MODELOS LINEARES COM ERROS SLASH-ELÍPTICOS: UMA ABORDAGEM EM INFLUÊNCIA LOCAL

IZABEL CRISTINA ALCANTARA DE SOUZA

Orientador: Prof. Dr. Francisco José de Azevêdo Cysneiros

Área de Concentração: Estatística Matemática

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

Recife, fevereiro de 2009

Souza, Izabel Cristina Alcantara de
Modelos lineares com erros slash-elípticos: uma abordagem em influência local / Izabel Cristina Alcantara de Souza - Recife : O Autor, 2009.
55 folhas : il., fig., tab.

Dissertação (mestrado) – Universidade Federal de Pernambuco. CCEN. Estatística, 2009.

Inclui bibliografia.

1. Estatística Matemática. I. Título.

519.9 CDD (22.ed.) MEI2009-028

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

26 de fevereiro de 2009
(data)

Nós recomendamos que a dissertação de mestrado de autoria de

Izabel Cristina Alcântara de Souza

intitulada

"Modelos Lineares com erros Slash-Elípticos: uma Abordagem em Influência Local"

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

Klaus Leite Pinto Vasconcellos
Coordenador da Pós-Graduação em Estatística

Prof. Klaus Leite P. Vasconcellos
Coordenador
Pós-graduação em
Estatística, UFPE

Banca Examinadora:

Francisco José de Azevedo Cysneiros.
Francisco José de Azevedo Cysneiros orientador

Juvêncio Santos Nobre
Juvêncio Santos Nobre (UFCE)

Getúlio José Amorim do Amaral
Getúlio José Amorim do Amaral

Este documento será anexado à versão final da dissertação.

Aos Professores
Francisco José de Azevêdo Cysneiros
e Ronei Marcos de Moraes
pela motivação e orientação
na minha carreira.
OFEREÇO

À minha mãe Cristina
e meu irmão Victor.
DEDICO

Agradecimentos

Agradeço a Deus,
por me oferecer saúde, paz, sabedoria e apoio familiar.

Aos meus familiares,
pela compreenção e paciência durante esse mestrado.

Aos amigos,
que sempre acreditaram no meu trabalho,
e incentivam no meu crescimento profissional.

A minha família pernambucana,
que não vou citar todos por ser uma família MUITO GRANDE...

Aos professores do Departamento de Estatística da UFPE,
pelos ensinamentos, motivação e incentivo.

Aos membros da Banca Examinadora,
pela contribuição, elogios e dedicação no exame desta dissertação.

À CAPES, pelo apoio financeiro.

Resumo

As distribuições de probabilidade geralmente utilizadas para modelagem de dados quando há simetria no comportamento dos erros do modelo são as pertencentes a família elíptica, sendo a distribuição normal a mais utilizada na literatura dentre todas as distribuições elípticas. Neste trabalho abordaremos uma outra classe de distribuições que apresenta a propriedade de simetria proposta recentemente por Gómez, Quintana e Torres (2007), denominada de distribuição slash-elíptica. A distribuição slash-elíptica apresenta como principal característica uma maior flexibilidade quanto ao grau da curtose frente a distribuição elíptica, além de conter a família elíptica como um caso limite. Propomos uma metodologia de estimação, testes de hipóteses, análise de resíduos e diagnóstico para a classe de modelos lineares com erros slash-elípticos com parâmetro q conhecido. Apresentamos duas aplicações para exemplificar a metodologia proposta. O primeiro conjunto de dados analisados refere-se aos dados de salinidade do rio *Pamlico Sound* na Carolina do Norte - EUA apresentado por Rupppert e Carroll (1980). O segundo conjunto de dados analisados refere-se a um experimento de 21 dias de observações de um planta sujeita a oxidação de amônia a ácido nítrico. Para os dois conjuntos de dados, consideramos os modelos elípticos normal e t-Student, e os modelos slash-normal e slash-t-Student, com o objetivo de realizar uma comparação empírica entre os mesmos.

Palavras-chaves: distribuição slash-elíptica, influência local, diagnóstico.

Abstract

We propose a linear regression model with distribution of errors slash-elliptical. The slash-elliptical distribution was recently proposed by Gómez, Quintana and Torres (2007). The main feature of the slash-elliptical distribution is to have greater flexibility in the degree of kurtosis front the elliptical distributions, and include the elliptical family as a limit case. We developed the methodology of estimation, hypothesis testing and analysis of residuals. Some diagnostic measures are introduced. Finally, practical applications that employ real data are presented to illustrate the proposed methodology.

Keywords: slash-elliptical distribution, local influence, diagnostic.

Sumário

Lista de Figuras

Lista de Tabelas

CAPÍTULO 1

Introdução

1.1 Formulação do problema e definição dos objetivos

A regressão linear é comumente utilizada como ferramenta estatística para modelagem de dados. Em geral, assume-se que os erros do modelo seguem distribuição normal. Mesmo quando essa suposição não é verificada, costuma-se realizar transformações, tal como a proposta por Box e Cox (1964), sobre a variável resposta para atingir a normalidade. Nem sempre a transformação de Box e Cox garante essa suposição, resultando ainda que esse tipo de modelagem pode ser altamente influenciada por observações extremas produzindo modelos inadequados, como é bem conhecido. Uma solução é utilizar modelos robustos supondo uma distribuição de probabilidade mais adequada para os erros do modelo. Quando há simetria no comportamento dos erros, as distribuições mais utilizadas numa abordagem robusta são as pertencentes a família de distribuição com contorno elíptico, também denominada de simétrica no caso univariado. Neste trabalho abordaremos uma classe de distribuições, mais geral, que também apresenta a propriedade de simetria denominada de distribuição slash-elíptica (GÓMEZ; QUINTANA; TORRES, 2007; GÓMEZ; VENEGAS, 2008).

Segundo Gómez, Quintana e Torres (2007), a distribuição slash-elíptica pode ser considerada como um caso particular de modelos de mistura-escala da distribuição elíptica com a distribuição $U(0,1)$, modificando a distribuição elíptica para flexibilizar o grau da curtose. Desta forma, a distribuição slash-elíptica apresenta uma maior flexibilidade quanto ao grau de curtose quando comparado a distribuição

elíptica, como foi observado por Genc (2007). Outra vantagem da classe de distribuições slash-elíptica é que esta contém a família elíptica como um caso limite.

Frente a essa nova família de distribuições, o objetivo geral deste estudo é o desenvolvimento do modelo de regressão linear com erros slash-elíptico, propondo uma metodologia de estimação e análise de diagnóstico sob o enfoque de influência local e alavancagem generalizada. A seguir relacionamos os objetivos específicos:

1. desenvolvimento da metodologia de estimação e testes assintóticos da significância dos parâmetros para a classe de modelos lineares com erros slash-elíptico;

2. propor resíduos e medidas de diagnóstico sob o enfoque de alavancagem generalizada e influência local sob vários esquemas de perturbação.

1.2 Apresentação dos capítulos

Nossa principal contribuição encontra-se nos capítulos 2 e 3. No capítulo 2 apresentamos o desenvolvimento do processo iterativo de estimação via método de máxima verossimilhança dos parâmetros no modelo de regressão linear com erros slash-elíptico. Testes da razão de verossimilhanças, Wald e escore para testar a significância dos parâmetros foram propostos. Um resíduo empírico e o resíduo componente de desvio foram propostos e um estudo de simulação sobre estes resíduos foi realizado com o objetivo de verificar as suas propriedades empíricas.

No capítulo 3 encontra-se o desenvolvimento da metodologia de análise de diagnóstico sobre o enfoque a medida de alavancagem proposta por Wei, Hu e Fung (1998) e influência local para a classe de modelos lineares com erros slash-elípticos. Consideramos duas medidas de influência local: uma baseado no afastamento da verossimilhança proposto por Cook (1986) e uma segunda distância baseada no resíduo de Pearson proposta por Thomas e Cook (1990). A análise de influência local baseada nesta última medida é denominada de influência local na predição e consideramos dois esquemas de perturbações: na resposta e nos regressores. Para análise de influência local pelo afastamento da verossimilhança, consideramos as perturbações na escala e nos casos.

No capítulo 4 apresentamos duas aplicações a conjunto de dados reais, para exemplificar a metodologia proposta. O primeiro conjunto de dados analisados refere-se aos dados de salinidade do rio *Pamlico Sound* na Carolina do Norte - EUA apresentado por Rupppert e Carroll (1980). O segundo conjunto de dados analisados refere-se a um experimento de 21 dias de observações de um planta sujeita a oxidação de amônia a ácido nítrico. Para os dois conjuntos de dados, consideramos os modelos elípticos normal e

t-Student, e os modelos slash-normal e slash-t-Student, a fim de fazer uma comparação empírica entre os mesmos.

1.3 Alguns resultados da família de distribuições elíptica

A família de distribuições elíptica pode ser considerada como uma generalização da distribuição normal, tanto no caso univariado como no multivariado. No caso univariado, está família de distribuição também é denominada de classe simétrica de distribuições. Neste estudo enfocaremos duas distribuições simétricas: normal e t-Student. O modelo de regressão linear, no qual a distribuição dos erros é normal, são os mais conhecidos na literatura de modelagem. Já o modelo de regressão linear com distribuição dos erros t-Student foi proposto por Lange, Little e Taylor (1989) e se enquadram na área de modelagem robusta. Mais detalhes sobre as distribuições simétricas podem ser encontradas em Fang e Zhang (1990) e sobre modelos de regressão linear com erros simétricos podem ser encontrados em Cysneiros (2004) e Cysneiros, Paula e Galea (2005), por exemplo.

Definição 1: Uma variável aleatória W com suporte em $\Re$ é denominada de variável aleatória simétrica com parâmetros de posição $\mu \in \Re$ e escala $\phi > 0$, denotada por $W \sim El(\mu, \phi; g(\cdot))$, se sua densidade é da forma

$$f_W(w; \mu, \phi) = \frac{1}{\sqrt{\phi}} g\left(\frac{(w-\mu)^2}{\phi}\right) \tag{1.1}$$

para alguma função geradora de densidade $g(\cdot)$, com $g(u) > 0$ para $u > 0$, $\int_0^\infty u^{-1/2} g(u) du = 1$ e $u = \frac{(w-\mu)^2}{\phi}$.

Considere a variável aleatória $W \sim El(\mu, \phi; g(\cdot))$, então os principais resultados da classe simétrica de distribuições são:

1. A distribuição de qualquer combinação linear de uma variável aleatória simétrica também é simétrica, isto é, se $W \sim El(\mu, \phi; g(\cdot))$ então $a + bW \sim El(a + b\mu, b^2\phi; g(\cdot))$, em que $a, b \in \Re$ com $b \neq 0$. Fazendo $a = -\frac{\mu}{\sqrt{\phi}}$ e $b = \frac{1}{\sqrt{\phi}}$, temos que $V = a + bW$ tem distribuição simétrica padrão denotada por $El(0, 1; g(\cdot))$;

2. A função característica da distribuição simétrica é

   $$\zeta_W(t) = E(e^{itw}) = e^{it\mu} \varphi^{(1)}(t^2\phi),$$

   em que $t \in \Re$ para alguma função φ, com $\varphi(u) \in \Re$ para $u > 0$, e $\varphi^{(m)}$ é a m-ésima derivada da função φ;

3. O k-ésimo momento central da distribuição $El(\mu,\phi;g(\cdot))$ é $M_{W(k)} = E(W^k) = i^{-k}\zeta_V^{(k)}(0)$, em que $\zeta_V^{(k)}(0)$ denota a k-ésima derivada de $\zeta_V(t)$ avaliada em $t=0$. Segundo Berkame e Bentler (1986), $M_W(k) = 0$ se k for ímpar e $M_W(2m) = \frac{(2m)!}{2^m m!} M_W(2)^m (h(m)+1)$ se k for par, em que $k = 2m$ para $m = 1,2,\ldots$ e $h(m) = \frac{\varphi^{(m)}(0)}{(\varphi^{(1)}(0))^m} - 1$.

4. Pelo item 3, quando existem, tem-se que o valor esperado, variância e curtose de W são dados respectivamente por:

$$\begin{aligned} E(W) &= \mu, \\ Var(W) &= -2\varphi^{(1)}(0)\phi, \\ \gamma_W &= 3(h(2)+1). \end{aligned}$$

Na Tabela 1.1, tem-se a função geradora de densidade $g(u)$, a função $-2\varphi^{(1)}(0)$ e o coeficiente de curtose das distribuições normal e t-Student, que serão consideradas neste estudo.

Tabela 1.1: Funções geradoras de densidade, função $-2\varphi^{(1)}(0)$ e Curtose das distribuições normal e t-Student

Distribuição	$g(u)$	$-2\varphi^{(1)}(0)$	γ
Normal	$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{u}{2}\right)$	1	3
t-Student	$\frac{\nu^{\nu/2}}{B(1/2,\nu/2)}(\nu+u)^{-\frac{\nu+1}{2}}$	$\frac{\nu}{(\nu-2)}$, para $\nu > 2$	$3+\frac{6}{\nu-4}$, para $\nu > 4$

Nota: $u > 0$, $\nu > 0$ e com $B(a,b)$ denotando a função beta.

1.4 Alguns resultados da família de distribuições slash-elíptica

Definição 2: Uma variável aleatória, Y, tem distribuição slash-elíptica com parâmetro de posição $\mu \in \Re$ e escala $\phi > 0$ se $Y = \mu + \sqrt{\phi}\frac{V}{U^{1/q}}$, em que V e U são variáveis aleatórias independentes, V tem distribuição simétrica padrão, U tem distribuição uniforme no intervalo $(0,1)$, e q é o parâmetro específico da distribuição slash-elíptica.

Definição 3: Uma variável aleatória Y com suporte nos $\Re$ é denominada de variável slash-elíptica com parâmetro de posição $\mu \in \Re$ e escala $\phi > 0$, denotada por $Y \sim SEl(\mu_i,\phi,q;g(\cdot))$, se sua densidade é da forma

$$f(y;\mu,\phi,q) = \frac{1}{\sqrt{\phi}} \begin{cases} \frac{q}{2|z|^{q+1}} H(z^2) & , z \neq 0 \\ \frac{q}{q+1} g(0) & , z = 0 \end{cases} \tag{1.2}$$

em que $z = (y-\mu)/\sqrt{\phi}$ e $H(z^2) = \int_0^{z^2} t^{(q-1)/2} g(t)dt$, para alguma função geradora de densidade $g(\cdot)$, com $g(t) > 0$ para $t > 0$ e $\int_0^{\infty} t^{-1/2} g(t)dt = 1$.

Qualquer uma das definições 2 e 3 determinam uma variável aleatória slash-elíptica, no entanto, a partir da definição 2 temos um método para gerar variáveis aleatórias slash-elíptica com base nas distribuições simétrica e uniforme. Já a definição 3 caracteriza uma variável aleatória slash-elíptica pela sua densidade.

Considere a variável aleatória $Y \sim SEl(\mu, \phi, q; g(\cdot))$, então os principais resultados da família de distribuições slash-elíptica são:

1. A distribuição de qualquer combinação linear de uma variável aleatória slash-elíptica também é slash-elíptica, isto é,

$$\text{se } Y \sim SEl(\mu, \phi, q; g(\cdot)) \text{ então } a + bY \sim SEl(a + b\mu, b^2\phi, q; g(\cdot)), \tag{1.3}$$

   em que $a, b \in \Re$ com $b \neq 0$. Fazendo $a = -\frac{\mu}{\sqrt{\phi}}$ e $b = \frac{1}{\sqrt{\phi}}$, temos que $Z = a + bY$ tem distribuição slash-elíptica padrão denotada por $SEl(0, 1, q; g(\cdot))$.

2. Caso o k-ésimo momento central da distribuição slash-elíptica padrão é dado por

$$M_Z(k) = E(Z^k) = \begin{cases} \frac{q}{q-k} a_{k/2} & \text{, se k é par} \\ 0 & \text{, se k é impar} \end{cases}, \quad k = 1, 2, \ldots \tag{1.4}$$

   em que $a_{k/2} = \int_{-\infty}^{\infty} t^k g(t^2)dt < \infty, \quad k = 1, 2, \ldots.$

3. O k-ésimo momento central da distribuição $SEl(\mu, \phi, q; g(\cdot))$ é dado por (1.5)

$$M_Y(k) = E(Y^k) = \sum_{r=0}^{k} \binom{k}{r} \sigma^r \mu^{k-r} M_Z(r) \tag{1.5}$$

4. De (1.4) e (1.5), temos que, quando existem, o valor esperado, a variância e a curtose de Y são dados respectivamente por:

$$E(Y) = \mu, \tag{1.6}$$

$$Var(Y) = \frac{q}{q-2} a_1 \phi, \quad q > 2, \tag{1.7}$$

$$\gamma_Y = \frac{(q-2)^2 a_2}{q(q-4)a_1^2}, \quad q > 4 \tag{1.8}$$

em que

$$a_1 = \int_{-\infty}^{\infty} t^2 g(t^2) dt \tag{1.9}$$

e

$$a_2 = \int_{-\infty}^{\infty} t^4 g(t^2) dt; \tag{1.10}$$

5. Segundo Gómez, Quintana e Torres (2007), Gómez e Venegas (2008), um estimador para q obtido pelo método dos momentos é dado por

$$\widehat{q} = 2\left(1 + a_1\sqrt{\frac{\widehat{\gamma}_Y}{\widehat{\gamma}_Y a_1^2 - a_2}}\right), \quad \text{para } q > 2 \text{ e } \widehat{\gamma}_Y a_1^2 - a_2 > 0 \tag{1.11}$$

em que $\widehat{\gamma}_Y$ é o estimador da curtose pelo método dos momentos. Logo q é uma função não linear da curtose e das constantes a_1 e a_2.

1.5 Comparação das densidades elíptica e slash-elíptica (caso univariado)

Para ilustrar a comparação entre as distribuições elíptica e slash-elíptica, considere as Figuras 1.1 (normal e slash-normal) e 1.2 (t-Student e slash-t-Student), todas com parâmetro de posição igual a zero e escala igual a um. Ao comparar as densidades elíptica e slash-elíptica para um mesmo núcleo, isto é, normal com slash-normal e t-Student com slash-t-Student, verifica-se que ambas possuem os mesmos parâmetros de posição e escala, que a forma das curvas são semelhantes, mas que a densidade slash-elíptica é sempre mais platicúrtica (achatada) que a densidade elíptica. Essa diferença no achatamento das densidades, que é caracterizada pela curtose, diminui a medida que o parâmetro q da slash-elíptica tende a infinito. Gómez, Quintana e Torres (2007), Gómez e Venegas (2008) afirma que a classe de distribuições slash-elíptica contém a família elíptica como um caso limite quando $q \to \infty$. Portanto, o parâmetro q da slash-elíptica é responsável pela diferença na curtose entre as distribuições elíptica e slash-elíptica e o modelo slash-elíptico permite uma maior flexibilidade no grau da curtose quando comparado ao modelo elíptico.

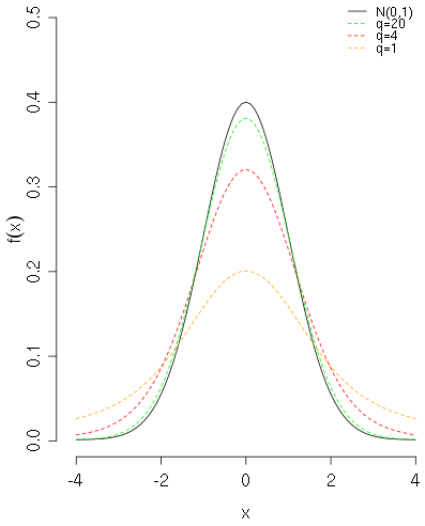
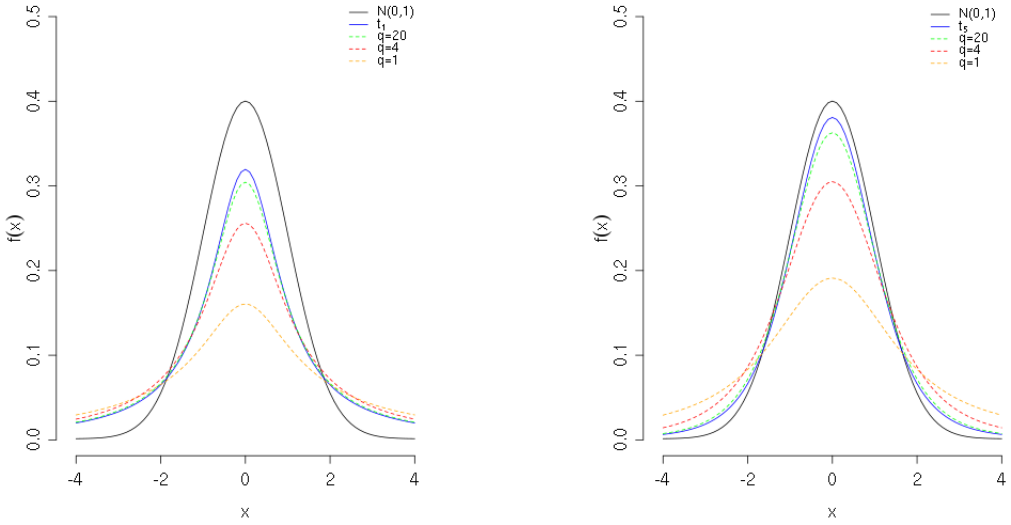


Figura 1.1: Distribuições normal (linha cheia preta) e slash-normal (linhas tracejadas laranja $q = 1$, vermelho $q = 5$, verde $q = 20$)



(a) Com 1 grau de liberdade

(b) Com 5 graus de liberdade

Figura 1.2: Distribuições normal (linha cheia preta), t-Student (linha cheia azul) e slash-t-Student (linhas tracejadas laranja $q = 1$, vermelho $q = 5$, verde $q = 20$)

1.6 Suporte Computacional

Todas as funções para estimação, testes de hipóteses, análise de resíduos e diagnóstico para a classe de modelos lineares com erros slash-elípticos, bem como as simulações e aplicações, foram desenvolvidas

a partir de programas construídos usando a linguagem de programação matricial Ox em sua versão 4.1 para sistema operacional Ubuntu GNU/Linux versão 8.10. O Ox foi desenvolvido por Jurgen Doornik em 1994 na Universidade de Oxford, Inglaterra e é distribuído gratuitamente para uso acadêmico em http://www.doornik.com (DOORNIK, 1999). Os gráficos apresentados foram produzidos utilizando o ambiente de programação R em sua versão 2.7.1 para o sistema operacional Ubuntu GNU/Linux versão 8.10. Esta linguagem foi criada por Ross Ihaka e Robert Gentleman na Universidade de Auckland (R Development Core Team, 2008). O R se encontra disponível gratuitamente em http://www.r-project.org.

CAPÍTULO 2

Modelo slash-elíptico

2.1 Introdução

A área de modelagem robusta para família de distribuições slash-elíptica está atualmente em desenvolvimento. Estimadores de máxima verossimilhança para os parâmetros de posição, escala e curtose da distribuição slash-exponencial potência foram propostos por Genc (2007). Apresentamos neste capítulo a metodologia de estimação e análise dos resíduos para a classe de modelos de regressão linear com distribuição dos erros pertencentes a família slash-elíptica com q conhecido ou fixado.

Sejam $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ variáveis aleatórias independentes, em que $\varepsilon_i \sim SEl(0, \phi, q; g(\cdot))$. O modelo de regressão linear com erros slash-elíptico é definido por

$$y_i = \mathbf{x}_i^t \beta + \varepsilon_i \quad i = 1, 2, \ldots, n, \tag{2.1}$$

em que para cada i-ésima observação, y_i é a variável resposta, $\mathbf{x}_i^t = (1, x_{i2}, ..., x_{ip})$ é um vetor $(p \times 1)$ de regressores e $\beta = (\beta_1, \ldots, \beta_p)^t$ é o vetor $(p \times 1)$ de parâmetros desconhecidos.

Outra forma de representar o modelo (2.1) é através da notação matricial

$$\mathbf{y} = \mathrm{X}\beta + \varepsilon,$$

em que

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathrm{X} = \begin{pmatrix} 1 & x_{12} & \cdots & x_{1p} \\ 1 & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n2} & \cdots & x_{np} \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Pelo resultado (1.3), y_i tem sua função densidade de probabilidade expressa por

$$f(y_i;\mu_i,\phi) = \frac{1}{\sqrt{\phi}} \begin{cases} \frac{q}{2|z_i|^{q+1}} H(z_i^2) & , z_i \neq 0 \\ \frac{q}{q+1} g(0) & , z_i = 0 \end{cases} \tag{2.2}$$

em que $\mu_i = \mathbf{x}_i^t \beta$, $z_i = (y_i - \mu_i)/\sqrt{\phi}$ e $H(z_i^2) = \int_0^{z_i^2} t^{(q-1)/2} g(t)dt$. Como as variáveis aleatórias $y_1, y_2, \ldots, y_n$ são independentes, temos que a sua função de densidade conjunta é dada por:

$$\begin{aligned} f(\mathbf{y};\mu,\phi) &= \prod_{i=1}^{n} f(y_i;\mu_i,\phi) \\ &= \phi^{-\frac{n}{2}} \left(\prod_{i \notin A} \frac{q}{q+1} g(0) \right) \left(\prod_{i \in A} \frac{q}{2|z_i|^{q+1}} H(z_i^2) \right) \\ &= \phi^{-\frac{n}{2}} \left(\frac{q}{q+1} g(0) \right)^{n-n_A} \left(\frac{q}{2} \right)^{n_A} \left(\prod_{i \in A} \frac{H(z_i^2)}{|z_i|^{q+1}} \right) \end{aligned}$$

em que $A = \{i : z_i \neq 0\}$ e n_A é o número de elementos do conjunto A.

2.2 Estimação dos parâmetros

Para estimar o vetor de parâmetros $\theta = (\beta^t, \phi)^t$, pode-se utilizar o método de máxima verossimilhança, que consiste em maximizar a função de verossimilhança do modelo slash-elíptico com respeito a θ. Trabalhamos com o logaritmo da função de verossimilhança, pois maximizar o logaritmo de uma função é em geral mais simples e produz os mesmos resultados da maximização da função original. A função de verossimilhança de θ é dada por

$$l(\theta) = \phi^{-\frac{n}{2}} \left(\frac{q}{q+1} g(0) \right)^{n-n_A} \left(\frac{q}{2} \right)^{n_A} \left(\prod_{i \in A} \frac{H(z_i^2)}{|z_i|^{q+1}} \right) \tag{2.3}$$

Aplicando o logaritmo a equação (2.3), obtemos a função de log-verossimilhança de θ para o modelo

slash-elíptico

$$\begin{aligned} L(\theta) &= \log(l(\theta)) \\ &= -\frac{n}{2}\log(\phi) + (n-n_A)\log\left(\frac{qg(0)}{q+1}\right) + n_A\log(\frac{q}{2}) + \sum_{i\in A}\left[\log(H(z_i^2)) - (q+1)\log|z_i|\right] \\ &= -\frac{n}{2}\log(\phi) + \sum_{i\in A} b(z_i,q) + C, \end{aligned} \tag{2.4}$$

em que $b(z_i,q) = \log(H(z_i^2)) - (q+1)\log|z_i|$ é uma função que depende de θ e $C = (n-n_A)\log\left(\frac{q}{q+1}g(0)\right) + n_A\log(\frac{q}{2})$ é constante em relação a θ.

Para encontrar o valor que maximiza (2.4), temos que resolver as equações $\mathrm{U}_\beta = \frac{\partial L(\theta)}{\partial \beta} = 0$ e $\mathrm{U}_\phi = \frac{\partial L(\theta)}{\partial \phi} = 0$, em que U_β e U_ϕ são as funções escore de β e ϕ, respectivamente. A função escore de β é dada por:

$$\begin{aligned} \mathrm{U}_\beta &= \frac{\partial}{\partial \beta}\left(-\frac{n}{2}\log(\phi) + \sum_{i\in A} b(z_i,q) + C\right) \\ &= \sum_{i\in A}\frac{\partial b(z_i,q)}{\partial z_i}\frac{\partial z_i}{\partial \beta} \\ &= \sum_{i\in A} b'(z_i,q)\left(-\frac{\mathbf{x}_i}{\sqrt{\phi}}\right) \\ &= -\frac{1}{\sqrt{\phi}}\mathrm{X}^t\mathbf{b}'_{(\mathbf{z},q)} \end{aligned} \tag{2.5}$$

em que $b'(z_i,q) = -\frac{(q+1)}{z_i} + \frac{2z_i^q g(z_i^2)}{H(z_i^2)}$ e $\mathbf{b}'_{(\mathbf{z},q)} = (b(z_1,q), b(z_2,q), \ldots, b(z_n,q))^t$ é um vetor $(n\times 1)$. A função escore de ϕ é dada por:

$$\begin{aligned} \mathrm{U}_\phi &= \frac{\partial}{\partial \phi}\left(-\frac{n}{2}\log(\phi) + \sum_{i\in A} b(z_i,q) + C\right) \\ &= -\frac{n}{2\phi} + \sum_{i\in A}\frac{\partial b(z_i,q)}{\partial z_i}\frac{\partial z_i}{\partial \phi} \\ &= -\frac{n}{2\phi} + \sum_{i\in A} b'(z_i,q)\left(-\frac{z_i}{2\phi}\right) \\ &= -\frac{n}{2\phi} - \frac{1}{2\phi}\sum_{i\in A} z_i b'(z_i,q). \end{aligned} \tag{2.6}$$

Como para a função de log-verossimilhança (2.4) não é possível encontrar expressões analíticas para solução das equações $\mathrm{U}_\beta = 0$ e $\mathrm{U}_\phi = 0$, podemos usar métodos iterativos de otimização não-linear tal como *Newton-Raphson*, Escore de *Fisher*, BFGS entre outros. Neste estudo usaremos o método quase-Newton BFGS, que é um dos mais conhecidos métodos quase-Newton e que foi proposto independente-

mente pelos autores Broyden (1970), Fletcher (1970), Goldfarb (1970) e Shanno (1970). A metodologia do método BFGS, bem como de outros métodos de otimização pode ser encontrada em Press, Vetterling e Flannery (1992), por exemplo.

Os métodos quase-Newton de otimização são um conjunto de procedimentos numéricos que visam minimizar uma função $f(\mathbf{x})$ em relação a $\mathbf{x}$, em que $f(\mathbf{x})$ é uma função real não-linear e $\mathbf{x} \in \Re^n$. Caso o objetivo seja maximizar a função $f(\mathbf{x})$, basta multiplicá-la por (-1), pois o problema de maximizar $f(\mathbf{x})$ é equivalente ao problema de minimizar $-f(\mathbf{x})$. Os métodos quase-Newton baseiam-se em algoritmos iterativos, no qual, em cada k-ésima iteração, uma aproximação de $\mathbf{x}_k$ e de uma matriz H_k de dimensão $(n \times n)$ são calculadas. Estes métodos requerem que H_k seja positiva definida e que satisfaça a equação:

$$\mathrm{H}_{k+1}\Delta\mathbf{g}_k = \lambda_k(\mathbf{x}_{k+1} - \mathbf{x}_k)$$

em que $\Delta\mathbf{g}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$, $\mathbf{g}_k = \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}|_{\mathbf{x}=\mathbf{x}_k}$ e λ_k é escolhido para satisfazer determinadas condições de busca em linha do método quase-Newton considerado. A seguir temos os procedimentos do algoritmo BFGS:

1. Faça $k = 0$ e atribua valores inciais para o vetor $\mathbf{x}_0$ e para matriz H_0;
2. Se $\mathbf{g}_k = 0$, o processo iterativo termina e $\mathbf{x}_k$ é o mínimo, senão $\mathbf{d}_k = -\mathrm{H}_k\mathbf{g}_k$;
3. Calcule:

$$\begin{aligned} \lambda_k &= \arg\min_{\lambda>0} f(\mathbf{x}_k + \lambda\mathbf{d}_k) \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \lambda\mathbf{d}a_k\mathbf{d}_k \end{aligned}$$

4. Atualize a matriz $\mathbf{H}$:

$$\mathrm{H}_{k+1} = \mathrm{H}_k + \left(1 + \frac{\Delta\mathbf{g}_k^t \mathrm{H}_k \Delta\mathbf{g}_k}{\Delta\mathbf{g}_k^t \Delta\mathbf{x}_k}\right)\left(\frac{\Delta\mathbf{x}_k \Delta\mathbf{x}_k^t}{\Delta\mathbf{x}_k^t \Delta\mathbf{g}_k}\right) - \frac{\mathrm{H}_k\Delta\mathbf{g}_k\Delta\mathbf{x}_k^t + (\mathrm{H}_k\Delta\mathbf{g}_k\Delta\mathbf{x}_k^t)^t}{\Delta\mathbf{g}_k^t\Delta\mathbf{x}_k} \tag{2.7}$$

   em que $\Delta\mathbf{x}_k = \lambda_k\mathbf{d}_k$;
5. Faça $k = k+1$ e volte ao passo 2.

2.3 Matriz de Informação de Fisher Observada

Para obter a matriz de informação de Fisher observada de θ como definida em

$$\ddot{L}_{\hat{\theta}\hat{\theta}} = -\left(\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta^t}\right)^{-1}|_{\theta=\widehat{\theta}} = -\begin{pmatrix} J_{\beta\beta} & J_{\beta\phi} \\ J_{\beta\phi}^t & J_{\phi\phi} \end{pmatrix}^{-1}|_{\theta=\widehat{\theta}} = \begin{pmatrix} \ddot{L}_{\hat{\beta}\hat{\beta}} & \ddot{L}_{\hat{\beta}\hat{\phi}} \\ \ddot{L}_{\hat{\phi}\hat{\beta}} & \ddot{L}_{\hat{\phi}\hat{\phi}} \end{pmatrix}, \tag{2.8}$$

calculamos as matrizes hessianas de segunda derivadas $J_{\beta\beta} = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial\beta\partial\beta^t}$, $J_{\phi\phi} = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial\phi\partial\phi}$ e $J_{\beta\phi} = \frac{\partial^2 L(\boldsymbol{\theta})}{\partial\beta\partial\phi}$, dadas a seguir.

$$\begin{aligned} J_{\beta\beta} &= \frac{\partial}{\partial\beta^t}\left(-\frac{1}{\sqrt{\phi}}\sum_{i\in A}\mathbf{x}_i b'(z_i,q)\right) \\ &= -\frac{1}{\sqrt{\phi}}\sum_{i\in A}\mathbf{x}_i\frac{\partial b'(z_i,q)}{\partial z_i}\frac{\partial z_i}{\partial\beta} \\ &= -\frac{1}{\sqrt{\phi}}\sum_{i\in A}\mathbf{x}_i b''(z_i,q)\left(-\frac{\mathbf{x}_i}{\sqrt{\phi}}\right) \\ &= -\frac{1}{\sqrt{\phi}}\sum_{i\in A}\mathbf{x}_i b''(z_i,q)\mathbf{x}_i^t \\ &= \frac{1}{\phi}\mathrm{X}^t B''_{(\mathbf{z},q)}\mathrm{X} \end{aligned} \tag{2.9}$$

em que $b''(z_i,q) = \frac{q+1}{z_i^2} + \frac{2qz_i^{q-1}g(z_i^2)}{H(z_i^2)} + \frac{4z_i^{q+1}g'(z_i^2)}{H(z_i^2)} - \frac{4z_i^{2q}g(z_i^2)^2}{H(z_i^2)^2}$ e $B''_{(\mathbf{z},q)}$ =diag$(b''(z_1,q), b''(z_2,q), \ldots, b''(z_n,q))$ na diagonal.

$$\begin{aligned} J_{\phi\phi} &= \frac{\partial}{\partial\phi}\left(-\frac{n}{2\phi} - \frac{1}{2\phi}\sum_{i\in A} z_i b'(z_i,q)\right) \\ &= \frac{n}{2\phi^2} + \frac{1}{2\phi^2}\sum_{i\in A} z_i b'(z_i,q) - \frac{1}{2\phi}\sum_{i\in A}\left(z_i b''(z_i,q) + b'(z_i,q)\right)\left(-\frac{z_i}{2\phi}\right) \\ &= \frac{n}{2\phi^2} + \frac{1}{4\phi^2}\sum_{i\in A} z_i\left[z_i b''(z_i,q) + 3b'(z_i,q)\right] \\ &= \frac{1}{4\phi^2}\left(2n + \mathbf{z}^t\mathbf{c}_{(\mathbf{z},q)}\right) \end{aligned} \tag{2.10}$$

em que $\mathbf{c}_{(\mathbf{z},q)} = (c(z_1,q), c(z_2,q), \ldots, c(z_n,q))^t$ e $c_{(z_i,q)} = z_i b''(z_i,q) + 3b'(z_i,q)$.

$$\begin{aligned} J_{\beta\phi} &= \frac{\partial}{\partial\phi}\left(-\frac{1}{\sqrt{\phi}}\sum_{i\in A}\mathbf{x}_i b'(z_i,q)\right) \\ &= \frac{1}{2\phi^{3/2}}\sum_{i\in A}\mathbf{x}_i b'(z_i,q) - \frac{1}{\sqrt{\phi}}\sum_{i\in A}\mathbf{x}_i b''(z_i,q)\left(-\frac{z_i}{2\phi}\right) \\ &= \frac{1}{2\phi^{3/2}}\sum_{i\in A}\mathbf{x}_i\left(b'(z_i,q) + b''(z_i,q)z_i\right) \\ &= \frac{1}{2\phi^{3/2}}\mathrm{X}^t\mathbf{m}_{(\mathbf{z},q)} \end{aligned} \tag{2.11}$$

em que $\mathbf{m}_{(\mathbf{z},q)} = (m(z_1,q), m(z_2,q), \ldots, m(z_n,q))^t$ e $m_{(z_i,q)} = b'(z_i,q) + b''(z_i,q)z_i$.

Substituindo as equações (2.9)-(2.11) em (2.8) e avaliando em $\theta = \hat{\theta}$, temos que

$$\begin{aligned}
\ddot{L}_{\hat{\theta}\hat{\theta}} &= \begin{pmatrix} -\frac{1}{\hat{\phi}}\mathrm{X}^t B''_{(\hat{\mathbf{z}},q)}\mathrm{X} & -\frac{1}{2\hat{\phi}^{3/2}}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)} \\ -\frac{1}{2\hat{\phi}^{3/2}}\mathbf{m}^t_{(\hat{\mathbf{z}},q)}\mathrm{X} & -\frac{1}{4\hat{\phi}^2}\left(2n+\hat{\mathbf{z}}^t\mathbf{c}_{(\hat{\mathbf{z}},q)}\right) \end{pmatrix}^{-1} \\
&= \begin{pmatrix} -\frac{1}{\hat{\phi}}D & -\frac{1}{2\hat{\phi}^{3/2}}\mathbf{d} \\ -\frac{1}{2\hat{\phi}^{3/2}}\mathbf{d}^t & -\frac{1}{4\hat{\phi}^2}e \end{pmatrix}^{-1} \\
&= \begin{pmatrix} -\frac{1}{\hat{\phi}}\left(D-\frac{1}{e}\mathbf{d}\mathbf{d}^t\right)^{-1} & \frac{2}{\sqrt{\hat{\phi}}e}\left(D-\frac{1}{e}\mathbf{d}\mathbf{d}^t\right)^{-1}\mathbf{d} \\ \frac{2}{\sqrt{\hat{\phi}}e}\mathbf{d}^t\left(D-\frac{1}{e}\mathbf{d}\mathbf{d}^t\right)^{-1} & -\frac{4\hat{\phi}^2}{e}\left(1+\frac{1}{e\hat{\phi}^2}\mathbf{d}^t\left(D-\frac{1}{e}\mathbf{d}\mathbf{d}^t\right)^{-1}\mathbf{d}\right) \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1}{\hat{\phi}}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1} & \frac{2}{\sqrt{\hat{\phi}}e}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)} \\ \frac{2}{\sqrt{\hat{\phi}}e}\mathbf{m}^t_{(\hat{\mathbf{z}},q)}\mathrm{X}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1} & -\frac{4\hat{\phi}^2}{e}\left(1+\frac{1}{e\hat{\phi}^2}\mathbf{m}^t_{(\hat{\mathbf{z}},q)}\mathrm{X}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)}\right) \end{pmatrix} \qquad (2.12)
\end{aligned}$$

em que $D = \mathrm{X}^t B''_{(\hat{\mathbf{z}},q)}\mathrm{X}$, $\mathbf{d}^t = \mathbf{d} = \mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)}$, $e = \left(2n+\hat{\mathbf{z}}^t\mathbf{c}_{(\hat{\mathbf{z}},q)}\right)$ e $V_{(\hat{\mathbf{z}},q)} = B''_{(\hat{\mathbf{z}},q)} - \frac{1}{e}\mathbf{m}_{(\hat{\mathbf{z}},q)}\mathbf{m}^t_{(\hat{\mathbf{z}},q)}$. Desta forma, as matrizes parciais de informação de Fisher observadas são:

$$\ddot{L}_{\hat{\beta}\hat{\beta}} = -\frac{1}{\hat{\phi}}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1} \qquad (2.13)$$

$$\ddot{L}_{\hat{\beta}\hat{\phi}} = \ddot{L}^t_{\hat{\phi}\hat{\beta}} = \frac{2}{\sqrt{\hat{\phi}}e}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)} \qquad (2.14)$$

$$\ddot{L}_{\hat{\phi}\hat{\phi}} = -\frac{4\hat{\phi}^2}{e}\left(1+\frac{1}{e\hat{\phi}^2}\mathbf{m}^t_{(\hat{\mathbf{z}},q)}\mathrm{X}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)}\right) \qquad (2.15)$$

2.4 Qualidade do Ajuste e Testes de Hipóteses

Uma vez estimado o modelo slash-elíptico para um q fixado ou conhecido, pode-se verificar a qualidade do ajuste através de critérios de informação, tais como o Critério de Informação de Akaike - *AIC* (AKAIKE, 1974) e o Critério de Informação Bayesiano - *BIC* (SCHWARZ, 1978), que são definidos pelas equações (2.16) e (2.17) respectivamente. Segundo (KUHA, 2004), o critério *BIC* apresenta uma melhor performace que o *AIC*, sendo o *BIC* o mais indicado em caso de discordância entre estes critérios. No caso do parâmetro q não ser conhecido, uma solução é estimar o modelo para um conjunto de valores de $q > 0$ e utilizar um dos critérios *AIC* ou *BIC* para selecionar o modelo com o q que possui o menor valor do critério selecionado.

$$AIC = -2\ln\left(L(\theta)\right) + 2(p+1) \qquad (2.16)$$

$$BIC = -2\ln\left(L(\theta)\right) + (p+1)\ln(n) \qquad (2.17)$$

Além das medidas de qualidade do ajuste, o teste de significância dos parâmetros também pode ser utilizado como uma forma de avaliar a inclusão ou exclusão de variáveis explicativas, o que pode refletir na qualidade do ajuste do modelo. Aplicamos os testes da razão de verossimilhanças, Wald e escore como apresentado em Li (2001). As hipóteses consideradas foram: $H_0 : \beta = \beta^0$ contra $H_1 : \beta \neq \beta^0$ para um determinado vetor β^0 conhecido de dimensão $(s \times 1)$. As estatísticas dos testes razão de verossimilhanças, Wald e escore são dadas pelas equações (2.18), (2.19) e (2.20) respectivamente.

$$\xi_{RV} = 2\{L(\hat{\beta},\hat{\phi}) - L(\beta^0,\hat{\phi})\} \tag{2.18}$$

$$\xi_{W} = [\hat{\beta} - \beta^0]^t \hat{Var}(\hat{\beta})^{-1}[\hat{\beta} - \beta^0] \tag{2.19}$$

$$\xi_{SR} = \mathrm{U}^t_{\hat{\beta}^0} \hat{Var}(\hat{\beta}^0) \mathrm{U}_{\hat{\beta}^0} \tag{2.20}$$

em que $\hat{Var}(\hat{\beta}) = -\frac{1}{\hat{\phi}}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}$ e $\hat{\beta}^0$ é o estimador de máxima verossimilhança de β sob H_0. Para n grande e sob a hipótese nula H_0, temos que as estatísticas ξ_{RV}, ξ_W e ξ_{SR} seguem distribuição qui-quadrado com s graus de liberdade.

2.5 Resíduos

2.5.1 Definição

Na análise de resíduos buscamos verificar se a discrepância entre os valores observados y_i's e os ajustados $\hat{y}_i$'s é estatísticamente relevante. Para isto, utilizamos medidas de discrepância denominadas de resíduos. Uma metodologia de análise de resíduos mais geral foi proposta por Cox e Snell (1968). Em Cysneiros (2004) e Cysneiros, Paula e Galea (2005) podemos encontrar a metodologia de análise de resíduos para classe de modelos simétricos. Detalhes teóricos sobre resíduos para classe de modelo de regressão simétrica não-lineares podem ser encontrados em Cysneiros e Vanegas (2008).

Propomos dois resíduos para a classe de modelos lineares com erros slash-elípticos. O primeiro é um resíduo empírico definido por

$$\begin{aligned} \hat{r}_i &= \frac{(y_i - \mathbf{x}_i^t\beta)}{\sqrt{\hat{Var}(\mathbf{y})}} \\ &= (y_i - \mathbf{x}_i^t\beta)\sqrt{\frac{(q-2)}{qa_1\hat{\phi}}}, \quad \text{para } q > 2 \end{aligned} \tag{2.21}$$

em que $\hat{Var}(\mathbf{y})$ é a variância de $\mathbf{y}$ obtida pelo método dos momentos, dada pela equação (1.7), substituindo o parâmetro ϕ pela seu estimador de máxima verossimilhança e a_1 é calculado pela equação (1.9).

O segundo é o resíduo componente de desvio proposto por Pregibon (1981). Este resíduo baseia-se na estatística da razão de verossimilhanças generalizada denotada por LR_i e expressa por

$$LR_i = 2\left[l(y_i;\widetilde{\mu}_i,\phi) - l(y_i;\widehat{\mu}_i,\phi)\right] \tag{2.22}$$

em que $\widetilde{\mu}_i$ é o valor de μ_i que maximiza $l(y_i;\mu_i,\phi)$ e $l(y_i;\mu_i,\phi)$ é obtida aplicando o logaritmos na função dada em (2.2), isto é,
$l(y_i;\mu_i,\phi) = \frac{1}{2}\log(\phi) + I_{(y_i=\mu_i)}\log\left(\frac{q}{q+1}g(0)\right) + I_{(y_i\neq\mu_i)}\log(\frac{q}{2}) + I_{(y_i\neq\mu_i)}\left[\log(H(z_i^2)) - (q+1)\log|z_i|\right]$.
Pode-se mostrar que o valor de μ_i que maximiza $l(y_i;\mu_i,\phi)$ é y_i. Assim sendo, substituindo $\widetilde{\mu}_i = y_i$ em (2.22) obtemos

$$LR_i = 2I_{(y_i\neq\hat{\mu}_i)}\left\{\log\left(\frac{q}{q+1}g(0)\right) - \log(\frac{q}{2}) - b(\hat{z}_i,q)\right\}$$

Pregibon (1981) define o resíduo componente de desvio como sendo

$$t_D(\hat{z}_i) = \mathsf{sinal}(\hat{z}_i)\sqrt{LR_i}$$

em que $\mathsf{sinal}(x)$ é uma função que retorna o sinal de x.

2.5.2 Simulação

Com o objetivo de verificar as propriedades empíricas dos resíduos $\hat{r}_i$ e $t_D(\hat{z}_i)$, realizamos um estudo de simulação, no qual, foram gerados 10.000 modelos da forma:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1,\ldots,20 \tag{2.23}$$

em que x_i tem distribuição normal padrão, $\beta_0 = 1$, $\beta_1 = 1$ e para ε_i foram consideradas as seguintes distribuições:

1. $\varepsilon_i \sim$ slash-normal$(\mu = 0, \phi = 4, q = 5)$;
2. $\varepsilon_i \sim$ slash-normal$(\mu = 0, \phi = 4, q = 10)$;
3. $\varepsilon_i \sim$ slash-t-Student$(\mu = 0, \phi = 4, q = 5, \nu = 3)$;
4. $\varepsilon_i \sim$ slash-t-Student$(\mu = 0, \phi = 4, q = 10, \nu = 3)$;

Para cada um dos 10.000 ajustes e cada um dos quatro modelos considerados acima foram calculados os resíduos $\hat{r}_i$ e $t_D(\hat{z}_i)$. Por fim, medidas descritivas foram calculadas para as 10.000 réplicas do i-ésimo resíduo $\hat{r}_i$, com $i = 1,\ldots,20$, e podem ser visualizadas nas Tabelas 2.1 e 2.2, enquanto que para o

resíduo $t_D(\hat{z}_i)$, com $i = 1, \ldots, 20$ podem ser visualizadas nas Tabelas 2.3 e 2.4. Para todas as amostras e modelos, o teste de normalidade de Kolmogorov-Smirnov também foi realizado e em todos os casos, os valores-p foram inferiores a $0,001$ e conseqüentemente, nenhum dos resíduos possuem a propriedade de normalidade.

Nas Tabelas 2.1 e 2.2 podemos verificar que as médias do resíduo $\hat{r}_i$ não estam próximas de zero, indicando uma possível existência de viés no ajuste dos modelos slash-normal e slash-t-Student. Quanto aos erros-padrão dos resíduos $\hat{r}_i$, observamos que estão bem próximos de um, tanto nos modelos slash-normal quanto nos modelos slash-t-Student. Os resíduos $\hat{r}_i$ dos modelos slash-normal são mais simétricos que os dos modelos t-Student e a assimetria diminui com o aumento de q em todos os modelos. Nos modelos slash-normal, a curtose dos resíduos $\hat{r}_i$ se aproxima de 3 a medida que q cresce. Já para os modelos slash-t-Student, a curtose dos resíduos são bem altas, mas também diminui com o aumento de q.

Tabela 2.1: Medidas descritivas para as simulações dos resíduos $\hat{r}_i$ para o modelo slash-normal

	slash-normal$(q = 5)$				slash-normal$(q = 10)$			
Observações	Média	EP	Assimetria	Curtose	Média	EP	Assimetria	Curtose
1	0,481	1,029	-0,04751	4,726	0,551	1,012	-0,00092	3,126
2	1,086	1,012	-0,06357	4,565	1,233	1,010	-0,00742	3,161
3	0,316	1,020	0,21750	8,640	0,359	1,011	0,02911	3,130
4	0,038	1,031	-0,01789	4,118	0,047	1,015	-0,03577	3,388
5	0,125	1,027	0,05690	5,880	0,143	1,015	-0,01962	3,091
6	0,274	1,012	0,03095	4,136	0,303	1,009	-0,01899	3,093
7	0,468	1,010	-0,01544	4,137	0,531	1,019	-0,00361	3,108
8	0,327	1,010	0,01020	3,929	0,383	1,018	0,01168	3,195
9	0,322	1,044	0,04879	5,584	0,352	1,016	-0,05356	3,070
10	0,661	1,043	-0,12964	6,784	0,742	1,015	-0,03854	3,244
11	0,439	1,020	-0,01835	4,100	0,497	1,013	-0,00645	3,186
12	0,069	1,035	-0,05493	4,926	0,060	1,007	0,06384	3,190
13	0,735	1,013	0,03764	4,375	0,827	1,008	-0,04032	3,143
14	0,847	1,035	-0,02171	4,372	0,978	1,020	0,03255	3,079
15	0,523	1,012	-0,17006	4,113	0,600	1,010	-0,00820	3,222
16	0,201	1,039	-0,01521	4,182	0,228	1,015	0,01423	3,256
17	-0,238	1,032	0,10323	5,012	-0,278	1,012	-0,02746	3,163
18	0,641	1,007	0,02983	4,263	0,736	1,007	0,01551	3,235
19	-0,188	1,014	-0,03367	3,902	-0,221	0,996	-0,01652	3,152
20	-0,319	1,027	-0,06902	4,404	-0,373	1,014	-0,02976	3,263

Tabela 2.2: Medidas descritivas para as simulações dos resíduos $\hat{r}_i$ para o modelo slash-t-Student

	slash-t-Student($\upsilon = 3, q = 5$)				slash-t-Student($\upsilon = 3, q = 10$)			
Observações	Média	EP	Assimetria	Curtose	Média	EP	Assimetria	Curtose
1	0,376	1,424	-3,59748	128,899	0,428	1,269	-0,65421	14,929
2	0,849	1,391	0,46573	43,107	0,982	1,373	3,10865	90,456
3	0,257	1,483	-0,47242	93,140	0,292	1,307	0,00957	20,271
4	-0,012	1,386	0,64800	86,848	0,063	1,350	5,16903	174,919
5	0,141	2,003	40,19570	3045,796	0,101	1,261	-0,22670	21,017
6	0,218	1,355	1,78609	49,288	0,217	1,435	-4,02657	200,122
7	0,375	1,315	0,35983	20,614	0,417	1,338	0,22896	32,816
8	0,260	1,346	-3,45709	122,853	0,335	1,307	0,93206	23,098
9	0,253	1,393	-3,08111	123,539	0,302	1,304	0,19577	15,717
10	0,524	1,383	1,30447	68,649	0,595	1,256	-0,05539	14,133
11	0,314	1,352	1,65776	58,175	0,382	1,303	-1,07804	30,502
12	0,053	1,351	0,34430	28,099	0,052	1,368	1,52780	51,826
13	0,543	1,362	2,03126	63,288	0,634	1,306	0,15366	24,465
14	0,621	1,490	-10,47648	424,191	0,730	1,402	1,96766	82,587
15	0,393	1,349	-0,06683	29,983	0,464	1,424	-1,22759	117,483
16	0,155	1,355	-0,77914	45,658	0,172	1,308	1,32807	63,619
17	-0,186	1,371	1,27244	36,203	-0,217	1,388	0,45394	39,619
18	0,494	1,382	2,58678	129,013	0,581	1,437	3,78277	105,357
19	-0,115	1,333	1,69535	67,596	-0,151	1,340	-0,40683	23,844
20	-0,263	1,371	-0,50659	40,330	-0,297	1,327	-1,58376	36,385

Nas Tabelas 2.3 e 2.4 temos as medidas descritivas dos resíduos $t_D(\hat{z}_i)$ para os modelos slash-normal e slash-t-Student respectivamente. Para este resíduo também observamos que as médias dos $t_D(\hat{z}_i)$ não estão satisfatoriamente próximas de zero, mas que os erros-padrão estão próximos de um em todos os modelos slash-normal e slash-t-Student. Os resíduos $t_D(\hat{z}_i)$ parecem ser mais simétricos que os resíduos $\hat{r}_i$ em todos os modelos, mas no caso dos modelos slash-normal, os coeficientes de assimetria estão mais próximos de zero para os modelos com $q = 10$, que para os com $q = 5$. Os resíduos $t_D(\hat{z}_i)$ também apresentam o coeficiente de curtose mais próximos de 3 que os resíduos $\hat{r}_i$ em todos os modelos.

Tabela 2.3: Medidas descritivas para as simulações dos resíduos $t_D(\hat{z}_i)$ para o modelo slash-normal

	slash-normal($q = 5$)				slash-normal($q = 10$)			
Observações	Média	EP	Assimetria	Curtose	Média	EP	Assimetria	Curtose
1	0,515	1,074	-0,180	3,236	0,559	1,024	-0,055	2,981
2	1,148	1,018	-0,407	3,524	1,243	1,002	-0,133	3,009
3	0,339	1,065	-0,127	3,212	0,365	1,024	-0,008	2,982
4	0,041	1,090	-0,036	3,107	0,048	1,029	-0,015	3,062
5	0,135	1,081	-0,068	3,104	0,145	1,031	-0,030	2,972
6	0,293	1,069	-0,069	3,107	0,308	1,024	-0,045	2,963
7	0,501	1,062	-0,153	3,121	0,538	1,031	-0,048	2,973
8	0,351	1,068	-0,102	3,057	0,389	1,031	-0,029	3,037
9	0,344	1,086	-0,089	3,255	0,358	1,031	-0,078	2,941
10	0,704	1,072	-0,218	3,271	0,752	1,022	-0,106	3,078
11	0,470	1,071	-0,142	3,159	0,504	1,025	-0,050	3,057
12	0,075	1,087	-0,046	3,198	0,061	1,022	0,048	3,018
13	0,783	1,048	-0,237	3,246	0,838	1,014	-0,117	2,996
14	0,899	1,060	-0,302	3,276	0,989	1,021	-0,066	2,893
15	0,562	1,064	-0,252	3,169	0,609	1,020	-0,063	3,048
16	0,216	1,095	-0,057	3,086	0,232	1,029	-0,014	3,070
17	-0,257	1,081	0,115	3,258	-0,282	1,026	0,003	3,026
18	0,684	1,051	-0,189	3,075	0,746	1,014	-0,061	3,061
19	-0,201	1,074	0,053	3,102	-0,224	1,011	0,004	3,002
20	-0,341	1,079	0,069	3,124	-0,379	1,027	0,011	3,067

Tabela 2.4: Medidas descritivas para as simulações dos resíduos $t_D(\hat{z}_i)$ para o modelo slash-t-Student

	slash-t-Student$(\upsilon = 3, q = 5)$				slash-t-Student$(\upsilon = 3, q = 10)$			
Observações	Média	EP	Assimetria	Curtose	Média	EP	Assimetria	Curtose
1	0,475	1,270	-0,420	3,164	0,519	1,242	-0,490	3,168
2	1,020	1,205	-0,971	4,315	1,106	1,145	-0,949	4,389
3	0,317	1,297	-0,269	3,056	0,340	1,260	-0,246	2,919
4	-0,011	1,282	-0,011	2,887	0,058	1,257	0,002	2,872
5	0,148	1,287	-0,068	2,908	0,123	1,254	-0,112	2,808
6	0,271	1,284	-0,244	2,900	0,275	1,255	-0,281	2,995
7	0,471	1,277	-0,436	3,078	0,501	1,241	-0,458	3,208
8	0,339	1,271	-0,328	2,996	0,389	1,248	-0,295	3,041
9	0,316	1,283	-0,252	2,958	0,355	1,252	-0,277	3,014
10	0,654	1,237	-0,611	3,496	0,705	1,205	-0,640	3,496
11	0,403	1,284	-0,405	3,021	0,456	1,245	-0,385	3,078
12	0,071	1,293	-0,085	2,897	0,058	1,265	-0,056	2,926
13	0,673	1,249	-0,633	3,397	0,739	1,216	-0,618	3,451
14	0,791	1,232	-0,792	3,775	0,856	1,214	-0,835	3,938
15	0,496	1,279	-0,465	3,170	0,555	1,226	-0,484	3,364
16	0,195	1,294	-0,171	2,873	0,208	1,254	-0,190	2,869
17	-0,248	1,291	0,259	2,984	-0,266	1,264	0,255	3,043
18	0,615	1,252	-0,563	3,306	0,675	1,218	-0,581	3,513
19	-0,153	1,288	0,159	2,843	-0,178	1,267	0,141	2,934
20	-0,330	1,287	0,292	3,005	-0,340	1,248	0,235	2,987

CAPÍTULO 3

Análise de Diagnóstico no modelo slash-elíptico

A metodologia para análise de diagnóstico sobre enfoque de influência local para família de modelos lineares com erros elípticos pode ser encontrada nos trabalhos de Galea, Paula e Bolfarine (1997), Galea, Bolfarine e Vilca-Labra (2002) e Galea, Paula e Uribe-Opazo (2003). Galea, Paula e Cysneiros (2005) apresentam métodos de diagnósticos para classe de modelos não-lineares simétricos. Apresentaremos aqui a metodologia de análise de diagnóstico sobre o enfoque de alavancagem generalizada e influência local para classe de modelos slash-elípticos.

Assumindo que o modelo ajustado está correto, a análise de diagnóstico visa investigar se existem observações que exercem forte influência no ajuste do modelo ou com valores atípicos nos regressores. Uma observação do conjunto de dados é dita ser influente se sob sua presença ou pequena perturbação ocorrer variações desproporcionais nas estimativas dos parâmetros sob o modelo ajustado.

A medida de alavancagem generalizada GL_{ij}, definida por

$$GL_{ij} = \frac{\partial \hat{y}_i}{\partial y_j}$$

reflete a taxa de mudança instantânea no i-ésimo valor predito $\hat{y}_i$, quando a j-ésima variável resposta y_j é acrescida por um infinitésimo. A alavancagem generalizada GL_{ii} é proposta por Wei, Hu e Fung (1998) como a medida de maior influência de y_i no seu próprio valor ajustado. Alta alavancagem sugere que a i-ésima observação pode ter valores atípicos dos regressores.

A análise de influência local é um método de diagnóstico proposto por Cook (1986) que visa avaliar o efeito de pequenas perturbações nos dados ou modelo segundo uma medida de influência. Seja $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ um vetor $(n \times 1)$ de perturbações restrito a algum conjunto aberto Ω e ω_0 o vetor de não-perturbação. Na prática, deseja-se comparar $\widehat{\theta}$ e $\widehat{\theta}_\omega$ segundo uma medida de influência quando ω varia em Ω. Consideráveis distâncias entre estas medidas podem indicar a presença de pontos influentes. Entre as medidas de influência local destaca-se o afastamento da verossimilhança proposta por Cook (1986) e uma medida de distância baseada no resíduo de Pearson proposta por Thomas e Cook (1990).

3.1 Alavancagem Generalizada

Denotemos o estimador de máxima verossimilhança de θ por $\hat{\theta}(\mathbf{y})$ e o de μ por $\hat{\mu} = \mathrm{X}\beta$, então Wei, Hu e Fung (1998) mostraram que a matriz $GL(\hat{\theta}) = \frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{y}} = \left(\frac{\partial \hat{y}_i}{\partial y_j}\right)_{(n\times n)}$ de alavancagem generalizada pode ser expressa na forma

$$GL(\hat{\theta}) = D_{\theta}\ddot{L}_{\theta\theta}\ddot{J}_{\theta\mathbf{y}}, \tag{3.1}$$

em que $D_{\theta} = \frac{\partial \hat{\mu}}{\partial \theta^t}|_{\theta=\hat{\theta}(\mathbf{y})} = \frac{\partial \mathrm{X}\beta}{\partial \theta^t}|_{\theta=\hat{\theta}(\mathbf{y})} = [\mathrm{X}, 0]$ é uma matriz de ordem $(n \times (p+1))$ e $\ddot{J}_{\theta\mathbf{y}} = \frac{\partial^2 L(\theta)}{\partial\theta\partial\mathbf{y}^t}|_{\theta=\hat{\theta}(\mathbf{y})}$ é uma matriz de ordem $((p+1)\times n)$. Então, podemos simplificar (3.1) e expressar da seguinte forma

$$\begin{aligned} GL(\hat{\theta}) &= \begin{pmatrix} \mathrm{X} & 0 \end{pmatrix} \begin{pmatrix} \ddot{L}_{\beta\beta} & \ddot{L}_{\beta\phi} \\ \ddot{L}_{\phi\beta} & \ddot{L}_{\phi\phi} \end{pmatrix} \begin{pmatrix} \ddot{J}_{\beta\mathbf{y}} \\ \ddot{J}_{\phi\mathbf{y}} \end{pmatrix} \\ &= \mathrm{X}\ddot{L}_{\beta\beta}\ddot{J}_{\beta\mathbf{y}} + \mathrm{X}\ddot{L}_{\beta\phi}\ddot{J}_{\phi\mathbf{y}}. \end{aligned} \tag{3.2}$$

Aplicando essa metodologia a classe de modelo lineares com erros slash-elíptico, temos que a matriz $\ddot{J}_{\beta\mathbf{y}} = -\frac{1}{\hat{\phi}}\mathrm{X}^t B''_{(\hat{\mathbf{z}},q)}$ e o vetor $\ddot{J}_{\phi\mathbf{y}} = -\frac{1}{2\hat{\phi}^{3/2}}\mathbf{m}^t_{(\hat{\mathbf{z}},q)}$ e desta forma, a matriz de alavancagem generalizada é dada por

$$\begin{aligned} GL(\hat{\theta}) &= \mathrm{X}\left(\frac{-\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}}{\hat{\phi}}\right)\left(\frac{-\mathrm{X}^t B''_{(\hat{\mathbf{z}},q)}}{\hat{\phi}}\right) + \mathrm{X}\left(\frac{2\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)}}{\sqrt{\hat{\phi}}e}\right)\left(\frac{-\mathbf{m}^t_{(\hat{\mathbf{z}},q)}}{2\hat{\phi}^{3/2}}\right) \\ &= \frac{1}{\hat{\phi}^2}\mathrm{X}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t B''_{(\hat{\mathbf{z}},q)} - \frac{1}{\hat{\phi}^2 e}\mathrm{X}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t\mathbf{m}_{(\hat{\mathbf{z}},q)}\mathbf{m}^t_{(\hat{\mathbf{z}},q)} \\ &= \frac{1}{\hat{\phi}^2}\mathrm{X}\left(\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}\mathrm{X}\right)^{-1}\mathrm{X}^t V_{(\hat{\mathbf{z}},q)}. \end{aligned} \tag{3.3}$$

Podemos re-escrever a equação (3.3) como expresso abaixo

$$\begin{aligned} GL(\hat{\theta}) &= \mathrm{X}^t \frac{1}{\sqrt{\hat{\phi}}} V_{(\hat{\mathbf{z}},q)}^{1/2} \left(\mathrm{X}^t \hat{\phi} V_{(\hat{\mathbf{z}},q)} \mathrm{X}\right)^{-1} \frac{1}{\sqrt{\hat{\phi}}} V_{(\hat{\mathbf{z}},q)}^{1/2} \mathrm{X} \\ &= \mathrm{X}^t {V^*}_{(\hat{\mathbf{z}},q)}^{1/2} \left(\mathrm{X}^t {V^*}_{(\hat{\mathbf{z}},q)} \mathrm{X}\right)^{-1} {V^*}_{(\hat{\mathbf{z}},q)}^{1/2} \mathrm{X}, \end{aligned} \tag{3.4}$$

em que ${V^*}_{(\hat{\mathbf{z}},q)}^{1/2} = \frac{1}{\sqrt{\hat{\phi}}} V_{(\hat{\mathbf{z}},q)}^{1/2}$. Desta forma, o $GL(\hat{\theta})$ pode ser interpretado como a matriz H de alavancagem de um modelo de regressão normal ponderado pela matriz ${V^*}_{(\hat{\mathbf{z}},q)}$.

Um gráfico dos elementos da diagonal de $GL(\hat{\theta})$ versus índices pode revelar pontos com alta influência no seu própio valor predito.

3.2 Influência local baseada no afastamento da verossimilhança

Seja $L(\theta)$ e $L(\theta_\omega)$ as funções de log-verossimilhança dos modelos postulado e perturbado respectivamente, então a medida de afastamento da verossimilhança é definida por

$$LD(\omega) = 2\{L(\widehat{\theta}) - L(\widehat{\theta}_\omega)\}$$

Pode-se notar que se $\omega = \omega_0$, então $LD(\omega) = 0$ e se $\omega \neq \omega_0$, então $LD(\omega) \geq 0$, isto é, ω_0 é um ponto de mínimo local de $LD(\omega)$.

A análise de influência local baseada no afastamento da verossimilhança consiste em estudar como a superfície $\alpha(\omega) = (\omega^t, LD(\omega))^t$, para $\omega \in \Omega$, desvia-se de seu plano tangente em torno de ω_0 através da análise da curvatura normal da superfície $\alpha(\omega)$ em alguma direção arbitrária d, com $\|d\| = 1$ (COOK; WEISBERG, 1982). Pode-se mostrar que essa curvatura normal na direção d pode ser expressa por

$$C_d = 2 \mid d^t \ddot{F} d \mid$$

em que $\ddot{F} = \Delta^t \ddot{L}_{\theta\theta} \Delta$ e $\Delta = \frac{\partial^2 L(\theta_\omega)}{\partial \theta \partial \omega^t}$ avaliada em $\omega = \omega_0$ e $\theta = \widehat{\theta}$.

Se o interesse for realizar a análise de influência para os parâmetros β e ϕ separadamente, a matriz $\ddot{F}$ pode ser adaptada para excluir de $\ddot{L}_{\theta\theta}$ a influência dos outros parâmetros. Desta forma, a curvatura normal na direção d para os parâmetros β e ϕ são respectivamente:

$$C_d(\beta) = 2 \mid d^t \Delta^t (\ddot{L}_{\theta\theta} - L_1) \Delta d \mid$$

e

$$C_d(\phi) = 2 \mid d^t \Delta^t (\ddot{L}_{\theta\theta} - L_2) \Delta d \mid,$$

em que $L_1 = \begin{bmatrix} 0 & 0 \\ 0 & \ddot{L}_{\phi\phi} \end{bmatrix}$ e $L_2 = \begin{bmatrix} \ddot{L}_{\beta\beta} & 0 \\ 0 & 0 \end{bmatrix}$.

O gráfico de índices $d_{\max}$, o auto-vetor correspondente ao maior auto=valor absoluto de $\ddot{F}$, pode revelar as observações mais influentes em $\hat{\theta}$. Outra possibilidade de análise foi proposta por Lesaffre e Verbeke (1998), que consiste na construção do gráfico dos C_i's, obtidos a partir do cálculo de C_{d_i} para o vetor d_i formado por zeros com um na i-ésima posição para cada $i = 1, 2, \ldots, n$. Lesaffre e Verbeke (1998) também sugere que as observações tais que $C_i > 2\overline{C}$, em que $\overline{C} = \frac{\sum_{i=1}^{n} C_i}{n}$, podem ser pontos influentes.

3.2.1 Perturbação na escala

Considere que $y_i \sim SEl(\mathbf{x}_i\beta, \phi/\omega_i, q; g(\cdot))$, isto é, que o modelo linear $y_i = \mathbf{x}_i\beta + \varepsilon_i$ é heteroscedástico com perturbação no parâmetro de escala $\phi_i = \phi/\omega_i$, para $\omega_i > 0$ e $i = 1, 2, \ldots, n$. Quando $\omega_i = 1$, $\phi_i = \phi$ e o modelo perturbado se reduz ao postulado; para $0 < \omega_i < 1$, há um inflacionamento de ϕ e quando $\omega_i > 1$, há uma redução de ϕ. A log-verossimilhança do modelo perturbado é dada por

$$L(\theta, \omega) = -\frac{n}{2}\log(\phi) + \sum_{i \in A} b(z_{i\omega}, q) + C,$$

em que $z_{i\omega} = \omega_i \dfrac{(y_i - \mathbf{x}_i\beta)}{\sqrt{\phi}} = \omega_i z_i$. A matriz de perturbação Δ com perturbação na escala é expressa por

$$\hat{\Delta} = \begin{pmatrix} \frac{1}{2\sqrt{\widehat{\phi}}} \mathrm{X}^t M_{(\widehat{\mathbf{z}}, q)} \\ \frac{1}{4\widehat{\phi}} \widehat{\mathbf{z}}^t M_{(\widehat{\mathbf{z}}, q)} \end{pmatrix},$$

em que $M_{(\widehat{\mathbf{z}}, q)}$ é uma matriz de zeros com os elementos $m(\widehat{z}_i, q) = -b'(\widehat{z}_i, q) - \widehat{z}_i b''(\widehat{z}_i, q)$, para $i = 1, 2, \ldots, n$, na diagonal.

3.2.2 Perturbação nos casos

Considere um esquema de perturbação do i-ésimo caso no qual a função de log-verossimilhança de θ do modelo perturbado seja expressa na forma

$$L(\theta, \omega) = \sum_{i=1}^{n} \omega_i \log(f_{y_i}(y_i)),$$

em que $0 \leq \omega_i \leq 1$. Para esse esquema de perturbação, a matriz Δ é dada por

$$\hat{\Delta} = \begin{pmatrix} -\frac{1}{\sqrt{\widehat{\phi}}} \mathrm{X}^t B'_{(\widehat{\mathbf{z}},q)} \\ \frac{1}{2\widehat{\phi}} \mathbf{r}^t_{(\widehat{\mathbf{z}},q)} \end{pmatrix},$$

em que $\mathbf{r}_{(\widehat{\mathbf{z}},q)}$ é um vetor $(n \times 1)$ com os elementos $\{-\widehat{z}_i b'(\widehat{z}_i, q) - 1\}$, para $i = 1, 2, \ldots, n$.

3.3 Influência local na predição

Considere o estudo do efeito de pequenas perturbações na predição de um particular vetor $\mathbf{v}$ de ordem $(p \times 1)$ de valores dos regressores, para o qual não se tem uma resposta observada. Sejam $\widehat{\mu}(\mathbf{v}) = \mathbf{v}^t \widehat{\beta}$ e $\widehat{\mu}(\mathbf{v}, \omega) = \mathbf{v}^t \widehat{\beta}_\omega$ as predições nos modelos postulado e perturbado respectivamente, onde $\widehat{\beta}_\omega$ é o vetor de estimativas de máxima verossimilhança de θ do modelo perturbado. Uma medida de influência local na predição proposta por Thomas e Cook (1990) baseada no resíduo de Pearson é definida por

$$f(\mathbf{v}, \omega) = \{\widehat{\mu}(\mathbf{v}) - \widehat{\mu}(\mathbf{v}, \omega)\}^2. \tag{3.5}$$

A partir do estudo do comportamento da superfície $\delta(\omega) = (\omega^t, f(\mathbf{v}, \omega))^t$ em torno de ω_0, pode-se mostrar que a curvatura normal desta superfície na direção d é $C_d(\mathbf{v}) = 2 \mid d^t \ddot{F} d \mid$, em que

$$\ddot{F} = \frac{\partial^2 f(\mathbf{v}, \omega)}{\partial \omega \partial \omega^t} \Big|_{\substack{\beta = \widehat{\beta} \\ \omega = \omega_0}} = \Delta^t (\ddot{L}_{\beta\beta} \mathbf{v}\mathbf{v}^t \ddot{L}_{\beta\beta}) \Delta.$$

Neste caso, o vetor $d_{\max}(\mathbf{v})$ para um dado vetor de valores dos regressores $\mathbf{v}$ se resume na expressão

$$d_{\max}(\mathbf{v}) = \Delta^t \ddot{L}_{\beta\beta} \mathbf{v} \tag{3.6}$$

Uma proposta de análise é considerar $\mathbf{v} = \mathbf{x}_i$ para cada observação i, com $i = 1, 2, \ldots, n$, e calcular o maior valor do i-ésimo vetor $d_{\max}(\mathbf{x}_i)$, que será denotado por $d_{\max_i}(\mathbf{x}_i)$. A partir do gráfico de índices do vetor $(d_{\max_1}(\mathbf{x}_1), d_{\max_2}(\mathbf{x}_2), \ldots, d_{\max_n}(\mathbf{x}_n))^t$ pode-se identificar quais observações tem substancial influência sobre $\widehat{y}$.

3.3.1 Perturbação na variável resposta

Suponha que a i-ésima resposta seja perturbada aditivamente com a perturbação $y_{i\omega} = y_i + s\omega_i$, em que s é o desvio-padrão estimado de y, de modo que quando $\omega_i = 0$, o modelo perturbado é igual ao postulado. A função de log-verossimilhança para o modelo perturbado aditivamente na resposta é dado

por

$$L(\theta,\omega) = -\frac{n}{2}\log(\phi) + \sum_{i\in A} b(z_{i\omega}, q) + C,$$

em que $z_{i\omega} = \dfrac{(y_i + s\omega_i - \mathbf{x}_i\beta)}{\sqrt{\phi}} = z_i + \dfrac{s\omega_i}{\sqrt{\phi}}$.

Para o esquema de perturbação aditiva na resposta, a matriz Δ obtida é dada por

$$\hat{\Delta} = -\frac{s}{\widehat{\phi}}\mathrm{X}^t B''_{(\widehat{\mathbf{z}},q)}.$$

3.3.2 Perturbação nos regressores

Suponha agora que a i-ésima observação de uma particular variável explicativa t, $t = 2,3,\ldots,p$, seja perturbada de forma aditiva pelo esquema $x_{it\omega} = x_{it} + \omega_i$, que também pode ser expresso em forma vetorial por $\mathbf{x}_{i\omega} = \mathbf{x}_i + \omega_i\mathbf{s}_t$, onde $\mathbf{s}_t$ é um vetor $(p\times 1)$ de zeros com um na t-ésima posição. Neste caso também, quando $\omega_i = 0$, o modelo perturbado é igual ao postulado. A função de log-verossimilhança para o modelo com perturbação aditiva na t-ésima variável explicativa é expressa por

$$L(\theta,\omega) = -\frac{n}{2}\log(\phi) + \sum_{i\in A} b(z_{i\omega}, q) + C,$$

em que $z_{i\omega} = \dfrac{(y_i - \mathbf{x}\beta - \omega_i\beta_t)}{\sqrt{\phi}} = z_i - \dfrac{\omega_i\beta_t}{\sqrt{\phi}}$.

Para o esquema de perturbação aditiva na t-ésima variável explicativa, a matriz Δ obtida é dada por

$$\hat{\Delta} = -\frac{\widehat{\beta}_t}{\widehat{\phi}}\mathrm{X}^t B''_{(\widehat{\mathbf{z}},q)} + \left(\frac{1}{\sqrt{\widehat{\phi}}}\sum_{i\in A} b'(\widehat{z}_i, q)\right)\mathbf{s}_t.$$

CAPÍTULO 4

Aplicações

4.1 Salinidade

Considere o conjunto de dados, com $n = 28$ observações, sobre a salinidade do rio *Pamlico Sound* na Carolina do Norte - EUA apresentado por Rupppert e Carroll (1980). Estamos interessados em modelar a média quinzenal da salinidade do rio (variável resposta $\mathbf{y}$) segundo a salinidade defasada de duas semanas (variável explicativa $\mathbf{x}_1$) e vazão de água do rio (variável explicativa $\mathbf{x_2}$). Como não conhecemos o valor do parâmetro q da slash-elíptica e dos graus de liberdade da t-Student e slash-t-Student, denotado por ν, escolhemos os modelos com os parâmetros q e ν correspondente ao modelo com menor AIC e BIC. O parâmetro q foi selecionado na seqüência 3; 3,5; 4; 4,5;... 50 e o parâmetro ν na seqüência 2,1; 2,5; 3; 3,5; 4. Os modelos escolhidos foram

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad i = 1, \ldots, 28 \tag{4.1}$$

em que

- $\varepsilon_i \sim$ normal(μ, ϕ);
- $\varepsilon_i \sim t$-Student$(\mu, \phi, \nu = 2,1)$;
- $\varepsilon_i \sim$ slash-normal$(\mu, \phi, q = 3)$;

- $\varepsilon_i \sim$ slash-t-Student$(\mu, \phi, q = 3, \nu = 2,1)$;

Podemos ver na Tabela 4.1 os valores dos critérios AIC e BIC dos modelos selecionados para análise. Na Tabela 4.2, podemos observar que os erros-padrão dos modelos slash-normal$(q = 3)$ e slash-t-Student$(q = 3, \nu = 2.1)$ são bem menores que os dos modelos normal e t-Student$(\nu = 2.1)$, sendo o modelo slash-t-Student$(q = 3, \nu = 2.1)$ o com os menores erros-padrão para todos os parâmetros. As estatísticas de testes marginais de razão de verossimilhanças, Wald e escore e seus respectivos valores p, omitidos aqui, foram calculados para todos os modelos considerados. Todos os valores p foram inferiores a $0,001$ em todos os casos, isto é, todos os coeficientes foram significativos ao nível de $0,1\%$ de significância.

Tabela 4.1: Qualidade do ajuste dos modelos para os dados de salinidade

Distribuições	AIC	BIC
normal	99,16	104,49
t-Student$(\nu = 2.1)$	97,44	102,77
slash-normal$(q = 3)$	98,84	104,17
slash-t-Student$(q = 3, \nu = 2.1)$	97,23	102,55

Tabela 4.2: Estimativas e seus erros padrão (em parêntesis) para os modelos ajustados dos dados de salinidade

Distribuições	β_0	β_1	β_2	ϕ
normal	9,32 (2,4900)	0,78 (0,0800)	-0,29 (0,0900)	1,52 (0,4058)
t-Student$(\nu = 2,1)$	14,12 (1,8100)	0,74 (0,0600)	-0,47 (0,0600)	0,51 (0,2115)
slash-normal$(q = 3)$	13,13 (0,0439)	0,75 (0,0004)	-0,44 (0,0001)	0,59 (0,0448)
slash-t-Student$(q = 3, \nu = 2,1)$	14,24 (0,0239)	0,73 (0,0002)	-0,47 (0,0001)	0,24 (0,0122)

Analisamos os resíduos $\hat{r}_i$ e $t_D(\hat{z}_i)$ para os modelos selecionados, com exceção do resíduo $\hat{r}_i$ para o modelo t-Student$(\nu = 2.1)$, uma vez que para calcular o estimador da variância de $\mathbf{y}$ pelo método dos momentos, ν tem que ser maior que 4. As medidas descritivas e o teste de hipótese de Kolmogorov Smirnov para o resíduo $\hat{r}_i$ e $t_D(\hat{z}_i)$ podem ser vistos nas Tabelas 4.3 e Tabela 4.4 respectivamente.

Considerando o resíduo $\hat{r}_i$, vemos que as médias dos resíduos dos modelos slash-elípticos não estão tão próximas de zero quanto as do modelo normal e o desvio padrão dos resíduos do modelo slash-t-Student$(q = 3, \nu = 2.1)$ é bem menor que nos demais modelos, que estão próximos de um. As curtoses dos resíduos $\hat{r}_i$ nos modelos slash-elípticos são maires que no modelo normal, que também é o que apresenta coeficiente de assimetria mais próximo de zero. A partir do teste de Kolmogorov-Smirnov, vemos que os resíduos $\hat{r}_i$ do modelo slash-t-Student$(q = 3, \nu = 2.1)$ não seguem distribuição normal.

Quanto aos resíduos $t_D(\hat{z}_i)$, vemos que o desvio padrão do modelo t-Student$(\nu = 2.1)$ é bem menor que nos demais modelos, que estão próximos de um. Somente a curtose dos resíduos $t_D(\hat{z}_i)$ do modelo normal está próxima de 3. Quanto a assimetria dos resíduos, os modelos normal e slash-t-Student$(q =$

$3, \nu = 2.1)$ são os mais simétricos. Segundo os testes de Kolmogorov-Smirnov, podemos observar que apenas os resíduos $t_D(\hat{z}_i)$ do modelo normal segue distribuição normal.

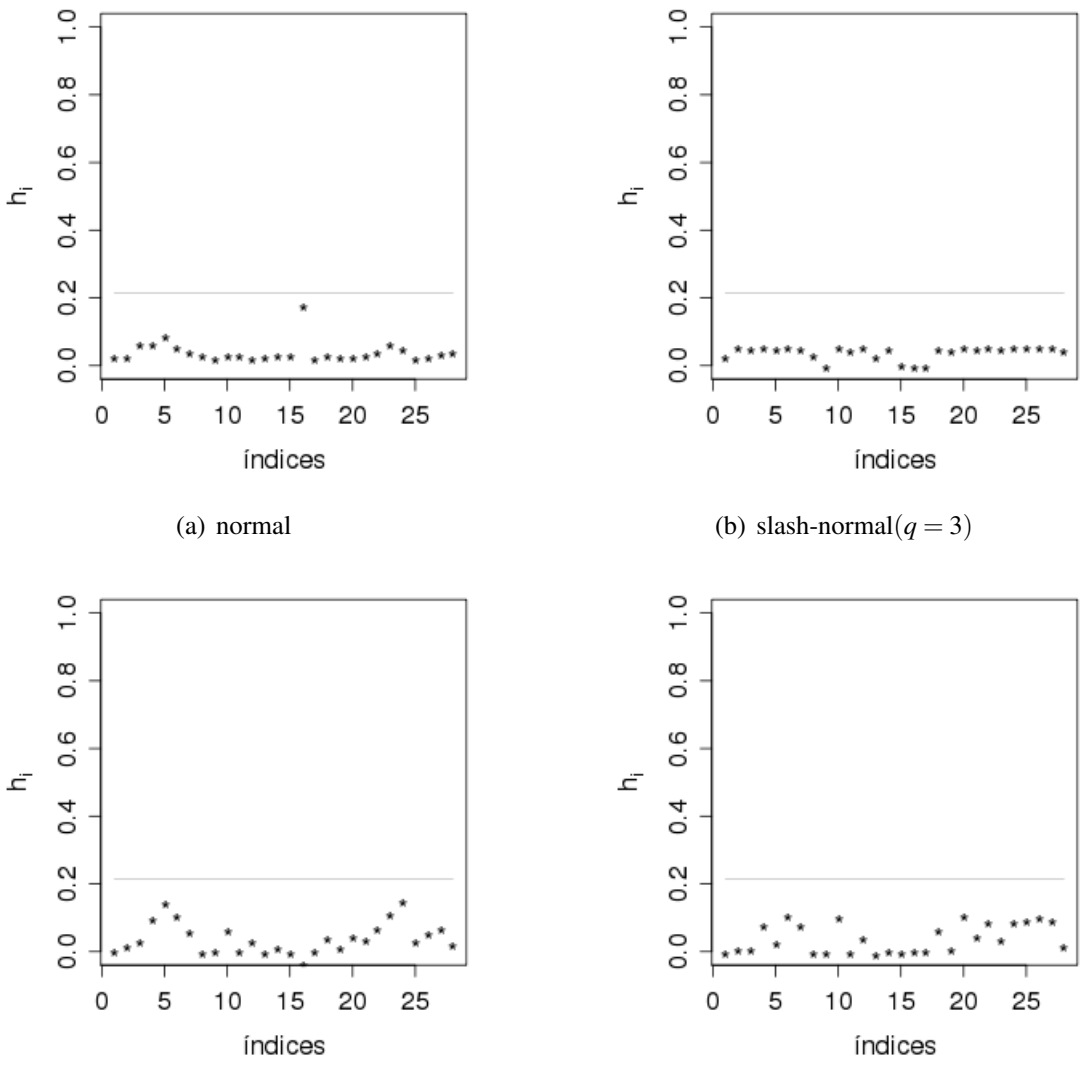
Tabela 4.3: Medidas descritivas dos resíduos $\hat{r}_i$ para os dados de salinidade

	normal	slash-normal $(q = 3)$	slash-t-Student $(q = 3, \nu = 2.1)$
Média	<0,001	0,048	-0,020
Desvio padrão	1,018	0,993	0,346
Assimetria	-0,053	0,741	0,961
Curtose	3,024	5,150	5,844
ks-valor	0,106	0,146	0,322
p-valor	0,879	0,545	0,004

Tabela 4.4: Medidas descritivas dos resíduos $t_D(\hat{z}_i)$ para os dados de salinidade

	normal	slash-normal $(q = 3)$	t-Student $(\nu = 2, 1)$	slash-t-Student $(q = 3, \nu = 2, 1)$
Média	<0,001	0,187	-0,020	-0,105
Desvio padrão	1,018	1,489	0,461	1,693
Assimetria	-0,053	-0,209	0,662	0,042
Curtose	3,024	2,046	4,816	1,718
ks-valor	0,106	0,423	0,285	0,316
p-valor	0,879	<0,001	0,016	0,005

Os gráficos de alavancagens generalizadas para os quatro modelos selecionados podem ser vistos na Figura 4.1. Observamos que o modelo slash-normal$(q = 3)$ apresentou as mais baixas alavancagens (Figura 4.1(b)), em relação aos demais modelos. Mesmo assim, não há indícios de ponto de alavanca nos quatro modelos.



(a) normal

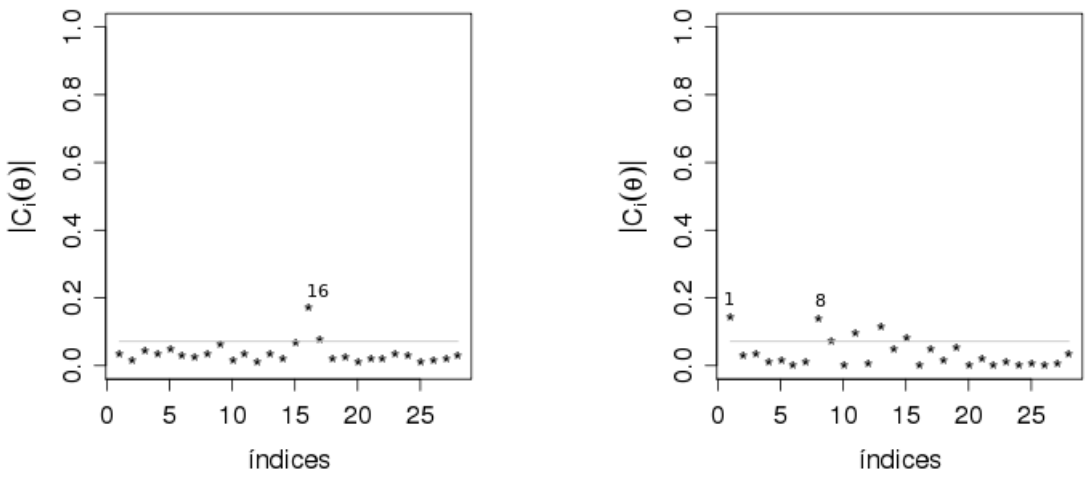
(b) slash-normal$(q = 3)$

(c) t-Student$(\nu = 2,1)$

(d) slash-t-Student$(q = 3, \nu = 2,1)$

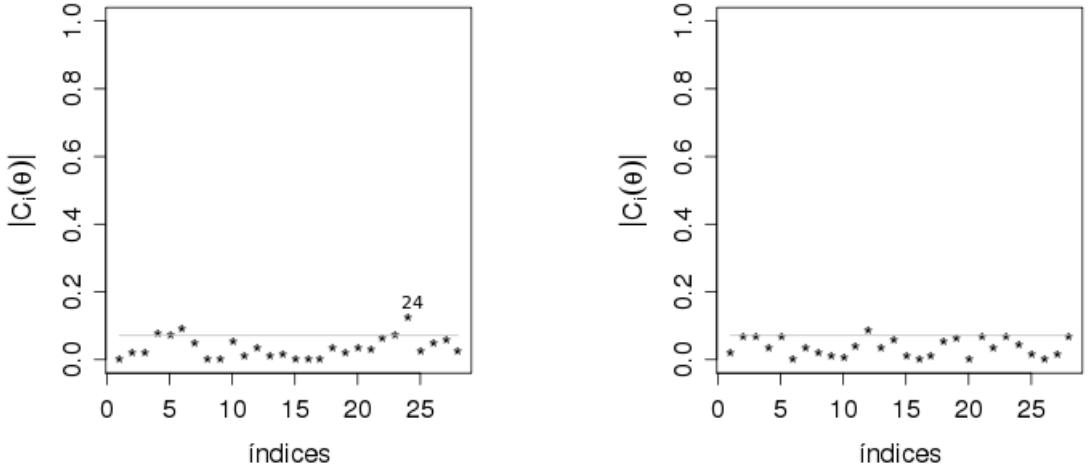
Figura 4.1: Gráficos de índices da alavancagem generalizada para os modelos ajustados aos dados de salinidade

Os gráficos C_i para os dados de salinidade sob perturbação na escala podem ser visto na Figura 4.2. Observamos que o modelo slash-t-Student$(q = 3, \nu = 2,1)$ é o único dos quatros modelos considerados que não apresentou pontos influentes quanto a perturbação na escala.



(a) normal

(b) slash-normal($q = 3$)

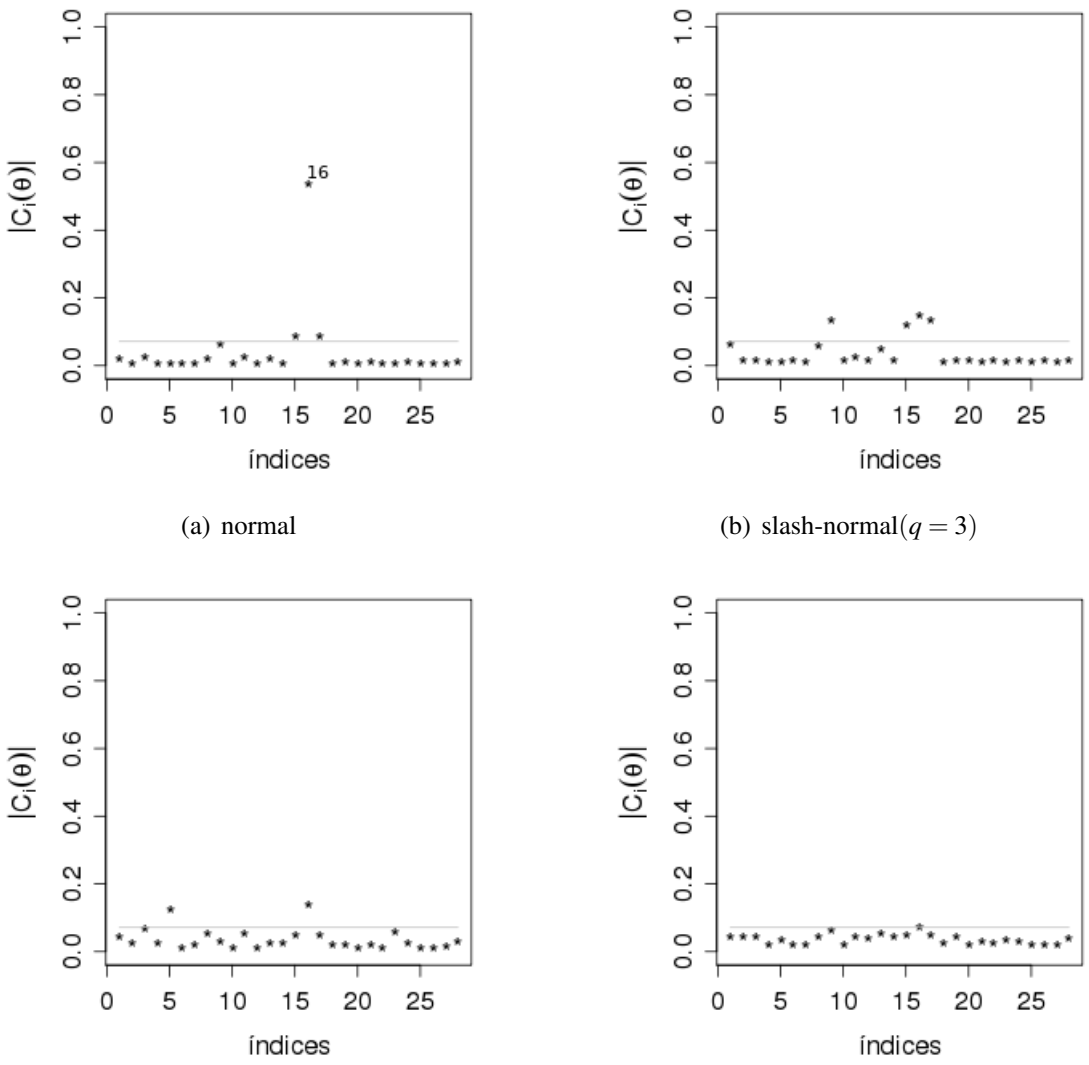


(c) t-Student($\nu = 2,1$)

(d) slash-t-Student($q = 3, \nu = 2,1$)

Figura 4.2: Gráficos de índices C_i para os modelos ajustados aos dados de salinidade sob perturbação na escala

Considere agora a análise de influência local com o esquema de perturbação nos casos. Os gráficos C_i podem ser visto na Figura 4.3. Podemos observar que a observação 16 é um ponto influente principalmente no modelo normal, mais também nos modelos slash-normal($q = 3$) e t-Student($\nu = 2,1$). O modelo slash-t-Student($q = 3, \nu = 2,1$) não possue pontos influentes quanto a perturbação nos casos.



(a) normal

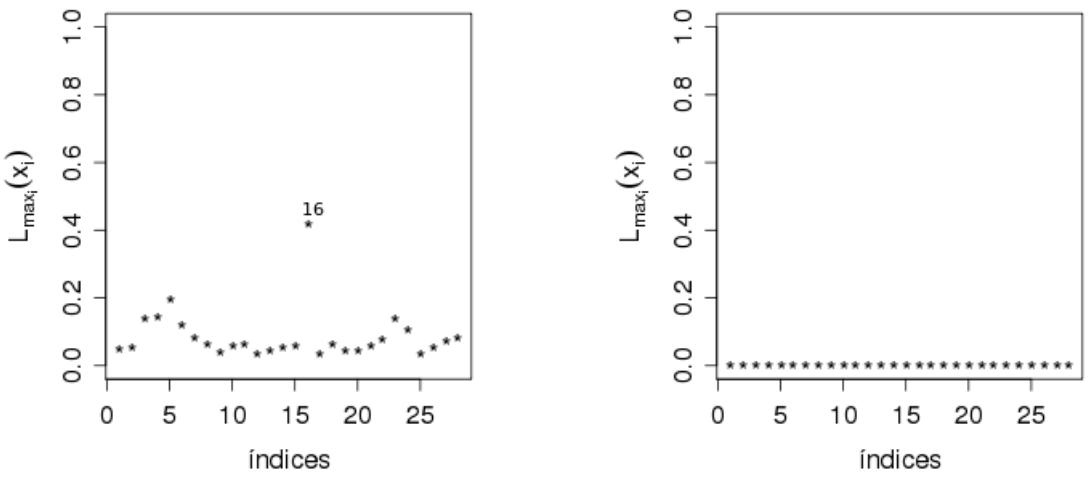
(b) slash-normal$(q = 3)$

(c) t-Student$(\nu = 2,1)$

(d) slash-t-Student$(q = 3, \nu = 2,1)$

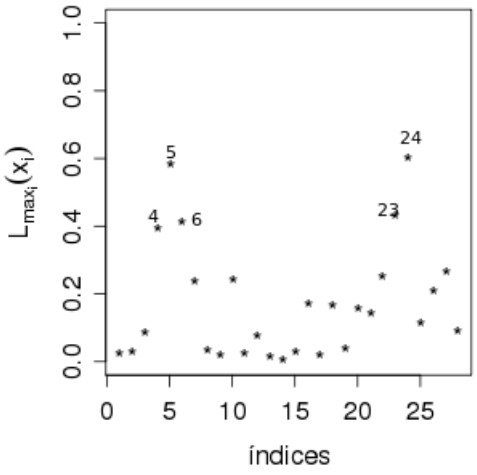
Figura 4.3: Gráficos de índices C_i para os modelos ajustados aos dados de salinidade sob perturbação nos casos

Na Figura 4.4 têm-se os gráficos L_{max} para os dados de salinidade sob perturbação na resposta, no qual, podemos observar que os modelos slash-elípticos não apresentam pontos influentes. O modelo normal apresenta um ponto influente, a observação 16, e o modelo t-Student$(\nu = 2,1)$ apresenta alguns pontos influentes.

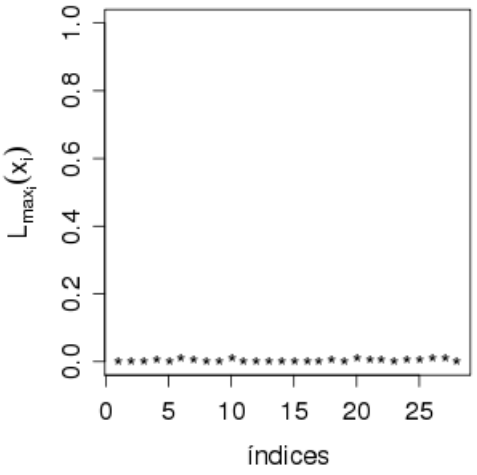


(a) normal

(b) slash-normal$(q = 3)$



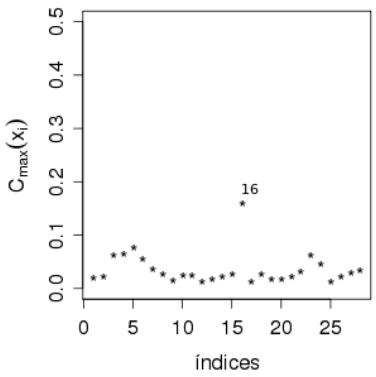
(c) t-Student$(\nu = 2,1)$



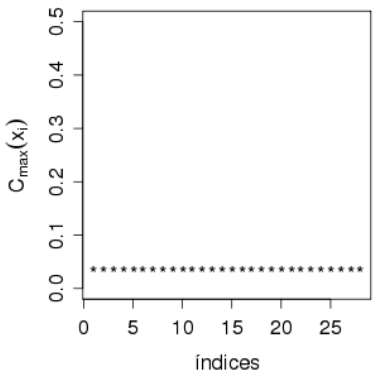
(d) slash-t-Student$(q = 3, \nu = 2,1)$

Figura 4.4: Gráficos de índices L_{max} para os modelos ajustados aos dados de salinidade sob perturbação na resposta

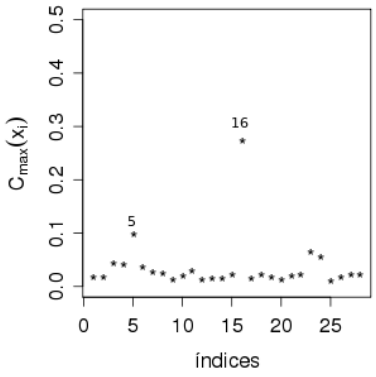
Na Figura 4.5 têm-se os gráficos C_{max} das variável x_1 e x_2 para os dados de salinidade sob perturbação nos regressores, nos quais, podemos observar que os modelos slash-elípticos não apresentam pontos influentes quanto a perturbação nas variáveis explicativas x_1 e x_2. Os modelos normal e t-Student$(\nu = 2,1)$ apresentam pontos influentes, que são as observações 5 e 16 para variável x_2 e a observação 16 para variável x_1.



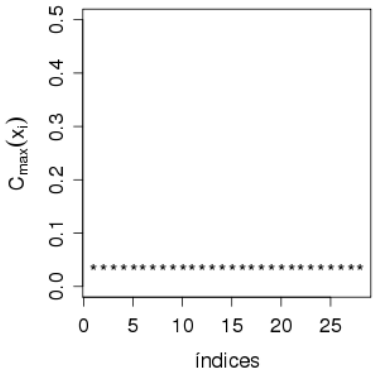
(a) x_1: normal



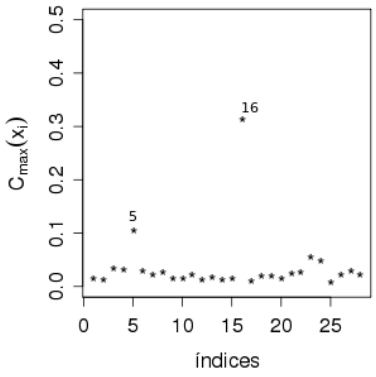
(b) x_1: slash-normal$_{(q=3)}$



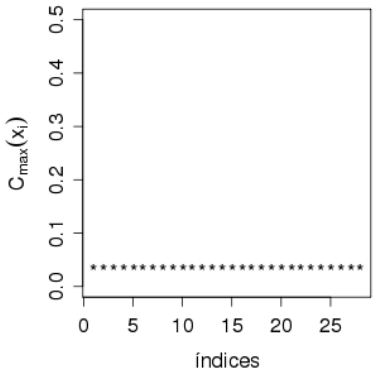
(c) x_1: t-Student$_{(\nu=2,1)}$



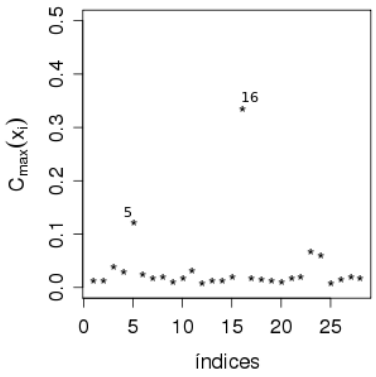
(d) x_1: slash-t-Student$_{(q=3,\nu=2,1)}$



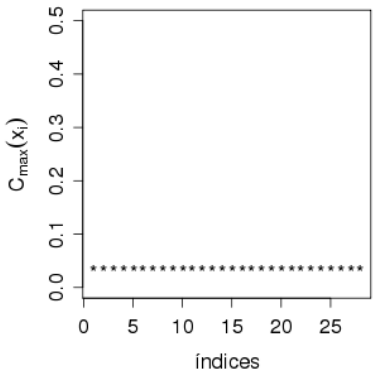
(e) x_2: normal



(f) x_2: slash-normal$_{(q=3)}$



(g) x_2: t-Student$_{(\nu=2,1)}$



(h) x_2: slash-t-Student$_{(q=3,\nu=2,1)}$

Para os dados de salinidade, podemos concluir que o modelo slash-t-Student apresentou o melhor ajuste dentre os modelos considerados. Além disso, não apresentou pontos de alavanca e pontos de influência quanto aos esquemas de perturbações considerados. No entanto, para os modelos elípticos, a observação 16 é influente.

4.2 Perda de amônia

Considere agora o conjunto de dados *stack-loss* apresentado por Becker, Chambers e Wilks (1988), que descreve um experimento de 21 dias de observações de um planta sujeita a oxidação de amônia a ácido nítrico. Estamos interessados em modelar a perda de amônia (variável resposta $\mathbf{y}$), que corresponde a 10 vezes o percentual de amônia perdida não convertida, em função do fluxo de ar (variável explicativa $\mathbf{x}_1$) e da temperatura da água (variável explicativa $\mathbf{x}_2$). Como não conhecemos o valor dos parâmetros q e ν, escolhemos os modelos com os parâmetros q e ν correspondente ao modelo com menor AIC e BIC. O parâmetro q foi selecionado na seqüência 3; 3,5; 4; 4,5;... 50 e o parâmetro ν na seqüência 2,1; 2,5; 3; 3,5; 4. Os modelos escolhidos foram

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad i = 1, \ldots, 28 \tag{4.2}$$

em que

- $\varepsilon_i \sim$ normal(μ, ϕ);
- $\varepsilon_i \sim t$-Student$(\mu, \phi, \nu = 2,1)$;
- $\varepsilon_i \sim$ slash-normal$(\mu, \phi, q = 3)$;
- $\varepsilon_i \sim$ slash-t-Student$(\mu, \phi, q = 3, \nu = 2,5)$;

Na Tabela 4.5 podemos observar os valores dos critérios AIC e BIC dos modelos selecionados para análise. As estimativas e erros-padrão dos parâmetros estimados podem ser vistos na Tabela 4.6. Observamos que os erros-padrão dos modelos slash-normal$(q = 3)$ e slash-t-Student$(q = 3, \nu = 2,5)$ são bem menores que os dos modelos normal e t-Student$(\nu = 2,1)$, com exceção do erro-padrão de $\hat{\phi}$ no modelo slash-normal$(q = 3)$ que é superior ao do modelo t-Student$(\nu = 2,1)$. De qualquer modo, o modelo slash-t-Student$(q = 3, \nu = 2,5)$ apresentou os menores erros-padrão para todos os parâmetros. As estatísticas de testes marginais de razão de verossimilhanças, Wald e escore e seus respectivos valores p, omitidos aqui, foram calculados para todos os modelos considerados. Todos os valores p foram inferiores a $0,001$ em todos os casos, isto é, todos os coeficientes foram significativos ao nível de $0,1\%$ de significância

Tabela 4.5: Qualidade do ajuste dos modelos para os dados de perda de amônia

Distribuições	AIC	BIC
normal	113,72	117,89
t-Student$(\nu = 2,1)$	110,50	114,68
slash-normal$(q = 3)$	112,59	116,77
slash-t-Student$(q = 3, \nu = 2,5)$	109,19	113,37

Tabela 4.6: Estimativas e seus erros padrão (em parêntesis) para os modelos ajustados dos dados de perda de amônia

Distribuições	β_0	β_1	β_2	ϕ
normal	-50,36 (4,7600)	0,67 (0,1200)	1,30 (0,3400)	8,99 (2,7744)
t-Student$(\nu = 2,1)$	-44,06 (3,0100)	0,82 (0,0700)	0,57 (0,2200)	2,19 (1,0551)
slash-sormal$(q = 3)$	-49,51 (0,3100)	0,83 (0,0004)	0,81 (0,0035)	3,13 (1,6800)
slash-t-Student$(q = 3, \nu = 2,5)$	-43,81 (0,1500)	0,82 (0,0001)	0,56 (0,0004)	1,21 (0,4077)

Na Tabela 4.7 temos as medidas descritivas quanto ao resíduo $\hat{r}_i$, onde podemos observar, baseado no teste de normalidade de Kolmogorov-Smirnov, que os resíduos dos modelos normal, slash-sormal$(q = 3)$ e slash-t-Student$(q = 3, \nu = 2,5)$ seguem distribuição normal. As curtoses dos resíduos dos modelos slash-elípticos são maiores que o da normal, que foi apróximadamente 3.

Na Tabela 4.8, podemos observar que apenas os resíduos $t_D(\hat{z}_i)$ do modelo normal seguem distribuição normal. A curtose dos resíduos $t_D(\hat{z}_i)$ para os modelos slash-elípticos são mais baixas que as dos modelos elípticos.

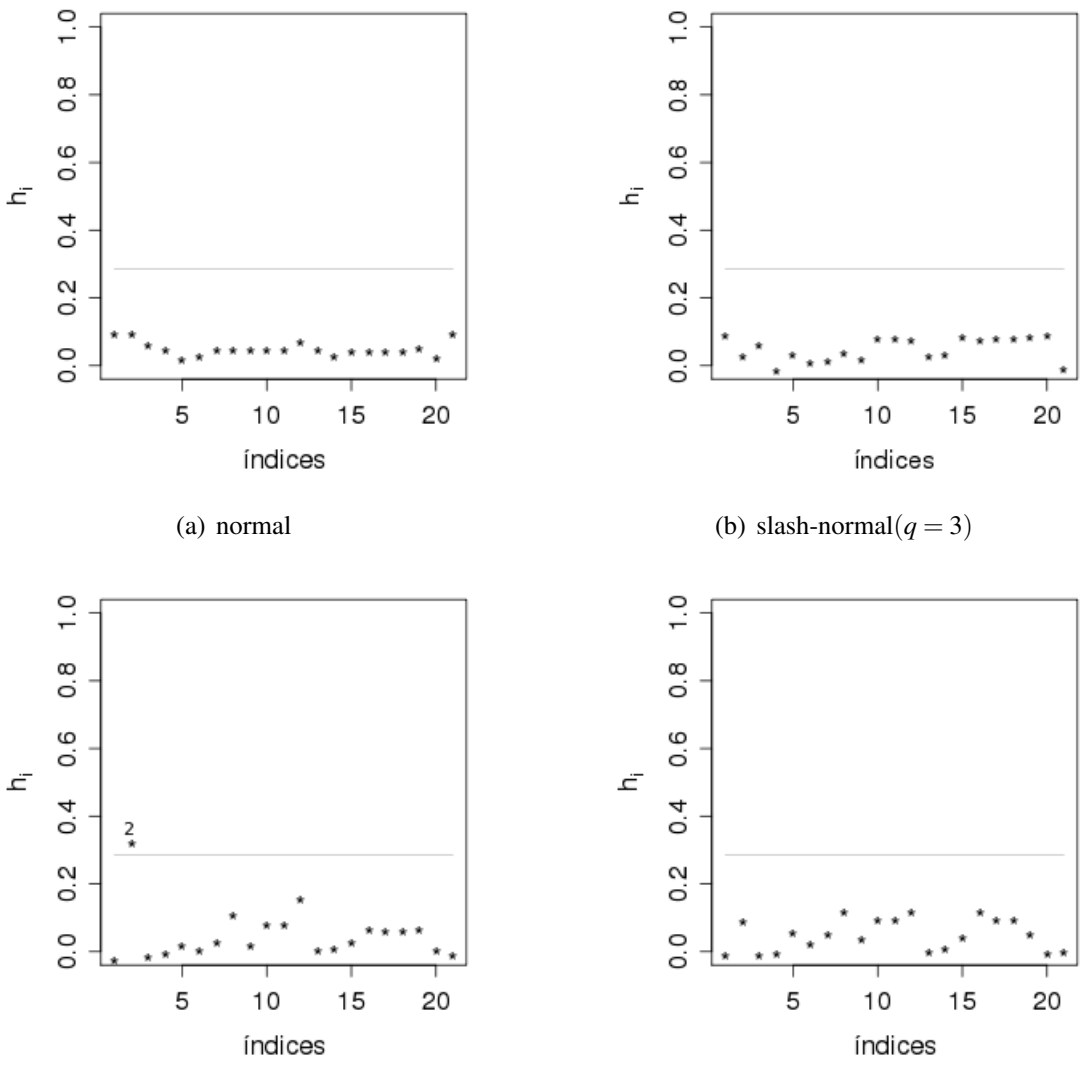
Tabela 4.7: Medidas descritivas dos resíduos $\hat{r}_i$ para os dados de perda de amônia

	normal	slash-sormal $(q = 3)$	slash-t-Student $(q = 3, \nu = 2,5)$
Média	<0,001	-0,048	0,053
Desvio padrão	1,025	1,053	0,808
Assimetria	-0,316	-0,732	-0,296
Curtose	3,159	5,409	5,270
ks-valor	0,084	0,158	0,224
p-valor	0,998	0,668	0,245

Tabela 4.8: Medidas descritivas dos resíduos $t_D(\hat{z}_i)$ para os dados de perda de amônia

	normal	t-Student ($\nu = 2,1$)	slash-sormal ($q = 3$)	slash-t-Student ($q = 3, \nu = 2,5$)
Média	<0,001	0,040	0,098	0,097
Desvio padrão	1,025	0,546	1,529	1,736
Assimetria	-0,316	-0,129	-0,312	0,017
Curtose	3,159	4,676	2,127	1,950
ks-valor	0,084	0,279	0,395	0,355
p-valor	0,998	0,075	0,003	0,010

Os gráficos de alavancagens generalizadas para os quatro modelos selecionados podem ser vistos na Figura 4.6. O modelo t-Student$(\nu = 2,1)$ foi o único dentre os quatros modelos considerados que apresentou um possível ponto de alavanca, a observação 2.



(a) normal

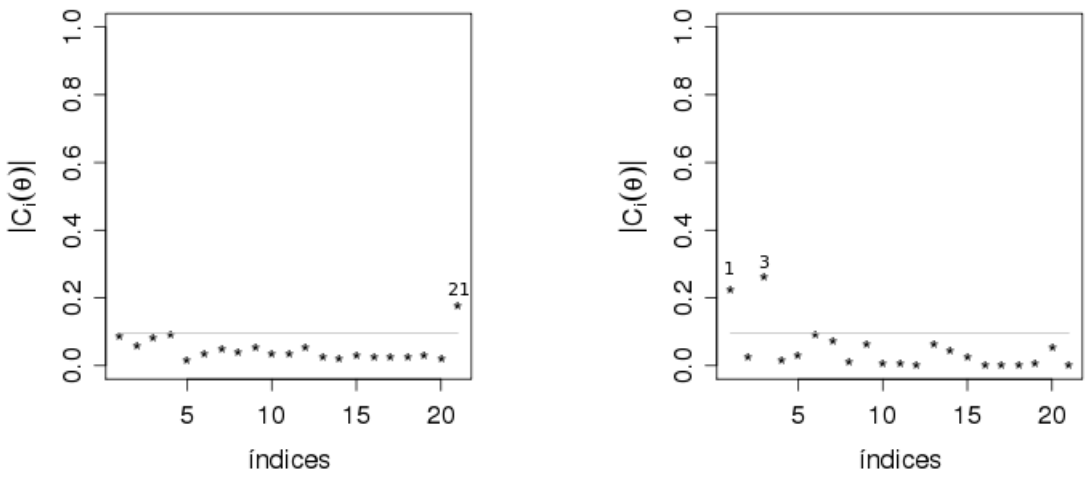
(b) slash-normal($q = 3$)

(c) t-Student($\nu = 2,1$)

(d) slash-t-Student($q = 3, \nu = 2,5$)

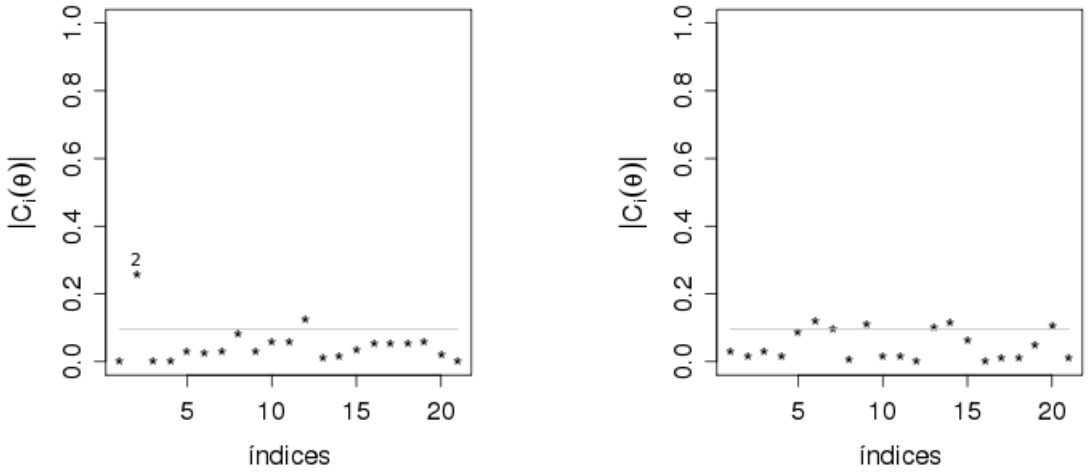
Figura 4.6: Gráficos de índices da alavancagem generalizada para os modelos ajustados aos dados de perda de amônia

Na Figura 4.7(c) observamos que no modelo t-Student($\nu = 2,1$) a observação 2 é um ponto influente quanto perturbação na escala, assim como foi observado na Figura 4.6(c), que esta observação é um possível ponto de alavanca. Há indícios que as observações 1 e 3 do modelo slash-normal($q = 3$) e 21 do modelo normal sejam pontos influentes. O modelo slash-t-Student($q = 3, \nu = 2,5$) não apresenta pontos influentes.



(a) normal

(b) slash-normal($q = 3$)

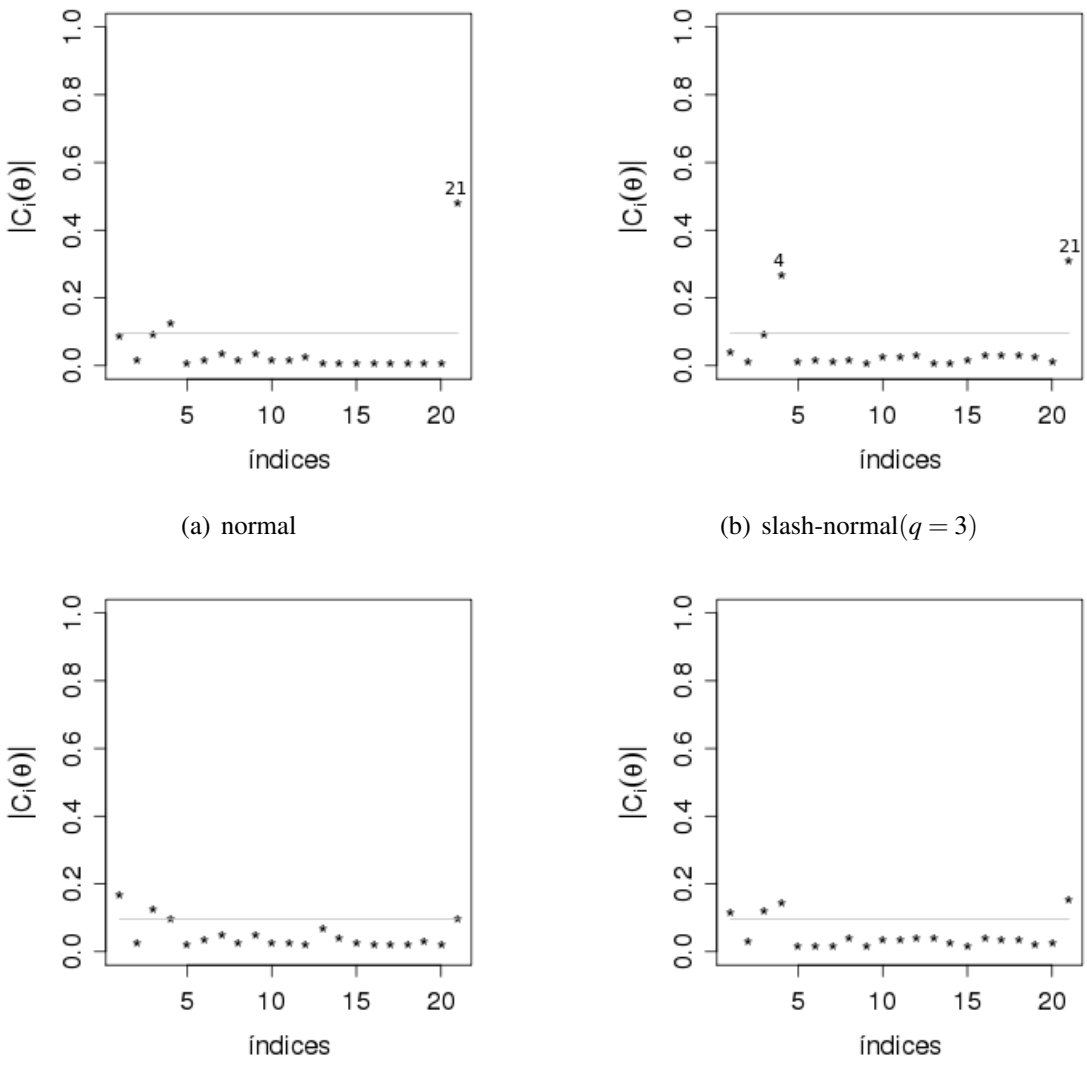


(c) t-Student($\nu = 2,1$)

(d) slash-t-Student($q = 3, \nu = 2,5$)

Figura 4.7: Gráficos de índices C_i para os modelos ajustados aos dados de perda de amônia sob perturbação na escala

Na Figura 4.8 têm-se os gráficos C_i para os dados de perda de amônia sob perturbação nos casos. Podemos notar que as observações 4 e 21 são pontos influentes nos modelos normal e slash-normal($q = 3$). Há indícios que os modelos t-Student($\nu = 2,1$) e slash-t-Student($q = 3, \nu = 2,5$) não possuam pontos influentes sob perturbações nos casos.



(a) normal

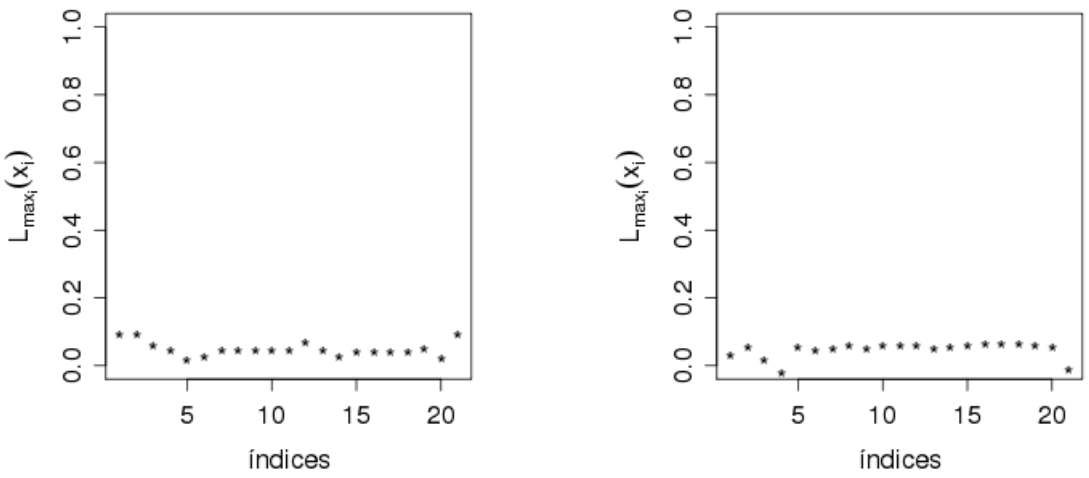
(b) slash-normal($q = 3$)

(c) t-Student($\nu = 2,1$)

(d) slash-t-Student($q = 3, \nu = 2,5$)

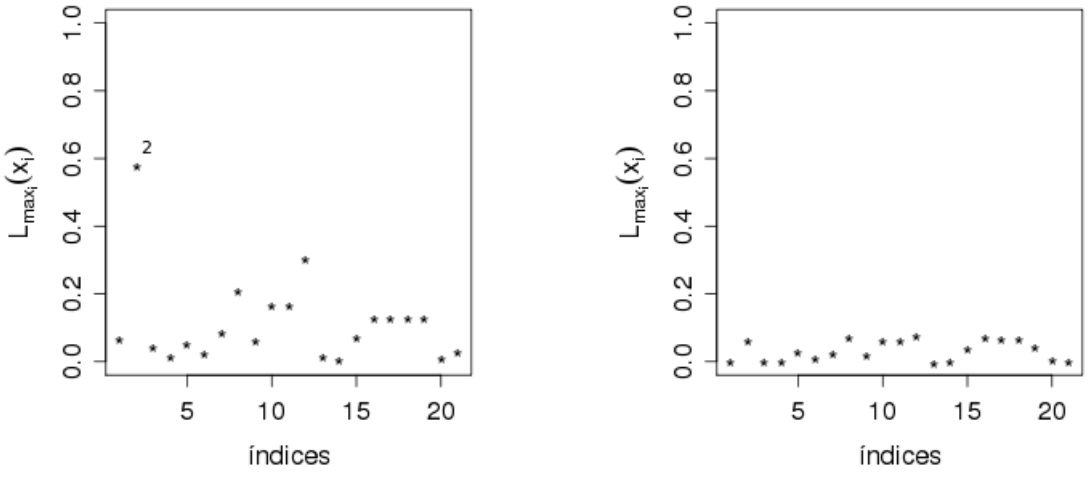
Figura 4.8: Gráficos de índices C_i para os modelos ajustados aos dados de perda de amônia sob perturbação nos casos

Na Figura 4.9 têm-se os gráficos de índices L_{max} para os dados de perda de amônia sob perturbação na resposta, nos quais, observamos que os modelos slash-elípticos e o modelo normal não apresentam pontos influentes quanto a perturbação na resposta. O modelo t-Student($\nu = 2,1$), no entanto, apresenta pelo menos um ponto influente, a observação 2. Estes gráficos estão de acordo com os gráficos de índices de alavancagem generalizada.



(a) normal

(b) slash-normal$(q=3)$



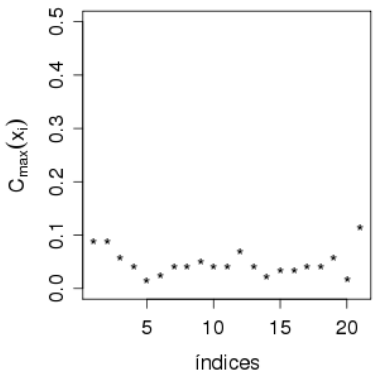
(c) t-Student$(\nu=2,1)$

(d) slash-t-Student$(q=3,\nu=2,5)$

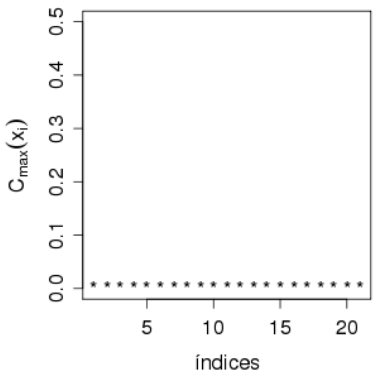
Figura 4.9: Gráficos de índices L_{max} para os modelos ajustados aos dados de perda de amônia sob perturbação na resposta

Na Figura 4.10 têm-se os gráficos C_{max} das variável x_1 e x_2 para os dados de perda de amônia sob perturbação nos regressores, nos quais, podemos observar que os modelos slash-elípticos não apresentam pontos influentes quanto a perturbação nas variáveis explicativas x_1 e x_2. Há indícios que as observações 1, 2, 3, 12 e 21 são influentes para as perturbações nas variáveis x_1 e x_2 nos modelos normal e t-Student$(\nu=2,1)$.

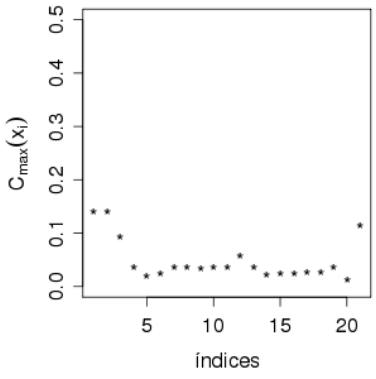
Para os dados de perda de amônia, podemos observar que os modelos slash-elípticos apresentaram melhores resultados frente aos modelos elípticos. No entanto, o modelo slash-normal apresentou possíveis pontos influentes sob perturbações na escala e nos casos, enquanto que o modelo slash-t-Student não apresenta pontos influentes para os quatros esquemas de perturbações considerados.



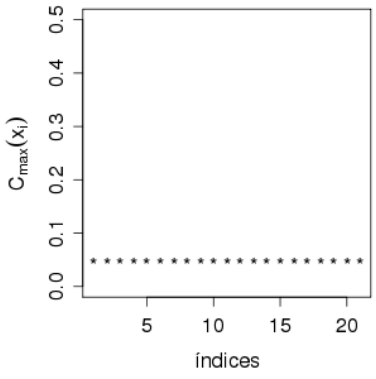
(a) x_1: normal



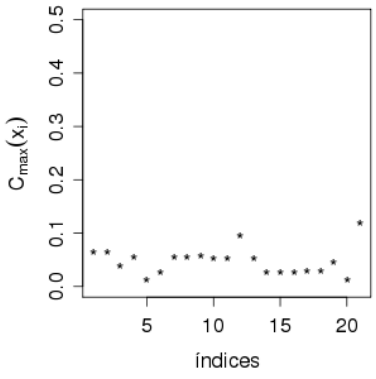
(b) x_1: slash-normal$_{(q=3)}$



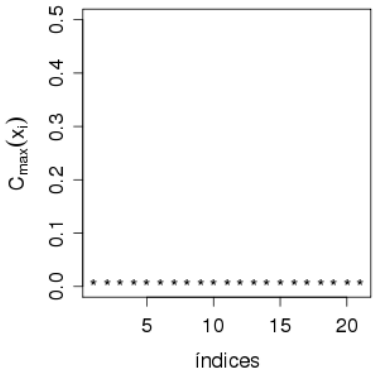
(c) x_1: t-Student$_{(\nu=2,1)}$



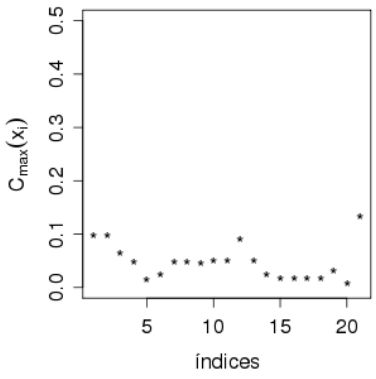
(d) x_1: slash-t-Student$_{(q=3,\nu=2,5)}$



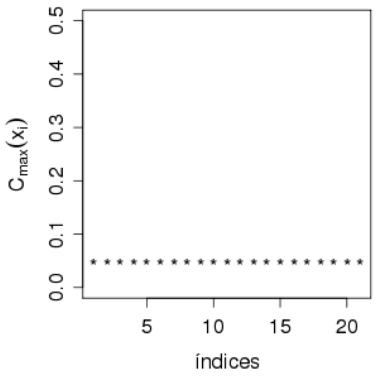
(e) x_2: normal



(f) x_2: slash-normal$_{(q=3)}$



(g) x_2: t-Student$_{(\nu=2,1)}$



(h) x_2: slash-t-Student$_{(q=3,\nu=2,5)}$

CAPÍTULO 5

Conclusões

Neste estudo propomos uma metodologia de estimação, testes de hipóteses, análise de resíduos e diagnóstico, sob o enfoque de alavancagem generalizada e influência local, para classe de modelos lineares com erros slash-elípticos com parâmetro q conhecido ou fixado. Utilizamos o algoritmo iterativo BFGS para obter as estimativas de máxima verossimilhança do vetor de parâmetros $\theta = (\beta^t, \phi)^t$. Propomos um resíduo empírico ($\hat{r}_i$) e o resíduo componente de desvio ($t_D(\hat{z}_i)$) para classe de modelos lineares com erros slash-elípticos e realizamos um estudo de simulação sobre estes resíduos para verificar as propriedades empíricas dos mesmos. Concluímos que estes resíduos não apresentam a propriedade de normalidade, mas os erros padrão estão próximos de um. O resíduo $t_D(\hat{z}_i)$ é mais simétrico e seu coeficiente de curtose está mais perto de 3 que o resíduo $\hat{r}_i$. O fato mais curioso para estes resíduos é que a média de ambos não está satisfatoriamente próxima de zero, o que sugere a existência de viés nos ajustes para esta classe de modelos. Apresentamos duas aplicações para exemplificar a metodologia proposta. O primeiro conjunto de dados analisados refere-se aos dados de salinidade do rio *Pamlico Sound* e o segundo conjunto de dados analisados refere-se aos dados de perda de amônia. Nas duas aplicações observamos que os modelos slash-t-Student apresentam melhores resultados quanto ao ajuste do modelo quando comparado com os modelos slash-normal e os modelos elípticos: normal e t-Student. Concluímos também que o modelo slash-t-Student não apresentou problemas de alavancagem ou pontos influentes nas duas aplicações.

Referências Bibliográficas

AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v. 19, n. 6, p. 716–723, 1974.

BECKER, R. A.; CHAMBERS, J. M.; WILKS, A. R. *The New S Language*. [S.l.]: Wadsworth & BrooksCole, 1988.

BERKAME, M.; BENTLER, P. M. Moments of elliptical distributed random variates. *Statistics & Probability Letters*, v. 4, p. 333–335, 1986.

BOX, G. E. P.; COX, D. R. An analysis of transformations (with discussion). *Journal of the Royal Statistical Society B*, v. 26, p. 211–252, 1964.

BROYDEN, C. G. The convergence of a class of double-rank minimization algorithms. *Journal of the Institute of Mathematical Applications*, v. 6, p. 76–90, 1970.

COOK, R. D. Assessment of local influence (with discussion). *Journal of the Royal Statistical Society B*, v. 48, n. 2, p. 133–169, 1986.

COOK, R. D.; WEISBERG, S. *Residuals and Influence in Regression*. London: Chapman and Hall, 1982.

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

CYSNEIROS, F. J. A. *Métodos Restritos e Validação de Modelos Simétricos de Regressão*. Tese (Doutorado) — Instituto de Matemática e Estatística, São Paulo, 2004.

CYSNEIROS, F. J. A.; PAULA, G. A.; GALEA, M. Modelos simétricos aplicados. In: *IX Escola de Modelos de Regressão*. São Paulo: ABE, 2005.

CYSNEIROS, F. J. A.; VANEGAS, L. H. Residuals and their statistical properties in symmetrical nonlinear models. *Statistics & Probability Letter*, v. 78, p. 3269–3273, 2008.

DOORNIK, J. A. *Object-oriented Matrix Programming Using Ox*. 3rd. ed. London: Timberlake Consultants, 1999.

FANG, K. T.; ZHANG, Y. T. *Generalized multivariate analysis*. New York: Springer-Verlag, 1990.

FLETCHER, R. A new approach to variable metric algorithm. *Computer Journal*, v. 13, p. 392–399, 1970.

GALEA, M.; BOLFARINE, H.; VILCA-LABRA, F. Influence diagnostics fo the structural error-in-variables model under the student-t distribution. *Journal of Applied Statistics*, v. 29, p. 1191–1204, 2002.

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

GALEA, M.; PAULA, G. A.; CYSNEIROS, F. J. A. On diagnostic in symmetrical nonlinear models. *Statistics & Probability Letter*, v. 73, p. 459–467, 2005.

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

GENC, A. I. A generalization of the univariate slash by a scale-mixtured exponential power distribution. *Communications in Statistics Simulation and Computation*, v. 36, n. 5, p. 937–947, 2007.

GOLDFARB, D. A family of variable metric methods derived by variational means. *Mathematics of Computations*, v. 24, p. 23–26, 1970.

GÓMEZ, H. W.; QUINTANA, F. A.; TORRES, F. J. A new family of slash-distributions with elliptical contours. *Statistics & Probability Letters*, v. 77, n. 7, p. 717–725, 2007.

GÓMEZ, H. W.; VENEGAS, O. Erratum to: "A new family of slash-distributions with elliptical contours". *Statistics & Probability Letters, 77 (2007) 717-725*, 2008.

KUHA, J. AIC and BIC: Comparisons of assumptions and performance. *Sociological Methods & Research*, v. 33, n. 3, p. 417–417, 2004.

LANGE, K. L.; LITTLE, R. J. A.; TAYLOR, J. M. G. Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, v. 84, n. 408, p. 881–896, 1989.

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

LI, B. Sensitivity of rao's score test, the wald test and the likelihood ratio test to nuisance parameters. *Journal of Statistical Planning and Inference*, v. 97, n. 1, p. 57–66, 2001.

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

PRESS, W. H.; VETTERLING, S. A.; FLANNERY, B. P. *Numerical recipes in C: the art of scientific computing*. New York: Cambridge University Press, 1992.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2008. ISBN 3-900051-07-0. Disponível em: <http://www.R-project.org>.

RUPPPERT, D.; CARROLL, R. J. Trimmed least squares estimation in the linear model. *Journal of the American Statistical Association*, v. 75, n. 372, p. 828–838, 1980.

SCHWARZ, G. Estimating the dimension of a model. *Annals of Statistics*, v. 6, n. 2, p. 461–464, 1978.

SHANNO, D. F. Conditioning of quasi-newton methods for function minimization. *Mathematics of Computations*, v. 24, p. 647–656, 1970.

THOMAS, W.; COOK, R. D. Assessing influence on predictions from generalized linear models. *Technometrics*, v. 32, n. 1, p. 59–65, 1990.

WEI, B. C.; HU, Y. Q.; FUNG, W. K. Generalized leverage and its applications. *Scandinavian Journal of Statistics*, v. 25, n. 1, p. 647–656, 1998.