Universidade de São Paulo
Escola Superior de Agricultura "Luiz de Queiroz"

Modelos para análise de dados discretos longitudinais com superdispersão

Fernanda Bührer Rizzato

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Estatística e Experimentação Agronômica

Piracicaba
2011

Fernanda Bührer Rizzato
Licenciada em Matemática

Modelos para análise de dados discretos longitudinais com superdispersão

Orientadora:
Profa. Dra. **CLARICE GARCIA BORGES DEMÉTRIO**

Tese apresentada para obtenção do título de Doutor em Ciências. Área de concentração: Estatística e Experimentação Agronômica

Piracicaba
2011

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - ESALQ/USP

Rizzato Fernanda Bührer
Modelos para análise de dados discretos longitudinais com superdispersão / Fernanda Bührer Rizzato. - - Piracicaba, 2011.
143 p. : il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2011.

1. Análise de dados longitudinais 2. Distribuição de Poisson 3. Distribuição de Bernoulli 4. Inferência Bayesiana 5. Modelos lineares generalizados 6. Modelos mistos I. Título

CDD 519.5
R627m

"Permitida a cópia total ou parcial deste documento, desde que citada a fonte – O autor"

DEDICATÓRIA

À Deus e à Nossa Senhora do Perpétuo
Socorro, minha fonte de fé.

Aos meus amados pais,
Suely *e* ***José Orlando****, pessoas essenciais em minha vida. Por todo o amor, o apoio incondicional, a torcida e a paciência a mim dedicados durante os anos de doutorado e durante todos os anos de minha vida. O amor que recebo de vocês é fundamental em minha vida, é o que me faz forte, assim como a presença de vocês.*

AGRADECIMENTOS

A Deus e a Nossa Senhora do Perpétuo Socorro, por guiarem meus passos, pela minha saúde física e mental, e por serem meu apoio e minha fonte de fé durante esses anos.

A Professora Dra. Clarice G. B. Demétrio a orientação, os momentos compartilhados, os ensinamentos e a paciência. Por ter me aceito como orientada já ao final do segundo ano do doutorado e assumido comigo os riscos e as dificuldades pelas quais passamos.

Aos demais professores do departamento pelos momentos compartilhados, sejam eles, de saber ou de conversas agradáveis nos corredores e eventos durante esses anos. Em especial a Professora Dra. Roseli Leandro, pela paciência e ajuda nos problemas Bayesianos e por tantas conversas e ao Professor Dr. Décio Barbin, simplesmente pela sua presença e pelo carinho e atenção com que sempre me tratou, para mim um exemplo de ser humano e de professor.

Aos funcionários do LCE, a dedicação e a atenção nos momentos solicitados.

Ao Conselho Nacional de Desenvolvimento Científico - CNPq, pela concessão de bolsa de estudo.

Aos meus pais Suely e José Orlando, meus lindinhos, por terem enfrentado comigo essa batalha, por serem minha retaguarda, por estarem ao meu lado em todos os momentos que precisei (que foram muitos!), por me incentivarem e não me deixarem desistir, por toda dedicação e por toda a educação baseada no amor profundo. Vocês são tudo em minha vida.

Às minhas irmãs do coração, Carina e Thais, por estarem há 25 anos presentes em minha vida, por dividirem comigo os momentos de felicidade, por estarem ao meu lado nas dificuldades, pela amizade verdadeira e incondicional.

Às queridas amigas Kelly Galesi e Renata Facco a amizade, o companheirismo e o apoio quando precisei e por tantos momentos inesquecíveis que dividimos.

À querida amiga Barbara Saleh pela amizade verdadeira, o apoio em momentos que eu realmente precisei e pela ligação telefônica mais cara da minha vida.

À querida amiga Joseane Padilha, pela sua presença em minha vida apesar da distância física, o apoio, a amizade e por tantos momentos de muitos risos.

Às amigas Juliana Fachinni, Giovana Oliveira, Andreia Meyer, Elisabeth Mie e

Michele Barbosa por ainda estarem presentes em minha vida apesar da distância, obrigada pelos momentos compartilhados, pelo apoio.

Aos amigos de Piracicaba, que me aguentaram durante crises de estresse, que tiraram de mim vários sorrisos, que me proporcionaram o divertimento e a companhia necessários para aguentar essa dificil jornada: Danila, Danão, Diego (um verdadeiro irmão para mim, obrigada pela companhia em tantos momentos e por todas as conversas!), Fernando, Gabriella, Lydia, Luis Henrique Big, Mantissa, Matola, Pamela, Paulão, Rafaela, Raphaela, Sara, Salim, Tones. Meus dias, com certeza, foram muito mais felizes e divertidos na presença de vocês.

Aos amigos do Programa de Pós-Graduação em Estatística e Experimentação Agronômica, durante esses anos de doutorado convivi com muitas pessoas especiais, com as quais pude compartilhar momentos de estudo e de divertimento.

Ao pessoal da salinha Lu, Vanderly, Simone Sartorio, Renata, Simone Werner, Marina, Wilson, Angela. Obrigada pelas conversas e pela ajuda certa quando precisei.

À minha chefe na UFPR Professora Dra. Nelva Sganzerla pela compreensão, permitindo tantas viagens à Piracicaba para resolver assuntos do doutorado.

Aos professores do LEG-UFPR, Elias T. Krainski, Paulo Justiniano Ribeiro Jr, Silvia E. Shimakura, Walmes M. Zeviani e Wagner H. Bonat pelas sugestões, discussões e conversas e ajuda em programas computacionais.

Ao amigo e Professor Cesar Taconeli, pela companheirismo no trabalho, mas, principalmente, pela presença amiga em Curitiba, pelos almoços, conversas, por escutar meus medos e angustias durante esse término do doutorado.

A toda minha família, tios e primos, em especial aos primos Rodrigo, Lucas e Thais, cada um em seu tempo dividiu comigo um período do doutorado morando na casa dos meus pais e permitindo a mim a convivência com "irmãos"com direito a tudo que irmãos compartilham: alegrias, risadas, divertimento, brigas, chateações, momentos de cantoria e conversas, enfim tudo, principalmente amor.

Enfim, a todas as pessoas que, de forma direta ou indireta, me apoiaram nessa decisão de abdicar de tantas coisas em minha vida, para me dedicar ao doutorado e que torceram por mim para que tudo desse certo.

SUMÁRIO

RESUMO 9
ABSTRACT 11
1 INTRODUÇÃO 13
2 REVISÃO DE LITERATURA 15
2.1 Distribuição de Poisson 15
2.1.1 Função de verossimilhança 15
2.2 Distribuição Bernoulli 16
2.2.1 Função de verossimilhança 16
2.3 Modelos Lineares Generalizados 17
2.4 Superdispersão 18
2.5 Dados longitudinais 21
2.6 Modelo Linear Generalizado Misto 22
2.6.1 Formulação do modelo 22
2.6.2 Estimação pelo método da máxima verossimilhança 22
2.7 Métodos de integração numérica 23
2.7.1 Aproximação do integrando: método de Laplace 23
2.7.2 Quadratura de Gauss-Hermite 24
2.8 Inferência Bayesiana 26
2.8.1 Distribuição a priori 26
2.8.2 Monte Carlo com Cadeias de Markov 27
2.8.3 Abordagem Bayesiana para ajuste do modelo 29
2.8.4 Predição de $\boldsymbol{b}_i$ via *empirical Bayes* 30
2.8.5 Critérios para comparação de modelos 30
2.8.5.1 *Deviance Information Criterion* 30
2.8.5.2 Avaliação do modelo Bayesiano utilizando quantidades pivotais 31
2.8.6 Qualidade de ajuste do modelo aos dados 32
2.9 Modelos para dados na forma de contagens longitudinais com superdispersão 33
2.9.1 Modelo 1: Poisson 34
2.9.2 Modelo 2: Poisson-gama 34

2.9.3 Modelo 3: Poisson-normal 35
2.9.4 Modelo 4: Poisson-normal-gama (*Combinado*) 36
2.10 Modelos para dados binários longitudinais com superdispersão 38
2.10.1 Modelo 1: Bernoulli 39
2.10.2 Modelo 2: Bernoulli-beta 39
2.10.3 Modelo 3: Bernoulli-normal 40
2.10.4 Modelo 4: Bernoulli-normal-beta (*Combinado*) 40
2.11 Método da clonagem de dados 42
3 METODOLOGIA 45
3.1 Material 45
3.1.1 Ensaio clínico para pacientes epilépticos 45
3.1.2 Ensaio clínico para tratamento de dermatite 46
3.2 Métodos 47
3.2.1 Análise Bayesiana para os dados de epilepsia 47
3.2.1.1 Obtenção das quantidades pivotais para avaliação do modelo Bayesiano 51
3.2.1.2 Implementação computacional 52
3.2.1.3 Estudo de simulação 53
3.2.2 Inferência via Dclone e estudo da estimabilidade dos modelos para os dados de epilepsia 55
3.2.3 Análise Bayesiana para os dados de dermatite 56
3.2.3.1 Implementação computacional 60
4 RESULTADOS E DISCUSSÃO 61
4.1 Análise Bayesiana para os dados de epilepsia 63
4.2 Estudo de simulação 78
4.3 Inferência frequentista via Dclone e estudo da estimabilidade dos modelos 84
4.4 Análise Bayesiana para os dados de dermatite 94
5 CONSIDERAÇÕES FINAIS 99
REFERÊNCIAS 101
ANEXOS 105
APÊNDICES 129

RESUMO

Modelos para análise de dados discretos longitudinais com superdispersão

Dados longitudinais na forma de contagens e na forma binária são muito comuns, os quais, frequentemente, podem ser analisados por distribuições de Poisson e de Bernoulli, respectivamente, pertencentes à família exponencial. Duas das principais limitações para modelar esse tipo de dados são: (1) a ocorrência de superdispersão, ou seja, quando a variabilidade dos dados não é adequadamente descrita pelos modelos, que muitas vezes apresentam uma relação pré-estabelecida entre a média e a variância, e (2) a correlação existente entre medidas realizadas repetidas vezes na mesma unidade experimental. Uma forma de acomodar a superdispersão é pela utilização das distribuições binomial negativa e beta binomial, ou seja, pela inclusão de um efeito aleatório com distribuição gama quando se considera dados provenientes de contagens e um efeito aleatório com distribuição beta quando se considera dados binários, ambos introduzidos de forma multiplicativa. Para acomodar a correlação entre as medidas realizadas no mesmo indivíduo podem-se incluir efeitos aleatórios com distribuição normal no preditor linear. Esses situações podem ocorrer separada ou simultaneamente. Molenberghs et al. (2010) propuseram modelos que generalizam os modelos lineares generalizados mistos Poisson-normal e Bernoulli-normal, incorporando aos mesmos a superdispersão. Esses modelos foram formulados e ajustados aos dados, usando-se o método da máxima verossimilhança. Entretanto, para um modelo de efeitos aleatórios, é natural pensar em uma abordagem Bayesiana. Neste trabalho, são apresentados modelos Bayesianos hierárquicos para dados longitudinais, na forma de contagens e binários que apresentam superdispersão. A análise Bayesiana hierárquica é baseada no método de Monte Carlo com Cadeias de Markov (MCMC) e para implementação computacional utilizou-se o software WinBUGS. A metodologia para dados na forma de contagens é usada para a análise de dados de um ensaio clínico em pacientes epilépticos e a metodologia para dados binários é usada para a análise de dados de um ensaio clínico para tratamento de dermatite.

Palavras-chave: Modelos lineares generalizados mistos; Inferência Bayesiana; Análise de dados longitudinais; Modelos mistos; Distribuição de Poisson; Distribuição Bernoulli

ABSTRACT

Models for analysis of longitudinal discrete data in the presence of overdispersion

Longitudinal count and binary data are very common, which often can be analyzed by Poisson and Bernoulli distributions, respectively, members of the exponential family. Two of the main limitations to model this data are: (1) the occurrence of overdispersion, i.e., the phenomenon whereby variability in the data is not adequately captured by the model, and (2) the accommodation of data hierarchies owing to, for example, repeatedly measuring the outcome on the same subject. One way of accommodating overdispersion is by using the negative-binomial and beta-binomial distributions, in other words, by the inclusion of a random, gamma-distributed effect when considering count data and a random, beta-distributed effect when considering binary data, both introduced by multiplication. To accommodate the correlation between measurements made in the same individual one can include normal random effects in the linear predictor. These situations can occur separately or simultaneously. Molenberghs et al. (2010) proposed models that simultaneously generalizes the generalized linear mixed models Poisson-normal and Bernoulli-normal, incorporating the overdispersion. These models were formulated and fitted to the data using maximum likelihood estimation. However, these models lend themselves naturally to a Bayesian approach as well. In this paper, we present Bayesian hierarchical models for longitudinal count and binary data in the presence of overdispersion. A hierarchical Bayesian analysis is based in the Monte Carlo Markov Chain methods (MCMC) and the software WinBUGS is used for the computational implementation. The methodology for count data is used to analyse a dataset from a clinical trial in epileptic patients and the methodology for binary data is used to analyse a dataset from a clinical trial in toenail infection named onychomycosis.

Keywords: Generalized linear mixed models; Bayesian inference; Longitudinal data analysis; Mixed models; Poisson distribution; Bernoulli distribution

1 INTRODUÇÃO

Medidas repetidas podem ser encontradas em diversos áreas, como medicina, agropecuária, entre outras. O termo medidas repetidas é utilizado para descrever medições da mesma característica, na mesma unidade experimental realizadas em mais de uma ocasião ou local. As medidas repetidas podem ser espaciais ou temporais. Como um caso especial de medidas repetidas, têm-se os dados longitudinais, para os quais a variável resposta é observada ao longo do tempo.

A análise estatística desse tipo de dados deve levar em conta a natureza da variável resposta e a correlação existente entre medidas realizadas em um mesmo indivíduo. As variáveis respostas a serem observadas podem ser contínuas, como peso de animais, ou discretas, como a contagem de ocorrências de algum evento ou presença e ausência de determinada característica.

Quando a variável resposta é contínua, admite-se que os dados são provenientes de uma distribuição normal, e a teoria existente para tratar dados longitudinais nesse caso é vasta. Uma metodologia bastante utilizada é a de modelos lineares mistos que proporciona uma grande versatilidade na modelagem da matriz de covariâncias, que poderá acomodar a correlação existente entre medidas repetidas por meio da introdução de efeitos aleatórios no modelo proposto.

Quando a variável resposta é discreta, faz-se o uso da teoria introduzida por Nelder e Wedderburn (1972), chamada de modelos lineares generalizados. O modelo linear generalizado misto, proposto por Breslow e Clayton (1993), permite, por meio da inclusão de efeitos aleatórios no modelo, acomodar a correlação existente entre medidas realizadas em um mesmo indivíduo, e assim é possível a análise de dados longitudinais discretos.

Neste trabalho, será feito o estudo de dados longitudinais na forma de contagens e na forma binária. A distribuição comumente utilizada para analisar dados na forma de contagens é a distribuição de Poisson, que apesar de bastante útil apresenta uma restrição forte, a igualdade da média e da variância dos dados. Quando essa restrição não é satisfeita tem-se falta de ajuste do modelo, chamada de superdispersão. Para tratar a superdispersão, geralmente, pode-se fazer uso da distribuição binomial negativa, que pode surgir pela inclusão de um efeito aleatório com distribuição gama multiplicando a média da distribuição de

Poisson. Dados na forma binária são um caso particular de dados na forma de proporções. A distribuição comumente utilizada para analisar dados na forma de proporções é a distribuição Binomial, para os quais também pode-se observar a presença de superdispersão. Para tratar a superdispersão, geralmente, pode-se fazer uso da distribuição beta binomial, que pode ser obtida pela inclusão de um efeito aleatório com distribuição beta multiplicando a média da distribuição de Binomial.

Portanto, são duas as situações que podem ser tratadas por meio da inclusão de efeitos aleatórios no modelo, a ocorrência da superdispersão e a existência de correlação entre medidas provenientes de um mesmo indivíduo. Essas situações podem ocorrer separada ou simultaneamente. O principal interesse está em se compararem modelos que tratem dessas situações separadamente com um modelo que as acomode ao mesmo tempo. Procedimentos bayesianos hierárquicos baseados em métodos MCMC serão utilizados para a estimação dos parâmetros.

Nos capítulos subsequentes deste trabalho, serão apresentados uma revisão da literatura existente sobre o assunto proposto, que inclui, a definição de um MLGM para dados na forma de contagens com superdispersão e para dados binários com superdispersão, considerando-se duas funções de ligação, logística e probit. Apresentar-se-ão, ainda, uma análise Bayesian hierárquica para esses MLGM com superdispersão baseada em Métodos de Monte Carlo com Cadeias de Markov (MCMC) e uma medida baseada em quantidades pivotais para comparação de modelos bayesianos hierárquicos baseada em Johnson (2007). Os modelos são aplicados a conjuntos de dados reais e um estudo de simulação é realizado para avaliar a eficiência comparação dos modelos através do método das quantidades pivotais. Serão apresentadas, ainda, as considerações finais sobre esse estudo, anexos e apêndices que auxiliarão na compreensão da pesquisa desenvolvida.

2 REVISÃO DE LITERATURA

2.1 Distribuição de Poisson

A distribuição de Poisson foi descoberta por Siméon-Denis Poisson e publicada, conjuntamente com sua teoria de probabilidade, em 1837, no seu trabalho *Recherches sur la probabilité des jugements en matières criminelles et matière civile.* A distribuição de Poisson é de grande importância tanto teórica como aplicada, principalmente, para análise de dados na forma de contagens.

Seja Y uma variável aleatória com distribuição de Poisson, $Y \sim \text{Poisson}(\lambda)$, $\lambda > 0$, com função de probabilidade de Y dada por:

$$P(Y = y) = \frac{e^{-\lambda}\lambda^y}{y!}, \quad \text{para} \quad y = 0, 1, 2, \ldots \tag{1}$$

A esperança de Y é dada por

$$\begin{aligned} \mathrm{E}(Y) &= \sum_{y=0}^{\infty} yP(Y = y) = \sum_{y=0}^{\infty} y\frac{e^{-\lambda}\lambda^y}{y!} \\ &= \lambda e^{-\lambda}\sum_{y=1}^{\infty}\frac{\lambda^{y-1}}{(y-1)!} = \lambda e^{-\lambda}e^{\lambda} = \lambda, \end{aligned}$$

pois $\sum_{y=a}^{\infty}\frac{\lambda^{y-a}}{(y-a)!} = e^{\lambda}$, em que a é uma constante.

A variância de Y pode ser calculada da seguinte forma:

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - [\mathrm{E}(Y)]^2 = \mathrm{E}(Y^2 - Y) + \mathrm{E}(Y) - [\mathrm{E}(Y)]^2.$$

Mas

$$\mathrm{E}(Y^2 - Y) = \sum_{y=0}^{\infty}(y^2 - y)P(Y = y) = \lambda^2 e^{-\lambda}\sum_{y=2}^{\infty}\frac{\lambda^{y-2}}{(y-2)!} = \lambda^2.$$

Assim, $\mathrm{Var}(Y) = \lambda$ e, portanto, para a distribuição de Poisson tem-se que a média e a variância são iguais. Entretanto, na prática, muitas vezes esse fato não é observado, o que pode caracterizar um problema de superdispersão, que será tratado mais adiante.

2.1.1 Função de verossimilhança

Seja $Y_1, Y_2, \ldots, Y_n$ uma amostra aleatória de uma distribuição de Poisson com média $\lambda > 0$. A função de verossimilhança para λ é dada por

$$L(\lambda) = \prod_{i=1}^{n}\frac{e^{-\lambda}\lambda^{y_i}}{y_i!} = \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n} y_i!}$$

cujo logaritmo é dado por:

$$l(\lambda) = -n\lambda + \sum_{i=1}^{n} y_i \log(\lambda) - \sum_{i=1}^{n} \log y_i!.$$

Resolvendo-se a equação $\dfrac{dl(\lambda)}{d\lambda} = 0$, obtém-se a estimativa de máxima verossimilhança

$$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

que é a média amostral dos dados observados.

2.2 Distribuição Bernoulli

A distribuição d Bernoulli, muito utilizada na análise de dados binários, foi introduzida em 1713 no trabalho *Ars conjectandi* escrito por Jakob Bernoulli e publicado por seu sobrinho, Niklaus Bernoulli, oito anos após sua morte.

Seja Y uma variável aleatória com distribuição Bernoulli, $Y \sim$ Bernoulli(π), $0 \leq \pi \leq 1$, com função de probabilidade de Y dada por:

$$P(Y = y) = \pi^y (1-\pi)^{1-y}, \quad \text{para} \quad y = 0 \quad \text{ou} \quad 1 \tag{2}$$

A esperança de Y é dada por

$$\mathrm{E}(Y) = \sum_{y=0}^{\infty} y P(Y = y) = \sum_{y=0}^{1} \pi^y (1-\pi)^{1-y} = \pi$$

e a sua variância é dada por

$$\mathrm{Var}(Y) = \mathrm{E}(Y^2) - [\mathrm{E}(Y)]^2 = \pi - \pi^2 = \pi(1-\pi),$$

pois $\mathrm{E}(Y^2) = \pi$.

Assim, para a distribuição Bernoulli, a média e a variância possuem uma relação estabelecida.

2.2.1 Função de verossimilhança

Seja $Y_1, Y_2, \ldots, Y_n$ uma amostra aleatória de uma distribuição Bernoulli com média π, $\quad 0 \leq \pi \leq 1$. A função de verossimilhança para π é dada por

$$L(\pi) \quad = \quad \prod_{i=1}^{n} \pi^{y_i} (1-\pi)^{1-y_i} = \pi^{\sum_{i=1}^{n} y_i} (1-\pi)^{n - \sum_{i=1}^{n} y_i}$$

cujo logaritmo é dado por:

$$l(\pi) = \frac{\sum_{i=1}^{n} y_i}{\pi} - \frac{n - \sum_{i=1}^{n} y_i}{1-\pi}.$$

Resolvendo-se a equação $\dfrac{dl(\pi)}{d\pi} = 0$, obtém-se a estimativa de máxima verossimilhança

$$\widehat{\pi} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

que é a média amostral dos dados observados.

2.3 Modelos Lineares Generalizados

Por muito tempo, a transformação de dados (BOX; COX, 1964) foi uma ferramenta na análise de dados provenientes de contagens a fim de se utilizar a teoria de modelos lineares clássicos, mas nem sempre os pressupostos de normalidade e homogeneidade de variâncias eram atendidos. Além disso, a utilização de uma teoria que considere a real natureza dos dados produz resultados mais consistentes.

O desenvolvimento da teoria de modelos lineares generalizados, por Nelder e Wedderburn (1972), foi de extrema importância para a análise de dados de contagens e/ou binários, pois representou a unificação de técnicas e metodologias estatísticas que, até então, eram estudadas separadamente.

A variável resposta do modelo deixava de ter, obrigatoriamente, distribuição normal, passando a ter distribuição pertencente a uma família de distribuições chamada família exponencial. Outra generalização encontrada nesse novo modelo é que a relação entre o valor médio da variável resposta e a combinação linear das variáveis explanatórias pode ser estabelecida por qualquer função monótona e diferenciável, o que antes era apenas obtida por meio da função identidade.

Considerando as variáveis aleatórias $Y_1, \ldots, Y_n$ independentes e identicamente distribuídas, com médias $\mu_1, \ldots, \mu_n$, ou seja, $\mathrm{E}(Y_i) = \mu_i$, $i = 1, \ldots, n$, o modelo linear generalizado apresenta três componentes básicos (CORDEIRO; DEMÉTRIO, 2011):

1. A distribuição da variável resposta Y_i pertence à família exponencial de distribuições, com função de probabilidade ou função densidade de probabilidade dada por:

$$f(y_i; \theta_i, \phi) = \exp\{\phi^{-1}[y_i\theta_i - b(\theta_i)] + c(y_i, \phi)\}$$

com média $\mathrm{E}(Y_i) = \mu_i = b'(\theta_i)$ e variância $\mathrm{Var}(Y_i) = \phi b''(\theta_i)$, sendo θ_i o parâmetro canônico e $\phi > 0$, conhecido, o parâmetro de dispersão do modelo.

Pertencem à família exponencial, as distribuições normal, gama e normal inversa para dados contínuos; binomial para proporções; Poisson e binomial negativa para contagens; entre outras.

2. Um componente sistemático, também chamado de preditor linear, ao qual estão relacionadas de forma aditiva as variáveis explanatórias do modelo

$$\eta_i = \boldsymbol{\beta}^T \boldsymbol{x}_i$$

sendo $i = 1, \ldots, n$, $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^T$ a matriz do modelo com $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})^T$, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ o vetor de parâmetros desconhecidos e $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)^T$ o preditor linear.

3. Uma função de ligação monotônica e diferenciável $g(.)$, que liga o preditor linear $\boldsymbol{\eta}$ à média de $\boldsymbol{Y}$, e, portanto,

$$g(\mu_i) = \boldsymbol{\beta}^T x_i \quad \Rightarrow \quad \mu_i = g^{-1}(\boldsymbol{\beta}^T x_i).$$

2.4 Superdispersão

Em algumas situações práticas, quando um modelo linear generalizado é ajustado a dados na forma de contagens (supondo distribuição de Poisson para Y) ou a dados na forma de proporções (supondo distribuição binomial para Y) e assumindo-se que o preditor linear resultante é adequado, pode-se observar a presença de superdispersão ou variabilidade extra-Poisson e variabilidade extra-binomial, respectivamente. Para dados na forma de contagens, em que $Y_i \sim \mathrm{Poisson}(\lambda_i)$, essa variabilidade extra-Poisson é observada quando $\mathrm{Var}(Y_i) > \lambda_i$. Para dados na forma de proporções, em que $Y_i \sim \mathrm{Binomial}(m_i, \pi_i)$, a variabilidade extra-binomial é observada quando $\mathrm{Var}(Y_i) > m_i\pi_i(1 - \pi_i)$. Isso ocorre devido a vários motivos, tais como (HINDE; DEMÉTRIO, 1998):

- variabilidade do material experimental;

- correlação entre indivíduos, por exemplo, quando se trabalha com animais, as medidas feitas em filhotes provenientes da mesma ninhada, provavelmente, são correlacionadas;

- agrupamentos amostrais;

- dados de nível agregado, sendo que o processo de agregação pode levar a distribuições compostas;

- variáveis não observáveis.

Entretanto, em geral, é difícil saber qual é a causa exata da superdispersão, mas é preciso estar atento para sua ocorrência, pois podem-se obter estimativas erradas dos erros padrões, e consequentemente, pode-se avaliar incorretamente a significância dos parâmetros de regressão.

McCullagh e Nelder (1989) e Lindsey (1995) fazem uma discussão geral a respeito da superdispersão enquanto que Breslow (1984) e Lawless (1987) estudam a superdispersão para dados na forma de contagens. Vieira, Hinde e Demétrio (2000) tratam a superdispersão gerada, particularmente, pela inflação de zeros em dados de contagens ou proporções.

Segundo Hinde e Demétrio (1998), as duas principais formas de se tratar a superdispersão são:

(i) admitir uma forma mais geral para a função de variância, possivelmente com a inclusão de novos parâmetros.

(ii) admitir um modelo de dois estágios para a variável resposta.

Hinde e Demétrio (1998) apresentam modelos dos tipos (i) e (ii) para dados na forma de proporções, cuja distribuição padrão é a binomial e para dados na forma de contagens, que têm como distribuição padrão a Poisson, os quais serão apresentados, a seguir.

Considerando que a variável resposta representa contagens com média λ_i e que a superdispersão foi identificada, a forma geral para variância em modelos do tipo (i) pode ser dada por

$$\mathrm{Var}(Y_i) = \lambda_i\big\{1 + \phi\lambda_i^{\delta}\big\}. \tag{3}$$

Várias funções de variância são definidas para diferentes valores de δ e ϕ em (3), tais como:

- se $\phi = 0$, então, tem-se a variância do tipo do modelo Poisson padrão com $\mathrm{Var}(Y_i) = \lambda_i$;

- se $\delta = 0$ tem-se que $\mathrm{Var}(Y_i) = \lambda_i\big\{1 + \phi\big\}$ que é o modelo de superdispersão constante, porém reparametrizado;

- se $\delta = 1$ tem-se $\mathrm{Var}(Y_i) = \lambda_i\big\{1 + \phi\lambda_i\big\}$ que é a função de variância do tipo da distribuição binomial negativa.

É possível perceber que modelos do tipo (i), geralmente, não correspondem à uma distribuição de probabilidade; são apenas vistos como uma extensão do modelo básico. Entretanto, os modelos do tipo (ii), ou seja, modelos em dois estágios, levam a modelos de probabilidade composta para a resposta. Como exemplo, tem-se a distribuição binomial negativa como será mostrado a seguir.

Para o modelo em dois estágios, assume-se que $Y_i|\lambda_i \sim \mathrm{Poisson}(\lambda_i)$ em que λ_i é uma variável aleatória com $\mathrm{E}(\lambda_i) = \mu_i$ e $\mathrm{Var}(\lambda_i) = \sigma_i^2$. Assim, tem-se que $\mathrm{E}(Y_i) = \mu_i$ e $\mathrm{Var}(Y_i) = \mu_i + \sigma_i^2$ o que representa um modelo de superdispersão. Um caso particular é assumir que $\lambda_i \sim \Gamma(k, \theta_i)$, o que resulta em Y_i ter distribuição binomial negativa com $\mathrm{E}(Y_i) = \dfrac{k}{\theta_i} = \mu_i$ e $\mathrm{Var}(Y_i) = \mu_i + \dfrac{\mu_i^2}{k}$. Para valores fixos de k, essa distribuição pertence à família exponencial.

Considerando, agora, que a variável resposta representa o número de sucessos de uma amostra de tamanho m_i e que a superdispersão foi identificada, a forma geral para variância em modelos do tipo (i) pode ser dada por

$$\mathrm{Var}(Y_i) = m_i\pi_i(1-\pi_i)\big[1 + \phi(m_i - 1)^{\delta_1}\{\pi_i(1-\pi_i)\}^{\delta_2}\big]. \tag{4}$$

Várias funções de variância são definidas para diferentes valores de δ_1, δ_2 e ϕ em (4), tais como:

- se $\phi = 0$, então, tem-se a variância do tipo do modelo binomial padrão com $\mathrm{Var}(Y_i) = mi_i\pi_i(1-\pi_i)$;

- se $\delta_1 = \delta_2 = 0$ tem-se que $\mathrm{Var}(Y_i) = mi_i\pi_i(1-\pi_i)\big[1+\phi\big]$ que é o modelo com superdispersão constante, porém reparametrizado;

- se $\delta_1 = 1$ e $\delta_2 = 0$ tem-se $\mathrm{Var}(Y_i) = m_i\pi_i(1-\pi_i)\big[1+\phi(m_i-1)\big]$ que é a função de variância do tipo da distribuição beta-binomial.

Para o modelo em dois estágios, assume-se que $Y_i|P_i \sim \mathrm{Binomial}(m_i, P_i)$ em que P_i é uma variável aleatória com $\mathrm{E}(P_i) = \pi_i$ e $\mathrm{Var}(P_i) = \phi\pi_i(1-\pi_i)$. Assim, tem-se que $\mathrm{E}(Y_i) = m_i\pi_i$ e $\mathrm{Var}(Y_i) = m_i\pi_i\big[1+\phi(m_i-1)\big]$ o que representa um modelo de superdispersão. Um caso particular é assumir que $P_i \sim \mathrm{Beta}(\alpha_i, \beta_i)$, com $\alpha_i + \beta_i$ constante, o que resulta em Y_i ter distribuição beta-binomial.

2.5 Dados longitudinais

Segundo Diggle (1988), o termo medidas repetidas destina-se a medidas feitas na mesma variável ou na mesma unidade experimental em mais de uma ocasião. Um estudo básico de medidas repetidas consiste em um experimento em que os tratamentos são completamente aleatorizados às unidades experimentais e os dados são coletados mais de uma vez para cada uma dessas unidades experimentais.

Dados longitudinais são um caso especial de medidas repetidas, ou seja, são medidas repetidas nas quais as observações são ordenadas pelo tempo ou pela posição no espaço. De forma mais geral, dados longitudinais podem ser definidos como medidas repetidas nas quais as observações nos indivíduos não foram aleatorizados para os níveis de um "tratamento"de interesse (geralmente, tempo ou posição no espaço), dando origem a correlações entre medidas.

Pode-se dizer que os dados longitudinais são regulares em relação ao tempo, quando o intervalo entre duas medidas consecutivas quaisquer é constante durante o estudo, e são balanceados em relação ao tempo, quando as observações são realizadas nos mesmos instantes de tempo em todas as unidades experimentais (HELMS, 1992).

Os ensaios longitudinais têm como objetivo estudar o comportamento da variável resposta ao longo do tempo e analisar a sua dependência em relação às covariáveis.

A principal característica dos estudos longitudinais é que, devido ao fato de as observações serem repetidas em um mesmo indivíduo, essas medidas tendem a ser correlacionadas. Tal correlação pode ser modelada por meio de uma estrutura de covariâncias, sendo que, para grande parte dos conjuntos de dados, é usual assumir independência. Para que as inferências sobre as médias sejam válidas, é essencial admitir uma estrutura de covariâncias adequada que pode ser criada pelo uso de efeitos aleatórios. Logo, podem-se usar os modelos de efeitos aleatórios em dois estágios, em que as variáveis respostas pertencem a uma mesma família mas alguns de seus parâmetros de efeitos aleatórios variam de indivíduo para indivíduo. Esses efeitos terão uma distribuição especificada no segundo estágio do modelo.

2.6 Modelo Linear Generalizado Misto

A teoria de modelos lineares generalizados considera apenas o estudo de variáveis com efeitos fixos e uma extensão natural são modelos que contemplem também efeitos aleatórios. Assim, surge o modelo linear generalizado misto (MLGM), proposto por Breslow e Clayton (1993) e muito utilizado para análise de dados discretos de medidas repetidas.

Devido ao constante avanço computacional, vários softwares permitem o ajuste desses modelos, dentre eles o SAS (SAS Institute, 2008) e o R (R DEVELOPMENT CORE TEAM, 2009)

2.6.1 Formulação do modelo

Seja Y_{ij} a j-ésima medida no i-ésimo indivíduo, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$ e $\boldsymbol{b}_i$ o vetor q-dimensional de efeitos aleatórios e com distribuição $N(\mathbf{0},\mathbf{D})$, sendo $\mathbf{0}$ um vetor de zeros e $\mathbf{D}$ uma matriz de variâncias-covariâncias. Então, $\boldsymbol{Y}_i$ é o vetor de dimensão n_i com todas as observações do indivíduo i, as medidas Y_{ij} são independentes com função (densidade) de probabilidade condicional a $\boldsymbol{b}_i$ pertencente à família exponencial

$$f(y_{ij}|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi) = \exp\{\phi^{-1}[y_{ij}\theta_{ij} - \psi(\theta_{ij})] + c(y_{ij}, \phi)\}$$

sendo $\eta(.)$ a função de ligação, ou seja,

$$\eta(\mu_{ij}) = \eta[\mathrm{E}(Y_{ij}|\boldsymbol{b}_i)] = \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i$$

em que $\boldsymbol{x}_{ij}$ é o vetor de dimensão p associado às covariáveis e $\boldsymbol{z}_{ij}$ é o vetor de dimensão q associado aos efeitos aleatórios, $\boldsymbol{\beta}$ é o vetor p-dimensional de efeitos fixos e ϕ o parâmetro de dispersão. Denota-se, também, por $f(\boldsymbol{b}_i|\boldsymbol{D})$ a função densidade dos efeitos aleatórios $\boldsymbol{b}_i$.

A função densidade de probabilidade condicional para $\boldsymbol{Y}_i$, denotada por $f(\boldsymbol{y}_i|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi)$ é dada por:

$$f(\boldsymbol{y}_i|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi) = \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi).$$

2.6.2 Estimação pelo método da máxima verossimilhança

A estimação do vetor de parâmetros $\boldsymbol{\beta}$ e dos parâmetros em $\boldsymbol{D}$ é feita pela maximização da função de verossimilhança marginal, obtida pela integração da função de

verossimilhança em relação aos efeitos aleatórios $\boldsymbol{b}_i$. A contribuição do i-ésimo indivíduo para a função de verossimilhança é

$$\begin{aligned} f(\boldsymbol{y}_i|\boldsymbol{\beta},\boldsymbol{D},\phi) &= \int f(\boldsymbol{y}_i|\boldsymbol{b}_i,\boldsymbol{\beta},\phi)f(\boldsymbol{b}_i|\boldsymbol{D})d\boldsymbol{b}_i \\ &= \int \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{b}_i,\boldsymbol{\beta},\phi)f(\boldsymbol{b}_i|\boldsymbol{D})d\boldsymbol{b}_i, \end{aligned} \tag{5}$$

e, portanto, a função de verossimilhança para $\boldsymbol{\beta}$, $\boldsymbol{D}$ e ϕ é

$$\begin{aligned} L(\boldsymbol{\beta},\boldsymbol{D},\phi) &= \prod_{i=1}^{N} f(\boldsymbol{y}_i|\boldsymbol{\beta},\boldsymbol{D},\phi) \\ &= \prod_{i=1}^{N} \int \prod_{j=1}^{n_i} f(y_{ij}|\boldsymbol{b}_i,\boldsymbol{\beta},\phi)f(\boldsymbol{b}_i|\boldsymbol{D})d\boldsymbol{b}_i \end{aligned} \tag{6}$$

O problema em maximizar a equação (6) é a presença de N integrais sob os efeitos aleatórios $\boldsymbol{b}_i$. Em alguns casos, essas integrais possuem solução analítica, mas, em outros não, e assim, faz-se necessário o uso de métodos numéricos. Uma outra opção a ser utilizada é a inferência Bayesiana, que garante o ajuste do modelo hierárquico.

2.7 Métodos de integração numérica

O cálculo de integrais é um problema que exige atenção em inferência estatística, pois em muitos problemas reais as integrais são bastente complexas e nem sempre possuem solução analítica. Assim, métodos numéricos são utilizados a fim de se encontrar uma forma fechada aproximada para as integrais em questão. Problemas bayesianos também demandam aproximações numéricas de integrais, uma vez que nem sempre as distribuições a posteriori conjunta possuem formas conhecidas.

2.7.1 Aproximação do integrando: método de Laplace

O método de Laplace, bastante utilizado na inferência Bayesiana, foi introduzido para aproximar integrais da forma

$$I = \int e^{Q(\boldsymbol{b})} d\boldsymbol{b} \tag{7}$$

em que $Q(\boldsymbol{b})$ é uma função conhecida, unimodal e limitada da variável q-dimensional $\boldsymbol{b}$. Seja $\widehat{\boldsymbol{b}}$ o valor de $\boldsymbol{b}$ que maximiza $Q(\boldsymbol{b})$. Aplicando a expansão em séries de Taylor de segunda

ordem a $Q(\boldsymbol{b})$, obtém-se

$$Q(\boldsymbol{b}) = Q(\widehat{\boldsymbol{b}}) + \frac{1}{2}(\boldsymbol{b} - \widehat{\boldsymbol{b}})' Q''(\widehat{\boldsymbol{b}})(\boldsymbol{b} - \widehat{\boldsymbol{b}}) \tag{8}$$

com $Q''(\widehat{\boldsymbol{b}})$ a matriz de derivadas de segunda ordem de Q calculadas no ponto $\widehat{\boldsymbol{b}}$. Substituindo (8) em (7), obtém-se

$$I \approx (2\pi)^{q/2} \mid -Q''(\widehat{\boldsymbol{b}}) \mid^{-1/2} e^{Q(\widehat{\boldsymbol{b}})}.$$

Cada integral na função de verossimilhança (6) é proporcional a (7) com $Q(\boldsymbol{b})$ dada por

$$Q(\boldsymbol{b}) = \phi^{-1} \sum_{j=1}^{n_i} [y_{ij}(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}) - \psi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b})] - \frac{1}{2}\boldsymbol{b}'\boldsymbol{D}^{-1}\boldsymbol{b}.$$

O valor de $\widehat{\boldsymbol{b}}$ depende dos parâmetros $\boldsymbol{\beta}$, ϕ e $\boldsymbol{D}$. Assim, em cada iteração da maximização numérica da função de verossimilhança, $\widehat{\boldsymbol{b}}$ será recalculado condicional aos valores atuais para que os parâmetros sejam estimados. A aproximação de Laplace é exata quando $Q(\boldsymbol{b})$ é igual ao núcleo da função densidade normal.

2.7.2 Quadratura de Gauss-Hermite

As quadraturas Gaussiana e Gaussiana adaptativa são utilizadas para aproximar integrais da forma

$$I = \int f(z)w(z)dz, \tag{9}$$

em que $f(z)$ é uma função conhecida e $w(z)$ é a função densidade da distribuição normal (multivariada).

Segundo Molenberghs e Verbeke (2005), para que (5) assuma a forma (9) é necessário apenas que seja feita uma padronização na matriz do efeito aleatório $\boldsymbol{b}_i$. Tomando o efeito aleatório $\boldsymbol{\delta}_i = \boldsymbol{D}^{-1/2}\boldsymbol{b}_i$, obtém-se um novo efeito aleatório com distribuição normal q-dimensional, com média $\mathbf{0}$ e matriz de variância-covariância $\mathbf{I}$. O preditor linear torna-se $\eta_{ij} = x_{ij}^T\boldsymbol{\beta} + z_{ij}^T\boldsymbol{D}^{1/2}\boldsymbol{\delta}_i$. Assim a função de verossimilhança (5) para o indivíduo i pode ser expressa da seguinte forma:

$$f_i(\boldsymbol{y}_i|\boldsymbol{\beta}, \boldsymbol{D}, \phi) = \int \prod_{j=1}^{n_i} f_{ij}(y_{ij}|\boldsymbol{\delta}_i, \boldsymbol{\beta}, \phi) f(\boldsymbol{\delta}_i) d\delta_i,$$

estando em condições de se utilizar o método de aproximação.

A aproximação pela quadratura de Gauss-Hermite é dada por

$$\int f(z)w(z)dz = \sum_{q=1}^{Q} w_q f(z_q),$$

em que Q é a ordem da aproximação, z_q são os nós, os pontos da quadratura nos quais a função será avaliada e w_q são os pesos escolhidos apropriadamente. Os valores de z_q e w_q são obtidos a partir dos polinômios ortogonais de Hermite. Para a quadratura de Gauss-Hermite $w(z) = \exp(-z^2/2)$.

Entretanto, a qualidade da aproximação da integral é melhor se $f(z)w(z)$ se comporta de forma similar a $w(z)$, o que nem sempre acontece. Assim, deve-se reescalonar e deslocar os pontos da quadratura de modo a aumentar a precisão.

Para o método da quadratura Gaussiana adaptativa, supõe-se que $f(z)w(z)$ segue, aproximadamente, uma distribuição normal com média igual à moda $\widehat{z}$ de $ln[f(z)w(z)]$ e variância dada por

$$\left[-\frac{\partial^2}{\partial z^2} ln[f(z)w(z)] \bigg|_{z=\widehat{z}} \right]^{-1}.$$

Assim, os novos pontos da quadratura (agora adaptativos) são dados por

$$z_q^+ = \widehat{z} + \left[-\frac{\partial^2}{\partial z^2} ln[f(z)w(z)] \bigg|_{z=\widehat{z}} \right]^{-1/2} z_q.$$

com seus respectivos pesos

$$w_q^+ = \left[-\frac{\partial^2}{\partial z^2} ln[f(z)w(z)] \bigg|_{z=\widehat{z}} \right]^{-1/2} \frac{w(z_q^+)}{w(z_q)} w_q.$$

Como anteriormente, a aproximação da integral é dada por

$$\int f(z)w(z)dz \approx \sum_{q=1}^{Q} w_q^+ f(z_q^+).$$

Usualmente, o método da quadratura Gaussiana adaptativa necessita de um número bem menor de pontos de quadratura, quando comparado ao método de Gauss-Hermite. Quando apenas um ponto de quadratura é assumido, ou seja, $Q = 1$, mostra-se que o método de quadratura Gaussiana adaptativa reduz-se ao método de Laplace (MOLENBERGHS; VERBEKE, 2005).

2.8 Inferência Bayesiana

A inferência Bayesiana paramétrica, assim como a frequentista, considera que um conjunto de observações y pode ser descrito por uma distribuição de probabilidade $f(y|\boldsymbol{\vartheta})$, em que $\boldsymbol{\vartheta}$ é o vetor desconhecido de parâmetros. Pode-se dizer que um dos principais objetivos da estatística é fazer inferências sobre $\boldsymbol{\vartheta}$, e muitas vezes o pesquisador possui alguma informação a priori a respeito dessa quantidade. A teoria Bayesiana diferencia-se da frequentista por permitir que essas informações prévias a respeito dos parâmetros de interesse para a análise sejam incorporadas ao modelo. Assim, a teoria Bayesiana permite a incorporação da incerteza relativa às quantidades desconhecidas na análise dos dados, enquanto a frequentista não.

A abordagem Bayesiana necessita, além da informação sobre os dados, que é incorporada na análise por meio da função de verossimilhança, de uma distribuição a priori para as quantidades desconhecidas no modelo (parâmetros e dados faltantes). Essas duas informações são usadas para obter a distribuição dos parâmetros, condicional aos dados observados, chamada distribuição a posteriori, da qual resultam todas as inferências estatísticas sobre os parâmetros do modelo (CARLIN; LOUIS, 1996). A distribuição a posteriori $\pi(\boldsymbol{\vartheta}|\boldsymbol{y})$ é dada por

$$\pi(\boldsymbol{\vartheta}|\boldsymbol{y}) = \frac{L(\boldsymbol{\vartheta}|\boldsymbol{y})\pi(\boldsymbol{\vartheta})}{f(\boldsymbol{y})}$$

em que

$$\begin{aligned} f(\boldsymbol{y}) &= \int_{\boldsymbol{\vartheta}} f(\boldsymbol{y},\boldsymbol{\vartheta})d\boldsymbol{\vartheta} = \int_{\boldsymbol{\vartheta}} L(\boldsymbol{\vartheta}|\boldsymbol{y})\pi(\boldsymbol{\vartheta})d\boldsymbol{\vartheta}, \quad \text{para o caso contínuo} \\ f(\boldsymbol{y}) &= \sum_{\boldsymbol{\vartheta}} f(\boldsymbol{y},\boldsymbol{\vartheta})d\boldsymbol{\vartheta} = \sum_{\boldsymbol{\vartheta}} L(\boldsymbol{\vartheta}|\boldsymbol{y})\pi(\boldsymbol{\vartheta}), \quad \text{para o caso discreto,} \end{aligned}$$

$\pi(\boldsymbol{\vartheta})$ é a distribuição a priori e $L(\boldsymbol{\vartheta}|\boldsymbol{y})$ é a função de verossimilhança. Como $f(\boldsymbol{y})$ não depende de $\boldsymbol{\vartheta}$, ela é apenas uma constante normalizadora que não tem influência sobre a estimação dos parâmetros. Assim, a distribuição a posteriori pode ser expressa como

$$\pi(\boldsymbol{\vartheta}|\boldsymbol{y}) \propto L(\boldsymbol{\vartheta}|\boldsymbol{y})\pi(\boldsymbol{\vartheta}).$$

2.8.1 Distribuição a priori

Segundo Box e Tiao (1992), a distribuição a priori representa o que se conhece a respeito do vetor de parâmetros de interesse, $\boldsymbol{\vartheta}$, antes da realização do experimento. Assim,

essa distribuição pode ser usada para representar o estado de conhecimento ou ignorância sobre o vetor de parâmetros $\boldsymbol{\vartheta}$ antes de se obterem os dados. Existem algumas formas de escolha das distribuições a priori.

A distribuição a priori não informativa é utilizada quando não existem informações a priori sobre os parâmetros de interesse ou quando as informações disponíveis são pouco significativas. Logo, a informação resultante da distribuição a posteriori será referente apenas aos dados. Quando não se tem informação suficiente sobre o modelo, seja por parte do estatístico ou por parte do pesquisador, utiliza-se a distribuição a priori subjetiva, que representa a expectativa subjetiva do pesquisador em relação a esses parâmetros. Para que se possa quantificar essa informação a priori a interação entre o estatístico e o pesquisador é muito importante. Diversos procedimentos para essa quantificação e consequente transformação em um distribuição a priori têm sido discutidos. Esses procedimentos são conhecidos na literatura Bayesiana como elicitação. Também, tem-se a distribuição a priori conjugada, o que ocorre quando as distribuições a priori e a posteriori pertencem à mesma família de distribuições. Entretanto, seu uso deve ser cuidadoso, pois apesar da facilidade analítica, muitas vezes não representa adequadamente a informação a priori do parâmetro.

2.8.2 Monte Carlo com Cadeias de Markov

Na abordagem Bayesiana, as inferências sobre os parâmetros de interesse são realizadas, usando-se a distribuição marginal a posteriori. Mas, na maioria das vezes, para obtenção dessa distribuição marginal a posteriori necessita-se da solução de integrais complexas, assim é possível recorrer aos métodos de Monte Carlo com cadeias de Markov (MCMC).

Os métodos MCMC para a obtenção de uma amostra de valores de $\boldsymbol{\vartheta}$, os quais não podem ser amostrados diretamente de $\pi(\boldsymbol{\vartheta}|\boldsymbol{y})$, são baseados na simulação de cadeias de Markov a partir de distribuições aproximadas que são atualizadas até a obtenção da distribuição estacionária $\pi(\boldsymbol{\vartheta}|\boldsymbol{y})$. As amostras são obtidas sequencialmente, com a distribuição dos valores dependendo apenas do último valor amostrado, formando assim, uma cadeia de Markov.

Assim, os métodos MCMC fazem uma associação entre os algoritmos para simulação de distribuições e o método Monte Carlo para aproximações de integrais. Entretanto, o sucesso do método não é devido à propriedade da cadeia de Markov, mas das

distribuições aproximadas que melhoram a cada passo da simulação a fim de convergir para a distribuição de interesse. Dois algoritmos que permitem gerar cadeias de Markov cuja distribuição de equilíbrio seja $\pi(\boldsymbol{\vartheta}|\boldsymbol{y})$ são Metropolis-Hasting e Gibbs sampling, que é um caso particular do anterior.

Gibbs sampling

O algoritmo Gibbs-sampling permite gerar amostras de uma distribuição conjunta a posteriori $\pi(\vartheta_1, \vartheta_2, \ldots, \vartheta_p|\boldsymbol{y})$ a partir das distribuições condicionais completas a posteriori de cada parâmetro $\pi(\vartheta_i|\vartheta_1, \vartheta_{i-1}, \ldots, \vartheta_{i+1}, \vartheta_p, \boldsymbol{y})$, desde que essas distribuições possuam forma fechada (CASELLA; GEORGE, 1992). Para implementação do algoritmo são necessários os seguintes passos:

(i) inicie o contador em $t = 1$ iterações;

(ii) especifique um valor inicial para $\boldsymbol{\vartheta} = (\vartheta_1^0, \vartheta_2^0, \ldots, \vartheta_p^0)$;

(iii) a partir de ϑ^{t-1}, obtenha um novo vetor $\boldsymbol{\vartheta}_t = (\vartheta_1^t, \vartheta_2^t, \ldots, \vartheta_p^t)$ em que:

$$\vartheta_1^t \sim \pi(\vartheta_1|\vartheta_2^{t-1}, \vartheta_3^{t-1}, \ldots, \vartheta_p^{t-1}, \boldsymbol{y})$$

$$\vartheta_2^t \sim \pi(\vartheta_2|\vartheta_1^t, \vartheta_3^{t-1}, \ldots, \vartheta_p^{t-1}, \boldsymbol{y})$$

$$\vdots$$

$$\vartheta_p^t \sim \pi(\vartheta_p|\vartheta_1^t, \vartheta_2^t, \ldots, \vartheta_{p-1}^t, \boldsymbol{y});$$

(iv) incremente o contador de t para $t + 1$ e volte ao passo (iii).

Metropolis-Hasting

O algoritmo Metropolis-Hasting (CHIB; GREENBERG, 1995) é um dos algoritmos usados para gerar amostras da distribuição conjunta a posteriori $\pi(\boldsymbol{\vartheta}|\boldsymbol{y})$, principalmente, quando a distribuição condicional completa a posteriori do parâmetro não possui forma fechada.

Considere uma distribuição $q(\vartheta, \phi)$ e seja a cadeia de Markov $(\vartheta^0, \vartheta^1, \ldots, \vartheta^t, \ldots)$, em que t é o estado atual da cadeia. Para gerar um valor que será candidato ao próximo

estado da cadeia, $(t+1)$, o algoritmo se baseia na distribuição proposta, ou seja, gera o candidato denominado ϕ a partir da distribuição $q(\phi|\vartheta^t)$. Esse candidato pode ser aceito com probabilidade $\alpha(\vartheta^t, \phi)$ ou rejeitado com probabilidade $1 - \alpha(\vartheta^t, \phi)$, sendo

$$\alpha(\vartheta^t, \phi) = \min\left\{1, \frac{\pi(\phi|\boldsymbol{y})q(\vartheta^t|\phi)}{\pi(\vartheta^t|\boldsymbol{y})q(\phi|\vartheta^t)}\right\}.$$

Se o candidato ϕ for aceito, então, ele será o valor no estado $(t+1)$ da cadeia, caso contrário, $\vartheta^{t+1} = \vartheta^t$.

Assim, para implementação do algoritmo Metropois-Hasting são necessários os seguintes passos:

(i) inicie o contador em $t = 0$ iterações;

(ii) especifique um valor inicial para ϑ^0;

(iii) a partir da distribuição proposta $q(\phi|\vartheta^t)$, gere um valor ϕ que é o candidato ao próximo estado da cadeia;

(iv) calcule a probabilidade de aceitação $\alpha(\vartheta^t, \phi)$;

(v) gere u a partir de uma distribuição $U(0, 1)$;

(vi) se $u < \alpha$, aceite o valor candidato, ou seja, $\vartheta^{t+1} = \phi$, caso contrário, $\vartheta^{t+1} = \vartheta^t$;

(vii) incremente o contador de t para $t+1$ e volte ao passo (iii).

2.8.3 Abordagem Bayesiana para ajuste do modelo

A abordagem Bayesiana se mostra bastante interessante para o ajuste de modelos lineares generalizados mistos, devido à sua formulação hierárquica, pois a distribuição de Y está condicionada ao valor do efeito aleatório $\boldsymbol{b}_i$. Assim, considera-se o modelo introduzido na seção 2.5.1 e assume-se que os parâmetros $\boldsymbol{\beta}$, ϕ e $\boldsymbol{D}$ são independentes com distribuições a priori dadas, respectivamente, por $f(\boldsymbol{\beta})$, $f(\phi)$ e $f(\boldsymbol{D})$. Admite-se para $\boldsymbol{\beta}$ a distribuição normal ou uma priori não informativa. Para ϕ e $\boldsymbol{D}$ são assumidas como priori a distribuição não informativa de Jeffreys (GELMAN et al., 1995). Entretanto, Fahrmeir e Tutz (1994, apud MOLENBERGHS; VERBEKE, 2005) reportam que funções de densidades a priori impróprias como a de Jeffreys, podem levar à obtenção de funções densidades a posteriori impróprias.

Tendo especificado as funções densidades a priori, a função densidade de pro-

babilidade a posteriori conjunta é

$$f(\boldsymbol{\beta}, \boldsymbol{D}, \phi, \boldsymbol{b}_1, \ldots, \boldsymbol{b}_N) \propto \prod_{i=1}^{N} \prod_{j=1}^{n_i} f_i(y_{ij}|\boldsymbol{\beta}, \phi, \boldsymbol{b}_i) \prod_{i=1}^{N} f(\boldsymbol{b}_i|\boldsymbol{D}) f(\boldsymbol{D}) f(\boldsymbol{\beta}) f(\phi).$$

2.8.4 Predição de $\boldsymbol{b}_i$ via *empirical Bayes*

Na prática, o interesse principal é obter estimativas para os parâmetros da distribuição marginal de $\boldsymbol{Y}_i$, ou seja, $\boldsymbol{\beta}$, ϕ e $\boldsymbol{D}$. Mas há, também, o interesse nas predições dos efeitos aleatórios $\boldsymbol{b}_i$, que mostram a variabilidade entre indivíduos, pois são necessários tanto para predição como para análise de diagnóstico. Para predição dos efeitos aleatórios $\boldsymbol{b}_i$ tem-se a sua distribuição a posteriori dada por:

$$f_i(\boldsymbol{b}_i|\boldsymbol{y}_i, \boldsymbol{\beta}, \boldsymbol{D}, \phi) = \frac{f_i(\boldsymbol{y}_i|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi) f(\boldsymbol{b}_i|\boldsymbol{D})}{\int f_i(\boldsymbol{y}_i|\boldsymbol{b}_i, \boldsymbol{\beta}, \phi) f(\boldsymbol{b}_i|\boldsymbol{D}) d\boldsymbol{b}_i}.$$

Assim, a predição $\widehat{\boldsymbol{b}}_i$ via *empirical Bayes* será o valor de $\boldsymbol{b}_i$ que maximiza a função $f_i(\boldsymbol{b}_i|\boldsymbol{y}_i, \boldsymbol{\beta}, \boldsymbol{D}, \phi)$, quando os valores de $\boldsymbol{\beta}$, $\boldsymbol{D}$ e ϕ são substituídos pelas suas respectivas estimativas de máxima verossimilhança.

2.8.5 Critérios para comparação de modelos

A fim de realizar uma comparação de modelos, podem-se utilizar critérios baseados em medidas de ajuste dos diferentes modelos aos dados que respeitem o princípio da parcimônia.

2.8.5.1 *Deviance Information Criterion*

Spiegelhalter et al. (2000) propuseram o critério DIC, *Deviance Information Criterion*, para a seleção de modelos. Essa medida pode ser vista como uma generalização do critério de informação de Akaike (AKAIKE, 1973). O DIC é um critério bayesiano baseado em métodos MCMC que pode ser usado para comparar modelos arbitrários, pois leva em consideração a adequabilidade do modelo e a sua complexidade, representados, respectivamente, pelas estatísticas $\overline{D}$ e p_D.

O cálculo do DIC é baseado no valor esperado da distribuição a posteriori de $D(\boldsymbol{\vartheta}) = -2\log[p(\boldsymbol{y}|\boldsymbol{\vartheta})] + 2\log[f(\boldsymbol{y})]$, que é a média a posteriori do desvio, e na estatística de penalização p_D associada à complexidade do modelo (representando o número efetivo de

parâmetros), que é a diferença entre a média a posteriori da *deviance* e a *deviance* das médias a posteriori dos parâmetros de interesse, como segue

$$\begin{aligned} p_D &= \mathrm{E}_{\boldsymbol{\vartheta}|\boldsymbol{y}}[D(\boldsymbol{\vartheta})] - D(\mathrm{E}_{\boldsymbol{\vartheta}|\boldsymbol{y}}[p(\boldsymbol{y}|\boldsymbol{\vartheta})]) \\ &= \overline{D(\boldsymbol{\vartheta})} - \overline{D}(\boldsymbol{\vartheta}). \end{aligned}$$

Baseado nessas estatísticas, o DIC é definido como

$$DIC = \overline{D(\boldsymbol{\vartheta})} + p_D,$$

sendo que o menor valor do DIC, indica um melhor ajuste do modelo.

2.8.5.2 Avaliação do modelo Bayesiano utilizando quantidades pivotais

A comparação de modelos bayesianos é, geralmente, baseada na análise do valor do DIC devido à facilidade de sua aplicação. Entretanto, Millar (2009) mostrou que o DIC é uma ferramenta perigosa para modelos bayesianos hierárquicos aplicados a dados de contagem com superdispersão. A mais conhecida alternativa ao DIC é o fator de Bayes, que é a razão entre densidades marginais dos dados sob dois modelos competitivos. Porém, o fator de Bayes oferece alguns inconvenientes: para se obter a densidade marginal necessita-se de integrações no espaço paramétrico que, geralmente, não podem ser resolvidas numericamente e o resultado obtido pode ser influenciado pela especificação das prioris atribuídas aos parâmetros. Uma outra alternativa ao DIC como método de avaliação do modelo é o uso da distribuição preditiva a posteriori. Foi proposta, inicialmente, por Guttman (1967) e Rubin (1984), e, posteriormente, estendida para uma função de discrepância por Gelman, Meng e Stern (1996). Apesar da facilidade de implementação computacional, Bayarri e Berger (2000), Robins, van der Vaart, e Ventura (2000) observaram que não é possível assumir que a distribuição preditiva a posteriori seja uniformemente distribuída, ainda que assintoticamente.

Como alternativa, Johnson (2007) propôs explorar a relação entre a distribuição das quantidades pivotais avaliadas no parâmetro verdadeiro e a distribuição das mesmas quantidades pivotais avaliadas nos valores das estimativas a posteriori dos parâmetros. Comumente, essas distribuições são idênticas possibilitando a definição de um grande número de medidas para avaliação de modelos Bayesianos.

Seja $\boldsymbol{Y}$ um vetor aleatório definido no espaço amostral $\mathcal{Y} \subset \mathbb{R}^n$ cuja função de densidade de probabilidade pertence à família paramétrica $\{f(\boldsymbol{y}|\boldsymbol{\vartheta}), \boldsymbol{\vartheta} \in \Theta \subset \mathbb{R}^s\}$. Johnson (2007) define uma *quantidade pivotal* como uma função $S(\mathcal{Y}, \Theta) \mapsto \mathbb{R}^t$ para a qual a integral

$$G(\boldsymbol{s}) \equiv \int_{\mathcal{Y}} \mathcal{I}_{A(\boldsymbol{\omega},\boldsymbol{y})} f(\boldsymbol{y}|\boldsymbol{\omega}) d\boldsymbol{y}, \quad A(\boldsymbol{\omega}, \boldsymbol{y}) = \{(\boldsymbol{y}, \boldsymbol{\omega}) : S(\boldsymbol{y}, \boldsymbol{\omega}) \leq \boldsymbol{s}\}$$

depende apenas do valor do vetor $\boldsymbol{s} \in \mathbb{R}^t$ para todo $\boldsymbol{\omega} \in \Theta$, ou seja, a distribuição amostrada de S é invariante quando o valor do seu segundo argumento $\boldsymbol{\vartheta}$ determina a densidade amostral do primeiro $\boldsymbol{Y}$.

Lema: Seja $S(\boldsymbol{Y}, \boldsymbol{\vartheta})$ uma quantidade pivotal e suponha que $\boldsymbol{\vartheta}_0$ é um vetor aleatório obtido a partir da densidade π. Dado $\boldsymbol{\vartheta}_0$, seja $\boldsymbol{Y}$ um vetor aleatório amostrado da densidade $f(\boldsymbol{y}|\boldsymbol{\vartheta}_0)$, e seja $\boldsymbol{\vartheta}_{\boldsymbol{Y}}$ um vetor de parâmetros obtido a partir da distribuição a posteriori de $\boldsymbol{\vartheta}$ dado $\boldsymbol{Y}$. Então, $S(\boldsymbol{Y}, \boldsymbol{\vartheta}_{\boldsymbol{Y}})$ e $S(\boldsymbol{Y}, \boldsymbol{\vartheta}_0)$ são identicamente distribuídas.

Com essa propriedade das quantidades pivotais, é possível definir medidas de diagnóstico para grande parte dos modelos bayesianos, principalmente, para modelos contínuos, mas podem ser definidas extensões para os casos discretos. Entretanto, existe uma limitação importante: a propriedade não é válida para a distribuição conjunta das quantidades pivotais baseadas em múltiplos valores de $(\boldsymbol{Y}, \boldsymbol{\omega}^i_{\boldsymbol{Y}})$, a partir da mesma distribuição a posteriori, em que $\{\boldsymbol{\omega}^i_{\boldsymbol{Y}}\}$ denota a amostra a posteriori de $\boldsymbol{\omega}$ baseada no mesmo vetor de dados $\boldsymbol{Y}$.

Uma aproximação sugerida por Johnson (2007) é basear a avaliação do modelo em comparações gráficas informais, ou seja, comparar a distribuição a posteriori das quantidades pivotais com suas distribuições amostrais. Johnson também sugere a utilização da medida de discrepância $X^2(\boldsymbol{y}; \boldsymbol{\vartheta})$ introduzida por Gelman, Meng e Stern (1996) em cada estágio do modelo hierárquico, em que

$$X^2(\boldsymbol{y}; \boldsymbol{\vartheta}) = \sum_{i=1}^{n} \frac{\big(y_i - \mathrm{E}(y_i|\boldsymbol{\vartheta})\big)^2}{\mathrm{Var}(y_i|\boldsymbol{\vartheta})} \sim \chi^2_n. \tag{10}$$

2.8.6 Qualidade de ajuste do modelo aos dados

Em uma análise estatística, verificar a qualidade de ajuste do modelo proposto aos dados é de extrema importância. Uma boa análise Bayesiana deve incluir, pelo menos, alguma forma de verificar se o modelo ajustado é adequado aos dados e se este modelo é

aceitável para os fins aos quais foi proposto (GELMAN et al., 1995). Uma das formas de se verificar o ajuste de um modelo proposto aos dados é por meio da distribuição preditiva a posteriori.

A verificação do ajuste do modelo aos dados por meio da distribuição preditiva a posteriori consiste em replicar o conjunto de dados a partir dos valores obtidos para os parâmetros do modelo, usando-se as distribuições a posteriori e, em seguida, comparar os conjuntos de dados replicados com os dados observados, que devem ser semelhantes. A discrepância entre o dados preditos pelo modelo e os dados observados será medida por testes a serem definidos.

Sejam $\boldsymbol{y}$ os dados observados, $\boldsymbol{\vartheta}$ o vetor de parâmetros do modelo proposto e $\boldsymbol{y}^{rep}$ os dados replicados. A distribuição de $\boldsymbol{y}^{rep}$, dado o atual estado de conhecimento é obtida por meio da distribuição preditiva a posteriori dada por:

$$p(\boldsymbol{y}^{rep}|\boldsymbol{y}) = \int p(\boldsymbol{y}^{rep}|\boldsymbol{\vartheta})p(\boldsymbol{\vartheta}|\boldsymbol{y})d\boldsymbol{\vartheta},$$

que, segundo Gilks, Richardson e Spiegelhalter (1998), pode ser estimada como

$$\widehat{p}(\boldsymbol{y}^{rep}|\boldsymbol{y}) = \frac{1}{m}\sum_{j=1}^{m} p(\boldsymbol{y}^{rep}|\boldsymbol{\vartheta}^{*})$$

em que $\boldsymbol{\vartheta}_j^*$, $j = 1, \ldots, m$, é um conjunto de valores que serão tratados como amostras de $p(\boldsymbol{\vartheta}|\boldsymbol{y})$.

2.9 Modelos para dados na forma de contagens longitudinais com superdispersão

Quando se consideram dados na forma de contagens que apresentam diversas medidas realizadas em um mesmo indivíduo ao longo do tempo, utiliza-se a teoria de modelos lineares generalizados mistos (MLGM) (BRESLOW; CLAYTON, 1993), em que a estrutura de dependência temporal é acomodada adicionando-se efeitos aleatórios no preditor linear.

Geralmente, a superdispersão e a estrutura longitudinal são modeladas separadamente. Mas existem dados de contagens, tanto na área médica como agronômica, para as quais os modelos lineares generalizados mistos não são suficientes. Assim, surge o interesse em estudar um modelo que unifique as teorias para análise de medidas longitudinais que apresentem o problema de superdispersão. Booth et al. (2003) propõem o modelo Poisson log-linear com efeitos aleatórios enquanto que Molenberghs, Verbeke e Demétrio (2007) e

Vangeneugden et al. (2009) usando a mesma estrutura, fornecem as expressões do vetor de médias, da matriz de variâncias-covariâncias e da função de probabilidade marginal conjunta para o modelo que leva em conta a superdispersão e a correlação de medidas longitudinais.

Seja Y_{ij} uma variável aleatória, representando a j-ésima medida longitudinal do i-ésimo indivíduo, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$ e $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{in_i})^T$ o vetor de todas as medidas do indivíduo i e $\boldsymbol{x}_{ij}$ um vetor de covariáveis associado a Y_{ij}.

A seguir, serão particularizados quatro modelos que podem ser considerados para análise de dados na forma de contagens longitudinais.

2.9.1 Modelo 1: Poisson

Supondo, inicialmente, que os Y_{ij} são independentes, com distribuição de Poisson com parâmetro λ_{ij}, isto é, $Y_{ij} \sim \text{Poisson}(\lambda_{ij})$, $\lambda_{ij} > 0$. Então,

$$P(Y_{ij} = y_{ij}) = \frac{e^{-\lambda_{ij}} \lambda_{ij}^{y_{ij}}}{y_{ij}!}, \quad \text{para} \quad y_{ij} = 0, 1, 2, \ldots \tag{11}$$

e admitindo-se que

$$\ln(\lambda_{ij}) = \boldsymbol{x}'_{ij}\boldsymbol{\beta} \tag{12}$$

tem-se que

$$\text{E}(Y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}) = \lambda_{ij} = \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}) \quad \text{e} \quad \text{Var}(Y_{ij}|\boldsymbol{x}_{ij}, \boldsymbol{\beta}) = \lambda_{ij} = \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}).$$

2.9.2 Modelo 2: Poisson-gama

Neste caso assume-se, inicialmente, a distribuição de Poisson dada por (11) para Y_{ij} e para acomodar a superdispersão considera-se um efeito aleatório θ_{ij} com distribuição gama, isto é,

$$\begin{aligned} Y_{ij}|\lambda_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\ \lambda_{ij} &= \theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}) \\ \theta_{ij} &\sim \text{gama}(\alpha_{1j}, \alpha_{2j}) \end{aligned} \tag{13}$$

sendo α_{1j} e α_{2j} parâmetros da distribuição gama.

Como $\mathrm{E}(\theta_{ij}) = \alpha_{1j}\alpha_{2j}$ e $\mathrm{Var}(\theta_{ij}) = \alpha_{1j}\alpha_{2j}^2$, então,

$$\begin{aligned} \mathrm{E}(Y_{ij}) &= \mathrm{E}[\mathrm{E}(Y_{ij}|\lambda_{ij})] = \mathrm{E}[\theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})] \\ &= \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\mathrm{E}[\theta_{ij}] = \alpha_{1j}\alpha_{2j}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}) \end{aligned} \tag{14}$$

e

$$\begin{aligned} \mathrm{Var}(Y_{ij}) &= \mathrm{E}[\mathrm{Var}(Y_{ij}|\lambda_{ij})] + \mathrm{Var}[\mathrm{E}(Y_{ij}|\lambda_{ij})] = \mathrm{E}[\theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})] + \mathrm{Var}[\theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})] \\ &= \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\mathrm{E}(\theta_{ij}) + [\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})]^2\mathrm{Var}(\theta_{ij}) = \alpha_{1j}\alpha_{2j}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}) + \alpha_{1j}\alpha_{2j}^2\left[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\right]^2 \\ &= \alpha_{1j}\alpha_{2j}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\Big[1 + \alpha_{2j}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\Big]. \end{aligned} \tag{15}$$

De (14) e (15) observa-se que a variância de Y_{ij} é maior do que a média. Assim, a presença da superdispersão no modelo assumido é dada pela adição do termo $\alpha_{1j}\alpha_{2j}^2\left[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\right]^2$. É importante ressaltar que quando se tem a presença do intercepto no preditor linear os parâmetros α_{1j} e α_{2j} da distribuição gama não são simultaneamente identificáveis. Assim, impõe-se a restrição $\alpha_{1j} \cdot \alpha_{2j} = 1$, ou seja, $\alpha_{1j} = \alpha_{2j}^{-1}$. Como pode ser visto no Anexo A, integrando-se a distribuição conjunta em relação ao efeito aleatório θ_{ij} com distribuição gama, este modelo é uma generalização do modelo binomial negativo.

2.9.3 Modelo 3: Poisson-normal

Considerando-se a distribuição de Poisson com função densidade de probabilidade dada por (11) para Y_{ij} e a distribuição normal para o efeito aleatório $\boldsymbol{b}_i$ para incorporar a correlação entre as medidas longitudinais, tem-se

$$\begin{aligned} Y_{ij}|\lambda_{ij} &\sim \mathrm{Poisson}(\lambda_{ij}) \\ \ln(\lambda_{ij}) &= \boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, \boldsymbol{D}) \end{aligned} \tag{16}$$

sendo que $\boldsymbol{D}$ é a matriz de variâncias-covariâncias de $\boldsymbol{b}_i$.

Logo, tem-se:

$$\begin{aligned} \mathrm{E}(Y_{ij}) &= \mathrm{E}[\mathrm{E}(Y_{ij}|\lambda_{ij})] = \mathrm{E}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)] \\ &= \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})\mathrm{E}[\exp(\boldsymbol{z}'_{ij}\boldsymbol{b}_i)] = \exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right) \end{aligned} \tag{17}$$

e

$$
\begin{aligned}
\mathrm{Var}(Y_{ij}) &= \mathrm{E}[\mathrm{Var}(Y_{ij}|\lambda_{ij})] + \mathrm{Var}[\mathrm{E}(Y_{ij}|\lambda_{ij})] \\
&= \mathrm{E}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)] + \mathrm{Var}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)] \\
&= \exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right) + \mathrm{E}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)^2] - \mathrm{E}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)]^2 \\
&= \exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right) + \exp(2\boldsymbol{x}'_{ij}\boldsymbol{\beta} + 2\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}) - \left\{\exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right)\right\}^2 \\
&= \exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right) + \exp(2\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij})\{\exp(\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}) - 1\}. \qquad (18)
\end{aligned}
$$

Como em Molenberghs, Verbeke e Demétrio (2007), o desenvolvimento para a distribuição de probabilidade de $\boldsymbol{Y}_i$ é mostrado no Anexo A.

$$
\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i = \boldsymbol{y}_i) &= \frac{1}{\prod_{j=1}^{n_i} y_{ij}!} \sum_{\boldsymbol{t}} \frac{(-1)^{\sum_{j=1}^{n_i} t_j}}{\prod_{j=1}^{n_i} t_j!} \exp\left[\sum_{j=1}^{n_i}(y_{ij} + t_j)\boldsymbol{x}'_{ij}\boldsymbol{\beta}\right] \times \\
&\times \exp\left\{\frac{1}{2}\left[\sum_{j=1}^{n_i}(y_{ij} + t_j)\boldsymbol{z}'_{ij}\right]\boldsymbol{D}\left[\sum_{j=1}^{n_i}(y_{ij} + t_j)\boldsymbol{z}_{ij}\right]\right\}, \qquad (19)
\end{aligned}
$$

em que $\boldsymbol{t} = (t_1, \ldots, t_{n_i})$.

2.9.4 Modelo 4: Poisson-normal-gama (*Combinado*)

Assumindo-se que variável aleatória Y_{ij} segue uma distribuição de Poisson com função densidade de probabilidade dada por (11), inicialmente, cada Y_{ij} será considerado independente e diferentes tipos de efeitos aleatórios serão incluídos para levar em conta a correlação entre medidas feitas no mesmo indivíduo e a superdispersão, conforme Molenberghs et al. (2010).

Supondo que

$$
\lambda_{ij} = \theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i) \qquad (20)
$$

em que $\boldsymbol{b}_i$ é o efeito aleatório adicionado ao preditor linear com a finalidade de incorporar a correlação entre as medidas repetidas realizadas em um mesmo indivíduo e θ_{ij} é o efeito aleatório que ao ser incluído no modelo acomodará a superdispersão.

Além disso, supõe-se que

$$
\boldsymbol{b}_i \sim N(\boldsymbol{0}, \boldsymbol{D})
$$

e que θ_{ij} tem distribuição com

$$\mathrm{E}(\boldsymbol{\theta}_i) = \mathrm{E}[(\theta_{i1}, \ldots, \theta_{in_i})'] = \boldsymbol{\Phi}_i \quad \text{e} \quad \mathrm{Var}(\boldsymbol{\theta}_i) = \boldsymbol{\Sigma}_i.$$

Seja $\boldsymbol{Y}_i$ o vetor de tamanho n_i com as medidas do i-ésimo indivíduo e $\boldsymbol{\lambda}_i = (\lambda_{i1}, \ldots, \lambda_{in_i})'$ com $\mathrm{E}(\boldsymbol{\lambda}_i) = \boldsymbol{\mu}_i$ e $\mathrm{Var}(\boldsymbol{\lambda}_i) = \boldsymbol{\Sigma}_i$. Portanto, seguindo Molenberghs, Verbeke e Demétrio (2007), o vetor de médias $\boldsymbol{\mu}_i = \mathrm{E}(\boldsymbol{Y}_i)$ tem como componentes:

$$\mu_{ij} = \phi_{ij}\exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right)$$

e a matriz de variâncias-covariâncias de $\boldsymbol{Y}_i$ é dada por

$$\mathrm{Var}(\boldsymbol{Y}_i) = M_i + M_i(P_i - J_{n_i}M_i),$$

em que M_i é uma matriz diagonal que tem como elementos o vetor $\boldsymbol{\mu}_i$, J_{n_i} é uma matriz de dimensão $n_i \times n_i$ composta apenas por 1's e P_i é uma matriz que tem como elementos

$$p_{i,jk} = \exp\left(\frac{1}{2}\boldsymbol{z}'_{ij}D\boldsymbol{z}_{ij}\right) \cdot \frac{\sigma_{i,jk} + \phi_{ij}\phi_{ik}}{\phi_j\phi_{ik}} \cdot \exp\left(\frac{1}{2}\boldsymbol{z}'_{ij}D\boldsymbol{z}_{ij}\right).$$

Como caso particular, assumindo a distribuição gama para o efeito aleatório θ_{ij} para acomodar a superdispersão, tem-se

$$\begin{aligned} Y_{ij}|\boldsymbol{b}_i, \theta_{ij} &\sim \mathrm{Poisson}(\lambda_{ij}) \\ \lambda_{ij} &= \theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i) \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, \boldsymbol{D}) \\ \theta_{ij} &\sim \mathrm{gama}(\alpha_{1j}, \alpha_{2j}) \end{aligned} \tag{21}$$

sendo que $\boldsymbol{D}$ é a matriz de variâncias-covariâncias referente a $\boldsymbol{b}_i$ e α_{1j} e α_{2j} são parâmetros da distribuição gama.

Como $\mathrm{E}(\theta_{ij}) = \alpha_{1j}\alpha_{2j}$ e $\mathrm{Var}(\theta_{ij}) = \alpha_{1j}\alpha_{2j}^2$, então,

$$\begin{aligned} \mathrm{E}(Y_{ij}) &= \mathrm{E}\{\mathrm{E}[\mathrm{E}(Y_{ij}|\boldsymbol{b}_i, \theta_{ij})]\} = \mathrm{E}\{\mathrm{E}[\theta_{ij}\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)]\} \\ &= \mathrm{E}\{\theta_{ij}\mathrm{E}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)]\} = \mathrm{E}(\theta_{ij})\mathrm{E}[\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)] \\ &= \alpha_{1j}\alpha_{2j}\exp\left(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}\right) \end{aligned} \tag{22}$$

e, considerando $\kappa_{ij} = \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)$, tem-se

$$\begin{aligned}
\mathrm{Var}(Y_{ij}) &= \mathrm{Var}\{\mathrm{E}[\mathrm{E}(Y_{ij}|\boldsymbol{b}_i,\theta_{ij})]\} + \mathrm{E}\{\mathrm{Var}[\mathrm{E}(Y_{ij}|\boldsymbol{b}_i,\theta_{ij})]\} + \mathrm{E}\{\mathrm{E}[\mathrm{Var}(Y_{ij}|\boldsymbol{b}_i,\theta_{ij})]\} \\
&= \mathrm{Var}\{\mathrm{E}[\theta_{ij}\kappa_{ij}]\} + \mathrm{E}\{\mathrm{Var}[\theta_{ij}\kappa_{ij}]\} + \mathrm{E}\{\mathrm{E}[\theta_{ij}\kappa_{ij}]\} \\
&= \mathrm{Var}\{\mathrm{E}(\theta_{ij})\kappa_{ij}\} + \mathrm{E}\{\mathrm{Var}(\theta_{ij})\kappa_{ij}^2\} + \mathrm{E}\{\mathrm{E}(\theta_{ij})\kappa_{ij}\} \\
&= \mathrm{E}(\theta_{ij})^2\mathrm{Var}(\kappa_{ij}) + \mathrm{Var}(\theta_{ij})^2\mathrm{E}(\kappa_{ij}^2) + \mathrm{E}(\theta_{ij})\mathrm{E}(\kappa_{ij}) \\
&= \alpha_{1j}^2\alpha_{2j}^2 e^{2\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}}\left(e^{\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}} - 1\right) + \alpha_{1j}\alpha_{2j}^2 e^{2\boldsymbol{x}'_{ij}\boldsymbol{\beta}+2\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}} + \\
&+ \alpha_{1j}\alpha_{2j} e^{\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\frac{1}{2}\boldsymbol{z}'_{ij}\boldsymbol{D}\boldsymbol{z}_{ij}}. \qquad (23)
\end{aligned}$$

Para esse modelo que combina os efeitos aleatórios normal e gama, também é possível obter a distribuição de probabilidade de $\boldsymbol{Y}_i$, que apresenta a seguinte forma:

$$\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i = \boldsymbol{y}_i) &= \sum_{\boldsymbol{t}}\left[\prod_{j=1}^{n_i}\binom{y_{ij}+t_j}{y_{ij}}\binom{\alpha_{1j}+y_{ij}+t_j-1}{\alpha_{1j}-1}(-1)^{t_j}\alpha_{2j}^{y_{ij}+t_j}\right] \times \\
&\times \exp\left(\sum_{j=1}^{n_i}(y_{ij}+t_j)\boldsymbol{x}'_{ij}\boldsymbol{\beta}\right) \times \\
&\times \exp\left\{\frac{1}{2}\left[\sum_{j=1}^{n_i}(y_{ij}+t_j)\boldsymbol{z}'_{ij}\right]\boldsymbol{D}\left[\sum_{j=1}^{n_i}(y_{ij}+t_j)\boldsymbol{z}_{ij}\right]\right\}, \qquad (24)
\end{aligned}$$

em que $\boldsymbol{t} = (t_1, \ldots, t_{n_i})$, como mostrado no Anexo A.

2.10 Modelos para dados binários longitudinais com superdispersão

Assim como quando se consideram dados na forma de contagens, quando dados binários estão sujeitos à superdispersão e são medidos longitudinalmente, necessita-se estender a teoria de modelos lineares generalizados mistos, que modela apenas a dependência temporal. Nesse caso, o interesse está em combinar efeitos aleatórios em um único modelo que permitam acomodar a superdispersão e a estrutura longitudinal, como em Molenberghs et al. (2010).

Seja Y_{ij} uma variável aleatória, representando a j-ésima medida longitudinal do i-ésimo indivíduo, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$ e $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{in_i})^T$ o vetor de todas as medidas do indivíduo i e $\boldsymbol{x}_{ij}$ um vetor de covariáveis associado a Y_{ij}.

A seguir, serão particularizados os modelos que podem ser usados na análise de dados binários longitudinais, sendo quatro ao se considerar a função de ligação logística e outros quatro ao se considerar a função de ligação probit.

2.10.1 Modelo 1: Bernoulli

Supondo, inicialmente, que os Y_{ij} são independentes, com distribuição Bernoulli com parâmetro π_{ij}, isto é, $Y_{ij} \sim \text{Bernoulli}(\pi_{ij})$, $0 \leq \pi_{ij} \leq 0$. Então,

$$P(Y_{ij} = y_{ij}) = \pi_{ij}^{y_{ij}}(1 - \pi_{ij})^{1-y_{ij}}, \quad \text{para} \quad y_{ij} = 0, 1 \tag{25}$$

em que, ao se considerar a função de ligação logística, tem-se

$$\pi_{ij} = \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})}, \tag{26}$$

e, ao se considerar a função de ligação probit, tem-se

$$\pi_{ij} = \Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta}), \tag{27}$$

em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$.

2.10.2 Modelo 2: Bernoulli-beta

Neste caso assume-se, inicialmente, a distribuição Bernoulli dada por (25) para Y_{ij} e para acomodar a superdispersão considera-se um efeito aleatório θ_{ij} com distribuição beta, isto é, para função de ligação logística

$$\begin{aligned} Y_{ij}|\pi_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \theta_{ij}\frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})}, \\ \theta_{ij} &\sim \text{beta}(\alpha_1, \alpha_2) \end{aligned} \tag{28}$$

e, para a função de ligação probit

$$\begin{aligned} Y_{ij}|\pi_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \theta_{ij}\Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta}) \\ \theta_{ij} &\sim \text{beta}(\alpha_1, \alpha_2) \end{aligned} \tag{29}$$

sendo α_1 e α_2 parâmetros da distribuição beta e $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$.

Para o modelo Bernoulli-beta a função de probabilidade condicional para $Y_{ij}|\theta_{ij}$ é dada por

$$f(y_{ij}|\theta_{ij}) = \big[\theta_{ij}\kappa_{ij}\big]^{y_{ij}}\big[1 - \theta_{ij}\kappa_{ij}\big]^{1-y_{ij}}, \quad y_{ij} \in \mathbb{R},$$

em que $\kappa_{ij} = \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})}{1+\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta})}$ para a função de ligação logística e $\kappa_{ij} = \Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta})$ para a função de ligação probit.

Assim, função de verossimilhança marginal é

$$L(\boldsymbol{\beta}, \alpha_1, \alpha_2) = \prod_{i=1}^{N}\prod_{j=1}^{n_i} \frac{1}{\alpha_1 + \alpha_2}(\alpha_1\kappa_{ij})^{y_{ij}}[(1-\kappa_{ij})\alpha_1 + \alpha_2]^{1-y_{ij}},$$

como mostrado no Anexo B. Um comentário importante é sobre a identificabilidade dos parâmetros. Os parâmetros α_1 e α_2 não são simultaneamente identificáveis, quando o preditor linear contém um intercepto. Uma solução é estabelecer uma restrição do tipo $\alpha_2/\alpha_1 = c$ ou $\alpha_1 + \alpha_2 = c$, que soluciona a identificabilidade (VIEIRA, 2008).

2.10.3 Modelo 3: Bernoulli-normal

Considerando-se a distribuição Bernoulli com função densidade de probabilidade dada por (25) para Y_{ij} e a distribuição normal para o efeito aleatório $\boldsymbol{b}_i$ para incorporar a correlação entre as medidas longitudinais, ao se considerar a função de ligação logística

$$\begin{aligned} Y_{ij}|\pi_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)}{1+\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)}, \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, \boldsymbol{D}) \end{aligned} \tag{30}$$

e, ao se considerar a função de ligação probit

$$\begin{aligned} Y_{ij}|\pi_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i) \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, \boldsymbol{D}) \end{aligned} \tag{31}$$

sendo que $\boldsymbol{D}$ é a matriz de variâncias-covariâncias referente a $\boldsymbol{b}_i$ e $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\left[-t^2/2\right]dt$.

2.10.4 Modelo 4: Bernoulli-normal-beta (*Combinado*)

Assumindo-se que variável aleatória Y_{ij} segue uma distribuição Bernoulli com função densidade de probabilidade dada por (25). Devido ao fato de se trabalhar com medidas na forma de contagens longitudinais, inicialmente, cada Y_{ij} será considerado independente e

diferentes tipos de efeitos aleatórios serão incluídos para levar em conta a correlação entre medidas feitas no mesmo indivíduo e a superdispersão, conforme Vieira (2008).

Seja $\boldsymbol{b}_i$ o efeito aleatório adicionado ao preditor linear com a finalidade de incorporar a correlação entre as medidas repetidas realizadas em um mesmo indivíduo e θ_{ij} o efeito aleatório que ao ser incluído no modelo acomodará a superdispersão. Assumindo uma função de ligação logística, tem-se

$$\pi_{ij} = \theta_{ij} \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)}{1 + \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)}, \tag{32}$$

e assumindo uma função de ligação probit, tem-se

$$\pi_{ij} = \theta_{ij}\Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i), \tag{33}$$

em que $\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\left[(1/2)u^2\right] du^4$.

Além disso, supõe-se que

$$\boldsymbol{b}_i \sim N(\boldsymbol{0}, \boldsymbol{D})$$

e que

$$\theta_{ij} \sim \text{Beta}(\alpha_1, \alpha_2),$$

sendo $\boldsymbol{D}$ a matriz de variâncias-covariâncias referente a $\boldsymbol{b}_i$ e α_1 e α_2 os parâmetros da distribuição beta, ou seja, ao se considerar a função de ligação logística tem-se

$$\begin{aligned} Y_{ij}|\boldsymbol{b}_i, \theta_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \theta_{ij} \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)}{1 + \exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)} \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, \boldsymbol{D}) \\ \theta_{ij} &\sim \text{beta}(\alpha_1, \alpha_2) \end{aligned} \tag{34}$$

e, ao se considerar a função de ligação probit tem-se

$$\begin{aligned} Y_{ij}|\boldsymbol{b}_i, \theta_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \theta_{ij}\Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i) \\ \boldsymbol{b}_i &\sim N(\boldsymbol{0}, \boldsymbol{D}) \\ \theta_{ij} &\sim \text{beta}(\alpha_1, \alpha_2) \end{aligned} \tag{35}$$

sendo que $\boldsymbol{D}$ é a matriz de variâncias-covariâncias referente a $\boldsymbol{b}_i$, α_1 e α_2 os parâmetros da distribuição beta e $\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\big[-t^2/2\big] dt$.

Para o modelo Bernoulli-normal-beta a função de probabilidade condicional para $Y_{ij}|\theta_{ij}, \boldsymbol{b}_i$ é dada por

$$f(y_{ij}|\theta_{ij}, \boldsymbol{b}_i) = \big[\theta_{ij}\kappa_{ij}\big]^{y_{ij}} \big[1 - \theta_{ij}\kappa_{ij}\big]^{1-y_{ij}}, \quad y_{ij} \in \mathbb{R},$$

em que $\kappa_{ij} = \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\boldsymbol{z}'_{ij}\boldsymbol{b}_i)}{1+\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\boldsymbol{z}'_{ij}\boldsymbol{b}_i)}$ para a função de ligação logística e $\kappa_{ij} = \Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)$ para a função de ligação probit.

A função de verossimilhança marginal condicionada apenas a $\boldsymbol{b}_i$ é

$$L(\boldsymbol{\beta}, \boldsymbol{D}, \alpha_1, \alpha_2) = \prod_{i=1}^{N} \int \prod_{j=1}^{n_i} \frac{1}{\alpha_1 + \alpha_2} (\alpha_1\kappa_{ij})^{y_{ij}} [(1 - \kappa_{ij})\alpha_1 + \alpha_2]^{1-y_{ij}} f(\boldsymbol{b}_i|\boldsymbol{D}) d\boldsymbol{b}_i,$$

como mostrado no Anexo B, o que necessita de algoritmos para solução numérica da integral. Para o modelo Bernoulli-normal-beta com função de ligação logística, assim como no modelo Bernoulli-beta com função de ligação logística, os parâmetros α_1 e α_2 não são simultaneamente identificáveis, quando o preditor linear contém um intercepto. Assim, impõe-se a restrição do tipo $\alpha_2/\alpha_1 = c$.

Para o modelo com função de ligação probit os parâmetros α_1 e α_2 também não são simultaneamente identificáveis, e para solucionar esse problema, impõe-se uma restrição do tipo $\alpha_1+\alpha_2 = c$. Entretanto, para esse modelo ainda é possível obter uma solução analítica para a integral

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{D}, \alpha_1, \alpha_2) &= \prod_{i=1}^{N} \int \prod_{j=1}^{n_i} \frac{\big[\Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)\alpha_1\big]^{y_{ij}}}{\alpha_1 + \alpha_2} \{[1 - \Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)]\alpha_1 + \alpha_2\}^{1-y_{ij}} \times \\ &\times \ (2\pi)^{-q/2} |\boldsymbol{D}|^{-1/2} \exp\Big\{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}\boldsymbol{b}_i\Big\} d\boldsymbol{b}_i, \end{aligned}$$

como mostrado no Anexo B.

2.11 Método da clonagem de dados

A idéia do método baseia-se no que se segue. Considere um modelo Bayesiano completo para o problema em questão, especificando prioris adequadas para os parâmetros, mas ao invés de utilizar a função de verossimilhança para os dados observados, utilize a função de verossimilhança para k cópias (clones) dos dados, em que k é suficientemente grande e

considere as cópias independentes entre si. As distribuições a posteriori são calculadas pelo método MCMC usual. As médias a posteriori dos parâmetros são iguais às estimativas de máxima verossimilhança dos parâmetros, e k vezes a variância a posteriori é igual à variância assintótica das estimativas de máxima verossimilhança (LELE et. al., 2007).

Seja $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)^T$ uma variável aleatória proveniente do seguinte modelo estatístico hierárquico:

$$\boldsymbol{Y} \sim f(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\varphi})$$

$$\boldsymbol{X} \sim g(\boldsymbol{x}|\boldsymbol{\delta})$$

em que f e g são distribuições de probabilidades conjuntas, $\boldsymbol{X}$ é um vetor de quantidades aleatórias ou processos que afetam as observações, $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_q)^T$ é um vetor de parâmetros fixos desconhecidos que afeta as observações e $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_p)^T$ é um vetor de parâmetros fixos desconhecidos relacionados a $\boldsymbol{X}$.

A função de verossimilhança para o modelo descrito é dada por:

$$L(\boldsymbol{\delta}, \boldsymbol{\varphi}; \boldsymbol{y}) = \int f(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\varphi}) g(\boldsymbol{X}|\boldsymbol{\delta}) d\boldsymbol{X}$$

em que $\boldsymbol{y}$ é o vetor observações. As estimativas de máxima verossimilhança dos parâmetros, denotadas por $(\widehat{\boldsymbol{\varphi}}, \widehat{\boldsymbol{\delta}}) = (\widehat{\varphi_1}, \widehat{\varphi_2}, \ldots, \widehat{\varphi_q}, \widehat{\delta_1}, \widehat{\delta_2}, \ldots, \widehat{\delta_p})$ são os valores de $(\boldsymbol{\varphi}, \boldsymbol{\delta}) = (\varphi_1, \varphi_2, \ldots, \varphi_q, \delta_1, \delta_2, \ldots, \delta_p)$ que maximizam conjuntamente a função de verossimilhança. Como ja foi dito, o cálculo da função e das estimativas de máxima verossimilhança necessita de um grande esforço computacional e cálculo de integrais em dimensões grandes.

A abordagem Bayesiana contorna o problema da integração em dimensões grandes. Para tanto, considera-se que os parâmetros desconhecidos $\boldsymbol{\varphi}$ e $\boldsymbol{\delta}$, ao invés de serem fixos, são variáveis aleatórias, aos quais são atribuídas as distribuições de probabilidades denominadas distribuição a priori. Assim, de acordo com o teorema de Bayes, a partir da função de verossimilhança e das distribuições a priori dos parâmetros obtém-se a distribuição a posteriori conjunta, a partir da qual resultam as inferências dos parâmetros. Seja $\pi(\boldsymbol{\varphi}, \boldsymbol{\delta})$ a distribuição conjunta a priori para os parâmetros $\boldsymbol{\varphi}$ e $\boldsymbol{\delta}$. Então, a distribuição conjunta a posteriori é dada por

$$h(\boldsymbol{\varphi}, \boldsymbol{\delta}, \boldsymbol{X}|\boldsymbol{y}) = \frac{f(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\varphi}) g(\boldsymbol{X}|\boldsymbol{\delta}) \pi(\boldsymbol{\varphi}, \boldsymbol{\delta})}{\int f(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\varphi}) g(\boldsymbol{X}|\boldsymbol{\delta}) \pi(\boldsymbol{\varphi}, \boldsymbol{\delta}) d\boldsymbol{X} d\boldsymbol{\delta} d\boldsymbol{\varphi}}.$$

A distribuição marginal a posteriori para os parâmetros $\pi(\boldsymbol{\varphi}, \boldsymbol{\delta}|\boldsymbol{y})$ é obtida, simplesmente, pela integração de $h(\boldsymbol{\varphi}, \boldsymbol{\delta}, \boldsymbol{X}|\boldsymbol{y})$ em relação a $\boldsymbol{X}$. Para o algoritmo MCMC, utiliza-se apenas o fato que $h(\boldsymbol{\varphi}, \boldsymbol{\delta}, \boldsymbol{X}|\boldsymbol{y}) \propto f(\boldsymbol{y}|\boldsymbol{X}, \boldsymbol{\varphi})g(\boldsymbol{X}|\boldsymbol{\delta})\pi(\boldsymbol{\varphi}, \boldsymbol{\delta})$, não sendo necessária a solução das integrais.

Suponha, agora, que um experimento estatístico foi realizado para se obterem as observações $\boldsymbol{y}$, mas não apenas uma vez e sim k vezes simultanea e independentemente. Suponha, ainda, que as k réplicas do experimento produzem exatamente o mesmo resultado y. A função de verossimilhança para os k "clones" é $[L(\boldsymbol{\varphi}, \boldsymbol{\delta}; \boldsymbol{y})]^k$, ou seja, a verossimilhança original elevada à k-ésima potência e apresenta o mesmo ponto de máximo que a função de verossimilhança original, ou seja, $(\widehat{\boldsymbol{\varphi}}, \widehat{\boldsymbol{\delta}})$ como estimativas de máxima verossimilhança.

A distribuição conjunta a posteriori $h^{(k)}(\boldsymbol{\varphi}, \boldsymbol{\delta}, \boldsymbol{X}|\boldsymbol{y})$, assim como a distribuição marginal a posteriori $\pi^{(k)}(\boldsymbol{\varphi}, \boldsymbol{\delta}|\boldsymbol{y})$ podem ser obtidas por meio da distribuição a priori $\pi(\boldsymbol{\varphi}, \boldsymbol{\delta})$ e da função de verossimilhança para os dados "clonados". Para k suficientemente grande, a distribuição marginal a posteriori $\pi^{(k)}(\boldsymbol{\varphi}, \boldsymbol{\delta}|\boldsymbol{y})$ converge para uma distribuição normal multivariada com média igual à estimativa de máxima verossimilhança $(\widehat{\boldsymbol{\varphi}}, \widehat{\boldsymbol{\delta}})$ e variância igual a $\frac{1}{k}I^{-1}(\widehat{\boldsymbol{\varphi}}, \widehat{\boldsymbol{\delta}})$, em que $I(\widehat{\boldsymbol{\varphi}}, \widehat{\boldsymbol{\delta}})$ é a matriz de informação de Fisher correspondente à função de máxima verossimilhança original (LELE et al., 2010).

3 METODOLOGIA

3.1 Material

Nesta seção, serão introduzidos três conjuntos de dados para os quais foram desenvolvidos os métodos que serão discutidos a seguir.

3.1.1 Ensaio clínico para pacientes epilépticos

Um dos conjuntos de dados a ser analisado neste trabalho é proveniente de um experimento conduzido para comparar uma nova droga anti-epiléptica (AED) com placebo, ambos utilizados como tratamento adjunto em pacientes que sofrem ataques epilépticos e recebem uma ou duas outras AED's em doses terapêuticas. O principal interesse da pesquisa é avaliar se o novo tratamento adjunto reduz, ou não, o número de ataques epilépticos sofridos na última semana de acompanhamento, que é a variável resposta em questão.

A aleatorização dos pacientes aos tratamentos ocorreu após 12 semanas, período que serviu para a estabilização das AEDs. Depois desse período, 45 pacientes passaram a receber o placebo, enquanto que 44 pacientes, o novo tratamento. Os pacientes foram observados semanalmente, sendo que o acompanhamento foi de 16 semanas e, depois disso, alguns pacientes continuaram a ser observados. O período máximo de observação foi de 27 semanas. Esses dados podem ser vistos detalhadamente em Faught et al. (1996).

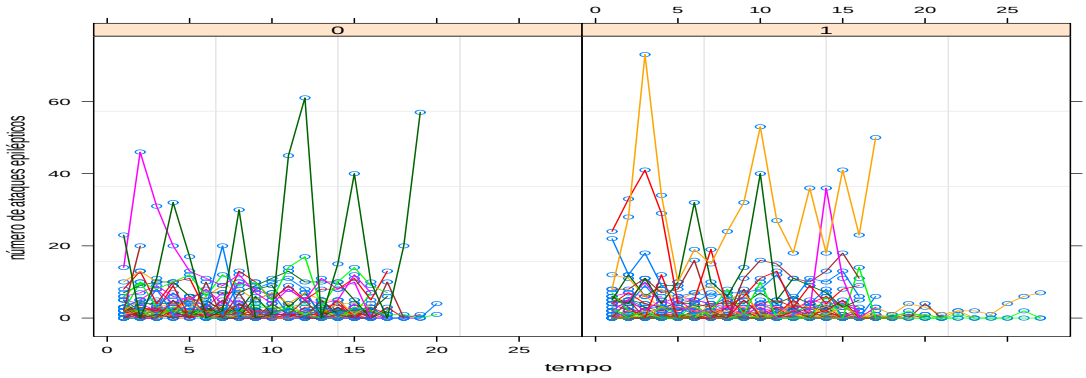


Figura 1 – Gráfico de perfis individuais para o número de ataques epilépticos sofridos em cada uma das 27 semanas pelos pacientes submetidos ao placebo (0) e à nova droga anti-epiléptica (1)

Na Figura 1, podem-se observar os perfis dos pacientes submetidos, respectivamente, ao placebo e ao novo tratamento. Verifica-se que os dados apresentam um comportamento muito instável com a presença de valores extremos. Além disso, há poucas observações disponíveis, principalmente, após a 20ª semana.

3.1.2 Ensaio clínico para tratamento de dermatite

Esses dados são provenientes de um estudo aleatorizado multi-centro, duplamente-cego e com grupos paralelos para comparar dois medicamentos administrados por via oral, aqui denominados por A e B, para tratamento da *toenail dermatophite onychomycosis* (TDO). Essa é uma infecção de unha bastante comum e de difícil tratamento. Os compostos antifúngicos tradicionais para tratamento dessa dermatite necessitam que o paciente administre o medicamento até que as unhas se desenvolvam por completo. Os novos medicamentos têm reduzido o tempo de tratamento para três meses.

Para o estudo descrito por De Backer et al. (1996), foram comparadas duas drogas quanto à sua eficácia e segurança, durante 12 semanas de terapia contínua. Os 378 pacientes associados a 36 centros de atendimento foram observados por 12 semanas durante o tratamento e, posteriormente, acompanhados até um total de 48 semanas (12 meses). As medidas foram tomadas no início do tratamento, a cada mês durante o tratamento e a cada 3 meses após o tratamento, resultando em, no máximo, 7 medidas por paciente. Antes do início do tratamento, o medico responsável selecionou a unha a ser observada durante o período. O estudo será limitado aos pacientes cuja unha selecionada é uma dos primeiros pododáctilos ("dedão do pé"), restringindo assim para um total de 294 pacientes. Os tratamentos foram aleatorizados aos pacientes, sendo 146 para o tratamento A e 148 para o tratamento B. Uma das variáveis respostas de interesse foi o grau de severidade, indentificadas como 0 (não-severo) e 1 (severo). Duas são as questões de interesse, identificar como o percentual de severidade decresce com o tempo e se os dois tratamentos diferem ou não.

Na Figura 2, pode-se observar a evolução do percentual de severidade para os tratamentos A e B, ao longo do tempo. É possível observar que existe influência do tempo no percentual de severidade para ambos os tratamentos, mas não fica claro se os mesmos diferem significativamente. Vale ressaltar que apenas 76% dos pacientes possuem medidas nos 7 tempos.

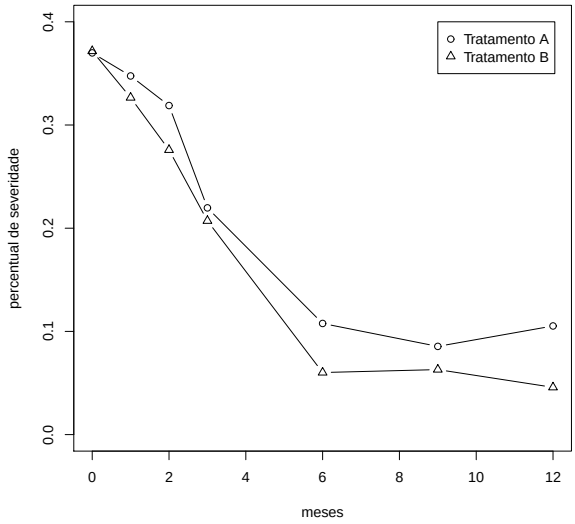


Figura 2 – Gráfico da evolução da proporção de severidade para os tratamentos A e B nos meses de observação dos pacientes submetidos ao tratamento de TDO

3.2 Métodos

3.2.1 Análise Bayesiana para os dados de epilepsia

Os modelos utilizados para análise dos dados de epilepsia são aqueles apresentados na seção 2.8, porém com enfoque bayesiano, como descritos a seguir.

Seja Y_{ij} o número de ataques epilépticos sofridos pelo paciente i na semana j, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$. Considere também, t_{ij} como a semana em que foi realizada a medida Y_{ij}, $t_{ij} = 1, 2, \ldots, 27$; $\boldsymbol{y} = (y_{11}, \ldots, y_{1n_1}, \ldots, y_{N1}, \ldots, y_{Nn_N})$ e $\boldsymbol{t} = (t_{11}, \ldots, t_{1n_1}, \ldots, t_{N1}, \ldots, t_{Nn_N})$.

Modelo 1: Poisson

A princípio, será assumido um modelo considerando que Y_{ij} segue a distribuição de Poisson com função densidade de probabilidade dada por (11) e parâmetro

$$\lambda_{ij} = \begin{cases} \exp\{\beta_{00} + \beta_{01} t_{ij}\} & \text{se} \quad i \in \text{placebo} \\ \exp\{\beta_{10} + \beta_{11} t_{ij}\} & \text{se} \quad i \in \text{tratamento.} \end{cases}$$

Assumindo-se independência entre os parâmetros $\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}$ a incerteza relativa aos mes-

mos é incorporada ao modelo, considerando-se as seguintes distribuições a priori:

$$\beta_{kl} \sim N(a, b^2); \quad a, b \quad \text{conhecidos} \tag{36}$$

com $k = 0, 1$ e $l = 0, 1$. Então, a distribuição conjunta a posteriori é proporcional a

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11} | \boldsymbol{y}, \boldsymbol{t}) &\propto \exp\left(-\sum_{i=1}^{N} \sum_{j=1}^{n_i} \lambda_{ij} \right) \prod_{i=1}^{N} \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \times \\ &\times \prod_{k=0}^{1} \prod_{l=0}^{1} \exp\Big\{ -\frac{1}{2b^2} (\beta_{kl} - a)^2 \Big\}. \end{aligned} \tag{37}$$

Modelo 2: Poisson-gama

Para esse modelo, será assumido que Y_{ij} tem distribuição de Poisson com função densidade de probabilidade dada por (11) e para acomodar a superdispersão considera-se um efeito aleatório θ_{ij} com distribuição gama, como a seguir:

$$\begin{aligned} Y_{ij} | \theta_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\ \lambda_{ij} &= \begin{cases} \theta_{ij} \exp\{\beta_{00} + \beta_{01} t_{ij}\} & \text{se} \quad i \in \text{placebo} \\ \theta_{ij} \exp\{\beta_{10} + \beta_{11} t_{ij}\} & \text{se} \quad i \in \text{tratamento} \end{cases} \\ \theta_{ij} &\sim \text{Gamma}(\alpha_2^{-1}, \alpha_2). \end{aligned}$$

Para uma análise bayesiana hierárquica, primeiramente, assume-se a distribuição a priori dada por (36) para os parâmetros β_{kl} , $k = 0, 1$ e $l = 0, 1$. Também, assume-se que os θ_{ij} são independentes e a distribuição a priori para o parâmtero α_2 é dada por

$$\alpha_2 \sim \text{Gamma}(c, d); \quad c, d \quad \text{conhecidos.} \tag{38}$$

Assumindo independência entre todos os parâmetros, a distribuição conjunta a posteriori é dada por

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \alpha_2 | \boldsymbol{y}, \boldsymbol{t}) &\propto \exp\left(-\sum_{i=1}^{N} \sum_{j=1}^{n_i} \lambda_{ij} \right) \prod_{i=1}^{N} \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \times \\ &\times \prod_{k=0}^{1} \prod_{l=0}^{1} \exp\Big\{ -\frac{1}{2b^2} (\beta_{kl} - a)^2 \Big\} \times \\ &\times \prod_{i=1}^{N} \prod_{j=1}^{n_i} \frac{\alpha_2^{-\alpha_2^{-1}}}{\Gamma(\alpha_2^{-1})} \theta_{ij}^{\alpha_2^{-1} - 1} e^{-\alpha_2^{-1} \theta_{ij}} \times \alpha_2^{c-1} e^{-\alpha_2/d} \end{aligned} \tag{39}$$

em que $\boldsymbol{\theta} = (\theta_{11}, \ldots, \theta_{1n_1}, \ldots, \theta_{N1}, \ldots, \theta_{Nn_N})$.

Modelo 3: Poisson-normal

Considerando que Y_{ij} tem distribuição de Poisson com função densidade de probabilidade dada por (11) e a distribuição normal para o efeito aleatório $\boldsymbol{b}_i$ para incorporar a correlação entre as medidas longitudinais segue:

$$\begin{aligned} Y_{ij}|b_i &\sim \text{Poisson}(\lambda_{ij}) \\ \lambda_{ij} &= \begin{cases} \exp\{(\beta_{00} + b_i) + \beta_{01} t_{ij}\} & \text{if} \quad i \in \text{placebo} \\ \exp\{(\beta_{10} + b_i) + \beta_{11} t_{ij}\} & \text{if} \quad i \in \text{tratamento} \end{cases} \\ b_i &\sim N(0, \sigma^2). \end{aligned}$$

Para a análise bayesiana hierárquica, primeiramente, assume-se a mesma distribuição a priori dada por (36) para os parâmetros β_{kl} , $k = 0, 1$ e $l = 0, 1$. Sob o modelo 3, considera-se que os b_i são variáveis aleatórias independentes com distribuição $N(0, \sigma^2)$. A seguir, considera-se que a distribuição a priori para σ^2 é dada por

$$\sigma^2 \sim \text{Gama Inversa}(g, h); \quad g, h \quad \text{conhecidos.} \tag{40}$$

Assumindo, também, indepedência a priori entre todos os parâmetros, a distribuição conjunta a posteriori é dada por

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{b}, \sigma^2 | \boldsymbol{y}, \boldsymbol{t}) &\propto \exp\left(- \sum_{i=1}^{N} \sum_{j=1}^{n_i} \lambda_{ij} \right) \prod_{i=1}^{N} \prod_{j=1}^{n_i} \lambda_{ij}^{y_{ij}} \times \\ &\times \prod_{k=0}^{1} \prod_{l=0}^{1} \exp\Big\{ - \frac{1}{2b^2} (\beta_{kl} - a)^2 \Big\} \times \\ &\times \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\Big\{ - \frac{b_i^2}{2\sigma^2} \Big\} \times (\sigma^2)^{-(g+1)} \exp\Big\{ - \frac{h}{\sigma^2} \Big\} \end{aligned} \tag{41}$$

em que $\boldsymbol{b} = (b_1, \ldots, b_N)$.

Modelo 4: Poisson-normal-gama (*Combinado*)

Por fim, considera-se a combinação dos modelos 2 e 3, isto é:

$$
\begin{aligned}
Y_{ij}|b_i,\theta_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\
\lambda_{ij} &= \begin{cases} \theta_{ij}\exp\{(\beta_{00}+b_i)+\beta_{01}t_{ij}\} & \text{se} \quad i \in \text{placebo} \\ \theta_{ij}\exp\{(\beta_{10}+b_i)+\beta_{11}t_{ij}\} & \text{se} \quad i \in \text{tratamento} \end{cases} \\
b_i &\sim N(0,\sigma^2) \\
\theta_{ij} &\sim \text{Gama}(\alpha_2^{-1},\alpha_2).
\end{aligned}
$$

Para a análise bayesiana hierárquica, primeiramente assume-se a mesma distribuição a priori dada por (36) para os parâmetros β_{kl} , $k = 0,1$ and $l = 0,1$. Considerando que b_i e θ_{ij} são independentes, assumem-se as seguintes distribuições a priori: (40) para o parâmetro σ^2 e (38) para o parâmetro α_2.

Assumindo, também, indepedência a priori entre todos os parâmetros, a distribuição conjunta a posteriori é dada por

$$
\begin{aligned}
\pi(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2|\boldsymbol{y},\boldsymbol{t}) &\propto \exp\left(-\sum_{i=1}^{N}\sum_{j=1}^{n_i}\lambda_{ij}\right)\prod_{i=1}^{N}\prod_{j=1}^{n_i}\lambda_{ij}^{y_{ij}} \times \\
&\times \prod_{k=0}^{1}\prod_{l=0}^{1}\exp\Big\{-\frac{1}{2b^2}(\beta_{kl}-a)^2\Big\} \times \\
&\times \prod_{i=1}^{N}\prod_{j=1}^{n_i}\frac{\alpha_2^{-\alpha_2^{-1}}}{\Gamma(\alpha_2^{-1})}\theta_{ij}^{\alpha_2^{-1}-1}e^{-\alpha_2^{-1}\theta_{ij}} \times \alpha_2^{c-1}e^{-\alpha_2/d} \times \\
&\times \prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big\{-\frac{b_i^2}{2\sigma^2}\Big\} \times (\sigma^2)^{-(g+1)}\exp\Big\{-\frac{h}{\sigma^2}\Big\}.
\end{aligned} \tag{42}
$$

Os resumos a posteriori para os parâmetros de interesse nas equações (37), (39), (41) e (42) serão obtidos, usando-se o método MCMC, por meio do algoritmo Metropolis-Hasting, pois as distribuições condicionais a posteriori para os parâmetros $(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2)$ não possuem forma fechada. As distribuições condicionais a posteriori para o algoritmo Metropolis-Hasting são apresentadas no Anexo C. Para a implementação computacional, será utilizado o *software* WinBUGS (Spiegelhalter et al., 1996) para o qual são necessários e suficientes a especificação estrutural do modelo, o conjunto de dados e os valores iniciais para cada parâmetro.

3.2.1.1 Obtenção das quantidades pivotais para avaliação do modelo Bayesiano

Considerando a teoria formulada por Johnson (2007) e a medida de discrepância (10) utilizada por Gelman, Meng e Stern (1996), ambas introduzidas na seção 2.7.5.2 e assumindo-se $\boldsymbol{y} = (y_{1;1}, \ldots, y_{1;n_1}, \ldots, y_{89;1}, \ldots, y_{89;n_{89}})$, $\boldsymbol{\lambda} = (\lambda_{1;1}, \ldots, \lambda_{1;n_1}, \ldots, \lambda_{89;1}, \ldots, \lambda_{89;n_{89}})$, $\boldsymbol{\theta} = (\theta_{1;1}, \ldots, \theta_{1;n_1}, \ldots, \theta_{89;1}, \ldots, \theta_{89;n_{89}})$ e $\boldsymbol{b} = (b_{1;1}, \ldots, b_{1;n_1}, \ldots, b_{89;1}, \ldots, b_{89;n_{89}})$, definem-se as seguintes quantidades pivotais para cada nível em cada modelo proposto em 3.2.1:

$$S(\boldsymbol{y}, \boldsymbol{\lambda}) = \sum_i \sum_j \frac{(y_{ij} - \lambda_{ij})^2}{\lambda_{ij}}, \tag{43}$$

$$S_b(\boldsymbol{b}, \sigma^2) = \sum_i \frac{(b_i - 0)^2}{\sigma^2} \tag{44}$$

e como, pela restrição imposta aos modelos com efeito gama, $\mathrm{E}(\theta_{ij}) = \alpha_1\alpha_2 = 1$ e $\mathrm{Var}(\theta_{ij}) = \alpha_1\alpha_2^2 = \alpha_2$

$$S_\theta(\boldsymbol{\theta}, \alpha_2) = \sum_i \sum_j \frac{(\theta_{ij} - 1)^2}{\alpha_2}. \tag{45}$$

Considerando-se que essas medidas são, aproximadamente, quantidades pivotais, a avaliação dos modelos será baseada na comparação da distribuição a posteriori dessas medidas com a suas respectivas distribuições χ^2_T, χ^2_{89} e χ^2_T em que $T = \sum_{i=1}^{89} n_i$, para o conjunto de dados de epilepsia $T = 1419$. Na prática, para comparação serão plotados os histogramas estimados de $S(\boldsymbol{y}, \boldsymbol{\lambda})$, $S_\theta(\boldsymbol{\theta}, \alpha_2)$ e $S_b(\boldsymbol{b}, \sigma^2)$ avaliados na amostra a posteriori de $(\boldsymbol{\lambda}, \boldsymbol{\theta}, \alpha_2, \boldsymbol{b}, \sigma^2)$ e sua respectiva distribuição χ^2 marginal.

Para o modelo Poisson será plotado o histograma estimado de (43) para comparação com a distribuição χ^2_{1419}. Para o modelo Poisson-gama serão plotados os histogramas estimados (43) e (44) para serem comparados, respectivamente, com χ^2_{1419} e χ^2_{1419}. Para o modelo Poisson-normal serão potados os histogramas estimados (43) e (45) para serem comparados, respectivamente, com χ^2_{1419} e χ^2_{89}. Finalmente, para o modelo Poisson-normal-gama serão plotados os histogramas estimados de (43), (44) e (45) para serem comparados, respectivamente, com χ^2_{1419}, χ^2_{89} and χ^2_{1419}.

3.2.1.2 Implementação computacional

A implementação computacional foi realizada usando-se o programa WinBUGS (SPIEGELHALTER et al., 1996). Para cada modelo, foram geradas 2 cadeias e em cada uma dessas cadeias foram simulados 55.000 valores para cada parâmetro. A fim de eliminar os efeitos dos valores iniciais do algoritmo, desprezaram-se os 5.000 primeiros valores ("burn-in") e para diminuir a autocorrelação entre os valores amostrados seleciona-ses um a cada 10 dos 50.000 valores restantes ("thin"), formando, assim, uma amostra de 5.000 valores para cada parâmetro. As cadeias foram inicializadas em pontos distintos e tiveram a convergência monitorada pelo critério de convergência de Gelman e Rubin (1992) existente no programa WinBUGS.

Os hiperparâmetros das distribuições a priori foram escolhidos de forma que as distribuições a priori fossem não informativas. Os valores utilizados foram $a = 0$, $b = 1000$, $c = d = 0,001$ e $g = h = 0,0001$. Os códigos dos programas utilizados no *software* WinBUGS podem ser vistos no Apêndice A.

Para comparar os modelos propostos, utilizou-se uma análise gráfica das quantidades pivotais, o critério DIC e também uma outra análise gráfica foi realizada comparando-se a variância dos dados com a variância dos valores preditos por modelo, separadamente para os indivíduos submetidos ao placebo e ao tratamento. Para poder comparar os tratamentos, além dos parâmetros dos modelos, duas outras medidas foram monitoradas nos programas, a diferença entre os parâmetros que representam os coeficientes angulares para cada tratamento, $ds = \beta_{11} - \beta_{01}$, e a razão entre os mesmos, $rs = \dfrac{\beta_{11}}{\beta_{01}}$. Se os valores zero e 1 estiverem contidos nos intervalos de credibilidades de cada uma dessas quantidades, conclui-se que os tratamentos não diferem entre si. Cada medida foi analisada separadamente.

Baseado na distribuição preditiva a posteriori, o ajuste do modelo selecionado aos dados foi verificado, graficamente. Considerando o modelo escolhido, as cadeias obtidas pelo software WinBUGS para cada parâmetro desse modelo e as cadeias dos valores dos efeitos aleatórios em questão, que também foram monitorados, calcularam-se os valores preditos de λ_{ij}, $i = 1, \ldots, 89$, $j = 1, \ldots, 27$, conforme definido pelo modelo em questão. Como cada cadeia possui 5.000 valores, foram obtidos 5.000 conjuntos de valores preditos. Assim, para cada uma das 1419 combinações indivíduo×tempo têm-se 5.000 valores preditos. Para verificar o

ajuste do modelo deveriam ser feitos gráficos de barras baseados nesses 5.000 valores para cada indivíduo em cada tempo, e no mesmo gráfico uma linha indicando o valor verdadeiro. Se o valor verdadeiro pertence à região de maior probabilidade conclui-se que o ajuste do modelo é consistente, caso contrário, conclui-se que o modelo não está bem ajustado aos dados. Entretanto, seriam 1419 gráficos a serem analisados, assim, calculou-se a média desses 5.000 valores preditos para cada combinação indivíduo×tempo, obtendo-se assim, a média dos valores preditos. E por fim, foram plotados os valores observados versus a média dos valores preditos. Para um bom ajuste do modelo, espera-se que esses valores sejam próximos, ou seja, que os pontos no gráfico estejam próximos à reta identidade.

3.2.1.3 Estudo de simulação

A fim de pesquisar a eficiência da utilização de quantidades pivotais como método de avaliação do modelo bayesiano foi realizado um estudo de simulação. Tomando como base os dados de epilepsia foram gerados três diferentes conjuntos dos seguintes modelos: Poisson-gama, Poisson-normal e Poisson-normal-gama (combinado), usando-se o software R (R DEVELOPMENT CORE TEAM, 2009). Foram considerados 89 pacientes, sendo 45 indivíduos tratados com placebo e 44 indivíduos tratados com a nova droga anti-epiléptica, os quais foram observados durante um período de 27 semanas, sem perda de observações. Os valores para os parâmetros utilizados na simulação foram obtidos de Molenberghs, Verbeke e Demétrio (2007). A implementação computacional para análise Bayesiana dos dados simulados, foi baseada nos procedimentos descritos na seção 3.2.1.2.

a. Modelo Poisson-gama

Para gerar dados a partir do modelo Poisson-gama, foram seguidos os seguintes passos:

1: Gerar 2403 valores para o efeito aleatório gama assumindo $\theta_{ij} \sim \text{Gama}(2.4640, 0.4059)$, $i = 1, \ldots, 89$, $j = 1, \ldots, 27$.

2: Calcular os valores para λ_{ij}, $i = 1, \ldots, 89$, $j = 1, \ldots, 27$ em que $\lambda_{ij} = \theta_{ij}\exp(1.2662 - 0.0134 * t_j)$, $i = 1, \ldots, 45$ e $t_j = 1, 2, \ldots, 27$ para indivíduos que receberam placebo e $\lambda_{ij} = \theta_{ij}\exp(1.4531 - 0.0328 * t_j)$, $i = 46, \ldots, 89$ e $t_j = 1, 2, \ldots, 27$ para indivíduos que receberam a nova droga anti-epiléptica.

3: Finalmente, gerar 2403 valores $Y_{ij} \sim \text{Poisson}(\lambda_{ij})$.

Para a análise desses dados foram considerados os modelos Poisson, Poisson-gama e Poisson-normal-gama, usando-se o enfoque Bayesiano.

b. Modelo Poisson-normal

Para gerar dados a partir do modelo Poisson-normal, foram seguidos os seguintes passos:

1: Gerar 89 valores para o efeito aleatório normal assumindo $b_i \sim \text{Normal}(0, 1.1289)$, $i = 1, \ldots, 89$.

2: Calcular os valores para λ_{ij}, $i = 1, \ldots, 89$, $j = 1, \ldots, 27$ em que $\lambda_{ij} = \exp(1.2662 + b_i - 0.0134 * t_j)$, $i = 1, \ldots, 45$ e $t_j = 1, 2, \ldots, 27$ para indivíduos que receberam placebo e $\lambda_{ij} = \exp(1.4531 + b_i - 0.0328 * t_j)$, $i = 46, \ldots, 89$ e $t_j = 1, 2, \ldots, 27$ para indivíduos que receberam a nova droga anti-epiléptica.

3: Finalmente, gerar 2403 valores $Y_{ij} \sim \text{Poisson}(\lambda_{ij})$.

Para a análise desse conjunto de dados foram considerados os modelos Poisson, Poisson-normal e Poisson-normal-gama, usando-se o enfoque Bayesiano.

c. Poisson-normal-gamma model

Para gerar dados a partir do modelo Poisson-normal-gama, foram seguidos os seguintes passos:

1: Gerar 2403 valores para o efeito aleatório gama assumindo $\theta_{ij} \sim \text{Gama}(2.4640, 0.4059)$, $i = 1, \ldots, 89$, $j = 1, \ldots, 27$.

2: Gerar 89 valores para o efeito aleatório normal assumindo $b_i \sim \text{Normal}(0, 1.1289)$, $i = 1, \ldots, 89$.

3: Calcular os valores para λ_{ij}, $i = 1, \ldots, 89$, $j = 1, \ldots, 27$ em que $\lambda_{ij} = \theta_{ij}\exp(1.2662 + b_i - 0.0134 * t_j)$, $i = 1, \ldots, 45$ e $t_j = 1, 2, \ldots, 27$ para indivíduos que receberam placebo e $\lambda_{ij} = \theta_{ij}\exp(1.4531 + b_i - 0.0328 * t_j)$, $i = 46, \ldots, 89$ e $t_j = 1, 2, \ldots, 27$ para indivíduos que receberam a nova droga anti-epiléptica.

4: Finalmente, gerar 2403 valores $Y_{ij} \sim \text{Poisson}(\lambda_{ij})$.

Para a análise deste conjunto de dados foram considerados os modelos Poisson, Poisson-gama, Poisson-normal e Poisson-normal-gama, usando-se o enfoque Bayesiano.

3.2.2 Inferência via Dclone e estudo da estimabilidade dos modelos para os dados de epilepsia

Para análise via o pacote Dclone do ambiente R, utilizou-se o conjunto de dados reais referente ao número de ataques epilépticos a fim de verificar a estimabilidade dos modelos e obter as estimativas dos parâmetros dos mesmos. Para implementação computacional foram utilizadas as mesmas especificações dadas na seção 3.2.1.2 para cada modelo. Entretanto, foram geradas 3 cadeias ao invés de 2. Apesar de a escolha das prioris não influenciar as estimativas de máxima verossimilhança obtidas pelo pacote Dclone (LELE; NADEEM; SCHMULAND, 2010), foram utilizadas as mesmas prioris citadas na seção 3.2.1.2 para facilidade computacional.

Os conjuntos de dados clonados foram obtidos de forma a multiplicar o número de indivíduos por k, porém, manter o número de tempos de cada indivíduo, ou seja, para obter os dados clonados copiam-se os dados originais no sentido de multiplicar apenas o número de indivíduos. Assim, também, multiplicou-se o número de efeitos aleatórios com distribuições normal e gama. Para a análise consideram-se $k = 1, 10, 25, 50$. A dimensão dos dados a ser multiplicada deve ser mencionada no *multiply* e a dimensão que deve permanecer igual deve ser mencionada no *unchanged*, ambos parâmetros da função dc.fit do pacote Dclone. A função dc.fit ainda permite escolher qual pacote será utilizado na implementação do método MCMC para cada clone selecionado. Nesse caso, foi utilizado o programa JAGS (PLUMMER, 2010a).

Os modelos a serem analisados são os mesmos da seção 2.8. Para implementação computacional desses modelos, inicialmente, define-se uma função que contém o modelo na linguagem BUGS, como os apresentados no Apêndice A. Os códigos dos programas utilizados no ambiente R podem ser vistos no Apêndice B.

3.2.3 Análise Bayesiana para os dados de dermatite

Os modelos utilizados para análise dos dados de dermatite são aqueles apresentados na seção 2.9, porém com enfoque bayesiano, como descritos a seguir.

Seja Y_{ij} o grau de severidade da unha em obseração do paciente i na semana j, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$. Considere também, t_{ij} como o mês em que foi realizada a medida Y_{ij}, $t_j = (0, 1, 2, 3, 6, 9, 12)$; $\boldsymbol{y} = (y_{11}, \ldots, y_{1n_1}, \ldots, y_{N1}, \ldots, y_{Nn_N})$ e $\boldsymbol{t} = (t_{11}, \ldots, t_{1n_1}, \ldots, t_{N1}, \ldots, t_{Nn_N})$.

Modelo 1: Bernoulli

A princípio, será assumido um modelo considerando que Y_{ij} segue a distribuição de Bernoulli com função densidade de probabilidade dada por (25) e parâmetro π_{ij} em que, ao se considerar a função de ligação logística, tem-se

$$\pi_{ij} = \begin{cases} \frac{\exp(\beta_{00}+\beta_{01}t_j)}{1+\exp(\beta_{00}+\beta_{01}t_j)} & \text{se} \quad i \in \text{Tratamento A} \\ \frac{\exp(\beta_{10}+\beta_{11}t_j)}{1+\exp(\beta_{10}+\beta_{11}t_j)} & \text{se} \quad i \in \text{Tratamento B} \end{cases}$$

e, ao se considerar a função de ligação probit, tem-se

$$\pi_{ij} = \begin{cases} \Phi(\beta_{00} + \beta_{01}t_j) & \text{se} \quad i \in \text{Tratamento A} \\ \Phi(\beta_{10} + \beta_{11}t_j) & \text{se} \quad i \in \text{Tratamento B} \end{cases}$$

em que $\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \exp\big[-t^2/2\big] dt$.

Assumindo-se independência entre os parâmetros $\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}$ a incerteza relativa aos mesmos é incorporada ao modelo, considerando-se as seguintes distribuições a priori:

$$\beta_{kl} \sim N(a, b^2); \quad a, b \quad \text{conhecidos} \tag{46}$$

com $k = 0, 1$ e $l = 0, 1$. Então, a distribuição conjunta a posteriori é proporcional a

$$\pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11} | \boldsymbol{y}, \boldsymbol{t}) \propto \prod_{i=1}^{N} \prod_{j=1}^{n_i} \pi_{ij}^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \times \prod_{k=0}^{1} \prod_{l=0}^{1} \exp\Big\{ -\frac{1}{2b^2} (\beta_{kl} - a)^2 \Big\} \tag{47}$$

Modelo 2: Bernoulli-beta

Para esse modelo, será assumido que Y_{ij} tem distribuição de Bernoulli com função densidade de probabilidade dada por (25) com parâmetro π_{ij} e a distribuição beta para o efeito aleatório θ_{ij} para incorporar a superdispersão, ao se considerar a função de ligação logística, tem-se

$$\begin{aligned} Y_{ij}|\theta_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \begin{cases} \theta_{ij}\frac{\exp(\beta_{00}+\beta_{01}t_j)}{1+\exp(\beta_{00}+\beta_{01}t_j)} & \text{se} \quad i \in \text{Tratamento A} \\ \theta_{ij}\frac{\exp(\beta_{10}+\beta_{11}t_j)}{1+\exp(\beta_{10}+\beta_{11}t_j)} & \text{se} \quad i \in \text{Tratamento B} \end{cases} \\ \theta_{ij} &\sim \text{Beta}(\alpha_1, \alpha_2) \end{aligned}$$

e, ao se considerar a função de ligação probit, tem-se

$$\begin{aligned} Y_{ij}|\theta_{ij} &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \begin{cases} \theta_{ij}\Phi(\beta_{00}+\beta_{01}t_j) & \text{se} \quad i \in \text{Tratamento A} \\ \theta_{ij}\Phi(\beta_{10}+\beta_{11}t_j) & \text{se} \quad i \in \text{Tratamento B} \end{cases} \\ \theta_{ij} &\sim \text{Beta}(\alpha_1, \alpha_2) \end{aligned}$$

em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$.

Para uma análise bayesiana hierárquica, primeiramente, assume-se a distribuição a priori dada por (46) para os parâmetros β_{kl} , $k = 0, 1$ e $l = 0, 1$. Sob o modelo 2, tanto para ligação logística como para a probit, considera-se que os θ_{ij} são variáveis aleatórias independentes com distribuição Beta(α_1, α_2). A seguir, considera-se que as distribuições a priori para os parâmetros α_1 e α_2 são dadas por

$$\alpha_q \sim \text{Gamma}(c, d); \quad c, d \quad \text{conhecidos}, \quad q = 1, 2. \tag{48}$$

Assumindo, também, indepedência a priori entre todos parâmetros, a distribuição conjunta a posteriori é dada por

$$\begin{aligned} \pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \alpha_1, \alpha_2|\boldsymbol{y}, \boldsymbol{t}) &\propto \prod_{i=1}^{N}\prod_{j=1}^{n_i}\pi_{ij}^{y_{ij}}(1-\pi_{ij})^{1-y_{ij}} \times \prod_{k=0}^{1}\prod_{l=0}^{1}\exp\Big\{-\frac{1}{2b^2}(\beta_{kl}-a)^2\Big\} \times \\ &\times \prod_{i=1}^{N}\prod_{j=1}^{n_i}\frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1,\alpha_2)} \times \alpha_1^{c-1}e^{-\alpha_1/d} \times \alpha_2^{c-1}e^{-\alpha_2/d} \end{aligned} \tag{49}$$

$\boldsymbol{\theta} = (\theta_{11}, \ldots, \theta_{1n_1}, \ldots, \theta_{N1}, \ldots, \theta_{Nn_N})$.

Modelo 3: Bernoulli-normal

Considerando que Y_{ij} tem distribuição de Bernoulli com função densidade de probabilidade dada por (25) com parâmetro π_{ij} e a distribuição normal para o efeito aleatório $\boldsymbol{b}_i$ para incorporar a correlação entre as medidas longitudinais, ao se considerar a função de ligação logística, tem-se

$$\begin{aligned} Y_{ij}|b_i &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \begin{cases} \frac{\exp(\beta_{00}+b_i+\beta_{01}t_j)}{1+\exp(\beta_{00}+b_i+\beta_{01}t_j)} & \text{se} \quad i \in \text{Tratamento A} \\ \frac{\exp(\beta_{10}+b_i+\beta_{11}t_j)}{1+\exp(\beta_{10}+b_i+\beta_{11}t_j)} & \text{se} \quad i \in \text{Tratamento B} \end{cases} \\ b_i &\sim N(0,\sigma^2) \end{aligned}$$

e, ao se considerar a função de ligação probit, tem-se

$$\begin{aligned} Y_{ij}|b_i &\sim \text{Bernoulli}(\pi_{ij}) \\ \pi_{ij} &= \begin{cases} \Phi(\beta_{00}+b_i+\beta_{01}t_j) & \text{se} \quad i \in \text{Tratamento A} \\ \Phi(\beta_{10}+b_i+\beta_{11}t_j) & \text{se} \quad i \in \text{Tratamento B} \end{cases} \\ b_i &\sim N(0,\sigma^2) \end{aligned}$$

em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$.

Para a análise bayesiana hierárquica, primeiramente, assume-se a mesma distribuição a priori dada por (46) para os parâmetros β_{kl} , $k = 0, 1$ e $l = 0, 1$. Sob o modelo 3, tanto para ligação logística como para a probit, considera-se que os b_i são variáveis aleatórias independentes com distribuição $N(0, \sigma^2)$. A seguir, considera-se que a distribuição a priori para σ^2 é dada por

$$\sigma^2 \sim \text{Gama Inversa}(g,h); \quad g,h \quad \text{conhecidos.} \tag{50}$$

Assumindo, também, indepedência a priori entre todos parâmetros, a distribuição conjunta a posteriori é dada por

$$\begin{aligned} \pi(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b},\sigma^2|\boldsymbol{y},\boldsymbol{t}) &\propto \prod_{i=1}^{N}\prod_{j=1}^{n_i}\pi_{ij}^{y_{ij}}(1-\pi_{ij})^{1-y_{ij}} \times \prod_{k=0}^{1}\prod_{l=0}^{1}\exp\Big\{-\frac{1}{2b^2}(\beta_{kl}-a)^2\Big\} \times \\ &\times \prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big\{-\frac{b_i^2}{2\sigma^2}\Big\} \times (\sigma^2)^{-(g+1)}\exp\Big\{-\frac{h}{\sigma^2}\Big\} \end{aligned} \tag{51}$$

em que $\boldsymbol{b} = (b_1, \ldots, b_N)$.

Modelo 4: Bernoulli-normal-beta (*Combinado*)

Por fim, considera-se a combinação dos modelos 2 e 3, ou seja, ao se considerar a função de ligação logística, tem-se

$$\begin{aligned}
Y_{ij}|\theta_{ij}, b_i &\sim \text{Bernoulli}(\pi_{ij}) \\
\pi_{ij} &= \begin{cases} \theta_{ij}\frac{\exp(\beta_{00}+b_i+\beta_{01}t_j)}{1+\exp(\beta_{00}+b_i+\beta_{01}t_j)} & \text{se} \quad i \in \text{Tratamento A} \\ \theta_{ij}\frac{\exp(\beta_{10}+b_i+\beta_{11}t_j)}{1+\exp(\beta_{10}+b_i+\beta_{11}t_j)} & \text{se} \quad i \in \text{Tratamento B} \end{cases} \\
b_i &\sim N(0, \sigma^2) \\
\theta_{ij} &\sim \text{Beta}(\alpha_1, \alpha_2)
\end{aligned}$$

e, ao se considerar a função de ligação probit, tem-se

$$\begin{aligned}
Y_{ij}|\theta_{ij}, b_i &\sim \text{Bernoulli}(\pi_{ij}) \\
\pi_{ij} &= \begin{cases} \theta_{ij}\Phi(\beta_{00} + b_i + \beta_{01}t_j) & \text{se} \quad i \in \text{Tratamento A} \\ \theta_{ij}\Phi(\beta_{10} + b_i + \beta_{11}t_j) & \text{se} \quad i \in \text{Tratamento B} \end{cases} \\
b_i &\sim N(0, \sigma^2) \\
\theta_{ij} &\sim \text{Beta}(\alpha_1, \alpha_2)
\end{aligned}$$

em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$.

Para a análise bayesiana hierárquica, primeiramente assume-se a mesma distribuição a priori dada por (46) para os parâmetros β_{kl} , $k = 0, 1$ and $l = 0, 1$. Considerando que b_i e θ_{ij} são independentes, assumem-se as seguintes distribuições a priori: (50) para o parâmetro σ^2 e (48) para os parâmetros α_1 e α_2.

Assumindo, também, indepedência a priori entre todos parâmetros, a distribuição conjunta a posteriori é dada por

$$\begin{aligned}
\pi(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \alpha_1, \alpha_2|\boldsymbol{y}, \boldsymbol{t}) \propto & \prod_{i=1}^{N}\prod_{j=1}^{n_i}\pi_{ij}^{y_{ij}}(1-\pi_{ij})^{1-y_{ij}} \times \prod_{k=0}^{1}\prod_{l=0}^{1}\exp\Big\{-\frac{1}{2b^2}(\beta_{kl}-a)^2\Big\} \times \\
\times & \prod_{i=1}^{N}\prod_{j=1}^{n_i}\frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1, \alpha_2)} \times \alpha_1^{c-1}e^{-\alpha_1/d} \times \alpha_2^{c-1}e^{-\alpha_2/d} \times \\
\times & \prod_{i=1}^{N}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\Big\{-\frac{b_i^2}{2\sigma^2}\Big\} \times (\sigma^2)^{-(g+1)}\exp\Big\{-\frac{h}{\sigma^2}\Big\}. \qquad (52)
\end{aligned}$$

Os resumos a posteriori para os parâmetros de interesse nas equações (47), (49), (51) e (52) serão obtidos, usando-se o método MCMC, por meio do algoritmo Metropolis-Hasting, pois as distribuições condicionais a posteriori para os parâmetros $(\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \boldsymbol{b}, \alpha_1, \alpha_2, \sigma^2)$ não possuem forma fechada. As distribuições condicionais a posteriori para o algoritmo Metropolis-Hasting são dadas no Anexo D. Para a implementação computacional será utilizado o *software* WinBUGS (Spiegelhalter et al., 1996) para o qual são necessários e suficientes a especificação estrutural do modelo, o conjunto de dados e os valores iniciais para cada parâmetro.

3.2.3.1 Implementação computacional

A implementação computacional foi realizada usando-se o programa WinBUGS (SPIEGELHALTER et al., 1996). Para cada modelo, foram geradas 2 cadeias e em cada uma dessas cadeias foram simulados 105.000 valores para cada parâmetro. A fim de eliminar os efeitos dos valores iniciais do algoritmo, desprezaram-se os 5.000 primeiros valores ("burn-in") e para diminuir a autocorrelação entre os valores amostrados seleciona-se um a cada 20 dos 50.000 valores restantes ("thin"), formando, assim, uma amostra de 5.000 valores para cada parâmetro. As cadeias foram inicializadas em pontos distintos e tiveram a convergência monitorada pelo critério de convergência de Gelman e Rubin (1992) existente no programa WinBUGS.

Os hiperparâmetros das distribuições a priori foram escolhidos de forma que as distribuições a priori fossem não informativas. Os valores utilizados foram $a = 0$, $b = 1000$, $c = d = 0,001$ e $g = h = 0,0001$. Os códigos dos programas utilizados no *software* WinBUGS podem ser vistos no Apêndice D. Na tentativa de introduzir as restrições necessárias nos modelos 2 e 4, conforme citado na seção 2.10, utiliza-se o código sugerido por Ntzoufras (2009) usando-se o "zero-ones trick".

4 RESULTADOS E DISCUSSÃO

Inicialmente, será apresentada uma análise exploratória dos dados de contagens considerando-se o modelo de Poisson, que é baseado em duas fortes hipóteses: i) independência entre os indivíduos e ii) a mesma média λ para todos os indivíduos tratados. A falha de um ou ambos os pressupostos básicos fará que a variância observada seja maior do que a média, a variância teórica estimada da distribuição de Poisson, acarretando a superdispersão.

Considerando um modelo Poisson para análise dos dados, se a *deviance* residual for maior do que o número de graus de liberdade a ela associados, então, temos um problema de falta de ajuste do modelo e, provavelmente, um problema de superdispersão. Para uma análise exploratória dos dados de epilepsia consideram-se duas opções de modelos ajustados:

i) Para cada indivíduo, um modelo de Poisson com preditor linear $\eta = \log(\mu) = \beta_0 + \beta_1 t$, resultando em 45 análises para indivíduos que receberam placebo e 44 análises para os indivíduos que receberam a nova droga anti-epiléptica.

Os valores das *deviances* e seus respectivos números de graus de liberdade foram plotados na Figura 3a, e, eliminando o indivíduo 38, observa-se a Figura 3b. Como esperado, devido à correlação entre medidas repetidas no mesmo indivíduo, observa-se que o modelo Poisson log-linear não se ajusta bem aos dados.

ii) Para cada tempo, considera-se um modelo Poisson com preditor linear $\eta = \log(\mu) = \beta_0 + \alpha_i$, $i = 1$ (placebo) e $i = 2$ (tratamento).

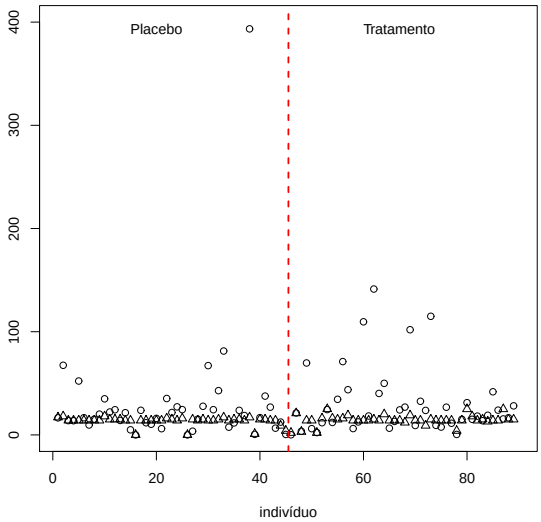
Na Figura 3c, observa-se o gráfico dos valores para as *deviances* por seus respectivos números de graus de liberdade. Como esperado, devido à variabilidade entre indivíduos, observa-se que o modelo Poisson log-linear não se ajusta bem aos dados, mostrando superdispersão.

Com essa análise exploratória inicial, percebe-se que para uma análise completa do conjunto de dados, necessita-se de dois diferentes tipos de efeitos aleatórios: i) um para acomodar no modelo a correlação entre as medidas repetidas no mesmo indivíduo e ii) um para incorporar ao modelo a variabilidade entre indivíduos.

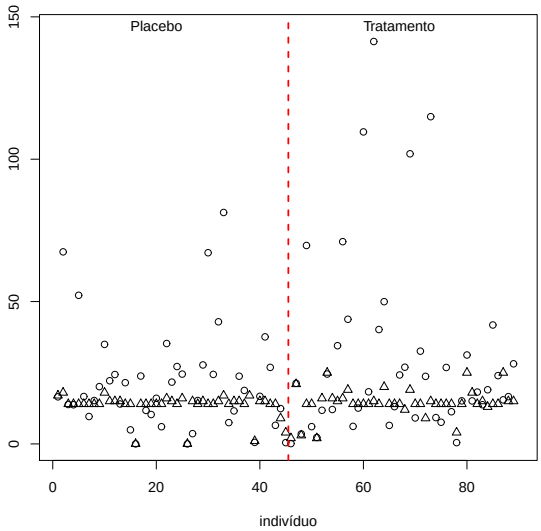
É importante notar que, além da superdispersão, esses dados apresentam perda de informações; para o grupo com placebo não houve observação a partir da 21ª semana. Além disso, na 17ª semana, o número de pacientes em observação para ambos os grupos de

tratamento se reduz a um pouco menos da metade do valor inicial, 18 pacientes para placebo e 17 para tratamento, e com o passar das semanas este número apenas diminui, como pode ser visto na Tabela 1.

a)



b)



c)

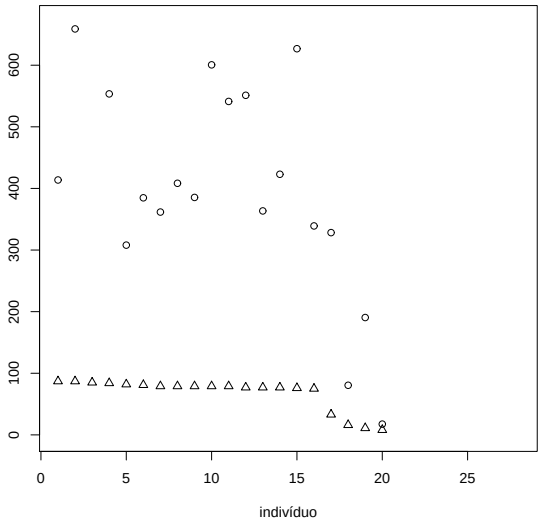


Figura 3 – a) Valores das *deviances* (○) e seus respectivos números de graus de liberdade (△) considerando um modelo Poisson log-linear para cada indivíduo. b) Valores das *deviances* (○) e seus respectivos números de graus de liberdade (△) considerando um modelo Poisson log-linear para cada indivíduo, eliminando o indivíduo 38. c) Valores das *deviances* (○) e seus respectivos números de graus de liberdade (△) considerando um modelo Poisson log-linear para cada tempo

Tabela 1 – Número de observações disponíveis em cada semana de observação para ambos os tratamentos

	Número de observações			Número de observações	
tempo(semanas)	Placebo	Tratamento	tempo(semanas)	Placebo	Tratamento
1	45	44	16	40	37
5	42	42	17	18	17
10	41	40	20	2	8
15	40	38	27	0	3

4.1 Análise Bayesiana para os dados de epilepsia

Os resultados obtidos para a metodologia desenvolvida na seção 3.2.1 e aplicada aos dados da seção 3.1.1 são apresentados e discutidos, a seguir. Primeiramente, serão apresentadas as análises gráficas das cadeias dos parâmetros de cada um dos quatro modelos, para verificação da convergência dos mesmos. Para que a convergência seja verificada, espera-se que os histogramas das distribuições a posteriori para cada um dos parâmetros, sejam unimodais e suaves; que a trajetória das cadeias dos parâmetros, sejam aleatórias e que a autocorrelação das cadeias dos parâmetros decaiam rapidamente ao longo das iterações. A seguir, será verificado qual modelo se ajusta melhor aos dados e, por fim, uma tabela com os resultados a posteriori para cada modelo.

Modelo 1: Poisson

As Figuras de 4 a 6, apresentam as análises gráficas das cadeias de todos os parâmetros do modelo Poisson. Os gráficos apresentados nessas figuras permitem verificar que há indícios de convergência das cadeias e que as mesmas são não correlacionadas.

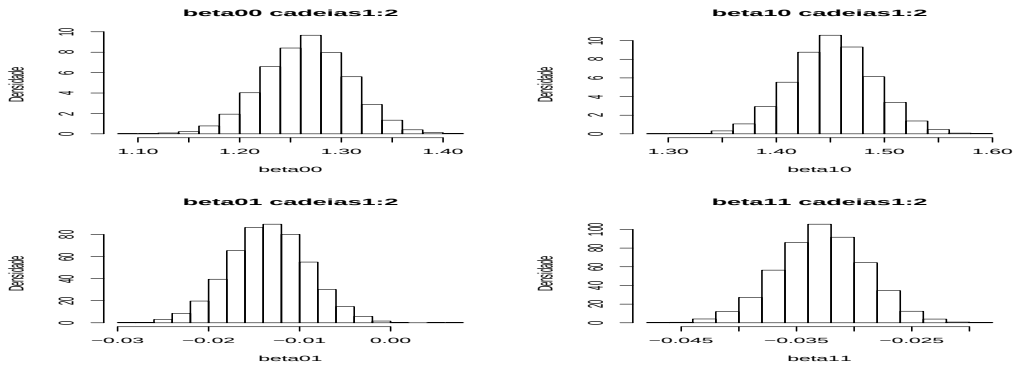


Figura 4 – Histograma dos valores gerados para β_{00}, β_{01}, β_{10} e β_{11}, considerando o conjunto de dados sob o modelo Poisson

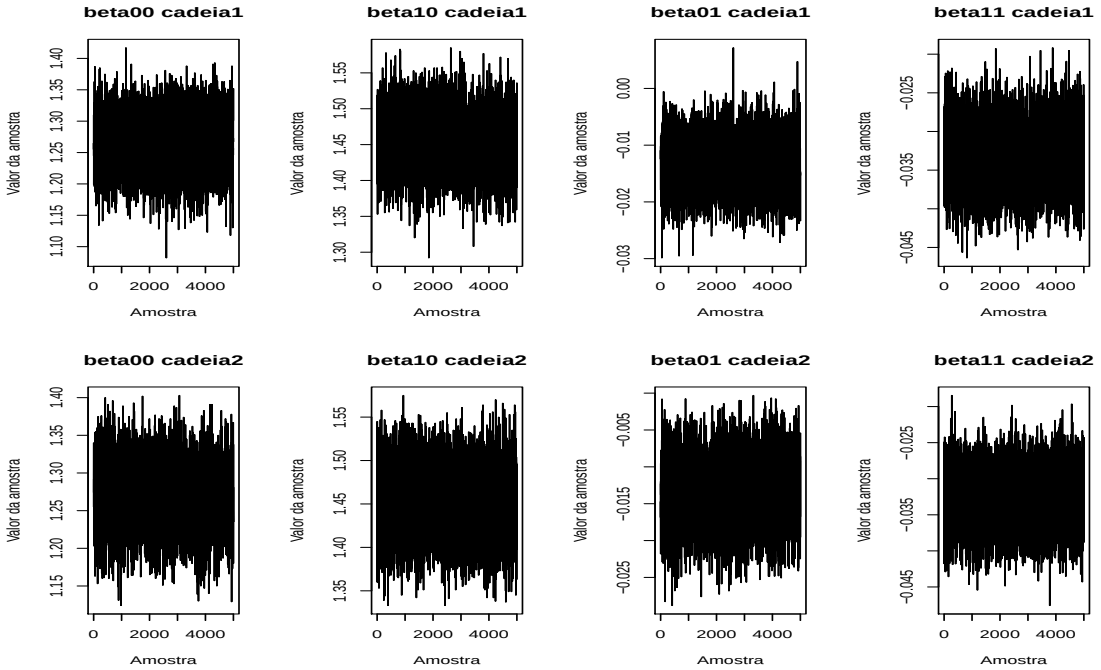


Figura 5 – Trajetória das cadeias geradas para β_{00}, β_{01}, β_{10} e β_{11}, considerando o conjunto de dados sob o modelo Poisson

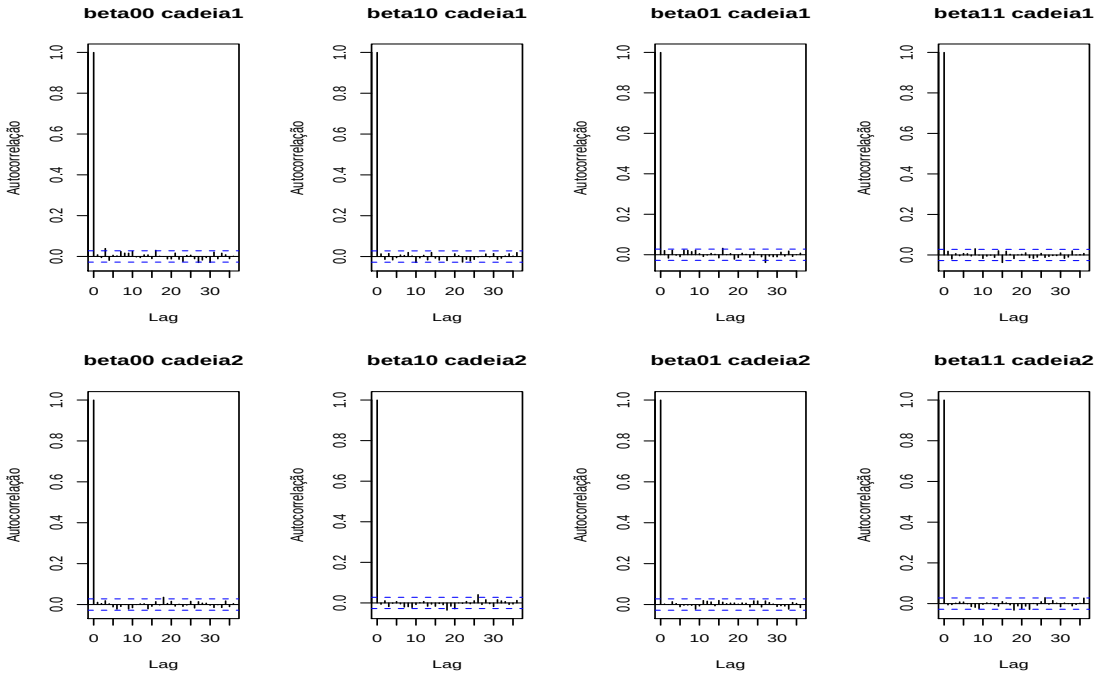


Figura 6 – Autocorrelação das cadeias geradas para β_{00}, β_{01}, β_{10} e β_{11}, considerando o conjunto de dados sob o modelo Poisson

Modelo 2: Poisson-gama

As Figuras de 7 a 9, apresentam as análises gráficas das cadeias de todos os parâmetros do modelo Poisson-gama. Os gráficos apresentados nessas figuras permitem verificar que há indícios de convergência das cadeias e que as mesmas são não correlacionadas.

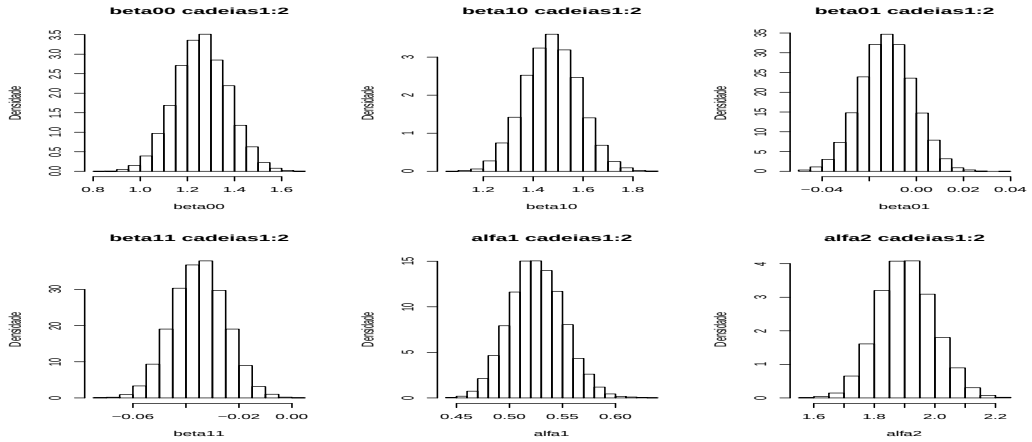


Figura 7 – Densidade dos valores gerados para β_{00}, β_{01}, β_{10}, β_{11}, α_1 e α_2 considerando o conjunto de dados sob o modelo Poisson-gama

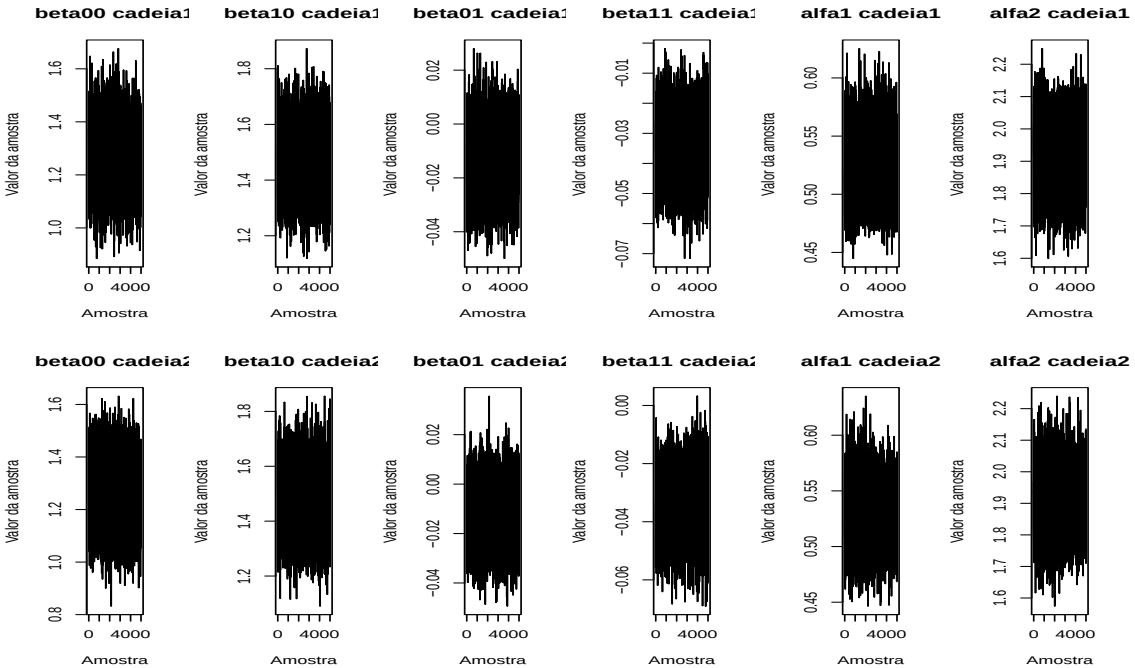


Figura 8 – Trajetória das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11}, α_1 e α_2 considerando o conjunto de dados sob o modelo Poisson-gama

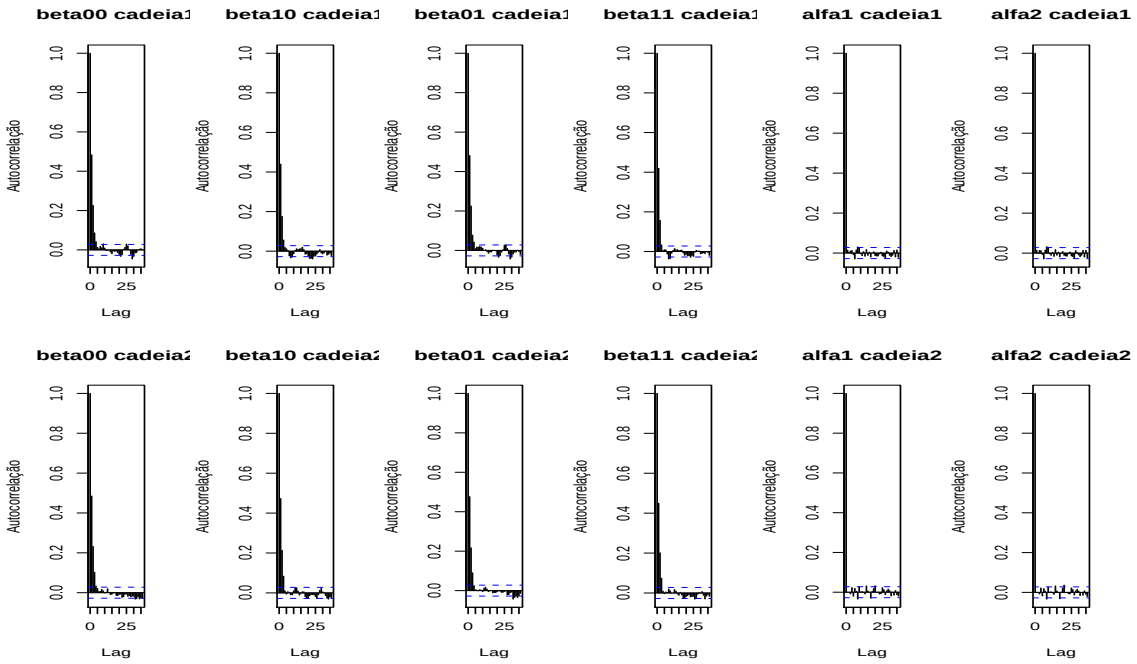


Figura 9 – Autocorrelação das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11}, α_1 e α_2 considerando o conjunto de dados sob o modelo Poisson-gama

Modelo 3: Poisson-normal

As Figuras de 10 a 12, apresentam as análises gráficas das cadeias de todos os parâmetros do modelo Poisson-normal. Os gráficos apresentados nessas figuras permitem verificar que há indícios de convergência das cadeias e que as mesmas são não correlacionadas.

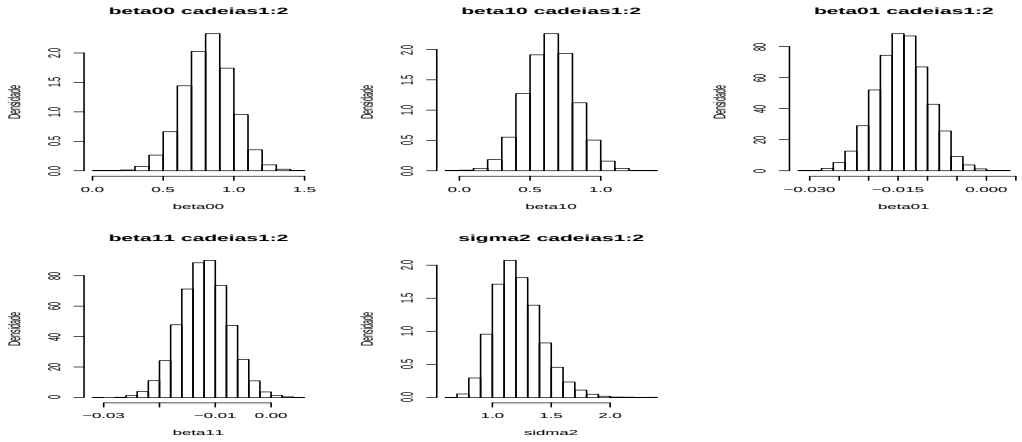


Figura 10 – Densidade dos valores gerados para β_{00}, β_{01}, β_{10}, β_{11} e σ^2 considerando o conjunto de dados sob o modelo Poisson-normal

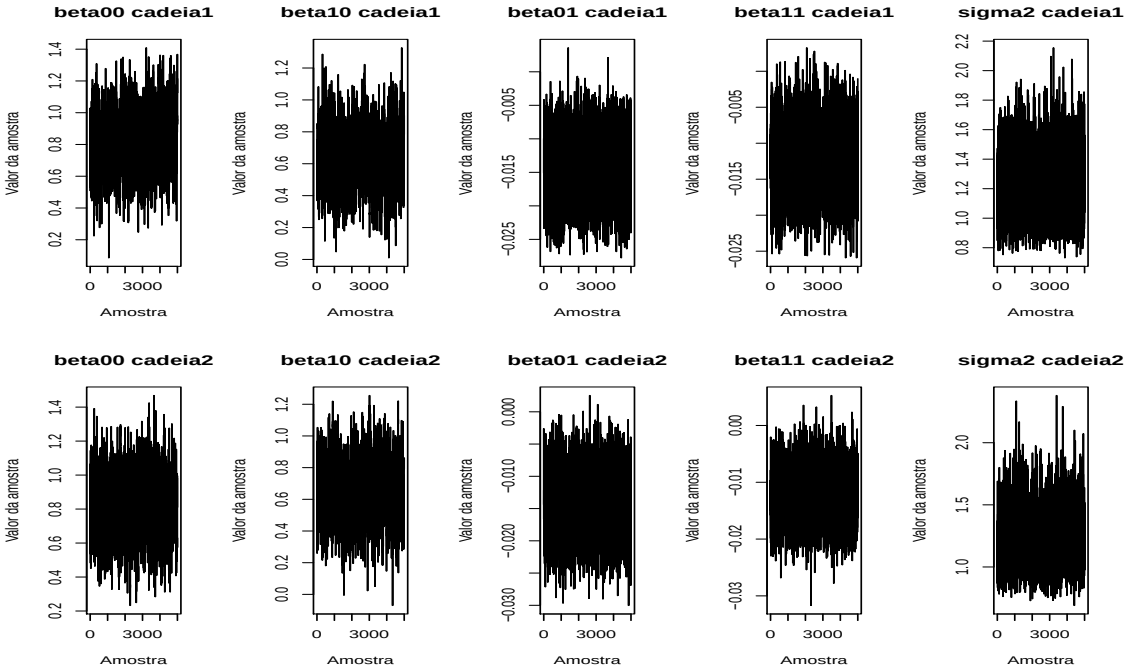


Figura 11 – Trajetória das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11} e σ^2 considerando o conjunto de dados sob o modelo Poisson-normal

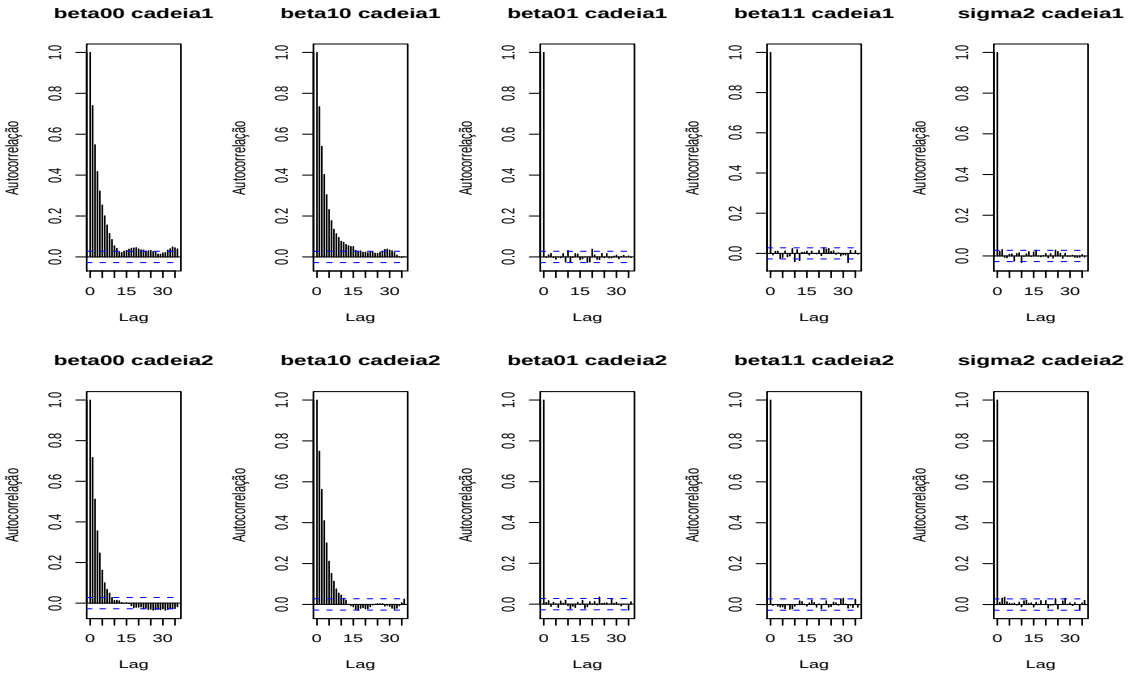


Figura 12 – Autocorrelação das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11} e σ^2 considerando o conjunto de dados sob o modelo Poisson-normal

Modelo 4: Poisson-normal-gama

As Figuras de 13 a 15, apresentam as análises gráficas das cadeias de todos os parâmetros do modelo Poisson-normal-gama. Os gráficos apresentados nessas figuras permitem verificar que há indícios de convergência das cadeias e que as mesmas são não correlacionadas.

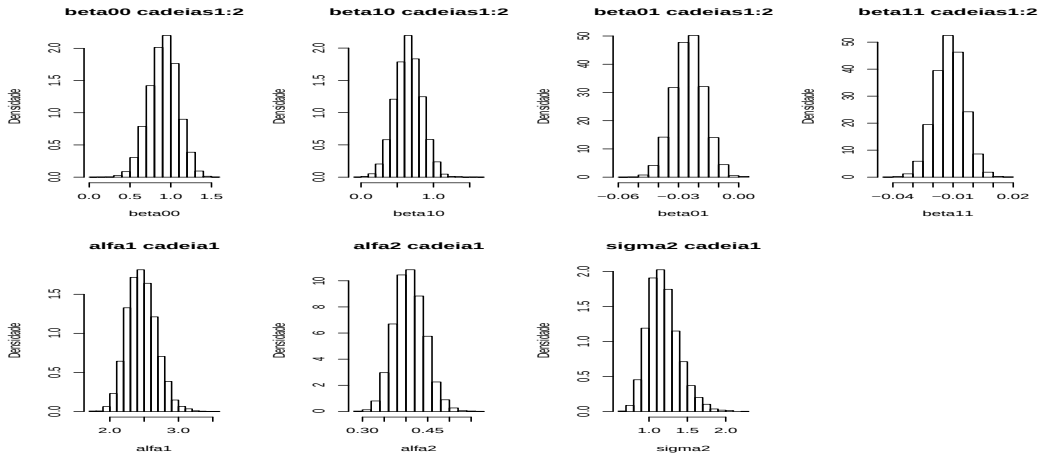


Figura 13 – Densidade dos valores gerados para β_{00}, β_{01}, β_{10}, β_{11}, α_1, α_2 e σ^2 considerando o conjunto de dados sob o modelo Poisson-normal-gama

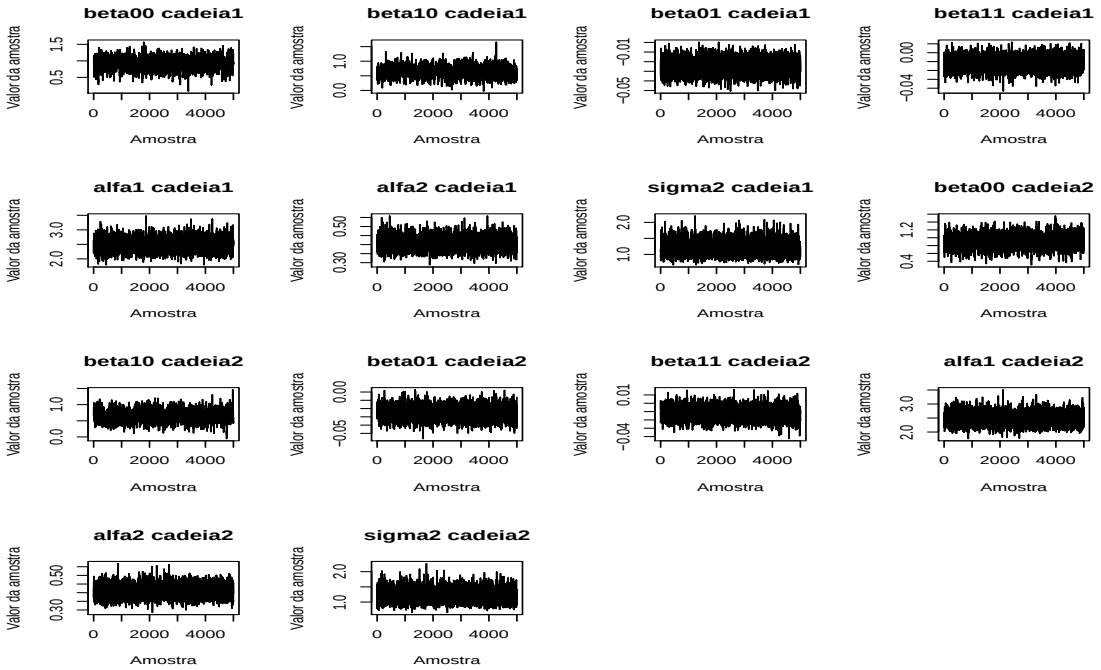


Figura 14 – Trajetória das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11}, α_1, α_2 e σ^2 considerando o conjunto de dados sob o modelo Poisson-normal-gama

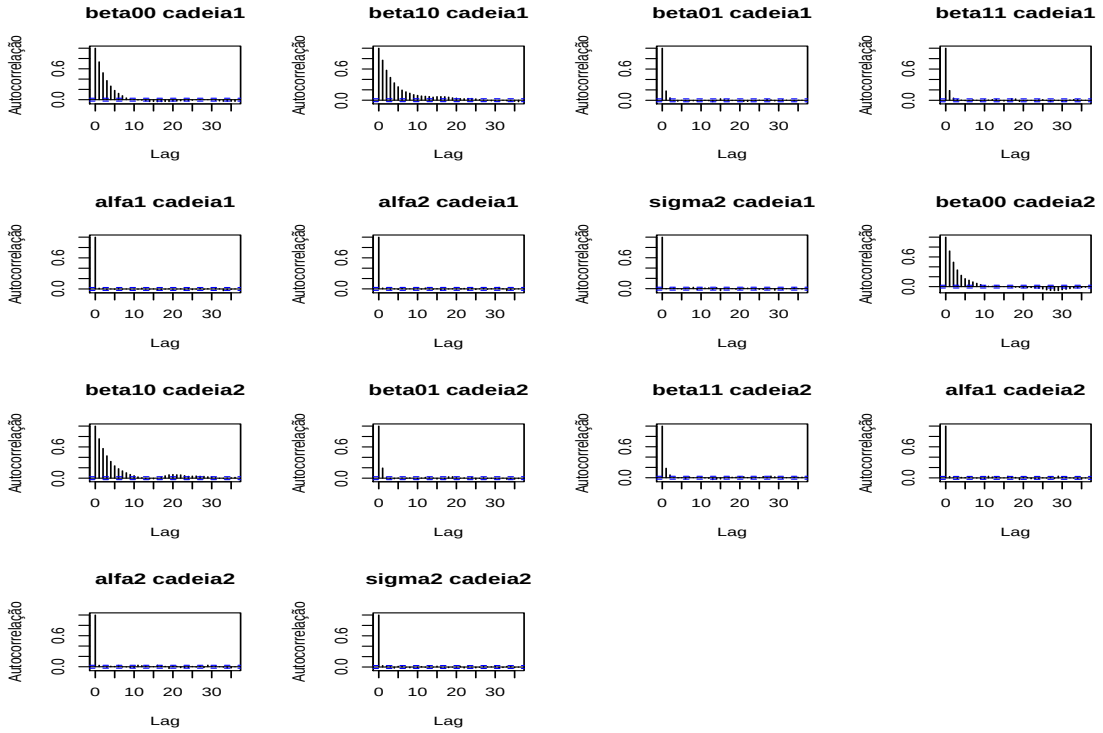


Figura 15 – Autocorrelação das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11}, α_1, α_2 e σ^2 considerando o conjunto de dados sob o modelo Poisson-normal-gama

Para diagnosticar qual modelo se ajusta melhor aos dados, nas Figuras 16 a 19, observam-se os histogranas estimados das distribuições a posteriori das medidas S, S_θ e S_b descritas anteriormente sob os modelos Poisson, Poisson-gama, Poisson-normal e combinado, respectivamente. Um bom ajuste do modelo é verificado quando os histogramas estimados das quantidades pivotais em cada nível do modelo se aproximam de sua respectiva distribuição de referência. A partir da análise gráfica, conclui-se que o modelo combinado é o que melhor se ajusta aos dados em todos os níveis do modelo hierárquico.

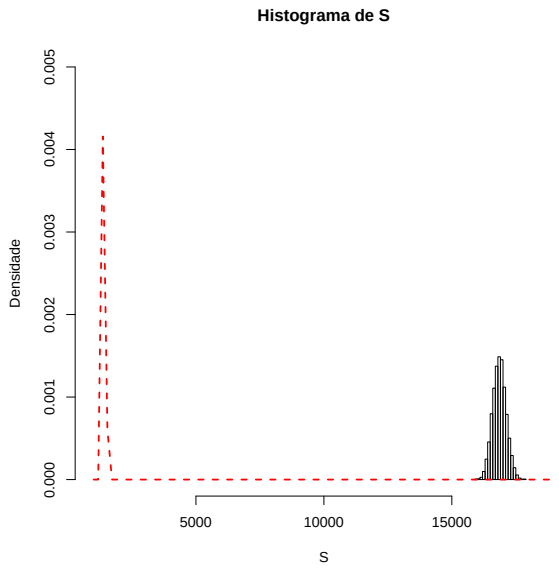


Figura 16 – Histograma estimado da distribuição a posteriori de $S(\boldsymbol{y}, \boldsymbol{\lambda}^i)$ sob o modelo Poisson e a respectiva distribuição de referência χ^2_{1419}

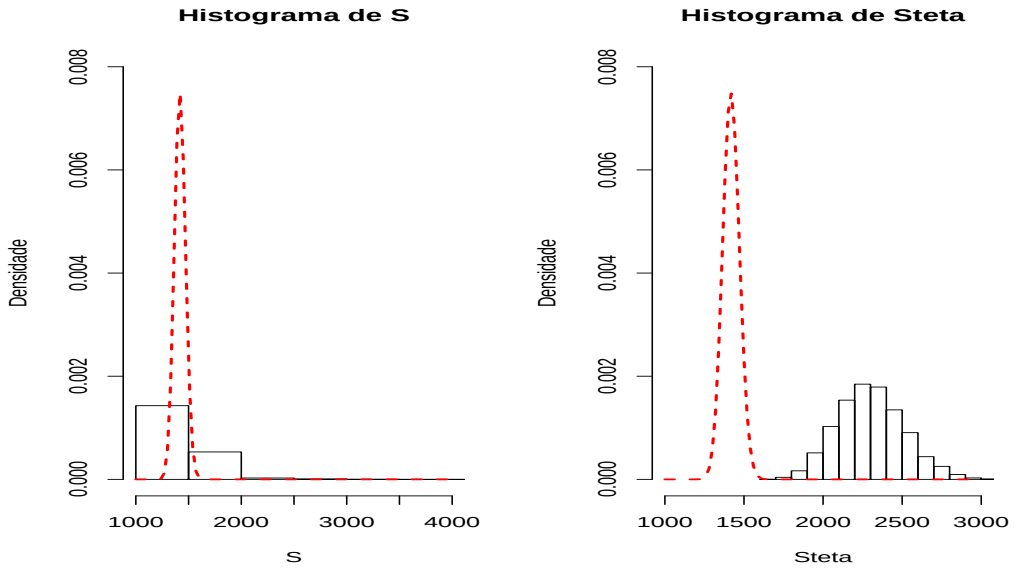


Figura 17 – Histograma estimado da distribuição a posteriori de $S(\boldsymbol{y}, \boldsymbol{\lambda}^i)$ e $S_\theta(\boldsymbol{\theta}^i, {\alpha_2}^i)$ sob o modelo Poisson-gama e as respectivas distribuições de referência χ^2_{1419} e χ^2_{1419}

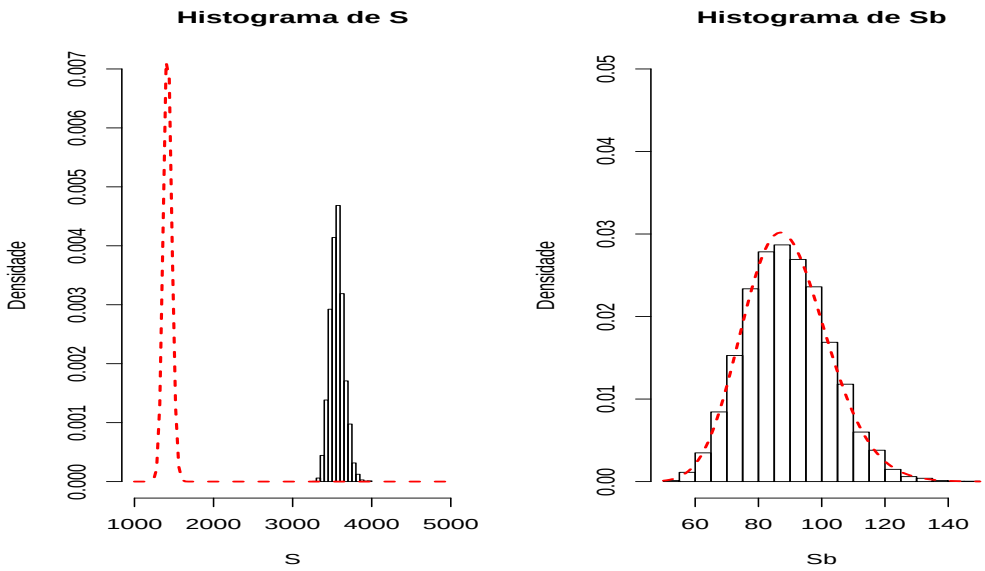


Figura 18 – Histograma estimado da distribuição a posteriori de $S(\boldsymbol{y}, \boldsymbol{\lambda}^i)$ e $S_b(\boldsymbol{b}^i, \sigma^{2^i})$ sob o modelo Poisson-normal e as respectivas distribuições de referência χ^2_{1419} e χ^2_{89}

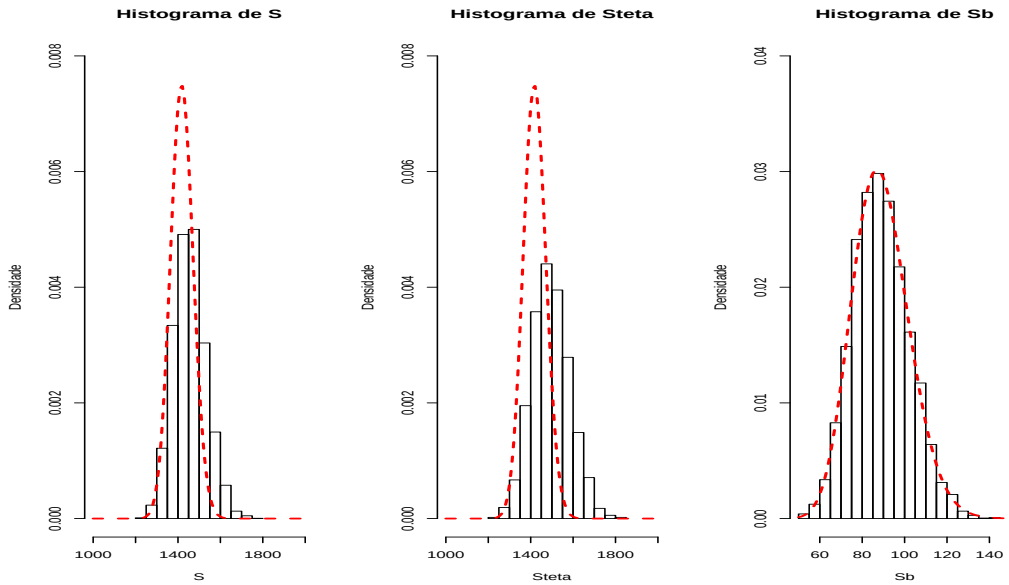


Figura 19 – Histograma estimado da distribuição a posteriori de $S(\boldsymbol{y}, \boldsymbol{\lambda}^i)$, $S_\theta(\boldsymbol{\theta}^i, \alpha_2{}^i)$ e $S_b(\boldsymbol{b}^i, \sigma^{2^i})$ sob o modelo Poisson-normal-gama e as respectivas distribuições de referência χ^2_{1419}, χ^2_{1419} e χ^2_{89}

Apesar de não fornecer um bom resultado para modelos mistos, na Tabela 2 tem-se o DIC (*"Deviance Information Criterion"*) e o respectivo valor de p_D, que representa

o número efetivo de parâmetros, para cada modelo, valores obtidos usando-se o software Winbugs. O menor valor do DIC indica o melhor modelo. Portanto, segundo o critério DIC o modelo Poisson-normal-gama é o que melhor se ajusta aos dados.

Tabela 2 – Valores obtidos, usando-se o software WinBUGS para o critério DIC

Modelo	DIC	p_D
Modelo Poisson	11597,700	4,002
Modelo Poisson-gama	4916,890	810,616
Modelo Poisson-normal	6047,630	86,845
Modelo Poisson-normal-gama	4838,770	558,346

Nas Figuras 20 e 21, é possível comparar a variância amostral com a variância capturada pelo modelo. Observa-se que o modelo Poisson-normal-gama, também, é o modelo que melhor captura a variabilidade dos dados. Na Figura 20, apresentam-se dois gráficos, em escalas diferentes, representando a variabilidade entre indivíduos que receberam o placebo capturada pelos diferentes modelos, para os diferentes tempos.

Portanto, a análise gráfica das quantidades pivotais, assim como os valores do DIC e a análise gráfica da variância capturada pelos modelos, indicam que o modelo combinado é o que melhor se ajusta aos dados.

a) b)

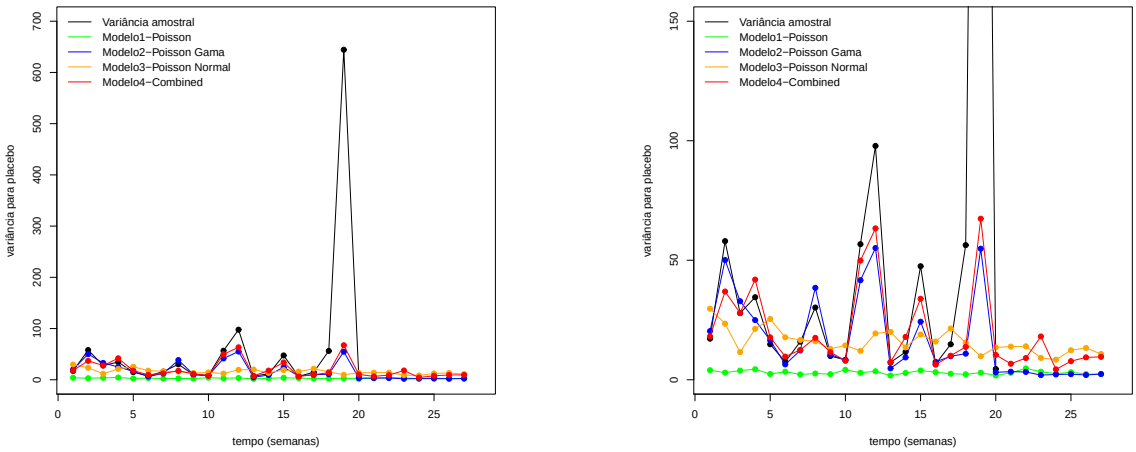


Figura 20 – a) Gráfico da variância entre indivíduos que receberam placebo capturada pelos diferentes modelos, para os diferentes tempos. b) Gráfico da variância entre indivíduos que receberam placebo capturada pelos diferentes modelos, para os diferentes tempos, com variância até 150

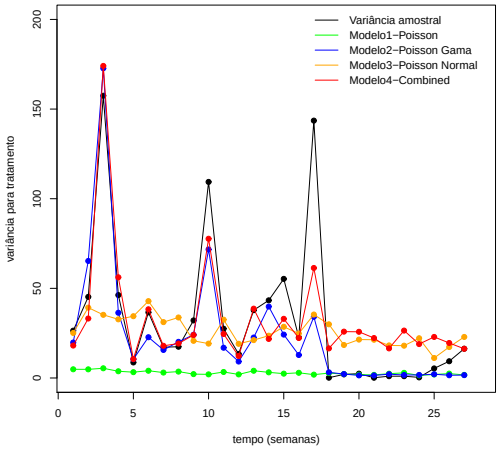


Figura 21 – Gráfico da variância entre indivíduos que receberam a nova droga capturada pelos modelos, para os diferentes tempos

A seguir, na Tabela 3, podem-se ver as estimativas dos parâmetros para os modelos Poisson, Poisson-gama, Poisson-normal e Poisson-normal-gama obtidas a partir das distribuições a posteriori. Observando-se os intervalos de credibilidade para a diferença entre os coeficientes angulares (ds) e o quociente entre os coeficientes angulares (rs) para cada um dos modelos, conclui-se que sob o modelo Poisson os tratamentos diferem entre si e sob os mo-

delos Poisson-gama, Poisson-normal e Poisson-normal-gama os tratamentos não diferem entre si. Entretanto, o intervalo de credibilidade para as medidas ds e rs no modelo selecionado, Poisson-normal-gama, é muito menor do que nos modelos Poisson-gama e Poisson-normal, mostrando uma melhor precisão nas estimativas.

Table 3 – Resumos a posteriori para os parâmetros dos modelos Poisson, Poisson-gama, Poisson-normal e combinado

Modelo Poisson				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,26600	0,04239	1,26500	(1,18200 ; 1,34700)
β_{01}	-0,01334	0,00432	-0,01335	(-0,02173 ; -0,00471)
β_{10}	1,45300	0,03800	1,45300	(1,37800 ; 1,52700)
β_{11}	-0,03282	0,00380	-0,03287	(-0,04024 ; -0,02535)
$ds = \beta_{11} - \beta_{01}$	-0,01948	0,00573	-0,01947	(-0,03083 ; -0,00825)
$rs = \frac{\beta_{11}}{\beta_{01}}$	2,89100	3,89800	2,45800	(1,41100 ; 6,99100)
Modelo Poisson-gama				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,26500	0,11400	1,26300	(1,04400 ; 1,49200)
β_{01}	-0,01277	0,01117	-0,01283	(-0,03453 ; 0,00931)
β_{10}	1,47500	0,10960	1,47400	(1,26400 ; 1,69200)
β_{11}	-0,03501	0,01019	-0,03493	(-0,05500 ; -0,01497)
α	0,52580	0,02554	0,52480	(0,47730 ; 0,57830)
β	1,90600	0,09253	1,90500	(1,72900 ; 2,09500)
$ds = \beta_{11} - \beta_{01}$	-0,02224	0,01492	-0,02221	(-0,05156 ; 0,00721)
$rs = \frac{\beta_{11}}{\beta_{01}}$	2,54600	64,53000	2,09400	(-21,16000 ; 27,86000)
Modelo Poisson-normal				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	0,80830	0,17000	0,80980	(0,46470 ; 1,13700)
β_{01}	-0,01430	0,00443	-0,01432	(-0,02312 ; -0,00561)
β_{10}	0,64140	0,16990	0,64310	(0,30990 ; 0,97280)
β_{11}	-0,01200	0,00431	-0,01202	(-0,02053 ; -0,00354)
σ^2	1,21100	0,20120	1,18900	(0,87670 ; 1,66300)
$ds = \beta_{11} - \beta_{01}$	0,00229	0,00622	0,00223	(-0,00986 ; 0,01467)
$rs = \frac{\beta_{11}}{\beta_{01}}$	0,95730	1,55600	0,84140	(0,22080 ; 2,42900)
Modelo combinado				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	0,89550	0,18290	0,89370	(0,53650 ; 1,25300)
β_{01}	-0,02484	0,00771	-0,02479	(-0,04007 ; -0,00956)
β_{10}	0,65750	0,18380	0,65900	(0,29270 ; 1,01400)
β_{11}	-0,01194	0,00757	-0,01197	(-0,02706 ; 0,00289)
α	2,46200	0,21130	2,45200	(2,08000 ; 2,90500)
β	0,40920	0,03496	0,40780	(0,34430 ; 0,48080)
σ^2	1,19100	0,20270	1,16800	(0,85490 ; 1,64900)
$ds = \beta_{11} - \beta_{01}$	0,01290	0,01077	0,01296	(-0,00828 ; 0,03417)
$rs = \frac{\beta_{11}}{\beta_{01}}$	0,56370	1,73000	0,48000	(-0,12830 ; 1,63300)

Como ilustração, foram construídos os intervalos de credibilidade a 95% de confiança para o número de ataques epilépticos para nove indivíduos sob o modelo combinado, para os dois tratamentos conforme mostrado na Figura 22. Verifica-se que a maior parte dos

pontos observados pertencem ao intervalo de credibilidade para ambos os tratamentos.

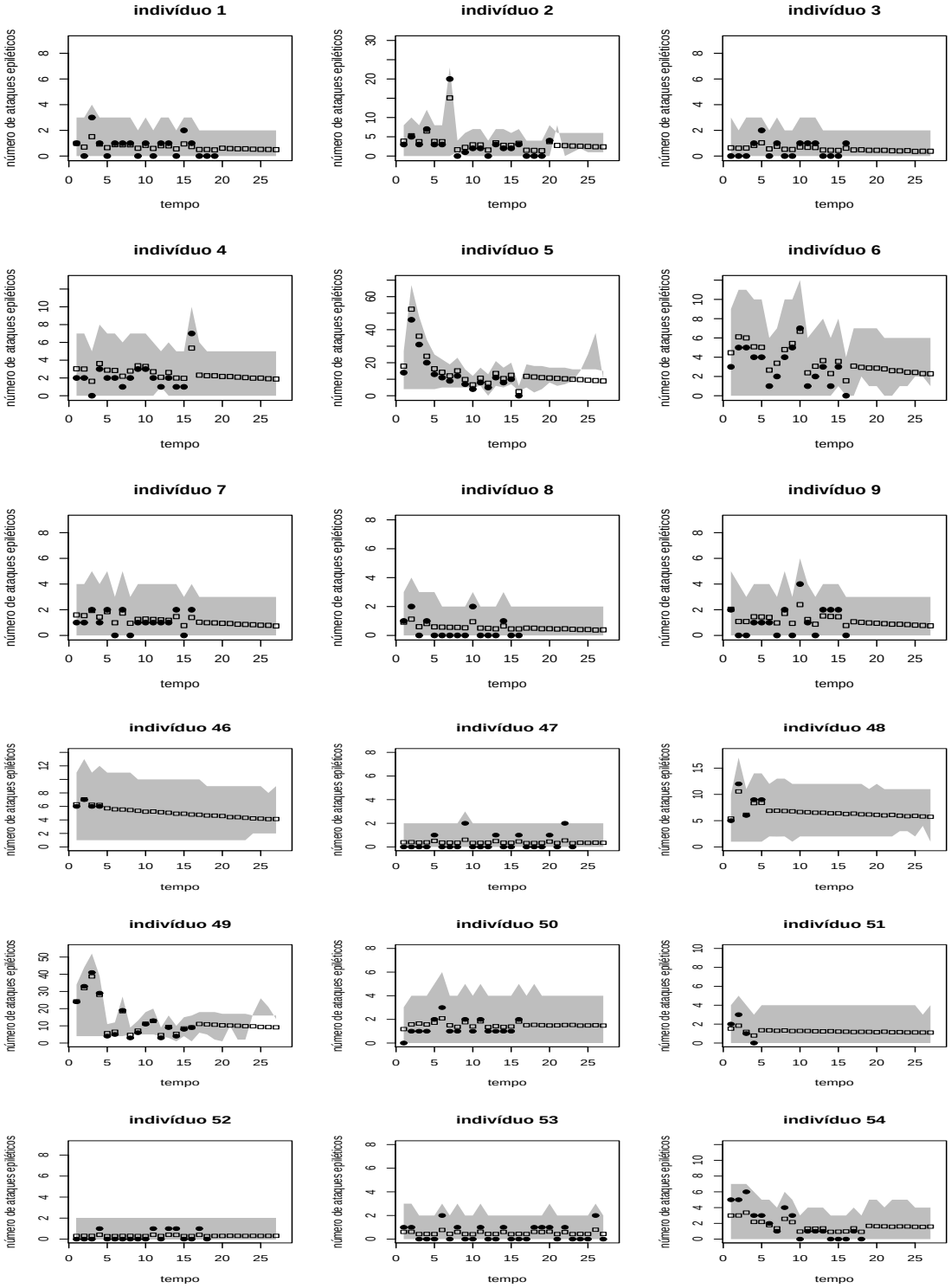
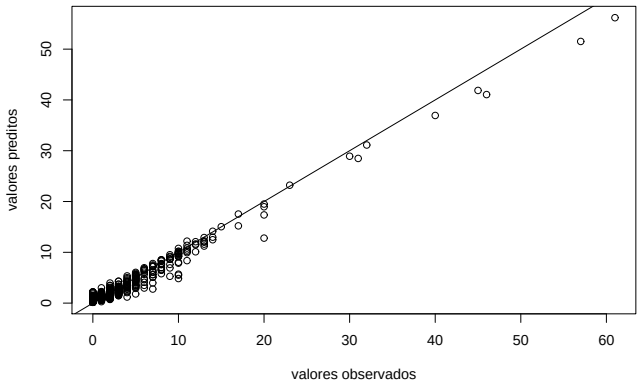


Figura 22 – Gráficos dos intervalos de credibilidade para o número de ataques epilépticos para indivíduos de 1 a 9 que receberam placebo e para indivíduos de 46 a 54 que receberam a nova droga anti-epiléptica, sob o modelo Poisson-normal-gama

Na Figura 23, têm-se os gráficos dos valores observados versus a média dos valores preditos pelo modelo Poisson-normal-gama para os indivíduos que foram submetidos ao placebo (a) e para os indivíduos que foram submetidos à nova droga anti-epiléptica (b). As Figuras 24 e 25 apresentam quatro gráficos, em diferentes escalas, dos valores observados versus a média dos valores preditos pelo modelo Poisson-normal-gama, respectivamente, para os indivíduos que foram submetidos ao placebo e à nova droga anti-epiléptica, para que se possa analisar com maior precisão a qualidade do ajuste do modelo aos dados para cada um dos tratamentos. Pelas Figuras 23, 24 e 25 observa-se um bom ajuste do modelo Poisson-normal-gama aos dados.

a)



b)

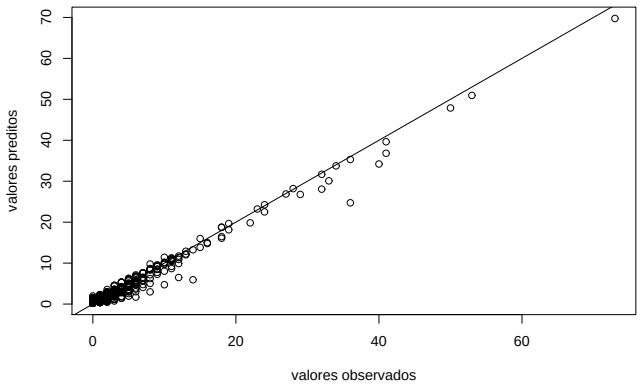


Figura 23 – a) Gráfico dos valores observados versus a média dos valores preditos pelo modelo Poisson-normal-gama para os indivíduos submetidos ao placebo. b) Gráfico dos valores observados versus a média dos valores preditos pelo modelo Poisson-normal-gama para os indivíduos submetidos à nova droga anti-epiléptica

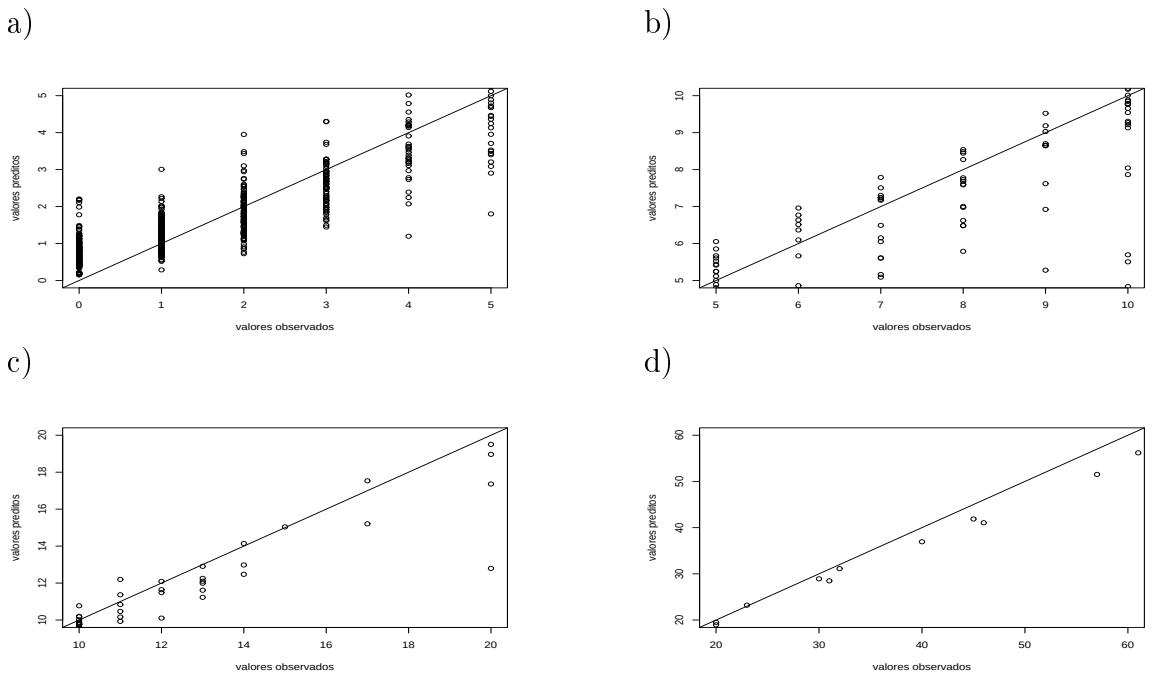


Figura 24 – Gráficos dos valores observados versus os valores preditos pelo modelo Poisson-normal-gama para os indivíduos submetidos ao placebo em diferentes intervalos: a) no intervalo de 0 a 5, b) no intervalo de 5 a 10, c) no intervalo de 10 a 20 e d) no intervalo de 20 a 60

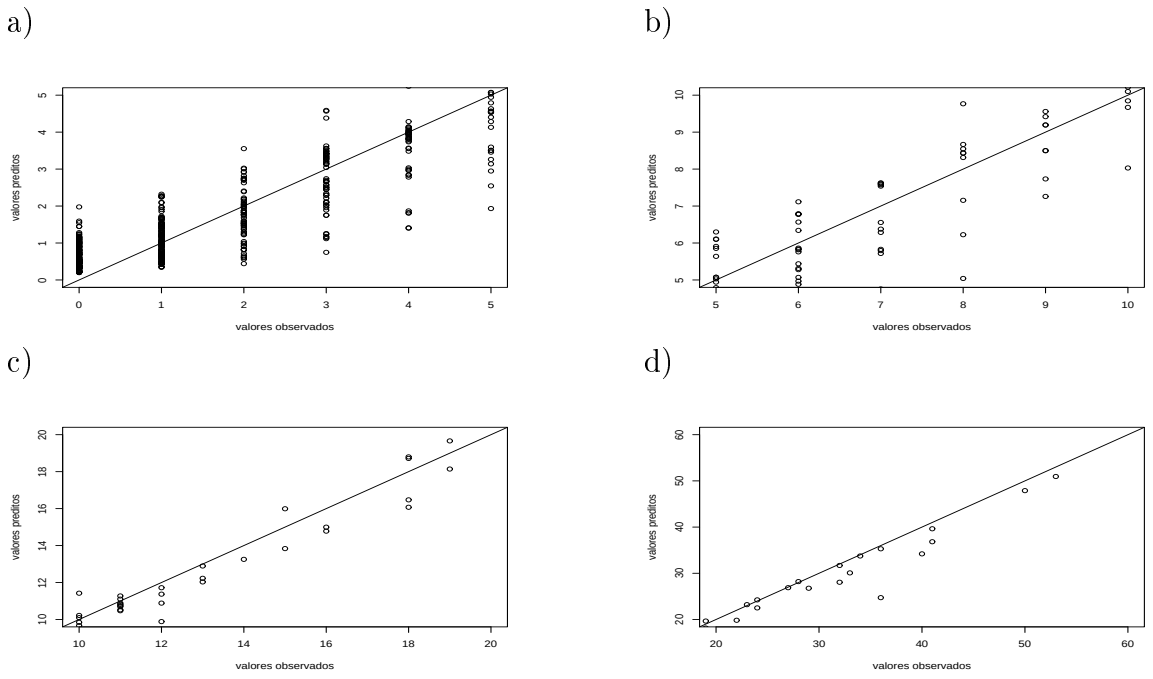


Figura 25 – Gráficos dos valores observados versus os valores preditos pelo modelo Poisson-normal-gama para os indivíduos submetidos à nova droga anti-epilética em diferentes intervalos: a) no intervalo de 0 a 5, b) no intervalo de 5 a 10, c) no intervalo de 10 a 20 e d) no intervalo de 20 a 60

4.2 Estudo de simulação

A seguir, serão apresentados os resultados referentes ao estudo de simulação para três conjutos de dados, simulados a partir dos modelos Poisson-gama, Poisson-normal e Poisson-normal-gama, com parâmetros $\beta_{00} = 1.2662$, $\beta_{01} = -0.0134$, $\beta_{10} = 1.4531$, $\beta_{11} = -0.0328$, $\alpha_1 = 2.4640$, $\alpha_2 = 0.4059$ e $\sigma^2 = 1.1289$, conforme seção 3.2.1.3. Para todos os modelos ajustados verificou-se a convergência da cadeia dos parâmetros e que as mesmas são não correlacionadas. Primeiramente, serão apresentadas as análises gráficas para seleção do modelo e a seguir uma tabela com os resumos a posteriori das estimativas dos parâmetros para cada modelo ajustado.

a. Análise dos dados simulados a partir do modelo Poisson-gama

Quando se observam os histogramas estimados das distribuições conjuntas a posteriori das medidas S, S_θ and S_b para os modelos Poisson, Poisson-gama e Poisson-normal-gama apresentados na Figura 26, verifica-se que o ajuste para os modelos Poisson-gama e Poisson-normal-gama é muito melhor do que para o modelo Poisson. Assim, pelo princípio da parcimônia, seleciona-se o modelo Poisson-gama. Pela Tabela 4, vê-se que os valores do DIC obtidos para os modelos Poisson-gama e Poisson-normal-gama são muito próximos e são valores inferiores ao do modelo Poisson, selecionando-se, assim, o modelo Poisson-gama. Observa-se, também, que os resumos a posteriori para os parâmetros do modelo Poisson-gama são similares aos valores utilizados para a simulação e que sob os modelos Poisson e Poisson-normal-gama também não há muita diferença entre esses valores. Vale ressaltar que o valor estimado para σ^2 no modelo Poisson-normal-gama é próximo a zero, reforçando, ainda mais a escolha feita.

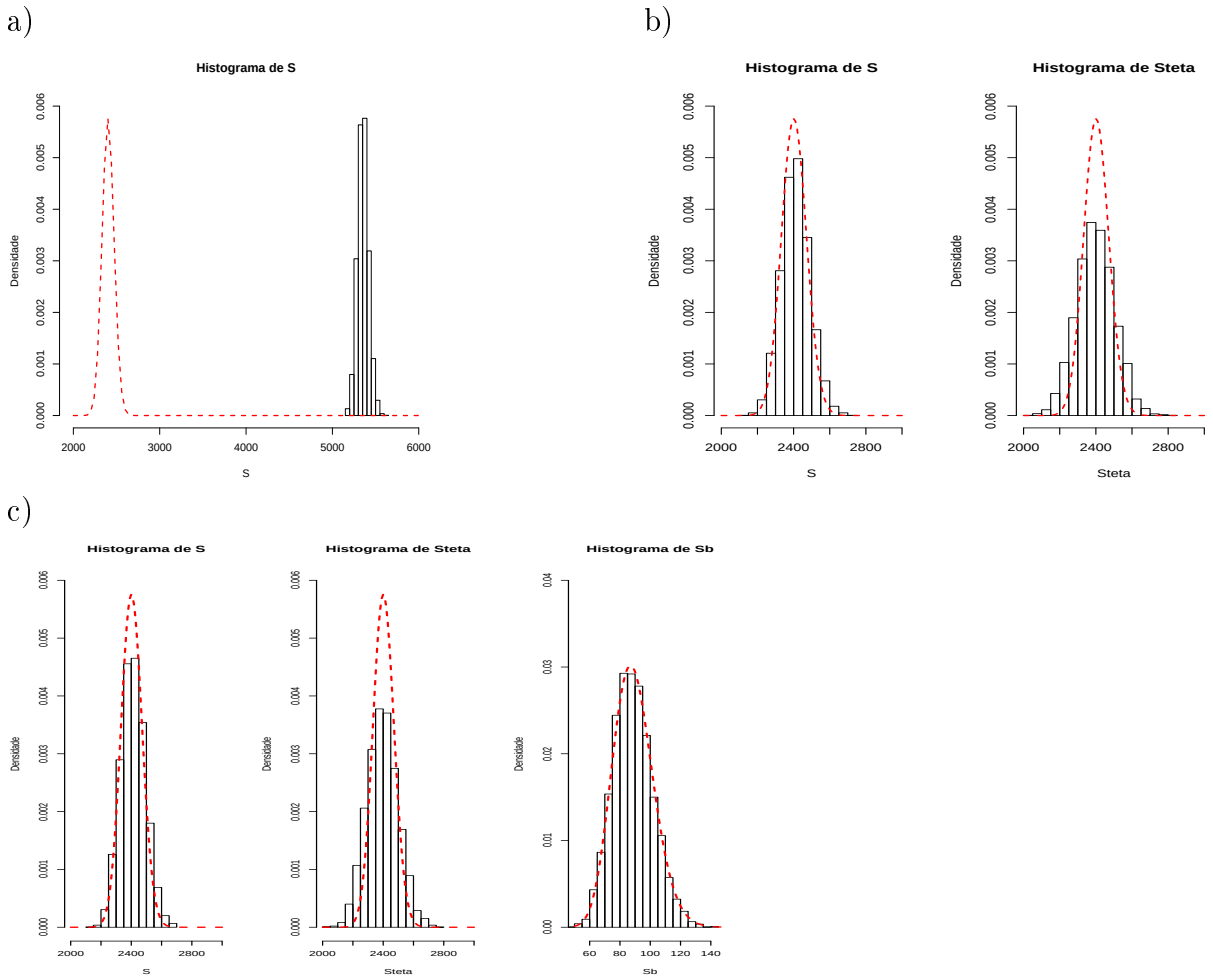


Figura 26 – a) Histogramas estimados da distribuição a posteriori das medidas S e sua respectiva distribuição de referência, sob modelo Poisson. b) Histogramas estimados da distribuição a posteriori das medidas S e S_θ e suas respectivas distribuições de referência, sob modelo Poisson-gama. c) Histogramas estimados da distribuição a posteriori das medidas S, S_θ e S_b e suas respectivas distribuições de referência, sob modelo Poisson-normal-gama

Tabela 4 – Resultados a posteriori considerando os modelos Poisson, Poisson-gama e Poisson-normal-gama para os dados simulados a partir do modelo Poisson-gama

Modelo Poisson				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,2180	0,0337	1,2180	(1,1510 ; 1,2840)
β_{01}	-0,0115	0,0022	-0,0115	(-0,01573 ; -0,0071)
β_{10}	1,4470	0,0328	1,4470	(1,3830 ; 1,5129)
β_{11}	-0,0327	0,0023	-0,0327	(-0,0373 ; -0,0282)
DIC=111962		$p_D = 4,018$		
Modelo Poisson-gama				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,2200	0,0519	12200	(1,1180 ; 1,3210)
β_{01}	-0,0116	0,0033	-0,0116	(-0,0179 ; -0,0051)
β_{10}	1,4480	0,0518	1,4480	(1,3460 ; 1,5500)
β_{11}	-0,0328	0,0034	-0,0327	(-0,0394 ; -0,0263)
α_1	2,2900	0,1317	2,2870	(2,0460 ; 2,5610)
α_2	0,4381	0,0251	0,4373	(0,3906 ; 0,4889)
DIC=9539,48		$p_D = 1105,13$		
Modelo Poisson-normal-gama				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,2190	0,0521	1,2190	(1,1170 ; 1,3200)
β_{01}	-0,0116	0,0033	-0,0116	(-0,0180 ; -0,0050)
β_{10}	1,4480	0,0518	1,4480	(1,3460 ; 1,5490)
β_{11}	-0,0328	0,0034	-0,0328	(-0,0395 ; -0,0263)
α_1	2,2970	0,1325	2,2910	(2,0510 ; 2,5740)
α_2	0,4368	0,0251	0,4364	(0,3885 ; 0,4877)
σ^2	0,0020	0,0025	0,0011	(6,29E-5 ; 0,0091)
DIC=9549,39		$p_D = 1105,37$		

b. Análise dos dados simulados a partir do modelo Poisson-normal

Quando se observam os histogramas estimados da distribuição a posteriori conjunta das medidas S, S_θ and S_b para os modelos Poisson, Poisson-normal e Poisson-normal-gama apresentados na Figura 27, verifica-se que o modelo Poisson-normal é o que melhor se ajusta aos dados, como esperado. Pela Tabela 5, vê-se que o valor do DIC para o modelo Poisson-normal é menor do que para os outros dois modelos, confirmando o resultado obtido na análise gráfica das quantidades pivotais. Observa-se, também, que os resultados a posteriori para os parâmetros do modelo Poisson-normal são similares aos valores considerados para a simulação e que sob o modelo Poisson-normal-gama não há muita diferença entre esses valores, o mesmo não ocorrendo para os interceptos do modelo Poisson.

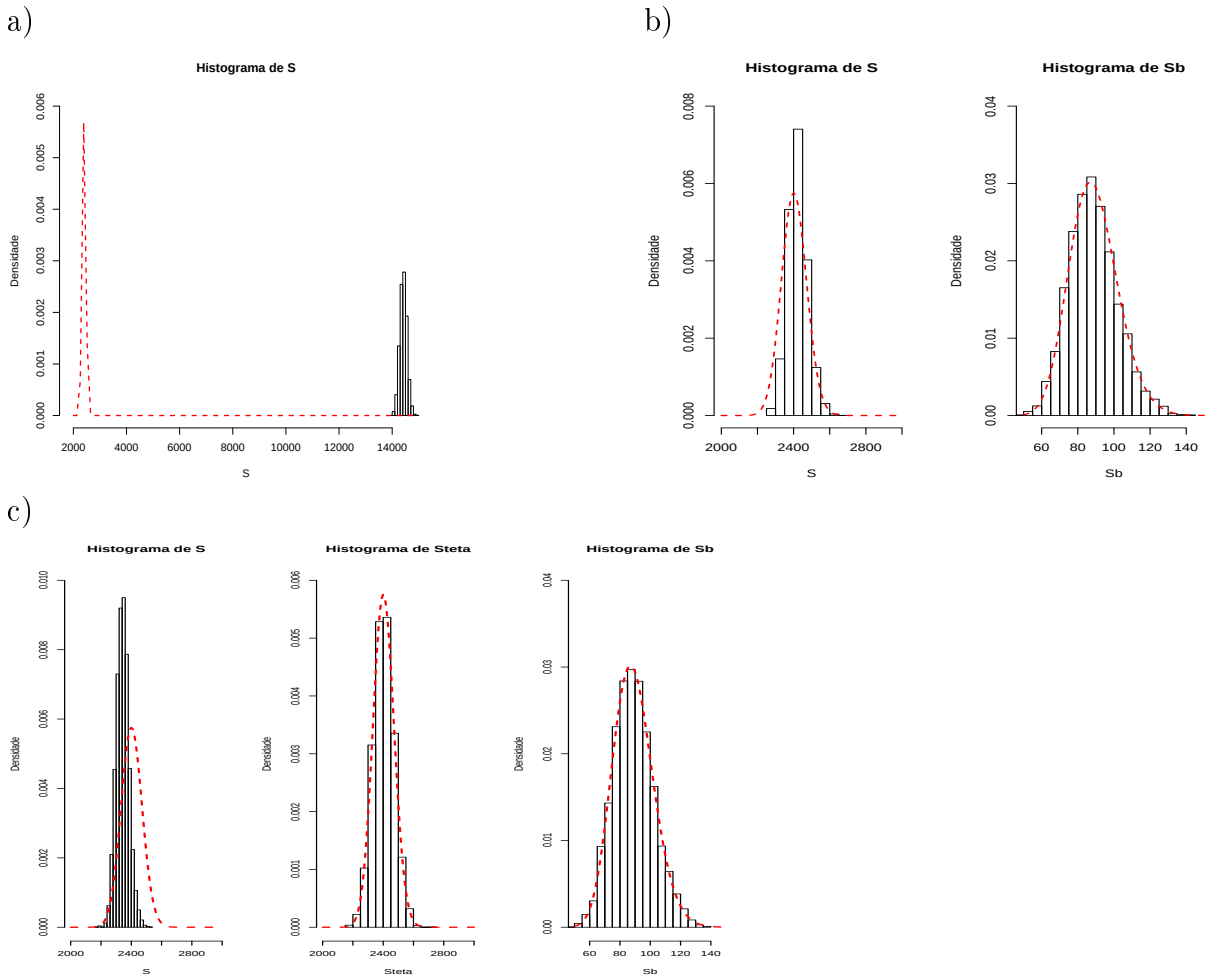


Figura 27 – a) Histogramas estimados da distribuição a posteriori das medidas S e sua respectiva distribuição de referência, sob modelo Poisson. b) Histogramas estimados da distribuição a posteriori das medidas S e S_b e suas respectivas distribuições de referência, sob modelo Poisson-normal. c) Histogramas estimados da distribuição a posteriori das medidas S, S_θ e S_b e suas respectivas distribuições de referência, sob modelo Poisson-normal-gama

Tabela 5 – Resultados a posteriori considerando os modelos Poisson, Poisson-normal e Poisson-normal-gama para os dados simulados a partir do modelo Poisson-normal

		Modelo Poisson		
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,7810	0,0253	1,7810	(1,7310 ; 1,8310)
β_{01}	-0,0139	0,0016	-0,0139	(-0,0171 ; -0,0106)
β_{10}	1,8640	0,0259	1,8640	(1,8120 ; 1,9140)
β_{11}	-0,0319	0,0018	-0,0319	(-0,0354 ; -0,0283)
DIC=18800,8		$p_D = 3,982$		
		Modelo Poisson-normal		
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,2990	0,1599	1,3010	(0,9710 ; 1,6050)
β_{01}	-0,0138	0,0017	-0,0138	(-0,0172 ; -0,0105)
β_{10}	1,3530	0,1573	1,3530	(1,0460 ; 1,6740)
β_{11}	-0,0319	0,0018	-0,0319	(-0,0355; -0,0283)
σ^2	1,0880	0,1735	1,0700	(0,8027 ; 1,4770)
DIC=7843,95		$p_D = 88,177$		
		Modelo Poisson-normal-gama		
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,2920	0,1595	1,2930	(0,9770 ; 1,5980)
β_{01}	-0,0139	0,0017	-0,0138	(-0,0172 ; -0,0106)
β_{10}	1,3560	0,1613	1,3560	(1,0390 ; 1,6720)
β_{11}	-0,0319	0,0018	-0,0319	(-0,0354 ; -0,0284)
α_1	$1,5150 \times 10^6$	$3,9990 \times 10^6$	155900,0	(731,6 ; $1,3030 \times 10^7$)
α_2	$1,6350 \times 10^{-4}$	$4,3940 \times 10^{-4}$	$6,4180 \times 10^{-6}$	($7,6940 \times 10^{-8}$; 0,0014)
σ^2	1,09	0,1742	1,073	(0,7983 ; 1,491)
DIC=9040,06		$p_D = 93,574$		

c. Análise dos dados simulados a partir do modelo Poisson-normal-gama

Quando se observam os historamas estimados das distribuições a posteriori conjunta das medidas S, S_θ e S_b para os modelos Poisson, Poisson-gama, Poisson-normal e Poisson-normal-gama apresentados na Figura 28, verifica-se que o modelo combinado é o que melhor se ajusta aos dados simulados, como esperado. Pela Tabela 6, vê-se que o valor do DIC é menor para o modelo Poisson-normal-gama, confirmando o resultado obtido através da análise gráfica das quantidades pivotais. Observa-se, também, que o resultado a posteriori para os parâmetros dos modelos Poisson-normal-gama e Poisson-normal são similares aos valores considerados para a simulação, o mesmo não ocorrendo quando são considerados os modelos Poisson e Poisson-gama.

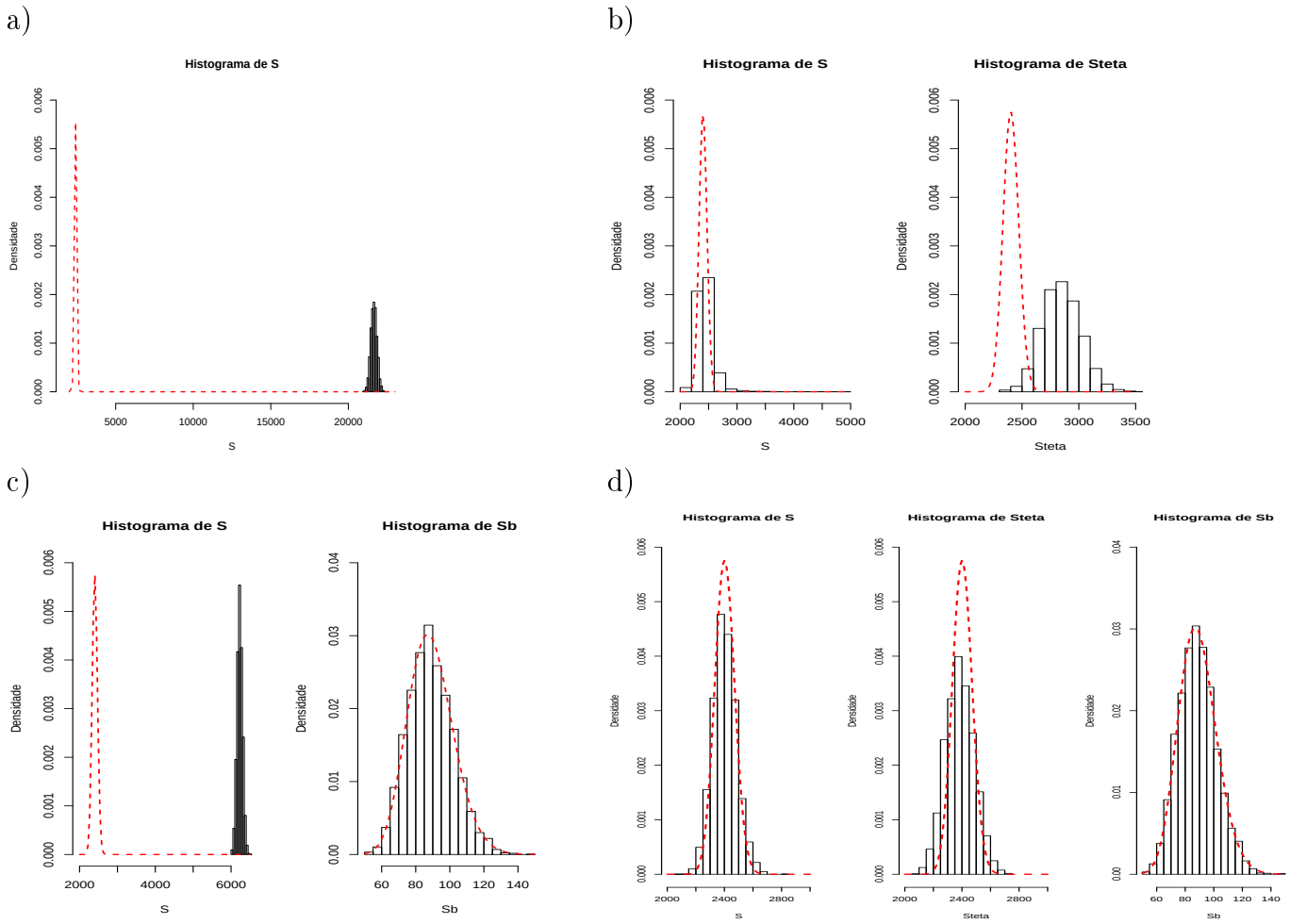


Figura 28 – a) Histogramas estimados da distribuição a posteriori das medidas S e sua respectiva distribuição de referência, sob modelo Poisson. b) Histogramas estimados da distribuição a posteriori das medidas S e S_θ e suas respectivas distribuições de referência, sob modelo Poisson-gama. c) Histogramas estimados da distribuição a posteriori das medidas S e S_b e suas respectivas distribuições de referência, sob modelo Poisson-normal. d) Histogramas estimados da distribuição a posteriori das medidas S, S_θ e S_b e suas respectivas distribuições de referência, sob modelo Poisson-normal-gama

Tabela 6 – Resultados a posteriori considerando os modelos Poisson, Poisson-gama, Poisson-normal e Poisson-normal-gama para os dados simulados a partir do modelo Poisson-normal-gama

Modelo Poisson				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,7810	0,0254	1,7810	(1,7310 ; 1,8310)
β_{01}	-0,0144	0,0017	-0,0144	(-0,0178 ; -0,0111)
β_{10}	1,8700	0,0265	1,8700	(1,8181 ; 1,9220)
β_{11}	-0,0354	0,0019	-0,0355	(-0,0391 ; -0,0317)
DIC=21576,8		$p_D = 4,007$		
Modelo Poisson-gama				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,7770	0,0764	1,7770	(1,6290 ; 1,9250)
β_{01}	-0,0142	0,0048	-0,0142	(-0,0234 ; -0,0049)
β_{10}	1,8810	0,0775	1,8790	(1,7360 ; 2,0380)
β_{11}	-0,0361	0,0049	-0,0361	(-0,0459 ; -0,0267)
α_1	0,7045	0,0247	0,7040	(0,6575 ; 0,7541)
α_2	1,4210	0,0498	1,4210	(1,3260 ; 1,5210)
DIC=9800,01		$p_D = 1546,28$		
Modelo Poisson-normal				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,3000	0,1549	1,302	(0,9966 ; 1,6010)
β_{01}	-0,0144	0,0017	-0,0144	(-0,0177 ; -0,0112)
β_{10}	1,3860	0,1611	1,3830	(1,0730 ; 1,6970)
β_{11}	-0,0354	0,0019	-0,0352	(-0,0390 ; -0,0317)
σ^2	1,0830	0,1730	1,0670	(0,7961 ; 1,4670)
DIC=12245,9		$p_D = 89,175$		
Modelo Poisson-normal-gama				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,2950	0,1655	1,2990	(0,9588 ; 1,6150)
β_{01}	-0,0140	0,0032	-0,0140	(-0,0203 ; -0,0078)
β_{10}	1,3830	0,1577	1,3850	(1,0780 ; 1,6950)
β_{11}	-0,0355	0,0034	-0,0355	(-0,0422 ; -0,0289)
α_1	2,5720	0,1440	2,5700	(2,3010 ; 2,8670)
α_2	0,3900	0,0218	0,3890	(0,3489 ; 0,4346)
σ^2	1,0710	0,1736	1,055	(0,7815 ; 1,4610)
DIC=9542,96		$p_D = 1110,56$		

4.3 Inferência frequentista via Dclone e estudo da estimabilidade dos modelos

Nesta seção, será apresentado o estudo da estimabilidade dos modelos Poisson-gama, Poisson-normal e combinado, usando-se o pacote dclone e os dados reais para o número de ataques epilépticos. Além disso, será feita uma comparação dos resultados baseados em estimativas de máxima verossimilhança para esses modelos. Vangeneugden et. al. (2011) obtiveram as estimativas de máxima verossimilhança usando o procedimento NLMIXED do

SAS, o qual disponibiliza diferentes métodos para aproximação numérica de integrais, entre eles a quadratura Gaussiana adaptativa, que foi utilizado no referido artigo quando necessário. Esses resultados serão comparados aos resultados obtidos pelo pacote dclone do ambiente R. Para o modelo combinado, também, serão apresentados os resultados da implementação no ambinete R utilizando como métodos de integração numérica o método de Laplace e a quadratura Gaussiana adaptativa. Os programas encontram-se no Apêndice B.

Modelo 1: Poisson

Para o estudo da estimabilidade do modelo Poisson, foi considerada a análise com k clones, $k = 1, 10, 25, 50$. Na Figura 29, são apresentadas as medidas descritivas a posteriori para cada parâmetro em cada passo, ou seja, para cada número determinado de clones em que os pontos representam as médias. Observa-se que para todos os parâmetros, eles convergem para a estimativa de máxima verossimilhança, as linhas verticais representam os erros padrões, que vão diminuindo com o aumento do número de clones.

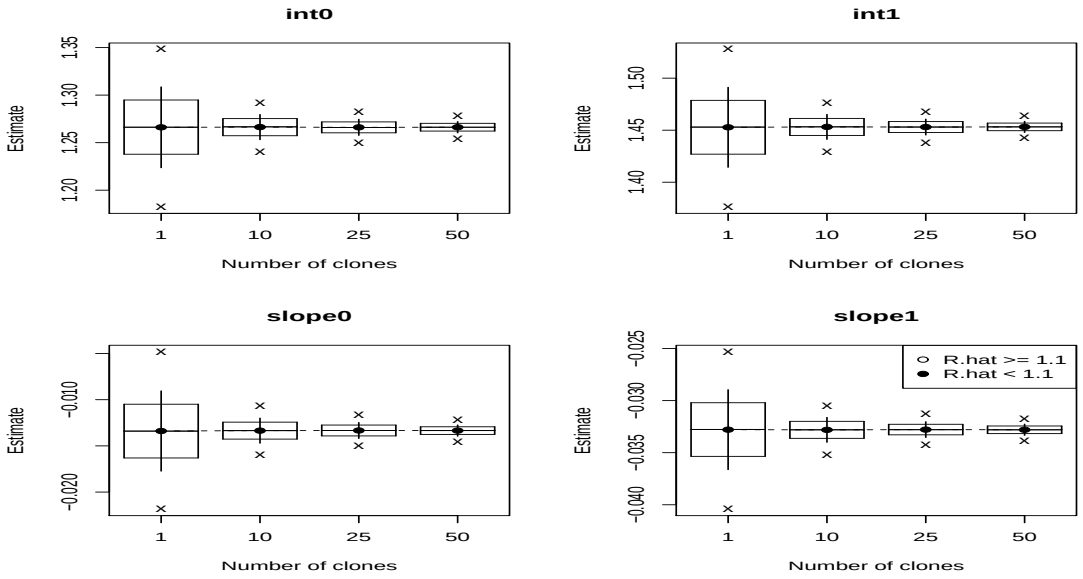


Figura 29 – Estatísticas para o modelo Poisson para 1, 10, 25 e 50 clones. Os pontos representam as médias e as linhas verticais representam os erros padrões. Também, são representados os quantis $2,5\%$, 50% e $97,5\%$

A Figura 30 apresenta o diagnóstico de convergência para os dados clonados baseado no modelo Poisson. Nela, são plotados gráficos do logaritmo da variância a posteriori para cada um dos parâmetros, que decrescem linearmente, garantindo assim a convergência

assintótica e com isso a estimabilidade do modelo.

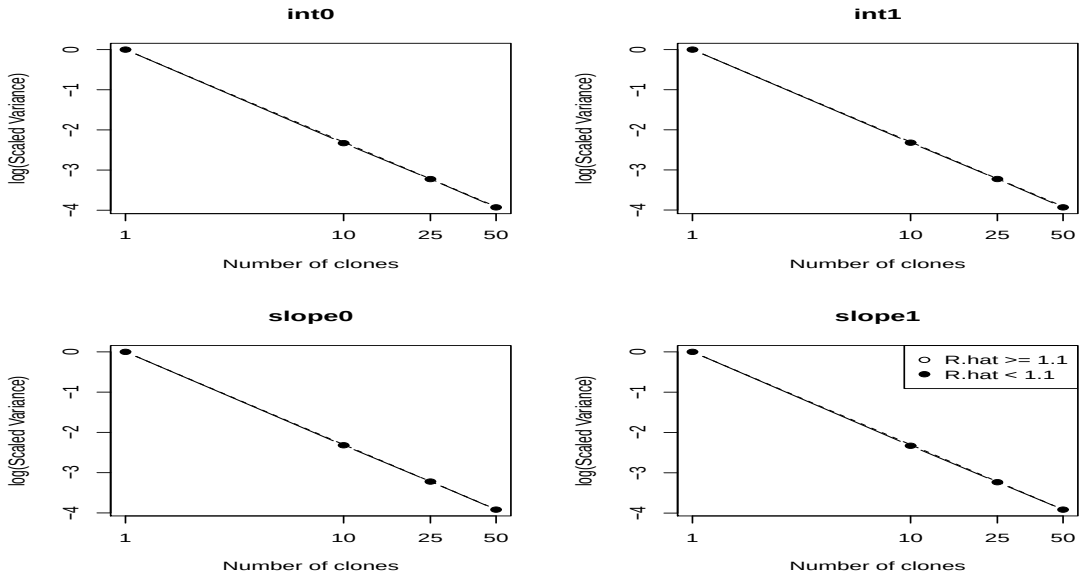


Figura 30 – Diagnóstico de convergência para os dados clonados sob o modelo Poisson para 1, 10, 25 e 50 clones

Na Tabela 7, observa-se que os resultados obtidos para as estimativas de máxima verossimilhança para os parâmetros do modelo Poisson, usando-se o software SAS e o pacote dclone do ambiente R são muito similares, apresentando diferenças a partir da terceira casa decimal, assim como os erros padrões.

Tabela 7 – Resultados das estimativas de máxima verossimilhança para os parâmetros do modelo Poisson, usando-se o pacote dclone do ambiente R e o procedimente NLMIXED do SAS

	Resultado via dclone		Resultado via NLMIXED	
Parâmetro	Estimativa	erro padrão	Estimativa	erro padrão
β_{00}	1,26623	0,042198	1,2662	0,0424
β_{01}	-0,01335	0,004333	-0,0134	0,0043
β_{10}	1,45316	0,038055	1,4531	0,0383
β_{11}	-0,03280	0,003838	-0,0328	0,0038

Modelo 2: Poisson-gama

Para o estudo da estimabilidade do modelo Poisson-gama foi considerada a análise com k clones, $k = 1, 10, 25, 50$. Na Figura 31, são apresentadas as medidas descritivas a posteriori para cada parâmetro em cada passo, ou seja, para cada número determinado de clones. Os pontos representam as médias, observa-se que para todos os parâmetros eles convergem para a estimativa de máxima verossimilhança, as linhas verticais representam os erros padrões, que vão diminuindo com o aumento do número de clones.

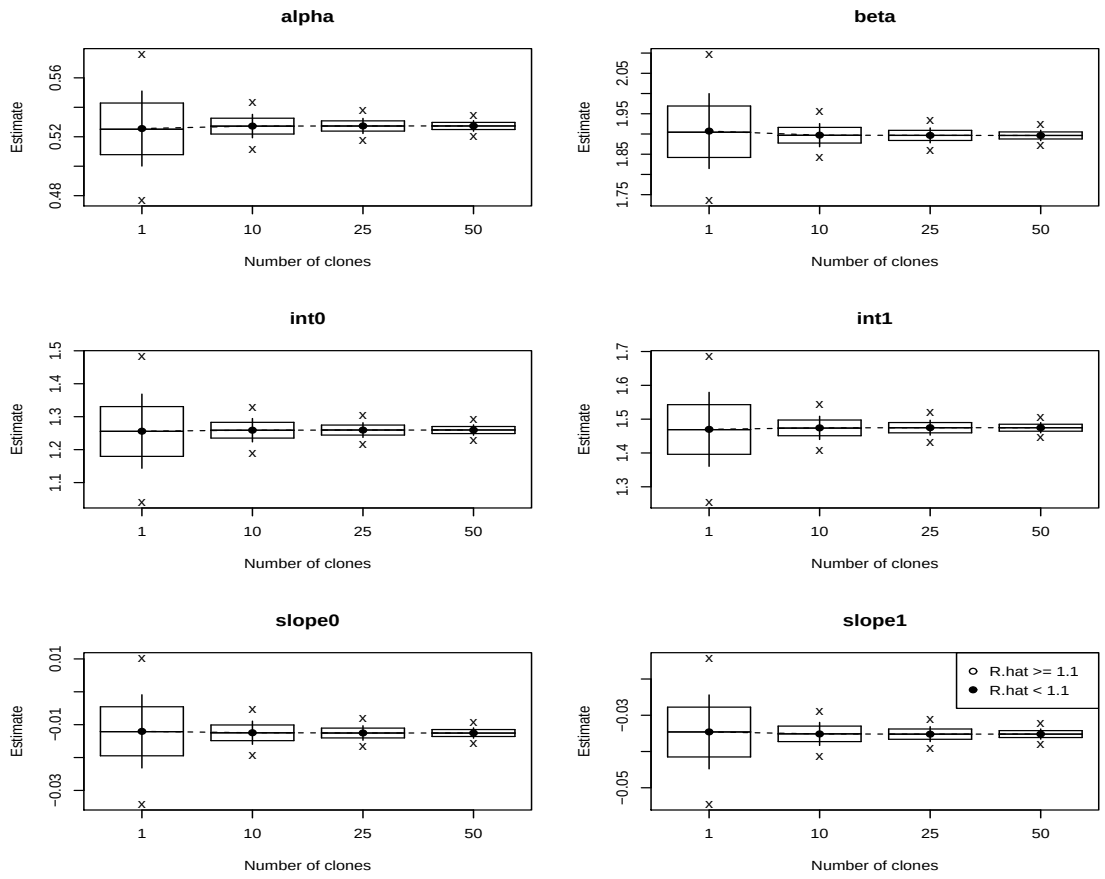


Figura 31 – Estatísticas para o modelo Poisson-gama para 1, 10, 25 e 50 clones. Os pontos representam as médias e as linhas verticais representam os erros padrões. Também, são representados os quantis $2,5\%$, 50% e $97,5\%$

A Figura 32 apresenta o diagnóstico de convergência para os dados clonados baseado no modelo Poisson-gama. Nela, são plotados gráficos do logaritmo da variância a posteriori para cada um dos parâmetros, que decrescem linearmente, garantindo assim a convergência assintótica e com isso a estimabilidade do modelo.

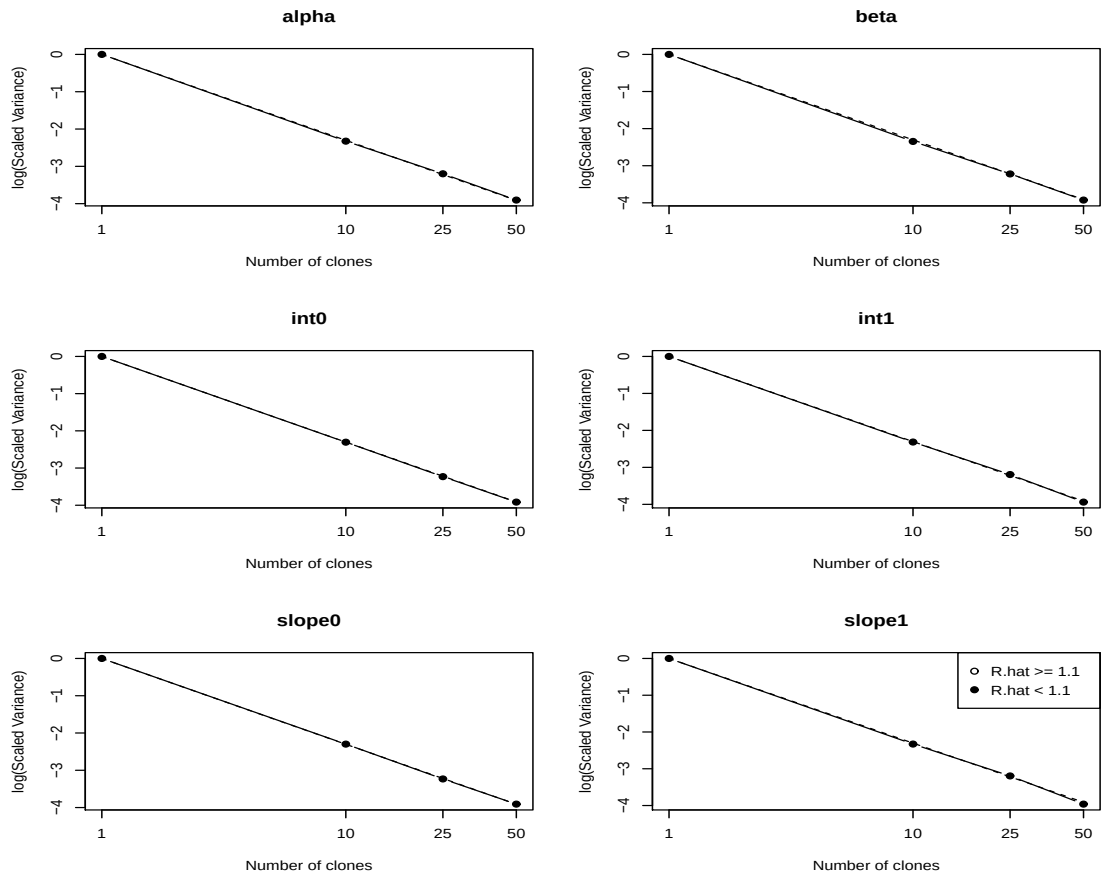


Figura 32 – Diagnóstico de convergência para os dados clonados sob o modelo Poisson-gama para 1, 10, 25 e 50 clones

Na Tabela 8, observa-se que os resultados obtidos para as estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-gama, usando-se o software SAS e o pacote dclone do ambiente R são muito similares, apresentando diferenças a partir da terceira casa decimal, assim como os erros padrões.

Tabela 8 – Resultados das estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-gama, usando-se o pacote dclone do ambiente R e o procedimente NLMIXED do SAS

	Resultado via dclone		Resultado via NLMIXED	
Parâmetro	Estimativa	erro padrão	Estimativa	erro padrão
β_{00}	1,25947	0,112427	1,2594	0,1119
β_{01}	-0,01254	0,011163	-0,0126	0,0111
β_{10}	1,47464	0,108016	1,4750	0,1093
β_{11}	-0,03519	0,009953	-0,0352	0,0101
α_1	0,52734	0,025586	0,5274	0,0255
α_2	1,89639	0,092018	1,8961	0,0918

Modelo 3: Poisson-normal

Para o estudo da estimabilidade do modelo Poisson-Normal foi considerada a análise com k clones, $k = 1, 10, 25, 50$. Na Figura 33, são apresentadas as medidas descritivas a posteriori para cada parâmetro em cada passo, ou seja, para cada número determinado de clones. Os pontos representam as médias, observa-se que para todos os parâmetros eles convergem para a estimativa de máxima verossimilhança, as linhas verticais representam os erros padrões, que vão diminuindo com o aumento do número de clones.

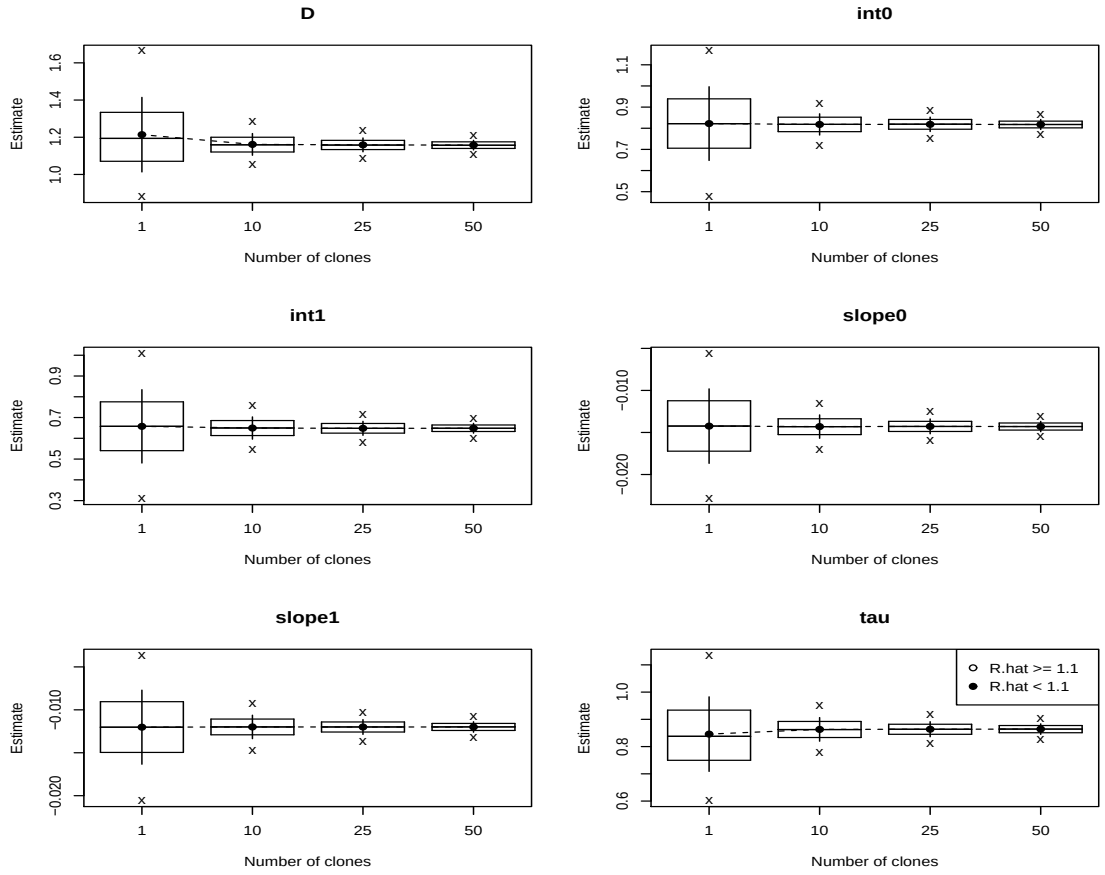


Figura 33 – Estatísticas para o modelo Poisson-Normal para 1, 10, 25 e 50 clones. Os pontos representam as médias e as linhas verticais representam os erros padrões. Também, são representados os quantis $2,5\%$, 50% e $97,5\%$

A Figura 34 apresenta o diagnóstico de convergência para os dados clonados baseado no modelo Poisson-gama. Nela, são plotados gráficos do logaritmo da variância a posteriori para cada um dos parâmetros, que decrescem de forma aproximadamente linear, garantindo assim a convergência assintótica e com isso a estimabilidade do modelo.

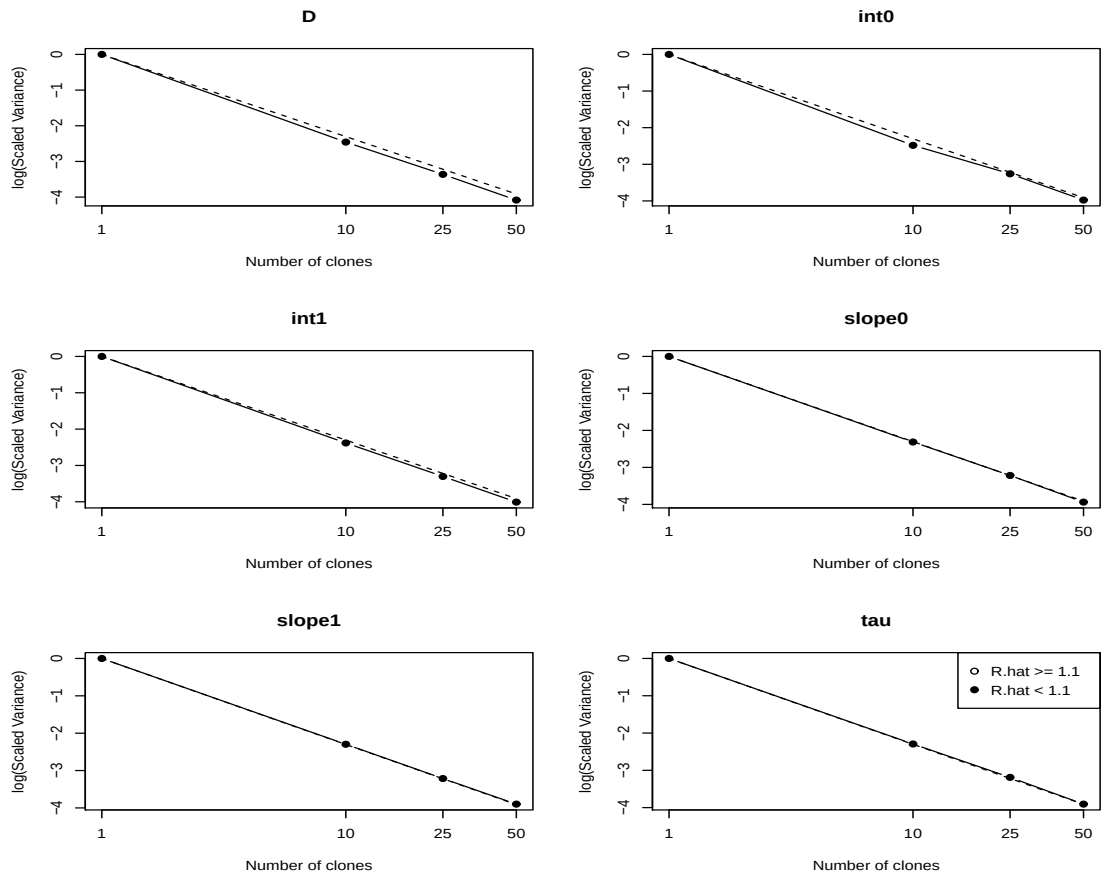


Figura 34 – Diagnóstico de convergência para os dados clonados sob o modelo Poisson-normal para 1, 10, 25 e 50 clones

Na Tabela 9, observa-se que os resultados obtidos para as estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-Normal, usando-se o software SAS e o pacote dclone do ambiente R são muito similares, apresentando diferenças a partir da terceira casa decimal, assim como os erros padrões.

Tabela 9 – Resultados das estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-Normal, usando-se o pacote dclone do ambiente R e o procedimente NLMIXED do SAS

	Resultado via dclone		Resultado via NLMIXED	
Parâmetro	Estimativa	erro padrão	Estimativa	erro padrão
β_{00}	0,8178	0,1689	0,8179	0,1677
β_{01}	-0,0143	0,0044	-0,0143	0,0044
β_{10}	0,6484	0,1685	0,6475	0,1701
β_{11}	-0,0120	0,0043	-0,0120	0,0043
σ^2	1,1577	0,1841	1,1568	0,1844

Modelo 4: Poisson-normal-gama

Para o estudo da estimabilidade do modelo Poisson-Normal-gama foi considerada a análise com k clones, $k = 1, 10, 25, 50$. Na Figura 35, são apresentadas as medidas descritivas a posteriori para cada parâmetro em cada passo, ou seja, para cada número determinado de clones. Os pontos representam as médias, observa-se que para todos os parâmetros eles convergem para a estimativa de máxima verossimilhança, as linhas verticais representam os erros padrões, que vão diminuindo com o aumento do número de clones.

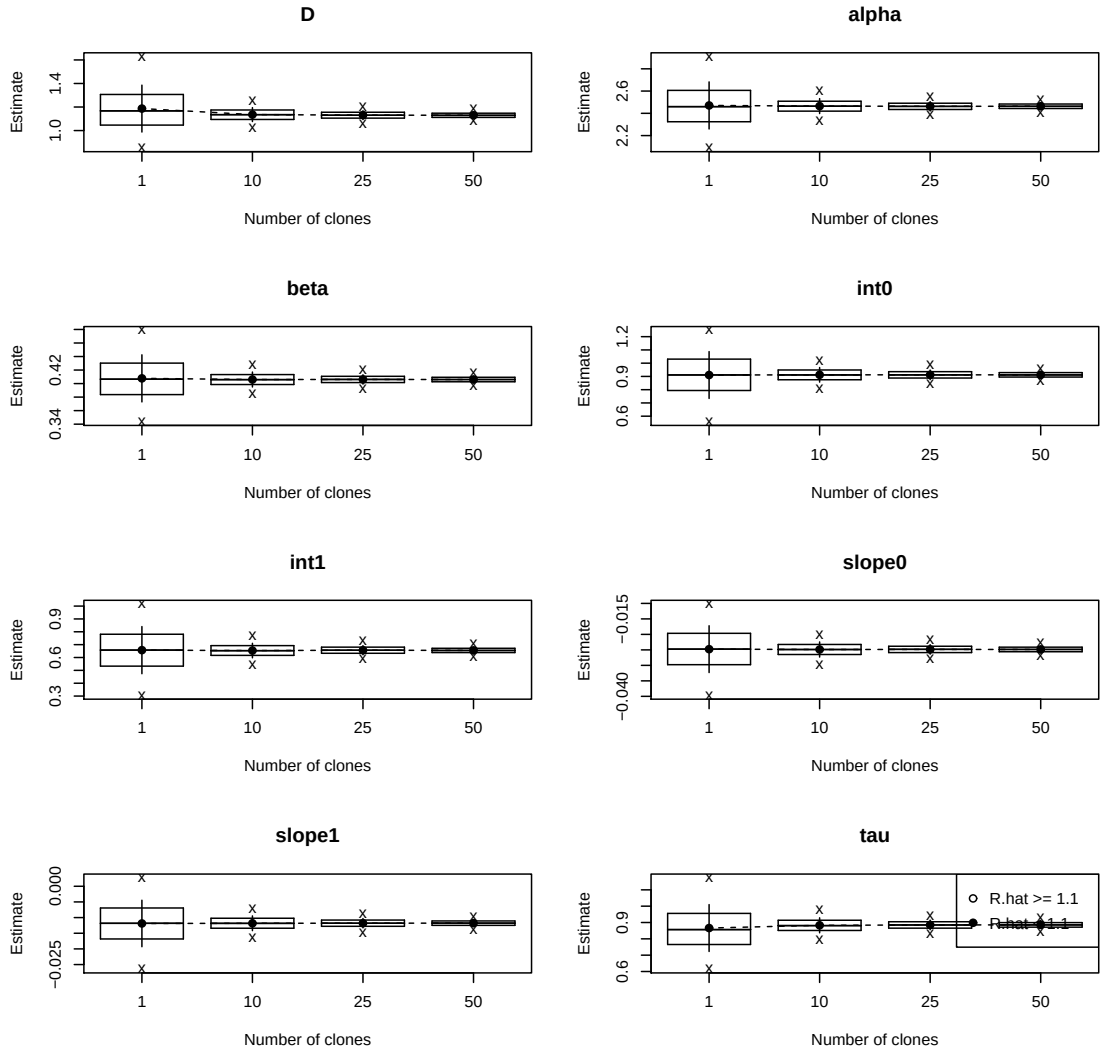


Figura 35 – Estatísticas para o modelo Poisson-Normal-gama para 1, 10, 25 e 50 clones. Os pontos representam as médias e as linhas verticais representam os erros padrões. Também, são representados os quantis $2,5\%$, 50% e $97,5\%$

A Figura 36 apresenta o diagnóstico de convergência para os dados clonados

baseado no modelo Poisson-Normal-gama. Nela, são plotados gráficos do logaritmo da variância a posteriori para cada um dos parâmetros, que decrescem de forma aproximadamente linear, garantindo assim a convergência assintótica e com isso a estimabilidade do modelo.

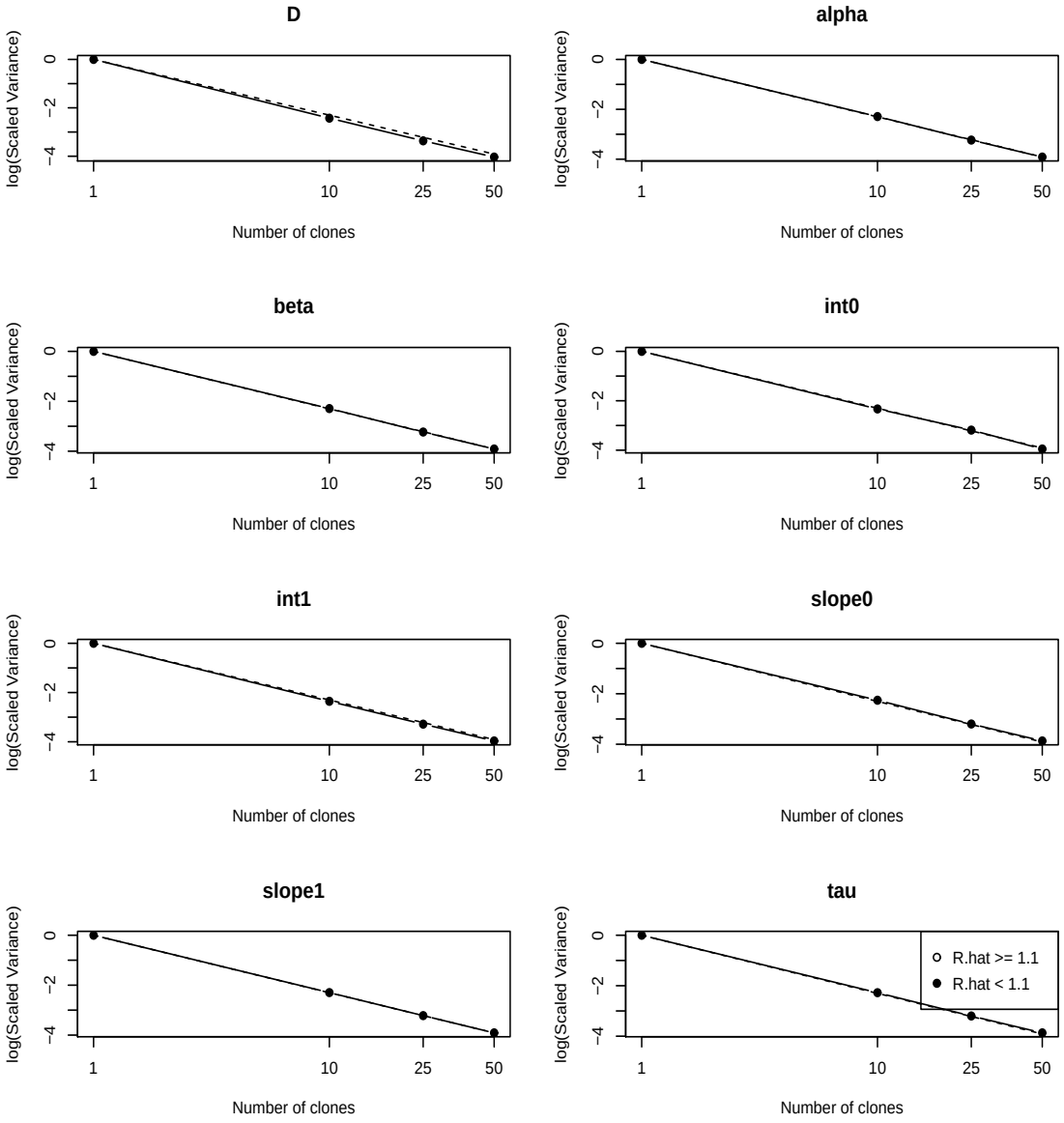


Figura 36 – Diagnóstico de convergência para os dados clonados sob o modelo Poisson-normal-gama para 1, 10, 25 e 50 clones

Na Tabela 10, observa-se que os resultados obtidos para as estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-Normal-gama, usando-se o software SAS e o pacote dclone do ambiente R são muito similares, apresentando diferenças a

partir da terceira casa decimal, assim como os erros padrões.

Tabela 10 – Resultados das estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-Normal-gama (combinado), usando-se o pacote dclone do ambiente R e o procedimente NLMIXED do SAS

	Resultado via dclone		Resultado via NLMIXED	
Parâmetro	Estimativa	erro padrão	Estimativa	erro padrão
β_{00}	0,9118	0,1734	0,9112	0,1755
β_{01}	-0,0249	0,0077	-0,0248	0,0077
β_{10}	0,6550	0,1786	0,6555	0,1782
β_{11}	-0,0118	0,0077	-0,0118	0,0075
α_1	2,4635	0,2113	2,4640	0,2113
α_2	0,4060	0,0348	0,4059	0,0348
σ^2	1,1299	0,1875	1,1289	0,1850

Na Tabela 11, observa-se que os resultados obtidos para as estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-normal-gama pela implementação, no ambiente R, do modelo considerando três aproximações numéricas da integral: método de Laplace, quadratura Gaussiana e quadratura Gaussiana adaptativa.

Tabela 11 – Resultados das estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-Normal-gama (combinado) pela implementação no ambiente R via aproximação numérica pelo método de Laplace, quadratura Gaussiana e pela quadratura Gaussiana adaptativa

	Método de Laplace		Quadratura Gaussiana		Quadratura Gaussiana adaptativa	
Parâmetro	Estimativa	erro padrão	Estimativa	erro padrão	Estimativa	erro padrão
β_1	0,9112	0,1162	0,9001	0,1715	0,9115	0,1754
β_2	-0,2506	0,2123	-0,2380	0,2410	-0,2547	0,2496
β_3	-0,0249	0,0051	-0,0250	0,0077	-0,0248	0,0077
β_4	0,0130	0,0062	0,0132	0,0107	0,0130	0,0107
ϕ	0,9027	0,1140	0,9024	0,0859	0,9018	0,0858
τ	-0,0598	0,0477	-0,0613	0,0819	-0,0615	0,0818
-2 log ver	5417,785		5416,996		5417,021	

Para melhor ajuste do modelo, o mesmo foi reparametrizado, sendo $\beta_{00} = \beta_1$, $\beta_{10} = \beta_1 + \beta_2$, $\beta_{01} = \beta_3$, $\beta_{11} = \beta_3 + \beta_4$, $\alpha_1 = e^{\phi}$, $\alpha_2 = 1/e^{\phi}$ e $\sigma^2 = 1/(e^{\tau})^2$. Assim, obtêm-se os valores da Tabela 12 para comparação com os resultados da Tabela **??**. As estimativas e os devios padrões calculados são muito similares. Os resultados obtidos pelo método da

quadratura Gaussiana são bem próximos aos obtidos pelo software SAS que utiliza o mesmo método para aproximação de integrais.

Tabela 12 – Resultados das estimativas de máxima verossimilhança para os parâmetros do modelo Poisson-normal-gama (combinado), usando-se o pacote dclone do ambiente R e o procedimente NLMIXED do SAS

	Método de Laplace		Quadratura Gaussiana		Quadratura Gaussiana adaptativa	
Parâmetro	Estimativa	erro padrão	Estimativa	erro padrão	Estimativa	erro padrão
β_{00}	0,9112	0,1162	0,9001	0,1715	0,9115	0,1754
β_{01}	-0,0249	0,0051	-0,0250	0,0077	-0,0248	0,0077
β_{10}	0,6606	-	0,6621	-	0,6568	-
β_{11}	-0,0119	-	-0,0118	-	0,0118	-
α_1	2,4663	-	2,4655	-	2,4640	-
α_2	0,4055	-	0,4056	-	0,4058	-
σ^2	1,1270	-	1,1304	-	1,1309	-

4.4 Análise Bayesiana para os dados de dermatite

Os resultados obtidos para a metodologia desenvolvida na seção 3.2.3 e aplicada aos dados da seção 3.1.2 são apresentados e discutidos, a seguir. Obteve-se convergência apenas para os modelos Bernoulli e Bernoulli-normal, para os quais serão apresentadas as análises gráficas das cadeias dos parâmetros para verificação da convergência dos mesmos e, por fim, uma tabela com os resultados a posteriori para cada um desses modelos. O estudo mais aprofundado dos outros modelos será tema de pesquisa futura.

Modelo 1: Bernoulli com função de ligação logística

As Figuras de 37 a 39, apresentam as análises gráficas das cadeias de todos os parâmetros do modelo Bernoulli com função de ligação logística. Os gráficos permitem verificar que há indícios de convergência das cadeias e que as mesmas são não correlacionadas.

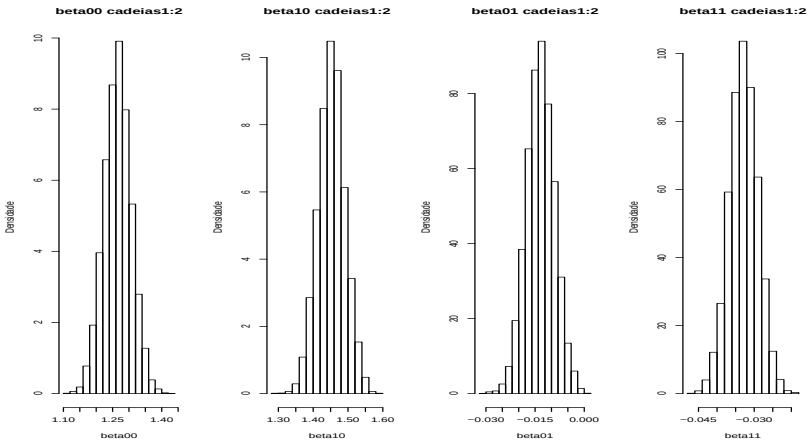


Figura 37 – Histograma dos valores gerados para β_{00}, β_{01}, β_{10} e β_{11}, considerando o conjunto de dados sob o modelo Bernoulli

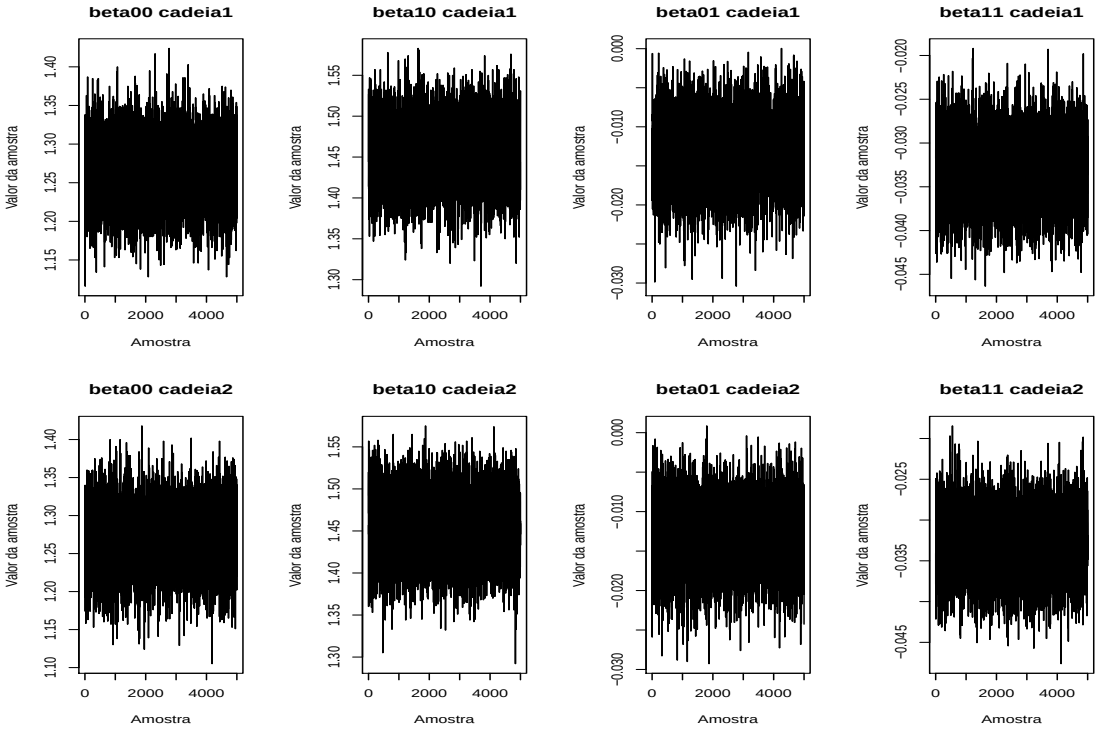


Figura 38 – Trajetória das cadeias geradas para β_{00}, β_{01}, β_{10} e β_{11}, considerando o conjunto de dados sob o modelo Bernoulli

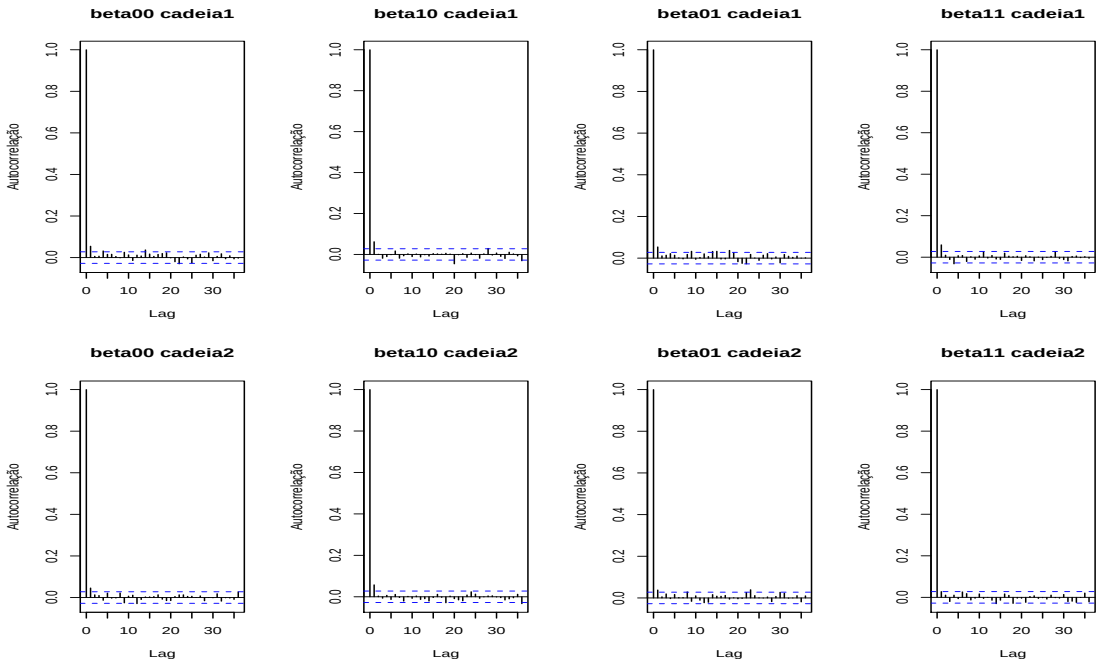


Figura 39 – Autocorrelação das cadeias geradas para β_{00}, β_{01}, β_{10} e β_{11}, considerando o conjunto de dados sob o modelo Bernoulli

Modelo 3: Bernoulli-normal com função de ligação logística

As Figuras de 40 a 42, apresentam as análises gráficas das cadeias de todos os parâmetros do modelo Bernoulli-normal com função de ligação logística. Os gráficos permitem verificar que há indícios de convergência das cadeias e que as mesmas são não correlacionadas.

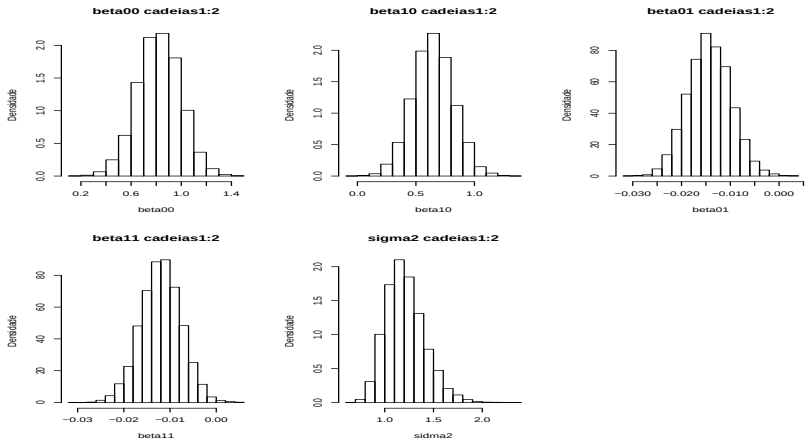


Figura 40 – Densidade dos valores gerados para β_{00}, β_{01}, β_{10}, β_{11} e σ^2 considerando o conjunto de dados sob o modelo Bernoulli-normal

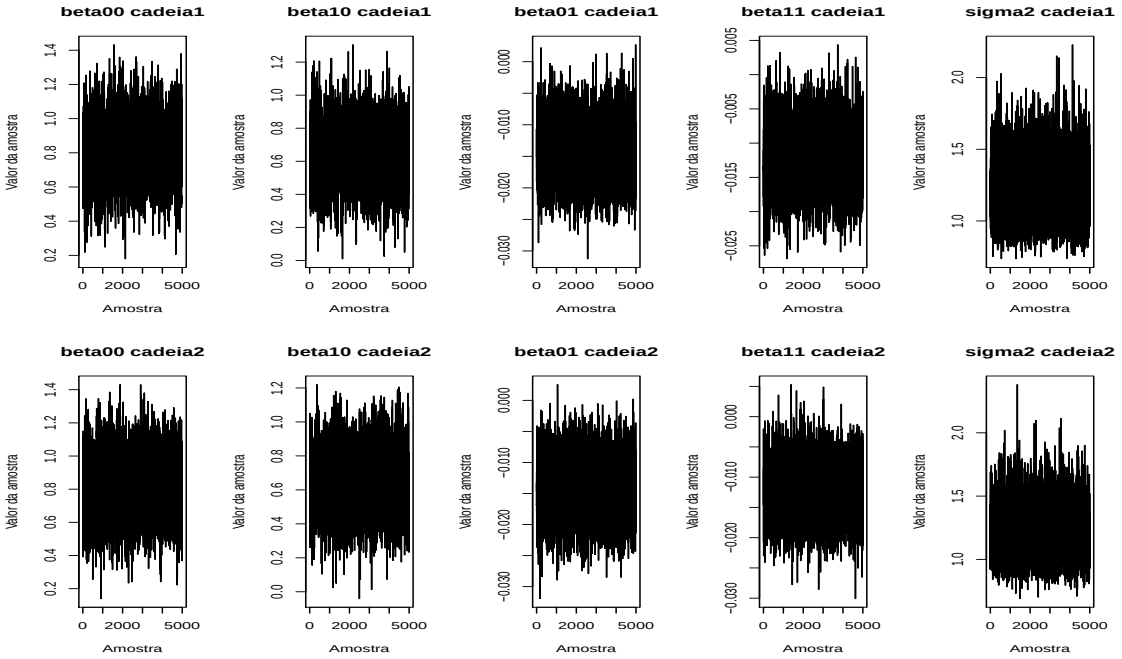


Figura 41 – Trajetória das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11} e σ^2 considerando o conjunto de dados sob o modelo Bernoulli-normal

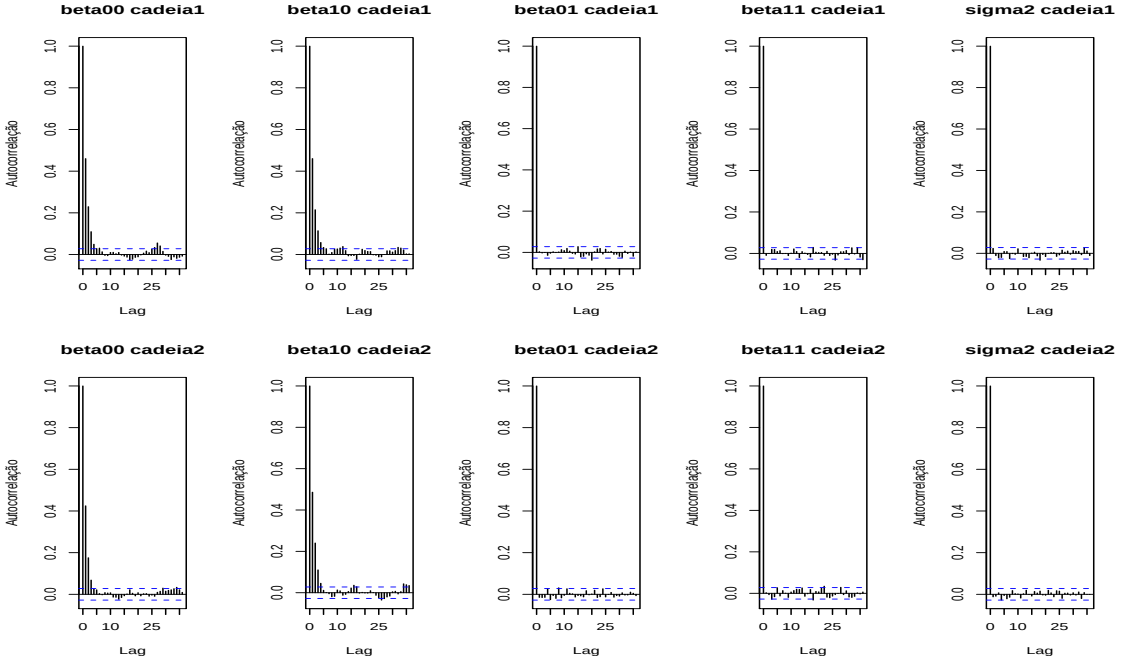


Figura 42 – Autocorrelação das cadeias geradas para β_{00}, β_{01}, β_{10}, β_{11} e σ^2 considerando o conjunto de dados sob o modelo Bernoulli-normal

A seguir, na Tabela 13 podem-se ver as estimativas obtidas para os parâmetros

para os modelos Bernoulli e Bernoulli-normal, a partir das distribuições a posteriori, sendo esses valores muito próximos das estimativas de máxima verossimilhança obtidas por Vieira (2008).

Table 13 – Resumos a posteriori para os parâmetros dos modelos Poisson, Poisson-gama, Poisson-normal e combinado

Modelo Bernoulli logístico				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	1,26600	0,04239	1,26500	(1,18200 ; 1,34700)
β_{01}	-0,01334	0,00432	-0,01335	(-0,02173 ; -0,00471)
β_{10}	1,45300	0,03800	1,45300	(1,37800 ; 1,52700)
β_{11}	-0,03282	0,00380	-0,03287	(-0,04024 ; -0,02535)
Modelo Bernoulli-normal logístico				
Parâmetro	Média	d.p.	Mediana	Intervalo de Credibilidade
β_{00}	0,80830	0,17000	0,80980	(0,46470 ; 1,13700)
β_{01}	-0,01430	0,00443	-0,01432	(-0,02312 ; -0,00561)
β_{10}	0,64140	0,16990	0,64310	(0,30990 ; 0,97280)
β_{11}	-0,01200	0,00431	-0,01202	(-0,02053 ; -0,00354)
σ^2	1,21100	0,20120	1,18900	(0,87670 ; 1,66300)

5 CONSIDERAÇÕES FINAIS

O principal objetivo deste trabalho foi apresentar uma proposta bayesiana dos modelos para análise de dados longitudinais na forma de contagens e binários com superdispersão, como alternativa aos modelos propostos por Molenberghs et al. (2010).

Para a análise de dados na forma de contagens considerou-se a distribuição de Poisson com a inclusão separada ou conjunta de efeitos aleatórios, sendo um efeito normal no preditor linear para acomodar a correlação de medidas longitudinais e um efeito gama multiplicativo para acomodar a superdispersão. Os modelos foram usados para a análise de um conjunto de dados reais para tratamento de pacientes epilépticos. Para selecionar o modelo que melhor se ajustava aos dados, utilizou-se a teoria das quantidades pivotais introduzida por Johnson (2007), a qual, assim como o DIC, direcionou para a seleção do modelo que possui os dois efeitos aleatórios simultaneamente. A partir da análise realizada neste trabalho, e pelos resultados apresentados em Molenberghs, Verbeke e Demétrio (2007) e Vangeneugden et al. (2009), baseados no mesmo conjunto de dados, observa-se que a teoria bayesiana e a frequentista apresentaram resultados semelhantes. Para ambos, o modelo que combina os efeitos aleatórios normal e gama foi o que melhor se ajustou aos dados, concluindo-se que os tratamentos não diferem entre si, sob o modelo Poisson-normal-gama.

Também foi realizado um estudo de simulação para avaliar o método das quantidades pivotais usado para verificação de ajuste do modelo. Foram simulados 3 conjuntos de dados baseados nos modelos Poisson-gama, Poisson-normal e Poisson-normal gama. Observou-se que os modelos utlizados na simulação foram os selecionados tanto pelo método das quantidades pivotais como pelo DIC, como esperado.

Para modelar dados binários longitudinais, considerou-se a distribuição Bernoulli com a inclusão separada ou conjunta de efeitos aleatórios, sendo um efeito normal no preditor linear para acomodar a correlação de medidas longitudinais e um efeito beta multiplicativo para acomodar a superdispersão. Foram consideradas duas funções de ligação, logística e probit. Os modelos foram usados para a análise de um conjunto de dados reais para tratamento de dermatite. Obteve-se convergência apenas nos modelos Bernoulli e Bernoulli-normal com função de ligação logística. Os resultados obtidos com a análise baseada na teoria Bayesiana apresentaram resultados muito próximos aos apresentados em Molenberghs et al.

(2010). O estudo mais aprofundado dos outros modelos será tema de pesquisa futura.

REFERÊNCIAS

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: INTERNACIONAL SYMPOSIUM OF INFORMATION THEORY, 2., Budapest, 1973. **Proceedings...** Budapest: Akademiai Kiàdio, p. 267-281, 1973.

BAYARRI, M.J.; BERGER, J.O. P values for composite null models. **Journal of the American Statistical Association**, Boston, v.95, p.1127-1142, 2000.

BERNOULLI, J. **Ars conjectandi, opus posthumum**: accedit tractatus de seriebus infinitis, et epistola gallicé scripta de ludo pilae reticularis. Basel: Thurneysen Brothers, 1713.

BOX, G.E.P.; COX, D.R. An analysis of transformation. **Journal of the Royal Statistical Society Series B**, London, v.26, p.211-52, 1964.

BOX, G.E.P.; TIAO, G.C. **Bayesian inference in statistical analysis**. New York: Wiley, 1992. 588p.

BOOTH, J.G.; CASELLA, G.; FRIEDL, H.; HOBERT, J.P. Negative binomial loglinear mixed models. **Statistical Modelling**, London, v.3, p.179–191, 2003.

BRESLOW, N.E. Extra-Poisson variation in log-linear models. **Journal of the Royal Statistical Society Series C (Applied Statistics)**, London, v. 33, p. 38-44, 1984.

BRESLOW, N.E.; CLAYTON, D.G. Approximate inference in generalized linear mixed models. **Journal of the American Statistical Association**, Boston, v. 88, p.9-25, 1993.

CARLIN, B.P.; LOUIS, T.A. **Bayes and empirical Bayes methods for data analysis**. London: Chapman & Hall, 1996. 399p.

CASELLA, G.; GEORGE, E.I. Explaining the Gibbs Sampler. **The American Statistician**, Washington, v.46, p. 167-174, 1992.

CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hasting algorithm. **The American Statistician**, Washington, v. 49, p. 326-335, 1995.

CORDEIRO, G.M.; DEMETRIO, C.G.B. **Modelos lineares generalizados e extensões**. Piracicaba, 2011. 246p.

DE BACKER, M.; DE KEYSER, P.; DE VROEY, C.; LESAFFRE, E. A 12-week treatment for dermatophyte toe onychomycosis: terbinafine 250mg/day vs itraconazole 200mg/day - a double-blind comparative trial. **British Journal of Dermatology**, London, v.134, p.16-17, 1996.

DIGGLE, P.J. An approach to the analysis of repeated measurements. **Biometrics**, Washington, v.44, p.959-971, 1988.

FAUGHT, E.; WILDER, B.J.; RAMSAY, R.E.; REIFE, R.A.; KRAMER, L.D.; PLEDGER, G.W.; KARIM, R.M. Topiramate placebo-controlled dose-ranging trial in refratory partial epilepsy using 200-, 400-, and 600-mg daily dosages. **Neurology**, New York, v.46, p.1684-1690, 1996.

GILKS, W.R.; RICHARDSON, S.; SPIEGELHALTER, D.J. **Markov chain Monte Carlo in practice**. London: Chapman and Hall, 1998. 486p.

GELMAN, A.; RUBIN, D.B. Inference from iterative simulation using multiple sequences. **Statistical Science**, Beachwood, v.7, n.4, p.457-511, 1992.

GELMAN, A.; MENG, X.L; STERN, H. Posterior Predictive assessment of model fitness via realized discrepancies. **Statistica Sinica**, Los Angeles, v.6, p.733-807, 1996.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. London: Chapman and Hall. 1995. 698p.

GUTTMAN, I. The use of the concept of a future observation in goodness-of-fit problems. **Journal of Royal Statistical Society Series B**, London, v.29, p.83-100, 1967.

HELMS, R.W. Intentionally incomplete longitudinal designs: I. Methodology and comparison of some full span designs. **Statistics in Medicine**, Chichester, v.11, p. 1889-1913, 1992.

HINDE, J.P.; DEMÉTRIO, C.G.B. Overdispersion: models and estimation. In: SIMPÓSIO NACIONAL de PROBABILIDADE e ESTATÍSTICA, 13., 1998, Caxambú. **Anais...**Caxambú: Associação Brasileira de Estatística, 1998. 73p.

JOHNSON, V.E. Bayesian model assessment using pivotal quantities. **Bayesian Analysis**, Pittsburgh, v.2, n.4, p.719-734, 2007.

KANASHIRO, S. **Efeitos de diferentes substratos na produção da espécie Aechmea fasciata (Lindley) baker em vasos.** 1999. 79p. Dissertação (Mestrado em Fitotecnia) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1999.

LAWLESS, J. Negative binomial and mixed poisson regression. **The Canadian Journal of Statistics**, Ottawa, v. 15, p. 209-225, 1987.

LELE, S.R.; DENNIS, B.; LUTSCHER, F. Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. **Ecology Letters**, Malden, v.10, p.551-563, 2007.

LELE, S.R.; NADEEM, K.; SCHUMULAND, B. Estimability and likelihood inference fo generalized linear mixed models using data cloning. **Journal of American Statistical Assocition**, Alexandria, v.105, n.492, p.1617-1625, 2010.

LINDSEY, J. **Modelling frequency and count data**. Oxford: Oxford University Press, 1995. 291p.

MCCULLAGH, P.; NELDER, J. **Generalized linear models**. 2nd ed. London: Chapman and Hall, 1989.

MILLAR, R.B. Comparison of hierarchical Bayesian models for overdispersed count data using DIC and Bayes'factors. **Biometrics**, Washington, v.65, p.962-969, 2009.

MOLENBERGHS, G.; VERBEKE, G. **Models for discrete longitudinal data**. New York: Springer, 2005.

MOLENBERGHS, G.; VERBEKE, G.; DEMÉTRIO, C.G.B. An extended random-effects approach to modeling repeated, overdispersion count data. **Lifetime Data Analysis**, Hinghan, v.13, n.4, 2007.

MOLENBERGHS, G.; VERBEKE, G,; DEMETRIO, C.G.B.; VIEIRA, A.M.C. A family of generalized linear models for repeated measures with normal and conjugate random effects. **Statistical Science**, Hayward, v.25, n.3, p.325-347, 2010.

NELDER, J.A.; WEDDERBURN, W.M. Generalized Linear Model. **Journal of the Royal Statistical Society**, London, v. 135, n.3, p. 370-384, 1972.

NTZOUFRAS, I. **Bayesian modeling using WinBUGS**. New Jersey: Wiley, 2009. 492p.

PLUMMER, M. **JAGS Version 2.0.0 manual** (Abril, 2010) 2010a. Disponível em: <http://mcmc-jags.sourceforge.net>. Acesso em: 25 nov. 2011.

POISSON, S. D. **Recherches sur la probabilité des jugements nn matière criminelle et en matière civile**: précédées des règles générales du calcul des probabilités. Paris: Bachelier, 1837.

R DEVELOPMENT CORE TEAM. **R:** a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2009. Disponível em: <http://www.R-project.org>. Acesso em: 25 nov. 2011.

ROBINS, J.M.; VAN DER VAART, A.; VENTURA, V. Asymptotic distribution of P values in composite null models. **Journal of the American Statistical Association**, Boston, v.95, p.1143-1159, 2000.

RUBIN, D.B. Bayesianly justifiable and relevant frequency calculations for the applied statistician. **The Annals of Statistics**, Philadelphia, v.12, p. 1151-1172, 1984.

SAS INSTITUTE. **SAS Software Release 9.1.3**. Cary: SAS Institute, 2008.

SPIEGELHALTER, D.J.; BEST, N.G.; CARLIN, B.P.; LINDE, A. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society Series B**, London, v.64, n.4, p.583-639, 2000.

SPIEGELHALTER, D.J.; THOMAS, A.; BEST, N.; GILKS, W. **Bugs**: Bayesian inference Using Gibbs Sampling Manual, 1996. Disponível em: <http://www.mrc-bsu.cam.ac.uk/bugs>. Acesso em: 25 nov. 2011.

THOMAS, A.; SPIEGELHALTER, D.J.; GILKS, W.R. BUGS: A program to perform Bayesian inference using Gibbs Sampling. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, editors. **Bayesian Statistics**, v.4, Oxford: Clarendon Press; 1992. p. 837-842.

VANGENEUGDEN, T.; MOLENBERGHS, G.; VERBEKE, G.; DEMÉTRIO, C. G. B. Marginal correlation from an extended random-effects model for repeated, overdispersed counts. **Journal of Applied Statistics**, Abingdon, v.38, n.2, p.215–232, 2011.

VIEIRA, A.M.C. **Modelagem simultânea de média e dispersão e aplicações na pesquisa agronômica**. 2008. 117p. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 2008.

VIEIRA, A.M.C.; HINDE, J.P.; DEMETRIO, C.G.B. Zero-inflated proportion data models applied to a biological control assay. **Journal of Applied Statistics**, Abingdon, v.27, n.3, p.373-389, 2000.

ANEXOS

Anexo A: Distribuições de probabilidade conjunta para modelos Poisson

A seguir, serão apresentadas as distribuições marginais de $\boldsymbol{Y}_i$ para os modelos Poisson-normal e Poisson-normal-gama conforme Molenberghs et al. (2010).

Modelo Poisson-normal

Para o modelo Poisson-normal, a contribuição do i-ésimo indivíduo para a função de verossimilhança é obtida por (5). Sabe-se que, para esse modelo, $\lambda_{ij} = \exp\{\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i\}$, considerando que $\gamma_{ij} = \boldsymbol{x}'_{ij}\boldsymbol{\beta}$ e $c_{ij} = \boldsymbol{z}'_{ij}\boldsymbol{b}_i$ têm-se:

$$\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i = \boldsymbol{y}_i) &= \int \prod_{j=1}^{n_i} \frac{e^{-\lambda_{ij}}\lambda_{ij}^{y_{ij}}}{y_{ij}!} \frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}} e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i} d\boldsymbol{b}_i \\
&= \int \prod_{j=1}^{n_i} \frac{1}{y_{ij}!} e^{-e^{\gamma_{ij}+c_{ij}}} (e^{\gamma_{ij}+c_{ij}})^{y_{ij}} \frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}} e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i} d\boldsymbol{b}_i.
\end{aligned}$$

Seja a série $\sum_{t=0}^{+\infty} \frac{(-1)^t e^{t\gamma_{ij}}}{t!} e^{tc_{ij}} = e^{-e^{\gamma_{ij}+c_{ij}}}$:

$$\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i = \boldsymbol{y}_i) &= \int \prod_{j=1}^{n_i} \left\{ \frac{1}{y_{ij}!} \sum_{t=0}^{+\infty} \left[\frac{(-1)^t e^{t\gamma_{ij}}}{t!} e^{tc_{ij}} \right] (e^{\gamma_{ij}+c_{ij}})^{y_{ij}} \right\} \frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}} e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i} d\boldsymbol{b}_i \\
&= \frac{1}{\prod_{j=1}^{n_i} y_{ij}!} \int \prod_{j=1}^{n_i} \left\{ \sum_{t=0}^{+\infty} \left[\frac{(-1)^t e^{(t+y_{ij})\gamma_{ij}}}{t!} e^{(t+y_{ij})c_{ij}} \right] \right\} \frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}} e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i} d\boldsymbol{b}_i \\
&= \frac{1}{\prod_{j=1}^{n_i} y_{ij}!} \int \sum_{\boldsymbol{t}} \left\{ \prod_{j=1}^{n_i} \frac{(-1)^{t_j} e^{(t_j+y_{ij})\gamma_{ij}}}{t_j!} e^{(t_j+y_{ij})c_{ij}} \right\} \frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}} e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i} d\boldsymbol{b}_i \\
&= \frac{1}{\prod_{j=1}^{n_i} y_{ij}!} \sum_{\boldsymbol{t}} \left\{ \prod_{j=1}^{n_i} \frac{(-1)^{t_j} e^{(t_j+y_{ij})\gamma_{ij}}}{t_j!} \int e^{\sum_{j=1}^{n_i}(t_j+y_{ij})c_{ij}} \frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}} e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i} d\boldsymbol{b}_i \right\}.
\end{aligned}$$

Mas, $\sum_{j=1}^{n_i}(t_j+y_{ij})c_{ij}=\sum_{j=1}^{n_i}(t_j+y_{ij})\boldsymbol{z}'_{ij}\boldsymbol{b}_i$, substituindo $w'_{ij}=\sum_{j=1}^{n_i}(t_j+y_{ij})\boldsymbol{z}'_{ij}$, tem-se:

$$\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i=\boldsymbol{y}_i) &= \frac{1}{\prod_{j=1}^{n_i}y_{ij}!}\sum_{\boldsymbol{t}}\left\{\prod_{j=1}^{n_i}\frac{(-1)^{t_j}e^{(t_j+y_{ij})\gamma_{ij}}}{t_j!}\int e^{w'_{ij}\boldsymbol{b}_i}\frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}}e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i}d\boldsymbol{b}_i\right\}\\
&= \frac{1}{\prod_{j=1}^{n_i}y_{ij}!}\sum_{\boldsymbol{t}}\left\{\prod_{j=1}^{n_i}\frac{(-1)^{t_j}e^{(t_j+y_{ij})\gamma_{ij}}}{t_j!}\times\right.\\
&\times \left.\int\frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}}e^{-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}_i+w'_{ij}\boldsymbol{b}_i-\frac{1}{2}w'_{ij}\boldsymbol{D}w_{ij}+\frac{1}{2}w'_{ij}\boldsymbol{D}w_{ij}}d\boldsymbol{b}_i\right\}\\
&= \frac{1}{\prod_{j=1}^{n_i}y_{ij}!}\sum_{\boldsymbol{t}}\left\{\prod_{j=1}^{n_i}\frac{(-1)^{t_j}e^{(t_j+y_{ij})\gamma_{ij}}}{t_j!}e^{\frac{1}{2}w'_{ij}\boldsymbol{D}w_{ij}}\times\right.\\
&\times \left.\int\frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}}e^{-\frac{1}{2}(\boldsymbol{b}_i-\boldsymbol{D}w_{ij})'\boldsymbol{D}^{-1}(\boldsymbol{b}_i-\boldsymbol{D}w_{ij})}d\boldsymbol{b}_i\right\}.
\end{aligned}$$

Como $\int\frac{1}{|\boldsymbol{D}|^{1/2}(2\pi)^{q/2}}e^{-\frac{1}{2}(\boldsymbol{b}_i-\boldsymbol{D}w_{ij})'\boldsymbol{D}^{-1}(\boldsymbol{b}_i-\boldsymbol{D}w_{ij})}d\boldsymbol{b}_i=1$, tem-se:

$$\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i=\boldsymbol{y}_i) &= \frac{1}{\prod_{j=1}^{n_i}y_{ij}!}\sum_{\boldsymbol{t}}\left\{\frac{(-1)^{\sum_{j=1}^{n_i}t_j}}{\prod_{j=1}^{n_i}t_j!}\exp\left[\sum_{j=1}^{n_i}(y_{ij}+t_j)\boldsymbol{x}'_{ij}\boldsymbol{b}_i\right]\times\right.\\
&\times \left.\exp\left[\frac{1}{2}\left(\sum_{j=1}^{n_i}(y_{ij}+t_j)\boldsymbol{z}'_{ij}\right)\boldsymbol{D}\left(\sum_{j=1}^{n_i}(y_{ij}+t_j)\boldsymbol{z}_{ij}\right)\right]\right\},
\end{aligned}$$

em que $\boldsymbol{t}=(t_1,t_2,\ldots,t_{n_i})$.

Modelo Poisson-normal-gama

Considerando o modelo com efeitos normal e gama, tem-se $\lambda_{ij}=\theta_{ij}\exp\{\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\boldsymbol{z}'_{ij}\boldsymbol{b}_i\}=\theta_{ij}e^{\gamma_{ij}+c_{ij}}=e^{\tilde{\gamma}_{ij}+c_{ij}}$. Dessa forma, $\mathrm{P}(\boldsymbol{Y}_i=\boldsymbol{y}_i|\boldsymbol{\eta}_i)$ assume a distribuição dada por (19). Assumindo que os efeitos aleatórios θ_{ij} são independentes com $\theta_{ij}\sim\mathrm{Gama}(\alpha_{1j},\alpha_{2j})$, há apenas a necessidade de integrar em relação a θ_{ij}. Reescrevendo os termos de $\tilde{\gamma}_{ij}$, o fator a ser integrado é:

$$\begin{aligned}
\int\frac{1}{\alpha_{2j}}\frac{1}{\Gamma(\alpha_{1j})}\theta_{ij}^{\alpha_{ij}-1}e^{-\theta_{ij}/\alpha_{2j}}\theta_{ij}^{y_{ij}+t_j}d\theta_{ij} &= \frac{1}{\alpha_{2j}}\frac{1}{\Gamma(\alpha_{1j})}\int\theta_{ij}^{\alpha_{1j}+y_{ij}+t_j-1}e^{-\theta_{ij}/\alpha_{2j}}d\theta_{ij}\\
&= \frac{1}{\alpha_{2j}}\frac{1}{\Gamma(\alpha_{1j})}\alpha_{2j}^{\alpha_{1j}+y_{ij}+t_j}\Gamma(\alpha_{1j}+y_{ij}+t_j). \quad (53)
\end{aligned}$$

Assim, para o modelo que combina os efeitos aleatórios normal e gama, a con-

tribuição do i-ésimo indivíduo para a função de verossimilhança é:

$$
\begin{aligned}
\mathrm{P}(\boldsymbol{Y}_i = \boldsymbol{y}_i) &= \sum_{\boldsymbol{t}} \left[\prod_{j=1}^{n_i} \frac{(-1)^{t_j}}{y_{ij}!t_j!} \frac{1}{\alpha_{2j}} \frac{1}{\Gamma(\alpha_{1j})} \alpha_{2j}^{\alpha_{1j}+y_{ij}+t_j} \Gamma(\alpha_{1j}+y_{ij}+t_j) e^{(y_{ij}+t_j)\boldsymbol{x}'_{ij}\boldsymbol{\beta}} \right] \times \\
&\times \exp\left[\frac{1}{2} \left(\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{z}'_{ij} \right) \boldsymbol{D} \left(\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{z}_{ij} \right) \right] \Bigg\} \\
&= \sum_{\boldsymbol{t}} \left[\prod_{j=1}^{n_i} \frac{(-1)^{t_j}}{y_{ij}!t_j!} \frac{1}{(\alpha_{1j}-1)!} \alpha_{2j}^{y_{ij}+t_j} (\alpha_{1j}+y_{ij}+t_j-1)! \right] \times \\
&\times \exp\left\{ \sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{x}'_{ij}\boldsymbol{\beta} \right\} \times \\
&\times \exp\left[\frac{1}{2} \left(\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{z}'_{ij} \right) \boldsymbol{D} \left(\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{z}_{ij} \right) \right] \Bigg\} \\
&= \sum_{\boldsymbol{t}} \left[\prod_{j=1}^{n_i} \binom{y_{ij}+t_j}{y_{ij}} \binom{\alpha_{1j}+y_{ij}+t_j-1}{\alpha_{1j}-1} (-1)^{t_j} \alpha_{2j}^{y_{ij}+t_j} \right] \times \\
&\times \exp\left(\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{x}'_{ij}\boldsymbol{\beta} \right) \times \\
&\times \exp\left\{ \frac{1}{2} \left[\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{z}'_{ij} \right] \boldsymbol{D} \left[\sum_{j=1}^{n_i} (y_{ij}+t_j)\boldsymbol{z}_{ij} \right] \right\},
\end{aligned}
$$

em que $\boldsymbol{t} = (t_1, t_2, \ldots, t_{n_i})$.

Observa-se pela equação (53) que quando tem-se apenas o efeito aleatório gama, obtém-se a distribuição Binomial Negativa.

Anexo B: Distribuições de probabilidade para modelos Bernoulli

A seguir, serão apresentadas as funções de verossimilhança dos parâmetros dos modelos Bernoulli-normal e Bernoulli-normal-beta conforme Vieira (2008).

Sejam os modelos dados por (34) e (35) e considere, respectivamente, $\kappa_{ij} = \frac{\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\boldsymbol{z}'_{ij}\boldsymbol{b}_i)}{1+\exp(\boldsymbol{x}'_{ij}\boldsymbol{\beta}+\boldsymbol{z}'_{ij}\boldsymbol{b}_i)}$ para a função de ligação logística e $\kappa_{ij} = \Phi(\boldsymbol{x}'_{ij}\boldsymbol{\beta} + \boldsymbol{z}'_{ij}\boldsymbol{b}_i)$ para a função de ligação probit. Considere, ainda a função densidade de probabilidade para o efeito aleatório $\boldsymbol{b}_i$ dada por

$$f(\boldsymbol{b}_i) = \frac{1}{\sqrt{(2\pi)^{n_i}}}\frac{1}{\sqrt{|\boldsymbol{D}|}}\exp\Big(-\frac{1}{2}\boldsymbol{b}'_i\boldsymbol{D}^{-1}\boldsymbol{b}'_i\Big), \quad \boldsymbol{b}_i \in \mathbb{R}^q$$

e a função densidade de probabilidade para o efeito aleatório θ_{ij} dada por

$$f(\theta_{ij}) = \frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1,\alpha_2)}, \quad \theta_{ij} \in [0,1].$$

A contribuição do (ij)-ésimo indivíduo para a função de verossimilhança, assumindo a independência entre θ_{ij} e $\boldsymbol{b}_i$ é dada por

$$f(y_{ij},\theta_{ij},\boldsymbol{b}_i) = f(y_{ij}|\theta_{ij},\boldsymbol{b}_i)f(\boldsymbol{b}_i)f(\theta_{ij}),$$

e, portanto a contribuição do i-ésimo indivíduo para a função de verossimilhança é

$$f(\boldsymbol{y}_i,\boldsymbol{\theta}_i,\boldsymbol{b}_i) = \prod_{j=1}^{n_i} f(y_{ij}|\theta_{ij},\boldsymbol{b}_i)f(\boldsymbol{b}_i)f(\theta_{ij}), \tag{54}$$

em que $\boldsymbol{y}_i = (y_{i1},\ldots,y_{in_i})$ e $\boldsymbol{\theta}_i = \theta_{i1},\ldots,\theta_{in_i}$.

Para a estimação por máxima verossimilhança, a função (54) necessita ser integrada em relação a $\boldsymbol{\theta}_i$ e $\boldsymbol{b}_i$. A integração em relação a $\boldsymbol{b}_i$ pode ser solucionada via integração numérica, assim deseja-se obter $f(y_{ij}|\boldsymbol{b}_i) = \int_0^1 f(y_{ij}|\theta_{ij},\boldsymbol{b}_i)f(\theta_{ij})d\theta_{ij}$. No caso Bernoulli, necessita-se considerar $y_{ij} = 0$ e $y_{ij} = 1$:

$$
\begin{aligned}
f(y_{ij}=0|\boldsymbol{b}_i) &= \int_0^1 f(y_{ij}=0|\theta_{ij},\boldsymbol{b}_i)f(\theta_{ij})d\theta_{ij} \\
&= \int_0^1 (1-\theta_{ij}\kappa_{ij})\frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1,\alpha_2)}d\theta_{ij} \\
&= \int_0^1 \frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1,\alpha_2)}d\theta_{ij} - \theta_{ij}\kappa_{ij}\int_0^1 \frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1,\alpha_2)}d\theta_{ij} \\
&= 1-\frac{\kappa_{ij}}{B(\alpha_1,\alpha_2)}\int_0^1 \theta_{ij}^{(\alpha_1+1)-1}(1-\theta_{ij})^{\alpha_2-1}d\theta_{ij} \\
&= 1-\frac{\kappa_{ij}}{B(\alpha_1,\alpha_2)}B(\alpha_1+1,\alpha_2) \\
&= \frac{(1-\kappa_{ij})\alpha_1+\alpha_2}{\alpha_1+\alpha_2} \qquad (55)
\end{aligned}
$$

$$
\begin{aligned}
f(y_{ij}=1|\boldsymbol{b}_i) &= \int_0^1 f(y_{ij}=1|\theta_{ij},\boldsymbol{b}_i)f(\theta_{ij})d\theta_{ij} \\
&= \int_0^1 (\theta_{ij}\kappa_{ij})\frac{\theta_{ij}^{\alpha_1-1}(1-\theta_{ij})^{\alpha_2-1}}{B(\alpha_1,\alpha_2)}d\theta_{ij} \\
&= \frac{\kappa_{ij}}{B(\alpha_1,\alpha_2)}\int_0^1 \theta_{ij}^{(\alpha_1+1)-1}(1-\theta_{ij})^{\alpha_2-1}d\theta_{ij} \\
&= \frac{\kappa_{ij}}{B(\alpha_1,\alpha_2)}B(\alpha_1+1,\alpha_2) \\
&= \frac{\kappa_{ij}\alpha_1}{\alpha_1+\alpha_2} \qquad (56)
\end{aligned}
$$

Assim, substituindo (55) e (56) em (6) obtém-se a função de verossimilhança para o modelo Bernoulli-normal-beta com função de ligação logit ou probit dada por

$$
\begin{aligned}
L(\boldsymbol{\beta},\boldsymbol{D},\alpha_1,\alpha_2) &= \prod_{i=1}^{N}\int\prod_{j=1}^{n_i}\Big(\frac{\kappa_{ij}\alpha_1}{\alpha_1+\alpha_2}\Big)^{y_{ij}}\Big(\frac{(1-\kappa_{ij})\alpha_1+\alpha_2}{\alpha_1+\alpha_2}\Big)^{1-y_{ij}}f(\boldsymbol{b}_i|\boldsymbol{D})d\boldsymbol{b}_i \\
&= \prod_{i=1}^{N}\int\prod_{j=1}^{n_i}\frac{1}{\alpha_1+\alpha_2}(\kappa_{ij}\alpha_1)^{y_{ij}}[(1-\kappa_{ij})\alpha_1+\alpha_2]^{1-y_{ij}}f(\boldsymbol{b}_i|\boldsymbol{D})d\boldsymbol{b}_i,
\end{aligned}
$$

e, considerando o modelo Bernoulli-beta como um caso particular do modelo Bernoulli-normal-beta, porém, sem a presença do efeito aleatório $\boldsymbol{b}_i$, a sua função de verossimilhança é dada por

$$
\begin{aligned}
L(\boldsymbol{\beta},\alpha_1,\alpha_2) &= \prod_{i=1}^{N}\prod_{j=1}^{n_i}\Big(\frac{\kappa_{ij}\alpha_1}{\alpha_1+\alpha_2}\Big)^{y_{ij}}\Big(\frac{(1-\kappa_{ij})\alpha_1+\alpha_2}{\alpha_1+\alpha_2}\Big)^{1-y_{ij}} \\
&= \prod_{i=1}^{N}\prod_{j=1}^{n_i}\frac{1}{\alpha_1+\alpha_2}(\kappa_{ij}\alpha_1)^{y_{ij}}[(1-\kappa_{ij})\alpha_1+\alpha_2]^{1-y_{ij}}.
\end{aligned}
$$

Anexo C: Distribuições condicionais a posteriori para modelos Poisson

A seguir, serão apresentadas as distribuições condicionais a posteriori para os modelos da seção 3.2.1.

Modelo 1: Poisson

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) \propto \psi_1(\beta_{00},\beta_{01},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_1(\beta_{00},\beta_{01},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{00}\sum_{i\in plac}\sum_{j=1}^{n_i} y_{ij} - \sum_{i\in plac}\sum_{j=1}^{n_i}\Big[\exp(\beta_{00}+\beta_{01}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) \propto \psi_2(\beta_{00},\beta_{01},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_2(\beta_{00},\beta_{01},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{01}\sum_{i\in plac}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in plac}\sum_{j=1}^{n_i}\Big[\exp(\beta_{00}+\beta_{01}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{y},\boldsymbol{t}) \propto \psi_3(\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_3(\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{10}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\exp(\beta_{10}+\beta_{11}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{y},\boldsymbol{t}) \propto \psi_4(\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_4(\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{11}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\exp(\beta_{10}+\beta_{11}t_{ij})\Big]\right\}.$$

Modelo 2: Poisson-gama

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_5(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_5(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{00}\sum_{i\in plac}\sum_{j=1}^{n_i} y_{ij} - \sum_{i\in plac}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{00}+\beta_{01}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_6(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_6(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{01}\sum_{i\in plac}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in plac}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{00}+\beta_{01}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_7(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_7(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{10}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{10}+\beta_{11}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{\theta},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_8(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_8(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{11}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{10}+\beta_{11}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para θ_{ij}, $i = 1,\ldots,N$, $j = 1,\ldots,n_i$:

$$\pi(\theta_{ij}|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta}_{(ij)},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \text{Gama}\Big(y_{ij}+\alpha_2^{-1},\alpha_2^{-1}+\exp(\beta_{00}+\beta_{01}t_{ij})\Big), & \text{se} \quad i\in\text{placebo} \\ \text{Gama}\Big(y_{ij}+\alpha_2^{-1},\alpha_2^{-1}+\exp(\beta_{10}+\beta_{11}t_{ij})\Big), & \text{se} \quad i\in\text{tratamento.} \end{cases}$$

Sendo $\boldsymbol{\theta}_{(ij)} = (\theta_{11}, \ldots, \theta_{1n_1}, \ldots, \theta_{i1}, \ldots, \theta_{ij-1}, \theta_{ij+1}, \ldots, \theta_{in_i}, \ldots, \theta_{N1}, \ldots, \theta_{Nn_N})$.

Distribuição condicional completa a posteriori para α_2:

$$\pi(\alpha_2|\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{t}) \propto \text{Gama}(c, d)\psi_9(\alpha_2, \boldsymbol{\theta})$$

em que

$$\begin{aligned}\psi_9(\alpha_2, \boldsymbol{\theta}) &= \exp\left\{ -\alpha_2^{-1}\ln\alpha_2 \sum_{i=1}^{N} n_i - \ln\Gamma(\alpha_2^{-1}) \sum_{i=1}^{N} n_i \right\} \\ &\times \exp\left\{ (\alpha_2^{-1} - 1) \sum_{i=1}^{N}\sum_{j=1}^{n_i} \ln\theta_{ij} - \alpha_2^{-1} \sum_{i=1}^{N}\sum_{j=1}^{n_i} \theta_{ij} \right\}.\end{aligned}$$

Modelo 3: Poisson-normal

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{b}, \sigma^2, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_{10}(\beta_{00}, \beta_{01}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t})N(a, b^2)$$

em que

$$\psi_{10}(\beta_{00}, \beta_{01}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t}) = \exp\left\{ \beta_{00} \sum_{i \in plac}\sum_{j=1}^{n_i} y_{ij} - \sum_{i \in plac}\sum_{j=1}^{n_i} \Big[\exp(\beta_{00} + b_i + \beta_{01}t_{ij})\Big] \right\}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00}, \beta_{10}, \beta_{11}, \boldsymbol{b}, \sigma^2, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_{11}(\beta_{00}, \beta_{01}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t})N(a, b^2)$$

em que

$$\psi_{11}(\beta_{00}, \beta_{01}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t}) = \exp\left\{ \beta_{01} \sum_{i \in plac}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i \in plac}\sum_{j=1}^{n_i} \Big[\exp(\beta_{00} + b_i + \beta_{01}t_{ij})\Big] \right\}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00}, \beta_{01}, \beta_{11}, \boldsymbol{b}, \sigma^2, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_{12}(\beta_{10}, \beta_{11}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t})N(a, b^2)$$

em que

$$\psi_{12}(\beta_{10}, \beta_{11}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t}) = \exp\left\{ \beta_{10} \sum_{i \in trat}\sum_{j=1}^{n_i} y_{ij} - \sum_{i \in trat}\sum_{j=1}^{n_i} \Big[\exp(\beta_{10} + b_i + \beta_{11}t_{ij})\Big] \right\}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00}, \beta_{01}, \beta_{10}, \boldsymbol{b}, \sigma^2, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_{13}(\beta_{10}, \beta_{11}, \boldsymbol{b}, \boldsymbol{y}, \boldsymbol{t})N(a, b^2)$$

em que

$$\psi_{13}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{11}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\exp(\beta_{10}+b_i+\beta_{11}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para b_i, $i = 1,\ldots,N$:

$$\pi(b_i|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b}_{(i)},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \exp\left\{-\frac{b_i^2}{2\sigma^2} + b_i\sum_{j=1}^{n_i} y_{ij} - \sum_{j=1}^{n_i}\Big[\exp(\beta_{00}+b_i+\beta_{01}t_{ij})\Big]\right\}, & \text{se} \quad i\in \text{placebo} \\ \exp\left\{-\frac{b_i^2}{2\sigma^2} + b_i\sum_{j=1}^{n_i} y_{ij} - \sum_{j=1}^{n_i}\Big[\exp(\beta_{10}+b_i+\beta_{11}t_{ij})\Big]\right\}, & \text{se} \quad i\in \text{tratamento.} \end{cases}$$

Sendo $\boldsymbol{b}_{(i)} = (b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_N)$.

Distribuição condicional completa a posteriori para σ^2:

$$\pi(\sigma^2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) \propto \text{GamaInversa}\left(\frac{N}{2}+g, h+\sum_{i=1}^{N}\frac{b_i^2}{2}\right).$$

Modelo 4: Poisson-normal-gama (*Combinado*)

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{14}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{14}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{00}\sum_{i\in plac}\sum_{j=1}^{n_i} y_{ij} - \sum_{i\in plac}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{00}+b_i+\beta_{01}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{15}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{15}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{01}\sum_{i\in plac}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in plac}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{00}+b_i+\beta_{01}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{16}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{16}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{10}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{10}+b_i+\beta_{11}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{17}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{17}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp\left\{\beta_{11}\sum_{i\in trat}\sum_{j=1}^{n_i} y_{ij}t_{ij} - \sum_{i\in trat}\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{10}+b_i+\beta_{11}t_{ij})\Big]\right\}.$$

Distribuição condicional completa a posteriori para θ_{ij}, $i = 1,\ldots,N$, $j = 1,\ldots,n_i$:

$$\pi(\theta_{ij}|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta}_{(ij)},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \text{Gama}\Big(y_{ij}+\alpha_2^{-1},\alpha_2^{-1}+\exp(\beta_{00}+b_i+\beta_{01}t_{ij})\Big), & \text{se} \quad i\in\text{placebo} \\ \text{Gama}\Big(y_{ij}+\alpha_2^{-1},\alpha_2^{-1}+\exp(\beta_{10}+b_i+\beta_{11}t_{ij})\Big), & \text{se} \quad i\in\text{tratamento.} \end{cases}$$

Sendo $\boldsymbol{\theta}_{(ij)} = (\theta_{11},\ldots,\theta_{1n_1},\ldots,\theta_{i1},\ldots,\theta_{ij-1},\theta_{ij+1},\ldots,\theta_{in_i},\ldots,\theta_{N1},\ldots,\theta_{Nn_N})$.

Distribuição condicional completa a posteriori para b_i, $i = 1,\ldots,N$:

$$\pi(b_i|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b}_{(i)},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \exp\left\{-\frac{b_i^2}{2\sigma^2}+b_i\sum_{j=1}^{n_i}y_{ij}-\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{00}+b_i+\beta_{01}t_{ij})\Big]\right\}, & \text{se} \quad i\in\text{placebo} \\ \exp\left\{-\frac{b_i^2}{2\sigma^2}+b_i\sum_{j=1}^{n_i}y_{ij}-\sum_{j=1}^{n_i}\Big[\theta_{ij}\exp(\beta_{10}+b_i+\beta_{11}t_{ij})\Big]\right\}, & \text{se} \quad i\in\text{tratamento.} \end{cases}$$

Sendo $\boldsymbol{b}_{(i)} = (b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_N)$.

Distribuição condicional completa a posteriori para α_2:

$$\pi(\alpha_2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \text{Gama}(c,d)\psi_{18}(\alpha_2,\boldsymbol{\theta})$$

em que

$$\begin{aligned} \psi_{18}(\alpha_2,\boldsymbol{\theta}) &= \exp\left\{-\alpha_2^{-1}\ln\alpha_2\sum_{i=1}^{N}n_i - \ln\Gamma(\alpha_2^{-1})\sum_{i=1}^{N}n_i\right\} \\ &\times \exp\left\{(\alpha_2^{-1}-1)\sum_{i=1}^{N}\sum_{j=1}^{n_i}\ln\theta_{ij} - \alpha_2^{-1}\sum_{i=1}^{N}\sum_{j=1}^{n_i}\theta_{ij}\right\}. \end{aligned}$$

Distribuição condicional completa a posteriori para σ^2:

$$\pi(\sigma^2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \text{GamaInversa}\left(\frac{N}{2}+g, h+\sum_{i=1}^{N}\frac{b_i^2}{2}\right).$$

Anexo D: Distribuições condicionais a posteriori para modelos Bernoulli

A seguir, serão apresentadas as distribuições condicionais a posteriori para os modelos da seção 3.2.3. Seja A o grupo de indivíduos que receberam o tratamento A e B o grupo de indivíduos que receberam o tratamento B.

Modelo 1: Bernoulli

Considerando a função de ligação logística tem-se:

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_1(\beta_{00}, \beta_{01}, \boldsymbol{y}, \boldsymbol{t}) N(a, b^2)$$

em que

$$\psi_1(\beta_{00}, \beta_{01}, \boldsymbol{y}, \boldsymbol{t}) = \exp(\beta_{00} \sum_{i \in A} \sum_{j=1}^{n_i} y_{ij}) \prod_{i \in A} \prod_{j=1}^{n_i} \left[1 + \exp(\beta_{00} + \beta_{01} t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00}, \beta_{10}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_2(\beta_{00}, \beta_{01}, \boldsymbol{y}, \boldsymbol{t}) N(a, b^2)$$

em que

$$\psi_2(\beta_{00}, \beta_{01}, \boldsymbol{y}, \boldsymbol{t}) = \exp(\beta_{01} \sum_{i \in A} \sum_{j=1}^{n_i} y_{ij} t_j) \prod_{i \in A} \prod_{j=1}^{n_i} \left[1 + \exp(\beta_{00} + \beta_{01} t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00}, \beta_{01}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_3(\beta_{10}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) N(a, b^2)$$

em que

$$\psi_3(\beta_{10}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) = \exp(\beta_{10} \sum_{i \in B} \sum_{j=1}^{n_i} y_{ij}) \prod_{i \in B} \prod_{j=1}^{n_i} \left[1 + \exp(\beta_{10} + \beta_{11} t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00}, \beta_{01}, \beta_{10}, \boldsymbol{y}, \boldsymbol{t}) \propto \psi_4(\beta_{10}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) N(a, b^2)$$

em que

$$\psi_4(\beta_{10}, \beta_{11}, \boldsymbol{y}, \boldsymbol{t}) = \exp(\beta_{11} \sum_{i \in B} \sum_{j=1}^{n_i} y_{ij} t_j) \prod_{i \in B} \prod_{j=1}^{n_i} \left[1 + \exp(\beta_{10} + \beta_{11} t_j)\right]^{-1}.$$

Considerando a função de ligação probit, em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$, tem-se:

Distribuição condicional completa a posteriori para β_{00} e β_{01}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) \propto \psi_5(\beta_{00},\beta_{01},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_5(\beta_{00},\beta_{01},\boldsymbol{y},\boldsymbol{t}) = \prod_{i\in A}\prod_{j=1}^{n_i}\big[\Phi(\beta_{00}+\beta_{01}t_j)\big]^{y_{ij}}\big[1-\Phi(\beta_{00}+\beta_{01}t_j)^{y_{ij}}\big]^{1-y_{ij}}$$

Distribuição condicional completa a posteriori para β_{10} e β_{11}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{y},\boldsymbol{t}) \propto \psi_6(\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_6(\beta_{10},\beta_{11},\boldsymbol{y},\boldsymbol{t}) = \prod_{i\in B}\prod_{j=1}^{n_i}\big[\Phi(\beta_{10}+\beta_{11}t_j)\big]^{y_{ij}}\big[1-\Phi(\beta_{10}+\beta_{11}t_j)^{y_{ij}}\big]^{1-y_{ij}}$$

Modelo 2: Bernoulli-beta

Considerando a função de ligação logística tem-se:

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_7(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_7(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\Big(\frac{e^{\beta_{00}}}{1+e^{\beta_{00}+\beta_{01}t_j}}\Big)^{y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{00}+\beta_{01}t_j}}{1+e^{\beta_{00}+\beta_{01}t_j}}\Big)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_8(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_8(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\Big(\frac{e^{\beta_{01}t_j}}{1+e^{\beta_{00}+\beta_{01}t_j}}\Big)^{y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{00}+\beta_{01}t_j}}{1+e^{\beta_{00}+\beta_{01}t_j}}\Big)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_9(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_9(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\Big(\frac{e^{\beta_{10}}}{1+e^{\beta_{10}+\beta_{11}t_j}}\Big)^{y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{10}+\beta_{11}t_j}}{1+e^{\beta_{10}+\beta_{11}t_j}}\Big)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{10}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{10}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\Big(\frac{e^{\beta_{11}t_j}}{1+e^{\beta_{10}+\beta_{11}t_j}}\Big)^{y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{10}+\beta_{11}t_j}}{1+e^{\beta_{10}+\beta_{11}t_j}}\Big)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para θ_{ij}, $i = 1,\ldots,N$, $j = 1,\ldots,n_i$:

$$\pi(\theta_{ij}|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta}_{(ij)},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{11}(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})\text{Beta}(y_{ij}+\alpha_1,\alpha_2)$$

em que

$$\psi_{11}(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) \propto \begin{cases} \Big(1-\theta_{ij}\frac{e^{\beta_{00}+\beta_{01}t_j}}{1+e^{\beta_{00}+\beta_{01}t_j}}\Big)^{1-y_{ij}}, & \text{se} \quad i\in \text{A} \\ \Big(1-\theta_{ij}\frac{e^{\beta_{10}+\beta_{11}t_j}}{1+e^{\beta_{10}+\beta_{11}t_j}}\Big)^{1-y_{ij}}, & \text{se} \quad i\in \text{B}. \end{cases}$$

Sendo $\boldsymbol{\theta}_{(ij)} = (\theta_{11},\ldots,\theta_{1n_1},\ldots,\theta_{i1},\ldots,\theta_{ij-1},\theta_{ij+1},\ldots,\theta_{in_i},\ldots,\theta_{N1},\ldots,\theta_{Nn_N})$.

Distribuição condicional completa a posteriori para α_1:

$$\pi(\alpha_1|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \text{Gama}(c,d)B(\alpha_1,\alpha_2)\psi_{12}(\alpha_1,\boldsymbol{\theta})$$

em que

$$\psi_{12}(\alpha_1,\boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_1-1}.$$

Distribuição condicional completa a posteriori para α_2:

$$\pi(\alpha_2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_1,\boldsymbol{y},\boldsymbol{t}) \propto \text{Gama}(c,d)B(\alpha_1,\alpha_2)\psi_{13}(\alpha_2,\boldsymbol{\theta})$$

em que

$$\psi_{13}(\alpha_2,\boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_2-1}.$$

Considerando a função de ligação probit, em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$, tem-se:

Distribuição condicional completa a posteriori para β_{00} e β_{01}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{14}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{14}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\big[\Phi(\beta_{00}+\beta_{01}t_j)\big]^{y_{ij}}\big[1-\theta_{ij}\Phi(\beta_{00}+\beta_{01}t_j)\big]^{y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{10} e β_{11}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{15}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{15}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\big[\Phi(\beta_{10}+\beta_{11}t_j)\big]^{y_{ij}}\big[1-\theta_{ij}\Phi(\beta_{10}+\beta_{11}t_j)\big]^{y_{ij}}.$$

Distribuição condicional completa a posteriori para θ_{ij}, $i = 1,\ldots,N$, $j = 1,\ldots,n_i$:

$$\pi(\theta_{ij}|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta}_{(ij)},\alpha_1,\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{16}(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t})\mathrm{Beta}(y_{ij}+\alpha_1,\alpha_2)$$

em que

$$\psi_{16}(\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{t}) \propto \begin{cases} \big[1-\theta_{ij}\Phi(\beta_{00}+\beta_{01}t_j)\big]^{1-y_{ij}}, & \text{se} \quad i\in \mathrm{A} \\ \big[1-\theta_{ij}\Phi(\beta_{10}+\beta_{11}t_j)\big]^{1-y_{ij}}, & \text{se} \quad i\in \mathrm{B}. \end{cases}$$

Sendo $\boldsymbol{\theta}_{(ij)} = (\theta_{11},\ldots,\theta_{1n_1},\ldots,\theta_{i1},\ldots,\theta_{ij-1},\theta_{ij+1},\ldots,\theta_{in_i},\ldots,\theta_{N1},\ldots,\theta_{Nn_N})$.

Distribuição condicional completa a posteriori para α_1:

$$\pi(\alpha_1|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \mathrm{Gama}(c,d)B(\alpha_1,\alpha_2)\psi_{17}(\alpha_1,\boldsymbol{\theta})$$

em que

$$\psi_{17}(\alpha_1,\boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_1-1}.$$

Distribuição condicional completa a posteriori para α_2:

$$\pi(\alpha_2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\alpha_1,\boldsymbol{y},\boldsymbol{t}) \propto \text{Gama}(c,d)B(\alpha_1,\alpha_2)\psi_{18}(\alpha_2,\boldsymbol{\theta})$$

em que

$$\psi_{18}(\alpha_2,\boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_2-1}.$$

Modelo 3: Bernoulli-normal

Considerando a função de ligação logística tem-se:

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{19}(\beta_{00},\beta_{01},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{19}(\beta_{00},\beta_{01},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp(\beta_{00}\sum_{i\in A}\sum_{j=1}^{n_i}y_{ij})\prod_{i\in A}\prod_{j=1}^{n_i}\left[1+\exp(\beta_{00}+b_i+\beta_{01}t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{20}(\beta_{00},\beta_{01},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{20}(\beta_{00},\beta_{01},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp(\beta_{01}\sum_{i\in A}\sum_{j=1}^{n_i}y_{ij}t_j)\prod_{i\in A}\prod_{j=1}^{n_i}\left[1+\exp(\beta_{00}+b_i+\beta_{01}t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{21}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{21}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp(\beta_{10}\sum_{i\in B}\sum_{j=1}^{n_i}y_{ij})\prod_{i\in B}\prod_{j=1}^{n_i}\left[1+\exp(\beta_{10}+b_i+\beta_{11}t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{22}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{22}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \exp(\beta_{11}\sum_{i\in B}\sum_{j=1}^{n_i} y_{ij}t_j)\prod_{i\in B}\prod_{j=1}^{n_i}\left[1+\exp(\beta_{10}+b_i+\beta_{11}t_j)\right]^{-1}.$$

Distribuição condicional completa a posteriori para b_i, $i=1,\ldots,N$:

$$\pi(b_i|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b}_{(i)},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \exp\left(-\frac{b_i^2}{2\sigma^2}+\sum_{j=1}^{n_i} b_i y_{ij}\right)\prod_{j=1}^{n_i}\left[1+\exp(\beta_{00}+b_i+\beta_{01}t_j)\right]^{-1}, & \text{se} \quad i\in A \\ \exp\left(-\frac{b_i^2}{2\sigma^2}+\sum_{j=1}^{n_i} b_i y_{ij}\right)\prod_{j=1}^{n_i}\left[1+\exp(\beta_{00}+b_i+\beta_{01}t_j)\right]^{-1}, & \text{se} \quad i\in B. \end{cases}$$

Sendo $\boldsymbol{b}_{(i)} = (b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_N)$.

Distribuição condicional completa a posteriori para σ^2:

$$\pi(\sigma^2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) \propto \text{GamaInversa}\left(\frac{N}{2}+g, h+\sum_{i=1}^{N}\frac{b_i^2}{2}\right).$$

Considerando a função de ligação probit, em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\left[-t^2/2\right]dt$, tem-se:

Distribuição condicional completa a posteriori para β_{00} e β_{01}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{23}(\beta_{00},\beta_{01},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{23}(\beta_{00},\beta_{01},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i\in A}\prod_{j=1}^{n_i}\left[\Phi(\beta_{00}+b_i+\beta_{01}t_j)\right]^{y_{ij}}\left[1-\Phi(\beta_{00}+b_i+\beta_{01}t_j)^{y_{ij}}\right]^{1-y_{ij}}$$

Distribuição condicional completa a posteriori para β_{10} e β_{11}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{b},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{24}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{24}(\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i\in A}\prod_{j=1}^{n_i}\left[\Phi(\beta_{10}+b_i+\beta_{11}t_j)\right]^{y_{ij}}\left[1-\Phi(\beta_{10}+b_i+\beta_{11}t_j)^{y_{ij}}\right]^{1-y_{ij}}$$

Distribuição condicional completa a posteriori para b_i, $i=1,\ldots,N$:

$$\pi(b_i|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b}_{(i)},\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \exp\big(-\frac{b_i^2}{2\sigma^2}+\sum_{j=1}^{n_i} b_i y_{ij}\big)\prod_{j=1}^{n_i}\big[1+\exp(\beta_{00}+b_i+\beta_{01}t_j)\big]^{-1}, & \text{se} \quad i\in A \\ \exp\big(-\frac{b_i^2}{2\sigma^2}+\sum_{j=1}^{n_i} b_i y_{ij}\big)\prod_{j=1}^{n_i}\big[1+\exp(\beta_{00}+b_i+\beta_{01}t_j)\big]^{-1}, & \text{se} \quad i\in B. \end{cases}$$

Sendo $\boldsymbol{b}_{(i)} = (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_N)$.

Distribuição condicional completa a posteriori para σ^2:

$$\pi(\sigma^2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) \propto \text{GamaInversa}\left(\frac{N}{2}+g, h+\sum_{i=1}^{N}\frac{b_i^2}{2}\right).$$

Modelo 4: Bernoulli-normal-beta

Considerando a função de ligação logística tem-se:

Distribuição condicional completa a posteriori para β_{00}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{25}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{25}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\left(\frac{e^{\beta_{00}}}{1+e^{\beta_{00}+b_i+\beta_{01}t_j}}\right)^{y_{ij}}\left(1-\theta_{ij}\frac{e^{\beta_{00}+b_i+\beta_{01}t_j}}{1+e^{\beta_{00}+b_i+\beta_{01}t_j}}\right)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{01}:

$$\pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{26}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{26}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\left(\frac{e^{\beta_{01}t_j}}{1+e^{\beta_{00}+b_i+\beta_{01}t_j}}\right)^{y_{ij}}\left(1-\theta_{ij}\frac{e^{\beta_{00}+b_i+\beta_{01}t_j}}{1+e^{\beta_{00}+b_i+\beta_{01}t_j}}\right)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{10}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{27}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{27}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\left(\frac{e^{\beta_{10}}}{1+e^{\beta_{10}+b_i+\beta_{11}t_j}}\right)^{y_{ij}}\left(1-\theta_{ij}\frac{e^{\beta_{10}+b_i+\beta_{11}t_j}}{1+e^{\beta_{10}+b_i+\beta_{11}t_j}}\right)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{11}:

$$\pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{10},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{28}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{28}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\Big(\frac{e^{\beta_{11}t_j}}{1+e^{\beta_{10}+b_i+\beta_{11}t_j}}\Big)^{y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{10}+b_i+\beta_{11}t_j}}{1+e^{\beta_{10}+b_i+\beta_{11}t_j}}\Big)^{1-y_{ij}}.$$

Distribuição condicional completa a posteriori para θ_{ij}, $i = 1,\ldots,N$, $j = 1,\ldots,n_i$:

$$\pi(\theta_{ij}|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta}_{(ij)},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{29}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})\text{Beta}(y_{ij}+\alpha_1,\alpha_2)$$

em que

$$\psi_{29}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) \propto \begin{cases} \Big(1-\theta_{ij}\frac{e^{\beta_{00}+b_i+\beta_{01}t_j}}{1+e^{\beta_{00}+b_i+\beta_{01}t_j}}\Big)^{1-y_{ij}}, & \text{se} \quad i\in \text{A} \\ \Big(1-\theta_{ij}\frac{e^{\beta_{10}+b_i+\beta_{11}t_j}}{1+e^{\beta_{10}+b_i+\beta_{11}t_j}}\Big)^{1-y_{ij}}, & \text{se} \quad i\in \text{B}. \end{cases}$$

Sendo $\boldsymbol{\theta}_{(ij)} = (\theta_{11},\ldots,\theta_{1n_1},\ldots,\theta_{i1},\ldots,\theta_{ij-1},\theta_{ij+1},\ldots,\theta_{in_i},\ldots,\theta_{N1},\ldots,\theta_{Nn_N})$.

Distribuição condicional completa a posteriori para b_i, $i = 1,\ldots,N$:

$$\pi(b_i|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b}_{(i)},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \prod_{j=1}^{n_i} e^{b_iy_{ij}-\frac{b_i^2}{2\sigma^2}}\left(1+e^{\beta_{00}+b_i+\beta_{01}t_j}\right)^{-y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{00}+b_i+\beta_{01}t_j}}{1+e^{\beta_{00}+b_i+\beta_{01}t_j}}\Big)^{1-y_{ij}}, & \text{se} \quad i\in \text{A} \\ \prod_{j=1}^{n_i} e^{b_iy_{ij}-\frac{b_i^2}{2\sigma^2}}\left(1+e^{\beta_{10}+b_i+\beta_{11}t_j}\right)^{-y_{ij}}\Big(1-\theta_{ij}\frac{e^{\beta_{10}+b_i+\beta_{11}t_j}}{1+e^{\beta_{10}+b_i+\beta_{11}t_j}}\Big)^{1-y_{ij}}, & \text{se} \quad i\in \text{B}. \end{cases}$$

Sendo $\boldsymbol{b}_{(i)} = (b_1,\ldots,b_{i-1},b_{i+1},\ldots,b_N)$.

Distribuição condicional completa a posteriori para α_1:

$$\pi(\alpha_1|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \text{Gama}(c,d)B(\alpha_1,\alpha_2)\psi_{30}(\alpha_1,\boldsymbol{\theta})$$

em que

$$\psi_{30}(\alpha_1,\boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_1-1}.$$

Distribuição condicional completa a posteriori para α_2:

$$\pi(\alpha_2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \text{Gama}(c,d)B(\alpha_1,\alpha_2)\psi_{31}(\alpha_2,\boldsymbol{\theta})$$

em que

$$\psi_{31}(\alpha_2,\boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_2-1}.$$

Distribuição condicional completa a posteriori para σ^2:

$$\pi(\sigma^2|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_2,\boldsymbol{y},\boldsymbol{t}) \propto \text{GamaInversa}\left(\frac{N}{2}+g, h+\sum_{i=1}^{N}\frac{b_i^2}{2}\right).$$

Considerando a função de ligação probit, em que $\Phi(x) = \int_{-\infty}^{x}(2\pi)^{-1/2}\exp\big[-t^2/2\big]dt$, tem-se:

Distribuição condicional completa a posteriori para β_{00} e β_{01}:

$$\pi(\beta_{00}|\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{01}|\beta_{00},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{32}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{32}(\beta_{00},\beta_{01},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\big[\Phi(\beta_{00}+b_i+\beta_{01}t_j)\big]^{y_{ij}}\big[1-\theta_{ij}\Phi(\beta_{00}+b_i+\beta_{01}t_j)\big]^{y_{ij}}.$$

Distribuição condicional completa a posteriori para β_{10} e β_{11}:

$$\pi(\beta_{10}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) = \pi(\beta_{11}|\beta_{00},\beta_{01},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\alpha_1,\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{33}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})N(a,b^2)$$

em que

$$\psi_{33}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\big[\Phi(\beta_{10}+b_i+\beta_{11}t_j)\big]^{y_{ij}}\big[1-\theta_{ij}\Phi(\beta_{10}+b_i+\beta_{11}t_j)\big]^{y_{ij}}.$$

Distribuição condicional completa a posteriori para θ_{ij}, $i = 1,\ldots,N$, $j = 1,\ldots,n_i$:

$$\pi(\theta_{ij}|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta}_{(ij)},\boldsymbol{b},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto \psi_{34}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t})\text{Beta}(y_{ij}+\alpha_1,\alpha_2)$$

em que

$$\psi_{34}(\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b},\boldsymbol{y},\boldsymbol{t}) \propto \begin{cases} \big[1-theta_{ij}\Phi(\beta_{00}+b_i+\beta_{01}t_j)\big]^{1-y_{ij}}, & \text{se} \quad i\in \text{A} \\ \big[1-theta_{ij}\Phi(\beta_{10}+b_i+\beta_{11}t_j)\big]^{1-y_{ij}}, & \text{se} \quad i\in \text{B}. \end{cases}$$

Sendo $\boldsymbol{\theta}_{(ij)} = (\theta_{11},\ldots,\theta_{1n_1},\ldots,\theta_{i1},\ldots,\theta_{ij-1},\theta_{ij+1},\ldots,\theta_{in_i},\ldots,\theta_{N1},\ldots,\theta_{Nn_N})$.

Distribuição condicional completa a posteriori para b_i, $i = 1,\ldots,N$:

$$\pi(b_i|\beta_{00},\beta_{01},\beta_{10},\beta_{11},\boldsymbol{\theta},\boldsymbol{b}_{(i)},\alpha_2,\sigma^2,\boldsymbol{y},\boldsymbol{t}) \propto$$

$$\propto \begin{cases} \exp\left\{-\frac{b_i}{2\sigma^2}\right\}\prod_{j=1}^{n_i}\big[\Phi(\beta_{00}+b_i+\beta_{01}t_j)\big]^{y_{ij}}\big[1-\theta_{ij}\Phi(\beta_{00}+b_i+\beta_{01}t_j)\big]^{y_{ij}}, & \text{se} \quad i\in \text{A} \\ \exp\left\{-\frac{b_i}{2\sigma^2}\right\}\prod_{j=1}^{n_i}\big[\Phi(\beta_{10}+b_i+\beta_{11}t_j)\big]^{y_{ij}}\big[1-\theta_{ij}\Phi(\beta_{10}+b_i+\beta_{11}t_j)\big]^{y_{ij}}, & \text{se} \quad i\in \text{B}. \end{cases}$$

Sendo $\boldsymbol{b}_{(i)} = (b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_N)$.

Distribuição condicional completa a posteriori para α_1:

$$\pi(\alpha_1|\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \boldsymbol{b}, \alpha_2, \sigma^2, \boldsymbol{y}, \boldsymbol{t}) \propto \text{Gama}(c, d)B(\alpha_1, \alpha_2)\psi_{35}(\alpha_1, \boldsymbol{\theta})$$

em que

$$\psi_{35}(\alpha_1, \boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_1-1}.$$

Distribuição condicional completa a posteriori para α_2:

$$\pi(\alpha_2|\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \boldsymbol{b}, \alpha_1, \sigma^2, \boldsymbol{y}, \boldsymbol{t}) \propto \text{Gama}(c, d)B(\alpha_1, \alpha_2)\psi_{36}(\alpha_2, \boldsymbol{\theta})$$

em que

$$\psi_{36}(\alpha_2, \boldsymbol{\theta}) = \prod_{i=1}^{N}\prod_{j=1}^{n_i}\theta_{ij}^{\alpha_2-1}.$$

Distribuição condicional completa a posteriori para σ^2:

$$\pi(\sigma^2|\beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}, \boldsymbol{\theta}, \boldsymbol{b}, \alpha_1, \alpha_2, \boldsymbol{y}, \boldsymbol{t}) \propto \text{GamaInversa}\left(\frac{N}{2} + g, h + \sum_{i=1}^{N}\frac{b_i^2}{2}\right).$$

APÊNDICES

Apêndice A: Programas WinBUGS para modelos Poisson

Modelo 1: Poisson

```
model {
    for (i in 1 : N) {
        for(j in 1 : M){
            y[i,j] ~ dpois(lambda[i,j])
            lambda[i,j]< -exp((int0+slope0*t[j])*(1-trt[i])+
            (int1+slope1*t[j])*(trt[i]))
        }
    }
int0 ~ dnorm(0.0,0.000001)
slope0 ~ dnorm(0.0,0.000001)
int1 ~ dnorm(0.0,0.000001)
slope1 ~ dnorm(0.0,0.000001)
ds< - slope1-slope0
rs< - slope1/slope0
}
```

Modelo 2: Poisson-gama

```
model {
    for (i in 1 : N) {
        for(j in 1 : M){
            y[i,j] ~ dpois(lambda[i,j])
            lambda[i,j]< -teta[i,j]*exp((int0+slope0*t[j])*(1-trt[i])+
            (int1+slope1*t[j])*(trt[i]))
            teta[i,j] ~dgamma(alpha,alpha)
        }
    }
int0 ~ dnorm(0.0,0.000001)
slope0 ~ dnorm(0.0,0.000001)
int1 ~ dnorm(0.0,0.000001)
```

```
slope1 ~ dnorm(0.0,0.000001)
alpha< -1/beta
beta ~ dgamma(0.001,0.001)
ds< - slope1-slope0
rs< - slope1/slope0
}
```

Modelo 3: Poisson-normal

```
model {
        for (i in 1 : N) {
                for(j in 1 : M){
                        y[i,j] ~ dpois(lambda[i,j])
                        lambda[i,j]< -exp((int0+b[i]+slope0*t[j])*(1-trt[i])+
                       (int1+b[i]+slope1*t[j])*(trt[i]))
                }
                b[i] ~ dnorm(0,tau)
        }
D < - 1 / tau
tau ~ dgamma(0.0001,0.0001)
int0 ~ dnorm(0.0,0.000001)
slope0 ~ dnorm(0.0,0.000001)
int1 ~ dnorm(0.0,0.000001)
slope1 ~ dnorm(0.0,0.000001)
ds< - slope1-slope0
rs< - slope1/slope0
}
```

Modelo 4: Poisson-normal-gama

```
model {
        for (i in 1 : N) {
                for(j in 1 : M){
                        y[i,j] ~ dpois(lambda[i,j])
```

```
                lambda[i,j]< −teta[i,j]*exp((int0+b[i]+slope0*t[j])*(1-trt[i])+
               (int1+b[i]+slope1*t[j])*(trt[i]))
                teta[i,j] ∼ dgamma(alpha,alpha)
            }
            b[i] ∼ dnorm(0,tau)
        }
alpha< −1/beta
beta ∼ dgamma(0.001,0.001)
D < − 1 / tau
tau ∼ dgamma(0.0001,0.0001)
int0 ∼ dnorm(0.0,0.000001)
slope0 ∼ dnorm(0.0,0.000001)
int1 ∼ dnorm(0.0,0.000001)
slope1 ∼ dnorm(0.0,0.000001)
ds< − slope1-slope0
rs< − slope1/slope0
}
```

Apêndice B: Programas R para modelos Poisson - pacote Dclone

Bibliotecas necessárias:

```
require(dclone)
require(rjags)
require(coda)
```

Definir N, M, y, trt e t

```
dat<-list(N=N,M=M,y=y,trt=trt,t=t)
```

A seguir os passos para cada um dos modelos:

Modelo 1: Poisson

```
p.model <- function(){
        for (i in 1 : N){
            for(j in 1 : M){
                y[i,j] ~ dpois(lambda[i,j])
                lambda[i,j]< −exp((int0+slope0*t[j])*(1-trt[i])+
                (int1+slope1*t[j])*(trt[i]))
                    }
                        }
int0 ~ dnorm(0.0,0.0001)
slope0 ~ dnorm(0.0,0.000001)
int1 ~ dnorm(0.0,0.0001)
slope1 ~ dnorm(0.0,0.000001)
}
```

```
pmod.jg< −dc.fit(dat,c("int0","int1","slope0","slope1","alpha1","alpha2"),p.model,
n.clones=c(1,10,25,50) , multiply="N", unchanged="M",n.adapt = 5000, n.update = 1000,
thin = 50, n.iter = 50000, flavour = "jags")
```

Modelo 2: Poisson-gama

```
pg.model <- function(){
        for (i in 1 : N){
            for(j in 1 : M){
                y[i,j]   dpois(lambda[i,j])
                lambda[i,j]< −teta[i,j]*exp((int0+slope0*t[j])*(1-trt[i])+
                (int1+slope1*t[j])*(trt[i]))
                teta[i,j]∼dgamma(alpha1,alpha1)
                    }
                        }
int0 ∼ dnorm(0.0 , 0.0001)
slope0 ∼ dnorm(0.0 , 0.000001)
int1 ∼ dnorm(0.0 , 0.0001)
slope1 ∼ dnorm(0.0 , 0.000001)
alpha1< −1/alpha2
alpha2 ∼ dgamma(0.001,0.001)
}

pgmod.jg<-dc.fit(dat, c("int0","int1","slope0","slope1","alpha1","alpha2"),pg.model,
n.clones=c(1,10,25,50) , multiply="N", unchanged="M",n.adapt = 5000, n.update = 1000,
thin = 50, n.iter = 50000, flavour = "jags")
```

Modelo 3: Poisson-normal

```
pn.model <- function(){
        for (i in 1 : N){
            for(j in 1 : M){
                y[i,j] ∼ dpois(lambda[i,j])
                lambda[i,j]< −exp((int0+b[i]+slope0*t[j])*(1-trt[i])+
                (int1+b[i]+slope1*t[j])*(trt[i]))
                    }
```

```
                b[i]~dnorm(0,tau)
                                }
int0 ~ dnorm(0.0 , 0.0001)
slope0 ~ dnorm(0.0 , 0.000001)
int1 ~ dnorm(0.0 , 0.0001)
slope1 ~ dnorm(0.0 , 0.000001)
D< -1/tau
tau ~ dgamma(0.0001,0.0001)
}

pnmod.jg<-dc.fit(dat, c("int0","int1","slope0","slope1","D","tau"),pn.model,
n.clones=c(1,10,25,50) , multiply="N", unchanged="M",n.adapt = 5000, n.update = 1000,
thin = 50, n.iter = 50000, flavour = "jags")
```

Modelo 4: Poisson-normal-gama

```
c.model <- function(){
        for (i in 1 : N){
            for(j in 1 : M){
                y[i,j] ~ dpois(lambda[i,j])
                lambda[i,j]< -teta[i,j]*exp((int0+b[i]+slope0*t[j])*(1-trt[i])+
                (int1+b[i]+slope1*t[j])*(trt[i]))
                teta[i,j] dgamma(alpha1,alpha1)
                    }
            b[i] dnorm(0,tau)
                                }
int0 ~ dnorm(0.0 , 0.0001)
slope0 ~ dnorm(0.0 , 0.000001)
int1 ~ dnorm(0.0 , 0.0001)
slope1 ~ dnorm(0.0 , 0.000001)
alpha1< -1/alpha2
```

```
alpha2 ~ dgamma(0.001,0.001)
D< -1/tau
tau ~ dgamma(0.0001,0.0001)
}

cmod.jg<-dc.fit(dat, c("int0","int1","slope0","slope1","alpha1","alpha2","D","tau"),
c.model,n.clones=c(1,10,25,50) , multiply="N", unchanged="M",n.adapt = 5000, n.update
= 1000,
thin = 50, n.iter = 50000, flavour = "jags")
```

Apêndice C: Programas R para modelo Poisson-normal-gama

```
### Funções para ajuste de um modelo Binomial Negativa com efeito Normal

## Função para simular do modelo

simula <- function(beta, phi,tau, X, n.bloco,n.rep){
b <- rep(rnorm(n.bloco,0,sd=tau),each=n.rep)
eta <- X%*%beta + b
mu <- exp(eta)
p <- phi/(mu+phi)
y <- rnbinom(length(p),size=phi,p=p)
saida = data.frame(y=y,X=X,id=rep(1:n.bloco,each=n.rep))
names(saida) <- c("y","Intercepto","Trat","Tempo","Trat:Tempo","id")
return(saida)
}


### Verossimilhança para um bloco
vero.bloco <- function(b,beta,phi,tau,X,y.vetor){
        phi <- exp(phi)
        tau <- exp(tau)
        eta <- as.matrix(X)%*%beta + b
mu <- exp(eta)
p <- phi/(mu+phi)
ll <- exp(sum(dnbinom(y.vetor,size=phi,p=p,log=TRUE)) + dnorm(b,0,sd=tau,log=TRUE))
return(ll)
}

## Verossimilhancao por GAUSS-HERMITE
vero.gauss <- function(b0,b1,b2,b3,phi,tau,dados,grid.pontos){
X.bloco <- split(dados,dados$id)
ll <- c()
for(i in 1:length(X.bloco)){
ll[i] = gauss.hermite(vero.bloco,grid.pontos=grid.pontos,
beta=c(b0,b1,b2,b3),phi=phi,tau=tau,X=X.bloco[[i]][-c(5,6)],y.vetor=X.bloco[[i]]$y)
}
saida <- sum(log(ll))
        print(round(c(b0,b1,b2,b3,phi,tau,saida),3))
return(-saida)
}


### Integração por Gauss Hermite
gauss.hermite <- function(funcao,grid.pontos,...){
  f.nodes <- c()
```

```
  nodes <- grid.pontos$nodes
  normalizacao <- exp(-nodes^2)
  for(i in 1:length(nodes)){
  f.nodes[i] <- funcao(b=nodes[i],...)}
  #f.nodes = apply(as.matrix(nodes),1,funcao,...)
  integral <- sum(grid.pontos$weights*f.nodes/normalizacao)
  return(integral)
}

#########################################################################
h.x <- function(b, beta, phi,tau,X,y.vetor){
  eta <- as.matrix(X)%*%beta + b
  mu <- exp(eta)
  phi <- exp(phi)
  tau <- exp(tau)
  p <- phi/(mu+phi)
  ll <- sum(dnbinom(y.vetor,size=phi,p=p,log=TRUE)) + dnorm(b,0,sd=tau,log=TRUE)
return(ll)
}

integra.laplace <- function(funcao,...){
  integral <- -9999
  temp <- try(optim(0,funcao,...,
method="BFGS", hessian=TRUE,control=list(fnscale=-1)))
  if(class(temp) != "try-error"){
  integral <- exp(temp$value)* (exp(0.5*log(2*pi) - 0.5*log(-temp$hessian)))}
   return(integral)
}

vero.laplace <- function(b0,b1,b2,b3,phi,tau,dados){
X.bloco <- split(dados,dados$id)
ll <- c()
for(i in 1:length(X.bloco)){
ll[i] = integra.laplace(h.x,beta=c(b0,b1,b2,b3),phi=phi,
tau=tau,X=X.bloco[[i]][-c(5,6)],y.vetor=X.bloco[[i]]$y) }
saida <- sum(log(ll))
return(-saida)
}

#################################################################
## Programa a função marginal ######################################
#################################################################
simula.poisson <- function(b0,tau,n.bloco,n.rep){
  b <- rep(rnorm(n.bloco,0,sd=tau))
  eta <- b0 + b
```

```
  lambda <- exp(eta)
  y <- matrix(NA,n.bloco,n.rep)
  for(i in 1:n.bloco){
  y[i,] <- rpois(n.rep,lambda[i])}
  dados <- data.frame(y=as.vector(y),X = rep(1,n.bloco*n.rep),
  id=rep(1:n.bloco,each=n.rep))
  return(dados)
}


dentro <- function(t,b0,tau,X,y){
saida <- c()
  for( i in 1:length(t)) {
  saida[i]<-(-1)^t[i]/factorial(t[i])*exp(X*b0*t[i])*exp(tau*y*t[i])*exp(0.5*tau*t[i]^2)
  }
  return(sum(saida))
}


vero <- function(b0,tau,X,y){
ll=(-log(factorial(y)))*(X*b0*y + 0.5*tau*y^2)+log(dentro(t=0:4,b0=b0,tau=tau,X=X,y=y))
return(ll)
}

vero.completa <- function(b0,tau,dados){
dados.bloco =   split(dados,dados$id)
n.rep <- dim(dados.bloco[[1]])[1]
n.bloco <- length(dados.bloco)
ll <- matrix(NA,ncol=n.rep,nrow=n.bloco)
for(j in 1:n.bloco){
  for(i in 1:n.rep){
     ll[j,i] <- vero(b0=b0,tau=tau,X=dados.bloco[[j]]$X[i],y=dados.bloco[[j]]$y[i])
}}
return(-sum(ll))}

##############################

## Carregando pacotes adicionais
require(statmod)
require(fOptions)
require(bbmle)

## Lendo os dados reais
dados <- read.table("dadosAJUSTE.txt",header=TRUE)
head(dados)
```

```
## Avaliando a Verossimilhança em alguns pontos

## Pontos para a integração por GH e QMC
pontos <- rnorm.halton(30,1)
grid.pontos <- gauss.quad(30,kind="hermite")

vero.gauss(b0=0.9,b1=-0.3,b2=-0.02,   b3=0.01, phi=2.5,tau=0.9,
 dados=dados,grid.pontos=grid.pontos)
vero.qmc(b0=0.9,b1=-0.3,b2=-0.02,     b3=0.01, phi=2.5,tau=0.9,
 dados=dados, pontos= pontos)
vero.laplace(b0=0.9,b1=-0.3,b2=-0.02, b3=0.01, phi=2.5,tau=0.9,
 dados=dados)

system.time(vero.gauss(b0=0.25,b1=-0.3,b2=-0.02,b3=0.01,phi=2.5,
tau=sqrt(1.1),dados=dados,grid.pontos=grid.pontos))
system.time(vero.qmc(b0=0.9,b1=-0.3,b2=-0.02,b3=0.01,phi=2.5,
tau=sqrt(1.1),dados=dados, pontos= pontos))
system.time(vero.laplace(b0=0.25,b1=-0.3,b2=-0.02,b3=0.01,
phi=2.5,tau=sqrt(1.1),dados=dados))

## Otimizando via bbmle
## Gauss-Hermite
modelo.gauss <- mle2(vero.gauss, start=list(b0=0, b1=0, b2=0,
 b3=0, phi=log(1), tau=log(1)),
                     data=list(dados=dados,grid.pontos=grid.pontos))
summary(modelo.gauss)
perf.gauss <- profile(modelo.gauss)
plot(perf.gauss)
ic.quad.gauss = confint(modelo.gauss,method="quad")

## Laplace
modelo.laplace <- mle2(vero.laplace, start=list(b0=0, b1=0,
 b2=0, b3=0, phi=log(1), tau=log(1)),data=list(dados=dados))
summary(modelo.laplace)
perf.laplace <- profile(modelo.laplace)
confint(modelo.laplace,method="quad")
```

Apêndice D: Programas WinBUGS para modelos Bernoulli

Modelo 1: Bernoulli

```
model {
        for (i in 1 : N) {
                for(j in 1 : M){
                        y[i,j] ~ dbern(pi[i,j])
                        logit(pi[i,j])<-((int0+slope0*t[j])*(1-trt[i])+(int1+slope1*t[j])*(trt[i]))
                }
        }
int0 ~ dnorm(0.0,0.000001)
slope0 ~ dnorm(0.0,0.000001)
int1 ~ dnorm(0.0,0.000001)
slope1 ~ dnorm(0.0,0.000001)
}
```

Modelo 3: Poisson-normal

```
model {
        for (i in 1 : N) {
                for(j in 1 : M){
                        y[i,j] ~ dbern(pi[i,j])
                        logit(pi[i,j])<-((int0+b[i]+slope0*t[j])*(1-trt[i])+(int1+b[i]+slope1*t[j])*(trt[i]))
                }
                b[i] ~ dnorm(0,tau)
        }
D < − 1 / tau
tau ~ dgamma(0.0001,0.0001)
int0 ~ dnorm(0.0,0.000001)
slope0 ~ dnorm(0.0,0.000001)
int1 ~ dnorm(0.0,0.000001)
slope1 ~ dnorm(0.0,0.000001)
}
```

Apêndice E: Código SAS para modelo Poisson com preditor não-linear logístico

```
proc nlmixed data=folha qpoints=50;
title '1. modelo Poisson';
parms beta11=3.2 beta12=3.6 beta13=4 beta14=3 beta15=3
beta21=-40 beta22=10 beta23=-10 beta24=1 beta25=-60
beta31=83 beta32=84 beta33=85 beta34=86 beta35=87;
eta1=((beta11)/(1+exp(-(tempo-beta21)/beta31)))*(trat1);
eta2=((beta12)/(1+exp(-(tempo-beta22)/beta32)))*(trat2);
eta3=((beta13)/(1+exp(-(tempo-beta23)/beta32)))*(trat3);
eta4=((beta14)/(1+exp(-(tempo-beta24)/beta34)))*(trat4);
eta5=((beta15)/(1+exp(-(tempo-beta25)/beta35)))*(trat5);
eta=eta1+eta2+eta3+eta4+eta5;
lambda = exp(eta);
model y poisson(lambda);
run;
```