
Diagnóstico de influência local no modelo de
calibração ultraestrutural com réplicas

Bruno Pinheiro de Andrade

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

BRUNO PINHEIRO DE ANDRADE

**DIAGNÓSTICO DE INFLUÊNCIA LOCAL NO MODELO DE CALIBRAÇÃO
ULTRAESTRUTURAL COM RÉPLICAS**

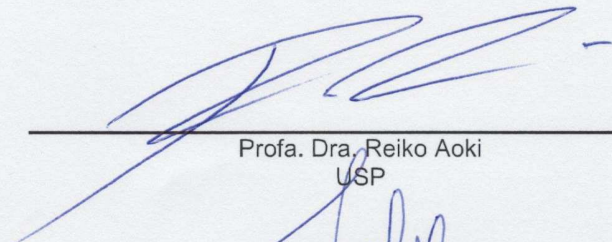
Tese apresentada ao Departamento de Estatística – Des/UFSCar e ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Estatística - Programa Interinstitucional de Pós-Graduação em Estatística UFSCar-USP.

Orientadora: Profa. Dra. Reiko Aoki

**São Carlos
Fevereiro de 2017**

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Bruno Pinheiro de Andrade, realizada em 09/12/2016:



Profa. Dra. Reiko Aoki
USP



Prof. Dr. Filidor Edilfonso Vilca Labra
UNICAMP



Profa. Dra. Roseli Aparecida Leandro
USP

Agradecimentos

A Deus e o Universo por sempre conspirarem ao meu favor.

A minhas mães Iseuda e Fátima, por terem sempre me dado muito amor, como também, a oportunidade de sonhar. Muito obrigado pelas incansáveis cobranças nos estudos, pois isso me ajudou a vencer e a superar os obstáculos.

Ao meu pai Isaias, pelo amor dedicado a mim.

À minha família, por estarem presentes em todos os momentos de minha vida, por sempre terem dado os melhores conselhos e por sempre acreditarem em mim. Esse exercito esteve comigo em todas as batalhas.

À minha namorada Renata, por compartilhar os momentos bons e ruins, por ter paciência, me incentivar, estudar e batalhar para o nosso sucesso profissional. Você é parte importante dessa conquista e sem você tudo teria sido mais difícil.

À professora Reiko Aoki, pela paciência e bom acompanhamento de todo o desenvolvimento do trabalho.

Aos professores Rafael Farias, Juvêncio e Ailton, pelos incentivos, conhecimentos técnicos e ensinamentos que vou levar para toda a vida

Aos professores Vera Tomazella e Mário de Castro, pela solicitude no âmbito acadêmico.

Aos professores que participaram da banca de qualificação e defesa, Cibele, Edwin, Roseli e Filidor pelas contribuições no trabalho.

À amizade de Eveliny e de William. Muito obrigado por minimizar minhas dificuldades em São Carlos, terei sempre um carinho especial por vocês.

Aos meus amigos do grupo de estudos composto por Aline, Juliana, Marcelo e Renata. Vocês contribuíram muito nessa caminhada e na conclusão do mestrado, obrigado pela força e apoio.

Ao grupo de estudos da UFC composto por Janaína, Kelly e Yuri. Vocês foram imprescindíveis para o meu sucesso. Lembro das nossas vitórias em cada desafio na graduação.

Agradeço a todos os meus amigos de Fortaleza-Ceará (Conjunto Mirassol/Veneza) e da Universidade Federal do Ceará, que sempre estiveram ao meu lado e torceram pelo meu sucesso.

Sumário

Agradecimentos	1
Lista de Figuras	5
Lista de Tabelas	7
Resumo	9
Abstract	11
1 Introdução	13
1.1 Revisão da literatura	15
1.1.1 Modelo de calibração comparativa	15
1.1.2 Diagnóstico de influência local	20
1.2 Formulação do problema	23
1.2.1 Modelagem sob distribuição normal	25
1.2.2 Modelagem sob distribuição t-Student	26
1.3 Organização do trabalho	27
2 Avaliação da influência local	28
2.1 Diagnóstico de influência local de Cook (1986)	29
2.2 Esquemas de perturbação no modelo normal	32
2.2.1 Ponderação de casos	33
2.2.2 Variável resposta	34
2.3 Esquemas de perturbação no modelo t-Student	40
2.3.1 Ponderação de casos	41

2.3.2 Variável resposta	41
3 Método para seleção apropriada da perturbação	45
3.1 Metodologia	46
3.2 Perturbação na variável resposta	47
3.2.1 Perturbação na variável resposta do laboratório de referência	49
3.2.2 Perturbação na variável resposta do i -ésimo laboratório	50
3.2.3 Perturbação nas variáveis respostas do j -ésimo patamar	50
4 Aplicação	52
4.1 Análise exploratória	52
4.2 Processo de modelagem normal e t-Student	56
4.2.1 Resultados do algoritmo EM para o modelo normal e t-Student	57
4.2.2 Avaliação da mudança relativa nas estimativas dos parâmetros.	59
4.3 Influência local de Cook (1986)	61
4.3.1 Perturbações sob o modelo normal	61
4.3.2 Perturbações sob o modelo t-Student	69
4.3.3 Conclusão	76
4.4 Método para a seleção apropriada	76
4.4.1 Perturbação na variável resposta do laboratório de referência	76
4.4.2 Perturbação na variável resposta do i -ésimo laboratório	78
4.4.3 Perturbação nas variáveis respostas do j -ésimo patamar	83
5 Conclusão	88
6 Propostas futuras	90
Referências Bibliográficas	91
Apêndice: A	98
Apêndice: B	103

Lista de Figuras

2.1	Curvatura normal para a superfície $\alpha(\omega)$ e direção unitária d (Verbeke & Molenberghs, 2000)	30
4.1	Boxplot da potência do motor por patamares dos laboratórios 2 a 6.	54
4.2	Dispersão das observações por patamares dos laboratórios.	55
4.3	Graus de liberdade versus AIC e BIC.	57
4.4	Ponderação de casos	61
4.5	Perturbação da variável resposta do laboratório de referência.	62
4.6	Perturbação da variável resposta do segundo laboratório.	63
4.7	Perturbação da variável resposta do terceiro laboratório.	63
4.8	Perturbação da variável resposta do quarto laboratório.	64
4.9	Perturbação da variável resposta do quinto laboratório.	64
4.10	Perturbação da variável resposta do sexto laboratório.	65
4.11	Perturbação na variável resposta no j -ésimo patamar.	69
4.12	Ponderação de casos	69
4.13	Perturbação na variável resposta do laboratório de referência	70
4.14	Gráficos de influência e $LD(w(a))$ quando a perturbação é no i -ésimo laboratório.	72
4.15	Gráficos de influência e $LD(w(a))$ quando a perturbação é no j -ésimo patamar.	76
4.16	Perturbação induzida da variável resposta do laboratório de referência.	77
4.17	Perturbação induzida da variável resposta do segundo laboratório.	79
4.18	Perturbação induzida da variável resposta do terceiro laboratório.	80
4.19	Perturbação induzida da variável resposta do quarto laboratório.	81
4.20	Perturbação induzida da variável resposta do quinto laboratório.	82

4.21	Perturbação induzida da variável resposta do sexto laboratório.	83
4.22	Perturbação em todos os laboratórios do patamar 1 após a correção adequada.	84
4.23	Perturbação em todos os laboratórios do patamar 2 após a correção adequada.	84
4.24	Perturbação em todos os laboratórios do patamar 3 após a correção adequada.	85
4.25	Perturbação em todos os laboratórios do patamar 4 após a correção adequada.	85
4.26	Perturbação em todos os laboratórios do patamar 5 após a correção adequada.	85
4.27	Perturbação em todos os laboratórios do patamar 6 após a correção adequada.	86
4.28	Perturbação em todos os laboratórios do patamar 7 após a correção adequada.	86
4.29	Perturbação em todos os laboratórios do patamar 8 após a correção adequada.	86
4.30	Perturbação em todos os laboratórios do patamar 9 após a correção adequada.	87
4.31	Perturbação em todos os laboratórios do patamar 10 após a correção adequada.	87

Lista de Tabelas

4.1	Resultados descritivos referente às observações do laboratório de referência por patamares.	53
4.2	Resultados descritivos referente às observações dos laboratórios restantes por patamares.	53
4.3	Estimativas de máxima verossimilhança do modelo normal.	58
4.4	Estimativas de máxima verossimilhança do modelo t-Student.	58
4.5	Mudança Relativa no modelo normal mediante retirada dos patamares. . . .	60
4.6	Mudança Relativa no modelo normal mediante retirada dos laboratórios. . .	60
4.7	Valores de C_{max} referente a perturbação em todos os laboratórios com o patamar fixo.	65
4.8	Valores de C_{max} referente a perturbação no i-ésimo laboratório.	70
4.9	C_{max} na perturbação em todos os laboratórios após o uso da correção apropriada.	72
4.10	C_{max} na perturbação em todos os laboratórios após o uso da correção apropriada.	84
6.1	Conjunto de dados usado na aplicação (dispostos na ordem Código (C), Faixa (F) e Potência Corrigida (PC))	103

Resumo

Este trabalho tem como proposta apresentar a metodologia de diagnóstico de influência local de [Cook \(1986\)](#) conjuntamente com a metodologia de seleção da perturbação adequada proposta por [Zhu et al. \(2007\)](#) no modelo de calibração ultraestrutural com réplicas. A metodologia de [Cook \(1986\)](#) será utilizada para investigar a robustez e a sensibilidade do modelo, onde os esquemas de perturbação adotados foram ponderação de casos e na variável resposta. Perturbar o modelo e/ou os dados de forma arbitrária pode conduzir a interpretações sobre a análise de diagnóstico e a conclusões equivocadas. Portanto, este trabalho irá avaliar as perturbações propostas segundo a metodologia de [Zhu et al. \(2007\)](#) e caso as perturbações não sejam adequadas, iremos propor uma nova forma de fazer as perturbações. Foi utilizado como aplicação a análise de um conjunto de dados com réplicas balanceadas e foram avaliadas quais patamares e laboratórios exercem um efeito desproporcional nas inferências feitas sob o modelo.

Palavras-chave: Modelo de calibração ultraestrutural, Avaliação da influência local de [Cook \(1986\)](#), Esquemas de perturbação, Seleção da perturbação adequada.

Abstract

This paper aims to present the local influence diagnostic methodology of [Cook \(1986\)](#) along with the selection of the appropriate perturbation schemes proposed by [Zhu et al. \(2007\)](#) in ultrastructural calibration model with replicas. The methodology of [Cook \(1986\)](#) will be used to investigate the robustness and sensitivity of model, where the adopted perturbation schemes were weighting cases and response variables. Perturbing the model and/or data in an arbitrary way can lead to miss interpretations of diagnostic analysis and wrong conclusions. Therefore, this study will evaluate the induced perturbations according to the methodology of [Zhu et al. \(2007\)](#) and if the perturbations are not suitable, we will propose a new way of perturbing the model or the data. As an application it was considered a data set with balanced repeated replication to evaluate which levels and laboratories exercise a disproportional effect on inferences made in the model.

Keywords: Ultrastructural calibration model, Assesment of local influence of [Cook \(1986\)](#), Perturbation schemes, Selection of the correct perturbation.

1

Introdução

O controle de qualidade é um procedimento comum nas empresas, pois visam garantir a qualidade dos produtos. Mas para que isso seja garantido, as medidas mensuradas pelo instrumento de medição, devem ser confiáveis para que as análises com as ferramentas de qualidade sejam corretas. Para tanto, a atenção com o instrumento de medição deve ser maior em virtude da importância de garantir o verdadeiro valor da medida. O instrumento de medição, na maioria dos casos, é um aparelho sensível que pode apresentar diversas fontes de variação que o conduzem a uma má medição. De modo geral, é possível dizer que os principais fatores responsáveis pela variabilidade associada ao processo de medição, segundo [Werkema \(1996\)](#), são: desgaste dos componentes do instrumento de medição, condições ambientais (temperatura, umidade, iluminação, poluição do ar, vibração, etc), treinamento insuficiente dos avaliadores e falta de calibração do aparelho de medição. Como visto, a qualidade da medição é algo relevante para garantir e assegurar uma boa qualidade dos dados, e caso ocorram medidas sem qualidade, haverá produtos sem qualidade e sem garantia de boa performance dos produtos que utilizam esses componentes.

O INMETRO (Instituto Nacional de Metrologia, Normalização e Qualidade industrial) e a sociedade desenvolveram diversos grupos de laboratórios que se reúnem periodicamente para avaliar e propor melhorias aos sistemas de medição. Um sistema de medição corresponde ao conjunto de equipamentos, métodos e pessoas, utilizados para obter os resultados da medição, onde entre as técnicas aplicadas, está o Ensaio de Proeficiência (EP). Os EP correspondem ao uso de comparações interlaboratoriais para fornecer a performance de laboratórios para a realização de medições específicas ou calibrações, de acordo com o [Guide43 \(1997\)](#) (ISO/IEC). O método de EP utilizado para gerar o conjunto de dados deste trabalho foi o esquema de comparação de medição, que consiste em um único item de teste distribuído

sequencialmente entre os laboratórios participantes. Cada laboratório envia o objeto para o próximo laboratório ou ao coordenador do estudo para a manutenção. Este procedimento é comum em comparações de padrões de calibração. Os resultados das medições juntamente com as incertezas padrões associadas, devem ser enviados ao coordenador.

Uma alternativa para comparar um instrumento de medição com um instrumento de referência, é usar o teste para a comparação de médias t-Student. Mas com esse teste apenas é possível concluir se um laboratório é significativamente igual ou não ao de referência, mas não é possível fazer comparações múltiplas. Desta forma, foi proposto o modelo com erro nas variáveis (modelo de calibração comparativa) que tem por objetivo relacionar medidas de diversos laboratórios e métodos para poder inferir a quantidade de vício que cada laboratório possui e através disso possibilitar melhorias para a calibração dos instrumentos. No artigo [Leão et al. \(2009\)](#), vemos o estudo de um modelo com erros nas variáveis, cujo proposta é testar a competência de laboratórios utilizando a classe de distribuições elípticas para modelar os dados e estabelecer testes estatísticos.

O modelo de calibração comparativa foi proposto por [Barnett \(1969\)](#), e tem por finalidade modelar experimentos que comparam quanto um instrumento de medição apresenta medidas distantes quando comparado com um instrumento de referência. Com isto, o modelo consegue quantificar quanto de vício aditivo, multiplicativo e ambos o instrumento está cometendo. Uma metodologia adequada para avaliar este tipo de problema é considerar a modelagem com erros nas variáveis ([Cheng & Van Ness, 1999](#)), que consiste em escrever as variáveis observadas do modelo como uma função de quantidades não observadas (medidas sem erro).

Neste trabalho iremos considerar uma extensão do modelo estrutural de [Barnett \(1969\)](#), o modelo de calibração ultraestrutural, que permite na modelagem a comparação de vários laboratórios com um de referência.

Uma das metodologias mais utilizadas para avaliar a robustez de um modelo estatístico, em modelos com erros de medição é o diagnóstico de influência local de [Cook \(1986\)](#), que consiste em introduzir uma pequena perturbação no modelo ou nos dados e através de gráficos avaliar se o modelo é robusto. Este método é baseado no afastamento da função de

verossimilhança. Esta metodologia será aplicada em um conjunto de dados reais que será modelado sob as distribuições Normal e t-Student multivariada, considerando a influência local dos casos de réplicas balanceada.

No caso em que as réplicas são desbalanceadas segundo [Giménez & Patat 2014](#) a técnica de influência local de [Cook \(1986\)](#) pode estar comprometida. Esta característica invalida o diagnóstico de influência convencional, pois tem como efeito mascarar as informações, em suma conduz a conclusões errôneas. Com o objetivo de resolver este problema [Zhu et al. \(2007\)](#) desenvolve em seu artigo os princípios básicos para selecionar o esquema de perturbação necessário. Esta questão será abordada ao longo do trabalho com mais detalhes no Capítulo 3.

1.1 Revisão da literatura

Nesta seção iremos abordar as principais referências bibliográficas dos estudos de modelos com erros nas variáveis, em específico aos modelos de calibração comparativa e aos estudos que utilizaram a metodologia de diagnóstico de [Cook \(1986\)](#) com foco principal nos modelos com erros nas variáveis.

1.1.1 Modelo de calibração comparativa

O modelo de calibração comparativa é muito importante para poder comparar diferentes laboratórios ou métodos de medição. Atualmente, os instrumentos de medição estão cada vez mais modernos, precisos, e os profissionais que realizam as medições estão cada vez mais capacitados. Logo, é necessário estudar a quantidade de vício que um instrumento tem quando comparado com um de referência e assim poder concluir se um método com um custo operacional menor é razoável, verificar se é hora de fazer manutenção no instrumento, ou a substituição por um equipamento moderno caso estejam produzindo um viés alto etc. Estes indicadores são importantes para todas as áreas, pois há cada vez mais necessidade de ter um instrumento de medição mais preciso com um baixo custo operacional.

Como motivação, segue adiante algumas referências que utilizaram modelos de calibração.

Grubbs (1973), por exemplo, compara três cronômetros, Barnett (1969) dá um exemplo onde quatro combinações de instrumento operador concebidos para medir a capacidade vital num grupo de pacientes são avaliados. Leurgans (1980) compara dois métodos para medir a concentração da glicose no sangue. Vários outros exemplos, na área médica, são mencionados por Kelly (1984), Kelly (1985) e Kaaks *et al.* (1994). Já Jaech (1985) dá vários exemplos na área industrial, Fuller (1987) apresenta exemplos na área agrônômica.

Como os modelos de calibração comparativa são uma classe dos modelos com erros nas variáveis, um estudo detalhado dos modelos com erros de medição é encontrado em Kendall & Stuart (1961), Moran (1971), Fuller (1987), Cheng & Van Ness (1999) e Carroll *et al.* (2006). Atualmente encontram-se na literatura estatística variadas formas de modelagem de regressão com erro nas variáveis, onde todos esses esforços abrangem os diversos casos que podem ocorrer na realidade. Os esforços citados englobam as modelagens funcionais, estruturais e ultraestruturais, lineares e não lineares, homocedásticos e heterocedásticos e entre outras formas generalizadas.

Como explicado anteriormente, o modelo com erros nas variáveis ocorre quando as variáveis estão sujeitas a erro de medição e o fato de não considerar isso na modelagem acarreta em resultados não confiáveis (Fuller, 1987). Para melhor compreensão, considere o exemplo com modelo de regressão linear simples com erros nas variáveis:

$$X_i = x_i + u_i \text{ e } Y_i = y_i + e_i, \quad i = 1, \dots, n, \quad (1.1)$$

onde Y_i e X_i são os valores observados de y_i e x_i , com erros e_i e u_i respectivamente. Assumindo a relação linear, temos:

$$y_i = \alpha + \beta x_i, \quad i = 1, \dots, n, \quad (1.2)$$

e portanto:

$$Y_i = \alpha + \beta x_i + e_i, \quad i = 1, \dots, n.$$

Se desconsiderar o erro em X_i e assumirmos:

$$Y_i = \alpha + \beta X_i + e_i, \quad i = 1, \dots, n,$$

o estimador de mínimos quadrados (EMQ) para β será inconsistente (Fuller, 1987).

Nestas condições podemos ter 3 modelos distintos. No primeiro caso a covariável não observada pode ser tratada como variável latente e assumindo uma distribuição de probabilidade, neste caso temos o modelo estrutural. O segundo caso considera a covariável como uma constante desconhecida, e é chamado de modelo funcional (Moran 1971 e Kendall & Stuart 1973). O terceiro caso considera variáveis latentes cujos parâmetros são incidentais, é o caso do modelo ultraestrutural (Dolby, 1976). O terceiro caso é uma forma generalizada dos modelos estrutural e funcional, e em ambos os casos devemos ter um cuidado especial, pois este tipo de modelagem pode possuir problemas como a falta de identificabilidade do modelo. O problema da falta da identificabilidade de um modelo é de suma importância em qualquer modelo paramétrico, pois um dos objetivos é estimar os parâmetros. Veja na tese de Galea-Rojas (1995) e em Cheng & Van Ness (1999), as definições e conceitos a cerca do problema da falta de identificabilidade de um modelo. Para retratar melhor essa questão, o exemplo a seguir será baseado no modelo (1.1) e (1.2).

Supondo, por exemplo, esta configuração: $x_i \sim N(\mu_x, \sigma_x^2)$, $u_i \sim N(0, \sigma_u^2)$ e $e_i \sim N(0, \sigma_e^2)$, temos que:

$$\begin{pmatrix} X_i \\ Y_i \end{pmatrix} \sim N_2 \left[\begin{pmatrix} \mu_x \\ \alpha + \beta\mu_x \end{pmatrix}; \begin{pmatrix} \sigma_x^2 + \sigma_u^2 & \beta\sigma_x^2 \\ \beta\sigma_x^2 & \beta^2\sigma_x^2 + \sigma_e^2 \end{pmatrix} \right], \quad i = 1, \dots, n.$$

Assim, o modelo $\mathbf{Z}_i = (X_i, Y_i)^\top \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, com

$$\boldsymbol{\mu} = (\mu_x, \alpha + \beta\mu_x)^\top$$

e

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_x^2 + \sigma_u^2 & \beta\sigma_x^2 \\ \beta\sigma_x^2 & \beta^2\sigma_x^2 + \sigma_e^2 \end{pmatrix},$$

onde o vetor de parâmetros é $\boldsymbol{\theta} = (\alpha, \beta, \mu_x, \sigma_x^2, \sigma_u^2, \sigma_e^2)$. O modelo normal bivariado é composto pelas seguintes estatísticas suficientes: $\sum_{i=1}^n y_i^2$, $\sum_{i=1}^n x_i^2$, $\sum_{i=1}^n y_i$, $\sum_{i=1}^n x_i$ e $\sum_{i=1}^n y_i x_i$. Note que a quantidade de parâmetros em $\boldsymbol{\theta}$ é maior do que a quantidade de estatísticas suficientes, logo o modelo é não identificável ou possui problema de identificabilidade, tendo como consequência um infinito número de soluções para os parâmetros.

Para contornar o problema da falta de identificabilidade a forma mais utilizada é assumir conhecido algum dos parâmetros. Na literatura é comum considerar o parâmetro λ conhecido, onde λ representa a razão das variâncias (Ripley & Thompson 1987, Riu & Rius 1996, Kulathinal [et al.](#) 2002, Galea-Rojas [et al.](#) 2003). Em alguns casos, pelas características naturais do modelo, considera-se α conhecido (Chan & Mak 1979 e Patefield 1978). Aoki [et al.](#) (2001), Aoki [et al.](#) (2002), Aoki [et al.](#) (2006), Vilca-Labra [et al.](#) (2005) e Russo [et al.](#) (2009) consideram o modelo de regressão com erros nas variáveis com intercepto nulo ($\alpha = 0$). Outras modelagens podem considerar que uma das variâncias é conhecida, ambas as variâncias são conhecidas ou o coeficiente de atenuação, $k_x = \frac{\sigma_u^2}{(\sigma_u^2 + \sigma_x^2)}$ é conhecida. Essas sugestões citadas podem ser utilizadas para resolver o problema de identificabilidade, e nenhuma suposição é feita a respeito do parâmetro β devido ao fato de ser um parâmetro de interesse primordial.

Cox (1976) estudou a relação linear estrutural, onde os dados eram de vários grupos, e testes de hipóteses foram feitos para testar o parâmetro β . Aoki [et al.](#) (2001) estudaram o modelo estrutural, a estimação foi feita pelo método de máxima verossimilhança usando o algoritmo Newton-Raphson. Russo (2006) considerou o modelo estrutural, onde as estimativas foram de máxima verossimilhança usando o algoritmo EM, e foram realizados também os testes de hipóteses de Wald, razão de máxima verossimilhança e score. Foi estudada a influência local em vários casos de perturbação. Patriota [et al.](#) (2009) estudaram o modelo estrutural heterocedástico, onde as estimativas foram obtidas pelos métodos de máxima verossimilhança e momentos. Vilca-Labra [et al.](#) (2011) estudaram o teste de hipóteses para o modelo de calibração estrutural. As estatísticas utilizadas foram de Wald, razão de verossimilhança e score. Castro & Galea-Rojas (2015) estudaram o modelo de calibração estrutural heterocedástico e testes de hipóteses foram realizados.

Kimura (1992) estudou o modelo de calibração comparativa funcional heterocedástico e usou o algoritmo EM para encontrar a aproximação das estimativas de máxima verossimilhança dos parâmetros do modelo. Giménez & Bolfarine (1997) e Giménez & Bolfarine (2000) fizeram a modelagem dos modelos de calibração comparativa funcional, onde as condições de regularidade foram estudadas e verificaram as propriedades de consistência e

normalidade assintótica. [Vilca-Labra et al. \(1998\)](#) estudaram o modelo funcional sob a classe de distribuições elípticas com ênfase na distribuição t-Student, e seus resultados englobam as propriedades de consistência e normalidade assintótica dos estimadores de máxima verossimilhança. [Rasekh \(2001\)](#) estudou a estimação do modelo com erros nas variáveis funcional, sua proposta era apresentar um estimador alternativo na presença de colineariedade entre os valores não observados, os resultados foram satisfatórios, pois apresentaram menor erro quadrático e menor erro padrão. [Giménez & Patat \(2005\)](#) estudaram as propriedades de consistência e normalidade assintótica dos estimadores do modelo de calibração comparativa funcional com réplicas. [Oliveira \(2011\)](#) estudou as inferências sob o modelo funcional com erros de medição heteroscedásticos na presença de réplicas das observações.

O modelo ultraestrutural foi proposto por [Dolby \(1976\)](#), neste trabalho Dolby apresenta uma síntese da relação funcional e estrutural nos casos com e sem réplicas. Neste artigo também foi considerado uma das variâncias conhecida, e o processo de estimação utilizado foi via máxima verossimilhança. [Patefield \(1978\)](#) analisou o artigo de Dolby e constatou que quando a variância é finita o estimador é inconsistente, ele ressalta que incorretamente Dolby propõe que o estimador da matriz de covariâncias assintóticas pode ser obtido pela matriz de informação de Fisher e a matriz de covariância assintótica correta e o estimador são sugeridos em seu artigo. [Gleser \(1985\)](#) escreveu uma nota a respeito do artigo de [Dolby \(1976\)](#) e concluiu que o ajuste do estimador de máxima verossimilhança de [Patefield \(1978\)](#) é fortemente consistente e assintoticamente normal. [Cheng & Van Ness \(1991\)](#) investigam o trabalho de [Dolby \(1976\)](#), e as principais referências que abordam este tema. [Arellano et al. \(1996\)](#) estudaram o modelo ultraestrutural elíptico. [Patriota et al. \(2011\)](#) verificaram as propriedades assintóticas do modelo ultraestrutural multivariado sob uma classe de distribuições elípticas. [Talarico \(2014\)](#) estuda o modelo de calibração ultraestrutural com réplicas desbalanceadas, propõe estimadores de máxima verossimilhança usando o algoritmo EM e faz testes de hipóteses assintóticos para os parâmetros.

Atualmente, pesquisadores tem se esforçado para estudar as configurações de situações reais dos modelos com erros nas variáveis. Por exemplo [Aoki et al. \(2006\)](#) estudaram um modelo com erros de medição multivariado considerando o intercepto nulo com misturas de

distribuições normais usando uma abordagem bayesiana. Os autores explicam que para o verdadeiro valor não observado da covariável, foi assumido uma mistura de duas distribuições normais. Testes de hipóteses foram estudados em, [Castro et al. \(2008\)](#) no modelo com erros nas variáveis heterocedástico e [Russo et al. \(2009\)](#) no modelo multivariado com intercepto nulo. [Castro & Galea-Rojas \(2010\)](#) propõem uma inferência robusta para a modelagem com erro nas variáveis heterocedástico, utilizando a distribuição t-Student. [Patriota \(2010\)](#) estuda a modelagem com erro nas variáveis para uma classe geral de distribuições. Ele estudou o modelo heterocedástico com erro nas variáveis, onde foram abordados a forma funcional e estrutural. [Borsoi \(2014\)](#) estudou os modelos mistos lineares elíptico com erros de medição sob a abordagem estrutural.

Em suma, esta sessão abordou as principais variantes a cerca da modelagem com erros nas variáveis. O apanhado de obras citadas dá uma visão geral do desenvolvimento técnico deste assunto ao longo do tempo.

1.1.2 Diagnóstico de influência local

A avaliação de diagnóstico de influência local, é uma proposta apresentada no artigo de [Cook \(1986\)](#), que visa mediante uma pequena perturbação, avaliar se existe observações que podem ser influentes e conseqüentemente causar um impacto nas estimativas dos parâmetros. Este método consiste em verificar o afastamento das verossimilhanças do modelo perturbado com o modelo sem perturbação, e a partir disso construir um gráfico de dispersão e identificar visualmente quais observações podem ser influentes. O método desenvolvido por Cook pode ser aplicado para uma grande variedade de modelos estatísticos, em particular para diversos modelos de regressão, porque o requisito básico é possuir a função de verossimilhança específica do modelo. A seguir algumas referências são citadas sobre a aplicação do diagnóstico de influência local de [Cook \(1986\)](#) em diversos tipos de modelagens considerando o erro nas variáveis.

Como este método requer o conhecimento da função de verossimilhança, ele se torna um método bastante comum em várias classes de modelagem estatística e principalmente em estudo com modelos com erros nas variáveis. [Thomas & Cook \(1989\)](#) desenvolveram o estudo

de influência local para os modelos lineares generalizados, e tinham o objetivo de verificar o efeito de uma pequena perturbação no modelo refletido nas estimativas dos parâmetros. Os primeiros artigos a avaliarem possíveis observações influentes nos modelos com erros nas variáveis foram propostos por [Kelly \(1984\)](#), no qual estuda a influência no modelo funcional e estrutural, e [Wellman & Gunst \(1991\)](#), que estuda a influência para o modelo linear com erros nas variáveis. Ambos se baseiam na referência de [Cook \(1977\)](#), que apresentou o estudo de diagnóstico de influência para o modelo de regressão linear.

[Zhao et al. \(1994\)](#), estudaram o diagnóstico de influência para os modelos lineares generalizados com erros nas variáveis, onde as verdadeiras covariáveis não são observadas, mas mensuradas com o erro. [Zhong et al. \(2000\)](#), estudaram a análise de influência local de [Cook \(1986\)](#) no modelo linear generalizado com erros de medida. A publicação apresentou os esquemas de perturbação de ponderação de casos, perturbação nas variáveis explicativas e perturbação na média.

[Carroll & Spiegelman \(1992\)](#) desenvolveram um estudo de diagnóstico baseado na análise de resíduos padronizados para o modelo com erros nas variáveis não lineares e heterocedástico. Vale salientar que o método de diagnóstico de análise de resíduos não é apropriado para a modelagem de dados com erros nas variáveis ver ([Fuller, 1987](#)). Neste artigo em questão foi considerado erro de medida na variável resposta correlacionado com o erro de medida das variáveis preditoras, surgindo um novo tipo de resíduo com essa configuração. [Zhao & Lee \(1995\)](#), estudaram diversos métodos de diagnóstico de influência para os modelos não lineares com erros nas variáveis, onde também foi utilizado o método de diagnóstico de influência local de [Cook \(1986\)](#), considerando a perturbação na variável resposta e covariáveis, ponderação de casos e perturbação na variância dos erros.

[Fung & Kwan \(1997\)](#) propõem outras funções como, um estimador de um parâmetro ou uma estatística de teste que podem ser usados para uma análise de influência local.

Novos enfoques para a investigação da influência local surgiram depois do diagnóstico proposto por [Cook \(1986\)](#). Todas essas novas propostas são baseadas em alguns princípios do diagnóstico de [Cook \(1986\)](#), e suas extensões e generalizações são descritas nos artigos de

Fung & Kwan (1997), Poon & Poon (1999), Loynes (2001), Suárez & Sierra (2001), Cadigan & Farrell (2002) e Zhu & Zhang (2004).

Rasekh & Fieller (2003) estudaram o diagnóstico do modelo com erros nas variáveis funcional com réplicas. O autor explica que para encontrar as estimativas dos parâmetros da regressão deve ser feito a suposição de que as variâncias dos erros são conhecidas e que as observações individuais, e as réplicas podem afetar as estimativas caso sejam influentes.

Vilca-Labra et al. (2005) desenvolveu um artigo de influência local onde estudou a influência no modelo de regressão t-Student com erro nas variáveis considerando o intercepto nulo ($\alpha = 0$). Eles tinham o objetivo de modelar um conjunto de dados que consiste em comparar duas escovas (experimental e convencional) e verificar por meio da técnica de influência local de Cook (1986) a robutez do modelo. Os esquemas de perturbação utilizados foram, ponderação de casos, perturbação na variável resposta, perturbação na variável explanatória e perturbação nos graus de liberdade.

Castro et al. (2007), estudaram o método de influência local de Cook (1986) no modelo com erros nas variáveis funcional heterocedástico. Eles tinham o objetivo de verificar o efeito das perturbações nas inferências em relação aos parâmetros e as perturbações utilizadas foram, ponderação de casos, perturbação tanto na variável resposta quanto na covariável observada na forma aditiva e multiplicativa.

Feng-Chang & Bo-Cheng (2009) estudaram a avaliação de diagnóstico de influência local de Cook (1986) para o modelo de regressão com erros nas variáveis com variável resposta Poisson. Eles investigaram a perturbação na variável resposta e ponderação de casos.

Galea-Rojas & Castro (2012), estudaram a influência local de Cook (1986) e Poon & Poon (1999) para o modelo com erros nas variáveis heterocedástico proposto por Kulathinal et al. (2002). Os esquemas de perturbação utilizados foram: ponderação de casos, perturbação da variável resposta e covariável, e perturbação da variância dos erros.

Zhu et al. (2007) estudaram o diagnóstico de influência local de Cook (1986) e a seleção do esquema de perturbação apropriada para acessar as observações que exercem uma influência desproporcional. Zhu descreve a metodologia necessária para verificar se a perturbação é adequada e caso não seja, ele explica qual a transformação necessária para tornar a perturbação

útil e assim estudar de forma correta a influência local. Baseado neste artigo citado, [Giménez & Galea-Rojas \(2013\)](#) estudaram o esquema de perturbação apropriado usando o método de influência local. [Giménez & Patat \(2014\)](#) explica que na modelagem considerando erros nas variáveis usando réplicas, acessar a observação influente através de um determinado esquema de perturbação pode ser um problema na presença de réplicas desbalanceadas. Neste artigo as autoras investigaram a influência local no modelo de calibração comparativa funcional com réplicas, sendo que este artigo aborda a seleção do esquema de perturbação adequado. Galea entende que o uso inadequado de esquemas de perturbação gera interpretações errôneas conduzindo o pesquisador a uma má conclusão do problema. Isso ocorre em virtude dos gráficos de influência local apontarem unidades influentes, quando na verdade não são ou omitirem pontos influentes, levando a concluir que o modelo é robusto. Mais detalhes a respeito da seleção do esquema de perturbação apropriado pode ser visto no Capítulo 3.

1.2 Formulação do problema

Este trabalho trata com a comparação de instrumentos ou métodos de medição, quando cada um deles é usado pra medir uma característica num grupo comum de unidades experimentais. Esse problema é conhecido na literatura estatística como calibração comparativa. O interesse é avaliar a qualidade de diferentes instrumentos de medição.

A motivação para utilizar modelos de calibração comparativa, surge da oportunidade de modelar um conjunto de dado real com réplicas balanceadas avaliadas pelo INMETRO. A empresa de motores *GM powertrain*, cedeu um artefato de motor a gasolina 1.0 ao INMETRO que repassou o mesmo exemplar para p laboratórios. Dentre estes laboratórios temos a Marelli, Delphi, Mahle, KSPG, Bosch e Mauá. Um desses laboratórios é um laboratório de referência credenciado pela INMETRO. Os outros $p - 1$ laboratórios terão a missão de apresentar uma boa compatibilidade quando comparado com o laboratório de referência, isto é, para serem bem avaliadas, terão que apresentar medidas próximas do real e isto ocorre quando o coeficiente do modelo é $\alpha = 0$ e $\beta = 1$, isso implica que o laboratório avaliado não apresenta medidas com vícios.

A potência do motor foi medida em m patamares de rotação pelos instrumentos de cada laboratório, assim temos que y_{ijk} é a k -ésima mensuração do verdadeiro valor da potência do motor no j -ésimo ponto de rotação medido pelo i -ésimo laboratório e x_j o verdadeiro valor não observável da potência do motor no j -ésimo ponto de rotação, $i = 1, \dots, p$, $j = 1, \dots, m$ e $k = 1, \dots, n_i$. Para o nosso conjunto de dados (Apêndice B, Tabela 6.1), foi assumido que y_{ijk} satisfaz a relação linear estrutural com o verdadeiro valor (não observável) x_j e denotando por Y_{ijk} o valor observado (sujeito a erro de medição) da k -ésima medição da potência do motor no j -ésimo ponto de rotação obtido pelo i -ésimo laboratório, o modelo proposto pode ser representado como

$$Y_{ijk} = y_{ijk} + e_{ijk},$$

$$y_{ijk} = \alpha_i + \beta_i x_j, \quad (1.3)$$

$i = 1, \dots, p$, $j = 1, \dots, m$ e $k = 1, \dots, n_i$ em que $E(e_{ijk}) = 0$, $V(e_{ijk}) = \sigma_{ij}^2$, $E(x_j) = \mu_{x_j}$ e $V(x_j) = \sigma_{x_j}^2$, x_j independente de e_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, m$ e $k = 1, \dots, n_i$. É importante ressaltar que esse conjunto de dados possui variâncias conhecidas, pois as incertezas das medições foram encaminhadas pelos laboratórios participantes.

A proposta deste modelo é comparar os vícios aditivos (α_i) e os vícios multiplicativos (β_i) de cada um dos laboratórios participantes tomando como medida padrão ($\alpha = 0$ e $\beta = 1$). [Barnett \(1969\)](#) assume um instrumento padrão para seu modelo que mede a covariável sem vícios. Aqui será considerado o primeiro laboratório como o laboratório padrão. Desta forma, tem-se que:

$$Y_{1jk} = x_j + e_{1jk} \quad \text{e} \quad Y_{ijk} = \alpha_i + \beta_i x_j + e_{ijk}, \quad (1.4)$$

$i = 2, \dots, p$, $j = 1, \dots, m$, e $k = 1, \dots, n_i$.

[Talarico \(2014\)](#) desenvolveu a modelagem deste conjunto de dados considerando as distribuições Normal e t-Student para efeito de comparação.

A distribuição t-Student tem forma geral simétrica e semelhante à forma da distribuição normal, com a vantagem de possuir caudas mais pesadas. O parâmetro γ representa o número de graus de liberdade. Quanto maior for o grau de liberdade, maior será a aproximação para a distribuição normal.

O modelo t-Student pode gerar estimativas mais robustas na presença de observações incomuns, além de ter uma melhor acomodação dos dados com as caudas mais pesadas do que as da distribuição normal. Com isto, podemos minimizar a influência dos outliers nos dados.

A seguir, descreveremos brevemente os dois modelos e para maiores detalhes ver a dissertação mencionada.

1.2.1 Modelagem sob distribuição normal

Assumindo o modelo definido em (1.4), onde $e_{ijk} \stackrel{ind}{\sim} N(0, \sigma_{ij}^2)$, $x_j \stackrel{ind}{\sim} N(\mu_{x_j}, \sigma_{x_j}^2)$, x_j independente de e_{ijk} , $i = 1, \dots, p$, $j = 1, \dots, m$ e $k = 1, \dots, n_i$, temos:

$$\mathbf{f}_{\mathbf{y}_j}(\mathbf{y}_j) = \frac{1}{(2\pi)^{\frac{n}{2}}} |\Sigma_j|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{y}_j - \boldsymbol{\mu}_j)^t \Sigma_j^{-1} (\mathbf{y}_j - \boldsymbol{\mu}_j) \right], \quad (1.5)$$

onde $\mathbf{Y}_j = (\mathbf{Y}_{1j}^\top, \dots, \mathbf{Y}_{pj}^\top)^\top$, com $\mathbf{Y}_{1j} = (Y_{1j1}, \dots, Y_{1jn_1})^\top$ e $\mathbf{Y}_{ij} = (Y_{ij1}, \dots, Y_{ijn_i})^\top$, $|\Sigma_j| = a_j |\mathbf{D}(\boldsymbol{\sigma}_j^2)|$, com $a_j = 1 + \boldsymbol{\sigma}_{x_j}^2 \boldsymbol{\beta}^t \mathbf{D}^{-1}(\boldsymbol{\sigma}_j^2) \boldsymbol{\beta}$ e $\boldsymbol{\sigma}_j^2 = (\sigma_{1j}^2 \mathbf{1}_{n_1}^\top, \dots, \sigma_{pj}^2 \mathbf{1}_{n_p}^\top)^\top$. Tem-se também que $n = \sum_{i=1}^p n_i$, $\boldsymbol{\mu}_j = (\boldsymbol{\mu}_{1j}^\top, \dots, \boldsymbol{\mu}_{pj}^\top)^\top = \boldsymbol{\alpha} + \mu_{x_j} \boldsymbol{\beta}$, com $\boldsymbol{\alpha} = (\mathbf{0}_{n_1}^\top, \alpha_2 \mathbf{1}_{n_2}^\top, \dots, \alpha_p \mathbf{1}_{n_p}^\top)^\top$, $\boldsymbol{\beta} = (\mathbf{1}_{n_1}^\top, \beta_2 \mathbf{1}_{n_2}^\top, \dots, \beta_p \mathbf{1}_{n_p}^\top)^\top$, $\mathbf{0}_{n_1}$ expressando um vetor composto por n_1 zeros, $\mathbf{1}_{n_i}$ expressando um vetor composto por n_i uns, $\Sigma_j = \mathbf{D}(\boldsymbol{\sigma}_j^2) + \sigma_{x_j}^2 \boldsymbol{\beta} \boldsymbol{\beta}^\top$ e $\mathbf{D}(\mathbf{a})$ denotando uma matriz diagonal com os elementos da diagonal dados pelos elementos do vetor \mathbf{a} .

Note que \mathbf{Y}_{1j} corresponde as medições da potência do motor pelo laboratório de referência no j -ésimo ponto de rotação e \mathbf{Y}_{ij} as medições do i -ésimo laboratório no j -ésimo ponto de rotação.

Chamaremos a forma quadrática $((\mathbf{y}_j - \boldsymbol{\mu}_j)^\top \Sigma_j^{-1} (\mathbf{y}_j - \boldsymbol{\mu}_j))$ de Q_j .

Seja $\mathbf{Y} = (\mathbf{Y}_1^\top, \dots, \mathbf{Y}_m^\top)^\top$. Da Equação (1.5), temos

$$f_{\mathbf{y}}(\mathbf{y}) = \prod_{j=1}^m f_{\mathbf{y}_j}(\mathbf{y}_j) = (2\pi)^{-\frac{mn}{2}} \left(\prod_{j=1}^m |\Sigma_j|^{-\frac{1}{2}} \right) \exp \left[-\frac{1}{2} \sum_{j=1}^m (\mathbf{y}_j - \boldsymbol{\mu}_j)^\top \Sigma_j^{-1} (\mathbf{y}_j - \boldsymbol{\mu}_j) \right] \quad (1.6)$$

Assumindo o logaritmo da Equação (1.6), temos que a função log-verossimilhança é dada por:

$$L(\boldsymbol{\theta}) = \log [f_{\mathbf{y}}(\mathbf{y})] = -\frac{mn}{2} [\log 2\pi] - \frac{1}{2} \sum_{j=1}^m \log |\Sigma_j| - \frac{1}{2} \sum_{j=1}^m (\mathbf{y}_j - \boldsymbol{\mu}_j)^\top \Sigma_j^{-1} (\mathbf{y}_j - \boldsymbol{\mu}_j), \quad (1.7)$$

em que $\boldsymbol{\theta} = (\mu_{x1}, \dots, \mu_{xm}, \alpha_2, \dots, \alpha_p, \beta_2, \dots, \beta_p)$.

No apêndice A, se encontram a função escore e os resultados para a construção da matriz de informação de fisher observada. Para estimar os parâmetros deste modelo foi usado o algoritmo EM, e os resultados referentes aos estimadores encontram-se na dissertação de [Talarico \(2014\)](#).

1.2.2 Modelagem sob distribuição t-Student

Assumindo que $\mathbf{Y}_j \sim t_n(\boldsymbol{\mu}_j, \boldsymbol{\Lambda}_j, \gamma)$, em que $t_r(\boldsymbol{\mu}, \boldsymbol{\Lambda}, \gamma)$ denota uma distribuição t de Student r -variada com vetor de locação $\boldsymbol{\mu}$, matriz escala $\boldsymbol{\Lambda}$ e γ graus de liberdade e $\boldsymbol{\Lambda}_j = \frac{\gamma-2}{\gamma} \Sigma_j$, com $\boldsymbol{\mu}_j$, Σ_j e n como definidos no caso normal, temos que:

$$Var(\mathbf{Y}_j) = \frac{\gamma}{\gamma-2} \boldsymbol{\Lambda}_j = \Sigma_j.$$

Desta forma,

$$f_{\mathbf{Y}_j}(\mathbf{y}_j) = \frac{\Gamma\left(\frac{\gamma+n}{2}\right)}{\Gamma\left(\frac{\gamma}{2}\right) \gamma^{\frac{n}{2}} \pi^{\frac{n}{2}}} \left(\frac{\gamma-2}{\gamma}\right)^{-\frac{n}{2}} |\Sigma_j|^{-\frac{1}{2}} \left[\frac{\gamma-2 + (\mathbf{Y}_j - \boldsymbol{\mu}_j)^\top \Sigma_j^{-1} (\mathbf{Y}_j - \boldsymbol{\mu}_j)}{\gamma-2} \right]^{-\left(\frac{\gamma+n}{2}\right)}.$$

Assim,

$$L(\boldsymbol{\theta}) = \log f_{\mathbf{Y}}(\mathbf{y}) = \sum_{j=1}^m \log f_{\mathbf{Y}_j}(\mathbf{y}_j) = cte - \frac{1}{2} \sum_{j=1}^m \log |\Sigma_j| - \left(\frac{\gamma+n}{2}\right) \sum_{j=1}^m \log(\gamma-2 + Q_j), \quad (1.8)$$

onde, $\mathbf{Y} = (\mathbf{Y}_1^\top, \dots, \mathbf{Y}_m^\top)^\top$, Q_j , $\boldsymbol{\mu}_j$ e Σ_j como foi apresentado na Seção 1.2.1.

No apêndice A se encontram a função score e as quantidades que compõem a matriz de informação de Fisher. Para a modelagem foi utilizado o algoritmo do tipo EM, e para mais detalhes recomendo a leitura de [Talarico \(2014\)](#).

Considerando este conjunto de dados e o modelo definido em (1.4), será desenvolvido um estudo de diagnóstico de influência local ([Cook, 1986](#)), assim como a metodologia para verificar se a perturbação é adequada ([Zhu et al., 2007](#)).

1.3 Organização do trabalho

Este trabalho de dissertação está organizado em seis capítulos, cujo a descrição é apresentada a seguir.

No Capítulo 2 será apresentado o diagnóstico de influência local de [Cook \(1986\)](#), onde será abordado os esquemas de perturbação para as duas classes de modelagens citadas. No Capítulo 3 será apresentado o estudo de seleção de esquema de perturbação apropriado segundo a metodologia proposta por [Zhu et al. \(2007\)](#). No Capítulo 4 é feita a aplicação contendo uma análise descritiva, modelagem dos dados com réplicas balanceadas, investigação da robustez do modelo pela técnica de diagnóstico de [Cook \(1986\)](#) e análise para verificar se a perturbação foi adequada segundo [Zhu et al. \(2007\)](#). No Capítulo 5 encontram-se as conclusões do trabalho e no Capítulo 6 encontram-se as propostas futuras.

2

Avaliação da influência local

Nos estudos de diagnósticos existem várias maneiras de investigar a influência, as duas formas mais comuns são a global e local. A eliminação de casos é uma técnica para avaliar o efeito de uma observação sobre o processo de estimação e teste de hipóteses, onde este método é chamado de análise de influência global. Alternativamente ao diagnóstico de influência global, [Cook \(1986\)](#) propõe o diagnóstico pela ótica local, que será abordada no decorrer dessa seção.

O conceito de influência local foi proposto por [Cook \(1986\)](#) com o objetivo de avaliar a mudança nos resultados da análise quando é introduzido pequenas perturbações no modelo ou no conjunto de dados, que implica na avaliação da qualidade do modelo. Neste contexto pode-se adotar a perturbação no modelo (ponderação de casos), perturbação nas variáveis resposta e explicativa, e por fim perturbação na matriz de covariância, por exemplo. Feita a modelagem do conjunto de dados e estudo de influência local, o procedimento consiste em avaliar por meio de gráficos, quais pontos são influentes no modelo. Identificado os pontos, as observações deverão ser avaliadas e portanto tomado as medidas cabíveis. Muitos autores têm utilizado esta abordagem, que pode ser aplicada a qualquer modelo estatístico com função de verossimilhança conhecida. Na literatura, muitos trabalhos encontrados apresentam o estudo de diagnóstico de [Cook \(1986\)](#) em diversos modelos estatísticos. Na Seção 1.1.2 foi descrito algumas das principais referências em que estudaram a influência local de Cook na classe dos modelos com erros nas variáveis. Outros resultados que consideram a influência local de [Cook \(1986\)](#) em outras classes de modelagem, podem ser conferidas em [Escobar & Meeker \(1992\)](#), [Lesaffre & Verbeke \(1998\)](#), [Osório *et al.* \(2007\)](#) e [Nobre & Singer \(2007\)](#), por exemplo.

Neste capítulo será apresentado a metodologia referente a avaliação da influência local

proposta por Cook (1986) aplicada no modelo estatístico de calibração ultraestrutural com réplicas. Os esquemas de perturbações adotados foram ponderação de casos e na variável resposta, onde a perturbação na variável resposta foi definida de 3 formas: no laboratório de referência, no i -ésimo laboratório, $i = 2, \dots, p$ e em todos os laboratórios fixando um patamar j , $j = 1, \dots, m$. Estes esquemas de perturbação foram desenvolvidos para o modelo normal e t-Student. Neste estudo iremos realizar a modelagem citada em um conjunto de dados em que as réplicas são balanceadas. Recentemente Giménez & Patat (2014) em seu artigo, explica que a influência local de Cook (1986) aplicada no modelo de calibração com réplicas desbalanceadas, pode conduzir o estudo a conclusões errôneas em virtude de não utilizar a perturbação adequada. Para contornar este problema, Giménez & Patat utiliza a metodologia proposta por Zhu et al. (2007), onde é estudado a seleção da perturbação adequada em cada esquema de perturbação. Esta metodologia será estudada no capítulo seguinte.

2.1 Diagnóstico de influência local de Cook (1986)

Seja $L(\boldsymbol{\theta})$ a função de log verossimilhança correspondente ao modelo de calibração ultraestrutural sob as distribuições de probabilidade normal e t de Student, onde $\boldsymbol{\theta}_{p \times 1}$ é um vetor de parâmetros desconhecidos. A perturbação do modelo é introduzida através de um vetor:

$$\boldsymbol{\omega} = \boldsymbol{\omega}_{q \times 1}, \text{ onde } \boldsymbol{\omega} \in \Omega \subseteq \mathbb{R}^q, \quad (2.1)$$

Ω um aberto, e seja $L(\boldsymbol{\theta}|\boldsymbol{\omega})$ a função de log-verossimilhança do modelo perturbado. É assumido que existe um $\boldsymbol{\omega}_0 \in \Omega$ tal que $L(\boldsymbol{\theta}) = L(\boldsymbol{\theta}|\boldsymbol{\omega}_0)$, $\forall \boldsymbol{\theta}$ e que $L(\boldsymbol{\theta}|\boldsymbol{\omega})$ é duas vezes continuamente diferenciável em $(\boldsymbol{\theta}^\top, \boldsymbol{\omega}^\top)^\top$. Seja $\hat{\boldsymbol{\theta}}$ e $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ os estimadores de máxima verossimilhança de $\boldsymbol{\theta}$ em $L(\boldsymbol{\theta})$ e $L(\boldsymbol{\theta}|\boldsymbol{\omega})$ respectivamente. Uma medida sugerida por Cook (1986) é dada na Equação (2.2):

$$LD = LD(\boldsymbol{\omega}) = 2 \left[L(\hat{\boldsymbol{\theta}}) - L(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}) \right], \quad (2.2)$$

e denominado o afastamento pela verossimilhança (“likelihood displacement”).

A ideia de influência local é estudar o comportamento da função $LD(\boldsymbol{\omega})$ numa vizinhança de $\boldsymbol{\omega}_0$. Para tanto, considera-se a superfície geométrica $(q + 1)$ -dimensional formada pelos valores do vetor

$$\alpha(\boldsymbol{\omega}) = \begin{bmatrix} \boldsymbol{\omega} \\ LD(\boldsymbol{\omega}) \end{bmatrix},$$

quando $\boldsymbol{\omega}$ varia em Ω , a superfície é denominada como gráfico de influência. O estudo de influência local consiste em analisar como a superfície $\alpha(\boldsymbol{\omega})$ desvia-se de seu plano tangente T_0 em $\boldsymbol{\omega}_0$. Essa análise pode ser feita estudando-se as curvaturas das seções normais da superfície $\alpha(\boldsymbol{\omega})$ em $\boldsymbol{\omega}_0$ que são intersecções de $\alpha(\boldsymbol{\omega})$ com planos contendo o vetor normal com seu plano tangente em $\boldsymbol{\omega}_0$. As curvaturas dessas seções são denominadas curvaturas normais, como no exemplo da Figura 2.1.

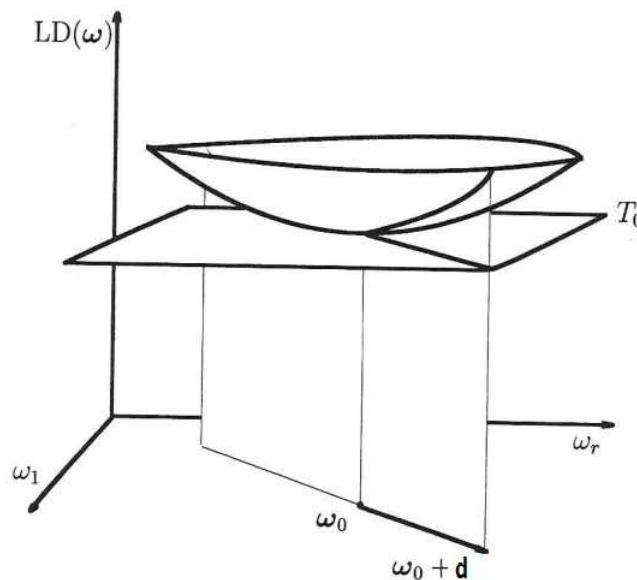


Figura 2.1 Curvatura normal para a superfície $\alpha(\boldsymbol{\omega})$ e direção unitária \mathbf{d} (Verbeke & Molenberghs, 2000)

De acordo com Cook (1986), uma forma de investigar o comportamento local de $LD(\boldsymbol{\omega})$ em torno de $\boldsymbol{\omega}_0$ é selecionar uma direção unitária \mathbf{d} e analisar o gráfico de $LD(\boldsymbol{\omega}_0 + a\mathbf{d})$ que tem um mínimo local em $a = 0$. Cada linha ajustada pode ser caracterizada considerando a curvatura normal C_d em $a = 0$, que é interpretada como a região de melhor ajuste em $\boldsymbol{\omega}_0$. Quanto maior o valor de C_d , maior é a sensibilidade a pequenas perturbações introduzidas na

direção \mathbf{d} . Cook (1986) sugere que a direção \mathbf{d}_{\max} , que corresponde à maior curvatura C_{\max} contém a informação de diagnóstico mais importante, ou seja, para detectar observações que mais influenciam em $LD(\boldsymbol{\omega})$, deve-se construir o gráfico de \mathbf{d}_{\max} pelo índice das observações e verificar quais pontos se destacam dos demais. Cook (1986) mostrou que:

$$C_d = 2|\mathbf{d}^\top \ddot{\mathbf{F}} \mathbf{d}|, \quad (2.3)$$

onde $\ddot{\mathbf{F}} = \boldsymbol{\Delta}^\top (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta}$, com $\boldsymbol{\Delta} = \frac{\partial^2}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}^\top} L(\boldsymbol{\theta} | \boldsymbol{\omega})$ avaliado em $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ e $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ e $-\ddot{\mathbf{L}}$ é a matriz de informação observada do modelo postulado. Desta forma, temos:

$$C_d = 2|\mathbf{d}^\top \boldsymbol{\Delta}^\top (\ddot{\mathbf{L}})^{-1} \boldsymbol{\Delta} \mathbf{d}|. \quad (2.4)$$

É de suma importância, após detectar os possíveis pontos influentes, investigar quais são as causas específicas que levaram estes pontos a produzir o efeito de influência desproporcional exercida no modelo ou nos dados.

Dependendo da aplicação o interesse pode ser somente em um sub vetor do vetor de parâmetros $\boldsymbol{\theta}$. Para acessar a influência em um sub vetor $\boldsymbol{\theta}_1$ de $\boldsymbol{\theta}$, sem perda de generalidade seja $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$. Particionando-se a matriz de informação de Fisher de acordo com a partição de $\boldsymbol{\theta}$, temos:

$$\ddot{\mathbf{L}} = \begin{pmatrix} \mathbf{L}_{11} & \mathbf{L}_{12} \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{pmatrix}.$$

Cook (1986) mostrou que:

$$C_d(\boldsymbol{\theta}_1) = 2|\mathbf{d}^\top \boldsymbol{\Delta}^\top (\ddot{\mathbf{L}}^{-1} - \mathbf{B}_{22}) \boldsymbol{\Delta} \mathbf{d}|, \quad (2.5)$$

onde,

$$\mathbf{B}_{22} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{L}_{22}^{-1} \end{pmatrix}.$$

A influência local para a i -ésima observação (C_i), segundo Lesaffre & Verbeke (1998),

é útil para construir o gráfico de influência local de C_i versus a i -ésima observação, como descrito a seguir:

$$C_i = 2|\mathbf{\Delta}_i^\top (\ddot{\mathbf{L}})^{-1} \mathbf{\Delta}_i|, \quad (2.6)$$

onde $\mathbf{\Delta}_i$ representa a i -ésima coluna da matriz $\mathbf{\Delta}$. A medida de corte estabelecida como critério foi valores acima de 2 vezes a média dos elementos de C_i , estes valores são classificados como influentes.

Apesar do alto crescimento em formas de abordar a influência local, as principais formas de perturbação continuam comum na maioria, como é apresentado a seguir:

- Ponderação de casos, $L(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \omega_i L_i(\boldsymbol{\theta})$, $0 \leq \omega_i \leq 1$ e L_i é a função de log-verossimilhança para a i -ésima observação;
- Perturbação na variável resposta, $y_{iw} = y_i + s_{y_i} \omega_i$, $\omega_i \in \mathbb{R}$;
- Perturbação na variável explicativa, $x_{iw} = x_i + s_{x_i} \omega_i$, $\omega_i \in \mathbb{R}$;
- Perturbação na matriz de variância-covariância, $\Sigma_{iw} = w_i^{-1} \Sigma_i$, $\omega_i \in \mathbb{R}^-$.

Nas seções 2.2 e 2.3 serão apresentados os esquemas de perturbação utilizado para aplicação de influência local e seus respectivos resultados que irão compor a matriz de perturbação $\mathbf{\Delta}$.

Novos enfoques de influência local surgiram após o artigo do Cook, e uma grande variedade de formas para acessar a influência local pode ser vista em [Fung & Kwan \(1997\)](#), [Poon & Poon \(1999\)](#), [Loynes \(2001\)](#), [Suárez & Sierra \(2001\)](#), [Cadigan & Farrell \(2002\)](#) e [Zhu & Zhang \(2004\)](#).

2.2 Esquemas de perturbação no modelo normal

Para investigar se o modelo proposto é robusto, foram propostas dois tipos de esquema de perturbação que são: ponderação de casos e perturbação na variável resposta. Na perturbação que envolve a variável resposta, o esquema de perturbação foi dividido em três

casos: no laboratório de referência, no i -ésimo laboratório e em todos os laboratórios do j -ésimo patamar para $j = 1, \dots, m$. A matriz de perturbação Δ foi particionada da seguinte forma:

$$\Delta = \frac{\partial^2 L(\boldsymbol{\theta}/\mathbf{w})}{\partial \boldsymbol{\theta} \partial \mathbf{w}^\top} = \begin{pmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 L(\boldsymbol{\theta}/\mathbf{w})}{\partial \boldsymbol{\theta}_1 \partial \mathbf{w}^\top} \\ \frac{\partial^2 L(\boldsymbol{\theta}/\mathbf{w})}{\partial \boldsymbol{\theta}_2 \partial \mathbf{w}^\top} \\ \frac{\partial^2 L(\boldsymbol{\theta}/\mathbf{w})}{\partial \boldsymbol{\theta}_3 \partial \mathbf{w}^\top} \end{pmatrix}, \quad (2.7)$$

onde $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top, \boldsymbol{\theta}_3^\top)^\top$, o vetor de parâmetros, foi subdivididos em $\boldsymbol{\theta}_1 = (\mu_{x_1}, \dots, \mu_{x_m})^\top$, $\boldsymbol{\theta}_2 = (\alpha_2, \dots, \alpha_p)^\top$, $\boldsymbol{\theta}_3 = (\beta_2, \dots, \beta_p)^\top$ e \mathbf{w} é o vetor de perturbação que será definido de acordo com cada esquema de perturbação utilizado.

2.2.1 Ponderação de casos

No esquema de ponderação de casos, a perturbação é feita em cada patamar e é expressa como na Equação (2.8):

$$L(\boldsymbol{\theta}|\mathbf{w}) = \sum_{j=1}^m L_j(\boldsymbol{\theta}|w_j) = \sum_{j=1}^m w_j L_j(\boldsymbol{\theta}) = \sum_{j=1}^m w_j \ln[f_{\mathbf{Y}_j}(\mathbf{Y}_j)], \quad (2.8)$$

onde, $L(\boldsymbol{\theta}|\mathbf{w})$ é a log-verossimilhança perturbada, $\mathbf{Y} = (\mathbf{Y}_1^\top, \dots, \mathbf{Y}_m^\top)^\top$, $\mathbf{w} = (w_1, \dots, w_m)^\top$ e $\boldsymbol{\omega}_0 = (1, \dots, 1)^\top$.

As expressões das derivadas, após manipulações algébricas, são explicitadas a seguir:

$$\begin{aligned} \Delta_{1_{jj}} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \mu_{x_j} \partial w_j} \right|_{\mathbf{w}=\boldsymbol{\omega}_0} = \frac{M_j}{a_j} - \frac{\mu_{x_j}}{\sigma_{x_j}^2}, \quad j = 1, \dots, m, \\ \Delta_{1_{hj}} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \mu_{x_h} \partial w_j} \right|_{\mathbf{w}=\boldsymbol{\omega}_0} = 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j, \\ \Delta_{2_{ij}} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \alpha_i \partial w_j} \right|_{\mathbf{w}=\boldsymbol{\omega}_0} = -\frac{1}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right\}, \quad j = 1, \dots, m \text{ e } i = \\ &2, \dots, p, \\ \Delta_{3_{ij}} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \beta_i \partial w_j} \right|_{\mathbf{w}=\boldsymbol{\omega}_0} = -\frac{\sigma_{x_j}^2}{\sigma_{ij}^2 a_j} \left\{ n_i \beta_i + M_j \left(\frac{n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right) \right\}, \quad j = \\ &1, \dots, m \text{ e } i = 2, \dots, p, \end{aligned}$$

$$\text{onde, } M_j = \frac{\mu_{xj}}{\sigma_{xj}^2} + \sum_{k=1}^{n_1} \frac{Y_{1jk}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i}{\sigma_{ij}^2} \left(\sum_{k=1}^{n_i} Y_{ijk} - \alpha_i n_i \right) \text{ e } a_j = 1 + \sigma_{xj}^2 \left(\frac{n_1}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{n_i \beta_i^2}{\sigma_{ij}^2} \right).$$

Na matriz Δ a seguir é disposto cada elemento que a compõe:

$$\Delta = \begin{bmatrix} \Delta_{111} & \Delta_{112} & \cdots & \Delta_{11m} \\ \Delta_{121} & \Delta_{122} & \cdots & \Delta_{12m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1m1} & \Delta_{1m2} & \cdots & \Delta_{1mm} \\ \Delta_{221} & \Delta_{222} & \cdots & \Delta_{22m} \\ \Delta_{231} & \Delta_{232} & \cdots & \Delta_{23m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{2p1} & \Delta_{2p2} & \cdots & \Delta_{2pm} \\ \Delta_{321} & \Delta_{322} & \cdots & \Delta_{32m} \\ \Delta_{331} & \Delta_{332} & \cdots & \Delta_{33m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{3p1} & \Delta_{3p2} & \cdots & \Delta_{3pm} \end{bmatrix}_{m+2(p-1) \times m}.$$

2.2.2 Variável resposta

1. Perturbação na variável resposta do laboratório de referência (Y_{1jk}):

No esquema de perturbação na variável resposta do laboratório de referência, a perturbação é feita diretamente nas medidas do laboratório de referência e é dada como na Equação (2.9):

$$Y_{1jk_w} = Y_{1jk} + w_{1j} s_{1j}, \quad j = 1, \dots, m \text{ e } k = 1, \dots, n_i. \quad (2.9)$$

Portanto o vetor $\mathbf{Y}_w = (\mathbf{Y}_{1_w}^\top, \mathbf{Y}_2^\top, \dots, \mathbf{Y}_p^\top)^\top$ apenas considera a perturbação nas medidas do laboratório de referência como especificado no vetor $\mathbf{Y}_{1_w} = (\mathbf{Y}_{11_w}^\top, \dots, \mathbf{Y}_{1m_w}^\top)^\top$, o vetor de perturbação para este esquema é $\mathbf{w}_1 = (w_{11}^\top, \dots, w_{1m}^\top)^\top$ com $\mathbf{w}_0 = (0, \dots, 0)^\top$ e s_{1j} representa o desvio padrão amostral dos valores observados do laboratório de referência no j -ésimo patamar, $j = 1, \dots, m$. Especificado o esquema de perturbação e

os vetores em questão, a função de log-verossimilhança perturbada é apresentada na Equação (2.10).

$$L(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{j=1}^m \left(-\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}_j| - \frac{1}{2} Q_{jw} \right), \quad (2.10)$$

note que Q_{jw} é a matriz quadrática considerando a perturbação \boldsymbol{w} .

As expressões resultantes das derivadas, após manipulações algébricas, são apresentadas abaixo:

$$\Delta_{1jj} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_j} \partial w_{1j}} \Big|_{\boldsymbol{w}=\boldsymbol{w}_0} = \frac{s_{1j} n_1}{a_j \sigma_{1j}^2}, \quad j = 1, \dots, m,$$

$$\Delta_{1hj} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_h} \partial w_{1j}} \Big|_{\boldsymbol{w}=\boldsymbol{w}_0} = 0, \quad j, h = 1, \dots, m, \quad \text{para } h \neq j,$$

$$\Delta_{2ij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \alpha_i \partial w_{1j}} \Big|_{\boldsymbol{w}=\boldsymbol{w}_0} = -\frac{n_i \beta_i \sigma_{x_j}^2 s_{1j} n_1}{a_j \sigma_{1j}^2 \sigma_{ij}^2}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\Delta_{3ij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \beta_i \partial w_{1j}} \Big|_{\boldsymbol{w}=\boldsymbol{w}_0} = -\left[\frac{\sigma_{x_j}^2 s_{1j} n_1}{a_j \sigma_{1j}^2 \sigma_{ij}^2} \right] \left\{ \frac{2n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right\},$$

$j = 1, \dots, m$ e $i = 2, \dots, p$,

$$\text{onde, } M_j = \frac{\mu_{x_j}}{\sigma_{x_j}^2} + \sum_{k=1}^{n_1} \frac{Y_{1jkw}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i}{\sigma_{ij}^2} \left(\sum_{k=1}^{n_i} Y_{ijk} - \alpha_i n_i \right).$$

Na matriz Δ a seguir é disposto cada elemento que a compõe e desta forma Δ é usada para os cálculos da influência local no esquema de perturbação na variável resposta do laboratório de referência.

$$\Delta = \begin{bmatrix} \Delta_{111} & \Delta_{112} & \cdots & \Delta_{11m} \\ \Delta_{121} & \Delta_{122} & \cdots & \Delta_{12m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1m1} & \Delta_{1m2} & \cdots & \Delta_{1mm} \\ \Delta_{221} & \Delta_{222} & \cdots & \Delta_{22m} \\ \Delta_{231} & \Delta_{232} & \cdots & \Delta_{23m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{2p1} & \Delta_{2p2} & \cdots & \Delta_{2pm} \\ \Delta_{321} & \Delta_{322} & \cdots & \Delta_{32m} \\ \Delta_{331} & \Delta_{332} & \cdots & \Delta_{33m} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{3p1} & \Delta_{3p2} & \cdots & \Delta_{3pm} \end{bmatrix}_{m+2(p-1) \times m}.$$

2. Perturbação na variável resposta do i -ésimo laboratório (Y_{ijk}):

Esse esquema de perturbação é feito na variável resposta do i -ésimo laboratório, a perturbação é feita diretamente nas medidas do laboratório e é dada como na Equação (2.11):

$$Y_{ijk_w} = Y_{ijk} + w_{ij}s_{ij}, \quad i = 2, \dots, p, \quad j = 1, \dots, m \text{ e } k = 1, \dots, n_i. \quad (2.11)$$

Portanto o vetor $\mathbf{Y}_w = (\mathbf{Y}_1^\top, \mathbf{Y}_2^\top, \dots, \mathbf{Y}_{(i-1)}^\top, \mathbf{Y}_{i_w}^\top, \mathbf{Y}_{(i+1)}^\top, \dots, \mathbf{Y}_p^\top)^\top$, onde desta forma é apenas considerado a perturbação nas medidas do i -ésimo laboratório como especificado no vetor $\mathbf{Y}_{i_w} = (Y_{i1_w}^\top, \dots, Y_{im_w}^\top)^\top$, o vetor de perturbação para este esquema é $\mathbf{w}_i = (w_{i1}^\top, \dots, w_{im}^\top)^\top$ com $\mathbf{w}_0 = (0, \dots, 0)^\top$ e s_{ij} representa o desvio padrão amostral dos valores observados do i -ésimo laboratório no j -ésimo patamar, onde $i = 1, \dots, p$ e $j = 1, \dots, m$. Abaixo a função de log-verossimilhança perturbada mediante o esquema citado.

$$L(\boldsymbol{\theta}|\mathbf{w}) = \sum_{j=1}^m \left(-\frac{n}{2} \log(2\pi) - \frac{1}{2} \log |\boldsymbol{\Sigma}_j| - \frac{1}{2} Q_{j_w} \right), \quad (2.12)$$

Especificado o esquema de perturbação e os vetores em questão, a função de log-verossimilhança perturbada é apresentada na Equação (2.12) e Q_{j_w} representa a forma quadrática afetada pela perturbação ω . As expressões das derivadas, após manipulações algébricas, são explicitadas a seguir:

$$\Delta_{1_{jij}} = -\frac{1}{2} \frac{\partial^2 Q_{j_w}}{\partial \mu_{x_j} \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{\beta_i s_{ij} n_i}{a_j \sigma_{ij}^2}, \quad j = 1, \dots, m,$$

$$\Delta_{1_{hij}} = -\frac{1}{2} \frac{\partial^2 Q_{j_w}}{\partial \mu_{x_h} \partial w_{ij}} \Big|_{\omega=\omega_0} = 0, \quad j, h = 1, \dots, m, \quad \text{para } h \neq j,$$

$$\Delta_{2_{iij}} = -\frac{1}{2} \frac{\partial^2 Q_{j_w}}{\partial \alpha_i \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{n_i s_{ij}}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} - 1 \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\Delta_{2_{lij}} = -\frac{1}{2} \frac{\partial^2 Q_{j_w}}{\partial \alpha_l \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{n_l \beta_l \sigma_{x_j}^2 n_i \beta_i s_{ij}}{a_j \sigma_{lj}^2 \sigma_{ij}^2}, \quad l, i = 2, \dots, p, \quad \text{para } l \neq i,$$

$$\Delta_{3_{iij}} = -\frac{1}{2} \frac{\partial^2 Q_{j_w}}{\partial \beta_i \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{\sigma_{x_j}^2 n_i s_{ij}}{a_j \sigma_{ij}^2} \left\{ \frac{2n_i \beta_i^2 \sigma_{x_j}^2 M_j}{\sigma_{ij}^2 a_j} + \frac{\beta_i}{\sigma_{ij}^2} \left(n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right) - M_j \right\},$$

$j = 1, \dots, m$ e $i = 2, \dots, p$,

$$\Delta_{3_{lij}} = -\frac{1}{2} \frac{\partial^2 Q_{j_w}}{\partial \beta_l \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{\sigma_{x_j}^2 \beta_i s_{ij} n_i}{a_j \sigma_{lj}^2 \sigma_{ij}^2} \left\{ \frac{2M_j n_l \beta_l \sigma_{x_j}^2}{a_j} + n_l \alpha_l - \sum_{k=1}^{n_l} Y_{ljk} \right\},$$

$l, i = 2, \dots, p$, para $l \neq i$,

$$\text{onde, } M_j = \frac{\mu_{x_j}}{\sigma_{x_j}^2} + \sum_{k=1}^{n_1} \frac{Y_{1jk}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i}{\sigma_{ij}^2} \left(\sum_{k=1}^{n_i} Y_{ijk} - \alpha_i n_i \right).$$

A seguir note como fica disposto os elementos que compõe a matriz de perturbação neste caso adotado.

$$\Delta = \begin{bmatrix} \Delta_{1i1} & 0 & \cdots & 0 \\ 0 & \Delta_{12i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Delta_{1mim} \\ \Delta_{22i1} & \Delta_{22i2} & \cdots & \Delta_{22im} \\ \Delta_{23i1} & \Delta_{23i2} & \cdots & \Delta_{23im} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{2(i-1)i1} & \Delta_{2(i-1)i2} & \cdots & \Delta_{2(i-1)im} \\ \Delta_{2ii1} & \Delta_{2ii2} & \cdots & \Delta_{2iim} \\ \Delta_{2(i+1)i1} & \Delta_{2(i+1)i2} & \cdots & \Delta_{2(i+1)im} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{2pi1} & \Delta_{2pi2} & \cdots & \Delta_{2pim} \\ \Delta_{32i1} & \Delta_{32i2} & \cdots & \Delta_{32im} \\ \Delta_{33i1} & \Delta_{33i2} & \cdots & \Delta_{33im} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{3(i-1)i1} & \Delta_{3(i-1)i2} & \cdots & \Delta_{3(i-1)im} \\ \Delta_{3ii1} & \Delta_{3ii2} & \cdots & \Delta_{3iim} \\ \Delta_{3(i+1)i1} & \Delta_{3(i+1)i2} & \cdots & \Delta_{3(i+1)im} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{3pi1} & \Delta_{3pi2} & \cdots & \Delta_{3pim} \end{bmatrix}_{m+2(p-1) \times m}.$$

3. Perturbação na variável resposta de todos os laboratórios no j-ésimo patamar:

Esse esquema de perturbação é na variável resposta em todos os laboratórios, onde será considerado o j fixo. Neste caso, o esquema de perturbação é definido como em (2.13) e a função de log-verossimilhança é definida como em (2.14).

$$Y_{ijk_w} = Y_{ijk} + w_{ij}s_{ij}, \quad i = 1, \dots, p \text{ e } j = 1, \dots, m. \quad (2.13)$$

$$L(\boldsymbol{\theta} | \mathbf{w}_j) = \sum_{\substack{l=1 \\ l \neq j}}^m \ln f_{\mathbf{Y}_l}(\mathbf{Y}_l) + \ln f_{\mathbf{Y}_{j_w}}(\mathbf{Y}_{j_w}). \quad (2.14)$$

Portanto neste esquema o vetor $\mathbf{Y}_{jw} = (\mathbf{Y}_{1jw}^\top, \mathbf{Y}_{2jw}^\top, \dots, \mathbf{Y}_{pjw}^\top)^\top$. Desta forma é considerado a perturbação nas medidas de todos os laboratórios no j -ésimo patamar, o vetor de perturbação para este esquema é dado por $\mathbf{w}_j = (w_{1j}^\top, \dots, w_{pj}^\top)^\top$, $\mathbf{w}_0 = (0, \dots, 0)^\top$ e s_{ij} o desvio padrão amostral das medidas no i -ésimo laboratório e j -ésimo patamar. Especificado o esquema de perturbação e os vetores em questão, as respectivas derivadas para este caso seguem abaixo.

$$\Delta_{1j1j} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_j} \partial w_{1j}} \Big|_{\omega=\omega_0} = \frac{s_{1j} n_1}{a_j \sigma_{1j}^2}, \quad j = 1, \dots, m,$$

$$\Delta_{1jij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_j} \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{\beta_i s_{ij} n_i}{a_j \sigma_{ij}^2}, \quad i = 2, \dots, p \text{ e } j = 1, \dots, m,$$

$$\Delta_{1hj} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_h} \partial w_{ij}} \Big|_{\omega=\omega_0} = 0, \quad i = 2, \dots, p \text{ e } j, h = 1, \dots, m, \text{ para } h \neq j,$$

$$\Delta_{2i1j} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \alpha_i \partial w_{1j}} \Big|_{\omega=\omega_0} = -\frac{n_i \beta_i \sigma_{x_j}^2 s_{1j} n_1}{a_j \sigma_{1j}^2 \sigma_{ij}^2}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\Delta_{2iij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \alpha_i \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{n_i s_{ij}}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} - 1 \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\Delta_{2lij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \alpha_l \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{n_l \beta_l \sigma_{x_j}^2 n_i \beta_i s_{ij}}{a_j \sigma_{lj}^2 \sigma_{ij}^2}, \quad l, i = 2, \dots, p, \text{ para } l \neq i,$$

$$\Delta_{3i1j} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \beta_i \partial w_{1j}} \Big|_{\omega=\omega_0} = -\left[\frac{\sigma_{x_j}^2 s_{1j} n_1}{a_j \sigma_{1j}^2 \sigma_{ij}^2} \right] \left\{ \frac{2n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\Delta_{3iij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \beta_i \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{\sigma_{x_j}^2 n_i s_{ij}}{a_j \sigma_{ij}^2} \left\{ \frac{2n_i \beta_i^2 \sigma_{x_j}^2 M_j}{\sigma_{ij}^2 a_j} + \frac{\beta_i}{\sigma_{ij}^2} \left(n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right) - M_j \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\Delta_{3lij} = -\frac{1}{2} \frac{\partial^2 Q_{jw}}{\partial \beta_l \partial w_{ij}} \Big|_{\omega=\omega_0} = -\frac{\sigma_{x_j}^2 \beta_i s_{ij} n_i}{a_j \sigma_{lj}^2 \sigma_{ij}^2} \left\{ \frac{2M_j n_l \beta_l \sigma_{x_j}^2}{a_j} + n_l \alpha_l - \sum_{k=1}^{n_l} Y_{ljk} \right\}, \quad l, i = 2, \dots, p, \text{ para } l \neq i.$$

Vale salientar que as demais derivadas já foram apresentadas nos demais casos de perturbação e que a diferença está na forma em que é organizada os valores na matriz de perturbação Δ .

$$\Delta = \begin{bmatrix} \Delta_{111j} & \Delta_{112j} & \cdots & \Delta_{1_{(i-1)j}} & \Delta_{1_{ij}} & \Delta_{1_{(i+1)j}} & \cdots & \Delta_{1_{pj}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_{1_{(j-1)1j}} & \Delta_{1_{(j-1)2j}} & \cdots & \Delta_{1_{(j-1)(i-1)j}} & \Delta_{1_{(j-1)ij}} & \Delta_{1_{(j-1)(i+1)j}} & \cdots & \Delta_{1_{(j-1)pj}} \\ \Delta_{1_{j1j}} & \Delta_{1_{j2j}} & \cdots & \Delta_{1_{j(i-1)j}} & \Delta_{1_{jij}} & \Delta_{1_{j(i+1)j}} & \cdots & \Delta_{1_{jpj}} \\ \Delta_{1_{(j+1)1j}} & \Delta_{1_{(j+1)2j}} & \cdots & \Delta_{1_{(j+1)(i-1)j}} & \Delta_{1_{(j+1)ij}} & \Delta_{1_{(j+1)(i+1)j}} & \cdots & \Delta_{1_{(j+1)pj}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_{1_{m1j}} & \Delta_{1_{m2j}} & \cdots & \Delta_{1_{m(i-1)j}} & \Delta_{1_{mij}} & \Delta_{1_{m(i+1)j}} & \cdots & \Delta_{1_{mpj}} \\ \Delta_{2_{21j}} & \Delta_{2_{22j}} & \cdots & \Delta_{2_{2(i-1)j}} & \Delta_{2_{2ij}} & \Delta_{2_{2(i+1)j}} & \cdots & \Delta_{2_{2pj}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_{2_{(i-1)1j}} & \Delta_{2_{(i-1)2j}} & \cdots & \Delta_{2_{(i-1)(i-1)j}} & \Delta_{2_{(i-1)ij}} & \Delta_{2_{(i-1)(i+1)j}} & \cdots & \Delta_{2_{(i-1)pj}} \\ \Delta_{2_{i1j}} & \Delta_{2_{i2j}} & \cdots & \Delta_{2_{i(i-1)j}} & \Delta_{2_{ij}} & \Delta_{2_{i(i+1)j}} & \cdots & \Delta_{2_{ipj}} \\ \Delta_{2_{(i+1)1j}} & \Delta_{2_{(i+1)2j}} & \cdots & \Delta_{2_{(i+1)(i-1)j}} & \Delta_{2_{(i+1)ij}} & \Delta_{2_{(i+1)(i+1)j}} & \cdots & \Delta_{2_{(i+1)pj}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_{2_{p1j}} & \Delta_{2_{p2j}} & \cdots & \Delta_{2_{p(i-1)j}} & \Delta_{2_{pij}} & \Delta_{2_{p(i+1)j}} & \cdots & \Delta_{2_{ppj}} \\ \Delta_{3_{21j}} & \Delta_{3_{22j}} & \cdots & \Delta_{3_{2(i-1)j}} & \Delta_{3_{2ij}} & \Delta_{3_{2(i+1)j}} & \cdots & \Delta_{3_{2pj}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_{3_{(i-1)1j}} & \Delta_{3_{(i-1)2j}} & \cdots & \Delta_{3_{(i-1)(i-1)j}} & \Delta_{3_{(i-1)ij}} & \Delta_{3_{(i-1)(i+1)j}} & \cdots & \Delta_{3_{(i-1)pj}} \\ \Delta_{3_{i1j}} & \Delta_{3_{i2j}} & \cdots & \Delta_{3_{i(i-1)j}} & \Delta_{3_{ij}} & \Delta_{3_{i(i+1)j}} & \cdots & \Delta_{3_{ipj}} \\ \Delta_{3_{(i+1)1j}} & \Delta_{3_{(i+1)2j}} & \cdots & \Delta_{3_{(i+1)(i-1)j}} & \Delta_{3_{(i+1)ij}} & \Delta_{3_{(i+1)(i+1)j}} & \cdots & \Delta_{3_{(i+1)pj}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_{3_{p1j}} & \Delta_{3_{p2j}} & \cdots & \Delta_{3_{p(i-1)j}} & \Delta_{3_{pij}} & \Delta_{3_{p(i+1)j}} & \cdots & \Delta_{3_{ppj}} \end{bmatrix}$$

2.3 Esquemas de perturbação no modelo t-Student

Nesta seção, vamos considerar os mesmos esquemas propostos para a modelagem sob distribuição normal. A diferença está na função de verossimilhança utilizada, onde no caso em questão é sob distribuição t-Student. Na Equação (1.8), encontra-se a função de log-verossimilhança utilizada para o desenvolvimento do diagnóstico de influência local de [Cook \(1986\)](#). Desta forma, somente iremos citar as expressões resultantes das derivadas nos respectivos esquemas de perturbação adotado.

2.3.1 Ponderação de casos

Este esquema de perturbação é similar ao adotado no modelo normal como expresso em (2.8). Vale salientar que este esquema utiliza a função de log-verossimilhança da distribuição t-Student. A seguir encontram-se as derivadas:

$$\begin{aligned}\Delta_{1jj} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \mu_{x_j} \partial w_j} \right|_{\omega=\omega_0} = \frac{(\gamma+n)}{\gamma-2+Q_j} \left\{ \frac{M_j}{a_j} - \frac{\mu_{x_j}}{\sigma_{x_j}^2} \right\}, \quad j = 1, \dots, m, \\ \Delta_{1hj} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \mu_{x_h} \partial w_j} \right|_{\omega=\omega_0} = 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j, \\ \Delta_{2ij} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \alpha_i \partial w_j} \right|_{\omega=\omega_0} = -\frac{(\gamma+n)}{\gamma-2+Q_j} \frac{1}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Z_{ijk} \right\}, \quad j = 1, \dots, m \\ \text{e } i &= 2, \dots, p, \\ \Delta_{3ij} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}|\mathbf{w}_j)}{\partial \beta_i \partial w_j} \right|_{\omega=\omega_0} = -\frac{\sigma_{x_j}^2}{a_j \sigma_{ij}^2} \left(n_i \beta_i + \frac{(\gamma+n)M_j}{\gamma-2+Q_j} \left\{ \frac{n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right\} \right), \\ j &= 1, \dots, m \text{ e } i = 2, \dots, p.\end{aligned}$$

2.3.2 Variável resposta

1. Perturbação na variável resposta do laboratório de referência (Y_{1jk}):

Neste esquema são usados as mesmas condições utilizadas no caso normal, e a função de log-verossimilhança perturbada é do modelo t-Student.

$$\begin{aligned}\Delta_{1jj} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}/\omega)}{\partial \mu_{x_j} \partial \omega_{1j}} \right|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \mu_{x_j}} \frac{\partial Q_{j\omega}}{\partial \omega_{1j}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_j} \partial \omega_{1j}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{1hj} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}/\omega)}{\partial \mu_{x_h} \partial \omega_{1j}} \right|_{\omega=\omega_0} = 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j, \\ \Delta_{2ij} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}/\omega)}{\partial \alpha_i \partial \omega_{1j}} \right|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \alpha_i} \frac{\partial Q_{j\omega}}{\partial \omega_{1j}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \alpha_i \partial \omega_{1j}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{3ij} &= \left. \frac{\partial^2 L(\boldsymbol{\theta}/\omega)}{\partial \beta_i \partial \omega_{1j}} \right|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \beta_i} \frac{\partial Q_{j\omega}}{\partial \omega_{1j}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \beta_i \partial \omega_{1j}} \right) \Big|_{\omega=\omega_0}.\end{aligned}$$

As expressões referente às derivadas de $\frac{\partial Q_{j\omega}}{\partial \omega_{1j}}$ e $\frac{\partial^2 Q_{j\omega}}{\partial \boldsymbol{\theta} \partial \omega_{1j}}$, após manipulações algébricas, são encontradas a seguir, e as expressões referente às derivadas de $\frac{\partial Q_{j\omega}}{\partial \boldsymbol{\theta}}$ encontram-se no apêndice A, porém estes resultados já estão aplicados no ponto ω_0 .

$$\left. \frac{\partial Q_{j\omega}}{\partial \omega_{1j}} \right|_{\omega=\omega_0} = \left(\frac{2s_{1j}}{\sigma_{1j}^2} \right) \left\{ \sum_{k=1}^{n_1} Y_{1jk} - \frac{n_1 \sigma_{x_j}^2 M_j}{a_j} \right\}, \quad j = 1, \dots, m,$$

$$\begin{aligned} \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_j} \partial w_{1j}} \Big|_{\omega=\omega_0} &= -\frac{2s_{1j}n_1}{a_j\sigma_{1j}^2}, \quad j = 1, \dots, m, \\ \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_h} \partial w_{1j}} \Big|_{\omega=\omega_0} &= 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j, \\ \frac{\partial^2 Q_{j\omega}}{\partial \alpha_i \partial w_{1j}} \Big|_{\omega=\omega_0} &= \frac{2n_i\beta_i\sigma_{x_j}^2 s_{1j}n_1}{a_j\sigma_{1j}^2\sigma_{ij}^2}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \frac{\partial^2 Q_{j\omega}}{\partial \beta_i \partial w_{1j}} \Big|_{\omega=\omega_0} &= \left[\frac{\sigma_{x_j}^2 s_{1j}n_1}{a_j\sigma_{1j}^2\sigma_{ij}^2} \right] \left\{ \frac{2n_i\beta_i\sigma_{x_j}^2 M_j}{a_j} + n_i\alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right\}, \\ &j = 1, \dots, m \text{ e } i = 2, \dots, p. \end{aligned}$$

2. Perturbação na variável resposta do i-ésimo laboratório (Y_{ijk}):

Neste esquema são usados as mesmas condições utilizadas no caso normal, considerando a função de log-verossimilhança perturbada do modelo t-Student.

$$\begin{aligned} \Delta_{1jij} &= \frac{\partial^2 L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \mu_{x_j} \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \mu_{x_j}} \frac{\partial Q_{j\omega}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_j} \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{1hij} &= \frac{\partial^2 L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \mu_{x_h} \partial w_{ij}} \Big|_{\omega=\omega_0} = 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j. \\ \Delta_{2iij} &= \frac{\partial^2 L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha_i \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \alpha_i} \frac{\partial Q_{j\omega}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \alpha_i \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{2lij} &= \frac{\partial^2 L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha_l \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \alpha_l} \frac{\partial Q_{j\omega}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \alpha_l \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{3iij} &= \frac{\partial^2 L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_i \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \beta_i} \frac{\partial Q_{j\omega}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \beta_i \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{3lij} &= \frac{\partial^2 L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_l \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{j\omega})^2} \frac{\partial Q_{j\omega}}{\partial \beta_l} \frac{\partial Q_{j\omega}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{j\omega})} \frac{\partial^2 Q_{j\omega}}{\partial \beta_l \partial w_{ij}} \right) \Big|_{\omega=\omega_0}. \end{aligned}$$

As expressões referente às derivadas $\frac{\partial Q_{j\omega}}{\partial w_{ij}}$ e $\frac{\partial^2 Q_{j\omega}}{\partial \boldsymbol{\theta} \partial w_{ij}}$, após manipulações algébricas, encontram-se a seguir. As expressões referentes às derivadas em $\frac{\partial Q_{j\omega}}{\partial \boldsymbol{\theta}}$, são encontradas no apêndice A, porém estes resultados já estão aplicado no ponto ω_0 .

$$\begin{aligned} \frac{\partial Q_{j\omega}}{\partial w_{ij}} \Big|_{\omega=\omega_0} &= \left(\frac{2s_{ij}}{\sigma_{ij}^2} \right) \left\{ \sum_{k=1}^{n_i} Y_{ijk} - n_i\alpha_i - \frac{n_i\beta_i\sigma_{x_j}^2 M_j}{a_j} \right\}, \quad j = 1, \dots, m, \\ \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_j} \partial w_{ij}} \Big|_{\omega=\omega_0} &= -\frac{2n_i\beta_i s_{ij}}{a_j\sigma_{ij}^2}, \quad j = 1, \dots, m, \\ \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_h} \partial w_{ij}} \Big|_{\omega=\omega_0} &= 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 Q_{jw}}{\partial \alpha_i \partial w_{ij}} \Big|_{\omega=\omega_0} &= \frac{2n_i s_{ij}}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} - 1 \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \frac{\partial^2 Q_{jw}}{\partial \alpha_l \partial w_{ij}} \Big|_{\omega=\omega_0} &= \frac{2n_l \beta_l \sigma_{x_j}^2 n_i \beta_i s_{ij}}{a_j \sigma_{lj}^2 \sigma_{ij}^2}, \quad l, i = 2, \dots, p, \text{ para } l \neq i, \\ \frac{\partial^2 Q_{jw}}{\partial \beta_i \partial w_{ij}} \Big|_{\omega=\omega_0} &= \frac{2\sigma_{x_j}^2 n_i s_{ij}}{a_j \sigma_{ij}^2} \left\{ \frac{2n_i \beta_i^2 \sigma_{x_j}^2 M_j}{\sigma_{ij}^2 a_j} + \frac{\beta_i}{\sigma_{ij}^2} \left(n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right) - M_j \right\}, \\ j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \frac{\partial^2 Q_{jw}}{\partial \beta_l \partial w_{ij}} \Big|_{\omega=\omega_0} &= \frac{2\sigma_{x_j}^2 \beta_i s_{ij} n_i}{a_j \sigma_{lj}^2 \sigma_{ij}^2} \left\{ \frac{2M_j n_l \beta_l \sigma_{x_j}^2}{a_j} + n_l \alpha_l - \sum_{k=1}^{n_l} Y_{ljk} \right\}, \\ l, i = 2, \dots, p, \text{ para } l \neq i. \end{aligned}$$

3. Perturbação na variável resposta de todos os laboratórios no j-ésimo pata-mar:

Neste esquema são usados as mesmas condições utilizadas no caso normal, com a função de log-verossimilhança perturbada do modelo t-Student.

$$\begin{aligned} \Delta_{1j1j} &= \frac{\partial^2 L(\theta|\omega)}{\partial \mu_{x_j} \partial w_{1j}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \mu_{x_j}} \frac{\partial Q_{jw}}{\partial w_{1j}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_j} \partial w_{1j}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{1jij} &= \frac{\partial^2 L(\theta|\omega)}{\partial \mu_{x_j} \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \mu_{x_j}} \frac{\partial Q_{jw}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \mu_{x_j} \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{1hij} &= \frac{\partial^2 L(\theta|\omega)}{\partial \mu_{x_h} \partial w_{ij}} \Big|_{\omega=\omega_0} = 0, \quad j, h = 1, \dots, m, \text{ para } h \neq j. \\ \Delta_{2i1j} &= \frac{\partial^2 L(\theta|\omega)}{\partial \alpha_i \partial w_{1j}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \alpha_i} \frac{\partial Q_{jw}}{\partial w_{1j}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \alpha_i \partial w_{1j}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{2iij} &= \frac{\partial^2 L(\theta|\omega)}{\partial \alpha_i \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \alpha_i} \frac{\partial Q_{jw}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \alpha_i \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{21ij} &= \frac{\partial^2 L(\theta|\omega)}{\partial \alpha_1 \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \alpha_1} \frac{\partial Q_{jw}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \alpha_1 \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{3i1j} &= \frac{\partial^2 L(\theta|\omega)}{\partial \beta_i \partial w_{1j}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \beta_i} \frac{\partial Q_{jw}}{\partial w_{1j}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \beta_i \partial w_{1j}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{3iij} &= \frac{\partial^2 L(\theta|\omega)}{\partial \beta_i \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \beta_i} \frac{\partial Q_{jw}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \beta_i \partial w_{ij}} \right) \Big|_{\omega=\omega_0}, \\ \Delta_{3lij} &= \frac{\partial^2 L(\theta|\omega)}{\partial \beta_l \partial w_{ij}} \Big|_{\omega=\omega_0} = \frac{(\gamma+n)}{2} \left(\frac{1}{(\gamma-2+Q_{jw})^2} \frac{\partial Q_{jw}}{\partial \beta_l} \frac{\partial Q_{jw}}{\partial w_{ij}} - \frac{1}{(\gamma-2+Q_{jw})} \frac{\partial^2 Q_{jw}}{\partial \beta_l \partial w_{ij}} \right) \Big|_{\omega=\omega_0}. \end{aligned}$$

Neste esquema de perturbação as expressões referentes às derivadas de $\frac{\partial Q_{j\omega}}{\partial \omega_{1j}}$, $\frac{\partial^2 Q_{j\omega}}{\partial \theta \partial \omega_{1j}}$, $\frac{\partial Q_{j\omega}}{\partial \omega_{ij}}$ e $\frac{\partial^2 Q_{j\omega}}{\partial \theta \partial \omega_{ij}}$, após manipulações algébricas, encontram-se a seguir. As expressões da derivada de $\frac{\partial Q_{j\omega}}{\partial \theta}$ são encontrados no apêndice A, porém estes resultados já estão aplicado no ponto ω_0 .

$$\begin{aligned} \left. \frac{\partial Q_{j\omega}}{\partial \omega_{1j}} \right|_{\omega=\omega_0} &= \left(\frac{2s_{1j}}{\sigma_{1j}^2} \right) \left\{ \sum_{k=1}^{n_1} Y_{1jk} - \frac{n_1 \sigma_{x_j}^2 M_j}{a_j} \right\}, \quad j = 1, \dots, m, \\ \left. \frac{\partial Q_{j\omega}}{\partial \omega_{ij}} \right|_{\omega=\omega_0} &= \left(\frac{2s_{ij}}{\sigma_{ij}^2} \right) \left\{ \sum_{k=1}^{n_i} Y_{ijk} - n_i \alpha_i - \frac{n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} \right\}, \quad i = 2, \dots, p \text{ e } j = 1, \dots, m, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_j} \partial \omega_{1j}} \right|_{\omega=\omega_0} &= -\frac{2s_{1j} n_1}{a_j \sigma_{1j}^2}, \quad j = 1, \dots, m, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_j} \partial \omega_{ij}} \right|_{\omega=\omega_0} &= -\frac{2n_i \beta_i s_{ij}}{a_j \sigma_{ij}^2}, \quad i = 2, \dots, p \text{ e } j = 1, \dots, m, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \mu_{x_h} \partial \omega_{ij}} \right|_{\omega=\omega_0} &= 0, \quad i = 2, \dots, p \text{ e } j, h = 1, \dots, m, \text{ para } h \neq j, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \alpha_i \partial \omega_{1j}} \right|_{\omega=\omega_0} &= \frac{2n_i \beta_i \sigma_{x_j}^2 s_{1j} n_1}{a_j \sigma_{1j}^2 \sigma_{ij}^2}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \alpha_i \partial \omega_{ij}} \right|_{\omega=\omega_0} &= \frac{2n_i s_{ij}}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} - 1 \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \alpha_l \partial \omega_{ij}} \right|_{\omega=\omega_0} &= \frac{2n_l \beta_l \sigma_{x_j}^2 n_i \beta_i s_{ij}}{a_j \sigma_{lj}^2 \sigma_{ij}^2}, \quad l, i = 2, \dots, p, \text{ para } l \neq i, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \beta_i \partial \omega_{1j}} \right|_{\omega=\omega_0} &= \left[\frac{2\sigma_{x_j}^2 s_{1j} n_1}{a_j \sigma_{1j}^2 \sigma_{ij}^2} \right] \left\{ \frac{2n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right\}, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \beta_i \partial \omega_{ij}} \right|_{\omega=\omega_0} &= \frac{2\sigma_{x_j}^2 n_i s_{ij}}{a_j \sigma_{ij}^2} \left\{ \frac{2n_i \beta_i^2 \sigma_{x_j}^2 M_j}{\sigma_{ij}^2 a_j} + \frac{\beta_i}{\sigma_{ij}^2} \left(n_i \alpha_i - \sum_{k=1}^{n_i} Y_{ijk} \right) - M_j \right\}, \\ & \quad j = 1, \dots, m \text{ e } i = 2, \dots, p, \\ \left. \frac{\partial^2 Q_{j\omega}}{\partial \beta_l \partial \omega_{ij}} \right|_{\omega=\omega_0} &= \frac{2\sigma_{x_j}^2 n_i \beta_i s_{ij}}{a_j \sigma_{lj}^2 \sigma_{ij}^2} \left\{ \frac{2M_j n_l \beta_l \sigma_{x_j}^2}{a_j} + n_l \alpha_l - \sum_{k=1}^{n_l} Y_{ljk} \right\}, \\ & \quad l, i = 2, \dots, p, \text{ para } l \neq i. \end{aligned}$$

3

Método para seleção apropriada da perturbação

Como mencionado no decorrer do trabalho, sabemos quão importante é detectar as observações influentes em uma modelagem estatística, e perturbar o modelo e/ou os dados de forma arbitrária pode conduzir a análise de diagnóstico e a conclusões equivocadas. Galea-Rojas (2014) estudou em seu artigo a influência local em um modelo de calibração comparativa balanceado na presença de medições *gold standard*, onde foi desenvolvido e aplicado a técnica de seleção apropriada. No modelo utilizado, é assumido para a matriz de variância-covariância a estrutura de independência entre os métodos de medições, e este modelo permite diferentes níveis de medições entre cada um dos métodos com o laboratório de referência (*gold standard*). Foi utilizado a técnica de Zhu et al. (2007) para verificar se as perturbações utilizadas são ou não apropriadas. Na avaliação da influência local, foi estudado as perturbações aditiva nos dados e na matriz de variância-covariância e, após estudar se as perturbações são adequadas, foram verificadas que são apropriadas, pois em ambos os casos as propriedades foram satisfeitas.

Segundo Zhu et al. (2007) e em uma recente publicação em modelagens considerando erro nas variáveis (Giménez & Patat 2014), consideram que tamanhos de grupos distintos (réplicas desbalanceadas), pode levar o atual esquema de perturbação a possuir um efeito sem equilíbrio entre os grupos, gerando assim conclusões equivocadas. Portanto, perturbar de forma arbitrária através de um esquema de perturbação pode mascarar o que de fato está acontecendo, levando a interpretações e inferências errôneas a respeito da causa de um efeito influente. Neste artigo foi estudado a influência local na modelagem de calibração comparativa funcional com dados replicados desbalanceados. Neste caso, a estrutura de

variância-covariância é independente e foram estudados oito esquemas de perturbação. Estes esquemas de perturbação utilizados poderão ser ou não apropriados, isto acontece pelo fato das réplicas serem desbalanceadas e estas conclusões somente serão observadas mediante aplicação da técnica. Neste artigo, a metodologia de influência local e método de seleção da perturbação apropriada foram aplicados em dois problemas distintos (*Egyptian pottery data* e *Oximetry data*). Em ambas as aplicações foi observado que cinco esquemas de perturbação avaliados não foram apropriados.

Na modelagem em questão existem situações em que as réplicas são desbalanceadas e a utilização da técnica se justifica por si só. Porém observa-se que nesta modelagem há estrutura de dependência na matriz de variância-covariância, onde essa estrutura de dependência pode ou não exercer um desequilíbrio no efeito sobre o estudo de influência local. A técnica da seleção apropriada neste caso poderá detectar se o efeito existe, e caso exista propor uma perturbação adequada para acessar corretamente a influência local e poder tomar a decisão certa a respeito das observações influentes.

Vale salientar que esta estrutura de dependência é decorrente das características inerentes do problema que toma uma única unidade experimental sendo explorada por diversos laboratórios. Todos os laboratórios encaminharam juntamente com as medidas, as incertezas referentes as medidas.

3.1 Metodologia

Seja $\boldsymbol{\omega}$ o vetor de perturbações como definido na Equação 2.1 e $L(\boldsymbol{\theta}/\boldsymbol{\omega})$ a função de log-verossimilhança do modelo perturbado. Para avaliar se a perturbação é adequada [Zhu et al. \(2007\)](#) sugerem calcular a função $G(\boldsymbol{\omega})$, que é a esperança da matriz de informação de Fisher com respeito ao vetor de perturbação $\boldsymbol{\omega}$, $G(\boldsymbol{\omega}) = \mathbb{E}_{\boldsymbol{\omega}}\{U(\boldsymbol{\omega})U^T(\boldsymbol{\omega})\}$, onde

$$U(\boldsymbol{\omega}) = \frac{\partial L(\boldsymbol{\theta}/\boldsymbol{\omega})}{\partial \boldsymbol{\omega}}$$

é a função score de $\boldsymbol{\omega}$ no modelo perturbado e $\mathbb{E}_{\boldsymbol{\omega}}$ é denotada a esperança com respeito a $L(\boldsymbol{\theta}/\boldsymbol{\omega})$.

Seja $g_{ij}(\boldsymbol{\omega})$ o elemento ij da matriz $G(\boldsymbol{\omega})$. Os elementos em $g_{ii}(\boldsymbol{\omega})$ indicam a quantidade

de perturbação introduzida pelo i -ésimo componente de $\boldsymbol{\omega}$. Os elementos fora da diagonal representam a associação entre diferentes componentes de $\boldsymbol{\omega}$. Uma maneira de medir o coeficiente de associação entre os componentes de $\boldsymbol{\omega}$, é através de $r_{ij} = \frac{g_{ij}(\boldsymbol{\omega})}{\sqrt{g_{ii}(\boldsymbol{\omega})g_{jj}(\boldsymbol{\omega})}}$. Quanto maior for o valor de r_{ij} , maior é o grau de associação.

Desta forma, uma perturbação é apropriada se a matriz $G(\boldsymbol{\omega})$ satisfaz as condições abaixo:

1. $G(\boldsymbol{\omega})$ é de posto completo numa vizinhança de $\boldsymbol{\omega}_0$ com propósito de evitar qualquer componente redundante de $\boldsymbol{\omega}$.
2. Os componentes fora da diagonal devem ser tão pequenos quanto possível com propósito de evitar uma forte associação entre eles e, em consequência, perturbações com efeitos ambíguos.
3. As diferenças entre os componentes da diagonal devem ser tão pequenas quanto possível tal que as quantidades introduzidas pelos elementos de $\boldsymbol{\omega}$ sejam semelhantes.

[Zhu et al. \(2007\)](#) afirmam que para que a perturbação seja apropriada atendendo as condições citadas, deve satisfazer a relação $G(\boldsymbol{\omega}_0) = c\mathbf{I}_q$, onde $c > 0$ (constante) e \mathbf{I}_q é a matriz identidade de ordem q .

Caso as condições 1, 2 e 3 não sejam satisfeita, deve-se assumir um novo vetor de perturbação como definido em (3.1).

$$\boldsymbol{\omega}^* = \boldsymbol{\omega}_0 + c^{-\frac{1}{2}}G(\boldsymbol{\omega}_0)^{\frac{1}{2}}(\boldsymbol{\omega} - \boldsymbol{\omega}_0), \quad (3.1)$$

tal que $G(\boldsymbol{\omega}^*)$, avaliado em $\boldsymbol{\omega}_0$ seja igual a $c\mathbf{I}_q$. O novo vetor de perturbação expresso em (3.1) é uma transformação do sistema de coordenadas de modo que seus elementos são ortonormais (independentes). Isto é, os elementos não são associados, ao menos localmente, e introduzem iguais quantidades de perturbação.

3.2 Perturbação na variável resposta

Neste estudo daremos enfoque a metodologia de [Zhu et al. \(2007\)](#). Esta técnica será utilizada no esquema de perturbação da variável resposta para todos os casos abordados

na Seção 2.2.2. Nestes casos a matriz utilizada para avaliar se os esquemas de perturbação adotados são adequados é a matriz $G(\boldsymbol{\omega}_0)$. Nas subseções a seguir são expressas as formulas úteis para calcular os elementos que compõe a matriz de avaliação da perturbação adequada para os esquemas de perturbação no laboratório de referência, no i -ésimo laboratório e em todos os laboratórios com patamar fixo respectivamente. Após manipulações algébricas, as equações das esperanças necessárias para obter as expressões utilizadas para calcular os valores da matriz $G(\boldsymbol{\omega}_0)$, encontram-se abaixo.

$$\mathbb{E} \left\{ \sum_{k=1}^{n_1} Y_{1jk} \right\} = n_1 \mu_{xj}, \quad j = 1, \dots, m,$$

$$\mathbb{E} \left\{ \left(\sum_{k=1}^{n_1} Y_{1jk} \right)^2 \right\} = n_1 \sigma_{1j}^2 + n_1^2 \sigma_{xj}^2 + (n_1 \mu_{xj})^2, \quad j = 1, \dots, m,$$

$$\mathbb{E} \left\{ \left(\sum_{k=1}^{n_1} Y_{1jk} \right) \left(\sum_{k=1}^{n_1} Y_{1lk} \right) \right\} = n_1^2 \mu_{xj} \mu_{xl}, \quad j, l = 1, \dots, m,$$

$$\mathbb{E} \left\{ \left(\sum_{k=1}^{n_1} Y_{1jk} \right) \left(\sum_{k=1}^{n_1} Y_{ilk} \right) \right\} = n_1 n_i \mu_{xj} (\alpha_i + \beta_i \mu_{xl}), \quad j, l = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\mathbb{E} \left\{ \sum_{k=1}^{n_i} Y_{ijk} \right\} = n_i (\alpha_i + \beta_i \mu_{xj}), \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\mathbb{E} \left\{ \left(\sum_{k=1}^{n_i} Y_{ijk} \right)^2 \right\} = n_i \sigma_{ij}^2 + n_i^2 \beta_i^2 \sigma_{xj}^2 + (n_i \alpha_i)^2 + 2n_i^2 \alpha_i \beta_i \mu_{xj} + (n_i \beta_i \mu_{xj})^2, \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\mathbb{E} \left\{ \left(\sum_{k=1}^{n_i} Y_{ijk} \right) \left(\sum_{k=1}^{n_i} Y_{ilk} \right) \right\} = (n_i \alpha_i)^2 + n_i^2 \alpha_i \beta_i (\mu_{xj} + \mu_{xl}) + n_i^2 \beta_i^2 \mu_{xj} \mu_{xl}, \quad j, l = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\mathbb{E} \left\{ \sum_{k=1}^{n_i} Y_{ilk} M_j \right\} = \frac{n_i \alpha_j \mu_{xj}}{\sigma_{xj}^2} (\alpha_i + \beta_i \mu_{xl}), \quad j, l = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\mathbb{E} \left\{ \sum_{k=1}^{n_i} Y_{ijk} M_j \right\} = \frac{n_i \alpha_i a_j \mu_{xj}}{\sigma_{xj}^2} + \frac{n_i \beta_i a_j \mu_{xj}^2}{\sigma_{xj}^2} + n_i \beta_i (a_j - 1), \quad j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$\mathbb{E} \{ M_j \} = \frac{\mu_{xj} a_j}{\sigma_{xj}^2}, \quad j, l = 1, \dots, m,$$

$$\mathbb{E} \{ M_j M_l \} = \frac{\mu_{xj} \mu_{xl} a_j a_l}{\sigma_{xj}^2 \sigma_{xl}^2}, \quad j, l = 1, \dots, m \text{ e}$$

$$\mathbb{E} \{ M_j^2 \} = \left(\frac{\mu_{xj} a_j}{\sigma_{xj}^2} \right)^2 + \frac{a_j (a_j - 1)}{\sigma_{xj}^2}, \quad j = 1, \dots, m.$$

Nas subseções seguintes, as expressões obtidas em $g_{ii}(\boldsymbol{\omega})$ e $g_{ij}(\boldsymbol{\omega})$, que representam respectivamente os elementos dentro e fora da diagonal da matrix $G(\boldsymbol{\omega})$ serão apresentados.

3.2.1 Perturbação na variável resposta do laboratório de referência

Neste esquema a perturbação foi utilizada da seguinte forma: $Y_{1jk\omega} = Y_{1jk} + s_{1j} \omega_{1j}$. Após manipulações algébricas, foram obtidas as expressões que compõem a matrix $G(\boldsymbol{\omega})$ e estão apresentadas a seguir.

$$g_{jj}(\boldsymbol{\omega}) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{1j}} \right)^2 \right\} = n_1 \left(\frac{s_{1j}}{\sigma_{1j}^2} \right)^2 \left\{ \sigma_{1j}^2 - \frac{n_1 \sigma_{xj}^2}{a_j} + n_1 (s_{1j} \omega_{1j})^2 \left(\frac{\sigma_{xj}^2 n_1}{a_j \sigma_{1j}^2} - 1 \right)^2 \right\},$$

$j = 1, \dots, m,$

$$g_{jl}(\boldsymbol{\omega}) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{1j}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{1l}} \right) \right\} = \frac{(n_1 s_{1j} s_{1l})^2 \omega_{1j} \omega_{1l}}{\sigma_{1j}^2 \sigma_{1l}^2} \left(1 - \frac{n_1 \sigma_{xj}^2}{a_j \sigma_{1j}^2} \right) \left(1 - \frac{n_1 \sigma_{xl}^2}{a_l \sigma_{1l}^2} \right),$$

$j, l = 1, \dots, m,$ com $j \neq l,$

Portanto, no ponto ω_0 temos:

$$g_{jj}(\boldsymbol{\omega}_0) = n_1 \left(\frac{s_{1j}}{\sigma_{1j}^2} \right)^2 \left\{ \sigma_{1j}^2 - \frac{n_1 \sigma_{xj}^2}{a_j} \right\}, \quad j = 1, \dots, m,$$

$$g_{jl}(\boldsymbol{\omega}_0) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{1j}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{1l}} \right) \right\} = 0, \quad j, l = 1, \dots, m, \text{ com } j \neq l.$$

E portanto, neste caso a perturbação não satisfaz as condições, pois os elementos da diagonal são diferentes entre si, apesar dos elementos fora da diagonal serem igual a zero. Assim, a nova perturbação é da seguinte forma: $Y_{1jk\omega} = Y_{1jk} + s_{1j} (g_{jj}(\omega_0))^{-\frac{1}{2}} \omega_{1j}^*$.

3.2.2 Perturbação na variável resposta do i -ésimo laboratório

Neste esquema a perturbação foi considerada da seguinte forma: $Y_{ijk\omega} = Y_{ijk} + s_{ij}\omega_{ij}$. Após manipulações algébricas, foram obtidas as expressões que compõem a matriz $G(\omega)$ que estão apresentadas a seguir.

$$g_{jj}(\omega) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{ij}} \right)^2 \right\} = \left(\frac{s_{ij}}{\sigma_{ij}^2} \right)^2 \left\{ n_i \sigma_{ij}^2 - \frac{(n_i \beta_i)^2 \sigma_{x_j}^2}{a_j} + \left(n_i s_{ij} \omega_{ij} \left(1 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} \right) \right)^2 \right\},$$

$i = 2, \dots, p$ e $j = 1, \dots, m$,

$$g_{jl}(\omega) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{ij}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{il}} \right) \right\} = \frac{(n_i s_{ij} s_{il})^2 \omega_{ij} \omega_{il}}{\sigma_{ij}^2 \sigma_{il}^2} \left(1 - \frac{n_i \beta_i^2 \sigma_{x_l}^2}{a_l \sigma_{il}^2} \right) \left(1 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} \right),$$

$i = 2, \dots, p$ e $j, l = 1, \dots, m$, com $j \neq l$,

Portanto, no ponto ω_0 temos:

$$g_{jj}(\omega_0) = n_i \left(\frac{s_{ij}}{\sigma_{ij}^2} \right)^2 \left\{ \sigma_{ij}^2 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j} \right\}, \quad i = 2, \dots, p \text{ e } j = 1, \dots, m,$$

$$g_{jl}(\omega_0) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{ij}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{il}} \right) \right\} = 0, \quad i = 2, \dots, p \text{ e } j, l = 1, \dots, m, \text{ com } j \neq l.$$

Desta forma a perturbação não satisfaz as condições, pois os elementos da diagonal são diferentes entre si, embora os elementos fora da diagonal foram igual a zero. Assim, a nova perturbação é da seguinte forma: $Y_{ijk\omega} = Y_{ijk} + s_{ij}(g_{jj}(\omega_0))^{-\frac{1}{2}} \omega_{1j}^*$.

3.2.3 Perturbação nas variáveis respostas do j -ésimo patamar

Neste esquema a perturbação foi dada da seguinte forma: $Y_{ijk\omega} = Y_{ijk} + s_{ij}\omega_{ij}$. Após manipulações algébricas, foram obtidas as expressões que compõem a matriz $G(\omega)$ que estão apresentadas a seguir.

$$g_{11}(\omega) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{1j}} \right)^2 \right\} = n_1 \left(\frac{s_{1j}}{\sigma_{1j}^2} \right)^2 \left\{ \sigma_{1j}^2 - \frac{n_1 \sigma_{x_j}^2}{a_j} + n_1 (s_{1j} \omega_{1j})^2 \left(\frac{\sigma_{x_j}^2 n_1}{a_j \sigma_{1j}^2} - 1 \right)^2 \right\},$$

$j = 1, \dots, m$,

$$g_{1l}(\omega) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{1j}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{1l}} \right) \right\} = \frac{(n_1 s_{1j} s_{1l})^2 \omega_{1j} \omega_{1l}}{\sigma_{1j}^2 \sigma_{1l}^2} \left(1 - \frac{n_1 \sigma_{x_j}^2}{a_j \sigma_{1j}^2} \right) \left(1 - \frac{n_1 \sigma_{x_l}^2}{a_l \sigma_{1l}^2} \right),$$

$j = 1, \dots, m$, e $l = 2, \dots, p$,

$$g_{ii}(\omega) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\omega)}{\partial \omega_{ij}} \right)^2 \right\} = \left(\frac{s_{ij}}{\sigma_{ij}^2} \right)^2 \left\{ n_i \sigma_{ij}^2 - \frac{(n_i \beta_i)^2 \sigma_{x_j}^2}{a_j} + \left(n_i s_{ij} \omega_{ij} \left(1 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} \right) \right)^2 \right\}, \quad i =$$

$2, \dots, p$ e $j = 1, \dots, m$,

$$g_{il}(\boldsymbol{\omega}) = \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{ij}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{lj}} \right) \right\} = \frac{(n_i s_{ij} s_{lj})^2 \omega_{ij} \omega_{lj}}{\sigma_{ij}^2 \sigma_{lj}^2} \left(1 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j \sigma_{ij}^2} \right) \left(1 - \frac{n_l \beta_l^2 \sigma_{x_j}^2}{a_j \sigma_{lj}^2} \right),$$

$i, l = 2, \dots, p$ e $j = 1, \dots, m$, com $i \neq l$,

Portanto, no ponto $\boldsymbol{\omega}_0$ temos:

$$\begin{aligned} g_{11}(\boldsymbol{\omega}_0) &= n_1 \left(\frac{s_{1j}}{\sigma_{1j}^2} \right)^2 \left\{ \sigma_{1j}^2 - \frac{n_1 \sigma_{x_j}^2}{a_j} \right\}, \quad j = 1, \dots, m, \\ g_{1l}(\boldsymbol{\omega}_0) &= \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{1j}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{lj}} \right) \right\} = 0, \quad j = 1, \dots, m \text{ e } l = 2, \dots, p, \\ g_{ii}(\boldsymbol{\omega}_0) &= n_i \left(\frac{s_{ij}}{\sigma_{ij}^2} \right)^2 \left\{ \sigma_{ij}^2 - \frac{n_i \beta_i^2 \sigma_{x_j}^2}{a_j} \right\}, \quad i = 2, \dots, p \text{ e } j = 1, \dots, m, \\ g_{il}(\boldsymbol{\omega}_0) &= \mathbb{E} \left\{ \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{ij}} \right) \left(\frac{\partial L(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_{lj}} \right) \right\} = 0, \quad i, l = 2, \dots, p \text{ com } i \neq l \text{ e } j = 1, \dots, m \end{aligned}$$

Neste caso todas as perturbações não satisfazem as condições, pois os elementos da diagonal são diferentes entre si, apesar de que os elementos fora da diagonal são iguais a zero. Assim, a nova perturbação é da seguinte forma: $Y_{ijk_\omega} = Y_{ijk} + s_{ij}(g_{jj}(\boldsymbol{\omega}_0))^{-\frac{1}{2}} \omega_{ij}^*$.

Vale salientar que em todos os casos, as perturbações não satisfazem as propriedades que são: ter os elementos da diagonal iguais e os elementos fora da diagonal iguais a zero. Desta forma, deve-se considerar a nova perturbação para que assim a técnica de influência local seja aplicada corretamente. Note que nas matrizes $G(\boldsymbol{\omega})$ em todos os casos, os elementos que compõem a diagonal são diferentes e os elementos fora da diagonal são zeros. Estas matrizes serão utilizadas para corrigir as perturbações. No Capítulo 4 é apresentado a aplicação do estudo de influência local utilizando a técnica apropriada para corrigir e obter os resultados coerentes.

4

Aplicação

Neste capítulo vamos aplicar a metodologia proposta por [Cook \(1986\)](#), na modelagem normal e t-Student, para um conjunto de dados com réplicas balanceadas. Para melhor compreender os dados, as modelagens e os resultados do diagnóstico, serão apresentados os resultados da análise exploratória dos dados. Em seguida serão apresentadas as estimativas referentes aos dois modelos e a mudança relativa (influência global) mediante a retirada por vez dos laboratórios e patamares na modelagem normal. Para inferir a respeito da qualidade do modelo e identificar se as modelagens são sensíveis a pequenas perturbações, será utilizada a avaliação de influência local. O conjunto de dados desta aplicação encontra-se no apêndice B e é composto pelas medições da potência de um motor 1.0 à gasolina, onde o número de laboratórios participantes foram seis ($p = 6$), o número de patamares considerados foram dez ($m = 10$) e o número de réplicas foram cinco ($k=5$).

4.1 Análise exploratória

Será feita a análise exploratória de dados com o objetivo de descrever a organização, apresentação e sintetização dos dados. Para isto serão apresentados tabelas com as medidas de posição e dispersão, gráficos do tipo dispersão e boxplot. As informações resultantes dessa análise poderão revelar indicadores de comportamentos e tendências que permitirão uma melhor compreensão a respeito dos dados e da modelagem.

Nas tabelas 4.1 e 4.2, encontram-se as medidas de tendência central e variabilidade, por patamares. Desta forma, pode-se observar, como se distribui os valores para cada patamar e assim identificar a ocorrência de algo fora do padrão.

Na descrição dos dados, foi considerado a análise de forma separada entre o laboratório de referência e os demais laboratórios. Isto ocorre pelo fato do laboratório de referência

obter boas medições e ao considerá-lo junto com os demais laboratórios pode mascarar as estatísticas reais destes dados.

A partir da análise da tabela 4.1, pode-se observar que de todos os patamares, o patamar 6000 apresenta a maior amplitude e variância, e os patamares 1500 e 6000 apresentam os maiores coeficientes de variação em relação aos demais patamares.

Tabela 4.1 *Resultados descritivos referente às observações do laboratório de referência por patamares.*

Pot. do motor	Mín.	Med	Média	Máx.	Amplitude	Variância	Coef. de variação (%)
1500	19,64	19,78	19,80	19,96	0,32	0,01	0,60
2000	27,18	27,25	27,26	27,33	0,15	0,00	0,22
2500	35,40	35,51	35,50	35,57	0,17	0,00	0,20
3000	45,34	45,54	45,53	45,61	0,27	0,01	0,24
3500	54,26	54,33	54,37	54,52	0,26	0,01	0,19
4000	63,08	63,17	63,29	63,68	0,60	0,06	0,39
4500	70,91	71,51	71,40	71,56	0,65	0,07	0,38
5000	76,78	76,99	76,92	77,01	0,23	0,01	0,14
5500	80,06	80,21	80,22	80,35	0,29	0,01	0,15
6000	79,13	80,13	79,96	80,36	1,22	0,23	0,61

Observando a tabela 4.2, os patamares 1 e 10, possuem respectivamente a menor e maior amplitude; o patamar 1 possui a menor variância, o patamar 7 possui a maior variância e o patamar 10 possui o menor coeficiente de variação.

Tabela 4.2 *Resultados descritivos referente às observações dos laboratórios restantes por patamares.*

Pot. do motor	Mín.	Med	Média	Máx.	Amplitude	Variância	Coef. de variação (%)
1500	19,06	19,60	19,56	20,02	0,96	0,06	1,29
2000	26,37	27,01	27,03	27,62	1,25	0,15	1,45
2500	34,21	35,26	35,13	36,08	1,87	0,28	1,52
3000	43,43	45,18	44,99	46,05	2,62	0,60	1,73
3500	51,92	53,92	53,75	55,31	3,39	0,80	1,66
4000	61,07	62,53	62,57	63,89	2,82	0,85	1,48
4500	68,48	70,36	70,14	71,67	3,19	1,13	1,52
5000	73,94	75,90	75,54	77,25	3,31	1,04	1,35
5500	77,53	78,60	78,73	80,84	3,31	0,67	1,04
6000	77,55	78,67	78,73	81,46	3,91	0,64	1,01

Para entender melhor o resumo descrito na Tabela 4.2, veja as Figuras 4.1 e 4.2, que são respectivamente os boxplots e a dispersão dos dados para os dez patamares.

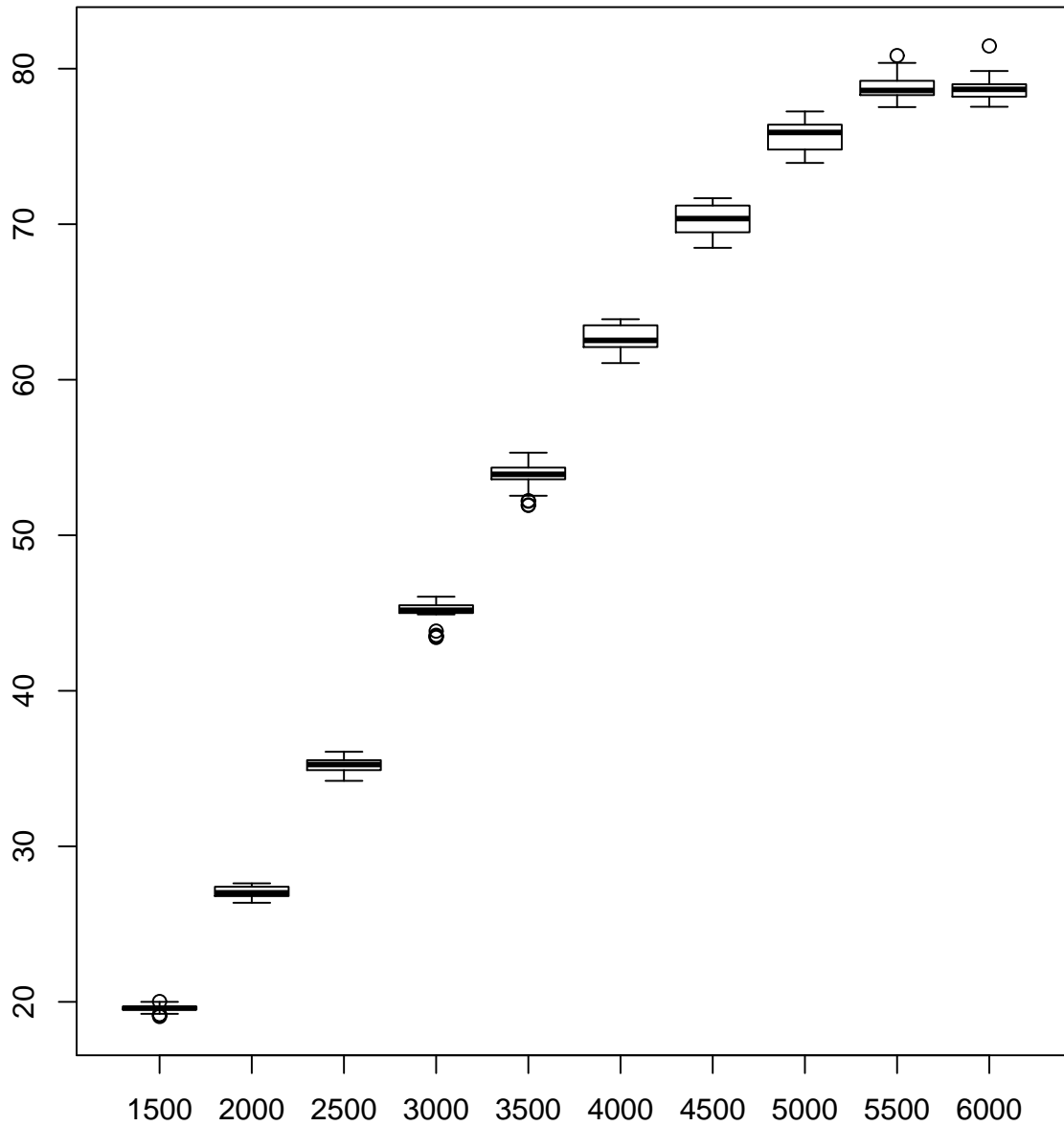


Figura 4.1 *Boxplot da potência do motor por patamares dos laboratórios 2 a 6.*

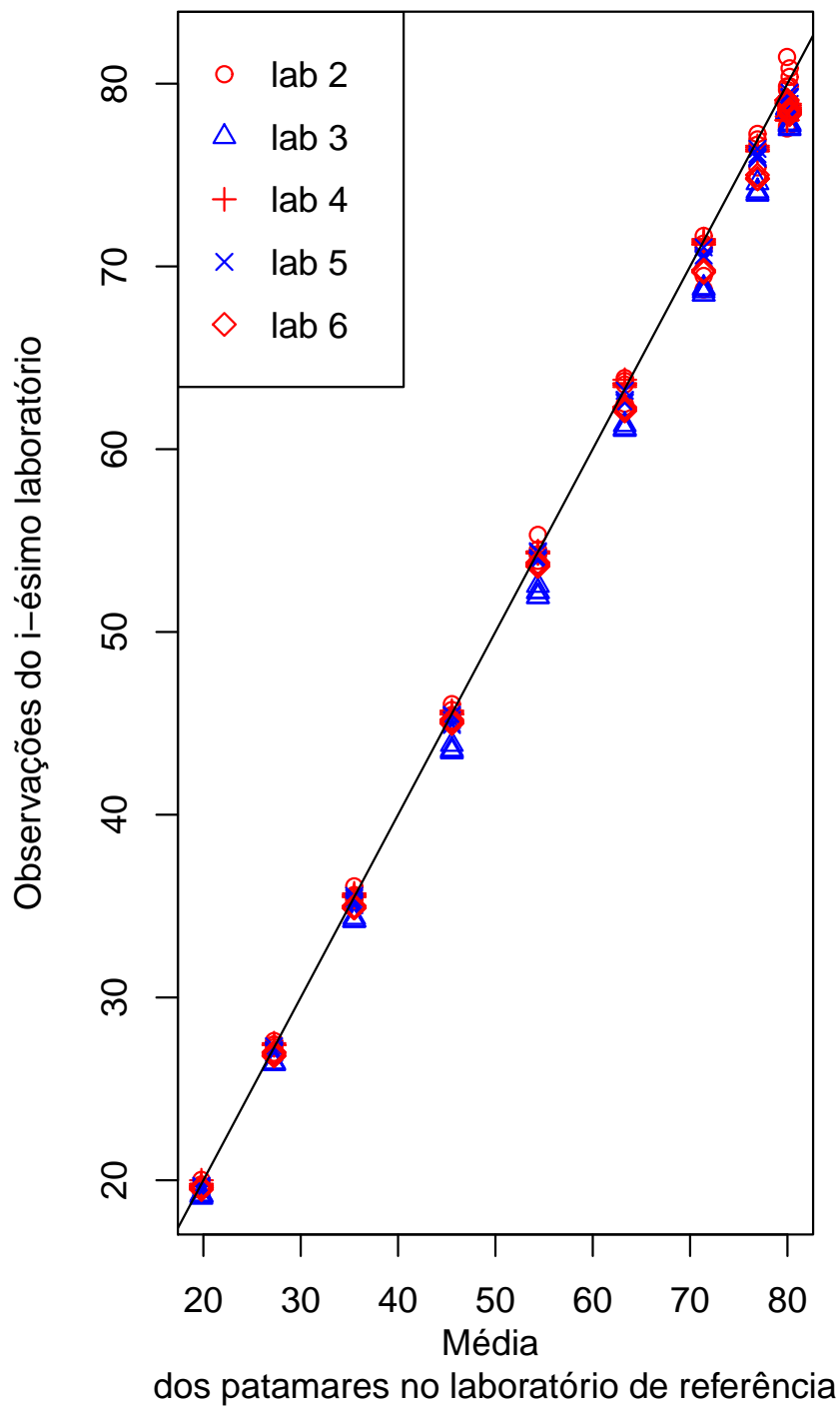


Figura 4.2 *Dispersão das observações por patamares dos laboratórios.*

No plano cartesiano da Figura 4.2, os pontos são referentes a dispersão dos valores das observações da potência do motor desconsiderando o laboratório de referência, eixo y, e a média dos patamares do laboratório de referência, eixo x. Uma alternativa de comparar as medidas dos laboratórios com o laboratório de referência é tomar como orientação os parâmetros $\alpha = 0$ e $\beta = 1$, implicando em uma reta de 45° , desta forma, pode-se perceber, em quais patamares os laboratórios são mais viesados. Nota-se que a medida que se pula de um patamar para o outro no sentido crescente, a dispersão aumenta, a variabilidade no primeiro patamar é bem pequena e nos últimos patamares a variabilidade é bem maior. As medidas do laboratório três se distanciam bastante em relação às medidas dos demais laboratórios, isso ocorrendo principalmente entre os patamares 4 ao 9. O mesmo acontece com o laboratório dois, onde suas medidas se destacam dos demais a partir do patamar 3 até o patamar 10.

Na Figura 4.1 encontram-se os *boxplot's* referentes a cada patamar onde foi identificado que existem algumas observações outlier. No laboratório dois foram detectadas uma observação outlier nos patamares 1, 9 e 10. Em relação ao laboratório três no patamar 1 foram detectados quatro observações outliers das cinco existentes, no patamar 4 foi detectado que todas as observações foram outliers, e no patamar 5 foi identificado quatro observações outliers. Em resumo, observamos que os laboratórios dois e três apresentam observações outliers que possivelmente podem exercer uma influência na modelagem.

Vale salientar que estas observações identificadas podem ou não ter um determinado efeito a ponto de influenciar na sensibilidade da modelagem utilizada.

4.2 Processo de modelagem normal e t-Student

Considerando o conjunto de dados das medições da potência do motor realizadas pelos diversos laboratórios participantes em m pontos de rotação, obtivemos as EMV dos parâmetros do modelo considerando o algoritmo EM descrito na dissertação da [Talarico \(2014\)](#).

O algoritmo EM foi implementado utilizando o programa R core team. Os valores iniciais usados no algoritmo para μ_{x_j} foram as médias dos laboratórios por patamares, para o

parâmetro α foi usado 0 e para o β foi usado 1. O critério de convergência utilizado foi calculado da seguinte forma:

$$\sum_{j=1}^m \left(\mu_{x_j}^{(r)} - \mu_{x_j}^{(r-1)} \right)^2 + \sum_{i=2}^p \left(\alpha_i^{(r)} - \alpha_i^{(r-1)} \right)^2 + \sum_{i=2}^p \left(\beta_i^{(r)} - \beta_i^{(r-1)} \right)^2 < \epsilon, \text{ onde}$$

$$\epsilon = 10^{-9}$$

A seguir a estimação do modelo de calibração para as distribuições de probabilidade normal e t-Student.

4.2.1 Resultados do algoritmo EM para o modelo normal e t-Student

Para estimar os parâmetros do modelo foi usado o algoritmo EM e seus resultados são expressos nas tabelas 4.3 e 4.4, com os desvios padrões assintóticos (dpa). Vale salientar que para a modelagem sob distribuição t o grau de liberdade utilizado foi 3,4, onde este valor baseia-se na Figura 4.3, pois garante o menor AIC e BIC.

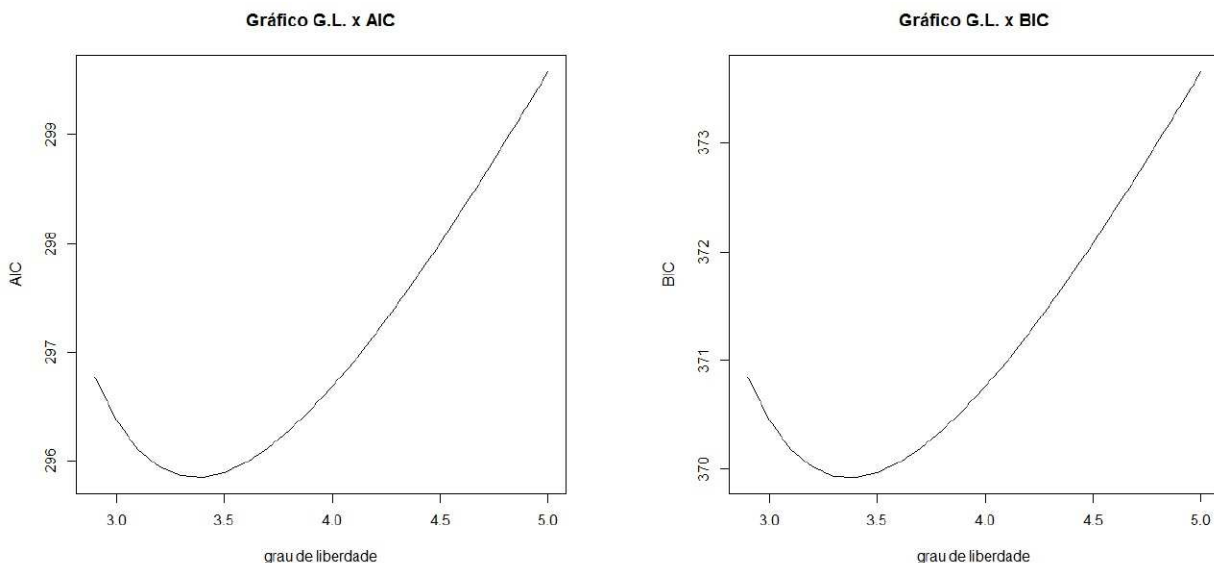


Figura 4.3 *Graus de liberdade versus AIC e BIC.*

Este tipo de modelagem tem a vantagem de avaliar, através das estimativas dos parâmetros, quais laboratórios têm menor viés em comparação ao laboratório de referência, desta forma,

a interpretação para o parâmetro α é quanto mais próximo de 0 for α , menor é o viés aditivo e para o parâmetro β será quanto mais próximo do valor 1 for, menor é o viés multiplicativo.

Tabela 4.3 *Estimativas de máxima verossimilhança do modelo normal.*

μ_{x_j}	dpa(μ_{x_j})	α_i	dpa(α_i)	β_i	dpa(β_i)
19,717	0,121	0,224	0,163	0,991	0,004
27,317	0,151	-0,187	0,188	0,970	0,004
35,562	0,189	0,418	0,132	0,987	0,003
45,679	0,242	0,074	0,446	0,989	0,008
54,539	0,289	0,243	0,133	0,978	0,003
63,552	0,337				
71,320	0,380				
76,692	0,409				
79,831	0,429				
79,789	0,432				

Tabela 4.4 *Estimativas de máxima verossimilhança do modelo t-Student.*

μ_{x_j}	dpa(μ_{x_j})	α_i	dpa(α_i)	β_i	dpa(β_i)
19,717	0,080	0,199	0,118	0,994	0,003
27,299	0,075	0,050	0,123	0,963	0,003
35,525	0,084	0,237	0,099	0,995	0,003
45,622	0,157	0,042	0,271	0,991	0,006
54,466	0,171	0,265	0,093	0,997	0,003
63,469	0,206				
71,225	0,292				
76,586	0,269				
79,715	0,455				
79,665	0,800				

Para avaliar se as modelagens feitas sofrem alguma influência de observações, iremos apresentar a seguir alguns resultados e interpretações referente a mudança relativa quando é retirado o laboratório e patamar um por vez.

4.2.2 Avaliação da mudança relativa nas estimativas dos parâmetros.

Uma alternativa para investigar se uma observação é influente globalmente, é utilizar o cálculo da mudança relativa (MR) expresso na Equação (4.1), onde l indica qual observação foi retirada (podendo ser retirados patamares, laboratórios ou ambas). Esta equação mede a porcentagem de variação na diferença da estimativa do vetor de parâmetros do modelo completo com a estimativa do vetor de parâmetros do modelo que desconsidera os pontos influentes em relação às estimativas do modelo completo. Neste caso, optamos por apenas observar as mudanças relativa na retirada dos patamares e laboratórios. Vale salientar que as retiradas das observações ocorrem uma por vez e em nenhum momento foi retirada mais de uma observação por vez.

$$MR = \left| \frac{\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_{(l)}}{\hat{\boldsymbol{\theta}}} \right| \times 100 \quad (4.1)$$

A tabela 4.5 é referente à mudança relativa nas estimativas dos parâmetros após a retirada de cada patamar e a tabela 4.6 é referente à mudança relativa após a retirada dos laboratórios. Percebe-se que a retirada em cada patamar causou uma influência desproporcional nas estimativas dos dados.

Observando o efeito da retirada de cada patamar sobre as estimativas dos parâmetros das médias, vemos que as retiradas dos patamares 1, 2 e 4 causou uma mudança relativa maior que a retirada dos demais patamares. Vale salientar que a retirada do patamar 4 causou a maior mudança relativa em μ_{x_8} , μ_{x_9} e $\mu_{x_{10}}$.

Observando o efeito produzido na retirada de cada patamar sobre as estimativas dos parâmetros α e β , vemos que as maiores variações relativas nas estimativas ocorrerem na retirada dos patamares 1, 4 e 10.

Tabela 4.5 *Mudança Relativa no modelo normal mediante retirada dos patamares.*

Parâmetros	μ_{x_1}	μ_{x_2}	μ_{x_3}	μ_{x_4}	μ_{x_5}	μ_{x_6}	μ_{x_7}	μ_{x_8}	μ_{x_9}	$\mu_{x_{10}}$
Modelo completo	19,717	27,317	35,562	45,679	54,539	63,552	71,320	76,692	79,831	79,789
Patamar #1	-	0,43%	0,23%	0,09%	0,02%	0,03%	0,07%	0,08%	0,10%	0,10%
Patamar #2	0,118%	-	0,052%	0,032%	0,019%	0,010%	0,004%	0,000%	0,002%	0,002%
Patamar #3	0,048%	0,037%	-	0,022%	0,018%	0,015%	0,014%	0,013%	0,013%	0,013%
Patamar #4	0,112%	0,074%	0,015%	-	0,003%	0,001%	0,018%	0,114%	0,199%	0,217%
Patamar #5	0,006%	0,023%	0,032%	0,038%	-	0,043%	0,044%	0,045%	0,045%	0,046%
Patamar #6	0,016%	0,017%	0,036%	0,049%	0,057%	-	0,063%	0,065%	0,069%	0,072%
Patamar #7	0,008%	0,001%	0,007%	0,013%	0,016%	0,021%	-	0,023%	0,020%	0,015%
Patamar #8	0,033%	0,002%	0,023%	0,038%	0,047%	0,055%	0,059%	-	0,058%	0,055%
Patamar #9	0,061%	0,001%	0,038%	0,064%	0,077%	0,084%	0,090%	0,094%	-	0,102%
Patamar #10	0,032%	0,003%	0,023%	0,032%	0,036%	0,030%	0,030%	0,034%	0,040%	-
Parâmetros	α_2	α_3	α_4	α_5	α_6	β_2	β_3	β_4	β_5	β_6
Modelo completo	0,224	-0,187	0,418	0,074	0,243	0,991	0,970	0,987	0,989	0,978
Patamar #1	158,97%	67,48%	93,97%	283,45%	48,22%	0,63%	0,18%	0,69%	0,37%	0,21%
Patamar #2	20,39%	48,77%	13,04%	49,72%	1,16%	0,06%	0,13%	0,07%	0,05%	0,00%
Patamar #3	11,09%	2,85%	4,63%	36,69%	1,27%	0,01%	0,02%	0,00%	0,02%	0,00%
Patamar #4	52,97%	8,96%	40,92%	152,35%	12,77%	0,33%	0,05%	0,44%	0,27%	0,15%
Patamar #5	8,59%	3,03%	3,81%	4,11%	5,77%	0,11%	0,07%	0,10%	0,06%	0,09%
Patamar #6	7,70%	0,50%	13,70%	26,12%	1,16%	0,08%	0,01%	0,25%	0,05%	0,01%
Patamar #7	23,79%	25,43%	10,10%	16,58%	14,61%	0,21%	0,18%	0,17%	0,05%	0,14%
Patamar #8	12,54%	17,76%	2,08%	10,57%	27,50%	0,11%	0,12%	0,03%	0,00%	0,25%
Patamar #9	8,19%	14,45%	23,03%	13,01%	11,32%	0,07%	0,07%	0,35%	0,01%	0,10%
Patamar #10	10,85%	98,21%	34,53%	90,46%	24,17%	0,09%	0,60%	0,52%	0,20%	0,21%

Analisando a retirada dos laboratórios, pode-se observar o efeito produzido sobre as estimativas dos parâmetros da média, α e β . A retirada do laboratório 4 resultou nas maiores mudanças relativas em 7 das 10 médias dos patamares e refletiu as maiores mudanças nas estimativas dos parâmetros α e β . Portanto, a retirada do laboratório 4 foi o que mais impactou na mudança das estimativas dos parâmetros em geral.

Tabela 4.6 *Mudança Relativa no modelo normal mediante retirada dos laboratórios.*

Parâmetros	μ_{x_1}	μ_{x_2}	μ_{x_3}	μ_{x_4}	μ_{x_5}	μ_{x_6}	μ_{x_7}	μ_{x_8}	μ_{x_9}	$\mu_{x_{10}}$
Modelo completo	19,717	27,317	35,562	45,679	54,539	63,552	71,320	76,692	79,831	79,789
Lab. #2	0,03%	0,01%	0,04%	0,01%	0,02%	0,01%	0,07%	0,01%	0,02%	0,01%
Lab. #3	0,05%	0,02%	0,01%	0,13%	0,12%	0,06%	0,07%	0,01%	0,09%	0,27%
Lab. #4	0,20%	0,10%	0,11%	0,07%	0,11%	0,33%	0,34%	0,19%	0,32%	0,78%
Lab. #5	<0,01%	<0,01%	<0,01%	<0,01%	<0,01%	0,01%	<0,01%	0,01%	0,01%	0,01%
Lab. #6	0,07%	0,09%	0,10%	0,16%	0,06%	0,14%	0,16%	0,21%	0,07%	0,38%
Parâmetros	α_2	α_3	α_4	α_5	α_6	β_2	β_3	β_4	β_5	β_6
Modelo completo	0,224	-0,187	0,418	0,074	0,243	0,991	0,970	0,987	0,989	0,978
Lab. #2	-	3,04%	0,25%	14,25%	0,26%	-	0,01%	<0,01%	0,02%	<0,01%
Lab. #3	1,89%	-	1,29%	82,93%	0,19%	0,01%	-	0,01%	0,12%	<0,01%
Lab. #4	5,14%	22,75%	-	194,00%	1,17%	0,03%	0,09%	-	0,29%	0,01%
Lab. #5	0,04%	0,04%	0,02%	-	<0,01%	<0,01%	<0,01%	<0,01%	-	<0,01%
Lab. #6	1,71%	3,95%	1,65%	34,79%	-	0,01%	0,02%	0,02%	0,05%	-

A seguir será apresentado os resultados da aplicação do estudo de influência local e os resultados da metodologia de seleção apropriada da perturbação, para os casos de perturbação na variável resposta.

4.3 Influência local de Cook (1986)

Neste estudo, iremos abordar a influência local nos esquemas de perturbações descritos anteriormente e iremos na Seção 4.4 aplicar a metodologia de seleção apropriada de [Zhu et al. \(2007\)](#).

4.3.1 Perturbações sob o modelo normal

1. Ponderação de casos:

Na Figura 4.4, apresentamos o gráfico de avaliação de influência local no esquema de perturbação ponderação de casos, a fim de detectar os pontos influentes no modelo normal, neste caso o valor de C_{max} foi 35,35.

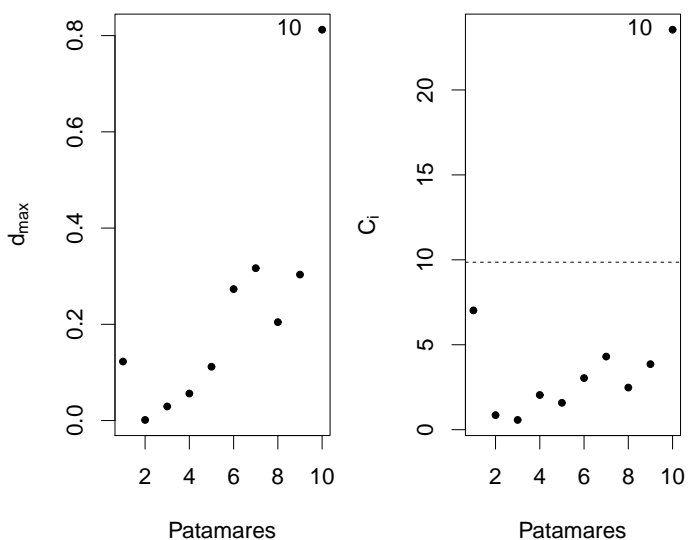


Figura 4.4 *Ponderação de casos*

O método da avaliação de influência local claramente detectou o patamar 10 como um patamar influente no modelo, pois o patamar se destacou no gráfico com um valor de C_{max} maior que dois.

2. Perturbação na variável resposta do laboratório de referência:

O gráfico de influência local sob a perturbação na variável resposta do laboratório de referência Figura 4.5 detectou o patamar 1 como influente, com um $C_{max} = 2,24$ maior que dois, concluí-se que a observação pode ser influente. Para acessar a influência individual de Cook (1986) de um i -ésimo laboratório ou um j -ésimo patamar, podemos construir o gráfico de LD como pode ser visto no ultimo gráfico da Figura 4.5. Considere um vetor \mathbf{d}_i ou \mathbf{d}_j , que consiste em um vetor de dimensão 1 por p ou m dependendo do caso adotado, com zeros em todas as posições exceto na i -ésima ou j -ésima posição, a qual usa-se -1. A partir disto, pode-se construir o gráfico de $LD(\boldsymbol{\omega}(a))$, com $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{d}_i$ ou $\boldsymbol{\omega}(a) = \boldsymbol{\omega}_0 + a\mathbf{d}_j$, variando-se os valores de a em um intervalo em torno de zero, por exemplo -1 e 1. No ultimo gráfico da Figura 4.5, temos o gráfico de $LD(\boldsymbol{\omega}(a)) = LD(\boldsymbol{\omega}_0 + a\mathbf{d}_1)$, onde o patamar 1 exerce bastante influência sobre as variações ocorridas na verossimilhança, pois as variações obtidas por $LD(\boldsymbol{\omega}(a)) = LD(\boldsymbol{\omega}_0 + a\mathbf{d}_{max})$ são bem próximas destas, como pode ser visto no gráfico. Logo, quando essas variações são bem inferiores em relação a direção \mathbf{d}_{max} , temos o indicativo de que a j -ésima observação não exerce uma grande influência.

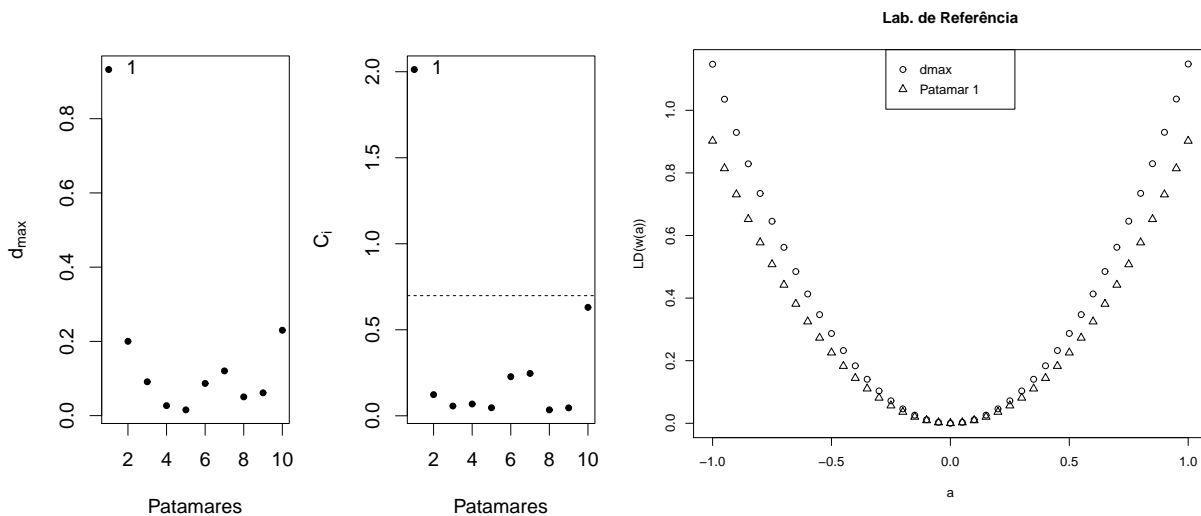


Figura 4.5 Perturbação da variável resposta do laboratório de referência.

3. Perturbação na variável resposta do i -ésimo laboratório:

Nesta seção, é apresentado a avaliação da influência local quando perturbado a variável resposta do i -ésimo laboratório.

Avaliando a Figura 4.6, quando perturbado o segundo laboratório, observa-se que em nenhuma das técnicas de influência foi detectado patamares influentes. Para este caso, o valor de C_{max} foi 23,78.

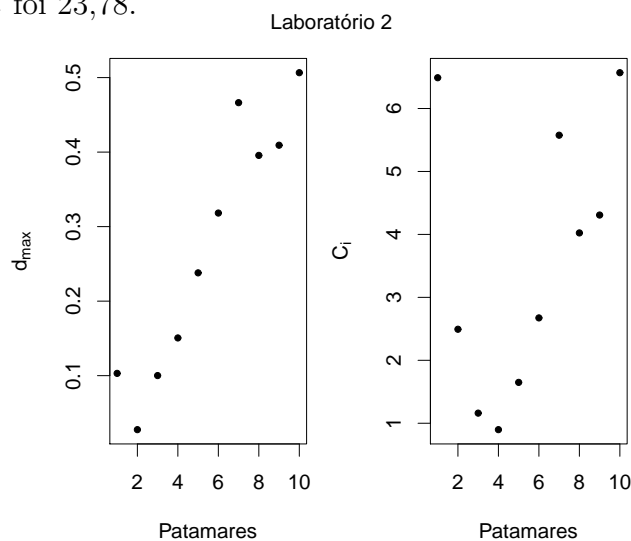


Figura 4.6 *Perturbação da variável resposta do segundo laboratório.*

Observando a Figura 4.7, quando a perturbação ocorre no terceiro laboratório, observa-se também que nenhum patamar se destacou nos gráficos de influência, onde o valor de C_{max} obtido foi 0,88.

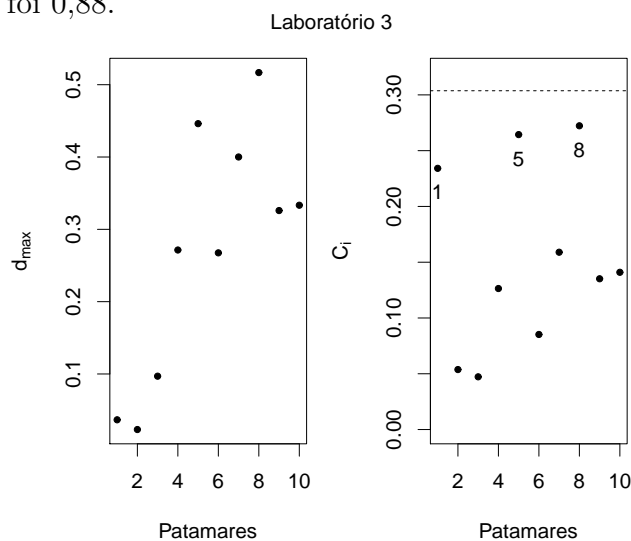


Figura 4.7 *Perturbação da variável resposta do terceiro laboratório.*

Estudando a influência local quando a perturbação ocorre no quarto laboratório,

observa-se que foi detectado o patamar 1 como possivelmente influente conjuntamente e individualmente, pois o valor de C_{max} sendo 5.13, entende-se que este patamar possivelmente é influente segundo os critérios de Cook (1986) que recomenda considerar influente valores de C_{max} maiores que 2.

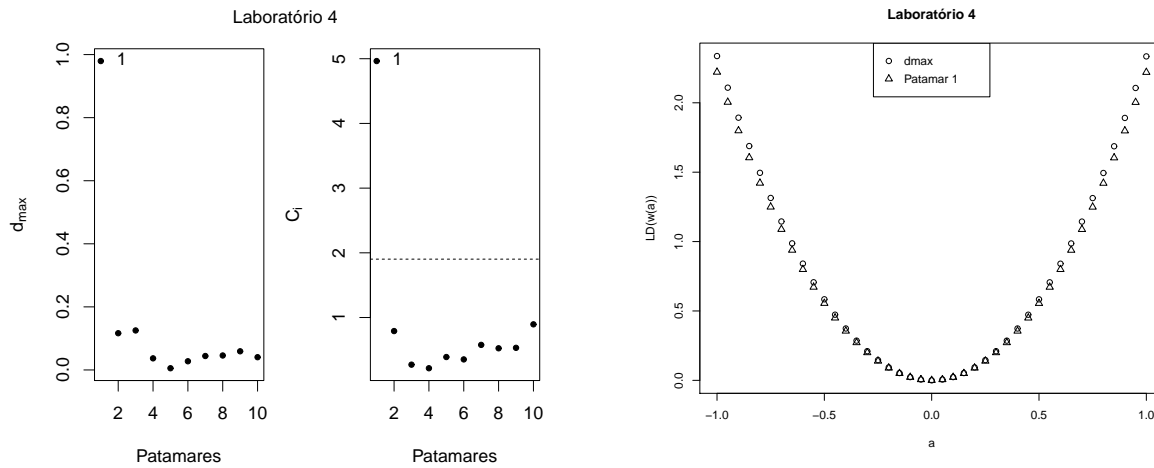


Figura 4.8 *Perturbação da variável resposta do quarto laboratório.*

Explorando a influência local quando a perturbação ocorre no quinto laboratório (Figura 4.9), observa-se que não foi detectado nenhum patamar influente conjuntamente ou individualmente (Cook (1986) e Lesaffre & Verbeke (1998), respectivamente). O valor de C_{max} foi igual a 0,38 e percebe-se neste caso que o valor de C_{max} foi bem baixo.

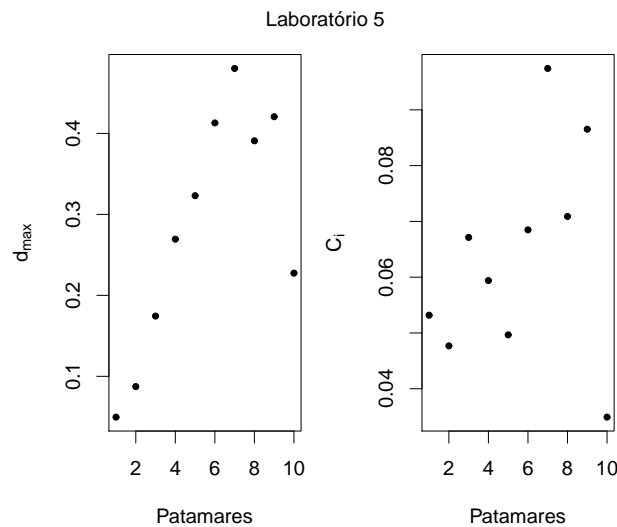


Figura 4.9 *Perturbação da variável resposta do quinto laboratório.*

Analisando a influência local quando a perturbação ocorre no sexto laboratório (Figura

4.10), observa-se que foi detectado o patamar 1 como influente conjuntamente e individualmente. Com um valor de C_{max} 3.49, entende-se que este patamar possivelmente é influente segundo os critérios de Cook (1986).

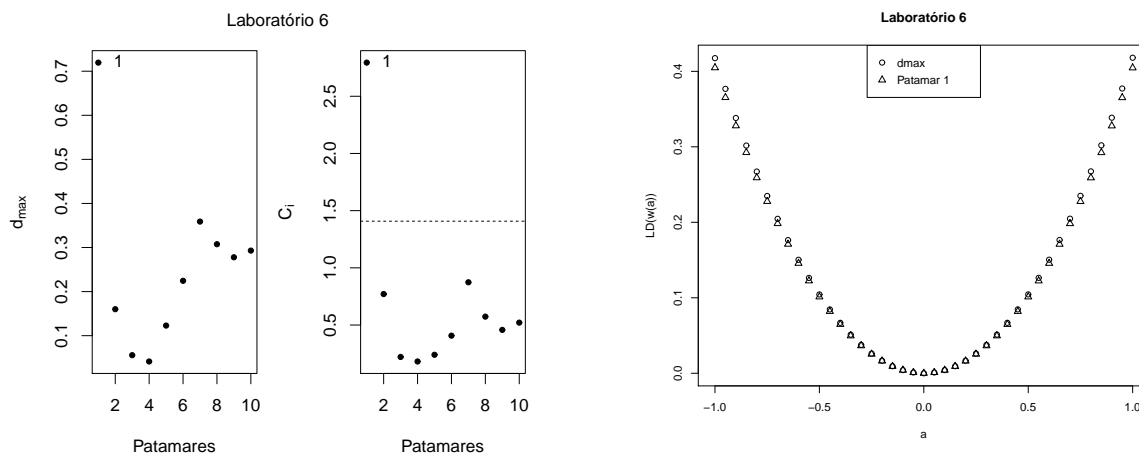


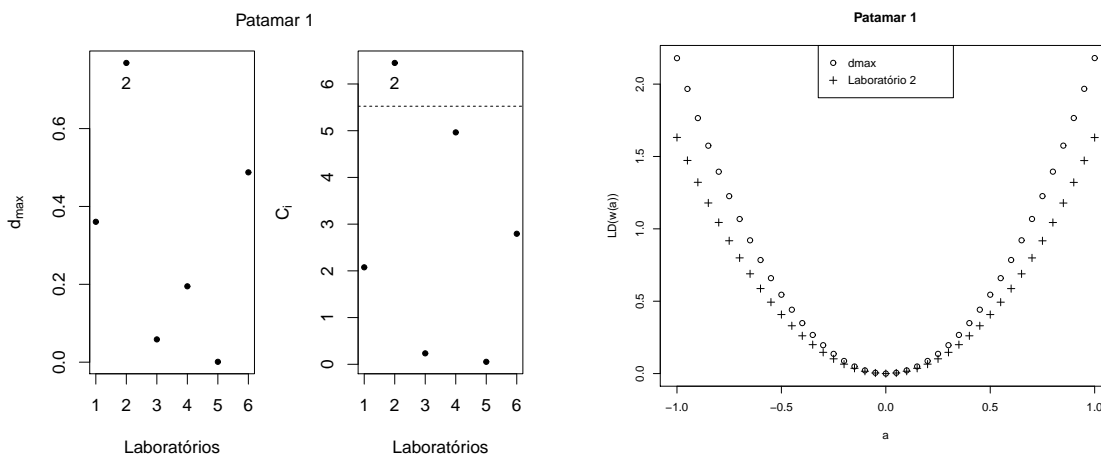
Figura 4.10 Perturbação da variável resposta do sexto laboratório.

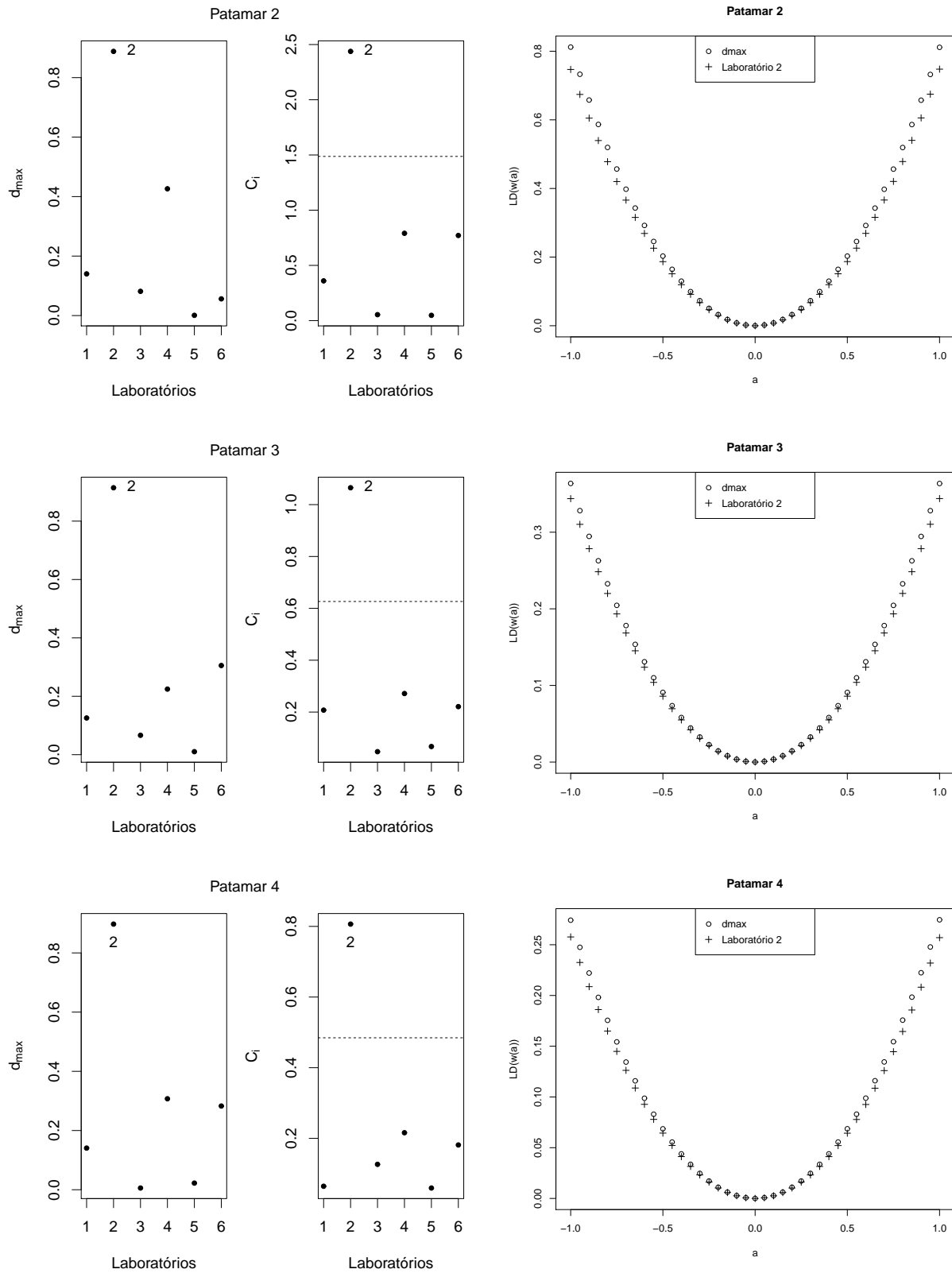
4. Perturbação na variável resposta de todos os laboratórios com o patamar fixo:

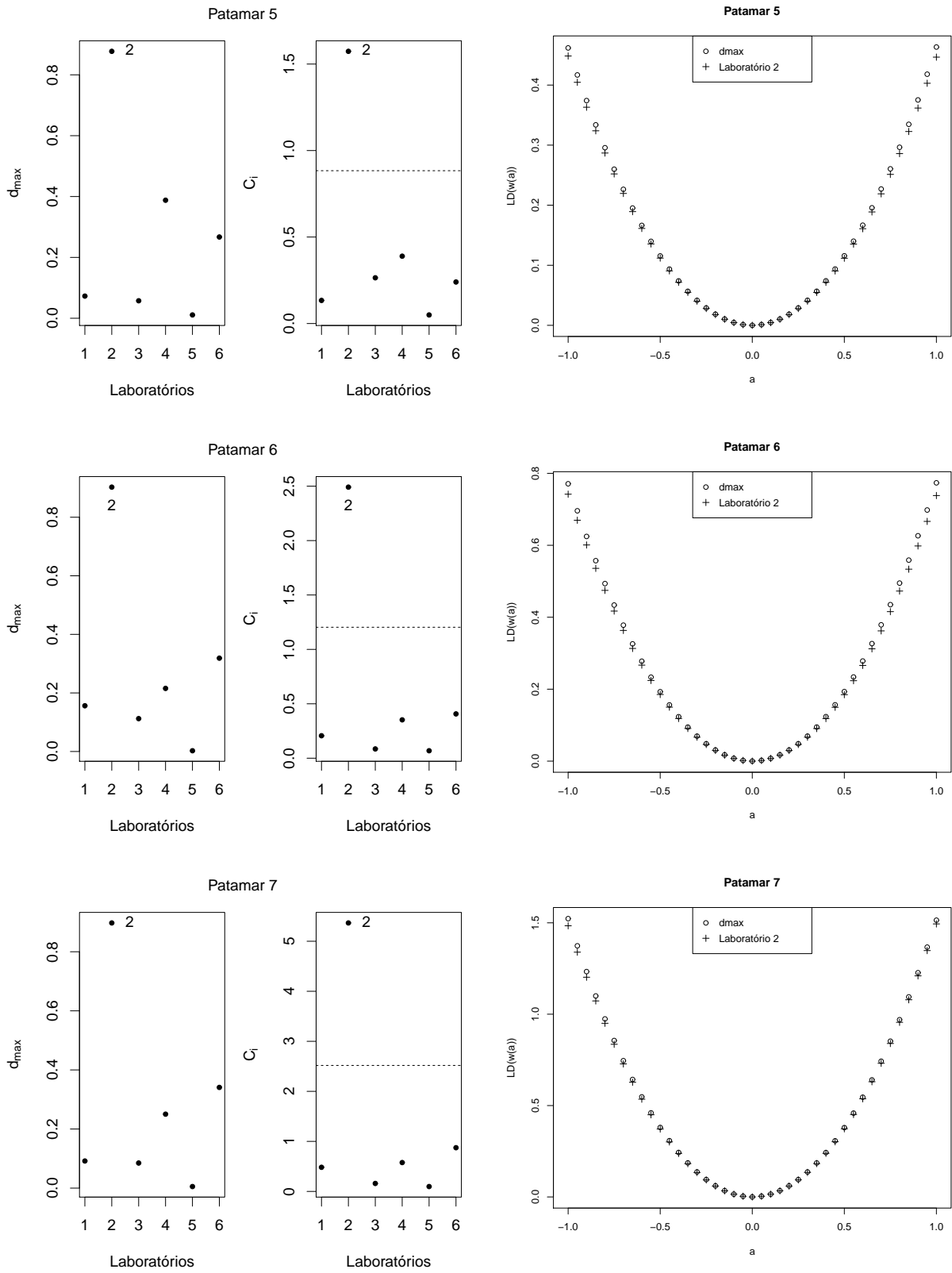
Neste caso identificou-se que o laboratório 2 foi detectado como influente em todos os patamares, e de acordo com os valores de C_{max} (tabela 4.7) pode-se concluir que estas observações possivelmente são influentes na maioria dos patamares.

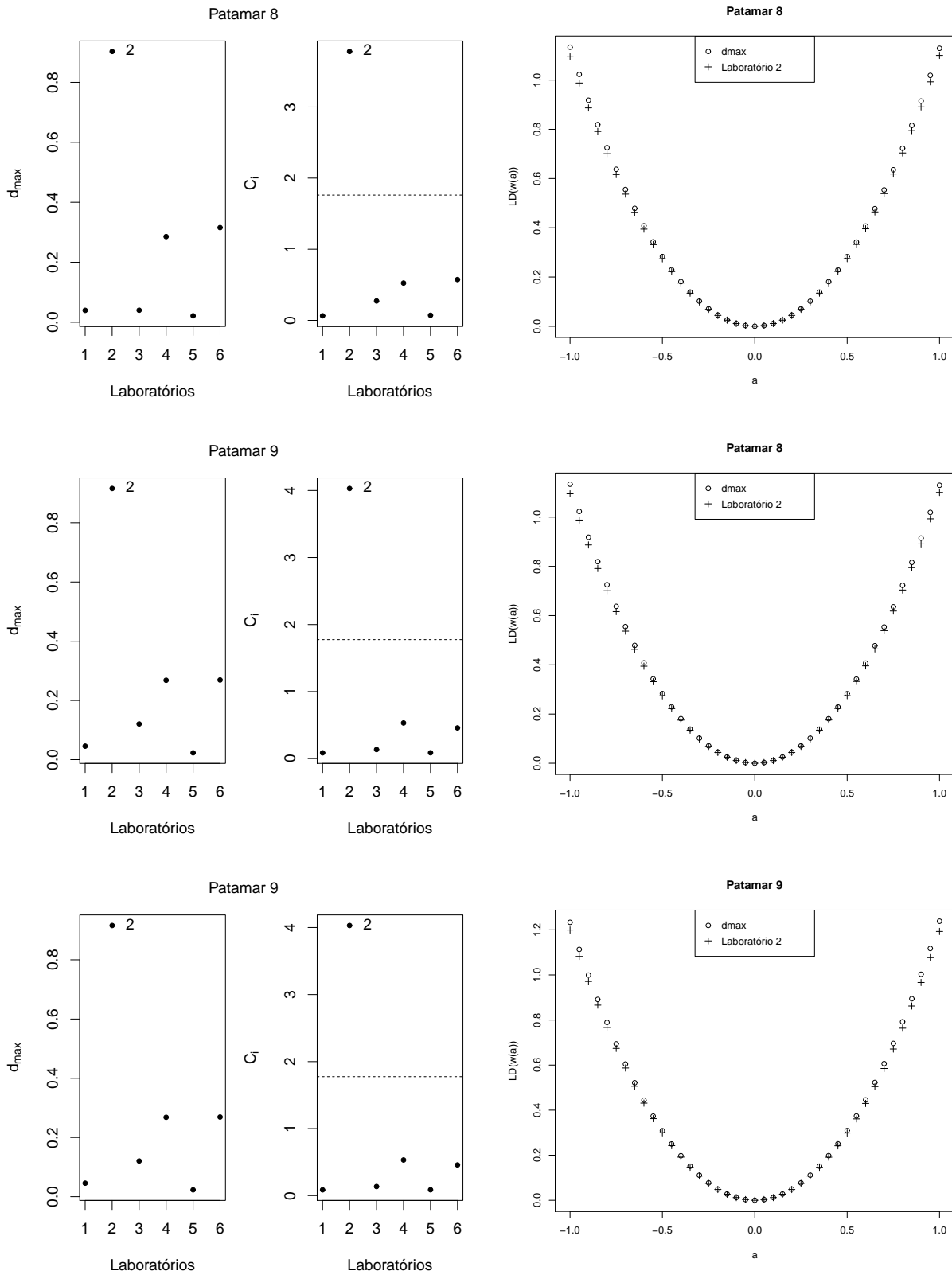
Tabela 4.7 Valores de C_{max} referente a perturbação em todos os laboratórios com o patamar fixo.

Patamares	1	2	3	4	5	6	7	8	9	10
Cmax	8,76	3,01	1,24	0,96	1,99	3,01	6,64	4,60	4,78	7,87









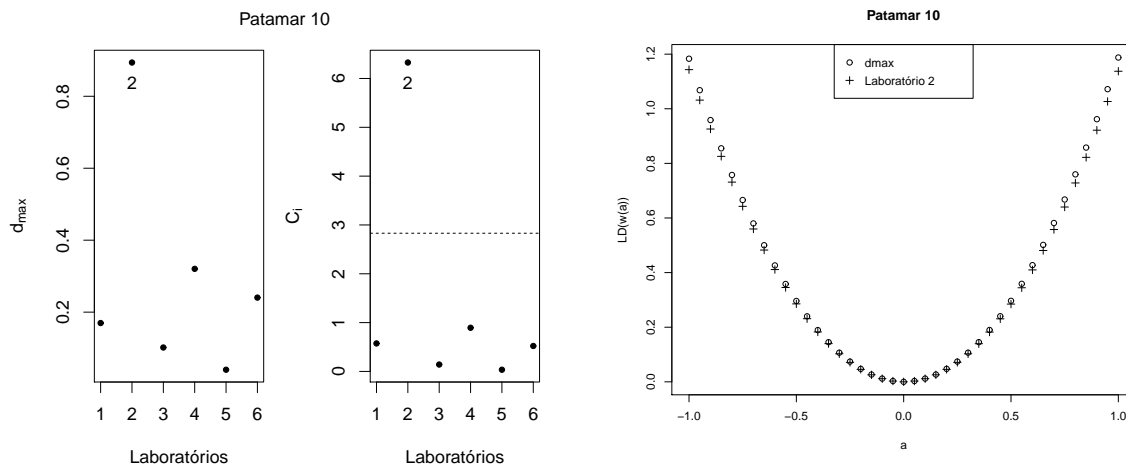


Figura 4.11 *Perturbação na variável resposta no j -ésimo patamar.*

4.3.2 Perturbações sob o modelo t-Student

1. Ponderação de casos:

A influência local, no esquema de perturbação ponderação de casos no modelo t-Student, obteve um C_{max} de 18,94. Nota-se na Figura 4.12 o gráfico de influência local resultante da aplicação da metodologia. Observa-se que em ambas as técnicas de avaliações utilizadas nenhum patamar se destacou nos gráficos de influência. Neste caso, o modelo t de Student com a perturbação no modelo é mais robusto que o modelo normal que foi sensível e teve suas inferências afetadas pelo patamar 10. Observou-se também que o valor de C_{max} neste caso é menor do que o C_{max} obtido no caso normal.

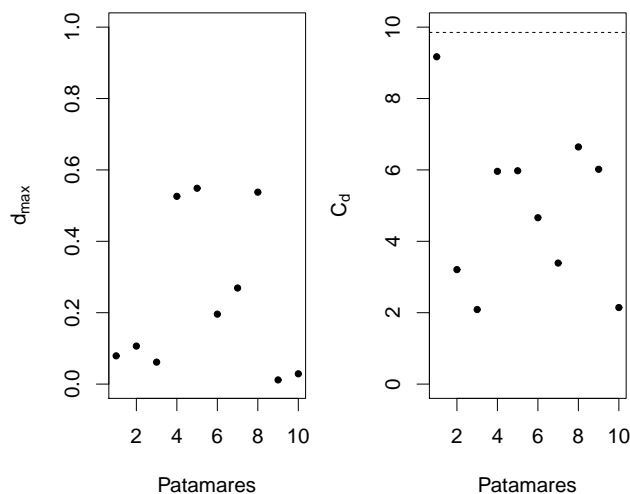


Figura 4.12 *Ponderação de casos*

2. Perturbação na variável resposta do laboratório de referência:

A técnica de influência local neste esquema de perturbação detectou o patamar 1 como influente como visto na Figura 4.13, porém com um valor de $C_{max} = 0,27$, entende-se com base na metodologia proposta em Cook (1986), que neste caso, o modelo t-Student é robusto e o patamar 1 não deve exercer influências.

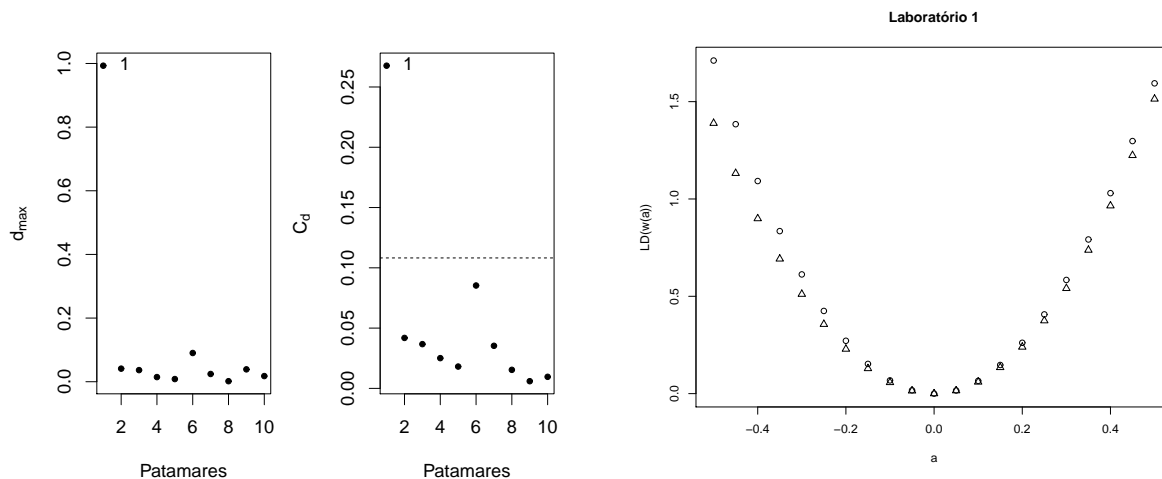


Figura 4.13 *Perturbação na variável resposta do laboratório de referência*

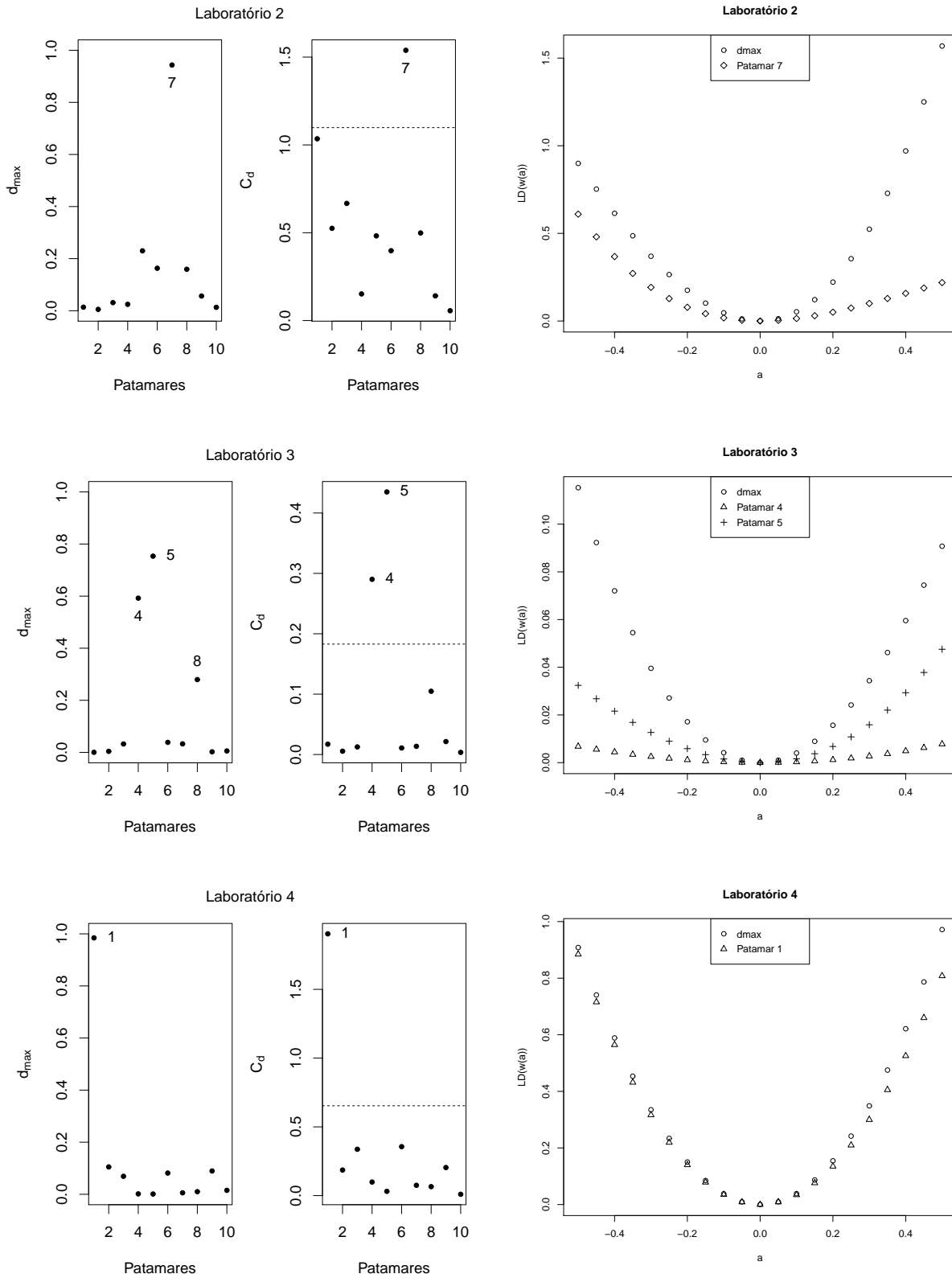
Note que o efeito causado ao perturbar na direção máxima se deve praticamente à observação do patamar 1, pois a dispersão dos pontos ficam bem próximas existindo uma diferença bem pequena.

3. Perturbação na variável resposta do i -ésimo laboratório:

Neste esquema, observa-se que há laboratórios que se destacam nos gráficos de influência. Porém, pode-se observar, que os valores de C_{max} na tabela 4.8 são menores que dois e segundo o critério proposto por Cook (1986) é o indicativo de que os patamares destacados nos gráficos não sejam influentes. Vale salientar que os valores de C_{max} neste caso, são menores que os obtidos neste mesmo esquema de perturbação no modelo normal em todos os casos.

Tabela 4.8 *Valores de C_{max} referente a perturbação no i -ésimo laboratório.*

Laboratórios	2	3	4	5	6
C_{max}	1,66	0,72	1,95	0,01	0,47



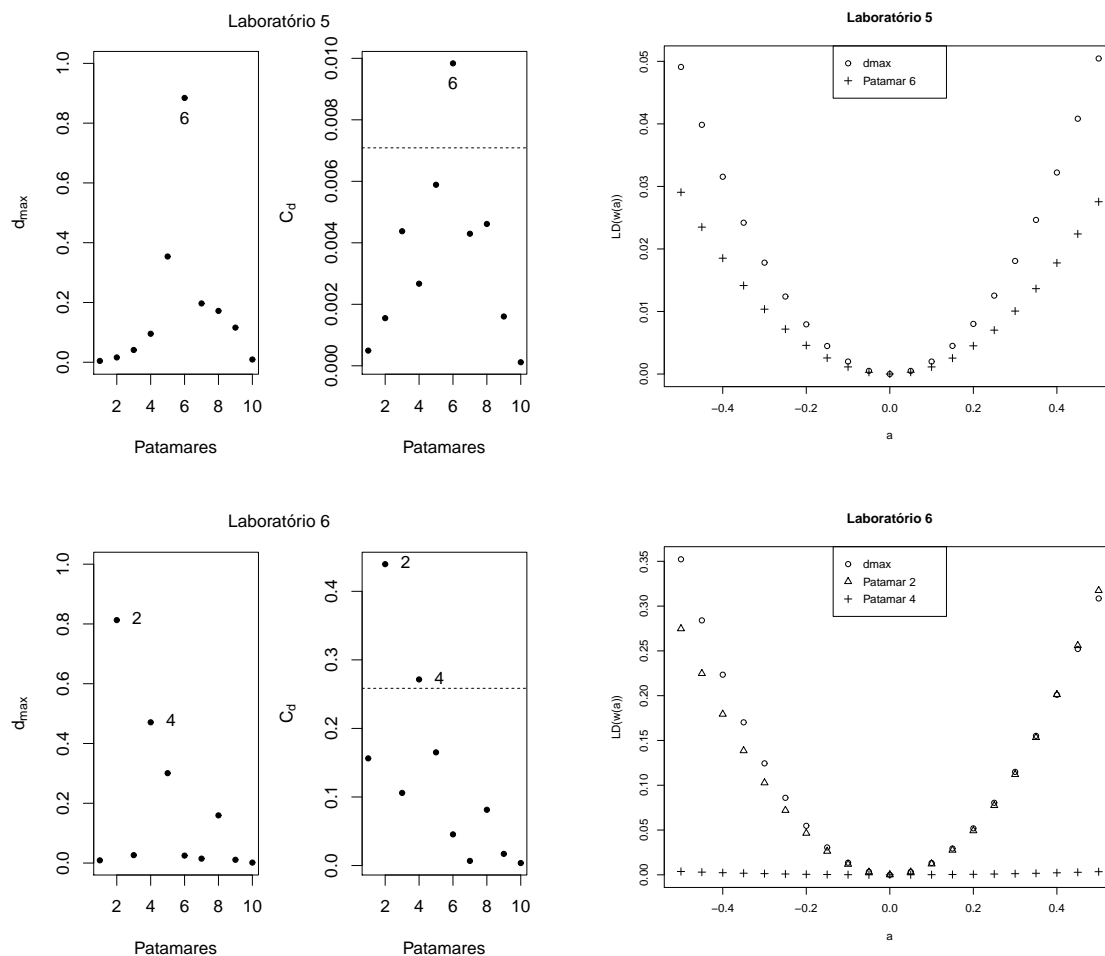


Figura 4.14 Gráficos de influência e $LD(w(a))$ quando a perturbação é no i -ésimo laboratório.

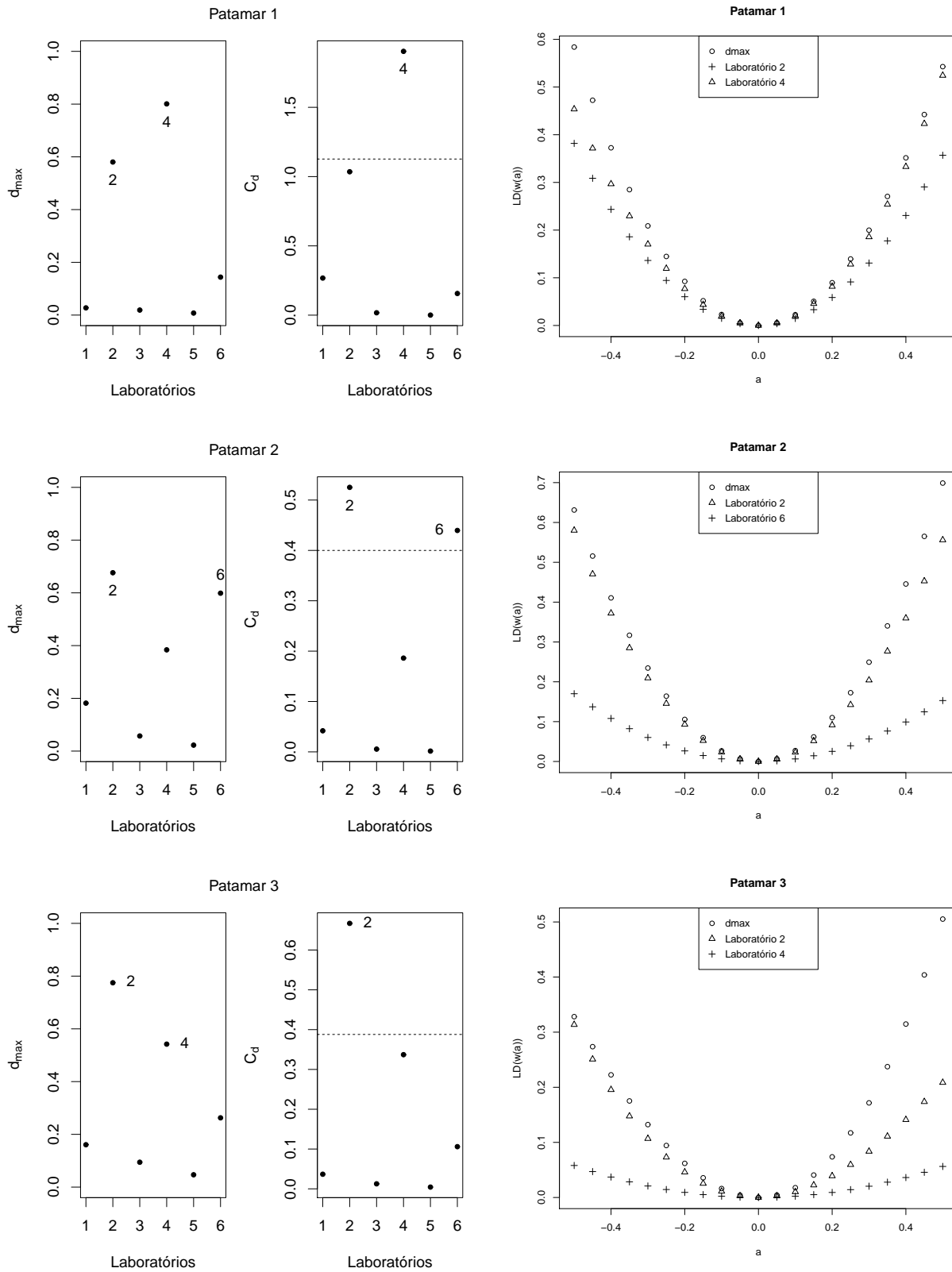
4. Perturbação na variável resposta de todos os laboratórios com o patamar

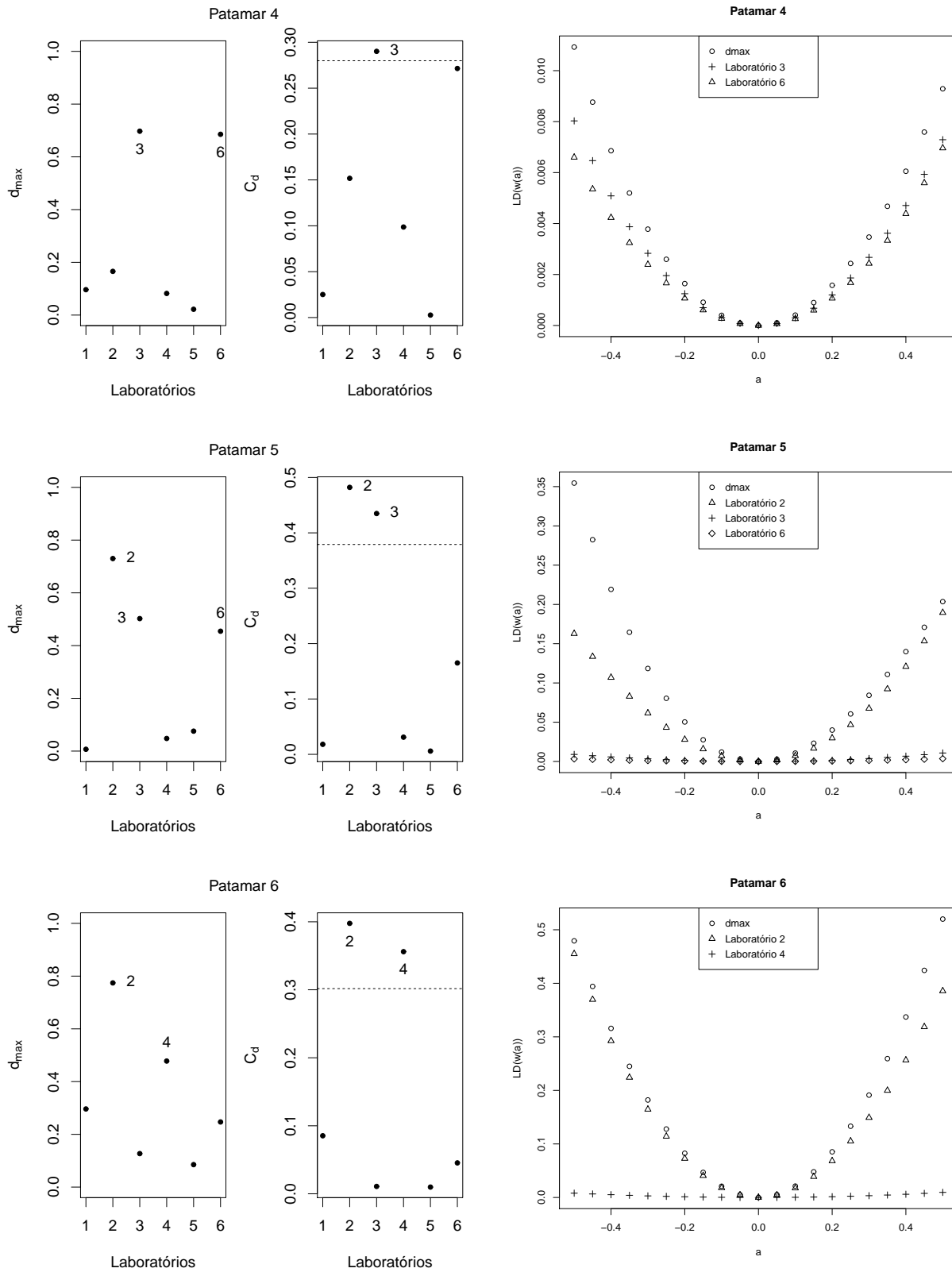
fixo:

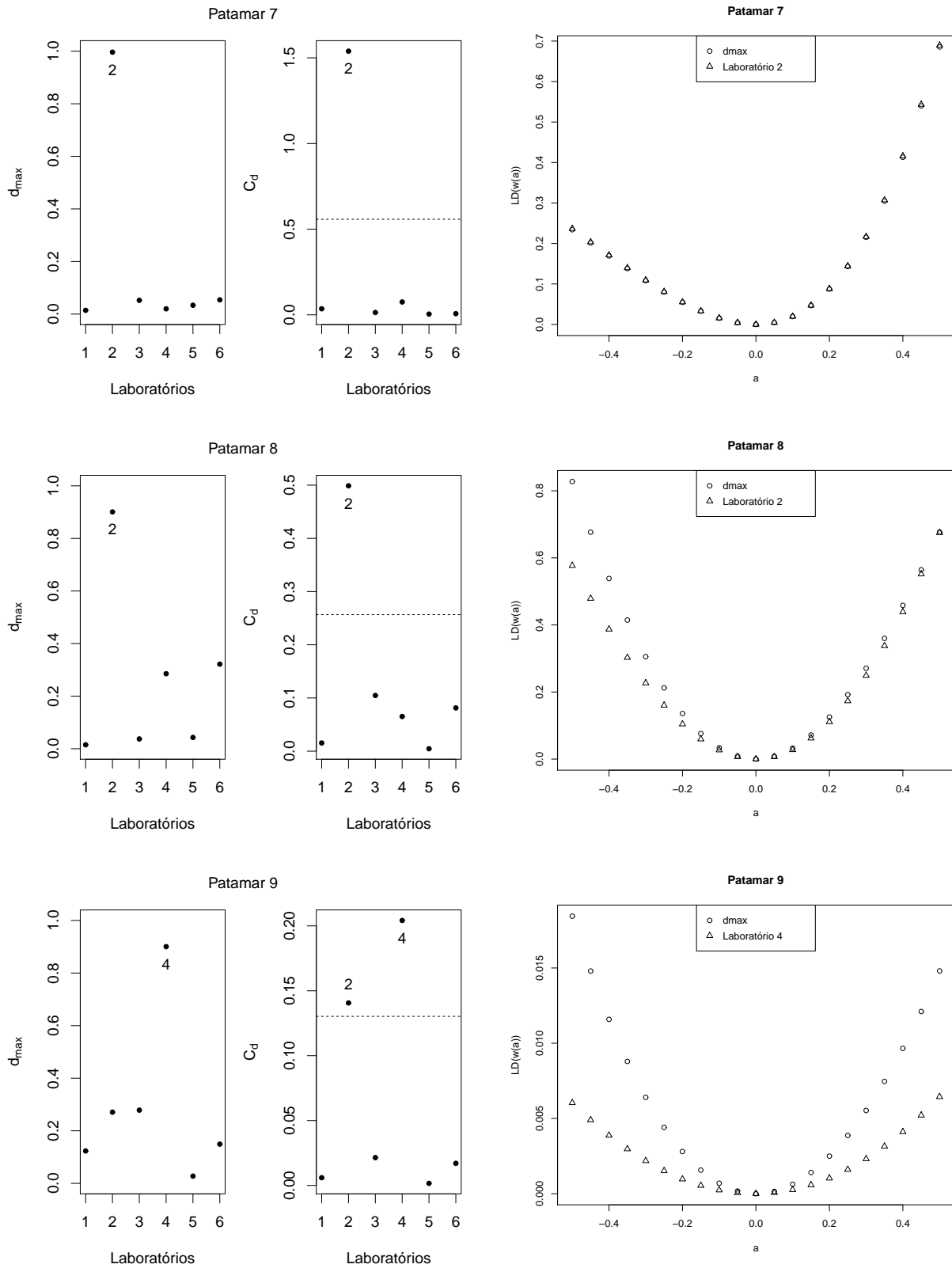
Neste esquema, observa-se que foi detectado alguns laboratórios influentes. Porém os valores de C_{max} na tabela 4.9 são menores que dois exceto no patamar 1, onde o C_{max} é 2,94, e possivelmente estas observações são influentes. Vale salientar que todos os valores de C_{max} neste esquema de perturbação foram menores que os valores obtidos de C_{max} do modelo normal. Como discutido nos casos anteriores, a interpretação do gráfico $LD(\omega)$ são semelhantes para todos os casos a seguir.

Tabela 4.9 C_{max} na perturbação em todos os laboratórios após o uso da correção apropriada.

Patamares	1	2	3	4	5	6	7	8	9	10
C_{max}	2,94	1,08	1,04	0,50	0,75	0,59	1,55	0,61	0,22	0,07







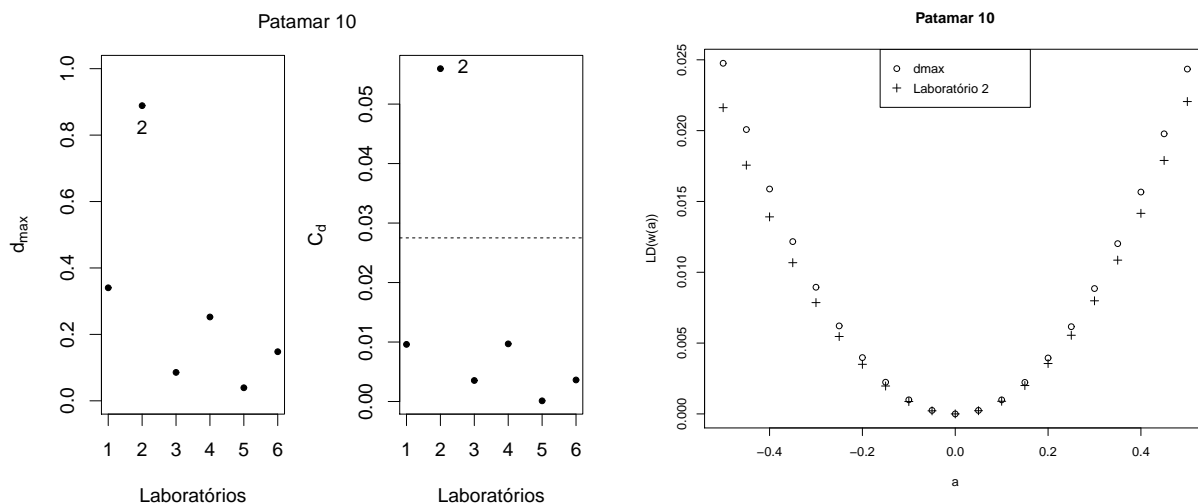


Figura 4.15 Gráficos de influência e $LD(w(a))$ quando a perturbação é no j -ésimo patamar.

4.3.3 Conclusão

A técnica de diagnóstico de influência local de Cook (1986) foi aplicada no modelo de calibração ultraestrutural com réplicas balanceadas sob as distribuições normal e t-Student. Comparando-se os resultados obtidos nas Seções 4.3.1 e 4.3.2, o modelo t-Student se mostrou mais robusto que o modelo normal, pois apesar de ambos terem tido observações que influenciavam o modelo, o modelo t-Student apresentou valores de C_{max} bem menores que o modelo normal em todos os casos. Além disso, o valor do C_{max} foi maior do que dois em somente dois tipos de perturbações, enquanto que no caso normal, o valor do C_{max} deu maior do que dois em doze tipos de perturbações, incluindo as duas perturbações onde o C_{max} deu maior do que dois no modelo t-Student. Desta forma, segundo o critério de Cook (1986), conclui-se que o modelo normal é mais sensível que o modelo t-Student.

4.4 Método para a seleção apropriada

Nesta seção vamos aplicar a metodologia de Zhu et al. (2007) como visto no Capítulo 3, considerando a perturbação na variável resposta do modelo normal e propor uma nova perturbação.

4.4.1 Perturbação na variável resposta do laboratório de referência

Através da matriz $G(\omega_0)$ abaixo podemos ver que as propriedades não foram satisfeitas e conseqüentemente a utilização da técnica de Cook (1986) não é razoável para investigar

a sensibilidade a pequenas perturbações. Note os elementos circulos na matriz $G(\omega_0)$. Estes elementos tem pesos maiores que os demais elementos. Estes elementos em destaque são responsáveis por aumentar a chance dos patamares 1 e 10 a se destacarem na técnica de influência local.

$$G(\omega_0) = \begin{pmatrix} \textcircled{1,62} & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,24 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,20 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,29 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,17 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,74 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,72 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,09 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,12 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & \textcircled{1,71} \end{pmatrix}_{10 \times 10}$$

Na Figura 4.16, após correção, observa-se que o primeiro patamar continua sendo detectado pela metodologia de Lesaffre & Verbeke (1998) como influente individualmente, porém na metodologia de Cook (1986) nenhum patamar foi detectado. Neste caso o valor obtido para o C_{max} foi 2,21.

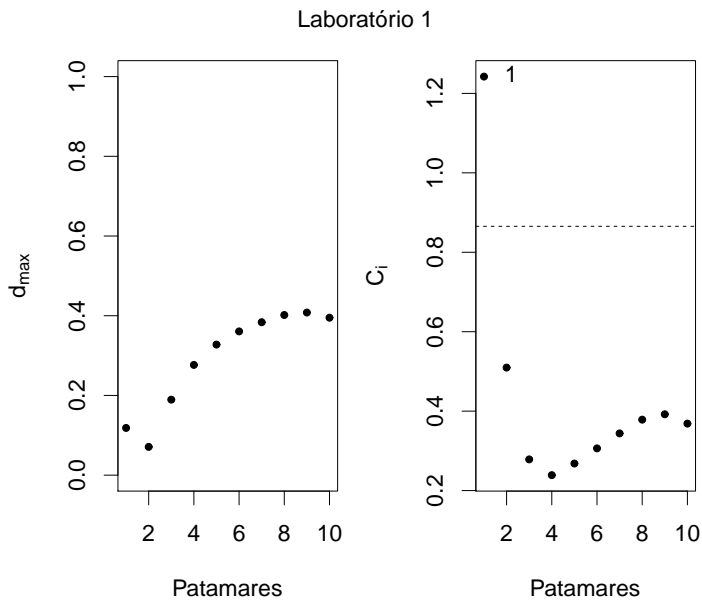


Figura 4.16 Perturbação induzida da variável resposta do laboratório de referência.

4.4.2 Perturbação na variável resposta do i-ésimo laboratório

Nesta seção estudamos a metodologia de [Zhu et al. \(2007\)](#), considerando a perturbação na variável resposta do i-ésimo laboratório, onde para cada caso será apresentado a matriz $G(\omega_0)$, e assim poderemos tirar conclusões e observar o que a matriz está indicando.

Analisando a perturbação no segundo laboratório, Figura 4.6, nenhum patamar tinha sido detectado. Porém, analisando a matriz $G(\omega_0)$ deste caso, vemos que os patamares 7 e 10 tem pesos maiores que os demais e estes pesos produzem um efeito nestes patamares, como pode ser observado na Figura 4.6.

$$G(\omega_0) = \begin{pmatrix} 4,23 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 3,68 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 2,95 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 2,68 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 4,52 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 6,19 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & \mathbf{12,57} & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 7,94 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 8,20 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & \mathbf{14,88} \end{pmatrix}_{10 \times 10}.$$

Nesta nova abordagem, usando uma perturbação adequada para acessar a influência local, a proposta de [Cook \(1986\)](#) não identificou nenhum patamar influente, embora a metodologia proposta por [Lesaffre & Verbeke \(1998\)](#) tenha destacado o patamar 1 como influente individualmente como pode ser visto na Figura 4.17. Neste caso o C_{max} foi igual a 2.81, sendo um valor bem inferior quando comparado com o valor de C_{max} desta mesma perturbação mas sem o uso da correção apropriada.

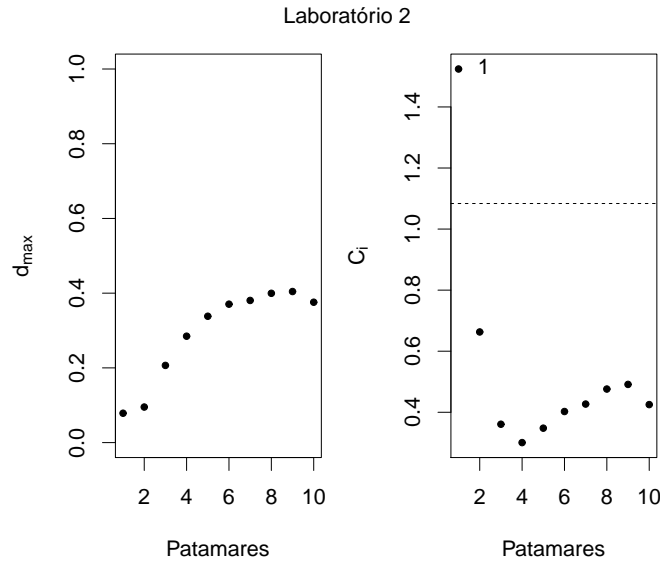


Figura 4.17 Perturbação induzida da variável resposta do segundo laboratório.

Analisando a perturbação no terceiro laboratório Figura 4.7, nenhum patamar tinha sido detectado. Porém, analisando a matriz $G(\omega_0)$ deste caso, vemos que o patamar 5 tem peso maior que os demais. E este peso produz um efeito neste patamar, sendo que não foi suficiente para que esta observação tenha sofrido uma alavancagem e assim fosse detectada pela técnica de influência local de Cook (1986), como pode ser observado na Figura 4.7.

$$G(\omega_0) = \begin{pmatrix} 0,19 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,05 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,12 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,54 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & \mathbf{1,02} & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,15 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,31 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,66 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,15 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,14 \end{pmatrix}_{10 \times 10}.$$

Nesta nova abordagem, usando uma perturbação adequada para acessar a influência local,

as propostas de influência local não identificaram nenhum patamar influente, como pode ser visto na Figura 4.18. Neste caso o valor do C_{max} obtido foi 3,33.

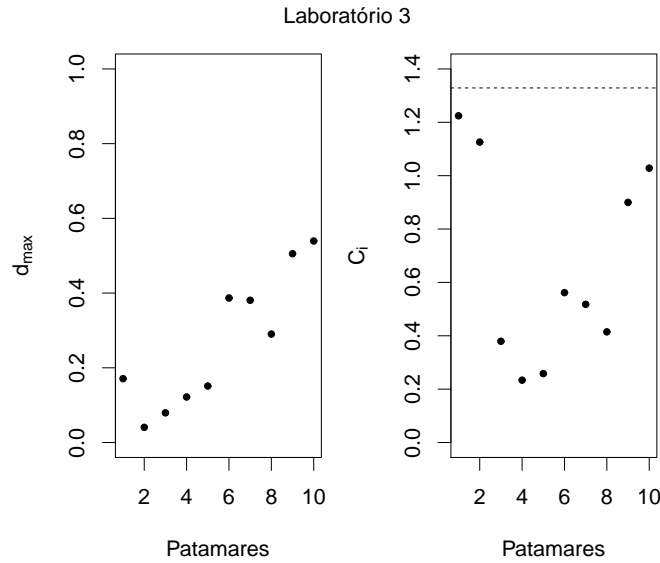


Figura 4.18 Perturbação induzida da variável resposta do terceiro laboratório.

Analisando a perturbação no quarto laboratório Figura 4.8, o patamar 1 tinha sido detectado como influente. Porém, ao analisar a matriz $G(\omega_0)$ neste caso, vemos que o patamar 1 tem peso maior que os demais. Isto implica que o peso produz um efeito ao acessar a influência local neste patamar, fazendo com que conclusões equivocadas a respeito do diagnóstico possam ser tomadas.

$$G(\omega_0) = \begin{pmatrix} \mathbf{3,49} & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,29 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,60 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,43 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,12 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,61 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,38 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,33 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,45 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,15 \end{pmatrix}_{10 \times 10}.$$

Verificando a matriz $G(\omega_0)$ note que os pesos são diferentes para cada patamar havendo a necessidade de usar uma perturbação correta para corrigir e estudar a influência local de forma adequada. Após a correção utilizando uma perturbação apropriada, o patamar 10 foi detectado como influente na Figura 4.19. Vale salientar que o valor do C_{max} foi 13.40, e quando comparado com o C_{max} deste mesmo caso sem o uso da correção, o C_{max} foi menor.

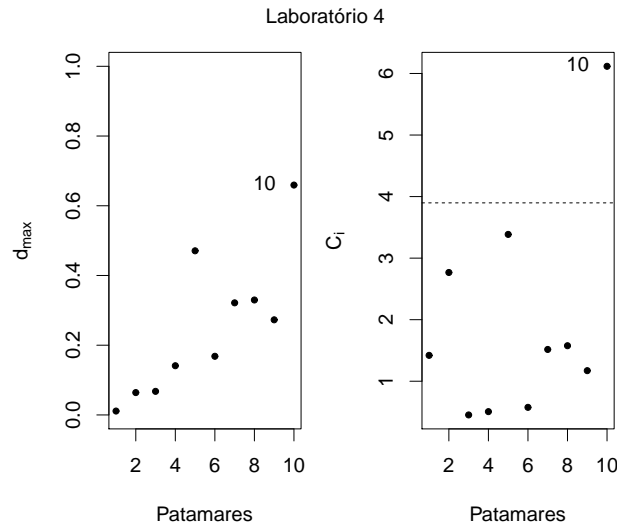


Figura 4.19 Perturbação induzida da variável resposta do quarto laboratório.

Explorando a perturbação no quinto laboratório na Figura 4.9, nenhum dos patamares tinha sido detectado como influente. Observando a matriz $G(\omega_0)$ neste caso, vemos que os pesos entre os patamares não são muito diferentes entre si, apesar de serem diferentes.

$$G(\omega_0) = \begin{pmatrix} 0,07 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,09 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,18 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,24 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,23 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,28 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,29 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,17 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,18 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,05 \end{pmatrix}_{10 \times 10}.$$

Após a correção na metodologia nenhum patamar se destacou no gráfico de influência local como pode ser observado na Figura 4.20 e o C_{max} obtido foi igual a 2,18.

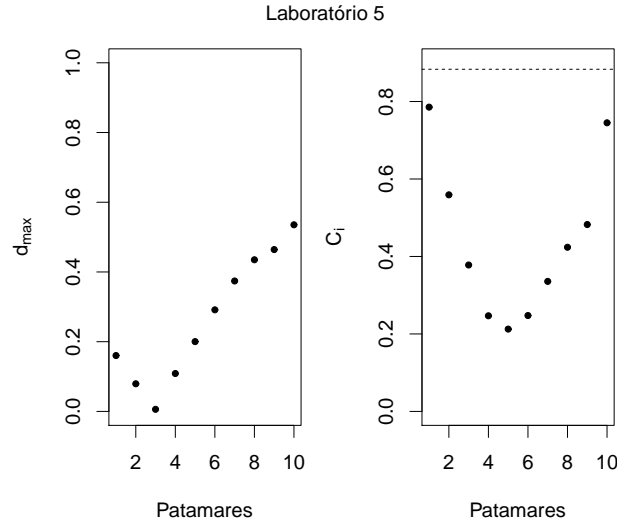


Figura 4.20 Perturbação induzida da variável resposta do quinto laboratório.

Investigando a perturbação no sexto laboratório na Figura 4.10, o patamar 1 tinha sido detectado como influente. Observando a matriz $G(\omega_0)$ neste caso, percebe-se que o patamar 2 tem peso maior e depois o patamar 1. Como visto neste caso, onde a componente 2 tem maior efeito, não é indicativo de que o respectivo patamar será influente, pois, o efeito produzido no patamar 2 não teve impacto a ponto de perturbar o patamar 2 no gráfico do d_{max} na Figura 4.10. Porém analisando os pesos na matriz $G(\omega_0)$ vemos que são diferentes e conseqüentemente não satisfazem as condições para que a perturbação seja adequada.

$$G(\omega_0) = \begin{pmatrix} \textcircled{0,69} & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & \textcircled{1,26} & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,24 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,35 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,29 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,19 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,04 & 0,00 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,15 & 0,00 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,22 & 0,00 & 0,00 \\ 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,13 & 0,00 \end{pmatrix}_{10 \times 10}.$$

Veja que após a correção feita na metodologia de influência local, o patamar 7 se destaca. Vale salientar que existe uma discordância entre os estudos com e sem correção, e caso não fosse utilizado a metodologia de [Zhu et al. \(2007\)](#), não seria possível acessar a influência local corretamente. Portanto, conclui-se que o patamar 7 pode ser influente conjuntamente e individualmente, com valor de C_{max} igual a 32,40.

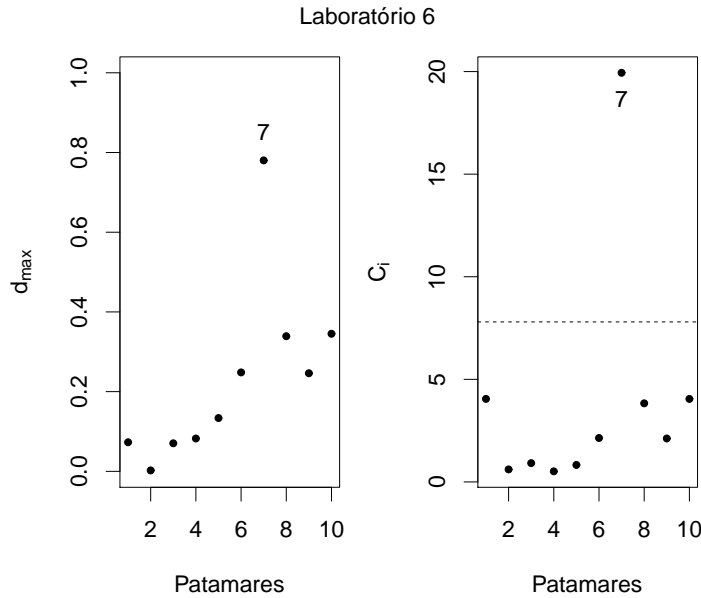


Figura 4.21 *Perturbação induzida da variável resposta do sexto laboratório.*

4.4.3 Perturbação nas variáveis respostas do j -ésimo patamar

Neste esquema, até o momento, a influência local tinha sido realizada a partir de uma perturbação arbitrária. A seguir apresentamos o estudo de influência local para o caso em que a seleção da perturbação foi induzida a partir de uma perturbação apropriada como mencionado no artigo proposto por [Zhu et al. \(2007\)](#), pois após obtido a matriz $G(\omega_0)$ foi constatado que o efeito em todos os patamares não são os mesmos. Vale salientar que na matriz $G(\omega_0)$ de cada patamar, o laboratório 2 tem um peso bem maior que os demais laboratórios e isto é o indicativo forte de interferência na influência local vista na Figura 4.11, onde este laboratório foi influente em todos os patamares. Na tabela 4.10, encontram-se os valores de C_{max} referente a nova influência.

Tabela 4.10 C_{max} na perturbação em todos os laboratórios após o uso da correção apropriada.

Patamares	1	2	3	4	5	6	7	8	9	10
Cmax	2,10	1,03	0,60	0,44	0,45	0,51	0,60	0,68	0,74	0,76

Dando continuidade, a seguir é apresentado as matrizes $G(\omega_0)$ e os gráficos de influência para os respectivos patamares.

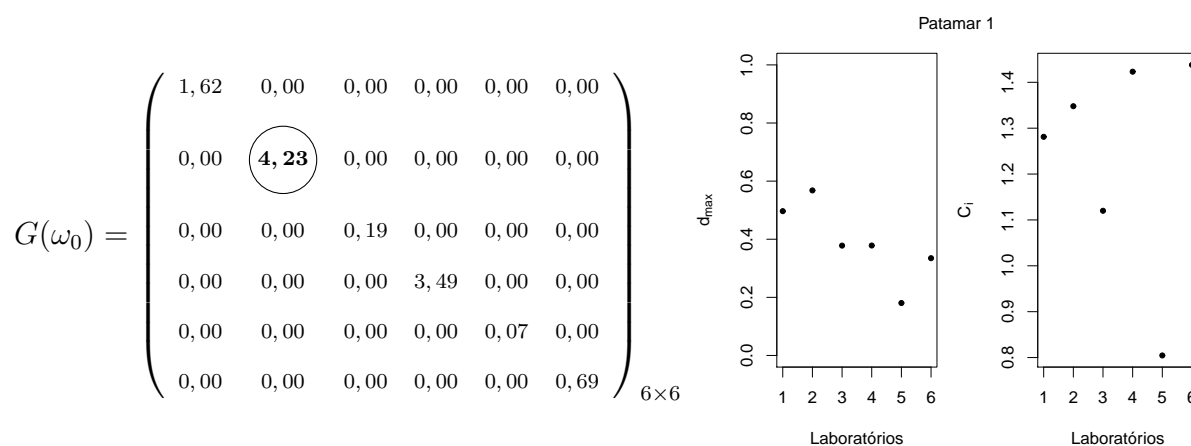


Figura 4.22 Perturbação em todos os laboratórios do patamar 1 após a correção adequada.

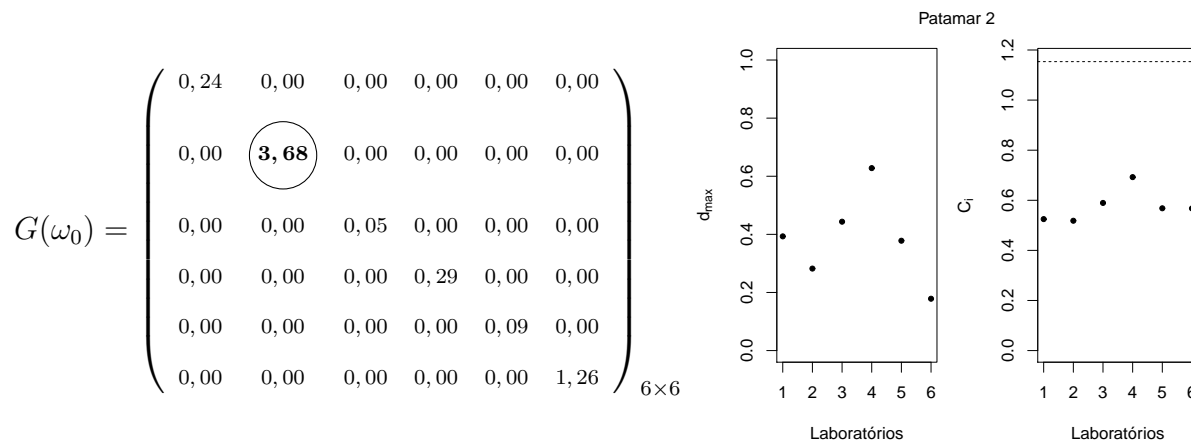


Figura 4.23 Perturbação em todos os laboratórios do patamar 2 após a correção adequada.

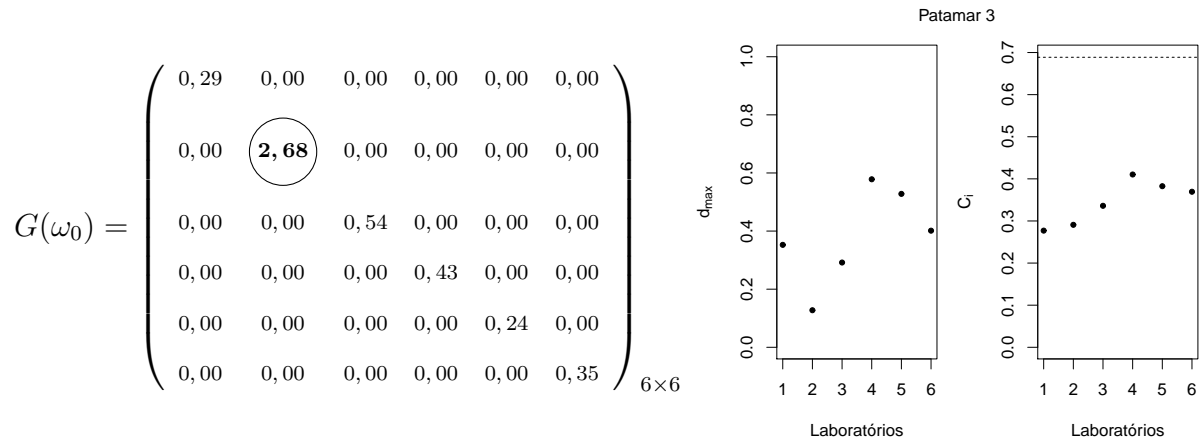


Figura 4.24 Perturbação em todos os laboratórios do patamar 3 após a correção adequada.

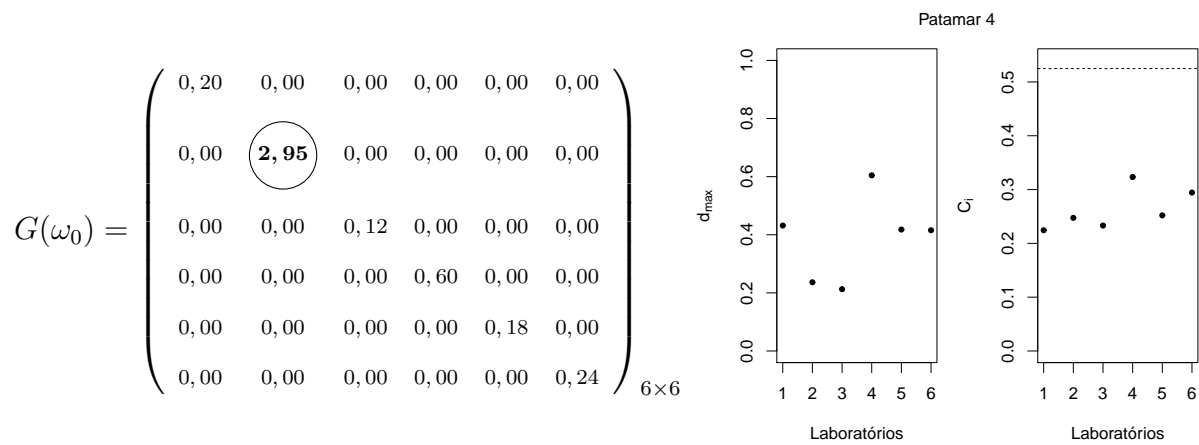


Figura 4.25 Perturbação em todos os laboratórios do patamar 4 após a correção adequada.

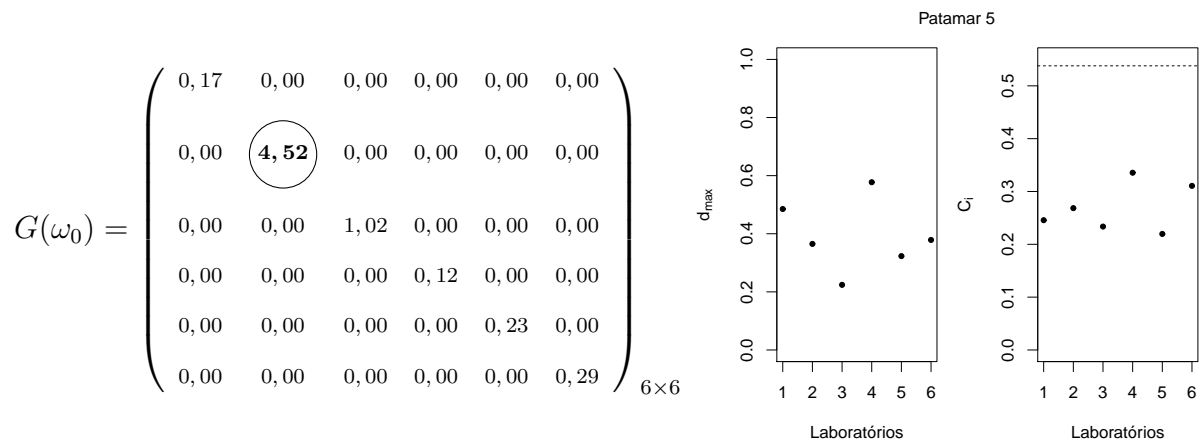


Figura 4.26 Perturbação em todos os laboratórios do patamar 5 após a correção adequada.

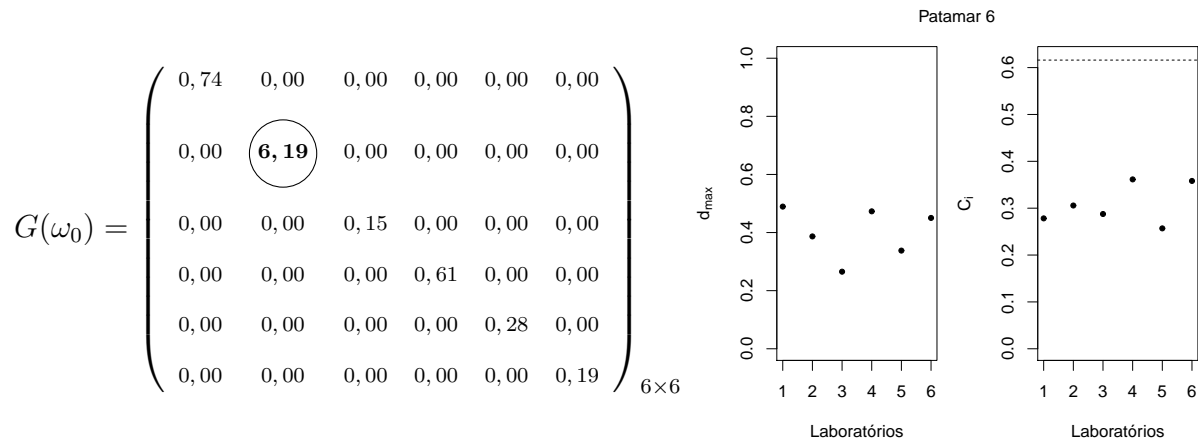


Figura 4.27 Perturbação em todos os laboratórios do patamar 6 após a correção adequada.

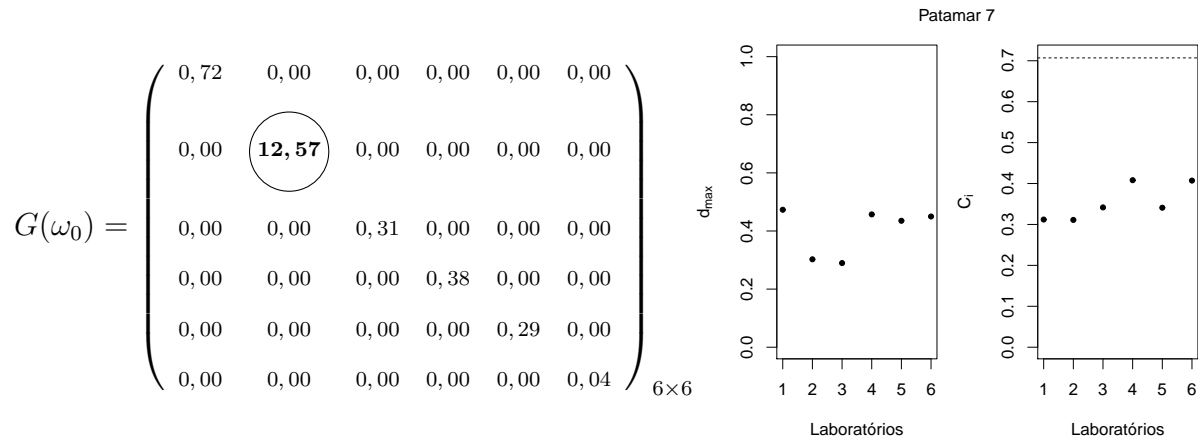


Figura 4.28 Perturbação em todos os laboratórios do patamar 7 após a correção adequada.

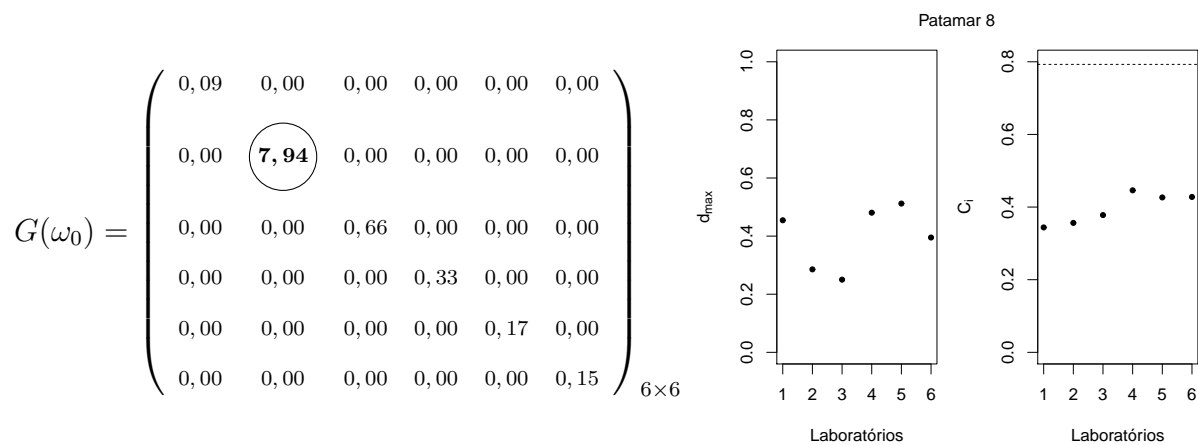


Figura 4.29 Perturbação em todos os laboratórios do patamar 8 após a correção adequada.

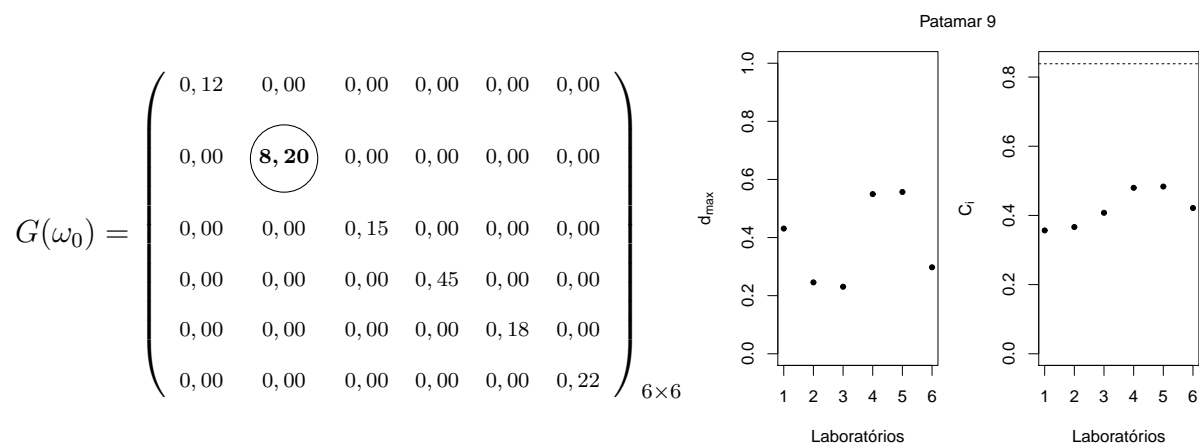


Figura 4.30 *Perturbação em todos os laboratórios do patamar 9 após a correção adequada.*

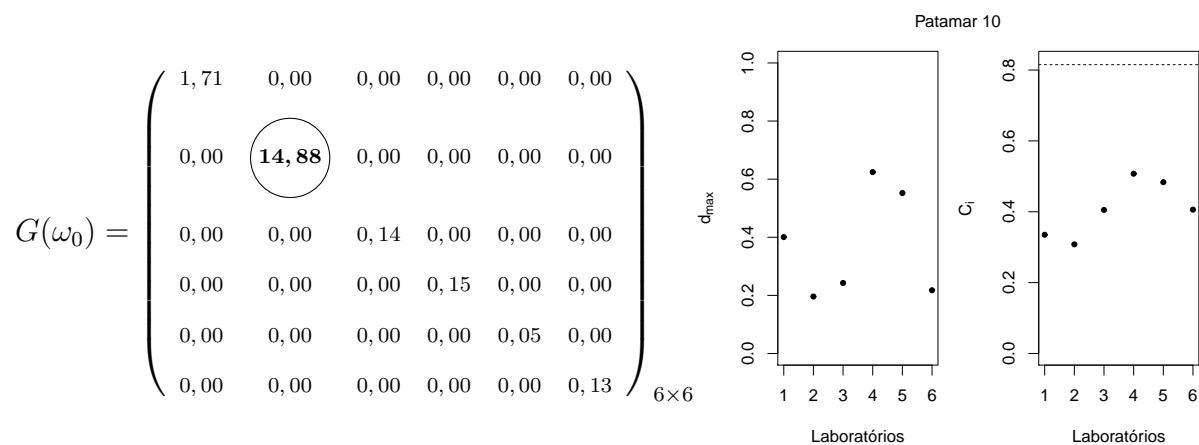


Figura 4.31 *Perturbação em todos os laboratórios do patamar 10 após a correção adequada.*

Analisando esses resultados de forma geral observa-se que, após igualar os pesos aplicados para cada laboratório, o laboratório 2 não sobressai como um ponto influente. Observando os gráficos, note que em nenhum patamar houve a detecção de laboratórios influente, tanto conjuntamente quanto individualmente. Pode-se observar que nenhum laboratório é influente no esquema onde se perturba todos os laboratórios ao mesmo tempo em cada patamar específico.

Vale salientar que sem essa proposta da correção, todas as conclusões e interpretações seriam equivocadas e esta informação vale para todos os casos abordados neste trabalho.

5

Conclusão

Neste trabalho, foi desenvolvido o estudo de diagnóstico de influência local de [Cook \(1986\)](#), esta técnica foi aplicada nas modelagem normal e t-Student. Foi feito uma análise exploratória dos dados e mudança relativa, onde percebemos que observações como patamar 1 e 10, foram identificadas como globalmente influentes, isso se deve pelo fato destes patamares estarem nas extremidades.

Foi desenvolvido o método da seleção da perturbação apropriada de [Zhu et al. \(2007\)](#) com o objetivo de identificar se a perturbação utilizada é adequada e caso não seja propor uma perturbação apropriada. A motivação de aplicar a técnica de [Zhu et al. \(2007\)](#), surge através do estudo de [Galea-Rojas \(2014\)](#) e [Giménez & Patat \(2014\)](#). Estes dois trabalhos possuem algumas características em comum, tais como: o modelo utilizado foi de calibração comparativa e a estrutura da matriz de variância e covariância são independentes, no entanto, no primeiro trabalho os dados são balanceados, enquanto que no segundo trabalho os dados são desbalanceados. Estes autores citados anteriormente propõem o uso da técnica pelo fato de identificar se as perturbações são adequadas e para corrigir o uso da técnica de diagnóstico de influência local. No nosso caso, temos um modelo de calibração comparativa com réplicas balanceadas, porém possuímos uma estrutura de variância e covariância dependente. Na modelagem deste trabalho, esta característica de dependência é o principal responsável por produzir um efeito que causa um desequilíbrio no estudo de influência local. Para contornar este problema, utilizamos uma nova perturbação através da técnica de [Zhu et al. \(2007\)](#) que pondera os pesos de tal forma que retome o equilíbrio no estudo de influência.

Após as correções feitas no diagnóstico da modelagem normal no esquema de perturbação na variável resposta, percebe-se que houve mudança em todos os casos, pois em todos os casos houve a necessidade de corrigir a perturbação. Anteriormente a correção, percebe-se

que haviam observações influentes em vários casos, e após a correção, acessando a influência de forma correta, foi detectado observações influentes apenas quando perturbado a variável resposta no laboratório 4 e 6.

O desenvolvimento deste estudo foi aplicado em um conjunto de dados reais, onde foi visto a importância de investigar a influência local no modelo ultraestrutural com réplicas, para poder estudar a robustez e sensibilidade do modelo a pequenas perturbações.

6

Propostas futuras

Nesta pesquisa, foi desenvolvido o diagnóstico de influência local de [Cook \(1986\)](#) e [Zhu et al. \(2007\)](#), considerando o modelo de calibração ultraestrutural com réplicas. Como propostas futuras, devemos fazer os estudos referente ao método de seleção apropriada de [Zhu et al. \(2007\)](#) nos esquemas de perturbação citados no Capítulo 3 para o modelo t de Student e após isto, poderemos comparar com o diagnóstico de influência do modelo normal na presença da perturbação adequada em cada caso.

Referências Bibliográficas

- Aoki, R., Bolfarine, H., & Sinjer, J. M. (2001). Null intercept measurement error regression models. Sociedad de Estadística e Investigación Operativa, **10**, 441–457.
- Aoki, R., Bolfarine, H., & Singer, J. M. (2002). Asymptotic efficiency of methods of moments estimators under null intercept measurement error regression models. Brasilian journal of probability and statistics, **16**, 157–166.
- Aoki, R., Leão, D. P. J., Achar, J. A., & Bolfarine, H. (2006). Mixture of normal distribution in multivariate null intercept measurement error models. Journal of Biopharmaceutical Statistics, **16**, 785–802.
- Arellano, R., Bolfarine, H., & Vilca-Labra, F. (1996). Ultrastructural elliptical models. The Canadian Journal of Statistics / La Revue Canadienne de Statistique, **24**(2), 207–216.
- Barnett, V. D. (1969). Simultaneous pairwise linear structural relationships. Biometrics, **25**, 125–142.
- Borssoi, J. A. (2014). Modelos mistos lineares elípticos com erros de medição. Tese de doutorado, IME-USP.
- Cadigan, N. G. & Farrell, P. J. (2002). Generalized local influence with applications to fish stock cohort analysis. Journal of the statistical society, **51**(4), 469–483.
- Carroll, R. J. & Spiegelman, C. H. (1992). Diagnostics for nonlinearity and heteroscedasticity in errors-in-variables regression. Technometrics, **34**(2), 186–196.
- Carroll, R. J., Ruppert, D., Stefanski, L. A., & Crainiceanu, C. M. (2006). Measurement error in nonlinear models. Chapman and Hall/CRC, 2 edition.

-
- Castro, M. & Galea-Rojas, M. J. (2010). Robust inference in an heteroscedastic measurement error model. Journal of the Korean Statistical Society, **39**, 439–447.
- Castro, M. & Galea-Rojas, M. J. (2015). Inference in a structural heteroskedastic calibration model. Stat Papers, **56**, 479–494.
- Castro, M., Galea-Rojas, M. J., & Bolfarine, H. (2007). Local influence assessment in heteroscedastic measurement error models. Computational Statistics & Data Analysis, **52**, 1132–1142.
- Castro, M., Galea-Rojas, M. J., & Bolfarine, H. (2008). Hypothesis testing in an errors-in-variables model with heteroscedastic measurement errors. Statistics in medicine.
- Chan, L. K. & Mak, T. K. (1979). Maximum likelihood estimation of a linear structural relationship with replication. Journal of the Royal Statistical Society, **41**(2), 263–268.
- Cheng, C. L. & Van Ness, J. W. (1991). On the unreplicated ultrastructural model. Biometrika, **78**(2), 442–445.
- Cheng, C. L. & Van Ness, J. W. (1999). Statistical regression with measurement error. Oxford University Press.
- Cook, D. (1977). Detection of influential observation in linear regression. Technometrics, **19**(1), 15–18.
- Cook, D. (1986). Assessment of local influence. Journal of the Royal Statistical Society, **48**, 133–169.
- Cox, N. R. (1976). The linear structural relation for several groups of data. Biometrika, **63**(2), 231–237.
- Dempster, A., Laird, N., & Rubin, D. (1977). Maximum likelihood from incomplete data via the em algorithm (with discussion). Journal of the Royal Statistical Society B, **39**, 1–38.

-
- Dolby, G. R. (1976). The ultrastructural relation: A synthesis of functional and structural relations. Biometrika, **63**, 39–50.
- Escobar, E. & Meeker, W. (1992). Assessing influence in regression analysis with censored data. Biometrics, **48**, 507–528.
- Feng-Chang, X. & Bo-Cheng, W. (2009). Diagnostics for generalized poisson regression models with errors in variables. Journal of Statistical Computation and Simulation, **79**(7), 909–922.
- Fuller, A. (1987). Measurement Error Models. 1 edition.
- Fung, W. K. & Kwan, C. W. (1997). A note on local influence based on normal curvature. Journal of the Royal Statistical Society, **59**(4), 839–843.
- Galea-Rojas, M. (1995). Calibração Comparativa Estrutural e Funcional. Tese de doutorado, IME-USP.
- Galea-Rojas, M. (2014). Comparación de métodos de medición en presencia de um gold estándar. Estadística, **65**(185), 49–65.
- Galea-Rojas, M. & Castro, M. (2012). Influence assessment in an heteroscedastic errors in variables model. Communications in Statistics-Theory and Methods, **41**, 1350–1363.
- Galea-Rojas, M., Castilho, M. V., Bolfarine, H., & Castro, M. (2003). Calibração comparativa estrutural e funcional. The Analyst, **128**(8), 1073–1081.
- Giménez, P. & Bolfarine, H. (1997). Corrected score functions in classical error in variables and incidental parameters models. Austral. J. Statist., **39**(3), 325–344.
- Giménez, P. & Bolfarine, H. (2000). Comparing consistent estimators in comparative calibration models. Journal of Statistical Planning and Inference, **86**, 143–155.
- Giménez, P. & Galea-Rojas, M. (2013). Influence measures on corrected score estimators in functional heteroscedastic measurement error models. Journal of Multivariate Analysis, **114**, 1–15.

-
- Giménez, P. & Patat, M. L. (2014). Local influence for functional comparative calibration models with replicated data. Stat Papers, **55**, 431–454.
- Giménez, P. & Patat, M. L. (2005). Estimation in comparative calibration models with replicate measurement. Statistics & Probability Letters, **71**, 155–164.
- Gleser, L. J. (1985). A note on g. r. dolby's unrepliated ultrastructural model. Biometrika, **72**(1), 117–124.
- Grubbs, F. E. (1973). Errors of measurement, precision, accuracy and the statistical comparison of measuring instruments. Technometrics, **15**, 53–66.
- Guide43 (1997). Proficiency testing by interlaboratory comparisons- Part I: Development and operation of proficiency testeing schemes.
- Jaech, J. L. (1985). Statistical analysis of measurement errors. Monographs, New York.
- Kaaks, R., Riboli, E., Estive, J., Kappel, V., & Staveren, V. (1994). Estimating the accuracy of questionnaire assessments: Validation in terms of structural equation models. Statistics in Medicine, **13**, 127–142.
- Kelly, G. (1984). The influence function in the errors in variable problem. Then Annals of Statistics, **12**, 87–100.
- Kelly, G. (1985). Use of structural equation model in assessing the reliability of a new measurement technique. Applied Statistics, **34**, 258–263.
- Kendall, M. & Stuart, A. (1961). The advanced theory of statistics, volume 1. London: Griffin.
- Kendall, M. & Stuart, A. (1973). The advanced theory of statistics, volume 2. London: Griffin.
- Kimura, D. K. (1992). Functional comparative calibration using an em algorithm. Biometrics, **48**(4), 1263–1271.

- Kulathinal, S. B., Kuulasmaa, K., & Gasbarra, D. (2002). Estimation of an errors in variables regression model when the variances of the measurement errors vary between the observations. Statistics in Medicine, **21**(8), 1089–1101.
- Leão, D. P. J., Aoki, R., & Silva, G. F. (2009). Statistical analysis of proficiency testing results under elliptical distributions. Computational Statistics & Data Analysis, **53**(4), 1427–1439.
- Lesaffre, E. & Verbeke, G. (1998). Local influence in linear mixed models. Biometrics, **54**, 570–582.
- Leurgans, S. (1980). Evaluating laboratory measurement techniques in Biostatistics.
- Loynes, R. M. (2001). A new measure in local influence. Journal of Statistical Planning and Inference, **92**, 47–53.
- McLachlan, G. & Krishnan, T. (1997). The EM algorithm and Extensions.
- Moran, P. A. P. (1971). Estimating structural and functional relationships. Journal of Multivariate Analysis, **1**, 232–255.
- Nobre, J. S. & Singer, J. M. (2007). Residual analysis for linear mixed models. Biometrical Journal, **49**, 863–875.
- Oliveira, W. L. (2011). Inferência em um modelo com erros de medição heteroscedásticos com observações replicadas. Dissertação de mestrado, ICMC-USP.
- Osório, F., Paula, G. A., & Galea-Rojas (2007). Assessment of local influence in elliptical linear models with longitudinal structure. Computacional Statistics and Data analysis, **51**, 4354–4368.
- Patefield, W. (1978). The unreplicated ultrastructural relation: large sample properties. Biometrika, **65**(3), 535.
- Patriota, A. G. (2010). Modelos heterocedásticos com erro nas variáveis. Tese de doutorado, IME-USP.

-
- Patriota, A. G., Bolfarine, H., & Castro, M. (2009). A heteroscedastic structural errors-in-variables model with equation error. Statistical Methodology, **6**, 408–423.
- Patriota, A. G., Bolfarine, H., & Arellano-Valle, R. B. (2011). A multivariate ultrastructural errors-in-variables model with equation error. Journal of Multivariate Analysis, **102**, 386–392.
- Poon, W. Y. & Poon, Y. S. (1999). Conformal normal curvature and assessment of local influence. Journal of the Royal Statistical Society, **61**(1), 51–61.
- Rasekh, A. R. (2001). Ridge estimation in functional measurement error models. Inst. Stat. Univ. Paris, **2**(3), 47–59.
- Rasekh, A. R. & Fieller, N. R. J. (2003). Influence functions in functional measurement error models with replicated data. Statistics, **37**(2), 169–178.
- Ripley, B. D. & Thompson, M. (1987). Regression techniques for the detection of analytical bias. The Analyst, **112**(4), 377–383.
- Riu, J. & Rius, F. X. (1996). Assessing the accuracy of analytical methods using linear regression with errors in both axes. Analytical Chemistry, **68**(11), 1851–1857.
- Russo, C. M. (2006). Análise de um modelo de regressão com erros nas variáveis multivariado com intercepto nulo. Dissertação de mestrado, ICMC-USP.
- Russo, C. M., Aoki, R., & Leão, D. P. J. (2009). Hypotheses testing on a multivariate null intercept errors-in-variables model. Communications in Statistics-Simulation and Computation, **38**, 1447–1469.
- Suárez, R. M. M. & Sierra, G. M. A. (2001). Diagnostic regression using local influence: a review. Comm. Statist. Theory Methods, **30**(5), 799–813.
- Talarico, A. (2014). Modelo de calibração ultraestrutural. Dissertação de mestrado, ICMC-USP.

-
- Thomas, W. & Cook, D. R. (1989). Trustassessing influence on regression coefficients in generalized linear models. Biometrika, **76**(4), 741–749.
- Verbeke, G. & Molenberghs, G. (2000). Linear Mixed Models for Longitudinal Data. New York.
- Vilca-Labra, F., Arellano-Valle, R. B., & Bolfarine, H. (1998). Elliptical functional models. Journal of Multivariate Analysis, **65**, 36–57.
- Vilca-Labra, F., Aoki, R., & Bolfarine, H. (2005). Local influence in null intercept measurement error regression under a student-t model. Journal of Applied Statistics, **32**(7), 723–739.
- Vilca-Labra, F., Aoki, R., & Zeller, C. B. (2011). Hypotheses testing for structural calibration model. Stat Papers, **52**, 553–565.
- Wellman, J. M. & Gunst, R. F. (1991). Trustinfluence diagnostics for linear measurement error models. Biometrika, **78**(2), 373–380.
- Werkema, M. C. C. (1996). Avaliação da qualidade de medidas, volume 13 of Ferramentas da qualidade. Belo Horizonte.
- Zhao, Y. & Lee, A. H. (1995). Assessment of influence in nonlinear measurement of error models. Journal of Applied Statistics, **22**(2), 215–225.
- Zhao, Y., Lee, A. H., & Hui, Y. Y. (1994). Influence diagnostics for generalized linear measurement error models. Inst. Statist. Math, **52**(2).
- Zhong, X. P., Wei, B. C., & Fung, W. K. (2000). Influence analysis for linear measurement error models. Biometrics, **50**(4), 1117–1128.
- Zhu, H. & Zhang, H. (2004). A diagnostic procedure based on local influence. Biometrika, **91**(3), 579–589.
- Zhu, H., Ibrahim, J. G., Lee, S., & Zhang, H. (2007). Perturbation selection and influence measures in local influence analysis. The Annals of Statistics, **35**(6), 2565–2588.

Apêndice: A

Neste apêndice iremos apresentar os resultados da função escore e a matriz de informação observada no que se refere a todas as expressões que irão alimentar cada elemento da matriz. Vale salientar que os resultados apresentados são respectivamente sob a modelagem normal e t-Student.

A1: Sob distribuição Normal

Função escore

A função escore é: $U(\boldsymbol{\theta}) = \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = (U(\boldsymbol{\theta}_1)^t, U(\boldsymbol{\theta}_2)^t, U(\boldsymbol{\theta}_3)^t)^t$, em que, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^t, \boldsymbol{\theta}_2^t, \boldsymbol{\theta}_3^t)^t$ com, $\boldsymbol{\theta}_1 = (\mu_{x_1}, \dots, \mu_{x_m})^t$, $\boldsymbol{\theta}_2 = (\alpha_2, \dots, \alpha_p)^t$, $\boldsymbol{\theta}_3 = (\beta_2, \dots, \beta_p)^t$ e

$$U(\boldsymbol{\theta}_1) = -\frac{1}{2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_1},$$

$$U(\boldsymbol{\theta}_2) = -\frac{1}{2} \sum_{j=1}^m \frac{\partial Q_j}{\partial \boldsymbol{\theta}_2},$$

$$U(\boldsymbol{\theta}_3) = -\frac{1}{2} \sum_{j=1}^m \frac{\partial \log |\boldsymbol{\Sigma}_j|}{\partial \boldsymbol{\theta}_3} - \frac{1}{2} \sum_{j=1}^m \frac{\partial Q_j}{\partial \boldsymbol{\theta}_3}.$$

Com isto, as derivadas primeiras com relação aos parâmetros são:

$$\text{a) } \frac{\partial Q_j}{\partial \mu_{x_j}} = \left[\frac{2}{a_j} \right] \left[\frac{\mu_{x_j}(a_j - 1)}{\sigma_{x_j}^2} - \frac{\sum_{k=1}^{n_1} y_{1jk}}{\sigma_{1j}^2} - \sum_{i=2}^p \frac{\beta_i}{\sigma_{ij}^2} \left(\sum_{k=1}^{n_i} y_{ijk} - \alpha_i n_i \right) \right], j = 1, \dots, m,$$

$$\text{b) } \frac{\partial Q_j}{\partial \mu_{x_h}} = 0, j, h = 1, \dots, m, \text{ para } h \neq j,$$

$$\text{c) } \frac{\partial Q_j}{\partial \alpha_i} = \frac{2}{\sigma_{ij}^2} \left\{ \frac{n_i \beta_i \sigma_{x_j}^2 M_j}{a_j} + n_i \alpha_i - \sum_{k=1}^{n_i} y_{ijk} \right\}, \text{ para } j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

$$d) \frac{\partial Q_j}{\partial \beta_i} = M_j \frac{\sigma_{xj}^2}{a_j} \left[\frac{\partial Q_j}{\partial \alpha_i} \right] \text{ onde, } j = 1, \dots, m \text{ e } i = 2, \dots, p,$$

onde,

$$M_j = \frac{\mu_{xj}}{\sigma_{xj}^2} + \sum_{k=1}^{n_1} \frac{Y_{1jk}}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{\beta_i}{\sigma_{ij}^2} \left(\sum_{k=1}^{n_i} Y_{ijk} - \alpha_i n_i \right) \text{ e} \quad (6.1)$$

$$a_j = 1 + \sigma_{xj}^2 \left(\frac{n_1}{\sigma_{1j}^2} + \sum_{i=2}^p \frac{n_i \beta_i^2}{\sigma_{ij}^2} \right). \quad (6.2)$$

Matriz de informação observada

A matriz de informação observada é $I(\boldsymbol{\theta}) = -\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^t}$, em que, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^t, \boldsymbol{\theta}_2^t, \boldsymbol{\theta}_3^t)^t$ com, $\boldsymbol{\theta}_1^t = (\mu_{x1}, \dots, \mu_{xm})^t$, $\boldsymbol{\theta}_2^t = (\alpha_2, \dots, \alpha_p)^t$, $\boldsymbol{\theta}_3^t = (\beta_2, \dots, \beta_p)^t$ e portanto,

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_1} & I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_2} & I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_3} \\ I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_1} & I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_2} & I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_3} \\ I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_1} & I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_2} & I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_3} \end{bmatrix},$$

em que,

$$\begin{aligned} I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_1} &= -\frac{1}{2} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1^t}, \quad I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_2} = I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_1}^t = -\frac{1}{2} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_2^t}, \quad I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_3} = I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_1}^t = -\frac{1}{2} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_3^t}, \\ I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_2} &= -\frac{1}{2} \sum_{j=1}^m \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2^t}, \quad I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_3} = I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_2}^t = -\frac{1}{2} \sum_{j=1}^m \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_3^t}, \quad I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_3} = -\frac{1}{2} \sum_{j=1}^m \frac{\partial^2 \log |\boldsymbol{\Sigma}_j|}{\partial \boldsymbol{\theta}_3 \partial \boldsymbol{\theta}_3^t} - \\ &\frac{1}{2} \sum_{j=1}^m \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_3 \partial \boldsymbol{\theta}_3^t}. \end{aligned}$$

Com isto,

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{\mu_{x1}\mu_{x1}} & \cdots & I_{\mu_{x1}\mu_{xm}} & I_{\mu_{x1}\alpha_2} & \cdots & I_{\mu_{x1}\alpha_p} & I_{\mu_{x1}\beta_2} & \cdots & I_{\mu_{x1}\beta_p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ I_{\mu_{xm}\mu_{x1}} & \cdots & I_{\mu_{xm}\mu_{xm}} & I_{\mu_{xm}\alpha_2} & \cdots & I_{\mu_{xm}\alpha_p} & I_{\mu_{xm}\beta_2} & \cdots & I_{\mu_{xm}\beta_p} \\ I_{\alpha_2\mu_{x1}} & \cdots & I_{\alpha_2\mu_{xm}} & I_{\alpha_2\alpha_2} & \cdots & I_{\alpha_2\alpha_p} & I_{\alpha_2\beta_2} & \cdots & I_{\alpha_2\beta_p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ I_{\alpha_p\mu_{x1}} & \cdots & I_{\alpha_p\mu_{xm}} & I_{\alpha_p\alpha_2} & \cdots & I_{\alpha_p\alpha_p} & I_{\alpha_p\beta_2} & \cdots & I_{\alpha_p\beta_p} \\ I_{\beta_2\mu_{x1}} & \cdots & I_{\beta_2\mu_{xm}} & I_{\beta_2\alpha_2} & \cdots & I_{\beta_2\alpha_p} & I_{\beta_2\beta_2} & \cdots & I_{\beta_2\beta_p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ I_{\beta_p\mu_{x1}} & \cdots & I_{\beta_p\mu_{xm}} & I_{\beta_p\alpha_2} & \cdots & I_{\beta_p\alpha_p} & I_{\beta_p\beta_2} & \cdots & I_{\beta_p\beta_p} \end{bmatrix},$$

em que,

$$I_{\theta_i\theta_h} = -\frac{\partial^2 L(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_h}.$$

A seguir, temos as derivadas de segunda ordem em relação aos parâmetros

- a) $I_{\mu_{xj}\mu_{xj}} = \frac{a_j - 1}{a_j\sigma_{xj}^2}$, para $j = 1, \dots, m$,
- b) $I_{\mu_{xj}\mu_{xh}} = 0$, para $j, h = 1, \dots, m, j \neq h$,
- c) $I_{\mu_{xj}\alpha_i} = \frac{n_i\beta_i}{a_j\sigma_{ij}^2}$, para $j = 1, \dots, m$ e $i = 2, \dots, p$,
- d) $I_{\mu_{xj}\beta_i} = \frac{1}{\sigma_{ij}^2 a_j} \left[\frac{2n_i\beta_i\sigma_{xj}^2 M_j}{a_j} - D_{ij} \right]$, para $j = 1, \dots, m$ e $i = 2, \dots, p$,
- e) $I_{\alpha_i\alpha_i} = \sum_{j=1}^m \frac{n_i}{\sigma_{ij}^2} \left(1 - \frac{n_i\beta_i^2\sigma_{xj}^2}{\sigma_{ij}^2 a_j} \right)$, para $i = 2, \dots, p$,
- f) $I_{\alpha_i\alpha_l} = -\sum_{j=1}^m \frac{n_i n_l \beta_i \beta_l \sigma_{xj}^2}{a_j \sigma_{ij}^2 \sigma_{lj}^2}$, para $i, l = 2, \dots, p$ e $i \neq l$,
- g) $I_{\alpha_i\beta_i} = \sum_{j=1}^m \frac{n_i \sigma_{xj}^2}{\sigma_{ij}^2 a_j} \left[M_j - \frac{\beta_i}{\sigma_{ij}^2} \left[\frac{2n_i\beta_i\sigma_{xj}^2 M_j}{a_j} - D_{ij} \right] \right]$, para $i = 2, \dots, p$,
- h) $I_{\alpha_i\beta_l} = \sum_{j=1}^m \frac{n_i \beta_l \sigma_{xj}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j} \left[D_{lj} - \frac{2n_l \beta_l \sigma_{xj}^2 M_j}{a_j} \right]$, para $i, l = 2, \dots, p$ e $i \neq l$,
- i) $I_{\beta_i\beta_i} = \sum_{j=1}^m \frac{\sigma_{xj}^2}{a_j \sigma_{ij}^2} \left[n_i - \frac{D_{ij}^2}{\sigma_{ij}^2} + \frac{n_i \sigma_{xj}^2}{a_j} \left[M_j^2 \left(1 - \frac{4n_i \sigma_{xj}^2 \beta_i^2}{a_j \sigma_{ij}^2} \right) - \frac{2n_i \beta_i^2}{\sigma_{ij}^2} + \frac{4M_j \beta_i D_{ij}}{\sigma_{ij}^2} \right] \right]$, para $i = 2, \dots, p$,

$$j) I_{\beta_i\beta_l} = - \sum_{j=1}^m \frac{\sigma_{xj}^2}{\sigma_{ij}^2 \sigma_{lj}^2 a_j} \left[D_{ij} D_{lj} + \frac{2\sigma_{xj}^2}{a_j} \left[n_i n_l \beta_i \beta_l \left(\frac{2\sigma_{xj}^2 M_j^2}{a_j} + 1 \right) n_i \beta_i M_j D_{lj} - n_l \beta_l M_j D_{ij} \right] \right],$$

para $i, l = 2, \dots, p$ e $i \neq l$.

A2: Sob distribuição t-Student

Função escore e Matriz de informação observada

A função escore para o modelo mencionado anteriormente é:

$$U(\boldsymbol{\theta}) = \frac{\partial L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = (U(\boldsymbol{\theta}_1)^t, U(\boldsymbol{\theta}_2)^t, U(\boldsymbol{\theta}_3)^t)^t \text{ onde, } \boldsymbol{\theta} = (\boldsymbol{\theta}_1^t, \boldsymbol{\theta}_2^t, \boldsymbol{\theta}_3^t)^t \text{ com, } \boldsymbol{\theta}_1 = (\mu_{x1}, \dots, \mu_{xm})^t, \\ \boldsymbol{\theta}_2 = (\alpha_2, \dots, \alpha_p)^t, \boldsymbol{\theta}_3 = (\beta_2, \dots, \beta_p)^t \text{ e}$$

$$U(\boldsymbol{\theta}_1) = -\frac{1}{2} (\gamma + n) \frac{1}{\gamma - 2 + Q_j} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_1},$$

$$U(\boldsymbol{\theta}_2) = -\frac{1}{2} (\gamma + n) \sum_{j=1}^m \frac{1}{\gamma - 2 + Q_j} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_2} \text{ e}$$

$$U(\boldsymbol{\theta}_3) = -\frac{1}{2} \sum_{j=1}^m \frac{\partial \log |\boldsymbol{\Sigma}_j|}{\partial \boldsymbol{\theta}_3} - \frac{1}{2} (\gamma + n) \sum_{j=1}^m \frac{1}{\gamma - 2 + Q_j} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_3}.$$

A matriz de informação observada é dada por $I(\boldsymbol{\theta}) = -\frac{\partial^2 L(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^t}$, em que, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^t, \boldsymbol{\theta}_2^t, \boldsymbol{\theta}_3^t)^t$ com, $\boldsymbol{\theta}_1^t = (\mu_{x1}, \dots, \mu_{xm})^t$, $\boldsymbol{\theta}_2^t = (\alpha_2, \dots, \alpha_p)^t$, $\boldsymbol{\theta}_3^t = (\beta_2, \dots, \beta_p)^t$ e portanto,

$$I(\boldsymbol{\theta}) = \begin{bmatrix} I_{\boldsymbol{\theta}_1\boldsymbol{\theta}_1} & I_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2} & I_{\boldsymbol{\theta}_1\boldsymbol{\theta}_3} \\ I_{\boldsymbol{\theta}_2\boldsymbol{\theta}_1} & I_{\boldsymbol{\theta}_2\boldsymbol{\theta}_2} & I_{\boldsymbol{\theta}_2\boldsymbol{\theta}_3} \\ I_{\boldsymbol{\theta}_3\boldsymbol{\theta}_1} & I_{\boldsymbol{\theta}_3\boldsymbol{\theta}_2} & I_{\boldsymbol{\theta}_3\boldsymbol{\theta}_3} \end{bmatrix},$$

onde,

$$I_{\boldsymbol{\theta}_1\boldsymbol{\theta}_1} = \frac{(\gamma + n)}{2} \left(-\frac{1}{(\gamma - 2 + Q_j)^2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_1} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_1^t} + \frac{1}{(\gamma - 2 + Q_j)} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_1^t} \right),$$

$$I_{\boldsymbol{\theta}_1\boldsymbol{\theta}_2} = I_{\boldsymbol{\theta}_2\boldsymbol{\theta}_1}^t = \frac{(\gamma + n)}{2} \left(-\frac{1}{(\gamma - 2 + Q_j)^2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_1} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_2^t} + \frac{1}{(\gamma - 2 + Q_j)} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_2^t} \right),$$

$$I_{\boldsymbol{\theta}_1 \boldsymbol{\theta}_3} = I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_1}^t = \frac{(\gamma + n)}{2} \left(-\frac{1}{(\gamma - 2 + Q_j)^2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_1} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_3^t} + \frac{1}{(\gamma - 2 + Q_j)} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_1 \partial \boldsymbol{\theta}_3^t} \right),$$

$$I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_2} = \frac{(\gamma + n)}{2} \left(\sum_{j=1}^m -\frac{1}{(\gamma - 2 + Q_j)^2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_2^t} + \sum_{j=1}^m \frac{1}{(\gamma - 2 + Q_j)} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_2^t} \right),$$

$$I_{\boldsymbol{\theta}_2 \boldsymbol{\theta}_3} = I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_2} = \frac{(\gamma + n)}{2} \left(\sum_{j=1}^m -\frac{1}{(\gamma - 2 + Q_j)^2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_3^t} + \sum_{j=1}^m \frac{1}{(\gamma - 2 + Q_j)} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_2 \partial \boldsymbol{\theta}_3^t} \right) e$$

$$I_{\boldsymbol{\theta}_3 \boldsymbol{\theta}_3} = \frac{1}{2} \sum_{j=1}^m \frac{\partial^2 \log |\boldsymbol{\Sigma}_j|}{\partial \boldsymbol{\theta}_3 \partial \boldsymbol{\theta}_3^t} - \frac{(\gamma + n)}{2} \sum_{j=1}^m \frac{1}{(\gamma - 2 + Q_j)^2} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_3} \frac{\partial Q_j}{\partial \boldsymbol{\theta}_3^t} +$$

$$+ \frac{(\gamma + n)}{2} \sum_{j=1}^m \frac{1}{(\gamma - 2 + Q_j)} \frac{\partial^2 Q_j}{\partial \boldsymbol{\theta}_3 \partial \boldsymbol{\theta}_3^t}.$$

Para mais detalhes, como por exemplo o passo a passo em relação as derivadas, devesse acessar a dissertação da [Talarico \(2014\)](#).

Apêndice: B

Tabela 6.1 *Conjunto de dados usado na aplicação (dispostos na ordem Código (C), Faixa (F) e Potência Corrigida (PC))*

C	F	PC	C	F	PC	C	F	PC
1	1500	19,76	1	4000	63,39	2	2000	27,27
1	1500	19,96	1	4500	71,48	2	2000	27,62
1	1500	19,78	1	4500	71,56	2	2500	35,47
1	1500	19,85	1	4500	71,51	2	2500	35,54
1	1500	19,64	1	4500	71,53	2	2500	35,35
1	2000	27,23	1	4500	70,91	2	2500	35,61
1	2000	27,33	1	5000	76,99	2	2500	36,08
1	2000	27,24	1	5000	77,01	2	3000	45,46
1	2000	27,18	1	5000	76,78	2	3000	45,40
1	2000	27,30	1	5000	76,99	2	3000	45,16
1	2500	35,57	1	5000	76,85	2	3000	45,73
1	2500	35,51	1	5500	80,06	2	3000	46,05
1	2500	35,45	1	5500	80,35	2	3500	54,25
1	2500	35,40	1	5500	80,33	2	3500	54,51
1	2500	35,55	1	5500	80,21	2	3500	53,89
1	3000	45,61	1	5500	80,14	2	3500	54,22
1	3000	45,60	1	6000	80,13	2	3500	55,31
1	3000	45,53	1	6000	80,35	2	4000	63,73
1	3000	45,54	1	6000	80,23	2	4000	63,54
1	3000	45,34	1	6000	79,13	2	4000	62,30
1	3500	54,52	1	6000	79,94	2	4000	62,49
1	3500	54,26	2	1500	19,77	2	4000	63,89
1	3500	54,42	2	1500	19,70	2	4500	71,01
1	3500	54,33	2	1500	19,51	2	4500	71,25
1	3500	54,32	2	1500	19,62	2	4500	68,75
1	4000	63,68	2	1500	20,02	2	4500	69,48
1	4000	63,17	2	2000	27,41	2	4500	71,67
1	4000	63,13	2	2000	27,43	2	5000	76,95
1	4000	63,08	2	2000	26,97	2	5000	76,66

C	F	PC	C	F	PC	C	F	PC
2	5000	74,85	3	4000	61,07	4	3500	54,30
2	5000	75,42	3	4500	68,78	4	3500	54,40
2	5000	77,25	3	4500	68,48	4	3500	54,40
2	5500	79,86	3	4500	68,83	4	3500	54,40
2	5500	80,37	3	4500	68,66	4	4000	63,40
2	5500	78,22	3	4500	68,48	4	4000	63,60
2	5500	78,71	3	5000	74,57	4	4000	63,50
2	5500	80,84	3	5000	73,99	4	4000	63,60
2	6000	79,66	3	5000	74,02	4	4000	63,80
2	6000	79,85	3	5000	74,12	4	4500	71,20
2	6000	77,55	3	5000	73,94	4	4500	71,40
2	6000	77,69	3	5500	77,83	4	4500	71,30
2	6000	81,46	3	5500	77,53	4	4500	71,20
3	1500	19,06	3	5500	77,73	4	4500	71,50
3	1500	19,14	3	5500	77,77	4	5000	76,30
3	1500	19,17	3	5500	77,59	4	5000	76,50
3	1500	19,23	3	6000	78,45	4	5000	76,40
3	1500	19,18	3	6000	78,41	4	5000	76,30
3	2000	26,43	3	6000	78,55	4	5000	76,60
3	2000	26,37	3	6000	78,67	4	5500	78,60
3	2000	26,41	3	6000	78,67	4	5500	78,80
3	2000	26,45	4	1500	19,80	4	5500	78,70
3	2000	26,37	4	1500	19,70	4	5500	78,50
3	2500	34,37	4	1500	19,70	4	5500	78,90
3	2500	34,25	4	1500	19,80	4	6000	78,00
3	2500	34,24	4	1500	20,00	4	6000	78,00
3	2500	34,21	4	2000	27,40	4	6000	78,00
3	2500	34,21	4	2000	27,40	4	6000	78,00
3	3000	43,84	4	2000	27,40	4	6000	78,20
3	3000	43,48	4	2000	27,40	5	1500	19,70
3	3000	43,57	4	2000	27,50	5	1500	19,67
3	3000	43,43	4	2500	35,60	5	1500	19,56
3	3000	43,52	4	2500	35,50	5	1500	19,38
3	3500	52,54	4	2500	35,50	5	1500	19,56
3	3500	52,22	4	2500	35,60	5	2000	27,06
3	3500	52,19	4	2500	35,70	5	2000	27,33
3	3500	51,95	4	3000	45,50	5	2000	27,01
3	3500	51,92	4	3000	45,50	5	2000	26,97
3	4000	61,16	4	3000	45,50	5	2000	27,13
3	4000	61,10	4	3000	45,60	5	2500	35,56
3	4000	61,34	4	3000	45,70	5	2500	35,43
3	4000	61,14	4	3500	54,30	5	2500	35,15

C	F	PC	C	F	PC	C	F	PC
5	2500	35,26	6	2000	26,80	6	6000	79,10
5	3000	45,47	6	2000	27,00	6	6000	79,10
5	3000	45,40	6	2000	26,80			
5	3000	45,18	6	2000	26,90			
5	3000	44,91	6	2500	35,00			
5	3000	45,00	6	2500	34,90			
5	3500	54,42	6	2500	35,00			
5	3500	54,35	6	2500	34,90			
5	3500	54,05	6	2500	34,90			
5	3500	53,90	6	3000	45,20			
5	3500	53,92	6	3000	45,00			
5	4000	63,20	6	3000	45,10			
5	4000	62,53	6	3000	45,00			
5	4000	62,58	6	3000	45,10			
5	4000	62,70	6	3500	53,80			
5	4000	62,67	6	3500	53,60			
5	4500	71,04	6	3500	53,60			
5	4500	70,60	6	3500	53,60			
5	4500	70,46	6	3500	53,70			
5	4500	70,39	6	4000	62,30			
5	4500	70,36	6	4000	62,10			
5	5000	76,43	6	4000	62,20			
5	5000	76,14	6	4000	62,10			
5	5000	76,00	6	4000	62,20			
5	5000	75,95	6	4500	69,80			
5	5000	75,90	6	4500	69,70			
5	5500	79,51	6	4500	69,70			
5	5500	78,89	6	4500	69,70			
5	5500	79,28	6	4500	69,70			
5	5500	79,26	6	5000	75,00			
5	5500	79,22	6	5000	74,80			
5	6000	78,85	6	5000	74,80			
5	6000	78,59	6	5000	74,80			
5	6000	78,79	6	5000	74,80			
5	6000	78,88	6	5500	78,60			
5	6000	78,76	6	5500	78,30			
6	1500	19,60	6	5500	78,40			
6	1500	19,50	6	5500	78,40			
6	1500	19,60	6	5500	78,50			
6	1500	19,50	6	6000	79,10			
6	1500	19,60	6	6000	79,00			
6	2000	27,00	6	6000	78,90			
