

**Modelos de sobrevivência
com fração de cura
e efeitos aleatórios**

Célia Mendes Carvalho Lopes

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

Área de Concentração: Estatística
Orientador: Prof. Dr. Heleno Bolfarine

São Paulo, maio de 2008

Modelos de sobrevivência com fração de cura e efeitos aleatórios

Este exemplar corresponde à redação
final da tese devidamente corrigida
e defendida por Célia Mendes Carvalho Lopes
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Heleno Bolfarine (orientador) - IME-USP
- Prof. Dr. Antonio Carlos Pedroso de Lima - IME-USP
- Prof. Dr. Francisco Louzada Neto - UFSCar
- Prof. Dr. Enrico Antonio Colosimo - UFMG
- Prof. Dr. Dalton Francisco de Andrade - UFSC

*Ao Gizelton,
Aos meus pais Ceci e Plínio*

Agradecimentos

Gostaria de agradecer a todos que, direta ou indiretamente, me ajudaram durante este longo período em que fiz o doutorado. Em especial, gostaria de agradecer:

Ao Prof. Dr. Heleno Bolfarine, meu orientador, pelos ensinamentos e pela paciência.

Ao Prof. Dr. Antonio Carlos Pedroso de Lima, pelas valiosas sugestões durante meu exame de qualificação.

Ao membros da Banca Examinadora, pelas preciosas sugestões, considerações, correções e incentivos que recebi durante a defesa desta Tese.

Ao Gizelton, pelo carinho, pela compreensão, por estar ao meu lado e por todo o super apoio e incentivo na confecção deste trabalho.

Aos amigos Lane Alencar e Marcelo Rocha, pelo super apoio, amizade, ensinamentos e paciência.

Aos meus pais Ceci e Plínio, minhas irmãs, meus avós (in memoriam), sogros, cunhados (valeu Ros!), sobrinhos e demais familiares, pelo apoio.

Às minhas amigas do Santa Cruz: Simone Hashimoto, Luciana Bertini, Priscila Bara e Renata Marson - pela amizade e pela torcida.

Aos amigos Solange, Alcides, Lucia, Prof. Marcel e demais colegas do Mackenzie, obrigada pela força!

Ao Prof. Dr. Plinio Simões, meu orientador de mestrado, pelo eterno apoio que nos proporciona.

Aos meus amigos imeanos, sempre vou lembrar-me de vocês com carinho.

Finalmente, agradeço a todos que, mesmo eu não tendo citado aqui nominalmente, me ajudaram com apoio, incentivo e torcida. Muito obrigada!

Resumo

Neste trabalho são apresentados dois modelos de sobrevivência com fração de cura e efeitos aleatórios, um baseado no modelo de Chen-Ibrahim-Sinha para fração de cura e o outro, no modelo de mistura. São estudadas abordagens clássica e bayesiana. Na inferência clássica são utilizados estimadores REML. Para a bayesiana foi utilizado Metropolis-Hastings. Estudos de simulação são feitos para avaliar a acurácia das estimativas dos parâmetros e seus respectivos desvios-padrão. O uso dos modelos é ilustrado com uma análise de dados de câncer na orofaringe.

Palavras-chave: Fração de cura, efeitos aleatórios, REML, Metropolis-Hastings.

Abstract

In this work, it is shown two survival models with long term survivors and random effects, one based on Chen-Ibrahim-Sinha model for models with surviving fraction and the other, on mixture model. We present bayesian and classical approaches. In the first one, we use Metropolis-Hastings. For the second one, we use the REML estimators. A simulation study is done to evaluate the accuracy of the applied techniques for the estimatives and their standard deviations. An example on orofaringe cancer is used to illustrate the models considered in the study.

Keywords: Surviving fraction, random effects, REML, Metropolis-Hastings.

Sumário

Lista de Figuras	xi
Introdução	1
1 Modelos de sobrevivência com fração de cura	3
1.1 Modelo de mistura	3
1.2 Modelo de Chen-Ibrahim-Sinha	4
1.3 Modelo de mistura com efeitos aleatórios	6
2 Modelo de Chen-Ibrahim-Sinha com efeitos aleatórios	7
2.1 Definindo o modelo	7
2.2 Inferência bayesiana	10
2.3 Análise clássica	15
3 Estudos de simulação - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios	19
3.1 Características	19
3.2 Simulação - abordagem bayesiana	20
3.2.1 Caso $n_k = 15$	20
3.2.2 Caso $n_k = 50$	26
3.3 Simulação - abordagem clássica	32
3.3.1 Caso $n_k = 15$	32
3.3.2 Caso $n_k = 50$	37
3.4 Conclusões	42
4 Análise de dados - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios	45
4.1 Dados utilizados	45
4.2 Inferência bayesiana via Metropolis-Hastings	48
4.3 Análise clássica	51

4.4	Conclusões	53
5	Modelo de mistura com efeitos aleatórios	55
5.1	Definindo o modelo	55
5.2	Inferência bayesiana	56
5.3	Análise clássica	58
6	Estudos de simulação - modelo de mistura com efeitos aleatórios	61
6.1	Características	61
6.2	Simulação - abordagem bayesiana	62
6.2.1	Caso $n_k = 15$	62
6.2.2	Caso $n_k = 50$	68
6.3	Simulação - abordagem clássica	74
6.3.1	Caso $n_k = 15$	74
6.3.2	Caso $n_k = 50$	79
6.4	Conclusões	84
7	Análise de dados - modelo de mistura com efeitos aleatórios	87
7.1	Inferência bayesiana via Metropolis-Hastings	87
7.2	Análise clássica	89
7.3	Conclusões	90
8	Conclusões	91
8.1	Resultados e conclusões	91
8.2	Perspectivas futuras	92
A	Programas	95
A.1	Gerando os dados para a simulação - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios	95
A.2	Comandos em Winbugs - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios	97
A.3	Comandos em Maple - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios	98
A.4	Gerando os dados para a simulação - modelo de mistura com efeitos aleatórios	101
A.5	Comandos em Winbugs - modelo de mistura com efeitos aleatórios	102
A.6	Comandos em Maple - modelo de mistura com efeitos aleatórios	103
	Referências Bibliográficas	109

Lista de Figuras

3.1	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – parte 1	23
3.2	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – parte 2	24
3.3	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – parte 3	25
3.4	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – parte 1	29
3.5	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – parte 2	30
3.6	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – parte 3	31
3.7	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – parte 1	35
3.8	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – parte 2	36
3.9	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – parte 3	37
3.10	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – parte 1	40
3.11	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – parte 2	41
3.12	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – parte 3	42
4.1	Gráficos de Kaplan-Meier para cada uma das seis clínicas	48
4.2	Histogramas – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings	49
4.3	Histogramas – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings - com alfa incluído	50
4.4	$\sqrt{\hat{R}}$ – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings	51
6.1	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – modelo de mistura – parte 1	65

6.2	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – modelo de mistura – parte 2	66
6.3	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – modelo de mistura – parte 3	67
6.4	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – modelo de mistura – parte 1	71
6.5	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – modelo de mistura – parte 2	72
6.6	Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – modelo de mistura – parte 3	73
6.7	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – modelo de mistura – parte 1	77
6.8	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – modelo de mistura – parte 2	78
6.9	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – modelo de mistura – parte 3	79
6.10	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – modelo de mistura – parte 1	82
6.11	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – modelo de mistura – parte 2	83
6.12	Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – modelo de mistura – parte 3	84
7.1	Histogramas – câncer na orofaringe – modelo de mistura com efeitos aleatórios – Metropolis-Hastings	88
7.2	$\sqrt{\hat{R}}$ – câncer na orofaringe – modelo de mistura com efeitos aleatórios – Metropolis-Hastings	89

Introdução

Em estudos de análise de sobrevivência, avalia-se o tempo até que certo evento ocorra. Entretanto, existem casos em que parte dos indivíduos pode não apresentar o evento. Por exemplo, se um grupo de pacientes que passou por um tratamento de um mesmo tipo de câncer é acompanhado para se conhecer o tempo até o reaparecimento da doença, uma parcela deles poderá ter recidivas e outra não apresentar a recorrência do evento. Por isso o interesse em se estudar a análise de sobrevivência com fração de cura, que inclui a presença de imunes no estudo.

Muitas vezes, ao se analisar os dados de sobrevivência, percebe-se que há um grande número de indivíduos censurados à direita. Esse fato pode ser uma indicação de que, em parte dos indivíduos em estudo, o evento de interesse nunca acontecerá, ou seja, pode se ter uma fração da população que é imune (ou que se cura).

Muitos pesquisadores vêm estudando modelos de sobrevivência com fração de cura.

Em 1952, os pesquisadores J. Berkson e R. Gage publicaram um artigo (Berkson; Gage, 1952) em que apresentam um modelo de mistura para modelar os casos em que há uma proporção p de indivíduos curados. Esse modelo será apresentado na seção 1.1.

Em 1999, foi publicado, por M.-H. Chen, J. G. Ibrahim e D. Sinha, um artigo (Chen et al., 1999) em que os modelos de sobrevivência com fração de cura são modelados baseado-se no número não observado de células susceptíveis a ter câncer que o indivíduo possui. Neste artigo, são abordadas algumas diferenças em relação ao modelo de mistura. Esse modelo, apresentado na seção 1.2 será denominado modelo de Chen-Ibrahim-Sinha.

Em muitos casos de análise de sobrevivência, os dados do estudo são provenientes de alguns grupos específicos, por exemplo, de várias clínicas, sendo que as respostas dos indivíduos em uma mesma clínica podem estar correlacionadas. Nesses casos, são adequados estudos via modelos com efeitos aleatórios, de modo que se possa avaliar o efeito da clínica no tratamento, por exemplo, do câncer.

Em 2001, foi publicado um trabalho de K. K. W. Yau e A. S. K. Ng (Yau; Ng, 2001) em que o modelo de mistura é analisado acrescentando-se efeitos aleatórios que pudessem medir o efeito do tratamento feito pelas diferentes clínicas nos pacientes que se curam e nos que não se curam. Tal modelo será relatado na seção 1.3.

Neste trabalho, é proposto o uso de efeitos aleatórios no modelo de Chen-Ibrahim-Sinha. O estudo é feito com enfoques bayesiano e clássico. Também é apresentado um estudo de simulação para ambas as abordagens. Para ilustrar a aplicação do modelo, é apresentada uma análise com os dados de Kalbfleisch e Prentice (1980, pp. 225-229), o mesmo conjunto de dados estudados no artigo de Yau e Ng (2001).

Também será analisado o modelo de mistura com efeitos aleatórios atribuídos somente à parcela de curados. Este segundo modelo será analisado com abordagens bayesiana e clássica. Do mesmo modo que para o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, será apresentado um estudo de simulação e uma aplicação a dados reais para ilustrar o uso do modelo.

Capítulo 1

Modelos de sobrevivência com fração de cura

Neste capítulo são apresentados modelos já existentes na literatura. Primeiro é apresentado o modelo de mistura (Berkson; Gage, 1952) para estudos de casos em que há presença de imunes (ou com fração de cura). Em seguida, tem-se o modelo de Chen-Ibrahim-Sinha (Chen et al., 1999), com uma abordagem alternativa para os casos com presença de imunes. Para finalizar, é apresentado o modelo de mistura com efeitos aleatórios para curados e para não curados (Yau; Ng, 2001).

1.1 Modelo de mistura

Berkson e Gage (1952) propõem um modelo de mistura para os casos em que uma parte da população apresenta a possibilidade de ser curada. Este modelo também é descrito de modo sucinto em Yau e Ng (2001).

Suponha que há uma fração p de curados na população e que os restantes $\pi = 1 - p$ são não curados. Seja T o tempo de ocorrência do evento (falha). Suponha que $T = \infty$ para os indivíduos curados, de modo que a correspondente função de sobrevivência condicional é 1, para todos os valores finitos de tempo, ou seja, se S_1 é a função de sobrevivência para o grupo de curados, tem-se que $S_1(t) = 1, \forall t \geq 0$.

Suponha ainda que a função de sobrevivência da população inteira seja dada por $S(t)$ e dos não curados seja dada por $S_2(t)$. Assim,

$$\begin{aligned} S(t) &= pS_1(t) + (1 - p)S_2(t) \\ &= 1 - \pi + \pi S_2(t). \end{aligned}$$

Defina uma variável indicadora de não-curados por $Y = 0$, se o indivíduo for, eventualmente, curado e $Y = 1$, caso contrário. Considere, ainda, o vetor de covariáveis \mathbf{x}_j para o indivíduo j . Assim, a proporção π de não curados pode ser modelada por uma função logística em \mathbf{x} , de modo que

$$\pi(\mathbf{x}_j) = P(Y = 1|\mathbf{x}_j) = \frac{1}{1 + e^{\mathbf{w}_j^T \boldsymbol{\gamma}}}$$

em que $\mathbf{w}_j = [1 \ \mathbf{x}_j^\top]^\top$ e $\boldsymbol{\gamma}$ é um vetor de parâmetros.

Se assumirmos que a função de risco condicional de T para o grupo de não curados, $h_2(t; \mathbf{x})$, segue o modelo de riscos proporcionais de Cox (Cox, 1972) então

$$h_2(t; \mathbf{x}) = h_{20}(t) \cdot e^{\mathbf{x}^\top \boldsymbol{\beta}}$$

em que $\boldsymbol{\beta}$ é o vetor de parâmetros da regressão e $h_{20}(t)$ é a função de riscos basal para o grupo de não curados. Assim, a função de risco não condicional de T é dada por

$$h(t; \mathbf{x}) = \pi(\mathbf{x}) h_{20}(t) e^{\mathbf{x}^\top \boldsymbol{\beta}} \frac{S_{20}(t)^{\exp(\mathbf{x}^\top \boldsymbol{\beta})}}{S(t; \mathbf{x})}$$

Neste caso, abordagens paramétricas e semi-paramétricas podem ser adotadas. No caso de uma abordagem paramétrica, a função de risco basal é modelada parametricamente, por exemplo, seguindo uma distribuição Weibull ou log-normal.

Se considerarmos ν_j o indicador de falha para a j -ésima observação, e se N é o número total de observações então a verossimilhança é dada por

$$\begin{aligned} L &= \prod_{j=1}^N f(t_j; \mathbf{x}_j)^{\nu_j} S(t_j; \mathbf{x}_j)^{1-\nu_j} \\ &= \prod_{j=1}^N S(t_j; \mathbf{x}_j)^{1-\nu_j} (h(t_j; \mathbf{x}_j) S(t_j; \mathbf{x}_j))^{\nu_j} \\ &= \prod_{j=1}^N S(t_j; \mathbf{x}_j)^{1-\nu_j} \left(\pi(\mathbf{x}) h_{20}(t) e^{\mathbf{x}^\top \boldsymbol{\beta}} S_{20}(t)^{\exp(\mathbf{x}^\top \boldsymbol{\beta})} \right)^{\nu_j} \end{aligned} \quad (1.1)$$

de modo que a log-verossimilhança é dada por

$$l = \sum_{j=1}^N \left\{ (1 - \nu_j) \ln(S(t_j; \mathbf{x}_j)) + \nu_j \ln \left(\pi(\mathbf{x}) h_{20}(t) e^{\mathbf{x}^\top \boldsymbol{\beta}} S_{20}(t)^{\exp(\mathbf{x}^\top \boldsymbol{\beta})} \right) \right\} \quad (1.2)$$

1.2 Modelo de Chen-Ibrahim-Sinha

O modelo proposto em Chen et al. (1999) é definido do seguinte modo: seja N o número de células susceptíveis ao aparecimento de câncer (chamadas células carcinogênicas). Suponha que N tem distribuição Poisson com média θ . Seja Z_c a variável aleatória para o tempo até que a c -ésima célula carcinogênica produza um câncer detectável. Supõe-se que as variáveis Z_c , $c = 1, 2, \dots$ são iid com função de distribuição $F(t) = 1 - S(t)$ e também são independentes de N . Assim, o tempo para o aparecimento do câncer pode ser definido pela variável aleatória $T = \min\{Z_c, 0 \leq c \leq N\}$ em que $P(Z_0 = \infty) = 1$ e N é independente da seqüência Z_1, Z_2, \dots . Desse modo, a função de sobrevivência

de T é dada por

$$\begin{aligned}
S_p(t) &= P(T > t) \\
&= \sum_{k=0}^{\infty} P(\min\{Z_1, Z_2, \dots, Z_k\} > t) \cdot P(N = k) = \\
&= P(N = 0) + \sum_{k=1}^{\infty} P(Z_1 > t, Z_2 > t, \dots, Z_k > t) \cdot P(N = k) = \\
&= P(N = 0) + \sum_{k=1}^{\infty} P(Z_1 > t)P(Z_2 > t) \cdots P(Z_k > t) \cdot P(N = k) = \\
&= e^{-\theta} + \sum_{k=1}^{\infty} S(t)^k \frac{e^{-\theta} \theta^k}{k!} = e^{-\theta} \left[1 + \sum_{k=1}^{\infty} \frac{(S(t)\theta)^k}{k!} \right] = \\
&= e^{-\theta} \left[\sum_{k=0}^{\infty} \frac{(\theta S(t))^k}{k!} \right] = e^{-\theta} e^{\theta S(t)} = e^{-\theta + \theta S(t)} = \\
&= e^{-\theta(1-S(t))} = e^{-\theta F(t)}
\end{aligned}$$

Note que, portanto, a fração de cura é calculada por

$$S_p(\infty) = \lim_{t \rightarrow \infty} S_p(t) = \lim_{t \rightarrow \infty} e^{-\theta F(t)} = e^{-\theta}.$$

Para o indivíduo j , com $j = 1, 2, \dots, n$, considere $y_j = \min\{T_j, C_j\}$ em que $T_j = \min\{Z_{j0}, Z_{j1}, \dots, Z_{jN_j}\}$, sendo $N_j \sim \text{Poisson}(\theta_j)$, e C_j a variável de censura à direita. Considere também que as censuras não informativas C_j são independentes. Considere, ainda, a função indicadora de falha ν_j . Para $N_j = n_j$, Chen et al. (1999) e Mizoi (2004) mostram que

$$f(y_j, \nu_j | n_j) = S(y_j)^{n_j - \nu_j} [n_j f(y_j)]^{\nu_j} \quad (1.3)$$

Para modelar θ_j , podemos utilizar covariáveis \mathbf{x}_j . Como θ_j é positivo, pode utilizar, por exemplo, $\theta_j = e^{\mathbf{w}_j^\top \boldsymbol{\gamma}}$.

Além disso, pode-se especificar um modelo paramétrico para o tempo de sobrevivência. Nesse caso, tem-se que $f(y_j) = f(y_j | \boldsymbol{\lambda}_j)$. Muitas vezes, segundo Chen et al. (1999), a distribuição escolhida é Weibull ou Gama. Note que nesses casos, $\boldsymbol{\lambda}_j = (\lambda, \alpha)$.

A distribuição Weibull utilizada em Chen et al. (1999) tem função densidade dada por $f(y) = \alpha y^{\alpha-1} \exp\{\lambda - y^\alpha \exp(\lambda)\}$, de modo que $\alpha > 0$ e $\lambda \in R$.

Para modelar λ_j , podemos utilizar covariáveis \mathbf{x}_j . Por exemplo, $\lambda_j = \mathbf{x}_j^\top \boldsymbol{\beta}$.

1.3 Modelo de mistura com efeitos aleatórios

Yau e Ng (2001) apresentam uma forma de se trabalhar o modelo de mistura com efeitos aleatórios, que muitas vezes são necessários em estudos quando um as respostas dos indivíduos em uma mesma clínica podem estar correlacionadas e um dos interesses do estudo é avaliar a diferença de tratamento entre clínicas diferentes.

as respostas dos indivíduos em uma mesma clínica podem estar correlacionadas

Yau e Ng (2001) consideram o modelo de mistura com fração de cura dada por $1 - \pi$:

$$S(t; \mathbf{x}) = 1 - \pi(\mathbf{x}) + \pi(\mathbf{x})S_2(t; \mathbf{x}),$$

em que S_2 é a função de sobrevivência da parcela da população que não se cura e \mathbf{x} são covariáveis.

Nesse caso, π é estimado via função logística do seguinte modo. Suponha que há M clínicas e que T_{jk} é tempo observado de falha/censura para o j -ésimo indivíduo da k -ésima clínica. Seja ainda n_k o número de observações na k -ésima clínica e $N = \sum_{k=1}^M n_k$ o número total de observações. Então,

$$\pi(\mathbf{x}_{jk}) = \frac{1}{1 + e^{\xi_{jk}}}$$

com

$$\xi_{jk} = \mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k$$

para $k = 1, 2, \dots, M$; $j = 1, 2, \dots, n_k$; $\mathbf{w}_{jk} = [1 \ \mathbf{x}_{jk}^\top]^\top$ e V_k é o efeito aleatório correspondente à k -ésima clínica e que está relacionado com a fração de cura. É suposto que $V_k, k = 1, \dots, M$ são iid $N(0, \theta_1)$.

Em relação ao modelo de riscos proporcionais de Cox, em que $h_2(t; \mathbf{x})$ é a função de riscos proporcionais de T , e h_{20} é a função de risco basal para o grupo de não curados, tem-se

$$h_2(t_{jk}; \mathbf{x}_{jk}) = h_{20}(t_{jk})e^{\eta_{jk}}$$

com

$$\eta_{jk} = \mathbf{x}_{jk}^\top \boldsymbol{\beta} + U_k$$

para $k = 1, 2, \dots, M$; $j = 1, 2, \dots, n_k$; e U_k é a variável não observável correspondente ao efeito aleatório da k -ésima clínica, que determina o risco de formação de tumor para indivíduos não curados. É suposto que U_k são iid $N(0, \theta_2)$.

Considerando $\mathbf{V} = [V_1, V_2, \dots, V_M]$ e $\mathbf{U} = [U_1, U_2, \dots, U_M]$, o estimador para $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{V}, \mathbf{U})$ é obtido maximizando-se a verossimilhança via BLUP (Henderson, 1975) em um primeiro passo e então estendido para se obter estimadores REML para $(\boldsymbol{\gamma}, \boldsymbol{\beta}, \mathbf{V}, \mathbf{U}, \theta_1, \theta_2)$.

Capítulo 2

Modelo de Chen-Ibrahim-Sinha com efeitos aleatórios

Neste capítulo é apresentado o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, que visam avaliar se as respostas de indivíduos em uma mesma clínica estão correlacionadas, por exemplo, no tratamento de câncer de certa população que tem presença de imunes. É feita a caracterização do modelo e apresentada uma abordagem bayesiana e uma clássica para a obtenção de estimadores.

2.1 Definindo o modelo

Suponha que foram estudados casos em M clínicas e que para a k -ésima clínica havia n_k indivíduos, $k = 1, 2, \dots, M$. Para $j = 1, 2, \dots, n_k$, considere N_{jk} o número de células susceptíveis a terem câncer que o indivíduo j da clínica k possui. Suponha ainda que N_{jk} são independentes e com distribuição de Poisson com média θ_{jk} . Para cada $k = 1, 2, \dots, M$, defina o vetor $\mathbf{N}_k = [N_{1k}, N_{2k}, \dots, N_{n_k k}]^\top$.

Seja Z_{cjk} a variável aleatória para o tempo até que a c -ésima célula carcinogênica do indivíduo j da clínica k produza um câncer detectável. Suponha que as variáveis Z_{cjk} , são iid com função de distribuição $F(t|\boldsymbol{\lambda}) = 1 - S(t|\boldsymbol{\lambda})$ e também são independentes de N_{jk} . Assim, o tempo para o aparecimento do câncer pode ser definido pela variável aleatória $T_{jk} = \min\{Z_{cjk}, 1 \leq c \leq N_{jk}\}$ mas $T_{jk} = \infty$ com probabilidade 1, se $N_{jk} = 0$.

Nesse trabalho, vamos utilizar formas paramétricas para $F(\cdot)$, como, por exemplo, Weibull. Neste caso, $\boldsymbol{\lambda} = (\lambda, \alpha)$ e $f(y|\boldsymbol{\lambda}) = \alpha y^{\alpha-1} e^{\lambda - y^\alpha} e^\lambda$ e $S(y|\boldsymbol{\lambda}) = e^{-y^\alpha} e^\lambda$, com $\alpha > 0$ e $\lambda \in R$.

Como mostrado na seção 1.2, a fração de cura é obtida por $e^{-\theta_{jk}}$. Aqui, os valores de θ_{jk} dependerão das covariáveis \mathbf{x}_{jk} por meio da relação $\theta_{jk}(\mathbf{w}_{jk}) = e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k}$, em que $\mathbf{w}_{jk} = [1 \ \mathbf{x}_{jk}^\top]^\top$ são covariáveis fixas, V_k é a variável correspondente ao efeito aleatório referente à clínica k , e vamos supor que V_1, \dots, V_M são independentes com $V_k \sim N(0, \sigma_V^2)$.

Para se determinar a verossimilhança, considere C_{jk} o tempo de censura (à direita) para o j -ésimo indivíduo da clínica k . Seja $Y_{jk} = \min\{T_{jk}, C_{jk}\}$ e seja ν_{jk} o indicador de falha, ou seja, $\nu_{jk} = 1$ se y_{jk} é falha e 0 se for censura.

Assim, os dados observados são $D_{obs} = (\mathbf{y}, \boldsymbol{\nu}, \mathbf{X})$ em que $\mathbf{y} = (y_{jk})_{j=1, \dots, n_k}^{k=1, \dots, M}$, $\boldsymbol{\nu} = (\nu_{jk})_{j=1, \dots, n_k}^{k=1, \dots, M}$ e \mathbf{X}

é a matriz de covariáveis.

Os dados completos são $D_c = (\mathbf{y}, \boldsymbol{\nu}, \mathbf{N}, \mathbf{V}, \mathbf{X})$ que incluem os dados observados e os dados não observados (\mathbf{N}, \mathbf{V}) , com $\mathbf{N} = [\mathbf{N}_1, \mathbf{N}_2, \dots, \mathbf{N}_M]$ e $\mathbf{V} = [V_1, V_2, \dots, V_M]$. O vetor de parâmetros é dado por $\boldsymbol{\phi} = (\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \sigma_V^2)$.

Assim, utilizando (1.3), temos

$$\begin{aligned}
f(\mathbf{y}, \mathbf{N}, \mathbf{V}) &= f(\mathbf{y}|\mathbf{N}, \mathbf{V}) \cdot f(\mathbf{N}|\mathbf{V}) \cdot f(\mathbf{V}) \\
&= f(\mathbf{y}|\mathbf{N}, \mathbf{V}) \cdot f(\mathbf{N}) \cdot f(\mathbf{V}) \\
&= \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ S(y_{jk}|\boldsymbol{\lambda})^{N_{jk}-\nu_{jk}} [N_{jk}f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \right\} \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ \frac{\theta_{jk}^{N_{jk}}}{N_{jk}!} e^{-\theta_{jk}} \right\} \\
&\quad \cdot \prod_{k=1}^M \left\{ \frac{1}{\sqrt{2\pi\sigma_V^2}} e^{\frac{-V_k^2}{2\sigma_V^2}} \right\} \\
&= \prod_{k=1}^M \left\{ \frac{1}{\sqrt{2\pi\sigma_V^2}} e^{\frac{-V_k^2}{2\sigma_V^2}} \right\} \\
&\quad \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ S(y_{jk}|\boldsymbol{\lambda})^{N_{jk}-\nu_{jk}} [N_{jk}f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \cdot \frac{\theta_{jk}^{N_{jk}}}{N_{jk}!} e^{-\theta_{jk}} \right\} \\
&= \prod_{k=1}^M \{*_k\} \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \{\star_{jk}\}
\end{aligned}$$

em que $*_k = \frac{1}{\sqrt{2\pi\sigma_V^2}} e^{\frac{-V_k^2}{2\sigma_V^2}}$ e $\star_{jk} = S(y_{jk}|\boldsymbol{\lambda})^{N_{jk}-\nu_{jk}} [N_{jk}f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \cdot \frac{\theta_{jk}^{N_{jk}}}{N_{jk}!} e^{-\theta_{jk}}$.

Assim, a verossimilhança completa é dada por

$$L(\boldsymbol{\phi}|D_c) = \prod_{k=1}^M \left\{ \frac{e^{\frac{-V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \right\} \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ S(y_{jk}|\boldsymbol{\lambda})^{N_{jk}-\nu_{jk}} [N_{jk}f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \frac{\theta_{jk}^{N_{jk}} e^{-\theta_{jk}}}{N_{jk}!} \right\}. \quad (2.1)$$

As variáveis correspondentes aos efeitos aleatórios (\mathbf{V}) e os números de células carcinogênicas (\mathbf{N}) são não observáveis. Então o estimador de máxima verossimilhança pode ser obtido maximizando-se a verossimilhança marginal, ou seja, integrando (somando) com relação às variáveis não observadas. Uma expressão simplificada pode ser obtida somando-se somente com relação ao vetor \mathbf{N} .

Desse modo, tem-se que

$$\begin{aligned} \sum_{\mathbf{N}} L(\phi|\mathbf{y}, \boldsymbol{\nu}, \mathbf{V}, \mathbf{X}) &= \sum_{\mathbf{N}} \left[\prod_{k=1}^M (*_k) \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} (*_{jk}) \right] \\ &= \prod_{k=1}^M (*_k) \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \sum_{N_{jk}=0}^{\infty} (*_{jk}) \end{aligned}$$

Mas,

$$\begin{aligned} \sum_{N_{jk}=0}^{\infty} (*_{jk}) &= \sum_{N_{jk}=0}^{\infty} \left\{ S(y_{jk}|\boldsymbol{\lambda})^{N_{jk}-\nu_{jk}} [N_{jk} f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \cdot \frac{\theta_{jk}^{N_{jk}}}{N_{jk}!} e^{-\theta_{jk}} \right\} \\ &= \sum_{N_{jk}=0}^{\infty} \left\{ S(y_{jk}|\boldsymbol{\lambda})^{N_{jk}} S(y_{jk}|\boldsymbol{\lambda})^{-\nu_{jk}} [N_{jk}]^{\nu_{jk}} [f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \frac{\theta_{jk}^{N_{jk}}}{N_{jk}!} e^{-\theta_{jk}} \cdot \right. \\ &\quad \left. \cdot e^{-\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})} e^{\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})} \right\} \\ &= \sum_{N_{jk}=0}^{\infty} \left\{ \frac{[N_{jk}]^{\nu_{jk}} [\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})]^{N_{jk}} e^{-\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})}}{N_{jk}!} \cdot \right. \\ &\quad \left. \cdot S(y_{jk}|\boldsymbol{\lambda})^{-\nu_{jk}} e^{-\theta_{jk}} e^{\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})} [f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \right\} \\ &= \theta_{jk}^{\nu_{jk}} e^{-\theta_{jk}} e^{\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})} [f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \end{aligned}$$

Portanto,

$$\begin{aligned} \sum_{\mathbf{N}} L(\phi|\mathbf{y}, \boldsymbol{\nu}, \mathbf{V}, \mathbf{X}) &= \prod_{k=1}^M (*_k) \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \sum_{N_{jk}=0}^{\infty} (*_{jk}) \\ &= \prod_{k=1}^M \left\{ \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \right\} \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ \theta_{jk}^{\nu_{jk}} e^{-\theta_{jk}} e^{\theta_{jk} S(y_{jk}|\boldsymbol{\lambda})} [f(y_{jk}|\boldsymbol{\lambda})]^{\nu_{jk}} \right\} \end{aligned}$$

Se $y_{jk} \sim Weibull(\alpha, \lambda_{jk})$, então

$$\begin{aligned} f(y_{jk}|\boldsymbol{\lambda}) &= \alpha y_{jk}^{\alpha-1} e^{\lambda_{jk} - y_{jk}^{\alpha}} e^{\lambda_{jk}}, \\ S(y_{jk}|\boldsymbol{\lambda}) &= e^{-y_{jk}^{\alpha}} e^{\lambda_{jk}}, \end{aligned}$$

em que

$$\lambda_{jk} = \mathbf{x}_{jk}^{\top} \boldsymbol{\beta}.$$

Assim,

$$\sum_{\mathbf{N}} L(\phi | \mathbf{y}, \boldsymbol{\nu}, \mathbf{V}, \mathbf{X}) = \prod_{k=1}^M \left\{ \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \right\} \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ (e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k})^{\nu_{jk}} e^{-e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k}} e^{(e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k})(e^{-y_{jk}^\alpha} e^{\lambda_{jk}})} \left[\alpha y_{jk}^{\alpha-1} e^{\lambda_{jk} - y_{jk}^\alpha} e^{\lambda_{jk}} \right]^{\nu_{jk}} \right\}. \quad (2.2)$$

2.2 Inferência bayesiana

Na abordagem bayesiana, a presença de quantidades não observáveis não representa um problema de inferência já que toda e qualquer quantidade desconhecida é uma variável aleatória. Assim, tanto para as quantidades de interesse (parâmetros), quanto para os efeitos aleatórios, é necessário obter sua distribuição a posteriori, ou seja, condicional aos dados observados. Basicamente, a distribuição a posteriori é proporcional ao produto da distribuição a priori para as quantidades desconhecidas e da função de verossimilhança. Uma completa apresentação da teoria bayesiana pode ser encontrada em Berger (1985).

Para o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, supondo prioris não informativas para $\boldsymbol{\gamma}$ e para $\boldsymbol{\beta}$, isto é, $\pi(\boldsymbol{\gamma}^\top) \propto 1$ e $\pi(\boldsymbol{\beta}^\top) \propto 1$, consideremos as distribuições a priori para o vetor de parâmetros $\boldsymbol{\phi}$:

$$\pi(\boldsymbol{\phi}) = \pi(\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \sigma_V^2) \propto \pi(\alpha) \pi(\sigma_V^2),$$

em que as distribuições a priori são tais que

1. $\alpha \sim \text{Gama}(\delta_0, \tau_0)$

$$\pi(\alpha) \propto \alpha^{\delta_0-1} e^{-\tau_0 \alpha};$$

2. $\sigma_V^2 \sim \text{Gama Inversa}(a_V, b_V)$

$$\pi(\sigma_V^2) \propto (\sigma_V^2)^{-(a_V+1)} e^{-\frac{b_V}{\sigma_V^2}}.$$

Considerando a função de verossimilhança em (2.2), a distribuição a posteriori é dada por

$$\begin{aligned}
\pi(\boldsymbol{\phi}, \mathbf{V} | D_{obs}) &= \sum_{\mathbf{N}} L(\boldsymbol{\phi} | D_c) \pi(\boldsymbol{\phi}) \\
&\propto \prod_{k=1}^M \left\{ \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \right\} \\
&\quad \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ (e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k})^{\nu_{jk}} e^{-e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k}} e^{(e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k}) (e^{-y_{jk}^\alpha} e^{\lambda_{jk}})} \left[\alpha y_{jk}^{\alpha-1} e^{\lambda_{jk} - y_{jk}^\alpha} e^{\lambda_{jk}} \right]^{\nu_{jk}} \right\} \\
&\quad \cdot \alpha^{\delta_0-1} e^{-\tau_0 \alpha} (\sigma_V^2)^{-(a_V+1)} e^{-\frac{b_V}{\sigma_V^2}}. \tag{2.3}
\end{aligned}$$

Uma forma de se obter estimativas dos parâmetros é via utilização de amostrador de Gibbs (Geman; Geman, 1984), método de simulação MCMC para aproximar as distribuições conjuntas e marginais, amostrando das distribuições condicionais. Assim, o próximo passo é tentar identificar as distribuições condicionais, a partir de (2.3).

Primeiro, vamos mostrar que $\sigma_V^2 | \alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{V} \sim \text{Gama Inversa} \left(a_V + \frac{1}{2}, b_V + \sum_{k=1}^M \frac{V_k^2}{2} \right)$. De fato:

$$\begin{aligned}
\pi(\sigma_V^2 | \alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \mathbf{V}, D_{obs}) &\propto \prod_{k=1}^M \left\{ \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \right\} \cdot (\sigma_V^2)^{-(a_V+1)} e^{-\frac{b_V}{\sigma_V^2}} \\
&= \frac{1}{\sqrt{2\pi\sigma_V^2}} \cdot (\sigma_V^2)^{-(a_V+1)} \cdot e^{-\frac{1}{\sigma_V^2} b_V} \cdot \prod_{k=1}^M e^{-\frac{1}{\sigma_V^2} \frac{V_k^2}{2}} \\
&\propto (\sigma_V^2)^{-1/2} (\sigma_V^2)^{-(a_V+1)} e^{-b_V (\sigma_V^2)^{-1}} e^{-(\sigma_V^2)^{-1} \sum_{k=1}^M \frac{V_k^2}{2}} \\
&= (\sigma_V^2)^{-(1/2+a_V+1)} e^{(\sigma_V^2)^{-1} \left[-b_V - \sum_{k=1}^M \frac{V_k^2}{2} \right]} \\
&= (\sigma_V^2)^{-(a_V+\frac{1}{2})} e^{-\frac{1}{\sigma_V^2} \left[b_V + \sum_{k=1}^M \frac{V_k^2}{2} \right]},
\end{aligned}$$

o que caracteriza a distribuição de $\sigma_V^2 | \alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \mathbf{V}$ como sendo uma Gama Inversa com parâmetros *Gama Inversa* $\left(a_V + \frac{1}{2}, b_V + \sum_{k=1}^M \frac{V_k^2}{2} \right)$.

Mas, nem todas as distribuições condicionais terão forma fechada. Nestes casos, uma alternativa é usar algum algoritmo de aceitação/rejeição ou "Metropolis-Hastings" (Chib & Greenberg, 1995).

Seja $d_{jk} = (e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k})^{\nu_{jk}} e^{-e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k}}$. A distribuição a posteriori condicional de $\alpha | \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \sigma_V^2, \mathbf{V}$, é dada por

$$\begin{aligned}
\pi\left(\alpha|\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \sigma_V^2, \mathbf{V}, D_{obs}\right) &\propto \prod_{k=1}^M \prod_{j=1}^{n_k} \left[d_{jk} \cdot e^{(e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k})(e^{-y_{jk}^\alpha} e^{\lambda_{jk}})} \alpha^{\nu_{jk}} y_{jk}^{(\alpha-1)\nu_{jk}} e^{\nu_{jk}(\lambda_{jk} - y_{jk}^\alpha e^{\lambda_{jk}})} \right] \\
&\cdot \alpha^{\delta_0-1} e^{-\tau_0 \alpha} \\
&= \prod_{k=1}^M \prod_{j=1}^{n_k} \left[d_{jk} \alpha^{\nu_{jk}} y_{jk}^{\alpha \nu_{jk}} y_{jk}^{-\nu_{jk}} e^{e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha} e^{\lambda_{jk}} + \nu_{jk} \lambda_{jk} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}}} \right] \\
&\cdot \alpha^{\delta_0-1} e^{-\tau_0 \alpha} \\
&\propto \alpha^{\delta_0-1} \cdot e^{-\tau_0 \alpha} \cdot \alpha^{\sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk}} \cdot y_{jk}^{\alpha \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk}} \\
&\cdot \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha} e^{\lambda_{jk}} + \nu_{jk} \lambda_{jk} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}} \right\} \\
&= \alpha^{\delta_0-1 + \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk}} \cdot \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk} \lambda_{jk} \right\} \\
&\cdot \exp \left\{ -\tau_0 \alpha + \alpha \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk} \cdot \ln y_{jk} \right. \\
&\left. + \sum_{k=1}^M \sum_{j=1}^{n_k} \left[e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha} e^{\lambda_{jk}} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}} \right] \right\} \\
&\propto \alpha^{(\delta_0 + \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk}) - 1} \cdot \exp \left\{ -\alpha \left(\tau_0 - \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk} \ln y_{jk} \right) \right\} \\
&\cdot \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} \left[e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha} e^{\lambda_{jk}} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}} \right] \right\} \\
&= \text{Gama} \left(\delta_0 + \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk} ; \tau_0 - \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk} \ln y_{jk} \right) \\
&\cdot \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} \left[e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha} e^{\lambda_{jk}} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}} \right] \right\},
\end{aligned}$$

desde que $\tau_0 - \sum_{k=1}^M \sum_{j=1}^{n_k} (\nu_{jk} \ln y_{jk}) > 0$.

Para a distribuição a posteriori condicional de $\mathbf{V}|\alpha, \boldsymbol{\beta}, \boldsymbol{\gamma}, \sigma_V^2$ tem-se:

$$\begin{aligned}
\pi(\mathbf{V}|\alpha, \beta, \gamma, \sigma_V^2, D_{obs}) &\propto \prod_{k=1}^M \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \cdot \prod_{k=1}^M \prod_{j=1}^{n_k} \left[e^{V_k \nu_{jk}} \cdot e^{-e^{\mathbf{w}_{jk}^\top \gamma + V_k}} \cdot e^{e^{\mathbf{w}_{jk}^\top \gamma + V_k - y_{jk}^\alpha} e^{\lambda_{jk}}} \right] \\
&\propto \prod_{k=1}^M \left\{ \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \cdot \prod_{j=1}^{n_k} \left[e^{V_k \nu_{jk}} \cdot e^{-e^{\mathbf{w}_{jk}^\top \gamma + V_k}} \cdot e^{e^{\mathbf{w}_{jk}^\top \gamma + V_k - y_{jk}^\alpha} e^{\lambda_{jk}}} \right] \right\} \\
&= \prod_{k=1}^M \pi(V_k|\alpha, \beta, \gamma, \sigma_V^2)
\end{aligned}$$

Note que a última igualdade é verdadeira visto que $\mathbf{V} = [V_1, V_2, \dots, V_M]$ e para $k = 1, 2, \dots, M$ tem-se V_k iid. Assim, para cada $k = 1, 2, \dots, M$

$$\begin{aligned}
\pi(V_k|\alpha, \beta, \gamma, \sigma_V^2, D_{obs}) &\propto \frac{e^{-\frac{V_k^2}{2\sigma_V^2}}}{\sqrt{2\pi\sigma_V^2}} \cdot \prod_{j=1}^{n_k} \left[e^{V_k \nu_{jk}} \cdot e^{-e^{\mathbf{w}_{jk}^\top \gamma + V_k}} \cdot e^{e^{\mathbf{w}_{jk}^\top \gamma + V_k - y_{jk}^\alpha} e^{\lambda_{jk}}} \right] \\
&= \exp \left\{ \frac{-V_k^2}{2\sigma_V^2} + \sum_{j=1}^{n_k} \left(V_k \nu_{jk} - e^{\mathbf{w}_{jk}^\top \gamma + V_k} + e^{\mathbf{w}_{jk}^\top \gamma + V_k} \cdot e^{-y_{jk}^\alpha} e^{\lambda_{jk}} \right) \right\} \\
&= \exp \left\{ \frac{-V_k^2}{2\sigma_V^2} \right\} \cdot \exp \left\{ \sum_{j=1}^{n_k} \left[V_k \nu_{jk} - e^{\mathbf{w}_{jk}^\top \gamma + V_k} \cdot \left(1 - e^{-y_{jk}^\alpha} e^{\lambda_{jk}} \right) \right] \right\} \\
&= \text{Normal}(0, \sigma_V^2) \cdot \exp \left\{ \sum_{j=1}^{n_k} \left[V_k \nu_{jk} - e^{\mathbf{w}_{jk}^\top \gamma + V_k} \cdot \left(1 - e^{-y_{jk}^\alpha} e^{\lambda_{jk}} \right) \right] \right\}
\end{aligned}$$

Para a distribuição a posteriori condicional de $\gamma^\top | \alpha, \beta^\top, \sigma_V^2, \mathbf{V}$ tem-se:

$$\begin{aligned}
\pi(\gamma^\top | \alpha, \beta^\top, \sigma_V^2, \mathbf{V}, D_{obs}) &\propto \prod_{k=1}^M \prod_{j=1}^{n_k} \left[e^{\nu_{jk}(\mathbf{w}_{jk}^\top \gamma + V_k)} e^{-e^{\mathbf{w}_{jk}^\top \gamma + V_k}} e^{(e^{\mathbf{w}_{jk}^\top \gamma + V_k - y_{jk}^\alpha} e^{\lambda_{jk}})} \right] = \\
&= \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} \nu_{jk}(\mathbf{w}_{jk}^\top \gamma + V_k) - e^{\mathbf{w}_{jk}^\top \gamma + V_k} + e^{\mathbf{w}_{jk}^\top \gamma + V_k - y_{jk}^\alpha} e^{\lambda_{jk}} \right\} = \\
&= \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} \left((\mathbf{w}_{jk}^\top \gamma + V_k) \nu_{jk} - e^{(\mathbf{w}_{jk}^\top \gamma + V_k)} (1 - e^{-y_{jk}^\alpha} e^{\lambda_{jk}}) \right) \right\}
\end{aligned}$$

Lembrando que $\lambda_{jk} = \mathbf{x}_{jk}^\top \boldsymbol{\beta}$, para a distribuição a posteriori condicional de $\boldsymbol{\beta}^\top | \alpha, \boldsymbol{\gamma}^\top, \sigma_V^2, \mathbf{V}$ tem-se:

$$\begin{aligned} \pi \left(\boldsymbol{\beta}^\top | \alpha, \boldsymbol{\gamma}^\top, \sigma_V^2, \mathbf{V}, D_{obs} \right) &\propto \prod_{k=1}^M \prod_{j=1}^{n_k} \exp \left\{ e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha e^{\lambda_{jk}}} \right\} e^{\nu_{jk} (\lambda_{jk} - y_{jk}^\alpha e^{\lambda_{jk}})} = \\ &= \prod_{k=1}^M \prod_{j=1}^{n_k} \exp \left\{ e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha e^{\lambda_{jk}}} + \nu_{jk} \lambda_{jk} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}} \right\} \\ &= \exp \left\{ \sum_{k=1}^M \sum_{j=1}^{n_k} \left(e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k - y_{jk}^\alpha e^{\mathbf{x}_{jk}^\top \boldsymbol{\beta}}} + \nu_{jk} \mathbf{x}_{jk}^\top \boldsymbol{\beta} - \nu_{jk} y_{jk}^\alpha e^{\mathbf{x}_{jk}^\top \boldsymbol{\beta}} \right) \right\} \end{aligned}$$

Para se avaliar se a convergência do modelo está adequada, Gelman e Rubin (1992) e Cowless e Carlin (1996) sugerem, dentre outras, a utilização do valor $\sqrt{\hat{R}}$. Para estes cálculos somente são considerados os últimos 50% valores da iteração, de modo a se obter uma melhor convergência.

Considere m o número de cadeias utilizadas com valores iniciais distintos, em que cada uma apresenta $2n$ valores, mas para o diagnóstico de convergência, serão utilizados os n últimos valores. Sejam B/n a variância entre as médias $(\bar{u}_i)_{i=1,2,\dots,m}$ das m cadeias, cada uma baseada em n valores de u e W a média das variâncias $(s_i^2)_{i=1,2,\dots,m}$ entre as cadeias, cada uma baseada em $n - 1$ graus de liberdade. Ou seja,

$$\begin{aligned} \frac{B}{n} &= \sum_{i=1}^m \frac{(\bar{u}_i - \bar{u}_{..})^2}{m-1} \\ W &= \sum_{i=1}^m \frac{s_i^2}{m} \end{aligned}$$

Considere ainda

$$\begin{aligned} \hat{\sigma}^2 &= \frac{(n-1)W}{n} + \frac{B}{n} \\ \hat{v} &= \hat{\sigma} + \frac{B}{mn} \\ \hat{\text{Var}}(\hat{v}) &= \left(\frac{n-1}{n} \right)^2 \frac{\hat{s}_i^2}{m} + \left(\frac{m+1}{mn} \right)^2 \frac{2B^2}{m-1} + \\ &\quad + \frac{2(m+1)(n-1)}{m^2 n} \left(\hat{\text{Cov}}(s_i^2, \bar{u}_i^2) - 2\bar{u}_{..} \hat{\text{Cov}}(s_i^2, \bar{u}_i) \right) \\ df &= \frac{2\hat{v}^2}{\hat{\text{Var}}(\hat{v})} \end{aligned}$$

Assim, define-se

$$\sqrt{\hat{R}} = \sqrt{\frac{n-1}{n} + \frac{m+1}{m} \frac{B/n}{W} \frac{df}{df-2}}.$$

Valores de $\sqrt{\hat{R}}$ próximos a 1 para todas as estimativas de interesse indicam que houve convergência da cadeia e que foi atingida a distribuição de cada estimador. Para maiores detalhes, ver Gelman e Rubin (1992).

2.3 Análise clássica

A equação (2.2) não é exatamente uma verossimilhança devido à presença de quantidades não observadas, dadas pelos efeitos aleatórios \mathbf{V} . Um modo de se encontrar estimadores para os parâmetros seria a utilização do algoritmo EM (McLachan e Krishnan, 1997). Mas, as integrais envolvidas não são tratáveis de um modo analítico direto, sendo necessário o uso de alguma técnica, por exemplo numérica, de aproximação de integrais.

Outro método possível de estimação se baseia em primeiro calcular o *melhor preditor linear não-viesado* (*best linear unbiased predictor* - BLUP) de Henderson (1975) e estendê-lo para se obter estimadores de máxima verossimilhança residual (residual maximum likelihood estimators - REML) aproximados (Thompson, 1980). Informações complementares podem ser encontradas em McGilchrist (1994) e em McGilchrist e Yau (1995).

Para se encontrar o BLUP, considere:

- $\phi_E = (\alpha, \beta^\top, \gamma^\top, \sigma_V^2, \mathbf{V}^\top)^\top$;
- l o logaritmo neperiano de (2.2), ou seja, $l = \ln(\sum_{\mathbf{N}} L(\phi|\mathbf{y}, \nu, \mathbf{V}, \mathbf{X}))$;
- \mathbf{U} o vetor escore de l em ϕ_E , ou seja, o vetor das primeiras derivadas de l em relação à ϕ_E ;
- H a matriz hessiana de l em ϕ_E , ou seja, a matriz das segundas derivadas de l em relação à ϕ_E .

O BLUP para \mathbf{V} e estimadores para ϕ serão encontrados a partir das raízes do vetor escore \mathbf{U} .

Para l tem-se:

$$\begin{aligned}
l &= \ln \left(\sum_{\mathbf{N}} L(\phi | \mathbf{y}, \boldsymbol{\nu}, \mathbf{V}, \mathbf{X}) \right) \\
&= \sum_{k=1}^M \left(\frac{-1}{2} \ln(2\pi\sigma_V^2) - \frac{V_k^2}{2\pi\sigma_V^2} \right) + \\
&\quad + \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ (\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \nu_{jk} - \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) + \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) + \right. \\
&\quad \left. + \nu_{jk} \cdot \left(\ln(\alpha) + (\alpha - 1) \ln(y_{jk}) + \lambda_{jk} - y_{jk}^\alpha e^{\lambda_{jk}} \right) \right\}
\end{aligned}$$

A derivada de l em α é dada por

$$\begin{aligned}
\frac{\partial l}{\partial \alpha} &= \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \cdot \exp(\lambda_{jk}) \cdot (-\alpha \cdot y_{jk}^{\alpha-1}) + \right. \\
&\quad \left. + \nu_{jk} \cdot \left[\frac{1}{\alpha} + \ln(y_{jk}) - e^{\lambda_{jk}} \cdot (-\alpha \cdot y_{jk}^{\alpha-1}) \right] \right\}
\end{aligned}$$

A derivada de l em $\boldsymbol{\beta}$ é dada por

$$\frac{\partial l}{\partial \boldsymbol{\beta}} = \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \cdot (-y_{jk}^\alpha e^{\lambda_{jk}} x_{jk}) + \nu_{jk} x_{jk} - \nu_{jk} y_{jk}^\alpha e^{\lambda_{jk}} x_{jk} \right\}$$

A derivada de l em $\boldsymbol{\gamma}$ é dada por

$$\begin{aligned}
\frac{\partial l}{\partial \boldsymbol{\gamma}} &= \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ \mathbf{w}_{jk} \cdot \nu_{jk} - \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \mathbf{w}_{jk} + \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \mathbf{w}_{jk} \cdot \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right\} \\
&= \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ \mathbf{w}_{jk} \cdot \nu_{jk} - \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \mathbf{w}_{jk} \cdot \left(1 - \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right) \right\} \\
&= \sum_{k=1}^M \sum_{j=1}^{n_k} \mathbf{w}_{jk} \cdot \nu_{jk} - \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \mathbf{w}_{jk} \cdot \left(1 - \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right) \right\}
\end{aligned}$$

A derivada de l em σ_V^2 é dada por

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_V^2} &= \sum_{k=1}^M \left\{ \frac{-1}{2} \frac{1}{\sigma_V^2} - \frac{V_k^2}{2} \left(\frac{-1}{(\sigma_V^2)^2} \right) \right\} = \sum_{k=1}^M \left\{ \frac{-1}{2\sigma_V^2} + \frac{V_k^2}{2(\sigma_V^2)^2} \right\} = \frac{1}{2\sigma_V^2} \sum_{k=1}^M \left\{ -1 + \frac{V_k^2}{\sigma_V^2} \right\} \\
&= \frac{1}{2\sigma_V^2} \left\{ \sum_{k=1}^M (-1) + \frac{1}{\sigma_V^2} \sum_{k=1}^M V_k^2 \right\} = \frac{1}{2\sigma_V^2} \left\{ -M + \frac{1}{\sigma_V^2} \sum_{k=1}^M V_k^2 \right\} = \frac{-M\sigma_V^2 + \sum_{k=1}^M V_k^2}{2\sigma_V^2}
\end{aligned}$$

Finalmente, a derivada de l em \mathbf{V} é o vetor das derivadas de l em V_k , em que cada componente é dada por

$$\begin{aligned}
\frac{\partial l}{\partial V_k} &= \frac{-1}{2\sigma_V^2} \cdot 2V_k + \sum_{j=1}^{n_k} \left\{ \nu_{jk} - \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot 1 + \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right\} \\
&= \frac{-V_k}{\sigma_V^2} + \sum_{j=1}^{n_k} \nu_{jk} - \sum_{j=1}^{n_k} \left\{ \exp(\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k) \cdot \left(1 - \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right) \right\} \\
&= \frac{-V_k}{\sigma_V^2} + \sum_{j=1}^{n_k} \nu_{jk} - \sum_{j=1}^{n_k} \left\{ e^{V_k} \cdot e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma}} \cdot \left(1 - \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right) \right\} \\
&= \frac{-V_k}{\sigma_V^2} + \sum_{j=1}^{n_k} \nu_{jk} - e^{V_k} \cdot \sum_{j=1}^{n_k} \left\{ e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma}} \cdot \left(1 - \exp(-y_{jk}^\alpha e^{\lambda_{jk}}) \right) \right\}
\end{aligned}$$

O vetor escore \mathbf{U} é dado por $\mathbf{U} = \left(\frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \boldsymbol{\beta}^\top}, \frac{\partial l}{\partial \boldsymbol{\gamma}^\top}, \frac{\partial l}{\partial \sigma_V^2}, \frac{\partial l}{\partial \mathbf{V}^\top} \right)^\top$.

Para se encontrar o BLUP, deve-se encontrar a solução de $\mathbf{U} = \vec{0}$. Mas, isto não é de fácil solução. Entretanto, note que

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_V^2} &= 0 \iff \frac{-M\sigma_V^2 + \sum_{k=1}^M V_k^2}{2\sigma_V^2} = 0 \iff -M\sigma_V^2 + \sum_{k=1}^M V_k^2 = 0 \iff \\
\sigma_V^2 &= \frac{\sum_{k=1}^M V_k^2}{M}
\end{aligned} \tag{2.4}$$

Como a solução das demais equações não são analiticamente tratáveis, uma alternativa é utilizar o método de Newton-Raphson, que diz que, se $\boldsymbol{\phi}_E(t+1)$ são as estimativas em $\boldsymbol{\phi}_E$ no passo $t+1$, se $\boldsymbol{\phi}_E(t)$ são as estimativas em $\boldsymbol{\phi}_E$ no passo t , se $\mathbf{U}(t)$ é o vetor escore calculado em $\boldsymbol{\phi}_E(t)$ e se $H(t)$ é a matriz hessiana calculada em $\boldsymbol{\phi}_E(t)$, então

$$\boldsymbol{\phi}_E(t+1) = \boldsymbol{\phi}_E(t) - H(t)^{-1} \cdot \mathbf{U}(t) \tag{2.5}$$

em que (2.5) é inicializado com valores para $\boldsymbol{\phi}_E$ e deve ser calculado até que $\|\boldsymbol{\phi}_E(t+1) - \boldsymbol{\phi}_E(t)\| < \varepsilon$, com ε suficientemente pequeno. Pode-se, por exemplo, usar $\varepsilon = 10^{-3}$.

Se $\tilde{\mathbf{V}}$ são os preditores para \mathbf{V} , e $\tilde{\boldsymbol{\phi}}_E$ são as estimativas para $\boldsymbol{\phi}_E$, obtidos por BLUP e se a matriz hessiana-inversa calculada em $\tilde{\boldsymbol{\phi}}_E$ é decomposta relativamente a $(\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top) | \sigma_V^2 | \mathbf{V}^\top$ por

$$H^{-1} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{2.6}$$

então, segundo McGilchrist e Yau (1995) e Yau e Ng (2001), as estimativas para σ_V^2 via REML serão calculadas por

$$\hat{\sigma}_V^2 = \frac{1}{M} \left(-tr(A_{33}) + \tilde{\mathbf{V}}^\top \tilde{\mathbf{V}} \right).$$

A variância de $\hat{\sigma}_V^2$ é determinada por

$$Var(\hat{\sigma}_V^2) = 2 \left(\frac{1}{\sigma_V^2} \left(M + 2 \frac{1}{\sigma_V^2} tr(A_{33}) \right) - \frac{1}{\sigma_V^4} tr(A_{33}^2) \right)^{-1}$$

As variâncias dos demais parâmetros são obtidas através da matriz informação de Fisher, ou seja, são dadas por

$$Var \left(\left[\alpha, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top \right]^\top \right) = -[A_{11}] \quad (2.7)$$

Capítulo 3

Estudos de simulação - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios

Neste capítulo são apresentados estudos de simulação feitos para o Modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, apresentado no Capítulo 2, tanto para a abordagem bayesiana quanto para a clássica.

Para este estudo foram utilizados os programas R (versão 2.5.1) (Venables et al., 2008), Winbugs (versão 1.4) (Spiegelhalter, 2007) e Maple (versão 10) (Maplesoft, 2005).

3.1 Características

Para este estudo, em uma primeira situação, são assumidas as mesmas condições básicas apresentadas em Yau-Ng (2001):

- $M = 10$, ou seja, 10 clínicas;
- $n_k = 15$ para todo $k = 1, 2, \dots, M$, ou seja, cada clínica possui 15 indivíduos;
- em cada clínica, 5 pacientes são aleatoriamente designados para o grupo de tratamento ($x_{jk} = 1$) e os 10 restantes, para o grupo controle ($x_{jk} = 0$);
- 12 grupos de parâmetros a serem analisados:
 1. $(\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5)$,
 2. $(\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5)$,
 3. $(\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5)$,
 4. $(\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5)$,
 5. $(\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5)$,
 6. $(\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5)$,
 7. $(\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5)$,

8. $(\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5)$,
 9. $(\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1)$,
 10. $(\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1)$,
 11. $(\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1)$,
 12. $(\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1)$;
- 50 simulações para cada grupo de parâmetros.

Em um segundo momento foi estudado o caso em $n_k = 50$ para todo $k = 1, 2, \dots, 15$ com as demais condições mantidas. Ou seja, a alteração feita nos permite analisar melhor resultados assintóticos, já que houve um aumento no número de indivíduos de cada clínica.

O que se espera é que:

- o desempenho das estimativas melhora, quando α aumenta de 1 para 1,3, porque quanto menor o valor de α , maior será a média da distribuição de Z_{cjk} , Weibull(α, λ).
- uma diminuição na censura (de 500 para 300) piora o desempenho das estimativas, já que esta diminuição poderá fazer com que se tenha um aumento de indivíduos não curados que formam censurados;
- um aumento em σ_V^2 (de 0,5 para 1) produza uma maior variabilidade entre as clínicas, acarretando estimativas não tão boas quanto as obtidas para valores menores de σ_V^2 ;
- o aumento no número de indivíduos por clínica (de 15 para 50) melhora a precisão dos resultados assintóticos.

3.2 Simulação - abordagem bayesiana

Em um primeiro momento foi feito um estudo de simulação com a utilização do programa Winbugs, versão 1.4. Os comandos utilizados encontram-se no apêndice A.2. Os dados utilizados foram gerados com a utilização do programa R e encontram-se no apêndice A.1.

Os dados foram compilados sempre a partir de três cadeias iniciadas em valores distintos, com 30 mil iterações cada.

3.2.1 Caso $n_k = 15$

Os resultados obtidos estão na Tabela 3.1, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a média dos desvios-padrão encontrados em cada uma das 50 simulações.

Tabela 3.1: Resultado de 50 simulações via Metropolis-Hastings para $n_k = 15$

Grupo 1: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0013	0,0729	0,0758
β_1	-0,5	-0,4761	0,2584	0,2280
γ_0	0	0,0298	0,2295	0,1913
γ_1	0,5	0,5203	0,2438	0,2326
σ_V^2	0,5	0,4018	0,1689	0,1878
Grupo 2: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,2782	0,0902	0,0988
β_1	-0,5	-0,6180	0,2329	0,2309
γ_0	0	-0,0323	0,2763	0,2019
γ_1	0,5	0,4625	0,2440	0,2350
σ_V^2	0,5	0,4764	0,2361	0,2234
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0,9914	0,0803	0,0755
β_1	-0,5	-0,4312	0,2756	0,2434
γ_0	0	0,0974	0,3264	0,2350
γ_1	0,5	0,4153	0,2657	0,2452
σ_V^2	1	0,7686	0,3940	0,3497
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0,9918	0,0712	0,0756
β_1	-0,5	-0,5121	0,2570	0,2255
γ_0	0	0,0049	0,1872	0,1912
γ_1	0,5	0,4752	0,2143	0,2310
σ_V^2	0,5	0,3985	0,1367	0,1878
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0124	0,0828	0,0878
β_1	-0,5	-0,4360	0,2357	0,2249
γ_0	-0,5	-0,5155	0,2291	0,2007
γ_1	0,5	0,4561	0,2149	0,2475
σ_V^2	0,5	0,3446	0,1114	0,1720
Grupo 6: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3221	0,1337	0,1140
β_1	-0,5	-0,5985	0,2450	0,2278
γ_0	-0,5	-0,4857	0,2461	0,1997
γ_1	0,5	0,4587	0,2358	0,2466
σ_V^2	0,5	0,3500	0,1774	0,1710

Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0123	0,0960	0,0860
β_1	-0,5	-0,4852	0,2662	0,2378
γ_0	-0,5	-0,3757	0,3325	0,2332
γ_1	0,5	0,4674	0,2294	0,2522
σ_V^2	1	0,6556	0,3618	0,3116
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0148	0,0937	0,0876
β_1	-0,5	-0,4553	0,2364	0,2276
γ_0	-0,5	-0,4264	0,2709	0,2060
γ_1	0,5	0,4434	0,2409	0,2461
σ_V^2	0,5	0,4148	0,2717	0,2006
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0078	0,0909	0,0958
β_1	-1	-0,9968	0,2373	0,2433
γ_0	-1	-0,9794	0,2831	0,2279
γ_1	1	0,9486	0,2905	0,2679
σ_V^2	0,5	0,3824	0,2114	0,1956
Grupo 10: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,2952	0,1154	0,1193
β_1	-1	-1,2092	0,2762	0,2479
γ_0	-1	-0,8451	0,3201	-0,2165
γ_1	1	0,9331	0,2662	0,2609
σ_V^2	0,5	0,3413	0,1461	0,1728
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0149	0,1164	0,0951
β_1	-1	-1,0114	0,2464	0,2509
γ_0	-1	-0,9043	0,3185	0,2587
γ_1	1	0,9532	0,2435	0,2712
σ_V^2	1	0,7230	0,5260	0,3477
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0280	0,0899	0,0959
β_1	-1	-0,9455	0,2482	0,2383
γ_0	-1	-0,8869	0,3019	0,2207
γ_1	1	0,9310	0,3103	0,2617
σ_V^2	0,5	0,3598	0,1823	0,1827

Note que a diferença entre os valores médios estimados e os valores verdadeiros são quase sempre menores do que um desvio-padrão, o que indica que os resultados obtidos via simulação estão adequados.

As Figuras de 3.1 a 3.3 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

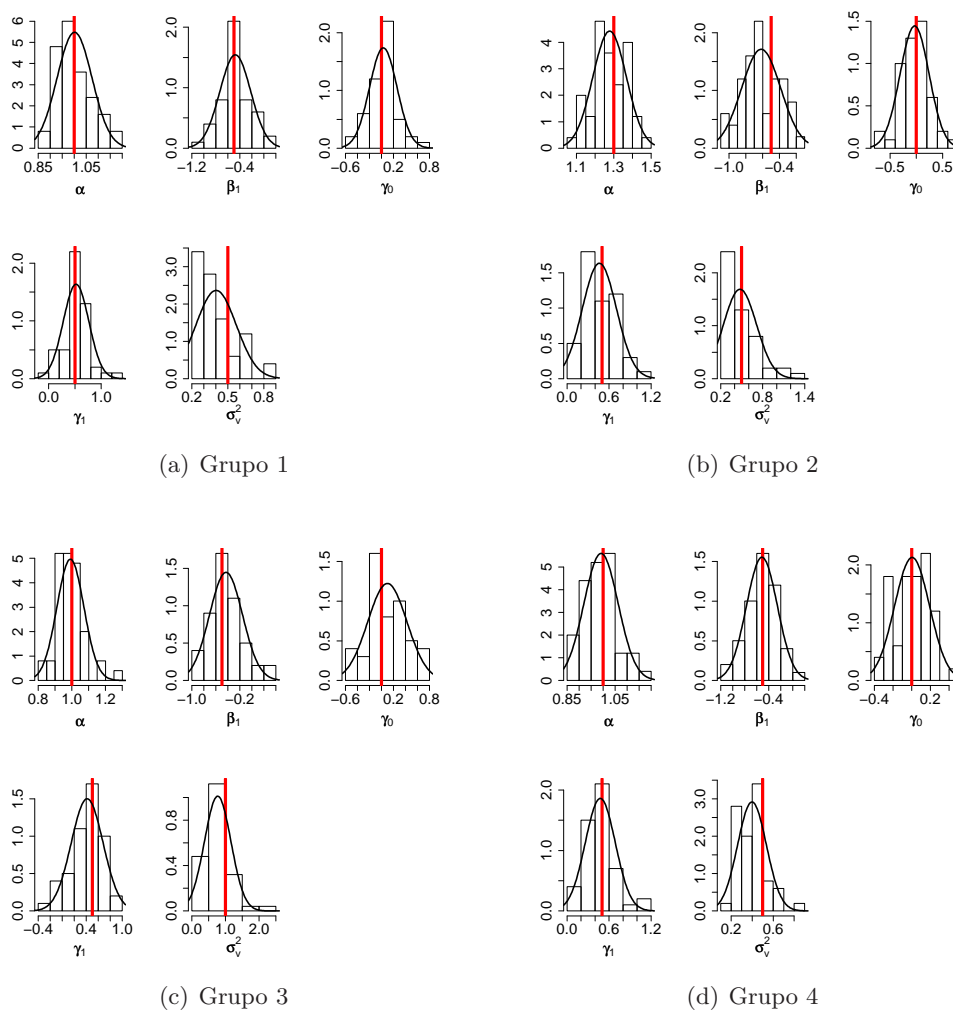


Figura 3.1: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – parte 1

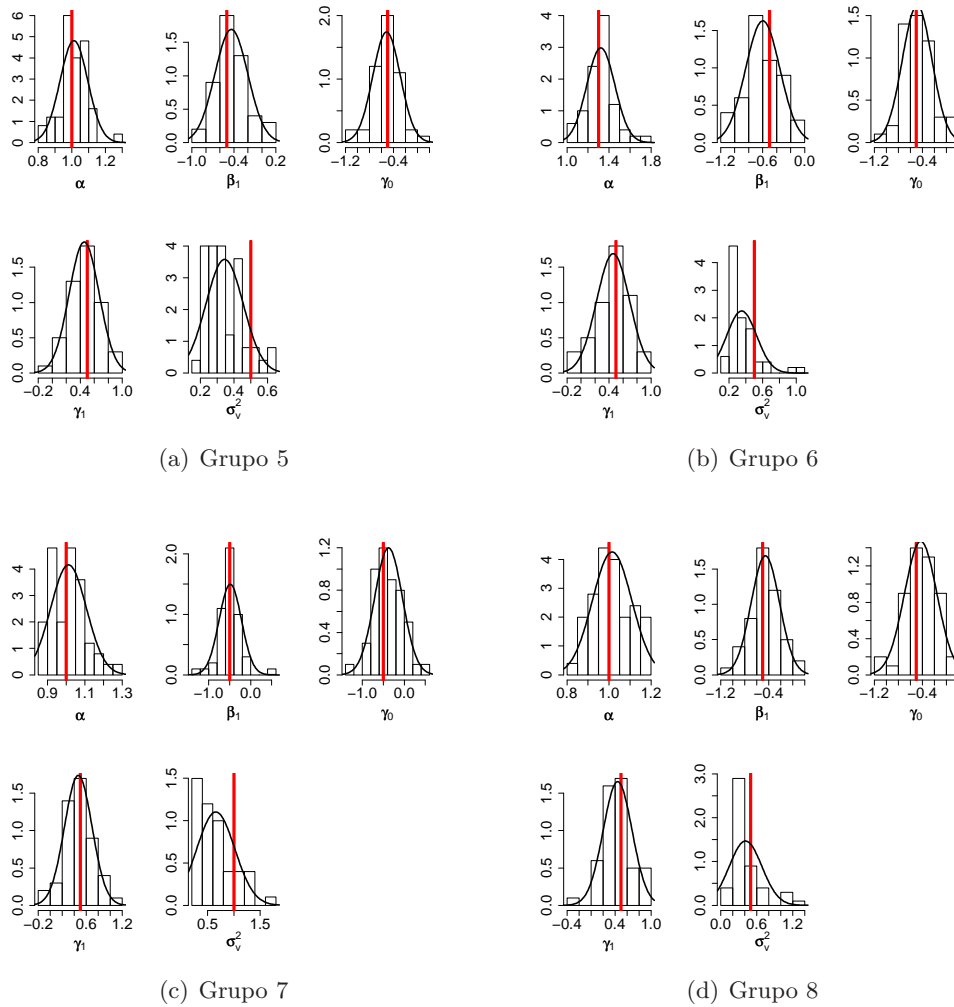


Figura 3.2: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – parte 2

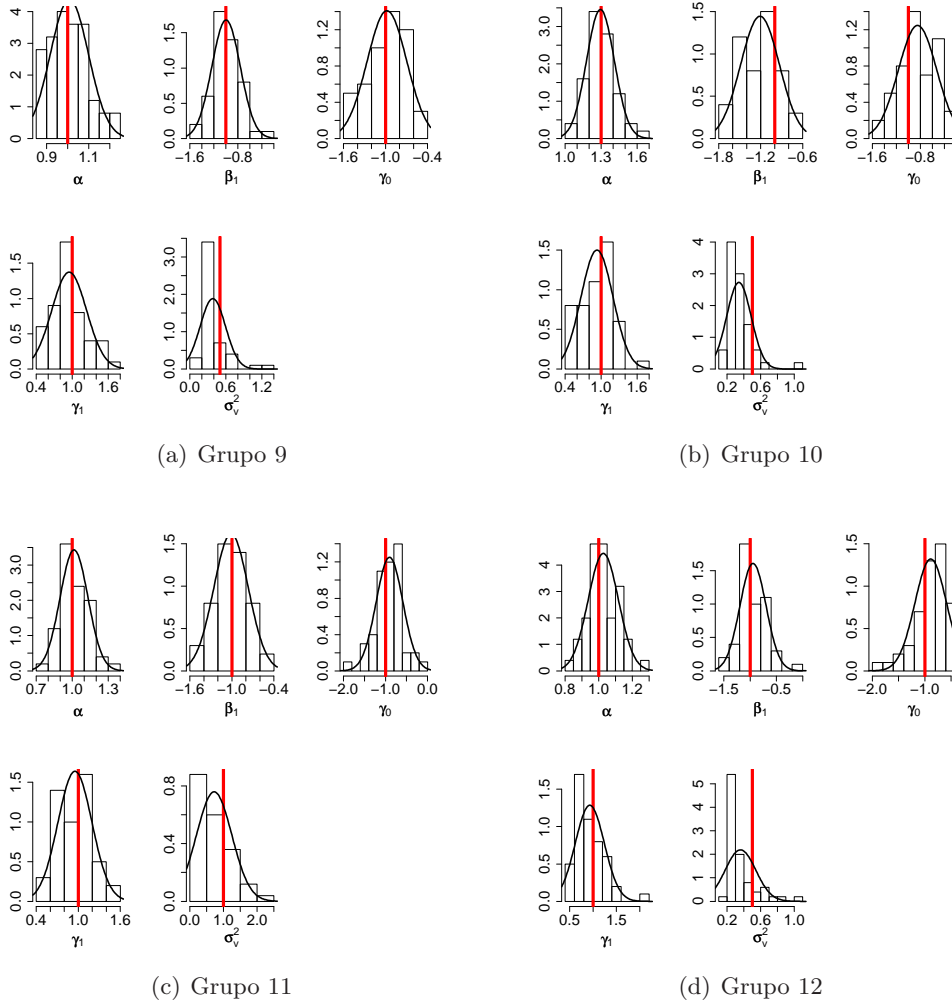


Figura 3.3: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, que estão representados pelas retas verticais. No caso de σ_v^2 , há um indício de que os valores estão sendo subestimados.

Para se avaliar se a convergência do modelo está adequada, calcula-se os valores de $\sqrt{\hat{R}}$, como definido na seção 2.2. Para isso, considere $m = 3$ o número de cadeias utilizadas com valores iniciais distintos, em que cada uma apresenta $2n = 30000$ valores, ou seja, $n = 15000$. Foram calculados os valores de $\sqrt{\hat{R}}$ para simulações em cada um dos 12 grupos. Os resultados obtidos encontram-se na Tabela 3.2.

Tabela 3.2: Valores de $\sqrt{\hat{R}}$ – caso $n_k = 15$

parâmetro	1	2	3	4	5	6
α	1,00156	1,00045	1,00134	1,00086	1,00142	1,00124
β_1	1,00030	1,00394	1,00410	1,00030	1,00229	1,00013
γ_0	1,00158	1,00045	1,00606	1,00292	1,00303	1,00151
γ_1	1,00051	1,00228	1,00071	1,00035	1,00345	1,00285
σ_V^2	1,00271	1,00074	1,00137	1,00277	1,00178	1,00201

parâmetro	7	8	9	10	11	12
α	1,00106	1,00026	1,00076	1,00158	1,00101	1,00174
β_1	1,00115	1,00164	1,00052	1,00017	1,00063	1,00094
γ_0	1,00176	1,00141	1,00419	1,00145	1,00148	1,00139
γ_1	1,00050	1,00104	1,00036	1,00026	1,00142	1,00070
σ_V^2	1,00196	1,00104	1,00668	1,00245	1,00034	1,00365

É perceptível que todos os valores estão bem próximos a 1 para todas as estimativas de interesse indicando que houve convergência da cadeia e que foi atingida a distribuição de cada estimador.

3.2.2 Caso $n_k = 50$

Os resultados obtidos estão na Tabela 3.3, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a média dos desvios-padrão encontrados em cada uma das 50 simulações.

Tabela 3.3: Resultado de 50 simulações via Metropolis-Hastings para $n_k = 50$

Grupo 1: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0.9980	0.0386	0.0404
β_1	-0.5	-0.5076	0.1101	0.1241
γ_0	0	0.0465	0.1894	0.1812
γ_1	0.5	0.4840	0.1052	0.1262
σ_V^2	0.5	0.5675	0.2428	0.2169
Grupo 2: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1.3	1.3011	0.0549	0.0534
β_1	-0.5	-0.6273	0.1299	0.1233
γ_0	0	-0.0123	0.2333	0.1743
γ_1	0.5	0.5099	0.1239	0.1255
σ_V^2	0.5	0.5215	0.2317	0.1991
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0.9923	0.0385	0.0410
β_1	-0.5	-0.4895	0.1489	0.1304
γ_0	0	0.0201	0.2876	0.2299
γ_1	0.5	0.4829	0.1293	0.1314
σ_V^2	1	1.0096	0.4277	0.3798
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1.0043	0.0427	0.0414
β_1	-0.5	-0.5092	0.1489	0.1212
γ_0	0	-0.0155	0.2228	0.1659
γ_1	0.5	0.4956	0.1064	0.1246
σ_V^2	0.5	0.4647	0.1786	0.1784
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0.9922	0.0444	0.0467
β_1	-0.5	-0.4811	0.1373	0.1229
γ_0	-0.5	-0.4965	0.2411	0.1806
γ_1	0.5	0.5050	0.1182	0.1335
σ_V^2	0.5	0.5314	0.2238	0.2094
Grupo 6: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1.3	1.3055	0.0692	0.0611
β_1	-0.5	-0.6222	0.1050	0.1239
γ_0	-0.5	-0.5114	0.2547	0.1783
γ_1	0.5	0.5190	0.1424	0.1341
σ_V^2	0.5	0.5084	0.2183	0.1994

Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0.9955	0.0464	0.0452
β_1	-0.5	-0.4896	0.1391	0.1287
γ_0	-0.5	-0.3946	0.3294	0.2287
γ_1	0.5	0.4964	0.1163	0.1355
σ_V^2	1	0.9379	0.4815	0.3619
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0.9990	0.0495	0.0465
β_1	-0.5	-0.5196	0.1332	0.1224
γ_0	-0.5	-0.4687	0.2476	0.1755
γ_1	0.5	0.5060	0.1156	0.1324
σ_V^2	0.5	0.4969	0.2288	0.1945
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0.9982	0.0579	0.0513
β_1	-1	-0.9847	0.1273	0.1305
γ_0	-1	-0.9924	0.2204	0.1770
γ_1	1	1.0009	0.1332	0.1435
σ_V^2	0.5	0.4385	0.1959	0.1781
Grupo 10: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1.3	1.3085	0.0880	0.0682
β_1	-1	-1.2937	0.1781	0.1388
γ_0	-1	-1.0285	0.2445	-0.1861
γ_1	1	1.0057	0.1617	0.1450
σ_V^2	0.5	0.5046	0.2381	0.2031
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1.0063	0.0493	0.0517
β_1	-1	-0.9732	0.1195	0.1356
γ_0	-1	-1.0086	0.3214	0.2327
γ_1	1	0.9694	0.1340	0.1476
σ_V^2	1	0.8984	0.4011	0.3593
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1.0063	0.0511	0.0528
β_1	-1	-0.9977	0.1368	0.1317
γ_0	-1	-1.0350	0.2439	0.1809
γ_1	1	0.9932	0.1480	0.1458
σ_V^2	0.5	0.4615	0.2147	0.1889

Note que a diferença entre as estimativas médias estimados e os valores verdadeiros são menores do que um desvio-padrão, em quase todos os casos, o que indica que os resultados obtidos via simulação estão bons.

Note ainda que as estimativas obtidas para este caso $n_k = 50$ estão mais próximos dos valores verdadeiros do que no caso $n_k = 15$, ou seja, o fato de se ter mais indivíduos em cada clínica faz com que as estimativas fiquem mais acuradas.

As Figuras de 3.4 a 3.6 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

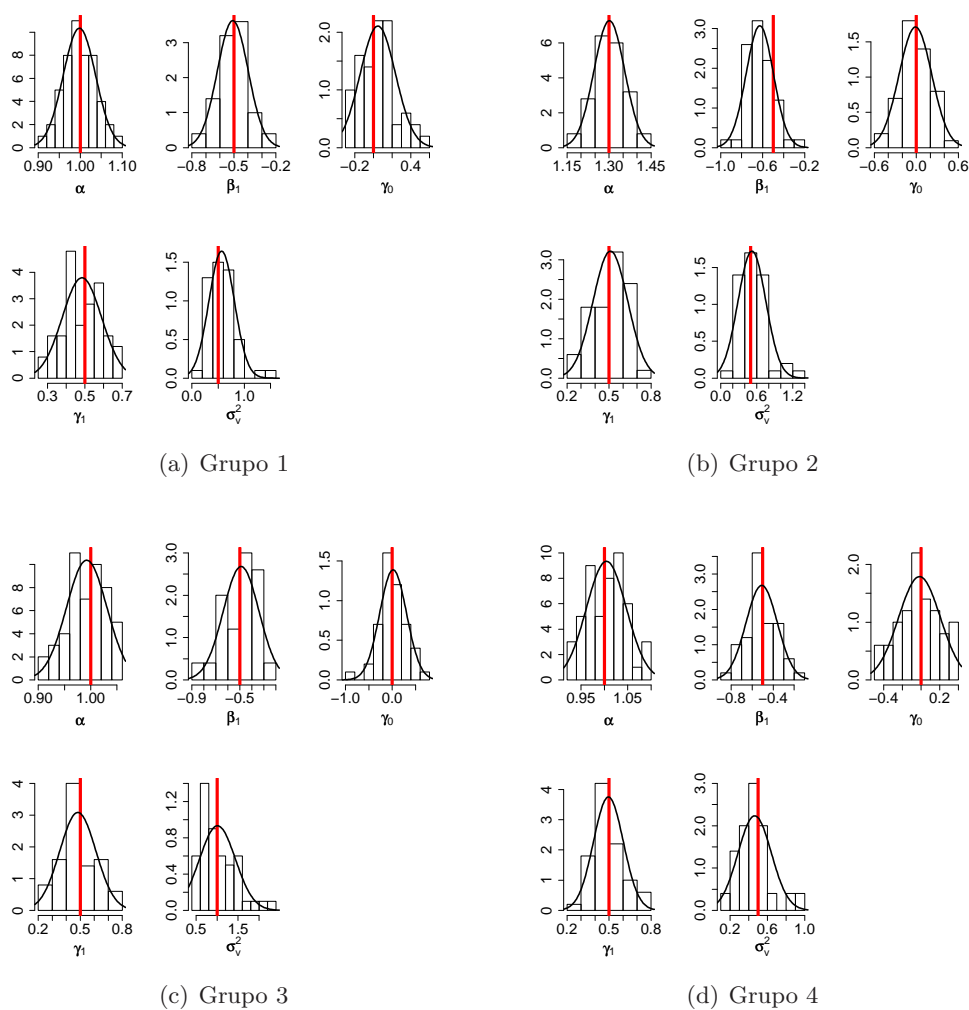


Figura 3.4: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – parte 1

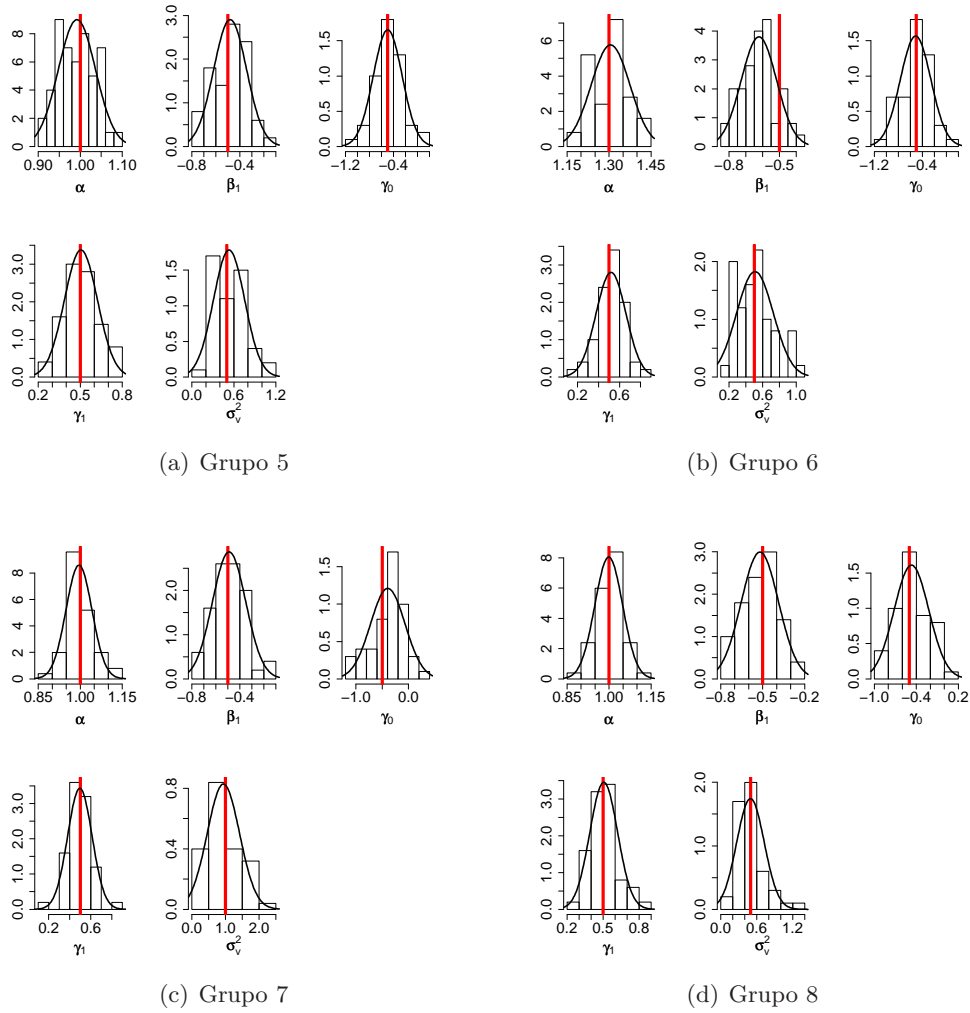


Figura 3.5: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – parte 2

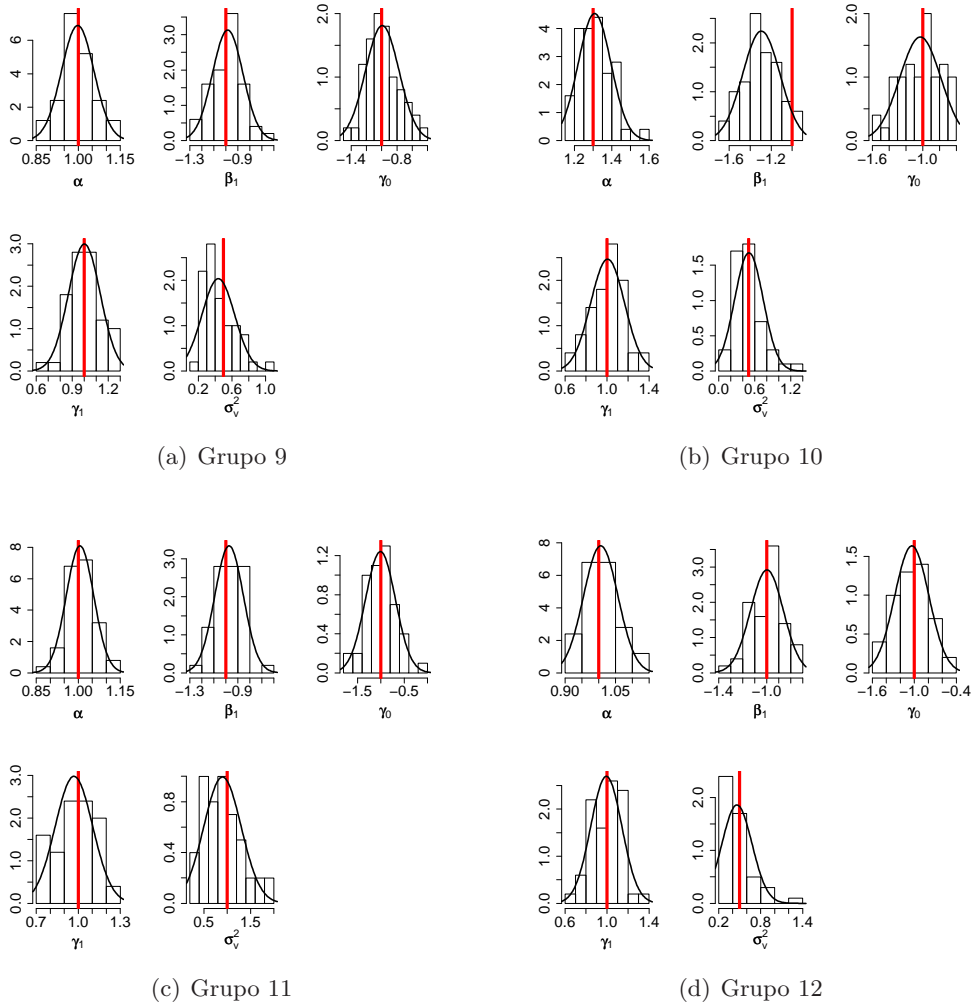


Figura 3.6: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, representados pelas retas verticais. Indicam também que as estimativas não estão muito longe da normalidade assintótica.

O histograma do grupo 10, aparentemente, não está muito bom para o parâmetro β_1 , mas a estimativa média difere do valor verdadeiro ($|-1,2937 - (-1)| = 0,2937$) por menos de dois desvios-padrão ($2 * 0,1781 = 0,3562$), o que, portanto, não é nenhum grande problema.

Para se avaliar se a convergência do modelo está adequada, calcula-se os valores de $\sqrt{\hat{R}}$, como definido na seção 2.2. Para isso, considere $m = 3$ o número de cadeias utilizadas com valores iniciais distintos, em que cada uma apresenta $2n = 30000$ valores, ou seja, $n = 15000$. Foram calculados os valores de $\sqrt{\hat{R}}$ para simulações em cada um dos 12 grupos. Os resultados obtidos encontram-se na Tabela 3.4.

Tabela 3.4: Valores de $\sqrt{\hat{R}}$ - caso $n_k = 50$

parâmetro	1	2	3	4	5	6
α	1,00033	1,00018	1,00002	1,00000	0,99998	1,00002
β_1	1,00083	1,00140	1,00040	1,00005	1,00018	1,00008
γ_0	1,00674	1,01705	1,01473	1,01240	1,00145	1,00542
γ_1	1,00152	1,00082	1,00055	1,00034	1,00042	1,00131
σ_V^2	1,00065	1,00036	1,00132	1,00050	1,00011	1,00087

parâmetro	7	8	9	10	11	12
α	1,00006	1,00026	1,00011	1,00010	1,00007	1,00012
β_1	1,00040	1,00060	1,00061	1,00029	1,00070	1,00028
γ_0	1,01302	1,00234	1,00499	1,00240	1,01053	1,00788
γ_1	1,00023	1,00200	1,00137	1,00059	1,00244	1,00007
σ_V^2	1,00069	1,00026	1,00012	1,00023	1,00090	1,00064

É perceptível que todos os valores estão bem próximos a 1 para todas as estimativas de interesse indicando que houve convergência da cadeia e que foi atingida a distribuição de cada estimador.

3.3 Simulação - abordagem clássica

Um estudo de simulação da parte clássica foi feito com o auxílio do programa Maple, versão 10. Os dados foram gerados com o programa R e são os mesmos utilizados na seção 3.2. Os comandos utilizados encontram-se no apêndice A.3.

3.3.1 Caso $n_k = 15$

Os resultados obtidos estão na Tabela 3.5, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a média dos desvios-padrão encontrados em cada uma das 50 simulações. Estes desvios-padrão foram obtidos pela raiz quadrada dos elementos da diagonal principal da matriz de informação de Fisher.

Tabela 3.5: Resultado de 50 simulações via REML para $n_k = 15$

Grupo 1: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0220	0,0736	0,0768
β_1	-0,5	-0,5302	0,2702	0,2365
γ_0	0	0,0356	0,2386	0,3122
γ_1	0,5	0,5744	0,2566	0,2398
σ_V^2	0,5	0,5914	0,1775	0,1621
Grupo 2: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3052	0,0942	0,1001
β_1	-0,5	-0,6689	0,2402	0,2378
γ_0	0	-0,0288	0,2850	0,3138
γ_1	0,5	0,5084	0,2596	0,2401
σ_V^2	0,5	0,6691	0,2185	0,1829
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0111	0,0813	0,0759
β_1	-0,5	-0,4710	0,2865	0,2493
γ_0	0	0,1150	0,3385	0,3135
γ_1	0,5	0,4577	0,2824	0,2499
σ_V^2	1	0,9118	0,2934	0,2401
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0125	0,0734	0,0765
β_1	-0,5	-0,5630	0,2640	0,2339
γ_0	0	0,0089	0,1917	0,3124
γ_1	0,5	0,5241	0,2304	0,2377
σ_V^2	0,5	0,5957	0,1541	0,1640
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0296	0,0848	0,0889
β_1	-0,5	-0,4746	0,2465	0,2299
γ_0	-0,5	-0,5379	0,2414	0,3232
γ_1	0,5	0,5049	0,2296	0,2503
σ_V^2	0,5	0,5964	0,1392	0,1746
Grupo 6: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3467	0,1369	0,1156
β_1	-0,5	-0,6374	0,2518	0,2327
γ_0	-0,5	-0,4990	0,2537	0,3223
γ_1	0,5	0,5003	0,2511	0,2497
σ_V^2	0,5	0,5763	0,1903	0,1675

Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0303	0,0987	0,0868
β_1	-0,5	-0,5235	0,2805	0,2428
γ_0	-0,5	-0,3825	0,3470	0,3221
γ_1	0,5	0,5103	0,2456	0,2552
σ_V^2	1	0,8610	0,2732	0,2362
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0315	0,0974	0,0887
β_1	-0,5	-0,4925	0,2428	0,2330
γ_0	-0,5	-0,4379	0,2807	0,3215
γ_1	0,5	0,4865	0,2521	0,2495
σ_V^2	0,5	0,6311	0,2436	0,1798
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0283	0,0942	0,0974
β_1	-1	-1,0647	0,2447	0,2504
γ_0	-1	-1,0201	0,2927	0,3391
γ_1	1	1,0323	0,3174	0,2723
σ_V^2	0,5	0,6599	0,2032	0,1976
Grupo 10: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3235	0,1197	0,1212
β_1	-1	-1,2813	0,2887	0,2548
γ_0	-1	-0,8808	0,3313	0,3340
γ_1	1	1,0142	0,2894	0,2652
σ_V^2	0,5	0,6123	0,1582	0,1830
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0329	0,1216	0,0962
β_1	-1	-1,0606	0,2655	0,2549
γ_0	-1	-0,9178	0,3302	0,3372
γ_1	1	1,0138	0,2545	0,2731
σ_V^2	1	0,9002	0,3870	0,2520
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0473	0,0927	0,0973
β_1	-1	-1,0102	0,2561	0,2456
γ_0	-1	-0,9244	0,3161	0,3357
γ_1	1	1,0093	0,3292	0,2658
σ_V^2	0,5	0,6279	0,1943	0,1871

Note que, como no caso bayesiano, a diferença entre as médias das estimativas e os valores verdadeiros são, quase sempre, menores do que um desvio-padrão, o que indica que os resultados obtidos via simulação estão bons.

Note ainda que os desvios-padrão das médias estão próximos aos desvios-médio em praticamente todos os casos. Isso é um bom indicício de que os desvios estão sendo bem estimados pelas segundas-derivadas da log-verossimilhança, ou seja, pela informação de Fisher.

As Figuras de 3.7 a 3.9 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

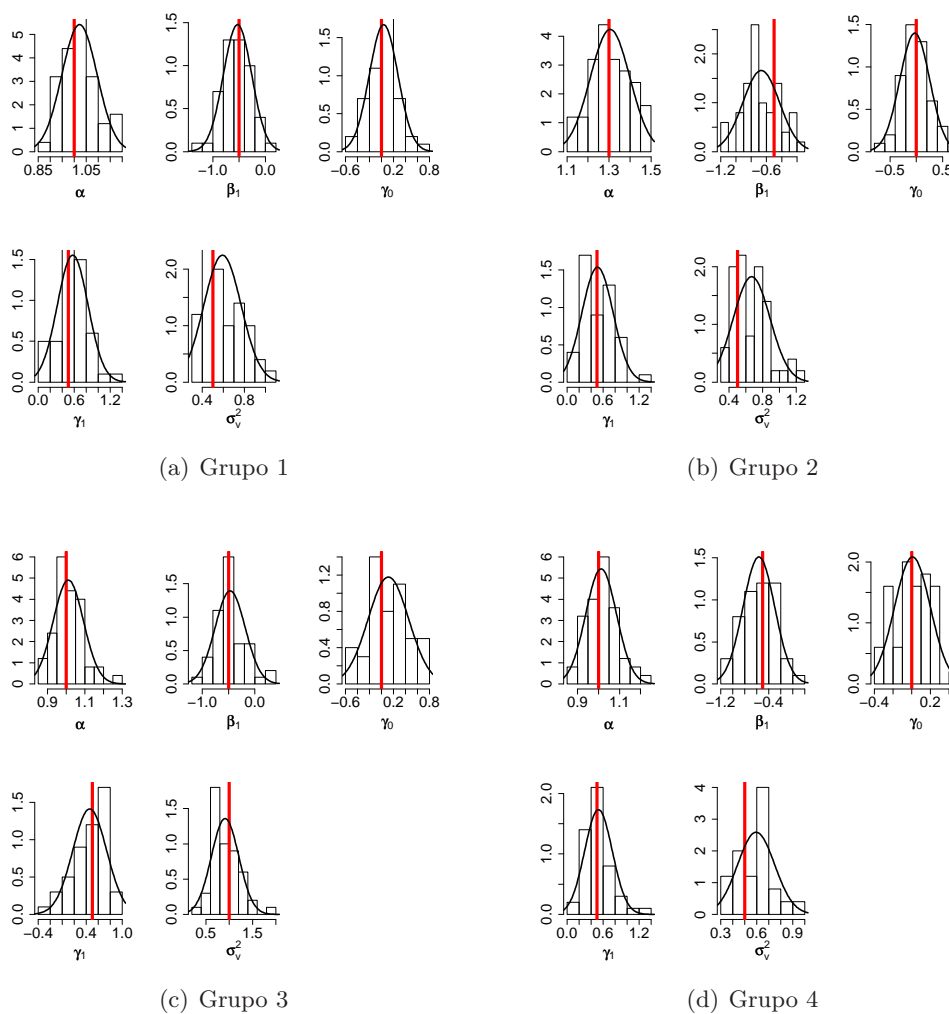


Figura 3.7: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – parte 1

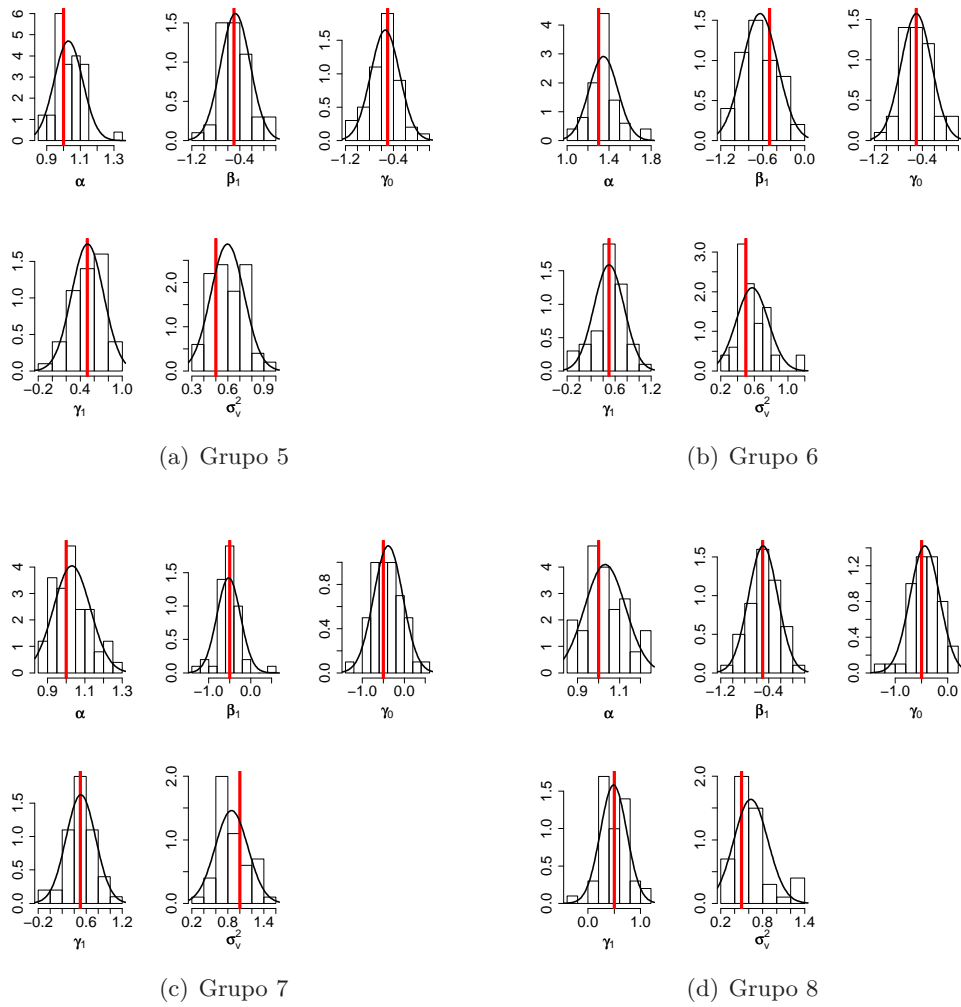


Figura 3.8: Histogramas das estimativas dos parâmetros estimados via REML - $n_k = 15$ - parte 2

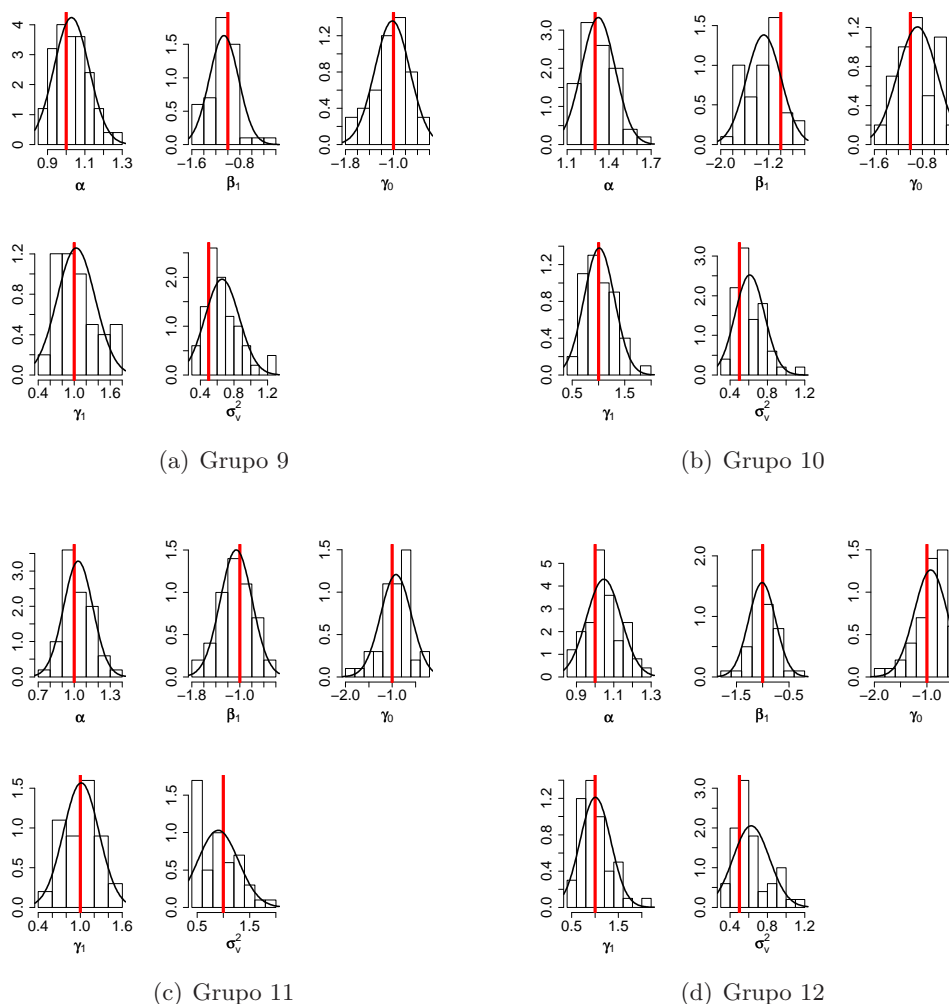


Figura 3.9: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, representados pelas retas verticais. Indicam também que as estimativas não estão muito longe da normalidade assintótica.

3.3.2 Caso $n_k = 50$

Os resultados obtidos estão na Tabela 3.6, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a média dos desvios-padrão encontrados em cada uma das 50 simulações. Estes desvios-padrão foram obtidos pela raiz quadrada dos elementos da diagonal principal da matriz de informação de Fisher.

Tabela 3.6: Resultado de 50 simulações via REML para $n_k = 50$

Grupo 1: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0027	0,0385	0,0405
β_1	-0,5	-0,5195	0,1106	0,1252
γ_0	0	0,0468	0,1924	0,2920
γ_1	0,5	0,4971	0,1059	0,1268
σ_V^2	0,5	0,5940	0,2235	0,1447
Grupo 2: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3079	0,0554	0,0536
β_1	-0,5	-0,6408	0,1311	0,1245
γ_0	0	-0,0138	0,2375	0,2923
γ_1	0,5	0,5232	0,1266	0,1264
σ_V^2	0,5	0,5499	0,2178	0,1335
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0,9963	0,0384	0,0410
β_1	-0,5	-0,4956	0,1486	0,1312
γ_0	0	0,0275	0,2867	0,2931
γ_1	0,5	0,4911	0,1301	0,1322
σ_V^2	1	0,9873	0,3615	0,2399
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0094	0,0432	0,0415
β_1	-0,5	-0,5219	0,1506	0,1223
γ_0	0	-0,0185	0,2279	0,2922
γ_1	0,5	0,5089	0,1097	0,1253
σ_V^2	0,5	0,4991	0,1759	0,1209
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0,9965	0,0446	0,0468
β_1	-0,5	-0,4912	0,1379	0,1236
γ_0	-0,5	-0,5062	0,2489	0,2957
γ_1	0,5	0,5172	0,1200	0,1341
σ_V^2	0,5	0,5851	0,2127	0,1453
Grupo 6: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3122	0,0698	0,0614
β_1	-0,5	-0,6343	0,1059	0,1249
γ_0	-0,5	-0,5215	0,2583	0,2958
γ_1	0,5	0,5331	0,1436	0,1342
σ_V^2	0,5	0,5582	0,2069	0,1384

Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	0,9992	0,0470	0,0454
β_1	-0,5	-0,4957	0,1401	0,1290
γ_0	-0,5	-0,3928	0,3272	0,2959
γ_1	0,5	0,5053	0,1181	0,1357
σ_V^2	1	0,9287	0,3993	0,2275
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0032	0,0497	0,0467
β_1	-0,5	-0,5298	0,1347	0,1232
γ_0	-0,5	-0,4773	0,2540	0,2954
γ_1	0,5	0,5178	0,1187	0,1329
σ_V^2	0,5	0,5470	0,2131	0,1353
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0030	0,0582	0,0515
β_1	-1	-1,0020	0,1279	0,1313
γ_0	-1	-1,0107	0,2245	0,3008
γ_1	1	1,0208	0,1367	0,1439
σ_V^2	0,5	0,5186	0,1883	0,1304
Grupo 10: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1,3	1,3160	0,0887	0,0685
β_1	-1	-1,3137	0,1803	0,1396
γ_0	-1	-1,0474	0,2483	0,3015
γ_1	1	1,0258	0,1656	0,1458
σ_V^2	0,5	0,5740	0,2247	0,1447
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0108	0,0497	0,0520
β_1	-1	-0,9855	0,1199	0,1363
γ_0	-1	-1,0215	0,3297	0,3022
γ_1	1	0,9861	0,1367	0,1478
σ_V^2	1	0,9252	0,3280	0,2313
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
α	1	1,0109	0,0515	0,0530
β_1	-1	-1,0146	0,1390	0,1327
γ_0	-1	-1,0572	0,2504	0,3016
γ_1	1	1,0147	0,1505	0,1459
σ_V^2	0,5	0,5416	0,2029	0,1371

Note que, a diferença entre os valores médios estimados e os valores verdadeiros são menores do que um desvio-padrão, o que indica que os resultados obtidos via simulação estão bons. Note também que esta diferença é menor do que no caso $n_k = 15$. Ou seja, como no caso bayesiano, o fato de termos mais indivíduos em cada clínica nos fornece valores estimados mais acurados.

Note ainda que, como no caso $n_k = 15$, os desvios-padrão das médias estão próximos aos desvios-médio em praticamente todos os casos. Isso é um bom indício de que os desvios estão sendo bem estimados pelas segundas-derivadas da log-verossimilhança, ou seja, pela informação de Fisher. Somente para o parâmetro σ_V^2 esta diferença é um pouco maior.

As Figuras de 3.10 a 3.12 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

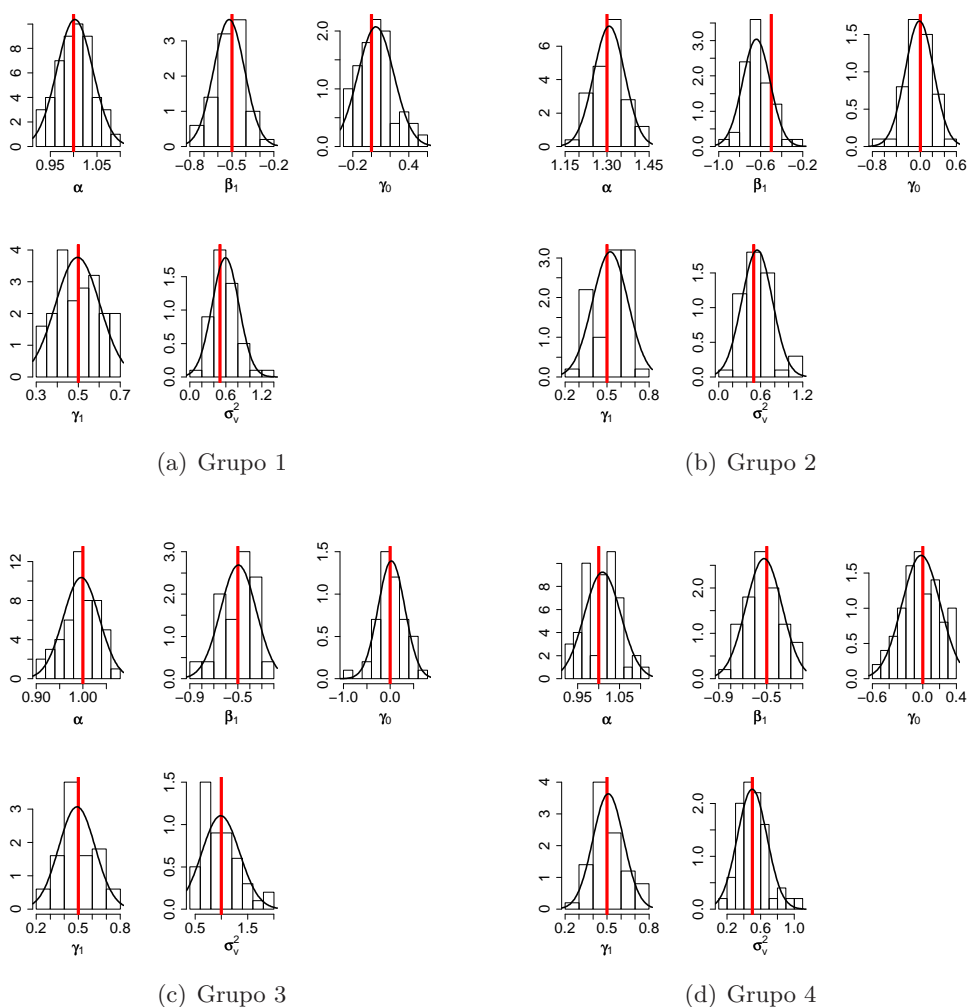


Figura 3.10: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – parte 1

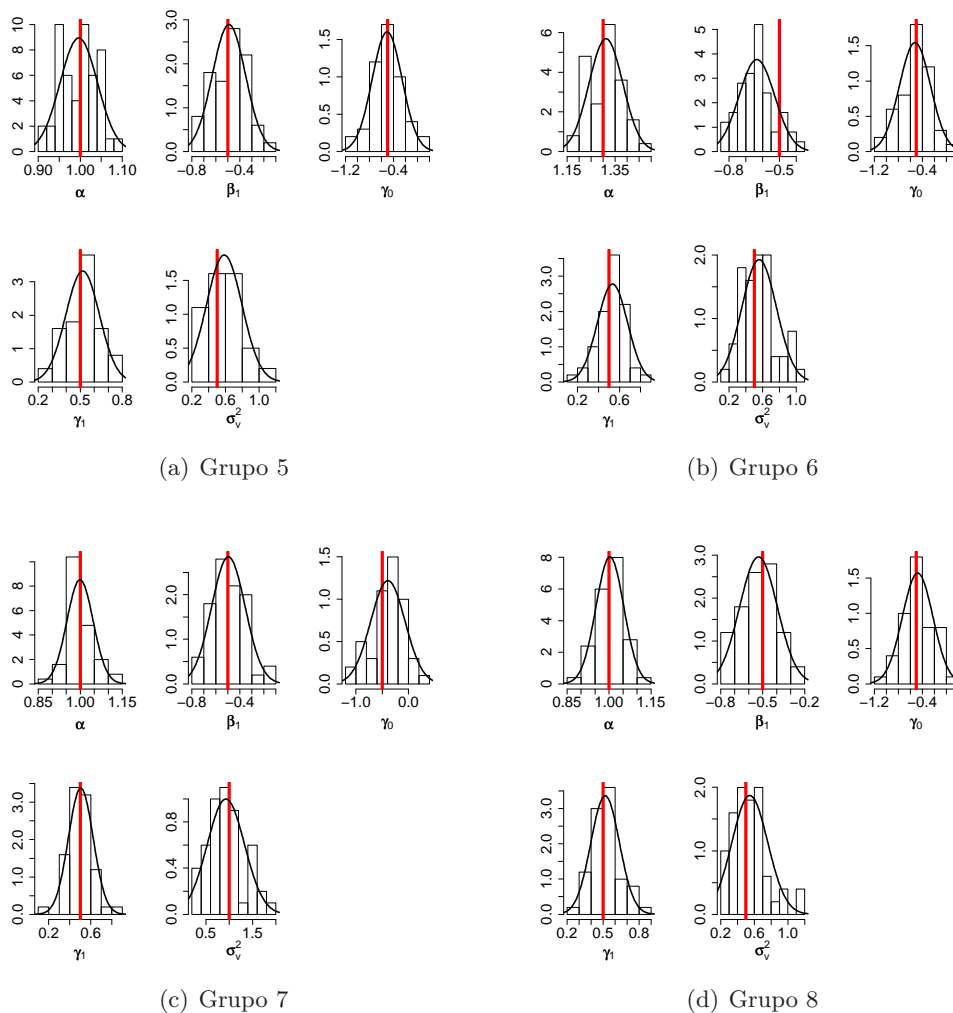


Figura 3.11: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – parte 2

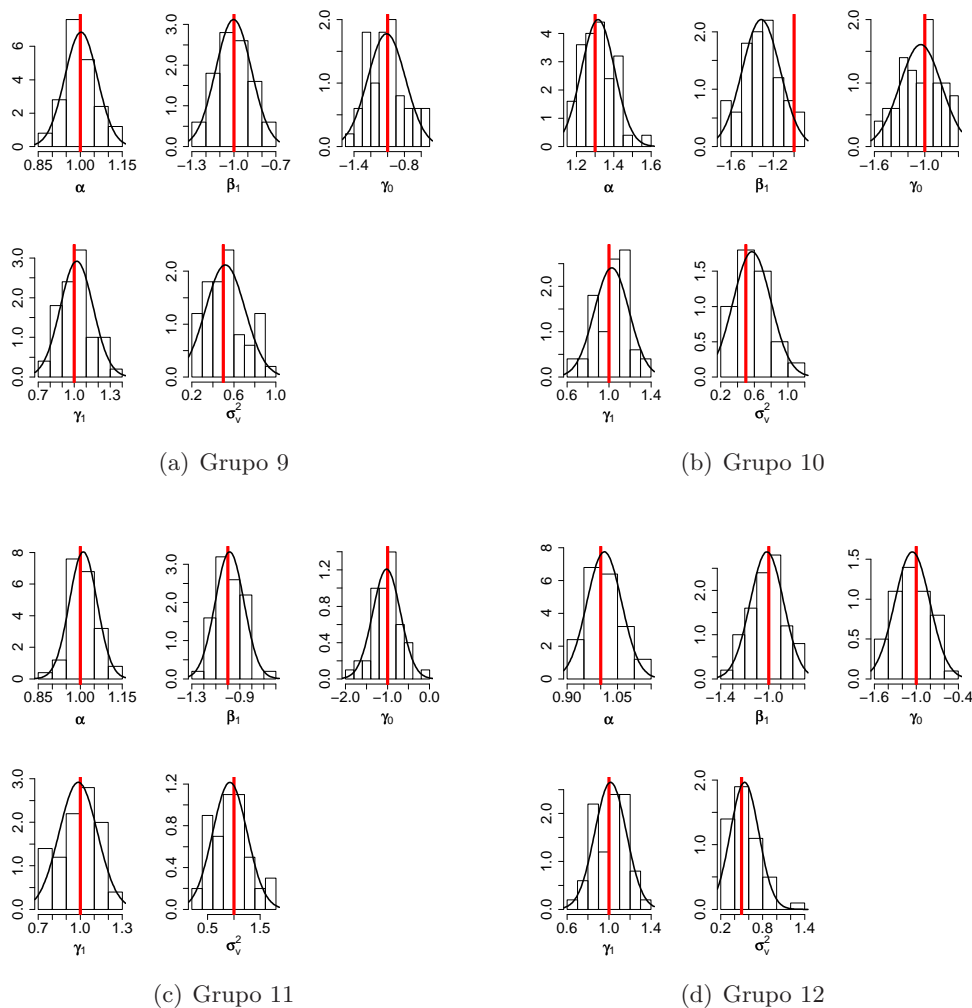


Figura 3.12: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, representados pelas retas verticais. Indicam, também, que as estimativas não estão muito longe da normalidade assintótica, como em todos os casos estudados.

3.4 Conclusões

Os resultados, tanto bayesianos quanto clássicos, mostram que os valores obtidos nas estimativas estão próximos aos valores reais, o que sugere que as estimações utilizando ambos os métodos de inferência implementados foram adequadas. Para se afirmar isso com mais segurança, o ideal seria fazer mais simulações, possivelmente para um número maior de valores de parâmetros.

Praticamente em todos os casos, a diferença entre a média do valor estimado e o valor verdadeiro é menor do que um desvio-padrão, o que sugere que os resultados estão razoavelmente precisos.

Os histogramas feitos com os valores estimados nas simulações ilustram que os valores verdadeiros, representados pelas retas verticais, estão dentro do esperado. E, que os valores têm distribuições aproximadamente normais.

Com relação à convergência das cadeias, na abordagem bayesiana, os valores de $\sqrt{\hat{R}}$ estão bem próximos a 1 para todas as estimativas de interesse indicando que houve a convergência das cadeias e que foi atingida a distribuição de cada estimador.

Nas simulações clássica e bayesiana, os valores estimados no caso $n_k = 50$ estão mais próximos dos valores verdadeiros do que no caso $n_k = 15$, mas em ambos os casos os valores estimados e verdadeiros diferem, quase sempre, por uma quantidade menor que um desvio-padrão.

Note, também, que, em ambas as abordagens, a diferença entre a média das estimativas e o valor verdadeiro para o caso $n_k = 50$ é menor do que no caso $n_k = 15$. Ou seja, o fato de termos mais indivíduos em cada clínica nos fornece valores estimados mais acurados.

Com relação aos desvios-padrão das estimativas, é perceptível que os valores estão, quase sempre, menores no caso $n_k = 50$ em comparação com $n_k = 15$. A exceção ocorre no caso de σ_v^2 em que essa situação se inverte, mas a acurácia do valor estimado é melhor em $n_k = 50$ do que em $n_k = 15$. Isso em ambas as inferências bayesiana e clássica.

Capítulo 4

Análise de dados - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios

Para ilustrar uma aplicação de uso do modelo de mistura com efeitos aleatórios, Yau e Ng (2001) utilizaram um conjunto de dados em Kalbfleisch e Prentice (1980, pp.225-229). Tais dados são parte de um estudo feito pelo *Radiation Therapy Oncology Group* nos Estados Unidos. A seguir, é apresentada a utilização deste mesmo conjunto de dados no modelo apresentado no Capítulo 2, para fração de cura com efeitos aleatórios, a partir do modelo de Chen-Ibrahim-Sinha.

4.1 Dados utilizados

Os dados são de pacientes, de seis clínicas, com carcinomas na orofaringe. Ao entrarem no estudo, os pacientes eram aleatoriamente encaminhados para um dos dois grupos de tratamentos: radioterapia somente ou radioterapia associada a um agente quimioterápico. Foram coletados dados de muitas covariáveis que poderiam estar relacionadas com a sobrevivência do paciente, dentre elas estão sexo, estágio do tumor (“estadio”), estágio de nódulos linfáticos, idade, grau de diferenciação do tumor, e condições funcionais do paciente na hora do diagnóstico.

O estudo visava identificar fatores de risco, determinar a extensão do efeito dos fatores de risco e comparar os dois tipos de tratamentos em relação à sobrevivência do paciente. Os tratamentos foram aplicados nos pacientes por um período de 90 dias, após os quais cada paciente recebeu cuidados médicos pelas instituições participantes do estudo. Neste período, após os 90 dias, a única restrição para as clínicas era o tipo de tratamento em estudo. Por não haver mais nenhuma limitação quanto à continuação no tratamento, as diferentes clínicas puderam apresentar diversos tratamentos e facilidades. Isto indica que possa ter havido uma diferença de resultados (efeitos) para as diferentes clínicas, sendo importante que o modelo a estudar os resultados considere o efeito da clínica no tempo de sobrevivência dos pacientes e no número de indivíduos que se curam, levando-se em consideração que há uma correlação entre os pacientes que estavam na mesma clínica.

Em Kalbfleisch e Prentice (1980) os efeitos de clínicas foram tratados como efeitos fixos. Em Yau e Ng (2001), os efeitos foram tratados com efeitos aleatórios, no modelo de mistura.

Como em Yau e Ng (2001), para a ilustração do uso do modelo apresentado no Capítulo 2, serão considerados somente os casos de carcinoma na parte faríngea da língua e o estágio do tumor para todos os pacientes, curados ou não.

O estágio do tumor mede a extensão do tumor em que

- T_1 se refere a tumores primários, em que o maior diâmetro é menor ou igual a 2 cm;
- T_2 se refere a tumores primários, em que o maior diâmetro mede de 2 cm a 4 cm;
- T_3 se refere a tumores primários, em que o maior diâmetro é maior ou igual a 4 cm;
- T_4 se refere a tumores maciços espalhados para tecidos adjacentes.

A partir dos dados de estágio do tumor, é considerada a função indicadora de estágio do tumor, dada por

$$X = \begin{cases} 0, & \text{se } T_1, T_2, \text{ ou } T_3 \\ 1, & \text{se } T_4 \end{cases} \quad (4.1)$$

O evento “falha” é considerado morte pelo carcinoma específico. Mortes por outras causas são tratadas como censuras. Por esta definição, alguns pacientes podem nunca apresentar o evento, principalmente aqueles que respondem de modo satisfatório ao tratamento. Estes pacientes podem ser considerados curados, sendo adequado o uso de um modelo que permita a inclusão de uma fração de curados. É razoável pensar que a fração de cura depende do estágio do tumor, sendo este fator levado em consideração na modelagem do problema.

Na Tabela 4.1 são apresentados os dados a serem utilizados, em que Tempo é o número de dias que o paciente sobreviveu desde a data do diagnóstico, X é a função definida em (4.1) e ν é a função indicadora de falha, ou seja

$$\nu = \begin{cases} 0, & \text{se censura} \\ 1, & \text{se falha} \end{cases}$$

Tabela 4.1: Dados de Tratamento de Carcinoma na Orofaringe

Clínica	Tempo	X	ν	Clínica	Tempo	X	ν
1	1312	0	0	3	1307	0	0
1	1219	0	0	3	800	0	1
1	264	1	1	3	666	0	1
1	11	1	1	3	1060	0	0
1	170	0	1	3	465	0	1
1	372	0	1	3	518	0	1
1	182	0	0	4	235	0	1
1	723	1	0	4	324	0	1
1	90	0	0	4	1565	0	0
2	270	0	1	4	1234	0	0
2	162	0	1	4	293	0	1
2	228	1	1	4	593	1	0
2	370	1	1	4	213	0	1
2	805	1	1	5	245	1	1
2	517	0	1	5	262	0	1
2	599	0	1	5	915	0	1
2	89	1	1	5	1377	1	0
2	99	0	1	5	172	1	1
2	219	1	1	5	15	1	1
2	154	1	1	5	382	1	1
2	266	1	1	5	363	0	1
2	407	0	1	5	107	1	1
2	751	0	0	5	757	1	1
3	184	0	1	5	38	0	1
3	1064	0	1	6	222	0	1
3	1472	0	0	6	1644	0	0
3	526	0	1	6	785	0	1
3	1495	0	0	6	11	1	1
3	661	0	1	6	1095	0	0
3	407	0	1	6	914	0	0
3	548	0	1	6	346	0	1
3	1317	0	0	6	608	0	1
3	1317	0	0	6	112	0	1

Fonte: Adaptado de Kalbfleisch e Prentice (1980)

A partir dos dados da Tabela 4.1, foram feitos os gráficos de Kaplan-Meier para cada uma das seis clínicas. O fato de as curvas estabilizarem em alturas diferentes de zero pode ser um indício de que haja uma parcela de pacientes curados. Outro indício de que é adequado um estudo com efeitos de clínica é o fato de que os gráficos de Kaplan-Meier (Figura 4.1) para cada clínica estabilizarem em alturas distintas.

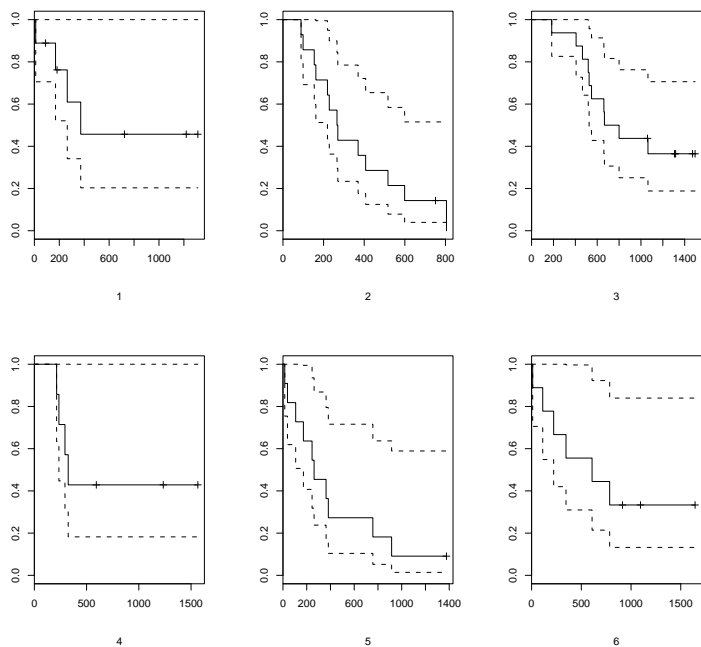


Figura 4.1: Gráficos de Kaplan-Meier para cada uma das seis clínicas

Para este conjunto de dados, o modelo apresentado no Capítulo 2 assume a forma dada em (2.2) em que

- $y_{jk} \sim \text{Weibull}(\alpha, \lambda_{jk})$;
- $\boldsymbol{\beta} = \beta_1$, portanto, $\lambda_{jk} = \mathbf{x}_{jk}^\top \boldsymbol{\beta} = \mathbf{x}_{jk} \beta_1$;
- $\boldsymbol{\gamma} = [\gamma_0, \gamma_1]^\top$
- os valores de \mathbf{x}_{jk} são lidos a partir da coluna X da Tabela 4.1, em que $M = 6$, (ou seja, $k = 1, 2, \dots, 6$), $j = 1, 2, \dots, n_k$ e $n_1 = 9$, $n_2 = 14$, $n_3 = 16$, $n_4 = 7$, $n_5 = 11$ e $n_6 = 9$.

4.2 Inferência bayesiana via Metropolis-Hastings

Uma primeira análise desses dados foi feita com a utilização do programa Winbugs (versão 1.4). Foi adotado $\alpha = 1$ fixado, ou seja, $y_{jk} \sim \text{Weibull}(\alpha, \lambda_{jk})$ ou, equivalentemente, $y_{jk} \sim \text{Exponencial}(\lambda_{jk})$. Os comandos utilizados encontram-se no apêndice A.2. Os resultados obtidos estão apresentados na Tabela 4.2, em que a coluna de média apresenta o valor da média entre os valores amostrados, DP apresenta o desvio-padrão e erro MC está mostrando os valores obtidos pelo erro padrão de Monte Carlo para a média.

Tabela 4.2: Resultados – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings

Resultados das posteriores, baseados em 30000 amostras								
parâmetro	média	DP	erro MC	2,50%	mediana	97,50%	início	amostras
β_1	-6,8400	0,7873	0,0358	-8,7420	-6,6950	-5,7040	4001	26000
γ_0	-0,4617	0,2365	0,0063	-0,9453	-0,4572	-0,0242	4001	26000
γ_1	1,4960	0,6874	0,0313	0,4221	1,3900	3,1560	4001	26000
σ_V^2	0,2559	0,1344	0,00162	0,1022	0,2241	0,5981	4001	26000

Note que σ_V^2 é muito pequeno o que indicaria que a componente de variância não é significativamente diferente de zero, de modo que não há efeito de clínica na parte da população de curados. Este resultado já tinha sido apontado em Yau e Ng (2001), em que modelaram os dados via modelo de mistura com efeitos aleatórios tanto na parcela da população que se cura quanto na que não se cura. Eles se utilizaram de uma abordagem clássica, em que os estimadores foram obtidos via máxima verossimilhança e via REML.

A Figura 4.2 mostra os histogramas feitos com os dados amostrados pelo programa Winbugs.

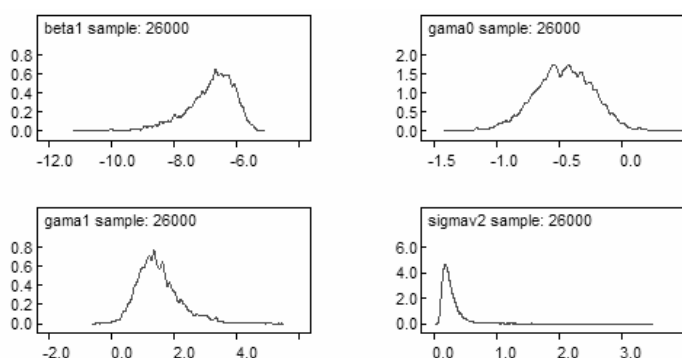


Figura 4.2: Histogramas – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings

O software Winbugs, neste exemplo, executou o algoritmo Metropolis-Hastings.

Para se incluir o caso em que $\alpha \neq 1$, ou seja, a distribuição é uma Weibull não exponencial, foram feitas alterações no programa para o Winbugs, que foi então recompilado. Os novos resultados estão apresentados na Tabela 4.3.

Tabela 4.3: Resultados – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings - com alfa incluído

Resultados das posteriores, baseados em 30000 amostras								
parâmetro	média	DP	erro MC	2,50%	mediana	97,50%	início	amostras
α	0,1501	0,0151	0,0001	0,1207	0,1498	0,1798	4001	26000
β_1	-3,8410	1,1330	0,0709	-6,4730	-3,7400	-1,9800	4001	26000
γ_0	-0,3496	0,2358	0,0059	-0,8307	-0,3405	0,0938	4001	26000
γ_1	3,1090	1,1330	0,0715	1,2210	3,0120	5,7090	4001	26000
σ_V^2	0,2529	0,1321	0,0015	0,1031	0,2214	0,5949	4001	26000

Novamente, o valor estimado para σ_V^2 é muito pequeno, indicando que pode não haver diferença de efeitos entre as clínicas, para os curados.

A Figura 4.3 mostra os histogramas feitos com os dados gerados pelo programa Winbugs, neste caso em que α não está fixado em 1.

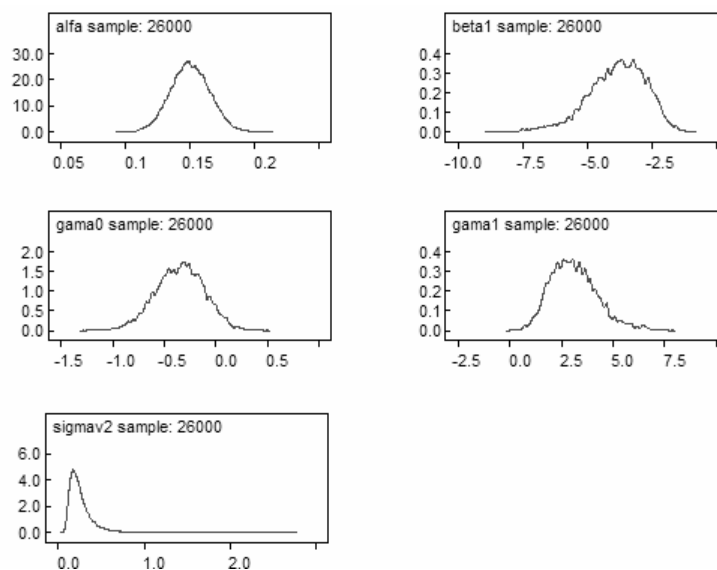


Figura 4.3: Histogramas – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings - com alfa incluído

A Figura 4.4 mostra os gráficos dos valores de $\sqrt{\hat{R}}$ (Gelman e Rubin, 1992), como definidos na seção 2.2.

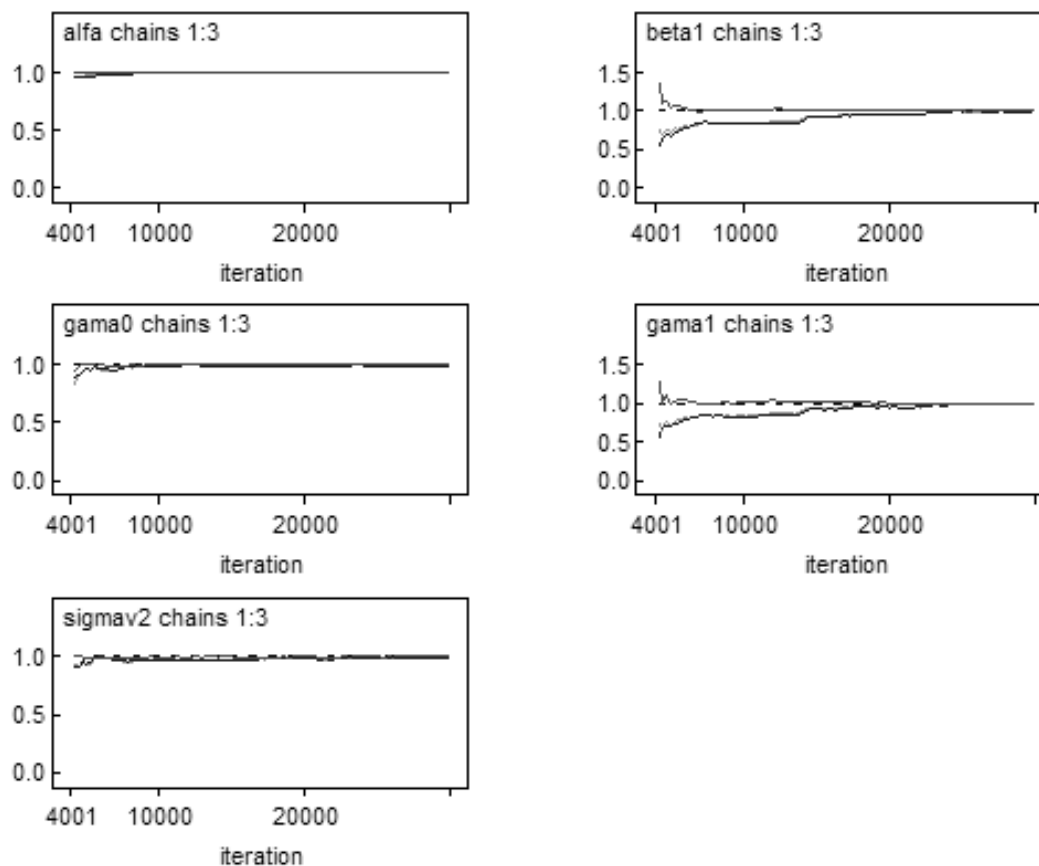


Figura 4.4: $\sqrt{\hat{R}}$ – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – Metropolis-Hastings

Note que para todos os parâmetros, os valores estão tendendo ao valor 1, o que indica que as cadeias estão convergindo e que foi atingida a distribuição de cada estimador.

4.3 Análise clássica

O conjunto de dados, apresentado na seção 4.1, foi novamente utilizado sendo feita a abordagem relatada na seção 2.3. Desta vez, o software utilizado para a compilação dos dados foi Maple versão 10. Os comandos estão apresentados no apêndice A.3.

Do mesmo modo que foi feito anteriormente, foram analisadas duas situações: $\alpha = 1$ fixado (distribuição exponencial) e α não fixado, a ser estimado pelo modelo.

No primeiro caso ($\alpha = 1$ fixado), os resultados obtidos estão na Tabela 4.4.

Tabela 4.4: Resultados – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – REML, com alfa fixado 1

parâmetros	estimativa	desvio-padrão
α (fixado)	1	–
β_1	-6,8388	0,9432
γ_0	-0,4273	0,3466
γ_1	1,4321	0,7809
σ_V^2 (BLUP)	0,0353	–
σ_V^2 (REML)	0,2187	0,0931

efeito de clínica	preditores (Blup)
V_1	-0,2834
V_2	0,2696
V_3	-0,0346
V_4	-0,1775
V_5	0,1270
V_6	0,0988

Note que os resultados da Tabela 4.4 estão bem próximos aos da Tabela 4.2.

Para o segundo caso, em que α é estimado, os resultados obtidos estão apresentados na Tabela 4.5.

Tabela 4.5: Resultados – câncer na orofaringe – modelo de Chen-Ibrahim-Sinha com efeitos aleatórios – REML

parâmetros	estimativa	desvio-padrão
α	0,1519	0,0153
β_1	-11,3699	52,8053
γ_0	-0,3280	0,3466
γ_1	10,6126	52,8050
σ_V^2 (BLUP)	0,0311	–
σ_V^2 (REML)	0,2129	0,0884

efeito de clínica	preditores (Blup)
V_1	-0,2680
V_2	-0,0589
V_3	-0,0589
V_4	-0,1281
V_5	0,1875
V_6	0,0245

Como no caso bayesiano, os valores estimados para σ_V^2 estão próximos de zero, indicando que pode não haver diferença entre as clínicas, para os indivíduos curados.

Note que, comparando-se os valores nas Tabelas 4.5 e 4.3, somente os valores estimados para α , γ_0 e σ_V^2 estão parecidos. Os valores estimados para β_1 e γ_1 estão muito diferentes.

Note, também, que os desvios-padrão para β_1 e γ_1 , na Tabela 4.5, estão muita elevadas, o que indica que a inferência não está boa. Um motivo para a obtenção destas variâncias é o fato de os dados serem muito heterogêneos entre as clínicas: o número de pacientes nas diferentes clínicas varia de 7 a 16; a clínica 2 possui somente 1 dos 14 pacientes censurado; a clínica 3 apresenta todos os 16 pacientes com tumor primário; os tempos de falha variam de 11 a 1644, sendo o valor médio aproximadamente 575 dias.

4.4 Conclusões

Os resultados obtidos com a estimação bayesiana, via Metropolis-Hastings estão próximos aos obtidos com a abordagem clássica, via REML.

É nítida uma diferença de estimativas para os parâmetros β_1 e γ_1 , quando estimados pelos dois métodos apresentados. Mas, para ambos, as variâncias estimadas classicamente estão muito elevadas, fato que pode ter ocorrido devido às características dos dados utilizados, como, por exemplo, baixo número de falhas em algumas clínicas, poucas observações com estágio avançado do tumor.

Capítulo 5

Modelo de mistura com efeitos aleatórios

Neste capítulo é apresentado o modelo de mistura com efeitos aleatórios na fração de curados, que visam avaliar o efeito que determinada clínica tem, por exemplo, no tratamento de câncer de certa população que tem presença de imunes. É feita a caracterização do modelo e apresentada uma abordagem bayesiana e uma clássica para a obtenção de estimadores.

5.1 Definindo o modelo

Considere o modelo de mistura apresentado na seção 1.1. Suponha que foram estudados casos em M clínicas e que para a k -ésima clínica havia n_k indivíduos, $k = 1, 2, \dots, M$. Para $j = 1, 2, \dots, n_k$ e $k = 1, 2, \dots, M$, considere T_{jk} o tempo observado de falha para o indivíduo j da clínica k . Considere ainda $N = \sum_{k=1}^M n_k$ o número total de observações. Seja \mathbf{x}_{jk} o vetor de covariáveis correspondentes a T_{jk} . Assim, a proporção p (curados), pode ser modelada por

$$p_{jk} = 1 - \pi(\mathbf{x}_{jk}) = 1 - \frac{1}{1 + e^{\mathbf{w}_{jk}^\top \boldsymbol{\gamma} + V_k}}$$

em que $\mathbf{w}_i = [1 \ \mathbf{x}_i^\top]^\top$, $\boldsymbol{\gamma}$ é um vetor de parâmetros, V_k é o efeito aleatório não observável da k -ésima clínica determinando a proporção de curados, sendo considerado que V_1, \dots, V_M são independentes com $V_k \sim N(0, \sigma_V^2)$. Note que valores maiores de V_k indicam que aqueles pacientes em tratamento pela k -ésima clínica terão uma probabilidade maior de cura.

Para $k = 1, 2, \dots, M$ e $j = 1, 2, \dots, n_k$, conforme apresentado na seção 1.1, para o modelo de riscos proporcionais de Cox, tem-se

$$h_2(t_{jk}; \mathbf{x}_{jk}) = h_{20}(t_{jk}) \cdot e^{\mathbf{x}_{jk}^\top \boldsymbol{\beta}}$$

em que $\boldsymbol{\beta}$ é o vetor de parâmetros da regressão e $h_{20}(t)$ é a função de riscos basal para os indivíduos não curados.

Neste trabalho, vamos supor que a função de risco basal para os não curados segue conforme a

distribuição Weibull dada por

$$h_{20}(t_{jk}) = \lambda \alpha (t_{jk})^{\alpha-1} \quad (5.1)$$

em que $\alpha > 0$ e $\lambda > 0$.

Para se determinar a verossimilhança, considere C_{jk} o tempo de censura (à direita) para o j -ésimo indivíduo da clínica k . Seja $Y_{jk} = \min\{T_{jk}, C_{jk}\}$ e seja ν_{jk} o indicador de falha, ou seja, $\nu_{jk} = 1$ se y_{jk} é falha e 0 se for censura.

Assim, os dados observados são $D_{obs} = (\mathbf{y}, \boldsymbol{\nu}, \mathbf{X})$ em que $\mathbf{y} = (y_{jk})_{j=1, \dots, n_k}^{k=1, \dots, M}$, $\boldsymbol{\nu} = (\nu_{jk})_{j=1, \dots, n_k}^{k=1, \dots, M}$ e \mathbf{X} são as covariáveis.

Os dados completos são $D_c = (\mathbf{y}, \boldsymbol{\nu}, \mathbf{V}, \mathbf{X})$ que incluem os dados observados e os dados não observados $\mathbf{V} = [V_1, V_2, \dots, V_M]$. O vetor de parâmetros é dado por $\boldsymbol{\phi} = (\alpha, \lambda, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top, \sigma_V^2)$.

Assim, utilizando (1.1), temos

$$\begin{aligned} f(\mathbf{y}, \mathbf{V}) &= f(\mathbf{y}|\mathbf{V}) \cdot f(\mathbf{V}) \\ &= \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ S(y_{jk}; \mathbf{x}_{jk})^{1-\nu_{jk}} \left(\pi(\mathbf{x}_{jk}) h_{20}(y_{jk}) e^{\mathbf{x}_{jk}^\top \boldsymbol{\beta}} S_{20}(y_{jk})^{\exp(\mathbf{x}_{jk}^\top \boldsymbol{\beta})} \right)^{\nu_{jk}} \right\} \\ &\quad \cdot \prod_{k=1}^M \left\{ \frac{1}{\sqrt{2\pi\sigma_V^2}} e^{-\frac{V_k^2}{2\sigma_V^2}} \right\} \end{aligned} \quad (5.2)$$

De modo que a log-verossimilhança completa é dada por

$$\begin{aligned} l(\boldsymbol{\phi}|D_c) &= \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ (1 - \nu_{jk}) \ln(S(y_{jk}; \mathbf{x}_{jk})) + \nu_{jk} \ln \left(\pi(\mathbf{x}_{jk}) h_{20}(y_{jk}) e^{\mathbf{x}_{jk}^\top \boldsymbol{\beta}} S_{20}(y_{jk})^{\exp(\mathbf{x}_{jk}^\top \boldsymbol{\beta})} \right) \right\} + \\ &\quad + \sum_{k=1}^M \left\{ \frac{-M}{2} \ln(2\pi\sigma_V^2) - \frac{V_k^2}{2\sigma_V^2} \right\} \end{aligned} \quad (5.3)$$

5.2 Inferência bayesiana

Com já afirmado anteriormente, na abordagem bayesiana, a presença de quantidades não observáveis não representa um problema de inferência já que toda e qualquer quantidade desconhecida é uma variável aleatória. Assim, tanto para as quantidades de interesse (parâmetros), quanto para os efeitos aleatórios, é necessário obter sua distribuição a posteriori, ou seja, condicional aos dados observados. Basicamente, a distribuição a posteriori é proporcional ao produto da distribuição a priori para as quantidades desconhecidas e da função de verossimilhança.

Para o modelo de mistura com efeitos aleatórios definido na seção 5.1, supondo prioris não

informativas para γ e para β , consideremos as distribuições a priori para o vetor de parâmetros ϕ :

$$\pi(\phi) = \pi(\alpha, \lambda, \beta^\top, \gamma^\top, \sigma_V^2) \propto \pi(\alpha)\pi(\lambda)\pi(\sigma_V^2),$$

em que as distribuições a priori são tais que

1. $\alpha \sim \text{Gama}(\delta_0, \tau_0)$

$$\pi(\alpha) \propto \alpha^{\delta_0-1} e^{-\tau_0 \alpha};$$

2. $\lambda \sim \text{Gama Inversa}(a, b)$

$$\pi(\lambda) \propto (\lambda)^{-(a+1)} e^{-\frac{b}{\lambda}}.$$

3. $\sigma_V^2 \sim \text{Gama Inversa}(a_V, b_V)$

$$\pi(\sigma_V^2) \propto (\sigma_V^2)^{-(a_V+1)} e^{-\frac{b_V}{\sigma_V^2}}.$$

Considerando a função de verossimilhança em (5.2), a distribuição a posteriori é dada por

$$\begin{aligned} \pi(\phi, \mathbf{V} | D_{obs}) &= \prod_{k=1}^M \prod_{j=1}^{n_k} \left\{ S(y_{jk}; \mathbf{x}_{jk})^{1-\nu_{jk}} \left(\pi(\mathbf{x}_{jk}) h_{20}(y_{jk}) e^{\mathbf{x}_{jk}^\top \beta} S_{20}(y_{jk})^{\exp(\mathbf{x}_{jk}^\top \beta)} \right)^{\nu_{jk}} \right\} \\ &\cdot \prod_{k=1}^M \left\{ \frac{1}{\sqrt{2\pi\sigma_V^2}} e^{-\frac{V_k^2}{2\sigma_V^2}} \right\} \\ &\cdot \alpha^{\delta_0-1} e^{-\tau_0 \alpha} (\lambda)^{-(a+1)} e^{-\frac{b}{\lambda}} (\sigma_V^2)^{-(a_V+1)} e^{-\frac{b_V}{\sigma_V^2}}. \end{aligned} \quad (5.4)$$

Uma forma de se obter estimativas dos parâmetros seria via utilização de amostrador de Gibbs (Geman; Geman, 1984), método de simulação MCMC para aproximar as distribuições conjuntas e marginais, amostrando das distribuições condicionais. Assim, o próximo passo seria tentar identificar as distribuições condicionais, a partir de (5.4). Do mesmo modo que no modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, nem todas as distribuições condicionais terão forma fechada. Novamente, a alternativa escolhida é utilização de "Metropolis-Hastings" (Chib & Greenberg, 1995).

Para a verificação da adequação da convergência do modelo, serão calculados os valores de $\sqrt{\hat{R}}$, já definido na seção 2.2, lembrando que valores de $\sqrt{\hat{R}}$ próximos a 1 para todas as estimativas de interesse indicam que houve convergência das cadeias e que foi atingida a distribuição de cada estimador.

5.3 Análise clássica

Devido à presença de quantidades não observadas, dadas pelos efeitos aleatórios \mathbf{V} , a equação (5.2) não é exatamente uma verossimilhança. O modo escolhido para se encontrar estimadores para os parâmetros se baseia em primeiro calcular o "melhor preditor linear não-viesado" (BLUP) e estendê-lo para se obter estimadores de máxima verossimilhança residual (REML) aproximados, como já ocorrido no modelo de Chen-Ibrahim-Sinha com efeitos aleatórios.

Para se encontrar o BLUP, considere:

- $\phi_E = (\alpha, \lambda, \beta^\top, \gamma^\top, \sigma_V^2, \mathbf{V}^\top)^\top$;
- l o logaritmo neperiano de (5.2), ou seja, a expressão dada em (5.3);
- \mathbf{U} o vetor escore de l em ϕ_E , ou seja, o vetor das primeiras derivadas de l em relação à ϕ_E ;
- H a matriz hessiana de l em ϕ_E , ou seja, a matriz das segundas derivadas de l em relação à ϕ_E .

O BLUP para \mathbf{V} e estimadores para ϕ serão encontrados a partir das raízes do vetor escore \mathbf{U} .

Lembre-se que l é dada por

$$\begin{aligned} l(\phi|D_c) &= \sum_{k=1}^M \sum_{j=1}^{n_k} \left\{ (1 - \nu_{jk}) \ln(S(y_{jk}; \mathbf{x}_{jk})) + \nu_{jk} \ln \left(\pi(\mathbf{x}_{jk}) h_{20}(y_{jk}) e^{\mathbf{x}_{jk}^\top \beta} S_{20}(y_{jk})^{\exp(\mathbf{x}_{jk}^\top \beta)} \right) \right\} + \\ &+ \sum_{k=1}^M \left\{ \frac{-M}{2} \ln(2\pi\sigma_V^2) - \frac{V_k^2}{2\sigma_V^2} \right\}. \end{aligned}$$

O vetor escore \mathbf{U} é dado por $\mathbf{U} = \left(\frac{\partial l}{\partial \alpha}, \frac{\partial l}{\partial \lambda}, \frac{\partial l}{\partial \beta^\top}, \frac{\partial l}{\partial \gamma^\top}, \frac{\partial l}{\partial \sigma_V^2}, \frac{\partial l}{\partial \mathbf{V}^\top} \right)^\top$.

Para se encontrar o BLUP, deve-se encontrar a solução de $\mathbf{U} = \vec{0}$. Mas, isto não é de fácil solução.

Entretanto, note que a derivada de l em σ_V^2 é dada por

$$\begin{aligned} \frac{\partial l}{\partial \sigma_V^2} &= \sum_{k=1}^M \left\{ \frac{-1}{2} \frac{1}{\sigma_V^2} - \frac{V_k^2}{2} \left(\frac{-1}{(\sigma_V^2)^2} \right) \right\} = \sum_{k=1}^M \left\{ \frac{-1}{2\sigma_V^2} + \frac{V_k^2}{2(\sigma_V^2)^2} \right\} = \frac{1}{2\sigma_V^2} \sum_{k=1}^M \left\{ -1 + \frac{V_k^2}{\sigma_V^2} \right\} \\ &= \frac{1}{2\sigma_V^2} \left\{ \sum_{k=1}^M (-1) + \frac{1}{\sigma_V^2} \sum_{k=1}^M V_k^2 \right\} = \frac{1}{2\sigma_V^2} \left\{ -M + \frac{1}{\sigma_V^2} \sum_{k=1}^M V_k^2 \right\} = \frac{-M\sigma_V^2 + \sum_{k=1}^M V_k^2}{2\sigma_V^2} \end{aligned}$$

Portanto,

$$\begin{aligned}\frac{\partial l}{\partial \sigma_V^2} &= 0 \iff \frac{-M\sigma_V^2 + \sum_{k=1}^M V_k^2}{2\sigma_V^2} = 0 \iff -M\sigma_V^2 + \sum_{k=1}^M V_k^2 = 0 \iff \\ \sigma_V^2 &= \frac{\sum_{k=1}^M V_k^2}{M}\end{aligned}$$

Como, analiticamente, as soluções das demais equações não são viáveis, uma alternativa é utilizar o método de Newton-Raphson, que já foi apresentado na seção 2.3 e é repetido aqui, para maior clareza do leitor.

Se $\phi_E(t+1)$ são as estimativas em ϕ_E no passo $t+1$, se $\phi_E(t)$ são as estimativas em ϕ_E no passo t , se $\mathbf{U}(t)$ é o vetor escore calculado em $\phi_E(t)$ e se $H(t)$ é a matriz hessiana calculada em $\phi_E(t)$, então

$$\phi_E(t+1) = \phi_E(t) - H(t)^{-1} \cdot \mathbf{U}(t) \quad (5.5)$$

em que (5.5) é inicializado com valores para ϕ_E e deve ser calculado até que $\|\phi_E(t+1) - \phi_E(t)\| < \varepsilon$, com ε suficientemente pequeno. Pode-se, por exemplo, usar $\varepsilon = 10^{-3}$.

Se $\tilde{\mathbf{V}}$ são os preditores para \mathbf{V} , e $\tilde{\phi}_E$ são as estimativas para ϕ_E , obtidos por BLUP e se a matriz hessiana-inversa calculada em $\tilde{\phi}_E$ é decomposta relativamente a $(\alpha, \lambda, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top) | \sigma_V^2 | \mathbf{V}^\top$ por

$$H^{-1} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \quad (5.6)$$

então, segundo McGilchrist e Yau (1995) e Yau e Ng (2001), as estimativas para σ_V^2 via REML serão calculadas por

$$\hat{\sigma}_V^2 = \frac{1}{M} \left(-tr(A_{33}) + \tilde{\mathbf{V}}^\top \tilde{\mathbf{V}} \right).$$

A variância de $\hat{\sigma}_V^2$ é determinada por

$$Var(\hat{\sigma}_V^2) = 2 \left(\frac{1}{\sigma_V^2} \left(M + 2 \frac{1}{\sigma_V^2} tr(A_{33}) \right) - \frac{1}{\sigma_V^4} tr(A_{33}^2) \right)^{-1}$$

As variâncias dos demais parâmetros são obtidas através da matriz informação de Fisher, ou seja, são dadas por

$$Var \left(\left[\alpha, \lambda, \boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top \right]^\top \right) = -[A_{11}] \quad (5.7)$$

Capítulo 6

Estudos de simulação - modelo de mistura com efeitos aleatórios

Neste capítulo são apresentados um estudo de simulação feito para o modelo de mistura com efeitos aleatórios, apresentado no Capítulo 5, tanto para a abordagem bayesiana quanto para a clássica.

Para este estudo foram utilizados os programas R (versão 2.5.1), Winbugs (versão 1.4) e Maple (versão 10).

6.1 Características

Para este estudo, em uma primeira situação, estão assumidas as mesmas condições básicas apresentadas em Yau-Ng (2001) e que já foram apresentadas no Capítulo 3:

- $M = 10$, ou seja, 10 clínicas;
- $n_k = 15$ para todo $k = 1, 2, \dots, M$, ou seja, cada clínica possui 15 indivíduos;
- em cada clínica, 5 pacientes são aleatoriamente designados para o grupo de tratamento ($x_{jk} = 1$) e os 10 restantes, para o grupo controle ($x_{jk} = 0$);
- 12 grupos de parâmetros a serem analisados:
 1. $(\alpha = 1; \sigma_V^2 = 0,5; Censura = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1)$,
 2. $(\alpha = 1,3; \sigma_V^2 = 0,5; Censura = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1)$,
 3. $(\alpha = 1; \sigma_V^2 = 1; Censura = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1)$,
 4. $(\alpha = 1; \sigma_V^2 = 0,5; Censura = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1)$,
 5. $(\alpha = 1; \sigma_V^2 = 0,5; Censura = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1)$,
 6. $(\alpha = 1,3; \sigma_V^2 = 0,5; Censura = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1)$,
 7. $(\alpha = 1; \sigma_V^2 = 1; Censura = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1)$,

8. $(\alpha = 1; \sigma_V^2 = 0,5; Censura = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1)$,
 9. $(\alpha = 1; \sigma_V^2 = 0,5; Censura = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1)$,
 10. $(\alpha = 1,3; \sigma_V^2 = 0,5; Censura = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1)$,
 11. $(\alpha = 1; \sigma_V^2 = 1; Censura = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1)$,
 12. $(\alpha = 1; \sigma_V^2 = 0,5; Censura = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1)$;
- 50 simulações para cada grupo de parâmetros.

Em um segundo momento foi estudado o caso em $n_k = 50$ para todo $k = 1, 2, \dots, 15$ com as demais condições mantidas.

O que se espera é que:

- o desempenho das estimativas melhora, quando α aumenta de 1 para 1,3, porque quanto menor o valor de α , maior será a média da distribuição;
- uma diminuição na censura (de 500 para 300) piora o desempenho das estimativas, já que esta diminuição poderá fazer com que se tenha um aumento de indivíduos não curados que formam censurados;
- um aumento em σ_V^2 (de 0,5 para 1) produza uma maior variabilidade entre as clínicas, acarretando estimativas não tão boas quanto as obtidas para valores menores de σ_V^2 ;
- o aumento no número de indivíduos por clínica (de 15 para 50) melhora a acurácia dos resultados assintóticos.

6.2 Simulação - abordagem bayesiana

Em um primeiro momento foi feito um estudo de simulação com a utilização do programa Winbugs, versão 1.4. Os comandos utilizados encontram-se no apêndice A.5. Os dados utilizados foram gerados com a utilização do programa R e encontram-se no apêndice A.4.

Os dados foram compilados sempre a partir de três cadeias iniciadas em valores distintos, com 30 mil iterações cada, seguindo o mesmo padrão já apresentado no Capítulo 3 para o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios.

6.2.1 Caso $n_k = 15$

Os resultados obtidos estão na Tabela 6.1, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a média dos desvios-padrão encontrados em cada uma das 50 simulações.

Tabela 6.1: Resultado de 50 simulações via Metropolis-Hastings para $n_k = 15$ para o modelo de mistura

Grupo 1: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4915	0,2267	0,2234
γ_0	0	-0,0074	0,3056	0,2352
γ_1	0,5	0,3519	0,3561	0,3599
σ_V^2	0,5	0,2645	0,0796	0,1481
Grupo 2: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5375	0,2572	0,2454
γ_0	0	0,0475	0,2794	0,2340
γ_1	0,5	0,5915	0,3927	0,3695
σ_V^2	0,5	0,2570	0,0687	0,1429
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5548	0,2621	0,2326
γ_0	0	0,0097	0,3242	0,2489
γ_1	0,5	0,4804	0,3751	0,3677
σ_V^2	1	0,3769	0,2027	0,2351
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5048	0,2446	0,2320
γ_0	0	-0,0033	0,2721	0,2379
γ_1	0,5	0,4787	0,3635	0,3651
σ_V^2	0,5	0,2897	0,1127	0,1681
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5352	0,1955	0,2061
γ_0	-0,5	-0,4479	0,3034	0,2401
γ_1	0,5	0,5038	0,3138	0,3598
σ_V^2	0,5	0,2690	0,0789	0,1527

Grupo 6: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5071	0,2085	0,2054
γ_0	-0,5	-0,4296	0,3422	0,2371
γ_1	0,5	0,4576	0,3863	0,3592
σ_V^2	0,5	0,2394	0,0550	0,1276
Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4753	0,2108	0,2018
γ_0	-0,5	-0,3909	0,3245	0,2516
γ_1	0,5	0,3354	0,3991	0,3651
σ_V^2	1	0,3769	0,2021	0,2353
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5054	0,1912	0,2003
γ_0	-0,5	-0,4829	0,2451	0,2415
γ_1	0,5	0,4115	0,2892	0,3605
σ_V^2	0,5	0,2834	0,1225	0,1619
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-1,0252	0,2245	0,2045
γ_0	-1	-0,8421	0,3188	0,2537
γ_1	1	0,8655	0,3110	0,3679
σ_V^2	0,5	0,2893	0,1789	0,1683
Grupo 10: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-0,9669	0,2247	0,2021
γ_0	-1	-0,9506	0,3294	-0,2528
γ_1	1	0,9116	0,4410	0,3701
σ_V^2	0,5	0,2430	0,0556	0,1321
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-0,9903	0,2395	0,2028
γ_0	-1	-0,8847	0,3715	0,2702
γ_1	1	0,8690	0,3170	0,3752
σ_V^2	1	0,4129	0,2366	0,2661
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-1,0042	0,2149	0,2055
γ_0	-1	-0,9379	0,3115	0,2525
γ_1	1	0,9779	0,4146	0,3698
σ_V^2	0,5	0,2459	0,0624	0,1346

Note que a diferença entre os valores médios estimados estão próximos dos valores verdadeiros, principalmente levando-se em consideração os valores dos desvios-padrão, o que indica que os resultados obtidos via simulação estão razoavelmente bons. Note que os valores estimados para σ_V^2 são sempre menores que o seu correspondente valor verdadeiro, e que em alguns casos, a média das

estimativa não está muito próxima do valor verdadeiro para σ_v^2 , podendo ser um indício de que esta abordagem pode não ser muito boa.

As Figuras de 6.1 a 6.3 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

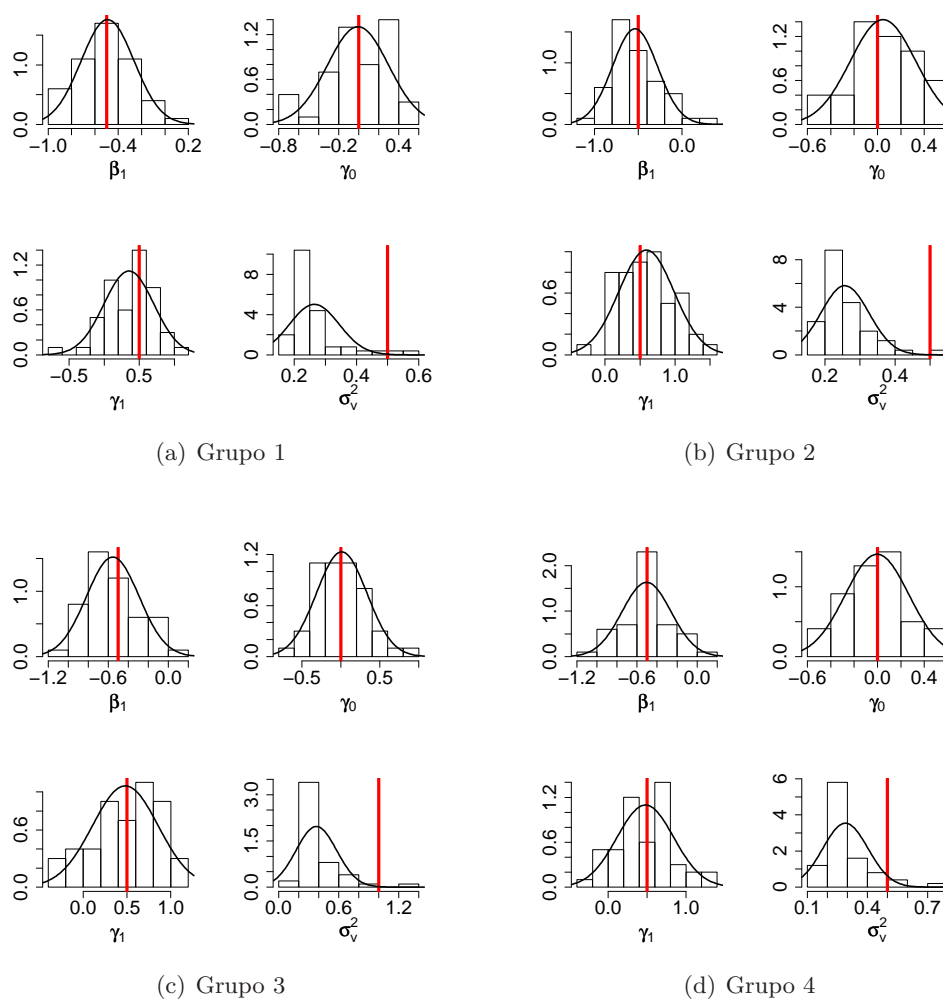


Figura 6.1: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – modelo de mistura – parte 1

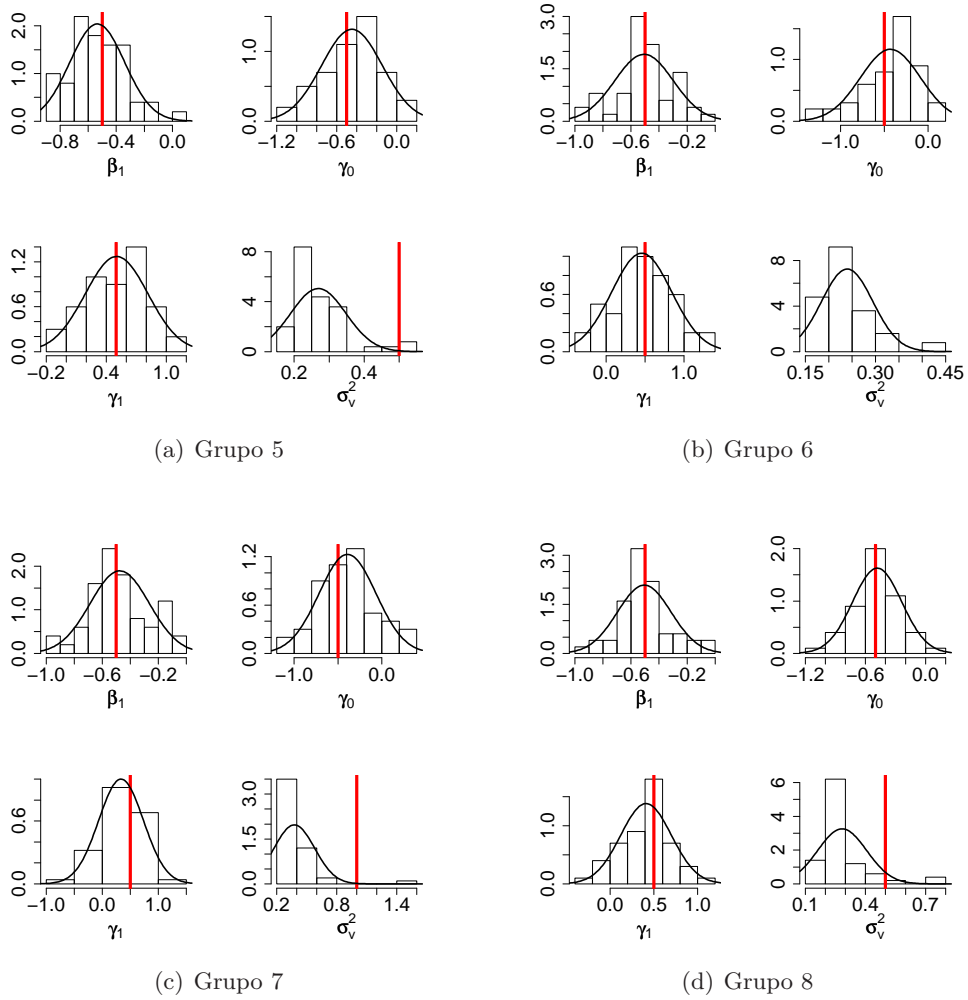


Figura 6.2: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – modelo de mistura– parte 2

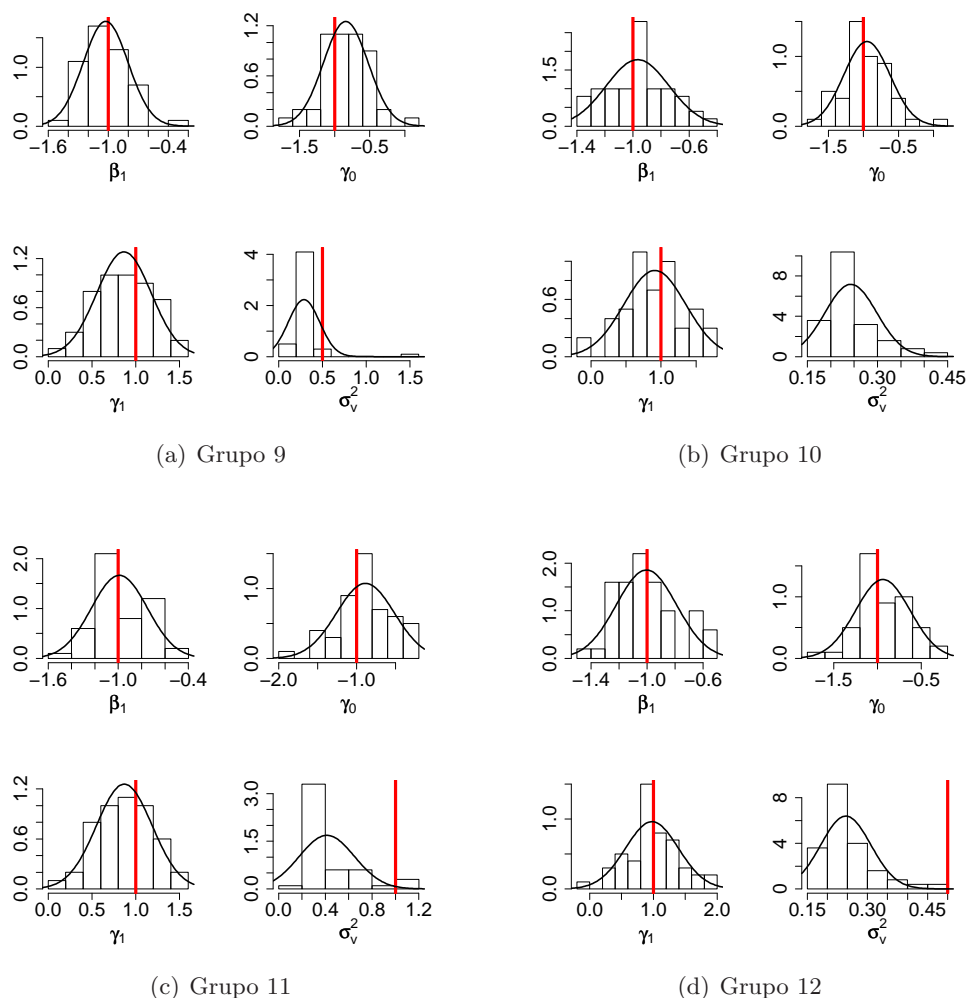


Figura 6.3: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 15$ – modelo de mistura – parte 3

Estes histogramas indicam que as estimativas obtidas para β_1 , γ_0 e γ_1 estão em torno dos valores verdadeiros, que estão representados pelas retas verticais. No caso de σ_V^2 , há um indício de que os valores estão sendo subestimados.

Para se avaliar se a convergência do modelo está adequada, calcula-se os valores de $\sqrt{\hat{R}}$, como definido ao final da seção 2.2. Para isso, considere $m = 3$ o número de cadeias utilizadas com valores iniciais distintos, em que cada uma apresenta $2n = 30000$ valores, ou seja, $n = 15000$. Foram calculados os valores de $\sqrt{\hat{R}}$ para simulações em cada um dos 12 grupos. Os resultados obtidos encontram-se na Tabela 6.2.

Tabela 6.2: Valores de $\sqrt{\hat{R}}$ - caso $n_k = 15$ - modelo de mistura

parâmetro	1	2	3	4	5	6
β_1	1,00060	1,00028	1,00013	1,00008	1,00003	1,00017
γ_0	1,00088	1,00158	1,00023	1,00011	1,00118	1,00046
γ_1	1,00005	1,00087	1,00036	1,00005	1,00146	1,00041
σ_V^2	1,00025	1,00090	1,00135	1,00195	1,00039	1,00039
parâmetro	7	8	9	10	11	12
β_1	1,00019	1,00023	1,00074	1,00009	1,00012	1,00010
γ_0	1,00002	1,00082	1,00054	1,00156	1,00215	1,00219
γ_1	1,00055	1,00088	1,00047	1,00213	1,00075	1,00093
σ_V^2	1,00120	1,00248	1,00035	1,00103	1,00281	1,00034

É perceptível que todos os valores estão bem próximos a 1 para todas as estimativas de interesse indicando que houve convergência da cadeia e que foi atingida a distribuição de cada estimador.

6.2.2 Caso $n_k = 50$

Os resultados obtidos estão na Tabela 6.3, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a média dos desvios-padrão encontrados em cada uma das 50 simulações.

Tabela 6.3: Resultado de 50 simulações via Metropolis-Hastings para $n_k = 50$ para o modelo de mistura

Grupo 1: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5029	0,1192	0,1245
γ_0	0	0,0070	0,2448	0,1790
γ_1	0,5	0,4848	0,1957	0,1996
σ_V^2	0,5	0,3889	0,1277	0,1746
Grupo 2: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,5240	0,1434	0,1269
γ_0	0	0,0744	0,2255	0,1827
γ_1	0,5	0,4841	0,1963	0,2012
σ_V^2	0,5	0,4266	0,2389	0,1881
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4797	0,1192	0,1228
γ_0	0	-0,0269	0,3378	0,2216
γ_1	0,5	0,4873	0,1905	0,2065
σ_V^2	1	0,7425	0,3809	0,3222
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = 0; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4950	0,1384	0,1246
γ_0	0	-0,0059	0,2548	0,1856
γ_1	0,5	0,4976	0,1630	0,2007
σ_V^2	0,5	0,4361	0,1706	0,1953
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4916	0,1097	0,1096
γ_0	-0,5	-0,4421	0,2599	0,1813
γ_1	0,5	0,4523	0,1915	0,1977
σ_V^2	0,5	0,3983	0,1571	0,1787
Grupo 6: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4979	0,1179	0,1098
γ_0	-0,5	-0,4696	0,2345	0,1874
γ_1	0,5	0,4872	0,2021	0,1986
σ_V^2	0,5	0,4383	0,1998	0,1966

Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4957	0,1099	0,1115
γ_0	-0,5	-0,4365	0,2922	0,2232
γ_1	0,5	0,5192	0,1987	0,2052
σ_V^2	1	0,7532	0,4462	0,3244
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-0,5	-0,4723	0,1109	0,1079
γ_0	-0,5	-0,5336	0,2339	0,1851
γ_1	0,5	0,4759	0,1990	0,1986
σ_V^2	0,5	0,4195	0,1823	0,1894
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-0,9885	0,1236	0,1100
γ_0	-1	-0,9702	0,2585	0,1888
γ_1	1	1,0073	0,2070	0,2046
σ_V^2	0,5	0,4014	0,1628	0,1836
Grupo 10: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-0,9896	0,1237	0,1096
γ_0	-1	-0,9566	0,2616	-0,1921
γ_1	1	0,9566	0,2257	0,2054
σ_V^2	0,5	0,4296	0,2217	0,1975
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-1,0013	0,1179	0,1096
γ_0	-1	-0,9973	0,3266	0,2416
γ_1	1	0,9934	0,2160	0,2141
σ_V^2	1	0,8491	0,4575	0,3723
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Média dos DPs
β_1	-1	-0,9682	0,0948	0,1084
γ_0	-1	-0,9580	0,2699	0,1884
γ_1	1	0,9214	0,2071	0,2040
σ_V^2	0,5	0,3949	0,1623	0,1815

Note que a diferença entre os valores médios estimados e os valores verdadeiros são menores do que um desvio-padrão, em quase todos os casos, o que indica que os resultados obtidos via simulação estão bons.

Note ainda que as estimativas obtidas para este caso $n_k = 50$ estão mais próximos dos valores verdadeiros do que no caso $n_k = 15$.

As Figuras de 6.4 a 6.6 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

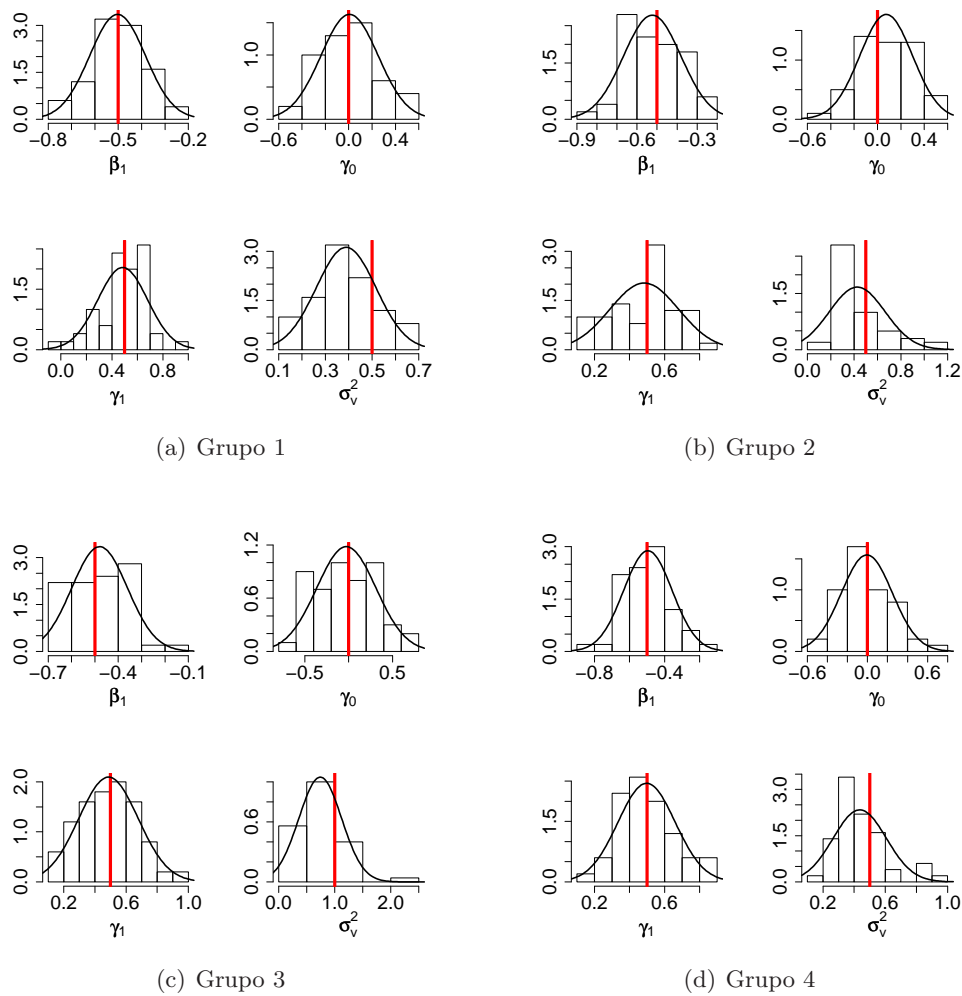


Figura 6.4: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – modelo de mistura – parte 1

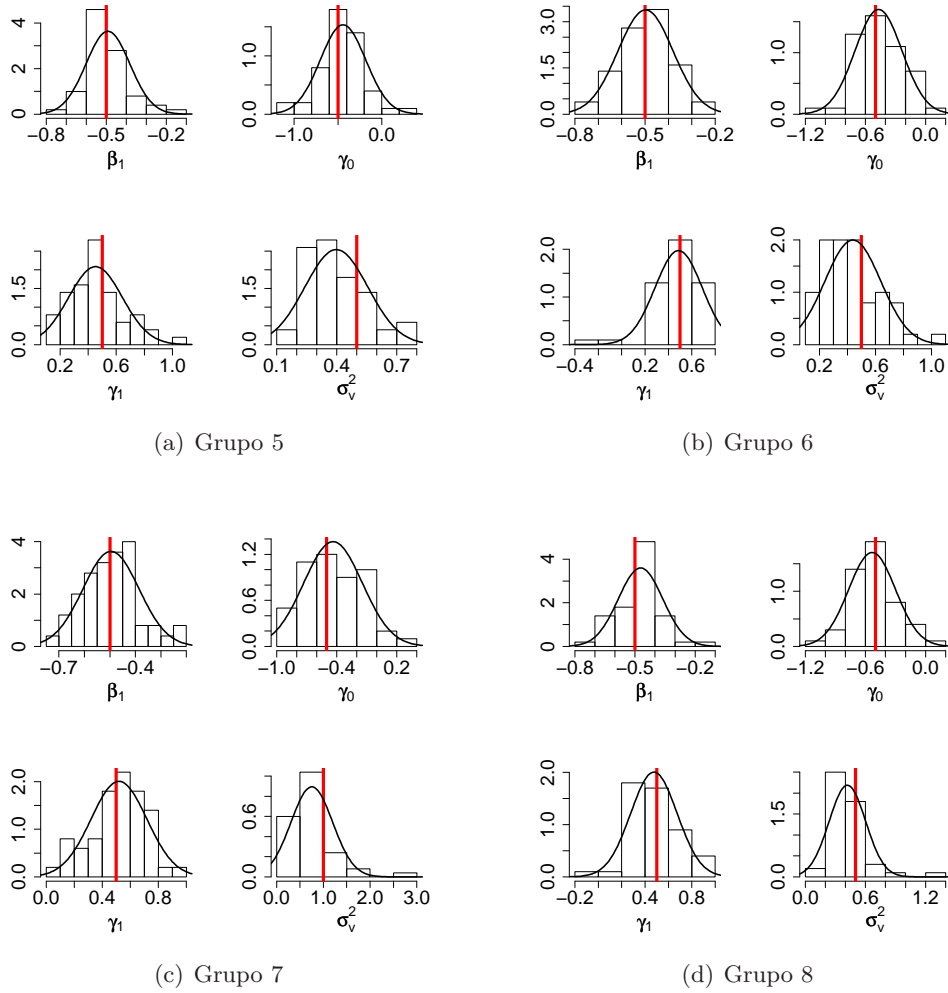


Figura 6.5: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – modelo de mistura – parte 2

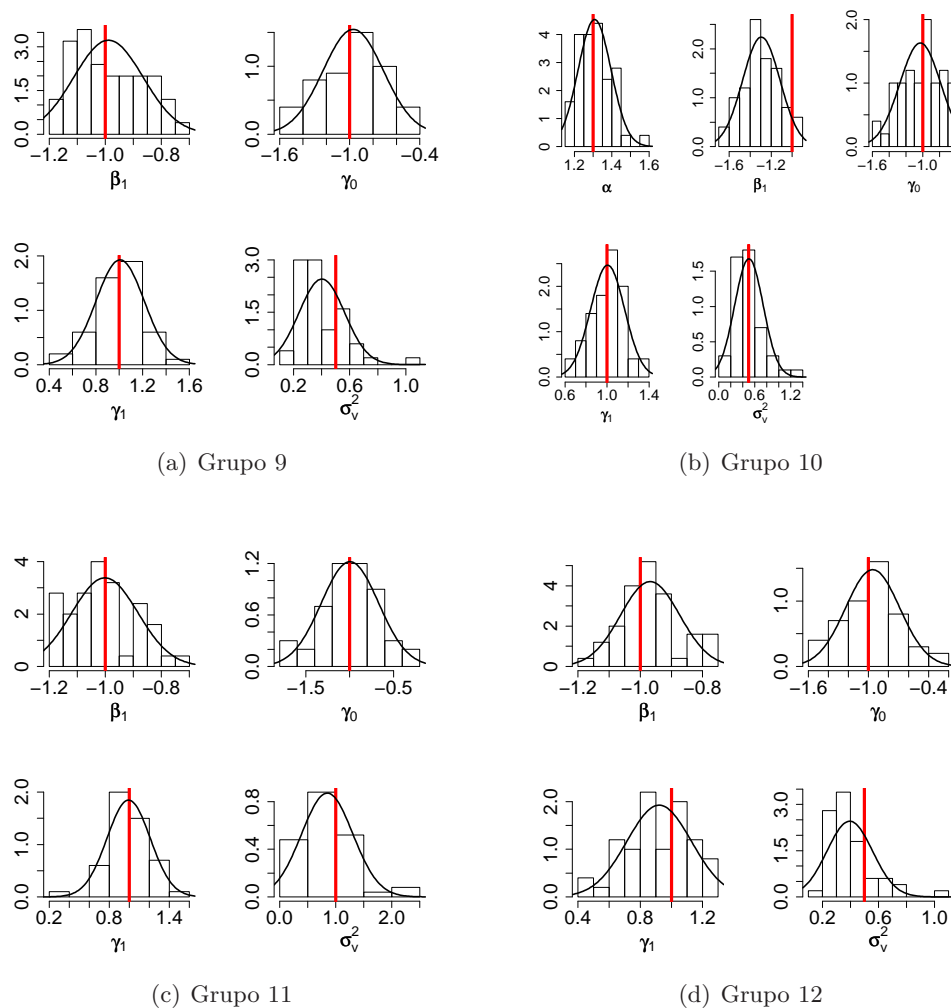


Figura 6.6: Histogramas das estimativas dos parâmetros estimados via Metropolis-Hastings – $n_k = 50$ – modelo de mistura – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, representados pelas retas verticais. Indicam também que os valores obtidos não estão muito longe da normalidade assintótica. Note ainda que os valores estimados para σ_v^2 estão melhores neste caso em que $n_k = 50$ do que no caso em que $n_k = 15$.

Para se avaliar se a convergência do modelo está adequada, calcula-se os valores de $\sqrt{\hat{R}}$, como definido ao final da seção 2.2. Para isso, considere $m = 3$ o número de cadeias utilizadas com valores iniciais distintos, em que cada uma apresenta $2n = 30000$ valores, ou seja, $n = 15000$. Foram calculados os valores de $\sqrt{\hat{R}}$ para simulações em cada um dos 12 grupos. Os resultados obtidos encontram-se na Tabela 6.4.

Tabela 6.4: Valores de $\sqrt{\hat{R}}$ – caso $n_k = 50$ – modelo de mistura

parâmetro	1	2	3	4	5	6
β_1	1,00026	1,00016	1,00078	1,00001	1,00007	1,00023
γ_0	1,00038	1,00232	1,00156	1,00030	1,00120	1,00003
γ_1	1,00028	1,00144	1,00127	1,00016	1,00035	1,00153
σ_V^2	1,00001	1,00025	1,00029	1,00047	1,00021	1,00004
parâmetro	7	8	9	10	11	12
β_1	1,00016	1,00085	1,00035	1,00037	1,00019	1,00107
γ_0	1,00062	1,00260	1,00539	1,00253	1,00007	1,00218
γ_1	1,00019	1,00024	1,00113	1,00044	1,00027	1,00061
σ_V^2	1,00006	1,00045	1,00027	1,00078	1,00046	1,00008

É perceptível que todos os valores estão bem próximos a 1 para todas as estimativas de interesse indicando que houve convergência da cadeia e que foi atingida a distribuição de cada estimador.

6.3 Simulação - abordagem clássica

Um estudo de simulação da parte clássica foi feito com o auxílio do programa Maple, versão 10. Os dados foram gerados com o programa R e são os mesmos utilizados na seção 6.2. Os comando utilizados encontram-se no apêndice A.6.

6.3.1 Caso $n_k = 15$

Os resultados obtidos estão na Tabela 6.5, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a diferença entre os valores verdadeiros e os médios-estimados.

Tabela 6.5: Resultado de 50 simulações via REML para $n_k = 15$ para o modelo de mistura

Grupo 1: ($\alpha = 1, 0; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4523	0,2011	-0,0477
γ_0	0	-0,0070	0,3056	0,0070
γ_1	0,5	0,3701	0,3607	0,1299
σ_V^2	0,5	0,6265	0,1435	-0,1265
Grupo 2: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4863	0,2467	-0,0137
γ_0	0	0,0459	0,2771	-0,0459
γ_1	0,5	0,6048	0,3917	-0,1048
σ_V^2	0,5	0,6129	0,1591	-0,1129
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,5075	0,2508	0,0075
γ_0	0	0,0102	0,3300	-0,0102
γ_1	0,5	0,5005	0,3805	-0,0005
σ_V^2	1	0,7787	0,2061	0,2213
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4576	0,2308	-0,0424
γ_0	0	-0,0066	0,2745	0,0066
γ_1	0,5	0,4977	0,3646	0,0023
σ_V^2	0,5	0,6609	0,1812	-0,1609
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4987	0,1912	-0,0013
γ_0	-0,5	-0,4590	0,3081	-0,0410
γ_1	0,5	0,5152	0,3227	-0,0152
σ_V^2	0,5	0,6351	0,1542	-0,1351
Grupo 6: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4651	0,1979	-0,0349
γ_0	-0,5	-0,4299	0,3402	-0,0701
γ_1	0,5	0,4771	0,3833	0,0229
σ_V^2	0,5	0,5741	0,1338	-0,0741

Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4385	0,2023	-0,0615
γ_0	-0,5	-0,3993	0,3304	-0,1007
γ_1	0,5	0,3652	0,4038	0,1348
σ_V^2	1	0,7950	0,1807	0,2050
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -0,5; \gamma_0 = -0,5; \gamma_1 = 0,5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4661	0,1828	-0,0339
γ_0	-0,5	-0,4828	0,2420	-0,0172
γ_1	0,5	0,4360	0,2923	0,0640
σ_V^2	0,5	0,6414	0,1858	-0,1414
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9991	0,2166	-0,0009
γ_0	-1	-0,8801	0,3234	-0,1199
γ_1	1	0,9175	0,3163	0,0825
σ_V^2	0,5	0,6741	0,1818	-0,1741
Grupo 10: ($\alpha = 1,3; \sigma_V^2 = 0,5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9434	0,2167	-0,0566
γ_0	-1	-0,9822	0,3285	-0,0178
γ_1	1	0,9606	0,4443	0,0394
σ_V^2	0,5	0,6097	0,1347	-0,1097
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9670	0,2316	-0,0330
γ_0	-1	-0,9280	0,3767	-0,0720
γ_1	1	0,9287	0,3252	0,0713
σ_V^2	1	0,8357	0,2166	0,1643
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0,5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9780	0,2071	-0,0220
γ_0	-1	-0,9679	0,3127	-0,0321
γ_1	1	1,0251	0,4179	-0,0251
σ_V^2	0,5	0,6081	0,1480	-0,1081

Note que, a diferença entre os valores médios estimados e os valores verdadeiros são relativamente pequenas, o que indica que os resultados obtidos via simulação estão bons.

As Figuras de 6.7 a 6.9 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

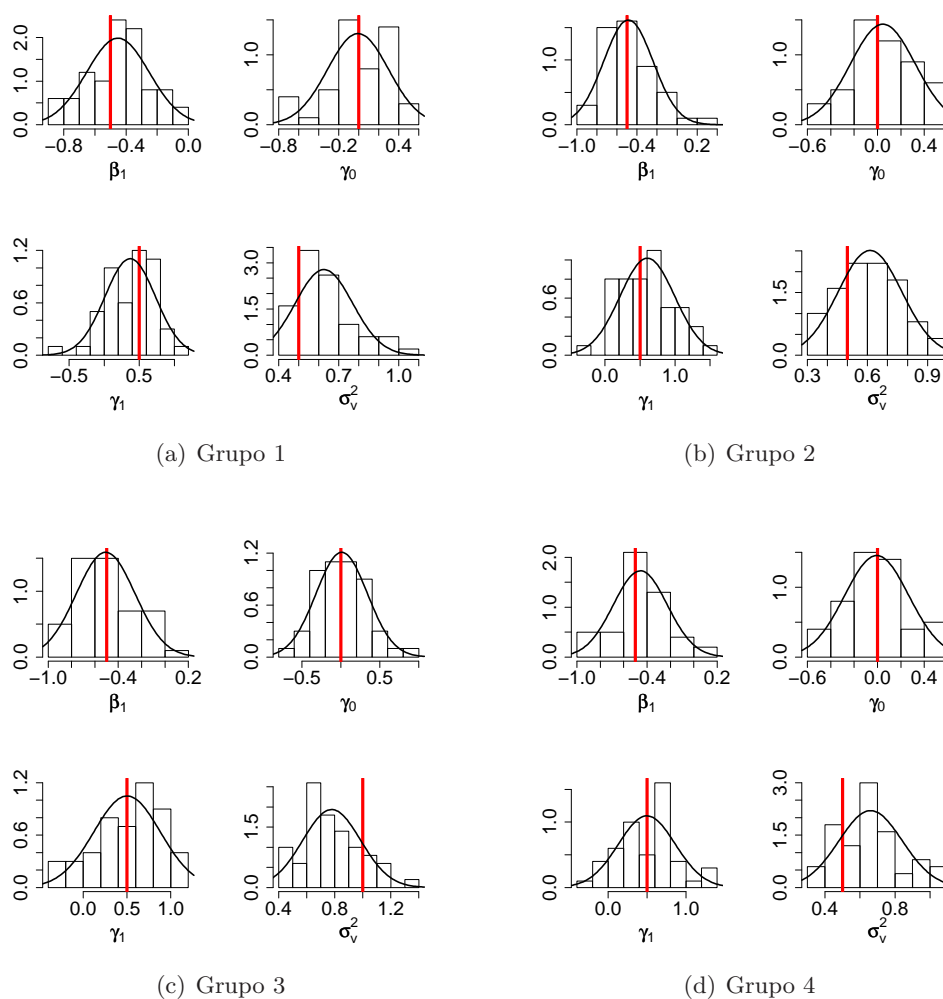


Figura 6.7: Histogramas das estimativas dos parâmetros estimados via REML - $n_k = 15$ - modelo de mistura - parte 1

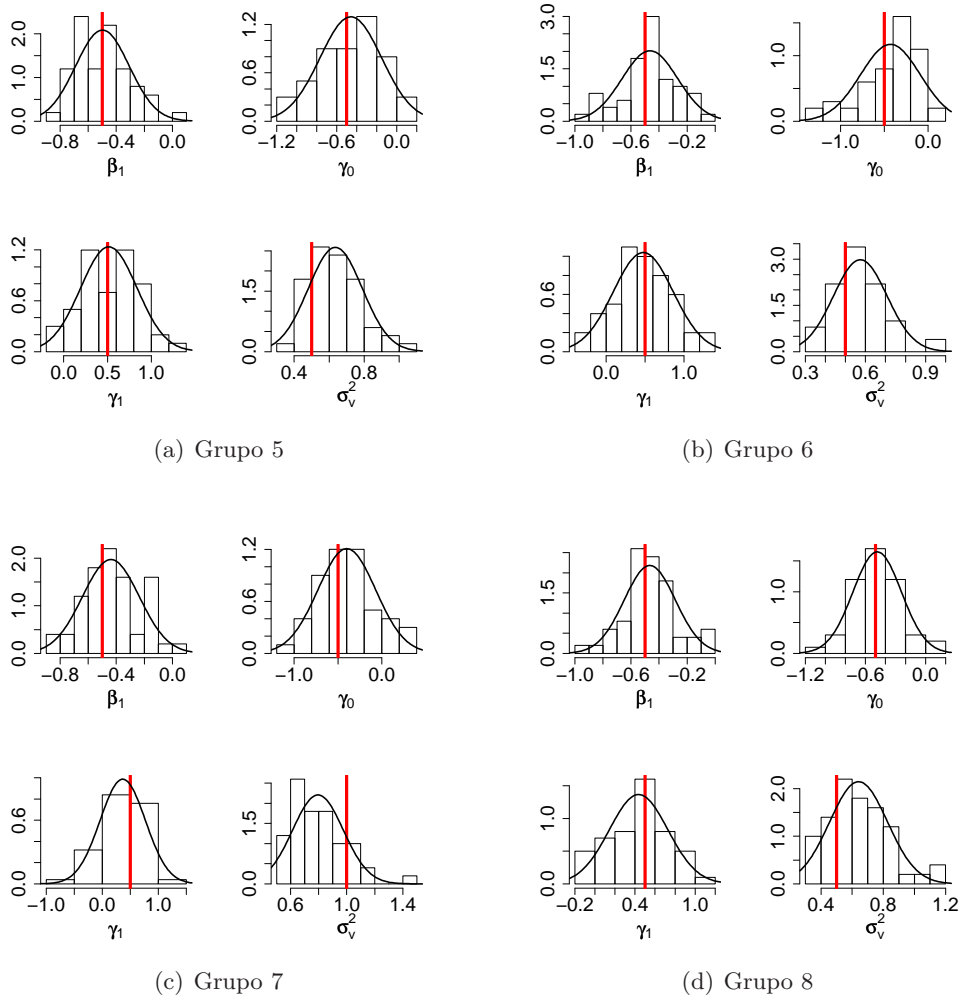


Figura 6.8: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – modelo de mistura – parte 2

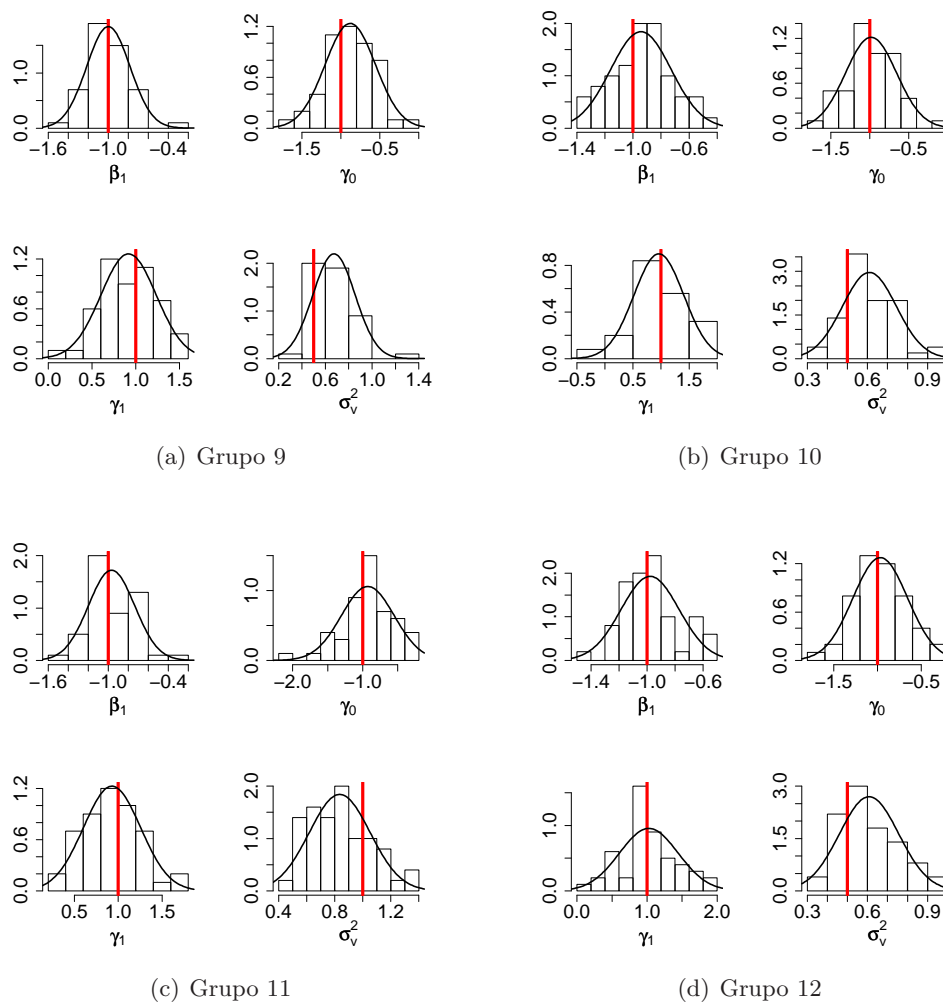


Figura 6.9: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 15$ – modelo de mistura – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, representados pelas retas verticais. Indicam também que os valores obtidos não estão muito longe da normalidade assintótica.

6.3.2 Caso $n_k = 50$

Os resultados obtidos estão na Tabela 6.6, em que a terceira coluna apresenta a média das estimativas obtidas nas 50 simulações. Na quarta coluna está o desvio padrão calculado a partir dos 50 valores de estimativas de cada parâmetro. E, a quinta coluna contém a diferença entre os valores verdadeiros e os médios-estimados.

Tabela 6.6: Resultado de 50 simulações via REML para $n_k = 50$ para o modelo de mistura

Grupo 1: ($\alpha = 1, 0; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4728	0,1135	-0,0272
γ_0	0	0,0060	0,2381	-0,0060
γ_1	0,5	0,4931	0,1923	0,0069
σ_V^2	0,5	0,5100	0,1360	-0,0100
Grupo 2: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4927	0,1368	-0,0073
γ_0	0	0,0713	0,2215	-0,0713
γ_1	0,5	0,4918	0,1933	0,0082
σ_V^2	0,5	0,5338	0,2160	-0,0338
Grupo 3: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4560	0,1133	-0,0440
γ_0	0	-0,0264	0,3328	0,0264
γ_1	0,5	0,4932	0,1836	0,0068
σ_V^2	1	0,8095	0,2842	0,1905
Grupo 4: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -0, 5; \gamma_0 = 0; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4667	0,1312	-0,0333
γ_0	0	-0,0069	0,2528	0,0069
γ_1	0,5	0,5060	0,1607	-0,0060
σ_V^2	0,5	0,5569	0,1603	-0,0569
Grupo 5: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4663	0,1049	-0,0337
γ_0	-0,5	-0,4360	0,2580	-0,0640
γ_1	0,5	0,4624	0,1889	0,0376
σ_V^2	0,5	0,5212	0,1569	-0,0212

Grupo 6: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4712	0,1119	-0,0288
γ_0	-0,5	-0,4638	0,2328	-0,0362
γ_1	0,5	0,4972	0,1973	0,0028
σ_V^2	0,5	0,5562	0,1985	-0,0562
Grupo 7: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4724	0,1068	-0,0276
γ_0	-0,5	-0,4281	0,2867	-0,0719
γ_1	0,5	0,5256	0,1957	-0,0256
σ_V^2	1	0,8070	0,3210	0,1930
Grupo 8: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -0, 5; \gamma_0 = -0, 5; \gamma_1 = 0, 5; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-0,5	-0,4487	0,1063	-0,0513
γ_0	-0,5	-0,5283	0,2304	0,0283
γ_1	0,5	0,4868	0,1968	0,0132
σ_V^2	0,5	0,5450	0,1728	-0,0450
Grupo 9: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9767	0,1186	-0,0233
γ_0	-1	-0,9859	0,2582	-0,0141
γ_1	1	1,0294	0,2051	-0,0294
σ_V^2	0,5	0,5422	0,1555	-0,0422
Grupo 10: ($\alpha = 1, 3; \sigma_V^2 = 0, 5; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9775	0,1191	-0,0225
γ_0	-1	-0,9732	0,2593	-0,0268
γ_1	1	0,9803	0,2235	0,0197
σ_V^2	0,5	0,5694	0,1981	-0,0694
Grupo 11: ($\alpha = 1; \sigma_V^2 = 1; \text{Censura} = 500; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9912	0,1143	-0,0088
γ_0	-1	-1,0061	0,3266	0,0061
γ_1	1	1,0072	0,2164	-0,0072
σ_V^2	1	0,9018	0,3301	0,0982
Grupo 12: ($\alpha = 1; \sigma_V^2 = 0, 5; \text{Censura} = 300; \beta_1 = -1; \gamma_0 = -1; \gamma_1 = 1; \lambda = 1$)				
Parâmetro	Valor verdadeiro	Média da estimativa	DP da estimativa	Verdadeiro - estimado
β_1	-1	-0,9572	0,0914	-0,0428
γ_0	-1	-0,9737	0,2682	-0,0263
γ_1	1	0,9450	0,2034	0,0550
σ_V^2	0,5	0,5374	0,1631	-0,0374

Note que, a diferença entre os valores médios estimados e os valores verdadeiros são menores do que um desvio-padrão, o que indica que os resultados obtidos via simulação estão precisos. Note também que esta diferença é menor do que no caso $n_k = 15$. Ou seja, como no caso bayesiano, o fato de termos mais indivíduos em cada clínica nos fornece valores estimados mais acurados.

As Figuras de 6.10 a 6.12 mostram os histogramas feitos a partir das estimativas calculadas nas 50 simulações, para cada um dos parâmetros.

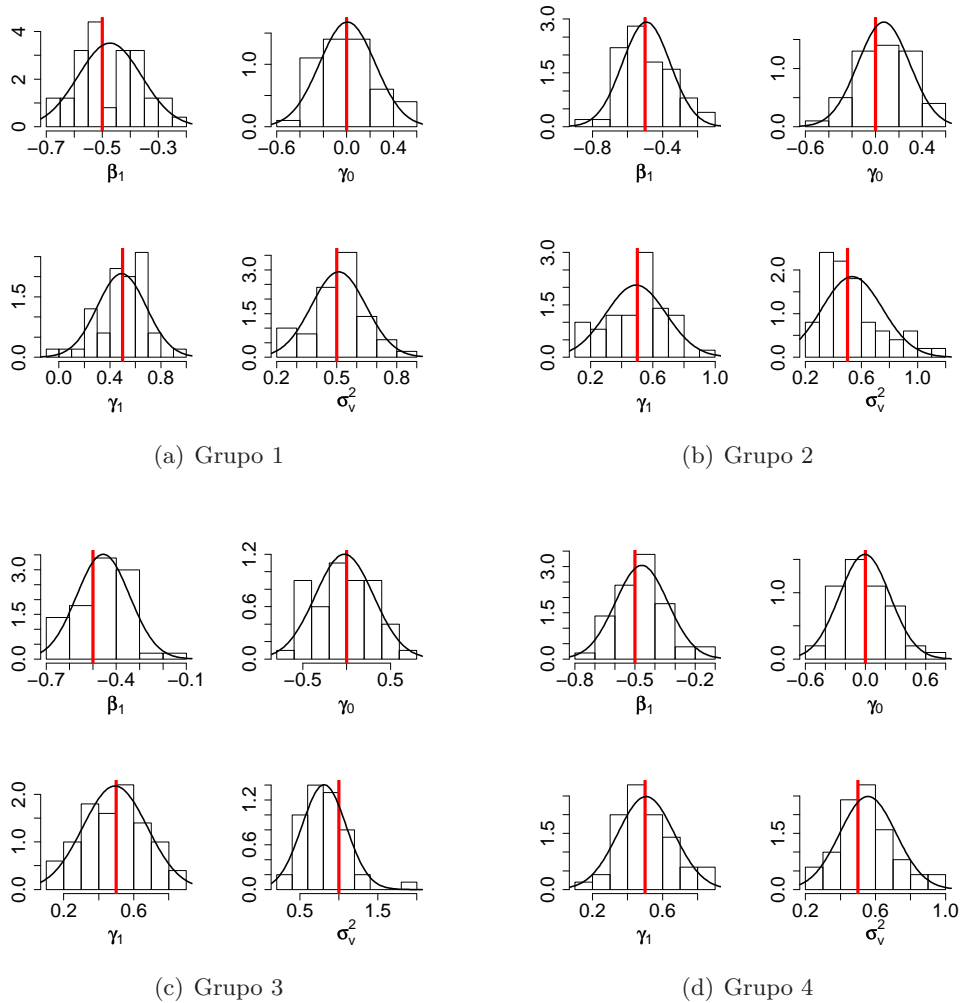


Figura 6.10: Histogramas das estimativas dos parâmetros estimados via REML - $n_k = 50$ - modelo de mistura - parte 1

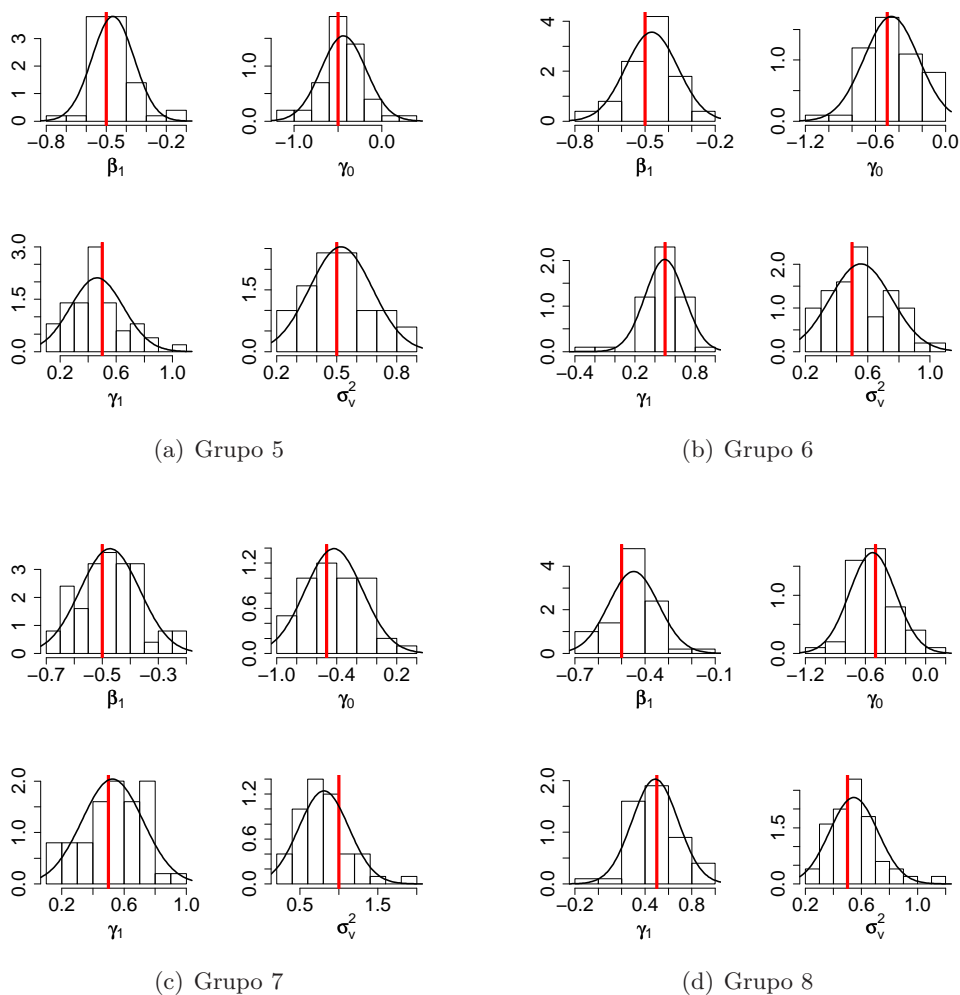


Figura 6.11: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – modelo de mistura – parte 2

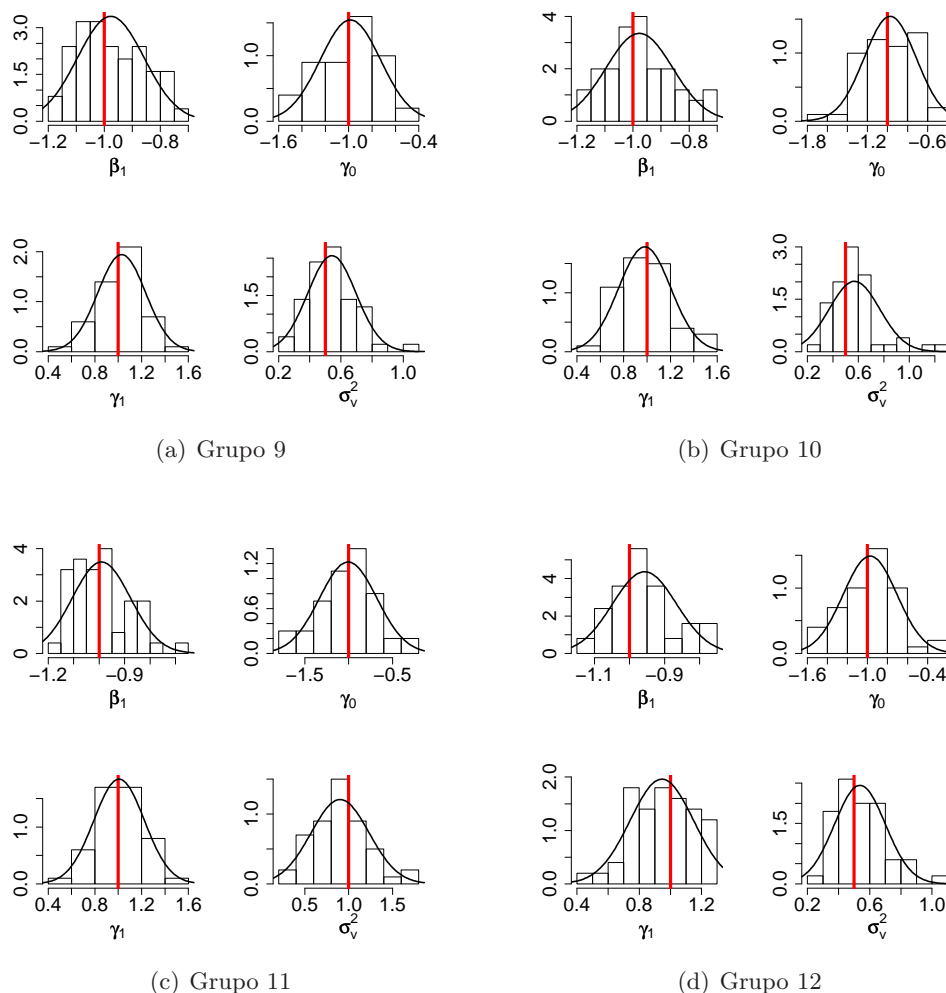


Figura 6.12: Histogramas das estimativas dos parâmetros estimados via REML – $n_k = 50$ – modelo de mistura – parte 3

Estes histogramas indicam que as estimativas obtidas estão em torno dos valores verdadeiros, representados pelas retas verticais. Indicam também que os valores obtidos não estão muito longe da normalidade assintótica. Este resultado é semelhante a todos os casos estudados.

6.4 Conclusões

Os resultados, tanto bayesiano quanto clássico, mostram que os valores obtidos nas estimativas estão próximos aos valores reais, o que sugere que as estimações utilizando ambos os métodos de inferência implementados foram adequadas. A única ressalva se deve quanto à estimação de σ_V^2 via Metropolis-Hastings, em que no caso $n_k = 15$, há uma subestimação e as estimativas não estão muito próximas dos valores verdadeiros.

Para β_1 , γ_0 e γ_1 , praticamente em todos os casos, a diferença entre a média do valor estimado e o

valor verdadeiro é menor do que um desvio-padrão, o que sugere que os resultados estão razoavelmente precisos. Os histogramas feitos com os valores estimados nas simulações ilustram que os valores verdadeiros, representados pelas retas verticais, estão dentro do esperado. E, que os valores têm distribuições aproximadamente normais.

Com relação à convergência das cadeias, na abordagem bayesiana, os valores de $\sqrt{\hat{R}}$ estão bem próximos a 1 para todas as estimativas de interesse indicando que houve a convergência das cadeias e que foi atingida a distribuição de cada estimador.

Nas simulações clássica e bayesiana, os valores estimados no caso $n_k = 50$ estão mais próximos dos valores verdadeiros do que no caso $n_k = 15$, mas em ambos os casos os valores estimados e verdadeiros diferem, quase sempre, por uma quantidade menor que um desvio-padrão. Note que, para σ_V^2 , com abordagem bayesiana, os resultados obtidos com $n_k = 50$ estão bem melhores do que os com $n_k = 15$.

Note, também, que, em ambas as abordagens, a diferença entre a média das estimativas e o valor verdadeiro para o caso $n_k = 50$ é menor do que no caso $n_k = 15$. Ou seja, o fato de termos mais indivíduos em cada clínica nos fornece valores estimados mais acurados.

Com relação aos desvios-padrão das estimativas, é perceptível que os valores estão, quase sempre, menores no caso $n_k = 50$ em comparação com $n_k = 15$. A exceção ocorre no caso de σ_v^2 em que essa situação se inverte, mas a acurácia do valor estimado é melhor em $n_k = 50$ do que em $n_k = 15$. Isso em ambas as inferências, bayesiana e clássica.

Capítulo 7

Análise de dados - modelo de mistura com efeitos aleatórios

Para ilustrar uma aplicação de uso do modelo de mistura com efeitos aleatórios nas frações de curados e não curados, Yau e Ng (2001) utilizaram um conjunto de dados em Kalbfleisch e Prentice (1980, pp.225-229). Tais dados são parte de um estudo feito pelo *Radiation Therapy Oncology Group* nos Estados Unidos.

A seguir, é apresentada a utilização deste mesmo conjunto de dados no modelo apresentado no Capítulo 5, para fração de cura com efeitos aleatórios na fração de curados, a partir do modelo de mistura. Os dados foram descritos na seção 4.1.

7.1 Inferência bayesiana via Metropolis-Hastings

Uma primeira análise desses dados foi feita com a utilização do programa Winbugs (versão 1.4). Os comandos utilizados encontram-se no apêndice A.5. Os resultados obtidos estão apresentados na Tabela 7.1, em que a coluna de média apresenta o valor da média entre os valores amostrados, DP apresenta o desvio-padrão e erro MC está mostrando os valores obtidos pelo erro padrão de Monte Carlo para a média.

Tabela 7.1: Resultados – câncer na orofaringe – modelo de mistura com efeitos aleatórios – Metropolis-Hastings

Resultados das posteriores, baseados em 30000 amostras								
parâmetro	média	DP	erro MC	2,50%	mediana	97,50%	início	amostras
α	0,67740	0,03343	0,00029	0,61040	0,67790	0,74150	4001	78000
β_1	-0,29330	0,27830	0,00249	-0,86610	-0,28540	0,22860	4001	78000
γ_0	-5,63400	1,60300	0,01496	-9,27800	-5,42700	-3,13000	4001	78000
γ_1	-1,43000	2,52100	0,02185	-6,73700	-1,26200	3,07100	4001	78000
λ	0,03366	0,00586	0,00005	0,02404	0,03303	0,04679	4001	78000
σ_V^2	0,32590	0,22260	0,00211	0,11350	0,26870	0,89020	4001	78000

Note que σ_V^2 é muito pequeno o que indicaria que a componente de variância não é significativamente diferente de zero, de modo que não há efeito de clínica na parte da população de curados.

88CAPÍTULO 7. ANÁLISE DE DADOS - MODELO DE MISTURA COM EFEITOS ALEATÓRIOS

Este resultado já tinha sido apontado na análise pelo modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, apresentado na seção 4.2.

O software Winbugs, neste exemplo, executou o algoritmo Metropolis-Hastings.

A Figura 7.1 mostra os histogramas feitos com os dados gerados pelo programa Winbugs.

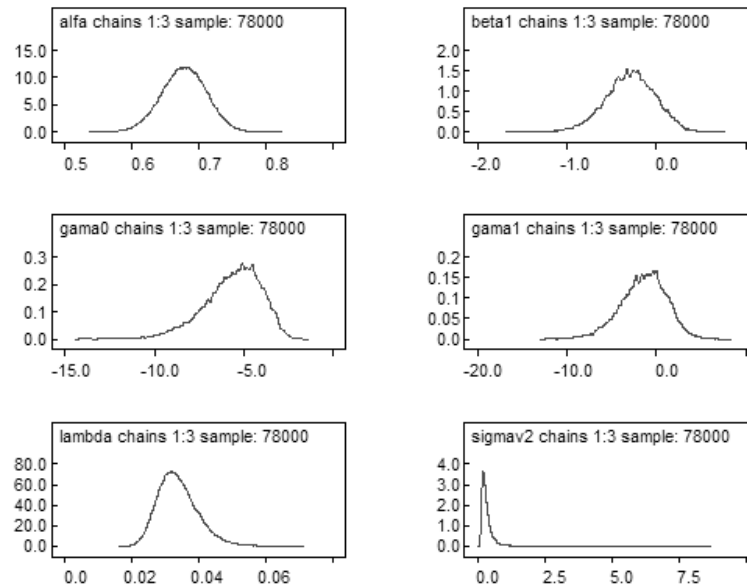


Figura 7.1: Histogramas – câncer na orofaringe – modelo de mistura com efeitos aleatórios – Metropolis-Hastings

A Figura 7.2 mostra os gráficos dos valores de $\sqrt{\hat{R}}$ (Gelman e Rubin, 1992), como definidos na seção 2.2.

Note que para todos os parâmetros, os valores estão tendendo ao valor 1, o que indica que as cadeias estão convergindo.

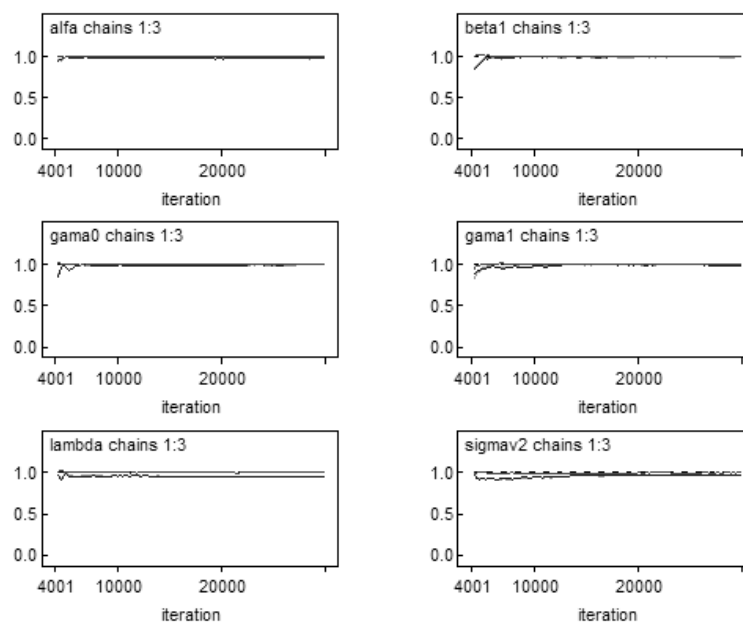


Figura 7.2: $\sqrt{\hat{R}}$ – câncer na orofaringe – modelo de mistura com efeitos aleatórios – Metropolis-Hastings

7.2 Análise clássica

O conjunto de dados, apresentado na seção 4.1, foi novamente utilizado sendo feita a abordagem relatada na seção 5.3. Desta vez, o software utilizado para a compilação dos dados foi Maple versão 10. Os comandos estão apresentados no apêndice A.6.

Os resultados obtidos estão na Tabela 7.2.

Tabela 7.2: Resultados – câncer na orofaringe – modelo de mistura com efeitos aleatórios – REML

parâmetros	estimativa	desvio-padrão
α	0,1412	0,0206
λ	0,4717	0,4356
β_1	-0,5736	0,4356
γ_0	-0,8482	0,4154
γ_1	-2,7615	7,1688
σ_V^2 (BLUP)	0,0468	–
σ_V^2 (REML)	0,2872	0,3674

efeito de clínica	preditores (Blup)
V_1	0,2912
V_2	-0,2933
V_3	0,1505
V_4	0,0461
V_5	-0,2794
V_6	0,0849

Note que os valores obtidos para σ_V^2 , apresentados nas Tabelas 7.1 e 7.2 estão bem próximos. Em ambos, tem-se valores pequenos para a variância dos efeitos aleatórios, indicando que pode não haver diferenças entre as clínicas.

As variâncias para α e λ estão subestimadas. Valores melhores para os desvios-padrão destes dois parâmetros podem ser obtidos por bootstrap (Efron, 1979; Efron, 1982).

7.3 Conclusões

Os resultados obtidos, para σ_V^2 com a estimação bayesiana, via Metropolis-Hastings estão próximos aos obtidos com a abordagem clássica, via REML.

É nítida uma diferença de estimativas para alguns parâmetros, fato que pode ter ocorrido devido às características dos dados utilizados, como, por exemplo, baixo número de falhas em algumas clínicas, poucas observações com estágio avançado do tumor.

Os resultados obtidos através do modelo de mistura com efeitos aleatórios para a variância dos efeitos aleatórios estão coerentes com os obtidos com o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, em termos da análise de efeito de clínica na fração de curados.

Capítulo 8

Conclusões

Os modelos propostos nos Capítulos 2 e 5 foram analisados por duas perspectivas: bayesiana e clássica. Para avaliarmos as técnicas estudadas, foram feitos estudos de simulação, apresentados nos Capítulos 3 e 6. Para exemplificar o uso dos modelos, foram feitas análises com dados reais de um estudo de câncer na orofaringe, apresentados na seção 4.1 e cujos resultados encontram-se nos Capítulos 4 e 7.

8.1 Resultados e conclusões

Em todos os casos simulados para o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, os resultados tanto para as estimativas dos parâmetros quanto para os desvios-padrão foram muito bons. Em quase todas as situações, a diferença entre os valores médios estimados e os valores verdadeiros foram menores do que um desvio-padrão, o que indica que os resultados obtidos via simulação, para os parâmetros estudados, estão bons.

Além disso, tal diferença apresentou-se menor nos casos em que tínhamos mais indivíduos em cada clínica (caso $n_k = 50$ em comparação com $n_k = 15$). Tanto no caso bayesiano, quanto no clássico, o fato de termos mais indivíduos em cada clínica nos fornece valores estimados mais acurados, como esperado.

Com relação às estimativas para os desvios-padrão, as simulações mostraram que os desvios-padrão das médias estão próximos aos desvios-médio em praticamente todos os casos. Isso é um bom indício de que os desvios estão sendo bem estimados pelas segundas-derivadas da log-verossimilhança, ou seja, pela informação de Fisher, no caso clássico. Somente para o parâmetro σ_V^2 esta diferença é um pouco maior. Mas, este é o parâmetro cuja variância está sendo estimada por REML.

Esta proximidade entre os valores de desvio da média e médio e entre as médias das estimativas e os valores verdadeiros são um indício de que as técnicas estudadas estão adequadas para a análise do modelo proposto no Capítulo 2.

Ainda no modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, é perceptível que os valores dos desvios-padrão das estimativas estão, quase sempre, menores no caso $n_k = 50$ em comparação com

$n_k = 15$. A exceção ocorre no caso de σ_V^2 em que essa situação se inverte, mas a acurácia do valor estimado é melhor em $n_k = 50$ do que em $n_k = 15$. Isso em ambas as inferências bayesiana e clássica.

Os histogramas das estimativas indicam que os valores parecem apresentar distribuições normal. No caso clássico, este resultado é condizente com uma análise via estimadores de máxima verossimilhança.

É interessante notar que os resultados obtidos via abordagens clássica e bayesiana estão próximos, ou seja, ambos estão apresentando estimativas próximas para o modelo em estudo.

Já com relação ao modelo de mistura com efeitos aleatórios, quando analisados por Metropolis-Hastings, há indícios de que os valores estimados para σ_V^2 estão sendo subestimados, isso é mais aparente no caso $n_k = 15$, em que, em alguns casos, o resultado não está bom. No caso $n_k = 50$, os resultados estão melhores. Para os demais parâmetros, os resultados das simulações estão satisfatórios.

Com relação à abordagem clássica, tem-se que os valores estimados para σ_V^2 estão, em média, subestimados quando o valor verdadeiro é 1, e superestimados, quando seu valor verdadeiro é 0,5. Essa diferença é mais aparente no caso $n_k = 15$, mas ainda assim os resultados estão satisfatórios. Como esperado, os resultados para todos os parâmetros estão melhores no caso $n_k = 50$ em que todos os grupos de parâmetros simulados apresentam resultados bons.

Levando-se em conta o modelo de mistura com efeitos aleatórios, o melhor resultado apresentado nas simulações foi obtido através da abordagem clássica, especialmente quando o número de indivíduos por clínica é maior.

Comparando as simulações tanto para o modelo de mistura quanto para o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios, percebe-se que os resultados estão melhores quando se tem mais indivíduos em cada clínica. De todos os casos estudados via simulação, o que apresentou o pior resultado foi o modelo de mistura com efeitos aleatórios, quando analisado por Metropolis-Hastings. Nos demais casos, os resultados estão dentro do esperado.

8.2 Perspectivas futuras

Devido às características do modelo estudado no Capítulo 2, os efeitos aleatórios não puderam entrar separadamente nos curados e não-curados, para um mesmo modelo, como feito em Yau e Ng (2001). Por uma questão de não identificabilidade do modelo, foram colocados efeitos aleatórios no modelo de Chen-Ibrahim-Sinha somente para a parcela de curados. Em um próximo passo, analisaremos o caso em que o efeito aleatório será estudado para toda a população.

Com relação à abordagem clássica uma alternativa aos BLUP e REML seria utilizar o algoritmo EM para a estimação. O novo resultado poderá ser utilizado para compararmos os resultados das estimativas. Esta nova abordagem poderá ser utilizada tanto no modelo de mistura quanto no modelo de Chen-Ibrahim-Sinha com efeitos aleatórios.

Com relação à abordagem bayesiana uma alternativa ao uso do programa BUGS seria a utilização do pacote UMACS do R ou a programação direta do Metropolis-Hastings. Os resultados poderão ser utilizados para compararmos a acurácia dos diversos softwares.

Para o modelo de mistura com efeitos aleatórios, em um próximo passo, pretendemos utilizar bootstrap de Efron (Efron, 1979; Efron, 1982) para a obtenção dos valores dos desvios-padrão para α e λ .

Ainda no modelo de mistura, pode ser feita uma análise semi-paramétrica para a estimação da função de risco basal $h_{20}(t)$, como alternativa ao uso da distribuição Weibull que adotamos.

Como alternativa ao uso dos resultados das simulações para a comparação dos modelos de mistura e de Chen-Ibrahim-Sinha com efeitos aleatórios, pretendemos estudar técnicas de comparação de modelos. Ressalto que estas comparações têm que ser feitas com muito cuidado, visto que os parâmetros têm interpretações diferentes nos dois modelos.

Há dois artigos de Cancho; Rodrigues & Castro (2008) e de Castro; Cancho & Rodrigues (2007), em desenvolvimento, em que o modelo de Chen-Ibrahim-Sinha é trabalhado utilizando distribuição binomial negativa no lugar da distribuição de Poisson. Pretendemos analisar como fica este modelo quando inserimos efeitos aleatórios.

Finalmente, pretendemos analisar como estes dois modelos apresentados aqui poderão ser analisados com a "teoria unificada para modelos de longa duração", que foi publicada em uma versão por Yin e Ibrahim (2005) e, em um outro enfoque, submetida para publicação por Rodrigues e Louzada-Neto (2006).

Apêndice A

Programas

A.1 Gerando os dados para a simulação - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios

A seguir está descrito o procedimento que foi feito para que o programa R (versão 1.19.2.3) gerasse os dados a serem utilizados na simulação para o modelo de Chen-Ibrahim-Sinha com efeitos aleatórios.

1. Foram atribuídos valores para os parâmetros, para o número de clínicas, para o número de pacientes por clínica e para a censura, conforme apresentados no Capítulo 3;
2. para as covariáveis X , foram atribuídos, aleatoriamente, valores 0 e 1, na proporção 2:1;
3. os efeitos aleatórios V_k , para cada uma das k clínicas foram sorteados de uma distribuição normal, com média 0 e variância σ_V^2 ;
4. para cada indivíduo de cada uma das clínicas, foram sorteados valores para N_{jk} , a partir da distribuição de Poisson de média $\exp(\gamma_0 + \gamma_1 * x_{jk} + V_k)$;
5. para cada célula de 0 a N_{jk} , foi armazenado um valor para z_{jk} , sorteado de uma Weibull, com parâmetros α e $\exp(-\beta * x_{jk})$;
6. T_{jk} armazenou o valor mínimo entre todos os z_{jk} ;
7. se o valor de censura fosse menor do que T_{jk} então T_{jk} passa a ter o valor da censura e, neste caso, o indicador de censura recebe o valor 0.

A seguir estão os principais comandos, feitos para o programa R, que foi utilizado para gerar os dados. A variável “listpar” recebeu os valores dos parâmetros.

```
for (ipar in 1:totallinhaslistpar) {
```

```

alfa <- listpar[ipar,1]
beta <- listpar[ipar,2] # lambda=beta x + u
                        #no winbugs é lambda = 1/lambda e aí o beta1=-beta
g <- matrix(0,2,1)
g <- c(listpar[ipar,3],listpar[ipar,4]) # gama
sigmav <- sqrt(listpar[ipar,5]) #ivv = 1/(sigmav)^2

# fixos
M<- 10 # clínicas
n<- 15 # pacientes
C<- 500 # censura à direita

y<- matrix(0,M,n)
x<- matrix(0,M,n)
V<- matrix(0,M)
N<- matrix(0,M,n) # número de celulas doentes
tempo<- matrix(0,M,n) # tempo que pode ser censurado
delta<- matrix(1,M,n) # indicador de falha
theta<- matrix(0,M,n) # parâmetro da Poisson - calcular prob cura exp(-theta)

for (i in 1:M) {
  V[i]<-rnorm(1,0,sigmav)
  trat<-c(rep(0,10),rep(1,5))
  trat<-sample(trat, 15, replace = FALSE, prob = NULL)
  x[i,]<-trat
}

for(i in 1:M) {
for(j in 1:n) {
  print(c(ipar,"i=",i," j=",j,M,n,ipar));
  theta[i,j]<-exp(g[1]+g[2]*x[i,j]+V[i])
  N[i,j]<- rpois(1, theta[i,j] )
  lambda<- exp(-(beta*x[i,j]))
  for(cel in 0:N[i,j]) {
    z<-rweibull(1,alfa,lambda)
    if (cel==1) tempo[i,j]<- z
    else tempo[i,j]<- min(z,tempo[i,j])
  }
  if (tempo[i,j] > C | N[i,j]==0) {

```


A.2. COMANDOS EM WINBUGS - MODELO DE CHEN-IBRAHIM-SINHA COM EFEITOS ALEATÓRIOS

```
tempo[i,j]<- C
delta[i,j]<- 0
}
}
}}
```

A.2 Comandos em Winbugs - modelo de Chen-Ibrahim-Sinha com efeitos aleatórios

Nesta seção estão apresentados os comandos para o modelo 2.2 utilizado no programa Winbugs (versão 1.4), relatados na seção 4.2 e na simulação apresentada na seção 3.2.

```
Model{
c <- 100000
for (i in 1:M) {
# random effects
v[i] ~ dnorm(0,ivv)
for (j in 1:Nk[1,i]) {
# likelihood
alam[i,j]<- (beta1*x[i,j])
s[i,j] <- exp(-exp(alam[i,j])*pow(y[i,j],1))
xg[i,j]<- gama0+gama1*x[i,j]+v[i]
theta[i,j]<- exp(xg[i,j])
L1[i,j]<- -theta[i,j]*(1-s[i,j])
L2[i,j]<- d[i,j]*(xg[i,j]+log(1)+alam[i,j]+(1-1)*log(y[i,j]) - exp(alam[i,j])*pow(y[i,j],1))
L3[i,j]<- ( (-pow(v[i],2)*0.5*ivv) + (0.5*log(ivv)) )/Nk[1,i]
zeros[i,j]<- -0
mu[i,j]<- -(L1[i,j]+L2[i,j]+L3[i,j]) + c
zeros[i,j] ~ dpois(mu[i,j])
}
}
sigmav <- 1/sqrt(ivv)
sigmav2 <- 1/ivv
# priors
beta1 ~ dnorm(0, 0.1)
gama0 ~ dnorm(0, 0.1)
gama1 ~ dnorm(0, 0.1)
ivv ~ dgamma(1,1)
}
```

Os dados da Tabela 4.1 entraram do seguinte modo:


```

> l1:=sum(f1,K=1..M);
> f2:=(gama0+gama1*x[j,K]+V[K])*nu[j,K]-exp(gama0+gama1*x[j,K]+V[K]) +
exp(gama0+gama1*x[j,K]+V[K])*exp(-(y[j,K])^alfa*exp(x[j,K]*beta1))+
nu[j,K]*(ln(alfa)+(alfa - 1)*ln(y[j,K])+x[j,K]*beta1-
(y[j,K])^alfa*exp(x[j,K]*beta1));
> l2:=sum(sum(f2,j=1..n[K]),K=1..M):
Aqui abaixo está a definição da log-verossimilhança completa, já somada em N
> l:=l1+l2:
Avisando que V e dV são vetores
> V:=array(1..M): dV := array(1..M):
Matriz com parâmetros iniciais e a serem atualizados
> phi := matrix(p,2,[10,1,10,-5,10,-0.5,10,+1.5,10,0.5,10,0.5,10,0.5,
10,-0.5,10,-0.5,10,-0.5]): sigmav2:=0.5:
Definindo vetor de derivadas, vetor escore
> U := matrix(p,1,[diff(1,alfa) , diff(1,beta1), diff(1,gama0),
diff(1,gama1), diff(1,V[1]), diff(1,V[2]), diff(1,V[3]), diff(1,V[4]),
diff(1,V[5]), diff(1,V[6])]):
> U2 := matrix(p,1):
Definindo a matriz hessiana
> H := hessian(1,[alfa,beta1,gama0,gama1,V[1],V[2],V[3],V[4],V[5],V[6]]):
> H2 := matrix(p,p);
Calculando a norma da diferença de phi em dois estágios consecutivos
> difphi:=evalf(norm(col(phi,1)-col(phi,2),2));
> passo := 0.1;
> for z from 1 to M do dV[z] := diff(1, V[z]); end do:
Programando o "laço" para a iteração
> for t from 1 to 1000 while (difphi > 0.001) do
for a from 1 to p do
phi[a,1] := phi[a,2]:
end do;

alfa:=phi[1,1];beta1:=phi[2,1];gama0:=phi[3,1];gama1:=phi[4,1];
V[1]:=phi[5,1];V[2]:=phi[6,1];V[3]:=phi[7,1];V[4]:=phi[8,1];
V[5]:=phi[9,1];V[6]:=phi[10,1];

for a from 1 to p do
U2[a,1] := evalf(subs({alfa=phi[1,1],beta1=phi[2,1],gama0=phi[3,1],
gama1=phi[4,1],sigmav2=sigmav2,V[1]=phi[5,1],V[2]=phi[6,1],
V[3]=phi[7,1],V[4]=phi[8,1],V[5]=phi[9,1],V[6]=phi[10,1]},U[a,1]));
end do;

for a from 1 to p do

```

```

for b from 1 to p do
H2[a,b] := evalf(subs({alfa=phi[1,1],beta1=phi[2,1],gama0=phi[3,1],
gama1=phi[4,1],sigmav2=sigmav2,V[1]=phi[5,1],V[2]=phi[6,1],
V[3]=phi[7,1],V[4]=phi[8,1],V[5]=phi[9,1],V[6]=phi[10,1]},H[a,b]));
end do;
end do;

for a from 1 to p do
phi[a,2] := phi[a,1] - passo*(evalm(inverse(H2) &* (U2))[a,1]);
end do;

sigmav2:= sum('V[z]^2','z'=1..M) / M;

# atualiza a condição de parada para a convergência
difphi:=evalf(norm(col(phi,1)-col(phi,2),2));
print(V, sigmav2);
print("t=",t);
end do;

Lendo da segunda coluna de phi quem são os "chapéus", as estimativas de  $\alpha$ ,  $\beta_1$ ,  $\gamma_0$ ,  $\gamma_1$  e os
BLUP para  $\mathbf{V}$ . Lendo também o valor estimado para  $\sigma_V^2$  por este método
> alfachapeu:=phi[1,2]; beta1chapeu:=phi[2,2]; gama0chapeu:=phi[3,2];
gama1chapeu:=phi[4,2]; V[1]:=phi[5,2]; V[2]:=phi[6,2]; V[3]:=phi[7,2];
V[4]:=phi[8,2]; V[5]:=phi[9,2]; V[6]:=phi[10,2];
sigmav2chapeublup:= sum('V[z]^2','z'=1..M) / M;
Atualizando a matriz hessiana para os valores dos chapéus
> for a from 1 to p do
for b from 1 to p do
H2[a,b] := evalf(subs({alfa=phi[1,2],beta1=phi[2,2],gama0=phi[3,2],
gama1=phi[4,2],sigmav2=sigmav2chapeublup,V[1]=phi[5,2],
V[2]=phi[6,2],V[3]=phi[7,2],V[4]=phi[8,2],V[5]=phi[9,2],
V[6]=phi[10,2]},H[a,b]));
end do;
end do;

REML: estimar o  $\sigma_V^2$  de modo menos viesado do que pelo método acima
Selecionando a parte da matriz hessiana correspondente a V
> A33:=matrix(6,6,[]);
> H2inv:=inverse(H2);
> A33:=evalm(-1*matrix([col(matrix([col(H2inv,5..10)]),5..10)]));
> sigmav2chapeu:=1/M*evalm( trace(A33) + (transpose(V)&*V) );
Variância de sigmav2chapeu, pelo REML
> varsigma2:=1/(2*(1/(sigmav2chapeu)^2*(M-2/sigmav2chapeu*trace(A33))+
1/(sigmav2chapeu)^4*trace(evalm(A33^2))));
Calculando as variâncias de  $\alpha$ ,  $\beta_1$ ,  $\gamma_0$ ,  $\gamma_1$ 

```

```
> var:=evalm(-1*matrix([col(matrix([col(H2inv,1..4)]),1..4)]));
```

A.4 Gerando os dados para a simulação - modelo de mistura com efeitos aleatórios

A seguir está descrito o procedimento que foi feito para que o programa R (versão 1.19.2.3) gerasse os dados a serem utilizados na simulação para o modelo de mistura com efeitos aleatórios.

1. Foram atribuídos valores para os parâmetros, para o número de clínicas, para o número de pacientes por clínica e para a censura, conforme apresentados no Capítulo 6;
2. para as covariáveis X , foram atribuídos, aleatoriamente, valores 0 e 1, na proporção 2:1;
3. os efeitos aleatórios V_k , para cada uma das k clínicas foram sorteados de uma distribuição normal, com média 0 e variância σ_V^2 ;
4. para cada indivíduo de cada uma das clínicas, foram armazenados valores para π_{jk} , dados por $1/(1 + \exp(\gamma_0 + \gamma_1 * x_{jk} + V_k))$;
5. para decidir se o indivíduo j da clínica k seria curado ou não, foi feito um sorteio a partir de uma distribuição binomial com probabilidade π_{jk} de o indivíduo ser não curado;
6. os indivíduos sorteados com não-curados receberam T_{jk} igual a censura;
7. os indivíduos sorteados com curados receberam T_{jk} um valor $(-\ln(1 - u)/\lambda * \exp(\eta_{jk}))^{1/\alpha}$ em que u foi sorteado da distribuição uniforme (0,1).

A seguir estão os principais comandos, feitos para o programa R, que foi utilizado para gerar os dados. A variável “listpar” recebeu os valores dos parâmetros.

```
# fixos
alfa <- listpar[ipar,1]           # alfa
sigmav <- sqrt(listpar[ipar,5])  # ivv = 1/(sigmav)^2
beta <- -listpar[ipar,2]         # lambda=beta x + u
                                #no winbugs é lambda = 1/lambda e ai o beta1=-beta
g <- matrix(0,2,1)
g <- c(listpar[ipar,3],listpar[ipar,4]) # gama
C <- listpar[ipar,6]             # censura à direita (300 ou 500)
lambda <- listpar[ipar,7]       # lambda

x<- matrix(0,M,nindiv)
V<- matrix(0,M)
```

```

tempo<- matrix(0,M,nindiv)           # tempo que pode ser censurado
delta<- matrix(1,M,nindiv)          # indicador de falha
eta <- matrix(0,M,nindiv)
pi <- matrix(0,M,nindiv)

for (i in 1:M) {
  V[i]<-rnorm(1,0,sigmav)
  trat<-c(rep(0,round(nindiv*2/3)),rep(1,nindiv-round(nindiv*2/3)))
  trat<-sample(trat, nindiv, replace = FALSE, prob = NULL)
  x[i,]<-trat
}

for(i in 1:M) {
for(j in 1:nindiv) {
  eta[i,j] <- beta * x[i,j]

  # gerar quem é curado e não curado
  pi[i,j] <- 1/(1+exp(g[1]+g[2]*x[i,j]+V[i])) #para não-curados; para os curados é 1-pi
  escolhanaocurado <- rbinom(1, 1, pi[i,j]) # gera 1 para não-curado com prob pi
  if (escolhanaocurado == 0) { #ou seja, é curado, então é censura à direita
    tempo[i,j]<- C
    delta[i,j]<- 0
  }
  if (escolhanaocurado == 1) {
    u <- runif(1, 0, 1)
    tempo[i,j]<- ( -log(1-u) / lambda*exp(eta[i,j]) )^(1/alfa)
  }
}
}
}

```

A.5 Comandos em Winbugs - modelo de mistura com efeitos aleatórios

Nesta seção estão apresentados os comandos utilizados na abordagem bayesiana para o modelo de mistura com efeitos aleatórios utilizado no programa Winbugs (versão 1.4), relatados na seção 5.2 e na simulação apresentada na seção 6.2.

```

Model{
c <- 100000 for (i in 1:M) {
# random effects

```

```

v[i] ~ dnorm(0,ivv)
for (j in 1:Nk[1,i]) {
# likelihood para modelo de mistura
eta[i,j] <- - beta1*x[i,j]
pi[i,j] <- 1/(1+exp(gama0+gama1*x[i,j]+v[i]))
s2[i,j] <- - exp(-exp(eta[i,j])/invlambda)*pow(y[i,j],alfa)
s[i,j] <- - 1-pi[i,j]+pi[i,j]*s2[i,j]
h20[i,j] <- - alfa*pow(y[i,j],alfa-1)/invlambda
s20[i,j] <- - exp(-pow(y[i,j],alfa)/invlambda)
L1[i,j] <- - (1-d[i,j])*log(s[i,j]) +
d[i,j]*log(pi[i,j]*h20[i,j]*exp(eta[i,j])*pow(s20[i,j],exp(eta[i,j])))
L2[i,j] <- - ( (-pow(v[i],2)*0.5*ivv) + (0.5*log(ivv)) )/Nk[1,i]
zeros[i,j] <- 0
mu[i,j] <- - (L1[i,j]+L2[i,j]) + c
zeros[i,j] ~ dpois(mu[i,j])
}
}
sigmav <- 1/sqrt(ivv)
sigmav2 <- 1/ivv
lambda <- 1/invlambda
# priors beta1 ~ dnorm(0, 0.1)
gama0 ~ dnorm(0, 0.1)
gama1 ~ dnorm(0, 0.1)
ivv ~ dgamma(1,1)
alfa ~ dgamma(1,1)
invlambda ~ dgamma(1,1)
}

```

A.6 Comandos em Maple - modelo de mistura com efeitos aleatórios

Nesta seção estão apresentados os comandos utilizados na abordagem clássica para o modelo de mistura com efeitos aleatórios utilizado no programa Maple (versão 10), relatados na seção 5.3 e na simulação apresentada na seção 6.3.

```

>
> with(linalg):
>
> M:=6;n:=[9,14,16,7,11,9]; p:=11;
> y:=transpose(matrix(M,16,[1312, 1219, 264, 11, 170, 372, 182, 723, 90, -1, -1,
-1, -1, -1, -1, -1, 270, 162, 228, 370, 805, 517, 599, 89, 99, 219, 154, 266, 407,
751, -1, -1, 184, 1064, 1472, 526, 1495, 661, 407, 548, 1317, 1317, 1307, 800, 666,

```

```
1060, 465, 518, 235, 324, 1565, 1234, 293, 593, 213, -1, -1, -1, -1, -1, -1, -1, -1, -1, -1,
-1, 245, 262, 915, 1377, 172, 15, 382, 363, 107, 757, 38, -1, -1, -1, -1, -1, 222,
1644, 785, 11, 1095, 914, 346, 608, 112, -1, -1, -1, -1, -1, -1, -1]]):
```

```
> x:=transpose(matrix(M,16,[0, 0, 1, 1, 0, 0, 0, 1, 0, -1, -1, -1, -1, -1, -1, -1,
0, 0, 1, 1, 1, 0, 0, 1, 0, 1, 1, 1, 0, 0, -1, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, -1, -1, -1, -1, -1, -1, -1, -1, -1, 1, 0, 0, 1,
1, 1, 1, 0, 1, 1, 0, -1, -1, -1, -1, -1, 0, 0, 0, 1, 0, 0, 0, 0, 0, -1, -1, -1, -1,
-1, -1, -1])):
```

```
> nu:=transpose(matrix(M,16,[0, 0, 1, 1, 1, 1, 0, 0, 0, -1, -1, -1, -1, -1, -1, -1,
1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, -1, -1, 1, 1, 0, 1, 0, 1, 1, 1, 0, 0, 0,
1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, -1, -1, -1, -1, -1, -1, -1, -1, -1, 1, 1, 1, 0,
1, 1, 1, 1, 1, 1, 1, -1, -1, -1, -1, -1, 1, 0, 1, 1, 0, 0, 1, 1, 1, -1, -1, -1, -1,
-1, -1, -1])):
```

Contas intermediárias para facilitar a definição de l, a log-verossimilhança completa, já somada em N

```
> f2:=-1/2*ln(2*Pi*sigmav2)-(V[K]^2)/(2*sigmav2):
> l2:=sum(f2,K=1..M);
> eta[j,K] := x[j,K]*beta1;
> pi[j,K] := 1/(1+exp(gama0+gama1*x[j,K]+V[K]));
> S2[j,K]:=exp((-lambda*exp(eta[j,K]))*y[j,K]^alfa);
> S[j,K]:=1-pi[j,K]+pi[j,K]*S2[j,K];
> h20[j,K]:=lambda*alfa*y[j,K]^(alfa-1);
> S20[j,K]:=exp(-lambda*y[j,K]^alfa);
> f1:=(1-nu[j,K])*ln(S[j,K])+(nu[j,K])*ln(pi[j,K]*h20[j,K]*exp(eta[j,K])*
S20[j,K]^exp(eta[j,K]));
> l1:=sum(sum(f1,j=1..n[K]),K=1..M):
```

Aqui abaixo está a definição da log-verossimilhança completa

```
> l:=l1+l2:
```

Definindo as dimensões de V e de sua derivada

```
> V:=array(1..M): dV := array(1..M):
```

Definir os valores iniciais para os parâmetros, que entram na segunda coluna; a primeira coluna tem valores "grandes" o suficiente para que o difphi entre no loop. Vou sair com alfa(primeiro) e lambda(último) pré fixados, conhecidos

```
> phi := matrix(p,2,[1.26,1.26,10,0.6,10,-1,10,-0.8,10,0.5,10,0.5,10,
0.5,10,-0.5,10,-0.5,10,-0.5, 0.0004,0.0004]): sigmav2:=0.3:
```

Definir vetor de derivadas

```
> U := matrix(p-2,1,[ diff(l,beta1), diff(l,gama0), diff(l,gama1),
diff(l,V[1]), diff(l,V[2]), diff(l,V[3]), diff(l,V[4]), diff(l,V[5]),
diff(l,V[6])]):
```

```
> U2 := matrix(p-2,1):
```

Definindo a matriz hessiana


```

> H := hessian(1, [beta1,gama0,gama1,V[1],V[2],V[3],V[4],V[5],V[6]]):
> H2 := matrix(p-2,p-2);
> difphi:=evalf(norm(col(phi,1)-col(phi,2),2));
> passo := 0.1;
> for z from 1 to M do dV[z] := diff(1, V[z]); end do:
Loop: calcular até que difphi seja menor que 0,001.
> for t from 1 to 50 while (difphi > 0.001) do

for a from 1 to p do
phi[a,1] := phi[a,2]:
end do;

alfa:=phi[1,1];beta1:=phi[2,1];gama0:=phi[3,1];gama1:=phi[4,1];
V[1]:=phi[5,1];V[2]:=phi[6,1];V[3]:=phi[7,1];V[4]:=phi[8,1];
V[5]:=phi[9,1]; V[6]:=phi[10,1];
lambda:=phi[11,1];

for a from 1 to p-2 do
U2[a,1] := evalf(subs({alfa=phi[1,1],beta1=phi[2,1],gama0=phi[3,1],
gama1=phi[4,1],sigmav2=sigmav2,V[1]=phi[5,1],V[2]=phi[6,1],V[3]=phi[7,1],
V[4]=phi[8,1],V[5]=phi[9,1],V[6]=phi[10,1],lambda=phi[11,1]},U[a,1])):
end do:

for a from 1 to p-2 do
for b from 1 to p-2 do
H2[a,b] := evalf(subs({alfa=phi[1,1],beta1=phi[2,1],gama0=phi[3,1],
gama1=phi[4,1],sigmav2=sigmav2,V[1]=phi[5,1],V[2]=phi[6,1],V[3]=phi[7,1],
V[4]=phi[8,1],V[5]=phi[9,1],V[6]=phi[10,1], lambda=phi[11,1]},H[a,b])):
end do:
end do:

for a from 2 to p-1 do
phi[a,2] := phi[a,1] - passo*(evalm(inverse(H2) &* (U2))[a-1,1]);
end do;

sigmav2:= sum('V[z]^2','z'=1..M) / M:

# atualiza a condição de parada para a convergência
difphi:=evalf(norm(col(phi,1)-col(phi,2),2)):
print("t=",t);
end do;

```

Lendo da segunda coluna de phi quem são os "chapéus", os estimadores de alfa, beta1, gama0, gama1 e os BLUP para os V.

```
> alfachapeu:=phi[1,2];beta1chapeu:=phi[2,2];gama0chapeu:=phi[3,2];
gama1chapeu:=phi[4,2]; V[1]:=phi[5,2];V[2]:=phi[6,2];V[3]:=phi[7,2];
V[4]:=phi[8,2]; V[5]:=phi[9,2];V[6]:=phi[10,2]; lambdachapeu:=phi[11,2];
sigmav2chapeublup:= sum('V[z]^2','z'=1..M) / M;
Atualizando a matriz hessiana para os valores dos "chapéus"
> for a from 1 to p-2 do
for b from 1 to p-2 do
H2[a,b] := evalf(subs({alfa=phi[1,2],beta1=phi[2,2],gama0=phi[3,2],
gama1=phi[4,2],sigmav2=sigmav2chapeublup,V[1]=phi[5,2],V[2]=phi[6,2],
V[3]=phi[7,2],V[4]=phi[8,2],V[5]=phi[9,2],V[6]=phi[10,2],
lambda=phi[11,2]},H[a,b]));
end do;
end do;
```

REML: estimar o sigmav2 de modo menos viesado do que pelo método acima

```
> A33:=matrix(M,M,[]);
> H2inv:=inverse(H2);
> A33:=evalm(-1*matrix([col(matrix([col(H2inv,4..p-2)]),4..p-2)]));
> sigmav2chapeu:=1/M*evalm( trace(A33) + (transpose(V)&*V) );
Substituindo os valores das estimativas no verossimilhança, que agora vai depender só de alfa e
lambda, para podermos estimar alfa e lambda
> #subs(beta1=beta1chapeu, subs(gama0=gama0chapeu, subs(gama1=
gama1chapeu, subs(V[1]=phi[5,2], subs(V[2]=phi[6,2], subs(V[3]=phi[7,2],
subs(V[4]=phi[8,2], subs(V[5]=phi[9,2], subs(V[6]=phi[10,2],
subs(sigmav2=sigmav2chapeu,1)))))))));
Definir a log-verossimilhança já com as estimativas substituídas, de modo a ficarmos somente com
alfa e lambda
> f2c:=-1/2*ln(2*Pi*sigmav2chapeu)-(V[K]^2)/(2*sigmav2chapeu):
> l2c:=sum(f2c,K=1..M);
> etac[j,K] := x[j,K]*beta1chapeu;
> pic[j,K] := 1/(1+exp(gama0chapeu+gama1chapeu*x[j,K]+V[K]));
> S2c[j,K]:=exp(-lambdac*exp(etac[j,K]))*y[j,K]^alfac;
> Sc[j,K]:=1-pic[j,K]+pic[j,K]*S2c[j,K];
> h20c[j,K]:=lambdac*alfac*y[j,K]^(alfac-1);
> S20c[j,K]:=exp(-lambdac*y[j,K]^alfac);
> f1c:=(1-nu[j,K])*ln(Sc[j,K])+(nu[j,K])*ln(pic[j,K]*h20c[j,K]*
exp(etac[j,K])*S20c[j,K]^exp(etac[j,K]));
> l1c:=sum(sum(f1c,j=1..n[K]),K=1..M);
Aqui abaixo está a definição da log-verossimilhança completa
```

```

> lc:=l1c+l2c:
> lcc:=subs(V[1]=phi[5,2], subs(V[2]=phi[6,2], subs(V[3]=phi[7,2],
subs(V[4]=phi[8,2], subs(V[5]=phi[9,2], subs(V[6]=phi[10,2],
subs(sigmav2=sigmav2chapeu,lc)))))):
Aqui tem que fazer o novo loop, para estimar alfa e lambda
> phi2:=matrix(2,2,[10,alfa,10,lambda]);
Definir vetor de derivadas
> Uc := matrix(2,1,[ diff(lcc,alfac), diff(lcc,lambdac)]):
> U2c := matrix(2,1):
Definindo a matriz hessiana
> Hc := hessian(lcc,[alfac,lambdac]):
> H2c := matrix(2,2);
> difphi2:=evalf(norm(col(phi2,1)-col(phi2,2),2));
> passo2 := 0.1;
> for t from 1 to 1000 while (difphi2 > 0.001) do

for a from 1 to 2 do
phi2[a,1] := phi2[a,2]:
end do;

alfac:=phi2[1,1];
lambdac:=phi2[2,1];

for a from 1 to 2 do
U2c[a,1] := evalf(subs({alfac=phi2[1,1],lambdac=phi2[2,1]},Uc[a,1]]):
end do:

for a from 1 to 2 do
for b from 1 to 2 do
H2c[a,b] := evalf(subs({alfac=phi2[1,1],lambdac=phi2[2,1]},Hc[a,b]]):
end do:
end do:

for a from 1 to 2 do
phi2[a,2] := phi2[a,1] - passo2*(evalm(inverse(H2c) &* (U2c))[a,1]);
end do;

phi2[1,2]:= 'if'(phi2[1,2]<=0,0.01,phi2[1,2]);
phi2[2,2]:= 'if'(phi2[2,2]<=0,0.01,phi2[2,2]);

# atualiza a condição de parada para a convergência
difphi2:=evalf(norm(col(phi2,1)-col(phi2,2),2)):
print("t=",t);
end do;

```

```
> alfachapeu:=phi2[1,2];lambdachapeu:=phi2[2,2];
```

Aqui estão as estimativas finais

```
> alfachapeu:=phi2[1,2];
```

```
lambdachapeu:=phi2[2,2];
```

```
beta1chapeu:=phi[2,2];
```

```
gama0chapeu:=phi[3,2];
```

```
gama1chapeu:=phi[4,2];
```

```
sigmav2chapeu:=1/M*evalm( trace(A33) + (transpose(V)&*V) );
```

Referências Bibliográficas

- [1] Berger, J. O. (1985). *Statistical decision theory and bayesian analysis*. New York: Springer Verlag.
- [2] Berkson, J.; Gage, R. (1952). Survival curve for cancer patients following treatment, *Journal of the American Statistical Association* **47**:501-515.
- [3] Cancho, V. G.; Rodrigues, J.; Castro, M. (2008). A bayesian model for survival data with a cure rate. Manuscrito a ser submetido.
- [4] Castro, M.; Cancho, V. G.; Rodrigues, J. (2007). A flexible model for survival data with a surviving fraction. Manuscrito a ser submetido para *Statistical Modelling*.
- [5] Chen, M.-H.; Ibrahim, J. G.; Sinha, D. (1999). A new bayesian model for survival data with a surviving fraction, *Journal of the American Statistical Association* **94**:909-919.
- [6] Chen, M.-H.; Ibrahim, J. G.; Sinha, D. (2002). Bayesian inference for multivariate survival data with a surviving fraction, *Journal of Multivariate Analysis* **80**:101-126.
- [7] Chib, S.; Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *The American Statistician*; **49(4)**: 327-335.
- [8] Cowless, M. K.; Carlin, B. P. (1996). Markov Chain Monte Carlo convergence diagnostics: a comparative review, *Journal of the American Statistical Association* **91(434)**:883-904.
- [9] Cox, D. R. (1972). Regression models and life tables (with discussion), *Journal of Royal Statistical Society, Ser. B* **34**:187-220.
- [10] Dempster, A.; Laird, N.; Rubin, D. (1977). Maximum likelihood from incomplete data via EM algorithm, *Journal of the Royal Statistical Society, Ser. B* **39**: 1-38.
- [11] Efron, B. (1979). Bootstrap methods: another look at the jackknife, *Annals of Statistics* **7**: 1-26.

- [12] Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM.
- [13] Feller, W. (1967). *An introduction to probability theory and its applications*, v.1, 3.ed. New York: Wiley.
- [14] Gelman, A.; Rubin, D. R. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**(4):457-511.
- [15] Geman, S.; Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, **6**, 721-741.
- [16] Henderson, C. R. (1975). Best linear unbiased estimation and prediction under a selection model, *Biometrics* **31**:423-447.
- [17] Ibrahim, J. G.; Chen, M.-H.; Sinha, D. (2001) *Bayesian survival analysis*. New York: Springer Verlag.
- [18] Kalbfleisch, J. D.; Prentice R. L. (1980). *The statistical analysis of failure data*. New York: Wiley.
- [19] Louis, T. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society Series B*, **44**: 226-233.
- [20] Maller, R. A.; Zhou, X. (1996). *Survival analysis with long-term survivors*. New York: Wiley.
- [21] Maplesoft. (2005). *Maple User Manual*. Maplesoft.
- [22] McGilchrist, C. A. (1994). Estimation in generalized mixed models, *Journal of the Royal Statistical Society Series B* **56**:61-69.
- [23] McGilchrist, C. A. e Yau, K. K. W. (1995). The derivation of BLUP, ML, REML estimation methods for generalised linear mixed models, *Communications in Statistics - Theory and Methods* **24**(12):2963-2980.
- [24] McLachan, G. J. e Krishnan, T. (1997). *The EM algorithm and extensions*. New York: Wiley.
- [25] Mizoi, M. F. (2004). *Influência local em modelos de sobrevivência com fração de cura*. Tese de Doutorado - IME - USP. São Paulo.

- [26] Rodrigues, J.; Louzada-Neto, F. (2006). On the unification of the long-term survival models. Manuscrito submetido para *Statistics and Probability Letters*.
- [27] Spiegelhalter, D.; Thomas, A.; Best, N.; Lunn, D. (2007). *Winbugs user manual*. Imperial College and MRC, UK.
- [28] Tanner, M.; Wong, W. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, **82**: 528-550.
- [29] Thompson, R. (1980). Maximum likelihood estimation of variance components. *Math. Operationsforsch. Statist., Ser. Statistics*, **11**: 545-561.
- [30] Venables, W. N.; Smith, D. M., R Development Core Team (2008). *An introduction to R*. Disponível em <http://cran.r-project.org/doc/manuals/R-intro.pdf>. Acesso em 23 maio 2008.
- [31] Yau, K. K. W.; Ng, A. S. K. (2001). Long-term survivor mixture model with random effects: application to a multi-centre clinical trial of carcinoma, *Statistics in Medicine* **20**: 1591-1607.
- [32] Yin, G.; Ibrahim, J. G. (2005). Cure rate models: a unified approach, *The Canadian Journal of Statistics* **33(4)**: 559-570.