

**Identificação de modelos lineares
mistos gaussianos**

Jairo Arturo Angel Guzmán

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

Programa: Estatística

Orientador: Prof. Dr. Francisco Marcelo Monteiro da Rocha

Coorientador: Prof. Dr. Julio da Motta Singer

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

São Paulo, Maio de 2021

Identificação de modelos lineares mistos gaussianos

Esta versão da tese contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho realizada em 01/04/2021, disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Francisco Marcelo Monteiro Da Rocha - UNIFESP - IME - Presidente
- Prof. Dr. Gilberto Alvarenga Paula - IME-USP
- Prof. Dr. Mario de Castro Andrade Filho - ICMC -USP
- Profa. Dra. Luzia Aparecida Trinca - UNESP
- Profa. Dra. Hildete Prisco Pinheiro - IMECC - UNIC

Dedicatória

Dedico este trabalho às minhas irmãs, aos meus sobrinho(a)s e cunhadas, às minhas filhas Sofia e Aura e à minha esposa Rosaura Santos.

Agradecimentos

A meu orientador e coorientador, Prof. Dr. Francisco Marcelo Monteiro da Rocha e a meu Prof. Dr. Julio da Motta Singer, pela orientação, apoio, paciência, compreensão e confiança depositada em mim durante a elaboração deste trabalho.

A minha esposa Rosaura Santos Tovar pelo amor e carinho, pelo apoio nos momentos difíceis e pelos momentos vividos durante estes anos.

A minha cunhada Loly María Santos Tovar pelos momentos em que estouvo, enquanto eu não fiquei em Colômbia e ela ajudou na assistência de minhas filhas Sofia e Aura.

Aos meus amigos Sergio, Estefania, Jocelânio e Yuri pela amizade, carinho e ajuda oferecida durante a estadia em São Paulo.

À professora Silvia Nagib Elian por ter contribuído na minha formação acadêmica e pessoal.

Resumo

Jairo Angel Guzmán. **Identificação de modelos lineares mistos gaussianos.** Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2021. O objetivo deste trabalho é avaliar uma possível má especificação de modelos lineares mistos gaussianos. Essa avaliação permite reconhecer quando o modelo é incorretamente especificado e identificar a fonte do erro de especificação, que pode estar na estrutura da média, na estrutura da matriz de covariâncias ou em ambas. Com esse propósito, propomos testes baseados na matriz de informação obtida da função de verossimilhança que, em conjunto com diagnósticos gráficos, são usados na identificação. Um estudo de simulação permite avaliar os testes em quanto ao poder e à taxa do erro tipo I. Concluimos com uma aplicação da estratégia da análise proposta num exemplo com dados reais.

Palavras-chave: má especificação, testes, diagnósticos gráficos.

Abstract

Jairo Angel Guzmán. **Identification of linear gaussian mixed models**. Tese (Doutorado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2021.

The objective of this work is to evaluate a possible misspecification in the gaussian linear mixed models. This evaluation makes it possible to recognize when the model is incorrectly specified and to identify the source of the specification error, which can be in the structure of the mean or in the structure of the covariances matrix of the vector of response, or in both cases. For this object, we propose tests based on the Fisher information matrix obtained from the likelihood function in joint together with graphics diagnostics for identification the model. A simulation study allows evaluating the tests in terms of power and type I error rate. We conclude with an application of the proposal developing examples with real data.

Keywords: misspecification, testes, graphics diagnostics.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiv
1 Introdução	1
2 Modelo linear misto	7
2.1 Inferência	8
2.2 Propriedades assintóticas do estimador de máxima verosimilhança	13
2.3 Detecção de erros de especificação num modelo linear misto gaussiano	15
2.3.1 Métodos analíticos	16
2.3.1.1 Teste Alternativo da Matriz de Informação	16
2.3.1.2 Teste do Estimador “Sanduíche” Modificado	18
2.3.1.3 Teste do Estimador “Sanduíche” Alternativo	21
2.3.2 Métodos gráficos	23
3 Estudo de Simulação	26
3.1 Modelos considerados para analisar erros de especificação em modelos lineares mistos gaussianos	27
3.2 Avaliação de um modelo corretamente especificado	28
3.3 Análise quando há presença de erro de especificação na estrutura da resposta média	33
3.3.1 Método analítico	33
3.3.2 Método gráfico	33
3.4 Análise quando há presença de erro de especificação na estrutura da matriz de covariâncias do vetor de respostas	37

3.4.1	Há presença de erro de especificação no número de efeitos aleatórios	38
3.4.1.1	Método analítico	39
3.4.2	Há presença de erro de especificação na matriz de covariâncias do vetor de efeitos aleatórios, os efeitos aleatórios são considerados independentes	44
3.4.2.1	Método analítico	44
3.4.2.2	Método gráfico	45
3.4.3	Há presença de erro de especificação na estrutura da matriz de covariâncias intraunidades amostrais	49
3.4.3.1	Método analítico	49
3.4.3.2	Método gráfico	50
3.4.4	Há presença de erro de especificação no número de efeitos aleatórios e na matriz de covariâncias intraunidades amostrais	54
3.4.4.1	Método analítico	54
3.4.4.2	Método gráfico	54
4	Aplicação	60
4.0.1	Método gráfico	63
4.0.2	Método gráfico	69
4.0.3	Método gráfico	73
4.1	Aplicação 2	77
5	Conclusão e estudos futuros	111
5.1	Conclusão	111
5.2	Estudos futuros	111
A		112
A.1	Derivadas do logaritmo da função de verossimilhança	112
A.2	Derivadas de primeira ordem	113
A.2.1	Derivada do vetor escore	113
A.2.2	Derivada primeira com relação ao vetor β	113
A.2.3	Derivada primeira com relação aos componentes do vetor θ	113
A.2.4	Derivada primeira com relação a σ^2	114
A.3	Derivadas de segunda ordem	114

A.3.1	Derivada segunda do vetor $\boldsymbol{\beta}^\top$ com relação ao vetor $\boldsymbol{\beta}$	114
A.3.2	Derivada segunda dos componentes do vetor $\boldsymbol{\theta}$ com relação ao vetor $\boldsymbol{\beta}$	114
A.3.3	Derivada segunda de σ^2 com relação ao vetor $\boldsymbol{\beta}$	114
A.3.4	Derivada segunda de cada componente do vetor $\boldsymbol{\theta}$ com relação a σ^2	115
A.3.5	Derivada segunda da função logaritmo da verossimilhança dos componentes do vetor $\boldsymbol{\theta}^\top$	115
A.3.6	Derivada segunda de σ^2 com relação a σ^2	115
A.4	Derivada terceira ordem	116
A.4.1	Derivada terceira com respeito a cada componente do vetor $\boldsymbol{\beta}$	116
A.4.2	Derivada com relação aos elementos do vetor $\boldsymbol{\beta}$, da derivada segunda dos elementos do vetor $\boldsymbol{\theta}$	117
A.4.3	Derivada com relação a β_t da derivada segunda de σ^2	118
A.4.4	Derivada terceira com relação aos elementos do vetor $\boldsymbol{\theta}$	118
A.4.5	Derivada da derivada segunda do logaritmo da função de verossimilhança de σ^2	118
A.4.6	Derivada com relação a σ^2 da derivada segunda do logaritmo da verossimilhança de β_k^2	119
A.4.7	Derivada com relação a σ^2 , da derivada segunda do logaritmo da verossimilhança de σ^2	119
A.5	Apêndice B	
	Expressões necessárias para os testes propostos	119
A.5.1	Matrizes	119
A.5.2	Calculando cada expressão em (A.25)	125
A.6	Analisando a expressão \mathbf{A}_{12} em (A.28)	126
A.6.1	Analisando a expressão \mathbf{A}_{13} em (A.28)	128
A.7	Analisando a expressão \mathbf{A}_{22} em (A.28)	129
A.7.1	Analisando a expressão \mathbf{A}_{22} em (A.28)	131
A.8	Apêndice C	
	Condições de regularidade	143
A.9	Apêndice D	145

A.9.1	Teste Alternativo da Matriz de Informação	145
A.9.2	Teste do Estimador “Sanduíche” Modificado	145
A.9.3	Teste do Estimador “Sanduíche” Alternativo	147
Referências Bibliográficas		148

Lista de Figuras

3.1	Perfis individuais para as 29 unidades amostrais, linha tracejada é o perfil loess, e a linha continua é o perfil médio dos dados simulados.	29
3.2	Gráfico da função de autocorrelação.	30
3.3	Gráfico de índice modificado de Lesaffre - Verbeke. A linha tracejada no gráfico representa 3o quartil + 1.5 distância interquartis.	30
3.4	Gráfico para os resíduos marginais padronizados.	31
3.5	Gráfico dos resíduos padronizados condicionados e histograma.	31
3.6	Gráfico QQ para a distância da Mahalanobis.	32
3.7	QQ plot e histograma para os resíduos minimamente confundidos padronizados.	32
3.8	Gráfico da função de autocorrelação	34
3.9	Gráfico do índice modificado de Lesaffre–Verbeke.	34
3.10	Resíduos marginais padronizados.	35
3.11	QQ plot e histograma para os resíduos minimamente confundidos padronizados.	35
3.12	Gráfico da distância dos quantis de Mahalanobis.	36
3.13	QQ plot e histograma para os resíduos minimamente confundidos padronizados.	36
3.14	Comportamento dos perfis individuais simulados ao redor do perfil médio e do perfil loess.	38
3.15	Gráfico de índice modificado de Lesaffre - Verbeke.	39
3.16	Gráfico dos resíduos marginais padronizados	40
3.17	Gráfico QQ para a distancia de Mahalanobis.	41
3.18	Gráfico da distancia de Mahalanobis.	41
3.19	Gráfico QQ e histograma para os resíduos minimamente confundidos padronizados.	42
3.20	Gráfico dos resíduos condicionados padronizados.	43

3.21	Gráfico da função de autocorrelação.	45
3.22	Gráfico do índice modificado de Lesaffre–Verbeke.	45
3.23	Gráfico dos resíduos marginais padronizados.	46
3.24	Gráfico da distância de Mahalanobis.	46
3.25	Gráfico da distância dos quantis de Mahalanobis.	47
3.26	Gráfico dos resíduos condicionais padronizados.	47
3.27	Gráfico QQ para os resíduos condicionais padronizados e histograma.	48
3.28	Gráfico da função de autocorrelação.	50
3.29	Gráfico do índice modificado de Lesaffre–Verbeke.	51
3.30	Gráfico dos resíduos marginais padronizados	51
3.31	Gráfico da distância dos quantis de Mahalanobis.	52
3.32	Gráfico dos resíduos condicionais padronizados.	52
3.33	QQ plot e histograma para os resíduos minimamente confundidos padronizados.	53
3.34	Gráfico da função de autocorrelação.	55
3.35	Gráfico do índice modificado de Lesaffre–Verbeke.	55
3.36	Gráfico dos resíduos marginais padronizados e histograma.	56
3.37	Gráfico da distância de Mahalanobis.	57
3.38	Gráfico da distância dos quantis de Mahalanobis.	57
3.39	QQ plot e histograma para os resíduos minimamente confundidos padronizados.	58
3.40	Gráfico dos resíduos condicionados padronizados e histograma.	59
4.1	Comportamento dos perfis individuais ao redor do perfil médio e do perfil loess.	61
4.2	Função de autocorrelação.	64
4.3	Gráfico do índice de Lesaffre–Verbeke modificado.	64
4.4	Gráfico QQ de resíduos marginais padronizados e histograma.	65
4.5	Gráfico da distância de Mahalanobis.	65
4.6	Gráfico da distância dos quantis de Mahalanobis.	66
4.7	Gráfico de resíduos condicionais padronizados e histograma.	67
4.8	Gráfico QQ e histograma dos resíduos condicionais minimamente confundidos.	67
4.9	Gráfico da função de autocorrelação.	69
4.10	Gráfico de índice de Lesaffre–Verbeke modificado.	69
4.11	Gráfico dos resíduos marginais padronizados e histograma.	70

4.12	Gráfico da distância de Mahalanobis.	70
4.13	Gráfico da distância dos quantis de Mahalanobis.	71
4.14	Gráfico dos resíduos condicionais padronizados.	71
4.15	Gráfico QQ e histograma para os resíduos condicionais minimamente confundidos.	72
4.16	Gráfico da função de autocorrelação.	74
4.17	Gráfico do índice de Lesaffre–Verbeke modificado.	74
4.18	Gráfico dos resíduos marginais padronizados e histograma.	75
4.19	Gráfico da distância de Mahalanobis.	75
4.20	Gráfico da distância dos quantis de Mahalanobis.	76
4.21	Gráfico dos resíduos condicionais padronizados.	76
4.22	Gráfico QQ e histograma para resíduos condicionais minimamente confundidos.	77
4.23	Perfis individuais ao redor do perfil médio e do perfil loess da concentração de bilirrubina.	80
4.24	Gráfico da função de autocorrelação.	83
4.25	Gráfico de índice modificado de Lesaffre - Verbeke.	83
4.26	Gráfico dos resíduos marginais padronizados.	84
4.27	Gráfico da distância dos quantis de Mahalanobis.	84
4.28	Gráfico da distância de Mahalanobis.	85
4.29	Gráfico dos resíduos condicionais padronizados.	85
4.30	Gráfico QQ para os resíduos condicionais padronizados e histograma.	86
4.31	Gráfico da função de autocorrelação.	88
4.32	Gráfico de índice modificado de Lesaffre - Verbeke.	88
4.33	Gráfico dos resíduos marginais padronizados.	89
4.34	Gráfico da distância dos quantis de Mahalanobis.	89
4.35	Gráfico da distância de Mahalanobis.	90
4.36	Gráfico dos resíduos condicionais padronizados.	90
4.37	Gráfico QQ para os resíduos condicionais padronizados e histograma.	91
4.38	Gráfico da função de autocorrelação.	93
4.39	Gráfico de índice modificado de Lesaffre - Verbeke.	93
4.40	Gráfico dos resíduos marginais padronizados.	94

4.41	Gráfico da distância dos quantis de Mahalanobis.	94
4.42	Gráfico da distância de Mahalanobis.	95
4.43	Gráfico dos resíduos condicionais padronizados.	95
4.44	Gráfico QQ para os resíduos condicionais padronizados e histograma.	96
4.45	Gráfico da função de autocorrelação.	98
4.46	Gráfico de índice modificado de Lesaffre - Verbeke.	98
4.47	Gráfico dos resíduos marginais padronizados.	99
4.48	Gráfico da distância dos quantis de Mahalanobis.	99
4.49	Gráfico da distância de Mahalanobis.	100
4.50	Gráfico dos resíduos condicionais padronizados.	100
4.51	Gráfico QQ para os resíduos condicionais padronizados e histograma.	101
4.52	Gráfico da função de autocorrelação.	102
4.53	Gráfico de índice modificado de Lesaffre - Verbeke.	103
4.54	Gráfico dos resíduos marginais padronizados.	103
4.55	Gráfico da distância dos quantis de Mahalanobis.	104
4.56	Gráfico da distância de Mahalanobis.	104
4.57	Gráfico dos resíduos condicionais padronizados.	105
4.58	Gráfico QQ para os resíduos condicionais padronizados e histograma.	105
4.59	Gráfico da função de autocorrelação.	107
4.60	Gráfico de índice modificado de Lesaffre - Verbeke.	107
4.61	Gráfico dos resíduos marginais padronizados.	108
4.62	Gráfico da distância dos quantis de Mahalanobis.	108
4.63	Gráfico da distância de Mahalanobis.	109
4.64	Gráfico dos resíduos condicionais padronizados.	109
4.65	Gráfico QQ para os resíduos condicionais padronizados e histograma.	110

Lista de Tabelas

1.1	Volume(ml) do ventrículo esquerdo em 29 crianças pré-termo, [Afiune (2000)].	2
2.1	Técnica de diagnóstico gráfico para detecção de fonte de erros de especificação em um modelo do tipo (2.1).	25
3.1	Modelos considerados para o processo de simulação.	28
3.2	Valores das estatísticas dos testes e valores p.	29
3.3	Estatísticas de cada teste e <i>valor-p</i>	33
3.4	Poder e taxa do erro tipo I (<i>E.tipoI</i>) para os testes: alternativo da matriz de informação (<i>I</i>), do estimador “sanduíche” modificado (<i>II</i>) e do estimador “sanduíche” alternativo (<i>III</i>).	37
3.5	Valores das estatísticas dos testes e valores p.	39
3.6	Poder e taxa do erro tipo I (<i>E.tipoI</i>) para os testes: da matriz de informação alternativo (<i>I</i>), do estimador “sanduíche” modificado (<i>II</i>) e do estimador “sanduíche” alternativo (<i>III</i>).	43
3.7	Estatística de cada teste, <i>valor-p</i> com nível de significância 5%.	44
3.8	Poder e taxa do erro tipo I (<i>E.tipoI</i>) para os testes: da matriz de informação alternativo (<i>I</i>), do estimador “sanduíche” modificado (<i>II</i>) e do estimador “sanduíche” alternativo (<i>III</i>).	48
3.9	Estatísticas de cada teste e <i>valor-p</i> , 5% significância.	49
3.10	Poder e taxa do erro tipo I (<i>E.tipoI</i>) para os testes: da matriz de informação alternativo (<i>I</i>), do estimador “sanduíche” modificado (<i>II</i>) e do estimador “sanduíche” alternativo (<i>III</i>).	53
3.11	Valores das estatísticas dos testes e valores p.	54

3.12 Poder e taxa do erro tipo I ($E.tipoI$) para os testes: da matriz de informação alternativo (I), do estimador “sanduíche” modificado (II) e do estimador “sanduíche” alternativo (III).	59
4.1 Resultados do modelo ajustado	63
4.2 Estatísticas de cada teste e $valor - p$	63
4.3 Resultados do ajuste do modelo alternativo	68
4.4 Estatística de cada teste e $valor - p$	68
4.5 Resultados do ajuste do modelo final	73
4.6 Estatísticas dos testes e $valor - p$	73
4.7 Concentração de bilirrubina ($\mu mol/L$) em recém - nascidos (RN) saudáveis em aleitamento materno.	78
4.8 Concentração de bilirrubina ($\mu mol/L$) em recém - nascidos (RN) saudáveis em aleitamento materno.	79
4.9 Resultados do modelo ajustado	82
4.10 Estatísticas de cada teste e $valor - p$	82
4.11 Resultados do modelo ajustado	87
4.12 Estatísticas de cada teste e $valor - p$	87
4.13 Resultados do modelo ajustado	92
4.14 Estatísticas de cada teste e $valor - p$	92
4.15 Resultados do modelo ajustado	97
4.16 Estatísticas de cada teste e $valor - p$	97
4.17 Resultados do modelo ajustado	101
4.18 Estatísticas de cada teste e $valor - p$	102
4.19 Resultados do modelo ajustado	106
4.20 Estatísticas de cada teste e $valor - p$	106

Capítulo 1

Introdução

Estudos longitudinais (ou de painel) têm como característica a observação de uma ou mais variáveis respostas sequencialmente ao longo do tempo na mesma unidade amostral. Na Tabela 1.1 são apresentados dados de um estudo longitudinal realizado na Faculdade de Medicina da Universidade de São Paulo, Brasil, que envolve a observação do volume (ml) do ventrículo esquerdo, num grupo de 29 crianças consideradas prematuras (pré-termo) e observadas ao longo de 15 semanas pós-concepção. Essas medidas foram tomadas usando métodos ecocardiográficos. Estudos longitudinais permitem analisar os perfis de resposta média e individual, assim como salientar possíveis mudanças populacionais e individuais. Estudos longitudinais constituem um caso especial de estudos de medidas repetidas. Detalhes sobre estudos longitudinais podem ser encontrados em [Diggle et al. \(2002\)](#), [Molenberghs & Verbeke \(2005\)](#) e [Hedeker & Gibbson \(2006\)](#).

A presença de dados omissos é comum em conjuntos de dados provenientes de estudos longitudinais. Os dados na Tabela 1.1 mostram uma estrutura irregular e desbalanceada em relação ao tempo. Os dados omissos são classificados de uma maneira geral como informativos ou não informativos. Neste trabalho consideramos apenas dados omissos não informativos, ou seja, em que a probabilidade de omissão depende somente dos valores observados, [[Little & Rubin \(1987\)](#)]. A análise de dados longitudinais é mais complexa quando há presença de dados omissos. Nesse caso, é mais difícil especificar um modelo linear misto para representar a variação da resposta média ao longo do tempo e a possível estrutura de covariância de cada unidade amostral.

Tabela 1.1: *Volume(ml) do ventrículo esquerdo em 29 crianças pré-termo, [Afune (2000)].*

ID	Semana pós-concepção															
	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	
1	1.69	0.94	1.03	1.52	0.91	0.91	1.60	2.74		2.57		2.74			2.92	
2					3.38	2.05	3.65		4.25	4.83		6.13				
3					1.33	1.60	1.60	2.41	3.65	4.83						
4				1.37	1.16		0.86		1.60		2.35	3.38	3.31		5.09	
5							1.60	2.52	2.10	2.63						
6						2.57		3.11	5.18		6.54					
7			1.95	2.10	1.86		2.05		2.99		2.41		2.99			
8			4.83	1.26	1.48			3.11		4.10						
9									4.25	2.92		3.44		4.25	7.88	
10							2.25	2.05	2.05		4.25					
11					1.26		2.80	2.92	4.41	5.18		8.24				
12			1.52		1.33	0.91										
13			3.44		1.44											
14							2.74	2.74		3.38	4.10	4.33		4.25		
15								1.60		2.20	2.35		5.18			
16						2.05	1.64	1.95	3.18		3.18	4.10				
17						2.10	2.41	2.63		4.25						
18				2.10	2.05	2.80			4.02							
19				2.92	1.37			1.33	2.10		1.95			3.31		
20							1.16		2.35	1.91	2.74		3.94			
21					2.41	2.41	1.73		2.74		5.36					
22			1.91	1.77	1.23	1.26	1.12	1.52								
23								1.77	2.63		3.65					
24					1.26	1.52	1.95		4.33							
25					2.99		2.92		5.18	5.93						
26					1.33	1.03	1.60		2.20		2.99					
27								2.35	2.35	2.05	2.57		3.38			
28							1.64	2.86		2.41			2.74		4.10	
29					1.77	2.35		2.35			2.74		3.51			

A análise de conjuntos de dados oriundos de estudos longitudinais pode ser realizada por meio de modelos lineares mistos. Duas fontes de variabilidade estão presentes neste tipo de modelos: a primeira permite modelar a variação entre os perfis individuais da variável resposta e a segunda está relacionada com a dispersão da resposta em torno dos perfis individuais. A correspondente matriz de covariâncias intraunidades amostrais num modelo linear misto pode ser simples ou complexa. A simplicidade reduz o número de parâmetros e os problemas de identificabilidade. Detalhes sobre modelos mistos podem ser encontrados em Diggle et al. (2002), Molenberghs & Verbeke (2005), Wang (2013) e Demidenko (2013).

A vantagem de analisar dados longitudinais considerando modelos lineares mistos é que eles incorporam características tais como a variação entreunidades e intraunidades amostrais, dados omissos e covariáveis que dependem do tempo. Quando um modelo linear misto é proposto para um conjunto de dados, o mecanismo gerador desses dados não é conhecido. Nesse caso, modelam-se os dados por meio de uma família paramétrica de densidades de probabilidade que pode ou não conter a verdadeira densidade dos dados. Se a densidade

de probabilidade dos dados pertence à família paramétrica adotada, então o modelo está corretamente especificado; em caso contrário, o modelo apresenta erro de especificação e, conseqüentemente, as propriedades de eficiência e consistência do estimador de máxima verossimilhança do vetor de parâmetros do modelo podem não valer.

Erros de especificação num modelo linear misto gaussiano podem ocorrer devido a uma incorreta especificação na forma funcional da resposta média ou a uma incorreta especificação na estrutura da matriz de covariâncias do vetor de respostas. Neste trabalho foram consideradas situações em que há incorreções na especificação da:

- (i) estrutura da média, há uma omissão de uma variável explicativa na forma funcional da resposta média,
- (ii) estrutura na matriz de covariâncias do vetor de efeitos aleatórios, o erro está no número de elementos no vetor de efeitos aleatórios,
- (iii) estrutura na matriz de covariâncias do vetor de efeitos aleatórios, os erros aleatórios são considerados independentes,
- (iv) especificação da matriz de covariâncias do vetor de erros aleatórios, concretamente os erros aleatórios seguem um processo autorregressivo de ordem 1, $AR(1)$,
- (v) estrutura no número de elementos no vetor de efeitos aleatórios e os erros aleatórios seguem um processo autorregressivo de ordem 1, $AR(1)$, conjuntamente.

Alguns autores utilizaram testes estatísticos para avaliar erros de especificação de modelos lineares. Por exemplo, [Hausman \(1978\)](#) propôs um teste para detectar erros de especificação na esperança condicional do vetor de erros aleatórios dado o vetor de variáveis explicativas. [White \(1982\)](#) propôs um teste denominado Teste da Matriz de Informação (*TMI*), baseado na diferença entre a matriz formada pelos elementos que estão na diagonal principal e acima da diagonal principal da matriz de informação de Fisher e pelos elementos na mesma posição na matriz de covariâncias do vetor escore. Um resultado diferente de zero, nesta diferença, é indicativo que o modelo linear apresenta erro de especificação. O autor, sugere não trabalhar com todos esses elementos nas matrizes descritas porque isso aumenta o número de parâmetros, os graus de liberdade e conseqüentemente o poder do teste. [Hall \(1987\)](#) decompôs

o *TMI* proposto por [White \(1982\)](#) na soma de três formas quadráticas independentes num modelo linear autorregressivo médias móveis (ARMA). A primeira forma considera um teste usando multiplicadores de Lagrange, o qual permite analisar heterocedasticidade. As outras duas formas, são expressas em termos do terceiro e do quarto momento dos resíduos do modelo, o que permite mostrar que o teste falha para detectar correlação e que não é assintoticamente ótimo para avaliar heterocedasticidade, assimetria e curtose conjuntamente. [Furno \(1996\)](#) mostrou que o teste proposto por [White \(1982\)](#) pode ser decomposto na soma de formas quadráticas para o caso de um modelo linear ARMA, a decomposição facilita a detecção de não linearidade na parte fixa e na estrutura do modelo. [Inoue et al. \(2014\)](#) propuseram um método que permite detectar e identificar erros de especificação na estrutura da matriz de covariâncias do vetor de erros aleatórios num modelo linear autorregressivo de ordem 1 (AR(1)), fazendo a análise da decomposição da matriz de covariâncias dos erros de previsão e fazendo a análise da função de verossimilhança marginal de modelos dinâmicos estocásticos.

Em todos os casos anteriores, só há presença de uma fonte de variação, devida ao erro aleatório no modelo linear. No caso dos modelos lineares mistos, a presença de duas fontes de variabilidade, requer mais atenção do que os modelos lineares. Para a detecção de erros de especificação de modelos lineares mistos, [Jiang \(2001\)](#) propôs um teste que permite avaliar erros de especificação na distribuição do vetor de efeitos aleatórios. [Huang \(2011\)](#) apresenta um teste que permite identificar erros de especificação na distribuição do vetor de efeitos aleatórios.

Em modelos lineares mistos, técnicas exploratórias permitem identificar a fonte dos erros. Por exemplo, [Lesaffre & Verbeke \(1998\)](#) sugeriram técnicas exploratórias para identificar observações discrepantes (*outliers*) que influenciam localmente nestes tipos de modelos. [Rocha \(2004\)](#) propôs um conjunto de técnicas gráficas e analíticas para o diagnóstico da matriz de covariâncias intraunidades amostrais em estudos com medidas repetidas. Essas técnicas abrangem a análise da matriz de covariâncias e correlações amostrais, dos gráficos de dispersão múltipla, do gráfico de perfil médio e de perfis individuais, dos gráficos de perfis das linhas das matrizes de correlações amostrais em função da defasagem e do variograma amostral das observações padronizadas. [Verbeke & Molenberghs \(2013\)](#) usaram a função gradiente como

uma ferramenta de diagnóstico para identificar erros de especificação na distribuição do vetor de efeitos aleatórios num modelo linear misto. [Drikvandi et al. \(2017\)](#) usaram a função gradiente com o fim de avaliar erros de especificação na distribuição do vetor de efeitos aleatórios. [Singer et al. \(2017\)](#) propuseram um conjunto de ferramentas gráficas para analisar violações nas suposições em modelos lineares mistos. Essas ferramentas são baseadas em análises de três tipos de resíduos: marginais, condicionais e de efeitos aleatórios. [Rocha & Singer \(2018\)](#) recomendaram métodos exploratórios baseados na análise dos perfis individuais e médio para identificar o número de efeitos fixos e aleatórios. Esses autores também aconselham a análise das linhas da matriz de covariâncias amostral para identificar o número de efeitos aleatórios no mesmo tipo de modelo.

Outros autores propuseram testes para detecção de erros de especificação em modelos lineares generalizados mistos. [Huang \(2009\)](#) propôs um teste que permite avaliar erros de especificação na distribuição do vetor de efeitos aleatórios neste tipo de modelos. [Abad et al. \(2010\)](#) propuseram dois testes com o objetivo de detectar erros de especificação na distribuição do vetor de efeitos aleatórios num modelo linear generalizado misto. Ambos os testes são modificações do teste apresentado por [White \(1982\)](#). O Teste da Matriz de Informação Modificada (*TMIM*) é baseado em uma estatística formada pela soma dos elementos da diagonal principal da matriz de informação esperada, e pelos elementos da diagonal principal da matriz do produto cruzado do vetor escore.

O segundo teste chamado Teste do Estimador “Sanduíche” (*TESD*), é baseado no vetor formado pela diagonal principal do estimador “sanduíche” da matriz de covariâncias do estimador de máxima verossimilhança do vetor de parâmetros do modelo. Este vetor é adicionado ao vetor formado pelos elementos diagonais da inversa da matriz de informação esperada. Os autores consideram que a estatística do teste tem distribuição normal assintótica, e sob algumas condições de regularidade apresentadas no apêndice C, constroem um estimador consistente para a matriz de covariâncias da estatística.

Nossa proposta contempla a construção de três testes de fácil manejo computacional a partir da proposta de [Abad et al. \(2010\)](#) em conjunto com diagnósticos gráficos, [Singer et al. \(2017\)](#). Os testes permitem identificar um modelo linear misto gaussiano com

erro de especificado, e os diagnósticos gráficos permitem localizar a fonte dos erros, quando houver. Em resumo, a proposta avalia a combinação de dois métodos um método analítico com outro gráfico.

No primeiro caso, propomos usar a estatística apresentada por [Abad et al. \(2010\)](#) no caso do *TMIM*, mas usar a matriz de informação observada avaliada no estimador de máxima verossimilhança do vetor de parâmetros do modelo como um estimador da matriz de covariâncias esperada, e propomos uma forma alternativa para calcular um estimador consistente para a matriz de covariâncias da estatística. O teste proposto é chamado Teste Alternativo da Matriz de Informação.

No segundo caso, propomos uma estatística baseada no estimador “sanduíche” da matriz de covariâncias do estimador de máxima verossimilhança do vetor de parâmetros, com o objetivo de substituir a matriz de informação esperada pela matriz de informação observada, o que permite capturar toda a informação sobre os parâmetros do modelo contida nos dados, o que permite fazer a diferencia com a proposta no teste *TESD* apresentado por [Abad et al. \(2010\)](#). Além disso, sugerimos um estimador consistente para a matriz de covariâncias da estatística, o teste construído é chamado, Teste do Estimador “Sanduíche” Modificado.

No terceiro caso, propomos usar a mesma estatística que no caso anterior, mudando a forma de calcular a matriz de covariâncias da estatística. Isto permite facilitar os cálculos computacionais. Para a construção da matriz de covariâncias é usada uma expansão em série de Taylor da estatística proposta e o método Delta. O teste é chamado de Teste do Estimador “Sanduíche” Alternativo. As propriedades estatísticas de todos os testes serão avaliadas por meio de estudos de simulação.

No Capítulo 2 descrevemos o modelo linear misto, especificando notação, estimação, incluindo as propostas dos testes e as técnicas de diagnóstico. No Capítulo 3, apresentamos o processo de simulação. No Capítulo 4 fazemos uma aplicação com dados reais, utilizando os testes propostos em conjunto com técnicas de diagnóstico com o fim de encontrar um modelo adequado. No Capítulo 5 são apresentadas as conclusões e estudos futuros. Detalhes técnicos são apresentados no Apêndice.

Capítulo 2

Modelo linear misto

A forma geral de um modelo linear misto é

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \mathbf{e}_i, \quad i = 1, \dots, n, \quad (2.1)$$

independentes em que $\mathbf{y}_i = (y_{i_1}, y_{i_2}, \dots, y_{i_{m_i}})^\top$ representa o vetor das m_i observações registradas para a i -ésima unidade amostral, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ denota o vetor de parâmetros de localização ou efeitos fixos, \mathbf{X}_i é a matriz correspondente à especificação dos termos fixos, $\mathbf{b}_i = (b_{1i}, \dots, b_{q_i})^\top$ representa o vetor de efeitos aleatórios, \mathbf{Z}_i é a matriz que corresponde à especificação do vetor de efeitos aleatórios e \mathbf{e}_i representa o vetor de erros aleatórios. Supõe-se independência entre \mathbf{b}_i e \mathbf{e}_i , $\mathbf{b}_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{G}(\boldsymbol{\theta})]$, $\mathbf{e}_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i(\boldsymbol{\theta})]$, $i = 1, \dots, n$. O vetor $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)^\top$ contém todos os componentes (parâmetros) não redundantes da matriz de covariâncias do vetor \mathbf{b}_i . O vetor $\boldsymbol{\phi} = (\boldsymbol{\beta}^\top, \boldsymbol{\theta}^\top)^\top (s \times 1)$ representa o vetor de todos os parâmetros no modelo (2.1). A matriz de covariâncias do vetor \mathbf{y}_i pode ser escrita como

$$\text{Var}(\mathbf{y}_i) = \mathbb{V}_i(\boldsymbol{\theta}) = \mathbb{V}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^\top + \mathbf{R}_i, \quad (2.2)$$

em que $\mathbf{G} = \mathbf{G}(\boldsymbol{\theta})$ e $\mathbf{R}_i = \mathbf{R}_i(\boldsymbol{\theta})$. Em resumo,

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}[\mathbf{X}_i\boldsymbol{\beta}, \mathbb{V}_i]. \quad (2.3)$$

Segundo Laird & Ware (1982), o modelo (2.1) pode ser expresso em dois estágios. No primeiro estágio, considera-se a distribuição condicional do vetor de respostas \mathbf{y}_i dado \mathbf{b}_i , ou seja,

$$\mathbf{y}_i | \mathbf{b}_i \sim \mathcal{N}_{m_i}[\mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i, \mathbf{R}_i], \quad (2.4)$$

e no segundo estágio é considerada a distribuição

$$\mathbf{b}_i \sim \mathcal{N}_q[\mathbf{0}, \mathbf{G}]. \quad (2.5)$$

Quando $\mathbf{R}_i = \sigma^2 \mathbf{I}_{m_i}$, o modelo resultante é conhecido como modelo de independência condicional homocedástico.

O modelo (2.1) pode ser expresso em forma compacta como

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \mathbf{e}, \quad (2.6)$$

em que $\mathbf{y} = (\mathbf{y}_1^\top, \mathbf{y}_2^\top, \dots, \mathbf{y}_n^\top)^\top (N \times 1)$, $N = \sum_{i=1}^n m_i$, contém as respostas das n unidades amostrais, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$, $\mathbf{X} = (\mathbf{X}_1^\top, \dots, \mathbf{X}_n^\top)^\top (N \times p)$, $\mathbf{Z} = \oplus_{i=1}^n \mathbf{Z}_i (N \times nq)$, $\mathbf{b} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_n^\top)^\top (nq \times 1)$, $\mathbf{e} = (\mathbf{e}_1^\top, \dots, \mathbf{e}_n^\top)^\top (N \times 1)$, $\mathbf{b} \sim \mathcal{N}[\mathbf{0}, \boldsymbol{\Psi}]$ e $\mathbf{e} \sim \mathcal{N}[\mathbf{0}, \mathbf{R}]$, com $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\boldsymbol{\theta}) = \mathbf{I}_n \otimes \mathbf{G}$, $\mathbf{R} = \mathbf{R}(\boldsymbol{\theta}) = \oplus_{i=1}^n \mathbf{R}_i$, e \mathbf{I}_n denota a matriz identidade de ordem n . Os símbolos \oplus e \otimes indicam a soma direita e o produto de Kronecker respectivamente, [Searle (1982, p.264)]. A distribuição do vetor aleatório \mathbf{y} é

$$\mathbf{y} \sim \mathcal{N}_N[\mathbf{X}\boldsymbol{\beta}, \mathbb{V}(\boldsymbol{\theta})], \quad (2.7)$$

em que

$$\text{Var}(\mathbf{y}) = \mathbb{V}(\boldsymbol{\theta}) = \mathbf{Z}\boldsymbol{\Psi}\mathbf{Z}^\top + \mathbf{R}.$$

2.1 Inferência

Os métodos de estimação mais usados para estimar os parâmetros do modelo (2.1) são o método de Máxima Verossimilhança (MV), o método de Máxima Verossimilhança Restrita (MVR) tratados por exemplo, por Patterson & Thompson (1977) e Robinson (1991), entre outros e o método bayesiano, abordado por Tountenburg (1982) e Jiang (2007), por exemplo.

A metodologia de máxima verossimilhança gera estimadores não enviesados para os efeitos fixos. No entanto, produz estimadores enviesados para os parâmetros da matriz de covariâncias intraunidades amostrais. Este viés surge devido à não consideração da perda de graus de liberdade na estimação dos termos fixos.

Os vetores aleatórios $\mathbf{y}_1, \dots, \mathbf{y}_n$ são independentes com distribuição indicada em (2.3). A função de densidade de probabilidade associada a cada vetor \mathbf{y}_i é denotada $f(\mathbf{y}_i; \boldsymbol{\phi})$. Levando em consideração as funções densidade de probabilidade das variáveis aleatórias (2.4) e (2.5), a função de verossimilhança de $\boldsymbol{\phi}$ é

$$L(\boldsymbol{\phi}; \mathbf{y}) = \prod_{1 \leq i \leq n} f(\mathbf{y}_i; \boldsymbol{\phi}) = \prod_{1 \leq i \leq n} \int_{\mathcal{R}^q} f(\mathbf{y}_i; \boldsymbol{\phi} | \mathbf{b}_i) f(\mathbf{b}_i; \boldsymbol{\phi}) d\mathbf{b}_i. \quad (2.8)$$

O logaritmo da função de verossimilhança é

$$l(\boldsymbol{\phi}; \mathbf{y}) = \sum_{i=1}^n l_i(\boldsymbol{\phi}; \mathbf{y}_i), \quad (2.9)$$

em que

$$l_i(\boldsymbol{\phi}; \mathbf{y}_i) = -\frac{m_i}{2} \log(2\pi) - \frac{1}{2} \log |\mathbb{V}_i(\boldsymbol{\theta})| - \frac{1}{2} \left\{ (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta})^\top [\mathbb{V}_i(\boldsymbol{\theta})]^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}) \right\} \quad (2.10)$$

representa a contribuição do i -ésimo indivíduo ao logaritmo da verossimilhança do modelo marginal. Alternativamente $l(\boldsymbol{\phi}; \mathbf{y})$ pode ser escrita como

$$l(\boldsymbol{\phi}; \mathbf{y}) = -\frac{1}{2} \left\{ N \log(2\pi) + \log |\mathbb{V}(\boldsymbol{\theta})| + (\mathbf{y} - \mathbf{X} \boldsymbol{\beta})^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}) \right\}. \quad (2.11)$$

A derivada parcial de $l(\boldsymbol{\phi}; \mathbf{y})$ com relação ao vetor $\boldsymbol{\phi}$ é o vetor escore de $\boldsymbol{\phi}$, cuja expressão é

$$\mathbf{s}(\boldsymbol{\phi}; \mathbf{y}) = \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\phi}} = \nabla l(\boldsymbol{\phi}; \mathbf{y}) = \begin{bmatrix} \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\beta}} \\ \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\theta}} \end{bmatrix}, \quad (2.12)$$

em que o operador $\nabla(\cdot) = \frac{\partial}{\partial \boldsymbol{\phi}}(\cdot)$. Agora,

$$\frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\beta}} = \begin{bmatrix} \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \beta_0} \\ \vdots \\ \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \beta_p} \end{bmatrix}, \quad (2.13)$$

e

$$\frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \theta_1} \\ \vdots \\ \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \theta_k} \end{bmatrix}, \quad (2.14)$$

em que

$$\frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \theta_t} = -\frac{1}{2} \text{tr} \left([\mathbb{V}(\boldsymbol{\theta})]^{-1} \frac{\partial \mathbb{V}(\boldsymbol{\theta})}{\partial \theta_t} \right) - \frac{1}{2} \mathbf{e}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \frac{\partial \mathbb{V}(\boldsymbol{\theta})}{\partial \theta_t} [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{e}, \quad (2.15)$$

$t = 1, \dots, k$, e $\text{tr}(\cdot)$ representa o traço da matriz (\cdot) . Resolvendo o sistema

$$\frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\beta}} = \mathbf{0} \quad (2.16)$$

com $\boldsymbol{\theta}$ conhecido, em seguida resolvendo

$$\frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \theta_t} = 0, \quad (2.17)$$

obtemos os estimadores dos efeitos fixos e os estimadores dos parâmetros da matriz de covariâncias [ver [Gumedze & Dunne \(2011\)](#), entre outros].

Para reduzir o viés no processo de estimação dos componentes do vetor $\boldsymbol{\theta}$, por máxima verossimilhança restrita, [Harville \(1977\)](#) e [Patterson & Thompson \(1977\)](#) propuseram usar uma transformação linear do tipo $\mathbf{y}^* = \mathbf{U}^\top \mathbf{y}$ com $\mathbb{E}(\mathbf{U} \mathbf{y}^*) = \mathbf{0}$. Os autores sugerem considerar a matriz \mathbf{U} , tal que $\mathbf{U}^\top \mathbf{U} = \mathbf{I}_n$ e $\mathbf{U} \mathbf{U}^\top = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$, de forma que,

$$\mathbf{y}^* \sim \mathcal{N}_{N-p} [\mathbf{0}, \mathbf{U}^\top \mathbb{V}(\boldsymbol{\theta}) \mathbf{U}]. \quad (2.18)$$

O logaritmo da função de verossimilhança marginal restrita é

$$l_R(\boldsymbol{\theta}; \mathbf{y}) = -\frac{(N-p)}{2} \log(2\pi) - \frac{1}{2} \log |\mathbb{V}(\boldsymbol{\theta})| - \frac{1}{2} \log |\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X}| - \frac{(N-p)}{2} \hat{\mathbf{e}}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \hat{\mathbf{e}}, \quad (2.19)$$

em que $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$, $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1} \mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{y}$, $\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})$ é o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ obtido sob a suposição de que $\boldsymbol{\theta}$ é conhecido. A maximização de (2.19) gera os estimadores $\hat{\boldsymbol{\theta}}_R$ de máxima verossimilhança restrita de $\boldsymbol{\theta}$. $\hat{\boldsymbol{\theta}}_R$ em conjunto com $\hat{\boldsymbol{\beta}}_R = \hat{\boldsymbol{\beta}}(\hat{\boldsymbol{\theta}}_R)$, constituem os estimadores de máxima verossimilhança restrita desejados. Detalhes do procedimento de maximização podem ser encontrado em Demidenko (2013). A função (2.19) também pode ser escrita como

$$l_R(\boldsymbol{\theta}; \mathbf{y}) = -\frac{1}{2} \left\{ (N-p) \log(2\pi) + \log |\mathbb{V}(\boldsymbol{\theta})| + \log |\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X}| + \mathbf{y}^\top \mathbf{P} \mathbf{y} \right\}, \quad (2.20)$$

com

$$\mathbf{P} = [\mathbb{V}(\boldsymbol{\theta})]^{-1} - [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X} (\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1} \mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1}.$$

Diferenciando (2.20) com respeito a cada um dos componentes do vetor $\boldsymbol{\theta}$, temos

$$\frac{\partial l_R(\boldsymbol{\theta}; \mathbf{y})}{\partial \theta_t} = -\frac{1}{2} \left[\frac{\partial \log |\mathbb{V}(\boldsymbol{\theta})|}{\partial \theta_t} + \frac{\partial \log |\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X}|}{\partial \theta_t} + \mathbf{y}^\top \frac{\partial \mathbf{P}}{\partial \theta_t} \mathbf{y} \right], \quad (2.21)$$

$t = 1, \dots, k$, em que

$$\begin{aligned} \frac{\partial \log |\mathbb{V}(\boldsymbol{\theta})|}{\partial \theta_t} &= -\text{tr} \left([\mathbb{V}(\boldsymbol{\theta})]^{-1} \frac{\partial \mathbb{V}(\boldsymbol{\theta})}{\partial \theta_t} \right), \\ \frac{\partial \mathbf{P}}{\partial \theta_t} &= -\mathbf{P} \frac{\partial \mathbb{V}(\boldsymbol{\theta})}{\partial \theta_t} \mathbf{P}, \end{aligned} \quad (2.22)$$

e

$$\begin{aligned} \frac{\partial \log |\mathbf{X}^\top \mathbb{V}(\boldsymbol{\theta})^{-1} \mathbf{X}|}{\partial \theta_t} &= -\text{tr} \left((\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1} \frac{\partial (\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})}{\partial \theta_t} \right), \\ &= -\text{tr} \left((\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1} \mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \frac{\partial \mathbb{V}(\boldsymbol{\theta})}{\partial \theta_t} [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X} \right). \end{aligned} \quad (2.23)$$

Procedimentos iterativos permitem encontrar as estimativas para os elementos (componentes) da matriz de covariâncias contidos no vetor $\boldsymbol{\theta}$. Detalhes desses procedimentos podem ser encontrados em [Zhu \(2016\)](#).

Preditores dos efeitos aleatórios podem ser obtidos conjuntamente com os estimadores para os parâmetros de localização por meio das equações de [Henderson \(1975\)](#). [Henderson \(1975\)](#) maximiza a distribuição de probabilidade conjunta das variáveis aleatórias \mathbf{y} e \mathbf{b} com respeito a $\boldsymbol{\beta}$ e com respeito a \mathbf{b} . Esse processo permite chegar às equações

$$\begin{pmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \boldsymbol{\Psi}^{-1} \end{pmatrix} \begin{pmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{y} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{y} \end{pmatrix}. \quad (2.24)$$

Considerando (2.24) e supondo $\boldsymbol{\Psi}$ e \mathbf{R} conhecidos de forma que $\mathbb{V}(\boldsymbol{\theta})$ é conhecida, o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ é

$$\hat{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\mathbf{X}^\top \mathbb{V}(\boldsymbol{\theta})^{-1} \mathbf{X})^{-1} \mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{y}. \quad (2.25)$$

Além disso, a predição de \mathbf{b} é dada por

$$\hat{\mathbf{b}}(\boldsymbol{\theta}) = \boldsymbol{\Psi} \mathbf{Z}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}). \quad (2.26)$$

Detalhes do processo de estimação podem ser encontrados em [Henderson \(1975\)](#), [McCulloch & Searle \(2001, p.259\)](#).

Pode-se mostrar que

$$\mathbb{E}(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})) = \boldsymbol{\beta}(\boldsymbol{\theta}),$$

e que $\mathbf{Var}(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})) = \mathbb{V}(\hat{\boldsymbol{\beta}}(\boldsymbol{\theta})) = (\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1}$, quando $\mathbf{Var}(\mathbf{y}) = \mathbb{V}(\boldsymbol{\theta})$. Além do anterior,

$$\mathbf{Var}(\hat{\mathbf{b}}(\boldsymbol{\theta})) = \mathbb{V}(\hat{\mathbf{b}}(\boldsymbol{\theta})) = \boldsymbol{\Psi} \mathbf{Z}^\top \mathbf{Q}(\boldsymbol{\theta}) \mathbf{Z} \boldsymbol{\Psi}, \quad (2.27)$$

em que

$$\mathbf{Q}(\boldsymbol{\theta}) = [\mathbb{V}(\boldsymbol{\theta})]^{-1} - [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X} (\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1} \mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1}$$

e

$$\text{Var}(\widehat{\mathbf{b}}(\boldsymbol{\theta}) - \mathbf{b}(\boldsymbol{\theta})) = \mathbb{V}(\widehat{\mathbf{b}}(\boldsymbol{\theta}) - \mathbf{b}(\boldsymbol{\theta})) = \boldsymbol{\Psi} - \mathbb{V}(\widehat{\mathbf{b}}(\boldsymbol{\theta})).$$

Como em geral as matrizes \mathbf{R} e $\boldsymbol{\Psi}$ são funções de $\boldsymbol{\theta}$, uma vez estimado o vetor $\boldsymbol{\theta}$, obtemos $\widehat{\mathbf{R}} = \mathbf{R}(\widehat{\boldsymbol{\theta}})$ e $\widehat{\boldsymbol{\Psi}} = \boldsymbol{\Psi}(\widehat{\boldsymbol{\theta}})$. Substituindo $\boldsymbol{\theta}$ por $\widehat{\boldsymbol{\theta}}$ e resolvendo o sistema das equações de Henderson (1975), obtemos o melhor estimador linear não enviesado empírico (Empirical best linear unbiased estimator - EBLUE) para $\boldsymbol{\beta}$ e o melhor preditor não enviesado empírico (Empirical best linear unbiased predictor - EBLUP) para o vetor \mathbf{b} , nomeadamente

$$\widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\theta}}) = \left(\mathbf{X}^\top [\mathbb{V}(\widehat{\boldsymbol{\theta}})]^{-1} \mathbf{X} \right)^{-1} \mathbf{X}^\top [\mathbb{V}(\widehat{\boldsymbol{\theta}})]^{-1} \mathbf{y} \quad (2.28)$$

e

$$\widehat{\mathbf{b}}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{\Psi}(\widehat{\boldsymbol{\theta}}) \mathbf{Z}^\top [\mathbb{V}(\widehat{\boldsymbol{\theta}})]^{-1} (\mathbf{y} - \mathbf{X} \widehat{\boldsymbol{\beta}}(\widehat{\boldsymbol{\theta}})). \quad (2.29)$$

Não é difícil mostrar que

$$\begin{aligned} \mathbb{V}(\widehat{\mathbf{b}}(\widehat{\boldsymbol{\theta}})) &= \mathbb{V}(\widehat{\mathbf{b}}(\widehat{\boldsymbol{\theta}}) - \mathbf{b}) \\ &= \boldsymbol{\Psi}(\boldsymbol{\theta}) \mathbf{Z}^\top \left([\mathbb{V}(\boldsymbol{\theta})]^{-1} - [\mathbb{V}(\widehat{\boldsymbol{\theta}})]^{-1} \mathbf{X} (\mathbf{X}^\top [\mathbb{V}(\boldsymbol{\theta})]^{-1} \mathbf{X})^{-1} \right) \mathbf{Z} \boldsymbol{\Psi}(\boldsymbol{\theta}), \end{aligned} \quad (2.30)$$

2.2 Propriedades assintóticas do estimador de máxima verossimilhança

Em geral não conhecemos o processo gerador do vetor \mathbf{y} , ou seja, não conhecemos qual é a verdadeira função densidade de probabilidade $g(\mathbf{y})$. O que fazemos é propor um modelo linear misto com as suposições a respeito da distribuição dos efeitos aleatórios e dos erros aleatórios.

O vetor aleatório \mathbf{y} terá uma função densidade $f(\mathbf{y}, \boldsymbol{\phi})$. Se existe um vetor $\boldsymbol{\phi}_0 \in \Theta$ tal que $g(\mathbf{y}) = f(\mathbf{y}, \boldsymbol{\phi}_0)$, pode-se concluir que o modelo está corretamente especificado; em caso contrário, o modelo terá erro de especificação. Yu et al. (2018) ilustram que quando o modelo está corretamente especificado, $\widehat{\boldsymbol{\phi}}_n$, obtido por máxima verossimilhança ou por

máxima verossimilhança restrita, é um estimador consistente para ϕ_0 , ou seja,

$$\hat{\phi}_n \xrightarrow{p} \phi_0. \quad (2.31)$$

Detalhes sobre convergência em probabilidade podem ser encontrados em [Sen & Singer \(1993, p.34\)](#).

Quando o modelo está incorretamente especificado, existe um vetor $\phi^* \in \Theta$, que minimiza o critério de informação (ou distância) de Kullback – Leibler [[Abad et al. \(2010\)](#)],

$$KL(g : f, \phi) = \mathbb{E}_g \left[\log \frac{g(\mathbf{y})}{f(\mathbf{y}, \phi)} \right] = \int_{\mathbb{R}^N} g(\mathbf{y}) \log \frac{g(\mathbf{y})}{f(\mathbf{y}, \phi)} d\mathbf{y}. \quad (2.32)$$

[White \(1982\)](#) prova que quando o modelo está corretamente especificado, o único valor que minimiza o critério de informação de KL é $\phi^* = \phi_0$. Suponhamos que as seguintes matrizes existem

$$\begin{aligned} \mathbf{A}(\phi) &= \mathbb{E} \left[\left(\frac{\partial^2 l(\phi; \mathbf{y})}{\partial \phi_k \partial \phi_l} \right) \right], & \mathbf{B}(\phi) &= \mathbb{E} \left[\left(\frac{\partial l(\phi; \mathbf{y})}{\partial \phi_k} \frac{\partial l(\phi; \mathbf{y})}{\partial \phi_l} \right) \right], \\ \mathbf{A}_n(\phi) &= \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \phi_k \partial \phi_l} \right), & \mathbf{B}_n(\phi) &= \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_k} \frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_l} \right), \\ \mathbf{A}(\phi_0) &= \mathbb{E} \left[\left(\frac{\partial^2 l(\phi; \mathbf{y})}{\partial \phi_k \partial \phi_l} \right) \right]_{\phi=\phi_0}, & \mathbf{B}(\phi_0) &= \mathbb{E} \left[\left(\frac{\partial l(\phi; \mathbf{y})}{\partial \phi_k} \frac{\partial l(\phi; \mathbf{y})}{\partial \phi_l} \right) \right]_{\phi=\phi_0}, \\ \mathbf{A}_n(\hat{\phi}_n) &= \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \phi_k \partial \phi_l} \right)_{\phi=\hat{\phi}_n}, & \mathbf{B}_n(\hat{\phi}_n) &= \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_k} \frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_l} \right)_{\phi=\hat{\phi}_n}. \end{aligned}$$

Em tal caso, de acordo com [White \(1982\)](#)

$$\mathbf{A}_n(\hat{\phi}_n) \xrightarrow{qc} \mathbf{A}(\phi).$$

Seja $\mathbf{H} = \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \phi_k \partial \phi_l}$ e $-\mathbf{H}$ a matriz de informação observada. Sob as condições de regularidade apresentadas no Apêndice C, se o modelo (2.1) está corretamente especificado,

$$\mathbf{A}(\phi_0) + \mathbf{B}(\phi_0) = \mathbf{0}, \quad (2.33)$$

[[White \(1982\)](#)]. A expressão (2.33) é chamada de igualdade da matriz de informação.

Sob as mesmas condições indicadas acima é possível demonstrar

$$\sqrt{n}[\hat{\phi}_n - \phi_0] \xrightarrow{d} \mathcal{N}_s[\mathbf{0}, \mathbb{V}(\phi_0)], \quad (2.34)$$

em que

$$\mathbb{V}(\phi_0) = [\mathbf{A}(\phi_0)]^{-1} \mathbf{B}(\phi_0) [\mathbf{A}(\phi_0)]^{-1}. \quad (2.35)$$

Detalhes podem ser encontrados em [Freedman \(2006\)](#).

2.3 Detecção de erros de especificação num modelo linear misto gaussiano

Em modelos lineares mistos, erros de especificação podem ocorrer na estrutura da média, ou na estrutura da matriz de covariâncias do vetor de respostas, nesse caso pode ocorrer

- Quando o número de efeitos aleatórios está incorrecto.
- Quando o número de efeitos aleatórios está correto, mas a matriz \mathbf{G} está incorretamente especificada.
- Quando a matriz de covariâncias intraunidades amostrais \mathbf{R}_i está incorretamente especificada.
- Quando o número de efeitos aleatórios está incorrecto e a matriz de covariâncias \mathbf{R}_i está incorretamente especificada.

Nas próximas seções apresentamos métodos analíticos que utilizam testes para identificar quando um modelo linear misto gaussiano tem erro de especificação. Esses testes não indicam a fonte do erro. Neste caso, métodos gráficos podem ser utilizados para localizar e identificar a fonte do erro.

2.3.1 Métodos analíticos

2.3.1.1 Teste Alternativo da Matriz de Informação

O Teste Alternativo da Matriz de Informação é baseado na hipótese H_0 de que se o modelo (2.1) estiver corretamente especificado, verifica-se $\mathbf{A}(\phi_0) + \mathbf{B}(\phi_0) = \mathbf{0}$.

Com base numa amostra aleatória $\mathbf{y}_1, \dots, \mathbf{y}_n$ as matrizes $\mathbf{A}(\phi_0)$ e $\mathbf{B}(\phi_0)$ são estimadas por $\mathbf{A}_n(\hat{\phi}_n)$ e por $\mathbf{B}_n(\hat{\phi}_n)$. Por simplicidade consideramos a estatística $\mathbf{d}_n(\mathbf{y}; \hat{\phi}_n)$ constituída pela diagonal principal da matriz $\mathbf{A}_n(\hat{\phi}_n) + \mathbf{B}_n(\hat{\phi}_n)$, ou seja

$$\begin{aligned} \mathbf{d}_n(\mathbf{y}; \hat{\phi}_n) &= \text{diag} \left[\mathbf{A}_n(\hat{\phi}_n) + \mathbf{B}_n(\hat{\phi}_n) \right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbf{d}(\mathbf{y}_i; \phi) \Big|_{\phi=\hat{\phi}_n} \end{aligned} \quad (2.36)$$

em que $\mathbf{d}(\mathbf{y}_i; \phi)$ denota o vetor de dimensão $(s \times 1)$ com componentes

$$d_l(\mathbf{y}_i; \phi) = \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \phi_l^2} + \left\{ \frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_l} \right\}^2, l = 1, \dots, s; \quad (2.37)$$

ou seja $\mathbf{d}(\mathbf{y}_i; \phi)$ é da forma

$$\mathbf{d}(\mathbf{y}_i; \phi) = \begin{pmatrix} \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \phi_1^2} \\ \vdots \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \phi_s^2} \end{pmatrix} + \begin{pmatrix} \left\{ \frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_1} \right\}^2 \\ \vdots \\ \left\{ \frac{\partial l_i(\phi; \mathbf{y}_i)}{\partial \phi_s} \right\}^2 \end{pmatrix} = \begin{pmatrix} d_1(\mathbf{y}_i; \phi) \\ \vdots \\ d_s(\mathbf{y}_i; \phi) \end{pmatrix}. \quad (2.38)$$

Abad et al. (2010) propõem a estatística $\mathbf{d}(\mathbf{y}; \phi) = \text{diag} [\mathbf{A}(\phi) + \mathbf{B}(\phi)]$ e estimam as matrizes $\mathbf{A}(\phi)$ e $\mathbf{B}(\phi)$ por meio das matrizes $\mathbf{A}_n(\phi_0)$ e $\mathbf{B}_n(\phi_0)$, isso faz a diferença entre a proposta dos autores e a nossa proposta.

Para a construção da matriz de covariâncias da estatística (2.36), denotada $\mathbb{M}(\phi_0)$, sob H_0 usamos uma expansão em Taylor de primeira ordem da estatística $\sqrt{n} \mathbf{d}_n(\mathbf{y}; \hat{\phi}_n)$ ao redor de ϕ_0 para obter

$$\mathbb{M}(\boldsymbol{\phi}_0) = \mathbb{E} \left[\left(\mathbf{a}(\boldsymbol{\phi}) \mathbf{a}(\boldsymbol{\phi})^T \right) \right] \Big|_{\boldsymbol{\phi}=\boldsymbol{\phi}_0}, \quad (2.39)$$

com

$$\mathbf{a}(\boldsymbol{\phi}) = \mathbf{d}(\mathbf{y}; \boldsymbol{\phi}) - \nabla \mathbf{d}(\boldsymbol{\phi}) [\mathbf{A}(\boldsymbol{\phi})]^{-1} \nabla l(\boldsymbol{\phi}; \mathbf{y}) \quad (2.40)$$

em que $\nabla \mathbf{d}(\boldsymbol{\phi}) = \frac{\partial \mathbf{d}(\mathbf{y}; \boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ e $\nabla l(\boldsymbol{\phi}; \mathbf{y}) = \frac{\partial l(\boldsymbol{\phi}; \mathbf{y})}{\partial \boldsymbol{\phi}}$, detalhes da demonstração são encontrados em [Gourieroux & Monfort \(1995, p.183\)](#) e [Zeigler \(2011, p.46\)](#). A matriz (2.39) pode ser estimada por

$$\widehat{\mathbb{M}}_n(\widehat{\boldsymbol{\phi}}_n) = \frac{1}{n} \sum_{i=1}^n \left[\mathbf{a}_i(\boldsymbol{\phi}) \mathbf{a}_i(\boldsymbol{\phi})^T \right] \Big|_{\boldsymbol{\phi}=\widehat{\boldsymbol{\phi}}_n}, \quad (2.41)$$

em que

$$\mathbf{a}_i(\boldsymbol{\phi}) = \mathbf{d}(\mathbf{y}_i; \boldsymbol{\phi}) - \nabla \mathbf{d}(\boldsymbol{\phi}) [\mathbf{A}(\boldsymbol{\phi})]^{-1} \nabla l_i(\boldsymbol{\phi}; \mathbf{y}_i). \quad (2.42)$$

Sob as condições de regularidade no apêndice C, e sob H_0 , a estatística $\sqrt{n} \mathbf{d}_n(\mathbf{y}; \widehat{\boldsymbol{\phi}}_n)$, é tal que

$$\sqrt{n} \mathbf{d}_n(\mathbf{y}, \widehat{\boldsymbol{\phi}}_n) \stackrel{A}{\sim} \mathcal{N}_s[\mathbf{0}, \mathbb{M}(\boldsymbol{\phi}_0)], \quad (2.43)$$

em que a expressão $\stackrel{A}{\sim}$ significa distribuído assintoticamente [[White \(1982\)](#)]. Considerando (2.43) e usando o Teorema de Cochran [[Sen & Singer \(1993, p.137\)](#)], a estatística do Teste Alternativo da Matriz de Informação, toma a forma

$$EAMI = n \mathbf{d}_n(\mathbf{y}, \widehat{\boldsymbol{\phi}}_n)^\top [\widehat{\mathbb{M}}_n(\widehat{\boldsymbol{\phi}}_n)]^{-1} \mathbf{d}_n(\mathbf{y}, \widehat{\boldsymbol{\phi}}_n), \quad (2.44)$$

e sob H_0 segue uma distribuição χ_s^2 , onde s são os graus de liberdade. Detalhes são indicados no apêndice D.

A taxa do erro tipo I associada com a região $EAMI > \chi_{\alpha,s}^2$, é

$$P(EAMI > \chi_{\alpha,s}^2 | H_0), \quad (2.45)$$

em que $\chi_{\alpha,s}^2$ denota o quantil α de uma distribuição qui quadrado com s graus de liberdade.

Se H_0 for falsa, com H_1 a hipótese alternativa,

$$1 - P(EAMI < \chi_{\alpha,s}^2 | H_1), \quad (2.46)$$

representa o poder do teste.

2.3.1.2 Teste do Estimador “Sanduíche” Modificado

Nesta seção desenvolvemos a proposta do Teste do Estimador “Sanduíche” Modificado baseado no estimador “sanduíche” da matriz de covariâncias do estimador $\hat{\phi}_n$, a saber $[\mathbf{A}_n(\hat{\phi}_n)]^{-1} \mathbf{B}_n(\hat{\phi}_n) [\mathbf{A}_n(\hat{\phi}_n)]^{-1}$ [Freedman (2006)], e na hipótese H_0 que se o modelo (2.1) é corretamente especificado, verifica-se

$$[\mathbf{A}_n(\hat{\phi}_n)]^{-1} \mathbf{B}_n(\hat{\phi}_n) [\mathbf{A}_n(\hat{\phi}_n)]^{-1} + [\mathbf{A}_n(\hat{\phi}_n)]^{-1} = \mathbf{0}.$$

Para a construção do teste propomos a estatística,

$$\begin{aligned} \mathbf{d}_n^*(\hat{\phi}_n) &= \text{diag} \left[[\mathbf{A}_n(\hat{\phi}_n)]^{-1} \mathbf{B}_n(\hat{\phi}_n) [\mathbf{A}_n(\hat{\phi}_n)]^{-1} + [\mathbf{A}_n(\hat{\phi}_n)]^{-1} \right] \\ &= \text{diag} \left[[\mathbf{A}_n(\hat{\phi}_n)]^{-1} \mathbf{B}_n(\hat{\phi}_n) [\mathbf{A}_n(\hat{\phi}_n)]^{-1} \right] + \text{diag} \left[[\mathbf{A}_n(\hat{\phi}_n)]^{-1} \right] \\ &= \frac{1}{n} \sum_{i=1}^n \mathbf{d}_i^*(\phi) \Big|_{\phi=\hat{\phi}_n} \end{aligned} \quad (2.47)$$

que é similar à proposta dos autores [Abad et al. \(2010\)](#). A diferença está em que nosso caso, utilizando a informação do vetor ϕ contida nos dados, estimamos a matriz $\mathbf{A}(\phi)$, por meio da matriz $\mathbf{A}_n(\hat{\phi}_n)$ enquanto os autores estimam a matriz $\mathbf{A}(\phi)$ por meio da matriz $\mathbf{A}_n(\phi_0)$.

Em (2.47)

$$\begin{aligned} \mathbf{d}_i^*(\boldsymbol{\phi}) = & \text{diag} \left[\left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j \partial \phi_h} \right)^{-1} \left(\frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j} \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_h} \right) \left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j \partial \phi_h} \right)^{-1} \right] \\ & + \text{diag} \left[\left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j \partial \phi_h} \right)^{-1} \right], \end{aligned}$$

agora,

$$\begin{aligned} \mathbf{d}_{1i}^*(\boldsymbol{\phi}) &= \text{diag} \left[\left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j \partial \phi_h} \right)^{-1} \left(\frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j} \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_h} \right) \left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j \partial \phi_h} \right)^{-1} \right] \text{ e} \\ \mathbf{d}_{2i}^*(\boldsymbol{\phi}) &= \text{diag} \left[\left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_j \partial \phi_h} \right)^{-1} \right] \text{ e seja} \end{aligned}$$

$$\mathbf{d}_1^*(\hat{\boldsymbol{\phi}}_n) = \text{diag} \left[[\mathbf{A}_n(\hat{\boldsymbol{\phi}}_n)]^{-1} \mathbf{B}_n(\hat{\boldsymbol{\phi}}_n) [\mathbf{A}_n(\hat{\boldsymbol{\phi}}_n)]^{-1} \right] = \sum_{i=1}^n \mathbf{d}_{1i}^*(\boldsymbol{\phi}) \Big|_{\boldsymbol{\phi}=\hat{\boldsymbol{\phi}}_n} \quad (2.48)$$

e

$$\mathbf{d}_2^*(\hat{\boldsymbol{\phi}}_n) = \text{diag} \left[[\mathbf{A}_n(\hat{\boldsymbol{\phi}}_n)]^{-1} \right] = \sum_{i=1}^n \mathbf{d}_{2i}^*(\boldsymbol{\phi}) \Big|_{\boldsymbol{\phi}=\hat{\boldsymbol{\phi}}_n}, \quad (2.49)$$

$i = 1, \dots, n$; $j = 1, \dots, s$ e $h = 1, \dots, s$. A estatística (2.47) é indicador potencial para detectar erro de especificação no modelo (2.1).

De maneira similar como no teste anterior, sob as condições de regularidade no Apêndice C e sob H_0 , aplicando o teorema central do limite, a estatística $\sqrt{n} \mathbf{d}_n^*(\hat{\boldsymbol{\phi}}_n)$ tem distribuição normal com média $\mathbf{0}$ e matriz de covariâncias denotada por $\mathbf{C}_v(\boldsymbol{\phi}_0)$, considerando que

$$\mathbb{E} \left(\mathbf{d}_{1i}^*(\hat{\boldsymbol{\phi}}_n) \right) = \boldsymbol{\mu}_1^*(\boldsymbol{\phi}), \text{ e } \mathbb{E} \left(\mathbf{d}_{2i}^*(\hat{\boldsymbol{\phi}}_n) \right) = \boldsymbol{\mu}_2^*(\boldsymbol{\phi}),$$

$$\begin{aligned}
 \mathbf{C}_v(\phi) &= \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right) \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right)^T \right) \\
 &+ \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right) \left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right)^T \right) \\
 &+ \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right) \left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right)^T \right) \\
 &+ \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right) \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right)^T \right).
 \end{aligned} \tag{2.50}$$

Detalhes desta matriz são encontrados no apêndice D. Computacionalmente a estimação da matriz (2.50) requer de derivadas parciais de ordem superior de (2.10), não obstante, o manejo computacional é de fácil implementação, considerando $\boldsymbol{\mu}_1^*(\hat{\phi}_n) = \frac{1}{n} \mathbf{d}_1^*(\hat{\phi}_n)$ e $\boldsymbol{\mu}_2^*(\hat{\phi}_n) = \frac{1}{n} \mathbf{d}_2^*(\hat{\phi}_n)$, um estimador consistente da matriz (2.50) é

$$\begin{aligned}
 \widehat{\mathbf{C}}_v(\hat{\phi}_n) &= \frac{1}{n-1} \sum_{i=1}^n \left(\left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\hat{\phi}_n) \right) \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\hat{\phi}_n) \right)^T \right) \\
 &+ \frac{1}{n-1} \sum_{i=1}^n \left(\left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\hat{\phi}_n) \right) \left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\hat{\phi}_n) \right)^T \right) \\
 &+ \frac{1}{n-1} \sum_{i=1}^n \left(\left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\hat{\phi}_n) \right) \left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\hat{\phi}_n) \right)^T \right) \\
 &+ \frac{1}{n-1} \sum_{i=1}^n \left(\left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\hat{\phi}_n) \right) \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\hat{\phi}_n) \right)^T \right).
 \end{aligned} \tag{2.51}$$

Agora, sob as condições de regularidade no apêndice C e sob H_0 , temos

$$\sqrt{n} \mathbf{d}_n^*(\hat{\phi}_n) \overset{A}{\rightsquigarrow} \mathcal{N}_s[\mathbf{0}, \mathbf{C}_v(\phi_0)], \tag{2.52}$$

detalhes no apêndice D. Tendo em conta (2.52), sob H_0 , aplicando o teorema de Cochran [Sen & Singer (1993, p.137)], a estatística do teste toma a forma

$$ESDM = n\mathbf{d}_n^*(\hat{\phi}_n)^\top [\widehat{\mathbf{C}}_v(\hat{\phi}_n)]^{-1} \mathbf{d}_n^*(\hat{\phi}_n) \quad (2.53)$$

que tem uma distribuição amostral assintótica χ_s^2 , detalhes no apêndice D. A taxa do erro tipo I associado à região $ESDM > \chi_{\alpha,s}^2$, é calculada por

$$P(ESDM > \chi_{\alpha,s}^2 | H_0). \quad (2.54)$$

Se H_0 for falsa, sendo H_1 a hipótese alternativa,

$$1 - P(ESDM < \chi_{\alpha,s}^2 | H_1), \quad (2.55)$$

representa o poder do teste.

2.3.1.3 Teste do Estimador “Sanduíche” Alternativo

O Teste do Estimador “Sanduíche” Alternativo é baseado no estimador “sanduíche” da matriz de covariâncias do estimador $\hat{\phi}_n$, e na igualdade da matriz de informação (2.33) sob a hipótese H_0 que o modelo (2.1) é corretamente especificado. Considerando a distribuição assintótica do estimador $\hat{\phi}_n$ e tendo em consideração as condições de regularidades no apêndice C, é conhecido que,

$$\sqrt{n} \left(\hat{\phi}_n - \phi_0 \right) \overset{A}{\approx} \mathcal{N}_s [\mathbf{0}, \mathbb{V}(\phi_0)], \quad (2.56)$$

em que $\mathbb{V}(\phi_0)$ é como em (2.35), [White (1982)]. Considerando o vetor

$$\mathbf{d}(\phi) = \text{diag} \left[[\mathbf{A}(\phi)]^{-1} \mathbf{B}(\phi) [\mathbf{A}(\phi)]^{-1} + [\mathbf{A}(\phi)]^{-1} \right], \quad (2.57)$$

sob a suposição de que é diferenciável em ϕ_0 , podemos construir

$$\nabla \mathbf{d}(\phi_0) = \nabla \mathbf{d}(\phi) |_{\phi=\phi_0} \neq \mathbf{0}.$$

O vetor $\mathbf{d}_n^*(\hat{\phi}_n)$ apresentado em (2.47), é um estimador do vetor (2.57), e é indicador

potencial para detectar erro de especificação no modelo (2.1). Seja o vetor

$$\mathbf{d}_n(\boldsymbol{\phi}) = \text{diag} \left[[\mathbf{A}_n(\boldsymbol{\phi})]^{-1} \mathbf{B}_n(\boldsymbol{\phi}) [\mathbf{A}_n(\boldsymbol{\phi})]^{-1} + [\mathbf{A}_n(\boldsymbol{\phi})]^{-1} \right] = \begin{pmatrix} d_{11}(\boldsymbol{\phi}) \\ \vdots \\ d_{ss}(\boldsymbol{\phi}) \end{pmatrix}, \quad (2.58)$$

com o fim de construir

$$\nabla \mathbf{d}_n(\boldsymbol{\phi}) = \nabla \left[\text{diag} \left[\mathbf{A}_n(\boldsymbol{\phi})^{-1} \mathbf{B}_n(\boldsymbol{\phi}) [\mathbf{A}_n(\boldsymbol{\phi})]^{-1} + [\mathbf{A}_n(\boldsymbol{\phi})]^{-1} \right] \right] \quad (2.59)$$

$$= \begin{pmatrix} \frac{\partial d_{11}(\boldsymbol{\phi})}{\partial \phi_1} & \cdots & \frac{\partial d_{11}(\boldsymbol{\phi})}{\partial \phi_s} \\ \vdots & & \vdots \\ \frac{\partial d_{ss}(\boldsymbol{\phi})}{\partial \phi_1} & \cdots & \frac{\partial d_{ss}(\boldsymbol{\phi})}{\partial \phi_s} \end{pmatrix} (s \times s). \quad (2.60)$$

Usando (2.56), sob a hipótese H_0 e o método Delta [Sen & Singer (1993, p.136)], obtemos

$$\sqrt{n} \left(\mathbf{d}_n^*(\hat{\boldsymbol{\phi}}_n) - \mathbf{d}(\boldsymbol{\phi}_0) \right) \xrightarrow{D} \mathcal{N}_s(\mathbf{0}, \nabla \mathbf{d}(\boldsymbol{\phi}_0) \mathbb{V}(\boldsymbol{\phi}_0) \nabla \mathbf{d}(\boldsymbol{\phi}_0)^\top), \quad n \rightarrow \infty, \quad (2.61)$$

detalhes no apêndice D. Sob a hipótese H_0 e com (2.61), aplicando o teorema de Cochran [Sen & Singer (1993, p.137)], a estatística do estimador “sanduíche” alternativo toma a forma

$$EESA = n \mathbf{b}_n^*(\hat{\boldsymbol{\phi}}_n)^\top \left[\nabla \mathbf{d}_n(\hat{\boldsymbol{\phi}}_n) \widehat{\mathbb{V}}(\hat{\boldsymbol{\phi}}_n) \nabla \mathbf{d}_n(\hat{\boldsymbol{\phi}}_n)^\top \right]^{-1} \mathbf{d}_n^*(\hat{\boldsymbol{\phi}}_n), \quad (2.62)$$

o qual tem uma distribuição assintótica χ_s^2 , detalhes são ilustrados no apêndice D.

$\widehat{\mathbb{V}}(\hat{\boldsymbol{\phi}}_n) = [\mathbf{A}_n(\hat{\boldsymbol{\phi}}_n)]^{-1} \mathbf{B}_n(\hat{\boldsymbol{\phi}}_n) [\mathbf{A}_n(\hat{\boldsymbol{\phi}}_n)]^{-1}$ representa um estimador da matriz de covariâncias do estimador “sanduíche” do vetor $\hat{\boldsymbol{\phi}}$, esse estimador é não enviesado e consistente [White (1982)]. Note que para calcular (2.62) foram necessárias derivadas parciais de grau 3 da função logaritmo da verosimilhança de $f(\mathbf{y}; \boldsymbol{\phi})$, pelo que expressões em forma analítica são complicadas de serem obtidas, sendo necessárias otimizar função não lineares usando a função `optim` de R.

Sob H_0 , a taxa estimada do erro tipo I é

$$P(EESA > \chi_{\alpha,s}^2 | H_0). \quad (2.63)$$

Se H_0 for falsa e H_1 é a hipótese alternativa,

$$1 - P(EESA < \chi_{\alpha,s}^2 | H_1), \quad (2.64)$$

representa uma estimativa do poder do teste.

2.3.2 Métodos gráficos

Os diagnósticos gráficos permitem identificar e localizar a fonte de erro de especificação em um modelo linear misto gaussiano, se este existir. [Hilden \(1995\)](#) e [Rocha & Singer \(2018\)](#), por exemplo, estendem a análise de resíduos para modelos lineares mistos. No modelo misto (2.1) há duas fontes de variação, \mathbf{e} e \mathbf{b} , que geram três tipos de resíduos:

(a) os resíduos marginais, $\hat{\boldsymbol{\xi}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}$, que predizem o erro marginal,

$$\boldsymbol{\xi} = \mathbf{y} - \mathbb{E}(\mathbf{y}) = \mathbf{y} - \mathbf{X}\boldsymbol{\beta},$$

(b) os resíduos condicionais, $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{Z}\hat{\mathbf{b}}$, que predizem o erro condicional,

$$\mathbf{y} - \mathbb{E}(\mathbf{y} | \mathbf{b}) = \mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b},$$

(c) os resíduos do efeito aleatório, $\mathbf{Z}\hat{\mathbf{b}}$, que predizem o efeito aleatório,

$$\mathbf{Z}\mathbf{b} = \mathbb{E}(\mathbf{y} | \mathbf{b}) - \mathbb{E}(\mathbf{y}).$$

Cada um deles permite estudar alguma suposição do modelo.

Para analisar a suposição de linearidade entre $\mathbb{E}(\mathbf{y})$ e as covariáveis dadas em \mathbf{X} , [Hilden \(1995\)](#) e [Singer et al. \(2017\)](#) sugerem analisar o gráfico dos resíduos marginais $\hat{\boldsymbol{\xi}}$ *versus* os valores ajustados. Um comportamento não sistemático em torno do valor 0 ilustra falta de evidência contra a suposição de linearidade. De acordo com [Lesaffre & Verbeke \(1998\)](#), se a estrutura da matriz de covariâncias intraunidades amostrais é adequada, $\mathbb{V}_i = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^\top + \mathbf{R}_i$ $v_i = \|\mathbf{I}_{m_i} - \boldsymbol{\zeta}_i \boldsymbol{\zeta}_i^\top\|^2$ com $\boldsymbol{\zeta}_i = \hat{\mathbb{V}}_i^{-1/2} \hat{\boldsymbol{\xi}}_i$ e $\|\cdot\|$ denotando a norma euclidiana, deveria ter valores próximos de 0. Um gráfico de v_i *versus* os índices das unidades amostrais ajuda a detectar casos em que a matriz de covariâncias intraunidades amostrais não é adequada. Uma

estrutura de covariâncias inadequada para o i -ésimo indivíduo, por exemplo, teria valores “grandes” de v_i . Dado que $\mathbb{V}(\hat{\boldsymbol{\xi}}_i) = \mathbb{V}_i - \mathbf{X}_i(\mathbf{X}_i^\top[\mathbb{V}_i]^{-1}\mathbf{X}_i)^{-1}\mathbf{X}_i^\top$, Singer et al. (2017) consideraram a substituição dos resíduos ζ_i na expressão v_i pelos resíduos marginais padronizados $\hat{\boldsymbol{\xi}}_i^* = [\widehat{\mathbb{V}}(\hat{\boldsymbol{\xi}}_i)]^{-1/2} \hat{\boldsymbol{\xi}}_i$, em que $\mathbb{V}(\hat{\boldsymbol{\xi}}_i)$ representa o i -ésimo elemento na matriz bloco diagonal $\mathbb{V} - \mathbf{X}(\mathbf{X}\mathbb{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^\top$. Os autores também propõem $v_i^* = \sqrt{v_i}/m_i$ como uma medida padronizada da adequação da estrutura de covariância intraunidades amostrais. O diagnóstico gráfico $\hat{\boldsymbol{\xi}}_i^* v_i^*$ versus o índice i , permite a identificação de unidades que têm uma estrutura de covariâncias possivelmente inadequada. Os autores também analisam a linearidade dos efeitos fixos em modelos do tipo (2.1), considerando o gráfico dos resíduos marginais padronizados construído com, $\hat{\xi}_{ij}^* = \hat{\xi}_{ij} / [\text{diag}_j(\widehat{\mathbb{V}}(\hat{\boldsymbol{\xi}}_i))]^{1/2}$, em que $\text{diag}_j(\widehat{\mathbb{V}}(\hat{\boldsymbol{\xi}}_i))$ representa o j -ésimo elemento da diagonal principal da matriz $\mathbb{V}(\hat{\boldsymbol{\xi}})$ versus os valores de cada variável explicativa bem como versus os valores ajustados.

$$\text{Dado que } \mathbb{V}(\hat{\boldsymbol{\epsilon}}) = \mathbf{R} \left[\mathbb{V}^{-1} - \mathbb{V}^{-1}\mathbf{X}(\mathbf{X}^\top\mathbb{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^\top\mathbb{V}^{-1} \right] \mathbf{R} = \mathbf{R}\mathbf{Q}\mathbf{R},$$

Nobre & Singer (2007), consideram que os resíduos condicionais podem ser heterocedásticos e sugerem a construção do gráfico dos resíduos condicionais padronizados,

$\hat{\epsilon}_{ij}^* = \hat{\epsilon}_{ij} / [\text{diag}_{ij}(\widehat{\mathbf{R}}\widehat{\mathbf{Q}}\widehat{\mathbf{R}})]^{1/2}$, em que $\text{diag}_{ij}(\mathbf{R}\mathbf{Q}\mathbf{R})$ representa a j -ésima observação da i -ésima unidade amostral na diagonal principal da matriz $\mathbf{R}\mathbf{Q}\mathbf{R}$ versus os valores ajustados para avaliar a homoscedasticidade dos erros condicionais (nos casos de modelos com homoscedasticidade condicional) ou contra os índices de observações, para detecção de observações discrepantes.

Na ausência de confundimento, para detalhes ver (Jiang, 2007) e (Hilden, 1995), sob a suposição de normalidade para os efeitos aleatórios, $M_i = \hat{\mathbf{b}}_i^\top \left\{ \mathbb{V}(\hat{\mathbf{b}}_i^\top - \mathbf{b}_i) \right\}^{-1} \hat{\mathbf{b}}_i$ (a distância de Mahalanobis entre \mathbf{b}_i e $\mathbb{E}(\mathbf{b}_i) = \mathbf{0}$) deveria seguir uma distribuição qui-quadrado com q graus de liberdade. Um gráfico QQ [Wilk & Gnanadesikan (1968)], baseado na distribuição χ_q^2 pode ser utilizado para verificar se os efeitos aleatórios seguem uma distribuição gaussiana q -variada. A Tabela 2.1 é reproduzida de Singer et al. (2017) e apresenta um resumo das

técnicas de diagnóstico usadas para avaliar o ajuste de um modelo linear misto gaussiano do tipo (2.1). Essas técnicas permitem localizar (se houver) fontes de erro de especificação neste tipo de modelos.

Tabela 2.1: *Técnica de diagnóstico gráfico para detecção de fonte de erros de especificação em um modelo do tipo (2.1).*

Diagnóstico para	Tipo de resíduo	Gráfico
Linealidade do efeito fixo ($\mathbb{E}(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$)	Marginal	$\widehat{\varepsilon}_{ij}^*$ versus valores ajustados ou variáveis explicativas.
Presença de observações aberrantes (y_{ij})	Marginal	$\widehat{\varepsilon}_{ij}^*$ versus índices de observações.
Matriz de covariância intraunidades amostrais (\mathbb{V}_i)	Marginal	\mathbf{v}_i^* versus índices de unidades.
Presença de observações aberrantes (y_{ij})	Condicional	$\widehat{\varepsilon}_{ij}^*$ versus índices de observações.
Homoscedasticidade dos erros condicionados (e_{ij})	Condicional	$\widehat{\varepsilon}_{ij}^*$ versus valores estimados.
Normalidade dos erros condicionados (e_{ij})	Condicional	gráfico QQ normal para uma transformação de $\widehat{\varepsilon}$.
Presença de unidades distantes (\mathbf{b}_i)	Efeitos aleatórios	M_i versus índices de unidades.
Normalidade do efeito aleatório (\mathbf{b}_i)	Efeitos aleatórios	χ_q^2 QQ plot para M_i .

Capítulo 3

Estudo de Simulação

O objetivo desse capítulo é ilustrar a análise conjunta dos testes desenvolvidos no Capítulo 2, assim como das ferramentas diagnósticas sugeridas por [Singer et al. \(2017\)](#), com o fim de avaliar se um modelo misto gaussiano está incorretamente especificado, e se estiver identificar as fontes dos erros de especificação.

Com os dados em [Afiune \(2000\)](#), descritos no Capítulo 1, usamos a técnica proposta por [Rocha & Singer \(2018\)](#), a qual consiste na análise dos perfis individuais para identificar tanto a estrutura da resposta média como dos efeitos aleatórios para um modelo linear misto gaussiano. O modelo identificado foi um polinômio de grau dois na parte fixa com intercepto e inclinação, na parte aleatória um polinômio de grau um com intercepto. Especificamente o modelo é

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_{0i} + b_{1i} x_{ij} + e_{ij}, \quad (3.1)$$

em que y_{ij} representa a j -ésima observação no indivíduo i -ésimo, x_{ij} representa o j -ésimo tempo registrado para o i -ésimo indivíduo. $\beta_0, \beta_1, \beta_2$ são os parâmetros de localização, b_{0i}, b_{1i} representam o intercepto, e inclinação aleatórias e e_{ij} representa o erro aleatório. Tal como foi descrito no Capítulo 2 o modelo pode ser expresso em forma compacta, considerando

$$\mathbf{X}_i = \begin{pmatrix} 1 & x_{i_1} & x_{i_1}^2 \\ \vdots & \vdots & \vdots \\ 1 & x_{i_{m_i}} & x_{i_{m_i}}^2 \end{pmatrix}, \mathbf{Z}_i = \begin{pmatrix} 1 & x_{i_1} \\ \vdots & \vdots \\ 1 & x_{i_{m_i}} \end{pmatrix}.$$

Supondo $\mathbf{G} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$ e $\mathbf{R}_i = \sigma^2 \mathbf{I}_n$, com independência entre \mathbf{b}_i e \mathbf{e}_i , $\mathbf{b}_i \sim \mathcal{N}_2[\mathbf{0}, \mathbf{G}]$,

$e_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i]$, $i = 1, \dots, n$. O vetor \mathbf{y}_i é tal que

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}(\mathbf{X}_i\boldsymbol{\beta}, \mathbb{V}_i), \quad (3.2)$$

$\mathbb{V}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^\top + \mathbf{R}_i$. Sendo $\boldsymbol{\phi} = (\boldsymbol{\beta}^\top, \boldsymbol{\theta}^\top)^\top$ o vetor de todos os parâmetros do modelo (3.1), por máxima verossimilhança obtemos $\hat{\boldsymbol{\phi}}_n$, a saber

$$\hat{\boldsymbol{\beta}} = \begin{pmatrix} 2.451 & 1.390 & 0.548 \end{pmatrix}^\top, \hat{\mathbf{G}} = \begin{pmatrix} 0.536 & 0.569 \\ 0.569 & 0.888 \end{pmatrix}, \hat{\sigma}^2 = 0.404 \quad (3.3)$$

esses valores são considerados como os verdadeiros valores dos parâmetros no processo de simulação.

Uma segunda parte deste mesmo capítulo contempla, a análise de diferentes tipos de erros de especificação em modelos lineares mistos gaussianos, também é exibido o gráfico da função de autocorrelação para detectar uma possível autocorrelação dos erros condicionais. Para isso dependendo do tipo de erro considerado simulamos um conjunto de dados de um modelo particular na Tabela 3.1 e ajustamos um modelo da mesma Tabela. Cada modelo permite estudar um tipo de erro específico, de tal maneira que em cada caso ilustramos os testes para identificar se o modelo está incorretamente especificado, e aplicamos os diagnósticos gráficos para localizar as fontes dos erros. Para o processo de simulação geramos o vetor de dados \mathbf{y}_i , $i = 1, \dots, n$; considerando que cada indivíduo tenha o mesmo número de observações, de tal forma que a estrutura do conjunto de dados \mathbf{y} é balanceada e completa. Além disso, é levando em consideração o poder e a taxa do erro tipo I, usando os tamanhos amostrais $n = 50, 100, 200, 400, 600, 800, 1000$ e 2000 , e um nível de significância 5%.

3.1 Modelos considerados para analisar erros de especificação em modelos lineares mistos gaussianos

Nesta seção são listados na Tabela 3.1 os modelos considerados para a aplicação dos testes e dos diagnósticos gráficos, assim como a análise do poder e da taxa do erro tipo I. Os valores dos parâmetros dos modelos listados foram definidos em (3.3), e para o caso *IV* e *VI* com $\rho=0.9$.

Tabela 3.1: Modelos considerados para o processo de simulação.

Caso	Modelo	\mathbf{G}	\mathbf{R}_i
<i>I</i>	$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_{0i} + b_{1i} x_{ij} + e_{ij}$	$\begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$	$\sigma^2 \mathbf{I}_n$
<i>II</i>	$y_{ij} = \beta_0 + \beta_1 x_{ij} + b_{0i} + b_{1i} x_{ij} + e_{ij}$	$\begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$	$\sigma^2 \mathbf{I}_n$
<i>III</i>	$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_{0i} + e_{ij}$	σ_0^2	$\sigma^2 \mathbf{I}_n$
<i>IV</i>	$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_{0i} + b_{1i} x_{ij} + e_{ij}$	$\begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$	$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{m_i-1} \\ & 1 & \rho & \dots & \rho^{m_i-2} \\ & & 1 & \rho & \vdots \\ & & & \ddots & \rho \\ & & & & 1 \end{pmatrix}, \rho < 1$
<i>V</i>	$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_{0i} + b_{1i} x_{ij} + e_{ij}$	$\begin{pmatrix} \sigma_0^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix}$	$\sigma^2 \mathbf{I}_n$
<i>VI</i>	$y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 x_{ij}^2 + b_{0i} + e_{ij}$	σ_0^2	$\sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{m_i-1} \\ & 1 & \rho & \dots & \rho^{m_i-2} \\ & & 1 & \rho & \vdots \\ & & & \ddots & \rho \\ & & & & 1 \end{pmatrix}, \rho < 1$

3.2 Avaliação de um modelo corretamente especificado

Nesta seção ilustramos um exemplo de um modelo corretamente especificado, para isso simulamos um conjunto de dados do modelo *I*, na Tabela 3.1. Com os dados simulados ajustamos o modelo *I*, sob a hipótese H_0 que o modelo não tem erro de especificação. Na Figura 4.24 são apresentados o comportamento dos perfis individuais simulados das unidades amostrais, o perfil médio e o perfil loess. Visualizamos variabilidades inter e intraunidades amostrais ao longo do tempo. Identificamos possíveis correlações intraunidades amostrais por meio da observação dos perfis dispostos acima ou abaixo do perfil loess e médio.

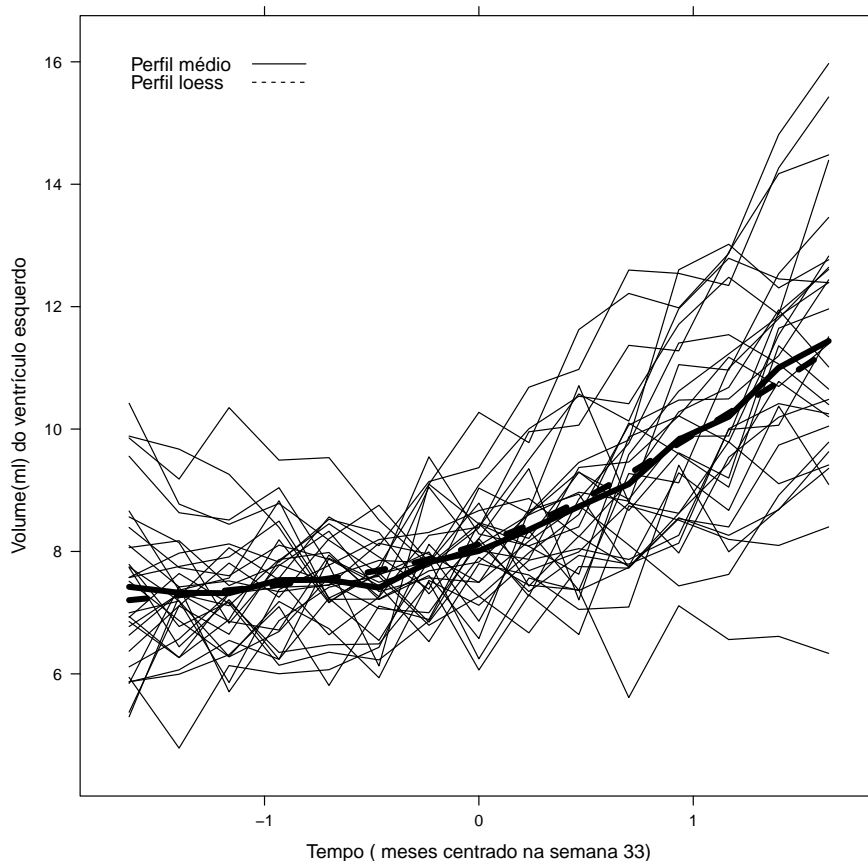


Figura 3.1: Perfis individuais para as 29 unidades amostrais, linha tracejada é o perfil loess, e a linha contínua é o perfil médio dos dados simulados.

Na Tabela 3.2 encontramos os resultados de cada um dos testes. Os testes não detectam erro de especificação no modelo ajustado aos dados simulados, o que já era de esperar, dado que o modelo ajustado aos dados foi o modelo que gero os dados.

Tabela 3.2: Valores das estatísticas dos testes e valores p .

$EAMI$	$valor-p$	$ESDM$	$valor-p$	$EESA$	$valor-p$
0.35	0.99	0.33	0.99	0.10	0.99

Na continuação são apresentados os diagnósticos gráficos. A função de autocorrelação ilustrada na Figura 3.2, não mostra evidência gráfica para suspeitar que os erros condicionais estejam correlacionados serialmente.

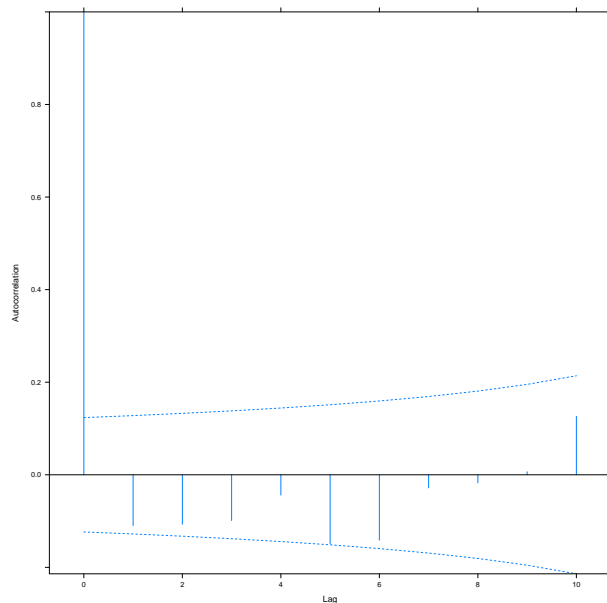


Figura 3.2: Gráfico da função de autocorrelação.

O diagnóstico gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 3.3, não detecta nenhum tipo de violação na suposição da matriz de covariâncias proposta.

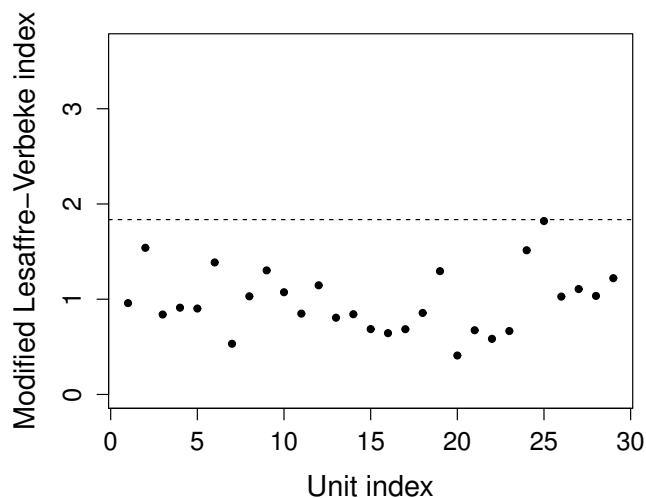


Figura 3.3: Gráfico de índice modificado de Lesaffre - Verbeke. A linha tracejada no gráfico representa 3o quartil + 1.5 distância interquartis.

Pode-se observar na Figura 3.4 que o comportamento dos resíduos marginais padronizados não apresentam tendência para levar à suspeita de alguma violação na estrutura da média.

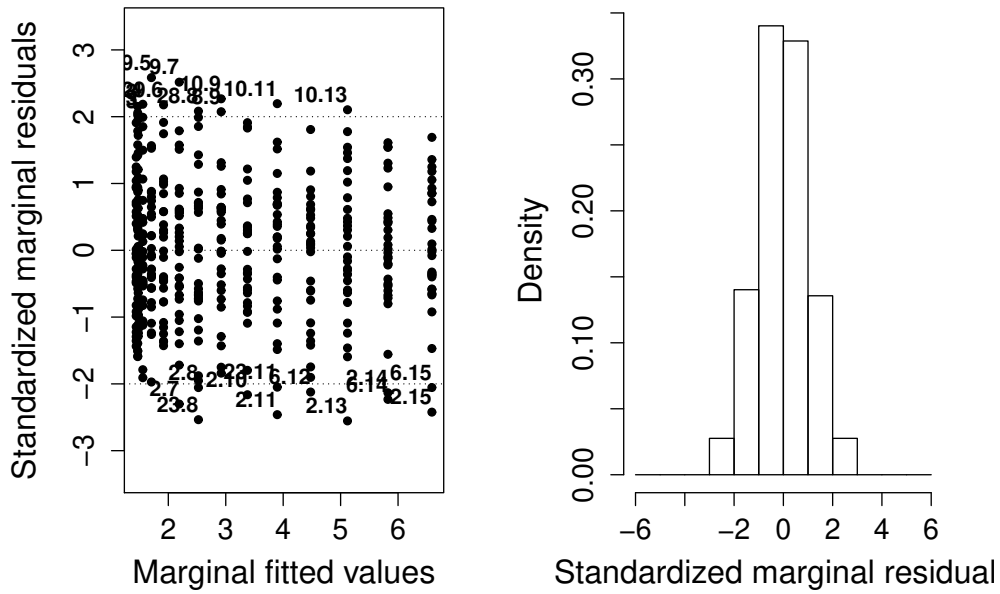


Figura 3.4: Gráfico para os resíduos marginais padronizados.

Pode-se observar na Figura 3.5 que o comportamento dos resíduos padronizados condicionados não apresentam tendência para se suspeitar de alguma violação nos erros condicionais.

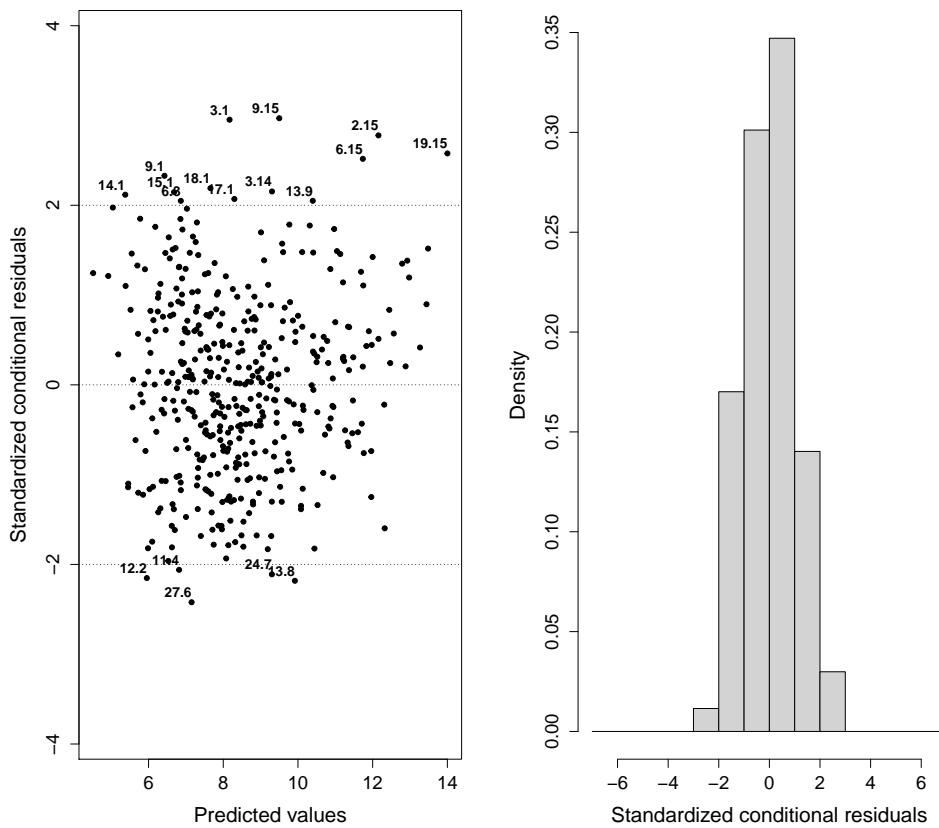


Figura 3.5: Gráfico dos resíduos padronizados condicionados e histograma.

A análise da Figura 3.6 para a distância da Mahalanobis permite concluir que a distribuição

do vetor de efeitos aleatórios não tem afastamento importante de uma distribuição normal.

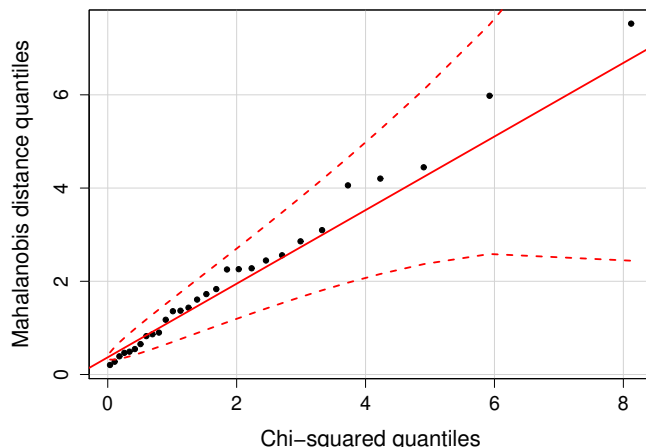


Figura 3.6: Gráfico QQ para a distância da Mahalanobis.

De acordo com os gráficos dos resíduos mínimamente confundidos e a densidade, exibidos na Figura 3.7, conclui-se que a distribuição do erro condicional não tem afastamento de uma distribuição normal.

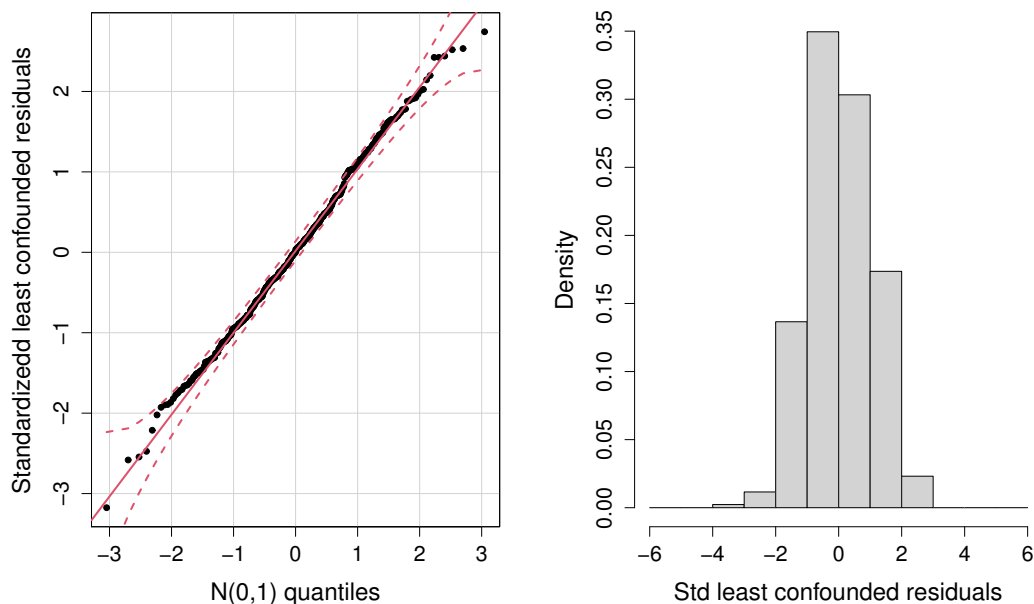


Figura 3.7: QQ plot e histograma para os resíduos mínimamente confundidos padronizados.

Nesta seção foi ilustrado o comportamento de um modelo corretamente especificado. Passaremos a analisar nas seguintes seções erros de especificação com relação à estrutura da

resposta média e com relação à matriz de covariâncias do vetor de respostas.

3.3 Análise quando há presença de erro de especificação na estrutura da resposta média

Neste caso vamos a gerar um único conjunto de dados do modelo I , descrito na Tabela 3.1, considerando os valores de β , \mathbf{G} , e σ^2 indicados em (3.3). O vetor de resposta é tal que

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}(\mathbf{X}_i\beta, \mathbb{V}_i),$$

em que $\mathbb{V}_i = \mathbf{Z}_i\mathbf{G}\mathbf{Z}_i^\top + \mathbf{R}_i$. Com os dados obtidos, ajustamos um modelo similar ao modelo I , mas o termo quadrático é omitido, essa omissão induz ao erro de especificação da parte fixa. O modelo a ajustar é apresentado na linha II da Tabela 3.1.

Na continuação ilustramos os resultados da aplicação de cada teste ao modelo ajustado. Cada teste permite identificar se o modelo apresenta erro de especificação.

3.3.1 Método analítico

Os testes detectam que o modelo tem erro de especificação, como pode ser observado nos resultados apresentados na Tabela 3.3.

Tabela 3.3: Estatísticas de cada teste e valor- p

<i>EAMI</i>	<i>valor-p</i>	<i>ESDM</i>	<i>valor-p</i>	<i>EESA</i>	<i>valor-p</i>
33.07	0.00	21.10	0.00	34.19	0.00

Agora identificamos a fonte do erro de especificação via diagnósticos gráficos.

3.3.2 Método gráfico

A função de autocorrelação é exibida na Figura 3.8 e não mostra evidência de correlação serial nos erros condicionais.

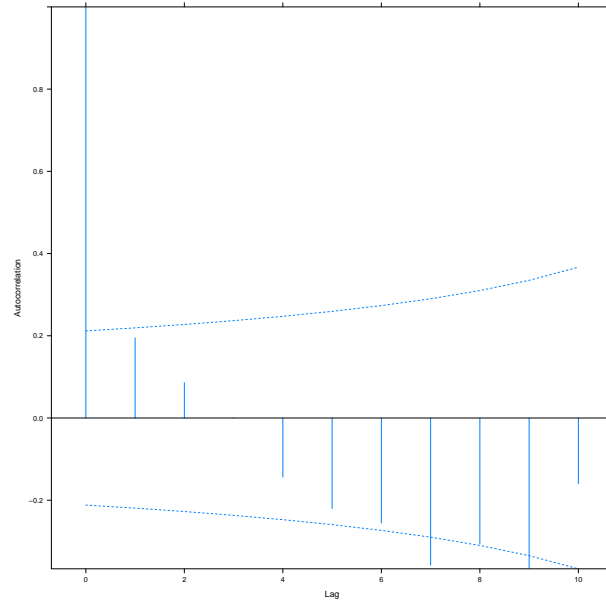


Figura 3.8: *Gráfico da função de autocorrelação*

O gráfico do índice de Lesaffre – Verbeke modificado apresentado na Figura 3.9, não ilustra evidência de algum tipo violação na suposição da matriz de covariâncias proposta.

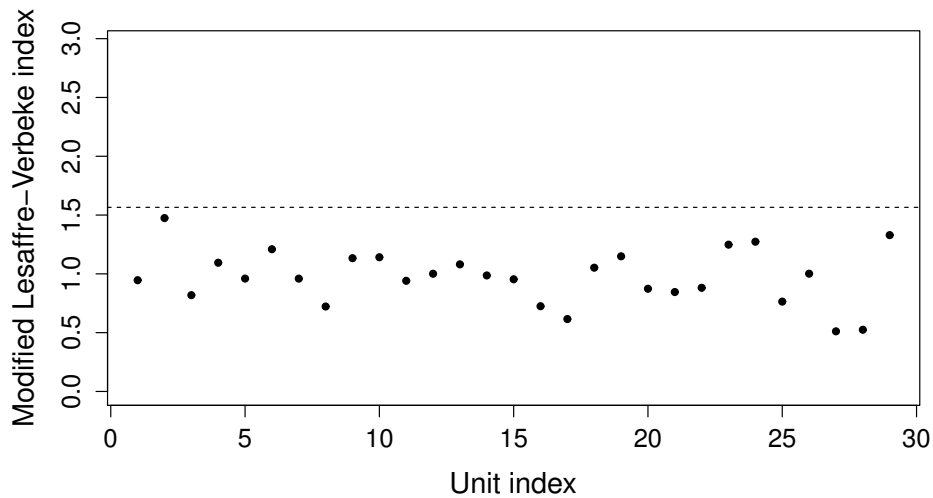


Figura 3.9: *Gráfico do índice modificado de Lesaffre–Verbeke.*

O gráfico para os resíduos marginais padronizados na Figura 3.10, sugere evidência de uma errônea especificação na estrutura da média. O gráfico que apresenta uma curvatura identifica omissão de um termo quadrático na estrutura da média.

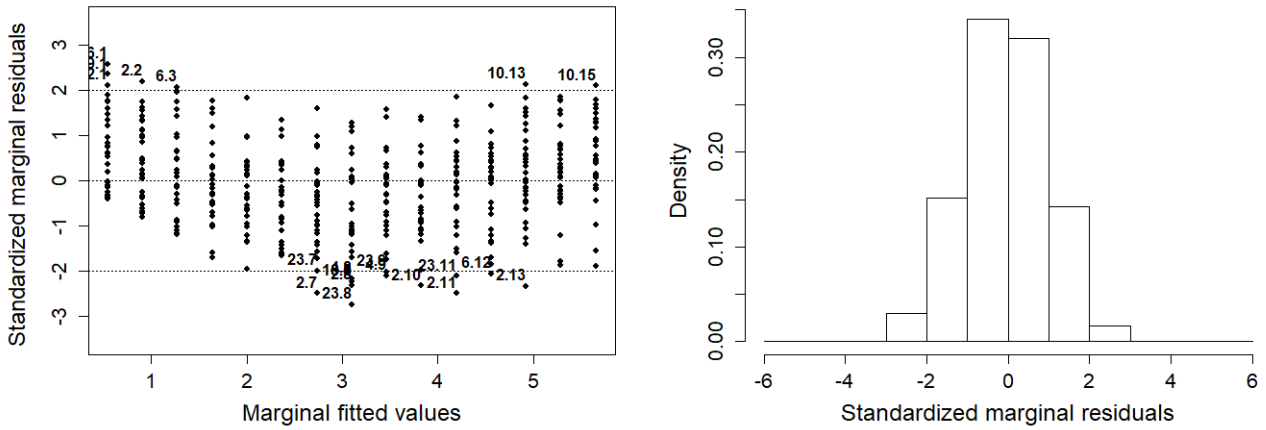


Figura 3.10: Resíduos marginais padronizados.

O gráfico dos resíduos minimamente confundidos padronizados exibidos na Figura 3.11, sugere que o vetor de erros condicionais tem uma estrutura homocedástica.

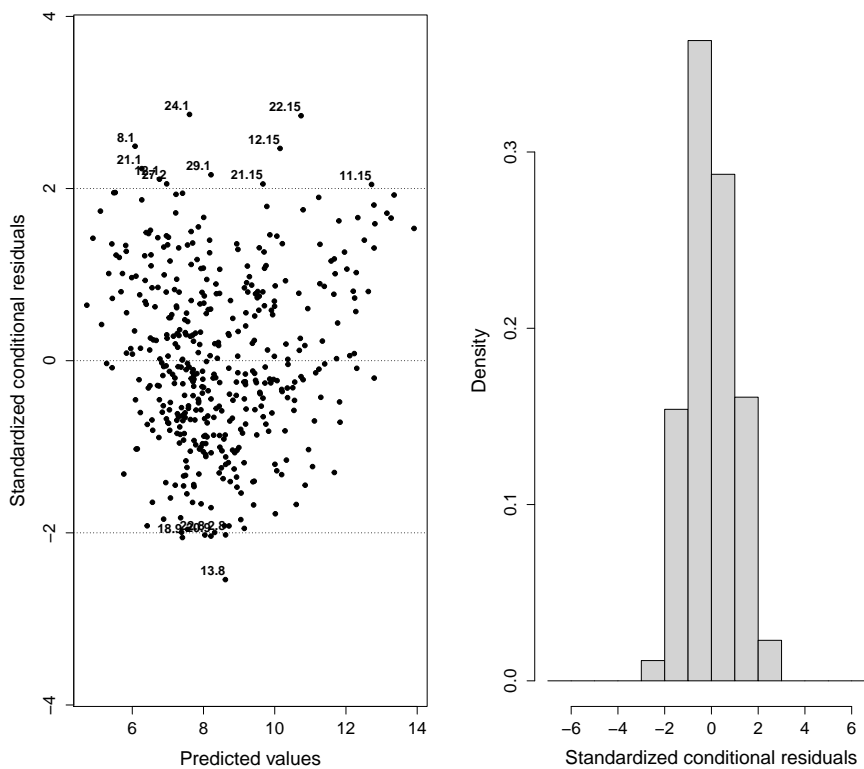


Figura 3.11: QQ plot e histograma para os resíduos minimamente confundidos padronizados.

Na Figura 3.12 o gráfico da distância dos quantis de Mahalanobis é exposto, não exibe evidência contra o comportamento gaussiano do vetor de efeitos aleatórios, e a Figura 3.13 não mostra evidência para suspeitar da normalidade para o erro condicional.

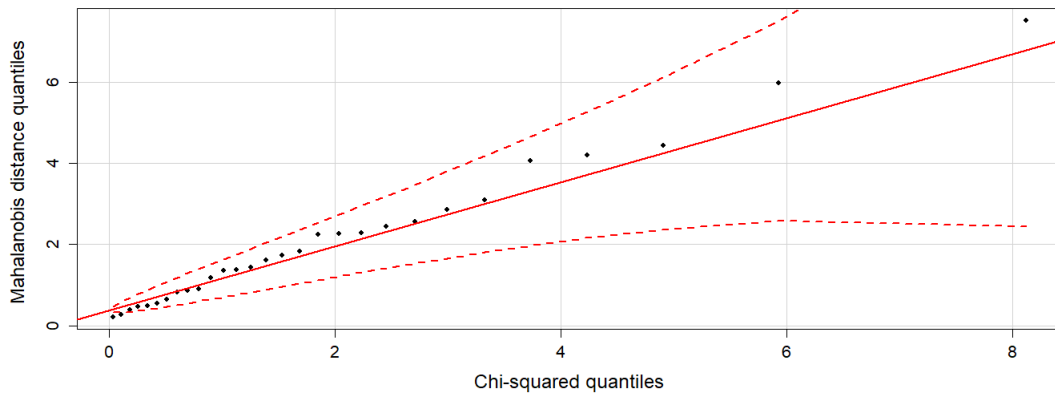


Figura 3.12: Gráfico da distância dos quantis de Mahalanobis.

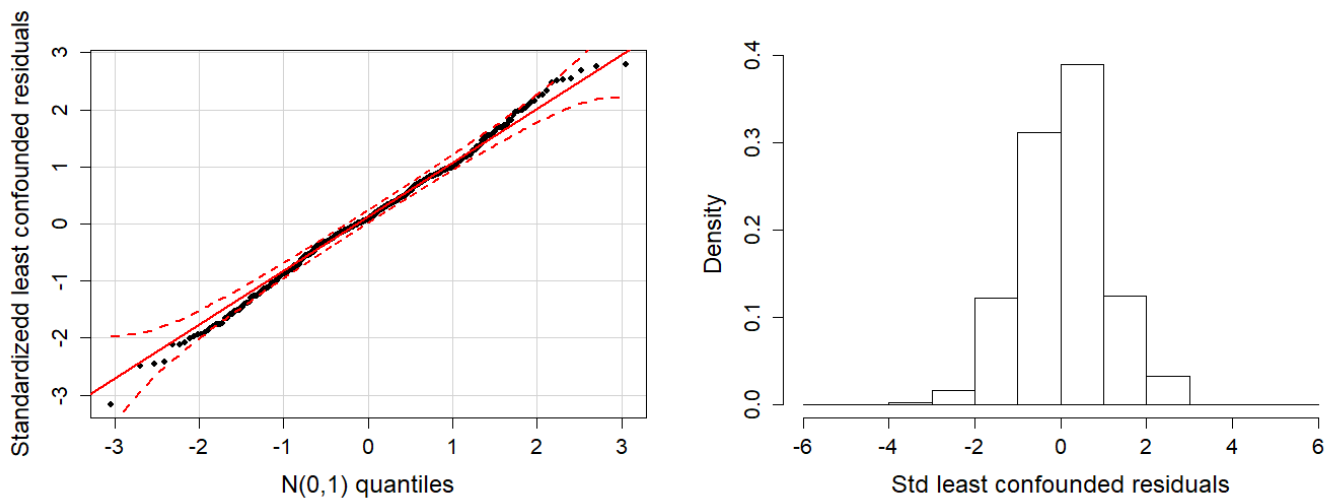


Figura 3.13: QQ plot e histograma para os resíduos minimamente confundidos padronizados.

Nesta seção foi analisado erro de especificação na estrutura da média. Na Tabela 3.4 são apresentados os resultados dos testes sob 1000 iterações quando é analisado, via simulação o poder e a taxa do erro tipo I.

Tabela 3.4: Poder e taxa do erro tipo I ($E.tipoI$) para os testes: alternativo da matriz de informação (I), do estimador “sanduíche” modificado (II) e do estimador “sanduíche” alternativo (III).

n	<i>I</i>		<i>II</i>		<i>III</i>	
	<i>Poder</i>	<i>E.tipoI</i>	<i>Poder</i>	<i>(E.tipoI)</i>	<i>Poder</i>	<i>E.tipoI</i>
50	0.18	0.25	0.84	0.42	0.83	0.32
100	0.19	0.18	0.94	0.25	0.88	0.22
200	0.26	0.12	0.99	0.18	0.90	0.19
400	0.43	0.11	1.00	0.12	1.00	0.11
600	0.54	0.08	1.00	0.10	1.00	0.09
800	0.84	0.07	1.00	0.09	1.00	0.08
1000	0.88	0.06	1.00	0.07	1.00	0.06
2000	0.99	0.05	1.00	0.05	1.00	0.05

Segundo os resultados na Tabela 3.4, o teste *I* tem baixo poder mesmo para amostras grandes $n = 600$. Comparando os resultado do teste *II* com o teste *III*, *III* apresenta melhor comportamento na taxa do erro na medida que aumentam os tamanhos da amostra, a partir da amostra de tamanho 400, o poder não apresenta diferenças.

3.4 Análise quando há presença de erro de especificação na estrutura da matriz de covariâncias do vetor de respostas

Nesta seção estudamos a situação em que o modelo apresenta erro de especificação na matriz de covariâncias do vetor de respostas, concretamente estudamos especificação incorreta em: o número de efeitos aleatórios. Neste caso há uma omissão na inclinação aleatória, os efeitos aleatórios são considerados independentes, quando em verdade não são independentes, os erros aleatórios seguem um processo autorregressivo de ordem 1, AR(1). O erro está na matriz de covariâncias intraunidades amostrais e finalmente consideramos o caso em que o número de efeitos aleatórios está errado e os erros aleatórios seguem um processo AR(1).

3.4.1 Há presença de erro de especificação no número de efeitos aleatórios

Nesta seção simulamos um conjunto de dados do modelo I , apresentado na linha I da Tabela 3.1, supondo que a estrutura da resposta média está representada por um intercepto, inclinação e efeito quadrático, o componente aleatório está representado por um intercepto e uma inclinação com \mathbf{X}_i , \mathbf{Z}_i definidos como na especificação para (3.1) e $\boldsymbol{\beta}$, \mathbf{G} e σ^2 como em (3.3). Supõe-se independência entre \mathbf{b}_i e \mathbf{e}_i , $\mathbf{b}_i \sim \mathcal{N}_2[\mathbf{0}, \mathbf{G}]$, $\mathbf{e}_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i]$, $i = 1, \dots, n$. O vetor \mathbf{y}_i é tal que

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}(\mathbf{X}_i\boldsymbol{\beta}, \mathbb{V}_i).$$

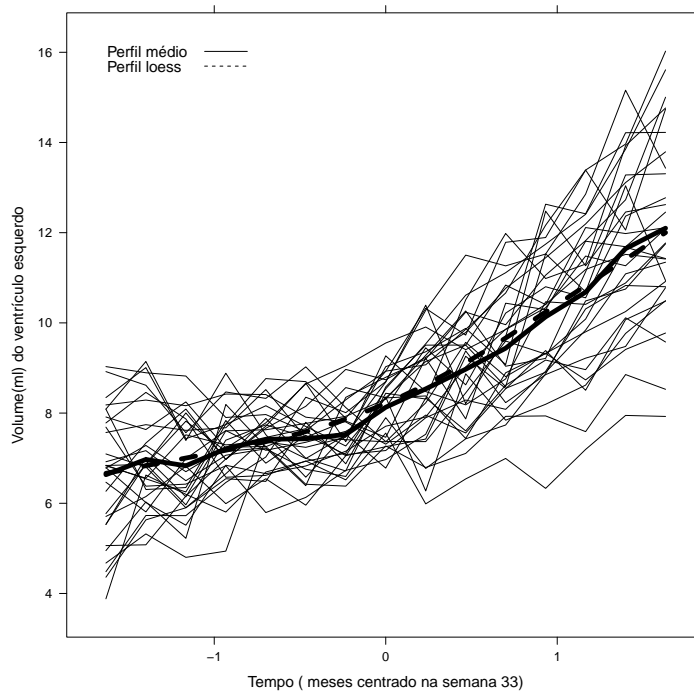


Figura 3.14: Comportamento dos perfis individuais simulados ao redor do perfil médio e do perfil loess.

Na Figura 3.14 ilustramos os perfis simulados de resposta, o perfil médio e o perfil loess. Os perfis permitem perceber um crescimento da variância com o tempo.

Com os dados simulados é ajustado o modelo III , apresentado na Tabela 3.1, linha III . Neste caso omitimos a inclinação aleatória que induz o erro de especificação no número de efeitos aleatórios, assim a matriz de covariâncias gerada por esse modelo induz uma estrutura

uniforme, em que as variâncias são constantes ao longo do tempo.

Para a análise do poder nesta situação, simulamos um único conjunto de dados do modelo I , e ajustamos o modelo III . Para a análise da taxa do erro tipo I , simulamos do modelo III e ajustamos o modelo III .

Na sequência são ilustrados os resultados do comportamento dos testes para identificar que o modelo apresenta erro de especificação.

3.4.1.1 Método analítico

Os testes indicam que o modelo está incorretamente especificado, tal como é observado na Tabela 3.5. Em qualquer dos casos, os testes detectam que o modelo tem erro de especificação.

Tabela 3.5: Valores das estatísticas dos testes e valores p .

$EAMI$	$valor-p$	$ESDM$	$valor-p$	$EESA$	$valor-p$
1485.81	0.00	135.27	0.00	12.29	0.03

Ao analisar o resultado gráfico, na Figura 3.15 pode-se observar que a estrutura para a matriz de covariâncias, pode não ser apropriada al menos para os indivíduos, #2 e #6.

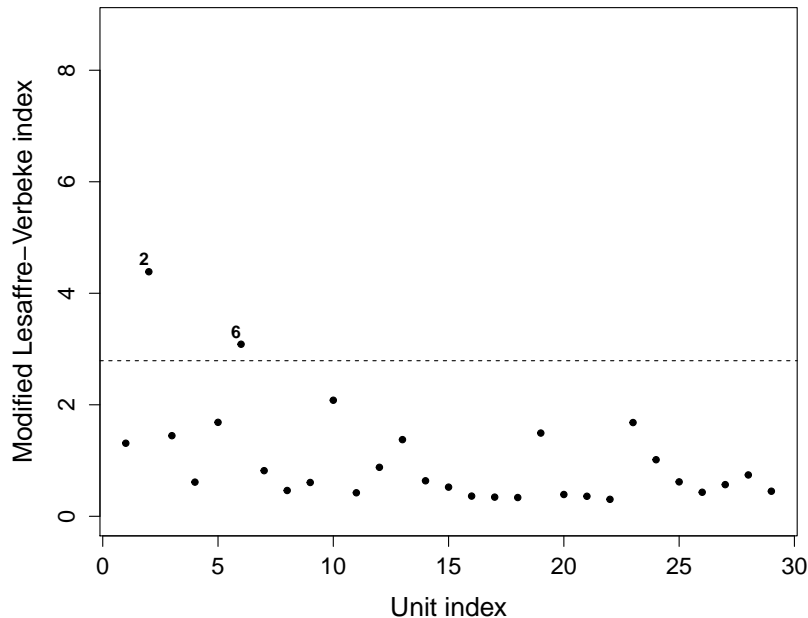


Figura 3.15: Gráfico de índice modificado de Lesaffre - Verbeke.

Na Figura 3.16 do gráfico dos resíduos marginais padronizados. Não é percebido uma

violação na suposição da estrutura da média no modelo ajustado. Não obstante no gráfico dos resíduos marginais padronizados é amostrado a existência de heterocedasticidade na distribuição do erro aleatório. A variabilidade aumenta com tempo indicando a ausência do efeito aleatório associado ao coeficiente angular.

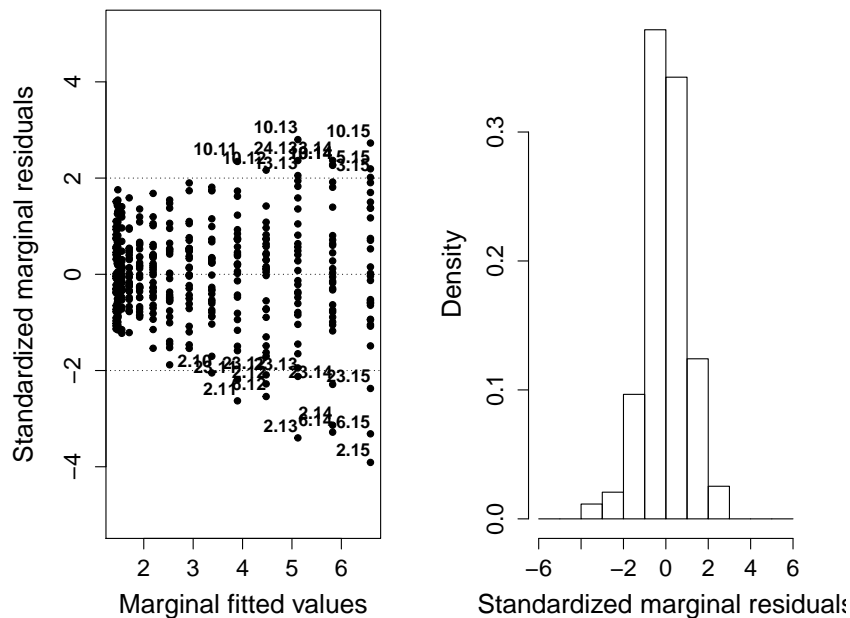


Figura 3.16: *Gráfico dos resíduos marginais padronizados*

Na Figura 3.17 o gráfico da distancia da malahanobis é exibido, nesse gráfico não há suspeita que o erro aleatório não tenha uma distribuição normal e na Figura 3.18 é percebido algumas observações influentes, por exemplos, #2, #10 e #23.

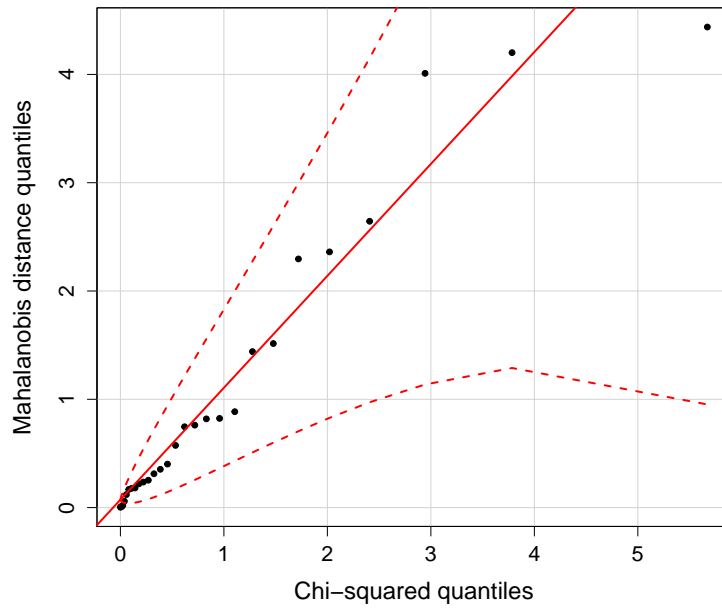


Figura 3.17: Gráfico QQ para a distancia de Mahalanobis.

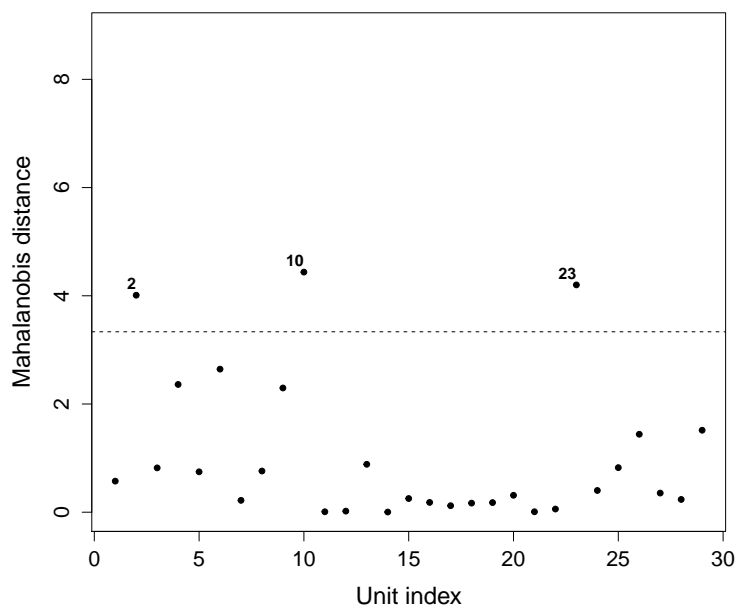


Figura 3.18: Gráfico da distancia de Mahalanobis.

Na Figura 3.19 o gráfico QQ para os resíduos minimamente confundidos padronizados é exibido, esse gráfico não ilustra suficiente informação sob alguma condição para suspeitar violação na suposição da distribuição normal para o erro condicional.

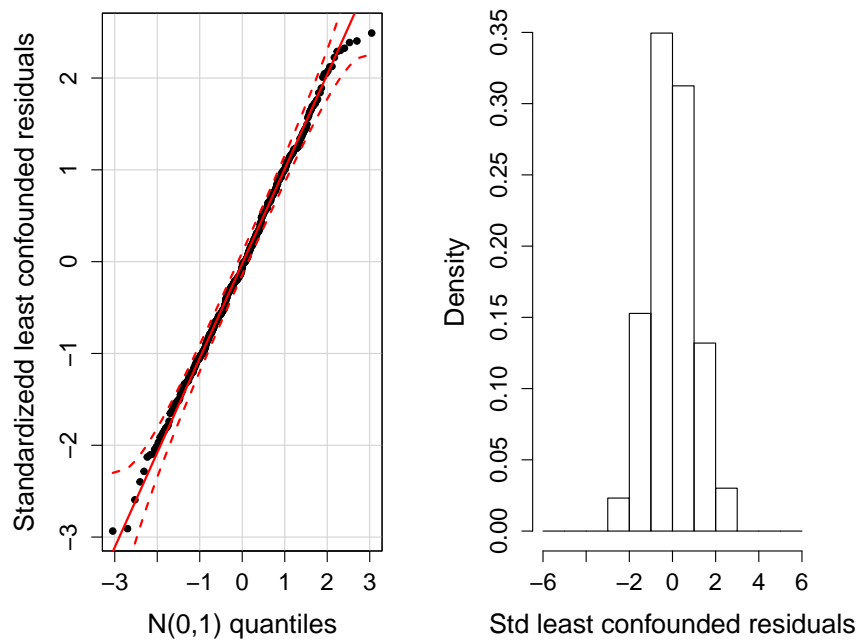


Figura 3.19: Gráfico *QQ* e histograma para os resíduos minimamente confundidos padronizados.

O gráfico dos resíduos padronizados apresentados na Figura 3.20 não ilustra evidência forte que indique que o erro condicional não tenha uma distribuição normal. Além do anterior, é observado que a observação 15 do indivíduo 2, #2.15, a observação 15 do indivíduo 6, #6.15, a observação 15 do indivíduo 10, #10.15 e a observação 13 do indivíduo 10, podem ser consideradas em destaque, possíveis valores atípicos com respeito às observações restantes.

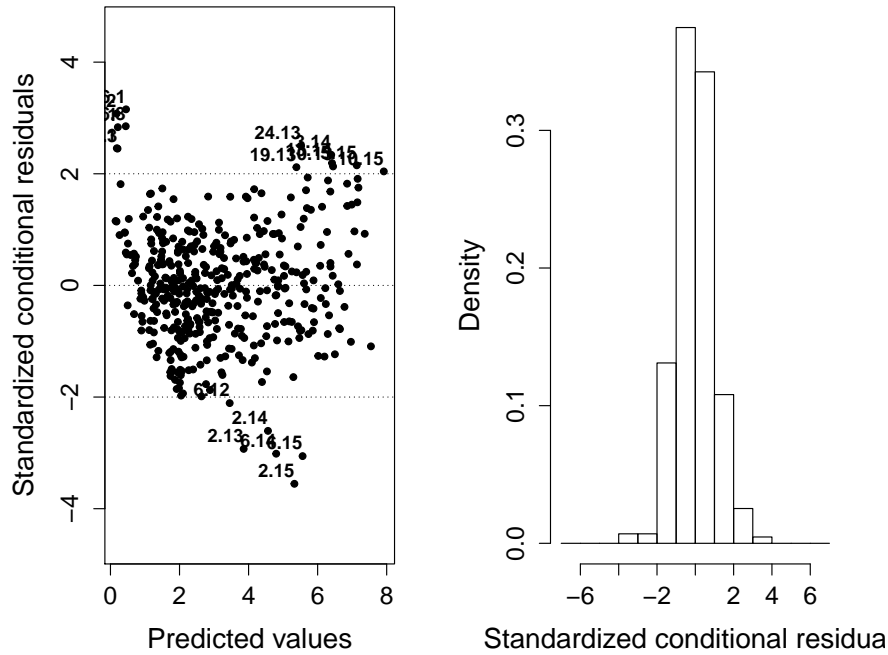


Figura 3.20: Gráfico dos resíduos condicionados padronizados.

Análise do poder e a taxa do erro tipo I.

Tabela 3.6: Poder e taxa do erro tipo I ($E.tipoI$) para os testes: da matriz de informação alternativo (I), do estimador “sanduíche” modificado (II) e do estimador “sanduíche” alternativo (III).

n	I		II		III	
	Poder	$E.tipoI$	Poder	$E.tipoI$	Poder	$E.tipoI$
50	0.85	0.42	0.92	0.33	0.35	0.41
100	0.98	0.21	0.99	0.21	0.39	0.30
200	1.00	0.18	1.00	0.19	0.87	0.23
400	1.00	0.11	1.00	0.10	0.91	0.18
600	1.00	0.09	1.00	0.08	0.95	0.10
1000	1.00	0.07	1.00	0.06	1.00	0.07
2000	1.00	0.06	1.00	0.05	1.00	0.05

Embora para os tamanhos da amostra entre 50 e 200 o poder dos testes do estimador “sanduíche” alternativo seja baixo comparado com os outros testes, para amostras superiores ou iguais a 400 não tem muita diferenças.

No caso da taxa do erro tipo I o comportamento não difere significativamente.

3.4.2 Há presença de erro de especificação na matriz de covariâncias do vetor de efeitos aleatórios, os efeitos aleatórios são considerados independentes

Neste caso vamos simular um único conjunto de dados do modelo apresentado na linha I da Tabela 3.1, esse modelo tem na parte fixa intercepto, inclinação e efeito quadrático e na parte aleatória intercepto e inclinação. Um modelo de independência condicional homocedástico, admitindo que os efeitos aleatórios estão correlacionados. Supõe-se independência entre \mathbf{b}_i e \mathbf{e}_i , $\mathbf{b}_i \sim \mathcal{N}_2[\mathbf{0}, \mathbf{G}]$, $\mathbf{e}_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i]$, $i = 1, \dots, n$. O vetor de dados \mathbf{y}_i é tal que

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}(\mathbf{X}_i\boldsymbol{\beta}, \mathbb{V}_i).$$

Com os dados simulados ajustou-se um modelo similar ao anterior, sob a hipótese H_0 de que o modelo ajustado não tem erro de especificação, considerando os efeitos aleatórios independentes, o modelo descrito está na linha V da Tabela 3.1.

Para o poder, simulamos do modelo I e ajustamos o modelo V e para a taxa do erro tipo I , simulamos modelo V e ajustamos o modelo V .

Na subseção 3.4.2.1 usamos os testes para identificar que o modelo apresenta erro de especificação.

3.4.2.1 Método analítico

Analisando a Tabela 3.7, os resultados dos testes indicam que o modelo apresenta erro de especificação. Em todos os casos cada um dos testes rejeitam a hipótese H_0 .

Tabela 3.7: Estatística de cada teste, valor- p com nível de significância 5%.

$EAMI$	valor- p	$ESDM$	valor- p	$EESA$	valor- p
18.26	0.01	1312.15	0.00	156.11	0.00

Na seguinte seção vamos a usar os diagnósticos gráficos para identificar a fonte do erro de especificação.

3.4.2.2 Método gráfico

Analisando a função de autocorrelação apresentada na Figura 3.21, não há suficiente evidência gráfica para concluir que o modelo ajustado tenha correlação serial nos erros condicionais.

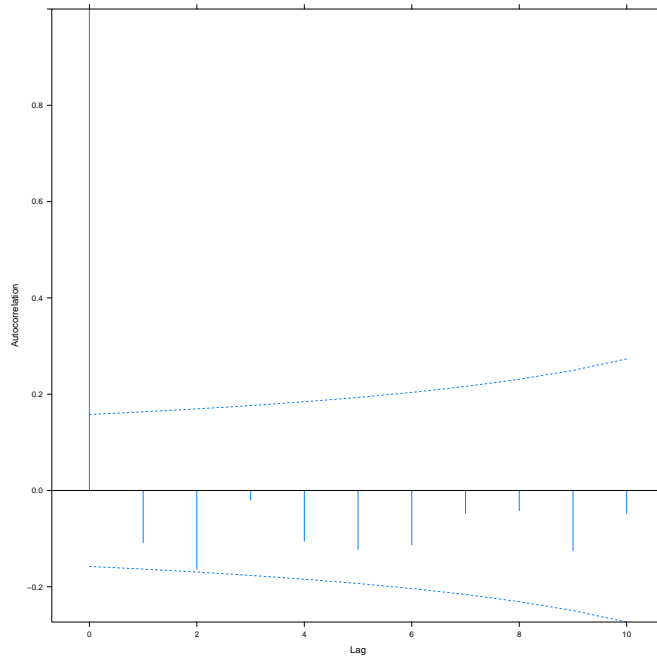


Figura 3.21: Gráfico da função de autocorrelação.

Na Figura 3.22 o gráfico do índice de Lesaffre – Verbeke modificado indica que pelo menos, para os indivíduos #2 e #25, a matriz de covariâncias proposta não é adequada.

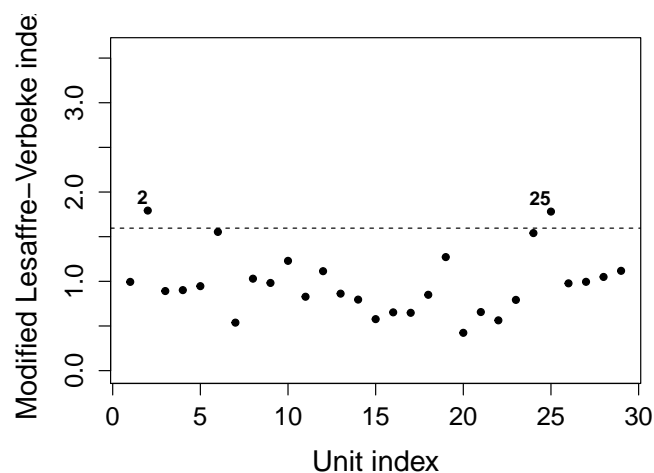


Figura 3.22: Gráfico do índice modificado de Lesaffre–Verbeke.

Analisando o gráfico dos resíduos marginais padronizados apresentado na Figura 3.23, é observada uma variabilidade que aumenta com o tempo. Essa variabilidade possivelmente é

devida à incorreta especificação na estrutura da matriz de covariâncias do vetor de efeitos aleatórios no modelo ajustado.

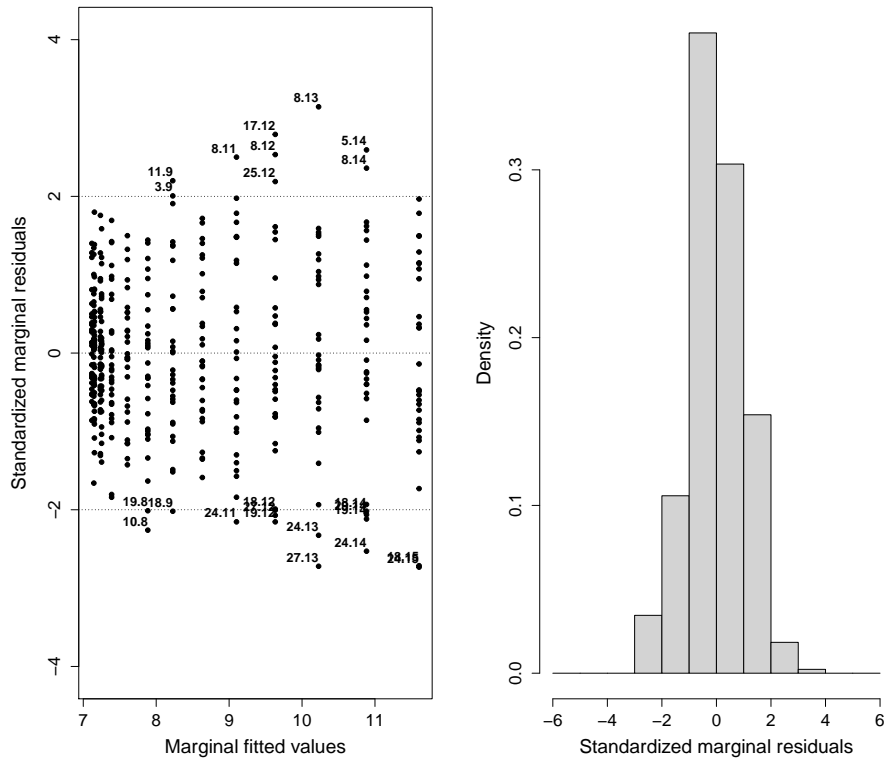


Figura 3.23: Gráfico dos resíduos marginais padronizados.

A análise da Figura 3.24, sugere que as observações #2, #6, #10 e #23 podem ser analisadas como outliers, comparadas com o resto.

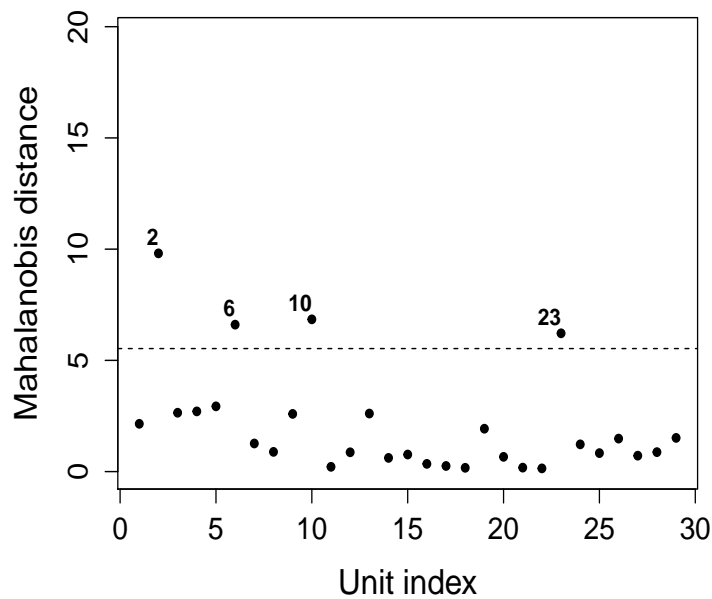


Figura 3.24: Gráfico da distância de Mahalanobis.

Os resultados apresentados na Figura 3.25 permitem evidenciar que a distribuição do vetor de efeitos aleatórios não tem um afastamento da distribuição normal.

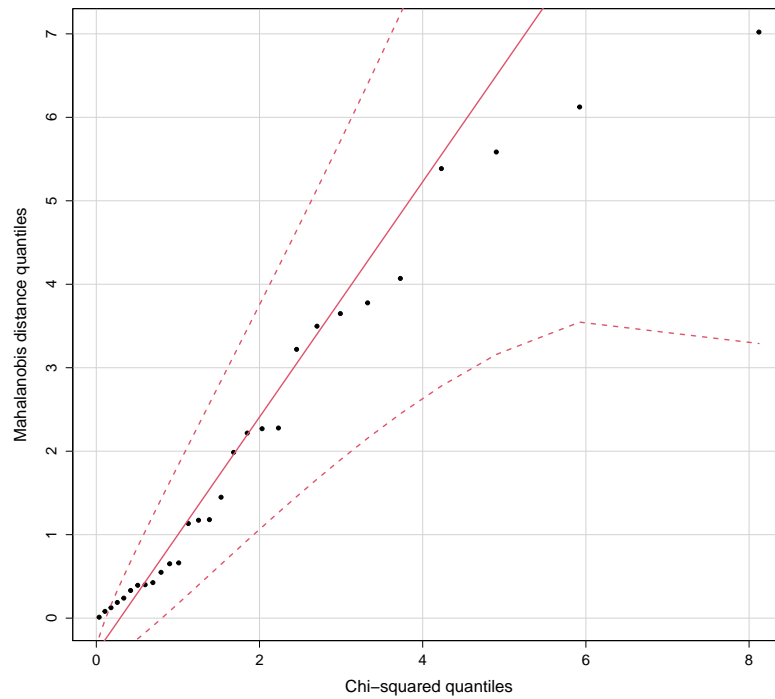


Figura 3.25: Gráfico da distância dos quantis de Mahalanobis.

Pela Figura 3.26 é percebido plausibilidade da suposição de normalidade para a distribuição nos erros condicionais.

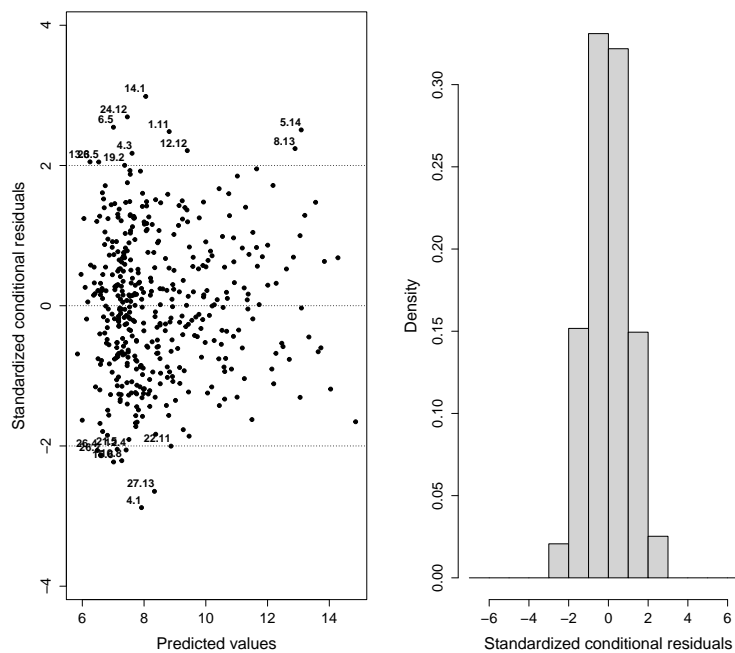


Figura 3.26: Gráfico dos resíduos condicionais padronizados.

Analisando a Figura 3.27, evidenciamos que a distribuição do vetor de erros aleatórios não tem um afastamento da distribuição normal.

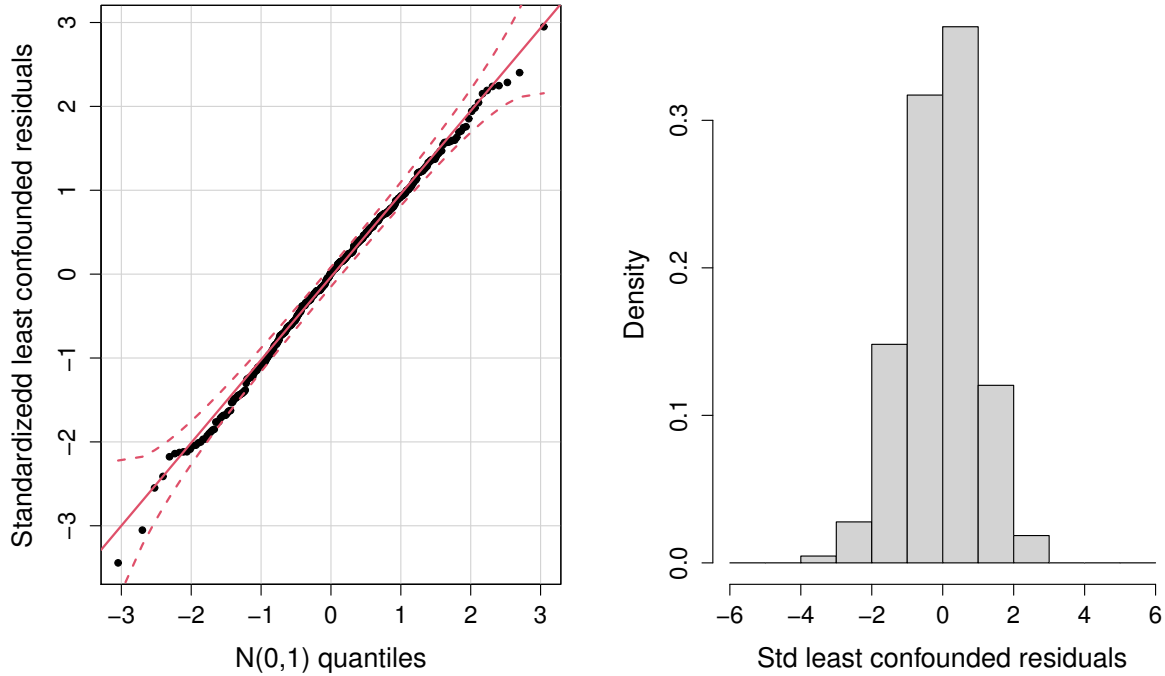


Figura 3.27: Gráfico *QQ* para os resíduos condicionais padronizados e histograma.

Agora vamos analisar o comportamento dos testes considerando o poder e a taxa do erro tipo *I*.

Tabela 3.8: Poder e taxa do erro tipo *I* ($E.tipoI$) para os testes: da matriz de informação alternativo (*I*), do estimador “sanduíche” modificado (*II*) e do estimador “sanduíche” alternativo (*III*).

n	<i>I</i>		<i>II</i>		<i>III</i>	
	Poder	$(E.tipoI)$	Poder	$(E.tipoI)$	Poder	$(E.tipoI)$
50	1.00	0.36	0.94	0.69	0.96	0.43
100	1.00	0.21	0.95	0.65	0.98	0.31
200	1.00	0.17	0.97	0.59	0.99	0.20
400	1.00	0.10	0.98	0.45	1.00	0.10
600	1.00	0.09	0.99	0.32	1.00	0.08
800	1.00	0.07	1.00	0.22	1.00	0.07
1000	1.00	0.06	1.00	0.10	1.00	0.06
2000	1.00	0.05	1.00	0.06	1.00	0.05

O poder mais alto e a taxa mais baixa foi determinado pelo teste da matriz de informação alternativo. Nessa ordem das ideias, o teste do estimador “sanduíche” modificado apresenta

melhor comportamento do poder comparado com o teste do estimador “sanduíche” alternativo.

3.4.3 Há presença de erro de especificação na estrutura da matriz de covariâncias intraunidades amostrais

Nesta seção com o fim de avaliar erro de especificação na estrutura da matriz de covariâncias no vetor de erros aleatórios num modelo linear misto gaussiano, usando os testes e os diagnósticos gráficos descritos no Capítulo 2, primeiramente simulamos um conjunto de dados de um modelo que tem na parte fixa inclinação, intercepto e efeito quadrático, na parte aleatória um polinômio de grau um com intercepto. Os erros aleatórios seguem um processo AR(1), $\rho = 0.9$, o modelo descrito está na linha IV da Tabela 3.1. Posteriormente, com os dados simulados é ajustado um modelo que tem na parte fixa inclinação, intercepto e efeito quadrático, na parte aleatória, intercepto e inclinação, é considerado um modelo de independência condicional, sob a hipótese H_0 que o modelo não tem erro de especificação. Este modelo é descrito na linha I da Tabela 3.1.

Para o poder, simulamos do modelo IV e ajustamos o modelo I, para a análise da taxa do erro tipo I, simulamos do modelo IV e ajustamos o modelo IV.

Na sequência ilustramos os resultados dos testes que permitem identificar se o modelo apresenta erro de especificação.

3.4.3.1 Método analítico

Os testes detectam que o modelo apresenta erro de especificação e, é suficiente analisar a Tabela 3.9 para verificar o resultado.

Tabela 3.9: Estatísticas de cada teste e valor- p , 5% significância.

<i>EAMI</i>	<i>valor-p</i>	<i>ESDM</i>	<i>valor-p</i>	<i>EESA</i>	<i>valor-p</i>
64.11	0.00	620.70	0.00	1233.40	0.00

Uma vez identificado que o modelo apresenta erro de especificação, usamos os diagnósticos gráficos para localizar a fonte do erro.

3.4.3.2 Método gráfico

Analisando o comportamento da função de autocorrelação apresentada na Figura 3.28, há suspeita para indicar que o vetor de erros aleatórios condicionais tenha uma estrutura AR(1).

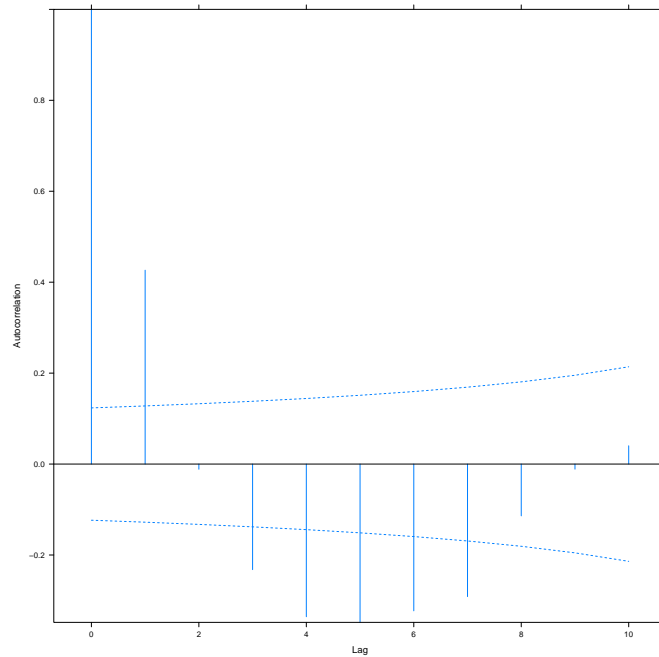


Figura 3.28: Gráfico da função de autocorrelação.

Na Figura 3.29 o gráfico do índice modificado de Lesaffre–Verbeke é analisado. Neste gráfico é observado que, pelo menos, para as unidades #2, #17 e #22 a matriz de covariâncias proposta é inapropriada.

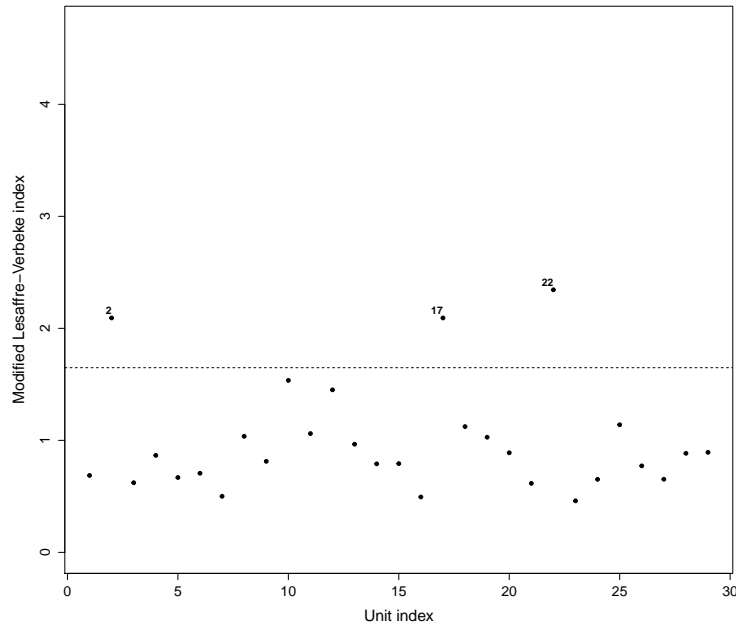


Figura 3.29: Gráfico do índice modificado de Lesaffre–Verbeke.

Na Figura 3.30 é analisado o gráfico dos resíduos marginais padronizados; na qual não se observa evidência gráfica para suspeitar que o modelo em estudo tenha erro de especificação na estrutura da média.

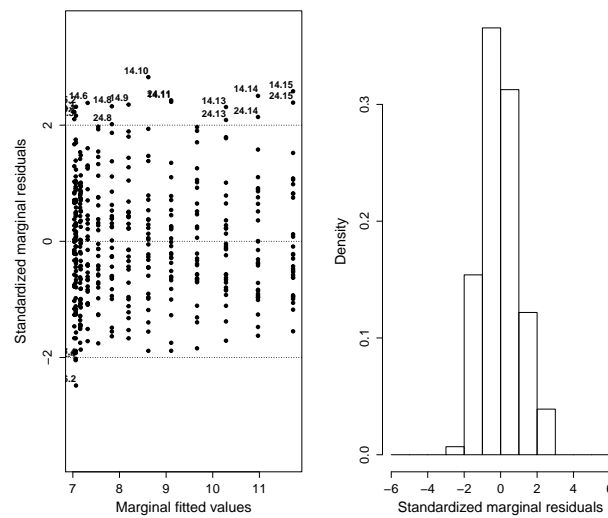


Figura 3.30: Gráfico dos resíduos marginais padronizados

Na Figura 3.31, observa-se que não existe evidência gráfica para suspeitar que o vetor de erros aleatórios não tenha uma distribuição normal. Por outro lado, na Figura 3.32 não há evidência para suspeitar de alguma violação nos erros condicionais.

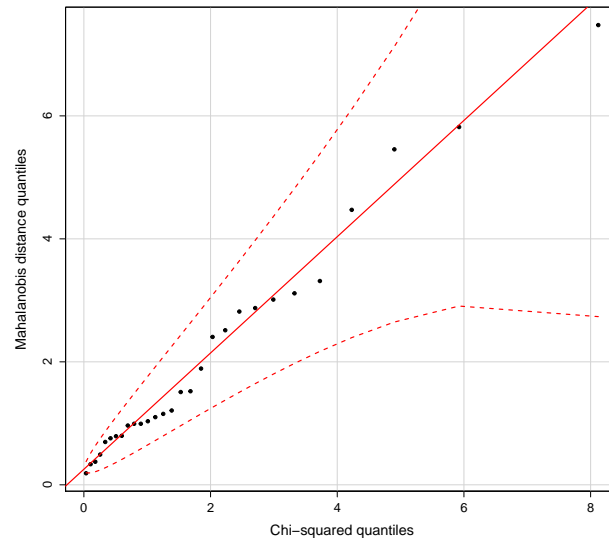


Figura 3.31: Gráfico da distância dos quantis de Mahalanobis.

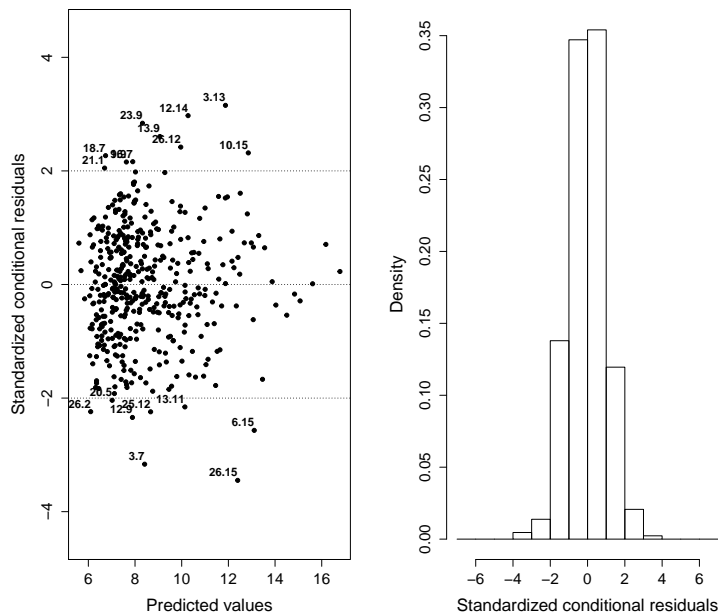


Figura 3.32: Gráfico dos resíduos condicionais padronizados.

Na Figura 3.33, são analisados os resíduos minimamente confundidos padronizados. A Figura não ilustra evidência para suspeitar que os erros aleatórios condicionais não tenha uma distribuição normal.

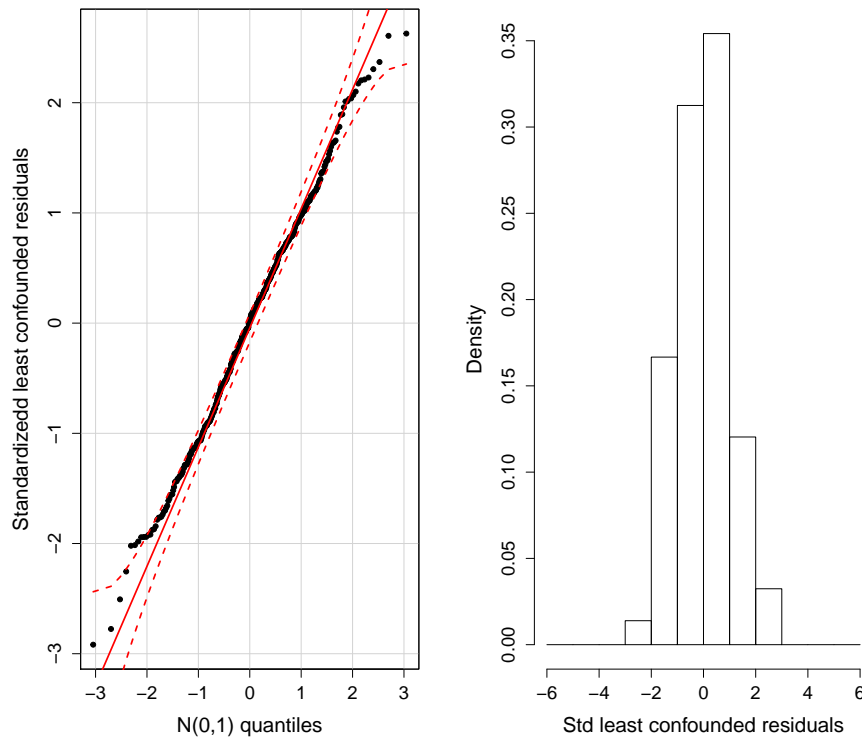


Figura 3.33: *QQ plot e histograma para os resíduos minimamente confundidos padronizados.*

Na continuação são analisados os resultados do poder e da taxa do erro tipo I.

Tabela 3.10: *Poder e taxa do erro tipo I (E.tipoI) para os testes: da matriz de informação alternativo (I), do estimador “sanduíche” modificado (II) e do estimador “sanduíche” alternativo (III).*

n	<i>I</i>		<i>II</i>		<i>III</i>	
	<i>Poder</i>	<i>E.tipoI</i>	<i>Poder</i>	<i>E.tipoI</i>	<i>Poder</i>	<i>E.tipoI</i>
50	0.99	0.48	1.00	0.37	0.93	0.36
100	1.00	0.34	1.00	0.21	0.95	0.32
200	1.00	0.28	1.00	0.19	0.96	0.21
400	1.00	0.21	1.00	0.11	0.97	0.12
600	1.00	0.14	1.00	0.11	0.98	0.10
800	1.00	0.10	1.00	0.09	1.00	0.09
1000	1.00	0.09	1.00	0.08	1.00	0.07
2000	1.00	0.06	1.00	0.05	1.00	0.05

Analisando o resultado do poder dos testes, não diferem notavelmente. Com relação à taxa do erro tipo I, o teste *II* exibe um resultado menor que os outros testes, não obstante, em geral a taxa tende para o nível de significância em todos os casos.

3.4.4 Há presença de erro de especificação no número de efeitos aleatórios e na matriz de covariâncias intraunidades amostrais

Neste caso simulamos dados de um modelo que tem na parte fixa intercepto, inclinação e efeito quadrático, na parte aleatória só tem um intercepto. Os erros aleatórios seguem um processo AR(1). Na Tabela 3.1 são ilustradas as estruturas das matrizes do vetor de efeitos aleatórios e do vetor de erros aleatórios, $b_i \sim \mathcal{N}[0, \sigma_0^2]$, $e_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i]$, $i = 1, \dots, n$. O vetor \mathbf{y}_i é tal que

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}(\mathbf{X}_i\boldsymbol{\beta}, \mathbb{V}_i),$$

o modelo descrito é apresentado na linha VI dessa Tabela. Com os dados simulados ajustamos um modelo que apresenta na parte fixa intercepto, inclinação e efeito quadrático, na parte aleatória um intercepto e uma inclinação, considerando que os efeitos aleatórios estão correlacionados e que o modelo é de independência condicional homocedástico. O modelo descrito é indicado na linha I da Tabela.

Para a análise do poder simulamos do modelo VI e ajustamos o modelo I, e para a taxa do erro tipo I, simulamos do modelo VI e ajustamos o modelo VI.

Na sequência usamos os testes para identificar que o modelo apresenta erros de especificação.

3.4.4.1 Método analítico

Os testes detectam que o modelo tem erro de especificação, ver Tabela 3.11.

Tabela 3.11: Valores das estatísticas dos testes e valores p .

<i>EAMI</i>	<i>valor-p</i>	<i>ESDM</i>	<i>valor-p</i>	<i>EESA</i>	<i>valor-p</i>
87.80	0.00	35.42	0.00	32.20	0.00

Agora vamos a localizar onde estão as fontes dos erros de especificação.

3.4.4.2 Método gráfico

Analisando a função de autocorrelação na Figura 3.34, é notado que existe evidência para concluir que os erros aleatórios apresentam uma estrutura autorregressiva.

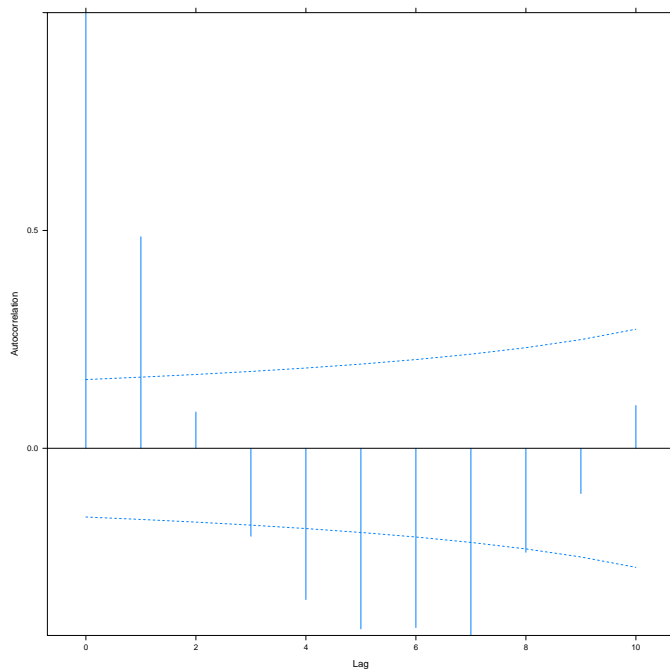


Figura 3.34: Gráfico da função de autocorrelação.

Ao analisar o resultado na Figura 3.35, gráfico do índice modificado de Lesaffre–Verbeke, pode-se observar que a matriz de covariâncias proposta, pode não ser a mais apropriada, pelo menos, para os indivíduos #6, #23 e #26.

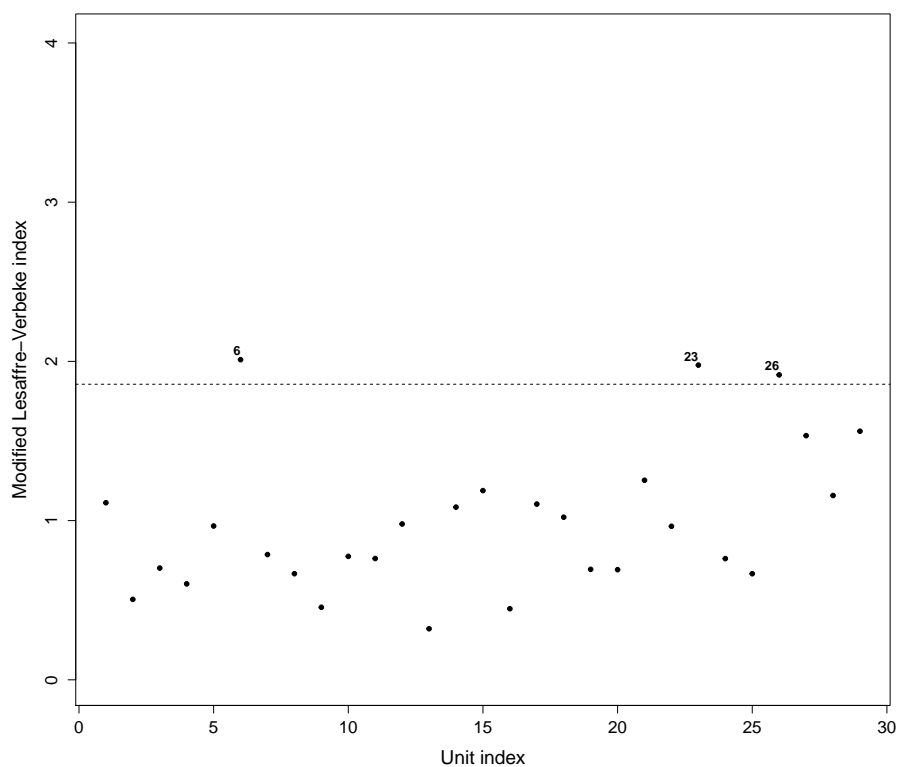


Figura 3.35: Gráfico do índice modificado de Lesaffre–Verbeke.

Na Figura 3.36, são analisados os resíduos marginais padronizados. Não é percebido uma violação na suposição da estrutura da média.

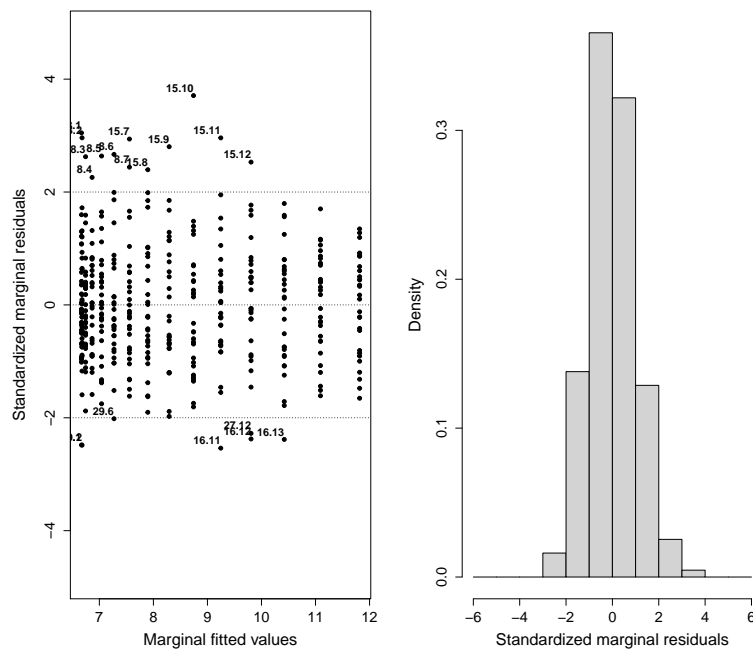


Figura 3.36: *Gráfico dos resíduos marginais padronizados e histograma.*

Tendo em conta a Figura 3.37, é captado que existem algumas observações outliers, por exemplo, as observações #8, #15, #16 e #29.

Segundo a Figura 3.38, há evidência gráfica para concluir que a distribuição do vetor de efeitos aleatórios não seja normal. Esse resultado provavelmente é percebido devido à omissão do efeito aleatório de intercepto no modelo ajustado.

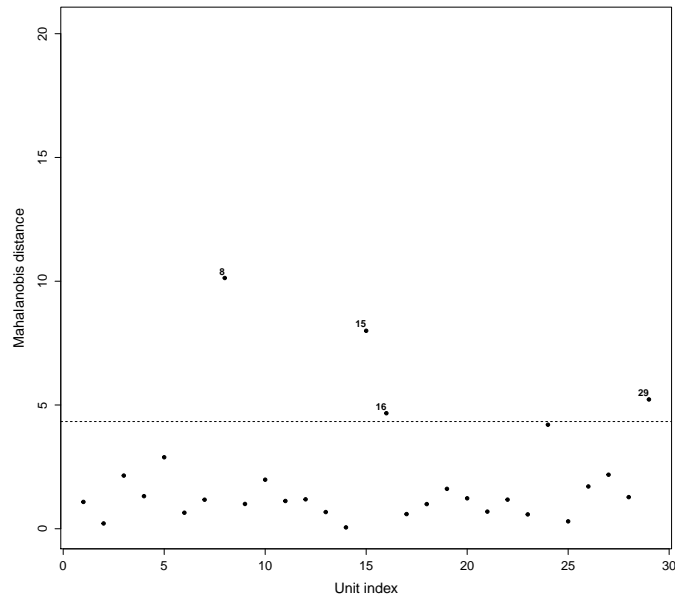


Figura 3.37: Gráfico da distância de Mahalanobis.

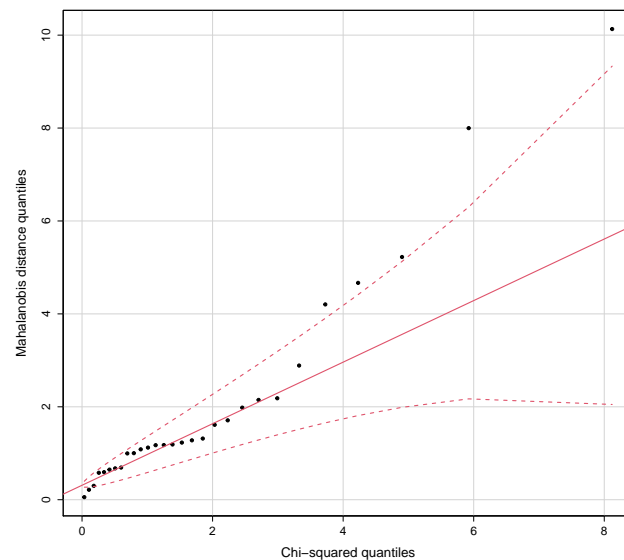


Figura 3.38: Gráfico da distância dos quantis de Mahalanobis.

Na Figura 3.39 é ilustrado o gráfico de resíduos minimamente confundidos padronizados, o qual exibe informação para suspeitar da violação na suposição da normalidade para o erro condicional. Essa detecção é devido ao fato que o modelo ajustado é de independência condicional homocedástico, mas os dados para ajustar o modelo, foram originados tendo presente que o vetor de erros aleatórios segue um processo AR(1).

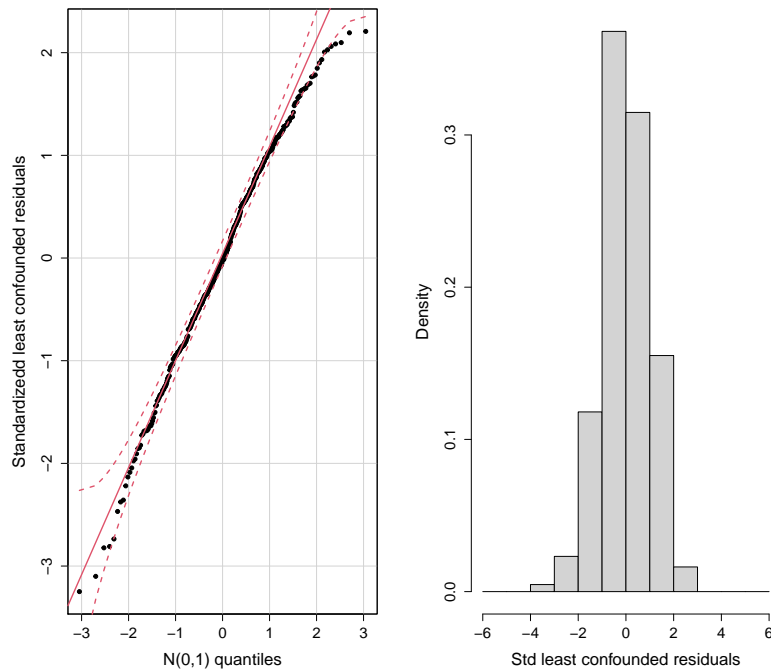


Figura 3.39: *QQ plot e histograma para os resíduos minimamente confundidos padronizados.*

A Figura 3.40 ilustra evidência para suspeitar que o vetor de erros aleatórios não é normal. Além do anterior, é evidenciado que, a observação quinze do indivíduo 11, escrita, #15.11, a observação 15 do indivíduo 10, #15.10, a observação vinte e oito do indivíduo 1, #28.1, podem ser consideradas em destaque, possíveis valores atípicos com respeito às observações restantes.

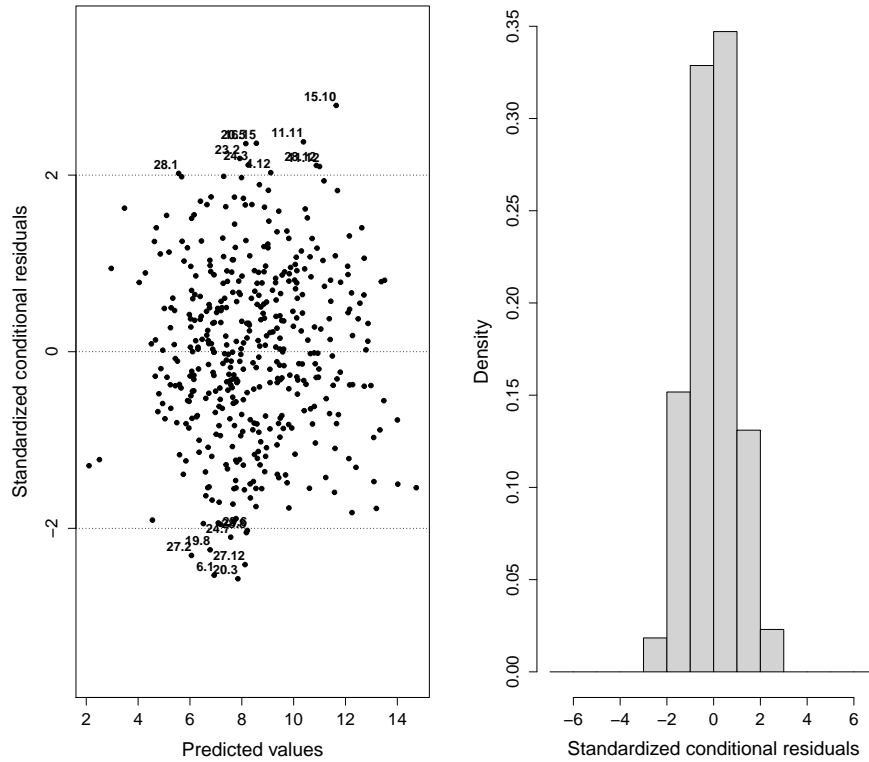


Figura 3.40: Gráfico dos resíduos condicionados padronizados e histograma.

Identificadas as fontes dos erros de especificação, analisamos o poder e a taxa do erro tipo I .

Tabela 3.12: Poder e taxa do erro tipo I ($E.tipoI$) para os testes: da matriz de informação alternativo (I), do estimador “sanduíche” modificado (II) e do estimador “sanduíche” alternativo (III).

n	I		II		III	
	Poder	$E.tipoI$	Poder	$E.tipoI$	Poder	$E.tipoI$
50	0.24	0.63	0.35	0.73	0.15	0.60
100	0.32	0.42	0.42	0.63	0.30	0.53
200	0.48	0.30	0.51	0.35	0.38	0.25
400	0.65	0.21	0.60	0.22	0.45	0.12
600	0.78	0.12	0.78	0.10	0.60	0.09
1000	0.88	0.09	0.89	0.07	0.82	0.08
2000	1.00	0.07	1.00	0.05	1.00	0.06

Analisando os resultados do poder dos testes, notamos que o melhor poder é registrado por o teste do estimador “sanduíche” modificado e o poder mais baixo é registrado por o teste do estimador “sanduíche” alternativo.

Em quanto à taxa do erro tipo I , a menor taxa é exibida por o teste do estimador “sanduíche” alternativo.

Capítulo 4

Aplicação

Nesta seção é analisado o conjunto de dados de [Afiune \(2000\)](#), exibidos na Tabela 1.1 e descritos no Capítulo 1. Os dados são desbalanceados e incompletos, no sentido que nem todos os indivíduos apresentam o mesmo número de observações durante as 15 semanas de observação. Os dados correspondem ao volume (ml) do ventrículo esquerdo de 29 crianças prematuras e constituem um estudo longitudinal realizado na Faculdade de Medicina da Universidade de São Paulo, Brasil.

Na Figura 4.1 podemos observar que não é simples propor um modelo fazendo uma análise visual dos perfis individuais e médio dada a natureza dos dados. O perfil loess sugere um polinômio do segundo grau para o perfil médio. Identificamos heterocedasticidade por intermédio da observação de diferentes variabilidades da resposta ao longo do tempo. Para analisar a variação da resposta média ao longo do tempo, neste caso sugerimos um modelo linear misto gaussiano com intercepto, inclinação e efeito quadrático (dado o comportamento dos perfis) na parte fixa, e intercepto e inclinação na parte aleatória. Com o fim de modelar a heterocedasticidade observada propomos que o modelo seja de independência condicional homocedástico para simplicidade.

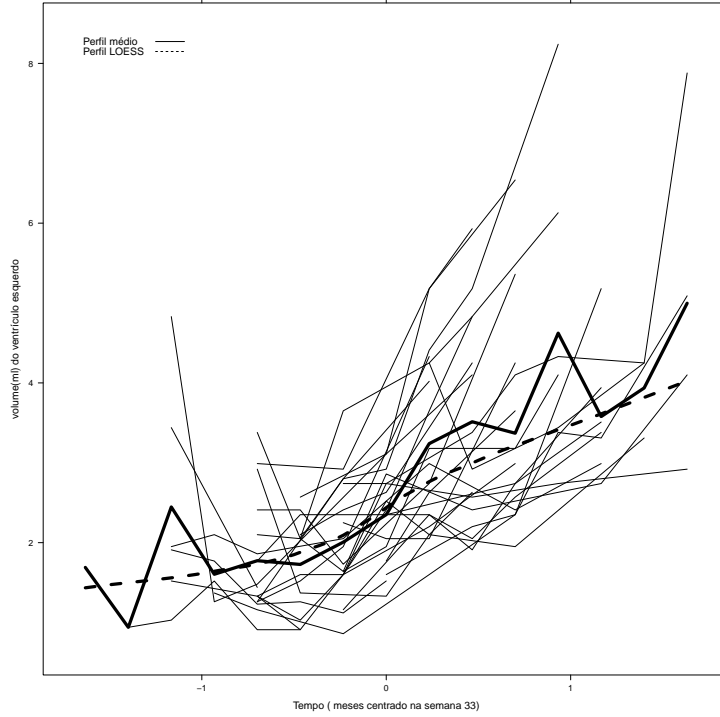


Figura 4.1: Comportamento dos perfis individuais ao redor do perfil médio e do perfil loess.

Consideramos o seguinte modelo

$$y_{ij} = \beta_0 + \beta_1 x_{ij}^* + \beta_2 x_{ij}^{2*} + b_{0i} + b_{1i} x_{ij}^* + e_{ij}. \quad (4.1)$$

$i = 1, \dots, 29$; $j = 1, \dots, m_i$ em que y_{ij} denota a j -ésima observação para o i -ésimo indivíduo, x_{ij}^* representa o tempo em que a observação ocorre. Para reduzir possíveis problemas relacionados com multicolinearidade, transformamos a variável tempo fazendo $x_{ij}^* = \frac{x_{ij} - 33}{4.29}$. Com

a transformação, a unidade de tempo é meses, observando que $4.29 = \frac{30 \text{ dias}}{7 \text{ dias por semana}}$ com

origem na semana 33. β_0, β_1 e β_2 , representam respectivamente o intercepto, a inclinação e o coeficiente quadrático da curva populacional. b_{0i}, b_{1i} são o intercepto e inclinação aleatória associados ao i -ésimo indivíduo e e_{ij} representa o erro aleatório associado à j -ésima observado para o i -ésimo indivíduo.

O modelo em (4.1), pode ser expresso em forma compacta como

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i, \quad (4.2)$$

em que $\mathbf{y}_i = (y_{i1}, \dots, y_{im_i})^\top$ representa as m_i observações registradas para a i -ésima unidade amostral,

$$\mathbf{X}_i = \begin{pmatrix} 1 & x_{i1}^* & x_{i1}^{*2} \\ \vdots & \vdots & \vdots \\ 1 & x_{im_i}^* & x_{im_i}^{*2} \end{pmatrix}, \boldsymbol{\beta} = (\beta_0, \dots, \beta_2)^\top \text{ contém os parâmetros de localização ou}$$

efeitos fixos. $\mathbf{b}_i = (b_{0i}, b_{1i})^\top$ representa o vetor de efeitos aleatórios, $\mathbf{Z}_i = \begin{pmatrix} 1 & x_{i1}^* \\ \vdots & \vdots \\ 1 & x_{im_i}^* \end{pmatrix}$,

corresponde à especificação do vetor de efeitos aleatórios e $\mathbf{e}_i = (e_{i1}, \dots, e_{im_i})^\top$ representa

o vetor de erros aleatórios. Assume-se que $\mathbf{G} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$ e $\mathbf{R}_i = \sigma^2 \mathbf{I}_{m_i}$, supõe-se

independência entre, $\mathbf{b}_i \sim \mathcal{N}_q[\mathbf{0}, \mathbf{G}]$ e $\mathbf{e}_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i]$. Logo

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}[\mathbf{X}_i \boldsymbol{\beta}, \mathbf{V}_i],$$

em que $\mathbf{V}_i = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^\top + \sigma^2 \mathbf{I}_{m_i}$. A estimação do vetor $\boldsymbol{\phi} = (\boldsymbol{\beta}^\top, \boldsymbol{\theta}^\top)^\top$ é realizada usando máxima verossimilhança dado que temos interesse em estudar erro de especificação tanto na estrutura da parte fixa como na matriz de covariancias do vetor de respostas no modelo linear misto gaussiano 2.1.

Apresentamos na Tabela 4.1 os resultados do modelo ajustado.

Tabela 4.1: *Resultados do modelo ajustado*

Parâmetro	<i>Estimativa(SE)</i>
β_0	2.4515 (0.1569)
β_1	1.3990 (0.2079)
β_2	0.5483 (0.1145)
σ_0^2	0.5362 (0.1737)
σ_1^2	0.8883 (0.3412)
σ_{01}	0.5691
σ^2	0.4042 (0.0590)
<i>AIC</i>	382.91
<i>BIC</i>	403.89

Para avaliar se esse modelo está bem especificado, empregamos os testes propostos no Capítulo 2. Os resultados dos testes, apresentados na Tabela 4.2 indicam que existe erro na especificação desse modelo, sem determinar qual ou quais são as característica responsáveis.

Tabela 4.2: *Estatísticas de cada teste e valor – p*

Estatística	<i>valor</i>	<i>valor – p</i>
<i>EAMI</i>	44.09	< 0.001
<i>ESDM</i>	71.14	< 0.001
<i>EESA</i>	69.10	< 0.001

Dado que foi identificado que o modelo apresenta erro de especificação, utilizaremos os diagnósticos gráficos propostos por Singer et al. (2017) para localizar a fonte do erro.

4.0.1 Método gráfico

A Figura 4.2 corresponde à função de autocorrelação para o modelo estimado. Analisando a figura, não existe evidência gráfica para suspeitar que os erros condicionais estejam autocorrelacionados serialmente. Detalhes sob a função de autocorrelação são dados em Pinheiro & Bates (2000, p.258).

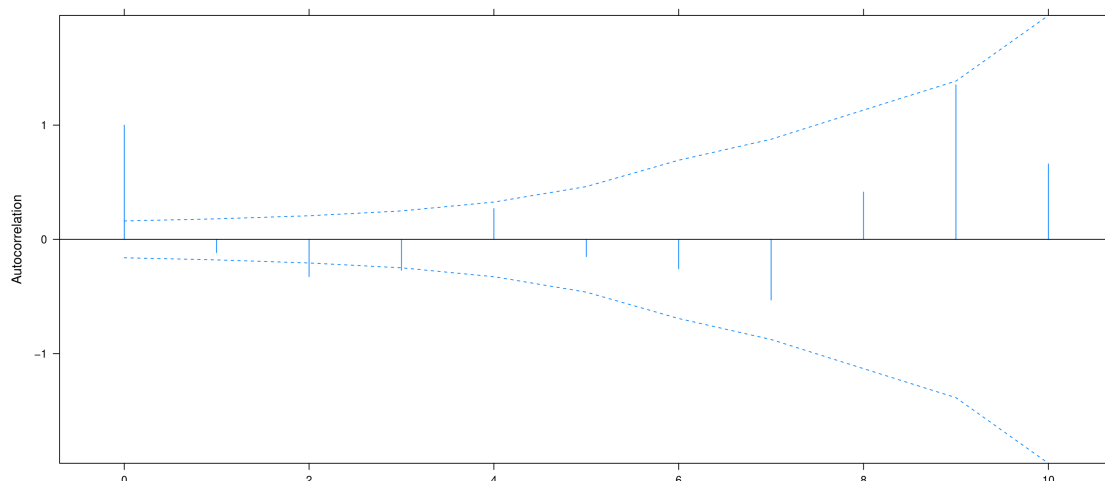


Figura 4.2: *Função de autocorrelação.*

De acordo com o gráfico do índice de Lesaffre–Verbeke modificado na Figura 4.3, a matriz de covariâncias proposta não é a mais adequada para os indivíduos #8 e #9.

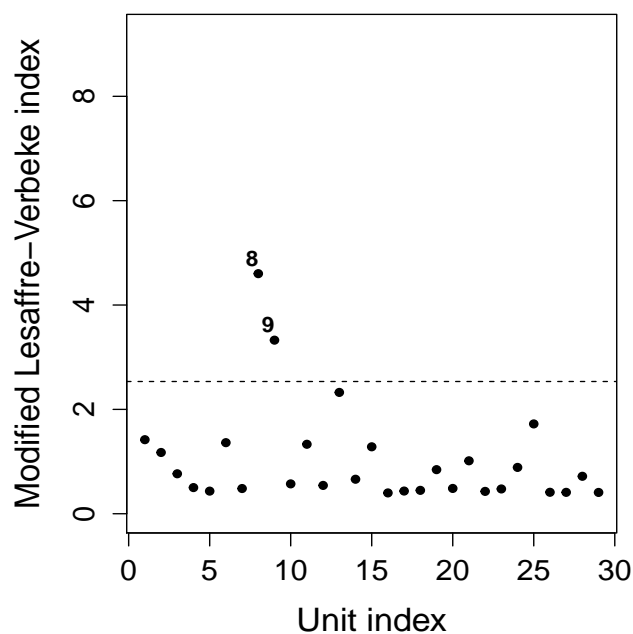


Figura 4.3: *Gráfico do índice de Lesaffre–Verbeke modificado.*

Examinando o resultado gráfico apresentado na Figura 4.4, os resíduos marginais padronizados, não ilustram evidência para suspeitar que a estrutura da parte fixa do modelo seja inadequada.

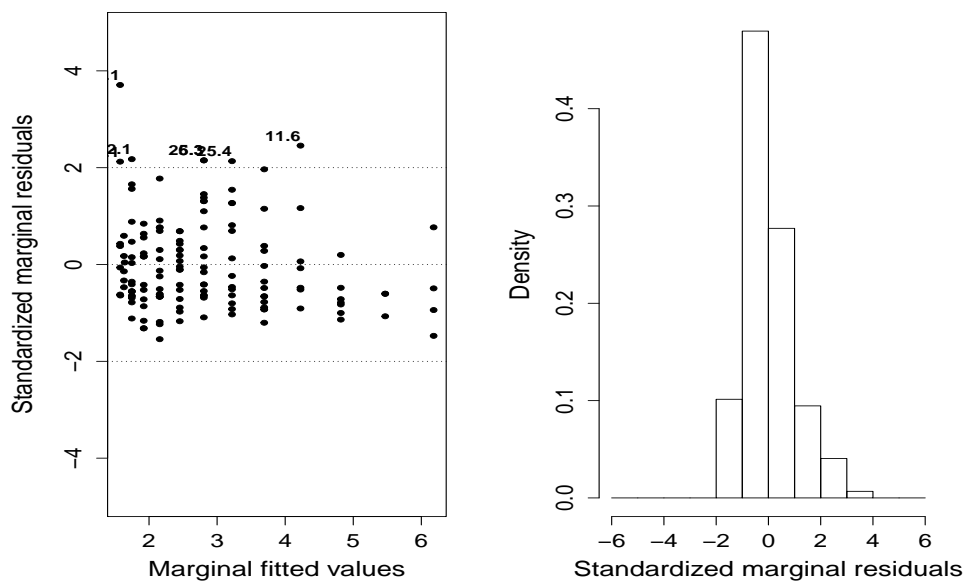


Figura 4.4: *Gráfico QQ de resíduos marginais padronizados e histograma.*

Na Figura 4.5 é ilustrado o gráfico da distância de Mahalanobis, pode – se observar a existência de um ponto sobre a linha tracejada, suspeitando que esse ponto é um possível indivíduo discrepante comparado com os outros.

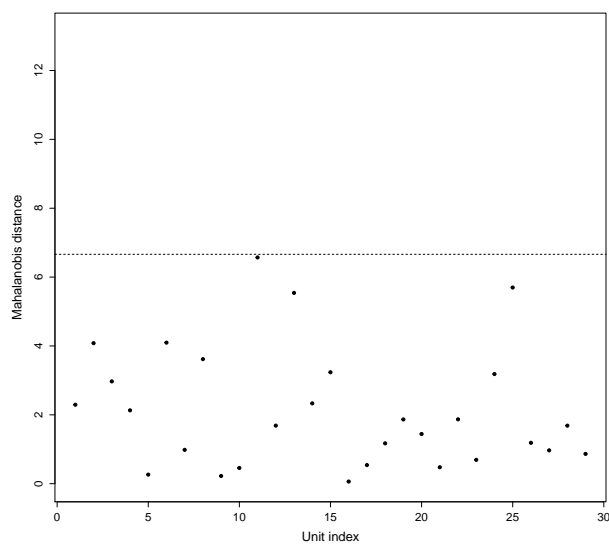


Figura 4.5: *Gráfico da distância de Mahalanobis.*

Na Figura 4.6 é analisado o gráfico da distância dos quantis de Mahalanobis que não exibe evidência gráfica para concluir que a distribuição do vetor de efeitos aleatórios não seja normal.

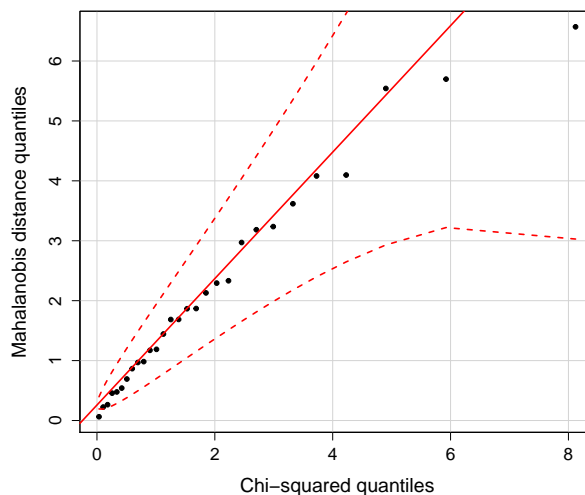


Figura 4.6: *Gráfico da distância dos quantis de Mahalanobis.*

Na Figura 4.7 é ilustrado o gráfico dos resíduos condicionais padronizados, em esse gráfico são destacados como possíveis valores atípicos a observação 1 do sujeito 8, #8.1, a observação 5 do sujeito 9, #9.5 e a observação 1 do sujeito 13, #13.1. Além do anterior, a figura não exibe uma tendência para suspeitar da presença de heterocedasticidade para os erros condicionais.

Na Figura 4.8 o gráfico dos resíduos condicionais minimamente confundidos é apresentado. No gráfico não há tendência para suspeitar que a distribuição do vetor de erros condicionais não seja normal.

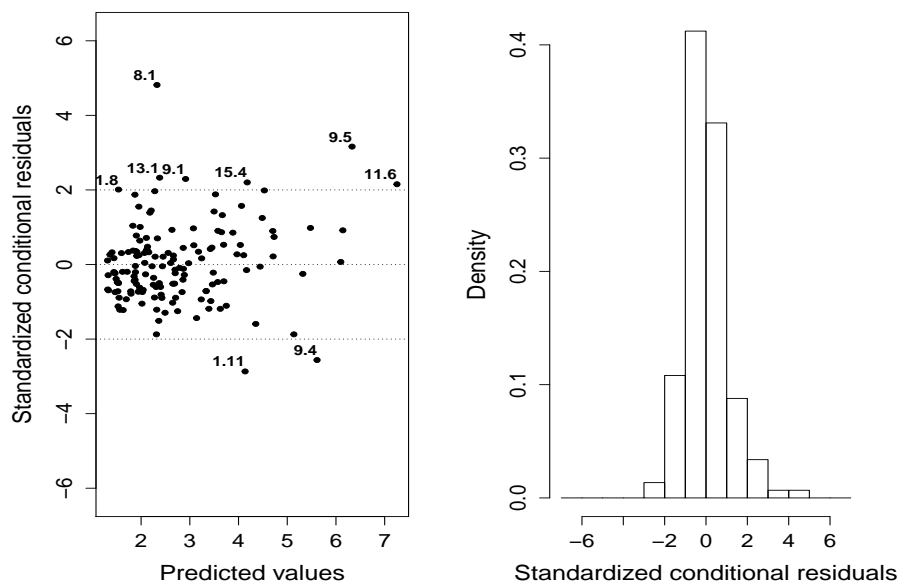


Figura 4.7: Gráfico de resíduos condicionais padronizados e histograma.

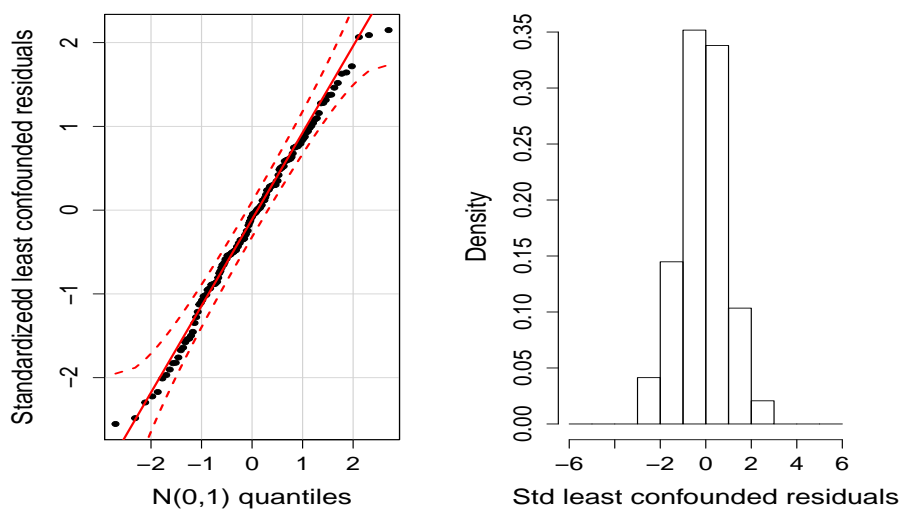


Figura 4.8: Gráfico QQ e histograma dos resíduos condicionais minimamente confundidos.

Analisando os diagnósticos propostos por Singer et al. (2017), dispostos nas Figuras 4.6 e 4.8, não é evidenciada má especificação nem da distribuição dos efeitos aleatórios nem da distribuição dos erros condicionais. No entanto, de acordo com o resultado do gráfico do índice modificado de Lesaffre–Verbeke ilustrado na Figura 4.3, a matriz de covariâncias adotada não está corretamente especificada para as unidades #8 e #9. Esses resultados sugerem

uma modificação do modelo concretizada pela alteração na estrutura de covariâncias com o acréscimo de variâncias diferentes para as unidades #8 e #9. Mais especificamente, propomos um modelo alternativo considerando o mesmo modelo em (4.1), mas com $e_{ij} \sim \mathcal{N}(0, \sigma_i^2)$;

$$\sigma_i^2 = \tau^2, i = 8, 9 \quad e \quad \sigma_i^2 = \sigma^2, \text{ em outro caso.}$$

Os resultados do ajuste estão indicados na Tabela 4.3

Tabela 4.3: *Resultados do ajuste do modelo alternativo*

Parâmetro	<i>Estimativa(SE)</i>
β_0	2.4307 (0.1593)
β_1	1.4410 (0.2190)
β_2	0.4922 (0.0996)
σ_0^2	0.5889 (0.1849)
σ_1^2	1.0864 (0.3676)
σ_{01}	0.6635
σ^2	0.2559 (0.0404)
F. Variância	$(2.879\sigma)^2$
<i>AIC</i>	357.80
<i>BIC</i>	381.78

Este modelo apresenta menor *AIC* e *BIC* que o modelo anterior.

Novamente, para avaliar se o modelo alternativo está bem especificado, empregamos os três testes desenvolvidos neste trabalho. Os resultados dos testes são apresentados na Tabela 4.4, indicam que o modelo apresenta erro de especificação, mas sem determinar a causa.

Tabela 4.4: *Estatística de cada teste e valor - p*

Estatística	<i>valor</i>	<i>valor - p</i>
<i>EAMI</i>	434.04	< 0.001
<i>ESDM</i>	17.11	< 0.001
<i>EESA</i>	1848.01	< 0.001

Dado que foi detectado que o modelo apresenta erro de especificação, o objetivo é localizar a fontes dos erros, para isso os diagnósticos gráficos são utilizados.

4.0.2 Método gráfico

Na Figura 4.9 é apresentado o gráfico da função de autocorrelação, não há tendência para concluir que o modelo final tenha correlação serial nos erros condicionais.

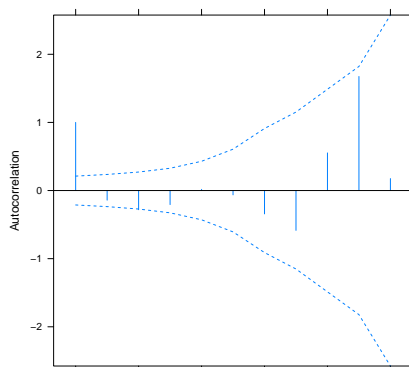


Figura 4.9: Gráfico da função de autocorrelação.

De acordo com o gráfico do índice de Lesaffre–Verbeke modificado na Figura 4.10, a matriz de covariâncias proposta não é a mais adequada, pelo menos, para o indivíduo #13.

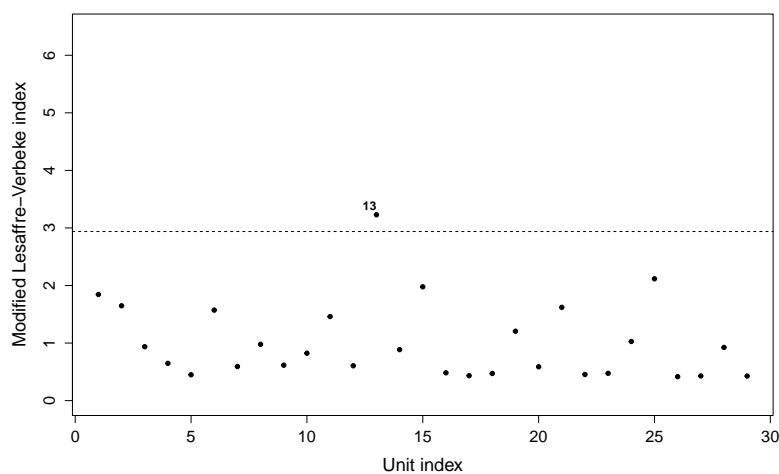


Figura 4.10: Gráfico de índice de Lesaffre–Verbeke modificado.

Na Figura 4.11 dos resíduos marginais padronizados, não há tendência gráfica que indique que a estrutura da parte fixa seja inadequada.

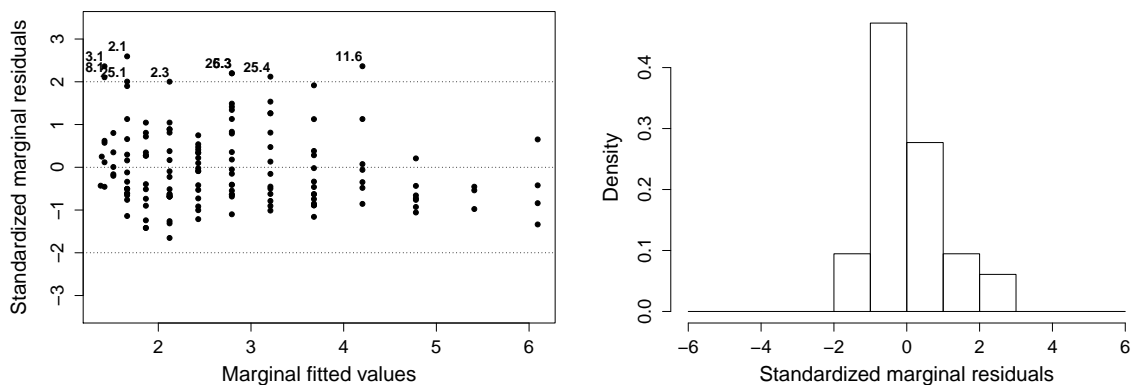


Figura 4.11: *Gráfico dos resíduos marginais padronizados e histograma.*

Na Figura 4.5 é exibida o gráfico da distância de Mahalanobis, pode-se observar a existência de dois pontos sobre a linha tracejada, suspeitando que os pontos poderiam ser valores atípicos.

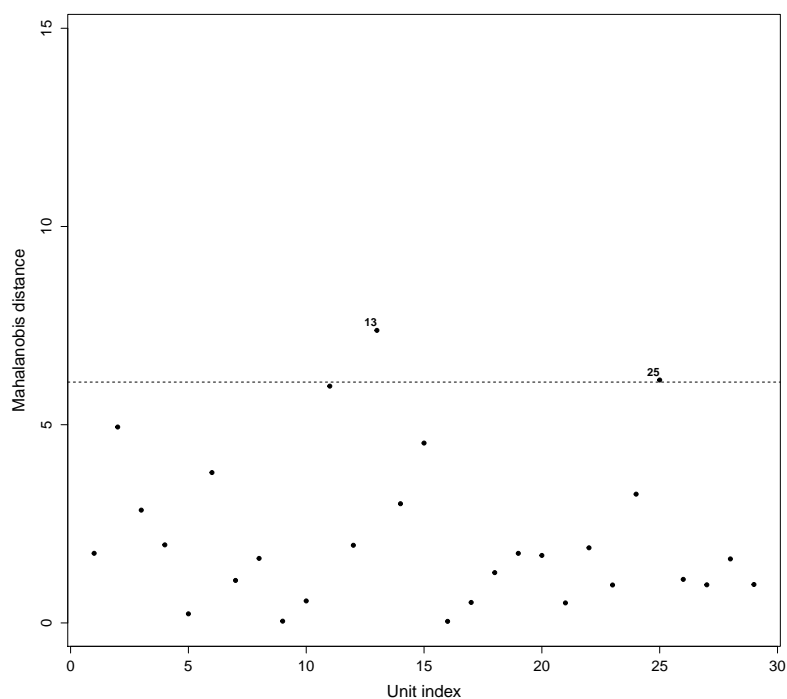


Figura 4.12: *Gráfico da distância de Mahalanobis.*

Segundo a Figura 4.13 da distância dos quantis de Mahalanobis, não há tendência gráfica para suspeitar que existe erro de especificação na distribuição para o vetor dos efeitos aleatórios.

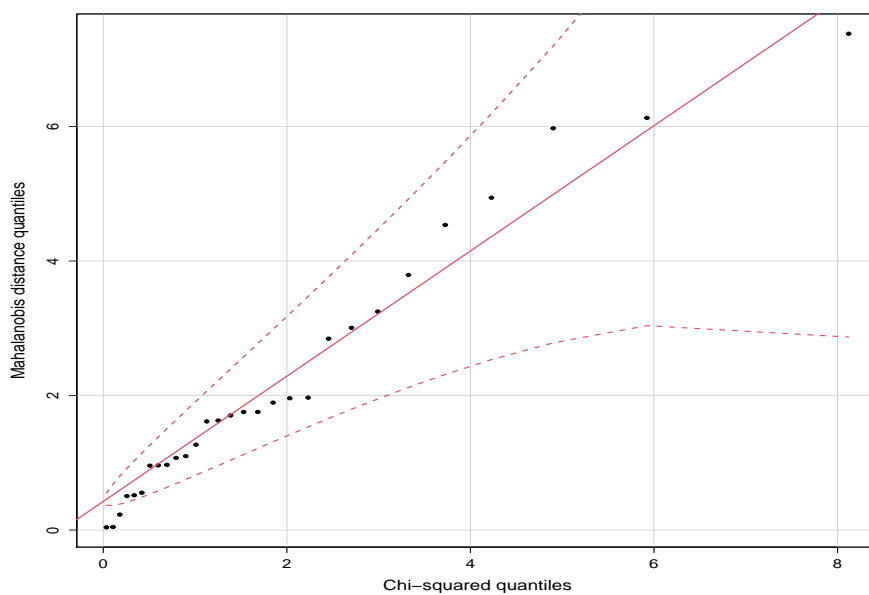


Figura 4.13: *Gráfico da distância dos quantis de Mahalanobis.*

Na figura 4.14 dos resíduos condicionais padronizados não ilustra a presença de uma heterocedasticidade dos erros condicionais. Na Figura 4.15 dos resíduos condicionais minimamente confundidos não existem tendências gráficas que permitam concluir que há erro de especificação na distribuição dos erros condicionais.

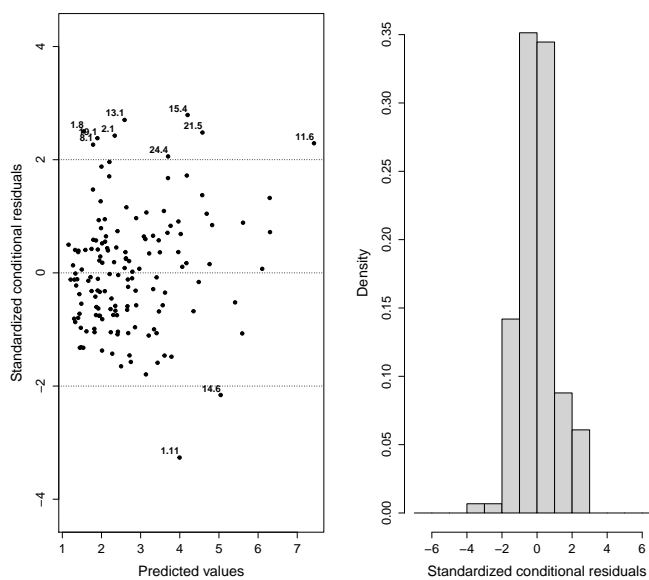


Figura 4.14: *Gráfico dos resíduos condicionais padronizados.*

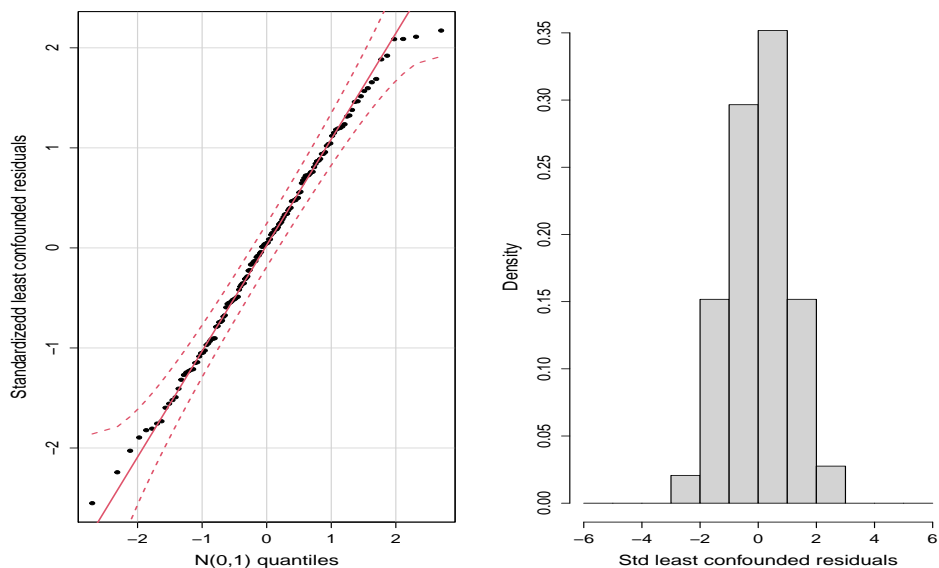


Figura 4.15: Gráfico *QQ* e histograma para os resíduos condicionais minimamente confundidos.

Similarmente como na análise do caso anterior, os diagnósticos dispostos nas Figuras 4.13 e 4.15, não evidenciam má especificação nem da distribuição dos efeitos aleatórios nem na distribuição dos erros condicionais. Não obstante, dado que o resultado gráfico do índice de Lesaffre–Verbeke modificado, Figura 4.10, e o resultado na distribuição dos erros marginais, Figura 4.11, sugerimos uma nova modificação do modelo fazendo uma alteração na estrutura de covariâncias definindo variâncias diferentes para as unidades #8, #9 e #13 e outra variância para o resto do grupo. O modelo final proposto é o mesmo modelo analisado em (4.1), mas com as condições,

$$\sigma_i^2 = \tau^2, i = 8, 9, 13 \quad e \quad \sigma_i^2 = \sigma^2, \text{ em outro caso.}$$

Os resultados do ajuste estão indicados na Tabela 4.5,

Tabela 4.5: *Resultados do ajuste do modelo final*

Parâmetro	<i>Estimativa</i> (<i>SE</i>)
β_0	2.4635 (0.1570)
β_1	1.5007 (0.2084)
β_2	0.4649 (0.098)
σ_0^2	0.5657 (0.1737)
σ_1^2	0.9411 (0.3219)
σ_{01}	0.6015
σ^2	0.2527 (0.0395)
F. Variância	$(2.893\sigma)^2$
<i>AIC</i>	354.95
<i>BIC</i>	378.93

Analisando os resultados do *AIC* e *BIC*, o melhor modelo considerado é o modelo final. Além do anterior, o coeficiente estimado da função variância para o modelo alternativo é $(2.879\sigma)^2$ e para o modelo final é $(2.893\sigma)^2$, ilustrando um decréscimo na variabilidade das medidas do volume.

Para avaliar se esse modelo está bem especificado, empregamos os testes propostos. Os resultados na Tabela 4.6 indicam que não há falha na especificação do modelo.

Tabela 4.6: *Estatísticas dos testes e valor - p*

Estatísticas	valor	<i>valor - p</i>
<i>EAMI</i>	6.20	0.62
<i>ESDM</i>	1.00	0.99
<i>EESA</i>	0.01	1.00

A continuação são empregados os diagnósticos gráficos com o fim avaliar os resultados graficamente.

4.0.3 Método gráfico

Na Figura 4.16 é apresentado o gráfico da função de autocorrelação, não há tendência para concluir que o modelo final tenha correlação serial nos erros condicionais.

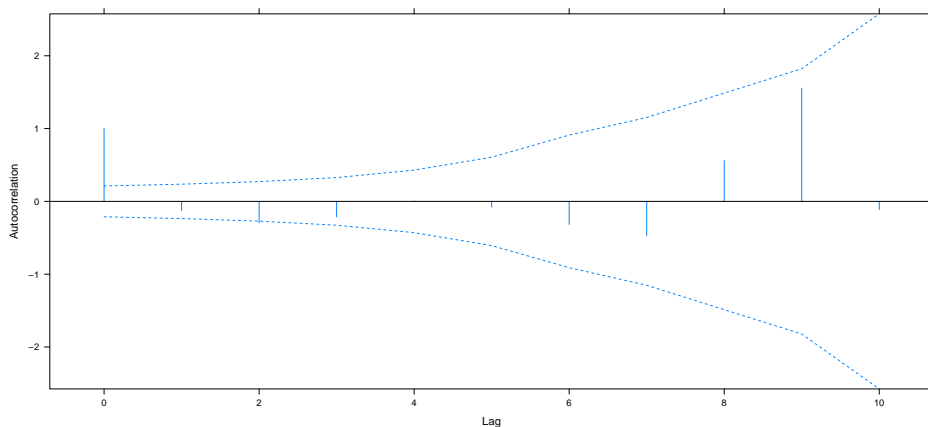


Figura 4.16: *Gráfico da função de autocorrelação.*

Na Figura 4.17 é apresentado o gráfico do índice de Lesaffre–Verbeke modificado, ele não mostra evidência que indique erro de especificação na matriz de covariâncias proposta.

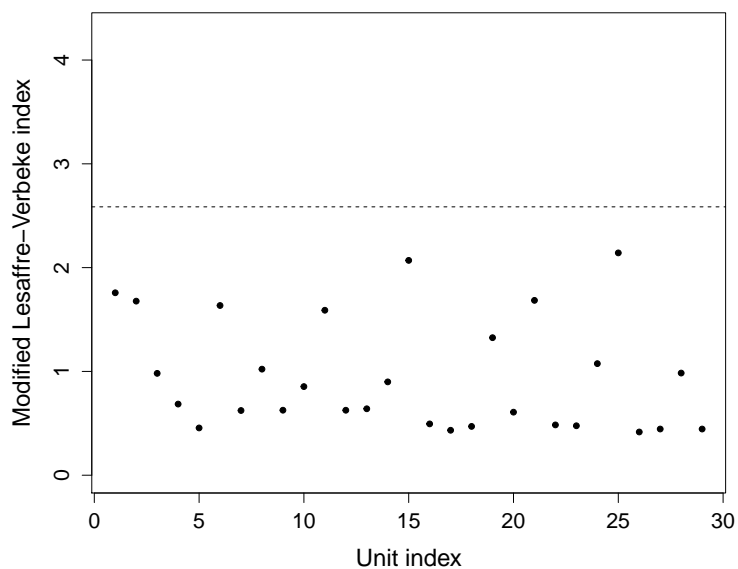


Figura 4.17: *Gráfico do índice de Lesaffre–Verbeke modificado.*

Na Figura 4.18 os resíduos marginais padronizados são ilustrados, não evidenciam graficamente suspeita que exista erro de especificação na especificação da parte fixa do modelo.

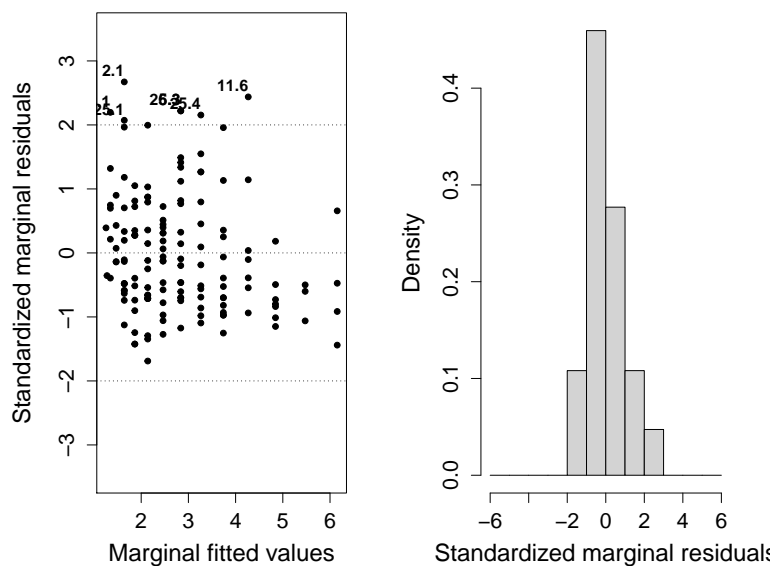


Figura 4.18: *Gráfico dos resíduos marginais padronizados e histograma.*

Na Figura 4.19 é apresentado o gráfico da distância de Mahalanobis. Sob a linha tracejada há dois pontos, indicando que eles poderiam ser possíveis indivíduos atípicos com respeito às restantes observações. O indivíduo #11 tem seis observações, a média de essas observações é 4.41 e o indivíduo #25 tem quatro observações, e a média é 4.22.

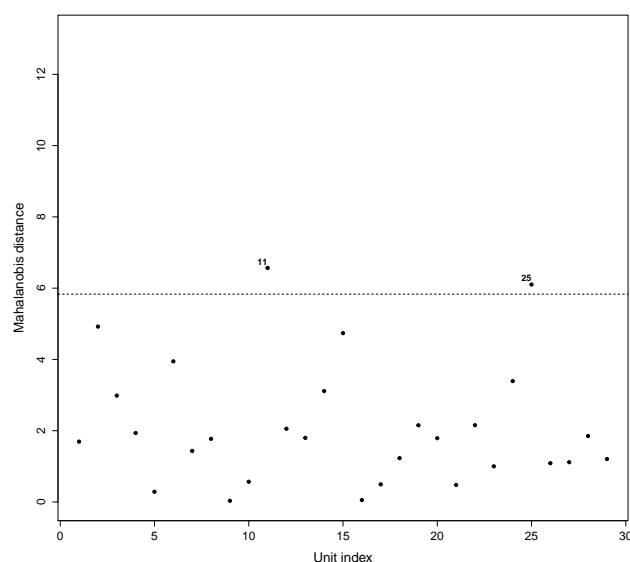


Figura 4.19: *Gráfico da distância de Mahalanobis.*

Na Figura 4.20 da distância dos quantis Mahalanobis não ilustra evidência gráfica para suspeitar que o vetor de efeitos aleatórios não tenha distribuição normal.

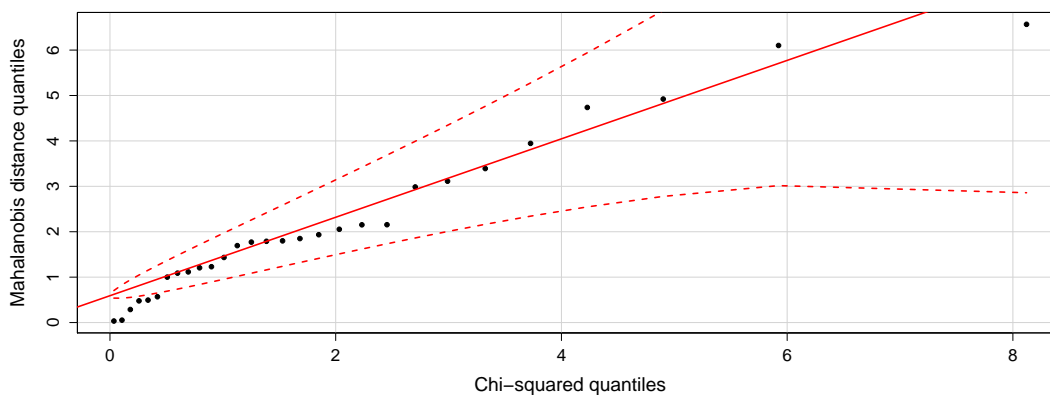


Figura 4.20: *Gráfico da distância dos quantis de Mahalanobis.*

Analisando a Figura 4.21 do gráfico dos resíduos condicionais padronizados e a Figura 4.22 para os resíduos condicionais minimamente confundidos, não existem evidências que indiquem que há erro de especificação na distribuição para os erros condicionais.

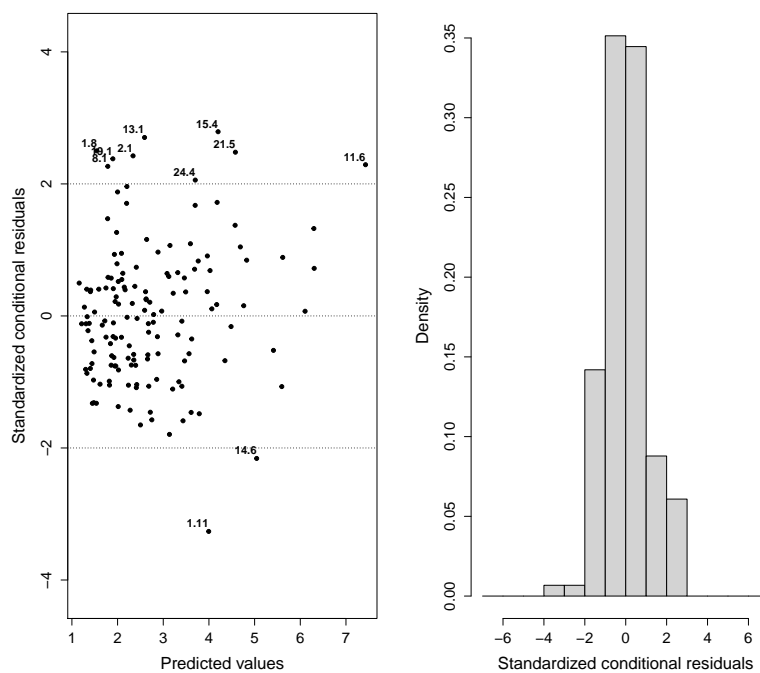


Figura 4.21: *Gráfico dos resíduos condicionais padronizados.*

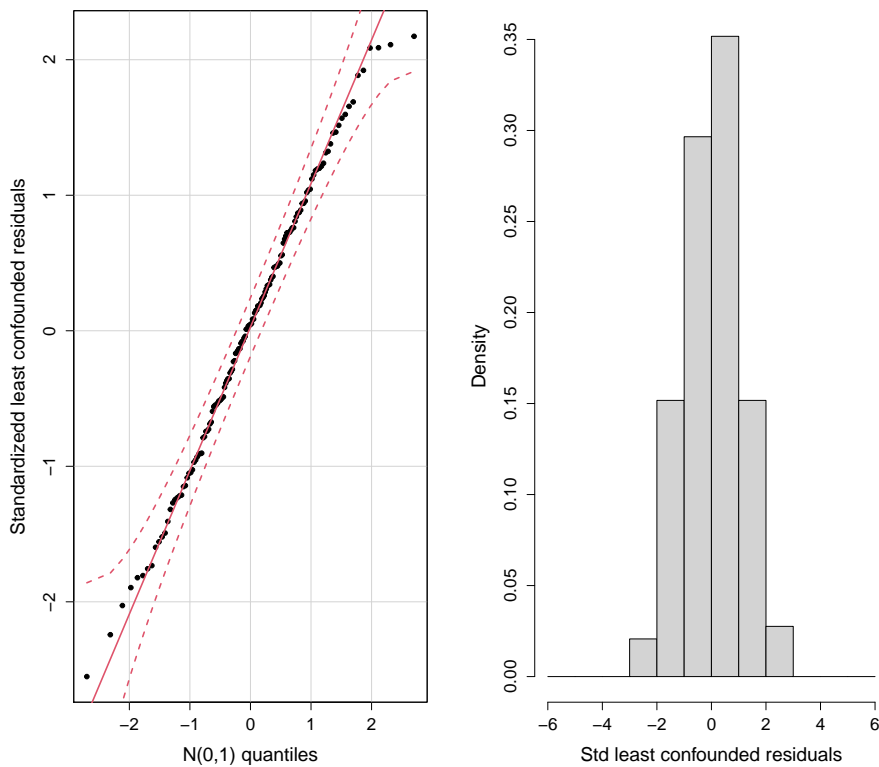


Figura 4.22: Gráfico QQ e histograma para resíduos condicionais minimamente confundidos.

4.1 Aplicação 2

Resultado de um estudo realizado na Escola Paulista de Medicina da Universidade Federal de São Paulo, Brasil um conjunto de dados com uma estrutura balanceada e completa são apresentados nas Tabelas 4.7 e 4.8. Os dados representam a concentração de bilirrubina ($\mu\text{mol}/L$) medida em 89 recém – nascidos a termo saudáveis em aleitamento materno durante os dias 1, 2, 3, 4, 5, 6, 8, 10 e 12. a pós do nascimento. O objetivo é explicar a variação média da concentração de bilirrubina com relação ao tempo.

Tabela 4.7: *Concentração de bilirrubina ($\mu\text{mol/L}$) em recém – nascidos (RN) saudáveis em aleitamento materno.*

RN	Dias após o nascimento								
	1	2	3	4	5	6	8	10	12
1	2.7	0.4	0.0	0.5	0.6	0.0	0.0	0.5	0.8
2	4.5	5.5	3.9	2.7	2.9	2.0	1.5	1.3	1.7
3	7.0	9.2	13.1	12.1	12.3	10.3	8.5	6.2	3.8
4	4.4	6.1	8.8	7.4	6.8	6.2	5.5	4.9	3.3
5	4.0	7.7	7.5	10.3	10.6	9.3	7.3	5.8	6.4
6	6.5	7.4	5.3	4.5	2.9	3.3	3.0	3.2	3.4
7	3.2	2.3	0.9	0.2	0.0	0.0	0.6	0.5	0.0
8	6.2	7.7	5.9	4.1	4.5	2.4	2.0	1.5	0.9
9	3.0	4.5	8.0	8.1	8.6	7.4	6.8	5.5	4.9
10	2.0	2.4	2.9	2.3	1.9	2.1	2.0	1.2	0.1
11	6.2	8.1	10.9	8.0	5.4	6.4	5.0	3.5	3.4
12	4.6	6.3	5.8	4.6	4.6	3.0	3.0	3.0	2.5
13	4.5	7.5	8.0	10.6	10.9	11.0	8.1	7.2	4.3
14	6.6	9.1	7.7	8.4	8.1	7.6	3.9	3.5	1.9
15	2.8	4.3	3.3	2.3	1.0	1.6	1.6	1.5	1.4
16	4.0	5.0	4.2	3.9	2.1	2.2	2.1	2.2	2.1
17	2.8	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
18	6.0	8.1	3.8	2.7	1.7	1.7	1.4	0.0	1.1
19	2.4	4.5	4.7	5.3	4.5	3.5	3.3	3.0	4.0
20	5.4	5.7	3.7	3.0	2.0	0.0	0.0	0.0	0.0
21	3.1	4.3	3.8	4.3	2.8	2.1	1.6	1.3	1.8
22	5.4	6.9	6.9	5.5	6.2	6.2	3.6	3.4	3.8
23	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.5	0.5
24	4.6	4.1	5.1	3.5	4.0	3.4	1.9	2.9	1.3
25	5.5	6.4	8.6	8.1	6.4	6.3	6.5	7.0	7.2
26	2.4	5.3	2.4	1.7	1.6	0.3	0.1	0.0	0.0
27	2.2	1.7	1.2	1.3	0.3	1.5	1.0	0.2	0.2
28	2.9	2.5	1.2	1.1	1.3	1.3	1.3	0.0	0.0
29	5.5	7.8	9.2	8.3	8.2	6.5	5.4	5.4	5.6
30	3.5	4.6	4.4	4.5	3.2	3.0	1.7	1.0	0.0
31	4.2	4.5	1.8	0.7	0.2	1.1	0.6	0.5	0.1
32	6.3	7.8	9.9	6.1	5.3	3.0	3.7	3.5	4.0
33	4.6	1.1	0.7	1.8	1.1	1.0	1.0	1.2	1.6
34	3.6	7.3	7.2	6.0	4.8	4.4	3.6	3.4	3.1
35	2.6	3.8	3.5	1.9	1.2	0.7	0.7	0.9	1.4
36	2.9	4.4	7.3	8.3	8.7	7.6	8.0	7.8	6.8
37	5.0	9.2	9.6	12.3	9.8	12.2	8.6	8.2	7.4
38	4.5	7.0	8.7	10.5	10.3	7.4	7.9	7.5	6.8
39	4.9	9.7	11.2	11.0	8.6	7.7	5.5	4.8	5.0
40	2.3	2.2	2.0	2.9	2.7	2.0	1.4	1.1	0.2
41	0.6	0.2	0.0	0.2	0.0	0.0	0.0	0.4	0.0
42	3.5	5.2	7.4	9.4	7.2	7.0	4.8	4.0	2.1
43	5.0	8.9	8.8	9.4	10.3	10.4	11.8	11.3	10.6
44	7.3	9.7	12.3	10.9	13.7	13.1	12.1	13.6	10.2
45	6.5	6.6	8.0	4.4	3.4	3.3	1.4	2.0	1.2

Tabela 4.8: Concentração de bilirrubina ($\mu\text{mol/L}$) em recém-nascidos (RN) saudáveis em aleitamento materno.

RN	Dias após o nascimento								
	1	2	3	4	5	6	8	10	12
46	5.4	7.3	7.8	8.6	9.1	9.0	9.7	6.3	6.0
47	3.5	5.0	4.8	4.5	3.5	3.5	3.6	3.0	2.8
48	1.5	4.8	4.1	5.0	4.1	3.5	2.1	3.5	3.3
49	3.9	2.9	1.5	1.0	1.5	1.1	0.7	0.5	0.1
50	2.8	2.2	1.5	1.2	0.9	0.3	0.5	0.0	0.2
51	6.2	6.8	4.6	3.5	2.4	3.1	2.7	3.0	3.5
52	5.9	10.0	8.6	9.3	10.5	8.9	10.0	9.7	7.2
53	8.8	9.9	7.5	11.0	7.3	5.0	5.1	4.0	2.9
54	5.4	8.2	8.0	8.2	6.7	4.6	3.2	2.7	4.0
55	5.1	6.0	2.7	1.0	0.8	0.5	0.4	0.5	0.5
56	4.1	9.1	10.5	12.8	14.8	11.8	13.5	11.2	7.3
57	8.0	10.2	9.4	7.6	5.7	2.9	3.0	3.4	2.0
58	4.0	6.5	7.8	8.6	5.5	6.7	3.8	2.2	1.2
59	6.3	5.6	5.5	4.6	3.2	0.9	1.4	1.2	1.2
60	1.4	1.3	0.2	0.0	0.3	0.4	0.6	0.8	1.0
61	3.6	7.3	7.3	8.0	4.2	5.0	2.8	2.5	2.0
62	2.6	3.4	1.3	0.5	1.5	1.0	0.2	0.5	1.1
63	6.7	11.4	14.6	14.1	13.6	9.6	9.0	7.8	3.9
64	2.6	8.6	9.0	9.8	9.6	8.6	6.2	5.9	5.5
65	4.0	6.1	5.6	3.7	3.1	3.9	2.3	2.3	2.2
66	5.5	7.5	8.3	6.6	6.0	4.0	3.0	2.7	3.3
67	6.3	7.5	9.0	8.3	8.7	9.2	9.2	7.6	7.4
68	9.1	11.3	14.4	14.3	9.8	10.0	12.6	13.2	12.4
69	5.9	7.3	10.6	8.7	8.8	7.1	6.9	4.2	4.0
70	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
71	6.7	4.0	2.0	1.2	1.5	1.2	0.3	0.9	0.4
72	6.4	5.7	2.0	1.5	0.5	0.7	1.2	0.6	0.3
73	1.2	0.9	0.2	0.6	0.0	0.0	0.0	0.6	0.4
74	5.1	4.5	4.2	3.6	3.3	2.5	1.7	1.6	1.6
75	2.5	1.5	2.4	1.1	1.9	2.0	1.6	0.8	0.6
76	3.6	5.0	4.1	2.6	1.3	2.4	2.0	2.2	2.4
77	4.3	5.0	6.6	3.1	2.2	2.8	2.5	1.7	1.2
78	7.5	10.6	12.0	13.5	9.6	11.5	9.2	7.6	7.7
79	6.5	8.1	11.0	10.7	10.0	9.7	7.9	7.4	5.2
80	3.6	6.6	9.9	8.8	11.5	12.0	12.0	11.3	9.7
81	3.6	3.7	2.8	2.0	1.5	0.0	1.2	1.6	0.5
82	2.6	1.4	1.3	1.0	1.6	0.4	0.0	0.3	0.6

Na Figura 4.23 são analisados os perfis individuais de acordo com os dados nas tabelas anteriores. Podemos observar que o loess sugere um polinômio de grau três para o perfil médio. Identificamos heterocedasticidade por intermédio da observação de diferentes variabilidades da resposta ao longo do tempo.

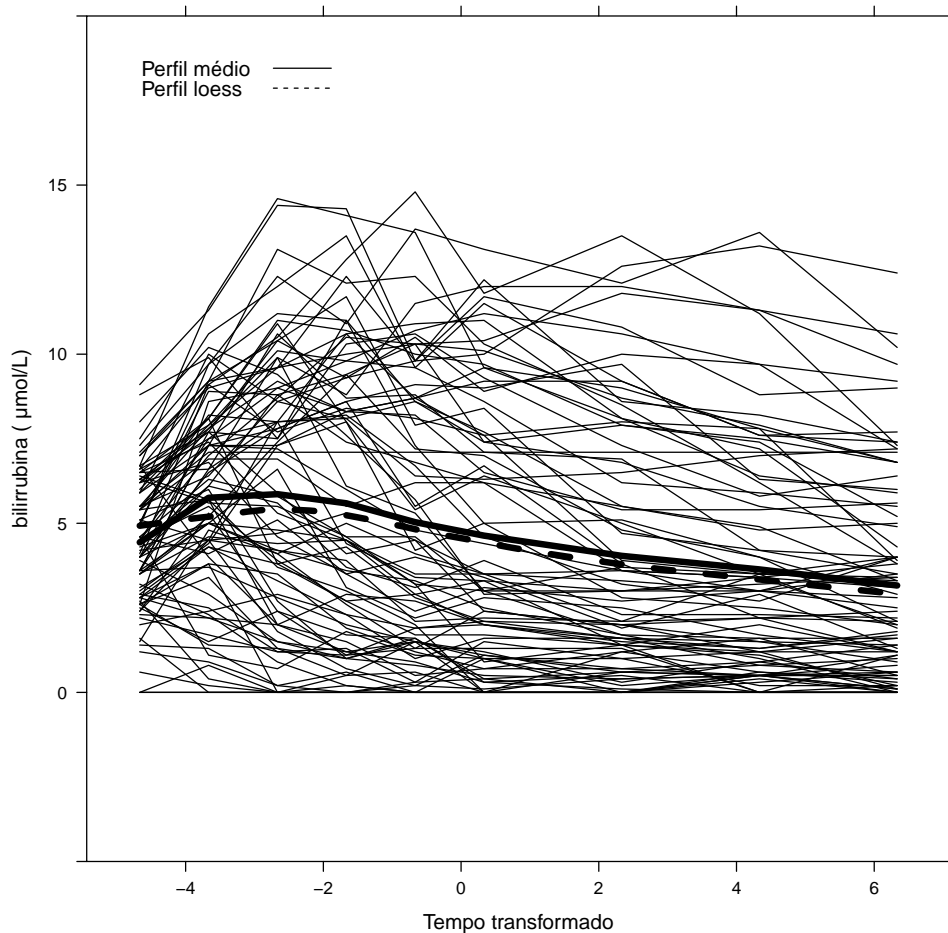


Figura 4.23: *Perfis individuais ao redor do perfil médio e do perfil loess da concentração de bilirrubina.*

Para analisar a variação da resposta média da concentração de bilirrubina ao longo do tempo usamos um modelo linear misto. Empregamos a proposta de Rocha & Singer (2018) para identificar o grau do polinômio da parte fixa e da parte aleatória do modelo. O resultado permite propor um modelo com intercepto, inclinação, efeito quadrático e efeito cúbico na parte fixa, na parte aleatória intercepto e inclinação, considerando um modelo de independência condicional homocedástico por simplicidade.

Modelo proposto

$$y_{ij} = \beta_0 + \beta_1 x_{ij}^{**} + \beta_2 x_{ij}^{2**} + \beta_3 x_{ij}^{3**} + b_{0i} + b_{1i} x_{ij}^{**} + e_{ij}, \quad (4.3)$$

$i = 1, \dots, 82$; $j = 1, \dots, 9$ em que y_{ij} denota a j -ésima observação para o i -ésimo indivíduo. Para reduzir possíveis problemas relacionados com multicolinearidade, transformamos a variável tempo x_{ij} fazendo $x_{ij}^{**} = x_{ij} - \bar{x}$. x_{ij}^{**} representa o j -ésimo tempo para o i -ésimo indivíduo. $\beta_0, \beta_1, \beta_2$ e β_3 , representam respectivamente o intercepto, a inclinação, o coeficiente quadrático e o o coeficiente cúbico da curva populacional. b_{0i}, b_{1i} são o intercepto e inclinação aleatória associados ao i -ésimo indivíduo e e_{ij} representa o erro aleatório associado à j -ésima observado para o i -ésimo indivíduo.

O modelo em (4.3) pode ser expresso em forma compacta como

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \mathbf{e}_i, \quad (4.4)$$

em que $\mathbf{y}_i = (y_{i1}, \dots, y_{im_i})^\top$ representa as m_i observações registradas para a i -ésima unidade amostral,

$$\mathbf{X}_i = \begin{pmatrix} 1 & x_{i1}^{**} & x_{i1}^{**2} & x_{i1}^{**3} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{im_i}^{**} & x_{im_i}^{**2} & x_{im_i}^{**3} \end{pmatrix}, \boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)^\top \text{ contém os parâmetros de localização}$$

ou efeitos fixos. $\mathbf{b}_i = (b_{0i}, b_{1i})^\top$ representa o vetor de efeitos aleatórios, $\mathbf{Z}_i = \begin{pmatrix} 1 & x_{i1}^{**} \\ \vdots & \vdots \\ 1 & x_{im_i}^{**} \end{pmatrix}$,

corresponde à especificação do vetor de efeitos aleatórios e $\mathbf{e}_i = (e_{i1}, \dots, e_{im_i})^\top$ representa

o vetor de erros aleatórios. Assume-se que $\mathbf{G} = \begin{pmatrix} \sigma_0^2 & \sigma_{01} \\ \sigma_{01} & \sigma_1^2 \end{pmatrix}$ e $\mathbf{R}_i = \sigma^2 \mathbf{I}_{m_i}$, supõe-se

independência entre, $\mathbf{b}_i \sim \mathcal{N}_q[\mathbf{0}, \mathbf{G}]$ e $\mathbf{e}_i \sim \mathcal{N}_{m_i}[\mathbf{0}, \mathbf{R}_i]$. Logo

$$\mathbf{y}_i \sim \mathcal{N}_{m_i}[\mathbf{X}_i \boldsymbol{\beta}, \mathbf{V}_i],$$

em que $\mathbf{V}_i = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^\top + \sigma^2 \mathbf{I}_{m_i}$. A estimação do vetor $\boldsymbol{\phi} = (\boldsymbol{\beta}^\top, \boldsymbol{\theta}^\top)^\top$ é realizada usando máxima verossimilhança.

Ajustamos o modelo (4.3) aos dados descritos nas Tabelas 4.7 e 4.8 e na Tabela 4.9 apresentamos os resultados.

Tabela 4.9: *Resultados do modelo ajustado*

Parâmetro	Estimativa (SE)
β_0	4.8670 (0.3427)
β_1	-0.4220 (0.0432)
β_2	-0.0443 (0.0058)
β_3	0.0106 (0.0015)
σ_0	3.0140
σ_1	0.1928
σ_{01}	0.1713
σ	1.3999
AIC	3002.41
BIC	3039.25

Para avaliar se o modelo está bem especificado, empregamos os testes propostos no Capítulo 2. Os resultados dos testes apresentados na Tabela 4.10 indicam que existe erro de especificação no modelo em estudo, sem determinar as característica responsáveis.

Tabela 4.10: *Estatísticas de cada teste e valor – p*

Estatística	valor	valor – p
<i>EAMI</i>	24.65	< 0.001
<i>ESDM</i>	59.22	< 0.001
<i>EESA</i>	32.01	< 0.001

Dado que foi identificado que o modelo apresenta erro de especificação, utilizaremos os diagnósticos gráficos para localizar a fonte do erro.

A função de autocorrelação ilustrada na Figura 4.24, mostra evidência gráfica para suspeitar que os erros condicionais estejam correlacionados serialmente.

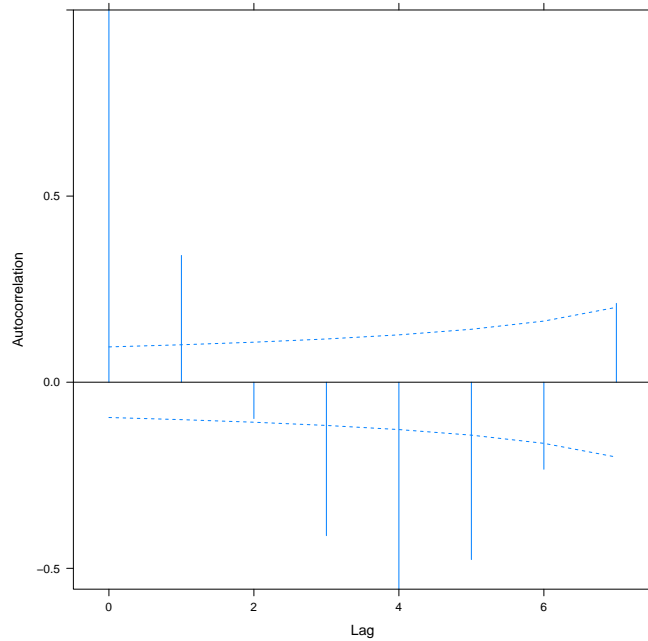


Figura 4.24: Gráfico da função de autocorrelação.

O gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 4.25, detecta que a matriz de covariâncias proposta é inapropriada al menos para os indivíduos #3, #56, #63 e #80.

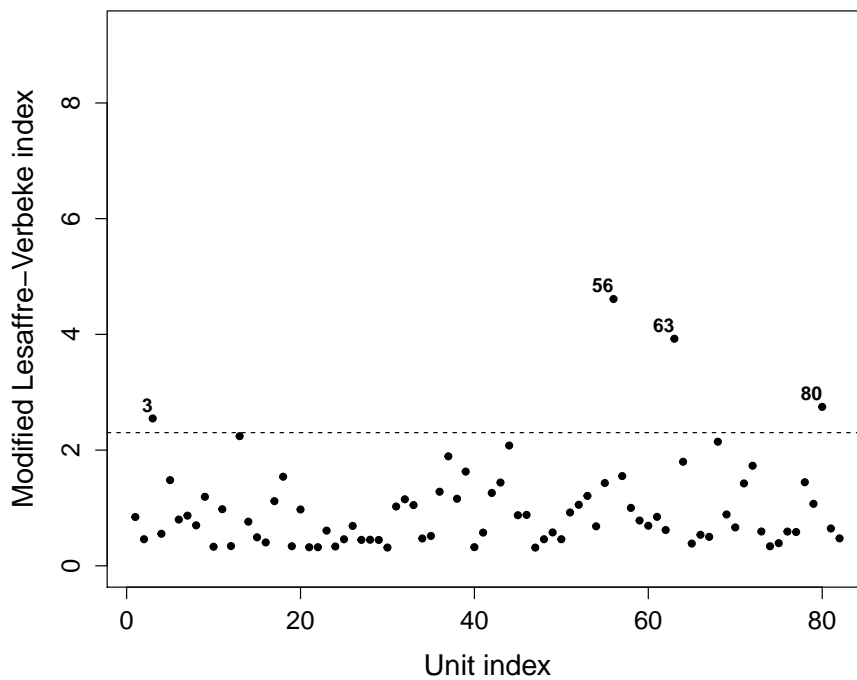


Figura 4.25: Gráfico de índice modificado de Lesaffre - Verbeke.

Analisando o gráfico dos resíduos marginais padronizados apresentado na Figura 4.26,

não existe evidência para suspeitar que a estrutura da média seja inadequada.

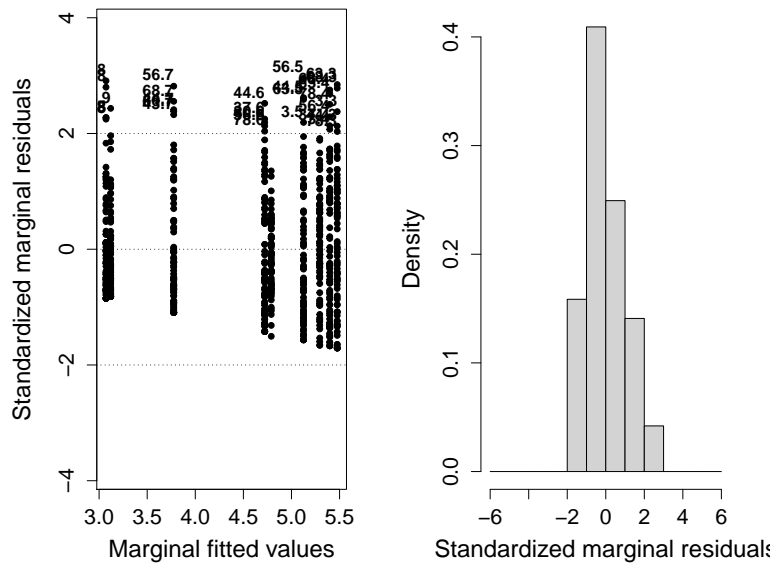


Figura 4.26: Gráfico dos resíduos marginais padronizados.

Os resultados apresentados na Figura 4.27 permitem evidenciar que a distribuição do vetor de efeitos aleatórios tem um afastamento da distribuição normal.

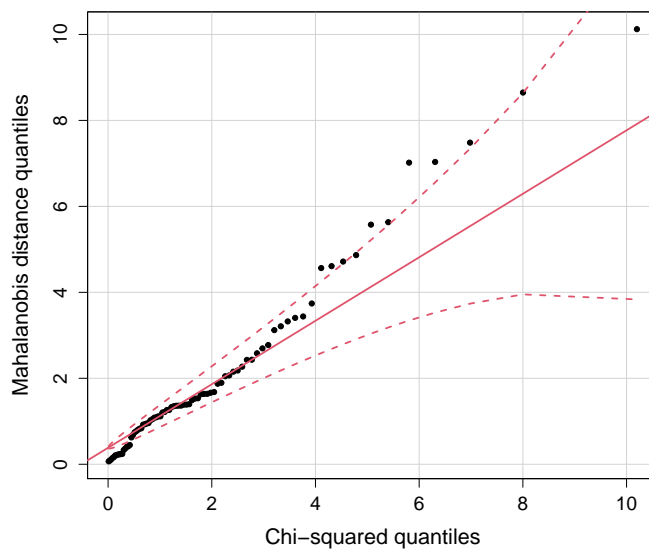


Figura 4.27: Gráfico da distância dos quantis de Mahalanobis.

Analisando a Figura 4.28 os indivíduos #3, #43, #44, #56, #57, #63, #68 e #80 podem ser considerados como observações outliers com respeito as restantes.

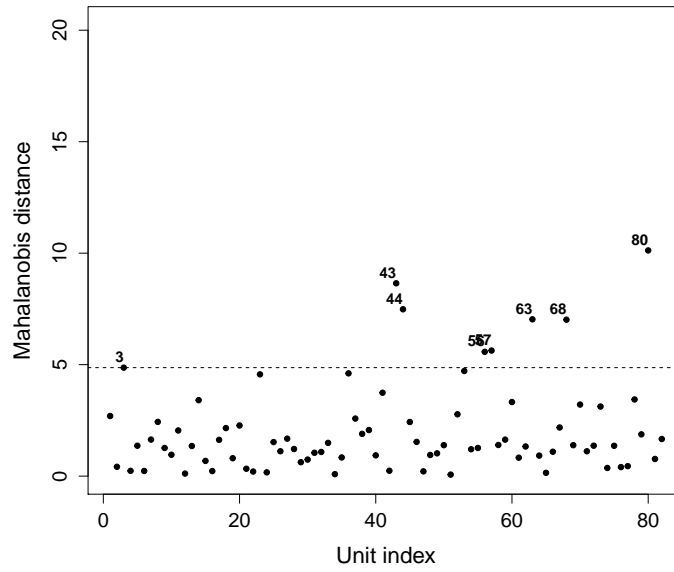


Figura 4.28: Gráfico da distância de Mahalanobis.

Pela Figura 4.29 é percebido plausibilidade na suposição da normalidade para a distribuição nos erros condicionais.

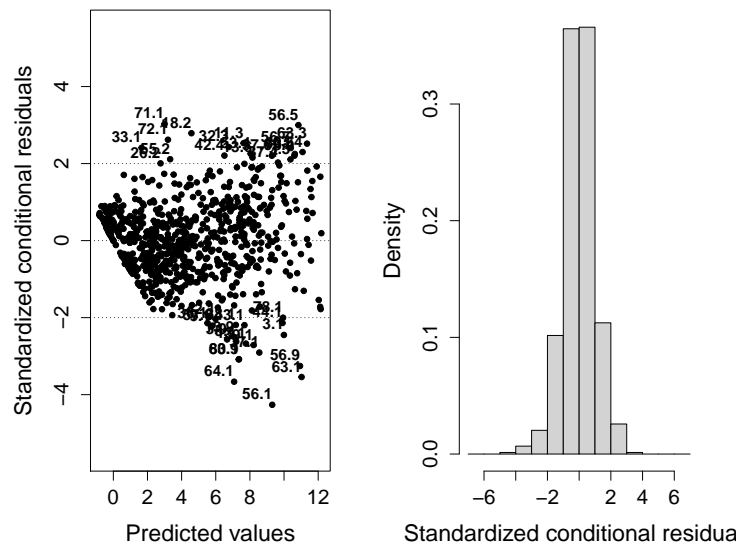


Figura 4.29: Gráfico dos resíduos condicionais padronizados.

Analisando a Figura 4.30, evidenciamos que a distribuição do vetor de erros aleatórios tem um afastamento da distribuição normal.

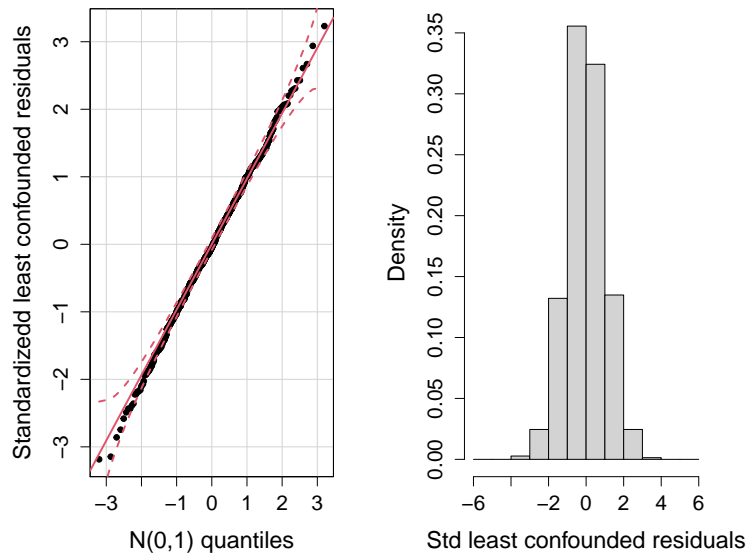


Figura 4.30: Gráfico *QQ* para os resíduos condicionais padronizados e histograma.

De acordo com o resultado do gráfico do índice modificado de Lesaffre–Verbeke ilustrado na Figura 4.25, é sugerido uma modificação do modelo (4.3) concretizada pela alteração na estrutura de covariâncias com o acréscimo de variâncias diferentes para as unidades #3, #56, #63 e #80. Mais especificamente, propomos um modelo alternativo considerando o mesmo modelo em (4.3), mas com $e_{ij} \sim \mathcal{N}(0, \sigma_i^2)$, em que

$$\sigma_i^2 = \tau^2, i = 3, 56, 63 \text{ e } 80 \quad \text{e} \quad \sigma_i^2 = \sigma^2, \text{ em outro caso.}$$

Ajustamos o modelo (4.3) aos dados descritos nas Tabelas 4.7 e 4.8. Na Tabela 4.11 os resultados do modelo ajustado são apresentados.

Tabela 4.11: *Resultados do modelo ajustado*

Parâmetro	Estimativa (SE)
β_0	4.7535 (0.3399)
β_1	-0.4224 (0.0409)
β_2	-0.0356 (0.0055)
β_3	0.0100 (0.0015)
σ_0	2.9917
σ_1	0.1813
σ_{01}	0.1635
σ	1.2993
F. Variância	$(2.2209\sigma)^2$
AIC	2954.73
BIC	2996.17

Segundo os resultados na Tabela 4.12 é identificado que o modelo apresenta erro de especificação.

Tabela 4.12: *Estatísticas de cada teste e valor – p*

Estatística	valor	valor – p
<i>EAMI</i>	58.67	< 0.001
<i>ESDM</i>	46.27	< 0.001
<i>EESA</i>	44.12	< 0.001

Utilizaremos os diagnósticos gráficos para localizar a fonte do erro. A função de autocorrelação ilustrada na Figura 4.31, mostra evidência gráfica para suspeitar que os erros condicionais estão correlacionados serialmente.

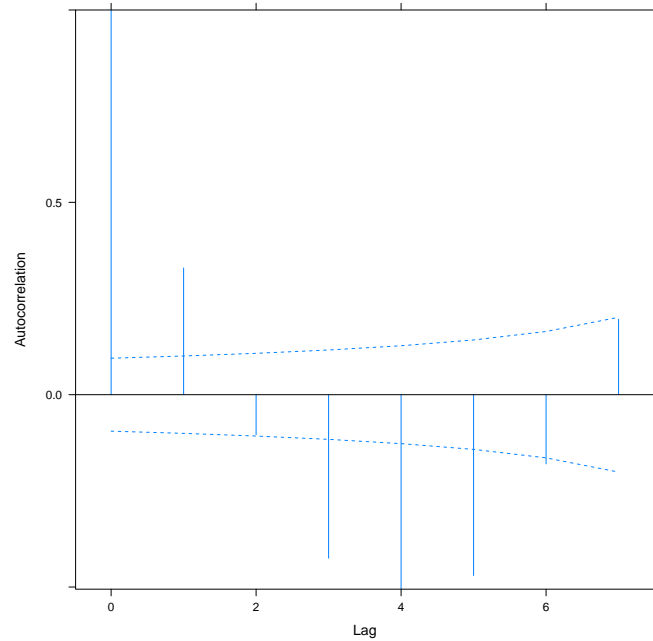


Figura 4.31: Gráfico da função de autocorrelação.

O gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 4.32, detecta que a matriz de covariâncias do vetor de respostas é inapropriada para o indivíduo #13.

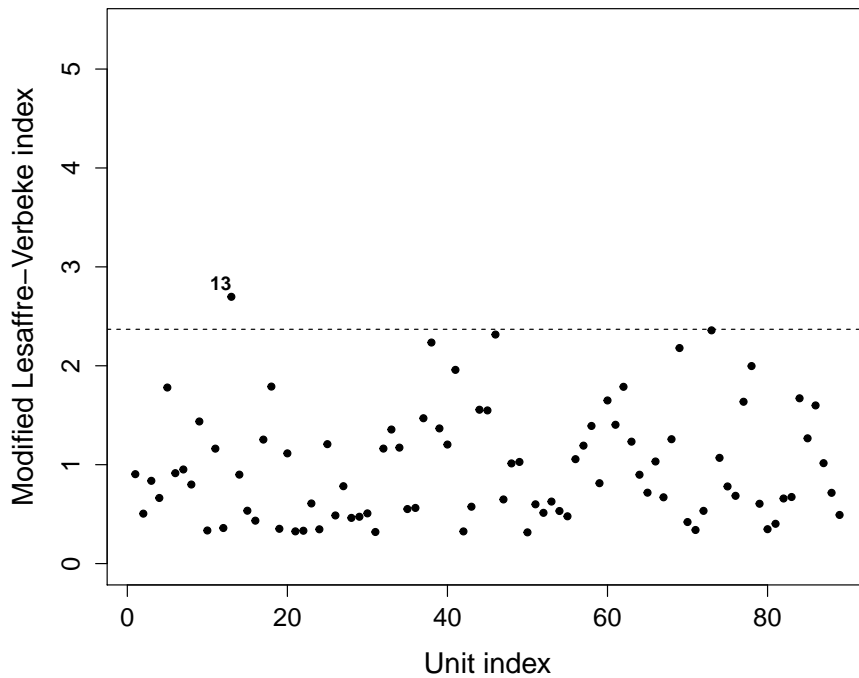


Figura 4.32: Gráfico de índice modificado de Lesaffre - Verbeke.

O gráfico dos resíduos marginais padronizados apresentado na Figura 4.33, não mostra evidência gráfica para suspeitar que existe violação na estrutura da média.

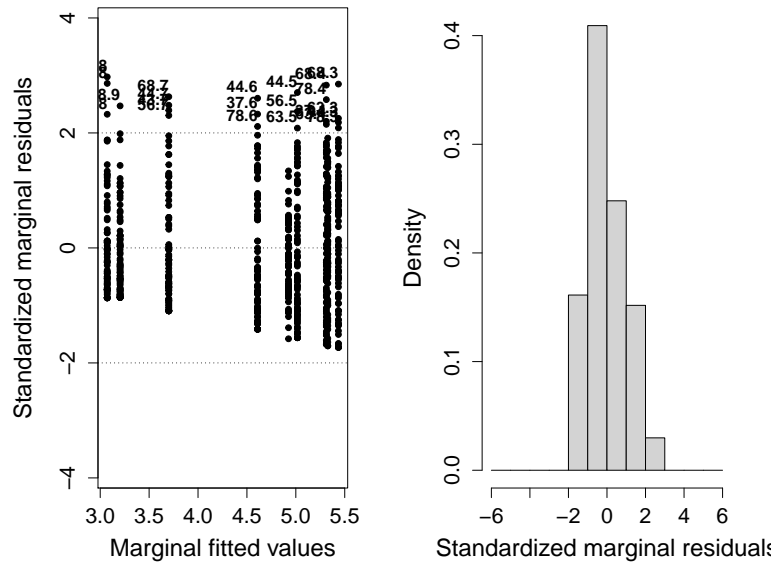


Figura 4.33: Gráfico dos resíduos marginais padronizados.

Os resultados apresentados na Figura 4.34 permitem evidenciar que a distribuição do vetor de efeitos aleatórios não tem um afastamento da distribuição normal.

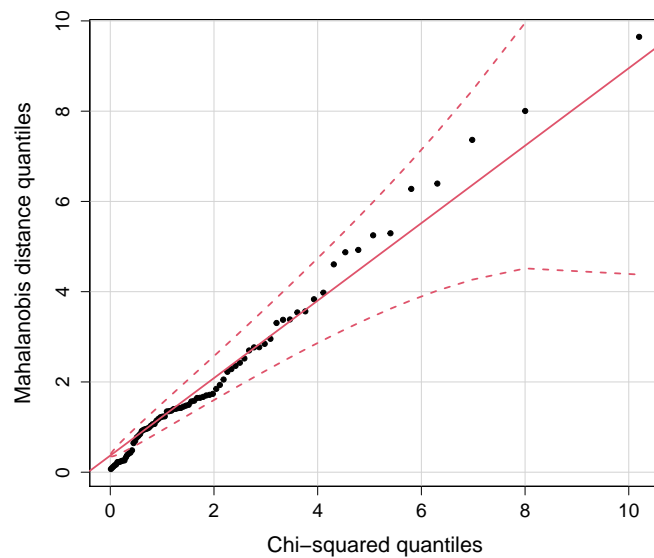


Figura 4.34: Gráfico da distância dos quantis de Mahalanobis.

Analisando a Figura 4.35 os indivíduos #43, #44, #57, #68 e #80 podem ser considerados como observações outliers com respeito as restantes.

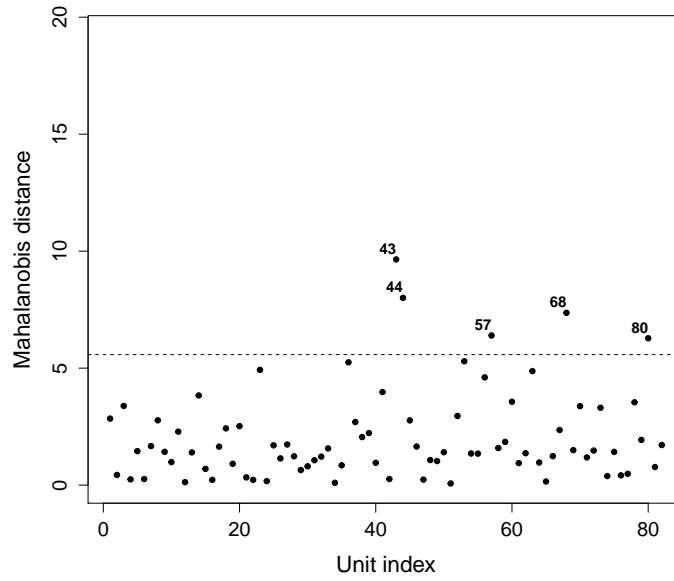


Figura 4.35: *Gráfico da distância de Mahalanobis.*

Pela Figura 4.36 é percebido plausibilidade na suposição de normalidade para a distribuição nos erros condicionais.

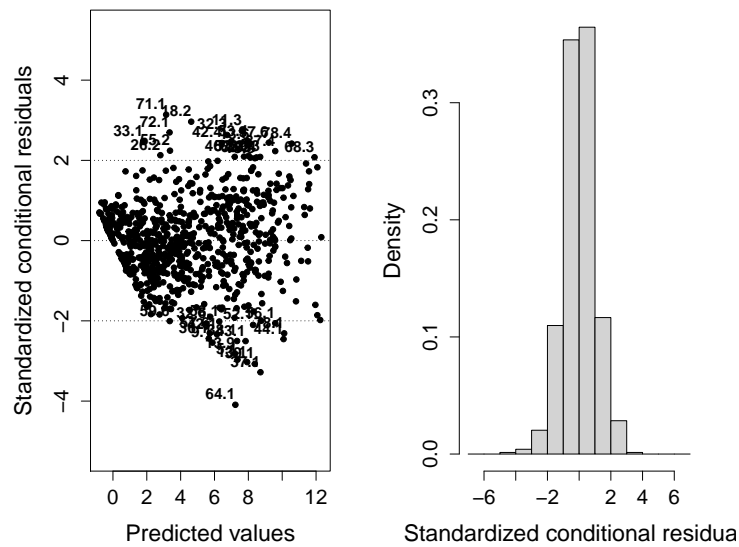


Figura 4.36: *Gráfico dos resíduos condicionais padronizados.*

Analisando a Figura 4.37 evidenciamos que a distribuição do vetor de erros aleatórios não tem um afastamento da distribuição normal.

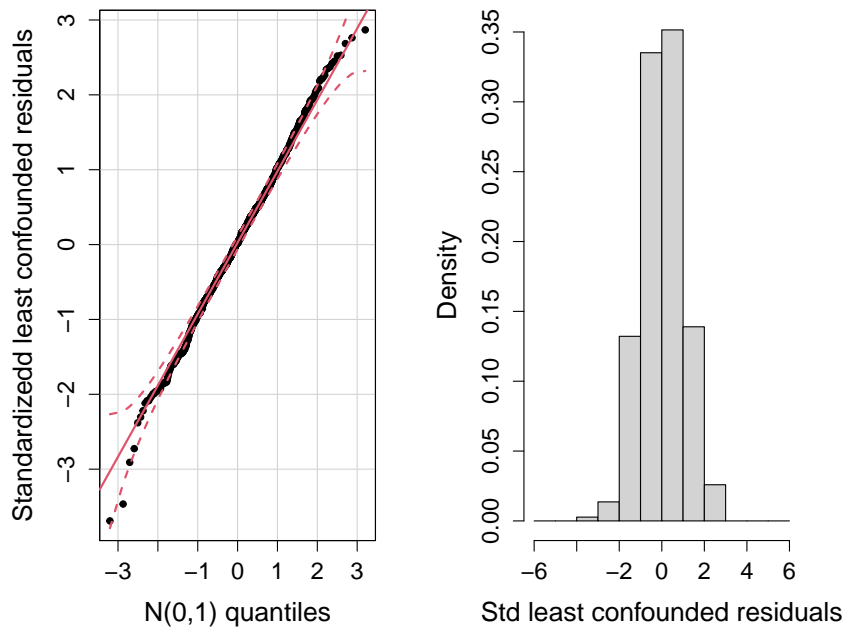


Figura 4.37: Gráfico *QQ* para os resíduos condicionais padronizados e histograma.

Em forma similar como no caso anterior considerando o resultado do gráfico do índice modificado de Lesaffre–Verbeke ilustrado na Figura 4.32, sugerimos uma modificação do modelo (4.3) com a alteração na estrutura de covariâncias, considerando variâncias diferentes para as unidades #3, #13 #56, #63 e #80. Mais especificamente, propomos um modelo alternativo considerando o mesmo modelo em (4.3), mas com $e_{ij} \sim \mathcal{N}(0, \sigma_i^2)$, em que

$$\sigma_i^2 = \tau^2, i = 3, 13, 56, 63 \text{ e } 80 \quad \text{e} \quad \sigma_i^2 = \sigma^2, \text{ em outro caso.}$$

Uma vez que o modelo (4.3) é ajustado aos dados, o objetivo é avaliar se o modelo está bem especificado, para isso empregamos os testes desenvolvidos.

Na Tabela 4.13 os resultados do modelo ajustado são apresentados.

Tabela 4.13: *Resultados do modelo ajustado*

Parâmetro	Estimativa (SE)
β_0	4.7292 (0.3399)
β_1	-0.4239 (0.0407)
β_2	-0.0336 (0.0054)
β_3	0.0099 (0.0014)
σ_0	2.9926
σ_1	0.1838
σ_{01}	0.1635
σ	1.2777
F. Variância	$(2.1816\sigma)^2$
AIC	2946.77
BIC	2988.20

Analisando a Tabela 4.14,

Tabela 4.14: *Estatísticas de cada teste e valor - p*

Estatística	valor	valor - p
<i>EAMI</i>	72.55	< 0.001
<i>ESDM</i>	60.45	< 0.001
<i>EESA</i>	69.22	< 0.001

concluimos que os testes detectam que o modelo apresenta erro de especificação. Agora usando os diagnósticos gráficos, ilustramos a fonte do erro.

A Figura 4.38 mostra evidência gráfica para suspeitar que os erros condicionais estejam correlacionados serialmente.

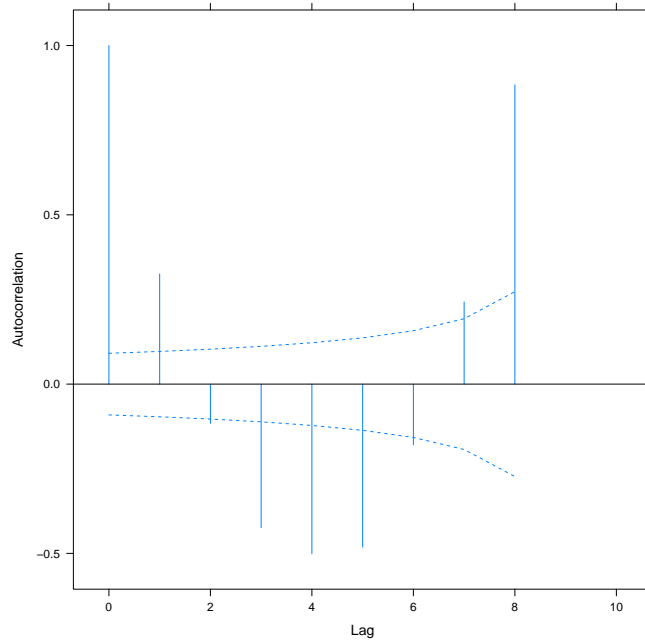


Figura 4.38: Gráfico da função de autocorrelação.

O gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 4.39, não detecta que a matriz de covariâncias proposta seja inapropriada.

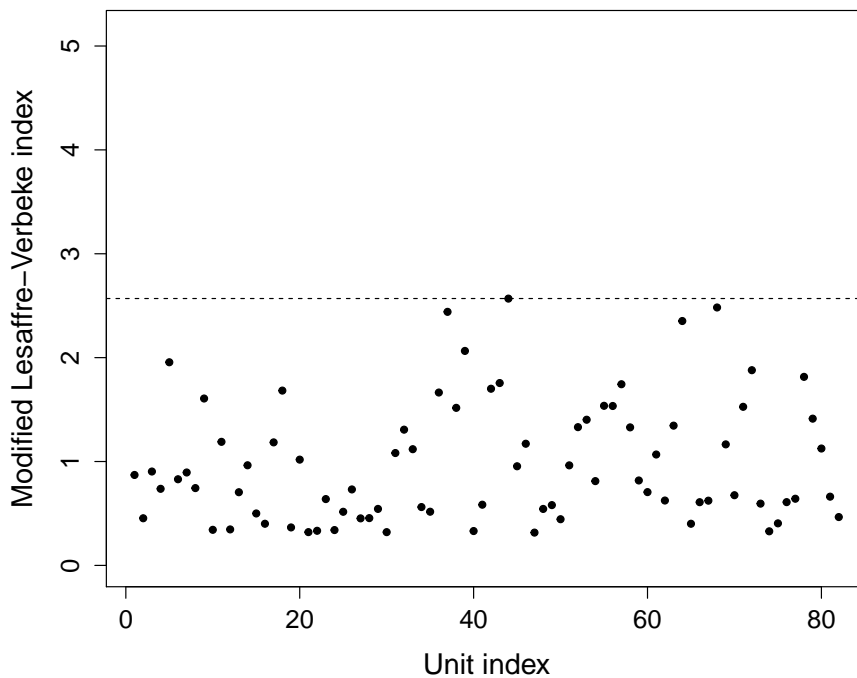


Figura 4.39: Gráfico de índice modificado de Lesaffre - Verbeke.

Analisando o gráfico dos resíduos marginais padronizados apresentado na Figura 4.40, observamos que esse gráfico não ilustra uma incorreta especificação na estrutura da média.

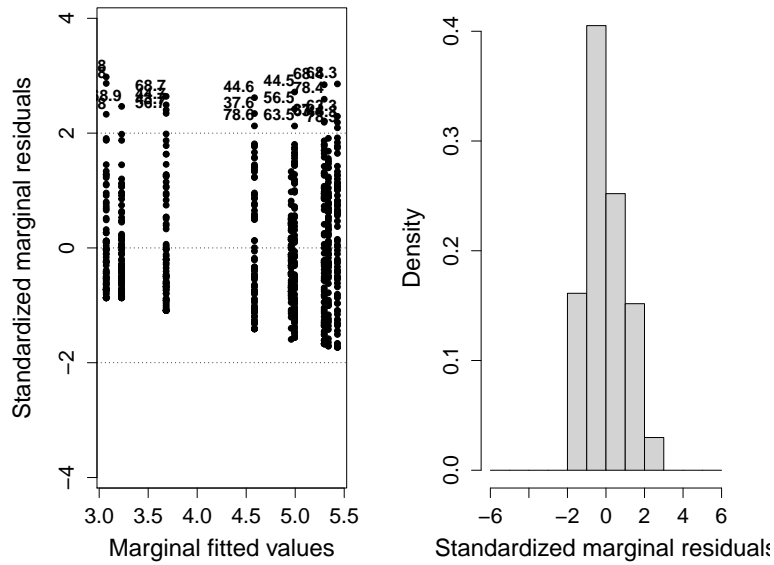


Figura 4.40: Gráfico dos resíduos marginais padronizados.

Os resultados apresentados na Figura 4.41 permitem evidenciar que a distribuição do vetor de efeitos aleatórios não tem um afastamento da distribuição normal.

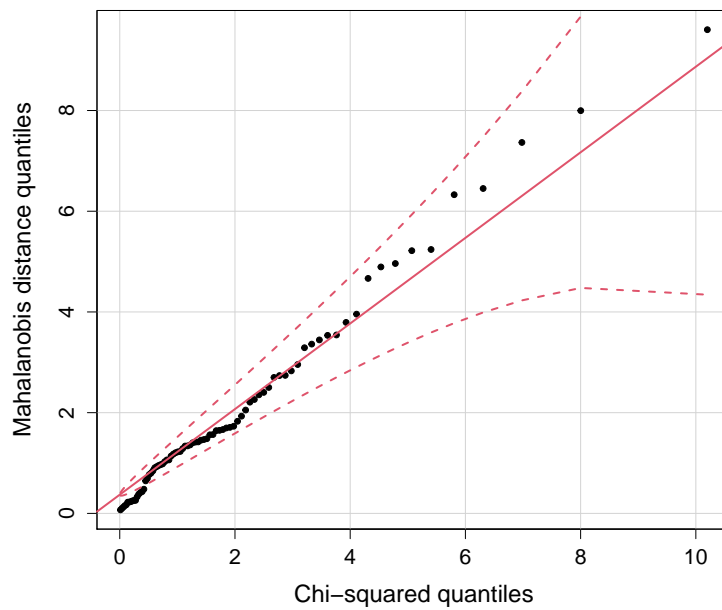


Figura 4.41: Gráfico da distância dos quantis de Mahalanobis.

Analisando a Figura 4.42 os indivíduos #43, #44, #57, #68 e #80 podem ser consideradas como observações outliers com respeito as restantes.

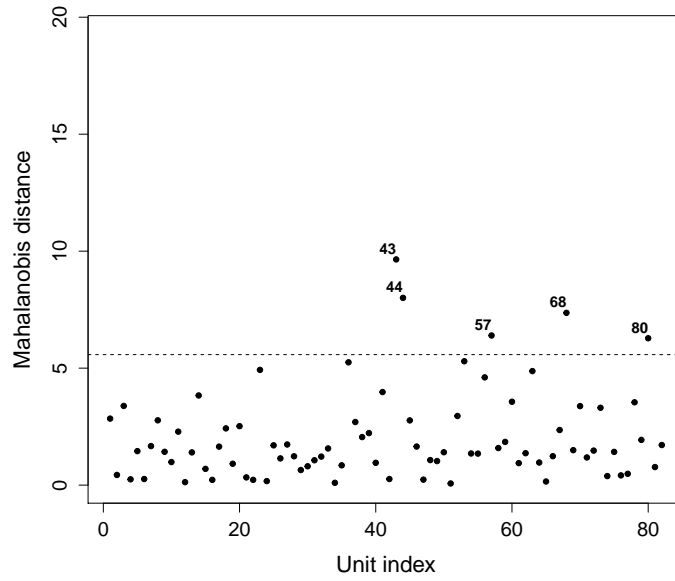


Figura 4.42: Gráfico da distância de Mahalanobis.

Pela Figura 4.43 é percebido plausibilidade da suposição de normalidade para a distribuição nos erros condicionais.

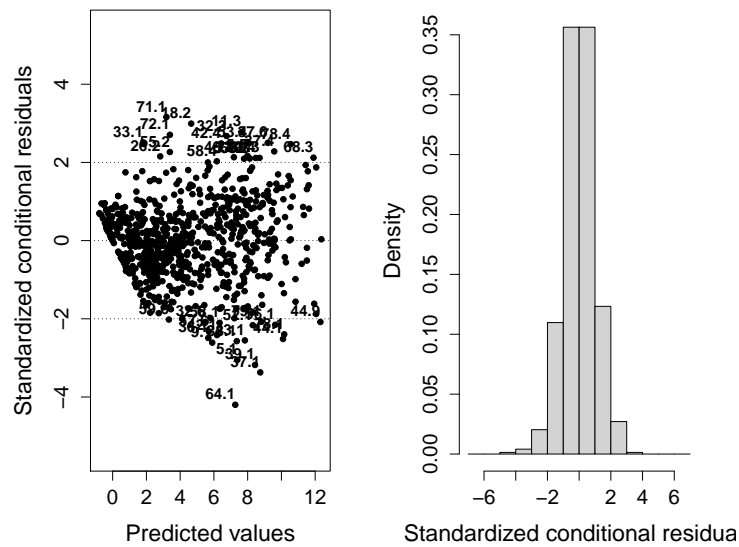


Figura 4.43: Gráfico dos resíduos condicionais padronizados.

Analisando a Figura 4.44, evidenciamos que a distribuição do vetor de erros aleatórios não tem um afastamento da distribuição normal.

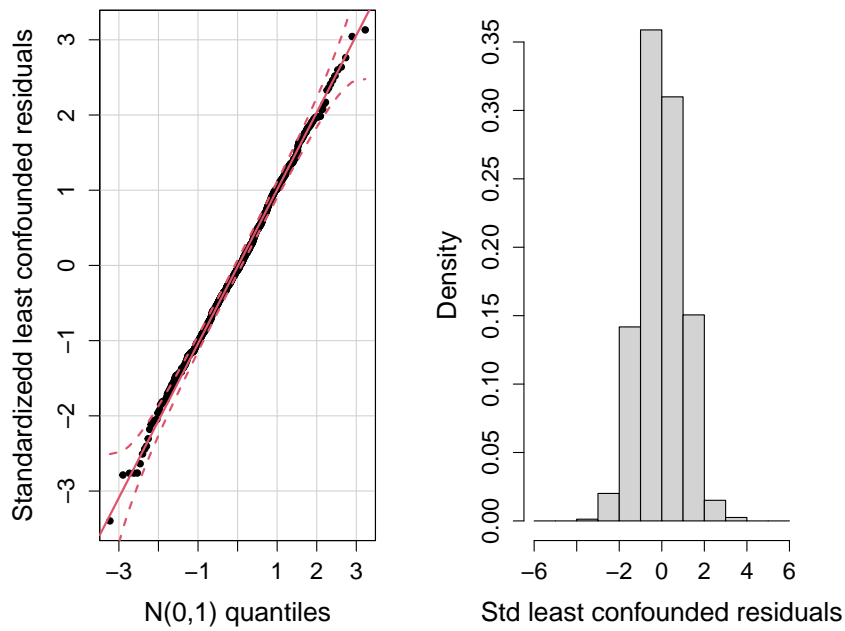


Figura 4.44: Gráfico *QQ* para os resíduos condicionais padronizados e histograma.

Com o fim de tentar modelar a autocorrelação ilustrada na Figura 4.38, propomos um modelo tal como em (4.3) com uma matriz de covariâncias não estruturada para o erro condicional, ou seja com uma estrutura de correlação geral e especificamos a forma dos termos das variâncias considerando uma variância para um grupo determinado e outra para o resto, ou seja, $e_{ij} \sim \mathcal{N}(0, \sigma_i^2)$ com

$$\sigma_i^2 = \tau^2, i = 3, 13, 56, 63 \text{ e } 80 \quad \text{e} \quad \sigma_i^2 = \sigma^2, \text{ em outro caso.}$$

Apresentamos na Tabela 4.15 os resultados do modelo ajustado.

Tabela 4.15: *Resultados do modelo ajustado*

Parâmetro	Estimativa (SE)
β_0	4.3166 (0.3112)
β_1	-0.2075 (0.0415)
β_2	-0.0500 (0.0104)
β_3	0.0075 (0.0014)
σ_0	1.6509
σ_1	4.67e-06
σ_{01}	2.4259
σ	1.2777
F. Variância	$(1.8334\sigma)^2$
AIC	2556.42
BIC	2763.60

Os resultados dos testes indicados na Tabela 4.16 ilustram que o modelo apresenta erro de especificação.

Tabela 4.16: *Estatísticas de cada teste e valor – p*

Estatística	valor	valor – p
<i>EAMI</i>	43.27	< 0.001
<i>ESDM</i>	35.12	< 0.001
<i>EESA</i>	22.58	< 0.001

Os diagnósticos gráficos ilustrados na continuação permitem detectar a fonte do erro. A Figura 4.45 não mostra evidência gráfica para suspeitar que os erros condicionais estejam correlacionados serialmente.

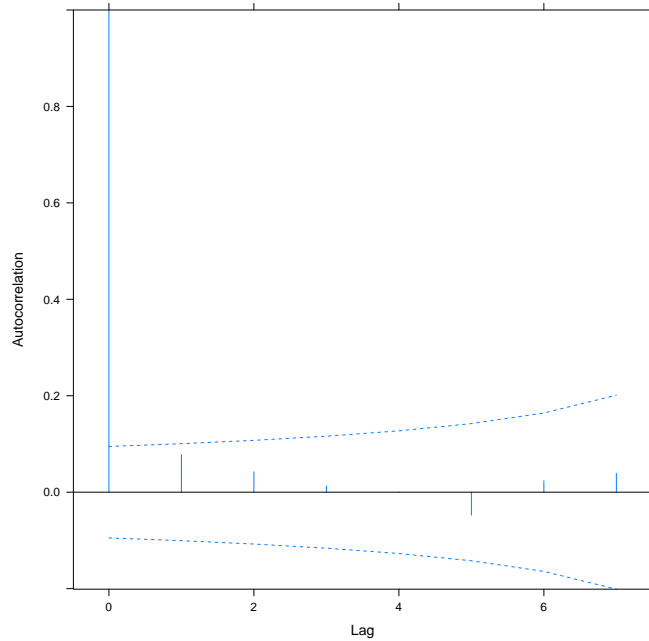


Figura 4.45: Gráfico da função de autocorrelação.

O gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 4.46, detecta que a matriz de covariâncias é inapropriada para os indivíduos #37, #44 e #53 e #68

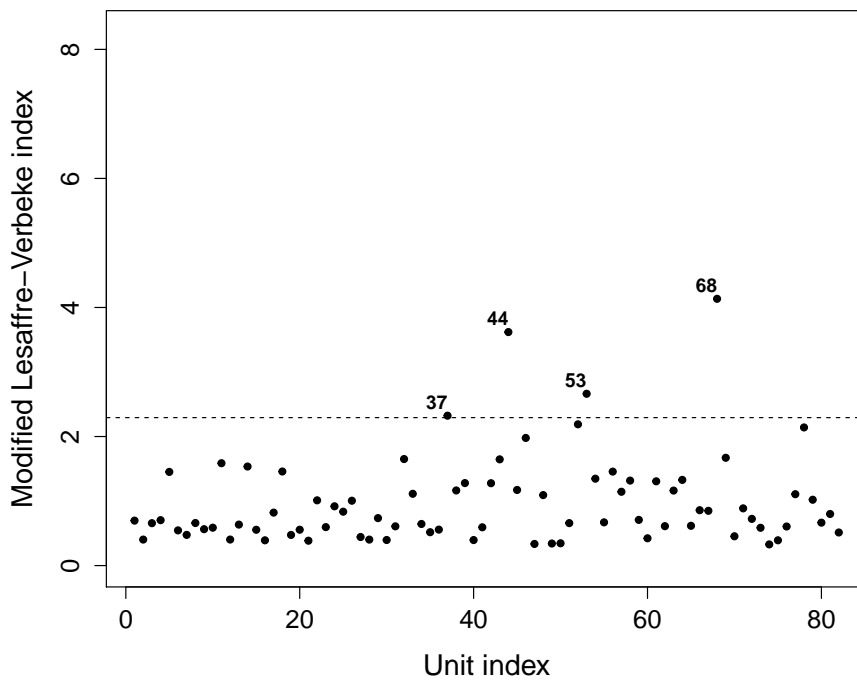


Figura 4.46: Gráfico de índice modificado de Lesaffre - Verbeke.

Analisando o gráfico dos resíduos marginais padronizados apresentado na Figura 4.47, não ilustra uma incorreta especificação na estrutura da matriz de covariâncias no vetor de

efeitos aleatórios.

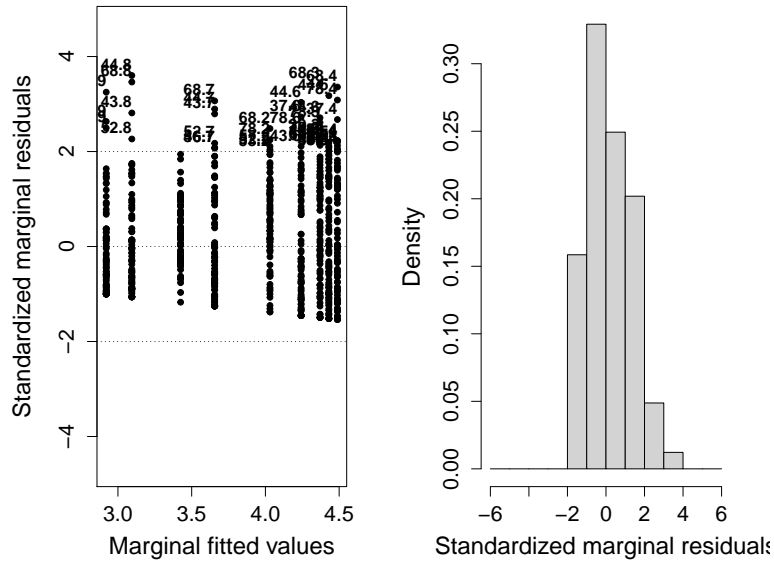


Figura 4.47: Gráfico dos resíduos marginais padronizados.

Os resultados apresentados na Figura 4.48 permitem evidenciar que a distribuição do vetor de efeitos aleatórios não tem um afastamento da distribuição normal.

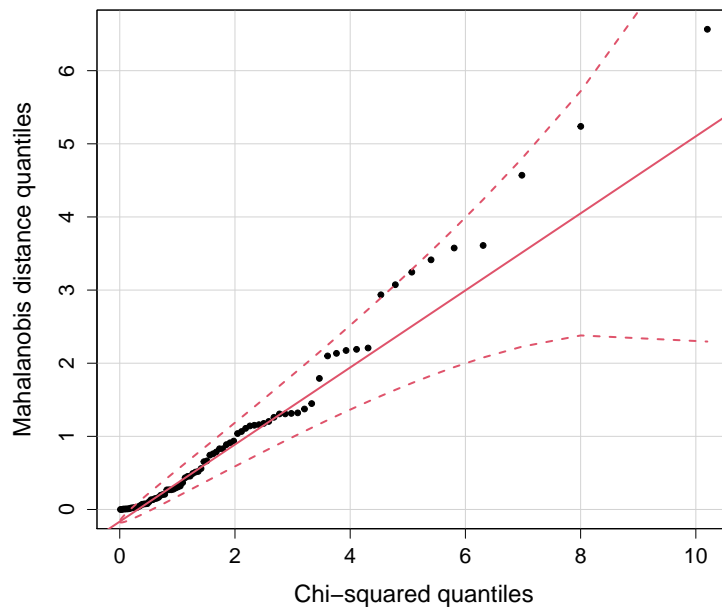


Figura 4.48: Gráfico da distância dos quantis de Mahalanobis.

A análise da Figura 4.49, sugere que as observações #10, #37, #41, #43, #44, #67, #68 e #78 podem ser analisadas como possíveis outliers comparadas com o resto.

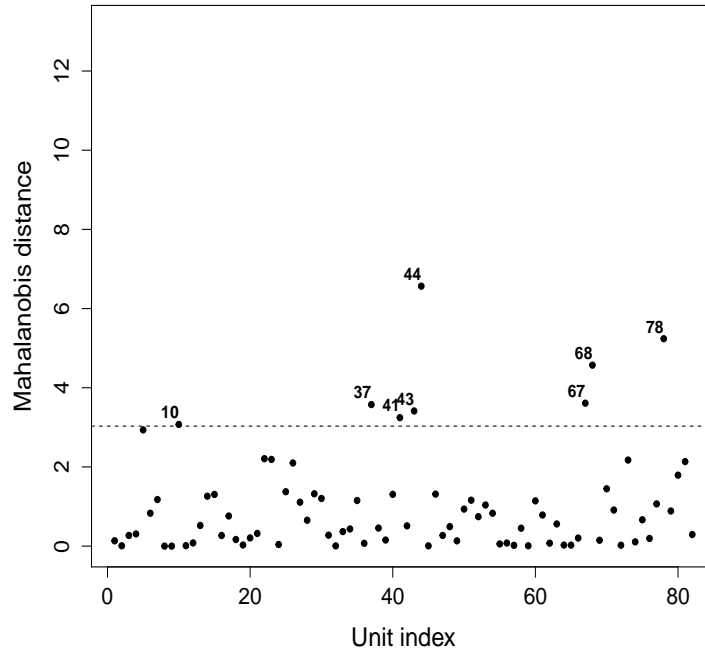


Figura 4.49: Gráfico da distância de Mahalanobis.

Pela Figura 4.50 é percebido plausibilidade da suposição de normalidade para a distribuição nos erros condicionais.

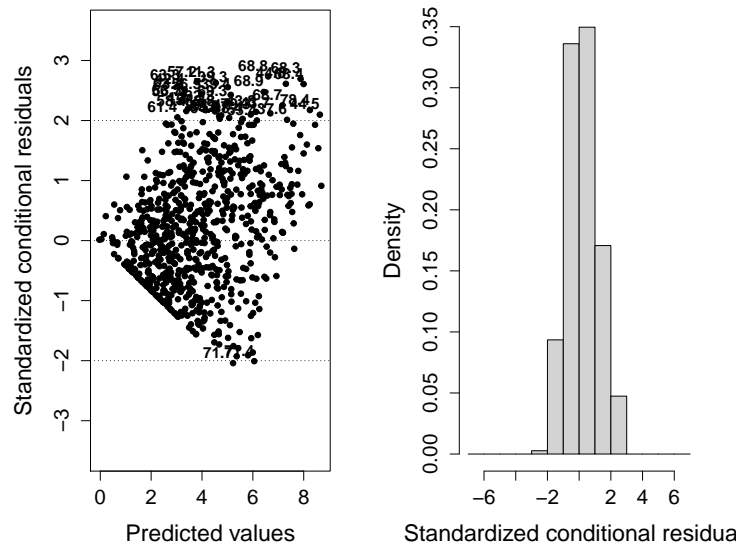


Figura 4.50: Gráfico dos resíduos condicionais padronizados.

Pela Figura 4.51 é percebido plausibilidade da suposição de normalidade para a distribuição nos erros condicionais.

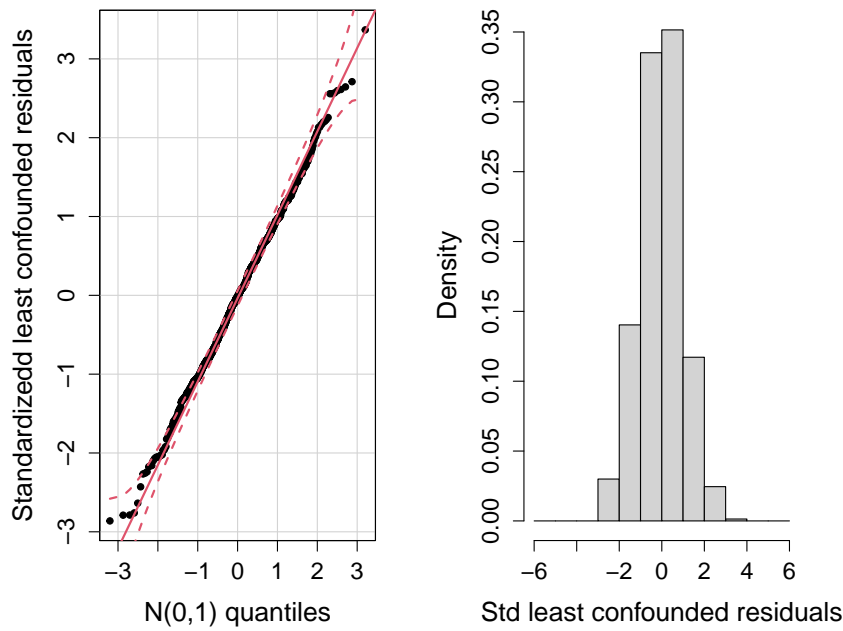


Figura 4.51: Gráfico *QQ* para os resíduos condicionais padronizados e histograma.

Nesta parte propomos um modelo tal como em (4.3) considerando uma matriz de covariâncias não estruturada para o erro condicional e considerando uma variância para un grupo determinado e outra para o resto, $e_{ij} \sim \mathcal{N}(0, \sigma_i^2)$; $\sigma_i^2 = \tau^2$, $i = 3, 13, 37, 44, 53, 56, 63, 68$ e 80 e $\sigma_i^2 = \sigma^2$, em outro caso.

Na Tabela 4.17 os resultados do modelo ajustado são ilustrados.

Tabela 4.17: Resultados do modelo ajustado

Parâmetro	Estimativa (SE)
β_0	3.9327 (0.2957)
β_1	-0.2328 (0.0429)
β_2	-0.0348 (0.0098)
β_3	0.0057 (0.0012)
σ_0	1.984375e-06
σ_1	3.734792e-05
σ_{01}	1.7034e-15
σ	1.2777
F. Variância	$(2.046\sigma)^2$
AIC	2504.41
BIC	2711.59

Os resultados dos testes são ilustrados na Tabela 4.18. Esses resultados indicam que o

modelo apresenta erro de especificação.

Tabela 4.18: *Estatísticas de cada teste e valor – p*

Estadística	valor	valor – p
<i>EAMI</i>	91.93	< 0.001
<i>ESDM</i>	76.17	< 0.001
<i>EESA</i>	51.16	< 0.001

Com os testes detectamos a fonte do erro de especificação. A Figura 4.59 não mostra evidência gráfica para suspeitar que os erros condicionais estejam correlacionados serialmente.

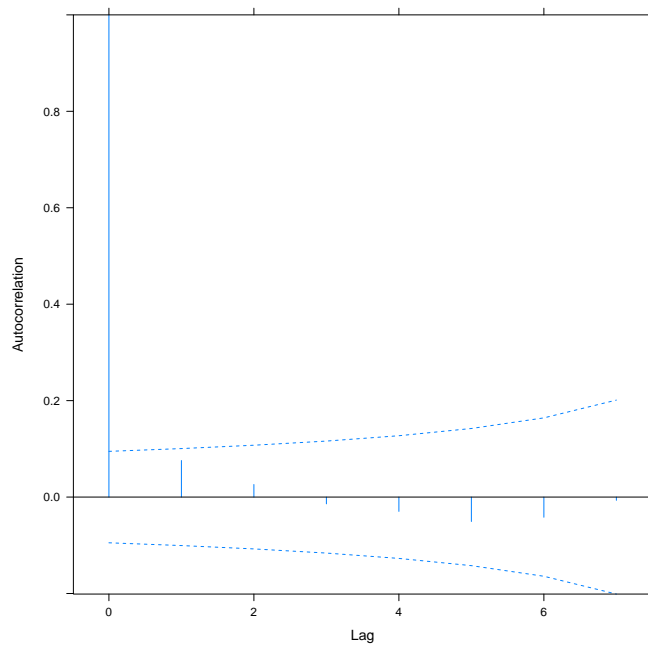


Figura 4.52: *Gráfico da função de autocorrelação.*

O gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 4.60, não detecta que a matriz de covariâncias proposta seja inapropriada.

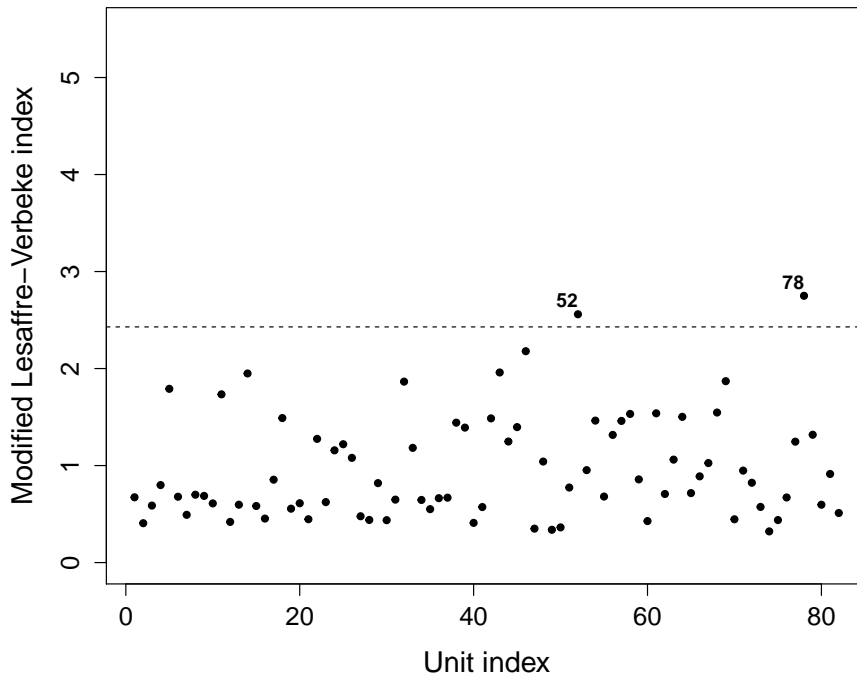


Figura 4.53: Gráfico de índice modificado de Lesaffre - Verbeke.

Analisando o gráfico dos resíduos marginais padronizados apresentado na Figura 4.54, observamos que esse gráfico não ilustra uma incorreta especificação na estrutura da média.

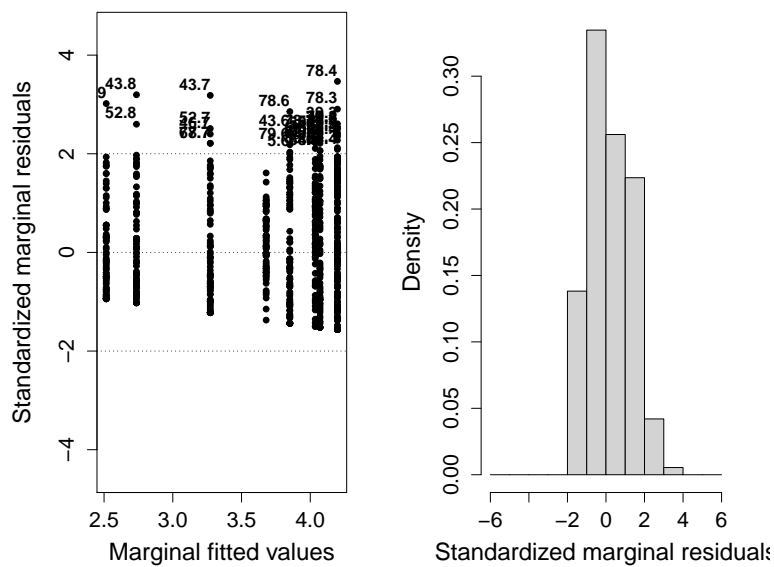


Figura 4.54: Gráfico dos resíduos marginais padronizados.

O resultado apresentado na Figura 4.55 permitem evidenciar que a distribuição do vetor de efeitos aleatórios não tem um afastamento da distribuição normal.

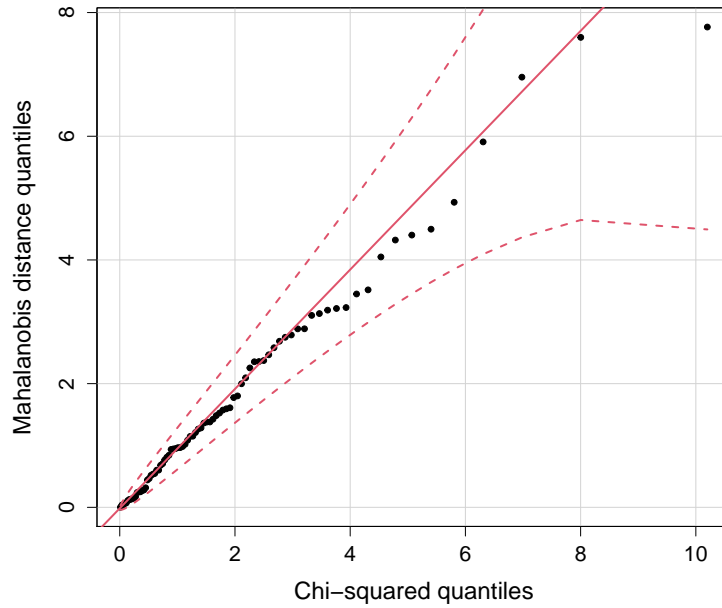


Figura 4.55: Gráfico da distância dos quantis de Mahalanobis.

A análise da Figura 4.56, sugere que as observações #14, #25, #43 e #78 podem ser analisadas como possíveis outliers comparadas com o resto.

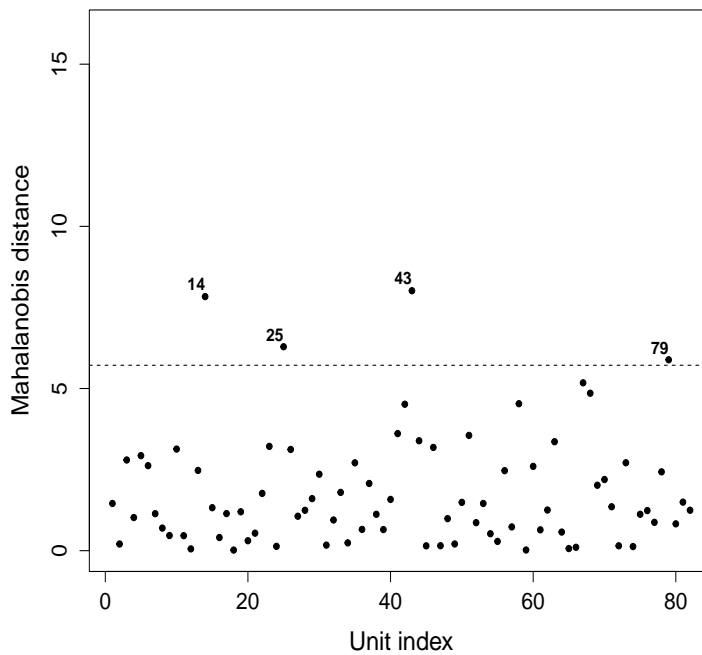


Figura 4.56: Gráfico da distância de Mahalanobis.

Pela Figura 4.57 é percebido plausibilidade da suposição de normalidade para a distribuição nos erros condicionais.

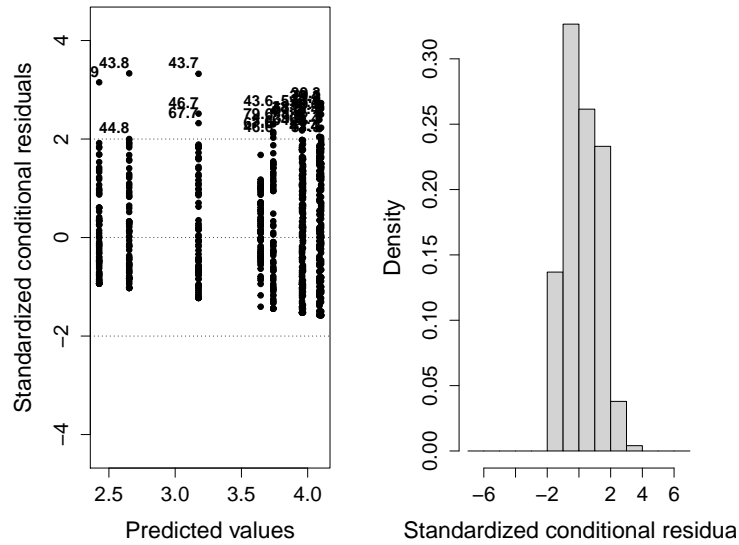


Figura 4.57: Gráfico dos resíduos condicionais padronizados.

Analisando a Figura 4.58, evidenciamos que a distribuição do vetor de erros aleatórios não tem um afastamento da distribuição normal.

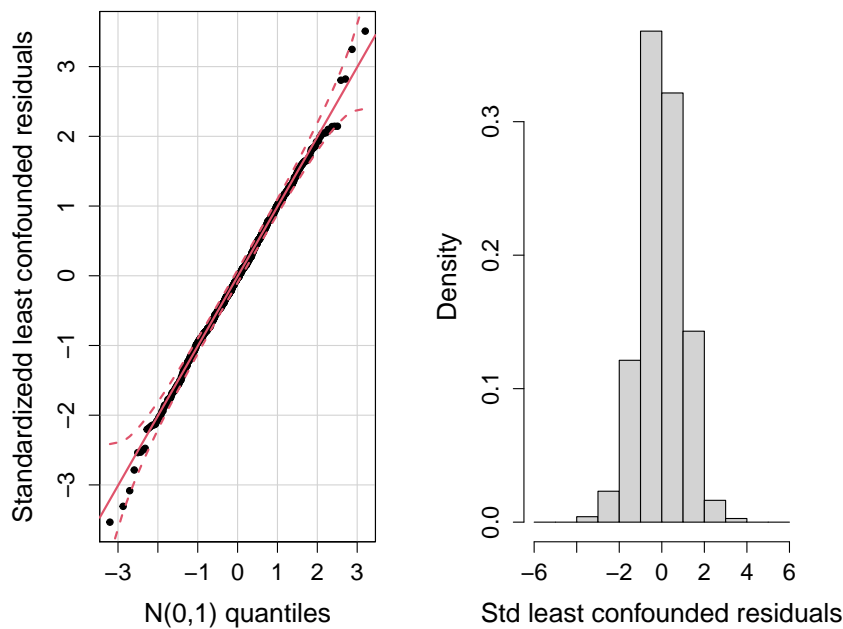


Figura 4.58: Gráfico QQ para os resíduos condicionais padronizados e histograma.

Finalmente considerando uma matriz de covariâncias não estruturada para o erro condicional e considerando uma variância para um grupo determinado e outra para o resto. Nesse caso, fazemos $e_{ij} \sim \mathcal{N}(0, \sigma_i^2)$; em que $\sigma_i^2 = \tau^2$, $i = 3, 13, 37, 44, 52, 53, 56, 63, 68, 78$ e 80 e

$\sigma_i^2 = \sigma^2$, em outro caso.

Na Tabela 4.19 os resultados do modelo ajustado são ilustrados.

Tabela 4.19: *Resultados do modelo ajustado*

Parâmetro	Estimativa (SE)
β_0	3.8203 (0.2870)
β_1	-0.2295 (0.0424)
β_2	-0.0323 (0.0096)
β_3	0.0053 (0.0012)
σ_0	1.966189e-06
σ_1	3.462905e-05
σ_{01}	1.7034e-15
σ	2.6090
F. Variância	$(2.098\sigma)^2$
AIC	2488.70
BIC	2695.88

Os resultados dos testes na Tabela 4.20, indicam que o modelo não apresenta erro de especificação.

Tabela 4.20: *Estatísticas de cada teste e valor – p*

Estatística	valor	valor – p
<i>EAMI</i>	1.40	0.99
<i>ESDM</i>	7.12	0.52
<i>EESA</i>	4.26	0.83

Dado que os testes não detectam erro de especificação, usamos os diagnósticos gráficos para ilustrar esse resultado.

A Figura 4.59 não mostra evidência gráfica para suspeitar que os erros condicionais estejam correlacionados serialmente.

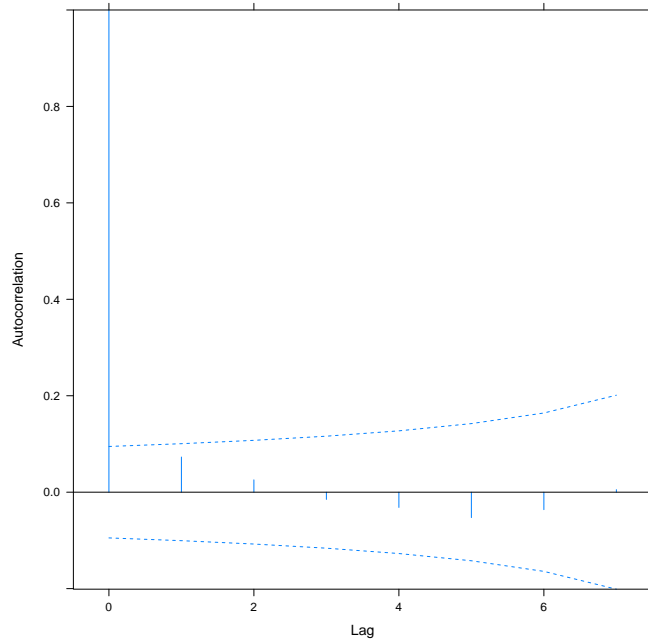


Figura 4.59: Gráfico da função de autocorrelação.

O gráfico do índice modificado de Lesaffre–Verbeke exibido na Figura 4.60, não detecta que a matriz de covariâncias proposta seja inapropriada.

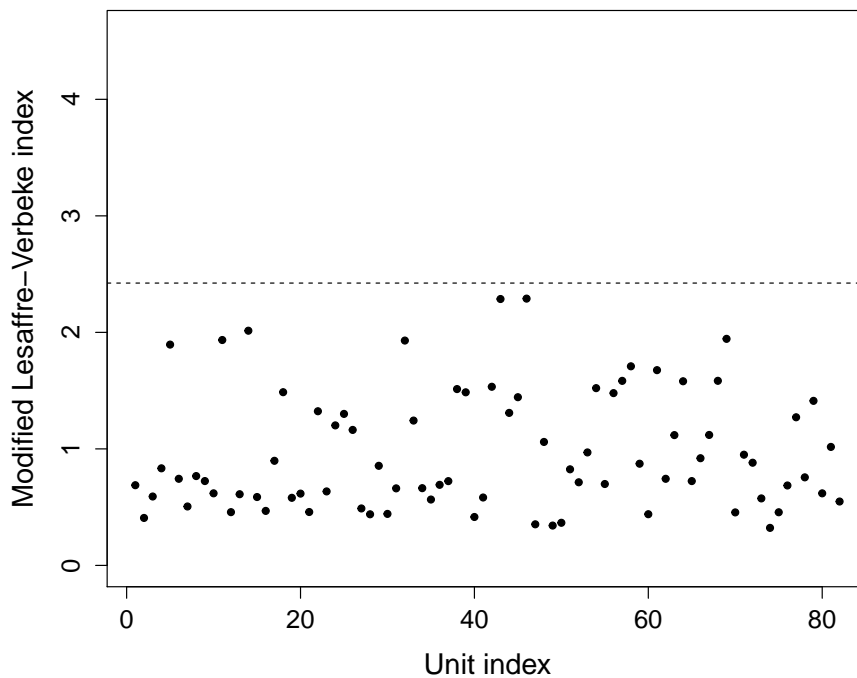


Figura 4.60: Gráfico de índice modificado de Lesaffre - Verbeke.

Analisando o gráfico dos resíduos marginais padronizados apresentado na Figura 4.61, observamos que esse gráfico não ilustra uma incorreta especificação na estrutura da média.

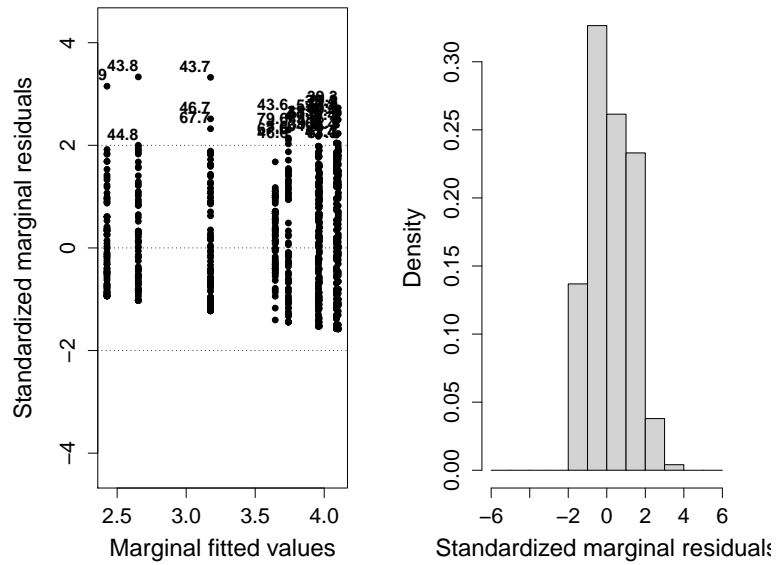


Figura 4.61: Gráfico dos resíduos marginais padronizados.

O resultado apresentados na Figura 4.62 permitem evidenciar que a distribuição do vetor de efeitos aleatórios não tem um afastamento da distribuição normal.

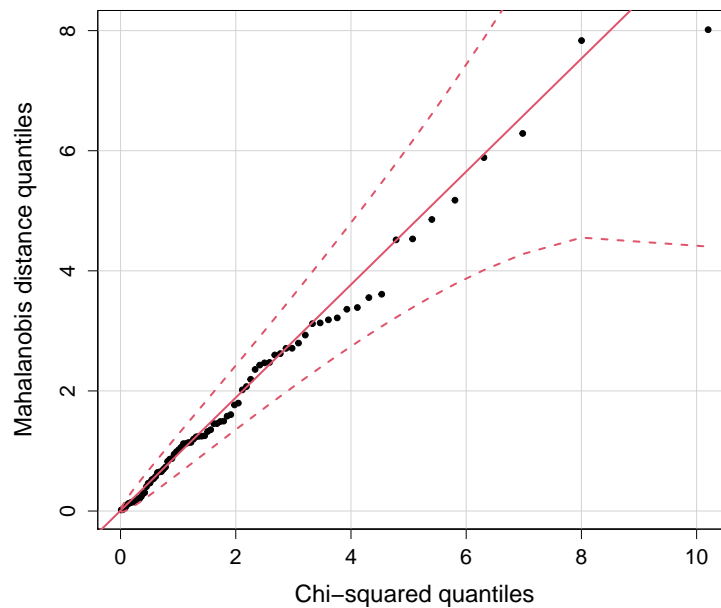


Figura 4.62: Gráfico da distância dos quantis de Mahalanobis.

A análise da Figura 4.63, sugere que as observações #14, #25, #43 e #79 podem ser analisadas como observações outliers comparadas com o resto.

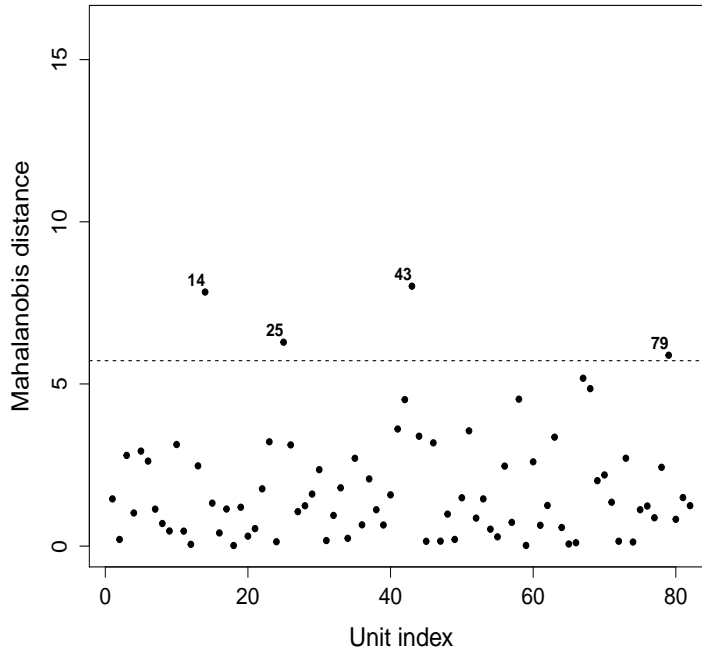


Figura 4.63: Gráfico da distância de Mahalanobis.

Pela Figura 4.64 é percebido plausibilidade da suposição de normalidade para a distribuição nos erros condicionais.

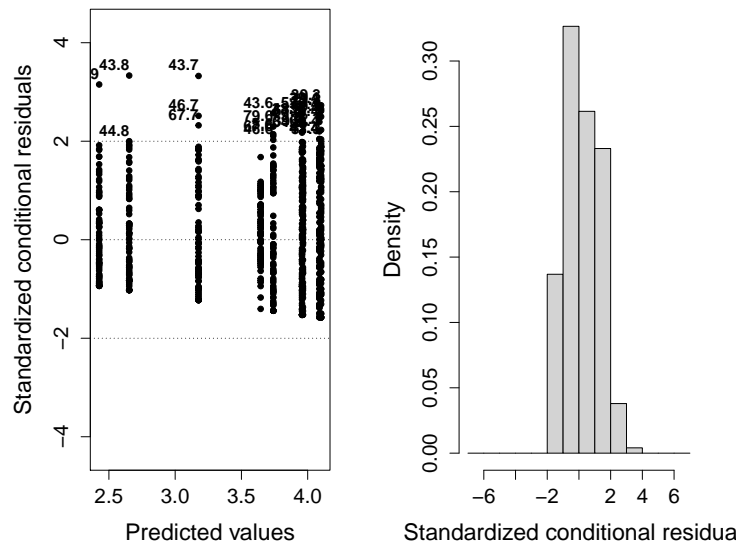


Figura 4.64: Gráfico dos resíduos condicionais padronizados.

Analisando a Figura 4.65 detectamos que o histograma dos resíduos minimamente confundidos tem um comportamento normal, situação que é evidenciada também no gráfico de QQ, ou seja que a distribuição do vetor de erros aleatórios não tem um afastamento da

distribuição normal.

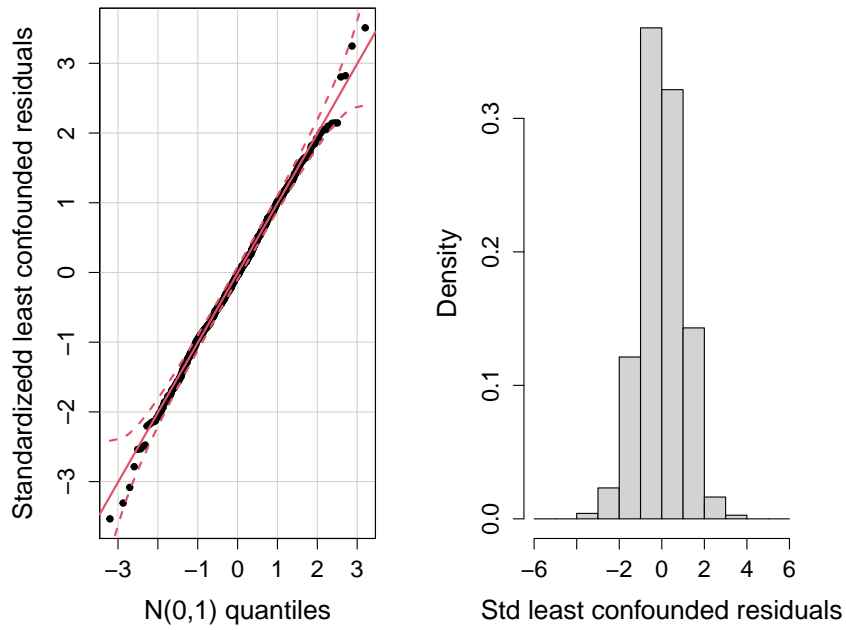


Figura 4.65: Gráfico QQ para os resíduos condicionais padronizados e histograma.

Capítulo 5

Conclusão e estudos futuros

5.1 Conclusão

- Os resultados do processo de simulação e das aplicações ilustram que os testes detectam erro de especificação num modelo linear misto gaussiano, e que os diagnósticos gráficos identificam a fontes dos erros.
- Os testes e os diagnósticos gráficos identificam erros de especificação em modelos lineares mistos gaussianos.

5.2 Estudos futuros

- Avaliar um modelo linear misto usando os testes e os diagnósticos gráficos quando:
 - O vetor de efeitos aleatórios tem uma distribuição diferente à gaussiana.
 - O vetor de erros aleatórios tem uma distribuição diferente à gaussiana.
 - As duas situações anteriores.
- Disponibilizar os códigos desenvolvidos num pacote de R.

Apêndice A

A.1 Derivadas do logaritmo da função de verossimilhança

Tendo em conta a função densidade de probabilidade para a distribuição do vetor aleatório (2.3), são apresentados os cálculos das primeira, segunda e terceira do logaritmo da função de verossimilhança do vetor $\boldsymbol{\phi}$. Estas derivadas foram usadas para a construção dos testes desenvolvidos neste trabalho.

$$f(\mathbf{y}_i, \boldsymbol{\phi}) = (2\pi)^{-\frac{m_i}{2}} |\mathbf{V}_i|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right\}, \quad (\text{A.1})$$

em que, $\mathbf{e}_i = \mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}$ e $\mathbf{V}_i = \mathbf{Z}_i \mathbf{G} \mathbf{Z}_i^\top + \sigma^2 \mathbf{I}_{m_i}$.

A função de log-verossimilhança é

$$l_i(\boldsymbol{\phi}, \mathbf{y}_i) = \log f(\mathbf{y}_i, \boldsymbol{\phi}) = -\frac{m_i}{2} \log(2\pi) - \frac{1}{2} \log |\mathbf{V}_i| - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{e}_i. \quad (\text{A.2})$$

A.2 Derivadas de primeira ordem

A.2.1 Derivada do vetor escore

$$\nabla \log f(\mathbf{y}_i, \boldsymbol{\phi}) = \frac{\partial \log f(\mathbf{y}_i, \boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\phi}} = \begin{pmatrix} \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta} \\ \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\theta}} \\ \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2} \end{pmatrix}, \quad (\text{A.3})$$

em que, $\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta}$ ($p \times 1$), $\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\theta}}$ ($q \times 1$) e $\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2}$ (1×1).

A.2.2 Derivada primeira com relação ao vetor β

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta} = -\frac{1}{2} \frac{\partial}{\partial \beta^\top} (\mathbf{e}_i^\top) \frac{\partial}{\partial \mathbf{e}_i} (\mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{e}_i), \quad (\text{A.4})$$

$$= \frac{1}{2} (\mathbf{X}_i^\top) (2\mathbf{V}_i^{-1} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{e}_i, \text{Demidenko (2013, p.174)}. \quad (\text{A.5})$$

$$\therefore \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta} = \begin{bmatrix} \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_0} \\ \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1} \end{bmatrix} = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{e}_i.$$

A.2.3 Derivada primeira com relação aos componentes do vetor $\boldsymbol{\theta}$

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_1} \\ \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_2} \\ \vdots \\ \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_q} \end{bmatrix}. \quad (\text{A.6})$$

O gradiente com respeito ao elemento k -ésimo, $k = 1, \dots, q$, da entrada do vetor $\boldsymbol{\theta}$, [Stroup \(2012, p.136-137\)](#) é

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_k} = -\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i. \quad (\text{A.7})$$

A.2.4 Derivada primeira com relação a σ^2

$$\frac{\partial l_i(\boldsymbol{\phi})}{\partial \sigma^2} = -\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1}) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i.$$

A.3 Derivadas de segunda ordem

A.3.1 Derivada segunda do vetor $\boldsymbol{\beta}^\top$ com relação ao vetor $\boldsymbol{\beta}$

$$\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} = \frac{\partial}{\partial \boldsymbol{\beta}} (\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{e}_i) = -\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i \quad (m_i \times m_i).$$

A.3.2 Derivada segunda dos componentes do vetor $\boldsymbol{\theta}$ com relação ao vetor $\boldsymbol{\beta}$

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\beta}} \left[\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\theta}^\top} \right] &= \frac{\partial}{\partial \boldsymbol{\beta}} \left[-\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \right], \\ &= \frac{\partial}{\partial \boldsymbol{\beta}} \left[\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \right] = -\mathbf{X}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i, \end{aligned}$$

$$k = 1, \dots, q.$$

A.3.3 Derivada segunda de σ^2 com relação ao vetor $\boldsymbol{\beta}$

$$\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \sigma^2} = -\mathbf{X}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \sigma^2} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \quad (p \times 1), \quad [\text{(Petersen \& Pedersen, 2012, p.11 ,Eq. (86))}].$$

A.3.4 Derivada segunda de cada componente do vetor θ com relação a σ^2

$$\begin{aligned} \frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \sigma^2 \partial \theta_k} &= \frac{\partial}{\partial \sigma^2} \left[-\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \right] \\ &= \frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-2} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left\{ \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) + \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \right\} \mathbf{V}_i^{-1} \mathbf{e}_i, \end{aligned} \quad (\text{A.8})$$

$k = 1, \dots, q.$

A.3.5 Derivada segunda da função logaritmo da verossimilhança dos componentes do vetor θ^\top

$$\begin{aligned} \frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \theta_{k_1} \partial \theta_{k_2}} &= \frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial^2 \mathbf{V}_i}{\partial \theta_{k_1} \partial \theta_{k_2}} + \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_1}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_2}} \right) \\ &+ \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left\{ \frac{\partial^2 \mathbf{V}_i}{\partial \theta_{k_1} \partial \theta_{k_2}} - \frac{\partial \mathbf{V}_i}{\partial \theta_{k_1}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_2}} - \frac{\partial \mathbf{V}_i}{\partial \theta_{k_2}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_1}} \right\} \mathbf{V}_i^{-1} \mathbf{e}_i \end{aligned} \quad (\text{A.9})$$

$$= \frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_1}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_2}} \right) - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left\{ \frac{\partial \mathbf{V}_i}{\partial \theta_{k_1}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_2}} + \frac{\partial \mathbf{V}_i}{\partial \theta_{k_2}} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_{k_1}} \right\} \mathbf{e}_i \mathbf{V}_i^{-1}, \quad (\text{A.10})$$

em que $\theta_{k_1} = 1, \dots, q$ e $\theta_{k_2} = 1, \dots, q.$

A.3.6 Derivada segunda de σ^2 com relação a σ^2

$$\begin{aligned} \frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma^2} &= \frac{\partial}{\partial \sigma^2} \left[-\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1}) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i \right] \\ &= \frac{1}{2} \text{tr} (\mathbf{V}_i^{-2}) - \mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i. \end{aligned} \quad (\text{A.11})$$

A.4 Derivada terceira ordem

A.4.1 Derivada terceira com respeito a cada componente do vetor

β

Nota-se que: para todo $t, k = 1, \dots, s$, $\frac{\partial}{\partial \beta_t} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta \partial \beta^\top} \right) = \frac{\partial}{\partial \beta_t} (-\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i) = \mathbf{0}$, pelo

que:

$$\frac{\partial}{\partial \beta_t} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_k^2} \right) = 0.$$

$$\begin{aligned} \frac{\partial d_k(\mathbf{y}_i, \phi)}{\partial \phi_l} &= \frac{\partial}{\partial \phi_l} \left[\left\{ \frac{\partial \log f(\mathbf{y}_i, \phi)}{\partial \phi_k} \right\}^2 + \frac{\partial^2 \log f(\mathbf{y}_i, \phi)}{\partial \phi_k^2} \right] \\ &= 2 \frac{\partial \log f(\mathbf{y}_i, \phi)}{\partial \phi_k} \frac{\partial^2 \log f(\mathbf{y}_i, \phi)}{\partial \phi_l \partial \phi_k} + \frac{\partial^3 \log f(\mathbf{y}_i, \phi)}{\partial \phi_l (\partial \phi_k)^2}, \end{aligned} \quad (\text{A.12})$$

(A.12) são elementos da matriz, $\nabla \mathbf{D}_n(\phi, \mathbf{y}_i) = \frac{1}{n} \sum_{i=1}^n \left[\frac{\partial d(\mathbf{y}_i, \phi)}{\partial \phi} \right]$,

$$\nabla \mathbf{D}_n(\phi, \mathbf{y}_i) = \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} \frac{\partial d_1}{\partial \beta_0} & \frac{\partial d_1}{\partial \beta_1} & \cdots & \frac{\partial d_1}{\partial \sigma_a^2} & \frac{\partial d_1}{\partial \sigma_b^2} & \cdots & \frac{\partial d_1}{\partial \sigma^2} \\ \frac{\partial d_2}{\partial \beta_0} & \frac{\partial d_2}{\partial \beta_1} & \cdots & \frac{\partial d_2}{\partial \sigma_a^2} & \frac{\partial d_2}{\partial \sigma_b^2} & \cdots & \frac{\partial d_2}{\partial \sigma^2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial d_s}{\partial \beta_0} & \frac{\partial d_s}{\partial \beta_1} & \cdots & \frac{\partial d_s}{\partial \sigma_a^2} & \frac{\partial d_s}{\partial \sigma_b^2} & \cdots & \frac{\partial d_s}{\partial \sigma^2} \end{bmatrix}.$$

Tendo em conta a expressão dois na equação (A.12), as terceiras derivadas parciais são:

$$\begin{bmatrix}
 \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \beta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \beta_p^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \theta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \theta_q^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 (\partial \sigma^2)^2} \\
 \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_p \partial \beta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_p \partial \beta_p^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_p \partial \theta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_p \partial \theta_q^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_p (\partial \sigma^2)^2} \\
 \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_1 \partial \beta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_1 \partial \beta_p^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_1 \partial \theta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_1 \partial \theta_q^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_1 (\partial \sigma^2)^2} \\
 \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
 \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_q \partial \beta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_q \partial \beta_p^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_q \partial \theta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_q \partial \theta_q^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_q (\partial \sigma^2)^2} \\
 \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2 \partial \beta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2 \partial \beta_p^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2 \partial \theta_1^2} & \dots & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2 \partial \theta_q^2} & \frac{\partial^3 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma^2 (\partial \sigma^2)^2}
 \end{bmatrix} \quad (s \times s)$$

A.4.2 Derivada com relação aos elementos do vetor $\boldsymbol{\beta}$, da derivada segunda dos elementos do vetor $\boldsymbol{\theta}$

Note-se que,

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_k} = -\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i$$

então,

$$\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_k^2} = \frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) - \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i$$

assim,

$$\begin{aligned}
 \frac{\partial}{\partial \beta_t} \left[\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_k^2} \right] &= \frac{\partial}{\partial \beta_t} \left[\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) - \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \right], \\
 &= \mathbf{X}_{it}^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i + \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{X}_{it}^\top,
 \end{aligned} \tag{A.13}$$

em que \mathbf{X}_{it}^\top representa a t -ésima coluna da matriz \mathbf{X}_i .

A.4.3 Derivada com relação a β_t da derivada segunda de σ^2

$$\begin{aligned} \frac{\partial}{\partial \beta_t} \left[\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{(\partial \sigma^2)^2} \right] &= \frac{\partial}{\partial \beta_t} \left[\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-2} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) - \mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i \right], t = 1, \dots, p. \\ &= \frac{\partial}{\partial \beta_t} [-\mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i] = \mathbf{X}_{it}^\top \mathbf{V}_i^{-3} \mathbf{e}_i + \mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{X}_{it}^\top. \end{aligned} \quad (\text{A.14})$$

A.4.4 Derivada terceira com relação aos elementos do vetor $\boldsymbol{\theta}$

$$\frac{\partial}{\partial \theta_l} \left[\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \theta_k^2} \right] = \frac{\partial}{\partial \theta_l} \left[\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right) - \mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \right], \quad (\text{A.15})$$

a primeira expressão em (A.15) é

$$\frac{\partial}{\partial \theta_l} \left[\text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \right] = -\text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} + \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \right). \quad (\text{A.16})$$

Analisando a expressão dois em (A.15), tendo em conta que, $\frac{\partial}{\partial \theta_l} \left[\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right] = \mathbf{0}$ temos,

$$\begin{aligned} &\frac{\partial}{\partial \theta_l} \left[\mathbf{e}_i^\top \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \left(\frac{\partial \mathbf{V}_i}{\partial \theta_k} \right) \mathbf{V}_i^{-1} \mathbf{e}_i \right], \\ &= -\mathbf{e}_i^\top \mathbf{V}_i^{-1} \left\{ \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} + \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} + \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_k} \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \right\} \mathbf{V}_i^{-1} \mathbf{e}_i. \end{aligned} \quad (\text{A.17})$$

A.4.5 Derivada da derivada segunda do logaritmo da função de verossimilhança de σ^2

$$\frac{\partial}{\partial \theta_l} \left[\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{(\partial \sigma^2)^2} \right] = \frac{\partial}{\partial \theta_l} \left[\frac{1}{2} \text{tr} (\mathbf{V}_i^{-2}) - \mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i \right]. \quad (\text{A.18})$$

Trabalhando com a primeira expressão em (A.18) temos:

$$\frac{\partial}{\partial \theta_l} \left[\frac{1}{2} \text{tr} (\mathbf{V}_i^{-2}) \right] = -\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-2} + \mathbf{V}_i^{-2} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-1} \right). \quad (\text{A.19})$$

Analisando a segunda expressão em (A.18),

$$\frac{\partial}{\partial \theta_l} [\mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i] = -\mathbf{e}_i^\top \mathbf{V}_i^{-1} \left\{ \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-2} + \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \mathbf{V}_i^{-1} + \mathbf{V}_i^{-2} \frac{\partial \mathbf{V}_i}{\partial \theta_l} \right\} \mathbf{V}_i^{-1} \mathbf{e}_i. \quad (\text{A.20})$$

A.4.6 Derivada com relação a σ^2 da derivada segunda do logaritmo da verossimilhança de β_k^2

$$\begin{aligned} \frac{\partial}{\partial \sigma^2} \left[\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_k^2} \right] &= \text{diag} \left\{ \frac{\partial}{\partial \sigma^2} \left(\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta \partial \beta^\top} \right) \right\} = \text{diag} \left\{ \frac{\partial}{\partial \sigma^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i) \right\} \\ &= \text{diag} \{ (\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_i) \}. \end{aligned} \quad (\text{A.21})$$

A.4.7 Derivada com relação a σ^2 , da derivada segunda do logaritmo da verossimilhança de σ^2

$$\begin{aligned} \frac{\partial}{\partial \sigma^2} \left[\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{(\partial \sigma^2)^2} \right] &= \frac{\partial}{\partial \sigma^2} \left[\frac{1}{2} \text{tr} (\mathbf{V}_i^{-2}) - \mathbf{e}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i \right] \\ &= -\text{tr}(\mathbf{V}_i^{-3}) + 3\mathbf{e}_i^\top \mathbf{V}_i^{-4} \mathbf{e}_i. \end{aligned} \quad (\text{A.22})$$

A.5 Apêndice B

Expressões necessárias para os testes propostos

A.5.1 Matrizes

Tendo em conta (2.1) a função logaritmo da verossimilhança para o i -ésimo indivíduo no modelo marginal é

$$l_i(\boldsymbol{\phi}, \mathbf{y}_i) = -\frac{m_i}{2} \log(2\pi) - \frac{1}{2} \log |V_i(\boldsymbol{\theta})| - \frac{1}{2} \left\{ (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta})^\top V_i(\boldsymbol{\theta})^{-1} (\mathbf{y}_i - \mathbf{X}_i \boldsymbol{\beta}) \right\}. \quad (\text{A.23})$$

De acordo com (A.23), a função logaritmo da verossimilhança para o vetor $\boldsymbol{\phi}$ é dada por

$$l(\boldsymbol{\phi}, \mathbf{y}) = \sum_{i=1}^n \log f(\mathbf{y}_i, \boldsymbol{\phi}) = \sum_{i=1}^n l_i(\boldsymbol{\phi}, \mathbf{y}_i). \quad (\text{A.24})$$

Tendo em conta algumas propriedades matriciais [Searle (1982)], $\mathbf{A}_n(\boldsymbol{\phi})$ e $\mathbf{B}_n(\boldsymbol{\phi})$ dadas por

$$\mathbf{A}_n(\boldsymbol{\phi}) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial^2 \log f(\mathbf{y}_i, \boldsymbol{\phi})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^\top} \right) = \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^\top} \right) \text{ e}$$

$$\mathbf{B}_n(\boldsymbol{\phi}) = \sum_{i=1}^n \left(\frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_k} \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \phi_l} \right); k, l = 1, \dots, s,$$

em que é considerado

$$\frac{\partial^2 l_i(\boldsymbol{\phi})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^\top} (s \times s) = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13} \\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23} \\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta \partial \beta^\top} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta \partial \boldsymbol{\theta}^\top} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta \partial \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \beta^\top} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \beta^\top} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \boldsymbol{\theta}; \mathbf{y}_i} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma^2} \end{pmatrix}, \quad (\text{A.25})$$

$$\mathbf{A}_n(\boldsymbol{\phi}) = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \beta_0)^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \sigma_1^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \beta_1)^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_1 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_1 \sigma_1^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \beta_0 \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_0^2 \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \beta_1} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \sigma_0)^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_1^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_1^2 \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_1^2 \partial \beta_1} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \sigma_1)^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_{01} \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_{01} \partial \beta_1} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_{01} \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma_{01} \partial \sigma_1^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \sigma_{01})^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \sigma^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \beta_0} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\sigma^2 \partial \beta_1} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma_1^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \sigma)^2} \end{pmatrix},$$

e

$$\frac{\partial \mathbf{A}_n(\phi)}{\partial \phi} = \frac{\partial}{\partial \phi_k} \begin{pmatrix} \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta \partial \beta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta \partial \boldsymbol{\theta}^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta \partial \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \beta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \beta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \boldsymbol{\theta}} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma^2} \end{pmatrix}, k = 1, \dots, s. \quad (\text{A.26})$$

Devemos encontrar, para todo k

$$\frac{\partial}{\partial \phi_k} \left(\frac{\partial l_i(\phi, \mathbf{y}_i)}{\partial \beta \partial \beta^\top} \right) = 0,$$

$k = 1, 2, 3.$

Note-se que, $\frac{\partial \mathbf{V}_i}{\partial \sigma_0^2} = \mathbf{Z}_0 \mathbf{Z}_0^\top$, $\frac{\partial \mathbf{V}_i}{\partial \sigma_{01}} = \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top$ e $\frac{\partial \mathbf{V}_i}{\partial \sigma_1^2} = \mathbf{Z}_1 \mathbf{Z}_1^\top$.

Por outro lado, para $k = 4$,

$$\mathbf{A}_n(\phi) = \frac{1}{n} \sum_{i=1}^n \begin{pmatrix} \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \beta_0)^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \partial \beta_1} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \sigma_0^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \sigma_1^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \sigma_{01}} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_1 \partial \beta_0} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \beta_1)^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_1 \partial \sigma_0^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_1 \sigma_1^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \sigma_{01}} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta_0 \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_0^2 \partial \beta_0} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \beta_1} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \sigma_0)^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_1^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_0^2 \sigma_{01}} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_0^2 \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_1^2 \partial \beta_0} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_1^2 \partial \beta_1} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \sigma_1)^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_1^2 \sigma_{01}} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_1^2 \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_{01} \partial \beta_0} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_{01} \partial \beta_1} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_{01} \partial \sigma_0^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma_{01} \partial \sigma_1^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \sigma_{01})^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma_{01} \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \beta_0} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\sigma^2 \partial \beta_1} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma_1^2} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \sigma)^2} \end{pmatrix},$$

e

$$\begin{aligned}
\text{diag} \left(\frac{\partial}{\partial \phi} [\mathbf{A}_n(\phi)^{-1}] \right) &= \text{diag} \left(\left(\frac{\partial \mathbf{A}_n(\phi)^{-1}}{\partial \phi_1} \quad \dots \quad \frac{\partial \mathbf{A}_n(\phi)^{-1}}{\partial \phi_s} \right) \right) \\
&= -\text{diag} \left(\left(\mathbf{A}_n(\phi)^{-1} \frac{\partial \mathbf{A}_n(\phi)}{\partial \phi_1} \mathbf{A}_n(\phi)^{-1} \quad \dots \quad \mathbf{A}_n(\phi)^{-1} \frac{\partial \mathbf{A}_n(\phi)}{\partial \phi_s} \mathbf{A}_n(\phi)^{-1} \right) \right) \\
&= -\text{diag} \left(\mathbf{A}_n(\phi)^{-1} \times \left(\frac{\partial \mathbf{A}_n(\phi)}{\partial \phi_1} \quad \dots \quad \frac{\partial \mathbf{A}_n(\phi)}{\partial \phi_s} \right) \times (\mathbf{I}_s \otimes \mathbf{A}_n(\phi)^{-1}) \right) \\
&= -\text{diag} \left(\mathbf{A}_n(\phi)^{-1} \times \frac{\partial \mathbf{A}_n(\phi)}{\partial \phi} \times (\mathbf{I}_s \otimes \mathbf{A}_n(\phi)^{-1}) \right).
\end{aligned}$$

em que \otimes representa o produto kronecker de matrizes. Por outro lado,

$$\begin{aligned}
\frac{\partial}{\partial \phi} \text{diag} (\mathbf{A}_n(\phi)^{-1} \mathbf{B}_n(\phi) \mathbf{A}_n(\phi)^{-1}) &= \text{diag} \left(\frac{\partial}{\partial \phi} (\mathbf{A}_n(\phi)^{-1} \mathbf{B}_n(\phi) \mathbf{A}_n(\phi)^{-1}) \right) \\
&= \text{diag} \left(\frac{\partial}{\partial \phi} \mathbf{A}_n(\phi)^{-1} \mathbf{B}_n(\phi) \mathbf{A}_n(\phi)^{-1} + \mathbf{A}_n(\phi)^{-1} \frac{\partial}{\partial \phi} \mathbf{B}_n(\phi) \mathbf{A}_n(\phi)^{-1} \right) \\
&+ \text{diag} \left(\mathbf{A}_n(\phi)^{-1} \mathbf{B}_n(\phi) \frac{\partial}{\partial \phi} \mathbf{A}_n(\phi)^{-1} \right).
\end{aligned}$$

Agora,

$$\frac{\partial}{\partial \phi} \mathbf{A}_n(\phi)^{-1} \mathbf{B}_n(\phi) \mathbf{A}_n(\phi)^{-1} = -\mathbf{A}_n(\phi)^{-1} \frac{\partial \mathbf{A}_n(\phi)}{\partial \phi} \mathbf{A}_n(\phi)^{-1} \mathbf{B}_n(\phi) \mathbf{A}_n(\phi)^{-1}$$

e temos inte-

resse em

$$\frac{\partial \mathbf{A}_n(\phi)}{\partial \phi} = \frac{\partial}{\partial \phi_k} \begin{pmatrix} \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta \partial \beta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta \partial \theta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \beta \partial \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \theta \partial \beta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \theta \partial \theta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \theta \partial \sigma^2} \\ \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \beta^\top} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \theta} & \frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{\partial \sigma^2 \partial \sigma^2} \end{pmatrix}, k = 1, \dots, s. \quad (\text{A.27})$$

A.5.2 Calculando cada expressão em (A.25)

Devemos encontrar, para todo k ,

$$\frac{\partial}{\partial \phi_k} \left(\frac{\partial l_i(\phi, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) = 0,$$

$k = 1, \dots, 3$.

Note que, $\frac{\partial \mathbf{V}_i}{\partial \sigma_0^2} = \mathbf{Z}_0 \mathbf{Z}_0^\top$, $\frac{\partial \mathbf{V}_i}{\partial \sigma_{01}} = \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top$ e $\frac{\partial \mathbf{V}_i}{\partial \sigma_1^2} = \mathbf{Z}_1 \mathbf{Z}_1^\top$.

Por outro lado, para $k = 4$,

$$\begin{aligned} \frac{\partial}{\partial \phi_k} \left(\frac{\partial l_i(\phi, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) &= \frac{\partial}{\partial \sigma_0^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_0^2} \mathbf{V}_i^{-1} \mathbf{X}_i, \\ &= \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_i \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \mathbf{Z}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i, \\ &= \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_i. \end{aligned}$$

$$\text{com } \mathbf{Z}_i = \begin{pmatrix} 1 & t_{i1} \\ \vdots & \vdots \\ 1 & t_{im_i} \end{pmatrix} = \begin{pmatrix} \mathbf{Z}_0 & \mathbf{Z}_1 \end{pmatrix} \text{ e } \mathbf{X}_i = \begin{pmatrix} 1 & t_{i1} & t_{i1}^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{im_i} & t_{im_i}^2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_{i0} & \mathbf{X}_{i1} & \mathbf{X}_{i2} \end{pmatrix}.$$

Para $k = 5$,

$$\frac{\partial}{\partial \phi_k} \left(\frac{\partial l_i(\phi, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) = \frac{\partial}{\partial \sigma_{01}} (-\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_i.$$

Para $k = 6$,

$$\frac{\partial}{\partial \phi_k} \left(\frac{\partial l_i(\phi, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) = \frac{\partial}{\partial \sigma_1^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_i.$$

Para $k = 7$,

$$\frac{\partial}{\partial \phi_k} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top} \right) = \frac{\partial}{\partial \sigma^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{X}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_i.$$

Temos interesse em,

$$\frac{\partial}{\partial \phi_k} \left(\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \boldsymbol{\theta}^\top} \right) = \frac{\partial}{\partial \phi_k} \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_1 \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \beta_2 \partial \sigma_1^2} \end{pmatrix} = \frac{\partial}{\partial \phi_k} (\mathbf{A}_{12}). \quad (\text{A.28})$$

A.6 Analisando a expressão \mathbf{A}_{12} em (A.28)

Para analisar a expressão $\frac{\partial}{\partial \beta_k} (\mathbf{A}_{12})$ para $k = 1, \dots, s$, lembremos que

$$\frac{\partial}{\partial \beta_k} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\theta}^\top} \right) = \frac{\partial}{\partial \beta_k} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_0^2} \quad \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_{01}} \quad \frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_1^2} \right),$$

cada expressão dentro da matriz \mathbf{A}_{12} pode ser encontrada calculando

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_0^2} = -\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_0^2} \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_0^2} \mathbf{V}_i^{-1} \mathbf{e}_i, \quad (\text{A.29})$$

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_{01}} = -\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_{01}} \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_{01}} \mathbf{V}_i^{-1} \mathbf{e}_i, \quad (\text{A.30})$$

$$\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_1^2} = -\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_1^2} \right) + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \frac{\partial \mathbf{V}_i}{\partial \sigma_1^2} \mathbf{V}_i^{-1} \mathbf{e}_i. \quad (\text{A.31})$$

Calculando a derivada segunda de (A.29), (A.30) e (A.31) com relação a cada β_k , $k = 1, 2, 3$.

$$\frac{\partial}{\partial \beta_0} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_0^2} \right) = -\frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0},$$

$$\frac{\partial}{\partial \beta_1} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_0^2} \right) = -\frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i1}$$

$$\frac{\partial}{\partial \beta_2} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_0^2} \right) = -\frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i2}.$$

$$\frac{\partial}{\partial \beta_0} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_{01}} \right) = -\frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i0},$$

$$\frac{\partial}{\partial \beta_1} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_{01}} \right) = -\frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i1},$$

$$\frac{\partial}{\partial \beta_2} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_{01}} \right) = -\frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i2}.$$

$$\frac{\partial}{\partial \beta_0} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_1^2} \right) = -\frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}.$$

$$\frac{\partial}{\partial \beta_1} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_1^2} \right) = -\frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i1}.$$

$$\frac{\partial}{\partial \beta_2} \left(\frac{\partial l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \sigma_1^2} \right) = -\frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i2}.$$

Calculando as derivadas terceiras com relação a ϕ_k , $k = 1, \dots, 7$ temos

$$\frac{\partial}{\partial \beta_0} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) = \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}^\top.$$

$$\frac{\partial}{\partial \beta_1} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) = \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i1} + \frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}.$$

$$\frac{\partial}{\partial \beta_2} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) = \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i2} + \frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}.$$

$$\frac{\partial}{\partial \sigma_0^2} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) = \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}.$$

$$\frac{\partial}{\partial \sigma_{01}} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) =$$

$$\begin{aligned} & \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i + \\ & \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0} + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i0}. \end{aligned}$$

$$\frac{\partial}{\partial \sigma_1^2} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) =$$

$$\begin{aligned} & \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \\ & \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0} + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0} \end{aligned}$$

$$\frac{\partial}{\partial \sigma^2} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta_0 \partial \sigma_0^2} \right) = \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \mathbf{e}_i +$$

$$\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0} + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \mathbf{X}_{i0}.$$

A.6.1 Analisando a expressão \mathbf{A}_{13} em (A.28)

Temos interesse em calcular a terceira derivada da expressão \mathbf{A}_{13} com relação a cada

elemento ϕ_k , isto é $\frac{\partial}{\partial \phi_k} \left(\frac{\partial^2 l_i(\phi, \mathbf{y}_i)}{\partial \beta \partial \sigma^2} \right) = \frac{\partial}{\partial \phi_k} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i)$, para $k = 1, \dots, 7$.

$$\frac{\partial}{\partial \beta_0} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_{i0}.$$

$$\frac{\partial}{\partial \beta_1} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_{i1}.$$

$$\frac{\partial}{\partial \beta_2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_{i2}.$$

$$\frac{\partial}{\partial \sigma_0^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \mathbf{e}_i + \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i.$$

$$\frac{\partial}{\partial \sigma_{01}} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-2} \mathbf{e}_i + \mathbf{X}_i^\top \mathbf{V}_i^{-2} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i,$$

$$\frac{\partial}{\partial \sigma_1^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-2} \mathbf{e}_i + \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i.$$

$$\frac{\partial}{\partial \sigma^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = 2\mathbf{X}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i.$$

A.7 Analisando a expressão \mathbf{A}_{22} em (A.28)

Vamos fazer o cálculo geral da segunda derivada do $l_i(\boldsymbol{\phi}; \mathbf{y}_i)$ com relação ao vetor $\boldsymbol{\theta}$, isto

é $\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}$,

$$\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \frac{\partial}{\partial \boldsymbol{\theta}} \left[\frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2} \quad \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01}^2} \quad \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2} \right] = \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_1^2} \end{pmatrix}.$$

Devemos calcular a terceira derivada da matriz \mathbf{A}_{12} com relação a ϕ_k para todo k, sendo assim, temos

$$\frac{\partial}{\partial \phi_k} \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_1^2} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}, \quad (\text{A.32})$$

Temos interesse em calcular a terceira derivada da expressão \mathbf{A}_{13} com relação a cada elemento ϕ_k , isto é $\frac{\partial}{\partial \phi_k} \left(\frac{\partial^2 l_i(\boldsymbol{\phi}, \mathbf{y}_i)}{\partial \boldsymbol{\beta} \partial \sigma^2} \right) = \frac{\partial}{\partial \phi_k} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i)$, para $k = 1, \dots, 7$.

$$\frac{\partial}{\partial \beta_0} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_{i0}.$$

$$\frac{\partial}{\partial \beta_1} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_{i1}.$$

$$\frac{\partial}{\partial \beta_2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{X}_{i2}.$$

$$\frac{\partial}{\partial \sigma_0^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \mathbf{e}_i + \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i.$$

$$\frac{\partial}{\partial \sigma_{01}} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-2} \mathbf{e}_i + \mathbf{X}_i^\top \mathbf{V}_i^{-2} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i,$$

$$\frac{\partial}{\partial \sigma_1^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-2} \mathbf{e}_i + \mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i.$$

$$\frac{\partial}{\partial \sigma^2} (-\mathbf{X}_i^\top \mathbf{V}_i^{-2} \mathbf{e}_i) = 2\mathbf{X}_i^\top \mathbf{V}_i^{-3} \mathbf{e}_i.$$

A.7.1 Analisando a expressão \mathbf{A}_{22} em (A.28)

Vamos fazer o cálculo geral da segunda derivada de $l_i(\boldsymbol{\phi}; \mathbf{y}_i)$ com relação ao vetor $\boldsymbol{\theta}$, isto

é $\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top}$,

$$\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} = \frac{\partial}{\partial \boldsymbol{\theta}} \left[\frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2} \quad \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01}^2} \quad \frac{\partial l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2} \right] = \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_1^2} \end{pmatrix}.$$

Calculando a terceira derivada da matriz \mathbf{A}_{12} com relação a ϕ_k para todo k, temos

$$\frac{\partial}{\partial \phi_k} \begin{pmatrix} \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_0^2 \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_{01} \partial \sigma_1^2} \\ \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_0^2} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_{01}} & \frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{\partial \sigma_1^2 \partial \sigma_1^2} \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{pmatrix}. \quad (\text{A.33})$$

Para resolver \mathbf{M}_{12} em (A.33) para todo k , devemos encontrar, $\frac{\partial}{\partial \phi_k} (\mathbf{M}_{12})$, isto é

$$\begin{aligned} \frac{\partial}{\partial \beta_0} (\mathbf{M}_{12}) &= \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \right) \\ &\quad - \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\ &\quad - \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right). \end{aligned}$$

Assim,

$$\begin{aligned} \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right) &= -\frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\ &\quad - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i0}. \end{aligned}$$

Por outro lado,

$$\begin{aligned} \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) &= \\ &\quad - \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i - \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}. \end{aligned}$$

Agora, para $\frac{\partial}{\partial \beta_1} (\mathbf{M}_{12})$,

$$\begin{aligned} \frac{\partial}{\partial \beta_1} (\mathbf{M}_{12}) &= \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \right) \\ &\quad - \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right). \end{aligned}$$

Derivando a segunda expressão na soma anterior temos:

$$\begin{aligned}
& \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\
&= -\frac{1}{2} \mathbf{X}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i - \\
& \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_i.
\end{aligned}$$

e

$$\begin{aligned}
& \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\
&= -\frac{1}{2} \mathbf{X}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
& -\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_i.
\end{aligned}$$

Agora bem, para o caso, $\frac{\partial}{\partial \beta_2} (\mathbf{M}_{12})$ temos:

$$\begin{aligned}
& \frac{\partial}{\partial \beta_2} (\mathbf{M}_{12}) = \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \right) \right) \\
& - \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right) - \\
& \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right),
\end{aligned}$$

temos,

$$\begin{aligned}
& \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\
&= -\frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i - \\
& \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i2}.
\end{aligned}$$

Por outro lado, para a segunda expressão na soma acima, temos

$$\begin{aligned} & \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\ &= -\frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\ & -\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i2}. \end{aligned}$$

Agora, para o caso, $\frac{\partial}{\partial \sigma_0^2} (\mathbf{M}_{12})$ temos

$$\begin{aligned} \frac{\partial}{\partial \sigma_0^2} (\mathbf{M}_{12}) &= -\text{tr} \left(\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \right) \\ & + \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\ & + \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i, \end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \sigma^2} (\mathbf{M}_{12}) &= -\frac{1}{2} \text{tr} (\mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \\
&\quad -\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \\
&\quad +\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad +\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad +\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-2} \mathbf{e}_i \\
&\quad +\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-2} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad +\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-2} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad +\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-2} \mathbf{e}_i.
\end{aligned}$$

O interesse está centrado em $\frac{\partial}{\partial \beta_0} (\mathbf{M}_{13})$, $\frac{\partial}{\partial \beta_1} (\mathbf{M}_{13})$, $\frac{\partial}{\partial \sigma_0^2} (\mathbf{M}_{13}) \dots, \frac{\partial}{\partial \sigma^2} (\mathbf{M}_{13})$. Sendo assim, encontramos cada um das derivadas indicadas ,

$$\begin{aligned}
\frac{\partial}{\partial \beta_0} (\mathbf{M}_{13}) &= \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \right) - \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\
&\quad - \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right),
\end{aligned}$$

derivando cada expressão tense,

$$\begin{aligned}
-\frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) &= \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}.
\end{aligned}$$

e

$$-\frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i$$

$$+ \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i0}.$$

Calculando a derivada segunda de M_{13} ,

$$\frac{\partial}{\partial \beta_1} (M_{13}) = \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \text{tr}(\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \right) - \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right)$$

$$- \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right),$$

$$-\frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i$$

$$+ \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i1},$$

e

$$-\frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i$$

$$+ \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i1}.$$

Agora para,

$$\frac{\partial}{\partial \beta_2} (M_{13}) = \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \text{tr}(\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \right) - \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right)$$

$$- \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right),$$

tense,

$$-\frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\ + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{X}_{i2},$$

e

$$-\frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\ + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{X}_{i2},$$

agora,

$$\frac{\partial}{\partial \sigma_0^2} (\mathbf{M}_{13}) = -\text{tr} \left(\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \right) \\ + \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\ + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\ + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\ + \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i,$$

para

$$\frac{\partial}{\partial \sigma_{01}} (\mathbf{M}_{13}) = \frac{\partial}{\partial \sigma_{01}} \left(\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \right) \right) - \frac{\partial}{\partial \sigma_{01}} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right).$$

Resolvendo o anterior, tense

$$\begin{aligned}
\frac{\partial}{\partial \sigma_{01}} (\mathbf{M}_{13}) &= -\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \\
&\quad - \frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \\
&\quad + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i.
\end{aligned}$$

Devemos encontrar agora,

$$\begin{aligned}
\frac{\partial}{\partial \sigma_1^2} (\mathbf{M}_{13}) &= \frac{\partial}{\partial \sigma_1^2} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \right) - \frac{\partial}{\partial \sigma_1^2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) \\
&= -\frac{\partial}{\partial \sigma_1^2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right),
\end{aligned}$$

assim,

$$\begin{aligned}
\frac{\partial}{\partial \sigma_1^2} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \right) &= -\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top) \\
&\quad - \frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top).
\end{aligned}$$

Agora temos,

$$\begin{aligned}
& -\frac{\partial}{\partial \sigma_1^2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \\
& \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i, \\
& = \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i.
\end{aligned}$$

$$\begin{aligned}
& -\frac{\partial}{\partial \sigma_1^2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i \right) = \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \\
& \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i, \\
& = \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{e}_i + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{Z}_0 \mathbf{Z}_0^\top \mathbf{V}_i^{-1} \mathbf{Z}_1 \mathbf{Z}_1^\top \mathbf{V}_i^{-1} \mathbf{e}_i.
\end{aligned}$$

Para o caso de $\frac{\partial}{\partial \beta_0} (\mathbf{M}_{22})$, se tem

$$\begin{aligned}
& \frac{\partial}{\partial \beta_0} \left(\frac{\partial^2 l_i(\phi; \mathbf{y}_i)}{(\partial \sigma_{01})^2} \right) = \frac{\partial}{\partial \beta_0} (\mathbf{M}_{22}) = \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \text{tr} \left(\mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \right) \right) \\
& - \frac{\partial}{\partial \beta_0} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right), \\
& = \frac{1}{2} \mathbf{X}_{i0}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\
& + \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i0}.
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \beta_1} \left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y})}{(\partial \sigma_{01})^2} \right) &= \frac{\partial}{\partial \beta_1} (\mathbf{M}_{22}) = \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \right) \\
&- \frac{\partial}{\partial \beta_1} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right), \\
&= \frac{1}{2} \mathbf{X}_{i1}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i + \\
&\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i1}.
\end{aligned}$$

Por outro lado,

$$\begin{aligned}
\frac{\partial}{\partial \beta_2} \left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \sigma_{01})^2} \right) &= \frac{\partial}{\partial \beta_2} (\mathbf{M}_{22}) = \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \right) \\
&- \frac{\partial}{\partial \beta_2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right), \\
&= \frac{1}{2} \mathbf{X}_{i2}^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \\
&+ \frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{X}_{i2}.
\end{aligned}$$

Analisando,

$$\begin{aligned}
\frac{\partial}{\partial \sigma_0^2} \left(\frac{\partial^2 l_i(\boldsymbol{\phi}; \mathbf{y}_i)}{(\partial \sigma_{01})^2} \right) &= \frac{\partial}{\partial \sigma_0^2} (\mathbf{M}_{22}) = \frac{\partial}{\partial \sigma_0^2} \left(\frac{1}{2} \text{tr} (\mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \}) \right) \\
&- \frac{\partial}{\partial \sigma_0^2} \left(\frac{1}{2} \mathbf{e}_i^\top \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \{ \mathbf{Z}_1 \mathbf{Z}_0^\top + \mathbf{Z}_0 \mathbf{Z}_1^\top \} \mathbf{V}_i^{-1} \mathbf{e}_i \right),
\end{aligned}$$

A.8 Apêndice C

Condições de regularidade

Consideramos a variável aleatória \mathbf{y}_i , $i = 1, \dots, n$, com função de distribuição conjunta G sob Θ , o espaço euclidiano mensurável com densidade Radon-Nikodim $g = \frac{dG}{d\vartheta}$. Se G é desconhecida a priori, então é selecionada uma família $(F(\mathbf{y}))$ de funções de distribuição que podem ou não conter a verdadeira estrutura de G . Essa família satisfaz as seguintes suposições.

1. A família de funções de distribuição $F(\mathbf{y})$ tem função de densidade Radon-Nikodim

$$f(\mathbf{y}, \phi) = \frac{dF(\mathbf{y}, \phi)}{d\vartheta}.$$

2. As seguintes propriedades são certas:

- $\mathbb{E}[\log(g(\mathbf{y}))]$ existe e $|\log f(\mathbf{y}, \phi)| \leq m(\mathbf{y})$ para toda $\mathbf{y} \in \Omega$

3. $\text{KL}(g : f; \phi)$ tem um único mínimo em $\phi^* \in \Theta$

4. As funções $\frac{\partial \log f(\mathbf{y}, \phi)}{\partial \phi_k}$, são mensuráveis com respeito a \mathbf{y} para cada $\phi \in \Theta$, contínuas e diferenciáveis de ϕ para cada $\mathbf{y} \in \Omega$

5. $\left| \frac{\partial^2 \log f(\mathbf{y}, \phi)}{\partial \phi_k \partial \phi_l} \right|$ e $\left| \frac{\partial \log f(\mathbf{y}, \phi)}{\partial \phi_k} \frac{\partial \log f(\mathbf{y}, \phi)}{\partial \phi_l} \right|$ são dominadas por funções integráveis com respeito a M para todo $\mathbf{y} \in \Omega$ e para todo $\phi \in \Theta$.

6. As seguintes propriedades são certas:

- (a) ϕ^* é um ponto interior de Θ
- (b) $\mathbf{B}(\phi^*)$ é não singular, e
- (c) ϕ^* é um ponto regular de $\mathbf{A}(\phi^*)$.

Usando as suposições de 1. a 6. [White \(1982\)](#) mostra que

Theorem 1 *Normalidade assintótica*

Dadas as suposições de 1. a 5.

$$\sqrt{n} \left(\hat{\phi}_n - \phi^* \right) \overset{A}{\sim} \mathcal{N}(\mathbf{0}, \mathbb{V}(\phi^*))$$

além, do que $\mathbb{V}_n(\hat{\phi}_n) \xrightarrow{c.s.} \mathbb{V}(\phi^)$.*

7. As funções

$$\left| \frac{\partial}{\partial \phi_l} \right| \left| \frac{\partial f(\mathbf{y}, \phi)}{\partial \phi_k} f(\mathbf{y}, \phi) \right|$$

são dominadas por funções integráveis com respeito a ϑ , para todo $\phi \in \Theta$ e o suporte mínimo de $f(\mathbf{y}, \phi)$ não depende de ϕ .

8. As funções

$$\frac{\partial d_k(\mathbf{y}, \phi)}{\partial \phi_l},$$

existem e são contínuas em ϕ para cada \mathbf{y} .

9. As funções

$$\left| \frac{\partial d_k(\mathbf{y}, \phi)}{\partial \phi_l} \right| \text{ e } \left| \partial d_k(\mathbf{y}, \phi) \frac{\partial f(\mathbf{y}, \phi)}{\partial \phi_l} \right|$$

são dominadas por funções integráveis com respeito a M para todo $\mathbf{y} \in \Omega$.

10. As funções

$$\frac{\partial d_k(\mathbf{y}, \phi)}{\partial \phi_l}$$

existem e são contínuas em ϕ para cada \mathbf{y} .

A.9 Apêndice D

A.9.1 Teste Alternativo da Matriz de Informação

- A ideia para a construção da matriz de covariâncias da estatística $\sqrt{n}\mathbf{d}_n(\mathbf{y}; \hat{\boldsymbol{\phi}}_n)$ é, primeiro usar uma expansão em Taylor de primeira ordem dessa estatística ao redor de $\boldsymbol{\phi}_0$, segundo usar o teorema do limite central, seguindo as indicações detalhadas em [Gourieroux & Monfort \(1995, p.183\)](#).
- Sabemos que sob as condições de regularidade no apêndice C, e sob a hipótese nula H_0 que o modelo é corretamente especificado a estatística $\sqrt{n}\mathbf{d}_n(\mathbf{y}; \hat{\boldsymbol{\phi}}_n)$, é tal que

$$\sqrt{n}\mathbf{d}_n(\mathbf{y}, \hat{\boldsymbol{\phi}}_n) \stackrel{A}{\sim} \mathcal{N}_s[\mathbf{0}, \mathbb{M}(\boldsymbol{\phi}_0)], \quad (\text{A.34})$$

[[White \(1982\)](#)]. Usando o Teorema de Cochran [[Sen & Singer \(1993, p.137\)](#)] considerando que $\hat{\mathbb{M}}_n(\hat{\boldsymbol{\phi}}_n)$ é uma estimação da matriz $\mathbb{M}(\boldsymbol{\phi}_0)$ e conhecendo que

$$\hat{\mathbb{M}}_n(\hat{\boldsymbol{\phi}}_n) \xrightarrow{c.s.} \mathbb{M}(\boldsymbol{\phi}_0)$$

Teorema 4.1 [White \(1982\)](#), então

$$EAMI = n\mathbf{d}_n(\mathbf{y}, \hat{\boldsymbol{\phi}}_n)^\top [\hat{\mathbb{M}}_n(\hat{\boldsymbol{\phi}}_n)]^{-1} \mathbf{d}_n(\mathbf{y}, \hat{\boldsymbol{\phi}}_n) \quad (\text{A.35})$$

é distribuído assintoticamente como χ_s^2 .

A.9.2 Teste do Estimador “Sanduíche” Modificado

- Considerando que $\mathbb{E}(\mathbf{d}_{1i}^*(\hat{\boldsymbol{\phi}}_n)) = \boldsymbol{\mu}_1^*(\boldsymbol{\phi})$, e $\mathbb{E}(\mathbf{d}_{2i}^*(\hat{\boldsymbol{\phi}}_n)) = \boldsymbol{\mu}_2^*(\boldsymbol{\phi})$.

$$\begin{aligned}
 Cov \left[\mathbf{d}_n^*(\hat{\phi}_n) \right] &= \sum_{i=1}^n Cov \left[\mathbf{d}_{1i}^*(\hat{\phi}_n) + \mathbf{d}_{2i}^*(\hat{\phi}_n) \right] \\
 &= \sum_{i=1}^n \left[\mathbb{V}(\mathbf{d}_{1i}^*(\hat{\phi}_n)) + \mathbb{V}(\mathbf{d}_{2i}^*(\hat{\phi}_n)) + Cov \left(\mathbf{d}_{1i}^*(\hat{\phi}_n), \mathbf{d}_{2i}^*(\hat{\phi}_n) \right) + Cov \left(\mathbf{d}_{2i}^*(\hat{\phi}_n), \mathbf{d}_{1i}^*(\hat{\phi}_n) \right) \right] \\
 &= \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right) \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right)^T \right) + \\
 &\quad \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right) \left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right)^T \right) + \\
 &\quad \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right) \left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right)^T \right) + \\
 &\quad \sum_{i=1}^n \mathbb{E} \left(\left(\mathbf{d}_{2i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_2^*(\phi) \right) \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) - \boldsymbol{\mu}_1^*(\phi) \right)^T \right).
 \end{aligned} \tag{A.36}$$

- Agora, sob as condições de regularidade e sob H_0 , tendo em conta o teorema central do limite, temos

$$\sqrt{n} \left[\mathbf{d}_n^*(\hat{\phi}_n) - \mathbb{E} \left(\mathbf{d}_n^*(\hat{\phi}_n) \right) \right] \stackrel{A}{\rightsquigarrow} \mathcal{N}_s(\mathbf{0}, \mathbf{C}_v(\phi_0))$$

e sob H_0 ,

$$\mathbb{E} \left(\mathbf{d}_n^*(\hat{\phi}_n) \right) = \mathbb{E} \left(\frac{1}{n} \sum_{i=1}^n \mathbf{d}_i^*(\hat{\phi}_n) \right) = \frac{1}{n} \mathbb{E} \left(\sum_{i=1}^n \left(\mathbf{d}_{1i}^*(\hat{\phi}_n) + \mathbf{d}_{2i}^*(\hat{\phi}_n) \right) \right) = \mathbf{0}, \tag{A.37}$$

então

$$\sqrt{n} \mathbf{d}_n^*(\hat{\phi}_n) \stackrel{A}{\rightsquigarrow} \mathcal{N}_s[\mathbf{0}, \mathbf{C}_v(\phi_0)]. \tag{A.38}$$

Aplicando o teorema de Cochran, [Sen & Singer (1993, p.137)], e tendo em conta que

$\widehat{\mathbf{C}}_v(\hat{\phi}_n)$ é um estimador consistente de $\mathbf{C}_v(\phi_0)$ temos que

$$n\mathbf{d}_n^*(\hat{\phi}_n)^\top [\widehat{\mathbf{C}}_v(\hat{\phi}_n)]^{-1} \mathbf{d}_n^*(\hat{\phi}_n) \stackrel{A}{\sim} \chi_s^2. \quad (\text{A.39})$$

A.9.3 Teste do Estimador “Sanduíche” Alternativo

- Considerando a distribuição assintótica do estimador $\hat{\phi}_n$ e tendo em consideração as condições de regularidades no apêndice C, então

$$\sqrt{n} (\hat{\phi}_n - \phi_0) \stackrel{A}{\sim} \mathcal{N}_s [\mathbf{0}, \mathbb{V}(\phi_0)], \quad (\text{A.40})$$

em que $\mathbb{V}(\phi_0)$ é como em (2.35), [White (1982)]. Além disso, $\nabla \mathbf{d}(\phi) |_{\phi=\phi_0}$ existe, então usando (A.40) e o método Delta, [Sen & Singer (1993, p.136)], obtemos

$$\sqrt{n} (\mathbf{d}_n^*(\hat{\phi}_n) - \mathbf{d}(\phi_0)) \xrightarrow{D} \mathcal{N}_s(\mathbf{0}, \nabla \mathbf{d}(\phi_0) \mathbb{V}(\phi_0) \nabla \mathbf{d}(\phi_0)^\top). \quad n \rightarrow \infty \quad (\text{A.41})$$

De acordo com (A.37), $\mathbb{E} (\mathbf{d}_n^*(\hat{\phi}_n)) = \mathbf{0}$, sob a hipótese H_0 , com (A.41), aplicando o teorema de Cochran, [Sen & Singer (1993, p.137)], a estatística do estimador “sanduíche” alternativo,

$$n\mathbf{b}_n^*(\hat{\phi}_n)^\top \left[\nabla \mathbf{d}_n(\hat{\phi}_n) \widehat{\mathbb{V}}(\hat{\phi}_n) \nabla \mathbf{d}_n(\hat{\phi}_n)^\top \right]^{-1} \mathbf{d}_n^*(\hat{\phi}_n) \stackrel{A}{\sim} \chi_s^2 \quad (\text{A.42})$$

Referências Bibliográficas

- Abad, A., Litiere, S., & Molenberghs, G. (2010). Testing for misspecification in generalized linear mixed models. *Biostatistics*, *11* (4), 771–786. 5, 6, 14, 16, 18
- Afiune, J. (2000). Avaliação ecocardiográfica evolutiva de recém-nascidos pré - termo, do nascimento até o termo. *Tese doutoral, Instituto da Criança da Faculdade de Medicina, Universidade de São Paulo, Brasil*. xiv, 2, 26, 60
- Demidenko, E. (2013). *Mixed Models: Theory and Applications with R*. John Wiley & Sons, New Jersey. 2, 11, 113
- Diggle, P., Heagerty, P., & Liang, K. (2002). *Analysis of Longitudinal Data*. Oxford University Press, Oxford. 1, 2
- Drikvandi, R., Verbeke, G., & Molenberghs, G. (2017). Diagnosing misspecification of the random-effects distribution in mixed models. *Biometrics*, *73* (1), 63–71. 5
- Freedman, D. (2006). On the so-called “huber sandwich estimator” and “robust standard errors”. *The American Statistician*, *60*(4), 299–302. 15, 18
- Furno, M. (1996). The information matrix test in the linear regression with arma errors. *Journal of the Italian Statistical Society*, *5*(3), 369–385. 4
- Gourieroux, C., & Monfort, A. (1995). *Statistics and Econometrics Models*. Cambridge University Press, Cambridge. 17, 145
- Gumedze, F., & Dunne, T. (2011). Parameter estimation and inference in the linear mixed model. *Linear Algebra and its Applications*, *435*(8), 1920–1944. 10

- Hall, A. (1987). The information matrix test for the linear model. *Review of Economic Studies*, 54(2), 257-263. 3
- Harville, D. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American statistical association*, 72(358). 10
- Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, 46(6), 1251-1271. 3
- Hedeker, D., & Gibbson, R. (2006). *Longitudinal Data Analysis*. John Wiley & Sons, New Jersey. 1
- Henderson, C. (1975). Best linear unbiased estimation and prediction under a selection model. *Biometrics*, 31(2), 423-447. 12, 13
- Hilden, M. (1995). Multilevel diagnostics for mixed and hierarchical linear models. *PhD Thesis, UCLA, Los Angeles..* 23, 24
- Huang, X. (2009). Diagnosis of random-effect model misspecification in generalized linear mixed models for binary response. *Biometrics*, 65(2), 361-366. 5
- Huang, X. (2011). Detecting random - effects model misspecification via coarsened data. *Computational Statistics & Data Analysis*, 55(1), 703-714. 4
- Inoue, H., Kuo, C., & Rossi, B. (2014). Identifying the sources of model misspecification. *Journal of Monetary Economics*, 110, 1-18. 4
- Jiang, J. (2001). Goodness-of-fit tests for mixed model diagnostics. *The Annals of Statistics*, 29(4), 1137-1164. 4
- Jiang, J. (2007). *Linear and Generalized Linear Mixed Models and Their Applications*. Springer Science & Business Media, New York. 8, 24
- Laird, N., & Ware, J. (1982). Random-effects models for longitudinal data. *Biometrics*, 38(4), 963-974. 7

- Lesaffre, E., & Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, *54*(2), 570–582. 4, 23
- Little, R., & Rubin, D. (1987). *Statistical Analysis with Missing Data*. John Wiley & Sons, New York. 1
- McCulloch, C., & Searle, S. (2001). *Generalized, Linear and Mixed Models*. John Wiley & Sons, New Jersey. 12
- Molenberghs, G., & Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer, New York. 1, 2
- Nobre, J. S., & Singer, J. M. (2007). Residual analysis for linear mixed models. *Biometrical Journal*, *49*(6), 863–875. 24
- Patterson, H., & Thompson, R. (1977). Recovery of inter block information when block sizes are unequal. *Biometrika*, *58*(3), 545–1971. 8, 10
- Petersen, K. B., & Pedersen, M. S. (2012). The matrix cookbook. URL <http://www2.imm.dtu.dk/pubdb/p.php, 3274>, 14. 114
- Pinheiro, J., & Bates, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer, New York. 63
- Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science*, *6*(1), 15–32. 8
- Rocha, F. M. (2004). *Seleção de Estruturas de Covariâncias Para Dados com Medidas Repetidas*. Mestre em Estatística , Universidade São Paulo, Brasil. 4
- Rocha, F. M., & Singer, J. M. (2018). Selection of terms in random coefficient regression models. *Journal of Applied Statistics*, *45*(2), 225–242. 5, 23, 26, 80
- Searle, S. R. (1982). *Matrix Algebra Useful for Statistics*. John Wiley & Sons, New York, 2nd Edition. 8, 120

- Sen, P. K., & Singer, J. M. (1993). *Large Sample Methods in Statistics. An Introduction with Applications..* Chapman & Hall, Flórida. 14, 17, 20, 22, 145, 146, 147
- Singer, J. M., Rocha, F. M., & Nobre, J. S. (2017). Graphical tools for detecting departures from linear mixed model assumptions and some remedial measures. *International Statistical Review*, 85 (2), 290–324. 5, 23, 24, 26, 63, 67
- Stroup, W. (2012). *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications..* CRC press, Taylor & Francis group, FlÃşrida. 113
- Tountenburg, H. (1982). *Prior Information in linear models.* John Wiley & Sons, New York. 8
- Verbeke, G., & Molenberghs, G. (2013). The gradient function as an exploratory assessment of the random-effects distribution in mixed models. *Biometrics*, 14(3), 447–490. 4
- Wang, W. (2013). Identifiability in linear mixed models. *Electronic Journal of Statistics*, 7(1), 244–263. 2
- White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica*, 50(1), 1–24. 3, 4, 5, 14, 17, 21, 22, 143, 145, 147
- Wilk, M., & Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. *Biometrikal*, 55(1), 1–17. 24
- Yu, D., Zhang, X., & Yau, K. (2018). Asymptotic properties and information criteria for misspecified generalized linear mixed models. *Journal of the Royal Statistical Society*, 80(4), 817 – 836. 13
- Zeigler, A. (2011). *Generalized Estimating Equations.* Springer, New York. 17
- Zhu, S. (2016). Computing log-likelihood and its derivatives for restricted maximum likelihood. *arXiv preprint arXiv:1608.07207, Ithaca.* 12