

NILTON TSUCHIYA

UTILIZAÇÃO DE TÉCNICAS BAYESIANAS EM MODELOS DE
REGRESSÃO DE POISSON PARA DADOS DE CONTAGEM
LONGITUDINAIS E DADOS DE CONTAGEM COM MEDIDAS
REPETIDAS APRESENTANDO EXCESSO DE ZEROS

São Carlos

Agosto de 2008

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Departamento de Estatística

NILTON TSUCHIYA

UTILIZAÇÃO DE TÉCNICAS BAYESIANAS EM MODELOS DE
REGRESSÃO DE POISSON PARA DADOS DE CONTAGEM
LONGITUDINAIS E DADOS DE CONTAGEM COM MEDIDAS
REPETIDAS APRESENTANDO EXCESSO DE ZEROS

Dissertação apresentada ao Departamento de Estatística da Universidade Federal de São Carlos - DEs/UFSCar, como parte dos requisitos para obtenção do título de Mestre em Estatística.

Orientador: Prof. Dr. Jorge Alberto Achcar

São Carlos

Agosto de 2008

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

T882ut

Tsuchiya, Nilton.

Utilização de técnicas bayesianas em modelos de regressão de Poisson para dados de contagem longitudinais e dados de contagem com medidas repetidas apresentando excesso de zeros / Nilton Tsuchiya. -- São Carlos : UFSCar, 2008.

129 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2008.

1. Análise de regressão. 2. Dados de contagem. 3. Modelos inflacionados de zeros. 4. Modelo bayesiano hierárquico. 5. MCMC. I. Título.

CDD: 519.536 (20^a)

Resumo

Nas pesquisas médicas e biológicas é comum encontrar dados de contagem. Por exemplo, as variáveis podem ser dadas pelo número de hospitalizações para cada paciente em unidades básicas de saúde. Para dados de contagem longitudinais, o uso de modelos de regressão de Poisson usuais, que assumem independência entre as observações, não é satisfatório, visto que as observações de um mesmo indivíduo são usualmente correlacionadas. Assim, efeitos aleatórios são considerados para capturar a possível correlação destas observações além de superdispersão associada a outros fatores. Neste trabalho são utilizados modelos Bayesianos hierárquicos considerando diferentes distribuições aos efeitos aleatórios para analisar tais tipos de dados. Aos efeitos aleatórios é atribuída uma distribuição normal, uma mistura de distribuições normais ou uma distribuição gama. Também são apresentados modelos de regressão de Poisson para dados com medidas repetidas apresentando excesso de zeros. Técnicas de simulação estocástica MCMC (Monte Carlo em Cadeias de Markov) são utilizadas para inferência e, em particular, para estimação dos parâmetros de interesse. Além disso, dados reais são considerados para ilustrar as metodologias propostas.

Palavras-chave: Dados de Contagem Longitudinais; Modelo de Regressão de Poisson; Modelos para Excesso de Zeros; Modelos Hierárquicos; Análise Bayesiana; Métodos MCMC.

Abstract

In medical and biological researches we often find count data. For longitudinal count data, usual Poisson regression models, assuming independence among observations, are not applicable because of the correlation of these measures. This work presents hierarchical Bayesian models considering random effects to analyze longitudinal count data. A Normal and a Gamma distribution are considered to these effects besides the mixture of Normal distributions. We also present zero inflated Poisson (ZIP) regression models for repeated measures. Markov Chain Monte Carlo (MCMC) is used to estimate the parameters.

Keywords: Longitudinal Count Data; Poisson Regression Model; Zero Inflated Model; Hierarchical Model; Bayesian Analysis; MCMC Methods.

Sumário

1	Introdução	1
2	Conceitos em Inferência Bayesiana	4
2.1	Uso da densidade preditiva para verificação da adequabilidade de um modelo	5
2.1.1	Critério AIC	7
2.1.2	Critério BIC	8
2.1.3	Critério DIC	8
2.2	Método de Monte Carlo via cadeias de Markov	9
2.2.1	Amostrador de Gibbs	9
2.2.2	Algoritmo de Metropolis-Hastings	9
2.2.3	Diagnósticos de Convergência	10
3	Metodologia para dados de contagem longitudinais	12
3.1	Modelagem	13
3.1.1	Modelo 1: Distribuição Normal para os efeitos aleatórios	13
3.1.2	Modelo 2: Mistura finita de distribuições Normais para os efeitos aleatórios	15
3.1.3	Modelo 3: Distribuição Gama para os efeitos aleatórios	18
3.2	Variâncias das respostas de contagem longitudinais sob diferentes modelos de efeitos aleatórios	20
3.2.1	Modelo 1	20
3.2.2	Modelo 2	21
3.2.3	Modelo 3	24
4	Metodologia para dados de contagem com medidas repetidas e excesso	

de zeros	25
4.1 Modelo 1: Distribuição Normal para os efeitos aleatórios	26
4.2 Modelo 2: Distribuição Gama para os efeitos aleatórios	29
5 Aplicação 1 - dados de contagem longitudinais	32
5.1 Apresentação e Descrição dos Dados	32
5.2 Ajustes dos modelos propostos no Capítulo 3	36
5.2.1 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Normal para os efeitos aleatórios (Modelo 1)	37
5.2.2 Comparação dos modelos na presença ou não de covariáveis assumindo uma mistura de duas distribuições Normais para os efeitos aleatórios (Modelo 2)	44
5.2.3 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Gama para os efeitos aleatórios (Modelo 3)	50
5.3 Discriminação de modelos	57
5.4 Variâncias não-condicionais das respostas	58
6 Aplicação 2 - dados de contagem inflacionados de zeros	62
6.1 Apresentação e descrição dos dados	62
6.2 Ajustes dos modelos propostos no Capítulo 4	65
6.2.1 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Normal para os efeitos aleatórios (Modelo 1)	66
6.2.2 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Gama para os efeitos aleatórios (Modelo 2)	70
6.3 Discriminação de modelos	73
7 Conclusões e propostas futuras	75
Referências Bibliográficas	79

Apêndices	80
A Código em Winbugs para o Modelo 1 do Capítulo 3	80
B Código em Winbugs para o Modelo 2 do Capítulo 3	92
C Código em Winbugs para o Modelo 3 do Capítulo 3	104
D Código em Winbugs para o Modelo 1 do Capítulo 4	116
E Código em Winbugs para o Modelo 2 do Capítulo 4	123

Capítulo 1

Introdução

Os modelos lineares normais foram utilizados durante muito tempo para a descrição dos fenômenos aleatórios. Mesmo quando a variável resposta parecia não seguir distribuição normal, recorria-se a uma transformação dos dados a fim de se obter este pressuposto. Com o avanço computacional, os métodos iterativos para estimação dos parâmetros começaram a ser utilizados. Assim, os modelos normais não-lineares, assumindo não-linearidade nos parâmetros, tiveram grande avanço. Porém, Nelder & Wedderburn (1972) propuseram uma teoria unificadora da modelagem estatística a que deram o nome de modelos lineares generalizados (MLGs), como uma extensão dos modelos lineares clássicos. Assim, a variável resposta que antes necessitava ser normal, agora poderia ser qualquer uma pertencente à família exponencial de distribuições. Os MLGs incluem, como casos especiais, os modelos de regressão linear e de análise de variância, os modelos logito e probito para dados de proporções, os modelos log-lineares, os modelos de resposta multinomial e alguns modelos comumente usados em dados de sobrevivência.

Dados de contagem são comuns em pesquisas médicas e biológicas. Por exemplo, as variáveis podem ser dadas pelos números de hospitalizações para cada paciente em unidades de saúde. Considerando estes tipos de dados, podemos utilizar um modelo de regressão de Poisson (GREENE, 1993; LINDSEY, 1997) quando as observações são independentes. Para dados de contagem longitudinais, o uso de modelos de regressão de Poisson usuais, que assumem independência entre as observações, não é satisfatório, visto que as observações de um mesmo indivíduo são usualmente correlacionadas. Vários modelos são propostos para dados desta natureza (GOLDSTEIN, 1995; BRYK & RAUN-

DENBUSH, 1992; LONGFORD, 1993). Além disso, modelos que incorporam superdispersão dos dados de contagem longitudinais também vêm sendo propostos (JOWAHEER & SUTRADHAR, 2002; LUCEÑO, 1995).

Para dados longitudinais, em que várias observações são realizadas para um mesmo indivíduo, o uso de modelos de regressão de efeitos mistos (efeitos fixo e aleatório) tem se tornado muito popular (SCHALL, 1991; LAIRD & WARE, 1982; LIANG & ZEGGER, 1986). Neste caso, o efeito aleatório representa o efeito do indivíduo.

Uma outra vertente dos modelos de contagem são os modelos de regressão de Poisson inflacionados de zeros (ZIP). Este modelo assume uma mistura de duas distribuições: uma para as observações de zero com determinada probabilidade e outra para as demais observações. Uma aplicação em defeitos de manufatura é mostrada por Lambert (1992). Hall (2000) apresenta um modelo ZIP com efeitos aleatórios para medidas repetidas.

Além disso, os métodos Bayesianos têm sido amplamente empregados na modelagem dos dados de contagem. Tunaru (2002) apresenta um modelo hierárquico Bayesiano para dados de contagem múltiplos. Vários tipos de modelos de Poisson incluindo efeitos aleatórios têm sido propostos (STUKEL, 1993). Thall (1988) introduz um modelo de regressão misto Poisson-gama para analisar dados discretos em intervalos longitudinais em que um único efeito aleatório é incluído para explicar as respostas correlacionadas dentro do indivíduo sendo que este apresenta distribuição gama na população de indivíduos (ALBERT, 1992; LAWLESS, 1987).

Visto que na abordagem Bayesiana os parâmetros são considerados quantidades aleatórias (BOX & TIAO, 1992; PAULINO et. al., 2003), ou seja, assumem distribuições de probabilidade, o objetivo desta metodologia é a obtenção de medidas resumo ou densidades *a posteriori* para os parâmetros de interesse. Estas densidades são resultantes da combinação de informações da amostra (função de verossimilhança) e de informações *a priori* sobre os parâmetros (densidades *a priori*).

Deste modo, uma grande vantagem da inferência Bayesiana é a possibilidade da incorporação de informações adicionais provenientes de especialistas (de FINETTI et. al., 1974). Isto é muito útil principalmente em estudos da área de saúde, pois, geralmente, o pesquisador tem conhecimento prévio sobre características de uma doença, por exemplo.

Neste trabalho, são propostos modelos Bayesianos com efeitos aleatórios que, além de capturar a correlação de medidas repetidas para um mesmo indivíduo, incorporam

a superdispersão dos dados de contagem longitudinais. A metodologia estatística será ilustrada através de um conjunto de dados reais.

O Capítulo 2 apresenta uma revisão de alguns conceitos importantes em Inferência Bayesiana. O Capítulo 3 apresenta os modelos estatísticos propostos para os dados de contagem longitudinais. O Capítulo 4 apresenta modelos para dados de contagem com medidas repetidas e excesso de zeros. A aplicação da metodologia proposta no Capítulo 3 é feita no Capítulo 5 e o Capítulo 6 apresenta uma aplicação para os modelos propostos no Capítulo 4. As conclusões e perspectivas futuras encontram-se no Capítulo 7.

Capítulo 2

Conceitos em Inferência Bayesiana

A Inferência Clássica busca fazer afirmações ou inferências sobre quantidades de interesse desconhecidas (parâmetros) associadas a um fenômeno aleatório a partir de uma amostra que da mesma foi selecionada. No enfoque Clássico o parâmetro θ , $\theta \in \Theta$ (espaço paramétrico), é tomado como um escalar ou vetor desconhecido, mas fixo. No enfoque Bayesiano o parâmetro θ , $\theta \in \Theta$, é tomado como um escalar ou vetor aleatório (não observável). Deste modo, a filosofia Bayesiana propõe: o que é desconhecido (no caso, o parâmetro θ) é incerto e toda incerteza deve ser quantificada em termos de probabilidade.

Assim, os Bayesianos defendem que a informação inicial ou *a priori* pode ser formalmente representada por uma distribuição de probabilidade para θ , chamada distribuição *a priori*, $\pi(\theta)$.

Seja $\mathbf{y}' = (y_1, \dots, y_n)$ o vetor dos dados observados gerados por uma variável aleatória Y com densidade $f(\mathbf{y} | \theta)$ (caso uniparamétrico). Na abordagem Bayesiana, toda inferência sobre θ é determinada a partir da distribuição *a posteriori* para θ , denotada por $\pi(\theta | \mathbf{y})$.

Utilizando o Teorema de Bayes, temos:

$$\begin{aligned} \pi(\theta | \mathbf{y}) &= \frac{\pi(\theta) f(\mathbf{y} | \theta)}{\int \pi(\theta) f(\mathbf{y} | \theta) d\theta} \\ &\propto \pi(\theta) f(\mathbf{y} | \theta) \end{aligned} \tag{2.1}$$

Em geral, não é necessário o cálculo da integral (uma constante) no denominador e, assim, a *posteriori* é proporcional à *priori* multiplicada pela verossimilhança. No caso

discreto, substitui-se a integral do denominador por \sum_{θ} .

2.1 Uso da densidade preditiva para verificação da adequabilidade de um modelo

Sejam:

$\mathbf{y}'_1 = (y_1, \dots, y_n)$: o vetor dos dados observados;

$\mathbf{y}'_2 = (y_{n+1}, \dots, y_m)$: o vetor das observações futuras;

$\boldsymbol{\theta}' = (\theta_1, \dots, \theta_k)$: o vetor de parâmetros.

Além disso, considere Y_i com função densidade $f_{\boldsymbol{\theta}}(y_i)$.

A densidade preditiva de \mathbf{y}_2 dado \mathbf{y}_1 pode ser escrita como:

$$\begin{aligned} f(\mathbf{y}_2 | \mathbf{y}_1) &= \int_{\Theta} f(\mathbf{y}_2, \boldsymbol{\theta} | \mathbf{y}_1) d\boldsymbol{\theta} = \int_{\Theta} \frac{f(\mathbf{y}_2, \boldsymbol{\theta}, \mathbf{y}_1)}{f(\mathbf{y}_1)} d\boldsymbol{\theta} = \\ &= \int_{\Theta} \frac{f(\mathbf{y}_2, \boldsymbol{\theta}, \mathbf{y}_1)}{f(\mathbf{y}_1, \boldsymbol{\theta})} \frac{f(\mathbf{y}_1, \boldsymbol{\theta})}{f(\mathbf{y}_1)} d\boldsymbol{\theta} = \int_{\Theta} f(\mathbf{y}_2 | \mathbf{y}_1, \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{y}_1) d\boldsymbol{\theta} \end{aligned} \quad (2.2)$$

Supondo \mathbf{Y}_2 independente de \mathbf{Y}_1 , tem-se:

$$f(\mathbf{y}_2 | \mathbf{y}_1) = \int_{\Theta} f(\mathbf{y}_2 | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{y}_1) d\boldsymbol{\theta} \quad (2.3)$$

em que:

$f(\mathbf{y}_2 | \boldsymbol{\theta})$ é a densidade para \mathbf{Y}_2 dado $\boldsymbol{\theta}$;

$f(\boldsymbol{\theta} | \mathbf{y}_1)$ é a posteriori para $\boldsymbol{\theta}$ dado \mathbf{y}_1 .

Sejam $\mathbf{y}' = (y_1, \dots, y_n)$ o vetor de dados observados e $\boldsymbol{\theta}' = (\theta_1, \dots, \theta_k)$ o vetor de parâmetros. Considere também $f(\mathbf{y} | \boldsymbol{\theta})$ a densidade conjunta dos dados, $\pi(\boldsymbol{\theta})$ e $\pi(\boldsymbol{\theta} | \mathbf{y})$ as distribuições *a priori* e *a posteriori* para $\boldsymbol{\theta}$, respectivamente.

Assim, a densidade preditiva para \mathbf{y} é dada por:

$$f(\mathbf{y}) = \int_{\Theta} f(\mathbf{y} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta} \quad (2.4)$$

Para comparação de dois modelos distintos, M_1 e M_2 , pode-se utilizar o fator de Bayes, definido por:

$$BF_{12} = \frac{f(\mathbf{y} | M_1)}{f(\mathbf{y} | M_2)} \quad (2.5)$$

em que $f(\mathbf{y} | M_i)$ representa a densidade preditiva para \mathbf{y} considerando o modelo M_i , $i = 1, 2$.

A Tabela 1 apresenta uma escala de evidência para escolha de M_1 proposta por Jeffreys (1961 apud DONATO & CHEN, 2005):

Tabela 1: Escala de Evidência do Modelo M_1 .

B_{12}	Evidência de M_1
(1, 3.2)	Difícil Decisão
(3.2, 10)	Substancial
(10, 100)	Forte
(100, ∞)	Decisiva

Nota-se que o fator de Bayes requer a especificação de uma distribuição *a priori* $\pi(\boldsymbol{\theta})$ sob os dois modelos. Com distribuições *a priori* impróprias, podemos ter problemas com o uso do fator de Bayes. Por exemplo, se $\pi(\boldsymbol{\theta}_1) \propto c_1$, c_1 arbitrário e se $\pi(\boldsymbol{\theta}_2) \propto c_2$, c_2 arbitrário, tem-se que a razão c_1/c_2 será arbitrária.

Uma sugestão para solucionar este problema é usar a distribuição *a posteriori* no lugar da *priori* no fator de Bayes. Assim, o fator de Bayes *a posteriori* (AITKIN, 1991) é dado por:

$$PostBF_{12} = \frac{\int_{\Theta} f(\mathbf{y} | \boldsymbol{\theta}_1) \pi(\boldsymbol{\theta}_1 | \mathbf{y}) d\boldsymbol{\theta}}{\int_{\Theta} f(\mathbf{y} | \boldsymbol{\theta}_2) \pi(\boldsymbol{\theta}_2 | \mathbf{y}) d\boldsymbol{\theta}} \quad (2.6)$$

Uma outra alternativa para a solução deste problema é a utilização do pseudo fator de Bayes (Geisser & Eddy, 1979; Gelfand et. al. 1992). O pseudo fator de Bayes (PsBF) é definido por:

$$PsBF = \prod_{i=1}^n \frac{f(y_{i,obs} | \mathbf{y}_{(i),obs}, M_1)}{f(y_{i,obs} | \mathbf{y}_{(i),obs}, M_2)} \quad (2.7)$$

em que $f(y_i | \mathbf{y}_{(i)})$ indica a densidade de validação cruzada aplicada em y_i condicional às demais observações com exceção desta. Esta quantidade é também chamada de ordenada preditiva condicional (CPO). Estes valores indicam a verossimilhança de cada observação dadas todas as outras e, portanto, baixos valores de CPO devem corresponder a observações mal ajustadas (PAULINO et. al., 2003). Gelfand et. al. (1992 apud SINHA et.

al. 1999) mostrou que a CPO pode ser calculada por:

$$CPO_i = \left\{ E_{\theta|\mathbf{y}} \left[\frac{1}{f(y_i | \boldsymbol{\theta})} \right] \right\}^{-1} \quad (2.8)$$

Assim, as estimativas das CPO's podem ser aproximadas (LI et. al, 2006) usando uma amostra de $\boldsymbol{\theta}$ de tamanho M , a partir da distribuição *a posteriori* $\pi(\boldsymbol{\theta} | \mathbf{y})$:

$$CPO_i \approx \frac{1}{M} \sum_{j=1}^M [f(y_i | \boldsymbol{\theta}^j)]^{-1} \quad (2.9)$$

em que $\boldsymbol{\theta}^j$ denota o valor de $\boldsymbol{\theta}$ na j -ésima iteração. Isto é, a CPO pode ser aproximada por uma média harmônica de verossimilhanças.

Pode-se também calcular uma medida sumário simples em termos das CPO's. Assim, define-se a pseudo verossimilhança marginal (PML) para o modelo M_k aplicada no logaritmo natural como $LPML_{M_k} = \sum_{i=1}^n \ln(CPO_{i,M_k})$. Escolhe-se o modelo com maior valor de $LPML$.

Além disso, existem alguns critérios baseados em aproximações ao fator de Bayes para comparações de dois modelos, os quais podem ser citados o critério AIC (*Akaike Information Criterion*), o critério BIC (*Bayesian Information Criterion*) e o critério DIC (*Deviance Information Criterion*). Estes critérios incorporam a complexidade do modelo no critério de seleção.

2.1.1 Critério AIC

Seja:

$$\Delta AIC = -2 \ln \left[\frac{\sup_{M_1} f(x | \theta_1, M_1)}{\sup_{M_2} f(x | \theta_2, M_2)} \right] - 2(p_2 - p_1) \quad (2.10)$$

em que p_i , $i = 1, 2$ representa o número de parâmetros de cada modelo. Este critério é baseado em considerações frequentistas de eficiência assintótica, mas só aproxima $-2 \ln BF$ se a informação contida na distribuição *a priori* aumentar na mesma razão que a informação contida na verossimilhança (PAULINO et. al., 2003). Assim, esta situação não é realista do ponto de vista da metodologia Bayesiana e, portanto, não é razoável a sua utilização neste contexto.

2.1.2 Critério BIC

Seja:

$$\Delta BIC = -2 \ln \left[\frac{\sup_{M_1} f(x | \theta_1, M_1)}{\sup_{M_2} f(x | \theta_2, M_2)} \right] - (p_2 - p_1) \ln(n) \quad (2.11)$$

em que n é a dimensão da amostra e p_i , $i = 1, 2$ é o número de parâmetros de cada modelo. Pode-se mostrar que para amostras de dimensão elevada, ΔBIC é uma boa aproximação a $-2 \ln BF$.

Carlin e Louis (2000 apud PAULINO et. al., 2003) sugerem uma modificação a este critério calculando para cada modelo M_i em competição:

$$BIC_i = 2E[\ln L(\theta_i | x, M_i)] - p_i \ln(n) \quad (2.12)$$

escolhendo o modelo que apresenta o maior BIC_i . O problema na utilização deste critério está na definição adequada de n (tamanho da amostra) e de p_i , o número de parâmetros do modelo M_i . Esta dificuldade é encontrada principalmente em modelos hierárquicos.

2.1.3 Critério DIC

A dificuldade encontrada anteriormente levou Spiegelhalter et. al. (2002) a sugerir uma generalização ao critério AIC baseada na distribuição *a posteriori* de $D_i(\theta) = -2 \ln \left[\frac{f(x|\theta, M_i)}{h(x)} \right]$, em que $h(x)$ é uma função apenas dos dados que não tem impacto na escolha do modelo. Propõem como medida da adequabilidade do modelo o valor esperado *a posteriori* de $D_i(\theta)$ e como penalização p_{D_i} (diferença entre este valor esperado e o valor de $D_i(\theta_i)$ calculado no valor esperado *a posteriori* de θ_i) associada à complexidade do modelo. Assim,

$$p_{D_i} = E_{(\theta_i|x, M_i)} [D_i(\theta_i)] - D_i(E_{(\theta_i|x, M_i)}[\theta_i]) \quad (2.13)$$

e

$$DIC_i = E_{(\theta_i|x, M_i)} [D_i(\theta_i)] + p_{D_i} = 2E_{(\theta_i|x, M_i)} [D_i(\theta_i)] - D_i(E_{(\theta_i|x, M_i)}[\theta_i]). \quad (2.14)$$

2.2 Método de Monte Carlo via cadeias de Markov

Deseja-se gerar uma amostra de uma distribuição *a posteriori* $\pi(\boldsymbol{\theta} | \mathbf{y})$, mas não é possível fazer isto diretamente. A idéia do método é baseada na substituição da expressão analítica da densidade por uma amostra gerada a partir desta densidade.

2.2.1 Amostrador de Gibbs

Deseja-se obter inferências da posteriori conjunta $\pi(\boldsymbol{\theta} | \mathbf{y})$. Para isso, simulam-se amostras de $\pi(\boldsymbol{\theta} | \mathbf{y})$ a partir das distribuições condicionais completas $\pi(\theta_i | \mathbf{y}, \boldsymbol{\theta}_{(i)})$, $\boldsymbol{\theta}_{(i)} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_k)$.

Dado um vetor de valores iniciais dos parâmetros $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$, atualizar $\boldsymbol{\theta}^{(0)}$ por $\boldsymbol{\theta}^{(1)}$ a partir do algoritmo:

- (i) Gerar $\theta_1^{(1)}$ de $\pi(\theta_1 | \mathbf{y}, \theta_2^{(0)}, \dots, \theta_k^{(0)})$;
- (ii) Gerar $\theta_2^{(1)}$ de $\pi(\theta_2 | \mathbf{y}, \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)})$;
- (iii) Gerar $\theta_3^{(1)}$ de $\pi(\theta_3 | \mathbf{y}, \theta_1^{(1)}, \theta_2^{(1)}, \theta_4^{(0)}, \dots, \theta_k^{(0)})$;
- ⋮
- (k) Gerar $\theta_k^{(1)}$ de $\pi(\theta_k | \mathbf{y}, \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{k-1}^{(1)})$.

Repetir (i) – (k) um número grande de vezes até a obtenção de convergência.

2.2.2 Algoritmo de Metropolis-Hastings

Quando as distribuições condicionais necessárias para o amostrador de Gibbs não são conhecidas, recorre-se ao algoritmo de Metropolis-Hastings. A idéia básica é definir um núcleo de transição $q(\theta; \beta)$ da distribuição $p(\theta)$ que represente $\pi(\theta_i | \boldsymbol{\theta}_{(i)})$, transformando θ em β .

Algoritmo:

- (i) Iniciar com um valor $\boldsymbol{\theta}^{(0)}$ e um indicador de estágio $j = 0$;
- (ii) Gerar um ponto β de acordo com o núcleo de transição $q(\theta^{(j)}; \beta)$;
- (iii) Atualizar $\theta^{(j)}$ por $\theta^{(j+1)} = \beta$ com probabilidade de transição

$$p = \min \left\{ 1, \frac{p(\beta) q(\theta^{(j)}; \beta)}{p(\theta^{(j)}) q(\beta; \theta^{(j)})} \right\}$$

Ficar com $\theta^{(j)}$ com probabilidade $1 - p$;

(iv) Repetir (i) - (iii) até alcançar uma distribuição estacionária.

O fundamental é escolher um bom $q(\theta^{(j)}; \beta)$. Uma das possibilidades é explorar a forma conhecida da distribuição condicional $\pi(\theta_i | \boldsymbol{\theta}_{(i)}, \mathbf{y})$ para especificar uma densidade geradora de amostras candidatas. (CHIB & GREENBERG, 1995). Por exemplo, $p(t) \propto \psi(t) h(t)$, onde $h(t)$ é uma densidade conhecida que pode ser simulada e $\psi(t)$ é uniformemente limitada; daí, considerar $q(\theta^{(j)}; \beta) = h(\beta)$ para gerar amostras candidatas.

Deste modo, a probabilidade de transição só exige o cálculo da função ψ sendo dada por:

$$p = \min \left\{ 1, \frac{\psi(\beta)}{\psi(\theta^{(j)})} \right\} = \min \left\{ 1, \frac{\psi(\theta^{(j+1)})}{\psi(\theta^{(j)})} \right\}$$

2.2.3 Diagnósticos de Convergência

Para verificar convergência do algoritmo pode-se utilizar:

(a) Gráficos de séries temporais;

A linha deste gráfico deve estar sempre em torno de uma faixa, não apresentando muitas oscilações.

(b) Gráficos de autocorrelação;

O gráfico de autocorrelação dos parâmetros somente deve ter a primeira correlação alta sendo as demais próximas de zero, indicando assim que as amostras geradas são não-correlacionadas.

(c) Índices de convergência, como o índice de Gelman & Rubin (1992).

Gelman e Rubin utilizam procedimentos de análise de variância (ANOVA) para verificar a convergência de cadeias paralelas (com valores iniciais diferentes). A idéia é comparar as variabilidades "entre" e "dentro" das cadeias geradas. O método funciona da seguinte maneira:

Simular $m \geq 2$ cadeias, cada uma de comprimento $2n$, descartando-se as n primeiras amostras;

Calcular a variabilidade entre cadeias

$$\frac{U}{n} = \sum_{i=1}^m (\bar{\theta}_i - \bar{\theta}_{..})^2 / (m-1)$$

em que $\bar{\theta}_i$ é a média das n amostras geradas na i -ésima cadeia; $\bar{\theta}_{..}$ é a média geral.

Define-se, também, a variabilidade dentro de cada cadeia,

$$W = \frac{1}{m} \sum_{i=1}^m s_i^2$$

em que $s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (\theta_{ij} - \bar{\theta}_i)^2$, $i = 1, \dots, m$.

Assim, a variância da distribuição estudada (*posteriori*) pode ser estimada por:

$$\hat{\sigma}^2 = \frac{n-1}{n} W + \frac{1}{n} U$$

Gelman e Rubin mostraram que a distribuição de θ dado y segue distribuição t de Student com centro em $\hat{\mu} = \bar{\theta}_{..}$, desvio-padrão $\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^2 + \frac{U}{mn}}$ e $\nu = \frac{2\hat{V}}{\text{var}(\hat{V})}$ graus de liberdade, em que:

$$\begin{aligned} \text{var}(\hat{V}) &= \left(\frac{n-1}{n}\right)^2 \frac{1}{m} \text{var}(s_i^2) + \left(\frac{m+1}{mn}\right)^2 \frac{2}{m-1} B^2 + \\ &+ 2 \frac{(m-1)(n-1)n}{mn^2} \frac{1}{m} [\text{cov}(s_i^2, \bar{u}_i^2) - 2\bar{u}_{..} \text{cov}(s_i^2, \bar{u}_i)] \end{aligned}$$

Estima-se, então, o fator de redução de escala como:

$$\sqrt{\hat{R}} = \sqrt{\frac{\hat{V}}{W} \frac{\nu}{\nu-2}}$$

Quando n é grande, o fator de redução pode ser simplificado para $\sqrt{\hat{R}} = \sqrt{\frac{\hat{V}}{W}}$. Nota-se, também, que $\sqrt{\hat{R}}$ decresce para 1 quando $n \rightarrow \infty$. Em geral, se $\sqrt{\hat{R}} < 1, 2$, tem-se convergência das cadeias.

Capítulo 3

Metodologia para dados de contagem longitudinais

Assumindo-se uma distribuição de Poisson para X_{li} , com $i = 1, \dots, n$ (indivíduos) e $l = 1, \dots, k$ (tempos), tem-se:

$$P(X_{li} = x_{li}) = \frac{\lambda_{li}^{x_{li}} e^{-\lambda_{li}}}{x_{li}!} \quad (3.1)$$

em que $x_{li} = 0, 1, 2, \dots$; $l = 1, \dots, k$ e $i = 1, 2, \dots, n$. Observa-se que $E(X_{li}) = \text{Var}(X_{li}) = \lambda_{li}$.

Na presença de p covariáveis, assume-se o modelo de regressão de Poisson, dado por:

$$\lambda_{li} = \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i) \quad (3.2)$$

em que $\mathbf{z}'_i = (z_{1i}, z_{2i}, \dots, z_{pi})$ é o vetor de covariáveis do i -ésimo indivíduo ($i = 1, \dots, n$), $\boldsymbol{\beta}'_l = (\beta_{1l}, \beta_{2l}, \dots, \beta_{pl})$ é o vetor de parâmetros do modelo de regressão no l -ésimo tempo ($l = 1, \dots, k$) e w_i o efeito aleatório na i -ésima observação que captura a correlação entre as medidas repetidas para cada indivíduo.

Assume-se que os efeitos aleatórios W_i são independentes e seguem uma distribuição a ser especificada.

Das equações (3.1) e (3.2), segue que a função de verossimilhança para $\boldsymbol{\beta} =$

$(\beta'_1, \beta'_2, \dots, \beta'_k)'$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_k)'$ e $\mathbf{w} = (w_1, w_2, \dots, w_n)'$ é dada por:

$$L(\beta, \alpha, \mathbf{w}) = \prod_{l=1}^k \prod_{i=1}^n \frac{e^{-\lambda_{li}} \lambda_{li}^{x_{li}}}{x_{li}!} \quad (3.3)$$

$$\propto \prod_{l=1}^k \prod_{i=1}^n \{ \exp[-\alpha_l \exp(\mathbf{z}'_i \beta_l + w_i)] \} \times [\alpha_l \exp(\mathbf{z}'_i \beta_l + w_i)]^{x_{li}}$$

3.1 Modelagem

Nesta seção são apresentadas três formas diferentes para a modelagem dos efeitos aleatórios.

3.1.1 Modelo 1: Distribuição Normal para os efeitos aleatórios

Assumindo-se que os efeitos aleatórios W_i seguem uma distribuição Normal com média zero e variância σ_w^2 , ou seja,

$$W_i \sim N(0, \sigma_w^2) \quad (3.4)$$

e, além disso, considerando-se independência *a priori* entre os parâmetros, as seguintes densidades *a priori* para β_{jl} , α_l e σ_w^2 são assumidas:

$$(i) \beta_{jl} \sim N(0, a_{jl}^2), \quad a_{jl} \text{ conhecido}, \quad l = 1, \dots, k \text{ e } j = 1, \dots, p; \quad (3.5)$$

$$(ii) \alpha_l \sim \Gamma(b_l, c_l), \quad b_l \text{ e } c_l \text{ conhecidos}, \quad l = 1, \dots, k;$$

$$(iii) \sigma_w^2 \sim I.G.(d, f), \quad d \text{ e } f \text{ conhecidos.}$$

em que $\Gamma(b, c)$ denota uma distribuição gama com média b/c e variância b/c^2 e $I.G.(d, f)$ denota uma distribuição gama invertida com média $f/(d-1)$ e variância $f^2/[(d-1)(d-2)]$.

Combinando-se as expressões (3.3), (3.4) e (3.5), obtém-se a distribuição *a poste-*

riori conjunta para $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\alpha}', \mathbf{w}', \sigma_w^2)'$, dada por:

$$\begin{aligned} \pi(\boldsymbol{\theta} | \mathbf{x}) &\propto \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\beta_{jl}^2}{2a_{jl}^2}} \right\} \times \left\{ \prod_{i=1}^n \frac{1}{\sqrt{\sigma_w^2}} e^{-\frac{w_i^2}{2\sigma_w^2}} \right\} \times \\ &\times \left\{ \prod_{l=1}^k \alpha_l^{b_l-1} e^{-c_l \alpha_l} \right\} \times (\sigma_w^2)^{-(d+1)} e^{-\frac{f}{\sigma_w^2}} \times \\ &\times \left\{ \prod_{i=1}^n \prod_{l=1}^k \left[\exp(-\alpha_l e^{\mathbf{z}'_i \boldsymbol{\beta}_l + w_i}) (\alpha_l e^{\mathbf{z}'_i \boldsymbol{\beta}_l + w_i})^{x_{li}} \right] \right\} \end{aligned} \quad (3.6)$$

sendo \mathbf{x} o vetor de dados observados.

As distribuições condicionais para o amostrador de Gibbs são dadas por:

$$(i) \pi(\beta_{jl} | \boldsymbol{\theta}_{(\beta_{jl})}, \mathbf{x}) \propto \exp \left\{ -\frac{\beta_{jl}^2}{2a_{jl}^2} + \beta_{jl} \sum_{i=1}^n z_{ji} x_{li} - \alpha_l \sum_{i=1}^n \exp(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i) \right\}, \quad (3.7)$$

em que $\boldsymbol{\theta}_{(\beta_{jl})}$ denota o vetor com todos os parâmetros exceto β_{jl} ; $j = 1, \dots, p$; $l = 1, \dots, k$;

$$(ii) \alpha_l | \boldsymbol{\theta}_{(\alpha_l)}, \mathbf{x} \sim \Gamma \left[b_l + \sum_{i=1}^n x_{li}; c_l + \sum_{i=1}^n \exp(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i) \right], \quad (3.8)$$

em que $\boldsymbol{\theta}_{(\alpha_l)}$ denota o vetor com todos os parâmetros exceto α_l , $l = 1, \dots, k$;

$$(iii) \pi(w_i | \boldsymbol{\theta}_{(w_i)}, \mathbf{x}) \propto \exp \left\{ -\frac{w_i^2}{2\sigma_w^2} + w_i \sum_{l=1}^k x_{li} - e^{w_i} \sum_{l=1}^k \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l) \right\}, \quad (3.9)$$

em que $\boldsymbol{\theta}_{(w_i)}$ denota o vetor com todos os parâmetros exceto w_i , $i = 1, \dots, n$;

$$(iv) \sigma_w^2 | \boldsymbol{\theta}_{(\sigma_w^2)}, \mathbf{x} \sim I.G. \left[d + \frac{n}{2}; f + \frac{1}{2} \sum_{i=1}^n w_i^2 \right], \quad (3.10)$$

em que $\boldsymbol{\theta}_{(\sigma_w^2)}$ denota o vetor com todos os parâmetros exceto σ_w^2 .

Nota-se que as distribuições condicionais em (3.7) e em (3.9) de β_{jl} e w_i apresentam forma desconhecida e, portanto, devem ser geradas utilizando o algoritmo *Metropolis-Hastings*.

3.1.2 Modelo 2: Mistura finita de distribuições Normais para os efeitos aleatórios

Como alternativa para o ajuste de dados longitudinais, considera-se agora o modelo de Poisson assumindo uma mistura finita de distribuições Normais (DIEBOLT & ROBERT, 1994) para os efeitos aleatórios W_i , $i = 1, \dots, n$, dada por:

$$\pi(w_i) = \sum_{r=1}^R p_r \phi_r(w_i | \mu_r, \sigma_r^2) \quad (3.11)$$

em que $\sum_{r=1}^R p_r = 1$ e ϕ_r denota uma densidade Normal com média μ_r e variância σ_r^2 , ou seja, $N(\mu_r, \sigma_r^2)$.

Aqui, consideram-se as mesmas densidades *a priori* (3.5) para β_{jl} e α_l , $l = 1, \dots, k$ e $j = 1, \dots, p$. Além disso, são assumidas também as seguintes distribuições *a priori* para μ_r , σ_r^2 e $\mathbf{p} = (p_1, \dots, p_R)'$:

$$\begin{aligned} (i) \quad \mu_r &\sim N(f_r, g_r^2), \quad f_r \text{ e } g_r \text{ conhecidos, } r = 1, \dots, R \\ (ii) \quad \sigma_r^2 &\sim I.G.(h_r, q_r), \quad h_r \text{ e } q_r \text{ conhecidos, } r = 1, \dots, R \\ (iii) \quad \mathbf{p} &\sim D(s_1, \dots, s_R), \quad s_1, \dots, s_R \text{ conhecidos} \end{aligned} \quad (3.12)$$

Nota: $D(s_1, \dots, s_R)$ denota uma distribuição de Dirichlet.

Assumindo-se independência *a priori* entre os parâmetros e combinando-se as expressões (3.3), (3.5), (3.11) e (3.12), obtém-se a distribuição *a posteriori* conjunta para $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\sigma}^2, \mathbf{p})'$ dada por:

$$\pi(\boldsymbol{\theta} | \mathbf{x}) \propto \psi(\boldsymbol{\theta}) \left\{ \prod_{i=1}^n \sum_{r=1}^R p_r \phi_r(w_i | \mu_r, \sigma_r^2) \right\} \quad (3.13)$$

em que:

$$\begin{aligned} \psi(\boldsymbol{\theta}) &= \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\beta_{jl}^2}{2a_{jl}^2}} \right\} \times \left\{ \prod_{l=1}^k \alpha_l^{b_l-1} e^{-c_l \alpha_l} \right\} \times \left\{ \prod_{r=1}^R e^{-\frac{1}{2g_r^2}(\mu_r - f_r)^2} \right\} \times \quad (3.14) \\ &\times \left\{ \prod_{r=1}^R (\sigma_r^2)^{-(h_r+1)} \exp\left(-\frac{q_r}{\sigma_r^2}\right) \right\} \times \left\{ \prod_{r=1}^R p_r^{s_r-1} \right\} \times \\ &\times \left\{ \prod_{i=1}^n \prod_{l=1}^k \left[\exp\left(-\alpha_l e^{\mathbf{z}'_i \boldsymbol{\beta}_l + w_i}\right) \left(\alpha_l e^{\mathbf{z}'_i \boldsymbol{\beta}_l + w_i}\right)^{x_{li}} \right] \right\} \end{aligned}$$

e $\boldsymbol{\mu} = (\mu_1, \dots, \mu_R)'$ e $\boldsymbol{\sigma}^2 = (\sigma_1^2, \dots, \sigma_R^2)'$.

Nota-se que em (3.13), independentemente da distribuição *a priori* adotada, o termo

$$\prod_{i=1}^n \sum_{r=1}^R p_r \phi_r(w_i | \mu_r, \sigma_r^2) \quad (3.15)$$

não permite expressar as densidades condicionais como um produto de componentes independentes.

Para a simplificação da *posteriori* conjunta e das distribuições condicionais completas para o amostrador de Gibbs, introduz-se uma variável latente (TANNER & WONG, 1987) $\mathbf{V}_i = (V_{i1}, \dots, V_{iR})'$ para cada observação. Assume-se também que as distribuições condicionais de \mathbf{V}_i dado $\boldsymbol{\theta}$ e \mathbf{x} têm distribuição Multinomial, $Mult[1, (u_{i1}, \dots, u_{iR})]$, com probabilidades dadas por

$$u_{ir} = \frac{s_r(p_r, w_i, \mu_r, \sigma_r^2)}{\sum_{r=1}^R s_r(p_r, w_i, \mu_r, \sigma_r^2)} \quad (3.16)$$

em que $s_r(p_r, w_i, \mu_r, \sigma_r^2) = p_r \phi_r(w_i | \mu_r, \sigma_r^2)$, $\sum_{r=1}^R u_{ir} = 1$ e $\sum_{i=1}^n v_{ir} = 1$, para $i = 1, \dots, n$, ou seja,

$$\pi(\mathbf{v}_1, \dots, \mathbf{v}_n | \boldsymbol{\theta}, \mathbf{x}) \propto \frac{\prod_{i=1}^n \prod_{r=1}^R [p_r \phi_r(w_i | \mu_r, \sigma_r^2)]^{v_{ir}}}{\prod_{i=1}^n \sum_{r=1}^R s_r(p_r, w_i, \mu_r, \sigma_r^2)} \quad (3.17)$$

Combinando-se as expressões (3.13) e (3.17), obtém-se a distribuição *a posteriori* conjunta para $\boldsymbol{\theta}$ e \mathbf{V} dado \mathbf{x} como um produto de componentes independentes, isto é,

$$\pi(\mathbf{v}_1, \dots, \mathbf{v}_n, \boldsymbol{\theta} | \mathbf{x}) \propto \psi(\boldsymbol{\theta}) \prod_{i=1}^n \prod_{r=1}^R [p_r \phi_r(w_i | \mu_r, \sigma_r^2)]^{v_{ir}} \quad (3.18)$$

sendo $\psi(\boldsymbol{\theta})$ dado em (3.14).

Assumindo-se um caso especial de (3.11), com $R = 2$, tem-se a seguinte distribuição *a priori* para p_1 ($p_1 + p_2 = 1$):

$$p_1 \sim \text{Beta}(s_1, s_2) \quad (3.19)$$

em que $\text{Beta}(s_1, s_2)$ denota uma distribuição Beta com média $s_1/(s_1 + s_2)$ e variância $s_1 s_2 / [(s_1 + s_2 + 1)(s_1 + s_2)^2]$.

A distribuição condicional de V_i dado $\boldsymbol{\theta}$ e \mathbf{x} é Bernoulli, $\text{Bin}(1, u_i)$, com probabilidades de sucesso

$$u_i = \frac{p_1 \phi_1(w_i | \mu_1, \sigma_1^2)}{p_1 \phi_1(w_i | \mu_1, \sigma_1^2) + p_2 \phi_2(w_i | \mu_2, \sigma_2^2)} \quad (3.20)$$

ou seja,

$$\pi(v_i | \boldsymbol{\theta}, \mathbf{x}) \propto u_i^{v_i} (1 - u_i)^{1 - v_i} \quad (3.21)$$

Para a geração de amostras da distribuição *a posteriori* conjunta em (3.17), considerando $R = 2$, utiliza-se o Amostrador de Gibbs. Considerando-se $\boldsymbol{\theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_p^{(0)})$ o vetor com os valores iniciais dos parâmetros, procede-se da seguinte maneira:

- (a) Gerar uma amostra de $\mathbf{v}^{(1)} = (v_1^{(1)}, v_2^{(1)}, \dots, v_n^{(1)})'$ da equação (3.20);
- (b) Gerar uma amostra de $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$ das distribuições condicionais $\pi(\theta_1^{(1)} | \theta_2^{(0)}, \dots, \theta_p^{(0)}, \mathbf{v}^{(1)}, \mathbf{x}), \dots, \pi(\theta_p^{(1)} | \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{p-1}^{(1)}, \mathbf{v}^{(1)}, \mathbf{x})$;
- (c) Repetir os passos (a) e (b) um número grande de vezes até obter convergência.

Deste modo, as distribuições condicionais para o algoritmo descrito acima são dadas por:

$$(i) \pi(w_i | \boldsymbol{\theta}_{(w_i)}, \mathbf{x}, \mathbf{v}) \propto N \left[\frac{v_i \mu_1 \sigma_2^2 + (1 - v_i) \mu_2 \sigma_1^2 + \sigma_1^2 \sigma_2^2 \sum_{l=1}^k x_{li}}{v_i \sigma_2^2 + (1 - v_i) \sigma_1^2}; \frac{\sigma_1^2 \sigma_2^2}{v_i \sigma_2^2 + (1 - v_i) \sigma_1^2} \right] \psi_{1i}(\boldsymbol{\theta}) \quad (3.22)$$

em que $\psi_{1i}(\boldsymbol{\theta}) = \exp \left\{ - \sum_{l=1}^k \alpha_l \exp(\mathbf{z}'_l \boldsymbol{\beta}_l + w_i) \right\}$ e $\boldsymbol{\theta}_{(w_i)}$ é o vetor com todos os parâmetros

exceto $w_i, i = 1, \dots, n$;

$$(ii) \sigma_r^2 \mid \boldsymbol{\theta}_{(\sigma_r^2)}, \mathbf{x}, \mathbf{v} \sim I.G. \left[h_r + \frac{1}{2} \sum_{i=1}^n \gamma_i; q_r + \frac{1}{2} \sum_{i=1}^n \gamma_i (w_i - \mu_r)^2 \right] \quad (3.23)$$

$$r = 1, 2; \text{ em que } \gamma_i = \begin{cases} v_i & \text{se } r = 1 \\ 1 - v_i & \text{se } r = 2 \end{cases};$$

$$(iii) \mu_r \mid \boldsymbol{\theta}_{(\mu_r)}, \mathbf{x}, \mathbf{v} \sim N \left[\frac{g_r^2 \sum_{i=1}^n \gamma_i w_i + \sigma_r^2 f_r}{g_r^2 \sum_{i=1}^n \gamma_i + \sigma_r^2}; \frac{\sigma_r^2 g_r^2}{g_r^2 \sum_{i=1}^n \gamma_i + \sigma_r^2} \right] \quad (3.24)$$

$r = 1, 2$;

$$(iv) p_1 \mid \boldsymbol{\theta}_{(\mu_r)}, \mathbf{x}, \mathbf{v} \sim Beta \left(s_1 + \sum_{i=1}^n v_i; s_2 + n - \sum_{i=1}^n v_i \right) \quad (3.25)$$

Observa-se que as distribuições condicionais para β_{jl} e α_l são as mesmas dadas em (3.7) e (3.8), $j = 1, \dots, p$ e $l = 1, \dots, k$.

3.1.3 Modelo 3: Distribuição Gama para os efeitos aleatórios

Modificando-se a média em (3.2), considera-se agora o seguinte modelo:

$$\lambda_{li} = w_i \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l) \quad (3.26)$$

em que w_i é o efeito aleatório ou "fragilidade", $l = 1, \dots, k$ e $i = 1, \dots, n$.

Assume-se que os efeitos aleatórios W_i são independentes e seguem distribuição gama com ambos os parâmetros iguais a ϕ^{-1} , ou seja,

$$W_i \sim \Gamma(\phi^{-1}, \phi^{-1}) \quad (3.27)$$

Isto é, $E(W_i) = \frac{\phi^{-1}}{\phi^{-1}} = 1$ e $Var(W_i) = \frac{\phi^{-1}}{(\phi^{-1})^2} = \phi$.

Das equações (3.3) e (3.26), segue que a função de verossimilhança para $\boldsymbol{\beta}, \boldsymbol{\alpha}$ e

\mathbf{w} é dada por:

$$\begin{aligned} L(\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w}) &= \prod_{l=1}^k \prod_{i=1}^n \frac{e^{-\lambda_{li}} \lambda_{li}^{x_{li}}}{x_{li}!} \\ &\propto \prod_{l=1}^k \prod_{i=1}^n \{ \exp[-w_i \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l)] \} \times [w_i \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l)]^{x_{li}} \end{aligned} \quad (3.28)$$

Assumindo-se independência *a priori* entre os parâmetros, consideram-se as seguintes densidades *a priori* para β_{jl} , α_l e ϕ :

$$\begin{aligned} (i) \quad \beta_{jl} &\sim N(0, a_{jl}^2), \quad a_{jl} \text{ conhecido}, \quad l = 1, \dots, k \text{ e } j = 1, \dots, p; \\ (ii) \quad \alpha_l &\sim \Gamma(b_l, c_l), \quad b_l \text{ e } c_l \text{ conhecidos}, \quad l = 1, \dots, k; \\ (iii) \quad \phi &\sim \Gamma(d, f), \quad d \text{ e } f \text{ conhecidos.} \end{aligned} \quad (3.29)$$

Combinando-se as expressões (3.27), (3.28) e (3.29), obtém-se a distribuição *a posteriori* conjunta para $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\alpha}, \mathbf{w}, \phi)'$, dada por:

$$\begin{aligned} \pi(\boldsymbol{\theta} \mid \mathbf{x}) &\propto \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\beta_{jl}^2}{2a_{jl}^2}} \right\} \times \left\{ \prod_{l=1}^k \alpha_l^{b_l-1} e^{-c_l \alpha_l} \right\} \times \\ &\times \left\{ \prod_{i=1}^n w_i^{\phi-1} e^{-\phi w_i} \right\} \times \{ \phi^{d-1} e^{-f\phi} \} \times \\ &\times \prod_{l=1}^k \prod_{i=1}^n \{ \exp[-w_i \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l)] \} \times [w_i \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l)]^{x_{li}} \end{aligned} \quad (3.30)$$

sendo \mathbf{x} o vetor dos dados observados.

As distribuições condicionais para o Amostrador de Gibbs são dadas por:

$$(i) \quad \pi(\beta_{jl} \mid \boldsymbol{\theta}_{(\beta_{jl})}, \mathbf{x}) \propto \exp\left\{-\frac{\beta_{jl}^2}{2a_{jl}^2}\right\} \exp\left\{\beta_{jl} \sum_{i=1}^n z_{ji} x_{li} - \alpha_l \sum_{i=1}^n w_i \exp(\mathbf{z}'_i \boldsymbol{\beta}_l)\right\} \quad (3.31)$$

em que $\boldsymbol{\theta}_{(\beta_{jl})}$ denota o vetor com todos os parâmetros exceto β_{jl} , $j = 1, \dots, p$ e $l = 1, \dots, k$.

$$(ii) \quad \alpha_l \mid \boldsymbol{\theta}_{(\alpha_l)}, \mathbf{x} \sim \Gamma\left(b_l + \sum_{i=1}^n x_{li}; \quad c_l + \sum_{i=1}^n w_i \exp(\mathbf{z}'_i \boldsymbol{\beta}_l)\right) \quad (3.32)$$

em que $\boldsymbol{\theta}_{(\alpha_l)}$ denota o vetor com todos os parâmetros exceto α_l , $l = 1, \dots, k$.

$$(iii) \quad w_i \mid \boldsymbol{\theta}_{(w_i)}, \mathbf{x} \sim \Gamma \left(\phi^{-1} + \sum_{l=1}^k x_{li}; \phi^{-1} + \sum_{l=1}^k \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l) \right) \quad (3.33)$$

em que $\boldsymbol{\theta}_{(w_i)}$ denota o vetor com todos os parâmetros exceto w_i , $i = 1, \dots, n$.

$$(iv) \quad \pi(\phi \mid \boldsymbol{\theta}_{(\phi)}, \mathbf{x}) \propto \left(\prod_{i=1}^n w_i \right)^{1/\phi} \phi^{d-1} \exp \left(-\frac{1}{\phi} \sum_{i=1}^n w_i - \phi f \right) \quad (3.34)$$

Nota-se que as distribuições condicionais em (3.31) e (3.34) de β_{jl} e ϕ apresentam forma desconhecida e, portanto, devem ser geradas utilizando o algoritmo *Metropolis-Hastings*.

3.2 Variâncias das respostas de contagem longitudinais sob diferentes modelos de efeitos aleatórios

Nesta seção são desenvolvidos os cálculos das esperanças e variâncias não-condicionais dos modelos considerados.

3.2.1 Modelo 1

As esperanças e variâncias condicionais das respostas são dadas, considerando-se a expressão (3.1), por:

$$\begin{aligned} E(X_{li} \mid \boldsymbol{\beta}_l, \alpha_l, \mathbf{z}_i, w_i) &= \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i) \\ Var(X_{li} \mid \boldsymbol{\beta}_l, \alpha_l, \mathbf{z}_i, w_i) &= \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i) \end{aligned} \quad (3.35)$$

Considerando esse modelo, a esperança não-condicional é dada por:

$$\begin{aligned} E(X_{li} \mid \boldsymbol{\beta}_l, \alpha_l, \mathbf{z}_i) &= E[E(X_{li} \mid \boldsymbol{\beta}_l, \alpha_l, \mathbf{z}_i, w_i)] \\ &= \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l) E[e^{W_i}] \end{aligned} \quad (3.36)$$

Resultado 1: Se W_i são variáveis aleatórias independentes e identicamente distribuí-

das $N(0, \sigma_w^2)$, então e^{W_i} tem distribuição log-normal com média $E(e^{W_i}) = \exp\left(\frac{\sigma_w^2}{2}\right)$ e variância $Var(e^{W_i}) = e^{\sigma_w^2} (e^{\sigma_w^2} - 1)$.

Através do Resultado 1, tem-se:

$$E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) = \alpha_l \exp(\mathbf{z}'_i \beta_l) \exp\left(\frac{\sigma_w^2}{2}\right) \quad (3.37)$$

para $l = 1, \dots, k$ e $i = 1, \dots, n$.

A variância não-condicional é calculada por:

$$\begin{aligned} Var(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) &= Var[E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] + E[Var(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] \\ &= \alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) Var(e^{W_i}) + \alpha_l \exp(\mathbf{z}'_i \beta_l) E[e^{W_i}] \\ &= \alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) e^{\sigma_w^2} (e^{\sigma_w^2} - 1) + \alpha_l \exp(\mathbf{z}'_i \beta_l) \exp\left(\frac{\sigma_w^2}{2}\right) \end{aligned} \quad (3.38)$$

De (3.37) e (3.38), verifica-se que a variabilidade extra-Poisson (superdispersão) é dada por:

$$\alpha_l^2 e^{\sigma_w^2} (e^{\sigma_w^2} - 1) \exp(2\mathbf{z}'_i \beta_l) \quad (3.39)$$

para $l = 1, \dots, k$ e $i = 1, \dots, n$.

3.2.2 Modelo 2

As esperanças e variâncias condicionais das respostas são as mesmas dadas em (3.35), com W_i sendo uma mistura finita de distribuições Normais com densidade $\pi(w_i) = \sum_{r=1}^R p_r \phi_r(w_i | \mu_r, \sigma_r^2)$.

Considerando esse modelo, a esperança não-condicional é dada por:

$$\begin{aligned} E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) &= E[E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] \\ &= \alpha_l \exp(\mathbf{z}'_i \beta_l) E[e^{W_i}] \end{aligned} \quad (3.40)$$

Resultado 2:

$$\begin{aligned}
\mu_M &= E[W_i] = \int_{-\infty}^{\infty} w_i \pi(w_i) dw_i \\
&= \int_{-\infty}^{\infty} w_i \sum_{r=1}^R p_r \phi_r(w_i | \mu_r, \sigma_r^2) dw_i \\
&= \sum_{r=1}^R p_r \int_{-\infty}^{\infty} w_i \phi_r(w_i | \mu_r, \sigma_r^2) dw_i \\
&= \sum_{r=1}^R p_r \mu_r
\end{aligned} \tag{3.41}$$

$$\begin{aligned}
E[W_i^2] &= \int_{-\infty}^{\infty} w_i^2 \pi(w_i) dw_i = \int_{-\infty}^{\infty} w_i^2 \sum_{r=1}^R p_r \phi_r(w_i | \mu_r, \sigma_r^2) dw_i \\
&= \sum_{r=1}^R p_r \int_{-\infty}^{\infty} w_i^2 \phi_r(w_i | \mu_r, \sigma_r^2) dw_i \\
&= \sum_{r=1}^R p_r (\sigma_r^2 + \mu_r^2)
\end{aligned} \tag{3.42}$$

$$\begin{aligned}
\sigma_M^2 &= Var(W_i) = E(W_i^2) - [E(W_i)]^2 \\
&= \sum_{r=1}^R p_r (\sigma_r^2 + \mu_r^2) - \mu_M^2
\end{aligned} \tag{3.43}$$

Assim, $W_i \sim (\mu_M, \sigma_M^2)$.

Resultado 3 (Método "delta"): Se W é uma variável aleatória seguindo distribuição com média μ e variância σ^2 , ou seja, $W \sim (\mu, \sigma^2)$, então $g(W) \stackrel{a}{\sim} (g(\mu); [g'(\mu)]^2 \sigma^2)$.

Como $W_i \sim (\mu_M, \sigma_M^2)$, utilizando-se o Resultado 3, tem-se:

$$g(W_i) = \exp(W_i) \stackrel{a}{\sim} \left(g(\mu_M); [g'(\mu_M)]^2 \sigma_M^2 \right) \tag{3.44}$$

Assim,

$$\begin{aligned} g(\mu_M) &= e^{\mu_M} = \exp\left(\sum_{r=1}^R p_r \mu_r\right) \\ g'(\mu_M) &= e^{\mu_M} = \exp\left(\sum_{r=1}^R p_r \mu_r\right) \end{aligned} \quad (3.45)$$

Com σ_M^2 dado em (3.43) tem-se:

$$g(W_i) = \exp(W_i) \stackrel{a}{\sim} (\exp(\mu_M); \exp(2\mu_M) \sigma_M^2) \quad (3.46)$$

isto é,

$$\begin{aligned} E[e^{W_i}] &\approx \exp(\mu_M) \\ Var[e^{W_i}] &\approx \sigma_M^2 \exp(2\mu_M) \end{aligned} \quad (3.47)$$

Portanto, a esperança não-condicional é dada por:

$$E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) \approx \alpha_l \exp(\mathbf{z}'_i \beta_l) \exp\left(\sum_{r=1}^R p_r \mu_r\right) \quad (3.48)$$

Utilizando-se os resultados acima, a variância não-condicional é dada por:

$$\begin{aligned} Var(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) &= Var[E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] + E[Var(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] \\ &= \alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) Var(e^{W_i}) + \alpha_l \exp(\mathbf{z}'_i \beta_l) E[e^{W_i}] \\ &\approx \alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) \sigma_M^2 \exp(2\mu_M) + \alpha_l \exp(\mathbf{z}'_i \beta_l) \exp(\mu_M) \end{aligned} \quad (3.49)$$

De (3.48) e (3.49), verifica-se que a variabilidade extra-Poisson é dada por:

$$\alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) \left[\sum_{r=1}^R p_r (\sigma_r^2 + \mu_r^2) - \mu_M^2 \right] \exp\left(2 \sum_{r=1}^R p_r \mu_r\right) \quad (3.50)$$

3.2.3 Modelo 3

As esperanças e variâncias condicionais das respostas são dadas por:

$$\begin{aligned} E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i) &= w_i \alpha_l \exp(\mathbf{z}'_i \beta_l) \\ \text{Var}(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i) &= w_i \alpha_l \exp(\mathbf{z}'_i \beta_l) \end{aligned} \quad (3.51)$$

Considerando esse modelo, a esperança não-condicional é dada por:

$$\begin{aligned} E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) &= E[E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] \\ &= \alpha_l \exp(\mathbf{z}'_i \beta_l) E[W_i] \end{aligned} \quad (3.52)$$

Como $W_i \sim \Gamma(\phi^{-1}, \phi^{-1})$, então $E(W_i) = \frac{\phi^{-1}}{\phi^{-1}} = 1$ e $\text{Var}(W_i) = \frac{\phi^{-1}}{(\phi^{-1})^2} = \phi$.

Logo,

$$E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) = \alpha_l \exp(\mathbf{z}'_i \beta_l) \quad (3.53)$$

A variância não-condicional é calculada da seguinte maneira:

$$\begin{aligned} \text{Var}(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i) &= \text{Var}[E(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] + E[\text{Var}(X_{li} | \beta_l, \alpha_l, \mathbf{z}_i, w_i)] \\ &= \alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) \text{Var}(W_i) + \alpha_l \exp(\mathbf{z}'_i \beta_l) E[W_i] \\ &= \alpha_l^2 \exp(2\mathbf{z}'_i \beta_l) \phi + \alpha_l \exp(\mathbf{z}'_i \beta_l) \end{aligned} \quad (3.54)$$

De (3.53) e (3.54), verifica-se que a variabilidade extra-Poisson é dada por:

$$\alpha_l^2 \phi \exp(2\mathbf{z}'_i \beta_l) \quad (3.55)$$

Capítulo 4

Metodologia para dados de contagem com medidas repetidas e excesso de zeros

Assumindo-se X_{li} , com $i = 1, \dots, n$ (unidades amostrais) e $l = 1, \dots, k$ (medidas repetidas para a mesma unidade), uma variável aleatória de valor inteiro não-negativo e supondo que $X_{li} = 0$ ocorre numa frequência acima do que é permitido modelar através de uma distribuição de Poisson usual, pode-se modelar estes tipos de dados através da seguinte função de probabilidade de X_{li} (ANGERS & BISWAS, 2003):

$$p(x_{li}) = P(X_{li} = x_{li}) = \begin{cases} \theta_{li} + (1 - \theta_{li}) e^{-\lambda_{li}}, & \text{se } x_{li} = 0 \\ (1 - \theta_{li}) \frac{e^{-\lambda_{li}} \lambda_{li}^{x_{li}}}{x_{li}!}, & \text{se } x_{li} = 1, 2, 3, \dots \end{cases} \quad (4.1)$$

em que $(1 - e^{-\lambda_{li}})^{-1} < \theta_{li} < 1$, $\lambda_{li} > 0$, para $l = 1, \dots, k$ e $i = 1, \dots, n$.

Pode-se interpretar este modelo da seguinte maneira: para um dado valor de λ_{li} , o parâmetro θ_{li} controla a probabilidade de excesso de zeros.

Na presença de p covariáveis, assume a função de ligação logito para θ_{li} :

$$\theta_{li} = \frac{\exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \quad (4.2)$$

em que $\mathbf{z}'_i = (z_{1i}, z_{2i}, \dots, z_{pi})$ é o vetor de covariáveis da i -ésima unidade amostral, $i = 1, \dots, n$, $\boldsymbol{\gamma}'_l = (\gamma_{1l}, \gamma_{2l}, \dots, \gamma_{pl})$ é o vetor de parâmetros associadas a essas covariáveis no l -ésimo tempo ($l = 1, \dots, k$) e w_i o efeito aleatório na i -ésima observação que captura a

correlação entre as medidas repetidas para cada unidade experimental.

Assim, tem-se a seguinte função de verossimilhança para os parâmetros:

$$L(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{w}, \boldsymbol{\gamma}) \propto \prod_{l=1}^k \prod_{i=1}^n \left\{ [\theta_{li} + (1 - \theta_{li}) e^{-\lambda_{li}}] \times I_{\{0\}}(x_{li}) \right\} \times \quad (4.3)$$

$$\times \left\{ (1 - \theta_{li}) [e^{-\lambda_{li}} \lambda_{li}^{x_{li}}] \times I_{\{1,2,\dots\}}(x_{li}) \right\}$$

em que $I_{\{m\}}(x_{li})$ denota uma função indicadora que assume 1, se x_{li} assumir o valor m e zero, em caso contrário. Nota-se que a verossimilhança é um produto de somas. Para a simplificação da mesma, será introduzida uma variável latente V_{li} para cada observação, sendo que a distribuição condicional desta variável dados todos os parâmetros $(\boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{w}, \boldsymbol{\gamma})$ e os valores observados \mathbf{x} , é Bernoulli com probabilidade de sucesso p_{li} , ou seja:

$$V_{li} \mid \boldsymbol{\beta}, \boldsymbol{\lambda}, \mathbf{w}, \boldsymbol{\gamma}, \mathbf{x} \sim Bin(1, p_{li}) \quad (4.4)$$

em que

$$p_{li} = \frac{\theta_{li}}{\theta_{li} + (1 - \theta_{li}) e^{-\lambda_{li}}} \quad (4.5)$$

Deste modo, a verossimilhança ampliada é dada por:

$$L_A \propto \prod_{l=1}^k \prod_{i=1}^n \left\{ [\theta_{li}^{v_{li}} (1 - \theta_{li})^{1-v_{li}} \exp[-(1 - v_{li}) \lambda_{li}] \times I_{\{0\}}(x_{li})] \times \quad (4.6)$$

$$\times [(1 - \theta_{li}) \lambda_{li}^{x_{li}} \exp(-\lambda_{li}) \times I_{\{1,2,\dots\}}(x_{li})] \right\}$$

4.1 Modelo 1: Distribuição Normal para os efeitos aleatórios

Para este modelo, é assumida a seguinte expressão para λ_{li} , dada por:

$$\lambda_{li} = \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i) \quad (4.7)$$

em que \mathbf{z}'_i é o mesmo definido anteriormente, $\boldsymbol{\beta}'_l = (\beta_{1l}, \beta_{2l}, \dots, \beta_{pl})$ é o vetor de parâmetros do modelo de regressão no l -ésimo tempo ($l = 1, \dots, k$) e w_i o efeito aleatório na i -ésima observação.

Além disso, assumindo-se que os efeitos aleatórios W_i seguem distribuição Normal com média zero e variância σ_w^2

$$W_i \sim N(0, \sigma_w^2) \quad (4.8)$$

e considerando-se independência *a priori* entre os parâmetros, as seguintes densidades *a priori* para β_{jl} , γ_{jl} , α_l e σ_w^2 :

$$\begin{aligned} (i) \quad & \beta_{jl} \sim N(0, a_{jl}^2), \quad a_{jl} \text{ conhecido}, \quad l = 1, \dots, k \text{ e } j = 1, \dots, p; \\ (ii) \quad & \gamma_{jl} \sim N(0, g_{jl}^2), \quad g_{jl} \text{ conhecido}, \quad l = 1, \dots, k \text{ e } j = 1, \dots, p; \\ (iii) \quad & \alpha_l \sim \Gamma(b_l, c_l), \quad b_l \text{ e } c_l \text{ conhecidos}, \quad l = 1, \dots, k; \\ (iv) \quad & \sigma_w^2 \sim I.G.(d, f), \quad d \text{ e } f \text{ conhecidos.} \end{aligned} \quad (4.9)$$

Combinando-se as expressões (4.6), (4.8) e (4.9), obtém-se a distribuição *a posteriori* conjunta para $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\alpha}', \mathbf{w}', \sigma_w^2)'$, dada por:

$$\begin{aligned} \pi(\boldsymbol{\theta} \mid \mathbf{v}, \mathbf{x}) & \propto \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\beta_{jl}^2}{2a_{jl}^2}} \right\} \times \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\gamma_{jl}^2}{2g_{jl}^2}} \right\} \times \left\{ \prod_{i=1}^{40} \frac{1}{\sqrt{\sigma_w^2}} e^{-\frac{w_i^2}{2\sigma_w^2}} \right\} \times \\ & \times \left\{ \prod_{l=1}^k \alpha_l^{b_l-1} e^{-c_l \alpha_l} \right\} \times (\sigma_w^2)^{-(d+1)} e^{-\frac{f}{\sigma_w^2}} \times \\ & \times \left[\prod_{l=1}^k \prod_{i=1}^n \frac{\exp\left\{(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i) v_{li} - \alpha_l e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)} (1 - v_{li})\right\}}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \times I_{\{0\}}(x_{li}) \right] \times \\ & \times \left[\prod_{l=1}^k \prod_{i=1}^n \frac{\alpha_l^{x_{li}} \exp(\mathbf{z}'_i \boldsymbol{\beta}_l x_{li} + w_i x_{li}) \exp(-\alpha_l e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)})}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \times I_{\{1,2,\dots\}}(x_{li}) \right] \end{aligned} \quad (4.10)$$

As distribuições condicionais para o amostrador de Gibbs são dadas por:

$$\begin{aligned} (i) \quad \pi(\beta_{jl} \mid \boldsymbol{\theta}_{(\beta_{jl})}, \mathbf{x}) & \propto \exp\left\{-\frac{\beta_{jl}^2}{2a_{jl}^2}\right\} \\ & \times \left\{ \exp\left[-\alpha_l \sum_{i=1}^n (1 - v_{li}) e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)}\right] I_{\{0\}}(x_{li}) \right\} \\ & \times \left\{ \exp\left[\beta_{jl} \sum_{i=1}^n z_{ji} x_{li} - \alpha_l \sum_{i=1}^n e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)}\right] I_{\{1,2,\dots\}}(x_{li}) \right\}, \end{aligned} \quad (4.11)$$

em que $\boldsymbol{\theta}_{(\beta_{jl})}$ denota o vetor com todos os parâmetros exceto β_{jl} ; $j = 1, \dots, p$; $l = 1, \dots, k$;

$$(ii) \alpha_l \mid \boldsymbol{\theta}_{(\alpha_l)}, \mathbf{x} \sim \Gamma [b^*; c^*], \quad (4.12)$$

em que $\boldsymbol{\theta}_{(\alpha_l)}$ denota o vetor com todos os parâmetros exceto α_l , $l = 1, \dots, k$; $b^* = b_l + \sum_{i=1}^n x_{li}$ e $c^* = c_l + \sum_{i=1}^n (1 - v_{li}) e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)} I_{\{0\}}(x_{li}) + \sum_{i=1}^n e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)} I_{\{1,2,\dots\}}(x_{li})$;

$$\begin{aligned} (iii) \pi(w_i \mid \boldsymbol{\theta}_{(w_i)}, \mathbf{x}) &\propto \exp \left\{ -\frac{w_i^2}{2\sigma_w^2} \right\} \times \prod_{l=1}^k \frac{1}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \\ &\times \exp \left\{ \sum_{l=1}^k \left[w_i v_{li} - (1 - v_{li}) \alpha_l e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)} \right] I_{\{0\}}(x_{li}) \right\} \\ &\times \exp \left\{ \sum_{l=1}^k \left[w_i x_{li} - \alpha_l e^{(\mathbf{z}'_i \boldsymbol{\beta}_l + w_i)} \right] I_{\{1,2,\dots\}}(x_{li}) \right\} \end{aligned} \quad (4.13)$$

em que $\boldsymbol{\theta}_{(w_i)}$ denota o vetor com todos os parâmetros exceto w_i , $i = 1, \dots, n$;

$$(iv) \sigma_w^2 \mid \boldsymbol{\theta}_{(\sigma_w^2)}, \mathbf{x} \sim I.G. \left[d + \frac{n}{2}; f + \frac{1}{2} \sum_{i=1}^n w_i^2 \right] \quad (4.14)$$

em que $\boldsymbol{\theta}_{(\sigma_w^2)}$ denota o vetor com todos os parâmetros exceto σ_w^2 .

$$\begin{aligned} (v) \pi(\gamma_{jl} \mid \boldsymbol{\theta}_{(\beta_{jl})}, \mathbf{x}) &\propto \exp \left\{ -\frac{\gamma_{jl}^2}{2g_{jl}^2} \right\} \times \prod_{i=1}^n \frac{1}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \\ &\times \left\{ \exp \left(\gamma_{jl} \sum_{i=1}^n z_{ji} v_{li} \right) I_{\{0\}}(x_{li}) \right\} \end{aligned} \quad (4.15)$$

em que $\boldsymbol{\theta}_{(\gamma_{jl})}$ denota o vetor com todos os parâmetros exceto γ_{jl} ; $j = 1, \dots, p$; $l = 1, \dots, k$;

Nota-se que as distribuições condicionais em (4.11), (4.13), (4.14), (4.15) apresentam forma desconhecida e, portanto, devem ser geradas utilizando o algoritmo *Metropolis-Hastings*.

4.2 Modelo 2: Distribuição Gama para os efeitos aleatórios

Para este modelo, é assumida a seguinte expressão para λ_{li} , dada por:

$$\lambda_{li} = w_i \alpha_l \exp(\mathbf{z}'_i \boldsymbol{\beta}_l) \quad (4.16)$$

em que \mathbf{z}'_i , $\boldsymbol{\beta}'_l$ e w_i são os mesmos definidos anteriormente, para $l = 1, \dots, k$ e $i = 1, \dots, n$.

Assume-se que os efeitos aleatórios W_i são independentes e seguem distribuição gama com ambos os parâmetros iguais a ϕ^{-1} , ou seja,

$$W_i \sim \Gamma(\phi^{-1}, \phi^{-1}) \quad (4.17)$$

Assumindo-se independência *a priori* entre os parâmetros, consideram-se as seguintes densidades *a priori* para β_{jl} , α_l e ϕ :

$$\begin{aligned} (i) \quad & \beta_{jl} \sim N(0, a_{jl}^2), \quad a_{jl} \text{ conhecido}, \quad l = 1, \dots, k \text{ e } j = 1, \dots, p; \\ (ii) \quad & \gamma_{jl} \sim N(0, g_{jl}^2), \quad g_{jl} \text{ conhecido}, \quad l = 1, \dots, k \text{ e } j = 1, \dots, p; \\ (iii) \quad & \alpha_l \sim \Gamma(b_l, c_l), \quad b_l \text{ e } c_l \text{ conhecidos}, \quad l = 1, \dots, k; \\ (iv) \quad & \phi \sim \Gamma(d, f), \quad d \text{ e } f \text{ conhecidos.} \end{aligned} \quad (4.18)$$

Combinando-se as expressões (4.6), (4.17) e (4.18), obtém-se a distribuição *a posteriori* conjunta para $\boldsymbol{\theta} = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \boldsymbol{\alpha}', \mathbf{w}', \phi)'$, dada por:

$$\begin{aligned} \pi(\boldsymbol{\theta} | \mathbf{x}) & \propto \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\beta_{jl}^2}{2a_{jl}^2}} \right\} \times \left\{ \prod_{j=1}^p \prod_{l=1}^k e^{-\frac{\gamma_{jl}^2}{2g_{jl}^2}} \right\} \times \left\{ \prod_{l=1}^k \alpha_l^{b_l-1} e^{-c_l \alpha_l} \right\} \times \\ & \times \left\{ \prod_{i=1}^n w_i^{\phi^{-1}-1} e^{-\phi^{-1} w_i} \right\} \times \{ \phi^{d-1} e^{-f\phi} \} \times \\ & \times \prod_{l=1}^k \prod_{i=1}^n \left\{ \left[\frac{\exp\left\{ (\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i) v_{li} - w_i \alpha_l e^{(\mathbf{z}'_i \boldsymbol{\beta}_l)} (1 - v_{li}) \right\}}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \times I_{\{0\}}(x_{li}) \right] \times \right. \\ & \left. \times \left[\frac{w_i^{x_{li}} \alpha_l^{x_{li}} \exp(\mathbf{z}'_i \boldsymbol{\beta}_l x_{li}) \exp(-w_i \alpha_l e^{(\mathbf{z}'_i \boldsymbol{\beta}_l)})}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \times I_{\{1,2,\dots\}}(x_{li}) \right] \right\} \end{aligned} \quad (4.19)$$

As distribuições condicionais para o amostrador de Gibbs são dadas por:

$$(i) \pi \left(\beta_{jl} \mid \boldsymbol{\theta}_{(\beta_{jl})}, \mathbf{v}, \mathbf{x} \right) \propto \exp \left\{ -\frac{\beta_{jl}^2}{2a_{jl}^2} \right\} \quad (4.20)$$

$$\times \left\{ \exp \left[-\alpha_l \sum_{i=1}^n w_i (1 - v_{li}) e^{(\mathbf{z}'_i \beta_l)} \right] I_{\{0\}}(x_{li}) \right\}$$

$$\times \left\{ \exp \left[\beta_{jl} \sum_{i=1}^n z_{ji} x_{li} - \alpha_l \sum_{i=1}^n w_i e^{(\mathbf{z}'_i \beta_l)} \right] I_{\{1,2,\dots\}}(x_{li}) \right\},$$

em que $\boldsymbol{\theta}_{(\beta_{jl})}$ denota o vetor com todos os parâmetros exceto β_{jl} ; $j = 1, \dots, p$; $l = 1, \dots, k$;

$$(ii) \alpha_l \mid \boldsymbol{\theta}_{(\alpha_l)}, \mathbf{v}, \mathbf{x} \sim \Gamma [b^*; c^*], \quad (4.21)$$

em que $\boldsymbol{\theta}_{(\alpha_l)}$ denota o vetor com todos os parâmetros exceto α_l , $l = 1, \dots, k$; $b^* = b_l + \sum_{i=1}^n x_{li}$ e $c^* = c_l + \sum_{i=1}^n w_i (1 - v_{li}) e^{(\mathbf{z}'_i \beta_l)} I_{\{0\}}(x_{li}) + \sum_{i=1}^n w_i e^{(\mathbf{z}'_i \beta_l)} I_{\{1,2,\dots\}}(x_{li})$;

$$(iii) \pi \left(w_i \mid \boldsymbol{\theta}_{(w_i)}, \mathbf{v}, \mathbf{x} \right) \propto w_i^{\phi-1} \exp(\phi^{-1} w_i) \times \prod_{l=1}^k \frac{1}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)}$$

$$\times \left[\exp \left\{ w_i \sum_{l=1}^k [v_{li} - (1 - v_{li}) \alpha_l e^{(\mathbf{z}'_i \beta_l)}] \right\} I_{\{0\}}(x_{li}) \right] \quad (4.22)$$

$$\times \left[\exp \left\{ w_i \sum_{l=1}^k [x_{li} - \alpha_l e^{(\mathbf{z}'_i \beta_l + w_i)}] \right\} I_{\{1,2,\dots\}}(x_{li}) \right]$$

em que $\boldsymbol{\theta}_{(w_i)}$ denota o vetor com todos os parâmetros exceto w_i , $i = 1, \dots, n$;

$$(iv) \pi \left(\phi \mid \boldsymbol{\theta}_{(\phi)}, \mathbf{v}, \mathbf{x} \right) \propto \left(\prod_{i=1}^n w_i \right)^{1/\phi} \phi^{d-1} \exp \left(-\frac{1}{\phi} \sum_{i=1}^n w_i - \phi f \right), \quad (4.23)$$

em que $\boldsymbol{\theta}_{(\phi)}$ denota o vetor com todos os parâmetros exceto ϕ .

$$(v) \pi \left(\gamma_{jl} \mid \boldsymbol{\theta}_{(\gamma_{jl})}, \mathbf{x} \right) \propto \exp \left\{ -\frac{\gamma_{jl}^2}{2g_{jl}^2} \right\} \times \prod_{i=1}^n \frac{1}{1 + \exp(\mathbf{z}'_i \boldsymbol{\gamma}_l + w_i)} \quad (4.24)$$

$$\times \left\{ \exp \left(\gamma_{jl} \sum_{i=1}^n z_{ji} v_{li} \right) I_{\{0\}}(x_{li}) \right\},$$

em que $\boldsymbol{\theta}_{(\gamma_{jl})}$ denota o vetor com todos os parâmetros exceto γ_{jl} ; $j = 1, \dots, p$; $l = 1, \dots, k$;

Nota-se que as distribuições condicionais em (4.20), (4.22), (4.23), (4.24) apresentam forma desconhecida e, portanto, devem ser geradas utilizando o algoritmo *Metropolis-Hastings*.

Capítulo 5

Aplicação 1 - dados de contagem longitudinais

Neste Capítulo é apresentado um conjunto de dados reais e realizada a modelagem através das três formas apresentadas no Capítulo 3.

5.1 Apresentação e Descrição dos Dados

O conjunto de dados (BRADSTREET, 1992) refere-se a um estudo sobre a toxicidade oral em cachorros. Trinta e dois cachorros (dezesesseis machos e dezesseis fêmeas) foram designados aleatoriamente a um grupo controle ou a uma das três doses graduadas (8, 25 e 75 mg/kg) de um composto sob investigação. A dosagem oral foi realizada diariamente, sete dias por semana, por um período de 12 semanas consecutivas. Amostras de sangue foram coletadas antes do tratamento e novamente nas semanas 4, 8 e 12 a partir da veia jugular de cada animal logo após ao jejum noturno. Os níveis de fosfatase alcalina (unidades por litro) foram examinados a partir dessas amostras.

O objetivo desta pesquisa era verificar a possível existência de diferenças na quantidade da enzima nos diferentes tempos, doses e conforme o gênero dos animais (machos e fêmeas). A Tabela 2 apresenta os dados deste estudo em que a variável resposta é a quantidade de unidades de fosfatase alcalina por litro e as covariáveis são as doses e o gênero. A Tabela 3 apresenta algumas estatísticas descritivas da variável resposta segundo gênero e tempo.

Tabela 2: Dados sobre o estudo sobre toxicidade em cachorros.

Dose (<i>mg/kg</i>)	Machos					Fêmeas				
	Tempos (Semanas)					Tempos (Semanas)				
	Animal	0	4	8	12	Animal	0	4	8	12
0	1	169	136	125	122	17	125	105	92	86
	2	291	282	218	154	18	138	94	81	72
	3	158	128	104	91	19	113	91	81	72
	4	122	113	91	82	20	137	94	79	75
8	5	203	171	141	109	21	170	150	125	121
	6	178	154	127	114	22	139	127	103	94
	7	141	104	90	80	23	131	152	112	99
	8	181	143	120	125	24	125	105	91	82
25	9	101	80	78	83	25	113	101	89	122
	10	199	149	147	142	26	150	113	105	90
	11	141	138	118	108	27	155	161	134	116
	12	149	131	133	134	28	133	197	218	159
75	13	135	218	204	226	29	113	148	122	143
	14	153	164	148	146	30	102	191	172	162
	15	147	192	205	167	31	128	226	261	279
	16	157	168	168	165	32	91	173	156	202

Tabela 3: Estatísticas descritivas para o conjunto de dados considerado.

Gênero	Tempo	Média	Variância	CV ^(a)	Mínimo	Máximo	Mediana
Machos	Basal	164,06	1849,13	0,26	101	291	155
	T ₂	154,44	2272,93	0,31	80	282	146
	T ₃	138,56	1775,86	0,30	78	218	130
	T ₄	128,00	1502,80	0,30	80	226	123,5
Fêmeas	Basal	128,94	402,46	0,16	91	170	131
	T ₂	139,25	1771,80	0,30	91	226	127
	T ₃	126,31	2730,63	0,41	79	261	105
	T ₄	123,38	3127,18	0,45	72	279	99
Geral	Basal	146,50	1407,87	0,26	291	91	140
	T ₂	146,84	2016,65	0,31	282	80	145,5
	T ₃	132,44	2219,29	0,36	261	78	123,5
	T ₄	125,69	2245,83	0,38	279	72	118,5

(a) CV: Coeficiente de variação; CV = desvio-padrão/média.

Através da Tabela 3, nota-se que para os cachorros machos, parece haver uma diminuição das médias das unidades de enzima à medida que o tempo passa. Entretanto, esse fato parece não ocorrer entre as fêmeas. Entre machos e fêmeas, parece existir uma diferença entre os valores médios da variável resposta, que são menores para as fêmeas. Analisando os coeficientes de variação, nota-se que parece não haver muita variabilidade entre os tempos (no caso geral). Nota-se, também, que a variância em todos os casos é maior que a média, caracterizando possível superdispersão dos dados.

A Figura 1 apresenta o boxplot dos dados segundo o gênero do animal e a Figura 2 apresenta os perfis das quantidades médias de unidades de fosfatase alcalina.

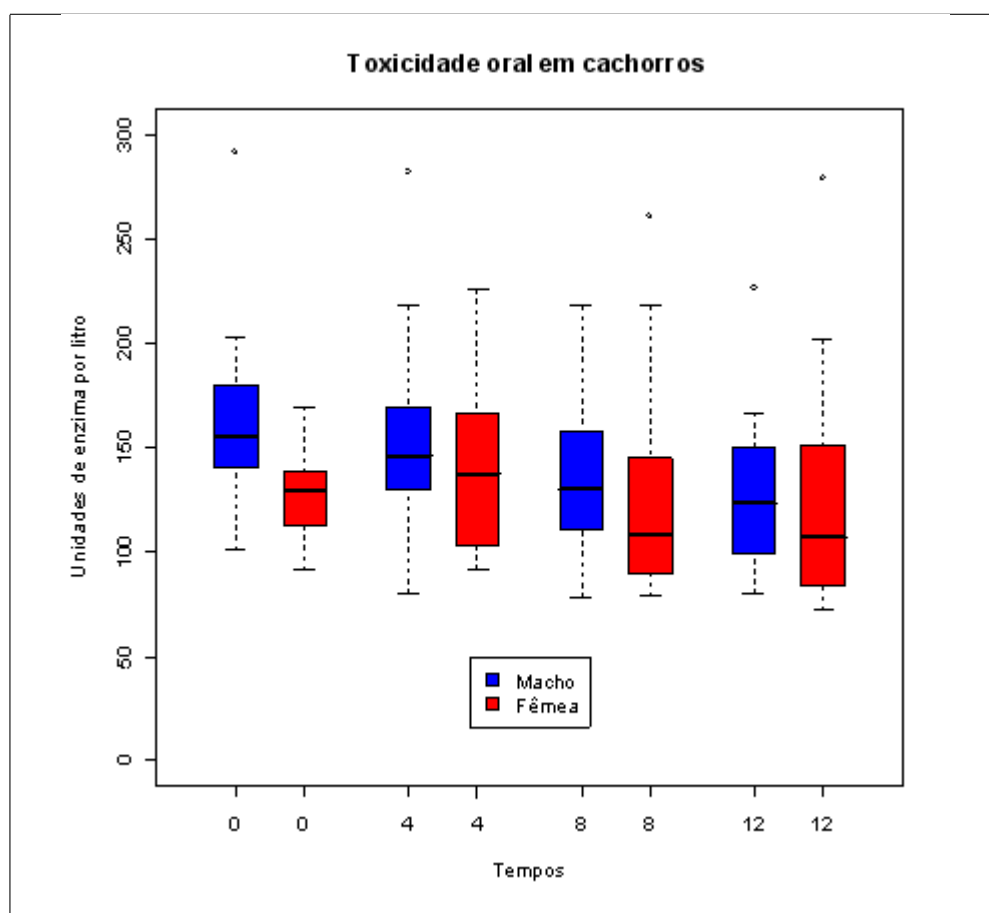


Figura 1: Boxplot para os dados sobre a toxicidade oral em cachorros.

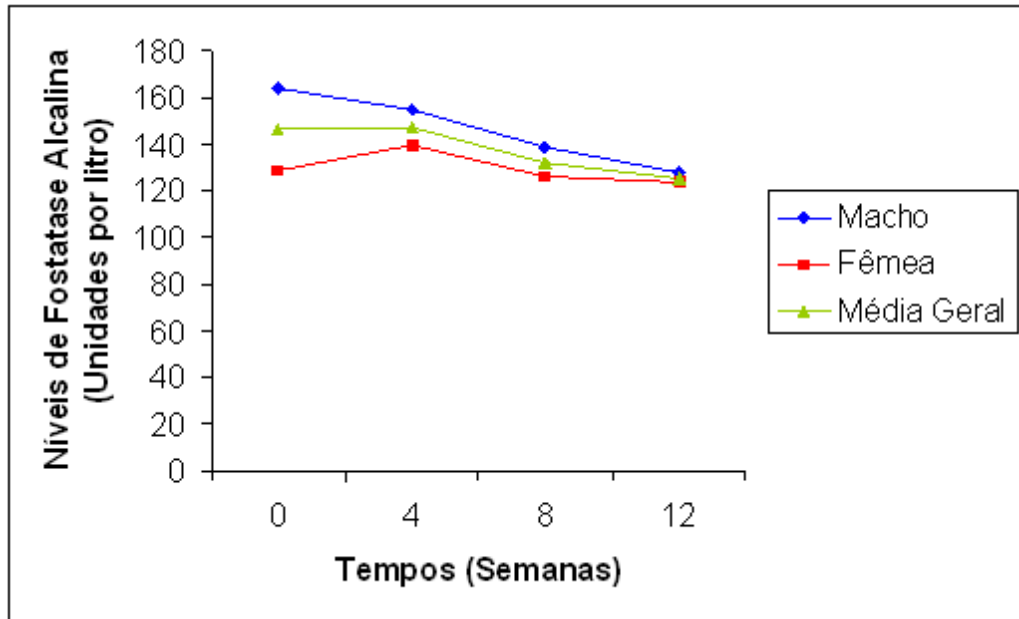


Figura 2: Perfil médio de unidades de fosfatase alcalina por litro.

Através das Figuras 2 e 3, nota-se que parece haver diferenças entre machos e fêmeas quanto à quantidade de unidades de enzima por litro. Nota-se, também, que parece haver uma diminuição destas unidades nos quatro tempos considerados.

5.2 Ajustes dos modelos propostos no Capítulo 3

Para analisar os dados da Tabela 2, foi considerada a distribuição de Poisson para X_{li} , $i = 1, \dots, 32$; $l = 1, 2, 3, 4$ (3.1) e os diferentes modelos de regressão introduzidos no Capítulo 3, na presença ou não de um vetor de covariáveis $\mathbf{Z}'_i = (Z_{1i}, Z_{2i})$ para cada animal, $i = 1, \dots, 32$, sendo que Z_{1i} denota a covariável gênero ($Z_{1i} = 1$ para machos e $Z_{1i} = 0$ para fêmeas) e Z_{2i} denota a covariável dose ($Z_{2i} = 0, 8, 25$ ou 75). Associados a cada covariável, tem-se os parâmetros de regressão $\beta_{11}, \beta_{12}, \beta_{13}$ e β_{14} (covariável Z_{1i} nos 4 tempos) e $\beta_{21}, \beta_{22}, \beta_{23}$ e β_{24} (covariável Z_{2i} nos 4 tempos).

5.2.1 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Normal para os efeitos aleatórios (Modelo 1)

Considerando-se as distribuições *a priori* em (3.5) com $a_{jl}^2 = 10^6$ e $b_l = c_l = d = f = 0,001$, $j = 1, 2$ e $l = 1, 2, 3, 4$, foram simuladas duas cadeias, cada uma com 330000 amostras das distribuições condicionais para cada parâmetro. Destas amostras, foram descartadas as 30000 primeiras para eliminar os efeitos dos valores iniciais ("*burn-in samples*") e consideradas as iterações $100^a, 200^a, 300^a, \dots$ ("*salto*" de tamanho 100), para evitar autocorrelações entre os valores gerados. Assim, a amostra final utilizada na obtenção dos sumários *a posteriori* dos parâmetros de interesse foi de tamanho 6000.

Uma grande simplificação na geração das amostras da distribuição *a posteriori* de interesse se dá através do uso do *software* Winbugs (SPIEGELHALTER et. al., 2003), que só requer a introdução da distribuição dos dados e das distribuições *a priori* para os parâmetros; Assim, optou-se pelo uso deste *software* na geração das amostras. A convergência do algoritmo foi verificada graficamente e através do critério de Gelman-Rubin.

A Tabela 4 apresenta as estimativas dos parâmetros e os respectivos intervalos de credibilidade (IC) do modelo sem a inclusão de covariáveis. Da mesma maneira, a Tabela 5 apresenta os resultados para o modelo com a inclusão de covariáveis.

Tabela 4: Estimativas *a posteriori* dos parâmetros do modelo sem covariáveis

parâmetros	Média	IC (95%)	
α_1	142,00	129,00	155,80
α_2	142,30	128,90	156,30
α_3	128,30	116,50	141,10
α_4	121,80	110,50	134,10
σ_w^2	0,07	0,04	0,12
$\alpha_1 - \alpha_2$	-0,34	-6,04	5,50
$\alpha_1 - \alpha_3$	13,61	7,85	19,40
$\alpha_1 - \alpha_4$	20,20	14,41	26,26
$\alpha_2 - \alpha_3$	13,95	8,08	19,77
$\alpha_2 - \alpha_4$	20,54	14,74	26,49
$\alpha_3 - \alpha_4$	6,60	1,29	12,01

Dos resultados da Tabela 4, nota-se que somente o parâmetro $\alpha_1 - \alpha_2$ não é significativo, ou seja, o intervalo de credibilidade contém o valor zero. Assim, a quantidade de enzimas no primeiro e segundo tempos é maior que a quantidade de enzimas dos outros tempos, isto é, houve diminuição da quantidade de fosfatase alcalina a partir de 8 semanas, continuando este decaimento na semana 12.

Tabela 5: Estimativas *a posteriori* dos parâmetros do modelo considerando a presença de covariáveis.

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
α_1	135,70	117,40	155,80	α_3	99,97	86,50	115,20
α_2	118,10	102,40	135,30	α_4	90,31	78,07	103,80
σ_w^2	0,05	0,03	0,08				
Gênero				Dose			
β_{11}	0,226	0,062	0,389	β_{21}	-0,002	-0,005	$4,9 \times 10^{-4}$
β_{12}	0,108	-0,055	0,280	β_{22}	0,005	0,002	0,008
β_{13}	0,103	-0,063	0,271	β_{23}	0,007	0,004	0,010
β_{14}	0,053	-0,111	0,220	β_{24}	0,009	0,006	0,012

Dos resultados da Tabela 5, observa-se que os parâmetros relacionados à covariável gênero são significativos apenas para o tempo 1, ou seja, o parâmetro β_{11} possui intervalo de credibilidade que não contém o valor zero. Para a covariável dose, os parâmetros significativos são para as doses 2, 3 e 4 (β_{22} , β_{23} e β_{24}).

Seja $\theta_{ljk} = \alpha_l \exp(\beta_{1l}z_{1j} + \beta_{2l}z_{2k})$ a quantidade média das unidades de enzimas por litro para o l -ésimo tempo ($l = 1, 2, 3, 4$), o j -ésimo gênero ($j = 1$ se machos e $j = 0$ se fêmeas, ou seja, $z_{11} = 1$ e $z_{10} = 0$) e a k -ésima dose ($k = 1, 2, 3, 4$ para as doses $z_{21} = 0$, $z_{22} = 8$, $z_{23} = 25$ e $z_{24} = 75$, respectivamente). A Tabela 6 apresenta as comparações destas quantidades para os todos os tempos dentro dos gêneros variando as doses.

Tabela 6: Comparações das estimativas *a posteriori* dos parâmetros considerando todos os tempos, dentro de cada gênero e variando as doses.

Tempo	Fêmeas				Machos			
	Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
1	$\theta_{101} - \theta_{102}$	2,70	-0,30	6,24	$\theta_{111} - \theta_{112}$	3,38	-0,37	7,71
	$\theta_{101} - \theta_{103}$	8,21	-0,95	18,66	$\theta_{111} - \theta_{113}$	10,26	-1,17	23,09
	$\theta_{101} - \theta_{104}$	22,73	-2,86	49,33	$\theta_{111} - \theta_{114}$	28,43	-3,54	61,04
	$\theta_{102} - \theta_{103}$	5,51	-0,64	12,43	$\theta_{112} - \theta_{113}$	6,89	-0,80	15,36
	$\theta_{102} - \theta_{104}$	20,03	-2,56	42,95	$\theta_{112} - \theta_{114}$	25,05	-3,16	53,34
	$\theta_{103} - \theta_{104}$	14,53	-1,92	30,53	$\theta_{113} - \theta_{114}$	18,16	-2,37	37,96
2	$\theta_{201} - \theta_{202}$	-4,72	-7,06	-2,12	$\theta_{211} - \theta_{212}$	-5,25	-7,82	-2,34
	$\theta_{201} - \theta_{203}$	-15,43	-23,57	-6,75	$\theta_{211} - \theta_{213}$	-17,17	-26,10	-7,46
	$\theta_{201} - \theta_{204}$	-53,23	-86,27	-21,40	$\theta_{211} - \theta_{214}$	-59,22	-95,31	-23,66
	$\theta_{202} - \theta_{203}$	-10,71	-16,51	-4,63	$\theta_{212} - \theta_{213}$	-11,92	-18,26	-5,12
	$\theta_{202} - \theta_{204}$	-48,51	-79,19	-19,28	$\theta_{212} - \theta_{214}$	-53,97	-87,55	-21,31
	$\theta_{203} - \theta_{204}$	-37,8	-62,66	-14,66	$\theta_{213} - \theta_{214}$	-42,05	-69,30	-16,20
3	$\theta_{301} - \theta_{302}$	-5,70	-7,69	-3,59	$\theta_{311} - \theta_{312}$	-6,32	-8,52	-3,97
	$\theta_{301} - \theta_{303}$	-18,98	-26,12	-11,62	$\theta_{311} - \theta_{313}$	-21,01	-28,89	-12,85
	$\theta_{301} - \theta_{304}$	-69,05	-102,10	-38,97	$\theta_{311} - \theta_{314}$	-76,44	-112,40	-42,91
	$\theta_{302} - \theta_{303}$	-13,28	-18,44	-8,04	$\theta_{312} - \theta_{313}$	-14,70	-20,37	-8,89
	$\theta_{302} - \theta_{304}$	-63,34	-94,42	-35,40	$\theta_{312} - \theta_{314}$	-70,13	-104,00	-38,98
	$\theta_{303} - \theta_{304}$	-50,07	-75,77	-27,35	$\theta_{313} - \theta_{314}$	-55,43	-83,6	-30,15
4	$\theta_{401} - \theta_{402}$	-6,84	-8,67	-4,94	$\theta_{411} - \theta_{412}$	-7,21	-9,12	-5,20
	$\theta_{401} - \theta_{403}$	-23,21	-29,94	-16,33	$\theta_{411} - \theta_{413}$	-24,44	-31,55	-17,20
	$\theta_{401} - \theta_{404}$	-89,70	-123,40	-58,15	$\theta_{411} - \theta_{414}$	-94,47	-130,00	-61,09
	$\theta_{402} - \theta_{403}$	-16,36	-21,28	-11,39	$\theta_{412} - \theta_{413}$	-17,23	-22,43	-11,99
	$\theta_{402} - \theta_{404}$	-82,86	-114,8	-53,24	$\theta_{412} - \theta_{414}$	-87,26	-121,10	-55,85
	$\theta_{403} - \theta_{404}$	-66,50	-93,62	-41,81	$\theta_{413} - \theta_{414}$	-70,03	-98,67	-43,92

Dos resultados da Tabela 6, nota-se que para o tempo 1, não existem diferenças

entre as médias das unidades de enzimas para as diferentes doses, tanto considerando machos quanto fêmeas (todos os intervalos de credibilidade contêm o valor zero). Para os demais tempos, todas as diferenças são significativas, ou seja, para doses maiores do composto, o número médio de unidades de enzima aumenta, tanto para machos quanto para fêmeas.

A Tabela 7 apresenta as comparações entre gêneros fixando-se as mesmas doses, considerando-se todos os tempos.

Tabela 7: Comparações das estimativas *a posteriori* dos parâmetros entre gêneros dentro dos tempos para cada dose.

Tempo	parâmetros	Média	IC (95%)	
1	$\theta_{111} - \theta_{101}$	34,16	8,43	59,83
	$\theta_{112} - \theta_{102}$	33,49	8,26	58,64
	$\theta_{113} - \theta_{103}$	32,11	8,03	56,05
	$\theta_{114} - \theta_{104}$	28,47	7,05	50,29
2	$\theta_{211} - \theta_{201}$	13,31	-7,58	33,80
	$\theta_{212} - \theta_{202}$	13,84	-7,89	35,08
	$\theta_{213} - \theta_{203}$	15,05	-8,56	38,28
	$\theta_{214} - \theta_{204}$	19,31	-10,84	49,25
3	$\theta_{311} - \theta_{301}$	10,71	-6,78	28,37
	$\theta_{312} - \theta_{302}$	11,32	-7,15	29,96
	$\theta_{313} - \theta_{303}$	12,74	-8,00	33,64
	$\theta_{314} - \theta_{304}$	18,10	-11,10	48,29
4	$\theta_{411} - \theta_{401}$	4,79	-11,16	20,48
	$\theta_{412} - \theta_{402}$	5,15	-12,08	21,94
	$\theta_{413} - \theta_{403}$	6,02	-13,98	25,73
	$\theta_{414} - \theta_{404}$	9,56	-21,94	40,90

Dos resultados da Tabela 7, nota-se que só existem diferenças significativas entre machos e fêmeas no tempo basal (tempo inicial do estudo) para todas as doses. Nos demais tempos, todas as diferenças são não-significativas, ou seja, possuem intervalos de

credibilidade que contêm o valor zero.

A Tabela 8 apresenta as comparações entre os tempos, dentro de cada dose para ambos os gêneros.

Tabela 8: Comparações das estimativas *a posteriori* dos parâmetros entre os tempos dentro de cada dose variando o gênero.

Dose	Fêmeas				Machos			
	Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
1	$\theta_{101} - \theta_{201}$	17,75	8,34	27,39	$\theta_{111} - \theta_{211}$	38,61	27,72	49,94
	$\theta_{101} - \theta_{301}$	35,85	26,15	46,39	$\theta_{111} - \theta_{311}$	59,31	47,64	72,02
	$\theta_{101} - \theta_{401}$	45,54	35,69	56,61	$\theta_{111} - \theta_{411}$	74,92	62,34	88,94
	$\theta_{201} - \theta_{301}$	18,10	10,09	26,87	$\theta_{211} - \theta_{311}$	20,70	12,15	29,88
	$\theta_{201} - \theta_{401}$	27,79	19,63	37,12	$\theta_{211} - \theta_{411}$	36,31	27,46	46,38
	$\theta_{301} - \theta_{401}$	9,69	2,32	17,37	$\theta_{311} - \theta_{411}$	15,61	7,94	23,62
2	$\theta_{102} - \theta_{202}$	10,34	1,79	19,08	$\theta_{112} - \theta_{212}$	29,98	20,16	39,98
	$\theta_{102} - \theta_{302}$	27,45	18,83	36,74	$\theta_{112} - \theta_{312}$	49,61	39,42	60,59
	$\theta_{102} - \theta_{402}$	36,00	27,33	45,74	$\theta_{112} - \theta_{412}$	64,33	53,44	76,42
	$\theta_{202} - \theta_{302}$	17,11	9,39	25,48	$\theta_{212} - \theta_{312}$	19,63	11,35	28,38
	$\theta_{202} - \theta_{402}$	25,66	17,85	34,54	$\theta_{212} - \theta_{412}$	34,35	25,80	43,66
	$\theta_{302} - \theta_{402}$	8,55	1,37	16,04	$\theta_{312} - \theta_{412}$	14,72	7,14	22,55
3	$\theta_{103} - \theta_{203}$	-5,88	-13,77	2,00	$\theta_{113} - \theta_{213}$	11,17	2,61	19,82
	$\theta_{103} - \theta_{303}$	8,66	0,95	16,51	$\theta_{113} - \theta_{313}$	28,03	19,37	37,21
	$\theta_{103} - \theta_{403}$	14,13	6,42	21,94	$\theta_{113} - \theta_{413}$	40,21	31,37	49,79
	$\theta_{203} - \theta_{303}$	14,55	6,79	22,46	$\theta_{213} - \theta_{313}$	16,86	8,79	25,32
	$\theta_{203} - \theta_{403}$	20,01	12,34	28,40	$\theta_{213} - \theta_{413}$	29,04	20,82	37,75
	$\theta_{303} - \theta_{403}$	5,46	-1,80	12,94	$\theta_{313} - \theta_{413}$	12,18	4,65	19,93
4	$\theta_{104} - \theta_{204}$	-58,21	-75,22	-43,52	$\theta_{114} - \theta_{214}$	-49,05	-66,41	-33,14
	$\theta_{104} - \theta_{304}$	-55,93	-72,39	-41,61	$\theta_{114} - \theta_{314}$	-45,57	-63,06	-30,40
	$\theta_{104} - \theta_{404}$	-66,90	-85,44	-50,61	$\theta_{114} - \theta_{414}$	-47,99	-65,03	-32,15
	$\theta_{204} - \theta_{304}$	2,27	-11,61	16,14	$\theta_{214} - \theta_{314}$	3,48	-12,26	19,08
	$\theta_{204} - \theta_{404}$	-8,69	-23,17	5,75	$\theta_{214} - \theta_{414}$	1,06	-14,71	17,01
	$\theta_{304} - \theta_{404}$	-10,96	-25,26	3,53	$\theta_{314} - \theta_{414}$	-2,42	-18,05	12,88

Dos resultados da Tabela 8, considerando-se as doses 1 e 2, todas as diferenças

são significativas, ou seja, houve diminuição do número de enzimas conforme o tempo aumenta. Para a dose 3, somente as diferenças entre os tempos 1 e 2 e entre os tempos 3 e 4 são não significativas, para as fêmeas. Nos machos, todas as diferenças são significativas. Para a dose 4, as diferenças entre o tempo 1 e os demais foram significativas, ou seja, houve aumento do número de enzimas somente comparando-se o tempo 1 com os restantes. Comparando-se os tempos 2, 3 e 4 entre si, não existem diferenças significativas, tanto para machos quanto para fêmeas.

5.2.2 Comparação dos modelos na presença ou não de covariáveis assumindo uma mistura de duas distribuições Normais para os efeitos aleatórios (Modelo 2)

Assumindo-se uma mistura de duas distribuições Normais em (3.8), $R = 2$, para os efeitos aleatórios e as distribuições *a priori* em (3.9) e (3.5) com $f_1 = 0, 5$, $f_2 = -0, 5$, $g_1^2 = g_2^2 = 30$, $h_1 = h_2 = q_1 = q_2 = 10$, $a_{jl}^2 = 10^6$ e $b_l = c_l = 0, 001$, $j = 1, 2$ e $l = 1, 2, 3, 4$, foram simuladas duas cadeias, cada uma com 31000000 amostras das distribuições condicionais para cada parâmetro. Destas amostras, foram descartadas as 1000000 primeiras para eliminar os efeitos dos valores iniciais ("*burn-in samples*") e consideradas as iterações 10000^a, 20000^a, 30000^a, ... ("*salto*" de tamanho 10000), para evitar autocorrelações entre os valores gerados. Assim, a amostra final utilizada na obtenção dos sumários *a posteriori* dos parâmetros de interesse foi de tamanho 6000.

A Tabela 9 apresenta as estimativas dos parâmetros e os respectivos intervalos de credibilidade (IC) do modelo sem a inclusão de covariáveis. Da mesma maneira, a Tabela 10 apresenta os resultados para o modelo com a inclusão de covariáveis.

Tabela 9: Estimativas *a posteriori* dos parâmetros do modelo sem covariáveis

parâmetros	Média	IC (95%)	
α_1	143,60	65,05	285,60
α_2	143,90	65,26	286,70
α_3	129,80	58,68	259,00
α_4	123,10	55,88	245,80
σ_1^2	0,87	0,08	5,19
σ_2^2	1,34	0,09	7,62
μ_1	0,44	0,07	0,80
μ_2	-0,42	-0,78	-0,05
$\alpha_1 - \alpha_2$	-0,28	-7,51	6,31
$\alpha_1 - \alpha_3$	13,74	5,11	30,34
$\alpha_1 - \alpha_4$	20,46	8,48	43,42
$\alpha_2 - \alpha_3$	14,02	5,26	31,28
$\alpha_2 - \alpha_4$	20,74	8,64	43,84
$\alpha_3 - \alpha_4$	6,726	0,97	18,05

Dos resultados da Tabela 9, nota-se que somente o parâmetro $\alpha_1 - \alpha_2$ não é significativo, ou seja, o intervalo de credibilidade contém o valor zero. Assim, a quantidade de enzimas no primeiro e segundo tempos é maior que a quantidade de enzimas dos outros tempos, isto é, houve diminuição da quantidade de fosfatase alcalina a partir de 8 semanas, continuando este decaimento na semana 12.

Tabela 10: Estimativas *a posteriori* dos parâmetros do modelo considerando a presença de covariáveis.

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
α_1	137,40	56,12	289,80	μ_1	0,42	0,07	0,78
α_2	119,50	48,95	250,20	μ_2	-0,41	-0,78	-0,05
α_3	101,30	41,22	211,60	σ_1^2	0,72	0,33	1,62
α_4	91,49	37,29	193,10	σ_2^2	0,79	0,35	1,70
Gênero				Dose			
β_{11}	0,207	-0,323	0,743	β_{21}	-0,001	-0,010	0,008
β_{12}	0,091	-0,435	0,629	β_{22}	0,006	-0,003	0,015
β_{13}	0,084	-0,435	0,614	β_{23}	0,008	$-9,4 \times 10^{-4}$	0,017
β_{14}	0,035	-0,493	0,571	β_{24}	0,010	0,001	0,019

Dos resultados da Tabela 10, observa-se que todos os parâmetros relacionados à covariável gênero são não-significativos (possuem intervalos de credibilidade que não contêm o valor zero. Para a covariável dose, o parâmetro significativo refere-se à dose 4 (β_{24}).

Seja a quantidade θ_{ijk} definida anteriormente. A Tabela 11 apresenta as comparações destas quantidades para os todos os tempos dentro dos gêneros variando as doses.

Tabela 11: Comparações das estimativas *a posteriori* dos parâmetros considerando todos os tempos, dentro de cada gênero e variando as doses.

Tempo	Fêmeas				Machos			
	Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
1	$\theta_{101} - \theta_{102}$	2,08	-7,57	16,41	$\theta_{111} - \theta_{112}$	2,59	-9,41	19,91
	$\theta_{101} - \theta_{103}$	5,80	-25,21	47,84	$\theta_{111} - \theta_{113}$	7,23	-30,86	58,02
	$\theta_{101} - \theta_{104}$	11,80	-90,49	115,90	$\theta_{111} - \theta_{114}$	14,81	-111,80	143,00
	$\theta_{102} - \theta_{103}$	3,72	-17,62	31,26	$\theta_{112} - \theta_{113}$	4,64	-21,40	38,10
	$\theta_{102} - \theta_{104}$	9,72	-82,86	99,59	$\theta_{112} - \theta_{114}$	12,22	-102,60	122,10
	$\theta_{103} - \theta_{104}$	6,00	-65,26	68,48	$\theta_{113} - \theta_{114}$	7,58	-81,19	83,56
2	$\theta_{201} - \theta_{202}$	-5,35	-16,74	3,36	$\theta_{211} - \theta_{212}$	-5,84	-18,50	3,70
	$\theta_{201} - \theta_{203}$	-18,13	-58,13	10,25	$\theta_{211} - \theta_{213}$	-19,80	-64,13	11,28
	$\theta_{201} - \theta_{204}$	-70,41	-246,40	28,68	$\theta_{211} - \theta_{214}$	-76,86	-269,10	31,55
	$\theta_{202} - \theta_{203}$	-12,78	-41,45	6,89	$\theta_{212} - \theta_{213}$	-13,96	-45,66	7,58
	$\theta_{202} - \theta_{204}$	-65,07	-229,90	25,33	$\theta_{212} - \theta_{214}$	-71,02	-251,50	27,82
	$\theta_{203} - \theta_{204}$	-52,28	-189,00	18,39	$\theta_{213} - \theta_{214}$	-57,06	-206,50	20,38
3	$\theta_{301} - \theta_{302}$	-6,28	-16,96	0,79	$\theta_{311} - \theta_{312}$	-6,81	-18,58	0,89
	$\theta_{301} - \theta_{303}$	-21,48	-59,66	2,43	$\theta_{311} - \theta_{313}$	-23,31	-65,15	2,75
	$\theta_{301} - \theta_{304}$	-86,25	-267,10	7,13	$\theta_{311} - \theta_{314}$	-93,57	-291,70	8,06
	$\theta_{302} - \theta_{303}$	-15,20	-42,71	1,65	$\theta_{312} - \theta_{313}$	-16,50	-46,90	1,86
	$\theta_{302} - \theta_{304}$	-79,98	-250,30	6,35	$\theta_{312} - \theta_{314}$	-86,76	-273,10	7,17
	$\theta_{303} - \theta_{304}$	-64,78	-208,90	4,70	$\theta_{313} - \theta_{314}$	-70,26	-227,40	5,30
4	$\theta_{401} - \theta_{402}$	-7,39	-17,98	-0,90	$\theta_{411} - \theta_{412}$	-7,64	-18,76	-0,99
	$\theta_{401} - \theta_{403}$	-25,66	-64,62	-2,88	$\theta_{411} - \theta_{413}$	-26,53	-66,90	-3,13
	$\theta_{401} - \theta_{404}$	-108,30	-305,20	-9,11	$\theta_{411} - \theta_{414}$	-111,90	-318,60	-9,71
	$\theta_{402} - \theta_{403}$	-18,27	-46,64	-1,97	$\theta_{412} - \theta_{413}$	-18,88	-48,07	-2,14
	$\theta_{402} - \theta_{404}$	-100,90	-287,40	-8,18	$\theta_{412} - \theta_{414}$	-104,20	-300,70	-8,70
	$\theta_{403} - \theta_{404}$	-82,64	-241,90	-6,17	$\theta_{413} - \theta_{414}$	-85,36	-253,00	-6,57

Dos resultados da Tabela 11, nota-se que para os tempos 1, 2 e 3 não existem

diferenças entre as médias das unidades de enzimas para as diferentes doses, tanto considerando machos quanto fêmeas (todos os intervalos de credibilidade contêm o valor zero). Para tempo 4, todas as diferenças são significativas, ou seja, para doses maiores do composto, o número médio de unidades de enzima aumenta, tanto para machos quanto para fêmeas.

A Tabela 12 apresenta as comparações entre gêneros fixando-se as mesmas doses, considerando-se todos os tempos.

Tabela 12: Comparações das estimativas *a posteriori* dos parâmetros entre gêneros dentro dos tempos para cada dose.

Tempo	parâmetros	Média	IC (95%)	
1	$\theta_{111} - \theta_{101}$	31,90	-51,48	141,90
	$\theta_{112} - \theta_{102}$	31,39	-50,02	139,10
	$\theta_{113} - \theta_{103}$	30,47	-48,59	135,50
	$\theta_{114} - \theta_{104}$	28,89	-46,36	133,50
2	$\theta_{211} - \theta_{201}$	11,54	-59,24	94,97
	$\theta_{212} - \theta_{202}$	12,04	-62,49	98,50
	$\theta_{213} - \theta_{203}$	13,22	-68,12	108,30
	$\theta_{214} - \theta_{204}$	17,99	-95,85	148,20
3	$\theta_{311} - \theta_{301}$	9,08	-52,35	79,53
	$\theta_{312} - \theta_{302}$	9,62	-55,18	82,80
	$\theta_{313} - \theta_{303}$	10,92	-61,86	93,26
	$\theta_{314} - \theta_{304}$	16,40	-93,79	144,10
4	$\theta_{411} - \theta_{401}$	3,31	-52,48	62,35
	$\theta_{412} - \theta_{402}$	3,57	-56,77	66,61
	$\theta_{413} - \theta_{403}$	4,18	-67,21	78,46
	$\theta_{414} - \theta_{404}$	6,90	-116,10	132,60

Dos resultados da Tabela 12, nota-se que todas as diferenças são não-significativas, ou seja, não existem diferenças entre machos e fêmeas.

A Tabela 13 apresenta as comparações entre os tempos, dentro de cada dose para

ambos os gêneros.

Tabela 13: Comparações das estimativas *a posteriori* dos parâmetros entre os tempos dentro de cada dose variando o gênero.

Dose	Fêmeas				Machos			
	Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
1	$\theta_{101} - \theta_{201}$	17,88	5,44	43,25	$\theta_{111} - \theta_{211}$	38,23	14,43	84,29
	$\theta_{101} - \theta_{301}$	36,05	13,94	79,21	$\theta_{111} - \theta_{311}$	58,87	23,45	127,00
	$\theta_{101} - \theta_{401}$	45,88	17,96	98,54	$\theta_{111} - \theta_{411}$	74,47	29,96	158,00
	$\theta_{201} - \theta_{301}$	18,18	6,03	42,53	$\theta_{211} - \theta_{311}$	20,64	7,14	47,71
	$\theta_{201} - \theta_{401}$	28,01	10,65	60,51	$\theta_{211} - \theta_{411}$	36,24	13,96	79,07
	$\theta_{301} - \theta_{401}$	9,83	1,53	25,23	$\theta_{311} - \theta_{411}$	15,60	4,70	36,84
2	$\theta_{102} - \theta_{202}$	10,45	1,58	27,77	$\theta_{112} - \theta_{212}$	29,80	11,13	65,72
	$\theta_{102} - \theta_{302}$	27,70	10,56	61,33	$\theta_{112} - \theta_{312}$	49,47	20,01	104,80
	$\theta_{102} - \theta_{402}$	36,42	14,51	77,78	$\theta_{112} - \theta_{412}$	64,24	26,29	134,70
	$\theta_{202} - \theta_{302}$	17,25	5,76	40,21	$\theta_{212} - \theta_{312}$	19,67	6,93	44,71
	$\theta_{202} - \theta_{402}$	25,96	9,92	55,25	$\theta_{212} - \theta_{412}$	34,44	13,53	74,25
	$\theta_{302} - \theta_{402}$	8,72	0,85	22,81	$\theta_{312} - \theta_{412}$	14,77	4,46	34,90
3	$\theta_{103} - \theta_{203}$	-6,05	-18,98	1,76	$\theta_{113} - \theta_{213}$	11,20	1,84	28,54
	$\theta_{103} - \theta_{303}$	8,78	0,91	23,53	$\theta_{113} - \theta_{313}$	28,33	10,96	60,95
	$\theta_{103} - \theta_{403}$	14,43	4,50	33,28	$\theta_{113} - \theta_{413}$	40,71	16,66	84,85
	$\theta_{203} - \theta_{303}$	14,83	4,51	34,66	$\theta_{213} - \theta_{313}$	17,13	5,74	39,89
	$\theta_{203} - \theta_{403}$	20,48	7,60	44,19	$\theta_{213} - \theta_{413}$	29,51	11,67	62,77
	$\theta_{303} - \theta_{403}$	5,65	-1,84	17,48	$\theta_{313} - \theta_{413}$	12,38	2,99	30,21
4	$\theta_{104} - \theta_{204}$	-64,34	-147,10	-22,72	$\theta_{114} - \theta_{214}$	-53,44	-123,80	-18,02
	$\theta_{104} - \theta_{304}$	-62,00	-143,60	-21,57	$\theta_{114} - \theta_{314}$	-49,51	-116,70	-16,80
	$\theta_{104} - \theta_{404}$	-74,22	-172,10	-26,08	$\theta_{114} - \theta_{414}$	-52,23	-121,20	-17,50
	$\theta_{204} - \theta_{304}$	2,34	-14,50	22,24	$\theta_{214} - \theta_{314}$	3,93	-13,77	27,00
	$\theta_{204} - \theta_{404}$	-9,88	-35,67	5,11	$\theta_{214} - \theta_{414}$	1,21	-18,03	22,22
	$\theta_{304} - \theta_{404}$	-12,22	-40,63	3,07	$\theta_{314} - \theta_{414}$	-2,72	-25,47	16,47

Dos resultados da Tabela 13, considerando-se as doses 1 e 2, todas as diferenças são significativas, ou seja, houve diminuição do número de enzimas conforme o tempo aumenta. Para a dose 3, somente as diferenças entre os tempos 1 e 2 e entre os tempos 3 e 4 são não significativas, para as fêmeas. Nos machos, todas as diferenças são significativas. Para a dose 4, as diferenças entre o tempo 1 e os demais foram significativas, ou seja, houve aumento do número de enzimas somente comparando-se o tempo 1 com os restantes. Comparando-se os tempos 2, 3 e 4 entre si, não existem diferenças significativas, tanto para machos quanto para fêmeas.

5.2.3 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Gama para os efeitos aleatórios (Modelo 3)

Considerando-se as distribuições *a priori* em (3.23) com $a_{jl}^2 = 10^6$ e $b_l = c_l = d = f = 0,001$, $j = 1, 2$ e $l = 1, 2, 3, 4$, foram simuladas duas cadeias, cada uma com 330000 amostras das distribuições condicionais para cada parâmetro. Destas amostras, foram descartadas as 30000 primeiras para eliminar os efeitos dos valores iniciais ("*burn-in samples*") e consideradas as iterações $100^a, 200^a, 300^a, \dots$ ("salto" de tamanho 100), para evitar autocorrelações entre os valores gerados. Assim, a amostra final utilizada na obtenção dos sumários *a posteriori* dos parâmetros de interesse foi de tamanho 6000.

A Tabela 14 apresenta as estimativas dos parâmetros e os respectivos intervalos de credibilidade (IC) do modelo sem a inclusão de covariáveis. Da mesma maneira, a Tabela 15 apresenta os resultados para o modelo com a inclusão de covariáveis.

Tabela 14: Estimativas *a posteriori* dos parâmetros do modelo sem covariáveis

parâmetros	Média	IC (95%)	
α_1	146,80	133,40	162,00
α_2	147,20	133,90	162,30
α_3	132,80	120,60	146,40
α_4	126,00	114,30	139,00
ϕ	0,069	0,041	0,115
$\alpha_1 - \alpha_2$	-0,35	-6,26	5,76
$\alpha_1 - \alpha_3$	14,09	8,37	20,11
$\alpha_1 - \alpha_4$	20,88	15,09	27,22
$\alpha_2 - \alpha_3$	14,44	8,521	20,62
$\alpha_2 - \alpha_4$	21,24	15,44	27,56
$\alpha_3 - \alpha_4$	6,793	1,195	12,44

Dos resultados da Tabela 14, nota-se que somente o parâmetro $\alpha_1 - \alpha_2$ não é significativo, ou seja, o intervalo de credibilidade contém o valor zero. Assim, a quantidade de enzimas no primeiro e segundo tempos é maior que a quantidade de enzimas dos outros tempos, isto é, houve diminuição da quantidade de fosfatase alcalina a partir de 8 semanas, continuando este decaimento na semana 12.

Tabela 15: Estimativas *a posteriori* dos parâmetros do modelo considerando a presença de covariáveis.

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
α_1	138,80	120,40	159,80	α_3	102,20	88,41	117,70
α_2	120,60	104,40	138,70	α_4	92,35	79,92	106,60
ϕ	0,050	0,029	0,084				
Gênero				Dose			
β_{11}	0,235	0,069	0,399	β_{21}	-0,003	-0,005	$3,1 \times 10^{-4}$
β_{12}	0,118	-0,049	0,282	β_{22}	0,005	0,002	0,008
β_{13}	0,112	-0,055	0,278	β_{23}	0,007	0,004	0,010
β_{14}	0,062	-0,105	0,228	β_{24}	0,009	0,006	0,012

Dos resultados da Tabela 15, observa-se que os parâmetros relacionados à covariável gênero são significativos apenas para o tempo 1, ou seja, o parâmetro β_{11} possui intervalo de credibilidade que não contém o valor zero. Para a covariável dose, os parâmetros significativos são para as doses 2, 3 e 4 (β_{22} , β_{23} e β_{24}).

Seja a quantidade θ_{ljk} definida anteriormente. A Tabela 16 apresenta as comparações destas quantidades para os todos os tempos dentro dos gêneros variando as doses.

Tabela 16: Comparações das estimativas *a posteriori* dos parâmetros considerando todos os tempos, dentro de cada gênero e variando as doses.

Tempo	Fêmeas				Machos			
	Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
1	$\theta_{101} - \theta_{102}$	2,90	-0,32	6,49	$\theta_{111} - \theta_{112}$	3,66	-0,40	8,16
	$\theta_{101} - \theta_{103}$	8,81	-1,01	19,38	$\theta_{111} - \theta_{113}$	11,10	-1,26	24,38
	$\theta_{101} - \theta_{104}$	24,33	-3,06	51,04	$\theta_{111} - \theta_{114}$	30,68	-3,82	64,02
	$\theta_{102} - \theta_{103}$	5,91	-0,69	12,90	$\theta_{112} - \theta_{113}$	7,45	-0,86	16,20
	$\theta_{102} - \theta_{104}$	21,43	-2,74	44,56	$\theta_{112} - \theta_{114}$	27,02	-3,42	55,90
	$\theta_{103} - \theta_{104}$	15,52	-2,05	31,71	$\theta_{113} - \theta_{114}$	19,57	-2,56	39,79
2	$\theta_{201} - \theta_{202}$	-4,66	-7,17	-2,08	$\theta_{211} - \theta_{212}$	-5,25	-8,09	-2,34
	$\theta_{201} - \theta_{203}$	-15,24	-23,86	-6,63	$\theta_{211} - \theta_{213}$	-17,17	-26,95	-7,41
	$\theta_{201} - \theta_{204}$	-52,42	-87,34	-20,94	$\theta_{211} - \theta_{214}$	-59,07	-99,33	-23,42
	$\theta_{202} - \theta_{203}$	-10,58	-16,68	-4,54	$\theta_{212} - \theta_{213}$	-11,92	-18,87	-5,08
	$\theta_{202} - \theta_{204}$	-47,76	-80,17	-18,86	$\theta_{212} - \theta_{214}$	-53,81	-91,28	-21,09
	$\theta_{203} - \theta_{204}$	-37,18	-63,51	-14,31	$\theta_{213} - \theta_{214}$	-41,90	-72,16	-15,99
3	$\theta_{301} - \theta_{302}$	-5,69	-7,76	-3,58	$\theta_{311} - \theta_{312}$	-6,37	-8,75	-3,96
	$\theta_{301} - \theta_{303}$	-18,91	-26,37	-11,61	$\theta_{311} - \theta_{313}$	-21,18	-29,78	-12,84
	$\theta_{301} - \theta_{304}$	-68,52	-102,30	-38,63	$\theta_{311} - \theta_{314}$	-76,76	-114,50	-42,90
	$\theta_{302} - \theta_{303}$	-13,22	-18,60	-8,02	$\theta_{312} - \theta_{313}$	-14,80	-20,99	-8,88
	$\theta_{302} - \theta_{304}$	-62,83	-94,53	-35,08	$\theta_{312} - \theta_{314}$	-70,38	-105,70	-38,92
	$\theta_{303} - \theta_{304}$	-49,61	-75,84	-27,01	$\theta_{313} - \theta_{314}$	-55,58	-84,94	-30,01
4	$\theta_{401} - \theta_{402}$	-6,87	-8,80	-5,03	$\theta_{411} - \theta_{412}$	-7,32	-9,36	-5,27
	$\theta_{401} - \theta_{403}$	-23,27	-30,39	-16,63	$\theta_{411} - \theta_{413}$	-24,78	-32,35	-17,39
	$\theta_{401} - \theta_{404}$	-89,56	-125,40	-58,87	$\theta_{411} - \theta_{414}$	-95,39	-133,70	-61,59
	$\theta_{402} - \theta_{403}$	-16,40	-21,59	-11,59	$\theta_{412} - \theta_{413}$	-17,46	-23,01	-12,12
	$\theta_{402} - \theta_{404}$	-82,69	-116,70	-53,87	$\theta_{412} - \theta_{414}$	-88,08	-124,30	-56,42
	$\theta_{403} - \theta_{404}$	-66,29	-95,01	-42,25	$\theta_{413} - \theta_{414}$	-70,62	-101,30	-44,30

Dos resultados da Tabela 16, nota-se que para o tempo 1, não existem diferenças

entre as médias das unidades de enzimas para as diferentes doses, tanto considerando machos quanto fêmeas (todos os intervalos de credibilidade contêm o valor zero). Para os demais tempos, todas as diferenças são significativas, ou seja, para doses maiores do composto, o número médio de unidades de enzima aumenta, tanto para machos quanto para fêmeas.

A Tabela 17 apresenta as comparações entre gêneros fixando-se as mesmas doses, considerando-se todos os tempos.

Tabela 17: Comparações das estimativas *a posteriori* dos parâmetros entre gêneros dentro dos tempos para cada dose.

Tempo	parâmetros	Média	IC (95%)	
1	$\theta_{111} - \theta_{101}$	36,68	10,55	63,46
	$\theta_{112} - \theta_{102}$	35,92	10,28	62,13
	$\theta_{113} - \theta_{103}$	34,38	9,88	59,46
	$\theta_{114} - \theta_{104}$	30,33	8,76	53,80
2	$\theta_{211} - \theta_{201}$	15,13	-6,30	36,23
	$\theta_{212} - \theta_{202}$	15,72	-6,62	37,60
	$\theta_{213} - \theta_{203}$	17,06	-7,29	41,14
	$\theta_{214} - \theta_{204}$	21,77	-9,28	53,79
3	$\theta_{311} - \theta_{301}$	12,15	-5,98	30,33
	$\theta_{312} - \theta_{302}$	12,83	-6,25	31,96
	$\theta_{313} - \theta_{303}$	14,42	-7,02	35,93
	$\theta_{314} - \theta_{304}$	20,38	-9,69	51,15
4	$\theta_{411} - \theta_{401}$	5,92	-10,06	21,81
	$\theta_{412} - \theta_{402}$	6,37	-10,85	23,47
	$\theta_{413} - \theta_{403}$	7,44	-12,63	27,63
	$\theta_{414} - \theta_{404}$	11,76	-19,88	43,90

Dos resultados da Tabela 17, nota-se que só existem diferenças significativas entre machos e fêmeas no tempo basal (tempo inicial do estudo) para todas as doses. Nos demais tempos, todas as diferenças são não-significativas, ou seja, possuem intervalos de

credibilidade que contêm o valor zero.

A Tabela 18 apresenta as comparações entre os tempos, dentro de cada dose para ambos os gêneros.

Tabela 18: Comparações das estimativas *a posteriori* dos parâmetros entre os tempos dentro de cada dose variando o gênero.

Dose	Fêmeas				Machos			
	Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
1	$\theta_{101} - \theta_{201}$	18,12	8,92	27,76	$\theta_{111} - \theta_{211}$	39,67	28,62	51,45
	$\theta_{101} - \theta_{301}$	36,52	26,71	47,29	$\theta_{111} - \theta_{311}$	61,06	49,04	74,00
	$\theta_{101} - \theta_{401}$	46,41	36,13	57,54	$\theta_{111} - \theta_{411}$	77,17	64,20	91,31
	$\theta_{201} - \theta_{301}$	18,41	10,24	27,49	$\theta_{211} - \theta_{311}$	21,39	12,28	30,98
	$\theta_{201} - \theta_{401}$	28,30	19,56	37,50	$\theta_{211} - \theta_{411}$	37,50	27,98	48,02
	$\theta_{301} - \theta_{401}$	9,89	2,21	17,43	$\theta_{311} - \theta_{411}$	16,11	8,00	24,50
2	$\theta_{102} - \theta_{202}$	10,55	2,04	19,35	$\theta_{112} - \theta_{212}$	30,76	20,87	40,99
	$\theta_{102} - \theta_{302}$	27,93	19,09	37,39	$\theta_{112} - \theta_{312}$	51,03	40,53	62,12
	$\theta_{102} - \theta_{402}$	36,64	27,55	46,25	$\theta_{112} - \theta_{412}$	66,19	55,21	78,55
	$\theta_{202} - \theta_{302}$	17,38	9,47	26,11	$\theta_{212} - \theta_{312}$	20,27	11,61	29,28
	$\theta_{202} - \theta_{402}$	26,09	17,72	34,74	$\theta_{212} - \theta_{412}$	35,43	26,25	45,30
	$\theta_{302} - \theta_{402}$	8,71	1,16	16,01	$\theta_{312} - \theta_{412}$	15,17	7,25	23,31
3	$\theta_{103} - \theta_{203}$	-5,93	-13,77	1,83	$\theta_{113} - \theta_{213}$	11,39	2,77	20,26
	$\theta_{103} - \theta_{303}$	8,81	0,95	16,92	$\theta_{113} - \theta_{313}$	28,78	19,93	38,12
	$\theta_{103} - \theta_{403}$	14,34	6,65	22,14	$\theta_{113} - \theta_{413}$	41,29	32,17	51,00
	$\theta_{203} - \theta_{303}$	14,74	6,89	23,02	$\theta_{213} - \theta_{313}$	17,38	9,00	26,09
	$\theta_{203} - \theta_{403}$	20,27	12,47	28,48	$\theta_{213} - \theta_{413}$	29,89	21,16	39,19
	$\theta_{303} - \theta_{403}$	5,53	-2,02	12,89	$\theta_{313} - \theta_{413}$	12,51	4,64	20,59
4	$\theta_{104} - \theta_{204}$	-58,63	-75,92	-43,75	$\theta_{114} - \theta_{214}$	-50,08	-68,35	-34,11
	$\theta_{104} - \theta_{304}$	-56,33	-73,60	-41,65	$\theta_{114} - \theta_{314}$	-46,38	-64,51	-30,78
	$\theta_{104} - \theta_{404}$	-67,48	-86,02	-51,70	$\theta_{114} - \theta_{414}$	-48,90	-67,06	-33,39
	$\theta_{204} - \theta_{304}$	2,31	-11,95	16,49	$\theta_{214} - \theta_{314}$	3,70	-11,89	19,64
	$\theta_{204} - \theta_{404}$	-8,84	-23,64	5,46	$\theta_{214} - \theta_{414}$	1,17	-14,50	17,39
	$\theta_{304} - \theta_{404}$	-11,15	-26,19	3,57	$\theta_{314} - \theta_{414}$	-2,53	-18,65	13,29

Dos resultados da Tabela 18, considerando-se as doses 1 e 2, todas as diferenças

são significativas, ou seja, houve diminuição do número de enzimas conforme o tempo aumenta. Para a dose 3, somente as diferenças entre os tempos 1 e 2 e entre os tempos 3 e 4 são não-significativas, para as fêmeas. Nos machos, todas as diferenças são significativas. Para a dose 4, as diferenças entre o tempo 1 e os demais foram significativas, ou seja, houve aumento do número de enzimas somente comparando-se o tempo 1 com os restantes. Comparando-se os tempos 2, 3 e 4 entre si, não existem diferenças significativas, tanto para machos quanto para fêmeas.

5.3 Discriminação de modelos

A Tabela 19 apresenta algumas medidas para a discriminação do melhor modelo entre os propostos.

Tabela 19: Medidas de discriminação dos modelos propostos

Modelo		Medida		
		BIC	DIC	LPML
Normal	Sem Covariáveis	-1399,00	1416,25	-828,48
	Com Covariáveis	-1106,00	1101,16	-573,78
Mistura	Sem Covariáveis	-1413,00	1225,53	-828,28
	Com Covariáveis	-1120,00	921,63	-572,84
Gama	Sem Covariáveis	-1399,00	1416,09	-827,21
	Com Covariáveis	-1106,00	1100,71	-573,22

Dos resultados da Tabela 19, segundo o critério BIC, os melhores modelos foram o Normal e o Gama com a inclusão de covariáveis, ambos com valores iguais a $-1106,00$ (maiores valores). Segundo o critério DIC e a medida LPML, o melhor modelo foi o de mistura de normais para os efeitos aleatórios, com a inclusão das covariáveis, apresentando menor valor do DIC (921,63) e maior valor de LPML($-572,84$). Assim, este foi considerado o melhor modelo entre os propostos, já que duas medidas de discriminação o indicaram.

5.4 Variâncias não-condicionais das respostas

A Tabela 20 apresenta o cálculo das esperanças e variâncias amostrais e não-condicionais das respostas sob os diferentes modelos, considerando os machos. A Tabela 21 apresenta os mesmos cálculos para as fêmeas.

Tabela 20: Esperanças e variâncias amostrais e não-condicionais calculadas para os machos.

Estatística	Tempo							
	0				4			
	Dose				Dose			
Média	0	8	25	75	0	8	25	75
Amostral	185,0	175,8	147,5	148,0	164,8	143,0	124,5	185,5
Modelo 1	174,3	170,8	163,7	145,1	134,7	140,1	152,3	195,4
Modelo 2	165,3	162,8	158,3	151,0	128,0	133,7	147,4	203,2
Modelo 3	175,4	171,8	164,3	144,8	135,8	141,0	152,9	194,8
Variância								
Amostral	5396,7	660,9	1619,7	92,0	6200,9	808,7	935,0	622,3
Modelo 1	1697,0	1632,0	1505,0	1204,0	1045,0	1124,0	1313,0	2113,0
Modelo 2	18650,0	17950,0	16850,0	16220,0	11210,0	12130,0	14610,0	29300,0
Modelo 3	1721,0	1652,0	1518,0	1200,0	1061,0	1139,0	1325,0	2108,0

Estatística	Tempo							
	8				12			
	Dose				Dose			
Média	0	8	25	75	0	8	25	75
Amostral	134,5	119,5	119,0	181,3	112,3	107,0	116,8	176,0
Modelo 1	113,5	120,0	135,0	191,8	97,5	104,9	122,6	194,3
Modelo 2	107,8	114,5	130,6	199,4	92,6	100,1	118,6	202,1
Modelo 3	114,4	120,8	135,6	191,1	98,3	105,6	123,0	193,7
Variância								
Amostral	3295,0	463,0	887,3	787,6	1068,3	368,7	716,9	1200,7
Modelo 1	759,8	841,3	1048,0	2040,0	574,5	656,3	874,2	2091,0
Modelo 2	7978,0	8914,0	11500,0	28200,0	5900,0	6826,0	9482,0	28970,0
Modelo 3	771,3	852,2	1057,0	2032,0	583,1	664,8	881,7	2083,0

Tabela 21: Esperanças e variâncias amostrais e não-condicionais calculadas para as fêmeas.

Estatística	Tempo							
	0				4			
	Dose				Dose			
Média	0	8	25	75	0	8	25	75
Amostral	128,3	141,3	137,8	108,5	96,0	133,5	143,0	184,5
Modelo 1	139,3	136,5	130,9	116,0	121,1	125,9	136,9	175,6
Modelo 2	134,2	132,2	128,6	122,9	116,8	122,0	134,5	185,9
Modelo 3	138,8	135,9	130,0	114,4	120,6	125,3	135,9	173,1
Variância								
Amostral	138,3	400,3	360,9	249,7	38,0	489,7	1968,0	1076,3
Modelo 1	1113,0	1071,0	988,6	792,7	857,0	920,9	1075,0	1727,0
Modelo 2	12270,0	11820,0	11120,0	10790,0	9303,0	10080,0	12170,0	24590,0
Modelo 3	1106,0	1062,0	976,4	773,3	852,0	913,5	1062,0	1681,0

Estatística	Tempo							
	8				12			
	Dose				Dose			
Média	0	8	25	75	0	8	25	75
Amostral	83,3	107,8	136,5	177,8	76,3	99,0	121,8	196,5
Modelo 1	102,5	108,4	122,0	173,3	92,6	99,6	116,4	184,5
Modelo 2	99,0	105,2	120,0	183,6	89,4	96,6	114,5	195,5
Modelo 3	102,2	107,9	121,1	170,8	92,4	99,2	115,6	181,9
Variância								
Amostral	34,9	206,3	3299,0	3514,9	44,3	266,0	809,6	3629,7
Modelo 1	630,4	697,5	867,3	1683,0	523,1	597,2	794,7	1896,0
Modelo 2	6698,0	7494,0	9693,0	23950,0	5475,0	6341,0	8827,0	27160,0
Modelo 3	627,2	692,4	856,7	1639,0	520,8	593,3	785,7	1848,0

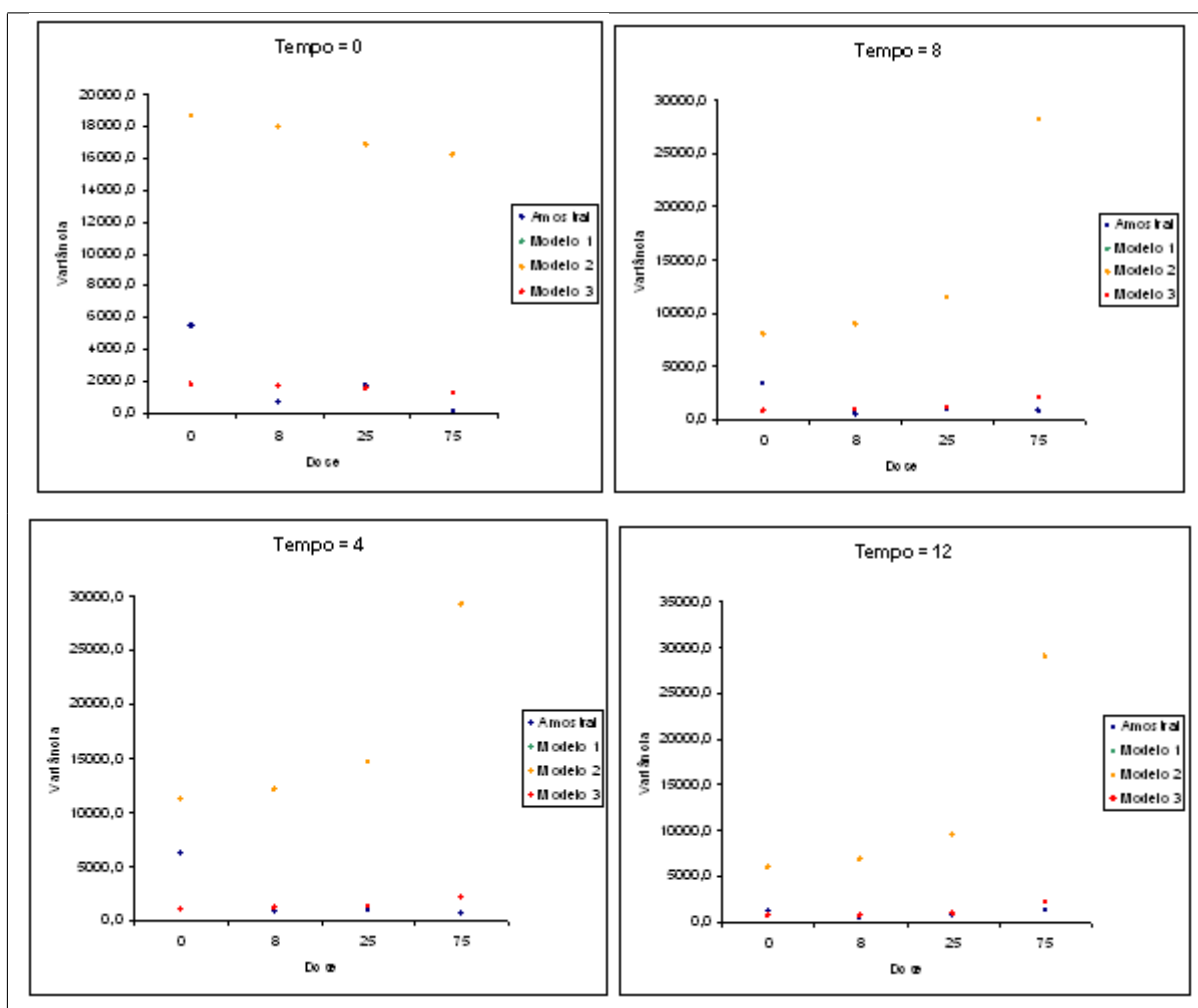


Figura 3: Gráfico das variâncias para o grupo dos machos, para todos os tempos.

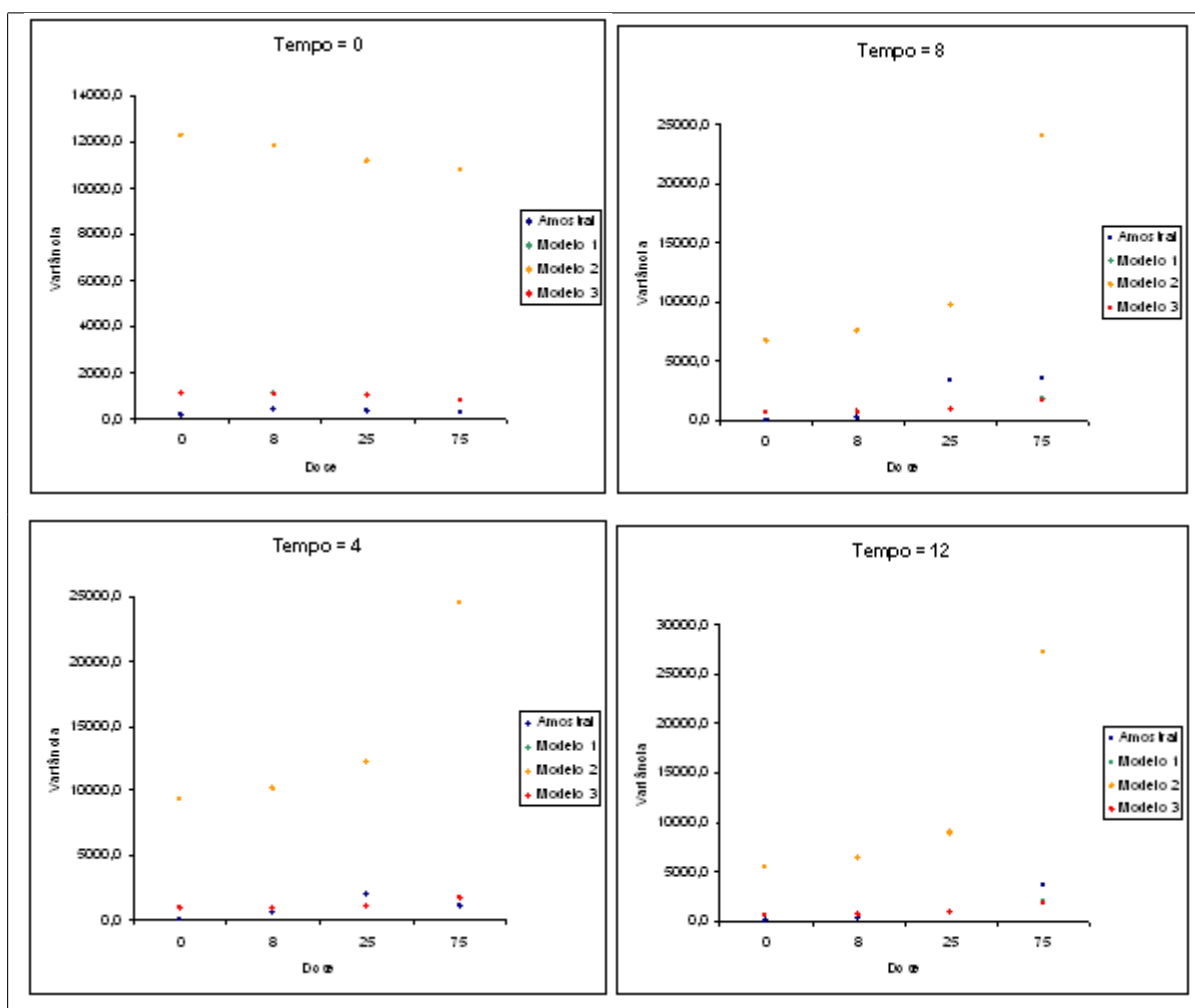


Figura 4: Gráfico das variâncias para o grupo das fêmeas, para todos os tempos.

Dos resultados das tabelas 20 e 21 e através das Figuras 3 e 4, nota-se que o modelo que leva à melhor aproximação com a variância amostral é o Modelo 3, cuja soma dos quadrados das diferenças entre as variâncias amostrais e as estimadas foi de 17636318,92.

Capítulo 6

Aplicação 2 - dados de contagem inflacionados de zeros

Neste Capítulo é apresentado um conjunto de dados reais e realizada a modelagem através das duas formas apresentadas no Capítulo 4.

6.1 Apresentação e descrição dos dados

Os dados do estudo em questão referem-se às quantidades de espécies de insetos (*Stenochironomus sp*, *Xestochironomus sp*, *Tipulidae Gen.1*, *Beardius sp* e *Oukuriella sp*) encontradas em troncos de madeira, coletados em alguns riachos (KOROIVA & FONSECA-GESSNER, 2008). O interesse do estudo era saber se estas quantidades estavam associadas a outras características observadas, tais como o volume do tronco, o tipo de superfície, a porcentagem de superfície coberta, o formato e consistência. A Tabela 22 apresenta os dados do estudo em questão.

Tabela 22: Dados do estudo sobre os insetos

Tronco	Espécies					Covariáveis				
	S1	S2	S3	S4	S5	Z ₁	Z ₂	Z ₃	Z ₄	Z ₅
1	0	0	0	0	0	389,45	1	3	1	3
2	0	0	0	0	0	448,50	0	3	2	2
3	0	0	0	0	0	204,17	1	2	1	3
4	0	0	0	0	0	139,73	1	3	2	3
5	0	0	0	0	0	196,94	1	3	3	2
6	0	0	0	0	0	288,48	1	3	2	2
7	0	0	0	0	0	31,40	1	3	2	3
8	0	0	0	0	0	17,27	1	2	2	2
9	0	0	0	0	0	592,00	0	3	1	1
10	0	0	0	0	0	92,31	1	3	2	3
11	0	0	0	0	0	21,19	1	1	2	3
12	0	0	0	0	0	2110,08	1	3	3	3
13	1	0	0	0	0	2100,00	0	3	3	1
14	1	0	0	0	0	841,52	1	1	1	3
15	1	0	0	0	0	1898,00	0	3	1	3
16	1	0	0	0	0	461,58	1	3	3	1
17	1	0	1	0	0	307,32	1	2	1	3
18	1	2	0	0	0	284,63	0	2	2	1
19	0	0	3	0	0	1560,00	0	3	1	1
20	0	0	3	0	0	59,66	1	3	1	3
21	4	2	0	0	0	309,72	1	1	2	3
22	5	1	0	0	0	682,50	0	3	1	1
23	0	0	10	0	0	376,80	1	3	1	1
24	6	3	0	1	1	2925,00	0	3	1	1
25	0	0	17	0	0	22,76	1	3	2	1

As covariáveis Z_1, Z_2, Z_3, Z_4 e Z_5 são, respectivamente, o volume do tronco (em

cm^3), o formato, a superfície, a consistência e a porcentagem de superfície coberta, cujos valores assumidos são:

$$\begin{aligned}
 Z_2 &= \begin{cases} 1, \text{ cilíndrico} \\ 0, \text{ paralelepípedo} \end{cases} \\
 Z_3 &= \begin{cases} 1, \text{ liso} \\ 2, \text{ superfície com alguns sulcos} \\ 3, \text{ superfície com muitos sulcos} \end{cases} \\
 Z_4 &= \begin{cases} 1, \text{ firme} \\ 2, \text{ macio} \\ 3, \text{ macio, mas interior firme} \end{cases} \\
 Z_5 &= \begin{cases} 1, \text{ menor ou igual a 25\%} \\ 2, \text{ entre 26 e 50\%} \\ 3, \text{ mais que 50\%} \end{cases}
 \end{aligned} \tag{6.1}$$

As covariáveis Z_3 , Z_4 e Z_5 foram transformadas nas seguintes variáveis dummies:

$$\begin{aligned}
 Z_{31} &= \begin{cases} 1, \text{ se liso} \\ 0, \text{ em caso contrário} \end{cases} & \text{e } Z_{32} &= \begin{cases} 1, \text{ se superfície com alguns sulcos} \\ 0, \text{ em caso contrário} \end{cases} \\
 Z_{41} &= \begin{cases} 1, \text{ se firme} \\ 0, \text{ em caso contrário} \end{cases} & \text{e } Z_{42} &= \begin{cases} 1, \text{ se macio} \\ 0, \text{ em caso contrário} \end{cases} \\
 Z_{51} &= \begin{cases} 1, \text{ se menor ou igual a 25\%} \\ 0, \text{ em caso contrário} \end{cases} & \text{e } Z_{52} &= \begin{cases} 1, \text{ se entre 26 a 50\%} \\ 0, \text{ em caso contrário} \end{cases}
 \end{aligned} \tag{6.2}$$

A Figura 5 mostra um gráfico de barras para examinar a estrutura dos dados, isto é, se estes possuem uma quantidade de valores zeros muito alta.

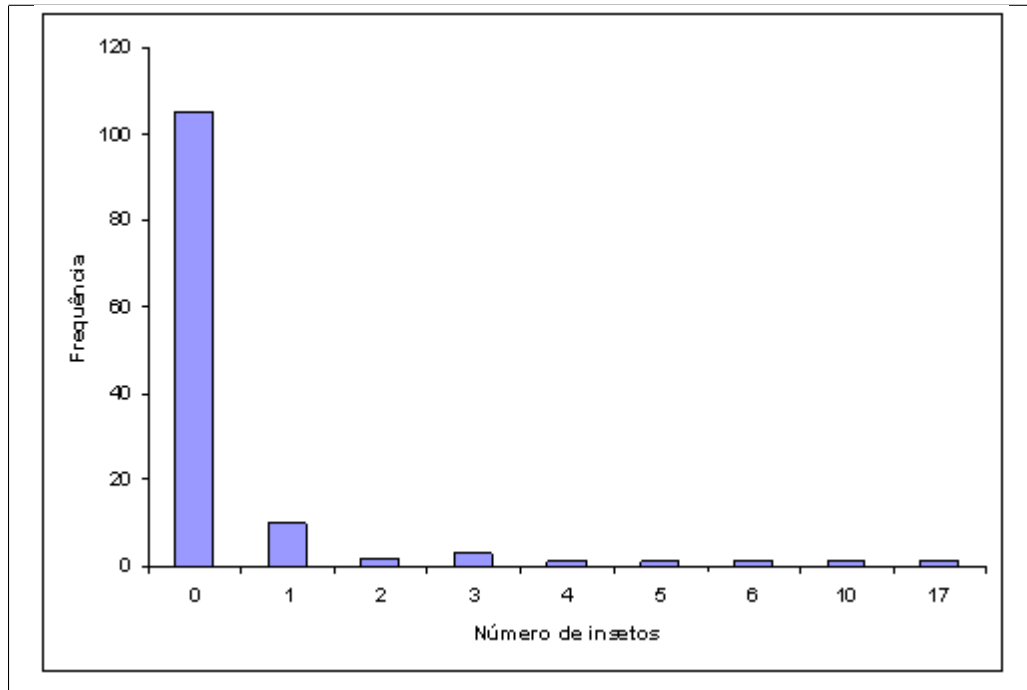


Figura 5: Gráfico de barras para os dados sobre a quantidade de insetos.

Nota-se, através da Figura 5, uma quantidade grande de zeros presentes no conjunto de dados. Assim, a modelagem apresentada no Capítulo 4 será aplicada aqui.

6.2 Ajustes dos modelos propostos no Capítulo 4

Para analisar os dados da Tabela 22, foi considerada a distribuição de Poisson para X_{li} , $i = 1, \dots, 25$ (troncos) e $l = 1, \dots, 5$ (espécies de inseto) dada em (4.1) e os diferentes modelos de regressão introduzidos no Capítulo 4, na presença ou não do vetor de covariáveis $\mathbf{Z}'_i = (Z_{1i}, Z_{2i}, Z_{31i}, Z_{32i}, Z_{41i}, Z_{42i}, Z_{51i}, Z_{52i})$ descrito anteriormente. Associados a cada covariável, tem-se os parâmetros do modelo de regressão $(\beta_{1l}, \beta_{2l}, \beta_{31l}, \beta_{32l}, \beta_{41l}, \beta_{42l}, \beta_{51l}$ e $\beta_{52l})$ e os parâmetros associados à probabilidade de excesso de zeros $(\gamma_{1l}, \gamma_{2l}, \gamma_{31l}, \gamma_{32l}, \gamma_{41l}, \gamma_{42l}, \gamma_{51l}$ e $\gamma_{52l})$.

6.2.1 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Normal para os efeitos aleatórios (Modelo 1)

Considerando-se as distribuições *a priori* em (4.9) com $a_{jl}^2 = g_{jl}^2 = 1000$, $b_l = c_l = 1$, $d = f = 0,01$, $l = 1, \dots, 5$, foram simuladas duas cadeias, cada uma com 400000 amostras das distribuições condicionais para cada parâmetro. Destas amostras, foram descartadas as 100000 primeiras para eliminar os efeitos dos valores iniciais ("*burn-in samples*") e consideradas as iterações $100^a, 200^a, 300^a, \dots$ ("salto" de tamanho 100), para evitar correlações entre os valores gerados. Assim, a amostra final utilizada na obtenção dos sumários *a posteriori* dos parâmetros de interesse foi de tamanho 6000.

A Tabela 23 apresenta as estimativas dos parâmetros e os respectivos intervalos de credibilidade (IC) do modelo sem a inclusão de covariáveis. Da mesma maneira, a Tabela 24 apresenta os resultados para o modelo com a inclusão de covariáveis. A Tabela 25 apresenta os coeficientes de regressão do modelo enquanto que a Tabela 26 apresenta os parâmetros para a probabilidade de excesso de zeros.

Tabela 23: Estimativas *a posteriori* dos parâmetros do modelo sem covariáveis

parâmetros	Média	IC (95%)	
α_1	0,738	0,437	1,130
α_2	0,281	0,116	0,514
α_3	1,203	0,778	1,737
α_4	0,036	0,001	0,137
α_5	0,036	0,001	0,136
σ_w^2	0,377	0,139	0,803

Tabela 24: Estimativas *a posteriori* dos parâmetros do modelo com covariáveis

parâmetros	Média	IC (95%)	
α_1	0,5948	0,0344	2,1810
α_2	0,2544	0,0009	1,4320
α_3	0,0289	0,0001	0,1841
α_4	0,4730	0,0017	2,4000
α_5	0,4744	0,0016	2,4530
σ_w^2	0,0917	0,0056	0,3831

Tabela 25: Estimativas *a posteriori* dos coeficientes de regressão do modelo

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
β_{11}	0,0004	-0,0003	0,0010	β_{411}	0,2908	-1,1000	1,9110
β_{12}	0,0007	-0,0004	0,0020	β_{412}	0,3637	-2,1100	3,6430
β_{13}	-0,0011	-0,0032	0,0003	β_{413}	2,9320	0,8261	6,0920
β_{14}	0,0011	-0,0015	0,0047	β_{414}	-1,9100	-7,5610	2,8680
β_{15}	0,0011	-0,0015	0,0047	β_{415}	-1,9770	-7,4940	3,0040
β_{21}	-1,9020	-4,0970	-0,1267	β_{421}	0,0298	-2,1370	2,1880
β_{22}	-26,7100	-60,0400	-5,2480	β_{422}	10,7500	-0,2886	31,5000
β_{23}	2,0110	0,4889	3,7120	β_{423}	2,5890	0,2828	5,8650
β_{24}	-32,9600	-78,0600	-4,3420	β_{424}	-24,6400	-71,0000	13,2800
β_{25}	-33,1400	-76,0800	-4,2900	β_{425}	-24,8400	-71,7600	12,6900
β_{311}	2,9410	0,8987	5,4540	β_{511}	0,8136	-0,6205	2,5400
β_{312}	18,0000	1,8590	49,6300	β_{512}	0,3654	-2,2790	3,7370
β_{313}	-26,1000	-72,4400	-1,7020	β_{513}	2,7900	1,7220	4,0110
β_{314}	-9,4300	-64,8600	37,8000	β_{514}	-2,0910	-8,1410	3,1100
β_{315}	-9,8910	-65,7700	37,2300	β_{515}	-2,0800	-8,3500	2,9840
β_{321}	0,2683	-2,0620	2,2640	β_{521}	-25,5700	-70,1800	-0,8814
β_{322}	-8,2270	-28,7900	2,8170	β_{522}	-31,3000	-74,2000	-3,8850
β_{323}	-1,6890	-4,9020	0,4211	β_{523}	-25,5500	-71,9700	-1,2750
β_{324}	-15,4500	-67,5600	26,6000	β_{524}	-16,1500	-68,2700	26,4000
β_{325}	-14,4700	-65,1900	27,1700	β_{525}	-15,9700	-65,7700	25,9900

Tabela 26: Estimativas *a posteriori* dos parâmetros relativos ao excesso de zeros.

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
γ_{11}	-0,0007	-0,0020	0,0004	γ_{411}	0,1547	-1,7070	2,1330
γ_{12}	-0,0001	-0,0010	0,0008	γ_{412}	-0,0542	-1,5940	1,4920
γ_{13}	0,0002	-0,0007	0,0011	γ_{413}	-0,4740	-2,0840	1,0770
γ_{14}	-0,0002	-0,0011	0,0007	γ_{414}	-0,0298	-1,5340	1,4410
γ_{15}	-0,0002	-0,0011	0,0007	γ_{415}	-0,0165	-1,4540	1,4700
γ_{21}	0,1923	-1,3360	1,6520	γ_{421}	0,0564	-1,7670	1,8970
γ_{22}	0,3902	-0,8980	1,6650	γ_{422}	-0,4228	-2,1220	1,2590
γ_{23}	-0,3460	-1,6810	0,9393	γ_{423}	0,0273	-1,6410	1,7240
γ_{24}	0,0664	-1,1460	1,2570	γ_{424}	-0,0588	-1,6580	1,5500
γ_{25}	0,0805	-1,1480	1,3050	γ_{425}	-0,0497	-1,6590	1,5850
γ_{311}	-1,5840	-5,0970	1,2040	γ_{511}	-0,5516	-2,4620	1,2550
γ_{312}	-0,5431	-3,0070	1,7370	γ_{512}	-0,3646	-1,8750	1,1580
γ_{313}	0,4095	-1,8500	2,6940	γ_{513}	-0,3957	-1,9960	1,1740
γ_{314}	0,0445	-2,1600	2,2310	γ_{514}	0,0126	-1,3960	1,4670
γ_{315}	0,0144	-2,1830	2,1910	γ_{515}	0,0218	-1,4380	1,4690
γ_{321}	-1,0720	-3,4460	1,1130	γ_{521}	0,2548	-1,7260	2,2650
γ_{322}	-0,4863	-2,5490	1,5620	γ_{522}	0,1958	-1,7670	2,1140
γ_{323}	0,1153	-1,9230	2,0610	γ_{523}	0,1515	-1,7830	2,0980
γ_{324}	0,0039	-1,7780	1,8040	γ_{524}	0,0344	-1,9010	1,9570
γ_{325}	-0,0047	-1,7960	1,7720	γ_{525}	0,0142	-1,8760	1,9010

Dos resultados da Tabela 25, nota-se que a covariável formato é significativa para todos os tipos de insetos: a quantidade de insetos é maior no formato cilíndrico em relação ao formato de paralelepípedo somente para a espécie 2; para as demais espécies, ocorre o processo inverso. Para o tipo de superfície, são encontrados mais insetos das espécies 1 e 2 nas superfícies lisas em relação às outras, enquanto que para a espécie 3 acontece o contrário. Para a covariável porcentagem de superfície coberta, o número de insetos das espécies 1, 2 e 3 é menor nas superfícies de tipo 2 (superfície coberta entre 26 e 50%)

em relação às demais. Dos resultados da Tabela 26, nota-se que nenhuma covariável é significativa para explicar a probabilidade de excesso de zeros.

6.2.2 Comparação dos modelos na presença ou não de covariáveis assumindo uma distribuição Gama para os efeitos aleatórios (Modelo 2)

Considerando-se as distribuições *a priori* em (4.18) com $a_{jl}^2 = g_{jl}^2 = 1000$, $b_l = c_l = d = f = 1$, $l = 1, \dots, 5$, foram simuladas duas cadeias, cada uma com 400000 amostras das distribuições condicionais para cada parâmetro. Para o modelo sem covariáveis, um "burn-in" de 10000 amostras foi considerado e, em seguida, foram geradas 240000 amostras, considerando um "salto" de tamanho 60. Assim, a amostra final utilizada na obtenção dos sumários *a posteriori* dos parâmetros de interesse foi de tamanho 8000. Para o modelo com covariáveis, foram descartadas as 100000 primeiras iterações para eliminar os efeitos dos valores iniciais e a amostra final resultante foi de 2000 amostras (considerando um "salto" de tamanho 1000).

A Tabela 27 apresenta as estimativas dos parâmetros e os respectivos intervalos de credibilidade (IC) do modelo sem a inclusão de covariáveis. Da mesma maneira, a Tabela 28 apresenta os resultados para o modelo com a inclusão de covariáveis. A Tabela 29 apresenta os coeficientes de regressão do modelo enquanto que a Tabela 30 apresenta os parâmetros para a probabilidade de excesso de zeros.

Tabela 27: Estimativas *a posteriori* dos parâmetros do modelo sem covariáveis

parâmetros	Média	IC (95%)	
α_1	3,5510	1,9580	5,8130
α_2	1,4530	0,6349	2,7280
α_3	5,6510	3,3930	8,7570
α_4	0,3252	0,0412	0,9433
α_5	0,3229	0,0379	0,9477
ϕ	0,2740	0,1531	0,4500

Tabela 28: Estimativas *a posteriori* dos parâmetros do modelo com covariáveis

parâmetros	Média	IC (95%)	
α_1	0,6960	0,0353	2,5800
α_2	0,2905	0,0012	1,6020
α_3	0,0233	0,0001	0,1584
α_4	0,4990	0,0017	2,3790
α_5	0,4900	0,0022	2,3410
ϕ	9,6560	0,3838	43,3400

Tabela 29: Estimativas *a posteriori* dos coeficientes de regressão do modelo

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
β_{11}	0,0000	-0,0008	0,0008	β_{411}	-0,1356	-1,5890	1,5410
β_{12}	0,0003	-0,0009	0,0017	β_{412}	-0,0957	-2,5900	3,2050
β_{13}	-0,0014	-0,0034	0,0002	β_{413}	2,8330	0,6071	6,1180
β_{14}	0,0008	-0,0018	0,0045	β_{414}	-2,3720	-7,9470	2,5250
β_{15}	0,0008	-0,0018	0,0045	β_{415}	-2,4210	-8,1810	2,3030
β_{21}	-2,3700	-4,5980	-0,5090	β_{421}	-0,7077	-3,0520	1,6400
β_{22}	-27,5200	-60,9900	-5,5000	β_{422}	10,2300	-1,0880	30,6900
β_{23}	1,8510	0,2179	3,6900	β_{423}	2,1780	-0,3558	5,5730
β_{24}	-32,8600	-75,8100	-4,6210	β_{424}	-24,9900	-74,0600	12,9400
β_{25}	-33,7400	-76,8000	-4,8150	β_{425}	-25,1300	-73,3700	11,8900
β_{311}	3,0650	0,8307	5,6590	β_{511}	0,5407	-0,9244	2,3180
β_{312}	18,1200	1,4350	49,2500	β_{512}	0,0327	-2,6040	3,6430
β_{313}	-26,0400	-71,0400	-1,8280	β_{513}	2,6350	1,3040	3,9830
β_{314}	-9,5530	-64,9300	36,6600	β_{514}	-2,4300	-8,6310	2,7220
β_{315}	-9,5320	-63,3300	39,1100	β_{515}	-2,4860	-8,5830	2,8580
β_{321}	0,3914	-1,8890	2,4920	β_{521}	-25,3800	-70,1300	-1,2170
β_{322}	-8,2510	-28,3200	3,0610	β_{522}	-31,9100	-74,7900	-4,1950
β_{323}	-1,7420	-5,0690	0,4435	β_{523}	-25,5400	-69,8200	-1,4850
β_{324}	-15,5900	-66,5600	25,7800	β_{524}	-16,1700	-68,1600	24,8200
β_{325}	-15,3600	-68,1500	26,5800	β_{525}	-15,3700	-65,7800	24,6500

Tabela 30: Estimativas *a posteriori* dos parâmetros relativos ao excesso de zeros.

Parâmetro	Média	IC (95%)		Parâmetro	Média	IC (95%)	
γ_{11}	-0,0010	-0,0024	0,0002	γ_{411}	-0,1442	-2,0290	1,7700
γ_{12}	-0,0004	-0,0014	0,0005	γ_{412}	-0,3773	-1,9960	1,2300
γ_{13}	0,0000	-0,0010	0,0009	γ_{413}	-0,8050	-2,4510	0,7475
γ_{14}	-0,0004	-0,0014	0,0005	γ_{414}	-0,3356	-1,8790	1,2390
γ_{15}	-0,0004	-0,0015	0,0005	γ_{415}	-0,3431	-1,9280	1,2850
γ_{21}	-0,1792	-1,7050	1,3260	γ_{421}	-0,3553	-2,3040	1,5270
γ_{22}	-0,0044	-1,2750	1,3460	γ_{422}	-0,8157	-2,5280	0,9237
γ_{23}	-0,7311	-2,0840	0,5881	γ_{423}	-0,3901	-2,0950	1,3000
γ_{24}	-0,3206	-1,5990	0,9540	γ_{424}	-0,4497	-2,0920	1,1730
γ_{25}	-0,3018	-1,5700	0,9605	γ_{425}	-0,4747	-2,1280	1,1750
γ_{311}	-1,6790	-5,3330	1,2320	γ_{511}	-0,8518	-2,7850	1,0120
γ_{312}	-0,5832	-3,1450	1,8110	γ_{512}	-0,6650	-2,2600	0,8368
γ_{313}	0,3094	-1,9870	2,7020	γ_{513}	-0,7332	-2,4370	0,8774
γ_{314}	-0,0427	-2,3260	2,1930	γ_{514}	-0,2736	-1,7750	1,1960
γ_{315}	-0,0681	-2,3360	2,2510	γ_{515}	-0,2658	-1,7470	1,2260
γ_{321}	-1,1270	-3,6010	1,1560	γ_{521}	0,0393	-1,9400	2,0890
γ_{322}	-0,6119	-2,8270	1,4270	γ_{522}	-0,0133	-2,0260	1,9730
γ_{323}	0,0707	-2,0820	2,1160	γ_{523}	-0,0567	-2,0500	2,0010
γ_{324}	-0,0859	-1,9890	1,8460	γ_{524}	-0,1622	-2,0840	1,7760
γ_{325}	-0,0820	-2,0520	1,7900	γ_{525}	-0,1863	-2,1050	1,7750

Nota-se que os resultados das Tabelas 29 e 30 são similares aos encontrados para as Tabelas 25 e 26.

6.3 Discriminação de modelos

A Tabela 31 apresenta o DIC para a discriminação do melhor modelo entre os propostos.

Tabela 31: Medidas de discriminação dos modelos propostos

	Modelo	DIC
Normal	Sem Covariáveis	586,832
	Com Covariáveis	472,259
Gama	Sem Covariáveis	526,482
	Com Covariáveis	436,660

Dos resultados da Tabela 31, o modelo com distribuição gama para os efeitos aleatórios com a inclusão de covariáveis é o melhor entre os propostos (menor valor de DIC: 526,482).

Capítulo 7

Conclusões e propostas futuras

A modelagem Bayesiana dos dados de contagem longitudinais, através do modelo de regressão de Poisson, conseguiu identificar alguns resultados que foram vistos, informalmente, na análise descritiva, como as diferenças entre o número médio de enzimas entre machos e fêmeas para o tempo inicial e também a diminuição do número de enzimas conforme o tempo aumenta. Além disso, os efeitos aleatórios conseguiram capturar a superdispersão dos dados, já que as variâncias estimadas foram maiores que as médias observadas.

Através dos resultados da discriminação de modelos (BIC, DIC, LMPL e Variância calculada e observada), o melhor dentre eles foi aquele em que os efeitos aleatórios eram modelados por uma mistura finita de distribuições Normais.

Para os dados com excesso de zeros, os modelos apresentaram estimativas parecidas, sendo que o melhor entre eles foi o de distribuição Gama para os efeitos aleatórios, incluindo covariáveis.

Para um trabalho futuro, seria interessante realizar o cálculo das variâncias e covariâncias de todos os modelos propostos além de um estudo de simulação para verificar robustez das estimativas dos parâmetros, por exemplo. Somente foi utilizado o critério DIC a discriminação dos modelos para excesso de zeros, ficando em aberto o cálculo de outras medidas para isto.

Referências Bibliográficas

- [1] ALBERT, J. A. (1992) Bayesian analysis of a Poisson random effect model for home-run hitters. *The American Statistician*, 16, p.246-253.
- [2] AITKIN, M. (1991) Posterior Bayes Factors. *Journal of the Royal Statistical. Society*, B, v.53, n.1, p.111-142.
- [3] ANGERS, J-F; BISWAS, A. (2003) A Bayesian analysis of zero-inflated generalized Poisson model. *Computational Statistics & Data Analysis*, v.42, p.37-46
- [4] BOX, G. E.; TIAO, G. C. (1992) *Bayesian inference in statistical analysis*. New York: John Wiley, 588p.
- [5] BRADSTREET, T. E.; (1992) “Favorite Data Sets form Early Phases of Drug Research – Part 2.” *Proceedings of the Section on Statistical Education of the American Statistical Association*.
- [6] BRYK, A. S.; RAUNDENBUSH, S. W. (1992) *Hierarchical linear models: applications and data analysis methods*. London. Sage Publications.
- [7] CARLIN, B. P.; LOUIS, T. (2000). *Bayes and empirical Bayes methods for data analysis*, 2nd ed., Chapman and Hall, London.
- [8] CASELLA, G.; GEORGE, E. I. (1992) Explaining the Gibbs Sampler. *The American Statistician*, v.46, n.3, p.167-174.
- [9] CHIB, S.; GREENBERG, E. (1995) Understanding the Metropolis-Hastings algorithm. *The American Statistician*, v.49, 4, p.327-335.
- [10] DeFINETTI, B.; MACHI, A.; SMITH, A. (1974) *Theory of Probability: A Critical Introductory Treatment*. John Wiley & Sons Inc.

-
- [11] DIEBOLT, J; ROBERT, C. P.(1994) Estimation of finite mixture distributions through Bayesian sampling. *Journal of the Royal Statistics, B*, v.56, n.2, p.363-375.
- [12] DONATO, G. G.; CHEN, M. H. (2005). Calibrating Bayes factor under prior predictive distributions,.*Statistica Sinica*, 15, p. 359-380
- [13] DOBSON, A. J. (2001) *An introduction to generalized linear models*. 2^a ed. London: Chapman & Hall.
- [14] GEISSER S; EDDY, W. (1979). A Predictive Approach to Model Selection. *Journal of the American Statistical Association*, v.74, p.153-160.
- [15] GELFAND, A. E.; SMITH, A. F. M. (1990) Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85, p.398-409.
- [16] GELFAND, A. E.; SMITH, A. F. M; LEE, T-M (1992) Bayesian Analysis of Constrained Parameter and Truncated Data Problems Using Gibbs Sampling. *Journal of the American Statistical Association*, v.87, n.418, p. 523-532
- [17] GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, B. D. (2003) *Bayesian Data Analysis*. 2nd ed. New York: Chapman & Hall.
- [18] GELMAN, A.; RUBIN, D. B (1992) Inference from Iterative Simulation Using Multiple Sequences, *Statistical Science*, v.7, n.4, p.457-472.
- [19] GOLDSTEIN, H. (1995) *Multilevel statistical models*. New York: Halsted Press
- [20] GREENE, W. H. (1993) *Econometric analysis*. 2nd edition. Englewood Cliffs: Prentice Hall.
- [21] HALL, D. B. (2000) Zero-Inflated Poisson and Binomial Regression with Random Effects: A Case Study. *Biometrics*, v.56, n.4 ,p.1030–1039.
- [22] JEFFREYS, H.(1961). *Theory of Probability*, 3rd edition, Clarendon Press, Oxford.
- [23] JOWAHEER, V; SUTRADHAR, B. C. (2002) Analysing longitudinal count data with overdispersion. *Biometrika*, 89, 2, p. 389-399.

- [24] KASS, R. E.; RAFTERY, A. E. (1995) Bayes Factor. *Journal of the American Statistical Association*, v.90, p.773-795.
- [25] KOROIVA, R.; FONSECA-GESSNER, A. A. Anexo II- BIGGER, MORE DETE-
RIORATION RATE AND RUGOSE, IS IT BETTER? INFLUENCE OF LOG
CHARACTER IN SAPROXYLIC AQUATIC INSECT COMMUNITY, IN: EFEITO
DO USO DO SOLO E DE DIFERENTES VEGETAIS NA COLONIZAÇÃO DE
TRONCOS POR MACROINVERTEBRADOS EM CÓRREGOS DO ESTADO DE
SAO PAULO. 3º Relatório Científico de Bolsa de Iniciação Científica .Proc. no.
2006/58449-0 FAPESP
- [26] LAIRD, N. M.; WARE, J. H. Random-effects models for longitudinal data. *Biomet-
rics*, v.38, n.4, p.963-974.
- [27] LAMBERT, D. (1992) Zero-Inflated Poisson Regression, with an Application to De-
fects in Manufacturing. *Technometrics*, v.34, n.1, p.1-14.
- [28] LAWLESS, J. F. (1987) Negative Binomial and mixed Poisson regression. *Canadian
Journal of Statistics*, 15, p.209-225.
- [29] LI, Y; BOLT, D. M.; FU, J. (2006) A comparison of alternative models for testlets.
Applied Psychological Measurement, v. 30, n.3.
- [30] LIANG, K. Y.; ZEGER, S. L. (1986) Longitudinal Data Analysis Using Generalized
Linear Models. *Biometrika*, v.73, n.1, p.13-22.
- [31] LINDSEY, J. K. (1997) *Applying generalized linear models*. New York: Springer-
Verlag.
- [32] LONGFORD, N. T. (1993) *Random effect models*. New York: Oxford University
Press.
- [33] LUCEÑO, A. (1995) A family of partially correlated Poisson models for overdispersion. *Computational Statistics & Data Analysis*, 20, p. 511-520.
- [34] McCULLAGH, P.; NELDER, J. A. (1989) *Generalized linear models*. 2ª ed. London:
Chapman & Hall.

- [35] NELDER, J. A.; WEDDERBURN, R. W. M. (1972) Generalized linear models. *Journal of the Royal Statistical Society, A*, v.135, n.3, p.370-384.
- [36] PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. (2003) *Estatística Bayesiana*. Lisboa: Fundação Calouste Gulbenkian.
- [37] R Development Core Team (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- [38] SCHALL, R. (1991) Estimation in generalized linear models with random effects. *Biometrika*, v.78, n.4, p.719-727.
- [39] SINHA, D.; IBRAHIM, J. G.; CHEN, M-H (1999) A New Bayesian Model for Survival Data with a Surviving Fraction Journal, *Journal of the American Statistical Association*, v.94, n.447, p.909-919.
- [40] SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; van der LINDE, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, B*, 64, p.583-639.
- [41] SPIEGELHALTER, D.; THOMAS, A.; BEST, N.; GILKS, W. R. (2003). WinBUGS: Bayesian inference using Gibbs version 1.4. Cambridge, UK: *MRC Biostatistics Unit*, Institute of Public Health.
- [42] STUKEL, T. A. (1993) Comparison of methods for the analysis of longitudinal interval count data. *Statistics in Medicine*, 12, p.1339-1351.
- [43] TANNER, M.; WONG, W. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, 82(398):528-540.
- [44] THALL, P. F. (1988) Mixed Poisson likelihood regression models for longitudinal interval data count. *Biometrics*, 44, p.197-209.
- [45] TUNARU, R. (2002) Hierarchical Bayesian Models for Multiple Count Data. *Austrian Journal of Statistics*, 31, n. 2&3, p. 221-229.

Apêndice A

Código em Winbugs para o Modelo 1 do Capítulo 3

```
#Modelo 1 (l=1,2,3,4 tempos e i=1,...,32 cachorros,  
#burn-in de 30000 (gerar 300 com salto de 100), 3000 salto 100)  
model {  
  for (l in 1:4) {  
    for (i in 1:32) {  
      y[l,i] ~ dpois(lambda[l,i])  
      lambda[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i] + w[i])  
      logvero[l,i] <- -lambda[l,i] + y[l,i]*log(lambda[l,i]) - logfact(y[l,i])  
      fy[l,i] <- exp(logvero[l,i])  
      CPO[l,i] <- 1/fy[l,i]  
    }  
  }  
  
  BIC <- -2*sum(logvero[,i]) - 13*log(32)  
  for (i in 1:32) {  
    w[i] ~ dnorm(0, tau.w)  
  }  
  for (l in 1:4) {  
    beta1[l] ~ dnorm(0, 1.0E-6)
```



```

beta2[l] ~ dnorm(0, 1.0E-6)
lamb[l] ~ dgamma(0.001, 0.001)
}
#lamb[1] ~ dgamma(21.46225, 0.1465)
#lamb[2] ~ dgamma(21.56309, 0.146844)
#lamb[3] ~ dgamma(17.53969, 0.132438)
#lamb[4] ~ dgamma(15.79735, 0.125688)
tau.w ~ dgamma(0.001, 0.001)
sigma2.w<-1/tau.w
##### Médias e variâncias não-condicionais #####
for (l in 1:4) {
  for (i in 1:32) {
    media[l,i] <- lamb[l]*exp(1/(2*tau.w))*exp(beta1[l]*z1[i] + beta2[l]*z2[i])
    var[l,i] <- lamb[l]*exp(1/(2*tau.w))*exp(beta1[l]*z1[i] + beta2[l]*z2[i])
+ pow(lamb[l],2)*exp(1/tau.w)*
(exp(1/tau.w)-1)*exp(2*(beta1[l]*z1[i] + beta2[l]*z2[i]))
  }
}
##### Médias #####
### theta[1, 0, 1]:
# 1 é o tempo (1,2,3,4)
# 0 é o sexo (1 macho, 0 femea)
# 1 é a dose (1:0, 2:8, 3:25, 4:75)
##### Média tempo 1 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta10[1]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*0)
# Dose 8
theta10[2]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*8)
# Dose 25
theta10[3]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*25)
# Dose 75

```

```

theta10[4]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta11[1]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*0)
# Dose 8
theta11[2]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*8)
# Dose 25
theta11[3]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*25)
# Dose 75
theta11[4]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*75)
##### Média tempo 2 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta20[1]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*0)
# Dose 8
theta20[2]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*8)
# Dose 25
theta20[3]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*25)
# Dose 75
theta20[4]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta21[1]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*0)
# Dose 8
theta21[2]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*8)
# Dose 25
theta21[3]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*25)
# Dose 75
theta21[4]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*75)
##### Média tempo 3 #####
##### covariável Z1=0 (feminino)
# Dose 0

```

```

theta30[1]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*0)
# Dose 8
theta30[2]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*8)
# Dose 25
theta30[3]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*25)
# Dose 75
theta30[4]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta31[1]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*0)
# Dose 8
theta31[2]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*8)
# Dose 25
theta31[3]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*25)
# Dose 75
theta31[4]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*75)
##### Média tempo 4 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta40[1]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*0)
# Dose 8
theta40[2]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*8)
# Dose 25
theta40[3]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*25)
# Dose 75
theta40[4]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta41[1]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*0)
# Dose 8
theta41[2]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*8)
# Dose 25

```

```

theta41[3]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*25)
# Dose 75
theta41[4]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*75)
##### Comparações #####
### delta 1 0 1 2:
# 1 é o tempo (1,2,3,4)
# 0 é o sexo (1 macho, 0 femea)
# comparação entre doses , p. ex, 1 com 2
##### DENTRO DE CADA TEMPO ##
#####
##### Dentro do tempo 1 #####
#####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta1012<-theta10[1] - theta10[2]
delta1013<-theta10[1] - theta10[3]
delta1014<-theta10[1] - theta10[4]
delta1023<-theta10[2] - theta10[3]
delta1024<-theta10[2] - theta10[4]
delta1034<-theta10[3] - theta10[4]
### Masculino #####
delta1112<-theta11[1] - theta11[2]
delta1113<-theta11[1] - theta11[3]
delta1114<-theta11[1] - theta11[4]
delta1123<-theta11[2] - theta11[3]
delta1124<-theta11[2] - theta11[4]
delta1134<-theta11[3] - theta11[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta1mf1<-theta11[1] - theta10[1]
delta1mf2<-theta11[2] - theta10[2]
delta1mf3<-theta11[3] - theta10[3]
delta1mf4<-theta11[4] - theta10[4]

```

```

#####
##### Dentro do tempo 2 #####
#####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta2012<-theta20[1] - theta20[2]
delta2013<-theta20[1] - theta20[3]
delta2014<-theta20[1] - theta20[4]
delta2023<-theta20[2] - theta20[3]
delta2024<-theta20[2] - theta20[4]
delta2034<-theta20[3] - theta20[4]
### Masculino #####
delta2112<-theta21[1] - theta21[2]
delta2113<-theta21[1] - theta21[3]
delta2114<-theta21[1] - theta21[4]
delta2123<-theta21[2] - theta21[3]
delta2124<-theta21[2] - theta21[4]
delta2134<-theta21[3] - theta21[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta2mf1<-theta21[1] - theta20[1]
delta2mf2<-theta21[2] - theta20[2]
delta2mf3<-theta21[3] - theta20[3]
delta2mf4<-theta21[4] - theta20[4]
#####
##### Dentro do tempo 3 #####
#####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta3012<-theta30[1] - theta30[2]
delta3013<-theta30[1] - theta30[3]
delta3014<-theta30[1] - theta30[4]
delta3023<-theta30[2] - theta30[3]

```

```

delta3024<-theta30[2] - theta30[4]
delta3034<-theta30[3] - theta30[4]
### Masculino #####
delta3112<-theta31[1] - theta31[2]
delta3113<-theta31[1] - theta31[3]
delta3114<-theta31[1] - theta31[4]
delta3123<-theta31[2] - theta31[3]
delta3124<-theta31[2] - theta31[4]
delta3134<-theta31[3] - theta31[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta3mf1<-theta31[1] - theta30[1]
delta3mf2<-theta31[2] - theta30[2]
delta3mf3<-theta31[3] - theta30[3]
delta3mf4<-theta31[4] - theta30[4]
#####
##### Dentro do tempo 4 #####
#####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta4012<-theta40[1] - theta40[2]
delta4013<-theta40[1] - theta40[3]
delta4014<-theta40[1] - theta40[4]
delta4023<-theta40[2] - theta40[3]
delta4024<-theta40[2] - theta40[4]
delta4034<-theta40[3] - theta40[4]
### Masculino #####
delta4112<-theta41[1] - theta41[2]
delta4113<-theta41[1] - theta41[3]
delta4114<-theta41[1] - theta41[4]
delta4123<-theta41[2] - theta41[3]
delta4124<-theta41[2] - theta41[4]
delta4134<-theta41[3] - theta41[4]

```

```

### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta4mf1<-theta41[1] - theta40[1]
delta4mf2<-theta41[2] - theta40[2]
delta4mf3<-theta41[3] - theta40[3]
delta4mf4<-theta41[4] - theta40[4]

##### ENTRE TEMPOS #####
##### Dose 1 #####
#### sexo feminino
eta0112<-theta10[1] - theta20[1]
eta0113<-theta10[1] - theta30[1]
eta0114<-theta10[1] - theta40[1]
eta0123<-theta20[1] - theta30[1]
eta0124<-theta20[1] - theta40[1]
eta0134<-theta30[1] - theta40[1]
#### sexo masculino
eta1112<-theta11[1] - theta21[1]
eta1113<-theta11[1] - theta31[1]
eta1114<-theta11[1] - theta41[1]
eta1123<-theta21[1] - theta31[1]
eta1124<-theta21[1] - theta41[1]
eta1134<-theta31[1] - theta41[1]

##### Dose 2 #####
#### sexo feminino
eta0212<-theta10[2] - theta20[2]
eta0213<-theta10[2] - theta30[2]
eta0214<-theta10[2] - theta40[2]
eta0223<-theta20[2] - theta30[2]
eta0224<-theta20[2] - theta40[2]
eta0234<-theta30[2] - theta40[2]
#### sexo masculino
eta1212<-theta11[2] - theta21[2]
eta1213<-theta11[2] - theta31[2]

```

```

eta1214<-theta11[2] - theta41[2]
eta1223<-theta21[2] - theta31[2]
eta1224<-theta21[2] - theta41[2]
eta1234<-theta31[2] - theta41[2]
##### Dose 3 #####
##### sexo feminino
eta0312<-theta10[3] - theta20[3]
eta0313<-theta10[3] - theta30[3]
eta0314<-theta10[3] - theta40[3]
eta0323<-theta20[3] - theta30[3]
eta0324<-theta20[3] - theta40[3]
eta0334<-theta30[3] - theta40[3]
##### sexo masculino
eta1312<-theta11[3] - theta21[3]
eta1313<-theta11[3] - theta31[3]
eta1314<-theta11[3] - theta41[3]
eta1323<-theta21[3] - theta31[3]
eta1324<-theta21[3] - theta41[3]
eta1334<-theta31[3] - theta41[3]
##### Dose 4 #####
##### sexo feminino
eta0412<-theta10[4] - theta20[4]
eta0413<-theta10[4] - theta30[4]
eta0414<-theta10[4] - theta40[4]
eta0423<-theta20[4] - theta30[4]
eta0424<-theta20[4] - theta40[4]
eta0434<-theta30[4] - theta40[4]
##### sexo masculino
eta1412<-theta11[4] - theta21[4]
eta1413<-theta11[4] - theta31[4]
eta1414<-theta11[4] - theta41[4]
eta1423<-theta21[4] - theta31[4]

```



```

eta1424<-theta21[4] - theta41[4]
eta1434<-theta31[4] - theta41[4]
}
#Dados
y[,1]  y[,2]  y[,3]  y[,4]  y[,5]  y[,6]  y[,7]  y[,8]  y[,9]  y[,10]
y[,11] y[,12] y[,13] y[,14] y[,15] y[,16] y[,17] y[,18] y[,19]
y[,20] y[,21] y[,22] y[,23] y[,24] y[,25] y[,26] y[,27] y[,28]
y[,29] y[,30] y[,31] y[,32]
169  291  158  122  203  178  141  181  101  199  141
149  135  153  147  157  125  138  113  137  170  139
131  125  113  150  155  133  113  102  128  91  136
282  128  113  171  154  104  143  80  149  138  131
218  164  192  168  105  94  91  94  150  127  152
105  101  113  161  197  148  191  226  173  125  218
104  91  141  127  90  120  78  147  118  133  204
148  205  168  92  81  81  79  125  103  112  91  89
105  134  218  122  172  261  156  122  154  91  82
109  114  80  125  83  142  108  134  226  146  167
165  86  72  72  75  121  94  99  82  122  90  116
159  143  162  279  202
END
z1[]  z2[]
0  1
0  1
0  1
0  1
8  1
8  1
8  1
8  1
25  1
25  1

```


)

Apêndice B

Código em Winbugs para o Modelo 2 do Capítulo 3

```
# Modelo 2 (l=1,2,3 tempos, j=1,2,3,4 tratamentos, i=1,...,40 pacientes:
# rodar com (1000 mil) burn-in (100 com salto de 1000) e
# 30000 com salto de (1000))
model {
  for (i in 1:32) {
    w[i] ~ dnorm(mu[T[i]], tau[T[i]])
    T[i] ~ dcat(P[])
  }
  for (l in 1:4) {
    for (i in 1:32) {
      y[l,i] ~ dpois(lambda[l,i])
      lambda[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i] + w[i])
      logvero[l,i] <- -lambda[l,i] + y[l,i]*log(lambda[l,i]) - logfact(y[l,i])
      fy[l,i] <- exp(logvero[l,i])
      CPO[l,i] <- 1/fy[l,i]
    }
  }

  BIC <- -2*sum(logvero[,]) - 17*log(32)
```

```

# prioris
#
mu[2] ~ dnorm(-0.5, 30)
mu[1] ~ dnorm(0.5, 30)

P[1:2] ~ ddirch(gamma[])
tau[1] ~ dgamma(10,10)
tau[2] ~ dgamma(10,10)

sigma2[1] <- 1/tau[1]
sigma2[2] <- 1/tau[2]

for (l in 1:4) {
beta1[l] ~ dnorm(0, 1.0E-6)
beta2[l] ~ dnorm(0, 1.0E-6)
lamb[l] ~ dgamma(0.001, 0.001)
}

#### Médias e variâncias não-condicionais ####
muM <- P[1]*mu[1] + P[2]*mu[2]
sig2M <- P[1]*(pow(mu[1],2)+1/tau[1]) + P[2]*
(pow(mu[2],2)+1/tau[2]) - pow(muM,2)
for (i in 1:32) {
for (l in 1:4) {
E[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i])*exp(muM)
Var[l,i] <- E[l,i] + pow(lamb[l],2)*
exp(2*(beta1[l]*z1[i] + beta2[l]*z2[i]))*exp(2*muM)*sig2M
}
}
}

#### Comparações ####

```

```

### beta 1
difb112<-beta1[1]-beta1[2]
difb113<-beta1[1]-beta1[3]
difb114<-beta1[1]-beta1[4]
difb123<-beta1[2]-beta1[3]
difb124<-beta1[2]-beta1[4]
difb134<-beta1[3]-beta1[4]
### beta 2
difb212<-beta2[1]-beta2[2]
difb213<-beta2[1]-beta2[3]
difb214<-beta2[1]-beta2[4]
difb223<-beta2[2]-beta2[3]
difb224<-beta2[2]-beta2[4]
difb234<-beta2[3]-beta2[4]
##### Médias #####
### theta[1, 0, 1]:
# 1 é o tempo (1,2,3,4)
# 0 é o sexo (1 macho, 0 femea)
# 1 é a dose (1:0, 2:8, 3:25, 4:75)
##### Média tempo 1 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta10[1]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*0)
# Dose 8
theta10[2]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*8)
# Dose 25
theta10[3]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*25)
# Dose 75
theta10[4]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta11[1]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*0)

```

```

# Dose 8
theta11[2]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*8)
# Dose 25
theta11[3]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*25)
# Dose 75
theta11[4]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*75)
##### Média tempo 2 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta20[1]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*0)
# Dose 8
theta20[2]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*8)
# Dose 25
theta20[3]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*25)
# Dose 75
theta20[4]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta21[1]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*0)
# Dose 8
theta21[2]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*8)
# Dose 25
theta21[3]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*25)
# Dose 75
theta21[4]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*75)
##### Média tempo 3 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta30[1]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*0)
# Dose 8
theta30[2]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*8)
# Dose 25

```

```

theta30[3]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*25)
# Dose 75
theta30[4]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta31[1]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*0)
# Dose 8
theta31[2]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*8)
# Dose 25
theta31[3]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*25)
# Dose 75
theta31[4]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*75)
##### Média tempo 4 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta40[1]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*0)
# Dose 8
theta40[2]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*8)
# Dose 25
theta40[3]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*25)
# Dose 75
theta40[4]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta41[1]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*0)
# Dose 8
theta41[2]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*8)
# Dose 25
theta41[3]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*25)
# Dose 75
theta41[4]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*75)
##### Comparações #####

```



```

### delta 1 0 1 2:
# 1 é o tempo (1,2,3,4)
# 0 é o sexo (1 macho, 0 femea)
# comparação entre doses , p. ex, 1 com 2
##### DENTRO DE CADA TEMPO ###
##### Dentro do tempo 1 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta1012<-theta10[1] - theta10[2]
delta1013<-theta10[1] - theta10[3]
delta1014<-theta10[1] - theta10[4]
delta1023<-theta10[2] - theta10[3]
delta1024<-theta10[2] - theta10[4]
delta1034<-theta10[3] - theta10[4]
### Masculino #####
delta1112<-theta11[1] - theta11[2]
delta1113<-theta11[1] - theta11[3]
delta1114<-theta11[1] - theta11[4]
delta1123<-theta11[2] - theta11[3]
delta1124<-theta11[2] - theta11[4]
delta1134<-theta11[3] - theta11[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1##
delta1mf1<-theta11[1] - theta10[1]
delta1mf2<-theta11[2] - theta10[2]
delta1mf3<-theta11[3] - theta10[3]
delta1mf4<-theta11[4] - theta10[4]
##### Dentro do tempo 2 #####
#####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta2012<-theta20[1] - theta20[2]
delta2013<-theta20[1] - theta20[3]

```

```

delta2014<-theta20[1] - theta20[4]
delta2023<-theta20[2] - theta20[3]
delta2024<-theta20[2] - theta20[4]
delta2034<-theta20[3] - theta20[4]
### Masculino #####
delta2112<-theta21[1] - theta21[2]
delta2113<-theta21[1] - theta21[3]
delta2114<-theta21[1] - theta21[4]
delta2123<-theta21[2] - theta21[3]
delta2124<-theta21[2] - theta21[4]
delta2134<-theta21[3] - theta21[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1###
delta2mf1<-theta21[1] - theta20[1]
delta2mf2<-theta21[2] - theta20[2]
delta2mf3<-theta21[3] - theta20[3]
delta2mf4<-theta21[4] - theta20[4]
#####
##### Dentro do tempo 3 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta3012<-theta30[1] - theta30[2]
delta3013<-theta30[1] - theta30[3]
delta3014<-theta30[1] - theta30[4]
delta3023<-theta30[2] - theta30[3]
delta3024<-theta30[2] - theta30[4]
delta3034<-theta30[3] - theta30[4]
### Masculino #####
delta3112<-theta31[1] - theta31[2]
delta3113<-theta31[1] - theta31[3]
delta3114<-theta31[1] - theta31[4]
delta3123<-theta31[2] - theta31[3]
delta3124<-theta31[2] - theta31[4]

```

```

delta3134<-theta31[3] - theta31[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1###
delta3mf1<-theta31[1] - theta30[1]
delta3mf2<-theta31[2] - theta30[2]
delta3mf3<-theta31[3] - theta30[3]
delta3mf4<-theta31[4] - theta30[4]
#####
##### Dentro do tempo 4 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta4012<-theta40[1] - theta40[2]
delta4013<-theta40[1] - theta40[3]
delta4014<-theta40[1] - theta40[4]
delta4023<-theta40[2] - theta40[3]
delta4024<-theta40[2] - theta40[4]
delta4034<-theta40[3] - theta40[4]
### Masculino #####
delta4112<-theta41[1] - theta41[2]
delta4113<-theta41[1] - theta41[3]
delta4114<-theta41[1] - theta41[4]
delta4123<-theta41[2] - theta41[3]
delta4124<-theta41[2] - theta41[4]
delta4134<-theta41[3] - theta41[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1###
delta4mf1<-theta41[1] - theta40[1]
delta4mf2<-theta41[2] - theta40[2]
delta4mf3<-theta41[3] - theta40[3]
delta4mf4<-theta41[4] - theta40[4]
##### ENTRE TEMPOS #####
##### Dose 1 #####
##### sexo feminino
eta0112<-theta10[1] - theta20[1]

```

```

eta0113<-theta10[1] - theta30[1]
eta0114<-theta10[1] - theta40[1]
eta0123<-theta20[1] - theta30[1]
eta0124<-theta20[1] - theta40[1]
eta0134<-theta30[1] - theta40[1]
#### sexo masculino
eta1112<-theta11[1] - theta21[1]
eta1113<-theta11[1] - theta31[1]
eta1114<-theta11[1] - theta41[1]
eta1123<-theta21[1] - theta31[1]
eta1124<-theta21[1] - theta41[1]
eta1134<-theta31[1] - theta41[1]
##### Dose 2 #####
#### sexo feminino
eta0212<-theta10[2] - theta20[2]
eta0213<-theta10[2] - theta30[2]
eta0214<-theta10[2] - theta40[2]
eta0223<-theta20[2] - theta30[2]
eta0224<-theta20[2] - theta40[2]
eta0234<-theta30[2] - theta40[2]
#### sexo masculino
eta1212<-theta11[2] - theta21[2]
eta1213<-theta11[2] - theta31[2]
eta1214<-theta11[2] - theta41[2]
eta1223<-theta21[2] - theta31[2]
eta1224<-theta21[2] - theta41[2]
eta1234<-theta31[2] - theta41[2]
##### Dose 3 #####
#### sexo feminino
eta0312<-theta10[3] - theta20[3]
eta0313<-theta10[3] - theta30[3]
eta0314<-theta10[3] - theta40[3]

```

```

eta0323<-theta20[3] - theta30[3]
eta0324<-theta20[3] - theta40[3]
eta0334<-theta30[3] - theta40[3]
##### sexo masculino
eta1312<-theta11[3] - theta21[3]
eta1313<-theta11[3] - theta31[3]
eta1314<-theta11[3] - theta41[3]
eta1323<-theta21[3] - theta31[3]
eta1324<-theta21[3] - theta41[3]
eta1334<-theta31[3] - theta41[3]
##### Dose 4 #####
##### sexo feminino
eta0412<-theta10[4] - theta20[4]
eta0413<-theta10[4] - theta30[4]
eta0414<-theta10[4] - theta40[4]
eta0423<-theta20[4] - theta30[4]
eta0424<-theta20[4] - theta40[4]
eta0434<-theta30[4] - theta40[4]
##### sexo masculino
eta1412<-theta11[4] - theta21[4]
eta1413<-theta11[4] - theta31[4]
eta1414<-theta11[4] - theta41[4]
eta1423<-theta21[4] - theta31[4]
eta1424<-theta21[4] - theta41[4]
eta1434<-theta31[4] - theta41[4]

}
#Dados
y[,1]   y[,2]   y[,3]   y[,4]   y[,5]   y[,6]   y[,7]   y[,8]   y[,9]   y[,10]
y[,11]  y[,12]  y[,13]  y[,14]  y[,15]  y[,16]  y[,17]  y[,18]  y[,19]
y[,20]  y[,21]  y[,22]  y[,23]  y[,24]  y[,25]  y[,26]  y[,27]  y[,28]
y[,29]  y[,30]  y[,31]  y[,32]

```

```
169 291 158 122 203 178 141 181 101 199 141
149 135 153 147 157 125 138 113 137 170 139
131 125 113 150 155 133 113 102 128 91 136
282 128 113 171 154 104 143 80 149 138 131
218 164 192 168 105 94 91 94 150 127 152
105 101 113 161 197 148 191 226 173 125 218
104 91 141 127 90 120 78 147 118 133 204
148 205 168 92 81 81 79 125 103 112 91 89
105 134 218 122 172 261 156 122 154 91 82
109 114 80 125 83 142 108 134 226 146 167
165 86 72 72 75 121 94 99 82 122 90 116
159 143 162 279 202
END
z1[] z2[]
0 1
0 1
0 1
0 1
8 1
8 1
8 1
8 1
25 1
25 1
25 1
25 1
75 1
75 1
75 1
75 1
0 0
0 0
```


Apêndice C

Código em Winbugs para o Modelo 3 do Capítulo 3

```
#Modelo 3 (l=1,2,3,4 tempos e i=1,...,32 cachorros,  
  #burn-in de 30000 (300 com salto de 100),  
  # 3000 com salto de 100, refresh 10)  
model {  
  for (l in 1:4) {  
    for (i in 1:32) {  
      y[l,i] ~ dpois(lambda[l,i])  
      lambda[l,i] <- w[i]*lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i])  
      logvero[l,i] <- lambda[l,i] + y[l,i]*log(lambda[l,i]) - logfact(y[l,i])  
      fy[l,i] <- exp(logvero[l,i])  
      CPO[l,i] <- 1/fy[l,i]  
    }  
  }  
  
  BIC <- -2*sum(logvero[,]) - 13*log(32)  
  eta <- 1/phi  
  for (i in 1:32) {  
    w[i] ~ dgamma(eta, eta)  
  }  
  
  for (l in 1:4) {
```



```

beta1[l] ~ dnorm(0, 1.0E-6)
beta2[l] ~ dnorm(0, 1.0E-6)
lamb[l] ~ dgamma(0.001, 0.001)
      }
phi ~ dgamma(0.001, 0.001)
#### Médias e variâncias não-condicionais #####
for (l in 1:4) {
  for (i in 1:32) {
    media[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i])
    var[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i]) +
pow(lamb[l],2)*phi*exp(2*(beta1[l]*z1[i] + beta2[l]*z2[i]))
      }
    }

### Comparações ###
difbeta112 <- beta1[1] - beta1[2]
difbeta113 <- beta1[1] - beta1[3]
difbeta114 <- beta1[1] - beta1[4]
difbeta123 <- beta1[2] - beta1[3]
difbeta124 <- beta1[2] - beta1[4]
difbeta134 <- beta1[3] - beta1[4]

difbeta212 <- beta2[1] - beta2[2]
difbeta213 <- beta2[1] - beta2[3]
difbeta214 <- beta2[1] - beta2[4]
difbeta223 <- beta2[2] - beta2[3]
difbeta224 <- beta2[2] - beta2[4]
difbeta234 <- beta2[3] - beta2[4]

##### Médias #####
### theta[1, 0, 1]:
# 1 é o tempo (1,2,3,4)
# 0 é o sexo (1 macho, 0 femea)
# 1 é a dose (1:0, 2:8, 3:25, 4:75)

```

```
##### Média tempo 1 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta10[1]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*0)
# Dose 8
theta10[2]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*8)
# Dose 25
theta10[3]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*25)
# Dose 75
theta10[4]<-lamb[1]*exp(beta2[1]*0 + beta1[1]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta11[1]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*0)
# Dose 8
theta11[2]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*8)
# Dose 25
theta11[3]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*25)
# Dose 75
theta11[4]<-lamb[1]*exp(beta2[1]*1 + beta1[1]*75)
##### Média tempo 2 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta20[1]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*0)
# Dose 8
theta20[2]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*8)
# Dose 25
theta20[3]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*25)
# Dose 75
theta20[4]<-lamb[2]*exp(beta2[2]*0 + beta1[2]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta21[1]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*0)
```

```
# Dose 8
theta21[2]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*8)
# Dose 25
theta21[3]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*25)
# Dose 75
theta21[4]<-lamb[2]*exp(beta2[2]*1 + beta1[2]*75)
##### Média tempo 3 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta30[1]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*0)
# Dose 8
theta30[2]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*8)
# Dose 25
theta30[3]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*25)
# Dose 75
theta30[4]<-lamb[3]*exp(beta2[3]*0 + beta1[3]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta31[1]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*0)
# Dose 8
theta31[2]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*8)
# Dose 25
theta31[3]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*25)
# Dose 75
theta31[4]<-lamb[3]*exp(beta2[3]*1 + beta1[3]*75)
##### Média tempo 4 #####
##### covariável Z1=0 (feminino)
# Dose 0
theta40[1]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*0)
# Dose 8
theta40[2]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*8)
# Dose 25
```

```

theta40[3]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*25)
# Dose 75
theta40[4]<-lamb[4]*exp(beta2[4]*0 + beta1[4]*75)
##### covariável Z1=1 (masculino)
# Dose 0
theta41[1]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*0)
# Dose 8
theta41[2]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*8)
# Dose 25
theta41[3]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*25)
# Dose 75
theta41[4]<-lamb[4]*exp(beta2[4]*1 + beta1[4]*75)
##### Comparações #####
### delta 1 0 1 2:
# 1 é o tempo (1,2,3,4)
# 0 é o sexo (1 macho, 0 femea)
# comparação entre doses , p. ex, 1 com 2
##### DENTRO DE CADA TEMPO ###
##### Dentro do tempo 1 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta1012<-theta10[1] - theta10[2]
delta1013<-theta10[1] - theta10[3]
delta1014<-theta10[1] - theta10[4]
delta1023<-theta10[2] - theta10[3]
delta1024<-theta10[2] - theta10[4]
delta1034<-theta10[3] - theta10[4]
### Masculino #####
delta1112<-theta11[1] - theta11[2]
delta1113<-theta11[1] - theta11[3]
delta1114<-theta11[1] - theta11[4]
delta1123<-theta11[2] - theta11[3]

```

```

delta1124<-theta11[2] - theta11[4]
delta1134<-theta11[3] - theta11[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta1mf1<-theta11[1] - theta10[1]
delta1mf2<-theta11[2] - theta10[2]
delta1mf3<-theta11[3] - theta10[3]
delta1mf4<-theta11[4] - theta10[4]
##### Dentro do tempo 2 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta2012<-theta20[1] - theta20[2]
delta2013<-theta20[1] - theta20[3]
delta2014<-theta20[1] - theta20[4]
delta2023<-theta20[2] - theta20[3]
delta2024<-theta20[2] - theta20[4]
delta2034<-theta20[3] - theta20[4]
### Masculino #####
delta2112<-theta21[1] - theta21[2]
delta2113<-theta21[1] - theta21[3]
delta2114<-theta21[1] - theta21[4]
delta2123<-theta21[2] - theta21[3]
delta2124<-theta21[2] - theta21[4]
delta2134<-theta21[3] - theta21[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta2mf1<-theta21[1] - theta20[1]
delta2mf2<-theta21[2] - theta20[2]
delta2mf3<-theta21[3] - theta20[3]
delta2mf4<-theta21[4] - theta20[4]
##### Dentro do tempo 3 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta3012<-theta30[1] - theta30[2]

```

```

delta3013<-theta30[1] - theta30[3]
delta3014<-theta30[1] - theta30[4]
delta3023<-theta30[2] - theta30[3]
delta3024<-theta30[2] - theta30[4]
delta3034<-theta30[3] - theta30[4]
### Masculino #####
delta3112<-theta31[1] - theta31[2]
delta3113<-theta31[1] - theta31[3]
delta3114<-theta31[1] - theta31[4]
delta3123<-theta31[2] - theta31[3]
delta3124<-theta31[2] - theta31[4]
delta3134<-theta31[3] - theta31[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta3mf1<-theta31[1] - theta30[1]
delta3mf2<-theta31[2] - theta30[2]
delta3mf3<-theta31[3] - theta30[3]
delta3mf4<-theta31[4] - theta30[4]
##### Dentro do tempo 4 #####
## Dentro de cada sexo, comparando doses #####
## Feminino#####
delta4012<-theta40[1] - theta40[2]
delta4013<-theta40[1] - theta40[3]
delta4014<-theta40[1] - theta40[4]
delta4023<-theta40[2] - theta40[3]
delta4024<-theta40[2] - theta40[4]
delta4034<-theta40[3] - theta40[4]
### Masculino #####
delta4112<-theta41[1] - theta41[2]
delta4113<-theta41[1] - theta41[3]
delta4114<-theta41[1] - theta41[4]
delta4123<-theta41[2] - theta41[3]
delta4124<-theta41[2] - theta41[4]

```

```

delta4134<-theta41[3] - theta41[4]
### Entre sexo variando doses (1,2,3,4) no final do delta fm 1####
delta4mf1<-theta41[1] - theta40[1]
delta4mf2<-theta41[2] - theta40[2]
delta4mf3<-theta41[3] - theta40[3]
delta4mf4<-theta41[4] - theta40[4]
##### ENTRE TEMPOS #####
##### Dose 1 #####
#### sexo feminino
eta0112<-theta10[1] - theta20[1]
eta0113<-theta10[1] - theta30[1]
eta0114<-theta10[1] - theta40[1]
eta0123<-theta20[1] - theta30[1]
eta0124<-theta20[1] - theta40[1]
eta0134<-theta30[1] - theta40[1]
#### sexo masculino
eta1112<-theta11[1] - theta21[1]
eta1113<-theta11[1] - theta31[1]
eta1114<-theta11[1] - theta41[1]
eta1123<-theta21[1] - theta31[1]
eta1124<-theta21[1] - theta41[1]
eta1134<-theta31[1] - theta41[1]
##### Dose 2 #####
#### sexo feminino
eta0212<-theta10[2] - theta20[2]
eta0213<-theta10[2] - theta30[2]
eta0214<-theta10[2] - theta40[2]
eta0223<-theta20[2] - theta30[2]
eta0224<-theta20[2] - theta40[2]
eta0234<-theta30[2] - theta40[2]
#### sexo masculino
eta1212<-theta11[2] - theta21[2]

```

```

eta1213<-theta11[2] - theta31[2]
eta1214<-theta11[2] - theta41[2]
eta1223<-theta21[2] - theta31[2]
eta1224<-theta21[2] - theta41[2]
eta1234<-theta31[2] - theta41[2]
##### Dose 3 #####
##### sexo feminino
eta0312<-theta10[3] - theta20[3]
eta0313<-theta10[3] - theta30[3]
eta0314<-theta10[3] - theta40[3]
eta0323<-theta20[3] - theta30[3]
eta0324<-theta20[3] - theta40[3]
eta0334<-theta30[3] - theta40[3]
##### sexo masculino
eta1312<-theta11[3] - theta21[3]
eta1313<-theta11[3] - theta31[3]
eta1314<-theta11[3] - theta41[3]
eta1323<-theta21[3] - theta31[3]
eta1324<-theta21[3] - theta41[3]
eta1334<-theta31[3] - theta41[3]
##### Dose 4 #####
##### sexo feminino
eta0412<-theta10[4] - theta20[4]
eta0413<-theta10[4] - theta30[4]
eta0414<-theta10[4] - theta40[4]
eta0423<-theta20[4] - theta30[4]
eta0424<-theta20[4] - theta40[4]
eta0434<-theta30[4] - theta40[4]
##### sexo masculino
eta1412<-theta11[4] - theta21[4]
eta1413<-theta11[4] - theta31[4]
eta1414<-theta11[4] - theta41[4]

```



```

eta1423<-theta21[4] - theta31[4]
eta1424<-theta21[4] - theta41[4]
eta1434<-theta31[4] - theta41[4]
}
#Dados
y[,1]  y[,2]  y[,3]  y[,4]  y[,5]  y[,6]  y[,7]  y[,8]  y[,9]  y[,10]
y[,11] y[,12] y[,13] y[,14] y[,15] y[,16] y[,17] y[,18] y[,19]
y[,20] y[,21] y[,22] y[,23] y[,24] y[,25] y[,26] y[,27] y[,28]
y[,29] y[,30] y[,31] y[,32]
169  291  158  122  203  178  141  181  101  199  141
149  135  153  147  157  125  138  113  137  170  139
131  125  113  150  155  133  113  102  128  91  136
282  128  113  171  154  104  143  80  149  138  131
218  164  192  168  105  94  91  94  150  127  152
105  101  113  161  197  148  191  226  173  125  218
104  91  141  127  90  120  78  147  118  133  204
148  205  168  92  81  81  79  125  103  112  91  89
105  134  218  122  172  261  156  122  154  91  82
109  114  80  125  83  142  108  134  226  146  167
165  86  72  72  75  121  94  99  82  122  90  116
159  143  162  279  202
END
z1[]  z2[]
0  1
0  1
0  1
0  1
8  1
8  1
8  1
8  1
25  1

```


Apêndice D

Código em Winbugs para o Modelo 1 do Capítulo 4

```
#Modelo 1 (l=1,2,3,4,5 espécies e i=1,...,25 troncos,  
  burn-in de 100000 , 3000 salto 100)  
model {  
  for (l in 1:5) {  
    for (i in 1:25) {  
      zeros[l,i] <- 0  
      lambda[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i] + beta31[l]*z31[i]  
+  
beta32[l]*z32[i] + beta41[l]*z41[i] + beta42[l]*z42[i] + beta51[l]*z51[i] +  
beta52[l]*z52[i] + w[i])  
      logit(phi[l,i]) <- gama1[l]*z1[i] + gama2[l]*z2[i] + gama31[l]*z31[i] +  
gama32[l]*z32[i] + gama41[l]*z41[i] + gama42[l]*z42[i] + gama51[l]*z51[i] +  
gama52[l]*z52[i] + w[i]  
    }  
  }  
  ### modelagem dos zeros  
  for (l in 1:5) {  
    for (i in 1:12) {  
      csi[l,i] <- -log(phi[l,i]) - log(1-phi[l,i]) + lambda[l,i]
```

```

        }
    }

for (l in 2:4) {
  for (i in 13:16) {
    csi[l,i]<-log(phi[l,i]) - log(1-phi[l,i]) + lambda[l,i]
  }
}

for (l in 5:5) {
  for (i in 13:23) {
    csi[l,i]<-log(phi[l,i]) - log(1-phi[l,i]) + lambda[l,i]
  }
}

csi[2,17]<-log(phi[2,17]) - log(1-phi[2,17]) + lambda[2,17]
csi[4,17]<-log(phi[4,17]) - log(1-phi[4,17]) + lambda[4,17]
csi[3,18]<-log(phi[3,18]) - log(1-phi[3,18]) + lambda[3,18]
csi[4,18]<-log(phi[4,18]) - log(1-phi[4,18]) + lambda[4,18]

csi[1,19]<-log(phi[1,19]) - log(1-phi[1,19]) + lambda[1,19]
csi[2,19]<-log(phi[2,19]) - log(1-phi[2,19]) + lambda[2,19]
csi[4,19]<-log(phi[4,19]) - log(1-phi[4,19]) + lambda[4,19]

csi[1,20]<-log(phi[1,20]) - log(1-phi[1,20]) + lambda[1,20]
csi[2,20]<-log(phi[2,20]) - log(1-phi[2,20]) + lambda[2,20]
csi[4,20]<-log(phi[4,20]) - log(1-phi[4,20]) + lambda[4,20]

csi[3,21]<-log(phi[3,21]) - log(1-phi[3,21]) + lambda[3,21]
csi[4,21]<-log(phi[4,21]) - log(1-phi[4,21]) + lambda[4,21]
csi[3,22]<-log(phi[3,22]) - log(1-phi[3,22]) + lambda[3,22]
csi[4,22]<-log(phi[4,22]) - log(1-phi[4,22]) + lambda[4,22]

```

```

csi[1,23]<-log(phi[1,23]) - log(1-phi[1,23]) + lambda[1,23]
csi[2,23]<-log(phi[2,23]) - log(1-phi[2,23]) + lambda[2,23]
csi[4,23]<-log(phi[4,23]) - log(1-phi[4,23]) + lambda[4,23]
csi[3,24]<-log(phi[3,24]) - log(1-phi[3,24]) + lambda[3,24]

csi[1,25]<-log(phi[1,25]) - log(1-phi[1,25]) + lambda[1,25]
csi[2,25]<-log(phi[2,25]) - log(1-phi[2,25]) + lambda[2,25]
csi[4,25]<-log(phi[4,25]) - log(1-phi[4,25]) + lambda[4,25]
csi[5,25]<-log(phi[5,25]) - log(1-phi[5,25]) + lambda[5,25]
### modelagem dos outros valores
for (l in 1:1) {
  for (i in 13:18) {
    csi[l,i]<-log(1-phi[l,i]) - y[l,i]*log(lambda[l,i]) + lambda[l,i] + logfact(y[l,i])

    }
  }

csi[3,17]<-log(1-phi[3,17]) - y[3,17]*log(lambda[3,17]) + lambda[3,17] +
logfact(y[3,17])

csi[2,18]<-log(1-phi[2,18]) - y[2,18]*log(lambda[2,18]) + lambda[2,18] +
logfact(y[2,18])

csi[3,19]<-log(1-phi[3,19]) - y[3,19]*log(lambda[3,19]) + lambda[3,19] +
logfact(y[3,19])

csi[3,20]<-log(1-phi[3,20]) - y[3,20]*log(lambda[3,20]) + lambda[3,20] +
logfact(y[3,20])

csi[1,21]<-log(1-phi[1,21]) - y[1,21]*log(lambda[1,21]) + lambda[1,21] +
logfact(y[1,21])

```

$$\text{csi}[2,21] < -\log(1-\text{phi}[2,21]) - y[2,21] * \log(\text{lambda}[2,21]) + \text{lambda}[2,21] + \log\text{fact}(y[2,21])$$

$$\text{csi}[1,22] < -\log(1-\text{phi}[1,22]) - y[1,22] * \log(\text{lambda}[1,22]) + \text{lambda}[1,22] + \log\text{fact}(y[1,22])$$

$$\text{csi}[2,22] < -\log(1-\text{phi}[2,22]) - y[2,22] * \log(\text{lambda}[2,22]) + \text{lambda}[2,22] + \log\text{fact}(y[2,22])$$

$$\text{csi}[3,23] < -\log(1-\text{phi}[3,23]) - y[3,23] * \log(\text{lambda}[3,23]) + \text{lambda}[3,23] + \log\text{fact}(y[3,23])$$

$$\text{csi}[1,24] < -\log(1-\text{phi}[1,24]) - y[1,24] * \log(\text{lambda}[1,24]) + \text{lambda}[1,24] + \log\text{fact}(y[1,24])$$

$$\text{csi}[2,24] < -\log(1-\text{phi}[2,24]) - y[2,24] * \log(\text{lambda}[2,24]) + \text{lambda}[2,24] + \log\text{fact}(y[2,24])$$

$$\text{csi}[4,24] < -\log(1-\text{phi}[4,24]) - y[4,24] * \log(\text{lambda}[4,24]) + \text{lambda}[4,24] + \log\text{fact}(y[4,24])$$

$$\text{csi}[5,24] < -\log(1-\text{phi}[5,24]) - y[5,24] * \log(\text{lambda}[5,24]) + \text{lambda}[5,24] + \log\text{fact}(y[5,24])$$

$$\text{csi}[3,25] < -\log(1-\text{phi}[3,25]) - y[3,25] * \log(\text{lambda}[3,25]) + \text{lambda}[3,25] + \log\text{fact}(y[3,25])$$

```

for (l in 1:5) {
  for (i in 1:25) {
    zeros[l,i] ~ dpois(csi[l,i])
  }
}

```

```

for (i in 1:n) {
  w[i] ~ dnorm(0, tau.w)
}

```

```
    }

    for (l in 1:5) {
      beta1[l] ~ dnorm(0, 1.0E-3)
      beta2[l] ~ dnorm(0, 1.0E-3)
      beta31[l] ~ dnorm(0, 1.0E-3)
      beta32[l] ~ dnorm(0, 1.0E-3)
      beta41[l] ~ dnorm(0, 1.0E-3)
      beta42[l] ~ dnorm(0, 1.0E-3)
      beta51[l] ~ dnorm(0, 1.0E-3)
      beta52[l] ~ dnorm(0, 1.0E-3)
      gama1[l] ~ dnorm(0, 1.0E-3)
      gama2[l] ~ dnorm(0, 1.0E-3)
      gama31[l] ~ dnorm(0, 1.0E-3)
      gama32[l] ~ dnorm(0, 1.0E-3)
      gama41[l] ~ dnorm(0, 1.0E-3)
      gama42[l] ~ dnorm(0, 1.0E-3)
      gama51[l] ~ dnorm(0, 1.0E-3)
      gama52[l] ~ dnorm(0, 1.0E-3)

      lamb[l] ~ dgamma(1, 1)
    }

    tau.w ~ dgamma(0.01, 0.01)
    sigma2.w <- 1/tau.w
  }

  }
```

Dados

```
list(n=25)
```

```
y[,1]  y[,2]  y[,3]  y[,4]  y[,5]  y[,6]  y[,7]  y[,8]
y[,9]  y[,10] y[,11] y[,12] y[,13] y[,14] y[,15]
y[,16] y[,17] y[,18] y[,19] y[,20] y[,21] y[,22]
```


y[,23]	y[,24]	y[,25]										
0	0	0	0	0	0	0	0	0	0	0	0	
1	1	1	1	1	1	0	0	4	5	0	6	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	2	0	0	2	1	0	3	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	3	3	0	0	10	0	17	0	
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0			

END

z1[]	z2[]	z31[]	z32[]	z41[]	z42[]	z51[]	z52[]
389.45	1	0	0	1	0	0	0
448.50	0	0	0	0	1	0	1
204.17	1	0	1	1	0	0	0
139.73	1	0	0	0	1	0	0
196.94	1	0	0	0	0	0	1
288.48	1	0	0	0	1	0	1
31.40	1	0	0	0	1	0	0
17.27	1	0	1	0	1	0	1
592.00	0	0	0	1	0	1	0
92.31	1	0	0	0	1	0	0
21.19	1	1	0	0	1	0	0
2110.08	1	0	0	0	0	0	0
2100.00	0	0	0	0	0	1	0
841.52	1	1	0	1	0	0	0
1898.00	0	0	0	1	0	0	0
461.58	1	0	0	0	0	1	0
307.32	1	0	1	1	0	0	0
284.63	0	0	1	0	1	1	0
1560.00	0	0	0	1	0	1	0

Apêndice E

Código em Winbugs para o Modelo 2 do Capítulo 4

Modelo 1 (l=1,2,3,4,5 espécies e i=1,...,25 troncos,

burn-in de 100000 , 3000 salto 100)

```
model {
  for (l in 1:5) {
    for (i in 1:25) {
      zeros[l,i] <- 0
      lambda[l,i] <- lamb[l]*exp(beta1[l]*z1[i] + beta2[l]*z2[i] +
beta31[l]*z31[i] + beta32[l]*z32[i] + beta41[l]*z41[i] + beta42[l]*z42[i] +
beta51[l]*z51[i] + beta52[l]*z52[i] + w[i])
      logit(phi[l,i]) <- gama1[l]*z1[i] + gama2[l]*z2[i] +
gama31[l]*z31[i] + gama32[l]*z32[i] + gama41[l]*z41[i] + gama42[l]*z42[i] +
gama51[l]*z51[i] + gama52[l]*z52[i] + w[i]
    }
  }
  ### modelagem dos zeros
  for (l in 1:5) {
    for (i in 1:12) {
      csi[l,i]<-log(phi[l,i]) - log(1-phi[l,i]) + lambda[l,i]
    }
  }
}
```

```

    }

for (l in 2:4) {
  for (i in 13:16) {
    csi[l,i]<-log(phi[l,i]) - log(1-phi[l,i]) + lambda[l,i]
  }
}

for (l in 5:5) {
  for (i in 13:23) {
    csi[l,i]<-log(phi[l,i]) - log(1-phi[l,i]) + lambda[l,i]
  }
}

csi[2,17]<-log(phi[2,17]) - log(1-phi[2,17]) + lambda[2,17]
csi[4,17]<-log(phi[4,17]) - log(1-phi[4,17]) + lambda[4,17]
csi[3,18]<-log(phi[3,18]) - log(1-phi[3,18]) + lambda[3,18]
csi[4,18]<-log(phi[4,18]) - log(1-phi[4,18]) + lambda[4,18]

csi[1,19]<-log(phi[1,19]) - log(1-phi[1,19]) + lambda[1,19]
csi[2,19]<-log(phi[2,19]) - log(1-phi[2,19]) + lambda[2,19]
csi[4,19]<-log(phi[4,19]) - log(1-phi[4,19]) + lambda[4,19]

csi[1,20]<-log(phi[1,20]) - log(1-phi[1,20]) + lambda[1,20]
csi[2,20]<-log(phi[2,20]) - log(1-phi[2,20]) + lambda[2,20]
csi[4,20]<-log(phi[4,20]) - log(1-phi[4,20]) + lambda[4,20]

csi[3,21]<-log(phi[3,21]) - log(1-phi[3,21]) + lambda[3,21]
csi[4,21]<-log(phi[4,21]) - log(1-phi[4,21]) + lambda[4,21]
csi[3,22]<-log(phi[3,22]) - log(1-phi[3,22]) + lambda[3,22]
csi[4,22]<-log(phi[4,22]) - log(1-phi[4,22]) + lambda[4,22]

```

```

csi[1,23]<-log(phi[1,23]) - log(1-phi[1,23]) + lambda[1,23]
csi[2,23]<-log(phi[2,23]) - log(1-phi[2,23]) + lambda[2,23]
csi[4,23]<-log(phi[4,23]) - log(1-phi[4,23]) + lambda[4,23]
csi[3,24]<-log(phi[3,24]) - log(1-phi[3,24]) + lambda[3,24]

csi[1,25]<-log(phi[1,25]) - log(1-phi[1,25]) + lambda[1,25]
csi[2,25]<-log(phi[2,25]) - log(1-phi[2,25]) + lambda[2,25]
csi[4,25]<-log(phi[4,25]) - log(1-phi[4,25]) + lambda[4,25]
csi[5,25]<-log(phi[5,25]) - log(1-phi[5,25]) + lambda[5,25]
### modelagem dos outros valores
for (l in 1:1) {
  for (i in 13:18) {
    csi[l,i]<-log(1-phi[l,i]) - y[l,i]*log(lambda[l,i]) + lambda[l,i] +
logfact(y[l,i])
  }
}

csi[3,17]<-log(1-phi[3,17]) - y[3,17]*log(lambda[3,17]) + lambda[3,17] +
logfact(y[3,17])

csi[2,18]<-log(1-phi[2,18]) - y[2,18]*log(lambda[2,18]) + lambda[2,18] +
logfact(y[2,18])

csi[3,19]<-log(1-phi[3,19]) - y[3,19]*log(lambda[3,19]) + lambda[3,19] +
logfact(y[3,19])

csi[3,20]<-log(1-phi[3,20]) - y[3,20]*log(lambda[3,20]) + lambda[3,20] +
logfact(y[3,20])

csi[1,21]<-log(1-phi[1,21]) - y[1,21]*log(lambda[1,21]) + lambda[1,21] +
logfact(y[1,21])
csi[2,21]<-log(1-phi[2,21]) - y[2,21]*log(lambda[2,21]) + lambda[2,21] +

```

```

logfact(y[2,21])

  csi[1,22]<-log(1-phi[1,22]) - y[1,22]*log(lambda[1,22]) + lambda[1,22] +
logfact(y[1,22])
  csi[2,22]<-log(1-phi[2,22]) - y[2,22]*log(lambda[2,22]) + lambda[2,22] +
logfact(y[2,22])
  csi[3,23]<-log(1-phi[3,23]) - y[3,23]*log(lambda[3,23]) + lambda[3,23] +
logfact(y[3,23])

  csi[1,24]<-log(1-phi[1,24]) - y[1,24]*log(lambda[1,24]) + lambda[1,24] +
logfact(y[1,24])
  csi[2,24]<-log(1-phi[2,24]) - y[2,24]*log(lambda[2,24]) + lambda[2,24] +
logfact(y[2,24])
  csi[4,24]<-log(1-phi[4,24]) - y[4,24]*log(lambda[4,24]) + lambda[4,24] +
logfact(y[4,24])
  csi[5,24]<-log(1-phi[5,24]) - y[5,24]*log(lambda[5,24]) + lambda[5,24] +
logfact(y[5,24])

  csi[3,25]<-log(1-phi[3,25]) - y[3,25]*log(lambda[3,25]) + lambda[3,25] +
logfact(y[3,25])

  for (l in 1:5) {
    for (i in 1:25) {
      zeros[l,i] ~ dpois(csi[l,i])
    }
  }

eta<-1/phi2

for (i in 1:n) {
  w[i] ~ dgamma(eta, eta)

```

```

    }

    for (l in 1:5) {
      beta1[l] ~ dnorm(0, 1.0E-3)
      beta2[l] ~ dnorm(0, 1.0E-3)
      beta31[l] ~ dnorm(0, 1.0E-3)
      beta32[l] ~ dnorm(0, 1.0E-3)
      beta41[l] ~ dnorm(0, 1.0E-3)
      beta42[l] ~ dnorm(0, 1.0E-3)
      beta51[l] ~ dnorm(0, 1.0E-3)
      beta52[l] ~ dnorm(0, 1.0E-3)
      gama1[l] ~ dnorm(0, 1.0E-3)
      gama2[l] ~ dnorm(0, 1.0E-3)
      gama31[l] ~ dnorm(0, 1.0E-3)
      gama32[l] ~ dnorm(0, 1.0E-3)
      gama41[l] ~ dnorm(0, 1.0E-3)
      gama42[l] ~ dnorm(0, 1.0E-3)
      gama51[l] ~ dnorm(0, 1.0E-3)
      gama52[l] ~ dnorm(0, 1.0E-3)

      lamb[l] ~ dgamma(1, 1)
    }
    phi2 ~ dgamma(1,1)
  }

```

Dados

```
list(n=25)
```

```

y[,1]  y[,2]  y[,3]  y[,4]  y[,5]  y[,6]  y[,7]  y[,8]
y[,9]  y[,10] y[,11] y[,12] y[,13] y[,14] y[,15]
y[,16] y[,17] y[,18] y[,19] y[,20] y[,21] y[,22]
y[,23] y[,24] y[,25]

```

```

0  0  0  0  0  0  0  0  0  0  0  0  1
1  1  1  1  1  0  0  4  5  0  6  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  2  0  0  2  1  0  3  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0
0  1  0  3  3  0  0  10  0  17  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  0  1  0  0  0  0
0  0  0  0  0  0  0  0  0  0  0  0  0
0  0  0  0  0  0  0  1  0

```

END

```

z1[]  z2[]  z31[]  z32[]  z41[]  z42[]  z51[]  z52[]
389.45  1  0  0  1  0  0  0
448.50  0  0  0  0  1  0  1
204.17  1  0  1  1  0  0  0
139.73  1  0  0  0  1  0  0
196.94  1  0  0  0  0  0  1
288.48  1  0  0  0  1  0  1
31.40   1  0  0  0  1  0  0
17.27   1  0  1  0  1  0  1
592.00  0  0  0  1  0  1  0
92.31   1  0  0  0  1  0  0
21.19   1  1  0  0  1  0  0
2110.08  1  0  0  0  0  0  0
2100.00  0  0  0  0  0  1  0
841.52  1  1  0  1  0  0  0
1898.00  0  0  0  1  0  0  0
461.58  1  0  0  0  0  1  0
307.32  1  0  1  1  0  0  0
284.63  0  0  1  0  1  1  0
1560.00  0  0  0  1  0  1  0
59.66   1  0  0  1  0  0  0

```