

UNIVERSIDADE FEDERAL DE MINAS GERAIS - UFMG  
INSTITUTO DE CIÊNCIAS EXATAS - ICE<sub>x</sub>  
DEPARTAMENTO DE ESTATÍSTICA

Dissertação de Mestrado:

**MODELOS LOGÍSTICOS COM CLASSES ESTENDIDAS DE  
DISTRIBUIÇÕES NORMAIS PARA OS EFEITOS  
ALEATÓRIOS**

Cristiano de Carvalho Santos  
Orientadora: Rosângela Helena Loschi

CRISTIANO DE CARVALHO SANTOS

MODELOS LOGÍSTICOS COM CLASSES ESTENDIDAS DE  
DISTRIBUIÇÕES NORMAIS PARA OS EFEITOS  
ALEATÓRIOS

Dissertação apresentada ao Programa de Pós-Graduação em Estatística do Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial a obtenção do título de Mestre em Estatística.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Rosângela Helena Loschi.

Universidade Federal de Minas Gerais  
Belo Horizonte  
Fevereiro de 2011

Dedico este trabalho  
aos meus pais, Lucia e Geraldo.

# Agradecimentos

Agradeço a Deus pelas conquistas que tem me proporcionado, pelas pessoas que tem colocado na minha vida e pela minha saúde.

Aos meus pais, Lucia e Geraldo, pelo apoio em minhas decisões, incentivo em meu projetos e por sempre acreditarem no meu sucesso. Por terem me ensinado a buscar meus objetivos com dedicação, honestidade, dignidade e humildade. Por terem me ensinado que nunca se perde tempo ao estudar.

À minha família e aos amigos pela confiança e apoio.

Aos amigos que fiz na UFMG, durante a graduação e o mestrado, pelas discussões estatísticas que, de forma direta ou indireta, me ajudaram nesta dissertação, pelos dias de estudo em conjunto para as inúmeras provas e listas de exercícios e, principalmente, pelos momentos de descontração tão importantes e necessários.

À minha orientadora, Professora Rosangela H. Loschi pela confiança, paciência, dedicação e comprometimento ao me orientar durante vários anos de estudos na UFMG.

Aos membros da banca examinadora, professores Reinaldo Boris Arellano-Valle (PUC-Chile), Sueli Aparecida Mingoti (UFMG) e Fábio Nogueira Demarqui (UFMG) pelas contribuições, correções e sugestões na dissertação.

Aos professores do departamento de Estatística pelos conhecimentos compartilhados durante estes seis anos de convívio.

À CAPES e à FAPEMIG pela bolsa de mestrado, uma vez que iniciei recebendo bolsa da primeira instituição e depois, para atender os interesses do departamento, passei a receber bolsa da outra instituição. À FAPEMIG por diversos apoios financeiros prestados para a participação em eventos e ao CNPq pela bolsa de iniciação científica que me colocou em contato com pesquisas e despertou meu interesse pela área acadêmica.

Enfim, agradeço a todos que contribuíram para esta conquista.

Muito Obrigado!  
Cristiano de Carvalho Santos

# Resumo

Esta dissertação apresenta uma extensão do modelo logístico misto ao considerar classes estendidas de distribuições normais para os efeitos aleatórios. Os efeitos aleatórios são introduzidos para modelar dados correlacionados, superdispersão nos dados ou para representar covariáveis que não foram medidas e, geralmente, assume-se a distribuição normal para representar o comportamento dos efeitos aleatórios. Uma das contribuições desta dissertação é a implementação através da abordagem bayesiana do modelo logístico misto considerando, para os efeitos aleatórios, as distribuições normais assimétricas univariada e multivariada.

Ao adotar o enfoque clássico, a razão de chances no modelo logístico misto é, em algumas situações, uma variável aleatória, pois depende dos efeitos aleatórios que são variáveis aleatórias. Larsen *et al.* (2000) e Larsen e Merlo (2005) encontraram a distribuição da razão de chances no modelo logístico misto com efeitos aleatórios normalmente distribuídos, propuseram a razão de chances mediana como uma medida para interpretar a razão de chances e encontram algumas propriedades desta medida. Esta dissertação também contribui ao estender estes resultados para as situações em que se utilizam as distribuições normais assimétricas univariada e multivariada para representar o comportamento dos efeitos aleatórios. Nesta última contribuição são obtidos alguns resultados originais referentes a distribuição de combinações lineares de variáveis aleatórias distribuídas segundo distribuições normais assimétricas e de funções destas combinações.

Um pequeno estudo com dados simulados é realizado para avaliar o efeito da má especificação da distribuição dos efeitos aleatórios na estimação dos efeitos fixos, dos efeitos aleatórios e também da razão de chances. Neste estudo verificamos que, de forma geral, o modelo assumindo distribuição normal assimétrica univariada para os efeitos aleatórios detectou assimetria na distribuição dos efeitos aleatórios quando ela existia, e nestes casos, melhorou um pouco as estimativas para os efeitos aleatórios. Os modelos propostos são aplicados para analisar os dados reais, apresentados em Liu e Dey (2008), de um experimento de atividade teratogênica com análogos de niacina.

**Palavras-chave:** Modelo logístico misto, razão de chances e distribuição normal assimétrica.

# Abstract

This work aims at extending the mixed logistic regression model (MLRM) by considering classes of univariate and multivariate skew-normal distributions for the random effects. The Bayesian paradigm is used. Random effects are introduced into the model to represent unmeasured covariates or as a way of modeling data correlation or overdispersion and, usually, it is assumed that the random effects are normally distributed.

As for the logistic regression, the odds ratio is used for the parameters interpretation. However, under the MLRM, for some cases, the odds ratio depends on the random effects thus it is a random quantity under both Bayesian and classical approaches. From the classical point of view, Larsen *et al.* (2000) find the distribution of the odds ratio under normality for the random effects. They suggest using the median odds ratio (MOR) to interpret the parameters. According to Larsen *et al.* (2000), the MOR quantifies appropriately the heterogeneity among the different groups. One of the main contributions of this work is to extend such results whenever both univariate and multivariate skew-normal distributions are considered to model the uncertainty about the random effects. As a byproduct, some results related to linear combination of skew-normal random variables are obtained.

Considering simulated data sets, a study is performed in order to evaluate the effect of the misspecification of the distributions of the random effects in the estimations of the fixed and random effects as well as the odds ratio. In general the model which assumes the univariate skew-normal for the random effects leads to the improvement of the estimates for the random effects and to the detection of the asymmetry in their distribution if it exists. Similar inferences are obtained if normal and the multivariate skew-normal are assumed. The proposed models are considered to analyze the data set reported in Liu e Dey (2008).

**Keywords:** Mixed logistic model, odds ratio and skew-normal distribution.

# Sumário

Página

<b>Introdução</b>	<b>11</b>
<b>1 Modelo de Regressão Logística com Efeitos Aleatórios Normalmente Distribuídos</b>	<b>14</b>
1.1 Introdução . . . . .	14
1.2 Modelo Logístico com Efeitos Aleatórios . . . . .	15
1.3 Inferência considerando Efeitos Aleatórios Normalmente Distribuídos . . . . .	15
<b>2 Inferência Bayesiana para o Modelo de Regressão Logística com Efeitos Aleatórios em Classes Estendidas da Distribuição Normal</b>	<b>18</b>
2.1 Introdução . . . . .	18
2.2 Definições e Propriedades da Distribuição Normal Assimétrica . . . . .	19
2.3 Modelo Logístico com Efeitos Aleatórios SN Independentes . . . . .	22
2.4 Modelo Logístico com Efeitos Aleatórios SN Correlacionados . . . . .	25
<b>3 Razão de Chances no Modelo Logístico Misto</b>	<b>28</b>
3.1 Efeitos Aleatórios Normalmente Distribuídos . . . . .	29
3.2 Efeitos Aleatórios Independentes com Distribuição Normal Assimétrica Univariada . . . . .	30
3.3 Efeitos Aleatórios Correlacionados com Distribuição Normal Assimétrica Multivariada . . . . .	32
3.4 Determinando a Razão de Chances <i>a posteriori</i> . . . . .	33
<b>4 Estudo com dados Simulados</b>	<b>35</b>
4.1 Efeitos Aleatórios com Distribuição Normal . . . . .	36
4.2 Efeitos Aleatórios com Distribuição Normal Assimétrica Univariada . . . . .	40
4.3 Efeitos Aleatórios com Distribuição Normal Assimétrica Multivariada . . . . .	44
<b>5 Aplicação a Dados Reais</b>	<b>49</b>
<b>Conclusões e Discussões</b>	<b>54</b>

<b>A Prova da Proposição 2.2.2</b>	<b>56</b>
<b>B Distribuições Condicionais Completas para Modelo Logístico Misto com distribuições assimétricas para os Efeitos Aleatórios</b>	<b>62</b>
B.1 Distribuição Normal Assimétrica Univariada para os Efeitos Aleatórios . . .	62
B.2 Distribuição Normal Assimétrica Multivariada para os Efeitos Aleatórios . .	63
<b>C Provas dos Teoremas e Proposições do Capítulo 3</b>	<b>64</b>
C.1 Provas dos Teoremas da Seção 3.1 . . . . .	64
C.2 Provas das Proposições da Seção 3.2 . . . . .	67
C.3 Provas das Proposições da Seção 3.3 . . . . .	70
<b>D Medidas de Qualidade de Ajuste dos Modelos</b>	<b>74</b>
D.1 Conditional Predictive Ordinate (CPO) . . . . .	74
D.2 Deviance Information Criterion (DIC) . . . . .	74
D.3 Índice C . . . . .	75
D.4 Taxa de Classificação Correta (TCC) . . . . .	75
<b>Referências Bibliográficas</b>	<b>78</b>



# Lista de Figuras

2.1	Funções de densidade da distribuição normal assimétrica padrão com diversos níveis de assimetria . . . . .	19
2.2	Funções de densidade da distribuição normal assimétrica com parâmetros de locação e escala . . . . .	20
4.1	Histograma e densidade teórica dos efeitos aleatórios $\gamma$ , com distribuição normal, gerados para os bancos de dados 1 e 2 . . . . .	36
4.2	Histograma e densidade teórica dos efeitos aleatórios $\gamma$ , com distribuição assimétrica univariada, gerados para os bancos de dados 1, 2, 3 e 4 . . . . .	40
4.3	Histograma dos efeitos aleatórios $\gamma$ , com distribuição normal assimétrica multivariada, gerados para os bancos de dados 1, 2, 3 e 4 . . . . .	45

# Lista de Tabelas

4.1	Estimativas <i>a posteriori</i> para os parâmetros . . . . .	37
4.2	Soma de quadrados dos erros para os efeitos aleatórios $\gamma$ . . . . .	38
4.3	Medidas de qualidade de ajuste do modelos . . . . .	38
4.4	Estimativas <i>a posteriori</i> para a razão de chances . . . . .	39
4.5	Estimativas <i>a posteriori</i> para os parâmetros . . . . .	41
4.6	Soma de quadrados dos erros para os efeitos aleatórios $\gamma$ . . . . .	42
4.7	Medidas de qualidade de ajuste dos modelos . . . . .	42
4.8	Estimativas <i>a posteriori</i> para a razão de chances . . . . .	43
4.9	Estimativas <i>a posteriori</i> para os parâmetros . . . . .	46
4.10	Soma de quadrados dos erros para os efeitos aleatórios $\gamma$ . . . . .	47
4.11	Medidas de qualidade de ajuste dos modelos . . . . .	47
4.12	Estimativas <i>a posteriori</i> para a razão de chances . . . . .	48
5.1	Dados experimentais. (Tabela Reformulada) . . . . .	49
5.2	Estimativas <i>a posteriori</i> para os parâmetros . . . . .	50
5.3	Medidas de qualidade ajuste dos modelos . . . . .	51
5.4	Estimativas <i>a posteriori</i> para a razão de chances . . . . .	52

# Introdução

Dados binários ocorrem frequentemente em diversas áreas da ciência, por exemplo, nas pesquisas médicas, sociais, econômicas e na agricultura. Nestes casos, em geral, o objetivo da modelagem estatística é explicar a probabilidade de ocorrência de um evento de interesse. A ocorrência de um evento de interesse é usualmente denominada por sucesso. O modelo de regressão logística (MRL) é um modelo muito utilizado para descrever a relação entre tal probabilidade de sucesso e as covariáveis observadas. Esta é uma relação não linear entre a resposta binária e o conjunto de variáveis explicativas.

Apesar de muito útil, os modelos logísticos não são apropriados para a análise de dados binários que apresentam estrutura de correlação, pois estes modelos assumem independência entre as observações. Dados correlacionados ocorrem, por exemplo, em estudos onde os indivíduos estão separados em *clusters* e também em estudos longitudinais. Em estudos com *clusters*, os indivíduos são observados dentro de grandes unidades, por exemplo, escolas, famílias, hospitais ou o local de trabalho. Em geral, as observações dentro de um *cluster* tendem a ser correlacionadas devido a várias características em comum. Exemplificando, se escolas são consideradas como *clusters*, observações de indivíduos dentro de uma mesma escola podem estar correlacionadas devido a características da prática de ensino que são particulares aquela escola. Já em estudos longitudinais, são feitas observações repetidas para cada indivíduo e, conseqüentemente, as observações para um mesmo indivíduo são correlacionadas. Estes estudos são conhecidos como modelos multinível ou hierárquicos, pois as observações no primeiro nível (observações) estão aninhadas dentro do segundo nível (*clusters*).

Uma estratégia usualmente adotada para modelar dados correlacionados é considerarmos Modelos Lineares Generalizados Mistos (GLMM), em que efeitos aleatórios são incluídos para introduzir a estrutura de correlação entre as observações. Essas observações podem, então, serem consideradas independentes quando condicionamos no conhecimento dos efeitos aleatórios. Estes modelos estendem a classe de Modelos Lineares Generalizados (GLM) através da adição dos termos de efeito aleatório no preditor linear do modelo correspondente.

A adição dos termos de efeitos aleatórios no modelo também pode ser utilizada como alternativa para levarmos em consideração a ocorrência de superdispersão nos dados (ver, por exemplo, Schall (1991)), ou ainda, para representar o efeito de variáveis que não foram medidas, mas que são importantes para explicar parte da variabilidade apresentada pela variável resposta. Souza e Migon (2010) utilizaram um termo de efeito aleatório no intercepto do modelo de regressão logística para acomodar e detectar *outliers*. Nesse caso, a introdução do termo de efeito aleatório visa acomodar uma eventual variabilidade extra que não é considerada na falta dos efeitos aleatórios. Em Gibbons *et al.* (1994), o efeito aleatório é

usado para representar as características não observáveis e/ou não medidas que levam um médico a ter maior risco de sofrer um processo judicial por erro médico.

Do ponto de vista clássico, fazer inferência em modelos com efeitos aleatórios não é simples na maioria dos casos. Para fazermos inferência sobre os efeitos fixos  $\beta$ , que são os parâmetros de interesse, assumimos uma distribuição para os efeitos aleatórios  $\gamma$  e integramos de forma a obter uma função de verossimilhança livre dos efeitos aleatórios. Feito isto, utilizamos os métodos habituais para obtermos as estimativas de  $\beta$ . Desta forma, é necessário calcularmos as integrais, algumas vezes de alta dimensão, na função de verossimilhança conjunta entre dados e efeitos aleatórios.

Uma das grandes dificuldades em trabalharmos com modelos de efeitos aleatórios do ponto de vista clássico é a resolução de tais integrais. A inferência torna-se factível apenas em casos bem simples, em que, em geral, assume-se que os efeitos aleatórios  $\gamma$  são normalmente distribuídos. Esta suposição foi considerada, por exemplo, em Larsen *et al.* (2000), Larsen e Merlo (2005), Gibbons *et al.* (1994), entre outros. Supondo normalidade, este problema de integração é resolvido através do uso de técnicas de integração numérica, como por exemplo, a quadratura de Gauss-Hermite (McCulloch e Searle (2001)) ou a aproximação de Laplace. Breslow e Clayton (1993) propõem um método baseado em pseudo quase verossimilhança (PQL), em que a estimação dos parâmetros é feita pela maximização da função de verossimilhança marginal após utilizarmos a aproximação de Laplace para resolvermos a integral nos efeitos aleatórios. Em Nelson *et al.* (2006) e Liu e Yu (2008), propõe-se transformações na função de verossimilhança conjunta entre dados e efeitos aleatórios para, ao assumirmos uma distribuição não-normal para os efeitos aleatórios, resolvermos o problema da integração através dos métodos numéricos habituais.

Do ponto de vista Bayesiano, como são desconhecidos, os efeitos aleatórios recebem o *status* de parâmetros a serem estimados (ver Paulino *et al.* (2005), Souza e Migon (2010), Zeger e Karim (1991)). Desta forma, métodos Monte Carlo via Cadeias de Markov (MCMC) podem tornar mais fácil o problema de inferência, uma vez que fornecem boas aproximações para a distribuição *a posteriori* dos parâmetros de interesse. Esta facilidade computacional viabiliza o uso de distribuições mais apropriadas para descrever o comportamento dos efeitos aleatórios. Por exemplo, no modelo Logístico Misto, Souza e Migon (2010) utilizam as distribuições *t*-Student e mistura de distribuições normais, independentes, para este fim. Anteriormente, Liu e Dey (2008) consideram efeitos aleatórios independentes e distribuídos segundo distribuições assimetrizadas (*t*-Student (Sahu *et al.* (2003)) e normal (Azzalini (1985))) fazendo uma comparação com o caso em que uma mistura de Processos Dirichlet (MDP) é considerada para modelar o efeito aleatório.

A interpretação dos parâmetros para modelos mistos não lineares, em particular no modelo logístico, não é fácil. No modelo logístico misto, a razão de chances (OR) é função tanto dos efeitos fixos quanto dos efeitos aleatórios (Larsen *et al.* (2000)), sendo, portanto, um objeto aleatório em ambas abordagens, Clássica e Bayesiana. A razão de chances mediana foi introduzida por Larsen *et al.* (2000) como uma medida eficiente de heterogeneidade e que, em algumas comparações, tem a simples e conveniente interpretação em termos da razão de chances usual.

Geralmente, utilizada para representar o comportamento dos efeitos aleatórios, a distribuição normal pode ser restritiva em situações em que estes efeitos sejam distribuídos

assimetricamente. Uma alternativa bastante atrativa para tais casos é considerar a distribuição normal assimétrica. Esta distribuição, proposta na versão univariada por Azzalini (1985) e na versão multivariada por Azzalini e Dalla-Valle (1996), é mais flexível no sentido de considerar uma possível assimetria. Além disso, tal distribuição apresenta algumas características interessantes, por exemplo, o fato de possuir domínio no conjunto dos números reais e possuir a distribuição normal como um caso particular.

Nesta dissertação, utilizamos o paradigma bayesiano para fazer inferências no modelo logístico com efeitos aleatórios. Estendemos os trabalhos de Larsen *et al.* (2000), Liu e Dey (2008) e Souza e Migon (2010) por considerarmos classes estendidas de distribuições normais para os efeitos aleatórios e por considerarmos que os efeitos aleatórios são tanto independentes quanto correlacionados. As distribuições para a “*odds ratio*” (razão de chances) quando tais distribuições são assumidas são obtidas estendendo os resultados de Larsen *et al.* (2000). Como sub-produto também obtemos alguns resultados originais referentes a distribuição de combinações lineares de variáveis aleatórias distribuídas segundo distribuições normais assimétricas e de funções destas. Apresentamos um pequeno estudo com dados simulados avaliando o efeito da má especificação da distribuição dos efeitos aleatórios na estimação dos efeitos fixos, dos efeitos aleatórios e também da razão de chances. Aplicaremos os modelos propostos para analisar os dados reais em Liu e Dey (2008).

Neste trabalho os modelos estatísticos descritos foram implementados no *software* gratuito Ox (Doornik (2007)) e os resultados obtidos foram analisados utilizando o *software* gratuito R (R Development Core Team (2009)).

A dissertação está organizada da seguinte forma. No Capítulo 1 apresentamos o modelo de regressão logística com efeitos aleatórios e inferências relacionadas a ele, usando a abordagem bayesiana, quando consideramos a distribuição normal para os efeitos aleatórios. O Capítulo 2 apresenta o modelo logístico misto quando consideramos classes estendidas da distribuição normal para descrever o comportamento dos efeitos aleatórios. Nesse capítulo, são consideradas as distribuições normal assimétrica univariada e multivariada para este fim. No Capítulo 3 apresentamos a interpretação dos parâmetros do modelo, em termos de razão de chances, ao assumir as distribuições normal, normal assimétrica univariada e normal assimétrica multivariada para os efeitos aleatórios. O Capítulo 4 mostra um estudo, com dados simulados, implementado para comparar os modelos estatísticos em situações em que os dados são gerados considerando as distribuições normal, normal assimétrica univariada e normal assimétrica multivariada para gerar os efeitos aleatórios. No Capítulo 5 analisamos um banco de dados reais em que a proposta é estudar a interação entre a atividade teratogênica de dois análogos de niacina 6-aminonicotinamida (6AN) e 3-acetilpyridina (6AP). O banco de dados para este estudo é apresentado em Liu e Dey (2008). Por fim, no Capítulo 5 são apresentadas algumas conclusões e possíveis trabalhos futuros a serem desenvolvidos.

# Capítulo 1

## Modelo de Regressão Logística com Efeitos Aleatórios Normalmente Distribuídos

### 1.1 Introdução

O modelo de regressão logística misto (MRLM) vem sendo amplamente discutido e utilizado na literatura para modelar o comportamento de dados binários correlacionados ou para descrever mais apropriadamente a variabilidade presente nos dados que não é capturada pelo modelo logístico usual. Para citar alguns exemplos, em Gibbons *et al.* (1994) o efeito aleatório é usado para representar as características não observáveis e/ou não medidas que levam um médico a ter um maior risco de sofrer um processo por erro médico e em Souza e Migon (2010) o termo de efeito aleatório é utilizado no intercepto do modelo de regressão logística para acomodar e detectar outliers, ou seja, neste artigo os efeitos aleatórios são utilizados para acomodar superdispersão nos dados. Em Larsen *et al.* (2000) é investigada a ocorrência de *Ascaris suum* em porcos, situação na qual, devido a condições igualitárias de higiene, as observações para porcos de um mesmo chiqueiro são correlacionadas. Diferenças dos chiqueiros quanto ao manejo dos porcos faz com que os porcos de chiqueiros diferentes venham a ter chances diferentes de ter *Ascaris suum*. Em Paulino *et al.* (2005) são considerados dados de um estudo da universidade da Califórnia, São Francisco, sobre infecção de HPV (*human papillomavirus*). Neste estudo os dados foram analisados através do modelo misto de regressão binomial com erro de classificação na variável resposta.

Geralmente, assume-se que os efeitos aleatórios incluídos no MRLM são independentes e normalmente distribuídos (ver Larsen *et al.* (2000), por exemplo). Esta escolha é, em geral, feita pela facilidade computacional que ela traz para o problema de inferência e não pela adequação desta distribuição para descrever o comportamento dos efeitos aleatórios. Neste capítulo, faremos uma breve revisão do MRLM com efeitos aleatórios normalmente distribuídos.

## 1.2 Modelo Logístico com Efeitos Aleatórios

Seja  $y_{ij}$  uma variável resposta binária observada para o  $j$ -ésimo indivíduo pertencente ao  $i$ -ésimo *cluster* e tal que  $y_{ij}$  é igual a 1 se o evento de interesse ocorreu para tal indivíduo e é igual a 0 caso contrário,  $i = 1, \dots, k$  e  $j = 1, \dots, n_i$ . Suponha que as unidades amostrais possam ser divididas em  $k$  *clusters* e que cada *cluster*  $i$  contenha  $n_i$  indivíduos. Além disso, para o indivíduo  $j$  do  $i$ -ésimo *cluster*, considere  $\mathbf{x}_{ij} = (1, x_{ij1}, x_{ij2}, \dots, x_{ijp})'$  o vetor de covariáveis observadas, de ordem  $(p + 1)$ , e  $\mathbf{z}_{ij}$  um vetor  $k \times 1$  que indica a qual *cluster* o  $j$ -ésimo indivíduo pertence, sendo que este vetor assume o valor um na  $i$ -ésima posição e zero nas demais. Defina por  $N = \sum_{i=1}^k n_i$  o número total de indivíduos. Seja  $\mathbf{X}$  a matriz, de dimensão  $N \times p + 1$ , contendo as covariáveis de todos os indivíduos de forma que cada linha contenha o vetor de covariáveis de um dado indivíduo e  $\mathbf{Z}$  a matriz, com dimensão  $N \times k$ , com os vetores indicadores de *clusters* para todos os indivíduos. Defina por  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_k)'$  o vetor  $k \times 1$  de efeitos aleatórios, onde  $\gamma_i$  representa o efeito aleatório atribuído ao  $i$ -ésimo *cluster* e por  $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)'$  o vetor de efeitos fixos com  $p + 1$  parâmetros desconhecidos.

Seja  $\pi_{ij} = P(y_{ij} = 1 | \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{x}_{ij}, \mathbf{z}_{ij})$ . Vamos adotar a introdução no preditor linear  $\eta_{ij}$  de um termo de efeito aleatório para cada um dos *clusters*, isto é, iremos considerar  $\eta_{ij} = \mathbf{x}'_{ij}\boldsymbol{\beta} + \mathbf{z}'_{ij}\boldsymbol{\gamma}$ . Utilizando a função de ligação logito para modelar  $\pi_{ij}$  temos que:

$$\pi_{ij} = \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}}. \quad (1.1)$$

Se consideramos um caso particular onde número de *clusters*  $k$  é igual ao número total de indivíduos  $N$ , estaremos no caso onde cada indivíduo da amostra pertence a um *cluster* diferente, e sendo assim, o efeito aleatório está sendo incluído no intercepto. Esse modelo é conhecido como modelo de intercepto aleatório e esse tipo de abordagem é adotado, por exemplo, em Souza e Migon (2010).

Suponha que, dado  $\pi_{ij}$ ,  $i = 1, \dots, k$  e  $j = 1, \dots, n_i$ , os  $y_{ij}$ 's são independentes e distribuídos seguindo uma distribuição Bernoulli( $\pi_{ij}$ ), denotada por  $y_{ij} | \pi_{ij} \sim Ber(\pi_{ij})$ . De (1.1) temos que a função de verossimilhança é dada por

$$f(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}) = \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp\{\eta_{ij}\}} \right]^{1-y_{ij}}, \quad (1.2)$$

onde  $\mathbf{y} = (y_{11}, \dots, y_{1n_1}, \dots, y_{k1}, \dots, y_{kn_k})'$  é um vetor de dimensão  $N \times 1$  contendo o valor observado para variável resposta em todos os indivíduos.

A distribuição de  $\mathbf{y}$ , bem como todas as distribuições condicionadas em  $\mathbf{y}$ , também é condicionada no conhecimento de  $\mathbf{X}$  e  $\mathbf{Z}$ . No entanto, para simplicidade de notação, este condicionamento será omitido ao longo de todo o texto.

No que segue, faremos inferências sobre  $\boldsymbol{\beta}$  e  $\boldsymbol{\gamma}$  utilizando o paradigma bayesiano. Para procedimentos clássicos de inferência no MRLM, ver McCulloch e Searle (2001), Breslow e Clayton (1993), entre outros.

## 1.3 Inferência considerando Efeitos Aleatórios Normalmente Distribuídos

Considerar que os efeitos aleatórios possuem distribuição normal multivariada é uma suposição usual em modelos mistos. Tal suposição é feita, por exemplo, nos trabalhos de Larsen *et al.* (2000),

Larsen e Merlo (2005), Breslow e Clayton (1993), entre outros. Nesta seção, assumimos que o vetor de efeitos aleatórios  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_k)'$  tem distribuição normal,  $k$ -variada, com vetor de médias  $\boldsymbol{\mu} = (\mu, \dots, \mu)'$  e matriz de covariância  $\sigma^2 I_k$ , em que  $I_k$  é a matriz identidade de ordem  $k$ .

Ao fazermos tal suposição, estamos considerando que, condicional aos parâmetros  $\boldsymbol{\mu}$  e  $\sigma$  que indexam a distribuição de cada um dos efeitos aleatórios  $\gamma_i$ 's, os efeitos aleatórios  $\gamma_i$ 's são independentes e identicamente distribuídos com distribuição normal com média  $\mu$  e variância  $\sigma^2$ . Sob a condição de independência condicional para os  $\gamma_i$ 's, obtemos que a distribuição conjunta do vetor  $\boldsymbol{\gamma}$  é tal que  $f_{\boldsymbol{\gamma}}(\boldsymbol{\gamma}|\boldsymbol{\mu}, \sigma^2) = \prod_{i=1}^k f_{\gamma_i}(\gamma_i|\mu, \sigma^2)$ .

Considerando a função de verossimilhança em (1.2), para completarmos a modelagem bayesiana faz-se necessário especificarmos as distribuições *a priori* para cada um dos parâmetros do modelo. Para o vetor  $\boldsymbol{\beta}$  de parâmetros associados aos efeitos fixos, em todos os casos, consideramos que o vetor  $\boldsymbol{\beta}$  é normalmente distribuído com vetor de médias  $\mathbf{m} = (m_0, \dots, m_p)'$  e matriz de covariâncias  $b^2 I_{p+1}$ , onde  $I_{p+1}$  denota a matriz identidade de ordem  $p + 1$ . Assumimos, para tal distribuição, a notação  $\boldsymbol{\beta} \sim N_{p+1}(\mathbf{m}, b^2 I_{p+1})$ . Ao fazer tal suposição estamos considerando, *a priori*, que os parâmetros  $\beta_0, \dots, \beta_p$  são independentes. A suposição da distribuição *a priori* normal para  $\boldsymbol{\beta}$  também foi considerada em Souza e Migon (2010).

Nesta dissertação, como usualmente feito, consideramos que, *a priori*,  $\sigma^2$  tem distribuição Gama Inversa com parâmetros  $a$  e  $d$ , denotada por  $\sigma^2 \sim GI(a, d)$ , com a função de densidade dada por:

$$f(\sigma^2|a, d) = \frac{d^a}{\Gamma(a)} (\sigma^2)^{-(a+1)} \exp\left\{-\frac{d}{\sigma^2}\right\} \quad (1.3)$$

onde  $\sigma^2$ ,  $a$  e  $d \in \mathbb{R}^+$ .

Considerando a função de verossimilhança em (1.2) e as distribuições *a priori* eliciadas para os parâmetros de interesse, temos que a distribuição *a posteriori* para  $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2)$  é dada por:

$$\begin{aligned} f(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2|\mathbf{y}) &= \frac{f(\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2)}{f(\mathbf{y})} \\ &= \frac{f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma})f(\boldsymbol{\beta})f(\boldsymbol{\gamma}|\sigma^2)f(\sigma^2)}{\int \int \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma})f(\boldsymbol{\beta})f(\boldsymbol{\gamma}|\sigma^2)f(\sigma^2)d\boldsymbol{\beta}d\sigma^2d\boldsymbol{\gamma}} \\ &= \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} (\pi_{ij})^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \right] \exp\left\{-\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} - \frac{\sum_{i=1}^k \gamma_i^2}{2\sigma^2} - \frac{d}{\sigma^2}\right\} \\ &\quad \left(\frac{1}{\sigma^2}\right)^{\frac{k+2a+2}{2}} [Kf(\mathbf{y})]^{-1}, \end{aligned} \quad (1.4)$$

em que  $\pi_{ij}$  é definido em (1.1) e  $Kf(\mathbf{y})$  é o núcleo da distribuição preditiva, sendo dado por:

$$\begin{aligned} Kf(\mathbf{y}) &= \int \int \int \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} (\pi_{ij})^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \right] \exp\left\{-\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} - \frac{\sum_{i=1}^k \gamma_i^2}{2\sigma^2} - \frac{d}{\sigma^2}\right\} \\ &\quad \left(\frac{1}{\sigma^2}\right)^{\frac{k+2a+2}{2}} d\boldsymbol{\beta}d\sigma^2d\boldsymbol{\gamma}. \end{aligned}$$

Como a distribuição *a posteriori* em (1.4) não possui forma fechada, utilizaremos procedimentos de Monte Carlo via cadeias de Markov (MCMC) para obtermos as estimativas para a tal distribuição. Detalhes sobre estes procedimentos podem ser encontrados em Casella e Robert (2005), Gamerman e Lopes (2006), Gilks *et al.* (1996), Marin e Robert (2007), entre outros. Um



procedimento MCMC que pode ser utilizado para obter estimativas da distribuição *a posteriori* em questão é o *Gibbs Sampler*. Proposto por Gelfand e Smith (1990), esse procedimento requer o conhecimento das distribuições condicionais completas e que seja possível gerar observações através destas distribuições. Com as considerações feitas para as distribuições *a priori* para os parâmetros de interesse e com a função de verossimilhança dada em (1.2), obtemos que a distribuição condicional completa para o vetor de efeitos fixos  $\beta$  é dada por:

$$\begin{aligned}
f(\beta|\mathbf{y}, \gamma, \sigma^2) &= \frac{f(\mathbf{y}, \beta, \gamma, \sigma^2)}{f(\mathbf{y}, \gamma, \sigma^2)} \\
&= \frac{f(\mathbf{y}|\beta, \gamma, \sigma^2)f(\beta)f(\gamma|\sigma^2)f(\sigma^2)}{f(\gamma|\sigma^2)f(\sigma^2) \int f(\mathbf{y}|\beta, \gamma, \sigma^2)f(\beta)d\beta} \\
&\propto f(\mathbf{y}|\beta, \gamma, \sigma^2)f(\beta) \\
&\propto \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp\{\eta_{ij}\}} \right]^{1-y_{ij}} \right] \exp \left\{ -\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} \right\}. \quad (1.5)
\end{aligned}$$

É importante notar que a distribuição condicional completa para o vetor de efeitos fixos  $\beta$  depende apenas da função de verossimilhança em (1.2) e da distribuição *a priori* considerada para  $\beta$ . Sendo assim, a distribuição condicional completa para  $\beta$  continuará a mesma ao assumir qualquer outra distribuição para os efeitos aleatórios  $\gamma$ .

Da mesma forma que para os efeitos fixos  $\beta$ , obtemos que a distribuição condicional completa para o vetor de efeitos aleatórios  $\gamma$  é dada por:

$$\begin{aligned}
f(\gamma|\mathbf{y}, \beta, \sigma^2) &\propto f(\mathbf{y}|\beta, \gamma, \sigma^2)f(\gamma|\sigma^2) \\
&\propto \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp\{\eta_{ij}\}} \right]^{1-y_{ij}} \right] \exp \left\{ -\frac{\sum_{i=1}^k \gamma_i^2}{2\sigma^2} \right\}.
\end{aligned}$$

Para o parâmetro de escala  $\sigma^2$  temos que distribuição condicional completa é

$$\begin{aligned}
f(\sigma^2|\mathbf{y}, \beta, \gamma) &\propto f(\gamma|\sigma^2)f(\sigma^2) \\
&\propto \left( \frac{1}{\sigma^2} \right)^{\frac{k+2a}{2}+1} \exp \left\{ -\frac{1}{\sigma^2} \frac{\sum_{i=1}^k \gamma_i^2 + 2d}{2} \right\}.
\end{aligned}$$

Logo,  $\sigma^2|\mathbf{y}, \beta, \gamma \sim GI \left( \frac{k+2a}{2}, \frac{1}{2} \sum_{i=1}^k \gamma_i^2 + d \right)$ .

Para gerarmos das distribuições condicionais completas dos vetores de parâmetros  $\beta$  e  $\gamma$  podemos utilizar o algoritmo *Adaptive Rejection Sampling* (ARS) proposto por Gilks e Wild (1992) (ver também Dellaportas e Smith (1993)), uma vez que suas respectivas distribuições completas são funções log-côncavas. O ARS é adequado para gerar dados de distribuições univariadas. Entretanto, no caso multivariado, como ocorre para  $\beta$  e para  $\gamma$ , podemos considerar um procedimento de *Gibbs Sampler* para gerar de cada um dos vetores, utilizando o ARS para gerar dados de distribuições univariadas de cada um dos componentes dos vetores de parâmetros em questão. Para o parâmetro  $\sigma^2$  podemos gerar valores diretamente da distribuição condicional completa, pois obtivemos forma fechada para tal distribuição.

## Capítulo 2

# Inferência Bayesiana para o Modelo de Regressão Logística com Efeitos Aleatórios em Classes Estendidas da Distribuição Normal

### 2.1 Introdução

A suposição de normalidade para os efeitos aleatórios pode ser restritiva, pois não considera uma possível multimodalidade e/ou assimetria para os efeitos aleatórios (ver Yang *et al.* (2010)). Se considerarmos uma possível assimetria nos efeitos aleatórios, uma opção razoável para modelarmos seu comportamento é considerarmos a distribuição normal assimétrica.

A distribuição normal assimétrica ou *skew-normal* (brevemente SN) define uma classe paramétrica de distribuições de probabilidade que inclui, como caso especial, a distribuição normal. Esta distribuição consiste de uma multiplicação da densidade da distribuição normal por uma função de distribuição acumulada (FDA). A assimetria é introduzida através da inclusão de um parâmetro de forma em tal FDA. O primeiro tratamento sistemático da classe SN no caso univariado foi dado por Azzalini (1985, 1986) e, posteriormente, Azzalini e Dalla-Valle (1996) introduziram a versão multivariada da densidade da distribuição Normal assimétrica.

Segundo Genton (2004), do ponto de vista de aplicação, a distribuição Normal assimétrica é uma escolha natural em situações em que há alguma assimetria presente, pois uma de suas propriedades particularmente importante é o fato de apresentar uma transição contínua do caso normal para o caso de distribuições assimétricas. Do ponto de vista teórico, a classe SN tem a vantagem de ser matematicamente tratável e de ter um bom número de propriedades em comum com a distribuição normal, como por exemplo, a densidade da distribuição normal assimétrica é unimodal e também tem o suporte no conjunto dos números reais.

Neste capítulo vamos considerar, primeiramente, que os efeitos aleatórios são independentes e possuem distribuição normal assimétrica univariada. Posteriormente consideramos que o vetor de efeitos aleatórios possui distribuição normal assimétrica multivariada. Tal suposição implica em considerarmos que os efeitos aleatórios são correlacionados. Essa afirmação será demonstrada no decorrer deste capítulo. De modo geral, podemos afirmar que uma das principais contribuições desta dissertação é considerarmos a distribuição normal assimétrica, tanto no caso univariado como

no caso multivariado, para modelar o comportamento dos efeitos aleatórios.

Para darmos seguimento, na próxima seção apresentamos as definições e algumas propriedades das distribuições normal assimétrica univariada e multivariada. Para estabelecer notações, ao longo deste trabalho  $\phi_m(y|B, C)$  e  $\Phi_m(y|B, C)$  denotam, respectivamente, as funções de densidade de probabilidade e de distribuição acumulada de uma distribuição normal  $m$ -variada com vetor de média  $B$  e matriz de covariâncias  $C$ , aplicada no ponto  $y$ . Se o vetor de médias é zero,  $B$  será omitida da notação. Se, além disto, a matriz de covariância for a matriz identidade,  $B$  e  $C$  serão omitidas da notação. No caso univariado assumiremos que  $\phi(\cdot)$  e  $\Phi(\cdot)$  denotam as funções de densidade de probabilidade e de distribuição acumulada da distribuição normal padrão, respectivamente.

## 2.2 Definições e Propriedades da Distribuição Normal Assimétrica

A distribuição Normal Assimétrica padrão no caso univariado, introduzida por Azzalini (1985), é definida como segue.

**Definição 2.2.1.** Dizemos que  $Z$  é uma variável aleatória com distribuição Normal Assimétrica padrão com parâmetro de assimetria  $\lambda \in \mathbb{R}$ , se sua função de densidade de probabilidade é dada por:

$$f_Z(z|\lambda) = 2\phi(z)\Phi(z\lambda), \quad -\infty < z < \infty. \quad (2.1)$$

Por brevidade, aqui será denotada por  $Z \sim SN(\lambda)$ .

Note que se  $\lambda = 0$  temos a densidade da distribuição normal padrão. A Figura 2.1 mostra as densidades para algumas distribuições normais assimétricas padrão com diferentes valores de  $\lambda$ .

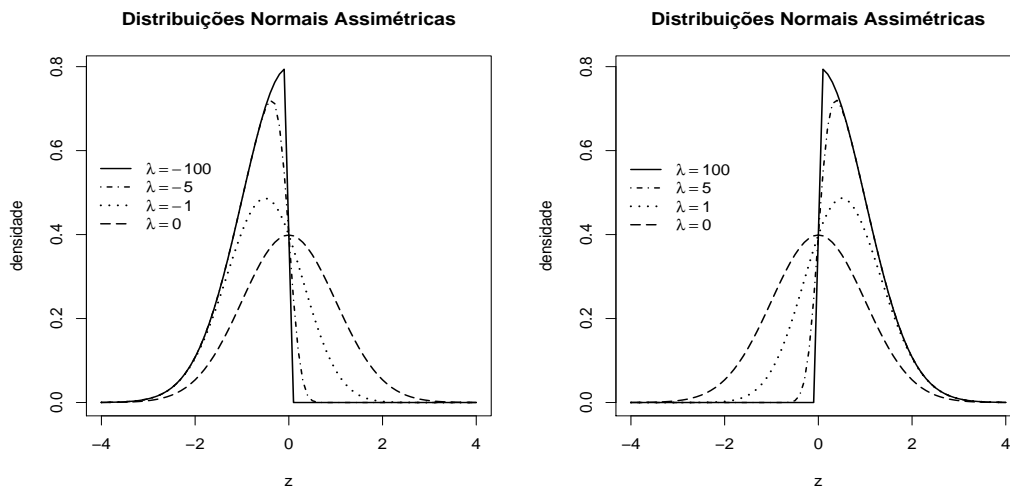


Figura 2.1: Funções de densidade da distribuição normal assimétrica padrão considerando diversos valores para  $\lambda$ .

A média e a variância da distribuição com densidade em (2.1) são dadas, respectivamente, por:

$$E[Z] = \left(\frac{2}{\pi}\right)^{\frac{1}{2}} \delta \text{ e } V[Z] = 1 - \frac{2}{\pi} \delta^2, \quad (2.2)$$

onde  $\delta$  é a seguinte função do parâmetro de assimetria  $\lambda$ :

$$\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}. \quad (2.3)$$

A representação estocástica para  $Z$ , introduzida por Henze (1986) e apresentada no Lema 2.2.1, é um resultado importante tanto do ponto de vista teórico quanto prático. Tal representação pode ser utilizada para gerar amostras da distribuição normal assimétrica e para facilitar a obtenção de alguns resultados analíticos referentes a distribuição  $SN$  como, por exemplo, seus momentos.

**Lema 2.2.1.** (*Representação Estocástica de Henze*) Se  $Y_0$  e  $Y_1$  são variáveis aleatórias independentes e identicamente distribuídas com distribuição  $N(0, 1)$  e  $\delta \in (-1, 1)$ , então

$$Z = \delta|Y_0| + (1 - \delta^2)^{\frac{1}{2}}Y_1$$

tem distribuição normal assimétrica padrão com parâmetro de assimetria  $\lambda = \frac{\delta}{\sqrt{1 - \delta^2}}$ .

Na prática, trabalhamos com a introdução de parâmetros de locação e escala que é feita através da transformação  $Y = \xi + \sigma Z$ , em que  $\xi \in \mathbb{R}$  e  $\sigma \in \mathbb{R}^+$ . A densidade da variável aleatória  $Y$ , com notação dada por  $Y \sim SN(\xi, \sigma^2, \lambda)$ , é escrita como:

$$f_Y(y|\xi, \sigma^2, \lambda) = \frac{2}{\sigma} \phi\left(\frac{y - \xi}{\sigma}\right) \Phi\left(\lambda \left(\frac{y - \xi}{\sigma}\right)\right), \quad -\infty < y < \infty. \quad (2.4)$$

Na Figura 2.2 mostramos as densidades para algumas distribuições normais assimétricas com  $\lambda = 2$  e diferentes parâmetros de locação e escala.

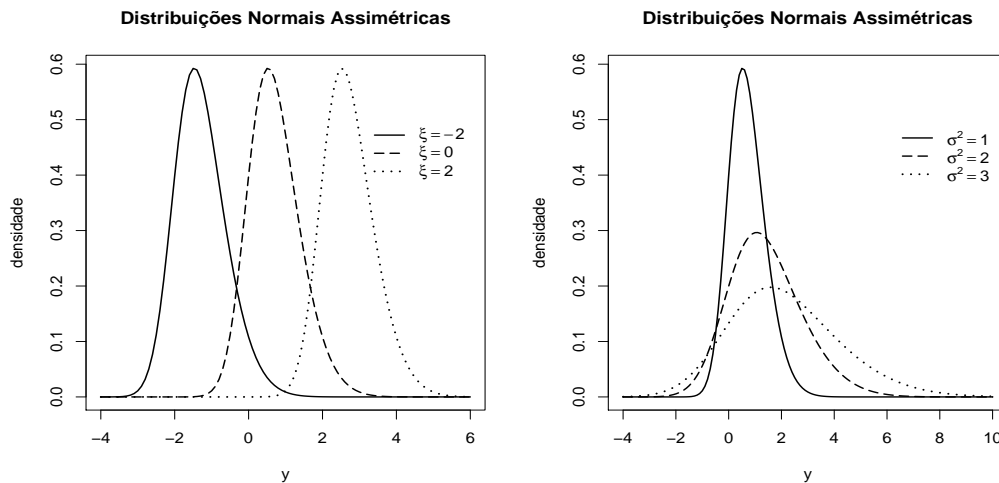


Figura 2.2: Funções de densidade da distribuição normal assimétrica:  $SN(\xi, 1, 2)$  variando  $\xi$  no gráfico da esquerda e  $SN(0, \sigma^2, 2)$  variando  $\sigma^2$  no gráfico da direita.

Para família  $SN$  de locação e escala definida em (2.4) temos que a média e a variância são, respectivamente, dadas por:

$$E[Y] = \xi + \sigma \left( \frac{2}{\pi} \right)^{\frac{1}{2}} \delta \quad \text{e} \quad V[Y] = \sigma^2 \left( 1 - \frac{2}{\pi} \delta^2 \right). \quad (2.5)$$

A distribuição normal assimétrica multivariada é uma extensão para o caso multivariado da distribuição normal assimétrica apresentada anteriormente. Proposta por Azzalini e Dalla-Valle (1996), esta classe de distribuições tem a propriedade de que cada uma de suas distribuições marginais univariadas tem distribuição normal assimétrica univariada, da mesma forma que ocorre com a distribuição normal. Sua definição é apresentada a seguir.

**Definição 2.2.2.** *Dizemos que o vetor aleatório  $\mathbf{Z} = (Z_1, \dots, Z_k)'$  tem distribuição normal assimétrica padrão  $k$ -dimensional, denotada por  $\mathbf{Z} \sim SN_k(\mathbf{\Omega}, \boldsymbol{\alpha})$ , se seus componentes são variáveis aleatórias contínuas e sua função de densidade de probabilidade é:*

$$f_{\mathbf{Z}}(\mathbf{z}|\mathbf{\Omega}, \boldsymbol{\alpha}) = 2\phi_k(z|\mathbf{\Omega}) \Phi(\boldsymbol{\alpha}'z), \quad \mathbf{z} \in \mathbb{R}^k, \quad (2.6)$$

em que  $\boldsymbol{\alpha} \in \mathbb{R}^k$  é um vetor de parâmetros de forma que regulam a assimetria,  $\mathbf{\Omega}$  é uma matriz de escala de ordem  $k \times k$  com características de matriz de correlação.

O vetor de médias e a matriz de variância-covariância da distribuição com densidade em (2.6) são dados, respectivamente, por

$$E[\mathbf{Z}] = \left( \frac{2}{\pi} \right)^{\frac{1}{2}} \boldsymbol{\Delta} \quad \text{e} \quad V[\mathbf{Z}] = \mathbf{\Omega} - \frac{2}{\pi} \boldsymbol{\Delta} \boldsymbol{\Delta}', \quad (2.7)$$

onde  $\boldsymbol{\Delta}$  é a seguinte função do parâmetro de assimetria  $\boldsymbol{\alpha}$ :

$$\boldsymbol{\Delta} = \frac{1}{(1 + \boldsymbol{\alpha}'\mathbf{\Omega}\boldsymbol{\alpha})^{1/2}} \mathbf{\Omega}\boldsymbol{\alpha}. \quad (2.8)$$

A correlação entre dois componentes genéricos do vetor aleatório  $\mathbf{Z} \sim SN_k(\mathbf{\Omega}, \boldsymbol{\alpha})$  é dada pela expressão:

$$\rho_{ij} = \text{Corr}(Z_i, Z_j) = \frac{\omega_{ij} - \frac{2}{\pi} \Delta_i \Delta_j}{\left[ (1 - \frac{2}{\pi} \Delta_i^2)(1 - \frac{2}{\pi} \Delta_j^2) \right]^{1/2}}, \quad i, j = 1, \dots, k, \quad (2.9)$$

em que  $\Delta_i$  é o  $i$ -ésimo componente do vetor  $\boldsymbol{\Delta}$  e  $\omega_{ij}$  é o componente da  $i$ -ésima linha e  $j$ -ésima coluna da matriz  $\mathbf{\Omega}$ .

O Lema 2.2.2 a seguir é uma forma útil de representar a distribuição normal assimétrica multivariada e também pode ser utilizada para gerar amostras de tal distribuição. Este resultado é uma particularização daquele obtido em Arellano-Valle *et al.* (2006) para a família SUN (*unified skew-normal*). Aqui apresentamos apenas o caso específico de interesse neste trabalho e sua prova pode ser encontrada no Apêndice A.

**Lema 2.2.2.** (*Representação Estocástica da normal assimétrica multivariada*)

*Se a variável aleatória  $Y_0$  e o vetor aleatório  $\mathbf{Y}_1$  são independentes de forma que  $Y_0 \sim N(0, 1)$  e  $\mathbf{Y}_1 \sim N_k(\mathbf{0}_k, \boldsymbol{\Psi})$ , em que  $\boldsymbol{\Psi} = \mathbf{\Omega} - c^2 \mathbf{\Omega} \boldsymbol{\alpha} \boldsymbol{\alpha}' \mathbf{\Omega}$ ,  $c = (1 + \boldsymbol{\alpha}' \mathbf{\Omega} \boldsymbol{\alpha})^{-\frac{1}{2}}$  e  $\boldsymbol{\Delta} = c \mathbf{\Omega} \boldsymbol{\alpha}$ , então*

$$\mathbf{Y} = \boldsymbol{\xi} + \boldsymbol{\Delta} |Y_0| + \mathbf{Y}_1$$

*é um vetor aleatório com distribuição  $SN_k(\boldsymbol{\xi}, \mathbf{\Omega}, \boldsymbol{\alpha})$ , em que  $\mathbf{\Omega}$  é uma matriz de escala de ordem  $k \times k$  com características de matriz de correlação.*

Também no caso multivariado, faz-se necessário trabalharmos com a introdução de parâmetros de locação e escala. Sejam  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)' \in \mathbb{R}^k$  e  $\mathbf{S} = \text{diag}\{\sigma_1, \dots, \sigma_k\}$ , de ordem  $k \times k$ , parâmetros de locação e escala, respectivamente. A função de densidade do vetor aleatório  $\mathbf{Y} = \boldsymbol{\xi} + \mathbf{S}\mathbf{Z}$ , em que  $\mathbf{Z} \sim SN_k(\boldsymbol{\Omega}, \boldsymbol{\alpha})$ , é

$$f_{\mathbf{Y}}(\mathbf{y}|\boldsymbol{\xi}, \boldsymbol{\Omega}, S, \lambda) = 2\phi_k(\mathbf{y}|\boldsymbol{\xi}, \mathbf{S}\boldsymbol{\Omega}\mathbf{S}) \Phi(\boldsymbol{\alpha}'\mathbf{S}^{-1}(\mathbf{y} - \boldsymbol{\xi})). \quad (2.10)$$

A distribuição em (2.10) é denominada distribuição normal assimétrica multivariada com locação e escala e será denotada por  $\mathbf{Y} \sim SN_k(\boldsymbol{\xi}, \mathbf{S}\boldsymbol{\Omega}\mathbf{S}, \boldsymbol{\alpha})$ .

Azzalini e Dalla-Valle (1996) também mostra que a distribuição marginal de cada componente de um vetor aleatório com distribuição normal assimétrica multivariada é a distribuição normal assimétrica univariada. Na Proposição 2.2.1 a seguir, apresentamos este resultado para o caso particular em que a matriz de escala é  $\sigma^2 I_k$  e o vetor de parâmetros de assimetria é  $\boldsymbol{\alpha} \mathbf{1}_k$ . Este caso particular será utilizado no decorrer do texto. A demonstração da Proposição 2.2.1 pode ser encontrada no Apêndice A.

**Proposição 2.2.1.** *Assuma que  $\boldsymbol{\gamma} \sim SN_k(\boldsymbol{\xi}, \sigma^2 I_k, \boldsymbol{\alpha} \mathbf{1}_k)$ , em que  $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_k)'$ ,  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)'$ ,  $\sigma^2 \in \mathbb{R}^+$  e  $\boldsymbol{\xi} \in \mathbb{R}$ . A distribuição marginal do componente  $\gamma_j$  do vetor  $\boldsymbol{\gamma}$  é  $SN(\xi_j, \sigma^2, \alpha^*)$ , com  $\alpha^* = \alpha/\sqrt{1 + (k-1)\alpha^2}$  em que  $-(k-1)^{-\frac{1}{2}} < \alpha^* < (k-1)^{-\frac{1}{2}}$ .*

## 2.3 Modelo Logístico com Efeitos Aleatórios SN Independentes

Consideremos a função de verossimilhança em (1.2). Estendendo trabalhos anteriores assuma que os efeitos aleatórios  $\gamma_1, \dots, \gamma_k$ , condicional a  $\sigma^2$  e  $\lambda$ , sejam independentes e identicamente distribuídos com  $\gamma_i|\sigma^2, \lambda \sim SN(-\delta\sigma\sqrt{2/\pi}, \sigma^2, \lambda)$ ,  $\sigma^2 \in \mathbb{R}^+$ ,  $\lambda \in \mathbb{R}$  e em que  $\delta$  é definido como em (2.3). Tal suposição quanto ao parâmetro de locação da distribuição dos efeitos aleatórios é necessária para que tenhamos  $E(\gamma_i) = 0$ ,  $\forall i = 1, \dots, k$ . Este fato evita o problema de falta de identificabilidade na estimação do parâmetro de intercepto  $\beta_0$  e do vetor de efeitos aleatórios  $\boldsymbol{\gamma}$ . Assuma também que  $\lambda \sim N(h, \tau^2)$ ,  $h \in \mathbb{R}$  e  $\tau^2 \in \mathbb{R}^+$  conhecidos. A distribuição *a priori* normal para os parâmetro de assimetria  $\lambda$  foi utilizada anteriormente em Arellano-Valle *et al.* (2009). Considerando as mesmas suposições sobre os efeitos fixos  $\boldsymbol{\beta}$  e sobre o parâmetro de escala  $\sigma^2$  que utilizamos na Seção 1.3, temos que a distribuição *a posteriori* conjunta para  $(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \lambda)$  é dada por:

$$\begin{aligned} f(\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \lambda|\mathbf{y}) &= \frac{f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma})f(\boldsymbol{\beta})f(\boldsymbol{\gamma}|\sigma^2, \lambda)f(\sigma^2)f(\lambda)}{\int \int \int \int f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma})f(\boldsymbol{\beta})f(\boldsymbol{\gamma}|\sigma^2, \lambda)f(\sigma^2)f(\lambda)d\lambda d\boldsymbol{\beta} d\sigma^2 d\boldsymbol{\gamma}} \\ &= \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} (\pi_{ij})^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \right] \left( \frac{1}{\sigma^2} \right)^{\frac{k+2a+2}{2}} \Phi_k \left( \frac{\lambda}{\sigma} \left( \boldsymbol{\gamma} + \mathbf{1}_k \delta \sigma \sqrt{\frac{2}{\pi}} \right) \right) \\ &\quad \exp \left\{ -\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} - \frac{\sum_{i=1}^k \left( \gamma_i + \delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{(\lambda - h)^2}{2\tau^2} - \frac{d}{\sigma^2} \right\} [Kf(\mathbf{y})]^{-1}, \end{aligned} \quad (2.11)$$

em que  $\delta$  é uma função de  $\lambda$  definida em (2.3),  $\pi_{ij}$  é definido em (1.1) e  $Kf(\mathbf{y})$  é o núcleo da distribuição preditiva, sendo dado por:

$$Kf(\mathbf{y}) = \int \int \int \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} (\pi_{ij})^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \right] \left( \frac{1}{\sigma^2} \right)^{\frac{k+2a+2}{2}} \Phi_k \left( \frac{\lambda}{\sigma} \left( \boldsymbol{\gamma} + \mathbf{1}_k \delta \sigma \sqrt{\frac{2}{\pi}} \right) \right)$$

$$\exp \left\{ -\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} - \frac{\sum_{i=1}^k \left( \gamma_i + \delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{(\lambda - h)^2}{2\tau^2} - \frac{d}{\sigma^2} \right\} d\beta d\sigma^2 d\gamma.$$

Como a distribuição em (2.11) não possui forma fechada, utilizaremos procedimentos MCMC para obtermos estimativas para distribuições *a posteriori* de interesse. O procedimento utilizado será *Gibbs Sampler*. Para utilizarmos tal procedimento é necessário encontrarmos as distribuições condicionais completas para cada um dos parâmetros de interesse. As distribuições condicionais completas para o caso em que trabalhamos diretamente com a densidade em (2.4) para os efeitos aleatórios  $\gamma_i$ 's são apresentadas na Seção B.1 do Apêndice B, no entanto esta abordagem não apresentou resultados computacionais satisfatórios. Então optamos por utilizar o Lema 2.2.1 para especificarmos a distribuição de  $\gamma_i|\sigma^2, \lambda$ . Utilizando este lema podemos escrever  $\Gamma_i \sim SN(\lambda)$  como

$$\Gamma_i = \delta|U_i| + (1 - \delta^2)^{\frac{1}{2}}V_i,$$

em que  $U_i$  e  $V_i$  são variáveis aleatórias com distribuição normal padrão e  $\delta$  é definido como em (2.3). Defina por  $\psi_i = |U_i|$ . Através do método do Jacobiano encontramos que a variável aleatória  $\psi_i$  tem distribuição *half-normal* com parâmetros de locação e escala iguais a 0 e a 1, respectivamente. Para a distribuição *half-normal* adotamos a notação  $\psi_i \sim HN(\theta_1, \theta_2)$  para a distribuição com função densidade de probabilidade é dada por

$$f(\psi_i) = \frac{1}{\sqrt{2\pi}\theta_2} \exp \left\{ -\frac{(\psi_i - \theta_1)^2}{2\theta_2} \right\} \mathbb{I}_{(0, \infty)}(\psi_i).$$

Para especificar  $\gamma_i|\sigma^2, \lambda \sim SN(-\delta\sigma\sqrt{2/\pi}, \sigma^2, \lambda)$  podemos utilizar uma transformação linear e escrever:

$$\begin{aligned} \gamma_i &= -\delta\sigma\sqrt{2/\pi} + \sigma\Gamma_i \\ &= -\delta\sigma\sqrt{2/\pi} + \sigma \left( \delta\psi_i + (1 - \delta^2)^{\frac{1}{2}}V_i \right) \\ &= \delta\sigma \left( \psi_i - \sqrt{2/\pi} \right) + \sigma (1 - \delta^2)^{\frac{1}{2}} V_i. \end{aligned}$$

Assim, sabendo que  $V_i$  tem distribuição normal padrão e condicionando em  $\psi_i$ , conseguimos escrever a distribuição desejada especificando  $\gamma_i|\psi_i, \sigma^2, \lambda \sim N(\sigma\delta(\psi_i - \sqrt{2/\pi}), \sigma^2(1 - \delta^2))$ , em que  $\delta$  é definido como em (2.3), e também  $\psi_i \sim HN(0, 1)$ .

Com as considerações feitas, obtemos as distribuições condicionais completas apresentadas a seguir. A distribuição condicional completa  $f(\boldsymbol{\beta}|\mathbf{y}, \boldsymbol{\gamma}, \boldsymbol{\psi}, \sigma^2, \lambda)$  para o vetor de efeitos fixos  $\boldsymbol{\beta}$  é a mesma obtida em (1.5), uma vez que não depende da suposição da distribuição para os efeitos aleatórios. As distribuições condicionais completas para o vetor de efeitos aleatórios  $\boldsymbol{\gamma}$ , para o componente  $\boldsymbol{\psi} = (\psi_1, \dots, \psi_k)'$ , que aparece no modelo ao utilizar o Lema 2.2.1 para especificar a distribuição normal assimétrica, e para os parâmetros de assimetria  $\lambda$  e escala  $\sigma^2$  são dadas, respectivamente, por:

$$\begin{aligned} f(\boldsymbol{\gamma}|\mathbf{y}, \boldsymbol{\beta}, \sigma^2, \boldsymbol{\psi}, \lambda) &\propto f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\psi}, \lambda) f(\boldsymbol{\gamma}|\sigma^2, \boldsymbol{\psi}, \lambda) \\ &\propto \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp\{\eta_{ij}\}} \right]^{1-y_{ij}} \right] \\ &\quad \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i - \sigma\delta(\psi_i - \sqrt{2/\pi}) \right)^2}{2\sigma^2(1 - \delta^2)} \right\}, \end{aligned}$$

$$\begin{aligned}
f(\boldsymbol{\psi}|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \lambda) &\propto f(\boldsymbol{\gamma}|\sigma^2, \boldsymbol{\psi}, \lambda)f(\boldsymbol{\psi}) \\
&\propto \prod_{i=1}^k f(\gamma_i|\sigma^2, \psi_i, \lambda)f(\psi_i) \\
&\propto \prod_{i=1}^k \exp \left\{ -\frac{(\gamma_i - \sigma\delta(\psi_i - \sqrt{2/\pi}))^2}{2\sigma^2(1-\delta^2)} - \frac{\psi_i^2}{2} \right\} \mathbb{I}_{(0,\infty)}(\psi_i) \\
&\propto \prod_{i=1}^k \exp \left\{ -\frac{(\psi_i - (\gamma_i + \sigma\delta\sqrt{2/\pi})\frac{\delta}{\sigma})^2}{2(1-\delta^2)} \right\} \exp \left\{ -\frac{(\gamma_i + \sigma\delta\sqrt{2/\pi})^2}{2\sigma^2} \right\} \mathbb{I}_{(0,\infty)}(\psi_i) \\
&\propto \prod_{i=1}^k \exp \left\{ -\frac{(\psi_i - (\gamma_i + \sigma\delta\sqrt{2/\pi})\frac{\delta}{\sigma})^2}{2(1-\delta^2)} \right\} \mathbb{I}_{(0,\infty)}(\psi_i), \tag{2.12}
\end{aligned}$$

$$\begin{aligned}
f(\lambda|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \boldsymbol{\psi}) &\propto f(\boldsymbol{\gamma}|\sigma^2, \boldsymbol{\psi}, \lambda)f(\lambda) \\
&\propto (1-\delta^2)^{-\frac{k}{2}} \exp \left\{ -\frac{\sum_{i=1}^n (\gamma_i - \sigma\delta(\psi_i - \sqrt{2/\pi}))^2}{2\sigma^2(1-\delta^2)} - \frac{(\lambda-h)^2}{2\tau^2} \right\} \\
&\propto (1+\lambda^2)^{\frac{k}{2}} \exp \left\{ -\frac{(1+\lambda^2)}{2\sigma^2} \sum_{i=1}^n \left( \gamma_i - \sigma \frac{\lambda}{\sqrt{1+\lambda^2}} (\psi_i - \sqrt{2/\pi}) \right)^2 - \frac{(\lambda-h)^2}{2\tau^2} \right\},
\end{aligned}$$

$$\begin{aligned}
f(\sigma^2|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\psi}, \lambda) &\propto f(\boldsymbol{\gamma}|\sigma^2, \boldsymbol{\psi}, \lambda)f(\sigma^2) \\
&\propto \left( \frac{1}{\sigma^2} \right)^{\frac{k}{2}+a+1} \exp \left\{ -\frac{\sum_{i=1}^n (\gamma_i - \sigma\delta(\psi_i - \sqrt{2/\pi}))^2}{2\sigma^2(1-\delta^2)} - \frac{d}{\sigma^2} \right\}.
\end{aligned}$$

Note de (2.12) que a distribuição condicional completa para  $\Psi_i$  possui núcleo conhecido e, assim, identificamos que  $\psi_i|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \lambda \sim HN \left( (\gamma_i + \sigma\delta\sqrt{2/\pi})\frac{\delta}{\sigma}, (1-\delta^2) \right)$ .

Por fim, é preciso gerar amostras das distribuições condicionais completas obtidas. Assim, como nas situações onde assumimos a distribuição normal para os efeitos aleatórios, podemos utilizar o método ARS para gerar das distribuições de  $\boldsymbol{\beta}$  e  $\boldsymbol{\gamma}$ , pois após alguns cálculos consegue-se provar suas distribuições são funções log-côncavas. Para gerarmos de  $\boldsymbol{\psi}$ , podemos gerar cada  $\psi_i$  quase que diretamente utilizando o método da rejeição. Amostras das distribuições *a posteriori* de  $\lambda$  e de  $\sigma^2$  são obtidas gerando das respectivas distribuições condicionais completas usando o algoritmo *Adaptive Rejection Metropolis Sampling* (ARMS) proposto por Gilks *et al.* (1995).



## 2.4 Modelo Logístico com Efeitos Aleatórios SN Correlacionados

Mais uma vez estendendo trabalhos anteriores, vamos considerar agora que o vetor de efeitos aleatórios  $\gamma$  tem distribuição normal assimétrica multivariada com vetor de parâmetros de locação  $\xi = -\Delta\sigma\sqrt{2/\pi}$ , matriz de escala  $\sigma^2 I_k$ ,  $\sigma^2 \in \mathbb{R}^+$ , e vetor de parâmetros de assimetria igual a  $\alpha$ ,  $\alpha \in \mathbb{R}$ , ou seja,  $\gamma \sim SN_k(-\Delta\sigma\sqrt{2/\pi}, \sigma^2 I_k, \mathbf{1}_k \alpha)$ , em que  $\mathbf{1}_k$  é o vetor unitário de ordem  $k$  e  $\Delta = \Delta \mathbf{1}_k = (\alpha/\sqrt{1+k\alpha^2})\mathbf{1}_k$  é definido em (2.8). Esta suposição quanto ao parâmetro de locação é feita em função de garantir que  $E(\gamma) = \mathbf{0}_k$ . Isto é necessário para evitar problemas de identificabilidade na estimação do parâmetro de intercepto  $\beta_0$  e dos efeitos aleatórios  $\gamma$ . Assumimos a função de verossimilhança em (1.2) e consideramos também que  $\alpha \sim N(h, \tau^2)$ , com  $h \in \mathbb{R}$  e  $\tau^2 \in \mathbb{R}^+$  conhecidos, suposição para o parâmetro assimetria igual a da seção 2.3. Considerando as mesmas suposições sobre os efeitos fixos  $\beta$  e sobre o parâmetro de escala  $\sigma^2$  que foram consideradas na Seção 1.3, temos que a distribuição *a posteriori* conjunta para  $(\beta, \gamma, \sigma^2, \alpha)$  é dada por

$$\begin{aligned} f(\beta, \gamma, \sigma^2, \alpha | \mathbf{y}) &= \frac{f(\mathbf{y} | \beta, \gamma) f(\beta) f(\gamma | \sigma^2, \alpha) f(\sigma^2) f(\alpha)}{\int \int \int \int f(\mathbf{y} | \beta, \gamma) f(\beta) f(\gamma | \sigma^2, \alpha) f(\sigma^2) f(\alpha) d\alpha d\beta d\sigma^2 d\gamma} \\ &= \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} (\pi_{ij})^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \right] \left( \frac{1}{\sigma^2} \right)^{\frac{k+2a+2}{2}} \Phi \left( \frac{\alpha}{\sigma} \sum_{i=1}^k \left( \gamma_i + \Delta\sigma\sqrt{\frac{2}{\pi}} \right) \right) \\ &\quad \exp \left\{ -\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} - \frac{\sum_{i=1}^k \left( \gamma_i + \Delta\sigma\sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{(\alpha - h)^2}{2\tau^2} - \frac{d}{\sigma^2} \right\} [Kf(\mathbf{y})]^{-1} \end{aligned} \quad (2.13)$$

em que  $\pi_{ij}$  é definido em (1.1) e  $Kf(\mathbf{y})$  é o núcleo da distribuição preditiva, sendo dado por

$$\begin{aligned} Kf(\mathbf{y}) &= \int \int \int \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} (\pi_{ij})^{y_{ij}} (1 - \pi_{ij})^{1-y_{ij}} \right] \left( \frac{1}{\sigma^2} \right)^{\frac{k+2a+2}{2}} \Phi \left( \frac{\alpha}{\sigma} \sum_{i=1}^k \left( \gamma_i + \Delta\sigma\sqrt{\frac{2}{\pi}} \right) \right) \\ &\quad \exp \left\{ -\frac{\sum_{i=0}^p (\beta_i - m_i)^2}{2b^2} - \frac{\sum_{i=1}^k \left( \gamma_i + \Delta\sigma\sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{(\alpha - h)^2}{2\tau^2} - \frac{d}{\sigma^2} \right\} d\beta d\sigma^2 d\gamma. \end{aligned}$$

A distribuição *a posteriori* em (2.13) também não possui forma fechada e, novamente, faz-se necessário implementar procedimentos MCMC para obtermos estimativas para os parâmetros do modelo em questão. Da mesma forma que na Seção 2.3, vamos utilizar uma representação estocástica para especificar a distribuição dos efeitos aleatórios. As distribuições condicionais completas para os parâmetros do modelo ao trabalhar diretamente com a densidade da distribuição normal assimétrica multivariada, e não utilizar representação estocástica, são apresentadas na Seção B.2 do Apêndice B. Para especificarmos a distribuição de  $\gamma | \sigma^2, \alpha$ , utilizamos o Lema 2.2.2. Inicialmente, através deste lema escrevemos a distribuição de  $\mathbf{\Gamma} \sim SN_k(\mathbf{0}_k, I_k, \alpha \mathbf{1}_k)$  como

$$\mathbf{\Gamma} = \mathbf{\Delta} |U| + V,$$

em que  $\mathbf{\Delta} = \alpha (1 + k\alpha^2)^{-\frac{1}{2}}$ ,  $U$  é uma variável aleatória com distribuição normal padrão e  $V$  é um vetor aleatório com distribuição  $N(\mathbf{0}_k, \mathbf{\Psi})$  com  $\mathbf{\Psi} = I_k - (1 + k\alpha^2)^{-1} \alpha^2 \mathbf{1}_k \mathbf{1}_k'$ . Se definimos  $T = |U|$ , através do método do Jacobiano sabemos que  $T \sim HN(0, 1)$ . Para introduzir parâmetros

de locação e escala utilizamos uma transformação linear como dito na Seção 2.2. Então, para especificar  $\boldsymbol{\gamma} \sim SN_k(-\boldsymbol{\Delta}\sigma\sqrt{2/\pi}, \sigma^2 I_k, \mathbf{1}_k \alpha)$  fazemos

$$\begin{aligned}\boldsymbol{\gamma} &= -\boldsymbol{\Delta}\sigma\sqrt{2/\pi} + \text{diag}\{\sigma, \dots, \sigma\}\boldsymbol{\Gamma} \\ &= -\boldsymbol{\Delta}\sigma\sqrt{2/\pi} + \sigma(\boldsymbol{\Delta}T + V) \\ &= \boldsymbol{\Delta}\sigma\left(T - \sqrt{2/\pi}\right) + \sigma V.\end{aligned}$$

Desta forma, se condicionarmos no conhecimento de  $T$  podemos escrever a distribuição desejada especificando  $\boldsymbol{\gamma}|T, \sigma^2, \alpha \sim N_k(\sigma\boldsymbol{\Delta}(T - \sqrt{2/\pi}), \sigma^2(I_k - \Delta^2\mathbf{1}_k\mathbf{1}'_k))$  e  $T \sim HN(0, 1)$ , em que  $\boldsymbol{\Delta} = \Delta\mathbf{1}_k = (\alpha/\sqrt{1+k\alpha^2})\mathbf{1}_k$  é definido como em (2.8). Com as considerações feitas, obtemos as distribuições condicionais completas apresentadas a seguir. A distribuição condicional completa  $f(\boldsymbol{\beta}|\mathbf{y}, \boldsymbol{\gamma}, T, \sigma^2, \alpha)$  para o vetor de efeitos fixos  $\boldsymbol{\beta}$  é a mesma obtida em (1.5), uma vez que não depende da suposição da distribuição para os efeitos aleatórios. As distribuições condicionais completas para o vetor de efeitos aleatórios  $\boldsymbol{\gamma}$ , para o componente  $T$ , que aparece no modelo ao utilizar o Lema 2.2.2 para especificar a distribuição normal assimétrica, e para os parâmetros de assimetria  $\alpha$  e escala  $\sigma^2$  são dadas, respectivamente, por

$$\begin{aligned}f(\boldsymbol{\gamma}|\mathbf{y}, \boldsymbol{\beta}, \sigma^2, T, \alpha) &\propto f(\mathbf{y}|\boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, T, \alpha)f(\boldsymbol{\gamma}|\sigma^2, T, \alpha) \\ &\propto \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp\{\eta_{ij}\}} \right]^{1-y_{ij}} \right] \\ &\quad \exp \left\{ -\frac{1}{2\sigma^2} \left[ \boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left( T - \sqrt{2/\pi} \right) \right]' \left[ I_k + \frac{\Delta^2}{1 - k\Delta^2} \mathbf{1}_k \mathbf{1}'_k \right] \left[ \boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left( T - \sqrt{2/\pi} \right) \right] \right\},\end{aligned}$$

$$\begin{aligned}f(T|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \alpha) &\propto f(\boldsymbol{\gamma}|\sigma^2, T, \alpha)f(T) \\ &\propto \exp \left\{ -\frac{1}{2\sigma^2} \left[ \boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left( T - \sqrt{2/\pi} \right) \right]' \left[ I_k + \frac{\Delta^2}{1 - k\Delta^2} \mathbf{1}_k \mathbf{1}'_k \right] \left[ \boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left( T - \sqrt{2/\pi} \right) \right] \right\} \\ &\quad \exp \left\{ -\frac{T^2}{2} \right\} \mathbb{I}_{(0, \infty)}(T) \\ &\propto \exp \left\{ -\frac{1}{2} \left[ T^2(1 - k\Delta^2)^{-1} - 2T \frac{\sigma\Delta}{\sigma^2(1 - k\Delta^2)} \left( \sum_{i=1}^k \gamma_i + k\sigma\Delta^2\sqrt{2/\pi} \right) \right] \right\} \mathbb{I}_{(0, \infty)}(T) \\ &\propto \exp \left\{ -\frac{1}{2(1 - k\Delta^2)} \left[ T^2 - 2T \left( \frac{\Delta}{\sigma} \sum_{i=1}^k \gamma_i + k\Delta^2\sqrt{2/\pi} \right) \right] \right\} \mathbb{I}_{(0, \infty)}(T),\end{aligned}\tag{2.14}$$

$$\begin{aligned}f(\alpha|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, T) &\propto f(\boldsymbol{\gamma}|\sigma^2, T, \alpha)f(\alpha) \\ &\propto (1 - k\Delta^2)^{-\frac{1}{2}} \exp \left\{ -\frac{(\alpha - h)^2}{2\tau^2} \right\} \\ &\quad \exp \left\{ -\frac{1}{2\sigma^2} \left[ \boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left( T - \sqrt{2/\pi} \right) \right]' \left[ I_k + \frac{\Delta^2}{1 - k\Delta^2} \mathbf{1}_k \mathbf{1}'_k \right] \left[ \boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left( T - \sqrt{2/\pi} \right) \right] \right\} \\ &\propto (1 + k\alpha^2)^{\frac{1}{2}} \exp \left\{ -\frac{(\alpha - h)^2}{2\tau^2} \right\} \\ &\quad \exp \left\{ -\frac{1}{2\sigma^2} \left[ \boldsymbol{\gamma} - \frac{\sigma\alpha(T - \sqrt{2/\pi})}{\sqrt{1 + k\alpha^2}} \mathbf{1}_k \right]' \left[ I_k + \alpha^2 \mathbf{1}_k \mathbf{1}'_k \right] \left[ \boldsymbol{\gamma} - \frac{\sigma\alpha(T - \sqrt{2/\pi})}{\sqrt{1 + k\alpha^2}} \mathbf{1}_k \right] \right\},\end{aligned}$$

$$\begin{aligned}
f(\sigma^2|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, T, \alpha) &\propto f(\boldsymbol{\gamma}|\sigma^2, T, \alpha)f(\sigma^2) \\
&\propto \left(\frac{1}{\sigma^2}\right)^{\frac{k}{2}+a+1} \exp\left\{-\frac{d}{\sigma^2}\right\} \\
&\quad \exp\left\{-\frac{1}{2\sigma^2} \left[\boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left(T - \sqrt{2/\pi}\right)\right]' \left[I_k + \frac{\Delta^2}{1 - k\Delta^2} \mathbf{1}_k \mathbf{1}_k'\right] \left[\boldsymbol{\gamma} - \sigma\boldsymbol{\Delta} \left(T - \sqrt{2/\pi}\right)\right]\right\}.
\end{aligned}$$

Note de (2.14) que a distribuição condicional completa para  $T$  possui núcleo conhecido e, assim, identificamos que  $T|\mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \alpha \sim HN\left(\frac{\Delta}{\sigma} \left[\sum_{i=1}^k \gamma_i + k\sigma\Delta\sqrt{2/\pi}\right], (1 - k\Delta^2)\right)$ .

Para implementar o modelo é necessário gerarmos amostras das distribuições condicionais completas obtidas. Da mesma forma que nos modelos anteriores, podemos utilizar o método ARS para gerarmos das distribuições de  $\boldsymbol{\beta}$  e  $\boldsymbol{\gamma}$ , pois suas distribuições são funções log-côncavas. A geração de  $T$  pode ser feita quase que diretamente utilizando o método da rejeição. Amostras das distribuições *a posteriori* de  $\alpha$  e de  $\sigma^2$  são obtidas gerando das respectivas distribuições condicionais completas usando o método ARMS.

# Capítulo 3

## Razão de Chances no Modelo Logístico Misto

A interpretação dos parâmetros no modelo logístico é usualmente feita utilizando-se a razão de chances ou “*Odds Ratio*”.

A *Odds Ratio* entre dois grupos é a razão entre a “chance” de ocorrência do evento de interesse em um grupo pela “chance” de ocorrência do evento no outro grupo. Por chance de ocorrência do evento em um grupo particular, entende-se a razão entre as probabilidades de ocorrência e não ocorrência deste evento em tal grupo. Ou seja, a razão de chances (brevemente, *OR*) entre dois indivíduos  $j_1$  e  $j_2$  pertencentes aos *clusters*  $i_1$  e  $i_2$ , respectivamente, é definida como

$$OR = \frac{\pi_{i_1 j_1} / (1 - \pi_{i_1 j_1})}{\pi_{i_2 j_2} / (1 - \pi_{i_2 j_2})}. \quad (3.1)$$

No caso do modelo logístico com efeitos aleatórios, a probabilidade  $\pi_{ij}$  é dada na Seção 2.2. Neste caso, como mostrado em Larsen *et al.* (2000), há uma diversidade de comparações possíveis de serem feitas, dependendo o interesse que se tem no problema. Além disto, contrário ao que ocorre no modelo logístico usual, a *Odds Ratio* depende não só dos efeitos fixos, como pode também depender dos efeitos aleatórios. Na situação que acabamos de descrever, a *OR* torna-se

$$\begin{aligned} OR &= \frac{\exp \left\{ \mathbf{x}'_{i_1 j_1} \boldsymbol{\beta} + \mathbf{z}'_{i_1 j_1} \boldsymbol{\gamma} \right\}}{\exp \left\{ \mathbf{x}'_{i_2 j_2} \boldsymbol{\beta} + \mathbf{z}'_{i_2 j_2} \boldsymbol{\gamma} \right\}} \\ &= \exp \left\{ (\mathbf{x}'_{i_1 j_1} - \mathbf{x}'_{i_2 j_2}) \boldsymbol{\beta} + (\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2}) \boldsymbol{\gamma} \right\}. \end{aligned} \quad (3.2)$$

Se considerarmos a situação na qual os dois indivíduos pertencem a um mesmo *cluster*, mas apresentam resultados distintos para as covariáveis, temos que a interpretação usual de razão de chances é aplicável pois, neste caso, a *OR* depende apenas dos efeitos fixos. Para um melhor entendimento, assumamos que os indivíduos  $j_1$  e  $j_2$  pertencentes ao *cluster*  $i_1$  apresentem valores distintos para as covariáveis. Neste caso, os vetores  $\mathbf{z}'_{i_1 j_1}$  e  $\mathbf{z}'_{i_1 j_2}$  são idênticos. Desta forma, a expressão para razão de chances em (3.1) é simplificada e assume a seguinte forma

$$OR = \exp \left\{ (\mathbf{x}'_{i_1 j_1} - \mathbf{x}'_{i_2 j_2}) \boldsymbol{\beta} \right\}. \quad (3.3)$$

Para quantificarmos os efeitos aleatórios, consideramos indivíduos com os mesmos valores de covariáveis para os efeitos fixos. Considere, por exemplo, indivíduos  $j_1$  e  $j_2$  com os mesmos valores

observados para as covariáveis, mas que pertencem a diferentes *clusters*  $i_1$  e  $i_2$ , respectivamente. Neste caso, os vetores  $\mathbf{x}'_{i_1j_1}$  e  $\mathbf{x}'_{i_2j_2}$  são iguais e a  $OR$  dependerá apenas dos efeitos aleatórios. Sendo assim, a razão de chances em (3.1), é dada por

$$OR = \exp \{(\mathbf{z}'_{i_1j_1} - \mathbf{z}'_{i_2j_2})\boldsymbol{\gamma}\}. \quad (3.4)$$

Em qualquer das três situações exibidas em (3.2), (3.3) e (3.4), do ponto de vista bayesiano, a  $OR$  é uma quantidade aleatória. Do ponto de vista clássico, no entanto, a  $OR$  é um objeto aleatório apenas nos casos em que é função dos efeitos aleatórios  $\boldsymbol{\gamma}$  como são os casos exibidos nas expressões (3.2) e (3.4).

Uma vez que a razão de chances é uma quantidade aleatória, Larsen *et al.* (2000) propõe estudá-la em termos de características relacionadas a sua distribuição de probabilidades. A proposta de Larsen *et al.* (2000) é considerarmos, para este fim, a Razão de Chances Mediana ( $MOR$ ) que quantifica apropriadamente a heterogeneidade entre os *clusters* ou grupos, permitindo uma avaliação mais apropriada dos efeitos fixos. No caso em que o interesse está em fazermos comparações entre um indivíduo com alto risco e um indivíduo com baixo risco, aleatoriamente escolhidos em grupos diferentes e com os mesmos valores observados para as covariáveis, tal medida é dada por

$$MOR = med\{\exp \{ |(\mathbf{z}'_{i_1j_1} - \mathbf{z}'_{i_2j_2})\boldsymbol{\gamma}| \}\}. \quad (3.5)$$

Para o caso onde indivíduos pertencem a diferentes *clusters* e também apresentam diferentes valores observados para as covariáveis, a  $MOR$  torna-se

$$MOR = med\{\exp \{(\mathbf{x}'_{i_1j_1} - \mathbf{x}'_{i_2j_2})\boldsymbol{\beta} + (\mathbf{z}'_{i_1j_1} - \mathbf{z}'_{i_2j_2})\boldsymbol{\gamma}\}\}. \quad (3.6)$$

A tarefa de eliciarmos distribuições *a priori* apropriadas para os parâmetros torna-se mais simples quando estes são interpretáveis. No caso dos efeitos aleatórios esta interpretação, geralmente, não é trivial, o que dificulta a construção de distribuições *a priori* razoáveis para descrevermos o seu comportamento. Assim, sua influência na quantificação da  $OR$ , como mostrado por Larsen *et al.* (2000), pode se tornar uma ferramenta valiosa na construção de distribuições *a priori* mais apropriadas. Como, *a priori*, estamos considerando a mesma distribuição para os efeitos fixos, os modelos adotados diferem por eliciarmos distribuições *a priori* diferentes para os efeitos aleatórios  $\boldsymbol{\gamma}$ . Assim, neste capítulo, propomos estudar o efeito inicialmente introduzido na  $OR$  por estas distribuições.

Iniciamos, apresentando os resultados obtidos por Larsen *et al.* (2000) para o caso em que os efeitos aleatórios são normalmente distribuídos. Posteriormente, nas Seções 3.2 e 3.3 fornecemos extensões originais destes resultados para os casos em que os efeitos aleatórios são independentes e correlacionados, respectivamente, e têm distribuições normal assimétrica.

### 3.1 Efeitos Aleatórios Normalmente Distribuídos

O artigo de Larsen *et al.* (2000) é o primeiro a fornecer uma interpretação para os efeitos fixos quando o modelo logístico com efeitos aleatórios é adotado. Apesar destes autores considerarem tal interpretação do ponto de vista da Estatística Clássica, os resultados obtidos por eles, fornecem o comportamento, *a priori*, da *odds ratio*, dados os efeitos fixos  $\boldsymbol{\beta}$  e os hiperparâmetros que indexam a distribuição dos efeitos aleatórios.

Larsen *et al.* (2000) mostrou que, se os efeitos aleatórios são normalmente distribuídos com vetor de médias  $\mathbf{0}$  e matriz de covariância  $\boldsymbol{\Sigma}$ , então a razão de chances, em (3.1), entre dois indivíduos

$j_1$  e  $j_2$  pertencentes, respectivamente, aos *clusters*  $i_1$  e  $i_2$ , pode ser escrita como:

$$OR = \exp\{k + \omega V\}, \quad (3.7)$$

em que  $k = (\mathbf{x}'_{i_1 j_1} - \mathbf{x}'_{i_2 j_2})\boldsymbol{\beta}$ ,  $\omega^2 = (\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})\boldsymbol{\Sigma}(\mathbf{z}_{i_1 j_1} - \mathbf{z}_{i_2 j_2})$  e  $V \sim N(0, 1)$ .

Tal expressão é válida pois, visto que o termo  $k$  é um escalar, aplicando o Teorema C.1.2, apresentado no Apêndice C.1, podemos mostrar que o termo  $(\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})\boldsymbol{\gamma}$  possui distribuição  $N(k, \omega^2)$ . Sendo assim, podemos reescrever tal termo como  $k + \omega V$ , em que  $V \sim N(0, 1)$ . Além disso, se forem observados os mesmos valores para as covariáveis para os indivíduos  $j_1$  e  $j_2$  e se estes forem pertencentes a *clusters*  $i_1$  e  $i_2$ , diferentes, então temos  $k = 0$  e a razão de chances é dada por:

$$OR = \exp\{\omega V\}.$$

Larsen e Merlo (2005) mostraram também que, sob normalidade para os efeitos aleatórios, a  $OR$  em (3.7) é uma quantidade que, dados  $\boldsymbol{\beta}$  e  $\boldsymbol{\Sigma}$ , tem distribuição lognormal. O resultado introduzido por Larsen e Merlo (2005) é apresentado no Teorema 3.1.1 a seguir e sua prova pode ser encontrada no Apêndice C.1.

**Teorema 3.1.1.** *Defina  $OR = \exp\{k + \omega V\}$ , onde  $k \in \mathbb{R}$  e  $\omega \in \mathbb{R}^+$  são dadas em (3.7). Se  $V$  tem distribuição normal padrão, então dados  $\boldsymbol{\beta}$  e  $\boldsymbol{\Sigma}$ ,  $OR$  tem distribuição lognormal( $k, \omega^2$ ) com densidade dada por*

$$\begin{aligned} f_{OR|k, \omega^2}(t) &= [\sqrt{2\pi t \omega}]^{-1} \exp\left\{-\frac{1}{2\omega^2}(\ln t - k)^2\right\}, \quad t \in \mathbb{R}^+ \\ &= t^{-1} \phi(\ln t | k, \omega^2). \end{aligned}$$

Larsen *et al.* (2000) estabelece outros resultados interessantes sobre a  $MOR$ . Estes resultados encontram-se no seguinte teorema cuja prova pode ser encontrada no Apêndice C.1.

**Teorema 3.1.2.** *Se  $V$  é uma variável aleatória com distribuição normal padrão então, dados  $\boldsymbol{\beta}$  e  $\boldsymbol{\Sigma}$ , segue que*

- (i)  $med\{\exp\{|\omega V|\}\} = \exp\{med\{|\omega V|\}\} = \exp\{\omega \Phi^{-1}(0, 75)\}$ ,
- (ii)  $med\{\exp\{-|\omega V|\}\} = [med\{\exp\{|\omega V|\}\}]^{-1}$ ,
- (iii)  $med\{\exp\{k + \omega V\}\} = \exp\{k\}$ ,

onde  $\omega \in \mathbb{R}^+$  e  $k \in \mathbb{R}$  são constantes.

Nas próximas seções apresentamos algumas das principais contribuições desta dissertação, onde obtemos resultados originais ao estender os resultados de Larsen *et al.* (2000) considerando famílias estendidas de distribuições normais para os efeitos aleatórios.

## 3.2 Efeitos Aleatórios Independentes com Distribuição Normal Assimétrica Univariada

A razão de chances em (3.2) também pode ser escrita como  $OR = \exp\{k + W\}$ , em que  $k = (\mathbf{x}'_{i_1 j_1} - \mathbf{x}'_{i_2 j_2})\boldsymbol{\beta}$  e  $W = (\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})\boldsymbol{\gamma}$ . Suponhamos agora que os efeitos aleatórios  $\gamma_i$ s são

independentes e identicamente distribuídos com distribuição normal assimétrica univariada com parâmetros de locação igual a zero, de escala  $\sigma^2 \in \mathbb{R}^+$  e de assimetria  $\lambda \in \mathbb{R}$ , ou seja, para o efeito aleatório  $\gamma_i$  associado ao  $i$ -ésimo *cluster* consideramos a distribuição  $SN(0, \sigma^2, \lambda)$ .

Sob estas suposições, condicionalmente ao vetor de parâmetros  $\beta$  associado aos efeitos fixos e nos hiperparâmetros  $\sigma^2$  e  $\lambda$  que indexam a distribuição dos efeitos aleatórios, vamos encontrar a distribuição de  $OR$  entre os indivíduos  $j_1$  e  $j_2$  pertencentes, respectivamente aos *clusters*  $i_1$  e  $i_2$ . Iniciamos por encontrar a distribuição da variável aleatória  $W = (\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})\gamma = \gamma_{i_1} - \gamma_{i_2}$  que será dada na Proposição 3.2.1 a seguir.

**Proposição 3.2.1.** *Se  $\gamma_i$  e  $\gamma_j$  são independentes e ambos têm distribuição  $SN(0, \sigma^2, \lambda)$  então, dados  $\sigma^2$  e  $\lambda$ , a distribuição de  $W = \gamma_i - \gamma_j$  tem a seguinte função densidade de probabilidade*

$$f_{W|\sigma^2, \lambda}(w) = 4\phi(w|2\sigma^2)\Phi_2\left(\lambda w[1, -1]'|4\sigma^2\left[I_2 + \frac{\lambda^2}{2}\mathbf{1}_2\mathbf{1}'_2\right]\right), \quad (3.8)$$

onde  $\mathbf{1}_2$  denota um vetor, de ordem  $2 \times 1$ , no qual todos os componentes são iguais ao valor 1.

Note que a distribuição em (3.8) pode ser escrita da seguinte forma

$$f_{W|\sigma^2, \lambda}(w) = \phi(w|2\sigma^2) \frac{\Phi_2(\lambda[1, -1]'w | 4\sigma^2 I_2 + \lambda \mathbf{1}_2 [2\sigma^2] \mathbf{1}'_2 \lambda)}{\Phi_2(\mathbf{0}|4\sigma^2 I_2)}. \quad (3.9)$$

Embora tenha forma bastante semelhante à da distribuição encontrada por Arellano-Valle *et al.* (2006), tal distribuição não é um membro da família unificada de distribuições normais assimétrizadas (SUN) definida por tais autores. No entanto, esta distribuição pertence a uma família que é uma pequena variação da família SUN. Por esta razão, a família de distribuições definida em (3.8) será denotada por  $W|\sigma^2, \lambda \sim SUNM\left(2\sigma^2, \lambda[1, -1]', 4\sigma^2\left[I_2 + \frac{\lambda^2}{2}\mathbf{1}_2\mathbf{1}'_2\right]\right)$ .

Note que se assumirmos que os efeitos aleatórios são definidos tal que  $\gamma_i \sim SN(\xi, \sigma^2, \lambda)$ , em que  $\xi \neq 0$ , como ocorre na seção 2.3, a Proposição 3.2.1 continua valendo, pois o efeito aleatório  $\gamma_i$  pode ser escrito como  $\gamma_i = \Gamma_i + \xi$ , em que  $\Gamma_i \sim SN(0, \sigma^2, \lambda)$ . Disso segue que

$$W = (\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})\gamma = \Gamma_{i_1} + \xi - \Gamma_{i_2} - \xi = \Gamma_{i_1} - \Gamma_{i_2}.$$

As provas da Proposição 3.2.1 e das demais proposições apresentadas nesta seção podem ser encontradas no Apêndice C.2. O resultado apresentado na Proposição 3.2.1 a respeito da distribuição de  $W = \gamma_{i_1} - \gamma_{i_2}$  será útil na prova da Proposição 3.2.2, apresentada a seguir, na qual encontramos a distribuição para a razão de chances ao assumirmos a distribuição normal assimétrica univariada para os efeitos aleatórios e que é um dos principais resultados deste capítulo.

**Proposição 3.2.2.** *Seja  $OR = \exp\{k + W\}$ , onde  $k = (\mathbf{x}'_{i_1 j_1} - \mathbf{x}'_{i_2 j_2})\beta$  e  $W = \gamma_i - \gamma_j$ . Se  $\gamma_i$  e  $\gamma_j$  são variáveis aleatórias que, dados  $\sigma^2$  e  $\lambda$ , são independentes e identicamente distribuídas com distribuição  $SN(0, \sigma^2, \lambda)$ , então condicionalmente em  $\beta$ ,  $\sigma^2$  e  $\lambda$ , a variável  $OR$  tem distribuição com densidade dada por*

$$f_{OR|\beta, \sigma^2, \lambda}(t) = \frac{4}{t}\phi(\ln t|k, 2\sigma^2)\Phi_2\left(\lambda[1, -1]'\ln t|\lambda[1, -1]'k, \left[I_2 + \frac{\lambda^2}{2}\mathbf{1}_2\mathbf{1}'_2\right]\right), \quad (3.10)$$

onde  $t \in \mathbb{R}^+$ .

Por simplicidade, chamaremos a distribuição em (3.10) de lognormal assimetrizada, embora esta não seja a mesma definida em Marchenko e Genton (2010).

Como para o caso normal, aqui também temos uma forma fechada para a *MOR*. Este resultado é fornecido na seguinte proposição.

**Proposição 3.2.3.** *Se os efeitos aleatórios são independentes e identicamente distribuídos tal que  $\gamma_i \sim SN(0, \sigma^2, \lambda)$ , então segue que a mediana da distribuição em (3.8) ou razão de chances mediana, dados  $\beta$ ,  $\sigma^2$ ,  $\lambda$  é*

$$MOR = med\{\exp\{k+W\}|\beta, \sigma^2, \lambda\} = \exp\left\{\Phi_{SUNM}^{-1}\left(0, 5|2\sigma^2, \lambda[1, -1]', 4\sigma^2\left[I_2 + \frac{\lambda^2}{2}\mathbf{1}_2\mathbf{1}_2'\right]\right) + k\right\},$$

onde  $\Phi_{SUNM}(\cdot)$  denota a f.d.a. da distribuição em (3.8).

### 3.3 Efeitos Aleatórios Correlacionados com Distribuição Normal Assimétrica Multivariada

Vamos considerar agora que o vetor de efeitos aleatórios  $\gamma = (\gamma_1, \dots, \gamma_k)'$  tem distribuição normal assimétrica multivariada com parâmetro de locação  $\xi = (0, \dots, 0)'$ , matriz de escala  $\sigma^2 I_k$ ,  $\sigma^2 \in \mathbb{R}^2$ , e vetor de parâmetros de forma que regulam a assimetria igual a  $\alpha \mathbf{1}_k$ ,  $\alpha \in \mathbb{R}$ , ou seja,  $\gamma \sim SN_k(\mathbf{0}, \sigma^2 I_k, \alpha \mathbf{1}_k)$ . Da Proposição 3.3.1, apresentada a seguir e cuja prova pode ser vista no Apêndice C.3, vemos que ao assumirmos esta distribuição conjunta para os efeitos aleatórios estamos impondo que estes sejam correlacionados. Neste caso, a distribuição conjunta de  $\gamma_i$  e  $\gamma_j$  é  $SN_2(\mathbf{0}_2, \sigma^2 I_2, \alpha^* \mathbf{1}_2)$ , onde  $\alpha^* = \alpha[1 + \alpha^2(k-2)]^{-1/2}$ , a qual não fatora no produto de duas densidades marginais. Aplicando a fórmula em (2.9), encontramos que a correlação entre  $\gamma_i$  e  $\gamma_j$  é

$$Corr(\gamma_i, \gamma_j) = \frac{-\frac{2}{\pi}(\Delta^*)^2}{1 - \frac{2}{\pi}(\Delta^*)^2},$$

em que  $\Delta^* = \frac{\alpha^*}{\sqrt{1+2(\alpha^*)^2}}$ .

**Proposição 3.3.1.** *Se o vetor de efeitos aleatórios  $\gamma \sim SN_k(\mathbf{0}_k, \sigma^2 I_k, \alpha \mathbf{1}_k)$ , então o vetor  $\gamma_{ij} = (\gamma_i, \gamma_j)'$  formado pelas componentes  $\gamma_i$  e  $\gamma_j$  de  $\gamma$  tem distribuição  $SN_2(\mathbf{0}_2, \sigma^2 I_2, \alpha^* \mathbf{1}_2)$ , em que  $\alpha^* = \alpha[1 + \alpha^2(k-2)]^{-1/2}$ .*

Sob estas suposições sobre os efeitos aleatórios, pode-se provar que a OR pode ser escrita com em (3.7), ou seja,

$$OR = \exp\{k + \omega V\},$$

em que  $k = (\mathbf{x}'_{i_1 j_1} - \mathbf{x}'_{i_2 j_2})\beta$ ,  $\omega^2 = \sigma^2(\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})I_k(\mathbf{z}_{i_1 j_1} - \mathbf{z}_{i_2 j_2})$  e  $V$  é uma variável aleatória tal que  $V \sim N(0, 1)$ . Para demonstrar este fato utilizamos o resultado a seguir cuja prova encontra-se no Apêndice C.3.

**Proposição 3.3.2.** *Se o vetor de efeitos aleatórios é tal que  $\gamma \sim SN_k(\mathbf{0}_k, \sigma^2 I_k, \alpha \mathbf{1}_k)$ , então a variável aleatória  $\gamma_i - \gamma_j$ , em que  $\gamma_i$  e  $\gamma_j$  são componentes de  $\gamma$ , tem distribuição Normal com média zero e variância  $2\sigma^2$ .*

O resultado dado na Proposição 3.3.2 pode ser estendido para situações mais gerais, como por exemplo quando assumimos  $\gamma \sim SN_k(\xi \mathbf{1}_k, \sigma^2 I_k, \alpha \mathbf{1}_k)$ , com  $\xi \neq 0$ , suposição feita no modelo implementado na Seção 2.4. Esta extensão encontra-se na Proposição 3.3.3.



**Proposição 3.3.3.** *Seja  $\mathbf{Y} \sim SN_k(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ , em que  $\boldsymbol{\Omega} = \text{diag}(\sigma_1^2, \dots, \sigma_k^2)$ ,  $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)'$  e  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)'$ . Considere a combinação linear dos elementos de  $\mathbf{Y}$  dada por  $C'\mathbf{Y}$ , em que  $C' = (c_1, \dots, c_k)$  é um vetor de constantes. Se  $\sum_{i=1}^k \alpha_i \sigma_i c_i = 0$  então  $C'\mathbf{Y} \sim N\left(\sum_{i=1}^k \xi_i c_i, \sum_{i=1}^k c_i^2 \sigma_i^2\right)$ .*

Se estamos calculando a razão de chances do evento de interesse ocorrer para indivíduos  $j_1$  e  $j_2$  que pertencem, respectivamente, aos clusters  $i_1$  e  $i_2$ , distintos, então o termo  $(\mathbf{z}'_{i_1 j_1} - \mathbf{z}'_{i_2 j_2})\boldsymbol{\gamma}$  torna-se  $\gamma_{i_1} - \gamma_{i_2}$ . Este termo  $\gamma_{i_1} - \gamma_{i_2}$  pode ser visto como uma combinação linear  $C'\boldsymbol{\gamma}$ , em que  $C' = (c_1, \dots, c_k)$ ,  $c_l = 1$ ,  $c_j = -1$  e  $c_i = 0$ ,  $\forall i \neq l, j$ . Sob as suposições feitas para os efeitos aleatórios, podemos aplicar a Proposição 3.3.3 considerando  $\xi_i = 0$ ,  $\sigma_i^2 = \sigma^2$  e  $\alpha_i = \alpha$ ,  $\forall i = 1, \dots, k$ . Neste caso, temos que  $\sum_{i=1}^k \alpha_i \sigma_i c_i = \sigma \alpha \sum_{i=1}^k c_i = 0$ , fato que atende a condição imposta pela proposição e então, conseqüentemente, temos que a variável aleatória  $C'\boldsymbol{\gamma} = \gamma_{i_1} - \gamma_{i_2}$  tem distribuição normal com média  $\sum_{i=1}^k \xi_i c_i = 0$  e variância  $\sum_{i=1}^k c_i^2 \sigma_i^2 = \sigma^2 \sum_{i=1}^k c_i^2 = 2\sigma^2$ . Desta forma, podemos escrever a expressão da razão de chances da mesma forma que no caso em que consideramos normalidade para os efeitos aleatórios, visto que  $\omega^2 = (\mathbf{z}'_{i_1} - \mathbf{z}'_{i_2})\sigma^2 I_k (\mathbf{z}_{i_1} - \mathbf{z}_{i_2}) = 2\sigma^2$ .

Assim, como conseguimos escrever a razão de chances em (3.2) na mesma forma que a obtida em (3.7), todos os resultados apresentados na Seção 3.1 para medidas propostas por Larsen *et al.* (2000) são também válidos para esta situação. Ou seja, como pode ser visto no Teorema 3.1.1, dados  $k$  e  $\omega^2$ , a variável aleatória  $OR$  tem distribuição lognormal( $k, \omega^2$ ) com densidade

$$f_{OR|k, \omega^2}(t) = [\sqrt{2\pi t \omega^2}]^{-1} \exp\left\{-\frac{1}{2\omega^2}(\ln t - k)^2\right\}, t \in \mathbb{R}^+,$$

onde  $\omega > 0$  e  $k > 0$ . Além disso, assumindo uma distribuição normal assimétrica multivariada para os efeitos aleatórios, a razão de chances mediana (MOR), como demonstrado no Teorema 3.1.2, é tal que

- (i)  $\text{med}\{\exp\{|\omega V|\}\} = \exp\{\text{med}\{|\omega V|\}\} = \exp\{\omega \Phi^{-1}(0, 75)\}$ ,
- (ii)  $\text{med}\{\exp\{-|\omega V|\}\} = [\text{med}\{\exp\{|\omega V|\}\}]^{-1}$ ,
- (iii)  $\text{med}\{\exp\{k + \omega V\}\} = \exp\{k\}$ .

### 3.4 Determinando a Razão de Chances *a posteriori*

Os resultados obtidos nas seções anteriores são úteis para a interpretação da  $OR$  do ponto de vista clássico. Do ponto de vista bayesiano, no entanto, esses resultados podem apenas auxiliar na construção de distribuições *a priori* mais apropriadas para os efeitos aleatórios. Tais distribuições *a priori* podem ser escolhidas atendendo a características conhecidas a respeito de  $OR$ . No entanto, *a posteriori*, para interpretarmos os resultados obtidos necessitamos determinar a distribuição *a posteriori* de  $OR$ . Tal distribuição depende tanto dos efeitos fixos  $\boldsymbol{\beta}$  quanto dos efeitos aleatórios  $\boldsymbol{\gamma}$ .

Como as distribuições *a posteriori* para tais parâmetros não têm forma fechada, utilizaremos o seguinte procedimento para aproximarmos a distribuição de  $OR$ .

Passo 1: Gere uma amostra da distribuição *a posteriori* dos parâmetros de interesse;

Passo 2: Calcule  $OR^{(r)} = \exp\{(\mathbf{x}_{i_1 j_1} - \mathbf{x}_{i_2 j_2})' \boldsymbol{\beta}^{(r)} + (\mathbf{z}_{i_1 j_1} - \mathbf{z}_{i_2 j_2})' \boldsymbol{\gamma}^{(r)}\}$ , em que  $(\boldsymbol{\beta}^{(r)}, \boldsymbol{\gamma}^{(r)})$  são os valores gerados no  $r$ -ésimo passo do processo MCMC para a distribuição *a posteriori* de  $\boldsymbol{\beta}$  e  $\boldsymbol{\gamma}$ .

A expressão no Passo 2 pode ser mudada de acordo com o interesse dos pesquisadores, como, por exemplo, para situações em que a razão de chances depende apenas dos efeitos aleatórios ou apenas dos efeitos fixos. A vantagem desta abordagem é que as medidas resumo *a posteriori* de OR, como, por exemplo, mediana, moda, média e intervalos de credibilidade de mais alta densidade (HPD), são facilmente obtidas. A moda *a posteriori* da razão de chances é um estimador interessante dado que ela está em uma região de alta densidade. Assim, informalmente falando, nós podemos concluir que quando comparamos dois sujeitos de dois *clusters* selecionados aleatoriamente a razão de chances mais provável é a razão de chances modal. A mediana tem uma interpretação natural em termos de probabilidades e ela é similar a MOR proposta por Larsen *et al.* (2000). Além de fornecer uma estimativa intervalar, os intervalos HPD podem ser utilizados para realizar testes de hipóteses sobre a razão de chances. Ao formular testes de hipóteses sobre a razão de chances, geralmente, o principal interesse é testar a hipótese nula  $H_0 : OR = 1$ . Esta hipótese pode ser aceita se o valor um estiver incluso no intervalo HPD *a posteriori* para OR.

# Capítulo 4

## Estudo com dados Simulados

Neste capítulo é apresentado um estudo de simulação com o objetivo de avaliar o comportamento dos modelos apresentados nos capítulos anteriores. Especificamente, queremos avaliar se a má especificação da distribuição dos efeitos aleatórios afeta significativamente as estimativas dos efeitos fixos e aleatórios, bem como as estimativas para a OR. Este é apenas um estudo exploratório e será aprimorado futuramente.

Em tal estudo as variáveis resposta foram geradas de uma distribuição Bernoulli( $\pi_{ij}$ ) com  $\pi_{ij}$  definido como em (1.1). As respostas foram divididas em 25 *clusters* ( $k = 25$ ) contendo 40 observações cada ( $n_i = 40$ ), totalizando um total de 1000 observações. Para gerar os dados, consideramos o modelo com apenas uma covariável, para a qual foram gerados 1000 valores da distribuição normal padrão. Como efeitos fixos consideramos  $\beta_0 = 0,5$  e  $\beta_1 = 2,5$ . Para gerar dos efeitos aleatórios consideramos três situações distintas, considerando as distribuições normal, normal assimétrica univariada e normal assimétrica multivariada. Os efeitos aleatórios gerados foram transformados de tal forma a terem média e variância exatamente iguais ao especificado para a distribuição da qual foram gerados. Em resumo, o processo de geração dos dados pode ser descrito pelos passos a seguir:

**Passo 1:** Definir número de *clusters*  $k$  e número  $n_i$  de observações em cada *cluster*  $i$ ;

**Passo 2:** Definir valores dos efeitos fixos  $\beta$ ;

**Passo 3:** Gerar efeitos aleatórios  $\gamma$  da distribuição especificada;

**Passo 4:** Calcular a probabilidade  $\pi_{ij}$ ,  $\forall i = 1, \dots, k$  e  $j = 1, \dots, n_i$ ;

**Passo 5:** Gerar os valores  $y_{ij}$ ,  $\forall i = 1, \dots, k$  e  $j = 1, \dots, n_i$ .

Para cada banco de dados gerado ajustamos três modelos, cujos nomes serão definidos pela distribuição adotada para os efeitos aleatórios. O modelo Normal, descrito na Seção 1.3, é o modelo logístico misto com efeitos aleatórios normalmente distribuídos. O modelo SNI, foi apresentado na Seção 2.3, é o modelo logístico misto com efeitos aleatórios SN independentes. Por fim, o modelo SND é o modelo logístico com efeitos aleatórios SN dependentes, modelo apresentado na Seção 2.4.

Para completar a especificação dos modelos eliciamos as distribuições *a priori*. Uma vez que não apresentaremos uma análise de sensibilidade do modelo, optamos por considerar distribuições *a priori* pouco informativas para todos os parâmetros. Em todos os modelos, para os parâmetros de efeitos fixos assumimos como distribuição *a priori* que  $\beta_i \sim N(0; 10)$ ,  $i = 0, 1$ , e para o parâmetro

de escala assumimos  $\sigma^2 \sim GI(2, 001; 1)$ . A distribuição *a priori* para  $\sigma^2$  foi escolhida de tal forma a garantirmos a existência dos momentos de primeira e segunda ordem da distribuição, assim temos, *a priori*, que  $E(\sigma^2) = 1$  e  $V(\sigma^2) = 1000$ . Assumimos também  $\lambda \sim N(0, 1000)$  e  $\alpha \sim N(0, 1000)$  nos modelos SNI e SND, respectivamente. Como já dito anteriormente, a distribuição *a priori* normal para os parâmetros de assimetria foi utilizada em Arellano-Valle *et al.* (2009).

Para gerar das distribuições *a posteriori* dos parâmetros consideramos *burn in* de 20000 para o modelo Normal e 40000 para os modelos SNI e SND. Além disso, os métodos MCMC fornecem amostras *a posteriori* com a autocorrelação elevada. Por isso consideramos para o *lag* os valores de 10, 20 e 30 para os modelos Normal, SNI e SND, respectivamente. No processo de inferência, para obter as estimativas da distribuição *a posteriori* dos parâmetros em cada um dos modelos, consideramos as amostras *a posteriori* com tamanho de 1000. A análise de convergência, que não será apresentada, foi realizada através do diagnóstico de convergência de Geweke (Geweke (1992)), do gráfico do traço e do gráfico da média ergódica das amostras das distribuições *a posteriori*.

Para cada banco de dados os modelos implementados foram comparados utilizando estatísticas que avaliam a qualidade de ajuste dos modelos. As estatísticas utilizadas são *conditional predictive ordinate* (CPO), *deviance information criterion* (DIC) e médias *a posteriori* do índice C, uma aproximação para a área abaixo da curva ROC, e da taxa de classificação correta (TCC). O critério CPO é utilizado através da estatística resumo denominada logaritmo da verossimilhança pseudo-marginal (LPML). No apêndice D são apresentados detalhes de implementação destas estatísticas.

Na próximas seções são apresentados mais detalhes a respeito da geração dos dados e a inferência realizada para cada um dos banco de dados com os modelos considerados.

## 4.1 Efeitos Aleatórios com Distribuição Normal

Nesta seção apresentamos o estudo de simulação com dados gerados de um modelo logístico misto com efeitos aleatórios normalmente distribuídos. Neste estudo, foram gerados dois bancos de dados com diferença na variância definida para a distribuição dos efeitos aleatórios. No primeiro, denominado banco de dados 1, os efeitos aleatórios foram gerados independentes e indenticamente distribuídos com distribuição normal padrão e a no segundo banco de dados, denominado banco de dados 2, os efeitos aleatórios foram gerados independentes e indenticamente distribuídos com distribuição normal com média 0 e variância 4. Na Figura 4.1 são apresentados o histograma dos efeitos aleatórios gerados e a respectiva distribuição teórica nos dois bancos de dados gerados.

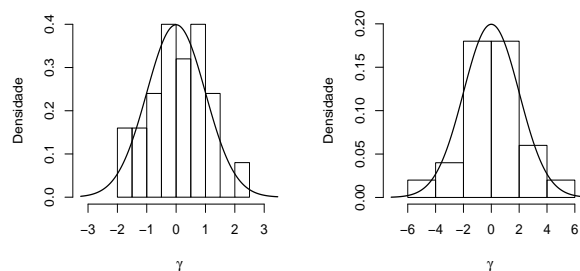


Figura 4.1: Histograma e densidade teórica dos efeitos aleatórios  $\gamma$  gerados para os bancos de dados 1 (esquerda) e 2 (direita).

A Tabela 4.1 apresenta as estimativas, para ambos bancos de dados gerados e obtidas com os três modelos que foram considerados, para os parâmetros de efeitos fixos  $\beta$ , para os parâmetros que indexam a distribuição assumida para os efeitos aleatórios e também para a variância dos efeitos aleatórios. Ao avaliar as estimativas, lembre que no modelo normal a variância dos efeitos aleatórios é igual ao próprio parâmetro  $\sigma^2$ . Os estimadores, *a posteriori*, utilizados são média (Me), mediana (Md), moda (Mo) e o intervalo de mais alta densidade com 95% de probabilidade (HPD).

Tabela 4.1: Estimativas *a posteriori* para os parâmetros

	Me	Md	Mo	HPD	Me	Md	Mo	HPD
	Banco de dados 1 - $V(\gamma_i) = 1$				Banco de dados 2 - $V(\gamma_i) = 4$			
modelo Normal								
$\beta_0$	0,448	0,446	0,391	[0,044;0,815]	0,497	0,493	0,515	[-0,210;1,398]
$\beta_1$	2,522	2,517	2,484	[2,241;2,839]	2,483	2,487	2,512	[2,141;2,818]
$\sigma^2$	0,883	0,827	0,748	[0,356;1,564]	3,375	3,117	2,786	[1,662;5,693]
modelo SNI								
$\beta_0$	0,455	0,458	0,433	[-0,005;0,825]	0,497	0,498	0,478	[-0,233;1,148]
$\beta_1$	2,510	2,504	2,499	[2,202;2,842]	2,477	2,471	2,457	[2,170;2,825]
$\sigma^2$	1,383	1,132	0,847	[0,373;2,961]	3,902	3,604	3,037	[1,538;6,871]
$V(\gamma_i)$	0,859	0,797	0,687	[0,301;1,444]	3,423	3,223	2,758	[1,491;5,552]
$\lambda$	-2,836	-0,459	-0,168	[-14,530;1,376]	-0,068	-0,035	-0,049	[-1,054;1,012]
modelo SND								
$\beta_0$	0,453	0,458	0,474	[0,165;0,712]	0,495	0,498	0,528	[0,059;1,013]
$\beta_1$	2,528	2,527	2,531	[2,252;2,866]	2,480	2,479	2,488	[2,137;2,821]
$\sigma^2$	0,895	0,843	0,759	[0,405;1,559]	3,309	3,118	2,868	[1,634;5,353]
$V(\gamma_i)$	0,873	0,824	0,740	[0,395;1,519]	3,229	3,040	2,795	[1,601;5,226]
$\alpha$	0,108	0,116	0,005	[-8,943;9,999]	0,110	0,115	0,133	[-9,999;10,258]

Analisando a Tabela 4.1 percebemos que para o banco de dados 1, se considerarmos as estimativas dadas por média e mediana *a posteriori*, o modelo SNI foi o que melhor estimou os parâmetros  $\beta_0$  e  $\beta_1$ , pois apresentou estimativas mais próximas dos valores reais 0,5 e 2,5. Se considerarmos a moda *a posteriori*, o modelos SND e SNI foram melhores para estimar  $\beta_0$  e  $\beta_1$ , respectivamente. Em todos os modelos os intervalos HPD *a posteriori*, com 95% de probabilidade, incluíram os verdadeiros valores de  $\beta_0$  e  $\beta_1$ . Os modelos SND e Normal apresentaram os intervalos HPD com menor amplitude para  $\beta_0$  e  $\beta_1$ , respectivamente. O modelo Normal foi o que melhor estimou a variância dos efeitos aleatórios  $V(\gamma_i)$  considerando as estimativas pontuais. Para a variância, todos os intervalos HPD incluíram o verdadeiro valor e o modelo SND apresentou o intervalo HPD com menor amplitude. O modelo SNI estimou valores negativos para o parâmetro de assimetria  $\lambda$ , a média *a posteriori* e intervalo HPD foram  $-2,836$  e  $[-14,530;1,376]$ . No entanto, os efeitos aleatórios foram gerados com a distribuição normal, ou seja, com  $\lambda = 0$ . Ao observarmos a Figura 4.1 podemos notar que, apesar dos efeitos aleatórios para o banco de dados 1 terem sido gerados da distribuição normal padrão, o histograma dos efeitos aleatórios é um pouco assimétrico. Para o modelo SND obtivemos estimativas próximas de zero para o parâmetro de assimetria  $\alpha$  e o intervalo HPD não parece indicar a presença de assimetria na distribuição dos efeitos aleatórios.

Analisando agora as estimativas obtidas para o banco de dados 2, vemos que o modelo Normal, modelo utilizado na geração dos dados, apresentou as estimativas mais próximas dos valores reais de  $\beta_0$  ao utilizar os estimadores de média e moda *a posteriori* e estimou valores mais próximos

do real de  $\beta_1$  com todos os estimadores. Em todos os modelos os intervalos HPD com 95% de probabilidade incluíram o valores reais. Os modelos que apresentam os intervalo HPD com menor amplitude foram o modelo SND para o parâmetro  $\beta_0$  e o modelo SNI para  $\beta_1$ . Avaliando as estimativa para variância dos efeitos aleatórios, o modelo SNI estimou valores mais próximos de 4, valor real, ao utilizar os estimadores de média e mediana *a posteriori*, enquanto o modelo SND estimou melhor a variância ao considerar a moda *a posteriori*. Novamente em todos os modelos o intervalo HPD incluiu o verdadeiro valor da variância e o modelo SND foi o que apresentou o intervalo HPD com menor amplitude. Os modelos SNI e SND estimaram valores próximos de zero para os parâmetros de assimétria  $\lambda$  e  $\alpha$ , indicando a ausência de assimetria na distribuição dos efeitos aleatórios.

Com o objetivo de avaliar a estimação dos efeitos aleatórios  $\gamma$ , vamos considerar a soma de quadrados dos erros (SQE) quando estimamos os efeitos aleatórios pela média *a posteriori*  $E(\gamma|\mathbf{y})$ . O SQE pode ser calculado por  $\sum_{i=1}^k (\gamma_i - E(\gamma_i|\mathbf{y}))^2$ . Naturalmente, o modelo que apresenta menor SQE estima melhor os efeitos aleatórios. Na Tabela 4.2 apresentamos os valores obtido para SQE considerando os modelos Normal, SNI e SND, para os dois bancos de dados considerados.

Tabela 4.2: Soma de quadrados dos erros para os efeitos aleatórios  $\gamma$

Modelo	Banco de dados	
	1	2
Normal	<b>3,991</b>	5,568
SNI	4,247	5,740
SND	4,186	<b>5,540</b>

Na Tabela 4.2 vemos que, para o banco de dados 1, o modelo Normal foi o que melhor estimou os efeitos aleatórios pois apresentou o menor valor de SQE. Isso era esperado uma vez que o modelo Normal é o modelo correto para estes dados. Para o banco de dados 2, o modelo SND foi o que melhor estimou os efeitos aleatórios, mas o modelo Normal apresentou SQE bem próximo.

Na Tabela 4.3 são apresentadas os valores para CPO, DIC, índice C e taxa de classificação correta obtidos para cada um dos modelos ajustados.

Tabela 4.3: Medidas de qualidade de ajuste do modelos

Modelo	CPO	DIC	Índice C	TCC	CPO	DIC	Índice C	TCC
	Banco de dados 1				Banco de dados 2			
Normal	<b>-394,199</b>	<b>788,023</b>	<b>0,906</b>	<b>0,820</b>	<b>-361,357</b>	<b>722,805</b>	<b>0,924</b>	<b>0,854</b>
SNI	-394,570	788,780	0,905	<b>0,820</b>	-361,980	724,094	0,923	0,853
SND	-394,269	788,231	0,905	<b>0,820</b>	-361,596	723,338	0,923	<b>0,854</b>

Para o banco de dados 1, a Tabela 4.3 mostra que o modelo Normal apresentou melhor ajuste aos dados segundo os critérios de CPO, DIC e índice C. No entanto os valores para o índice C foram muito próximos indicando que, por esse critério, os modelos são praticamente iguais. Através da taxa de classificação correta não houve diferença entre os modelos. Para o banco de dados 2 observamos que, de acordo com os critérios CPO, DIC e índice C, o modelo Normal apresentou melhor ajuste aos dados do que os demais modelos. Segundo a taxa de classificação correta os modelos Normal e SND foram um pouco melhores do que o modelo SNI. De forma geral, as medida de qualidade de ajuste de modelo apontaram o modelo utilizado na geração dos dados como sendo o melhor.

Vamos calcular a razão de chances *a posteriori*, como descrito na Seção 3.4, para alguns casos. Inicialmente consideramos a razão de chances entre indivíduos que apresentam o mesmo valor  $x$  observado para a covariável, mas pertencem a *clusters* distintos, em que razão de chances depende apenas dos efeitos aleatórios  $\gamma$ . Consideramos também a razão de chances entre indivíduos de um mesmo *cluster* com diferença entre os valores observados para a covariável igual a uma unidade. Nesta situação a razão de chances depende apenas do parâmetro  $\beta_1$ . Por fim, consideramos a razão de chances entre indivíduos que pertençam a diferentes *clusters* e apresentem diferença igual a 1 para o valor observado para a covariável. As situações abordadas são:

$OR_1$ : razão de chances entre um indivíduo que pertence ao primeiro *cluster* dos 25 gerados e um indivíduo do segundo *cluster* com o mesmo valor observado  $x$  para a covariável;

$OR_2$ : razão de chances entre indivíduos do mesmo *cluster* e com diferença igual a uma unidade na covariável;

$OR_3$ : razão de chances entre um indivíduo do *cluster* 1 com valor da covariável igual a  $x + 1$  e um indivíduo do *cluster* 2 com valor da covariável igual a  $x$ .

Na Tabela 4.4 apresentamos os valores reais e as respectivas estimativas para a razão de chances nas situações abordadas para cada um dos bancos de dados gerados.

Tabela 4.4: Estimativas *a posteriori* para a razão de chances

	Real	Me	Md	Mo	HPD	Real	Me	Md	Mo	HPD
	Banco de dados 1					Banco de dados 2				
modelo Normal										
$OR_1$	5,611	2,725	2,236	1,794	[0,500; 6,146]	1,969	5,127	3,832	2,542	[0,482; 14,829]
$OR_2$	12,182	12,614	12,388	11,864	[9,398; 17,098]	12,182	12,160	12,030	12,204	[8,267; 16,459]
$OR_3$	68,358	34,621	28,124	21,349	[5,825; 80,623]	23,989	63,363	47,068	28,291	[5,738; 172,804]
modelo SNI										
$OR_1$	5,611	2,555	2,210	1,719	[0,491; 5,224]	1,969	4,869	3,642	2,778	[0,552; 12,021]
$OR_2$	12,182	12,471	12,226	12,035	[8,479; 16,308]	12,182	12,090	11,833	11,444	[8,547; 16,643]
$OR_3$	68,358	32,161	27,429	20,457	[7,158; 69,981]	23,989	60,216	43,086	30,171	[6,289; 152,850]
modelo SND										
$OR_1$	5,611	2,679	2,211	1,755	[0,566; 6,087]	1,969	4,876	3,755	2,514	[0,487; 12,397]
$OR_2$	12,182	12,681	12,512	12,286	[8,827; 16,404]	12,182	12,115	11,927	11,731	[8,193; 16,322]
$OR_3$	68,358	34,429	27,150	20,422	[4,572; 80,372]	23,989	59,853	43,800	28,171	[4,797; 152,493]

A Tabela 4.4 nos mostra que, nos dois bancos de dados, os três modelos estimaram muito bem a razão de chances  $OR_2$ , aquela que depende apenas de  $\beta_1$ , indicando a boa estimação de  $\beta_1$ . Para  $OR_2$  o modelo SNI foi o melhor para os dois bancos de dados ao considerar a média *a posteriori*, mas, de forma geral, as estimativas foram próximas com qualquer um dos estimadores e em ambos bancos de dados. As estimativas pontuais para  $OR_1$  e  $OR_3$  foram ruins se compararmos com os valores reais. Para o banco de dados 1, o modelo Normal obteve melhor desempenho ao estimar  $OR_1$  e  $OR_3$ , no entanto foi o modelo que estimou o intervalo HPD com maior amplitude para estas quantidades. Note também que o valor real de  $OR_1$  (5,611) não está contido no intervalo HPD para o modelo SNI, [0,491; 5,224]. Para o banco de dados 2, o modelo SNI estimou um pouco melhor  $OR_1$  quando consideramos média e mediana *a posteriori*. O modelo SND foi o melhor na estimação de  $OR_3$  quando consideramos média e moda *a posteriori*, mas a estimação foi ruim.

## 4.2 Efeitos Aleatórios com Distribuição Normal Assimétrica Univariada

Nesta seção apresentamos o estudo de simulação com dados gerados de um modelo logístico misto com efeitos aleatórios com distribuição normal assimétrica univariada. Neste estudo, foram gerados quatro bancos de dados diferentes nos valores especificados para  $\lambda$  e variância dos efeitos aleatórios. Os valores especificados para o parâmetro de escala  $\sigma^2$  foram escolhidos da forma que  $V(\gamma_i)$  fosse igual a 1 nos bancos de dados 1 e 3, e igual a 4 nos bancos de dados 2 e 4. O parâmetro de locação da distribuição normal assimétrica univariada foi escolhido para que  $E(\gamma_i) = 0$ . As distribuições utilizadas para gerar os efeitos aleatórios  $\gamma$  foram

**Banco de dados 1:**  $\gamma_i \sim SN(-1, 158; 2, 342; 3)$ ;

**Banco de dados 2:**  $\gamma_i \sim SN(-2, 317; 9, 367; 3)$ ;

**Banco de dados 3:**  $\gamma_i \sim SN(-1, 306; 2, 705; 10)$ ;

**Banco de dados 4:**  $\gamma_i \sim SN(-2, 612; 10, 820; 10)$ .

Na Figura 4.2 são apresentados o histograma dos efeitos aleatórios gerados e a respectiva distribuição teórica nos bancos de dados gerados.

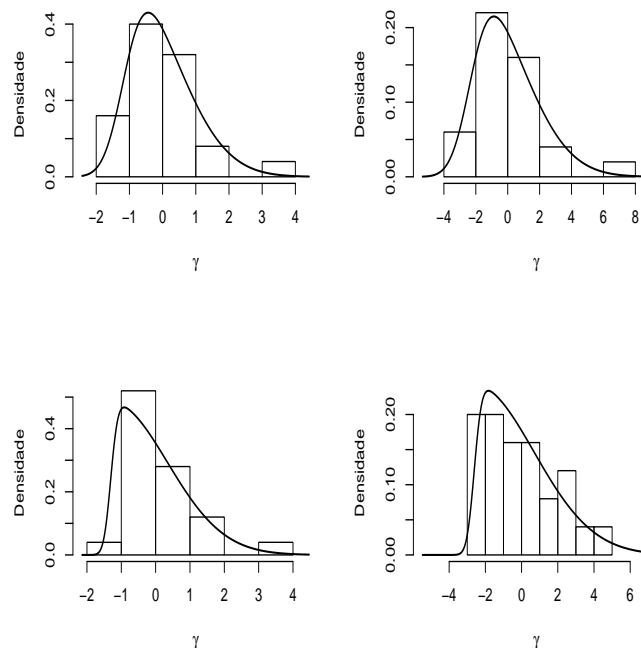


Figura 4.2: Histograma e densidade teórica dos efeitos aleatórios  $\gamma$  gerados para os bancos de dados 1 (esquerda em cima), 2 (direita em cima), 3 (esquerda em baixo) e 4 (direita em baixo).



A Tabela 4.5 apresenta as estimativas, obtidas nos três modelos que foram considerados, para os parâmetros de efeitos fixos  $\beta$ , para os parâmetros da distribuição assumida para os efeitos aleatórios e para a variância dos efeitos aleatórios.

Tabela 4.5: Estimativas *a posteriori* para os parâmetros

	Me	Md	Mo	HPD	Me	Md	Mo	HPD
	Banco de dados 1 - $\lambda = 3$ e $V(\gamma_i) = 1$				Banco de dados 2 - $\lambda = 3$ e $V(\gamma_i) = 4$			
modelo Normal								
$\beta_0$	0,478	0,473	0,460	[0,079;0,809]	0,437	0,435	0,472	[-0,189;1,181]
$\beta_1$	2,373	2,363	2,319	[2,087;2,672]	2,411	2,404	2,416	[2,123;2,755]
$\sigma^2$	0,629	0,592	0,491	[0,220;1,094]	3,165	2,950	2,557	[1,362;5,323]
modelo SNI								
$\beta_0$	0,483	0,475	0,440	[0,169,0,840]	0,416	0,406	0,377	[-0,259;1,046]
$\beta_1$	2,370	2,362	2,353	[2,057;2,662]	2,415	2,413	2,405	[2,090;2,717]
$\sigma^2$	1,057	0,912	0,694	[0,286;2,123]	6,729	6,581	6,896	[1,577;11,987]
$V(\gamma_i)$	0,600	0,552	0,467	[0,199;1,079]	3,015	2,862	2,711	[1,419;5,086]
$\lambda$	3,739	1,472	0,399	[-1,160;15,635]	8,035	7,017	0,572	[-0,875;21,322]
modelo SND								
$\beta_0$	0,486	0,486	0,447	[0,203; 0,735]	0,430	0,433	0,524	[-0,055; 0,923]
$\beta_1$	2,378	2,379	2,386	[2,076; 2,649]	2,412	2,411	2,395	[2,114; 2,721]
$\sigma^2$	0,641	0,592	0,482	[0,272; 1,197]	3,239	3,040	2,895	[1,638; 5,663]
$V(\gamma_i)$	0,626	0,578	0,470	[0,266; 1,166]	3,160	2,964	2,826	[1,597; 5,519]
$\alpha$	0,014	0,180	0,243	[-10,258; 9,224]	0,117	-0,202	-1,958	[-8,769; 9,323]
	Banco de dados 3 - $\lambda = 10$ e $V(\gamma_i) = 1$				Banco de dados 4 - $\lambda = 10$ e $V(\gamma_i) = 4$			
modelo Normal								
$\beta_0$	0,367	0,366	0,359	[0,040;0,699]	0,408	0,415	0,420	[-0,356;1,208]
$\beta_1$	2,423	2,420	2,427	[2,117;2,726]	2,623	2,608	2,568	[2,282;2,969]
$\sigma^2$	0,520	0,475	0,422	[0,193;0,944]	3,360	3,158	3,008	[1,664;5,657]
modelo SNI								
$\beta_0$	0,391	0,391	0,426	[0,075;0,664]	0,411	0,381	0,339	[-0,218;1,211]
$\beta_1$	2,404	2,400	2,375	[2,115;2,698]	2,628	2,621	2,610	[2,286;3,001]
$\sigma^2$	0,983	0,899	0,746	[0,295;1,878]	5,852	5,105	3,668	[1,668;12,013]
$V(\gamma_i)$	0,402	0,366	0,308	[0,136;0,743]	3,257	3,041	2,688	[1,459;5,611]
$\lambda$	8,680	8,045	7,126	[-0,939;20,054]	4,473	1,340	0,266	[-1,094;17,695]
modelo SND								
$\beta_0$	0,372	0,375	0,396	[0,135; 0,603]	0,405	0,398	0,298	[-0,021; 0,909]
$\beta_1$	2,422	2,419	2,421	[2,137; 2,757]	2,625	2,614	2,584	[2,294; 2,995]
$\sigma^2$	0,515	0,475	0,422	[0,188; 0,910]	3,354	3,163	3,003	[1,697; 5,696]
$V(\gamma_i)$	0,503	0,463	0,411	[0,181; 0,887]	3,272	3,087	2,927	[1,654; 5,551]
$\alpha$	-0,033	-0,017	-0,353	[-9,427; 9,533]	0,049	0,196	0,467	[-10,126; 9,643]

Analisando a Tabela 4.5 podemos perceber que utilizar o modelo SNI, modelo utilizado na geração dos dados, em geral não houve melhora na estimação dos parâmetros. Se compararmos os modelos ajustados, tal modelo obteve um melhor comportamento na estimação de  $\beta_0$  apenas para o banco de dados 3, ao considerar qualquer um dos estimadores e para o banco de dados 4 quando consideramos a média *a posteriori*. Avaliando a estimação de  $\beta_1$ , o modelo SND foi o melhor para o banco de dados 1, o modelo SNI foi um pouco melhor para o banco de dados 2 e o modelo normal foi

o melhor para os bancos de dados 3 e 4. Todos os intervalos HPD, com 95% de probabilidade, para  $\beta_0$  e  $\beta_1$  incluíram os valores verdadeiros dos parâmetros e, em geral, o modelo SND apresentou os intervalos HPD com menor amplitude. De forma geral, o modelo que melhor estimou a variância dos efeitos aleatórios foi o modelo Normal. No entanto, para o banco de dados 3, em nenhum dos modelos o verdadeiro valor de variância dos efeitos aleatórios ficou incluso no intervalo HPD. Quanto ao parâmetro de assimetria  $\lambda$ , o modelo SNI estimou valores positivos em todos os bancos de dados, sendo que as estimativas estiveram mais próximas do valor real nos bancos de dados 1 e 3, bancos que foram gerados com menor variância para a distribuição dos efeitos aleatórios. O modelo SND estimou valores para  $\alpha$  próximos de zero nos quatro bancos de dados e os intervalos HPD são quase simétricos ao redor de zero.

Da mesma forma que na seção anterior, vamos utilizar o SQE para avaliar a estimação dos efeitos aleatórios  $\gamma$ . Na Tabela 4.6 apresentamos os valores obtidos para SQE considerando os modelos Normal, SNI e SND, para os bancos de dados considerados.

Tabela 4.6: Soma de quadrados dos erros para os efeitos aleatórios  $\gamma$

Modelo	Banco de dados			
	1	2	3	4
Normal	8,802	9,034	6,207	3,037
SNI	<b>8,710</b>	<b>7,892</b>	<b>5,924</b>	<b>3,011</b>
SND	8,778	8,958	6,280	3,016

Através da Tabela 4.6 notamos que o modelo utilizado na geração dos dados, o modelo SNI, foi o que melhor estimou os efeitos aleatórios em todos os quatro bancos de dados gerados, pois apresentou o menor valor de SQE em todos os casos.

Na Tabela 4.7 são apresentadas medidas que avaliam a qualidade de ajuste dos modelos utilizados.

Tabela 4.7: Medidas de qualidade de ajuste dos modelos

Modelo	CPO	DIC	Índice C	TCC	CPO	DIC	Índice C	TCC
Banco de dados 1				Banco de dados 2				
Normal	<b>-402,507</b>	<b>804,479</b>	0,899	<b>0,813</b>	-375,230	750,645	<b>0,917</b>	<b>0,826</b>
SNI	-403,451	806,380	<b>0,900</b>	<b>0,813</b>	<b>-374,197</b>	<b>748,489</b>	<b>0,917</b>	<b>0,826</b>
SND	-402,634	804,695	<b>0,900</b>	<b>0,813</b>	-375,545	751,016	<b>0,917</b>	<b>0,826</b>
Banco de dados 3				Banco de dados 4				
Normal	-409,904	819,734	<b>0,897</b>	<b>0,822</b>	<b>-355,920</b>	<b>712,281</b>	<b>0,926</b>	<b>0,842</b>
SNI	<b>-408,421</b>	<b>816,824</b>	<b>0,897</b>	0,821	-356,305	712,847	<b>0,926</b>	<b>0,842</b>
SND	-409,978	819,902	<b>0,897</b>	<b>0,822</b>	-356,634	713,605	<b>0,926</b>	<b>0,842</b>

Analisando a Tabela 4.7 vemos que o modelo utilizado na geração dos dados, modelo SNI, ajustou melhor aos dados quando consideramos os critérios de CPO e DIC nos bancos de dados 2 e 3. Utilizando os mesmos critérios, o modelo Normal foi melhor para os demais bancos de dados. Note que o modelo SNI foi o que pior se ajustou aos dados no banco de dados 1, banco em que os efeitos foram gerados considerando menores valores para variância e para assimetria. O índice C e a taxa de classificação correta pouco discriminaram os modelos. Estas medidas apresentam diferenças mínimas apenas no banco de dados 1 em que, segundo o índice C, o modelo normal

teve o pior ajuste aos dados, e no banco de dados 2 em que, de acordo com a taxa de classificação correta, o modelo SNI teve o pior ajuste.

Para avaliar a estimação da razão de chances vamos considerar  $OR_1$ ,  $OR_2$  e  $OR_3$  como descritos na Seção 4.1. Os valores reais para OR nestas situações e as respectivas estimativa com cada modelo são apresentados na Tabela 4.8.

Tabela 4.8: Estimativas *a posteriori* para a razão de chances

	Real	Me	Md	Mo	HPD	Real	Me	Md	Mo	HPD
	Banco de dados 1					Banco de dados 2				
modelo Normal										
$OR_1$	6,079	1,738	1,513	1,182	[0,397; 3,691]	131,824	93,129	33,447	21,876	[2,230; 310,365]
$OR_2$	12,182	10,853	10,626	10,126	[8,058; 14,467]	12,182	11,306	11,066	10,344	[7,916; 15,103]
$OR_3$	74,063	18,892	16,446	12,397	[4,074; 41,594]	1605,951	1078,415	374,202	182,784	[19,351; 3681,392]
modelo SNI										
$OR_1$	6,079	1,792	1,519	1,215	[0,409; 3,919]	131,824	349,914	66,394	193,157	[1,816; 596,061]
$OR_2$	12,182	10,827	10,610	10,362	[7,821; 14,329]	12,182	11,342	11,168	10,945	[8,009; 15,014]
$OR_3$	74,063	19,489	16,427	12,249	[3,680; 42,586]	1605,951	4453,537	745,051	3063,459	[13,395; 6724,695]
modelo SND										
$OR_1$	6,079	1,736	1,526	1,201	[0,366; 3,685]	131,824	86,555	38,452	21,570	[1,728; 300,228]
$OR_2$	12,182	10,908	10,790	10,548	[7,976; 14,134]	12,182	11,307	11,140	10,662	[8,281; 15,202]
$OR_3$	74,063	18,985	16,415	13,028	[3,173; 40,946]	1605,951	1012,204	434,816	213,705	[13,130; 3575,617]
	Banco de dados 3					Banco de dados 4				
modelo Normal										
$OR_1$	1,189	1,178	0,995	0,799	[0,209; 2,523]	6,103	5,389	4,159	3,066	[0,729; 13,104]
$OR_2$	12,182	11,418	11,242	11,210	[8,283; 15,254]	12,182	14,005	13,566	12,930	[9,512; 19,021]
$OR_3$	14,488	13,526	10,963	8,299	[3,145; 31,143]	74,350	76,693	57,448	40,815	[5,551; 187,993]
modelo SNI										
$OR_1$	1,189	1,073	1,002	0,929	[0,363; 1,894]	6,103	5,337	4,230	2,859	[0,808; 13,229]
$OR_2$	12,182	11,196	10,984	10,561	[8,011; 14,518]	12,182	14,094	13,748	13,227	[9,537; 19,603]
$OR_3$	14,488	12,006	11,008	9,601	[4,384; 22,754]	74,350	76,404	60,920	33,977	[9,537; 189,208]
modelo SND										
$OR_1$	1,189	1,148	1,007	0,716	[0,197; 2,472]	6,103	5,447	4,292	2,803	[0,582; 13,607]
$OR_2$	12,182	11,406	11,238	11,206	[8,116; 15,179]	12,182	14,051	13,658	12,740	[9,537; 19,490]
$OR_3$	14,488	13,159	11,213	8,579	[1,914; 28,944]	74,350	77,876	59,124	37,739	[6,825; 201,200]

Vemos na Tabela 4.8 que os três exemplos de razão de chances foram bem estimados com qualquer dos modelos apenas nos bancos de dados 3 e 4. Em qualquer dos bancos de dados, as estimativas para  $OR_2$  são melhores do que as estimativas para os outros exemplos de razão de chances. Isto se deve ao fato de que  $OR_2$  depender apenas de  $\beta_0$  e  $\beta_1$ , parâmetros que foram melhor estimados que os efeitos aleatórios. Para o banco de dados 1, o modelo SNI estimou melhor  $OR_1$  e  $OR_3$  quando o estimador considerado foi a média *a posteriori*. O modelo SND foi o melhor para estimar  $OR_1$  com a mediana *a posteriori* e  $OR_3$  com a moda *a posteriori*. Em geral, os intervalos HPD dos três modelos foram próximos e o valor real para  $OR_1$  e  $OR_3$  não foram incluídos entre os limites dos intervalos, caracterizando uma estimação intervalar ruim. Com o banco de dados 2 percebemos que o modelo SNI parece ter estimado melhor  $OR_2$ , independente do estimador considerado. Quanto a  $OR_1$  e  $OR_3$  percebemos que suas estimativas pontuais foram ruins com todos os três modelos, mas as estimativas intervalares incluíram os valores reais. Para o banco de dados 3 vemos que o estimador com estimativas mais próximas dos valores reais é a média *a posteriori*, e com este estimador o modelo Normal é o melhor para estimar  $OR_1$ ,  $OR_2$  e  $OR_3$ . Os intervalos HPD dos três modelos contiveram os valores reais e o modelo SNI forneceu os intervalos mais precisos. Para o banco de dados 4, novamente a estimação foi melhor ao considerar a média *a posteriori*. De forma geral o modelo SND foi o melhor para estimar  $OR_1$ , o modelo Normal foi

o melhor para estimar  $OR_2$  e o modelo SNI foi o melhor para estimar  $OR_3$ . Os intervalos HPD estimaram bem a razão de chances nos casos abordados. Em resumo podemos dizer que utilizar o modelo SNI, que foi considerado na geração dos dados, não melhorou a estimação da razão de chances.

### 4.3 Efeitos Aleatórios com Distribuição Normal Assimétrica Multivariada

Nesta seção apresentamos o estudo de simulação com dados gerados de um modelo logístico misto com efeitos aleatórios com distribuição normal assimétrica multivariada. Neste estudo, foram gerados quatro banco de dados, sendo que estes bancos diferem nos valores usados para o parâmetro de assimetria  $\alpha$  e para a variância da distribuição utilizada para gerar dos efeitos aleatórios. Os valores especificados para o parâmetro de escala  $\sigma^2$  foram escolhidos da forma a determinar que  $V(\gamma_i)$  fosse igual a 1 nos banco de dados 1 e 3, e igual a 4 nos bancos de dados 2 e 4. O parâmetro de locação da distribuição normal assimétrica univariada foi escolhido para que  $E(\gamma_i) = 0$ . Com as especificações utilizadas, em todos os quatro bancos de dados, o vetor de efeitos aleatórios foi gerado com correlação entre dois componentes genéricos do vetor sendo aproximadamente -0,026. Esta correlação parece ser muito pequena, mas é um valor próximo a maior correlação negativa possível entre dois componentes genéricos do vetor quando  $k = 25$ , o valor para o parâmetro de assimetria da distribuição normal assimétrica é a mesmo para todos os componentes, ou seja, consideramos  $\alpha = \alpha \mathbf{1}_{25}$ , e a matriz de escala é diagonal. As distribuições utilizadas para gerar os efeitos aleatórios  $\gamma$  foram

**Banco de dados 1:**  $\gamma \sim SN_k(-0,162; 1,026; 5);$

**Banco de dados 2:**  $\gamma \sim SN_k(-0,323; 4,104; 5);$

**Banco de dados 3:**  $\gamma \sim SN_k(-0,162; 1,026; 10);$

**Banco de dados 4:**  $\gamma \sim SN_k(-0,323; 4,104; 10).$

Na Figura 4.3 apresentamos o histograma dos efeitos aleatórios gerados. Note que neste caso não é possível traçar a função de densidade teórica juntamente com o histograma pelo fato de se tratar de uma distribuição multivariada. Sendo assim, neste caso, o histograma não serve para avaliar a qualidade da geração dos efeitos aleatórios e é útil apenas para ilustrar a presença de assimetria nos dados.

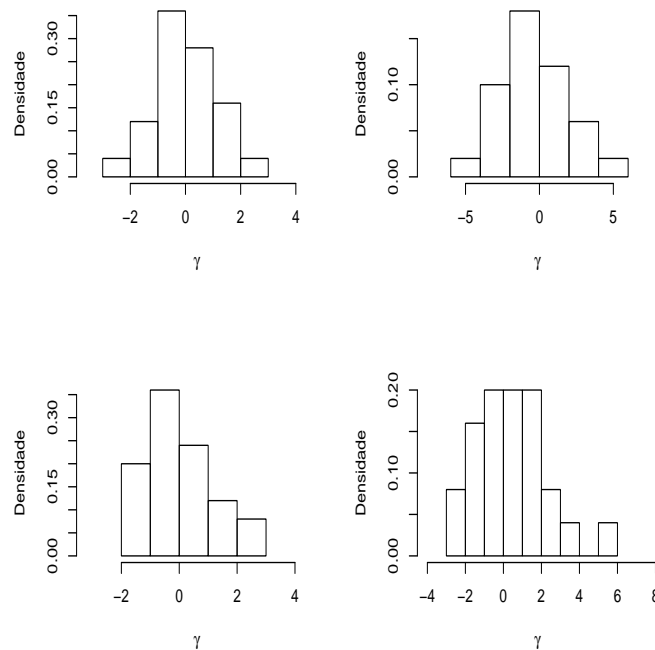


Figura 4.3: Histograma dos efeitos aleatórios  $\gamma$  gerados para os bancos de dados 1 (esquerda em cima), 2 (direita em cima), 3 (esquerda em baixo) e 4 (direita em baixo).

A Tabela 4.9 apresenta as estimativas, obtidas nos três modelos que foram considerados, para os parâmetros de efeitos fixos  $\beta$ , para os parâmetros da distribuição assumida para os efeitos aleatórios e para a variância dos efeitos aleatórios. Analisando esta tabela percebemos que utilizar o modelo utilizado na geração dos dados, ou seja, utilizar o modelo SND, não está melhorando a estimação dos parâmetros de efeitos fixos  $\beta$ . Apenas no banco de dados 2 o modelo SND foi o que melhor estimou os parâmetros  $\beta_0$  e  $\beta_1$ . Além disso notamos que no banco de dados 4 todos os modelos estimaram muito mal  $\beta_0$  e  $\beta_1$ . Assim como nas seções anteriores, para todos os bancos de dados e em todos os modelos, os intervalos HPD com 95% de probabilidade para  $\beta_0$  e  $\beta_1$  incluíram os verdadeiros valores e o modelo SND é o modelo que fornece os intervalos HPD com menor amplitude para  $\beta_0$ . Quando avaliamos a estimação da variância dos efeitos aleatórios, se utilizamos a média *a posteriori*, percebemos que o modelo SND teve melhor desempenho nos bancos de dados 1, 2 e 3. Se utilizarmos a mediana *a posteriori*, o modelo SND é o melhor apenas nos bancos de dados 1 e 2. Utilizando a moda *a posteriori*, o modelo SND tem melhor desempenho nos bancos de dados 1 e 3. Para a variância dos efeitos aleatórios, em todos os bancos de dados e com todos os modelos, o verdadeiro valor ficou contido nos intervalos HPD. O modelo SNI estimou valores positivos para o parâmetro de variância  $\lambda$ , estimando, com a média *a posteriori*, os valores de 1,001 e 2,593 nos bancos de dados 1 e 3, respectivamente. Para os demais bancos de dados os valores estimados foram próximos de zero. Aparentemente, utilizar valores altos para  $\alpha$  na geração do vetor de efeitos aleatórios  $\gamma$  com distribuição normal assimétrica multivariada não significa ter valores altos para o parâmetro de assimetria  $\lambda$  quando consideramos a distribuição marginal de  $\gamma_i$ . Apesar dos verdadeiros valores de  $\alpha$  serem 5 (nos bancos de dados 1 e 2) e 10 (nos bancos de dados 3 e 4),

percebemos que o modelo SND sempre estima valores de  $\alpha$  próximos de zero, ou seja, o modelo SND não está sendo capaz de estimar assimetria.

Tabela 4.9: Estimativas *a posteriori* para os parâmetros

	Me	Md	Mo	HPD	Me	Md	Mo	HPD
	Banco de dados 1 - $\alpha = 5$ e $V(\gamma_i) = 1$				Banco de dados 2 - $\alpha = 5$ e $V(\gamma_i) = 4$			
modelo Normal								
$\beta_0$	0,619	0,620	0,608	[0,165; 1,161]	0,287	0,280	0,369	[-0,616; 1,556]
$\beta_1$	2,598	2,597	2,594	[2,278; 2,947]	2,660	2,657	2,679	[2,292; 3,019]
$\sigma^2$	1,342	1,251	1,153	[0,499; 2,217]	6,141	5,740	4,911	[2,734; 10,151]
modelo SNI								
$\beta_0$	0,628	0,630	0,640	[0,159; 1,118]	0,278	0,265	0,318	[-0,742; 1,338]
$\beta_1$	2,602	2,597	2,587	[2,290; 2,964]	2,655	2,645	2,639	[2,301; 3,006]
$\sigma^2$	1,793	1,545	1,234	[0,552; 3,639]	7,467	6,740	5,624	[3,035; 14,257]
$V(\gamma_i)$	1,347	1,259	1,094	[0,605; 2,390]	6,429	6,014	5,263	[2,698; 11,174]
$\lambda$	1,001	0,255	0,064	[-1,540; 6,380]	0,167	0,110	0,090	[-0,972; 1,287]
modelo SND								
$\beta_0$	0,633	0,631	0,615	[0,281; 0,946]	0,302	0,299	0,210	[-0,400; 0,875]
$\beta_1$	2,601	2,593	2,582	[2,294; 2,978]	2,645	2,641	2,631	[2,263; 2,998]
$\sigma^2$	1,325	1,239	1,093	[0,533; 2,249]	6,146	5,806	5,651	[2,714; 10,204]
$V(\gamma_i)$	1,292	1,210	1,065	[0,520; 2,191]	5,997	5,662	5,496	[2,662; 9,964]
$\alpha$	0,044	-0,096	-0,672	[-9,533; 9,998]	0,276	0,155	-1,049	[-9,127; 9,533]
	Banco de dados 3 - $\alpha = 10$ e $V(\gamma_i) = 1$				Banco de dados 4 - $\alpha = 10$ e $V(\gamma_i) = 4$			
modelo Normal								
$\beta_0$	0,587	0,592	0,659	[0,091; 1,139]	1,012	1,007	0,934	[0,331; 1,850]
$\beta_1$	2,568	2,571	2,583	[2,228; 2,913]	2,952	2,948	2,955	[2,580; 3,366]
$\sigma^2$	1,412	1,335	1,267	[0,635; 2,345]	3,082	2,873	2,532	[1,318; 5,373]
modelo SNI								
$\beta_0$	0,548	0,550	0,597	[0,012; 1,051]	0,990	0,970	0,925	[0,303; 1,739]
$\beta_1$	2,556	2,552	2,542	[2,242; 2,898]	2,952	2,943	2,910	[2,588; 3,387]
$\sigma^2$	2,123	1,755	1,471	[0,505; 4,530]	3,928	3,446	2,957	[1,402; 7,910]
$V(\gamma_i)$	1,382	1,288	1,227	[0,578; 2,409]	3,124	2,966	2,735	[1,397; 5,199]
$\lambda$	2,593	0,350	0,024	[-1,594; 14,521]	0,679	0,128	-0,022	[-1,431; 3,785]
modelo SND								
$\beta_0$	0,572	0,569	0,586	[0,242; 0,936]	1,006	1,005	1,059	[0,520; 1,420]
$\beta_1$	2,564	2,563	2,581	[2,203; 2,865]	2,947	2,946	2,934	[2,563; 3,322]
$\sigma^2$	1,410	1,326	1,169	[0,632; 2,400]	3,035	2,842	2,523	[1,292; 5,179]
$V(\gamma_i)$	1,376	1,296	1,132	[0,602; 2,327]	2,961	2,774	2,477	[1,267; 5,064]
$\alpha$	-0,100	-0,174	-0,810	[-10,689; 8,943]	0,019	0,049	1,012	[-9,876; 9,426]

A Tabela 4.10 apresenta os valores de SQE para os efeitos aleatórios obtidos com os modelos Normal, SNI e SND nos bancos de dados gerados com distribuição normal assimétrica para os efeitos aleatórios.

Tabela 4.10: Soma de quadrados dos erros para os efeitos aleatórios  $\gamma$

Modelo	Banco de dados			
	1	2	3	4
Normal	<b>3,382</b>	13,034	3,329	13,787
SNI	3,504	13,406	<b>3,231</b>	<b>12,880</b>
SND	3,406	<b>12,866</b>	3,324	13,478

Ao analisar a Tabela 4.10 vemos que o modelo SND, modelo com o qual os dados foram gerados, foi o melhor quanto a estimação dos efeitos aleatórios apenas no banco de dados 2. O modelo Normal foi o melhor no banco de dados 1 e o modelo SNI obteve melhores estimativas nos demais bancos de dados, bancos de dados em que a assimetria da distribuição dos efeitos aleatórios é maior.

A Tabela 4.11 apresenta as medidas de CPO, DIC, índice C e taxa de classificação correta para os modelos Normal, SNI e SND em todos os bancos de dados gerados com distribuição normal assimétrica multivariada para os efeitos aleatórios.

Tabela 4.11: Medidas de qualidade de ajuste dos modelos

Modelo	CPO	DIC	Índice C	TCC	CPO	DIC	Índice C	TCC
Banco de dados 1				Banco de dados 2				
Normal	-375,949	751,854	<b>0,915</b>	<b>0,832</b>	-320,298	640,736	<b>0,942</b>	0,863
SNI	-376,628	753,186	<b>0,915</b>	<b>0,832</b>	-320,139	640,222	<b>0,942</b>	0,863
SND	<b>-375,711</b>	<b>751,418</b>	<b>0,915</b>	<b>0,832</b>	<b>-319,914</b>	<b>639,836</b>	<b>0,942</b>	<b>0,864</b>
Banco de dados 3				Banco de dados 4				
Normal	<b>-385,564</b>	<b>770,811</b>	<b>0,912</b>	<b>0,829</b>	-333,121	666,036	<b>0,934</b>	<b>0,859</b>
SNI	-386,070	771,939	<b>0,912</b>	<b>0,829</b>	<b>-332,753</b>	<b>665,218</b>	<b>0,934</b>	<b>0,859</b>
SND	-385,745	771,086	<b>0,912</b>	0,828	-333,155	666,164	<b>0,934</b>	<b>0,859</b>

Através da Tabela 4.11, notamos que ao adotar os critérios de CPO e DIC o modelo SND teve o melhor ajuste nos bancos de dados 1 e 2, bancos de dados em que os efeitos aleatórios foram gerados com menor assimetria. O modelo normal apresentou melhor ajuste no banco de dados 3, enquanto o modelo SNI foi o melhor para o banco de dados 3. Ao utilizar o índice C, em todos os bancos de dados os modelos foram iguais. Se consideramos a taxa de classificação correta dizemos que em todos os bancos de dados praticamente não houve diferença entre os modelos.

A avaliação da estimação da razão de chances será realizada da mesma forma que nas Seções 4.2 e 4.3. Os casos abordados para OR são os descritos na Seção 4.1. Para cada um dos quatro bancos de dados analisados nesta seção, os valores reais e as estimativas *a posteriori* para razão de chances nos casos abordados são apresentadas na Tabela 4.12.

Tabela 4.12: Estimativas *a posteriori* para a razão de chances

	Real	Me	Md	Mo	HPD	Real	Me	Md	Mo	HPD
	Banco de dados 1					Banco de dados 2				
	modelo Normal									
$OR_1$	4,557	5,189	4,306	3,124	[0,926; 11,534]	0,241	0,159	0,130	0,097	[0,018; 0,382]
$OR_2$	12,182	13,633	13,428	13,111	[9,305; 18,530]	12,182	14,543	14,259	14,389	[9,615; 20,173]
$OR_3$	55,515	71,741	58,478	39,072	[9,220; 166,119]	2,940	2,286	1,840	1,265	[0,313; 5,284]
	modelo SNI									
$OR_1$	4,557	5,073	4,288	3,528	[0,994; 11,635]	0,241	0,165	0,139	0,085	[0,021; 0,390]
$OR_2$	12,182	13,698	13,423	13,099	[9,352; 18,575]	12,182	14,465	14,087	13,891	[9,890; 20,039]
$OR_3$	55,515	70,623	57,923	39,362	[8,685; 163,120]	2,940	2,372	1,908	1,368	[0,293; 5,763]
	modelo SND									
$OR_1$	4,557	5,302	4,427	3,058	[0,945; 11,641]	0,241	0,161	0,133	0,091	[0,022; 0,383]
$OR_2$	12,182	13,681	13,364	13,077	[9,235; 18,615]	12,182	14,360	14,024	13,577	[9,568; 20,002]
$OR_3$	55,515	74,177	59,023	42,309	[11,115; 175,292]	2,940	2,284	1,886	1,405	[0,370; 5,520]
	Banco de dados 3					Banco de dados 4				
	modelo Normal									
$OR_1$	0,802	0,520	0,440	0,342	[0,099; 1,191]	0,112	0,061	0,050	0,031	[0,008; 0,145]
$OR_2$	12,182	13,244	13,073	12,878	[8,677; 17,643]	12,182	19,537	19,062	18,582	[12,277; 27,265]
$OR_3$	9,769	6,805	5,807	4,097	[1,088; 15,711]	1,368	1,156	0,935	0,613	[0,144; 2,827]
	modelo SNI									
$OR_1$	0,802	0,534	0,457	0,394	[0,104; 1,157]	0,112	0,063	0,050	0,032	[0,006; 0,151]
$OR_2$	12,182	13,064	12,838	12,610	[9,076; 17,657]	12,182	19,556	18,964	17,884	[12,592; 28,065]
$OR_3$	9,769	6,934	5,867	4,690	[1,331; 15,688]	1,368	1,193	0,954	0,630	[0,171; 2,899]
	modelo SND									
$OR_1$	0,802	0,505	0,418	0,336	[0,110; 1,150]	0,112	0,061	0,048	0,033	[0,005; 0,148]
$OR_2$	12,182	13,183	12,979	11,747	[8,679; 17,153]	12,182	19,408	19,031	18,600	[12,671; 27,149]
$OR_3$	9,769	6,618	5,364	4,248	[1,127; 14,879]	1,368	1,159	0,915	0,585	[0,144; 2,840]

Vemos na Tabela 4.12 que os três modelos ajustados fornecem estimativas próximas para a razão de chances no casos abordados, mas em alguns casos estas estimativas são distantes dos valores verdadeiros. Além disso, em poucos casos utilizar o modelo SND, com o qual foram gerados os dados, trouxe uma melhoria na estimação da razão de chances. Em geral, as melhores estimativas são obtidas com os modelos Normal e SNI, variando de acordo com o banco de dados e com o estimador *a posteriori* considerado. Considerando a média *a posteriori*, o modelo SNI obteve estimativas um pouco melhores para  $OR_1$  e  $OR_3$ . No entanto parece não haver um padrão definindo em que situação cada modelo é melhor. Quanto aos intervalos HPD com 95% de probabilidade, notamos que estes contém os valores verdadeiros em cada um dos casos considerados.



# Capítulo 5

## Aplicação a Dados Reais

Como aplicação a dados reais, consideramos os dados apresentados em Liu e Dey (2008), descritos como obtidos em um experimento de atividade teratogênica conduzido por Landauer e Clark (1962). Neste exemplo, a proposta é estudar a interação entre a atividade teratogênica de dois análogos de niacina 6-aminonicotinamida (*AN6*) e 3-acetilpyridina (*AP3*). Todos os ovos utilizados no experimento vieram de frangos da linhagem *White Leghorn* do Departamento de Ciência Galinácea da Universidade de Connecticut, Storrs. Os ovos foram incubados em uma incubadora com mecanismo giratório automático. Todas as injeções foram dadas 96 horas após a incubação e foram aplicadas dentro da gema de cada ovo. As soluções de 6AN e 3AP foram feitas em água desmineralizada e esterizada por filtração através de filtros bacteriológicos ultrafinos. Depois, ambos os componentes foram injetados ao mesmo tempo e dissolvidos juntos. O volume de fluido injetado foi de 0,05 ml em todos os casos. Os experimentos foram realizados em quatro meses independentes e, assim como em Liu e Dey (2008), vamos tratar estes meses como *clusters*. Os dados consistem de 2154 observações. Os dados são apresentados na Tabela 5.1, na qual as observações estão separadas em *slots*. Cada *slot* contém  $n$  observações, sendo em que  $r$  delas observou-se frangos anormais ( $y = 1$ ).

Tabela 5.1: Dados experimentais. (Tabela Reformulada)

Data	Slot	Dosagem AN6	Dosagem AP3	Sobreviventes ( $n$ )	Frangos anormais ( $r$ )	$r/n$
Nov. 1961	(1)	2,5	0	110	17	0,155
	(2)	2,5	375	108	28	0,259
	(3)	0	375	107	25	0,234
	(4)	2,5	500	107	43	0,402
	(5)	0	500	103	36	0,350
Dez. 1961	(1)	4	0	98	22	0,224
	(2)	4	375	102	39	0,382
	(3)	4	500	106	80	0,755
Jan. 1962	(1)	10	0	203	198	0,975
	(2)	10	500	175	153	0,875
	(3)	0	500	180	90	0,500
Abr. 1962	(1)	5	0	156	93	0,596
	(2)	5	375	152	41	0,270
	(3)	0	375	148	32	0,216
	(4)	5	500	151	106	0,702
	(5)	0	500	148	70	0,473

Neste estudo, ao ajustar os modelos, consideramos as covariáveis *AN6*, *AP3*, a interação entre elas e o termo de intercepto. O preditor linear considerado para os modelos ajustados é dado

$\eta_{ij} = \beta_0 + \beta_1 AN6_{ij} + \beta_2 AP3_{ij} + \beta_3 (AN6 \times AP3)_{ij} + Z_{ij}\gamma$ , em que  $i = 1, \dots, 4$ , representa o mês em que a observação foi coletada *cluster*,  $j$  representa a ordem da observação nesse mês,  $\gamma = (\gamma_1, \gamma_2, \gamma_3, \gamma_4)'$  e  $Z_{ij}$  é um vetor, de ordem  $4 \times 1$ , contendo o valor um na posição equivalente a ordem do mês que foi observado este indivíduo e zero na demais posições. Exemplificando, se a observação  $y_{ij}$  foi coletada em dezembro de 1961 (segundo mês do estudo), temos  $i = 2$  e  $Z_{2j} = (0, 1, 0, 0)'$ . Além disso, transformamos as covariáveis  $AN6$  e  $AP3$  para que tivessem média e variância iguais a zero e um, respectivamente. Neste trabalho as observações para a variável resposta são tratadas como observações de uma variável aleatória Bernoulli, enquanto em Liu e Dey (2008) trabalha-se supondo distribuição binomial para a variável resposta, ou seja, neste último cada *slot* é visto como sendo apenas um valor observado. Em decorrência disto, apesar de que em ambos os trabalhos foram utilizadas transformações para padronizar as variáveis explicativas, estas transformações levam a diferentes valores para as variáveis explicativas. Por isso as estimativas obtidas nesta dissertação para os parâmetros de efeitos fixos  $\beta$  não são comparáveis às estimativas obtidas em Liu e Dey (2008).

As distribuição *a priori* utilizadas foram as mesmas do estudo de simulação, a saber, consideramos  $\beta_i \sim N(0, 10)$ , para  $i = 0, \dots, 3$ ,  $\sigma^2 \sim GI(2, 001; 1)$ ,  $\lambda \sim N(0, 1000)$  e  $\alpha \sim N(0, 1000)$ . Na implementação dos modelos consideramos os mesmos valores para *burn in*, *lag* e tamanho das amostras *a posteriori* que foram utilizados no estudo de simulação. A análise de convergência também foi realizada utilizando as mesmas ferramentas que foram citadas no Capítulo 4.

A Tabela 5.2 apresenta as estimativas *a posteriori* obtidas para média (Me), mediana (Md) e moda (Mo) das distribuições marginais *a posteriori* de cada um dos parâmetros, além dos respectivos intervalos de mais alta densidade (HPD) com 95% de probabilidade.

Tabela 5.2: Estimativas *a posteriori* para os parâmetros

	Me	Md	Mo	HPD		Me	Md	Mo	HPD
modelo Normal									
$\beta_0$	-0,093	-0,075	-0,028	[-0,683; 0,539]	$\gamma_1$	-0,358	-0,377	-0,432	[-0,908; 0,334]
$\beta_1$	1,154	1,150	1,130	[ 0,998; 1,330]	$\gamma_2$	-0,082	-0,095	-0,080	[-0,693; 0,540]
$\beta_2$	0,479	0,480	0,503	[ 0,360; 0,599]	$\gamma_3$	0,388	0,378	0,412	[-0,234; 1,079]
$\beta_3$	-0,759	-0,757	-0,734	[-0,968;-0,581]	$\gamma_4$	-0,113	-0,125	-0,110	[-0,766; 0,455]
$\sigma^2$	0,444	0,370	0,267	[ 0,091; 0,979]					
modelo SNI									
$\beta_0$	-0,089	-0,089	-0,077	[-0,474; 0,425]	$\gamma_1$	-0,345	-0,346	-0,344	[-0,854; 0,105]
$\beta_1$	1,166	1,163	1,157	[ 0,992; 1,324]	$\gamma_2$	-0,090	-0,084	-0,034	[-0,596; 0,369]
$\beta_2$	0,479	0,480	0,496	[ 0,357; 0,618]	$\gamma_3$	0,368	0,364	0,364	[-0,171; 0,827]
$\beta_3$	-0,762	-0,790	-0,567	[-0,968;-0,567]	$\gamma_4$	-0,110	-0,104	-0,084	[-0,562; 0,370]
$\sigma^2$	0,533	0,442	0,304	[ 0,121; 1,258]					
$V(\gamma_i)$	0,226	0,185	0,129	[ 0,045; 0,562]					
$\lambda$	2,169	2,255	2,914	[-15,694;21,017]					
modelo SND									
$\beta_0$	-0,122	-0,134	-0,092	[-0,550; 0,371]	$\gamma_1$	-0,329	-0,315	-0,291	[-0,797; 0,150]
$\beta_1$	1,156	1,154	1,128	[ 0,978; 1,323]	$\gamma_2$	-0,056	-0,043	-0,013	[-0,517; 0,446]
$\beta_2$	0,477	0,478	0,472	[ 0,354; 0,597]	$\gamma_3$	0,422	0,425	0,410	[-0,017; 0,944]
$\beta_3$	-0,787	-0,753	-0,759	[-0,966;-0,580]	$\gamma_4$	-0,082	-0,069	-0,017	[-0,581; 0,326]
$\sigma^2$	0,476	0,354	0,273	[ 0,096; 1,054]					
$V(\gamma_i)$	0,405	0,301	0,237	[ 0,096; 0,924]					
$\alpha$	-0,812	-0,766	-1,932	[-21,200;16,849]					

Analisando a Tabela 5.2 vemos que as estimativas dadas por cada modelo para os parâmetros  $\beta$  são ligeiramente diferentes. Para  $\beta_0$ , vemos que o modelo SND estimou um efeito um pouco

maior com todos os estimadores e as estimativas foram negativas em todos os modelos. Para esse parâmetro o modelo SNI apresentou intervalo HPD com menor amplitude. As estimativas pontuais para  $\beta_1$  foram bem próximas com os modelos Normal e SND. O efeito estimado foi positivo, implicando que um aumento na covariável *AN6* aumenta a probabilidade dos frangos nascerem anormais. Para  $\beta_2$  as estimativas foram muito próximas com todos os modelos e foram positivas, implicando em efeito de aumento da probabilidade dos frangos nascerem anormais para a covariável *AP3*. As estimativas para a interação entre *AN6* e *AP3* foram negativas, sinalizando que a interação entre estas covariáveis diminui a probabilidade dos frangos nascerem anormais. O modelo Normal estimou valores maiores para a variância dos efeitos aleatórios com todos os estimadores, inclusive com o intervalo HPD, enquanto o modelo SNI estimou os menores valores. Com a média *a posteriori*, o modelo SNI estimou o valor de 2,169 para  $\lambda$  e o intervalo HPD estimado foi [-15,694;21,017]. Isto indica que a distribuição dos efeitos aleatórios é assimétrica, mas a assimetria parece não ser muito alta. O modelo SND estima, com a média *a posteriori*, o valor de -0,812 para  $\alpha$ . Para este parâmetro, o intervalo HPD estimado é [-21,200;16,849]. Para os efeitos aleatórios, notamos que as estimativas obtidas nos três modelos são um pouco diferentes. Os modelos estimou valor positivo para  $\gamma_3$  e valores negativos para  $\gamma_1$ ,  $\gamma_2$  e  $\gamma_4$ , ou seja, os *clusters* 1, 2 e 4 tem um efeito de diminuir a probabilidade do frango ser anormal. Se mantermos fixos os valores da covariáveis, os frangos do *cluster 3* são os mais prováveis de apresentarem anormalidade, enquanto os frangos do *cluster 1* são os menos prováveis. Visto que o modelo Normal apresenta as maiores estimativas para a variância dos efeitos aleatórios, este modelo também apresenta os intervalos HPD para  $\gamma$  com maior amplitude.

Na Tabela 5.3 apresentamos obtidos para as medidas que avaliam a qualidade de ajuste dos modelos implementados.

Tabela 5.3: Medidas de qualidade ajuste dos modelos

Modelo	CPO	DIC	Índice C	TCC
Normal	<b>-1226,553</b>	<b>2453,040</b>	<b>0,735</b>	<b>0,670</b>
SNI	-1226,795	2453,458	<b>0,735</b>	<b>0,670</b>
SND	-1226,701	2453,276	<b>0,735</b>	<b>0,670</b>

Analisando a Tabela 5.3 concluímos que embora o modelo Normal tenha sido indicado como o que melhor ajustou aos dados tanto quando consideramos a medida de CPO quanto quando consideramos o DIC, os modelos são praticamente equivalentes. Este resultado era esperado uma vez que, se não houver assimetria na distribuição dos efeitos aleatórios, ao adotarmos distribuições *a priori* centradas em zero para os parâmetros de assimetria estamos especificado corretamente a distribuição *a priori* para todos os parâmetros e, assim, modelos menos parcimoniosos tenderiam a ter desvantagens. As estimativas *a posteriori* foram iguais em todos os modelos tanto para o índice C quanto para a taxa de classificação correta.

Vamos considerar algumas situações distintas para avaliar a estimação da razão de chances. Entre estas situações vamos considerar a razão entre indivíduos de *clusters* distintos mas com os mesmos valores observados para as covariáveis. Consideramos também a razão de chances entre indivíduos do mesmo *cluster* com diferentes valores observados para as covariáveis e a razão de chances entre indivíduos de *clusters* distintos com diferentes valores observados para covariáveis. A razão de chances será abordada nos seis casos descritos a seguir:

$OR_1$ : razão de chances entre frangos dos *slots* 4 e 5 do mês de novembro de 1961;

$OR_2$ : razão de chances entre frangos dos *slots* 2 e 3 do mês de abril de 1962;

$OR_3$ : razão de chances entre frangos do *slot* 5 do mês de novembro de 1961 e do *slot* 3 do mês de janeiro de 1962;

$OR_4$ : razão de chances entre frangos do *slot* 3 do mês de novembro de 1961 e do *slot* 3 do mês de abril de 1962;

$OR_5$ : razão de chances entre frangos do *slot* 1 do mês de dezembro de 1961 e do *slot* 1 do mês de janeiro de 1962;

$OR_6$ : razão de chances entre frangos do *slot* 2 do mês de dezembro de 1961 e do *slot* 5 do mês de abril de 1962.

Note que as quantidades  $OR_1$  e  $OR_2$  dependem apenas das estimativas de  $\beta$ , uma vez que são comparados frangos do mesmo *cluster*. As quantidades  $OR_3$  e  $OR_4$  dependem somente de  $\gamma$ , pois comparamos frangos com as mesmas covariáveis. Por fim, as quantidades  $OR_5$  e  $OR_6$  dependem tanto de  $\beta$  quanto de  $\gamma$ .

As estimativas *a posteriori* para a razão de chances, nos casos abordados, são apresentadas na Tabela 5.4.

Tabela 5.4: Estimativas *a posteriori* para a razão de chances

	Me	Md	Mo	HPD		Me	Md	Mo	HPD
modelo Normal									
$OR_1$	1.476	1.475	1.468	[1.316; 1.621]	$OR_4$	0.790	0.781	0.744	[0.595; 1.002]
$OR_2$	4.265	4.234	4.164	[3.410; 5.221]	$OR_5$	0.011	0.010	0.008	[0.004; 0.019]
$OR_3$	0.481	0.476	0.458	[0.338; 0.646]	$OR_6$	1.497	1.471	1.421	[1.085; 1.975]
modelo SNI									
$OR_1$	1.486	1.484	1.463	[1.329; 1.631]	$OR_4$	0.796	0.791	0.751	[0.616; 0.998]
$OR_2$	4.334	4.303	4.210	[3.525; 5.356]	$OR_5$	0.011	0.010	0.009	[0.004; 0.019]
$OR_3$	0.497	0.489	0.483	[0.339; 0.648]	$OR_6$	1.497	1.479	1.390	[1.065; 1.954]
modelo SND									
$OR_1$	1.480	1.477	1.478	[1.330; 1.626]	$OR_4$	0.788	0.783	0.808	[0.597; 0.988]
$OR_2$	4.281	4.258	4.240	[3.405; 5.236]	$OR_5$	0.011	0.010	0.009	[0.004; 0.019]
$OR_3$	0.479	0.471	0.460	[0.329; 0.642]	$OR_6$	1.499	1.477	1.437	[1.048; 1.943]

Analisando as estimativas apresentadas na Tabela 5.4 percebemos que os modelos Normal, SNI e SND forneceram estimativas semelhantes para a razão de chances nos casos abordados, principalmente quando o estimador considerado é a média *a posteriori*. As estimativas dos modelos são um pouco mais distantes quando o estimador considerado é a moda *a posteriori*. De forma geral, vemos que as estimativas intervalares também foram muito próximas. As estimativas  $OR_3$  e  $OR_4$ , casos em que a razão de chances depende apenas dos efeitos aleatórios, foram um pouco distantes do valor 1 e os intervalos HPD não incluíram este valor. Por isso,  $OR_3$  indica que, se mantermos fixas as covariáveis, a chance do frango nascer anormal não foi a mesma para os meses de novembro de 1961 e janeiro de 1962. Com  $OR_4$  chegamos a mesma conclusão comparando os meses de novembro de 1961 e abril de 1962 utilizando os modelos SNI e SND. Observe que o intervalo de credibilidade obtido para  $OR_4$  no modelo Normal inclui o valor 1, logo a hipótese  $OR_4 = 1$  pode ser aceita, o que significa que a chance de um frango nascer anormal não é diferente para os

meses de novembro de 1961 e abril de 1962 se os valores das covariáveis são iguais. Em termos de exemplificar a interpretação das estimativas obtidas considere o modelo SNI e a razão de chances  $OR_6$ . Falando informalmente, o valor mais provável para  $OR_6$  é 1,390, pois este é o valor para a moda *a posteriori*. Através da mediana vemos que, *a posteriori*,  $OR_6$  tem 50% de probabilidade ser maior ou igual a 1,479. Com 95% de probabilidade, *a posteriori*,  $OR_6$  está entre [1,065; 1,954]. Para  $OR_6$ , o valor médio *a posteriori* é 1,197. Para os demais modelos e demais casos para a razão de chances temos que a interpretação é a mesma.

# Conclusões e Discussões

Nesta dissertação apresentamos o modelo logístico misto com efeitos aleatórios normalmente distribuídos. Estendemos o modelo logístico misto ao considerar classes estendidas de distribuições normais para os efeitos aleatórios. Esta extensão é realizada considerando as distribuições normal assimétrica univariada e normal assimétrica multivariada propostas por Azzalini (1985) e Azzalini e Dalla-Valle (1996), respectivamente. Ao utilizarmos a distribuição normal assimétrica univariada assumimos que os efeitos aleatórios são independentes enquanto estamos impondo uma estrutura de dependência entre os efeitos aleatórios ao utilizar a distribuição normal assimétrica multivariada. Para realizar a inferência nestes modelos utilizamos a abordagem Bayesiana. Apresentamos a razão de chances mediana proposta por Larsen *et al.* (2000). Tal medida é útil para a interpretação dos parâmetros no modelo logístico misto. Os resultados em Larsen *et al.* (2000) foram obtidos segundo o enfoque clássico, mas também podem ser utilizados para ajudar a eliciar as distribuições *a priori* para os parâmetros. Neste trabalho também estendemos os resultados obtidos por Larsen *et al.* (2000) para os modelos considerando, para os efeitos aleatórios, as distribuições normal assimétrica univariada e multivariada. Avaliamos a implementação dos modelos em um estudo de simulação, em que obtemos estimativas para os parâmetros e razão de chances em alguns bancos de dados gerados. Foram gerados bancos de dados considerando diferentes distribuições para os efeitos aleatórios, diferentes valores para variância e assimetria destas distribuições utilizadas. Por fim, realizamos inferência em um banco de dados reais.

Quanto a extensão da razão de chances mediana proposta por Larsen *et al.* (2000), conseguimos provar resultados muito interessantes. Um deles é o fato de que ao assumirmos a distribuição normal assimétrica multivariada para os efeitos aleatórios, continuam válidos todos os teoremas obtidos ao termos assumido a distribuição normal para os efeitos aleatórios. Outra contribuição interesse é o fato de que, sob certas condições, a combinação linear dos elementos de um vetor com distribuição normal assimétrica tem distribuição normal. Ao assumirmos a distribuição normal assimétrica univariada para os efeitos aleatórios obtivemos a distribuição de probabilidades para a razão de chances e a fórmula analítica para a razão de chances mediana. É interessante lembrar que tais resultados são úteis tanto ao implementarmos os modelos com a abordagem clássica quanto para eliciar as distribuições *a priori* ao adotarmos a abordagem bayesiana.

No estudo com dados simulados verificamos que, de forma geral, o modelo assumindo distribuição normal assimétrica univariada para os efeitos aleatórios foi capaz de detectar assimetria na distribuição dos efeitos aleatórios quando ela existia, e nestes casos, este modelo forneceu estimativas um pouco melhores para os efeitos aleatórios do que os demais modelos. No entanto, de acordo com os critérios de avaliação de ajuste, o modelo SNI foi o melhor modelo ajustado em apenas algumas das situações nas quais havia assimetria.

Vimos também que o modelo logístico misto com efeitos aleatórios normalmente distribuídos forneceu boas estimativas para os parâmetros de efeitos fixos até em situações em que a distribuição verdadeira dos efeitos aleatórios era muito assimétrica. Além disso, apesar do modelo Normal ser

um caso particular dos modelos SNI e SND nos qual os parâmetros de assimetria são iguais a zero, o modelo Normal foi o melhor, de acordo com as medidas de ajuste, quando os dados foram gerados com distribuições sem assimetria para os efeitos aleatórios. De certa forma isso era esperado pelo fato do modelo Normal ser mais parcimonioso do que os demais.

Ainda no estudo com dados simulados, notamos que o modelo logístico misto com distribuição normal assimétrica multivariada para os efeitos aleatórios, modelo em que é imposta uma estrutura de correlação entre os efeitos aleatórios, não foi capaz de fornecer boas estimativas para a assimetria na distribuição dos efeitos. Uma possível explicação para este problema é que, como vimos na Seção 2.2, ao assumirmos tal distribuição *a priori* segue de Azzalini e Dalla-Valle (1996) que as distribuições marginais para cada efeito  $\gamma_i$ , dados  $\sigma^2$  e  $\alpha$ , é  $SN(\xi_i, \sigma^2, \alpha^*)$  em que  $\xi_i = \alpha[1 + k\alpha^2]^{-\frac{1}{2}}\sigma\sqrt{2/\pi}$  e  $\alpha^* = \alpha[1 + (k-1)\alpha^2]^{-\frac{1}{2}}$ . Perceba que se o número de *clusters*  $k$  tende a infinito tal distribuição marginal tenderá a uma distribuição normal. Além disto, essa distribuição conjunta não permite um grau de assimetria marginal muito elevado uma vez que o parâmetro  $\alpha^*$  está contido no intervalo entre  $-(k-1)^{-\frac{1}{2}}$  e  $(k-1)^{-\frac{1}{2}}$ . Exemplificando, em um banco de dados com  $k = 25$ , o parâmetro  $\alpha^*$  terá domínio apenas no intervalo entre -0,204 e 0,204. Percebemos também que o maior valor possível para o parametro de assimetria  $\alpha^*$  ocorre quando  $k = 2$ , situação na qual  $\alpha^* \in (-1, 1)$ .

No capítulo de aplicação a dados reais, verificamos que os modelos apresentaram estimativas relativamente próximas. As estimativas para os parâmetros de efeitos fixos confirmaram a influência dos dois análogos de niacina na formação dos frangos. As estimativas para os efeitos aleatórios confirmaram que, fixados os valores para as covariáveis, ainda assim, frangos com ovos incubados em meses diferentes tiveram chances diferentes de nascerem anormais.

Para trabalhos futuros, sugerimos implementar os modelos utilizados neste trabalho assumindo a distribuição *a priori* de Jeffreys para os parâmetros de efeitos fixos. A justificativa para tal extensão é possibilitar a implementação do modelo logístico misto considerando uma distribuição não informativa para os efeitos fixos que não requer a elicitação de hiperparâmetros e depende apenas da função de verossimilhança. Pensamos também em implementar os modelos considerados aqui utilizando o enfoque da Estatística clássica. Além disso, adotando as abordagens bayesiana e clássica, pretendemos considerar classes de distribuições mais flexíveis para representar o comportamento dos efeitos aleatórios no modelo logístico misto. As classes de distribuições a serem consideradas seriam mais flexíveis no sentido de terem diferentes padrões de assimetria, curtose e caudas. Tais distribuições serão consideradas tanto quando os efeitos aleatórios são independente entre si quanto quando estes são dependentes. Entre as classes a serem consideradas estão as distribuições *t*-assimétrica, normal-*t*-assimétrica e *t*-normal-assimétrica (ver Arellano-Valle *et al.* (2008)). Acredita-se que, em alguns casos, tais classes poderiam representar mais adequadamente o comportamento dos efeitos aleatórios, e então melhorar o ajuste e capacidade de predição do modelo logístico misto aos dados.

# Apêndice A

## Prova da Proposição 2.2.2

Neste apêndice são apresentadas as demonstrações de resultados exibidos na Seção 2.2. Primeiro apresentamos a demonstração para a representação estocástica da distribuição normal assimétrica multivariada apresentada na Proposição 2.2.2 e depois a demonstração da Proposição 2.2.1.

### Demonstração da Proposição 2.2.2.

Inicialmente considere  $T = |Y_0|$ . Visto que  $Y_0 \sim N(0, 1)$ , então  $T \sim HN(0, 1)$ . Condicional ao conhecimento de  $T = t$  temos que

$$\mathbf{Y}|T = t \sim N_k(\boldsymbol{\xi} + \boldsymbol{\Delta}t, \boldsymbol{\Psi}).$$

A distribuição do vetor aleatório  $\mathbf{Y}$  pode ser obtido por

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \int_0^{\infty} f_{\mathbf{Y},T}(\mathbf{y}, t) dt \\ &= \int_0^{\infty} f_{\mathbf{Y}|T=t}(\mathbf{y}) f_T(t) dt \\ &= \int_0^{\infty} 2(2\pi)^{-\frac{k+1}{2}} \det\{\boldsymbol{\Psi}\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \boldsymbol{\xi} - \boldsymbol{\Delta}t)' \boldsymbol{\Omega}^{-1}(\mathbf{y} - \boldsymbol{\xi} - \boldsymbol{\Delta}t) - \frac{t^2}{2}\right\} dt. \end{aligned} \quad (\text{A.1})$$

Fazendo  $\mathbf{z} = \mathbf{y} - \boldsymbol{\xi}$  e substituindo em (A.1) temos:

$$\begin{aligned} f_{\mathbf{Y}}(\mathbf{y}) &= \int_0^{\infty} 2(2\pi)^{-\frac{k+1}{2}} \det\{\boldsymbol{\Psi}\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(\mathbf{z} - \boldsymbol{\Delta}t)' \boldsymbol{\Omega}^{-1}(\mathbf{z} - \boldsymbol{\Delta}t) - \frac{t^2}{2}\right\} dt \\ &= 2(2\pi)^{-\frac{k+1}{2}} \det\{\boldsymbol{\Psi}\}^{-\frac{1}{2}} \int_0^{\infty} \exp\left\{-\frac{1}{2}[(\mathbf{z} - \boldsymbol{\Delta}t)' \boldsymbol{\Omega}^{-1}(\mathbf{z} - \boldsymbol{\Delta}t) + t^2]\right\} dt. \end{aligned} \quad (\text{A.2})$$

Sabendo que  $\boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \mathbf{z} = \mathbf{z}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta}$  e denotando por  $A_1$  o termo  $(\mathbf{z} - \boldsymbol{\Delta}t)' \boldsymbol{\Omega}^{-1}(\mathbf{z} - \boldsymbol{\Delta}t) + t^2$ , obtemos

$$\begin{aligned} A_1 &= (\mathbf{z} - \boldsymbol{\Delta}t)' \boldsymbol{\Omega}^{-1}(\mathbf{z} - \boldsymbol{\Delta}t) + t^2 \\ &= \mathbf{z}' \boldsymbol{\Psi}^{-1} \mathbf{z} - \mathbf{z}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta} t - \boldsymbol{\Delta}' t \boldsymbol{\Psi}^{-1} \mathbf{z} + t^2 \boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta} + t^2 \\ &= t^2(1 + \boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta}) - 2t(\boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \mathbf{z}) + \mathbf{z}' \boldsymbol{\Psi}^{-1} \mathbf{z} \\ &= (1 + \boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta}) \left[ t^2 - 2t \frac{\boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \mathbf{z}}{1 + \boldsymbol{\Delta}' \boldsymbol{\Psi}^{-1} \boldsymbol{\Delta}} \right] + \mathbf{z}' \boldsymbol{\Psi}^{-1} \mathbf{z} \end{aligned}$$



$$\begin{aligned}
&= (1 + \Delta' \Psi^{-1} \Delta) \left[ t^2 - 2t \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta} + \left( \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta} \right)^2 - \left( \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta} \right)^2 \right] + \mathbf{z}' \Psi^{-1} \mathbf{z} \\
&= (1 + \Delta' \Psi^{-1} \Delta) \left( t - \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta} \right)^2 - \frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta} + \mathbf{z}' \Psi^{-1} \mathbf{z}. \tag{A.3}
\end{aligned}$$

Substituindo o termo em (A.3) em (A.2) e tirando fora da integral a parte que não tem  $t$  obtemos

$$\begin{aligned}
f_{\mathbf{Y}}(\mathbf{y}) &= 2(2\pi)^{-\frac{k+1}{2}} \det \{ \Psi \}^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[ -\frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta} + \mathbf{z}' \Psi^{-1} \mathbf{z} \right] \right\} \\
&\quad \times \int_0^\infty \exp \left\{ -\frac{(1 + \Delta' \Psi^{-1} \Delta)}{2} \left( t - \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta} \right)^2 \right\} dt \\
&= 2(2\pi)^{-\frac{k}{2}} \det \{ \Psi \}^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[ -\frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta} + \mathbf{z}' \Psi^{-1} \mathbf{z} \right] \right\} (1 + \Delta' \Psi^{-1} \Delta)^{-\frac{1}{2}} \\
&\quad \times \int_0^\infty (2\pi)^{-\frac{1}{2}} (1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}} \exp \left\{ -\frac{(1 + \Delta' \Psi^{-1} \Delta)}{2} \left( t - \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta} \right)^2 \right\} dt \\
&= 2(2\pi)^{-\frac{k}{2}} \det \{ \Psi \}^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[ -\frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta} + \mathbf{z}' \Psi^{-1} \mathbf{z} \right] \right\} (1 + \Delta' \Psi^{-1} \Delta)^{-\frac{1}{2}} \\
&\quad \times \left[ 1 - \Phi \left( 0 \middle| \frac{\Delta' \Psi^{-1} \mathbf{z}}{1 + \Delta' \Psi^{-1} \Delta}, (1 + \Delta' \Psi^{-1} \Delta)^{-1} \right) \right] \\
&= 2(2\pi)^{-\frac{k}{2}} \det \{ \Psi \}^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[ -\frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta} + \mathbf{z}' \Psi^{-1} \mathbf{z} \right] \right\} (1 + \Delta' \Psi^{-1} \Delta)^{-\frac{1}{2}} \\
&\quad \times \Phi \left( \frac{\Delta' \Psi^{-1} \mathbf{z}}{(1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}}} \right). \tag{A.4}
\end{aligned}$$

Para continuar vamos calcular o termo  $\Psi^{-1}$ . Para isto vamos utilizar uma propriedade do calculo de matrizes inversas que diz que  $[\mathbf{A} - \mathbf{u}\mathbf{v}']^{-1} = \mathbf{A}^{-1} + \frac{\mathbf{A}^{-1}\mathbf{u}\mathbf{v}'\mathbf{A}^{-1}}{1 - \mathbf{v}'\mathbf{A}^{-1}\mathbf{u}}$ , em que  $\mathbf{A}$  é uma matriz de dimensão  $k$  por  $k$ ,  $\mathbf{u}$  e  $\mathbf{v}$  são vetores de ordem  $k$  por 1.

$$\begin{aligned}
\Psi^{-1} &= [\Omega - c^2 \Omega \alpha \alpha' \Omega]^{-1} \\
&= \Omega^{-1} + \frac{c^2 \Omega^{-1} \Omega \alpha \alpha' \Omega \Omega^{-1}}{1 - c^2 \alpha' \Omega \Omega^{-1} \Omega \alpha} \\
&= \Omega^{-1} + \frac{c^2 \alpha \alpha'}{1 - c^2 \alpha' \Omega \alpha}. \tag{A.5}
\end{aligned}$$

Sabendo que  $\Delta = c\Omega\alpha$  e utilizando (A.5), segue que

$$\begin{aligned}
\Delta' \Psi^{-1} &= c\alpha' \Omega \left[ \Omega^{-1} + \frac{c^2 \alpha \alpha'}{1 - c^2 \alpha' \Omega \alpha} \right] \\
&= c \left[ \alpha' + \frac{c^2 \alpha' \Omega \alpha \alpha'}{1 - c^2 \alpha' \Omega \alpha} \right] \\
&= \frac{c\alpha'}{1 - c^2 \alpha' \Omega \alpha}. \tag{A.6}
\end{aligned}$$

Substituindo  $c = (1 + \alpha' \Omega \alpha)^{-\frac{1}{2}}$  em (A.6) temos

$$\begin{aligned} \Delta' \Psi^{-1} &= \frac{(1 + \alpha' \Omega \alpha)^{-\frac{1}{2}} \alpha'}{1 - (1 + \alpha' \Omega \alpha)^{-1} \alpha' \Omega \alpha} \\ &= \frac{(1 + \alpha' \Omega \alpha)^{-\frac{1}{2}} \alpha'}{(1 + \alpha' \Omega \alpha)^{-1}} \\ &= (1 + \alpha' \Omega \alpha)^{\frac{1}{2}} \alpha'. \end{aligned} \quad (\text{A.7})$$

Segue também que

$$\begin{aligned} (1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}} &= \left[ 1 + (1 + \alpha' \Omega \alpha)^{\frac{1}{2}} \alpha' c \Omega \alpha \right]^{\frac{1}{2}} \\ &= \left[ 1 + (1 + \alpha' \Omega \alpha)^{\frac{1}{2}} \alpha' (1 + \alpha' \Omega \alpha)^{-\frac{1}{2}} \Omega \alpha \right]^{\frac{1}{2}} \\ &= \left[ 1 + \alpha' \Omega \alpha \right]^{\frac{1}{2}}. \end{aligned} \quad (\text{A.8})$$

Logo utilizando (A.7) e (A.8) segue que

$$\begin{aligned} \frac{\Delta' \Psi^{-1}}{(1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}}} &= \frac{(1 + \alpha' \Omega \alpha)^{\frac{1}{2}} \alpha'}{\left[ 1 + \alpha' \Omega \alpha \right]^{\frac{1}{2}}} \\ &= \alpha'. \end{aligned} \quad (\text{A.9})$$

Utilizando (A.5) e substituindo  $c = (1 + \alpha' \Omega \alpha)^{-\frac{1}{2}}$  obtemos:

$$\begin{aligned} \Psi^{-1} &= \Omega^{-1} + \frac{(1 + \alpha' \Omega \alpha)^{-1} \alpha \alpha'}{1 - (1 + \alpha' \Omega \alpha)^{-1} \alpha' \Omega \alpha} \\ &= \Omega^{-1} + \frac{(1 + \alpha' \Omega \alpha)^{-1} \alpha \alpha'}{(1 + \alpha' \Omega \alpha - \alpha' \Omega \alpha)(1 + \alpha' \Omega \alpha)^{-1}} \\ &= \Omega^{-1} + \alpha \alpha'. \end{aligned} \quad (\text{A.10})$$

Agora vamos calcular o termo  $\mathbf{z}' \Psi^{-1} \mathbf{z} - \frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta}$ . Para isto, substituímos os termos calculados em (A.9) e (A.10).

$$\begin{aligned} \mathbf{z}' \Psi^{-1} \mathbf{z} - \frac{(\Delta' \Psi^{-1} \mathbf{z})^2}{1 + \Delta' \Psi^{-1} \Delta} &= \mathbf{z}' [\Omega^{-1} + \alpha \alpha'] \mathbf{z} - \frac{(\Delta' \Psi^{-1} \mathbf{z})'}{(1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}}} \frac{\Delta' \Psi^{-1} \mathbf{z}}{(1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}}} \\ &= \mathbf{z}' [\Omega^{-1} + \alpha \alpha'] \mathbf{z} - \mathbf{z}' \frac{(\Delta' \Psi^{-1})'}{(1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}}} \frac{\Delta' \Psi^{-1}}{(1 + \Delta' \Psi^{-1} \Delta)^{\frac{1}{2}}} \mathbf{z} \\ &= \mathbf{z}' \Omega^{-1} \mathbf{z} + \mathbf{z}' \alpha \alpha' \mathbf{z} - \mathbf{z}' \alpha \alpha' \mathbf{z} \\ &= \mathbf{z}' \Omega^{-1} \mathbf{z} \end{aligned} \quad (\text{A.11})$$

Para calcular o termo  $\det \{\Psi\}$  é necessário utilizar a propriedade de determinantes que diz que  $\det \{\mathbf{A} - \mathbf{u} \mathbf{u}'\} = \det \{\mathbf{A}\} (1 - \mathbf{u}' \mathbf{A}^{-1} \mathbf{u})$ , em que  $\mathbf{A}$  é uma matriz de dimensão  $k$  por  $k$ ,  $\mathbf{u}$  é um vetor de ordem  $k$  por 1. Utilizando tal propriedade e considerando  $\mathbf{A} = \Omega$  e  $\mathbf{u} = c \Omega \alpha$  temos que

$$\det \{\Psi\} = \det \{\Omega - c^2 \Omega \alpha \alpha' \Omega\}$$

$$\begin{aligned}
&= \det \{\Omega\} (1 - c^2 \alpha' \Omega \Omega^{-1} \Omega \alpha) \\
&= \det \{\Omega\} (1 - c^2 \alpha' \Omega \alpha) \\
&= \det \{\Omega\} (1 - (1 + \alpha' \Omega \alpha)^{-1} \alpha' \Omega \alpha) \\
&= \det \{\Omega\} (1 + \alpha' \Omega \alpha)^{-1}
\end{aligned} \tag{A.12}$$

Por fim, substituímos os termos calculados em (A.9), (A.11) e (A.12) em (A.4), e fazemos  $\mathbf{z} = \mathbf{y} - \boldsymbol{\xi}$ , obtendo

$$\begin{aligned}
f_{\mathbf{Y}}(\mathbf{y}) &= 2(2\pi)^{-\frac{k}{2}} [\det \{\Omega\} (1 + \alpha' \Omega \alpha)^{-1}]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \mathbf{z}' \Omega^{-1} \mathbf{z} \right\} (1 + \Delta' \Psi^{-1} \Delta)^{-\frac{1}{2}} \Phi(\alpha' \mathbf{z}) \\
&= 2(2\pi)^{-\frac{k}{2}} \det \{\Omega\}^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \mathbf{z}' \Omega^{-1} \mathbf{z} \right\} \Phi(\alpha' \mathbf{z}) \\
&= 2\phi(\mathbf{z} | \mathbf{0}_k, \Omega) \Phi(\alpha' \mathbf{z}) \\
&= 2\phi(\mathbf{y} - \boldsymbol{\xi} | \mathbf{0}_k, \Omega) \Phi(\alpha' (\mathbf{y} - \boldsymbol{\xi})) \\
&= 2\phi(\mathbf{y} | \boldsymbol{\xi}, \Omega) \Phi(\alpha' (\mathbf{y} - \boldsymbol{\xi})).
\end{aligned}$$

Portanto  $\mathbf{Y} \sim SN_k(\boldsymbol{\xi}, \Omega, \alpha)$ .

□

### Demonstração da Proposição 2.2.1.

A função de densidade conjunta do vetor  $\boldsymbol{\gamma}$  pode ser reescrita da seguinte forma

$$f(\boldsymbol{\gamma} | \sigma^2, \alpha) = 2 \left[ \prod_{i=1}^k \phi(\gamma_i | \xi_i, \sigma^2) \right] P \left( Z \leq \frac{\alpha}{\sigma} \sum_{i=1}^k \gamma_i - \xi_i \right), \tag{A.13}$$

onde  $Z \sim N(0, 1)$ . Segue de (A.13) que

$$\begin{aligned}
f(\boldsymbol{\gamma} | \sigma^2, \alpha) &= 2 \left[ \prod_{i=1}^k \phi(\gamma_i | \xi_i, \sigma^2) \right] P \left( Z\sigma \leq \alpha \left( \gamma_j - \xi_j + \sum_{\forall i \neq j} \gamma_i - \xi_i \right) \right) \\
&= 2 \left[ \prod_{i=1}^k \phi(\gamma_i | \xi_i, \sigma^2) \right] P \left( Z\sigma - \alpha(\gamma_j - \xi_j) \leq \alpha \sum_{\forall i \neq j} \gamma_i - \xi_i \right) \\
&= 2 \left[ \prod_{i=1}^k \phi(\gamma_i | \xi_i, \sigma^2) \right] \Phi \left( \alpha \sum_{\forall i \neq j} \gamma_i - \xi_i - \alpha(\gamma_j - \xi_j), \sigma^2 \right).
\end{aligned} \tag{A.14}$$

Faça  $\boldsymbol{\gamma}_{-\{j\}} = (\gamma_1, \dots, \gamma_{j-1}, \gamma_{j+1}, \dots, \gamma_k)$ ,  $\boldsymbol{\xi}_{-\{j\}} = (\xi_1, \dots, \xi_{j-1}, \xi_{j+1}, \dots, \xi_k)$  e considere a seguinte notação  $d\boldsymbol{\gamma}_{-\{j\}} = d\gamma_1 \dots d\gamma_{j-1} d\gamma_{j+1} \dots d\gamma_k$ . Integrando (A.14) segue que a função de densidade conjunta de  $\gamma_j$  e  $\boldsymbol{\gamma}_{-\{j\}}$  é

$$\begin{aligned}
f(\gamma_j | \boldsymbol{\xi}_{-\{j\}}, \sigma^2, \alpha) &= \int \dots \int f(\boldsymbol{\gamma} | \boldsymbol{\xi}_{-\{j\}}, \sigma^2, \alpha) d\boldsymbol{\gamma}_{-\{j\}} \\
&= 2\phi(\gamma_j | \boldsymbol{\xi}_{-\{j\}}, \sigma^2)
\end{aligned}$$

$$\begin{aligned}
& \int \dots \int \phi_{k-1}(\boldsymbol{\gamma}_{-\{j\}} | \boldsymbol{\xi}_{\{-j\}}, \sigma^2 I_{k-1}) \Phi \left( \alpha \sum_{\forall i \neq j} \gamma_i - \xi_i - \alpha(\gamma_j - \xi_j), \sigma^2 \right) d\boldsymbol{\gamma}_{-\{j\}} \\
&= 2\phi(\gamma_j | \xi_j, \sigma^2) E_{\boldsymbol{\gamma}_{-\{j\}}} \left[ \Phi \left( \alpha \sum_{\forall i \neq j} \gamma_i - \xi_i - \alpha(\gamma_j - \xi_j), \sigma^2 \right) \right], \tag{A.15}
\end{aligned}$$

onde  $E_{\boldsymbol{\gamma}_{-\{j\}}}[\cdot]$  denota a esperança com respeito a distribuição de  $\boldsymbol{\gamma}_{-\{j\}}$  a qual é  $N_{k-1}(\boldsymbol{\xi}_{\{-j\}}, \sigma^2 I_{k-1})$ .

Aplicando o Lema C.2.1 para calcularmos  $E_{\boldsymbol{\gamma}_{-\{j\}}} \left[ \Phi \left( \alpha \sum_{\forall i \neq j} \gamma_i - \xi_i - \alpha(\gamma_l + \gamma_j - \xi_j), \sigma^2 \right) \right]$  em (A.15), temos que, desde que  $\boldsymbol{\gamma}_{-\{j\}} \sim N_{k-1}(\boldsymbol{\xi}_{\{-j\}}, \sigma^2 I_{k-1})$  e assumindo  $A_{1 \times (k-1)} = \alpha \mathbf{1}'_{k-1}$ ,  $\mathbf{c} = \boldsymbol{\xi}_{\{-j\}}$ ,  $a = -\alpha \sum_{\forall i \neq j} \xi_i$ ,  $\mathbf{b} = -\alpha(\gamma_j - \xi_j)$ ,  $B = \sigma^2$  e  $C = \sigma^2 I_{k-1}$  tal esperança torna-se

$$\begin{aligned}
E_{\boldsymbol{\gamma}_{-\{j\}}} \left[ \Phi \left( \alpha \sum_{\forall i \neq j} \gamma_i - \xi_i - \alpha(\gamma_j - \xi_j), \sigma^2 \right) \right] &= \Phi(0 | -\alpha(\gamma_j - \xi_j), \sigma^2 (1 + \alpha^2(k-1))) \\
&= \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-1)}} (\gamma_j - \xi_j) | \sigma^2 \right), \tag{A.16}
\end{aligned}$$

uma vez que

$$\begin{aligned}
\mathbf{A}\mathbf{c} + \mathbf{a} &= \alpha \mathbf{1}'_{k-1} \boldsymbol{\xi}_{\{-j\}} + (-\alpha \sum_{\forall i \neq j} \xi_i) = 0, \text{ e} \\
\mathbf{B} + \mathbf{A}\mathbf{C}\mathbf{A}' &= \sigma^2 + \alpha \mathbf{1}'_{k-1} \sigma^2 I_{k-1} \alpha \mathbf{1}_{k-1} = \sigma^2 (1 + \alpha^2(k-1)).
\end{aligned}$$

Portanto, substituindo (A.16) na equação em (A.15) decorre que

$$\begin{aligned}
f(\gamma_j | \xi_j, \sigma^2, \alpha) &= 2\phi(\gamma_j | \xi_j, \sigma^2) \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-1)}} (\gamma_j - \xi_j) | \sigma^2 \right) \\
&= 2\phi \left( \frac{\gamma_j - \xi_j}{\sigma} \right) \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-1)}} \left( \frac{\gamma_j - \xi_j}{\sigma} \right) \right).
\end{aligned}$$

Logo,  $\gamma_j \sim SN(\xi_j, \sigma^2, \alpha^*)$  em que  $\alpha^* = \alpha \left( \sqrt{1 + \alpha^2(k-1)} \right)^{-1}$ .

Agora vamos estudar qual é o domínio do parâmetro de assimetria  $\alpha^*$  nesta situação. Note que, mantendo  $k$  fixo,  $\alpha^*$  é uma função de  $\alpha$  monótona crescente. Além disso sabemos que  $-\infty < \alpha < \infty$ . Então  $\alpha^*$  terá seu limite inferior quando  $\alpha$  estiver em seu limite inferior e o mesmo vale para o limite superior. Inicialmente vamos avaliar o limite inferior do domínio de  $\alpha^*$ . Segue que

$$\begin{aligned}
\lim_{\alpha \rightarrow -\infty} \alpha^* &= \lim_{\alpha \rightarrow -\infty} \frac{\alpha}{\sqrt{1 + \alpha^2(k-1)}} \\
&= \lim_{\alpha \rightarrow -\infty} \frac{\alpha}{\sqrt{1 + \alpha^2(k-1)}} \frac{\sqrt{1 + \alpha^2}}{\sqrt{1 + \alpha^2}} \\
&= \lim_{\alpha \rightarrow -\infty} \frac{\sqrt{1 + \alpha^2}}{\sqrt{1 + \alpha^2(k-1)}} \lim_{\alpha \rightarrow -\infty} \frac{\alpha}{\sqrt{1 + \alpha^2}}
\end{aligned}$$

$$\begin{aligned}
&= \lim_{\alpha \rightarrow -\infty} \sqrt{\frac{1 + \alpha^2}{1 + \alpha^2(k-1)}} \lim_{\alpha \rightarrow -\infty} \frac{\alpha}{\sqrt{1 + \alpha^2}} \\
&= \left( \lim_{\alpha \rightarrow -\infty} \frac{1 + \alpha^2}{1 + \alpha^2(k-1)} \right)^{1/2} \lim_{\alpha \rightarrow -\infty} \frac{\alpha}{\sqrt{1 + \alpha^2}}. \tag{A.17}
\end{aligned}$$

Sabemos que

$$\lim_{\alpha \rightarrow -\infty} \frac{\alpha}{\sqrt{1 + \alpha^2}} = -1. \tag{A.18}$$

e utilizando a regra de L'Hopital temos que

$$\lim_{\alpha \rightarrow -\infty} \frac{1 + \alpha^2}{1 + \alpha^2(k-1)} = \lim_{\alpha \rightarrow -\infty} \frac{2\alpha}{2(k-1)\alpha} = \frac{1}{k-1}. \tag{A.19}$$

Substituindo (A.18) e (A.19) em (A.17) temos

$$\lim_{\alpha \rightarrow -\infty} \alpha^* = -\sqrt{\frac{1}{k-1}}. \tag{A.20}$$

Para avaliar o limite superior de variação de  $\alpha^*$  fazemos o mesmo procedimento sabendo que  $\lim_{\alpha \rightarrow \infty} \frac{\alpha}{\sqrt{1 + \alpha^2}} = 1$  e obtemos que

$$\lim_{\alpha \rightarrow \infty} \alpha^* = \sqrt{\frac{1}{k-1}} \tag{A.21}$$

Por fim, de (A.20) e (A.21) obtemos que  $-\sqrt{\frac{1}{k-1}} < \alpha^* < \sqrt{\frac{1}{k-1}}$ , concluindo a prova.  $\square$

# Apêndice B

## Distribuições Condicionais Completas para Modelo Logístico Misto com distribuições assimétricas para os Efeitos Aleatórios

Neste apêndice são apresentadas as distribuições completas para os casos em que assumimos as distribuições normal assimétrica univariada e normal assimétrica multivariada para os efeitos aleatórios  $\gamma$  e não utilizamos as representações estocásticas para especificar tais distribuições. As distribuições *a priori* consideradas são as mesmas já citadas nos Capítulos 1 e 2, ou seja, consideramos  $\sigma^2 \sim GI(a, d)$  e  $\lambda \sim N(h, \tau^2)$ .

### B.1 Distribuição Normal Assimétrica Univariada para os Efeitos Aleatórios

As distribuições condicionais completas para o vetor de efeitos aleatórios  $\gamma$ , para os parâmetros de assimetria  $\lambda$  e escala  $\sigma^2$ , na situação em que  $\gamma_i \sim SN(-\delta\sigma\sqrt{2/\pi}, \sigma^2, \lambda)$  e não é utilizada a representação estocástica para especificar a distribuição de  $\gamma$ , são dadas, respectivamente, por:

$$\begin{aligned} f(\gamma|\mathbf{y}, \beta, \sigma^2, \lambda) &\propto f(\mathbf{y}|\beta, \gamma, \sigma^2, \lambda)f(\gamma|\sigma^2, \lambda) \\ &\propto \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp\{\eta_{ij}\}}{1 + \exp\{\eta_{ij}\}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp\{\eta_{ij}\}} \right]^{1-y_{ij}} \right] \\ &\quad \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i + \delta\sigma\sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} \right\} \Phi_k \left( \frac{\lambda}{\sigma} \left( \gamma + \mathbf{1}_k \delta\sigma\sqrt{\frac{2}{\pi}} \right) \right), \end{aligned}$$

$$f(\lambda|\mathbf{y}, \beta, \gamma, \sigma^2) \propto f(\gamma|\sigma^2, \lambda)f(\lambda)$$

$$\propto \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i + \delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{(\lambda - h)^2}{2\tau^2} \right\} \Phi_k \left( \frac{\lambda}{\sigma} \left( \gamma + \mathbf{1}_k \delta \sigma \sqrt{\frac{2}{\pi}} \right) \right),$$

$$\begin{aligned} f(\sigma^2 | \mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \lambda) &\propto f(\boldsymbol{\gamma} | \sigma^2, \lambda) f(\sigma^2) \\ &\propto \left( \frac{1}{\sigma^2} \right)^{\frac{k}{2} + a + 1} \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i + \delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{d}{\sigma^2} \right\} \Phi_k \left( \frac{\lambda}{\sigma} \left( \gamma + \mathbf{1}_k \delta \sigma \sqrt{\frac{2}{\pi}} \right) \right), \end{aligned}$$

no qual  $\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}$ .

## B.2 Distribuição Normal Assimétrica Multivariada para os Efeitos Aleatórios

As distribuições condicionais completas para o vetor de efeitos aleatórios  $\boldsymbol{\gamma}$ , para os parâmetros de assimetria  $\alpha$  e escala  $\sigma^2$ , na situação em que  $\boldsymbol{\gamma} \sim SN_k(-\boldsymbol{\Delta} \sigma \sqrt{2/\pi}, \sigma^2 I_k, \alpha \mathbf{1}_k)$  e não utilizamos a representação estocástica para especificar a distribuição de  $\boldsymbol{\gamma}$ , são dadas, respectivamente, por:

$$\begin{aligned} f(\boldsymbol{\gamma} | \mathbf{y}, \boldsymbol{\beta}, \sigma^2, \alpha) &\propto f(\mathbf{y} | \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2, \alpha) f(\boldsymbol{\gamma} | \sigma^2, \alpha) \\ &\propto \left[ \prod_{i=1}^k \prod_{j=1}^{n_i} \left[ \frac{\exp \{ \eta_{ij} \}}{1 + \exp \{ \eta_{ij} \}} \right]^{y_{ij}} \left[ \frac{1}{1 + \exp \{ \eta_{ij} \}} \right]^{1 - y_{ij}} \right] \\ &\quad \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i + \Delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} \right\} \Phi \left( \frac{\alpha}{\sigma} \sum_{i=1}^k \left( \gamma_i + \Delta \sigma \sqrt{\frac{2}{\pi}} \right) \right), \end{aligned}$$

$$\begin{aligned} f(\alpha | \mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma^2) &\propto f(\boldsymbol{\gamma} | \sigma^2, \alpha) f(\alpha) \\ &\propto \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i + \Delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{(\alpha - h)^2}{2\tau^2} \right\} \Phi \left( \frac{\alpha}{\sigma} \sum_{i=1}^k \left( \gamma_i + \Delta \sigma \sqrt{\frac{2}{\pi}} \right) \right), \end{aligned}$$

$$\begin{aligned} f(\sigma^2 | \mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha) &\propto f(\boldsymbol{\gamma} | \sigma^2, \alpha) f(\sigma^2) \\ &\propto \left( \frac{1}{\sigma^2} \right)^{\frac{k}{2} + a + 1} \exp \left\{ -\frac{\sum_{i=1}^k \left( \gamma_i + \Delta \sigma \sqrt{\frac{2}{\pi}} \right)^2}{2\sigma^2} - \frac{d}{\sigma^2} \right\} \Phi \left( \frac{\alpha}{\sigma} \sum_{i=1}^k \left( \gamma_i + \Delta \sigma \sqrt{\frac{2}{\pi}} \right) \right), \end{aligned}$$

no qual  $\Delta = \frac{\alpha}{\sqrt{1 + k\alpha^2}}$ .

# Apêndice C

## Provas dos Teoremas e Proposições do Capítulo 3

Neste apêndice vamos apresentar as demonstrações relacionadas aos teoremas e proposições utilizados no decorrer do Capítulo 3. Estas demonstrações são apresentadas nas Seções C.1, C.2 e C.3 que contém as demonstrações dos teoremas e proposições das Seções 3.1, 3.2 e 3.3, respectivamente.

### C.1 Provas dos Teoremas da Seção 3.1

Neste apêndice são apresentadas as demonstrações dos teoremas da Seção 3.1. Para tal, os seguintes resultados se fazem necessários.

**Teorema C.1.1.** *O vetor  $Y$  tem distribuição normal multivariada se, e somente se,  $\mathbf{a}'Y$  tem distribuição normal univariada para todo vetor real não nulo  $\mathbf{a}$ .*

A demonstração deste teorema pode ser encontrada em Seber e Lee (2003), página 30.

**Teorema C.1.2.** *Se o vetor aleatório  $\mathbf{X} = (X_1, \dots, X_n)'$ , de dimensão  $n \times 1$ , tem distribuição normal multivariada com vetor de médias  $\boldsymbol{\mu}$  e matriz de covariância  $\boldsymbol{\Sigma}$ , então a variável aleatória  $Y = \mathbf{a}'\mathbf{X} + c$ , em que  $\mathbf{a}$  é um vetor real de ordem  $n \times 1$  e  $c \in \mathbb{R}$  é uma constante, tem distribuição Normal com média  $\mathbf{a}'\boldsymbol{\mu} + c$  e variância  $\mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}$ .*

*Demonstração.*

Inicialmente, consideremos a variável aleatória  $S = \mathbf{a}'\mathbf{X}$ . Pelo Teorema C.1.1, sabemos que  $S$  tem distribuição normal e de resultados usuais de cálculos de probabilidade segue que a com média e a variância são dadas, respectivamente, por:

$$\begin{aligned} E[S] &= E[\mathbf{a}'\mathbf{X}] = \mathbf{a}'E[\mathbf{X}] = \mathbf{a}'\boldsymbol{\mu}, \\ V[S] &= V[\mathbf{a}'\mathbf{X}] = \mathbf{a}'V[\mathbf{X}]\mathbf{a} = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}. \end{aligned}$$

O resultado segue por notarmos que  $Y = S + c$ . □



**Prova do Teorema 3.1.1.**

Primeiramente vamos calcular a função de distribuição acumulada de  $OR$  e a partir dela encontrarmos a função densidade. Seja  $t \in \mathbb{R}^+$  qualquer. Daí,

$$\begin{aligned} F_{OR|k,\omega^2}(t) &= P(\exp\{k + \omega V\} \leq t) \\ &= P(k + \omega V \leq \ln t) \\ &= P\left(V \leq \frac{\ln(t) - k}{\omega}\right) \\ &= F_V\left(\frac{\ln(t) - k}{\omega}\right). \end{aligned}$$

Derivando com respeito a  $t$  tem-se que a função densidade da variável  $OR$  é:

$$\begin{aligned} f_{OR|k,\omega^2}(t) &= \frac{d}{dt} F_{OR|k,\omega^2}(t) \\ &= f_V\left(\frac{\ln(t) - k}{\omega}\right) \frac{1}{t\omega} \\ &= \frac{1}{\sqrt{2\pi t\omega}} \exp\left\{-\frac{1}{2}\left(\frac{\ln t - k}{\omega}\right)^2\right\} \\ &= [\sqrt{2\pi t\omega}]^{-1} \exp\left\{-\frac{1}{2\omega^2}(\ln t - k)^2\right\}. \end{aligned}$$

Concluindo a prova. □

**Prova do Teorema 3.1.2.**

- (i) Denote por  $\Phi(\cdot)$  a função distribuição acumulada da distribuição normal padrão. Seja  $a_1 = \text{med}\{\exp\{|\omega V|\}\}$ . Assim, usando a definição de mediana e equivalência de eventos, segue que

$$\begin{aligned} P(\exp\{|\omega V|\} \leq a_1) &= 0,5 \\ P(|\omega V| \leq \ln a_1) &= 0,5 \\ P(-\ln a_1 \leq \omega V \leq \ln a_1) &= 0,5 \\ P\left(V \leq \frac{\ln a_1}{\omega}\right) - P\left(V \leq \frac{-\ln a_1}{\omega}\right) &= 0,5 \\ 2P\left(V \leq \frac{\ln a_1}{\omega}\right) - 1 &= 0,5 \\ P\left(V \leq \frac{\ln a_1}{\omega}\right) &= 0,75 \\ \Phi\left(\frac{\ln a_1}{\omega}\right) &= 0,75. \end{aligned} \tag{C.1}$$

Consequentemente, tem-se que

$$\begin{aligned} \frac{\ln a_1}{\omega} &= \Phi^{-1}(0,75) \\ a_1 &= \exp\{\omega \Phi^{-1}(0,75)\}. \end{aligned}$$

Agora, faça  $b = \text{med}\{|\omega V|\}$ . Similarmente temos que

$$\begin{aligned} P(|\omega V| \leq b) &= 0,5 \\ P\left(\frac{-b}{\omega} \leq V \leq \frac{b}{\omega}\right) &= 0,5 \\ P\left(V \leq \frac{b}{\omega}\right) - P\left(V \leq \frac{-b}{\omega}\right) &= 0,5 \\ 2P\left(V \leq \frac{b}{\omega}\right) - 1 &= 0,5 \\ P\left(V \leq \frac{b}{\omega}\right) &= 0,75 \\ \Phi\left(\frac{b}{\omega}\right) &= 0,75. \end{aligned}$$

Segue como consequência que

$$\begin{aligned} \frac{b}{\omega} &= \Phi^{-1}(0,75) \\ b &= \omega\Phi^{-1}(0,75). \end{aligned}$$

O que implica que  $\exp\{b\} = \exp\{\omega\Phi^{-1}(0,75)\}$  e conclui a prova do item (i).

(ii) Seja  $\text{med}\{\exp\{|\omega V|\}\} = h$ . Como em (i), segue que

$$\begin{aligned} P(\exp\{-|\omega V|\} \leq h) &= 0,5 \\ P(-|\omega V| \leq \ln h) &= 0,5 \\ P(|\omega V| \geq -\ln h) &= 0,5 \\ P\left(|\omega V| \geq \ln\left(\frac{1}{h}\right)\right) &= 0,5 \\ P\left(\omega V \geq \ln\left(\frac{1}{h}\right)\right) + P\left(\omega V \leq -\ln\left(\frac{1}{h}\right)\right) &= 0,5 \\ P(\omega V \geq \ln 1 - \ln h) + P(\omega V \leq -\ln 1 + \ln h) &= 0,5 \\ P(\omega V \geq -\ln h) + P(\omega V \leq \ln h) &= 0,5 \\ P\left(V < \frac{\ln h}{\omega}\right) + P\left(V \leq \frac{\ln h}{\omega}\right) &= 0,5 \\ 2P\left(V \leq \frac{\ln h}{\omega}\right) &= 0,5 \\ \Phi\left(\frac{\ln h}{\omega}\right) &= 0,25. \end{aligned}$$

Consequentemente, tem-se que

$$\begin{aligned} \frac{\ln h}{\omega} &= \Phi^{-1}(0,25) \\ \frac{\ln h}{\omega} &= -\Phi^{-1}(0,75) \\ h &= \exp\{-\omega\Phi^{-1}(0,75)\} \\ h &= [\text{med}\{\exp\{|\omega V|\}\}]^{-1}. \end{aligned}$$

O que conclui a prova do item (ii).

(iii) Defina  $a = \text{med}\{\exp\{k + \omega V\}\}$ . Logo temos que

$$\begin{aligned} F_{OR|k,\omega^2}(a) &= 0,5 \\ P(\exp\{k + \omega V\} \leq a) &= 0,5 \\ P\left(V \leq \frac{\ln(a) - k}{\omega}\right) &= 0,5 \\ \Phi\left(\frac{\ln(a) - k}{\omega}\right) &= 0,5. \end{aligned} \tag{C.2}$$

De (C.2) segue então que

$$\begin{aligned} \frac{\ln(a) - k}{\omega} &= \Phi^{-1}(0,5) \\ \frac{\ln(a) - k}{\omega} &= 0 \\ a &= \exp\{k\}. \end{aligned}$$

Portanto,  $\text{med}\{\exp\{k + \omega V\}\} = \exp\{k\}$ , concluindo, assim, a prova do Teorema 3.1.2.

□

## C.2 Provas das Proposições da Seção 3.2

Neste apêndice são apresentadas as demonstrações das proposições da Seção 3.2.

Para provarmos os resultados a seguir e também outros do Apêndice C.3 o seguinte lema será utilizado.

**Lema C.2.1.** *Se  $\mathbf{U} \sim N_k(\mathbf{c}, C)$  é um vetor não-singular, então para algum vetor fixo  $\mathbf{a}$   $m$ -dimensional e alguma matriz  $A$  com dimensão  $m \times k$ , temos que:*

$$E[\Phi_m(A\mathbf{U} + \mathbf{a}|\mathbf{b}, B)] = \Phi_m(A\mathbf{c} + \mathbf{a}|\mathbf{b}, B + ACA').$$

Mais sobre este lema pode ser visto em Arellano-Valle e Genton (2005) e Gupta *et al.* (2004).

### Demonstração da Proposição 3.2.1.

Como  $\gamma_l$  e  $\gamma_j$  são independentes, a função de densidade de probabilidade conjunta do vetor  $(\gamma_l, \gamma_j)'$  é dada por:

$$\begin{aligned} f_{\gamma_l, \gamma_j}(\gamma_l, \gamma_j) &= f(\gamma_l)f(\gamma_j) \\ &= \frac{2}{\sigma} \phi\left(\frac{\gamma_l}{\sigma}\right) \Phi\left(\frac{\gamma_l}{\sigma}\lambda\right) \frac{2}{\sigma} \phi\left(\frac{\gamma_j}{\sigma}\right) \Phi\left(\frac{\gamma_j}{\sigma}\lambda\right) \\ &= \frac{4}{\sigma^2} \phi\left(\frac{\gamma_l}{\sigma}\right) \phi\left(\frac{\gamma_j}{\sigma}\right) \Phi_2\left(\lambda \begin{bmatrix} \gamma_l \\ \gamma_j \end{bmatrix} \middle| \sigma^2 I_2\right). \end{aligned}$$

Para encontrarmos a distribuição de  $\gamma_l - \gamma_j$  faremos uma transformação de variáveis. Defina  $W = \gamma_l - \gamma_j$  e  $Y = \gamma_l$ . Então, utilizando o método do Jacobiano segue que:

$$\begin{aligned}
f_{W,Y}(w, y) &= f_{\gamma_l, \gamma_j}(y, y-w) \left| \begin{array}{cc} \frac{d\gamma_l}{dy} & \frac{d\gamma_l}{dw} \\ \frac{d\gamma_j}{dy} & \frac{d\gamma_j}{dw} \end{array} \right| \\
&= \frac{4}{\sigma^2} \phi\left(\frac{y}{\sigma}\right) \phi\left(\frac{y-w}{\sigma}\right) \Phi_2\left(\lambda \begin{bmatrix} y \\ y-w \end{bmatrix} \mid \sigma^2 I_2\right) \\
&= \frac{4}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}(2y^2 - 2yw + w^2)\right\} \Phi_2\left(\lambda y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} 0 \\ -w \end{bmatrix} \mid \sigma^2 I_2\right) \\
&= \frac{4}{2\pi\sigma^2} \exp\left\{-\frac{2}{2\sigma^2}\left(y - \frac{w}{2}\right)^2\right\} \exp\left\{-\frac{w^2}{4\sigma^2}\right\} \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right) \\
&= 4 \frac{\sqrt{2}}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{2}{2\sigma^2}\left(y - \frac{w}{2}\right)^2\right\} \frac{1}{\sqrt{4\pi\sigma^2}} \exp\left\{-\frac{w^2}{4\sigma^2}\right\} \times \\
&\quad \times \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right) \\
&= 4\phi\left(y \mid \frac{w}{2}, \frac{\sigma^2}{2}\right) \phi(w \mid 2\sigma^2) \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right).
\end{aligned}$$

A distribuição marginal de  $W = \gamma_l - \gamma_j$  é obtida como segue

$$\begin{aligned}
f_w(w) &= \int f_{W,Y}(w, y) dy \\
&= 4\phi(w \mid 2\sigma^2) \int \phi\left(y \mid \frac{w}{2}, \frac{\sigma^2}{2}\right) \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right) dy \\
&= 4\phi(w \mid 2\sigma^2) E_y \left[ \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right) \right], \tag{C.3}
\end{aligned}$$

onde  $E_y[\cdot]$  denota a esperança com respeito a distribuição  $N(y \mid \frac{w}{2}, \frac{\sigma^2}{2})$ .

Para calcularmos  $E_y \left[ \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right) \right]$  na expressão (C.3), utilizaremos o Lema C.2.1, onde assumimos  $y \sim N(\frac{w}{2}, \frac{\sigma^2}{2})$ ,  $A_{2 \times 1} = \frac{\lambda}{\sigma} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ,  $\mathbf{a} = \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}$ ,  $\mathbf{b} = \mathbf{0}_2$ ,  $B = I_2$ ,  $\mathbf{c} = \frac{w}{2}$  e  $C = \frac{\sigma^2}{2}$ . Consequentemente, temos que

$$\begin{aligned}
\mathbf{A}\mathbf{c} + \mathbf{a} &= \frac{\lambda}{\sigma} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{w}{2} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix} = \frac{\lambda w}{2\sigma} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \\
B + \mathbf{A}\mathbf{C}\mathbf{A}' &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \frac{\sigma^2}{2} \frac{\lambda}{\sigma} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 + \frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1 + \frac{\lambda^2}{2} \end{bmatrix}.
\end{aligned}$$

Daí, segue do Lema C.2.1 que

$$E_y \left[ \Phi_2\left(\frac{\lambda}{\sigma} y \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{\lambda}{\sigma} \begin{bmatrix} 0 \\ -w \end{bmatrix}\right) \right] = \Phi_2\left(\frac{\lambda w}{2\sigma} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \mid \begin{bmatrix} 1 + \frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1 + \frac{\lambda^2}{2} \end{bmatrix}\right). \tag{C.4}$$

Substituindo o termo em (C.4) na equação em (C.3), segue que a distribuição de  $W$  é

$$\begin{aligned} f_W(w) &= 4\phi(w|2\sigma^2)\Phi_2\left(\frac{\lambda w}{2\sigma}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|\begin{bmatrix} 1+\frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1+\frac{\lambda^2}{2} \end{bmatrix}\right) \\ &= 4\phi(w|2\sigma^2)\Phi_2\left(\lambda w\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|4\sigma^2\begin{bmatrix} 1+\frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1+\frac{\lambda^2}{2} \end{bmatrix}\right) \\ &= 4\phi(w|2\sigma^2)\Phi_2\left(\lambda w\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|4\sigma^2\left[I_2+\frac{\lambda^2}{2}\mathbf{1}_2\mathbf{1}'_2\right]\right), \end{aligned}$$

onde  $\mathbf{1}_2$  denota um vetor de uns de ordem  $2 \times 1$ .

□

### Demonstração da Proposição 3.2.2.

Uma vez que  $OR = \exp\{k + W\}$ , segue que  $W = \ln(OR) - k$ . Usando resultados do cálculo de probabilidades e o resultado da Proposição 3.2.1 segue que

$$\begin{aligned} f_{OR|\beta,\sigma^2,\lambda}(t) &= f_W(\ln t - k)\left|\frac{dw}{dt}\right| \\ &= 4\phi(\ln t - k|2\sigma^2)\Phi_2\left(\lambda(\ln t - k)\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|4\sigma^2\begin{bmatrix} 1+\frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1+\frac{\lambda^2}{2} \end{bmatrix}\right)\frac{1}{t} \\ &= \frac{4}{t}\phi(\ln t|k, 2\sigma^2)\Phi_2\left(\lambda \ln t\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|\lambda k\begin{bmatrix} 1 \\ -1 \end{bmatrix}, 4\sigma^2\begin{bmatrix} 1+\frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1+\frac{\lambda^2}{2} \end{bmatrix}\right), \end{aligned}$$

concluindo a prova.

□

### Demonstração da Proposição 3.2.3.

Seja  $a$  a mediana da distribuição de  $OR = \exp\{k + W\}$ . Assim, por definição de mediana e da Proposição 3.2.2 segue que

$$\begin{aligned} F_{OR|\beta,\sigma^2,\lambda}(a) &= \int_0^a f_{OR|\beta,\sigma^2,\lambda}(t)dt = 0,5 \\ \Leftrightarrow \int_0^a \frac{4}{t}\phi(\ln t - k|2\sigma^2)\Phi_2\left(\lambda(\ln t - k)\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|4\sigma^2\begin{bmatrix} 1+\frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1+\frac{\lambda^2}{2} \end{bmatrix}\right)dt &= 0,5. \end{aligned}$$

Usando a transformação  $z = \ln t - k$ , temos que  $dz = \frac{1}{t}dt$ , Assim, temos como consequência que

$$\begin{aligned} F_{OR|\beta,\sigma^2,\lambda}(a) &= \int_{-\infty}^{\ln a - k} 4\phi(z|2\sigma^2)\Phi_2\left(\lambda z\begin{bmatrix} 1 \\ -1 \end{bmatrix}\middle|4\sigma^2\begin{bmatrix} 1+\frac{\lambda^2}{2} & \frac{\lambda^2}{2} \\ \frac{\lambda^2}{2} & 1+\frac{\lambda^2}{2} \end{bmatrix}\right)dz = 0,5; \\ &= \Phi_{SUNM}\left(\ln a - k|2\sigma^2, \lambda[1, -1]', 4\sigma^2\left[I_2+\frac{\lambda^2}{2}\mathbf{1}_2\mathbf{1}'_2\right]\right) = 0,5. \end{aligned}$$

Disto, segue que:

$$\begin{aligned} \ln a - k &= \Phi_{SUNM}^{-1} \left( 0, 5|2\sigma^2, \lambda[1, -1]', 4\sigma^2 \left[ I_2 + \frac{\lambda^2}{2} \mathbf{1}_2 \mathbf{1}_2' \right] \right); \\ a &= \exp \left\{ \Phi_{SUNM}^{-1} \left( 0, 5|2\sigma^2, \lambda[1, -1]', 4\sigma^2 \left[ I_2 + \frac{\lambda^2}{2} \mathbf{1}_2 \mathbf{1}_2' \right] \right) + k \right\}, \end{aligned}$$

o que conclui a prova. □

### C.3 Provas das Proposições da Seção 3.3

Neste apêndice são apresentadas as demonstrações das proposições utilizadas na Seção 3.3.

#### Demonstração da Proposição 3.3.1.

Note de (2.10) que a função de densidade conjunta do vetor  $\gamma$  pode ser reescrita da seguinte forma

$$f(\gamma|\sigma^2, \alpha) = 2 \left[ \prod_{i=1}^k \phi(\gamma_i|\sigma^2) \right] P \left( Z \leq \frac{\alpha}{\sigma} \sum_{i=1}^k \gamma_i \right), \quad (\text{C.5})$$

onde  $Z \sim N(0, 1)$ . Segue de (C.5) que

$$\begin{aligned} f(\gamma|\sigma^2, \alpha) &= 2 \left[ \prod_{i=1}^k \phi(\gamma_i|\sigma^2) \right] P \left( Z\sigma \leq \alpha \left( \gamma_l + \gamma_j + \sum_{\forall i \neq l, j} \gamma_i \right) \right) \\ &= 2 \left[ \prod_{i=1}^k \phi(\gamma_i|\sigma^2) \right] P \left( Z\sigma - \alpha(\gamma_l + \gamma_j) \leq \alpha \sum_{\forall i \neq l, j} \gamma_i \right) \\ &= 2 \left[ \prod_{i=1}^k \phi(\gamma_i|\sigma^2) \right] \Phi \left( \alpha \sum_{\forall i \neq l, j} \gamma_i - \alpha(\gamma_l + \gamma_j), \sigma^2 \right). \end{aligned} \quad (\text{C.6})$$

Faça  $\gamma_{-\{l, j\}} = (\gamma_1, \dots, \gamma_{l-1}, \gamma_{l+1}, \dots, \gamma_{j-1}, \gamma_{j+1}, \dots, \gamma_k)$  e considere a seguinte notação  $d\gamma_{-\{l, j\}} = d\gamma_1 \dots d\gamma_{l-1} d\gamma_{l+1} \dots d\gamma_{j-1} d\gamma_{j+1} \dots d\gamma_k$ . Integrando (C.6) segue que a função de densidade conjunta de  $\gamma_l$  e  $\gamma_j$  é

$$\begin{aligned} f(\gamma_l, \gamma_j|\sigma^2, \alpha) &= \int \dots \int f(\gamma|\sigma^2, \alpha) d\gamma_{-\{l, j\}} \\ &= 2\phi(\gamma_l|\sigma^2)\phi(\gamma_j|\sigma^2) \times \\ &\quad \times \int \dots \int \phi_{k-2}(\gamma_{-\{l, j\}}|\sigma^2 I_{k-2}) \Phi \left( \alpha \sum_{\forall i \neq l, j} \gamma_i - \alpha(\gamma_l + \gamma_j), \sigma^2 \right) d\gamma_{-\{l, j\}} \\ &= 2\phi(\gamma_l|\sigma^2)\phi(\gamma_j|\sigma^2) E_{\gamma_{-\{l, j\}}} \left[ \Phi \left( \alpha \sum_{\forall i \neq l, j} \gamma_i - \alpha(\gamma_l + \gamma_j), \sigma^2 \right) \right], \end{aligned} \quad (\text{C.7})$$

onde  $E_{\gamma_{-\{l, j\}}}[\cdot]$  denota a esperança com respeito a distribuição de  $\gamma_{-\{l, j\}}$  a qual é  $N_{k-2}(\mathbf{0}_{k-2}, \sigma^2 I_{k-2})$ .

Aplicando o Lema C.2.1 para calcularmos  $E\gamma_{-\{l,j\}} \left[ \Phi \left( \alpha \sum_{\forall i \neq l,j} \gamma_i | - \alpha(\gamma_l + \gamma_j), \sigma^2 \right) \right]$  em (C.7), temos que, desde que  $\gamma_{-\{l,j\}} \sim N_{k-2}(\mathbf{0}_{k-2}, \sigma^2 I_{k-2})$  e assumindo  $A_{1 \times (k-2)} = \alpha \mathbf{1}'_{k-2}$ ,  $\mathbf{c} = \mathbf{0}_{k-2}$ ,  $\mathbf{a} = 0$ ,  $\mathbf{b} = -\alpha(\gamma_l + \gamma_j)$ ,  $B = \sigma^2$  e  $C = \sigma^2 I_{k-2}$  tal esperança torna-se

$$\begin{aligned} E\gamma_{-\{l,j\}} \left[ \Phi \left( \alpha \sum_{\forall i \neq l,j} \gamma_i | - \alpha(\gamma_l + \gamma_j), \sigma^2 \right) \right] &= \Phi(0 | -\alpha(\gamma_l + \gamma_j), \sigma^2 (1 + \alpha^2(k-2))) \\ &= \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-2)}} (\gamma_l + \gamma_j) | \sigma^2 \right), \end{aligned} \quad (\text{C.8})$$

uma vez que

$$\begin{aligned} \mathbf{A}\mathbf{c} + \mathbf{a} &= \alpha \mathbf{1}'_{k-2} \mathbf{0}_{k-2} + 0 = 0, \text{ e} \\ B + \mathbf{A}\mathbf{C}\mathbf{A}' &= \sigma^2 + \alpha \mathbf{1}'_{k-2} \sigma^2 I_{k-2} \alpha \mathbf{1}_{k-2} = \sigma^2 (1 + \alpha^2(k-2)). \end{aligned}$$

Portanto, substituindo (C.8) na equação em (C.7) decorre que

$$\begin{aligned} f(\gamma_l, \gamma_j | \sigma^2, \alpha) &= 2\phi(\gamma_l | \sigma^2) \phi(\gamma_j | \sigma^2) \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-2)}} (\gamma_l + \gamma_j) | \sigma^2 \right) \\ &= 2\phi_2 \left( \begin{bmatrix} \gamma_l \\ \gamma_j \end{bmatrix} | \sigma^2 I_2 \right) \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-2)}} (\gamma_l + \gamma_j) | \sigma^2 \right) \\ &= 2\phi_2 \left( \begin{bmatrix} \gamma_l \\ \gamma_j \end{bmatrix} | \sigma^2 I_2 \right) \Phi \left( \frac{\alpha}{\sqrt{1 + \alpha^2(k-2)}} \mathbf{1}'_2 \frac{1}{\sigma} I_2 \gamma_{lj} \right). \end{aligned}$$

Logo,  $\gamma_{lj} \sim SN_2 \left( \mathbf{0}_2, \sigma^2 I_2, \frac{\alpha}{\sqrt{1 + \alpha^2(k-2)}} \mathbf{1}_2 \right)$ , fato que conclui a prova. □

### Demonstração para a Proposição 3.3.2.

Utilizando a Proposição 3.3.1 temos que  $\gamma_{lj} \sim SN_2(\mathbf{0}_2, \sigma^2 I_2, \alpha^* \mathbf{1}_2)$ , em que  $\alpha^* = \alpha[1 + \alpha^2(k-2)]^{-1}$ . Para encontrarmos a distribuição de  $\gamma_l - \gamma_j$  utilizaremos o método do Jacobiano. Defina  $W = \gamma_l - \gamma_j$  e  $Y = \gamma_l$ . Então, a distribuição conjunta de  $W$  e  $Y$  é

$$\begin{aligned} f_{W,Y}(w, y) &= f_{\gamma_l, \gamma_j}(y, y - w) \left| \begin{array}{cc} \frac{d\gamma_l}{dy} & \frac{d\gamma_l}{dw} \\ \frac{d\gamma_j}{dy} & \frac{d\gamma_j}{dw} \end{array} \right| \\ &= 2\phi_2 \left( \begin{bmatrix} y \\ y - w \end{bmatrix} | \sigma^2 I_2 \right) \Phi(\alpha^* (2y - w) | \sigma^2) \\ &= 2\phi(y | \sigma^2) \phi(y - w | \sigma^2) \Phi(\alpha^* (2y - w) | \sigma^2). \end{aligned} \quad (\text{C.9})$$

Considerando apenas o termo  $\phi(y | \sigma^2) \phi(y - w | \sigma^2)$ , da equação (C.9) temos que

$$\begin{aligned} \phi(y | \sigma^2) \phi(y - w | \sigma^2) &= \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2\sigma^2} (y^2 + (y - w)^2) \right\} \\ &= \frac{1}{2\pi\sigma^2} \exp \left\{ -\frac{1}{2\sigma^2} (2y^2 - 2yw + w^2) \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{2}{2\sigma^2}\left(y-\frac{w}{2}\right)^2\right\} \exp\left\{-\frac{1}{2\sigma^2}w^2 + \frac{1}{4\sigma^2}w^2\right\} \\
&= \frac{1}{\sqrt{2\pi}} \frac{\sqrt{2}}{\sigma} \exp\left\{-\frac{2}{2\sigma^2}\left(y-\frac{w}{2}\right)^2\right\} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2}\sigma} \exp\left\{-\frac{1}{4\sigma^2}w^2\right\} \\
&= \phi\left(y\left|\frac{w}{2}, \frac{\sigma^2}{2}\right.\right) \phi(w|2\sigma^2). \tag{C.10}
\end{aligned}$$

Desta forma, substituindo o termo em (C.10) na equação em (C.9), temos:

$$f_{W,Y}(w, y) = 2\phi\left(y\left|\frac{w}{2}, \frac{\sigma^2}{2}\right.\right) \phi(w|2\sigma^2) \Phi(\alpha^*(2y-w)|\sigma^2).$$

Para encontrarmos a distribuição de marginal de  $W = \gamma_l - \gamma_j$  fazemos:

$$\begin{aligned}
f_w(w) &= \int f_{W,Y}(w, y) dy \\
&= 2\phi(w|2\sigma^2) \int \phi\left(y\left|\frac{w}{2}, \frac{\sigma^2}{2}\right.\right) \Phi(\alpha^*(2y-w)|\sigma^2) dy \\
&= 2\phi(w|2\sigma^2) E_y[\Phi(\alpha^*(2y-w)|\sigma^2)], \tag{C.11}
\end{aligned}$$

onde  $E_y[\cdot]$  denota a esperança com respeito a distribuição  $N(y|\frac{w}{2}, \frac{\sigma^2}{2})$ .

Utilizando o Lema C.2.1 para calcular  $E_y[\Phi(\alpha^*(2y-w)|\sigma^2)]$ , onde agora temos  $y \sim N(\frac{w}{2}, \frac{\sigma^2}{2})$ ,  $A_{1 \times 1} = 2\alpha^*$ ,  $\mathbf{a} = -w\alpha^*$ ,  $\mathbf{b} = 0$ ,  $B = \sigma^2$ ,  $\mathbf{c} = \frac{w}{2}$  e  $C = \frac{\sigma^2}{2}$ , segue que

$$E_y[\Phi(\alpha^*(2y-w)|\sigma^2)] = \Phi\left(0|\sigma^2\left(1 + 2\frac{\alpha^2}{1 + \alpha^2(k-2)}\right)\right) \tag{C.12}$$

uma vez que

$$\begin{aligned}
Ac + a &= 2\alpha^* \frac{w}{2} - w\alpha^* = 0, \text{ e} \\
B + ACA' &= \sigma^2 + 2\alpha^* \frac{\sigma^2}{2} 2\alpha^* = \sigma^2 (1 + 2(\alpha^*)^2).
\end{aligned}$$

Substituindo o termo em (C.12) na equação em (C.11), temos, como consequência, que

$$\begin{aligned}
f_w(w) &= 2\phi(w|2\sigma^2) \Phi(0|\sigma^2(1 + 2(\alpha^*)^2)) \\
&= \phi(w|2\sigma^2).
\end{aligned}$$

Portanto, concluímos que  $W = \gamma_l - \gamma_j \sim N(0, 2\sigma^2)$ , finalizando a prova. □

### Demonstração da Proposição 3.3.3.

A função geradora de momentos (f.g.m.) da distribuição  $SN_k(\boldsymbol{\xi}, \boldsymbol{\Omega}, \boldsymbol{\alpha})$ , com densidade dada em (2.10), aplicada no ponto  $\mathbf{t} = (t_1, \dots, t_k)'$  é dada por

$$M(\mathbf{t}) = 2 \exp\left\{\boldsymbol{\xi}'\mathbf{t} + \frac{1}{2}\mathbf{t}'\boldsymbol{\Omega}\mathbf{t}\right\} \Phi(\boldsymbol{\Delta}'\boldsymbol{\omega}\mathbf{t}), \tag{C.13}$$



em que

$$\mathbf{\Delta} = (1 + \boldsymbol{\alpha}'\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha})^{-\frac{1}{2}}\bar{\boldsymbol{\Omega}}\boldsymbol{\alpha}, \quad (\text{C.14})$$

$$\bar{\boldsymbol{\Omega}} = \boldsymbol{\omega}^{-1}\boldsymbol{\Omega}\boldsymbol{\omega}^{-1}. \quad (\text{C.15})$$

Supondo que  $\boldsymbol{\Omega} = \text{diag}(\sigma_1^2, \dots, \sigma_k^2)$  a f.g.m. pode ser simplificada. Neste caso temos que  $\boldsymbol{\omega} = \text{diag}(\sigma_1, \dots, \sigma_k)$ . Desta forma segue que a expressão em (C.15) resulta em

$$\bar{\boldsymbol{\Omega}} = \text{diag}(\sigma_1^{-1}, \dots, \sigma_k^{-1}) \text{diag}(\sigma_1^2, \dots, \sigma_k^2) \text{diag}(\sigma_1^{-1}, \dots, \sigma_k^{-1}) = I_k. \quad (\text{C.16})$$

Substituindo (C.16) em (C.14) obtemos

$$\mathbf{\Delta} = (1 + \boldsymbol{\alpha}'I_k\boldsymbol{\alpha})^{-\frac{1}{2}}I_k\boldsymbol{\alpha} = (1 + \boldsymbol{\alpha}'\boldsymbol{\alpha})^{-\frac{1}{2}}\boldsymbol{\alpha} = \left(1 + \sum_{i=1}^k \alpha_i^2\right)^{-\frac{1}{2}}\boldsymbol{\alpha}. \quad (\text{C.17})$$

Consequentemente, obtemos as seguintes expressões:

$$\mathbf{\Delta}'\boldsymbol{\omega}\mathbf{t} = \left(1 + \sum_{i=1}^k \alpha_i^2\right)^{-\frac{1}{2}} \sum_{i=1}^k \alpha_i \sigma_i t_i, \quad (\text{C.18})$$

$$\boldsymbol{\xi}'\mathbf{t} = \sum_{i=1}^k \xi_i t_i, \quad (\text{C.19})$$

$$\mathbf{t}'\boldsymbol{\Omega}\mathbf{t} = \sum_{i=1}^k t_i^2 \sigma_i^2. \quad (\text{C.20})$$

Substituindo (C.18), (C.19) e (C.20) em (C.13), temos que a f.g.m. é dada por

$$M(\mathbf{t}) = 2 \exp \left\{ \sum_{i=1}^k \xi_i t_i + \frac{1}{2} \sum_{i=1}^k t_i^2 \sigma_i^2 \right\} \Phi \left( \frac{1}{\sqrt{1 + \sum_{i=1}^k \alpha_i^2}} \sum_{i=1}^k \alpha_i \sigma_i t_i \right). \quad (\text{C.21})$$

Considere agora a seguinte combinação linear  $C'\mathbf{Y} = \sum_{i=1}^k c_i Y_i$  dos elementos do vetor aleatório  $\mathbf{Y}$ . Por propriedades de f.g.m. segue que  $M_{C'\mathbf{Y}}$  é dada por:

$$\begin{aligned} M_{C'\mathbf{Y}}(t) &= E[\exp\{tC'\mathbf{Y}\}] \\ &= M_{\mathbf{Y}}(tC) \\ &= 2 \exp \left\{ \sum_{i=1}^k \xi_i t c_i + \frac{1}{2} \sum_{i=1}^k t^2 c_i^2 \sigma_i^2 \right\} \Phi \left( \frac{1}{\sqrt{1 + \sum_{i=1}^k \alpha_i^2}} \sum_{i=1}^k \alpha_i \sigma_i t c_i \right) \\ &= 2 \exp \left\{ t \sum_{i=1}^k \xi_i c_i + \frac{t^2}{2} \sum_{i=1}^k c_i^2 \sigma_i^2 \right\} \Phi \left( \frac{t}{\sqrt{1 + \sum_{i=1}^k \alpha_i^2}} \sum_{i=1}^k \alpha_i \sigma_i c_i \right). \end{aligned}$$

Se, por hipótese,  $\sum_{i=1}^k \alpha_i \sigma_i c_i = 0$ , temos que a função geradora de momentos é simplificada para

$$M_{C'\mathbf{Y}}(t) = \exp \left\{ t \sum_{i=1}^k \xi_i c_i + \frac{t^2}{2} \sum_{i=1}^k c_i^2 \sigma_i^2 \right\}.$$

Portanto, conclui-se que  $C'\mathbf{Y} \sim N\left(\sum_{i=1}^k \xi_i c_i, \sum_{i=1}^k c_i^2 \sigma_i^2\right)$ , finalizando a prova.  $\square$

# Apêndice D

## Medidas de Qualidade de Ajuste dos Modelos

Para comparar os modelos ajustados vamos utilizar as medidas CPO, DIC, índice C e taxa de classificação correta. Para efeito de simplificação, vamos definir tais medidas de forma genérica, considerando um vetor de respostas binárias  $\mathbf{y} = (y_1, \dots, y_n)'$  com dimensão  $n \times 1$ . Destas observações binárias,  $n_1$  foram iguais a 1 e  $n_2$  são iguais a 0, em que  $n = n_1 + n_2$ . Além disso, considere que a distribuição de  $y_i$ ,  $i = 1, \dots, n$ , é indexada pelo vetor de parâmetros  $\boldsymbol{\theta}$  e que dispomos de uma amostra de tamanho  $L$  da distribuição *a posteriori* de  $\boldsymbol{\theta}$ .

### D.1 Conditional Predictive Ordinate (CPO)

A ordenada preditiva condicional (CPO) é uma medida que reflete a qualidade de ajuste do modelo e é definida como

$$CPO_i = f(y_i | \mathbf{y}_{(-i)}) = \int f(y_i | \boldsymbol{\theta}) f(\boldsymbol{\theta} | \mathbf{y}_{(-i)}) d\boldsymbol{\theta}, \quad (\text{D.1})$$

em que  $f(\boldsymbol{\theta} | \mathbf{y}_{(-i)})$  representa a distribuição *a posteriori* dado todas as respostas exceto  $y_i$ . A medida CPO pode ser interpretada como a altura da densidade marginal *a posteriori* de  $y_i$  dado  $\mathbf{y}_{(-i)}$  e maior CPO indica melhor modelo. Dey *et al.* (1997) propõe estimar a CPO por

$$C\hat{P}O_i = \left[ \frac{1}{L} \sum_{l=1}^L \frac{1}{f(y_i | \boldsymbol{\theta}^{(l)})} \right]^{-1}, \quad (\text{D.2})$$

em que  $\boldsymbol{\theta}^{(l)}$ ,  $l = 1, 2, \dots, L$ , são amostras das distribuições *a posteriori*, baseadas nos dados completos, para os parâmetros que indexam a distribuição de  $y_i$ . Uma alternativa para interpretar o CPO é a estatística resumo chamada de logaritmo da verossimilhança pseudo-marginal (LPML) definida por Geisser e Eddy (1979) como  $LPML = \sum_{i=1}^n \log(CPO_i)$ , em que  $n$  é o número de observações. Maior  $LPML$  indica melhor modelo.

### D.2 Deviance Information Criterion (DIC)

Defina por deviance  $D = -2 \log(f(\mathbf{y} | \boldsymbol{\theta}))$ . Uma amostra *a posteriori* para a deviance pode ser obtida fazendo  $D^{(l)} = -2 \log(f(\mathbf{y} | \boldsymbol{\theta}^{(l)}))$ . Seja  $\bar{D}$  a média a posteriori da deviance e considere

$\hat{D} = -2\log(f(\mathbf{y}|\bar{\boldsymbol{\theta}}))$ , em que  $\bar{\boldsymbol{\theta}}$  é média *a posteriori* de  $\boldsymbol{\theta}$ .

O critério de qualidade de ajuste DIC pode ser calculado por

$$DIC = \bar{D} + pD = \hat{D} + 2pD, \quad (\text{D.3})$$

em que  $pD = \bar{D} - \hat{D}$ . Menor DIC indica melhor modelo.

### D.3 Índice C

O índice C (ver Hanley e McNeil (1982)) é uma aproximação para a área sob a curva ROC que assume valores entre zero e um. A curva ROC representa uma relação entre probabilidades condicionais denominadas por sensibilidade e especificidade. Sensibilidade é a probabilidade do modelo prever o valor um para a observação dado que o verdadeiro valor observado é um, ou seja,  $P(\hat{y} = 1|y = 1)$ . Especificidade é a probabilidade do modelo prever valor zero para a observação dado que o verdadeiro valor observado é zero, ou seja,  $P(\hat{y} = 0|y = 0)$ . Quanto maior for o valor observado para o índice C melhor é capacidade de predição do modelo considerado. Em Souza e Migon (2010) encontramos como calcular tal índice. Primeiro separamos todos os indivíduos da amostra em dois grupos. O primeiro grupo, com tamanho  $n_1$ , é formado pelos indivíduos para os quais ocorreu o evento de interesse ( $y = 1$ ) e o outro grupo, com tamanho  $n_2$ , inclui os indivíduos para os quais o evento de interesse não ocorreu ( $y = 0$ ). Defina por  $\pi_{i1}^{(l)}$  a probabilidade, no passo  $l$ , do  $i$ -ésimo indivíduo do primeiro grupo ter resposta igual a um e por  $\pi_{j2}^{(l)}$  a probabilidade, no passo  $l$ , do  $j$ -ésimo indivíduo do segundo grupo ter resposta igual a um, em que  $i = 1, \dots, n_1$  e  $j = 1, \dots, n_2$ .

Um amostra da distribuição *a posteriori* do índice  $C$  pode ser obtida por

$$C^{(l)} = \frac{1}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} C_{ij}^{(l)},$$

em que  $C_{ij}^{(l)}$  é igual a um se  $\pi_{i1}^{(l)} > \pi_{j2}^{(l)}$  e igual a zero caso contrário.

### D.4 Taxa de Classificação Correta (TCC)

A taxa de classificação correta (TCC) é a proporção de indivíduos corretamente alocados pelo modelo segundo a ocorrência, ou não, do evento de interesse. Uma amostra da distribuição *a posteriori* da taxa de classificação correta pode ser obtida por

$$TCC^{(l)} = \frac{1}{n} \sum_{i=1}^n I(y_i, \hat{y}_i^{(l)}),$$

em que  $I(y_i, \hat{y}_i^{(l)})$  é igual a 1 se  $y_i = \hat{y}_i^{(l)}$  e igual a 0 caso contrário. O valor de  $\hat{y}_i^{(l)}$  é 1 se  $P(y_1 = 1) > c$  e 0 caso contrário, em que  $c$  é um ponto de corte pré-especificado.

Naturalmente, por este critério, o melhor modelo é aquele que apresenta maior TCC.

# Referências Bibliográficas

- Arellano-Valle, R., e Azzalini, A. (2006). On the unification of families of skew-normal distributions, *Scand. J. Statist.* **33**: 561–574.
- Arellano-Valle, R. B., Castro, L. M., Genton, M. G. e Gómez, H. W. (2008). Bayesian inference for shape mixtures of skewed distributions, with application to regression analysis, *Bayesian Analysis* **3**: 513–540.
- Arellano-Valle, R. B., Genton, M. G. e Loschi, R. H. (2009). Shape mixtures of multivariate skew-normal distributions, *Journal of Multivariate Analysis* **100**: 91–101.
- Arellano-Valle, R. e Genton, M. G. (2005). On fundamental skew distributions, *Journal of Multivariate Analysis* **96**: 93–116.
- Azzalini, A. (1985). A class of distributions which includes the normal ones, *Scand. J. Statist.* **12**: 171–178.
- Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica* **46**: 199–208.
- Azzalini, A. e Dalla-Valle, A. (1996). The multivariate skew-normal distribution, *Biometrika* **83**: 715–726.
- Breslow, N. E. e Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association* **88**: 9–25.
- Casella, G. e Robert, C. P. (2005). *Monte Carlo Statistical Methods*, Springer-Verlag, New York.
- Dellaportas, P. e Smith, A. F. M. (1993). Bayesian inference for generalized linear and proportional hazards models via gibbs sampling, *Appl Stat* **42**: 443–459.
- Dey, D. K., Chen, M. H. e Chang, H. (1997). Bayesian approach for nonlinear random effects models, *Biometrics* **53**: 1239–1252.
- Doornik, J. A. (2007). *Ox 5 - An Object-oriented Matrix Programming Language*, Timberlake Consultants Ltd.
- Gamerman, D. e Lopes, H. F. (2006). *Markov Chain Monte Carlo*, Chapman & Hall.
- Geisser, S. e Eddy, W. (1979). A predictive approach to model selection, *Journal of the American Statistical Association* **74**: 153–160.

- Gelfand, A. E. e Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities, *J Am Stat Assoc* **85**: 398–409.
- Genton, M. G. (ed.) (2004). *Skew-Elliptical Distributions and Their Applications. A Journey Beyond Normality*, Chapman & Hall.
- Geweke, J. (1992). *Bayesian Statistics*, Clarendon Press, Oxford, UK, chapter Evaluating the accuracy of sampling-based approaches to calculating posterior moments.
- Gibbons, R. D., Hedeker, D., Charles, S. C. e Frisch, P. (1994). A random-effects probit model for predicting medical malpractice claims, *Journal of the American Statistical Association* **89**: 760–767.
- Gilks, W. R., Best, N. G. e Tan, K. K. C. (1995). Adaptive rejection metropolis sampling within gibbs sampling, *Appl Stat* **44**: 455–472.
- Gilks, W. R. e Wild, P. (1992). Adaptive rejection sampling for gibbs sampling, *Appl Stat* **41**: 337–348.
- Gilks, W. R., Richardson, S. e Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics*, Chapman & Hall.
- Gupta, A. K., Gonzalez-Farias, G. e Dominguez-Molina, J. A. (2004). A multivariate skew normal distribution, *Journal of Multivariate Analysis* **89**: 181–190.
- Hanley, J. A. e McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (roc) curve, *Radiology* **143**: 29–36.
- Henze, N. (1986). A probabilistic representaton of the “skew-normal” distribution, *Scand. J. Statist.* **13**: 271–275.
- Landauer, W. e Clark, E. M. (1962). The interaction in teratogenic activity of the two niacin analogs 3-acetylpyridine and 6-aminonicotinamide, *Journal of Experimental Zoology* **151**: 253–258.
- Larsen, K. e Merlo, J. (2005). Appropriate assessment of neighborhood effects on individual health: Integrating random and fixed effects in multilevel logistic regression, *American Journal of Epidemiology* **161**: 81–88.
- Larsen, K., Petersen, J. H., Budtz-Jorgensen, E. e Endahl, L. (2000). Interpreting parameters in the logistic regression model with random effects, *Biometrics* **56**: 909–914.
- Liu, J. e Dey, D. K. (2008). Skew random effects in multilevel binomial models: an alternative to nonparametric approach, *statistical Modelling* **8**: 221–241.
- Liu, L. e Yu, Z. (2008). A likelihood reformulation method in non-normal random effects models, *Statistics in Medicine* **27**: 3105–3124.
- Marchenko, Y. V. e Genton, M. G. (2010). Multivariate log-skew-elliptical distributions with applications to precipitation data, *Environmetrics* **21**: 318–340.
- Marin, J.-M. e Robert, C. P. (2007). *Bayesian Core: A Practical Approach to Computational Bayesian Statistics*, Springer-Verlag, New York.

- McCulloch, C. e Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*, Wiley: New York.
- Nelson, K. P., Lipsitz, S. R., Fitzmaurice, G. M., Ibrahim, J., Parzen, M. e Strawderman, R. (2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects, *Journal of Computational and Graphical Statistics* **15**: 39–57.
- Paulino, C. D., Silva, G. e Achcar, J. A. (2005). Bayesian analysis of correlated misclassified binary data, *Computational Statistics & Data Analysis* **49**: 1120–1131.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Sahu, S. K., Dey, D. K. e Branco, M. D. (2003). A new class of multivariate skew distributions with applications to bayesian regression models, *The Canadian Journal of Statistics* **31**: 129–150.
- Schall, R. (1991). Estimation in generalized linear models with random effects, *Biometrika* **78**: 719–727.
- Seber, G. A. F. e Lee, A. J. (2003). *Linear Regression Analysis*, Wiley-Interscience.
- Souza, A. D. P. e Migon, H. S. (2010). Bayesian outlier analysis in binary regression, *Journal of Applied Statistics* **37**: 1355–1368.
- Yang, Y., Müller, P. e Rosner, G. L. (2010). Semiparametric bayesian inference for repeated fractional measurement data, *Chilean Journal of Statistics* **1**: 59–74.
- Zeger, S. L. e Karim, M. R. (1991). Generalized linear models with random effects; a gibbs sampling approach, *Journal of the American Statistical Association* **86**: 79–86.