

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PÓS-GRADUAÇÃO EM ESTATÍSTICA

MODELAGEM PARA DADOS LONGITUDINAIS DE CONTAGEM

DANIELE DE BRITO TRINDADE

Orientador: Prof. Dr. Raydonal Ospina Martínez
Co-orientador: Prof^a. Dr^a. Leila Denise Alves Ferreira Amorim

Área de concentração: Estatística Aplicada

Recife, fevereiro de 2014

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
PÓS-GRADUAÇÃO EM ESTATÍSTICA

MODELAGEM PARA DADOS LONGITUDINAIS DE CONTAGEM

DANIELE DE BRITO TRINDADE

Orientador: Prof. Dr. Raydonal Ospina Martínez
Co-orientador: Prof^a. Dr^a. Leila Denise Alves Ferreira Amorim

Área de concentração: Estatística Aplicada

Trabalho apresentado ao Programa de Pós-Graduação em Estatística do
Departamento de Estatística da Universidade Federal de Pernambuco como
requisito parcial para obtenção do grau de Mestre em Estatística.

Recife, fevereiro de 2014

Catálogo na fonte
Bibliotecária Jane Souto Maior, CRB4-571

Trindade, Daniele de Brito

**Modelagem para dados longitudinais de contagem /
Daniele de Brito Trindade. - Recife: O Autor, 2014.
190 f.: il., fig., tab.**

Orientador: Raydonal Ospina Martínez.

**Dissertação (mestrado) - Universidade Federal de Pernambuco.
CCEN, Estatística, 2014.**

Inclui referências e apêndices.

**1. Estatística aplicada. I. Martínez, Raydonal Ospina (orientador).
II. Título.**

310

CDD (23. ed.)

MEI2014 – 034

Universidade Federal de Pernambuco
Pós-Graduação em Estatística

19 de fevereiro de 2014

Nós recomendamos que a Dissertação de Mestrado de autoria de

Daniele de Brito Trindade

Intitulada:

“Modelagem para Dados Longitudinais de Contagem”

Seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.

Coordenador da Pós-Graduação em Estatística

Banca Examinadora:

Leila Denise Alves Amorim	Co-orientadora/UFBA
---------------------------	---------------------

Claudia Regina Oliveira de Paiva Lima	UFPE
---------------------------------------	------

Juvêncio Santos Nobre	UFC
-----------------------	-----

Este documento será anexado à versão final da tese.

*Dedico esta dissertação à minha mãe Maria Terezinha e ao meu pai Luiz Carlos
(in memoriam), pelo amor incondicional e apoio em todos os momentos da minha
vida.*

Agradecimentos

Primeiramente, agradeço a Deus pela constante presença na minha vida, dando-me tranquilidade e equilíbrio para vencer todos os obstáculos que encontrei durante esta longa e árdua caminhada.

Agradeço à minha mãe Maria Terezinha Vieira de Brito pelo apoio, pela força, pela superação da distância e principalmente pelas palavras de conforto.

Ao meu orientador Raydonal Ospina Martínez e a minha co-orientadora Leila Denise Alves Ferreira Amorim, pela orientação, apoio, paciência, compreensão e confiança depositada em mim durante a elaboração deste trabalho.

Aos meus colegas e amigos de pós-graduação que me apoiaram e me ajudaram durante as dificuldades. Em especial, Jéssica Rivas que esteve junto comigo desde a graduação superando todas as dificuldades, Cleber Xavier e Sérgio Oliveira por me ajudarem e apoiarem no início desta caminhada, Cristiany Barros, Maria Priscila Souza, Giannini Italino e Pedro Rafael por estarem presentes em momentos de superação, incentivando e apoiando nas horas difíceis.

Agradeço às minhas amigas Ana Clara Paixão, Danila Silva e Natália Pires, por me ajudarem nas horas que mais necessitei e me acolherem em Recife. Serei eternamente grata pelo apoio e amor que vocês têm por mim.

Aos meus amigos Evandro Santos, Rosemeire Fiaccone, Nívea Bispo, Mariese Alves, Liane e Michele Sacramento, que mesmo distante sempre se preocuparam comigo, trazendo palavras de conforto e de incentivo.

Aos meus professores da graduação e pós-graduação, por me ajudarem a superar as dificuldades e por dividir os seus conhecimentos comigo.

À Valéria Bittencourt por todo carinho, dedicação e gentileza. Muito obrigada pelas palavras de conforto e todo amor que teve por mim durante estes dois anos de muita luta e dedicação.

Ao Instituto de Saúde Coletiva (ISC - UFBA) pelo espaço físico e materiais disponíveis para a execução do meu trabalho.

Agradeço ao Ministério da Saúde pelo fornecimento dos dados do Sistema de Controle de Exames Laboratoriais (SISCEL).

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro concedido por meio da bolsa de mestrado.

*“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar.
Mas o mar seria menor se lhe faltasse uma gota”.*
Madre Teresa de Calcuta

Resumo

A modelagem para dados de contagem é bastante utilizada em diversas áreas do conhecimento, como nas ciências biológicas, educação e saúde pública. O modelo comumente utilizado para analisar dados de contagem é o modelo Poisson. Contudo, quando os dados apresentam superdispersão o modelo Poisson não é mais indicado. Existem extensões do modelo Poisson que podem ser usados nesta situação, como o modelo Poisson inflacionado de zeros (*ZIP*, em inglês). Porém, neste trabalho, é considerado o modelo Binomial Negativo, que é adequado para esta situação, além de ser um modelo simples e bastante conhecido. Uma suposição do modelo de regressão tradicional é a independência entre as observações. Contudo, quando as unidades amostrais são medidas repetidamente ao longo do tempo, os estudos longitudinais permite a verificação das taxas de mudança que ocorrem ao longo do tempo e os fatores que podem motivar tal variação. Estes estudos são de particular interesse quando o objetivo é avaliar variações globais ou individuais da resposta ao longo do tempo. Este tipo de estudo considera a correlação entre as respostas dentro das unidades amostrais e a ordenação cronológica das respostas. Duas abordagens de regressão comumente utilizadas para analisar dados longitudinais são os modelos condicionais e os marginais. O modelo condicional assume a existência de efeitos aleatórios que descrevem o comportamento de um indivíduo específico, sendo este modelo também chamado de multinível. No modelo marginal a variável resposta é modelada independentemente da correlação existente entre as medidas de cada unidade amostral (denotada por correlação intra-indivíduo), modelando a expectativa marginal como uma função das variáveis explicativas. Neste trabalho as duas abordagens foram aplicadas à análise de dados de contagem longitudinais. Estudos de simulação foram realizados para avaliar a performance dos estimadores provenientes destas metodologias. Aplicações com bases de dados reais são apresentadas.

Palavras-chave: Dados de Contagem Longitudinais. Distribuição Poisson. Distribuição Binomial Negativa. Equações de Estimação Generalizadas. Estudos de Simulação. Modelos Multiníveis. Modelos Marginais via Cópulas Gaussianas.

Abstract

Models for count data are widely used in various areas of knowledge, as in biological sciences, education and public health. The commonly used model to analyze count data is the Poisson model. However, when the data exhibit overdispersion the Poisson model is no indicated. There are extensions of the Poisson model that can be used in this situation, as the zero-inflated Poisson model (ZIP). However, this work is considered the Negative Binomial model, which is suitable for this situation, besides being a well known and simple model. An assumption of traditional regression models is the independence between observations. However, when the sample units are measured repeatedly over time, longitudinal studies allows the verification of rates of change occurring over time and the factors that may motivate such a change. This type of study considers the correlation between responses within the sample units and the chronological ordering of answers. Two regression approaches commonly used to analyze longitudinal data are conditional and marginal models. The conditional model assumes the existence of random effects that describe the behavior of a specific individual. This model is also called multilevel model. In the marginal model the response variable is modeled independently of the correlation between the measurements of each sample unit (denoted by intra - individual correlation), modeling the marginal expectation as a function of the explanatory variables. In this work the two approaches were applied to longitudinal count data. Simulation studies were conducted to evaluate the estimators' performance from these methodologies. Application with real data are provided.

Keywords: Generalized Estimating Equations. Longitudinal Count Data. Multilevel Models. Marginal Models by Gaussian Copulas. Negative Binomial Distribution. Poisson Distribution. Simulation Studies.

Lista de Figuras

1.1	Gráficos do (a) perfil individual e (b) perfil médio do número de ataques epilépticos.	24
5.1	Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$	95
5.2	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$	96
5.3	Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0, 1$	102
5.4	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0, 1$	103
5.5	Valor absoluto do viés relativo para diferentes especificações no modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação AR1.	107
5.6	Diferença relativa dos estimadores para diferentes especificações no modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação AR1.	108
5.7	Valor absoluto do viés relativo para diferentes especificações nos modelos Poisson e BN considerando o modelo BN multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$	113

5.8	Diferença relativa das estimativas para diferentes especificações nos modelos Poisson e BN considerando o modelo BN multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$	113
5.9	Valor absoluto do viés relativo para diferentes especificações nos modelos Poisson e BN marginais com cópulas Gaussianas, com dados gerados pelo modelo BN marginal via cópulas Gaussianas com correlação AR1.	116
5.10	Diferença relativa das estimativas para diferentes especificações nos modelos Poisson e BN marginais via cópulas Gaussianas, com dados gerados pelo modelo BN marginal via cópulas Gaussianas com correlação AR1.	117
6.1	Gráficos do (a) perfil individual e (b) perfil médio do número de ataques epilépticos.	121
6.2	Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson multiníveis (a.1 e a.2) com intercepto aleatório e (b.1 e b.2) com dois efeitos aleatórios.	126
6.3	Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos BN multiníveis (a.1 e a.2) com intercepto aleatório e (b.1 e b.2) com dois efeitos aleatórios.	127
6.4	Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson marginais, segundo a abordagem EEG, com estrutura de correlação (a.1 e a.2) AR1 e (b.1 e b.2) permutável. . .	128
6.5	Gráficos com a distância de Cook e os resíduos de Pearson padronizados do modelo Poisson marginal considerando cópulas Gaussianas com estrutura de correlação (a.1 e a.2) autorregressiva de ordem 1 (AR1), (b.1 e b.2) permutável e (c.1 e c.2) não estruturada.	129
6.6	Gráficos com a distância de Cook e os resíduos de Pearson padronizados do modelo Binomial Negativo marginal considerando cópulas Gaussianas com estrutura de correlação (a.1 e a.2) autorregressiva de ordem 1 (AR1), (b.1 e b.2) permutável e (c.1 e c.2) não estruturada. . .	130
6.7	Envelopes simulados do modelo Poisson marginal (cópulas Gaussianas) com estruturas de correlação (a) AR1, (b) permutável, (c) não estruturada e (d) independente.	131

6.8	Envelopes simulados do modelo BN marginal (cópu- las Gaussianas) com estruturas de correlação (a) AR1, (b) permutável, (c) não estru- turada e (d) independente.	131
6.9	Gráficos do (a) perfil individual e (b) perfil médio do número de ataques epiléticos sem três indivíduos com valores discrepantes. . .	132
6.10	Gráficos com a distância de Cook e os resíduos de Pearson padro- nizados dos modelos Poisson multinível (a.1 e a.2) com dois efeitos aleatórios.	134
6.11	Gráficos com a distância de Cook e os resíduos de Pearson padroni- zados dos modelos BN multinível (a.1 e a.2) com intercepto aleatório.	135
6.12	Gráficos com a distância de Cook e os resíduos de Pearson padroniza- dos dos modelos Poisson marginal, segundo a abordagem EEG, com estrutura de correlação (a.1 e a.2) AR1.	135
6.13	Distância de Cook e Resíduos de Pearson Padronizados do modelo BN marginal com cópu- las Gaussianas e estrutura de correlação per- mutável sem os três indivíduos discrepantes.	136
6.14	Envelope simulado do modelo BN marginal via cópu- las Gaussianas e estrutura de correlação permutável sem os três indivíduos discrepantes.	137
6.15	Gráficos do (a) perfil individual para dez pacientes e do (b) perfil médio, global e por sexo, do número de consultas ao médico.	138
6.16	Gráfico de perfis individuais para o número de células CD4+ em 10 indivíduos selecionados aleatoriamente.	142
B.1	Valor absoluto do viés relativo para diferentes especificações no mo- delo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,5$	168
B.2	Valor absoluto do viés relativo para diferentes especificações no mo- delo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,7$	169
B.3	Valor absoluto do viés relativo para diferentes especificações no mo- delo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 1,0$	170
B.4	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com inter- cepto aleatório para a geração dos dados e $\tau_0 = 0,5$	171

B.5	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,7$	172
B.6	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 1,0$	173
B.7	Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,5$	174
B.8	Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,7$	175
B.9	Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 1,0$	176
B.10	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,5$	177
B.11	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,7$	178
B.12	Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 1,0$	179
B.13	Valor absoluto do viés relativo considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada.	186
B.14	Valor absoluto do viés relativo considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação permutável.	188
B.15	Diferença relativa dos estimadores considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada.	189

B.16 Diferença relativa dos estimadores considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação permutável.	190
---	-----

Lista de Tabelas

3.1	Ilustração da estrutura de um banco de dados longitudinais	40
3.2	Uso dos resíduos na análise de diagnóstico para modelos lineares mistos. Tipo de resíduo: 1-Marginal, 2- Condicional e 3-EBLUP.	54
5.1	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	93
5.2	Valores do AIC para diferentes especificações no modelo Poisson com 50 indivíduos e gerando os dados utilizando o modelo Poisson multinível com intercepto aleatório.	98
5.3	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	99
5.4	Valores do AIC para diferentes especificações no modelo Poisson considerando 50 indivíduos e gerando os dados utilizando o modelo Poisson multinível com dois efeitos aleatórios.	104
5.5	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) com 50 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação AR1. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	106

5.6	Valores do critério de quasi-verossimilhança (QIC) para a escolha do melhor modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação autorregressiva de ordem 1.	110
5.7	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações nos modelos Poisson e BN com 50 indivíduos e dados gerados pelo modelo BN multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	112
5.8	Valores do AIC para diferentes especificações dos modelos Poisson e BN considerando 50 indivíduos e gerando os dados utilizando o modelo BN multinível com intercepto aleatório.	114
5.9	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações nos modelos Poisson e BN com 50 indivíduos e dados gerados pelo modelo BN marginal via cópulas Gaussianas com estrutura de correlação AR1. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	115
5.10	Valores do AIC para diferentes especificações dos modelos Poisson e BN marginais via cópulas Gaussianas considerando 50 indivíduos e gerando os dados pelo modelo BN multinível com cópulas Gaussianas com correlação AR1.	118
6.1	Número médio de ataques epilépticos e desvio-padrão (DP) por grupo de tratamento.	121
6.2	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão Poisson e BN independentes, multiníveis com intercepto aleatório (Multinível 1) e multiníveis com dois efeitos aleatórios (Multinível 2) para o número de ataques epilépticos.	122
6.3	Estimativas dos coeficientes de regressão e erros-padrão (entre parênteses) usando os modelos de regressão de Poisson e BN na abordagem EEG com diferentes estruturas de correlação para o número de ataques epilépticos.	123

6.4	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Poisson e BN na abordagem marginal via cópulas Gaussianas com diferentes estruturas de correlação para o número de ataques epiléticos.	124
6.5	Número médio de ataques epiléticos e desvio-padrão (DP) por grupo de tratamento.	133
6.6	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão Poisson multinível com dois efeitos aleatórios (Multinível 2) e BN multinível com intercepto aleatório (Multinível 1) para o número de ataques epiléticos sem três indivíduos discrepantes.	133
6.7	Estimativas dos coeficientes de regressão e erros-padrão (entre parênteses) usando os modelos de regressão de Poisson e BN na abordagem EEG com estruturas de correlação AR1 para o número de ataques epiléticos sem indivíduos discrepantes.	134
6.8	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão Poisson e BN marginais via cópulas Gaussianas com estrutura de correlação permutável para o número de ataques epiléticos sem indivíduos discrepantes.	134
6.9	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Poisson independente, multinível com intercepto aleatório e multinível com dois efeitos aleatórios para o número de consultas ao médico.	139
6.10	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando a abordagem EEG para o modelo Poisson marginal com diferentes estruturas de correlação para o número de consultas ao médico.	139
6.11	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Binomial Negativo tradicional e Binomial Negativo multinível com intercepto aleatório para o número de consultas ao médico.	140
6.12	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando a abordagem EEG para o modelo Binomial Negativo com diferentes estruturas de correlação para o número de consultas ao médico.	140

6.13	Sumarização das variáveis pertencentes ao estudo sobre contagem de células CD4+ para pacientes residentes em Salvador-Bahia.	142
6.14	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Poisson independente, multinível com intercepto aleatório e multinível com dois efeitos aleatórios para o número de células CD4+ para pacientes residentes em Salvador-Bahia.	143
6.15	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) para o modelo Poisson marginal usando a abordagem EEG com diferentes estruturas de correlação para o número de células CD4+ para pacientes residentes em Salvador-Bahia.	143
6.16	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Binomial Negativo independente e Binomial Negativo multinível com intercepto aleatório (Beta) para o número de células CD4+ para pacientes residentes em Salvador-Bahia.	144
6.17	Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando a abordagem EEG para o modelo Binomial Negativo com diferentes estruturas de correlação para o número de células CD4+ para pacientes residentes em Salvador-Bahia.	144
B.1	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	158
B.2	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	159
B.3	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	160

B.4	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	161
B.5	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	162
B.6	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	163
B.7	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	164
B.8	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	165
B.9	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	166
B.10	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	167
B.11	Valores do AIC para diferentes especificações no modelo Poisson considerando 150 indivíduos e gerando os dados utilizando o modelo Poisson multinível com intercepto aleatório.	180
B.12	Valores do AIC para diferentes especificações no modelo Poisson considerando 250 indivíduos e gerando os dados utilizando o modelo Poisson multinível com intercepto aleatório.	180

B.13	Valores do AIC para diferentes especificações no modelo Poisson considerando 150 indivíduos e gerando os dados utilizando o modelo Poisson multinível com dois efeitos aleatórios.	181
B.14	Valores do AIC para diferentes especificações no modelo Poisson considerando 250 indivíduos e gerando os dados utilizando o modelo Poisson multinível com dois efeitos aleatórios.	181
B.15	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 150 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação AR1. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	182
B.16	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 250 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação AR1. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	182
B.17	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 50 indivíduos dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	183
B.18	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 150 indivíduos dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	183
B.19	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 250 indivíduos dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada. $\beta^T = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	184

B.20	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 50 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação permutável. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	184
B.21	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) com 150 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação permutável. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	185
B.22	Valores médios das estimativas, vieses médios e <i>REQM</i> das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) com 250 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação permutável. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$	185
B.23	Valores do critério de quasi-verossimilhança (QIC) para a escolha do melhor modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada.	187
B.24	Valores do critério de quasi-verossimilhança (QIC) para a escolha do melhor modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação permutável.	187

Sumário

1	Introdução	23
1.1	Objetivos	25
1.2	Estrutura da Dissertação	25
1.3	Suporte Computacional	26
2	Modelagem para Dados de Contagem	27
2.1	Modelos Lineares Generalizados (MLG's)	28
2.2	Modelo Poisson	31
2.3	Modelo Binomial Negativo	32
2.4	Diagnóstico para MLG's	33
2.4.1	Resíduos	34
2.4.2	Distância de Cook	35
2.4.3	Gráfico Normal de Probabilidade (<i>Normal Plots</i>) com Envelopes	36
2.5	Seleção de Modelos	36
2.6	Programas Estatísticos	37
3	Modelagem para Dados Longitudinais	38
3.1	Abordagens para Análise de Dados Longitudinais	39
3.1.1	Modelos Multiníveis	44
3.1.2	Equações de Estimação Generalizadas (EEG)	48
3.2	Diagnóstico	53
3.2.1	Modelo Linear Misto	53
3.2.2	Equações de Estimação Generalizadas (EEG)	54
3.3	Seleção de Modelos	56
3.3.1	Modelos Multiníveis	56
3.3.2	Equações de Estimação Generalizadas (EEG)	57

3.4	Programas Estatísticos	58
3.4.1	R	58
3.4.2	STATA	59
4	Modelagem para Dados de Contagem Longitudinais	61
4.1	Quadratura de Gauss-Hermite	62
4.2	Modelo Poisson Multinível	63
4.2.1	Caso 1: Efeito Aleatório no Intercepto	63
4.2.2	Caso 2: Dois Efeitos Aleatórios	65
4.3	Modelo Poisson Marginal (Abordagem EEG)	68
4.3.1	Estimação	68
4.4	Modelo Binomial Negativo Multinível	69
4.4.1	Caso 1: Efeito Aleatório no Intercepto	70
4.4.2	Caso 2: Dois Efeitos Aleatórios	73
4.5	Modelo Binomial Negativo Marginal (Abordagem EEG)	75
4.5.1	Estimação	75
4.6	Modelo Marginal (Cópulas Gaussianas)	76
4.7	Programas Estatísticos	80
4.7.1	R	80
4.7.2	STATA	82
5	Estudos de Simulação	83
5.1	Geração de Variáveis Aleatórias Poisson e Binomial Negativa Correlacionadas	83
5.2	Cenários de Simulação	84
5.2.1	Geração: Modelos Multiníveis	85
5.2.2	Geração: Modelos Poisson Marginais (EEG)	86
5.2.3	Geração: Modelos BN Marginais (Cópulas Gaussianas)	88
5.3	CrITÉrios de Avaliação dos Estimadores	88
5.4	Suporte Computacional	89
5.4.1	Geração de Variáveis Aleatórias Correlacionadas	89
5.4.2	Modelo Poisson Multinível	89
5.4.3	Modelo Poisson Marginal (EEG)	90
5.4.4	Modelo Binomial Negativo Multinível	90
5.4.5	Modelo Binomial Negativo Marginal (Cópulas Gaussianas)	90
5.4.6	Problemas de Convergência	90

5.5	Resultados	91
5.5.1	Geração: Modelo Poisson Multinível com Efeito Aleatório no Intercepto	91
5.5.2	Geração: Modelo Poisson Multinível com Dois Efeitos Aleatórios	98
5.5.3	Geração: Modelo Poisson Marginal (EEG)	105
5.5.4	Geração: Modelo Binomial Negativo Multinível com Efeito Aleatório no Intercepto	110
5.5.5	Geração: Modelo Binomial Negativo Marginal (Cópulas Gaussianas)	114
6	Aplicações	119
6.1	Aplicação 1: Ataques de epilepsia	120
6.2	Aplicação 2: Estudo sobre a reforma no sistema de saúde alemão	137
6.3	Aplicação 3: Contagem de células CD4+ em pacientes portadores do vírus HIV residentes em Salvador-Bahia	141
7	Conclusões	146
	Referências	149
A	Anexo: Modelo Binomial Negativo	155
B	Estudo de Simulação	157
B.1	Modelo Poisson Multinível: Tabelas	157
B.2	Modelo Poisson Multinível: Gráficos	168
B.3	Modelo Poisson Multinível: Critério de Seleção	180
B.4	Modelo Poisson Marginal (EEG): Tabelas	181
B.5	Modelo Poisson Marginal (EEG): Gráficos	186
B.6	Modelo Poisson Marginal (EEG): Critério de Seleção	187

Capítulo 1

Introdução

Bebês engatiam e andam, crianças aprendem a ler e escrever e idosos tornam-se frágeis e esquecidos são exemplos de mudanças, tanto físicas como mentais, dos indivíduos ao longo do tempo. Além destas mudanças naturais, intervenções específicas também podem causar, por exemplo, a diminuição dos níveis de colesterol por uso de um novo medicamento, ou ainda, a alteração de aprendizado após introdução de novas técnicas de ensino. Ao medir e mapear tais mudanças, tanto naturalmente quanto induzidas experimentalmente, pode-se avaliar como as mesmas foram desencadeadas (Singer & Willett, 2003).

A avaliação das mudanças ao longo do tempo pode ser feita por meio de estudos experimentais ou observacionais. Além disso, os dados podem ser coletados prospectivamente ou retrospectivamente e o tempo pode ser mensurado em várias unidades: semanas, semestres, meses, anos, etc. O período da coleta dos dados pode ainda ser fixo, todos possuem o mesmo tempo, ou flexível, cada pessoa tem um período (Singer & Willett, 2003).

Pode-se medir as mudanças individuais quando se tem dados longitudinais. Uma das características mais importantes que define um estudo longitudinal é que os indivíduos são medidos repetidamente ao longo do tempo (Diggle et al., 1994) ou em qualquer escala ordenada. Através desse tipo de desenho de estudo é possível caracterizar o comportamento da variável resposta ao longo do tempo, investigar o efeito de variáveis independentes sobre o perfil da variável resposta, fazer predição para a variável resposta ou, ainda, avaliar mudanças globais ou individuais nesta resposta ao longo do tempo (Singer & Andrade, 1986).

Uma etapa importante na avaliação dos dados é a análise descritiva que, no caso de estudos longitudinais, pode revelar o comportamento de cada indivíduo que foi mensurado ao longo do tempo. A maneira mais simples de visualizar como um

indivíduo específico se comporta ao longo do tempo é examinar o gráfico de perfil individual (Figura 1.1-(a)). Por outro lado, para verificar o comportamento conjunto de todos indivíduos deve-se analisar o gráfico de perfil médio, visualizado na Figura 1.1-(b).

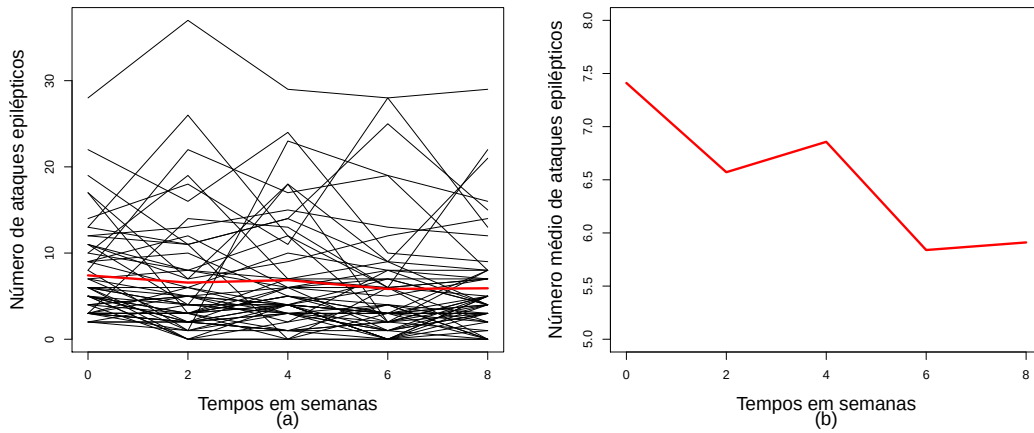


Figura 1.1: Gráficos do (a) perfil individual e (b) perfil médio do número de ataques epiléticos.

A investigação dessas mudanças tem fascinado pesquisadores por gerações. No entanto, foi apenas a partir da década de 80 que uma classe de modelos estatísticos apropriados foi desenvolvida tornando os pesquisadores capazes de estudar melhor as mudanças individuais ao longo do tempo. Essa classe de modelos é conhecida como modelos multiníveis ou modelos lineares hierárquicos.

O modelo multinível pode apresentar dois ou mais níveis e está associado à efeitos aleatórios que descrevem um indivíduo específico e definem a estrutura de correlação entre as medidas repetidas. Desta forma a interpretação destes modelos é individual e está condicionada aos efeitos aleatórios (por este motivo pode ser considerado um modelo condicional). Os modelos multiníveis são modelos mais flexíveis e apresentam maior facilidade no ajuste para estruturas mais complexas.

Para ilustrar a ideia dos efeitos aleatórios, considere a Figura 1.1-(a). É possível notar claramente que cada indivíduo possui um número médio de ataques epiléticos. Além disso, a variabilidade da variável de interesse ao longo do tempo difere de indivíduo para indivíduo. Isto pode ser explicado pela presença de efeitos aleatórios associado ao intercepto (b_{0i}) e de efeitos aleatórios relacionado ao parâmetro associado a covariável tempo (b_{1i}).

Uma segunda abordagem para analisar dados longitudinais é conhecida como

modelo marginal. O termo marginal, neste contexto, é utilizado para enfatizar que o modelo para a resposta média em cada ocasião depende apenas das covariáveis de interesse e não incorporam a dependência das respostas anteriores e/ou efeitos aleatórios. Isso é um contraste com os modelos multiníveis, em que a resposta média é modelada não apenas como uma função das covariáveis, mas é condicionada a efeitos aleatórios, que proporcionam a variação tanto para o valor médio da resposta para cada indivíduo, como ao longo do tempo (Fitzmaurice et al., 2009).

Quando a variável de interesse é uma contagem, as distribuições Poisson e Binomial Negativa são as frequentemente usadas para a sua modelagem. Em um modelo de regressão tradicional para variáveis discretas, por exemplo, um modelo de regressão Poisson, assume-se que as observações são independentes. Porém, em dados provenientes de estudos longitudinais com interesse em avaliar a mudança na variável resposta ao longo do tempo, a suposição de independência entre as observações não é adequada. Portanto, os modelos multiníveis e marginais podem ser estratégias alternativas para lidar com a análise de dados de contagem longitudinais.

1.1 Objetivos

Os objetivos desta dissertação são:

1. Revisar os principais resultados teóricos acerca dos modelos multiníveis e marginais para dados contínuos longitudinais;
2. Apresentar algumas extensões dos modelos multiníveis e marginais para dados de contagem longitudinais;
3. Avaliar alguns modelos multiníveis e marginais para o ajuste de dados de contagem longitudinais através de estudos de simulação;
4. Aplicar os modelos multiníveis e marginais à análise de dados reais;
5. Desenvolver um material didático para usuários da área de saúde.

1.2 Estrutura da Dissertação

Esta dissertação está estruturada da seguinte forma. No Capítulo 2 alguns modelos para dados de contagem são discutidos, bem como técnicas de diagnóstico utilizadas para verificar o ajuste dos modelos. Os modelos utilizados na análise de dados longitudinais com resposta contínua são apresentados sucintamente no

Capítulo 3. Seguidamente, no Capítulo 4 são estendidas estas ideias para dados longitudinais com resposta de contagem. Os estudos de simulação e as aplicações com dados reais são apresentados nos Capítulos 5 e 6, respectivamente. As conclusões são discutidas e apresentadas no Capítulo 7.

1.3 Suporte Computacional

Os programas estatísticos utilizados para as análises dos dados reais e os estudos de simulação foram o **R** versão 3.0.2 e o STATA versão 10. O **R** é uma linguagem e um ambiente de desenvolvimento integrado para cálculos estatísticos e gráficos, sendo gratuito e disponível em www.r-project.org. O STATA (www.stata.com) é um programa estatístico que foi desenvolvido em C e lançado em 1985. A versão mais atual do STATA é a 13. Há versões do **R** e STATA para Windows, Macintosh, Linux e Unix.

Os códigos utilizados para os estudos de simulação, para a construção dos gráficos e para a construção dos envelopes simulados encontram-se disponíveis em www.de.ufpe.br/~raydonal/Tese/Supplementary-Materials/Trindade-2014.

Capítulo 2

Modelagem para Dados de Contagem

Dados de contagem apresentam apenas valores inteiros não negativos, isto é, assumem os valores 0, 1, 2, Pode-se citar, como exemplos, a contagem de ataques de pânico ocorridos durante um intervalo de tempo, ou ainda o número de parceiros sexuais em um período de três meses de indivíduos registrados em um programa de prevenção ao HIV (Diggle et al., 1994).

Para o ajuste dos dados de contagem, modelos de regressão usualmente não lineares são os mais adequados uma vez que eles levam em consideração as propriedades específicas dos dados, incluindo a falta de continuidade e a não negatividade na variável resposta.

Os dados de contagem são comumente modelados com a distribuição Poisson. Uma característica importante desta distribuição é que o valor esperado e a variância das respostas são iguais, ou seja, $E(Y) = \text{Var}(Y)$. Porém, na vida real isto não acontece (ou acontece raramente). Na presença, por exemplo, de superdispersão (um fenômeno que aparece quando a variabilidade dos dados é maior do que a média e que pode surgir da heterogeneidade entre os indivíduos e da dependência entre os eventos) não é mais adequado utilizar esta distribuição. Existem extensões do modelo Poisson para lidar com a presença de superdispersão, como o modelo Poisson Inflacionado de Zeros (*ZIP*, em inglês). Um modelo mais simples que pode ser utilizado nesta situação é o modelo Binomial Negativo (BN) que pode ser representado como uma mistura das distribuições Gama e Poisson, denotadas respectivamente por $G(\cdot, \cdot)$ e $Po(\cdot)$. Vale ressaltar que o modelo de Série de Potências (MSP) e o MSP com excesso de zeros generalizam os modelos Poisson (e extensões) e BN.

É possível encontrar na literatura uma vasta discussão sobre dados de contagem e o uso das distribuições Poisson e BN para o ajuste de modelos de regressão.

Por exemplo, Barron (1992) descreveu vários métodos utilizados para analisar dados de contagem, realizou estudos de simulação para avaliar as propriedades dos estimadores na presença de autocorrelação e superdispersão e investigou os métodos discutidos em conjuntos de dados reais. Lindsey (1995) apresentou técnicas de modelagem para dados de contagem e frequências. Gupta et al. (1996) desenvolveram um modelo para analisar dados de contagem que apresentam excesso de zeros. Winkelmann (2003) discutiu sobre a análise de dados de contagem em contextos econométricos. Cameron & Trivedi (2013) escreveram sobre diferentes aspectos da análise de regressão para dados de contagem.

Ao lidar com dados de contagem, uma abordagem possível é ajustar um Modelo Linear Generalizado (MLG) com distribuição Poisson e uma função de ligação canônica (neste caso o logaritmo) a fim de modelar a média. Neste capítulo, na Seção 2.1 é realizada uma revisão sucinta sobre os Modelos Lineares Generalizados (MLG's). Nas Seções 2.2 e 2.3 são discutidos os modelos Poisson e BN, respectivamente, e suas principais características. Na Seção 2.4 são apresentadas algumas técnicas de diagnóstico utilizadas para avaliar o ajuste desta classe de modelos. As bibliotecas e algumas funções utilizadas nos programas estatísticos **R** e STATA para ajustar tais modelos são destacadas na Seção 2.6.

2.1 Modelos Lineares Generalizados (MLG's)

Modelos para a análise de dados não normais usando modelos não lineares vêm sendo utilizados com frequência em diferentes áreas de conhecimento. Em 1934, Bliss descreveu um experimento no qual se aplicava nicotina em insetos e a proporção de mortos era anotada, realizando assim uma das primeiras análises de regressão probit. Um ano depois, Fisher utilizou o método de máxima verossimilhança para obter os estimadores do modelo probit (McCulloch & Searle, 2001).

Uma nova classe de modelos de regressão que permitisse incorporar essas novas ideias de regressão não linear foi definida por Nelder & Wedderburn (1972) e denotada como Modelos Lineares Generalizados (MLG's). Os MLG's são caracterizados por

- Um componente aleatório, especificando a distribuição condicional da variável resposta. Desta forma, tem-se que $\mathbf{Y}_i = (Y_1, \dots, Y_m)^\top$ é um vetor de variáveis aleatórias independentes seguindo uma densidade na família exponencial linear, isto é, $\mathbf{Y}_i \sim f_{\mathbf{Y}_i}(y_i; \theta_i, \phi)$ em que

$$f_{\mathbf{Y}_i}(y_i; \theta_i, \phi) = \exp \{ \phi^{-1} [y_i \theta_i - b(\theta_i)] + c(y_i, \phi) \}, \quad (2.1)$$

com ϕ^{-1} sendo o parâmetro de dispersão, θ_i o parâmetro de posição, $b(\theta_i)$ uma função que depende apenas de θ_i e $c(y_i, \phi)$ uma função não constante de y_i e/ou ϕ . Alguns exemplos de distribuições que pertencem à família exponencial são Normal, Binomial, Poisson, Binomial Negativa entre outras.

- Um preditor linear, que é uma função dos regressores

$$\eta_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_p X_{ip},$$

em que $\mathbf{X}_i^\top = (X_{i1}, \dots, X_{ip})$ denota o vetor de variáveis explicativas para o i -ésimo indivíduo, $i = 1, \dots, m$ (em geral, $X_{i1} = 1, \forall i$) e $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ como o vetor de parâmetros a ser estimado.

- Uma função de ligação, $h(\cdot)$, duplamente diferenciável, que relaciona o parâmetro de posição θ_i com as variáveis explicativas. Pode-se também modelar uma transformação da média, μ_i , que é alguma função de θ_i , como uma função linear nos parâmetros, isto é,

$$h(E(\mathbf{Y}_i)) = h(\mu_i) = \mathbf{X}_i^\top \boldsymbol{\beta}.$$

Segundo Paula (2013), é possível mostrar, sob algumas condições usuais de regularidade (Cox & Hinkley, 1979), que

$$\begin{aligned} E \left[\frac{\partial \ell_i(\theta_i)}{\partial \theta_i} \right] &= 0 \quad \text{e} \\ \text{Var} \left[\frac{\partial \ell_i(\theta_i)}{\partial \theta_i} \right] &= -E \left[\frac{\partial^2 \ell_i(\theta_i)}{\partial \theta_i^2} \right], \end{aligned} \quad (2.2)$$

em que $\ell_i(\theta_i) = \log f_{\mathbf{Y}_i}(y_i; \theta_i, \phi)$ é o logaritmo da função de densidade.

Usando (2.1) em (2.2), tem-se

$$E \left[\phi \left(\mathbf{Y}_i - \frac{\partial b(\theta_i)}{\partial \theta_i} \right) \right] = 0$$

o que implica

$$E[\mathbf{Y}_i] = \mu_i = \frac{\partial b(\theta_i)}{\partial \theta_i}.$$

Pode-se verificar que $\text{Var}(\mathbf{Y}_i) = \phi^{-1} v(\mu_i)$, em que $v(\mu_i) = \frac{\partial^2 b(\theta_i)}{\partial \theta_i^2}$ é chamada de “função de variância” e $\phi^{-1} > 0$ é o parâmetro de dispersão.

Após escolher um determinado modelo distribucional na família exponencial, é necessário estimar seus parâmetros e avaliar a precisão das estimativas. No caso dos MLG's, os parâmetros podem ser estimados pelo método da máxima verossimilhança (MV). Para obter as estimativas dos parâmetros, deve-se maximizar o logaritmo da função de verossimilhança, em relação aos parâmetros, supondo fixos os dados observados. O logaritmo da função de verossimilhança para (2.1) é dado por

$$\ell(\boldsymbol{\theta}) = \log \prod_{i=1}^m f(y_i) = \ell(\boldsymbol{\theta}_i, \phi; \mathbf{y}) = \sum_{i=1}^m \phi^{-1} [y_i \theta_i - b(\theta_i)] + \sum_{i=1}^m c(y_i, \phi), \quad (2.3)$$

em que $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi)^\top$ denota uma partição (Paula, 2013).

As funções escores de $\boldsymbol{\beta}$ e ϕ são dadas, respectivamente, por

$$\begin{aligned} U_{\boldsymbol{\beta}}(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \phi \mathbf{X}^\top \mathbf{W}^{1/2} \mathbf{V}^{-1/2} (\mathbf{y} - \boldsymbol{\mu}), \\ U_{\phi}(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \phi} = \sum_{i=1}^m [y_i \theta_i - b(\theta_i)] + \sum_{i=1}^m \frac{\partial c(y_i, \phi)}{\partial \phi}. \end{aligned} \quad (2.4)$$

em que $\mathbf{X}^\top = (X_1, \dots, X_m)$ é uma matriz $m \times p$ de posto completo, $\mathbf{y} = (y_1, \dots, y_m)^\top$ é o vetor de respostas, $\mathbf{W} = \text{diag}\{\omega_1, \dots, \omega_m\}$ denota a matriz diagonal de pesos denotados por w_i , $\mathbf{V} = \text{diag}\{v(\mu_1), \dots, v(\mu_m)\}$ é uma matriz diagonal contendo as funções de variância e $\boldsymbol{\mu} = (\mu_1, \dots, \mu_m)^\top$ são as respectivas médias.

Maximizar (2.3) é equivalente a solucionar o sistema (2.4), que não possui forma analítica fechada, na maioria dos casos, sendo, desta forma, necessário utilizar métodos iterativos para obter a estimativa de máxima verossimilhança de $\boldsymbol{\beta}$ e ϕ . No caso da $\boldsymbol{\beta}$, um processo usualmente utilizado é o método iterativo de Newton-Raphson que consiste em expandir a função escore $U_{\boldsymbol{\beta}}$ em torno de um valor inicial $\boldsymbol{\beta}^0$ de tal forma que

$$U_{\boldsymbol{\beta}} \cong \mathbf{U}_{\boldsymbol{\beta}}^{(0)} + \mathbf{U}_{\boldsymbol{\beta}}^{'(0)} (\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)}),$$

em que U' denota a primeira derivada de $U_{\boldsymbol{\beta}}$ com respeito a $\boldsymbol{\beta}^\top$, sendo $\mathbf{U}_{\boldsymbol{\beta}}^{'(0)}$ e $\mathbf{U}_{\boldsymbol{\beta}}^{(0)}$, respectivamente, quantidades avaliadas em $\boldsymbol{\beta}^{(0)}$. Repetindo-se o procedimento acima tem-se o processo iterativo

$$\boldsymbol{\beta}^{(k+1)} = \boldsymbol{\beta}^{(k)} + [(-U'_{\boldsymbol{\beta}})^{-1}]^{(k)} U_{\boldsymbol{\beta}}^{(k)}, \quad k = 0, 1, 2, \dots$$

Note que $U_{\boldsymbol{\beta}}^{(k)}$ denota o vetor escore com elementos $(\partial \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta})$ e $(-U'_{\boldsymbol{\beta}})^k$ é a matriz de informação observada com elementos $(-\partial^2 \ell(\boldsymbol{\theta}) / \partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^\top)$, avaliados em $\boldsymbol{\beta} = \boldsymbol{\beta}^{(k)}$.

Como a matriz U'_{β} pode não ser positiva definida, a aplicação do método escore de Fisher substituindo a matriz $-U'_{\beta}$ pelo correspondente valor esperado $K_{\beta}\beta$ pode ser mais adequada. Assim,

$$\beta^{(k+1)} = \beta^{(k)} + [K_{\beta}^{-1}]^{(k)} U_{\beta}^{(k)}, \quad k = 0, 1, 2, \dots$$

Maiores detalhes ver Paula (2013).

2.2 Modelo Poisson

A distribuição Poisson é considerada uma das distribuições discretas mais importantes. Ela é comumente utilizada para analisar o número de ocorrências de um evento durante um intervalo fixado de tempo, distância, área ou volume. Maiores detalhes em Poisson (1837).

Suponha que Y_i seja o número de eventos ocorridos para o i -ésimo indivíduo em um intervalo de tempo. Considere que Y_i é uma variável aleatória que segue a distribuição Poisson com parâmetro μ_i , denotado por $Y_i \sim Po(\mu_i)$, então a função de probabilidade (f.p.) é dada por

$$P(Y_i = y_i) = \frac{\exp(-\mu_i) \mu_i^{y_i}}{y_i!}, \quad \text{para } y_i = 0, 1, 2, \dots \quad (2.5)$$

em que $\mu_i > 0$ é o valor esperado de Y_i . É possível reescrever (2.5) na forma da família exponencial (2.1), ou seja,

$$P(Y_i = y_i) = \exp\{[y_i \log \mu_i - \mu_i] - \log(y_i!)\},$$

em que o parâmetro de dispersão é $\phi = 1$, a função de ligação canônica é $\log(\mu_i)$, $b(\theta_i) = e^{\theta_i} = e^{\log(\mu_i)} = \mu_i$, $\frac{\partial b(\theta_i)}{\partial \theta_i} = e^{\theta_i} = \mu_i$, $c(y_i, \phi) = -\log(y_i!)$. Desta forma $\text{Var}(Y_i) = \phi^{-1} \frac{\partial^2 b(\theta_i)}{\partial \theta_i^2} = \mu_i$. O preditor linear é $\eta_i = X_i^\top \beta$ e desta forma $\mu_i = \exp(X_i^\top \beta)$, $i = 1, \dots, m$, para $X_i^\top = (X_{i1}, \dots, X_{ip})$ e $\beta = (\beta_1, \dots, \beta_p)^\top$.

Uma das características da distribuição Poisson é que $E(Y_i) = \text{Var}(Y_i) = \mu_i$. No caso de superdispersão, $E(Y_i) < \text{Var}(Y_i)$, existem técnicas alternativas de estimação dos parâmetros do modelo a fim de corrigir esse problema (Hinde & Demétrio, 1998). Neste caso, a suposição inicial de que o parâmetro de dispersão $\phi = 1$ não é mais verificada, surgindo então a necessidade de lidar com a variabilidade extra.

Algumas soluções podem ser consideradas: (i) uma abordagem bayesiana, na qual pode ser assumido que o parâmetro do modelo pode ser representado por uma distribuição de probabilidade; (ii) a estimação por quase-verossimilhança, incluindo um fator de dispersão diferente da unidade ou uma função de variância alternativa ou (iii) a utilização do modelo Binomial Negativo, entre outras.

2.3 Modelo Binomial Negativo

Vários artigos sobre a utilização da distribuição Binomial Negativa para análise de dados de contagem estão disponíveis na literatura. Miaou (1994) avaliou a performance dos modelos Poisson e Binomial Negativo para estabelecer a relação entre acidentes de caminhões e o desenho geométrico das estradas. Gardner et al. (1995) discutiram a aplicação dos modelos para dados de contagem na psicologia.

Em teoria da probabilidade e estatística, a distribuição Binomial Negativa (BN) ou distribuição de Pascal é uma distribuição de probabilidade discreta do número de sucessos em uma sequência de ensaios de Bernoulli antes de r número de falhas ocorrerem.

Suponha que haja uma sequência de ensaios de Bernoulli independentes, em cada experimento, tendo dois resultados potenciais definidos com sucesso, com probabilidade p e fracasso com probabilidade $(1 - p)$. O interesse é observar esta sequência até que um número definido previamente de r falhas ocorram. Então, o número de sucessos, Y , tem distribuição BN, denotada por $Y \sim BN(r, p)$.

A função de probabilidade (f.p.) da distribuição BN é dada por:

$$f_Y(y; r, p) = P(Y = y) = \binom{y + r - 1}{y} p^y (1 - p)^r, \quad y = 0, 1, 2, \dots,$$

em que $\binom{y + r - 1}{y} = \frac{(y + r - 1)!}{y!(r - 1)!}$.

Desta forma, o valor esperado e a variância da resposta são dados, respectivamente, por: $E(Y) = \frac{pr}{1-p}$ e $\text{Var}(Y) = \frac{pr}{(1-p)^2}$.

Quando se verifica a presença da superdispersão é possível considerar a suposição de que, condicionada aos parâmetros, a variável resposta, Y_i , segue a distribuição Poisson. Assim, os parâmetros da distribuição Poisson são tratados como variáveis aleatórias com distribuições conhecidas. Portanto, a distribuição marginal de Y_i é Binomial Negativa (BN) se a distribuição a priori dos parâmetros for Gama.

Para evitar notação sobrecarregada considere que $Y_i = Y$. Sejam duas variáveis aleatórias (v.a.'s) Y e Z , tal que $Y|Z$ segue uma distribuição Poisson, $Y|Z \sim Po(Z)$, e que a distribuição a priori de Z é Gama com parâmetros $\phi\mu$ e ϕ . Logo,

$$Y|Z \sim Po(Z), \quad \text{sendo} \quad f_{Y|Z}(y|z) = e^{-z} \frac{z^y}{y!} 1_{\{0,1,\dots\}}(y) \quad \text{e}$$

$$Z \sim G(\phi\mu, \phi), \quad \text{sendo} \quad g_Z(z) = \frac{\phi^{\phi\mu}}{\Gamma(\phi\mu)} z^{\phi\mu-1} e^{-\phi z} 1_{(0,\infty)}(z).$$

A distribuição marginal de Y é dada pela integração da distribuição conjunta de Y e Z em relação a Z . Desta forma $Y \sim BN(\phi\mu, \phi/(\phi+1))$ com f.p. dada por:

$$f_Y(y) = \frac{(y + \phi\mu - 1)!}{y!(\phi\mu - 1)!} \left(\frac{\phi}{\phi + 1} \right)^{\phi\mu} \left(\frac{1}{\phi + 1} \right)^y 1_{\{0,1,\dots\}}(y). \quad (2.6)$$

Portanto, $E(Y) = \mu$ e $\text{Var}(Y) = \mu(\phi + 1)/\phi$.

Outra possibilidade é considerar que $Y|Z \sim Po(Z)$ e Z tem distribuição a priori Gama com parâmetros ν e μ/ν , isto é

$$Y|Z \sim Po(Z), \quad \text{sendo} \quad f_{Y|Z}(y|z) = e^{-z} \frac{z^y}{y!} 1_{\{0,1,\dots\}}(y) \quad \text{e}$$

$$Z \sim G\left(\nu, \frac{\nu}{\mu}\right), \quad \text{sendo} \quad g_Z(z) = \frac{\left(\frac{\mu}{\nu}\right)^\nu}{\Gamma(\nu)} z^{\nu-1} e^{-z\frac{\nu}{\mu}} 1_{(0,\infty)}(z)$$

Assim, a distribuição marginal de Y é BN, com uma função de variância quadrática, é dada por:

$$f_Y(y) = \frac{(y + \nu - 1)!}{y!(\nu - 1)!} \left(\frac{\frac{\nu}{\mu}}{\frac{\nu}{\mu} + 1} \right)^\nu \left(\frac{1}{\frac{\nu}{\mu} + 1} \right)^y 1_{\{0,1,\dots\}}(y), \quad (2.7)$$

de tal forma que $E(Y) = \mu$ e $\text{Var}(Y) = \mu + \frac{\mu^2}{\nu}$.

Note que as distribuições (2.6) e (2.7) têm variância maior do que a média, o que é vantajoso na modelagem para o caso de dados que apresentam superdispersão. Os cálculos algébricos encontram-se disponíveis no Anexo A. Maiores detalhes podem ser encontrados em Lawless (1987).

2.4 Diagnóstico para MLG's

Uma etapa importante no ajuste de um modelo de regressão é a verificação de possíveis afastamentos das suposições feitas para o modelo, especialmente para o componente aleatório e para a parte sistemática do modelo, bem como a avaliação do efeito de observações discrepantes nos resultados do ajuste. Tal etapa, conhecida como análise de diagnóstico é geralmente inicializada com a análise de resíduos a fim de detectar a presença de possíveis pontos discrepantes (*outliers*, em inglês) e avaliar a adequação da distribuição proposta para a variável resposta. As técnicas usadas para análise de resíduos e diagnóstico para os MLG's são semelhantes às utilizadas no modelo clássico de regressão.

Os resíduos devem expressar a discrepância entre a observação y_i e o seu valor ajustado $\hat{\mu}_i$, sendo mensurado por

$$r_i = \varrho_i(y_i, \hat{\mu}_i),$$

em que ϱ_i é uma função conhecida e de fácil interpretação.

Uma outra quantidade de interesse é a matriz de projeção $\mathbf{H}$ a qual é definida nos MLG's por:

$$\mathbf{H} = \mathbf{W}^{1/2} \mathbf{X} (\mathbf{X}^\top \mathbf{W} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{W}^{1/2}, \quad (2.8)$$

em que $0 \leq h_{ii} \leq 1$ e $\text{tr}(\mathbf{H}) = p$.

Na prática $\mathbf{H}$ é função de μ e ϕ . Logo deve ser estimada e assim temos $\hat{\mathbf{H}}$.

Paula (2013) apresenta algumas técnicas de diagnóstico utilizadas em MLG's, como, por exemplo, os elementos da diagonal principal da matriz de projeção, a distância de Cook e os resíduos do modelo ajustado, que podem ser usados para detectar a influência dos regressores sobre a resposta.

2.4.1 Resíduos

O resíduo de Pearson é o mais simples e mais comum nos MLG's, sendo definido por:

$$r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\widehat{\text{Var}}(y_i)}}, \quad i = 1, \dots, m, \quad (2.9)$$

em que $\hat{\mu}_i$ e $\widehat{\text{Var}}(y_i)$ são a média ajustada e a função de variância ajustada de y_i , obtidas com o ajuste do modelo em estudo. Os resíduos de Pearson são muito utilizados para identificação de possíveis valores discrepantes.

Este resíduo pode ser visto com uma componente da estatística de Pearson generalizada $X_p^2 = \sum_{i=1}^m r_i^{p^2}$. A desvantagem do resíduo de Pearson é que sua distribuição é bastante assimétrica para modelos não-normais.

No caso da distribuição Poisson o resíduo de Pearson é dado por

$$r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i}}, \quad i = 1, \dots, m.$$

Considerando o modelo Binomial Negativo apresentado em (2.7) os resíduos de Pearson são dados por

$$r_i^P = \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i + \hat{\mu}_i/\nu}}, \quad i = 1, \dots, m,$$

em que ν pode ser fixo ou estimado.

Outros resíduos bastante utilizados são os resíduos de Pearson Studentizados, os quais são definidos por:

$$r_i^{P'} = \frac{y_i - \hat{\mu}_i}{\sqrt{\widehat{\text{Var}}(y_i)(1 - h_{ii})}}, \quad i = 1, \dots, m,$$

em que h_{ii} é o i -ésimo elemento da diagonal da matriz ortogonal de projeção dada em (2.8). Vale ressaltar que, na prática, $\mathbf{H}$ é função de μ e ϕ , portanto deve ser estimada. Portanto, h_{ii} é na verdade $\hat{h}_{ii}$.

Para as distribuições Poisson e BN os resíduos de Pearson Studentizados, são dados, respectivamente, por:

$$\begin{aligned} r_i^{P'} &= \frac{y_i - \hat{\mu}_i}{\sqrt{\hat{\mu}_i(1 - h_{ii})}} \quad \text{e} \\ r_i^{P'} &= \frac{y_i - \hat{\mu}_i}{\sqrt{(\hat{\mu}_i + \hat{\mu}_i/\nu)(1 - h_{ii})}} \quad i = 1, \dots, m. \end{aligned}$$

Um outro tipo de resíduo bastante utilizado nos MLG's é o resíduo componente do desvio (resíduo do desvio), definido pela raiz quadrada da diferença entre o logaritmo da função de verossimilhança sob o modelo saturado ($\hat{\ell}_{sat}$) e a função de verossimilhança sob o modelo corrente ($\hat{\ell}_{cor}$) para cada uma das observações, com sinal dado pelo sinal de $(y_i - \hat{\mu}_i)$. O resíduo componente do desvio para o i -ésimo caso é a contribuição da i -ésima observação na soma do desvio do modelo, ou seja,

$$\begin{aligned} r_i^D &= \text{sinal}(y_i - \hat{\mu}_i) \sqrt{2(\hat{\ell}_{sat} - \hat{\ell}_{cor})} \\ &= \text{sinal}(y_i - \hat{\mu}_i) \left\{ 2y_i[\tilde{\theta}_i - \hat{\theta}_i] - 2b(\tilde{\theta}_i) + 2b(\hat{\theta}_i) \right\}^{1/2}, \end{aligned} \quad (2.10)$$

em que $\tilde{\theta}_i = \theta(y_i)$ e $\hat{\theta}_i = \theta(\hat{\mu}_i)$ representando, respectivamente, as estimativas de máxima verossimilhança de θ para o modelo saturado (m parâmetros) e para o modelo com p parâmetros ($p < m$).

É possível observar que o resíduo (2.10) representa a distância da observação y_i ao seu valor ajustado $\hat{\mu}_i$, medida na escala do logaritmo da função de verossimilhança. Finalmente, os resíduos componentes do desvio Studentizados são definidos a partir de (2.10), como se segue

$$r_i^{D'} = \frac{r_i^D}{\sqrt{1 - h_{ii}}}, \quad i = 1, \dots, m. \quad (2.11)$$

2.4.2 Distância de Cook

A distância de Cook (Cook, 1977) originalmente desenvolvida para modelos normais lineares, foi rapidamente assimilada e estendida para diversas classes de

modelos. Ela mede o efeito na estimação do vetor de parâmetros ao se excluir uma dada observação, de modo que os pontos com grande distância de Cook devem ser considerados para verificação.

A distância de Cook no modelo de regressão linear é definida por

$$D_i = \frac{(\hat{\beta} - \hat{\beta}_i)^\top (\mathbf{X}^\top \mathbf{X}) (\hat{\beta} - \hat{\beta}_i)}{ps^2}, \quad (2.12)$$

em que, $\hat{\beta} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$, $\hat{\beta}_i$ denota o estimador de β sem o i -ésimo indivíduo e $s^2 = (r_i^p)^2 / (m - p)$. Portanto, D_i será grande quando o i -ésimo ponto for aberrante e/ou quando h_{ii} for próximo de 1 (Paula, 2013).

Uma versão da distância de Cook para MLG's é dada por

$$D_i = \frac{\hat{h}_{ii}}{(1 - \hat{h}_{ii})^2} (r_i^P)^2.$$

2.4.3 Gráfico Normal de Probabilidade (*Normal Plots*) com Envelopes

O gráfico normal de probabilidades destaca-se por dois aspectos: a identificação da distribuição originária dos dados (hipótese distribucional) e a identificação de valores que se destacam no conjunto de observações. Os envelopes, no caso de MLG's com distribuições diferentes da normal, são construídos com os resíduos sendo gerados a partir do modelo ajustado (Paula, 2013).

O principal objetivo do gráfico de envelope simulado é verificar a adequabilidade do modelo ajustado. Se muitos pontos estiverem fora da faixa de confiança o ajuste não é recomendado.

Em <http://www.ime.usp.br/~giapaula/textoregressao.htm> estão disponíveis os códigos em **R** para a construção do gráfico normal de probabilidade com envelopes para vários modelos, inclusive para os modelos Poisson e BN apresentados neste capítulo. Venezuela (2003) apresenta o algoritmo para a construção do gráfico de probabilidade meio-normal com envelope simulado.

2.5 Seleção de Modelos

A escolha do melhor modelo de regressão tradicional pode ser feita através do Critério de Informação de Akaike (AIC, em inglês), sendo que aparentemente o modelo mais adequado será aquele com menor valor de AIC.

Este critério é definido como

$$AIC = -2L + 2p,$$

em que L é o logaritmo da função de verossimilhança e p denota o número de parâmetros no modelo (Cui, 2010). Este tipo de critério pode ser usado para os modelos lineares, logísticos e de contagem.

2.6 Programas Estatísticos

A linguagem de programação **R** apresenta bibliotecas e funções para o ajuste dos modelos Poisson e Binomial Negativo. A biblioteca **stats** do **R** contém funções para cálculos estatísticos e geração de números aleatórios. A densidade, a função de distribuição, a função quantílica e a geração dos números aleatórios para a distribuição Poisson com parâmetro μ são dadas, respectivamente, por **dpois**, **ppois**, **qpois** e **rpois**. O mesmo pode ser considerado para a distribuição Binomial Negativa substituindo **pois** por **nbinom**, ou seja, **dnbinom**, **pnbinom**, **qnbinom** e **rnbinom** que denotam, respectivamente, a densidade, a função de distribuição, a função quantílica e a geração dos números aleatórios para a distribuição Binomial Negativa com parâmetros **size** (parâmetro de dispersão) e **prob** (probabilidade de sucesso). Para maiores detalhes sobre as funções do pacote **stats** digitar no PROMPT de comando do **R** `library(help = "stats")`.

Para ajustar modelos de regressão Poisson e Binomial Negativo no **R** pode-se usar a função **glm**. Esta função ajusta os MLG's utilizando métodos de otimização não linear como Newton-Raphson (*default*). A distribuição da variável resposta é definida no argumento **family**, sendo **poisson** e **nbinomial** para as distribuições Poisson e Binomial Negativa, respectivamente. Nesta função o argumento **link** indica a função de ligação que, por *default*, é logaritmo para ambas distribuições. Para maiores detalhes desta função digitar no PROMPT de comando do **R** `help(glm)`.

Também é possível ajustar os modelos Poisson e Binomial Negativo utilizando o *software* STATA através das funções **poisson** e **nbreg**, respectivamente. Maiores detalhes ver o manual do STATA em <http://www.stata.com/help.cgi?poisson> e <http://www.stata.com/help.cgi?nbreg> para distribuições Poisson e Binomial Negativa, respectivamente.

Capítulo 3

Modelagem para Dados Longitudinais

Os estudos longitudinais desempenham um papel importante em diversas áreas do conhecimento como nas ciências sociais, ciências comportamentais, saúde pública e ciências biológicas, bem como na agricultura, educação, economia e marketing, pois nestas áreas podem ser estudadas as variações ao longo do tempo de uma determinada variável de interesse. Em bioestatística estes estudos são conhecidos como estudos de coorte, ao passo que em outros campos do conhecimento, como sociologia, economia e administração são chamados de estudos de painéis (Singer et al., 2012).

Se os indivíduos são observados de forma repetida ao longo do tempo, um estudo longitudinal permite a verificação das taxas de mudança que ocorrem ao longo do tempo e dos fatores que podem influenciar em tal variação. Duas importantes características de dados longitudinais são a correlação entre as respostas dentro das unidades amostrais e a ordenação cronológica das respostas (Skrondal & Rabe-Hesketh, 2004).

Contudo, um dos inconvenientes associados aos estudos longitudinais é o alto custo, tanto financeiro quanto computacional, para a sua execução. O custo financeiro está relacionado aos recursos necessários para o acompanhamento dos participantes ao longo do tempo a fim de realizar a mensuração das medidas de interesse. Já o custo computacional está relacionado à utilização de métodos estatísticos mais complexos que levam em consideração a correlação entre as medidas repetidas em cada indivíduo.

Os estudos com dados longitudinais apresentam alguns benefícios, como a possibilidade de registrar os eventos incidentes, ou seja, um determinado estudo longitudinal prospectivo (que ocorre baseado em um protocolo de pesquisa e inclui indivíduos

após a idealização do protocolo) mensura, por exemplo, uma nova ocorrência de um evento de interesse e o momento do seu aparecimento pode estar relacionado com as mudanças recentes na exposição do indivíduo a um fator de risco. Além disso, existe a mensuração da variação individual nas respostas de interesse.

Em estudos longitudinais, os dados associados a cada unidade amostral podem ser expressos na forma de um vetor, cujos elementos são os valores da variável resposta em cada instante de observação, e de uma matriz, cujos elementos correspondem aos valores das variáveis explicativas, que podem variar entre unidades amostrais (denotadas por variáveis tempo-independentes) ou podem variar também dentro das unidades amostrais (denotadas por variáveis tempo-dependentes) (Singer et al., 2012).

Em relação à coleta de dados, tem-se que os dados longitudinais são balanceados quando todas as unidades amostrais são medidas no mesmo conjunto de pontos do tempo t_j , $j = 1, \dots, n$ e desbalanceados se as unidades são mensuradas em diferentes pontos no tempo t_{ij} , $i = 1, \dots, m$ e $j = 1, \dots, n_i$, em que j é o índice para as medidas ao longo do tempo e i é o índice para os indivíduos.

Para efeito de análise, os dados longitudinais devem ser dispostos de forma a conter uma indicação da unidade amostral à qual estão associados. Na Tabela 3.1 é apresentada, como ilustração, a disposição de um conjunto de dados em que as diferentes unidades amostrais são observadas em conjuntos de instantes de tempo possivelmente distintos. Em cada unidade amostral, tanto a resposta (Y) quanto as variáveis explicativas (X , K e W) são observadas em cada instante do tempo. As variáveis X e K representam características fixas da unidade amostral, sendo por isso covariáveis independentes do tempo ao passo que a variável W é uma covariável dependente do tempo, pois seus valores não são constantes ao longo do estudo (Singer et al., 2012).

Nos estudos longitudinais o i -ésimo indivíduo possui j observações, sendo estas correlacionadas. Logo, não se deve utilizar as técnicas para ajuste de modelos tradicionais pois não levam em consideração tal informação. Na Seção 3.1 são discutidas algumas abordagens utilizadas nas análises de dados longitudinais.

3.1 Abordagens para Análise de Dados Longitudinais

Nos últimos 30 anos tem-se observado muitos avanços na metodologia estatística para analisar dados longitudinais. A introdução de novas ferramentas computacionais e novos pacotes estatísticos para a análise de dados longitudinais fizeram com

Tabela 3.1: Ilustração da estrutura de um banco de dados longitudinais

Unidade amostral	Resposta	Tempo	Covariáveis		
			X	K	W
1	y_{11}	t_{11}	x_1	k_1	w_{11}
1	y_{12}	t_{12}	x_1	k_1	w_{12}
.	.	.	.	.	.
1	y_{1n_1}	t_{1n_1}	x_1	k_1	w_{1n_1}
2	y_{21}	t_{21}	x_2	k_2	w_{21}
2	y_{22}	t_{22}	x_2	k_2	w_{22}
.	.	.	.	.	.
2	y_{2n_2}	t_{2n_2}	x_2	k_2	w_{2n_2}
.	.	.	.	.	.
m	y_{m1}	t_{m1}	x_m	k_m	w_{m1}
m	y_{m2}	t_{m2}	x_m	k_m	w_{m2}
.	.	.	.	.	.
m	y_{mn_m}	t_{mn_m}	x_m	k_m	w_{mn_m}

que muitos métodos se tornassem acessíveis aos pesquisadores (Fitzmaurice et al., 2009).

Muitos dos primeiros métodos estatísticos utilizados na análise de dados longitudinais foram baseados na análise de variância (ANOVA), desenvolvida por Fisher nos anos 20, e aplicada inicialmente em experimentos agrícolas. Um dos métodos mais antigos proposto para a análise de dados longitudinais foi a ANOVA com um único efeito aleatório individual (cada indivíduo tem um determinado comportamento que difere de indivíduo para indivíduo), induzindo uma correlação positiva entre as medidas repetidas do mesmo indivíduo (Fitzmaurice et al., 2009). Este modelo foi utilizado por muito tempo na análise de dados longitudinais, muitas vezes referida como ANOVA com medidas repetidas univariada. Quando se tem um vetor de respostas multivariadas, deve-se utilizar a análise de variância multivariada (MANOVA), que, em 1950, teve uma variação desenvolvida por Box para análise de medidas repetidas, sendo conhecida como análise de perfis (Fitzmaurice et al., 2009).

Apesar da ANOVA ter sido uma das metodologias iniciais para análise de dados longitudinais, uma de suas limitações é a restrição da estrutura da matriz de covariância das medidas repetidas no mesmo indivíduo (Fitzmaurice et al., 2009). A matriz de covariância permutável (em que as correlações entre as medidas repetidas de um determinado indivíduo são iguais) não é adequada para dados longitudinais, pois espera-se que as correlações entre as medidas decaiam ao longo do tempo (Fitzmaurice et al., 2009). Além disso, a suposição sobre a variância constante ao longo

do tempo muitas vezes não é verificada, pois, em muitos estudos longitudinais, a variabilidade da resposta no início do estudo é visivelmente diferente da variabilidade apresentada no final do estudo, especialmente quando a primeira observação é no *baseline*, isto é, no início do estudo (Fitzmaurice et al., 2009).

Em contraste, a análise de medidas repetidas utilizando a MANOVA não faz suposições sobre a estrutura de covariância entre as respostas longitudinais no mesmo indivíduo, portanto, as correlações podem assumir qualquer padrão e a variabilidade pode mudar ao longo do tempo. No entanto, a MANOVA tem uma série de características que também limitam a sua utilidade como: a exigência de que as covariáveis sejam mensuradas no mesmo conjunto de pontos do tempo, que as covariáveis não sejam tempo-dependentes e que não hajam dados faltantes.

Em resumo, os métodos ANOVA fornecem uma base razoável para a análise longitudinal quando o desenho do estudo é muito simples. Porém, em muitos estudos longitudinais existe uma variação considerável entre os indivíduos, tanto no número de mensurações quanto no tempo em que eles foram medidos, tornando assim os dados bastante complexos, sendo, inviável a utilização destes métodos para a análise. Desta forma, foram desenvolvidas técnicas que pudessem lidar com os problemas mais comuns encontrados neste tipo de dados, como dados faltantes, dados desbalanceados e dados desigualmente espaçados, bem como variáveis tempo-dependentes ou tempo-independentes, ou ainda, quando as respostas são discretas, em vez de contínuas (Fitzmaurice et al., 2009).

Duas abordagens comumente utilizadas para analisar dados longitudinais são os modelos condicionais e os marginais. Em um modelo condicional, também conhecido como modelo misto, os coeficientes têm interpretações específicas no conglomerado (*cluster*, em inglês). O modelo multinível, considerado um modelo de regressão linear hierárquico (Little et al., 2000), está condicionado à efeitos aleatórios que descrevem um indivíduo específico, portanto pode ser considerado um modelo condicional. Este modelo pode apresentar dois ou mais níveis. Para ilustrar essa ideia, considere como exemplo, várias escolas (Nível 3) e dentro dessas escolas várias turmas (Nível 2), que, por sua vez, apresentam diferentes alunos (Nível 1) e o principal interesse é avaliar o coeficiente de rendimento (CR) em matemática dos alunos. Desta forma, o CR do aluno pode está sendo influenciado pela turma que ele se encontra e pela escola que ele estuda (efeitos aleatórios).

Já em um modelo marginal a variável resposta é modelada separadamente da correlação entre as medidas de cada unidade amostral (denotada por correlação

intra-unidade ou intra-indivíduo). Nesta abordagem, modela-se o valor esperado marginal, $E(Y_{ij})$ $i = 1, \dots, m$ e $j = 1, \dots, n_i$ em que o índice i denota os indivíduos e j as observações, como uma função das variáveis explicativas (Diggle et al., 2002).

O modelo multinível é provavelmente o método mais frequentemente adotado para analisar dados longitudinais. A sua utilidade foi destaque na década de 1980 em um artigo escrito por Laird & Ware (1982). Uma aplicação natural deste modelo é em dados longitudinais em que os efeitos aleatórios variam entre os indivíduos e induzem uma dependência entre as medidas repetidas dentro de um mesmo indivíduo após serem condicionadas as covariáveis observadas (Fitzmaurize et al., 2009).

A utilização do modelo multinível foi difundida por bioestatísticos que trabalhavam no Instituto Nacional de Saúde dos EUA (NIH, em inglês). Eles propuseram um método para análise de dados com medidas repetidas, em que, inicialmente, os coeficientes de regressão especificados para os indivíduos foram estimados por meio de uma regressão por mínimos quadrados ordinários (MQO) e, em uma segunda etapa, os coeficientes de regressão estimados foram então analisados utilizando métodos paramétricos (ou não paramétricos). Curiosamente, este método para análise de dados com medidas repetidas tornou-se conhecido como o “método NIH”. Rowell & Walters, em 1976, apresentaram uma abordagem semelhante para aplicações nas ciências agrícolas (Fitzmaurize et al., 2009).

No início de 1980, Laird & Ware (1982) baseando-se em uma classe geral de modelos mistos apresentada anteriormente por Harville (1976; 1977), propuseram os modelos lineares de efeitos mistos para dados longitudinais. O modelo linear de efeitos mistos é definido através da relação

$$Y_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \mathbf{b}_i + \epsilon_i, \quad i = 1, \dots, m, \quad (3.1)$$

em que $\mathbf{X}_i$ é uma matriz $(n_i \times p)$ de desenho relacionada aos efeitos fixos, $\boldsymbol{\beta}$ é um vetor $(p \times 1)$ de parâmetros também conhecido como vetor de efeitos fixos, $\mathbf{Z}_i$ é uma matriz $(n_i \times q)$ de desenho relacionada aos efeitos aleatórios e $\mathbf{b}_i$ denota um vetor $(q \times 1)$ de efeitos aleatórios. $\mathbf{b}_i$ e ϵ_i são termos aleatórios, tal que $\mathbf{b}_i \sim \mathcal{N}_q(0, \mathbf{G})$ e $\epsilon_i \sim \mathcal{N}_{n_i}(0, \boldsymbol{\Sigma}_i)$, sendo $\mathbf{G}$ e $\boldsymbol{\Sigma}_i$ matrizes de covariância apropriadas.

Seja $\mathbf{Y} = (\mathbf{Y}_1^\top, \dots, \mathbf{Y}_m^\top)^\top$, $\mathbf{X} = (\mathbf{X}_1^\top, \dots, \mathbf{X}_m^\top)^\top$, $\mathbf{Z} = \bigoplus_{i=1}^m \mathbf{Z}_i$ ($\bigoplus$ representa a soma direta) e $\mathbf{b} = (\mathbf{b}_1^\top, \dots, \mathbf{b}_m^\top)^\top$ e $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1^\top, \dots, \boldsymbol{\epsilon}_m^\top)^\top$, é possível reescrever (3.1) como

$$\mathbf{Y} = \mathbf{X} \boldsymbol{\beta} + \mathbf{Z} \mathbf{b} + \boldsymbol{\epsilon}. \quad (3.2)$$

Desta forma,

$$\begin{bmatrix} \mathbf{b} \\ \boldsymbol{\epsilon} \end{bmatrix} \sim \mathcal{N}_{qm+n} \begin{bmatrix} 0_{qm} \\ 0_n \end{bmatrix}, \begin{bmatrix} \mathbf{D} & 0_{qm \times n} \\ 0_{n \times qm} & \mathbf{R} \end{bmatrix},$$

em que $n = \sum_{i=1}^m n_i$, $\mathbf{D} = \mathbf{I}_m \otimes \mathbf{G}$ ($\otimes$ denota o produto Kronecker e $\mathbf{I}_m$ representa a matriz identidade de ordem m) e $\mathbf{R} = \bigoplus_{i=1}^m \boldsymbol{\Sigma}_i$. Portanto, a matriz de covariância de $\mathbf{Y}$ é dada por

$$\mathbf{V} = (\mathbf{Z}\mathbf{D}\mathbf{Z}^\top + \mathbf{R}).$$

Vários métodos de estimação dos parâmetros do modelo (3.1) estão disponíveis na literatura, dentre eles vale ressaltar os métodos Bayesianos (Tountenburg, 1982; Maritz & Lwin, 1989 e Searle et al., 1992), o método de Máxima Verossimilhança (MV) e Máxima Verossimilhança Restrita (MVR) (Patterson & Thompson, 1971; Harville, 1977; Robinson, 1991; Searle et al., 1992 e Jiang, 1996) e o método de Mínimos Quadrados (MQ) (Searle et al., 1992; Draper & Smith, 1998 e Hoffman & Vieira, 1998).

Em contraste ao modelo de efeitos mistos, que considera a resposta média como função tanto das covariáveis quanto do vetor de efeitos aleatórios, tem-se o modelo marginal, em que a média da variável resposta depende apenas das covariáveis de interesse.

Os modelos marginais fornecem uma maneira simples para estender os Modelos Lineares Generalizados (MLG's) para dados longitudinais. Eles modelam diretamente a resposta média em cada instante do tempo, $E(Y_{ij}|X_{ij})$, utilizando uma função de ligação apropriada (Fitzmaurice et al., 2009). O modelo marginal possui as seguintes suposições:

1. A esperança marginal da resposta para o i -ésimo indivíduo com j medidas repetidas, $E(Y_{ij}|X_{ij}) = \mu_{ij}$, depende das variáveis explanatórias, X_{ij} , na forma $h(\mu_{ij}) = X_{ij}^\top \beta$, sendo h uma função de ligação adequada e duas vezes degenerada, $i = 1, \dots, m$ e $j = 1, \dots, n_i$.
2. A variância marginal depende da esperança marginal de acordo com $\text{Var}(Y_{ij}|X_{ij}) = \phi v(\mu_{ij})$, em que $v(\mu_{ij})$ é uma “função da variância” conhecida e ϕ é um parâmetro de escala que precisa ser estimado.
3. A associação (condicional intra-indivíduo) do vetor de respostas repetidas, y_{ij} e $y_{ij'}$, dadas as covariáveis, é considerada como uma função de um conjunto

adicional de parâmetros de associação, por exemplo α (e pode também depender de μ_{ij}), que indica o grau de correlação, por exemplo.

É possível notar que as duas primeiras suposições do modelo marginal representam as especificações de um MLG. A terceira suposição está vinculada à associação entre o vetor de respostas repetidas dentro do indivíduo, sendo assim a principal extensão dos MLG's para dados longitudinais. Portanto, os modelos marginais apresentam um MLG para as respostas longitudinais em cada ocasião e adicionam um modelo para a associação entre o vetor de respostas repetidas dentro de um mesmo indivíduo (Fitzmaurice et al., 2009).

Por volta de 1980 notáveis avanços na metodologia para a análise longitudinal de dados discretos foram desenvolvidos quando Liang & Zeger (1986) propuseram o uso de Equações de Estimação Generalizadas (EEG). Os autores reconheceram que era possível estimar os parâmetros de regressão sem fazer suposições sobre a distribuição da variável resposta. Isto é potencialmente vantajoso, pois não há especificação conveniente da distribuição multivariada conjunta de Y_i para modelos marginais quando as respostas são discretas (Fitzmaurice et al., 2009). Assim, a abordagem EEG requer apenas a especificação das médias marginais.

A seguir são apresentadas e discutidas essas duas abordagens no contexto do modelo linear para dados longitudinais.

3.1.1 Modelos Multiníveis

O modelo multinível tornou-se bastante popular na análise de dados longitudinais por causa da sua flexibilidade e ampla aplicabilidade. Além disso, muitos pacotes estatísticos disponíveis no mercado oferecem procedimentos para o seu ajuste. Eles assumem que as medidas de um único indivíduo compartilham um conjunto de variáveis observadas e efeitos aleatórios que são utilizados para gerar uma estrutura de associação entre as medidas repetidas (Verbeke et al., 2010).

Com dados normais repetidos, uma classe geral e bastante flexível de modelos paramétricos é obtida a partir da abordagem com efeitos aleatórios (Verbeke et al., 2010). Suponha que uma variável resposta contínua, denotada por Y_i , é observada repetidamente ao longo do tempo em um conjunto de indivíduos. Um modelo de regressão linear com intercepto e efeito linear do tempo parece plausível para descrever os dados de cada indivíduo separadamente. No entanto, diferentes indivíduos tendem a ter diferentes interceptos e diferentes inclinações. Desta forma, a j -ésima

observação ($j = 1, \dots, n_i$) do i -ésimo indivíduo ($i = 1, \dots, m$), Y_{ij} , mensurada no tempo t_{ij} , é dada por

$$Y_{ij} = \gamma_{0i} + \gamma_{1i}t_{ij} + \epsilon_{ij}, \quad (3.3)$$

que representa o primeiro nível do modelo multinível. Aqui $\boldsymbol{\gamma}_i = (\gamma_{0i}, \gamma_{1i})^\top$ é o vetor de parâmetros específicos por indivíduo e ϵ_{ij} é distribuído normalmente com média 0 e variância $\sigma_{\epsilon_{ij}}^2$ (Verbeke et al., 2010). Considere agora o segundo nível do modelo multinível dado por

$$\begin{aligned} \gamma_{0i} &= \beta_0 + b_{0i}, \\ \gamma_{1i} &= \beta_1 + b_{1i}, \end{aligned} \quad (3.4)$$

em que $\mathbf{b}_i = (b_{0i}, b_{1i})^\top$ denotam os efeitos aleatórios e assumem médias iguais a zero e matriz de covariância $\mathbf{G} = \begin{bmatrix} \tau_0 & \tau_{01} \\ \tau_{01} & \tau_1 \end{bmatrix}$. Assim, substituindo (3.4) em (3.3), tem-se que

$$Y_{ij} = (\beta_0 + b_{0i}) + (\beta_1 + b_{1i})t_{ij} + \epsilon_{ij}. \quad (3.5)$$

O modelo (3.5) é um caso especial do modelo misto linear geral (3.1). É possível notar a relação entre os modelos (3.5) e (3.1), uma vez que $\mathbf{X}_i^\top = \begin{bmatrix} 1 & t_{ij} \end{bmatrix}$, $\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$, $\mathbf{Z}_i^\top = \begin{bmatrix} 1 & t_{ij} \end{bmatrix}$ e $\mathbf{b}_i = \begin{bmatrix} b_{0i} \\ b_{1i} \end{bmatrix}$.

Nesta classe de modelos existe o interesse na estimação dos efeitos fixos, na predição dos efeitos aleatórios e na estimação dos componentes de variância. A obtenção das equações de modelos mistos é feita pela minimização da soma de quadrados dos resíduos ou pela maximização da função densidade de probabilidade conjunta de $\mathbf{Y}$ e $\mathbf{b}$ (Freitas, 2007).

A função densidade de probabilidade de $\mathbf{Y}$ e a função de probabilidade conjunta de $\mathbf{Y}$ e $\mathbf{b}$, considerando o modelo (3.2) são dadas por

$$f_Y(\mathbf{y}) = \frac{1}{2\pi^{n_i/2}(|\mathbf{ZDZ}^\top + \mathbf{R}|)^{1/2}} \exp \frac{-1}{2}[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})^\top (\mathbf{ZDZ}^\top + \mathbf{R})^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})]$$

e

$$\begin{aligned} f(\mathbf{y}, \mathbf{b}) &= f(\mathbf{y}|\mathbf{b}).f(\mathbf{b}) \\ &= \frac{1}{2\pi^{n_i/2}(|\mathbf{R}|)^{1/2}} \exp \left\{ -\frac{1}{2}[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Zb})^\top \mathbf{R}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Zb})] \right\} \times \\ &\quad \frac{1}{2\pi^{q/2}(|\mathbf{R}|)^{1/2}} \exp \left\{ -\frac{1}{2}[\mathbf{b}^\top \mathbf{D}^{-1}\mathbf{b}] \right\}, \end{aligned} \quad (3.6)$$

em que n_i e q são os postos das matrizes $\mathbf{D}$ e $\mathbf{R}$, respectivamente.

Considerando (3.6), o logaritmo da função de verossimilhança é dada por

$$\begin{aligned} \ell(\boldsymbol{\beta}, \mathbf{b}; \mathbf{Y}) &= \frac{(m+n)}{2} \log(2\pi) - \frac{1}{2}(\log \mathbf{R} + \log \mathbf{D}) - \frac{1}{2} \left(\mathbf{Y}^\top \mathbf{R}^{-1} \mathbf{Y} - 2\mathbf{Y}^\top \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta} - 2\mathbf{Y}^\top \mathbf{R} \mathbf{Z} \mathbf{b} \right. \\ &\quad \left. + 2\boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Z} \mathbf{b} + \boldsymbol{\beta}^\top \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X} \boldsymbol{\beta} + \mathbf{b}^\top \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} \mathbf{b} + \mathbf{b}^\top \mathbf{D}^{-1} \mathbf{b} \right). \end{aligned} \quad (3.7)$$

Derivando a equação (3.7) em relação a $\boldsymbol{\beta}$ e $\mathbf{b}$, tomando-se as derivadas parciais iguais a 0, obtém-se:

$$\begin{bmatrix} \frac{\partial \ell(\boldsymbol{\beta}, \mathbf{b}; \mathbf{Y})}{\partial \boldsymbol{\beta}} \\ \frac{\partial \ell(\boldsymbol{\beta}, \mathbf{b}; \mathbf{Y})}{\partial \mathbf{b}} \end{bmatrix} = \begin{bmatrix} -\mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Y} + \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Z} \hat{\mathbf{b}} \\ -\mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Y} + \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{X} \hat{\boldsymbol{\beta}} + \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} \hat{\mathbf{b}} + \mathbf{D}^{-1} \mathbf{b} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Desenvolvendo-se as expressões e agrupando-se de modo conveniente, tem-se

$$\begin{bmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Y} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Y} \end{bmatrix}.$$

As equações apresentadas acima são conhecidas como equações de Henderson. Através delas são obtidas as soluções para os efeitos fixos $\hat{\boldsymbol{\beta}}$ e a predição para os efeitos aleatórios $\hat{\mathbf{b}}$. As soluções das equações são dadas por:

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Z} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{X} & \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \mathbf{D}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}^\top \mathbf{R}^{-1} \mathbf{Y} \\ \mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Y} \end{bmatrix}. \quad (3.8)$$

Henderson et al. (1959) provaram que $\hat{\boldsymbol{\beta}}$, dado pela equação (3.8), coincide com a solução obtida pelo método de mínimos quadrados generalizados, cuja expressão é

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{Y}. \quad (3.9)$$

A predição de $\hat{\mathbf{b}}$, segundo Henderson (1975), é dada por

$$\hat{\mathbf{b}} = \mathbf{D} \mathbf{Z} \mathbf{V}^{-1} (\mathbf{y} - \mathbf{X} \hat{\boldsymbol{\beta}}), \quad (3.10)$$

em que $\mathbf{V}^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1} \mathbf{Z} (\mathbf{Z}^\top \mathbf{R}^{-1} \mathbf{Z} + \mathbf{D}^{-1})^{-1} \mathbf{Z}^\top \mathbf{R}^{-1}$.

Segundo Searle (1971), há desvantagem em não utilizar esta opção, pois o cálculo de $\mathbf{V}^{-1}$ pode se tornar difícil, uma vez que a dimensão de $\mathbf{V}$ é igual ao número de observações que, geralmente, é bem alto. Sendo as matrizes $\mathbf{D}$ e $\mathbf{R}$ conhecidas, $\hat{\boldsymbol{\beta}}$ é o melhor estimador linear não viesado (BLUE) e $\hat{\mathbf{b}}$ é o melhor preditor linear não viesado (BLUP) (Littell et al., 2006).

Para prever e estimar os parâmetros da equação (3.8) é necessário conhecer a estrutura de $\mathbf{V}$, no entanto, a mesma depende das estruturas de $\mathbf{R}$ e $\mathbf{D}$ que, em geral, são desconhecidas. Portanto, deve-se estimar os parâmetros de $\mathbf{R}$ e $\mathbf{D}$.

A estimação de $\mathbf{R}$ e $\mathbf{D}$ pode ser realizada por vários métodos, dentre os quais destacam-se os métodos de Máxima Verossimilhança (MV) e Máxima Verossimilhança Restrita (MVR) (Patterson & Thompson; 1971; Harville, 1977; Robinson, 1991; Searle et al., 1992 e Jiang, 1996).

O método MV consiste em obter a função de verossimilhança das observações e maximizá-la em função dos parâmetros de $\mathbf{R}$, $\mathbf{D}$ e dos efeitos fixos, assumindo que $\mathbf{Y}$ do modelo (3.2) segue distribuição normal multivariada. Já os estimadores de MVR são obtidos maximizando-se a parte da função de verossimilhança que é invariante ao parâmetro de locação, ou seja, em termos do modelo misto, é invariante para $\mathbf{X}\boldsymbol{\beta}$. É preferível utilizar a função MVR, pois ela mantém as propriedades da função MV, é não viesada e permite a imposição de não negatividade. Maiores detalhes em Pinheiro & Bates (2000).

Em ambos os métodos são geradas equações cujas soluções fornecem estimativas para os parâmetros de $\mathbf{D}$ e $\mathbf{R}$. Essas estimativas são obtidas por processos iterativos (Newton-Raphson, por exemplo), uma vez que não é possível a obtenção de expressões analíticas para os parâmetros.

Substituindo as matrizes de covariâncias $\mathbf{R}$ e $\mathbf{D}$ pelas respectivas matrizes de covariâncias estimadas $\hat{\mathbf{R}}$ e $\hat{\mathbf{D}}$, o sistema de equações dado em (3.8) passa a ser

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^\top \hat{\mathbf{R}}^{-1} \mathbf{X} & \mathbf{X}^\top \hat{\mathbf{R}}^{-1} \mathbf{Z} \\ \mathbf{Z}^\top \hat{\mathbf{R}}^{-1} \mathbf{X} & \mathbf{Z}^\top \hat{\mathbf{R}}^{-1} \mathbf{Z} + \hat{\mathbf{D}}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}^\top \hat{\mathbf{R}}^{-1} \mathbf{Y} \\ \mathbf{Z}^\top \hat{\mathbf{R}}^{-1} \mathbf{Y} \end{bmatrix}. \quad (3.11)$$

O estimador de $\hat{\boldsymbol{\beta}}$ e preditor de $\mathbf{b}$ são dados, respectivamente, por

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{X}^\top \hat{\mathbf{V}}^{-1} \mathbf{Y} \quad (3.12)$$

e

$$\hat{\mathbf{b}} = \hat{\mathbf{D}} \mathbf{Z} \hat{\mathbf{V}}^{-1} (\mathbf{Y} - \mathbf{X} \hat{\boldsymbol{\beta}}). \quad (3.13)$$

Os estimadores $\hat{\boldsymbol{\beta}}$ e $\hat{\mathbf{b}}$ são denotados BLUE empírico (EBLUE) e BLUP empírico (EBLUP). Portanto, um problema relacionado com a estimação dos parâmetros das matrizes de covariâncias consiste no fato de optar pelo método mais conveniente, bem como a seleção das estruturas mais adequadas.

Vale ressaltar que a escolha da estrutura de covariância atinge diretamente as estimativas de efeitos fixos e erros padrão, diagnósticos e inferências. Para maiores detalhes sobre a escolha da estrutura da matriz de covariâncias ver Verbeke & Molenberghs (1997), Pinheiro & Bates (2000) e Rocha (2004).

3.1.2 Equações de Estimação Generalizadas (EEG)

As Equações de Estimação Generalizadas (EEG) foram propostas por Liang & Zeger (1986) exigindo apenas a especificação correta das distribuições marginais univariadas desde que se esteja disposto a adotar uma matriz de correlação de “trabalho”, que indica o grau de dependência entre as medidas repetidas.

Uma característica básica da abordagem EEG é que a distribuição do vetor de resposta dos indivíduos, $\mathbf{Y}_i = (Y_{i1}, \dots, Y_{in_i})^\top$ não precisa ser especificada. Uma facilidade relacionada aos modelos na abordagem EEG é que a estrutura de covariância é tratada como um parâmetro de incômodo ou ruído, levando o foco da análise aos parâmetros de regressão de X em Y . Neste contexto, as EEG produzem estimativas consistentes e assintoticamente normais para os coeficientes de regressão $\boldsymbol{\beta}$, mesmo sob má especificação da estrutura de covariância dos dados longitudinais (Hedeker & Gibbons, 2006).

Os modelos utilizando a abordagem EEG podem ser considerados como uma extensão dos MLG's para dados correlacionados, portanto, sua especificação é baseada nestes modelos. Assim, primeiramente, o preditor linear é especificado como

$$\eta_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\beta},$$

em que $\mathbf{X}_{ij}^\top = (X_{i1}, \dots, X_{ip})$ é o vetor p -dimensional de covariáveis para o i -ésimo indivíduo no j -ésimo tempo. Depois uma função de ligação que relaciona a média marginal com o preditor linear, dada por

$$h(\mu_{ij}) = \eta_{ij}$$

é escolhida. Assim como nos MLG's, as funções de ligação comumente usadas são a identidade, o logit e o log para dados contínuos, binários e de contagem, respectivamente. A variância é então descrita como uma função da média, denotada por

$$\text{Var}(Y_{ij}) = \phi v(\mu_{ij}),$$

em que $v(\mu_{ij})$ é conhecida como “função de variância” e ϕ é um parâmetro de dispersão que pode ser conhecido ou estimado dependendo do caso (Hedeker & Gibbons, 2006).

A especificação adicional para o modelo na abordagem EEG está relacionada à estrutura de correlação de “trabalho” das medidas repetidas. Esta matriz de correlação de “trabalho”, definida por R , tem dimensão $(n \times n)$, pois assume-se que há um número fixo de pontos no tempo, n , em que os indivíduos são mensurados. Se o indivíduo não foi mensurado em todos os n pontos do tempo, cada matriz de correlação individual, R_i , é de tamanho $n_i \times n_i$ com as linhas e colunas removidas apropriadamente se $n_i < n$. Supõe-se que a matriz de correlação de “trabalho”, R , bem como R_i , depende de um vetor de parâmetros de associação denotado por α . Assume-se que estes parâmetros α são os mesmos para todos os indivíduos e representam a dependência média entre as observações repetidas e os indivíduos (Hedeker & Gibbons, 2006).

Geralmente é recomendado que a escolha de R seja consistente com as correlações observadas. No entanto, as EEG produzem estimativas consistentes dos coeficientes de regressão e dos erros-padrão mesmo sob uma má especificação da estrutura de correlação (Hedeker & Gibbons, 2006). As matrizes de correlação usualmente consideradas são a independente, a permutável, a autorregressiva de ordem 1 (AR1) e a não estruturada.

A estimação utilizando EEG é dada como segue. Considerando-se dados balanceados para simplificar a notação, seja $\mathbf{A}_i$ uma matriz diagonal $n \times n$ com elementos da diagonal dados por $\text{Var}(Y_{ij})$, em que $i = 1, \dots, m$ e $j = 1, \dots, n$. Assim,

$$\mathbf{A}_i = \begin{bmatrix} \text{Var}(Y_{i1}) & 0 & \dots & 0 \\ 0 & \text{Var}(Y_{i2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \text{Var}(Y_{in}) \end{bmatrix}.$$

Seja $R_i(\alpha)$ a matriz de correlação de “trabalho” $n \times n$ para o i -ésimo indivíduo, em que α é o vetor de parâmetros de associação. Então, a matriz de covariância para $\mathbf{Y}_i$ é dada por

$$V(\alpha) = \phi \mathbf{A}_i^{1/2} R_i(\alpha) \mathbf{A}_i^{1/2}, \quad i = 1, \dots, m. \quad (3.14)$$

Para variáveis respostas distribuídas normalmente com variância homogênea ao longo do tempo, tem-se que $V(\alpha) = \phi R_i(\alpha)$. Park (1993) estendeu (3.14) para o caso que as variâncias fossem heterogêneas ao longo do tempo, isto é, permitindo que o parâmetro de escala ϕ_j variasse ao longo do tempo, com $j = 1, \dots, n$ (Hedeker & Gibbons, 2006).

O estimador EEG de β é obtido através da solução da equação

$$\sum_{i=1}^m \mathbf{D}_i^\top [V(\hat{\boldsymbol{\alpha}})]^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_i) = 0, \quad (3.15)$$

em que $\hat{\boldsymbol{\alpha}}$ é um estimador consistente de $\boldsymbol{\alpha}$ e $\mathbf{D}_i = \partial \boldsymbol{\mu}_i / \partial \boldsymbol{\beta}$ (Hedeker & Gibbons, 2006). No caso normal tem-se $\boldsymbol{\mu}_i = \mathbf{X}_i \boldsymbol{\beta}$, $\mathbf{D}_i = \mathbf{X}_i$ e $V(\hat{\boldsymbol{\alpha}}) = R_i(\hat{\boldsymbol{\alpha}})$. Então, a equação de estimação é da forma $\sum_{i=1}^m \mathbf{X}_i^\top [\mathbf{R}_i(\hat{\boldsymbol{\alpha}})]^{-1} (y_i - \mathbf{X}_i \boldsymbol{\beta}) = 0$ e, portanto, $\boldsymbol{\beta}$ é estimado através de

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}_i^\top \mathbf{R}_i(\hat{\boldsymbol{\alpha}})^{-1} \mathbf{X}_i \right)^{-1} \left(\mathbf{X}_i^\top \mathbf{R}_i(\hat{\boldsymbol{\alpha}})^{-1} \mathbf{y}_i \right). \quad (3.16)$$

Note que (3.16) é um estimador de mínimos quadrados ponderados (MQP), em que a matriz de pesos é dada por $[R_i(\hat{\boldsymbol{\alpha}})]^{-1}$. Uma limitação desta abordagem é que em um estimador obtido por MQP tipicamente se conhece a matriz de pesos. No entanto esta matriz depende do parâmetro $\boldsymbol{\alpha}$ que deve ser estimado. Assim, a solução de (3.15) pode ser encontrada através do estimador de mínimos quadrados ponderados iterativo (MQPI), em que as estimativas de $\boldsymbol{\alpha}$ obtidas iterativamente são usadas para produzir novas estimativas de $\boldsymbol{\beta}$, através de um processo alternado até a convergência. Como (3.15) depende apenas da média e da variância de Y , estes estimadores são denominados de estimadores de quase-verossimilhança.

A abordagem EEG utiliza a iteração entre a solução da quasi-verossimilhança para a estimação de $\boldsymbol{\beta}$ e um método robusto para estimar $\boldsymbol{\alpha}$ como uma função de $\boldsymbol{\beta}$. Essencialmente, ela envolve os seguintes passos que são repetidos até a convergência:

1. Assumindo independência entre as observações, obtêm-se a estimativa inicial para $\boldsymbol{\beta}$;
2. Dadas as estimativas iniciais, pelo método dos momentos, para $\boldsymbol{\alpha}$ e ϕ , estima-se $\text{Var}(Y_i)$, e uma estimativa atualizada para $\boldsymbol{\beta}$ é obtida como solução da equação de estimação em (3.15);
3. Dada a estimativa para $\boldsymbol{\beta}$, obtida no passo 2, encontram-se estimativas atualizadas para $\boldsymbol{\alpha}$ e ϕ através dos resíduos. Assim,

$$\hat{\phi} = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} \hat{r}_{ij}^2}{\sum_{i=1}^m n_i - p}, \quad (3.17)$$

em que $\hat{r}_{ij}^2 = \frac{(y_{ij} - \hat{\mu}_{ij})^2}{\sqrt{[V(\hat{\boldsymbol{\alpha}})]_{ij}}}$, $\hat{\mu}_{ij} = h^{-1}(X_{ij}^\top \hat{\boldsymbol{\beta}})$ é a média do i -ésimo indivíduo e $[V(\hat{\boldsymbol{\alpha}})]_{ij}$ representa a matriz de variância.

As estimativas para α dependerão da estrutura escolhida para a matriz $R_i(\alpha)$. A seguir, estas matrizes são apresentadas assumindo-se quatro medidas repetidas para um i -ésimo indivíduo.

- Independente: Assume-se que as correlações para as diferentes medidas de um determinado indivíduo são zero, conhecida também por $R_i(\alpha) = I$, uma matriz identidade ($n \times n$). Esta forma não é adequada para estudos longitudinais, uma vez que estes dados são altamente correlacionados.

$$R_i(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- Permutável: Quando se assume que as correlações entre todas as medidas repetidas dos indivíduos são iguais. Esta estrutura é dada por

$$R_i(\alpha) = \begin{bmatrix} 1 & \alpha & \alpha & \alpha \\ \alpha & 1 & \alpha & \alpha \\ \alpha & \alpha & 1 & \alpha \\ \alpha & \alpha & \alpha & 1 \end{bmatrix}.$$

O estimador $\hat{\alpha}$ é único para todos os tempos, sendo dado por

$$\hat{\alpha} = \frac{\sum_{i=1}^m \sum_{j \neq j'} \hat{r}_{ij}^2 \hat{r}_{ij'}^2}{\hat{\phi}(N^* - p)},$$

em que $N^* = \sum_{i=1}^m \frac{n_i(n_i-1)}{2}$.

- Autoregressiva de ordem 1 (AR1): Assume-se que correlações adjacentes são maiores em magnitude. A correlação intra-indivíduos ao longo do tempo é uma função exponencial do comprimento. No caso de dados longitudinais esta estrutura é a mais parcimoniosa, uma vez que depende apenas de um termo e ainda permite que as correlações diminuam ao longo do tempo.

$$R_i(\alpha) = \begin{bmatrix} 1 & \alpha & \alpha^2 & \alpha^3 \\ \alpha & 1 & \alpha & \alpha^2 \\ \alpha^2 & \alpha & 1 & \alpha \\ \alpha^3 & \alpha^2 & \alpha & 1 \end{bmatrix},$$

em que $\alpha \in (0, 1)$.

O estimador $\hat{\alpha}$ é dado por

$$\hat{\alpha} = \frac{\sum_{i=1}^m \sum_{j \leq n_i-1} \hat{r}_{ij}^2 \hat{r}_{i,j+1}^2}{\hat{\phi}(K^* - p)},$$

em que $K^* = \sum_{i=1}^m (n_i - 1)$.

- Indeterminada ou não estruturada: Todas $n(n-1)/2$ correlações de R_i são estimadas. Esta forma é mais eficiente e útil quando existem relativamente poucos pontos no tempo. Quando há muitas observações, a estimação das $n(n-1)/2$ correlações é bastante complicada. Além disso, a presença de dados faltantes dificulta a estimação de R .

$$R_i(\boldsymbol{\alpha}) = \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & \alpha_3 \\ \alpha_1 & 1 & \alpha_4 & \alpha_5 \\ \alpha_2 & \alpha_4 & 1 & \alpha_6 \\ \alpha_3 & \alpha_5 & \alpha_6 & 1 \end{bmatrix}.$$

O estimador $\hat{\boldsymbol{\alpha}}$ é dado para cada tempo, ou seja

$$\hat{\boldsymbol{\alpha}} = \frac{\sum_{i=1}^m \hat{r}_{ij}^2 \hat{r}_{ij'}^2}{\hat{\phi}(m - p)}.$$

Em relação aos critérios para a seleção da estrutura de correlação de “trabalho”, recomenda-se que: se o número de medidas no indivíduo é pequeno e os dados são balanceados e completos (todos os indivíduos com o mesmo número de medidas), deve-se utilizar a matriz de correlação não estruturada; se o número de medidas no indivíduo é grande, recomenda-se a utilização da matriz de correlação permutável; se as observações são coletadas ao longo do tempo, é necessário utilizar uma estrutura que considere a correlação em função do tempo, como a autorregressiva. Neste trabalho, assume-se que as covariáveis e a matriz de covariâncias já foram especificadas previamente. Para maiores detalhes referentes a métodos de especificação da matriz de covariâncias ver Rocha (2004).

Quando os dados se enquadram em mais de uma das situações citadas anteriormente, gerando dúvida sobre qual estrutura utilizar, pode-se optar por um critério estatístico para selecionar a estrutura de correlação mais adequada. O critério de seleção utilizado é o critério de quase-verossimilhança sob o modelo de independência (QIC, em inglês) que será apresentado na Seção 3.3.

3.2 Diagnóstico

Como discutido no Capítulo 2, Seção 2.4, a análise de diagnóstico é utilizada para avaliar a validade das suposições dos modelos. Nesta seção serão apresentadas algumas técnicas utilizadas na análise de diagnóstico no contexto de dados com medidas repetidas.

3.2.1 Modelo Linear Misto

Cox & Snell (1968) definiram uma forma geral para calcular os resíduos para modelos com um única fonte de variação. No modelo 3.1 existe mais de uma fonte de variação, portanto mais de um tipo de resíduo. Desta forma a análise de diagnóstico apresentada na Seção 2.4 não pode ser utilizada.

É possível encontrar na literatura uma vasta discussão sobre análise de diagnóstico para modelos com mais de uma fonte de variação. Hilden-Minton (1995) e Verbeke & Lesaffre (1996a), por exemplo, estendem algumas ideias para definir três tipos de resíduos que acomodam a fonte de variabilidade extra presente nos modelos lineares mistos. Nobre & Singer (2007) revisam algumas técnicas de análise de resíduos que têm sido utilizadas no contexto de modelos lineares mistos e propõem uma padronização para o resíduo condicional usado para identificar observações discrepantes e conglomerados (*clusters*, em inglês). Nobre & Singer (2011) apresentam a análise de pontos de alavanca para modelos lineares mistos.

Considere o modelo linear misto 3.2. Condicionado as matrizes de covariância $\mathbf{D}$ e $\mathbf{R}$, o melhor estimador linear não viesado (BLUE, em inglês) de $\boldsymbol{\beta}$ e o melhor preditor linear não-viesado (BLUP, em inglês) de $\mathbf{b}$ são dados, respectivamente, por

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{V}^{-1} \mathbf{X})^\top \mathbf{X}^\top \mathbf{V}^{-1} \mathbf{Y} = \mathbf{T} \mathbf{Y} \quad \text{e} \quad (3.18)$$

$$\hat{\mathbf{b}} = \mathbf{D} \mathbf{Z}^\top \mathbf{Q} \mathbf{Y}, \quad (3.19)$$

em que $\mathbf{T} = (\mathbf{X}^\top \mathbf{M} \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{M}$, $\mathbf{Q} = \mathbf{M}(\mathbf{I} - \mathbf{X} \mathbf{T})$, em que $\mathbf{M} = \mathbf{V}^{-1} = (\mathbf{Z} \mathbf{D} \mathbf{Z}^\top + \mathbf{R})^{-1}$.

Na prática, as matrizes $\mathbf{D}$ e $\mathbf{V}$ em (3.18) e (3.19) devem ser convenientemente estimadas, chamadas de BLUE empírico (EBLUE) e BLUP empírico (EBLUP). Jiang (1998, 1999) discute propriedades para o EBLUE e EBLUP.

Os três tipos de resíduos úteis para avaliar alguma suposição desse modelo são

1. Resíduos marginais: $\hat{\xi} = \mathbf{Y} - \mathbf{X}\hat{\beta} = \mathbf{M}^{-1}\mathbf{QY}$, que é o preditor dos erros marginais, $\xi = \mathbf{Y} - \mathbf{E}(\mathbf{Y}) = \mathbf{Y} - \mathbf{X}\beta = \mathbf{Zb} - \epsilon$.
2. Resíduos condicionais: $\hat{\epsilon} = \mathbf{Y} - \mathbf{X}\hat{\beta} - \mathbf{Z}\hat{\mathbf{b}} = \Sigma\mathbf{QY}$, que é o preditor dos erros condicionais, $\epsilon = \mathbf{Y} - \mathbf{E}[\mathbf{Y}|\mathbf{b}] = \mathbf{Y} - \mathbf{X}\beta - \mathbf{Zb}$.
3. O BLUP: $\mathbf{Zb}$, o preditor dos efeitos aleatórios, $\mathbf{Zb} = \mathbf{E}[\mathbf{Y}|\mathbf{b}] - \mathbf{E}[\mathbf{Y}]$.

De acordo com Hilden-Minton (1995), um resíduo é considerado “puro” para um tipo específico de erro se ele depende apenas dos componentes fixos e do erro que é considerado como preditor. Os resíduos que dependem de outros tipos de erro são chamados resíduos com “confundimento”. Desta forma, sob (3.2)-(3.19), tem-se que

$$\hat{\xi} = [\mathbf{I} - \mathbf{X}(\mathbf{X}^\top \mathbf{M}\mathbf{X})^\top \mathbf{X}^\top \mathbf{M}]\xi, \quad (3.20)$$

$$\hat{\epsilon} = \mathbf{RQ}\epsilon + \mathbf{RQZb}, \quad (3.21)$$

$$\mathbf{Z}\hat{\mathbf{b}} = \mathbf{ZDZ}^\top \mathbf{ZDZ} + \mathbf{ZDZ}^\top \mathbf{Q}\epsilon, \quad (3.22)$$

os resíduos $\hat{\epsilon}$ e $\mathbf{Z}\hat{\mathbf{b}}$ são confundimentos com $\mathbf{b}$ e ϵ , respectivamente. Isto implica, por exemplo, que $\hat{\epsilon}$ não pode ser adequado para checar a normalidade de ϵ quando $\mathbf{b}$ é não-normal (Nobre & Singer, 2007). A Tabela 3.2 sumariza a utilização de cada tipo de resíduo.

Tabela 3.2: Uso dos resíduos na análise de diagnóstico para modelos lineares mistos. Tipo de resíduo: 1-Marginal, 2- Condicional e 3-EBLUP.

Diagnóstico para	Tipo	Gráfico
Linearidade dos efeitos ($\mathbf{E}(\mathbf{Y}) = \mathbf{X}\beta$)	1	$\hat{\xi}_j$ vs. variáveis aleatórias
Matriz de covariância dentro dos indivíduos ($\mathbf{V}_i$)	1	$\ I_{n_i} - \mathcal{R}_i \mathcal{R}_i^\top\ ^2$ vs. índices dos indivíduos
Presença de observações discrepantes	2	$\hat{\epsilon}_j^*$ vs. índices das observações
Homoscedasticidade dos erros condicionais	2	$\hat{\epsilon}_j^*$ vs. valores ajustados
Normalidade dos erros condicionais (ϵ_i)	2	QQ para o último resíduo de confundimento
Presença de indivíduos discrepantes	3	$\hat{\omega}_i$ vs. índice dos indivíduos
Estrutura de covariância dos efeitos aleatórios ($\mathbf{D}$)	3	$ d_{max} $ vs. índice dos indivíduos
Normalidade dos efeitos aleatórios ($\mathbf{b}_i$)	3	QQ ponderado para $\hat{\mathbf{b}}_i$

em que $\mathcal{R}_i = \hat{\mathbf{V}}_i^{-1/2} \hat{\xi}_i$, $\hat{\epsilon}_j^* = \frac{\hat{\epsilon}_j}{\hat{\sigma} \sqrt{\hat{p}_{jj}}}$ ($\hat{p}_{jj}$ é uma estimativa do j -ésimo elemento da diagonal principal de $\Sigma\mathbf{Q}\Sigma$) e $\hat{\omega}_i = \hat{b}_i^\top \widehat{\text{Var}}[\hat{b}_i - b_i] \hat{b}_i$.

3.2.2 Equações de Estimação Generalizadas (EEG)

Para os Modelos Lineares Generalizados (MLG's), as técnicas de diagnóstico já estão bastante difundidas. Há na literatura algumas técnicas de diagnóstico para

modelos com estrutura de dependência entre as observações da mesma unidade experimental. Por exemplo, Tan et al. (1997) propõem técnicas de diagnósticos para modelos de regressão logística, para dados com estrutura de dependência. Preisser & Qaqish (1996) apresentam medidas para detectar observações e/ou unidades experimentais influentes em MLG's com medidas repetidas. Chang (2000) apresenta um teste não-paramétrico para avaliar a aleatoriedade dos resíduos, considerando a abordagem EEG. Pan (2001) apresenta medidas para escolher a melhor matriz de correlação de trabalho e para selecionar covariáveis, baseadas no critério de informação de Akaike (AIC). Venezuela (2003) desenvolve uma generalização das técnicas de diagnóstico comumente utilizadas para os MLG's considerando dados com medidas repetidas.

Considere o estimador de β dado por

$$\hat{\beta}_G = \left[\sum_{i=1}^m \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \mathbf{X}_i \right]^{-1} \left[\sum_{i=1}^m \mathbf{X}_i^\top \widehat{\mathbf{W}}_i \mathbf{z}_i \right] \quad (3.23)$$

em que $\widehat{\mathbf{W}}_i = \widehat{\Delta}_i \widehat{\mathbf{A}}_i \widehat{\Omega}_i^\top \widehat{\mathbf{A}}_i \widehat{\Delta}_i$, $z_i = \hat{\eta}_i + (\widehat{\Delta}_i \widehat{\mathbf{A}}_i)^{-1}(\mathbf{y}_i - \hat{\mu}_i)$, $\Delta_i = \text{diag} \partial \theta_{ij} / \partial \eta_{ij}$, com dimensão $n \times n$ e $\mathbf{A}_i = \text{diag} \partial \mu_{ij} / \partial \theta_{ij} = \text{diag} b''(\theta_{ij})$ com dimensão $n \times n$.

Para simplificar a expressão acima, considere $\widehat{\mathbf{W}} = \text{diag}(\widehat{\mathbf{W}}_1, \dots, \widehat{\mathbf{W}}_m)$, com dimensão $(N \times N)$, sendo $N = mn$. Assim a equação (3.23) é reescrita por

$$\hat{\beta}_G = (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X})^{-1} (\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{z}), \quad (3.24)$$

em que $\mathbf{X} = (\mathbf{X}_1^\top, \dots, \mathbf{X}_m^\top)^\top$ e $\mathbf{z} = (z_1^\top, \dots, z_m^\top)^\top$, com dimensões $(N \times p)$ e $(N \times p)$, respectivamente.

O resíduo ordinário, que é a diferença entre os valores observado e ajustado, é definido por

$r^* = \widehat{\mathbf{W}}^{1/2}(z - \hat{\eta}) = \widehat{\mathbf{W}}^{1/2}(\widehat{\mathbf{A}}\widehat{\Delta})^{-1}(\mathbf{y} - \hat{\mu})$, em que $\widehat{\mathbf{A}} = \text{diag}(\widehat{\mathbf{A}}_1, \dots, \widehat{\mathbf{A}}_m)$ e $\widehat{\Delta} = \text{diag}(\widehat{\Delta}_1, \dots, \widehat{\Delta}_m)$, ambas com dimensão $(N \times N)$ e $\mathbf{y} = (y_1^\top, \dots, y_m^\top)^\top$ e $\hat{\mu} = (\hat{\mu}_1^\top, \dots, \hat{\mu}_m^\top)^\top$, ambos com dimensão $(N \times 1)$.

Como os elementos de r^* possuem variâncias diferentes, o que dificulta compará-los entre si, definimos o resíduo padronizado associado à observação y_{ij} por

$$(r_{SD})_{ij} = \frac{\mathbf{e}_{(ij)}^\top \widehat{\mathbf{W}}_i^{1/2} (\widehat{\mathbf{A}}\widehat{\Delta})^{-1} (\mathbf{y}_i - \hat{\mu}_i)}{\sqrt{1 - h_{ij}}}, \quad (3.25)$$

em que $\mathbf{e}_{(ij)}$ é um vetor de tamanho n com a posição referente à observação y_{ij} contendo o valor 1 e as demais posições contendo o valor zero, h_{ij} é o j -ésimo elemento da diagonal principal de $\mathbf{H}_i$, $i = 1, \dots, m$ e $j = 1, \dots, n$. Maiores detalhes ver Venezuela (2003).

Um ponto influente ocorre quando este possui um perfil diferente dos demais no que tange aos valores da variável resposta, porém apresenta valor alto na matriz de projeção $\mathbf{H}$. Esse tipo de ponto tem grande peso na estimação dos parâmetros do modelo e para detectá-lo, a medida mais conhecida é a distância de Cook. Essa distância mede o afastamento entre a estimativa do vetor paramétrico utilizando todas as observações ($\hat{\boldsymbol{\beta}}_G$) e sem a observação y_{ij} ($\hat{\boldsymbol{\beta}}_{G(ij)}$), $i = 1, \dots, m$ e $j = 1, \dots, n_i$ (Venezuela, 2013).

Como, em geral, não é possível obter uma forma fechada para $\hat{\boldsymbol{\beta}}_{G(ij)}$, tem sido utilizada uma aproximação de um passo, que consiste em tomar a primeira iteração do processo iterativo pelo método de *scoring* de Fisher quando o mesmo é iniciado em $\hat{\boldsymbol{\beta}}_G$. Para os MLG's com medidas repetidas essa aproximação é dada por

$$\hat{\boldsymbol{\beta}}_{G(ij)}^{(1)} = \hat{\boldsymbol{\beta}}_G - \frac{[\mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X}]^{-1} [\mathbf{X}^\top \widehat{\mathbf{W}}^{1/2} \mathbf{e}_{(ij)}] [\mathbf{e}_{(ij)}^\top \widehat{\mathbf{W}}^{1/2} (\widehat{\mathbf{A}} \widehat{\boldsymbol{\Delta}})^{-1} (\mathbf{y} - \widehat{\boldsymbol{\mu}})]}{1 - h_{ij}}. \quad (3.26)$$

Desta forma, a distância de Cook, quando se exclui a observação y_{ij} , é dada por

$$D_{ij} = \frac{1}{p} (\hat{\boldsymbol{\beta}}_G^{(1)} - \hat{\boldsymbol{\beta}}_{G(ij)}^{(1)})^\top \mathbf{X}^\top \widehat{\mathbf{W}} \mathbf{X} (\hat{\boldsymbol{\beta}}_G^{(1)} - \hat{\boldsymbol{\beta}}_{G(ij)}^{(1)}). \quad (3.27)$$

Maiores detalhes ver Venezuela (2003).

3.3 Seleção de Modelos

Nesta seção são apresentados os critérios utilizados para seleção de modelos considerando as abordagens multinível e marginal.

3.3.1 Modelos Multiníveis

Como apresentado no Capítulo 2, Seção 2.5, o critério de seleção para o modelo de regressão comumente utilizado é o Critério de Informação de Akaike (AIC). Porém, para os modelos multiníveis este critério de seleção não é consistente. Desta forma, é preferível utilizar o Critério de Informação de Akaike Consistente (CAIC) ou o Critério de Informação Bayesiano (BIC).

O Critério de Informação de Akaike Consistente (CAIC), proposto por Bozdogan (1987), é dado por

$$\text{CAIC} = -2\log L + p[(\log n) + 1],$$

em que L é o logaritmo da função de verossimilhança, n é o tamanho da amostra e p número de parâmetros. Maiores informações ver Bozdogan (1987).

O Critério de Informação Bayesiano (BIC), proposto por Schwarz (1978), é consistente, sendo portanto mais indicado. Assumindo que um número fixo de modelos são avaliados e que um deles é o verdadeiro modelo, o critério de seleção consistente é o que irá selecionar o verdadeiro modelo com probabilidade aproximadamente de 100% quando $n \rightarrow \infty$. O BIC é dado por

$$\text{BIC} = -2\log L + p[\log(n)]$$

em que L é o logaritmo da função de verossimilhança, n é o tamanho da amostra e p número de parâmetros. Maiores informações ver Schwarz (1978).

3.3.2 Equações de Estimação Generalizadas (EEG)

Pan (2001) propôs um método de seleção da estrutura de correlação para EEG semelhante ao AIC, mas que substitui a função de verossimilhança pela função de quase-verossimilhança. Este critério foi denominado critério de quase-verossimilhança sob o modelo de independência (QIC, em inglês).

O QIC é calculado a partir da comparação de um modelo com uma determinada estrutura de correlação de “trabalho” com aquele gerado utilizando a estrutura independente. A estrutura de correlação com o menor valor de QIC é a escolhida para ajuste do modelo (Barnett et al., 2010).

Este critério é definido como

$$\text{QIC} = -2QL(\mu) + 2q,$$

em que a função de quase-verossimilhança $QL(\mu)$ é tal que

$$QL(\mu) = \int_y^\mu \frac{y-t}{\phi v(t)} dt,$$

em que $q = \text{tr}(\hat{\Omega}_I^{-1} \hat{V}_R)$, $\mu = E(Y)$, $\text{Var}(Y) = \phi v(\mu)$, ϕ é um parâmetro de dispersão e $v(t)$ é a função de variância avaliada no ponto t (Cui, 2007; Cui, 2010).

O subscrito I denota a estrutura de correlação independente com base na qual a quase-verossimilhança é calculada. $\hat{\Omega}_I^{-1}$ representa o estimador da variância sob

a estrutura de correlação independente e $\widehat{V}_R$ é o estimador de covariância robusto obtido através do modelo contendo a matriz de correlação de “trabalho” R (Cui, 2010).

3.4 Programas Estatísticos

Nesta seção são apresentados alguns dos comandos utilizados para os ajustes dos modelos multinível e marginal para dados contínuos nos programas estatísticos **R** e STATA. Considere que y denota a variável resposta, x_1 , x_2 e x_3 são covariáveis, id representa a variável que identifica os indivíduos e $tempo$ denota o período em que as observações foram mensuradas.

3.4.1 R

Para o ajuste dos modelos multiníveis utiliza-se a biblioteca **lme4** (Boeck & Wilson, 2004). Esta biblioteca é utilizada para ajustar modelos lineares mistos e modelos lineares generalizados mistos. O comando utilizado é o **lmer** e as funções são dadas por

```
mod1 = lmer(y ~ x1 + x2 + x3 + tempo + (1 | id), data=dados)
```

```
mod2 = lmer(y ~ x1 + x2 + x3 + tempo + (tempo | id), data= dados).
```

Para ajustar o modelo multinível com intercepto aleatório (**mod1**) é necessário considerar o argumento $(1 | id)$. O modelo multinível com efeitos aleatórios no intercepto e no tempo (**mod2**) é identificado através do argumento $(tempo | id)$.

Maiores detalhes podem ser encontrados no manual da biblioteca **lme4** em <http://cran.r-project.org/web/packages/lme4/lme4.pdf>.

Para ajustar os dados utilizando a abordagem de Equações de Estimação Generalizadas (EEG) deve-se utilizar a biblioteca **geepack** criada por Jun Yan em 2012. O comando da biblioteca **geepack** utilizada é a **geeglm** e a função pode ser escrita por

```
mod3 = geeglm(y ~ x1 + x2 + x3 + tempo, data = dados, family =  
gaussian, id = id, corstr = “ ”).
```

No argumento **family** utiliza-se **gaussian**, pois a variável de interesse é contínua. A matriz de correlação de “trabalho” é determinada no argumento **corstr** e são dadas por

- **ar1** que representa a matriz de correlação de “trabalho” autorregressiva de primeira ordem;
- **exchangeable** que denota a matriz de correlação de “trabalho” permutável (*default*);
- **independence** que representa a matriz de correlação de “trabalho” independente;
- **unstructured** que define a matriz de correlação de “trabalho” não estruturada.

Em <http://cran.r-project.org/web/packages/geepack/geepack.pdf> encontram-se maiores detalhes sobre o pacote **geepack**.

3.4.2 STATA

O programa estatístico STATA também ajusta o modelo multinível para dados contínuos utilizando a função **regress** e os comandos são dados por:

```
regress y x1 x2 x3 tempo || id:
regress y x1 x2 x3 tempo || id:tempo, variance cov(un).
```

Para o modelo multinível com intercepto aleatório deve-se adicionar o termo `|| id`. No caso do modelo multinível com intercepto e tempo aleatórios utiliza-se o termo `|| id:tempo, variance cov(un)`.

Em <http://www.stata.com/help.cgi?regress> maiores detalhes podem ser encontrados .

Os modelos usando as Equações de Estimação Generalizadas para dados contínuos podem ser ajustados através do comando **xtgee** e o comando é

```
xtgee y x1 x2 x3 tempo, family(gaussian) link(identity) corr( ) i(id)
t(tempo)
```

No caso dos dados contínuos, o argumento **family** deve conter o termo **gaussian**. Como no **R**, as matrizes de correlação de “trabalho” são especificadas em **corr** através das expressões

- **exc** que denota a matriz de correlação de “trabalho” permutável;
- **ar1** que representa a matriz de correlação de “trabalho” autorregressiva de primeira ordem;

- `uns` que define a matriz de correlação de “trabalho” não estruturada.

A função de ligação é determinada no objeto `link` e neste caso é a função de ligação identidade (`identity`). Nos objetos `i( )` e `t( )` devem ser alocadas as variáveis que denotam os indivíduos e o tempo, respectivamente.

Em http://www.stata-press.com/manuals/stata12/xt_xtreg.pdf são encontrados maiores detalhes do comando `xtreg`.

Capítulo 4

Modelagem para Dados de Contagem Longitudinais

O modelo de regressão tradicional para respostas de contagem assume que as observações são independentes. Contudo, em muitos casos os dados estão agrupados, por exemplo, por conglomerados ou por desenho de estudo longitudinal. Em situações como estas a suposição de independência entre as observações não é mais razoável (Hedeker & Gibbons, 2006).

Considere, por exemplo, os estudos de utilização de serviços de saúde mental em que o foco principal da investigação é a mudança no número de ataques psicóticos com a utilização de medicamentos ao longo do tempo. Ou seja, é importante o acompanhamento do indivíduo para verificar se o medicamento está influenciando o número de ataques, sendo portanto necessário um estudo longitudinal. Da mesma forma, em estudos transversais, em que as observações são agrupadas dentro de unidades geográficas (por exemplo, o número de suicídios em municípios), as observações dentro das unidades geográficas são mais propensas a serem correlacionadas do que as observações entre as unidades geográficas. Assim, não se deve analisar estes tipos de dados sem levar em consideração a correlação existente entre as observações dentro do mesmo conglomerado.

Neste capítulo são apresentados os modelos condicionais e marginais (EEG e via cópulas Gaussianas) para modelagem de dados longitudinais de contagem. Inicialmente é discutido sucintamente o conceito de quadratura de Gauss-Hermite. Na Seção 4.2 o modelo Poisson multinível é definido. O modelo Poisson marginal considerando a abordagem EEG é discutido na Seção 4.3. O modelo Binomial Negativo multinível é apresentado na Seção 4.4. A abordagem com Equações de Estimação Generalizadas (EEG) para a distribuição Binomial Negativa é discutida na Seção 4.5. A Seção 4.6 apresenta informações sobre os modelos marginais utilizando cópulas

Gaussianas. Os pacotes e funções utilizadas para o ajuste dos modelos apresentados neste capítulo são dispostos na Seção 4.7.

4.1 Quadratura de Gauss-Hermite

O método da quadratura de Gauss-Hermite unidimensional aproxima integrais da forma

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx,$$

por uma soma ponderada, ou seja,

$$\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{q=1}^Q \omega_q f(x_q), \quad (4.1)$$

em que $f(x)$, $x \in \mathbb{R}$, representa uma função regular, o nó x_q denota a q -ésima raiz do polinômio Hermite de ordem Q , $H_Q(x)$,

$$H_Q(x) = (-1)^Q e^{x^2/2} \frac{d^Q e^{-x^2/2}}{dx^Q},$$

e os pesos ω_q dependem do número de nós Q e do polinômio Hermite $H_{Q-1}(x)$ avaliado em x_q , tal que

$$\omega_q(x_q) = \frac{2^{Q-1} Q! \sqrt{\pi}}{Q^2 [H_{Q-1}(x_q)]^2}.$$

O método de Gauss-Hermite pode ser estendido para o caso multidimensional considerando um vetor k -dimensional $x = (x_1, \dots, x_k)$, uma função regular $g(x) : \mathbb{R}^k \rightarrow \mathbb{R}$, uma função de peso $\omega(x) : \mathbb{R}^k \rightarrow \mathbb{R}_+$, tal que $g(x) = w(x)f(x)$. A expressão (4.1) pode ser utilizada para aproximar a integral k -dimensional

$$\int_{\mathbb{R}^k} g(x) dx \approx \int_{\mathbb{R}} \dots \int_{\mathbb{R}} \omega(x_1, \dots, x_k) f(x_1, \dots, x_k) dx_1 \dots dx_k, \quad (4.2)$$

em que a função de peso é dada por

$$\omega(x) = \exp(-x^\top x) = \exp(-x_1^2) \dots \exp(-x_k^2).$$

A quadratura de Gauss-Hermite unidimensional da forma (4.1) é aplicada a cada um dos componentes de x . Assim, utilizando Q_r nós na r -ésima dimensão, $r = 1, 2, \dots, k$, obtém-se a aproximação

$$\int_{\mathbb{R}^k} g(x) dx \approx \sum_{q_1=1}^{Q_1} \omega_{q_1}^{(1)} \dots \sum_{q_k=1}^{Q_k} \omega_{q_k}^{(k)} f(x_{q_1}^{(1)}, \dots, x_{q_k}^{(k)}), \quad (4.3)$$

em que $x_{q_r}^r$ é a Q_r -ésima raiz do polinômio Hermite de ordem Q_r e $\omega_{q_r}^r$ é o peso correspondente. Maiores informações podem ser encontradas em Manco (2013).

4.2 Modelo Poisson Multinível

Para mostrar o desenvolvimento dos modelos multiníveis para a distribuição Poisson estão sendo considerados dois modelos: o primeiro com o intercepto aleatório e o segundo com o intercepto e o efeito do tempo aleatórios .

Suponha que y_{ij} é a j -ésima ($j = 1, \dots, n_i$) observação inteira e não negativa de um i -ésimo indivíduo ($i = 1, \dots, m$), isto é, y_{ij} assume os valores $0, 1, 2, \dots$. Portanto, a função de densidade condicional das n_i observações individuais para o i -ésimo indivíduo é escrita como (Hedeker & Gibbons, 2006)

$$\begin{aligned} f(\mathbf{Y}_i | \mu_i) &= \prod_{j=1}^{n_i} f(y_{ij}; \mu_{ij}) \\ &= \prod_{j=1}^{n_i} \frac{\exp(-\mu_{ij}) \mu_{ij}^{y_{ij}}}{y_{ij}!}. \end{aligned} \quad (4.4)$$

O logaritmo da função de verossimilhança correspondente a (4.4) é dado por

$$\ell(\mathbf{Y}_i | \mu_i) = \sum_{j=1}^{n_i} [-\mu_{ij} + y_{ij} \log \mu_{ij} - \log(y_{ij}!)].$$

A seguir são apresentados os modelos Poisson multinível com efeito aleatório no intercepto (caso 1) e com dois efeitos aleatórios (caso 2).

4.2.1 Caso 1: Efeito Aleatório no Intercepto

Na primeira situação considere um modelo com o único efeito aleatório associado ao intercepto, b_{0i} . Vale ressaltar que b_{0i} pode seguir qualquer distribuição. No STATA, por exemplo, o b_{0i} segue distribuição Gama (*default*), quando se considera o modelo Poisson multinível. No **R** a distribuição de b_{0i} é Normal. Neste trabalho serão apresentados resultados assumindo que b_{0i} é distribuído normalmente.

Considere que o preditor linear é dado por:

$$\begin{aligned} \eta_{ij} &= \gamma_{0i} + \gamma_{1i} t_{ij} + \mathbf{X}_{ij}^\top \boldsymbol{\beta}, \quad \text{em que} \\ \gamma_{0i} &= \beta_1 + b_{0i} \quad \text{e} \\ \gamma_{1i} &= \beta_p, \end{aligned} \quad (4.5)$$

em que $\mathbf{X}_{ij}^\top = (X_{ij2}, X_{ij3}, \dots, X_{ij(p-1)})$, t_{ij} é o tempo da j -ésima medida do i -ésimo indivíduo, $\boldsymbol{\beta} = (\beta_2, \dots, \beta_{p-1})^\top$ e $b_{0i} \sim \mathcal{N}(0, \tau_0)$, sendo independente do vetor de covariáveis $\mathbf{X}_{ij}$. Padronizando a variável b_{0i} , tem-se que $\psi_i = \frac{b_{0i}}{\tau_0}$. Desta forma (4.5) pode ser reescrito como

$$\eta_{ij} = \mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i, \quad (4.6)$$

em que $\boldsymbol{\beta}^* = (\beta_1, \dots, \beta_p)^\top$ é um vetor p -dimensional de parâmetros fixos associados ao vetor de covariáveis $\mathbf{Z}_{ij}^\top = (X_{ij1}, X_{ij2}, \dots, X_{ij(p-1)}, t_{ij})$, tal que $X_{ij1} = 1, \forall i, j$ e $\psi_i \sim \mathcal{N}(0, 1)$. Desta forma, o parâmetro da distribuição Poisson, μ_{ij} , é dado por:

$$\mu_{ij} = \exp(\eta_{ij}) = \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i), \quad (4.7)$$

em que $i = 1, \dots, m$ denota os indivíduos e $j = 1, \dots, n_i$ as medidas repetidas.

A função de densidade condicional (4.4) pode ser reescrita substituindo μ_{ij} pelo resultado de (4.7). Assim $f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)$ é dada por

$$\begin{aligned} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) &= \prod_{j=1}^{n_i} f(y_{ij}; \boldsymbol{\beta}^*, \psi_i) \\ &= \prod_{j=1}^{n_i} \frac{\exp[-\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)] \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)^{y_{ij}}}{y_{ij}!}. \end{aligned} \quad (4.8)$$

Portanto, o logaritmo da função de verossimilhança correspondente a (4.8) é dado por

$$\ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) = \sum_{j=1}^{n_i} [-\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) + y_{ij}(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) - \log(y_{ij}!)].$$

Aqui $\boldsymbol{\beta}^*$ e τ_0 devem ser estimados. As primeiras e segundas derivadas em relação aos parâmetros $\boldsymbol{\beta}^*$ e τ_0 são

$$\begin{aligned} \frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \boldsymbol{\beta}^*} &= \sum_{j=1}^{n_i} [y_{ij} - \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)] \mathbf{Z}_{ij}, \\ \frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \tau_0} &= \sum_{j=1}^{n_i} [y_{ij} - \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)] \psi_i \end{aligned}$$

e

$$\begin{aligned} \frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \boldsymbol{\beta}^* \partial \boldsymbol{\beta}^{*\top}} &= - \sum_{j=1}^{n_i} \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) \mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top, \\ \frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \tau_0 \partial \tau_0^\top} &= - \sum_{j=1}^{n_i} \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) \psi_i^2. \end{aligned}$$

Para determinar a densidade marginal de $\mathbf{Y}_i$ (integração da conjunta), pode-se usar a quadratura de Gauss-Hermite unidimensional apresentada na Seção 4.1. Desta forma, a densidade marginal de $\mathbf{Y}_i$ é dada por

$$\begin{aligned}
h(\boldsymbol{\beta}^*, \tau_0, \psi; \mathbf{Y}_i) &= h(\mathbf{Y}_i) \\
&= \int_{\psi} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi) g(\psi) d\psi \\
&= \int_{-\infty}^{\infty} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi) \frac{1}{\sqrt{2\pi}} e^{-\psi^2} d\psi \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\psi^2} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi) d\psi \\
&\approx \sum_{q=1}^Q \left[\prod_{j=1}^{n_i} \frac{\exp(-\mu_{ijq}) \mu_{ijq}^{y_{ij}}}{y_{ij}!} \right] A(B_q),
\end{aligned}$$

em que $\mu_{ijq} = \exp(\mathbf{Z}_{ijq}^\top \boldsymbol{\beta}^* + \tau_0 B_q)$, B_q e $A(B_q)$ são os valores da quadratura de Gauss-Hermite e o peso no ponto q , respectivamente, com $q = 1, \dots, Q$. Maiores detalhes ver (Stroud & Sechrest, 1996). As primeiras derivadas para o logaritmo da função de verossimilhança podem ser reescritas como:

$$\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \boldsymbol{\beta}^*} = \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_q^Q \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq}) \mathbf{Z}_{ij} \frac{\exp(-\mu_{ijq}) \mu_{ijq}^{y_{ij}}}{y_{ij}} A(B_q) \quad (4.9)$$

e

$$\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \tau_0} = \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_q^Q \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq}) \psi_i \frac{\exp(-\mu_{ijq}) \mu_{ijq}^{y_{ij}}}{y_{ij}} A(B_q), \quad (4.10)$$

em que $h(\mathbf{Y}_i) \approx \sum_{q=1}^Q \left[\prod_{j=1}^{n_i} \frac{\exp(-\mu_{ijq}) \mu_{ijq}^{y_{ij}}}{y_{ij}!} \right] A(B_q)$.

Os parâmetros do modelo podem ser estimados igualando as derivadas (4.9) e (4.10) a zero e resolvendo de forma iterativa essas equações utilizando os métodos de escore de Fisher ou Newton-Raphson (Hedeker & Gibbons, 2006).

4.2.2 Caso 2: Dois Efeitos Aleatórios

Considere um modelo multinível mais complexo que contém dois efeitos aleatórios, ou seja, o preditor linear é dado por:

$$\begin{aligned}
\eta_{ij} &= \gamma_{0i} + \mathbf{X}_{ij}^\top \boldsymbol{\beta} + \gamma_{1i} t_{ij} \quad \text{tal que,} \\
\gamma_{0i} &= \beta_1 + b_{0i} \quad \text{e} \\
\gamma_{1i} &= \beta_p + b_{1i},
\end{aligned} \quad (4.11)$$

em que $b_i = [b_{0i} \ b_{1i}]^\top \sim \mathcal{N}_2\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_0 & \tau_{01} \\ \tau_{01} & \tau_1 \end{bmatrix}\right)$, t_{ij} é o tempo da j -ésima medida do i -ésimo indivíduo e $\mathbf{X}_{ij}^\top = (X_{ij2}, X_{ij3}, \dots, X_{ij(p-1)})$ é um vetor de covariáveis. A função de distribuição normal bivariada de b_{0i} e b_{1i} é dada por

$$g(b_{0i}, b_{1i}) = \frac{1}{2\pi\sqrt{\tau_{0i}}\sqrt{\tau_{1i}}\sqrt{(1-\tau_{01}^2)}} \exp\left\{-\frac{1}{2(1-\tau_{01}^2)}M\right\},$$

$$\text{em que } M = \left[\left(\frac{b_{0i}}{\tau_0}\right)^2 + \left(\frac{b_{1i}}{\tau_1}\right)^2 - 2\tau_{01}\left(\frac{b_{0i}}{\tau_0}\right)\left(\frac{b_{1i}}{\tau_1}\right)\right].$$

Padronizando os dois efeitos aleatórios, tem-se que $\psi_i = b_{0i}/\tau_0$ e $\zeta_i = b_{1i}/\tau_1$. A função acima passa a ser

$$g(\psi_i, \zeta_i) = \frac{1}{2\pi\sqrt{(1-\tau_{01}^2)}} \exp\left\{-\left[\frac{\psi_i^2 + \zeta_i^2 - 2\tau_{01}\psi_i\zeta_i}{2(1-\tau_{01}^2)}\right]\right\}.$$

Desta forma, o preditor linear (4.11) pode ser reescrito como

$$\begin{aligned} \eta_{ij} &= \beta_1 + \mathbf{X}_{ij}^\top \boldsymbol{\beta} + \beta_p t_{ij} + b_{0i} + b_{1i} t_{ij} \\ &= \mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i} + b_{1i} t_{ij} \\ &= \mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}, \end{aligned} \quad (4.12)$$

em que $\mathbf{Z}_{ij}^\top = (X_{ij1}, X_{ij2}, \dots, X_{ij(p-1)}, t_{ij})$ é um vetor de dimensão $n_i \times p$ de covariáveis, sendo $X_{ij1} = 1, \forall \ i, j$ e $\boldsymbol{\beta}^* = (\beta_1, \beta_2, \dots, \beta_{p-1}, \beta_p)^\top$ é um vetor $(p \times 1)$ de parâmetros. Portanto, considerando o preditor linear (4.12), obtém-se

$$\mu_{ij} = \exp(\eta_{ij}) = \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}). \quad (4.13)$$

Novamente é possível reescrever a função de densidade condicional apresentada em (4.4) substituindo μ_{ij} pelo resultado de (4.13). Portanto $f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)$ é dada por

$$\begin{aligned} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i) &= \prod_{j=1}^{n_i} f(y_{ij}; \boldsymbol{\beta}^*, \psi_i, \zeta_i) \\ &= \prod_{j=1}^{n_i} \frac{e^{[-\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})]} e^{(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) y_{ij}}}{y_{ij}!}. \end{aligned} \quad (4.14)$$

Desta forma, o logaritmo da função de verossimilhança correspondente à (4.14), é dado por

$$\begin{aligned} \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i) &= \sum_{j=1}^{n_i} [-\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) \\ &\quad + y_{ij}(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) - \log(y_{ij}!)]. \end{aligned}$$

Aqui β^* , τ_0 e τ_1 devem ser estimados. Portanto, as primeiras e segundas derivadas em relação a β^* , τ_0 e τ_1 são

$$\begin{aligned}\frac{\partial \ell(\mathbf{Y}_i | \beta^*, \psi_i, \zeta_i)}{\partial \beta^*} &= \sum_{j=1}^{n_i} [y_{ij} - \exp(\mathbf{Z}_{ij}^\top \beta^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})] \mathbf{Z}_{ij}, \\ \frac{\partial \ell(\mathbf{Y}_i | \beta^*, \psi_i, \zeta_i)}{\partial \tau_0} &= \sum_{j=1}^{n_i} [y_{ij} - \exp(\mathbf{Z}_{ij}^\top \beta^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})] \psi_i, \\ \frac{\partial \ell(\mathbf{Y}_i | \beta^*, \psi_i, \zeta_i)}{\partial \tau_1} &= \sum_{j=1}^{n_i} [y_{ij} - \exp(\mathbf{Z}_{ij}^\top \beta^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})] \zeta_i t_{ij}\end{aligned}$$

e

$$\begin{aligned}\frac{\partial^2 \ell(\mathbf{Y}_i | \beta^*, \psi_i, \zeta_i)}{\partial \beta^* \partial \beta^{*\top}} &= - \sum_{j=1}^{n_i} \exp(\mathbf{Z}_{ij}^\top \beta^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) \mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top, \\ \frac{\partial^2 \ell(\mathbf{Y}_i | \beta^*, \psi_i, \zeta_i)}{\partial \tau_0 \partial \tau_0} &= - \sum_{j=1}^{n_i} (\exp(\mathbf{Z}_{ij}^\top \beta^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})) \psi_i^2, \\ \frac{\partial^2 \ell(\mathbf{Y}_i | \beta^*, \psi_i, \zeta_i)}{\partial \tau_1 \partial \tau_1} &= - \sum_{j=1}^{n_i} \exp(\mathbf{Z}_{ij}^\top \beta^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) \zeta_i^2 t_{ij}^2.\end{aligned}$$

A densidade marginal de $\mathbf{Y}_i$ é determinada utilizando a quadratura de Gauss-Hermite bidimensional, que é um caso particular da quadratura de Gauss-Hermite k -dimensional apresentada na Seção 4.1 quando $k = 2$. Desta forma, a densidade marginal de $\mathbf{Y}_i$ é dada por

$$\begin{aligned}h(\beta^*, \tau_0, \psi, \tau_1, \zeta; \mathbf{Y}_i) &= h(\mathbf{Y}_i) \\ &= \int_{\psi} \int_{\zeta} f(\mathbf{Y}_i | \beta^*, \psi, \zeta) g(\psi, \zeta) d\psi d\zeta \\ &\approx \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \left[\prod_{j=1}^{n_i} \frac{\exp(-\mu_{ijq_1q_2}) \mu_{ijq_1q_2}^{y_{ij}}}{y_{ij}!} \right] A(B_{q_1}) A(B_{q_2}),\end{aligned}$$

em que $\mu_{ijq_1q_2} = \exp(\mathbf{Z}_{ijq_1q_2}^\top \beta^* + \tau_0 B_{q_1} + \tau_1 B_{q_2})$, B_{q_1} e $A(B_{q_1})$ são os valores da quadratura de Gauss-Hermite e o peso em q_1 , respectivamente, com $q_1 = 1, \dots, Q_1$ e B_{q_2} e $A(B_{q_2})$ são os valores da quadratura de Gauss-Hermite e o peso em q_2 , respectivamente, com $q_2 = 1, \dots, Q_2$. As primeiras derivadas para o logaritmo da verossimilhança são

$$\begin{aligned}\frac{\partial \ell(\mathbf{Y}_i | \beta^*, \psi_i, \tau_{1i})}{\partial \beta^*} &= \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_{q_1}^{Q_1} \sum_{q_2}^{Q_2} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ij}^*) \mathbf{Z}_{ij} \frac{\exp(-\mu_{ij}^*) (\mu_{ij}^*)^{y_{ij}}}{y_{ij}} A(B_{q_1}) A(B_{q_2}), \\ \frac{\partial \ell(\mathbf{Y}_i | \beta^*, \psi_i, \tau_{1i})}{\partial \tau_0} &= \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_{q_1}^{Q_1} \sum_{q_2}^{Q_2} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ij}^*) \psi_i \frac{\exp(-\mu_{ij}^*) (\mu_{ij}^*)^{y_{ij}}}{y_{ij}} A(B_{q_1}) A(B_{q_2}),\end{aligned}$$

e

$$\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \tau_{1i})}{\partial \tau_1} = \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ij}^*) \zeta_i \frac{\exp(-\mu_{ij}^*) (\mu_{ij}^*)^{y_{ij}}}{y_{ij}} A(B_{q_1}) A(B_{q_2}),$$

em que $\mu_{ij}^* = \mu_{ijq_1q_2}$ e $h(\mathbf{Y}_i) \approx \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \left[\prod_{j=1}^{n_i} \frac{\exp(-\mu_{ij}^*) (\mu_{ij}^*)^{y_{ij}}}{y_{ij}!} \right] A(B_{q_1}) A(B_{q_2})$.

Como no Caso 1, é possível estimar os parâmetros igualando as derivadas acima à zero e resolvendo de forma iterativa o sistema de equações usando um algoritmo Newton-Raphson ou Quasi-Newton.

4.3 Modelo Poisson Marginal (Abordagem EEG)

Considerando a distribuição Poisson, tem-se que, para o i -ésimo indivíduo na j -ésima medida repetida, com $i = 1, \dots, m$ e $j = 1, \dots, n_i$, o preditor linear é dado por

$$\eta_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\beta}, \quad (4.15)$$

em que $\mathbf{X}_{ij}$ é o vetor de covariáveis para o i -ésimo indivíduo na j -ésima medida repetida. A função de ligação é o logaritmo (log), isto é, $\log(\mu_{ij}) = \mathbf{X}_{ij}^\top \boldsymbol{\beta}$, ou seja $\mu_{ij} = \exp(\mathbf{X}_{ij}^\top \boldsymbol{\beta})$.

Quando se considera uma resposta de contagem Y_{ij} , a distribuição Poisson especifica que a média e a variância de Y_{ij} são iguais. Assim,

$$\text{Var}(Y_{ij}) = \mu_{ij}, \quad (4.16)$$

e $\phi = 1$, ou seja, não é estimado (Hedeker & Gibbons, 2006).

4.3.1 Estimação

Seja $\mathbf{A}_i$ uma matriz diagonal de dimensão $n_i \times n_i$, com $\text{Var}(Y_{ij})$ como o j -ésimo elemento diagonal, ou seja,

$$\mathbf{A}_i = \begin{bmatrix} \text{Var}(Y_{i1}) & 0 & \cdots & 0 \\ 0 & \text{Var}(Y_{i2}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{Var}(Y_{in_i}) \end{bmatrix}. \quad (4.17)$$

A matriz de correlação de “trabalho” de dimensão $n_i \times n_i$ para o i -ésimo indivíduo é denotada por $R_i(\boldsymbol{\alpha})$, em que $\boldsymbol{\alpha}$ é um vetor de parâmetros de associação. Então, a matriz de covariância para $\mathbf{Y}_i$ é dada por

$$V_i(\boldsymbol{\alpha}) = \phi \mathbf{A}_i^{1/2} R_i(\boldsymbol{\alpha}) \mathbf{A}_i^{1/2}. \quad (4.18)$$

O estimador EEG de $\boldsymbol{\beta}$ é obtido através da solução da equação de estimação

$$\sum_{i=1}^m D_i^\top [V(\hat{\boldsymbol{\alpha}})]^{-1} (\mathbf{Y}_i - \boldsymbol{\mu}_i) = 0, \quad (4.19)$$

em que $\hat{\boldsymbol{\alpha}}$ é um estimador consistente de α e $D_i = \partial \boldsymbol{\mu}_i / \partial \boldsymbol{\beta}$.

Dado que $\text{Var}(Y_i) = \mu_i$, a matriz diagonal $\mathbf{A}_i$ pode ser escrita como

$$\mathbf{A}_i = \begin{bmatrix} \mu_{i1} & 0 & \cdots & 0 \\ 0 & \mu_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mu_{in_i} \end{bmatrix}.$$

No caso da distribuição Poisson, tem-se que $D_i = \partial \boldsymbol{\mu}_i / \partial \boldsymbol{\beta} = \exp(\mathbf{X}_i^\top \boldsymbol{\beta}) \mathbf{X}_i$. Logo, o estimador EEG de $\boldsymbol{\beta}$ é dado pela solução da equação

$$\sum_{i=1}^m [\exp(\mathbf{X}_i^\top \boldsymbol{\beta}) \mathbf{X}_i]^\top [\mathbf{A}_i^{1/2} R_i(\hat{\boldsymbol{\alpha}}) \mathbf{A}_i^{1/2}]^{-1} [\mathbf{Y}_i - \exp(\mathbf{X}_i^\top \boldsymbol{\beta})] = 0, \quad (4.20)$$

a qual é uma equação não linear em $\boldsymbol{\beta}$ e desta forma sua solução pode ser alcançada usando algum método iterativo, como, por exemplo, Newton-Raphson.

4.4 Modelo Binomial Negativo Multinível

Para mostrar o desenvolvimento dos modelos multiníveis para a distribuição Binomial Negativa serão considerados também dois modelos: o primeiro com o intercepto aleatório e o segundo com o intercepto e o efeito do tempo aleatórios. Além disso, assume-se que o parâmetro ν_{ij} é conhecido.

Suponha que y_{ij} é a j -ésima observação do i -ésimo indivíduo, em que $i = 1, \dots, m$ e $j = 1, \dots, n_i$. Considere que o vetor $\mathbf{Y}_i = (y_{i1}, \dots, y_{in_i})^\top$ segue distribuição Binomial Negativa com parâmetros ν e $(\frac{\nu/\mu}{1+\nu/\mu})$. Assim a função de densidade condicional da n_i observações individuais para o i -ésimo indivíduo é dado por

$$\begin{aligned} f(\mathbf{Y}_i | \mu_i, \nu_i) &= \prod_{j=1}^{n_i} f\left(y_{ij}; \nu_{ij}, \frac{\nu_{ij}/\mu_{ij}}{1 + \nu_{ij}/\mu_{ij}}\right) \\ &= \prod_{j=1}^{n_i} \frac{(y_{ij} + \nu_{ij} - 1)!}{y_{ij}! (\nu_{ij} - 1)!} \left(\frac{\nu_{ij}/\mu_{ij}}{\frac{\nu_{ij}}{\mu_{ij}} + 1}\right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ij}} + 1}\right)^{y_{ij}} \\ &= \prod_{j=1}^{n_i} c_1 \left(\frac{\nu_{ij}/\mu_{ij}}{\frac{\nu_{ij}}{\mu_{ij}} + 1}\right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ij}} + 1}\right)^{y_{ij}}, \end{aligned} \quad (4.21)$$

em que $c_1 = \frac{(y_{ij} + \nu_{ij} - 1)!}{y_{ij}!(\nu_{ij} - 1)!}$.

Note que $E(\mathbf{Y}_i) = \mu_i$ e $\text{Var}(\mathbf{Y}_i) = \mu_i + \frac{\mu_i^2}{\nu_i}$. O logaritmo da função de verossimilhança correspondente a (4.21) é dado por

$$\begin{aligned}\ell(\mathbf{Y}_i | \mu_i, \nu_i) &= \sum_{j=1}^{n_i} \left\{ \log[(y_{ij} + \nu_{ij} + 1)!] - \log[y_{ij}!(\nu_{ij} - 1)!] + y_{ij} \left[-\log\left(\frac{\nu_{ij}}{\mu_{ij}} + 1\right) \right] \right. \\ &\quad \left. + \nu_{ij} \left[\log\left(\frac{\nu_{ij}}{\mu_{ij}}\right) - \log\left(\frac{\nu_{ij}}{\mu_{ij}} + 1\right) \right] \right\} \\ &= \sum_{j=1}^{n_i} \left\{ \log[(y_{ij} + \nu_{ij} + 1)!] - \log(y_{ij}!) - \log[(\nu_{ij} - 1)!] - \nu_{ij} \log\left(\frac{\nu_{ij}}{\mu_{ij}} + 1\right) \right. \\ &\quad \left. + \nu_{ij} \log(\nu_{ij}) - \nu_{ij} \log(\mu_{ij}) - y_{ij} \left[\log\left(\frac{\nu_{ij}}{\mu_{ij}} + 1\right) \right] \right\} \\ &= \sum_{j=1}^{n_i} \left\{ c_2 - \nu_{ij} \log(\mu_{ij}) - \nu_{ij} \log\left(\frac{\nu_{ij}}{\mu_{ij}} + 1\right) - y_{ij} \left[\log\left(\frac{\nu_{ij}}{\mu_{ij}} + 1\right) \right] \right\}, \quad (4.22)\end{aligned}$$

em que $c_2 = \log[(y_{ij} + \nu_{ij} + 1)!] - \log(y_{ij}!) - \log[(\nu_{ij} - 1)!] + \nu_{ij} \log(\nu_{ij})$.

Como no modelo Poisson multinível, qualquer distribuição pode ser considerada para o efeito aleatório do intercepto (b_{0i}) para modelo Binomial Negativo multinível. No STATA, por exemplo, o b_{0i} segue distribuição Beta. No **R** a distribuição de b_{0i} é Normal. Neste trabalho serão apresentados resultados para os dois casos. Para o modelo Binomial Negativo multinível com dois efeitos aleatórios considera-se que a distribuição conjunta dos efeitos b_{0i} e b_{1i} é normal bivariada.

4.4.1 Caso 1: Efeito Aleatório no Intercepto

4.4.1.1 b_{0i} segue distribuição normal

Considere o mesmo preditor linear apresentado na equação (4.6). Neste caso, o logaritmo da função de verossimilhança é dada por:

$$\begin{aligned}\ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) &= \sum_{j=1}^{n_i} \left\{ c_2 - \nu_{ij} \log\left(\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)\right) - \nu_{ij} \log\left(\frac{\nu_{ij}}{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)} + 1\right) \right. \\ &\quad \left. - y_{ij} \left[\log\left(\frac{\nu_{ij}}{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)} + 1\right) \right] \right\}, \\ &= \sum_{j=1}^{n_i} \left\{ c_2 - \nu_{ij} (\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) - \nu_{ij} \log\left(\nu_{ij} [\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)]^{-1} + 1\right) \right. \\ &\quad \left. - y_{ij} \log\left(\nu_{ij} [\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)]^{-1} + 1\right) \right\},\end{aligned}$$

em que $c_2 = \log[(y_{ij} + \nu_{ij} + 1)!] - \log(y_{ij}!) - \log[(\nu_{ij} - 1)!] + \nu_{ij} \log(\nu_{ij})$. Aqui $\boldsymbol{\beta}^*$ e τ_0 são os parâmetros a serem estimados. As primeiras e segundas derivadas em relação a $\boldsymbol{\beta}^*$ e τ_0 são

$$\begin{aligned}\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \boldsymbol{\beta}^*} &= \sum_{j=1}^{n_i} \left\{ -\nu_{ij} + (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)}{[\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) + 1]} \right\} \mathbf{Z}_{ij}, \\ \frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \tau_0} &= \sum_{j=1}^{n_i} \left\{ -\nu_{ij} + (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)}{[\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) + 1]} \right\} \psi_i\end{aligned}$$

e

$$\begin{aligned}\frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \boldsymbol{\beta}^* \partial \boldsymbol{\beta}^{*\top}} &= \sum_{j=1}^{n_i} \left\{ (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)}{[\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) + 1]^2} \right\} \mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top, \\ \frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \tau_0 \partial \tau_0^\top} &= \sum_{j=1}^{n_i} \left\{ (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i)}{[\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i) + 1]^2} \right\} \psi_i^2.\end{aligned}$$

Assim como no modelo Poisson multinível, para determinar a densidade marginal de $\mathbf{Y}_i$ no modelo Binomial Negativo multinível pode-se usar a quadratura de Gauss-Hermite unidimensional. Desta forma, a densidade marginal de $\mathbf{Y}_i$ é dada por

$$\begin{aligned}h(\boldsymbol{\beta}^*, \tau_0, \psi; \mathbf{Y}_i) &= h(\mathbf{Y}_i) \\ &= \int_{\psi} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) g(\psi) d\psi \\ &= \int_{-\infty}^{\infty} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) \frac{1}{\sqrt{2\pi}} e^{-\psi^2} d\psi \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\psi^2} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) d\psi \\ &\approx \sum_{q=1}^Q \left[\prod_{j=1}^{n_i} c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq}}}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{y_{ij}} \right] A(B_q),\end{aligned}$$

em que $c_1 = \frac{(y_{ij} + \nu_{ij} - 1)!}{y_{ij}! (\nu_{ij} - 1)!}$, $\mu_{ijq} = \exp(\mathbf{Z}_{ijq}^\top \boldsymbol{\beta}^* + \tau_0 B_q)$ e B_q e $A(B_q)$ são os valores da quadratura de Gauss-Hermite e o peso em q , respectivamente, com $q = 1, \dots, Q$.

As primeiras derivadas para o logaritmo da verossimilhança são

$$\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \boldsymbol{\beta}^*} = \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_q \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq}) \mathbf{Z}_{ij} \left[c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq}}}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{y_{ij}} \right] A(B_q)$$

e

$$\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i)}{\partial \tau_0} = \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_q \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq}) \theta_i \left[c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq}}}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{y_{ij}} \right] A(B_q),$$

em que

$$h(\mathbf{Y}_i) = \sum_{q=1}^Q \left[\prod_{j=1}^{n_i} c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq}}}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq}} + 1} \right)^{y_{ij}} \right] A(B_q).$$

Os parâmetros do modelo podem ser estimados igualando essas derivadas a zero e resolvendo de forma iterativa utilizando métodos iterativos como Escore de Fisher ou Newton-Raphson.

4.4.1.2 b_{0i} segue distribuição Beta

Suponha que $\mathbf{Y}_i$ segue distribuição BN com função de probabilidade dada por (4.21). O modelo com o único efeito aleatório associado ao intercepto, b_{0i} , ou seja, o preditor linear é dado por:

$$\begin{aligned}\eta_{ij} &= \gamma_{0i} + \gamma_{1i}t_{ij} + X_{ij}^\top \boldsymbol{\beta}, \quad \text{em que} \\ \gamma_{0i} &= \boldsymbol{\beta}_1 + b_{0i} \quad \text{e} \\ \gamma_{1i} &= \boldsymbol{\beta}_p.\end{aligned}\tag{4.23}$$

Aqui, $\mathbf{X}_{ij}^\top = (X_{ij2}, X_{ij3}, \dots, X_{ij(p-1)})$, t_{ij} é o tempo da j -ésima medida do i -ésimo indivíduo, $\boldsymbol{\beta} = (\boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_{p-1})^\top$ e $b_{0i} \sim \text{Beta}(\lambda, \tau_0)$, cuja função densidade de probabilidade é dada por

$$g(b_{0i}; \lambda, \tau_0) = \frac{\Gamma(\tau_0)}{\Gamma(\lambda\tau_0)\Gamma((1-\lambda)\tau_0)} b_{0i}^{\lambda\tau_0-1} (1-b_{0i})^{(1-\lambda)\tau_0-1},$$

em que $0 < b_{0i} < 1$, $0 < \lambda < 1$ e $\tau_0 > 0$. A média e a variância de b_{0i} são: $E(b_{0i}) = \lambda$ e $\text{Var}(b_{0i}) = \lambda(1-\lambda)/(1+\tau_0)$.

Pode-se reescrever (4.23) como

$$\eta_{ij} = \mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i},\tag{4.24}$$

em que $\boldsymbol{\beta}^* = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_p)^\top$ é um vetor p -dimensional de parâmetros fixos associados ao vetor de covariáveis $\mathbf{Z}_{ij}^\top = (X_{ij1}, X_{ij2}, \dots, X_{ij(p-1)}, t_{ij})$, tal que $X_{ij1} = 1, \forall i, j$ e $b_{0i} \sim \text{Beta}(\lambda, \tau_0)$. Desta forma, o parâmetro da distribuição Poisson, μ_{ij} , é dado por:

$$\mu_{ij} = \exp(\eta_{ij}) = \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i}),\tag{4.25}$$

em que $i = 1, \dots, m$ denota os indivíduos e $j = 1, \dots, n_i$ as medidas repetidas e o logaritmo da função de verossimilhança é dada por

$$\begin{aligned}\ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i) &= \sum_{j=1}^{n_i} \left\{ c_2 - \nu_{ij} \log \left(\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i}) \right) - \nu_{ij} \log \left(\frac{\nu_{ij}}{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i})} + 1 \right) \right. \\ &\quad \left. - y_{ij} \left[\log \left(\frac{\nu_{ij}}{\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i})} + 1 \right) \right] \right\}, \\ &= \sum_{j=1}^{n_i} \left\{ c_2 - \nu_{ij} (\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i}) - \nu_{ij} \log \left(\nu_{ij} [\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i})]^{-1} + 1 \right) \right. \\ &\quad \left. - y_{ij} \log \left(\nu_{ij} [\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + b_{0i})]^{-1} + 1 \right) \right\},\end{aligned}$$

em que $c_2 = \log[(y_{ij} + \nu_{ij} + 1)!] - \log(y_{ij}!) - \log[(\nu_{ij} - 1)!] + \nu_{ij} \log(\nu_{ij})$. A densidade marginal de $\mathbf{Y}_i$ é dada por

$$\begin{aligned} h(\boldsymbol{\beta}^*, \lambda, \tau_0; \mathbf{Y}_i) &= h(\mathbf{Y}_i) \\ &= \int_{b_{0i}} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, b_{0i}) g(b_{0i}; \lambda, \tau_0) db_{0i} \\ &= \int_0^1 f(\mathbf{Y}_i | \boldsymbol{\beta}^*, b_{0i}) \frac{\Gamma(\tau_0)}{\Gamma(\lambda\tau_0)\Gamma((1-\lambda)\tau_0)} b_{0i}^{\lambda\tau_0-1} (1-b_{0i})^{(1-\lambda)\tau_0-1} db_{0i}. \end{aligned} \quad (4.26)$$

em que $f(\mathbf{Y}_i | \boldsymbol{\beta}^*)$ é dada na Equação (4.21).

Suponha que é feita a seguinte mudança de variável, $b_{0i} = \frac{1}{1+e^{-a}}$. Então, (4.26) pode ser reescrito como:

$$\begin{aligned} h(\boldsymbol{\beta}^*, \lambda, \tau_0; \mathbf{Y}_i) &= \int_{-\infty}^{\infty} f(\mathbf{Y}_i | \boldsymbol{\beta}^*) g\left(\frac{1}{1+e^{-a}}\right) \frac{e^{-a}}{1+e^{-a}} da \\ &= \int_{-\infty}^{\infty} f(\mathbf{Y}_i | \boldsymbol{\beta}^*) g\left(\frac{1}{1+e^{-a}}\right) \frac{e^{-a}}{1+e^{-a}} \frac{e^{-a^2}}{e^{-a^2}} da \\ &= \int_{-\infty}^{\infty} e^{-a^2} \left[f(\mathbf{Y}_i | \boldsymbol{\beta}^*) g\left(\frac{1}{1+e^{-a}}\right) \frac{e^{-a^2}}{1+e^{-a}} \right] da \\ &= \int_{-\infty}^{\infty} e^{-a^2} g(a) da. \end{aligned}$$

Portanto, é possível utilizar a quadratura de Gauss-Hermite para aproximar $h(\boldsymbol{\beta}^*, \lambda, \tau_0; \mathbf{Y}_i)$. A função (4.26) não possui expressão analítica de forma fechada e precisa-se usar algum método numérico para aproximá-la.

4.4.2 Caso 2: Dois Efeitos Aleatórios

Considere agora um modelo Binomial Negativo multinível mais complexo no qual há dois efeitos aleatório, cuja distribuição conjunta é normal bivariada. Desta forma, o preditor linear é dado pela equação (4.12). O logaritmo da função de verossimilhança é dado por

$$\begin{aligned} \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i) &= \sum_{j=1}^{n_i} \left\{ c_2 - \nu_{ij} (\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) \right. \\ &\quad - \nu_{ij} \log \left(\nu_{ij} [\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})]^{-1} + 1 \right) \\ &\quad \left. - y_{ij} \log \left(\nu_{ij} [\exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})]^{-1} + 1 \right) \right\} \end{aligned}$$

em que $c_2 = \log[(y_{ij} + \nu_{ij} + 1)!] - \log(y_{ij}!) - \log[(\nu_{ij} - 1)!] + \nu_{ij} \log(\nu_{ij})$.

As primeiras e segundas derivadas da função acima em relação a $\boldsymbol{\beta}^*$, τ_0 e τ_1 são

$$\begin{aligned}
\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \boldsymbol{\beta}^*} &= \sum_{j=1}^{n_i} \left\{ -\nu_{ij} + (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})}{[\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) + 1]} \right\} \mathbf{Z}_{ij}, \\
\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \tau_0} &= \sum_{j=1}^{n_i} \left\{ -\nu_{ij} + (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})}{[\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) + 1]} \right\} \psi_i, \\
\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \tau_1} &= \sum_{j=1}^{n_i} \left\{ -\nu_{ij} + (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})}{[\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) + 1]} \right\} \zeta_i t_{ij}
\end{aligned}$$

e

$$\begin{aligned}
\frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \boldsymbol{\beta}^* \partial \boldsymbol{\beta}^{*\top}} &= \sum_{j=1}^{n_i} \left\{ (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})}{[\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) + 1]^2} \right\} \mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top, \\
\frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \tau_0 \partial \tau_0^\top} &= \sum_{j=1}^{n_i} \left\{ (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})}{[\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) + 1]^2} \right\} \psi_i^2, \\
\frac{\partial^2 \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \tau_1 \partial \tau_1^\top} &= \sum_{j=1}^{n_i} \left\{ (\nu_{ij} + y_{ij}) \frac{\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij})}{[\exp(\mathbf{Z}_{ij} \boldsymbol{\beta}^* + \tau_0 \psi_i + \tau_1 \zeta_i t_{ij}) + 1]^2} \right\} \zeta_i^2 t_{ij} t_{ij}^\top.
\end{aligned}$$

Utilizando a quadratura de Gauss-Hermite bidimensional, é possível calcular a densidade marginal $h(\boldsymbol{\beta}^*, \tau_0, \theta, \tau_1, \zeta; \mathbf{Y}_i) = h(\mathbf{Y}_i)$ através de

$$\begin{aligned}
h(\mathbf{Y}_i) &= \int_{\psi} \int_{\zeta} f(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi, \zeta) g(\psi, \zeta) d\psi d\zeta \\
&\approx \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \left[\prod_{j=1}^{n_i} c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq_1q_2}}}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{y_{ij}} \right] A(B_{q_1}) A(B_{q_2}),
\end{aligned}$$

em que $c_1 = \frac{(y_{ij} + \nu_{ij} - 1)!}{y_{ij}! (\nu_{ij} - 1)!}$, $\mu_{ijq_1q_2} = \exp(\mathbf{Z}_{ij}^\top \boldsymbol{\beta}^* + \tau_0 B_{q_1} + \tau_1 B_{q_2})$, B_{q_1} e $A(B_{q_1})$ são os valores da quadratura de Gauss-Hermite e o peso em q_1 , respectivamente, com $q_1 = 1, \dots, Q_1$ e B_{q_2} e $A(B_{q_2})$ são os valores da quadratura de Gauss-Hermite e o peso em q_2 , respectivamente, com $q_2 = 1, \dots, Q_2$.

As primeiras derivadas para o logaritmo da função de verossimilhança são

$$\begin{aligned}
\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \boldsymbol{\beta}^*} &= \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq_1q_2}) \\
&\quad \mathbf{Z}_{ij} \left[c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq_1q_2}}}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{y_{ij}} \right] A(B_{q_1}) A(B_{q_2}), \\
\frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \tau_0} &= \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq_1q_2}) \\
&\quad \psi_i \left[c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq_1q_2}}}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{y_{ij}} \right] A(B_{q_1}) A(B_{q_2}),
\end{aligned}$$

e

$$\begin{aligned} \frac{\partial \ell(\mathbf{Y}_i | \boldsymbol{\beta}^*, \psi_i, \zeta_i)}{\partial \tau_1} &= \sum_{i=1}^m \frac{1}{h(\mathbf{Y}_i)} \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \sum_{j=1}^{n_i} (y_{ij} - \mu_{ijq_1q_2}) \\ &\quad \zeta_i \left[c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq_1q_2}}}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{y_{ij}} \right] A(B_{q_1}) A(B_{q_2}), \end{aligned}$$

em que

$$h(\mathbf{Y}_i) \approx \sum_{q_1=1}^{Q_1} \sum_{q_2=1}^{Q_2} \left[\prod_{j=1}^{n_i} c_1 \left(\frac{\frac{\nu_{ij}}{\mu_{ijq_1q_2}}}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{\nu_{ij}} \left(\frac{1}{\frac{\nu_{ij}}{\mu_{ijq_1q_2}} + 1} \right)^{y_{ij}} \right] A(B_{q_1}) A(B_{q_2}).$$

Novamente estima-se os parâmetros do modelo igualando as equações acima à zero e utilizando os métodos de escore de Fisher ou Newton-Raphson na resolução do sistema de equações.

4.5 Modelo Binomial Negativo Marginal (Abordagem EEG)

Considerando a distribuição Binomial Negativa apresentada em (??) e o preditor linear dado por $\eta_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\beta}$, tem-se que $\mu_{ij} = \exp(\mathbf{X}_{ij}^\top \boldsymbol{\beta})$.

Como apresentado na Seção 4.4, se Y_{ij} tem distribuição Binomial Negativa, a média e a variância de Y_{ij} são dadas, respectivamente, por $E(Y_{ij}) = \mu_{ij}$ e $\text{Var}(Y_{ij}) = \mu_{ij} + \frac{\mu_{ij}^2}{\nu_{ij}}$. Assim, no modelo Binomial Negativo EEG, tem-se que:

$$\text{Var}(Y_{ij}) = \phi \left[\mu_{ij} + \frac{\mu_{ij}^2}{\nu_{ij}} \right], \quad (4.27)$$

em que $i = 1, \dots, m$, $j = 1, \dots, n_i$ e ϕ pode ser conhecido ou estimado.

4.5.1 Estimação

Considerando a matriz diagonal $\mathbf{A}_i$ apresentada em (4.17) e sabendo que a variância de Y_{ij} é dada em (4.27), tem-se que o estimador EEG de $\boldsymbol{\beta}$ é dado pela solução da equação

$$\sum_{i=1}^m [\exp(\mathbf{X}_i^\top \boldsymbol{\beta}) \mathbf{X}_i]^\top [\mathbf{A}_i^{1/2} R_i(\hat{\boldsymbol{\alpha}}) \mathbf{A}_i^{1/2}]^{-1} [\mathbf{Y}_i - \exp(\mathbf{X}_i^\top \boldsymbol{\beta})] = 0, \quad (4.28)$$

em que

$$\mathbf{A}_i = \begin{bmatrix} \phi \left[\mu_{i1} + \frac{\mu_{i1}^2}{\nu_{i1}} \right] & 0 & \cdots & 0 \\ 0 & \phi \left[\mu_{i2} + \frac{\mu_{i2}^2}{\nu_{i2}} \right] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \phi \left[\mu_{in_i} + \frac{\mu_{in_i}^2}{\nu_{in_i}} \right] \end{bmatrix},$$

em que $\nu_{i1}, \nu_{i2}, \dots, \nu_{in_i}$ são conhecidos e ϕ pode ser conhecido ou estimado.

A equação (4.28) é uma equação não linear em $\boldsymbol{\beta}$ e, desta forma, sua solução pode ser alcançada usando algum método iterativo, como, por exemplo, Newton-Raphson.

Por questão de dificuldades computacionais não foi possível utilizar o modelo BN marginal utilizando a abordagem EEG para a realização dos estudos de simulação. Por este motivo, utilizou-se uma segunda abordagem para o modelo marginal que é apresentada a seguir.

4.6 Modelo Marginal (Cópulas Gaussianas)

Modelos marginais para respostas correlacionadas não-normais são tipicamente ajustados pela abordagem de Equações de Estimação Generalizadas proposta por Liang & Zeger (1986). Apesar das várias vantagens teóricas e práticas, análise de verossimilhança para modelos marginais não-normal não é muito difundido, pois é verificável uma grande dificuldade na identificação das distribuições multivariadas para respostas categóricas e discretas. No entanto a análise verossimilhança tem sido usada por muitos autores (Masarotto & Varin, 2012).

Em probabilidade e estatística, cópula é definida como uma distribuição de probabilidade multivariada em que a distribuição de probabilidade marginal de cada variável é conhecida, sendo utilizada para descrever a dependência entre variáveis aleatórias.

Os modelos de regressão marginais via cópulas Gaussianas constituem uma classe de modelos flexíveis e com estrutura geral para a análise de regressão com dados dependentes de qualquer tipo. As cópulas Gaussianas combinam a simplicidade de interpretação na modelagem marginal com a flexibilidade na especificação a estrutura de dependência. Apesar disso, a regressão via cópulas Gaussianas ainda tem um uso limitado, uma vez que, para as respostas dependentes não contínuas, a função de verossimilhança exige uma aproximação de integrais de grande dimensão.

Masarotto & Varin (2012) desenvolveram análise de verossimilhança para modelos de regressão marginais para respostas correlacionadas não-normais baseados em cópulas Gaussianas.

Seja $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ um vetor com n variáveis aleatórias correlacionadas contínuas, discretas ou categóricas e seja $\mathbf{y} = (y_1, \dots, y_n)^\top$ as realizações (valores observados) correspondentes. Considere que o objetivo é avaliar com a distribuição de Y_j varia de acordo com a mudança do vetor p -dimensional $\mathbf{x}_j = (x_{j1}, \dots, x_{jp})^\top$.

Denote por $p_j(y_j; \lambda) = p(y_j|x_j; \lambda)$ a função de Y_j dado x_j , de modo que as covariáveis podem afetar não apenas a média de Y_j , mas toda a distribuição marginal univariada de Y . Dentro deste contexto, a densidade de $p_j(y_j; \lambda)$ identifica o modelo de regressão.

Em termos mais gerais, o modelo de regressão é expresso como

$$Y_j = g(x_j, \epsilon; \lambda), \quad j = 1, \dots, n, \quad (4.29)$$

em que $g(\cdot)$ é uma função adequada dos regressores x_j e de uma variável estocástica ϵ_j , comumente chamada de componente de erro, e λ é um vetor de parâmetros desconhecidos. Existem várias especificações para $g(\cdot)$, sendo que uma escolha bastante útil é a função quantil

$$Y_j = F_j^{-1}(\Phi(\epsilon_j); \lambda), \quad j = 1, \dots, n, \quad (4.30)$$

em que ϵ_j é uma variável normal padrão e $F_j(\cdot; \lambda) = F(\cdot|x_j; \lambda)$ e $\Phi(\cdot)$ são as funções de distribuição de Y_j dado x_j e da variável normal padrão, respectivamente. Pelo teorema da transformação da integral, o modelo de regressão (4.30) garante a distribuição marginal desejada para a resposta Y_j e especifica ϵ_j nos termos de um erro normal. A especificação (4.30) inclui todos os possíveis modelos de regressão paramétricos para qualquer tipo de resposta (Masarotto & Varin, 2012).

Por exemplo, o modelo de regressão linear $Y_j = x_j^\top \beta + \sigma \epsilon_j$ corresponde o conjunto $F_j(Y_j; \lambda) = \Phi\{(Y_j - x_j^\top \beta)/\sigma\}$ em (4.30), com $\lambda = (\beta^\top, \sigma)^\top$. Para o modelo log-linear Poisson é obtido

$$F_j(Y_j; \lambda) = \sum_{k=0}^{Y_j} \frac{e^{-\mu_j} \mu_j^k}{k!}$$

em que $\mu_j = \exp(x_j^\top \beta)$, $\lambda \equiv \beta$.

Este modelo trata da análise de regressão na presença de dependência. Desta forma, a especificação do modelo é então completada assumindo que o vetor de erros $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_n)^\top$ segue distribuição é normal multivariada (NMV),

$$\epsilon \sim (NMV)(0, \Omega), \quad (4.31)$$

em que Ω é uma matriz de correlação (Song, 2000; Masarotto & Varin, 2012). O caso especial em que as observações são independentes a matriz de correlação é $\Omega = I_n$, ou seja a matriz identidade $n \times n$.

A especificação do modelo convenientemente separa a componente marginal (4.30) da componente dependente (4.31), o último sendo descritos em termo de um processo normal multivariado. Várias formas de dependência nos dados podem ser modelada através de uma padronização adequada da matriz de correlação Ω como uma função do vetor de parâmetros ς . O vetor de parâmetro é denotado por $\vartheta = (\lambda^\top, \varsigma^\top)^\top$, em que λ denota o vetor de parâmetros do modelo marginal e ς é o vetor de parâmetros dependentes (Masarotto & Varin, 2012).

A estrutura de dependência no modelo marginal via cópulas Gaussianas é modelada através da especificação de uma matriz de correlação apropriada Ω do vetor de erros ϵ .

Para dados longitudinais, suponha que as observações são agrupadas em m agrupamentos de dimensões $n_i, i = 1, \dots, m$, em que $\sum_{i=1}^m n_i = n$, ou seja, m indivíduos são observados em cada n_i ocasiões. Sob a suposição de independência entre os diferentes indivíduos ou grupos, matrizes de correlação adequadas para os erros são bloco diagonal,

$$\begin{bmatrix} \Omega_1 & 0 & \dots & 0 \\ 0 & \Omega_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Omega_m \end{bmatrix},$$

em que Ω_i é uma matriz de correlação $n_i \times n_i$. Similarmente ao método EEG, os blocos Ω_i podem ser especificados de diferentes maneiras, como:

- independente, $\text{cor}(\epsilon_{ij}, \epsilon_{ij'}) = 0$;
- permutável, $\text{cor}(\epsilon_{ij}, \epsilon_{ij'}) = \varsigma$;
- autorregressiva, $\text{cor}(\epsilon_{ij}, \epsilon_{ij'}) = \varsigma^{|j-j'|}$;
- Não estruturada, a matriz de correlação sem qualquer restrição.

Os modelos de regressão via cópulas Gaussianas difere de outros modelos marginais na forma de distribuições conjuntas bivariada e de dimensão de ordem superior.

No caso contínuo, a relação entre as respostas Y_i e erros ϵ_i produz a seguinte expressão de forma fechada para a verossimilhança de ϑ

$$L(\vartheta; y) = \prod_{j=1}^n p_i(y_i; \lambda) \frac{p(\epsilon_1, \dots, \epsilon_n; \vartheta)}{p(\epsilon_1; \lambda) \cdots p(\epsilon_n; \lambda)}. \quad (4.32)$$

Para os casos discretos e categóricos tem-se que as distribuições marginais conjuntas são expressas como integrais de normais multivariadas. Por exemplo, considerando duas observações, a distribuição é uma integral bidimensional dada por

$$p_{jj'}(y_j, y_{j'}; \vartheta) = \int_{\mathcal{D}_j(y_j; \lambda)} \int_{\mathcal{D}_{j'}(y_{j'}; \lambda)} p(\epsilon_j, \epsilon_{j'}; \vartheta) d\epsilon_j d\epsilon_{j'},$$

cujo domínio da integral acima é o produto cartesiano dos intervalos

$$\mathcal{D}_j(y_j; \lambda) = [\Phi^{-1}(F_j(y_j^-; \lambda)), \Phi^{-1}(F_j(y_j; \lambda))], \quad j = 1, \dots, n,$$

em que $F_j(y_j^-; \lambda)$ é o limite inferior de $F_j(\cdot; \lambda)$ em y_j , sendo $y_j^- = y_j - 1$. Maiores detalhes podem ser encontrados em Masarotto & Varin (2012).

No caso de variáveis respostas discreta ou categóricas, a verossimilhança é dada pelo cálculo da integral

$$L(\vartheta; y) = \int_{\mathcal{D}_1(y_1; \lambda)} \cdots \int_{\mathcal{D}_n(y_n; \lambda)} p(\epsilon_1, \dots, \epsilon_n; \vartheta) d\epsilon_1 \cdots d\epsilon_n. \quad (4.33)$$

É conveniente reescrever a integral acima considerando a mudança da variável $\mathcal{D}_1(y_1; \lambda) \times \cdots \times \mathcal{D}_n(y_n; \lambda)$ para $(0, 1)^n$ de tal forma que

$$\epsilon_j(u_j) = \Phi^{-1}\{F_j(y_j; \lambda) - u_j p_i(y_i; \lambda)\}, j = 1, \dots, n.$$

Então, a verossimilhança (4.33) assume a forma

$$L(\vartheta; y) = L_{ind}(\lambda; y) \int_{[0,1]} \cdots \int_{[0,1]} q\{\epsilon_1(u_1), \dots, \epsilon_n(u_n); \vartheta\} du_1 \cdots du_n. \quad (4.34)$$

Para maiores detalhes ver Masarotto & Varin (2012).

4.7 Programas Estatísticos

Nesta Seção são apresentados os comandos utilizados para os ajustes dos modelos multinível e marginal (EEG e via cópulas Gaussianas) para dados de contagem nos programas estatísticos **R** e **STATA**. Considere que y denota a variável resposta de contagem, x_1 , x_2 e x_3 são covariáveis, id representa a variável que identifica os indivíduos e $tempo$ denota o período em que as observações foram mensuradas.

4.7.1 R

Para o ajuste dos modelos Poisson multiníveis utiliza-se a mesma biblioteca apresentada na Seção 3.4, **lme4**. O comando utilizado é o **glmer** e no argumento **family** deve ser considerada a distribuição Poisson dada por **poisson**. As funções são definidas para os casos com apenas o intercepto aleatório e com intercepto e inclinação aleatórios, respectivamente, como

```
mod1 = glmer(y ~ x1 + x2 + x3 + tempo + (1|id), data=dados,  
family=poisson)
```

```
mod2 = glmer(y ~ x1 + x2 + x3 + tempo + (tempo | id), data= dados,  
family=poisson).
```

Os modelos Binomial Negativo multiníveis são ajustados utilizando a biblioteca **glmmADMB** através da função **glmmadmb** e os comandos são mostrados a seguir para os casos com apenas o intercepto aleatório e com intercepto e inclinação aleatórios:

```
mod1 = glmmadmb(y ~ x1 + x2 + x3 + tempo + (1|id), data=dados,  
family="nbinom")
```

```
mod2 = glmmadmb(y ~ x1 +x2 + x3 + tempo + (tempo|id), data=dados,  
family="nbinom").
```

Para ajustar o modelo multinível com intercepto aleatório é necessário considerar o argumento $(1 | id)$. O modelo multinível com efeitos aleatórios no intercepto e tempo é identificado através do argumento $(tempo | id)$. O argumento **family** deve conter a família Binomial Negativa dada por **nbinom**. Maiores detalhes podem ser encontrados em <http://glmmadmb.r-forge.r-project.org/>.

Para ajustar os dados utilizando a abordagem de equações de estimação generalizadas (EEG) considerando a distribuição Poisson, deve-se utilizar a biblioteca **geepack** usando a função **geeglm**. A função a ser implementada é:

```
mod1 = geeglm(y ~ x1 + x2 + x3 + tempo, data = dados, family =
poisson, id = id, corstr = " ").
```

O argumento `family` deve conter a distribuição Poisson dada por `poisson` e as matrizes de correlação de trabalho são identificadas no argumento `corstr` e dadas por

- `exchangeable` que denota a matriz de correlação de “trabalho” permutável;
- `ar1` que representa a matriz de correlação de “trabalho” autorregressiva de primeira ordem;
- `unstructured` que define a matriz de correlação de “trabalho” não estruturada.

A função `geeglm` não tem suporte para ajustar o modelo Binomial Negativo. Uma biblioteca que ajusta os dados segundo esta distribuição é a `gcmr`. Esta biblioteca utiliza cópula normal para ajustar modelos marginais descritos por Song (2000) e Masarotto & Varin (2012) que permitem modelar respostas dependentes de qualquer tipo. Além disso a interpretação na modelagem marginal é simplificada quando se utilizam as cópulas Gaussianas. O comando utilizado é

```
mod1 = gcmr(y ~ x1 + x2 + x3 + tempo, data = dados, marginal =
negbin.marg, id = id, cormat = cluster.cormat(id, type = )).
```

A família da distribuição é identificada no objeto `marginal` e a matriz de correlação é informada no objeto `type`, dentro de `cormat`, cujas estruturas podem ser:

- `exchangeable` que denota a matriz de correlação de “trabalho” permutável;
- `ar1` que representa a matriz de correlação de “trabalho” autorregressiva de primeira ordem;
- `independence` que indica a matriz de correlação de “trabalho” independente.
- `unstructured` que define a matriz de correlação de “trabalho” não estruturada.

O modelo Poisson marginal via cópulas gaussianas também é ajustado nesta biblioteca, sendo modificado apenas a família para `poisson.marg`. Através dos exemplos apresentados no manual, verificou-se que os resultados das estimativas usando a biblioteca `geepack` são similares aos do `gcmr`. Porém, o critério de seleção

utilizado nas bibliotecas é distinto, uma vez que a **geepack** considera o QIC e a **gcmr** utiliza o AIC.

Em <http://cran.r-project.org/web/packages/gcmr/gcmr.pdf> é possível encontrar maiores detalhes sobre o pacote **gcmr**.

4.7.2 STATA

É possível ajustar os modelos Poisson e BN multiníveis utilizando os comandos **xtpoisson** e **xtnbreg**. Como exemplo tem-se:

```
xtpoisson y x1 x2 x3 tempo, nolog re i(id)
```

```
xtpoisson y x1 x2 x3 tempo, nolog normal re i(id)
```

```
xtnbreg y x1 x2 x3 tempo, nolog re i(id)
```

A distribuição do efeito aleatório para o modelo Poisson multinível considerada como *default* é a Gama, mas é possível utilizar a distribuição normal para o efeito aleatório usando o argumento **normal** (**xtpoisson**). Para o modelo Binomial Negativo multinível, a distribuição do efeito aleatório no intercepto que se encontra implementada é beta (**xtnbreg**).

Em relação à abordagem EEG, as funções utilizadas para ajustar os modelos Poisson e Binomial Negativo são, respectivamente, **xtgee** e **xtnbreg** e os comandos são dados por

```
xtgee y x1 x2 x3 tempo, nolog fam(poisson) corr( ) i(id)
```

```
xtnbreg y x1 x2 x3 tempo, nolog pa corr( ) i(id)
```

O objeto **fam** no comando **xtgee** indica a família de distribuição que está sendo ajustada. No caso da primeira função, a distribuição Poisson é considerada. O objeto **pa** no comando **xtnbreg** significa que foi utilizado o estimador para a média populacional. As matrizes de correlação de “trabalho” são definidas no objeto **corr** e são dadas por

- **exch** que denota a matriz de correlação de “trabalho” permutável;
- **ar1** que representa a matriz de correlação de “trabalho” autorregressiva de primeira ordem;
- **ind** que indica a matriz de correlação de “trabalho” independente;
- **unstr** que define a matriz de correlação de “trabalho” não estruturada.

Capítulo 5

Estudos de Simulação

Neste capítulo são apresentados alguns estudos de simulação que consideram diferentes cenários para os modelos Poisson e Binomial Negativo a fim de avaliar o desempenho do processo de estimação nas abordagens condicional e marginal na análise de dados longitudinais.

Inicialmente os dados de contagem correlacionados foram gerados e o método de geração de dados é apresentado na Seção 5.1. Os diferentes cenários de simulação para as abordagens condicional e marginal são descritos na Seção 5.2. Na Seção 5.3 são apresentados os critérios de avaliação dos estimadores. Para a realização das simulações utilizou-se o programa estatístico **R** versão 3.0.2 para windows e as bibliotecas utilizadas são apresentadas na Seção 5.4. Os resultados são sumarizados e discutidos na Seção 5.5.

5.1 Geração de Variáveis Aleatórias Poisson e Binomial Negativa Correlacionadas

Yahav e Shmueli (2008) propuseram um método computacionalmente eficiente e conceitualmente atraente para gerar variáveis aleatórias (v.a.'s) Poisson multivariadas. A ideia básica é, primeiramente, gerar um vetor de dimensão $p \times 1$ de variáveis que tem distribuição normal multivariada e então transformá-lo em um vetor de variáveis com distribuição Poisson usando as funções de distribuição acumulada (f.d.a.) Gaussiana e Poisson.

O algoritmo pode ser sumarizado da seguinte forma:

1. Gere um vetor Normal p -dimensional $\mathbf{X}^\top = (X_1, \dots, X_p)$ com vetor de médias $\boldsymbol{\mu} = 0$, vetor de variância com elementos iguais a 1 e uma matriz de correlação R .

2. Para cada valor de X_k , $k = 1, 2, \dots, p$, calcule a f.d.a. Gaussiana: $\Phi(X_k)$.
3. Para cada $\Phi(X_k)$ calcule a f.d.a. inversa da Poisson definida por Ξ^{-1} , também conhecida como função quantílica, com razão μ dada por:

$$Y_k = \Xi^{-1}(\Phi(X_k))$$

O vetor $\mathbf{Y}^\top = (Y_1, \dots, Y_p)$ é um vetor de dimensão $p \times 1$ com matriz de correlação R e vetor de médias $\mathbf{\Lambda} = (\mu_1, \dots, \mu_p)^\top$.

Para a geração das v.a.'s com distribuição Binomial Negativa propomos uma metodologia baseada na proposta de Yahav e Shmueli, em que substituímos a f.d.a. inversa da Poisson pela f.d.a. da Binomial Negativa. O algoritmo é idêntico ao apresentado anteriormente, mas a função quantílica da Poisson Ξ^{-1} é trocada pela função quantílica da Binomial Negativa definida por Υ^{-1} . Desta forma

$$\tilde{Y}_k = \Upsilon^{-1}(\Phi(X_k)),$$

em que $\tilde{\mathbf{Y}}^\top = (\tilde{Y}_1, \dots, \tilde{Y}_p)$ é um vetor de dimensão $p \times 1$ com matriz de correlação R e vetor de médias $\mathbf{\Psi} = (\tilde{\mu}_1, \dots, \tilde{\mu}_p)^\top$.

Vale ressaltar que aqui propomos um método mais simples de geração de dados correlacionados usando a distribuição BN. Contudo, outras abordagens encontram-se disponíveis na literatura, como, por exemplo, a geração de BN multivariadas usando cópulas (Lee, 2009) ou geração de BN multivariadas usando a distribuição exponencial multivariada (Kopocinski, 2009).

5.2 Cenários de Simulação

Nesta seção são apresentados diferentes cenários de simulação pelo método de Monte Carlo a fim de avaliar a modelagem de regressão para dados de contagem longitudinais. Foram utilizadas 2000 réplicas para as abordagens considerando a distribuição Poisson e apenas 500 réplicas para as abordagens considerando a distribuição BN, por causa das dificuldades computacionais encontradas para esta distribuição (tempo de execução do algoritmo e sensibilidade dos resultados). Foram utilizados diferentes números de indivíduos ($m = 50; 150; 250$) e diferentes números de medidas ($n_i = 3; 5; 8; 10$). A variável explicativa X_i^* foi gerada considerando uma distribuição normal com média 0 e variância 0.2, ou seja, $X_i^* \sim \mathcal{N}(0, 0.2)$. Em todas as simulações foram gerados dados balanceados e igualmente espaçados.

5.2.1 Geração: Modelos Multiníveis

Para o modelo Poisson multinível foram considerados dois cenários, que são descritos a seguir. Para o modelo Binomial Negativo multinível foi considerado apenas o primeiro destes cenários, pois quando os dados foram gerados considerando o cenário mais complexo, os algoritmos de ajuste dos modelos apresentaram problemas de convergência.

Para as v.a.'s geradas considerando a distribuição Poisson, em ambos cenários, foram ajustados os seguintes modelos: (i) **Modelo 1**: modelo Poisson independente, (ii) **Modelo 2**: modelo Poisson multinível com intercepto aleatório e (iii) **Modelo 3**: modelo Poisson multinível com dois efeitos aleatórios.

Para as v.a.'s geradas considerando a distribuição Binomial Negativa foram ajustados os seguintes modelos: (i) **Modelo 1**: modelo Poisson independente, (ii) **Modelo 2**: modelo Poisson multinível com intercepto aleatório, (iii) **Modelo 3**: modelo Poisson multinível com dois efeitos aleatórios, (iv) **Modelo 4**: modelo Binomial Negativo independente, (v) **Modelo 5**: modelo Binomial Negativo multinível com intercepto aleatório e (vi) **Modelo 6**: modelo Binomial Negativo multinível com dois efeitos aleatórios.

5.2.1.1 Cenário 1: Modelo Multinível com Efeito Aleatório no Intercepto

Os dados foram gerados considerando apenas o intercepto aleatório, ou seja, o modelo multinível foi dado por

$$\mu_{ij} = \exp(\beta_1 + \beta_2 X_i^* + \beta_3 t_{ij} + b_{0i}), \quad (5.1)$$

em que $\beta^\top = (\beta_1, \beta_2, \beta_3) = (0, 3; 2, 3; 1, 1)$ é o vetor de parâmetros e $t = 1, \dots, T$ denota o tempo em que as medidas foram obtidas ($T = 3; 5; 8; 10$). Como os dados são igualmente espaçados tem-se que o número de medidas repetidas é igual ao último ponto em que as medidas foram mensuradas, ou seja, $T = n_i$. O efeito aleatório do intercepto é representado por b_{0i} que é distribuído normalmente com média 0 e variância τ_0 , isto é, $b_{0i} \sim \mathcal{N}(0, \tau_0)$, sendo $\tau_0 = (0, 1; 0, 5; 0, 7; 1, 0)$. Além disso, tem-se que b_{0i} e X_i^* são independentes.

5.2.1.2 Cenário 2: Modelo Multinível com Dois Efeitos Aleatórios

Neste cenário os dados foram gerados considerando dois efeitos aleatórios, isto é, o modelo multinível foi

$$\mu_{ij} = \exp(\beta_1 + \beta_2 X_i^* + \beta_3 t_{ij} + b_{0i} + b_{1i} t_{ij}), \quad (5.2)$$

em que $\beta^\top = (\beta_1, \beta_2, \beta_3) = (0, 3; 2, 3; 1, 1)$ é o vetor de parâmetros e t denota o tempo em que as medidas foram obtidas, em que $t = 1, \dots, T$ ($T = 3; 5; 8; 10$). Novamente os dados são igualmente espaçados, portanto o número de medidas repetidas é igual ao último ponto em que as medidas foram mensuradas, ou seja, $T = n_i$. Os efeitos aleatórios do intercepto e do tempo, representados por b_{0i} e b_{1i} , respectivamente, seguem distribuição normal bivariada com vetor de médias igual a $\begin{bmatrix} 0 & 0 \end{bmatrix}^\top$ e matriz de covariância $\begin{bmatrix} \tau_0 & \tau_{01} \\ \tau_{01} & \tau_1 \end{bmatrix}$, em que $\tau_0 = (0, 1; 0, 5; 0, 7; 1, 0)$, $\tau_1 = 0, 3$ e $\tau_{01} = 0, 01$.

5.2.2 Geração: Modelos Poisson Marginais (EEG)

Como discutido no Capítulo 4, Seção 4.3, os modelos marginais ajustam diretamente a resposta média em cada ocasião. Para a geração dos dados, utilizou-se um modelo da forma

$$\mu_{ij} = \exp(\beta_1 + \beta_2 X_i^* + \beta_3 t_{ij}),$$

em que $\beta^\top = (\beta_1, \beta_2, \beta_3) = (0, 3; 2, 3; 1, 1)$ é o vetor de parâmetros e t denota o tempo em que as medidas foram obtidas, em que $t_{ij} = 1, \dots, T$ ($T = 3; 5; 8; 10$). Como foram gerados dados balanceados, todos os indivíduos possuem o mesmo número de medidas repetidas. As matrizes de correlação, R , utilizadas para a geração dos dados foram dadas pelas matrizes de correlação de “trabalho” na forma autorregressiva de ordem 1 (AR1), permutável e não estruturada (NE) apresentadas no Capítulo 3, Subseção 3.1.2, sendo aqui o parâmetro de associação, α , igual a 0,5. As matrizes de correlação de “trabalho” não estruturada considerando 3, 5, 8 e 10 medidas repetidas foram, respectivamente, definidas como sendo (unicamente apresentamos o bloco superior, uma vez que $R(\alpha)$ é simétrica)

$$R(\alpha) = \begin{bmatrix} 1 & 0,9\alpha & 0,7\alpha \\ & 1 & 0,5\alpha \\ & & 1 \end{bmatrix}_{n=3},$$

$$R(\alpha) = \begin{bmatrix} 1 & 0,9\alpha & 0,7\alpha & 0,4\alpha & 0,8\alpha \\ & 1 & 0,5\alpha & 0,1\alpha & 0,3\alpha \\ & & 1 & 0,2\alpha & 0,4\alpha \\ & & & 1 & 0,4\alpha \\ & & & & 1 \end{bmatrix}_{n=5},$$

$$\begin{aligned}
R(\alpha) &= \begin{bmatrix} 1 & 0,9\alpha & 0,7\alpha & 0,4\alpha & 0,8\alpha & 0,6\alpha & 0,5\alpha & 0,3\alpha \\ & 1 & 0,5\alpha & 0,1\alpha & 0,3\alpha & 0,2\alpha & 0,2\alpha & 0,1\alpha \\ & & 1 & 0,2\alpha & 0,4\alpha & 0,1\alpha & 0,7\alpha & 0,4\alpha \\ & & & 1 & 0,4\alpha & 0,8\alpha & 0,5\alpha & 0,7\alpha \\ & & & & 1 & 0,7\alpha & 0,6\alpha & 0,5\alpha \\ & & & & & 1 & 0,5\alpha & 0,9\alpha \\ & & & & & & 1 & 0,4\alpha \\ & & & & & & & 1 \end{bmatrix}_{n=8} \quad \text{e} \\
R(\alpha) &= \begin{bmatrix} 1 & 0,9\alpha & 0,7\alpha & 0,4\alpha & 0,8\alpha & 0,6\alpha & 0,5\alpha & 0,3\alpha & 0,4\alpha & 0,2\alpha \\ & 1 & 0,5\alpha & 0,1\alpha & 0,3\alpha & 0,2\alpha & 0,2\alpha & 0,1\alpha & 0,8\alpha & 0,4\alpha \\ & & 1 & 0,2\alpha & 0,4\alpha & 0,1\alpha & 0,7\alpha & 0,4\alpha & 0,9\alpha & 0,3\alpha \\ & & & 1 & 0,4\alpha & 0,8\alpha & 0,5\alpha & 0,7\alpha & 0,2\alpha & 0,1\alpha \\ & & & & 1 & 0,7\alpha & 0,6\alpha & 0,5\alpha & 0,3\alpha & 0,2\alpha \\ & & & & & 1 & 0,5\alpha & 0,9\alpha & 0,4\alpha & 0,8\alpha \\ & & & & & & 1 & 0,4\alpha & 0,6\alpha & 0,9\alpha \\ & & & & & & & 1 & 0,2\alpha & 0,1\alpha \\ & & & & & & & & 1 & 0,6\alpha \\ & & & & & & & & & 1 \end{bmatrix}_{n=10}.
\end{aligned}$$

Três diferentes cenários foram considerados na abordagem marginal, considerando EEG, e os mesmos são descritos a seguir:

1. **Cenário 1:** geração das v.a.'s correlacionadas utilizando a matriz de correlação autorregressiva de ordem 1 (AR1);
2. **Cenário 2:** geração das v.a.'s correlacionadas utilizando a matriz de correlação não estruturada (NE);
3. **Cenário 3:** geração das v.a.'s correlacionadas utilizando a matriz de correlação permutável.

Para o modelo Poisson marginal utilizando a abordagem EEG, em todos os cenários, foram ajustados quatro modelos: (i) **Modelo 1:** modelo Poisson marginal considerando a abordagem EEG com estrutura de correlação de “trabalho” autorregressiva de ordem 1; (ii) **Modelo 2:** modelo Poisson marginal considerando a abordagem EEG com estrutura de correlação de “trabalho” permutável; (iii) **Modelo 3:** modelo Poisson marginal considerando a abordagem EEG com estrutura de correlação de “trabalho” independente e (iv) **Modelo 4:** modelo Poisson marginal considerando a abordagem EEG com estrutura de correlação de “trabalho” não estruturada.

5.2.3 Geração: Modelos BN Marginais (Cópulas Gaussianas)

Por dificuldades computacionais não foi possível ajustar o modelo Binomial Negativo (BN) marginal utilizando a abordagem EEG. Então, utilizou-se o modelo marginal via cópulas Gaussianas (CG) apresentado no Capítulo 4, Seção 4.6.

O modelo ajustado, as matrizes de correlação $R(\alpha)$ e os cenários de simulação apresentados na seção anterior foram os mesmos utilizados para o modelo BN marginal via cópulas Gaussianas e para todos os cenários foram ajustados oito modelos: (i) **Modelo 1**: modelo Poisson marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” autorregressiva de ordem 1; (ii) **Modelo 2**: modelo Poisson marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” permutável; (iii) **Modelo 3**: modelo Poisson marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” independente, (iv) **Modelo 4**: modelo Poisson marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” não estruturada; (v) **Modelo 5**: modelo BN marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” autorregressiva de ordem 1; (vi) **Modelo 6**: modelo BN marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” permutável; (vii) **Modelo 7**: modelo BN marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” independente e (viii) **Modelo 8**: modelo BN marginal via cópulas Gaussianas com estrutura de correlação de “trabalho” não estruturada.

5.3 Critérios de Avaliação dos Estimadores

Para avaliar o desempenho do ajuste dos modelos, foram calculados a média das estimativas ($\widehat{\beta}$), o viés médio do estimador (VM), o valor absoluto do viés relativo estimado ($|VR|$), a diferença relativa estimada ($diff.rel$) entre o desvio padrão (DP) e erro padrão (EP) estimados e a raiz do erro quadratico médio estimado ($REQM$).

A média das estimativas e o viés médio são dados, respectivamente, por

$$\widehat{\beta} = \frac{1}{N} \sum_{l=1}^N \widehat{\beta}_l \quad \text{e} \quad VM = \frac{1}{N} \sum_{l=1}^N (\beta - \widehat{\beta}_l),$$

em que N denota o número de réplicas de Monte Carlo.

O valor absoluto do viés relativo estimado ($|VR|$) é dado por:

$$|VR| = \left| \frac{\widehat{\beta} - \beta}{\beta} \right|$$

em que $\widehat{\beta}$ é a média das estimativas dos parâmetros (Cribari-Neto e Gois, 2002).

A diferença relativa estimada entre o desvio padrão e o erro padrão estimados é dada por:

$$dif.rel(\widehat{\beta}) = \left| \frac{DP(\widehat{\beta}) - EP(\widehat{\beta})}{EP(\widehat{\beta})} \right|, \quad (5.3)$$

em que $DP(\widehat{\beta}) = \sqrt{\text{Var}(\widehat{\beta})}$ é o desvio padrão do estimador e $EP(\widehat{\beta}) = \sqrt{\widehat{\text{Var}}(\widehat{\beta})}$ denota o erro padrão do estimador.

O estimador da variância tem boa performance quando o erro padrão médio das estimativas e o desvio padrão de $\widehat{\beta}$ são próximos, ou seja, quando a diferença relativa entre o desvio padrão e o erro padrão está próxima de zero.

A raiz do erro quadrático médio estimada é definida por

$$REQM(\widehat{\beta}) = \sqrt{E(\widehat{\beta} - \beta)^2} = \sqrt{(\text{Var}(\widehat{\beta}) + [VM]^2)},$$

em que $\widehat{\beta}$ refere-se ao valor estimado pelo modelo, β é o parâmetro e $\text{Var}(\widehat{\beta}) = DP(\widehat{\beta})^2$.

5.4 Suporte Computacional

O programa utilizado para a realização das simulações foi o **R** versão 3.0.2. Nesta Seção são apresentadas as bibliotecas utilizadas para cada abordagem considerando os modelos Poisson e Binomial Negativo.

5.4.1 Geração de Variáveis Aleatórias Correlacionadas

Para a geração das variáveis aleatórias (v.a's) com distribuição Normal Multivariada utilizou-se a função `mvrnorm` do pacote **MASS**.

As funções `pnorm`, `qpois` e `qnbinom` calculam a função de distribuição acumulada (f.d.a.) Normal, a f.d.a. inversa da distribuição Poisson e a f.d.a. inversa da distribuição Binomial Negativa (BN), respectivamente. No caso da distribuição BN o parâmetro de dispersão, definido por `size`, foi igual a 1.

A matriz de covariância foi determinada a partir da matriz de correlação R utilizando o comando `cor2cov` do pacote **MBESS**.

5.4.2 Modelo Poisson Multinível

Para o ajuste do modelo Poisson multinível utilizou-se a biblioteca **lme4** e o comando usado foi o `glmer`, em que o argumento `family` foi definido como sendo `poisson`.

5.4.3 Modelo Poisson Marginal (EEG)

Para ajustar os modelo Poisson segundo a abordagem de equações de estimação generalizadas (EEG) utilizou-se a biblioteca **geepack** usando a função **geeglm**.

5.4.4 Modelo Binomial Negativo Multinível

O modelo Binomial Negativo multinível foi ajustado utilizando a biblioteca **glmmADMB** através da função **glmmadmb**.

5.4.5 Modelo Binomial Negativo Marginal (Cópulas Gaussianas)

O pacote **geepack** não tem suporte para ajustar o modelo Binomial Negativo marginal utilizando a abordagem EEG. A biblioteca utilizada para ajustar o modelo BN marginal via cópulas Gaussianas foi **gcmr**, sendo a função dada por **gcmr**. O argumento **marginal** define qual distribuição está sendo utilizada para ajustar os dados. No caso da BN utiliza-se **negbin.marg**.

O modelo Poisson também é ajustado nesta biblioteca, sendo modificada apenas a família (**marginal**) para **poisson.marg**.

5.4.6 Problemas de Convergência

Durante a realização das simulações, os modelos considerando a distribuição Binomial Negativa (BN) tiveram problemas de convergência e um custo computacional muito elevado. Em relação ao tempo de execução de cada simulação, quando foram consideradas 2000 réplicas as simulações para cenários mais simples demoraram aproximadamente 2 semanas. Por este motivo foi necessário a diminuição do número de réplicas para 500 réplicas.

Para o cenário 1, em que os dados são gerados a partir do modelo BN multinível com intercepto aleatório, o custo computacional foi muito intenso, especialmente com o aumento do número de medidas repetidas. Considerando, por exemplo, 50 indivíduos com 8 medidas repetidas apenas 3 das 500 réplicas (0,6%) não convergiram. Em relação ao tempo de simulação, foi necessário 66,31 horas, aproximadamente 3 dias, para a obtenção dos resultados. Não houve convergência para os modelos Poisson e BN quando se considerou o modelo BN multinível com dois efeitos aleatórios.

Não foi possível ajustar o modelo BN marginal considerando a abordagem EEG, pois não havia suporte computacional no **R**. Por este motivo utilizou-se a abordagem das cópulas Gaussianas para ajustar o modelo BN marginal. Como houve um custo

computacional muito alto em relação ao tempo de execução de cada simulação, foi necessário a diminuição do número de réplicas.

Além disso, os ajustes apresentaram problemas de convergência para os modelos BN marginais via cópulas Gaussianas. Considerando 8 e 10 medidas repetidas, os modelos não convergiram, para 50, 150 e 250 indivíduos. Para 3 e 5 medidas repetidas o número de não convergência foi muito alto. Por exemplo, para 5 medidas repetidas e considerando 50 e 150 indivíduos, 188 (37,6%) e 333 (66,6%) das 500 réplicas, respectivamente, não convergiram.

5.5 Resultados

5.5.1 Geração: Modelo Poisson Multinível com Efeito Aleatório no Intercepto

A Tabela 5.1 apresenta os valores médios das estimativas ($\widehat{\beta}$), os vieses médios (VM) das estimativas e os valores da raiz do erro quadrático médio ($REQM$) para os 2000 conjuntos de dados simulados considerando 50 indivíduos. Os modelos ajustados foram: Modelo Poisson independente (**Modelo 1**), Modelo Poisson multinível com efeito aleatório no intercepto (**Modelo 2**) e Modelo Poisson multinível com dois efeitos aleatórios (**Modelo 3**).

Pode-se verificar, de modo geral, que os valores da raiz do erro quadrático médio diminuíram com o aumento do número de medidas repetidas, exceto para o **Modelo 1** considerando 8 e 10 medidas repetidas, o estimador $\widehat{\beta}_1$ e $\tau_0 = 1, 0$. Em relação ao viés médio (VM) das estimativas é notável a diminuição dos valores com o aumento do número de medidas repetidas, exceto para 8 e 10 medidas, o **Modelo 1** e o estimador $\widehat{\beta}_1$.

É possível, através dos resultados apresentados na Tabela 5.1, comparar o impacto da variância do efeito aleatório associado ao intercepto, denotado por τ_0 . Verificou-se que a variância do efeito aleatório proporciona, de modo geral, um aumento no viés médio dos estimadores. Considerando o modelo Poisson independente (**Modelo 1**) é possível notar claramente este aumento, especialmente no estimador do intercepto $\widehat{\beta}_1$. Por exemplo, para 10 medidas, os valores do viés são 0,006 e 0,525 para $\tau_0 = 0, 1$ e $\tau_0 = 1, 0$, respectivamente.

Para o estimador $\widehat{\beta}_2$, o impacto da variância do efeito aleatório τ_0 é maior para os modelos Poisson multinível com efeito aleatório no intercepto (**Modelo 2**) e com dois efeitos aleatórios (**Modelo 3**). Os valores do viés para $\widehat{\beta}_2$ no **Modelo 2**, considerando-se 3, 5, 8 e 10 medidas são, respectivamente, 0,065, 0,066, 0,057 e

0,056 para $\tau_0 = 0,1$; e 0,595, 0,569, 0,561 e 0,561 para $\tau_0 = 1,0$. Nota-se também para $\hat{\beta}_1$ que o **Modelo 2** apresentou os menores valores para o viés quando $\tau_0 = 1,0$, sendo -0,029, -0,017, -0,011 e -0,011, respectivamente, para 3, 5, 8 e 10 medidas.

Para o estimador $\hat{\beta}_3$, o impacto de τ_0 é contrário, isto é, com o aumento da variância do intercepto aleatório houve uma diminuição dos valores do viés médio (*VM*) e da raiz do erro quadrático médio (*REQM*) para todos os modelos ajustados. Além disso, pode-se notar que os valores da média, do viés e da raiz do erro quadrático médio do estimador $\hat{\beta}_3$ para os **Modelos 1 e 2** são idênticos.

Em relação à raiz do erro quadrático médio (*REQM*) é possível verificar que há uma diminuição dos seus valores quando se aumenta o número de medidas repetidas, exceto para o estimador de β_1 do **Modelo 1**, considerando $\tau_0 = 1,0$ (Tabela 5.1). Os valores da raiz do erro quadrático médio considerando esta configuração são 0,511, 0,520, 0,524 e 0,525 para 3, 5, 8 e 10 medidas, respectivamente.

Verifica-se, novamente, o aumento substancial dos valores da raiz do erro quadrático médio quando se aumenta o valor da variância do efeito aleatório. Considere, como exemplo, o **Modelo 2** para o estimador $\hat{\beta}_2$, em que os valores da raiz do erro quadrático médio são 0,069, 0,067, 0,057 e 0,056 para $\tau_0 = 0,1$; e 0,598, 0,569, 0,561 e 0,561 para $\tau_0 = 1,0$.

Desta forma, pode-se concluir que a variância do intercepto aleatório tem uma forte influência no comportamento dos estimadores. Além disso, o tipo da variável modifica o resultado deste impacto. Para a covariável tempo-independente (*X*) há um aumento e para a variável tempo-dependente (*tempo*) há uma diminuição nos valores do *VM* e da *REQM*.

Com o aumento do número de observações, resultados similares aos apresentados foram encontrados. As Tabelas B.1 e B.2 apresentam os resultados para 150 e 250 indivíduos, respectivamente, e encontram-se no Apêndice B.1. Também foram considerados diferentes valores para a variância dos efeitos aleatórios e comportamentos similares foram encontrados. As Tabelas B.3, B.4 e B.5 apresentam os valores médios das estimativas, os vieses médios e os valores da *REQM* das estimativas dos parâmetros dos **Modelos 1, 2 e 3** com $\tau_0 = 0,5$ e $\tau_0 = 0,7$ para 50, 150 e 250 indivíduos, respectivamente.

Tabela 5.1: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\tau_0 = 0, 1$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,282309	-0,017691	0,036849	0,279339	-0,020661	0,038432	0,278349	-0,021651	0,039039
	5	0,298653	-0,001347	0,017010	0,289656	-0,010344	0,020085	0,290834	-0,009166	0,019974
	8	0,305352	0,005352	0,007798	0,297472	-0,002528	0,006234	0,297537	-0,002463	0,006397
	10	0,306547	0,006547	0,007024	0,298813	-0,001187	0,002819	0,298795	-0,001205	0,002840
$\widehat{\beta}_2$	3	2,359484	0,059484	0,063936	2,364854	0,064854	0,069301	2,364557	0,064557	0,068920
	5	2,349951	0,049951	0,050706	2,366071	0,066071	0,067274	2,365908	0,065908	0,067395
	8	2,348287	0,048287	0,048316	2,357013	0,057013	0,057111	2,356712	0,056712	0,056849
	10	2,348185	0,048185	0,048188	2,356192	0,056192	0,056202	2,356075	0,056075	0,056120
$\widehat{\beta}_3$	3	1,105754	0,005754	0,011042	1,105754	0,005754	0,011042	1,106117	0,006117	0,011259
	5	1,101455	0,001455	0,003393	1,101455	0,001455	0,003393	1,101261	0,001261	0,003362
	8	1,100158	0,000158	0,000701	1,100158	0,000158	0,000701	1,100152	0,000152	0,000726
	10	1,100010	0,000010	0,000247	1,100010	0,000010	0,000247	1,100013	0,000013	0,000248
$\tau_0 = 1, 0$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,810478	0,510478	0,511013	0,270617	-0,029383	0,043226	0,261950	-0,038050	0,052712
	5	0,820099	0,520099	0,520248	0,282533	-0,017467	0,023016	0,279146	-0,020854	0,027404
	8	0,823915	0,523915	0,523932	0,288378	-0,011622	0,012453	0,287852	-0,012148	0,013344
	10	0,824616	0,524616	0,524619	0,289273	-0,010727	0,010908	0,289218	-0,010782	0,011097
$\widehat{\beta}_2$	3	2,485371	0,185371	0,186101	2,895480	0,595480	0,598073	2,891768	0,591768	0,594395
	5	2,479504	0,179504	0,179606	2,868891	0,568891	0,569292	2,865665	0,565665	0,566210
	8	2,478514	0,178514	0,178518	2,861243	0,561243	0,561261	2,860700	0,560700	0,560952
	10	2,478453	0,178453	0,178453	2,860658	0,560658	0,560660	2,860038	0,560038	0,560328
$\widehat{\beta}_3$	3	1,103267	0,003267	0,007618	1,103267	0,003267	0,007618	1,106654	0,006654	0,011310
	5	1,100821	0,000821	0,002415	1,100820	0,000820	0,002415	1,101609	0,001609	0,003397
	8	1,100089	0,000089	0,000517	1,100089	0,000089	0,000517	1,100163	0,000163	0,000683
	10	1,100003	0,000003	0,000184	1,100003	0,000003	0,000184	1,100012	0,000012	0,000239

A Figura 5.1 apresenta o valor absoluto do viés relativo ($|VR|$), considerando a variância do efeito aleatório igual a 0,1 ($\tau_0 = 0,1$), para os modelos Poisson independente (**Modelo 1**), Poisson multinível com intercepto aleatório (**Modelo 2**) e Poisson multinível com dois efeitos aleatórios (**Modelo 3**).

É possível notar que as estimativas diminuem quando se aumenta o número de medidas repetidas para os estimadores $\hat{\beta}_2$ e $\hat{\beta}_3$, o que implica que o estimador está se aproximando do verdadeiro valor do parâmetro. Em relação ao estimador $\hat{\beta}_1$ é possível verificar que o valor absoluto do viés relativo diminui com aumento do número de medidas para os **Modelos 2 e 3** quando se consideram 50 e 150 indivíduos. O mesmo não é verificado para o **Modelo 1**, em que para 50 indivíduos os valores absolutos do viés relativo são 0,059, 0,004, 0,0178 e 0,022 para 3, 5, 8 e 10 medidas, respectivamente.

Verifica-se também na Figura 5.1 que o **Modelo 1** tem os valores absolutos do viés relativo menores do que os **Modelos 2 e 3** para o estimador de β_2 . Em relação ao estimador $\hat{\beta}_3$, observa-se que os três modelos apresentam um comportamento similar.

Também foi avaliado o comportamento do valor absoluto do viés relativo para os diferentes valores da variância do efeito aleatório e os resultados encontram-se disponíveis no Apêndice B.2, mostrando um comportamento similar ao analisado.

Para avaliar o comportamento do estimador da variância foram calculadas os valores da diferença relativa entre o desvio padrão (DP) e o erro padrão (EP) dos estimadores para os modelos Poisson independente (**Modelo 1**), Poisson multinível com intercepto aleatório (**Modelo 2**) e Poisson multinível com dois efeitos aleatórios (**Modelo 3**) e os resultados encontram-se sumarizados na Figura 5.2.

Considerando o estimador de β_1 é possível notar que os valores da diferença relativa entre o DP e o EP para os 3 modelos são bem próximos quando se considera 3 medidas repetidas, independente do número de indivíduos. Por exemplo, para 50 indivíduos tem-se que os valores são 0,476, 0,481 e 0,484 para os **Modelos 1, 2 e 3**, respectivamente. Porém, com o aumento do número de medidas isto não é mais observado, uma vez que os valores para a diferença relativa diminuem quando se considera o **Modelo 1** e aumentam para os **Modelos 2 e 3** (Figura 5.2).

Em relação à diferença relativa para o estimador $\hat{\beta}_2$, é possível notar que o **Modelo 1** apresenta as menores diferenças relativas. Além disso elas diminuem com o aumento do número de medidas, visualizado na Figura 5.2.

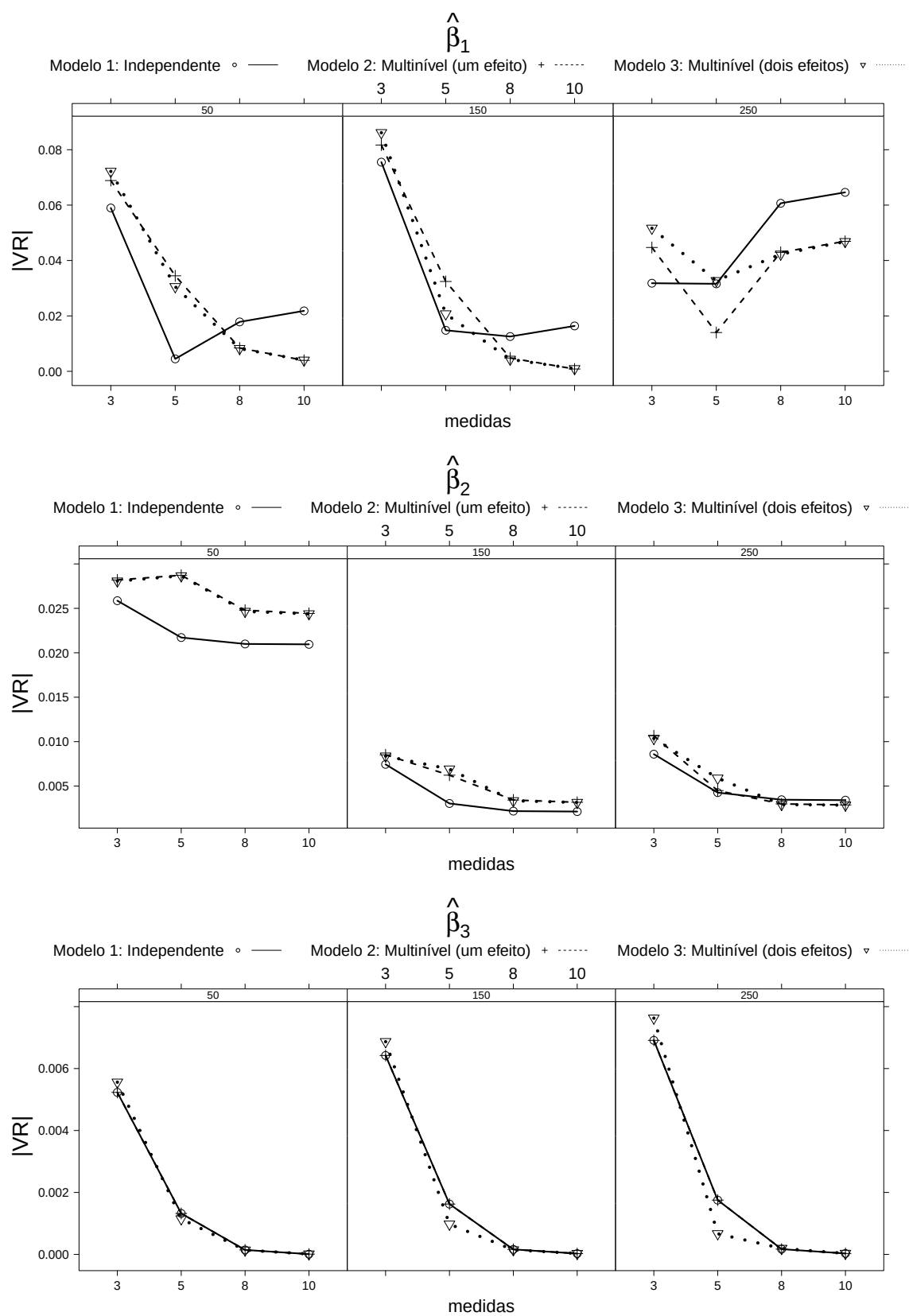


Figura 5.1: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$.

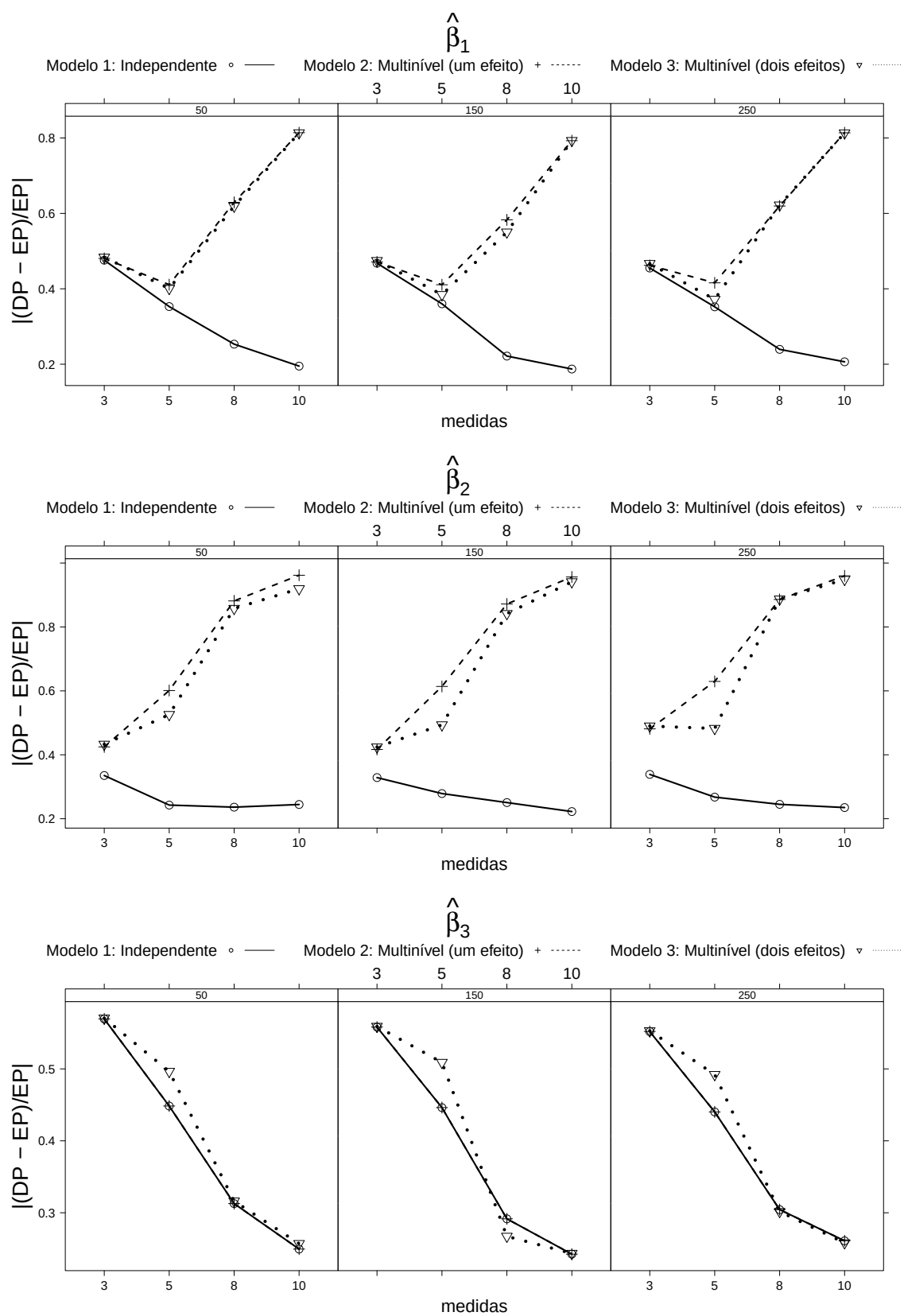


Figura 5.2: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$.

O mesmo não é observado para os **Modelos 2 e 3**, ou seja, as diferenças relativas aumentam com o aumento do número de medidas repetidas. Para o estimador de β_3 os valores para a diferença relativa diminuem com o aumento do número de medidas em todos os modelos ajustados.

De modelo geral, pode-se verificar que a performance do estimador da variância está associada ao tipo de variável, ou seja, quando se considera a variável tempo-independente os valores da diferença relativa entre o DP e o EP aumentam com o aumento do número de medidas repetidas. Já para a variável tempo-dependente os valores da diferença relativa diminuem com o aumento do número de mediadas repetidas.

Outros valores para a variância do efeito aleatório associado ao intercepto foram considerados e os gráficos encontram-se disponíveis nas Figuras B.1, B.2 e B.6 no Apêndice B.2.

A Tabela 5.2 apresenta os valores do AIC para o ajuste dos **Modelos 1, 2 e 3** considerando 50 indivíduos. É possível notar que o modelo Poisson multinível com efeito aleatório no intercepto (**Modelo 2**) ajusta melhor os dados, segundo o critério AIC e os seus valores encontram-se destacados. Isto é esperado uma vez que os dados foram gerados considerando o **Modelo 2**. Também é notável que os valores do AIC para o **Modelo 1** são muito maiores do que os **Modelos 2 e 3**. Além disso, verifica-se também que os valores do AIC dos **Modelos 2 e 3** são bem próximos. Considere, como exemplo, 10 medidas repetidas e variância do efeito aleatório igual a 0,5, os valores do AIC são 2.515.134,6; **4.833,9** e 4.837,4 para os **Modelos 1, 2 e 3**, respectivamente.

Pode-se verificar também que a variância do efeito aleatório τ_0 não tem muito efeito no critério de seleção AIC para os modelos multiníveis. Porém, para o modelo Poisson independente, com o aumento de τ_0 , há uma aumento muito grande nos valores do AIC. Considerando 8 medidas, os valores do AIC para o **Modelo 1** são 12.231,7; 281.187,8; 617.028,3 e 1.544.070,2 para τ_0 igual a 0,1; 0,5; 0,7 e 1,0 respectivamente (Tabela 5.2).

O mesmo comportamento para o critério de seleção de modelo AIC foi verificado para 150 e 250 indivíduos e os resultados encontram-se nas Tabelas B.11 e B.12 no Apêndice B.3.

De modo geral, verificou-se que o viés médio, o valor absoluto do viés relativo, a raiz do erro quadrático médio e a diferença relativa diminuíram com o aumento do número de medidas repetidas. O aumento da variabilidade de τ_0 , que implica em

Tabela 5.2: Valores do AIC para diferentes especificações no modelo Poisson com 50 indivíduos e gerando os dados utilizando o modelo Poisson multinível com intercepto aleatório.

τ_0	Modelos ajustados	Cenário 1/ Modelo 2 50 indivíduos			
		3	5	8	10
0,1	Modelo 1	750,4	1.799,8	12.231,7	87.669,0
	Modelo 2	748,9	1.573,1	3.262,5	4.668,5
	Modelo 3	752,7	1.585,2	3.268,9	4.671,9
0,5	Modelo 1	1.815,5	11.689,6	281.187,8	2.515.134,6
	Modelo 2	877,7	1.730,5	3.426,9	4.833,9
	Modelo 3	886,7	1.734,1	3.430,4	4.837,4
0,7	Modelo 1	3.143,4	24.035,2	617.028,3	5.546.330,0
	Modelo 2	911,6	1.765,3	3.462,5	4.869,9
	Modelo 3	915,3	1.768,9	3.466,0	4.873,4
1,0	Modelo 1	6.804,5	58.109,5	1.544.070,2	13.913.719,4
	Modelo 2	948,7	1.803,2	3.501,3	4.909,3
	Modelo 3	952,5	1.806,8	3.504,8	4.912,8

aumento da correlação intraclasse, afeta substancialmente o viés e erro quadrático médio dos resultados provenientes do modelo que assume independência das observações, quaisquer que sejam o número de medidas repetidas e número de indivíduos acompanhados longitudinalmente. Além disso, o critério de seleção de modelo AIC selecionou o modelo corretamente, não sofrendo influência da variância do efeito aleatório.

5.5.2 Geração: Modelo Poisson Multinível com Dois Efeitos Aleatórios

Nesta seção são apresentados os resultados para o ajuste dos modelos Poisson independente (**Modelo 1**), Poisson multinível com efeito aleatório no intercepto (**Modelo 2**) e Poisson multinível com dois efeitos aleatórios (**Modelo 3**), considerando que os dados foram gerados pelo modelo Poisson multinível com dois efeitos aleatórios.

A Tabela 5.3 apresenta os valores médios das estimativas, os vieses médios das estimativas e os valores médios da raiz do erro quadrático médio para os 2000 conjuntos de dados simulados considerando 50 indivíduos.

Tabela 5.3: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\tau_0 = 0, 1$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,397507	0,097507	0,102778	0,044159	-0,255841	0,258578	0,269157	-0,030843	0,055657
	5	0,327209	0,027209	0,031698	-0,542125	-0,842125	0,842370	0,291341	-0,008659	0,029804
	8	-0,165227	-0,465227	0,465244	-2,054321	-2,354321	2,354329	0,297361	-0,002639	0,013307
	10	-0,559640	-0,859640	0,859641	-3,232549	-3,532549	3,532550	0,289672	-0,010328	0,012187
$\widehat{\beta}_2$	3	1,827794	-0,472206	0,472626	2,227024	-0,072976	0,086959	2,372501	0,072501	0,099773
	5	1,364615	-0,935385	0,935398	2,146108	-0,153892	0,157069	2,368904	0,068904	0,082390
	8	0,815388	-1,484612	1,484612	1,992148	-0,307852	0,308055	2,369124	0,069124	0,073301
	10	0,582239	-1,717761	1,717761	1,890348	-0,409652	0,409688	2,380677	0,080677	0,081436
$\widehat{\beta}_3$	3	1,162005	0,062005	0,062899	1,162005	0,062005	0,062899	1,066689	-0,033311	0,036115
	5	1,250241	0,150241	0,150277	1,250241	0,150241	0,150277	1,058017	-0,041983	0,042361
	8	1,375440	0,275440	0,275440	1,375440	0,275440	0,275440	1,057025	-0,042975	0,043007
	10	1,432293	0,332293	0,332293	1,432293	0,332293	0,332293	1,057840	-0,042160	0,042165
$\tau_0 = 1, 0$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,930756	0,630756	0,631227	0,082155	-0,217845	0,220787	0,273101	1,054473	0,058598
	5	0,910437	0,610437	0,610563	-0,386706	-0,686706	0,687006	0,263510	-0,036490	0,053504
	8	0,522000	0,222000	0,222023	-1,670634	-1,970634	1,970645	0,269889	-0,030111	0,036075
	10	0,147041	-0,152959	0,152963	-2,739794	-3,039794	3,039796	0,279489	-0,020511	0,024893
$\widehat{\beta}_2$	3	1,997803	-0,302197	0,302545	2,784481	0,484481	0,488685	2,989524	0,689524	0,695418
	5	1,544205	-0,755795	0,755805	2,666273	0,366273	0,368693	2,973951	0,673951	0,678611
	8	0,970291	-1,329709	1,329709	2,499970	0,199970	0,200497	2,940019	0,640019	0,642168
	10	0,704564	-1,595436	1,595436	2,396295	0,096295	0,096555	2,905586	0,605586	0,606869
$\widehat{\beta}_3$	3	1,138150	0,038150	0,038957	1,138150	0,038150	0,038957	1,054473	-0,045527	0,047861
	5	1,213159	0,113159	0,113187	1,213159	0,113159	0,113187	1,059987	-0,040013	0,040656
	8	1,323828	0,223828	0,223829	1,323828	0,223828	0,223829	1,058753	-0,041247	0,041312
	10	1,380318	0,280318	0,280318	1,380318	0,280318	0,280318	1,057697	-0,042303	0,042323

É possível notar que o valor do viés médio do estimador $\hat{\beta}_1$ aumenta consideravelmente com o aumento da variância do efeito aleatório do intercepto, especialmente para o **Modelo 1**. Para 5 medidas os valores do viés são 0,027 e 0,610 considerando $\tau_0 = 0,1$ e $\tau_0 = 1,0$. Quando se aumenta o número de medidas para 8 e 10 medidas, os valores do viés diminuem, quando se comparam os valores de τ_0 .

Verifica-se que o estimador $\hat{\beta}_2$ está sendo subestimado no **Modelos 2** quando se considera a variância do efeito aleatório do intercepto τ_0 igual a 0,1. Porém o mesmo resultado não é verificado quando se considera $\tau_0 = 1,0$ (Tabela 5.3). Observa-se também aumento substancial do valor do viés médio do estimador $\hat{\beta}_2$ no **Modelo 3** quando se aumenta a variância de b_{0i} .

Para o estimador $\hat{\beta}_3$ é notado que os valores do viés para os **Modelos 1 e 2** são iguais. Além disso, o **Modelo 3** apresenta menores vieses para o estimador de β_3 (Tabela 5.3). Isso é esperado pois os dados foram gerados considerando dois efeitos aleatórios, um no intercepto (b_{0i}) e outro no tempo (b_{1i}). É possível verificar também que os valores da raiz do erro quadrático médio (*REQM*) diminuem com o aumento do número de indivíduos apenas para o **Modelo 3**. Considerando, por exemplo, o estimador $\hat{\beta}_2$ e $\tau_0 = 1,0$ tem-se que os valores para a *REQM* são 0,695, 0,679, 0,642 e 0,607 para 3, 5, 8 e 10 medidas, respectivamente, no **Modelo 3**, como visualizado na Tabela 5.3.

Com o aumento do número de indivíduos, resultados similares foram observados e os mesmos encontram-se disponíveis nas Tabelas B.6 e B.7 do Apêndice B.1. Também foi avaliado o impacto da variância do efeito aleatório para o ajuste dos modelos. Verificou-se um comportamento similar ao já descrito anteriormente e os resultados encontram-se disponíveis nas Tabelas B.8, B.9 e B.10 no Apêndice B.1.

A Figura 5.3 apresenta os valores absolutos do viés relativo ($|VR|$) dos estimadores $\hat{\beta}_1$, $\hat{\beta}_2$ e $\hat{\beta}_3$ considerando que a variância do efeito aleatório do intercepto é 0,1 para os modelos Poisson independente (**Modelo 1**), Poisson com efeito aleatório associado ao intercepto (**Modelo 2**) e Poisson com dois efeitos aleatórios (**Modelo 3**).

É visível que o **Modelo 3** apresenta a melhor performance para os três estimadores, ou seja, com o aumento do número de medidas repetidas os valores absolutos do viés relativo diminuem. Além disso, estes valores estão próximos de zero.

Vale ressaltar que para o estimador $\hat{\beta}_1$ é notável que o **Modelo 2** apresenta maiores valores absolutos do viés relativo. No entanto, quando se observa o estimador $\hat{\beta}_2$ o **Modelo 1** possui maiores valores para o valor absoluto do viés relativo,

como mostrado na Figura 5.3. Para o estimador $\hat{\beta}_3$ verifica-se que a performance dos estimadores para os **Modelos 1 e 2** são iguais.

Para ver o comportamento da variância do intercepto aleatório (τ_0) foram calculados os valores absolutos do viés relativo dos estimadores para $\tau_0 = 0,5; 0,7$ e $1,0$ e os resultados encontram-se sumarizados nas Figuras B.7, B.8 e B.9, respectivamente, no Apêndice B.2. Verifica-se que há uma influência da variância nos valores absolutos do viés relativo, mas, de modo geral, o **Modelo 3** apresenta menores valores, especialmente quando se aumenta o número de indivíduos.

Para avaliar a performance do estimador da variância foram avaliadas as diferenças relativas entre o desvio padrão e o erro padrão das estimativas e os resultados estão sumarizados na Figura 5.4, considerando a variância do efeito aleatório do intercepto τ_0 igual a 0,1. Verifica-se também que o número de indivíduos não afeta, de modo geral, na diferença relativa entre o desvio padrão e o erro padrão dos estimadores.

É possível notar que o tipo da variável influencia na performance do estimador da variância. Para o estimador $\hat{\beta}_2$ (estimador associado a variável tempo-independente) o **Modelo 2** apresenta a pior performance. Para o estimador de β_3 , os valores da diferença relativa para os **Modelos 1 e 2** são idênticos e convergem para zero. O mesmo não é observado no **Modelo 3**, uma vez que com o aumento do número de medidas a diferença relativa converge para 1 (Figura 5.4). Isso implica que a diferença entre o desvio padrão e o erro padrão é grande, o que não é um resultado positivo, pois o valor absoluto do viés relativo deveria diminuir com o aumento do número de medidas.

Com o aumento da variância do efeito aleatório, os valores absolutos do viés relativo diminuíram apenas para o **Modelo 1**. Os resultados encontram-se sumarizados nas Figuras B.10, B.11 e B.12, no Apêndice B.2.

A Tabela 5.4 apresenta os valores do AIC para os **Modelos 1, 2 e 3** considerando 50 indivíduos. É possível notar que o modelo Poisson multinível com dois efeitos aleatórios (**Modelo 3**) ajusta melhor os dados segundo o critério AIC. Este resultado é esperado, uma vez que os dados foram gerados considerando o **Modelo 3**. Também é notável que os valores do AIC para o **Modelo 1** são muito maiores do que os **Modelos 2 e 3**. Diferente dos valores do AIC encontrados no **Cenário 1**, os valores do AIC dos **Modelos 2 e 3** não são tão próximos, especialmente quando se aumenta o número de medidas repetidas.

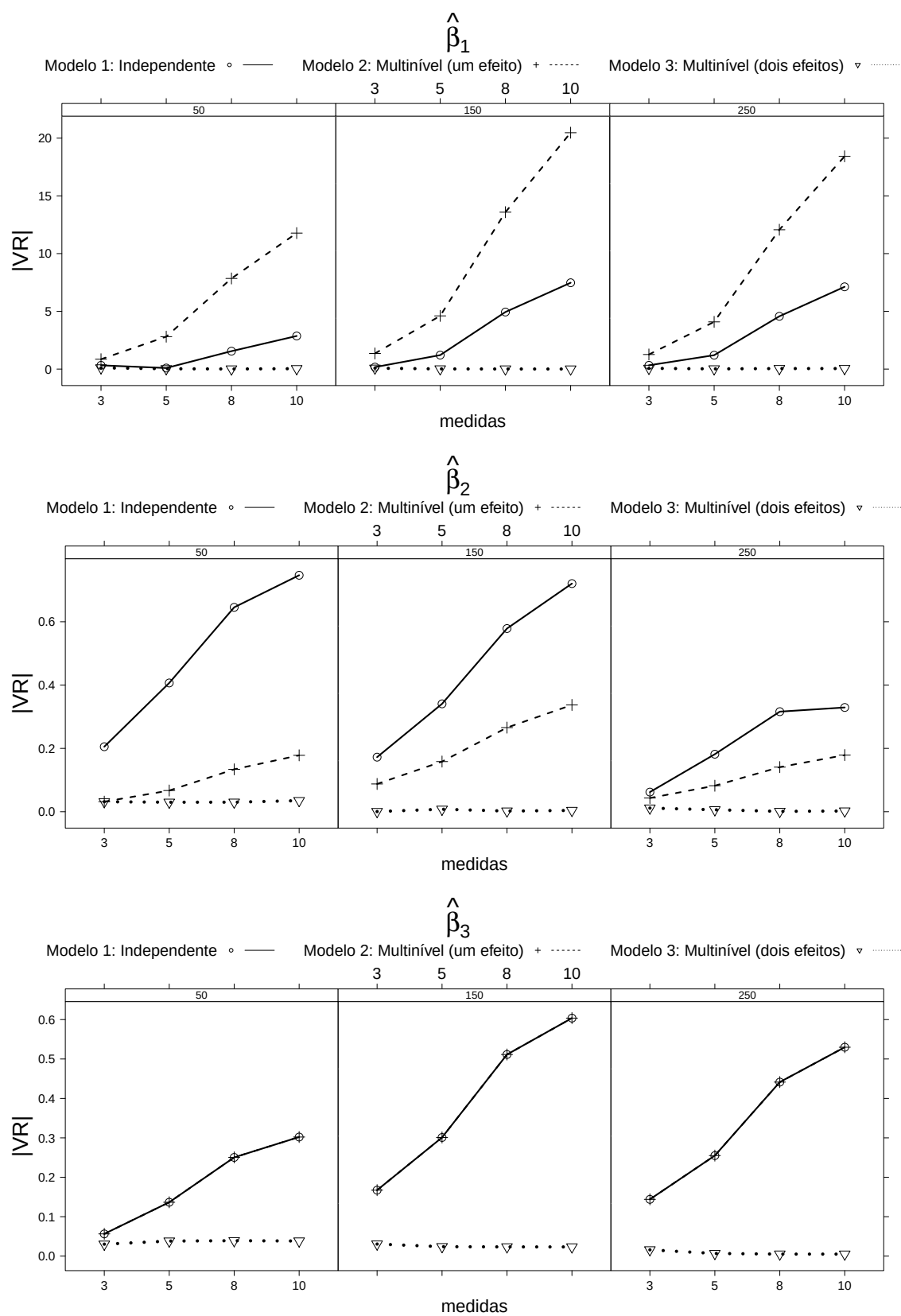


Figura 5.3: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0, 1$.

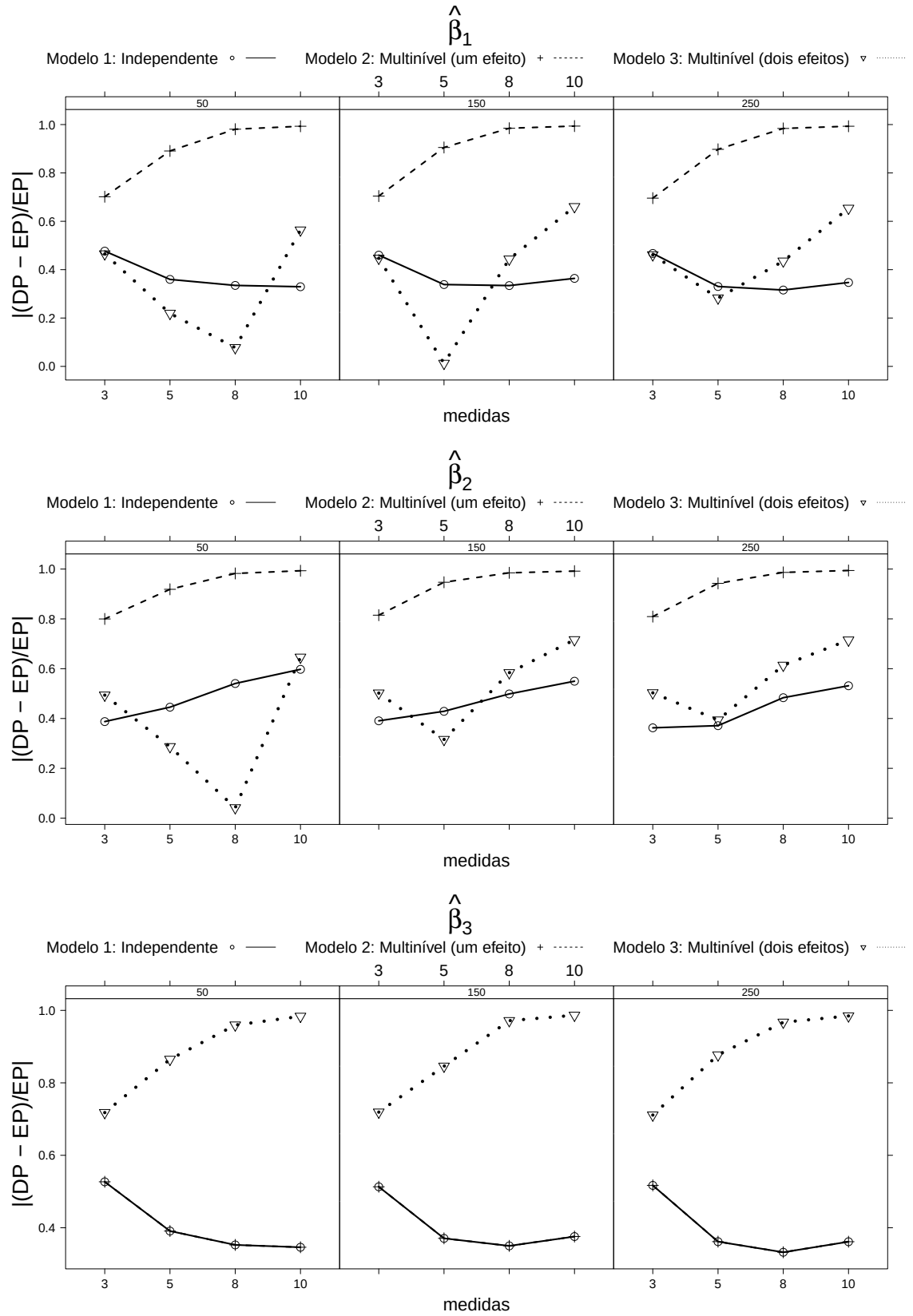


Figura 5.4: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0, 1$.

Considere, como exemplo, 10 medidas repetidas e variância do efeito aleatório igual a 0,5, quando os valores do AIC são 175.241.716,1; 761.584,3 e **5.178,4** para os **Modelos 1, 2 e 3**, respectivamente.

Pode-se verificar também que a variância do efeito aleatório τ_0 não tem muito efeito no critério de seleção AIC para os modelos multiníveis. Porém, para o modelo Poisson independente, com o aumento de τ_0 há uma aumento substancial nos valores do AIC (Tabela 5.4). Considere 8 medidas, os valores do AIC para o **Modelo 1** são 8.139.271,2; 8.248.977,1; 8.742.147,6 e 10.224.337,1, respectivamente, para τ_0 igual a 0,1; 0,5; 0,7 e 1,0.

Comportamentos similares para o critério de seleção AIC foram encontrados com o aumento do número de indivíduos e os resultados encontram-se dispostos nas Tabelas B.13 e B.14 no Apêndice B.3.

Tabela 5.4: Valores do AIC para diferentes especificações no modelo Poisson considerando 50 indivíduos e gerando os dados utilizando o modelo Poisson multinível com dois efeitos aleatórios.

τ_0	Modelos ajustados	Cenário 2/ Modelo 3 50 indivíduos			
		3	5	8	10
0,1	Modelo 1	3.033,3	71.676,7	8.139.271,2	183.123.733,6
	Modelo 2	1.031,4	3.699,4	67.768,0	685.424,4
	Modelo 3	918,4	1.810,4	3.603,6	5.021,9
0,5	Modelo 1	4.210,7	83.945,3	8.248.977,1	175.241.716,1
	Modelo 2	1.065,9	3.906,0	74.710,3	761.584,3
	Modelo 3	973,1	1.904,6	3.710,1	5.178,4
0,7	Modelo 1	5.476,4	97.949,3	8.742.147,6	178.969.515,0
	Modelo 2	1.093,9	4.106,1	80.943,9	826.795,5
	Modelo 3	996,9	1.937,6	3.745,6	5.214,6
1,0	Modelo 1	8.779,8	134.365,4	10.224.337,1	196.216.599,4
	Modelo 2	1.149,3	4.566,4	94.853,9	969.028,7
	Modelo 3	1.033,9	1.975,9	3.785,0	5.254,8

Pode-se concluir que com o aumento do número de indivíduos o valor absoluto do viés relativo diminuiu apenas para o modelo Poisson multinível com dois efeitos aleatórios (**Modelo 3**), independente do valor da variância (τ_0). Os resultados apresentados para a diferença relativa entre o desvio padrão e o erro padrão das estimativas não são favoráveis ao **Modelo 3** como era o esperado, uma vez que os dados foram gerados utilizando este modelo. Com o aumento da variância (τ_0) não foi verificado resultados melhores.

Em relação ao critério de seleção, tem-se que o AIC escolhe corretamente o

melhor modelo para ajustar os dados, independentemente do valor da variância de b_{0i} .

5.5.3 Geração: Modelo Poisson Marginal (EEG)

Como apresentado na Seção 5.2, na abordagem EEG foram ajustados quatro modelos em três cenários. Para cada número de indivíduos foram avaliados os modelos marginais de Poisson, segundo a abordagem EEG, com diferentes estruturas de correlação, sendo o **Modelo 1**: matriz de correlação de “trabalho” autorregressiva de ordem 1 (AR1), **Modelo 2**: matriz de correlação de “trabalho” permutável, **Modelo 3**: matriz de correlação de “trabalho” independente e **Modelo 4**: matriz de correlação de “trabalho” não estruturada (NE).

Considerando que as v.a.’s Poisson correlacionadas foram geradas utilizando a matriz de correlação R autorregressiva de ordem 1 (AR1), a Tabela 5.5 apresenta os valores médios das estimativas ($\bar{\beta}$), os vieses médios (VM) das estimativas e os valores da raiz do erro quadrático médio ($REQM$) para os 2000 conjuntos de dados simulados considerando 50 indivíduos.

É possível notar que os valores do viés médio vão convergindo para zero quando o número de medidas repetidas aumenta, considerando todas as matrizes de correlação, ou seja, os estimadores tornam-se não viesados com o aumento do número de medidas repetidas. Pode-se verificar também que o modelo com estrutura de correlação independente (**Modelo 3**) possui os valores do viés médio para 3 e 5 medidas um pouco maior em relação às outras estruturas de correlação. Na Tabela 5.5, por exemplo, para 5 medidas e considerando $\hat{\beta}_2$, os valores do viés são 0,0005, -0,0005, **0,0012** e 0,0006 para os **Modelos 1, 2, 3** e **4**, respectivamente. Por outro lado, quando se considera 8 e 10 medidas, os valores do viés médio são praticamente zero para todos os modelos ajustados.

Em relação à raiz do erro quadrático médio ($REQM$), é possível observar uma diminuição quando o número de medidas repetidas aumenta. Considerando, por exemplo, a matriz de correlação autorregressiva de ordem 1 (AR1) e o estimador $\hat{\beta}_3$, os valores da raiz do erro quadrático médio são 0,0118, 0,0032, 0,0006 e 0,0002, respectivamente, para 3, 5, 8 e 10 medidas, visualizado na Tabela 5.5.

O mesmo comportamento pode ser verificado quando se aumenta o número de indivíduos e os resultados encontram-se disponíveis nas Tabelas B.15 e B.16, no Apêndice B.4. Também foram avaliados as matrizes de correlação de “trabalho” não estruturada (**Cenário 2**) e permutável (**Cenário 3**) utilizadas para geração dos

Tabela 5.5: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) com 50 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação AR1. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	<i>VM</i>	<i>REQM</i>	$\hat{\beta}$	<i>VM</i>	<i>REQM</i>
$\hat{\beta}_1$	3	0,277317	-0,022683	0,040418	0,278215	-0,021785	0,039980
	5	0,294860	-0,005140	0,016413	0,295744	-0,004256	0,016397
	8	0,299526	-0,000474	0,004826	0,300356	0,000356	0,005045
	10	0,299940	-0,000060	0,001991	0,300364	0,000364	0,002232
$\hat{\beta}_2$	3	2,307013	0,007013	0,022723	2,304602	0,004602	0,022494
	5	2,300509	0,000509	0,004433	2,299510	-0,000490	0,004769
	8	2,300036	0,000036	0,001381	2,299684	-0,000316	0,001574
	10	2,299988	-0,000012	0,000457	2,299891	-0,000109	0,000533
$\hat{\beta}_3$	3	1,105695	0,005695	0,011883	1,105934	0,005934	0,012023
	5	1,100904	0,000904	0,003208	1,100953	0,000953	0,003268
	8	1,100056	0,000056	0,000599	1,099972	-0,000028	0,000221
	10	1,100007	0,000007	0,000199	1,099988	-0,000012	0,000623
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	<i>VM</i>	<i>REQM</i>	$\hat{\beta}$	<i>VM</i>	<i>REQM</i>
$\hat{\beta}_1$	3	0,274473	-0,025527	0,042457	0,277464	-0,022536	0,040412
	5	0,293070	-0,006930	0,017560	0,294831	-0,005169	0,016687
	8	0,298990	-0,001010	0,005215	0,299093	-0,000907	0,014570
	10	0,299748	-0,000252	0,002209	0,299947	-0,000053	0,003233
$\hat{\beta}_2$	3	2,311122	0,011122	0,025892	2,306749	0,006749	0,022889
	5	2,301151	0,001151	0,004939	2,300565	0,000565	0,004744
	8	2,300116	0,000116	0,001493	2,300076	0,000076	0,003956
	10	2,300004	0,000004	0,000506	2,299976	0,000024	0,001119
$\hat{\beta}_3$	3	1,106035	0,006035	0,012104	1,105691	0,005691	0,011921
	5	1,101214	0,001214	0,003432	1,100898	0,000898	0,003256
	8	1,100123	0,000123	0,000652	1,100122	0,000122	0,002243
	10	1,100026	0,000026	0,000223	1,100006	0,000006	0,000332

dados e resultados similares foram encontrados, sendo disponibilizados no Apêndice B.4 nas Tabelas B.17 - B.22.

A Figura 5.5 mostra os resultados do valor absoluto do viés relativo para as estimativas dos coeficientes do intercepto (β_1), da covariável X_i (β_2) e da covariável tempo, t (β_3), considerando 50, 150 e 250 indivíduos e diferentes números de medidas repetidas ($n = 3; 5; 8; 10$).

De modo geral, verificou-se que, com o aumento do número de medidas repetidas, o valor absoluto do viés relativo ($|VR|$) diminuiu substancialmente, considerando todas as matrizes de correlação de “trabalho”.

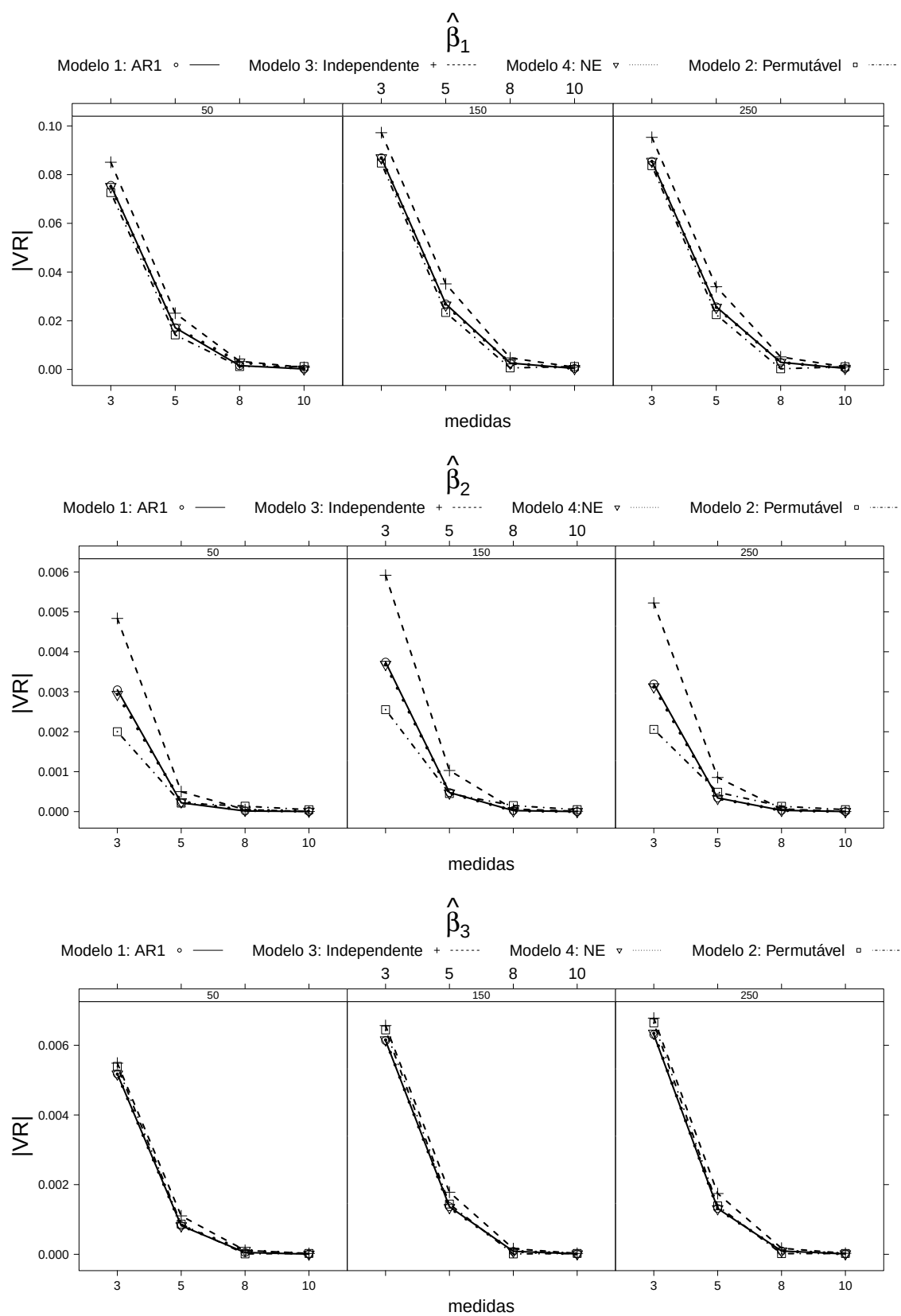


Figura 5.5: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação AR1.

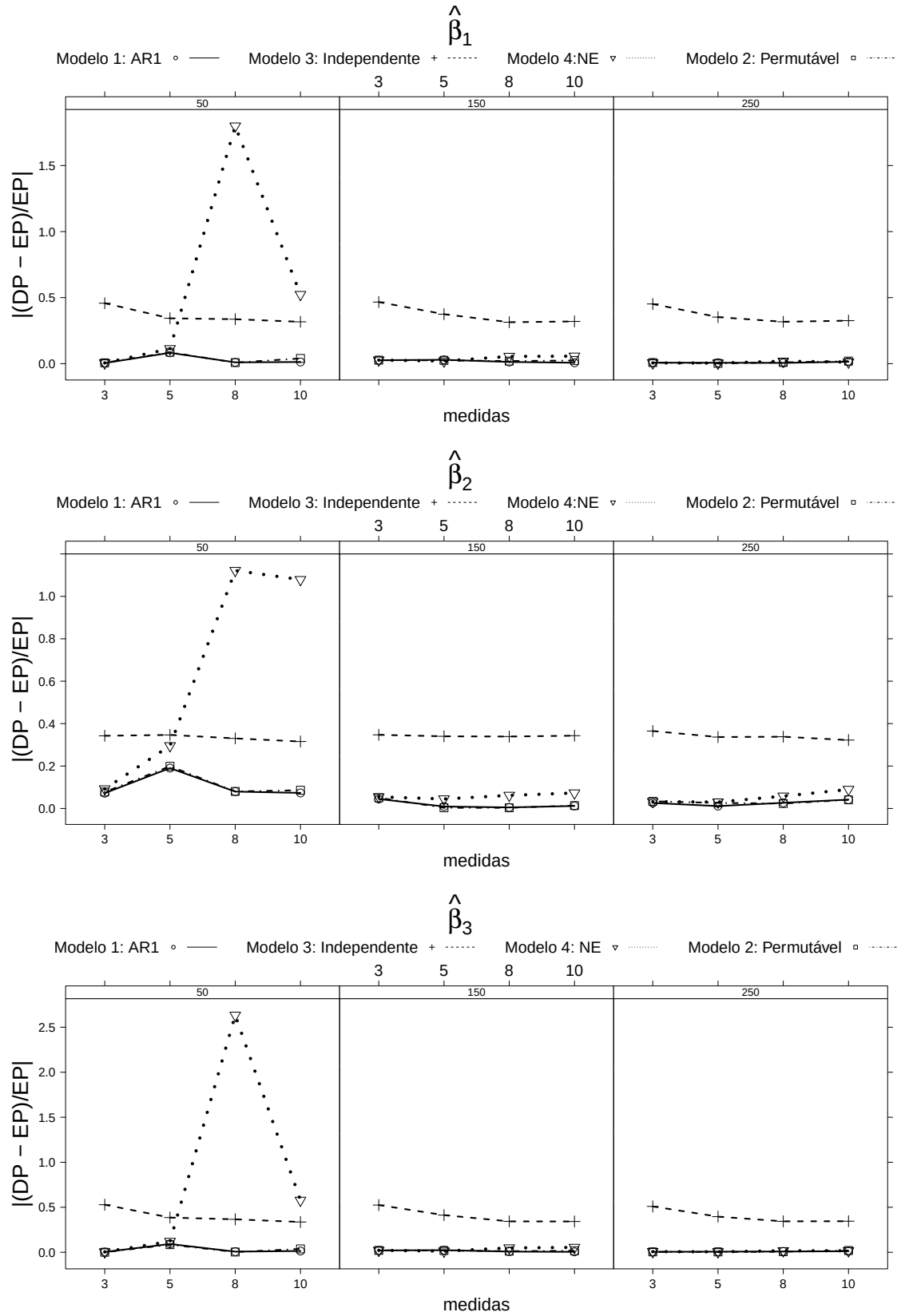


Figura 5.6: Diferença relativa dos estimadores para diferentes especificações no modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação AR1.

É possível notar que a diminuição do valor absoluto do viés relativo é mais visível quando se aumenta o número de medidas repetidas do que quando se aumenta o número de indivíduos. Verifica-se também que o número de indivíduos não interfere na estimação de β_3 , que representa o coeficiente angular da variável tempo (Figura 5.5). Observe que o maior valor do módulo do viés relativo é apresentado quando se considera a matriz com estrutura de correlação independente.

Quando se considera um menor número de medidas repetidas ($n = 3$) é notável que o valor absoluto do viés relativo considerando a matriz de correlação de “trabalho” permutável é o menor. No entanto, com o aumento do número de medidas a diferença entre os valores absolutos do viés relativo diminui substancialmente.

Também é possível verificar que, com o aumento do número de medidas, os valores absolutos do viés relativo convergem para zero. Isto pode ser explicado pelo fato do modelo marginal apresentar estimadores consistentes mesmo sob má especificação da matriz de correlação de “trabalho”.

Resultados similares para o valor absoluto do viés relativo foram encontrados para os cenários 2 e 3, ou seja, as respostas correlacionadas com distribuição de Poisson multivariada foram geradas considerando as estruturas de correlação NE e permutável, respectivamente. Os gráficos com estes resultados encontram-se disponíveis no Apêndice B.5, Figuras B.13 e B.14.

Para verificar a performance da variância dos estimadores foram calculadas as distâncias relativas entre o desvio padrão e o erro padrão de cada estimador através da equação (5.3) e os resultados encontram-se sumarizados na Figura 5.6, sendo considerada a mesma configuração usada para o valor absoluto do viés relativo.

Pode-se notar que as estimativas obtidas pelo modelo Poisson marginal com estrutura de correlação de “trabalho” independente apresentam a pior performance, exceto para 50 indivíduos em que os valores da diferença relativa para o modelo Poisson marginal com a estrutura de correlação não estruturada apresenta um aumento nos valores da diferença relativa para 8 e 10 medidas repetidas, como visualizado na Figura 5.6.

Também foram ajustados os modelos em que a geração considerou as matrizes de correlação de “trabalho” não estruturada e permutável. Os resultados foram similares aos apresentados nas Figuras B.15 e B.16 são apresentados no Apêndice B.5.

Para a escolha do melhor modelo foi considerado o critério de quasi-verossimilhança (QIC). É possível verificar na Tabela 5.6 que, mesmo gerando os dados com a

Tabela 5.6: Valores do critério de quasi-verossimilhança (QIC) para a escolha do melhor modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação autorregressiva de ordem 1.

n	Estrutura de correlação dos modelos ajustados	Matriz de correlação de “trabalho” AR1		
		m		
		50	150	250
3	AR1	-34.902,5	-81.147,5	-128.937,9
	Permutável	-34.902,3	-81.147,3	-128.937,6
	NE	-34.902,5	-81.147,5	-128.937,9
5	AR1	-676.030,0	-1.270.081,7	-2.027.052,9
	Permutável	-676.029,4	-1.270.080,7	-2.027.051,8
	NE	-676.030,0	-1.270.081,7	-2.027.052,9
8	AR1	-23.176.863,5	-56.319.984,5	-90.144.723,1
	Permutável	-23.176.861,9	-56.319.982,6	-90.144.721,1
	NE	-23.176.162,4	-56.319.984,6	-90.144.723,2
10	AR1	-261.937.052,8	-639.933.424,3	-1.025.218.744,3
	Permutável	-261.937.050,9	-639.933.422,2	-1.025.218.742,0
	NE	-261.937.022,3	-639.933.424,6	-1.025.218.744,4

matriz de correlação de “trabalho” autorregressiva de ordem 1 (AR1), os modelos ajustados com estrutura de correlação autorregressiva de ordem 1 e não estruturada apresentaram os mesmos valores para o QIC quando o número de medidas é pequeno (3 e 5 medidas). Quando o número de medidas aumenta para 8 e 10 e se considere 50 indivíduos o modelo 1 (AR1) apresenta menores valores para o QIC, **-23.176.863,5** e **-261.937.052,8**, respectivamente.

A diferença entre os valores do QIC é muito pequena, portanto este critério não é bom para selecionar o melhor modelo para dados de contagem longitudinal. Comportamentos similares para o critério de seleção de modelo QIC foram encontrados para os cenários 2 e 3 e os resultados encontram-se disponíveis nas Tabelas B.23 e B.24, no Apêndice B.6.

5.5.4 Geração: Modelo Binomial Negativo Multinível com Efeito Aleatório no Intercepto

Nesta seção são apresentados os resultados para o ajuste dos modelos Poisson independente (**Modelo 1**), Poisson multinível com efeito aleatório no intercepto (**Modelo 2**), Poisson multinível com dois efeitos aleatórios (**Modelo 3**), BN independente (**Modelo 4**), BN com efeito aleatório no intercepto (**Modelo 5**) e BN com dois efeitos aleatórios (**Modelo 6**) considerando que os dados foram gerados pelo modelo BN multinível com efeito aleatório no intercepto.

Por causa do alto tempo de execução necessário para a realização das simulações considerando a distribuição Binomial Negativa, estão sendo apresentados apenas os resultados considerando 50 indivíduos e a variância do efeito aleatório igual a 0,1 ($\tau_0 = 0,1$).

A Tabela 5.7 apresenta os valores médios das estimativas ($\widehat{\beta}$), os vieses médios das estimativas (VM) e os valores médios da raiz do erro quadrático médio ($REQM$) para os 500 conjuntos de dados simulados.

É possível verificar, avaliando o viés médio, que a média das estimativas está muito distante do verdadeiro valor do parâmetro para o estimador $\widehat{\beta}_1$, em todos os modelos ajustados. De modelo geral, com o aumento do número de medidas repetidas o viés converge para 0, especialmente para o estimador $\widehat{\beta}_3$.

As Figuras 5.7 e 5.8 apresentam os valores absolutos do viés relativo e a distância relativa das estimativas, respectivamente. Neste gráfico, b1, b2 e b3 representam, respectivamente, os estimadores $\widehat{\beta}_1$, $\widehat{\beta}_2$ e $\widehat{\beta}_3$.

Pode-se verificar que com o aumento do número de medidas repetidas houve uma diminuição no valor absoluto do viés relativo (Figura 5.7). Considerando o estimador $\widehat{\beta}_1$, o modelo Poisson independente apresentou os menores valores absolutos do viés relativo. Para os outros dois estimadores, verificou-se que todos os modelos apresentaram valores bem próximos, além disso os valores absolutos do viés relativo estavam próximos de zero.

Em relação à diferença relativa, notou-se que, com o aumento do número de medidas repetidas, a diferença relativa aumenta apenas para os diferentes modelos Poisson, especialmente para o estimador $\widehat{\beta}_2$. Já os modelos Binomial Negativo se mantêm em torno do zero. Com isso, pode-se concluir que os modelos Poisson não são adequados para ajustar os dados gerados a partir do modelo BN multinível com intercepto aleatório, além disso, vale ressaltar que estes dados apresentam superdispersão.

Tabela 5.7: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações nos modelos Poisson e BN com 50 indivíduos e dados gerados pelo modelo BN multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\tau_0 = 0, 1$											
	n	Modelo 1			Modelo 2			Modelo 3			
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	
$\widehat{\beta}_1$	3	-0,004672	-0,304672	0,364147	-0,087578	-0,387578	0,423769	-0,112861	-0,412861	0,441864	
	5	0,017556	-0,282444	0,366317	-0,018346	-0,318346	0,345883	-0,016229	-0,316229	0,344153	
	8	0,047511	-0,252489	0,397787	0,397787	-0,372405	0,457020	-0,083087	-0,383087	0,438050	
	10	0,046850	-0,253150	0,477583	0,024293	-0,275707	0,296254	0,024531	-0,275469	0,296245	
$\widehat{\beta}_2$	3	2,365709	0,065709	0,250434	2,391727	0,091727	0,219580	2,414868	0,114868	0,228107	
	5	2,341299	0,041299	0,253775	2,308765	0,008765	0,241086	2,308777	0,008777	0,241057	
	8	2,314569	0,014569	0,291340	2,334022	0,034022	0,283495	2,334320	0,034320	0,296175	
	10	2,334033	0,034033	0,260291	2,296889	-0,003111	0,242590	2,296889	-0,003111	0,242590	
$\widehat{\beta}_3$	3	1,111898	0,011898	0,071640	1,109747	0,009747	0,062716	1,106238	0,006238	0,055700	
	5	1,104590	0,004590	0,048611	1,103349	0,003349	0,024702	1,102991	0,002991	0,024881	
	8	1,100246	0,000246	0,094323	1,105110	0,005110	0,092626	1,104235	0,004235	0,089761	
	10	1,099809	-0,000191	0,064995	1,100870	0,000870	0,049854	1,100892	0,000892	0,049913	
	n	Modelo 4			Modelo 5			Modelo 6			
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	
$\widehat{\beta}_1$	3	-0,028729	-0,328729	0,352587	-0,033730	-0,333730	0,357660	-0,017542	-0,317542	0,342600	
	5	-0,005029	-0,305029	0,318718	-0,007478	-0,307478	0,321166	-0,006610	-0,306610	0,320162	
	8	0,016290	-0,283710	0,295093	0,003198	-0,296802	0,400283	-0,009692	-0,309692	0,484892	
	10	0,019864	-0,280136	0,289272	0,001466	-0,298534	0,496304	0,015884	-0,284116	0,296426	
$\widehat{\beta}_2$	3	2,408411	0,108411	0,187876	2,407769	0,107769	0,187814	2,404656	0,104656	0,189405	
	5	2,375105	0,075105	0,159324	2,373756	0,073756	0,158907	2,372138	0,072138	0,158953	
	8	2,359489	0,059489	0,229545	2,360254	0,060254	0,232510	2,357366	0,057366	0,229982	
	10	2,368870	0,068870	0,174882	2,367056	0,067056	0,176071	2,361108	0,061108	0,201898	
$\widehat{\beta}_3$	3	1,118290	0,018290	0,049178	1,117503	0,017503	0,048748	1,098766	-0,001234	0,046619	
	5	1,109621	0,009621	0,021779	1,109544	0,009544	0,021695	1,108942	0,008942	0,021318	
	8	1,103911	0,003911	0,085734	1,104148	0,004148	0,086071	1,103050	0,003050	0,085633	
	10	1,102764	0,002764	0,049787	1,100334	0,000334	0,187270	1,094499	-0,005501	0,105938	

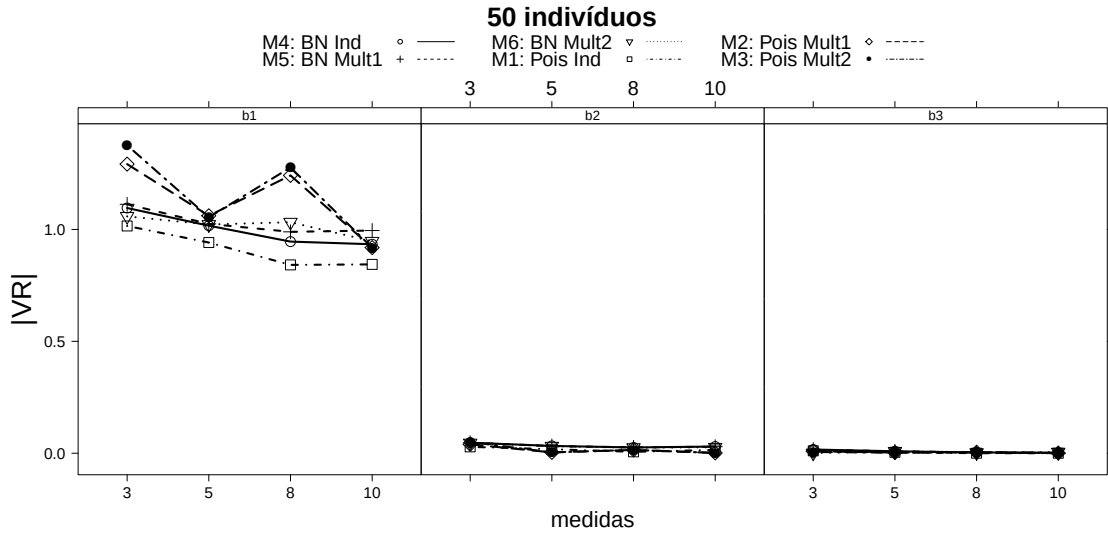


Figura 5.7: Valor absoluto do viés relativo para diferentes especificações nos modelos Poisson e BN considerando o modelo BN multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$.

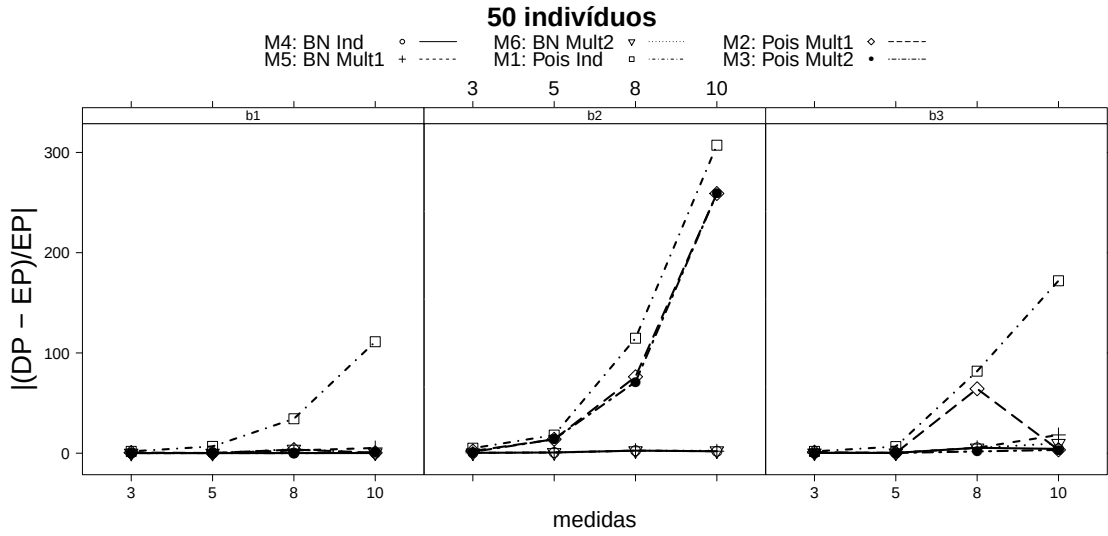


Figura 5.8: Diferença relativa das estimativas para diferentes especificações nos modelos Poisson e BN considerando o modelo BN multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0, 1$.

A Tabela 5.8 apresenta os valores do AIC para as diferentes especificações dos modelos Poisson e BN considerando que o modelo BN multinível com intercepto aleatório para a geração dos dados.

Claramente pode-se observar que os modelos Poisson não são adequados para ajustar os dados gerados. Além disso, o número de medidas influencia bastante nos valores do AIC. Considerando os modelos Binomial Negativo, notou-se que, em geral, os valores são próximos. O modelo selecionado para ajustar os dados foi o Binomial Negativo independente, para todos os número de medidas considerados. Este resultado não era o esperado, uma vez que os dados foram gerados a partir do modelo Binomial Negativo com intercepto aleatório.

Tabela 5.8: Valores do AIC para diferentes especificações dos modelos Poisson e BN considerando 50 indivíduos e gerando os dados utilizando o modelo BN multinível com intercepto aleatório.

τ_0	Modelos ajustados	Modelo 5 50 indivíduos			
		3	5	8	10
0,1	Modelo 1	1.803,9	12.104,0	292.402,1	2.650.281,2
	Modelo 2	1.182,6	8.770,4	126.170,4	2.216.786,4
	Modelo 3	1.039,0	8.773,1	98.802,7	2.216.789,3
	Modelo 4	966,7	2.135,2	4.713,1	6.979,2
	Modelo 5	968,1	2.136,6	4.721,9	7.412,6
	Modelo 6	970,2	2143,7	4.727,2	7.074,3

5.5.5 Geração: Modelo Binomial Negativo Marginal (Cópulas Gaussianas)

Nesta seção são apresentados os resultados para o ajuste dos modelos Poisson marginal via cópulas Gaussianas considerando a estrutura de correlação autorregressiva de ordem 1 (**Modelo 1**), Poisson marginal via cópulas Gaussianas considerando a estrutura de correlação permutável (**Modelo 2**), Poisson marginal via cópulas Gaussianas considerando a estrutura de correlação independente (**Modelo 3**), Poisson marginal via cópulas Gaussianas considerando a estrutura de correlação não estruturada (**Modelo 4**), BN marginal via cópulas Gaussianas considerando a estrutura de correlação autorregressiva de ordem 1 (**Modelo 5**), BN marginal via cópulas Gaussianas considerando a estrutura de correlação permutável (**Modelo 6**), BN marginal via cópulas Gaussianas considerando a estrutura de correlação independente (**Modelo 7**) e BN marginal via cópulas Gaussianas considerando a estrutura de correlação não estruturada (**Modelo 8**).

Tabela 5.9: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações nos modelos Poisson e BN com 50 indivíduos e dados gerados pelo modelo BN marginal via cópulas Gaussianas com estrutura de correlação AR1. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	n	Modelo 1: Poisson AR1			Modelo 2: Poisson Permutável			Modelo 3: Poisson Independente			Modelo 4: Poisson Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,039337	-0,260663	0,351255	0,043033	-0,256967	0,345490	-0,000870	-0,300870	0,376389	0,045233	-0,254767	0,345386
	5	0,079087	-0,220913	0,323607	0,079313	-0,220687	0,316916	0,071007	-0,228993	0,350211	0,075425	-0,224575	0,328595
$\hat{\beta}_2$	3	2,249439	-0,050561	0,245913	2,245280	-0,054720	0,245858	2,293177	-0,006823	0,237590	2,246322	-0,053678	0,248714
	5	2,241908	-0,058092	0,208303	2,243188	-0,056812	0,207518	2,254656	-0,045344	0,223938	2,247400	-0,052600	0,205393
$\hat{\beta}_3$	3	1,074986	-0,025014	0,096948	1,074282	-0,025718	0,095556	1,109781	0,009781	0,081151	1,071976	-0,028024	0,097066
	5	1,059192	-0,040808	0,085361	1,058769	-0,041231	0,082501	1,091603	-0,008397	0,058126	1,059305	-0,040695	0,084563
			Modelo 5: BN AR1		Modelo 6: BN Permutável			Modelo 7: BN Independente			Modelo 8: BN Não Estruturada		
	n	$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	-0,035772	-0,335772	0,365376	-0,028913	-0,328913	0,359373	-0,037439	-0,337439	0,367068	-0,032070	-0,332070	0,362157
	5	-0,003744	-0,303744	0,324667	0,012084	-0,287916	0,311650	0,006808	-0,293192	0,315473	0,005218	-0,294782	0,317172
$\hat{\beta}_2$	3	2,350293	0,050293	0,153847	2,342688	0,042688	0,154057	2,335896	0,035896	0,153389	2,349313	0,049313	0,153952
	5	2,316005	0,016005	0,116624	2,306177	0,006177	0,121649	2,306428	0,006428	0,117975	2,311673	0,011673	0,120498
$\hat{\beta}_3$	3	1,119636	0,019636	0,059536	1,117512	0,017512	0,058719	1,121715	0,021715	0,060542	1,118229	0,018229	0,059622
	5	1,107500	0,007500	0,029917	1,103761	0,003761	0,029980	1,105356	0,005356	0,030181	1,105098	0,005098	0,030761

Por causa do alto tempo de execução necessário para a realização das simulações considerando o modelo Binomial Negativo marginal via cópulas Gaussianas, estão sendo apresentados apenas os resultados considerando 50 indivíduos para 3 e 5 medidas repetidas. Quando se considerou 8 e 10 medidas, o número de não convergência foi muito alto. Por exemplo, considerando 5 medidas repetidas, das 500 réplicas utilizadas, 188 não convergiram. Isto causa um viés muito grande nas estimativas e os resultados obtidos não são mais confiáveis.

A Tabela 5.9 apresenta os valores médios das estimativas, os vieses médios e os valores da *REQM* para os modelos apresentados acima. É possível notar que para os modelos Poisson marginais via cópulas Gaussianas o viés médio aumenta com o aumento do número de medidas repetidas. O mesmo não é verificado para os modelos BN marginais via cópulas Gaussianas para todas as estruturas de correlação de trabalho.

Em relação à raiz do erro quadrático médio é possível verificar que com o aumento do número de medidas repetidas houve uma diminuição dos valores da raiz do erro quadrático médio. Observa-se também que os valores da raiz do erro quadrático médio para os modelos BN marginais via cópulas Gaussianas são menores do que os valores para os modelos Poisson marginais via cópulas Gaussianas (Tabela 5.9).

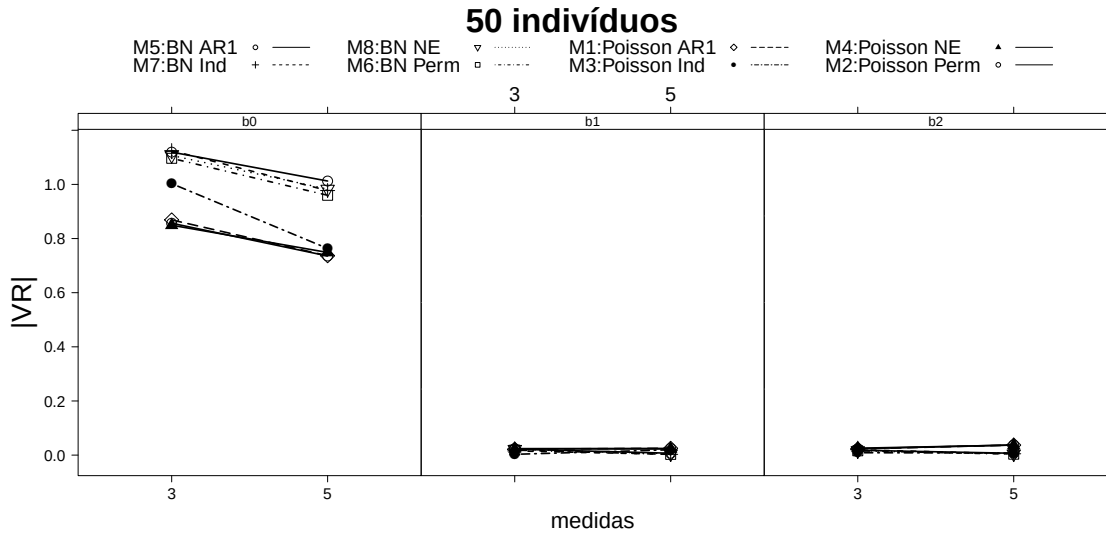


Figura 5.9: Valor absoluto do viés relativo para diferentes especificações nos modelos Poisson e BN marginais com cópulas Gaussianas, com dados gerados pelo modelo BN marginal via cópulas Gaussianas com correlação AR1.

A Figura 5.9 apresenta os valores absolutos do viés relativo para os modelos ajustados. Pode-se verificar que os modelos Poisson marginais com cópulas Gaussi-

anas apresentaram os menores valores absolutos do viés relativo, especialmente para o estimador $\hat{\beta}_1$, representado por (b0). Vale a pena ressaltar que o modelo Poisson marginal via cópulas Gaussianas considerando estrutura de correlação independente apresentou o maior valor absoluto do viés relativo. Porém com o aumento do número de medidas houve uma diminuição considerável.

Em relação aos modelos BN marginais via cópulas Gaussianas, verificou-se que, para o estimador $\hat{\beta}_1$, os valores absolutos do viés relativos são maiores que para os modelos Poisson. Para os estimadores $\hat{\beta}_2$ (b1) e $\hat{\beta}_3$ (b2) os valores absolutos do viés relativo são próximos de zero para todos os modelos ajustados.

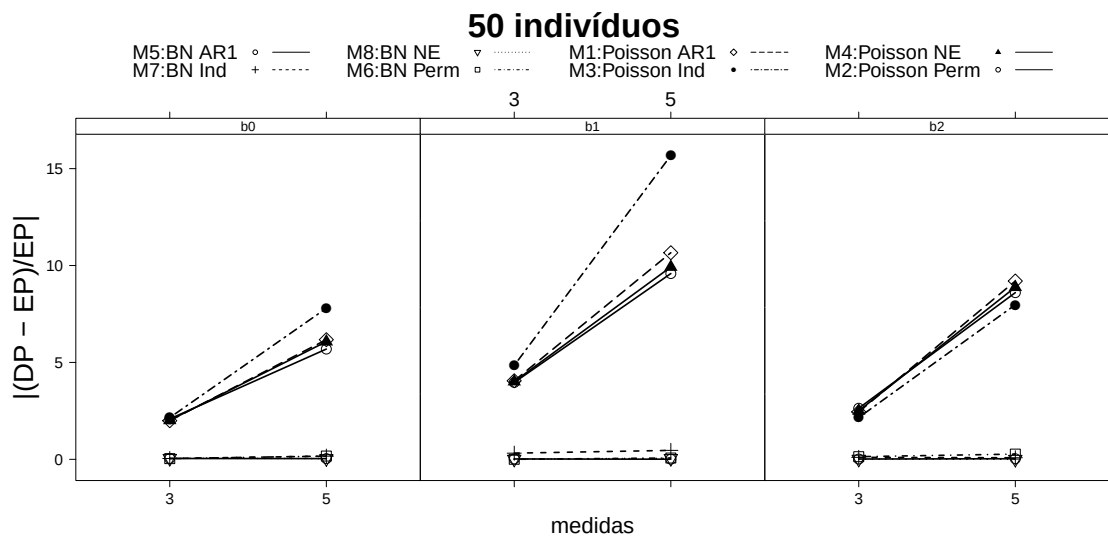


Figura 5.10: Diferença relativa das estimativas para diferentes especificações nos modelos Poisson e BN marginais via cópulas Gaussianas, com dados gerados pelo modelo BN marginal via cópulas Gaussianas com correlação AR1.

Em relação à diferença relativa, pode-se verificar que com o aumento do número de medidas repetidas, os valores da diferença relativa aumenta apenas para os modelos Poisson marginais, em especial para o modelo Poisson marginal via cópulas Gaussianas com estrutura de correlação independente, visualizado na Figura 5.10.

Considerando o critério de seleção AIC, verificou-se que os modelos Poisson marginal via cópulas Gaussianas não são indicados para ajustar os dados gerados nesta simulação. O modelo escolhido foi o modelo Binomial Negativo marginal com cópulas Gaussianas com estrutura de correlação AR1, com AIC iguais a 935,8 e 2074,0 para 3 e 5 medidas repetidas, repectivamente (Tabela 5.10).

Pode-se concluir que, em geral os modelos BN marginais via cópulas Gaussianas são mais indicados para ajustar os dados gerados neste cenário. Além disso, o

Tabela 5.10: Valores do AIC para diferentes especificações dos modelos Poisson e BN marginais via cópulas Gaussianas considerando 50 indivíduos e gerando os dados pelo modelo BN multinível com cópulas Gaussianas com correlação AR1.

Modelos ajustados	Modelo 5	
	50 indivíduos	
	3	5
Modelo 1	1.327,8	3.612,1
Modelo 2	1.333,0	3.630,1
Modelo 3	1.736,6	10.305,7
Modelo 4	1.327,1	3.602,9
Modelo 5	935,8	2.074,0
Modelo 6	940,3	2.095,8
Modelo 7	959,8	2.124,6
Modelo 8	937,5	2.082,0

critério de seleção AIC é um boa maneira de escolher o melhor modelo, uma vez que ele selecionou o modelo corretamente, pois os dados foram gerados a partir da estrutura de correlação autorregressiva de ordem 1.

Capítulo 6

Aplicações

Neste capítulo são apresentadas três aplicações com o objetivo de ilustrar a utilização dos modelos longitudinais para análise de dados de contagem considerando as abordagens multinível e marginal. A principal motivação para apresentar essas aplicações é mostrar diferentes contextos em que os métodos podem ser utilizados desde situações que envolvem dados com estrutura mais simples até dados que envolvem maior complexidade.

A primeira aplicação, apresentada na Seção 6.1, é considerada a mais simples, uma vez que os dados são balanceados e igualmente espaçados no tempo. Os dados são provenientes de um estudo clínico que avalia 59 pacientes com epilepsia. Os pacientes foram alocados de forma aleatória em dois grupos e o número de ataques epiléticos foi mensurado ao longo de quatro períodos igualmente espaçados. Neste conjunto de dados três covariáveis foram consideradas e nenhuma delas variava no tempo.

Na segunda aplicação, 6.127 indivíduos foram acompanhados com a finalidade de avaliar a reforma no sistema de saúde alemão (Seção 6.2). Nesta aplicação o conjunto de dados é desbalanceado, uma vez que os indivíduos não possuem o mesmo número de medidas repetidas. Contudo as mensurações são igualmente espaçadas porque as observações foram obtidas anualmente no período de 1984 a 1988. Foram avaliadas seis covariáveis, em que uma delas é tempo-dependente.

A Seção 6.3 traz uma terceira aplicação que apresenta dados com estrutura mais complexa. Os dados são totalmente desbalanceados e desigualmente espaçados. Este conjunto de dados inéditos refere-se a informações sobre o número de células CD4+ de pacientes com HIV residentes em Salvador-Bahia que foram acompanhados no período de janeiro-2002 a agosto/2012. Todas as covariáveis consideradas são tempo-independentes.

A seguir são apresentados maiores detalhes sobre cada uma das aplicações, bem como resultados das análises conduzidas utilizando os modelos Poisson e Binomial Negativo sob as abordagens multinível e marginal (utilizando EEG e cópulas Gaussianas). As análises foram realizadas nos programas estatísticos **R** e STATA.

6.1 Aplicação 1: Ataques de epilepsia

Neste estudo clínico são avaliados 59 pacientes com epilepsia que foram distribuídos aleatoriamente em dois grupos: um recebendo uma droga anti-epiléptica (tratamento) e o outro placebo (Thall & Vail, 1990). O número de ataques epiléticos foram contados ao longo de quatro períodos, portanto é um estudo longitudinal. Além disso, uma taxa de ataque inicial foi registrada para cada paciente (*baseline*, $t = 0$). As variáveis estudadas foram: (i) o tempo, t , de mensuração em semanas (0, 2, 4, 6 e 8); (ii) a idade dos pacientes no *baseline*, em anos completos; (iii) grupo de tratamento (1 = droga anti-epiléptica e 0 = placebo) e (iv) o número de ataques epiléticos (Y) como a variável resposta. Para cada um dos quatro períodos os pacientes foram avaliados durante duas semanas. No *baseline* foi contabilizada a soma do número de ataques epiléticos ocorridos durante 8 semanas anteriores ao início do estudo e esta soma foi dividida por 4 para tornar possível a comparação com os outros 4 períodos. Se este número médio for decimal, ele é arredondado para mais, por exemplo 4,5 ataques considera-se 5 ataques no *baseline*. Note que não há variáveis tempo-dependentes neste conjunto de dados.

Como os dados são balanceados e igualmente espaçados, é possível construir os gráficos de perfis individuais e do perfil médio. O gráfico de perfis individuais, visualizado na Figura 6.1-(a), apresenta as trajetórias de todos os pacientes no período do estudo. Pode-se notar que a maioria dos pacientes tiveram entre 0 e 20 ataques epiléticos. No entanto, um indivíduo apresentou um grande número de ataques no primeiro período de acompanhamento e um outro paciente teve um aumento considerável no número de ataques no final do estudo causando, portanto, uma elevação no perfil médio dos indivíduos, apresentado na Figura 6.1-(b).

A Tabela 6.1 apresenta informações sobre o número de ataques epiléticos em relação ao tratamento recebido pelos pacientes. Pode-se verificar que 28 pacientes receberam o placebo e 31 o tratamento anti-epiléptico no *baseline*. O número médio de ataques epiléticos para os grupos placebo e tratamento foi 8,49 e 8,01, respectivamente, em todo o período do estudo. Além disso, ao longo do período do estudo, o número médio de ataques epiléticos diminuiu a partir da segunda semana, exceto

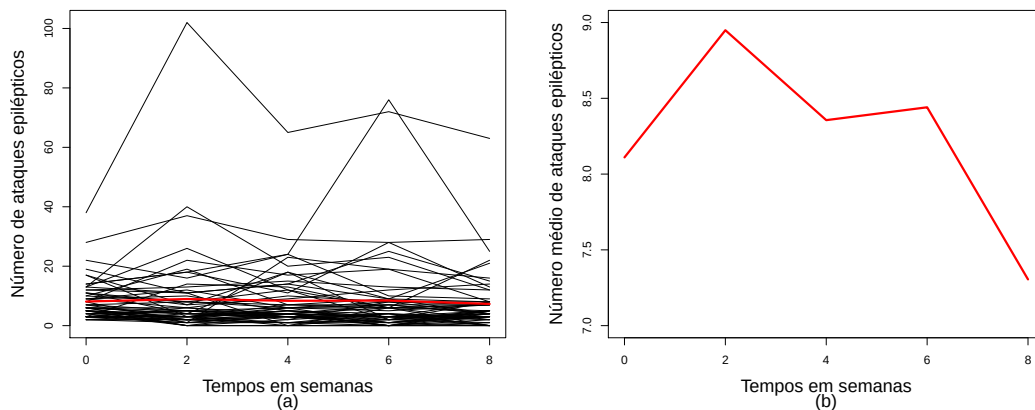


Figura 6.1: Gráficos do (a) perfil individual e (b) perfil médio do número de ataques epiléticos.

para o grupo placebo que apresentou um pequeno aumento entre as semanas 4 e 6, passando de 8,29 ataques para 8,79 ataques.

Tabela 6.1: Número médio de ataques epiléticos e desvio-padrão (DP) por grupo de tratamento.

Tempo em semanas	Placebo (m=28)		Tratamento (m = 31)		Total (m = 59)	
	Média	DP	Média	DP	Média	DP
Baseline	8,04	6,49	8,23	6,97	8,14	6,69
2	9,36	10,14	8,58	18,24	8,95	14,84
4	8,29	8,16	8,42	11,86	8,36	10,19
6	8,79	14,67	8,13	13,89	8,44	14,15
8	7,96	7,63	6,71	11,26	7,31	9,65
Média	8,49	9,42	8,01	12,45	8,24	11,11

Foram ajustados modelos de regressão Poisson e Binomial Negativo (BN) para avaliar o efeito do tratamento e outros fatores no número médio de ataques epiléticos. Estes modelos foram ajustados sobre três abordagens: independente, multinível e marginal (EEG e cópulas Gaussianas). Na Tabela 6.2 encontram-se as estimativas para o ajuste dos modelos independente, multinível com efeito aleatório no intercepto (Multinível 1) e multinível com efeitos aleatórios no intercepto e no tempo (Multinível 2). O efeito aleatório no intercepto, denotado por b_{0i} , segue distribuição normal e o efeito aleatório no intercepto e o efeito aleatório no tempo, b_{1i} , seguem distribuição conjunta normal bivariada.

Considerando a distribuição Poisson, verificou-se que a variável idade foi significativa apenas no modelo Poisson independente, ao nível de 5% de significância.

Tabela 6.2: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão Poisson e BN independentes, multiníveis com intercepto aleatório (Multinível 1) e multiníveis com dois efeitos aleatórios (Multinível 2) para o número de ataques epiléticos.

Modelos	Abordagem	Parâmetros				AIC
		Intercepto	Tratamento	Idade	Tempo	
Poisson	Independente	2,634 (0,103)*	-0,078 (0,041)	-0,015 (0,003)*	-0,013 (0,007)	3739,4
	Multinível 1	2,212 (0,549)*	-0,203 (0,226)	-0,012 (0,018)	-0,013 (0,007)	1762,8
	Multinível 2	2,256 (0,508)*	-0,085 (0,209)	-0,012 (0,017)	-0,045 (0,011)*	1731,9
BN	Independente	2,717 (0,325)*	-0,127 (0,126)	-0,017 (0,010)	-0,014 (0,022)	1874,5
	Multinível 1	2,293 (0,553)*	-0,206 (0,227)	-0,012 (0,018)	-0,030 (0,012)*	1641,7
	Multinível 2	2,288 (0,537)*	-0,183 (0,221)	-0,012 (0,018)	-0,035 (0,013)*	1642,2

* p-valor < 5%

Observou-se também que a variável tempo tem estimativas idênticas considerando-se as primeiras 3 casas decimais para os modelos Poisson independente e Multinível 1 (Tabela 6.2). Além disso, verificou-se que a variável tempo é significativa, ao nível de 5% de significância, apenas para o modelo Multinível 2. Segundo o critério AIC, o melhor modelo Poisson foi o modelo Multinível 2 (AIC = 1731,9).

Em relação à distribuição BN, notou-se que a variável tempo não foi significativa apenas para o modelo BN independente, ao nível de 5% de significância (Tabela 6.2). Foi possível notar também que o tratamento com a droga anti-epilética não teve efeito estatisticamente significativo em quaisquer modelos. Porém houve uma diminuição no número de ataques epiléticos no grupo tratamento comparado com o grupo placebo. O AIC do modelo BN multinível com efeito aleatório no intercepto foi o menor dentre todos os modelos ajustados (AIC = 1641,7).

Também foram ajustados modelos Poisson e BN marginais utilizando a abordagem EEG com diferentes estruturas de correlação e os resultados podem ser verificados na Tabela 6.3. Quando se considerou a matriz de correlação de “trabalho” não estruturada o modelo não apresentou convergência para nenhum dos modelos ajustados. Não se calcula o valor do QIC para o modelo com estrutura de correlação

independente, pois os outros valores do QIC são calculados em relação ao mesmo.

Tabela 6.3: Estimativas dos coeficientes de regressão e erros-padrão (entre parênteses) usando os modelos de regressão de Poisson e BN na abordagem EEG com diferentes estruturas de correlação para o número de ataques epiléticos.

Modelos	Estrutura de dependência	Parâmetros				QIC
		Intercepto	Tratamento	Idade	Tempo	
Poisson	AR1	2,689 (0,195)*	-0,089 (0,080)	-0,020 (0,007)*	-0,015 (0,012)	-5047,6
	Permutável	2,608 (0,198)*	-0,069 (0,081)	-0,016 (0,007)*	-0,019 (0,007)*	-4974,5
	Independente	2,634 (0,103)*	-0,078 (0,041)	-0,015 (0,003)*	-0,013 (0,007)	
BN	AR1	2,769 (0,578)*	-0,146 (0,236)	-0,022 (0,019)	-0,017 (0,578)	347,4
	Permutável	2,694 (0,600)*	-0,126 (0,247)	-0,017 (0,198)	-0,020 (0,022)	355,8
	Independente	2,717 (0,314)*	-0,127 (0,124)	-0,017 (0,010)	-0,014 (0,022)	

* p-valor < 5%

Observa-se que a variável tempo foi significativa apenas no modelo Poisson com estrutura de correlação de “trabalho” permutável. Pelo QIC, o melhor modelo Poisson EEG é o que apresenta estrutura de correlação de “trabalho” AR1 (QIC = -5047,6). Para este modelo, o logaritmo do número médio de ataques epiléticos diminui 0,016 quando se aumenta 1 ano na idade do paciente, considerando as outras variáveis fixas no modelo.

Considerando o modelo Binomial Negativo, o QIC demonstrou que a melhor estrutura de correlação é a AR1 (QIC= 347,4). Pode-se verificar também na Tabela 6.3 que nenhuma das variáveis foram significativas para o modelo, considerando a abordagem EEG.

Também foram ajustados modelos Poisson e BN utilizando a abordagem marginal via cópulas Gaussianas e utilizando diferentes estruturas de correlação e os resultados podem ser verificados na Tabela 6.4.

Considerando o modelo Poisson marginal via cópulas Gaussianas, verificou-se que a variável tempo não foi significativa apenas para a estrutura de correlação independente, ao nível de 5% de significância (Tabela 6.4). No modelo BN marginal via cópulas Gaussianas, a variável idade não foi significativa, ao nível de 5%, para

Tabela 6.4: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Poisson e BN na abordagem marginal via cópulas Gaussianas com diferentes estruturas de correlação para o número de ataques epiléticos.

Modelos	Estruturas de dependência	Parâmetros				AIC
		Intercepto	Tratamento	Idade	Tempo	
Poisson	AR1	2,364 (0,154)*	-0,381 (0,054)*	-0,003 (0,005)	-0,031 (0,008)*	2331,0
	Permutável	2,331 (0,166)*	-0,369 (0,060)*	-0,002 (0,005)	-0,032 (0,006)*	2256,0
	Não Estruturada	2,307 (0,171)*	-0,274 (0,066)*	-0,005 (0,006)	-0,032 (0,007)*	2304,0
	Independente	2,634 (0,103)*	-0,078 (0,041)	-0,015 (0,003)*	-0,013 (0,007)	3739,0
BN	AR1	2,920 (0,450)*	-0,148 (0,176)	-0,024 (0,014)	-0,018 (0,020)	1708,0
	Permutável	2,878 (0,498)*	-0,171 (0,196)	-0,020 (0,016)	-0,024 (0,012)*	1676,0
	Não Estruturada	2,860 (0,491)*	-0,163 (0,198)	-0,020 (0,015)	-0,023 (0,010)*	1683,0
	Independente	2,717 (0,326)*	-0,127 (0,127)	-0,017 (0,010)	-0,014 (0,022)	1875,0

* p-valor < 5%

nenhuma das estruturas de correlação.

Dentre os modelos Poisson marginais via cópulas Gaussianas, observou-se que o melhor modelo é aquele com estrutura de correlação permutável segundo o critério AIC (AIC = 2256,0). Em relação ao modelo BN marginal via cópulas Gaussianas notou-se que o melhor modelo foi também aquele com estrutura de correlação permutável, segundo o critério AIC (AIC = 1676,0), como mostrado na Tabela 6.4.

Verificou-se ainda que as estimativas dos coeficientes relacionados ao intercepto e à idade para o modelo BN marginal via cópulas Gaussianas são maiores do que as estimativas para o modelo Poisson marginal correspondente (Tabela 6.4). Notou-se também que os erros padrão para o modelo BN marginal são maiores do que os erros padrão para o modelo Poisson marginal via cópulas Gaussianas, para todas as estruturas de correlação de “trabalho”.

Realizou-se uma análise de diagnóstico para verificar a bondade de ajuste dos

modelos multiníveis e marginais (EEG e cópulas Gaussianas). Em relação aos modelos marginais segundo a abordagem EEG, o diagnóstico foi realizado apenas para os modelos Poisson, não sendo possível construir os gráficos da distância de Cook e dos resíduos padronizados considerando a estrutura de correlação não estruturada, pois não foi possível calcular a matriz $\mathbf{H}$. Os envelopes simulados foram feitos apenas para os modelos Binomial Negativo marginais via cópulas Gaussianas.

Os gráficos foram obtidos através de uma adaptação baseada nos códigos fornecidos por Paula em www.ime.usp.br/~giapaula/textoregressao.htm e encontram-se disponíveis em www.de.ufpe.br/~raydonal/Tese/Supplementary-Materials/Trindade-2014.

A Figura 6.2 apresenta os gráficos contendo a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson multiníveis com intercepto aleatório e com dois efeitos aleatórios. É verificado que existem observações discrepantes no conjunto de dados e estas são destacadas nos gráficos da distância de Cook.

A análise de diagnóstico foi também realizada para os modelos BN multiníveis com intercepto aleatório e com dois efeitos aleatórios e os resultados encontram-se na Figura 6.3. Pode-se observar que a maioria dos valores destacados nos modelos Poisson multiníveis são encontrados no diagnóstico dos modelos BN multiníveis. Os resíduos de Pearson padronizados encontram-se dentro da banda de confiança e a maioria dos pontos estão em torno de zero.

Em relação aos modelos marginais (EEG), a Figura 6.4 apresenta os gráficos da distância de Cook e dos resíduos padronizados para os modelos Poisson marginais (EEG) com estrutura de correlação autorregressiva de ordem 1 (AR1) e permutável, respectivamente. O diagnóstico para o modelo BN marginal (EEG) não foi realizado, pois o mesmo não pode ser ajustado no $\mathbf{R}$.

É possível notar que, novamente, houve detecção de pontos atípicos no conjunto de dados, tanto pelo gráfico da distância de Cook quanto pelo gráfico dos resíduos de Pearson Padronizados.

A Figura 6.5 apresenta os gráficos contendo a distância de Cook e os resíduos de Pearson padronizados do modelo Poisson marginal via cópulas Gaussianas considerando as estruturas de correlação AR1, permutável e não estruturada, respectivamente.

É notável que existem observações discrepantes no conjunto de dados e estas são destacadas nos gráficos da distância de Cook e dos resíduos de Pearson padronizados para todas as estruturas de correlação consideradas (Figura 6.5).

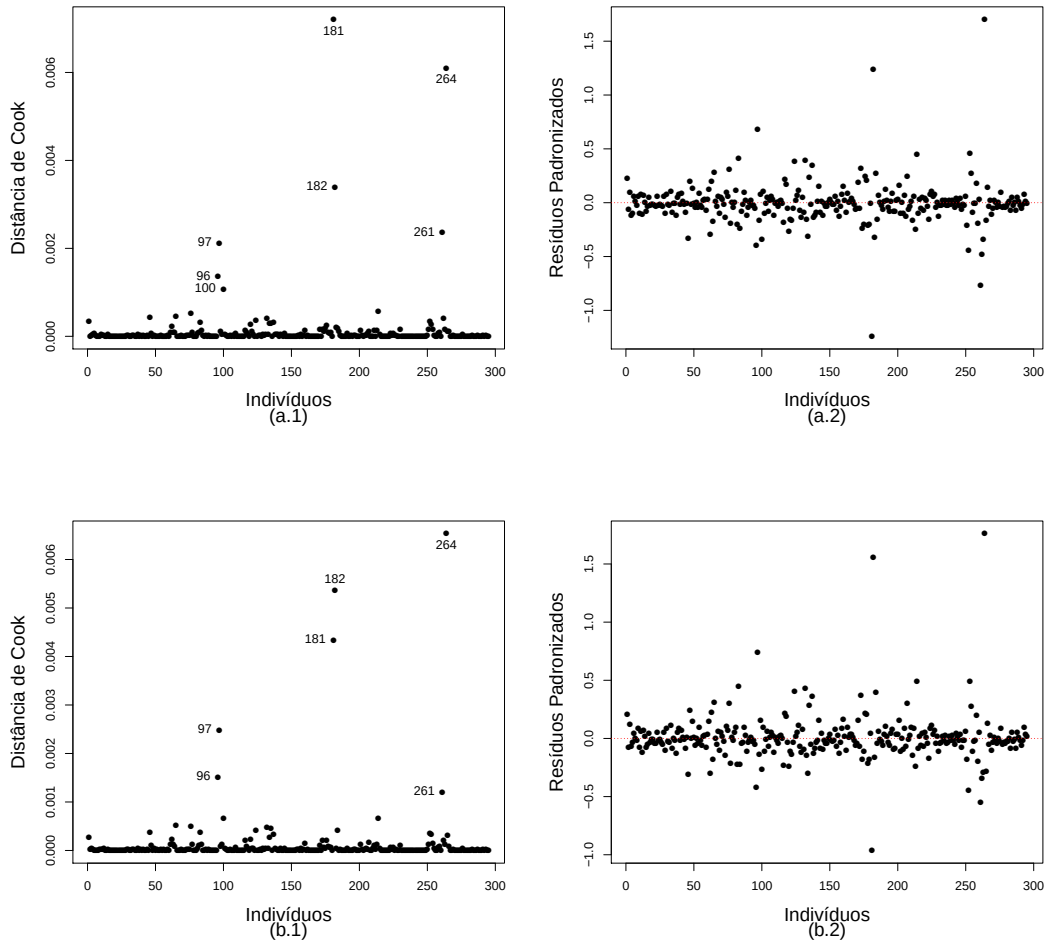


Figura 6.2: Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson multiníveis (a.1 e a.2) com intercepto aleatório e (b.1 e b.2) com dois efeitos aleatórios.

A análise de diagnóstico foi também realizada para o modelo BN marginal via cópulas Gaussianas e os resultados encontram-se na Figura 6.6. Novamente verificou-se que há pontos influentes no conjunto de dados através dos gráficos da distância de Cook e dos resíduos de Pearson Padronizados para todas as estruturas de correlação. Além disso, há uma maior variação nos valores da distância de Cook e dos resíduos de Pearson padronizados quando se considera distribuição BN marginal.

Vale salientar que a maioria dos pontos foram destacados em todos os gráficos, independentemente do modelo que foi ajustado. Isto é uma forte evidência de que estes pontos são realmente atípicos e devem ser retirados do conjunto de dados para um novo ajuste.

O gráfico de envelope simulado tem por finalidade verificar a adequabilidade

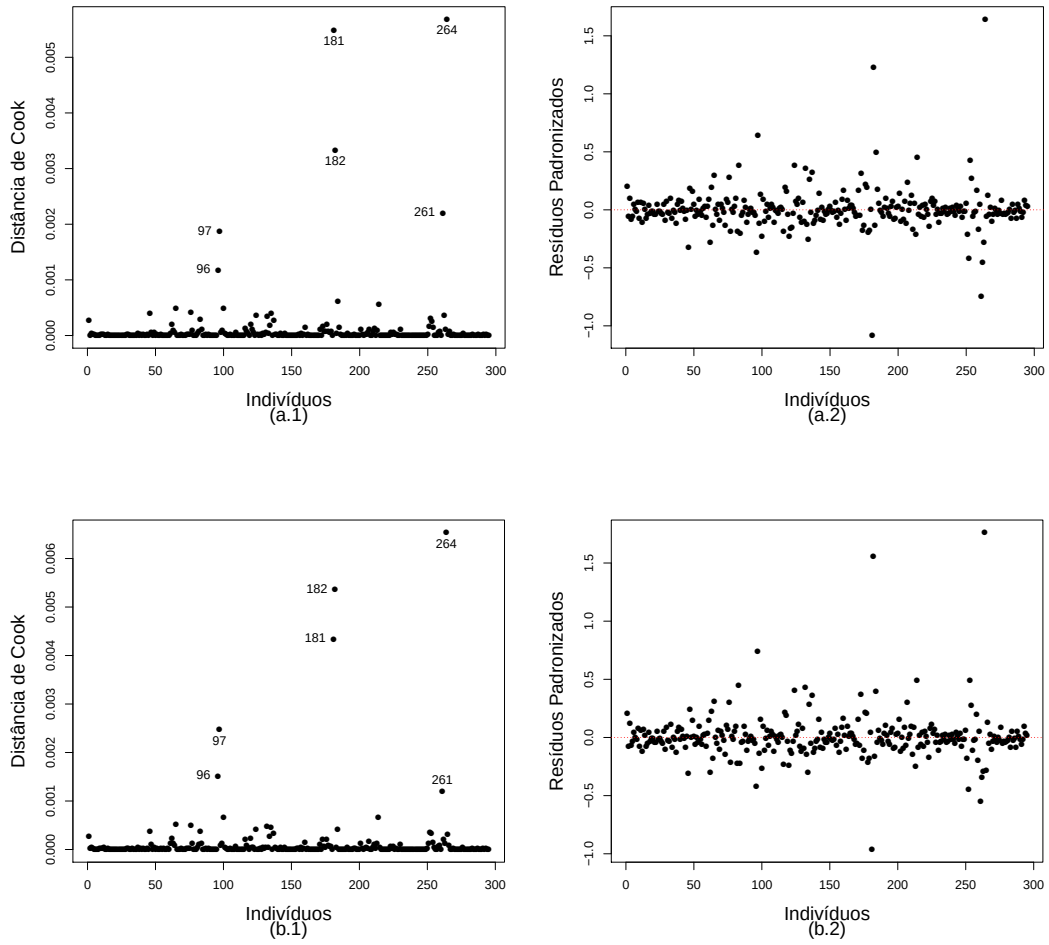


Figura 6.3: Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos BN multiníveis (a.1 e a.2) com intercepto aleatório e (b.1 e b.2) com dois efeitos aleatórios.

do modelo ajustado. Se muitos pontos estiverem fora do intervalo de credibilidade o ajuste não é recomendado. A Figura 6.7 apresenta o envelope simulado para o modelo Poisson marginal via cópulas Gaussianas considerando diferentes matrizes de correlação de “trabalho”. Verificou-se que nenhuma das estruturas de correlação consideradas são adequadas para estes dados, uma vez que a maioria dos pontos encontram-se fora do intervalo de credibilidade.

A Figura 6.8 apresenta o envelope simulado para o modelo BN marginal via cópulas Gaussianas. Verificou-se que houve uma melhora substancial quando se compara com a Figura 6.7, no entanto nenhuma das estruturas de correlação consideradas são adequadas para estes dados.

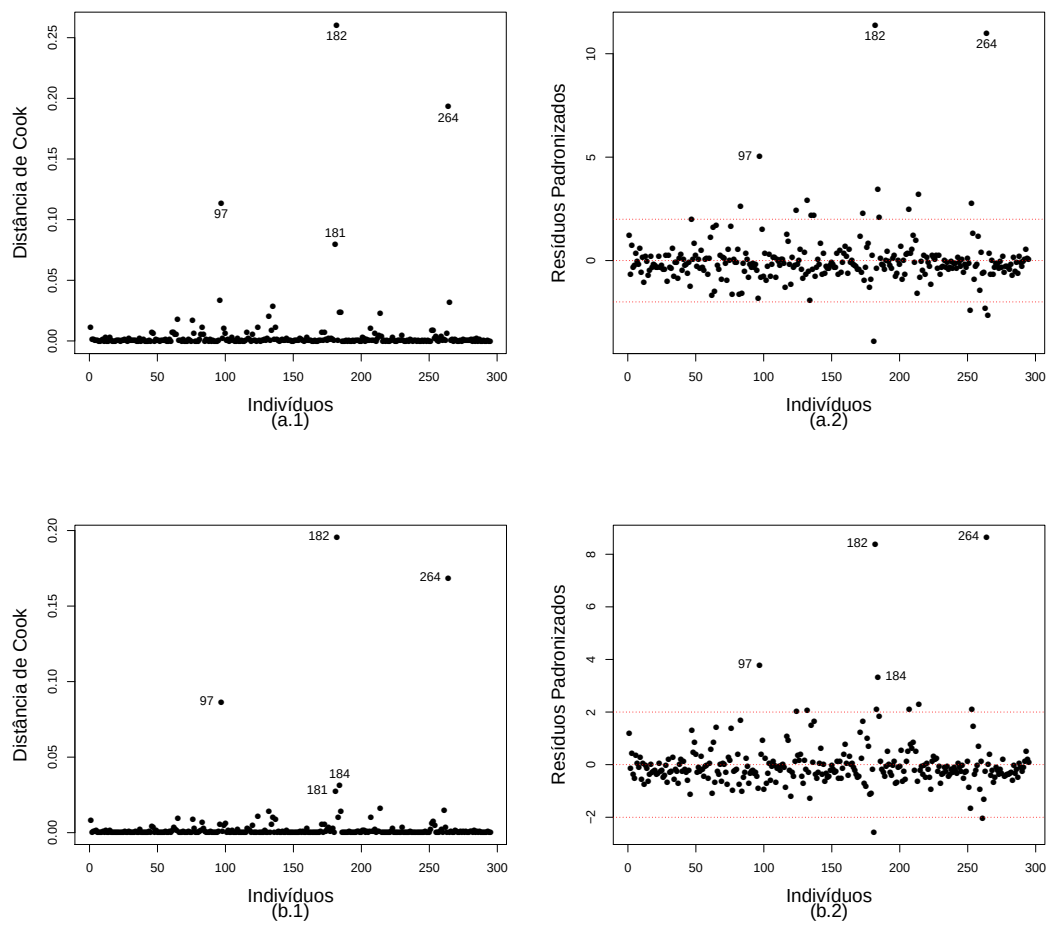


Figura 6.4: Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson marginais, segundo a abordagem EEG, com estrutura de correlação (a.1 e a.2) AR1 e (b.1 e b.2) permutável.

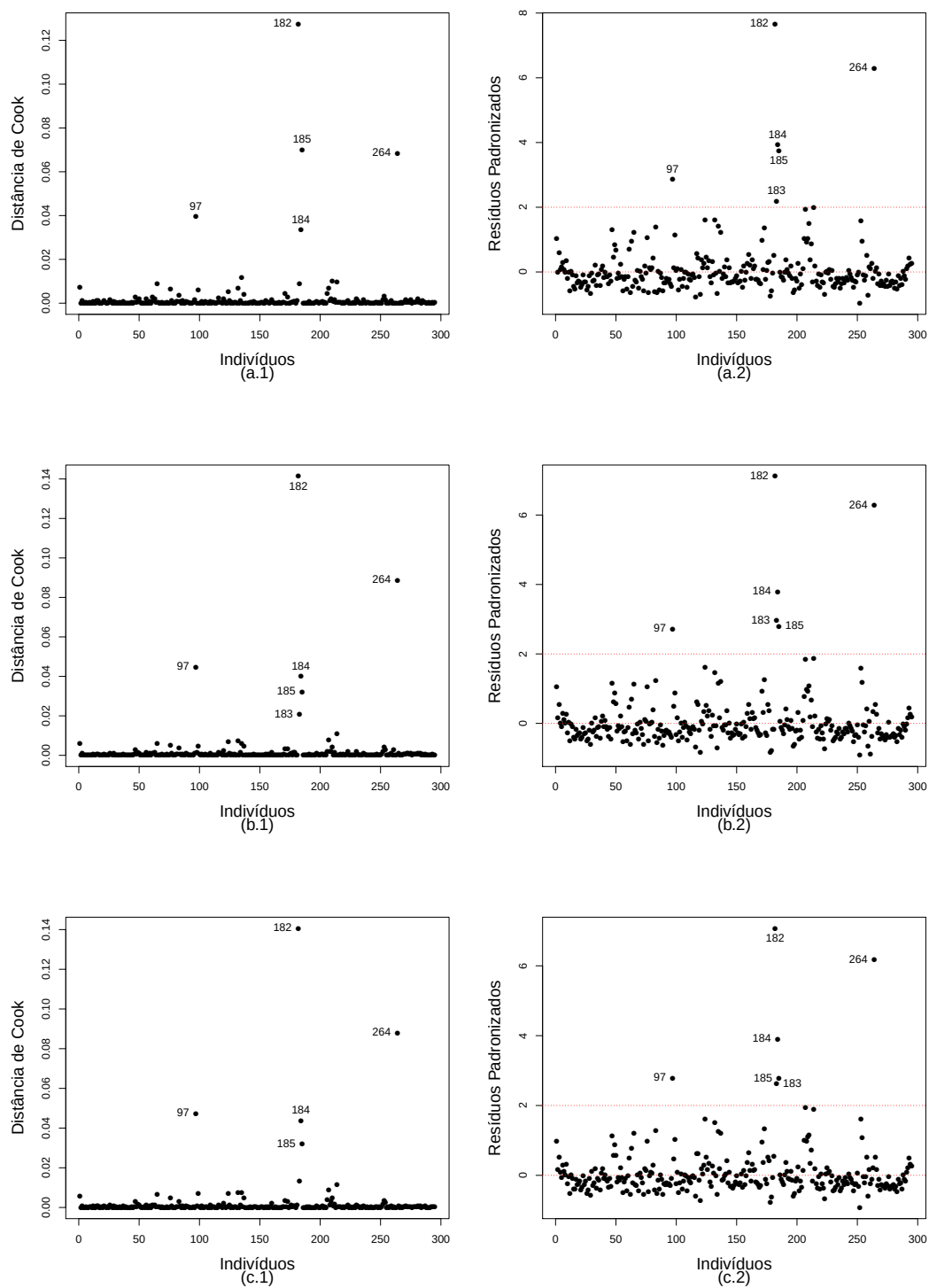


Figura 6.5: Gráficos com a distância de Cook e os resíduos de Pearson padronizados do modelo Poisson marginal considerando cópulas Gaussianas com estrutura de correlação (a.1 e a.2) autorregressiva de ordem 1 (AR1), (b.1 e b.2) permutável e (c.1 e c.2) não estruturada.

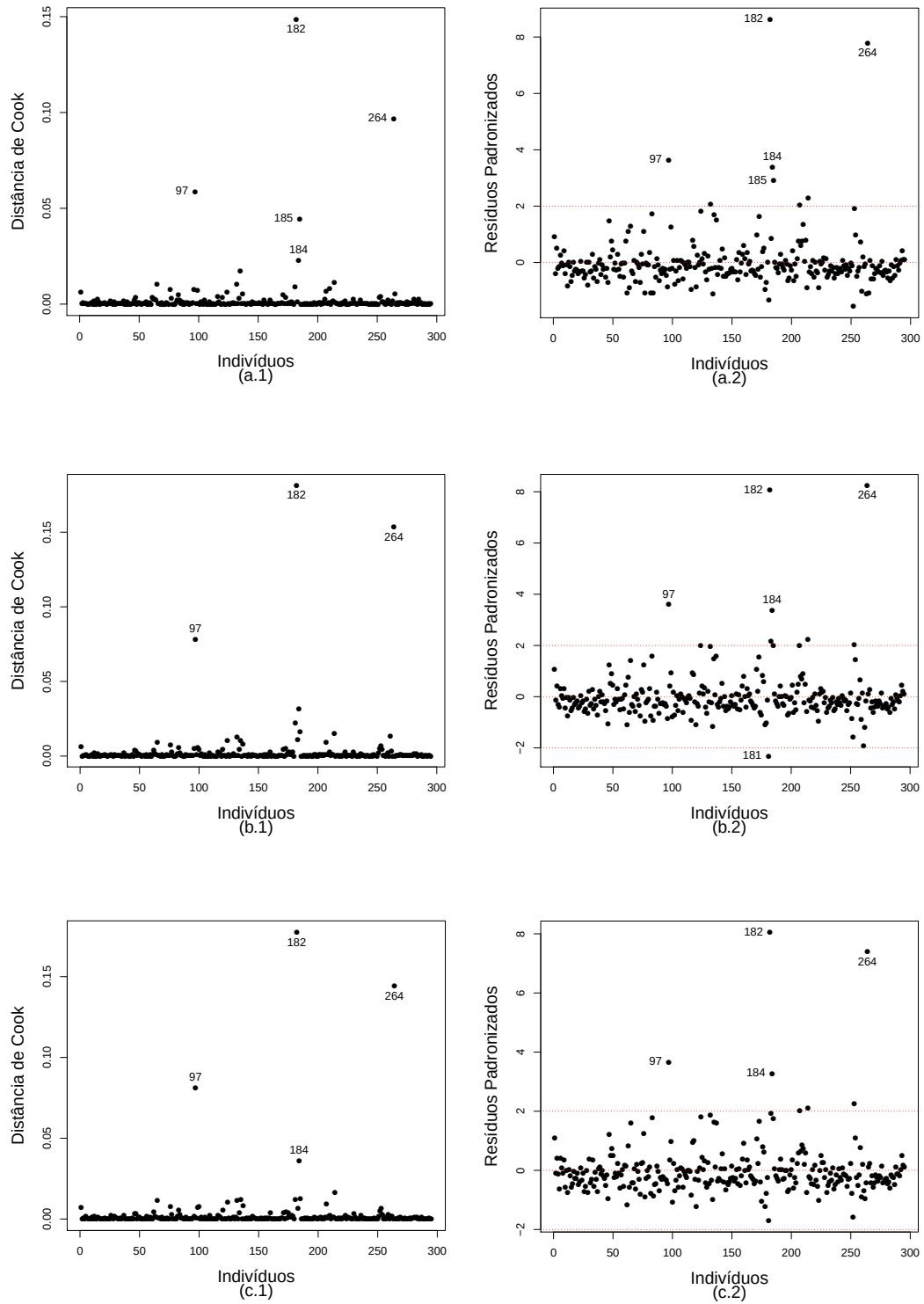


Figura 6.6: Gráficos com a distância de Cook e os resíduos de Pearson padronizados do modelo Binomial Negativo marginal considerando cópulas Gaussianas com estrutura de correlação (a.1 e a.2) autorregressiva de ordem 1 (AR1), (b.1 e b.2) permutável e (c.1 e c.2) não estruturada.

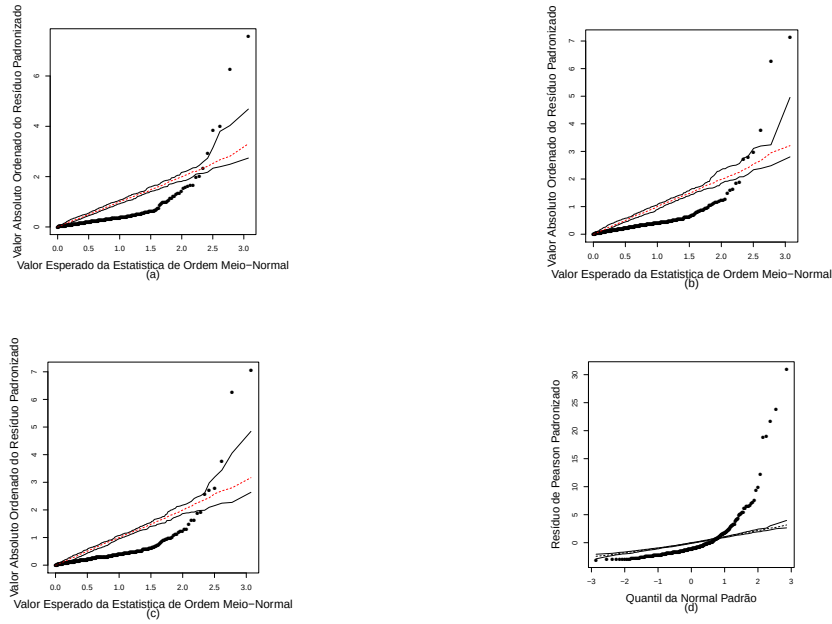


Figura 6.7: Envelopes simulados do modelo Poisson marginal (cópuas Gaussianas) com estruturas de correlação (a) AR1, (b) permutável, (c) não estruturada e (d) independente.

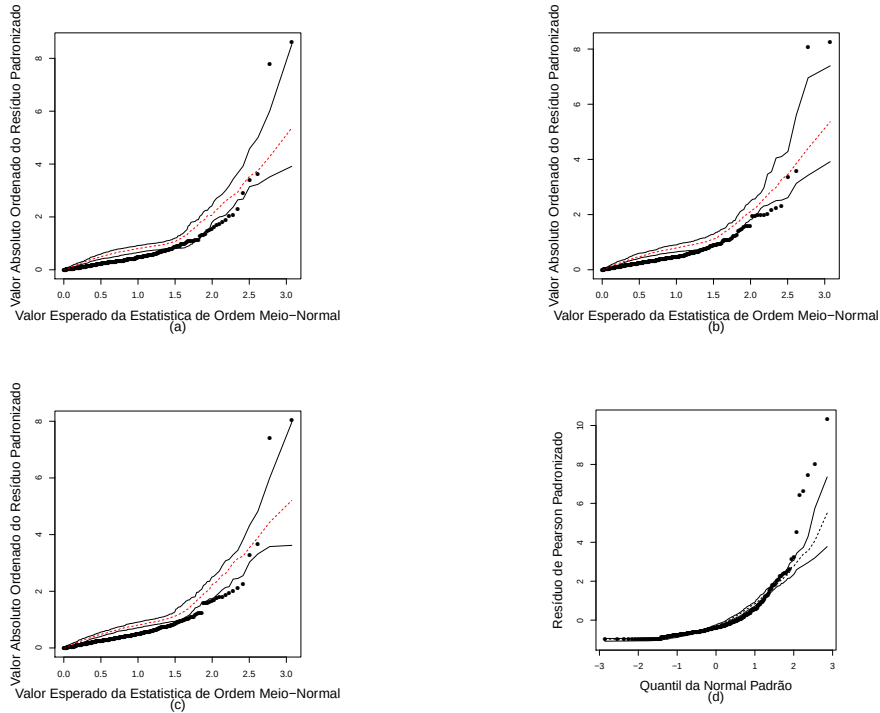


Figura 6.8: Envelopes simulados do modelo BN marginal (cópuas Gaussianas) com estruturas de correlação (a) AR1, (b) permutável, (c) não estruturada e (d) independente.

Foi possível notar que as mesmas observações foram destacadas pelos gráficos da distância de cook e dos resíduos de Pearson padronizados, tanto na distribuição Poisson, quanto na distribuição BN. As observações 97, 181, 182 e 184 e 264 referem-se aos indivíduos 20, 37 e 53. Estes indivíduos foram retirados e uma nova análise foi realizada, sendo os novos resultados apresentados a seguir. É possível notar que houve uma considerável diminuição na variabilidade do número de ataques epilépticos com a retirada dos três pontos influentes, sendo visualizado na Figura 6.9. Além disso, é possível verificar o comportamento global dos indivíduos através do gráfico de perfil médio. Note que há uma diminuição do número médio de ataques epilépticos com o aumento do número de semanas.

A Tabela 6.5 apresenta o número médio de ataques epilépticos e desvio padrão por tratamento. Observou-se que os pacientes que estavam sendo tratados com a droga apresentaram um número médio de ataques epilépticos menor quando comparados com os pacientes que receberam o placebo, exceto na sexta semana. Comparando as Tabelas 6.1 e 6.5, é visível que os três pacientes retirados na análise influenciavam fortemente a média dos ataques de epilepsia.

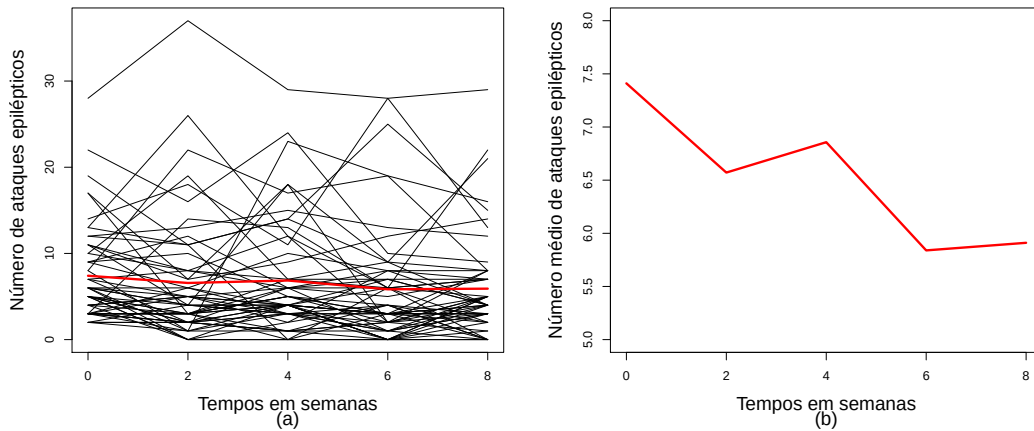


Figura 6.9: Gráficos do (a) perfil individual e (b) perfil médio do número de ataques epilépticos sem três indivíduos com valores discrepantes.

Os modelos selecionados anteriormente foram reajustados e os resultados estão dispostos na Tabela 6.6. Considerando o modelo Poisson multinível com dois efeitos aleatórios (Multinível 2) é possível notar que as variáveis idade e tratamento não foram significantes ao nível de 5% de significância. Em relação ao modelo BN multinível com intercepto aleatório foi observado que a variável tempo é significativa, ao nível de 5% de significância. Através do critério de seleção de modelo pode-se

Tabela 6.5: Número médio de ataques epiléticos e desvio-padrão (DP) por grupo de tratamento.

Tempo em semanas	Placebo (m=26)		Tratamento (m = 30)		Total (m = 56)	
	Média	DP	Média	DP	Média	DP
Baseline	7,62	6,55	7,23	4,31	7,41	5,42
2	7,85	8,25	5,47	5,76	6,57	7,06
4	7,23	7,45	6,53	5,61	6,86	6,47
6	5,65	5,79	6,00	7,37	5,84	6,63
8	7,15	7,06	4,83	4,28	5,91	5,81
Média	7,10	7,02	6,01	5,50	6,50	6,30

concluir que o modelo que ajustou melhor este conjunto de dados foi o modelo BN multinível com intercepto aleatório (AIC = 1483,8).

Tabela 6.6: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão Poisson multinível com dois efeitos aleatórios (Multinível 2) e BN multinível com intercepto aleatório (Multinível 1) para o número de ataques epiléticos sem três indivíduos discrepantes.

Modelos	Abordagem	Parâmetros				AIC
		Intercepto	Tratamento	Idade	Tempo	
Poisson						
	Multinível 2	2,321* (0,461)	-0,107 (0,188)	-0,017 (0,015)	-0,049* (0,011)	1511,0
BN						
	Multinível 1	2,288* (0,492)	-0,169 (0,200)	-0,016 (0,017)	-0,036* (0,011)	1483,8

* p-valor < 5%

A Tabela 6.7 apresenta os modelos Poisson e BN marginais segundo a abordagem EEG. É possível verificar que nem a variável tempo nem a idade foram significantes para o modelo BN marginal, ao nível de 5% de significância. Em relação ao critério de seleção QIC, verificou-se que o melhor modelo para ajustar os dados é o modelo BN marginal EEG com estrutura de correlação AR1 (QIC = 258,0).

Em relação à abordagem marginal via cópulas Gaussianas, os resultados dos ajustes estão disponíveis na Tabela 6.8. É possível verificar que as estimativas dos modelos Poisson e BN nesta abordagem são próximas, mas o modelo BN apresenta o AIC menor (AIC = 1499,1).

As Figuras 6.10, 6.11, 6.12 e 6.13 apresentam os gráficos da distância de Cook e dos resíduos de Pearson padronizados para os modelos selecionados. Em todos os modelos, verificou-se a existência de outros valores discrepantes, porém sem muita influência nos dados.

Tabela 6.7: Estimativas dos coeficientes de regressão e erros-padrão (entre parênteses) usando os modelos de regressão de Poisson e BN na abordagem EEG com estruturas de correlação AR1 para o número de ataques epilépticos sem indivíduos discrepantes.

Modelos	Estrutura de dependência	Parâmetros				QIC
		Intercepto	Tratamento	Idade	Tempo	
Poisson EEG	AR1	2,467 (0,200)*	-0,184 (0,081)*	-0,013 (0,007)	-0,029 (0,008)*	1388,0
	BN EEG	2,500 (0,555)*	-0,213 (0,224)	-0,014 (0,018)	-0,029 (0,022)	258,0

* p-valor < 5%

Tabela 6.8: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão Poisson e BN marginais via cópulas Gaussianas com estrutura de correlação permutável para o número de ataques epilépticos sem indivíduos discrepantes.

Modelos	Estrutura de dependência	Parâmetros				AIC
		Intercepto	Tratamento	Idade	Tempo	
Poisson	Permutável	2,612 (0,204)*	-0,044 (0,078)	-0,023 (0,009)*	-0,043 (0,007)*	1955,5
BN	Permutável	2,581 (0,435)*	-0,126 (0,168)	-0,018 (0,014)	-0,034 (0,012)*	1499,1

* p-valor < 5%

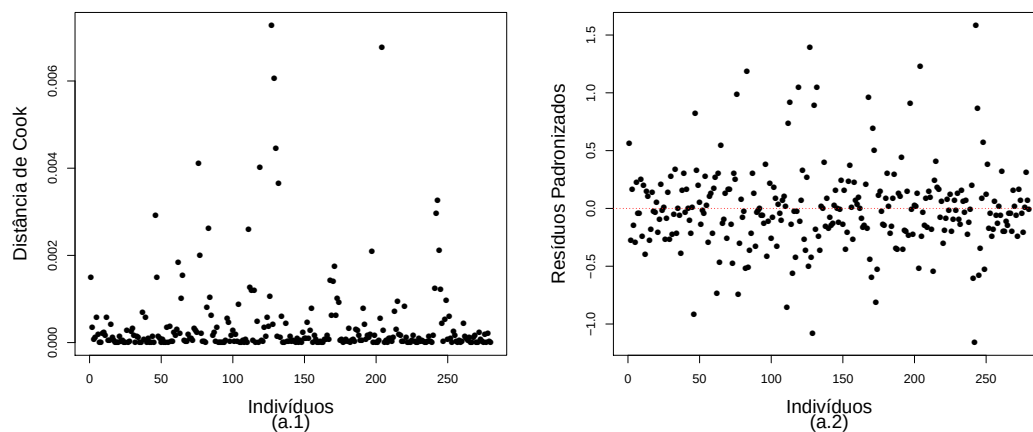


Figura 6.10: Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson multinível (a.1 e a.2) com dois efeitos aleatórios.

A Figura 6.10 apresenta os resultados referentes ao modelo Poisson multinível com dois efeitos aleatórios. Verifica-se que houve um aumento na variabilidade dos resíduos, especialmente para os resíduos de Pearson padronizados.

O mesmo é verificado para o modelo Binomial Negativo multinível com intercepto aleatório (Figura 6.11). Note que os resíduos de Pearson padronizados continuam dentro da banca de confiança, porém os resíduos já não se encontram tão concentrados em zero.

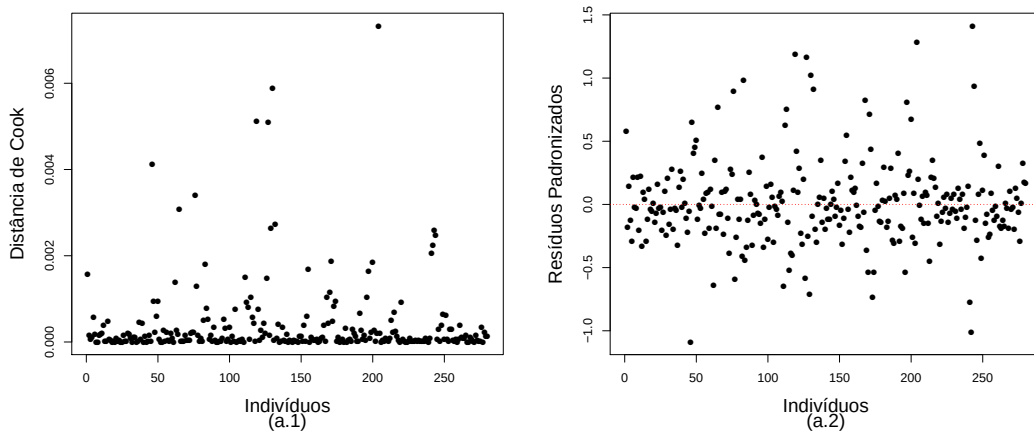


Figura 6.11: Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos BN multinível (a.1 e a.2) com intercepto aleatório.

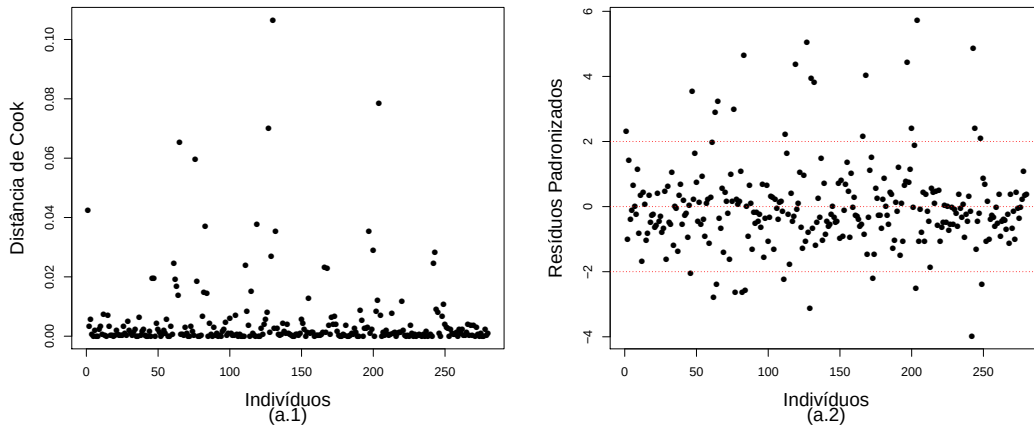


Figura 6.12: Gráficos com a distância de Cook e os resíduos de Pearson padronizados dos modelos Poisson marginal, segundo a abordagem EEG, com estrutura de correlação (a.1 e a.2) AR1.

A Figura 6.12 apresenta os gráficos da distância de Cook e dos resíduos de Pearson padronizados para o modelo Poisson marginal (EEG) com matriz de correlação

AR1. É verificado pontos discrepantes, porém os mesmos não afetam na análise dos dados.

Também foi realizado a análise de diagnóstico para o modelo Binomial Negativo via cópulas Gaussianas com estrutura de correlação permutável e os resultados encontram-se na Figura 6.13. Novamente, pontos discrepantes foram observados, mas não têm tanta influência.

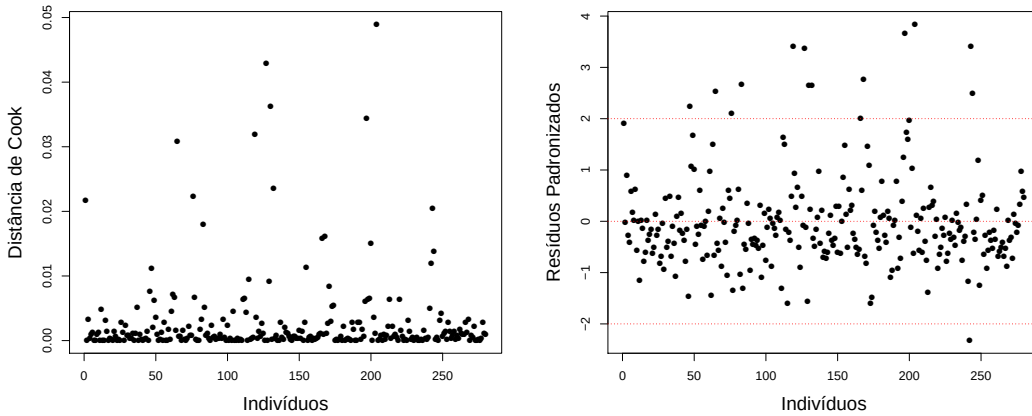


Figura 6.13: Distância de Cook e Resíduos de Pearson Padronizados do modelo BN marginal com cópulas Gaussianas e estrutura de correlação permutável sem os três indivíduos discrepantes.

Analisando a Figura 6.14, observou-se que o gráfico de envelope simulado teve uma melhora substancial quanto à disposição dos pontos no gráfico quando se eliminou as observações discrepantes. Aparentemente, o modelo BN marginal com cópulas Gaussianas e estrutura de correlação permutável é indicado para ajustar o conjunto de dados. Este modelo também aparece como o melhor modelo para o ajuste dos dados segundo o critério AIC, que é apresentado na Tabela 6.8 (AIC = 1499,1).

Em geral, verificou-se que os modelos BN multinível e marginal (EEG e cópulas Gaussianas) ajustaram melhor os dados do que os modelos Poisson para as três abordagens.

Uma observação importante em relação às abordagens marginais utilizadas (EEG e cópulas Gaussianas) é que elas podem levar a diferentes resultados sob diferentes estruturas de correlação.

Além disso, uma vantagem da análise de dados longitudinais via cópulas Gaussianas com respeito a EEG é que os critérios de informação padrão (AIC) podem ser utilizados para selecionar o melhor o modelo.

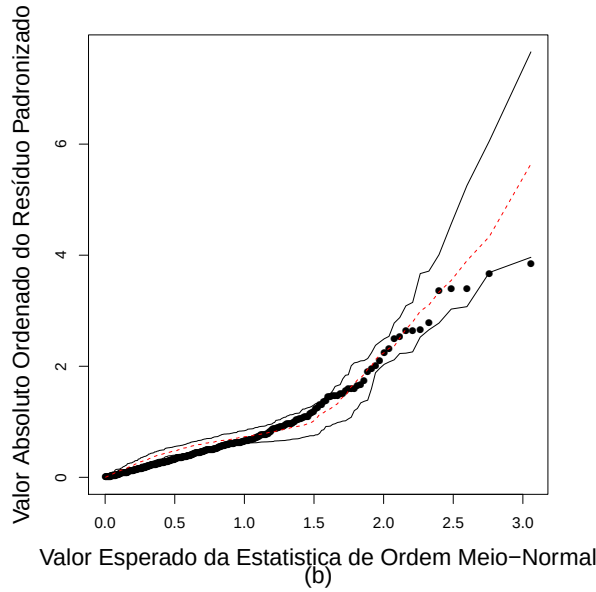


Figura 6.14: Envelope simulado do modelo BN marginal via cópulas Gaussianas e estrutura de correlação permutável sem os três indivíduos discrepantes.

6.2 Aplicação 2: Estudo sobre a reforma no sistema de saúde alemão

Nesta aplicação foram consideradas informações referentes a 6.127 indivíduos que foram acompanhados a fim de avaliar a reforma no sistema de saúde alemão (Rabe-Hesketh & Skrondal, 2005). As seguintes variáveis foram analisadas: (i) gênero do indivíduo (1 = feminino, 0 = masculino), (ii) tempo em que a variável resposta foi mensurada (de 1984 a 1988), (iii) idade no ano da mensuração, (iv) nível educacional, que é definido pelos anos de escolaridade, (v) estado civil (casado: sim = 1, não = 0), (vi) indicador de emprego (1=não trabalha, 0=trabalha) e (vii) número de consultas ao médico (resposta de interesse). A única variável tempo-dependente deste conjunto de dados foi a variável idade. As medidas foram realizadas anualmente, sendo portanto os dados igualmente espaçados. Porém, alguns indivíduos não possuem todas as observações, tornando o conjunto de dados desbalanceados.

Realizando a análise descritiva, verificou-se que 50,1% dos indivíduos são do sexo feminino e que 74,5% são casados. Além disso, 36,8% dos indivíduos não trabalham. O tempo médio de escolaridade é de 11,4 anos (DP = 2,4 anos). Além disso, o número de consultas ao médico variou de 0 a 121 consultas no período do estudo (1984 - 1988).

Para ilustrar o comportamento dos pacientes ao longo do estudo foi construído

o gráfico do perfil individual para dez pacientes selecionados aleatoriamente (Figura 6.15-(a)). O gráfico de perfil médio apresentado na Figura 6.15-(b) apontou uma diminuição no número médio de consultas com o passar dos anos, tanto globalmente quanto por sexo. Verificou-se que as mulheres apresentam uma média maior do número de consultas e também tiveram uma maior diminuição ao longo dos anos.

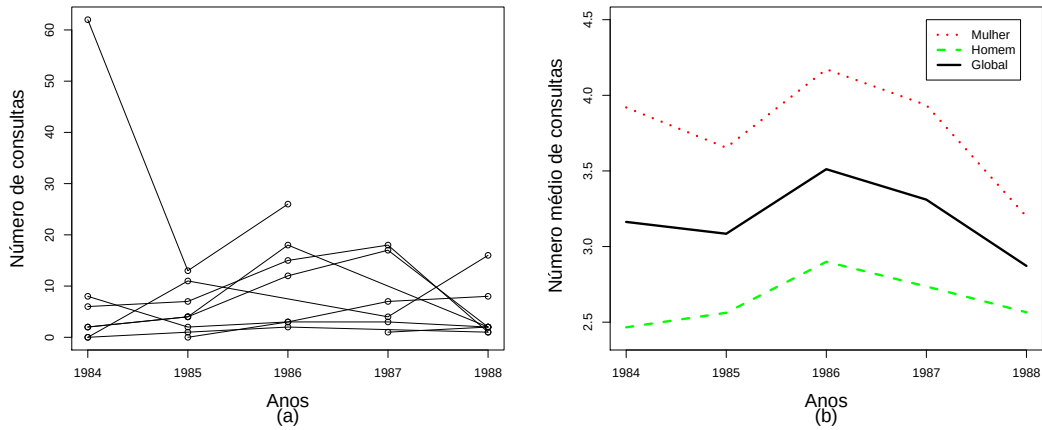


Figura 6.15: Gráficos do (a) perfil individual para dez pacientes e do (b) perfil médio, global e por sexo, do número de consultas ao médico.

Os modelos Poisson independente, multinível com intercepto aleatório (Multinível 1) e multinível com dois efeitos aleatórios (Multinível 2) foram ajustados e os resultados encontram-se na Tabela 6.9. Os efeitos aleatórios no intercepto (b_{0i}) e no tempo (b_{1i}) são distribuídos normalmente. Verificou-se que todas as variáveis foram significativas, ao nível de 5% de significância. Considerando o modelo Poisson multinível com dois interceptos aleatórios, observou-se que a variável tempo tem uma influência positiva no número de consultas ao médico, diferente dos modelos Poisson independente e multinível com intercepto aleatório. Pelo critério AIC, o modelo Multinível 2 é o melhor modelo Poisson para ajustar estes de dados (AIC = 93122,3).

Os modelos Poisson considerando a abordagem EEG também foram ajustados e os resultados encontram-se na Tabela 6.10. Pelo critério QIC é possível notar que o melhor modelo para ajustar os dados é o modelo marginal EEG com estrutura de correlação autorregressiva de ordem 1 (AR1).

O modelo Binomial Negativo considerando as abordagens multiníveis e EEG também foram ajustados. A Tabela 6.11 apresenta os modelos Binomial Negativo independente e multinível Binomial Negativo com intercepto aleatório com distri-

Tabela 6.9: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Poisson independente, multinível com intercepto aleatório e multinível com dois efeitos aleatórios para o número de consultas ao médico.

Variáveis	Modelo Poisson independente	Modelo Poisson Multinível	
		intercepto aleatório	dois efeitos aleatórios
Intercepto	0,569 (0,033)*	-0,272 (0,115)*	-0,583 (0,122)*
Gênero	0,187 (0,009)*	0,380 (0,034)*	0,381 (0,035)*
Tempo	-0,007 (0,003)*	-0,018 (0,003)*	0,033 (0,007)*
Idade	0,019 (0,001)*	0,024 (0,001)*	0,025 (0,001)*
Educação	-0,035 (0,002)*	-0,030 (0,007)*	-0,032 (0,008)*
Estado Civil	-0,090 (0,010)*	-0,133 (0,025)*	-0,144 (0,031)*
Ocupação	0,215 (0,009)*	0,072 (0,018)*	0,106 (0,023)*
AIC	152.652,0	99.881,8	93.122,3

* p-valor < 5%

Tabela 6.10: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando a abordagem EEG para o modelo Poisson marginal com diferentes estruturas de correlação para o número de consultas ao médico.

Variáveis	Modelo Poisson EEG			
	AR1	Permutável	Não Estruturada	Independente
Intercepto	35,784 (7,548)*	23,879 (4,822)*	26,265 (5,012)*	15,180 (5,525)*
Sexo	0,204 (0,013)*	0,206 (0,011)*	0,203 (0,012)*	0,187 (0,009)*
Tempo	-0,018 (0,003)*	-0,012 (0,002)*	-0,013 (0,003)*	-0,007 (0,002)*
Idade	0,018 (0,001)*	0,020 (0,000)*	0,020 (0,000)*	0,019 (0,000)*
Educação	-0,053 (0,003)*	-0,032 (0,003)*	-0,031 (0,002)*	-0,035 (0,002)*
Estado Civil	-0,048 (0,015)*	-0,114 (0,012)*	-0,112 (0,011)*	-0,090 (0,010)*
Ocupação	0,176 (0,013)*	0,172 (0,011)*	0,171 (0,011)*	0,215 (0,009)*
QIC	71.297,1	71.342,5	71.623,0	

* p-valor < 5%

buição Beta. Não foi possível ajustar o modelo Binomial Negativo Multinível com efeito aleatório normal no programa **R**, pois não houve convergência. Os modelos apresentados na Tabela 6.11 foram ajustados no programa estatístico STATA. Pelo critério de seleção de modelo AIC é verificado que o modelo Binomial Negativo multinível com intercepto aleatório ajusta melhor os dados (AIC = 82.478,0).

Os modelos Binomial Negativo marginal com diferentes estruturas de correlação

Tabela 6.11: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Binomial Negativo tradicional e Binomial Negativo multinível com intercepto aleatório para o número de consultas ao médico.

Variáveis	Modelo BN Tradicional	Modelo BN Multinível
		Intercepto Aleatório (Beta)
Intercepto	2,843 (15,506)	-20,196 (11,120)
Gênero	0,238 (0,024)*	0,355 (0,027)*
Tempo	-0,001 (0,008)	0,010 (0,006)
Idade	0,020 (0,001)*	0,019 (0,001)*
Educação	-0,042 (0,005)*	-0,013 (0,006)*
Estado Civil	-0,091 (0,027)*	-0,007 (0,028)*
Ocupação	0,203 (0,026)*	0,144 (0,025)*
AIC	85.715,1	82.478,0

* p-valor < 5%

na abordagem EEG também foram ajustados e os resultados encontram-se na Tabela 6.12. Verificou-se que a variável estado civil foi significativa apenas quando se considera a estrutura de correlação autorregressiva de ordem 1. Segundo o critério QIC, a melhor estrutura de correlação é a permutável (QIC = 31.524,1).

Tabela 6.12: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando a abordagem EEG para o modelo Binomial Negativo com diferentes estruturas de correlação para o número de consultas ao médico.

Variáveis	Modelo Binomial Negativo EEG			
	AR1	Permutável	Não Estruturada	Independente
Intercepto	24,007 (15,722)	14,227 (10,041)	16,546 (10,401)	4,657 (11,411)*
Gênero	0,251 (0,028)*	0,245 (0,023)*	0,242 (0,024)*	0,231 (0,019)*
Tempo	-0,012 (0,008)*	-0,006 (0,005)*	-0,008 (0,005)*	-0,002 (0,006)*
Idade	0,018 (0,001)*	0,020 (0,001)*	0,020 (0,001)*	0,020 (0,001)*
Educação	-0,061 (0,006)*	-0,038 (0,005)*	-0,037 (0,005)*	-0,041 (0,004)*
Estado Civil	-0,029 (0,0302)	-0,114 (0,025)*	-0,113 (0,025)*	-0,091 (0,020)*
Ocupação	0,167 (0,027)*	0,164 (0,022)*	0,163 (0,022)*	0,204 (0,019)*
QIC	31.590,0	31.524,1	31.524,9	

* p-valor < 5%

Pode-se concluir que a distribuição BN ajusta melhor os dados do que a distribuição Poisson para as duas abordagens consideradas. Os modelos Poisson e Binomial

Negativo marginais via cópulas Gaussianas não convergiram.

6.3 Aplicação 3: Contagem de células CD4+ em pacientes portadores do vírus HIV residentes em Salvador-Bahia

O objetivo da terceira aplicação é modelar a contagem de células CD4+ em pacientes portadores do vírus HIV que residem em Salvador-Bahia no período de janeiro/2002 a agosto/2012. Estes dados são provenientes do Sistema de Controle de Exames Laboratoriais (SISCEL) do Ministério da Saúde. O estudo é desbalanceado e desigualmente espaçado, ou seja, os indivíduos não possuem o mesmo número de medidas e as mensurações não ocorreram nos mesmos instantes de tempo. Os 5116 portadores de HIV foram acompanhados desde sua entrada no SISCEL até 2012. O número de observações por paciente variou de 1 a 24.

As variáveis analisadas foram: (i) o número de células CD4+ no tempo t , (ii) o sexo do paciente (0 = masculino, 1 = feminino), (iii) o indicador que informa se o paciente esteve em tratamento durante o período do estudo (0 = não esteve em tratamento, 1 = esteve em tratamento). Esta variável não varia no tempo, (iv) o número de células CD4+ no momento do cadastro do paciente (denotada por CD4-cadastro) e (v) a carga viral no momento do cadastro do paciente (CV-cadastro).

Através da análise descritiva verificou-se que apenas 1715 (33,52%) pacientes não estiveram em tratamento no período de acompanhamento. Também é possível observar que a maioria dos pacientes (59,58%) dos paciente é do sexo masculino.

De acordo com a literatura, o número máximo de células CD4+ de um indivíduo portador do vírus HIV é 1500 células (www.aids.gov.com.br). Assim, valores superiores a este critério foram retirados do conjunto de dados. Para evitar problemas computacionais, as variáveis foram padronizadas e nomeadas de CD4-cadastro padronizado e CV-cadastro padronizado. A Tabela 6.13 apresenta a sumarização das variáveis contínuas do conjunto de dados. É possível verificar ainda superdispersão na variável de interesse, pois a variância do número de células de CD4+ é muito maior do que a sua média.

O gráfico de perfis individuais para 10 pacientes escolhidos aleatoriamente é apresentado na Figura 6.16. Para a construção desse gráfico foi considerado o tempo em que a medida foi mensurada para cada indivíduo. Esta variável tempo foi construída da seguinte forma: considerou-se que a primeira mensuração do indivíduo seria o *baseline* ($t = 0$) e as outras medidas seriam baseadas no mesmo (em anos). Por

Tabela 6.13: Sumarização das variáveis pertencentes ao estudo sobre contagem de células CD4+ para pacientes residentes em Salvador-Bahia.

Variáveis	Média	Desvio-padrão	Mínimo	Máximo
CD4+	500,607	289,191	0	1500
CD4-cadastro	324,931	326,225	0	6000,000
CV-cadastro	277320,500	1085396,000	0	$2,10^{10}$
CD4-cadastro padronizado	-0,015	0,970	-0,996	17,396
CV-cadastro padronizado	0,002	1,007	-0,256	19,092

exemplo, se o indivíduo foi mensurado em fev/2002 ($t = 0$) e a segunda medida foi em abril/2003, então o tempo é dado por $t = 1,2$ anos.

Apesar da superdispersão ter sido identificada na análise descritiva foram ajustados, além do modelo Binomial Negativo (BN), o modelo Poisson para comparação dos resultados.

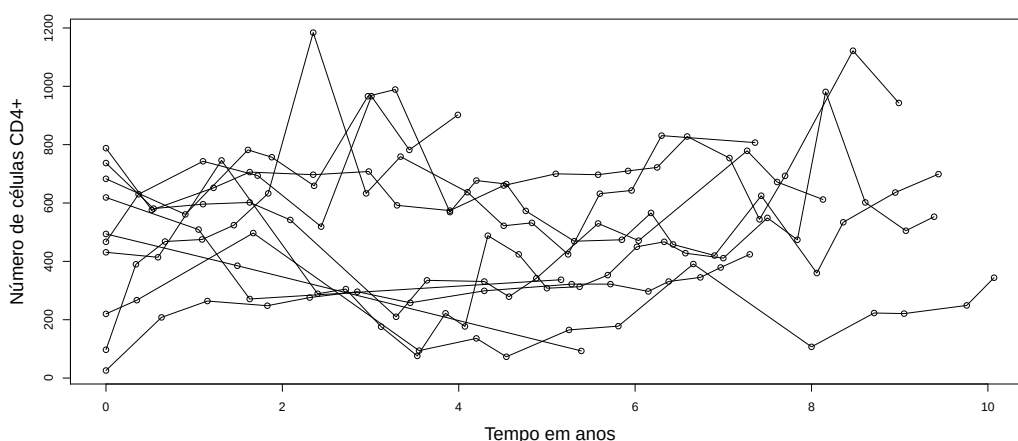


Figura 6.16: Gráfico de perfis individuais para o número de células CD4+ em 10 indivíduos selecionados aleatoriamente.

A Tabela 6.14 apresenta os resultados dos ajustes para os modelos Poisson independente, Poisson multinível com intercepto aleatório e Poisson multinível com dois efeitos aleatórios. Pode-se verificar que a carga viral mensurada no cadastro e padronizada (CV-cadastro padronizado) só é significativa para o modelo Poisson independente, ao nível de 5% de significância. Pelo critério AIC, tem-se que o melhor modelo para ajustar os dados, considerando a distribuição Poisson, é o modelo Poisson multinível com dois efeitos aleatórios ($AIC = 2.421.835,0$).

A Tabela 6.15 apresenta os ajustes dos modelos EEG considerando a distribuição

Tabela 6.14: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Poisson independente, multinível com intercepto aleatório e multinível com dois efeitos aleatórios para o número de células CD4+ para pacientes residentes em Salvador-Bahia.

Variáveis	Modelo Poisson independente	Modelo Poisson Multinível	
		intercepto aleatório	dois efeitos aleatórios
Intercepto	5,817 (0,001)*	5,332 (0,070)*	5,370 (0,070)*
Tempo	0,0469 (0,000)*	0,045 (0,000)*	0,043 (0,004)*
Gênero	0,073 (0,000)*	0,099 (0,228)*	0,111 (0,023)*
Tratamento	-0,056 (0,000)*	0,228 (0,245)*	0,181 (0,025)*
CD4-cadastro padronizado	0,000 (0,000)*	0,001 (0,000)*	0,001 (0,000)*
CV-cadastro padronizado	-0,000 (0,000)*	-0,000 (0,000)	-0,000 (0,000)
AIC	8.590.068,0	3.559.437,0	2.421.835,0

* p-valor 5%

Poisson. Verifica-se que todas as variáveis foram significantes, ao nível de 5% de significância, em todos os modelos ajustados. Pelo critério QIC, tem-se que o melhor modelo para ajustar os dados é o modelo Poisson marginal EEG com estrutura de correlação permutável ($QIC = 1.435.573,9$). Não houve convergência para o modelo Poisson marginal via cópulas Gaussianas.

Tabela 6.15: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) para o modelo Poisson marginal usando a abordagem EEG com diferentes estruturas de correlação para o número de células CD4+ para pacientes residentes em Salvador-Bahia.

Variáveis	Modelo Poisson EEG			
	AR1	Permutável	Independente	Não Estruturada
Intercepto	5,859 (0,005)*	5,621 (0,006)*	5,900 (0,002)*	5,701 (0,005)*
Tempo	0,029 (0,000)*	0,030 (0,000)*	0,034 (0,000)*	0,026 (0,000)*
Gênero	0,056 (0,002)*	0,084 (0,003)*	0,074 (0,001)*	0,076 (0,003)*
Tratamento	0,163 (0,005)*	0,452 (0,006)*	0,164 (0,002)*	0,356 (0,005)*
CD4-cadastro padronizado	0,074 (0,001)*	0,071 (0,001)*	0,083 (0,000)*	0,089 (0,001)*
CV-cadastro padronizado	-0,019 (0,001)*	-0,030 (0,002)*	-0,021 (0,001)*	-0,039 (0,002)*
QIC	1.456.184,3	1.435.573,9		1.445.648,1

* p-valor 5%

Não foi possível ajustar os modelos multiníveis Binomial Negativo com intercepto aleatório e com dois efeitos aleatórios com distribuição normal, pois não houve con-

vergência. A Tabela 6.16 apresenta os resultados do ajuste do modelo Binomial Negativo independente e o modelo Binomial Negativo com intercepto aleatório com distribuição Beta. Verificou-se que a variável gênero só foi significativa para o modelo Binomial Negativo independente. Considerando o critério de avaliação AIC, o modelo que ajusta melhor os dados é o modelo Binomial Negativo multinível com intercepto aleatório ($AIC = 117.588,8$).

Tabela 6.16: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando os modelos de regressão de Binomial Negativo independente e Binomial Negativo multinível com intercepto aleatório (Beta) para o número de células CD4+ para pacientes residentes em Salvador-Bahia.

Variáveis	Modelo BN independente	Modelo BN Multinível intercepto aleatório (Beta)
Intercepto	5,870 (0,028)*	1,165 (0,052)*
Tempo	0,037 (0,002)*	0,026 (0,002)*
Gênero	0,057 (0,014)*	0,041 (0,028)
Tratamento	0,176 (0,028)*	0,412 (0,051)*
CD4-cadastro padronizado	0,179 (0,009)*	0,098 (0,009)*
CV-cadastro padronizado	-0,012 (0,007)	-0,020 (0,012)
AIC	122.338,8	117.588,8

* p-valor < 5%

Também foi considerada a abordagem EEG no modelo Binomial Negativo e os resultados encontram-se na Tabela 6.17.

Tabela 6.17: Estimativas dos coeficientes de regressão e erros padrão (entre parênteses) usando a abordagem EEG para o modelo Binomial Negativo com diferentes estruturas de correlação para o número de células CD4+ para pacientes residentes em Salvador-Bahia.

Variáveis	Modelo EEG Binomial Negativo		
	AR1	Permutável	Independente
Intercepto	5,836 (0,094)*	5,671 (0,098)*	5,870 (0,045)*
Tempo	0,032 (0,006)*	0,031 (0,003)*	0,037 (0,004)*
Gênero	0,042 (0,053)	0,050 (0,064)	0,057 (0,022)*
Tratamento	0,170 (0,094)*	0,384 (0,101)*	0,176 (0,044)*
CD4-cadastro padronizado	0,190 (0,025)*	0,189 (0,027)*	0,179 (0,010)
CV-cadastro padronizado	-0,006 (0,023)	-0,018 (0,028)	-0,012 (0,011)*
QIC	3.771,0	3.736,8	

* p-valor < 5%

Não foi possível obter convergência para o modelo Binomial Negativo com estrutura de correlação não estruturada. Observa-se também que a variável CD4-cadastro padronizado não foi significativa para o modelo com estrutura de correlação independente. Pelo critério QIC, tem-se que melhor estrutura de correlação é a permutável ($\text{QIC} = 3.736,8$).

Não foi possível ajustar os modelos Binomial Negativo marginais via cópulas Gaussianas considerando as diferentes estruturas de correlação, pois não houve convergência. Não foi possível realizar a análise de diagnóstico para esse conjunto de dados, o volume de informações é muito extenso e necessita de uma memória computacional muito grande para alocar todas as informações necessárias para a construção dos gráficos avaliados neste trabalho.

Capítulo 7

Conclusões

A modelagem para dados de contagem é bastante utilizada em diversas áreas do conhecimento, como nas ciências biológicas, educação e saúde pública, bem como na agricultura, economia e marketing, pois nelas podem ser estudadas as variações ao longo do tempo de uma determinada variável de interesse. O modelo frequentemente usado para analisar dados de contagem é o modelo de Poisson. No entanto quando os dados apresentam superdispersão um modelo indicado para ajustar os dados é o modelo Binomial Negativo.

Duas abordagens de regressão comumente utilizadas para analisar dados longitudinais são os modelos condicionais e os marginais. É importante considerar as metodologias apropriadas na análise de dados longitudinais, pois a correlação entre as respostas dentro das unidades amostrais e a ordenação cronológica das respostas devem ser consideradas nos modelos.

Para o modelo Binomial Negativo marginal foram considerados dois tipos de modelagem, as Equações de Estimação Generalizadas (EEG) e as cópulas Gaussianas. A abordagem EEG requer apenas a especificação das médias marginais, além disso, as EEG produzem estimativas consistentes dos coeficientes de regressão e dos erros-padrão mesmo sob uma má especificação da estrutura de correlação. Já as cópulas Gaussianas utilizam verossimilhança e possibilitam a utilização do critério de informação AIC para a seleção do modelo.

As abordagens multinível e marginal (EEG e cópulas Gaussianas) foram aplicadas a dados de contagem longitudinais, com aplicações a dados reais e realização de estudos de simulação.

Com os estudos de simulação pode-se verificar a diminuição dos valores do viés médio (VM), dos valores absolutos do viés relativo ($|VR|$), valores da raiz do erro quadrático médio ($REQM$) e os valores da diferença relativa ($diff.rel$) entre o desvio

padrão (DP) e erro padrão (EP) com o aumento do número de medidas repetidas.

Vale ressaltar que a variabilidade de b_{0i} afeta substancialmente o VM e a $REQM$ obtidos no modelo Poisson independente. Além disso, a performance do estimador da variância ($dif.rel$) está associada ao tipo de covariável considerada.

Em relação ao critério de seleção AIC, verificou-se que o modelo utilizado para a geração dos dados foi escolhido como o melhor modelo para o ajuste dos dados. Além disso, o critério de seleção AIC é consistente e não sofreu influência da variabilidade de b_{0i} .

Para o modelo Poisson marginal considerando a abordagem EEG, verificou-se que a estrutura de correlação independente apresentou a pior performance. Em relação ao critério de seleção QIC, notou-se que os valores são bem parecidos para todas as estruturas de correlação. Isto pode ser justificado pelo fato dos modelos marginais (EEG) produzirem estimativas consistentes dos coeficientes de regressão e dos erros-padrão mesmo sob uma má especificação da matriz de correlação de “trabalho”.

Para os modelos considerando a distribuição Binomial Negativa, verificou-se que os modelos multiníveis são computacionalmente intensivo. Não houve convergência para os dados gerados considerando o modelo BN multinível com dois efeitos aleatórios.

A abordagem EEG não está implementada na linguagem de programação **R**. Logo, uma solução encontrada para ajustar o modelo BN marginal foi utilizar cópulas Gaussianas. Esta abordagem também é computacionalmente intensiva e apresenta sérios problemas de convergência sendo limitado em relação ao número de indivíduos e ao número de medidas repetidas. Portanto, o modelo BN marginal via cópulas Gaussianas não é indicado para cenários mais complexos.

Em relação às aplicações, observou-se que para os conjuntos de dados mais complexos não foi possível ajustar os modelos marginais via cópulas Gaussianas, pois os mesmos não convergiram.

Em relação à análise de diagnóstico, verificou-se que existe na literatura materiais disponíveis para dados com medidas repetidas quando a variável resposta é contínua. No entanto, para dados de contagem longitudinal há pouco material disponível na literatura.

Para estudos futuros, ainda há bastante a ser explorado, principalmente quando se trata de técnicas de diagnóstico para análise de dados longitudinais. Desta forma, temos como proposta futura estudar as propriedades das medidas de diagnóstico

para as distribuições Poisson e BN, no contexto de dados longitudinais.

Além disso, é necessário estudar o comportamento dos estimadores em cenários mais complexos através de estudos de simulação. Portanto é necessário considerar conjuntos com dados desbalanceados e/ou desigualmente espaçados, covariáveis tempo-dependentes, presença de pontos aberrantes, por exemplo.

Referências

- [1] BARNETT, A. G., KOPER, N., DOBSON, A. J., SCHMIEGELOW, F & MANSEAU, M. (2010). Using Information Criteria to Select the Correct Variance Covariance Structure for Longitudinal Data in Ecology. *Methods in Ecology and Evolution*, **1**, 15-24.
- [2] BARRON, D. N. (1992). The Analysis of Count Data. Overdispersion and Autocorrelation. *American Sociological Association*.
- [3] BOECK, P. & WILSON, M. (2004). *Explanatory Item Response Models*. Springer.
- [4] BOZDOGAN, H. (1987). Model Selection and Akaike's Information Criterion (AIC): The General Theory and its Analytical Extensions. *Psychometrika*, **52**, 345-370.
- [5] CAMERON, A. C. & TRIVEDI, P. K. (2013). *Regression Analysis of Count Data*. Second Edition. Cambridge.
- [6] CHANG, Y.-C. (2000). Residuals Analysis of the Generalized Linear Modelos for Longitudinal Data. *Statistics in Medicine*, **19**, 1277-1293.
- [7] COOK, R. D. (1977). Detection of Influential Observation in Linear Regression. *Technometrics*, **19**, 15-18.
- [8] COX, D. R. & HINKLEY, D. V. (1979). *Theoretical statistics*. Chapman and Hall.
- [9] COX, D. R. & SNELL, E. J. (1968). A general Definition of Residuals (with discussion). *Journal Royal Statistical Society B*, **30**, 248-275.

- [10] CRIBARI-NETO, F, & GOIS, M. C. A. (2002). Uma Análise de Monte Carlo do Desempenho de Estimadores de Matrizes de Covariância sob Heterocedasticidade de Forma Desconhecida. *RBE*, **56**, 309-334.
- [11] CUI, J. (2007). QIC Program and Model Selection in GEE Analyses. *The STATA Journal*, **7**, 209-220.
- [12] CUI, J. (2010). Information Criterion and Computation for GEE Model Selection, with an Application to a Longitudinal Study of Skin Cancer. *Biometrics: Methods, Applications and Analysis*, 167-175.
- [13] DRAPER, N. R. & SMITH, H. (1988). *Applied Regression Analysis*. 3rd Edition. New York: John Wiley & Sons.
- [14] DIGGLE, P. J., HEAGERTY, P., LIANG KY. & ZEGGER, S. (2002). *Analysis of Longitudinal Data*. Oxford.
- [15] DIGGLE, D.J., LIANG K.Y & ZEGGER, S.L. (1994). *Analysis Longitudinal Data*. Oxford.
- [16] FITZMAURICE, G., DAVIDIAN, M., VERBEKE, G. & MOLENBERGHS G. (2009). *Longitudinal Data Analysis*. Chapman & Hall/CRC.
- [17] FREITAS, E. G (2007). *Análise de Dados Longitudinais em Experimentos com cana-de-açúcar*. Dissertação de Mestrado. Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba.
- [18] GARDNER, W., MULVEY, E. P. & SHAW, E. C. (1995). Regression Analyses of Counts and Rates: Poisson, Overdispersed Poisson, and Negative Binomial Models. *Psychological Bulletin*, **118**, 392-404.
- [19] GUPTA, P. L., GUPTA, R. C. & TRIPATHI, R. C. (1996). Analysis of zero-adjusted count data. *Computational Statistics & Data Analysis*. **23**, 207-218.
- [20] HARVILLE, D. (1976). Extension of the Gauss-Markov Theorem to Include the Estimation of Random Effects. *The Annals of Statistics*, **4**, 384-395.
- [21] HARVILLE, D. (1977). Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *Journal of the American Statistical Association*, **72**, 320-338.
- [22] HEDEKER, D. & GIBBONS, R. D. (2006). *Longitudinal data analysis*. WILEY

- [23] HENDERSON, C.R; KEMPTHORNE, O & SEARLE, S.R. (1959). The Estimation of Environmental and Genetic Trends from Records Subject to Culling. *Biometrics*, **15**, 192-193.
- [24] HENDERSON, C. R. (1975). Best Linear Unbiased Estimation and Prediction under a Selection Model. *Biometrics*, **31**, 423-449.
- [25] HILDEN-MINTON, J. A. (1995). *Multilevel Diagnostics for Mixed and Hierarchical Linear Models*. PhD Thesis. University of California, Los Angeles.
- [26] HINDE, J. & DEMÉTRIO, C. G. B. (1998). Overdispersion model and estimation. *Computational Statistics and Data Analysis*, **27**, 151-170.
- [27] HOFFMAN, R. & VIEIRA, S. (1988). *Análise de Regressão: Uma Introdução à Econometria*. 3ª Edição. São Paulo: Editora Hucitec.
- [28] JIANG, J. (1996). REML Estimation: Asymptotic Behavior and Related Topics. *The Annals of Statistics*, **24**, 255-286.
- [29] JIANG, J. (1998). Asymptotic Properties of the Empirical BLUP and BLUE in Mixed Linear Models. *Statistica Sinica*, **8**, 861-885.
- [30] JIANG, J. (1999). On Unbiasedness of the Empirical BLUE and BLUP. *Statistics & Probability Letters*, **41**, 19-24.
- [31] KOPOCINSKI, B. (2009). Multivariate Negative Binomial Distributions Generated by Multivariate Exponential Distributions. *Applicationes Mathematicae*, **25**, 463-472.
- [32] LAIRD, N. M. & WARE, J. H. (1982). Random-Effects Models for Longitudinal Data. *Biometrics*, **38**, 963-974.
- [33] LAWLESS, J. F. (1987). Negative Binomial and Mixed Poisson Regression. *The Canadian Journal of Statistics*, **15**, 209-225.
- [34] LEE, E. H. (2009). *Copula Analysis of Correlated Counts*. University of California, Irvine.
- [35] LIANG, K. & ZEGER, S. L. (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, **73**, 13-22.
- [36] LINDSEY, J. (1995). *Modelling Frequency and Count Data*. Oxford

- [37] LITTELL, R.C. et al. (2006). *SAS Mixed Models*. 2nd Edition. Cary: SAS Institute.
- [38] LITTLE, T. D., SCHNABEL, K. U. & BAUMERT, J. (2000). *Modeling Longitudinal and Multilevel Data*. LEA.
- [39] MANCO, O. C., U. (2013). *Modelos de Regressão Beta com Efeitos Aleatórios Normais e Não-normais para Dados Longitudinais*. 2013. 123f Tese (Doutorado em Ciências). Instituto de Matemática e Estatística - Universidade de São Paulo, SP, 2013.
- [40] MASAROTTO, G. & VARIN, C. (2012). Gaussian Copula Marginal Regression. *Electronic Journal of Statistics*, **6**, 1517-1549.
- [41] MARITZ, J.S. & LWIN, T. (1989). *Empirical Bayes Methods*. 2nd Edition. London: Chapman & Hall.
- [42] MCCULLOCH & SEARLE, (2001). *Generalizer, Linear, and Mixed Models*. Wiley.
- [43] MIAOU, S. P. (1994). The Relationship Between Truck Accidents and Geometric Design of Road Sections: Poisson versus Negative Binomial Regressions. *Accid. Anal. and Prev.*, **26**, 471-482.
- [44] NELDER, J. A. & WEDDERBURN, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society*, **135**, 370-384.
- [45] NOBRE, J. S. & SINGER, J. M. (2007). Residual Analysis for Linear Mixed Models. *Biometrical Journal*, **49**, 863-875.
- [46] NOBRE, J. S. & SINGER, J. M. (2011). Leverage analysis for linear mixed models. *Journal of Applied Statistics*.
- [47] PAN, W. (2001). Akaike's Information Criterion in Generalized Estimating Equations. *Biometrics*, **57**, 120-125.
- [48] PARK (1993). A Comparison of the Generalized Estimating Equation Approach with the Maximum Likelihood Approach for Repeated Measurements. *Statistics in Medicine*, **12**, 1723-1732.
- [49] PATTERSON, H. D., & THOMPSON, R. (1971). Recovery of Inter-block Information when Block Sizes are Unequal. *Biometrika*, **58**, 545-554.

- [50] PAULA, G. A. (2013). *Modelos de Regressão com Apoio Computacional*. IME/USP.
- [51] PINHEIRO, J. C. & BATES, D. M. (2000). *Mixed-Effects in S and S-PLUS*. New York: Springer.
- [52] POISSON, S. D. (1837). Recherches sur la Probabilité des Jugements en Matière Criminelle et en Matière Civile. *Bachelier*.
- [53] PREISSER, J. S. & QAQISH, B. F. (1996). Deletion Diagnostics for Generalized Estimating Equations. *Biometrika*, **83**, 551-562.
- [54] RABE-HESKETH & SKRONDAL (2005). *Multilevel and Longitudinal Modeling Using Stata*. Stata Press.
- [55] ROBINSON, G. K. (1991). That BLUP is a Good Thing: The Estimation of Random Effects (with discussion). *Statistical Science*, **6**, 15-51.
- [56] ROCHA, F. M. M. (2004). *Seleção de Estruturas de Covariância para Dados com Medidas Repetidas*. Dissertação de Mestrado. IME/USP, São Paulo.
- [57] SCHWARZ, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, **6**, 461-464.
- [58] SEARLE, S. R. (1971). *Linear Models*. New York: Wiley.
- [59] SEARLE, S. R., CASSELA, G., MCCULLOGH, C.E. (1992). *Variance Components*. New York: John Wiley & Sons.
- [60] SINGER, J. M. & ANDRADE, D. F. (1986). *Análise de Dados Longitudinais*. In: VII Simpósio Nacional de Probabilidade e Estatística. Anais, 19-26. Campinas/ SP.
- [61] SINGER, J. D & WILLETT, J. B. (2003) *Applied Longitudinal Data Analysis. Modeling Change and Event Occurrence*. Oxford.
- [62] SINGER, J. M., NOBRE, J. S & ROCHA, F. M. M. (2012). *Análise de dados longitudinais*.
- [63] SKRONDAL A. & RABE-HESKETH, S. (2004). *Generalized Latent Variable Modeling. Multilevel, Longitudinal and Structural Equation Models*. Chapman & Hall/CRC.

- [64] SONG, P. X.-K. (2000). Multivariate Dispersion Models Generated from Gaussian Copula. *Scandinavian Journal of Statistics*, **27**, 305-320.
- [65] STROUD, A. H. & SECHREST, D. (1966). *Gaussian Quadrature Formulas*. Prentice Hall, Englewood Cliffs, NJ.
- [66] TAN, M., QU, Y. & KUTNER, M. H. (1997). Model Diagnostics for Marginal Regression Analysis of Correlated Binary Data. *Commun. Statist. - Simula*, **26**, 539-558.
- [67] THALL, P. F. & VAIL, S. C. (1990). Some Covariance Models for Longitudinal Count Data with Over-dispersion. *Biometrics*, **46**, 657-671.
- [68] TOUNTENBURG, H. (1982). *Prior Information in Linear Models*. Chichester. John Wiley & Sons.
- [69] VENEZUELA, M. K. (2003). *Modelos Lineares Generalizados para Análise de Dados com Medidas Repetidas*. Dissertação de Mestrado. IME/USP, São Paulo.
- [70] VERBEKE, G. & LESAFFRE, E. (1996a). A Linear Mixed-effects Model with Heterogeneity in the Random-effects Population. *Journal of the American Statistical Association*, **91**, 217-221.
- [71] VERBEKE, G. & MOLENBERGHS, G. (1997). *Linear Mixed Models In Practice: A SAS Oriented Approach*. Lecture Notes in Statistics 126. New York: Springer-Verlag.
- [72] VERBEKE, G., MOLENBERGHS, G. & RIZOPOULOS, D. (2010). *Longitudinal Research with Latent Variables*. Springer.
- [73] WINKELMANN, R. (2003). *Econometric Analysis of Count Data*. Springer.
- [74] YAHAV, I. & SHMUELI, G. (2008). An Elegant Method for Generating Multivariate Poisson Random Variables.

Apêndice A

Anexo: Modelo Binomial Negativo

Sejam duas variáveis aleatórias (v.a.'s) Y e Z , tal que $Y|Z$ segue uma distribuição Poisson, $Y|Z \sim Po(Z)$, e que a distribuição a priori de Z é Gama com parâmetros $\phi\mu$ e ϕ . Logo,

$$Y|Z \sim Po(Z), \quad \text{sendo} \quad f_{Y|Z}(y|z) = e^{-z} \frac{z^y}{y!} 1_{\{0,1,\dots\}}(y) \quad \text{e}$$
$$Z \sim G(\phi\mu, \phi), \quad \text{sendo} \quad g_Z(z) = \frac{\phi^{\phi\mu}}{\Gamma(\phi\mu)} z^{\phi\mu-1} e^{-\phi z} 1_{(0,\infty)}(z).$$

A distribuição marginal de Y é dada pela integração da distribuição conjunta de Y e Z em relação a Z , ou seja,

$$\begin{aligned} f_Y(y) &= \int_0^\infty f(y|z)g(z)dz \\ &= \int_0^\infty e^{-z} \frac{z^y}{y!} \frac{\phi^{\phi\mu}}{\Gamma(\phi\mu)} z^{\phi\mu-1} e^{-\phi z} dz \\ &= \frac{\phi^{\phi\mu}}{y! \Gamma(\phi\mu)} \int_0^\infty z^{y+\phi\mu-1} e^{-(\phi+1)z} dz \\ &= \frac{\phi^{\phi\mu}}{y! \Gamma(\phi\mu)} \frac{\Gamma(y+\phi\mu)}{(\phi+1)^{y+\phi\mu}} \int_0^\infty \frac{(\phi+1)^{y+\phi\mu}}{\Gamma(y+\phi\mu)} z^{y+\phi\mu-1} e^{-(\phi+1)z} dz. \end{aligned} \tag{A.1}$$

Note que a solução da integral (A.1) é igual a 1, pois está sendo integrada uma variável aleatória cuja distribuição é $G(y+\phi\mu, \phi+1)$. Após manipulações algébrica, chega-se à seguinte expressão

$$f_Y(y) = \frac{(y+\phi\mu-1)!}{y! (\phi\mu-1)!} \left(\frac{\phi}{\phi+1} \right)^{\phi\mu} \left(\frac{1}{\phi+1} \right)^y 1_{\{0,1,\dots\}}(y). \tag{A.2}$$

Portanto, $Y \sim BN(\phi\mu, \phi/(\phi+1))$, em que $E(Y) = \mu$ e $\text{Var}(Y) = \mu(\phi+1)/\phi$.

Outra possibilidade é considerar que $Y|Z \sim Po(Z)$ e Z tem distribuição a priori Gama com parâmetros ν e μ/ν , isto é

$$Y|Z \sim Po(Z), \quad \text{sendo} \quad f_{Y|Z}(y|z) = e^{-z} \frac{z^y}{y!} 1_{\{0,1,\dots\}}(y) \quad \text{e}$$

$$Z \sim G\left(\nu, \frac{\nu}{\mu}\right), \quad \text{sendo} \quad g_Z(z) = \frac{\left(\frac{\mu}{\nu}\right)^\nu}{\Gamma(\nu)} z^{\nu-1} e^{-z\frac{\nu}{\mu}} 1_{(0,\infty)}(z)$$

Assim, a distribuição marginal de Y é dada por

$$\begin{aligned} f_Y(y) &= \int_0^\infty e^{-z} \frac{z^y}{y!} \frac{1}{\Gamma(\nu) \left(\frac{\mu}{\nu}\right)^\nu} z^{\nu-1} e^{-z\frac{\nu}{\mu}} dz \\ &= \frac{\left(\frac{\nu}{\mu}\right)^\nu}{y! \Gamma(\nu)} \int_0^\infty z^{y+\nu-1} e^{-z\frac{\nu}{\mu}} dz \\ &= \frac{\left(\frac{\nu}{\mu}\right)^\nu}{y! \Gamma(\nu)} \frac{\Gamma(y+\nu)}{\left(\frac{\nu}{\mu} + 1\right)^{y+\nu}} \int_0^\infty \frac{\left(\frac{\nu}{\mu} + 1\right)^{y+\nu}}{\Gamma(y+\nu)} z^{y+\nu-1} e^{-z\left(\frac{\nu}{\mu} + 1\right)} dz. \end{aligned} \tag{A.3}$$

Novamente a solução da integral em (A.3) é igual a 1, uma vez que está sendo integrada uma variável aleatória cuja distribuição é $G(y+\nu, (1+\nu/\mu))$. Através de manipulações algébricas tem-se que a distribuição marginal de Y é BN, com uma função de variância quadrática, isto é

$$f_Y(y) = \frac{(y+\nu-1)!}{y!(\nu-1)!} \left(\frac{\frac{\nu}{\mu}}{\frac{\nu}{\mu} + 1}\right)^\nu \left(\frac{1}{\frac{\nu}{\mu} + 1}\right)^y 1_{\{0,1,\dots\}}(y), \tag{A.4}$$

de tal forma que $E(Y) = \mu$ e $\text{Var}(Y) = \mu + \frac{\mu^2}{\nu}$.

Apêndice B

Estudo de Simulação

Neste apêndice são apresentados os resultados referentes aos estudos de simulação realizados. Na Seção B.1 encontram-se os resultados referentes ao modelo Poisson multinível considerando 150 e 250 indivíduos e a variância do intercepto aleatório iguais a 0,1 e 1,0, para os cenários 1 e 2. Além disso, são sumarizados nas tabelas os resultados para o modelo Poisson multinível considerando 50, 150 e 250 indivíduos e a variância do intercepto aleatório iguais a 0,5 e 0,7, para os cenários 1 e 2.

O Apêndice B.2 apresenta os gráficos do valor absoluto do viés relativo ($|VR|$) e da diferença relativa entre o desvio-padrão e o erro padrão (*dif.rel*) das estimativas para os modelos ajustados considerando a distribuição Poisson, segundo a abordagem multinível, para diferentes valores da variância do efeito do intercepto (τ_0) e considerando os cenários 1 e 2.

No Apêndice B.3 encontram-se disponíveis os resultados do critério de seleção AIC para a escolha do melhor modelo Poisson considerando 150 e 250 indivíduos, para os cenários 1 e 2.

Em relação à abordagem marginal, as tabelas e gráficos para os cenários 2 e 3, considerando a distribuição Poisson, encontram-se dispostos nos Apêndices B.4 e B.5. Os resultados que se referem ao critério de seleção QIC estão disponíveis no Apêndice B.6.

B.1 Modelo Poisson Multinível: Tabelas

Tabela B.1: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\tau_0 = 0, 1$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,277346	-0,022654	0,030712	0,275492	-0,024508	0,032123	0,274159	-0,025841	0,033171
	5	0,295558	-0,004442	0,011541	0,290272	-0,009728	0,014456	0,293781	-0,006219	0,012659
	8	0,303768	0,003768	0,005333	0,298558	-0,001442	0,004052	0,298698	-0,001302	0,004297
	10	0,304916	0,004916	0,005183	0,299750	-0,000250	0,001667	0,299718	-0,000282	0,001678
$\widehat{\beta}_2$	3	2,317125	0,017125	0,023074	2,319590	0,019590	0,025189	2,319262	0,019262	0,024831
	5	2,306997	0,006997	0,008852	2,314319	0,014319	0,016189	2,315901	0,015901	0,018258
	8	2,305018	0,005018	0,005133	2,307893	0,007893	0,008197	2,307699	0,007699	0,008147
	10	2,304905	0,004905	0,004919	2,307308	0,007308	0,007345	2,307256	0,007256	0,007325
$\widehat{\beta}_3$	3	1,107065	0,007065	0,009403	1,107065	0,007065	0,009403	1,107555	0,007555	0,009784
	5	1,101789	0,001789	0,002663	1,101789	0,001789	0,002663	1,101075	0,001075	0,002267
	8	1,100178	0,000178	0,000485	1,100178	0,000178	0,000485	1,100160	0,000160	0,000520
	10	1,100028	0,000028	0,000162	1,100028	0,000028	0,000162	1,100032	0,000032	0,000163
$\tau_0 = 1, 0$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,713912	0,413912	0,414256	0,283258	-0,016742	0,025643	0,267564	-0,032436	0,040177
	5	0,726130	0,426130	0,426224	0,292757	-0,007243	0,012094	0,286471	-0,013529	0,018124
	8	0,731825	0,431825	0,431836	0,299272	-0,000728	0,003323	0,298387	-0,001613	0,004419
	10	0,732622	0,432622	0,432624	0,300157	0,000157	0,001411	0,299958	-0,000042	0,001794
$\widehat{\beta}_2$	3	2,156642	-0,143358	0,143885	2,393656	0,093656	0,098997	2,388455	0,088455	0,094037
	5	2,150296	-0,149704	0,149766	2,378178	0,078178	0,079176	2,373434	0,073434	0,074708
	8	2,149061	-0,150939	0,150941	2,372440	0,072440	0,072487	2,371253	0,071253	0,071658
	10	2,148991	-0,151009	0,151009	2,372138	0,072138	0,072143	2,371565	0,071565	0,071973
$\widehat{\beta}_3$	3	1,104930	0,004930	0,007093	1,104930	0,004930	0,007093	1,110914	0,010914	0,012858
	5	1,101258	0,001258	0,002081	1,101258	0,001258	0,002081	1,102647	0,002647	0,003423
	8	1,100124	0,000124	0,000397	1,100124	0,000124	0,000397	1,100242	0,000242	0,000544
	10	1,100018	0,000018	0,000135	1,100018	0,000018	0,000135	1,100039	0,000039	0,000177

Tabela B.2: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\tau_0 = 0, 1$									
	n	Modelo 1				Modelo 2			
		$\hat{\beta}$	VM	REQM		$\hat{\beta}$	VM	REQM	Modelo 3
$\hat{\beta}_1$	3	0,290458	-0,009542	0,019104		0,286589	-0,013411	0,021338	0,284507 -0,015493 0,022722
	5	0,309478	0,009478	0,012677		0,304202	0,004202	0,009459	0,309818 0,009818 0,013284
	8	0,318199	0,018199	0,018426		0,312884	0,012884	0,013204	0,312687 0,012687 0,013016
	10	0,319383	0,019383	0,019423		0,314077	0,014077	0,014134	0,314042 0,014042 0,014099
$\hat{\beta}_2$	3	2,319761	0,019761	0,022830		2,324521	0,024521	0,027488	2,323828 0,023828 0,026777
	5	2,309810	0,009810	0,010645		2,310326	0,010326	0,012214	2,313546 0,013546 0,015779
	8	2,307969	0,007969	0,008011		2,306887	0,006887	0,007108	2,306713 0,006713 0,006935
	10	2,307854	0,007854	0,007859		2,306619	0,006619	0,006647	2,306566 0,006566 0,006613
$\hat{\beta}_3$	3	1,107604	0,007604	0,009060		1,107604	0,007604	0,009060	1,108392 0,008392 0,009737
	5	1,101929	0,001929	0,002481		1,101929	0,001929	0,002481	1,100733 0,000733 0,001803
	8	1,100188	0,000188	0,000394		1,100188	0,000188	0,000394	1,100214 0,000214 0,000409
	10	1,100033	0,000033	0,000126		1,100033	0,000033	0,000126	1,100036 0,000036 0,000128
$\tau_0 = 1, 0$									
	n	Modelo 1				Modelo 2			
		$\hat{\beta}$	VM	REQM		$\hat{\beta}$	VM	REQM	Modelo 3
$\hat{\beta}_1$	3	0,969618	0,669618	0,669723		0,427403	0,127403	0,128175	0,411172 0,111172 0,112624
	5	0,979463	0,679463	0,679490		0,437508	0,137508	0,137671	0,431015 0,131015 0,131332
	8	0,984071	0,684071	0,684074		0,443075	0,143075	0,143091	0,442105 0,142105 0,142138
	10	0,984703	0,684703	0,684704		0,443793	0,143793	0,143796	0,443553 0,143553 0,143559
$\hat{\beta}_2$	3	2,255786	-0,044214	0,044927		2,390631	0,090631	0,094103	2,385803 0,085803 0,089483
	5	2,250586	-0,049414	0,049497		2,371551	0,071551	0,072306	2,367548 0,067548 0,068459
	8	2,249640	-0,050360	0,050363		2,366277	0,066277	0,066311	2,364982 0,064982 0,065274
	10	2,249577	-0,050423	0,050423		2,365952	0,065952	0,065955	2,365584 0,065584 0,065808
$\hat{\beta}_3$	3	1,103972	0,003972	0,005315		1,103972	0,003972	0,005315	1,110162 0,010162 0,011372
	5	1,101022	0,001022	0,001524		1,101022	0,001022	0,001524	1,102455 0,002455 0,002956
	8	1,100097	0,000097	0,000270		1,100098	0,000098	0,000270	1,100226 0,000226 0,000428
	10	1,100015	0,000015	0,000090		1,100015	0,000015	0,000090	1,100040 0,000040 0,000134

Tabela B.3: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\sigma^2 = 0, 5$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,417633	0,117633	0,121163	0,272584	-0,027416	0,042041	0,281625	-0,018375	0,060807
	5	0,431550	0,131550	0,132442	0,287096	-0,012904	0,020785	0,284962	-0,015038	0,022976
	8	0,437109	0,137109	0,137206	0,293483	-0,006517	0,008350	0,293200	-0,006800	0,008797
	10	0,438118	0,138118	0,138137	0,294584	-0,005416	0,005900	0,294560	-0,005440	0,006010
$\widehat{\beta}_2$	3	2,479215	0,179215	0,180467	2,616887	0,316887	0,319725	2,607737	0,307737	0,311465
	5	2,470855	0,170855	0,171034	2,587500	0,287500	0,288049	2,585819	0,285819	0,286425
	8	2,469395	0,169395	0,169402	2,580827	0,280827	0,280851	2,580560	0,280560	0,280710
	10	2,469301	0,169301	0,169302	2,580358	0,280358	0,280361	2,579958	0,279958	0,280116
$\widehat{\beta}_3$	3	1,104756	0,004756	0,009720	1,104756	0,004756	0,009720	1,101929	0,001929	0,018418
	5	1,101193	0,001193	0,003024	1,101193	0,001193	0,003024	1,101685	0,001685	0,003477
	8	1,100131	0,000131	0,000633	1,100131	0,000131	0,000633	1,100170	0,000170	0,000692
	10	1,100007	0,000007	0,000224	1,100007	0,000007	0,000224	1,100012	0,000012	0,000243
$\sigma^2 = 0, 7$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,546870	0,246870	0,248333	0,272444	-0,027556	0,042039	0,265420	-0,034580	0,049130
	5	0,559309	0,259309	0,259702	0,285454	-0,014546	0,021399	0,282830	-0,017170	0,024505
	8	0,564218	0,264218	0,264262	0,291447	-0,008553	0,009869	0,291067	-0,008933	0,010521
	10	0,565114	0,265114	0,265122	0,292462	-0,007538	0,007852	0,292426	-0,007574	0,007999
$\widehat{\beta}_2$	3	2,498610	0,198610	0,199565	2,726310	0,426310	0,428953	2,723546	0,423546	0,426203
	5	2,491060	0,191060	0,191195	2,699550	0,399550	0,399999	2,697309	0,397309	0,397852
	8	2,489759	0,189759	0,189764	2,692960	0,392960	0,392979	2,692565	0,392565	0,392758
	10	2,489677	0,189677	0,189677	2,692473	0,392473	0,392475	2,691971	0,391971	0,392182
$\widehat{\beta}_3$	3	1,104215	0,004215	0,008900	1,104215	0,004215	0,008900	1,106954	0,006954	0,011506
	5	1,101051	0,001051	0,002796	1,101051	0,001051	0,002796	1,101658	0,001658	0,003453
	8	1,100115	0,000115	0,000590	1,100115	0,000115	0,000590	1,100169	0,000169	0,000691
	10	1,100005	0,000005	0,000209	1,100005	0,000005	0,000209	1,100012	0,000012	0,000241

Tabela B.4: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\sigma^2 = 0, 5$													
	n	Modelo 1			Modelo 2			Modelo 3					
		$\widehat{\beta}$	VM	$REQM$	$\widehat{\beta}$	VM	$REQM$	$\widehat{\beta}$	VM	$REQM$			
$\widehat{\beta}_1$	3	0,388780	0,088780	0,090938	0,279495	-0,020505	0,029014	0,389303	0,089303	0,108891			
	5	0,405221	0,105221	0,105712	0,291438	-0,008562	0,013490	0,287933	-0,012067	0,016440			
	8	0,412656	0,112656	0,112713	0,298866	-0,001134	0,003800	0,298444	-0,001556	0,004206			
	10	0,413696	0,113696	0,113707	0,299924	-0,000076	0,001580	0,299836	-0,000164	0,001710			
$\widehat{\beta}_2$	3	2,290645	-0,009355	0,017316	2,356779	0,056779	0,062187	2,336991	0,036991	0,047675			
	5	2,281598	-0,018402	0,019101	2,341696	0,041696	0,043093	2,339481	0,039481	0,040970			
	8	2,279854	-0,020146	0,020172	2,336330	0,036330	0,036405	2,335843	0,035843	0,036089			
	10	2,279752	-0,020248	0,020251	2,336090	0,036090	0,036098	2,335850	0,035850	0,036061			
$\widehat{\beta}_3$	3	1,106414	0,006414	0,008738	1,106414	0,006414	0,008738	1,067157	-0,032843	0,039439			
	5	1,101625	0,001625	0,002487	1,101625	0,001625	0,002487	1,102400	0,002400	0,003151			
	8	1,100161	0,000161	0,000459	1,100161	0,000161	0,000459	1,100218	0,000218	0,000514			
	10	1,100025	0,000025	0,000154	1,100025	0,000025	0,000154	1,100034	0,000034	0,000168			
$\sigma^2 = 0, 7$													
	n	Modelo 1			Modelo 2			Modelo 3					
		$\widehat{\beta}$	VM	$REQM$	$\widehat{\beta}$	VM	$REQM$	$\widehat{\beta}$	VM	$REQM$			
$\widehat{\beta}_1$	3	0,495419	0,195419	0,196310	0,281169	-0,018831	0,027577	0,269118	-0,030882	0,038317			
	5	0,510408	0,210408	0,210632	0,292022	-0,007978	0,012898	0,287394	-0,012606	0,017061			
	8	0,517205	0,217205	0,217232	0,299023	-0,000977	0,003624	0,298418	-0,001582	0,004295			
	10	0,518162	0,218162	0,218167	0,300014	0,000014	0,001517	0,299883	-0,000117	0,001745			
$\widehat{\beta}_2$	3	2,249900	-0,050100	0,051956	2,370347	0,070347	0,075803	2,366734	0,066734	0,072344			
	5	2,241824	-0,058176	0,058376	2,355936	0,055936	0,057097	2,352789	0,052789	0,054115			
	8	2,240266	-0,059734	0,059742	2,350751	0,050751	0,050809	2,349997	0,049997	0,050302			
	10	2,240176	-0,059824	0,059825	2,350509	0,050509	0,050516	2,350134	0,050134	0,050428			
$\widehat{\beta}_3$	3	1,105905	0,005905	0,008163	1,105904	0,005904	0,008163	1,110498	0,010498	0,012418			
	5	1,101491	0,001491	0,002340	1,101491	0,001491	0,002340	1,102514	0,002514	0,003272			
	8	1,100148	0,000148	0,000438	1,100148	0,000148	0,000438	1,100229	0,000229	0,000527			
	10	1,100023	0,000023	0,000148	1,100023	0,000023	0,000148	1,100037	0,000037	0,000172			

Tabela B.5: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com intercepto aleatório. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\sigma^2 = 0, 5$										
		Modelo 1			Modelo 2			Modelo 3		
	n	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,484879	0,184879	0,185490	0,349933	0,049933	0,052335	0,339676	0,039676	0,043215
	5	0,500590	0,200590	0,200737	0,363360	0,063360	0,063847	0,359598	0,059598	0,060240
	8	0,507782	0,207782	0,207798	0,370719	0,070719	0,070768	0,370226	0,070226	0,070289
	10	0,508761	0,208761	0,208764	0,371724	0,071724	0,071733	0,371604	0,071604	0,071616
$\widehat{\beta}_2$	3	2,324528	0,024528	0,026617	2,352505	0,052505	0,056379	2,349895	0,049895	0,053882
	5	2,316257	0,016257	0,016677	2,337794	0,037794	0,038882	2,335818	0,035818	0,036947
	8	2,314727	0,014727	0,014746	2,333249	0,033249	0,033301	2,332726	0,032726	0,032876
	10	2,314629	0,014629	0,014631	2,332980	0,032980	0,032986	2,332796	0,032796	0,032910
$\widehat{\beta}_3$	3	1,106266	0,006266	0,007695	1,106266	0,006266	0,007695	1,110176	0,010176	0,011345
	5	1,101588	0,001588	0,002133	1,101588	0,001588	0,002133	1,102419	0,002419	0,002904
	8	1,100154	0,000154	0,000350	1,100154	0,000154	0,000350	1,100219	0,000219	0,000418
	10	1,100026	0,000026	0,000114	1,100026	0,000026	0,000114	1,100038	0,000038	0,000131
$\sigma^2 = 0, 7$										
		Modelo 1			Modelo 2			Modelo 3		
	n	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,647705	0,347705	0,347980	0,381125	0,081125	0,082508	0,368080	0,068080	0,070320
	5	0,661082	0,361082	0,361151	0,393060	0,093060	0,093355	0,388103	0,088103	0,088555
	8	0,667270	0,367270	0,367278	0,399664	0,099664	0,099694	0,398961	0,098961	0,099007
	10	0,668114	0,368114	0,368115	0,400549	0,100549	0,100555	0,400380	0,100380	0,100388
$\widehat{\beta}_2$	3	2,308586	0,008586	0,012790	2,367887	0,067887	0,071648	2,364392	0,064392	0,068293
	5	2,301485	0,001485	0,003719	2,351233	0,051233	0,052132	2,348470	0,048470	0,049447
	8	2,300184	0,000184	0,000708	2,346453	0,046453	0,046492	2,345638	0,045638	0,045834
	10	2,300098	0,000098	0,000246	2,346166	0,046166	0,046171	2,345907	0,045907	0,046064
$\widehat{\beta}_3$	3	1,105353	0,005353	0,006746	1,105353	0,005353	0,006746	1,110326	0,010326	0,011513
	5	1,101368	0,001368	0,001897	1,101368	0,001368	0,001897	1,102462	0,002462	0,002953
	8	1,100132	0,000132	0,000320	1,100132	0,000132	0,000320	1,100225	0,000225	0,000425
	10	1,100022	0,000022	0,000105	1,100022	0,000022	0,000105	1,100040	0,000040	0,000133

Tabela B.6: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\tau_0 = 0, 1$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,248791	-0,051209	0,054937	-0,105500	-0,405500	0,406119	0,276925	-0,023075	0,034912
	5	-0,062512	-0,362512	0,362629	-1,081874	-1,381874	1,381917	0,295809	-0,004191	0,018582
	8	-1,181951	-1,481951	1,481952	-3,776189	-4,076189	4,076190	0,297868	-0,002132	0,005585
	10	-1,942875	-2,242875	2,242875	-5,841148	-6,141148	6,141148	0,298503	-0,001497	0,003261
$\widehat{\beta}_2$	3	1,903627	-0,396373	0,396544	2,098370	-0,201630	0,203558	2,301174	0,001174	0,041684
	5	1,516296	-0,783704	0,783708	1,935729	-0,364271	0,364515	2,318017	0,018017	0,030478
	8	0,969271	-1,330729	1,330729	1,689011	-0,610989	0,611022	2,304795	0,004795	0,009958
	10	0,642754	-1,657246	1,657246	1,524431	-0,775569	0,775582	2,309282	0,009282	0,010748
$\widehat{\beta}_3$	3	1,284016	0,184016	0,184129	1,284016	0,184016	0,184129	1,133644	0,033644	0,034667
	5	1,431036	0,331036	0,331041	1,431036	0,331036	0,331041	1,126317	0,026317	0,026610
	8	1,662540	0,562540	0,562540	1,662540	0,562540	0,562540	1,125747	0,025747	0,025757
	10	1,764081	0,664081	0,664081	1,764081	0,664081	0,664081	1,125648	0,025648	0,025650
$\tau_0 = 1, 0$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,026597	0,352259	0,352635	-0,189143	-0,489143	0,489625	0,306497	0,006497	0,029883
	5	0,281044	-0,018956	0,020336	-1,214632	-1,514632	1,514664	0,280088	-0,019912	0,028601
	8	-0,68429	-0,984290	0,984291	-3,674932	-3,974932	3,974932	0,282342	-0,017658	0,020333
	10	-1,244265	-1,544265	1,544265	-5,456812	-5,756812	5,756812	0,289067	-0,010933	0,012633
$\widehat{\beta}_2$	3	1,767180	-0,532820	0,532894	2,189104	-0,110896	0,116826	2,430376	0,130376	0,140618
	5	1,560288	-0,739712	0,739714	2,009627	-0,290373	0,290781	2,456386	0,156386	0,163524
	8	1,349039	-0,950961	0,950961	1,754243	-0,545757	0,545786	2,444701	0,144701	0,147392
	10	1,175096	-1,124904	1,124904	1,588953	-0,711047	0,711056	2,427660	0,127660	0,129041
$\widehat{\beta}_3$	3	1,311013	0,211013	0,211080	1,311013	0,211013	0,211080	1,119032	0,019032	0,021031
	5	1,458447	0,358447	0,358450	1,458447	0,358447	0,358450	1,129275	0,029275	0,029571
	8	1,649488	0,549488	0,549488	1,649488	0,549488	0,549488	1,127947	0,027947	0,027978
	10	1,724848	0,624848	0,624848	1,724848	0,624848	0,624848	1,126741	0,026741	0,026750

Tabela B.7: Valores médios das estimativas, vieses médios e $REQM$ das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$. $\sigma^2 = 0,1$

	n	Modelo 1		Modelo 2		Modelo 3	
		VM	$REQM$	VM	$REQM$	VM	$REQM$
$\widehat{\beta}_1$	3	0,202088	-0,097912	0,099122	-0,078565	-0,378565	0,378949
	5	-0,060475	-0,360475	0,360553	-0,926109	-1,226109	1,226139
	8	-1,071161	-1,371161	1,371161	-3,32007	-3,620070	3,620071
	10	-1,837192	-2,137192	2,137192	-5,527401	-5,527401	5,527401
$\widehat{\beta}_2$	3	2,157394	-0,142606	0,142932	2,200558	-0,099442	0,101845
	5	1,882163	-0,417837	0,417844	2,111170	-0,188830	0,189150
	8	1,573016	-0,726984	0,726984	1,976056	-0,323944	0,323972
	10	1,542704	-0,757296	0,757296	1,888654	-0,411346	0,411352
$\widehat{\beta}_3$	3	1,258380	0,158380	0,158459	1,258380	0,158380	0,158459
	5	1,380484	0,280484	0,280488	1,380484	0,280484	0,280488
	8	1,585555	0,485555	0,485555	1,585555	0,485555	0,485555
	10	1,682821	0,582821	0,582821	1,682821	0,582821	0,582821
$\sigma^2 = 1,0$							
	n	Modelo 1		Modelo 2		Modelo 3	
		VM	$REQM$	VM	$REQM$	VM	$REQM$
$\widehat{\beta}_1$	3	0,880799	0,580799	0,580911	0,165406	-0,134594	0,135342
	5	0,110398	0,364494	0,364540	-0,527545	-0,827545	0,827577
	8	-0,283466	-0,583466	0,583467	-2,706008	-3,006008	3,006009
	10	-1,168745	-1,468745	1,468745	-4,662354	-4,962354	4,962354
$\widehat{\beta}_2$	3	2,132806	-0,167194	0,167348	2,263159	-0,036841	0,045176
	5	1,858489	-0,441511	0,441515	2,170362	-0,129638	0,130193
	8	1,397881	-0,902119	0,902119	2,036062	-0,263938	0,263984
	10	1,264230	-1,035770	1,035770	1,948111	-0,351889	0,351898
$\widehat{\beta}_3$	3	1,215452	0,115452	0,115511	1,215452	0,115452	0,115511
	5	1,322500	0,222500	0,222503	1,322500	0,222500	0,222503
	8	1,522835	0,422835	0,422835	1,522835	0,422835	0,422835
	10	1,638223	0,538223	0,538223	1,638223	0,538223	0,538223

Tabela B.8: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 50 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\sigma^2 = 0, 5$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,551792	0,251792	0,253485	0,054443	-0,245557	0,248169	0,333793	0,033793	0,102434
	5	0,515499	0,215499	0,216004	-0,479456	-0,779456	0,779707	0,262183	-0,037817	0,051307
	8	0,079393	-0,220607	0,220638	-1,890501	-2,190501	2,190509	0,278237	-0,021763	0,026887
	10	-0,305108	-0,605108	0,605109	-3,022794	-3,322794	3,322795	0,286661	-0,013339	0,016917
$\widehat{\beta}_2$	3	1,946769	-0,353231	0,353694	2,472505	0,172505	0,180160	2,608657	0,308657	0,329027
	5	1,474228	-0,825772	0,825785	2,374852	0,074852	0,082489	2,690168	0,390168	0,329027
	8	0,898945	-1,401055	1,401055	2,217331	-0,082669	0,083591	2,633292	0,333292	0,335218
	10	0,645537	-1,654463	1,654463	2,115200	-0,184800	0,184900	2,607273	0,307273	0,308327
$\widehat{\beta}_3$	3	1,154103	0,054103	0,054927	1,154102	0,054102	0,054927	1,039421	-0,060579	-0,060579
	5	1,235342	0,135342	0,135375	1,235342	0,135342	0,135375	1,062238	-0,037762	0,038320
	8	1,353442	0,253442	0,253442	1,353442	0,253442	0,253442	1,058729	-0,041271	0,041315
	10	1,410178	0,310178	0,310178	1,410178	0,351025	0,310178	1,057682	-0,042318	0,042329
$\sigma^2 = 0, 7$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,680157	0,380157	0,381152	0,064315	-0,235685	0,238331	0,280369	-0,019631	0,052438
	5	0,653519	0,353519	0,353791	-0,443638	-0,743638	0,743900	0,264226	-0,035774	0,050775
	8	0,239156	-0,060844	0,060944	-1,803771	-2,103771	2,103780	0,276487	-0,023513	0,029223
	10	-0,141318	-0,441318	0,441320	-2,911553	-3,211553	3,211555	0,284475	-0,015525	0,019536
$\widehat{\beta}_2$	3	1,978214	-0,321786	0,322221	2,596310	0,296310	0,301552	2,801843	0,501843	0,508397
	5	1,509662	-0,790338	0,790349	2,490904	0,190904	0,194561	2,805500	0,505500	0,510637
	8	0,930947	-1,369053	1,369053	2,330218	0,030218	0,032904	2,753758	0,453758	0,455755
	10	0,671025	-1,628975	1,628975	2,227606	-0,072394	0,072678	2,726494	0,426494	0,427663
$\widehat{\beta}_3$	3	1,148409	0,048409	0,049221	1,148409	0,048409	0,049221	1,055523	-0,044477	0,046628
	5	1,226812	0,126812	0,126843	1,226812	0,126812	0,126843	1,061047	-0,038953	0,039529
	8	1,341777	0,241777	0,241778	1,341777	0,241777	0,241778	1,058539	-0,041461	0,041511
	10	1,398446	0,298446	0,298446	1,398446	0,298446	0,298446	1,057612	-0,042388	0,042403

Tabela B.9: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 150 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\sigma^2 = 0, 5$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,352113	0,052113	0,055363	-0,143361	-0,443361	0,443904	0,329808	0,029808	0,067830
	5	0,013258	-0,286742	0,286872	-1,151503	1,445621	1,451540	0,291220	-0,008780	0,019386
	8	-1,028832	-1,328832	1,328833	-3,747000	-4,047000	4,047000	0,288974	-0,011026	0,013543
	10	-1,688699	-1,988699	1,988699	-5,683228	-5,983228	5,983228	0,292027	-0,007973	0,009356
$\widehat{\beta}_2$	3	1,870513	-0,429487	0,429622	2,135902	-0,164098	0,166981	2,338651	0,038651	0,069822
	5	1,554145	-0,745855	0,745859	1,967607	-0,332393	0,332683	2,364801	0,064801	0,074235
	8	1,147576	-1,152424	1,152424	1,717792	-0,582208	0,582238	2,375693	0,075693	0,078288
	10	0,875713	-1,424287	1,424287	1,553005	-0,746995	0,747006	2,370740	0,070740	0,071985
$\widehat{\beta}_3$	3	1,296868	0,196868	0,196962	1,296868	0,196868	0,196962	1,113365	0,013365	0,025580
	5	1,445621	0,345621	0,345625	1,445621	0,345621	0,345625	1,127270	0,027270	0,027502
	8	1,658790	0,558790	0,558790	1,658790	0,558790	0,558790	1,127023	0,027023	0,027043
	10	1,747972	0,647972	0,647972	1,747972	0,647972	0,647972	1,126399	0,026399	0,026404
$\sigma^2 = 0, 7$										
	n	Modelo 1			Modelo 2			Modelo 3		
		$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM	$\widetilde{\beta}$	VM	REQM
$\widehat{\beta}_1$	3	0,450462	0,150462	0,151519	-0,162340	-0,462340	0,462860	0,311763	0,011763	0,029934
	5	0,098186	-0,201814	0,201979	-1,180032	-1,480032	1,480067	0,286721	-0,013279	0,022711
	8	-0,909471	-1,209471	1,209471	-3,720942	-4,020942	4,020943	0,285564	-0,014436	0,016886
	10	-1,524877	-1,824877	1,824877	-5,593093	-5,893093	5,893093	0,290373	-0,009627	0,011116
$\widehat{\beta}_2$	3	1,837299	-0,462701	0,462812	2,156938	-0,143062	0,146910	2,381371	0,081371	0,095190
	5	1,563180	-0,736820	0,736823	1,984127	-0,315873	0,316196	2,401365	0,101365	0,109506
	8	1,233702	-1,066298	1,066298	1,732337	-0,567663	0,567691	2,406449	0,106449	0,109017
	10	0,997604	-1,302396	1,302396	1,567366	-0,732634	0,732643	2,395522	0,095522	0,096862
$\widehat{\beta}_3$	3	1,302907	0,202907	0,202990	1,302907	0,202907	0,202990	1,119038	0,019038	0,020869
	5	1,451491	0,351491	0,351495	1,451491	0,351491	0,351495	1,128135	0,028135	0,028386
	8	1,655430	0,555430	0,555430	1,655430	0,555430	0,555430	1,127503	0,027503	0,027527
	10	1,738769	0,638769	0,638769	1,738769	0,638769	0,638769	1,126590	0,026590	0,026597

Tabela B.10: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson com 250 indivíduos e dados gerados pelo modelo Poisson multinível com dois efeitos aleatórios. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

$\sigma^2 = 0, 5$									
	n	Modelo 1				Modelo 2			
		$\tilde{\beta}$	VM	REQM		$\tilde{\beta}$	VM	REQM	Modelo 3
$\hat{\beta}_1$	3	0,401996	0,101996	0,102976		0,020162	-0,279838	0,280299	0,360279 0,060279 0,063556
	5	0,162299	-0,137701	0,137876		-0,772929	-1,072929	1,072959	0,360052 0,060052 0,061422
	8	-0,835475	-1,135475	1,135476		-3,108381	-3,408381	3,408382	0,366064 0,066064 0,066324
	10	-1,652176	-1,952176	1,952176		-5,046782	-5,346782	-5,346782	0,367710 0,067710 0,067810
$\hat{\beta}_2$	3	2,161358	-0,138642	0,138928		2,227636	-0,072364	0,076002	2,362100 0,062100 0,070952
	5	1,866884	-0,433116	0,433122		2,138319	-0,161681	0,162074	2,356014 0,056014 0,062325
	8	1,488832	-0,811168	0,811168		2,002414	-0,297586	0,297614	2,345015 0,045015 0,047552
	10	1,428448	-0,871552	0,871552		1,914930	-0,385070	0,385076	2,341204 0,041204 0,042481
$\hat{\beta}_3$	3	1,242926	0,142926	0,143000		1,242926	0,142926	0,143000	1,109657 0,009657 0,011608
	5	1,359860	0,259860	0,259864		1,359860	0,259860	0,259864	1,107941 0,007941 0,008378
	8	1,565639	0,465639	0,465639		1,565639	0,465639	0,465639	1,106286 0,006286 0,006335
	10	1,670237	0,570237	0,570237		1,670237	0,570237	0,570237	1,105938 0,005938 0,005951
$\sigma^2 = 0, 7$									
	n	Modelo 1				Modelo 2			
		$\tilde{\beta}$	VM	REQM		$\tilde{\beta}$	VM	REQM	Modelo 3
$\hat{\beta}_1$	3	0,563735	0,263735	0,264063		0,076230	-0,223770	0,224292	0,386354 0,086354 0,088754
	5	0,334629	0,034629	0,035241		-0,680190	-0,980190	0,980221	0,388596 0,088596 0,089600
	8	-0,647810	-0,947810	0,947811		-2,964009	-3,264009	3,264010	0,395353 0,095353 0,095564
	10	-1,49296	-1,79296	1,792960		-4,914286	-5,214286	5,214286	0,397425 0,097425 0,097507
$\hat{\beta}_2$	3	2,157025	-0,142975	0,143216		2,242484	-0,057516	0,062393	2,380295 0,080295 0,087692
	5	1,864063	-0,435937	0,435942		2,151623	-0,148377	0,148823	2,376539 0,076539 0,082238
	8	1,449739	-0,850261	0,850261		2,015755	-0,284245	0,284276	2,359441 0,059441 0,061986
	10	1,364433	-0,935567	0,935567		1,928139	-0,371861	0,371867	2,352744 0,052744 0,053969
$\hat{\beta}_3$	3	1,232705	0,132705	0,132774		1,232705	0,132705	0,132774	1,110709 0,010709 0,012552
	5	1,346080	0,246080	0,246083		1,346080	0,246080	0,246083	1,107942 0,007942 0,008410
	8	1,550680	0,450680	0,450680		1,550680	0,450680	0,450680	1,106188 0,006188 0,006246
	10	1,659617	0,559617	0,559617		1,659617	0,559617	0,559617	1,105820 0,005820 0,005837

B.2 Modelo Poisson Multinível: Gráficos

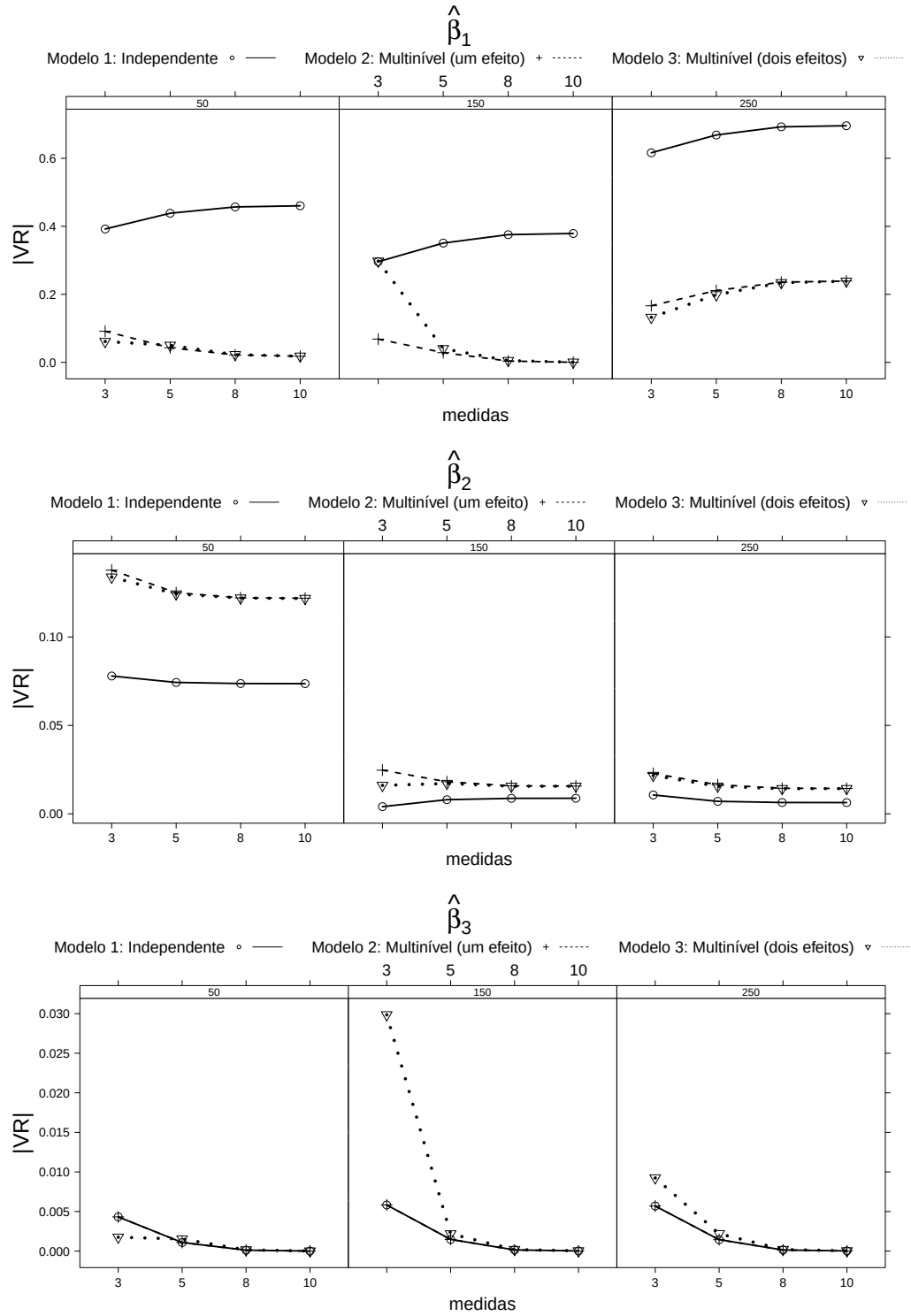


Figura B.1: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,5$.

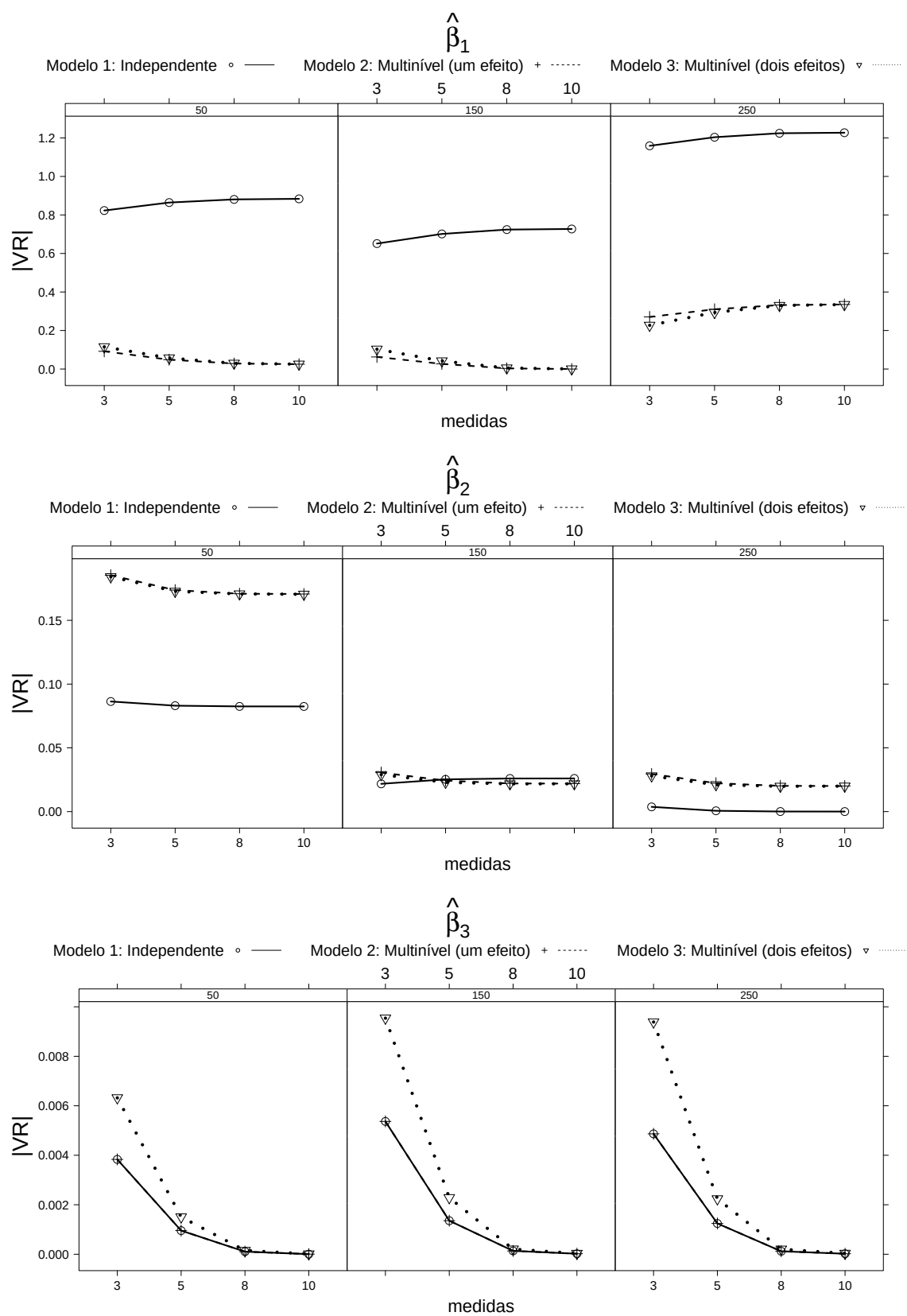


Figura B.2: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,7$.

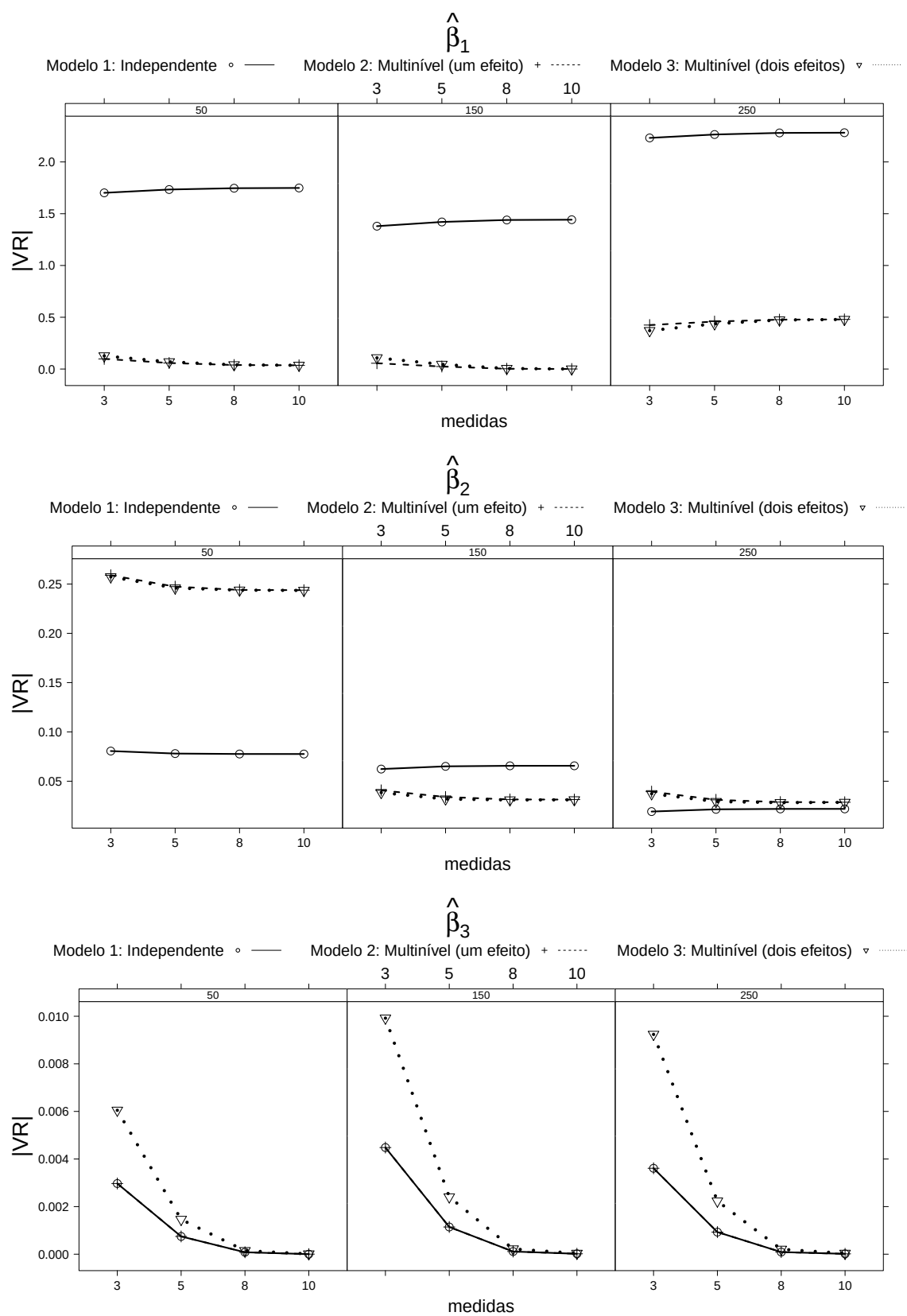


Figura B.3: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 1, 0$.

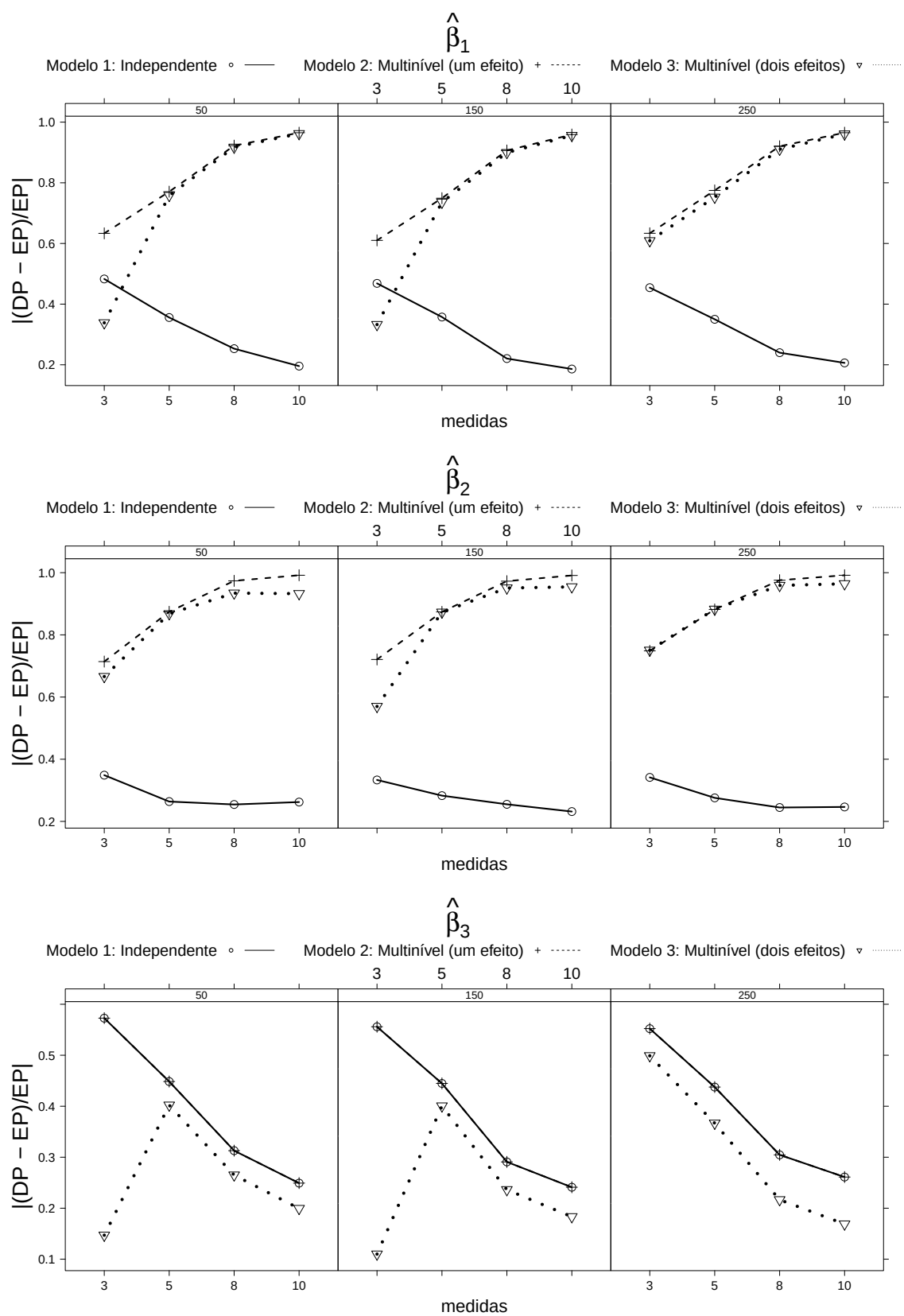


Figura B.4: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,5$.

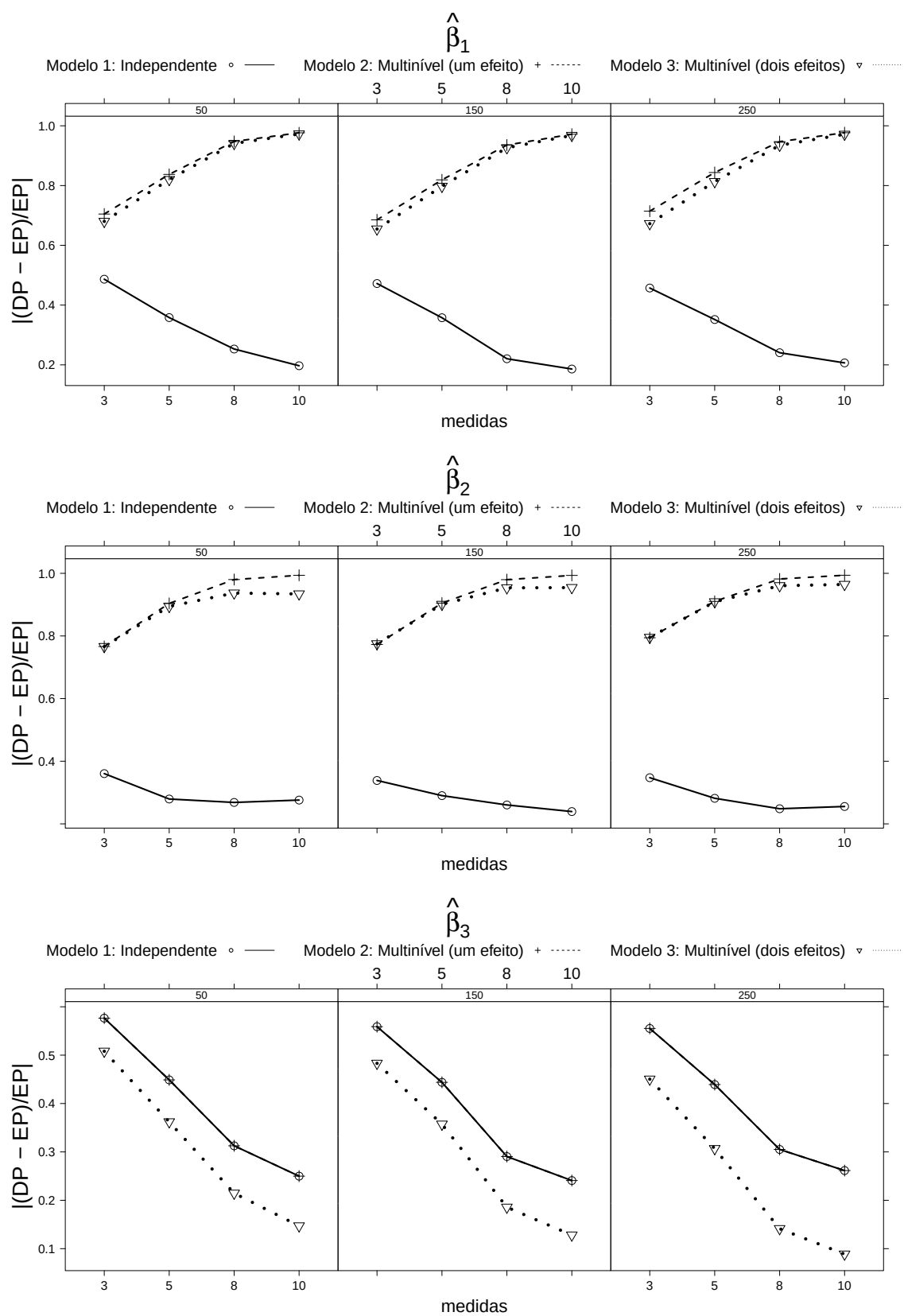


Figura B.5: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 0,7$.

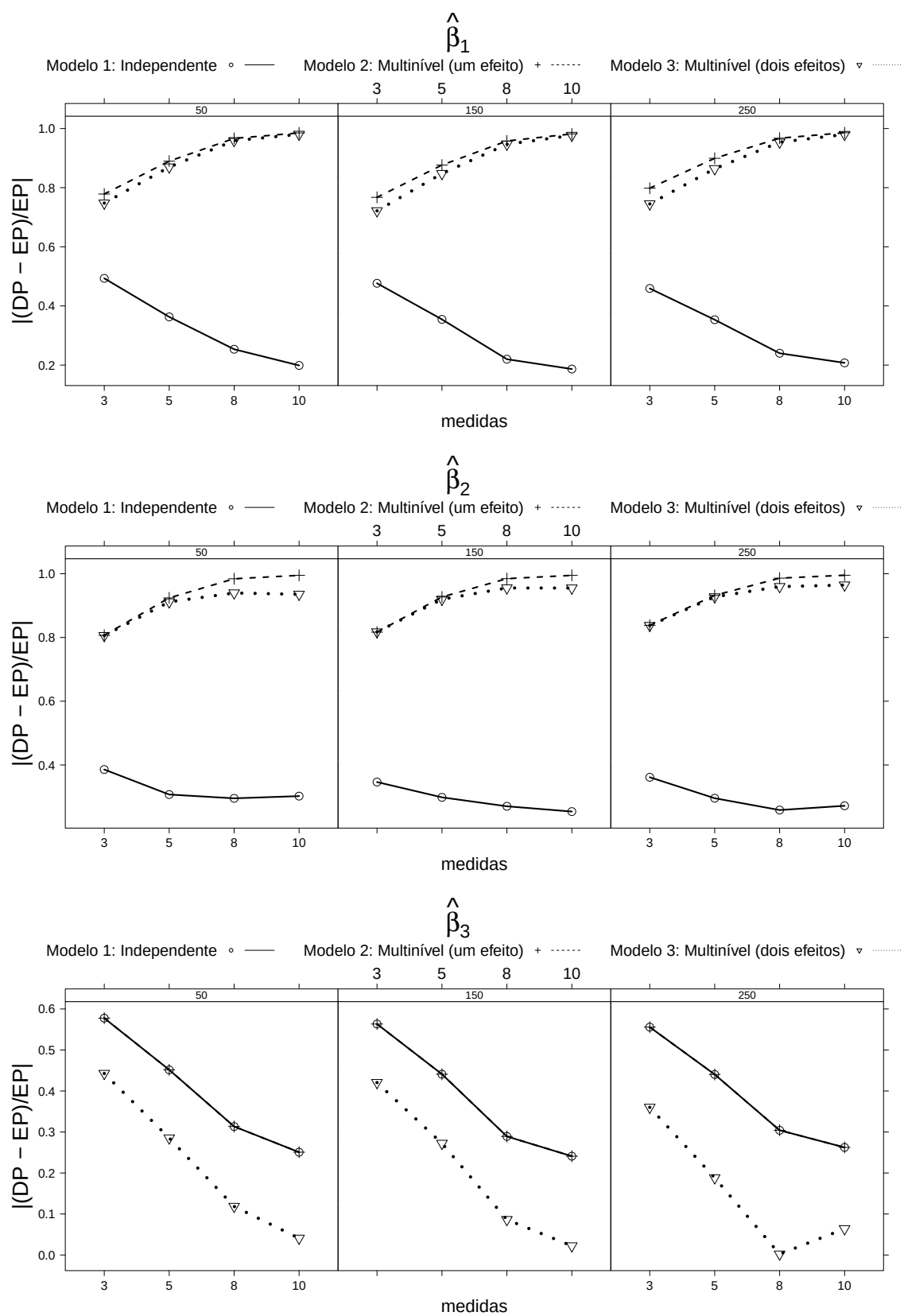


Figura B.6: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com intercepto aleatório para a geração dos dados e $\tau_0 = 1,0$.

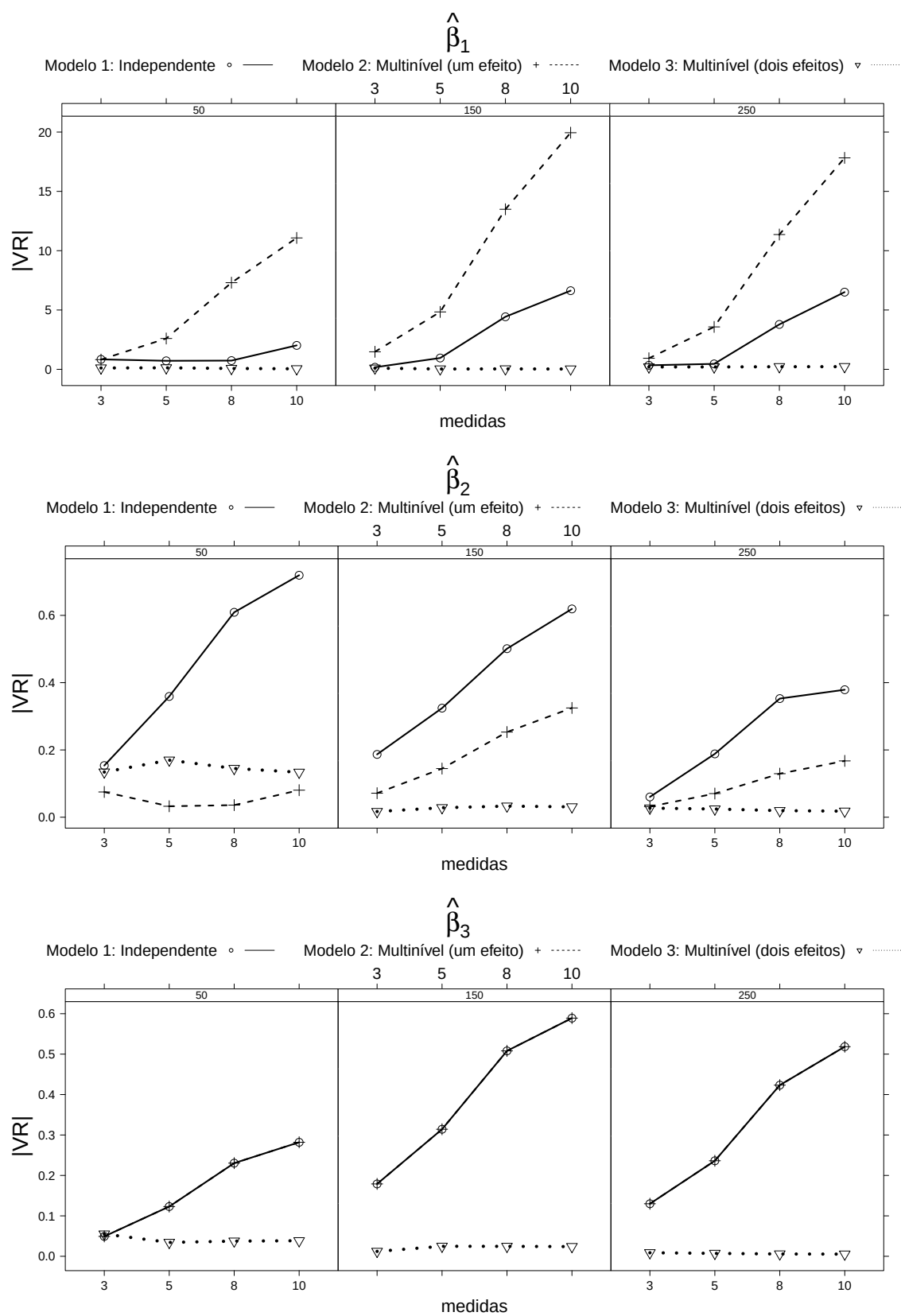


Figura B.7: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,5$.

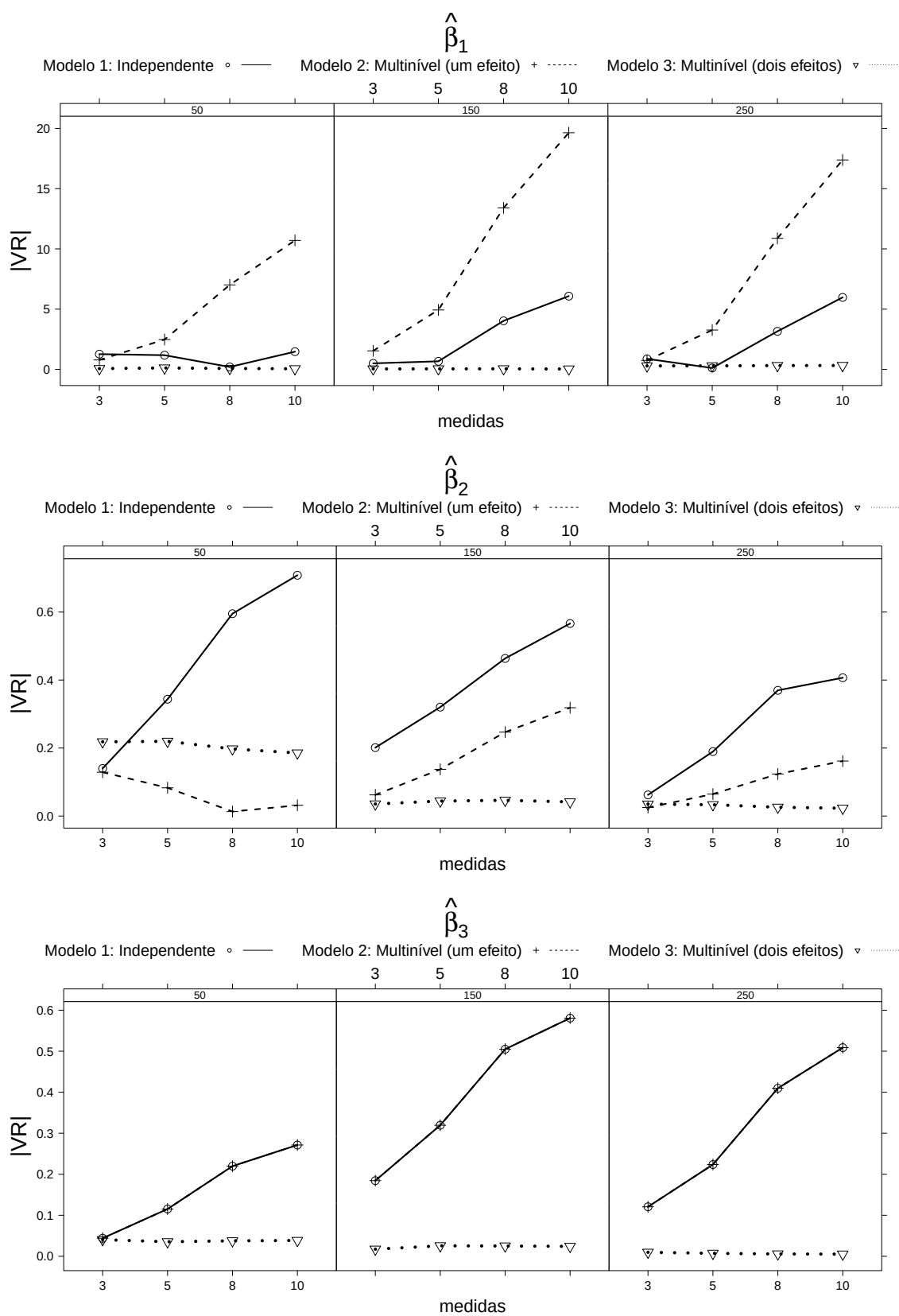


Figura B.8: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,7$.

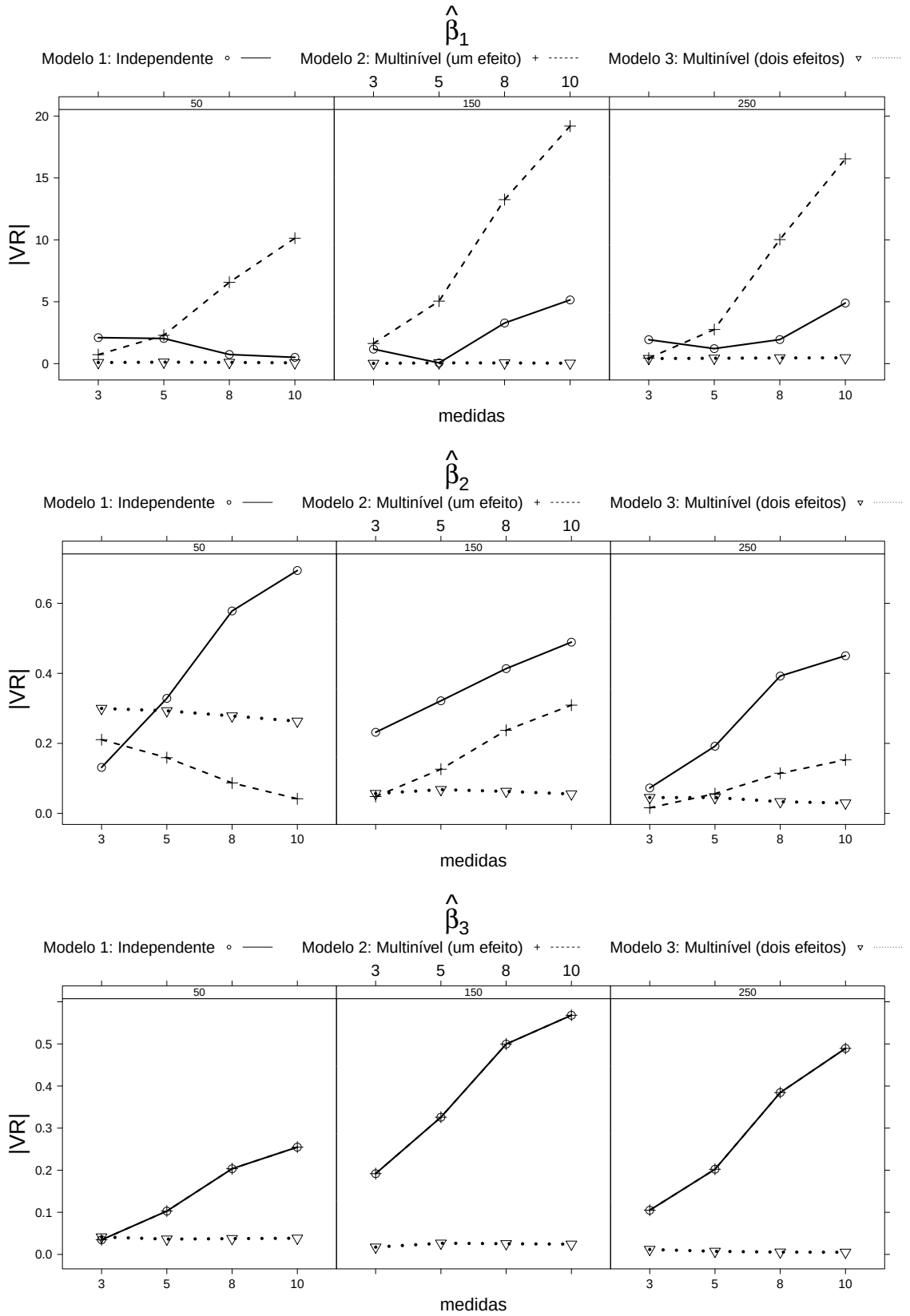


Figura B.9: Valor absoluto do viés relativo para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 1,0$.

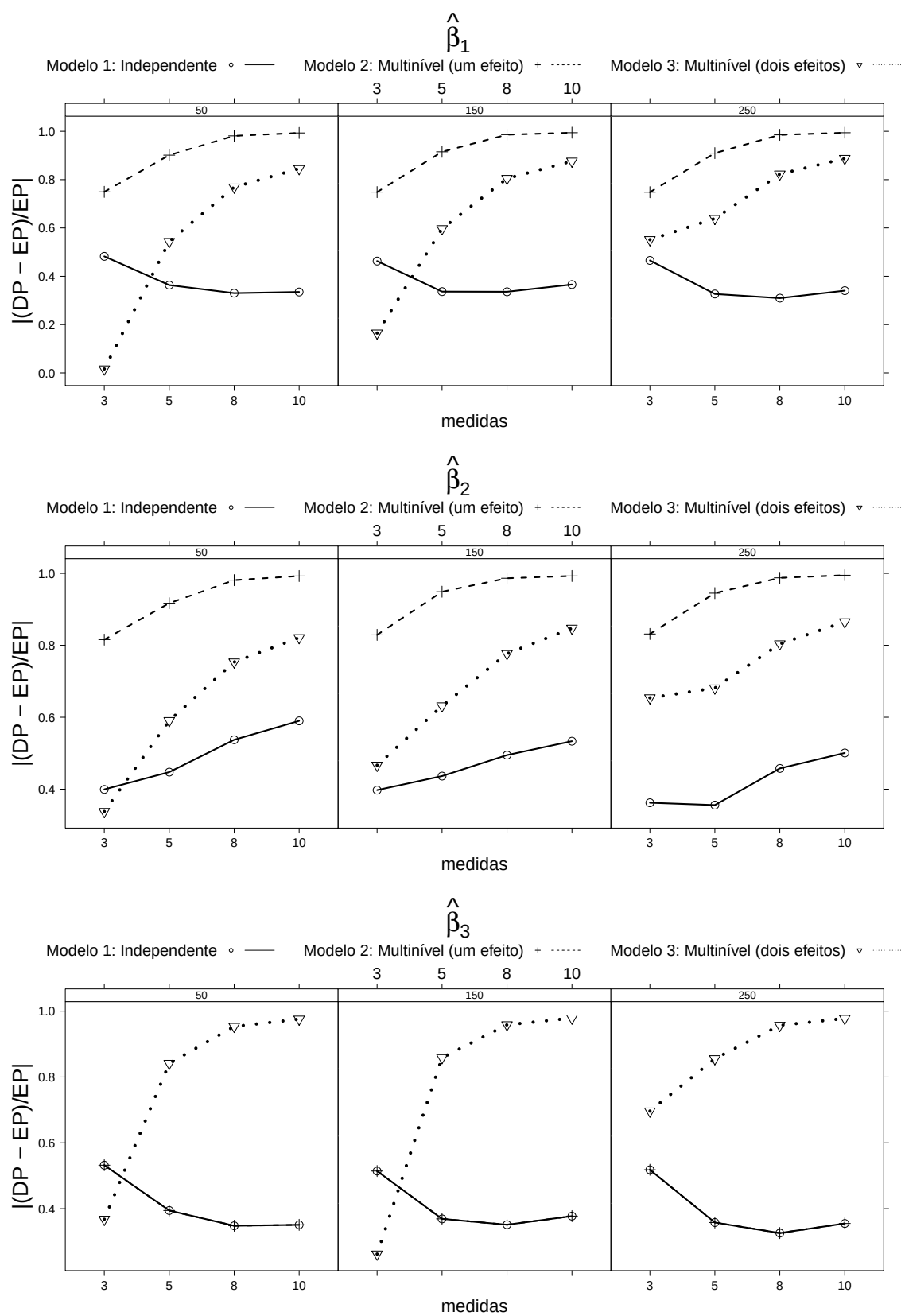


Figura B.10: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,5$.

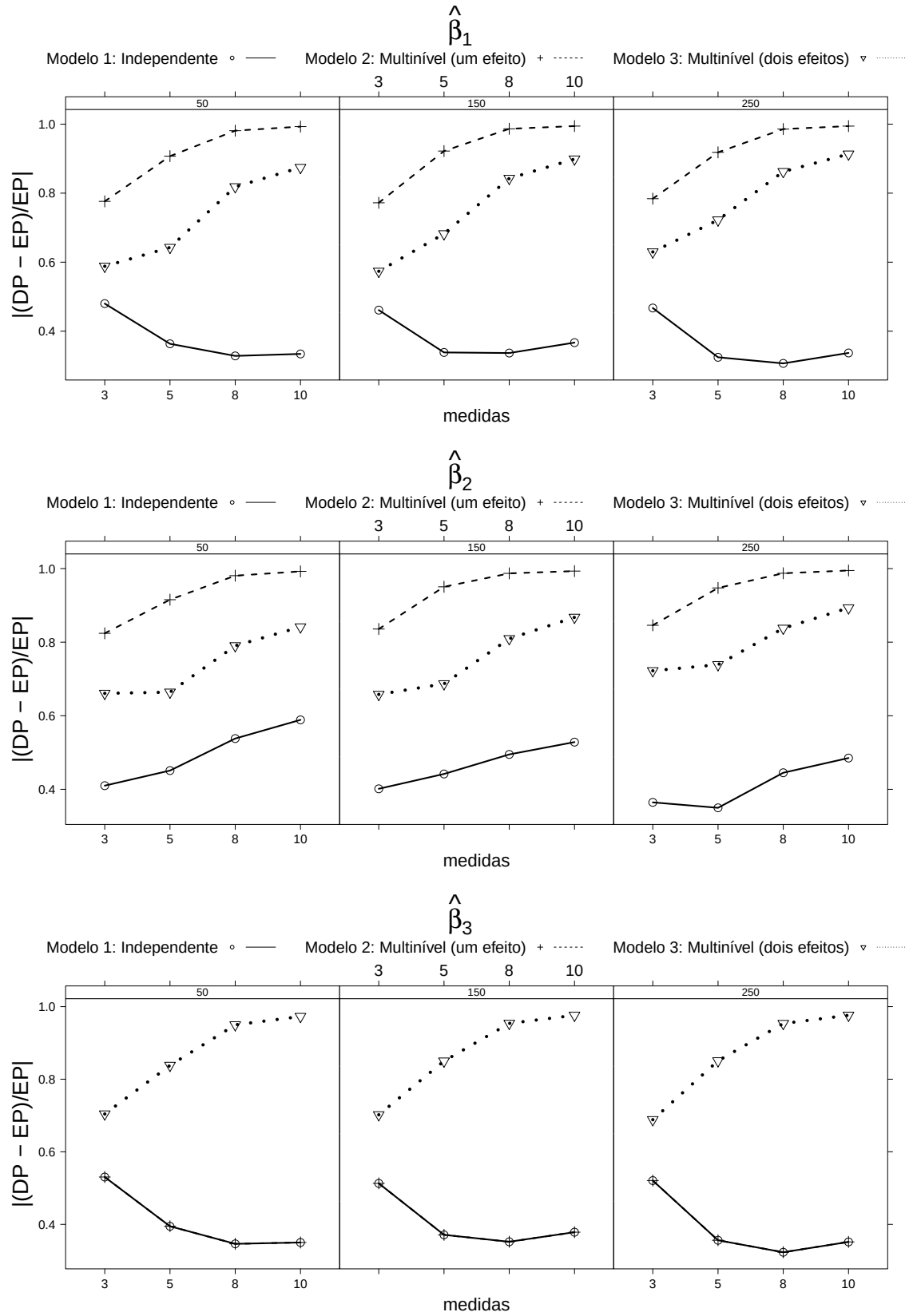


Figura B.11: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 0,7$.

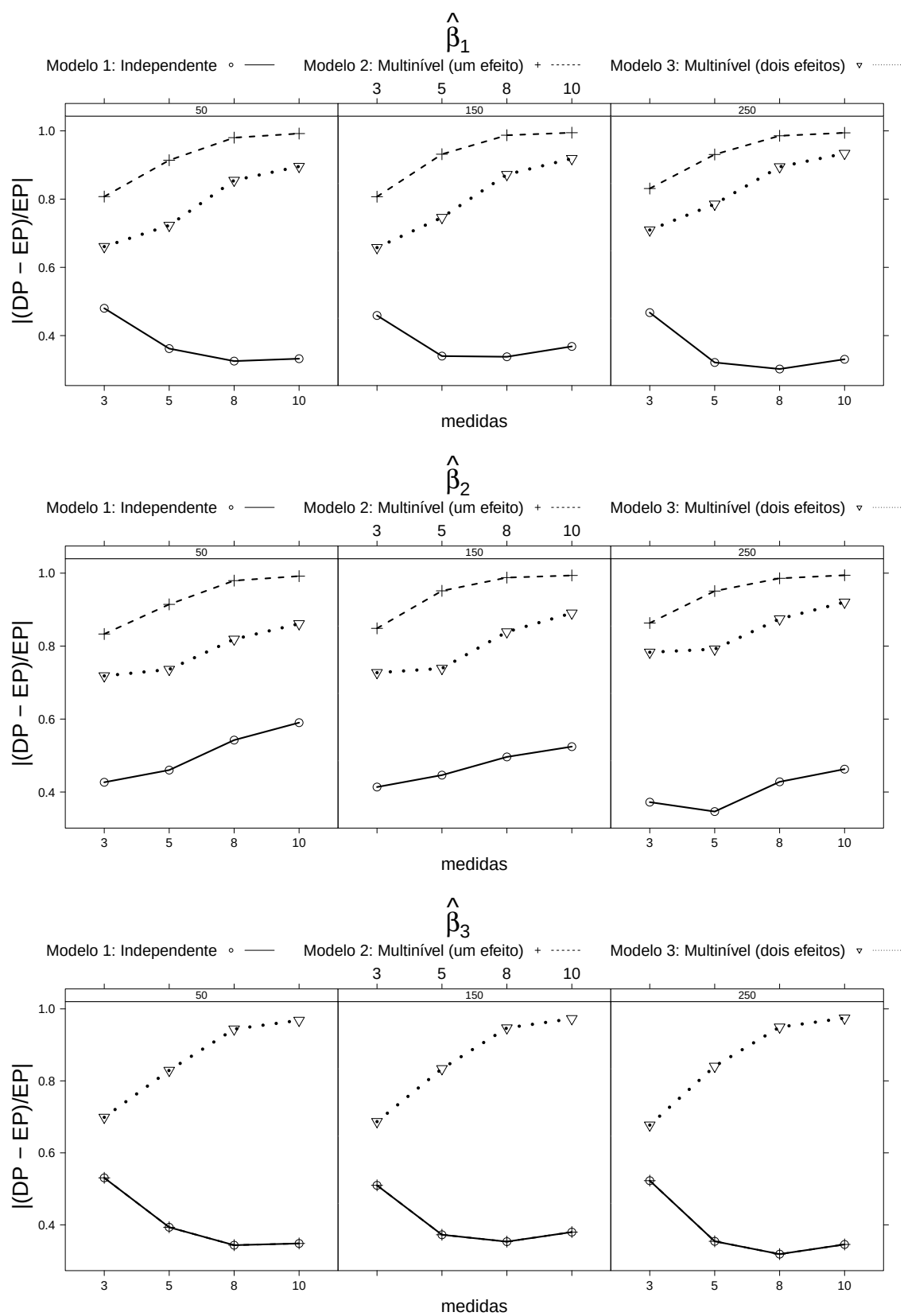


Figura B.12: Diferença relativa das estimativas para diferentes especificações no modelo Poisson considerando o modelo Poisson multinível com dois efeitos aleatórios para a geração dos dados e $\tau_0 = 1, 0$.

B.3 Modelo Poisson Multinível: Critério de Seleção

Tabela B.11: Valores do AIC para diferentes especificações no modelo Poisson considerando 150 indivíduos e gerando os dados utilizando o modelo Poisson multinível com intercepto aleatório.

τ_0	Modelos ajustados	Modelo 2 150 indivíduos			
		3	5	8	10
0,1	Modelo 1	2169,0	5162,9	33003,1	230943,0
	Modelo 2	2164,0	4589,6	9599,4	13778,8
	Modelo 3	2167,5	4627,7	9617,9	13782,2
0,5	Modelo 1	4519,1	26983,6	626418,4	5586838,4
	Modelo 2	2537,0	5053,4	10081,6	14261,9
	Modelo 3	2650,9	5056,8	10085,1	14265,4
0,7	Modelo 1	7132,4	51256,8	1286561,1	11545090,8
	Modelo 2	2635,5	5154,2	10182,8	14363,1
	Modelo 3	2638,9	5157,7	10186,3	14366,6
1,0	Modelo 1	13842,6	113651,6	2983719,2	26863270,5
	Modelo 2	2742,3	5261,9	10290,7	14471,1
	Modelo 3	2745,7	5265,3	10294,1	14474,5

Tabela B.12: Valores do AIC para diferentes especificações no modelo Poisson considerando 250 indivíduos e gerando os dados utilizando o modelo Poisson multinível com intercepto aleatório.

τ_0	Modelos ajustados	Modelo 2 250 indivíduos			
		3	5	8	10
0,1	Modelo 1	3680,0	9199,0	70944,6	528759,4
	Modelo 2	3643,6	7707,5	16061,1	23032,2
	Modelo 3	3646,7	7780,8	16064,4	23035,6
0,5	Modelo 1	9910,3	66790,3	1635586,8	14650273,3
	Modelo 2	4336,7	8564,2	16993,3	23994,5
	Modelo 3	4339,9	8567,5	16996,8	23997,9
0,7	Modelo 1	17710,9	139204,4	3604657,5	32422595,3
	Modelo 2	4528,9	8775,5	17226,6	24242,2
	Modelo 3	4532,1	8778,8	17230,0	24245,6
1,0	Modelo 1	40414,6	350369,3	9348817,1	84268867,9
	Modelo 2	4749,9	9020,3	17504,1	24541,0
	Modelo 3	4753,1	9023,6	17507,5	24544,4

Tabela B.13: Valores do AIC para diferentes especificações no modelo Poisson considerando 150 indivíduos e gerando os dados utilizando o modelo Poisson multinível com dois efeitos aleatórios.

τ_0	Modelos ajustados	Modelo 3 150 indivíduos			
		3	5	8	10
0,1	Modelo 1	13479,2	568335,1	158588728,8	6841761036,9
	Modelo 2	3233,1	16251,2	730138,9	13265556,9
	Modelo 3	2763,1	5527,4	11013,0	15580,8
0,5	Modelo 1	19831,2	759286,6	189201363,8	7638227739,6
	Modelo 2	3369,6	17624,5	760484,0	13680700,9
	Modelo 3	2911,2	5804,9	11429,6	16027,5
0,7	Modelo 1	25797,6	922543,2	215153998,2	8367747288,2
	Modelo 2	3472,5	18807,7	802349,0	14274095,8
	Modelo 3	2981,6	5893,6	11526,0	16126,3
1,0	Modelo 1	40032,8	1293233,3	273504589,5	10063718702,0
	Modelo 2	3673,1	21413,7	907783,0	15749881,1
	Modelo 3	3077,0	5993,6	11631,1	16232,9

Tabela B.14: Valores do AIC para diferentes especificações no modelo Poisson considerando 250 indivíduos e gerando os dados utilizando o modelo Poisson multinível com dois efeitos aleatórios.

τ_0	Modelos ajustados	Modelo 3 250 indivíduos			
		3	5	8	10
0,1	Modelo 1	18234,0	691916,4	167428216,2	6429926828,1
	Modelo 2	5153,4	23068,2	897391,6	14164768,5
	Modelo 3	4528,4	9090,5	18112,8	25609,8
0,5	Modelo 1	24906,8	769669,7	170430200,7	6478934563,8
	Modelo 2	5396,9	24875,1	1022251,7	16295605,9
	Modelo 3	4868,6	9671,6	18962,5	26546,3
0,7	Modelo 1	33392,6	879495,1	177270931,4	6624694931,9
	Modelo 2	5583,4	26567,9	1132074,4	18181035,2
	Modelo 3	5036,9	9876,6	19201,5	26801,5
1,0	Modelo 1	57615,0	1190215,2	197219880,7	7046205064,6
	Modelo 2	5944,2	30468,1	1374919,8	22359817,2
	Modelo 3	5257,7	10130,0	19491,4	27113,0

B.4 Modelo Poisson Marginal (EEG): Tabelas

Tabela B.15: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 150 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação AR1. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3, 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,273918	-0,026082	0,033294	0,274580	-0,02542	0,032787
	5	0,291923	-0,008077	0,012840	0,292985	-0,007015	0,012370
	8	0,299214	-0,000786	0,003202	0,300186	0,000186	0,003296
	10	0,299878	-0,000122	0,001287	0,300357	0,000357	0,001454
$\hat{\beta}_2$	3	2,308608	0,008608	0,016538	2,305873	0,005873	0,015559
	5	2,301099	0,001099	0,004648	2,298958	-0,001042	0,004921
	8	2,300060	0,000060	0,000854	2,299650	-0,000350	0,001025
	10	2,300004	0,000004	0,000283	2,299885	-0,000115	0,000347
$\hat{\beta}_3$	3	1,106749	0,006749	0,009461	1,107085	0,007085	0,009730
	5	1,101503	0,001503	0,002487	1,101586	0,001586	0,002581
	8	1,100095	0,000095	0,000397	1,100011	0,000011	0,000408
	10	1,100012	0,000012	0,000129	1,099972	-0,000028	0,000143
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,270833	-0,029167	0,035863	0,273930	-0,026070	0,033281
	5	0,289459	-0,010541	0,014845	0,292074	-0,007926	0,012778
	8	0,298576	-0,001424	0,003629	0,299225	-0,000775	0,003284
	10	0,299717	-0,000283	0,001409	0,299881	-0,000119	0,001325
$\hat{\beta}_2$	3	2,313608	0,013608	0,020322	2,308477	0,008477	0,016554
	5	2,302369	0,002369	0,005512	2,301070	0,001070	0,004752
	8	2,300146	0,000146	0,000966	2,300056	0,000056	0,000893
	10	2,300022	0,000022	0,000317	2,300007	0,000007	0,000299
$\hat{\beta}_3$	3	1,107223	0,007223	0,009853	1,106771	0,006771	0,009477
	5	1,101959	0,001959	0,002872	1,101475	0,001475	0,002479
	8	1,100177	0,000177	0,000456	1,100094	0,000094	0,000407
	10	1,100029	0,000029	0,000142	1,100012	0,000012	0,000132

Tabela B.16: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 250 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação AR1. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,274367	-0,025633	0,030554	0,274894	-0,025106	0,030117
	5	0,292282	-0,007718	0,011176	0,293250	-0,006750	0,010663
	8	0,299122	-0,000878	0,002566	0,300077	0,000077	0,002576
	10	0,299868	-0,000132	0,000999	0,300337	0,000337	0,001118
$\hat{\beta}_2$	3	2,307354	0,007354	0,012785	2,304734	0,004734	0,011683
	5	2,300788	0,000788	0,003507	2,298882	-0,001118	0,003881
	8	2,300073	0,000073	0,000671	2,299691	-0,000309	0,000806
	10	2,299993	-0,000007	0,000225	2,299886	-0,000114	0,000279
$\hat{\beta}_3$	3	1,106939	0,006939	0,008789	1,107307	0,007307	0,009103
	5	1,101447	0,001447	0,002157	1,101528	0,001528	0,002243
	8	1,100105	0,000105	0,000318	1,100021	0,000021	0,000321
	10	1,100013	0,000013	0,000100	1,099974	-0,000026	0,000110
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,271406	-0,028594	0,033179	0,274365	-0,025635	0,030574
	5	0,289811	-0,010189	0,013277	0,292345	-0,007655	0,011180
	8	0,298454	-0,001546	0,003039	0,299130	-0,000870	0,002600
	10	0,299688	-0,000312	0,001123	0,299864	-0,000136	0,001015
$\hat{\beta}_2$	3	2,312014	0,012014	0,016356	2,307184	0,007184	0,012726
	5	2,301966	0,001966	0,004266	2,300758	0,000758	0,003546
	8	2,300155	0,000155	0,000740	2,300070	0,000070	0,000683
	10	2,300006	0,000006	0,000247	2,299992	-0,000008	0,000232
$\hat{\beta}_3$	3	1,107458	0,007458	0,009240	1,106980	0,006980	0,008832
	5	1,101918	0,001918	0,004266	1,101438	0,001438	0,002162
	8	1,100191	0,000191	0,000382	1,100104	0,000104	0,000323
	10	1,100032	0,000032	0,000114	1,100014	0,000014	0,000102

Tabela B.17: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 50 indivíduos dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,276811	-0,023189	0,040612	0,277691	-0,022309	0,039936
	5	0,291812	-0,008188	0,017273	0,293049	-0,006951	0,016537
	8	0,299407	-0,000593	0,004638	0,300465	0,000465	0,004342
	10	0,299797	-0,000203	0,001935	0,300343	0,000343	0,001827
$\hat{\beta}_2$	3	2,307917	0,007917	0,022796	2,305727	0,005727	0,021524
	5	2,301124	0,001124	0,007174	2,299470	-0,000530	0,006728
	8	2,300050	0,000050	0,001320	2,299669	-0,000331	0,001297
	10	2,300011	0,000011	0,000447	2,299898	-0,000102	0,000422
$\hat{\beta}_3$	3	1,105695	0,005695	0,012005	1,105894	0,005894	0,012011
	5	1,101516	0,001516	0,003334	1,101508	0,001508	0,003329
	8	1,100071	0,000071	0,000574	1,099976	-0,000024	0,000530
	10	1,100020	0,000020	0,000194	1,099974	-0,000026	0,001827
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,274663	-0,025337	0,041944	0,277625	-0,022375	0,039997
	5	0,290640	-0,009360	0,018257	0,294654	-0,005346	0,016985
	8	0,299003	-0,000997	0,004838	0,300876	0,000876	0,004713
	10	0,299690	-0,00031	0,002081	0,300229	0,000229	0,002823
$\hat{\beta}_2$	3	2,310977	0,010977	0,024846	2,304974	0,004974	0,021567
	5	2,301710	0,001710	0,007386	2,298759	-0,001241	0,006883
	8	2,300118	0,000118	0,001434	2,299858	-0,000142	0,001393
	10	2,300020	0,000020	0,000474	2,300015	0,000015	0,000589
$\hat{\beta}_3$	3	1,105979	0,005979	0,012033	1,106118	0,006118	0,012114
	5	1,101730	0,001730	0,003526	1,101317	0,001317	0,003476
	8	1,100122	0,000122	0,000591	1,099905	-0,000095	0,000579
	10	1,100031	0,000031	0,000208	1,099977	0,000023	0,000294

Tabela B.18: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 150 indivíduos dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,272854	-0,027146	0,034214	0,273565	-0,026435	0,033583
	5	0,291137	-0,008863	0,012919	0,292512	-0,007488	0,011967
	8	0,298985	-0,001015	0,003153	0,300220	0,000220	0,002825
	10	0,299817	-0,000183	0,001254	0,300429	0,000429	0,001234
$\hat{\beta}_2$	3	2,309924	0,009924	0,017089	2,307377	0,007377	0,015286
	5	2,301577	0,001577	0,004674	2,299569	-0,000431	0,004173
	8	2,300060	0,000060	0,000882	2,299604	-0,000396	0,000913
	10	2,300020	0,000020	0,000294	2,299879	-0,000121	0,000292
$\hat{\beta}_3$	3	1,106955	0,006955	0,009745	1,107233	0,007233	0,009909
	5	1,101631	0,001631	0,002502	1,101607	0,001607	0,002493
	8	1,100126	0,000126	0,000390	1,100011	0,000011	0,000344
	10	1,100018	0,000018	0,000126	1,099966	-0,000034	0,000122
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,270537	-0,029463	0,036093	0,273269	-0,026731	0,033806
	5	0,289844	-0,010156	0,014006	0,294057	-0,005943	0,011031
	8	0,298501	-0,001499	0,003411	0,300574	0,000574	0,002803
	10	0,299686	-0,000314	0,001394	0,300358	0,000358	0,001149
$\hat{\beta}_2$	3	2,313610	0,013610	0,019777	2,306557	0,006557	0,014859
	5	2,302295	0,002295	0,005049	2,298879	-0,001121	0,004237
	8	2,300136	0,000136	0,000964	2,299827	-0,000173	0,000814
	10	2,300031	0,000031	0,000317	2,299995	-0,000005	0,000227
$\hat{\beta}_3$	3	1,107346	0,007346	0,009986	1,107551	0,007551	0,010144
	5	1,101879	0,001879	0,002708	1,101413	0,001413	0,002368
	8	1,100188	0,000188	0,000421	1,099946	-0,000054	0,000341
	10	1,100032	0,000032	0,000140	1,099965	-0,000035	0,000115

Tabela B.19: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 250 indivíduos dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	Modelo 1: AR1				Modelo 2: Permutável			
	n	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	
$\widehat{\beta}_1$	3	0,273467	-0,026533	0,031233	0,274049	-0,025951	0,030693	
	5	0,290758	-0,009242	0,011802	0,292088	-0,007912	0,010762	
	8	0,298886	-0,001114	0,002583	0,300133	0,000133	0,002199	
	10	0,299791	-0,000209	0,000983	0,300413	0,000413	0,000976	
$\widehat{\beta}_2$	3	2,308481	0,008481	0,013263	2,306169	0,006169	0,011607	
	5	2,301460	0,001460	0,003647	2,299624	-0,000376	0,003157	
	8	2,300084	0,000084	0,000656	2,299663	-0,000337	0,000695	
	10	2,300004	0,000004	0,000218	2,299879	-0,000121	0,000233	
$\widehat{\beta}_3$	3	1,107129	0,007129	0,008983	1,107400	0,007400	0,009166	
	5	1,101729	0,001729	0,002271	1,101692	0,001692	0,002249	
	8	1,100136	0,000136	0,000322	1,100018	0,000018	0,000272	
	10	1,100021	0,000021	0,000099	1,099968	-0,000032	0,000095	
	Modelo 3: Independente				Modelo 4: Não Estruturada			
	n	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	
$\widehat{\beta}_1$	3	0,271292	-0,028708	0,033099	0,273716	-0,026284	0,030958	
	5	0,289593	-0,010407	0,012879	0,293529	-0,006471	0,009672	
	8	0,298385	-0,001615	0,002869	0,300513	0,000513	0,002194	
	10	0,299647	-0,000353	0,001082	0,300345	0,000345	0,000909	
$\widehat{\beta}_2$	3	2,312015	0,012015	0,015949	2,305450	0,005450	0,011210	
	5	2,302049	0,002049	0,003994	2,298975	-0,001025	0,003234	
	8	2,300153	0,000153	0,000720	2,299870	-0,000130	0,000589	
	10	2,300012	0,000012	0,000229	2,299986	-0,000014	0,000168	
$\widehat{\beta}_3$	3	1,107509	0,007509	0,009247	1,107707	0,007707	0,009405	
	5	1,101962	0,001962	0,002485	1,101517	0,001517	0,002108	
	8	1,100201	0,000201	0,000356	1,099950	-0,000050	0,000270	
	10	1,100036	0,000036	0,000109	1,099967	-0,000033	0,000091	

Tabela B.20: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) considerando 50 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correlação permutável. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	Modelo 1: AR1				Modelo 2: Permutável			
	n	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	
$\widehat{\beta}_1$	3	0,278628	-0,021372	0,038271	0,279624	-0,020376	0,037546	
	5	0,295466	-0,004534	0,017228	0,296792	-0,003208	0,016325	
	8	0,300277	0,000277	0,004986	0,301507	0,001507	0,004498	
	10	0,299940	-0,000060	0,001991	0,300364	0,000364	0,002232	
$\widehat{\beta}_2$	3	2,305359	0,005359	0,022613	2,302379	0,002379	0,021500	
	5	2,299949	-0,000051	0,007189	2,297698	-0,002302	0,006733	
	8	2,299953	-0,000047	0,001332	2,299484	-0,000516	0,001151	
	10	2,299988	-0,000012	0,000457	2,299891	-0,000109	0,000533	
$\widehat{\beta}_3$	3	1,105509	0,005509	0,010842	1,105857	0,005857	0,010957	
	5	1,100860	0,000860	0,003247	1,100966	0,000966	0,003141	
	8	1,099972	-0,000028	0,000612	1,099866	-0,000134	0,000538	
	10	1,100007	0,000007	0,000199	1,099972	0,000028	0,000221	
	Modelo 3: Independente				Modelo 4: Não Estruturada			
	n	$\widehat{\beta}$	VM	REQM	$\widehat{\beta}$	VM	REQM	
$\widehat{\beta}_1$	3	0,274646	-0,025354	0,041056	0,279641	-0,020359	0,037622	
	5	0,291074	-0,008926	0,01953	0,297314	-0,002686	0,016691	
	8	0,298974	-0,001026	0,005834	0,301570	0,001570	0,005025	
	10	0,299748	-0,000252	0,002209	0,299947	-0,000053	0,003233	
$\widehat{\beta}_2$	3	2,310846	0,010846	0,026587	2,302292	0,002292	0,021907	
	5	2,301758	0,001758	0,008781	2,297536	-0,002464	0,007283	
	8	2,300120	0,000120	0,001680	2,299479	-0,000521	0,001357	
	10	2,300004	0,000004	0,000506	2,299976	-0,000024	0,001119	
$\widehat{\beta}_3$	3	1,106017	0,006017	0,011061	1,105868	0,005868	0,010986	
	5	1,101606	0,001606	0,003559	1,100884	0,000884	0,003186	
	8	1,100127	0,000127	0,000699	1,099857	-0,000143	0,000600	
	10	1,100026	0,000026	0,000223	1,100006	0,000006	0,000332	

Tabela B.21: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) com 150 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correção permutável. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3, 2, 3; 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,274508	-0,025492	0,032249	0,275210	-0,024790	0,031639
	5	0,295059	-0,004941	0,011482	0,296097	-0,003903	0,010706
	8	0,300127	0,000127	0,003153	0,301589	0,001589	0,003118
	10	0,299878	-0,000122	0,001287	0,300357	0,000357	0,001454
$\hat{\beta}_2$	3	2,307138	0,007138	0,015903	2,303724	0,003724	0,014260
	5	2,300209	0,000209	0,004583	2,297588	-0,002412	0,004683
	8	2,299927	-0,000073	0,000890	2,299376	-0,000624	0,000920
	10	2,300004	0,000004	0,000283	2,299885	-0,000115	0,000347
$\hat{\beta}_3$	3	1,106765	0,006765	0,009065	1,107224	0,007224	0,009389
	5	1,100928	0,000928	0,002178	1,101101	0,001101	0,002181
	8	1,099993	-0,000007	0,000387	1,099860	-0,000140	0,000358
	10	1,100012	0,000012	0,000129	1,099972	-0,000028	0,000143
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,270284	-0,029716	0,035849	0,275183	-0,024817	0,031664
	5	0,289963	-0,010037	0,014662	0,296208	-0,003792	0,010717
	8	0,298529	-0,001471	0,003968	0,301584	0,001584	0,003180
	10	0,299717	-0,000283	0,001409	0,299881	-0,000119	0,001325
$\hat{\beta}_2$	3	2,313675	0,013675	0,020678	2,303689	0,003689	0,014369
	5	2,302456	0,002456	0,006066	2,297555	-0,002445	0,004778
	8	2,300133	0,000133	0,001134	2,299376	-0,000624	0,000942
	10	2,300022	0,000022	0,000317	2,300007	0,000007	0,000299
$\hat{\beta}_3$	3	1,107428	0,007428	0,009558	1,107246	0,007246	0,009411
	5	1,101826	0,001826	0,002704	1,101085	0,001085	0,002181
	8	1,100184	0,000184	0,000479	1,099861	-0,000139	0,000367
	10	1,100029	0,000029	0,000142	1,100012	0,000012	0,000132

Tabela B.22: Valores médios das estimativas, vieses médios e *REQM* das estimativas dos parâmetros para diferentes especificações no modelo Poisson marginal (EEG) com 250 indivíduos e dados gerados pelo modelo Poisson marginal EEG com correção permutável. $\beta^\top = (\beta_1; \beta_2; \beta_3) = (0, 3; 2, 3; 1, 1)$.

	n	Modelo 1: AR1			Modelo 2: Permutável		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,275117	-0,024883	0,029327	0,275635	-0,024365	0,028847
	5	0,294846	-0,005154	0,009496	0,295875	-0,004125	0,008706
	8	0,300110	0,000110	0,002466	0,301568	0,001568	0,002659
	10	0,299868	-0,000132	0,000999	0,300337	0,000337	0,001118
$\hat{\beta}_2$	3	2,305794	0,005794	0,011992	2,302744	0,002744	0,010506
	5	2,300146	0,000146	0,003454	2,297721	-0,002279	0,003760
	8	2,299959	-0,000041	0,000666	2,299453	-0,000547	0,000746
	10	2,299993	-0,000007	0,000225	2,299886	-0,000114	0,000279
$\hat{\beta}_3$	3	1,106921	0,006921	0,008418	1,107369	0,007369	0,008765
	5	1,100987	0,000987	0,001822	1,101139	0,001139	0,001851
	8	1,099993	-0,000007	0,000307	1,099858	-0,000142	0,000301
	10	1,100013	0,000013	0,000100	1,099974	-0,000026	0,000110
	n	Modelo 3: Independente			Modelo 4: Não Estruturada		
		$\hat{\beta}$	VM	REQM	$\hat{\beta}$	VM	REQM
$\hat{\beta}_1$	3	0,271191	-0,028809	0,032798	0,275613	-0,024387	0,028880
	5	0,289869	-0,010131	0,013112	0,295925	-0,004075	0,008742
	8	0,298376	-0,001624	0,003288	0,301550	0,001550	0,002688
	10	0,299688	-0,000312	0,001123	0,299864	-0,000136	0,001015
$\hat{\beta}_2$	3	2,311955	0,011955	0,016553	2,302706	0,002706	0,010573
	5	2,302105	0,002105	0,004665	2,297702	-0,002298	0,003797
	8	2,300154	0,000154	0,000845	2,299454	-0,000546	0,000756
	10	2,300006	0,000006	0,000247	2,299992	-0,000008	0,000232
$\hat{\beta}_3$	3	1,107566	0,007566	0,008935	1,107386	0,007386	0,008788
	5	1,101889	0,001889	0,002441	1,101132	0,001132	0,001854
	8	1,100203	0,000203	0,000401	1,099860	-0,000140	0,000306
	10	1,100032	0,000032	0,000114	1,100014	0,000014	0,000102

B.5 Modelo Poisson Marginal (EEG): Gráficos

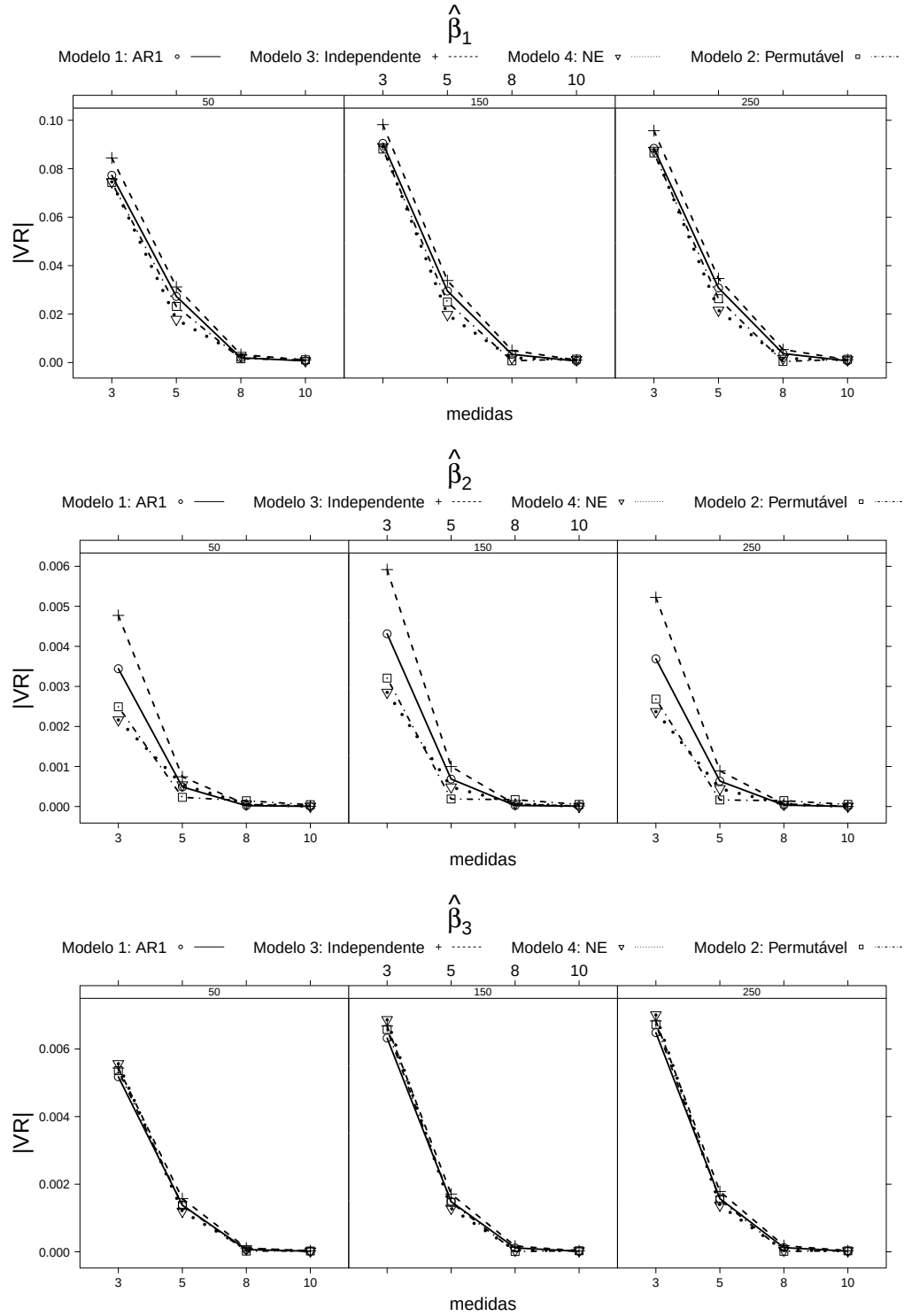


Figura B.13: Valor absoluto do viés relativo considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada.

B.6 Modelo Poisson Marginal (EEG): Critério de Seleção

Tabela B.23: Valores do critério de quasi-verossimilhança (QIC) para a escolha do melhor modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada.

n	Estrutura de correlação dos modelos ajustados	Matriz de correlação de “trabalho” NE		
		m		
		50	150	250
3	AR1	-34900,0	-81151,5	-128942,3
	Permutável	-34900,2	-81151,7	-128942,5
	NE	-34900,3	-81151,7	-128942,5
5	AR1	-531195,8	-1270048,6	-2027111,5
	Permutável	-531196,2	-1270048,8	-2027111,8
	NE	-531196,3	-1270049,0	-2027111,8
8	AR1	-23176966,2	-56320384,6	-90145345,5
	Permutável	-23176966,8	-56320385,0	-90145345,7
	NE	-23176966,1	-56320386,0	-90145346,6
10	AR1	-261935720,8	-639935643,6	-1025223291,9
	Permutável	-261935721,6	-639935644,3	-1025223292,4
	NE	-261928670,5	-639935646,0	-1025223294,2

Tabela B.24: Valores do critério de quasi-verossimilhança (QIC) para a escolha do melhor modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação permutável.

n	Estrutura de correlação dos modelos ajustados	Matriz de correlação de “trabalho” Permutável		
		m		
		50	150	250
3	AR1	-34900,9	-81152,1	-128947,0
	Permutável	-34901,1	-81152,1	-128947,0
	NE	-34901,1	-81152,1	-128947,0
5	AR1	-531125,4	-1269967,3	-2027026,2
	Permutável	-531126,4	-1269968,1	-2027026,8
	NE	-531126,3	-1269968,1	-2027026,8
8	AR1	-23177196,9	-56320221,5	-90145405,1
	Permutável	-23177199,3	-56320223,5	-90145406,7
	NE	-23177197,6	-56320223,6	-90145406,7
10	AR1	-261937052,8	-639933424,3	-1025218744,3
	Permutável	-261937050,9	-639933422,2	-1025218742,0
	NE	-261937022,3	-639933424,6	-1025218744,4

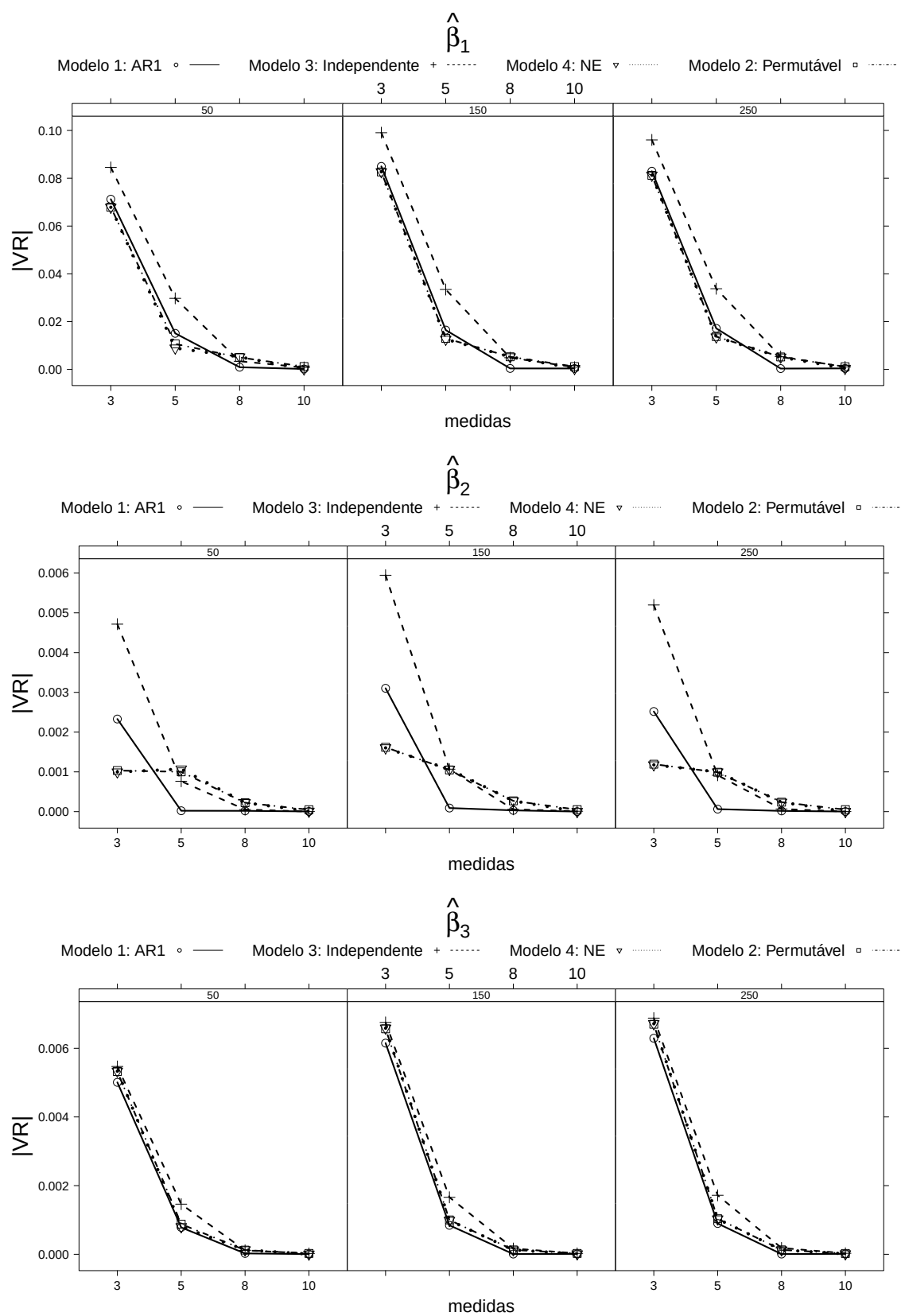


Figura B.14: Valor absoluto do viés relativo considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação permutável.

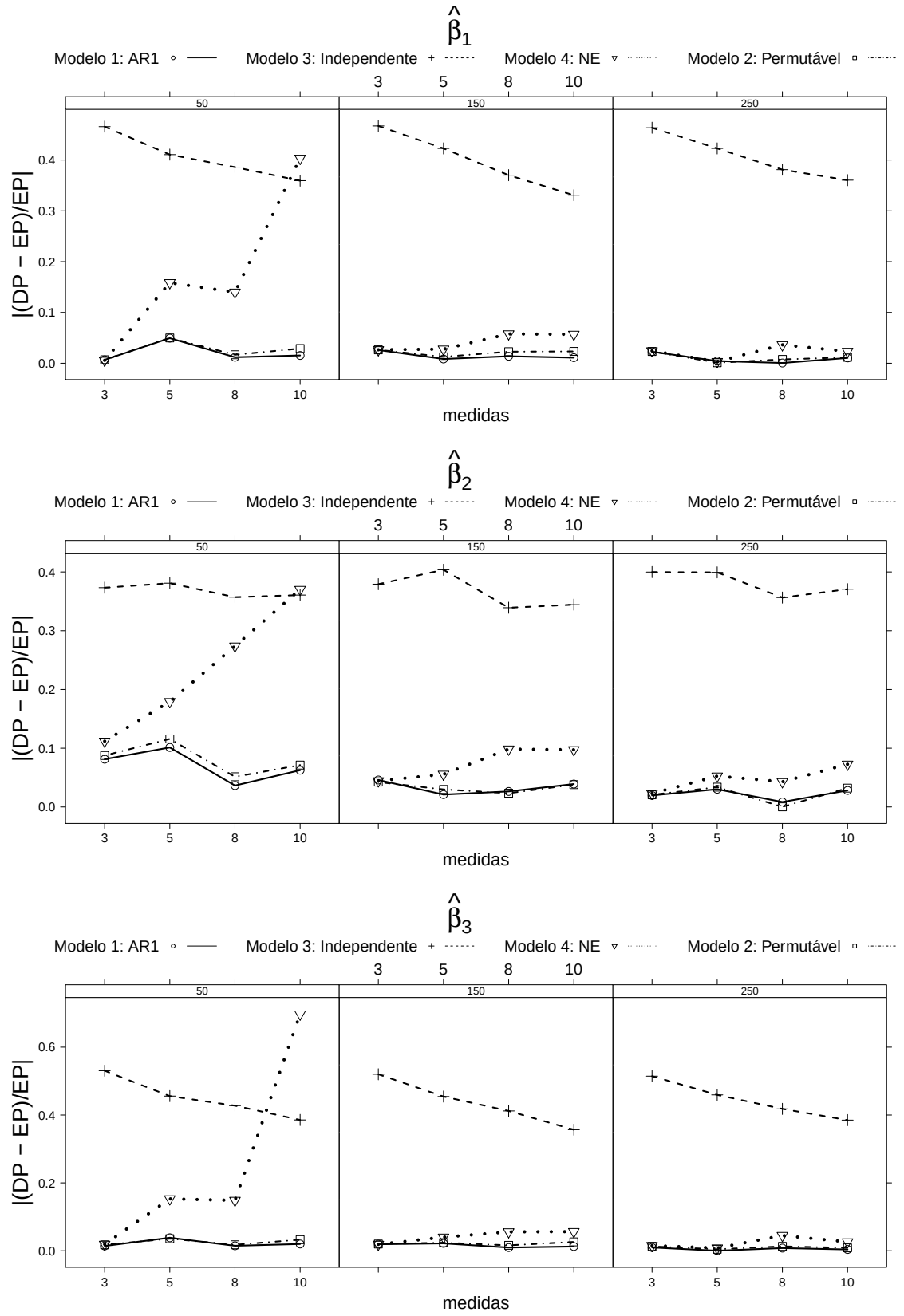


Figura B.15: Diferença relativa dos estimadores considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação não estruturada.

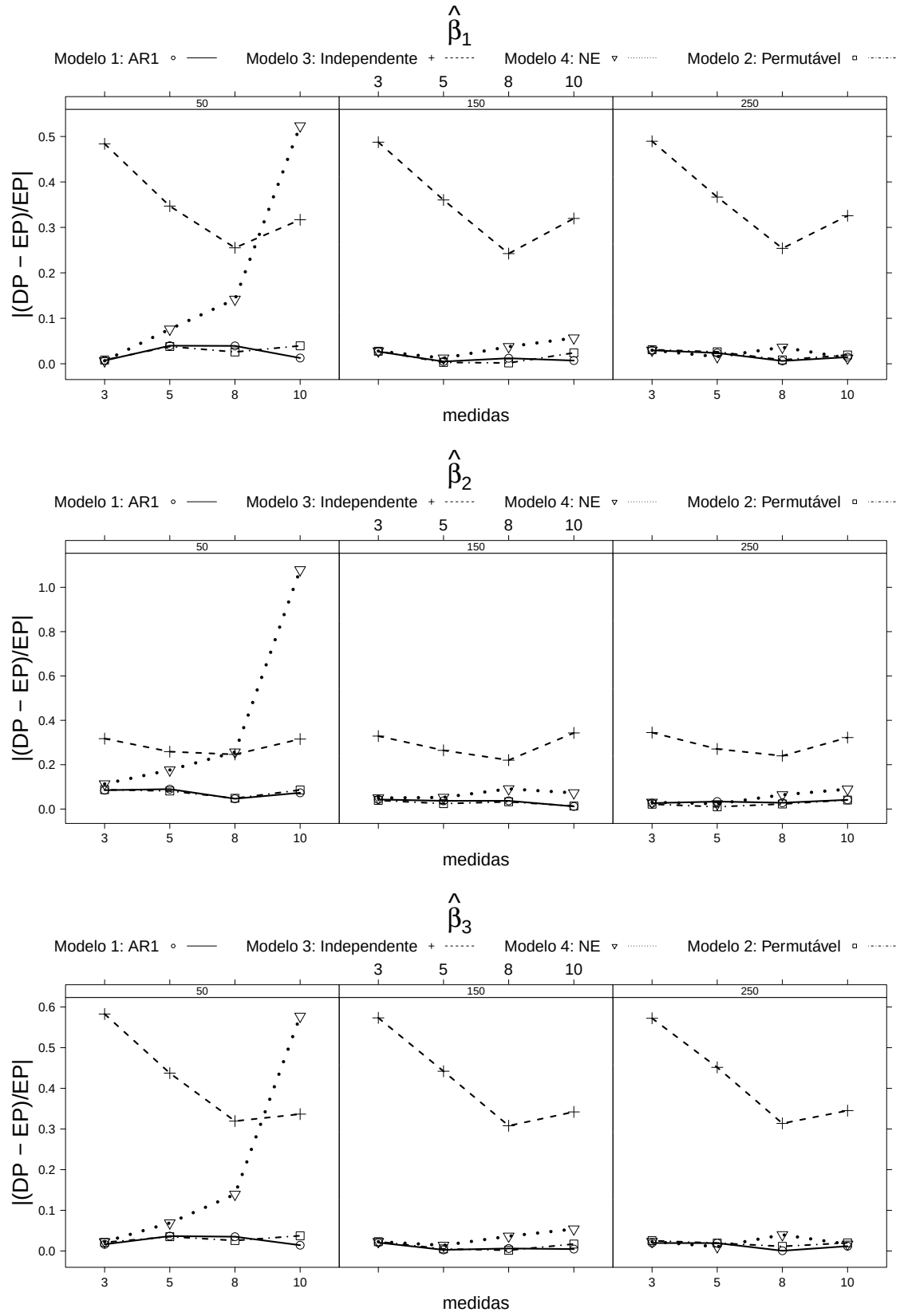


Figura B.16: Diferença relativa dos estimadores considerando o modelo Poisson marginal EEG, com dados gerados pelo modelo Poisson marginal EEG com correlação permutável.