

Equações de estimação
para dados com medidas repetidas
em mais de um fator

Patrícia Viana da Silva

DISSERTAÇÃO APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientador: Profa. Dra. Denise A. Botter

Ao longo deste trabalho a autora recebeu auxílio financeiro da CAPES.

São Paulo, 05 de junho de 2009

Equações de estimação
para dados com medidas repetidas
em mais de um fator

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Patrícia Viana da Silva
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Profa. Dra. Denise Aparecida Botter (orientadora) - IME-USP.
- Prof. Dr. Rinaldo Artes - Insper.
- Profa. Dra. Maria Kelly Venezuela - Insper.

Agradecimentos

Pela conclusão desse trabalho tenho muito a agradecer e a muitas pessoas.

Agradeço às professoras doutoras **Denise Aparecida Botter** e **Mônica Carneiro Sandoval** pela disponibilidade, paciência e dedicação conjuntas na orientação e no desenvolvimento desse projeto.

Aos meus pais **Antonio** e **Terezinha** a quem devo quase tudo o que sou. Aos meus irmãos, **Henrique** e **José Mário**, os quais aprendi a admirar.

Aos meus grandes amigos **Jony** e **Márcio** pelo apoio incondicional, sendo a família que escolhi ou talvez a que me escolheu.

Ao amigo **Caio**, atualmente distante, mas que não deixa de se preocupar comigo. Às amigas **Gleiciani**, **Nubia**, **Roberta** e **Fernanda** que me re-apresentaram à felicidade da amizade feminina.

A todos do meu convívio próximo no **Laboratório de Epidemiologia e Estatística** pela compreensão e oportunidade de crescimento.

Às minhas eternas professoras **Antonia**, pelo carinho maternal, e **Sissi**, por me mostrar o caminho da estatística.

Aos meus professores da Universidade Federal do Ceará: **Julio**, pelo incentivo nas primeiras pesquisas; **Silvia**, pela paciência com as minhas inúmeras dúvidas, e principalmente ao professor **Mauricio**, pelos grandes ensinamentos, pela sua incansável dedicação aos alunos e por ter acreditado em mim.

À **CAPES** pela bolsa de mestrado concedida que possibilitou não só o desenvolvimento desse projeto como a viabilização do meu mestrado.

Resumo

Nesse trabalho desenvolvemos equações de estimação generalizadas para análise de dados com medidas repetidas em dois fatores. As equações foram concebidas segundo a proposta de Liang e Zeger (1986). O método não pressupõe o conhecimento da estrutura real de dependência entre as observações sob os fatores, no entanto, utiliza uma *matriz de correlação de trabalho* a fim de modelar essa estrutura de dependência. Apresentamos estimadores consistentes para os parâmetros que modelam a média e para os parâmetros de correlação considerando-se sub-estruturas na matriz de correlação de trabalho. Apresentamos também rotinas para o cálculo das estimativas desses parâmetros. Analisamos três conjuntos de dados e avaliamos a qualidade dos modelos ajustados por meio das técnicas de diagnóstico desenvolvidas por Venezuela et al. (2007).

Palavras-chave: Equações de estimação generalizadas; estrutura de correlação; família exponencial de distribuições; medidas repetidas.

Abstract

This work develops generalized estimating equations for analysis of data with repeated measures in two factors. The equations had been developed according to Liang and Zeger (1986)'s proposal. The method does not assume the knowledge of the real structure of dependence between subjects under factors, it use a *working correlation matrix* to model this structure of dependence. Consistent estimators are thus presented for the parameters that shape the average and for the parameters of correlation considering substructures in the working correlation matrix. Also, routines are presented for the calculation of the estimates of these parameters. Three data sets are analysed and the quality of the adjusted models is evaluated applying the diagnostic techniques developed by Venezuela et al. (2007).

Keywords: Correlation structure; exponencial distribution family; generalized estimating equation; repeated measures.

Sumário

Lista de Figuras	xi
Lista de Tabelas	xiii
1 Introdução	1
1.1 Organização do trabalho	4
2 Medidas repetidas em um único fator	7
2.1 Equações de estimação	13
2.1.1 Equação de estimação sob independência	14
2.1.2 Equações de estimação com matriz de correlação conhecida	17
2.1.3 Equação de estimação generalizada	19
2.2 Métodos de estimação	22
2.2.1 Estruturas de correlação para medidas repetidas em um único fator . .	23
2.2.2 Testes de hipóteses	27
3 Medidas repetidas em mais de um fator	29
3.1 Experimento com medidas repetidas em dois fatores	30
3.2 Equações de estimação generalizadas	32

3.3	Estruturas de correlação considerando sub-estruturas	33
3.3.1	Estrutura de correlação não-estruturada	33
3.3.2	Sub-estruturas de correlação padrão uniforme	34
3.3.3	Sub-estruturas de correlação AR-1	36
3.3.4	Estrutura de correlação mista (sub-estrutura AR-1 e sub-estrutura padrão uniforme)	37
3.4	Estimadores de β e de sua matriz de covariâncias	39
4	Aplicações	41
4.1	Aplicação 1	41
4.2	Aplicação 2	46
4.3	Aplicação 3	52
5	Conclusões e estudos futuros	63
A	Conjuntos de dados e rotinas computacionais	65
A.1	Aplicação 1	65
A.1.1	Rotina para estimação dos parâmetros de correlação, de dispersão e da média com matriz de correlação de trabalho com sub-estruturas padrão uniforme em ambos os fatores:	69
A.2	Aplicação 2	74
A.2.1	Rotina para estimação dos parâmetros de correlação, de dispersão e de média com matriz de correlação de trabalho com sub-estruturas AR-1 em ambos os fatores:	76
A.3	Aplicação 3	80

<i>SUMÁRIO</i>	ix
A.3.1 Rotina para estimação dos parâmetros de correlação, de dispersão e de média com matriz de correlação de trabalho com sub-estruturas mista (AR-1 e padrão uniforme):	82
Referências Bibliográficas	87

Lista de Figuras

4.1	Histograma para a variável <i>Força</i>	42
4.2	Distância de Cook para o modelo de regressão gama ajustado com sub-estruturas de correlação padrão uniforme para a variável <i>Força</i>	44
4.3	Resíduos padronizados para o modelo de regressão gama ajustado com sub-estruturas de correlação padrão uniforme para a variável <i>Força</i>	45
4.4	Gráfico de probabilidade meio-normal com envelope simulado para o modelo de regressão gama ajustado com sub-estruturas de correlação padrão uniforme para a variável <i>Força</i>	46
4.5	Distância de Cook para o modelo de regressão gama ajustado com sub-estruturas de correlação AR-1 para a <i>variável resposta</i>	50
4.6	Resíduos Padronizados para o modelo de regressão gama ajustado com sub-estruturas de correlação AR-1 para a <i>variável resposta</i>	51
4.7	Gráfico de probabilidade meio-normal com envelope simulado para o modelo de regressão gama ajustado com sub-estruturas de correlação AR-1 para a <i>variável resposta</i>	52
4.8	Histograma para a variável <i>Escore de acurácia</i>	53
4.9	Distância de Cook para o modelo de regressão normal ajustado com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável <i>Escore</i>	57

4.10	Resíduos Padronizado para o modelo de regressão normal ajustado com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável Escore. . .	58
4.11	Gráfico de probabilidade meio-normal com envelope simulado para o modelo de regressão normal ajustado com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável escore.	59
4.12	Comparação das distâncias de Cook entre o ajuste com estrutura de correlação mista (AR-1 e uniforme) e ajuste com ambas as sub-estruturas de correlação uniforme.	60
4.13	Comparação dos resíduos padronizados entre ajuste com estrutura de correlação mista (AR-1 e uniforme) e ajuste com ambas as sub-estruturas de correlação uniforme.	60
4.14	Comparação dos gráficos de probabilidade meio-normal com envelope simulado entre ajuste com estrutura de correlação mista (AR-1 e uniforme) e ajuste com ambas as sub-estruturas de correlação uniforme.	61

Lista de Tabelas

3.1	Esquema de um experimento com medidas repetidas em dois fatores.	30
4.1	Estimativas e erros padrões dos parâmetros do modelo de regressão gama com sub-estruturas de correlação padrão uniforme para a variável <i>Força</i>	43
4.2	Estimativas e erros padrões dos parâmetros do modelo de regressão gama com estrutura de correlação AR-1 para a <i>variável resposta</i>	48
4.3	ANOVA (Teste de Wald) para os dados da Aplicação 2.	49
4.4	Estimativas e erros padrões dos parâmetros do modelo de regressão normal com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável <i>score</i>	55
4.5	ANOVA (Teste de Wald) para os dados da Aplicação 3.	55
A.1	Conjunto de dados da Aplicação 1.	65
A.2	Conjunto de dados da Aplicação 2.	74
A.3	Conjunto de dados da Aplicação 3.	80

Capítulo 1

Introdução

Experimentos com medidas repetidas envolvem a realização de observações sob diferentes condições de avaliação para cada unidade experimental. Esse tipo de delineamento, apesar de apresentar, em geral, dificuldade de acompanhamento das unidades experimentais ao longo do estudo, é muito comum em diversas áreas de pesquisa, já que exige tamanhos de amostra menores, e fornece melhores condições para o estudo de covariáveis, pois características externas ao estudo se mantêm controladas pelo fato de as medidas serem tomadas no mesmo indivíduo. Além disso, permite acompanhar a mudança de comportamento da resposta média ao longo das condições de avaliação e muitas vezes melhora a precisão de contrastes associados a diferenças entre valores médios da variável resposta.

A literatura estatística é muito vasta quanto à análise de dados com medidas repetidas quando a distribuição de probabilidades da variável resposta pode ser considerada normal. Como exemplo, podemos citar Neter et al. (1996), Singer e Andrade (1986), Laird e Ware (1982) e Andreoni (2003). Isso se deve à existência da distribuição normal multivariada e ao seu fácil manuseio teórico inferencial. Quando a variável não segue a distribuição normal, várias dificuldades surgem. Por exemplo, a inviabilidade do uso do método de máxima verossimilhança para estimação de parâmetros de interesse, causada pela falta de conhecimento sobre a distribuição conjunta dos dados.

Quando há interesse em analisar o comportamento médio da variável resposta, pode-se contornar o problema utilizando-se a abordagem de Equações de Estimação Generalizadas -

EEG para dados com medidas repetidas (Generalized Estimating Equations - GEE) desenvolvida por Liang e Zeger (1986). Nessa abordagem, considera-se que cada observação segue uma distribuição marginal pertencente à família exponencial de distribuições (Nelder e Wedderburn, 1972) para a qual definimos um modelo de regressão da mesma forma que no *modelo linear generalizado - MLG*.

Dessa forma, o modelo probabilístico multivariado não precisa ser explicitado e a estimação dos parâmetros de interesse é realizada sem ignorar a estrutura de dependência entre as observações da mesma unidade experimental que é incorporada sob a forma de uma *matriz de correlação de trabalho*, $\mathbf{R}(\boldsymbol{\rho})$, totalmente especificada pelo vetor $\boldsymbol{\rho}$ considerado parâmetro de perturbação. As raízes da equação de estimação generalizada são estimadores consistentes para os parâmetros que modelam a média desde que o estimador do vetor $\boldsymbol{\rho}$ seja consistente.

A *matriz de correlação de trabalho*, $\mathbf{R}(\boldsymbol{\rho})$, objetiva definir estruturas que são, supostamente, mas não garantidamente seguidas pela matriz de correlação dos dados. As propriedades assintóticas dos estimadores dos parâmetros da regressão não são afetadas se a estrutura de correlação de trabalho não coincide com a verdadeira, desde que $\boldsymbol{\rho}$ seja consistentemente estimado. Contudo, podem sofrer uma grande perda de eficiência (Fitzmaurice, 1995). É necessário, pois, propor estruturas adequadas e obter estimadores consistentes para os parâmetros de correlação.

Em alguns problemas a estrutura de correlação não se adequa diretamente às usuais (padrão uniforme, M-dependente, AR-1, etc). É o caso em que a variável resposta é avaliada sob um experimento fatorial em que as medidas se repetem em ambos os fatores conjuntamente aplicados na mesma unidade experimental. Nesse caso, temos uma estrutura peculiar na matriz de covariâncias para a dependência intra indivíduos, *entre* e *intra* fatores. Assim, existe a necessidade de mudanças no modelo para que este considere esse tipo de dependência das medidas dentro de cada unidade experimental. Uma das mudanças é a assimilação das estruturas usuais como sub-estruturas na *matriz de correlação de trabalho*.

A seguir, apresentamos um exemplo obtido de Winer et al. (1991).

Exemplo 1 (Winer et al., 1991) *Um médico deseja estudar os efeitos de dois medicamentos*

(A e B) na circulação sanguínea de seres humanos. Os níveis dos dois fatores são:

- Fator A: placebo e uso do medicamento A;
- Fator B: placebo e uso do medicamento B.

Doze homens de meia-idade participaram do estudo, sendo que cada um deles recebeu os quatro tratamentos formados pelos cruzamentos dos níveis dos dois fatores. Os quatro tratamentos são:

A_1B_1 - placebo (sem uso de medicamento A e sem uso do medicamento B);

A_1B_2 - uso do medicamento B isolado;

A_2B_1 - uso do medicamento A isolado; e

A_2B_2 - uso dos medicamentos A e B em conjunto.

Dessa forma, temos dois fatores com dois níveis cada e a matriz de correlação entre as observações de um indivíduo considerando sub-estruturas pode ser descrita como:

$$\mathbf{R}(\rho) = \begin{bmatrix} 1 & \rho_{11,12} & \vdots & \rho_{11,21} & \rho_{11,22} \\ \rho_{12,11} & 1 & \vdots & \rho_{12,21} & \rho_{12,22} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_{21,11} & \rho_{21,12} & \vdots & 1 & \rho_{21,22} \\ \rho_{22,11} & \rho_{22,12} & \vdots & \rho_{22,21} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_2 & \mathbf{R}_1 \end{bmatrix},$$

em que $\rho_{jk,j'k'}$, $j, j' = 1, 2$ e $k, k' = 1, 2$, é a correlação entre a observação do indivíduo no nível j do fator A e no nível k do fator B e a observação do indivíduo no nível j' do fator A e no nível k' do fator B. E ainda, a sub-matriz \mathbf{R}_1 contém as correlações entre as observações fixado cada nível do fator A e \mathbf{R}_2 contém as correlações entre as observações fixado cada nível do fator B e diferentes níveis de ambos os fatores, A e B. Quando o número de níveis de cada fator é maior que dois essas sub-matrizes podem assumir um tipo de sub-estrutura conhecida, por exemplo, padrão uniforme, autoregressiva de ordem 1 ou não-estruturada.

Para o caso de um único fator, são dadas na literatura algumas sugestões para a estimação da estrutura de correlação como em Artes e Jorgensen (2000), Huajun e Jianxin (2006) e Wang

e Carey (2004), por exemplo. Isso não ocorre, no entanto, quando as medidas se repetem em mais de um fator.

Nesse trabalho são considerados dados balanceados, ou seja, originados de experimentos em que as unidades experimentais são observadas em todos os cruzamentos entre níveis de fatores. Além disso, pressupondo as condições gerais da teoria de *equações de estimação generalizadas*, são definidas estruturas para a matriz de correlação $R(\rho)$ e propostos estimadores para ρ que garantem consistência ao estimar os parâmetros que modelam a média da variável resposta.

Como em qualquer modelo de regressão, faz-se necessária a verificação da qualidade do seu ajuste, diagnóstico, pelo qual são examinados possíveis afastamentos das suposições do modelo e são identificadas observações atípicas que podem influenciar desproporcionalmente a inferência. As técnicas de diagnósticos são bem definidas para modelos lineares generalizados (Paula, 2004) e foram estendidas para dados com medidas repetidas por Venezuela et al. (2007). Utilizaremos esta extensão introduzindo a estimação dos parâmetros de correlação no ajuste.

Aplicações a dados reais são apresentadas considerando-se desde a estimação dos parâmetros até a verificação da adequação do ajuste do modelo.

1.1 Organização do trabalho

O corpo do trabalho apresenta-se organizado em capítulos conforme resumo abaixo.

O Capítulo 2 traz a teoria de funções de estimação, definições e conceitos necessários para a obtenção de estimadores consistentes e assintoticamente normais a partir de equações de estimação generalizadas para medidas repetidas em um único fator.

No Capítulo 3 apresentamos as condições e especificações do modelo e, ainda, as propostas de estruturas de dependência para dados com medidas repetidas em mais de um fator.

As aplicações são descritas no Capítulo 4, em que também são avaliados os ajustes com as técnicas de diagnóstico adequadas.

As conclusões encontram-se no Capítulo 5, assim como sugestões para estudos futuros.

1.1. ORGANIZAÇÃO DO TRABALHO

5

No Apêndice A constam os conjuntos de dados utilizados e as rotinas computacionais desenvolvidas para analisá-los.

Capítulo 2

Funções e equações de estimação para medidas repetidas em um único fator

Neste capítulo são apresentados alguns conceitos e definições da teoria de Funções de Estimação (Godambe, 1991).

Uma função de estimação é uma função da amostra, \mathbf{y} , e do parâmetro de interesse θ . Um exemplo bem conhecido de função de estimação é a *função score*. A seguir temos a definição formal encontrada em Artes e Botter (2005).

Definição 1 *Seja $\mathcal{X} \in \mathcal{R}^t$ um espaço amostral no qual se define uma família*

$\mathcal{P} = \{P_\theta : \theta \in \Theta \subseteq \mathcal{R}^p\}$ de distribuições de probabilidades indexadas por um parâmetro θ desconhecido. Por definição, uma função $\psi : \mathcal{X} \times \Theta \rightarrow \mathcal{R}^p$, é uma função de estimação se para cada $\theta \in \Theta$, $\psi(\cdot, \theta)$ é um vetor aleatório.

Considerando uma amostra de n vetores aleatórios independentes: $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{it})^T$, $i = 1, 2, \dots, n$, com uma função de estimação ψ_i associada a cada um deles, o conceito pode ser estendido como *função de estimação para amostra* por

$$\Psi_n(\mathbf{y}; \theta) = \sum_{i=1}^n \psi_i(\mathbf{y}_i; \theta), \quad (2.1)$$

em que $\mathbf{y} = (\mathbf{y}_1^T, \mathbf{y}_2^T, \dots, \mathbf{y}_n^T)^T$ é o vetor de observações com dimensão $(N \times 1)$, com $N = nt$, e $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$ é o vetor de parâmetros com dimensão $(p \times 1)$.

O conceito de função de estimação é muito geral, abrangendo diversos tipos de funções. No entanto, para nosso estudo são interessantes apenas aquelas que possuem como raízes estimadores dos parâmetros envolvidos no modelo, ou seja,

$$\Psi_n(\hat{\boldsymbol{\theta}}_n) = 0. \quad (2.2)$$

A expressão (2.2) é chamada *equação de estimação* relacionada a $\boldsymbol{\theta}$ e tem como raiz $\hat{\boldsymbol{\theta}}_n$ estimador do parâmetro.

Adotando a notação $\Psi_n(\boldsymbol{\theta})$ em lugar de $\Psi_n(\mathbf{y}; \boldsymbol{\theta})$ são apresentados a seguir alguns importantes conceitos da teoria de *Funções de Estimação*.

Definição 2 *As funções de estimação $\Psi_n(\boldsymbol{\theta})$ e $\Phi_n(\boldsymbol{\theta})$ serão ditas funções de estimação equivalentes se*

$$\Psi_n(\boldsymbol{\theta}) = C(\boldsymbol{\theta})\Phi_n(\boldsymbol{\theta}),$$

em que $C(\boldsymbol{\theta})$ é uma matriz quadrada não estocástica e não singular. Isso implica que $\Psi_n(\boldsymbol{\theta})$ e $\Phi_n(\boldsymbol{\theta})$ terão as mesmas raízes.

Definição 3 *A função de estimação $\Psi_n(\boldsymbol{\theta})$ será considerada não viciada se*

$$E_{\boldsymbol{\theta}}(\Psi_n(\boldsymbol{\theta})) = \mathbf{0}, \forall \boldsymbol{\theta} \in \Theta.$$

De (2.1) se depreende que, se $\psi_i(\boldsymbol{\theta})$ é não viciada, para todo $i = 1, 2, \dots, n$, então a função $\Psi_n(\boldsymbol{\theta})$ com base numa amostra de tamanho n também será não viciada.

Definição 4 A função $\Psi_n(\boldsymbol{\theta})(\mathbf{y}, \boldsymbol{\theta}) = (\psi_1(\mathbf{y}, \boldsymbol{\theta}), \psi_2(\mathbf{y}, \boldsymbol{\theta}), \dots, \psi_p(\mathbf{y}, \boldsymbol{\theta}))^T : \mathcal{X} \times \Theta \rightarrow \mathcal{R}^p$ será considerada função de estimação regular caso satisfaça para todo $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)^T \in \Theta$

- i. $E_{\boldsymbol{\theta}} \{ \Psi_n(\mathbf{y}, \boldsymbol{\theta}) \} = \mathbf{0}$, ou seja, a função é não viciada;
- ii. A derivada parcial de $\partial \Psi(\mathbf{y}; \boldsymbol{\theta}) / \partial \theta_i$ existe quase certamente para $\mathbf{y} \in \mathcal{X}$;
- iii. É possível permutar o sinal de integração e diferenciação da seguinte forma:

$$\frac{\partial}{\partial \theta_i} \int_{\mathcal{X}} \Psi(\mathbf{y}; \boldsymbol{\theta}) p(\mathbf{y}, \boldsymbol{\theta}) d\mathbf{y} = \int_{\mathcal{R}} \frac{\partial}{\partial \theta_i} \{ \psi(\mathbf{y}; \boldsymbol{\theta}) p(\mathbf{y}, \boldsymbol{\theta}) \} d\mathbf{y};$$

- iv. $E_{\boldsymbol{\theta}} [\Psi_i(\mathbf{y}; \boldsymbol{\theta}) \Psi_j(\mathbf{y}; \boldsymbol{\theta})] \in \mathcal{R}$, para $i, j = 1, 2, \dots, p$ e

$$\mathbf{V}_{\Psi}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}} \{ \Psi(\mathbf{y}; \boldsymbol{\theta}) \Psi_n^T(\mathbf{y}; \boldsymbol{\theta}) \}$$

é positiva definida; e ainda

- v. $E_{\boldsymbol{\theta}} \left\{ \frac{\partial}{\partial \theta_r} \psi_i(\mathbf{y}; \boldsymbol{\theta}) \frac{\partial}{\partial \theta_s} \psi_j(\mathbf{y}; \boldsymbol{\theta}) \right\} \in \mathcal{R}$, e $\mathbf{S}_{\Psi}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}} \{ \nabla_{\boldsymbol{\theta}} \Psi(\mathbf{y}; \boldsymbol{\theta}) \}$ é não singular, em que $\nabla_{\boldsymbol{\theta}}$ representa o operador gradiente em relação a $\boldsymbol{\theta}$, ou seja, $\nabla_{\boldsymbol{\theta}} \Psi(\mathbf{y}; \boldsymbol{\theta}) = \partial \Psi(\mathbf{y}; \boldsymbol{\theta}) / \partial \boldsymbol{\theta}^T$, sendo

$$\frac{\partial \Psi(\mathbf{y}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} = \begin{bmatrix} \frac{\partial \psi_1(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_1} & \cdots & \frac{\partial \psi_1(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial \psi_p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_1} & \cdots & \frac{\partial \psi_p(\mathbf{y}; \boldsymbol{\theta})}{\partial \theta_p} \end{bmatrix} = \left(\frac{\partial \Psi(\mathbf{y}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T.$$

Seja $\Psi_n(\boldsymbol{\theta})$ uma função de estimação regular. A matriz de informação de Godambe de $\boldsymbol{\theta}$ associada a Ψ_n é definida como

$$\mathbf{J}_{\Psi}(\boldsymbol{\theta}) = \mathbf{S}_{\Psi}(\boldsymbol{\theta})^T \mathbf{V}_{\Psi}^{-1}(\boldsymbol{\theta}) \mathbf{S}_{\Psi}(\boldsymbol{\theta})$$

com dimensão $(p \times p)$.

A *matriz de informação de Godambe* desempenha para as funções de estimação regulares o mesmo papel da matriz de informação de Fisher para a função escore. Temos para a função escore (função de estimação regular) $S_{\Psi}(\theta) = -V_{\Psi}(\theta)$ e, então, a matriz de informação de Godambe igual à matriz de informação de Fisher.

$S_{\Psi}(\theta)$ é denominada *matriz de sensibilidade* de uma função de estimação e está relacionada à derivada parcial da função de estimação em relação aos parâmetros. Essa matriz carrega informação sobre a eficiência da função na estimação dos parâmetros e no caso unidimensional espera-se que assumam alto valor em módulo.

Por outro lado, $V_{\Psi}(\theta)$ é conhecida como *matriz de variabilidade* e espera-se que a função de estimação apresente baixa variabilidade para que no verdadeiro valor do parâmetro possa se aproximar do seu valor médio, qual seja, zero.

Godambe (1960) desenvolveu o critério de *otimalidade* de uma função de estimação. No caso de θ ser unidimensional, por exemplo, pode-se considerar uma *função de estimação ótima* se suas raízes possuem variância assintótica mínima. Uma extensão do conceito para o caso multiparamétrico exige ordenação de matrizes de covariância assintóticas (Chandrasekar e Kale, 1984).

Neste estudo serão consideradas as *funções de estimação aditivas* ou *funções de estimação lineares* (Crowder, 1987). A classe dessas funções geradas por $\mathbf{u}_i = \mathbf{u}_i(\mathbf{y}_i; \theta)$, $i = 1, 2, \dots, n$ e vetores mutuamente independentes com média zero é definida por

$$\mathcal{L}(\mathbf{u}) = \left\{ \Psi_n(\theta) \in \mathcal{R} : \Psi_n(\theta) = \sum_{i=1}^n Q_i(\theta_i) \mathbf{u}_i(\mathbf{y}_i; \theta) \right\}, \quad (2.3)$$

com $Q_i(\theta)$, matriz de posto completo não estocástica, e \mathcal{R} contendo todas as funções regulares relacionadas a θ .

Crowder (1987) mostra que sob à classe de funções de estimação aditivas, $\mathcal{L}(\mathbf{u})$, a função de estimação ótima existe e pode ser obtida por

$$\Psi_n^*(\theta) = \sum_{i=1}^n Q_i^*(\theta) \mathbf{u}_i(\mathbf{y}_i; \theta) \quad (2.4)$$

em que

$$Q_i^*(\theta) = E_{\theta} \left(\frac{\partial \mathbf{u}_i}{\partial \theta^T} \right)^T Cov_{\theta}^{-1}(\mathbf{u}_i). \quad (2.5)$$

Funções de estimação *equivalentes* a $\Psi_n^*(\theta)$ também serão ótimas.

No caso de um único parâmetro de interesse, para que a distribuição assintótica da raiz de uma equação de estimação seja considerada normal, é necessário apenas mostrar sua consistência. Para θ vetor, no entanto, é necessário atender a algumas condições estabelecidas no teorema¹ a seguir:

Teorema 1 *Considerando-se*

- a. \mathbf{y}_i , $i = 1, 2, \dots, n$, vetores aleatórios t -dimensionais independentes;
- b. $\psi_i(\theta) = (\psi_{i1}, \psi_{i2}, \dots, \psi_{ip})^T$, $i = 1, 2, \dots, n$, funções de estimação regulares;

c. $\Psi_n(\theta) = \sum_{i=1}^n \psi_i(\theta)$;

- d. para $\delta > 0$

$$E_{\theta} \left\{ \sup_{\mathbf{h}: \|\mathbf{h}\| \leq \delta} \left\| \frac{\partial}{\partial \theta} \psi_i(\theta + \mathbf{h}) - \frac{\partial}{\partial \theta} \psi_i(\theta) \right\| \right\} = \phi_{\delta_i},$$

$$\frac{1}{n} \sum_{i=1}^n \phi_{\delta_i} \longrightarrow \phi_{\delta} \text{ e } \phi_{\delta} \longrightarrow 0 \text{ quando } \delta \longrightarrow 0;$$

¹A prova completa de todos os itens do Teorema 1 pode ser conferida em Artes (1997).

e. para $\delta > 0$, existe ϵ , $0 < \epsilon < 1$, tal que

$$n^{-1-\epsilon} \sum_{i=1}^n E_{\theta} \left\{ \sup_{\mathbf{h}: \|\mathbf{h}\| \leq \delta} \left\| \frac{\partial}{\partial \theta} \psi_i(\theta + \mathbf{h}) - \frac{\partial}{\partial \theta} \psi_i(\theta) - \phi_{\delta_i} \right\| \right\}^{1+\epsilon} \rightarrow 0;$$

f. que para $\dot{\psi}_i(\theta) = \frac{\partial}{\partial \theta^T} \psi_i(\theta)$,

$$\frac{1}{n} \sum_{i=1}^n E \left\{ \dot{\psi}_i(\theta) \right\} = \frac{1}{n} \sum_{i=1}^n S_i(\theta) \rightarrow S(\theta)$$

com S não singular e $S_i = E \left\{ \dot{\psi}_i(\theta) \right\}$;

g. que existe ϵ , $0 < \epsilon < 1$, tal que

$$n^{-1-\epsilon} \sum_{i=1}^n E \left| \dot{\psi}_i(\theta) - S_i(\theta) \right|^{1+\epsilon} \rightarrow 0;$$

h. que existe ϵ , $0 < \epsilon < 1$, tal que $\max_{1 \leq i \leq n} \max_{1 \leq j \leq t} E |\psi_{ij}|^{2+\epsilon} < \infty$;

i. $n^{-1} \sum_{i=1}^n \text{Cov}(\psi_i) \rightarrow V(\theta)$, matriz positiva definida;

j. $\hat{\theta}_n$ como a solução de $\Psi(\omega) = \mathbf{0}$, $\omega \in \Theta$,

e, ainda, sob condições de existência de uma sequência de raízes $\Psi_n(\omega)$ que seja limitada em probabilidade ou restrita a um conjunto compacto q . c., quando n tende a infinito, temos que

$$\hat{\theta}_n \xrightarrow{P} \theta.$$

Sob essas hipóteses,

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{D} N_p(\mathbf{0}, J_{\psi}^{-1}(\theta)),$$

em que

$$\begin{aligned} J_{\psi}(\theta) &= \left\{ \lim_{n \rightarrow \infty} S_n^T(\theta) \right\} \left\{ \lim_{n \rightarrow \infty} V_n(\theta) \right\}^{-1} \left\{ \lim_{n \rightarrow \infty} S_n(\theta) \right\} \\ &= \left\{ \lim_{n \rightarrow \infty} S_n^T(\theta) \right\} \left\{ \lim_{n \rightarrow \infty} V_n^{-1}(\theta) \right\} \left\{ \lim_{n \rightarrow \infty} S_n(\theta) \right\} \\ &= \lim_{n \rightarrow \infty} \left\{ S_n^T(\theta) V_n^{-1}(\theta) S_n(\theta) \right\}, \end{aligned}$$

com

$$S_n(\theta) = \sum_{i=1}^n \frac{S_{\psi_i}(\theta)}{n}, \quad S_{\psi_i}(\theta) = E_{\theta} \left\{ \frac{\partial}{\partial \theta^T} \psi_i(\mathbf{y}_i, \theta) \right\}$$

e

$$V_n(\theta) = \sum_{i=1}^n \frac{V_{\psi_i}(\theta)}{n}, \quad V_{\psi_i}(\theta) = E_{\theta} \left\{ \Psi_n(\mathbf{y}_i, \theta)^T \Psi_n(\mathbf{y}_i, \theta) \right\}.$$

2.1 Equações de estimação

A teoria de *equações de estimação* foi desenvolvida por Liang e Zeger (1986) como uma extensão da teoria de quase-verossimilhança multivariada, admitindo que cada observação tenha distribuição de probabilidade marginal pertencente à *Família Exponencial* - FE. Nesta seção são apresentados conceitos dessa teoria.

Seja $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{it})^T$, o vetor de respostas ($t \times 1$) associado ao i -ésimo elemento de uma amostra aleatória, $i = 1, 2, \dots, n$, e $\mathbf{X}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{it})^T$, uma matriz de planejamento ($t \times p$), em que $\mathbf{x}_{ij} = (x_{ij1}, x_{ij2}, \dots, x_{ijp})^T$ é o vetor $p \times 1$ de covariáveis para o indivíduo i na condição j , $j = 1, 2, \dots, t$. Assumimos a *função densidade de probabilidade* ou *função de probabilidade* de y_{ij} dada por

$$f(y_{ij}; \theta_{ij}, \phi) = \exp[\phi \{y_{ij}\theta_{ij} - b(\theta_{ij})\} + c(y_{ij}, \phi)] \quad (2.6)$$

para a qual $E[y_{ij}] = \mu_{ij} = b'(\theta_{ij})$ e $\text{Var}[y_{ij}] = \phi^{-1}V_{ij}$, em que $b'(\theta_{ij}) = db(\theta_{ij})/d\theta_{ij}$ e $V_{ij} = d\mu_{ij}/d\theta_{ij} = b''(\theta_{ij})$ é denominada *função de variância*, a qual caracteriza a distribuição na FE e $\phi^{-1} > 0$ é o parâmetro de dispersão assumido conhecido.

O modelo de regressão para a média μ_{ij} é estabelecido por

$$g(\mu_{ij}) = \eta_{ij} = \sum_{k=1}^p x_{ijk}\beta_k = \mathbf{x}_{ij}^T \boldsymbol{\beta}, \quad (2.7)$$

sendo $g(\cdot)$ uma função qualquer monótona e diferenciável denominada *função de ligação*, $\mathbf{x}_{ij}^T \boldsymbol{\beta}$ é o preditor linear e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$, com $p < n$, é o vetor de parâmetros modeladores da média.

2.1.1 Equação de estimação sob independência

Apresentamos inicialmente a função de estimação ótima, que ignora a dependência entre as observações. Serão consideradas as suposições e a notação expostas no começo deste capítulo. De (2.4) e (2.5) advém que a função de estimação ótima é dada por

$$\boldsymbol{\Psi}_n^{I*}(\boldsymbol{\beta}) = \phi \sum_{i=1}^n \mathbf{D}_i^T \mathbf{A}_i^{-1} \mathbf{u}_i, \quad (2.8)$$

com $\mathbf{u}_i = (\mathbf{y}_i - \boldsymbol{\mu}_i)$,

$$E_{\boldsymbol{\beta}} \left(\frac{\partial \mathbf{u}_i}{\partial \boldsymbol{\beta}^T} \right)^T = E_{\boldsymbol{\beta}} \left(\frac{\partial}{\partial \boldsymbol{\beta}^T} (\mathbf{y}_i - \boldsymbol{\mu}_i) \right)^T = - \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \boldsymbol{\beta}^T} \right)^T = -\mathbf{X}_i^T \boldsymbol{\Delta}_i \mathbf{A}_i = -\mathbf{D}_i^T$$

e

$$Cov_{\beta}(\mathbf{u}_i) = Cov(\mathbf{y}_i) = diag \{Var(y_{ij})\} = \phi^{-1} diag \{b''(\theta_{ij})\} = \phi^{-1} \mathbf{A}_i,$$

em que

$$\mathbf{X}_i = \frac{\partial \boldsymbol{\eta}_i}{\partial \boldsymbol{\beta}} = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{it})^T, \text{ matriz } (t \times p),$$

$$\boldsymbol{\Delta}_i = diag \left\{ \frac{\partial \theta_{ij}}{\partial \eta_{ij}} \right\}, \text{ matriz } (t \times t)$$

e

$$\mathbf{A}_i = diag \left\{ \frac{\partial \mu_{ij}}{\partial \theta_{ij}} \right\} = diag \{b''(\theta_{ij})\} = diag \{V_{ij}\}, \text{ matriz } (t \times t),$$

com $\boldsymbol{\eta}_i = (\eta_{i1}, \eta_{i2}, \dots, \eta_{it})^T$.

É razoável salientar que \mathbf{A}_i é definida pela função de variância e $\boldsymbol{\Delta}_i$ depende da função de ligação.

Igualando a função de estimação a zero, obtemos a *equação de estimação sob independência* - EEI - relacionada a $\boldsymbol{\beta}$, dada por

$$\Psi_n^{I*}(\boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{D}_i^T \mathbf{A}_i^{-1} \mathbf{u}_i = 0. \quad (2.9)$$

Atendendo às condições de regularidade do Teorema 1, é possível demonstrar que o estimador $\hat{\boldsymbol{\beta}}_{I*}$, raiz da EEI, é consistente para $\boldsymbol{\beta}$ e ainda

$$\sqrt{n}(\hat{\beta}_{I^*} - \beta) \xrightarrow{D} N_p(\mathbf{0}, J_{I^*}^{-1})$$

em que

$$J_{I^*} = \lim_{n \rightarrow \infty} \frac{J_{nI^*}}{n},$$

em que J_{nI^*} , matriz de informação de Godambe de β associada a $\Psi_n^{I^*}$, é igual à matriz de informação de Fisher, dada por

$$\begin{aligned} J_{nI^*} &= \left\{ \sum_{i=1}^n S_i \right\} \left\{ \sum_{i=1}^n V_i \right\}^{-1} \left\{ \sum_{i=1}^n S_i \right\} \\ &= \phi \sum_{i=1}^n D_i^T A_i D_i, \end{aligned}$$

com as matrizes de sensibilidade e de variabilidade da forma

$$\begin{aligned} S_i &= E_\beta \left[\frac{\partial}{\partial \theta^T} D_i^T A_i^{-1} \mathbf{u}_i \right] \\ &= D_i^T A_i^{-1} E_\beta \left[\frac{\partial}{\partial \theta^T} \mathbf{u}_i \right] = -D_i^T A_i^{-1} D_i \end{aligned}$$

e

$$\begin{aligned} V_i &= E_\beta [D_i^T A_i^{-1} \mathbf{u}_i \mathbf{u}_i^T A_i^{-1} D_i] \\ &= D_i^T A_i^{-1} Cov(\mathbf{y}_i) A_i^{-1} D_i = \phi D_i^T A_i^{-1} A_i A_i^{-1} D_i = -\phi^{-1} S_i. \end{aligned}$$

Um estimador consistente para a matriz de covariâncias de $\hat{\beta}_{I^*}$ é dado por

$$\hat{J}_{nI^*}^{-1} = \left\{ \phi \sum_{i=1}^n \hat{D}_i^T \hat{A}_i \hat{D}_i \right\}^{-1},$$

sendo as quantidades \hat{D} e \hat{A} avaliadas em $\hat{\beta}_{I^*}$.

Embora o estimador $\hat{\beta}_{I^*}$ seja consistente, ignorar a dependência entre as observações de um mesmo indivíduo gera perda de eficiência e as estimativas de erros padrões não são confiáveis.

2.1.2 Equações de estimação com matriz de correlação conhecida

Para que a função de estimação incorpore informação sobre a dependência dos dados, algumas alterações são necessárias. Assim, de (2.4) e (2.5), depreende-se que a equação de estimação ótima de β , quando a verdadeira matriz de correlações entre as observações de um mesmo indivíduo é conhecida - EEC, tem a seguinte forma

$$\Psi_n^{G^*} = \sum_{i=1}^n D_i \Sigma_i^{-1} \mathbf{u}_i = \mathbf{0},$$

com $D_i = \mathbf{X}_i^T \Delta_i \mathbf{A}_i$. Notar que aqui

$$E_{\beta} \left(\frac{\partial \mathbf{u}_i}{\partial \beta^T} \right)^T = E_{\beta} \left(\frac{\partial}{\partial \theta^T} (\mathbf{y}_i - \boldsymbol{\mu}_i) \right)^T = - \left(\frac{\partial \boldsymbol{\mu}_i}{\partial \beta^T} \right)^T = - \mathbf{X}_i^T \Delta_i \mathbf{A}_i = -D_i^T$$

e

$$\begin{aligned} Cov_{\beta}(\mathbf{u}_i) &= Cov(\mathbf{y}_i) = diag\{Var(y_{ij})^{1/2}\} \mathbf{R}_i diag\{Var(y_{ij})^{1/2}\} \\ &= \phi^{-1} \mathbf{A}_i^{1/2} \mathbf{R}_i \mathbf{A}_i^{1/2} = \Sigma_i, \end{aligned}$$

na qual \mathbf{R}_i é a verdadeira matriz de correlações de \mathbf{y}_i , $i = 1, \dots, n$.

Sob as condições de regularidade apresentadas no Teorema 1, $\hat{\beta}_{G^*}$, a raiz da equação

apresentada anteriormente é um estimador consistente para β e

$$\sqrt{n} (\hat{\beta}_{G^*} - \beta) \xrightarrow{\mathcal{D}} N_p(\mathbf{0}, J_{G^*}^{-1}),$$

com $J_{G^*} = \lim_{n \rightarrow \infty} \frac{J_{nG^*}}{n}$,

em que

$$\begin{aligned} J_{nG^*} &= \left\{ \sum_{i=1}^n S_i \right\} \left\{ \sum_{i=1}^n V_i \right\}^{-1} \left\{ \sum_{i=1}^n S_i \right\} = \\ &= \sum_{i=1}^n D_i^T \Sigma_i D_i \end{aligned}$$

é a matriz de informação de Godambe de β associada a $\Psi_n^{G^*}$.

Neste caso,

$$\begin{aligned} S_i &= E_\beta \left[\frac{\partial}{\partial \theta^T} D_i^T \Sigma_i^{-1} \mathbf{u}_i \right] \\ &= D_i^T \Sigma_i^{-1} E_\beta \left[\frac{\partial}{\partial \theta^T} \mathbf{u}_i \right] = -D_i^T \Sigma_i^{-1} D_i \end{aligned}$$

e

$$\begin{aligned} V_i &= E_\beta [D_i^T \Sigma_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \Sigma_i^{-1} D_i] \\ &= D_i^T \Sigma_i^{-1} Cov(\mathbf{y}_i) \Sigma_i^{-1} D_i = D_i^T \Sigma_i^{-1} D_i = -S_i. \end{aligned}$$

Um estimador consistente para a matriz de covariâncias de $\hat{\beta}_{G^*}$ é dado por

$$\hat{\mathbf{J}}_{nG*}^{-1} = \left\{ \sum_{i=1}^n \hat{\mathbf{D}}_i^T \hat{\Sigma}_i^{-1} \hat{\mathbf{D}}_i \right\}^{-1}, \quad (2.10)$$

que na literatura é conhecido como estimador “*naïve*” ou “*model-based*”.

Em geral, a verdadeira matriz de correlações é desconhecida, sendo que a equação Ψ_n^{G*} é pouco utilizada.

2.1.3 Equação de estimação generalizada

A teoria desenvolvida por Liang e Zeger (1986) consiste em contornar o problema da falta de conhecimento sobre a matriz de correlações entre as observações de um mesmo indivíduo. Os autores utilizam uma matriz simétrica ($t \times t$), definida sob condições de ser uma matriz de correlações, denominada *matriz de correlação de trabalho*, $\mathbf{R}(\boldsymbol{\rho})$, caracterizada pelo vetor de parâmetros de perturbação $\boldsymbol{\rho}$ com dimensão ($s \times 1$). Uma grande vantagem em relação a outras teorias, é que $\mathbf{R}(\boldsymbol{\rho})$ não precisa ser a verdadeira matriz de correlações dos vetores \mathbf{y}_i 's.

Assim, a equação de estimação generalizada - EEG de $\boldsymbol{\beta}$ é dada por:

$$\Psi_n^G(\boldsymbol{\beta}) = \Psi_n^G(\boldsymbol{\beta}, \hat{\boldsymbol{\rho}}(\boldsymbol{\beta}, \phi)) = \sum_{i=1}^n \mathbf{D}_i^T \Omega_i^{-1} \mathbf{u}_i = \mathbf{0}, \quad (2.11)$$

em que

$$\Omega_i = \Omega_i(\hat{\boldsymbol{\rho}}, \phi) = \phi^{-1} \mathbf{A}_i^{1/2} \mathbf{R}(\hat{\boldsymbol{\rho}}) \mathbf{A}_i^{1/2} \quad (2.12)$$

e $\hat{\boldsymbol{\rho}}$ é um estimador consistente de $\boldsymbol{\rho}$.

Nesse caso, (2.11) não é uma equação de estimação ótima, podendo inclusive ser viciada.

Dessa forma, precisamos obter um estimador de ρ de tal modo que o estimador de β , solução de (2.11), seja consistente e assintoticamente normal e, ainda, um estimador para ϕ , no caso de ele ser desconhecido.

As condições impostas aos estimadores de ρ e ϕ para que o estimador de β continue com as propriedades citadas acima são dadas pelo Teorema 2² a seguir.

Teorema 2 *Seja $\hat{\beta}_G$ raiz de (2.11). Sob condições de regularidade e assumindo que*

- a. $\hat{\rho}(\beta, \phi)$ é um estimador \sqrt{n} -consistente de ρ dados β e ϕ ;
- b. $\hat{\phi}(\beta)$ é um estimador \sqrt{n} -consistente de ϕ dado β ; e
- c. $\left| \frac{\partial \hat{\rho}(\beta, \phi)}{\partial \phi} \right| \leq H(\mathbf{y}, \beta)$, sendo H uma função $Op(1)$,

temos que $\hat{\beta}_G$ é um estimador consistente de β e

$$\sqrt{n} (\hat{\beta}_G - \beta) \xrightarrow{D} N_p(\mathbf{0}, \mathbf{J}_G^{-1}),$$

em que

$$\mathbf{J}_G = \lim_{n \rightarrow \infty} \frac{\mathbf{J}_{nG}}{n},$$

e a matriz de informação de Godambe de β associada a Ψ_n^G é dada por

$$\mathbf{J}_{nG} = \left\{ \sum_{i=1}^n \mathbf{S}_i \right\} \left\{ \sum_{i=1}^n \mathbf{V}_i \right\}^{-1} \left\{ \sum_{i=1}^n \mathbf{S}_i \right\}$$

em que

$$\mathbf{S}_i = E_\beta \left[\frac{\partial}{\partial \beta^T} \mathbf{D}_i^T \Omega_i^{-1} \mathbf{u}_i \right] = \mathbf{D}_i^T \Omega_i^{-1} E_\beta \left[\frac{\partial}{\partial \beta^T} \mathbf{u}_i \right] = -\mathbf{D}_i^T \Omega_i^{-1} \mathbf{D}_i$$

²A prova do Teorema 2 pode ser consultada em Liang e Zeger (1986).

e

$$V_i = E_{\beta} [D_i^T \Omega_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \Omega_i^{-1} D_i] = D_i^T \Omega_i^{-1} Cov(\mathbf{y}_i) \Omega_i^{-1} D_i.$$

Os resultados do Teorema 2 são válidos mesmo quando $\mathbf{R}(\boldsymbol{\rho})$ não é a verdadeira matriz de correlação de \mathbf{y}_i . Um estimador consistente para a matriz de covariâncias de $\hat{\boldsymbol{\beta}}_G$ é dado por

$$\hat{J}_{nG}^{-1} = \left\{ \sum_{i=1}^n \hat{S}_i \right\}^{-1} \left\{ \sum_{i=1}^n \hat{D}_i^T \hat{\Omega}_i^{-1} \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^T \hat{\Omega}_i^{-1} \hat{D}_i \right\} \left\{ \sum_{i=1}^n \hat{S}_i \right\}^{-1}, \quad (2.13)$$

conhecido na literatura como estimador “robusto”, “empírico” ou “sanduíche” e sendo as quantidades \hat{S}_i , \hat{D}_i e $\hat{\Omega}_i$ obtidas pelas estimativas consistentes de $\boldsymbol{\rho}$, $\boldsymbol{\beta}$ e ϕ .

A partir dos conceitos e da teoria expostos até agora, é possível fazer as seguintes observações:

- Se $\mathbf{R}(\boldsymbol{\rho})$ se aproxima da verdadeira matriz de correlação dos \mathbf{y}_i 's, então as estimativas com base no estimador “naïve” e com base no estimador “robusto” também são próximas. Isso pode auxiliar na verificação da correta especificação de $\mathbf{R}(\boldsymbol{\rho})$. Além disso, temos um aumento na eficiência dos estimadores de $\boldsymbol{\beta}$ [Zeger et al. (1992) e Fitzmaurice et al. (1993)].
- Para atender à propriedade de consistência, o estimador “naïve” necessita de correta especificação do modelo e da matriz de correlação de trabalho, enquanto o estimador “robusto” depende apenas da correta especificação do modelo (Prentice, 1988).
- Em pequenas amostras o estimador “robusto” pode ser altamente viesado, pois só atende a essa propriedade assintoticamente. Nesse caso, o estimador “naïve” pode apresentar melhores resultados contanto que $\hat{\boldsymbol{\rho}}$ seja consistente (Prentice, 1988).

2.2 Métodos de estimação

Não é costumeiro encontrar uma forma analítica para o estimador de β . Em razão disso, é proposto um processo iterativo para estimar β com base no método *scoring de Fisher* combinado com o *método de momentos* para estimar ϕ e ρ , esses últimos sendo tratados como parâmetros de perturbação.

Ao expandirmos a EEG dada em (2.11) em torno de um valor inicial $\hat{\beta}_G^{[0]}$, temos o algoritmo de estimação abaixo

$$\begin{aligned}\hat{\beta}_G^{[m+1]} &= \hat{\beta}_G^{[m]} - \left\{ E_{\beta} \left[\frac{\partial}{\partial \theta^T} \Psi_n^G \left(\hat{\beta}_G^{[m]} \right) \right] \right\}^{-1} \Psi_n^G \left(\hat{\beta}_G^{[m]} \right) \\ &= \hat{\beta}_G^{[m]} + \left\{ \left[\sum_{i=1}^n \hat{D}_i^T \hat{\Omega}^{-1} \hat{D}_i \right]^{-1} \left[\sum_{i=1}^n \hat{D}_i^T \hat{\Omega}^{-1} \hat{u}_i \right] \right\}^{[m]},\end{aligned}\quad (2.14)$$

em que $m = 0, 1, 2, \dots$ é o numero da iteração, $m + 1$ é a iteração atual e m a iteração anterior pela qual se obtiveram as estimativas de β , ρ e ϕ para atualizar as quantidades \hat{D} , $\hat{\Omega}$ e \hat{u} .

Quando substituimos $\hat{D}_i = \hat{A}_i \hat{\Delta}_i \hat{X}_i$ na expressão (2.14), chegamos a

$$\hat{\beta}_G^{[m+1]} = \left\{ \left[\sum_{i=1}^n \mathbf{X}_i^T \hat{\mathbf{W}}_i \mathbf{X}_i \right]^{-1} \left[\sum_{i=1}^n \mathbf{X}_i^T \hat{\mathbf{W}}_i \mathbf{z}_i \right] \right\}^{[m]}, \quad (2.15)$$

com $\hat{\mathbf{W}}_i = \hat{\Delta}_i \hat{A}_i \hat{\Omega}_i^{-1} \hat{A}_i \hat{\Delta}_i$ e $\mathbf{z}_i = \hat{\eta}_i + (\hat{A}_i \hat{\Delta}_i)^{-1}$.

Em seguida, desenvolvemos um processo de *mínimos quadrados ponderados* em que \mathbf{z} é a variável dependente modificada e \mathbf{W} é a matriz de pesos ambos mudando a cada passo do processo iterativo.

O *resíduo de Pearson* para a observação y_{ij} (Paula, 2004) é definido como

$$\hat{r}_{ij}^{[m]} = \frac{y_{ij} - \hat{\mu}_{ij}^{[m]}}{\sqrt{\hat{a}_{ij}^{[m]}}}, \quad (2.16)$$

em que $\hat{a}_{ij} = \hat{V}_{ij}$ é o j -ésimo elemento da diagonal principal de \hat{A}_i . O resíduo de Pearson possui as seguintes propriedades: $E[r_{ij}] = 0$, $Var[r_{ij}^2] = \phi^{-1}$ e $E[r_{ij}r_{ij'}] = \phi^{-1}\rho_{jk}$, $j \neq j'$.

Podemos, portanto, obter um estimador \sqrt{n} -consistente de ϕ em função de \hat{r}_{ij} , ou seja, dado β

$$\hat{\phi}^{[m]} = \left\{ \sum_{i=1}^n \sum_{j=1}^t \frac{(\hat{r}_{ij}^{[m]})^2}{nt - p} \right\}^{-1}. \quad (2.17)$$

2.2.1 Estruturas de correlação para medidas repetidas em um único fator

As EEG's permitem diferentes especificações para a estrutura de correlação entre as observações da mesma unidade amostral e os estimadores de ρ , dado ϕ , podem ser definidos em função de \hat{r}_{ij} 's. Utilizaremos, pois, $\mathbf{r}_i = (r_{i1}, r_{i2}, \dots, r_{it})$ vetor de resíduos de Pearson associados a \mathbf{y}_i a fim de especificar estimadores para algumas estruturas de correlação.

Tendo-se por premissa uma *matriz de correlação de trabalho não-estruturada*, não se impõe nenhum padrão específico para ρ , ou seja, $\mathbf{R}(\rho)$ é dada por

$$\mathbf{R}(\rho) = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1t} \\ \vdots & 1 & \rho_{23} & \cdots & \rho_{2t} \\ & \vdots & 1 & \cdots & \rho_{3t} \\ & & \vdots & \vdots & \vdots \\ & & & \cdots & \rho_{t,t-1} \\ & & & \cdots & 1 \end{bmatrix}. \quad (2.18)$$

Para a matriz não-estruturada apresentada em (2.18), um estimador \sqrt{n} -consistente (Liang e Zeger, 1986) pode ser obtido por

$$\mathbf{R}(\hat{\rho}) = \frac{\hat{\phi}}{n} \sum_{i=1}^n \hat{\mathbf{A}}_i^{-1/2} \hat{\mathbf{u}}_i \hat{\mathbf{u}}_i^T \hat{\mathbf{A}}_i^{-1/2}. \quad (2.19)$$

A *matriz de correlação de trabalho não estruturada* é especialmente útil quando não há informação sobre o comportamento das correlações. Contudo, ela não é muito prática, afinal, necessitamos de um n maior que garanta a convergência do método iterativo, pois o número de parâmetros de correlação a ser estimado é o maior possível.

Podemos obter melhores resultados quando uma *estrutura de correlação uniforme* é presumida, na qual, consideramos a mesma correlação entre todas as observações do mesmo indivíduo, ou seja, $\rho_{jk} = \rho$ para $i \neq j$, e a matriz de correlação de trabalho fica definida por

$$\mathbf{R}(\rho) = \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \vdots & 1 & \rho & \cdots & \rho \\ & \vdots & \vdots & \vdots & \vdots \\ & & & \cdots & \rho \\ & & & \cdots & 1 \end{bmatrix}. \quad (2.20)$$

Um estimador \sqrt{n} -consistente para ρ (Laird e Ware, 1982), nesse caso, é dado por

$$\hat{\rho} = \hat{\phi} \frac{\sum_{i=1}^n \sum_{j>k} \hat{r}_{ij} \hat{r}_{ik}}{2t(t-1)/n-p}. \quad (2.21)$$

Outra alternativa é presumirmos uma *estrutura de correlação autoregressiva de ordem 1*, AR-1, considerando a dependência entre as observações consecutivas, com $\rho_{jk} = \rho^{|j-k|}$. Nesse caso, a matriz de correlação de trabalho é dada por

$$R(\rho) = \begin{bmatrix} 1 & \rho^1 & \rho^2 & \cdots & \rho^{t-1} \\ \vdots & 1 & \rho^1 & \cdots & \rho^{t-2} \\ & \vdots & 1 & \cdots & \rho^{t-3} \\ & & \vdots & \vdots & \vdots \\ & & & \cdots & \rho^1 \\ & & & \cdots & 1 \end{bmatrix}, \quad (2.22)$$

em que o estimador \sqrt{n} -consistente para ρ (Artes e Botter, 2005) é dado por

$$\hat{\rho} = \hat{\phi} \frac{\sum_{i=1}^n \sum_{j=1}^{t-1} \hat{r}_{ij} \hat{r}_{i,j+1}}{n(t-1)}. \quad (2.23)$$

Podemos, ainda, estimar as correlações para uma estrutura AR-1 ao utilizarmos o coeficiente angular de uma regressão linear simples em que consideramos $\log(\hat{r}_{jk} \hat{r}_{jl})$ como variável dependente e $\log|k-l|$ como a variável independente (Liang e Zeger, 1986).

Há além das anteriores a *estrutura de correlação 1-dependente* que define dependência apenas entre observações consecutivas, em que $\rho_{j,j+1} = \rho_j$ com $j = 1, 2, \dots, t-1$. Assim,

$$R(\rho) = \begin{bmatrix} 1 & \rho_1 & 0 & \cdots & 0 \\ \vdots & 1 & \rho_2 & \cdots & 0 \\ & \vdots & 1 & \cdots & 0 \\ & & \vdots & \vdots & \vdots \\ & & & \cdots & \rho_{t-1} \\ & & & \cdots & 1 \end{bmatrix}, \quad (2.24)$$

com estimador \sqrt{n} -consistente dado por

$$\hat{\rho}_j = \hat{\phi} \frac{\sum_{i=1}^n \hat{r}_{ij} \hat{r}_{i,j+1}}{n-p}. \quad (2.25)$$

Para essa estrutura temos um caso especial quando $\rho_j = \rho$, isto é,

$$\mathbf{R}(\rho) = \begin{bmatrix} 1 & \rho & 0 & \cdots & 0 \\ \vdots & 1 & \rho & \cdots & 0 \\ & \vdots & \vdots & \vdots & \vdots \\ & & & \cdots & \rho \\ & & & \cdots & 1 \end{bmatrix}, \quad (2.26)$$

com um estimador \sqrt{n} -consistente para ρ definido a partir dos $\hat{\rho}_j$'s anteriores, ou seja,

$$\hat{\rho} = \frac{1}{t-1} \sum_{j=1}^{t-1} \hat{\rho}_j. \quad (2.27)$$

Em seguida, especificamos alguns passos a ser cumpridos para facilitar a estimação dos parâmetros.

Etapas do processo de estimação

1. Inicialmente, podemos desconsiderar a dependência entre as observações da mesma unidade amostral a fim de estimar $\beta^{[0]}$, valor inicial, por um MLG em que não são necessárias estimativas para ϕ ou ρ , ou seja, Ω sendo uma matriz identidade na equação (2.15);
2. Especificamos uma estrutura para a matriz de correlação de trabalho;
3. Caso a estrutura escolhida seja a de independência as estimativas de β são as obtidas no Passo 1;

4. Caso contrário, utilizamos o valor inicial $\beta^{[0]}$ para calcular as quantidades necessárias à equação (2.15) na próxima etapa:
 - a. Resíduos de Pearson $\hat{r}_{ij}^{[0]}$;
 - b. Parâmetro de escala $\hat{\phi}^{[0]}$; e
 - c. Parâmetro de correlação $\hat{\rho}^{[0]}$;
5. Nessa etapa, calculamos $\beta^{[m+1]}$ a partir das quantidades da etapa anterior e atualizamos $\hat{r}_{ij}^{[m]}$, $\hat{\phi}^{[m]}$ e $\hat{\rho}^{[m]}$, sendo $[m + 1]$ a iteração atual do processo;
6. Repetimos o Passo 5 até a convergência do vetor β .

2.2.2 Testes de hipóteses

Queremos testar a hipótese

$$\begin{cases} H_0 : \lambda_1 = \lambda_0 \\ H_1 : \lambda_1 \neq \lambda_0 \end{cases}, \quad (2.28)$$

em que λ_1 é um subvetor ($q \times 1$) contendo quaisquer q componentes de β , $1 \leq q \leq p$, ou seja, $\beta = (\lambda_1, \lambda_2)$ é uma partição do vetor de parâmetros da regressão com λ_1 contendo os q parâmetros de interesse na hipótese.

A estatística de Wald generalizada é dada por

$$Q_W = (\hat{\lambda}_1 - \lambda_0)^T \hat{J}_{n\lambda_G} (\hat{\lambda}_1 - \lambda_0), \quad (2.29)$$

com $\hat{J}_{n\lambda_G}^{-1}$, submatriz ($q \times q$) do estimador robusto \hat{J}_{nG}^{-1} , apresentado em (2.13), referente aos parâmetros de interesse da hipótese.

A estatística score generalizada é dada por

$$Q_S = \Psi_n^G(\lambda_0)^T \hat{V}_{n\lambda_0}^{-1} \Psi_n^G(\lambda_0), \quad (2.30)$$

em que $\Psi_n^G(\lambda_0)$ é o subvetor $(q \times 1)$ de Ψ_n^G e $\hat{V}_{n\lambda_0}$ é a submatriz $(q \times q)$ da matriz de variabilidade que consta no estimador robusto dado em (2.13) avaliados em λ_0 .

Consideramos a distribuição assintótica de $\hat{\beta}_G$ e, ainda, sob a hipótese nula dada em (2.28), quando $n \rightarrow \infty$, temos que

$$Q_W \sim \chi_q^2$$

e

$$Q_S \sim \chi_q^2,$$

ou seja, as duas estatísticas seguem distribuição assintótica qui-quadrado com q graus de liberdade.

Como o estimador “robusto” \hat{J}_{nG}^{-1} pode ser muito viciado em amostras pequenas, Rottitzky e Jewell (1990) fornecem uma alternativa ao escreverem as estatísticas (2.29) e (2.30), utilizando o estimador “naïve”.

Capítulo 3

Equações de estimação generalizadas para medidas repetidas em mais de um fator

Nesse capítulo são apresentadas as estruturas das equações de estimação generalizadas adequadas ao modelo com mais de uma medida repetida, em que as repetições se apresentam em pelo menos dois fatores intra-indivíduos. Os efeitos dos fatores envolvidos e da interação entre eles influenciam não só a matriz de planejamento mas também a estrutura de correlação.

Sem perda de generalidade, e para facilitar a notação, apresentaremos a abordagem de medidas repetidas em apenas dois fatores.

Seja $\mathbf{y}_i = (y_{i11}, y_{i12}, \dots, y_{i1t}, y_{i21}, y_{i22}, \dots, y_{i2t}, \dots, y_{is1}, y_{is2}, \dots, y_{ist})^T$ o vetor de respostas ($st \times 1$) do i -ésimo elemento amostral, $i = 1, 2, \dots, n$, com st observações ao longo de duas dimensões (dois fatores cruzados) ditos A (com s condições de observação) e B (com t condições de observação). Seja ainda, $\mathbf{X}_i = (\mathbf{x}_{i11}, \mathbf{x}_{i12}, \dots, \mathbf{x}_{i1t}, \mathbf{x}_{i21}, \mathbf{x}_{i22}, \dots, \mathbf{x}_{i2t}, \dots, \mathbf{x}_{is1}, \mathbf{x}_{is2}, \dots, \mathbf{x}_{ist})^T$, matriz de planejamento ($st \times p$), com $\mathbf{x}_{ijk} = (x_{ijk1}, x_{ijk2}, \dots, x_{ijkp})^T$, $j = 1, 2, \dots, s$ e $k = 1, 2, \dots, t$, vetores de covariáveis ($p \times 1$) para o indivíduo i na condição j do fator A e na condição k do fator B . Admitimos a *função de distribuição ou densidade de probabilidade* marginal de cada observação y_{ijk} conhecida e pertencente à família exponencial e, ainda, dada por

$$f(y_{ijk}; \theta_{ijk}, \phi) = \exp[\phi \{y_{ijk}\theta_{ijk} - b(\theta_{ijk})\} + c(y_{ijk}, \phi)], \quad (3.1)$$

com $E[y_{ijk}] = \mu_{ijk} = b'(\theta_{ijk})$ e $Var[y_{ijk}] = \phi^{-1}V_{ijk}$, representando a esperança e a variância de y_{ijk} , respectivamente, $\phi^{-1} > 0$ o parâmetro de dispersão e $V_{ijk} = d\mu_{ijk}/d\theta_{ijk} = b''(\theta_{ijk})$ a função de variância. A média é modelada por

$$g(\mu_{ijk}) = \eta_{ijk} = \sum_{l=1}^p x_{ijkl}\beta_l = \mathbf{x}_{ijk}^T \boldsymbol{\beta}, \quad (3.2)$$

em que $g(\cdot)$ é a função de ligação, $\eta_{ijk} = \mathbf{x}_{ijk}^T \boldsymbol{\beta}$ é o preditor linear e $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^T$, com $p < n$ é o vetor de parâmetros desconhecido.

3.1 Experimento com medidas repetidas em dois fatores

Na Tabela 3.1 é representado o esquema de um experimento com dois fatores, sendo A com níveis a_1, a_2, \dots, a_s e B com níveis b_1, b_2, \dots, b_t .

Tabela 3.1: Esquema de um experimento com medidas repetidas em dois fatores.

	a_1			\dots	a_s		
indivíduo	b_1	\dots	b_t	\dots	b_1	\dots	b_t
1	y_{111}	\dots	y_{11t}	\dots	y_{1s1}	\dots	y_{1st}
2	y_{211}	\dots	y_{21t}	\dots	y_{2s1}	\dots	y_{2st}
\vdots	\vdots		\vdots		\vdots		\vdots
n	y_{n11}	\dots	y_{n1t}	\dots	y_{ns1}	\dots	y_{nst}

Cada indivíduo é observado em todas as combinações dos $s \times t$ níveis e y_{ijk} é a observação do i -ésimo indivíduo no j -ésimo nível do fator A e k -ésimo nível do fator B .

A partir das características do problema e do modelo adotado é necessário definir as ma-

trizes e vetores utilizados no modelo e, conseqüentemente, redefinir estruturas para a matriz de correlação de trabalho.

Supondo um experimento em que as medidas se repetem em dois fatores cada um com 3 níveis, ou seja, $s = t = 3$, para cada indivíduo teremos o seguinte vetor de observações, matriz de planejamento e vetor de parâmetros:

$$\mathbf{y}_i = \left[\begin{array}{ccccccccc} y_{i11} & y_{i12} & y_{i13} & y_{i21} & y_{i22} & y_{i23} & y_{i31} & y_{i32} & y_{i33} \end{array} \right]^T,$$

$$\mathbf{X}_i = \left[\begin{array}{ccccccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array} \right] \text{ e } \boldsymbol{\beta} = \left[\begin{array}{ccccccccc} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & \beta_6 & \beta_7 & \beta_8 & \beta_9 \end{array} \right].$$

A matriz de planejamento X_i segue a *parametrização casela de referência* em que x_{il} representa a coluna l , $l = 1, \dots, 9$, da seguinte forma

$$\mathbf{x}_{i1} = \begin{cases} 1, & \text{para todas as observações do indivíduo;} \end{cases}$$

$$\mathbf{x}_{i2} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 2 \text{ do fator A} \\ 0, & \text{caso contrário;} \end{cases}$$

$$\mathbf{x}_{i3} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 3 \text{ do fator A,} \\ 0, & \text{caso contrário;} \end{cases}$$

$$\mathbf{x}_{i4} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 2 \text{ do fator B,} \\ 0, & \text{caso contrário;} \end{cases}$$

$$\mathbf{x}_{i5} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 3 \text{ do fator B,} \\ 0, & \text{caso contrário;} \end{cases}$$

$$\mathbf{x}_{i6} = \mathbf{x}_{i2} * \mathbf{x}_{i4};$$

$$\mathbf{x}_{i7} = \mathbf{x}_{i2} * \mathbf{x}_{i5};$$

$$\mathbf{x}_{i8} = \mathbf{x}_{i3} * \mathbf{x}_{i4};$$

$$\mathbf{x}_{i9} = \mathbf{x}_{i3} * \mathbf{x}_{i5};$$

O operador (*) denota o produto dos vetores elemento a elemento.

Mesmo $\mathbf{R}(\rho)$ não influenciando as propriedades assintóticas dos estimadores de β , é necessário que $\hat{\phi}$ e $\hat{\rho}$ sejam consistentes para atender às condições do Teorema 2.

A matriz de correlação de trabalho, nesse caso, tem determinada estrutura *dentro* dos níveis de cada fator e *entre* os fatores. Por exemplo, podemos ter todas as três sub-estruturas com *padrão uniforme* ou a sub-estrutura sob os níveis do fator A pode ser *padrão uniforme*, sob os níveis do fator B pode ser *AR-1* e entre os dois fatores pode ser *não-estruturada*.

3.2 Equações de estimação generalizadas

Apresentamos nesta seção a teoria de Liang e Zeger (1986) de EEG para o caso de dados com medidas repetidas em mais de um fator. As mudanças ocorrem principalmente na *matriz de correlação de trabalho* para que $\mathbf{R}(\rho)$ incorpore informação sobre a estrutura de dependência intra e entre-fatores.

A equação de estimação generalizada (EEG) para β , nesse caso, é a mesma dada em (2.11):

$$\Psi_n^G(\beta) = \Psi_n^G(\beta, \hat{\rho}(\beta, \phi)) = \sum_{i=1}^n D_i^T \Omega_i^{-1} \mathbf{u}_i = \mathbf{0},$$

com $\Omega_i = \Omega_i(\hat{\rho}, \phi) = \phi^{-1} \mathbf{A}_i^{1/2} \mathbf{R}(\hat{\rho}) \mathbf{A}_i^{1/2}$ e $\hat{\rho}$ estimador consistente de ρ .

Novamente, a equação de estimação não é ótima. Precisamos obter estimadores de ϕ , no caso de ele ser desconhecido, e de ρ , de tal modo que o estimador de β solução de EEG seja consistente e assintoticamente normal, atendendo às condições do Teorema 2. Para obter estes estimadores, utilizamos o mesmo método do Capítulo 2. Definimos o *Resíduo de Pearson* pela expressão

$$r_{ijk} = \frac{y_{ijk} - \mu_{ijk}}{\sqrt{a_{ijk}}}, \quad (3.3)$$

em que $\hat{a}_{ijk} = \hat{V}_{ijk}$, com as mesmas propriedades citadas no capítulo anterior.

Dessa forma, podemos propor o estimador de momentos de ϕ baseado nos resíduos

$$\hat{\phi}^{-1} = \sum_{i=1}^n \sum_{j=1}^s \sum_{k=1}^t \frac{r_{ijk}^2}{(nst - p)}, \quad (3.4)$$

para o qual p é a dimensão do vetor β , utilizado implicitamente no cálculo dos resíduos de Pearson. A correção “ $-p$ ” é utilizada para evitar o problema de viés nos estimadores.

Note-se que $\hat{\phi}$ e r_{ijk} mantêm sua definição como em (2.17) e (2.17), apenas sendo apresentado na notação de dois fatores para medidas repetidas.

3.3 Estruturas de correlação considerando sub-estruturas

Nesta seção apresentamos diferentes estruturas para a matriz de correlação $R(\rho)$ que levam em consideração as sub-estruturas intra-fatores (efeitos principais) e entre-fatores (interações). Para cada estrutura apresentada, propomos estimadores consistentes para o parâmetro ρ que se baseiam nos resíduos de Pearson definidos em (3.3) e são obtidos pelo método dos momentos como em Crowder (1995).

As sub-estruturas na matriz de correlação de trabalho representam as matrizes de correlação entre observações de um mesmo indivíduo, mas *dentro* de cada nível dos fatores ou *entre* diferentes níveis de ambos os fatores.

3.3.1 Estrutura de correlação não-estruturada

Sem perda de generalidade, consideramos o caso em que $s = t = 3$ para exemplificar a matriz $R(\rho)$:

$$R(\rho) = \begin{bmatrix} 1 & \rho_{11,12} & \rho_{11,13} & \rho_{11,21} & \rho_{11,22} & \rho_{11,23} & \rho_{11,31} & \rho_{11,32} & \rho_{11,33} \\ \rho_{12,11} & 1 & \rho_{12,13} & \rho_{12,21} & \rho_{12,22} & \rho_{12,23} & \rho_{12,31} & \rho_{12,32} & \rho_{12,33} \\ \rho_{13,11} & \rho_{13,12} & 1 & \rho_{13,21} & \rho_{13,22} & \rho_{13,23} & \rho_{13,31} & \rho_{13,32} & \rho_{13,33} \\ \rho_{21,11} & \rho_{21,12} & \rho_{21,13} & 1 & \rho_{21,22} & \rho_{21,23} & \rho_{21,31} & \rho_{21,32} & \rho_{21,33} \\ \rho_{22,11} & \rho_{22,12} & \rho_{22,13} & \rho_{22,21} & 1 & \rho_{22,23} & \rho_{22,31} & \rho_{22,32} & \rho_{22,33} \\ \rho_{23,11} & \rho_{23,12} & \rho_{23,13} & \rho_{23,21} & \rho_{23,22} & 1 & \rho_{23,31} & \rho_{23,32} & \rho_{23,33} \\ \rho_{31,11} & \rho_{31,12} & \rho_{31,13} & \rho_{31,21} & \rho_{31,22} & \rho_{31,23} & 1 & \rho_{31,32} & \rho_{31,33} \\ \rho_{32,11} & \rho_{32,12} & \rho_{32,13} & \rho_{32,21} & \rho_{32,22} & \rho_{32,23} & \rho_{32,31} & 1 & \rho_{32,33} \\ \rho_{33,11} & \rho_{33,12} & \rho_{33,13} & \rho_{33,21} & \rho_{33,22} & \rho_{33,23} & \rho_{33,31} & \rho_{33,32} & 1 \end{bmatrix}.$$

A partir da equação de momentos

$$E[r_{ijk}r_{ij'k'}] = \phi^{-1}\rho_{jk,j'k'}, \quad (3.5)$$

oriunda das propriedades do resíduo de Pearson, obtemos um estimador \sqrt{n} -consistente de $\rho_{jk,j'k'}$, para $j = 1, 2, \dots, s$ e $k = 1, 2, \dots, t$ dado por

$$\hat{\rho}_{jk,j'k'} = \hat{\phi} \sum_{i=1}^n \frac{r_{ijk}r_{ij'k'}}{n-p}, \quad (3.6)$$

em que $j \neq j'$ ou $k \neq k'$.

3.3.2 Sub-estruturas de correlação padrão uniforme

Nesse caso, propomos que a matriz de correlação de trabalho possua duas *sub-estruturas de correlação padrão uniforme*. Assim, para $s = t = 3$, temos a seguinte estrutura na *matriz de correlação de trabalho* para cada indivíduo:

$$\mathbf{R}(\boldsymbol{\rho}) = \begin{bmatrix} 1 & \rho_2 & \rho_3 & \rho_1 & \rho_3 & \rho_3 & \rho_1 & \rho_3 & \rho_3 \\ \rho_2 & 1 & \rho_2 & \rho_3 & \rho_1 & \rho_3 & \rho_3 & \rho_1 & \rho_3 \\ \rho_2 & \rho_2 & 1 & \rho_3 & \rho_3 & \rho_1 & \rho_3 & \rho_3 & \rho_1 \\ \rho_1 & \rho_3 & \rho_3 & 1 & \rho_2 & \rho_2 & \rho_1 & \rho_3 & \rho_3 \\ \rho_3 & \rho_1 & \rho_3 & \rho_2 & 1 & \rho_2 & \rho_3 & \rho_1 & \rho_3 \\ \rho_3 & \rho_3 & \rho_1 & \rho_2 & \rho_2 & 1 & \rho_3 & \rho_3 & \rho_1 \\ \rho_1 & \rho_3 & \rho_3 & \rho_1 & \rho_3 & \rho_3 & 1 & \rho_2 & \rho_2 \\ \rho_3 & \rho_1 & \rho_3 & \rho_3 & \rho_1 & \rho_3 & \rho_2 & 1 & \rho_2 \\ \rho_3 & \rho_3 & \rho_1 & \rho_3 & \rho_3 & \rho_1 & \rho_2 & \rho_2 & 1 \end{bmatrix},$$

em que

$\rho_1 = \mathbf{Cor}(y_{ijk}, y_{ij'k})$, é a correlação (constante) entre as observações do indivíduo sob os níveis do fator A fixado o nível do fator B , com $k = 1, 2, \dots, t$ e $1 \leq j \neq j' \leq s$;

$\rho_2 = \mathbf{Cor}(y_{ijk}, y_{ijk'})$ é a correlação (constante) entre as observações do indivíduo sob os níveis do fator B fixado o nível do fator A , com $j = 1, 2, \dots, s$ e $1 \leq k \neq k' \leq t$;

$\rho_3 = \mathbf{Cor}(y_{ijk}, y_{ij'k'})$, a correlação entre as observações do indivíduo entre diferentes níveis de ambos os fatores, com $j \neq j'$ e $k \neq k'$, assumindo que as correlações *entre* diferentes níveis de ambos os fatores também sejam equicorrelatas.

Os estimadores de ρ_1 , ρ_2 e ρ_3 , com base nas propriedades dos resíduos de Pearson, são dados por

$$\hat{\rho}_1 = \hat{\phi} \sum_{i=1}^n \sum_{k=1}^t \sum_{j>j'}^s \frac{\hat{r}_{ijk} \hat{r}_{ij'k}}{nts(s-1)/2 - p}, \quad (3.7)$$

$$\hat{\rho}_2 = \hat{\phi} \sum_{i=1}^n \sum_{j=1}^s \sum_{k>k'}^t \frac{\hat{r}_{ijk} \hat{r}_{ijk'}}{nst(t-1)/2 - p} \quad (3.8)$$

e

$$\hat{\rho}_3 = \hat{\phi} \frac{\sum_{i=1}^n \sum_{j>j'}^s \sum_{k \neq k'}^t \hat{r}_{jk} \hat{r}_{j'k'}}{nst(s-1)(t-1)/2 - p}. \quad (3.9)$$

3.3.3 Sub-estruturas de correlação AR-1

Nessa proposta, temos duas *sub-estruturas de correlação autoregressiva de ordem 1* diferentes, uma para cada fator. Apresentamos abaixo a *matriz de correlação de trabalho* $\mathbf{R}(\rho)$ em que $s = t = 3$ para cada indivíduo:

$$\mathbf{R}(\rho) = \begin{bmatrix} 1 & \rho_2 & \rho_2^2 & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & \rho_1^2 & \rho_1^2\rho_2 & \rho_1^2\rho_2^2 \\ \rho_2 & 1 & \rho_2 & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \rho_1^2\rho_2 & \rho_1^2 & \rho_1^2\rho_2 \\ \rho_2^2 & \rho_2 & 1 & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \rho_1^2\rho_2^2 & \rho_1^2\rho_2 & \rho_1^2 \\ \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & 1 & \rho_2 & \rho_2^2 & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 \\ \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \rho_2 & 1 & \rho_2 & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 \\ \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \rho_2^2 & \rho_2 & 1 & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 \\ \rho_1^2 & \rho_1^2\rho_2 & \rho_1^2\rho_2^2 & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & 1 & \rho_2 & \rho_2^2 \\ \rho_1^2\rho_2 & \rho_1^2 & \rho_1^2\rho_2 & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \rho_2 & 1 & \rho_2 \\ \rho_1^2\rho_2^2 & \rho_1^2\rho_2 & \rho_1^2 & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \rho_2^2 & \rho_2 & 1 \end{bmatrix},$$

em que

$\rho_1^{|j-j'|} = \mathbf{Cor}(y_{ijk}, y_{ij'k})$ é a correlação entre as observações do indivíduo sob os níveis do fator A fixado o nível do fator B , com $k = 1, 2, \dots, t$ e $1 \leq j \neq j' \leq s$, presumindo que a correlação entre as observações de um indivíduo *dentro* de um nível do segundo fator diminui à medida que aumenta a distância entre os níveis do primeiro fator;

$\rho_2^{|k-k'|} = \mathbf{Cor}(y_{ijk}, y_{ijk'})$ é a correlação entre as observações do indivíduo sob os níveis do fator B fixado cada nível do fator A , com $j = 1, 2, \dots, s$ e $1 \leq k' \neq k \leq t$, presumindo que a

correlação entre as observações de um indivíduo *dentro* de um nível do primeiro fator diminui à medida que aumenta a distância entre os níveis do segundo fator;

e

$\rho_2^{|j-j'|} \rho_1^{|k-k'|} = \mathbf{Cor}(y_{ijk}, y_{ij'k'})$ é a correlação *entre* as observações do indivíduo em diferentes níveis de ambos os fatores, com $j \neq j'$ e $k \neq k'$, assumindo que a correlação *entre* diferentes níveis de ambos os fatores diminui com o aumento da distância entre os dois níveis.

Os estimadores para ρ_1 e ρ_2 , com base nas propriedades dos resíduos de Pearson, são dados por

$$\hat{\rho}_1 = \hat{\phi} \sum_{i=1}^n \sum_{k=1}^t \sum_{j=1}^{s-1} \frac{\hat{r}_{ijk} \hat{r}_{i(j+1)k}}{nt(s-1)}. \quad (3.10)$$

e

$$\hat{\rho}_2 = \hat{\phi} \sum_{i=1}^n \sum_{j=1}^s \sum_{k=1}^{t-1} \frac{\hat{r}_{ijk} \hat{r}_{ij(k+1)}}{ns(t-1)} \quad (3.11)$$

3.3.4 Estrutura de correlação mista (sub-estrutura AR-1 e sub-estrutura padrão uniforme)

Em uma situação em que a estrutura de dependência não é mesma para os dois fatores, pode-se propor sub-estruturas diferentes para cada um deles. Aqui presumimos uma *sub-estrutura de correlação autoregressiva de ordem 1* para o primeiro fator, enquanto para o segundo fator será assumida uma *sub-estrutura de correlação uniforme*. Apresentamos a seguir a *matriz de correlação de trabalho* $\mathbf{R}(\rho)$ com $s = t = 3$ para cada indivíduo.

$$\mathbf{R}(\boldsymbol{\rho}) = \begin{bmatrix} 1 & \rho_2 & \rho_2^2 & \vdots & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & \vdots & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 \\ \rho_2 & 1 & \rho_2 & \vdots & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \vdots & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 \\ \rho_2^2 & \rho_2 & 1 & \vdots & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \vdots & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 \\ \hline \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & \vdots & 1 & \rho_2 & \rho_2^2 & \vdots & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 \\ \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \vdots & \rho_2 & 1 & \rho_2 & \vdots & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 \\ \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \vdots & \rho_2^2 & \rho_2 & 1 & \vdots & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 \\ \hline \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & \vdots & \rho_1 & \rho_1\rho_2 & \rho_1\rho_2^2 & \vdots & 1 & \rho_2 & \rho_2^2 \\ \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \vdots & \rho_1\rho_2 & \rho_1 & \rho_1\rho_2 & \vdots & \rho_2 & 1 & \rho_2 \\ \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \vdots & \rho_1\rho_2^2 & \rho_1\rho_2 & \rho_1 & \vdots & \rho_2^2 & \rho_2 & 1 \end{bmatrix},$$

em que

$\rho_1 = \text{Cor}(y_{ijk}, y_{ij'k})$ é a correlação entre as observações do indivíduo sob os níveis do fator A , fixando o nível do fator B , com $k = 1, 2, \dots, t$ e $1 \leq j \neq j' \leq s$, sub-estrutura de correlação padrão uniforme;

$\rho_2^{|k-k'|} = \text{Cor}(y_{ijk}, y_{ij'k'})$ é a correlação entre as observações do indivíduo sob os níveis do fator B , fixado cada nível do fator A , com $j = 1, 2, \dots, s$ e $1 \leq k' \neq k \leq t$, sub-estrutura de correlação AR-1;

e

$\rho_2^{|j-j'|} \rho_1 = \text{Cor}(y_{ijk}, y_{ij'k'})$ é a correlação *entre* as observações do indivíduo entre diferentes níveis de ambos os fatores, com $j \neq j'$ e $k \neq k'$, propondo que a correlação *entre* os diferentes níveis de ambos os fatores diminui com o aumento da distância entre os níveis do segundo fator, contudo um pouco mais fraca pois é multiplicada pela correlação entre os níveis do primeiro fator.

Os estimadores de ρ_1 e ρ_2 , com base nos réiduos de Pearson, são dados por

$$\hat{\rho}_1 = \hat{\phi} \sum_{i=1}^n \sum_{k=1}^t \sum_{j>j'}^s \frac{\hat{r}_{ijk} \hat{r}_{ij'k}}{nts(s-1)/2 - p}. \quad (3.12)$$

e

$$\hat{\rho}_2 = \hat{\phi} \sum_{i=1}^n \sum_{j=1}^s \sum_{k=1}^{t-1} \frac{\hat{r}_{ijk} \hat{r}_{ij(k+1)}}{ns(t-1)} \quad (3.13)$$

No caso da estrutura mista a proposta pode apresentar muitas variações, por exemplo, o inverso do exposto anteriormente com sub-estrutura de correlação uniforme para o primeiro fator e sub-estrutura de correlação AR-1 para o segundo fator, dependendo, é claro, da natureza dos dados. Pode-se ainda, propor uma correlação constante entre as observações de diferentes níveis de ambos o fatores.

3.4 Estimadores de β e de sua matriz de covariâncias

Dados os estimadores \sqrt{n} -consistentes de ϕ e de ρ , temos que $\hat{\beta}_G$ é um estimador consistente de β com distribuição assintótica dada por

$$\sqrt{n} \left(\hat{\beta}_G - \beta \right) \xrightarrow{\mathcal{D}} N_p \left(\mathbf{0}, J_G^{-1} \right)$$

em que

$$J_G = \lim_{n \rightarrow \infty} \frac{J_{nG}}{n},$$

com

$$J_{nG} = \left\{ \sum_{i=1}^n S_i \right\}^T \left\{ \sum_{i=1}^n V_i \right\}^{-1} \left\{ \sum_{i=1}^n S_i \right\},$$

$$S_i = E_{\beta} \left[\frac{\partial}{\partial \beta^T} D_i^T \Omega_i^{-1} \mathbf{u}_i \right] = D_i^T \Omega_i^{-1} E_{\beta} \left[\frac{\partial}{\partial \beta^T} \mathbf{u}_i \right] = -D_i^T \Omega_i^{-1} D_i$$

e

$$V_i = E_{\beta} [D_i^T \Omega_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \Omega_i^{-1} D_i] = D_i^T \Omega_i^{-1} Cov(\mathbf{y}_i) \Omega_i^{-1} D_i$$

.

Os resultados do Teorema 2 estão garantidos pela consistência dos estimadores de $\boldsymbol{\rho}$ e de ϕ . O estimador “robusto” para a matriz de covariâncias de $\hat{\boldsymbol{\beta}}_G$ é dado por:

$$\hat{\mathbf{J}}_{nG}^{-1} = \left\{ \sum_{i=1}^n S_i \right\}^{-1} \left\{ \sum_{i=1}^n D_i^T \Omega_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \Omega_i^{-1} D_i \right\} \left\{ \sum_{i=1}^n S_i \right\}^{-1}. \quad (3.14)$$

Uma estimativa de (3.14) é obtida quando se avaliam as quantidades nos valores das estimativas consistentes de $\boldsymbol{\rho}$, $\boldsymbol{\beta}$ e ϕ .

A estimação dos parâmetros que modelam a média é realizada por método iterativo adaptado do algoritmo apresentado no Capítulo 2. Apenas a mudança na estrutura da matriz de correlação de trabalho precisa ser considerada na atualização do estimador a cada interação.

As estatísticas dos testes de hipóteses, tanto de *Wald* quanto de *Score* mantêm a mesma forma, pois já consideram a matriz $\mathbf{R}(\boldsymbol{\rho})$ quando necessário.

Um método alternativo de estimação para os parâmetros de correlação é dada por Prentice e Zhao (1991), que sugere o uso de *EEG* para estimar conjuntamente $\boldsymbol{\beta}$ e $\boldsymbol{\rho}$.

Capítulo 4

Aplicações

Neste capítulo são analisados três conjuntos de dados utilizando-se o modelo de equações de estimação para problemas com medidas repetidas em mais de um fator apresentado no Capítulo 3 e uma rotina desenvolvida no software livre R (<http://www.r-project.org/>) a partir do pacote *gee* apresentada no Apêndice A. Para verificação da qualidade do ajuste dos modelos utilizamos as técnicas de diagnósticos disponíveis em <http://www.ime.usp.br/~mkelly/> (Venezuela et al., 2007).

4.1 Aplicação 1

O conjunto de dados utilizado nessa aplicação faz parte do estudo de Rodrigues (2007). O objetivo desse estudo é verificar o efeito dos fatores de utilização de *Barra*, instrumento odontológico de retenção (*níveis: com e sem*) e *Lado* da mastigação (*níveis: direito e esquerdo*) sobre a variável intensidade da *Força* (Kgf) aplicada durante a mordida realizada pelo paciente. Cada paciente foi avaliado nas quatro combinações dos níveis dos fatores.

Optou-se pelo ajuste de um modelo considerando a distribuição *Gama* na expressão (3.1), com função de ligação “*logaritmica*” em (3.2) pelo fato de a variável resposta ser positiva e apresentar uma assimetria à direita identificada pelo histograma apresentado na Figura 4.1.

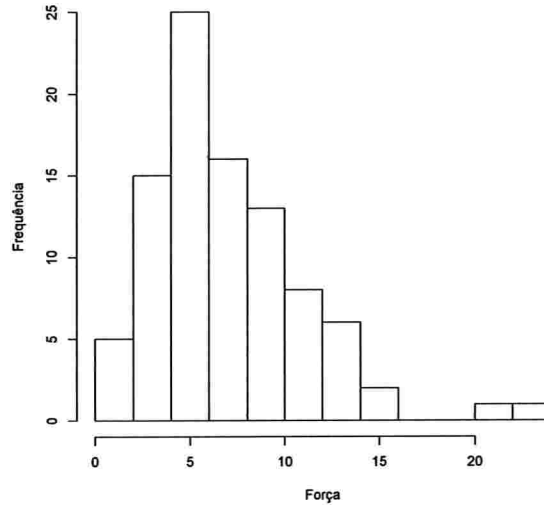


Figura 4.1: Histograma para a variável *Força*.

A média da variável *Força* é dada por

$$\log(\mu_{ijk}) = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4},$$

$i = 1, 2, \dots, 23$, $j = 1, 2$ e $k = 1, 2$, em que

$$x_{i1} = \begin{cases} 1, & \text{para todas as observações do indivíduo;} \end{cases}$$

$$x_{i2} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 2 \text{ (com) do fator } \textit{Barra}, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i3} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 2 \text{ (direito) do fator } \textit{Lado}, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i4} = x_{i2} \times x_{i3}.$$

Adotamos sub-estruturas de correlação de trabalho *padrão uniforme* em ambos os fatores (ver seção 3.3.2), devido ao fato de que possuem apenas dois níveis cada e outra sub-estrutura não agregaria informação. Dessa forma, temos três parâmetros de correlação, $\rho = (\rho_1, \rho_2, \rho_3)^T$, em que

ρ_1 é a correlação entre as observações do paciente nos níveis do fator *Barra*, fixados os níveis do fator *Lado*,

ρ_2 é a correlação entre as observações do paciente nos níveis do fator *Lado*, fixados os níveis do fator *Barra*, e

ρ_3 é a correlação entre as observações do paciente em diferentes níveis do fator *Barra* e do fator *Lado*.

Na Tabela 4.1 são apresentados os coeficientes obtidos pelo ajuste do modelo. Ao nível de 5% de significância verificamos somente a presença do efeito de *Barra* (valor-p < 0,001). Os valores-p foram calculados com base no estimador “robusto”.

Tabela 4.1: Estimativas e erros padrões dos parâmetros do modelo de regressão gama com sub-estruturas de correlação padrão uniforme para a variável *Força*.

Parâmetro	Estimativa	E. P. Naïve	E. P. Robusto	z (Naïve)	z (Robusto)	valor-p
β_1	1,638	0,100	0,086	16,358	19,150	< 0,001
β_2	0,621	0,096	0,081	6,473	7,639	< 0,001
β_3	-0,106	0,094	0,101	-1,122	-1,043	0,297
β_4	-0,035	0,124	0,108	-0,286	-0,330	0,742
ϕ	0,231					

Em (4.1) é apresentada uma estimativa da matriz de correlação de trabalho do modelo de regressão gama para a variável *Força*, ajustado com sub-estruturas de correlação padrão uniforme:

$$\mathbf{R}(\boldsymbol{\rho}) = [\rho_{jk}] = \begin{bmatrix} 1,00 & 0,56 & 0,54 & 0,48 \\ 0,56 & 1,00 & 0,48 & 0,54 \\ 0,54 & 0,48 & 1,00 & 0,56 \\ 0,48 & 0,54 & 0,56 & 1,00 \end{bmatrix}. \quad (4.1)$$

Aplicando as técnicas de diagnóstico desenvolvidas por Venezuela et al. (2007), calculamos a *Distância de Cook* e os *Resíduos padronizados* para identificar pontos influentes ou aberrantes. E ainda, o *Gráfico de probabilidades meio normal* para verificar a adequação do modelo. Os resultados são apresentados nas Figuras 4.2, 4.3 e 4.4, respectivamente.

Na Figura 4.2 identificamos como possível ponto influente uma observação do oitavo indivíduo com valor da distância de Cook maior que os demais. No entanto, após deletarmos as observações desse indivíduo, verificamos que a inferência sobre os parâmetros do modelo não se altera.

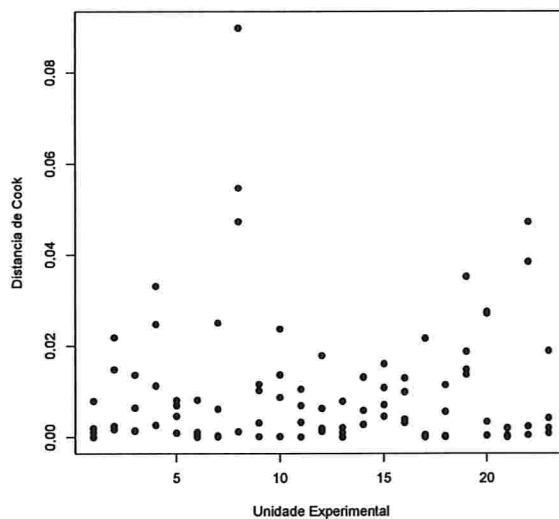


Figura 4.2: Distância de Cook para o modelo de regressão gama ajustado com sub-estruturas de correlação padrão uniforme para a variável *Força*.

Na Figura 4.3 não identificamos possíveis pontos aberrantes com resíduo maior que os demais.

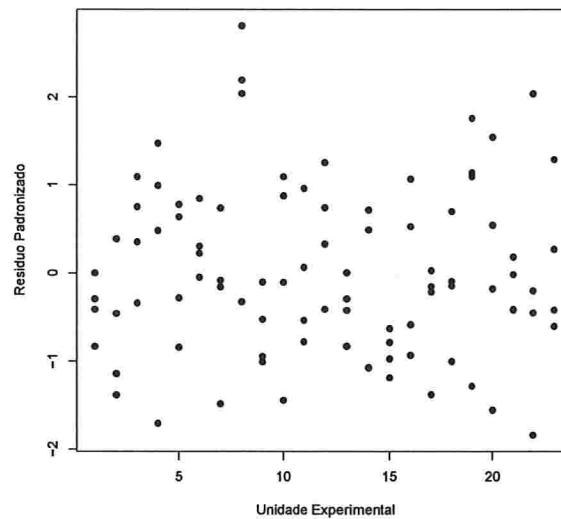


Figura 4.3: Resíduos padronizados para o modelo de regressão gama ajustado com sub-estruturas de correlação padrão uniforme para a variável *Força*.

Além disso, o *gráfico de probabilidades meio-normal com envelope simulado* apresentado na Figura 4.4 apresenta todos os pontos dentro da banda de confiança de 95%. Assim, concluímos que o ajuste do modelo é adequado à variável *Força*.

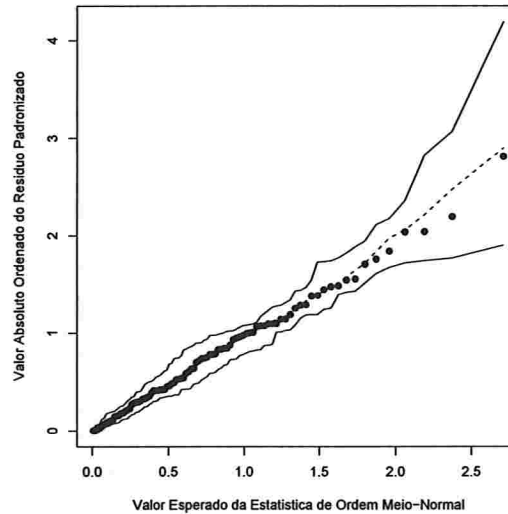


Figura 4.4: Gráfico de probabilidade meio-normal com envelope simulado para o modelo de regressão gama ajustado com sub-estruturas de correlação padrão uniforme para a variável *Força*.

4.2 Aplicação 2

O conjunto de dados utilizado nessa aplicação foi extraído de Winer et al. (1991), pp. 582. O objetivo é avaliar o efeito do fator *B* e do fator *C*, ambos com 3 níveis sobre a *variável resposta*. Cada um dos seis indivíduos foi avaliado nas nove combinações dos fatores num total de 54 observações.

Inicialmente, optamos por ajustar um modelo presumindo distribuição *Normal* com função de ligação “*identidade*”, principalmente porque, na referência, os dados eram tratados sob essa distribuição. No entanto, o algoritmo não convergiu nessa situação para nenhuma estrutura de correlação. Dessa forma, propusemos um ajuste com distribuição gama em (3.1) e uma função de ligação “*logaritmica*” em (3.2).

A média da *variável resposta* é dada por

$$\log(\mu_{ijk}) = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} + \beta_9 x_{i9},$$

$i = 1, \dots, 6$, $j = 1, 2, 3$ e $k = 1, 2, 3$, em que

$$x_{i1} = \begin{cases} 1, & \text{para todas as observações do indivíduo;} \end{cases}$$

$$x_{i2} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 2 \text{ do fator } B, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i3} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 3 \text{ do fator } B, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i4} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 2 \text{ do fator } C, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i5} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 3 \text{ do fator } C, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i6} = x_{i2} \times x_{i4};$$

$$x_{i7} = x_{i2} \times x_{i5};$$

$$x_{i8} = x_{i3} \times x_{i4};$$

$$x_{i9} = x_{i3} \times x_{i5}.$$

Como os dados são apenas teóricos e não apresentam descrição alguma sobre estrutura de dependência, com o objetivo de ilustração, admitimos uma sub-estrutura de correlação de trabalho *AR-1* dentro dos níveis de ambos os fatores (ver seção 3.3.3). Dessa forma, temos dois parâmetros de correlação, $\boldsymbol{\rho} = (\rho_1, \rho_2)^T$, em que

ρ_1 é a correlação entre as observações consecutivas do indivíduo nos níveis do fator B , fixados os níveis do fator C ,

ρ_2 é a correlação entre as observações consecutivas do indivíduo nos níveis do fator C , fixados os níveis do fator B .

Na Tabela 4.2 apresentamos as estimativas para os parâmetros do modelo. Os valores-p foram obtidos com base no estimador robusto e no estimador naïve, no entanto, como nossa amostra é pequena a inferência foi considerada a partir do estimador “naïve”. Assim, ao nível de significância de 5%, podemos verificar apenas a presença do efeito do fator B .

Tabela 4.2: Estimativas e erros padrões dos parâmetros do modelo de regressão gama com estrutura de correlação AR-1 para a *variável resposta*.

Parâmetro	Estimativa	E. P.		valor-z		valor-p	
		Naïve	Robusto	Naïve	Robusto	Naïve	Robusto
β_1	3,546	0,124	0,142	28,548	24,871	< 0,001	< 0,001
β_2	-0,244	0,176	0,148	-1,388	-1,644	0,165	0,100
β_3	-0,619	0,176	0,109	-3,524	-5,656	< 0,001	< 0,001
β_4	0,192	0,176	0,089	1,092	2,133	0,275	0,033
β_5	0,425	0,176	0,083	2,417	5,137	0,016	< 0,001
β_6	-0,023	0,248	0,100	-0,092	-0,229	0,926	0,818
β_7	0,061	0,248	0,116	0,247	0,529	0,805	0,596
β_8	-0,005	0,248	0,056	-0,021	-0,091	0,984	0,927
β_9	0,130	0,248	0,063	0,523	2,063	0,601	0,039
ϕ	0,093						

Resumimos os resultados da Tabela 4.2 construindo a Tabela 4.3 de Análise de Variância (ANOVA), na qual também observa-se apenas o efeito do fator B ao nível de significância de 5%. Como o tamanho da amostra é pequeno, consideramos as estatísticas com base no estimador “naïve”.

Tabela 4.3: ANOVA (Teste de Wald) para os dados da Aplicação 2.

Efeito	gl	Robusto		Naïve	
		QW	valor-p	QW	valor-p
Fator B	2	110,090	< 0,001	12,607	0,002
Fator C	2	130,043	< 0,001	5,859	0,053
Interação	4	29,327	< 0,001	0,390	0,983

Em (4.2) apresentamos a matriz de correlação de trabalho estimada para o modelo de regressão ajustado para a *variável resposta*. Podemos verificar que as estimativas seguem as sub-estruturas propostas, mas os valores são baixos, com resultados menores do que 0,0001. Dessa forma, podemos concluir que os dados poderiam ser analisados sob estrutura de correlação independente. No entanto, uma análise com essa estrutura não mostrou modificação na inferência.

$$R(\rho) = \begin{bmatrix} 1 & 7,4e-05 & 5,6e-09 & 7,7e-05 & 5,8e-09 & 4,3e-13 & 6,0e-09 & 4,5e-13 & 3,3e-17 \\ 7,4e-05 & 1 & 7,4e-05 & 5,8e-09 & 7,7e-05 & 5,8e-09 & 4,5e-13 & 6,0e-09 & 4,5e-13 \\ 5,6e-09 & 7,4e-05 & 1 & 4,3e-13 & 5,8e-09 & 7,7e-05 & 3,3e-17 & 4,5e-13 & 6,0e-09 \\ 7,7e-05 & 5,8e-09 & 4,3e-13 & 1 & 7,4e-05 & 5,6e-09 & 7,7e-05 & 5,8e-09 & 4,3e-13 \\ 5,8e-09 & 7,7e-05 & 5,8e-09 & 7,4e-05 & 1 & 7,4e-05 & 5,8e-09 & 7,7e-05 & 5,8e-09 \\ 4,3e-13 & 5,8e-09 & 7,7e-05 & 5,6e-09 & 7,4e-05 & 1 & 4,3e-13 & 5,8e-09 & 7,7e-05 \\ 6,0e-09 & 4,5e-13 & 3,3e-17 & 7,7e-05 & 5,8e-09 & 4,3e-13 & 1 & 7,4e-05 & 5,6e-09 \\ 4,5e-13 & 6,0e-09 & 4,5e-13 & 5,8e-09 & 7,7e-05 & 5,8e-09 & 7,4e-05 & 1 & 7,4e-05 \\ 3,3e-17 & 4,5e-13 & 6,0e-09 & 4,3e-13 & 5,8e-09 & 7,7e-05 & 5,6e-09 & 7,4e-05 & 1 \end{bmatrix} \quad (4.2)$$

Técnicas de diagnósticos também foram aplicadas ao modelo para verificação da qualidade do ajuste.

Na Figura 4.5, na qual são apresentadas as distâncias de Cook, identificamos a observação 22 que se refere ao indivíduo 3 como possível ponto influente.

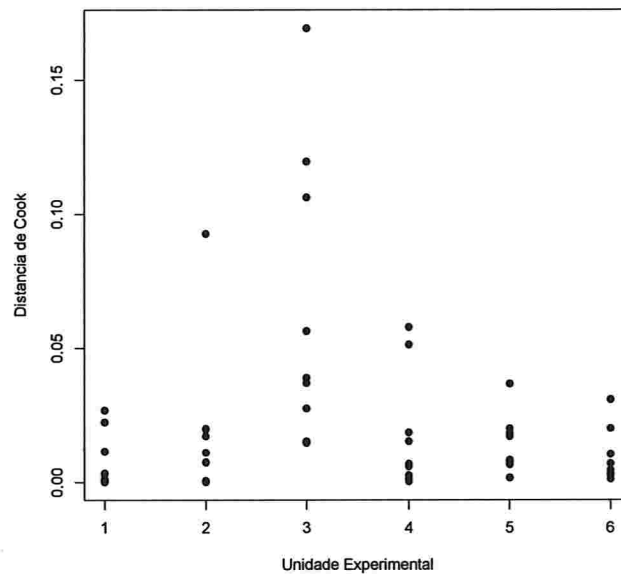


Figura 4.5: Distância de Cook para o modelo de regressão gama ajustado com sub-estruturas de correlação AR-1 para a *variável resposta*.

Na Figura 4.6, com os resíduos padronizados, identificamos as observações 19, 22 e 25 também do indivíduo 3 como possíveis pontos aberrantes. No entanto, temos uma amostra de tamanho reduzido e optamos pela não delegação de observações.

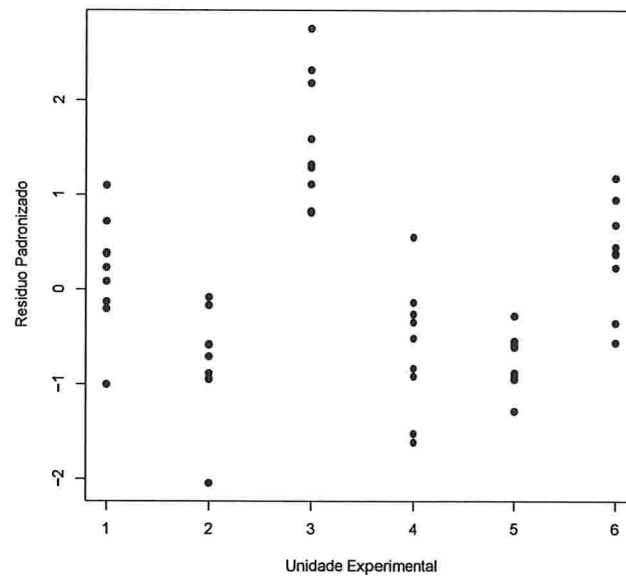


Figura 4.6: Resíduos Padronizados para o modelo de regressão gama ajustado com sub-estruturas de correlação AR-1 para a *variável resposta*.

Além disso, o gráfico de probabilidades meio-normal com envelope simulado, Figura 4.7, não mostra nenhum ponto fora da banda de confiança de 95%, indicando-nos um ajuste adequado do modelo proposto.

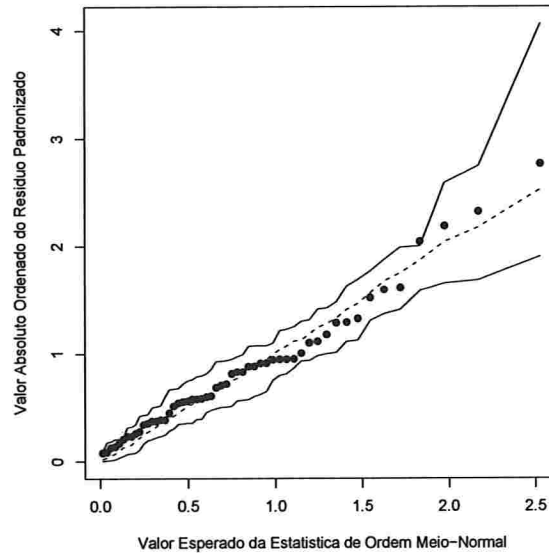


Figura 4.7: Gráfico de probabilidade meio-normal com envelope simulado para o modelo de regressão gama ajustado com sub-estruturas de correlação AR-1 para a *variável resposta*.

4.3 Aplicação 3

O conjunto de dados apresentado nessa aplicação foi retirado de Winer et al. (1991), pp. 546. Seu objetivo é avaliar o efeito dos fatores *Período de tempo*, com dez minutos de duração cada (*níveis: 10, 20 e 30*) e a monitoração de *Canal de Frequência* (*níveis: c1, c2 e c3*) sobre a variável *Escore de acurácia* em realizar os ajustes solicitados. Seis indivíduos foram avaliados nas nove combinações dos fatores num total de 54 observações.

Optou-se pelo ajuste de um modelo considerando a distribuição *Normal* em (3.1) com função de ligação “identidade” em (3.2), pois, pelo histograma, apresentado na Figura 4.8, a variável *Escore de acurácia* apresenta simetria e a hipótese de ajuste à essa distribuição não é rejeitada quando aplicado teste de *Kolmogorov-Smirnov*.

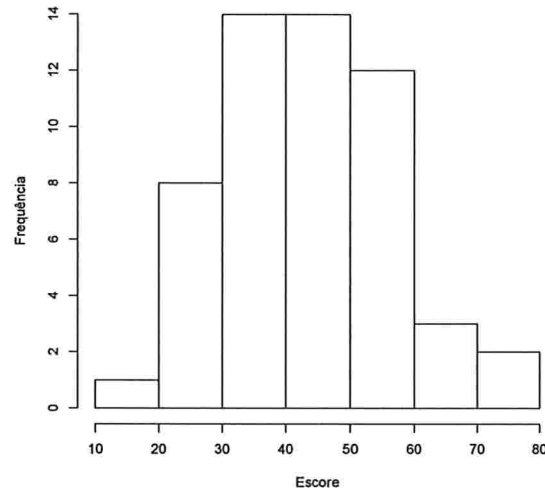


Figura 4.8: Histograma para a variável *Escore de acurácia*.

A média da variável *Escore de acurácia* é dada por

$$\mu_{ijk} = \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + \beta_5 x_{i5} + \beta_6 x_{i6} + \beta_7 x_{i7} + \beta_8 x_{i8} + \beta_9 x_{i9},$$

$i = 1, \dots, 6$, $j = 1, 2, 3$ e $k = 1, 2, 3$, em que

$$x_{i1} = \begin{cases} 1, & \text{para todas as observações do indivíduo;} \end{cases}$$

$$x_{i2} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 2, \text{ } 20 \text{ minutos, do fator } \textit{Período}, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i3} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } j = 3, \text{ } 30 \text{ minutos, do fator } \textit{Período}, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i4} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 2, \text{ } c2, \text{ do fator } \textit{Canal de frequência}, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i5} = \begin{cases} 1, & \text{se a observação é obtida sob o nível } k = 3, c3, \text{ do fator } \textit{Canal de frequência}, \\ 0, & \text{caso contrário;} \end{cases}$$

$$x_{i6} = x_{i2} \times x_{i4};$$

$$x_{i7} = x_{i2} \times x_{i5};$$

$$x_{i8} = x_{i3} \times x_{i4};$$

$$x_{i9} = x_{i3} \times x_{i5}.$$

Devido a natureza dos dados, em que o fator *Período* por referir-se a tempo pode gerar desgaste do indivíduo ao ajustar o *Canal de Frequência* parecendo adequar-se à uma sub-estrutura de correlação *AR-1*, enquanto o fator *Canal de Frequência* não aparenta justificar alteração na correlação para os diferentes *Períodos* sendo razoável a utilização da sub-estrutura *padrão uniforme*, presumimos a estrutura de correlação *mista* (ver seção 3.3.4). Dessa forma, temos dois parâmetros de correlação, $\rho = (\rho_1, \rho_2)^T$, em que

ρ_1 é a correlação entre as observações consecutivas do indivíduo nos níveis do fator *Canal de frequência*, fixando o fator *Período*, e

ρ_2 é a correlação entre as observações do indivíduo nos níveis do fator *Período*, fixando o fator *Canal de frequência*.

A Tabela 4.4 apresenta as estimativas para os parâmetros do modelo. Ao nível de significância de 5% verificamos que os parâmetros referentes ao fator *Período* e ao fator *Canal de frequência* são significativos, mas os parâmetros referentes as iterações não são. Os valores-p foram obtidos com base no estimador *robusto*, contudo, a inferência permanece a mesma

quando consideramos o estimador “naïve”.

Tabela 4.4: Estimativas e erros padrões dos parâmetros do modelo de regressão normal com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável *score*.

Parâmetro	Estimativa	E. P. Naïve	E. P. Robusto	z (Naïve)	z (Robusto)	valor-p
β_1	48,000	3,749	3,440	12,803	13,954	< 0,001
β_2	-21,000	4,384	3,000	-4,790	-7,000	< 0,001
β_3	-10,833	4,384	3,227	-2,471	-3,357	0,001
β_4	15,000	3,905	1,247	3,840	12,027	< 0,001
β_5	4,000	3,016	1,269	1,326	3,151	0,002
β_6	0,500	4,567	1,954	0,109	0,256	0,798
β_7	3,167	3,527	2,397	0,898	1,321	0,186
β_8	2,000	4,567	2,549	0,438	0,784	0,433
β_9	1,000	3,527	2,656	0,283	0,376	0,706
ϕ	0,012					

Resumimos os resultados da Tabela 4.4 construindo a Tabela 4.5 de Análise de Variância (ANOVA), na qual não observamos efeito de interação entre os fatores *Canal de frequência* e *Período*, mas observamos os efeitos marginais destes fatores, ao nível de significância de 5%. Como o número de indivíduos da amostra é pequeno, consideramos também as estatísticas com base no estimador “naïve”, mas a inferência não se modifica.

Tabela 4.5: ANOVA (Teste de Wald) para os dados da Aplicação 3.

Efeito	gl	Robusto		Naïve	
		QW	valor-p	QW	valor-p
Fator <i>Canal de frequência</i>	2	76,504	< 0,001	25,009	< 0,001
Fator <i>Período</i>	2	151,530	< 0,001	25,786	< 0,001
Interação	4	6,654	0,155	3,205	0,524

Em (4.3) apresentamos a matriz de correlação de trabalho estimada para o modelo de regressão ajustado para a variável *Escore de acurácia* com estrutura de correlação mista, AR-1 no fator *Período* e padrão uniforme no fator *Canal de frequência*.

$$R(\rho) = \begin{bmatrix} 1,00 & 0,68 & 0,46 & 0,32 & 0,21 & 0,14 & 0,32 & 0,21 & 0,14 \\ 0,68 & 1,00 & 0,68 & 0,21 & 0,32 & 0,21 & 0,21 & 0,32 & 0,21 \\ 0,46 & 0,68 & 1,00 & 0,14 & 0,21 & 0,32 & 0,14 & 0,21 & 0,32 \\ 0,32 & 0,21 & 0,14 & 1,00 & 0,68 & 0,46 & 0,32 & 0,21 & 0,14 \\ 0,21 & 0,32 & 0,21 & 0,68 & 1,00 & 0,68 & 0,21 & 0,32 & 0,21 \\ 0,14 & 0,21 & 0,32 & 0,46 & 0,68 & 1,00 & 0,14 & 0,21 & 0,32 \\ 0,32 & 0,21 & 0,14 & 0,32 & 0,21 & 0,14 & 1,00 & 0,68 & 0,46 \\ 0,21 & 0,32 & 0,21 & 0,21 & 0,32 & 0,21 & 0,68 & 1,00 & 0,68 \\ 0,14 & 0,21 & 0,32 & 0,14 & 0,21 & 0,32 & 0,46 & 0,68 & 1,00 \end{bmatrix} \quad (4.3)$$

Novamente, as técnicas de diagnóstico foram aplicadas ao modelo para verificação da qualidade do ajuste.

Na Figura 4.9, a qual apresenta as distâncias de Cook para as unidades experimentais, identificamos uma observação do indivíduo 3 com valor maior do que as demais, o que pode significar um possível ponto influente.

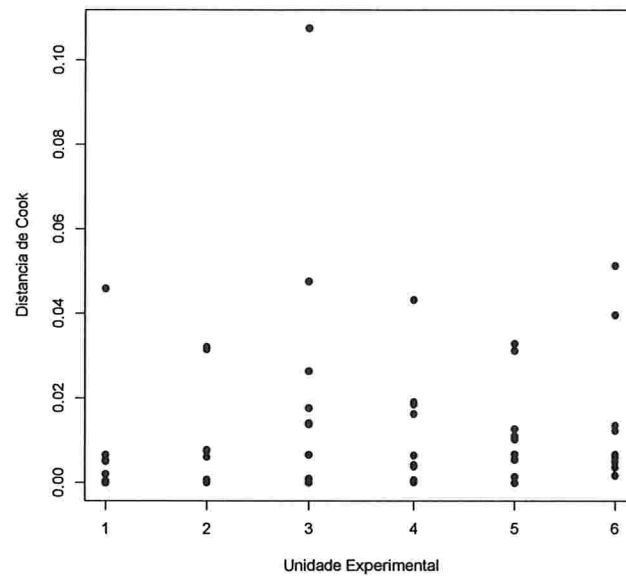


Figura 4.9: Distância de Cook para o modelo de regressão normal ajustado com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável *Escore*.

Já na Figura 4.10, na qual são expostos os resíduos de padronizados, podemos verificar que existe um ponto fora dos limites referente a mesma observação do indivíduo 3 que apresentou maior distância de Cook. No entanto, novamente contamos com uma amostra de tamanho reduzido e optamos por não deletar as observações do indivíduo.

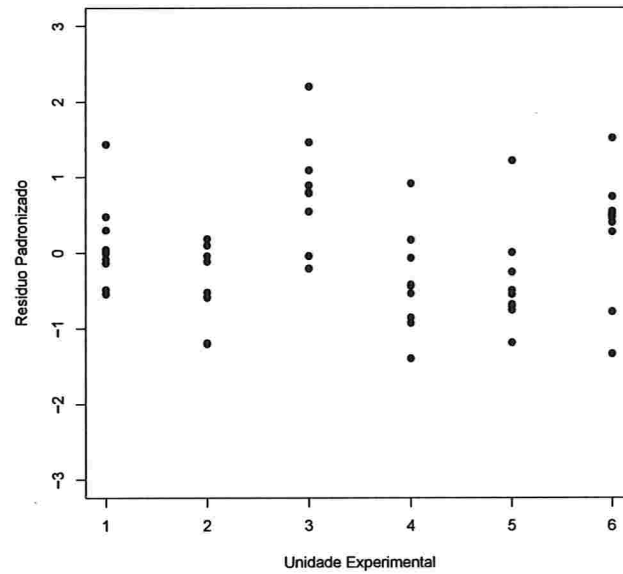


Figura 4.10: Resíduos Padronizado para o modelo de regressão normal ajustado com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável *Score*.

Contudo, o gráfico de probabilidades meio-normal com envelope simulado, Figura 4.11, apresenta todos os pontos dentro da banda de confiança de 95%, o que nos aponta um ajuste adequado do modelo proposto.

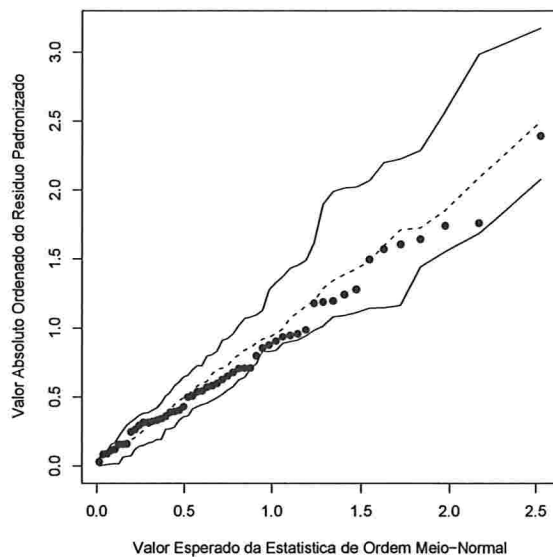


Figura 4.11: Gráfico de probabilidade meio-normal com envelope simulado para o modelo de regressão normal ajustado com estrutura de correlação mista (AR-1 e padrão uniforme) para a variável escore.

A referência da qual foram retirados os dados tratava a estrutura de correlação como *padrão uniforme*. Dessa forma, ajustamos os dados com sub-estruturas de correlação *padrão uniforme* para uma comparação com os resultados da estrutura de correlação *mista (AR-1 e uniforme)*. Nesse caso, a inferência não se modificou, no entanto, os gráficos de diagnósticos se mostraram piores como podemos observar nas figuras 4.12, 4.13 e 4.14.

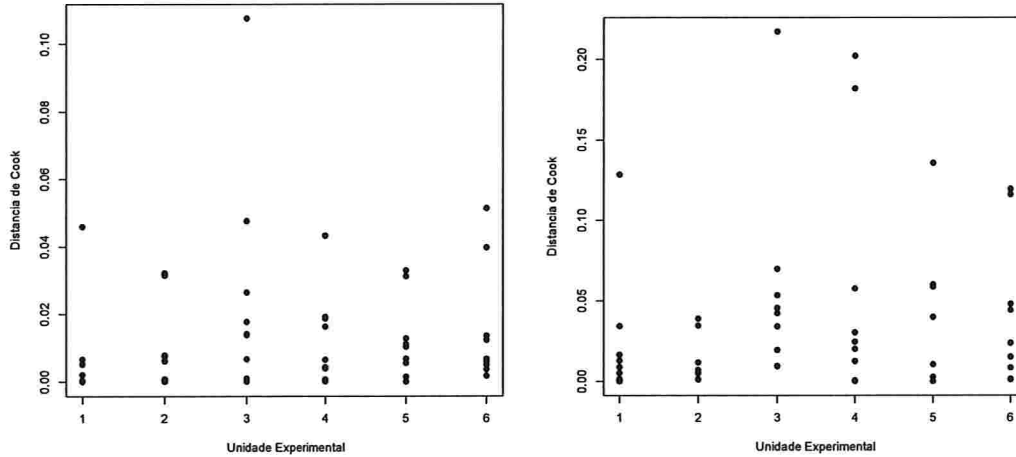


Figura 4.12: Comparação das distâncias de Cook entre o ajuste com estrutura de correlação mista (AR-1 e uniforme) e ajuste com ambas as sub-estruturas de correlação uniforme.

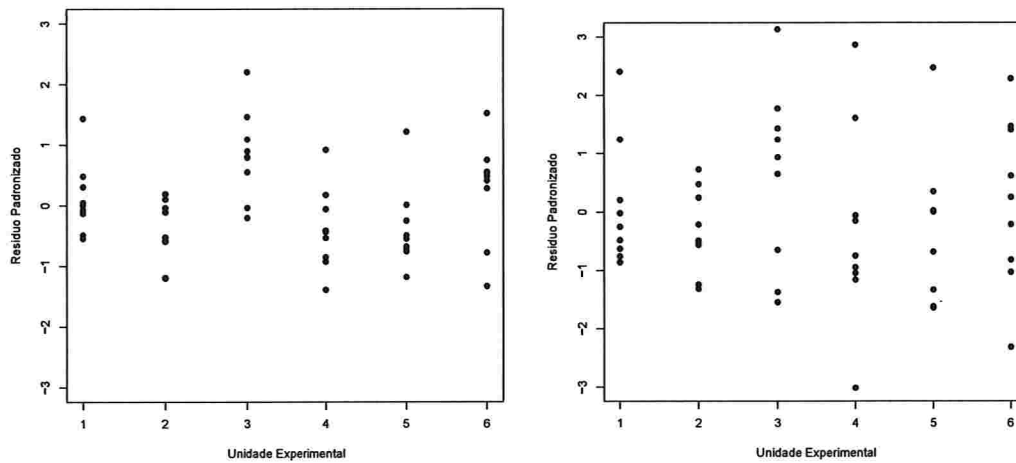


Figura 4.13: Comparação dos resíduos padronizados entre ajuste com estrutura de correlação mista (AR-1 e uniforme) e ajuste com ambas as sub-estruturas de correlação uniforme.

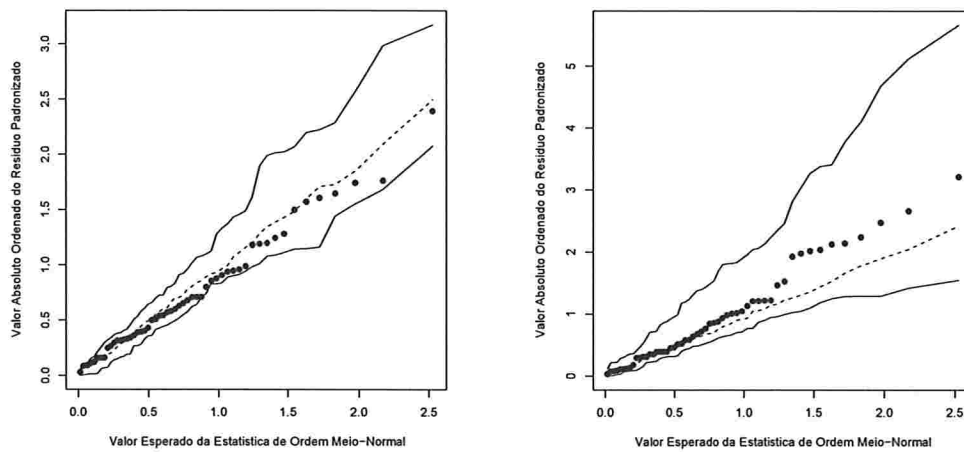


Figura 4.14: Comparação dos gráficos de probabilidade meio-normal com envelope simulado entre ajuste com estrutura de correlação mista (AR-1 e uniforme) e ajuste com ambas as sub-estruturas de correlação uniforme.

Capítulo 5

Conclusões e estudos futuros

Nesse trabalho estendemos as equações de estimação generalizadas desenvolvidas por Liang e Zeger (1986) para medidas repetidas em mais de um fator. Além disso, apresentamos estruturas de correlação de trabalho adequadas a medidas repetidas em dois fatores, bem como estimadores para os parâmetros de correlação dessas estruturas, colaborando, dessa forma, para a melhor especificação da estrutura de dependência entre os dados. Desenvolvemos na linguagem R rotinas para o cálculo das estimativas dos modelos propostos.

Como estudos futuros, muitas propostas podem ser exploradas quando os dados possuem medidas repetidas em mais de um fator, especialmente em relação à estrutura de dependência. Outros métodos de estimação dos parâmetros de correlação, como o de *quase-mínimos-quadrados* citados em Wang e Carey (2003), podem ser abordados e a avaliação do impacto desses métodos nas estimativas dos parâmetros de modelagem da média pode ser considerada.

Apêndice A

Conjuntos de dados e rotinas computacionais

Para a consideração da estrutura de correlação em ambos os fatores no processo iterativo responsável pela estimação dos parâmetros da regressão, foi necessário criar rotinas em que utilizamos implicitamente o pacote *gee* do software livre R (<http://www.r-project.org/>).

Não houve necessidade de alterações nos algoritmos de criação dos gráficos de *distância de Cook* e de *resíduos padronizados* pois o método iterativo utilizado para estimar os parâmetros modeladores da média e a matriz de correlação fornece todas as informações. Apenas na rotina responsável pela criação do gráfico de *envelope meio-normal* precisamos utilizar internamente a rotina de estimação do modelo com as sub-estruturas na matriz de correlação de trabalho.

Nesse capítulo apresentamos os conjuntos de dados referentes às aplicações e as rotinas utilizadas nas análises.

A.1 Aplicação 1

Tabela A.1: Conjunto de dados da Aplicação 1.

Observação	Indivíduo	Barra	Lado	Força
1	1	sem	esquerdo	4,33
2	1	sem	direito	3,43
3	1	com	esquerdo	6,67

Observação	Indivíduo	Barra	Lado	Força
4	1	com	direito	4,65
5	2	sem	esquerdo	1,30
6	2	sem	direito	1,60
7	2	com	esquerdo	5,60
8	2	com	direito	7,10
9	3	sem	esquerdo	7,93
10	3	sem	direito	5,10
11	3	com	esquerdo	12,77
12	3	com	direito	11,88
13	4	sem	esquerdo	8,13
14	4	sem	direito	2,77
15	4	com	esquerdo	12,77
16	4	com	direito	11,88
17	5	sem	esquerdo	4,75
18	5	sem	direito	3,48
19	5	com	esquerdo	12,03
20	5	com	direito	10,72
21	6	sem	esquerdo	6,28
22	6	sem	direito	6,58
23	6	com	esquerdo	10,72
24	6	com	direito	10,33
25	7	sem	esquerdo	2,07
26	7	sem	direito	3,97
27	7	com	esquerdo	8,08
28	7	com	direito	9,57
29	8	sem	esquerdo	8,40
30	8	sem	direito	12,07
31	8	com	esquerdo	23,08
32	8	com	direito	20,17

Observação	Indivíduo	Barra	Lado	Força
33	9	sem	esquerdo	2,12
34	9	sem	direito	1,78
35	9	com	esquerdo	6,63
36	9	com	direito	4,65
37	10	sem	esquerdo	2,92
38	10	sem	direito	4,58
39	10	com	esquerdo	12,77
40	10	com	direito	11,88
41	11	sem	esquerdo	6,63
42	11	sem	direito	4,65
43	11	com	esquerdo	7,58
44	11	com	direito	5,80
45	12	sem	esquerdo	8,22
46	12	sem	direito	5,03
47	12	com	esquerdo	12,77
48	12	com	direito	11,88
49	13	sem	esquerdo	4,33
50	13	sem	direito	3,43
51	13	com	esquerdo	6,63
52	13	com	direito	4,65
53	14	sem	esquerdo	5,85
54	14	sem	direito	2,53
55	14	com	esquerdo	10,20
56	14	com	direito	4,50
57	15	sem	esquerdo	2,13
58	15	sem	direito	1,67
59	15	com	esquerdo	2,82
60	15	com	direito	1,87
61	16	sem	esquerdo	6,85

Observação	Indivíduo	Barra	Lado	Força
62	16	sem	direito	3,38
63	16	com	esquerdo	8,08
64	16	com	direito	9,57
65	17	sem	esquerdo	4,33
66	17	sem	direito	3,43
67	17	com	esquerdo	7,22
68	17	com	direito	3,05
69	18	sem	esquerdo	4,52
70	18	sem	direito	4,30
71	18	com	esquerdo	6,00
72	18	com	direito	9,63
73	19	sem	esquerdo	4,57
74	19	sem	direito	8,50
75	19	com	esquerdo	15,70
76	19	com	direito	14,03
77	20	sem	esquerdo	8,05
78	20	sem	direito	5,67
79	20	com	esquerdo	9,23
80	20	com	direito	4,28
81	21	sem	esquerdo	5,10
82	21	sem	direito	4,30
83	21	com	esquerdo	7,58
84	21	com	direito	6,53
85	22	sem	esquerdo	4,80
86	22	sem	direito	7,42
87	22	com	esquerdo	7,22
88	22	com	direito	3,05
89	23	sem	esquerdo	4,78
90	23	sem	direito	6,85

Observação	Indivíduo	Barra	Lado	Força
91	23	com	esquerdo	8,08
92	23	com	direito	9,57

A.1.1 Rotina para estimação dos parâmetros de correlação, de dispersão e da média com matriz de correlação de trabalho com sub-estruturas padrão uniforme em ambos os fatores:

```
library(MASS)
# Estimação de beta sob independência
fit.inic <- glm(y ~ X + (-1), family=Gamma(link="log"))#(link = linkmu))

beta <- fit.inic$coefficients
p <- ncol(X)
n <- length(y)/(s*t)
N <- length(y)

# Cálculo dos resíduos e de phi
mu <- fit.inic$fitted.values
if (opFamily=="poisson")          vmu <- as.vector(mu)
if (opFamily=="Gamma")            vmu <- as.vector(mu^2)
if (opFamily=="gaussian")         vmu <- as.vector(mu^0)
if (opFamily=="binomial")         vmu <- as.vector(mu*(1-mu))
if (opFamily=="inverse.gaussian") vmu <- as.vector(mu^3)

res <- as.vector((fit.inic$residuals)*(1/sqrt(vmu)))
phi <- (res%*%res)/(n*s*t - p)
if(opFamily == "poisson") phi <- summary(fit.inic)$dispersion
```

```
# Estimação das correlações
if (optionR=="excfac2") {
  rr <- as.vector(rep(0,3))
  i <- 0
  a <- 0
  while (i < N-1) {
    for (j in (i+1):(i+t-1)){
      for(k in (j+1):(i+t)) {
        rr[2] <- rr[2]+res[j]*res[k]
      }
    }
    i <- i+t
  }
  i <- 0
  while (i < N-t*(s-1)) {
    j <- i+1
    while (j < (i+(s-1)*t+1)){
      k <- j+t
      while(k < i+s*t+1) {
        rr[1] <- rr[1]+res[j]*res[k]
        k <- k+t
      }
      j <- j+1
    }
    i <- i+t*s
  }
  i <- 0
  while (i < N) {
    for (s1 in 1:(s-1)) {
```

```

    for (j in (i+1+t*(s1-1)):(i+t*s1)) {
      for (k in (i+t*s1+1):(i+t*s)) {
        rr[3] <- rr[3]+res[j]*res[k];a=a+1
        #cat(j,k,i,"\n")
      }
    }
  }
  i <- i+t*s
}
rr[3] <- rr[3]-rr[1]
# Estimativas iniciais das correlações
rho <- phi*rr/c(N*(t-1)/2-p,N*(s-1)/2-p,N*(s-1)*(s-1)/2-p)

# Estimativa inicial da matriz de correlação de trabalho
R1 <- diag(alpha[1],t,t) + (matrix(1,t,t)-diag(t))*alpha[3]
R2 <- (matrix(1,t,t)-diag(t))*alpha[2] # dentro do fator 1
R3 <- kronecker(diag(s), R2)
R4 <- kronecker(matrix(1,s,s)-diag(s), R1)
Rn <- diag(s*t) + R3 + R4
}

# Início do processo iterativo para estimar beta
library(gee)
ite <- 1
dif <- 1

while (dif>0.00001 && ite<50){
fitn = gee(y ~ X + (-1), id = individuos, corstr="fixed",

```

```

family=Gamma(link="log"), R = Rn, b = beta, maxiter = 1)

betan <- fitn$coefficients

# Cálculo dos resíduos e phi
mu <- fitn$fitted.values          # eta
if (opFamily=="poisson")         vmu <- as.vector(mu)
if (opFamily=="Gamma")          vmu <- as.vector(mu^2)
if (opFamily=="gaussian")       vmu <- as.vector(mu^2)
if (opFamily=="binomial")       vmu <- as.vector(mu*(1-mu))
if (opFamily=="inverse.gaussian") vmu <- as.vector(mu^3)

res <- (fitn$residuals)*(1/sqrt(vmu))
phi <- summary(fitn)$scale

# Estimativa das correlações
{
rr <- as.vector(rep(0,3))
i <- 0
while (i < N-1) {
  for (j in (i+1):(i+t-1)){
    for(k in (j+1):(i+t)) {
      rr[2] <- rr[2]+res[j]*res[k]
    }
  }
  i <- i+t
}
i <- 0
while (i < N-t*(s-1)) {

```

```

j <- i+1
while (j < (i+(s-1)*t+1)){
  k <- j+t
  while(k < i+s*t+1) {
    rr[1] <- rr[1]+res[j]*res[k]
    k <- k+t
  }
  j <- j+1
}
i <- i+t*s
}
i <- 0
while (i < N) {
  for (s1 in 1:(s-1)) {
    for (j in (i+1+t*(s1-1)):(i+t*s1)) {
      for (k in (i+t*s1+1):(i+t*s)) {
        rr[3] <- rr[3]+res[j]*res[k]
      }
    }
  }
  i <- i+t*s
}
rr[3] <- rr[3]-rr[1]
rho <- phi*rr/c(N*(t-1)/2-p,N*(s-1)/2-p,N*(s-1)*(s-1)/2-p)

# Estimativa da matriz de correlação de trabalho na iteração
R1 <- diag(rho[1],t,t) + (matrix(1,t,t)-diag(t))*rho[3]
R2 <- (matrix(1,t,t)-diag(t))*rho[2]
R3 <- kronecker(diag(s), R2)
R4 <- kronecker(matrix(1,s,s)-diag(s), R1)

```



```

Rho <- diag(s*t) + R3 + R4
}

#Atualizations
dif <- abs(sum((betan-beta)))
beta <- betan
ite <- ite+1
Rn <- Rho
}
summary(fitn)

```

A.2 Aplicação 2

Tabela A.2: Conjunto de dados da Aplicação 2.

Observação	Indivíduos	B	C	V. Resposta
1	1	b1	c1	25
2	1	b1	c2	43
3	1	b1	c3	50
4	1	b2	c1	30
5	1	b2	c2	42
6	1	b2	c3	47
7	1	b3	c1	18
8	1	b3	c2	27
9	1	b3	c3	36
10	2	b1	c1	15
11	2	b1	c2	31
12	2	b1	c3	40
13	2	b2	c1	20
14	2	b2	c2	27
15	2	b2	c3	37

Observação	Indivíduos	B	C	V. Resposta
16	2	b3	c1	15
17	2	b3	c2	22
18	2	b3	c3	31
19	3	b1	c1	50
20	3	b1	c2	55
21	3	b1	c3	65
22	3	b2	c1	48
23	3	b2	c2	44
24	3	b2	c3	60
25	3	b3	c1	30
26	3	b3	c2	37
27	3	b3	c3	40
28	4	b1	c1	40
29	4	b1	c2	38
30	4	b1	c3	51
31	4	b2	c1	15
32	4	b2	c2	24
33	4	b2	c3	41
34	4	b3	c1	16
35	4	b3	c2	13
36	4	b3	c3	25
37	5	b1	c1	32
38	5	b1	c2	35
39	5	b1	c3	45
40	5	b2	c1	20
41	5	b2	c2	27
42	5	b2	c3	33
43	5	b3	c1	12
44	5	b3	c2	17

Observação	Indivíduos	B	C	V. Resposta
45	5	b3	c3	27
46	6	b1	c1	46
47	6	b1	c2	50
48	6	b1	c3	67
49	6	b2	c1	30
50	6	b2	c2	29
51	6	b2	c3	47
52	6	b3	c1	21
53	6	b3	c2	19
54	6	b3	c3	36

A.2.1 Rotina para estimação dos parâmetros de correlação, de dispersão e de média com matriz de correlação de trabalho com sub-estruturas AR-1 em ambos os fatores:

```
library(MASS)

# Estimação de beta sob independência

fit.inic <- glm(y ~ X + (-1), family=Gamma(link="log"))
beta <- fit.inic$coefficients
p <- ncol(X)
n <- length(y)/(s*t)
N <- length(y)

# Cálculo de resíduos e estimativa inicial de phi

mu <- fit.inic$fitted.values
```

```

if (opFamily=="poisson")          vmu <- as.vector(mu)
if (opFamily=="Gamma")           vmu <- as.vector(mu^2)
if (opFamily=="gaussian")        vmu <- as.vector(mu^0)
if (opFamily=="binomial")        vmu <- as.vector(mu*(1-mu))

res <- as.vector((fit.inic$residuals)*(1/sqrt(vmu)))
phi <- summary(fit.inic)$dispersion
cat("phi estimate under glm:", phi^(-1), "\n")

# Estimativas iniciais de rho

if (optionR=="ar1fac2") {
  rr <- as.vector(rep(0,2))

  i <- 0
  while (i < N-1) {
    for (j in (i+1):(i+t-1)){
      rr[2] <- rr[2]+res[j]*res[j+1]}
    i <- i+t}

  i <- 0; a <- 0
  while (i < N-t*(s-1)) {
    for(j in (i+1):(i+t*(s-1))){
      rr[1] <- rr[1]+res[j]*res[j+t]; a=a+1}
    i <- i+t*s}

  alpha <- rr*phi/c(n*s*(t-1),n*t*(s-1))

# Estimativa inicial da matriz de correlação de trabalho

```

```

S1 <- matrix(0,t,t)
  for(i in 1:t){
    for(j in 1:t){S1[i,j]=abs(i-j)}}
R1 <- matrix(1,t,t)*alpha[1]
R2 <- alpha[2]^S1
Rn <- kronecker(matrix(1,s,s),R2)*kronecker(R1^S1,matrix(1,s,s))
}

# Processo iterativo para estimação de beta

library(gee)
ite <- 1
dif <- 1

while (dif>0.00001 && ite<5) {
  fitn = gee(y ~ X + (-1), id = individuos, corstr="fixed",
            family=Gamma(link="log"), R = Rn, b = beta, maxiter = 1)
  betan <- fitn$coefficients

  mu <- fitn$fitted.values
  if (opFamily=="poisson")          vmu <- as.vector(mu)
  if (opFamily=="Gamma")           vmu <- as.vector(mu^2)
  if (opFamily=="gaussian")        vmu <- as.vector(mu^2)
  if (opFamily=="binomial")        vmu <- as.vector(mu*(1-mu))

  res <- (fitn$residuals)*(1/sqrt(vmu))
  phi <- summary(fitn)$scale
{
  rr <- as.vector(rep(0,2))
  i <- 0

```

```

while (i < N-1) {
  for (j in (i+1):(i+t-1)){
    rr[2] <- rr[2]+res[j]*res[j+1]}
  i <- i+t}
i <- 0
while (i < N-t*(s-1)) {
  for(j in (i+1):(i+t*(s-1))){
    rr[1] <- rr[1]+res[j]*res[j+t]}
  i <- i+t*s}
alpha <- rr/phi/c(n*s*(t-1),n*t*(s-1))

S1 <- matrix(0,t,t)
  for(i in 1:t){
    for(j in 1:t){S1[i,j]=abs(i-j)}}

R1 <- matrix(1,t,t)*alpha[1]
R2 <- alpha[2]^S1
Rho <- kronecker(matrix(1,s,s),R2)*kronecker(R1^S1,matrix(1,s,s))
}

#Atualizações

dif <- abs(sum((betan-beta)))
beta <- betan
ite <- ite+1
Rn <- Rho
}

summary(fitn)

```

A.3 Aplicação 3

Tabela A.3: Conjunto de dados da Aplicação 3.

Observação	Individuo	Canal de frequência	Período	Escore
1	1	c1	10	45
2	1	c1	20	40
3	1	c1	30	28
4	1	c2	10	53
5	1	c2	20	52
6	1	c2	30	37
7	1	c3	10	60
8	1	c3	20	57
9	1	c3	30	46
10	2	c1	10	35
11	2	c1	20	30
12	2	c1	30	25
13	2	c2	10	41
14	2	c2	20	37
15	2	c2	30	32
16	2	c3	10	50
17	2	c3	20	47
18	2	c3	30	41
19	3	c1	10	60
20	3	c1	20	58
21	3	c1	30	40
22	3	c2	10	65
23	3	c2	20	54
24	3	c2	30	47
25	3	c3	10	75

Observação	Individuo	Canal de frequência	Período	Escore
26	3	c3	20	70
27	3	c3	30	50
28	4	c1	10	50
29	4	c1	20	25
30	4	c1	30	16
31	4	c2	10	48
32	4	c2	20	34
33	4	c2	30	23
34	4	c3	10	61
35	4	c3	20	51
36	4	c3	30	35
37	5	c1	10	42
38	5	c1	20	30
39	5	c1	30	22
40	5	c2	10	45
41	5	c2	20	37
42	5	c2	30	37
43	5	c3	10	55
44	5	c3	20	43
45	5	c3	30	37
46	6	c1	10	56
47	6	c1	20	40
48	6	c1	30	31
49	6	c2	10	60
50	6	c2	20	39
51	6	c2	30	29
52	6	c3	10	77
53	6	c3	20	57
54	6	c3	30	46

A.3.1 Rotina para estimação dos parâmetros de correlação, de dispersão e de média com matriz de correlação de trabalho com sub-estruturas mista (AR-1 e padrão uniforme):

```

library(MASS)
# Estimação de beta sob independência
fit.inic <- glm(y ~ X + (-1), family=gaussian(link="identity"))

beta <- fit.inic$coefficients
p <- ncol(X)
n <- length(y)/(s*t)
N <- length(y)

# Cálculo dos resíduos e estimação de phi
mu <- fit.inic$fitted.values # eta
if (opFamily=="poisson") vmu <- as.vector(mu)
if (opFamily=="Gamma") vmu <- as.vector(mu^2)
if (opFamily=="gaussian") vmu <- as.vector(mu^0)
if (opFamily=="binomial") vmu <- as.vector(mu*(1-mu))

# Estimação das correlações
res <- as.vector(fit.inic$residuals*(1/sqrt(vmu)))
phi <- 1/((res%*%res)/(n*s*t - p))
if(opFamily == "poisson") {phi <- summary(fit.inic)$dispersion
if (optionR=="mistfac2") {
rr <- as.vector(rep(0,2))
i <- 0
while (i < N-1) {
for (j in (i+1):(i+t-1)){rr[2] <- rr[2]+res[j]*res[j+1]}
}
}
}

```

```

    i <- i+t
  }
  i <- 0
  while (i < N-t*(s-1)){j <- i+1
    while (j < (i+(s-1)*t+1)){k <- j+t
      while(k < i+s*t+1){
        rr[1] <- rr[1]+res[j]*res[k]
        k <- k+t}
      j <- j+1}
    i <- i+t*s}
# Estimativa inicial das correlações
alpha <- phi*rr/c(N*(s-1)*(s-1)/2-p,n*s*(t-1))

# Estimativa inicial da matriz de correlação de trabalho
S1 <- matrix(0,t,t)
for(i in 1:t){for(j in 1:t){S1[i,j]=abs(i-j)}}
R1 <- matrix(1,s,s)-diag(s)
S <- alpha[2]^S1
R <- alpha[1]^R1
Rn <- kronecker(R,S)
}

# Início do processo de estimação de beta
library(gee)
ite <- 1
dif <- 1

while (dif>0.00001 && ite<50) {
  fitn = gee(y ~ X + (-1), id = cluster, corstr="fixed",
    family=gaussian(link="identity"), R = Rn, b = beta, maxiter = 1)

```

```

betan <- fitn$coefficients
mu <- fitn$fitted.values
if (opFamily=="poisson")          vmu <- as.vector(mu)
if (opFamily=="Gamma")           vmu <- as.vector(mu^2)
if (opFamily=="gaussian")        vmu <- as.vector(mu^0)
if (opFamily=="binomial")        vmu <- as.vector(mu*(1-mu))

res <- as.vector(fit.inic$residuals*(1/sqrt(vmu)))
phi <- 1/((res%*%res)/(n*s*t - p))

# Estimação das correlações na iteração
{
  rr <- as.vector(rep(0,2))
  i <- 0
  while (i < N-1){
    for (j in (i+1):(i+t-1))
      {rr[2]<- rr[2]+res[j]*res[j+1];i <- i+t}}
  i <- 0
  while (i < N-t*(s-1)){j <- i+1
    while (j < (i+(s-1)*t+1)){k <- j+t
      while(k < i+s*t+1){rr[1] <- rr[1]+res[j]*res[k]; k <- k+t}
      j <- j+1}
    i <- i+t*s}
  alpha <- phi*rr/c(N*(s-1)*(s-1)/2-p,n*s*(t-1))
S1 <- matrix(0,t,t)
  for(i in 1:t){for(j in 1:t){S1[i,j]=abs(i-j)}}
R1 <- matrix(1,s,s)-diag(s)
S <- alpha[2]^S1
R <- alpha[1]^R1
Rho <- kronecker(R,S)

```

```
}  
  
# Atualizations  
dif <- abs(sum(betan-beta))  
beta <- betan  
ite <- ite+1  
Rn <- Rho  
}
```


Referências Bibliográficas

- Andreoni, S. (2003). *Modelos de efeitos aleatórios para análise de dados longitudinais não balanceados em relação ao tempo*, Master's thesis, São Paulo: IME - Universidade de São Paulo.
- Artes, R. (1997). *Extensões da Teoria das Equações de Estimação Generalizadas a Dados Circulares e Modelos de Dispersão*, PhD thesis, São Paulo: IME - Universidade de São Paulo.
- Artes, R. e Botter, D. A. (2005). *Funções de Estimação em Modelos de Regressão*, São Paulo: Associação Brasileira de Estatística.
- Artes, R. e Jorgensen, B. (2000). Longitudinal data estimating equations for dispersion models, *Scandinavian Journal of Statistics* **27**: 321–334.
- Chandrasekar, B. e Kale, B. K. (1984). Unbiased statistical estimation functions in presence of nuisance parameter, *Journal of Statistical Planning and Inference* **9**: 45–54.
- Crowder, M. (1987). On linear and quadratic estimating function, *Biometrika* **74**: 591–7.

- Crowder, M. (1995). On the use of a working correlation matrix in using generalised linear models for repeated measures, *Biometrika* **82**: 407–10.
- Fitzmaurice, G. M. (1995). A caveat concerning independence estimating equations with multivariate binary data, *Biometrics* **51**: 309–317.
- Fitzmaurice, G. M., Laird, N. M. e Rotnitzky, A. G. (1993). Regression models for discrete longitudinal responses (with discussion), *Statistical Science* **8**: 284–309.
- Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation, *Annals of Mathematical Statistics* **31**: 1208–1211.
- Godambe, V. P. (1991). *Estimating Function (Oxford Statistical Science Series - 7)*, Oxford: Oxford University Press.
- Huajun, Y. E. e Jianxin, P. A. N. (2006). Modelling of covariance structures in generalised estimating equations for longitudinal data, *Biometrika* **93**(4): 927–941.
- Laird, N. M. e Ware, J. H. (1982). Random-effects models for longitudinal data, *Biometrics* **38**: 963–74.
- Liang, K.-Y. e Zeger, S. L. (1986). Longitudinal analysis using generalized linear model, *Biometrika* **73**: 13–22.
- Nelder, J. A. e Wedderburn, R. W. M. (1972). Generalized linear models, *Journal of the Royal Statistical Society A* **135**: 370–384.
- Neter, J., Kutner, M. H., Nachtsheim, C. e Wasserman, W. (1996). *Applied Linear Statistical Models*, Chicago: McGraw-Hill. 4^a ed.7.

- Paula, G. (2004). *Modelos de Regressão com Apoio Computacional*, São Paulo: IME - Universidade de São Paulo.
- Prentice, R. (1988). Correlated binary regression with covariates specific to each binary observation, *Biometrics* **44**: 1033–48.
- Prentice, R. e Zhao, C. (1991). Estimating equations for parameters in means and covariances of multivariate discrete and continuous responses, *Biometrics* **47**: 825–39.
- Rodrigues, R. S. (2007). *Estudo comparativo da intensidade da força de mordida de próteses totais convencionais e overdentures inferiores retidas por sistema barra/clip sobre implantes*, PhD thesis, São Paulo: Faculdade de Odontologia - Universidade de São Paulo.
- Rotnitzky, A. e Jewell, N. (1990). Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data, *Biometrika* **77**: 485–97.
- Singer, J. M. e Andrade, D. F. (1986). *Análise de dados Longitudinais*, IV Simpósio Nacional de Probabilidade e Estatística.
- Venezuela, M. K., Botter, D. A. e Sandoval, M. C. (2007). Diagnostic techniques in generalized estimating equation, *Journal of Statistical Computation and Simulation* **77**(10): 879–888.
- Wang, Y.-G. e Carey, V. (2003). Working correlation structure misspecification, estimation and covariate desing: Implications for generalised estimating equations performance, *Biometrika* **90**: 29–41.

- Wang, Y.-G. e Carey, V. (2004). Unbiased estimating equations from working correlation models for irregularly timed repeated measures, *Journal of the American Statistical Association* **99**: 845–853.
- Winer, B. J., Brown, D. R. e Michels, K. (1991). *Statistical principles in experimental design*, 3 edn, McGraw-Hill.
- Zeger, S. L., Liang, K.-Y. e Qaqish, B. (1992). Multivariate regression analysis for categorical data, *Journal of the Royal Statistical Society* **54**: 3–40. (com discussões).