

UNIVERSIDADE FEDERAL DA PARAÍBA  
CENTRO DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM  
MODELAGEM MATEMÁTICA E COMPUTACIONAL

**Análise de Influência para o modelo de  
Regressão de Valor Extremo sob Censura  
do Tipo II**

por

Victor Carvalho

Dissertação submetida como requisito parcial  
para a obtenção do grau de  
Mestre em Modelagem Matemática e Computacional

Prof<sup>ª</sup>. Dr<sup>a</sup>. Andréa Rocha  
Orientadora

João Pessoa, fevereiro de 2016.

Victor José Araújo de Carvalho

# Análise de Influência para o modelo de Regressão de Valor Extremo sob Censura do Tipo II

Dissertação submetida ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do Centro de Informática da Universidade Federal da Paraíba, como requisito parcial para a obtenção do grau de Mestre em Modelagem Matemática e Computacional.  
**Orientador:** Prof. Dr<sup>a</sup>. Andréa Vanessa Rocha.

João Pessoa – PB  
Março de 2016

C331a Carvalho, Victor José Araujo de.  
Análise de influência para o Modelo de Regressão de Valor Extremo sob Censura do tipo II / Victor José Araujo de Carvalho.- João Pessoa, 2016.  
47f.  
Orientadora: Andréa Rocha  
Dissertação (Mestrado) - UFPB/CI  
1. Matemática computacional - modelagem. 2. Modelo de Regressão de Valor Extremo. 3. Censura Tipo II.  
4. Diagnóstico de influência. 5. Alavancagem generalizada.

UFPB/BC

CDU: 519.6(043)

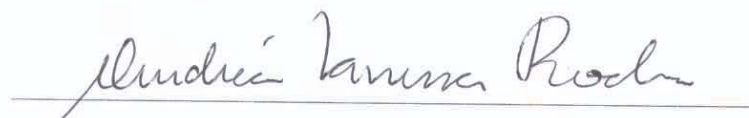
# Análise de Influência para o modelo de Regressão de Valor Extremo sob Censura do Tipo II

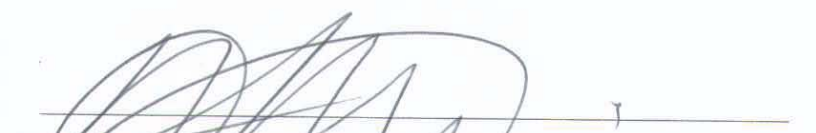
Dissertação submetida ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do Centro de Informática da Universidade Federal da Paraíba, como requisito para a obtenção do grau de Mestre em Modelagem Matemática e Computacional.

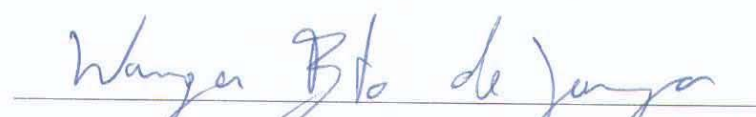
**Orientadora:** Profa. Dra. Andréa Vanessa Rocha

Aprovado em: 01 / 03 / 2016

**COMISSÃO EXAMINADORA:**

  
\_\_\_\_\_  
Profa. Dra. Andréa Vanessa Rocha – UFPB  
Orientadora

  
\_\_\_\_\_  
Prof. Dr. Alberto Masayoshi Faria Ohashi – UFPB  
Examinador

  
\_\_\_\_\_  
Prof. Dr. Wagner Barreto de Souza – UFMG  
Examinador

## AGRADECIMENTOS

- À Deus.
- A Professora Andréa Vanessa Rocha por ter orientado este trabalho, pelo conhecimento transmitido, por me motivar, transmitindo sua experiência de forma sábia e precisa.
- Ao Professor Alexandre Simas pela ajuda nos momentos de dificuldade.
- Aos professores da UFPB.
- A minha família, em especial a minha mãe, minha esposa Vanessa Furtado e meu irmão e grande amigo Vinicius, pela confiança, carinho, apoio e ajuda em todas as decisões.
- Aos meus amigos, em especial à Josenildo, Leon Tarquino, Maria Aparecida, Ricardo Pinheiro, Suelena, pela amizade e companheirismo .
- A todas as pessoas que contribuíram de forma significativa, durante todo o trajeto desde a graduação até o presente momento.
- A banca examinadora: Prof. Dr. Alberto Masayoshi Faria Ohashi e Prof. Dr. Wagner Barreto de Souza por aceitarem participar da avaliação deste trabalho.
- A Capes ao CNPq pelo apoio financeiro.

# Sumário

AGRADECIMENTOS . . . . .	v
RESUMO . . . . .	viii
ABSTRACT . . . . .	ix
1 INTRODUÇÃO . . . . .	1
2 MODELO DE REGRESSÃO DE VALOR EXTREMO . . . . .	3
3 ESTIMAÇÕES E INFERÊNCIA . . . . .	5
4 ANÁLISE DE DIAGNÓSTICO PARA O MODELO DE REGRESSÃO DO VALOR EXTREMO . . . . .	8
4.1 Influência Global . . . . .	8
4.2 Influência Local . . . . .	9
4.3 Cálculo de Curvaturas . . . . .	11
4.3.1 Perturbação de casos Ponderados . . . . .	12
4.3.2 Perturbação da variável resposta . . . . .	12
4.3.3 Perturbação da Variável Explicativa do Parâmetro de Localização ( $x_j^T$ ) . . . . .	13
4.3.4 Perturbação da Variável Explicativa do Parâmetro de Escala ( $z_j^T$ ) . . . . .	14
4.3.5 Perturbações simultâneas das variáveis explicativas $(x_j^T, z_l^T)^T$ . . . . .	14
4.4 Alavancagem Generalizada . . . . .	15
5 CENSURA DO TIPO II . . . . .	17
6 ANÁLISE DE DIAGNOSTICO . . . . .	24
6.1 Influência Global . . . . .	24
6.2 Influência Local . . . . .	25

<b>6.3</b>	<b>Cálculo de Curvaturas</b>	<b>27</b>
6.3.1	Perturbação de casos Ponderados	28
6.3.2	Perturbação da variável resposta	28
6.3.3	Perturbação da Variável Explicativa do Parâmetro de Localização ( $x_j^T$ )	29
6.3.4	Perturbação da Variável Explicativa do Parâmetro de Escala ( $z_j^T$ )	30
6.3.5	Perturbações simultâneas das variáveis explicativas ( $x_j^T, z_l^T$ ) <sup>T</sup>	30
<b>6.4</b>	<b>Alavancagem Generalizada</b>	<b>31</b>
<b>7</b>	<b>APLICAÇÃO</b>	<b>33</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>37</b>

## RESUMO

Neste trabalho, analisaremos o problema de avaliar a Influência de observações no Modelo de Regressão de Valor Extremo sob Censura tipo II. Tal modelo é bastante importante no estudo de dados de tempo de vida. Primeiramente obteremos a função log-verossimilhança, a função escore e a matriz de informação. Em seguida discutiremos alguns métodos de influência, tais como a influência global e a influência local. Na análise de influência local derivaremos as expressões para curvaturas normais sob diferentes esquemas de perturbações. Finalizaremos obtendo um expressão de forma fechada para a alavancagem generalizada.

**Palavras-chave:** Modelo de Regressão de Valor Extremo. Censura Tipo II. Diagnóstico de Influência. Alavancagem Generalizada.

## ABSTRACT

In this work, we analyze the problem of evaluating the Influence of observations in the Extreme Value Regression model under censor type II. Such a model is very important in the analysis of lifetime data. First we obtain the log-likelihood function, the function score and information matrix. Then we will discuss some methods of influence, such as global influence and local influence. In the local influence analysis we will derive expressions for normal bends under different disturbances schemes. Finalize obtaining a closed-form expression for the overall leverage.

**Palavras-chave:** Extreme value regression model. Censorship type II. Diagnosis of Influence. Leverage Generalized.

# 1 INTRODUÇÃO

Ao realizar inferência estatística em dados obtidos a partir de teste de confiabilidade, existe um problema, que é o de que não se saber a informação exata de algumas observações. Apenas temos informações parciais. Como informação sempre é importante, esta informação parcial é usada no modelo, e estas observações parciais são chamadas de dados censurados. Mais precisamente os dados censurados, são observações que por alguma razão foram interrompidas, sendo assim nem todas as observações foram obtidas.

Censuras são extremamente importantes nos processos de análise de sobrevivência, onde o tempo por diversas razões é limitado. Um caso amplamente aplicável em situações como essa é a censura do tipo II, ver Park and Lee [9].

Nosso objetivo nesse trabalho é avaliar a influência global e local das observações para o modelo de regressão de valor extremo sob censura do tipo II.

Tal análise de diagnóstico, segundo Rocha e Simas [11], é uma forma eficiente para detectar quais observações são influentes ou não no modelo. A primeira ferramenta utilizada para a análise do impacto individual das observações no processo de estimação é a exclusão de observações, essa ferramenta é conhecida como análise de influência global. No entanto é importante observar que as vezes estamos interessados em observar a influência de uma observação em algum aspecto específico e utilizando a ferramenta anterior isso não é possível uma vez que o caso de exclusão exclui todas as informações que a observação excluída fornece.

Para contornar tal problema utiliza-se uma abordagem da influência local, onde agora investiga-se o modelo de sensibilidade sob pequenas perturbações. Vários autores consideraram o método de influência local nos modelos de regressão;

ver, por exemplo, Lawrance [5], Thomas e Cook [12], Paula [10], Lesaffre e Verbeke [7] e, mais recentemente, Espinheira et al.[4], Lemonte [6], entre outros.

Nesse trabalho, apresentaremos métodos de diagnósticos baseados em influência global, influência local e alavancagem generalizada para o modelo de regressão de valor extremo mínimo, com censura do tipo II que foram inspirados em Cook (1986)[3].

## 2 MODELO DE REGRESSÃO DE VALOR EXTREMO

Uma variável aleatória  $Y$  tem distribuição de valor extremo se sua função densidade de probabilidade é dada por

$$g(y, \mu, \phi) = \phi^{-1} \exp\left(\frac{y - \mu}{\phi}\right) \exp\left(-\exp\left(\frac{y - \mu}{\phi}\right)\right) \text{ com } y \in \mathbb{R}, \quad (2.1)$$

onde  $\mu \in \mathbb{R}$  e  $\phi > 0$  são os parâmetros de localização e dispersão, respectivamente. A função geradora de momentos de  $Y$  é dada por

$$E(e^{ty}) = e^{t\mu} \Gamma(1 - t\phi), \text{ com } t < \phi^{-1}.$$

Daí, obtemos a esperança e a variância

$$E(Y) = \mu - \gamma\phi$$

e

$$Var(Y) = \frac{\pi^2}{6} \phi^2,$$

onde  $\gamma$  é a constante de Euler,  $\gamma = \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \log n \right) \approx 0,5772$ .

Seja  $Y_1, Y_2, \dots, Y_n$  uma amostra aleatória, onde cada  $Y_i$  tem função densidade de probabilidade dado como em (2.1), com parâmetros de locação  $\mu_i$  e parâmetros de escala  $\phi_i$ .

Assumimos que as componentes de ambos os vetores de parâmetros  $\mu = (\mu_1, \mu_2, \dots, \mu_n)^T$  e  $\phi = (\phi_1, \phi_2, \dots, \phi_n)^T$  variam de acordo com as observações.

O modelo de regressão com covariáveis de dispersão é definido por (2.1) e por duas componentes sistemáticas que são parametrizadas como

$$g_1(\mu) = \eta_1 = f_1(X; \beta) \quad , \quad g_2(\phi) = \eta_2 = f_2(Z; \theta), \quad (2.2)$$

onde  $g_1(\cdot)$  e  $g_2(\cdot)$  são funções de  $\mathbb{R}$  e  $\mathbb{R}^+$  respectivamente, estritamente monótonas e duas vezes diferenciáveis,  $f_1(\cdot : \beta)$  e  $f(\cdot : \theta)$  são funções não necessariamente lineares e de classe  $\mathcal{C}^2$ . Onde  $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$  e  $\theta = (\theta_1, \theta_2, \dots, \theta_q)^T$  pertencem a  $\mathbb{R}^p$  e  $\mathbb{R}^q$  respectivamente são vetores de parâmetros desconhecidos a serem estimados e  $X$  e  $Z$  são matrizes  $n \times p$  e  $n \times q$  respectivamente, cujas colunas representam diferentes covariáveis e estas matrizes tem posto  $p$  e  $q$  respectivamente.

### 3 ESTIMAÇÕES E INFERÊNCIA

A log-verossimilhança para a amostra  $Y_1, Y_2, \dots, Y_n$  com valores observados  $y_1, y_2, \dots, y_n$  é dada por

$$l(\beta, \theta) = - \sum_{i=1}^n \log \phi_i + \sum_{i=1}^n \frac{y_i - \mu_i}{\phi_i} - \sum_{i=1}^n \exp\left(\frac{y_i - \mu_i}{\phi_i}\right),$$

onde  $\mu_i$  e  $\phi_i$  são como foram definidos na capítulo anterior.

Assim como feito em Barreto-Souza[2] definimos a função escore  $U(\beta, \theta) = \left(\frac{\partial l}{\partial \beta^T}, \frac{\partial l}{\partial \theta^T}\right)^T$ .

Assim temos que

$$U_j(\beta, \theta) = \frac{\partial l}{\partial \beta_j} = - \sum_{i=1}^n \frac{1}{\phi_i} \frac{\partial \mu_i}{\partial \eta_{1i}} \frac{\partial \eta_{1i}}{\partial \beta_j} + \sum_{i=1}^n \frac{1}{\phi_i} \exp\left(\frac{y_i - \mu_i}{\phi_i}\right) \frac{\partial \mu_i}{\partial \eta_{1i}} \frac{\partial \eta_{1i}}{\partial \beta_j}$$

e

$$\begin{aligned} U_J(\beta, \theta) &= \frac{\partial l}{\partial \theta_J} = - \sum_{i=1}^n \frac{1}{\phi_i} \frac{\partial \phi_i}{\partial \eta_{2i}} \frac{\partial \eta_{2i}}{\partial \theta_J} - \sum_{i=1}^n \frac{y_i - \mu_i}{\phi_i^2} \frac{\partial \phi_i}{\partial \eta_{2i}} \frac{\partial \eta_{2i}}{\partial \theta_J} \\ &+ \sum_{i=1}^n \frac{y_i - \mu_i}{\phi_i^2} \exp\left(\frac{y_i - \mu_i}{\phi_i}\right) \frac{\partial \phi_i}{\partial \eta_{2i}} \frac{\partial \eta_{2i}}{\partial \theta_J} \end{aligned}$$

Em notação matricial temos

$$U_j(\beta, \theta) = \frac{\partial l}{\partial \beta_j} = \tilde{X}^T \Phi^{-1} M_1 u$$

e

$$U_J(\beta, \theta) = \frac{\partial l}{\partial \theta_J} = \tilde{S} \Phi^{-1} M_2 v,$$

onde definimos os vetores  $u = (-1 + \exp(\frac{y_1 - \mu_1}{\phi_1}), -1 + \exp(\frac{y_2 - \mu_2}{\phi_2}), \dots, -1 + \exp(\frac{y_n - \mu_n}{\phi_n}))^T$  e  $v = (-1 - (\frac{y_1 - \mu_1}{\phi_1}) + (\frac{y_1 - \mu_1}{\phi_1}) \exp(\frac{y_1 - \mu_1}{\phi_1}), -1 - (\frac{y_2 - \mu_2}{\phi_2}) + (\frac{y_2 - \mu_2}{\phi_2}) \exp(\frac{y_2 - \mu_2}{\phi_2}), \dots, -1 - (\frac{y_n - \mu_n}{\phi_n}) + (\frac{y_n - \mu_n}{\phi_n}) \exp(\frac{y_n - \mu_n}{\phi_n}))^T$ , as matrizes  $\tilde{X} = \left(\frac{\partial \eta_{1i}}{\partial \beta_j}\right)_{i,j}$   $n \times p$ ,  $\tilde{S} = \left(\frac{\partial \eta_{2i}}{\partial \theta_J}\right)_{i,J}$   $n \times q$ ,  $\Phi = \text{diag}(\phi_i)$ ,  $M_1 = \text{diag}\left(\frac{\partial \mu_i}{\partial \eta_{1i}}\right)$  e  $M_2 = \text{diag}\left(\frac{\partial \phi_i}{\partial \eta_{2i}}\right)$   $n \times n$ .

Os estimadores de máxima verossimilhança para os parâmetros  $\beta$  e  $\theta$  são obtidos resolvendo um sistema de equações não lineares  $U = 0$  e não ha uma forma fechada para tal resolução.

Portanto, utilizam-se algoritmos de otimização não-linear tais como o algoritmo de Newton ou quase-Newton, para encontrar os estimadores de máxima verossimilhança.

Agora obtemos um expressão para a matriz de informação observada.

$$C = C(\beta, \theta) = \begin{pmatrix} C_{\beta\beta} & C_{\beta\theta} \\ C_{\theta\beta} & C_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \tilde{X}^T Q_{\beta\beta} \tilde{X} + Q_{\beta\beta}^1 & \tilde{X}^T Q_{\beta\theta} \tilde{S} \\ \tilde{S}^T Q_{\theta\beta} \tilde{X} & \tilde{S}^T Q_{\theta\theta} \tilde{S} + Q_{\theta\theta}^4 \end{pmatrix},$$

onde,

$$\begin{aligned} Q_{\beta\beta} &= \text{diag} \left( \left[ -\frac{1}{\phi_i^2} \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) \right] \left( \frac{\partial \mu_i}{\partial \eta_{1i}} \right)^2 + \left[ \frac{1}{\phi_i} \left( \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) - 1 \right) \right] \frac{\partial^2 \mu_i}{\partial \eta_{1i}^2} \right), \\ Q_{\beta\beta}^1 &= \left( \sum_{i=1}^n \frac{1}{\phi_i} \left[ \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) - 1 \right] \frac{\partial \mu_i}{\partial \eta_{1i}} \frac{\partial^2 \eta_{1i}}{\partial \beta_j \partial \beta_l} \right)_{jl}, \\ Q_{\theta\theta} &= \text{diag} \left( \left[ \frac{1}{\phi_i^2} + 2 \left[ \left( \frac{y_i - \mu_i}{\phi_i^3} \right) - \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) \left( \frac{y_i - \mu_i}{\phi_i^3} \right) \right] - \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) \left( \frac{y_i - \mu_i}{\phi_i^2} \right)^2 \right] \left( \frac{\partial \phi_i}{\partial \eta_{2i}} \right)^2 \right) \\ &\quad - \text{diag} \left( \left[ \frac{1}{\phi_i} + \left( \frac{y_i - \mu_i}{\phi_i} \right) - \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) \left( \frac{y_i - \mu_i}{\phi_i^2} \right) \right] \frac{\partial^2 \phi_i}{\partial \eta_{2i}^2} \right), \\ Q_{\theta\theta}^4 &= \left( - \sum_{i=1}^r \left[ \frac{1}{\phi_i} + \left( \frac{y_i - \mu_i}{\phi_i} \right) - \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) \left( \frac{y_i - \mu_i}{\phi_i^2} \right) \right] \frac{\partial \phi_i}{\eta_{2i}} \frac{\partial^2 \eta_{2i}}{\partial \theta_J \partial \theta_L} \right)_{JL}, \\ Q_{\beta\theta} &= \text{diag} \left( \frac{1}{\phi_i^2} \left[ 1 - \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) - \exp \left( \frac{y_i - \mu_i}{\phi_i} \right) \left( \frac{y_i - \mu_i}{\phi_i} \right) \right] \frac{\partial \phi_i}{\partial \eta_{2i}} \frac{\partial \mu_i}{\partial \eta_{1i}} \right). \end{aligned}$$

Agora obtemos a expressão para a matriz de informação de Fisher.

Definimos as matrizes diagonais  $W_{\beta\beta} = \text{diag}(\frac{1}{\phi^2} (\frac{\partial \mu_i}{\partial \eta_{1i}})^2)$ ,  $W_{\theta\theta} = \text{diag}(c(\frac{1}{\phi} \frac{\partial \phi_i}{\partial \eta_{2i}})^2)$  e  $W_{\beta\theta} = \text{diag}(\frac{\gamma-1}{\phi_i^2} (\frac{\partial \mu_i}{\partial \eta_{1i}} \frac{\partial \phi_i}{\partial \eta_{2i}}))$ , onde a constante  $c = (1 - \gamma)^2 + \frac{\pi^2}{6} \approx 1,82368$ . A

matriz de informação de Fisher é então obtida de maneira analoga ao que foi feito em Barreto-Souza [2]

$$\begin{aligned}
 K = K(\beta, \theta) &= \begin{pmatrix} K_{\beta\beta} & K_{\beta\theta} \\ K_{\theta\beta} & K_{\theta\theta} \end{pmatrix} \\
 &= \begin{pmatrix} \tilde{X}^T W_{\beta\beta} \tilde{X} & \tilde{X}^T W_{\beta\theta} \tilde{S} \\ \tilde{S}^T W_{\theta\beta} \tilde{X} & \tilde{S}^T W_{\theta\theta} \tilde{S} \end{pmatrix}.
 \end{aligned}$$

Definimos as matrizes  $\mathbb{X}$  e  $\tilde{W}$  com dimensões  $2n \times (p + q)$  e  $2n \times 2n$  respectivamente, como  $\mathbb{X} = \begin{pmatrix} \tilde{X} & 0 \\ 0 & \tilde{S} \end{pmatrix}$  e  $\tilde{W} = \begin{pmatrix} W_{\beta\beta} & W_{\beta\theta} \\ W_{\theta\beta} & W_{\theta\theta} \end{pmatrix}$ . Então podemos escrever a matriz de informação de Fisher como:

$$K = \mathbb{X}^T \tilde{W} \mathbb{X}$$

## 4 ANÁLISE DE DIAGNÓSTICO PARA O MODELO DE REGRESSÃO DO VALOR EXTREMO

Neste capítulo, vamos avaliar a influência global e local para o Modelo de Regressão de Valor Extremo.

### 4.1 Influência Global

Tal como referido na Introdução, o primeiro passo envolvido na execução de análise de sensibilidade é concentrado em influências globais, de acordo com o caso de eliminação. O caso de eliminação é uma abordagem comum para estudar o efeito de abandonar o  $i$ -ésimo caso a partir do conjunto de dados. O caso de eliminação da  $i$ -ésima observação para o modelo é dado por

$$g_1(\mu_j) = \eta_{1j} = f_1(x_j^T; \beta) \quad \text{e} \quad g_2(\phi_j) = \eta_{2j} = f_1(z_j^T; \theta), \quad j = 1, 2, \dots, n; j \neq i$$

No que segue, um quantidade com índice “  $i$  ” indica os dados originais com a  $i$ -ésima observação excluída. Para o modelo anterior a função de log-verossimilhança é denotada por  $l_{(i)}(\zeta)$ .

Seja  $\hat{\zeta}_{(i)} = (\hat{\beta}_{(i)}^T, \hat{\theta}_{(i)}^T)^T$  o estimador de máxima verossimilhança baseado em  $\hat{\zeta}$ , obtido a partir de maximizar  $l_{(i)}(\zeta)$ . Para avaliar a influência da  $i$ -ésima observação no EMV.  $\hat{\zeta} = (\hat{\beta}^T, \hat{\theta}^T)^T$ , comparamos a diferença entre  $\hat{\zeta}_{(i)}$  e  $\hat{\zeta}$ . Se a eliminação de uma observação influencia seriamente uma estimativa, mais atenção deve ser dada a essa observação em particular. Portanto, se  $\hat{\zeta}_{(i)}$  está longe de ser  $\hat{\zeta}$ , então neste caso é considerado como uma observação influente. Uma medida inicial de influência global com base no teste de Wald é definida como a norma padronizada

de  $\widehat{\zeta}_{(i)}$  que é conhecida como a distância generalizada de Cook é dada pela seguinte expressão

$$GD_i(\zeta) = (\widehat{\zeta}_{(i)} - \widehat{\zeta})^T [-C] (\widehat{\zeta}_{(i)} - \widehat{\zeta})$$

Outra alternativa é o de avaliar os valores  $GD_i(\beta)$  e  $GD_i(\theta)$ , revelam que o impacto da  $i$ -ésima observação sobre a previsão de  $\beta$  e  $\theta$ , respectivamente. Outra medida popular é a da diferença da verossimilhança entre  $\widehat{\zeta}_{(i)}$  e  $\widehat{\zeta}$

$$LD_i(\zeta) = 2[l(\widehat{\zeta}) - l(\widehat{\zeta}_{(i)})],$$

Além disso, também podemos calcular  $\widehat{\beta}_j - \widehat{\beta}_{(i)j}$  ( $j = 1, 2, \dots, p$ ), para obter a diferença entre  $\widehat{\beta} - \widehat{\beta}_{(i)}$  e calcular  $\widehat{\theta}_j - \widehat{\theta}_{(i)j}$  ( $j = 1, 2, \dots, p$ ), para saber a diferença entre  $\widehat{\theta} - \widehat{\theta}_{(i)}$ . Outras medidas de influência global também são possíveis. Podemos observar o comportamento de uma estatística de teste sob um esquema de casos de eliminação. Tais estatísticas podem ser, por exemplo o teste de Wald para as covariáveis explicativas.

## 4.2 Influência Local

O método de influência local é recomendado quando a preocupação está relacionada em investigar a sensibilidade do modelo em algumas perturbações menores nos modelos ou dados. Seja  $w$  um vetor de perturbações  $t$ -dimensional, então, denotamos a função de log-verossimilhança perturbada por  $l(\zeta|w)$ . Suponhamos que existe um vetor não perturbado, ou seja  $w_0$  tal que  $l(\zeta|w_0) = l(\zeta)$ . A influência de perturbações menores no estimador  $\widehat{\zeta}$  de máxima verossimilhança pode ser avaliado usando o deslocamento da log-verossimilhança

$$LD_w = 2(l(\widehat{\zeta}) - l(\widehat{\zeta}_w))$$

onde  $\widehat{\zeta}_w$  denota o estimado de máxima verossimilhança em  $l(\widehat{\zeta}|w)$ .

A ideia de [3] para avaliar a influência local é essencialmente analisar o comportamento local de  $LD_w$  em torno de  $w_0$  avaliando a curvatura da curva de

$LD_{w_0+ad}$ , onde  $a \in \mathbb{R}$  e  $d$  é uma direção normal unitária. Uma das medidas de particular interesse é a direção  $d_{\max}$  correspondente a maior curvatura  $C_{d_{\max}}$ . O índice da curva de  $d_{\max}$  pode indicar se essa observação tem uma influência considerável sobre  $LD_w$  sob perturbações menores. Além disso a curva de  $d_{\max}$  em relação aos valores de variáveis independentes pode ser útil para identificar padrões atípicos. [3] mostrou que a curvatura normal na direção de  $d$  é dada pela seguinte expressão

$$C_d(\zeta) = 2|d^T \Delta^T (-C)^{-1} \Delta d|$$

onde  $\Delta = \frac{\partial^2 l(\zeta|w)}{\partial \zeta \partial w^T}$ , e ambos  $\Delta$  e  $K$  são avaliados em  $\hat{\zeta}$  e  $w_0$ .

Assim  $C_{d_{\max}}$  é o maior autovalor de  $B = -\Delta^T C^{-1} \Delta$  e  $d_{\max}$  é o autovetor normal unitário correspondente. O índice da curva de  $d_{\max}$  para a matriz  $B$  pode mostrar como perturbar o modelo ou dados para se obter grandes mudanças no estimador  $\zeta$ .

Portanto, se o interesse está em calcular a influência para  $\beta$ , a curvatura normal na direção do vetor  $d$  é

$$C_{d;\beta}(\zeta) = 2|d^T \Delta^T (C^{-1} - C_{\theta\theta}) \Delta d|,$$

onde  $C_{\theta\theta} = \begin{pmatrix} 0 & 0 \\ 0 & C_{\theta\theta}^{-1} \end{pmatrix}$ , e  $d_{\max;\beta}$  aqui é o autovetor normal unitário correspondente ao maior autovalor de  $B_1 = -\Delta^T (C^{-1} - C_{\theta\theta}) \Delta$ .

O índice da curva do maior autovetor de  $B_1$  pode revelar a influência desta observação sobre  $\hat{\beta}$ .

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

$$C_{d;\theta}(\zeta) = 2|d^T \Delta^T (C^{-1} - C_{\beta\beta}) \Delta d|,$$

onde  $C_{\beta\beta} = \begin{pmatrix} C_{\beta\beta}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ , e  $d_{\max;\theta}$  aqui é o vetor normal unitário correspondente ao autovalor de  $B_2 = -\Delta^T(C^{-1} - C_{\beta\beta})\Delta$ . Portanto o índice da curva do maior autovetor de  $B_2$  revela a influência desta observação sobre  $\hat{\theta}$ .

Outro procedimento é a curvatura local total correspondente ao  $i$ -ésimo elemento, que segue, tomando  $d_i$  como sendo o vetor  $n \times 1$  de zeros com 1 na  $i$ -ésima coordenada. Assim a curvatura na direção de  $d_i$  assume a forma

$$C_i(\zeta) = 2|\Delta_i^T(-C^{-1})\Delta_i|,$$

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

$$C_{i;\beta}(\zeta) = 2|\Delta_i^T(-C^{-1} - C_{\theta\theta})\Delta_i|,$$

enquanto que se o interesse esta em  $\theta$ , temos

$$C_{i;\theta}(\zeta) = 2|\Delta_i^T(-C^{-1} - C_{\beta\beta})\Delta_i|.$$

### 4.3 Cálculo de Curvaturas

No que segue, calcularemos cinco sistemas diferentes de perturbações para a matriz,

$$\Delta = \{\Delta_{ri}\}_{(p+q) \times n} = \left\{ \frac{\partial^2 l(\zeta|w)}{\partial \zeta_r \partial w_i} \right\} = \begin{pmatrix} \Delta\beta \\ \Delta\theta \end{pmatrix},$$

onde  $r = 1, 2, \dots, p + q$  e  $i = 1, 2, \dots, n$ .

Considerando o modelo definido em (2.2) e sua função de log-verossimilhança dada por (5.1). No que segue as quantidades, distingue-se pela adição de “ $\hat{\phantom{x}}$ ” são avaliadas em  $\hat{\zeta} = (\hat{\beta}^T, \hat{\theta}^T)^T$ .

### 4.3.1 Perturbação de casos Ponderados

A perturbação dos casos é feita através da definição de alguns pesos para cada observação na função de log-verossimilhança da seguinte forma

$$l(\zeta|w) = \sum_{i=1}^n w_i l_i(\zeta|w)$$

onde  $w = (w_1, w_2, \dots, w_n)^T$  é o vetor de peso total, com  $0 \leq w_i \leq 1$ , para  $i = 1, 2, \dots, n$ , e  $w_0 = (1, 1, \dots, 1)^T$  é o vetor de não perturbação.

As matrizes  $\Delta\beta$  e  $\Delta\theta$  são dadas por

$$\begin{aligned}\Delta\beta &= \tilde{X}^T \Phi^{-1} M_1 \text{diag}(u) \\ \Delta\theta &= \tilde{S}^T \Phi^{-1} M_2 \text{diag}(v)\end{aligned}$$

onde todas as matrizes foram anteriormente definidas.

### 4.3.2 Perturbação da variável resposta

Vamos considerar que cada  $y_i$  é perturbado como  $y_{iw} = y_i + w_i S_y^i$ , onde  $S_y$  é um escalar que pode ser por exemplo, em  $i$  constante igual ao desvio padrão estimado em  $y = (y_1, y_2, \dots, y_n)^T$ . Neste caso, a perturbação da log-verossimilhança é dada por

$$l(\zeta|w) = \sum_{i=1}^n l_i^r(\mu_i, \phi_i, w_i)$$

onde,

$$l_i^r(\mu_i, \phi_i, w_i) = -\log \phi_i + \left( \frac{y_{iw} - \mu_i}{\phi_i} \right) - \exp \left( \frac{y_{iw} - \mu_i}{\phi_i} \right)$$

e neste caso,  $w_0 = (0, 0, \dots, 0)^T$  é o vetor de não perturbação. As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem as formas

$$\begin{aligned}\Delta\beta &= \tilde{X}^T Q_y \Phi^{-1} M_1 T_\beta \\ \Delta\theta &= \tilde{S}^T Q_y \Phi^{-1} M_2 T_\theta\end{aligned}$$

onde  $y_w = (y_{1w}, y_{2w}, \dots, y_{nw})^T$ ,  $Q_y = \text{diag}(S_y^1, \dots, S_y^r, \dots, S_y^r)$ ,

$$T_\beta = \text{diag} \left( \exp \left( \frac{y_i + w_i S_y^i - \mu_i}{\phi_i} \right) \frac{1}{\phi} \right)$$

e

$$T_\theta = \text{diag} \left( \frac{1}{\phi_i} + \exp \left( \frac{y_i + w_i S_y^i - \mu_i}{\phi_i} \right) \left( \frac{y_i + w_i S_y^i - \mu_i}{\phi_i^2} \right) + \exp \left( \frac{y_i + w_i S_y^i - \mu_i}{\phi_i} \right) \frac{1}{\phi_i} \right).$$

### 4.3.3 Perturbação da Variável Explicativa do Parâmetro de Locação $(x_j^T)$

Considerando agora uma perturbação aditiva em uma variável explicativa, em particular, seja  $x_j$ , dada por  $x_{ijw} = x_{ij} + w_i S_x$ , onde  $S_x$  é um escalar que pode por exemplo ser o desvio padrão estimado de  $x_j$ . Este esquema de perturbação leva à seguinte expressão para a função de log-verossimilhança

$$l(\zeta|w) = \sum_{i=1}^n l_i^{e^x}(\mu_i, \phi_i, w_i),$$

onde

$$l_i^{e^x}(\mu_i, \phi_i, w_i) = -\log \phi_i + \left( \frac{y_i - \mu_{iw}}{\phi_i} \right) - \exp \left( \frac{y_i - \mu_{iw}}{\phi_i} \right),$$

onde,  $\mu_{iw} = g_1^{-1}(\eta_{1i})$ ,  $\eta_{1i} = f_1(x_{iw}^T; \beta)$  com  $x_{iw} = (x_{i1w}, x_{i2w}, \dots, x_{ipw})$ . Consequentemente,  $w_0 = (0, 0, \dots, 0)^T$  é o vetor de não perturbação. Seja  $\tilde{X}(w) = \frac{\partial \eta_{1i}}{\partial \beta}$ , e observe que  $\tilde{X}(w_0) = \tilde{X}$ . Portanto, seja  $\tilde{X}_w = \frac{\partial \eta_{1i}}{\partial w}$ . No entanto, os parenteses  $(w)$  querem dizer que as quantidades são avaliadas em  $\mu_{iw}$  em vez de  $\mu_i$ . As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem as formas

$$\begin{aligned} \Delta\beta &= \hat{\tilde{X}}(w)^T Q_{\beta\beta}(w) \hat{\tilde{X}}_w + Q_{\beta\beta w}^1(w) \\ \Delta\theta &= \hat{\tilde{S}}^T(w) Q_{\theta\beta}(w) \hat{\tilde{X}}_w. \end{aligned}$$

#### 4.3.4 Perturbação da Variável Explicativa do Parâmetro de Escala ( $z_j^T$ )

Uma vez que estamos lidando com covariáveis de dispersão, pode ser de interesse para certificar se uma observação particular é influente para o parâmetro de escala do modelo. Conseqüentemente, uma perturbação aditiva em uma variável explicativa do parâmetro de precisão da regressão, ou seja,  $z_j$  dada por  $z_{ijw} = z_{ij} + w_i S_z$ , onde  $S_z$  é um escalar que pode ser por exemplo, o desvio padrão de  $z_j$ . Este esquema de perturbação leva a seguinte expressão para a função de log-verossimilhança

$$l(\zeta|w) = \sum_{i=1}^n l_i^{e^z}(\mu_i, \phi_i, w_i)$$

onde

$$l_i^{e^z}(\mu_i, \phi_i, w_i) = -\log \phi_{iw} + \left( \frac{y_i - \mu_i}{\phi_{iw}} \right) - \exp \left( \frac{y_i - \mu_i}{\phi_{iw}} \right),$$

onde,  $\phi_{iw} = g_2^{-1}(\eta_{2i})$ ,  $\eta_{2i} = f_2(z_{iw}^T; \theta)$  com  $z_{iw} = (z_{i1w}, z_{i2w}, \dots, z_{iqw})$ . Conseqüentemente,  $w_0 = (0, 0, \dots, 0)^T$  é o vetor de não perturbação. Seja  $\tilde{S}(w) = \frac{\partial \eta_{2i}}{\partial \theta}$ , e observe que  $\tilde{S}(w_0) = \tilde{S}$ . Portanto, seja  $\tilde{S}_w = \frac{\partial \eta_{2i}}{\partial w}$ . No entanto, os parenteses ( $w$ ) querem dizer que as quantidades são avaliadas em  $\phi_{iw}$  em vez de  $\phi_i$ . As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem as formas

$$\begin{aligned} \Delta\beta &= \hat{X}^T(w) Q_{\beta\theta}(w) \hat{S}_w \\ \Delta\theta &= \hat{S}^T(w) Q_{\theta\theta}(w) \hat{S}_w + Q_{\theta\theta w}^4(w). \end{aligned}$$

#### 4.3.5 Perturbações simultâneas das variáveis explicativas ( $x_j^T, z_l^T$ )<sup>T</sup>

Agora consideramos uma perturbação aditiva em uma variável explicativa particular do parâmetro de média da regressão e uma perturbação aditiva em uma variável explicativa do parâmetro de precisão da regressão, ou seja,  $x_j$  e  $z_l$ , fazendo  $x_{ijw} = x_{ij} + w_i S_x$  e  $z_{ilw} = z_{il} + w_i S_z$ , respectivamente, onde  $S_x$  e  $S_z$  são fatores escalares que podem ser, por exemplo, os desvios padrão estimados de  $x_j$  e

$z_l$ , respectivamente. Este esquema de perturbação usa à seguinte expressão para a função de log-verossimilhança:

$$l(\zeta|w) = \sum_{i=1}^n l_i^{e^{xz}}(\mu_i, \phi_i, w_i),$$

onde,

$$l_i^{e^{xz}}(\mu_i, \phi_i, w_i) = -\log \phi_{iw} + \left( \frac{y_i - \mu_{iw}}{\phi_{iw}} \right) - \exp \left( \frac{y_i - \mu_{iw}}{\phi_{iw}} \right),$$

onde,  $\mu_{iw} = g_1^{-1}(\eta_{1iw})$ ,  $\eta_{1iw} = f_1(x_{iw}^T; \beta)$ , com  $x_{iw} = (x_{i1w}, x_{i2w}, \dots, x_{ipw})^T$  e  $\phi_{iw} = g_2^{-1}(\eta_{2iw})$ ,  $\eta_{2iw} = f_2(z_{iw}^T; \theta)$ , com  $z_{iw} = (z_{i1w}, z_{i2w}, \dots, z_{iqw})^T$ . Aqui  $w_0 = (0, 0, \dots, 0)^T$  é o vetor de não perturbação. Tudo isso, usando que os parenteses ( $w$ ) significam que as quantidades são avaliadas em  $\mu_{iw}$  e  $\phi_{iw}$  em vez de  $\mu_i$  e  $\phi_i$ . As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem as formas

$$\begin{aligned} \Delta\beta &= \widehat{X}^T(w) Q_{\beta\beta}(w) \widehat{X}_w + Q_{\beta\beta}^1(w) + \widehat{X}^T(w) Q_{\beta\theta}(w) \widehat{S}_w \\ \Delta\theta &= \widehat{S}^T(w) Q_{\theta\beta}(w) \widehat{X}_w + \widehat{S}^T(w) Q_{\theta\theta}(w) \widehat{S}_w + Q_{\theta\theta}^4(w). \end{aligned}$$

#### 4.4 Alavancagem Generalizada

Agora calcularemos a alavancagem generalizada proposta por [14], que é definida como:

$$GL(\zeta) = \frac{\partial \widehat{y}}{\partial y^T},$$

onde  $\zeta$  é um r-vetor tal que  $E(y) = \mu(\zeta)$  e  $\widehat{\zeta}$  é um estimador de  $\zeta$ , com  $\widehat{y} = \mu(\widehat{\zeta})$ . Aqui, o  $(i, l)$  elemento de  $GL(\zeta)$ , ou seja, a alavancagem generalizada do estimador  $\widehat{\zeta}$  pelo  $(i, l)$  é o valor da taxa de variação instantânea no  $i$ -ésimo valor previsto em relação ao  $l$ -ésimo valor resposta. Como observamos pelas reparametrização e observações onde  $GL_{iJ}$  são grandes pontos de alavancagem. Wei mostrou que o efeito da alavanca generalizado é obtido através

$$GL(\zeta) = D_{\zeta}(-C)C_{\zeta y},$$

em  $\zeta = \hat{\zeta}$ , onde  $D_\zeta = \frac{\partial \mu}{\partial \zeta}$  e  $C_{\zeta y} = \frac{\partial^2 l(\zeta)}{\partial \zeta \partial y}$ . De acordo com o modelo definido em (2.2), temos

$$D_\zeta = \begin{pmatrix} M_1 \tilde{X} & 0 \end{pmatrix} \quad \text{e} \quad K_{\zeta y} = \begin{pmatrix} \tilde{X} \Phi^{-1} M_1 T_\beta \\ \tilde{S} \Phi^{-1} M_2 T_\theta \end{pmatrix}$$

## 5 CENSURA DO TIPO II

Como dito na introdução desejamos trabalhar com o Modelo de Regressão de Valor Extremo sob Censura do Tipo II.

Seja  $Y_1, Y_2, \dots, Y_k$  uma amostra aleatória, onde cada  $Y_i$  tem função de densidade de probabilidade dada como em (2.1), com parâmetros de locação  $\mu_i$ , e parâmetros de escala  $\phi_i$ .

O modelo de regressão com covariáveis de dispersão é definido como no Capítulo 2.

Obtemos  $Y_{1,i}, Y_{2,i}, \dots, Y_{n_i,i}$  observações independentes e identicamente distribuídas. A censura tipo II é feita em cima de cada coluna da "matriz"  $Y = \{Y_{si}\}$  com  $s = 1, 2, \dots, n_i$  e  $i = 1, 2, \dots, k$ .

Assim ordenamos,

$$Y_{(1,n_i)} \leq Y_{(2,n_i)} \leq \dots \leq Y_{(r,n_i)} \leq Y_{(r,n_i)} \leq \dots \leq Y_{(n_i,n_i)}.$$

Para todo  $i=1,2,\dots,k$ , os vetores

$$Y_i = (Y_{1,i}, Y_{2,i}, \dots, Y_{n_i,i}) \quad \text{são independentes.}$$

Como  $Y_{(1,n_i)}, Y_{(2,n_i)}, \dots, Y_{(r,n_i)}, Y_{(r,n_i)}, \dots, Y_{(n_i,n_i)}$ , estão sob censura do tipo II, sua log-verossimilhança é dada como segue

$$L_i(\beta; \theta) = \frac{n_i!}{(n_i - r)!} \prod_{s=1}^{n_i} L_{si}(\beta; \theta),$$

onde,  $L_{si} = g(y_{(s,n_i)}; \beta; \theta)$ , se  $s \leq r$  e  $L_{si} = [1 - G(y_{(r,n_i)}; \beta; \theta)]$ , se  $s > r$ .

Como  $Y_1, Y_2, \dots, Y_k$  são independentes, temos

$$L(\beta; \theta) = \prod_{i=1}^k \frac{n_i!}{(n_i - r)!} \prod_{s=1}^{n_i} L_{si}(\beta; \theta).$$

e

$$l(\beta; \theta) = \sum_{i=1}^k \left[ \log \left( \frac{n_i!}{(n_i - r)!} \right) + l_{si}(\beta; \theta) \right],$$

onde

$$l_{si}(\beta; \theta) = \sum_{i=1}^r \log(g(y_{(s,n_i)}; \beta; \theta)) + (n_i - r) \log[1 - G(y_{(r,n_i)}; \beta; \theta)].$$

Daí

$$\begin{aligned} l(\beta; \theta) &= \sum_{i=1}^k \left[ \log \left( \frac{n_i!}{(n_i - r)!} \right) - r \log(\phi_i) + \sum_{s=1}^r \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i} \right) \right. \\ &\quad \left. - \sum_{s=1}^r \exp \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i} \right) - (n_i - r) \exp \left( \frac{y_{(r,n_i)} - \mu_i}{\phi_i} \right) \right]. \end{aligned} \quad (5.1)$$

Temos que para  $i = 1, 2, \dots, k$ , fixado,  $Y_{si} \sim EV(\mu_{si}, \phi_{si})$ , onde  $\mu_{si}$  e  $\phi_{si}$  satisfazendo

$$g_1(\mu_{si}) = g_1(\mu_i) = \eta_{1si} = f_1(x_i^T; \beta)$$

e

$$g_2(\phi_{si}) = g_2(\phi_i) = \eta_{2si} = f_2(z_i^T; \theta)$$

Observe que  $\mu_{si} = \mu_i$  e  $\phi_{si} = \phi_i$  são constantes em  $i$  para  $s$  fixado.

Agora, como anteriormente definimos a função escore por  $U \equiv U(\beta, \theta) = \left( \frac{\partial l}{\partial \beta^T}, \frac{\partial l}{\partial \theta^T} \right)$ , onde

$$\begin{aligned} U_j(\beta; \theta) = \frac{\partial l}{\partial \beta_j} &= \left[ - \sum_{i=1}^k \sum_{s=1}^r \frac{1}{\phi_i} + \sum_{i=1}^k \sum_{s=1}^r \frac{1}{\phi_i} \exp \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i} \right) + \right. \\ &\quad \left. + \sum_{i=1}^k \frac{n_i - r}{\phi_i} \exp \left( \frac{y_{(r,n_i)} - \mu_i}{\phi_i} \right) \right] \frac{\partial \mu_{si}}{\partial \eta_{1si}} \frac{\partial \eta_{1si}}{\partial \beta_j} \end{aligned}$$

$$U_J(\beta; \theta) = \frac{\partial l}{\partial \theta_J} = \left[ - \sum_{i=1}^k \sum_{s=1}^r \frac{1}{\phi_i} - \sum_{i=1}^k \sum_{s=1}^r \left( \frac{y_{(s,ni)} - \mu_i}{\phi_i^2} \right) + \sum_{i=1}^k \sum_{s=1}^r \left( \frac{y_{(s,ni)} - \mu_i}{\phi_i^2} \right) \exp \left( \frac{y_{(s,ni)} - \mu_i}{\phi_i} \right) + \sum_{i=1}^k (n_i - r) \left( \frac{y_{(r,ni)} - \mu_i}{\phi_i} \right) \exp \left( \frac{y_{(r,ni)} - \mu_i}{\phi_i} \right) \right] \frac{\partial \phi_{si}}{\partial \eta_{2si}} \frac{\partial \eta_{2si}}{\partial \theta_J},$$

assim obtemos as seguintes expressões matriciais,

$$U_j = \tilde{X}^T M_1 u_s$$

e

$$U_J = \tilde{S}^T M_2 v_s,$$

onde definimos os vetores  $u_s^T = \left( \left( -\frac{1}{\phi_1} + \frac{1}{\phi_1} \exp\left(\frac{y_{(1,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} + \frac{1}{\phi_2} \exp\left(\frac{y_{(1,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( -\frac{1}{\phi_1} + \frac{1}{\phi_1} \exp\left(\frac{y_{(2,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} + \frac{1}{\phi_2} \exp\left(\frac{y_{(2,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( -\frac{1}{\phi_1} + \frac{1}{\phi_1} \exp\left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} + \frac{1}{\phi_2} \exp\left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( -\frac{1}{\phi_1} + \frac{1}{\phi_1} \exp\left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} + \frac{1}{\phi_2} \exp\left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( -\frac{1}{\phi_k} + \frac{1}{\phi_k} \exp\left(\frac{y_{(r,n_k)} - \mu_k}{\phi_k}\right) \right) \right), v_s^T = \left( \left( -\frac{1}{\phi_1} - \left(\frac{y_{(1,n_1)} - \mu_1}{\phi_1^2}\right) + \left(\frac{y_{(1,n_1)} - \mu_1}{\phi_1^2}\right) \exp\left(\frac{y_{(1,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} - \left(\frac{y_{(1,n_2)} - \mu_2}{\phi_2^2}\right) + \left(\frac{y_{(1,n_2)} - \mu_2}{\phi_2^2}\right) \exp\left(\frac{y_{(1,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( -\frac{1}{\phi_1} - \left(\frac{y_{(2,n_1)} - \mu_1}{\phi_1^2}\right) + \left(\frac{y_{(2,n_1)} - \mu_1}{\phi_1^2}\right) \exp\left(\frac{y_{(2,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} - \left(\frac{y_{(2,n_2)} - \mu_2}{\phi_2^2}\right) + \left(\frac{y_{(2,n_2)} - \mu_2}{\phi_2^2}\right) \exp\left(\frac{y_{(2,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( -\frac{1}{\phi_1} - \left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1^2}\right) + \left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1^2}\right) \exp\left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1}\right) \right), \left( -\frac{1}{\phi_2} - \left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2^2}\right) + \left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2^2}\right) \exp\left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( \left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1^2}\right) \exp\left(\frac{y_{(r,n_1)} - \mu_1}{\phi_1}\right) \right), \left( \left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2^2}\right) \exp\left(\frac{y_{(r,n_2)} - \mu_2}{\phi_2}\right) \right), \dots, \left( \left(\frac{y_{(r,n_k)} - \mu_k}{\phi_k^2}\right) \exp\left(\frac{y_{(r,n_k)} - \mu_k}{\phi_k}\right) \right) \right)^T$ , definimos  $\tilde{X} = \left( \frac{\partial \eta_{1si}}{\partial \beta_j} \right)$  onde  $\tilde{X}$  é uma array  $n \times k \times p$  que pode ser pensado como  $n$  matrizes  $k \times p$ . Para simplificar as operação computacionais, considere  $\tilde{X}$  como a matriz  $(n * k) \times p$  formada pela concatenação dessas  $n$  matrizes. De maneira analoga definimos  $\tilde{S} = \left( \frac{\partial \eta_{2si}}{\partial \theta_j} \right)$  e interpretamos como uma matriz  $(n * k) \times q$ , obtida por concatenação como anteriormente, definimos ainda  $M_1 = \left( \frac{\partial \mu_{si}}{\partial \eta_{si}} \right)$  e  $M_2 = \left( \frac{\partial \phi_{si}}{\partial \eta_{2si}} \right)$  que intertretaremos como matrizes  $(n * k) \times (n * k)$ , obtidas por uma concatenação.

Calcularemos as derivadas de segunda ordem para obtermos a matriz de informação observada e depois a matriz de informação de esperada.





$$Q_{\beta\beta}^1 = \left( \left[ \sum_{i=1}^k \sum_{s=1}^r \left[ \frac{1}{\phi_i} \left( \exp \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i} \right) - 1 \right) \right] + \sum_{i=1}^k \frac{n_i - r}{\phi_i} \exp \left( \frac{y_{(r,n_i)} - \mu_i}{\phi_i} \right) \right] \frac{\partial \mu_i}{\partial \eta_{1i}} \frac{\partial^2 \eta_{1i}}{\partial \beta_j \partial \beta_l} \right)_{jl},$$

$$Q_{\theta\theta}^4 = \left( \left[ - \sum_{i=1}^k \sum_{s=1}^r \left[ \frac{1}{\phi_i} + \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i^2} \right) - \exp \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i} \right) \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i^2} \right) \right] + \sum_{i=1}^k \left( \frac{y_{(r,n_i)} - \mu_i}{\phi_i^2} \right) \exp \left( \frac{y_{(r,n_i)} - \mu_i}{\phi_i} \right) \right] \frac{\partial \phi_i}{\partial \eta_{2i}} \frac{\partial^2 \eta_{2i}}{\partial \theta_J \partial \theta_L} \right)_{JL},$$

Agora obtemos um expressão para a matriz de informação observada.

$$C = C(\beta, \theta) = \begin{pmatrix} C_{\beta\beta} & C_{\beta\theta} \\ C_{\theta\beta} & C_{\theta\theta} \end{pmatrix}$$

$$= \begin{pmatrix} \tilde{X}^T Q_{\beta\beta} \tilde{X} + Q_{\beta\beta}^1 + \tilde{X}^T Q_{\beta\beta}^{11} \tilde{X} & \tilde{X}^T Q_{\beta\theta} \tilde{S} \\ \tilde{S}^T Q_{\theta\beta} \tilde{X} & \tilde{S}^T Q_{\theta\theta} \tilde{S} + Q_{\theta\theta}^4 + \tilde{S}^T Q_{\theta\theta}^{44} \tilde{S} \end{pmatrix}.$$

Agora iremos mostrar a expressão para a matriz de informação de Fisher,

$$K = K(\beta, \theta) = \begin{pmatrix} K_{\beta\beta} & K_{\beta\theta} \\ K_{\theta\beta} & K_{\theta\theta} \end{pmatrix}.$$

Primeiramente considere  $Z_{(s,n_i)} = \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_i} \right)$ , onde  $s = 1, 2, \dots, n_i$  e  $i = 1, 2, \dots, k$ .

Dai temos que

$$K_{\beta\beta} = E(-U_{lj}) = \sum_{i=1}^k \left[ \sum_{s=1}^r \frac{1}{\phi_i^2} E(\exp(Z_{(s,n_i)}) + \left( \frac{n_i - r}{\phi_i^2} \right) E(\exp(Z_{(r,n_i)})) \right] \left[ \left( \frac{\partial \mu_i}{\partial \eta_{1i}} \right)^2 \left( \frac{\partial \eta_{1i}}{\partial \beta_l} \frac{\partial \eta_{1i}}{\partial \beta_j} \right) \right],$$

$$\begin{aligned}
K_{\theta\theta} = E(-U_{LJ}) &= \sum_{i=1}^k \left[ \sum_{s=1}^r \frac{1}{\phi_i^2} - \frac{2}{\phi_i^2} \sum_{s=1}^r E(Z_{(s,n_i)}) + \frac{2}{\phi_i^2} \sum_{s=1}^r E(Z_{(s,n_i)}^1 \exp(Z_{(s,n_i)})) + \right. \\
&+ \frac{1}{\phi_i^2} \sum_{s=1}^r E(Z_{(s,n_i)}^2 \exp(Z_{(s,n_i)})) + 2 \left( \frac{n_i - r}{\phi_i^2} \right) E(Z_{(r,n_i)}^1 \exp(Z_{(r,n_i)})) + \\
&\left. + \left( \frac{n_i - r}{\phi_i^2} \right) E(Z_{(r,n_i)}^2 \exp(Z_{(r,n_i)})) \right] \left[ \left( \frac{\partial \phi_i}{\partial \eta_{2i}} \right)^2 \frac{\partial \eta_{2i}}{\partial \theta_L} \frac{\partial \eta_{2i}}{\partial \theta_J} \right],
\end{aligned}$$

e

$$\begin{aligned}
K_{\theta\beta} = E(-U_{Lj}) &= \sum_{i=1}^k \left[ - \sum_{s=1}^r \frac{1}{\phi_i^2} + \sum_{s=1}^r \frac{1}{\phi_i^2} E(\exp(Z_{(s,n_i)})) + \frac{1}{\phi_i^2} \sum_{s=1}^r E(Z_{(s,n_i)}^1 \exp(Z_{(s,n_i)})) + \right. \\
&+ \left. \left( \frac{n_i - r}{\phi_i^2} \right) E(\exp(Z_{(r,n_i)})) + \left( \frac{n_i - r}{\phi_i^2} \right) E(Z_{(r,n_i)}^1 \exp(Z_{(r,n_i)})) \right] \times \\
&\times \left[ \frac{\partial \phi_i}{\partial \eta_{2i}} \frac{\partial \eta_{2i}}{\partial \theta_L} \frac{\partial \mu_i}{\partial \eta_{1i}} \frac{\partial \eta_{1i}}{\partial \beta_j} \right].
\end{aligned}$$

Resta-nos calcular as seguintes expressões

$$\begin{aligned}
&E(Z_{(s,n_i)}), E(\exp(Z_{(s,n_i)})), E(Z_{(r,n_i)}), E(Z_{(s,n_i)}^1 \exp(Z_{(s,n_i)})), E(Z_{(r,n_i)}^1 \exp(Z_{(r,n_i)})), \\
&E(Z_{(s,n_i)}^2 \exp(Z_{(s,n_i)})) \text{ e } E(Z_{(r,n_i)}^2 \exp(Z_{(r,n_i)})).
\end{aligned}$$

Que são calculadas da seguinte forma como em Balakrishnan and Chan[1]

$$\begin{aligned}
E(Z_{(s,n_i)}) &= \frac{n_i!}{(s-1)!(n_i-s)!} \sum_{a=0}^{n_i-s} (-1)^{a+1} \binom{n_i-s}{a} \frac{\gamma + \log(s+a)}{s+a}, \\
E(Z_{(s,n_i)}^b \exp(Z_{(s,n_i)})) &= \frac{n_i!}{(s-1)!(n_i-s)!} \sum_{a=0}^{n_i-s} (-1)^a \binom{n_i-s}{a} g_b(s+a),
\end{aligned}$$

onde  $b = 0, 1, 2$  e  $g_0(c) = \frac{1}{c^2}$ ,  $g_1(c) = \frac{1}{c^2}(1 - \gamma - \log(c))$  e  $g_3(c) = \frac{1}{c^2}[(\log(c))^2 - 2(1 - \gamma) \log(c) + \gamma^2 - 2\gamma + \frac{\pi^2}{6}]$

Os estimadores de máxima verossimilhança para os parâmetros  $\beta$  e  $\theta$  são obtidos resolvendo um sistema de equações não lineares  $U = 0$  e não ha uma forma fechada para tal resolução.

Portanto, utiliza-se algoritmos de otimização não-linear como o algoritmo de Newton ou quase-Newton, para encontrar os estimadores de máxima verossimilhança.

## 6 ANÁLISE DE DIAGNOSTICO

Neste capítulo vamos, avaliar a influência global e local para o Modelo de Regressão de Valor Extremo sob Censura do Tipo II.

### 6.1 Influência Global

Tal como referido na introdução, o primeiro passo envolvido na execução de análise de sensibilidade é se concentrar em influências globais de acordo com o caso de eliminação. O caso de eliminação é uma abordagem comum para estudar o efeito de abandonar o  $i$ -ésimo caso a partir do conjunto de dados. O caso de eliminação da  $i$ -ésima observação para o modelo é dado por

$$g_1(\mu_j) = \eta_{1j} = f_1(x_j^T; \beta) \quad \text{e} \quad g_2(\phi_j) = \eta_{2j} = f_1(z_j^T; \theta), \quad j = 1, 2, \dots, n; j \neq i$$

No que segue, um quantidade com índice “  $i$  ” indica os dados originais com a  $i$ -ésima observação excluída. Para o modelo anterior a função de log-verossimilhança é denotada por  $l_{(i)}(\zeta)$ .

Seja  $\widehat{\zeta}_{(i)} = (\widehat{\beta}_{(i)}^T, \widehat{\theta}_{(i)}^T)^T$  o estimador de máxima verossimilhança baseado em  $\widehat{\zeta}$ , obtido a partir de maximizar  $l_{(i)}(\zeta)$ . Para avaliar a influência da  $i$ -ésima observação no EMV.  $\widehat{\zeta} = (\widehat{\beta}^T, \widehat{\theta}^T)^T$ , comparamos a diferença entre  $\widehat{\zeta}_{(i)}$  e  $\widehat{\zeta}$ . Se a eliminação de uma observação influência seriamente uma estimava, maior atenção deve ser dada a essa observação em particular. Portanto, se  $\widehat{\zeta}_{(i)}$  esta longe se ser  $\widehat{\zeta}$ , então, neste caso, é considerado como uma observação influente. Uma medida inicial de influência global com base no teste de Wald é definida como a norma padronizada de  $\widehat{\zeta}_{(i)}$  que é conhecida como a distância generalizada de Cook e é dada pela seguinte

expressão

$$GD_i(\zeta) = (\widehat{\zeta}_{(i)} - \widehat{\zeta})^T [-C] (\widehat{\zeta}_{(i)} - \widehat{\zeta})$$

Outra alternativa é avaliar os valores  $GD_i(\beta)$  e  $GD_i(\theta)$ , que revelam o impacto da  $i$ -ésima observação sobre a previsão de  $\beta$  e  $\theta$ , respectivamente. Outra medida popular é a da diferença da verossimilhança entre  $\widehat{\zeta}_{(i)}$  e  $\widehat{\zeta}$

$$LD_i(\zeta) = 2[l(\widehat{\zeta}) - l(\widehat{\zeta}_{(i)})],$$

Além disso, também podemos calcular  $\widehat{\beta}_j - \widehat{\beta}_{(i)j}$  ( $j = 1, 2, \dots, p$ ), para calcular a diferença entre  $\widehat{\beta} - \widehat{\beta}_{(i)}$  e calcular  $\widehat{\theta}_j - \widehat{\theta}_{(i)j}$  ( $j = 1, 2, \dots, p$ ), para obter a diferença entre  $\widehat{\theta} - \widehat{\theta}_{(i)}$ . Outras medidas de influência global também são possíveis. Podemos observar o comportamento de uma estatística de teste sob um esquema de casos de eliminação, tais estatísticas podem ser, por exemplo o teste de Wald para as covariáveis explicativas.

## 6.2 Influência Local

O método de influência local é recomendado quando a preocupação está relacionada em investigar a sensibilidade do modelo em algumas perturbações menores no modelos ou dados. Seja  $w$  um vetor de perturbações  $t$ -dimensional. Então, denotamos a função de log-verossimilhança perturbada por  $l(\zeta|w)$ . Suponhamos que existe um vetor não perturbado, ou seja  $w_0$  tal que  $l(\zeta|w_0) = l(\zeta)$ . A influência de perturbações menores no estimador  $\widehat{\zeta}$  de máxima verossimilhança pode ser avaliado usando o deslocamento da log-verossimilhança

$$LD_w = 2(l(\widehat{\zeta}) - l(\widehat{\zeta}_w))$$

onde  $\widehat{\zeta}_w$  denota o estimado de máxima verossimilhança em  $l(\widehat{\zeta}|w)$ .

A ideia de [3] para avaliar a influência local é essencialmente analisar o comportamento local de  $LD_w$  em torno de  $w_0$  avaliando a curvatura da curva de

$LD_{w_0+ad}$ , onde  $a \in \mathbb{R}$  e  $d$  é uma direção normal unitária. Uma das medidas de particular interesse é a direção  $d_{\max}$  correspondente a maior curvatura  $C_{d_{\max}}$ . O índice da curva de  $d_{\max}$  pode indicar se essa observação tem uma influência considerável sobre  $LD_w$  sob perturbações menores. Além disso a curva de  $d_{\max}$  em relação aos valores de variáveis independentes pode ser útil para identificar padrões atípicos. [3] mostrou que a curvatura normal na direção de  $d$  é dada pela seguinte expressão

$$C_d(\zeta) = 2|d^T \Delta^T C^{-1} \Delta d|$$

onde  $\Delta = \frac{\partial^2 l(\zeta|w)}{\partial \zeta \partial w^T}$ , e ambos  $\Delta$  e  $C$  são avaliados em  $\hat{\zeta}$  e  $w_0$ .

Assim  $C_{d_{\max}}$  é o maior autovalor de  $B = -\Delta^T C^{-1} \Delta$  e  $d_{\max}$  é o autovetor normal unitário correspondente. O índice da curva de  $d_{\max}$  para a matriz  $B$  pode mostrar como perturbar o modelo ou dados para se obter grandes mudanças no estimador  $\zeta$ .

Portanto, se o interesse esta em calcular a influência para  $\beta$ , a curvatura normal na direção do vetor  $d$  é

$$C_{d;\beta}(\zeta) = 2|d^T \Delta^T (C^{-1} - C_{\theta\theta}) \Delta d|,$$

onde  $C_{\theta\theta} = \begin{pmatrix} 0 & 0 \\ 0 & C_{\theta\theta}^{-1} \end{pmatrix}$ , e  $d_{\max;\beta}$  aqui é o autovetor normal unitário correspondente ao maior autovalor de  $B_1 = -\Delta^T (C^{-1} - C_{\theta\theta}) \Delta$ .

O índice da curva do maior autovetor de  $B_1$  pode revelar a influência desta observação sobre  $\hat{\beta}$ .

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

$$C_{d;\theta}(\zeta) = 2|d^T \Delta^T (C^{-1} - C_{\beta\beta}) \Delta d|,$$

onde  $C_{\beta\beta} = \begin{pmatrix} C_{\beta\beta}^{-1} & 0 \\ 0 & 0 \end{pmatrix}$ , e  $d_{\max;\theta}$  aqui é o vetor normal unitário correspondente ao autovalor de  $B_2 = -\Delta^T(C^{-1} - C_{\beta\beta})\Delta$ . Portanto o índice da curva do maior autovetor de  $B_2$  revela a influência desta observação sobre  $\hat{\theta}$ .

Outro procedimento é a curvatura local total correspondente ao  $i$ -ésimo elemento, que segue, tomando  $d_i$  como sendo o vetor  $n \times 1$  de zeros com 1 na  $i$ -ésima coordenada. Assim a curvatura na direção de  $d_i$  assume a forma

$$C_i(\zeta) = 2|\Delta_i^T C^{-1} \Delta_i|,$$

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

$$C_{i;\beta}(\zeta) = 2|\Delta_i^T (C^{-1} - C_{\theta\theta}) \Delta_i|,$$

enquanto que se o interesse esta em  $\theta$ , temos

$$C_{i;\theta}(\zeta) = 2|\Delta_i^T (C^{-1} - C_{\beta\beta}) \Delta_i|.$$

[13] seguem indicando os casos de tem forma que  $C_1 \geq 2\bar{C}$  onde  $\bar{C} = \sum_{i=1}^n \frac{C_i}{n}$ .

### 6.3 Cálculo de Curvaturas

No que segue, calcularemos, cinco sistemas diferentes de perturbações para a matriz,

$$\Delta = \{\Delta_{gh}\}_{(p+q) \times (n*k)} = \left\{ \frac{\partial^2 l(\zeta|w)}{\partial \zeta_r \partial w_i} \right\} = \begin{pmatrix} \Delta\beta \\ \Delta\theta \end{pmatrix},$$

onde  $g = 1, 2, \dots, p + q$  e  $h = 1, 2, \dots, n * k$ .

Considerando o modelo definido em (2.2) e sua função de log-verossimilhança dada por (5.1). No que segue as quantidades distinguem-se pela adição de “ $\hat{\phantom{x}}$ ” são avaliadas em  $\hat{\zeta} = (\hat{\beta}^T, \hat{\theta}^T)^T$ .

### 6.3.1 Perturbação de casos Ponderados

A perturbação dos casos é feita através da definição de alguns pesos para cada observação na função de log-verossimilhança da seguinte forma

$$l(\zeta|w) = \sum_{i=1}^n w_{si} l_{si}(\zeta|w)$$

onde  $w$  é uma matriz  $n \times k$  de peso, com  $0 \leq w_{si} \leq 1$ , para todo  $s = 1, 2, \dots, n$ , e para todo  $i = 1, 2, \dots, k$  e  $w_0$  é a matriz nula de não perturbação.

As matrizes  $\Delta\beta$  e  $\Delta\theta$  são dadas por

$$\begin{aligned}\Delta\beta &= \tilde{X}^T M_1 \text{diag}(u_s), \\ \Delta\theta &= \tilde{S}^T M_2 \text{diag}(v_s),\end{aligned}$$

onde todas as matrizes foram anteriormente definidas.

### 6.3.2 Perturbação da variável resposta

Vamos considerar que cada  $y_{si}$  é perturbado como  $y_{siw} = y_{si} + w_{si} S_y^{si}$ , onde  $S_y^{si}$  é um escalar que pode ser por exemplo, em  $i$  constante igual ao desvio padrão estimado em  $y$ . Neste caso a perturbação da log-verossimilhança é dada por

$$\begin{aligned}l(\zeta|w) &= \sum_{i=1}^k \left[ \log \left( \frac{n_i!}{(n_i - r)!} \right) - r \log(\phi_i) + \sum_{s=1}^r \left( \frac{y_{(siw, n_i)} - \mu_i}{\phi_i} \right) \right. \\ &\quad \left. - \sum_{s=1}^r \exp \left( \frac{y_{(siw, n_i)} - \mu_i}{\phi_i} \right) - (n_i - r) \exp \left( \frac{y_{(riw, n_i)} - \mu_i}{\phi_i} \right) \right],\end{aligned}$$

onde,  $w_0$  é a matriz nula de não perturbação. As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem a forma

$$\begin{aligned}\Delta\beta &= \tilde{X}^T M_1 T_\beta, \\ \Delta\theta &= \tilde{S}^T M_2 T_\theta,\end{aligned}$$

onde,  $\alpha_\beta = (\exp(\frac{y_{(1,n_1)} + w_1 S_y^s - \mu_1}{\phi_1}) \frac{S_y^1}{\phi_1^2}, \dots, \exp(\frac{y_{(r,n_k)} + w_k S_y^r - \mu_k}{\phi_k}) \frac{S_y^r}{\phi_k^2})$ ,  $T_\beta = \text{diag}(\alpha_\beta)$ ,  $\alpha_\theta = ([-\frac{S_y^1}{\phi_1^2} + \exp(\frac{y_{(1,n_1)} + w_1 S_y^1 - \mu_1}{\phi_1}) (\frac{y_{(1,n_1)} + w_1 S_y^1 - \mu_1}{\phi_1^2}) \frac{S_y^1}{\phi_1} + \exp(\frac{y_{(1,n_1)} + w_1 S_y^1 - \mu_1}{\phi_1}) \frac{S_y^1}{\phi_1^2}], \dots, [-\frac{S_y^r}{\phi_k^2} + \exp(\frac{y_{(r,n_k)} + w_k S_y^r - \mu_k}{\phi_k}) (\frac{y_{(r,n_k)} + w_k S_y^r - \mu_k}{\phi_k^2}) \frac{S_y^r}{\phi_k} + \exp(\frac{y_{(r,n_k)} + w_k S_y^r - \mu_k}{\phi_k}) \frac{S_y^r}{\phi_k^2}])$ , e  $T_{\theta_s} = \text{diag}(\alpha_\theta)$ .

### 6.3.3 Perturbação da Variável Explicativa do Parâmetro de Localização ( $x_j^T$ )

Considerando agora uma perturbação aditiva em uma variável explicativa, em particular, seja  $x_j$ , dada por  $x_{siw} = x_{si} + w_{si} S_x$ , onde  $S_x$  é um escalar que pode por exemplo ser o desvio padrão estimado de  $x_j$ . Este esquema de perturbação leva à seguinte expressão para a função de log-verossimilhança

$$\begin{aligned}l(\zeta|w) &= \sum_{i=1}^k \left[ \log \left( \frac{n_i!}{(n_i - r)!} \right) - r \log(\phi_i) + \sum_{s=1}^r \left( \frac{y_{(s,n_i)} - \mu_{iw}}{\phi_i} \right) \right. \\ &\quad \left. - \sum_{s=1}^r \exp \left( \frac{y_{(s,n_i)} - \mu_{iw}}{\phi_i} \right) - (n_i - r) \exp \left( \frac{y_{(r,n_i)} - \mu_{iw}}{\phi_i} \right) \right],\end{aligned}$$

onde,  $\mu_{iw} = g_1^{-1}(\eta_{1i})$ ,  $\eta_{1i} = f_1(x_{iw}^T; \beta)$  com  $x_{iw} = (x_{i1w}, x_{i2w}, \dots, x_{ipw})$ . Conseqüentemente,  $w_0$  é a matriz nula de não perturbação. Seja  $\tilde{X}(w) = \frac{\partial \eta_{1i}}{\partial \beta}$ , e observe que  $\tilde{X}(w_0) = \tilde{X}$ . Portanto, seja  $\tilde{X}_w = \frac{\partial \eta_{1i}}{\partial w}$ . Mas geralmente, os parenteses ( $w$ ) querem dizer que as quantidades são avaliadas em  $\mu_{iw}$  em vez de  $\mu_i$ . As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem a forma

$$\begin{aligned}\Delta\beta &= \tilde{X}_w^T Q_{\beta\beta}(w) \tilde{X}(w) + Q_{\beta\beta}^1(w) + \tilde{X}_w^T Q_{\beta\beta}^{11}(w) \tilde{X}(w), \\ \Delta\theta &= \tilde{X}_w^T Q_{\beta\theta}(w) \tilde{S}.\end{aligned}$$

### 6.3.4 Perturbação da Variável Explicativa do Parâmetro de Escala ( $z_j^T$ )

Uma vez que estamos lidando com co-variáveis de dispersão, pode ser interessante verificar se uma observação em particular é influente para o parâmetro de escala do modelo. Conseqüentemente, uma perturbação aditiva em uma variável explicativa do parâmetro de precisão da regressão, ou seja,  $z_j$  dada por  $z_{ijw} = z_{ij} + w_i S_z$ , onde  $S_z$  é um escalar que pode ser por exemplo, o desvio padrão de  $z_j$ . Este esquema de perturbação leva a seguinte expressão para a função de log-verossimilhança

$$l(\zeta|w) = \sum_{i=1}^k \left[ \log \left( \frac{n_i!}{(n_i - r)!} \right) - r \log(\phi_{iw}) + \sum_{s=1}^r \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_{iw}} \right) - \sum_{s=1}^r \exp \left( \frac{y_{(s,n_i)} - \mu_i}{\phi_{iw}} \right) - (n_i - r) \exp \left( \frac{y_{(r,n_i)} - \mu_i}{\phi_{iw}} \right) \right],$$

onde,  $\phi_{iw} = g_2^{-1}(\eta_{2i})$ ,  $\eta_{2i} = f_2(z_{iw}^T; \theta)$  com  $z_{iw} = (z_{i1w}, z_{i2w}, \dots, z_{iqw})$ . Conseqüentemente,  $w_0$  é a matriz nula de não perturbação. Seja  $\tilde{S}(w) = \frac{\partial \eta_{2i}}{\partial \theta}$ , e observe que  $\tilde{S}(w_0) = \tilde{S}$ . Portanto, seja  $\tilde{S}_w = \frac{\partial \eta_{2i}}{\partial w}$ . Mas geralmente, os parenteses ( $w$ ) querem dizer que as quantidades são avaliadas em  $\phi_{iw}$  em vez de  $\phi_i$ . As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem a forma

$$\Delta\beta = \tilde{S}_w^T Q_{\beta\theta}(w) \tilde{X},$$

$$\Delta\theta = \tilde{S}_w^T Q_{\theta\theta}(w) \tilde{S}(w) + Q_{\theta\theta}^4(w) + \tilde{S}_w^T Q_{\theta\theta}^{44}(w) \tilde{S}(w).$$

### 6.3.5 Perturbações simultâneas das variáveis explicativas ( $x_j^T, z_i^T$ )<sup>T</sup>

Agora consideramos uma perturbação aditiva em uma variável explicativa particular do parâmetro de média da regressão e uma perturbação aditiva em

uma variável explicativa do parâmetro de precisão da regressão, ou seja,  $x_j$  e  $z_l$ , fazendo  $x_{ijw} = x_{ij} + w_i S_x$  e  $z_{ilw} = z_{il} + w_i S_z$ , respectivamente, onde  $S_x$  e  $S_z$  são fatores escalares que podem ser, por exemplo, os desvios padrão estimados de  $x_j$  e  $z_l$ , respectivamente. Este esquema de perturbação usa a seguinte expressão para a função de log-verossimilhança

$$l(\zeta|w) = \sum_{i=1}^k \left[ \log \left( \frac{n_i!}{(n_i - r)!} \right) - r \log(\phi_{iw}) + \sum_{s=1}^r \left( \frac{y_{(s,n_i)} - \mu_{iw}}{\phi_{iw}} \right) - \sum_{s=1}^r \exp \left( \frac{y_{(s,n_i)} - \mu_{iw}}{\phi_{iw}} \right) - (n_i - r) \exp \left( \frac{y_{(r,n_i)} - \mu_{iw}}{\phi_{iw}} \right) \right],$$

onde,  $\mu_{iw} = g_1^{-1}(\eta_{1iw})$ ,  $\eta_{1iw} = f_1(x_{iw}^T; \beta)$ , com  $x_{iw} = (x_{i1w}, x_{i2w}, \dots, x_{ipw})^T$  e  $\phi_{iw} = g_2^{-1}(\eta_{2iw})$ ,  $\eta_{2iw} = f_2(z_{iw}^T; \theta)$ , com  $z_{iw} = (z_{i1w}, z_{i2w}, \dots, z_{iqw})^T$ . Aqui  $w_0$  é a matriz de não perturbação. Tudo isso, usando que os paranteses ( $w$ ) significam que as quantidades são avaliadas em  $\mu_{iw}$  e  $\phi_{iw}$  em vez de  $\mu_i$  e  $\phi_i$ . As matrizes  $\Delta\beta$  e  $\Delta\theta$  assumem a forma:

$$\Delta\beta = \tilde{X}_w^T Q_{\beta\theta}(w) \tilde{S}(w) + \tilde{X}_w^T Q_{\beta\beta}(w) \tilde{X}(w) + Q_{\beta\beta}^1(w) + \tilde{X}_w^T Q_{\beta\beta}^{11}(w) \tilde{X}(w),$$

$$\Delta\theta = \tilde{S}_w^T Q_{\theta\theta}(w) \tilde{S}(w) + Q_{\theta\theta}^4(w) + \tilde{S}_w^T Q_{\theta\theta}^{44}(w) \tilde{S}(w) + \tilde{S}_w^T Q_{\beta\theta}(w) \tilde{X}(w).$$

## 6.4 Alavancagem Generalizada

A partir desse momento calcularemos a alavancagem generalizada proposta por Wei [14], que é definido como

$$GL(\zeta) = \frac{\partial \hat{y}}{\partial y^T},$$

onde  $\zeta$  é uma  $r$ -vetor tal que  $E(y) = \mu(\zeta)$  e  $\hat{\zeta}$  é um estimador de  $\zeta$ , com  $\hat{y} = \mu(\hat{\zeta})$ . Aqui, o  $(i, l)$  elemento de  $GL(\zeta)$ , ou seja, a alavancagem generalizada do estimador

$\hat{\zeta}$  pelo  $(i, l)$ , é o valor da taxa de variação instantânea no  $i$ -ésimo valor previsto em relação ao  $l$ -ésimo valor resposta. Como observamos pela reparametrização e observações  $GL_{i,j}$  são grandes pontos de alavancagem. Wei mostrou que o efeito da alavanca generalizado é obtido através de

$$GL(\zeta) = D_{\zeta}(-C)C_{\zeta y},$$

em  $\zeta = \hat{\zeta}$ , onde  $D_{\zeta} = \frac{\partial \mu}{\partial \zeta}$  e  $C_{\zeta y} = \frac{\partial^2 l(\zeta)}{\partial \zeta \partial y}$ . De acordo com o modelo definido em (2.2), temos

$$D_{\zeta} = \begin{pmatrix} M_1 \tilde{X} & 0 \end{pmatrix} \quad \text{e} \quad C_{\zeta y} = \begin{pmatrix} \tilde{X} M_1 T_{\beta} \\ \tilde{S} M_2 T_{\theta} \end{pmatrix}$$

## 7 APLICAÇÃO

Neste Capítulo faremos uma aplicação a dados reais, para isto foi utilizado o *software* R ([www.r-project.org](http://www.r-project.org)) na sua versão 3.2.2.

O estudo foi baseado no banco de dados de tempo de falha apresentado no trabalho de Mc Cool[8]. Ele nos deu os tempos de falha para amostras de aço endurecido num ensaio de fadiga constante de um rolamento, 10 observações independentes foram tomadas em cada um dos 4 valores de tensão de contato ilustrado na tabela 7.1. A Engenharia sugere que pelo nível de tensão  $s$ , o tempo de falha deve ter aproximadamente uma distribuição Weibull com um parâmetro de escala  $\alpha$  relacionado com  $s$  pela lei de potência dada por  $\alpha = cs^p$ . Reparametrizando obtemos que nosso problema pode ser tratado como um problema que segue aproximadamente o Modelo de Regressão de Valor Extremo onde  $\mu(x) = \log(\alpha)$ ,  $x = \log(s)$ ,  $v_0 = \log(c)$  e  $v_1 = p$ , com um parâmetro  $\phi$  independente de  $s$ .

Tensão (psi <sup>2</sup> +10 <sup>6</sup> )	Tempo de falha (ordenados)									
0.87	1.67	2.20	2.51	3.00	3.90	4.70	7.53	14.70	27.80	37.40
0.99	0.80	1.00	1.37	2.25	2.95	3.70	6.07	6.65	7.05	7.37
1.09	0.012	0.18	0.20	0.24	0.26	0.32	0.32	0.42	0.44	0.08
1.18	0.073	0.098	0.117	0.135	0.175	0.262	0.270	0.350	0.386	0.456

Figura 7.1: Dados

No modelo estudado nós consideramos como função de ligação a função identidade que satisfaz nossa hipótese, consideramos ainda para o nosso problema tomar uma censura de 30% da amostra e ainda segunda Mc Cool [8] temos o seguinte estimador de máxima verossimilhança

$$\hat{\zeta} = (\beta, \theta) = (v_0, v_1, \sigma) = (0.8394, -12.5250, 0, 9309)$$

Considerando os dados anteriores analizaremos a influência local. A figura 7.2 apresenta os gráficos de  $d_{max}$  e  $C_i$  que correspondem a influência local e influência local total, respectivamente, sob o regime de perturbação dos casos ponderados.

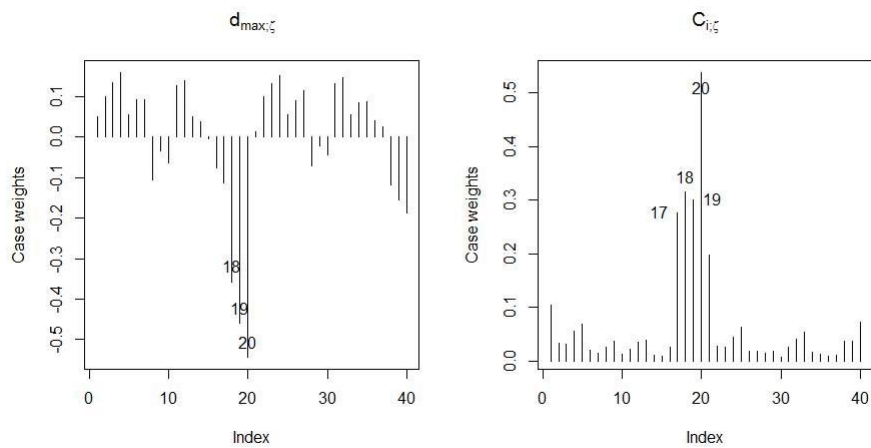


Figura 7.2: Perturbação dos Casos Ponderados

Com base na figura 7.2 no gráfico de  $d_{max}$  observamos que as observações 18, 19 e 20 exercem grande influencia sobre o vetor de parametros  $\zeta$ . Ainda na figura 7.2 mas agora no gráfico de  $C_i$  as observações 18, 19 e 20 aparecem como influêntes. No entanto com base no método de influência local total detectamos também como influênte a observação 17 que não foi detectada pelo método de influência local.

Passando agora para o próximo caso trataremos da perturbação da Variável Resposta.

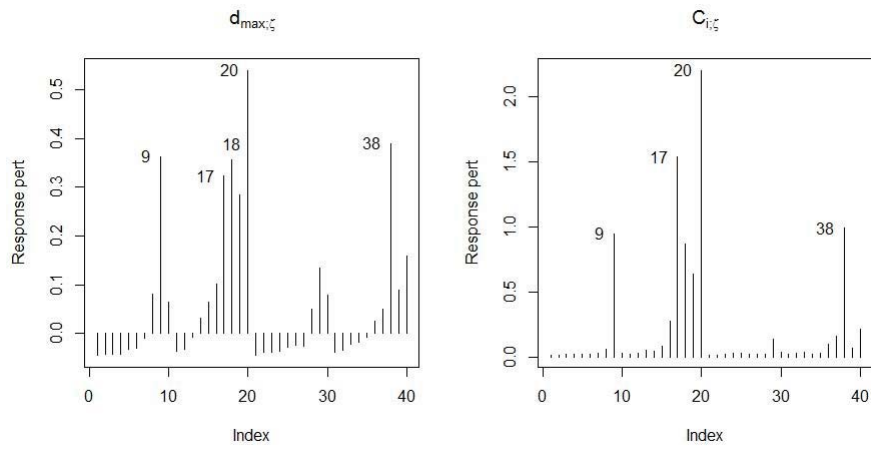


Figura 7.3: Perturbação da Variável Resposta

Com base na figura 7.3 no gráfico de  $d_{max}$  observamos que as observações 9, 17, 18, 20 e 38 exercem grande influência sobre o vetor de parâmetros  $\zeta$ . Ainda na figura 7.3 mas agora no gráfico de  $C_i$  as observações 9, 17, 20 e 38 aparecem como influêntes.

Por fim trataremos da perturbação da Variável Explicativa do Parâmetro de Localização.

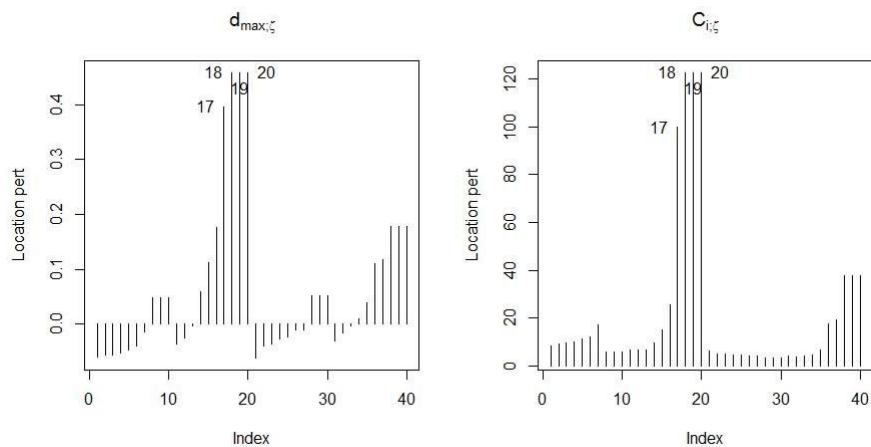


Figura 7.4: Perturbação da Variável Explicativa de Localização

Com base na figura 7.4 nos gráficos de  $d_{max}$  e  $C_i$  observamos que as observações 18, 19 e 20 exercem grande influência sobre o vetor de parâmetros  $\zeta$ .

## Referências Bibliográficas

- [1] BALAKRISHNAN, N., AND CHAN, P. Order statistics from extreme value distribution, i: Tables of means, variances and covariances. *Communications in Statistics-Simulation and Computation* 21, 4 (1992), 1199–1217.
- [2] BARRETO-SOUZA, W., AND VASCONCELLOS, K. L. Bias and skewness in a general extreme-value regression model. *Computational statistics & data analysis* 55, 3 (2011), 1379–1393.
- [3] COOK, R. D. Assessment of local influence. *Journal of the Royal Statistical Society. Series B (Methodological)* (1986), 133–169.
- [4] ESPINHEIRA, P. L., FERRARI, S. L., AND CRIBARI-NETO, F. On beta regression residuals. *Journal of Applied Statistics* 35, 4 (2008), 407–419.
- [5] LAWRENCE, A. J. Regression transformation diagnostics using local influence. *Journal of the American Statistical Association* 83, 404 (1988), 1067–1072.
- [6] LEMONTE, A. J., AND PATRIOTA, A. G. Influence diagnostics in birnbaum–saunders nonlinear regression models. *Journal of Applied Statistics* 38, 5 (2011), 871–884.
- [7] LESAFFRE, E., AND VERBEKE, G. Local influence in linear mixed models. *Biometrics* (1998), 570–582.
- [8] MCCOOL, J. I. Confidence limits for weibull regression with censored data. *IEEE Trans. Reliab* 29 (1980), 145–15.
- [9] PARK, C., AND LEE, S. B. Parameter estimation from censored samples using the expectation-maximization algorithm. *arXiv preprint arXiv:1203.3880* (2012).

- [10] PAULA, G. A. Assessing local influence in restricted regression models. *Computational statistics & data analysis* 16, 1 (1993), 63–79.
- [11] ROCHA, A. V., AND SIMAS, A. B. Influence diagnostics in a general class of beta regression models. *TEST* 20, 1 (2011), 95–119.
- [12] THOMAS, W., AND COOK, R. D. Assessing influence on predictions from generalized linear models. *Technometrics* 32, 1 (1990), 59–65.
- [13] VERBEKE, G., AND MOLENBERGHS, G. *Linear mixed models for longitudinal data*. Springer Series in Statistics. Springer-Verlag, New York, 2000.
- [14] WEI, B.-C., HU, Y.-Q., AND FUNG, W.-K. Generalized leverage and its applications. *Scandinavian Journal of statistics* 25, 1 (1998), 25–37.