509
13	$-0,0197$	1,5613	$-0,5969$	4,3143
14	$-0,0347$	1,5560	0,1437	4,0623
15	$-0,0399$	1,5749	$-0,4646$	4,2825
16	$-0,0360$	1,5555	$-0,4700$	4,3334
17	0,0548	1,5838	$-0,2657$	4,6119
18	0,0225	1,5859	$-0,0184$	4,1870
19	0,0101	1,5613	0,3716	4,0608
20	0,0042	1,4847	$-0,5623$	3,9788

Tabela 3.2.5 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,35; 0,76)$, $\lambda = 1,25$.

t	média	desvio-padrão	assimetria	curtose
1	0,0019	0,9125	0,2673	3,2532
2	0,0066	0,9040	0,1492	2,9623
3	0,0059	0,8822	$-0,0063$	3,0381
4	0,0058	0,8804	0,1274	3,1379
5	$-0,0045$	0,8880	0,2276	2,9006
6	$-0,0125$	0,9026	$-0,1348$	3,0871
7	0,0171	0,8854	$-0,2787$	3,0934
8	0,0004	0,9020	$-0,1997$	3,1005
9	$-0,0004$	0,8769	$-0,2405$	2,8581
10	0,0028	0,8821	0,1380	3,0740
11	$-0,0058$	0,9055	0,0450	2,9744
12	$-0,0003$	0,8852	$-0,0117$	3,0052
13	$-0,0014$	0,8845	$-0,2173$	2,9657
14	0,0080	0,9035	0,0864	3,0979
15	0,0041	0,9100	$-0,2100$	3,1805
16	$-0,0022$	0,9000	$-0,1983$	3,1269
17	$-0,0052$	0,9127	$-0,1747$	3,2223
18	$-0,0054$	0,8927	$-0,0771$	3,0303
19	$-0,0061$	0,8930	0,1561	3,0343
20	$-0,0082$	0,8929	$-0,3250$	3,0338

Tabela 3.2.6 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,35; 0,76)$, $\lambda = 3,5$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0073$	0,5393	0,0900	2,8562
2	0,0066	0,5437	0,0779	2,8424
3	$-0,0032$	0,5456	$-0,0056$	2,8446
4	$-0,0097$	0,5367	0,0719	2,7816
5	0,0040	0,5262	0,0906	2,8117
6	$-0,0070$	0,5381	$-0,1097$	2,9251
7	$-0,0082$	0,5286	$-0,1051$	2,7915
8	0,0031	0,5377	$-0,1239$	2,7888
9	0,0020	0,5408	$-0,1159$	2,8722
10	$-0,0004$	0,5347	0,0436	2,7936
11	0,0040	0,5260	0,0122	2,8254
12	0,0076	0,5395	$-0,0119$	2,8743
13	0,0026	0,5350	$-0,1076$	2,7668
14	$-0,0037$	0,5391	0,0335	2,8680
15	$-0,0078$	0,5329	$-0,0468$	2,8031
16	0,0089	0,5422	$-0,1087$	2,7725
17	$-0,0015$	0,5354	$-0,0363$	2,8356
18	0,0062	0,5391	$-0,0698$	2,8113
19	0,0008	0,5412	0,0865	2,8382
20	0,0030	0,5396	$-0,1081$	2,8714

Tabela 3.2.7 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,81; 0,95)$, $\lambda = 0,5$.

t	média	desvio-padrão	assimetria	curtose
1	0,0254	1,3790	$-0,7348$	3,5682
2	0,0181	1,4057	$-0,8644$	3,7940
3	$-0,0514$	1,4474	$-1,1084$	4,7680
4	$-0,0167$	1,4401	$-1,0212$	4,2944
5	0,0487	1,3706	$-0,6478$	3,2410
6	0,0176	1,4252	$-1,1248$	4,8069
7	0,0395	1,4268	$-0,9308$	4,2913
8	$-0,0068$	1,4313	$-1,0537$	4,5529
9	0,0156	1,4273	$-0,7670$	3,6932
10	0,0057	1,3929	$-0,9405$	4,0501
11	$-0,0408$	1,4391	$-1,0463$	4,4624
12	0,0017	1,4397	$-1,0483$	4,4261
13	$-0,0095$	1,4517	$-1,0201$	4,1970
14	$-0,0196$	1,4519	$-1,0658$	4,4748
15	0,0289	1,4247	$-1,0474$	4,5075
16	$-0,0172$	1,4591	$-1,0372$	4,3982
17	$-0,0187$	1,4593	$-1,1162$	4,6671
18	$-0,0226$	1,4878	$-1,3173$	5,9367
19	$-0,0012$	1,4478	$-0,9684$	4,1977
20	0,0135	1,4206	$-0,7770$	3,7177

Tabela 3.2.8 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,81; 0,95)$, $\lambda = 1,25$.

t	média	desvio-padrão	assimetria	curtose
1	0,0120	0,9117	$-0,3561$	2,9174
2	0,0060	0,9025	$-0,4421$	3,1032
3	0,0200	0,8996	$-0,4764$	3,1857
4	$-0,0226$	0,8850	$-0,3921$	3,0836
5	0,0211	0,8979	$-0,2875$	2,8901
6	0,0096	0,9009	$-0,5131$	3,2190
7	0,0147	0,9063	$-0,3814$	2,9849
8	0,0079	0,9187	$-0,4579$	3,0528
9	0,0153	0,8870	$-0,3099$	2,9828
10	0,0247	0,8881	$-0,4474$	3,1394
11	$-0,0292$	0,9093	$-0,4249$	3,0908
12	$-0,0052$	0,9211	$-0,5423$	3,3345
13	0,0040	0,9124	$-0,4884$	3,2550
14	$-0,0069$	0,9110	$-0,4517$	3,2230
15	$-0,0131$	0,9147	$-0,5019$	3,3568
16	$-0,0282$	0,8986	$-0,4633$	3,2119
17	$-0,0054$	0,9047	$-0,4468$	3,1198
18	$-0,0006$	0,9054	$-0,4701$	3,1801
19	$-0,0235$	0,9021	$-0,3272$	2,9433
20	0,0037	0,9058	$-0,3936$	3,0760

Tabela 3.2.9 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,81; 0,95)$, $\lambda = 3,5$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0044$	0,5499	$-0,1183$	2,7819
2	$-0,0058$	0,5468	$-0,1502$	2,8122
3	0,0031	0,5469	$-0,1622$	2,8147
4	0,0033	0,5521	$-0,2046$	2,8463
5	0,0142	0,5370	$-0,1199$	2,8731
6	0,0078	0,5396	$-0,2084$	2,8906
7	0,0101	0,5364	$-0,1746$	2,7959
8	$-0,0233$	0,5439	$-0,1726$	2,7799
9	0,0124	0,5396	$-0,1435$	2,8046
10	$-0,0115$	0,5426	$-0,2179$	2,8348
11	$-0,0009$	0,5420	$-0,2162$	2,8453
12	0,0018	0,5501	$-0,2108$	2,7977
13	0,0005	0,5411	$-0,2109$	2,8364
14	$-0,0027$	0,5497	$-0,2171$	2,8822
15	$-0,0189$	0,5347	$-0,1720$	2,8678
16	0,0087	0,5401	$-0,1870$	2,8597
17	$-0,0029$	0,5452	$-0,1544$	2,8194
18	0,0112	0,5470	$-0,1437$	2,7263
19	0,0037	0,5532	$-0,0892$	2,8111
20	$-0,0044$	0,5411	$-0,1181$	2,8402

3.2.2 Avaliação numérica: modelo simplex linear com precisão variável

Objetivando avaliar o comportamento da distribuição empírica do resíduo r_t^{pp} em (3.4) sob a modelagem do parâmetro de precisão, consideramos adicionalmente a (3.5) que

$$h(\lambda_t) = \gamma_1 + \gamma_2 z_t, \quad t = 1, \ldots, 20,$$

em que $h(\cdot)$ é a função de ligação logarítmica e a variável explicativa z_t é conhecida para cada observação. Os valores da covariável z_t foram obtidos através de realizações independentes de uma variável aleatória com distribuição $\mathcal{U}(0,1)$ e mantidos fixos para cada réplica de Monte Carlo. Para esta investigação, consideramos os três cenários apresentados anteriormente para a média da variável resposta. Definimos o grau de heterogeneidade da dispersão dos dados como

$$\alpha = \max\{\lambda_1, \ldots, \lambda_n\}/\min\{\lambda_1, \ldots, \lambda_n\},$$

$t = 1, \ldots, 20$. Dessa forma, $\alpha = 1$ indica que a precisão é constante ao longo das observações. Para os três cenários descritos, admitimos que $\gamma_1 = 2,2$ e variamos o valor de γ_2 para obter diferentes níveis de heterogeneidade. Assim, fixamos $\gamma_2 = 1$, $\gamma_2 = 2,5$ e $\gamma_2 = 3,2$, o que conduz a valores de $\alpha \approx 2$, $\alpha \approx 10$ e $\alpha \approx 21$, respectivamente. Os resultados da simulação são apresentados nas Tabelas 3.2.10–3.2.18

Nota-se, de modo geral, que a distribuição do resíduo ponderado padronizado sob a modelagem da precisão apresenta média próxima de zero e desvio-padrão significativamente

inferior ao da distribuição $\mathcal{N}(0,1)$. Quanto a assimetria e curtose da distribuição de r_t^{pp}, verifica-se uma assimetria desprezível e uma curtose que acompanha a da distribuição normal padrão. Observa-se ainda que a distribuição do resíduo tende a ser menos assimétrica quando comparada a distribuição de r_t^{pp} sob a suposição de dispersão constante. Dessa forma, ao que parece, a distribuição do resíduo apresenta comportamento similar para todos os cenários considerados para a média da resposta, independentemente do nível de heterogeneidade considerado.

Tabela 3.2.10 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,03; 0,18)$, $\alpha \approx 2$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0042$	0,3010	0,0278	2,7149
2	0,0025	0,2490	0,0352	2,6982
3	$-0,0020$	0,2057	0,0757	2,5848
4	0,0011	0,2352	0,0659	2,6955
5	0,0002	0,2105	0,0598	2,7545
6	$-0,0107$	0,3129	0,0848	2,5142
7	$-0,0026$	0,2655	0,0115	2,8400
8	0,0060	0,2625	0,0359	2,6984
9	0,0050	0,2187	0,0245	2,6253
10	0,0031	0,2548	$-0,0152$	2,7168
11	$-0,0031$	0,2125	0,0293	2,5874
12	0,0003	0,2251	0,0401	2,7121
13	$-0,0028$	0,3199	0,0962	2,5054
14	0,0007	0,2673	0,0378	2,7466
15	0,0058	0,3215	0,0947	2,5076
16	0,0012	0,3183	0,0659	2,4835
17	$-0,0029$	0,2914	$-0,0318$	2,6680
18	$-0,0038$	0,2448	0,0253	2,6370
19	0,0017	0,3070	0,0339	2,5819
20	$-0,0017$	0,2009	0,0297	2,6549

Tabela 3.2.11 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,03; 0,18)$, $\alpha \approx 10$.

t	média	desvio-padrão	assimetria	curtose
1	0,0023	0,2521	0,0331	2,5649
2	$-0,0035$	0,1590	0,0077	2,6738
3	0,0025	0,0993	$-0,0457$	2,5729
4	$-0,0028$	0,1381	$-0,0137$	2,5858
5	0,0018	0,1035	0,0129	2,6526
6	$-0,0049$	0,2795	0,0734	2,5414
7	$-0,0003$	0,1805	0,0405	2,7523
8	0,0041	0,1701	0,0598	2,6876
9	$-0,0004$	0,1203	$-0,0109$	2,5414
10	0,0012	0,1628	$-0,0232$	2,6719
11	$-0,0013$	0,1067	0,0076	2,5865
12	$-0,0012$	0,1255	0,0012	2,6140
13	$-0,0002$	0,3099	0,0837	2,5127
14	$-0,0042$	0,1857	0,0596	2,6940
15	0,0015	0,3145	0,0742	2,4652
16	0,0015	0,3051	0,0954	2,5074
17	$-0,0033$	0,2265	0,0617	2,7075
18	$-0,0002$	0,1499	0,0439	2,6712
19	0,0012	0,2681	0,0112	2,5640
20	0,0003	0,0944	$-0,0155$	2,6194

Tabela 3.2.12 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,03; 0,18)$, $\alpha \approx 21$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0007$	0,2240	0,0418	2,6333
2	$-0,0009$	0,1247	0,0028	2,7239
3	$-0,0002$	0,0722	$-0,0133$	2,5313
4	$-0,0006$	0,1062	$-0,0071$	2,6612
5	$-0,0003$	0,0763	$-0,0017$	2,6271
6	$-0,0003$	0,2682	0,0432	2,5498
7	0,0015	0,1537	0,0111	2,7127
8	0,0001	0,1371	0,0554	2,8032
9	$-0,0015$	0,0909	0,0161	2,4921
10	$-0,0013$	0,1295	0,0078	2,7570
11	0,0009	0,0781	0,0342	2,4790
12	0,0014	0,0935	$-0,0012$	2,5673
13	$-0,0056$	0,3167	0,0845	2,4540
14	$-0,0035$	0,1568	0,0077	2,8246
15	$-0,0008$	0,3118	0,0974	2,4954
16	0,0034	0,2976	0,0252	2,4610
17	0,0009	0,2020	0,0542	2,6139
18	0,0027	0,1183	$-0,0254$	2,6321
19	0,0009	0,2523	0,0429	2,5452
20	$-0,0004$	0,0659	0,0335	2,6129

Tabela 3.2.13 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,35; 0,76)$, $\alpha \approx 2$.

t	média	desvio-padrão	assimetria	curtose
1	0,0042	0,3010	0,0305	2,6397
2	0,0015	0,2468	0,0130	2,7124
3	0,0002	0,2038	0,0005	2,5310
4	$-0,0057$	0,2333	0,0348	2,6488
5	0,0033	0,2115	0,0307	2,7031
6	$-0,0018$	0,3109	$-0,0231$	2,5716
7	$-0,0031$	0,2666	$-0,0299$	2,7923
8	0,0005	0,2589	$-0,0243$	2,7881
9	0,0040	0,2230	$-0,0643$	2,8620
10	0,0032	0,2545	0,0057	2,8327
11	0,0008	0,2089	$-0,0532$	2,5808
12	$-0,0022$	0,2260	$-0,0095$	2,6680
13	$-0,0001$	0,3168	$-0,0257$	2,5342
14	$-0,0053$	0,2665	0,0577	2,7710
15	0,0022	0,3221	$-0,0420$	2,4842
16	0,0043	0,3129	$-0,0468$	2,5058
17	$-0,0041$	0,2833	0,0562	2,6381
18	0,0004	0,2478	$-0,0317$	2,6430
19	0,0015	0,3028	0,0726	2,6246
20	$-0,0007$	0,1985	$-0,0409$	2,7067

Tabela 3.2.14 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,35; 0,76)$, $\alpha \approx 10$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0014$	0,2460	0,0206	2,6169
2	0,0009	0,1547	0,0199	2,7236
3	0,0039	0,1005	$-0,0433$	2,5478
4	$-0,0012$	0,1362	0,0169	2,6351
5	$-0,0010$	0,1057	0,0216	2,6277
6	0,0117	0,2824	$-0,0096$	2,4698
7	$-0,0009$	0,1802	0,0107	2,7779
8	$-0,0014$	0,1672	0,0186	2,6079
9	0,0002	0,1209	0,0007	2,5436
10	0,0019	0,1644	$-0,0041$	2,7139
11	$-0,0023$	0,1067	0,0337	2,4776
12	$-0,0003$	0,1252	0,0327	2,5940
13	$-0,0004$	0,3102	$-0,0657$	2,4550
14	$-0,0013$	0,1870	0,0381	2,8107
15	$-0,0032$	0,3075	$-0,0268$	2,4561
16	$-0,0046$	0,3029	$-0,0051$	2,4228
17	0,0014	0,2219	$-0,0430$	2,7253
18	0,0005	0,1475	0,0398	2,7297
19	$-0,0014$	0,2665	0,0465	2,6182
20	$-0,0003$	0,0926	$-0,0067$	2,6418

Tabela 3.2.15 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,35; 0,76)$, $\alpha \approx 21$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0025$	0,2236	0,0266	2,6403
2	0,0021	0,1259	$-0,0329$	2,7333
3	0,0001	0,0710	$-0,0053$	2,5149
4	0,0011	0,1054	0,0072	2,6051
5	0,0007	0,0795	$-0,0337$	2,6152
6	$-0,0040$	0,2720	$-0,0579$	2,5008
7	$-0,0008$	0,1519	$-0,0009$	2,8051
8	$-0,0012$	0,1388	$-0,0067$	2,6456
9	0,0014	0,0933	$-0,0165$	2,5013
10	$-0,0022$	0,1336	0,0139	2,7835
11	$-0,0006$	0,0760	0,0319	2,5117
12	$-0,0010$	0,0936	$-0,0033$	2,5447
13	$-0,0011$	0,3039	$-0,0500$	2,4297
14	0,0051	0,1584	0,0277	2,7020
15	$-0,0032$	0,3073	$-0,0576$	2,4810
16	0,0050	0,2973	$-0,0400$	2,5005
17	$-0,0034$	0,1977	$-0,0057$	2,7711
18	$-0,0028$	0,1172	0,0171	2,6789
19	0,0070	0,2485	0,0484	2,6857
20	0,0001	0,0642	$-0,0056$	2,6346

Tabela 3.2.16 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,81; 0,95)$, $\alpha \approx 2$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0007$	0,3022	$-0,0877$	2,6328
2	$-0,0027$	0,2473	$-0,0174$	2,7496
3	$-0,0034$	0,2066	$-0,0521$	2,5474
4	0,0013	0,2365	$-0,0381$	2,6926
5	0,0008	0,2047	$-0,0240$	2,7301
6	0,0082	0,3109	$-0,1020$	2,5908
7	0,0023	0,2676	$-0,0220$	2,6977
8	$-0,0020$	0,2594	0,0022	2,7915
9	0,0022	0,2169	0,0035	2,6208
10	$-0,0010$	0,2593	$-0,0379$	2,7527
11	0,0002	0,2112	$-0,0381$	2,6347
12	0,0068	0,2271	$-0,0093$	2,7051
13	$-0,0005$	0,3198	$-0,0520$	2,4495
14	0,0029	0,2691	$-0,0741$	2,7565
15	$-0,0031$	0,3231	$-0,0652$	2,5601
16	0,0028	0,3198	$-0,0871$	2,4978
17	0,0020	0,2935	$-0,0080$	2,6536
18	$-0,0040$	0,2425	0,0123	2,6723
19	0,0082	0,3079	$-0,1014$	2,6311
20	$-0,0015$	0,2039	0,0282	2,6570

Tabela 3.2.17 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,81; 0,95)$, $\alpha \approx 10$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0010$	0,2487	$-0,0240$	2,6367
2	$-0,0011$	0,1580	0,0402	2,7014
3	0,0024	0,1007	$-0,0078$	2,5464
4	$-0,0022$	0,1367	0,0124	2,6673
5	0,0002	0,1056	$-0,0094$	2,7064
6	0,0014	0,2826	$-0,0603$	2,5605
7	0,0003	0,1803	$-0,0227$	2,6904
8	0,0024	0,1701	0,0139	2,8190
9	0,0017	0,1203	0,0267	2,5359
10	$-0,0012$	0,1660	$-0,0472$	2,7021
11	0,0012	0,1077	0,0231	2,5068
12	$-0,0015$	0,1229	0,0595	2,6134
13	$-0,0085$	0,3130	$-0,0213$	2,4885
14	$-0,0001$	0,1883	0,0072	2,7983
15	$-0,0071$	0,3131	$-0,0240$	2,4608
16	$-0,0048$	0,3053	$-0,0813$	2,5294
17	$-0,0034$	0,2236	$-0,0296$	2,7435
18	0,0009	0,1508	$-0,0411$	2,7153
19	0,0002	0,2702	$-0,0181$	2,6018
20	$-0,0020$	0,0911	$-0,0236$	2,6280

Tabela 3.2.18 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\mu \in (0,81; 0,95)$, $\alpha \approx 21$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0063$	0,2250	$-0,0004$	2,5599
2	$-0,0001$	0,1278	0,0207	2,6921
3	0,0008	0,0723	0,0101	2,4824
4	$-0,0010$	0,1071	$-0,0306$	2,6159
5	$-0,0005$	0,0756	0,0295	2,5376
6	0,0065	0,2723	-0,0376	2,5092
7	$-0,0017$	0,1503	0,0227	2,7250
8	0,0018	0,1390	$-0,0439$	2,7866
9	0,0003	0,0917	$-0,0125$	2,5139
10	$-0,0020$	0,1340	0,0513	2,6786
11	0,0022	0,0773	$-0,0155$	2,5386
12	$-0,0003$	0,0931	0,0096	2,6246
13	$-0,0073$	0,3080	$-0,0381$	2,4886
14	$-0,0014$	0,1604	$-0,0334$	2,8041
15	$-0,0023$	0,3107	$-0,0679$	2,4449
16	$-0,0075$	0,2989	$-0,0556$	2,4912
17	0,0032	0,1992	$-0,0207$	2,6208
18	$-0,0011$	0,1164	$-0,0381$	2,7380
19	0,0031	0,2523	$-0,0761$	2,5583
20	$-0,0006$	0,0650	0,0055	2,5364

3.2.3 Avaliação numérica: modelo simplex não linear com precisão variável

Avaliamos agora as propriedades empíricas do resíduo r_t^{pp} para o modelo simplex não linear com dispersão variável. Consideramos que $y_t \sim \mathcal{S}^{-1}(\mu_t, \lambda_t)$ com

$$\begin{aligned} g(\mu_t) &= \beta_1 + \exp(\beta_2 x_t), \\ h(\lambda_t) &= \gamma_1 + \gamma_2 z_t, \end{aligned} \tag{3.6}$$

em que $t = 1, \ldots, 20$, $g(\cdot)$ e $h(\cdot)$ são as funções de ligação logito e logarítmica, respectivamente, e as variáveis explicativas x_t e z_t são conhecidas para cada observação. Destacamos que as covariáveis x_t e z_t foram geradas através de realizações independentes da distribuição uniforme $\mathcal{U}(0, 1)$ e mantidas fixas para as 5000 réplicas de Monte Carlo. Admitimos ainda que $\beta_1 = -1,5$ e $\beta_2 = 1,2$, o que conduz a $\mu \in (0,39; 0,80)$. Além disso, consideramos $\gamma_1 = 2,2$ e variamos o valor de γ_2 para obter diferentes graus de heterogeneidade. Especificamente, fixamos $\gamma_2 = 1$, $\gamma_2 = 2,5$ e $\gamma_2 = 3,2$, o que produz $\alpha \approx 2$, $\alpha \approx 10$ e $\alpha \approx 21$, respectivamente.

Os resultados da simulação são apresentados nas Tabelas 3.2.19–3.2.21. Para os diferentes graus de heterogeneidade considerados, nota-se, de modo geral, que r_t^{pp} apresenta média aproximadamente zero e desvio-padrão inferior ao da distribuição normal padrão. Além disso, a distribuição de r_t^{pp} tem assimetria desprezível e uma curtose próxima a da distribuição $\mathcal{N}(0, 1)$. Dessa forma, os resultados sugerem, que o resíduo proposto apresenta propriedades similares para o modelo simplex com precisão variável e o modelo apresentado em (3.6). No entanto, nota-se que quando comparado aos resultados obtidos para o modelo simplex com precisão constante a distribuição do resíduo tende a ser menos assimétrica nos dois últimos casos.

Tabela 3.2.19 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\alpha \approx 2$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0040$	0,2913	$-0,0349$	2,6260
2	0,0002	0,2435	$-0,0291$	2,6910
3	$-0,0054$	0,2128	$-0,0049$	2,4715
4	0,0048	0,2351	$-0,0286$	2,6704
5	0,0010	0,1936	$-0,0100$	2,8941
6	0,0013	0,3138	$-0,0489$	2,5099
7	$-0,0010$	0,2580	0,0130	2,7540
8	0,0049	0,2549	0,0224	2,7892
9	$-0,0018$	0,2482	0,0371	2,6995
10	0,0001	0,2511	$-0,0558$	2,7155
11	$-0,0004$	0,2188	$-0,0029$	2,4996
12	$-0,0010$	0,2284	0,0037	2,6191
13	0,0019	0,3214	0,0240	2,4832
14	$-0,0005$	0,2687	0,0056	2,7732
15	$-0,0013$	0,3229	0,0391	2,4656
16	0,0013	0,3210	0,0190	2,5115
17	$-0,0014$	0,2925	$-0,0025$	2,7070
18	$-0,0004$	0,2517	0,0028	2,6813
19	0,0025	0,3042	$-0,0439$	2,5952
20	0,0022	0,1860	0,0109	2,7385

Tabela 3.2.20 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\alpha \approx 10$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0012$	0,2429	$-0,0126$	2,5787
2	$-0,0010$	0,1538	0,0024	2,7653
3	0,0020	0,1051	$-0,0535$	2,5126
4	0,0018	0,1360	0,0137	2,5695
5	$-0,0001$	0,0901	0,0223	2,7657
6	0,0033	0,2852	$-0,0240$	2,5167
7	$-0,0002$	0,1779	$-0,0361$	2,7603
8	$-0,0021$	0,1699	0,0082	2,7744
9	0,0012	0,1365	0,0478	2,5567
10	$-0,0005$	0,1634	$-0,0028$	2,6643
11	$-0,0003$	0,1092	$-0,0073$	2,5098
12	$-0,0002$	0,1292	0,0259	2,4728
13	0,0008	0,3084	$-0,0350$	2,4846
14	0,0028	0,1857	$-0,0128$	2,7560
15	0,0026	0,3126	$-0,0184$	2,4585
16	$-0,0048$	0,3083	0,0023	2,4662
17	$-0,0034$	0,2275	$-0,0150$	2,7338
18	$-0,0045$	0,1531	0,0138	2,6138
19	$-0,0042$	0,2602	$-0,0180$	2,5610
20	$-0,0007$	0,0847	0,0240	2,6388

Tabela 3.2.21 – Média, desvio-padrão, assimetria e curtose do resíduo ponderado padronizado. Cenário $\alpha \approx 21$.

t	média	desvio-padrão	assimetria	curtose
1	$-0,0023$	0,2239	$-0,0469$	2,5771
2	$-0,0011$	0,1240	$-0,0295$	2,6804
3	0,0002	0,0766	0,0178	2,4773
4	0,0006	0,1046	$-0,0147$	2,7027
5	0,0004	0,0642	$-0,0226$	2,7184
6	$-0,0011$	0,2794	$-0,0263$	2,5239
7	0,0024	0,1459	0,0078	2,7763
8	$-0,0039$	0,1384	0,0003	2,6787
9	$-0,0016$	0,1033	0,0359	2,5183
10	$-0,0001$	0,1333	0,0166	2,7818
11	0,0005	0,0809	$-0,0063$	2,4475
12	$-0,0020$	0,0971	0,0286	2,5371
13	$-0,0114$	0,3012	$-0,0195$	2,5805
14	0,0017	0,1597	$-0,0266$	2,6820
15	0,0042	0,3074	$-0,0150$	2,4954
16	0,0015	0,2954	0,0112	2,5051
17	$-0,0046$	0,1983	0,0340	2,6136
18	0,0017	0,1211	$-0,0370$	2,6969
19	$-0,0097$	0,2512	$-0,0247$	2,5726
20	0,0012	0,0581	$-0,0277$	2,4814

3.3 Influência local

Para um determinado conjunto de dados observados, seja $\ell(\theta)$ o logaritmo da função de verossimilhança correspondente ao modelo postulado, em que θ é um vetor $(k+q) \times 1$ de parâmetros desconhecidos. Considere ainda um vetor de perturbação δ, em geral $n \times 1$, restrito a algum conjunto aberto $\mathcal{D}$ de $\mathbb{R}^n$, introduzido no modelo postulado. Geralmente, δ pode refletir qualquer esquema de perturbação bem definido e não se restringir apenas ao esquema de ponderação de casos. Por exemplo, δ pode ser usado para induzir pequenas modificações nas covariáveis em modelos lineares generalizados, ou para perturbar a matriz de covariâncias dos erros em modelos normais lineares (Cook, 1986).

Seja $\ell_\delta(\theta)$ o logaritmo da função de verossimilhança do modelo perturbado para um dado $\delta \in \mathcal{D}$. Assumimos que existe um vetor de não perturbação δ_0 em $\mathcal{D}$, de tal forma que $\ell(\theta) = \ell(\theta|\delta_0)$ para todo θ. Finalmente, sejam $\widehat{\theta}$ e $\widehat{\theta}_\delta$ os estimadores de máxima verossimilhança de θ para o modelo postulado e para o modelo perturbado, respectivamente. O deslocamento pela verossimilhança ("likelihood displacement"), que neste caso mais geral é expresso por

$$LD_\delta = 2\{\ell(\widehat{\theta} - \ell(\widehat{\theta}_\delta))\}, \tag{3.7}$$

pode ser utilizado como medida para avaliar a influência sobre a estimativa de θ ao se variar δ através de $\mathcal{D}$. No entanto, avaliar o comportamento de LD_δ para todo $\delta \in \mathcal{D}$ pode ser inviável. Neste sentido, Cook (1986) propôs avaliar o comportamento local de LD_δ em torno do vetor de não perturbação δ_0, tal que $LD_{\delta_0} = 0$. Uma vez que qualquer valor δ em uma vizinhança

de δ_0 representa um incremento em LD_{δ_0}, este procedimento descreve a sensibilidade de $\ell(\widehat{\theta})$ quando pequenas modificações são introduzidas em $\ell(\theta)$. Cook (1986) sugere então avaliar a curvatura do gráfico LD_{δ_0+aI}, em que $a \in \mathbb{R}$ e I é um vetor de norma unitária, ou seja, $\|I\| = 1$; esse gráfico é denominado de linha projetada. O autor sugere inspecionar a direção I_{max} correspondente à linha projetada de maior curvatura C_{max}.

Cook (1986) mostra através de conceitos de geometria diferencial que ao utilizarmos como métrica (3.7), a curvatura normal na direção de algum vetor I pode ser expressa por

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

em que $-\ddot{L}_{\theta\theta}$ é a matriz de informação observada e Δ é uma matriz $(k+q) \times n$ dada por $\Delta = \partial^2 \ell_\delta(\theta)/\partial\theta\partial\delta^\top$ que depende do esquema de perturbação, avaliados em $\theta = \widehat{\theta}$ e $\delta = \delta_0$. Dessa forma, C_{max} é o maior autovalor da matriz $\Delta^\top(-\ddot{L}_{\theta\theta})^{-1}\Delta$ e I_{max} é o autovetor de norma unitária corresponde a C_{max}.

Também é possível calcular a influência local para uma partição do vetor de parâmetros. Por exemplo, suponha que possamos particionar o vetor de parâmetros como $\theta = (\theta_1^\top, \theta_2^\top)^\top$, tal que

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

em que $\ddot{L}_{\theta_1\theta_1} = \partial\ell(\theta)/\partial\theta_1\partial\theta_1^\top$, $\ddot{L}_{\theta_1\theta_2} = \partial\ell(\theta)/\partial\theta_1\partial\theta_2^\top$, $\ddot{L}_{\theta_2\theta_1} = \partial\ell(\theta)/\partial\theta_2\partial\theta_1^\top$ e $\ddot{L}_{\theta_2\theta_2} = \partial\ell(\theta)/\partial\theta_2\partial\theta_2^\top$. Dessa forma, se o interesse é avaliar a influência local apenas para θ_1, Cook (1986) mostra que a curvatura normal na direção do vetor I é dada por

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

em que

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

e $I_{max;\theta_1}$ é o autovetor de norma igual 1 correspondente ao maior autovalor $C_{max;\theta_1}$ da matriz $-\Delta^\top(\ddot{L}_{\theta\theta}^{-1} - \ddot{L}_{22})\Delta$. O procedimento para avaliar a influência local apenas para θ_2 é análogo.

Cook (1986) propõe avaliar os componentes de I_{max}, uma vez que isto pode revelar observações conjuntamente influentes. Entretanto, podem existir casos individualmente influentes e que podem não ser destacados através da análise dos elementos de I_{max}. Nesse sentido, Lesaffre e Verbeke (1998) sugerem avaliar a curvatura na direção do t-ésimo indivíduo, ou seja, na direção do vetor de norma unitária I em que o t-ésimo componente assume o valor 1 e os demais o valor zero. Neste caso, a curvatura normal é expressa por

$$C_t = 2|\Delta_t^\top \ddot{L}_{\theta\theta}^{-1} \Delta_t|,$$

em que Δ_t é a t-ésima coluna da matriz Δ. Os autores denotam tal curvatura por influência local total do t-ésimo indivíduo. É possível ainda calcular a influência local total do t-ésimo indivíduo na estimação de parte do vetor θ. Se o interesse recai em θ_1, temos que

$$C_{t;\theta_1} = 2|\Delta_t^\top(\ddot{L}_{\theta\theta}^{-1} - \ddot{L}_{22})\Delta_t|.$$

Segundo Lesafre e Verbeke (1998) os componentes de I_{max} e as medidas de influência local total C_t, $t = 1, \ldots, n$ contêm informações diferentes sobre a influência das observações no ajuste do modelo. Sejam $C_{max}/2 \equiv \xi_1 \geq \xi_2 \cdots \geq \xi_{(k+q)} > 0$ autovalores não nulos da matriz $-\Delta^\top \ddot{L}_{\theta\theta}^{-1}\Delta$, e sejam $I_{max} \equiv \vartheta_1, \ldots, \vartheta_n$ os correspondentes autovetores ortogonais de norma unitária. Os autores mostram que

$$C_t = 2\sum_{p=1}^{k+q} \xi_p \vartheta_{pt}^2 \tag{3.8}$$

em que ϑ_{pt} é o t-ésimo componente do vetor ϑ_p. Assim, a partir de (3.8) observa-se que casos individuais podem ter grandes valores de C_t, sem ter o t-ésimo componente expressivo na direção I_{max} de máxima curvatura. Isso ocorrerá para indivíduos com valor expressivo em qualquer autovetor ϑ_p, $p \neq 1$, correspondente a um autovalor relativamente grande (por exemplo, ξ_2, ξ_3). Gráficos dos componentes de I_{max} contra os índices das observações podem ser utilizados para destacar observações conjuntamente influentes, enquanto gráficos C_t contra os índices das observações podem sugerir casos individualmente influentes. Cook (1986) destaca ainda que gráficos de I_{max} contra os índices das observações ou contra covariáveis podem revelar tendências nos dados (por exemplo, não linearidade, heteroscedasticidade, etc.). Nos gráficos de influência local total C_t, Lesafre e Verbeke (1998) sugerem como ponto de corte duas vezes o valor médio desta medida, ou seja, observações com C_t maior que $2\bar{C}$, com $\bar{C} = \sum_{t=1}^{n} C_t/n$ são consideradas como influentes.

3.4 Esquemas de perturbação

Para aplicação do método de influência local proposto por Cook (1986) é necessário inicialmente obter a matriz de informação observada $-\ddot{L}_{\theta\theta}$. Considerando o modelo definido em (2.4), temos que

$$-\ddot{L}_{\theta\theta} = \begin{pmatrix} -\ddot{L}_{\beta\beta} & -\ddot{L}_{\beta\gamma} \\ -\ddot{L}_{\gamma\beta} & -\ddot{L}_{\gamma\gamma} \end{pmatrix}, \tag{3.9}$$

em que $-\ddot{L}_{\beta\beta} = \tilde{X}^\top \Lambda Q \tilde{X} - [\ b_\beta^\top\ ][\ \tilde{X}_\beta\ ]$, $-\ddot{L}_{\beta\gamma} = (-\ddot{L}_{\gamma\beta})^\top = -\tilde{X}^\top THU\mathcal{E}\tilde{Z}$ e $-\ddot{L}_{\gamma\gamma} = \tilde{Z}^\top \mathcal{V}\tilde{Z} - [\ b_\gamma^\top\ ][\ \tilde{Z}_\gamma\ ]$, $Q = \text{diag}\{q_1, \ldots, q_n\}$, com

$$q_t = \left\{ u_t - (y_t - \mu_t)u_t' + (y_t - \mu_t)u_t \frac{g''(\mu_t)}{g'(\mu_t)} \right\} \frac{1}{\{g'(\mu_t)\}^2}, \tag{3.10}$$

sendo

$$u'_t = -\left\{\frac{2(y_t-\mu_t)u_t}{\mu_t(1-\mu_t)} + \frac{3(1-2\mu_t)}{\mu_t^4(1-\mu_t)^4} + \frac{(1-2\mu_t)d(y_t;\mu_t)}{\mu_t^2(1-\mu_t)^2}\right\},$$

$\mathcal{V} = \text{diag}\{\nu_1, \ldots, \nu_n\}$, com

$$\nu_t = d_t + a_t\frac{h''(\lambda_t)}{\{h'(\lambda_t)^3}, \tag{3.11}$$

e

$$\mathcal{E} = \text{diag}\{(y_1 - \widehat{\mu}_1), \ldots, (y_n - \widehat{\mu}_n)\} \tag{3.12}$$

As matrizes Λ, T, H e U estão definidas em (2.11), (2.12), (2.15) e (2.9), respectivamente, enquanto as quantidades a_t e d_t são definidas em (2.14) e (2.18), respectivamente. Além disso, $b_\beta = \Lambda TU(y-\mu)$, $\tilde{X}_\beta = (\tilde{X}_t)$ é um array de dimensão $n \times k \times k$, sendo $\tilde{X}_t$ uma matriz $k \times k$ com elementos dados por $\partial^2\eta_t/\partial\beta_i\partial\beta_p$, $b_\gamma = Ha$ e $\tilde{Z}_\gamma = (\tilde{Z}_t)$ é um array de dimensão $n \times q \times q$, sendo $\tilde{Z}_t$ uma matriz $q \times q$ com elementos dados por $\partial^2\zeta_t/\partial\gamma_j\partial\gamma_l$. Finalmente, $[\,\cdot\,][\,\cdot\,]$ representa o produto colchete de uma matriz por um array como definido por Wei (1998, p. 188). Ressaltamos que nas Seções 3.4.1–3.4.5 as quantidades assinaladas com " $\widehat{\ }$ " são avaliadas em $(\widehat{\beta}^\top, \widehat{\gamma}^\top)^\top$. A seguir, apresentamos diferentes esquemas de perturbação, a saber: ponderação de casos, perturbação da variável resposta, perturbação de covariáveis da média, perturbação de covariáveis da precisão e perturbação simultânea de covariáveis.

3.4.1 Ponderação de casos

Considere o esquema de perturbação dado por

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

em que $0 \leq \delta_t \leq 1$. Para este esquema, temos que $\Delta_t = \partial\ell_t(\theta)/\partial\theta$, $t = 1, \ldots, n$ e $\delta_0 = (1, 1, \ldots, 1)^\top$. Assim, segue da Seção B.1 do Apêndice B que

$$\Delta = \begin{pmatrix} \widehat{\tilde{X}}^\top\widehat{\Lambda}\widehat{U}\widehat{T}\mathcal{E} \\ \widehat{\tilde{Z}}^\top\widehat{H}\mathcal{A} \end{pmatrix}, \tag{3.13}$$

em que

$$\mathcal{A} = \text{diag}\{a_1, \ldots, a_n\}, \tag{3.14}$$

e $\mathcal{E}$ está definida em (3.12).

O esquema de ponderação de casos tem sido um dos mais utilizados para acessar a influência, podendo ser interpretado no caso de modelos normais lineares como uma perturbação na variância do t-ésimo caso (Thomas e Cook, 1989). Baseado neste tipo de perturbação, Cook (1986) analisa o comportamento de I_{max} em modelos de regressão linear simples sob violação da suposição de homoscedasticidade. O autor conclui que gráficos de I_{max} versus valores de covariáveis podem sugerir quais covariáveis devem ser utilizadas para modelar a variância.

3.4.2 Perturbação da variável resposta

Considere uma perturbação aditiva na resposta y_t, tal que

$$y_t(\delta) = y_t + \delta_t s(y_t), \tag{3.15}$$

em que $s(y_t)$, $t = 1, \ldots, n$ é um fator de escala utilizado para padronizar os componentes de δ em situações em que cada y_t apresenta variância diferente. Para o modelo simplex não linear, consideramos como fator de escala a raiz quadrada da estimativa da função de variância dada por $V(\mu_t) = \mu_t^3(1-\mu_t)^3$. Esta é uma nova proposta para o fator de escala de padronização, dado que em geral utiliza-se o desvio-padrão de y. Vale destacar, que a nova proposta apresenta um menor custo computacional, uma vez que a variância de uma variável aleatória com distribuição simplex depende da função gama incompleta. Para este esquema, temos que o vetor de não perturbação é dado por $\delta_0 = (0, 0, \ldots, 0)^\top$ e da Seção B.2 do Apêndice B segue que

$$\Delta = \begin{pmatrix} \widehat{\tilde{X}}^\top \widehat{\Lambda} \widehat{T} B S_y \\ \widehat{\tilde{Z}}^\top \widehat{H} C S_y \end{pmatrix}, \tag{3.16}$$

em que $B = \mathrm{diag}\{\widehat{b}_1, \ldots, \widehat{b}_n\}$, com

$$b_t = \frac{1}{y_t(1-y_t)} \left\{ \frac{2}{y_t(1-\mu_t)^3} + \frac{(1-3\mu_t)}{\mu_t^2(1-\mu_t)^3} - \frac{1}{2} \frac{\partial d(y_t;\mu_t)}{\partial \mu_t} \right\},$$

$v(y_t) = \mathrm{diag}\{v(y_1), \ldots, v(y_n)\}$ e $C = \mathrm{diag}\{\widehat{c}_1, \ldots, \widehat{c}_n\}$, com

$$c_t = -\frac{1}{2y_t(1-y_t)} \left\{ d(y_t;\mu_t) + \frac{2(y_t-\mu_t)}{y_t\mu_t(1-\mu_t)^2} \right\}. \tag{3.17}$$

A perturbação da variável resposta está fortemente relacionada com o conceito de alavanca generalizada. Espinheira et al. (2008b) apresenta uma comparação formal entre medidas de influência local baseadas na perturbação aditiva da resposta e alavanca para o modelo de regressão beta. Dessa forma, os componentes de I_{max} e as medidas de influência local total C_t podem ser utilizadas para destacar observações que exercem influência desproporcional sobre o próprio valor ajustado.

3.4.3 Perturbação de covariáveis da média ($x_p^\top$)

Thomas e Cook (1990) sugerem modificar a p-ésima covariável da estrutura de regressão da média, x_p, $p = 2, \ldots, k$, adicionando um vetor de perturbação δ, de forma que

$$x_{tp}(\delta) = x_{tp} + \delta_t s_{x_p},$$

em que s_{x_p} é um fator de escala, que pode ser, por exemplo, o desvio-padrão da covariável modificada. Neste caso, temos que

$$\eta_t(\delta) = f_1(x_{t1}, \ldots, x_{tp}(\delta), \ldots, x_{tk}; \beta) \tag{3.18}$$

e $\mu_t(\delta)$ é tal que $g(\mu_t(\delta)) = \eta_t(\delta)$. Como as covariáveis que determinam a média não interferem na precisão, $\lambda_t(\delta) = \lambda_t$. Para este tipo de perturbação $\delta_0 = (0, 0, \ldots, 0)^\top$ e da Seção B.3 do Apêndice B, temos que

$$\Delta = \begin{pmatrix} -\widehat{\tilde{X}}^\top \widehat{\Lambda}\widehat{Q}\widehat{\tilde{X}}_\delta + [\ \widehat{b}_\beta^\top\ ][\ \widehat{\tilde{X}}_{\beta\delta}\ ] \\ \widehat{\tilde{Z}}^\top \widehat{H}\widehat{T}\widehat{U}\mathcal{E}\widehat{\tilde{X}}_\delta \end{pmatrix}, \tag{3.19}$$

em que $\tilde{X}_{\beta\delta} = (\tilde{X}_{\delta t})$ é um array $n \times k \times n$, $\tilde{X}_{\delta t}$ é uma matriz $k \times n$ com elementos $\partial^2\eta_t(\delta)/\partial\beta_i\partial\delta_t$, $\tilde{X}_\delta = \partial\eta(\delta)/\partial\delta$ e $b_\beta = \Lambda T U(y - \mu)$.

3.4.4 Perturbação de covariáveis da precisão ($z_{p'}^\top$)

No caso em que a média da variável resposta e o parâmetro de precisão são modelados simultaneamente, pode haver o interesse em checar a presença de observações que exercem influência desproporcional sobre a estimativa do vetor γ. Nesse sentido, considere uma perturbação aditiva na p'-ésima covariável que determina a precisão, de forma que

$$z_{tp'}(\delta) = z_{tp'} + \delta_t s_{z_{p'}},$$

em que $s_{z_{p'}}$ é o desvio-padrão da covariável modificada. Neste caso, temos que

$$\zeta_t(\delta) = f_2(z_{t1}, \ldots, z_{tp'}(\delta), \ldots, z_{tq}; \gamma) \tag{3.20}$$

e $\lambda_t(\delta)$ é tal que $h(\lambda_t(\delta)) = \zeta_t(\delta)$. Uma vez que as covariáveis que determinam a precisão não interferem na média da variável resposta, segue que $\mu_t(\delta) = \mu_t$. Para este esquema de perturbação $\delta_0 = (0, 0, \ldots, 0)^\top$ e da Seção B.4 do Apêndice B temos que

$$\Delta = \begin{pmatrix} \widehat{\tilde{X}}^\top \widehat{T}\widehat{H}\widehat{U}\mathcal{E}\widehat{\tilde{Z}}_\delta \\ -\widehat{\tilde{Z}}^\top \mathcal{V}\widehat{\tilde{Z}}_\delta + [\ \widehat{b}_\gamma^\top\ ][\ \widehat{\tilde{Z}}_{\gamma\delta}\ ], \end{pmatrix}, \tag{3.21}$$

em que $\tilde{Z}_{\gamma\delta} = (\tilde{Z}_{\delta t})$ é um array $n \times q \times n$, $\tilde{Z}_{\delta t}$ é uma matriz $q \times n$ com elementos $\partial^2\zeta_t(\delta)/\partial\gamma_j\partial\delta_t$, $\tilde{Z}_\delta = \partial\zeta(\delta)/\partial\delta$ e $b_\gamma = Ha$.

3.4.5 Perturbação simultânea de covariáveis ($x_p^\top, z_{p'}^\top$)

Considere agora, o caso em que a p-ésima covariável da média e a p'-ésima covariável da precisão são perturbadas simultaneamente através de um vetor δ, de modo que

$$\begin{aligned} x_{tp}(\delta) &= x_{tp} + \delta_t s_{x_p}, \\ z_{tp'}(\delta) &= z_{tp'} + \delta_t s_{z_{p'}}, \end{aligned}$$

em que s_{x_p} e $s_{z_{p'}}$ são os desvios padrão de x_p e $z_{p'}$, respectivamente. Neste caso,

$$\begin{aligned} \eta_t(\delta) &= f_1(x_{t1}, \ldots, x_{tp}(\delta), \ldots, x_{tk}; \beta), \\ \zeta_t(\delta) &= f_2(z_{t1}, \ldots, z_{tp'}(\delta), \ldots, z_{tq}; \gamma), \end{aligned} \tag{3.22}$$

$\mu_t(\delta)$ é tal que $g(\mu_t(\delta)) = \eta_t(\delta)$ e $\lambda_t(\delta)$ é tal que $h(\lambda_t(\delta)) = \zeta_t(\delta)$. Para este esquema, o vetor de não perturbação é $\delta_0 = (0, 0, \ldots, 0)^\top$ e da Seção B.5 do Apêndice B obtemos que

$$\Delta = \begin{pmatrix} \widehat{\tilde{X}}^\top \widehat{T}\widehat{H}\widehat{U}\mathcal{E}\hat{\tilde{Z}}_\delta - \widehat{\tilde{X}}^\top \widehat{\Lambda}\widehat{Q}\widehat{\tilde{X}}_\delta + [\ \widehat{b}_\beta^\top\ ][\ \widehat{\tilde{X}}_{\beta\delta}\ ] \\ \hat{\tilde{Z}}^\top \widehat{T}\widehat{H}\widehat{U}\mathcal{E}\widehat{\tilde{X}}_\delta - \hat{\tilde{Z}}^\top \mathcal{V}\hat{\tilde{Z}}_\delta + [\ \widehat{b}_\gamma^\top\ ][\ \hat{\tilde{Z}}_{\gamma\delta}\ ] \end{pmatrix}, \tag{3.23}$$

sendo que $\tilde{X}_{\beta\delta}$ e $\tilde{Z}_{\gamma\delta}$ foram definidos nas seções anteriores. Os componentes da matriz diagonal Q estão definidos em (3.10), e as matrizes $\mathcal{E}$ e $\mathcal{V}$ estão definidas em (3.12) e (3.11), respectivamente.

3.5 Alavanca generalizada

Em modelos normais lineares, observações com alta alavancagem podem ser acessadas através dos elementos diagonais da matriz de projeção $H = X(X^\top X)^{-1}X^\top$, em que X é uma matriz $n \times k$ de covariáveis. Baseados na essência da definição de pontos de alavanca, que são caracterizados por exercerem influência sobre o próprio valor ajustado $\widehat{y}$, Wei et al. (1998) apresentaram uma generalização desse conceito para métodos de estimação e modelos de regressão mais gerais. Segundo os autores, a matriz de alavanca generalizada $\partial\widehat{y}/\partial y^\top$ pode ser obtida da forma geral para o método de máxima verossimilhança como

$$GL(\theta) = \frac{\partial \widehat{y}}{\partial y^\top} = \left\{ D_\theta(-\ddot{L}_{\theta\theta})^{-1}\ddot{L}_{\theta y} \right\}\Big|_{\theta=\widehat{\theta}}, \tag{3.24}$$

em que $D_\theta = \partial\mu/\partial\theta^\top$, $\ddot{L}_{\theta\theta} = -\partial^2\ell(\theta)/\partial\theta\partial\theta^\top$ e $\ddot{L}_{\theta y} = \partial^2\ell(\theta)/\partial\theta\partial y^\top$. A expressão (3.24) generaliza a definição apresentada em St. Laurent e Cook (1992).

Para o modelo definido em (2.4), temos que

$$D_\theta = \begin{pmatrix} T\tilde{X} & 0 \end{pmatrix}. \tag{3.25}$$

Além disso, a matriz $\ddot{L}_{\theta y}$ pode ser expressa por

$$\ddot{L}_{\theta y} = \begin{pmatrix} \ddot{L}_{\beta y} \\ \ddot{L}_{\gamma y} \end{pmatrix},$$

em que $\ddot{L}_{\beta y} = \partial^2\ell(\theta)/\partial\beta\partial y^\top$ e $\ddot{L}_{\gamma y} = \partial^2\ell(\theta)/\partial\gamma\partial y^\top$. Desse modo, o (i,t)-ésimo elemento de $\ddot{L}_{\beta y}$ é dado por

$$\frac{\partial^2\ell(\beta,\gamma)}{\partial\beta_i\partial y_t} = \sum_{t=1}^{n} \lambda_t m_t \frac{1}{g'(\mu_t)}\frac{\partial\eta_t}{\partial\beta_i},$$

em que

$$m_t = u_t + \frac{(y_t - \mu_t)}{y_t(1-y_t)\mu_t(1-\mu_t)}\left\{ d(y_t;\mu_t) + \frac{2(y_t-\mu_t)}{y_t\mu_t(1-\mu_t)^2} \right\}, \tag{3.26}$$

com u_t e $d(y_t;\mu_t)$ definidos em (2.9) e (2.3), respectivamente.

Analogamente, o (j,t)-ésimo elemento de $\ddot{L}_{\gamma y}$ é dado por

$$\frac{\partial^2 \ell(\beta,\gamma)}{\partial \gamma_j \partial y_t} = \sum_{t=1}^{n} c_t \frac{1}{h'(\lambda_t)} \frac{\partial \zeta_t}{\partial \gamma_j}.$$

Em notação matricial temos

$$\ddot{L}_{\theta y} = \begin{pmatrix} \tilde{X}^\top \Lambda T M \\ \tilde{Z}^\top H C \end{pmatrix}, \tag{3.27}$$

em que $M = \text{diag}\{m_1, \ldots, m_n\}$ e $C = \text{diag}\{c_1, \ldots, c_n\}$, com m_t e c_t dados em (3.26) e (3.17), respectivamente.

Substituindo (3.9), (3.25) e (3.27) em (3.24), temos

$$\begin{aligned} GL(\theta) &= [T\tilde{X} \;\; 0] \begin{pmatrix} -\ddot{L}_{\beta\beta} & -\ddot{L}_{\beta\gamma} \\ -\ddot{L}_{\gamma\beta} & -\ddot{L}_{\gamma\gamma} \end{pmatrix}^{-1} \begin{pmatrix} \tilde{X}^\top \Lambda T M \\ \tilde{Z}^\top H C \end{pmatrix} \\ &= [T\tilde{X} \;\; 0] \begin{pmatrix} -\ddot{L}_{\beta\beta}^{-1} + F E^{-1} F^\top & -F E^{-1} \\ -E^{-1} F^\top & E^{-1} \end{pmatrix} \begin{pmatrix} \tilde{X}^\top \Lambda T M \\ \tilde{Z}^\top H C \end{pmatrix}, \end{aligned}$$

em que $F = \ddot{L}_{\beta\beta}^{-1} \ddot{L}_{\beta\gamma}$, $F^\top = \ddot{L}_{\gamma\beta} \ddot{L}_{\beta\beta}^{-1}$, $E = -\ddot{L}_{\gamma\gamma} + \ddot{L}_{\gamma\beta} \ddot{L}_{\beta\beta}^{-1} \ddot{L}_{\beta\gamma}$, e $\ddot{L}_{\beta\beta}$, $\ddot{L}_{\beta\gamma} = (\ddot{L}_{\gamma\beta})^\top$ e $\ddot{L}_{\gamma\gamma}$ são componentes da matriz de informação observada dada em (3.9). Logo,

$$\begin{aligned} GL(\theta) &= \begin{pmatrix} T\tilde{X}(-\ddot{L}_{\beta\beta}^{-1} + F E^{-1} F^\top) & -T\tilde{X} F E^{-1} \end{pmatrix} \begin{pmatrix} \tilde{X}^\top \Lambda T M \\ \tilde{Z}^\top H C \end{pmatrix} \\ &= T\tilde{X}(-\ddot{L}_{\beta\beta}^{-1} + F E^{-1} F^\top) \tilde{X}^\top \Lambda T M - T\tilde{X} F E^{-1} \tilde{Z}^\top H C. \end{aligned}$$

Gráficos dos elementos diagonais de $GL(\widehat{\theta})$ contra a ordem das observações ou contra os valores preditos são sugeridos para identificar possíveis pontos de alavanca.

CAPÍTULO 4

APLICAÇÕES

4.1 Aplicação I: dados simulados

Nesta seção, simulamos um exemplo com uma única amostra, para investigar o comportamento dos gráficos de diagnóstico sob a presença de observações que exercem efeitos desproporcionais sobre o ajuste do modelo e sob a presença de erro na especificação da componente sistemática. Inicialmente, geramos $n-1$ observações independentes de uma covariável $x_t \sim \mathcal{U}(0,1)$ e adicionamos um n-ésimo valor à amostra, $x_n = 1,2$, induzindo a ocorrência de uma observação alavanca. Considerando o modelo de regressão simplex não linear $\mathcal{S}^{-1}(\mu_t, \lambda)$ com

$$\log[\mu_t/(1-\mu_t)] = \beta_1 + \exp(\beta_2 x_t), \tag{4.1}$$

$t = 1, \ldots, n$, $\beta_1 = -1,8$, $\beta_2 = 1,5$ e $\lambda = 12,18$, geramos $n = 30$ valores da variável resposta y, formando o conjunto de dados que denominados de dados corretos.

Modificamos o conjunto de dados corretos, simulando um erro de entrada dos dados, de forma que o valor da variável resposta correspondente a $x_{30} = 1,2$ fosse uma réplica de outro valor de y presente nos dados. Especificamente, assumimos que $y_{30} = \min\{y_1, \ldots, y_{29}\}$, induzindo a ocorrência de um ponto de alavanca e influente. Chamamos esse conjunto de dados de "dados com erro". Para avaliar o comportamento das técnicas de diagnóstico sob erro de especificação do modelo não linear em (4.1), ajustamos ao conjunto de "dados com erro" um modelo de regressão simplex linear constante. As estimativas dos parâmetros, erros-padrão e p-valores para o modelo estimado são apresentados na Tabela **??**. Nota-se que a covariável x_t é significativa ao nível de 5% para explicar a média da variável resposta e que mesmo sob a presença de uma obervação influente e sob erro de especificação do modelo a estimava de β_2 não foi afetada consideravelmente, quando comparada ao verdadeiro valor do parâmetro. Já a estimativa do intercepto e do parâmetro de precisão são significativamente subestimadas. Além disso, o intercepto não foi significativo aos níveis usuais.

Para verificar desvios das suposições assumidas inicialmente, construímos gráficos do resíduo ponderado padronizado contra os índices das observações, contra os valores da covariável x_t, contra os valores preditos e o gráfico normal de probabilidades com envelopes simulados (Figura 4.1.1a–d). Nota-se a partir desses gráficos que a observação 30 é destacada como um possível ponto aberrante. Além disso, os gráficos de resíduos contra os valores de x_t e contra os valores preditos evidenciam uma tendência não linear, o que sugere falta na qualidade do ajuste do modelo, ou ainda, falta de algum componente não linear na parte sistemática. Com relação ao gráfico normal de probabilidades com envelopes (Figura 4.1.1d), observa-se que há indícios de afastamento da suposição de que o modelo simplex proposto é adequado.

Tabela 4.1.1 – Resultados inferenciais. Dados simulados.

Parâmetros	Estimativa	Erro-padrão	p-valor
β_1	$-0,2984$	0,2139	0,1630
β_2	1,2564	0,3451	0,0003
λ	0,6315	0,1631	·

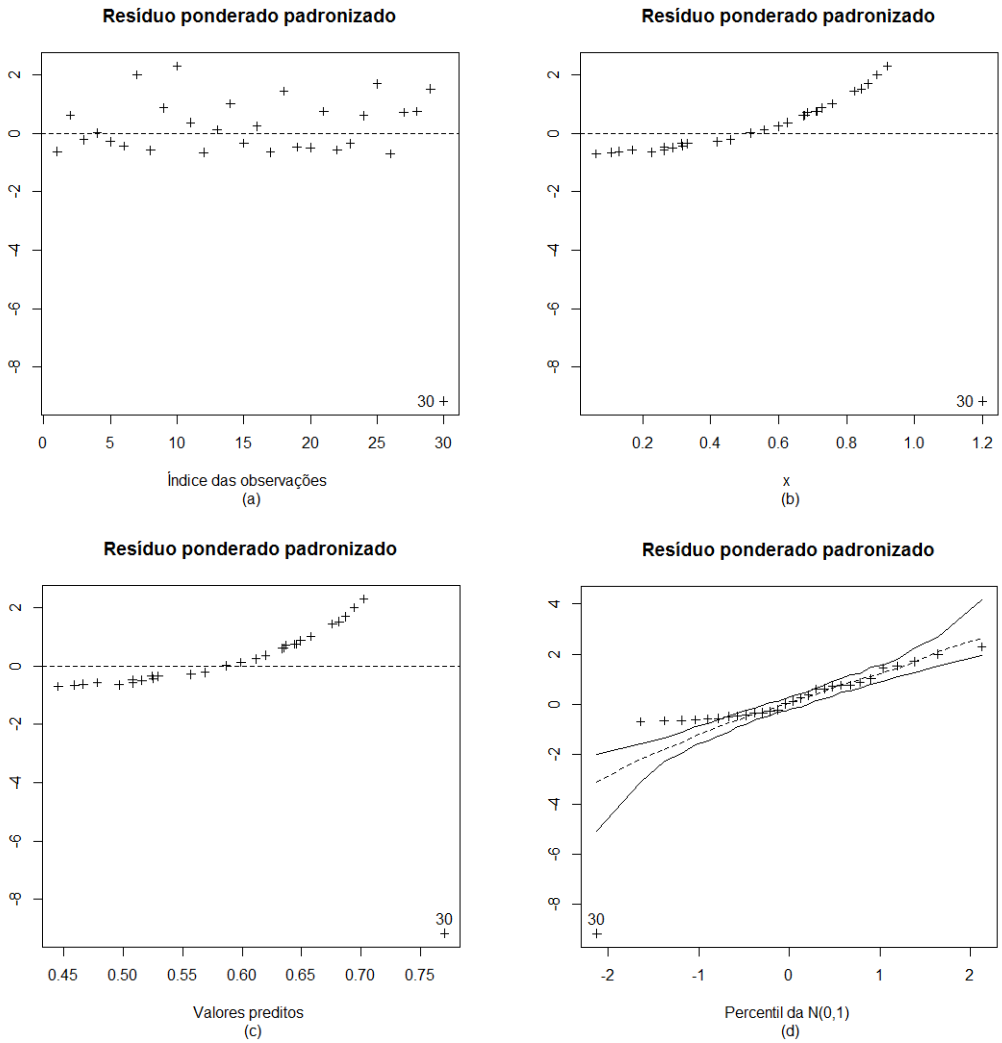

Figura 4.1.1 – Gráficos de resíduos. Dados simulados.

Objetivando identificar observações influentes, construímos os gráficos de I_{max} contra os índices das observações (Figura 4.1.2) e os gráficos de influência local total C_t contra os índices das observações (Figura 4.1.3). Para o esquema de perturbação da covariável, em particular, consideramos a variável explicativa x_t.

De modo geral, verificamos que a observação 30 é destacada como influente para o vetor de parâmetros β e para o parâmetro de precisão λ, considerando todos os esquemas de perturbação avaliados (Figura 4.1.2a–i). Em particular, a perturbação da covariável x_t destaca de forma mais evidente essa observação, quando comparado aos demais esquemas de perturbação. No entanto, é através dos gráficos da influência local total, que o caso 30 é evidenciado fortemente, já que é caracterizado por ser individualmente influente sobre os parâmetros do modelo 4.1.3a–i). Em particular, a perturbação da variável resposta confirma o fato da observação 30 ter influência sobre o próprio valor ajustado, já que este esquema tem relação direta com o conceito de alavanca generalizada como enfatizado por Espinheira et al. (2008b). Ainda com relação ao esquema de perturbação da resposta, são destacadas as observações 7, 10, 12, 25, 26 e 29 como

individualmente influentes. Investigando detalhadamente essas observações, verificamos que os casos 7, 10, 25 e 29 apresentam os maiores valores para a variável resposta, enquanto os casos 12 e 26 apresentaram os menores valores observados para y.

Segundo Cook (1986), gráficos dos componentes de I_{max} contra os índices das observações ou contra covariáveis podem revelar tendências importantes do conjunto de dados, como por exemplo, heteroscedasticidade, componente não linear negligenciada no modelo, etc. Desse modo, objetivando detectar a necessidade de especificação de um modelo não linear para os dados, construímos gráficos de I_{max} contra os valores da covariável x_t para os esquemas de ponderação de casos, perturbação da variável resposta e perturbação da variável explicativa x_t (Figura 4.1.4a–c). A partir desses gráficos, observa-se uma relação não linear entre a medida de influência local e a covariável x_t, o que sugere falta de algum componente não linear na especificação inicial do modelo. De fato, na estimação do modelo consideramos um modelo simplex linear simulando um erro de especificação. Dessa forma, gráficos de I_{max} contra valores de covariáveis podem ser uma ferramenta adicional para validação do modelo postulado.

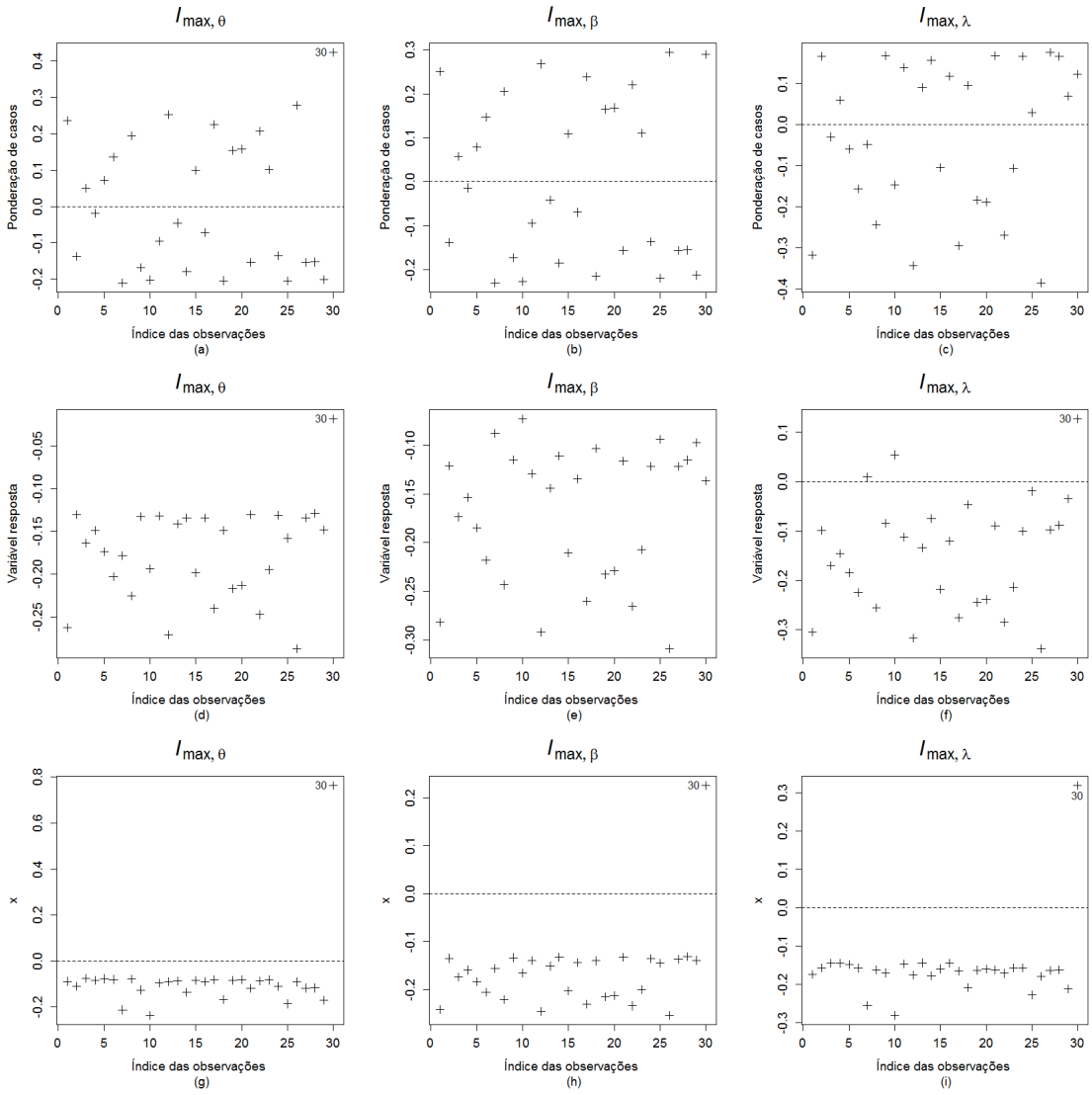

Figura 4.1.2 – Gráficos de influência local. Dados simulados.

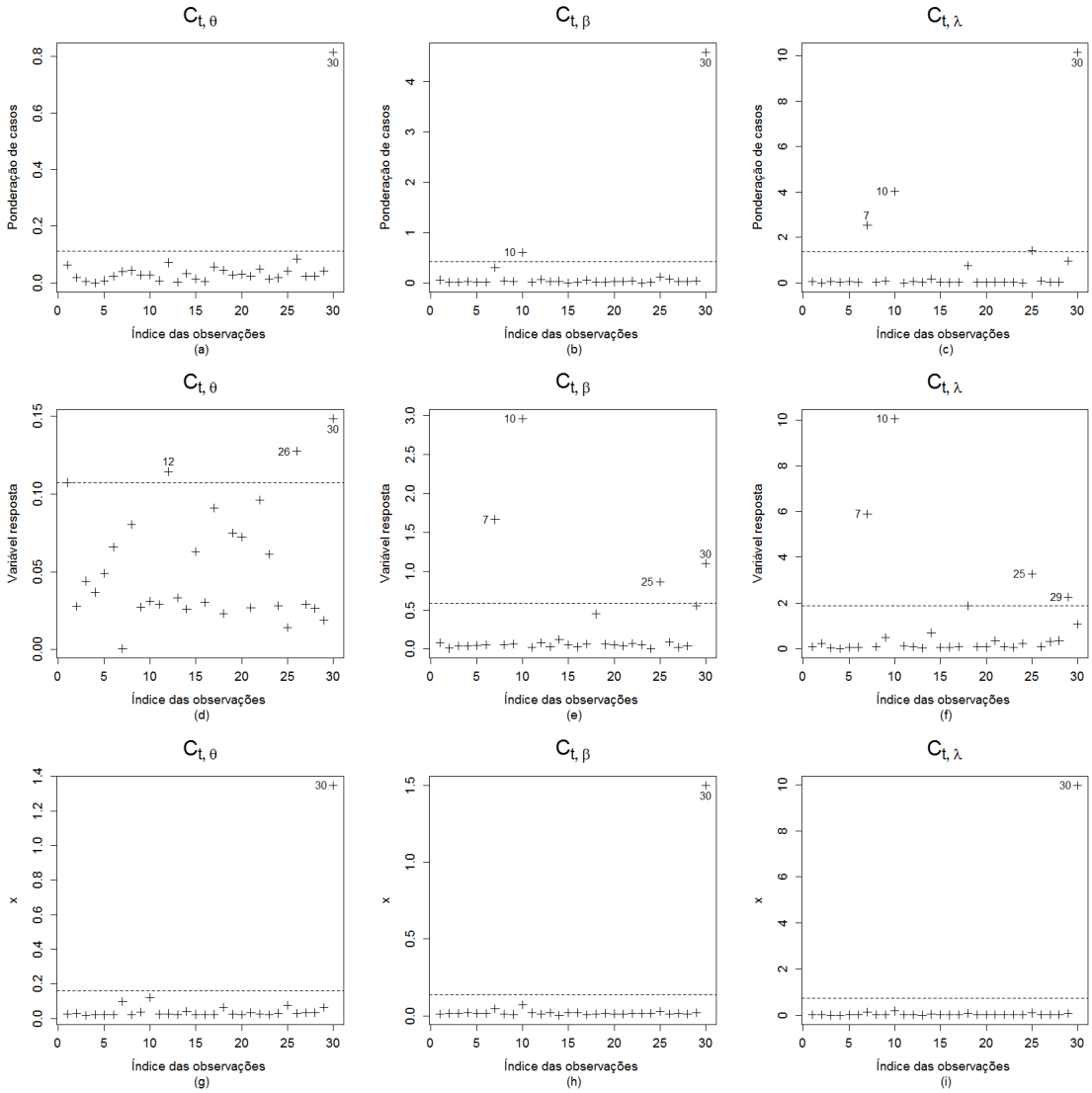

Figura 4.1.3 – Gráficos de influência local total. Dados simulados.

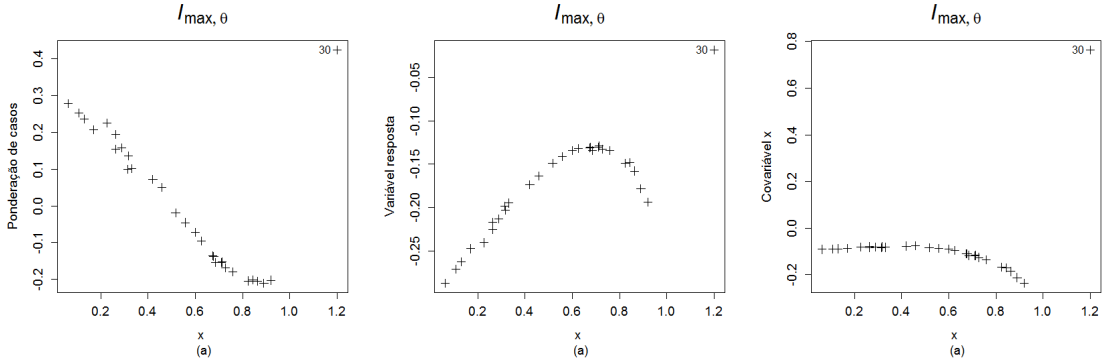

Figura 4.1.4 – Gráficos de influência local contra os valores da covariável x_t.

Para avaliar o efeito da observação 30, destacada nos gráficos de influência, ajustamos o modelo linear aos "dados com erro" sem essa observação. Os resultados da estimação são apresentados na Tabela 4.1.2. Nota-se com a exclusão do caso 30, que tanto a estimativa do

intercepto quanto da covariável x_t foram significativas aos níveis usuais. No entanto, os gráficos de resíduos contra o preditor linear e contra a covariável (Figura 4.1.5b–c), evidenciam ainda falta de qualidade do ajuste do modelo mesmo sob a exclusão da observação influente. Com relação ao gráfico normal de probabilidade com envelopes (Figura 4.1.5d), nota-se uma tendência dos resíduos entre (0;0,5) estarem fora das bandas de confiança, sugerindo que o modelo simplex linear constante não é adequado para modelar os dados.

Tabela 4.1.2 – Resultados inferenciais. Dados simulados excluindo o caso 30.

Parâmetros	Estimativa	Erro-padrão	p-valor
β_1	$-1,1870$	0,1131	0,0000
β_2	3,3757	0,1812	0,0000
λ	2,7060	0,2626	·

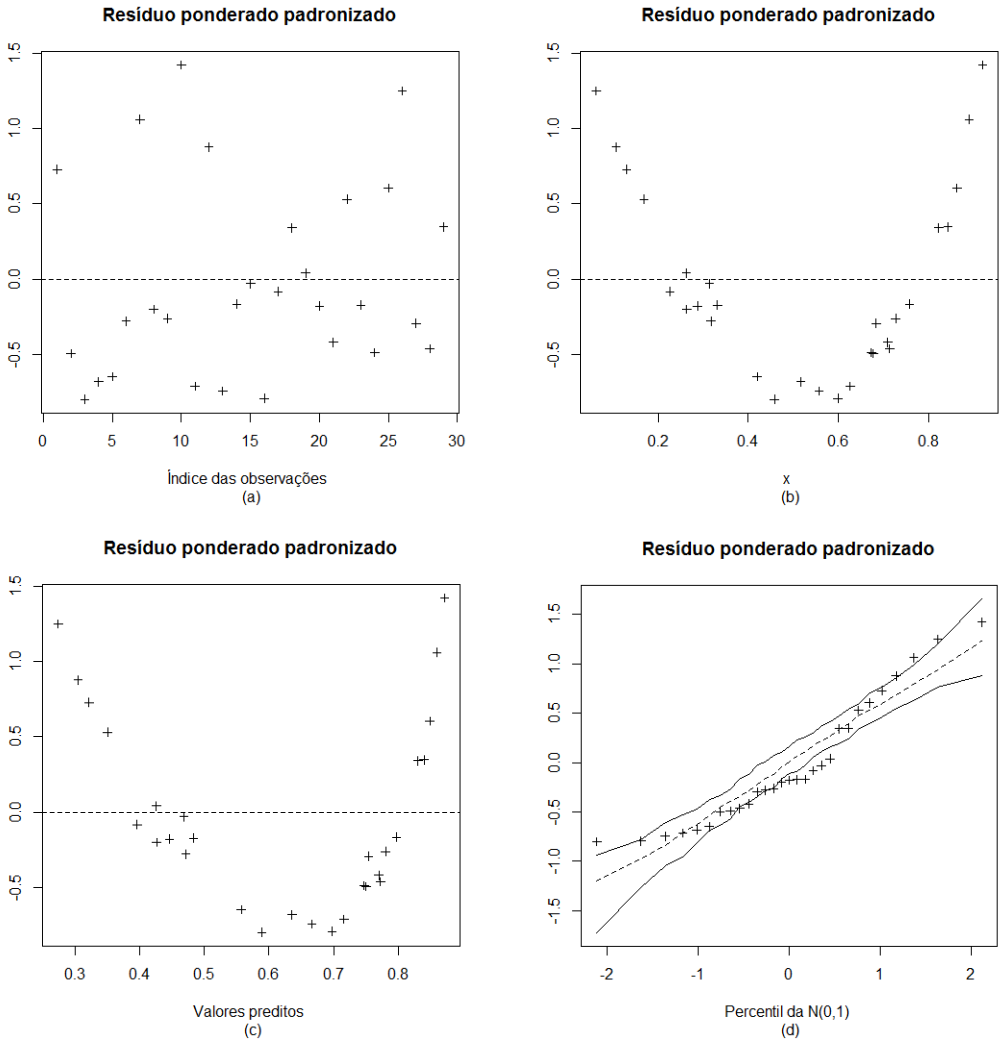

Figura 4.1.5 – Gráficos de resíduos. Dados simulados sem o caso 30.

Dessa forma, ajustamos o modelo simplex não linear constante aos "dados com erro" e aos dados sem a observação influente 30. Os resultados inferenciais dos modelos são apresentados na Tabela 4.1.3. Observa-se que mesmo sob a especificação correta do modelo aos dados, as estimativas dos parâmetros são significativamente subestimadas na presença da observação 30, principalmente o coeficiente da covariável x_t e o parâmetro de precisão λ. Com a exclusão dessa observação, as estimativas do intercepto e da covariável encontram-se próximas dos verdadeiros valores dos parâmetros. O viés relativo das estimativas dos parâmetros β_1 e β_2 foram 0,25% e 0,55%, respectivamente. Com relação ao parâmetro de precisão λ, nota-se que há um viés significativamente maior, quando comparado aos parâmetros da regressão. Na Figura 4.1.6, apresentamos os gráficos dos resíduos contra os índices das observações e o gráfico normal de probabilidades com envelopes para os modelos ajustados. Nota-se que os gráficos baseados no resíduo proposto destacam de forma evidente a observação 30 (Figura 4.1.6a–b). Além disso, o gráficos normal de probabilidades com envelopes, evidencia falta de qualidade de ajuste do modelo não linear constante ajustado aos "dados com erro" (Figura 4.1.6b), uma vez que os resíduos encontram-se fora das bandas de confiança. Esse fato reforça o impacto que a observação 30 exerce sobre o ajuste do modelo, já que este foi especificado corretamente. Considerando o modelo ajustado aos dados sem o caso influente, nota-se claramente que os gráficos de resíduos (Figura 4.1.6c–d) sugerem que o modelo proposto é adequado aos dados, já que os resíduos encontram-se dispersos dentro dos envelopes simulados.

Este exemplo mostra que o resíduo ponderado padronizado proposto para o modelo simplex não linear é útil na identificação de observações aberrantes e que adicionalmente exercem influências desproporcionais sobre o ajuste do modelo. Além disso, gráficos normais de probabilidades com envelopes simulados evidenciam a qualidade do resíduo proposto na investigação de possíveis afastamentos das suposições assumidas para a componente aleatória do modelo. Adicionalmente, destaca-se a importância dos gráficos de influência local contra os índices das observações ou contra valores de covariáveis para revelar tendências do conjunto de dados, auxiliando, desse modo, na validação do modelo postulado aos dados.

Tabela 4.1.3 – Resultados inferenciais para o modelo não linear. Dados simulados.

Dados	Parâmetros	β_1	β_2	λ
Com erro	Estimativas	$-1,0384$	0,6444	0,6001
	p-valor	0,0000	0,0002	
Sem o caso 30	Estimativas	$-1,8044$	1,5082	14,7694
	p-valor	0,0000	0,0000	

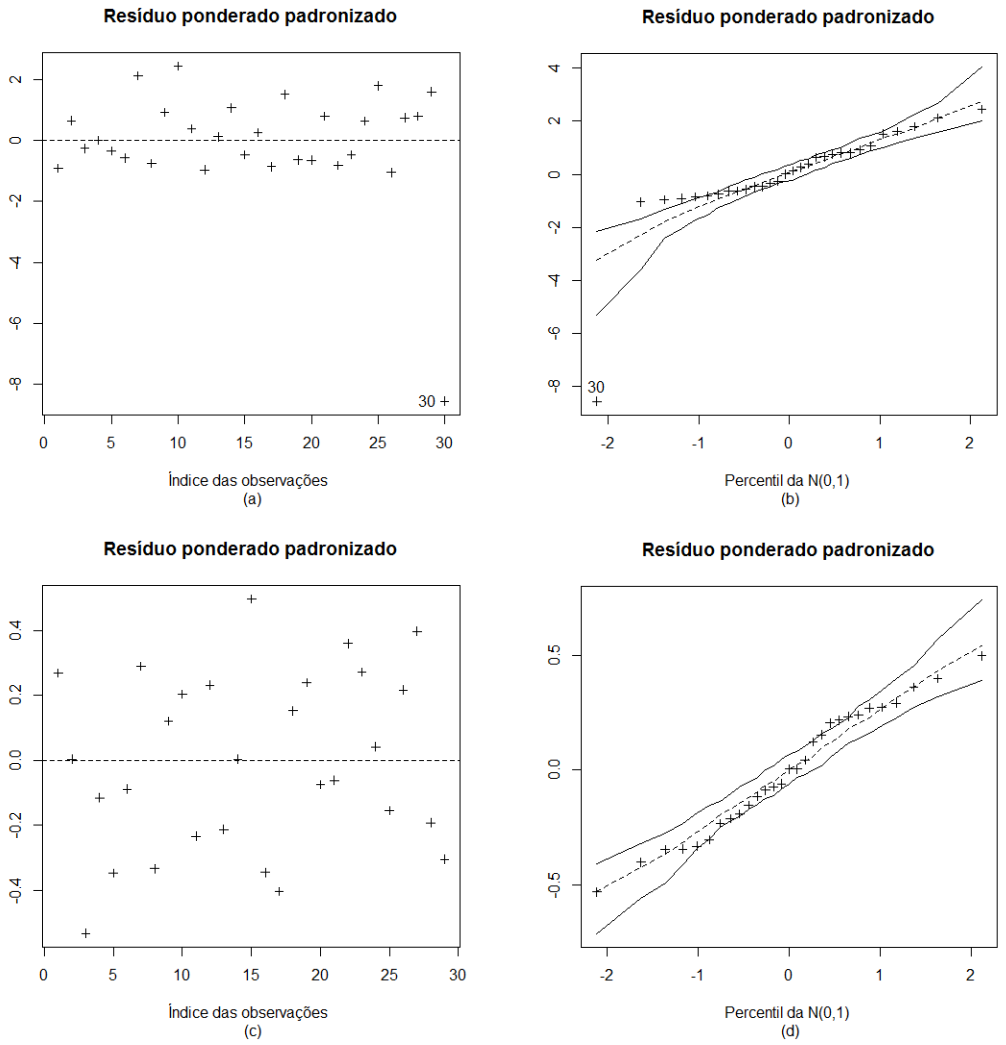

Figura 4.1.6 – Gráficos de resíduos para o modelo não linear: (a)-(b) Dados com erro, (c)-(d) Dados sem o caso 30.

4.2 Aplicação II: dados de oxidação de amônia

Na segunda aplicação, consideramos os dados de oxidação de amônia, analisado originalmente por Brownlee (1965) e apresentado em Miyashiro (2008). O interesse do estudo recai em analisar a perda na conversão de amônia em ácido nítrico durante 21 dias de processos de produção de ácido nítrico em uma planta industrial. O ácido nítrico é utilizado na fabricação de diversos produtos, tais como fertilizantes, corantes, medicamentos, etc. Durante o processo de conversão, o gás amônia reage com o oxigênio presente no ar formando o óxido nítrico, que por sua vez reage novamente com o oxigênio produzindo dióxido de nitrogênio. Este último, reage com a água usada no resfriamento do processo, formando o ácido e o óxido nítricos. O óxido nítrico é absorvido e reutilizado no processo.

Para analisar este conjunto de dados, supomos que a variável resposta y_t, que corresponde a proporção de amônia não convertida em ácido nítrico, segue distribuição simplex $\mathcal{S}^{-1}(\mu_t, \lambda)$

com

$$\log\left(\frac{\mu_t}{1-\mu_t}\right) = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3}, \quad t = 1, \ldots, 21,$$

em que x_{t2} é a corrente de ar e x_{t3} é a temperatura da água utilizada no resfriamento do processo. Na Tabela 4.2.4 são apresentados os resultados da estimação do modelo. Nota-se que a estimativa do parâmetro de precisão λ é significativamente pequena. Além disso, as covariáves corrente de ar e temperatura da água foram significativas ao nível nominal de 5%. O sinal positivo da estimativa do parâmero correspondente a variável corrente de ar indica que um incremento nesta covariável implica em um aumento na proporção de amônia não convertida em ácido nítrico. De maneira similar, o sinal positivo de β_3 sugere que quanto maior a temperatura da água utilizada no resfriamento do processo maior será a perda na conversão de amônia em ácido nítrico.

Com o objetivo de verificar possíveis afastamentos das suposições feitas para o modelo, construímos os gráficos de resíduos contra os índices das observações (Figura 4.2.7a), contra os valores das covariáveis corrente de ar (Figura 4.2.7b) e temperatura da água (Figura 4.2.7c) e o gráfico normal de probabilidades com envelopes simulados (Figura 4.2.7d). Com base nesses gráficos, observa-se que os resíduos encontram-se dispersos aleatoriamente em torno de zero, o que sugere uma boa adequação do modelo simplex aos dados. Com relação ao gráfico normal de probabilidades com envelopes simulados (Figura 4.2.7d), verifica-se, de forma geral, que os resíduos encontram-se dentro das bandas de confiança, indicando que não há fortes indícios de afastamento da suposição de que o modelo simplex é adequado para os dados.

Tabela 4.2.4 – Resultados inferenciais. Dados de oxidação de amônia.

Parâmetros	Estimativa	Erro-padrão	p-valor
β_1	$-8,1660$	0,2570	0,0000
β_2	0,0494	0,0056	0,0000
β_3	0,0497	0,0156	0,0015
λ	0,7844	0,2421	·

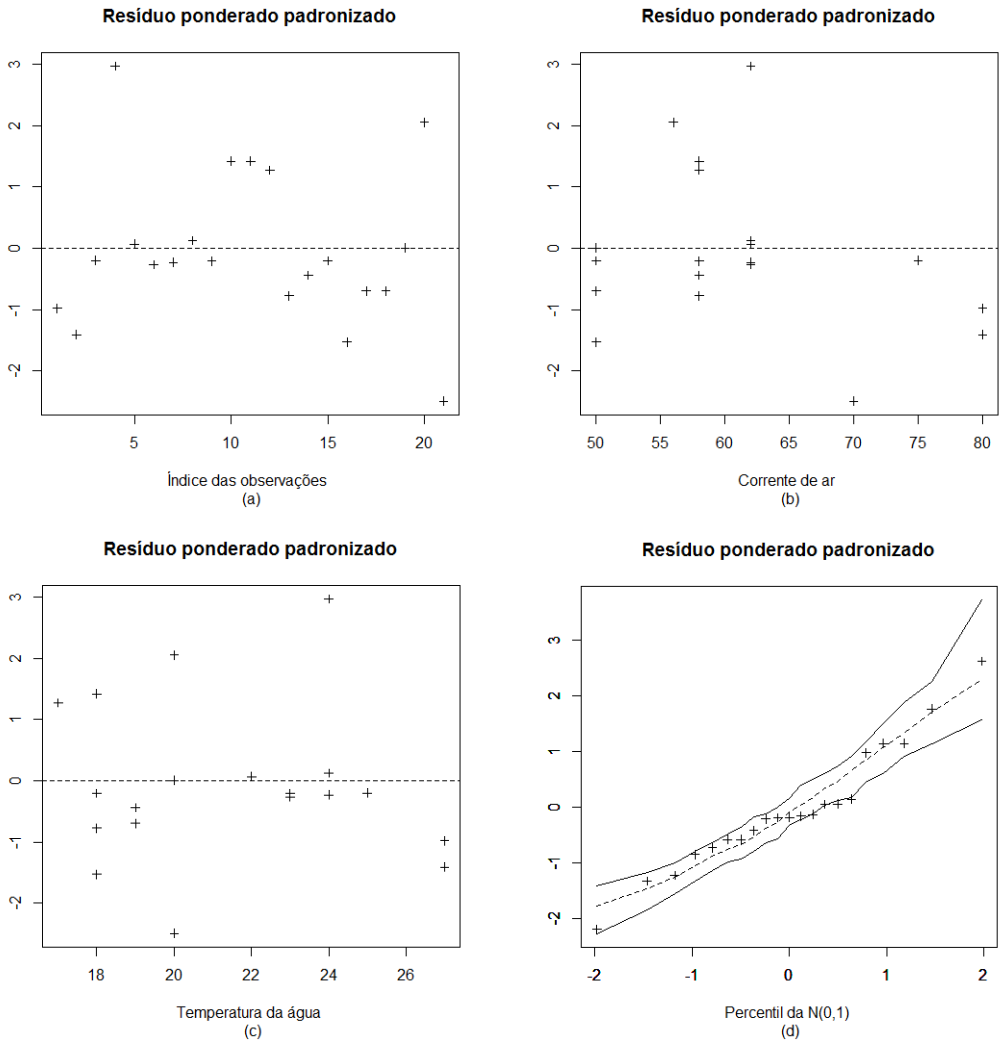

Figura 4.2.7 – Gráficos de resíduos. Dados de oxidação de amônia.

Dando continuidade à análise de diagnóstico construímos os gráficos da alavanca generalizada contra os valores preditos (Figura 4.2.8) e os gráficos de influência local. Nas Figuras 4.2.9 e 4.2.10 encontram-se, respectivamente, os gráficos de I_{max} contra os índices das observações e os gráficos de influência local total C_t contra os índices das observações. Em particular, para a perturbação individual de covariáveis utilizamos as variáveis corrente de ar e temperatura da água.

Na Figura 4.2.8, nota-se que o gráfico de alavanca generalizada destaca as observações 4, 7, 12, 15 e 16. Considerando o esquema de ponderação de casos (Figura 4.2.9a–c), verificamos que a observação 4 é destacada como sendo influente para o vetor de parâmetros β e para o parâmetro λ. Além disso, os gráficos relacionados ao esquema de perturbação individual de covariáveis destacam de modo geral as observações 10, 11, 12 e 15 como conjuntamente influentes tanto para a estimativa de β quanto para a de λ. Entretanto, a perturbação da variável temperatura da água trás ainda o caso 20 que não foi destacado considerando a perturbação da variável corrente de ar. Quanto ao esquema de perturbação da resposta, que está relacionada

ao conceito de alavanca generalizada, não foram observados casos conjuntamente influentes de forma evidente.

Considerando agora a influência individual das observações, verificamos segundo o esquema de ponderação de casos que as observações 4 e 16 exercem influências individuais sobre a estimativa do vetor $\theta = (\beta^\top, \lambda)^\top$. Considerando apenas o vetor de coeficientes da regressão β são destacadas as observações 4 e 21, enquanto que para o parâmetro de precisão λ são evidenciados os casos 4, 20 e 21. Considerando a perturbação das covariáveis temperatura da água e corrente de ar, destacam-se as observações 4, 10, 11, 12, 15, 20 e 21 como individualmente influentes. Já para o esquema de perturbação da resposta, destaca-se apenas a observação 21 como sendo influente.

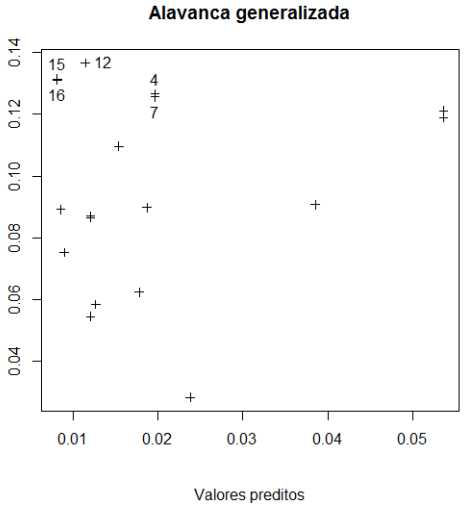

Figura 4.2.8 – Gráfico de alavanca generalizada. Dados de oxidação de amônia.

Para avaliar a influencia das observações destacadas nos resultados inferenciais do modelo, excluímos individualmente os casos 4, 10, 12, 15, 16, 20 e 21, e conjuntamente os casos {10, 11, 12, 15} e {10, 11, 12 , 20}. Outras exclusões foram avaliadas, porém, não apresentaram variações significativas nas estimativas dos coeficientes da regressão. Na Tabela 4.2.5 apresentamos as variações percentuais nas estimativas dos parâmetros relativas às exclusões mais expressivas e os p-valores para os testes de significância associados aos parâmetros. Nota-se que o caso 4 influencia consideravelmente a estimativa de β_3 e em menor proporção a estimativa da precisão λ. De maneira similar, a exclusão do caso 21 resulta em mudanças mais expressivas apenas sobre a estimativa de β_2, quando comparada à exclusão do caso 4. Já a exclusão dos casos 10, 12, 15, 16 e 20 não resultaram em mudanças substanciais nas estimativas dos parâmetros do modelo. Adicionalmente, as exclusões conjuntas dos casos {10, 11, 12, 15} e {10, 11, 12 , 20} afetam consideravelmente a estimativa do coeficiente da variável temperatura da água e de forma menos expressiva as estimativas de β_2 e λ. No entanto, vale destacar que para esta aplicação dispõe-se de um tamanho de amostra de $n = 21$, o que pode ser um fator decisivo na

exclusão de muitas observações. Nenhuma das exclusões alterou a significância dos coeficientes do modelo, nem os sinais das estimativas.

Investigamos detalhadamente os casos 4 e 21, por apresentarem mudanças substanciais nos resultados inferenciais do modelo. Inicialmente, temos que os valores da variável resposta concentram-se no intervalo (0,007;0,042). O caso 4 apresenta uma proporção de amônia não convertida em ácido nítrico de 0,028. Espera-se que quanto menor a corrente de ar menor seja a perda na conversão de amônia em ácido nítrico. No entanto, essa observação contradiz esta relação por apresentar uma corrente de ar de 62, considerada pequena, uma vez que essa covariável concentra-se no intervalo (50, 80), mas com uma perda significativa na conversão de amônia. Já o caso 21, aponta na direção oposta, uma vez que apresenta um valor de corrente de ar de 70, com uma proporção de amônia não convertida consideravelmente pequena de 0,015.

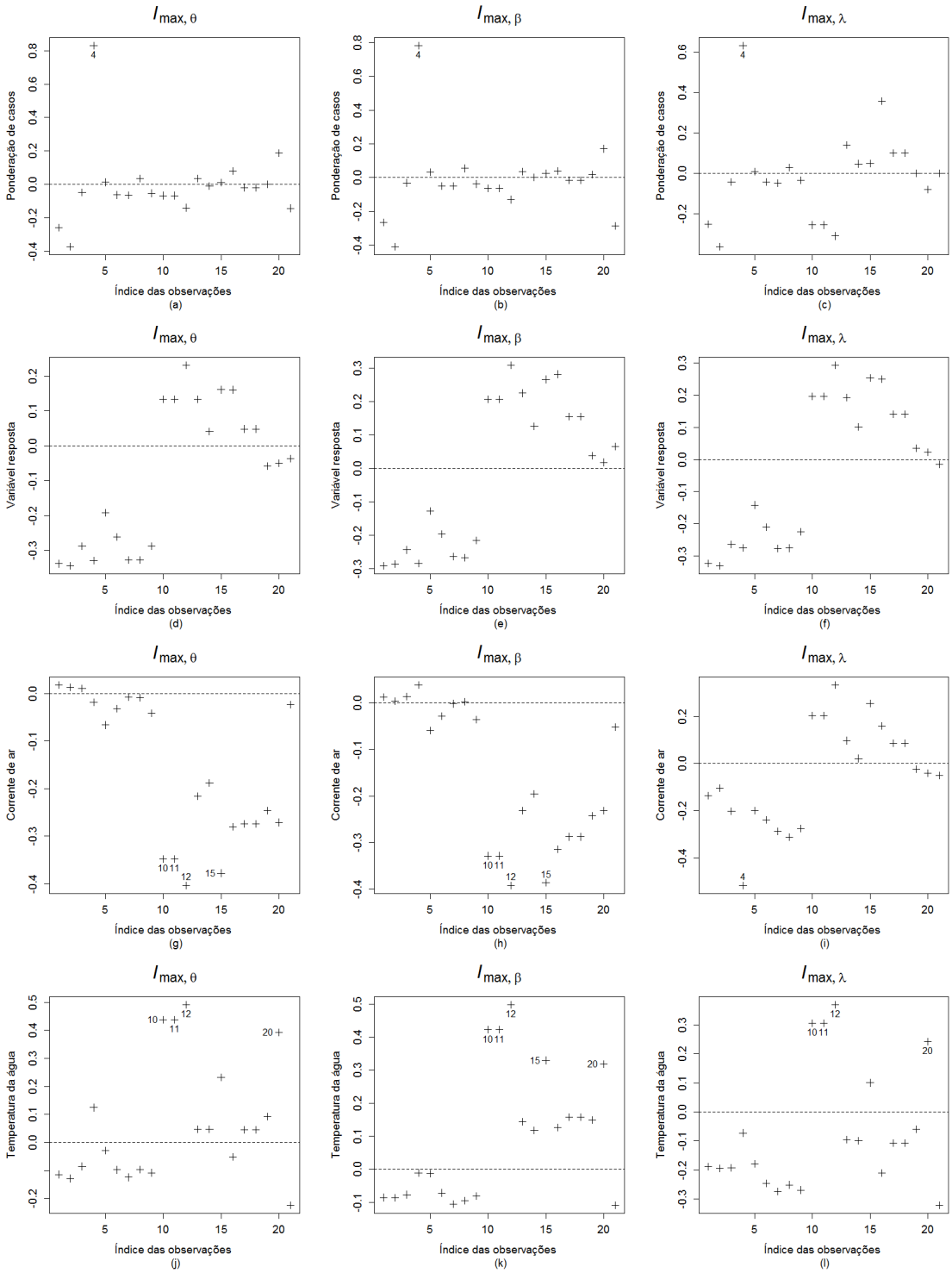

Figura 4.2.9 – Gráficos de influência local. Dados de oxidação de amônia.

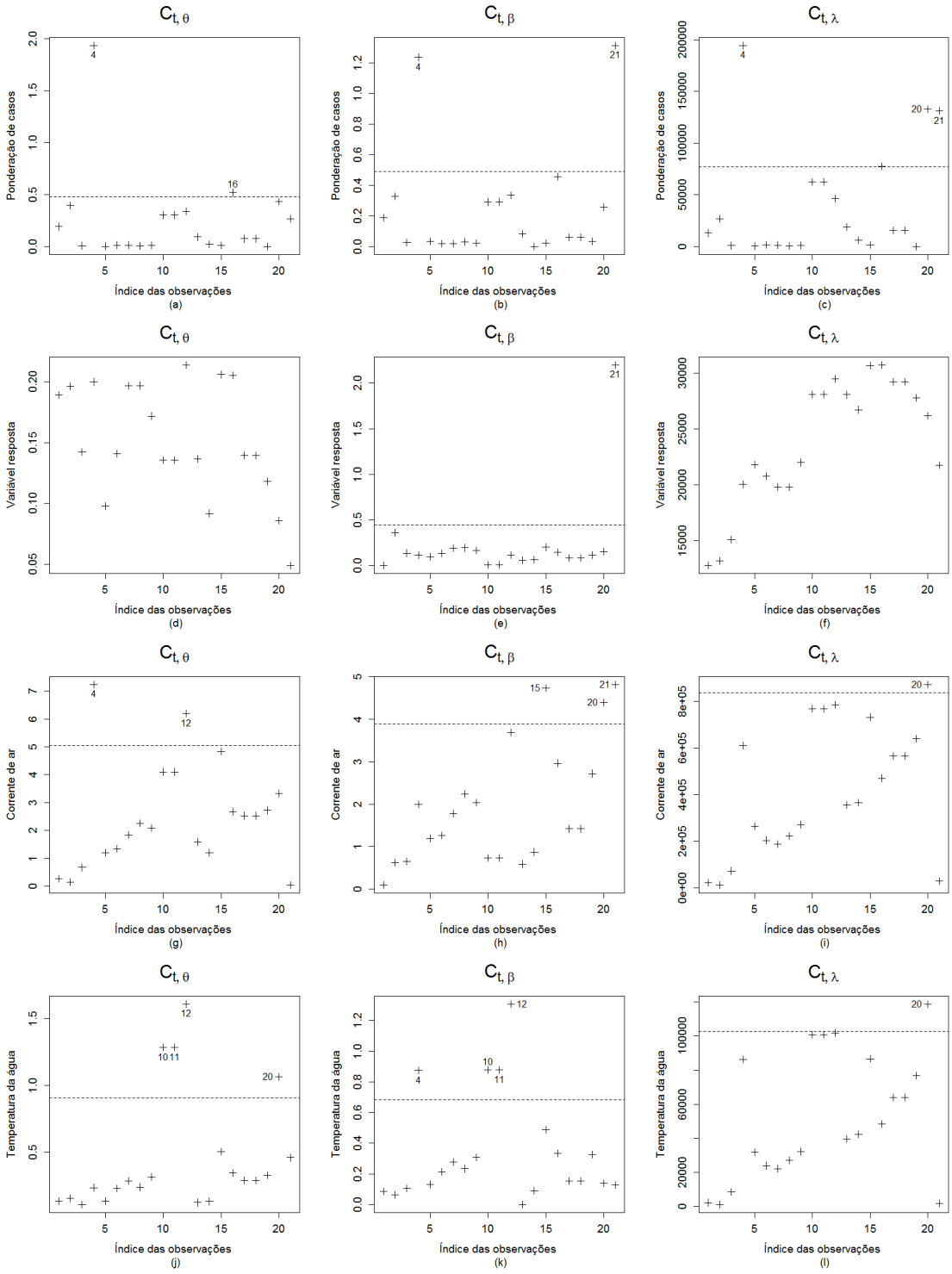

Figura 4.2.10 – Gráficos de influência local total. Dados de oxidação de amônia.

Tabela 4.2.5 – Variação percentual das estimativas dos parâmetros e p-valores retirando observações influentes. Dados de oxidação de amônia.

Casos	β_1	β_2	β_3	λ
4	$-3,2623$	1,6194	$-32,7968$	14,7374
p-valor	0,0000	0,0000	0,0271	·
10	0,1188	$-2,2267$	6,4386	$-1,0071$
p-valor	0,0000	0,0000	0,0012	·
12	$-0,2376$	$-4,4534$	9,8592	0,4972
p-valor	0,0000	0,0000	0,0015	·
15	$-0,3380$	$-1,417$	1,2072	$-1,9378$
p-valor	0,0000	0,0000	0,0015	·
16	$-2,0463$	$-5,4656$	$-0,2012$	3,4421
p-valor	0,0000	0,0000	0,0013	·
20	0,0269	$-0,4049$	$-0,2012$	$-2,5752$
p-valor	0,0000	0,0000	0,0016	·
21	2,3500	17,4089	$-28,3702$	13,9597
p-valor	0,0000	0,0000	0,0224	·
10, 11, 12, 15	1,4279	$-27,5304$	83,7022	16,1907
p-valor	0,0000	0,0000	0,0000	·
10, 11, 12, 20	2,0916	$-30,9717$	96,5795	31,3233
p-valor	0,0000	0,0000	0,0000	·

4.3 Aplicação III: dados de inseticida

A terceira aplicação refere-se aos dados apresentados em McCullagh e Nelder (1989, pg. 384). O interesse do estudo recai na estimação de uma mistura de menor custo de inseticida e sinérgico, através da modelagem da proporção de gafanhotos mortos. Simas et al. (2010) e Rocha e Simas (2010) utilizaram esses dados como ilustração dos modelos de regressão beta não linear. Aqui, assumimos que a variável resposta y_t, $t = 1, \ldots, 15$, é a proporção de gafanhotos mortos e as variáveis independentes são dose de inseticida (x_{t2}) e dose de sinérgico (x_{t3}). Além disso, consideramos que y_t segue distribuição simplex $\mathcal{S}^{-1}(\mu_t, \lambda_t)$ com

$$\begin{aligned} g(\mu_t) &= \beta_1 + \beta_2 \log(x_{t2} - \beta_3) + \beta_4 x_{t3}/(\beta_5 + x_{t3}), \\ h(\lambda_t) &= \gamma_1 + \gamma_2 x_{t2} + \gamma_3 x_{t3}, \end{aligned}$$

em que $g(\cdot)$ e $h(\cdot)$ são as funções de ligação logito e raiz quadrada, respectivamente. As estimativas dos parâmetros foram obtidas por máxima verossimilhança através do método de otimização BFGS e são apresentadas na Tabela 4.3.6. Nota-se que apenas o parâmetro β_5 não é significativamente diferente de zero ao nível nominal de 5% para o submodelo da média (p-valor

> 0.05). Além disso, as covariáveis dose de inseticida e sinérgico não foram significativas para o submodelo da precisão, considerando o nível de significância de 5%. No entanto, para efeito da análise de diagnóstico realizada a seguir, consideremos o modelo completo. Na Figura 4.3.11, apresentamos os gráficos de resíduos contra os índices das observações (a), contra as covariáveis dose de inseticida (b) e dose de sinérgico (c) e o gráfico normal de probabilidades com envelopes simulados (d). A partir desses gráficos observa-se que os resíduos encontram-se dispersos aleatoriamente em torno do zero, sugerindo uma boa adequação da componente sistemática do modelo. No entanto, com relação ao gráfico normal de probabilidades (Figura 4.3.11d), nota-se uma leve falta de qualidade de ajuste do modelo simplex não linear com precisão variável, já que os resíduos encontram-se fora das bandas de confiança.

Tabela 4.3.6 – Resultados inferenciais. Dados de inseticida.

Parâmetro	Estimativa	Erro-padrão	p-valor
β_1	$-3,4612$	0,4764	0,0000
β_2	1,5672	0,1767	0,0000
β_3	1,5031	0,1676	0,0000
β_4	1,8587	0,1734	0,0000
β_5	1,9115	1,0373	0,0654
γ_1	0,2151	0,4271	0,6145
γ_2	0,1257	0,0653	0,0542
γ_3	0,0419	0,0233	0,0724

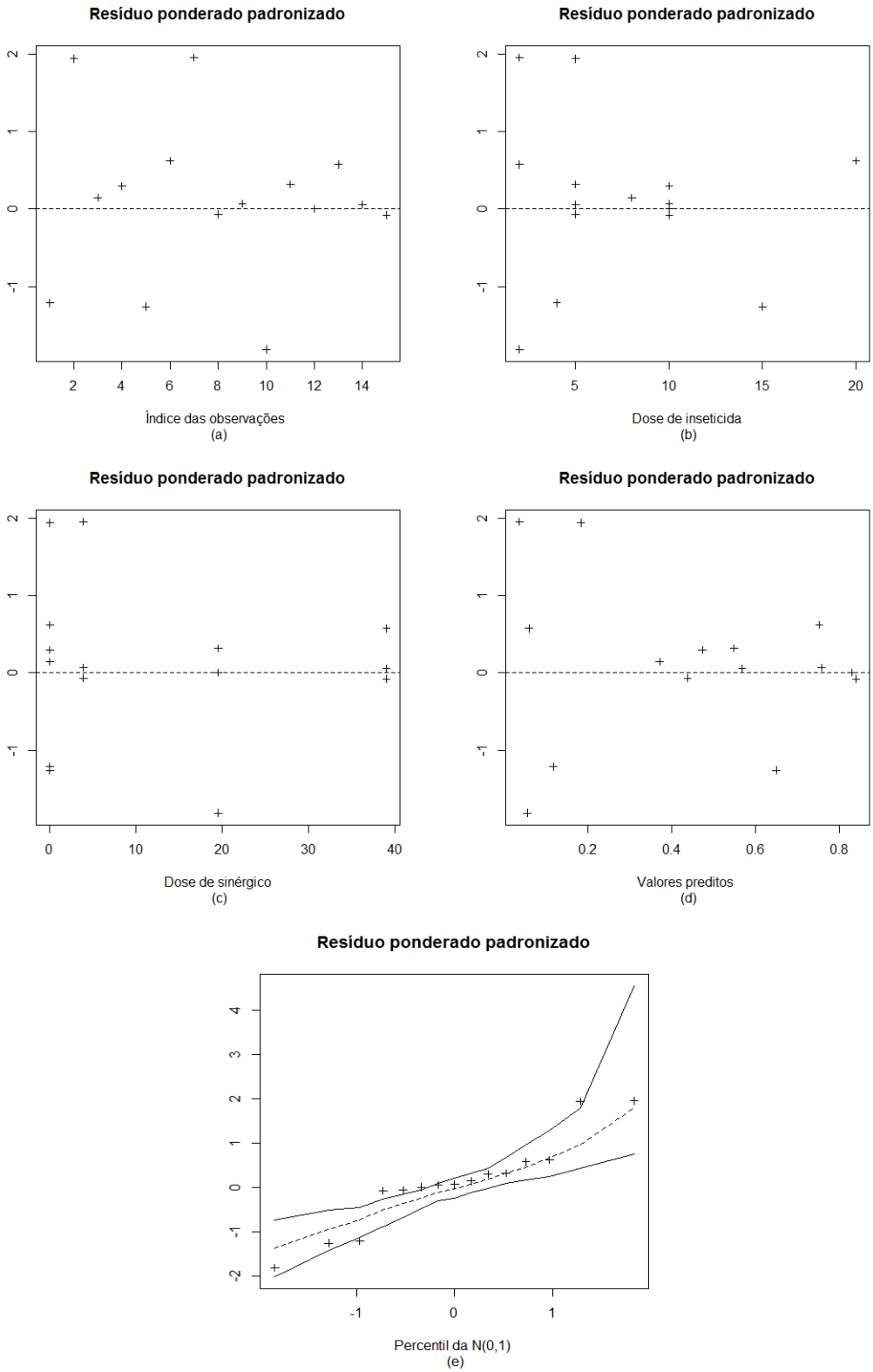

Figura 4.3.11 – Gráficos de resíduos. Dados de inseticida.

Com o objetivo de identificar observações com efeitos desproporcionais sobre os resultados inferenciais do modelo, construímos o gráfico de alavanca generalizada contra os valores preditos (Figura 4.3.12) e os gráficos de influência local. Na Figura 4.3.13 apresentamos os

gráficos de I_{max} contra os índices das observações e na Figura 4.3.14 estão apresentados os gráficos de influência local total C_t contra os índices das observações. Para a perturbação individual de covariáveis, em particular, consideramos as covariáveis dose de inseticida e dose de sinérgico. Adicionalmente, como as covariáveis determinam ambas a média e a precisão, consideramos o esquema de perturbação simultânea apresentado na Seção 3.4.5. A partir da Figura 4.3.12 temos que as observações 6, 14 e 15 são destacadas como pontos de alavanca. Para o esquema de ponderação de casos, temos que a observação 5 é identificada como sendo influente para os vetores β e γ. Considerando a perturbação aditiva da resposta, nota-se a influência de dois subconjuntos de observações na estimativa de β, em direções opostas, sendo o primeiro formado unicamente pelo caso 6 e o segundo pelas observações 14 e 15. Para o vetor γ, em particular, são destacados os casos 5 e 6 como sendo influentes. Como a perturbação da variável resposta está relacionada ao conceito de alavanca generalizada, essas observações possivelmente são responsáveis por exercerem influências desproporcionais sobre o próprio valor ajustado, com destaque para a observação 6, que configura-se como um ponto alavanca e influente. Cabe destacar, que apenas o esquema de perturbação aditiva da resposta destacou os casos 14 e 15 como conjuntamente influentes, mostrando a eficiência do fator de escala de padronização proposto na identificação dessas observações. Note, que apesar do gráfico de alavanca generalizada destacar as observações 6, 14 e 15, os gráficos de influência para perturbação da resposta, identificaram a direção contrária dessas observações, permitindo a retirada conjunta dos casos 14 e 15. Considerando a perturbação individual de covariáveis, nota-se, de modo geral, que as observações 5 e 6 são destacadas como sendo conjuntamente influentes. Entretanto, a perturbação da covariável dose de inseticida trás ainda a observação 10, que não foi destacada no gráfico de I_{max} para a perturbação da variável explicativa dose de sinérgico (Figura 4.3.13j–l).

Avaliando agora a influência individual das observações sobre os vetores β e γ, nota-se, para o esquema de ponderação de casos que as observações 5 e 10 são destacadas como individualmente influentes. Considerando a perturbação aditiva da resposta, são evidenciados os casos 6 ,10 e 15 como individualmente influentes para a estimativa do vetor β, enquanto apenas o caso 10 é destacado como influente para o vetor de parâmetro da precisão. Já para a perturbação individual de covariáveis, destacam-se, de modo geral, as observações 2, 5, 6 e 10. Nota-se ainda, que a perturbação da variável explicativa dose de inseticida evidencia a observação 7, que não é destaca no gráfico de C_t para a variável dose de sinérgico.

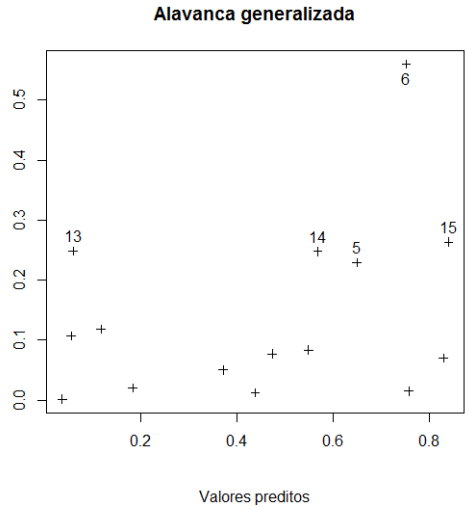

Figura 4.3.12 – Gráfico de alavanca generalizada. Dados de inseticida.

Para avaliar a influência das observações destacadas nos gráficos de I_{max} e C_t, excluímos individualmente os casos 2, 5, 6, 10, 13 e 15, e conjuntamente as observações $\{5, 6\}$, $\{5, 6, 10\}$ e $\{14, 15\}$. As variações percentuais nas estimativas dos parâmetros e os p-valores para os testes de significância associados aos parâmetros são apresentados na Tabela 4.3.7 para as exclusões mais relevantes.

Nota-se, que a exclusão individual das observações 5 e 6 afetam significativamente as estimativas dos parâmetros do modelo da precisão. Na ausência desses casos isoladamente, os parâmetros tornam-se substancialmente significativos. Além disso, com a exclusão do caso 5 o parâmetro β_5 passa a ser significativo aos níveis usuais. Por outro lado, a exclusão conjunta dos casos $\{14, 15\}$ conduzem a uma conclusão contrária, ou seja, o parâmetro β_5 torna-se ainda menos significativo assim como todos os parâmetros do modelo da precisão.

Diante de tais fatos, ajustamos inicialmente alguns modelos competidores, tais como o modelo com precisão variável sem o intercepto e o modelo com precisão constante ao longo das observações. Não apresentaremos estas análises aqui, no entanto, relatamos que em todos os modelos o parâmetro β_5 ou se mostrou diretamente não significativo ou extremamente sensível aos casos influentes.

Ao analisarmos as observações influentes, notamos que os casos 5 e 6 apresentam a mesma característica, são ocorrências em que a dose de sinérgico é zero e as doses de inseticida são, respectivamente, 15 e 20, conduzindo a uma proporção de gafanhotos mortos relativamente alta, 60% e 80%, evidenciando a importância da dose de inseticida na resposta média. Quando essas observações são retiradas a importância da dose de inseticida é diminuída e sugere, talvez erroneamente, que a dose de sinérgico é importante para a explicação da resposta tornando β_5 significativo. Esse fato pode ser uma evidência que a dose de sinérgico não interfere na proporção de gafanhotos mortos. A característica comum dos casos 14 e 15 reforçam essa ideia. Pois essas

observações tratam-se de ocorrências em que a dose de sinérgico é a maior possível, 39, mas o que parece definir a proporção de gafanhotos mortos é a dose de inseticida iguais a 5 e 10 conduzindo a proporções de gafanhotos mortos iguais a 60% e 80%. Assim, a nossa sugestão é que apenas a dose de inseticida é importante para a explicação da proporção média de gafanhotos mortos. Portanto, uma possibilidade futura seria considerar o modelo

$$\begin{aligned} g(\mu_t) &= \beta_1 + \beta_2 \log(x_{t2} - \beta_3), \\ h(\lambda_t) &= \gamma_1 + \gamma_2 x_{t2}. \end{aligned}$$

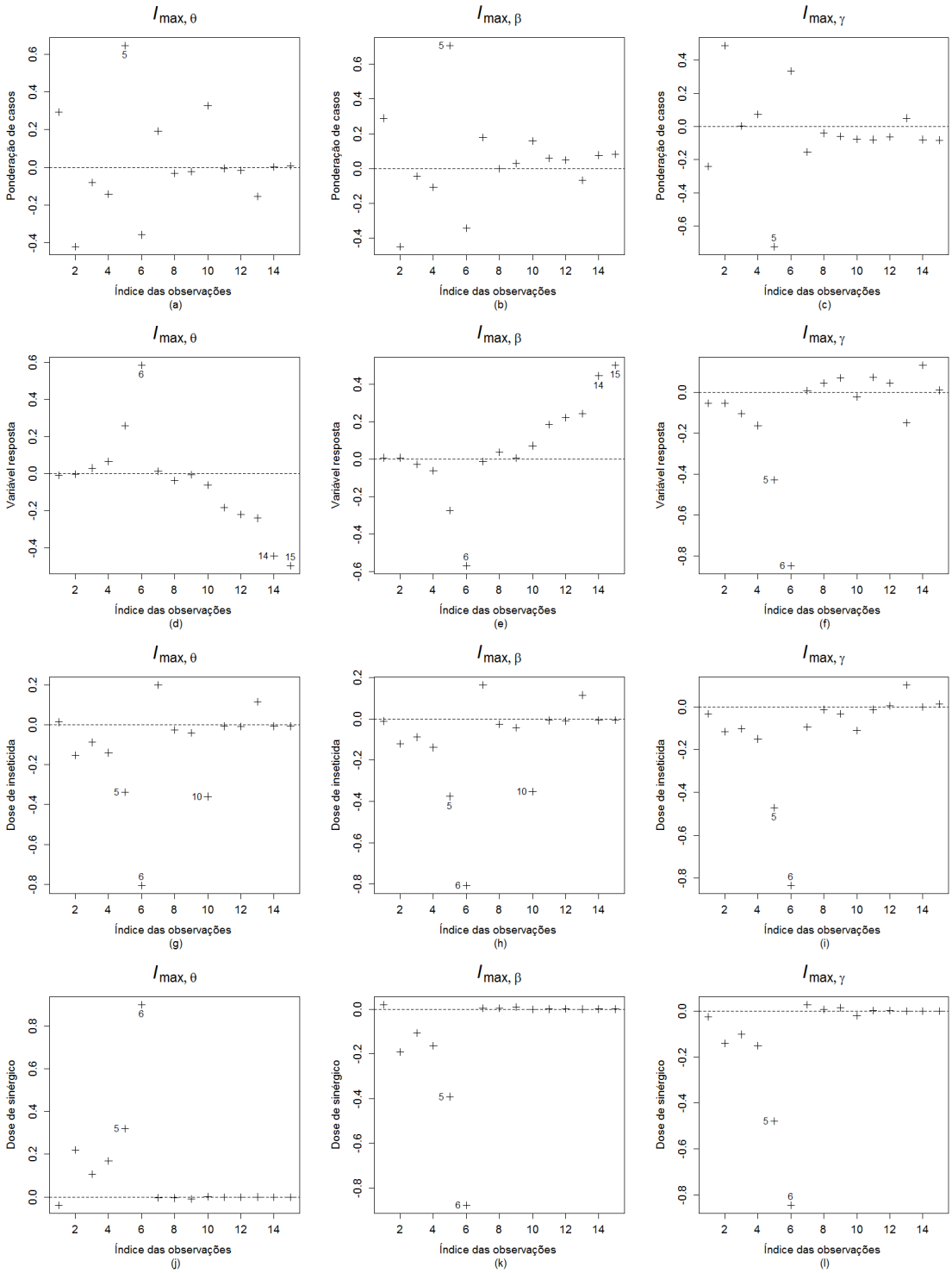

Figura 4.3.13 – Gráficos de influêncial local. Dados de inseticida.

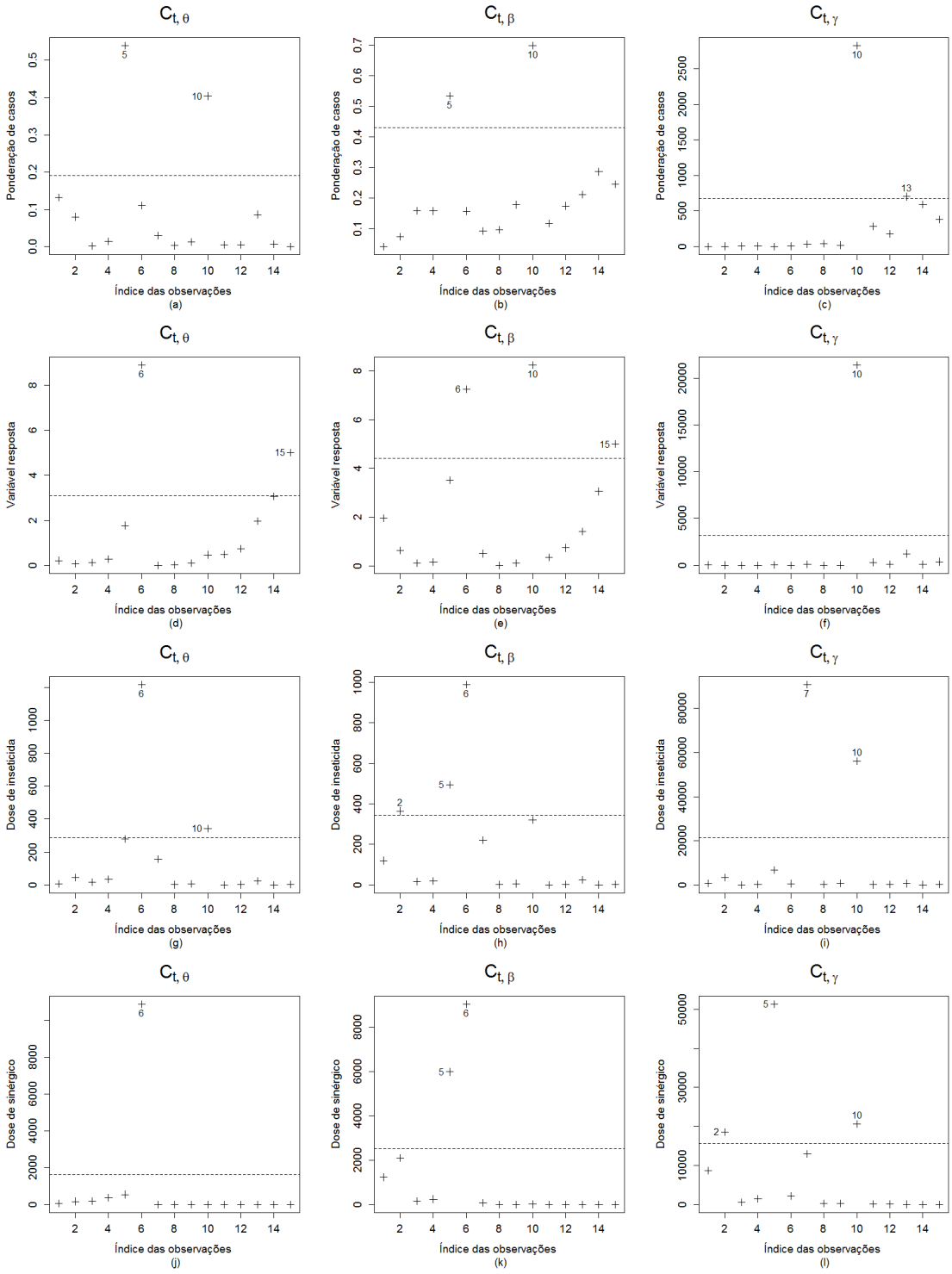

Figura 4.3.14 – Gráficos de influência local total. Dados de inseticida.

Tabela 4.3.7 – Variação percentual das estimativas dos parâmetros e p-valores retirando observações influentes. Dados de inseticida.

Casos	β_1	β_2	β_3	β_4	β_5	γ_1	γ_2	γ_3
Completo	$-3,46$	1,57	1,50	1,86	1,91	0,22	0,13	0,04
p-valor	0,0000	0,0000	0,0000	0,0000	0,0654	0,6145	0,0545	0,0724
5	$-3,27$	1,52	1,54	1,76	1,97	$-1,00$	0,47	0,07
p-valor	0,0000	0,0000	0,0000	0,0000	0,0022	0,0019	0,0001	0,0298
6	$-2,05$	0,95	1,92	1,82	1,87	$-1,92$	0,60	$-0,01$
p-valor	0,0000	0,0000	0,0000	0,0000	0,0029	0,0003	0,0000	0,7129
14, 15	$-3,50$	1,58	1,49	1,86	1,78	0,32	0,11	0,03
p-valor	0,0000	0,0000	0,0000	0,0000	0,0994	0,4874	0,0848	0,2280

CAPÍTULO 5

CONSIDERAÇÕES FINAIS

5.1 Conclusões

Nesta dissertação propomos uma extensão do modelo de regressão simplex proposto por Miyashiro (2008), em que tanto a média da variável resposta quanto o parâmetro precisão estão relacionados às covariáveis por meio de preditores não lineares. No Capítulo 2 definimos o modelo simplex não linear e apresentamos expressões em forma fechada para o vetor escore, matriz de informação de Fisher e sua inversa.

No Capítulo 3, propomos o resíduo ponderado padronizado para o modelo simplex não linear com base no processo iterativo escore de Fisher. Avaliamos numericamente o comportamento da distribuição empírica do resíduo proposto para diferentes modelos simplex, a saber: modelo simplex constante, modelo simplex com precisão variável e modelo simplex não linear com precisão variável. Os resultados evidenciaram, de modo geral, que o resíduo proposto tem média aproximadamente zero e desvio-padrão inferior a um. Além disso, a distribuição do resíduo apresenta leve assimetria e curtose que acompanha a da distribuição $\mathcal{N}(0, 1)$. Ainda no Capítulo 3, desenvolvemos medidas de influência baseadas no método de influência local desenvolvido por Cook (1986), para diferentes esquemas de perturbação. Para o esquema de perturbação aditiva da resposta, em particular, propomos um novo fator de escala de padronização, útil em situações em que o cálculo da variância da resposta apresenta alto custo computacional.

Finalmente, no capítulo 4, aplicamos a metodologia proposta a dados reais e dados simulados. Na primeira aplicação, geramos uma única amostra para avaliar o comportamento das técnicas de diagnósticos sob a presença de observações que exercem efeitos desproporcionais sobre o ajuste do modelo e sob erro de especificação da componente sistemática. Na segunda aplicação, consideramos os dados de oxidação de amônia, analisados por Miyashiro (2008) e na terceira, os dados apresentados em McCullagh e Nelder (1989, pg.384). A partir dos resultados, recomendamos medidas de influência local baseadas nos seguintes esquemas de perturbação:

ponderação de casos, perturbação da variável resposta, perturbação de covariáveis da média, perturbação de covariáveis da precisão e perturbação simultânea de covariáveis.

APÊNDICE A

PROPRIEDADES DA DISTRIBUIÇÃO SIMPLEX

Apresentamos a seguir alguns resultados para a distribuição simplex apresentada em (2.2). Para mais detalhes ver Song e Tan (2000) e Song (2007).

Proposição A.1. *Seja* y_t, $t = 1, \ldots, n$, *uma variável aleatória com distribuição simplex de parâmetros* $\mu_t \in (0,1)$ *e* $\lambda_t > 0$. *Então,*

(a) $\mathrm{E}\{d(y_t;\mu_t)\} = \frac{1}{\lambda_t}$;

(b) $\mathrm{E}\left\{(y_t-\mu_t)\frac{\partial d(y_t;\mu_t)}{\partial \mu_t}\right\} = -\frac{2}{\lambda_t}$;

(c) $\mathrm{E}\{(y_t-\mu_t)d(y_t;\mu_t)\} = 0$;

(d) $\mathrm{E}\left(\frac{\partial^2 d(y_t;\mu_t)}{\partial \mu_t^2}\right) = 0$;

(e) $\frac{1}{2}\mathrm{E}\left(\frac{\partial^2 d(y_t;\mu_t)}{\partial \mu_t^2}\right) = \frac{3}{\lambda_t\mu_t(1-\mu_t)} + \frac{1}{\mu_t^3(1-\mu_t)^3}$;

(f) $\mathrm{Var}\{d(y_t;\mu_t)\} = \frac{2}{\lambda_t^2}$.

O componente do desvio para a distribuição simplex é dado por

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

Pode-se mostrar que

$$\frac{\partial d(y_t;\mu_t)}{\partial \mu_t} = -2(y_t-\mu_t)u_t.$$

em que

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

Assim,

$$-\frac{1}{2}\frac{\partial d(y_t;\mu_t)}{\partial \mu_t} = (y_t - \mu_t)u_t. \tag{A.1}$$

Pela Proposição A.1 (c), e considerando que $\mathrm{E}(y_t) = \mu_t$, seque que

$$\mathrm{E}\{(y_t - \mu_t)u_t\} = \frac{\mathrm{E}\{(y_t - \mu_t)d(y_t;\mu_t)\}}{\mu_t(1-\mu_t)} + \frac{\mathrm{E}(y_t - \mu_t)}{\mu_t^3(1-\mu_t)^3} = 0. \tag{A.2}$$

APÊNDICE B

INFLUÊNCIA LOCAL

Considerando o modelo de regressão simplex não linear definido em (2.4), com $\theta = (\beta^\top, \gamma^\top)^\top$, temos que Δ é uma matriz $(k+q) \times n$ dada por

$$\Delta = \begin{pmatrix} \frac{\partial^2 \ell_\delta(\theta)}{\partial\beta_1\partial\delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\theta)}{\partial\beta_1\partial\delta_n} \\ \frac{\partial^2 \ell_\delta(\theta)}{\partial\beta_2\partial\delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\theta)}{\partial\beta_2\partial\delta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ell_\delta(\theta)}{\partial\beta_k\partial\delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\theta)}{\partial\beta_k\partial\delta_n} \\ \frac{\partial^2 \ell_\delta(\theta)}{\partial\gamma_1\partial\delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\theta)}{\partial\gamma_1\partial\delta_n} \\ \frac{\partial^2 \ell_\delta(\theta)}{\partial\gamma_2\partial\delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\theta)}{\partial\gamma_2\partial\delta_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \ell_\delta(\theta)}{\partial\gamma_q\partial\delta_1} & \cdots & \frac{\partial^2 \ell_\delta(\theta)}{\partial\gamma_q\partial\delta_n} \end{pmatrix},$$

avaliada em $\widehat{\theta} = (\widehat{\beta}^\top, \widehat{\gamma}^\top)^\top$ e δ_0. Podemos expressar Δ como

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

em que $\Delta_\beta = \Delta_\beta(\theta, \delta)$ é uma matriz $k \times n$ e $\Delta_\gamma = \Delta_\gamma(\theta, \delta)$ é uma matriz $q \times n$, dadas, respectivamente, por

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

Nas seções seguintes, apresentamos medidas de influência local para o modelo proposto considerando diferentes esquemas de perturbação.

B.1 Ponderação de casos

Para o esquema de ponderação de casos, temos que $\ell_\delta(\beta,\gamma) = \sum_{t=1}^{n} \delta_t \ell_t(\mu_t,\lambda_t)$ e, assim, segue que a t-ésima coluna de Δ é dada por

$$\Delta_t = \frac{\partial \ell_t(\theta}{\partial \theta} = \left(\frac{\partial \ell_t(\theta)}{\partial \beta_1},\ldots,\frac{\partial \ell_t(\theta)}{\partial \beta_k},\frac{\partial \ell_t(\theta)}{\partial \gamma_1},\ldots,\frac{\partial \ell_t(\theta)}{\partial \gamma_q}\right)^\top\Bigg|_{\theta=\widehat{\theta}}.$$

Com base em (2.5), temos que o (i,t)-ésimo elemento de Δ_β e o (j,t)-ésimo elemento de Δ_γ são dados, respectivamente, por

$$\frac{\partial \ell_t(\beta,\gamma)}{\partial \beta_i} = \lambda_t(y_t-\mu_t)u_t\frac{1}{g'(\mu_t)}\frac{\partial \eta_t}{\partial \beta_i} \tag{B.1}$$

e

$$\frac{\partial \ell_t(\beta,\gamma)}{\partial \gamma_j} = a_t\frac{1}{h'^{(\lambda_t)}}\frac{\partial \zeta_t}{\partial \gamma_j}, \tag{B.2}$$

em que u_t e a_t são dados, respectivamente, em (2.9) e (2.14), e $g(\mu_t)$ e $h(\lambda_t)$ estão definidas em (2.4).

Avaliando (B.1) e (B.2) em $\delta_0 = (0,0,\ldots,0)^\top$ e $\widehat{\theta} = (\widehat{\beta}^\top,\widehat{\gamma}^\top)^\top$, podemos escrever $\Delta_\beta = \widehat{\tilde{X}}^\top\widehat{\Lambda}\widehat{U}\widehat{T}\mathcal{E}$ e $\Delta_\gamma = \widehat{\tilde{Z}}^\top\widehat{H}\mathcal{A}$, em que $\mathcal{E}$ e $\mathcal{A}$ estão definidas em (3.12) e (3.14), respectivamente. Dessa forma, a matriz Δ fica expressa na forma apresentada em (3.13).

B.2 Perturbação da resposta

Considerando o esquema de perturbação aditiva na resposta apresentada em (3.15), e com base em (2.5), temos que o logaritmo da função de verossimilhança do modelo perturbado é dado por

$$\begin{aligned}\ell_\delta(\beta,\gamma) &= \sum_{t=1}^{n}\{(1/2)\log\lambda_t - (1/2)\log 2\pi - (3/2)\log\{y_t(\delta)(1-y_t(\delta))\}\\ &- (\lambda_t/2)d(y_t(\delta);\mu_t)\},\end{aligned}$$

em que $y_t(\delta)$ é a resposta modificada definida em (3.15). Assim, temos que

$$\begin{aligned}\frac{\partial \ell_\delta(\beta,\gamma)}{\partial \delta_t} &= -\frac{\lambda_t}{2}s(y_t)\frac{1}{y_t(\delta)(1-y_t(\delta))}\left\{d(y_t(\delta);\mu_t) + \frac{2(y_t(\delta)-\mu_t)}{y_t(\delta)\mu_t(1-\mu_t)^2}\right\}\\ &- \frac{3}{2}s(y_t)\frac{(1-2y_t(\delta))}{y_t(\delta)(1-y_t(\delta))}.\end{aligned}$$

Para $i = 1, \ldots, k$

$$\begin{aligned}\frac{\partial^2 \ell_\delta(\beta,\gamma)}{\partial\beta_i \partial\delta_t} &= \lambda_t \frac{s(y_t)}{y_t(\delta)(1-y_t(\delta))}\left\{\frac{2}{y_t(\delta)(1-\mu_t)^3} + \frac{1-3\mu_t}{\mu_t^2(1-\mu_t)^3} - \frac{1}{2}\frac{\partial d(y_t(\delta);\mu_t)}{\partial\mu_t}\right\} \\ &\times \frac{1}{g'(\mu_t)}\frac{\partial\eta_t}{\partial\beta_i} \end{aligned} \tag{B.3}$$

e para $j = 1, \ldots, q$

$$\begin{aligned}\frac{\partial^2 \ell_\delta(\beta,\gamma)}{\partial\gamma_j \partial\delta_t} &= -\frac{1}{2}\frac{s(y_t)}{y_t(\delta)(1-y_t(\delta))}\left\{d(y_t(\delta);\mu_t) + \frac{2(y_t(\delta)-\mu_t)}{y_t(\delta)(1-y_t(\delta))}\right\} \\ &\times \frac{1}{h'(\lambda_t)}\frac{\partial\zeta_t}{\partial\gamma_j}. \end{aligned} \tag{B.4}$$

Avaliando as expressões (B.3) e (B.4), em $\delta_0 = (0, 0, \ldots, 0)^\top$ e $\widehat{\theta} = (\widehat{\beta}^\top, \widehat{\gamma}^\top)^\top$, temos que $\Delta_\beta = \widehat{\tilde{X}}^\top \widehat{\Lambda}\widehat{T}BS_y$ e $\Delta_\gamma = \widehat{\tilde{Z}}^\top \widehat{H}CS_y$, em que B e C estão definidas na Seção 3.4.2. Logo, a matriz Δ fica expressa na forma apresentada em (3.16).

B.3 Perturbação de covariável da média ($x_p^\top$)

Neste caso, temos que o logaritmo da função de verossimilhança do modelo perturbado é dado por

$$\begin{aligned}\ell_\delta(\beta,\gamma) &= \sum_{t=1}^{n}\{(1/2)\log\lambda_t - (1/2)\log 2\pi - (3/2)\log\{y_t(1-y_t)\} \\ &- (\lambda_t/2)d(y_t;\mu_t(\delta))\}, \end{aligned}$$

em que $\mu_t(\delta)$ é tal que $g(\mu_t(\delta)) = \eta_t(\delta)$, com $\eta_t(\delta)$ dado em (3.18). Dessa forma, temos que

$$\frac{\partial\ell_\delta(\beta,\gamma)}{\partial\delta_t} = -\frac{\lambda_t}{2}\frac{\partial d(y_t;\mu_t(\delta))}{\partial\mu_t(\delta)}\frac{1}{g'(\mu_t(\delta))}\frac{\partial\eta_t(\delta)}{\partial\delta_t}.$$

Como $-(1/2)\partial d(y_t;\mu_t(\delta))/\partial\mu_t(\delta) = (y_t - \mu_t(\delta))u_t(\delta)$, obtemos

$$\frac{\partial\ell_\delta(\beta,\gamma)}{\partial\delta_t} = \lambda_t(y_t - \mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{\partial\eta_t(\delta)}{\partial\delta_t}, \tag{B.5}$$

em que, para este esquema de perturbação,

$$u_t(\delta) = \frac{1}{\mu_t(\delta)(1-\mu_t(\delta))}\left\{d(y;\mu_t(\delta)) + \frac{1}{\mu_t^2(\delta)(1-\mu_t(\delta))^2}\right\}$$

De (B.5) segue que para $i = 1, \ldots, k$

$$\begin{aligned}\frac{\partial \ell_\delta(\beta,\gamma)}{\partial \beta_i \partial \delta_t} &= -\lambda_t \left\{ u_t(\delta) - (y_t - \mu_t(\delta))u_t'(\delta) + (y_t - \mu_t(\delta))u_t(\delta)\frac{g''(\mu_t(\delta))}{g'(\mu_t(\delta))} \right\} \frac{1}{\{g'(\mu_t(\delta))\}^2} \\ &\times \frac{\partial \eta_t(\delta)}{\partial \delta_t}\frac{\partial \eta_t(\delta)}{\partial \beta_i} + \lambda_t(y_t - \mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{\partial^2 \eta_t(\delta)}{\partial \beta_i \partial \delta_t},\end{aligned}$$

que pode ser reescrita como

$$\frac{\partial \ell_\delta(\beta,\gamma)}{\partial \beta_i \partial \delta_t} = -\lambda_t q_t(\delta)\frac{\partial \eta_t(\delta)}{\partial \delta_t}\frac{\partial \eta_t(\delta)}{\partial \beta_i} + \lambda_t(y_t - \mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{\partial^2 \eta_t(\delta)}{\partial \beta_i \partial \delta_t}, \tag{B.6}$$

em que

$$q_t(\delta) = \left\{ u_t(\delta) - (y_t - \mu_t(\delta))u_t'(\delta) + (y_t - \mu_t(\delta))u_t(\delta)\frac{g''(\mu_t(\delta))}{g'(\mu_t(\delta))} \right\} \frac{1}{\{g'(\mu_t(\delta))\}^2},$$

com

$$u_t'(\delta) = -\left\{ \frac{2(y_t - \mu_t(\delta))u_t(\delta)}{\mu_t(\delta)(1-\mu_t(\delta))} + \frac{3(1-2\mu_t(\delta))}{\mu_t^4(\delta)(1-\mu_t(\delta))^4} + \frac{(1-2\mu_t(\delta))d(y_t;\mu_t(\delta))}{\mu_t^2(\delta)(1-\mu_t(\delta))^2} \right\}.$$

Ainda de (B.5), temos para $j = 1, \ldots, q$

$$\frac{\partial \ell_\delta(\beta,\gamma)}{\partial \gamma_j \partial \delta_t} = (y_t - \mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{1}{h'(\lambda_t)}\frac{\partial \eta_t(\delta)}{\partial \delta_t}\frac{\partial \zeta_t}{\partial \gamma_j}. \tag{B.7}$$

Avaliando as expressões (B.6) e (B.7) em $\delta_0 = (0, 0, \ldots, 0)^\top$ e $\widehat{\theta} = (\widehat{\beta}^\top, \widehat{\gamma}^\top)^\top$, temos que $\Delta_\beta = -\widehat{\widetilde{X}}^\top \widehat{\Lambda}\widehat{Q}\widehat{\widetilde{X}}_\delta + [\ \widehat{b}_\beta^\top\ ][\ \widehat{\widetilde{X}}_{\beta\delta}\ ]$ e $\Delta_\gamma = \widehat{\tilde{Z}}^\top \widehat{H}\widehat{T}\widehat{U}\mathcal{E}\widehat{\widetilde{X}}_\delta$, em que $\tilde{X}_{\beta\delta}$, $\tilde{X}_\delta$ e b_β estão definidos na Seção 3.4.3. Assim, a matriz Δ fica expressa na forma apresentada em (3.19).

B.4 Perturbação de covariável da precisão ($z_{p'}^\top$)

Neste caso, temos que o logaritmo da função de verossimilhança do modelo perturbado é dado por

$$\begin{aligned}\ell_\delta(\beta,\gamma) &= \sum_{t=1}^{n} \{(1/2)\log\lambda_t(\delta) - (1/2)\log 2\pi - (3/2)\log\{y_t(1-y_t)\} \\ &- (\lambda_t(\delta)/2)d(y_t;\mu_t)\},\end{aligned}$$

em que $\lambda_t(\delta)$ é tal que $h(\lambda_t(\delta)) = \zeta_t(\delta)$, com $\zeta_t(\delta)$ dado em (3.20). Logo, temos que

$$\frac{\partial \ell_\delta(\beta,\gamma)}{\partial \delta_t} = \left\{ \frac{1}{2\lambda_t(\delta)} - \frac{d(y_t;\mu_t)}{2} \right\} \frac{1}{h'(\lambda_t(\delta))} \frac{\partial \zeta_t(\delta)}{\partial \delta_t},$$

que pode ser reescrita como

$$\frac{\partial \ell_\delta(\beta,\gamma)}{\partial \delta_t} = a_t(\delta) \frac{1}{h'(\lambda_t(\delta))} \frac{\partial \zeta_t(\delta)}{\partial \delta_t}, \tag{B.8}$$

em que

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

De (B.8), seque que para $i = 1, \ldots, k$

$$\frac{\partial^2 \ell_\delta(\beta,\gamma)}{\partial \beta_i \partial \delta_t} = (y_t - \mu_t) u_t \frac{1}{g'(\mu_t)} \frac{1}{h'(\lambda_t(\delta))} \frac{\partial \zeta_t(\delta)}{\partial \delta_t} \frac{\partial \eta_t}{\partial \beta_i}. \tag{B.9}$$

Considerando agora o vetor de parâmetros γ, temos que para $j = 1, \ldots, q$

$$\begin{aligned} \frac{\partial^2 \ell_\beta(\beta,\gamma)}{\partial \gamma_j \partial \delta_t} &= -\left\{ \frac{1}{2\lambda_t^2(\delta)} + a_t(\delta) \frac{h''(\lambda_t(\delta))}{h'(\lambda_t(\delta))} \right\} \frac{1}{\{h'(\lambda_t(\delta))\}^2} \frac{\partial \zeta_t(\delta)}{\partial \delta_t} \frac{\partial \zeta_t(\delta)}{\partial \gamma_j} \\ &+ a_t(\delta) \frac{1}{h'(\lambda_t(\delta))} \frac{\partial^2 \zeta_t(\delta)}{\partial \gamma_j \partial \delta_t}. \end{aligned}$$

Definindo

$$\nu_t(\delta) = \left\{ \frac{1}{2\lambda_t^2(\delta)} + a_t(\delta) \frac{h''(\lambda_t(\delta))}{h'(\lambda_t(\delta))} \right\} \frac{1}{\{h'(\lambda_t(\delta))\}^2},$$

obtemos

$$\frac{\partial^2 \ell_\delta(\beta,\gamma)}{\partial \gamma_j \partial \delta_t} = -\nu_t(\delta) \frac{\partial \zeta_t(\delta)}{\partial \delta_t} \frac{\partial \zeta_t(\delta)}{\partial \gamma_j} + a_t(\delta) \frac{1}{h'(\lambda_t(\delta))} \frac{\partial^2 \zeta_t(\delta)}{\partial \gamma_j \partial \delta_t}. \tag{B.10}$$

Avaliando as expressões apresentadas em (B.9) e (B.10) no vetor de não perturbação $\delta_0 = (0, 0, \ldots, 0)^\top$ e no estimador de máxima verossimilhança $\widehat{\theta}$, obtemos $\Delta_\beta = \widehat{\tilde{X}}^\top \widehat{T} \widehat{H} \widehat{U} \mathcal{E} \widehat{\tilde{Z}}_\delta$ e $\Delta_\gamma = -\widehat{\tilde{Z}}^\top \mathcal{V} \widehat{\tilde{Z}}_\delta + [\ \widehat{b}_\gamma^\top\ ][\ \widehat{\tilde{Z}}_{\gamma\delta}\ ]$, em que $\tilde{Z}_{\gamma\delta}$, $\tilde{Z}_\delta$ e b_γ estão definidos na Seção 3.4.4. Assim, a matriz Δ fica expressa na forma apresentada em (3.21).

B.5 Perturbação simultânea de covariáveis $(x_p^\top, z_{p'}^\top)$

Para este tipo de perturbação, temos que o logaritmo da função de verossimilhança é dado por

$$\begin{aligned}\ell_\delta(\beta,\gamma) &= \sum_{t=1}^{n}\{(1/2)\log\lambda_t(\delta) - (1/2)\log 2\pi - (3/2)\log\{y_t(1-y_t)\} \\ &- (\lambda_t(\delta)/2)d(y_t;\mu_t(\delta))\}\,,\end{aligned}$$

em que $\mu_t(\delta) = g^{-1}(\eta_t(\delta))$ e $\lambda_t(\delta) = h^{-1}(\zeta_t(\delta))$, com $\eta_t(\delta)$ e $\zeta_t(\delta)$ dados em (3.22). Assim, temos que

$$\frac{\partial\ell_\delta(\beta,\gamma)}{\partial\delta_t} = a_t(\delta)\frac{1}{h'(\lambda_t(\delta))}\frac{\partial\zeta_t(\delta)}{\partial\delta_t} + \lambda_t(\delta)(y_t-\mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{\partial\eta_t(\delta)}{\partial\delta_t}. \tag{B.11}$$

De (B.11) segue que para $i = 1,\ldots,k$

$$\begin{aligned}\frac{\partial^2\ell_\delta(\beta,\gamma)}{\partial\beta_t\partial\delta_t} &= (y_t-\mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{1}{h'(\lambda_t(\delta))}\frac{\partial\eta_t(\delta)}{\partial\beta_i}\frac{\partial\zeta_t(\delta)}{\partial\delta_t} \\ &- \lambda_t(\delta)q_t(\delta)\frac{\partial\eta_t(\delta)}{\partial\beta_i}\frac{\partial\eta_t(\delta)}{\partial\delta_t} + \lambda_t(\delta)(y-\mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{\partial^2\eta_t(\delta)}{\partial\beta_i\partial\delta_t}\end{aligned} \tag{B.12}$$

Considerando as segundas derivadas com respeito ao vetor γ, temos que para $j = 1,\ldots,q$

$$\begin{aligned}\frac{\partial^2\ell_\delta(\beta,\gamma)}{\partial\gamma_j\partial\delta_t} &= (y_t-\mu_t(\delta))u_t(\delta)\frac{1}{g'(\mu_t(\delta))}\frac{1}{h'(\lambda_t(\delta))}\frac{\partial\zeta_t(\delta)}{\partial\gamma_j}\frac{\partial\eta_t(\delta)}{\partial\delta_t} \\ &- \nu_t(\delta)\frac{\partial\zeta_t(\delta)}{\partial\gamma_j}\frac{\partial\zeta_t(\delta)}{\partial\delta_t} + a_t(\delta)\frac{1}{h'(\lambda_t(\delta))}\frac{\partial^2\zeta_t(\delta)}{\partial\gamma_j\partial\delta_t}.\end{aligned} \tag{B.13}$$

Avaliando as expressões (B.12) e (B.13) em $\delta_0 = (0,0,\ldots,0)^\top$ e no estimador de máxima verosimilhança $\widehat{\theta} = (\widehat{\beta}^\top,\widehat{\gamma}^\top)^\top$, obtemos $\Delta_\beta = \widehat{\tilde{X}}^\top\widehat{T}\widehat{H}\widehat{U}\mathcal{E}\widehat{\tilde{Z}}_\delta - \widehat{\tilde{X}}^\top\widehat{\Lambda}\widehat{Q}\widehat{\tilde{X}}_\delta + [\,\widehat{b}_\beta^\top\,][\,\widehat{\tilde{X}}_{\beta\delta}\,]$ e $\Delta_\gamma = \widehat{\tilde{Z}}^\top\widehat{T}\widehat{H}\widehat{U}\mathcal{E}\widehat{\tilde{X}}_\delta - \widehat{\tilde{Z}}^\top\mathcal{V}\widehat{\tilde{Z}}_\delta + [\,\widehat{b}_\gamma^\top\,][\,\widehat{\tilde{Z}}_{\gamma\delta}\,]$, em que os elementos das matrizes diagonais Q, $\mathcal{V}$ estão definidos em (3.10) e (3.11), respectivamente, e $\mathcal{E}$ está definida em (3.12). Assim, a matriz Δ fica expressa na forma apresentada em (3.23).

REFERÊNCIAS

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.; JORGENSEN, B. Some parametric models on the simplex. **Journal of Multivariate Analysis**, 39, 106–116, 1991.

BAYER, F. M.; CRIBARI–NETO, F. Model selection criteria in beta regression with varying dispersion. **ArXiv paper**, 2014. Disponível em: <http://arxiv.org/abs/1405.3718>. Acesso em: 5 jan. 2015.

BROWNLEE, K. A. **Statistical Theory and Methodology in Science and Engineering**, John Wiley and Sons, 2nd ed. London, 1965.

COOK, R. D. Assessment of local influence (with discussion). **Journal of the Royal Statistical Society B**, 48, 133–169, 1986.

COPAS, J. B. Binary regression models for contaminated data. **Journal of the Royal Statistical Society B**, 50, 225–265, 1988.

COX, D.; SNELL, E. A general definition of residuals. **Journal of the Royal Statistical Society B**, 30, 248–275, 1968.

CRIBARI–NETO, F.; ZARKOS, S. G. R: Yet another econometric programming environment. **Journal of Applied Econometrics**, 14, 319–329, 1999.

DAVISON, A. C.; GIGLI, A. Deviance residuals and normal scores plots. **Biometrika**, 76, 211–221, 1989.

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

DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals. **Journal of Computational and Graphical Statistics**, 5, 1–10, 1996.

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

ESPINHEIRA, P. L.; FERRARI, S. L. P; CRIBARI–NETO, F. Influence diagnostics in beta regression. **Computational Statistics & Data Analysis**, 52, 4417–4431, 2008b.

ESPINHEIRA, P. L.; FERRARI, S. L. P.; CRIBARI–NETO, F. Bootstrap prediction intervals in beta regressions. **Computational Statistics (Zeitschrift)**, 29, 1263–1277, 2014.

FARHRMEIR, L.; TUTZ, G. **Multivariate Statistical Modelling based on Generalized Linear Models**. New York: Springer, 1994.

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

GALEA, M.; PAULA, G. A.; BOLFARINE, H. Local influence in elliptical linear regression models. **The Statistician**, 46, 71–79, 1997.

GALEA, M.; PAULA, G. A.; URIBE–OPAZO, M. On influence diagnostics in univariate elliptical linear regression models. **Statistical Papers**, 44, 23–45, 2003.

JORGENSEN, B. **The Theory of Dispersion Models**. Chapman Hall, London, 1997.

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

LAMPORT, L . **A Document Preparation System**, 2nd. edn, Addison–Wesley, Massachusetts, 1994.

LESAFRE, E.; VERBEKE, G. Local influence in linear mixed models. **Biometrics**, 54, 570–582, 1998.

LIU, S. Z. On local influence for elliptical linear models. **Statistical Papers**, 41, 211–224, 2000.

LIU, S. Z. Local influence in multivariate elliptical linear regression models. **Linear Algebra and its Applications**, 354, 159–174, 2002.

LÓPEZ, F. O. A Bayesian Approach to Parameter Estimation in Simplex Regression Model: A compararison with Beta Regression. **Revista Colombiana de Estadística**, 36, 1–21, 2013.

MCCULLAGH, P. **Tensor Methods in Statistics**. London: Chapman and Hall, 1987.

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

MIYASHIRO, E. S. **Modelos de regressão beta e simplex para análise de proporções**. 76 p. Dissertação (Mestrado em Estatística)–Instituto de Matemática e Estatística, Universidade de São Paulo. São Paulo, 2008.

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

NOCEDAL, J.; WRIGHT, S. J. **Numerical Optimization**. New York: Springer–Verlag, 1999.

ORTEGA, E. M. M.; BOLFARINE, H.; PAULA, G. A. Influence diagnostic in generalized log–gamma regression models. **Computational Statistics & Data Analysis**, 42, 165–186, 2003.

OSPINA, R. **Modelos de regressão beta inflacionados**. Tese (Doutorado em Estatística)–Instituto de Matemática e Estatística, Universidade de São Paulo. São Paulo, 2008.

OSPINA, R.; CRIBARI–NETO, F.; VASCONCELLOS, K. L. Improved point and interval estimation for a beta regression model. **Computational Statistics & Data Analysis**, 51, 960–981, 2006.

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

PAULA, G. A. Influence diagnostics in proper dispersion models. **Australian Journal of Statistics**, 38, 307–316, 1996.

PAULA, G. A. **Modelos de Regressão com apoio computacional**. São Paulo: Instituto de Matemática e Estatística?USP, 2013.

PIERCE , D. A.; SCHAFER, D. W. Residuals in Generalized Linear Models. **Journal of the American Statistical Association**, 81, 977–986, 1986.

PREGIBON, D. Logistic regression diagnostics. **Annals of Statistics**, 9, 705–724, 1981.

PRESS *et al.* **Numerical recipes in C: The art of scientific computing**. 3nd edition. Cambridge University Press, 2007.

ROCHA, A. V.; SIMAS, A. B. Influence diagnostics in a general class of beta regression models. **Test (Madrid)**, 20, 95–119, 2010.

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

SMITHSON, M.; VERKUILEN, J. A better lemon–squeezer? Maximum likelihood regression with beta–distribuited dependent variables. **Psychological Methods**, 11, 54–71, 2006.

SONG, X. K. **Correlated Data Analysis: Modeling, Analytics, and Applications**, Springer, New York, 2007.

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

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

ST. LAURENT, R. T.; COOK, R. D. Leverage and superleverage in nonlinear regression. **Journal of the American Statistical Association**. 87, 985–990, 1992.

THOMAS, W.; COOK, R. D. Assessing influence on regression coefficients in generalized linear models. **Biometrika**, 76, 741–749, 1989.

THOMAS, W.; COOK, R. D. Assessing influence on predictions from generalized linear models. **Biometrika**, 76, 741–749, 1990.

TSAI, C. H.; WU, X. (1992). Assessing local influence in linear regression models with first–order autoregressive or heteroscedastic error struture. **Statistics and Probability Letters**, 14, 247–252.

VENEZUELA, M. K. **Equação de estimação generalizada e influência local para modelos de regressão beta com medidas repetidas**. 153 p. Tese (Doutorado em Estatística)–Instituto de Matemática e Estatística, Universidade de São Paulo. São Paulo, 2007.

WEI, B. -C. **Exponential Family Nonlinear Models**. Singapore: Springer, 1998.

WEI, B. -C.; HU, Y. -Q.; FUNG, W. -K. Generalized leverage and its applications. **Scandinavian Journal of Statistics**, 25, 25–37, 1998.

WILLIAMS, D. A. Residuals in generalized linear models. In: **Proceedings of the 12th International Conference**, Tokyo, 59–68, 1984.

WILLIAMS, D. A. Generalized linear models diagnostic using the deviance and single case deletion. **Applied Statistics**, 36, 181–191, 1987.