

---

**Diagnóstico de Influência Local para a  
obtenção de dados mascarados influentes em  
modelos de regressão com erros nas variáveis  
e propriedades assintóticas do modelo de  
calibração ultraestrutural**

*Juan Pablo Mamani Bustamante*

---



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA INTERINSTITUCIONAL DE  
PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

**JUAN PABLO MAMANI BUSTAMANTE**

DIAGNÓSTICO DE INFLUÊNCIA LOCAL PARA A OBTENÇÃO DE  
DADOS MASCARADOS INFLUENTES EM MODELOS DE REGRESSÃO  
COM ERROS NAS VARIÁVEIS E PROPRIEDADES ASSINTÓTICAS DO  
MODELO DE CALIBRAÇÃO ULTRAESTRUTURAL

Tese apresentada ao Departamento de Estatística - Des/UFSCar e ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Mestre ou Doutor em Estatística - Programa Interinstitucional de Pós-Graduação em Estatística UFSCar-USP.

Orientadora: Profa. Dra. Reiko Aoki

**São Carlos**  
**Junho de 2020**



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA INTERINSTITUCIONAL DE  
PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

**JUAN PABLO MAMANI BUSTAMANTE**

LOCAL INFLUENCE DIAGNOSTICS FOR DETECTING MASKED  
INFLUENTIAL OBSERVATIONS IN MEASUREMENT ERROR MODELS  
AND ASYMPTOTIC PROPERTIES OF THE ULTRASTRUCTURAL  
CALIBRATION MODEL

Doctoral dissertation submitted to the Department of Statistics - Des/UFS-Car and to the Institute of Mathematics and Computer Sciences - ICMC-USP, in partial fulfillment of the requirements for PhD degree Statistics - Interinstitutional Program Graduation in Statistics UFSCar-USP.

Advisor: Prof. Dr. Reiko Aoki

**São Carlos**  
**June 2020**



Mamani Bustamante, Juan Pablo

Diagnóstico de Influência Local para a obtenção de dados mascarados influentes em modelos de regressão com erros nas variáveis e propriedades assintóticas do modelo de calibração ultraestrutural / Juan Pablo Mamani Bustamante. -- 2020.  
181 f. : 30 cm.

Tese (doutorado)-Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador: Reiko Aoki

Banca examinadora: Reiko Aoki, (ICMC-USP), Gilberto Alvarenga Paula (IME-USP), Filidor E. Labra (Unicamp), Silvia Nagib Elian (IME-USP), Cibele Maria Russo Novelli, (ICMC-USP)

Bibliografia

1. teoria assintótica. 2. procura passo a frente. 3. influência local. I. Orientador. II. Universidade Federal de São Carlos. III. Título.

Ficha catalográfica elaborada pelo Programa de Geração Automática da Secretaria Geral de Informática (SIn).

DADOS FORNECIDOS PELO(A) AUTOR(A)

Bibliotecário(a) Responsável: Ronildo Santos Prado – CRB/8 7325



**UNIVERSIDADE FEDERAL DE SÃO CARLOS**


Centro de Ciências Exatas e de Tecnologia  
Programa Interinstitucional de Pós-Graduação em Estatística

---

**Folha de Aprovação**

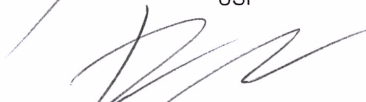
---

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Juan Pablo Mamani Bustamante, realizada em 08/05/2020:



---

Prôfa. Dra. Reiko Aoki  
USP




---

Prof. Dr. Gilberto Alvarenga Paula  
IME-USP




---

Prof. Dr. Filidor Edilson Vilca Labra  
UNICAMP



---


Prôfa. Dra. Silvia Nagib Elian  
IME-USP



---

Prôfa. Dra. Cibele Maria Russo Novelli  
USP

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Gilberto Alvarenga Paula, Filidor Edilson Vilca Labra, Silvia Nagib Elian, Cibele Maria Russo Novelli e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ao) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.



---

Prôfa. Dra. Reiko Aoki

*À minha  
querida família.*



# AGRADECIMENTOS

---

---

Em primeiro lugar agradeço a Deus pela saúde.

À Professora Reiko Aoki, minha querida orientadora, por me entender e me ensinar muitas coisas de estatística, pelas orientações na elaboração deste trabalho e também pela paciência, apoio e amizade durante todo o doutorado. Faz quatro anos nunca pensei que começaria e terminaria o doutorado, por tal sacrifício feito, os intensos diálogos com minha orientadora.

Ao Professor Leão, que eu considero como meu Co-orientador e um dos melhores professores que já tive.

Ao Marcelo Hartmann pela amizade, pela ajuda na disciplina de inferência estatística no mestrado. Ter aquelas discussões de cada sábado de noite sobre inferência estatística na Toca do Castor.

Ao Rogério Alves Santana meu colega e amigo que me ajudou em muitas das disciplinas e pela paciência em resolver meus problemas.

Ao Marcos Antônio Pereira, um amigo e colega nas muitas matérias que nós levamos.

Ao Rafael Paixão meu amigo, por resolver muitas questões de estatística, quando eu tinha dúvidas sempre encontrei nele uma ajuda.

Ao Marco Pollo um amigo de cerveja.

Ao Grober amigo Languenho da UNSAAC, que tivemos muitas histórias juntos.

À Veronica Maita que me apoiou com uma passagem por quando estava doente.

Ao Oscar Pacha meu amigo del Instituto Gregoria Santos.

À Elizabeth Pacompia por me falar em aqueles dias de solidão que estava trabalhando.

Ao Thales Matsumoto meu eterno amigo, meu quase irmão por me salvar em muitas disciplinas.

À Marina e Vinícius (cabeça) que foram amigos de estudos de estudos e melhores amigos, por muitos momentos de risadas que tivemos.

Ao Efrain Candia um amigo incondicional quase meu irmão que não tive, por meio de ti amigo fui a trabalhar em Puerto Maldonado e logo vim estudar no Brasil.

À Sayda Chahuasoncco quase uma mãe para mim.

Ao Luciano um grande amigo pois se não fosse por ele não teria feito o mestrado por conseguinte nem o doutorado.

À Lorena e Miguel amigos que me ajudaram em aprovar na disciplina de Tópicos de Pesquisa.

Ao professor Vicente Cancho que me ajudou e me deu muitas dicas

Ao Alcides o irmão da igreja Presbiteriana com seu conselho e sabedoria me ajudou muito.

Às minhas amigas Yuri e Elizabeth as únicas amigas verdadeiras que se lembraram do meu aniversário, daqueles momentos em que nós comíamos pizza.

Ao Carlos Franklin amigo que me fez ficar bravo e rir muitas vezes.

Aos meus amigos Susan, Hélio, Fabiano Coelho, Fernando e José da sala de estudos, que compartilhei muitas coisas.

Ao Charlie e Fiorela meus colegas de quarto, por arrumar muitas coisas em casa.

Aos amigos Alex, Abraham e Cubas sinceros.

Agradeço à Ritha e a Rafael por sua compreensão aquele verão inesquecível.

Agradeço à Noemi por ter me levado a emergência, uma amiga de toda a vida.

À Maria Eugenia Baldassari, melhor amiga e irmã que me deu muitas caronas e sempre lembrarei da piada do sequestro.

À Mariana, amiga que tive muitas conversas agradáveis.

À Karina minha melhor amiga que me ajudou no momento em que mais necessitava, durante o doutorado.

Agradecimento a Cristel Ecaterin e a Amélia Fernandes as melhores amigas que me ensinaram a estudar no mestrado e animaram a fazer o doutorado, à turma da estatística, aos amigos da escola e a tantos outros que não mencionei aqui mas que me ajudaram de alguma forma, me incentivaram e acreditaram na possibilidade de ter esse sonho realizado.

À minha irmã Silvia Eugenia Mamani Bustamante.

Ao meu pai Feliciano Mamani Astete, à minha mãe Juana Bustamante Aragón por todo apoio.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

*“ Lo terrible es saber que te va a pasar,  
lo maravilloso, lo vital es no saber lo que te va a pasar.”  
(César Hildebrandt)*



# RESUMO

BUSTAMANTE, J. P. M. **Diagnóstico de Influência Local para a obtenção de dados mascarados influentes em modelos de regressão com erros nas variáveis e propriedades assintóticas do modelo de calibração ultraestrutural**. 2020. 180 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Testes de proficiência determinam a performance de laboratórios individuais em relação a testes específicos de medição e é usado para monitorar a confiabilidade das medições dos laboratórios. Considerando um modelo de regressão multivariado com erros nas variáveis e com réplicas, o modelo de calibração ultraestrutural com réplicas, iremos desenvolver as propriedades assintóticas associada a este modelo para poder testar a equivalência entre as medições obtidas por diferentes laboratórios em testes de proficiência. Após ajustado um modelo, a análise de diagnóstico é uma etapa importante na análise de um conjunto de dados. Cook (1986) introduziu um método geral de diagnóstico de influência para avaliar a influência local de pequenas perturbações no modelo estatístico, usando diferentes tipos de perturbações. Como complemento às técnicas de detecção de observações discrepantes, é proposto o método procura passo a frente, por Atkinson e Riani (2000), que é uma metodologia para detectar observações atípicas mascaradas. Propomos uma metodologia baseada na influência local de Cook (1986) e Poon e Poon (1999) com a procura passo a frente de Atkinson e Riani (2000) em modelos de regressão com erros de medição que será utilizada na obtenção de observações mascaradas influentes.

**Palavras-chave:** regressão, ultraestrutural, diagnóstico, influência local, curvatura normal, conformal, procura passo a frente, teoria assintótica.



# ABSTRACT

BUSTAMANTE, J. P. M. **Local Influence Diagnostics for detecting masked influential observations in measurement error models and asymptotic properties of the ultrastructural calibration model.** 2020. 180 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Proficiency testing determines the performance of individual laboratories with respect to specific measurement tests and it is used to monitor the reliability of laboratory measurements. Considering a multivariate measurement error model with replicates, the ultrastructural calibration model with replicates, we will develop the asymptotic properties associated with this model in order to test the equivalence between the measurements obtained by different laboratories in proficiency testing. After adjusting a model, the diagnostic analysis is an important step in the analysis of a data set. Cook (1986) introduced a general methodology to assess the local influence of minor perturbations in the statistical model, using different types of perturbations. As a complement to the techniques for detecting discrepant observations, it is proposed the forward search by Atkinson and Riani (2000), which is a methodology to detect atypical masked observations. We propose a methodology based on the local influence of Cook (1986) and Poon and Poon (1999) with the forward search of Atkinson and Riani (2000) considering the measurement error regression models to obtain masked influential observations.

**Keywords:** regression, ultrastructural, assessment, local influence, influence graph, normal curvature, conformal, forward search, asymptotic theory.



# LISTA DE ILUSTRAÇÕES

---

---

Figura 1 – $H_0 : \alpha_2 = \dots = \alpha_5 = 0$ e $\beta_2 = \dots = \beta_5 = 1$ - poder do teste com $\sigma_{ij}^a$ . . . . .	51
Figura 2 – $H_0 : \alpha_2 = \dots = \alpha_5 = 0$ e $\beta_2 = \dots = \beta_5 = 1$ - poder do teste com $\sigma_{ij}^b$ . . . . .	51
Figura 3 – $H_0 : \alpha_2 = \dots = \alpha_5 = 0$ e $\beta_2 = \dots = \beta_5 = 1$ - poder do teste. . . . .	52
Figura 4 – $H_0 : \alpha_2 = 0, \beta_2 = 1$ - poder do teste com $\sigma_{ij}^a$ . . . . .	53
Figura 5 – $H_0 : \alpha_2 = 0, \beta_2 = 1$ - poder do teste com $\sigma_{ij}^b$ . . . . .	53
Figura 6 – $H_0 : \alpha_2 = 0$ e $\beta_2 = 1$ - poder do teste com $\sigma_{ij}^d$ . . . . .	54
Figura 7 – Regiões de confiança para os laboratorios 2 a 5. . . . .	55
Figura 8 – Regiões de confiança para os laboratorios 6 a 8. . . . .	56
Figura 9 – Perturbação na variável explanatória: gráfico do valor absoluto do autovetor, $l_{max}$ , associado ao maior autovalor com o índice das observações. Líquido de bochecho A após 6 meses. . . . .	85
Figura 10 – Perturbação na variável explanatória: gráfico de ILPPF. Líquido de bochecho A após 6 meses. . . . .	85
Figura 11 – Perturbação na variável explanatória: gráfico de afastamento pela verossimilhança. Líquido de bochecho A após 6 meses. . . . .	87
Figura 12 – Perturbação na variável explanatória: boxplot da ILPPF. Líquido de bochecho A após 6 meses. . . . .	88
Figura 13 – Perturbação na variável explanatória: heatmap de ILPPF. Líquido de bochecho A após 6 meses. . . . .	89
Figura 14 – Ponderação de casos: gráfico do valor absoluto do autovetor, associado ao maior autovalor com o índice das observações. Líquido de controle após 3 meses. . . . .	90
Figura 15 – Ponderação de casos: boxplot de ILPPF. Líquido de bochecho controle após 3 meses. . . . .	90
Figura 16 – Ponderação de casos: gráfico de ILPPF. Líquido de bochecho controle após 3 meses. . . . .	91
Figura 17 – Ponderação de casos: gráfico da função de afastamento pela verossimilhança. Líquido de bochecho controle após 3 meses. . . . .	92
Figura 18 – Ponderação de casos: gráfico do valor absoluto do autovetor, $l_{max}$ , associado ao maior autovalor com o índice das observações. Líquido de controle após 6 meses. . . . .	93
Figura 19 – Ponderação de casos: gráfico de ILPPF. Líquido de bochecho controle após 6 meses. . . . .	94

Figura 20 – Ponderação de casos: boxplot de ILPPF. Líquido de bochecho controle após 6 meses. . . . .	95
Figura 21 – Ponderação de casos: gráfico da função de afastamento pela verossimilhança. Líquido de bochecho controle após 6 meses. . . . .	95
Figura 22 – Ponderação de casos: gráfico da função de afastamento pela verossimilhança. Líquido de bochecho controle após 6 meses. . . . .	96
Figura 23 – Perturbação na resposta: gráfico do valor absoluto do autovetor, associado ao maior autovalor com o índice das observações. Líquido de bochecho controle após 6 meses. . . . .	97
Figura 24 – Perturbação na resposta: gráfico de ILPPF. Líquido de bochecho controle após 6 meses. . . . .	98
Figura 25 – Perturbação na resposta: gráfico de afastamento pela verossimilhança. Líquido de bochecho controle após 6 meses. . . . .	100
Figura 26 – Perturbação na variância: gráfico do valor absoluto do autovetor, associado ao maior autovalor com o índice das observações. Líquido de bochecho B após 6 meses. . . . .	101
Figura 27 – Perturbação na variância: gráfico de ILPPF. Líquido de bochecho B após 6 meses. . . . .	101
Figura 28 – Perturbação na variância: boxplot da ILPPF. Líquido de bochecho B após 6 meses. . . . .	102
Figura 29 – Perturbação na variância: Heatmap de ILPPF. Líquido de bochecho B após 6 meses. . . . .	102
Figura 30 – Perturbação na variância: gráfico de afastamento pela verossimilhança. Líquido de bochecho B após 6 meses. . . . .	104
Figura 31 – Perturbação nas variâncias dos erros: gráfico do valor absoluto do autovetor correspondente ao maior autovalor, dados do líquido de bochecho. . . . .	107
Figura 32 – Perturbação nas variâncias dos erros: gráfico de ILPPF, dados do líquido de bochecho. . . . .	108
Figura 33 – Perturbação nas variâncias dos erros: Heat Map, dados de líquido de bochecho. . . . .	108
Figura 34 – Perturbação nas variâncias dos erros: gráfico de ILPPF, dados de líquido de bochecho. . . . .	109
Figura 35 – Perturbação nas variâncias dos erros: gráfico de $LD(\boldsymbol{\omega}(a))$ vs $a$ , dados do líquido de bochecho. . . . .	109
Figura 36 – Dados de escova de dente: gráfico de $m(q)$ para Ponderação de casos, Perturbação nas variáveis respostas e Perturbação nas variáveis explanatórias. Contribuição total ( $q = 0$ ). . . . .	114
Figura 37 – Dados de escova de dente: gráfico da CNCPPF. Ponderação de casos. Contribuição total ( $q = 0$ ). . . . .	114

Figura 38 – Dados de escova de dente: gráfico da CNCPPF. Perturbação nas variáveis respostas. Contribuição total ( $q = 0$ ). . . . .	115
Figura 39 – Dados de escova de dente: gráfico da CNCPPF. Perturbação nas variáveis explanatórias. Contribuição total ( $q = 0$ ). . . . .	116
Figura 40 – Dados de escova de dente: gráfico de $LD(\boldsymbol{\omega}(a))$ versus $a$ com $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$ para (a) ponderação de casos, (b) perturbação nas variáveis respostas e (c) perturbação nas variáveis explanatórias. . . . .	116
Figura 41 – Dados de escova de dente: gráfico $\mathbf{m}(q)$ para (a) ponderação de casos, (b) perturbação nas variáveis respostas e (c) perturbação nas variáveis explanatórias. Contribuição do autovetor associado ao maior autovalor. . . . .	118
Figura 42 – Dados de escova de dente: gráfico da CNCPPF para ponderação de casos. Contribuição do autovetor associado ao maior autovalor. . . . .	118
Figura 43 – Dados de escova de dente: gráfico da CNCPPF para perturbação na variável resposta. Contribuição do autovetor associado ao maior autovalor. . . . .	119
Figura 44 – Dados de escova de dente: gráfico da CNCPPF para perturbação na variável explanatória. Contribuição do autovetor associado ao maior autovalor. . . . .	120
Figura 45 – Dados de escova de dente: ponderação de casos - gráfico de $LD(\boldsymbol{\omega}(a))$ versus $a$ com $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$ . . . . .	121
Figura 46 – Dados de líquido de bochecho: gráfico da CNCPPF para o esquema de perturbação da ponderação de casos. Contribuição agregada dos dois maiores autovalores e dos autovetores associados e contribuição do autovetor associado ao maior autovalor. . . . .	123
Figura 47 – Dados de líquido de bochecho: ponderação de casos - gráfico de $LD(\boldsymbol{\omega}(a))$ versus $a$ com $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$ . . . . .	124
Figura 48 – Dados de líquido de bochecho: gráfico da CNCPPF para esquema de perturbação nas variáveis explanatórias. Contribuição agregada dos dois maiores autovalores e dos autovetores associados. . . . .	124
Figura 49 – Dados de líquido de bochecho: gráfico da CNCPPF para o esquema de perturbação nas variáveis respostas. Contribuição do autovetor associado ao maior autovalor. . . . .	125
Figura 50 – Gráfico do valor absoluto do autovetor associado ao maior autovalor na metodologia CNCPPF para os dados de bochecho. . . . .	126
Figura 51 – Dados de líquido de bochecho: perturbação na variável resposta - gráfico de $LD(\boldsymbol{\omega}(a))$ versus $a$ com $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$ . . . . .	126
Figura 52 – Dados de líquido de bochecho: gráfico da CNCPPF para esquema de perturbação na variância. Contribuição do autovetor associado ao maior autovalor. . . . .	127
Figura 53 – Dados de líquido de bochecho: perturbação na variância - gráfico de $LD(\boldsymbol{\omega}(a))$ versus $a$ com $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$ . . . . .	130
Figura 54 – Gráfico de dispersão dos dados simulados. . . . .	131

Figura 55 – Perturbação na variância. . . . .	131
Figura 56 – Gráfico de dispersão dos dados simulados alterados. . . . .	132
Figura 57 – Perturbação na variância: gráfico de $m(q)$ com o índice das observações. Contribuição do autovetor associado ao maior autovalor. Dados alterados. . .	133
Figura 58 – Perturbação na variância: gráfico da CNCPPF. Dados alterados. . . . .	133
Figura 59 – Perturbação na variância: mudanças relativas nas estimativas dos parâmetros na direção do autovetor associado ao maior autovalor. Dados alterados. . . .	134
Figura 60 – Perturbação na variância: mudança relativa nas estimativas dos parâmetros na direção de (a) observação 7, (b) observação 12 e (c) observação 24. Dados alterados. . . . .	134
Figura 61 – Exemplo de forma geral da elipse: $x^2 + 3y^2 + 2xy + 12x - 5y + 50 = 0$ . . . .	173

# LISTA DE TABELAS

Tabela 1 – $H_0 : \alpha_2 = \dots = \alpha_5 = 0, \beta_2 = \dots = \beta_5 = 1$ - tamanho empírico. . . . .	50
Tabela 2 – $H_0 : \alpha_i = 0, \beta_i = 1, i = 2, \dots, 5$ - tamanho empírico. . . . .	50
Tabela 3 – Estimativas de máxima verossimilhança do vetor de parâmetros $\theta$ . . . . .	56
Tabela 4 – $Q_W$ para os 8 testes com os respectivos p-valores entre parênteses. . . . .	57
Tabela 5 – Subconjuntos, começando com $m_0 = 2$ , onde $n = 4$ , temos um total de 6 subconjuntos $(C_{m=m_0}^n)$ . $l_{m_0}^* = \min(l_{med_1}, l_{med_2}, l_{med_3}, l_{med_4}, l_{med_5}, l_{med_6})$ e o autovetor correspondente. . . . .	69
Tabela 6 – Subconjuntos, começando com $m = m_0 + 1 = 3$ , onde $n = 4$ , temos um total de 4 subconjuntos $(C_{m=m_0+1}^n)$ . $l_{m_0+1}^* = \min(l_{med_1}, l_{med_2}, l_{med_3}, l_{med_4})$ . . . . .	69
Tabela 7 – Neste caso, $m = n$ e temos apenas um conjunto que é o próprio conjunto de dados $(C_{m=n}^n)$ . . . . .	70
Tabela 8 – Perturbação na variável explanatória: mudança relativa na estimativa dos parâmetros. Líquido de bochecho A após 6 meses. . . . .	88
Tabela 9 – Ponderação de casos: mudança relativa na estimativa dos parâmetros. Líquido de bochecho controle após 3 meses. . . . .	92
Tabela 10 – Ponderação de casos: mudança relativa na estimativa dos parâmetros. Líquido de bochecho controle após 6 meses. . . . .	96
Tabela 11 – Perturbação na resposta: mudança relativa na estimativa dos parâmetros ao tirar algumas observações. Líquido de bochecho controle após 6 meses. . . . .	99
Tabela 12 – Perturbação na variância: mudança relativa na estimativa dos parâmetros. Líquido de bochecho B após 6 meses. . . . .	103
Tabela 13 – Mudança relativa nas estimativas dos parâmetros dos dados de líquido de bochecho . . . . .	110
Tabela 14 – Dados de escova de dente: parâmetros estimados e mudança relativa. . . . .	121
Tabela 15 – EMV e mudanças relativas nas estimativas dos parâmetros. . . . .	128
Tabela 16 – EMV e mudanças relativas nas estimativas dos parâmetros. . . . .	129
Tabela 17 – Teste de hipóteses: $H_0 : \beta_{1_1} = \beta_{1_2}$ . . . . .	129
Tabela 18 – Teste de hipóteses: $H_0 : \beta_{2_2} = \beta_{2_3}$ . . . . .	129
Tabela 19 – Perturbação na variância: mudança nas estimativas dos parâmetros dos dados simulados. . . . .	132
Tabela 26 – Variâncias do erro de medição $(\sigma_{ij}^2)$ para o $i$ -ésimo laboratório no $j$ -ésimo valor de rotação do motor. . . . .	159

Tabela 27 – Variância do verdadeiro valor da potência do motor ( $\sigma_{x_j}^2$ ) no  $j$ -ésimo ponto de rotação do motor. . . . . 159

# SUMÁRIO

---

---

1	INTRODUÇÃO . . . . .	23
1.1	Dados de escovas de dentes . . . . .	32
1.2	Dados do líquidos de bochecho . . . . .	34
1.3	Dados da Potência do Motor . . . . .	35
2	PROPRIEDADES ASSINTÓTICAS DO MODELO DE CALIBRAÇÃO ULTRAESTRUTURAL COM RÉPLICAS . . . . .	39
2.1	O modelo . . . . .	39
2.2	Propriedades assintóticas . . . . .	40
2.3	Equivalência entre laboratórios participantes . . . . .	46
2.3.1	<i>Teste de hipótese</i> . . . . .	46
2.3.2	<i>Região de Confiança Conjunta</i> . . . . .	48
2.4	Estudos de simulação . . . . .	49
2.4.1	<i>Simulação do tamanho dos testes assintóticos</i> . . . . .	50
2.4.2	<i>Simulação do poder dos testes assintóticos</i> . . . . .	50
2.4.2.1	$H_0 : \alpha_2 = \dots = \alpha_5 = 0$ e $\beta_2 = \dots = \beta_5 = 1$ . . . . .	50
2.4.2.2	$H_0 : \alpha_2 = 0$ e $\beta_2 = 1$ . . . . .	54
2.5	Aplicação . . . . .	55
3	MÉTODOS DE DIAGNÓSTICO . . . . .	59
3.1	Influência local de Cook (1986) . . . . .	59
3.2	Curvatura Normal Conformal de Poon e Poon (1999) . . . . .	60
3.3	Procura passo a frente de Atkinson e Riani (2000) . . . . .	63
4	INFLUÊNCIA LOCAL COM PROCURA PASSO A FRENTE E CURVATURA NORMAL CONFORMAL COM PROCURA PASSO A FRENTE . . . . .	67
4.1	Influência local com procura passo a frente: ILPPF . . . . .	67
4.2	Curvatura Normal Conformal com Procura Passo a Frente: CNCPPF . . . . .	71
5	ILPPF NO MODELO DE REGRESSÃO COM ERROS NAS VARIÁVEIS COM INTERCEPTO NULO . . . . .	73
5.1	Estimação dos parâmetros . . . . .	77
5.1.1	<i>Algoritmo EM</i> . . . . .	78

<b>5.2</b>	<b>Esquemas de perturbações</b>	<b>79</b>
<b>5.2.1</b>	<b><i>Ponderação de casos</i></b>	<b>80</b>
<b>5.2.2</b>	<b><i>Perturbação na covariável</i></b>	<b>81</b>
<b>5.2.3</b>	<b><i>Perturbação na variável resposta</i></b>	<b>81</b>
<b>5.2.4</b>	<b><i>Perturbação nas variâncias dos erros</i></b>	<b>82</b>
<b>5.2.5</b>	<b><i>Aplicação</i></b>	<b>83</b>
5.2.5.1	<i>Dados do líquido de bochecho (antisséptico bucal)</i>	83
5.2.5.2	<i>Líquido de bochecho A após 6 meses do início do estudo - perturbação na variável explanatória.</i>	84
5.2.5.3	<i>Líquido de bochecho controle após 3 meses do início do estudo - Ponderação de casos</i>	89
5.2.5.4	<i>Líquido de bochecho controle após 6 meses do início do estudo - Ponderação de casos</i>	93
5.2.5.5	<i>Líquido de bochecho controle após 6 meses do início do estudo - Perturbação na variável resposta</i>	97
5.2.5.6	<i>Líquido de bochecho B após 6 meses do início do estudo - Perturbação na variância</i>	100
<b>6</b>	<b>ILPPF NO MODELO DE REGRESSÃO MULTIVARIADO COM ERROS NAS VARIÁVEIS COM INTERCEPTO NULO</b>	<b>105</b>
<b>6.1</b>	<b>Dados do líquido de bochecho</b>	<b>105</b>
<b>6.1.1</b>	<b><i>Aplicação</i></b>	<b>107</b>
<b>7</b>	<b>CURVATURA NORMAL CONFORMAL COM PROCURA PASSO A FRENTE NO MODELO DE REGRESSÃO MULTIVARIADO COM ERROS NAS VARIÁVEIS COM INTERCEPTO NULO</b>	<b>111</b>
<b>7.1</b>	<b>Dados de escovas de dentes</b>	<b>111</b>
<b>7.1.1</b>	<b><i>Aplicação</i></b>	<b>113</b>
<b>7.2</b>	<b>Dados do líquido de bochecho</b>	<b>122</b>
<b>7.2.1</b>	<b><i>Aplicação</i></b>	<b>122</b>
<b>7.3</b>	<b>Dados simulados</b>	<b>130</b>
<b>8</b>	<b>CONCLUSÕES</b>	<b>137</b>
<b>9</b>	<b>ESTUDOS FUTUROS</b>	<b>139</b>
	<b>REFERÊNCIAS</b>	<b>141</b>
	<b>APÊNDICE A CONJUNTO DE DADOS</b>	<b>151</b>
<b>A.1</b>	<b>Dados de Escovas de dentes</b>	<b>151</b>
<b>A.2</b>	<b>Dados de Líquidos de bochecho</b>	<b>153</b>

A.3	Dados da Potência do Motor . . . . .	156
APÊNDICE B	CONVERGÊNCIA DOS ELEMENTOS DA MATRIZ DE INFORMAÇÃO . . . . .	161
APÊNDICE C	OUTROS . . . . .	171
C.1	Identificabilidade . . . . .	171
C.2	Exemplo da forma paramétrica da elipse . . . . .	171
C.3	Base ortonormal . . . . .	173
C.4	Modelo de calibração Ultraestrutural com réplicas . . . . .	174
C.4.1	<i>Algoritmo EM</i> . . . . .	174
C.4.2	<i>Função score e matriz de informação observada.</i> . . . . .	175
C.5	Modelo de regressão multivariado com erros nas variáveis com intercepto nulo . . . . .	176
C.5.1	<i>Matriz de informação observada</i> . . . . .	176
C.5.2	<i>Matriz delta para os dados de líquidos de bochecho</i> . . . . .	179



---

## INTRODUÇÃO

---

No cotidiano existem situações em que a covariável (variável explicativa) é observada com erro de medição. Os **modelos de regressão com erros nas variáveis** são modelos aplicáveis em muitos problemas experimentais práticos, quando a covariável é medida com erro. [Sprent \(1990\)](#) descreve o desenvolvimento histórico e uma ampla bibliografia pode ser encontrada em [Moran \(1971\)](#), [Fuller \(1987\)](#) e [Cheng e Ness \(1999\)](#).

O modelo linear simples com erros de medição é definido da seguinte forma: dado um conjunto finito de pares  $(\xi_1, \eta_1), \dots, (\xi_n, \eta_n)$ , as variáveis  $\xi$  e  $\eta$  são relacionadas como:

$$\eta_i = \alpha + \beta \xi_i, \quad i = 1, \dots, n, \quad (1.1)$$

porém, os verdadeiros valores,  $\xi_i$  e  $\eta_i$ , não são observados diretamente. Os pares de valores observados  $(X_i, Y_i)$ , estão sujeitos aos pares de erros de medição  $(\delta_i, \varepsilon_i)$ , respectivamente, ou seja,

$$X_i = \xi_i + \delta_i \quad e \quad Y_i = \eta_i + \varepsilon_i, \quad i = 1, \dots, n. \quad (1.2)$$

É assumido que  $\delta_1, \dots, \delta_n$  e  $\varepsilon_1, \dots, \varepsilon_n$  possuem variâncias finitas  $\sigma_\delta^2$  e  $\sigma_\varepsilon^2$ , são não correlacionadas e com média zero, ou seja,

$$\begin{aligned} E(\delta_i) = E(\varepsilon_i) = 0, \quad Var(\delta_i) = \sigma_\delta^2 > 0, \quad Var(\varepsilon_i) = \sigma_\varepsilon^2 > 0, \quad (1.3) \\ Cov(\delta_i, \delta_j) = Cov(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j, \quad Cov(\delta_i, \varepsilon_j) = 0, \end{aligned}$$

para todo  $i = 1, \dots, n$  e  $j = 1, \dots, n$ .

Este modelo pode ser classificado como um modelo funcional, estrutural ou ultraestrutural de acordo com as suposições feitas nas covariáveis  $\xi_i$ ,  $i = 1, \dots, n$  ([Cheng e Ness \(1999\)](#)).

**Modelo funcional:** quando  $\xi_i$  é considerado constante desconhecido;

**Modelo estrutural:** quando  $\xi_i$  é uma variável aleatória com média  $\mu$ , variância  $\sigma_x^2$  e ind. para  $i = 1, \dots, n$ ;

**Modelo ultraestrutural:** neste caso,  $\xi_i$ 's são variáveis aleatórias independentes como em um modelo estrutural, mas não são identicamente distribuídas, podendo ter diferentes médias  $\mu_i$  e variâncias iguais  $\sigma_x^2$ . O modelo ultraestrutural foi proposto por [Dolby \(1976\)](#) e é uma generalização dos modelos funcional e estrutural. Se  $\mu_1 = \mu_2 = \dots = \mu_n = \mu$ , então o modelo ultraestrutural reduz-se ao modelo estrutural. Se  $\sigma_x^2 = 0$ , o modelo ultraestrutural reduz-se ao modelo funcional.

Definido os modelos, surgem problemas ligados à consistência das estimativas ao estimar os parâmetros dos modelos com erros nas variáveis. Em modelos funcionais, devido à presença dos parâmetros incidentais (secundários), a estimação de máxima verossimilhança (MV) pode não existir, ou caso exista, pode não ser consistente. Nos modelos estruturais normais, os problemas de inconsistência ocorrem devido à falta de identificabilidade do modelo (veja [Apêndice C.1](#)). Modelos não identificáveis permitem que diferentes conjuntos de valores para os parâmetros originem a mesma distribuição para  $X$  e  $Y$ . A falta de identificabilidade implica ausência de estimadores consistentes, o que teria sérias implicações na teoria assintótica de estimação e testes de hipóteses.

É bem conhecido que o modelo estrutural normal não é identificável. Uma das formas de contornar o problema de identificabilidade é assumir que algum dos parâmetros é conhecido. [Cheng e Ness \(1999\)](#) apresentam as seguintes suposições:

- a) a razão entre as variâncias dos erros é conhecida  $\lambda = \sigma_e^2 / \sigma_\delta^2$ ;
- b)  $k_x = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\delta^2}$ , é conhecida;
- c)  $\sigma_\delta^2$  é conhecida;
- d)  $\sigma_e^2$  é conhecida;
- e) ambas variâncias dos erros são conhecidas,  $\sigma_\delta^2$  e  $\sigma_e^2$ ; ou
- f) o intercepto  $\alpha$  é conhecido e  $E(X) \neq 0$ .

Nas condições enumeradas acima, não assumimos o  $\beta$  conhecido porque a estimação deste parâmetro é, geralmente, o principal objetivo numa análise de regressão. O parâmetro  $\mu$  é o único parâmetro identificável sem nenhuma suposição adicional, isto porque  $\bar{X}$  é um estimador consistente para  $\mu$ .

Nesta tese vamos considerar o modelo de regressão estrutural com erros nas variáveis com intercepto nulo (univariado e multivariado), neste caso, como o intercepto é conhecido

o modelo é identificável. Iremos desenvolver também a teoria assintótica de um modelo de calibração ultraestrutural com réplicas.

Após ajustar um modelo estatístico deve ser realizada uma avaliação do modelo adotado, pois, podem existir anomalias no ajuste do modelo. Para detectar essas anomalias é preciso realizar um diagnóstico como análise de resíduos, análise da suposição de normalidade, análise de sensibilidade, análise da suposição de correlação nula e análise de dados atípicos [Singer, Nobre e Rocha \(2012, pag. 52\)](#).

Existe uma vasta literatura de métodos ou técnicas de análise de diagnóstico de dados atípicos em diferentes contextos de regressão linear normal, por exemplo, o DFBETA, o DFFIT [Belsley, Kuh e Welsch \(1980\)](#) e a distância de [Cook \(1977\)](#) são extensamente usados para a identificação de dados atípicos, os quais avaliam o impacto da retirada de uma observação particular nas estimativas da regressão. Essa metodologia foi rapidamente assimilada, estendida para diversas classes de modelos e muitos trabalhos lidam com o esquema de influência global como em [Cook e Weisberg \(1982\)](#), [Atkinson \(1985\)](#), [Chatterjee e Hadi \(1988\)](#) e [Weisberg \(2005\)](#).

[Brys, Hubert e Struyf \(2004\)](#) investigaram o “**medcouple**”, uma alternativa robusta para o coeficiente de assimetria clássico. [Hubert e Vandervieren \(2008a\)](#) propõem um ajuste do boxplot que inclui uma medida de assimetria na determinação da cauda. [Hubert e Vandervieren \(2008a\)](#) investigaram e usaram a medida robusta de assimetria “medcouple” (MC) introduzido em [Brys, Hubert e Struyf \(2004\)](#) para comparar com o octil e o quartil assimétrico. [Corbellini, Riani e Atkinson \(2015\)](#) fazem a discussão do artigo de [Hubert, Rousseeuw e Segaert \(2015\)](#), introduzindo um boxplot bivariado robusto na análise de detecção de dados atípicos.

[Hubert, Rousseeuw e Segaert \(2015\)](#) construíram novas técnicas numéricas e gráficas para a detecção de valores atípicos em dados funcionais multivariados. Para isso introduziram a distancia “bagdistance”, e também o “heatmap” de “bagdistance” que ajuda a distinguir entre dados atípicos isolados e persistentes.

Modelos cuidadosamente construídos em combinação com métodos de diagnóstico apropriados fornecem uma base útil para a análise estatística, no entanto, um simples diagnóstico de eliminação de casos pode ser computacionalmente intensivo e sofrer do problema de mascaramento. **O problema de mascaramento** ocorre quando um subconjunto anômalo não é detectado devido à presença de um outro, geralmente um subconjunto adjacente, ou seja, quando um subconjunto de dados atípicos e/ou influentes não é detectado por métodos convencionais e tais dados influenciam nas estimativas dos parâmetros. Por tanto, ocasiona má escolha do modelo.

[Cook e Weisberg \(1982, p.-31\)](#) descrevem que os resíduos de sinal oposto podem mascarar um ao outro de modo que nenhum aparece como dado atípico no método de diagnóstico usando resíduos. [Hawkins, Bradu e Kass \(1984\)](#) descrevem que os dados atípicos em uma regressão linear múltipla com  $p$  variáveis preditoras podem ser identificados extraindo todos os subconjuntos de tamanho  $p$  dos casos restantes e ajustando o modelo. Apesar de poder localizar

e testar a significância de um dado atípico, vários valores atípicos continuam a ser um problema. A dificuldade em identificá-los, mesmo em amostras aleatórias simples, tem sido reconhecida como problema de mascaramento. [Atkinson \(1985, p. 31\)](#) escreve que o mascaramento não seria revelado pelo simples cálculo das medidas de diagnóstico de exclusões individuais de cada observação, embora possa ser detectado por outras medidas de eliminação. A importância de uma observação particular pode não ser aparente até que alguma outra observação seja excluída ([Atkinson \(1985, p. 212\)](#)) na presença do efeito de mascaramento. [Leroy e Rousseeuw \(1987, p. viii\)](#) constataram que certos métodos robustos podem suportar pontos de alavanca, ao passo que outros não podem, e que alguns diagnósticos nos permitem detectar vários dados atípicos, enquanto outros são facilmente mascarados. [Chatterjee e Hadi \(1988\)](#) discutem métodos para a detecção de observações que em conjunto são dados influentes e individualmente, podem não ser considerados atípicos, pontos de alavanca, ou pontos influentes, esta situação é considerado efeito de mascaramento. [Rousseeuw e Zomeren \(1990\)](#) propõem calcular distâncias baseadas em estimativas muito robustas de localização e escala. Essas distâncias robustas são mais adequadas para expor os dados atípicos e evitar o efeito de mascaramento. Um dos trabalhos que tratam da identificação de dados atípicos múltiplos pode ser visto em [Hadi \(1992\)](#). Em particular, [Hadi \(1992\)](#) utilizou a distância de Mahalanobis para detectar os dados atípicos. A distância clássica de Mahalanobis não é claramente efetiva na identificação de dados atípicos múltiplos já que sofre de problemas de mascaramento [Hadi \(1992\)](#).

**Arrastamento** ocorre quando as observações “boas” são incorretamente identificadas como valores atípicos por causa da presença de um outro subconjunto de observações. Métodos que são menos suscetíveis aos problemas de mascaramento e arrastamento, (resíduos padronizados e pontos de alavanca) são dadas em [Hadi e Simonoff \(1993\)](#).

[Hadi e Simonoff \(1993\)](#) propõem um procedimento para a detecção de dados atípicos, considerando o problema de identificação e teste de hipótese de dados atípicos múltiplos. Os métodos de identificação de dados atípicos disponíveis, muitas vezes, não conseguem detectar os dados atípicos múltiplos porque são afetados pelas próprias observações que deveriam ser identificadas, “o problema de mascaramento”. [Atkinson e Mulira \(1993\)](#) usam a distância de Mahalanobis e constroem sequencialmente um subconjunto livre de dados atípicos, a partir de um pequeno subconjunto aleatório. O gráfico de “estalactite” fornece um resumo convincente de suspeitas de valores atípicos conforme o tamanho do subconjunto aumenta e também identifica valores extremos multivariados, mesmo na presença de **mascaramento** apreciável. [Cerioli e Riani \(1999\)](#) usam a técnica do algoritmo de **procura passo a frente** que ordena as observações de acordo com um modelo de auto-correlação especificado. Isso leva à identificação de dados atípicos. Em particular, o foco da análise é sobre modelos de previsão espacial. [Cerioli e Riani \(1999\)](#) apresentam exemplos que revelam claramente o poder do seu método em relação a procedimento de diagnóstico padrão (de exclusão).

[Atkinson e Riani \(2000\)](#) desenvolveram uma metodologia baseada nos artigos de [Hadi](#)

(1992) e Hadi e Simonoff (1993), o método **procura passo a frente** para a detecção de dados atípicos. O método **procura passo a frente** tem por finalidade descrever explicitamente na evolução, através de gráficos, o impacto que cada observação tem sobre o ajuste do modelo e respectivas análises de diagnósticos. Considerando modelos de regressão, o conjunto de dados é dividido em subconjuntos de tamanho  $m$  e o modelo ajustado aos dados de cada subconjunto. É selecionado um dos subconjuntos de tamanho  $m$  (conjunto com  $m$  dados) livre de dados atípicos, depois são selecionados subconjuntos de tamanho  $m + 1$  (conjunto com  $m + 1$  dados) e ajustado o modelo estatístico aos dados do subconjunto e assim sucessivamente até chegar a  $n$ . O propósito do método é observar as características do modelo (resíduos, estimativas dos parâmetros etc) e os dados na evolução do método, ver quais deles se destacam, são dados atípicos, quais sofrem do problema de mascaramento, etc.

Atkinson e Riani (2001) desenvolvem o procedimento de procura passo a frente que é um método simples e robusto para a detecção de observações atípicas e influentes que ordena as observações de acordo com a proximidade destes em relação ao modelo proposto, veja também Atkinson e Riani (2002). Atkinson e Riani (2004) estenderam a procura passo a frente para análise multivariada e novos desenvolvimentos foram realizados nos últimos anos como em Johansen e Nielsen (2010), Johansen, Nielsen *et al.* (2016) e Johansen e Nielsen (2016) desenvolveram a teoria assintótica para os estimadores obtidos no algoritmo procura passo a frente. Johansen, Nielsen *et al.* (2016) desenvolveram a análise da procura passo a frente usando alguns novos resultados para martingales e processos empíricos. Salini *et al.* (2016) estudaram o comportamento dos procedimentos de diagnóstico obtidos a partir de estimadores de regressão de pontos de ruptura quando nenhum dado aberrante está presente nos dados. Riani, Perrotta e Cerioli (2015) propuseram algumas melhorias computacionais do algoritmo procura passo a frente. Shi *et al.* (2015) usaram o método de influência local “stepwise” para detectar dados atípicos em modelo autorregressivo condicionalmente heterocedástico generalizado.

Bertaccini e Varriale (2007) apresentaram um método robusto simples para a detecção de observações atípicas e a análise de seus efeitos na estrutura ANOVA. Eles propuseram o uso da procura passo a frente que ordena as observações de acordo com a proximidade com o modelo. Atkinson e Riani (2007) usam a **procura passo a frente** para dados atípicos no modelo de regressão e usam a metodologia “backward” padrão para a seleção de variáveis (critério AIC) no processo de seleção de variáveis. Pesentia e Pirasa (2008) usam uma abordagem modificada da procura passo a frente aplicada à análise da série no tempo. Atkinson (2008) ilustra o uso da procura passo a frente para fornecer análises robustas de dados econométricos. Mavridis e Moustaki (2008) estenderam e implementaram o algoritmo de procura passo a frente para identificação de observações atípicas mascaradas em modelos de análise fatorial, enquanto que Mavridis e Moustaki (2009) implementaram o algoritmo de procura passo a frente para identificar dados atípicos mascarados em modelos de análise fatorial para dados binários.

Atkinson, Riani e Cerioli (2010) analisaram o desenvolvimento teórico do algoritmo

de procura passo a frente e ilustraram um exemplo de regressão com 330 observações e 9 variáveis explicativas potenciais. Foram feitos os procedimentos para dados multivariados, incluindo agrupamento, análise de séries temporais e detecção de fraude. Bellini (2010) aplicou o algoritmo de procura passo a frente para copulas elípticas em aplicações financeiras. Li e Valliant (2011) adaptaram o método procura passo a frente modificado para o diagnóstico de modelo de regressão linear para detecção de grupos de dados influentes. Zijlstra, Ark e Sijtsma (2011) aplicaram o algoritmo de procura passo a frente adaptado para a análise de escala de Mokken. Bellini (2012) propôs um procedimento que combina a super-eficiência de análise por envelopamento de dados e a procura passo a frente no problema de detecção de dados atípicos. Bellini e Riani (2012) analisaram modelos multivariados, incorporando o algoritmo de filtro de Kalman em um contexto de procura passo a frente. Torti *et al.* (2012) revelam que procura passo a frente têm o melhor desempenho ao comparar o poder dos estimadores de **procura passo a frente (FS: Forward Search)**, menor valor mediano quadrático (LMS: Least Median Squares) e mínimos quadrados aparados (LTS: Least Trimmed Squares). Riani *et al.* (2013) explicaram a falta de homogeneidade da regressão em dados de coordenadas espaciais fazendo gráficos de resíduos da regressão robusta da procura passo a frente que indicaram que alguns conjuntos de alocações são implausíveis.

Riani *et al.* (2014a) modelos de regressão são excepcionalmente suscetíveis à presença de valores atípicos. Em vez disso, existem vários métodos muito robustos, que resistem assintoticamente a 50% dos valores discrepantes nos dados, são preferidos. Riani *et al.* (2014a) apresentaram um quadro sistemático e parametrizado para a comparação não assintótica desses métodos. A robustez diz respeito à adequação de um único modelo aos dados que são gerados por dois, ou talvez mais, modelos. Suponhamos que a maior parte dos dados,  $1 - \varepsilon$ , onde  $0 < \varepsilon < 0.5$ , é gerado pelo modelo  $M_1(\theta_1)$  e a parte restante  $\varepsilon$  dos dados é gerada pelo modelo  $M_2(\theta_2)$ . Na ausência de dados atípicos, ou seja, quando  $\varepsilon = 0$ , um estimador robusto ideal teria uma variância que atinge o limite inferior de Cramer-Rao. Riani *et al.* (2014a) compararam e contrastaram as propriedades do que atualmente são considerados os cinco melhores métodos para uma regressão muito robusta (FS, LTS, LTSr: Least Trimmed Squares reweighted, S: estimator e MM: estimador). Os algoritmos que usaram estão todos publicamente disponíveis na biblioteca do Matlab de Análise de Dados de Pesquisa Avançada (FSDA). Veja Riani, Perrotta e Torti (2012). Riani *et al.* (2014a) apresentaram e descreveram um parâmetro  $\lambda$  que define um caminho paramétrico no espaço dos modelos e nos permite estudar, de forma sistemática, as propriedades dos estimadores à medida que os grupos de dados passam a estar muito afastados entre si. Riani *et al.* (2014b) combinaram o estimador S com regressão robusta usando a procura passo a frente, de tal modo a obter uma compreensão da importância das observações individuais.

Toman (2014) usa o algoritmo de procura passo a frente em análise fatorial confirmatória. Cerioli, Farcomeni e Riani (2014) trabalham no método de procura passo a frente para dados multivariados em um sólido conceito estatístico e mostram que os estimadores da procura passo a frente são fortemente consistentes no modelo normal multivariado. Atkinson *et al.* (2015)

propõem a procura passo a frente confirmatória, que estabelece uma associação precisa do cluster, o método funciona bem como métodos robustos, como TCLUS: Trimmed Cluster definido por [García-Escudero \*et al.\* \(2008\)](#), no entanto, não requer especificação prévia do número de clusters, nem do nível de corte de dados atípicos.

[Atkinson, Riani e Torti \(2016\)](#) desenvolveram a procura passo a frente heterocedástico em um exemplo que contém 1100 observações dos dados de comércio internacional. [Atkinson, Cerioli e Riani \(2016\)](#) fazem a discussão do artigo de [Johansen e Nielsen \(2016\)](#). [Bellini \(2016\)](#) estendeu a metodologia de procura passo a frente para o modelo autorregressivo do vetor cointegrado usando estimadores robustos de máxima pseudo-verossimilhança fazendo uma análise de dados contaminados. Com o objetivo de detectar observações atípicas mascaradas. [Cerasa e Cerioli \(2017\)](#) focaram no estimador-MM e procura passo a frente no estudo do problema da fusão de grupos homogêneos de observações pré-classificadas, motivada pela análise antifraude de dados de comércio internacional. [Atkinson, Corbellini e Riani \(2017\)](#) utilizaram a procura passo a frente como o método de regressão robusta para introduzir uma priori informativa.

A eliminação de pontos é o método mais tradicional de sensibilidade que consiste em avaliar o impacto da retirada de uma observação nas estimativas dos parâmetros de um modelo. Como foi visto, existe uma vasta literatura sobre o assunto. Contudo, o método de eliminação individual de pontos não verifica a influência conjunta das observações nas estimativas dos parâmetros. Nesse sentido, outra alternativa pode ser o método de **influência local**, que tem se constituído numa ferramenta importante na análise da influência conjunta das observações nas estimativas dos parâmetros do modelo. Durante as últimas três décadas, entre as muitas metodologias para a detecção de dados atípicos, foi proposto a metodologia de avaliação de influência local de [Cook \(1986\)](#), que é baseada na análise de influência de pequenas perturbações no modelo ou no conjunto de dados utilizando a curvatura normal. Hoje em dia, existem inúmeras aplicações sobre o método proposto por [Cook \(1986\)](#). Ele sugere que examinemos a direção de maior curvatura normal, pois a partir dela identificamos observações potencialmente influentes sob o esquema de perturbação ao qual sujeitamos o modelo.

[Beckman, Nachtsheim e Cook \(1987\)](#) descrevem o método de influência local para avaliar o modelo misto, [Lawrance \(1988\)](#) desenvolveu o diagnóstico de influência local em relação aos parâmetros estimados da transformação de Box-Cox na variável resposta no modelo linear, [Pettitt e Daud \(1989\)](#) investigaram a influência local em modelos de regressão de Cox com riscos proporcionais. [Thomas e Cook \(1990\)](#) avaliaram o efeito de pequenas perturbações nos dados contínuos (dados de usina nuclear) num ponto específico de predição de um modelo linear generalizado. [Escobar e Jr \(1992\)](#) adaptaram o método de influência local numa classe paramétrica de modelos para análise de sobrevivência, [Hines, Lawless e Carter \(1992\)](#) aplicaram os métodos de influência local em um modelo linear generalizado multinomial, [Tsai e Wu \(1992\)](#) investigaram a influência local em modelos autorregressivos de primeira ordem e modelos

heteroscedásticos, Paula (1993) aplicou influência local em modelos lineares com restrições nos parâmetros.

Wu e Luo (1993) avaliaram a segunda derivada ou a curvatura da superfície correspondente a inclinação que é relacionada com a curvatura da superfície de afastamento pela verossimilhança de Cook (1986), a abordagem de Lawrance (1988) é de primeira ordem comparado com a abordagem de Wu e Luo (1993) que produz uma abordagem de segunda ordem para o modelo de transformação de Box e Cox. Laurent e Cook (1993) mostram as conexões entre as medidas de alavanca e a influência local em modelos de regressão não linear.

Kim (1995) aplicou métodos de influência local em regressão multivariada, Pan, Fang e Rosen (1997) aplicaram métodos de influência local em regressão multivariada com covariância não estruturadas para dados biológicos, Shi (1997) usou a influência local em componentes principais, Lesaffre e Verbeke (1998) desenvolveram métodos de diagnóstico para modelos lineares mistos com base no conceito de influência local.

Mais tarde, Poon e Poon (1999) introduziram a **curvatura normal conformal** baseando-se no trabalho de Cook (1986), com a mesma finalidade de Cook (1986). Esta nova curvatura fornece uma medida de influência local que varia entre 0 e 1, com o objetivo de ter um ponto de corte ou uma faixa limite para detectar dados atípicos. Poon e Poon (2002b) adotaram uma perspectiva diferente e desenvolveram medidas que descrevem o comportamento total de uma função de afastamento pela verossimilhança, duas medidas, a saber, o comprimento de arco padronizado e o deslocamento médio, são construídas e suas aplicações são ilustradas com a ponderação de casos no modelo de regressão linear.

Galea, Paula e Bolfarine (1997), Liu (2000), Galea, Paula e Uribe-Opazo (2003) e Osorio, Paula e Galea (2007) apresentaram estudos de influência local em modelos de contornos elípticos. Zhu *et al.* (2007) mostraram que o tensor métrico da perturbação da variedade (múltipla entrada) fornece informação importante sobre a seleção de uma perturbação apropriada no diagnóstico de influência local de um modelo estatístico. Cadigan (2006) apresentou diagnósticos de influência local para medir a sensibilidade de um ponto de referência limite (biológico) estimado ao ajustar um modelo para dados de estoque e recrutamento, descrevendo efeitos das mudanças na ponderação de casos. Jung (2008) investigou a influência das observações nos coeficientes de regressão das equações de estimação generalizada usando a influência local. Shi e Chen (2008) estudaram a influência local das observações em modelos de regressão multinível. Zhu, Ibrahim e Tang (2011) desenvolveram um quadro geral de análise de influência bayesiano para avaliar vários esquemas de perturbação para os dados. Shi, Zhu e Ibrahim (2009) desenvolveram uma metodologia de influência local geral para realizar a análise de sensibilidade nos modelos lineares generalizados com covariáveis faltantes, ele examinou dois tipos de esquemas de perturbação (esquema de perturbação caso único e global).

Shi e Huang (2011) propuseram uma nova metodologia chamada de influência local “stepwise”. Lu, Shi e Chen (2012) propuseram a técnica de influência local “stepwise” em

modelo ARIMA. [Zhu, Ibrahim e Tang \(2014\)](#) desenvolveram métodos de análise de sensibilidade Bayesiano para avaliar várias perturbações (chamado perturbação Bayesiana de “manifold”) para métodos estatísticos com dados faltantes. [Zhu, Liu e Shi \(2016\)](#) usaram a metodologia de influência local para estudar dados de séries de tempo de contagem ajustando um modelo autorregressivo Poisson. [Dai et al. \(2016\)](#) estudaram a influência local no modelo espacial geral incluindo o modelo autorregressivo espacial e o modelo espacial com erro, como dois casos especiais, eles utilizaram o método de influência local “stepwise” para a análise de diagnóstico. Na literatura outras diversas referências em relação a influência local de [Cook \(1986\)](#) podem ser encontradas.

Considerando a aplicação da **metodologia de influência local de Cook (1986) aos modelos com erros de medição (MEM)**, temos por exemplo, [Zhao e Lee \(1995\)](#) que fizeram a avaliação de dados atípicos em modelos de regressão com erros de medição não lineares, [Russo \(2006\)](#) faz uma análise de diagnóstico para o modelo de regressão com erros nas variáveis para os dados de líquidos de bochecho ([Hadgu e Koch \(1999\)](#)).

[Castro, Galea-Rojas e Bolfarine \(2007\)](#) consideram a influência local no modelo de regressão com erros de medição para dados sobre doenças cardiovasculares. [Lachos, Montenegro e Bolfarine \(2008\)](#) observaram aspectos inferenciais de modelos com erros de medição com intercepto nulo quando a variável latente  $x$  segue uma distribuição normal-assimétrica avaliando a influência local nas estimativas dos parâmetros sob diferentes esquemas de perturbações utilizando os dados apresentados em [Hadgu e Koch \(1999\)](#). Outros exemplos como [Lachos et al. \(2010\)](#) utilizam a distribuição normal-assimétrica para a variável não observada e uma mistura de escala da distribuição normal para modelos multivariados com erros de medição. Isto fornece uma alternativa robusta para estimação de parâmetro em modelos de regressão com erros de medição multivariada.

[Ma e Li \(2010\)](#) desenvolveram uma estrutura para seleção de variáveis em modelos com erros de medição através da estimação de equações penalizadas. [Lachos, Angolini e Abanto-Valle \(2011\)](#) usam o método de influência local para a avaliação do aspecto de robustez das estimativas dos parâmetros seguindo o enfoque de [Zhu e Lee \(2001\)](#) que usam o algoritmo EM para dados incompletos nos modelos com erros de medição. [Giménez e Patat \(2014\)](#) pesquisaram a análise de influência local em modelos de calibração comparativa com dados replicados.

Considerando os modelos de regressão com erros nas variáveis, a influência local de [Cook \(1986\)](#), a curvatura normal conformal de [Poon e Poon \(1999\)](#) e a metodologia **procura passo a frente** de [Atkinson e Riani \(2000\)](#) podemos dividir a tese em duas partes.

Primeiro, considerando o modelo de calibração ultraestrutural com réplicas definido em [Talarico \(2014\)](#) para avaliar a equivalência de medições dos laboratórios, a proposta é fazer um estudo detalhado das propriedades assintóticas associada a este modelo para a obtenção da distribuição limite dos dados baseado no trabalho de [Sweeting \(1980\)](#), onde é provado a normalidade assintótica do estimador de máxima verossimilhança para amostras dependentes

com matriz de informação de Fisher aleatória, o que significa que o modelo é não ergódico (Bhat (1974), Keiding (1974), Keiding (1975)) porém positiva definida. Outra proposta é provar que um subconjunto dos estimadores de máxima verossimilhança do modelo definido em Talarico (2014) é consistente. A partir de condições de regularidade na verossimilhança, vamos estabelecer o teste assintótico de Wald.

A segunda parte, se refere a metodologia de diagnóstico. Baseado nas metodologias descritas acima, Cook (1986), Poon e Poon (1999), Atkinson e Riani (2000), vamos propor uma nova metodologia conjunta do tipo procura passo a frente de Atkinson e Riani (2000) para a obtenção de pontos influentes mascarados considerando a influência local de Cook (1986) e a curvatura normal conformal de Poon e Poon (1999). A metodologia proposta será aplicada a modelos de regressão com erros nas variáveis com intercepto nulo para a obtenção de dados mascarados influentes.

Há situações onde o uso dos modelos com intercepto nulo são adequados (Chan e Mak (1979)). Aoki (2001) propõe um modelo de regressão com erros nas variáveis com intercepto nulo onde o interesse é avaliar a eficácia de uma escova de dente experimental em relação a uma escova de dente convencional na remoção de placa bacteriana. Aoki *et al.* (2003) propõe um modelo de regressão com erros nas variáveis multivariado com intercepto nulo onde o interesse é avaliar dois tipos de líquidos de bochecho experimentais com um líquido de bochecho controle na remoção de placas bacterianas. Estes conjuntos de dados serão utilizados nesta tese e apresentados a seguir com os respectivos modelos.

## 1.1 Dados de escovas de dentes

Aoki (2001) considera o uso do modelo com intercepto nulo e erros de medição para reanálise dos dados de um estudo pré-teste/pós-teste destinado para comparar dois tipos de escovas de dentes com respeito a sua eficácia em remover a placa bacteriana. Daqui para frente este conjunto de dados será chamado de dados de escovas de dentes e estes dados se encontram no Apêndice A.1. A proposta do modelo sem o intercepto é devido ao fato de que se um indivíduo possui o índice de placa bacteriana nulo no início do estudo, após a escovação o índice de placa bacteriana esperada é nula. Como a covariável (placa bacteriana) é medida com erro, uma forma alternativa para a análise deste conjunto de dados é considerar os modelos com erros de medição. Adicionalmente, como os indivíduos foram avaliados sob duas condições experimentais diferentes, o modelo considera a possível dependência existente nas medidas tomadas no mesmo indivíduo (ver Aoki, Bolfarine e Singer (2001), Aoki, Bolfarine e Singer (2002), Aoki, Singer e Bolfarine (2007)).

Em Aoki, Bolfarine e Singer (2001) é definido o modelo com  $p$  tratamentos (no caso do

conjunto de dados de escovas de dentes,  $p$  é igual a 2) dado por:

$$\begin{cases} X_{ij} = x_{ij} + u_{ij}, \\ Y_{ij} = \beta_i x_{ij} + e_{ij}, \quad i = 1, \dots, p, j = 1, \dots, n, \end{cases} \quad (1.4)$$

onde  $X_{ij}$  e  $Y_{ij}$ , são os valores observados da variável explanatória e da variável resposta, respectivamente, para o sujeito  $j$  que utilizou o tratamento  $i$ ,  $x_{ij}$  corresponde ao verdadeiro valor não observado da variável explanatória e  $\beta_i$  representa a inclinação (desconhecida). Em que  $u_{ij} \stackrel{ind.}{\sim} N(0, \sigma_u^2)$ ,  $e_{ij} \stackrel{ind.}{\sim} N(0, \lambda_i \sigma_u^2)$ ,  $x_{ij} \stackrel{ind.}{\sim} N(\mu, \sigma_x^2)$ ,  $u_{ij}$ ,  $e_{ij}$  e  $x_{ij}$  independentes,  $\forall i = 1, \dots, p, j = 1, \dots, n$ . Definindo,

$$x_{ij} = \mu + a_j, i = 1, \dots, p, j = 1, \dots, n, \quad (1.5)$$

em que  $a_j \stackrel{ind.}{\sim} N(0, \sigma_x^2)$ , e

$$\begin{pmatrix} e_{ij} \\ u_{ij} \\ x_{ij} \end{pmatrix} \stackrel{ind.}{\sim} N_3 \left( \begin{bmatrix} 0 \\ 0 \\ \mu \end{bmatrix}, \begin{bmatrix} \lambda_i \sigma_u^2 & 0 & 0 \\ 0 & \sigma_u^2 & 0 \\ 0 & 0 & \sigma_x^2 \end{bmatrix} \right), i = 1, \dots, p, j = 1, \dots, n, \quad (1.6)$$

temos:

$$\rho_{i,i'} = \text{corr}(X_{ij}, X_{i'j}) = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2}, i, i' = 1, \dots, p, j = 1, \dots, n, i \neq i'.$$

Seja,  $\mathbf{Z}_j = (\mathbf{X}_j^T, \mathbf{Y}_j^T)^T = (X_{1j}, \dots, X_{pj}, Y_{1j}, \dots, Y_{pj})^T$ , os valores observados do individuo  $j$  e considerando

$$\mathbf{b} = (\mathbf{1}_p^T, \boldsymbol{\beta}^T)^T, \text{ com } \boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T \text{ e } \mathbf{m} = \boldsymbol{\mu} \mathbf{b} \quad (1.7)$$

temos:

$$\mathbf{Z}_j \stackrel{iid}{\sim} N_{2p}(\mathbf{m}, \mathbf{V}), \quad j = 1, \dots, n, \quad (1.8)$$

em que,

$$\mathbf{V} = \sigma^2 \mathbf{A} + \sigma_x^2 \mathbf{b} \mathbf{b}^T, \text{ com } \mathbf{A} = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_p \\ \mathbf{0}_p & \mathbf{D}(\boldsymbol{\lambda}) \end{pmatrix}, \text{ e } \boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p)^T. \quad (1.9)$$

O determinante e a inversa de  $\mathbf{V}$  são dados por:

$$|\mathbf{V}| = (\sigma^2)^{2p} |\mathbf{D}(\boldsymbol{\lambda})| c, \quad \text{com } c = 1 + \frac{\sigma_x^2}{\sigma^2} (p + \boldsymbol{\beta}^T \mathbf{D}(\boldsymbol{\lambda})^{-1} \boldsymbol{\beta}) \text{ e}$$

$$\mathbf{V}^{-1} = \sigma^{-2} \mathbf{A}^{-1} - c^{-1} \mathbf{M}, \quad \text{com } \mathbf{M} = \frac{\sigma_x^2}{\sigma^4} \mathbf{A}^{-1} \mathbf{b} \mathbf{b}^T \mathbf{A}^{-1}.$$

O logaritmo da função verossimilhança do modelo (1.4) é dada por

$$\ell(\boldsymbol{\theta}, \mathbf{Z}) = -np \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{j=1}^n (\mathbf{Z}_j - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_j - \mathbf{m}). \quad (1.10)$$

Sendo  $\boldsymbol{\theta} = (\boldsymbol{\beta}, \mu, \sigma_x^2, \sigma_u^2, \boldsymbol{\lambda})^T$ , o conjunto de dados pode ser visto no Apêndice A.1.

## 1.2 Dados do líquidos de bochecho

Um outro ensaio clínico dentário é apresentado em um artigo escrito por [Hadgu e Koch \(1999\)](#), onde 109 voluntários adultos de ambos os sexos foram escolhidos aleatoriamente para utilizar entre dois líquidos de bochecho experimentais ( $A$  e  $B$ ) ou um líquido de bochecho controle e avaliado sob estas três condições experimentais, com respeito ao índice de placa bacteriana no início do estudo, depois de três meses e depois de seis meses do início do estudo com o uso do líquido de bochecho  $A$ , líquido de bochecho  $B$ , ou o líquido de bochecho controle. Como a covariável (índice de placa bacteriana no início do estudo) é medida com erro, uma forma alternativa para analisar este conjunto de dados é considerar os modelos com erros de medição. Em adição, como o índice de placa bacteriana foi coletado no início do estudo e após 3 e 6 meses do início do estudo, o modelo apresentado leva em consideração a possível dependência nos dados coletados no mesmo indivíduo e o modelo considerado é com intercepto nulo, pelas mesmas razões descritas nos dados da escova de dente. Desta forma, foi considerado um modelo de regressão com erros nas variáveis multivariado com intercepto nulo em [Aoki et al. \(2003\)](#) e [Russo \(2006\)](#).

Neste estudo cada indivíduo usou um dos três líquidos de bochecho:

- líquido de bochecho controle ( $i = 1$ ),
- líquido de bochecho  $A$  ( $i = 2$ ),
- líquido de bochecho  $B$  ( $i = 3$ ).

Neste caso, temos  $p = 3$  líquidos de bochecho e  $n = 105$ , onde  $n_1 = 36$ ,  $n_2 = 33$ , e  $n_3 = 36$ . Daqui para frente será chamado de dados de líquidos de bochecho e o conjunto de dados pode ser encontrado no Apêndice [A.2](#). O modelo definido em [Aoki et al. \(2003\)](#) para analisar este conjunto de dados é dado por:

$$\begin{aligned} \mathbf{x}_i &= \boldsymbol{\xi}_i + \boldsymbol{\delta}_i, \\ \mathbf{y}_i &= \mathbf{X}_i \boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \dots, p, \end{aligned} \tag{1.11}$$

onde:

- i) no caso dos dados de líquidos de bochecho,  $p = 3$ ,
- ii) o vetor observado  $\mathbf{x}_i = (x_{i1}, \dots, x_{in_i})^T$  representa o índice de placa dentária no início do estudo, antes do uso do líquido de bochecho  $i$ ,  $i = 1, 2, 3$ ,
- iii) o vetor observado  $\mathbf{y}_i = (\mathbf{y}_{1i}^T, \mathbf{y}_{2i}^T)^T$ , com  $\mathbf{y}_{1i} = (y_{1i1}, \dots, y_{1in_i})^T$  e  $\mathbf{y}_{2i} = (y_{2i1}, \dots, y_{2in_i})^T$ , representam, respectivamente, os índices de placa dentária após 3 e 6 meses do início do estudo, com o uso do líquido de bochecho  $i$ ,  $i = 1, 2, 3$ ,

iv) o vetor  $\boldsymbol{\xi}_i = (\xi_{i_1}, \dots, \xi_{i_{n_i}})^T$ , representa os verdadeiros valores do índice de placa dentária no início do estudo, antes do uso do líquido de bochecho  $i$ ,  $i = 1, 2, 3$ ,

v)

$$\mathbf{X}_i = \begin{pmatrix} \boldsymbol{\xi}_i & 0 \\ 0 & \boldsymbol{\xi}_i \end{pmatrix}, i = 1, 2, 3, \quad (1.12)$$

vi)  $\boldsymbol{\delta}_i = (\delta_{i_1}, \dots, \delta_{i_{n_i}})^T$ ,  $\boldsymbol{\varepsilon}_i = (\boldsymbol{\varepsilon}_{1i}^T, \boldsymbol{\varepsilon}_{2i}^T)^T$ , com  $\boldsymbol{\varepsilon}_{1i} = (\varepsilon_{1i_1}, \dots, \varepsilon_{1i_{n_i}})^T$  e  $\boldsymbol{\varepsilon}_{2i} = (\varepsilon_{2i_1}, \dots, \varepsilon_{2i_{n_i}})^T$  representam os erros de medição no início do estudo e após 3 e 6 meses do início do estudo, respectivamente, com o uso do líquido de bochecho  $i$ ,  $i = 1, 2, 3$ ,

vii)  $\boldsymbol{\beta}_i = (\beta_{1i}, \beta_{2i})^T$  representa a proporção de placa dentária restante após o uso do líquido de bochecho  $i$ , decorridos 3 e 6 meses respectivamente.

Assumindo que  $\delta_{ij} \stackrel{ind.}{\sim} N(0, \sigma_\delta^2)$ ,  $\varepsilon_{1ij} \stackrel{ind.}{\sim} N(0, \sigma_{e_{1i}}^2)$ ,  $\varepsilon_{2ij} \stackrel{ind.}{\sim} N(0, \sigma_{e_{2i}}^2)$ ,  $\xi_{ij} \stackrel{ind.}{\sim} N(\mu, \sigma_x^2)$ ,  $\delta_{ij}$ ,  $\varepsilon_{ij}$  e  $\xi_{ij}$  independentes, para  $i = 1, \dots, p$ ,  $j = 1, \dots, n_i$ , o vetor observado é dado por:

$$\mathbf{z}_{ij} = \begin{pmatrix} x_{ij} \\ y_{1ij} \\ y_{2ij} \end{pmatrix} \sim N_3 \left( \mathbf{m}_i = \begin{bmatrix} \mu \\ \beta_{1i}\mu \\ \beta_{2i}\mu \end{bmatrix}, \mathbf{V}_i = \begin{bmatrix} \sigma_x^2 + \sigma_\delta^2 & \beta_{1i}\sigma_x^2 & \beta_{2i}\sigma_x^2 \\ \beta_{1i}\sigma_x^2 & \beta_{1i}^2\sigma_x^2 + \sigma_{e_{1i}}^2 & \beta_{1i}\beta_{2i}\sigma_x^2 \\ \beta_{2i}\sigma_x^2 & \beta_{1i}\beta_{2i}\sigma_x^2 & \beta_{2i}^2\sigma_x^2 + \sigma_{e_{2i}}^2 \end{bmatrix} \right)$$

para  $i = 1, \dots, p$ ,  $j = 1, \dots, n_i$  em que  $\mathbf{z}_i = (\mathbf{z}_{i_1}^T, \dots, \mathbf{z}_{i_{n_i}}^T)^T$ ,  $\mathbf{z} = (\mathbf{z}_1^T, \dots, \mathbf{z}_p^T)^T$ , com  $N = \sum_{i=1}^p n_i$ . O logaritmo da função de verossimilhança do modelo definido em (1.11) é dado por

$$\ell(\mathbf{z}, \boldsymbol{\theta}) = -\frac{3N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^p n_i \log |\mathbf{V}_i| - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} (\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{V}_i^{-1} (\mathbf{z}_{ij} - \mathbf{m}_i). \quad (1.13)$$

Todas as expressões do modelo se encontram na dissertação de [Russo \(2006\)](#).

## 1.3 Dados da Potência do Motor

Na dissertação de [Talarico \(2014\)](#) foi proposto um modelo de regressão ultraestrutural multivariado com réplicas, para poder avaliar a equivalência entre as medições feitas por diferentes laboratórios credenciados pela INMETRO (Instituto Nacional de Metrologia, Normalização e Qualidade Industrial). Neste estudo, a potência do motor foi medida em diversos pontos de rotação por  $p$  laboratórios.

Sejam:

- $y_{ijk}$  a  $k$ -ésima réplica da medição do verdadeiro valor não observado da potência do motor no,  $j$ -ésimo ponto de rotação ou patamar pelo  $i$ -ésimo laboratório,  $i = 1, \dots, p$ ,  $j = 1, \dots, m$  e  $k = 1, \dots, n_i$ ,

- $x_j$  o verdadeiro valor não observado da potência do motor no  $j$ -ésimo ponto de rotação,
- $Y_{ijk}$  o valor observado da  $k$ -ésima medição da potencia do motor no  $j$ -ésimo ponto de rotação obtido pelo  $i$ -ésimo laboratório.

Assumindo que  $y_{ijk}$  satisfaz a relação linear ultraestrutural com  $x_j$ , o modelo proposto pode ser representado como:

$$\begin{aligned} y_{ijk} &= \alpha_i + \beta_i x_j, \\ Y_{ijk} &= y_{ijk} + e_{ijk}, \end{aligned} \quad (1.14)$$

em que  $\mathbb{E}(e_{ijk}) = 0$ ,  $\text{Var}(e_{ijk}) = \sigma_{ij}^2$ ,  $\mathbb{E}(x_j) = \mu_{x_j}$ ,  $\text{Var}(x_j) = \sigma_{x_j}^2$ ,  $x_j$  independente de  $e_{ijk}$ , para cada  $k = 1, \dots, n_i$ ,  $i = 1, \dots, p$  e  $j = 1, \dots, m$ .

Neste modelo,  $\mu_{x_j}$  é a média e  $\sigma_{x_j}^2$  a variabilidade natural associado com o valor não observado do mensurando (valor a ser medido) no  $j$ -ésimo ponto de rotação. No geral, o parâmetro  $\sigma_{x_j}^2$  é determinado por um laboratório especializado durante o estudo de estabilidade que é conduzido para garantir a estabilidade do item em teste. Em nosso exemplo, a GM motor train desenvolveu um motor padrão e, durante o estudo de estabilidade, avaliou sua variabilidade natural ( $\sigma_{x_j}^2$ ).

Considera-se, sem perda de generalidade, que o primeiro laboratório é o laboratório de referência e neste caso,  $\alpha_1 = 0$  e  $\beta_1 = 1$ . Desta forma, temos que:

$$Y_{1jk} = x_j + e_{1jk}, \quad j = 1, \dots, m \text{ e } k = 1, \dots, n_1. \quad (1.15)$$

Aqui  $e_{1jk}$  é o erro de medição que corresponde ao laboratório de referência. A variância  $\sigma_{1j}^2$  é determinada pela variância combinada calculada e fornecida pelo laboratório de referência seguindo o protocolo proposto por [ISO GUM \(1995\)](#).

$$Y_{ijk} = \alpha_i + \beta_i x_j + e_{ijk}, \quad i = 2, \dots, p, \quad j = 1, \dots, m \text{ e } k = 1, \dots, n_i,$$

em que  $\alpha_i$  descreve o viés aditivo e  $\beta_i$  descreve o viés multiplicativo do laboratório  $i = 1, \dots, p$ . Da mesma forma,  $e_{ijk}$  representa o erro de medição associado com o  $i$ -ésimo laboratório. A variância  $\sigma_{ij}^2$  é determinada pela variância combinada calculada e fornecida pelo  $i$ -ésimo laboratório seguindo o protocolo proposto por [ISO GUM \(1995\)](#).

Considerando o modelo definido por (1.14) e (1.15) a covariância entre as observações tomados no mesmo ponto de rotação do motor pelo laboratório de referência ( $i$ -ésimo laboratório) é dado por  $\sigma_{x_j}^2$  ( $\beta_i^2 \sigma_{x_j}^2$ ) e a covariância entre as observações do laboratório de referência e o  $i$ -ésimo laboratório no mesmo ponto de rotação é dado por  $\beta_i \sigma_{x_j}^2$ , enquanto que a covariância entre as observações do  $i$ -ésimo e  $h$ -ésimo laboratório no  $j$ -ésimo ponto de rotação do motor é dado por  $\beta_i \beta_h \sigma_{x_j}^2$ , isto é:

$$\text{cov}(Y_{1jk}, Y_{1jl}) = \sigma_{x_j}^2, \quad \text{cov}(Y_{ijk}, Y_{ijl}) = \beta_i^2 \sigma_{x_j}^2, \quad \text{cov}(Y_{1jk}, Y_{ijl}) = \beta_i \sigma_{x_j}^2, \quad \text{cov}(Y_{ijk}, Y_{hjl}) = \beta_i \beta_h \sigma_{x_j}^2,$$

$i, h = 2, \dots, p; j = 1, \dots, m; k, l = 1, \dots, n_i$ .

Seja  $\mathbf{Y}_{1j} = (Y_{1j1}, \dots, Y_{1jn_1})^T$  o vetor observado da potência do motor do laboratório de referência no  $j$ -ésimo ponto de rotação,  $\mathbf{Y}_{ij} = (Y_{ij1}, \dots, Y_{ijn_i})^T$  o vetor observado da potência do motor do  $i$ -ésimo laboratório no  $j$ -ésimo ponto de rotação,  $\mathbf{Y}_j^n = (\mathbf{Y}_{1j}^\top, \dots, \mathbf{Y}_{pj}^\top)^\top$ , os valores observados de todos os laboratórios, 1 até  $p$ , no  $j$ -ésimo ponto de rotação e finalmente,  $\mathbf{Y}^n = (\mathbf{Y}_1^{n\top}, \dots, \mathbf{Y}_m^{n\top})^\top$  os dados observados com  $n = \sum_{i=1}^p n_i$ . Então, assumindo que  $\mathbf{Y}_j^n \sim N_n(\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ , nós temos:  $\mathbf{Y}_{1j} \sim N_{n_1}(\boldsymbol{\mu}_{1j}, \boldsymbol{\Sigma}_{11j})$ ,  $\mathbf{Y}_{ij} \sim N_{n_i}(\boldsymbol{\mu}_{ij}, \boldsymbol{\Sigma}_{ij})$ ,  $i = 2, \dots, p; j = 1, \dots, m$ ; onde  $\boldsymbol{\mu}_{1j} = \mu_{x_j} \mathbf{1}_{n_1}$ ,  $\boldsymbol{\Sigma}_{11j} = \sigma_{1j}^2 \mathbf{I}_{n_1} + \sigma_{x_j}^2 \mathbf{1}_{n_1} \mathbf{1}_{n_1}^T$ ,  $\boldsymbol{\mu}_{ij} = (\alpha_i + \beta_i \mu_{x_j}) \mathbf{1}_{n_i}$ ,  $\boldsymbol{\Sigma}_{ij} = \sigma_{ij}^2 \mathbf{I}_{n_i} + \beta_i^2 \sigma_{x_j}^2 \mathbf{1}_{n_i} \mathbf{1}_{n_i}^T$ , com  $\mathbf{I}_{n_i}$  e  $\mathbf{1}_{n_i}$  denotando, respectivamente, a matriz identidade de dimensão  $n_i$  e um vetor composto por  $n_i$  uns. Além disso,

$$\boldsymbol{\mu}_j = (\boldsymbol{\mu}_{1j}^T, \dots, \boldsymbol{\mu}_{pj}^T)^T = \boldsymbol{\alpha} + \mu_{x_j} \boldsymbol{\beta} \quad \text{e} \quad \boldsymbol{\Sigma}_j = D(\boldsymbol{\sigma}_j^2) + \boldsymbol{\sigma}_{x_j}^2 \boldsymbol{\beta} \boldsymbol{\beta}^\top,$$

com  $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1^\top, \boldsymbol{\alpha}_2 \mathbf{1}_{n_2}^\top, \dots, \boldsymbol{\alpha}_p \mathbf{1}_{n_p}^\top)^\top$ ,  $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2 \mathbf{1}_{n_2}^\top, \dots, \boldsymbol{\beta}_p \mathbf{1}_{n_p}^\top)^\top$ ,  $\boldsymbol{\sigma}_j^2 = (\sigma_{1j}^2 \mathbf{1}_{n_1}^\top, \dots, \sigma_{pj}^2 \mathbf{1}_{n_p}^\top)^\top$ ,  $\mathbf{0}_{n_1}$  denotando um vetor composto por  $n_1$  zeros e  $D(\boldsymbol{a})$  denotando a matriz diagonal com elementos na diagonal dados por  $\boldsymbol{a}$ . Além do mais,

$$f_{\mathbf{Y}_j^n}(\mathbf{y}_j^n, \boldsymbol{\theta}) = (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Sigma}_j|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\mathbf{y}_j^n - \boldsymbol{\mu}_j)^\top \boldsymbol{\Sigma}_j^{-1} (\mathbf{y}_j^n - \boldsymbol{\mu}_j) \right\}, \quad j = 1, \dots, m$$

e

$$f_{\mathbf{Y}^n}(\mathbf{y}^n, \boldsymbol{\theta}) = \prod_{j=1}^m f_{\mathbf{Y}_j^n}(\mathbf{y}_j^n, \boldsymbol{\theta}) \quad \text{com} \quad \boldsymbol{\theta} = (\mu_{x_1}, \dots, \mu_{x_m}, \alpha_2, \dots, \alpha_p, \beta_2, \dots, \beta_p)^T = (\theta_1, \dots, \theta_{m+2(p-1)})^T.$$

O logaritmo da função de verossimilhança é dado por

$$L^n(\boldsymbol{\theta}) = \log f_{\mathbf{Y}^n}(\mathbf{y}^n, \boldsymbol{\theta}) = -\frac{mn}{2} \log(2\pi) - \frac{1}{2} \sum_{j=1}^m \log a_j^n - \frac{1}{2} \sum_{j=1}^m \sum_{i=1}^p n_i \log(\sigma_{ij}^2) - \frac{1}{2} \sum_{j=1}^m Q_j^n, \quad (1.16)$$

onde  $a_j^n = 1 + \boldsymbol{\sigma}_{x_j}^2 \boldsymbol{\beta}^\top D^{-1}(\boldsymbol{\sigma}_j^2) \boldsymbol{\beta}$  e  $Q_j^n = (\mathbf{y}_j^n - \boldsymbol{\mu}_j)^\top \boldsymbol{\Sigma}_j^{-1} (\mathbf{y}_j^n - \boldsymbol{\mu}_j)$ ,  $j = 1, \dots, m$ .

Para mais detalhes em relação às expressões do modelo, consultar a dissertação de [Talarico \(2014\)](#).



---

# PROPRIEDADES ASSINTÓTICAS DO MODELO DE CALIBRAÇÃO ULTRAESTRUTURAL COM RÉPLICAS

---

---

A calibração é uma metodologia que usa técnicas estatísticas para comparar diferentes sistemas de medição. Neste capítulo desenvolveremos a teoria assintótica necessária para provar a consistência e a normalidade assintótica do estimador de máxima verossimilhança de um subconjunto dos parâmetros considerando o modelo definido em (1.14). Além disso obtemos a estatística do teste de Wald para testar as hipóteses de interesse e fornecemos uma análise gráfica para avaliar a equivalência das medições dos laboratórios com respeito à medição do laboratório de referencia. Finalmente, foi feito um estudo de simulação e aplicamos os resultados ao conjunto de dados reais descrito na Introdução (Seção 1.3).

## 2.1 O modelo

Estudos de proficiência são conduzidos para avaliar a equivalência das medições dos laboratórios (ver [ISO 17043 \(2010\)](#)). Nesses estudos, um valor de referência de algum mensurando (a quantidade a ser medida) é determinado e os resultados de todos os laboratórios são comparados com esse valor de referência. Várias técnicas estatísticas foram adotadas para avaliar a equivalência entre as medições dos laboratórios. Isso inclui técnicas estatísticas clássicas, como teste t pareado, z-escore, erro normalizado, análise de variância de medidas repetidas, gráfico de Bland-Altman, ver [Bland e Altman \(1986\)](#) e [ISO 17043 \(2010\)](#), e suas referências.

Um ponto crítico dessas técnicas é o fato de que a fonte de variabilidade do tipo B ([ISO GUM, 1995](#)) não é considerada. Nesse sentido, [Pinto, Aoki e Silva \(2009\)](#) estenderam o modelo de [Jaech \(1985\)](#) para abranger a fonte de variação do tipo B e avaliaram esse modelo sob distribuições elípticas. [Toman \(2007\)](#) propôs um modelo hierárquico bayesiano para incluir a

fonte de variação do tipo B no modelo.

Apesar da quantidade de técnicas para avaliar a equivalência entre as medições dos laboratórios, essas abordagens consideram a referência como um valor único. Em alguns estudos de proficiência, o item em teste é medido em diferentes pontos. Como uma ilustração, considere o sistema de medição para avaliar a potência do motor descrito na introdução. Nesse caso, a potência (torque vezes rotação) é medida em diferentes pontos de rotação e, como resultado, temos um valor de referência para cada ponto. Desta forma, em Talarico (2014) foi definido um modelo multivariado para avaliar a equivalência entre as medições dos laboratórios, considerando que o item em teste é medido em diferentes pontos (modelo definido em 1.14 e 1.15) e foram obtidas as expressões da função score e matriz de informação observada que podem ser encontradas no Apêndice C.4.2, assim como o algoritmo EM para obtenção dos EMV.

## 2.2 Propriedades assintóticas

Nesta seção, desenvolveremos a teoria assintótica necessária para provar a consistência e a normalidade assintótica do EMV em relação aos parâmetros de viés. Na sequência, aplicaremos as propriedades de regularidade da função de probabilidade para estabelecer os resultados assintóticos, conforme proposto por Sweeting (1980) e Weiss (1973). Para um dado  $\theta$ , definimos

$$\mathbb{P}_{\theta, (x_1, \dots, x_m)}^n(B^n) = \int_{B^n} f_{Y_c^n}(\mathbf{y}_c^n, \theta) d\mathbf{y}_c^n,$$

para todo  $B^n \in \beta(\mathbb{R}^{m(n+1)})$ , em que  $\beta(\mathbb{R}^{m(n+1)})$  é a  $\sigma$ -álgebra de Borel. Aplicando o teorema de extensão de Kolmogorov, existe uma única probabilidade  $\mathbb{P}_{\theta, (x_1, \dots, x_m)}$  definido em  $(\mathbb{R}^\infty, \beta(\mathbb{R}^\infty))$  tal que

$$\mathbb{P}_{\theta, (x_1, \dots, x_m)}(B^n \times \mathbb{R} \times \mathbb{R} \times \dots) = \mathbb{P}_{\theta, (x_1, \dots, x_m)}^n(B^n),$$

para todo  $B^n \in \beta(\mathbb{R}^{m(n+1)})$ . A distribuição marginal do dado observado será denotado por  $\mathbb{P}_\theta$  e a distribuição marginal das variáveis não observadas  $(x_1, \dots, x_m)$  será denotado por  $\mathbb{P}_{(x_1, \dots, x_m)}$ . Dizemos que  $n \rightarrow \infty$  quando  $n_i \rightarrow \infty$  e  $\frac{n_i}{n} \rightarrow w_i$  em que  $w_i$  é uma constante positiva para todo  $i = 1, \dots, p$ .

Seja  $M^\ell$  o espaço de todas as matrizes  $\ell \times \ell$ . A norma  $\|A\|$  da matriz  $A$  é  $\|A\| = \max\{|A_{hs}| : h, s = 1, \dots, \ell\}$ . Uma sequência de matrizes  $\{A^u : u = 1, 2, \dots\}$  converge para um limite  $A$  se, e somente se,  $\|A^u - A\| \rightarrow 0$ . Se a matriz  $A$  é definida positiva, escrevemos  $A > 0$ . Neste caso,  $A^{1/2}$  denota a raiz quadrada positiva simétrica de  $A$ . Na mesma forma, para um dado vetor  $\mathbf{v} = (v_1, \dots, v_\ell) \in \mathbb{R}^\ell$ , consideramos a norma de  $\mathbf{v}$  como segue  $|\mathbf{v}| = \max\{|v_1|, \dots, |v_\ell|\}$ .

Denotamos  $\mathbf{1}_{m+2(p-1)} = (1, \dots, 1)_{m+2(p-1)}$  e por  $B_c(n, \theta)$  o conjunto de todos os vetores  $\boldsymbol{\psi} \in \mathbb{R}^{m+2(p-1)}$  tal que  $\sqrt{n} |\boldsymbol{\psi} - \theta| \leq c \mathbf{1}_{m+2(p-1)}$ , em que  $c$  é uma constante positiva. Além disso, denotamos por  $R_c(n, \theta)$  o conjunto de todos os vetores aleatórios  $\boldsymbol{\phi}$  com valores em  $\mathbb{R}^{m+2(p-1)}$  tal que  $\sqrt{n} |\boldsymbol{\phi} - \theta| \leq c \mathbf{1}_{m+2(p-1)}$ . Aqui, tomamos o vetor aleatório  $\boldsymbol{\phi}$  como função

do dado observado  $\mathbf{Y}^n$ . Temos a  $L^n$  logaritmo da função de verossimilhança definido em (1.16). A matriz de informação observada  $J^n(\boldsymbol{\theta}) = -\frac{\partial^2 L^n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$ .

**Lema 1.** Para alguma sequencia  $\{\boldsymbol{\psi}^n : \boldsymbol{\psi}^n \in B_c(n, \boldsymbol{\theta}), n \geq 1\}$  e  $\{\boldsymbol{\phi}^n : \boldsymbol{\phi}^n \in R_c(n, \boldsymbol{\theta}), n \geq 1\}$  existe uma matriz aleatória positiva semidefinida  $W(\boldsymbol{\theta})$  tal que

$$\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)} \left[ \left\| \frac{1}{n} J^n(\boldsymbol{\phi}^n) - W(\boldsymbol{\theta}) \right\| \geq \varepsilon \right] \rightarrow 0, \quad n \rightarrow \infty.$$

*Demonstração.* Veja o Apêndice B. □

A matriz aleatória  $W(\boldsymbol{\theta})$  tem duas características importantes. Primeiro, todo componente associado a  $\mu_{x_j}$  é nulo, isto significa que não temos informações suficientes para estimar  $\mu_{x_j}$  de uma forma consistente. Isso é uma consequência do fato de que, para qualquer ponto  $j = 1, \dots, m$ , o mesmo artefato (motor) é medido por todos os laboratórios nas mesmas condições. Segundo, a matriz  $W(\boldsymbol{\theta})$  é aleatória e o modelo é considerado não-ergódico. Como consequência, o processo aleatório correspondente a função score  $U^n(\boldsymbol{\theta})$  é não-ergódico. Além disso os componentes associados com  $\mu_{x_j}$  satisfazem,

$$\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{1}{\sqrt{n}} U_{\mu_{x_j}}^n(\boldsymbol{\theta}) \right| \geq \varepsilon \right] \rightarrow 0, \quad \varepsilon > 0, \quad j = 1, 2, \dots, m. \quad (2.1)$$

Denotaremos por  $\tilde{U}^n(\boldsymbol{\theta})$  e  $\tilde{J}^n(\boldsymbol{\theta})/n$  o vetor score e a matriz de informação observada sem os componentes que envolvem  $\mu_{x_j}$ , respectivamente. Também denotamos por  $\tilde{W}(\boldsymbol{\theta})$  a matriz aleatória  $W(\boldsymbol{\theta})$  sem os componentes que envolvem  $\mu_{x_j}$ . Além disso, denotamos os componentes de viés do vetor de parâmetros por  $\tilde{\boldsymbol{\theta}} = (\alpha_2, \dots, \alpha_p, \beta_2, \dots, \beta_p)^T \in \mathbb{R}^{2(p-1)}$  e  $\hat{\boldsymbol{\theta}}^n$  o EMV correspondente. Além disso, segue-se do apêndice que as matrizes aleatórias  $W(\boldsymbol{\theta})$  e  $\tilde{W}(\boldsymbol{\theta})$  dependem apenas dos componentes de viés do vetor de parâmetros. Como consequência, a partir de agora, denotamos  $W(\boldsymbol{\theta})$  e  $\tilde{W}(\boldsymbol{\theta})$  por  $W(\tilde{\boldsymbol{\theta}})$  e  $\tilde{W}(\tilde{\boldsymbol{\theta}})$ , respectivamente.

Seja  $\{g^n : n \geq 1\}$  uma sequência de funções contínuas reais definidas em um espaço métrico, dizemos que  $g^n(\tau)$  converge uniformemente em  $\tau$  para  $g(\tau)$  se  $g^n(\tau^n) \rightarrow g(\tau)$  para toda sequência  $\tau^n \rightarrow \tau$ . Seja  $\lambda_\tau$  e  $\{\lambda_\tau^n : n \geq 1\}$  probabilidades definidas em subconjuntos de Borel de um espaço métrico, dependendo do parâmetro arbitrário  $\tau$ , e seja  $C$  o espaço das funções uniformemente contínuas e limitadas na reta. Diremos que  $\lambda_\tau^n \Rightarrow_u \lambda_\tau$  uniformemente se

$$\int f d\lambda_\tau^n \rightarrow \int f d\lambda_\tau \quad \text{uniformemente em } \tau, \text{ para todo } f \in C.$$

Se  $Q$  for um espaço métrico e  $\tau \in Q$ , a família  $\lambda_\tau$  de probabilidades será contínua em  $\tau$  se  $\lambda_{\tau^n} \Rightarrow \lambda_\tau$  sempre que  $\tau^n \rightarrow \tau$  em  $Q$ .

Conforme descrito no apêndice, a matriz aleatória  $W(\tilde{\boldsymbol{\theta}})$  é uma função das variáveis não observadas  $(x_1, \dots, x_m)$  e do parâmetro  $\tilde{\boldsymbol{\theta}}$ . E portanto, é definido no espaço de probabilidade  $(\mathbb{R}^m, \beta(\mathbb{R}^m), \mathbb{P}_{(x_1, \dots, x_m)})$ . Denotamos por  $\mathbb{G}_{\tilde{\boldsymbol{\theta}}}$  a distribuição da matriz aleatória  $W(\tilde{\boldsymbol{\theta}})$ .

**Lema 2.** Dadas as sequências  $\{\boldsymbol{\psi}^n : \boldsymbol{\psi}^n \in B_c(n, \boldsymbol{\theta}), n \geq 1\}$  e  $\{\boldsymbol{\phi}^n : \boldsymbol{\phi}^n \in R_c(n, \boldsymbol{\theta}), n \geq 1\}$  e  $g : M^{m+2(p-1)} \rightarrow \mathbb{R}$  uma função contínua limitada, então

$$\mathbb{E}_{\boldsymbol{\psi}^n} \left[ g \left( \frac{1}{n} J^n(\boldsymbol{\phi}^n) \right) \right] = \int g \left( \frac{1}{n} J^n(\boldsymbol{\phi}^n) \right) d\mathbb{P}_{\boldsymbol{\psi}^n} \rightarrow \int g(W(\tilde{\boldsymbol{\theta}})) d\mathbb{P}_{(x_1, \dots, x_m)} = \mathbb{E}_{(x_1, \dots, x_m)} [g(W(\tilde{\boldsymbol{\theta}}))],$$

para todo  $\boldsymbol{\theta} \in \mathbb{R}^{m+2(p-1)}$ . Além disso, a distribuição  $\mathbb{G}_{\tilde{\boldsymbol{\theta}}}$  de  $W(\tilde{\boldsymbol{\theta}})$  é contínuo em  $\tilde{\boldsymbol{\theta}}$ .

*Demonstração.* O fato de que  $\mathbb{G}_{\tilde{\boldsymbol{\theta}}}$  é contínuo segue de [Sweeting \(1980, Lemma 3\)](#). Segue-se do [Lemma 1](#) que  $\frac{1}{n} J^n(\boldsymbol{\phi}^n)$  converge uniformemente em probabilidade para  $W(\tilde{\boldsymbol{\theta}})$ . Como convergência uniforme em probabilidade implica convergência uniforme em distribuição [Sweeting \(1980, Lema 2\)](#), segue-se do [Lema 1 em Sweeting \(1980\)](#) que

$$\int g \left( \frac{1}{n} J^n(\boldsymbol{\phi}^n) \right) d\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)} \rightarrow \int g(W(\tilde{\boldsymbol{\theta}})) d\mathbb{P}_{\tilde{\boldsymbol{\theta}}, (x_1, \dots, x_m)}, \quad (2.2)$$

para qualquer função contínua limitada  $g$ . Como  $J^n(\boldsymbol{\phi}^n)$  depende apenas de  $\boldsymbol{Y}^n$ , concluímos que

$$\int g \left( \frac{1}{n} J^n(\boldsymbol{\phi}^n) \right) d\mathbb{P}_{\boldsymbol{\psi}^n} = \int g \left( \frac{1}{n} J^n(\boldsymbol{\phi}^n) \right) d\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)}.$$

O fato de que  $W(\tilde{\boldsymbol{\theta}})$  depende apenas das variáveis não-observáveis  $(x_1, \dots, x_m)$  faz com que

$$\int g(W(\tilde{\boldsymbol{\theta}})) d\mathbb{P}_{(x_1, \dots, x_m)} = \int g(W(\tilde{\boldsymbol{\theta}})) d\mathbb{P}_{\tilde{\boldsymbol{\theta}}, (x_1, \dots, x_m)}.$$

Como consequência, concluímos o [Lemma](#).  $\square$

Seja  $\boldsymbol{s} \in \mathbb{R}^{m+2(p-1)}$  um vetor e seja  $\{\boldsymbol{\theta}^n : \boldsymbol{\theta}^n \in B_c(n, \boldsymbol{\theta}), n \geq 1\}$  uma sequência de parâmetros. Definimos  $\boldsymbol{\psi}^n = \boldsymbol{\theta}^n + \frac{1}{\sqrt{n}} \boldsymbol{s}$  um vetor em  $\mathbb{R}^{m+2(p-1)}$  tal que  $\boldsymbol{\psi}^n \rightarrow \boldsymbol{\theta}$  quando  $n \rightarrow \infty$ . Como  $L^n(\boldsymbol{\theta}^n)$  é uma função suave, nós podemos escrever

$$\begin{aligned} L^n(\boldsymbol{\psi}^n) &= L^n(\boldsymbol{\theta}^n) + (\boldsymbol{\psi}^n - \boldsymbol{\theta}^n)^T U^n(\boldsymbol{\theta}^n) - \frac{1}{2} (\boldsymbol{\psi}^n - \boldsymbol{\theta}^n)^T J^n(\boldsymbol{\phi}^n) (\boldsymbol{\psi}^n - \boldsymbol{\theta}^n) \\ &= L^n(\boldsymbol{\theta}^n) + \left( \boldsymbol{\theta}^n + \frac{1}{\sqrt{n}} \boldsymbol{s} - \boldsymbol{\theta}^n \right)^T U^n(\boldsymbol{\theta}^n) - \frac{1}{2} \left( \boldsymbol{\theta}^n + \frac{1}{\sqrt{n}} \boldsymbol{s} - \boldsymbol{\theta}^n \right)^T J^n(\boldsymbol{\phi}^n) \left( \boldsymbol{\theta}^n + \frac{1}{\sqrt{n}} \boldsymbol{s} - \boldsymbol{\theta}^n \right) \\ &= L^n(\boldsymbol{\theta}^n) + \frac{1}{\sqrt{n}} \boldsymbol{s}^T U^n(\boldsymbol{\theta}^n) - \frac{1}{2n} \boldsymbol{s}^T J^n(\boldsymbol{\phi}^n) \boldsymbol{s}, \end{aligned} \quad (2.3)$$

em que  $\boldsymbol{\phi}^n = (1 - \delta^n) \boldsymbol{\theta} + \delta^n \boldsymbol{\psi}^n$ ,  $0 < \delta^n < 1$  e  $\delta^n$  é aleatório. Como  $\delta^n$  é uma função do dado observado  $\boldsymbol{Y}^n$  e  $0 < \delta^n < 1$ , concluímos que  $\boldsymbol{\phi}^n \in R_c(n, \boldsymbol{\theta})$ . Como uma consequência, obtemos que  $\boldsymbol{\phi}^n \rightarrow \boldsymbol{\theta}$  quando  $n \rightarrow \infty$ .

**Teorema 1.** Seja  $\{\boldsymbol{\theta}^n : \boldsymbol{\theta}^n \in B_c(n, \boldsymbol{\theta}), n \geq 1\}$  uma sequência de parâmetros e seja  $\{\boldsymbol{h}^n : \boldsymbol{h}^n \in R_c(n, \boldsymbol{\theta}), n \geq 1\}$  uma sequência de vetores aleatórios. Então, temos que

$$\left( \frac{1}{\sqrt{n}} U^n(\boldsymbol{\theta}^n), \frac{1}{n} J^n(\boldsymbol{h}^n) \right) \Rightarrow_u (H(\tilde{\boldsymbol{\theta}}), W(\tilde{\boldsymbol{\theta}}))$$

em que  $H(\tilde{\boldsymbol{\theta}}) = (\mathbf{0}_m^T, ((\tilde{W}(\tilde{\boldsymbol{\theta}}))^{1/2} \boldsymbol{z}))^T$  tal que

- $\mathbf{0}_m$  é o vetor nulo de dimensão  $m$ ;
- $\mathbf{z}$  é um vetor aleatório normal padrão em  $\mathbb{R}^{2(p-1)}$ , independente da matriz aleatória  $\tilde{W}(\tilde{\boldsymbol{\theta}})$ .

Além disso, a matriz aleatória  $\tilde{W}(\tilde{\boldsymbol{\theta}})$  é positiva definida com probabilidade 1 e depende apenas dos componentes de viés  $\tilde{\boldsymbol{\theta}}$  e das variáveis não-observáveis  $(x_1, \dots, x_m)$ .

*Demonstração.* Tomamos exponenciais na equação (2.3) e reorganizando temos:

$$\exp \left[ \frac{1}{2n} \mathbf{s}^T \mathbf{J}^n(\boldsymbol{\phi}^n) \mathbf{s} \right] f_{\mathbf{Y}^n}(\mathbf{y}^n, \boldsymbol{\psi}^n) = \exp \left[ \frac{1}{\sqrt{n}} \mathbf{s}^T \mathbf{U}^n(\boldsymbol{\theta}^n) \right] f_{\mathbf{Y}^n}(\mathbf{y}^n, \boldsymbol{\theta}^n). \quad (2.4)$$

Seja  $0 < \varepsilon < 1$  e escolhamos uma constante positiva  $v$  tal que  $\mathbb{P}_{(x_1, \dots, x_m)}[\|W(\tilde{\boldsymbol{\theta}})\| \geq v] \leq \varepsilon$ . Como uma consequência do Lemma 2, temos que  $(1/n)\mathbf{J}^n(\boldsymbol{\phi}^n)$  converge uniformemente em distribuição para  $W(\tilde{\boldsymbol{\theta}})$  sob a família de probabilidades  $\{\mathbb{P}_{\boldsymbol{\psi}^n} : n \geq 1\}$  e  $\{\mathbb{P}_{\boldsymbol{\theta}^n} : n \geq 1\}$ . Como  $\{A \in M^{m+2(p-1)} : \|A\| < v\}$  é um conjunto  $\mathbb{G}_{\boldsymbol{\theta}}$ -contínuo, segue de Lemmas 1 e 3 em Sweeting (1980) que

$$\mathbb{P}_{\boldsymbol{\theta}^n} \left[ \left\| \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \right\| < v \right] \rightarrow \mathbb{P}_{(x_1, \dots, x_m)} [\|W(\tilde{\boldsymbol{\theta}})\| < v]. \quad (2.5)$$

Seja  $\mathbb{Q}_{\boldsymbol{\theta}^n}$  a probabilidade  $\mathbb{P}_{\boldsymbol{\theta}^n}$  condicionada em  $\{\| \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \| < v\}$ . Neste caso, o componente finito dimensional de  $\mathbb{Q}_{\boldsymbol{\theta}^n}$  tem a seguinte densidade

$$q_{\mathbf{Y}^n}(\mathbf{y}^n, \boldsymbol{\theta}^n) = \begin{cases} \frac{f_{\mathbf{Y}^n}(\mathbf{y}^n, \boldsymbol{\theta}^n)}{\mathbb{P}_{\boldsymbol{\theta}^n}[\| \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \| < v]}, & \| \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \| < v \\ 0, & \text{caso contrario.} \end{cases}$$

Seja  $g$  uma função limitada em  $M^{m+2(p-1)}$ , contínua em  $\|A\| < v$  e com  $g(A) = 0$  para todo  $\|A\| \geq v$ . Seja  $\mathbb{E}_{\boldsymbol{\theta}^n}^*$  a esperança em relação a probabilidade  $\mathbb{Q}_{\boldsymbol{\theta}^n}$  e  $\mathbb{E}_{(x_1, \dots, x_m)}^*$  a esperança com respeito a probabilidade  $\mathbb{P}_{(x_1, \dots, x_m)}$  condicionada no conjunto  $\{\|W(\tilde{\boldsymbol{\theta}})\| < v\}$ . Multiplicando equação (2.4) por  $g((1/n)\mathbf{J}^n(\boldsymbol{\phi}^n))$  e integrando com relação à medida de Lebesgue no conjunto  $\{\| (1/n)\mathbf{J}^n(\boldsymbol{\phi}^n) \| < v\}$  temos:

$$\begin{aligned} \mathbb{E}_{\boldsymbol{\theta}^n}^* \left[ g \left( \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \right) \exp \left( \frac{1}{\sqrt{n}} \mathbf{s}^T \mathbf{U}^n(\boldsymbol{\theta}^n) \right) \right] &= \frac{\mathbb{E}_{\boldsymbol{\psi}^n} \left[ g \left( \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \right) \exp \left( \frac{1}{2n} \mathbf{s}^T \mathbf{J}^n(\boldsymbol{\phi}^n) \mathbf{s} \right) \right]}{\mathbb{P}_{\boldsymbol{\theta}^n} [\| \frac{1}{n} \mathbf{J}^n(\boldsymbol{\phi}^n) \| < v]} \\ &\rightarrow \frac{\mathbb{E}_{(x_1, \dots, x_m)} [g(W(\tilde{\boldsymbol{\theta}})) \exp \left( \frac{1}{2} \mathbf{s}^T W(\tilde{\boldsymbol{\theta}}) \mathbf{s} \right)]}{\mathbb{P}_{(x_1, \dots, x_m)} [\|W(\tilde{\boldsymbol{\theta}})\| < v]} \\ &= \mathbb{E}_{(x_1, \dots, x_m)}^* \left[ g(W(\tilde{\boldsymbol{\theta}})) \exp \left( \frac{1}{2} \mathbf{s}^T W(\tilde{\boldsymbol{\theta}}) \mathbf{s} \right) \right], \end{aligned}$$

como uma consequência da equação (2.5), Lemma 2 e o fato de que  $g(A) \exp((1/2)\mathbf{s}^T A \mathbf{s})$  é uma função  $\mathbb{G}_{\boldsymbol{\theta}}$ -contínua limitada (Veja, Billingsley (1968), Theorem 5.2).

Decompondo o vetor  $\mathbf{s}$  em  $\mathbf{s} = (\mathbf{s}_1^T, \mathbf{s}_2^T)^T$  em que  $\mathbf{s}_1 = (s_1, \dots, s_m)^T$  e  $\mathbf{s}_2 = (s_{m+1}, \dots, s_{m+2(p-1)})^T$ , como todo componente de  $W(\tilde{\boldsymbol{\theta}})$  associado com  $\mu_{x_j}$  é nulo, obtivemos que

$$\mathbb{E}_{(x_1, \dots, x_m)}^* \left[ g(W(\tilde{\boldsymbol{\theta}})) \exp \left( \frac{1}{2} \mathbf{s}^T W(\tilde{\boldsymbol{\theta}}) \mathbf{s} \right) \right] = \mathbb{E}_{(x_1, \dots, x_m)}^* \left[ g(W(\tilde{\boldsymbol{\theta}})) \exp \left( \mathbf{s}_1^T \mathbf{0}_m + \frac{1}{2} \mathbf{s}_2^T \tilde{W}(\tilde{\boldsymbol{\theta}}) \mathbf{s}_2 \right) \right] =$$

$$\mathbb{E}_{(x_1, \dots, x_m)}^* \left[ g(W(\tilde{\boldsymbol{\theta}})) \exp(\mathbf{s}^T H(\tilde{\boldsymbol{\theta}})) \right],$$

tal que  $H(\tilde{\boldsymbol{\theta}}) = (\mathbf{0}_m, (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{1/2} \mathbf{z})$  em que  $\mathbf{0}_m$  é o vetor nulo de dimensão  $m$  e  $\mathbf{z}$  é um vetor aleatório normal padrão em  $\mathbb{R}^{2(p-1)}$ , independente da matriz aleatória  $\tilde{W}(\tilde{\boldsymbol{\theta}})$ . Neste paragrafo em particular  $\mathbb{1}$  é a função indicadora. Pela unicidade da função geradora de momento e pelo teorema da compacidade fraca, concluímos que

$$\left( \frac{1}{n} J^n(\boldsymbol{\phi}^n), \frac{1}{\sqrt{n}} U^n(\boldsymbol{\theta}^n) \right) \Rightarrow (W(\tilde{\boldsymbol{\theta}}), H(\tilde{\boldsymbol{\theta}})) \mathbb{1}_{\{\|W(\tilde{\boldsymbol{\theta}})\| < v\}}$$

sob a família  $\{\mathbb{Q}_{\boldsymbol{\theta}^n} : n \geq 1\}$  de probabilidades. Como  $v$  é arbitrário, segue que

$$\left( \frac{1}{n} J^n(\boldsymbol{\phi}^n), \frac{1}{\sqrt{n}} U^n(\boldsymbol{\theta}^n) \right) \Rightarrow (W(\tilde{\boldsymbol{\theta}}), H(\tilde{\boldsymbol{\theta}}))$$

sob a família  $\{\mathbb{P}_{\boldsymbol{\theta}^n} : n \geq 1\}$  de probabilidades. Aplicando o Lema 1, concluímos que

$$\left\| \frac{1}{n} J^n(\boldsymbol{\phi}^n) - \frac{1}{n} J^n(\mathbf{h}^n) \right\| \rightarrow_u 0$$

uniformemente em probabilidade. Portanto, obtemos que

$$\left( \frac{1}{n} J^n(\mathbf{h}^n), \frac{1}{\sqrt{n}} U^n(\boldsymbol{\theta}^n) \right) \Rightarrow_u (W(\tilde{\boldsymbol{\theta}}), H(\tilde{\boldsymbol{\theta}})).$$

□

Na sequência, mostraremos a normalidade assintótica do EMV em relação aos parâmetros de viés  $\tilde{\boldsymbol{\theta}}$ . Para cada vetor  $\mathbf{s} = (\mathbf{s}_1, \mathbf{s}_2) \in \mathbb{R}^{m+2(p-1)}$  tal que  $\mathbf{s}_1 = (s_1, \dots, s_m)$  e  $\mathbf{s}_2 = (s_{m+1}, \dots, s_{m+2(p-1)})$ , segue da equação (2.3) que

$$\begin{aligned} L^n(\boldsymbol{\psi}^n) - L^n(\boldsymbol{\theta}) &= \frac{1}{\sqrt{n}} \mathbf{s}^T U^n(\boldsymbol{\theta}) - \frac{1}{2n} \mathbf{s}^T J^n(\boldsymbol{\phi}^n) \mathbf{s} = \frac{1}{\sqrt{n}} \mathbf{s}_2^T \tilde{U}^n(\boldsymbol{\theta}) - \frac{1}{2n} \mathbf{s}_2^T \tilde{J}^n(\boldsymbol{\theta}) \mathbf{s}_2 \\ &+ \frac{1}{\sqrt{n}} \mathbf{s}_1^T U_{\mu_{x_j}}^n(\boldsymbol{\theta}) - \frac{1}{2n} \sum_{i=1}^m s_i^2 J_{ii}^n(\boldsymbol{\phi}^n) - \frac{1}{2n} \sum_{i=1}^m \sum_{j=m+1}^{m+2(p-1)} s_i s_j J_{ij}^n(\boldsymbol{\phi}^n) \\ &- \frac{1}{2n} \sum_{i=m+1}^{m+2(p-1)} \sum_{j=1}^m s_i s_j J_{ij}^n(\boldsymbol{\phi}^n) + \frac{1}{2n} \mathbf{s}_2^T [\tilde{J}^n(\boldsymbol{\theta}) - \tilde{J}^n(\boldsymbol{\phi}^n)] \mathbf{s}_2, \end{aligned}$$

em que  $\boldsymbol{\theta} \in \mathbb{R}^{m+2(p-1)}$ ,  $\boldsymbol{\psi}^n = \boldsymbol{\theta} + \frac{1}{\sqrt{n}} \mathbf{s}$  e  $\boldsymbol{\phi}^n = \delta^n \boldsymbol{\theta} + (1 - \delta^n) \boldsymbol{\psi}^n$ ,  $0 < \delta^n < 1$  tal que  $\delta^n$  é aleatório. Como uma consequência do Lema 1 e da equação (2.1), chegamos ao lema seguinte.

**Lema 3.** Temos que

$$\sup_{\mathbf{s} \in \mathbb{R}^{m+2(p-1)}} [L^n(\boldsymbol{\psi}^n) - L^n(\boldsymbol{\theta})] = \sup_{\mathbf{s}_2 \in \mathbb{R}^{2(p-1)}} \left[ \frac{1}{\sqrt{n}} \mathbf{s}_2^T \tilde{U}^n(\boldsymbol{\theta}) - \frac{1}{2n} \mathbf{s}_2^T \tilde{J}^n(\boldsymbol{\theta}) \mathbf{s}_2 \right] + o_p(1), \quad (2.6)$$

para  $\boldsymbol{\theta} \in \mathbb{R}^{m+2(p-1)}$ .

Esse lema é importante para entender o comportamento da função de probabilidade em relação aos parâmetros de valor real  $(\mu_{x_1}, \dots, \mu_{x_m})$ . Para  $n$  suficientemente grande, o impacto dos verdadeiros valores desaparece. Além disso, o máximo em relação a  $\mathbf{s}_2$  do lado direito da equação (2.6) satisfaz

$$\frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \hat{\mathbf{s}}_2 = \frac{1}{\sqrt{n}} \tilde{U}^n(\boldsymbol{\theta}) + o_p(1). \quad (2.7)$$

Aplicando a equação (2.6), para  $n$  suficientemente grande, concluímos que  $(\mathbf{s}_1^T, \hat{\mathbf{s}}_2^T)^T$  corresponde ao valor de  $\hat{\mathbf{s}}$  que maximiza  $L^n(\boldsymbol{\theta} + n^{-1/2}(\mathbf{s}_1^T, \hat{\mathbf{s}}_2^T)^T)$  independente do vetor  $\mathbf{s}_1$ . Por construção, o máximo de  $L^n(\boldsymbol{\theta} + n^{-1/2}(\mathbf{s}_1^T, \hat{\mathbf{s}}_2^T)^T)$  é atingido próximo ao EMV  $\hat{\boldsymbol{\theta}}^n$ . Então, obtemos que

$$\hat{\boldsymbol{\theta}}^n = \tilde{\boldsymbol{\theta}} + \frac{1}{\sqrt{n}} \hat{\mathbf{s}}_2 + o_p(1) \quad \text{e} \quad \sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) = \hat{\mathbf{s}}_2 + o_p(1),$$

em que  $\tilde{\boldsymbol{\theta}}^n$  corresponde ao EMV dos parâmetros de viés  $\tilde{\boldsymbol{\theta}}$ . Como consequência, concluímos que

$$\left( \frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \right) \sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) = \frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \hat{\mathbf{s}}_2 + o_p(1). \quad (2.8)$$

Resumindo os resultados obtidos das equações (2.7) e (2.8), chegamos ao seguinte teorema.

**Teorema 2.** O EMV  $\hat{\boldsymbol{\theta}}^n$  de  $\tilde{\boldsymbol{\theta}}$  satisfaz

$$\frac{1}{\sqrt{n}} \tilde{U}^n(\boldsymbol{\theta}) - \left( \frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \right) \sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) \rightarrow_u 0,$$

uniformemente em probabilidade, em que  $\boldsymbol{\theta} = (\mu_{x_1}, \dots, \mu_{x_m}, \tilde{\boldsymbol{\theta}}) \in \mathbb{R}^{m+2(p-1)}$ .

Como uma consequência dos Teoremas 1 e 2 e o teorema do mapeamento contínuo, concluímos que

$$\left( \left[ \frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \right]^{1/2} \sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}), \frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \right) \Rightarrow_u (\mathbf{z}, \tilde{W}(\tilde{\boldsymbol{\theta}})). \quad (2.9)$$

De (2.9) e o teorema do mapeamento contínuo, temos:

$$\sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) \Rightarrow_u [\tilde{W}(\tilde{\boldsymbol{\theta}})]^{-1/2} \mathbf{z}. \quad (2.10)$$

Da equação (2.10), sabemos que a distribuição assintótica do EMV em relação aos parâmetros de viés  $\tilde{\boldsymbol{\theta}}$  não é normal, porque a matriz  $\tilde{W}(\tilde{\boldsymbol{\theta}})$  é aleatória.

**Corolário 1.** O EMV  $\hat{\boldsymbol{\theta}}^n$  relacionados aos parâmetros de viés satisfaz

$$\left(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}\right) \rightarrow_u 0.$$

Na sequência, derivaremos as estatísticas usuais de Wald para realizar testes de hipóteses sobre os parâmetros de viés. Aplicando o Teorema de Prohorov, obtemos que a sequência  $\{\sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) : n \geq 1\}$  é uniformemente “tight”. Então, para cada  $\varepsilon > 0$ , existe uma constante  $c > 0$  tal que

$$\mathbb{P}_{\boldsymbol{\theta}^n, (x_1, \dots, x_m)} \left[ \left| \sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) \right| > c \right] < \varepsilon, \quad n \geq 1.$$

Como consequência, com probabilidade tendendo a um,  $\hat{\boldsymbol{\theta}}^n \in R_c(n, \boldsymbol{\theta})$ . Assim, chegamos aos seguintes Corolários.

**Corolário 2.** Condicional em  $\tilde{J}^n(\hat{\boldsymbol{\theta}}^n)$ , a distribuição assintótica de  $\sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}})$  é dado por

$$N_{2(p-1)} \left[ \mathbf{0}_{2(p-1)}, \left( \frac{1}{n} \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} \right].$$

Aplicando novamente a equação (2.9) e o Teorema do mapeamento contínuo, chegamos às estatísticas de Wald.

**Corolário 3.** Temos que

$$\left(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}\right)^T \left[ \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right] \left(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}\right) \Rightarrow_u \mathbf{z}^T \mathbf{z}.$$

onde os componentes de viés do vetor de parâmetros por  $\tilde{\boldsymbol{\theta}} = (\alpha_2, \dots, \alpha_p, \beta_2, \dots, \beta_p)^T \in \mathbb{R}^{2(p-1)}$  e  $\hat{\boldsymbol{\theta}}^n$  o EMV correspondente; em que  $\boldsymbol{\theta} = (\mu_{x_1}, \dots, \mu_{x_m}, \tilde{\boldsymbol{\theta}}) \in \mathbb{R}^{m+2(p-1)}$  e  $\hat{\boldsymbol{\theta}}^n$  o EMV correspondente.

## 2.3 Equivalência entre laboratórios participantes

### 2.3.1 Teste de hipótese

Nesta seção, proporemos o teste Wald assintótico para avaliar a equivalência entre as medições dos laboratórios em relação ao laboratório de referência. Inicialmente, testaremos a equivalência de todos os laboratórios em relação ao laboratório de referência,

$$H_0 : \alpha_2 = \dots = \alpha_p = 0 \quad \text{e} \quad \beta_2 = \dots = \beta_p = 1. \quad (2.11)$$

Para testar a hipótese (2.11), podemos aplicar a estatística de Wald, utilizando o corolário estabelecido em Corolário 3.

Seja:

$$Q_w = \left(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}_0\right)^T \left[ \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right] \left(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}_0\right) \quad (2.12)$$

sob as condições do Corolário 3,  $Q_w$  possui distribuição assintótica  $\chi_{2(p-1)}^2$  sob  $H_0$ .

Para testar individualmente cada laboratório,

$$H_0 : \alpha_i = 0 \quad \text{e} \quad \beta_i = 1, \quad i = 2, \dots, p, \quad (2.13)$$

desenvolveremos a seguir os testes de hipóteses composto

$$H_0 : h(\tilde{\boldsymbol{\theta}}) = 0, \quad (2.14)$$

em que  $h : \mathbb{R}^{2(p-1)} \rightarrow \mathbb{R}^r$  é uma função com valor-vetorial, tal que a matriz das derivadas  $H(\tilde{\boldsymbol{\theta}}) = (\partial/\partial\tilde{\boldsymbol{\theta}})h(\tilde{\boldsymbol{\theta}})^T$  é contínua em  $\tilde{\boldsymbol{\theta}}$  e o  $\text{posto}(H(\tilde{\boldsymbol{\theta}})) = r$ . Para desenvolver esses testes compostos, considere a expansão de Taylor

$$h(\tilde{\boldsymbol{\theta}} + n^{-1/2}\mathbf{u}) = h(\tilde{\boldsymbol{\theta}}) + n^{-1/2}H^T(\tilde{\boldsymbol{\theta}}^*)\mathbf{u} = h(\tilde{\boldsymbol{\theta}}) + n^{-1/2}H^T(\tilde{\boldsymbol{\theta}})\mathbf{u} + n^{-1/2} \left[ H^T(\tilde{\boldsymbol{\theta}}^*) - H^T(\tilde{\boldsymbol{\theta}}) \right] \mathbf{u},$$

em que  $\tilde{\boldsymbol{\theta}}^* = \tilde{\boldsymbol{\theta}} + n^{-1/2}\gamma\mathbf{u}$ ,  $0 < \gamma < 1$  e  $\mathbf{u} \in \mathbb{R}^{2(p-1)}$ . Como consequência da suposição de continuidade de  $H(\tilde{\boldsymbol{\theta}})$ , chegamos à seguinte expressão

$$\sqrt{n} \left[ h(\tilde{\boldsymbol{\theta}} + n^{-1/2}\mathbf{u}) - h(\tilde{\boldsymbol{\theta}}) \right] = H^T(\tilde{\boldsymbol{\theta}})\mathbf{u} + n^{-1/2} \left[ H^T(\tilde{\boldsymbol{\theta}}^*) - H^T(\tilde{\boldsymbol{\theta}}) \right] \mathbf{u}.$$

Seja  $\mathbf{u} = \sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}})$ , obtivemos

$$\sqrt{n} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] - H^T(\tilde{\boldsymbol{\theta}})\sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}) = o_p(1). \quad (2.15)$$

**Teorema 3.** Temos que

$$\left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \left[ H^T(\hat{\boldsymbol{\theta}}^n) \left( \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} H(\hat{\boldsymbol{\theta}}^n) \right]^{-1} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] \Rightarrow_u \mathbf{z}_r^T \mathbf{z}_r.$$

*Demonstração.* Aplicando a equação (2.9), temos que

$$\left( H(\tilde{\boldsymbol{\theta}})\sqrt{n}(\hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}}), \frac{1}{n}\tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right) \Rightarrow_u \left( H(\tilde{\boldsymbol{\theta}}) (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{-1/2} \mathbf{z}, \tilde{W}(\tilde{\boldsymbol{\theta}}) \right).$$

Como a distribuição de  $\mathbf{z}$  é independente de  $\tilde{W}(\tilde{\boldsymbol{\theta}})$ , obtivemos que

$$\begin{aligned} & \mathbb{P}_{(x_1, \dots, x_m)} \left[ H(\tilde{\boldsymbol{\theta}}) (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{-1/2} \mathbf{z} \in D \mid \tilde{W}(\tilde{\boldsymbol{\theta}}) \right] \\ &= \mathbb{P}_{(x_1, \dots, x_m)} \left[ \left[ H^T(\tilde{\boldsymbol{\theta}}) (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{-1} H(\tilde{\boldsymbol{\theta}}) \right]^{-1/2} \mathbf{z}_r \in D \mid \tilde{W}(\tilde{\boldsymbol{\theta}}) \right], \end{aligned}$$

para cada  $D \in \beta(\mathbb{R}^r)$ , onde  $\mathbf{z}_r$  é um vetor aleatório normal padrão em  $\mathbb{R}^r$ , independente da matriz aleatória  $\tilde{W}(\tilde{\boldsymbol{\theta}})$ . Então, os vetores aleatórios

$$H(\tilde{\boldsymbol{\theta}}) (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{-1/2} \mathbf{z} \quad \text{e} \quad \left[ H^T(\tilde{\boldsymbol{\theta}}) (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{-1} H(\tilde{\boldsymbol{\theta}}) \right]^{-1/2} \mathbf{z}_r$$

têm a mesma distribuição. Como consequência, obtemos que

$$\left( H(\tilde{\boldsymbol{\theta}}) \sqrt{n} \left( \hat{\boldsymbol{\theta}}^n - \tilde{\boldsymbol{\theta}} \right), \frac{1}{n} \tilde{J}^n(\boldsymbol{\theta}) \right) \Rightarrow_u \left( \left[ H^T(\tilde{\boldsymbol{\theta}}) (\tilde{W}(\tilde{\boldsymbol{\theta}}))^{-1} H(\tilde{\boldsymbol{\theta}}) \right]^{1/2} \mathbf{z}_r, \tilde{W}(\tilde{\boldsymbol{\theta}}) \right).$$

Aplicando Corolário 1 e a Equação (2.15), concluímos que

$$\left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \left[ H^T(\hat{\boldsymbol{\theta}}^n) \left( \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} H(\hat{\boldsymbol{\theta}}^n) \right]^{-1} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] \Rightarrow_u \mathbf{z}_r^T \mathbf{z}_r.$$

□

Desta forma, seja:

$$Q_w = \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \left[ H^T(\hat{\boldsymbol{\theta}}^n) \left( \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} H(\hat{\boldsymbol{\theta}}^n) \right]^{-1} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right], \quad (2.16)$$

sob  $H_0$ ,  $Q_w$  possui distribuição assintótica  $\chi_2^2$ .

**Corolário 4.** Sob as condições do teorema (3) a estatística

$$Q_w = \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \left[ H^T(\hat{\boldsymbol{\theta}}^n) \left( \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} H(\hat{\boldsymbol{\theta}}^n) \right]^{-1} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right], \quad (2.17)$$

possui distribuição assintótica  $\chi_2^2$ .

### 2.3.2 Região de Confiança Conjunta

Uma outra forma de avaliar se os laboratórios estão conformes é utilizando regiões de confiança para  $\alpha_i$  e  $\beta_i$ ,  $i = 1, \dots, p$ . Seja,

$$\mathbf{V} = \left[ H^T(\hat{\boldsymbol{\theta}}^n) \left( \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} H(\hat{\boldsymbol{\theta}}^n) \right]^{-1} = \begin{pmatrix} V_{\alpha_i \alpha_i} & V_{\alpha_i \beta_i} \\ V_{\beta_i \alpha_i} & V_{\beta_i \beta_i} \end{pmatrix}_{2 \times 2} = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}. \text{ Do teorema}$$

(3), temos que:

$$\left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \left[ H^T(\hat{\boldsymbol{\theta}}^n) \left( \tilde{J}^n(\hat{\boldsymbol{\theta}}^n) \right)^{-1} H(\hat{\boldsymbol{\theta}}^n) \right]^{-1} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] \sim \mathbf{z}_r^T \mathbf{z}_r,$$

$$\left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \mathbf{V} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] \sim \chi_2^2,$$

e isso implica que:

$$\mathbb{P} \left\{ \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \mathbf{V} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] \leq \chi_{1-\alpha, 2}^2 \right\} = 1 - \alpha.$$

Em consequência, uma região de confiança, para o parâmetro  $(\alpha_i, \beta_i)$ ,  $i = 1, \dots, p$  é dado por:

$$\left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right]^T \mathbf{V} \left[ h(\hat{\boldsymbol{\theta}}^n) - h(\tilde{\boldsymbol{\theta}}) \right] \leq \chi_{1-\alpha, 2}^2, \quad (2.18)$$

$$\left[ \begin{pmatrix} \alpha_i - \hat{\alpha}_i \\ \beta_i - \hat{\beta}_i \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right]^T \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} \left[ \begin{pmatrix} \alpha_i - \hat{\alpha}_i \\ \beta_i - \hat{\beta}_i \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right] \leq \chi_{1-\alpha,2}^2.$$

Após algumas manipulações obtemos:

$$V_{11}\alpha_i^2 + V_{22}\beta_i^2 + r\alpha_i\beta_i + (-2\hat{\alpha}_iV_{11} - \hat{\beta}_i r)\alpha_i + (-2\hat{\beta}_iV_{22} - \hat{\alpha}_i r)\beta_i \\ + (\hat{\alpha}_i^2V_{11} + \hat{\beta}_i^2V_{22} + r\hat{\alpha}_i\hat{\beta}_i - \chi_{\alpha,2}^2) \leq 0$$

Seja  $r = V_{12} + V_{21}$ , então segundo a equação geral da elipse (C.2) temos que os elementos da elipse são dados por:  $Ax^2 + By^2 + Cxy + Dx + Ey + F = 0$  é  $A = V_{11}$ ,  $B = V_{22}$ ,  $C = r$ ,  $D = (-2\hat{\alpha}_iV_{11} - \hat{\beta}_i r)$ ,  $E = (-2\hat{\beta}_iV_{22} - \hat{\alpha}_i r)$ , e  $F = (\hat{\alpha}_i^2V_{11} + \hat{\beta}_i^2V_{22} + r\hat{\alpha}_i\hat{\beta}_i - \chi_{1-\alpha,2}^2)$ .

Desta forma, podemos construir uma região de confiança para  $(\alpha_i, \beta_i)$ ,  $i = 1, \dots, p$ . Para podermos aplicar o coeficiente de confiança simultaneamente para todos os conjuntos de regiões de confiança, dividimos  $\alpha$  pelo número de regiões de confiança que serão utilizados para avaliar quais laboratórios estão conformes. De tal forma que a taxa de erro de família de testes é menor do que  $\alpha$ .

## 2.4 Estudos de simulação

Apresentaremos, nesta seção, estudos de simulação para verificarmos o comportamento do teste assintóticos de Wald obtido em (2.12) e (2.17) em relação ao número de réplicas e os desvios dos erros de medição de cada laboratório  $i$  e em cada patamar  $j$ ,  $i = 1, \dots, p$ ;  $j = 1, \dots, m$ . Considerando o modelo definido em (1.14) e (1.15) foram geradas 10 mil amostras com 3, 7, 15 e 30 réplicas. Em todos os casos simulados utilizou-se  $p = 5$  (número de laboratórios) e  $m = 5$  (número de patamares). Consideramos para a média do verdadeiro valor do mensurando em cada patamar ( $\mu_{x_j}$ ), os valores  $\mu_{x_1} = 10$ ,  $\mu_{x_2} = 20$ ,  $\mu_{x_3} = 30$ ,  $\mu_{x_4} = 40$ ,  $\mu_{x_5} = 50$ , e o desvio do verdadeiro valor não observado do mensurando em cada patamar ( $\sigma_{x_j}$ ) os valores  $\sigma_{x_1} = 0.24$ ,  $\sigma_{x_2} = 0.31$ ,  $\sigma_{x_3} = 0.38$ ,  $\sigma_{x_4} = 0.45$ ,  $\sigma_{x_5} = 0.52$ . Para os desvios dos erros de medição de cada laboratório  $i$  em cada patamar  $j$  ( $\sigma_{ij}$ ) foram considerados os seguintes conjuntos de valores:

1.  $\sigma_{ij}^a$ :  $\sigma_{i1} = 0.1$ ,  $\sigma_{i2} = 0.2$ ,  $\sigma_{i3} = 0.3$ ,  $\sigma_{i4} = 0.4$ ,  $\sigma_{i5} = 0.5$
2.  $\sigma_{ij}^b$ :  $\sigma_{i1} = 0.2$ ,  $\sigma_{i2} = 0.4$ ,  $\sigma_{i3} = 0.6$ ,  $\sigma_{i4} = 0.8$ ,  $\sigma_{i5} = 1.0$
3.  $\sigma_{ij}^c$ :  $\sigma_{i1} = 0.3$ ,  $\sigma_{i2} = 0.6$ ,  $\sigma_{i3} = 0.9$ ,  $\sigma_{i4} = 1.2$ ,  $\sigma_{i5} = 1.5$
4.  $\sigma_{ij}^d$ :  $\sigma_{i1} = 0.5$ ,  $\sigma_{i2} = 1.0$ ,  $\sigma_{i3} = 1.5$ ,  $\sigma_{i4} = 2.0$ ,  $\sigma_{i5} = 2.5$

Primeiramente foi feito um estudo em relação ao nível de significância do teste, e depois foi feito um estudo do poder do teste.

### 2.4.1 Simulação do tamanho dos testes assintóticos

No estudo de simulação foram geradas 10 mil amostras e obtidas as porcentagens dos valores das estatísticas que foram maiores do que os quantis de 1%, 5% e 10% de uma distribuição Chi-quadrado com 8(2) graus de liberdade para o teste conjunto (teste individual). O objetivo destas simulações é avaliar numericamente o comportamento das estatísticas quanto ao tamanho das réplicas e quanto aos diferentes valores considerados para os parâmetros. Estes resultados estão apresentados nas Tabelas 1 e 2.

Tabela 1 –  $H_0 : \alpha_2 = \dots = \alpha_5 = 0, \beta_2 = \dots = \beta_5 = 1$  - tamanho empírico.

$n_i$	$\sigma_{ij}^a$			$\sigma_{ij}^b$			$\sigma_{ij}^c$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%
3	0.012	0.059	0.114	0.023	0.084	0.15	0.043	0.126	0.202
7	0.011	0.053	0.106	0.015	0.065	0.127	0.019	0.076	0.140
15	0.011	0.053	0.102	0.011	0.058	0.109	0.017	0.068	0.124
30	0.010	0.053	0.107	0.011	0.053	0.102	0.012	0.056	0.110

Tabela 2 –  $H_0 : \alpha_i = 0, \beta_i = 1, i = 2, \dots, 5$  - tamanho empírico.

$n_i$	$\sigma_{ij}^a$			$\sigma_{ij}^b$			$\sigma_{ij}^c$			$\sigma_{ij}^d$		
	1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
3	0.016	0.065	0.126	0.024	0.081	0.147	0.035	0.114	0.189	0.056	0.154	0.234
7	0.010	0.051	0.101	0.017	0.070	0.129	0.023	0.088	0.151	0.038	0.121	0.201
15	0.010	0.051	0.101	0.013	0.061	0.113	0.016	0.068	0.126	0.024	0.089	0.159
30	0.008	0.048	0.102	0.012	0.053	0.102	0.013	0.063	0.120	0.018	0.074	0.135

Em todos os casos (Tabela 1 e 2) podemos observar que conforme o número de réplicas aumenta a porcentagem de rejeição se aproxima do nível nominal. Além disso, conforme aumenta a variância dos erros de medição dos laboratórios ( $\sigma_{ij}^2$ ), as porcentagens de rejeição se distanciam do nível nominal, ou seja, fixados  $\mu_{x_j}$ ,  $\sigma_{x_j}$  e o tamanho da réplica, o comportamento da estatística apresentou tendência a melhorar quando diminui o valor de  $\sigma_{ij}^2$  (variância do erro de medição do valor observado no  $i$ -ésimo laboratório no  $j$ -ésimo patamar).

### 2.4.2 Simulação do poder dos testes assintóticos

Além das simulações do tamanho do teste, realizamos simulações para verificar o comportamento da estatística de Wald quanto ao poder do teste. Para avaliar a estatística neste sentido, foram geradas 10 mil amostras, sob as hipóteses alternativas  $H_1$ .

#### 2.4.2.1 $H_0 : \alpha_2 = \dots = \alpha_5 = 0$ e $\beta_2 = \dots = \beta_5 = 1$

Considerando os valores dos parâmetros descritos no início desta seção e o conjunto de valores definidos como  $\sigma_{ij}^a$  e  $\sigma_{ij}^b$ , foram calculados as porcentagens de rejeição comparando-se com o quantil de 95% de uma  $\chi_8^2$ . Neste caso, adotou-se um distanciamento gradual nos

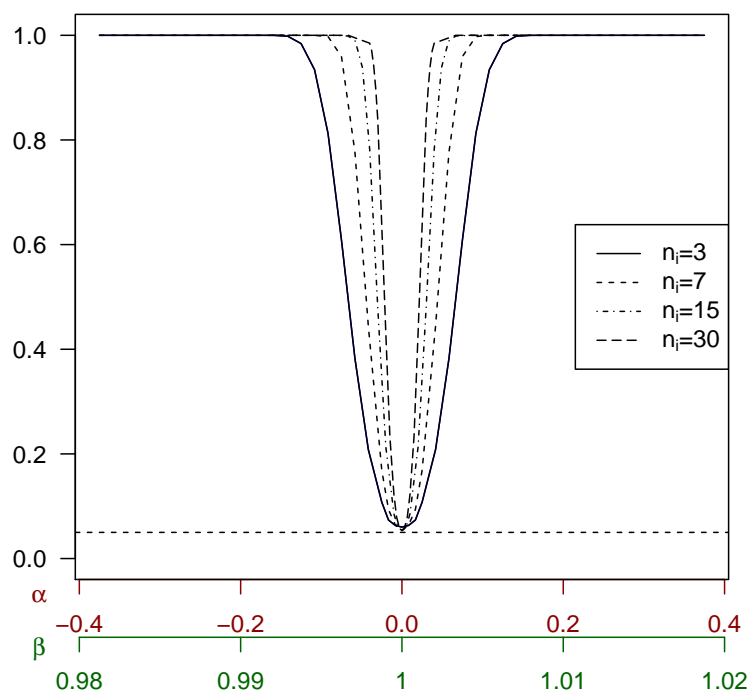


Figura 1 –  $H_0 : \alpha_2 = \dots = \alpha_5 = 0$  e  $\beta_2 = \dots = \beta_5 = 1$  - poder do teste com  $\sigma_{ij}^a$

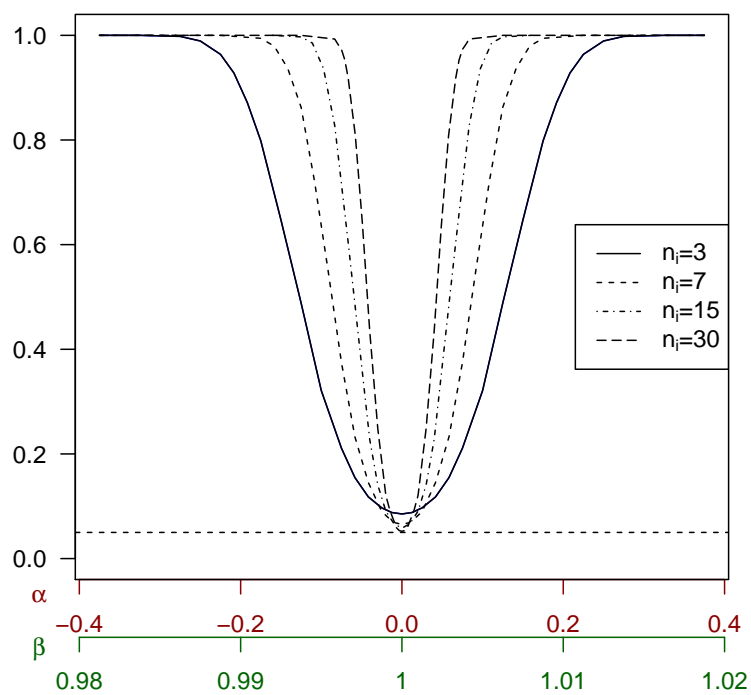


Figura 2 –  $H_0 : \alpha_2 = \dots = \alpha_5 = 0$  e  $\beta_2 = \dots = \beta_5 = 1$  - poder do teste com  $\sigma_{ij}^b$ .

valores dos parâmetros de viés aditivo e multiplicativo, partindo-se de  $H_0 : \alpha_2 = \dots = \alpha_5 = 0$  e  $\beta_2 = \dots = \beta_5 = 1$ , como pode ser visto nas Figuras 1, 2 e 3. Os parâmetros que sofreram este distanciamento foram:  $\alpha_2$ ,  $\beta_2$ ,  $\alpha_4$  e  $\beta_4$ .

Na Figura 1(2) foi construído o gráfico do poder do teste considerando 3, 7, 15 e 30 réplicas com o desvio padrão dos erros de medição dos laboratórios dados por  $\sigma_{ij}^a$  ( $\sigma_{ij}^b$ ). Em ambos os casos verifica-se que conforme o numero de réplicas aumenta, o poder aumenta.

Na Figura 3 foram construídos os gráficos do poder do teste para comparar o poder do teste quando o desvio padrão dos erros de medição dos laboratórios é  $\sigma_{ij}^a$  com o caso quando é  $\sigma_{ij}^b$ .

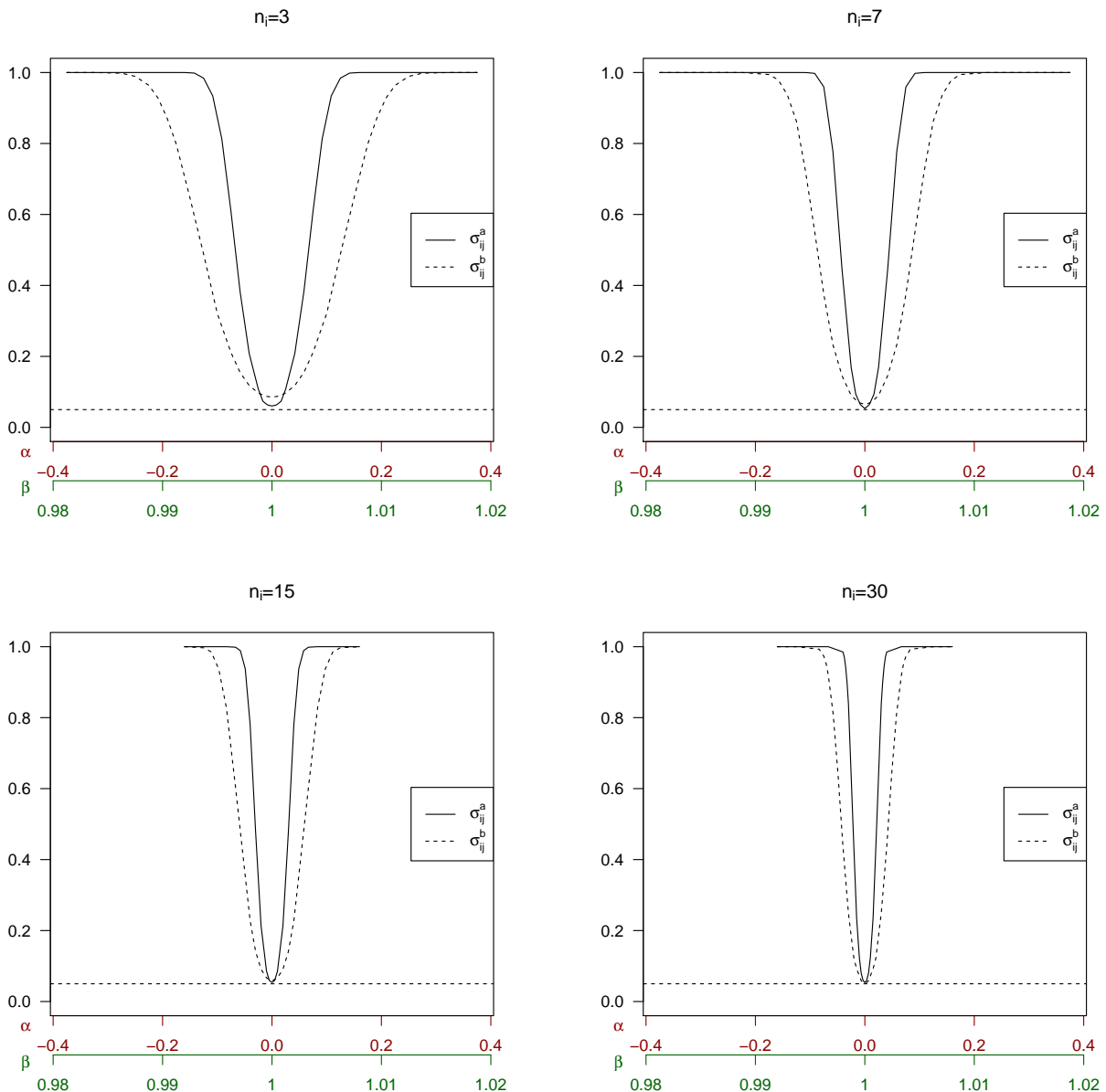
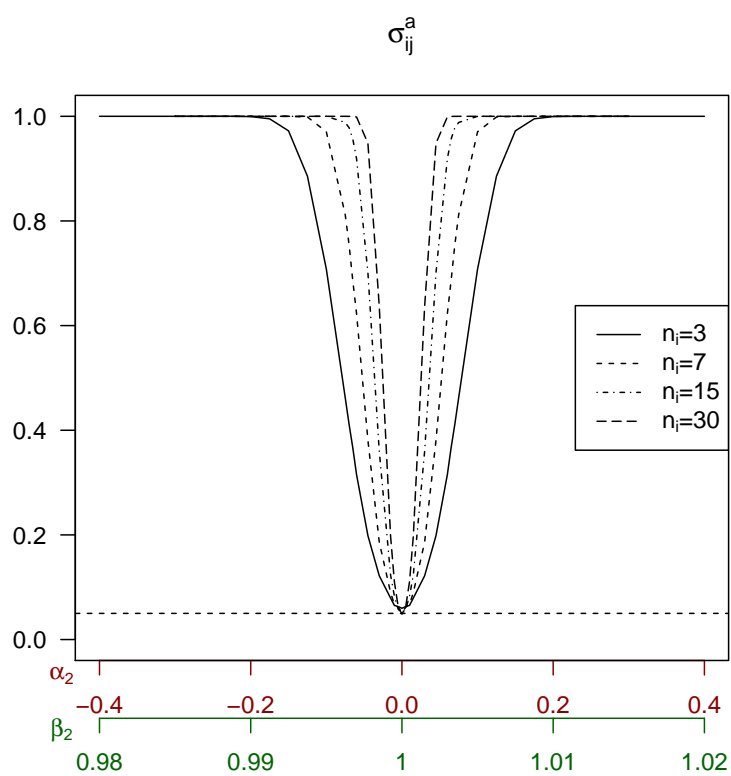
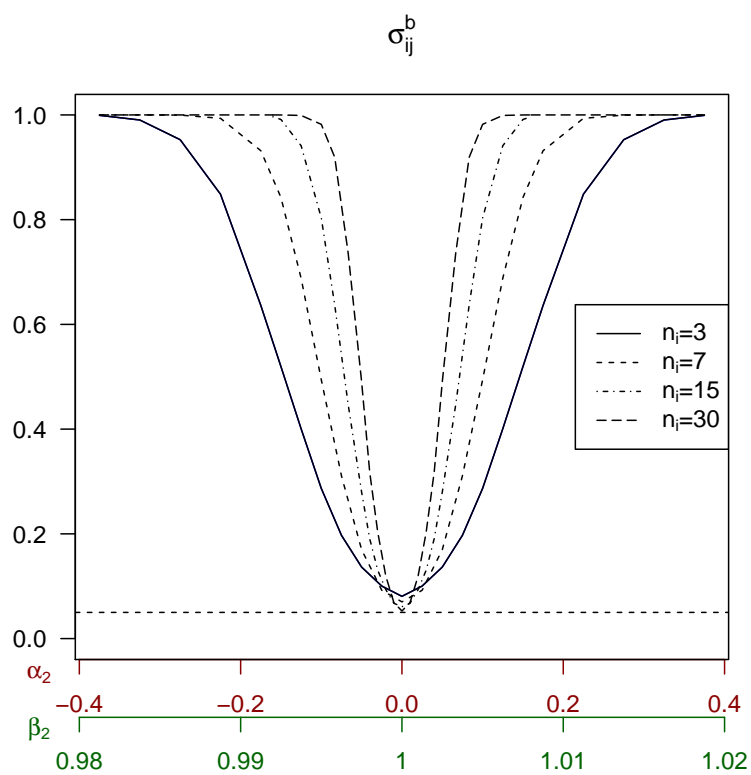


Figura 3 –  $H_0 : \alpha_2 = \dots = \alpha_5 = 0$  e  $\beta_2 = \dots = \beta_5 = 1$  - poder do teste.

Figura 4 –  $H_0 : \alpha_2 = 0, \beta_2 = 1$  - poder do teste com  $\sigma_{ij}^a$ .Figura 5 –  $H_0 : \alpha_2 = 0, \beta_2 = 1$  - poder do teste com  $\sigma_{ij}^b$ .

Observa-se que para os dados simulados com erros de medição com desvio menor o teste é mais poderoso principalmente quando temos poucas réplicas ( $n_i = 3$ ).

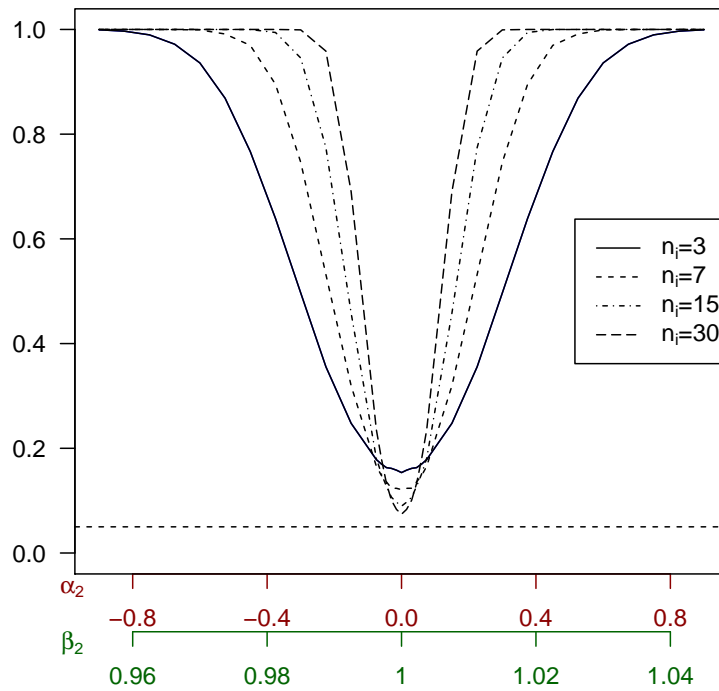


Figura 6 –  $H_0: \alpha_2 = 0$  e  $\beta_2 = 1$  - poder do teste com  $\sigma_{ij}^d$ .

#### 2.4.2.2 $H_0: \alpha_2 = 0$ e $\beta_2 = 1$

Neste caso, foram considerados os valores dos parâmetros descritos no início desta seção e para os valores dos desvios padrões dos erros de medição dos laboratórios foram considerados os conjuntos definidos como  $\sigma_{ij}^a$ ,  $\sigma_{ij}^b$  e  $\sigma_{ij}^d$ . Em todos os casos foram obtidas as porcentagens de rejeição comparando-se com o quantil 95% de uma  $\chi_2^2$ .

As Figuras 4, 5 e 6 mostram o poder de teste para a hipóteses  $H_0: \alpha_2 = 0$  e  $\beta_2 = 1$  com os valores dos desvios dos erros de medição dos laboratórios dados por  $\sigma_{ij}^a$ ,  $\sigma_{ij}^b$  e  $\sigma_{ij}^d$ , respectivamente. Em cada gráfico foram construídas as curvas da função poder para  $n_2 = 3, 7, 15$  e 30 réplicas.

Observa-se em todos eles que conforme o número de réplicas aumenta, aumenta o poder.

Fixado o numero de réplicas e os valores dos parâmetros, exceto  $\sigma_{ij}$   $i = 1, \dots, 5$  e  $j = 1, \dots, 5$ , observa-se que quanto maior a variância do erro de medição menos poderoso é o teste.

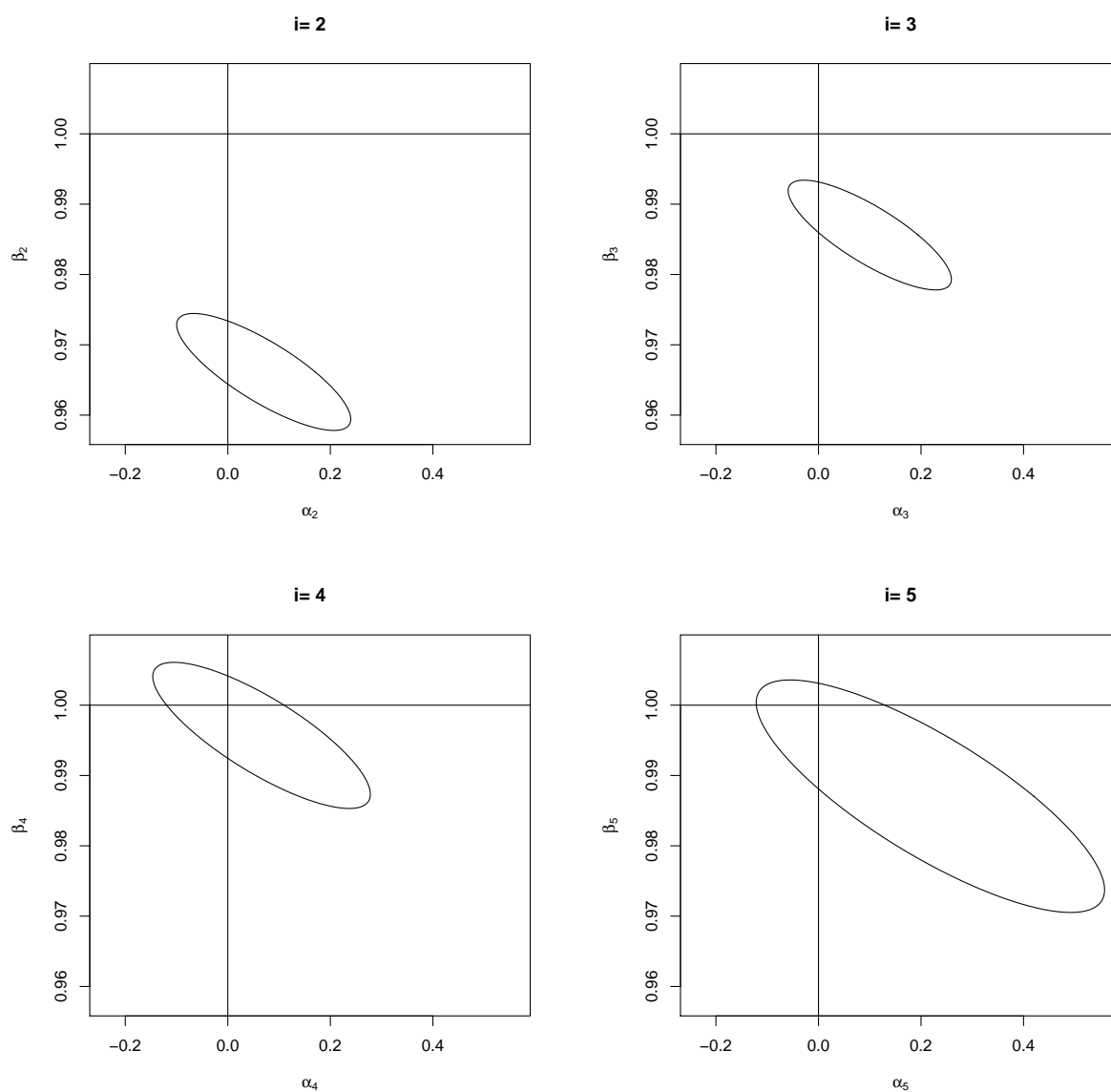


Figura 7 – Regiões de confiança para os laboratórios 2 a 5.

## 2.5 Aplicação

Nesta seção, aplicaremos a metodologia desenvolvida nas seções anteriores ao conjunto de dados reais apresentado no Apêndice A.3. Primeiro, obtivemos as estimativas de máxima verossimilhança dos parâmetros, apresentados na Tabela 3.

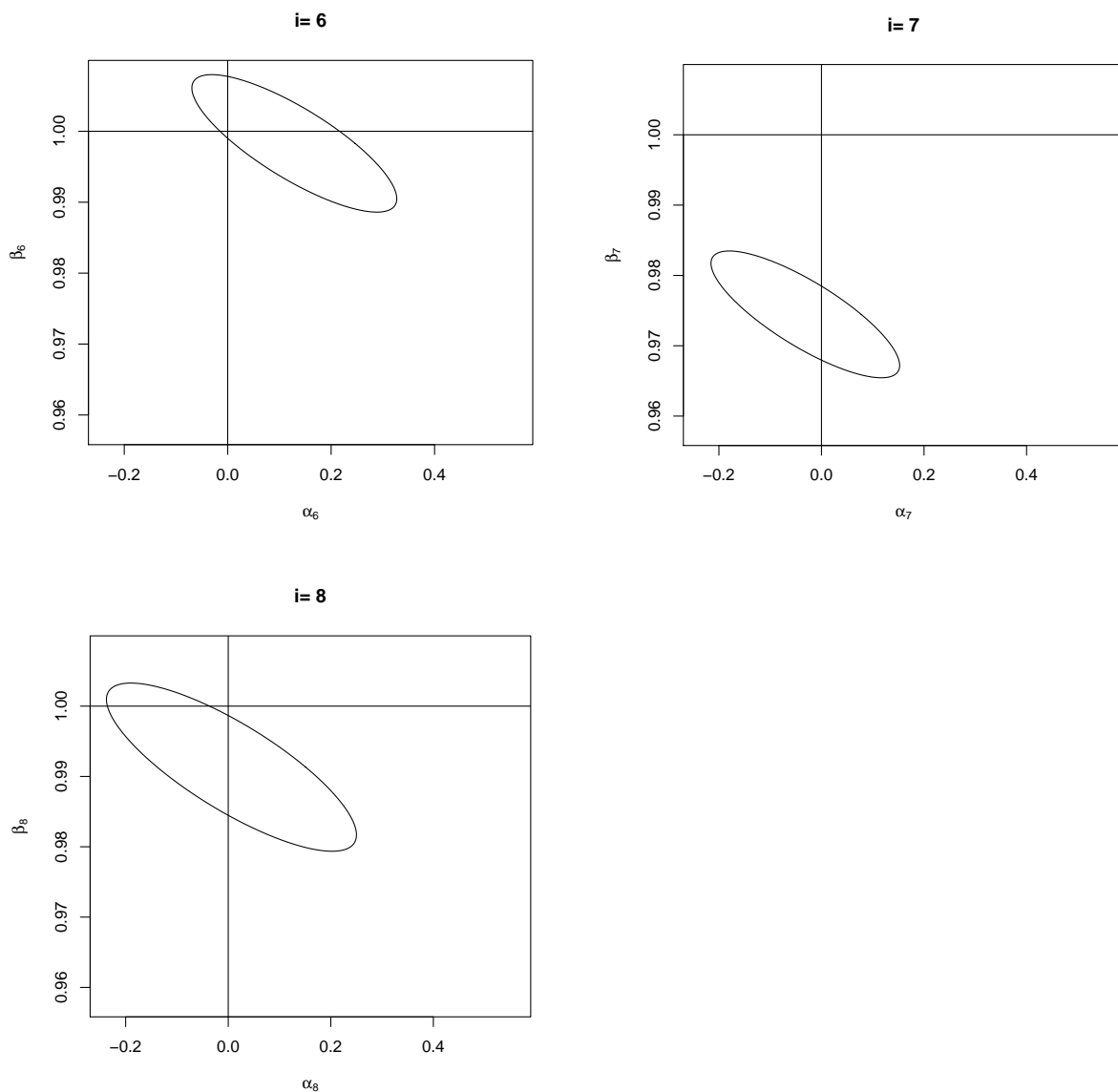


Figura 8 – Regiões de confiança para os laboratorios 6 a 8.

Tabela 3 – Estimativas de máxima verossimilhança do vetor de parâmetros  $\theta$ .

$i$	$\hat{\alpha}_i$	$\hat{\beta}_i$	$j$	$\hat{\mu}_{x_j}$
2	0.0700	0.9661	1	8.8306
3	0.1000	0.9856	2	15.9425
4	0.0658	0.9957	3	26.9652
5	0.2183	0.9871	4	31.5969
6	0.1288	0.9983	5	37.3500
7	-0.0315	0.9745	6	44.3796
8	0.0063	0.9913	7	47.5788
			8	49.6742
			9	50.6601

Em seguida consideramos a estatística do teste de Wald desenvolvida na seção anterior para testar as hipóteses de interesse. Considerando a hipótese  $H_0 : \alpha_2 = \dots = \alpha_8 = 0$  e  $\beta_2 = \dots = \beta_8 = 1$ , obtivemos  $Q_W = 2043.898$ . Concluimos então que o grupo de laboratórios não

são consistentes; há pelo menos um laboratório com viés multiplicativo ou aditivo significativo. Para o teste definido em (2.17) com as hipóteses  $H_0 : \alpha_i = 0$  e  $\beta_i = 1$ ,  $i = 2, \dots, 8$  os valores das estatísticas de teste, juntamente com os  $p$ -valores, estão listados na Tabela 4.

Fixado  $\alpha$  em 1%, concluímos que o laboratório 4 está conforme, assim como o laboratório 5 está conforme. No entanto, não estamos interessados em testar individualmente se um determinado laboratório está conforme. Mas sim quais laboratórios estão conformes a um determinado nível de confiança. Desta forma, foram aplicadas as correções de Hochberg (Hochberg (1988)), Holm (Holm (1979)) e Hommel (Hommel (1988)). Nos três casos podemos concluir a nível de confiança 1% que os laboratórios 4, 5 e 6 estão conformes. Neste caso, a taxa de erro da família dos testes é menor do que 1%.

Tabela 4 –  $Q_W$  para os 8 testes com os respectivos  $p$ -valores entre parênteses.

	$H_0 : \alpha_i = 0$ e $\beta_i = 1, i = 2, \dots, 8$						
	i=2	i=3	i=4	i=5	i=6	i=7	i=8
$Q_{Wi}$	517.267900	69.357334	1.968156	6.639442	10.940891	324.554420	17.563404
$p$ valor	0.000000	0.000000	0.373784	0.036163	0.004209	0.000000	0.000153
$p$ valor Holm	0.000000	0.000000	0.373784	0.072326	0.012628	0.000000	0.000614
$p$ valor Hochberg	0.000000	0.000000	0.373784	0.072326	0.012628	0.000000	0.000614
$p$ valor Hommel	0.000000	0.000000	0.373784	0.072326	0.012628	0.000000	0.000614

Utilizando o resultado (2.18) da seção 2.3.2 foram construídas as regiões de confiança nas Figuras 7 e 8. Foi considerado o valor  $1 - \alpha = 99\%$ . Observa-se que os laboratórios 4, 5 e 6 estão conformes com uma taxa de erro da família de testes menor do que 1%. Um fato interessante é que em relação ao vício aditivo, todos os laboratórios estão conformes.



## MÉTODOS DE DIAGNÓSTICO

Neste capítulo descreveremos as metodologias de influência local de [Cook \(1986\)](#), a curvatura normal conformal de [Poon e Poon \(1999\)](#) e a procura passo a frente de [Atkinson e Riani \(2000\)](#) que serão utilizados na proposta a ser desenvolvida nesta tese.

### 3.1 Influência local de [Cook \(1986\)](#)

[Cook \(1986\)](#) propõe um procedimento para avaliar a influência local de pequenas perturbações em componentes de um modelo. Sejam  $\boldsymbol{\theta}_{h \times 1}$  um vetor de parâmetros desconhecidos, em que o domínio é um subconjunto aberto do espaço euclidiano  $\mathbb{R}^h$ .  $L(\boldsymbol{\theta})$  é uma função objetivo (por exemplo, função de verossimilhança).  $\boldsymbol{\omega} = (w_1, w_2, \dots, w_n)^T$  o vetor de perturbações (onde  $n$  não necessariamente é a dimensão dos dados), por exemplo, a ponderação ou uma pequena mudança nos dados. A função objetivo do modelo perturbado é  $L(\boldsymbol{\theta}|\boldsymbol{\omega})$ ,  $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$  é o EMV do modelo perturbado e  $\hat{\boldsymbol{\theta}}$  o EMV de  $\boldsymbol{\theta}$ .

Sejam  $\ell(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$  e  $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \log L(\boldsymbol{\theta}|\boldsymbol{\omega})$ . Dado  $\boldsymbol{\omega}_0$  o vetor de não perturbação tal que

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = \ell(\boldsymbol{\theta}), \quad (3.1)$$

em que  $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$  tem derivadas parciais contínuas de segunda ordem, [Cook \(1986\)](#) propôs o afastamento da verossimilhança

$$LD(\boldsymbol{\omega}) = 2\{\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})\}, \quad (3.2)$$

para acessar a influência de  $\boldsymbol{\omega}$  ao variar  $\boldsymbol{\omega}$  em  $\boldsymbol{\Omega}$ .

Como a análise de  $LD(\boldsymbol{\omega})$  para todos os possíveis valores de  $\boldsymbol{\omega}$  é impossível, [Cook \(1986\)](#) propôs o estudo do comportamento local em torno de  $\boldsymbol{\omega}_0$  considerando uma superfície

formada pelos elementos do vetor  $\alpha(\boldsymbol{\omega})$ , denotado por gráfico de influência, em que

$$\alpha(\boldsymbol{\omega}) = \begin{pmatrix} \boldsymbol{\omega} \\ LD(\boldsymbol{\omega}) \end{pmatrix}. \quad (3.3)$$

A ideia básica é analisar como  $\alpha(\boldsymbol{\omega})$  desvia-se do plano tangente em  $\boldsymbol{\omega}_o$  e também como a função se comporta em torno de  $\boldsymbol{\omega}_o$ . A taxa de variação em  $\boldsymbol{\omega}_o$  do gráfico de influência reflete a sensibilidade do modelo, em que  $\boldsymbol{\omega}_o$  corresponde ao modelo primário. Maiores detalhes podem ser vistos em Cook (1986). A ideia da representação do gráfico de influência (gráfico da superfície) e o enfoque de influência local é dado no livro de Verbeke e Molenberghs (2000, p. 155, fig. 11.1). Cook (1986) apontou que a curvatura de influência de  $\alpha(\boldsymbol{\omega})$  é dada por

$$C_l = 2|\mathbf{l}^T \boldsymbol{\Delta}^T \ddot{L}^{-1} \boldsymbol{\Delta} \mathbf{l}|, \quad (3.4)$$

em que a direção unitária  $\mathbf{l}$  é um vetor de dimensão  $n \times 1$  que é um autovetor de

$$\ddot{F} = \boldsymbol{\Delta}^T \ddot{L}^{-1} \boldsymbol{\Delta}. \quad (3.5)$$

$-\ddot{L}$  é a matriz de informação observada com dimensão  $h \times h$ , em que

$$\ddot{L} = \left. \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \quad e \quad (3.6)$$

$$\boldsymbol{\Delta} = \left. \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_o}. \quad (3.7)$$

$\boldsymbol{\Delta}$  é uma matriz  $h \times n$ , no ponto  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ ,  $\boldsymbol{\omega} = \boldsymbol{\omega}_o$ . Sejam  $\lambda_1, \dots, \lambda_n$  autovalores da matriz  $\ddot{F}$  em que  $\lambda_1 > \dots > \lambda_n$ . A curva de influência máxima  $C_{max} = 2\lambda_1$ , em que  $\lambda_1$  é o autovalor de  $\ddot{F}$  cujo valor absoluto é máximo e  $\mathbf{l}_{max} = \mathbf{l}_1$  é o autovetor correspondente que é chamado de direção de curvatura máxima de influência.

## 3.2 Curvatura Normal Conformal de Poon e Poon (1999)

A curvatura normal utilizada por Cook (1986) pode assumir um valor positivo qualquer, para facilitar a interpretação, Poon e Poon (1999) sugerem uma curvatura que pertence a uma faixa ou intervalo limitado  $(0, 1)$ . Uma outra vantagem da sua proposta é que esta medida de influência é invariante, por reparametrizações de escala. Outras referências sobre a proposta de Poon e Poon (1999) podem ser encontradas em Poon, Lew e Poon (2000) e Poon e Poon (2002a). Seja  $LD$  a medida de afastamento pela verossimilhança como foi definido em (3.2),  $\boldsymbol{\omega}$  o vetor de perturbações,  $\mathbf{I}$  a primeira forma fundamental de um mapa da superfície e  $\boldsymbol{\Pi}$  a segunda forma fundamental de um mapa da superfície representados por

$$\mathbf{I}_{ij} = \delta_{ij} + \frac{\partial f}{\partial w_i} \frac{\partial f}{\partial w_j} \quad (3.8)$$

e

$$\mathbf{\Pi}_{ij} = \frac{1}{(1 + |\nabla_f|^2)^{1/2}} \frac{\partial^2 f}{\partial w_i \partial w_j},$$

em que,  $f = LD(\boldsymbol{\omega})$ ,  $\delta_{ij}$  é igual a 1 se  $i = j$  e zero caso contrário, e  $|\nabla_f|$  representa a norma do vetor gradiente de  $f$ . Essas duas formas são avaliadas nos vetores  $\mathbf{v}$  e  $\mathbf{w}$  por  $\mathbf{I}(\mathbf{v}, \mathbf{w}) = \mathbf{v}^T \mathbf{I} \mathbf{w}$  e  $\mathbf{\Pi}(\mathbf{v}, \mathbf{w}) = \mathbf{v}^T \mathbf{\Pi} \mathbf{w}$ . Sejam  $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$  uma reta definida em  $\Omega$  passando por  $\boldsymbol{\omega}_0$ , em que  $a \in \mathbb{R}$ ,  $\boldsymbol{\omega}_0$  o vetor de não perturbação e  $\mathbf{l} \in \mathbb{R}^n$ . Então a curvatura normal do gráfico  $\boldsymbol{\alpha}$  (vetor 3.3) na direção  $\mathbf{l}$  no ponto  $\boldsymbol{\omega}_0$  é dada por:

$$C_{\mathbf{l}} = C(\mathbf{l}, \mathbf{l}) = \frac{\mathbf{\Pi}(\mathbf{l}, \mathbf{l})}{\mathbf{I}(\mathbf{l}, \mathbf{l})} = \frac{\mathbf{l}^T H_f \mathbf{l}}{\mathbf{l}^T (\mathbf{I}_n + \nabla_f \nabla_f^T) \mathbf{l} (1 + |\nabla_f|^2)^{1/2}} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}, \quad (3.9)$$

em que  $\mathbf{I}_n$  é a matriz identidade  $n \times n$  e

$$H_f = \frac{\partial^2 f}{\partial w_i \partial w_j}$$

é a matriz hessiana. Cook (1986) propõe usar a curvatura normal para estudar características do gráfico de influência e deduziu que se  $\mathbf{l}^T \mathbf{l} = 1$ , a equação (3.9) é reduzida a

$$C_{\mathbf{l}} = \mathbf{l}^T H_f \mathbf{l} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}. \quad (3.10)$$

Da equação (3.10), temos que:

$$C_{\mathbf{l}} = -2(\mathbf{l}^T \ddot{F} \mathbf{l}) \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}, \quad (3.11)$$

em que,  $\ddot{F}$  é uma matriz  $n \times n$  com elementos dados por  $\partial^2 \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}) / \partial \omega_i \partial \omega_j$ . Seja  $\Delta$  uma matriz  $h \times n$  como definido em (3.7) e  $\check{L}$  uma matriz  $h \times h$  como definido em (3.6) e avaliada em  $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$  e  $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ . Então, a equação (3.11) pode ser escrita como Cook (1986):

$$C_{\mathbf{l}} = -2\{\mathbf{l}^T \Delta^T (\check{L})^{-1} \Delta \mathbf{l}\} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}. \quad (3.12)$$

Na direção  $\mathbf{l}_{max}$  ocorre a maior mudança na função de afastamento pela verossimilhança e podem ser identificadas as observações mais influentes. A direção  $\mathbf{l}_{max}$  é dado por  $C_{max} = \max_{\mathbf{l}}(C_{\mathbf{l}})$ , e  $C_{max}$  junto com  $\mathbf{l}_{max}$  representam o maior autovalor associado ao autovetor da matriz simétrica positiva semidefinida  $-2\ddot{F}$  em (3.11).

A curvatura normal conformal no ponto  $\boldsymbol{\omega}_0$  de um gráfico  $\boldsymbol{\alpha}$  na direção  $\mathbf{l}$  é dado por:

$$B_{\mathbf{l}} = \frac{\mathbf{\Pi}(\mathbf{l}, \mathbf{l})}{\mathbf{I}(\mathbf{l}, \mathbf{l}) \{tr(\mathbf{\Pi}^2)\}^{1/2}} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}. \quad (3.13)$$

A matriz  $\mathbf{\Pi}$  é simétrica, quando os autovalores da segunda forma fundamental são  $\lambda_i$ ,  $i = 1, \dots, n$ ,  $tr(\mathbf{\Pi}^2) = \sum_{i=1}^n \lambda_i^2$ . Em que  $\|H_f\| = \sqrt{tr(H_f^2)}$  então da equação (3.9) e (3.13) temos que

$$B_{\mathbf{l}} = \frac{1}{\mathbf{l}^T (\mathbf{I}_n + \nabla_f \nabla_f^T) \mathbf{l} \|H_f\|} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0}. \quad (3.14)$$

Da equação (3.10)-(3.12) podemos deduzir que a curvatura normal conformal é

$$B_l = \frac{\mathbf{l}^T H_f \mathbf{l}}{\|H_f\|} \Big|_{\boldsymbol{\omega}=\boldsymbol{\omega}_0} = - \frac{\mathbf{l}^T \ddot{F} \mathbf{l}}{\sqrt{\text{tr}(\ddot{F}^2)}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \quad (3.15)$$

ou seja, com as expressões necessárias para calcular a curvatura normal podemos obter a curvatura normal conformal, sem precisar obter novas expressões.

Poon e Poon (1999) mostraram que para qualquer direção  $\mathbf{l}$ ,  $B_l$  satisfaz  $0 \leq |B_l| \leq 1$ .

Se  $\mathbf{l}_i : 1 \leq i \leq n$  é uma coleção de vetores ortonormais (C.3) de  $\boldsymbol{\Pi}$ , então  $\sum_i B_{l_i}^2 = 1$ .  $B_{l_i}$  é igual ao autovalor normalizado  $\hat{\lambda}_i$ , dado por:

$$\hat{\lambda}_i = \frac{\lambda_i}{(\sum_{k=1}^n \lambda_k^2)^{1/2}}. \quad (3.16)$$

Poon e Poon (1999) definem que um autovetor  $\mathbf{l}$  é  $q$  influente se  $|B_l| \geq q/\sqrt{n}$ .

Seja  $\mathbf{e}_t$  o vetor coluna cuja  $t$ -ésima entrada é 1 e todas as outras são 0. Chamamos  $\mathbf{e}_t$  de vetor de perturbação básica do espaço de perturbação. Quando  $\mathbf{l}_i = \sum_{t=1}^n a_{it} \mathbf{e}_t$ ,  $\sum_{t=1}^n a_{it}^2 = 1$  e  $i$  fixo, se a contribuição de todos os  $a_{it}$  forem uniforme, então  $|a_{it}| = 1/\sqrt{n}$ . Com isso, pode-se construir, inclusive, o valor de referência. Esse método pode ser aplicado para estudar  $\mathbf{l}_{max}$  (Poon e Poon (1999)).

Definindo-se  $u_i = |\hat{\lambda}_i|$  os valores absolutos dos autovalores próprios normalizados

$$u_{max} = u_1 \geq \dots \geq u_k \geq q/\sqrt{n} > u_{k+1} \dots u_n \geq 0$$

e usando  $a_{it}$  para denotar o  $t$ -ésimo elemento do autovetor normalizado correspondente a  $u_i$  tal que  $\mathbf{l}_i = \sum_{t=1}^n a_{it} \mathbf{e}_t$ , então a contribuição agregada do  $t$ -ésimo vetor básico de perturbação para todos os autovetores  $q$ -influentes é dado por

$$m(q)_t = \sqrt{\sum_{i=1}^k \mu_i a_{it}^2}. \quad (3.17)$$

Poon e Poon (1999) mostram que se a contribuição de todos os vetores básicos de perturbação forem uniforme, então cada um é igual a

$$\bar{m}(q) = \sqrt{\frac{1}{n} \sum_{i=1}^k \mu_i}. \quad (3.18)$$

Desta forma,  $\bar{m}(q)$  pode ser usado como ponto de corte. Baseado neste ponto de corte, Zhu e Lee (2001) propuseram alguns pontos de corte, aqui consideramos

$$\bar{m}(q) + 2sd, \quad (3.19)$$

onde  $sd$  é o desvio padrão dos elementos do vetor  $\mathbf{m}(q)$ .

Existem dois casos extremos, a primeira é fazer com que  $q$  seja suficientemente grande para que na contribuição individual dos vetores de perturbação básico seja considerada apenas o  $l_{max}$ . A outra é fazer  $q = 0$  de modo que todos os autovalores sejam incluídos na análise.

Quando  $q = 0$ , a contribuição  $m(q)_t$  é chamada de contribuição total e é dada por

$$m_t = m(0)_t = \sqrt{\sum_{i=1}^n \mu_i a_{it}^2}.$$

A contribuição total  $m_t$  e a curvatura normal conformal  $B_{e_t}$  do vetor de perturbação básico são altamente relacionadas.

$$B_{e_t} = \sum_{i=1}^n \hat{\lambda}_i a_{ij}^2.$$

Se todos os autovalores forem não-negativos, então  $B_{e_t}$  será igual ao quadrado da contribuição total do  $t$ -ésimo vetor de perturbação básico. Resumindo, se a matriz Hessiana  $\mathbf{H}_f$  é positiva semidefinida e todos os autovalores são não-negativos então  $m_t^2 = m_t^2(0) = B_{e_t}$  para todo  $t$  segundo *Poon e Poon (1999, Teorema 4)*.

### 3.3 Procura passo a frente de *Atkinson e Riani (2000)*

O algoritmo de procura passo a frente inicialmente desenvolvido para identificar dados atípicos em dados multivariados (*Hadi (1992)*), foi implementado depois para o modelo de regressão linear em *Atkinson (1994)* e *Atkinson e Riani (2000)*, em *Atkinson, Riani et al. (2013)* para a análise de componentes principais, análise discriminante, análise de agrupamento e modelos lineares espaciais. A chave do algoritmo de procura passo a frente é ordenar os dados com base nos resíduos observados.

Em modelos de regressão linear, o processo de seleção de variáveis explicativas e a estimação dos parâmetros é muito sensível à presença de dados atípicos (*Atkinson e Riani (2000)* e *Montgomery, Jennings e Kulahci (2011)*), tais valores podem vir a camuflar (mascarar) a significância das covariáveis e homoscedasticidade feitas à componente do erro.

O propósito da procura passo a frente é identificar observações, as quais são diferentes da maioria dos dados e determinar o efeito destas observações para poder fazer inferências no modelo correto. Pode haver alguns dados atípicos, ou pode ser que as observações possam ser divididas em grupos, de modo que é apropriado ajustar um modelo diferente a cada grupo. Embora seja conveniente referir-se a tais observações como dados atípicos (outliers), eles poderiam ser compostos por uma grande parte dos dados e indicar uma estrutura não suspeita. Quando os dados atípicos são mascarados, métodos “backward” usando a eliminação de observações omitem algumas características importantes. Muitos métodos para a detecção de dados atípicos divide os dados em duas partes, um grupo limpo de dados atípicos e um outro grupo com os dados atípicos. O grupo com os dados limpos são usados para a estimação. A procura passo a

frente fornece subconjuntos de tamanho cada vez maior que são projetados para excluir os dados atípicos até que não exista dados limpos restantes fora do subconjunto.

A ênfase da metodologia *procura passo a frente* é o fato de usarmos informações tais como a estimativa dos parâmetros e o gráfico dos resíduos a cada passo para guiar-nos a um modelo adequado.

Rousseeuw (1984) propõe tomar de forma aleatória subconjuntos iniciais de tamanho  $k$ . Se  $\binom{n}{k}$  (combinatória) for muito grande, Atkinson e Riani (2000) recomendam tomar 1000 subconjuntos iniciais de amostras ou um grande número de candidatos, ajustar o modelo de regressão que seja robusta (Leroy e Rousseeuw (1987)) aos grupos e tomar como o conjunto limpo de dados atípicos aquele que produzir o menor resíduo mediano quadrático.

O procedimento não é sensível ao método utilizado para selecionar um subconjunto inicial. O subconjunto inicial deve ser livre de dados atípicos ou algum tipo de mascaramento. No entanto, mesmo se dados atípicos são incluídos no início, eles são muitas vezes removidos nos primeiros passos.

A procura evita a inclusão inicial de dados atípicos e fornece uma ordenação natural para o modelo especificado.

Em algum estado da *procura passo a frente* é definido o conjunto de  $m$  observações,  $S_*^{(m)}$ , usado para ajustar o modelo ( $m$  é o tamanho do subconjunto  $S_*^{(m)}$ ). Seja  $n$  o número total de observações. O  $m$  varia como:

$$m = m_0, m_0 + 1, \dots, n - 1, n, \quad (3.20)$$

em que o  $m_0$  é o tamanho do grupo inicial  $S_*^{(m_0)}$ ,  $m_0 + 1$  é o tamanho do grupo seguinte e assim por diante até  $n$ . O vetor de parâmetros estimados deste subconjunto é  $\hat{\theta}_m^*$ , com estas estimativas dos parâmetros podemos calcular um conjunto de  $n$  resíduos observados ao quadrado  $r_{im}^{*2}, i = 1, \dots, n$ .

Suponha que o subconjunto  $S_*^{(m)}$  é limpo de dados atípicos, então teríamos  $n - m$  observações não usadas no ajuste que poderiam conter dados atípicos. A identificação destes dados atípicos não é através de um teste formal. O interesse é na evolução, conforme  $m$  vai de  $m_0$  a  $n$ , de quantidades como resíduos, por exemplo. É recomendável observar se existe algum pico ou salto no gráfico da sequência das estimativas dos parâmetros e monitorar as mudanças que acontecem, o que pode ser associada com a introdução de um grupo particular de observações. Na prática, geralmente é uma observação, no entanto, mais de uma observação podem entrar ou sair deste subconjunto de tamanho  $m$  usado para montagem. A interpretação dessas mudanças é complementada através da análise das mudanças no gráfico da *procura passo a frente* dos resíduos.

Dado que foi escolhido um subconjunto de dimensão  $m \geq m_0$ , a *procura passo a frente* move-se para a dimensão  $m + 1$  ordenando os resíduos ao quadrado,  $r_{im}^{*2}, i = 1, \dots, n$ , de cada subconjunto e selecionando o subconjunto com o menor valor mediano. Em que  $r_{[k]m}^{*2}$  é o

$k$ -ésimo resíduo ao quadrado ordenado entre os  $r_{im}^*$ ,  $i = 1, \dots, n$  e como definido em *Atkinson e Riani (2000, p. 29)* e *Rousseeuw e Leroy (1987, p. 133 Teorema 6)*, procura-se o

$$\min_m r_{[med]m}^*{}^2 \text{ com } med = \left\lceil \left\lfloor \frac{n+p+1}{2} \right\rfloor \right\rceil. \quad (3.21)$$

Seja  $\hat{\boldsymbol{\theta}}_{m_0}^*$  o estimador de  $\boldsymbol{\theta}_{m_0}^*$  dos dados de tamanho  $m_0$  e seja o estimador  $\hat{\boldsymbol{\theta}}_n^* = \hat{\boldsymbol{\theta}}$  dos dados completos. Na ausência de dados atípicos a seguinte relação deve ocorrer:

$$E(\hat{\boldsymbol{\theta}}_{m_0}^*) = E(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta}, \quad (3.22)$$

isto é, ambos estimadores serão não viciados (*Atkinson e Riani (2000)*) espera-se a mesma propriedade para a sequência de estimativas de  $\hat{\boldsymbol{\theta}}_m^*$  produzido na procura passo a frente com  $m$  definido em (3.20).



---

# INFLUÊNCIA LOCAL COM PROCURA PASSO A FRENTE E CURVATURA NORMAL CONFORMAL COM PROCURA PASSO A FRENTE

---

## 4.1 Influência local com procura passo a frente: **ILPPF**

Uma metodologia bastante utilizada para a análise de diagnóstico em modelos de regressão com erros nas variáveis é a influência local de [Cook \(1986\)](#). Veja por exemplo, [Lee e Zhao \(1996\)](#), [Galea-Rojas, Bolfarine e Castro \(2002\)](#), [Lee e Tang \(2004\)](#), [Lee, Lu e Song \(2006\)](#), [Rasekh \(2006\)](#), [Aoki, Singer e Bolfarine \(2007\)](#).

Neste capítulo vamos propor uma nova metodologia com o uso da procura passo a frente de [Atkinson e Riani \(2000\)](#) conjuntamente com o método de influência local de [Cook \(1986\)](#) e aplicar em modelos de regressão com erros de medição.

Na metodologia de influência global a verificação das mudanças decorrentes nas estimativas dos parâmetros de regressão ao retirar uma observação (feito para cada caso) ocasiona um alto custo computacional na detecção de dados atípicos e/ou influentes. Além disso, podem ainda haver dados que não foram detectadas pelos métodos de influência global devido à presença de algum outro dado(s) (problema de mascaramento), ou seja, a existência de dados atípicos mascarados que não foram detectadas pela medida de influência global, para isso, [Atkinson e Riani \(2000\)](#) propuseram a metodologia de procura passo a frente.

Por outro lado, a metodologia de influência local de [Cook \(1986\)](#) é uma medida de influência local que introduz pequenas perturbações nos dados ou no modelo para avaliar se o modelo é sensível a essas perturbações, enquanto que a influência global utiliza alguma medida como DFBeta, DFFitS e D-Cook [Cook \(1977\)](#), [Belsley, Kuh e Welsch \(1980\)](#), [Cook e Weisberg](#)

(1982). Diferentemente da metodologia de influência global, esta metodologia permite avaliar a influência conjunta de todas as observações. No entanto, neste caso também pode haver dados influentes mascarados que não foram detectados com o uso da influência local. Desta forma, a nossa proposta é utilizar a influência local de Cook (1986) e propor uma nova metodologia do tipo procura passo a frente de Atkinson e Riani (2000) para a obtenção de dados mascarados localmente influentes.

Seja:

$$LD_i(\omega) = 2\{\ell_i(\hat{\theta}) - \ell_i(\hat{\theta}_\omega)\} \quad i = 1, \dots, n, \quad (4.1)$$

em que,  $\ell(\theta) = \sum_{i=1}^n \log L_i(\theta) = \sum_{i=1}^n \ell_i(\theta)$ ;  $\hat{\theta}$  e  $\hat{\theta}_\omega$  os estimadores de máxima verossimilhança de  $\theta$  em  $L(\theta)$  e  $L(\theta|\omega)$ , respectivamente.  $L(\theta)$  e  $L(\theta|\omega)$  como definido em (3.2 e 3.3).

A metodologia proposta consiste em obter subconjuntos de tamanho  $m$  livre de pontos discrepantes ou influentes, estimar os parâmetros e obter o  $LD_i(\omega)$  definido em (4.1) considerando todos os indivíduos. Ordenar, e tomar o valor mediano e considerar como o subconjunto seguinte de tamanho  $(m + 1)$  aquele que tiver o menor valor mediano.

Desta forma, vamos considerar um subconjunto de observações que sejam resistentes a pequenas perturbações no conjunto de dados ou no modelo. No entanto, observe que ao considerar a direção  $\mathbf{l}_{max}$  correspondente ao  $C_{max}$  da matriz simétrica positiva semidefinida  $\tilde{F}$ , se o  $i$ -ésimo elemento de  $\mathbf{l}_{max}$  for relativamente grande é porque a perturbação no peso  $\omega_i$  do  $i$ -ésimo caso pode levar a mudanças substanciais no resultado da análise, ou seja, vamos ordenar então o vetor  $\mathbf{l}_{max}$  e considerar o subconjunto que obtiver o menor valor mediano considerando o conjunto todo. Vamos utilizar esta metodologia em modelos de regressão com erros nas variáveis.

**Passos:** Vamos definir os passos desta metodologia que estamos propondo.

1. No primeiro passo começamos ajustando modelos de regressão com erros de medição aos pequenos subconjuntos formados por  $m = m_0$  elementos. Obtemos as estimativas de máxima verossimilhança dos parâmetros em cada um dos  $C_m^n = \binom{n}{m}$  subconjuntos.

Considerando agora o conjunto de dados completo (com as  $n$  observações) e o modelo ajustado com as  $m$  observações, obtemos os autovetores associados ao maior autovalor. A direção  $\mathbf{l}_{max}$  é dado por  $C_{max} = \max_{\mathbf{l}}(C_{\mathbf{l}})$ , e  $C_{max}$  junto com  $\mathbf{l}_{max}$  são o maior autovalor associado ao autovetor da matriz simétrica positiva semidefinida  $\tilde{F}$  (segundo o enfoque de Cook (1986)). Depois, ordenamos estes autovetores e finalmente, obtemos as medianas  $med$  destes autovetores (3.21) (autovetores com  $n$  elementos) associados ao maior autovalor para cada um dos subconjuntos de dados (subconjuntos com  $m$  observações). Ordenamos e tomamos o subconjunto referente ao menor valor mediano do autovetor correspondente ao maior autovalor. Na Tabela 5, representamos este passo para uma amostra de tamanho  $n = 4$  (como um exemplo simples, pequeno e pouco convencional), onde iniciamos com  $m_0 = 2$ .

Possíveis subconjuntos	1	2	3	4	5	6
Elementos no subconjunto	1 2	1 3	1 4	2 3	2 4	3 4
Estimativas	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
Autovetores	$l_1 =$	$l_2 =$	$l_3 =$	$l_4 =$	$l_5 =$	$l_6 =$
	$l_{11}$ $l_{21}$ ...	$l_{12}$ $l_{22}$ ...	$l_{13}$ $l_{23}$ ...	$l_{14}$ $l_{24}$ ...	$l_{15}$ $l_{25}$ ...	$l_{16}$ $l_{26}$ ...
	$l_{med_1}$	$l_{med_2}$	$l_{med_3}$	$l_{med_4}$	$l_{med_5}$	$l_{med_6}$
	...	...	...	...	...	...
	$l_{n1}$	$l_{n2}$	$l_{n3}$	$l_{n4}$	$l_{n5}$	$l_{n6}$

Tabela 5 – Subconjuntos, começando com  $m_0 = 2$ , onde  $n = 4$ , temos um total de 6 subconjuntos ( $C_{n=m_0}^n$ ).  $l_{m_0}^* = \min(l_{med_1}, l_{med_2}, l_{med_3}, l_{med_4}, l_{med_5}, l_{med_6})$  e o autovetor correspondente.

$l_{m_0}^*$

- No segundo passo, subconjuntos de tamanho  $m + 1$  são considerados de tal forma que uma ou mais observações podem entrar ou sair do subconjunto anterior.

Da mesma forma como no passo anterior, são obtidas as estimativas dos parâmetros para cada subconjunto de tamanho  $m + 1$ , depois considerando o conjunto de dados todo, são obtidos os autovetores associados ao maior autovalor para cada modelo ajustado e finalmente estes autovetores são ordenados. As medianas de cada um desses autovetores são obtidas e é escolhido o subconjunto que tiver o menor valor mediano. Como representado na Tabela 6.

Possíveis subconjuntos	1	2	3	4
Elementos no subconjunto	1 2 3	1 2 4	2 3 4	3 4 1
Estimativas	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
Autovetores	$l_1 =$	$l_2 =$	$l_3 =$	$l_4 =$
	$l_{11}$ $l_{21}$ ...	$l_{12}$ $l_{22}$ ...	$l_{13}$ $l_{23}$ ...	$l_{14}$ $l_{24}$ ...
	$l_{med_1}$	$l_{med_2}$	$l_{med_3}$	$l_{med_4}$
	...	...	...	...
	$l_{n1}$	$l_{n2}$	$l_{n3}$	$l_{n4}$

Tabela 6 – Subconjuntos, começando com  $m = m_0 + 1 = 3$ , onde  $n = 4$ , temos um total de 4 subconjuntos ( $C_{n=m_0+1}^n$ ).

$l_{m_0+1}^* = \min(l_{med_1}, l_{med_2}, l_{med_3}, l_{med_4})$

$l_{m_0+1}^*$

- No terceiro passo é obtido o subconjunto de tamanho  $m + 2$  da mesma forma, são obtidas as estimativas dos parâmetros para cada subconjunto de tamanho  $m + 2$ , são calculados

os autovalores com todas as observações e escolhido o subconjunto com o menor valor mediano, da mesma forma que foram obtidos os subconjuntos de tamanhos  $m$  e  $m + 1$ , e assim por diante, até chegar a última observação ( $n$ -ésima). Veja Tabela 7

Possíveis subconjuntos	1
Elementos no subconjunto	1 2 3 4
Estimativas	$\hat{\beta}$
Autovetores	$\mathbf{l} =$
	$l_1$ $l_2$ ... $l_{med}$ ... $l_n$

Tabela 7 – Neste caso,  $m = n$  e temos apenas um conjunto que é o próprio conjunto de dados ( $C_{m=n}^n$ ).  
 $\mathbf{l}_n^*$

**Observação:** Se  ${}^n C_m$  for muito grande, pode ser considerado no lugar disso algum número grande de subconjuntos aleatórios, por exemplo, de 1000 amostras como máximo.

$$\text{Número de subconjuntos} = \begin{cases} {}^n C_m, & \text{se } {}^n C_m < 1000, \\ 1000, & \text{se } {}^n C_m \geq 1000, \end{cases} \quad (4.2)$$

segundo [Atkinson e Riani \(2000\)](#).

Denotando por  $\mathbf{l}_m^*$  o autovetor selecionado a cada passo da iteração de influência local de [Cook \(1986\)](#) com procura passo a frente (ILPPF) e  $\mathbf{l}_{ILPPF}$  a coleção desses autovetores associados ao maior autovalor produzidos a cada passo do processo, temos:

$$\mathbf{l}_{ILPPF} = (\mathbf{l}_{m_0}^*, \mathbf{l}_{m_0+1}^*, \dots, \mathbf{l}_n^*). \quad (4.3)$$

Podemos construir o gráfico passo a frente destes vetores, onde no eixo da ordenada temos os vetores  $\mathbf{l}_m^*$  obtido a cada passo  $m$ ,  $m = m_0, \dots, n$ .

De forma resumida, a ideia é considerar a metodologia de procura passo a frente conforme [Atkinson e Riani \(2000\)](#) e utilizar como critério de escolha do grupo limpo a metodologia de influência local de [Cook \(1986\)](#), como descrito anteriormente.

Vamos aplicar a metodologia proposta no modelo de regressão univariado com erros nas variáveis com intercepto nulo, e em seguida, ao modelo de regressão multivariado com erros nas variáveis com intercepto nulo. Observe que essa metodologia pode ser aplicada a outros modelos estatísticos (regressão, series temporais, componentes principais, etc).

## 4.2 Curvatura Normal Conformal com Procura Passo a Frente: CNCPPF

Nesta seção, apresentamos a metodologia proposta de CNCPPF. O objetivo do algoritmo de procura passo a frente é detectar vários dados atípicos mascarados. Por outro lado, [Cook \(1986\)](#) propôs uma metodologia geral para avaliar a influência local de pequenas perturbações no conjunto de dados ou no modelo. Além disso, [Poon e Poon \(1999\)](#) introduziram a curvatura normal conformal para estabelecer um ponto de corte. Nesta seção, propomos o uso da curvatura normal conformal e a procura passo a frente para detectar observações influentes mascaradas ou grupo de observações.

Para aplicar a curvatura normal conformal, foram consideradas quatro esquemas de perturbação. As matrizes necessárias para calcular a contribuição agregada do  $j$ -ésimo vetor de perturbação básico,  $j = 1, \dots, r$  para o modelo definido em (1.4) podem ser encontradas em [Russo \(2006\)](#), enquanto que para o modelo definido em (1.4) essas expressões podem ser encontradas em [Aoki, Singer e Bolfarine \(2007\)](#).

Muitas das metodologias desenvolvidas para obter dados atípicos mascarados dividem o conjunto de dados em um subconjunto limpo, livre de dados discrepantes, o conjunto básico e outro conjunto de dados composto pelo restante das observações com possíveis dados discrepantes. É definido um critério pelo qual novas observações são introduzidas no conjunto básico e é incrementado até que todas as observações sejam incluídas. Portanto, na última etapa, obtemos as estimativas dos parâmetros e a análise considerando todo o conjunto de dados. A escolha do subconjunto inicial não influencia muito a procura e não afeta as etapas finais em que os resultados mais importantes da análise estão concentrados.

Depois de definir o esquema de perturbação a ser usado, a metodologia começa em  $s = s_0$  no primeiro passo e termina em  $s = n$  no último passo. De modo que

$$s = s_0, s_0 + 1, \dots, n - 1, n.$$

Em cada etapa é amostrado todos os subconjuntos possíveis de  $\binom{n}{s}$  de tamanho  $s$  ou se  $\binom{n}{s} \geq 1000$ , são amostrados 1000 subconjuntos, vamos indicar o número de subconjuntos na  $s$ -ésima iteração como  $b_s$ . Portanto, na primeira etapa, existem  $b_{s_0}$  subconjuntos de tamanho  $s_0$  e na última iteração existem  $b_n$  subconjuntos de tamanho  $n$  ( $b_n = 1$ ).

O algoritmo começa amostrando  $b_{s_0}$  amostras de tamanho  $s_0$  e, em seguida, é calculado o EMV dos parâmetros para cada um dos  $b_{s_0}$  subconjuntos. Posteriormente, para cada um dos  $b_{s_0}$  subconjuntos, é escolhido o valor  $q$ , de forma que a contribuição agregada do  $j$ -ésimo vetor de perturbação básica (em que  $s$  se refere ao tamanho do subconjunto) inclua os  $k$  maiores autovalores. Além disso, para cada subconjunto é calculado  $\mathbf{m}(q)_{s_0}^h$ ,  $h = 1, \dots, b_{s_0}$ , considerando todas as  $n$  observações (todo o conjunto de dados), mas com o EMV obtido para esse subconjunto. Posteriormente, cada um dos  $b_{s_0}$  vetores  $\mathbf{m}(q)_{s_0}^h$  é ordenado e é escolhido o vetor com a menor

mediana, que será indicado por  $\mathbf{m}(q)_{s_0}^*$ .

A procura passo a frente move para a próxima iteração com  $s = s_0 + 1$  e assim por diante até que todas as observações sejam incluídas. No final, é obtido

$$\mathbf{m}(q)^* = (\mathbf{m}(q)_{s_0}^*, \mathbf{m}(q)_{s_0+1}^*, \dots, \mathbf{m}(q)_n^*)^T.$$

$\mathbf{m}(q)_n^*$  é a contribuição agregada obtida considerando todo o conjunto de dados, isto é, a contribuição agregada usual onde é considerado o conjunto de dados todo para obter o EMV dos parâmetros.

Outras quantidades de interesse podem ser obtidas, como por exemplo

$$\hat{\boldsymbol{\theta}}^* = (\hat{\boldsymbol{\theta}}_{s_0}^*, \dots, \hat{\boldsymbol{\theta}}_n^*)^T,$$

onde  $\hat{\boldsymbol{\theta}}_s^*$ ,  $s = s_0, \dots, n$  representa os parâmetros estimados do conjunto escolhido em cada iteração.

Essas quantidades podem ser resumidas usando o gráfico passo a frente.

Zhu e Lee (2001) propuseram alguns pontos de corte, considerando a curvatura normal conformal de Poon e Poon (1999), neste trabalho introduzimos  $\bar{m}(q) + 2sd$  com  $sd$  denotando o desvio padrão dos elementos do vetor  $\mathbf{m}(q)$  na contribuição agregada.

As duas metodologias serão aplicadas nos dados apresentados na introdução.

# ILPPF NO MODELO DE REGRESSÃO COM ERROS NAS VARIÁVEIS COM INTERCEPTO NULO

Neste Capítulo aplicaremos a metodologia desenvolvida na Seção 4.1 aos conjuntos de dados que motivaram o uso do modelo de regressão com erros nas variáveis com intercepto nulo. Iremos definir o modelo de regressão com erros nas variáveis com intercepto nulo e as perturbações que serão utilizadas. Depois, iremos aplicar a metodologia proposta de ILPPF aos dois conjuntos de dados descritos na introdução (1.1 e 1.2) considerando apenas uma variável resposta.

Sejam:

$$\begin{cases} X_i = \xi_i + \delta_i, \\ Y_i = \beta \xi_i + \varepsilon_i, \end{cases} \quad i = 1, \dots, n, \quad (5.1)$$

onde,  $\delta_i \stackrel{ind.}{\sim} N(0, \sigma_\delta^2)$ ,  $\varepsilon_i \stackrel{ind.}{\sim} N(0, \sigma_\varepsilon^2)$ ,  $\xi_i \stackrel{ind.}{\sim} N(\mu, \sigma_x^2)$ ,  $\delta_i$ ,  $\varepsilon_i$  e  $\xi_i$  independentes, para todo  $i = 1, \dots, n$ .

Da equação (5.1) podemos reescrever o modelo com erros da seguinte forma:

$$\begin{aligned} Y_i &= \beta(X_i - \delta_i) + \varepsilon_i \\ &= \beta X_i - \beta \delta_i + \varepsilon_i \\ &= \beta X_i + \varepsilon_i - \beta \delta_i, \quad i = 1, \dots, n. \end{aligned} \quad (5.2)$$

Da última expressão, mostra-se a diferença do modelo com erros de medição (5.1) com o modelo de regressão usual, o erro de (5.2) depende de  $\beta$ . E o termo  $(\varepsilon_i - \beta \delta_i)$  é correlacionado

com  $X_i$

$$\begin{aligned}
Cov(X_i, (\varepsilon_i - \beta \delta_i)) &= Cov(\xi_i + \delta_i, \varepsilon_i - \beta \delta_i) \\
&= Cov(\xi_i, \varepsilon_i) + Cov(\xi_i, -\beta \delta_i) + Cov(\delta_i, \varepsilon_i) + Cov(\delta_i, -\beta \delta_i) \\
&= -\beta Cov(\delta_i, \delta_i) \\
&= -\beta \sigma_\delta^2, \quad i = 1, \dots, n
\end{aligned} \tag{5.3}$$

e será zero se  $\beta$  ou  $\sigma_\delta^2$  é zero. Quando  $\sigma_\delta^2 = 0$ , a expressão se reduz ao modelo de regressão usual. Se usarmos estimadores de mínimos quadrados, obteremos estimadores inconsistentes (Fuller (1987)).

Seja  $\mathbf{Z}_i$  o vetor de observações do  $i$ -ésimo indivíduo, ou seja,  $\mathbf{Z}_i = (X_i, Y_i)^T$  e considerando que  $\mathbf{Z}_i$  tem distribuição normal bivariada e

$$\mathbf{a} = (1, \beta)^T, \quad \mathbf{m} = \mu \mathbf{a} \quad \text{e} \quad \tau = \text{diag}(\sigma_\delta^2, \sigma_\varepsilon^2), \tag{5.4}$$

temos:

$$\begin{aligned}
\mathbf{Z}_i &= \xi_i \mathbf{a} + \begin{pmatrix} \delta_i \\ \varepsilon_i \end{pmatrix} \quad \text{e} \\
\mathbf{Z}_i &\stackrel{iid}{\sim} N_2(\mathbf{m}, \mathbf{V}), \quad i = 1, \dots, n,
\end{aligned} \tag{5.5}$$

em que

$$\mathbf{V} = \sigma_x^2 \mathbf{a} \mathbf{a}^T + \tau = \begin{pmatrix} \sigma_x^2 + \sigma_\delta^2 & \beta \sigma_x^2 \\ \beta \sigma_x^2 & \beta^2 \sigma_x^2 + \sigma_\varepsilon^2 \end{pmatrix}. \tag{5.6}$$

O determinante e a inversa de  $\mathbf{V}$  são dados por:

$$|\mathbf{V}| = c^{-1} \sigma_\delta^2 \sigma_\varepsilon^2 \sigma_x^2 \quad \text{e} \quad \mathbf{V}^{-1} = \tau^{-1} - c \tau^{-1} \mathbf{a} \mathbf{a}^T \tau^{-1}, \tag{5.7}$$

em que

$$c = \sigma_x^2 (1 + \sigma_x^2 \mathbf{a}^T \tau^{-1} \mathbf{a})^{-1}. \tag{5.8}$$

Seja  $\boldsymbol{\theta} = (\beta, \mu, \sigma_\delta^2, \sigma_x^2, \sigma_\varepsilon^2)^T \in \Theta \subseteq \mathbb{R}^5$  (modelo 5.1), em que  $\sigma_\delta^2 > 0$ ,  $\sigma_x^2 > 0$  e  $\sigma_\varepsilon^2 > 0$ . Temos que a função densidade de probabilidade para a  $i$ -ésima observação é dada por

$$f(\mathbf{Z}_i, \boldsymbol{\theta}) = (2\pi)^{-1} |\mathbf{V}|^{-\frac{1}{2}} e^{-\frac{1}{2} (\mathbf{Z}_i - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i - \mathbf{m})}.$$

A função densidade de probabilidade conjunta para todas as observações,  $i = 1, \dots, n$ , é dada por:

$$\begin{aligned}
f(\mathbf{Z}, \boldsymbol{\theta}) &= \prod_{i=1}^n (2\pi)^{-1} |\mathbf{V}|^{-\frac{1}{2}} e^{-\frac{1}{2} (\mathbf{Z}_i - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i - \mathbf{m})} \\
&= (2\pi)^{-n} |\mathbf{V}|^{-\frac{n}{2}} e^{-\frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i - \mathbf{m})}.
\end{aligned} \tag{5.9}$$

O logaritmo da função verossimilhança do modelo (5.5) é dado por

$$\begin{aligned}
\ell(\boldsymbol{\theta}; \mathbf{Z}) &= \log f(\mathbf{Z}, \boldsymbol{\theta}) \\
&= \sum_{i=1}^n \log f(\mathbf{Z}_i; \boldsymbol{\theta}) \\
&= -n \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i - \mathbf{m}) \\
&= -n \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{i=1}^n q_i,
\end{aligned} \tag{5.10}$$

em que  $\boldsymbol{\theta} = (\beta, \mu, \sigma_\delta^2, \sigma_x^2, \sigma_e^2)^T$  denota o vetor de parâmetros do modelo (5.5),

$$\begin{aligned}
q_i &= (\mathbf{Z}_i - \boldsymbol{\mu}\mathbf{a})^T \mathbf{V}^{-1} (\mathbf{Z}_i - \boldsymbol{\mu}\mathbf{a}) = q_{i1} - cq_{i2}^2, \quad \text{com} \\
q_{i1} &= (\mathbf{Z}_i - \boldsymbol{\mu}\mathbf{a})^T \boldsymbol{\tau}^{-1} (\mathbf{Z}_i - \boldsymbol{\mu}\mathbf{a}) = \sigma_\delta^{-2} (X_i - \mu)^2 + \sigma_e^{-2} (Y_i - \mu\beta)^2 \quad \text{e} \\
q_{i2} &= \mathbf{a}^T \boldsymbol{\tau}^{-1} (\mathbf{Z}_i - \mathbf{m}) = \sigma_\delta^{-2} (X_i - \mu) + \beta \sigma_e^{-2} (Y_i - \mu\beta), \quad i = 1, \dots, n.
\end{aligned} \tag{5.11}$$

O modelo estrutural normal com intercepto nulo possui 5 parâmetros  $(\beta, \mu, \sigma_x^2, \sigma_\delta^2, \sigma_e^2)^T$  e cinco estatísticas suficientes  $(\bar{X}, \bar{Y}, S_{XY}, S_{XX}, S_{YY})^T$ , dadas por  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ ;  $\bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i$ ;  $S_{XX} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$ ;  $S_{YY} = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$ ; e  $S_{XY} = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$ .

Dessa forma, os estimadores de máxima verossimilhança dos parâmetros são dados por:

$$\begin{aligned}
\hat{\mu} &= \bar{X}, \quad E(X) \neq 0; \\
\hat{\beta} &= \frac{\bar{Y}}{\bar{X}}, \quad \text{assumindo que } \bar{X} \neq 0; \\
\hat{\sigma}_x^2 &= \frac{S_{XY}}{\hat{\beta}}; \\
\hat{\sigma}_\delta^2 &= S_{XX} - \frac{S_{XY}}{\hat{\beta}}; \\
\hat{\sigma}_e^2 &= S_{YY} - \hat{\beta} S_{XY}.
\end{aligned} \tag{5.12}$$

A seguir foram obtidos a função escore e a matriz de informação observada.

Para isso foram calculados as derivadas do  $\log |\mathbf{V}|$ ,  $q_{i1}$ ,  $q_{i2}$ , e  $c$  em relação aos elementos de  $\boldsymbol{\theta}$ . Os cálculos das derivadas foram **conferidas por meio do programa Wolfram Research, Inc. (2012)**.

As derivadas do logaritmo de  $|\mathbf{V}|$ , dado em (5.7), em relação aos componentes de  $\boldsymbol{\theta}$  são dados por:

$$\begin{aligned}
\frac{\partial \log |\mathbf{V}|}{\partial \beta} &= 2c\sigma_e^{-2}\beta, & \frac{\partial \log |\mathbf{V}|}{\partial \mu} &= 0, & \frac{\partial \log |\mathbf{V}|}{\partial \sigma_\delta^2} &= \sigma_\delta^{-2}(1 - c\sigma_\delta^{-2}), \\
\frac{\partial \log |\mathbf{V}|}{\partial \sigma_x^2} &= \sigma_x^{-2}(1 - c\sigma_x^{-2}) & \text{e} & \frac{\partial \log |\mathbf{V}|}{\partial \sigma_e^2} &= \sigma_e^{-2}(1 - c\beta^2\sigma_e^{-2}).
\end{aligned}$$

As derivadas de  $q_{i1}$ ,  $q_{i2}$  e  $c$ , dados em (5.8 e 5.11), em relação aos componentes de  $\boldsymbol{\theta}$  são dados por:

$$\begin{aligned} \frac{\partial q_{i1}}{\partial \beta} &= \mu(-2\sigma_e^{-2}(y_i - \beta\mu)), & \frac{\partial q_{i1}}{\partial \mu} &= -2\sigma_\delta^{-2}(x_i - \mu) - 2\beta\sigma_e^{-2}(y_i - \beta\mu), \\ \frac{\partial q_{i1}}{\partial \sigma_\delta^2} &= -\sigma_\delta^{-4}(x_i - \mu)^2, & \frac{\partial q_{i1}}{\partial \sigma_x^2} &= 0, & \frac{\partial q_{i1}}{\partial \sigma_e^2} &= -\sigma_e^{-4}(y_i - \beta\mu)^2, \\ \frac{\partial q_{i2}}{\partial \beta} &= \sigma_e^{-2}(y_i - 2\beta\mu), & \frac{\partial q_{i2}}{\partial \mu} &= -(\sigma_\delta^{-2} + \beta^2\sigma_e^{-2}), & \frac{\partial q_{i2}}{\partial \sigma_\delta^2} &= -\sigma_\delta^{-4}(x_i - \mu), \\ \frac{\partial q_{i2}}{\partial \sigma_x^2} &= 0, & \frac{\partial q_{i2}}{\partial \sigma_e^2} &= -\beta\sigma_e^{-4}(y_i - \beta\mu), & \frac{\partial c}{\partial \beta} &= -2c^2\sigma_e^{-2}\beta, & \frac{\partial c}{\partial \mu} &= 0, \\ \frac{\partial c}{\partial \sigma_\delta^2} &= c^2\sigma_\delta^{-4}, & \frac{\partial c}{\partial \sigma_x^2} &= c^2\sigma_x^{-4} & \text{e} & \frac{\partial c}{\partial \sigma_e^2} &= \beta^2c^2\sigma_e^{-4}. \end{aligned}$$

Dessa forma, após manipulações algébricas a função escore é dada por:

$$\begin{aligned} \frac{\partial \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta} &= \sum_{i=1}^n \{ \sigma_e^{-2} [-c\beta + \mu(y_i - \beta\mu) + q_{i2}(y_i - 2\beta\mu)c - q_{i2}^2c^2\beta] \}; \\ \frac{\partial \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu} &= \sum_{i=1}^n \{ q_{i2} - cq_{i2}(\sigma_\delta^{-2} + \beta^2\sigma_e^{-2}) \}; \\ \frac{\partial \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2} &= \sum_{i=1}^n \{ -(\sigma_\delta^{-4}/2) [\sigma_\delta^2(1 - c\sigma_\delta^{-2}) - (x_i - \mu)^2 + 2cq_{i2}(x_i - \mu) - c^2q_{i2}^2] \}; \\ \frac{\partial \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2} &= \sum_{i=1}^n \{ -(1/2) [\sigma_x^{-2}(1 - c\sigma_x^{-2}) - c^2q_{i2}^2\sigma_x^{-4}] \}; \\ \frac{\partial \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2} &= \sum_{i=1}^n \{ -(\sigma_e^{-4}/2) [\sigma_e^2(1 - c\beta^2\sigma_e^{-2}) - (y_i - \beta\mu)^2 + 2c\beta q_{i2}(y_i - \beta\mu) \\ &\quad - \beta^2c^2q_{i2}^2] \}. \end{aligned}$$

Depois, obtivemos a matriz de informação observada dada por:

$$\begin{aligned} I_0(\boldsymbol{\theta}) &= -\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \\ &= - \begin{pmatrix} \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \beta} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \mu} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \sigma_\delta^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \sigma_x^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \sigma_e^2} \\ \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \beta} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \mu} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_\delta^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_x^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_e^2} \\ \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \beta} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \mu} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_\delta^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_x^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_e^2} \\ \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \beta} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \mu} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_\delta^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_x^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_e^2} \\ \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2 \partial \beta} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2 \partial \mu} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2 \partial \sigma_\delta^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2 \partial \sigma_x^2} & \frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2 \partial \sigma_e^2} \end{pmatrix}, \quad (5.13) \end{aligned}$$

em que,

$$\begin{aligned}
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \beta} &= -\sigma_e^{-2} [c - 2\beta^2 c^2 \sigma_e^{-2} - c \sigma_e^{-2} (y_i - 2\beta \mu - 2\beta c q_{i2})^2 + (q_{i2} c + \mu)^2]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \mu} &= c \sigma_e^{-2} \sigma_x^{-2} [y_i - 2\beta \mu - 2\beta c q_{i2}]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \sigma_\delta^2} &= -\beta c^2 \sigma_\delta^{-4} \sigma_e^{-2} [1 + \beta^{-1} c^{-1} (x_i - \mu - c q_{i2}) (y_i - 2\beta \mu - 2\beta c q_{i2})]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \sigma_x^2} &= -c^2 \sigma_x^{-4} \sigma_e^{-2} [\beta - q_{i2} (y_i - 2\beta \mu - 2\beta c q_{i2})]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \beta \partial \sigma_e^2} &= -\sigma_e^{-4} [\beta c (c q_{i2}^2 - (1 - \beta^2 c \sigma_e^{-2})) + \mu (y_i - \beta \mu) + c (y_i - 2\beta \mu \\
&\quad - 2\beta c q_{i2}) (\beta \sigma_e^{-2} (y_i - \beta \mu) + q_{i2} (1 - \beta^2 c \sigma_e^{-2}))]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \mu} &= -c \sigma_x^{-2} [\sigma_\delta^{-2} + \beta^2 \sigma_e^{-2}]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_\delta^2} &= -c \sigma_x^{-2} \sigma_\delta^{-4} [x_i - \mu - c q_{i2}]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_x^2} &= -c^2 q_{i2} \sigma_x^{-4} [\sigma_\delta^{-2} + \beta^2 \sigma_e^{-2}]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_e^2} &= -\beta c \sigma_x^{-2} \sigma_e^{-4} [y_i - \beta \mu - q_{i2} \beta c]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_\delta^2} &= \frac{\sigma_\delta^{-6}}{2} [\sigma_\delta^2 (1 - c \sigma_\delta^{-2})^2 - 2(1 - c \sigma_\delta^{-2}) (x_i - \mu - c q_{i2})^2]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_x^2} &= -\frac{\sigma_x^{-4} \sigma_\delta^{-4}}{2} [-c^2 + 2c^2 q_{i2} (x_i - \mu) - 2c^3 q_{i2}^2]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_e^2} &= -\frac{\sigma_\delta^{-4} \sigma_e^{-4}}{2} [-\beta^2 c^2 - 2(\beta c (y_i - \beta \mu) (x_i - \mu) - q_{i2} \beta^2 c^2 (x_i - \mu) \\
&\quad + \beta^2 c^3 q_{i2}^2 - \beta c^2 q_{i2} (y_i - \beta \mu))]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_x^2} &= \frac{\sigma_x^{-8}}{2} [(c - \sigma_x^2)^2 + 2c^2 q_{i2}^2 (c - \sigma_x^2)]; \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_e^2} &= -\frac{c^2 \sigma_e^{-4} \sigma_x^{-4}}{2} [-\beta^2 + 2\beta q_{i2} (y_i - \beta \mu) - 2q_{i2}^2 \beta^2 c] \quad e \\
\frac{\partial^2 \ell(\mathbf{Z}, \boldsymbol{\theta})}{\partial \sigma_e^2 \partial \sigma_e^2} &= -\frac{\beta^2 c \sigma_e^{-8} - \sigma_e^{-6}}{2} [\sigma_e^2 - \beta^2 c - 2(y_i - \beta \mu)^2 + 4\beta c q_{i2} (y_i - \beta \mu) - 2\beta^2 c^2 q_{i2}^2].
\end{aligned}$$

Na subseção seguinte, calcularemos os estimadores dos parâmetros.

## 5.1 Estimación dos parâmetros

Frequentemente na prática o logaritmo da função de máxima verossimilhança não pode ser maximizada analiticamente, em tais casos, poderia ser possível calcular iterativamente o estimador de máxima verossimilhança do vetor de parâmetros usando o procedimento de

maximização de Newton-Raphson ou alguma variante, desde que o número total de parâmetros do modelo não seja muito grande. Aqui, iremos utilizar o algoritmo EM descrito a seguir.

### 5.1.1 Algoritmo EM

O algoritmo de estimação “estimation-maximization”, EM, têm sido uma ferramenta popular na estatística aplicada proposto por [Dempster, Laird e Rubin \(1977\)](#). No seu artigo o algoritmo é usado com um exemplo para dados com observações faltantes distribuídos multinomialmente.

A ideia do algoritmo EM no modelo de regressão com erros de medição é considerar os verdadeiros valores das covariáveis como dados perdidos. Isto é, tomar como vetor de dados completos o vetor  $\mathbf{W}_i = (\xi_i, X_i, Y_i)^T = (\xi_i, \mathbf{Z}_i^T)^T$  e com o algoritmo estimando os valores de  $E[\xi_i | \mathbf{Z}, \boldsymbol{\theta}]$  e  $E[\xi_i^2 | \mathbf{Z}, \boldsymbol{\theta}]$  e maximizando a esperança condicional do logaritmo da função de verossimilhança dos dados completos dadas as observações para obter as estimativas dos parâmetros.

O modelo definido em (5.1) com os dados completos é dado da seguinte forma:

$$\begin{aligned} \mathbf{W}_i &\stackrel{ind.}{\sim} N_3(\boldsymbol{\mu}_W, \boldsymbol{\Sigma}), \quad i = 1, \dots, n, \text{ ou seja,} \\ \begin{pmatrix} \xi_i \\ X_i \\ Y_i \end{pmatrix} &\stackrel{ind.}{\sim} N_3 \left( \begin{bmatrix} \mu \\ \mu \\ \beta\mu \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_x^2 & \beta\sigma_x^2 \\ \sigma_x^2 & \sigma_x^2 + \sigma_\delta^2 & \beta\sigma_x^2 \\ \beta\sigma_x^2 & \beta\sigma_x^2 & \beta^2\sigma_x^2 + \sigma_e^2 \end{bmatrix} \right), \quad i = 1, \dots, n \end{aligned} \quad (5.14)$$

e com o determinante e a inversa da matriz de variâncias e covariâncias, dados por,

$$|\boldsymbol{\Sigma}| = \sigma_\delta^2 \sigma_e^2 \sigma_x^2 \quad e \quad \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \frac{1}{\sigma_\delta^2} + \frac{\beta^2}{\sigma_e^2} + \frac{1}{\sigma_x^2} & -\frac{1}{\sigma_\delta^2} & -\frac{\beta}{\sigma_e^2} \\ -\frac{1}{\sigma_\delta^2} & \frac{1}{\sigma_\delta^2} & 0 \\ -\frac{\beta}{\sigma_e^2} & 0 & -\frac{\beta}{\sigma_e^2} \end{pmatrix}. \quad (5.15)$$

Desta forma, a log-verossimilhança dos dados completos é dado por:

$$\begin{aligned} \ell(\mathbf{W}, \boldsymbol{\theta} | \mathbf{Z}) &= \log L(\mathbf{W}, \boldsymbol{\theta} | \mathbf{Z}) \\ &= \log \prod_{i=1}^n (2\pi)^{-\frac{3}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{W}_i - \boldsymbol{\mu}_W)^T \boldsymbol{\Sigma}^{-1} (\mathbf{W}_i - \boldsymbol{\mu}_W)} \\ &= \sum_{i=1}^n \log(2\pi)^{-\frac{3}{2}} + \sum_{i=1}^n \log |\boldsymbol{\Sigma}|^{-\frac{1}{2}} + \sum_{i=1}^n \log e^{-\frac{1}{2}(\mathbf{W}_i - \boldsymbol{\mu}_W)^T \boldsymbol{\Sigma}^{-1} (\mathbf{W}_i - \boldsymbol{\mu}_W)} \\ &= -\frac{3n}{2} \log(2\pi) - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{W}_i - \boldsymbol{\mu}_W)^T \boldsymbol{\Sigma}^{-1} (\mathbf{W}_i - \boldsymbol{\mu}_W) \\ &= -\frac{3n}{2} \log(2\pi) - \frac{n}{2} \log |\boldsymbol{\Sigma}| \\ &\quad - \frac{1}{2} \sum_{i=1}^n (\mathbf{W}_i - \boldsymbol{\mu}_W)^T \frac{Adj(Cof(\boldsymbol{\Sigma}))}{|\boldsymbol{\Sigma}|} (\mathbf{W}_i - \boldsymbol{\mu}_W), \text{ e portanto} \end{aligned}$$

$$\begin{aligned} \ell(\mathbf{W}, \boldsymbol{\theta} | \mathbf{Z}) &= -\frac{3n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma_\delta^2 \sigma_e^2 \sigma_x^2) \\ &\quad - \frac{1}{2} \sum_{i=1}^n \left[ \frac{(\xi_i - \mu)^2}{\sigma_x^2} + \frac{(\xi_i - X_i)^2}{\sigma_\delta^2} + \frac{(Y_i - \beta \xi_i)^2}{\sigma_e^2} \right] \end{aligned} \quad (5.16)$$

Passo-E: esperança condicional. Considerando as propriedades da distribuição normal multivariada, o passo-E consiste na obtenção de:

$$\begin{aligned} \hat{\xi}_i^{(r+1)} &= \mathbb{E}[\xi_i | \mathbf{Z}, \boldsymbol{\theta}] = \frac{\frac{\mu^{(r)}}{\hat{\sigma}_x^2} + \frac{X_i}{\hat{\sigma}_\delta^2} + \frac{\hat{\beta}^{(r)} Y_i}{\hat{\sigma}_e^2}}{\frac{1}{\hat{\sigma}_x^2} + \frac{1}{\hat{\sigma}_\delta^2} + \frac{(\hat{\beta}^{(r)})^2}{\hat{\sigma}_e^2}} \quad e \\ \hat{\xi}_i^2{}^{(r+1)} &= \mathbb{E}[\xi_i^2 | \mathbf{Z}, \boldsymbol{\theta}] = (\hat{\xi}_i^{(r+1)})^2 + \frac{1}{\frac{1}{\hat{\sigma}_x^2} + \frac{1}{\hat{\sigma}_\delta^2} + \frac{(\hat{\beta}^{(r)})^2}{\hat{\sigma}_e^2}}. \end{aligned}$$

Passo-M: maximização da esperança condicional. Do logaritmo da verossimilhança dos dados completos, temos:

$$\begin{aligned} \hat{\mu}^{(r+1)} &= \frac{\sum_{i=1}^n \hat{\xi}_i^{(r+1)}}{n} \\ \hat{\sigma}_x^2{}^{(r+1)} &= \frac{\sum_{i=1}^n ((\hat{\xi}_i^2)^{(r+1)} - 2\hat{\mu}^{(r+1)} \hat{\xi}_i^{(r+1)} + (\hat{\mu}^{(r+1)})^2)}{n} \\ \hat{\sigma}_\delta^2{}^{(r+1)} &= \frac{\sum_{i=1}^n ((\hat{\xi}_i^2)^{(r+1)} - 2X_i \hat{\xi}_i^{(r+1)} + X_i^2)}{n} \\ \hat{\beta}^{(r+1)} &= \frac{\sum_{i=1}^n \hat{\xi}_i^{(r+1)} Y_i}{\sum_{i=1}^n (\hat{\xi}_i^2)^{(r+1)}} \\ \hat{\sigma}_e^2{}^{(r+1)} &= \frac{\sum_{i=1}^n ((\hat{\beta}^{(k+1)})^2 (\hat{\xi}_i^2)^{(r+1)} - 2\hat{\beta}^{(k+1)} \hat{\xi}_i^{(r+1)} + Y_i^2)}{n} \end{aligned}$$

Na subsecção seguinte, definiremos alguns tipos de perturbações.

## 5.2 Esquemas de perturbações

Nesta subsecção foram obtidos as matrizes  $\Delta = \left. \frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T} \right|_{\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}, \boldsymbol{\omega} = \boldsymbol{\omega}_0}$  que são matrizes necessárias para a utilização do método de influência local de [Cook \(1986\)](#). As seguintes perturbações foram consideradas: ponderação de casos, perturbação na resposta, perturbação na covariável e perturbação na variância. Estes resultados serão utilizados na metodologia de [ILPPF](#).

### 5.2.1 Ponderação de casos

Nesta subsecção vamos considerar o esquema de perturbação do modelo dado pela ponderação de casos no modelo definido em (5.1). Para isto seja  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$  o vetor de perturbações, em que  $\boldsymbol{\theta} = (\beta, \mu, \sigma_\delta^2, \sigma_x^2, \sigma_e^2)^T$  denota o vetor de parâmetros do modelo. O logaritmo da função verossimilhança perturbada é dado por:

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\ &= \sum_{i=1}^n \omega_i \log f(\mathbf{Z}_i, \boldsymbol{\theta}) \\ &= -\log(2\pi) \sum_{i=1}^n \omega_i - \frac{1}{2} \log |\mathbf{V}| \sum_{i=1}^n \omega_i - \frac{1}{2} \sum_{i=1}^n \omega_i (\mathbf{Z}_i - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i - \mathbf{m}). \end{aligned}$$

Diferenciando  $l_{\boldsymbol{\omega}} = \ell(\boldsymbol{\theta}|\boldsymbol{\omega})$  em relação ao vetor  $\boldsymbol{\theta}$ , temos que:

$$\frac{\partial l_{\boldsymbol{\omega}}}{\partial \boldsymbol{\theta}} = \left( \frac{\partial l_{\boldsymbol{\omega}}}{\partial \beta}, \frac{\partial l_{\boldsymbol{\omega}}}{\partial \mu}, \frac{\partial l_{\boldsymbol{\omega}}}{\partial \sigma_\delta^2}, \frac{\partial l_{\boldsymbol{\omega}}}{\partial \sigma_x^2}, \frac{\partial l_{\boldsymbol{\omega}}}{\partial \sigma_e^2} \right)^T, \quad \text{onde} \quad (5.17)$$

$$\begin{aligned} \frac{\partial l_{\boldsymbol{\omega}}}{\partial \beta} &= \sum_{i=1}^n \omega_i \{ \sigma_e^{-2} [-c\beta + \mu(y_i - \beta\mu) + q_{i2}(y_i - 2\beta\mu)c - q_{i2}^2 c^2 \beta] \}; \\ \frac{\partial l_{\boldsymbol{\omega}}}{\partial \mu} &= \sum_{i=1}^n \omega_i \{ q_{i2} - cq_{i2}(\sigma_\delta^{-2} + \beta^2 \sigma_e^{-2}) \}; \\ \frac{\partial l_{\boldsymbol{\omega}}}{\partial \sigma_\delta^2} &= \sum_{i=1}^n \omega_i \{ -(\sigma_\delta^{-4}/2) [\sigma_\delta^2(1 - c\sigma_\delta^{-2}) - (x_i - \mu)^2 + 2cq_{i2}(x_i - \mu) - c^2 q_{i2}^2] \}; \\ \frac{\partial l_{\boldsymbol{\omega}}}{\partial \sigma_x^2} &= \sum_{i=1}^n \omega_i \{ -(1/2) [\sigma_x^{-2}(1 - c\sigma_x^{-2}) - c^2 q_{i2}^2 \sigma_x^{-4}] \} \text{ e} \\ \frac{\partial l_{\boldsymbol{\omega}}}{\partial \sigma_e^2} &= \sum_{i=1}^n \omega_i \{ -(\sigma_e^{-4}/2) [\sigma_e^2(1 - c\beta^2 \sigma_e^{-2}) - (y_i - \beta\mu)^2 + 2c\beta q_{i2}(y_i - \beta\mu) \\ &\quad - \beta^2 c^2 q_{i2}^2] \}. \end{aligned} \quad (5.18)$$

Diferenciando agora, o resultado acima, em relação ao vetor  $\boldsymbol{\omega}^T$ , segue que

$$\begin{aligned} \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \beta \partial \omega_i} &= \sigma_e^{-2} [-c\beta + \mu(y_i - \beta\mu) + q_{i2}(y_i - 2\beta\mu)c - q_{i2}^2 c^2 \beta]; \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \mu \partial \omega_i} &= q_{i2} - cq_{i2}(\sigma_\delta^{-2} + \beta^2 \sigma_e^{-2}); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_\delta^2 \partial \omega_i} &= -(\sigma_\delta^{-4}/2) [\sigma_\delta^2(1 - c\sigma_\delta^{-2}) - (x_i - \mu)^2 + 2cq_{i2}(x_i - \mu) - c^2 q_{i2}^2]; \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_x^2 \partial \omega_i} &= -(1/2) [\sigma_x^{-2}(1 - c\sigma_x^{-2}) - c^2 q_{i2}^2 \sigma_x^{-4}] \text{ e} \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_e^2 \partial \omega_i} &= -(\sigma_e^{-4}/2) [\sigma_e^2(1 - c\beta^2 \sigma_e^{-2}) - (y_i - \beta\mu)^2 + 2c\beta q_{i2}(y_i - \beta\mu) - \beta^2 c^2 q_{i2}^2]; \\ & i = 1, \dots, n, \end{aligned}$$

e por tanto,  $\Delta = \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}$ , com os elementos da  $\frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^T}$  dada acima e  $\boldsymbol{\omega}_0 = (1, \dots, 1)$ .

Temos também que  $\ddot{L} = I_0(\boldsymbol{\theta})$  (5.13).

### 5.2.2 Perturbação na covariável

Consideramos a perturbação na covariável  $X_i$  com

$$X_{i\omega} = X_i + S_X \omega_i,$$

$i = 1, \dots, n$ , onde  $S_X$  é um fator de escala, por exemplo, o desvio padrão amostral de  $X$ :

$$S_X = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2}.$$

A log-verossimilhança do modelo perturbado neste caso é dada por:

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\ &= \sum_{i=1}^n \log f(\mathbf{Z}_i^*, \boldsymbol{\theta}) \\ &= -n \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i^* - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i^* - \mathbf{m}), \end{aligned}$$

com  $\mathbf{Z}_i^* = (X_{i\omega}, Y_i)^T = (X_i + S_X \omega_i, Y_i)^T$ ,  $i = 1, \dots, n$  e o vetor de não perturbação é dado por  $\boldsymbol{\omega}_0 = (0, \dots, 0)_n^T$ . Utilizando  $\mathbf{a}$ ,  $\mathbf{m}$ ,  $\boldsymbol{\tau}$ ,  $\mathbf{V}$ ,  $c$  e  $q_i^* = (\mathbf{Z}_i^* - \boldsymbol{\mu}\mathbf{a})^T \mathbf{V}^{-1} (\mathbf{Z}_i^* - \boldsymbol{\mu}\mathbf{a})$  como definidos em (5.4), (5.6), (5.8) e (5.11) obtemos os elementos da matriz  $\Delta$ , com  $i = 1, \dots, n$ , dados por:

$$\begin{aligned} \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \beta \partial \omega_i} &= S_X \sigma_e^{-2} \sigma_{\delta}^{-2} (c(Y_i - 2\beta u) - 2\beta c^2 q_{i2}); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \mu \partial \omega_i} &= S_X \sigma_{\delta}^{-2} (1 - c(\sigma_{\delta}^{-2} + \beta^2 \sigma_e^{-2})); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_{\delta}^2 \partial \omega_i} &= -\sigma_{\delta}^{-4} S_X (-(-u + S_X \omega_i + Xi) + c(\sigma_{\delta}^{-2}(-u + S_X \omega_i + Xi) + q_{i2}) \\ &\quad - c^2 q_{i2} \sigma_{\delta}^{-2}); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_x^2 \partial \omega_i} &= c^2 q_{i2} S_X \sigma_{\delta}^{-2} \sigma_x^{-4} \text{ e} \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_e^2 \partial \omega_i} &= -\sigma_e^{-4} \sigma_{\delta}^{-2} S_X (\beta c(Y_i - \beta u) - \beta^2 c^2 q_{i2}). \end{aligned}$$

### 5.2.3 Perturbação na variável resposta

Consideramos a perturbação na variável resposta  $Y_i$  com

$$Y_{i\omega} = Y_i + S_Y \omega_i,$$

$i = 1, \dots, n$ , onde  $S_Y$  é um fator de escala, por exemplo, o desvio padrão amostral de  $Y$ :

$$S_Y = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2}.$$

A log-verossimilhança do modelo perturbado neste caso é dada por:

$$\begin{aligned}\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\ &= \sum_{i=1}^n \log f(\mathbf{Z}_i^*, \boldsymbol{\theta}) \\ &= -n \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i^* - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_i^* - \mathbf{m}),\end{aligned}$$

com  $\mathbf{Z}_i^* = (X_i, Y_{i\omega})^T = (X_i, Y_i + S_Y \omega_i)^T$ ,  $i = 1, \dots, n$  e o vetor de não perturbação é dado por  $\boldsymbol{\omega}_0 = (0, \dots, 0)_n^T$ . Utilizando  $\mathbf{a}$ ,  $\mathbf{m}$ ,  $\boldsymbol{\tau}$ ,  $\mathbf{V}$ ,  $c$  e  $q_i^* = (\mathbf{Z}_i^* - \boldsymbol{\mu}\mathbf{a})^T \mathbf{V}^{-1} (\mathbf{Z}_i^* - \boldsymbol{\mu}\mathbf{a})$  como definidos em (5.4), (5.6), (5.8) e (5.11) obtemos os elementos da matriz  $\Delta$ , com  $i = 1, \dots, n$ , dados por:

$$\begin{aligned}\frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \beta \partial \omega_i} &= S_Y \sigma_e^{-2} (\boldsymbol{\mu} + \beta c \sigma_e^{-2} (Y_i + S_Y \omega_i - 2\beta \boldsymbol{\mu}) + c q_{i2}) \\ &\quad - 2\beta^2 c^2 \sigma_e^{-2} q_{i2}); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \boldsymbol{\mu} \partial \omega_i} &= \beta S_Y \sigma_e^{-2} (1 - c(\sigma_{\delta}^{-2} + \beta^2 \sigma_e^{-2})); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_{\delta}^2 \partial \omega_i} &= -\beta S_Y \sigma_{\delta}^{-4} \sigma_e^{-2} (c(-\boldsymbol{\mu} + X_i) - c^2 q_{i2}); \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_x^2 \partial \omega_i} &= \beta c^2 S_Y q_{i2} \sigma_x^{-4} \sigma_e^{-2} \text{ e} \\ \frac{\partial^2 l_{\boldsymbol{\omega}}}{\partial \sigma_e^2 \partial \omega_i} &= -S_Y \sigma_e^{-4} (-(Y_i + S_Y \omega_i - \beta \boldsymbol{\mu}) + \beta^2 c \sigma_e^{-2} (Y_i + S_Y \omega_i \\ &\quad - \beta \boldsymbol{\mu}) + \beta c q_{i2} - \beta^3 c^2 q_{i2} \sigma_e^{-2}).\end{aligned}$$

### 5.2.4 Perturbação nas variâncias dos erros

Consideramos a perturbação na variância dos erros de medição  $\sigma_{\delta}^2$  com

$$\begin{aligned}\sigma_{\delta\omega_i}^2 &= \frac{\sigma_{\delta}^2}{\omega_i} \\ \text{e} \\ \sigma_{e\omega_i}^2 &= \frac{\sigma_e^2}{\omega_i},\end{aligned}\tag{5.19}$$

$i = 1, \dots, n$ .

A log-verossimilhança do modelo perturbado neste caso é dada por

$$\begin{aligned}\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\ &= -n \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log |\mathbf{V}_i^*| - \frac{1}{2} \sum_{i=1}^n (\mathbf{Z}_i - \mathbf{m})^T (\mathbf{V}_i^*)^{-1} (\mathbf{Z}_i - \mathbf{m}).\end{aligned}$$

com  $i = 1, \dots, n$  e o vetor de não perturbação é dado por  $\boldsymbol{\omega}_0 = (1, \dots, 1)_n^T$ . Utilizando  $\mathbf{a}$ ,  $\mathbf{m}$ ,  $\boldsymbol{\tau}_i^* = \begin{pmatrix} \frac{\sigma_{\delta}^2}{\omega_i} & 0 \\ 0 & \frac{\sigma_e^2}{\omega_i} \end{pmatrix}$ ,  $\mathbf{V}_i^* = \sigma_x^2 \mathbf{a}\mathbf{a}^T + \boldsymbol{\tau}_i^* = \begin{pmatrix} \sigma_x^2 + \frac{\sigma_{\delta}^2}{\omega_i} & \beta \sigma_x^2 \\ \beta \sigma_x^2 & \beta^2 \sigma_x^2 + \frac{\sigma_e^2}{\omega_i} \end{pmatrix}$ ,  $c_i^* = \sigma_x^2 (1 + \sigma_x^2 \mathbf{a}^T (\boldsymbol{\tau}_i^*)^{-1} \mathbf{a})^{-1}$

e  $q_i^* = (Z_i - \mu \mathbf{a})^T (\mathbf{V}_i^*)^{-1} (Z_i - \mu \mathbf{a})$  como definidos em (5.4), (5.6), (5.8) e (5.11) obtemos os elementos da matriz  $\Delta$ , com  $i = 1, \dots, n$ , dados por:

$$\begin{aligned} \frac{\partial^2 l_{\omega}}{\partial \beta \partial \omega_i} &= -\sigma_e^{-2} (\{\beta c - \mu(Y_i - \beta \mu) + \beta c^2 q_{i2}^2 - c q_{i2}(Y_i - 2\beta \mu)\} + \omega_i \{-\beta(1/\sigma_{\delta}^2 + \beta^2/\sigma_e^2)c^2 + \{2\beta c^2 q_{i2}^2 - (Y_i - 2\beta \mu)c q_{i2}\} \{-c(1/\sigma_{\delta}^2 + \beta^2/\sigma_e^2) + 1/\omega_i\}\}); \\ \frac{\partial^2 l_{\omega}}{\partial \mu \partial \omega_i} &= q_{i2}/\omega_i - c q_{i2}(1/\sigma_{\delta}^2 + b^2/\sigma_e^2) \{2 - c \omega_i(1/\sigma_{\delta}^2 + b^2/\sigma_e^2)\}; \\ \frac{\partial^2 l_{\omega}}{\partial \sigma_{\delta}^2 \partial \omega_i} &= -(\sigma_{\delta}^{-4}/2)(c^2 q_{i2}^2 - (X_i - \mu)^2 + c - \{c - 2c q_{i2}(X_i - \mu) + 2c^2 q_{i2}^2\} \{2 - c \omega_i(1/\sigma_{\delta}^2 + b^2/\sigma_e^2)\}); \\ \frac{\partial^2 l_{\omega}}{\partial \sigma_x^2 \partial \omega_i} &= -(c/2) \sigma_x^{-4} \sigma_{\delta}^{-2} (\sigma_x^2 \{-c(\beta^2 \sigma_e^{-2} + \omega_i \sigma_{\delta}^{-2}) + 1\} - 2c q_{i2}(X_i - \mu - c q_{i2})) \text{ e} \\ \frac{\partial^2 l_{\omega}}{\partial \sigma_e^2 \partial \omega_i} &= -(c/2) \sigma_e^{-4} \sigma_{\delta}^{-2} \{-c \sigma_e^2 (\sigma_x^{-2} + \omega_i \sigma_{\delta}^{-2}) + \sigma_e^2 - 2\beta(\beta c q_{i2} - (Y_i - \beta \mu))(X_i - \mu - c q_{i2})\}. \end{aligned}$$

### 5.2.5 Aplicação

A seguir aplicaremos a metodologia proposta de **ILPPF** nos dois conjuntos de dados odontológicos, Líquido de bochecho A, B e controle e Escova de dente experimental e convencional, definidos na Introdução (1.1.1 e 1.2.2) utilizando as perturbações definidas na subsecção 5.2. No processo da metodologia de influência local com **procura passo a frente** descrito no Capítulo 4.1, em cada passo do algoritmo, foram considerados subconjuntos de tamanho  $m_0$ , no passo seguinte foram considerados subconjuntos de tamanho  $m_0 + 1$  e assim sucessivamente foram adicionadas observações até chegar a completar o número total de dados que é  $n$ , ou seja, foi ajustado o modelo para cada um dos  $\binom{n}{m_0}$  subconjuntos de tamanho  $m_0$ , para cada um dos  $\binom{n}{m_0+1}$  subconjuntos de tamanho  $m_0 + 1$ , e assim sucessivamente até  $\binom{n}{n}$  conjunto de tamanho  $n$ , desde que o número de subconjuntos fosse menores ou iguais a 1000. Em cada passo, foi selecionado o subconjunto de acordo com a metodologia proposta e descrita na Seção 4.1.

#### 5.2.5.1 Dados do líquido de bochecho (antisséptico bucal)

Considerando o conjunto de dados do líquido de bochecho apresentado em **Hadgu e Koch (1999)** (Apêndice 1.2), foram coletados os índices de placa bacteriana no início do estudo, após 3 meses do início do estudo e após 6 meses do início do estudo com a utilização dos líquidos de bochecho A, B e controle.

Podemos construir 6 modelos de regressão com erros nas variáveis com intercepto nulo, onde a variável resposta em cada um dos modelos são dadas pelo índice de placa bacteriana

após 3(6) meses do início do estudo com o uso dos líquidos de bochecho A, B e controle. A covariável é o índice de placa bacteriana no início do estudo. Vamos fazer a análise de alguns destes conjuntos de dados.

Primeiro, vamos considerar o líquido de bochecho experimental A após 6 meses do início do estudo e a perturbação na variável explanatória.

### 5.2.5.2 Líquido de bochecho A após 6 meses do início do estudo - perturbação na variável explanatória.

Sejam:

$x$ : índice de placa bacteriana no início do estudo e

$y$ : índice de placa bacteriana após 6 meses do início do estudo com o uso do líquido de bochecho A.

Primeiramente, foi feito o diagnóstico de influência local de Cook (1986). Obtivemos a curvatura máxima  $C_{max} = 7.68$  e foi feito o gráfico do valor absoluto do autovetor correspondente ao maior autovalor e o índice das observações. Na Figura 9 podemos observar que não tem pontos que se destacam das demais observações, ou seja, não há ponto influente.

A seguir foi feito o gráfico de ILPPF. Seja  $\mathbf{Z} = (\mathbf{x}, \mathbf{y})$  com  $n = 33$ , assim  $\mathbf{Z}$  é  $33 \times 2$  e a escolha do subconjunto inicial pode ser realizada pela numeração exaustiva de todas as  $\binom{n}{m_0}$  possíveis subconjuntos de tamanho  $m_0 = 10$ ,  $\mathbf{S}_{i_1, \dots, i_{10}}^{(10)} \equiv \{\mathbf{z}_{i_1}, \dots, \mathbf{z}_{i_{10}}\}$  para estimar os parâmetros, em que  $\mathbf{z}_{i_1}^T$  é a  $i_1$ -ésima linha de  $\mathbf{Z}$ , para  $1 \leq i_1, \dots, i_{10} \leq 33$  e  $i_j \neq i_l$ . E seja  $\mathbf{S}^{(10)}$  o conjunto de todas as amostras  $\mathbf{S}_{i_1, \dots, i_{10}}^{(10)}$  obtidas. Foram calculadas as estimativas dos parâmetros para cada subconjunto de tamanho  $m_0 = 10$  contida em  $\mathbf{S}^{(10)}$  e utilizamos estas estimativas de cada subconjunto e o conjunto de dados todo ( $n$ ) para obtermos o autovetor associado ao maior autovalor da matriz simétrica semidefinida positiva  $\tilde{\mathbf{F}}$  como foi descrito na seção 3.1, para cada um dos subconjuntos. Depois, foi ordenado os autovetores obtidos e tomado o subconjunto  $\mathbf{S}_{i_1, \dots, i_{10}}^{(10)*}$  com a menor mediana de todos os autovetores,  $l_{m_0}^*$ , e o correspondente autovetor  $l_{m_0}^*$ . No passo seguinte, consideramos  $\mathbf{S}_{i_1, \dots, i_{11}}^{(11)}$  subconjuntos de tamanho  $m = 11$  e estimamos os parâmetros de cada um dos subconjuntos e utilizamos estas estimativas e os  $n$  dados para obtermos os autovetores associados ao maior autovalor da matriz simétrica semidefinida positiva

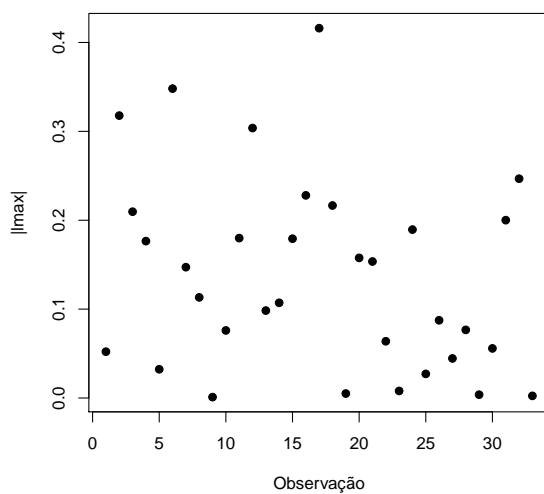


Figura 9 – Perturbação na variável explanatória: gráfico do valor absoluto do autovetor,  $l_{max}$ , associado ao maior autovalor com o índice das observações. Líquido de bochecho A após 6 meses.

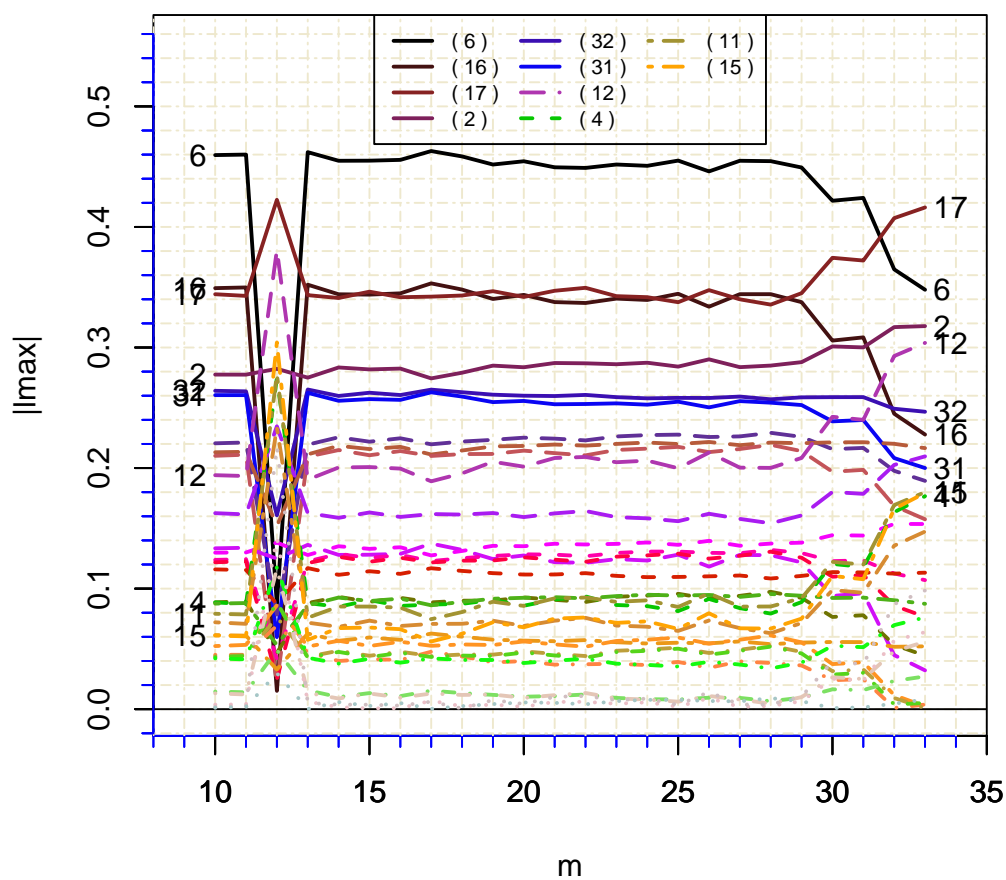


Figura 10 – Perturbação na variável explanatória: gráfico de ILPPF. Líquido de bochecho A após 6 meses.

$\ddot{F}$ , de cada subconjunto e ordenamos estes autovetores. Tomamos ou escolhemos o subconjunto com a menor mediana,  $\mathbf{l}_{m_0+1}^*$  e o autovetor correspondente  $\mathbf{l}_{m_0+1}^*$ , e assim sucessivamente até chegar a  $n = 33$  observações. Desta forma, obtemos  $\mathbf{l}^* = (\mathbf{l}_{m_0}^*, \mathbf{l}_{m_0+1}^*, \dots, \mathbf{l}_n^*)$ . Foi feito o gráfico com todos os autovetores escolhidos (Figura 10) e podemos observar este gráfico para verificar se existe algum dado mascarado durante a evolução do algoritmo de *ILPPF*.

No gráfico da Figura 10 foi feito o gráfico do valor absoluto de  $\mathbf{l}^*$  em cada passo da *ILPPF*. Na abscisa, temos os valores de  $m$  começando com o valor inicial que é  $m_0 = 10$ , até completar os dados que é  $n = 33$  observações. Na ordenada, colocamos os valores absolutos dos componentes de  $\mathbf{l}^*$  referente a cada subconjunto obtido.

Observe que, se fizermos um gráfico onde na abscisa colocamos as observações e na ordenada o último passo do algoritmo de *ILPPF*, ou seja, os pontos que estão na ordenada da Figura 10 para o valor 33 na abscissa ( $\mathbf{l}_n^*$ ), vamos obter a Figura 9. Desta forma, podemos acompanhar o que acontece com o valor absoluto do  $\mathbf{l}_{max}$  conforme o tamanho da amostra considerada aumenta.

No gráfico da Figura 10 observamos que existe um dado na iteração inicial que é a observação 6 que se destaca das demais observações e depois em quase toda a evolução da metodologia de *ILPPF*, no entanto, no final das iterações a influência dessa observação decresce. A observação 6 corresponde a observação com o maior índice de placa bacteriana no início do estudo.

Apesar da observação 17 se destacar em primeiro lugar no gráfico de influência de Cook (1986), durante a evolução do gráfico de *ILPPF* ela se encontra como o segundo ou o terceiro valor mais influente e no final quando  $m = 29$  começa a subir e na última iteração se encontra como o mais influente. Por outro lado, apesar da observação 6 se destacar em primeiro lugar em quase toda a evolução do processo de *ILPPF* ela começa a decair em  $m = 29$  e passa a ser o segundo ponto a mais influente.

A observação 16 se encontra na maior parte da evolução do processo de *ILPPF* como sendo a observação que está em segundo ou terceiro lugar como mais influente, no entanto, a partir de  $m = 29$  decai e fica mascarada.

Na Figura 11 temos o gráfico de afastamento pela verossimilhança. Este gráfico foi feito com  $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{d}$ , vetor de perturbação, onde  $\boldsymbol{\omega}_0$  é o vetor de não perturbação,  $a$  é uma constante e  $\mathbf{d}$  é o vetor representando a direção. Neste caso, foi construído na direção  $\mathbf{d}_{max} = \mathbf{l}_{max}$  e também na direção  $\mathbf{d} = \mathbf{d}_j$  com  $j = 6, 16$  e  $17$ , ou seja,  $\mathbf{d}_6, \mathbf{d}_{16}$  e  $\mathbf{d}_{17}$ , onde  $\mathbf{d}_j$  é um vetor de zeros exceto na  $j$ -ésima posição onde é 1 e para valores de  $a$  que pertencem aos números reais, tendo um valor mínimo local em  $a = 0$ . No gráfico da Figura 11 para  $LD(\boldsymbol{\omega}(a))$  considerou-se  $-1 \leq a \leq 1$ . Fazendo o gráfico com os 4 vetores  $LD(\boldsymbol{\omega}(a)) = LD(\boldsymbol{\omega}_0 + a\mathbf{d}_6), LD(\boldsymbol{\omega}_0 + a\mathbf{d}_{16}), LD(\boldsymbol{\omega}_0 + a\mathbf{d}_{17})$  e  $LD(\boldsymbol{\omega}_0 + a\mathbf{d}_{max})$  no eixo das ordenadas, obtivemos o gráfico da Figura 11.

Observe que o dado que produz a maior alteração no valor de  $LD$ , conforme passamos de

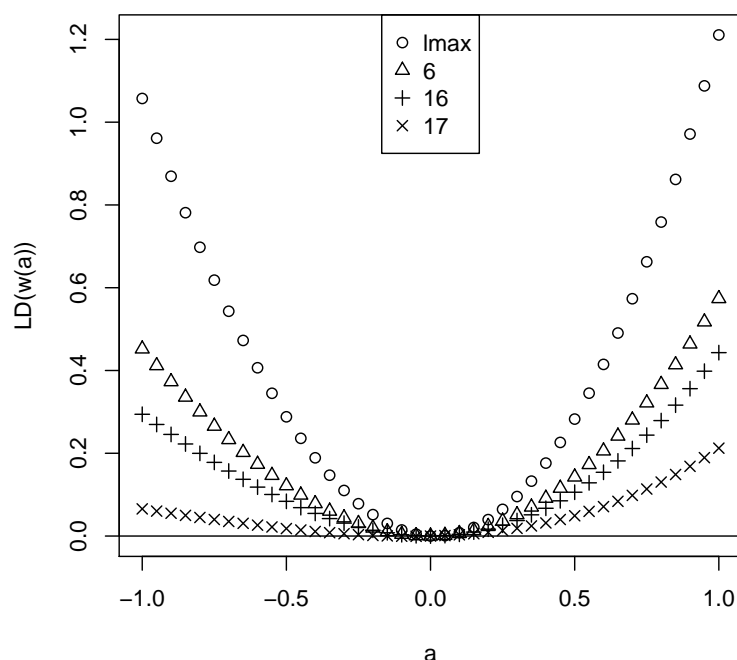


Figura 11 – Perturbação na variável explanatória: gráfico de afastamento pela verossimilhança. Líquido de bochecho A após 6 meses.

$\boldsymbol{\omega} = \boldsymbol{\omega}_0$  para  $\boldsymbol{\omega} = \boldsymbol{\omega}_0 + ad$  é a observação 6 (dentre as observações 6, 16 e 17 que estão sendo analisadas) que é a observação que se destacou durante quase toda a evolução como sendo o mais influente. Mas terminou como sendo o ponto que se destaca em 2º lugar. Analisando o gráfico de influência local de Cook (1986) só teríamos esta informação de que se destaca em segundo lugar.

Considerando o gráfico  $LD$ , podemos ver que a observação que aparece como sendo o 2º mais influente é a observação 16 que foi mascarada. Através do gráfico de influência local de Cook (1986) não seria possível detectar esta observação.

Note também que a observação 17 é o que possui o terceiro menor índice de placa bacteriana e aparece como ponto influente no gráfico de influência local de Cook (1986). Esta observação é a que produz a menor mudança em  $LD$  dentre as observações 6, 16 e 17.

Nos boxplots seguintes serão utilizados o boxplot assimétrico ou boxplot ajustado que foi proposto por Hubert e Vandervieren (2008b) que usa o *medcouple* (MC) introduzido por Brys, Hubert e Struyf (2004). Hubert e Vandervieren (2008b) concluem com base em resultados obtidos usando dados reais e simulados que o boxplot ajustado em distribuições assimétricas possui a vantagem de melhorar a distinção entre observações regulares e valores atípicos. Aqui, iremos utilizar para ter uma melhor visualização das observações que se destacam no decorrer da evolução da ILPPF.

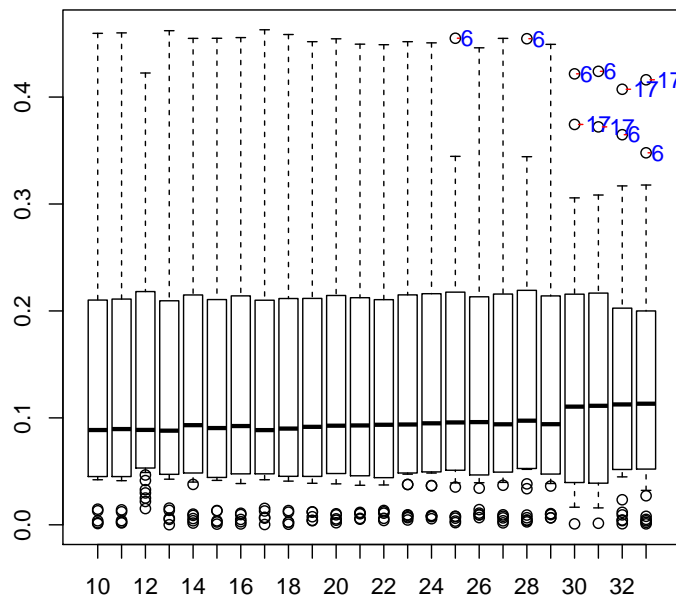


Figura 12 – Perturbação na variável explanatória: boxplot da *ILPPF*. Líquido de bochecho A após 6 meses.

EMV	$\hat{\beta}$	$\hat{\mu}$	$\hat{\sigma}_\delta^2$	$\hat{\sigma}_x^2$	$\hat{\sigma}_e^2$
Dados completos	0.501	2.57	0.007	0.12	0.432
sem a observação 6	0.491	2.539	0.007	0.091	0.433
	2	1	1	24	0
sem a observação 16	0.488	2.544	0.007	0.101	0.423
	3	1	6	16	2
sem a observação 17	0.518	2.558	0.006	0.12	0.4
	3	0	14	0	8

Tabela 8 – Perturbação na variável explanatória: mudança relativa na estimativa dos parâmetros. Líquido de bochecho A após 6 meses.

Na Figura 12, foram feitos os boxplots dos valores que compõem  $l^*$  obtidos em cada iteração da metodologia de *ILPPF*. Observa-se que os dados 6 e 17 aparecem como dados atípicos. Na figura 13 foi construído o heatmap. O heatmap é um gráfico onde valores maiores são representados por cores mais fortes. Podemos ver que as observações 6, 16 e 17 são as mais escuras e se destacam das demais observações em quase toda a evolução da *procura passo a frente* (possui cores mais fortes em quase toda a evolução). A vantagem do heatmap é que dá para ter uma visão geral de cada uma das observações no decorrer da evolução de *ILPPF*. Por exemplo, as observações 23 e 25 são observações que menos se destacam em quase toda a evolução de *ILPPF*. Considerando agora a análise de influência global, temos a Tabela 8.

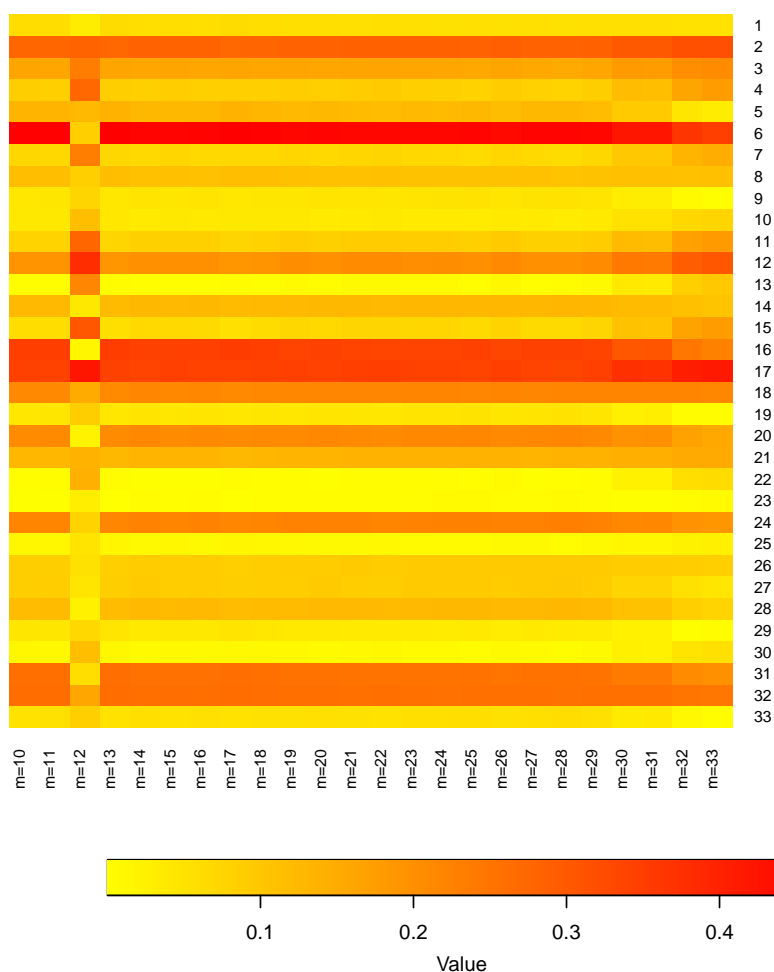


Figura 13 – Perturbação na variável explanatória: heatmap de ILPPF. Líquido de bochecho A após 6 meses.

Observando a Tabela 8, as observações 6 e 16 ocasionam a maior mudança global na estimativa do parâmetro  $\sigma_x^2$ , é importante enfatizar que a ordem das observações que causam a maior mudança local e global, não necessariamente são as mesmas.

### 5.2.5.3 Líquido de bochecho controle após 3 meses do início do estudo - Ponderação de casos

Sejam:

$x$ : índice de placa bacteriana no início do estudo e

$y$ : índice de placa bacteriana após 3 meses do início do estudo com o uso do líquido de bochecho controle.

Primeiramente, foi feito o diagnóstico de influência local de Cook (1986). Obtivemos a curvatura máxima  $C_{max} = 8.4$  e foi feito o gráfico do valor absoluto do autovetor correspondente

ao maior autovalor e o índice das observações.

Nenhuma observação se destaca em particular (Figura 14).

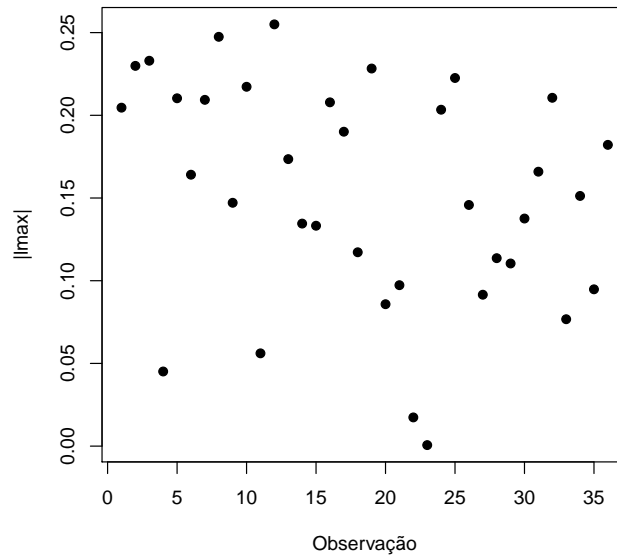


Figura 14 – Ponderação de casos: gráfico do valor absoluto do autovetor, associado ao maior autovalor com o índice das observações. Líquido de controle após 3 meses.

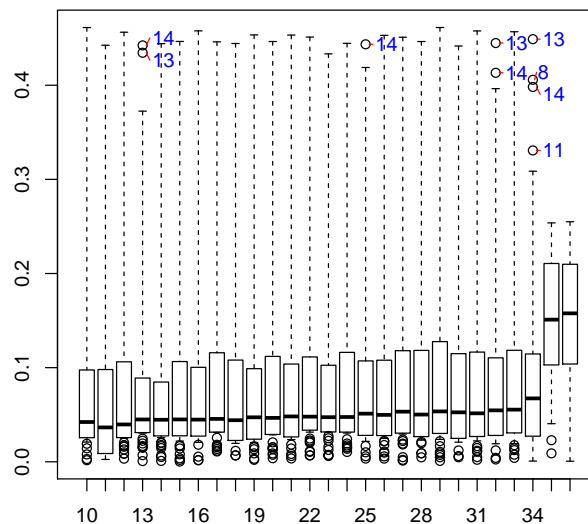


Figura 15 – Ponderação de casos: boxplot de *ILPPF*. Líquido de bochecho controle após 3 meses.

No boxplot assimétrico da Figura 15 observamos que no último passo da procura passo a frente não existe nenhum dado atípico e nas outras iterações, eventualmente aparecem as

observações 8, 11, 13 e 14. Na Figura 17, estas observações são as observações que causam as maiores mudanças no LD comparada com as outras observações no gráfico exceto a observação 12.

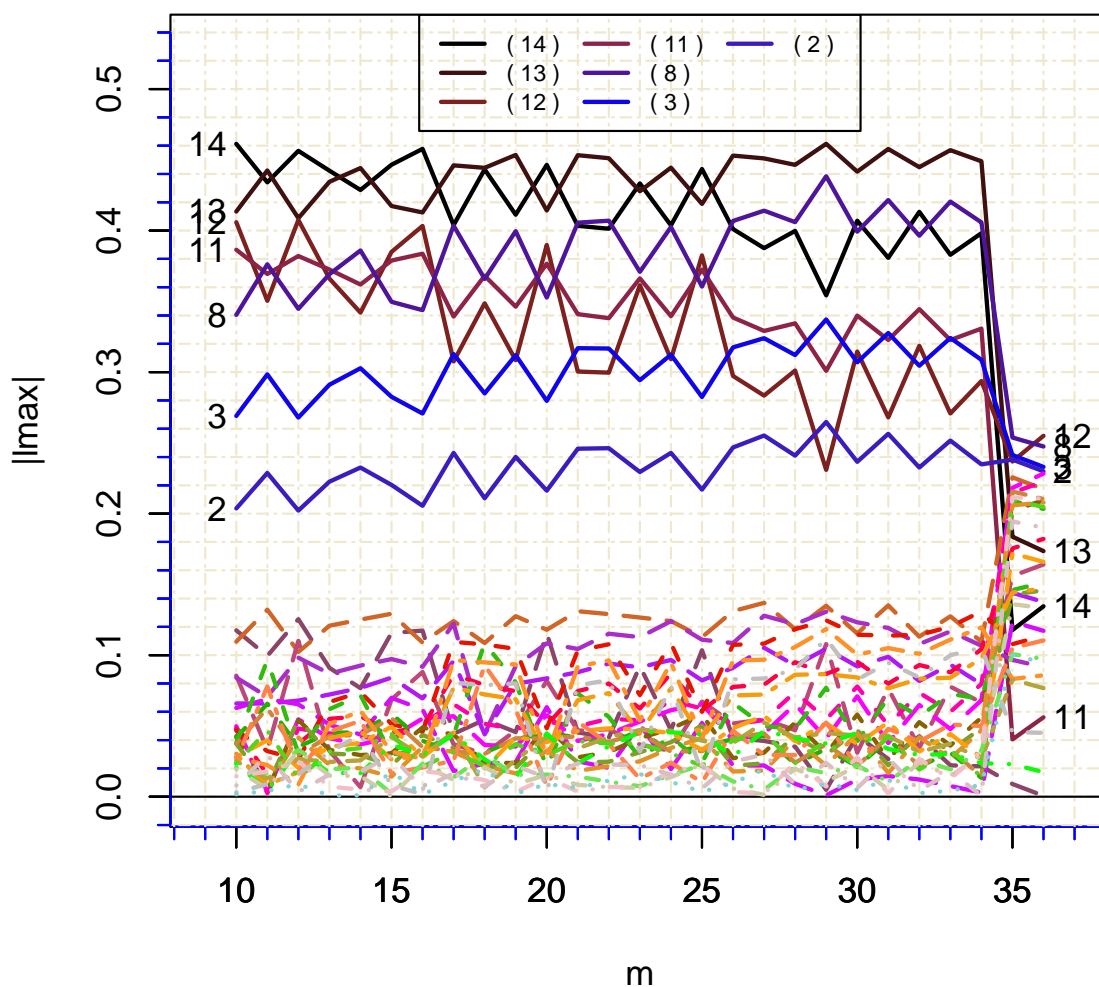


Figura 16 – Ponderação de casos: gráfico de ILPPF. Líquido de bochecho controle após 3 meses.

Na Figura 16 temos o gráfico de ILPPF. Claramente os dados são divididos em 2 grupos. Observamos que existe uma proporção de observações (14, 13, 12, 11, 8, 3 e 2) que se destacam das demais, em todo o processo da metodologia de ILPPF e no final do procedimento eles são mascarados. Sendo que as observações 14, 13 e 8 são as que mais se destacam. No entanto, como houve uma “grande” mudança no comportamento do gráfico ILPPF quando passam do passo  $m = 34$  para  $m = 35$  com a entrada das observações 12 e 14, estas observações devem ser influentes. Estas observações correspondem aos indivíduos com o maior índice de placa bacteriana no início do estudo (observação 14) e o segundo maior índice de placa bacteriana no

## Afastamento pela verossimilhança

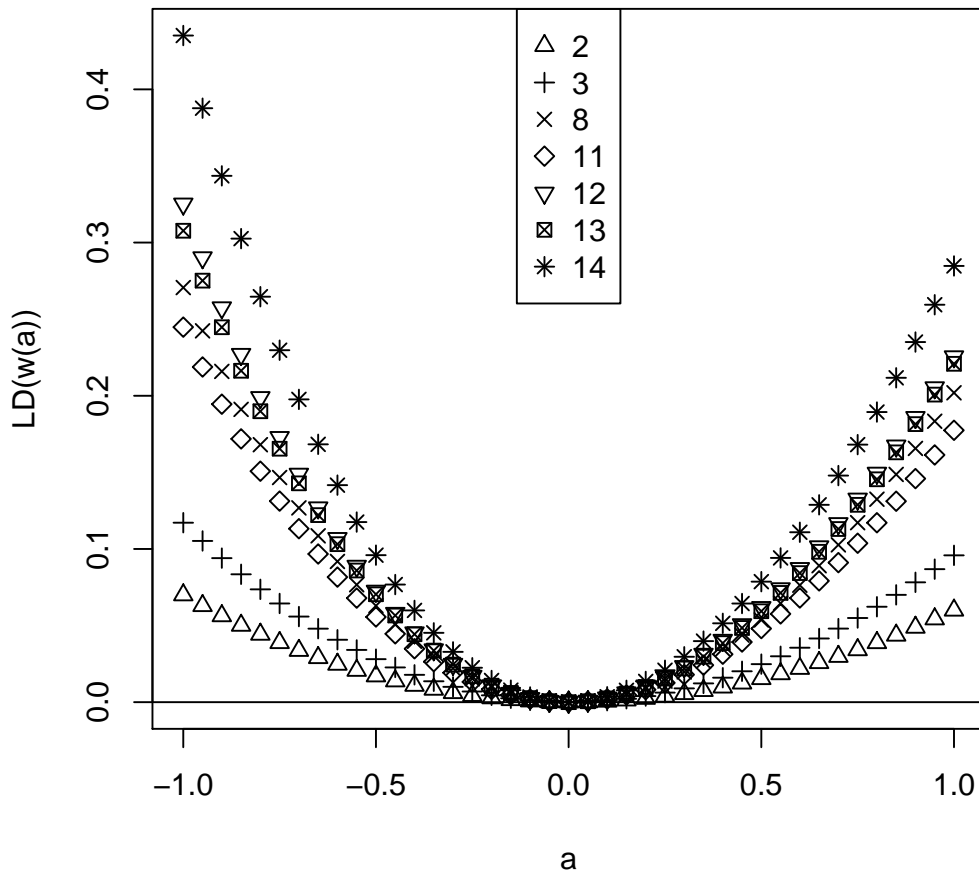


Figura 17 – Ponderação de casos: gráfico da função de afastamento pela verossimilhança. Líquido de bochecho controle após 3 meses.

	$\beta$	$\mu$	$\sigma_{\delta}^2$	$\sigma_x^2$	$\sigma_e^2$
Dados Completos	0.702	2.557	0.005	0.113	0.314
sem a(s) observação(s)					
12,14	0.7	2.514	0.004	0.088	0.332
porcentagem	0	2	15	22	6
14,13,8	0.676	2.498	0.005	0.08	0.303
porcentagem	4	2	0	29	4
14,13,12,11,8,3,2	0.648	2.417	0.004	0.035	0.306
porcentagem	8	5	13	69	3

Tabela 9 – Ponderação de casos: mudança relativa na estimativa dos parâmetros. Líquido de bochecho controle após 3 meses.

início do estudo (observação 12). No gráfico de afastamento pela verossimilhança da Figura 17 a observação 14 é a que mais se destaca e é influente dentre as outras. Observe também que as

observações 3 e 2, que são as observações que menos se destacam dentro o grupo de observações que se destacam (14, 13, 12, 11, 8, 3 e 2), são as que causam o menor impacto no gráfico LD da Figura 17.

Na Tabela 9 foram obtidas as mudanças relativas na estimativa dos parâmetros quando retiramos as observações (12, 14) que foram as observações que entram no passo  $m = 35$ , as observações (14, 13, 8) que são as observações que mais se destacam no gráfico de ILPPF e finalmente as observações (14, 13, 12, 11, 8, 3, 2) que se destacam das demais. Observamos que a maior mudança nas estimativas dos parâmetros foi no parâmetro  $\sigma_x^2$ , tem a maior porcentagem de mudança, comparado com os demais parâmetros.

#### 5.2.5.4 Líquido de bochecho controle após 6 meses do início do estudo - Ponderação de casos

Sejam:

$x$ : índice de placa bacteriana no início do estudo e

$y$ : índice de placa bacteriana após 6 meses do início do estudo com o uso do líquido de bochecho controle.

Foi feito o diagnóstico de influência local de Cook (1986) e obtivemos a curvatura máxima  $C_{max} = 3.23$ . Foi feito o gráfico do valor absoluto do autovetor correspondente ao maior autovalor e o índice das observações que se encontra na Figura 18.

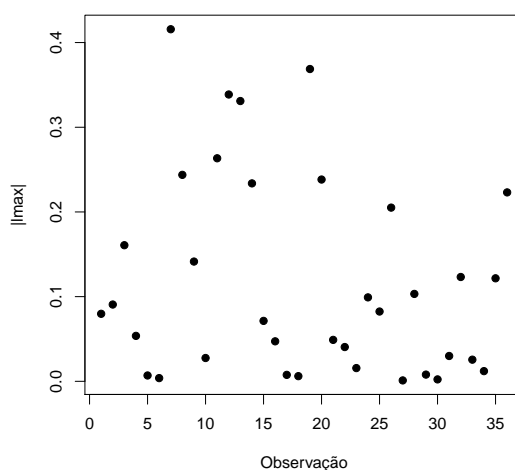


Figura 18 – Ponderação de casos: gráfico do valor absoluto do autovetor,  $l_{max}$ , associado ao maior autovalor com o índice das observações. Líquido de controle após 6 meses.

Podemos observar na Figura 18 que nenhuma observação se destaca das demais. Considerando a Figura 19, a observação 7 está afastada da maioria do conjunto de observações ou dados restantes, seguido das observações 19, 12 e 13. No caso da observação 7, o índice de placa bacteriana após 6 meses do uso do líquido de bochecho controle foi 0.35 que é o menor índice.

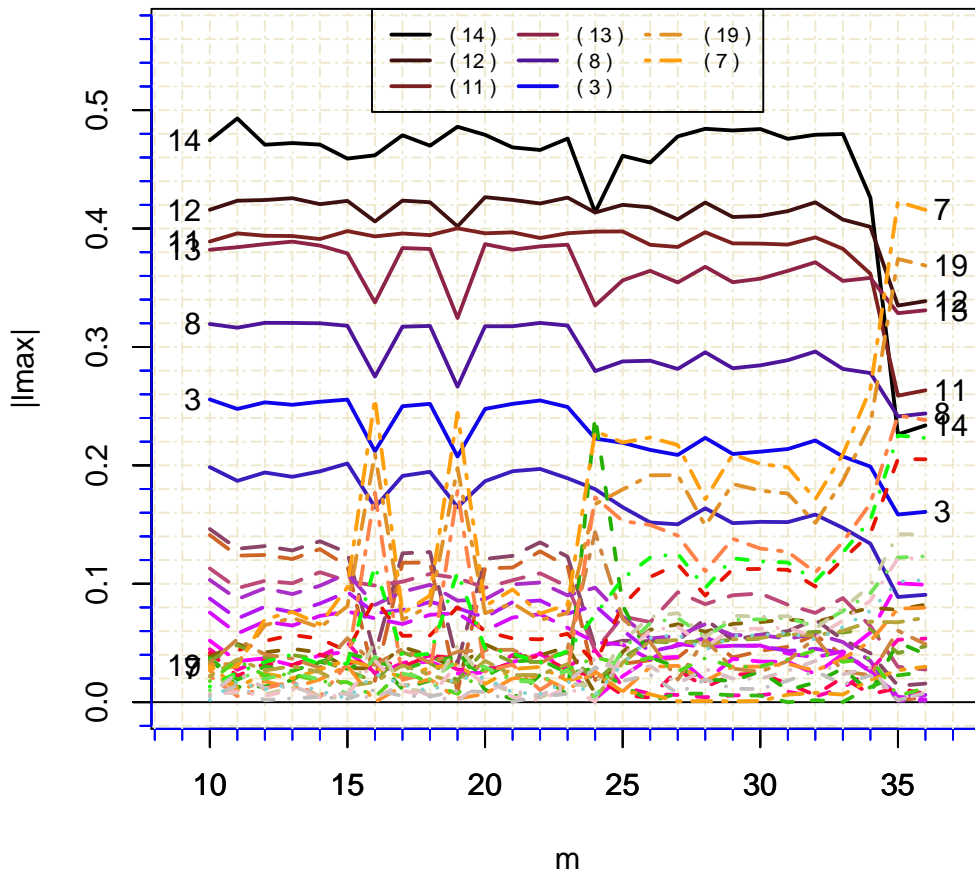


Figura 19 – Ponderação de casos: gráfico de *ILPPF*. Líquido de bochecho controle após 6 meses.

No boxplot ajustado da Figura 20 as observações 11, 12, 13 e 14 aparecem eventualmente como dados atípicos, no entanto na última iteração do gráfico da Figura 19, as observações 7 e 19 são as que mais se destacam e estas não aparecem no boxplot.

No gráfico da Figura 19 no passo  $m = 34$  entram as observações (13, 14) e no passo  $m = 35$  entram as observações (11, 12). No passo  $m = 34$  sai a observação 12 e no passo  $m = 35$  sai a observação 10. Observe na Figura 19 que apesar das observações 7, 19, 12 e 13 terem se destacado no final, a observação 14 é a que mais se destaca durante a evolução de *ILPPF* passando a não se destacar somente na parte final da evolução de *ILPPF*. E a observação 11 é a que se destaca em terceiro lugar em quase toda a evolução de *ILPPF* e somente no final, esta observação passa a não se destacar. Enquanto que as observações 7 e 19 não se destaca até a parte final da evolução de *ILPPF*. Na Tabela 10, a observação 14 é a que provoca a maior porcentagem de mudança relativa nas estimativas dos parâmetros dentre as observações individuais e no gráfico de afastamento pela verossimilhança da Figura 21 a observação 14 é a mais influente, desta forma, a observação 14 deve ser uma observação influente mascarada.

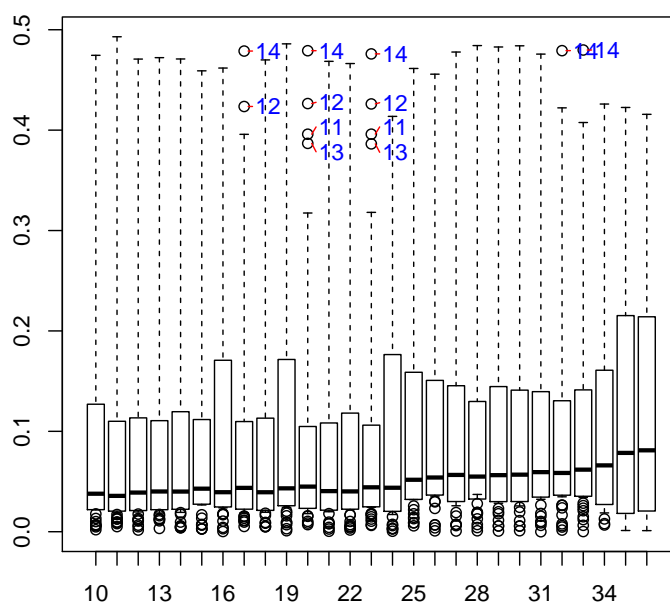


Figura 20 – Ponderação de casos: boxplot de ILPPF. Líquido de bochecho controle após 6 meses.

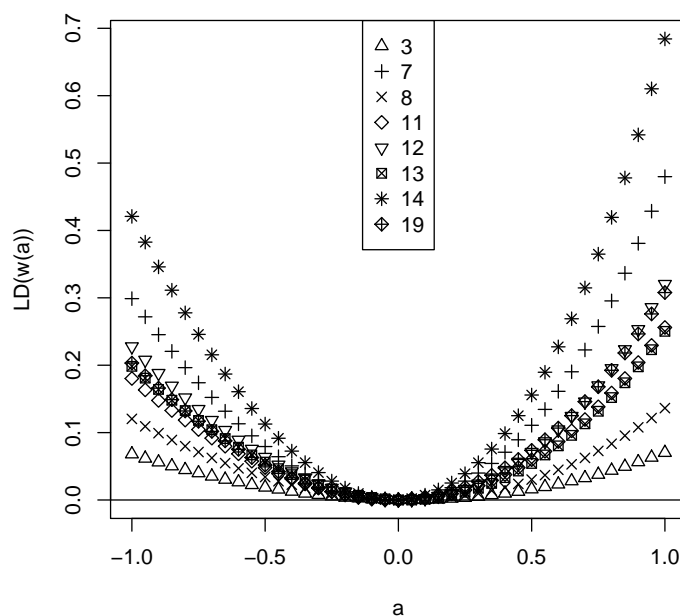


Figura 21 – Ponderação de casos: gráfico da função de afastamento pela verossimilhança. Líquido de bochecho controle após 6 meses.

Considerando o grupo de observações que foram excluídas, temos; (7, 19) são as observações que mais se destacam no final, (12, 14) são as observações que mais se destacam

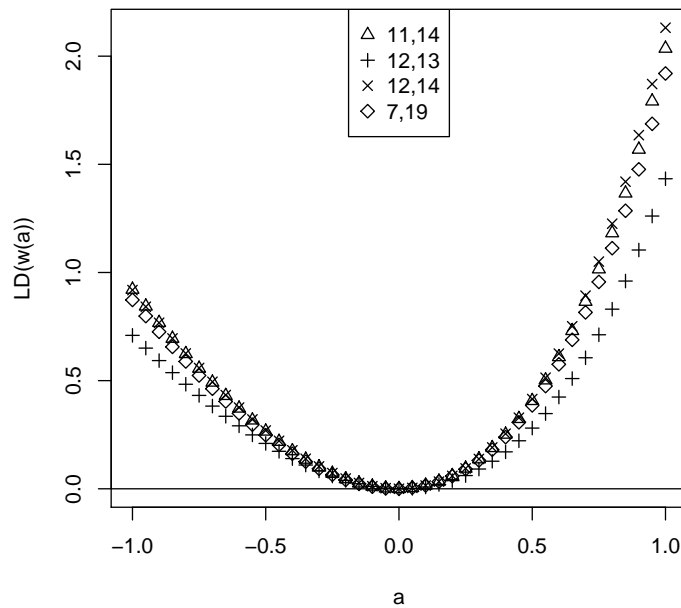


Figura 22 – Ponderação de casos: gráfico da função de afastamento pela verossimilhança. Líquido de bochecho controle após 6 meses.

Estimativas	$\beta$	$\mu$	$\sigma_{\delta}^2$	$\sigma_x^2$	$\sigma_e^2$
<b>Dados completos</b>	0.686	2.556	0.01	0.108	0.238
<b>Sem a obs.</b>					
(7)	0.696	2.57	0.011	0.103	0.209
Mudança relativa	0.014	0.005	0.054	-0.044	-0.121
(11)	0.682	2.538	0.011	0.098	0.242
Mudança relativa	-0.006	-0.007	0.054	-0.09	0.017
(12)	0.686	2.537	0.01	0.097	0.245
Mudança relativa	0	-0.008	-0.042	-0.099	0.03
(14)	0.674	2.535	0.012	0.093	0.222
Mudança relativa	-0.018	-0.008	0.15	-0.137	-0.067
(19)	0.695	2.57	0.011	0.103	0.217
Mudança relativa	0.013	0.005	0.054	-0.044	-0.088
(7, 19)	0.705	2.585	0.014	0.096	0.185
Mudança relativa	0.028	0.011	0.342	-0.109	-0.222
(11, 14)	0.67	2.516	0.012	0.082	0.225
Mudança relativa	-0.023	-0.016	0.15	-0.239	-0.054
(12, 14)	0.674	2.514	0.011	0.081	0.229
Mudança relativa	-0.018	-0.017	0.054	-0.248	-0.037
(7, 12, 13, 19)	0.71	2.545	0.011	0.078	0.196
Mudança relativa	0.035	-0.004	0.054	-0.276	-0.176
(11, 12, 13, 14)	0.671	2.471	0.01	0.056	0.24
Mudança relativa	-0.022	-0.033	-0.042	-0.48	0.009

Tabela 10 – Ponderação de casos: mudança relativa na estimativa dos parâmetros. Líquido de bochecho controle após 6 meses.

durante a evolução da *ILPPF* (ver Figura 22), (7,12,13,19) são as 4 observações que se destacam

no final e (11,12,13,14) as observações que se destacam durante a evolução de ILPPF. As maiores mudanças foram obtidas na estimativa de  $\sigma_x^2$  e neste caso o grupo de observações que destaca no final (7,19) ((7,12,13,19)) ocasiona menor mudança do que o grupo de observações que se destacam durante a evolução (12,14) ((11,12,13,14)).

No gráfico LD da Figura 21, notamos que a observação 3 é a observação que causa a menor influência dentre as observações consideradas no gráfico após a observação 8. Estas são as observações com os menores valores das curvas dentre as observações numeradas na Figura 19 de ILPPF. A observação com a maior influência (Figura 21) é a 14 que é a observação que mais se destaca no gráfico de ILPPF, no entanto, ela é mascarada no final. Na parte positiva de  $a$  (Figura 21) observa-se que a segunda observação mais influente é a 7 seguida de 19, que são as observações que se destacam no final da evolução de ILPPF (Figura 19) seguida das observações 12, 11 e 13 (Figura 21) que são as observações que mais se destacam durante a evolução de ILPPF após a observação 14.

#### 5.2.5.5 Líquido de bochecho controle após 6 meses do início do estudo - Perturbação na variável resposta

Sejam:

$x$ : Índice de placa bacteriana no início do estudo e

$y$ : Índice de placa bacteriana após 6 meses do início do estudo com o uso do líquido de bochecho controle.

No diagnóstico de influência local de Cook (1986) obtivemos a curvatura máxima  $C_{max} = 6.59$  e foi feito o gráfico do valor absoluto do autovetor correspondente ao maior

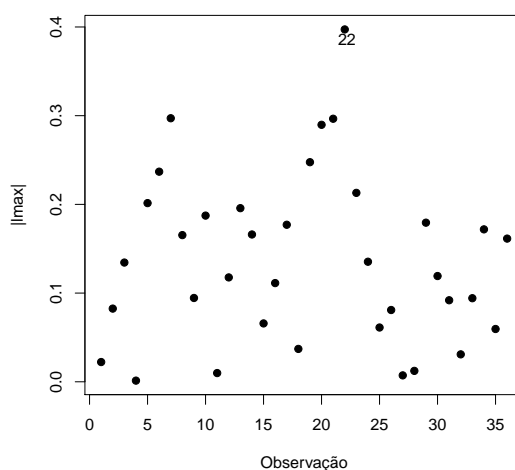


Figura 23 – Perturbação na resposta: gráfico do valor absoluto do autovetor, associado ao maior autovalor com o índice das observações. Líquido de bochecho controle após 6 meses.

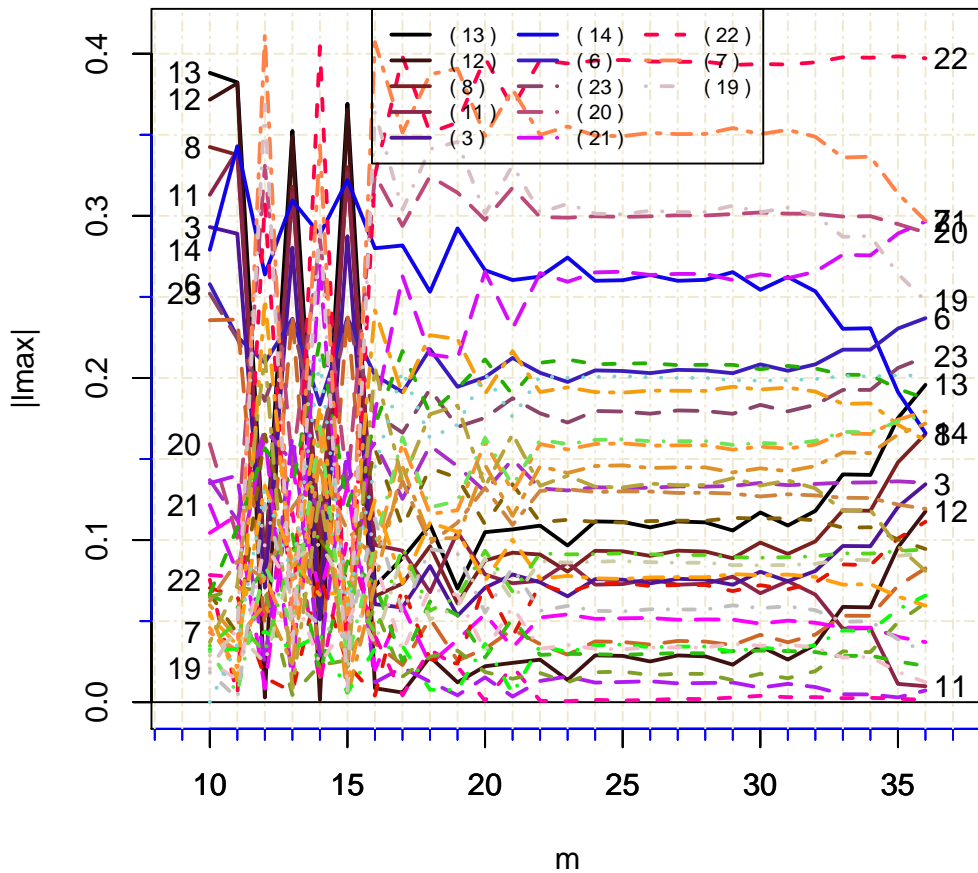


Figura 24 – Perturbação na resposta: gráfico de *ILPPF*. Líquido de bochecho controle após 6 meses.

autovalor e o índice das observações.

Podemos observar na Figura 23 que a observação 22 está um pouco afastado da maioria do conjunto de observações ou dados restantes, o índice de placa bacteriana deste indivíduo no início do estudo é 2.52 e após 6 meses do uso do líquido de bochecho controle o índice de placa bacteriana foi 2.85, ou seja, após o uso do líquido de bochecho o índice de placa bacteriana aumentou.

Na Figura 24 podemos notar que 6 observações se destacam das demais durante a evolução da *ILPPF* após a iteração  $m = 16$  até  $m = 34$  que são as observações 22, 7, 19, 20, 14 e 21. No gráfico LD da Figura 25 a observação 22 é a mais influente seguida de observação 7, dentre as analisadas e estas são as observações que se destacam primeiro e segundo lugar durante a evolução da *ILPPF* a partir de  $m = 22$ .

As observações 14, 19 e 20 tem impacto próximo no gráfico de afastamento pela verossimilhança da Figura 25. No gráfico da Figura 24 observamos que o dado 19 e 14 são possíveis

dados mascarados segundo a metodologia de ILPPF considerando a perturbação da variável resposta.

Estimativas	$\beta$	$\mu$	$\sigma_{\delta}^2$	$\sigma_x^2$	$\sigma_e^2$
<b>Dados completos</b>	0.686	2.556	0.01	0.108	0.238
<b>Sem a obs.</b>					
(7)	0.696	2.57	0.011	0.103	0.209
Mudança relativa	0.014	0.005	0.054	-0.044	-0.121
(14)	0.674	2.535	0.012	0.093	0.222
Mudança relativa	-0.018	-0.008	0.15	-0.137	-0.067
(19)	0.695	2.57	0.011	0.103	0.217
Mudança relativa	0.013	0.005	0.054	-0.044	-0.088
(20)	0.695	2.56	0.01	0.111	0.221
Mudança relativa	0.013	0.001	-0.042	0.031	-0.071
(21)	0.679	2.565	0.01	0.11	0.231
Mudança relativa	-0.01	0.003	-0.042	0.021	-0.029
(22)	0.674	2.557	0.01	0.112	0.208
Mudança relativa	-0.018	0	-0.042	0.04	-0.126
(21, 22)	0.666	2.566	0.009	0.114	0.198
Mudança relativa	-0.029	0.004	-0.137	0.058	-0.168
(7, 19)	0.705	2.585	0.014	0.096	0.185
Mudança relativa	0.028	0.011	0.342	-0.109	-0.222
(7, 20, 21)	0.699	2.583	0.011	0.108	0.183
Mudança relativa	0.019	0.01	0.054	0.003	-0.231
(7, 19, 20)	0.715	2.59	0.015	0.097	0.163
Mudança relativa	0.042	0.013	0.438	-0.099	-0.315
(7, 14, 20)	0.694	2.552	0.015	0.088	0.174
Mudança relativa	0.012	-0.002	0.438	-0.183	-0.269

Tabela 11 – Perturbação na resposta: mudança relativa na estimativa dos parâmetros ao tirar algumas observações. Líquido de bochecho controle após 6 meses.

Na Tabela 11 observamos que as maiores mudanças relativas considerando a exclusão de uma única observação ocorrem nas observações 14 e 22 e além disso, a observação 22 tem maior impacto no grafico de afastamento pela verossimilhança da Figura 25 e também na procura passo a frente da Figura 24. Na Tabela 11 temos os grupos (21,22), ((7,19)) onde as observações 21(7) e 22(19) possuem comportamentos parecidos, (7, 20, 21) que destacam em segundo lugar no final do algoritmo, (7,19,20) destacam no meio da evolução do algoritmo após  $m = 22$  até  $m = 34$ .

Vamos considerar agora a perturbação na variância definida em (5.19).

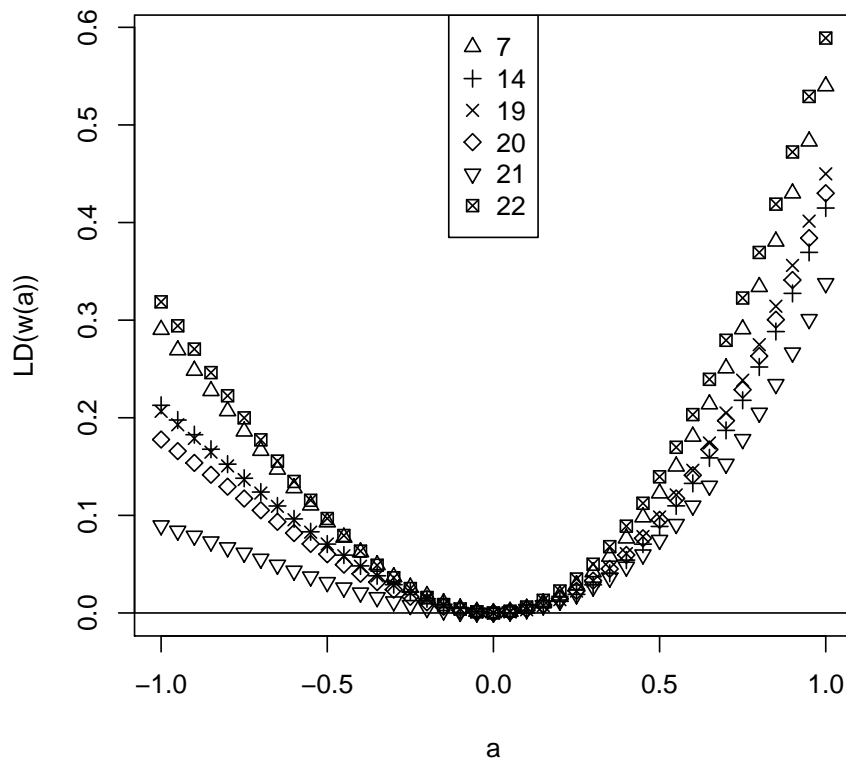


Figura 25 – Perturbação na resposta: gráfico de afastamento pela verossimilhança. Líquido de bochecho controle após 6 meses.

#### 5.2.5.6 Líquido de bochecho B após 6 meses do início do estudo - Perturbação na variância

Sejam:

$x$ : índice de placa no início do estudo e

$y$ : índice de placa após 6 meses do início do estudo com o uso do líquido de bochecho B.

Foi feito o diagnóstico de influência local de Cook (1986) considerando a perturbação na variância. Obtivemos a curvatura máxima  $C_{max} = 0.56$  e foi feito o gráfico do valor absoluto do autovetor correspondente ao maior autovalor e o índice das observações. Podemos observar na Figura 26 que a observação 12 está afastada da maioria do conjunto de observações ou dados restantes. No gráfico da Figura 27 observamos que o dado 12 é a que mais se destaca durante a evolução de *ILPPF* seguida da observação 11 e o 5.

As observações 5 e 11 chegam juntos no final da *ILPPF* no gráfico da Figura 27, então a observação 11 é mascarada pois produz maior influência comparada com a observação 5 no gráfico de afastamento pela verossimilhança da Figura 30 e por outro lado na influencia global ao retirar a observação 11 a mudança relativa no parâmetro  $\sigma_\delta^2$  e  $\sigma_x^2$  é a maior comparada com os outros.

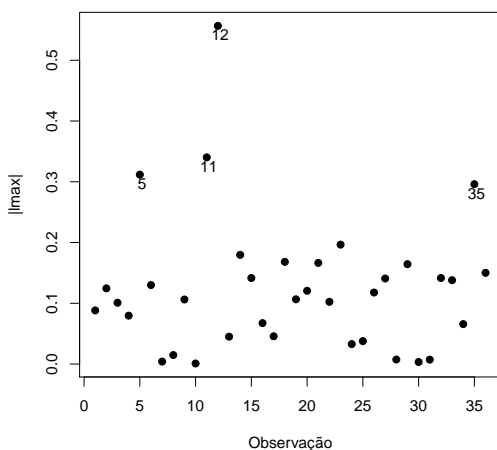


Figura 26 – Perturbação na variância: gráfico do valor absoluto do autovetor, associado ao maior autovalor com o índice das observações. Líquido de bochecho B após 6 meses.

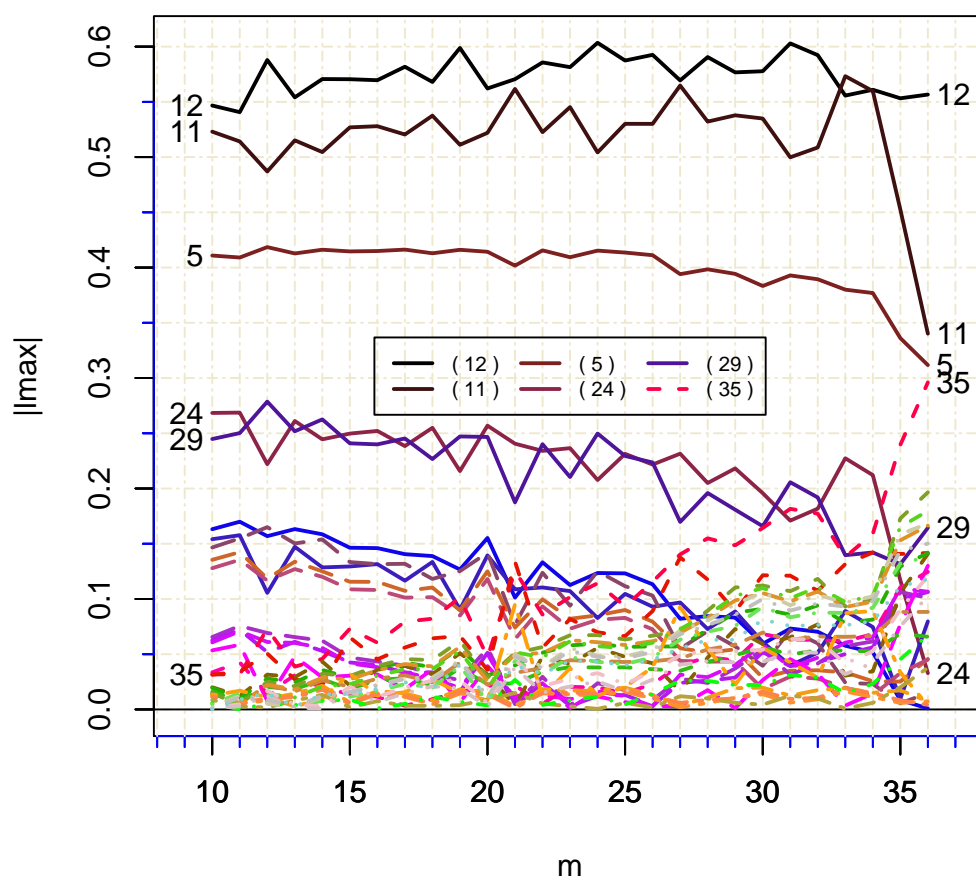


Figura 27 – Perturbação na variância: gráfico de ILPPF. Líquido de bochecho B após 6 meses.

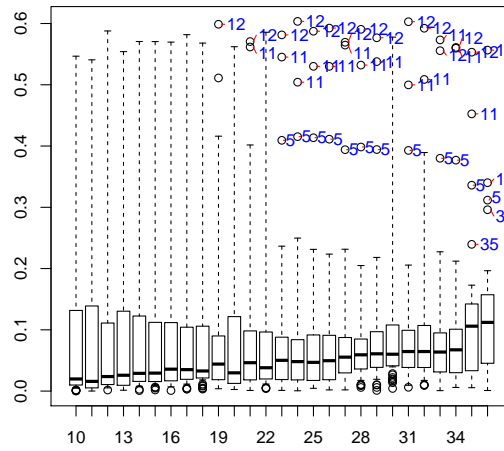


Figura 28 – Perturbação na variância: boxplot da *ILPPF*. Líquido de bochecho B após 6 meses.

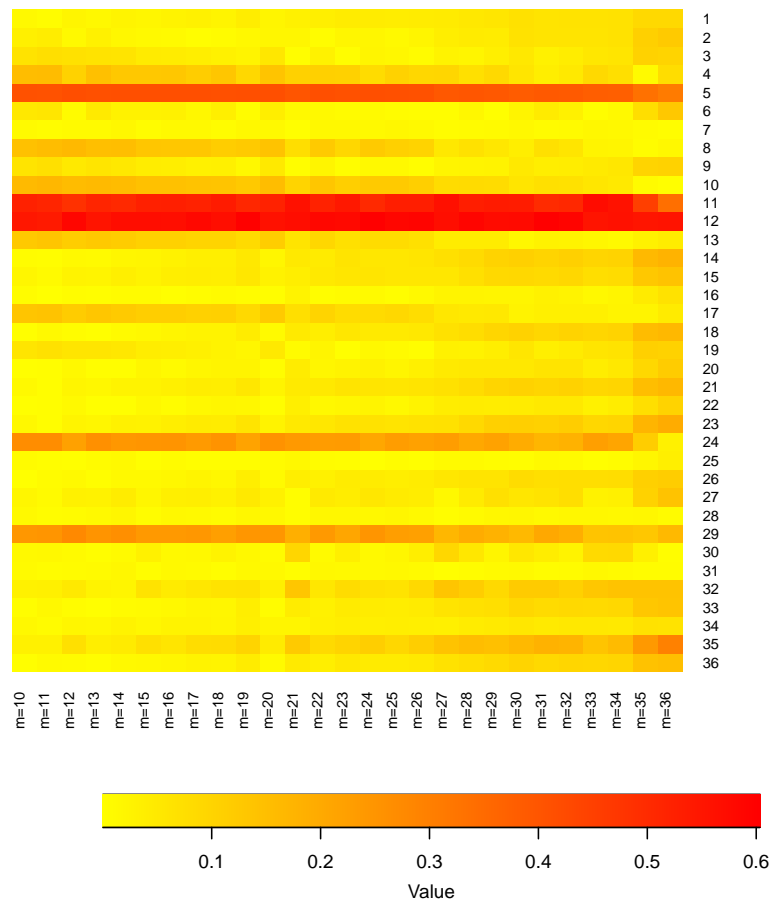


Figura 29 – Perturbação na variância: Heatmap de *ILPPF*. Líquido de bochecho B após 6 meses.

Estimativas	$\hat{\beta}$	$\hat{\mu}$	$\hat{\sigma}_{\delta}^2$	$\hat{\sigma}_x^2$	$\hat{\sigma}_e^2$
<b>Dados completos</b>	0.416	2.479	0.04	0.046	0.189
<b>Sem a obs.</b>					
(5)	0.419 0.006	2.464 -0.006	0.03 -0.247	0.049 0.068	0.195 0.031
(11)	0.425 0.02	2.461 -0.007	0.007 -0.824	0.069 0.504	0.185 -0.022
(12)	0.417 0.001	2.461 -0.007	0.035 -0.121	0.041 -0.106	0.195 0.031
(24)	0.424 0.018	2.468 -0.005	0.02 -0.498	0.063 0.374	0.185 -0.022
(29)	0.411 -0.013	2.468 -0.005	0.049 0.23	0.035 -0.237	0.187 -0.011
(35)	0.424 0.018	2.491 0.005	0.056 0.406	0.027 -0.411	0.18 -0.048
(24, 29)	0.419 0.006	2.456 -0.009	0.032 -0.196	0.049 0.068	0.183 -0.032
(11, 12)	0.425 0.02	2.441 -0.015	0.004 -0.9	0.061 0.33	0.191 0.01
(5, 11, 35)	0.435 0.044	2.456 -0.009	0.018 -0.548	0.048 0.047	0.18 -0.048
(11, 24)	0.434 0.042	2.449 -0.012	0.017 -0.573	0.056 0.221	0.177 -0.064
(12, 29)	0.412 -0.011	2.449 -0.012	0.048 0.205	0.024 -0.477	0.192 0.015

Tabela 12 – Perturbação na variância: mudança relativa na estimativa dos parâmetros. Líquido de bochecho B após 6 meses.

Observamos que a partir da metade da evolução da procura passo a frente os boxplots da Figura 28 detectam como dados atípicos os dados 12, 11 e 5, mas no heatmap da Figura 29 observa-se que o dado 11 no final da evolução fica laranja o que indica que pode ser um dado mascarado.

As observações 11 e 12 tem maior impacto dentre as observações perturbadas individualmente. No caso dos grupos, o grupo de observações que tem maior impacto é (5, 11, 12), seguido de (5, 11, 35) e (11, 12). como pode ser visto na Figura 30.

Considerando a influencia global, a maior mudança ocorre na estimativa das variâncias do verdadeiro valor da covariável e na variância do erro de medição da covariável.

Em relação a mudança na variância do erro de medição da covariável, o grupo de observações (11, 12) seguido de 11 e (11, 24) foram o que sofreram as maiores mudanças, enquanto que em relação a variância do verdadeiro valor da covariável, as observações que sofreram maiores mudanças foram a observação 11, seguido de (12, 29) e 35.

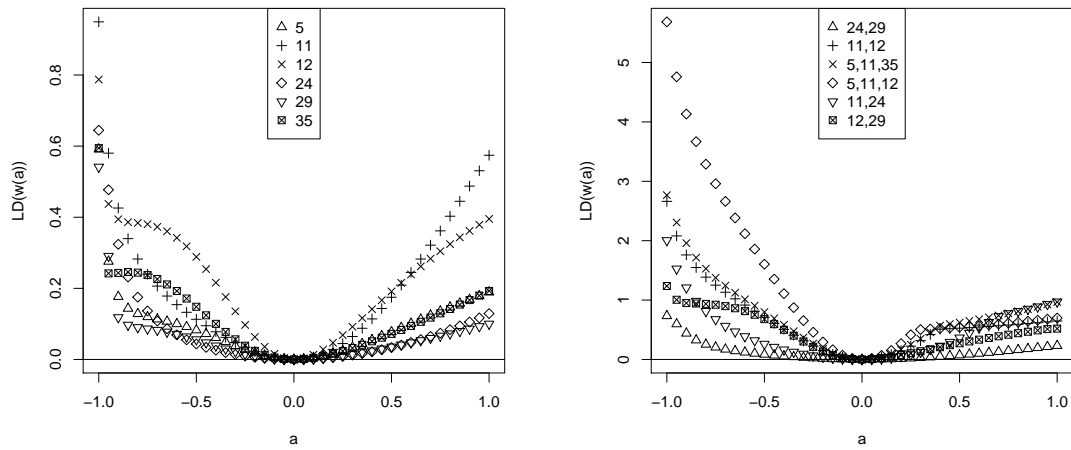


Figura 30 – Perturbação na variância: gráfico de afastamento pela verossimilhança. Líquido de bochecho B após 6 meses.

# ILPPF NO MODELO DE REGRESSÃO MULTIVARIADO COM ERROS NAS VARIÁVEIS COM INTERCEPTO NULO

## 6.1 Dados do líquido de bochecho

Aoki *et al.* (2003) e Russo, Aoki e Jr (2009) propuseram um modelo de regressão com erros de medição multivariado com intercepto nulo para modelar os dados de Hadgu e Koch (1999). A descrição dos dados e do modelo podem ser encontrados na Introdução. Considerando este conjunto de dados (1.2 e A.2) iremos aplicar a metodologia proposta de ILPPF.

Foram consideradas 4 tipos de perturbações para o modelo multivariado.

### 1. Ponderação de casos

Seja  $\mathbf{z}_{i_j}$  o vetor observado e  $\boldsymbol{\omega}$  o vetor de perturbação de dimensão  $N \times 1$ , com  $N = \sum_{i=1}^p n_i$ , tal que  $\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_p)^T = (\omega_{1_1}, \dots, \omega_{1_{n_1}}, \dots, \omega_{p_1}, \dots, \omega_{p_{n_p}})^T$ . A função do logaritmo de verossimilhança do modelo perturbado neste caso é escrita como:

$$\ell(\boldsymbol{\theta} | \boldsymbol{\omega}) = \sum_{i=1}^p \sum_{j=1}^{n_i} \omega_{ij} \log f(\mathbf{z}_{i_j}, \boldsymbol{\theta}),$$

e por tanto:

$$\ell(\boldsymbol{\theta} / \boldsymbol{\omega}) = -\frac{3}{2} \log(2\pi) \sum_{i=1}^p \sum_{j=1}^{n_i} \omega_{ij} - \frac{1}{2} \sum_{i=1}^p \log |\mathbf{V}_i| \sum_{j=1}^{n_i} \omega_{ij} - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} \omega_{ij} (\mathbf{z}_{i_j} - \mathbf{m}_i)^T \mathbf{V}_i^{-1} (\mathbf{z}_{i_j} - \mathbf{m}_i), \quad (6.1)$$

com o vetor de não-perturbação dado por  $\boldsymbol{\omega}_0 = \mathbf{1}_N = (1, \dots, 1)^T$ .

### 2. Perturbação nas covariáveis

As variáveis explanatórias  $x_{ij}$  foram perturbadas da seguinte maneira:

$$x_{ij\omega} = x_{ij} + S_x \omega_{ij},$$

$i = 1, \dots, p, j = 1, \dots, n_i$ , onde  $S_x$  é um fator de escala, por exemplo o desvio padrão amostral de  $x$ . A função do logaritmo de verossimilhança do modelo perturbado neste caso é escrita como

$$\ell(\boldsymbol{\theta}/\boldsymbol{\omega}) = -\frac{3N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^p n_i \log |\mathbf{V}_i| - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} (\mathbf{z}_{ij}^* - \mathbf{m}_i)^T \mathbf{V}_i^{-1} (\mathbf{z}_{ij}^* - \mathbf{m}_i), \quad (6.2)$$

com o vetor de não-perturbação dado por  $\boldsymbol{\omega}_0 = \mathbf{0}_N = (0, \dots, 0)^T$ . Onde  $\mathbf{z}_{ij}^* = (x_{ij} + S_x \omega_{ij}, y_{1ij}, y_{2ij})^T$ .

### 3. Perturbação nas variáveis respostas

As variáveis respostas,  $\mathbf{y}_{ij}$  foram perturbadas da seguinte maneira

$$\mathbf{y}_{ij\omega} = \mathbf{y}_{ij} + S_{\mathbf{y}_i} \omega_{ij},$$

onde  $S_{\mathbf{y}_i} = (S_{y_{1i}}, S_{y_{2i}})^T$  são fatores de escala e  $\mathbf{y}_{ij} = (y_{1ij}, y_{2ij})^T$ , em que  $i = 1, \dots, p, j = 1, \dots, n_i$ . A função do logaritmo de verossimilhança do modelo perturbado neste caso é dada por

$$\ell(\boldsymbol{\theta}/\boldsymbol{\omega}) = -\frac{3N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^p n_i \log |\mathbf{V}_i| - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} (\mathbf{z}_{ij}^{**} - \mathbf{m}_i)^T \mathbf{V}_i^{-1} (\mathbf{z}_{ij}^{**} - \mathbf{m}_i), \quad (6.3)$$

com o vetor de não-perturbação dado por  $\boldsymbol{\omega}_0 = \mathbf{0}_N = (0, \dots, 0)^T$ , onde  $\mathbf{z}_{ij}^{**} = (x_{ij}, y_{1ij} + S_{y_{1i}} \omega_{ij}, y_{2ij} + S_{y_{2i}} \omega_{ij})^T$ .

### 4. Perturbação nas variâncias dos erros

Seja o vetor de perturbação  $\boldsymbol{\omega} = (\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_p)^T = (\omega_{11}, \dots, \omega_{1n_1}, \dots, \omega_{p1}, \dots, \omega_{pn_p})^T$ , com o vetor de não-perturbação dado por  $\boldsymbol{\omega}_0 = \mathbf{1}_N = (1, \dots, 1)^T$ .

$$\sigma_{\omega_{ij}}^2 = \frac{\sigma^2}{\omega_{ij}} \text{ e } \boldsymbol{\sigma}_{\mathbf{e}_i \omega_{ij}}^2 = \frac{\boldsymbol{\sigma}_{\mathbf{e}_i}^2}{\omega_{ij}}.$$

A função do logaritmo de verossimilhança do modelo perturbado neste caso é dada por

$$\ell(\boldsymbol{\theta}/\boldsymbol{\omega}) = -\frac{3N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} \log |\mathbf{V}_{ij}^{**}| - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} (\mathbf{z}_{ij} - \mathbf{m}_i)^T (\mathbf{V}_{ij}^{**})^{-1} (\mathbf{z}_{ij} - \mathbf{m}_i), \quad (6.4)$$

com o vetor de não-perturbação dado por  $\boldsymbol{\omega}_0 = \mathbf{1}_N = (1, \dots, 1)^T$ , onde

$$\mathbf{V}_{ij}^{**} = \begin{bmatrix} \sigma_x^2 + \frac{\sigma^2}{\omega_{ij}} & \beta_{1i} \sigma_x^2 & \beta_{2i} \sigma_x^2 \\ \beta_{1i} \sigma_x^2 & \beta_{1i}^2 \sigma_x^2 + \frac{\sigma_{e_{1i}}^2}{\omega_{ij}} & \beta_{1i} \beta_{2i} \sigma_x^2 \\ \beta_{2i} \sigma_x^2 & \beta_{1i} \beta_{2i} \sigma_x^2 & \beta_{2i}^2 \sigma_x^2 + \frac{\sigma_{e_{2i}}^2}{\omega_{ij}} \end{bmatrix} \quad (6.5)$$

Estas perturbações serão utilizadas também no Capítulo seguinte para podermos obter a CNCPPF.

### 6.1.1 Aplicação

Considerando o conjunto de dados do líquido de bochecho, foi aplicado a metodologia descrita de ILPPF para a obtenção de dados influentes mascarados. Ao fazer a perturbação na variância do erro foi feito o diagnóstico de influência local de Cook (1986) e obtivemos a curvatura máxima  $C_{max} = 5.33$ . Foi feito o gráfico do valor absoluto do autovetor correspondente ao maior autovalor e o índice das observações se encontram na Figura 31. Na Figura 31 observa-se que o dado 42 está distante das demais observações seguido da observação 22.

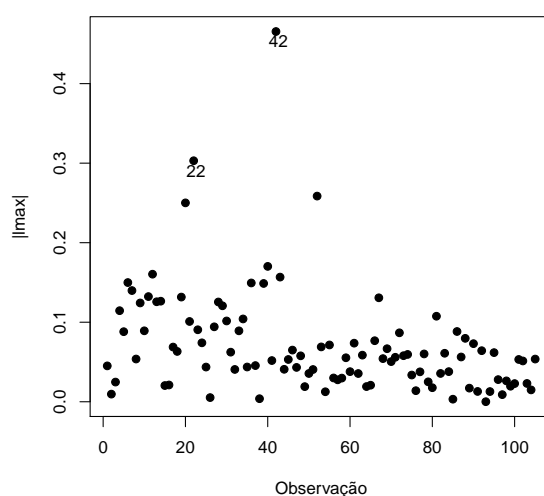


Figura 31 – Perturbação nas variâncias dos erros: gráfico do valor absoluto do autovetor correspondente ao maior autovalor, dados do líquido de bochecho.

Considerando a metodologia de ILPPF na perturbação da variância do erro, obtivemos o gráfico da Figura 32.

Na Figura 32, temos o gráfico da evolução da metodologia proposta, onde no eixo horizontal temos os tamanhos dos subconjuntos utilizados ( $m$ ) e no eixo vertical o gráfico de influência local de Cook (1986) obtido para cada subconjunto escolhido de tamanho  $m$ . Na última etapa do procedimento, é utilizado o conjunto de dados todo na estimação dos parâmetros, de tal forma que os pontos do gráfico de Cook da Figura 31 se encontra neste gráfico na última iteração da evolução. Podemos ver que existe um grupo de observações que se destacam no decorrer do processo, porém como o conjunto de dados é grande, é difícil detectar quais são estas observações (Figura 32).

Desta forma, foi construído o gráfico "Heat Map", onde cada observação fica em uma linha, sendo que quando as observações se destacam em alguma iteração, a cor fica mais escura. Quanto maior o destaque, mais escura é a cor.

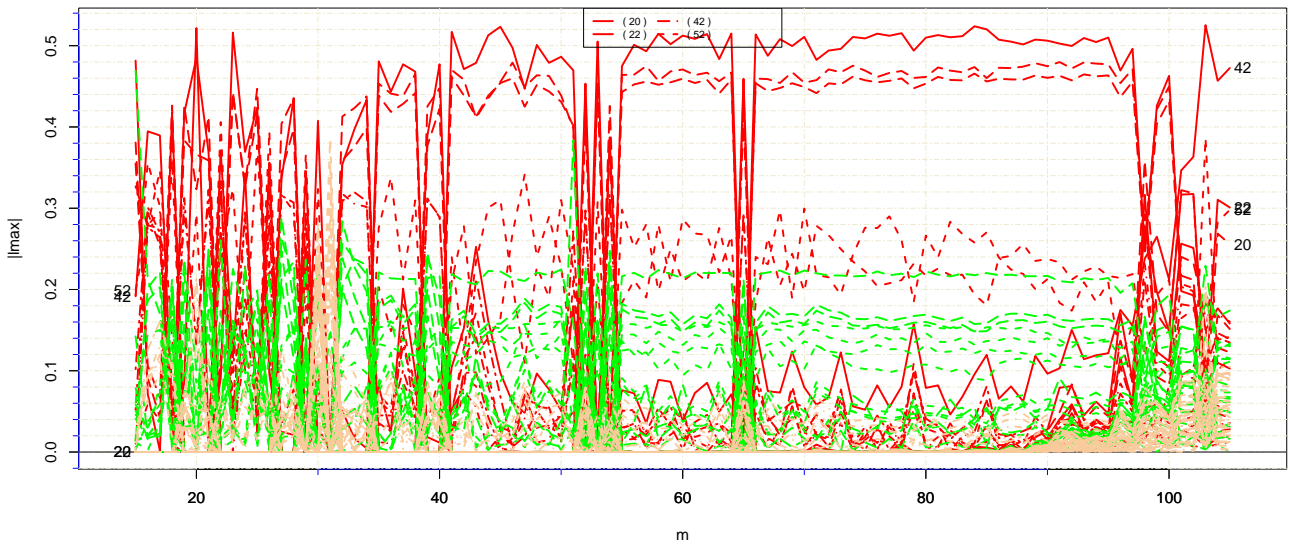


Figura 32 – Perturbação nas variâncias dos erros: gráfico de ILPPF, dados do líquido de bochecho.

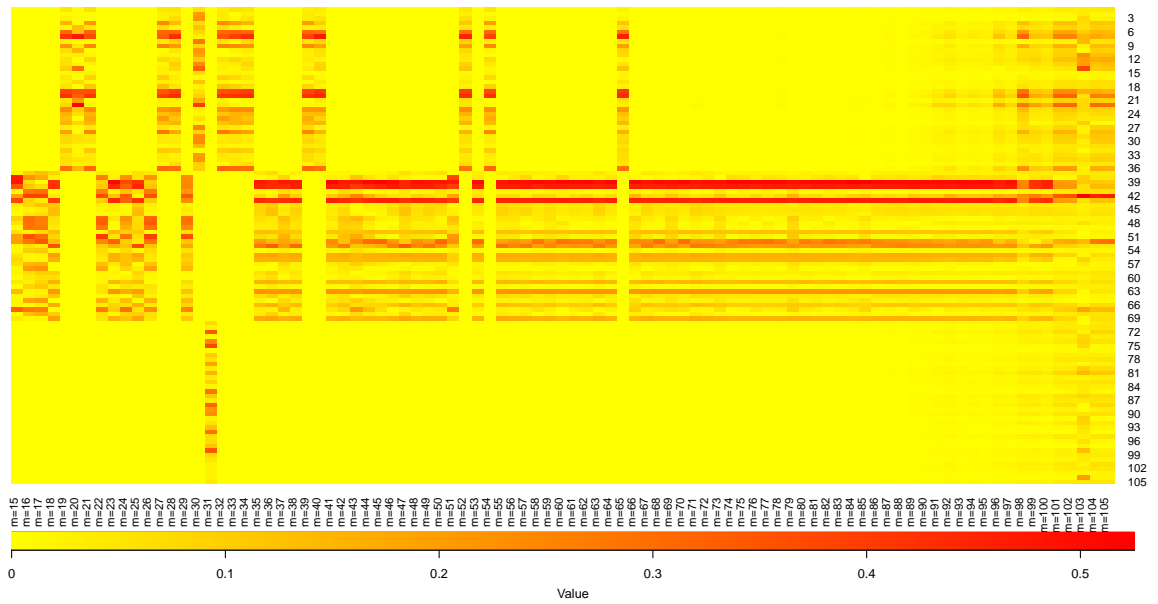


Figura 33 – Perturbação nas variâncias dos erros: Heat Map, dados de liquido de bochecho.

Com o auxílio do "Heat Map" podemos detectar que as observações 39, 40 e 43 se destacam em quase toda a evolução do procedimento, enquanto que a observação 42 se destaca a partir de  $m=98$ . Fazendo o gráfico da evolução da metodologia proposta (Figura 34) somente para as observações que detectamos (39, 40, 43 e 42) e comparando com o gráfico da Figura 32 (com todas as observações) fica claro que as observações 39, 40 e 43 se destacaram praticamente na evolução toda, enquanto que a observação 42, se destaca somente no final do procedimento.

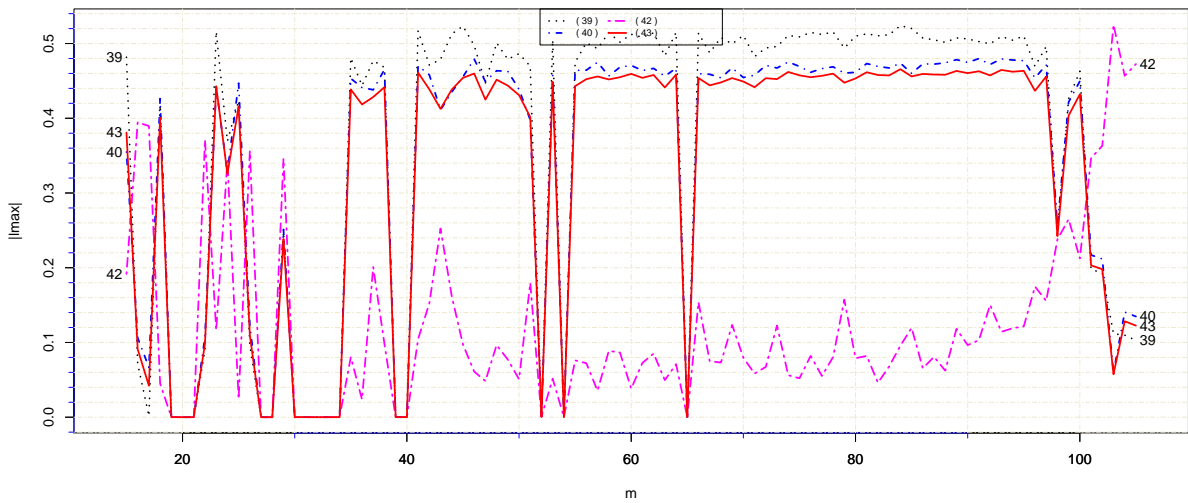


Figura 34 – Perturbação nas variâncias dos erros: gráfico de ILPPF, dados de líquido de bochecho.

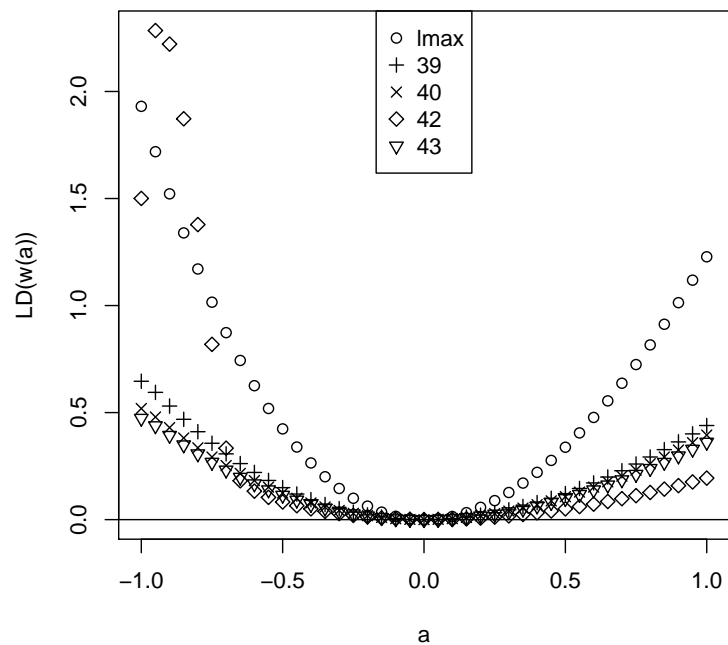


Figura 35 – Perturbação nas variâncias dos erros: gráfico de  $LD(\boldsymbol{w}(a))$  vs  $a$ , dados do líquido de bochecho.

Para podermos analisar a influência destas observações, vamos construir o gráfico de  $LD(\boldsymbol{w}(a))$  onde  $\boldsymbol{w}(a) = \boldsymbol{w}_0 + a\boldsymbol{l}$ , com  $a$  um escalar e  $\boldsymbol{l}$  um vetor fixo de norma 1. Foi feito o gráfico de  $LD(\boldsymbol{w}(a))$  vs  $a$  na direção  $\boldsymbol{l}_{max}$  e  $\boldsymbol{l}_{39}$ ,  $\boldsymbol{l}_{40}$ ,  $\boldsymbol{l}_{43}$  e  $\boldsymbol{l}_{42}$ , onde  $\boldsymbol{l}_{39}$  representa um vetor nulo

	$\hat{\beta}_{11}$	$\hat{\beta}_{21}$	$\hat{\beta}_{12}$	$\hat{\beta}_{22}$	$\hat{\beta}_{13}$	$\hat{\beta}_{23}$	$\hat{\mu}$	$\hat{\sigma}_{\delta}^2$
Dados.completos	0.703	0.6872	0.5253	0.5024	0.5082	0.4139	2.5337	0.0107
sem obs. 42	0.7033	0.6875	0.5092	0.4914	0.5084	0.414	2.5238	0.0105
	0	0	3	2	0	0	0	1
sem obs. 39	0.703	0.6872	0.543	0.5119	0.5083	0.4139	2.5321	0.0104
	0	0	3	2	0	0	0	3
sem obs. 40	0.703	0.6872	0.5392	0.5154	0.5082	0.4139	2.5342	0.0107
	0	0	3	3	0	0	0	0
sem obs. 43	0.703	0.6872	0.5398	0.5137	0.5082	0.4139	2.5342	0.0107
	0	0	3	2	0	0	0	0
sem obs. 39,43,40	0.703	0.6871	0.5743	0.5384	0.5083	0.4139	2.5331	0.0103
	0	0	9	7	0	0	0	3
	$\hat{\sigma}_x^2$	$\hat{\sigma}_{e_{11}}^2$	$\hat{\sigma}_{e_{21}}^2$	$\hat{\sigma}_{e_{12}}^2$	$\hat{\sigma}_{e_{22}}^2$	$\hat{\sigma}_{e_{13}}^2$	$\hat{\sigma}_{e_{23}}^2$	
Dados.completos	0.1027	0.3121	0.2335	0.4298	0.4316	0.255	0.1922	
sem obs. 42	0.0931	0.3133	0.2342	0.4124	0.4313	0.2552	0.1922	
	9	0	0	4	0	0	0	
sem obs. 39	0.1038	0.312	0.2335	0.3793	0.4263	0.255	0.1923	
	1	0	0	12	1	0	0	
sem obs. 40	0.1037	0.312	0.2335	0.3967	0.4042	0.2549	0.1922	
	1	0	0	8	6	0	0	
sem obs. 43	0.1037	0.312	0.2335	0.3928	0.4143	0.2549	0.1922	
	1	0	0	9	4	0	0	
sem obs. 39,43,40	0.1061	0.3117	0.2334	0.2903	0.3726	0.2549	0.1923	
	3	0	0	32	14	0	0	

Tabela 13 – Mudança relativa nas estimativas dos parâmetros dos dados de líquido de bochecho

de tamanho 105 com 1 na posição 39 (Figura 35). Para as demais observações, a notação é análoga.

Podemos ver que a curva da observação 39 se encontra mais perto da curva da direção  $l_{max}$ , seguida das observações 40, 43 e 42. Desta forma, as observações 39, 40 e 43 são observações mascaradas, pois no final do procedimento de ILPPF, as observações não se destacam apesar de estarem destacados durante quase toda a evolução da metodologia.

Na Tabela 13 considerando a influência global, observamos que ao retirar as observações (39, 40 e 43) em conjunto existe uma variação considerável na estimativa do parâmetro  $\sigma_{e_{12}}^2$ .

No Capítulo seguinte utilizamos as outras perturbações e este conjunto de dados para fazer o estudo da CNCPF.

---

# CURVATURA NORMAL CONFORMAL COM PROCURA PASSO A FRENTE NO MODELO DE REGRESSÃO MULTIVARIADO COM ERROS NAS VARIÁVEIS COM INTERCEPTO NULO

---

Neste capítulo, aplicamos a metodologia de CNCPPF aos dois conjuntos de dados reais descritos na Introdução (dados de escova de dentes e dados do líquido de bochecho) e a um conjunto de dados simulados para mostrar a utilidade da metodologia.

## 7.1 Dados de escovas de dentes

A seguir considerando o modelo definido em (1.4 e 1.5) iremos descrever as perturbações que foram utilizadas para a aplicação da CNCPPF nos dados da escova de dentes obtidos de um estudo conduzido na faculdade de odontologia da Universidade de São Paulo e analisado no CEA - Centro de Estatística Aplicada (IME-USP). Mais tarde, estes dados foram modelados em [Singer e Andrade \(1997\)](#) e [Aoki, Bolfarine e Singer \(2001\)](#).

### Esquemas de perturbação

Vamos considerar as 4 perturbações definidas em [Aoki, Singer e Bolfarine \(2007\)](#).

1. **Ponderação de casos** O logaritmo da função de verossimilhança perturbada neste caso é

dado por:

$$\begin{aligned}
\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\
&= \sum_{j=1}^n \omega_j \log f(\mathbf{Z}_j, \boldsymbol{\theta}) \\
&= -p \log(2\pi) \sum_{j=1}^n \omega_j - \frac{1}{2} \log |\mathbf{V}| \sum_{j=1}^n \omega_j - \frac{1}{2} \sum_{j=1}^n \omega_j (\mathbf{Z}_j - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_j - \mathbf{m}).
\end{aligned} \tag{7.1}$$

onde  $\boldsymbol{\omega} = (\omega_1, \dots, \omega_n)^T$  é o vetor de perturbações e o vetor de não perturbações é dado por  $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$  (definido em 3.1).

2. **Perturbação nas covariáveis** A perturbação na  $i$ -ésima covariável  $\mathbf{X}_i$  foi definida da seguinte forma:

$$X_{ij\omega} = X_{ij} + S_{X_i} \omega_j,$$

$j = 1, \dots, n$ , e  $i = 1, \dots, p$ , onde  $S_{X_i}$  é um fator de escala, por exemplo, o desvio padrão amostral de  $\mathbf{X}_i$ .

$$S_{X_i} = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2}.$$

A log-verossimilhança do modelo perturbado neste caso é dada por:

$$\begin{aligned}
\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\
&= \sum_{j=1}^n \log f(\mathbf{Z}_j^*, \boldsymbol{\theta}) \\
&= -np \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{j=1}^n (\mathbf{Z}_j^* - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_j^* - \mathbf{m}),
\end{aligned} \tag{7.2}$$

com  $\mathbf{Z}_j^* = (X_{1j} + S_{X_1} \omega_j, \dots, X_{pj} + S_{X_p} \omega_j, Y_{1j}, \dots, Y_{pj})^T$ ,  $j = 1, \dots, n$  e o vetor de não perturbação é dado por  $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$

3. **Perturbação nas variáveis respostas** A perturbação na variável resposta  $\mathbf{Y}_i$  foi definida como:

$$Y_{ij\omega} = Y_{ij} + S_{Y_i} \omega_j,$$

$j = 1, \dots, n$ , onde  $S_{Y_i}$  é um fator de escala, por exemplo, o desvio padrão amostral de  $Y_i$ .

$$S_{Y_i} = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (Y_{ij} - \bar{Y}_i)^2}.$$

A log-verossimilhança do modelo perturbado neste caso é dada por:

$$\begin{aligned}\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= \log L(\boldsymbol{\theta}|\boldsymbol{\omega}) \\ &= \sum_{j=1}^n \log f(\mathbf{Z}_j^*, \boldsymbol{\theta}) \\ &= -\frac{np}{2} \log(2\pi) - \frac{n}{2} \log |\mathbf{V}| - \frac{1}{2} \sum_{j=1}^n (\mathbf{Z}_j^* - \mathbf{m})^T \mathbf{V}^{-1} (\mathbf{Z}_j^* - \mathbf{m}),\end{aligned}\tag{7.3}$$

com  $\mathbf{Z}_j^* = (X_{1j}, \dots, X_{pj}, Y_{1j} + S_{Y_1} \boldsymbol{\omega}_j, \dots, Y_{pj} + S_{Y_p} \boldsymbol{\omega}_j)^T$ ,  $j = 1, \dots, n$  e o vetor de não perturbação é dado por  $\boldsymbol{\omega}_0 = (0, \dots, 0)^T$ .

**4. Perturbação na variância do erro** A perturbação utilizada na variância dos erros de medição  $\sigma_u^2$  foi definida como:

$$\sigma_{u\omega}^2 = \frac{\sigma_u^2}{\omega_j},$$

$j = 1, \dots, n$ . A log-verossimilhança perturbada é dada da seguinte forma:

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{np}{2} \log(2\pi) - \frac{n}{2} \log |\mathbf{V}_j^*| - \frac{1}{2} \sum_{j=1}^n (\mathbf{Z}_j - \mathbf{m})^T \mathbf{V}_j^{*-1} (\mathbf{Z}_j - \mathbf{m}),\tag{7.4}$$

com  $\mathbf{V}_j^* = \frac{\sigma_u^2}{\omega_j} \mathbf{A} + \sigma_x^2 \mathbf{b}\mathbf{b}^T$ . O determinante e a inversa de  $\mathbf{V}_j^*$  são dados por:

$$|\mathbf{V}_j^*| = \left( \frac{\sigma_u^2}{\omega_j} \right)^{2p} |\mathbf{D}(\boldsymbol{\lambda})| c_j^*, \quad \text{com} \quad c_j^* = 1 + \frac{\sigma_x^2}{\frac{\sigma_u^2}{\omega_j}} (p + \boldsymbol{\beta}^T \mathbf{D}(\boldsymbol{\lambda})^{-1} \boldsymbol{\beta}) \quad \text{e}$$

$$(\mathbf{V}_j^*)^{-1} = \left( \frac{\sigma_u^2}{\omega_j} \right)^{-1} \mathbf{A}^{-1} - (c_j^*)^{-1},$$

onde o vetor de não perturbações é dado por  $\boldsymbol{\omega}_0 = (1, \dots, 1)^T$ .

A matriz de informação de Fisher e os elementos da matriz  $\Delta$  necessárias para a aplicação de CNCPPF de todas as perturbações podem-se encontrados em [Aoki, Singer e Bolfarine \(2007\)](#).

[Zhu e Lee \(2001\)](#) propuseram alguns pontos de corte, considerando a curvatura normal conformal de [Poon e Poon \(1999\)](#), neste trabalho introduzimos  $\bar{m}(q) + 2sd$  com  $sd$  denotando o desvio padrão dos elementos do vetor  $\mathbf{m}(q)$  na contribuição agregada. Em seguida, aplicamos a metodologia proposta ao conjunto de dados da escova.

### 7.1.1 Aplicação

[Aoki, Bolfarine e Singer \(2001\)](#) propuseram o uso do modelo de regressão com erros nas variáveis com intercepto nulo para comparar a eficácia de dois tipos de escovas de dentes, a escova experimental e a escova convencional, descritas na Introdução. Considerando o modelo

definido em (1.4) e (1.10), foram aplicadas a metodologia de CNCPPF a cada um dos esquemas de perturbação definidos anteriormente. Aqui, consideramos dois extremos neste método. A contribuição total com  $q = 0$  para que todos os autovalores e os autovetores associados sejam incluídos e o outro extremo que permita  $q$  suficientemente grande para que apenas a contribuição do autovetor correspondente ao maior autovalor seja considerada. Os EMV dos parâmetros em cada iteração foram obtidos usando o algoritmo EM descrito em Aoki, Singer e Bolfarine (2007).

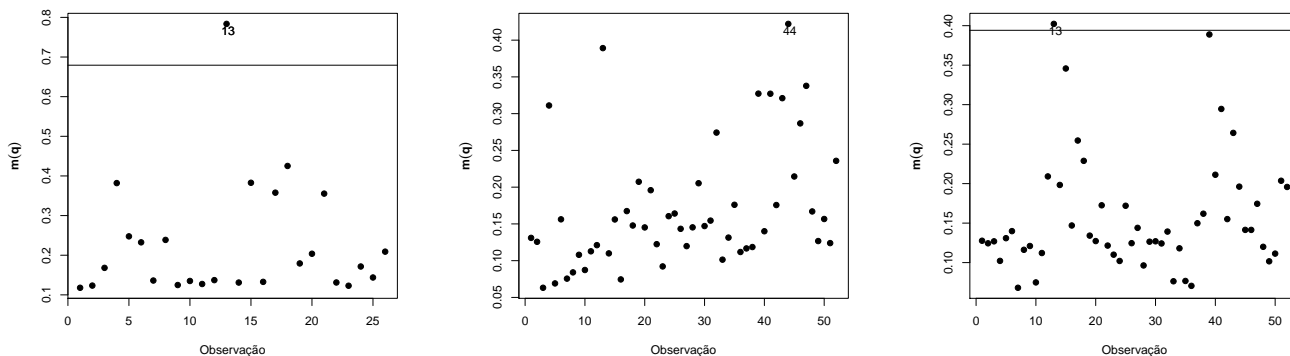


Figura 36 – Dados de escova de dente: gráfico de  $m(q)$  para Ponderação de casos, Perturbação nas variáveis respostas e Perturbação nas variáveis explanatórias. Contribuição total ( $q = 0$ ).

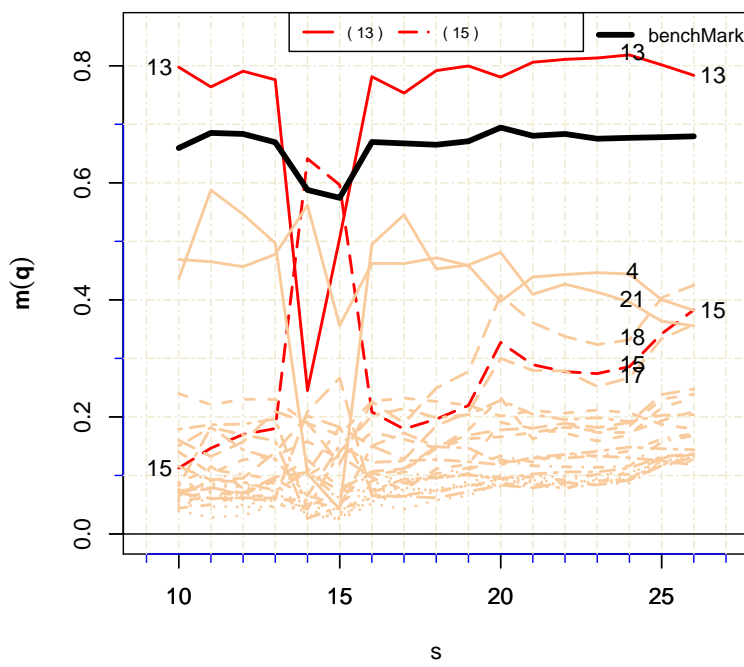


Figura 37 – Dados de escova de dente: gráfico da CNCPPF. Ponderação de casos. Contribuição total ( $q = 0$ ).

A Figura 36 mostra o gráfico do índice de  $m(q)$  quando  $q = 0$ , isto é, quando todos os

autovalores e os autovetores associados são incluídos na análise. O gráfico do índice considerando o esquema de ponderação de casos e o esquema de perturbação na variância é semelhante, chegando quase às mesmas conclusões, portanto apenas o primeiro é apresentado. No gráfico do meio (esquema de perturbação nas variáveis respostas) e no gráfico do lado direito (esquema de perturbação nas variáveis explicativas) os índices 1 a 26 no eixo  $x$ , referem-se aos dados associados à escova de dente experimental, enquanto que os índices 27 a 52 referem-se aos dados correspondentes à escova de dentes convencional para os mesmos indivíduos na mesma ordem. Considerando o gráfico a esquerda (esquema ponderação de casos) e o gráfico a direita (esquema de perturbação nas variáveis explicativas) a observação 13 está acima do ponto de corte. Essa observação corresponde ao pré-escolar com o menor índice de placa dentária antes da escovação com a escova convencional e também a única criança que não teve redução no tratamento convencional, enquanto que no tratamento experimental é a criança com o maior índice de placa dentária pós-escovação. O gráfico do meio não mostra observações influentes.

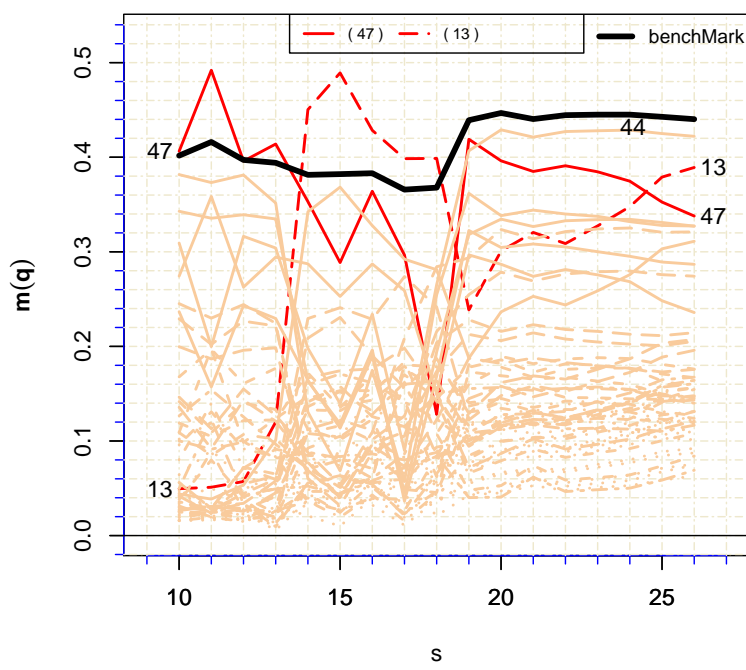


Figura 38 – Dados de escova de dente: gráfico da CNCPPF. Perturbação nas variáveis respostas. Contribuição total ( $q = 0$ ).

Nas Figuras 37, 38 e 39 mostramos os gráficos da CNCPPF para os três casos mostrados na Figura 36. O algoritmo começa com um subconjunto de tamanho  $m = 10$  e em cada etapa do algoritmo o tamanho da amostra é incrementado em 1, de tal forma que, na última etapa, a análise e a estimativa dos parâmetros são feitas com toda a amostra de tamanho  $m = 26$ , o que significa que os pontos na última iteração são iguais ao gráfico do índice da contribuição agregada da Figura 36. A observação 13 na Figura 37, está acima do ponto de corte em quase toda a evolução

da CNCPPF, ou seja, claramente a observação 13 é influente. Enquanto que a observação 15 ultrapassa o ponto de corte nas iterações  $m = 14$  e  $m = 15$ . Além disso, existem duas observações com linha sólida abaixo do ponto de corte que se destacam das demais observações durante quase toda a evolução da CNCPPF (observações 4 e 21). Existem também três observações com linha pontilhada abaixo do ponto de corte que se destacam do restante das observações, da iteração 20 até o final da evolução, e essas três observações apresentam comportamento semelhante durante essas iterações (observações 15, 17, 18).

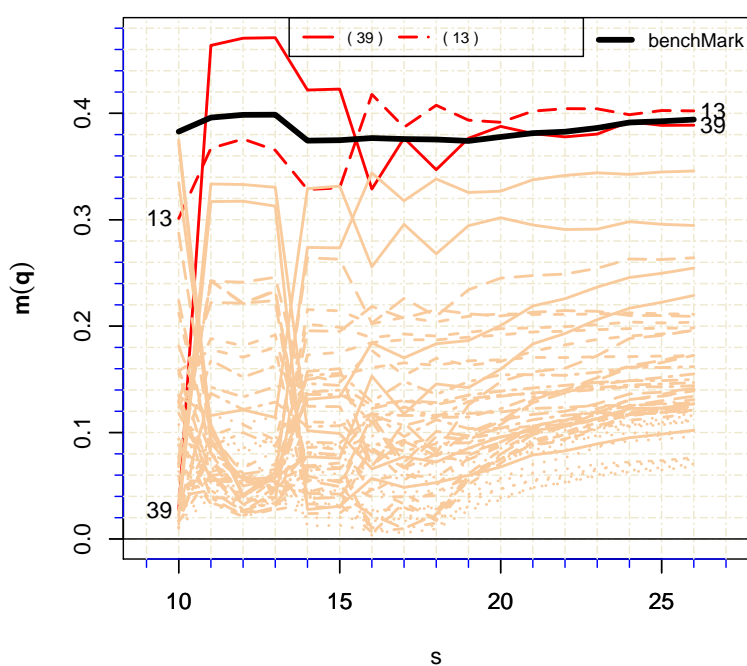


Figura 39 – Dados de escova de dente: gráfico da CNCPPF. Perturbação nas variáveis explanatórias. Contribuição total ( $q = 0$ ).

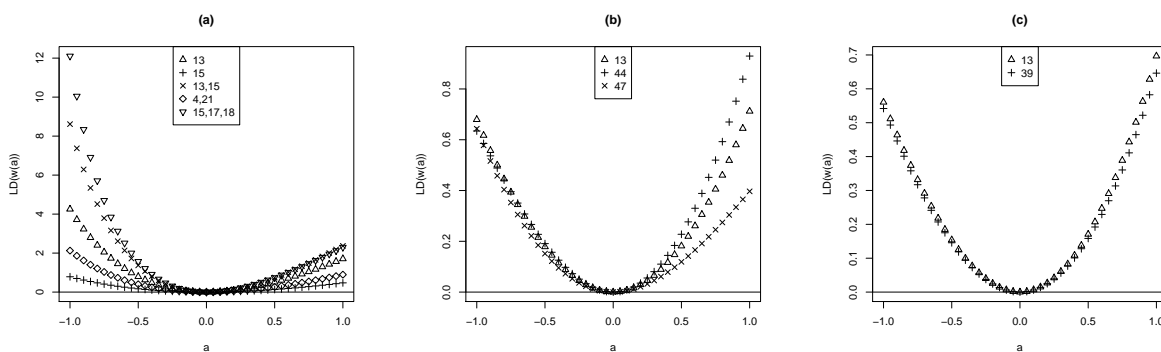


Figura 40 – Dados de escova de dente: gráfico de  $LD(\omega(a))$  versus  $a$  com  $\omega(a) = \omega_0 + a\mathbf{l}$  para (a) ponderação de casos, (b) perturbação nas variáveis respostas e (c) perturbação nas variáveis explanatórias.

A Figura 40a mostra o gráfico de  $LD(\boldsymbol{\omega}_0 + a\mathbf{l})$  versus  $a \in [-1, 1]$  para  $\mathbf{l} = \mathbf{l}_k$ , com  $k = 13, 15, (13, 15), (4, 21)$  e  $(15, 17, 18)$  sob o esquema de ponderação de casos, com  $\mathbf{l}_k$  denotando um vetor nulo de comprimento 26 com o  $k$ -ésimo elemento substituído por 1. As observações  $(13, 15)$  que são as duas observações que excedem o ponto de corte e as observações  $(15, 17, 18)$  que possuem comportamento similar são igualmente influente no lado positivo de  $a$ , enquanto que no lado negativo as observações  $(15, 17, 18)$  provoca uma mudança ligeiramente maior no afastamento pela verossimilhança do que as observações  $(13, 15)$ , seguido das observações 13,  $(4, 21)$  e 15. Considerando a influência global, podemos ver na Tabela 14 que a exclusão da observação 13 impacta na estimativa das variâncias do erro de medição. Foram excluídas as observações  $(13, 15)$  que ultrapassam o benchmark durante a evolução da CNCPPF e também as observações  $(4, 21)$  e  $(15, 17, 18)$ .

Na Figura 38, a observação 21 com o uso da escova de dente convencional (47 no gráfico) está acima do ponto de corte nos passos iniciais. Este pré-escolar teve a maior redução no índice de placa dentária após o uso da escova convencional e além disso possui o segundo maior índice de placa dentária antes do uso da escova de dente convencional. No meio da evolução da CNCPPF aparece a observação 13. Após a iteração 19 não aparecem observações acima do ponto de corte, porém temos a observação 18 (44 no gráfico) que está perto do ponto de corte até o fim da evolução e é separado do resto das observações. A observação 18 é de um pré-escolar que teve o maior valor de índice de placa dentária após o uso da escova de dente convencional. A Figura 40b mostra que para os valores positivos de  $a$  a observação 18 (44 no gráfico) é a mais influente, seguido pelas observações 13 e 21 (47 no gráfico). Do lado negativo de  $a$  as influências destas observações estão bastante próximos.

A Figura 39 mostra que a observação 13 com o uso da escova de dente experimental está acima do ponto de corte a partir de  $m = 16$  até o final da evolução da CNCPPF, enquanto que as mesmas observações com o uso da escova de dente convencional (observações 39 no gráfico) está acima do ponto de corte de  $m = 11$  até  $m = 15$ , a partir de  $m = 19$  a curva fica em torno do ponto de corte. A observação 39 corresponde a criança que possui o menor índice de placa dentária dentre as crianças que usaram a escova de dente convencional e além disso essa criança não teve redução do índice de placa no tratamento convencional. No último gráfico da Figura 40 podemos ver que estas duas observações possuem praticamente a mesma influência nas estimativas dos parâmetros, sendo que a observação 13 provoca uma mudança ligeiramente maior no LD,

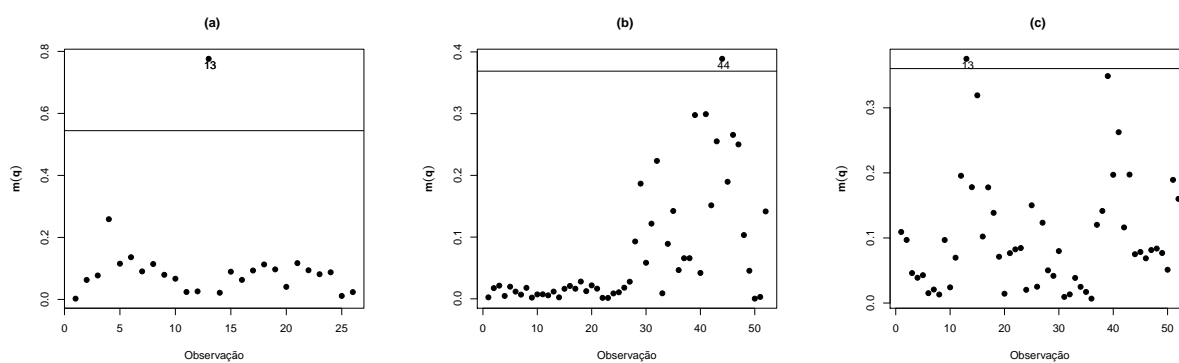


Figura 41 – Dados de escova de dente: gráfico  $m(q)$  para (a) ponderação de casos, (b) perturbação nas variáveis respostas e (c) perturbação nas variáveis explanatórias. Contribuição do autovetor associado ao maior autovalor.

A Figura 41 mostra o gráfico de  $m(q)$  com  $q$  suficientemente grande tal que só a contribuição do autovetor associado com o maior autovalor é considerado. Como o gráfico

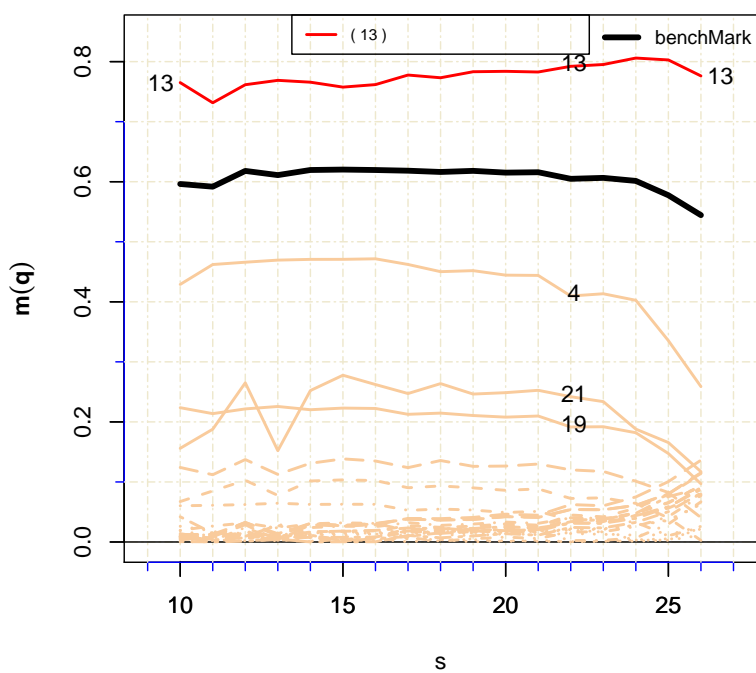


Figura 42 – Dados de escova de dente: gráfico da CNCPPF para ponderação de casos. Contribuição do autovetor associado ao maior autovalor.

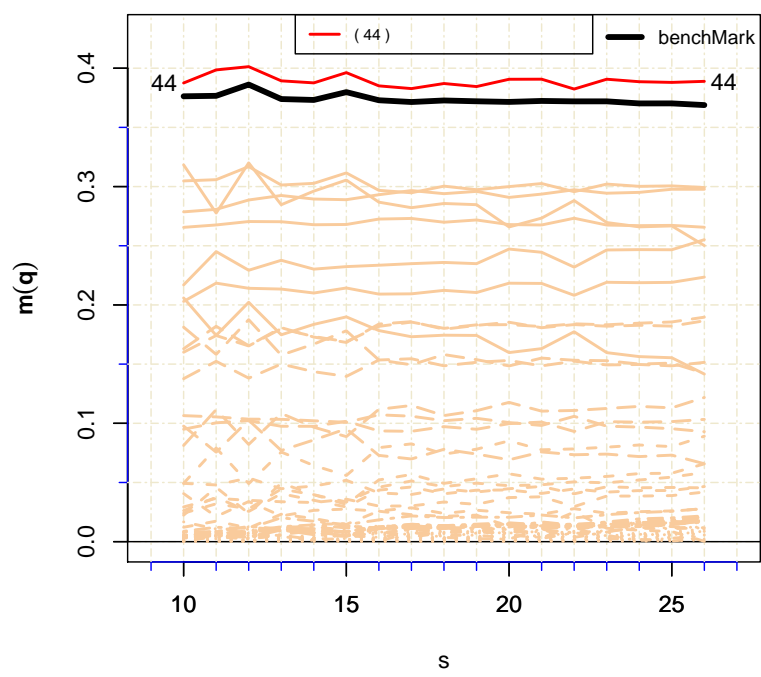


Figura 43 – Dados de escova de dente: gráfico da CNCPPF para perturbação na variável resposta. Contribuição do autovetor associado ao maior autovalor.

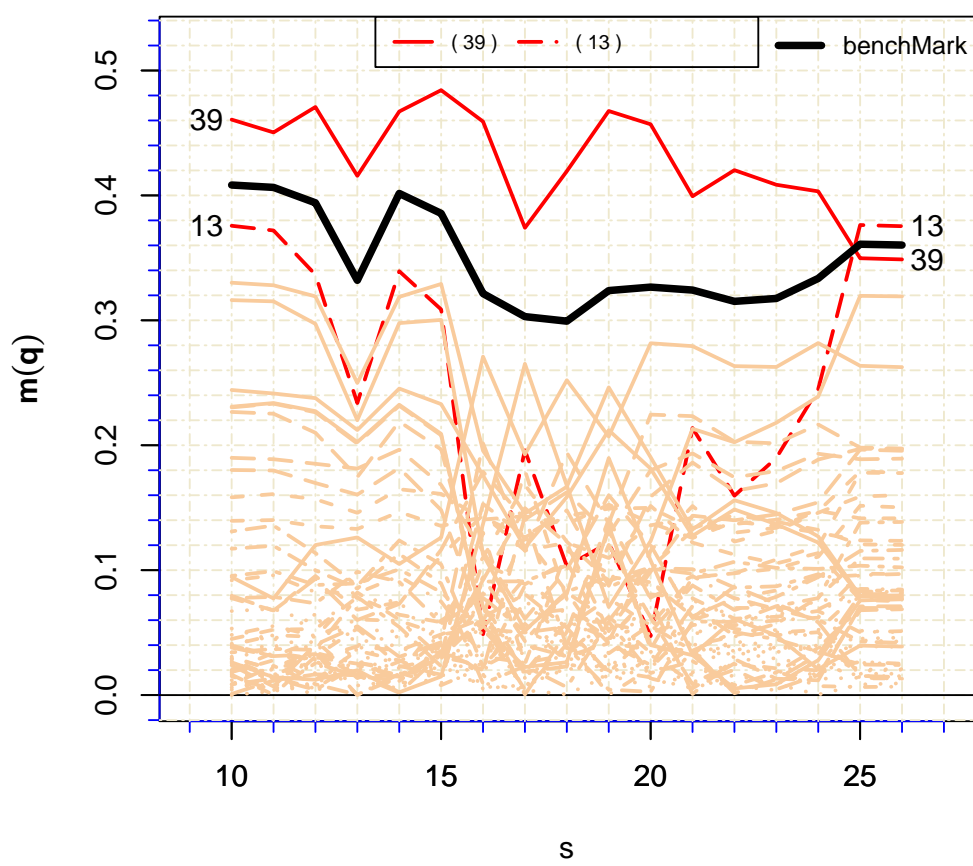


Figura 44 – Dados de escova de dente: gráfico da CNCPPF para perturbação na variável explanatória. Contribuição do autovetor associado ao maior autovalor.

da ponderação de casos e da perturbação na variância são similares com as mesmas conclusões, só o primeiro vai ser apresentado.

Considerando a Figura 41a (ponderação de casos) e a Figura 41c (perturbação na variável explanatória), a observação 13 está acima do ponto de corte, enquanto que na Figura 41b (perturbação na variável resposta) a observação 18 (44 no gráfico) esta acima do ponto de corte. Além disso, claramente uma pequena mudança na variável resposta associada com a escova de dente convencional tem um efeito maior na estimativa dos parâmetros.

As Figuras 42, 43 e 44 mostram o gráfico da CNCPPF considerando a ponderação de casos, perturbação nas variáveis respostas e perturbação nas variáveis explanatórias. O gráfico da CNCPPF para a perturbação na variância é omitida por ser similar a ponderação de casos chegando a mesmas conclusões. Considerando os gráficos 42 e 43 é claro que a observação 13 e a observação 18 (44 no gráfico) respectivamente, devem ser influentes. Enquanto que a Figura 44 mostra que a observação 13 sobre o tratamento convencional (39 no gráfico) está acima do ponto de corte durante quase toda a evolução da CNCPPF mas é mascarado nos passos finais. Por outro

lado, a observação 13 sobre o tratamento experimental aparece nos passos finais. Este gráfico refere-se ao esquema de perturbação na variável explanatória. No tratamento convencional o índice de placa bacteriana no início do estudo para a observação 13 é o menor valor, enquanto que no tratamento experimental este valor é o segundo maior.

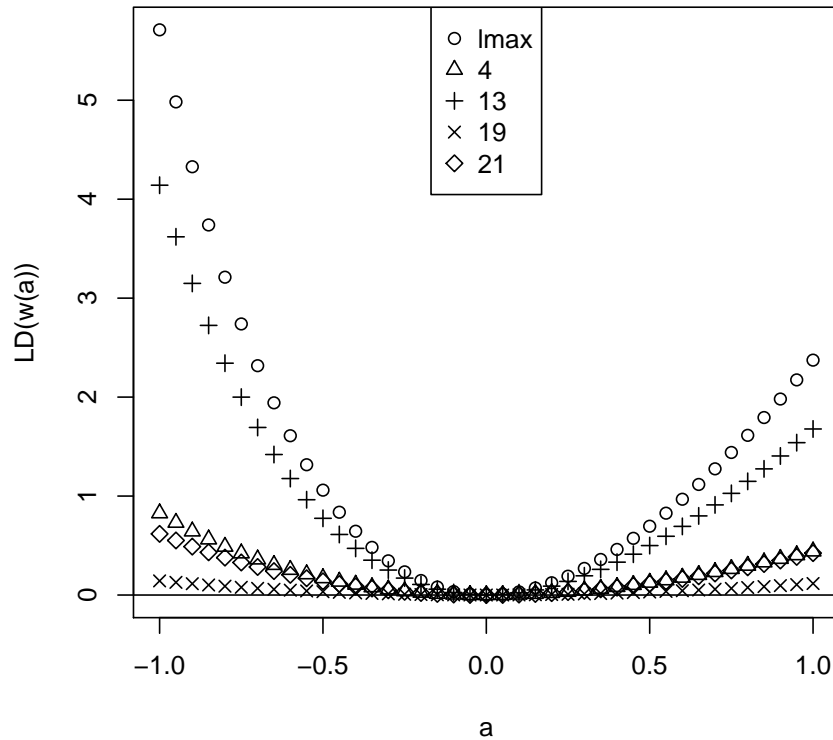


Figura 45 – Dados de escova de dente: ponderação de casos - gráfico de  $LD(\omega(a))$  versus  $a$  com  $\omega(a) = \omega_0 + aI$ .

	EMV	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\mu}$	$\hat{\sigma}_x^2$	$\hat{\sigma}_u^2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Dados Completos		0.147	0.454	1.759	0.539	0.482	0.102	0.267
Sem a obs. 13		0.135	0.464	1.761	0.593	0.366	0.091	0.312
		-0.082	0.022	0	0.100	-0.241	-0.108	0.169
Sem a obs. 18		0.148	0.431	1.724	0.523	0.486	0.107	0.211
		0.007	-0.051	-0.02	-0.03	0.008	0.049	-0.210
Sem a obs. (13, 15)		0.137	0.448	1.748	0.638	0.286	0.119	0.355
		-0.068	-0.013	-0.006	0.184	-0.407	0.167	0.330
Sem a obs. (4, 21)		0.130	0.480	1.682	0.501	0.489	0.080	0.258
		-0.116	0.057	-0.044	-0.054	0.015	-0.216	-0.034
Sem a obs. (15, 17, 18)		0.158	0.392	1.645	0.464	0.396	0.138	0.172
		0.075	-0.137	-0.065	-0.139	-0.178	0.353	-0.356

Tabela 14 – Dados de escova de dente: parâmetros estimados e mudança relativa.

Na Figura 45 foi construído o gráfico de  $LD(\omega(a))$  considerando a ponderação de casos.

Na Figura 42 a observação 13 está acima do ponto de corte, seguida das observações 4, 21 e 19 que estão destacadas das demais observações porém abaixo do ponto de corte. Observa-se na Figura 45 que claramente a observação 13 tem uma grande influência seguida das observações 4, 21 e 19 que provocam uma menor mudança no  $LD(\omega(a))$ .

A seguir vamos aplicar a metodologia de CNCPPF no modelo definido para os dados de líquido de bochecho.

## 7.2 Dados do líquido de bochecho

Considerando o modelo definido em (1.11) vamos utilizar as perturbações que foram definidas em 6.1 para a aplicação da CNCPPF nos dados do líquido de bochecho.

### 7.2.1 Aplicação

Considerando os dados de líquido de bochecho descritos na Introdução, a metodologia de CNCPPF foi aplicada aos quatro esquemas de perturbação descritos anteriormente. O EMV dos parâmetros foram obtidos usando o algoritmo EM descrito em Russo, Aoki e Jr (2009). Na Figura 46, o primeiro e o segundo gráfico referem-se ao gráfico da CNCPPF com  $q$  suficientemente grande para que a contribuição dos autovetores associados aos dois maiores autovalores e a contribuição do autovetor associado ao maior autovalor sejam considerados, respectivamente. A última iteração fornece o gráfico de  $m(q)$  com os dados completos, o que mostra claramente que apenas a observação 42 está acima do ponto de corte nos dois casos, o que significa que, com o gráfico de índice usual, a conclusão seria que a observação 42 pode ser influente.

Analisando o primeiro gráfico (CNCPPF), embora a observação 42 seja a que aparece na última iteração com todo o conjunto de dados como a possível observação influente, as observações (39,40,43) estão acima do ponto de corte na maior parte das iterações de  $m = 62$  a  $m = 96$  e eles são mascarados nas iterações finais. Eles também têm comportamento semelhante durante a evolução da procura passo a frente. Além disso, a observação 52 está próxima do ponto de corte. Além disso, as observações (7,22) destacam-se na parte central da evolução da procura passo a frente, enquanto que as observações (7,19) apresentam comportamento semelhante. O segundo gráfico leva à mesma conclusão em relação à última iteração. As observações (39,40,43) apresentam também comportamentos semelhantes durante a evolução da procura passo a frente e são mascaradas nas etapas finais. A observação 67 aparece como influente em algumas etapas.

No primeiro gráfico da Figura 47 a observação 42 fornece a maior mudança no afastamento pela verossimilhança seguido das observações 39 e 40. No segundo gráfico da Figura 47, as observações (39,40,43) ocasionam quase a mesma mudança no afastamento pela verossimilhança que as observações (39,40,43,42) quando os valores de  $a$  é negativo, enquanto que com valores positivos de  $a$  o primeiro dá uma mudança ligeiramente maior no afastamento pela verossimilhança. As observações (39,40,43,52) são as observações que dão a maior mudança

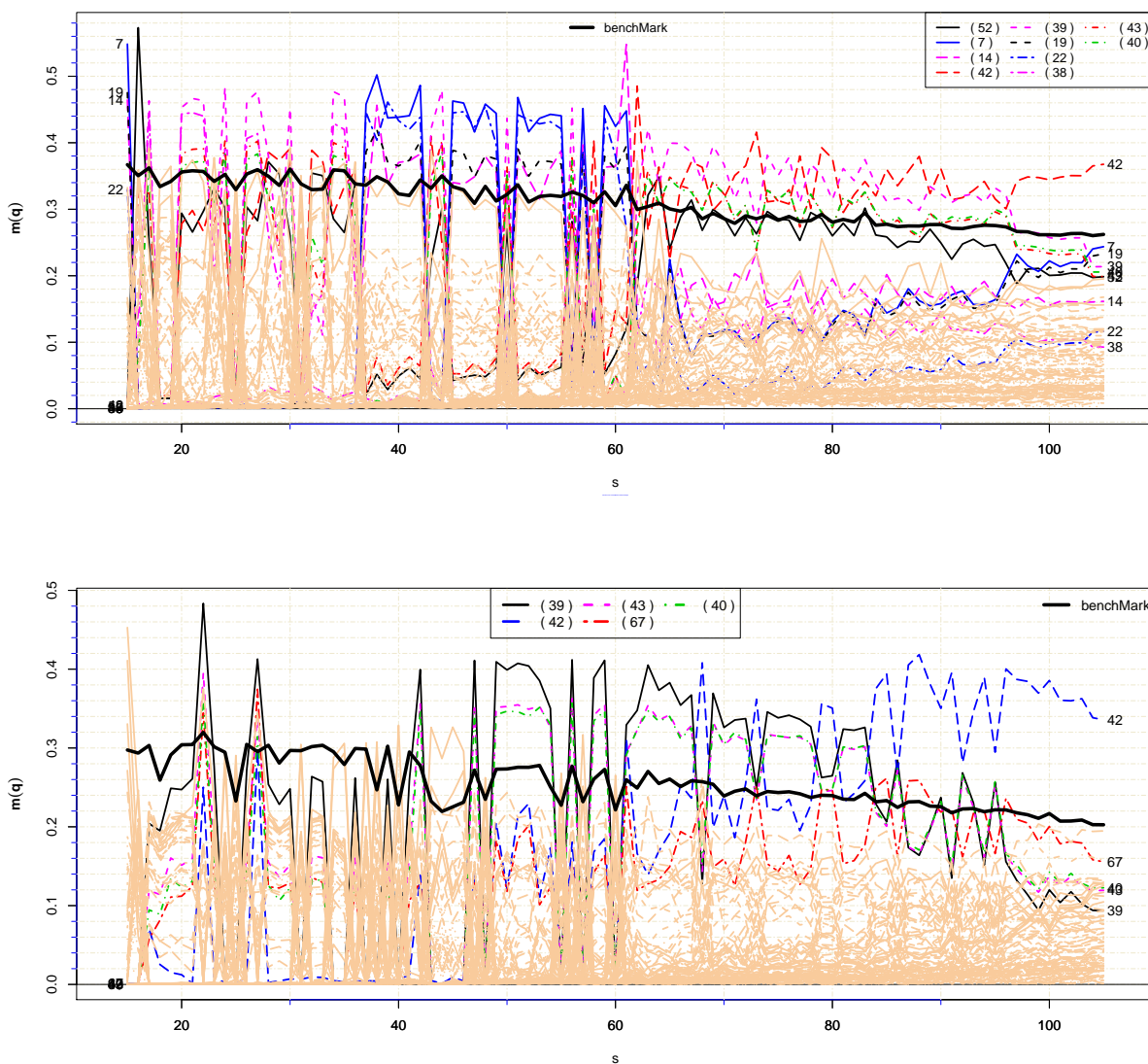


Figura 46 – Dados de líquido de bochecho: gráfico da CNCPPF para o esquema de perturbação da ponderação de casos. Contribuição agregada dos dois maiores autovalores e dos autovetores associados e contribuição do autovetor associado ao maior autovalor.

no afastamento pela verossimilhança. As observações (39,40,43,52) que são conjuntamente influente são todas observações de indivíduos que usaram o líquido de bochecho experimental A, bem como as observações 42, 52 e 67. As observações 7, 19 e 22 pertencem ao grupo controle. A Tabela 15 e a Tabela 16 fornecem os EMV dos parâmetros considerando todo o conjunto de dados e sem a observação 42 e o grupo de observações analisados anteriormente. O grupo de observações (39,40,43,52) é a que ocasiona maior mudança relativa na maioria das estimativas.

A Figura 48 mostra o gráfico passo a frente da contribuição agregada dos dois maiores autovalores e dos autovetores associados, considerando o esquema de perturbação nas variáveis explanatórias. Nesse caso, o gráfico de índice usual de  $m(q)$ , que é a última iteração do gráfico de passo a frente, levaria à conclusão de que nenhuma das observações são influentes. Todas as

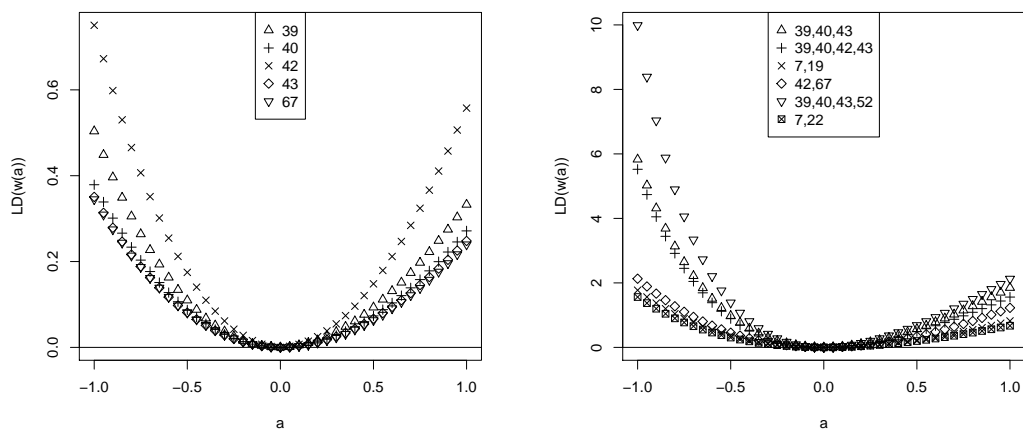


Figura 47 – Dados de líquido de bochecho: ponderação de casos - gráfico de  $LD(\boldsymbol{\omega}(a))$  versus  $a$  com  $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{l}$ .

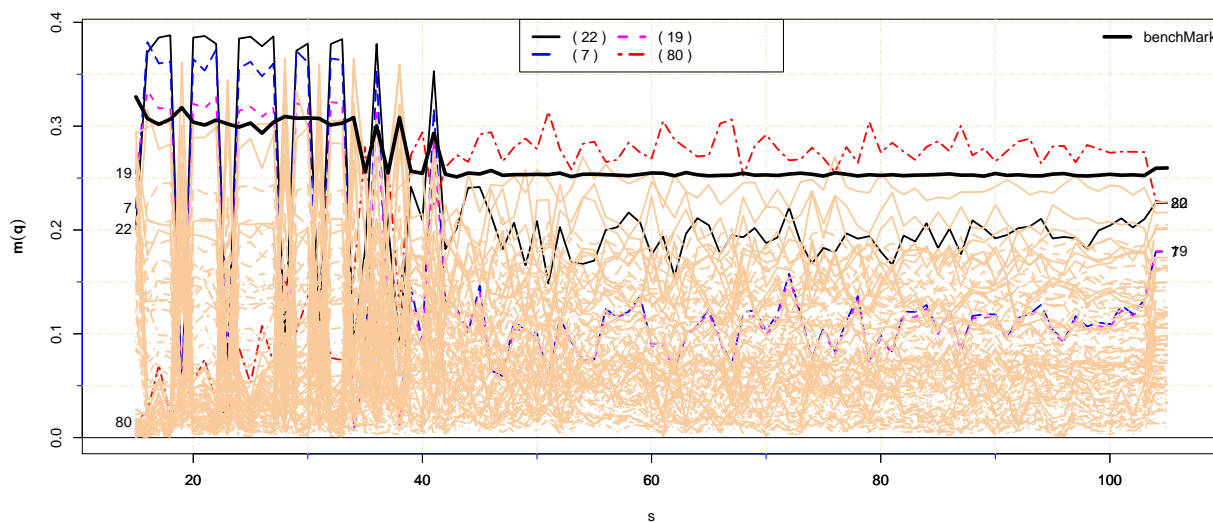


Figura 48 – Dados de líquido de bochecho: gráfico da CNCPPF para esquema de perturbação nas variáveis explanatórias. Contribuição agregada dos dois maiores autovalores e dos autovetores associados.

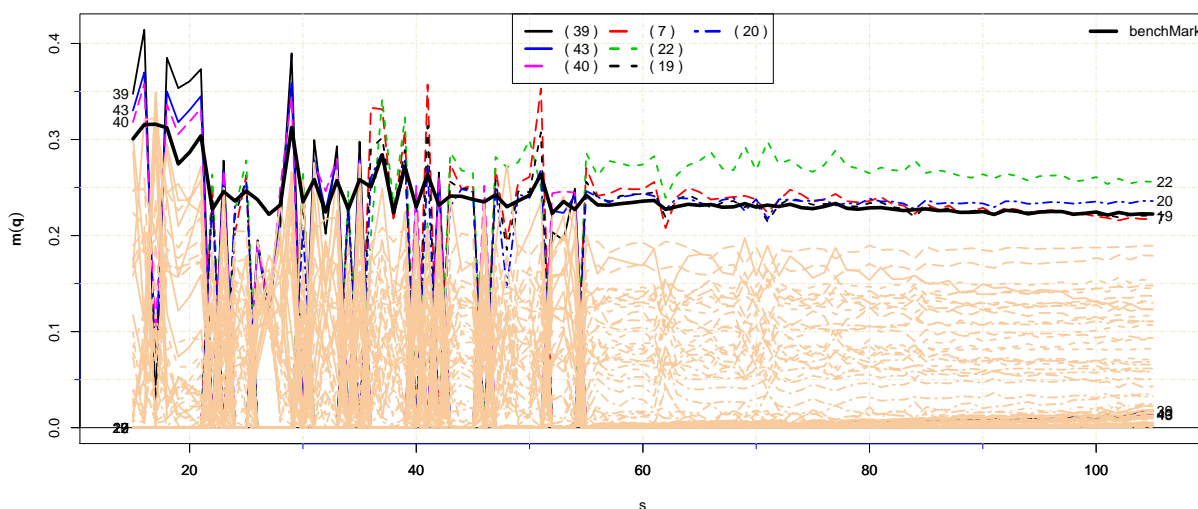


Figura 49 – Dados de líquido de bochecho: gráfico da CNCPPF para o esquema de perturbação nas variáveis respostas. Contribuição do autovetor associado ao maior autovalor.

observações estão abaixo do ponto de corte. No entanto, a observação 80 está acima do ponto de corte a partir da iteração  $m = 40$  e é mascarada nas etapas finais. A observação 80 pertence ao grupo que utilizou o líquido de bochecho experimental B e, entre esses indivíduos, foi o que apresentou o segundo maior valor do índice de placa no início do tratamento e também o que apresentou a menor redução no índice de placa dentária do início do estudo até o final do estudo. A Tabela 18 mostra que, com nível de significância 10 %, rejeitamos a hipótese de que a taxa de redução do índice de placa dentária após 6 meses usando o líquido de bochecho experimental A e B seja a mesma, enquanto que com a remoção da observação 80 leva à conclusão oposta. A Figura 49 fornece o gráfico passo a frente da CNCPPF considerando o esquema de perturbação nas variáveis respostas e a contribuição do autovetor associado ao maior autovalor. As observações (20,22) estão acima do ponto de corte na etapa final da evolução do gráfico passo a frente, mas as observações (7,19) se mantêm acima do ponto de corte na maior parte das iterações entre  $m = 36$  a  $m = 88$  e depois mantêm-se muito próximo do ponto de corte. O gráfico à esquerda na Figura 51 mostra que a observação 22 apresenta a maior mudança no afastamento pela verossimilhança. Para valores positivos de  $a$ , esta observação é seguida pela observação 7 e pela observação 20, juntamente com a observação 19. No lado negativo de  $a$  as observações 7, 19 e 20 apresentam a mesma alteração no afastamento pela verossimilhança. O gráfico da direita mostra que as observações (7,19) apresentam a mesma alteração no LD que as observações (7, 19,20,22) no lado positivo de  $a$ . Considerando a Tabela 15 e a Tabela 16 a maior mudança na estimativa dos parâmetros quando as observações (7,19), (20,22) ou (7,19,20,22) são descartadas do conjunto de dados está no parâmetro de variância do erro de medição do índice de placa após 6 meses para o grupo controle. A Figura 50 mostra o gráfico do índice de  $m(q)$  correspondente à última iteração da Figura 49, em que os índices 1 a 36 no eixo  $x$ , referem-se aos dados associados aos indivíduos que usaram o líquido de bochecho controle, enquanto os índices 37 a 69 (70 a 105) se referem às

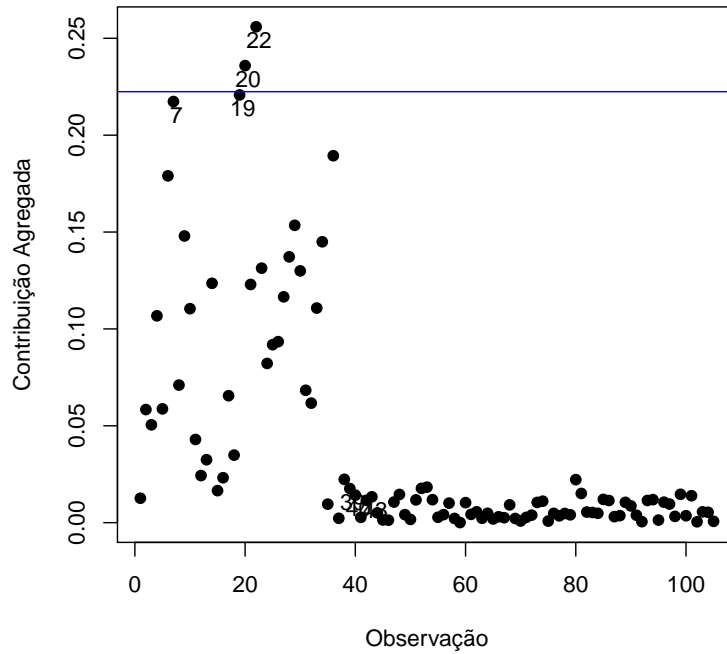


Figura 50 – Gráfico do valor absoluto do autovetor associado ao maior autovalor na metodologia CNCPF para os dados de bochecho.

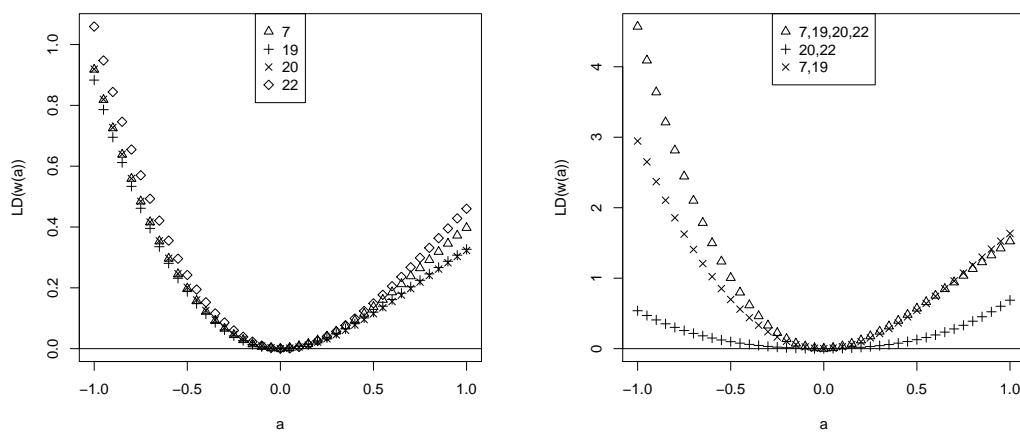


Figura 51 – Dados de líquido de bochecho: perturbação na variável resposta - gráfico de  $LD(\omega(a))$  versus  $a$  com  $\omega(a) = \omega_o + aI$ .

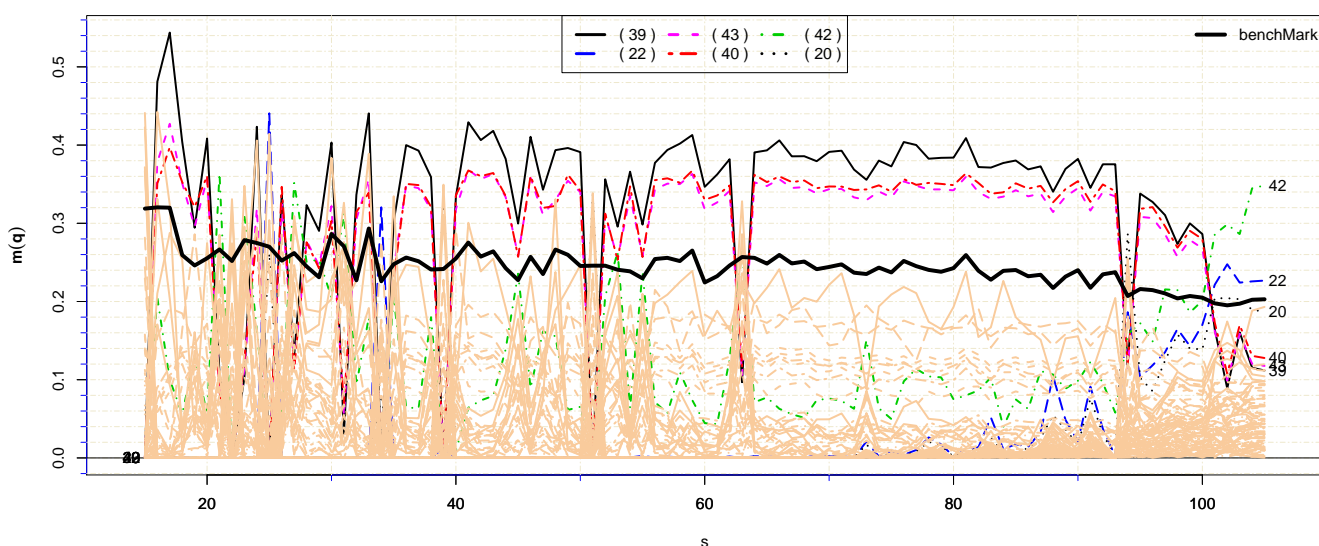


Figura 52 – Dados de líquido de bochecho: gráfico da CNCPPF para esquema de perturbação na variância. Contribuição do autovetor associado ao maior autovalor.

observações correspondentes aos indivíduos que usaram o líquido de bochecho experimental A (B). Pode-se observar que pequenas alterações locais na variável resposta que correspondem ao tratamento de líquido de bochecho controle tem um efeito maior nas estimativas dos parâmetros.

Considerando o esquema de perturbação na variância (Figura 52), as observações 42 e 22 estão acima do ponto de corte no último passo a CNCPPF, embora as observações (22,42) seja o grupo de observações que apresenta a menor alteração no afastamento pela verossimilhança entre o grupo de observações considerado na Figura 53. Porém, a Tabela 18 mostra que a remoção dessas observações altera a conclusão do teste de hipóteses ao nível de significância 10%. Além disso, a Figura 52 mostra que, embora as observações 42 e 22 surjam nas etapas finais do gráfico passo a frente, as observações 39, 40 e 43 estão acima do ponto de corte em quase toda a evolução do algoritmo de procura passo a frente. A Figura 53 mostra o gráfico do afastamento pela verossimilhança na direção das observações (39,40,43), (39,40,43,42), (39,40,43,22) e (39,40,43,22,42). Pode-se observar que a mudança no  $LD(\omega(a))$  no lado positivo de  $a$  é a mesma na direção de (39,40,43) e (39,40,43,42) e também o mesmo ocorre na direção de (39,40,43,22) e (39,40,43,22,42), o que significa que a inclusão de 42 não fez diferença nos dois casos. A Tabela 17 mostra que a exclusão do grupo de observações (39,40,43), (39,40,43,22) ou (39,40,43,22,42) altera a conclusão da hipótese testada com nível de significância 1%, enquanto que a exclusão de (39,40,43,42) não faria diferença. Considerando a hipótese apresentada na Tabela 18, a exclusão de um dos quatro grupos descritos altera o resultado do teste de hipótese ao nível de significância 5%.

	completo	sem as observações					
		42	(39, 40, 43)	(39, 40, 43, 42)	(39, 40, 43, 52)	(7, 19)	(7, 22)
$\hat{\beta}_{11}$	0.703	0.703	0.703	0.703	0.698	0.716	0.700
Mudança relat.		0.000	0.000	0.000	-0.007	0.018	-0.004
$\hat{\beta}_{21}$	0.687	0.688	0.687	0.687	0.689	0.706	0.685
Mudança relat.		0.001	0.000	0.000	0.003	0.028	-0.003
$\hat{\beta}_{12}$	0.525	0.509	0.574	0.560	0.588	0.525	0.525
Mudança relat.		-0.030	0.093	0.067	0.120	0.000	0.000
$\hat{\beta}_{22}$	0.502	0.491	0.538	0.529	0.520	0.502	0.502
Mudança relat.		-0.022	0.072	0.054	0.036	0.000	0.000
$\hat{\beta}_{13}$	0.508	0.508	0.508	0.508	0.497	0.508	0.508
Mudança relat.		0.000	0.000	0.000	-0.022	0.000	0.000
$\hat{\beta}_{23}$	0.414	0.414	0.414	0.414	0.409	0.414	0.414
Mudança relat.		0.000	0.000	0.000	-0.012	0.000	0.000
$\hat{\mu}$	2.534	2.524	2.533	2.523	2.525	2.543	2.538
Mudança relat.		-0.004	0.000	-0.004	-0.004	0.004	0.002
$\hat{\sigma}_{\delta}^2$	0.011	0.011	0.010	0.010	0.103	0.012	0.011
Mudança relat.		0.000	-0.091	-0.091	8.364	0.091	0.000
$\hat{\sigma}_x^2$	0.103	0.093	0.106	0.096	0.002	0.099	0.103
Mudança relat.		-0.097	0.029	-0.068	-0.981	-0.039	0.000
$\hat{\sigma}_{e_{11}}^2$	0.312	0.313	0.312	0.313	0.477	0.301	0.304
Mudança relat.		0.003	0.000	0.003	0.529	-0.035	-0.026
$\hat{\sigma}_{e_{21}}^2$	0.234	0.234	0.233	0.234	0.339	0.181	0.176
Mudança relat.		0.000	-0.004	0.000	0.449	-0.226	-0.248
$\hat{\sigma}_{e_{12}}^2$	0.430	0.412	0.29	0.277	0.364	0.430	0.430
Mudança relat.		-0.042	-0.326	-0.356	-0.153	0.000	0.000
$\hat{\sigma}_{e_{22}}^2$	0.432	0.431	0.373	0.376	0.454	0.432	0.432
Mudança relat.		-0.002	-0.137	-0.130	0.051	0.000	0.000
$\hat{\sigma}_{e_{13}}^2$	0.255	0.255	0.255	0.255	0.292	0.255	0.255
Mudança relat.		0.000	0.000	0.000	0.145	0.000	0.000
$\hat{\sigma}_{e_{23}}^2$	0.192	0.192	0.192	0.192	0.197	0.192	0.192
Mudança relat.		0.000	0.000	0.000	0.026	0.000	0.000

Tabela 15 – EMV e mudanças relativas nas estimativas dos parâmetros.

	completo	sem as observações						
		(42, 67)	(80)	(20, 22)	(7, 19, 20, 22)	(22, 42)	(39, 40, 43, 22)	(39, 40, 43, 42, 22)
$\hat{\beta}_{11}$	0.703	0.703	0.703	0.704	0.718	0.695	0.695	0.695
Mudança relat.		0.000	0.000	0.001	0.021	-0.011	-0.011	-0.011
$\hat{\beta}_{21}$	0.687	0.688	0.687	0.684	0.704	0.675	0.675	0.675
Mudança relat.		0.001	0.000	-0.004	0.025	-0.017	-0.017	-0.017
$\hat{\beta}_{12}$	0.525	0.490	0.525	0.525	0.525	0.509	0.574	0.560
Mudança relat.		-0.067	0.000	0.000	0.000	-0.030	0.093	0.067
$\hat{\beta}_{22}$	0.502	0.486	0.502	0.502	0.502	0.491	0.538	0.529
Mudança relat.		-0.032	0.000	0.000	0.000	-0.022	0.072	0.054
$\hat{\beta}_{13}$	0.508	0.508	0.517	0.508	0.508	0.508	0.508	0.508
Mudança relat.		0.000	0.018	0.000	0.000	0.000	0.000	0.000
$\hat{\beta}_{23}$	0.414	0.414	0.424	0.414	0.414	0.414	0.414	0.414
Mudança relat.		0.000	0.024	0.000	0.000	0.000	0.000	0.000
$\hat{\mu}$	2.534	2.518	2.528	2.535	2.544	2.524	2.533	2.523
Mudança relat.		-0.006	-0.002	0.000	0.004	-0.004	0.000	-0.004
$\hat{\sigma}_{\delta}^2$	0.011	0.010	0.010	0.010	0.012	0.010	0.010	0.010
Mudança relat.		-0.091	-0.091	-0.091	0.091	-0.091	-0.091	-0.091
$\hat{\sigma}_x^2$	0.103	0.091	0.101	0.105	0.101	0.095	0.108	0.098
Mudança relat.		-0.117	-0.019	0.019	-0.019	-0.078	0.049	-0.049
$\hat{\sigma}_{e_{11}}^2$	0.312	0.314	0.312	0.294	0.281	0.307	0.305	0.307
Mudança relat.		0.006	0.000	-0.058	-0.099	-0.016	-0.022	-0.016
$\hat{\sigma}_{e_{21}}^2$	0.234	0.234	0.234	0.188	0.131	0.205	0.204	0.205
Mudança relat.		0.000	0.000	-0.197	-0.440	-0.124	-0.128	-0.124
$\hat{\sigma}_{e_{12}}^2$	0.430	0.373	0.430	0.430	0.430	0.413	0.290	0.277
Mudança relat.		-0.133	0.000	0.000	0.000	-0.040	-0.326	-0.356
$\hat{\sigma}_{e_{22}}^2$	0.432	0.442	0.431	0.431	0.432	0.431	0.372	0.376
Mudança relat.		0.023	-0.002	-0.002	0.000	-0.002	-0.139	-0.130
$\hat{\sigma}_{e_{13}}^2$	0.255	0.255	0.251	0.255	0.255	0.255	0.255	0.255
Mudança relat.		0.000	-0.016	0.000	0.000	0.000	0.000	0.000
$\hat{\sigma}_{e_{23}}^2$	0.192	0.192	0.184	0.192	0.192	0.192	0.192	0.192
Mudança relat.		0.000	-0.042	0.000	0.000	0.000	0.000	0.000

Tabela 16 – EMV e mudanças relativas nas estimativas dos parâmetros.

Dados	sem as observações						
	completos	(39,40,43)	(39,40,43,42)	(39,40,43,52)	(39,40,43,52,42)	(39,40,43,22)	(39,40,43,22,42)
p-value	0.002	0.015	0.007	0.082	0.018	0.022	0.011

Tabela 17 – Teste de hipóteses:  $H_0 : \beta_{11} = \beta_{12}$ 

Dados	sem as observações											
	completos	42	52	67	80	(39,40,43)	(42,22)	(42,67)	(40,43)	(39,40,43,22)	(39,40,43,42)	(39,40,43,22,42)
p-value	0.096	0.154	0.163	0.122	0.139	0.017	0.154	0.194	0.030	0.017	0.032	0.032

Tabela 18 – Teste de hipóteses:  $H_0 : \beta_{22} = \beta_{23}$

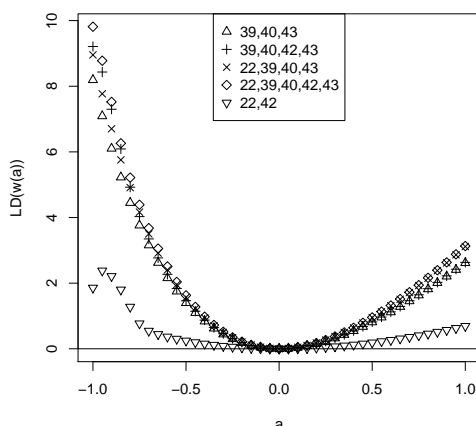


Figura 53 – Dados de líquido de bochecho: perturbação na variância - gráfico de  $LD(\boldsymbol{w}(a))$  versus  $a$  com  $\boldsymbol{w}(a) = \boldsymbol{w}_o + a\boldsymbol{l}$ .

### 7.3 Dados simulados

Considerando o modelo definido em (1.4) e (1.5) foram geradas 26 observações com  $\beta_1 = 0.156$ ,  $\beta_2 = 0.436$ ,  $\mu = 1.769$ ,  $\sigma_x^2 = 0.408^2$ ,  $\sigma_u^2 = 0.562^2$ ,  $\lambda_1 = 0.074^2$  e  $\lambda_2 = 0.363^2$ . Na Figura 54 temos o gráfico de dispersão do dados gerados. Considerando as perturbações na variância foi construído o gráfico da CNCPPF na Figura 55. Observe que não temos pontos influentes. Sejam  $\mathbf{Z}_7 = (X_{1,7}, X_{2,7}, Y_{1,7}, Y_{2,7})^T$ ,  $\mathbf{Z}_{12} = (X_{1,12}, X_{2,12}, Y_{1,12}, Y_{2,12})^T$ , e  $\mathbf{Z}_{24} = (X_{1,24}, X_{2,24}, Y_{1,24}, Y_{2,24})^T$ , ou seja, as observações 7, 12 e 24. Estas observações foram alterados da seguinte forma para podermos introduzir os dados atípicos:  $\mathbf{Z}_{7(\text{alterado})} = \mathbf{Z}_7 + (0.4, 2.6, 0.5, 2.5)^T$ ;  $\mathbf{Z}_{12(\text{alterado})} = \mathbf{Z}_{12} + (0.2, 0.3, 1.3, 0.7)^T$  e  $\mathbf{Z}_{24(\text{alterado})} = \mathbf{Z}_{24} + (1.0, 1.8, 1.8, 0.5)^T$ .

O gráfico de dispersão com as observações alteradas está na Figura 56. Observamos que os dados 7 e 24 podem ser atípicos nos três gráficos e o dado 12 no primeiro gráfico.

Foi construído o gráfico de contribuição agregada considerando o maior autovalor (Figura 57). A observação 7 aparece como dado influente.

Na Figura 58 foi construído o gráfico da CNCPPF na direção do maior autovalor. A observação 12 se encontra acima do ponto de corte até a penúltima iteração, no entanto é mascarado na última iteração. Enquanto que a observação 24 fica em torno do ponto de corte e também é mascarado na última iteração. Por outro lado, a observação 7 aparece abaixo do ponto de corte e na última iteração quando entram as observações 15 e 24 e sai a observação 12 dá um salto e fica acima do ponto de corte.

No gráfico da Figura 59 observamos que o parâmetro  $\lambda_2$  apresenta a maior variação na estimativa dos parâmetros ao perturbar a variância na direção máxima seguida do parâmetro  $\sigma_u^2$ . No gráfico da Figura 60 foram construídos os gráficos de mudança relativa na estimativa dos parâmetros na direção da observação 7 (Figura 60a) quando a variância é perturbada. O

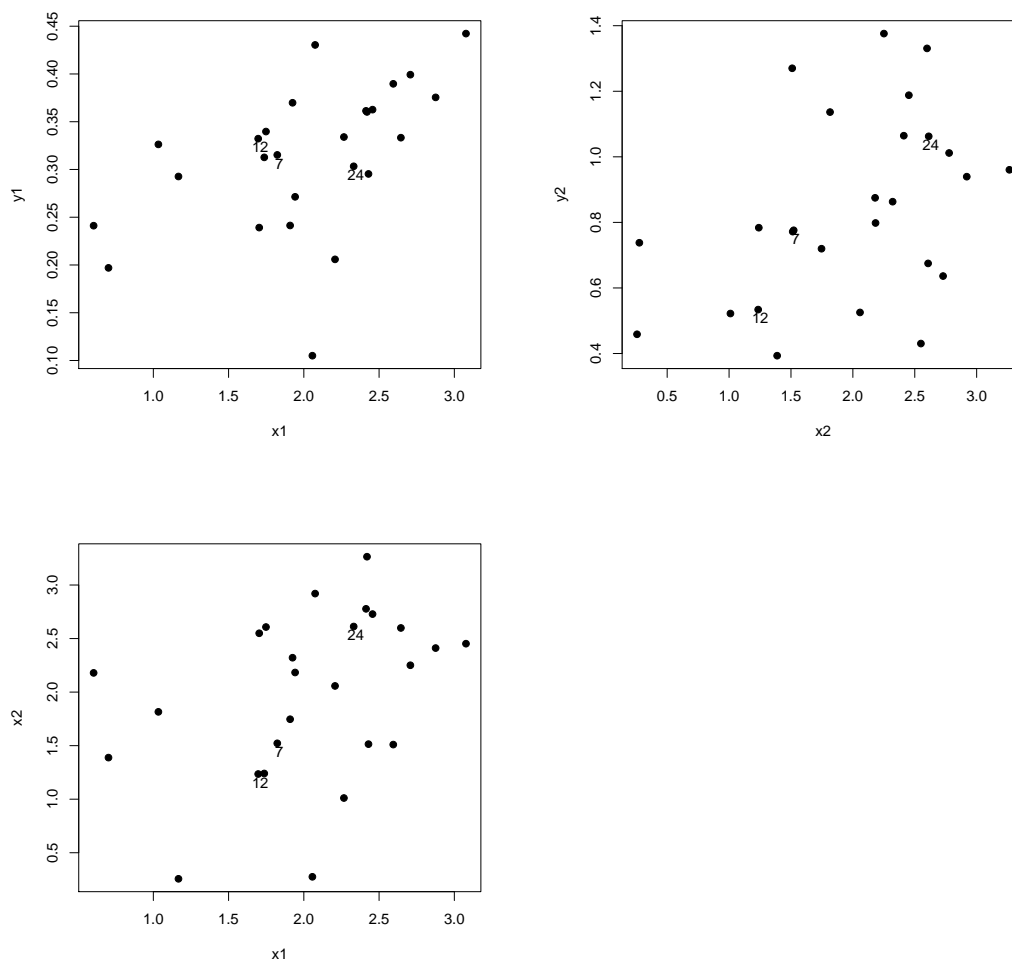


Figura 54 – Gráfico de dispersão dos dados simulados.

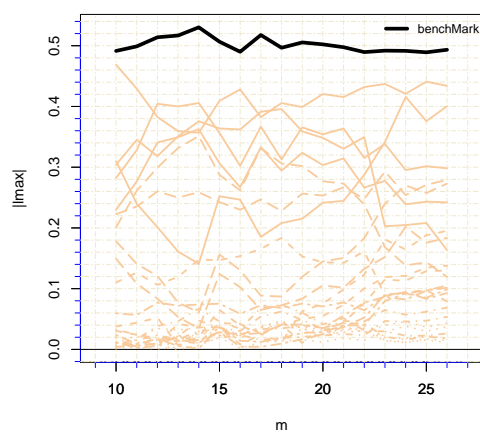


Figura 55 – Perturbação na variância.

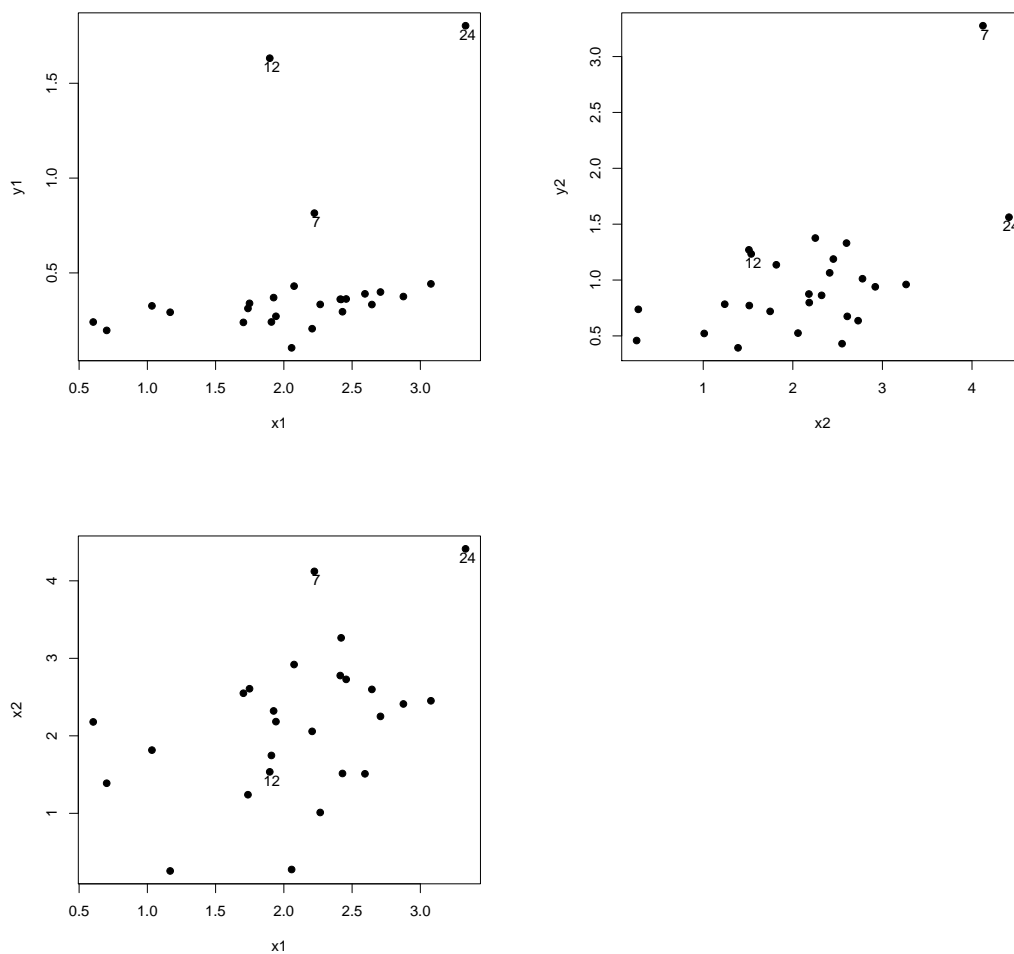


Figura 56 – Gráfico de dispersão dos dados simulados alterados.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\mu}$	$\hat{\sigma}_x^2$	$\hat{\sigma}_u^2$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Dados completos	0.2196	0.4759	2.1046	0.3872	0.3525	0.3225	0.5074
<b>sem a observação 7</b>	0.2187	0.4307	2.069	0.3185	0.3378	0.3355	0.1129
	0	9	2	18	4	4	78
<b>sem a observação 12</b>	0.1993	0.4689	2.1125	0.448	0.346	0.1681	0.5202
	9	1	0	16	2	48	3
<b>sem a observação 24</b>	0.1905	0.4832	2.0386	0.2572	0.3513	0.1982	0.5488
	13	2	3	34	0	39	8
<b>sem a observação 7,12,24</b>	0.1561	0.4212	2.0134	0.1979	0.3207	0.0047	0.1322
	29	12	4	49	9	99	74

Tabela 19 – Perturbação na variância: mudança nas estimativas dos parâmetros dos dados simulados.

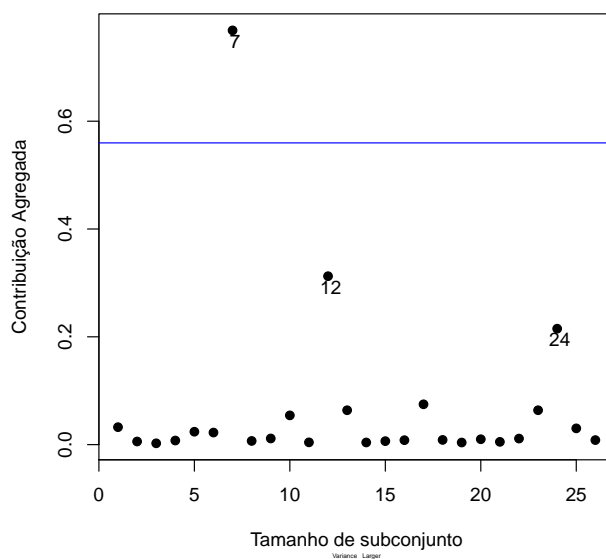


Figura 57 – Perturbação na variância: gráfico de  $m(q)$  com o índice das observações. Contribuição do autovetor associado ao maior autovalor. Dados alterados.

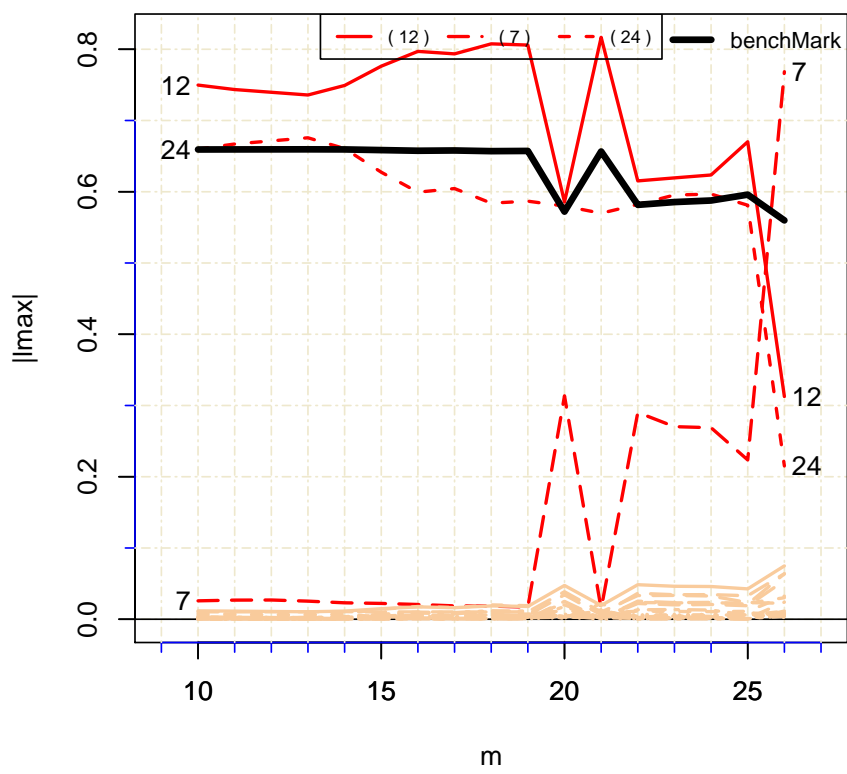


Figura 58 – Perturbação na variância: gráfico da CNCPPF. Dados alterados.

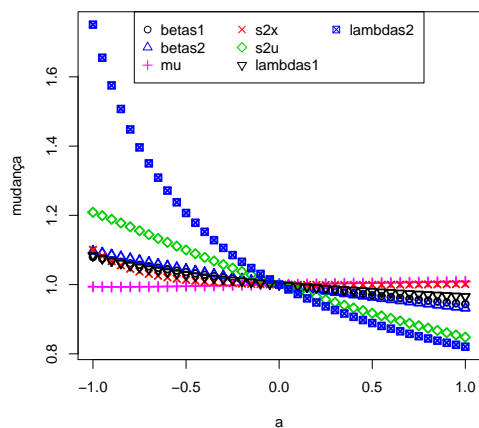


Figura 59 – Perturbação na variância: mudanças relativas nas estimativas dos parâmetros na direção do autovetor associado ao maior autovalor. Dados alterados.

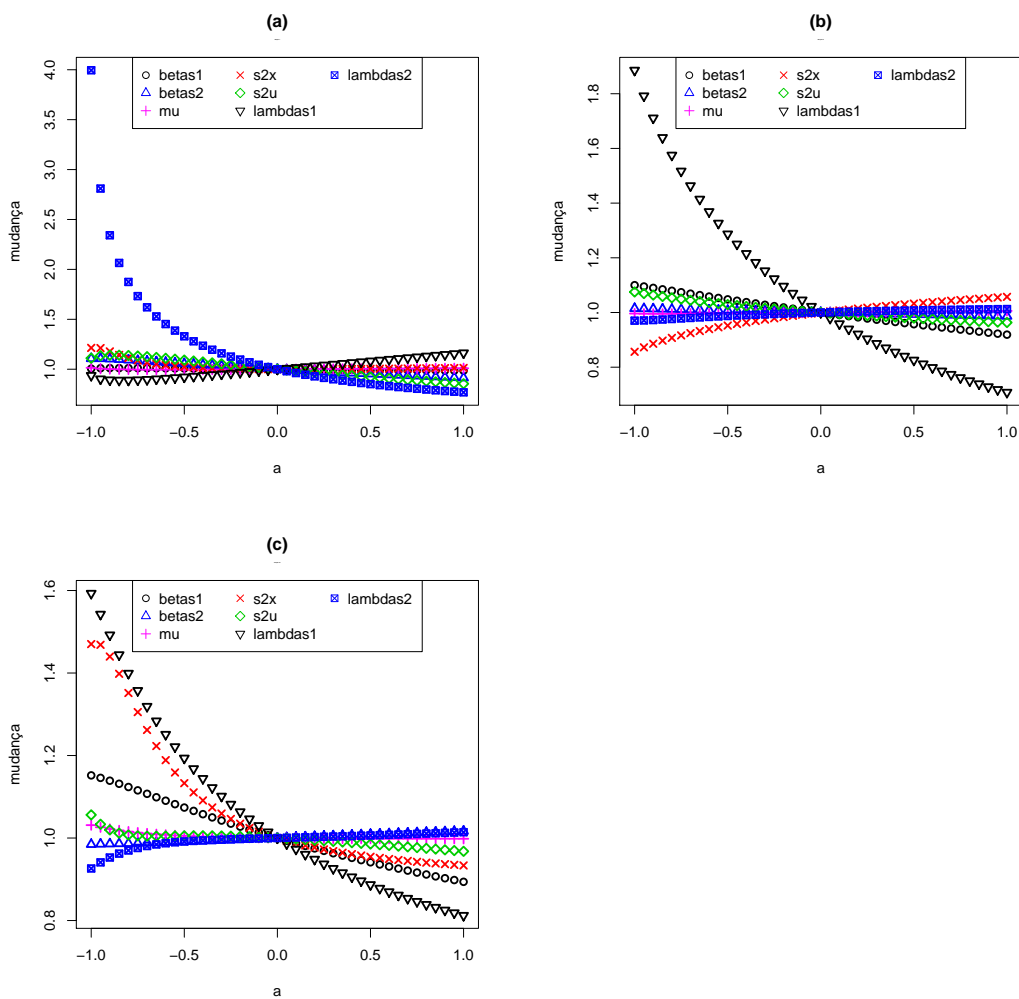


Figura 60 – Perturbação na variância: mudança relativa nas estimativas dos parâmetros na direção de (a) observação 7, (b) observação 12 e (c) observação 24. Dados alterados.

parâmetro  $\lambda_2$  é a que sofre a maior alteração quando é perturbada na direção da observação 7, e na direção da observação 12 (Figura 60b) o parâmetro  $\lambda_1$  sofre a maior alteração. Quando é perturbado na direção da observação 24 (Figura 60c), os parâmetros  $\lambda_1$  e  $\sigma_x^2$  são os que mais mudam, seguido de  $\beta_1$  na direção negativa de  $a$ .

Na Tabela 19 observamos que ao excluir a observação 7 o parâmetro que sofre a maior mudança relativa na estimativa é o  $\lambda_2$ , que é o parâmetro que sofre a maior mudança relativa quando perturbamos na direção da observação 7 (Figura 60a). O mesmo acontece com as observações 12 e 24, ou seja ao excluirmos a observação 12(24) a maior mudança relativa ocorre na estimativa do parâmetro  $\lambda_1$ , que é o parâmetro que sofrem a maior mudança relativa quando foi perturbada na direção da observação 12(24). Nota-se também que o parâmetro que sofre a maior mudança relativa na estimativa é a variância do verdadeiro valor do índice de placa dentária no início do estudo.



---

## CONCLUSÕES

---

Nesta tese apresentamos as propriedades assintóticas do modelo de calibração ultraestrutural com réplicas, e provamos que o estimador de máxima verossimilhança dos parâmetros não é assintoticamente Normal. Com base nisso, desenvolvemos os testes de hipótese assintótico para testar os parâmetros do viés aditivo e multiplicativo do modelo proposto. Os testes de hipótese de Wald propostos no Capítulo 2, apresentaram resultados satisfatórios nos estudos de simulação, nos itens 2.4.1 e 2.4.2.

Foram propostos também novas metodologias de análise de diagnóstico, a ILPPF e CNCPPF para os modelos de regressão com erros nas variáveis. Estes foram capazes de detectar dados mascarados influentes em três modelos diferentes de regressão com erros nas variáveis.

Para o caso da CNCPPF concluímos que ela facilitou a detecção de dados mascarados devido ao benchmark que ajuda nesse propósito. Ao calcular a contribuição agregada para três cenários; com o maior autovalor, os dois maiores autovalores e todos os autovalores; a CNCPPF apresentou melhor performance com os dois maiores autovalores na detecção de observações influentes mascarados.



---

## ESTUDOS FUTUROS

---

---

- a) Uma extensão a ser considerada é desenvolver as propriedades assintóticas do modelo de calibração ultraestrutural com réplicas para outros testes assintóticos como score, gradiente e razão de verossimilhança;
- b) Uma outra extensão a ser considerada é desenvolver as propriedades assintóticas do modelo de calibração ultraestrutural com réplicas para a distribuição t-student e/ou distribuição elíptica;
- c) Como uma continuação deste trabalho, vamos considerar os modelos de efeitos mistos e utilizar a metodologia de influencia local com procura passo a frente e/ou a metodologia de curvatura normal conformal com procura passo a frente neste modelo;
- d) Aplicar a metodologia proposta de [ILPPF](#) e [CNCPPF](#) no modelo de calibração ultraestrutural com réplicas;
- e) Utilizando a [ILPPF](#) fazer a definição de arrastamento e analisar o problema de arrastamento para os modelos utilizados.
- f) Poderia criar um índice de estabilidade do processo ou evolução de [ILPPF](#).



## REFERÊNCIAS

---

- AOKI, R. **Modelos de regressão com erros nas variáveis com intercepto nulo**. Tese (Doutorado) — Departamento de Estatística, Universidade de São Paulo, 2001. Citado na página [32](#).
- AOKI, R.; BOLFARINE, H.; ACHCAR, J. A.; JR, L. P. D. Bayesian analysis of a multivariate null intercept errors-in-variables regression model. **Journal of biopharmaceutical statistics**, Taylor & Francis, v. 13, n. 4, p. p. 767–775, 2003. Citado nas páginas [32](#), [34](#) e [105](#).
- AOKI, R.; BOLFARINE, H.; SINGER, J. Asymptotic efficiency of method of moments estimators under null intercept measurement error regression models. **Brazilian Journal of Probability and Statistics**, v. 16, p. p. 157–166, 2002. Citado na página [32](#).
- AOKI, R.; BOLFARINE, H.; SINGER, J. M. Null intercept measurement error regression models. **Test**, Springer, v. 10, n. 2, p. p. 441–457, 2001. Citado nas páginas [32](#), [111](#) e [113](#).
- AOKI, R.; SINGER, J. da M.; BOLFARINE, H. Local influence for measurement error regression models for the analysis of pretest/posttest data. **Journal of Applied Statistical Science**, v. 15, n. 3, p. p. 317–330, 2007. Citado nas páginas [32](#), [67](#), [71](#), [111](#), [113](#) e [114](#).
- ATKINSON, A. Fast very robust methods for the detection of multiple outliers. **Journal of the American Statistical Association**, Taylor & Francis, v. 89, n. 428, p. 1329–1339, 1994. Citado na página [63](#).
- ATKINSON, A.; MULIRA, H.-M. The stalactite plot for the detection of multivariate outliers. **Statistics and Computing**, Springer, v. 3, n. 1, p. p. 27–35, 1993. Citado na página [26](#).
- ATKINSON, A.; RIANI, M. **Robust diagnostic regression analysis**. [S.l.]: Springer, 2000. Citado nas páginas [19](#), [26](#), [31](#), [32](#), [59](#), [63](#), [64](#), [65](#), [67](#), [68](#) e [70](#).
- \_\_\_\_\_. Cerioli a. **Exploring Multivariate Data with the Forward Search**, 2004. Citado na página [27](#).
- \_\_\_\_\_. Building regression models with the forward search. **CIT. Journal of Computing and Information Technology**, SRCE-Sveučilišni računski centar, v. 15, n. 4, p. 287–294, 2007. Citado na página [27](#).
- ATKINSON, A.; RIANI, M. *et al.* **Exploring multivariate data with the forward search**. [S.l.]: Springer Science & Business Media, 2013. Citado na página [63](#).
- ATKINSON, A. C. **Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis**. [S.l.]: Clarendon Press Oxford, 1985. Citado nas páginas [25](#) e [26](#).
- \_\_\_\_\_. Econometric applications of the forward search in regression: Robustness, diagnostics, and graphics. **Econometric reviews**, Taylor & Francis, v. 28, n. 1-3, p. 21–39, 2008. Citado na página [27](#).

ATKINSON, A. C.; CERIOLI, A.; MORELLI, G.; RIANI, M. Finding the number of disparate clusters with background contamination. In: **Data Science, Learning by Latent Structures, and Knowledge Discovery**. [S.l.]: Springer, 2015. p. 29–42. Citado na página 28.

ATKINSON, A. C.; CERIOLI, A.; RIANI, M. Discussion of “Asymptotic Theory of Outlier Detection Algorithms for Linear Time Series Regression Models” by johansen and nielsen. **Scandinavian Journal of Statistics**, Wiley Online Library, v. 43, n. 2, p. 349–352, 2016. Citado na página 29.

ATKINSON, A. C.; CORBELLINI, A.; RIANI, M. Robust bayesian regression with the forward search: theory and data analysis. **TEST**, Springer, p. 1–18, 2017. Citado na página 29.

ATKINSON, A. C.; RIANI, M. Regression diagnostics for binomial data from the forward search. **Journal of the Royal Statistical Society: Series D (The Statistician)**, Wiley Online Library, v. 50, n. 1, p. 63–78, 2001. Citado na página 27.

\_\_\_\_\_. Forward search added-variable t-tests and the effect of masked outliers on model selection. **Biometrika**, Oxford University Press, v. 89, n. 4, p. 939–946, 2002. Citado na página 27.

ATKINSON, A. C.; RIANI, M.; CERIOLI, A. The forward search: Theory and data analysis. **Journal of the Korean Statistical Society**, Elsevier, v. 39, n. 2, p. 117–134, 2010. Citado na página 27.

ATKINSON, A. C.; RIANI, M.; TORTI, F. Robust methods for heteroskedastic regression. **Computational Statistics & Data Analysis**, Elsevier, v. 104, p. 209–222, 2016. Citado na página 29.

BECKMAN, R. J.; NACHTSHEIM, C. J.; COOK, R. D. Diagnostics for mixed–model analysis of variance. **Technometrics**, Taylor & Francis Group, v. 29, n. 4, p. 413–426, 1987. Citado na página 29.

BELLINI, T. Detecting atypical observations in financial data: the forward search for elliptical copulas. **Advances in data analysis and classification**, Springer, v. 4, n. 4, p. 287–299, 2010. Citado na página 28.

\_\_\_\_\_. Forward search outlier detection in data envelopment analysis. **European Journal of Operational Research**, Elsevier, v. 216, n. 1, p. 200–207, 2012. Citado na página 28.

\_\_\_\_\_. The forward search interactive outlier detection in cointegrated VAR analysis. **Advances in Data Analysis and Classification**, Springer, v. 10, n. 3, p. 351–373, 2016. Citado na página 29.

BELLINI, T.; RIANI, M. Robust analysis of default intensity. **Computational Statistics & Data Analysis**, Elsevier, v. 56, n. 11, p. 3276–3285, 2012. Citado na página 28.

BELSLEY, D. A.; KUH, E.; WELSCH, R. E. **Regression diagnostics: Identifying influential data and sources of collinearity**. New York: John Wiley and Sons, 1980. Citado nas páginas 25, 67 e 68.

BERTACCINI, B.; VARRIALE, R. Robust analysis of variance: An approach based on the forward search. **Computational statistics & data analysis**, Elsevier, v. 51, n. 10, p. 5172–5183, 2007. Citado na página 27.

BHAT, B. On the method of maximum-likelihood for dependent observations. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 48–53, 1974. Citado na página 32.

BILLINGSLEY, P. **Convergence of probability measures**. New York [u.a.]: Wiley, 1968. (Wiley series in probability and mathematical statistics). Citado na página 43.

BLAND, J. M.; ALTMAN, D. Statistical methods for assessing agreement between two methods of clinical measurement. **The lancet**, Elsevier, v. 327, n. 8476, p. 307–310, 1986. Citado na página 39.

BRYS, G.; HUBERT, M.; STRUYF, A. A robust measure of skewness. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 13, n. 4, p. 996–1017, 2004. Citado nas páginas 25 e 87.

CADIGAN, N. Local influence diagnostics for quasi-likelihood and lognormal estimates of a biological reference point from some fish stock and recruitment models. **Biometrics**, Wiley Online Library, v. 62, n. 3, p. 713–720, 2006. Citado na página 30.

CASTRO, M. D.; GALEA-ROJAS, M.; BOLFARINE, H. Local influence assessment in heteroscedastic measurement error models. **Computational Statistics & Data Analysis**, Elsevier, v. 52, n. 2, p. p. 1132–1142, 2007. Citado na página 31.

CERASA, A.; CERIOLI, A. Outlier-free merging of homogeneous groups of pre-classified observations under contamination. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 87, n. 15, p. 2997–3020, 2017. Citado na página 29.

CERIOLI, A.; FARCOMENI, A.; RIANI, M. Strong consistency and robustness of the forward search estimator of multivariate location and scatter. **Journal of Multivariate Analysis**, Elsevier, v. 126, p. 167–183, 2014. Citado na página 28.

CERIOLI, A.; RIANI, M. The ordering of spatial data and the detection of multiple outliers. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 8, n. 2, p. p. 239–258, 1999. Citado na página 26.

CHAN, L. K.; MAK, T. K. On the maximum likelihood estimation of a linear structural relationship when the intercept is known. **Journal of Multivariate Analysis**, Elsevier, v. 9, n. 2, p. p. 304–313, 1979. Citado na página 32.

CHATTERJEE, S.; HADI, A. S. **Sensitivity Analysis in Linear Regression**. [s.n.], 1988. XIV. pp. 315 p. (Wiley Series in Probability and Statistics, XIV). ISBN 9780471822165. Disponível em: <<http://books.google.com.br/books?id=D-IfpBnR0IEC>>. Citado nas páginas 25 e 26.

CHENG, C.-L.; NESS, J. W. V. **Statistical regression with measurement error**. [S.l.]: John Wiley & Sons, 1999. Citado nas páginas 23, 24 e 171.

COOK, R. D. Detection of influential observation in linear regression. **Technometrics**, JSTOR, v. 19, n. 1, p. p. 15–18, 1977. Citado nas páginas 25, 67 e 68.

\_\_\_\_\_. Assessment of local influence. **Journal of the Royal Statistical Society. Series B (Methodological)**, Wiley-Blackwell for the Royal Statistical Society, v. 48, n. 2, p. p. 133–169, 1986. ISSN 00359246. Citado nas páginas 19, 29, 30, 31, 32, 59, 60, 61, 67, 68, 70, 71, 79, 84, 86, 87, 89, 93, 97, 100 e 107.

- COOK, R. D.; WEISBERG, S. Residuals and influence in regression. **Monographs on Statistics and Applied Probability**., New York: Chapman and Hall, 1982. Citado nas páginas 25, 67 e 68.
- CORBELLINI, A.; RIANI, M.; ATKINSON, A. C. Hubert, rousseeuw and segaert: multivariate functional outlier detection. **Statistical Methods & Applications**, Springer, v. 24, n. 2, p. 257–261, 2015. Citado na página 25.
- DAI, X.; JIN, L.; SHI, L.; YANG, C.; LIU, S. Local influence analysis in general spatial models. **AStA Advances in Statistical Analysis**, Springer, v. 100, n. 3, p. 313–331, 2016. Citado na página 31.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. **Journal of the royal statistical society. Series B (methodological)**, JSTOR, p. 1–38, 1977. Citado nas páginas 78 e 174.
- DOLBY, G. R. The ultrastructural relation: A synthesis of the functional and structural relations. **Biometrika**, [Oxford University Press, Biometrika Trust], v. 63, n. 1, p. 39–50, 1976. Citado na página 24.
- ESCOBAR, L. A.; JR, W. Q. M. Assessing influence in regression analysis with censored data. **Biometrics**, JSTOR, p. 507–528, 1992. Citado na página 29.
- FULLER, W. A. **Measurement Error Models**. New York: John Wiley and Sons, 1987. Citado nas páginas 23 e 74.
- GALEA, M.; PAULA, G. A.; BOLFARINE, H. Local influence in elliptical linear regression models. **Journal of the Royal Statistical Society: Series D (The Statistician)**, Wiley Online Library, v. 46, n. 1, p. 71–79, 1997. Citado na página 30.
- GALEA, M.; PAULA, G. A.; URIBE-OPAZO, M. On influence diagnostic in univariate elliptical linear regression models. **Statistical Papers**, Springer, v. 44, n. 1, p. 23–45, 2003. Citado na página 30.
- GALEA-ROJAS, M.; BOLFARINE, H.; CASTRO, M. de. Local influence in comparative calibration models. **Biometrical journal**, Wiley Online Library, v. 44, n. 1, p. p. 59–81, 2002. Citado na página 67.
- GARCÍA-ESCUADERO, L. A.; GORDALIZA, A.; MATRÁN, C.; MAYO-ISCAR, A. *et al.* A general trimming approach to robust cluster analysis. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 36, n. 3, p. 1324–1345, 2008. Citado na página 29.
- GIMÉNEZ, P.; PATAT, M. L. Local influence for functional comparative calibration models with replicated data. **Statistical Papers**, Springer, v. 55, n. 2, p. 431–454, 2014. Citado na página 31.
- HADGU, A.; KOCH, G. Application of generalized estimating equations to a dental randomized clinical trial. **Journal of Biopharmaceutical Statistics**, Taylor & Francis, v. 9, n. 1, p. p. 161–178, 1999. Citado nas páginas 31, 34, 83 e 105.
- HADI, A. S. Identifying multiple outliers in multivariate data. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. p. 761–771, 1992. Citado nas páginas 26, 27 e 63.

HADI, A. S.; SIMONOFF, J. S. Procedures for the identification of multiple outliers in linear models. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 88, n. 424, p. p. 1264–1272, 1993. Citado nas páginas 26 e 27.

HAWKINS, D. M.; BRADU, D.; KASS, G. V. Location of several outliers in multiple regression data using elemental sets. **Technometrics**, Taylor & Francis Group, v. 26, n. 3, p. p. 197–208, 1984. Citado na página 25.

HINES, R. O.; LAWLESS, J.; CARTER, E. Diagnostics for a cumulative multinomial generalized linear model, with applications to grouped toxicological mortality data. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 87, n. 420, p. 1059–1069, 1992. Citado na página 29.

HOCHBERG, Y. A sharper bonferroni procedure for multiple tests of significance. **Biometrika**, Oxford University Press, v. 75, n. 4, p. 800–802, 1988. Citado na página 57.

HOLM, S. A simple sequentially rejective multiple test procedure. **Scandinavian journal of statistics**, JSTOR, p. 65–70, 1979. Citado na página 57.

HOMMEL, G. A stagewise rejective multiple test procedure based on a modified bonferroni test. **Biometrika**, Oxford University Press, v. 75, n. 2, p. 383–386, 1988. Citado na página 57.

HUBERT, M.; ROUSSEEUW, P. J.; SEGAERT, P. Multivariate functional outlier detection. **Statistical Methods & Applications**, Springer, v. 24, n. 2, p. 177–202, 2015. Citado na página 25.

HUBERT, M.; VANDERVIJEREN, E. An adjusted boxplot for skewed distributions. **Computational statistics & data analysis**, Elsevier, v. 52, n. 12, p. 5186–5201, 2008. Citado na página 25.

\_\_\_\_\_. An adjusted boxplot for skewed distributions. **Computational statistics & data analysis**, Elsevier, v. 52, n. 12, p. 5186–5201, 2008. Citado na página 87.

ISO 17043, E. Iec 17043. **Conformity assessment-General requirements for proficiency testing**, v. 1, 2010. Citado na página 39.

ISO GUM. **Guide to the Expression of Uncertainty in Measurement, (GUM), BIPM, IEC, IFCC, IUPAC, IUPAP, OIML**. 1995. Citado nas páginas 36 e 39.

JAECH, J. L. Generalized models for making inferences on measurement error parameters. In: COMMISSION OF THE EUROPEAN COMMUNITIES, JOINT RESEARCH CENTRE. **7th Annual Symposium on Safeguards and Nuclear Material Management: Congress Hall, Liège, Belgium, 21st/23rd May 1985**. [S.l.], 1985. v. 19, p. 321. Citado na página 39.

JOHANSEN, S.; NIELSEN, B. Discussion: The forward search: Theory and data analysis. **Journal of the Korean Statistical Society**, Elsevier, v. 39, n. 2, p. 137–145, 2010. Citado na página 27.

\_\_\_\_\_. Asymptotic theory of outlier detection algorithms for linear time series regression models. **Scandinavian Journal of Statistics**, Wiley Online Library, v. 43, n. 2, p. 321–348, 2016. Citado nas páginas 27 e 29.

- JOHANSEN, S.; NIELSEN, B. *et al.* Analysis of the forward search using some new results for martingales and empirical processes. **Bernoulli**, Bernoulli Society for Mathematical Statistics and Probability, v. 22, n. 2, p. 1131–1183, 2016. Citado na página 27.
- JUNG, K.-M. Local influence in generalized estimating equations. **Scandinavian Journal of Statistics**, Wiley Online Library, v. 35, n. 2, p. 286–294, 2008. Citado na página 30.
- KEIDING, N. Estimation in the birth process. **Biometrika**, Oxford University Press, v. 61, n. 1, p. 71–80, 1974. Citado na página 32.
- \_\_\_\_\_. Maximum likelihood estimation in the birth-and-death process. **The Annals of Statistics**, JSTOR, p. 363–372, 1975. Citado na página 32.
- KIM, M. G. Local influence in multivariate regression. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 24, n. 5, p. 1271–1278, 1995. Citado na página 30.
- LACHOS, V.; ANGOLINI, T.; ABANTO-VALLE, C. On estimation and local influence analysis for measurement errors models under heavy-tailed distributions. **Statistical Papers**, Springer, v. 52, n. 3, p. p. 567–590, 2011. Citado na página 31.
- LACHOS, V.; LABRA, F.; BOLFARINE, H.; GHOSH, P. Multivariate measurement error models based on scale mixtures of the skew-normal distribution. **Statistics**, Taylor & Francis, v. 44, n. 6, p. p. 541–556, 2010. Citado na página 31.
- LACHOS, V.; MONTENEGRO, L.; BOLFARINE, H. Inference and local influence assessment in skew-normal null intercept measurement error model. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 78, n. 3, p. p. 395–419, 2008. Citado na página 31.
- LAURENT, R. T. S.; COOK, R. D. Leverage, local influence and curvature in nonlinear regression. **Biometrika**, Biometrika Trust, v. 80, n. 1, p. 99–106, 1993. Citado na página 30.
- LAWRANCE, A. J. Regression transformation diagnostics using local influence. **Journal of the American Statistical Association**, Taylor & Francis, v. 83, n. 404, p. 1067–1072, 1988. Citado nas páginas 29 e 30.
- LEE, A. H.; ZHAO, Y. Assessing local influence in measurement error models. **Biometrical journal**, Wiley Online Library, v. 38, n. 7, p. p. 829–841, 1996. Citado na página 67.
- LEE, S.-Y.; LU, B.; SONG, X.-Y. Assessing local influence for nonlinear structural equation models with ignorable missing data. **Computational statistics & data analysis**, Elsevier, v. 50, n. 5, p. p. 1356–1377, 2006. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167947304003883>>. Citado na página 67.
- LEE, S.-Y.; TANG, N.-S. Local influence analysis of nonlinear structural equation models. **Psychometrika**, Springer, v. 69, n. 4, p. p. 573–592, 2004. Disponível em: <<http://link.springer.com/article/10.1007/BF02289856>>. Citado na página 67.
- LEROY, A. M.; ROUSSEEUW, P. J. **Robust regression and outlier detection**. [S.l.: s.n.], 1987. v. 1. Citado nas páginas 26 e 64.
- LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. **Biometrics**, JSTOR, v. 54, n. 2, p. p. 570–582, 1998. Citado na página 30.

- LI, J.; VALLIANT, R. Detecting groups of influential observations in linear regression using survey data-adapting the forward search method. **Pakistan Journal of Statistics**, v. 27, n. 4, 2011. Citado na página 28.
- LIU, S. On local influence for elliptical linear models. **Statistical Papers**, Springer, v. 41, n. 2, p. 211–224, 2000. Citado na página 30.
- LU, J.; SHI, L.; CHEN, F. Outlier detection in time series models using local influence method. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 41, n. 12, p. 2202–2220, 2012. Citado na página 30.
- MA, Y.; LI, R. Variable selection in measurement error models. **Bernoulli: official journal of the Bernoulli Society for Mathematical Statistics and Probability**, NIH Public Access, v. 16, n. 1, p. p. 274, 2010. Citado na página 31.
- MAVRIDIS, D.; MOUSTAKI, I. Detecting outliers in factor analysis using the forward search algorithm. **Multivariate behavioral research**, Taylor & Francis, v. 43, n. 3, p. 453–475, 2008. Citado na página 27.
- \_\_\_\_\_. The forward search algorithm for detecting aberrant response patterns in factor analysis for binary data. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 18, n. 4, p. 1016–1034, 2009. Citado na página 27.
- MONTGOMERY, D.; JENNINGS, C. L.; KULAHCI, M. **Introduction to time series analysis and forecasting**. [S.l.]: John Wiley and Sons, 2011. v. 526. Citado na página 63.
- MORAN, P. Estimating structural and functional relationships. **Journal of Multivariate Analysis**, Elsevier, v. 1, n. 2, p. p. 232–255, 1971. Citado na página 23.
- OSORIO, F.; PAULA, G. A.; GALEA, M. Assessment of local influence in elliptical linear models with longitudinal structure. **Computational Statistics & Data Analysis**, Elsevier, v. 51, n. 9, p. 4354–4368, 2007. Citado na página 30.
- PAN, J.-X.; FANG, K.-T.; ROSEN, D. von. Local influence assessment in the growth curve model with unstructured covariance. **Journal of Statistical Planning and Inference**, Elsevier, v. 62, n. 2, p. 263–278, 1997. Citado na página 30.
- PAULA, G. A. Assessing local influence in restricted regression models. **Computational statistics & data analysis**, Elsevier, v. 16, n. 1, p. p. 63–79, 1993. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0167947393902450>>. Citado na página 30.
- PESENTIA, M.; PIRASA, M. **A MODIFIED FORWARD SEARCH APPROACH APPLIED TO TIME SERIES ANALYSIS**. 2008. Citado na página 27.
- PETTITT, A.; DAUD, I. B. Case-weighted measures of influence for proportional hazards regression. **Applied Statistics**, JSTOR, p. 51–67, 1989. Citado na página 29.
- PINTO, D. L.; AOKI, R.; SILVA, G. Statistical analysis of proficiency testing results under elliptical distributions. **Computational Statistics & Data Analysis**, v. 53, n. 4, p. 1427 – 1439, 2009. ISSN 0167-9473. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167947308005720>>. Citado na página 39.

- POON, W.-Y.; LEW, S.-F.; POON, Y. S. A local influence approach to identifying multiple multivariate outliers. **British Journal of Mathematical and Statistical Psychology**, Wiley Online Library, v. 53, n. 2, p. p. 255–273, 2000. Citado na página 60.
- POON, W.-Y.; POON, Y. S. Conformal normal curvature and assessment of local influence. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 61, n. 1, p. p. 51–61, 1999. Citado nas páginas 19, 30, 31, 32, 59, 60, 61, 62, 63, 71, 72 e 113.
- \_\_\_\_\_. Influential observations in the estimation of mean vector and covariance matrix. **British Journal of Mathematical and Statistical Psychology**, Wiley Online Library, v. 55, n. 1, p. 177–192, 2002. Citado na página 60.
- \_\_\_\_\_. Total behavior of likelihood displacement. **Statistica Sinica**, JSTOR, p. 599–607, 2002. Citado na página 30.
- RASEKH, A. Local influence in measurement error models with ridge estimate. **Computational statistics & data analysis**, Elsevier, v. 50, n. 10, p. p. 2822–2834, 2006. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167947305001015>>. Citado na página 67.
- RIANI, M.; ATKINSON, A. C.; FANTI, G.; CROSILLA, F. Regression analysis with partially labelled regressors: carbon dating of the shroud of turin. **Statistics and Computing**, Springer, v. 23, n. 4, p. 551–561, 2013. Citado na página 28.
- RIANI, M.; ATKINSON, A. C.; PERROTTA, D. *et al.* A parametric framework for the comparison of methods of very robust regression. **Statistical Science**, Institute of Mathematical Statistics, v. 29, n. 1, p. 128–143, 2014. Citado na página 28.
- RIANI, M.; CERIOLI, A.; ATKINSON, A. C.; PERROTTA, D. *et al.* Monitoring robust regression. **Electronic Journal of Statistics**, The Institute of Mathematical Statistics and the Bernoulli Society, v. 8, n. 1, p. 646–677, 2014. Citado na página 28.
- RIANI, M.; PERROTTA, D.; CERIOLI, A. The forward search for very large datasets. **Journal of Statistical Software**, v. 67, n. 1, p. 1–20, 2015. Citado na página 27.
- RIANI, M.; PERROTTA, D.; TORTI, F. FSDA: A matlab toolbox for robust analysis and interactive data exploration. **Chemometrics and Intelligent Laboratory Systems**, Elsevier, v. 116, p. 17–32, 2012. Citado na página 28.
- ROUSSEEUW, P.; LEROY, A. Robust regression and outlier detection john wiley & sons. **Inc., New York**, 1987. Citado na página 65.
- ROUSSEEUW, P. J. Least median of squares regression. **Journal of the American statistical association**, Taylor & Francis, v. 79, n. 388, p. p. 871–880, 1984. Citado na página 64.
- ROUSSEEUW, P. J.; ZOMEREN, B. C. V. Unmasking multivariate outliers and leverage points. **Journal of the American Statistical Association**, Taylor & Francis, v. 85, n. 411, p. p. 633–639, 1990. Citado na página 26.
- RUSSO, C. M. **Análise de um modelo de regressão com erros nas variáveis multivariado com intercepto nulo**. Dissertação (Mestrado) — Universidade de São Paulo, 2006. Citado nas páginas 31, 34, 35, 71, 176 e 180.

- RUSSO, C. M.; AOKI, R.; JR, D. L.-P. Hypotheses testing on a multivariate null intercept errors-in-variables model. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 38, n. 7, p. p. 1447–1469, 2009. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/03610910902972310>>. Citado nas páginas 105 e 122.
- SALINI, S.; CERIOLI, A.; LAURINI, F.; RIANI, M. Reliable robust regression diagnostics. **International Statistical Review**, Wiley Online Library, v. 84, n. 1, p. 99–127, 2016. Citado na página 27.
- SHI, L. Local influence in principal components analysis. **Biometrika**, Biometrika Trust, v. 84, n. 1, p. 175–186, 1997. Citado na página 30.
- SHI, L.; CHEN, G. Local influence in multilevel models. **Canadian Journal of Statistics**, Wiley Online Library, v. 36, n. 2, p. 259–275, 2008. Citado na página 30.
- SHI, L.; HUANG, M. Stepwise local influence analysis. **Computational Statistics & Data Analysis**, Elsevier, v. 55, n. 2, p. 973–982, 2011. Citado na página 30.
- SHI, L.; RAHMAN, M. M.; GAN, W.; ZHAO, J. Stepwise local influence in generalized autoregressive conditional heteroskedasticity models. **Journal of Applied Statistics**, Taylor & Francis, v. 42, n. 2, p. 428–444, 2015. Citado na página 27.
- SHI, X.; ZHU, H.; IBRAHIM, J. G. Local influence for generalized linear models with missing covariates. **Biometrics**, Wiley Online Library, v. 65, n. 4, p. 1164–1174, 2009. Citado na página 30.
- SINGER, J.; NOBRE, J.; ROCHA, F. **Análise de dados longitudinais**. [S.l.: s.n.], 2012. Citado na página 25.
- SINGER, J. M.; ANDRADE, D. F. Regression models for the analysis of pretest/posttest data. **Biometrics**, JSTOR, p. 729–735, 1997. Citado na página 111.
- SPRENT, P. Some history of functional and structural relationships. **Contemp. Math**, v. 112, p. p. 3–15, 1990. Citado na página 23.
- SWEETING, T. J. Uniform asymptotic normality of the maximum likelihood estimator. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 8, n. 6, p. 1375–1381, 1980. ISSN 00905364. Disponível em: <<http://www.jstor.org/stable/2240949>>. Citado nas páginas 31, 40, 42 e 43.
- TALARICO, A. M. **Modelo de calibração ultraestrutural**. Dissertação (Mestrado) — Universidade de São Paulo, 2014. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/55/55134/tde-26032014-161411/en.php>>. Citado nas páginas 31, 32, 35, 37, 40, 174, 175 e 176.
- THOMAS, W.; COOK, R. D. Assessing influence on predictions from generalized linear models. **Technometrics**, Taylor & Francis, v. 32, n. 1, p. 59–65, 1990. Citado na página 29.
- TOMAN, A. Robust confirmatory factor analysis based on the forward search algorithm. **Statistical papers**, Springer, v. 55, n. 1, p. 233–252, 2014. Citado na página 28.
- TOMAN, B. Bayesian approaches to calculating a reference value in key comparison experiments. **Technometrics**, Taylor & Francis, v. 49, n. 1, p. 81–87, 2007. Citado na página 39.

- TORTI, F.; PERROTTA, D.; ATKINSON, A. C.; RIANI, M. Benchmark testing of algorithms for very robust regression: Fs, lms and lts. **Computational statistics & data analysis**, Elsevier, v. 56, n. 8, p. 2501–2512, 2012. Citado na página 28.
- TSAI, C.-L.; WU, X. Assessing local influence in linear regression models with first-order autoregressive or heteroscedastic error structure. **Statistics & probability letters**, Elsevier, v. 14, n. 3, p. 247–252, 1992. Citado na página 29.
- VERBEKE, G.; MOLENBERGHS, G. **Linear mixed models for longitudinal data**. [S.l.]: Springer, 2000. Citado na página 60.
- WEISBERG, S. **Applied linear regression**. [S.l.]: John Wiley and Sons, 2005. v. 528. Citado na página 25.
- WEISS, L. Asymptotic properties of maximum likelihood estimators in some nonstandard cases, ii. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 68, n. 342, p. 428–430, 1973. Citado na página 40.
- Wolfram Research, Inc. **Mathematica**. Champaign, Illinois: [s.n.], 2012. Disponível em: <<https://www.wolfram.com>>. Citado na página 75.
- WU, X.; LUO, Z. Second-order approach to local influence. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 929–936, 1993. Citado na página 30.
- ZHAO, Y.; LEE, A. H. Assessment of influence in non-linear measurement error models. **Journal of Applied Statistics**, Taylor & Francis, v. 22, n. 2, p. p. 215–225, 1995. Citado na página 31.
- ZHU, F.; LIU, S.; SHI, L. Local influence analysis for poisson autoregression with an application to stock transaction data. **Statistica Neerlandica**, Wiley Online Library, v. 70, n. 1, p. 4–25, 2016. Citado na página 31.
- ZHU, H.; IBRAHIM, J. G.; LEE, S.; ZHANG, H. *et al.* Perturbation selection and influence measures in local influence analysis. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 35, n. 6, p. 2565–2588, 2007. Citado na página 30.
- ZHU, H.; IBRAHIM, J. G.; TANG, N. Bayesian influence analysis: a geometric approach. **Biometrika**, NIH Public Access, v. 98, n. 2, p. 307, 2011. Citado na página 30.
- \_\_\_\_\_. Bayesian sensitivity analysis of statistical models with missing data. **Statistica Sinica**, NIH Public Access, v. 24, n. 2, p. 871, 2014. Citado na página 31.
- ZHU, H.-T.; LEE, S.-Y. Local influence for incomplete data models. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 63, n. 1, p. p. 111–126, 2001. Citado nas páginas 31, 62, 72 e 113.
- ZIJLSTRA, W. P.; ARK, L. A. van der; SIJTSMA, K. Robust mokken scale analysis by means of the forward search algorithm for outlier detection. **Multivariate behavioral research**, Taylor & Francis, v. 46, n. 1, p. 58–89, 2011. Citado na página 28.

---

## CONJUNTO DE DADOS

---

---

### A.1 Dados de Escovas de dentes

Tabela 20 – Dados da Escova de dente convencional.

Número	$x_{c_d}$	$y_{c_d}$
1	1.20	0.75
2	1.43	0.55
3	0.68	0.08
4	1.45	0.75
5	0.50	0.05
6	2.75	1.60
7	1.25	0.65
8	0.40	0.13
9	1.18	0.83
10	1.43	0.58
11	0.45	0.38
12	1.60	0.63
13	0.25	0.25
14	2.98	1.03
15	3.35	1.58
16	1.50	0.20
17	4.08	1.88
18	3.15	2.00
19	0.90	0.25
20	1.78	0.18
21	3.50	0.85

Continua na próxima página.

---

Tabela 20: Dados Escova de dentes convencional (continuação).

Número	$x_{c_a}$	$y_{c_d}$
22	2.50	1.15
23	2.18	0.93
24	2.68	1.05
25	2.73	0.85
26	3.43	0.88

$x_{c_a}$ : índice de placa bacteriana no início do estudo (antes do uso da escova de dente convencional),

$y_{c_d}$ : índice de placa bacteriana após o uso da escova de dente convencional.

Tabela 21 – Dados da Escova de dente experimental.

Número	$x_{e_a}$	$y_{e_d}$
1	2.18	0.43
2	2.05	0.08
3	1.05	0.18
4	1.95	0.78
5	0.28	0.03
6	2.63	0.23
7	1.50	0.20
8	0.45	0.00
9	0.70	0.05
10	1.30	0.30
11	1.25	0.33
12	0.18	0.00
13	3.30	0.90
14	1.40	0.24
15	0.90	0.15
16	0.58	0.10
17	2.50	0.33
18	2.25	0.33
19	1.53	0.53
20	1.43	0.43
21	3.48	0.65
22	1.80	0.20
23	1.50	0.25
24	2.55	0.15

Continua na próxima página.

Tabela 21: Dados Escova de dentes experimental (continuação).

Número	$x_{e_a}$	$y_{e_d}$
25	1.30	0.05
26	2.65	0.25

$x_{e_a}$ : índice de placa bacteriana no início do estudo (antes do uso da escova de dente experimental),

$y_{e_d}$ : índice de placa bacteriana após o uso da escova de dente experimental.

## A.2 Dados de Líquidos de bochecho

Tabela 22 – Líquido de bochecho controle.

Número	$x_c$	$y_{c_3}$	$y_{c_6}$
1	2.46	1.73	1.58
2	2.89	2.61	1.91
3	3.00	2.75	1.89
4	2.52	0.98	1.71
5	2.52	1.61	2.29
6	2.78	1.25	1.34
7	2.08	0.88	0.35
8	3.11	3.04	1.93
9	2.38	0.90	1.28
10	2.70	2.11	2.46
11	3.20	2.39	2.48
12	3.25	2.23	2.23
13	3.21	2.88	1.96
14	3.30	2.50	3.10
15	2.27	1.61	1.61
16	2.21	1.15	1.67
17	2.40	1.69	2.08
18	2.64	1.45	1.96
19	2.07	0.70	0.48
20	2.43	0.85	0.78
21	2.27	1.71	2.27
22	2.52	2.50	2.85
23	2.79	1.54	1.42
24	2.54	1.58	1.35

Continua na próxima página.

Tabela 22: Líquido de bochecho controle (continuação).

Número	$x_c$	$y_{c_3}$	$y_{c_6}$
25	2.17	0.77	1.48
26	2.09	1.16	0.98
27	2.68	2.74	1.88
28	2.35	0.72	1.48
29	2.36	2.33	2.04
30	2.61	2.43	2.16
31	2.48	2.00	1.93
32	2.31	1.33	1.38
33	2.50	2.32	1.96
34	2.31	2.27	1.96
35	2.38	1.98	1.38
36	2.27	0.75	0.96

$x_c$ : índice de placa bacteriana no início do estudo (antes do uso do líquido de bochecho controle),

$y_{c_3}$ : índice de placa bacteriana após 3 meses com o uso do líquido de bochecho controle,

$y_{c_6}$ : índice de placa bacteriana após 6 meses com o uso do líquido de bochecho controle.

Tabela 23 – Líquido de bochecho A.

Número	$x_a$	$y_{a_3}$	$y_{a_6}$
1	2.65	1.42	1.25
2	2.13	2.48	1.63
3	2.70	0.00	0.59
4	2.48	0.10	0.12
5	2.98	1.57	2.31
6	3.57	2.80	2.41
7	2.48	0.05	0.27
8	2.73	1.44	1.17
9	2.35	1.35	0.65
10	2.52	1.59	0.75
11	2.61	1.98	2.32
12	2.43	1.63	2.43
13	2.75	1.60	2.31
14	2.27	0.63	0.96
15	2.68	2.24	2.52
16	3.40	0.85	2.53

Continua na próxima página.

Tabela 23: Líquido de bochecho A (continuação).

Número	$x_a$	$y_{a_3}$	$y_{a_6}$
17	2.96	1.48	0.29
18	2.16	1.46	1.20
19	2.34	0.70	0.59
20	2.04	0.46	0.55
21	2.35	1.78	1.43
22	2.68	1.43	1.93
23	2.50	1.35	1.04
24	2.07	1.29	0.80
25	2.52	0.79	1.00
26	2.37	1.50	1.15
27	2.27	0.39	0.64
28	2.21	1.27	0.63
29	2.71	1.80	1.71
30	2.45	1.10	0.65
31	3.11	2.81	1.83
32	3.02	1.94	1.33
33	2.32	0.73	0.57

$x_a$ : índice de placa bacteriana no início do estudo (antes do uso do líquido de bochecho A),

$y_{a_3}$ : índice de placa bacteriana após 3 meses com o uso do líquido de bochecho A,

$y_{a_6}$ : índice de placa bacteriana após 6 meses com o uso do líquido de bochecho A.

Tabela 24 – Líquido de bochecho B.

Número	$x_b$	$y_{b_3}$	$y_{b_6}$
1	2.33	0.56	0.90
2	2.53	1.00	0.89
3	2.61	2.11	1.61
4	2.74	1.02	0.48
5	3.02	2.04	1.06
6	2.56	2.04	0.29
7	2.25	0.96	0.58
8	2.74	1.54	1.65
9	2.60	0.88	1.17
10	2.76	1.82	1.14
11	3.11	0.93	0.59

Continua na próxima página.

Tabela 24: Líquido de bochecho B (continuação).

Número	$x_b$	$y_{b_3}$	$y_{b_6}$
12	3.13	1.85	1.25
13	2.71	1.36	1.09
14	2.42	1.36	1.62
15	2.32	0.92	1.36
16	2.19	1.85	0.54
17	2.72	0.50	0.98
18	2.45	1.20	1.55
19	2.60	2.06	1.52
20	2.17	1.60	0.35
21	2.38	1.78	1.48
22	2.18	0.55	0.39
23	2.38	0.33	1.73
24	2.88	1.53	0.55
25	2.21	1.90	0.60
26	2.42	0.67	0.88
27	2.21	1.17	1.60
28	2.22	1.46	0.83
29	2.86	2.20	1.64
30	2.06	1.13	1.48
31	2.23	0.61	0.68
32	2.02	1.00	1.11
33	2.45	0.98	1.14
34	2.29	1.13	0.89
35	2.06	0.27	0.23
36	2.44	0.88	1.31

$x_b$ : índice de placa bacteriana no início do estudo (antes do uso do líquido de bochecho B),

$y_{b_3}$ : índice de placa bacteriana após 3 meses com o uso do líquido de bochecho B,

$y_{b_6}$ : índice de placa bacteriana após 6 meses com o uso do líquido de bochecho B.

### A.3 Dados da Potência do Motor

Nesta seção, apresentamos o conjunto de dados usado para ilustrar a metodologia desenvolvida que consiste nas medições da potência do motor em 9 pontos de rotação para 8 laboratórios. Os laboratórios participantes foram Mauá , Bosch , Ford , KSPG , Mahle , Delphi , GM , Marelli . Não vamos identificar o número de laboratórios no conjunto de dados com os nomes dos laboratórios fornecidos acima, pois é confidencial.

Tabela 25 – Dados da Potência do motor.

	$Y_{1j1}$	$Y_{1j2}$	$Y_{1j3}$	$Y_{1j4}$	$Y_{1j5}$	$Y_{2j1}$	$Y_{2j2}$	$Y_{2j3}$	$Y_{2j4}$	$Y_{2j5}$	$Y_{2j6}$	$Y_{2j7}$	$Y_{2j8}$
1200	8.89	8.83	8.86	8.85	8.88	8.57	8.56	8.58	8.64	8.64	8.60	8.59	8.56
2000	15.84	15.80	15.80	15.79	15.81	15.28	15.27	15.35	15.36	15.38	15.29	15.24	15.27
3000	26.84	26.61	26.86	26.85	26.92	26.12	26.12	26.09	26.12	26.22	26.12	25.95	26.14
3600	31.41	31.31	31.40	31.40	31.50	30.48	30.45	30.41	30.52	30.48	30.41	30.25	30.32
4400	37.19	37.12	37.24	37.17	37.32	36.46	36.33	36.25	36.36	36.45	36.41	36.21	36.20
5200	44.35	44.35	44.28	44.31	44.40	43.29	43.22	43.11	43.21	43.34	43.28	43.05	43.05
5600	47.49	47.33	47.56	47.60	47.79	46.29	46.21	46.11	46.17	46.18	45.94	46.01	45.97
6000	49.92	49.76	49.98	49.90	49.96	48.01	47.92	47.90	48.00	47.88	47.80	47.68	47.71
6400	50.92	50.74	50.89	50.84	50.94	48.85	48.84	48.75	48.86	48.85	48.66	48.57	48.45
	$Y_{2j9}$	$Y_{2j10}$	$Y_{2j11}$	$Y_{2j12}$	$Y_{2j13}$	$Y_{2j14}$	$Y_{2j15}$	$Y_{2j16}$	$Y_{2j17}$	$Y_{2j18}$	$Y_{2j19}$	$Y_{2j20}$	$Y_{2j21}$
1200	8.57	8.64	8.59	8.60	8.67	8.72	8.79	8.79	8.79	8.73	8.60	8.60	8.46
2000	15.31	15.34	15.32	15.37	15.35	15.57	15.60	15.60	15.54	15.53	15.26	15.26	15.10
3000	26.06	25.93	26.03	26.08	26.13	26.34	26.27	26.25	26.38	26.49	26.05	26.19	25.76
3600	30.42	30.52	30.45	30.44	30.65	30.91	30.93	30.97	30.94	30.86	30.26	30.57	29.98
4400	36.16	36.33	36.17	36.17	36.34	36.58	36.44	36.52	36.82	36.66	36.23	36.32	35.80
5200	42.98	43.11	43.06	43.06	43.21	43.45	43.37	43.50	43.65	43.65	43.23	43.20	42.54
5600	46.08	46.03	46.19	46.29	46.58	46.64	46.69	46.71	46.73	46.03	46.03	45.34	45.36
6000	47.64	47.90	47.79	47.85	47.99	48.13	48.35	48.41	48.51	48.41	47.60	47.67	46.93
6400	48.58	48.66	48.86	48.81	49.00	49.09	49.29	49.33	49.42	49.33	48.44	48.38	47.79
	$Y_{2j22}$	$Y_{2j23}$	$Y_{3j1}$	$Y_{3j2}$	$Y_{3j3}$	$Y_{3j4}$	$Y_{3j5}$	$Y_{3j6}$	$Y_{3j7}$	$Y_{3j8}$	$Y_{3j9}$	$Y_{3j10}$	$Y_{3j11}$
1200	8.47	8.52	8.85	8.76	8.77	8.80	8.77	8.77	8.78	8.78	8.76	8.77	8.77
2000	15.05	15.18	15.85	15.78	15.81	15.71	15.76	15.75	15.76	15.75	15.73	15.76	15.73
3000	25.61	25.78	26.68	26.68	26.68	26.52	26.59	26.61	26.52	26.50	26.61	26.56	26.73
3600	29.90	30.13	31.33	31.22	31.23	31.21	31.27	31.25	31.28	31.26	31.31	31.26	31.25
4400	35.67	35.73	37.02	36.88	36.93	36.89	36.89	36.88	36.92	36.94	36.93	36.91	36.91
5200	42.56	42.56	43.90	43.81	43.88	43.75	43.79	43.74	43.80	43.86	43.82	43.84	43.82
5600	45.55	45.70	47.00	46.99	47.01	46.93	46.99	46.94	46.99	47.06	47.05	47.03	47.02
6000	46.84	47.20	48.78	48.81	48.87	48.86	48.83	48.79	48.81	48.90	48.86	48.84	48.82
6400	47.71	48.01	49.86	49.87	49.91	49.88	49.89	49.85	49.90	49.92	49.90	49.88	49.85
	$Y_{3j12}$	$Y_{3j13}$	$Y_{3j14}$	$Y_{3j15}$	$Y_{3j16}$	$Y_{3j17}$	$Y_{3j18}$	$Y_{4j1}$	$Y_{4j2}$	$Y_{4j3}$	$Y_{4j4}$	$Y_{4j5}$	$Y_{4j6}$
1200	8.84	8.79	8.79	8.80	8.81	8.89	8.81	8.73	8.84	8.85	8.77	8.78	8.88
2000	15.81	15.78	15.81	15.82	15.74	15.82	15.82	15.68	15.88	15.89	15.82	16.00	16.00
3000	26.81	26.72	26.76	26.73	26.77	26.82	26.82	26.60	26.91	26.82	26.84	27.04	27.06
3600	31.29	31.28	31.33	31.15	31.26	31.26	31.28	31.65	31.49	31.49	31.31	31.51	31.60
4400	36.90	36.86	36.87	36.88	36.95	36.98	36.96	37.51	37.33	37.32	37.13	37.31	37.35
5200	43.83	43.80	43.78	43.78	43.88	43.87	43.80	44.37	44.29	44.29	44.05	44.23	44.25
5600	47.06	46.98	46.96	46.99	47.06	47.09	47.00	47.44	47.43	47.47	47.25	47.30	47.32
6000	48.88	48.78	48.76	48.78	48.86	48.88	48.74	49.30	49.31	49.36	49.13	49.19	49.14
6400	49.84	49.76	49.73	49.75	49.83	49.85	49.73	50.20	50.24	50.23	50.05	50.09	50.08
	$Y_{4j7}$	$Y_{4j8}$	$Y_{4j9}$	$Y_{5j1}$	$Y_{5j2}$	$Y_{5j3}$	$Y_{5j4}$	$Y_{5j5}$	$Y_{5j6}$	$Y_{5j7}$	$Y_{5j8}$	$Y_{5j9}$	$Y_{5j10}$
1200	8.90	8.92	8.93	8.70	8.68	9.13	9.07	9.04	9.03	8.82	8.75	8.92	8.95
2000	15.94	16.00	16.05	15.46	15.50	16.24	16.35	16.14	16.22	15.58	15.58	16.05	15.96
3000	26.85	26.93	26.96	25.94	25.91	27.24	27.41	27.38	27.41	26.16	25.96	27.18	27.05

Continua na próxima página.

Tabela 25: Dados da Potência do motor (continuação).

3600	31.56	31.56	31.60	30.71	30.85	32.15	32.27	32.26	32.26	30.99	30.83	31.95	31.87
4400	37.32	37.36	37.45	35.93	35.92	38.02	38.10	37.99	38.16	36.54	36.26	37.59	37.57
5200	44.34	44.40	44.39	42.47	42.56	45.16	45.27	44.98	45.28	43.08	42.92	44.67	44.70
5600	47.54	47.52	47.57	45.65	45.70	48.38	48.40	48.13	48.40	46.12	45.94	47.82	47.96
6000	49.41	49.44	49.43	47.30	47.42	49.95	50.08	49.89	50.36	47.71	47.63	49.51	49.40
6400	50.33	50.35	50.34	48.11	48.19	50.76	50.24	50.49	50.51	48.51	48.33	49.99	50.38
	$Y_{5j11}$	$Y_{5j12}$	$Y_{6j1}$	$Y_{6j2}$	$Y_{6j3}$	$Y_{6j4}$	$Y_{6j5}$	$Y_{6j6}$	$Y_{6j7}$	$Y_{6j8}$	$Y_{6j9}$	$Y_{6j10}$	$Y_{6j11}$
1200	8.85	8.87	9.00	8.90	8.90	9.00	9.00	8.90	8.90	8.90	9.00	8.90	8.90
2000	15.94	15.95	16.00	15.90	16.10	16.10	16.10	16.00	16.00	15.90	16.10	16.00	16.00
3000	26.81	26.84	27.10	27.00	27.10	27.10	27.10	27.10	27.00	27.00	27.30	27.10	27.10
3600	31.73	31.67	31.70	31.60	31.70	31.70	31.80	31.80	31.50	31.60	31.80	31.70	31.70
4400	37.41	37.58	37.60	37.50	37.70	37.70	37.70	37.60	37.50	37.60	37.90	37.70	37.60
5200	44.18	44.15	44.80	44.60	44.90	44.70	44.80	44.80	44.60	44.60	44.90	44.90	44.70
5600	47.39	47.41	47.60	47.40	47.70	47.70	47.60	47.80	47.30	47.30	47.80	47.90	47.50
6000	49.23	49.33	49.10	49.10	49.30	49.40	49.30	49.30	49.00	49.00	49.40	49.50	49.20
6400	49.68	49.72	49.80	49.90	50.00	50.10	50.20	50.10	49.70	49.70	50.20	50.20	49.90
	$Y_{6j12}$	$Y_{6j13}$	$Y_{6j14}$	$Y_{6j15}$	$Y_{6j16}$	$Y_{7j1}$	$Y_{7j2}$	$Y_{7j3}$	$Y_{7j4}$	$Y_{7j5}$	$Y_{7j6}$	$Y_{7j7}$	$Y_{7j8}$
1200	8.90	8.90	8.90	8.90	8.90	8.70	8.70	8.50	8.50	8.50	8.60	8.60	8.60
2000	16.00	16.10	16.00	16.00	16.00	15.60	15.60	15.40	15.40	15.40	15.40	15.50	15.40
3000	27.10	27.20	27.10	27.00	27.10	26.10	25.80	25.60	25.60	25.60	25.70	26.00	25.80
3600	31.70	31.80	31.70	31.60	31.70	30.70	30.60	30.20	30.10	30.40	30.60	30.50	30.50
4400	37.60	37.70	37.60	37.70	37.70	36.50	36.60	36.30	36.30	36.40	36.50	36.60	36.50
5200	44.70	44.90	44.70	44.80	44.70	43.80	43.60	43.40	43.10	42.80	43.30	43.50	43.50
5600	47.60	47.80	47.80	47.50	47.50	46.20	46.10	46.00	45.70	46.40	46.60	46.50	46.60
6000	49.30	49.50	49.50	49.20	49.10	48.70	48.30	47.70	48.10	47.40	48.30	48.30	48.50
6400	50.10	50.20	50.20	50.00	49.90	49.70	49.60	48.60	49.00	49.10	49.10	49.40	49.40
	$Y_{7j9}$	$Y_{7j10}$	$Y_{7j11}$	$Y_{7j12}$	$Y_{7j13}$	$Y_{7j14}$	$Y_{7j15}$	$Y_{7j16}$	$Y_{7j17}$	$Y_{7j18}$	$Y_{7j19}$	$Y_{7j20}$	$Y_{7j21}$
1200	8.60	8.60	8.60	8.60	8.60	8.60	8.70	8.60	8.60	8.60	8.50	8.60	8.60
2000	15.40	15.50	15.40	15.50	15.40	15.50	15.60	15.50	15.40	15.50	15.50	15.40	15.40
3000	25.80	25.90	26.00	25.70	25.70	25.90	25.90	25.90	26.00	25.80	25.90	25.80	25.80
3600	30.60	30.50	30.40	30.40	30.40	30.60	30.60	30.50	30.50	30.60	30.60	30.60	30.60
4400	36.60	36.50	36.60	36.60	36.70	36.70	36.60	36.70	36.70	36.70	36.60	36.60	36.70
5200	43.60	44.00	44.20	43.50	43.50	43.70	43.70	43.70	43.80	43.70	43.70	43.70	43.60
5600	47.00	47.50	46.70	46.70	46.60	46.80	46.50	47.00	46.50	46.60	46.10	46.80	46.70
6000	48.30	49.20	49.10	48.60	48.40	48.50	48.60	48.50	48.70	48.50	48.70	48.40	48.50
6400	49.30	50.00	50.20	49.30	49.20	49.40	49.10	49.10	49.30	49.50	49.40	49.40	49.10
	$Y_{7j22}$	$Y_{7j23}$	$Y_{7j24}$	$Y_{7j25}$	$Y_{7j26}$	$Y_{8j1}$	$Y_{8j2}$	$Y_{8j3}$	$Y_{8j4}$	$Y_{8j5}$	$Y_{8j6}$	$Y_{8j7}$	$Y_{8j8}$
1200	8.70	8.60	8.60	8.60	8.60	8.70	8.70	8.60	8.60	8.80	8.70	8.70	8.60
2000	15.50	15.50	15.50	15.50	15.40	16.00	15.80	15.80	15.80	16.00	15.80	15.80	15.90
3000	25.90	25.80	25.50	25.60	25.60	27.10	27.00	26.70	26.70	27.10	27.10	26.90	26.90
3600	30.10	30.30	30.40	30.40	30.30	31.80	31.70	31.50	31.50	31.70	31.80	31.70	31.70
4400	36.70	36.50	36.20	36.30	36.30	37.50	37.40	36.90	37.00	37.40	37.40	36.80	37.00
5200	43.70	43.60	42.70	43.00	43.20	44.10	44.10	43.50	43.70	44.00	44.10	43.50	43.60
5600	46.60	45.80	46.30	46.50	46.20	47.00	47.10	46.60	46.80	46.80	47.00	46.60	46.80

Continua na próxima página.

Tabela 25: Dados da Potência do motor (continuação).

6000	48.50	48.40	47.90	47.70	48.30	48.80	48.70	48.20	48.50	48.60	48.80	48.40	48.70
6400	49.40	49.30	49.10	49.30	49.10	49.80	49.50	49.10	49.40	49.60	49.70	49.10	49.40
	$Y_{8j9}$	$Y_{8j10}$	$Y_{8j11}$	$Y_{8j12}$	$Y_{8j13}$	$Y_{8j14}$	$Y_{8j15}$	$Y_{8j16}$					
1200	8.80	8.70	8.80	8.70	8.70	8.70	8.80	8.70					
2000	16.00	15.90	16.00	16.00	15.90	15.90	16.00	15.80					
3000	27.20	27.30	27.00	26.90	26.60	26.60	26.50	26.30					
3600	31.90	31.90	31.90	31.80	31.70	31.60	31.50	31.50					
4400	37.50	37.50	37.30	37.20	37.30	37.10	36.90	36.90					
5200	44.30	44.40	43.80	43.90	44.00	43.80	43.50	43.50					
5600	47.10	47.20	47.20	47.30	47.10	47.10	46.80	46.90					
6000	49.00	49.10	49.10	49.20	49.00	49.00	48.60	48.60					
6400	49.70	49.70	49.60	50.00	49.60	49.70	49.30	49.40					

Tabela 26 – Variâncias do erro de medição ( $\sigma_{ij}^2$ ) para o  $i$ -ésimo laboratório no  $j$ -ésimo valor de rotação do motor.

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$	$j = 7$	$j = 8$	$j = 9$
$i = 1$	0.0068	0.0215	0.0618	0.0848	0.1190	0.1690	0.1944	0.2141	0.2225
$i = 2$	0.0054	0.0170	0.0491	0.0671	0.0949	0.1343	0.1535	0.1650	0.1711
$i = 3$	0.0005	0.0018	0.0050	0.0069	0.0097	0.0136	0.0157	0.0169	0.0176
$i = 4$	0.0081	0.0263	0.0750	0.1031	0.1446	0.2035	0.2333	0.2521	0.2615
$i = 5$	0.0498	0.1587	0.4509	0.6270	0.8680	1.2158	1.3936	1.4954	1.5341
$i = 6$	0.0101	0.0327	0.0935	0.1280	0.1806	0.2552	0.2888	0.3091	0.3186
$i = 7$	0.0114	0.0372	0.1029	0.1435	0.2061	0.2919	0.3307	0.3591	0.3760
$i = 8$	0.0249	0.0830	0.2371	0.3300	0.4543	0.6319	0.7243	0.7811	0.8060

Tabela 27 – Variância do verdadeiro valor da potência do motor ( $\sigma_{x_j}^2$ ) no  $j$ -ésimo ponto de rotação do motor.

$\sigma_{x_1}^2$	$\sigma_{x_2}^2$	$\sigma_{x_3}^2$	$\sigma_{x_4}^2$	$\sigma_{x_5}^2$	$\sigma_{x_6}^2$	$\sigma_{x_7}^2$	$\sigma_{x_8}^2$	$\sigma_{x_9}^2$
0.0077	0.0256	0.0740	0.0999	0.1414	0.2007	0.2266	0.2500	0.2581



## CONVERGÊNCIA DOS ELEMENTOS DA MATRIZ DE INFORMAÇÃO

A seguir iremos demonstrar a convergência da matriz de informação observada para a matriz  $\tilde{W}(\tilde{\theta})$ . Seja  $\boldsymbol{\psi}^n \in \Theta$  um vetor de parâmetros tais que  $\boldsymbol{\psi}^n = \boldsymbol{\theta} + \frac{1}{\sqrt{n}}\mathbf{s}$  para algum vetor  $\mathbf{s} \in \mathbb{R}^{m+2(p-1)}$  fixo e, seja  $\boldsymbol{\phi}^n$  um vetor aleatório que satisfaça

$$\boldsymbol{\phi}^n = (1 - \delta^n)\boldsymbol{\theta} + \delta^n\boldsymbol{\psi}^n = (1 - \delta^n)\boldsymbol{\theta} + \delta^n\left(\boldsymbol{\theta} + \frac{1}{\sqrt{n}}\mathbf{s}\right) = \boldsymbol{\theta} + \delta^n\frac{1}{\sqrt{n}}\mathbf{s},$$

em que  $0 < \delta^n < 1$  e  $\delta^n$  é uma variável aleatória. Seja  $c = \max\{|s_1|, \dots, |s_{m+2(p-1)}|\}$  a norma do vetor  $\mathbf{s}$ . Para simplificar a notação, sem perda de generalidade assumimos que  $\boldsymbol{\psi}^n = \boldsymbol{\theta} + \frac{c}{\sqrt{n}}\mathbf{1}_{m+2(p-1)}$  e  $\boldsymbol{\phi}^n = \boldsymbol{\theta} + \delta^n\frac{c}{\sqrt{n}}\mathbf{1}_{m+2(p-1)}$ . Dizemos que  $n \rightarrow \infty$  quando  $n_i \rightarrow \infty$  e  $\frac{n_i}{n} \rightarrow w_i$  em que  $w_i$  é uma constante positiva para todo  $i = 2, \dots, p$ . Enfatizamos que os resultados são válidos para todo  $\boldsymbol{\theta}^n$  tal que

$$\sqrt{n} \|\boldsymbol{\theta}^n - \boldsymbol{\theta}\| \leq c\mathbf{1}_{m+2(p-1)}, \quad n \geq 1.$$

Para todo  $\boldsymbol{\theta} = (\mu_{x_1}, \dots, \mu_{x_m}, \alpha_2, \dots, \alpha_p, \beta_2, \dots, \beta_p) \in \mathbb{R}^{m+2(p-1)}$ , os componentes da matriz de informação observada dependem dos seguintes elementos

$$a_j^n(\boldsymbol{\theta}) = 1 + \sigma_{x_j}^2 \left[ \frac{n_1}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{n_i \beta_i^2}{\sigma_{ij}^2} \right], \quad M_j^n(\boldsymbol{\theta}) = \frac{\mu_{x_j}}{\sigma_{x_j}^2} + \sum_{k=1}^{n_1} \frac{Y_{1jk}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i}{\sigma_{ij}^2} \left( \sum_{k=1}^{n_i} Y_{ijk} - n_i \alpha_i \right),$$

e

$$D_{1j}(\boldsymbol{\theta}) = \sum_{k=1}^{n_1} Y_{1jk} \quad \text{e} \quad D_{ij}(\boldsymbol{\theta}) = \sum_{k=1}^{n_i} Y_{ijk} - n_i \alpha_i,$$

para todo  $i = 2, \dots, p$  e  $j = 1, \dots, m$ . Neste apêndice, provaremos que

$$\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{1}{n} \mathbf{s}^T \mathbf{J}^n(\boldsymbol{\phi}^n) \mathbf{s} - \mathbf{s}^T \mathbf{W}(\boldsymbol{\theta}) \mathbf{s} \right| \geq \varepsilon \right] \rightarrow 0, \quad n \rightarrow \infty, \quad \varepsilon > 0.$$

Para provar a convergência uniforme em probabilidade, provaremos que cada componente da matriz de informação observada converge uniformemente em probabilidade. Temos que

$$\begin{aligned} \frac{a_j^n(\boldsymbol{\phi}^n)}{n} &= \frac{1}{n} + \sigma_{x_j}^2 \left[ \sum_{i=2}^p \frac{n_i}{n} \frac{(\beta_i + \delta^n \frac{c}{\sqrt{n}})^2}{\sigma_{ij}^2} + \frac{n_1}{n} \frac{1}{\sigma_{1j}^2} \right] = \\ &= \frac{1}{n} + \sigma_{x_j}^2 \left\{ \sum_{i=2}^p \frac{n_i}{n} \frac{1}{\sigma_{ij}^2} \left[ \beta_i^2 + 2\beta_i \frac{\delta^n c}{\sqrt{n}} + \left( \frac{\delta^n c}{\sqrt{n}} \right)^2 \right] + \frac{n_1}{n} \frac{1}{\sigma_{1j}^2} \right\} = \\ &= \frac{1}{n} + \sigma_{x_j}^2 \sum_{i=2}^p \frac{n_i}{n} \frac{1}{\sigma_{ij}^2} \left[ 2\beta_i \frac{\delta^n c}{\sqrt{n}} + \left( \frac{\delta^n c}{\sqrt{n}} \right)^2 \right] + \frac{n_1}{n} \frac{\sigma_{x_j}^2}{\sigma_{1j}^2} + \sigma_{x_j}^2 \sum_{i=2}^p \frac{n_i}{n} \frac{\beta_i^2}{\sigma_{ij}^2}. \end{aligned}$$

Como  $\delta^n$  é uma variável aleatória tal que  $0 < \delta^n < 1$ , obtivemos que

$$\begin{aligned} \left| \frac{a_j^n(\boldsymbol{\phi}^n)}{n} - w_1 \frac{\sigma_{x_j}^2}{\sigma_{1j}^2} - \sigma_{x_j}^2 \sum_{i=2}^p w_i \frac{\beta_i^2}{\sigma_{ij}^2} \right| &\leq \frac{1}{n} + \sigma_{x_j}^2 \sum_{i=2}^p \frac{n_i}{n} \frac{1}{\sigma_{ij}^2} \left[ \frac{c^2}{n} + 2 \frac{c}{\sqrt{n}} |\beta_i| \right] + \\ &+ \left| \frac{n_1}{n} - w_1 \right| \frac{\sigma_{x_j}^2}{\sigma_{1j}^2} + \sigma_{x_j}^2 \sum_{i=2}^p \left| \frac{n_i}{n} - w_i \right| \frac{\beta_i^2}{\sigma_{ij}^2} \end{aligned}$$

Como uma consequência, obtemos que

$$\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{a_j^n(\boldsymbol{\phi}^n)}{n} - \sigma_{x_j}^2 \left( w_1 \frac{1}{\sigma_{1j}^2} + \sum_{i=2}^p w_i \frac{\beta_i^2}{\sigma_{ij}^2} \right) \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad \varepsilon > 0,$$

e  $j = 1, \dots, m$ . Na sequência, tomamos os dados observados  $\mathbf{Y}^n$  com distribuição  $\mathbb{P}_{\boldsymbol{\psi}^n, (x_1, \dots, x_m)}$ , para todo  $n \geq 1$ . Por definição, temos que para todo  $i = 2, \dots, p$ ,

$$\frac{\sum_{k=1}^{n_i} Y_{ijk}}{n} = \frac{\sum_{k=1}^{n_i} \left[ (\alpha_i + \frac{c}{\sqrt{n}}) + (\beta_i + \frac{c}{\sqrt{n}}) x_j + e_{ijk} \right]}{n_i} = \frac{n_i}{n} \frac{\sum_{k=1}^{n_i} e_{ijk}}{n_i} + \frac{n_i}{n} \frac{c}{\sqrt{n}} (1 + x_j).$$

Então, concluímos que

$$\left| \frac{\sum_{k=1}^{n_i} Y_{ijk}}{n} - w_i (\alpha_i + \beta_i x_j) \right| \leq \left| \frac{n_i}{n} - w_i \right| |\alpha_i + \beta_i x_j| + \frac{n_i}{n} \left| \frac{\sum_{k=1}^{n_i} e_{ijk}}{n_i} \right| + \frac{n_i}{n} \frac{c}{\sqrt{n}} |1 + x_j|, \quad i = 2, \dots, p.$$

Aplicando os mesmos argumentos, obtemos que

$$\left| \frac{\sum_{k=1}^{n_i} Y_{1jk}}{n} - w_1 x_j \right| \leq \left| \frac{n_1}{n} - w_1 \right| |x_j| + \frac{n_1}{n} \left| \frac{\sum_{k=1}^{n_1} e_{1jk}}{n_1} \right| + \frac{n_1}{n} \frac{c}{\sqrt{n}} |1 + x_j|.$$

Como a distribuição do erro aleatório  $\{e_{ij\ell} : \ell = 1, 2, \dots\}$  é independente do parâmetro, pela lei dos grandes números, temos:

$$\mathbb{P}_{\Psi^n, (x_1, \dots, x_m)} \left[ \left| \frac{\sum_{k=1}^{n_i} e_{ijk}}{n_i} \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad \varepsilon > 0, \quad i = 1, \dots, p \quad \text{e} \quad j = 1, \dots, m.$$

Portanto, concluímos que

$$\mathbb{P}_{\Psi^n, (x_1, \dots, x_m)} \left[ \left| \frac{\sum_{k=1}^{n_i} Y_{ijk}}{n} - w_i (\alpha_i + \beta_i x_j) \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad \varepsilon > 0,$$

$i = 2, \dots, p$ ; e  $j = 1, \dots, m$ , e

$$\mathbb{P}_{\Psi^n, (x_1, \dots, x_m)} \left[ \left| \frac{\sum_{k=1}^{n_1} Y_{1jk}}{n} - w_1 x_j \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad \varepsilon > 0,$$

para todo  $j = 1, \dots, m$ . Além disso, para todo  $i = 2, \dots, p$ , temos que

$$\frac{D_{ij}^n(\phi^n)}{n} = \frac{\sum_{k=1}^{n_i} Y_{ijk}}{n} - \frac{n_i (\alpha_i + \frac{\delta^n c}{\sqrt{n}})}{n} \quad \text{e} \quad \frac{D_{1j}^n(\phi^n)}{n} = \frac{\sum_{k=1}^{n_1} Y_{1jk}}{n}.$$

Então, obtivemos que

$$\mathbb{P}_{\Psi^n, (x_1, \dots, x_m)} \left[ \left| \frac{D_{ij}^n(\phi^n)}{n} - w_i \beta_i x_j \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0 \quad \text{e} \quad \mathbb{P}_{\Psi^n, (x_1, \dots, x_m)} \left[ \left| \frac{D_{1j}^n(\phi^n)}{n} - w_1 x_j \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0.$$

Aplicando os mesmos argumentos, temos que

$$\frac{M_j^n(\phi^n)}{n} = \frac{\mu_{x_j} + \frac{\delta^n c}{\sqrt{n}}}{n \sigma_{x_j}^2} + \sum_{i=2}^p \left( \beta_i + \frac{\delta^n c}{\sqrt{n}} \right) \frac{D_{ij}^n(\phi^n)}{n \sigma_{ij}^2} + \frac{D_{1j}^n(\phi^n)}{n \sigma_{1j}^2}, \quad j = 1, \dots, m.$$

Então, concluímos que

$$\mathbb{P}_{\boldsymbol{\Psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{M_j^n(\boldsymbol{\phi}^n)}{n} - x_j \left( \sum_{i=2}^p \frac{w_i \beta_i^2}{\sigma_{ij}^2} + \frac{w_1}{\sigma_{1j}^2} \right) \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad j = 1, \dots, m.$$

Como consequência do teorema do mapeamento contínuo, concluímos que

$$\mathbb{P}_{\boldsymbol{\Psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{M_j^n(\boldsymbol{\phi}^n)}{a_j^n(\boldsymbol{\phi}^n)} - \frac{x_j \left( \sum_{i=2}^p \frac{w_i \beta_i^2}{\sigma_{ij}^2} + \frac{w_1}{\sigma_{1j}^2} \right)}{\sigma_{x_j}^2 \left( w_1 \frac{1}{\sigma_{1j}^2} + \sum_{i=2}^p w_i \frac{\beta_i^2}{\sigma_{ij}^2} \right)} \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad j = 1, \dots, m.$$

Na sequência, mostraremos a convergência uniforme para cada componente da matriz de informação observada. Temos que:

$$1. \frac{J_{(\mu_{x_j}, \mu_{x_j})}(\boldsymbol{\phi}^n)}{n} = -\frac{(1-a_j^n(\boldsymbol{\phi}^n))}{n\sigma_{x_j}^2 a_j^n(\boldsymbol{\phi}^n)} = -\frac{1}{\sigma_{x_j}^2} \left( \frac{1}{n} - \frac{a_j^n(\boldsymbol{\phi}^n)}{n} \right) \frac{1}{a_j^n(\boldsymbol{\phi}^n)}. \text{ Então, obtemos que}$$

$$\mathbb{P}_{\boldsymbol{\Psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{J_{(\mu_{x_j}, \mu_{x_j})}(\boldsymbol{\phi}^n)}{n} \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad \varepsilon > 0, \quad j = 1, \dots, m.$$

Nesse caso, temos que  $W_{(\mu_{x_j}, \mu_{x_j})}(\boldsymbol{\theta}) = 0$ , para todo  $j = 1, \dots, m$ .

$$2. \frac{J_{(\mu_{x_j}, \mu_{x_h})}(\boldsymbol{\phi}^n)}{n} = \frac{0}{n} \xrightarrow{n \rightarrow \infty} 0 = W_{(\mu_{x_j}, \mu_{x_h})}(\boldsymbol{\theta}), \text{ para todo } j \neq h = 1, \dots, m.$$

$$3. \frac{J_{(\mu_{x_j}, \alpha_i)}(\boldsymbol{\phi}^n)}{n} = \frac{1}{n} \frac{n_i \left( \beta_i + \delta^n \frac{c}{\sqrt{n}} \right)}{\sigma_{ij}^2 a_j^n(\boldsymbol{\phi}^n)} = \frac{n_i \left( \beta_i + \delta^n \frac{c}{\sqrt{n}} \right)}{n \sigma_{ij}^2} \frac{1}{a_j^n(\boldsymbol{\phi}^n)},$$

para todo  $i = 2, \dots, p$  e  $j = 1, \dots, m$ . Além disso, temos que

$$\frac{J_{(\mu_{x_j}, \alpha_1)}(\boldsymbol{\phi}^n)}{n} = \frac{n_1}{n} \frac{1}{\sigma_{1j}^2} \frac{1}{a_j^n(\boldsymbol{\phi}^n)}, \quad j = 1, \dots, m.$$

Como  $a_j^n(\boldsymbol{\phi}^n) \rightarrow \infty$ , obtemos que

$$\mathbb{P}_{\boldsymbol{\Psi}^n, (x_1, \dots, x_m)} \left[ \left| \frac{J_{(\mu_{x_j}, \alpha_i)}(\boldsymbol{\phi}^n)}{n} \right| \geq \varepsilon \right] \xrightarrow{n \rightarrow \infty} 0, \quad \varepsilon > 0, \quad i = 1, \dots, p, \quad j = 1, \dots, m.$$

4.

$$\begin{aligned} \frac{J_{(\mu_{x_j}, \beta_i)}(\boldsymbol{\phi}^n)}{n} &= \frac{1}{n} \frac{1}{\sigma_{ij}^2 a_j^n} \left( \frac{2n\beta_i^n \sigma_{x_j}^2 M_j^n}{a_j^n} - D_{ij}^n \right) = \frac{1}{\sigma_{ij}^2} \left( \frac{2n\beta_i^n \sigma_{x_j}^2 M_j^n}{na_j^n} - \frac{D_{ij}^n}{n} \right) \frac{1}{a_j^n} \\ &= \frac{1}{\sigma_{ij}^2} \left( 2\beta_i^n \sigma_{x_j}^2 \frac{M_j^n}{a_j^n} - \frac{D_{ij}^n}{n} \right) \frac{1}{a_j^n} \xrightarrow{n \rightarrow \infty} 0 = W_{(\mu_{x_j}, \beta_i)} \end{aligned}$$

5.

$$\begin{aligned} \frac{J(\alpha_i \alpha_i)}{k} &= \frac{1}{k} \sum_{j=1}^m \frac{k}{\sigma_{ij}^2} \left( 1 - \frac{k\beta_i^2 \sigma_{x_j}^2}{\sigma_{ij}^2 a_j^k} \right) = \sum_{j=1}^m \frac{k}{k\sigma_{ij}^2} \left( 1 - \frac{\beta_i^2 \sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \right) = \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \left( 1 - \frac{\beta_i^2 \sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \right) \\ &\xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \left( 1 - \frac{\beta_i^2 \sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) = \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} - \beta_i^2 \sum_{j=1}^m \frac{1}{(\sigma_{ij}^2)^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} = W_{(\alpha_i \alpha_i)} \end{aligned}$$

6.

$$\begin{aligned} \frac{J(\alpha_i \alpha_l)}{k} &= -\frac{1}{k} \sum_{j=1}^m \frac{k\beta_i \beta_l \sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j^k} = -\sum_{j=1}^m \frac{\beta_i \beta_l \sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \xrightarrow{k \rightarrow \infty} -\sum_{j=1}^m \frac{\beta_i \beta_l \sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \\ &= -\beta_i \beta_l \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sigma_{lj}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} = W_{(\alpha_i \alpha_l)} \end{aligned}$$

7.

$$\begin{aligned} \frac{J(\alpha_i \beta_i)}{k} &= \frac{1}{k} \sum_{j=1}^m \frac{k\sigma_{x_j}^2}{\sigma_{ij}^2 a_j^k} \left[ M_j^k - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2k\beta_i \sigma_{x_j}^2 M_j^k}{a_j^k} - D_{ij}^k \right) \right] \\ &= \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \left[ \frac{M_j^k}{k} - \frac{1}{k} \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2k\beta_i \sigma_{x_j}^2 M_j^k}{a_j^k} - D_{ij}^k \right) \right] \\ &= \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2} \left[ \frac{\frac{M_j^k}{k}}{\frac{a_j^k}{k}} - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2\beta_i \sigma_{x_j}^2 \frac{M_j^k}{k} - \frac{D_{ij}^k}{k}}{\frac{a_j^k}{k}} \right) \right] = \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2} \left[ \frac{M_j^k}{a_j^k} - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2\beta_i \sigma_{x_j}^2 \frac{M_j^k}{a_j^k} - \frac{D_{ij}^k}{a_j^k} \right) \right] \\ &\xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2} \left[ \frac{x_j}{\sigma_{x_j}^2} - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2\beta_i \sigma_{x_j}^2 \frac{x_j}{\sigma_{x_j}^2} - \beta_i x_j}{\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) \right] = \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2} \left[ \frac{x_j}{\sigma_{x_j}^2} - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2\beta_i x_j - \beta_i x_j}{\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) \right] \\ &= \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2} \left[ \frac{x_j}{\sigma_{x_j}^2} - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{\beta_i x_j}{\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) \right] = \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \left[ x_j - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{\beta_i x_j}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) \right] \\ &= \sum_{j=1}^m \frac{x_j}{\sigma_{ij}^2} - \beta_i^2 \sum_{j=1}^m \frac{x_j}{(\sigma_{ij}^2)^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} = W_{(\alpha_i \beta_i)} \end{aligned}$$

8.

$$\begin{aligned}
\frac{J_{(\alpha_i\beta_l)}}{k} &= \frac{1}{k} \sum_{j=1}^m \frac{k\beta_i\sigma_{x_j}^2}{\sigma_{ij}^2\sigma_{lj}^2a_j^k} \left( D_{lj}^k - \frac{2k\beta_l\sigma_{x_j}^2M_j^k}{a_j^k} \right) = \sum_{j=1}^m \frac{\beta_i\sigma_{x_j}^2}{\sigma_{ij}^2\sigma_{lj}^2\frac{a_j^k}{k}} \left( \frac{D_{lj}^k}{k} - 2\beta_l\sigma_{x_j}^2\frac{M_j^k}{a_j^k} \right) \\
&= \sum_{j=1}^m \frac{\beta_i\sigma_{x_j}^2}{\sigma_{ij}^2\sigma_{lj}^2\frac{a_j^k}{k}} \left( \frac{D_{lj}^k}{k} - 2\beta_l\sigma_{x_j}^2\frac{M_j^k}{a_j^k} \right) \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{\beta_i\sigma_{x_j}^2}{\sigma_{ij}^2\sigma_{lj}^2\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left( \beta_l x_j - 2\beta_l\sigma_{x_j}^2\frac{x_j}{\sigma_{x_j}^2} \right) \\
&= \sum_{j=1}^m \frac{\beta_i}{\sigma_{ij}^2\sigma_{lj}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} (\beta_l x_j - 2\beta_l x_j) = -\beta_i\beta_l \sum_{j=1}^m \frac{x_j}{\sigma_{ij}^2\sigma_{lj}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} = W_{(\alpha_i\beta_l)}
\end{aligned}$$

9.

$$\begin{aligned}
\frac{J_{(\beta_i\beta_i)}}{k} &= \frac{1}{k} \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 a_j^k} \left\{ k - \frac{D_{ij}^k{}^2}{\sigma_{ij}^2} + \frac{k\sigma_{x_j}^2}{a_j^k} \left[ M_j^{k2} \left( 1 - \frac{4k\sigma_{x_j}^2\beta_i^2}{\sigma_{ij}^2 a_j^k} \right) + \frac{4\beta_i M_j^k D_{ij}^k}{\sigma_{ij}^2} - \frac{2k\beta_i^2}{\sigma_{ij}^2} \right] \right\} \\
&= \frac{1}{k^2} \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \left\{ k - \frac{D_{ij}^k{}^2}{\sigma_{ij}^2} + \frac{k\sigma_{x_j}^2}{a_j^k} \left[ M_j^{k2} \left( 1 - \frac{4k\sigma_{x_j}^2\beta_i^2}{\sigma_{ij}^2 a_j^k} \right) + \frac{4\beta_i M_j^k D_{ij}^k}{\sigma_{ij}^2} - \frac{2k\beta_i^2}{\sigma_{ij}^2} \right] \right\} \\
&= \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \left\{ \frac{k}{k^2} - \frac{1}{\sigma_{ij}^2} \left( \frac{D_{ij}^k}{k} \right)^2 + \frac{\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \left( \frac{M_j^k}{k} \right)^2 \left( 1 - \frac{4\sigma_{x_j}^2\beta_i^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \right) + \frac{4\beta_i M_j^k D_{ij}^k}{\sigma_{ij}^2 k} - \frac{2k\beta_i^2}{k^2 \sigma_{ij}^2} \right] \right\} \\
&= \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \left\{ \frac{1}{k} - \frac{1}{\sigma_{ij}^2} \left( \frac{D_{ij}^k}{k} \right)^2 + \frac{\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \left( \frac{a_j^k}{k} \right)^2 \left( \frac{M_j^k}{k} \right)^2 \left( 1 - \frac{4\sigma_{x_j}^2\beta_i^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \right) + \right. \right. \\
&\quad \left. \left. \frac{4\beta_i M_j^k D_{ij}^k}{\sigma_{ij}^2 \frac{a_j^k}{k} k} - \frac{2\beta_i^2}{\sigma_{ij}^2 k} \right] \right\} \\
&= \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \left\{ \frac{1}{k} - \frac{1}{\sigma_{ij}^2} \left( \frac{D_{ij}^k}{k} \right)^2 + \frac{\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \left( \frac{a_j^k}{k} \right)^2 \left( \frac{M_j^k}{k} \right)^2 \left( 1 - \frac{4\sigma_{x_j}^2\beta_i^2}{\sigma_{ij}^2 \frac{a_j^k}{k}} \right) + \right. \right. \\
&\quad \left. \left. \frac{4\beta_i M_j^k D_{ij}^k}{\sigma_{ij}^2 a_j^k k} - \frac{2\beta_i^2}{\sigma_{ij}^2 k} \right] \right\} \\
&\xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ 0 - \frac{(\beta_i x_j)^2}{\sigma_{ij}^2} + \frac{\sigma_{x_j}^2}{\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ \left( \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)^2 \left( \frac{x_j}{\sigma_{x_j}^2} \right)^2 \times \right. \right. \\
&\quad \left. \left. \left( 1 - \frac{4\sigma_{x_j}^2\beta_i^2}{\sigma_{ij}^2\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) + \frac{4\beta_i x_j}{\sigma_{ij}^2 \sigma_{x_j}^2} \beta_i x_j \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} - \frac{2\beta_i^2}{\sigma_{ij}^2} 0 \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ -\frac{(\beta_i x_j)^2}{\sigma_{ij}^2} + \frac{1}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)^2 (x_j)^2 \left( 1 - \frac{4\beta_i^2}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \right) \right. \right. \\
& \quad \left. \left. + \frac{4\beta_i^2 x_j^2}{\sigma_{ij}^2} \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right] \right\} \\
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ -\frac{(\beta_i x_j)^2}{\sigma_{ij}^2} + \frac{1}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ x_j^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)^2 - \frac{4\beta_i^2}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)^2 \right. \right. \\
& \quad \left. \left. + \frac{4\beta_i^2 x_j^2}{\sigma_{ij}^2} \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right] \right\} \\
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ -\frac{(\beta_i x_j)^2}{\sigma_{ij}^2} + \frac{1}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ x_j^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)^2 - \frac{4\beta_i^2 x_j^2}{\sigma_{ij}^2} \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right. \right. \\
& \quad \left. \left. + \frac{4\beta_i^2 x_j^2}{\sigma_{ij}^2} \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right] \right\} \\
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ -\frac{(\beta_i x_j)^2}{\sigma_{ij}^2} + \frac{1}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ x_j^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)^2 \right] \right\} \\
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ -\frac{(\beta_i x_j)^2}{\sigma_{ij}^2} + x_j^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right\} \\
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \left\{ -\frac{(\beta_i x_j)^2}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} + \frac{x_j^2}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right\} \\
& \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \left\{ -\frac{(\beta_i x_j)^2}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} + x_j^2 \right\} \xrightarrow{k \rightarrow \infty} \sum_{j=1}^m \frac{x_j^2}{\sigma_{ij}^2} \left\{ 1 - \frac{\beta_i^2}{\sigma_{ij}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} \right\} = W_{(\beta_i \beta_i)}
\end{aligned}$$

10.

$$\begin{aligned}
\frac{J_{(\beta_i \beta_l)}}{k} &= -\frac{1}{k} \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j^k} \left\{ D_{ij}^k D_{lj}^k + \frac{2\sigma_{x_j}^2}{a_j^k} \left[ n_i n_l \beta_i \beta_l \left( 1 + \frac{2\sigma_{x_j}^2 M_j^{k^2}}{a_j^k} \right) \right. \right. \\
& \quad \left. \left. - n_i \beta_i M_j^k D_{lj}^k - n_l \beta_l M_j^k D_{ij}^k \right] \right\}
\end{aligned}$$

$$\begin{aligned}
&= -\frac{1}{k^2} \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \left\{ D_{ij}^k D_{lj}^k + \frac{2\sigma_{x_j}^2}{a_j^k} \left[ kk\beta_i\beta_l \left( 1 + \frac{2\sigma_{x_j}^2 M_j^{k^2}}{a_j^k} \right) \right. \right. \\
&\quad \left. \left. - k\beta_i M_j^k D_{lj}^k - k\beta_l M_j^k D_{ij}^k \right] \right\} \\
&= -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \left\{ \frac{D_{ij}^k D_{lj}^k}{k} + \frac{2\sigma_{x_j}^2}{a_j^k} \left[ \frac{k^2}{k^2} \beta_i\beta_l \left( 1 + \frac{2\sigma_{x_j}^2 M_j^{k^2}}{a_j^k} \right) \right. \right. \\
&\quad \left. \left. - k\beta_i \frac{M_j^k D_{lj}^k}{k} - k\beta_l \frac{M_j^k D_{ij}^k}{k} \right] \right\} \\
&= -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \left\{ \frac{D_{ij}^k D_{lj}^k}{k} + \frac{2\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \beta_i\beta_l \left( \frac{1}{k} + \frac{2\sigma_{x_j}^2 M_j^{k^2}}{a_j^k k} \right) \right. \right. \\
&\quad \left. \left. - \frac{k}{k} \beta_i \frac{M_j^k D_{lj}^k}{k} - \frac{k}{k} \beta_l \frac{M_j^k D_{ij}^k}{k} \right] \right\} \\
&= -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \left\{ \frac{D_{ij}^k D_{lj}^k}{k} + \frac{2\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \beta_i\beta_l \left( \frac{1}{k} + \frac{2\sigma_{x_j}^2 M_j^{k^2}}{\frac{a_j^k}{k} k^2} \right) \right. \right. \\
&\quad \left. \left. - \beta_i \frac{M_j^k D_{lj}^k}{k} - \beta_l \frac{M_j^k D_{ij}^k}{k} \right] \right\} \\
&= -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \left\{ \frac{D_{ij}^k D_{lj}^k}{k} + \frac{2\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \beta_i\beta_l \left( \frac{1}{k} + 2\sigma_{x_j}^2 \left( \frac{M_j^k}{\frac{a_j^k}{k}} \right)^2 \frac{a_j^k}{k} \right) \right. \right. \\
&\quad \left. \left. - \beta_i \frac{\frac{M_j^k}{\frac{a_j^k}{k}} D_{lj}^k a_j^k}{\frac{a_j^k}{k} k} - \beta_l \frac{\frac{M_j^k}{\frac{a_j^k}{k}} D_{ij}^k a_j^k}{\frac{a_j^k}{k} k} \right] \right\} \\
&= -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \frac{a_j^k}{k}} \left\{ \frac{D_{ij}^k D_{lj}^k}{k} + \frac{2\sigma_{x_j}^2}{\frac{a_j^k}{k}} \left[ \beta_i\beta_l \left( \frac{1}{k} + 2\sigma_{x_j}^2 \left( \frac{M_j^k}{a_j^k} \right)^2 \frac{a_j^k}{k} \right) \right. \right. \\
&\quad \left. \left. - \beta_i \frac{M_j^k D_{lj}^k a_j^k}{a_j^k k} - \beta_l \frac{M_j^k D_{ij}^k a_j^k}{a_j^k k} \right] \right\} \\
&\xrightarrow{k \rightarrow \infty} -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ \beta_i x_j \beta_l x_j + \frac{2\sigma_{x_j}^2}{\sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ \beta_i\beta_l \left( 0 + 2\sigma_{x_j}^2 \left( \frac{x_j}{\sigma_{x_j}^2} \right)^2 \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right) \right. \right. \\
&\quad \left. \left. - \beta_i \frac{x_j}{\sigma_{x_j}^2} \beta_l x_j \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} - \beta_l \frac{x_j}{\sigma_{x_j}^2} \beta_i x_j \sigma_{x_j}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right] \right\} \\
&\xrightarrow{k \rightarrow \infty} -\sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sigma_{lj}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ \beta_i\beta_l x_j^2 + \frac{2}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ \beta_i\beta_l \left( 2x_j^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right) \right. \right.
\end{aligned}$$

$$\begin{aligned}
& \left. -\beta_i\beta_l x_j^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} - \beta_l\beta_i x_j^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right\} \\
\stackrel{k \rightarrow \infty}{\longrightarrow} & - \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sigma_{lj}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ \beta_i\beta_l x_j^2 + \frac{2}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left[ 2\beta_i\beta_l x_j^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} - 2\beta_l\beta_i x_j^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right] \right\} \\
\stackrel{k \rightarrow \infty}{\longrightarrow} & - \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sigma_{lj}^2 \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} \left\{ \beta_i\beta_l x_j^2 + \frac{2}{\sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2}} [0] \right\} \\
\stackrel{k \rightarrow \infty}{\longrightarrow} & -\beta_i\beta_l \sum_{j=1}^m \frac{x_j^2}{\sigma_{ij}^2 \sigma_{lj}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} = W_{(\beta_i\beta_l)}
\end{aligned}$$

Por conseguinte:

$$\begin{aligned}
W_{(\mu_{x_j}\mu_{x_h})} &= W_{(\mu_{x_j}\mu_{x_h})} = W_{(\mu_{x_j}\alpha_i)} = W_{(\mu_{x_j}\beta_i)} = 0, \quad j, h = 1, \dots, m; \quad j \neq h, \quad i = 2, \dots, p; \\
W_{(\alpha_i\alpha_i)} &= \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} - \beta_i^2 \sum_{j=1}^m \frac{1}{(\sigma_{ij}^2)^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)}, \quad W_{(\alpha_i\alpha_l)} = -\beta_i\beta_l \sum_{j=1}^m \frac{1}{\sigma_{ij}^2 \sigma_{lj}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)}, \\
W_{(\alpha_i\beta_i)} &= \sum_{j=1}^m \frac{x_j}{\sigma_{ij}^2} - \beta_i^2 \sum_{j=1}^m \frac{x_j}{(\sigma_{ij}^2)^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)}, \quad W_{(\alpha_i\beta_l)} = -\beta_i\beta_l \sum_{j=1}^m \frac{x_j}{\sigma_{ij}^2 \sigma_{lj}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)}, \\
W_{(\beta_i\beta_i)} &= \sum_{j=1}^m \frac{x_j^2}{\sigma_{ij}^2} \left\{ 1 - \frac{\beta_i^2}{\sigma_{ij}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)} \right\}, \quad \text{e } W_{(\beta_i\beta_l)} = -\beta_i\beta_l \sum_{j=1}^m \frac{x_j^2}{\sigma_{ij}^2 \sigma_{lj}^2 \left( \sum_{q=1}^p \frac{\beta_q^2}{\sigma_{qj}^2} \right)}, \\
& \quad \quad \quad i, l = 2, \dots, p; \quad i \neq l
\end{aligned}$$



---



---

## OUTROS

---



---

### C.1 Identificabilidade

**Definição:** Um modelo é identificável se, e somente se,  $\theta_1 \neq \theta_2$  implica que  $F_{\theta_1} \neq F_{\theta_2}$  para todo  $\theta_1, \theta_2 \in \Theta$ . O vetor de parâmetros  $\theta$  é identificável se, e somente se, todos os seus componentes são identificáveis [Cheng e Ness \(1999\)](#).

**Exemplo** Considere o modelo definido em (1.1 e 1.2) e sejam dois conjuntos de parâmetros dados por:

$$\theta_1 = (\mu, \sigma_x^2, \sigma_e^2, \sigma_\delta^2, \alpha, \beta)^T = (3, 2, 5, 2, 4, 1)^T$$

e

$$\theta_2 = (\mu, \sigma_x^2, \sigma_e^2, \sigma_\delta^2, \alpha, \beta)^T = (3, 1, 3, 3, 1, 2)^T.$$

Ambos vetores de parâmetros conduzem à mesma distribuição no modelo definido em (1.1 e 1.2)

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N \left( \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 & 2 \\ 2 & 7 \end{bmatrix} \right), \quad (\text{C.1})$$

e por tanto, o modelo não é identificável.

### C.2 Exemplo da forma paramétrica da elipse

Uma elipse com centro  $(h, k)$  e semieixos  $a$  e  $b$  é dado por:

$$\left( \frac{x-h}{a} \right)^2 + \left( \frac{y-k}{b} \right)^2 = 1,$$

i.e., o conjunto de pontos  $(x(t), y(t))$  em que,

$$x(t) = h + a \cos(t),$$

$$y(t) = k + b \sin(t),$$

e  $0 \leq t \leq 2\pi$

A forma paramétrica da elipse, com inclinação gira no sentido anti-horário por  $\tau$  radians é dado por:

$$x(t) = h + \cos(\tau) [a \cos(t)] - \sin(\tau) [b \sin(t)] \quad \text{e}$$

$$y(t) = k + \sin(\tau) [a \cos(t)] + \cos(\tau) [b \sin(t)].$$

O conjunto de pontos  $(x, y)$  do plano cartesiano que satisfaz

$$Ax^2 + By^2 + Cxy + Dx + Ey + F = 0 \quad (\text{C.2})$$

com  $C^2 - 4AB < 0$  define uma elipse. Definindo a matriz

$$M_0 = \begin{bmatrix} F & \frac{D}{2} & \frac{E}{2} \\ \frac{D}{2} & A & \frac{C}{2} \\ \frac{E}{2} & \frac{C}{2} & B \end{bmatrix} \quad \text{e} \quad M = \begin{bmatrix} A & \frac{C}{2} \\ \frac{C}{2} & B \end{bmatrix},$$

temos que:

$$a = \sqrt{-\frac{\det(M_0)}{\det(M)\lambda_1}}, \quad b = \sqrt{-\frac{\det(M_0)}{\det(M)\lambda_2}},$$

$$h = \frac{CE - 2BD}{4AB - C^2}, \quad k = \frac{CD - 2AE}{4AB - C^2} \quad \text{e}$$

$$\tau = \frac{\operatorname{arccot}\left(\frac{A-B}{C}\right)}{2}$$

onde,  $\lambda_1$  e  $\lambda_2$  são os autovalores de  $M$  ordenado tal que  $|\lambda_1 - A| \leq |\lambda_1 - B|$  (isso garante que  $|\lambda_2 - B| \leq |\lambda_2 - A|$ ).

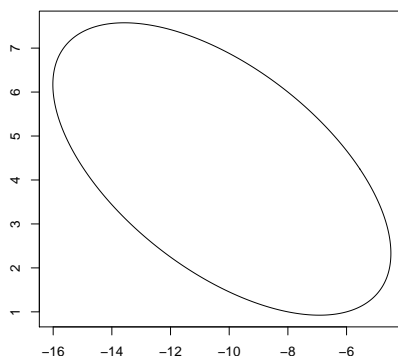


Figura 61 – Exemplo de forma geral da elipse:  $x^2 + 3y^2 + 2xy + 12x - 5y + 50 = 0$ .

### C.3 Base ortonormal

Um conjunto de vetores diferentes de zero  $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$  em um subespaço vetorial  $W$  de  $\mathbb{R}^n$  é uma base ortogonal, se for uma base, e se os vetores forem mutuamente ortogonais, i.e.,  $\langle \mathbf{w}_i, \mathbf{w}_j \rangle = 0$  para todos os  $i \neq j$  e  $1 \leq i, j \leq r$ . Se além disso, o comprimento de cada  $\mathbf{w}_j$  é igual a um, então a base é considerada uma base ortonormal.

Por exemplo, a base padrão  $\{\mathbf{e}_1, \dots, \mathbf{e}_r\}$  é uma base ortonormal para  $\mathbb{R}^n$ . Note que se  $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$  é uma base ortogonal para um subespaço vetorial  $W$ , então é possível normalizar o comprimento de cada vetor para obter uma base ortonormal. Para ser mais preciso, a coleção de vetores

$$\frac{1}{\|\mathbf{w}_1\|} \mathbf{w}_1, \dots, \frac{1}{\|\mathbf{w}_r\|} \mathbf{w}_r$$

é uma base ortonormal.

Além disso, se a coleção de vetores diferentes de zero  $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$  for mutuamente ortogonal, eles serão linearmente independentes, porque se houver constantes  $\lambda_j$ , tais que

$$\lambda \mathbf{w}_1 + \dots + \lambda_r \mathbf{w}_r = \mathbf{0},$$

então, para cada  $j$ ,  $1 \leq j \leq r$ .

$$0 = \langle \mathbf{w}_j, \mathbf{0} \rangle = \langle \mathbf{w}_j, \lambda_1 \mathbf{w}_1 + \lambda_2 \mathbf{w}_2 + \dots + \lambda_r \mathbf{w}_r \rangle = \lambda_j \langle \mathbf{w}_j, \mathbf{w}_j \rangle.$$

isso só é possível quando  $\lambda_j = 0$  para cada  $j$ . Em particular, a dimensão da extensão linear de  $\{\mathbf{w}_1, \dots, \mathbf{w}_r\}$  é igual a  $r$ .

## C.4 Modelo de calibração Ultraestrutural com réplicas

### C.4.1 Algoritmo EM

Nesta seção vamos apresentar o algoritmo EM (Dempster, Laird e Rubin (1977)) usado para obter a estimativa dos parâmetros descrito em Talarico (2014). A ideia é a utilização dos dados latentes  $x_j$ ,  $j = 1, \dots, m$  para aumentar os dados observados de tal forma que as estimativas de máxima verossimilhança (EMV) dos parâmetros baseados nos dados aumentados (dados completos) sejam fáceis de obter. Considerando o modelo definido por (1.14) e (1.15) e os dados observados no  $j$ -ésimo valor de rotação do motor,  $\mathbf{Y}_j^n = (\mathbf{Y}_{1j}^\top, \dots, \mathbf{Y}_{pj}^\top)^\top$ , nós aumentamos  $\mathbf{Y}_j^n$  considerando os dados não observados  $x_j$ . Então, os dados completos para o  $j$ -ésimo valor de rotação do motor é dado por  $\mathbf{Y}_{jc} = (x_j, \mathbf{Y}_j^{nT})^\top$ , com  $E(\mathbf{Y}_{jc}) = \boldsymbol{\mu}_{jc} = (\mu_{x_j}, \boldsymbol{\mu}_j^T)^\top$  e a matriz de variâncias e covariâncias dado por  $\boldsymbol{\Sigma}_{jc} = \begin{pmatrix} \sigma_{x_j}^2 & \boldsymbol{\Sigma}_{12j} \\ \boldsymbol{\Sigma}_{21j} & \boldsymbol{\Sigma}_j \end{pmatrix}$ ,  $j = 1, \dots, m$ , com  $\boldsymbol{\Sigma}_{12j} = \boldsymbol{\Sigma}_{21j}^\top = \sigma_{x_j}^2 \boldsymbol{\beta}^T$ ,  $\boldsymbol{\mu}_j$  e  $\boldsymbol{\Sigma}_j$  como definido em (1.3). Além do mais,  $\mathbf{Y}_{jc} \sim N_{n+1}(\boldsymbol{\mu}_{jc}, \boldsymbol{\Sigma}_{jc})$  e seja  $\mathbf{Y}_c = (\mathbf{Y}_{1c}^\top, \dots, \mathbf{Y}_{mc}^\top)^\top$ , então

$$f_{\mathbf{Y}_c}(\mathbf{Y}_c) = \prod_{j=1}^m f_{\mathbf{Y}_{jc}}(\mathbf{Y}_{jc})$$

$$= (2\pi)^{-\frac{m(n+1)}{2}} \left[ \prod_{j=1}^m \left( \sigma_{x_j}^2 (\sigma_{1j}^2)^{n_1} \dots (\sigma_{pj}^2)^{n_p} \right)^{-\frac{1}{2}} \right] \exp \left\{ - \sum_{j=1}^m \frac{1}{2} (\mathbf{Y}_{jc} - \boldsymbol{\mu}_{jc})^\top \boldsymbol{\Sigma}_{jc}^{-1} (\mathbf{Y}_{jc} - \boldsymbol{\mu}_{jc}) \right\}.$$

Segue que o logaritmo da função de verossimilhança dos dados completos é dado por

$$L_c(\boldsymbol{\theta}) = cte - \frac{1}{2} \sum_{j=1}^m (\log \sigma_{x_j}^2 + \sum_{i=1}^p n_i \log \sigma_{ij}^2) - \frac{1}{2} \left[ \sum_{j=1}^m \frac{(x_j - \mu_{x_j})^2}{\sigma_{x_j}^2} + \sum_{j=1}^m \sum_{k=1}^{n_1} \frac{(Y_{1jk} - x_j)^2}{\sigma_{1j}^2} + \sum_{j=1}^m \sum_{i=2}^p \sum_{k=1}^{n_i} \frac{(Y_{ijk} - \alpha_i - \beta_i x_j)^2}{\sigma_{ij}^2} \right],$$

que é muito mais simples do que (1.16). Dada as estimativas de  $\boldsymbol{\theta}$  na  $r$ -ésima iteração,  $\boldsymbol{\theta}^{(r)}$ , o passo E consiste na obtenção da esperança do logaritmo da função de verossimilhança do modelo com dados completos,  $L_c(\boldsymbol{\theta})$ , com respeito à distribuição condicional de  $\mathbf{x} = (x_1, \dots, x_m)^\top$  dado os dados observados  $\mathbf{Y}^n$  e  $\boldsymbol{\theta}^{(r)}$ . O passo M consiste na maximização da função obtida no passo E com respeito a  $\boldsymbol{\theta}$ , obtendo-se assim as estimativas dos parâmetros da iteração seguinte,  $\boldsymbol{\theta}^{(r+1)}$ . A cada iteração o algoritmo EM incrementa o logaritmo da função de verossimilhança dos dados observados,  $L^n(\boldsymbol{\theta})$ , i.e.,  $L^n(\boldsymbol{\theta}^{(r)}) \leq L^n(\boldsymbol{\theta}^{(r+1)})$ . Quando a função dos dados completos pertence a família exponencial, a implementação do algoritmo EM é usualmente simples. Em nosso caso, o passo E consiste na obtenção de  $\mathbb{E}_{\boldsymbol{\theta}^{(r)}}[x_j/\mathbf{Y}^n]$  e  $\mathbb{E}_{\boldsymbol{\theta}^{(r)}}[x_j^2/\mathbf{Y}^n]$ ,  $j = 1, \dots, m$ . No passo M, maximizamos o logaritmo da função de verossimilhança dos dados completos em que os valores da estatística suficiente foram substituídos pelos valores esperados obtidos no passo E.

O algoritmo EM para o modelo definido por (1.14) e (1.15) pode ser resumido da seguinte forma:

Passo-E: considerando as propriedades da distribuição normal multivariada, o passo-E consiste na obtenção de

$$\hat{x}_j^{(r)} = \mathbb{E}_{\boldsymbol{\theta}^{(r-1)}} [x_j / \mathbf{Y}_j^n] = \frac{\sigma_{x_j}^2 M_j^{n(r-1)}}{a_j^{n(r-1)}},$$

$$\text{em que } M_j^{n(r-1)} = \left[ \frac{\mu_{x_j}^{(r-1)}}{\sigma_{x_j}^2} + \sum_{k=1}^{n_1} \frac{Y_{1jk}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i^{(r-1)}}{\sigma_{ij}^2} (\sum_{k=1}^{n_i} Y_{ijk} - n_i \alpha_i^{(r-1)}) \right] e$$

$$\hat{x}_j^2^{(r)} = \mathbb{E}_{\boldsymbol{\theta}^{(r-1)}} [x_j^2 / \mathbf{Y}_j^n] = \frac{\sigma_{x_j}^2}{a_j^{n(r-1)}} + \left( \hat{x}_j^{(r)} \right)^2,$$

com  $a_j^{n(r-1)}$  representando o valor de  $a_j^n$  em  $\boldsymbol{\theta}^{(r-1)}$ .

Passo-M: o passo-M consiste na obtenção de

$$\hat{\mu}_{x_j}^{(r)} = \hat{x}_j^{(r)}, \quad j = 1, \dots, m,$$

$$\hat{\beta}_i^{(r)} = \frac{\left( \sum_{j=1}^m \frac{\hat{x}_j^{(r)}}{\sigma_{ij}^2} \sum_{k=1}^{n_i} Y_{ijk} \right) \left( \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \right) - \left( \sum_{j=1}^m \frac{\hat{x}_j^{(r)}}{\sigma_{ij}^2} \right) \left( \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \sum_{k=1}^{n_i} Y_{ijk} \right)}{n_i \left[ \left( \sum_{j=1}^m \frac{\hat{x}_j^2^{(r)}}{\sigma_{ij}^2} \right) \left( \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \right) - \left( \sum_{j=1}^m \frac{\hat{x}_j^{(r)}}{\sigma_{ij}^2} \right)^2 \right]},$$

e

$$\hat{\alpha}_i^{(r)} = \frac{\left( \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \sum_{k=1}^{n_i} Y_{ijk} - n_i \hat{\beta}_i^{(r)} \sum_{j=1}^m \frac{\hat{x}_j^{(r)}}{\sigma_{ij}^2} \right)}{n_i \sum_{j=1}^m \frac{1}{\sigma_{ij}^2}}, \quad i = 1, \dots, p.$$

Serão apresentados a função score e a matriz de informação observada.

### C.4.2 Função score e matriz de informação observada.

Em Talarico (2014) podem ser encontrados os elementos da função score,  $U^n(\boldsymbol{\theta}) = \frac{\partial L^n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ , denotado por  $U_{\theta_q}^n$ ,  $q = 1, \dots, m + 2(p-1)$ , que é dado por:

$$U_{\mu_{x_j}}^n = \frac{M_j^n}{a_j^n} - \frac{\mu_{x_j}}{\sigma_{x_j}^2}, \quad j = 1, \dots, m; \quad U_{\alpha_i}^n = - \sum_{j=1}^m \frac{1}{\sigma_{ij}^2} \left[ \frac{n_i \beta_i M_j^n \sigma_{x_j}^2}{a_j^n} - D_{ij}^n \right],$$

e

$$U_{\beta_i}^n = - \sum_{j=1}^m \frac{\sigma_{x_j}^2}{a_j^n \sigma_{ij}^2} \left\{ n_i \beta_i + M_j^n \left[ \frac{n_i \beta_i \sigma_{x_j}^2 M_j^n}{a_j^n} - D_{ij}^n \right] \right\}, \quad i = 2, \dots, p,$$

com  $M_j^n = \frac{\mu_{x_j}}{\sigma_{x_j}^2} + \frac{\mathbf{Y}_{1j}^T \mathbf{1}_{n_1}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i D_{ij}^n}{\sigma_{ij}^2}$ ,  $D_{ij}^n = (\mathbf{Y}_{ij} - \alpha_i \mathbf{1}_{n_i})^T \mathbf{1}_{n_i}$ . e  $a_j^n = 1 + \sigma_{x_j}^2 \boldsymbol{\beta}^\top D^{-1}(\sigma_{x_j}^2) \boldsymbol{\beta}$  como definido em (1.16)

A matriz de informação observada,  $J^n(\boldsymbol{\theta}) = -\frac{\partial^2 L^n(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$  com  $\boldsymbol{\theta} = (\mu_{x_1}, \dots, \mu_{x_m}, \alpha_2, \dots, \alpha_p, \beta_2, \dots, \beta_p)^T = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_{m+2(p-1)})^T$ , foi obtida em Talarico (2014) com os elementos dados por  $J^n_{\theta_r, \theta_h}, r, h = 1, \dots, m+2(p-1)$ ,

$$J^n_{\mu_{x_j}, \mu_{x_j}} = -\frac{(1-a_j^n)}{\sigma_{x_j}^2 a_j^n}, J^n_{\mu_{x_j}, \mu_{x_q}} = 0, J^n_{\mu_{x_j}, \alpha_i} = \frac{n_i \beta_i}{\sigma_{ij}^2 a_j^n}, J^n_{\mu_{x_j}, \beta_i} = \frac{1}{\sigma_{ij}^2 a_j^n} \left( \frac{2n_i \beta_i \sigma_{x_j}^2 M_j^n}{a_j^n} - D_{ij}^n \right),$$

$$J^n_{\alpha_i, \alpha_i} = \sum_{j=1}^m \frac{n_i}{\sigma_{ij}^2} \left( 1 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{\sigma_{ij}^2 a_j^n} \right), J^n_{\alpha_i, \alpha_l} = -\sum_{j=1}^m \frac{n_i n_l \beta_i \beta_l \sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j^n},$$

$$J^n_{\alpha_i, \beta_i} = \sum_{j=1}^m \frac{n_i \sigma_{x_j}^2}{\sigma_{ij}^2 a_j^n} \left[ M_j^n - \frac{\beta_i}{\sigma_{ij}^2} \left( \frac{2n_i \beta_i \sigma_{x_j}^2 M_j^n}{a_j^n} - D_{ij}^n \right) \right], J^n_{\alpha_i, \beta_l} = \sum_{j=1}^m \frac{n_i \beta_i \sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j^n} \left( D_{lj}^n - \frac{2n_l \beta_l \sigma_{x_j}^2 M_j^n}{a_j^n} \right),$$

$$J^n_{\beta_i, \beta_i} = \sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 a_j^n} \left\{ n_i - \frac{(D_{ij}^n)^2}{\sigma_{ij}^2} + \frac{n_i \sigma_{x_j}^2}{a_j^n} \left[ (M_j^n)^2 \left( 1 - \frac{4n_i \sigma_{x_j}^2 \beta_i^2}{\sigma_{ij}^2 a_j^n} \right) + \frac{4\beta_i M_j^n D_{ij}^n}{\sigma_{ij}^2} - \frac{2n_i \beta_i^2}{\sigma_{ij}^2} \right] \right\},$$

$$J^n_{\beta_i, \beta_l} = -\sum_{j=1}^m \frac{\sigma_{x_j}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j^n} \left\{ D_{ij}^n D_{lj}^n + \frac{2\sigma_{x_j}^2}{a_j^n} \left[ n_i n_l \beta_i \beta_l \left( 1 + \frac{2\sigma_{x_j}^2 (M_j^n)^2}{a_j^n} \right) - n_i \beta_i M_j^n D_{lj}^n - n_l \beta_l M_j^n D_{ij}^n \right] \right\},$$

$j \neq q, i \neq l, j, q = 1, \dots, m, i, l = 2, \dots, p$  e com  $a_j^n, D_{ij}^n$  e  $M_j^n$  como dado na Seção 2.

## C.5 Modelo de regressão multivariado com erros nas variáveis com intercepto nulo

A seguir serão apresentadas a matriz de informação observada e a matriz  $\Delta$  para cada tipo de perturbação considerando o modelo definido em (1.11).

### C.5.1 Matriz de informação observada

Considerando,  $\ell(\mathbf{z}, \boldsymbol{\theta})$  como definido em (1.13) onde  $\boldsymbol{\theta}_{(4p+3) \times 1} = (\boldsymbol{\beta}_1^T, \dots, \boldsymbol{\beta}_p^T, \mu, \sigma_\delta^2, \sigma_x^2, \sigma_{e_1}^2, \dots, \sigma_{e_p}^2)^T$  com  $\boldsymbol{\beta}_i = (\beta_{1i}, \beta_{2i})^T$  e  $\boldsymbol{\sigma}_{e_i}^2 = (\sigma_{e_{1i}}^2, \sigma_{e_{2i}}^2)^T$ , a matriz de informação observada, que se encontra em Russo (2006, Apêndice), é dado por:

$$\begin{aligned}
\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i^T} &= \sum_{j=1}^{n_i} \sigma_x^2 b_i^{-1} (\mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu) \cdot (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2)) \\
&\quad - 2\sigma_x^4 b_i^{-2} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] (\mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot [\boldsymbol{\beta}_i \cdot (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu)^T \\
&\quad + (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu) \cdot \boldsymbol{\beta}_i^T] \cdot \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2)) \\
&\quad + 2\sigma_x^4 b_i^{-2} (1 + 2\sigma_x^2 b_i^{-1} ((\mathbf{z}_{ij} - \mathbf{m}_i)^T \cdot \mathbf{B}_i \cdot (\mathbf{z}_{ij} - \mathbf{m}_i))) \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i \cdot \boldsymbol{\beta}_i^T \cdot \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \\
&\quad - \left\{ \mu^2 + \sigma_x^2 b_i^{-1} (1 + 2\mu (\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)) \right. \\
&\quad \left. + \sigma_x^2 b_i^{-1} ((\mathbf{z}_{ij} - \mathbf{m}_i)^T \cdot \mathbf{B}_i \cdot (\mathbf{z}_{ij} - \mathbf{m}_i))) \right\} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2);
\end{aligned}$$

$$\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \mu} = \sum_{j=1}^{n_i} -2b_i^{-2} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \sigma_x^2 \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i + b_i^{-1} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu);$$

$$\begin{aligned}
\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \sigma_\delta^2} &= \sum_{j=1}^{n_i} -\sigma_x^4 \sigma_\delta^{-4} b_i^{-2} \left\{ 1 - 2(x_{ij} - \mu) [\mathbf{a}_i^T \mathbf{A}_i^{-1} (\mathbf{z}_{ij} - \mathbf{m}_i)] \right. \\
&\quad \left. + 2\sigma_x^2 b_i^{-1} (\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i) \right\} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i \\
&\quad - \sigma_x^2 \sigma_\delta^{-4} b_i^{-1} \left\{ (x_{ij} - \mu) - \sigma_x^2 b_i^{-1} [\mathbf{a}_i^T \mathbf{A}_i^{-1} (\mathbf{z}_{ij} - \mathbf{m}_i)] \right\} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu);
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \sigma_x^2} &= \sum_{j=1}^{n_i} -b_i^{-2} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i - 2b_i^{-3} \sigma_x^2 [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i \\
&\quad + b_i^{-2} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu);
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \boldsymbol{\beta}_i \partial \sigma_{e_i}^2} &= \sum_{j=1}^{n_i} \sigma_x^2 b_i^{-1} \left\{ -\mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right. \\
&\quad \left. + \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) - \sigma_x^2 b_i^{-1} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i \cdot \boldsymbol{\beta}_i^T \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right\} \\
&\quad - \mu \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) - \sigma_x^4 b_i^{-2} [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] \\
&\quad \left\{ 2\sigma_x^2 b_i^{-1} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i \cdot \boldsymbol{\beta}_i^T \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) - \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right\} \\
&\quad - \sigma_x^2 b_i^{-1} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \left\{ \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu) \right. \\
&\quad \left. - \sigma_x^2 b_i^{-1} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot [2\boldsymbol{\beta}_i (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T + (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu) \boldsymbol{\beta}_i^T] \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right\};
\end{aligned}$$

$$\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \mu \partial \mu} = - \sum_{i=1}^p \sum_{j=1}^{n_i} \left\{ (1 - b_i^{-1}) \sigma_x^{-2} \right\};$$

$$\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_\delta^2} = - \sum_{i=1}^p \sum_{j=1}^{n_i} b_i^{-1} \sigma_\delta^{-4} \left\{ (x_{ij} - \mu) - \sigma_x^2 b_i^{-1} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \right\};$$

$$\frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_x^2} = - \sum_{i=1}^p \sum_{j=1}^{n_i} (b_i^{-1} - b_i^{-2}) \sigma_x^{-2} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)];$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \mu \partial \sigma_{e_i}^2} &= \sum_{j=1}^{n_i} b_i^{-2} \sigma_x^2 [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \boldsymbol{\beta}_i^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \\ &\quad - b_i^{-1} (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i); \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_\delta^2} &= - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} \left\{ -\sigma_\delta^{-4} + (2 - \sigma_x^2 \sigma_\delta^{-2} b_i^{-1}) \sigma_\delta^{-6} \sigma_x^2 b_i^{-1} \right. \\ &\quad - 4 \sigma_x^2 \sigma_\delta^{-6} b_i^{-1} (1 - \sigma_x^2 \sigma_\delta^{-2} b_i^{-1}) (x_{ij} - \mu) [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \\ &\quad + 2 \sigma_x^4 \sigma_\delta^{-6} b_i^{-2} (1 - \sigma_x^2 \sigma_\delta^{-2} b_i^{-1}) [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] \\ &\quad \left. + 2 \sigma_\delta^{-6} (1 - \sigma_x^2 \sigma_\delta^{-2} b_i^{-1}) (x_{ij} - \mu)^2 \right\}; \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_x^2} &= - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} \left\{ -\sigma_\delta^{-4} b_i^{-2} - 2 b_i^{-3} \sigma_x^2 \sigma_\delta^{-4} [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] \right. \\ &\quad \left. + 2 b_i^{-2} \sigma_\delta^{-4} (x_{ij} - \mu) [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \right\}; \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \sigma_\delta^2 \partial \sigma_{e_i}^2} &= - \frac{1}{2} \sum_{j=1}^{n_i} \left\{ b_i^{-2} \sigma_\delta^{-4} \left\{ \sigma_x^4 \left[ 2(x_{ij} - \mu) [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \right. \right. \right. \\ &\quad \left. \left. - 2 \sigma_x^2 [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] b_i^{-1} - 1 \right] \boldsymbol{\beta}_i^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right\} \\ &\quad \left. - b_i^{-1} \sigma_\delta^{-4} (2 \sigma_x^2 (x_{ij} - \mu) - \sigma_x^2 b_i^{-1} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)]) \right. \\ &\quad \left. (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right\}; \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_x^2} &= - \frac{1}{2} \sum_{i=1}^p \sum_{j=1}^{n_i} \left\{ -\sigma_x^{-4} [1 - b_i^{-1} (2 - b_i^{-1})] \right. \\ &\quad \left. + 2 \sigma_x^{-2} b_i^{-2} (1 - b_i^{-1}) [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] \right\}; \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \sigma_x^2 \partial \sigma_{e_i}^2} &= \frac{1}{2} \sum_{j=1}^{n_i} \left\{ b_i^{-2} (1 + 2 b_i^{-1} \sigma_x^2 [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)]) \boldsymbol{\beta}_i^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right. \\ &\quad \left. - 2 b_i^{-2} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \right\}; \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{z}, \boldsymbol{\theta})}{\partial \sigma_{e_i}^2 \partial \sigma_{e_i}^2} &= \frac{1}{2} \sum_{j=1}^{n_i} \left\{ -2 \mathbf{D}(\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \cdot \mathbf{D}^{-3}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \right. \\ &\quad + 2 \sigma_x^2 b_i^{-1} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \\ &\quad - \sigma_x^2 b_i^{-1} [2 \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-3}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) - \sigma_x^2 b_i^{-1} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i \cdot \boldsymbol{\beta}_i^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i)] \\ &\quad - 2 \sigma_x^4 b_i^{-2} [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)] [\mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-3}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) \\ &\quad - \sigma_x^2 b_i^{-1} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i \cdot \boldsymbol{\beta}_i^T \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i)] \\ &\quad \left. + 4 \sigma_x^2 b_i^{-1} [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] (\mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-3}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \right. \end{aligned}$$

$$-\frac{1}{2}\sigma_x^2 b_i^{-1} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\boldsymbol{\beta}_i \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T + (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \cdot \boldsymbol{\beta}_i^T) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{D}(\boldsymbol{\beta}_i) + \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \};$$

onde  $\mathbf{a}_i$  e  $\mathbf{A}_i$  tal que  $\mathbf{m}_i = \mu \mathbf{a}_i$  e  $\mathbf{V}_i = \mathbf{A}_i + \sigma_x^2 \mathbf{a}_i \mathbf{a}_i^T$ , e seja,

$$\mathbf{a}_i = (1, \boldsymbol{\beta}_i^T)^T = (1, \beta_{1i}, \beta_{2i})^T \quad e \quad \mathbf{A}_i = \mathbf{D}(\sigma^2, \boldsymbol{\sigma}_{e_i}^2)^T = \mathbf{D}(\sigma^2, \sigma_{e_{1i}}^2, \sigma_{e_{2i}}^2)^T$$

$$b_i = 1 + \sigma_x^2 \mathbf{a}_i^T \mathbf{A}_i^{-1} \mathbf{a}_i = 1 + \sigma_x^2 \left[ \sigma^{-2} + \boldsymbol{\beta}_i^T \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i \right] \quad e \quad \mathbf{B}_i = \mathbf{A}_i^{-1} \mathbf{a}_i \mathbf{a}_i^T \mathbf{A}_i^{-1}$$

### C.5.2 Matriz delta para os dados de líquidos de bochecho

Seja

$$C_{ij} = \sigma_\delta^{-2} (x_{ij} - \mu) + ((\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \cdot \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i)$$

$$= [\mathbf{a}_i^T \cdot \mathbf{A}_i^{-1} \cdot (\mathbf{z}_{ij} - \mathbf{m}_i)] \quad e$$

$$G_{ij} = \sigma_\delta^{-4} (x_{ij} - \mu)^2 + 2\sigma_\delta^{-2} (x_{ij} - \mu) (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i$$

$$+ (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)^T \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \boldsymbol{\beta}_i \boldsymbol{\beta}_i^T \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)$$

$$= [(\mathbf{z}_{ij} - \mathbf{m}_i)^T \mathbf{B}_i (\mathbf{z}_{ij} - \mathbf{m}_i)],$$

a matriz  $\Delta$  necessária, para fazer a análise de influência local, se encontra abaixo para cada tipo de perturbação definida, considerando o modelo (1.11); **Ponderação de casos**

$$\Delta_{\boldsymbol{\beta}_{ij}}^T = (\mu + \sigma_x^2 b_i^{-1} C_{ij}) \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) - \sigma_x^2 b_i^{-1} (1 + \sigma_x^2 b_i^{-1} G_{ij} + \mu C_{ij}) \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i;$$

$$\Delta_{\mu ij} = b_i^{-1} C_{ij};$$

$$\Delta_{\sigma_\delta^2 ij} = -\frac{1}{2} \{ \sigma_\delta^{-2} [1 - \sigma_\delta^{-2} (\sigma_x^2 b_i^{-1} + (x_{ij} - \mu)^2)] + \sigma_x^2 \sigma_\delta^{-4} b_i^{-1} [2(x_{ij} - \mu) C_{ij} - \sigma_x^2 b_i^{-1} G_{ij}] \};$$

$$\Delta_{\sigma_x^2 ij} = -\frac{1}{2} \{ \sigma_x^{-2} (1 - b_i^{-1}) - b_i^{-2} G_{ij} \};$$

$$\Delta_{\boldsymbol{\sigma}_{e_i}^2 ij}^T = \frac{1}{2} \mathbf{D}(\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) - \sigma_x^2 b_i^{-1} C_{ij} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu)$$

$$+ \frac{1}{2} \sigma_x^2 b_i^{-1} \{ \sigma_x^2 b_i^{-1} G_{ij} + 1 \} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i - \frac{1}{2} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot \mathbf{1}_2.$$

**Perturbação na covariável**

$$\Delta_{\boldsymbol{\beta}_{ij}}^T = \sigma_x^2 \sigma_\delta^{-2} b_i^{-1} S_x \{ -2\sigma_x^2 b_i^{-1} C_{ij} \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i + \mathbf{D}^{-1}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - 2\boldsymbol{\beta}_i \mu) \};$$

$$\Delta_{\mu ij} = S_x \sigma_\delta^{-2} b_i^{-1};$$

$$\Delta_{\sigma_\delta^2 ij} = \sigma_\delta^{-4} S_x (1 - \sigma_\delta^{-2} \sigma_x^2 b_i^{-1}) [(x_{ij} - \mu) - \sigma_x^2 C_{ij} b_i^{-1}];$$

$$\Delta_{\sigma_x^2 ij} = \sigma_\delta^{-2} b_i^{-2} S_x C_{ij};$$

$$\Delta_{\boldsymbol{\sigma}_{e_i}^2 ij}^T = S_x \sigma_x^2 b_i^{-1} \sigma_\delta^{-2} \left\{ \sigma_x^2 b_i^{-1} C_{ij} \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot \boldsymbol{\beta}_i - \mathbf{D}(\boldsymbol{\beta}_i) \cdot \mathbf{D}^{-2}(\boldsymbol{\sigma}_{e_i}^2) \cdot (\mathbf{y}_{ij} - \boldsymbol{\beta}_i \mu) \right\}.$$

**Perturbação na variável resposta**

$$\begin{aligned}
\Delta_{\beta_{ij}}^T &= \sigma_x^2 b_i^{-1} (\beta_i^T \cdot D^{-1}(\sigma_{e_i}^2) \cdot S_{y_i}) (D^{-1}(\sigma_{e_i}^2) \cdot (y_{ij} - 2\beta_i \mu) - 2b_i^{-1} \sigma_x^2 C_{ij} D^{-1}(\sigma_{e_i}^2) \cdot \beta_i) \\
&\quad + (\sigma_x^2 b_i^{-1} C_{ij} + \mu) D^{-1}(\sigma_{e_i}^2) \cdot S_{y_i}; \\
\Delta_{\mu_{ij}} &= b_i^{-1} (\beta_i^T \cdot D^{-1}(\sigma_{e_i}^2) \cdot S_{y_i}); \\
\Delta_{\sigma_{\delta}^2 ij} &= \sigma_x^2 \sigma_{\delta}^{-4} b_i^{-1} (\sigma_x^2 b_i^{-1} C_{ij} - (x_{ij} - \mu)) (\beta_i^T \cdot D^{-1}(\sigma_{e_i}^2) \cdot S_{y_i}); \\
\Delta_{\sigma_x^2 ij} &= b_i^{-2} C_{ij} (\beta_i^T \cdot D^{-1}(\sigma_{e_i}^2) \cdot S_{y_i}); \\
\Delta_{\sigma_{e_i}^2 ij}^T &= D(S_{y_i}) \cdot D^{-2}(\sigma_{e_i}^2) \cdot (y_{ij} - \beta_i \mu) - \sigma_x^2 b_i^{-1} C_{ij} D(\beta_i) \cdot D^{-2}(\sigma_{e_i}^2) \cdot S_{y_i} \\
&\quad + \sigma_x^2 b_i^{-1} (\beta_i^T \cdot D^{-1}(\sigma_{e_i}^2) \cdot S_{y_i}) (\sigma_x^2 b_i^{-1} C_{ij} D(\beta_i) \cdot D^{-2}(\sigma_{e_i}^2) \cdot \beta_i \\
&\quad - D(\beta_i) \cdot D^{-2}(\sigma_{e_i}^2) \cdot (y_{ij} - \beta_i \mu)).
\end{aligned}$$

**Perturbação na variância do erro**

$$\begin{aligned}
\Delta_{\beta_{ij}}^T &= -\sigma_x^2 b_i^{-2} D^{-1}(\sigma_{e_i}^2) \beta_i + \mu D^{-1}(\sigma_{e_i}^2) (y_{ij} - \beta_i \mu) + \sigma_x^2 \left[ -\sigma_x^2 b_i^{-3} (2 + b_i) C_{ij}^2 D^{-1}(\sigma_{e_i}^2) \beta_i \right. \\
&\quad \left. + b_i^{-2} (1 + b_i) C_{ij} D^{-1}(\sigma_{e_i}^2) (y_{ij} - 2\beta_i \mu) \right]; \\
\Delta_{\mu_{ij}} &= b_i^{-2} C_{ij}; \\
\Delta_{\sigma_{\delta}^2 ij} &= -\frac{1}{2} \left\{ -\sigma_{\delta}^{-4} \sigma_x^2 b_i^{-2} - \sigma_{\delta}^{-4} [(x_{ij} - \mu)^2 + \sigma_x^2 b_i^{-3} C_{ij} (\sigma_x^2 C_{ij} (2 + b_i) \right. \\
&\quad \left. - 2b_i(b_i + 1)(x_{ij} - \mu))] \right\}; \\
\Delta_{\sigma_x^2 ij} &= -\frac{b_i^{-3}}{2} \left\{ \sigma_x^{-2} b_i (b_i - 1) - 2C_{ij}^2 \right\}; \\
\Delta_{\sigma_{e_i}^2 ij}^T &= -\frac{1}{2} \left\{ -\sigma_x^2 b_i^{-2} D(\beta_i) \cdot D^{-2}(\sigma_{e_i}^2) \cdot \beta_i - D(y_{ij} - \beta_i \mu) D^{-2}(\sigma_{e_i}^2) \cdot (y_{ij} - \beta_i \mu) \right. \\
&\quad \left. - \sigma_x^2 b_i^{-3} C_{ij} \left[ \sigma_x^2 (b_i + 2) C_{ij} D(\beta_i) \cdot D^{-2}(\sigma_{e_i}^2) \cdot \beta_i \right. \right. \\
&\quad \left. \left. - 2b_i(b_i + 1) D(\beta_i) \cdot D^{-2}(\sigma_{e_i}^2) \cdot (y_{ij} - \beta_i \mu) \right] \right\}.
\end{aligned}$$

As expressões dos esquemas de perturbações descritas acima, foram retiradas da dissertação de [Russo \(2006\)](#).

