

**DADOS DE SOBREVIVÊNCIA MULTIVARIADOS NA
PRESENÇA DE COVARIÁVEIS E OBSERVAÇÕES
CENSURADAS: UMA ABORDAGEM BAYESIANA**

Carlos Aparecido dos Santos

Orientador: Prof. Dr. Jorge Alberto Achcar

São Carlos-SP
Março - 2010

**DADOS DE SOBREVIVÊNCIA MULTIVARIADOS NA
PRESENÇA DE COVARIÁVEIS E OBSERVAÇÕES
CENSURADAS: UMA ABORDAGEM BAYESIANA**

Carlos Aparecido dos Santos

Orientador: Prof. Dr. Jorge Alberto Achcar

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

São Carlos-SP
Março - 2010

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

S237ds

Santos, Carlos Aparecido dos.

Dados de sobrevivência multivariados na presença de covariáveis e observações censuradas : uma abordagem bayesiana / Carlos Aparecido dos Santos. -- São Carlos : UFSCar, 2010.

145 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2010.

1. Análise de sobrevivência. 2. Inferência bayesiana. 3. Cópula. 4. Algoritmos MCMC. 5. Dados censurados. I. Título.

CDD: 519.9 (20^a)

Carlos Aparecido dos Santos

**Dados de sobrevivência multivariados na presença de covariáveis e
observações censuradas: uma abordagem bayesiana**

Tese apresentada à Universidade Federal de São Carlos, como parte dos requisitos
para obtenção do título de Doutor em Estatística.

Aprovada em 04 de março de 2010.

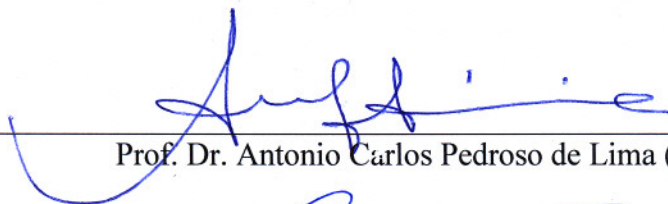
BANCA EXAMINADORA

Presidente



Prof. Dr. Jorge Alberto Achcar (FMRP-USP/Orientador)

1º Examinador



Prof. Dr. Antonio Carlos Pedroso de Lima (IME-USP)

2º Examinador



Prof. Dr. Carlos Alberto de Bragança Pereira (IME-USP)

3º Examinador



Profa. Dra. Eliane Regina Rodrigues (UNAM/México)

4º Examinador



Prof. Dr. Luiz Koodi Hotta (IMECC-UNICAMP)

Dedico esta tese, com todo amor e carinho do mundo, a minha esposa e companheira Daniele Cristina Tita Granzotto, que sempre me incentivou e me ajudou a lutar pelas coisas em que acredito; e a minha linda princesa Iasmim Granzotto dos Santos, que é minha grande fonte de inspiração.

*Quando estiver em dificuldade e pensar em desistir,
Lembre-se dos obstáculos que já superou,
Olhe para trás.
Se tropeçar e cair, levante,
Não fique prostrada(o),
Esqueça o passado.
Olhe para frente.
Ao sentir-se orgulhosa(o),
Por alguma realização pessoal,
Sonde suas motivações,
Olhe para dentro.
Antes que o egoísmo o domine,
Enquanto seu coração é sensível socorra aos que o cercam,
Olhe para os lados.
Na escalada rumo às altas posições,
No afã de concretizar seus sonhos,
Observe se não está pisando em alguém,
Olhe para baixo.
Em todos os momentos da vida,
Seja qual for sua atividade,
Busque a aprovação de DEUS e agradeça a Ele, seu amor!
Olhe para cima.*

(Charles Chaplin - Direcione seu olhar)

Agradecimentos

Agradeço...

Inicialmente Deus, pois sem ele nem minha própria existência seria possível.

Ao meu orientador Prof. Dr. Jorge Alberto Achcar, por ter acreditado em mim e estar presente em cada etapa deste trabalho. Sua dedicação, entusiasmo e incentivo para com a pesquisa é impar e me permitiu que trilhasse um caminho seguro durante todo o processo de elaboração deste trabalho, portanto, tem minha eterna gratidão e admiração.

Aos professores, Dr. Carlos Fogo e Dr. Edson Martinez pelas valiosas sugestões e comentários apresentados durante o exame de qualificação.

Aos professores do programa de pós-graduação do Departamento de Estatística da UFSCar, pelos ensinamentos e apoio as minhas dificuldades neste período.

A todos os professores e funcionários do Departamento de Estatística - Universidade Estadual de Maringá por terem me propiciado a oportunidade de realizar mais este projeto de minha vida.

A toda minha família, em especial meus pais Mario e Maria; meus irmãos Sueli e Marcos, meu sogros Benedicto e Sirlei, meu cunhado Marcos e Roberta.

A todos os amigos e amigas que, de alguma maneira contribuíram para a realização deste trabalho, em especial o Ferreira, a Cristina, a Tereza Cristina, a Vera Lúcia, o João Tomazella, o Silvio Marra e o Sérgio Borges, pela amizade e carinho e pela ajuda no decorrer do trabalho.

RESUMO

Nesta tese introduzimos uma análise Bayesiana para dados de sobrevivência multivariados, na presença de um vetor de covariáveis e observações censuradas. Diferentes “fragilidades” ou variáveis latentes são consideradas para capturar a correlação existente entre os tempos de sobrevivência, para o mesmo indivíduo. Também apresentamos uma análise Bayesiana para algumas das mais populares distribuições exponenciais bivariadas introduzidas na literatura. Uma análise Bayesiana também é introduzida para a distribuição exponencial bivariada de Block & Basu, usando métodos MCMC (Monte Carlo em Cadeias de Markov) e considerando os tempos de sobrevivência na presença de covariáveis e dados censurados. Em outro tópico, introduzimos uma análise Bayesiana para dados de sobrevivência bivariados na presença de covariáveis e observações censuradas, assumindo diferentes distribuições bivariadas Weibull derivadas de algumas funções cópulas existentes. Uma grande simplificação computacional para simular amostras da distribuição a posteriori conjunta de interesse é obtida usando o *software WinBUGS*. Ilustrações numéricas são introduzidas considerando conjunto de dados reais, para cada uma das metodologias propostas.

Palavras chaves: Distribuições bivariadas Weibull, distribuição exponencial bivariada de Block & Basu, funções cópulas, métodos MCMC, análise Bayesiana, dados censurados, covariáveis.

ABSTRACT

In this work, we introduce a Bayesian Analysis for survival multivariate data in the presence of a covariate vector and censored observations. Different “frailties” or latent variables are considered to capture the correlation among the survival times for the same individual. We also introduce a Bayesian analysis for some of the most popular bivariate exponential distributions introduced in the literature. A Bayesian analysis is also introduced for the Block & Basu bivariate exponential distribution using Markov Chain Monte Carlo (MCMC) methods and considering lifetimes in presence of covariates and censored data. In another topic, we introduce a Bayesian Analysis for bivariate lifetime data in the presence of covariates and censoring data assuming different bivariate Weibull distributions derived from some existing copula functions. A great computational simplification to simulate samples for the joint posterior distribution is obtained using the WinBUGS software. Numerical illustrations are introduced considering real data sets considering every proposed methodology.

Key words: Bivariate Weibull distributions, Block & Basu bivariate exponential distribution, Copula functions, MCMC methods, Bayesian analysis, censoring data, covariates.

Sumário

1	Introdução	1
2	Análise de Sobrevivência	3
2.1	Dados de Sobrevivência	4
2.1.1	Censuras	6
2.1.2	Presença de Variáveis Explicativas	8
2.2	Tempos de Falhas Multivariados	9
2.2.1	A Função de Verossimilhança	10
2.3	Modelos Paramétricos de Sobrevivência	11
2.3.1	Modelo Exponencial	12
2.3.2	A Distribuição Weibull	12
2.3.3	As Distribuições Gama e Gama Generalizada	14
2.4	Dados de Eventos Recorrentes	17
2.5	Modelos de Regressão	20
2.6	Modelando as Fragilidades	22
3	Métodos Bayesianos	26
3.1	Monte Carlo via Cadeias de Markov (MCMC)	28
3.1.1	Amostrador de <i>Gibbs</i> (<i>Gibbs Sampling</i>)	28
3.1.2	Algoritmo de Metropolis-Hastings	29
3.2	CrITÉrios de Comparação de Modelos	30
3.3	Modelos Hierárquicos	32
3.4	Diagnósticos de Convergência	34
3.4.1	Método de Gelman e Rubin	35
4	Alguns Modelos de Regressão para Dados de Sobrevivência Multivariados Obtidos pela Inclusão de Fatores Aleatórios	38
4.1	Modelo de Regressão Weibull	38
4.1.1	Análise Bayesiana para o Modelo	42
4.2	Modelo de Regressão Gama Generalizada	46

4.3	Aplicação aos Dados de McGilchrist e Aisbett	47
4.3.1	Análise Descritiva dos Dados	48
4.3.2	Análise Bayesiana dos dados	53
5	Modelos Paramétricos Exponenciais Bivariados para Dados de Sobre-	
	vivência sem Censuras	63
5.1	Algumas Distribuições Exponenciais Bivariadas, num contexto Bayesiano .	65
5.1.1	A Distribuição Exponencial Bivariada de Block & Basu	65
5.1.2	A Distribuição Exponencial Correlacionada da Família Farlie-Gumbel-	
	Morgenstern	67
5.1.3	Distribuição Exponencial Bivariada de Gumbel	69
5.1.4	Distribuição Exponencial Bivariada de Freund	71
5.1.5	Distribuição Bivariada de Marshall & Olkin	73
5.1.6	Presença de “Fragilidades” para Capturar a Correlação entre Distri-	
	buições Exponenciais	75
5.2	Análise dos dados da Liga dos Campeões da UEFA	77
6	Análise Bayesiana da Distribuição Exponencial Bivariada de Block &	
	Basu com Censuras e Covariáveis	85
6.1	Uma Análise Bayesiana na Presença de Dados Censurados	87
6.2	Presença de Covariáveis e Observações Censuradas	89
6.3	Exemplos	91
6.3.1	Tempos de Recorrência para a Infecção de rins de Pacientes	91
6.3.2	Transplante de Medula Óssea para Leucemia	95
7	Modelos Bivariados de Sobrevivência Derivados de Funções Cópulas	102
7.1	Funções Cópulas	103
7.2	Algumas Distribuições Bivariadas Weibull Derivadas de	
	Funções Cópulas	105
7.3	Uma Análise Bayesiana para o “Modelo 1” e “Modelo 2” na Presença de	
	Dados Censurados	107
7.3.1	“Modelo 1”	108
7.3.2	“Modelo 2”	109
7.4	Presença de Covariáveis e Dados Censurados	111
7.5	Exemplos	112
7.5.1	Tempos de Recorrência para a Infecção de rins de Pacientes	112
7.5.2	Transplante de Medula Óssea para Leucemia	117

8	Conclusões e Perspectivas Futuras	124
8.1	Estudos Futuros	128
A	Medidas condicionais e não-condicionais para a distribuição Gama	138
B	Alguns Programas Utilizados	141

Capítulo 1

Introdução

O objetivo principal desta tese é modelar, via métodos Bayesianos, dados de sobrevivência multivariados, na presença de covariáveis e observações censuradas.

Um conjunto de dados de sobrevivência multivariados geralmente apresenta associação entre os tempos de sobrevivência, onde essa possível associação é modelada frequentemente através de um modelo de *fragilidade*. Nesse modelo, um efeito aleatório, denominado fragilidade, é introduzido na função de risco para descrever essa possível heterogeneidade entre as unidades em estudo.

Outra forma de modelar esse tipo de dados, em especial dados bivariados, é a utilização dos populares modelos de tempos de sobrevivência dados pelas distribuições exponenciais bivariadas, onde podemos destacar os modelos de: Block & Basu (1974); Farlie-Gumbel-Morgenstern (Farlie, 1960); Gumbel (1960); Freund (1961) e Marshall & Olkin (1967a,b).

Também podemos utilizar diferentes formulações para a distribuição Weibull ou outras distribuições de sobrevivência, para tempos de sobrevivência bivariados, considerando o uso de funções cópulas (veja por exemplo, Nelsen, 1999; Trivedi & Zimmer, 2005a,b).

Problemas relacionados com tempos de sobrevivência e ocorrência de eventos sucessivos como, por exemplo, dois ou mais tempos de sobrevivência associados a cada unidade (podem ser componentes ou indivíduos), são aplicações comuns em várias áreas, como medicina e engenharia. Nesse caso uma suposição de independência para os tempos de sobrevivência pode não ser adequada, sendo razoável supor, quando existe um agrupamento natural ou artificial de unidades, que haja alguma associação entre os tempos de um mesmo grupo. Como alguns desses eventos de interesse não são terminais (não há morte do indivíduo) e podem ocorrer mais de uma vez para uma mesma unidade, temos os conhecidos eventos recorrentes.

No caso de dois tempos de sobrevivência podemos adotar um modelo de riscos proporcionais de Cox (Cox, 1972), considerando os tempos independentes e incorporando as covariáveis e a presença de um efeito aleatório que captura a possível correlação entre os

tempos de vida; ou ainda utilizando um modelo de sobrevivência paramétrico bivariado (ver por exemplo, Block & Basu, 1974). Quando são considerados mais de dois tempos de sobrevivência ou mesmo dados bivariados associados a um mesmo paciente, a abordagem clássica pode se tornar muito complexa e imprecisa e a metodologia Bayesiana surge como boa alternativa.

Nesta tese, vamos explorar alguns modelos de sobrevivência multivariados propostos na literatura, usualmente considerados para dados sem censuras e generalizar esses resultados na presença de dados censurados e covariáveis. Isso será considerado para a distribuição exponencial bivariada de Block & Basu. Além disso, vamos propor diferentes estruturas de fragilidade para capturar a correlação existente entre os tempos de sobrevivência multivariados e propor novos modelos de sobrevivência derivados de funções cópulas, em especial considerando dados de sobrevivência bivariados. Para a análise estatística dos modelos propostos, vamos usar a metodologia Bayesiana onde sumários a posteriori de interesse serão obtidos via simulação MCMC (Monte Carlo em Cadeias de Markov).

A tese é organizada da seguinte forma: no Capítulo 2, apresentamos alguns conceitos básicos em análise de sobrevivência; no Capítulo 3, introduzimos alguns conceitos básicos sobre métodos Bayesianos e em especial sobre métodos MCMC (Monte Carlo em Cadeias de Markov); no Capítulo 4, introduzimos alguns modelos de regressão para dados de sobrevivência multivariados na presença de diferentes estruturas de fragilidade; no Capítulo 5, introduzimos uma análise Bayesiana para alguns modelos paramétricos usuais para dados de sobrevivência bivariados; no Capítulo 6, apresentamos uma análise Bayesiana para a distribuição exponencial bivariada de Block & Basu na presença de dados censurados e covariáveis; no Capítulo 7, introduzimos alguns modelos de sobrevivência bivariados derivados de funções cópulas, com especial destaque para a distribuição bivariada de Weibull na presença de censuras e covariáveis; finalmente no Capítulo 8, apresentamos algumas conclusões e perspectivas futuras.

Alguns dos programas desenvolvidos para os modelos propostos, são apresentados no Apêndice B; os demais podem ser obtidos diretamente com o autor desta tese.

Capítulo 2

Análise de Sobrevivência

A análise de sobrevivência é uma das áreas da Estatística que experimentou um rápido crescimento durante a última metade do século XX (Colosimo & Giolo, 2006). Um fator que contribuiu para a aplicação intensa desses métodos em diversas áreas do conhecimento, foi sua implementação nos atuais programas computacionais, mais rápidos e flexíveis. Uma indicação desse sucesso é o número elevado de publicações na área de medicina, por exemplo, que cresceu de 11% em 1979, para 32%, em 1989, no número de artigos do conceituado periódico *The New England Journal of Medicine*.

A análise de sobrevivência pode ser definida, de forma ampla, como a análise do tempo até a ocorrência de um dado evento. Este tempo é denominado tempo de falha e pode ser o tempo até a morte do paciente, bem como até a cura ou recidiva de uma doença.

A principal característica de dados de sobrevivência é a presença de censuras, que é a observação parcial da resposta. Isto se refere a situações em que, por alguma razão, o acompanhamento do paciente foi interrompido, seja porque o mesmo mudou de cidade, o estudo terminou para a análise dos dados ou, o paciente morreu de causa diferente da estudada. Sem a presença de censuras, técnicas estatísticas clássicas usuais podem ser usadas para analisar os dados (ver por exemplo, Kalbfleisch & Prentice, 2002).

Por outro lado, alguns eventos de interesse não são terminais e podem ocorrer, para o mesmo indivíduo, mais de uma vez; originando os eventos recorrentes.

Para eventos recorrentes os dados são observados para cada indivíduo, sendo verificados nas áreas de estudos clínicos, demografia, criminologia, confiabilidade industrial e produção. Esse tipo de dados surge naturalmente em estudos longitudinais envolvendo tempos de sobrevivência múltiplos. Diversos artigos e trabalhos têm abordado problemas

de eventos recorrentes, como por exemplo, Lawless & Nadeau (1995) e Tomazella (2003), que utilizam um Processo de Poisson para desenvolver modelos que enfocam o número esperado de eventos ocorridos em um determinado intervalo de tempo.

Apresentamos a seguir, alguns conceitos que são importantes em análise de sobrevivência.

2.1 Dados de Sobrevivência

Quando se trata de análise de sobrevivência, as unidades de estudo são usualmente indivíduos. Essas unidades que compõem os dados de sobrevivência, são formados essencialmente pelos tempos de vida das mesmas até a ocorrência do evento de interesse. Os tempos de sobrevivência, como normalmente são chamados, podem também ser referidos como tempos de falhas ou tempos de vida.

Os dados de sobrevivência incorporam tanto os tempos de sobrevivência como o conjunto de outras variáveis observáveis que podem estar relacionadas com esses tempos. Essas variáveis são chamadas covariáveis ou variáveis explicativas (explanatórias). Quando os tempos de sobrevivência estão relacionados com essas covariáveis, dizemos que a população é heterogênea; caso contrário ela é homogênea.

As covariáveis normalmente são medidas uma única vez ao longo do tempo de estudo e por conta disso, são fixas. Em alguns casos, onde as covariáveis não são fixas, estas variam em função do tempo, modificando-se durante o período de observação (ver por exemplo, Tomazella, 2003).

A variável aleatória contínua e não-negativa denotando o tempo de sobrevivência ($T \geq 0$), pode ser expressa através de funções matematicamente equivalentes, onde a especificação de uma delas é suficiente para derivar as outras. Então temos as funções: densidade de probabilidade, $f(t)$, a de sobrevivência, $S(t)$ e a de risco, $h(t)$; que são apresentadas nesta seção.

Estas funções são utilizadas na prática com o objetivo de descrever os aspectos apresentados pelo conjunto de dados.

A função densidade de probabilidade (fdp) é expressa como o limite da probabilidade de um indivíduo vir a experimentar o evento de interesse no intervalo de tempo $[t, t + dt)$ por unidade de tempo e é expressa como (ver por exemplo, Lee & Wang, 2003),

$$f(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{dt}, \quad (2.1)$$

no qual $f(t) \geq 0$ para todo t . Por se tratar de uma fdp, temos a área abaixo da curva igual a 1.

A função que descreve a forma distribucional dos tempos de sobrevivência através da probabilidade de um indivíduo não falhar (ou do evento de interesse não ocorrer), pelo menos até um instante de tempo t , é dada por (ver por exemplo, Lawless, 1982),

$$\begin{aligned} S(t) &= P(T > t) \\ &= 1 - P(T \leq t) \\ &= 1 - \int_0^t f(u) du, \end{aligned} \quad (2.2)$$

em que $f(\cdot)$ é a fdp; $S(t)$ é a função de sobrevivência no tempo fixado t .

Alternativamente, (2.2) pode ser escrita na forma

$$S(t) = 1 - F(t), \quad (2.3)$$

sendo $F(t)$ a probabilidade de um indivíduo experimentar o evento de interesse até o tempo t ; então a partir de (2.3), temos $S(t) + F(t) = 1$ ($F(t)$ é a função de distribuição acumulada).

Também observamos em (2.2), uma função monótona não-crescente no intervalo de tempo $\mathcal{T} = [0, \infty)$, tal que $S(0) = 1$ e $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$.

A função de risco fornece a taxa instantânea de falha, por unidade de tempo, isto é, o limite da probabilidade de um indivíduo falhar no intervalo de tempo $[t, t + dt)$ com $dt \rightarrow 0$, dado que ele tenha sobrevivido até o instante t e é expressa por (ver por exemplo, Cox & Oakes, 1984)

$$h(t) = \lim_{dt \rightarrow 0^+} \frac{P(t \leq T < t + dt | T \geq t)}{dt}. \quad (2.4)$$

Devido a sua interpretação, a função de risco (2.4) tem sido preferida por muitos autores para descrever o comportamento do tempo de sobrevivência e sua importância é descrita por Klein & Moeschberger (1997).

A função de risco também é conhecida como força de mortalidade ou taxa de mortalidade condicional. Além disso, através da função de risco pode-se caracterizar classes especiais de distribuições de tempos de sobrevivência, de acordo com o seu comportamento em relação ao tempo. A função de risco pode ser constante, crescente, decrescente ou mesmo não monótona. Algumas distribuições usuais de tempo de sobrevivência são por

exemplo, a exponencial, que apresenta a função de risco constante e a Weibull (Weibull, 1951), que apresenta função de risco crescente, decrescente ou ainda constante. Estas funções, bem como outras, serão apresentadas no decorrer desta tese.

Da função de risco, dada em (2.4), temos que

$$\begin{aligned}
 h(t) &= \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt | T \geq t)}{dt} \\
 &= \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt)}{dt P(T \geq t)} \\
 &= \lim_{dt \rightarrow 0} \left[\frac{F(t + dt) - F(t)}{dt} \right] \frac{1}{S(t)} \\
 &= \left(\frac{d}{dt} F(t) \right) \frac{1}{S(t)} \\
 &= \frac{f(t)}{S(t)}. \tag{2.5}
 \end{aligned}$$

Usando a relação apresentada em (2.3), temos que (2.5) pode ainda ser escrita como sendo

$$h(t) = -\frac{S'(t)}{S(t)} = -\frac{d}{dt} \log[S(t)]. \tag{2.6}$$

A partir de (2.6) temos,

$$S(t) = \exp \left[- \int_0^t h(u) du \right], \tag{2.7}$$

já que $\int_0^t h(u) du$ é a função de risco acumulada e é finita para algum tempo $t > 0$ e $\int_0^\infty h(u) du = \infty$.

Desta forma, temos duas expressões alternativas para representar a fdp da variável aleatória T :

$$f(t) = h(t) S(t) \quad \text{ou} \quad f(t) = h(t) \exp \left[- \int_0^t h(u) du \right]. \tag{2.8}$$

2.1.1 Censuras

Um complicador presente nos dados de sobrevivência relaciona-se ao fato da variável de interesse, normalmente tempo de sobrevivência, não ser medida instantaneamente e in-

dependentemente do tamanho da resposta. Para obter valores grandes da variável tempo, necessitamos de mais tempo e persistência nas observações. Em situações extremas, este fato pode comprometer a observação do valor da variável para alguns indivíduos, uma vez que o evento de interesse pode não ocorrer até o final do tempo de estudo. Também o indivíduo pode abandonar o estudo antes mesmo da ocorrência do evento de interesse ou vir a óbito devido a outras causas, que não seja a estudada (Achcar *et al.*, 2002). Esse tipo de dados são denominados tempos censurados.

Os tempos censurados devem ser usados na análise estatística, pois mesmo incompletas essas observações nos fornecem informações sobre o tempo de falha dos pacientes; sendo que sua omissão levará certamente a conclusões errôneas.

Desta forma existe a necessidade de introduzir uma variável extra na análise, indicando se o indivíduo teve ou não seu tempo de sobrevivência exatamente observado. Na literatura de análise de sobrevivência essa variável é conhecida, como variável indicadora de falha e assume os valores,

$$\delta = \begin{cases} 1 & \text{para observação completa} \\ 0 & \text{para observação censurada.} \end{cases}$$

Podemos citar quatro tipos de censuras (ver por exemplo, Lee & Wang, 2003):

- **Censura do tipo I:** ocorre quando o estudo termina em um tempo pré-estabelecido e alguns dos tempos de sobrevivência dos pacientes não puderam ser observados, tendo seus tempos censurados à direita. Por exemplo, suponha que seis ratos tenham sido expostos a um agente carcinogênico, pela injeção de células contaminadas nos seus pés. Os tempos de desenvolvimento de um tumor, de um dado tamanho, são observados. O pesquisador decide terminar o experimento depois de 30 semanas. A Figura 2.1(a) é um gráfico dos tempos de desenvolvimento dos tumores. Os Ratos A, B e D desenvolveram tumores após 10, 15 e 25 semanas, respectivamente. Os ratos C e E não desenvolveram tumores no final do estudo (seus tempos livres de tumores ocorrem após 30 semanas). O rato F morreu acidentalmente, sem tumores, após 19 semanas de observação, portanto antes do tempo de censura.
- **Censura do tipo II:** ao invés do tempo final do estudo ser pré-estabelecido, o estudo termina após um determinado número de pacientes experimentar o evento de interesse. Neste caso, se não há perda acidental, as observações censuradas são de comprimento igual a das observações não-censuradas. Do exemplo anterior, podemos observar na Figura 2.1(b), que o pesquisador pode decidir terminar o estudo depois que quatro, dos seis ratos, tenham desenvolvido tumores (no gráfico, as "??" significam observações perdidas).

- **Censura aleatória:** diferentemente das outras censuras, ela é mais difícil do experimentador ter controle. Geralmente ocorre quando o paciente abandona o estudo sem ter experimentado o evento de interesse. Por exemplo, o rato F da Figura 2.1 (a).
- **Censura intervalar:** ocorre quando não se conhece o tempo exato em que ocorreu o evento de interesse, mas sim, que ele ocorreu dentro de um intervalo de tempo especificado.

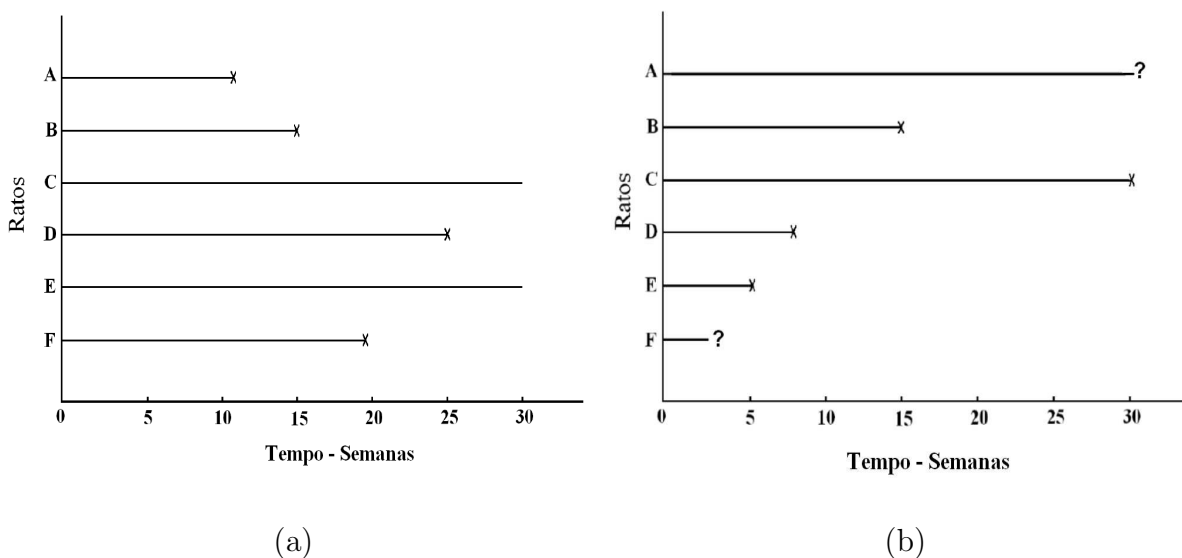


Figura 2.1: (a): Dados censurados do tipo I. (b): Dados censurados do tipo II.

2.1.2 Presença de Variáveis Explicativas

Além do tempo de sobrevivência e da variável indicadora de falha, podemos ter observado, nos dados, variáveis (por exemplo, sexo e idade) que representam tanto a heterogeneidade existente na população, como também possíveis tratamentos aos quais os indivíduos são submetidos.

Tais variáveis são conhecidas como covariáveis ou variáveis explicativas. Muitas vezes, o objetivo na análise de sobrevivência está centrado na relação entre o tempo de sobrevivência e algumas variáveis explicativas de interesse (ver por exemplo, Dias, 2002).

Para uma análise estatística, temos as variáveis tempo de sobrevivência, variável indicadora de censuras e um vetor de variáveis explicativas disponíveis para essa análise.

Um complicador adicional que também pode ocorrer na análise de sobrevivência é encontrar variáveis explicativas dependentes do tempo, ou seja, os valores das covariáveis no final do estudo podem diferir dos valores iniciais. Por exemplo, podemos ter um experimento em que a dose de um medicamento é modificada ao longo do experimento, uma vez que o paciente pode sofrer algum tipo de efeito colateral da droga (ver por exemplo, Tomazella, 2003).

2.2 Tempos de Falhas Multivariados

Existem diferentes classificações para dados multivariados de sobrevivência (ver por exemplo, Hougaard, 2000; Kalbfleisch & Prentice, 2002). Os tempos de falha multivariados são obtidos quando:

- (a) Dois ou mais eventos, do mesmo tipo ou de tipos diferentes, ocorrem no mesmo indivíduo;
- (b) Eventos ocorrem em indivíduos que estão agrupados pelo planejamento usado no estudo, e existe razão para assumir uma possível estrutura de correlação entre os tempos de falha de indivíduos do mesmo grupo, sendo que há independência entre os grupos.

Segundo Therneau & Grambsch (2000), as falhas podem ser classificadas como sendo:

- (a) **Falhas ordenadas:** são os eventos recorrentes que surgem quando cada indivíduo em observação está sujeito a experimentar o mesmo evento várias vezes, como um estudo que registre os tempos de ocorrência do primeiro e segundo infarto do miocárdio. São falhas ordenadas porque um indivíduo só estará em risco de experimentar a segunda falha se apresentou a primeira;
- (b) **Falhas não-ordenadas de tipos diferentes:** são diferentes tipos de eventos, em que cada indivíduo pode passar por diferentes estados em qualquer ordem. Podem ser encontrados em estudos que acompanham a evolução da diabetes em que se registram os tempos até as ocorrências de algumas sequelas de interesse, como retinopatia, nefropatia, microangiopatia e neuropatia. Uma característica desses dados é que cada tipo de falha pode ocorrer somente uma vez para o mesmo indivíduo, não permitindo eventos recorrentes;

- (c) **Falhas não-ordenadas de mesmo tipo:** ocorrem em indivíduos que pertencem a algum tipo de agrupamento, como os estudos de famílias, em que cada membro da mesma está sob o risco de apresentar o evento de interesse. Os tempos de falha são eventos ocorrendo em órgãos similares de um mesmo indivíduo, como o tempo até o início de perda visual em cada olho;
- (d) **Modelo de riscos competitivos:** quando um único evento ocorre no indivíduo dentre vários possíveis eventos, ou seja, há interesse em mais de um tipo de falha mas a ocorrência da primeira inibe as outras falhas, como o estudo de mortalidade por causa do óbito;
- (e) **Modelo multiestado:** quando permite a ocorrência de eventos de diferentes tipos ou iguais, no mesmo indivíduo, ou seja, a ocorrência de um tipo de falha não inibe a ocorrência do mesmo tipo de falha e nem a ocorrência de outro tipo de falha. Um exemplo é o estudo para analisar os tempos até a infecção por fungos, bactérias ou vírus em transplantados.

2.2.1 A Função de Verossimilhança

Como a presença de tempos censurados é muito comum no contexto da análise de sobrevivência, suponha dados censurados à direita. Suponha também um esquema de censuras onde as mesmas ocorrem em diferentes tempos, de indivíduo para indivíduo, e que tal esquema sempre possa ser considerado estatisticamente independente do mecanismo que causa o evento de interesse no indivíduo (ou ocorrência da característica em observação), ou seja, um esquema de censuras independentes dos tempos de sobrevivência e das covariáveis.

Nesse esquema, ou o tempo exato de ocorrência Y_i da característica para o i -ésimo cliente é conhecido, ou apenas tem-se a informação de que este tempo é maior do que o tempo pré-fixado de censuras (L_i). Desta forma o conjunto \mathcal{L} agrupa essencialmente os valores observados para os tempos de sobrevivência T_i e as funções indicadoras de censuras δ_i , $i = 1, 2, \dots, n$ definidas por,

$$T_i = \begin{cases} Y_i, & \text{se } Y_i \leq L_i \\ L_i, & \text{se } Y_i > L_i \end{cases} \quad \text{e} \quad \delta_i = \begin{cases} 1, & \text{se } Y_i \leq L_i \\ 0, & \text{se } Y_i > L_i \end{cases}. \quad (2.9)$$

Nas populações homogêneas o conjunto \mathcal{L} é dado pelos pares (t_i, δ_i) , $i = 1, 2, \dots, n$; ao passo que para as populações heterogêneas o conjunto \mathcal{L} é dado por

$$\mathcal{L} = \{(t_i, \delta_i, \mathbf{x}_i), i = 1, 2, \dots, n\}, \quad (2.10)$$

com t_i o valor observado do tempo de sobrevivência T_i e δ_i a função indicadora de falhas da unidade i , com vetor de covariáveis \mathbf{x}_i , $i = 1, 2, \dots, n$.

Através das equações (2.1) e (2.7) e na presença de censuras, pelo esquema descrito acima, temos a função de verossimilhança dada por,

$$L(\theta; \mathbf{t}) = \prod_{i=1}^n [f(t_i; \theta)]^{\delta_i} [S(t_i; \theta)]^{(1-\delta_i)}. \quad (2.11)$$

Utilizando a representação (2.8), a função de verossimilhança (2.11) pode ser reescrita como

$$L(\theta; \mathbf{t}) = \prod_{i=1}^n [h(t_i)]^{\delta_i} \exp \left[- \int_0^{t_i} h(u) du \right]. \quad (2.12)$$

Para estimar os parâmetros do modelo, podemos fazer via maximização direta da equação (2.12), utilizando métodos numéricos, como o método de Newton-Raphson.

2.3 Modelos Paramétricos de Sobrevivência

Embora existam vários modelos probabilísticos, alguns ocupam maior destaque por sua comprovada adequação a várias situações reais, ou seja, por modelar de forma plausível os tempos de sobrevivência. Em particular, nesta seção, serão abordadas quatro das principais distribuições de probabilidade utilizadas na modelagem de dados de sobrevivência: a exponencial, a Weibull, a gama e a gama generalizada.

2.3.1 Modelo Exponencial

A distribuição exponencial é uma das mais simples e importantes distribuições de probabilidade utilizada na modelagem de dados que representam o tempo até a ocorrência de algum evento de interesse. Essa distribuição tem sido utilizada amplamente em pesquisas envolvendo tempo de sobrevivência ou tempo de remissão de doenças crônicas (ver por exemplo, Feigl & Zelen, 1965).

Uma variável aleatória não-negativa contínua T possui distribuição exponencial com parâmetro $v \geq 0$ (tempo médio de vida), se a sua função densidade de probabilidade e de sobrevivência forem, respectivamente,

$$f(t|v) = v \exp(-tv) \quad \text{e} \quad S(t) = \exp(-tv), \quad t \geq 0, v > 0. \quad (2.13)$$

Essa distribuição é caracterizada pela propriedade da “falta de memória”, ou seja, apresenta uma função de risco constante ao longo do tempo (Lawless, 1982). Portanto a função de risco é dada por,

$$h(t) = v, \quad \forall t \geq 0. \quad (2.14)$$

Na área de confiabilidade significa que um componente com 10000 horas de uso tem o mesmo risco de falha que um componente recém fabricado. Em muitas aplicações esse fato pode não ser realista, já que o envelhecimento do componente ou do equipamento pode influenciar na função de risco. Por exemplo, um computador pode ter risco de falha crescente, após um período inicial de funcionamento, com o tempo de uso. Nesse sentido surge então, a necessidade de obter uma distribuição que contemple este fato, que é o caso da distribuição Weibull.

2.3.2 A Distribuição Weibull

A distribuição Weibull foi proposta originalmente por Weibull (1951) e tem sido amplamente usada como modelo em confiabilidade e aplicações biométricas e industriais, por exemplo, em estudo do tempo de ocorrência de tumores em população humana (Whittmore & Altschuler, 1976) ou estudo em laboratórios de animais (Peto *et al.*, 1972), entre tantas outras situações.

Sua função densidade de probabilidade, de sobrevivência e de risco, são dadas respectivamente por,

$$f(t|v, \gamma) = v\gamma(vt)^{\gamma-1} \exp[-(vt)^\gamma], \quad S(t) = \exp[-(vt)^\gamma] \quad (2.15)$$

e

$$h(t) = v\gamma(vt)^{\gamma-1}, \quad t \geq 0, \quad \gamma \text{ e } v > 0, \quad (2.16)$$

onde γ é o parâmetro de forma e v é o parâmetro de escala.

Uma das características importantes desta distribuição, na modelagem de tempos de sobrevivência, está relacionada à sua flexibilidade em acomodar diferentes formas para a função de risco. Temos funções de risco monótonas decrescentes para o parâmetro de forma $\gamma < 1$; funções de risco monótonas crescentes para $\gamma > 1$; e risco constante para $\gamma = 1$ (caso particular, distribuição exponencial).

A Figura 2.2 (a) apresenta gráficos que mostram a grande variedade de formas da função de risco e a Figura 2.2 (b) mostra várias formas para a função de sobrevivência para o modelo Weibull, considerando diferentes valores de γ .

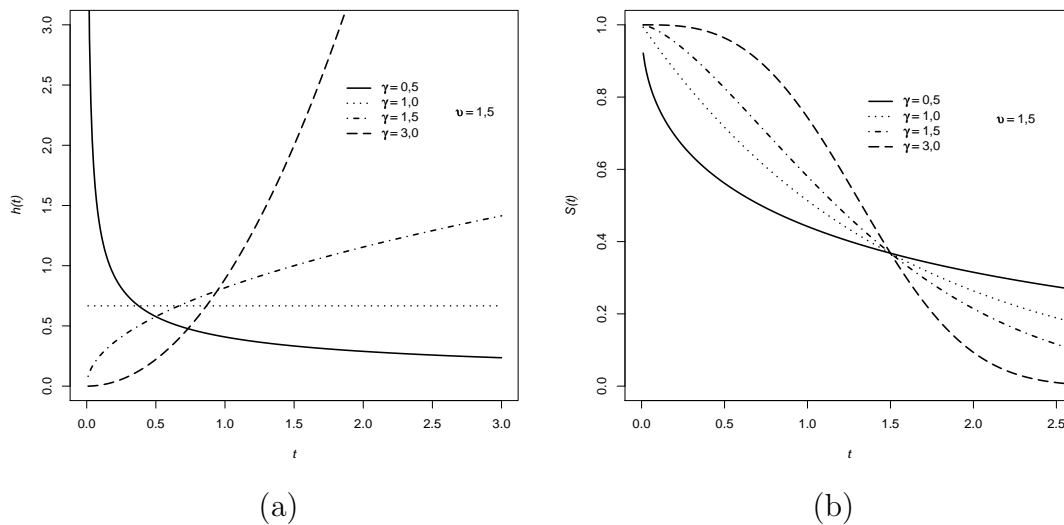


Figura 2.2: (a): Função de risco para diferentes valores de γ , (b): Função de Sobrevivência para diferentes valores de γ .

É importante ressaltar que o modelo de regressão Weibull é um caso particular do modelo de Cox (Cox, 1972), quando se toma $\lambda_0(t) = \frac{\gamma}{v^\gamma} t^{\gamma-1}$ na expressão:

$$\lambda(t) = \lambda_0(t)g(\mathbf{x}'\boldsymbol{\beta}), \quad (2.17)$$

em que g é uma função não-negativa que deve ser especificada, tal que $g(0) = 1$.

O modelo Weibull, na presença de covariáveis, é estendido a partir da densidade de probabilidade (2.15), assumindo que o parâmetro de escala v é dado por $G(\mathbf{x}_i, \boldsymbol{\beta}) = \exp(\mathbf{x}'_i \boldsymbol{\beta})$, com $\boldsymbol{\beta}$ o vetor de coeficientes desconhecidos associados ao vetor de covariáveis \mathbf{x} . Então o modelo Weibull, com covariáveis é definido pela função de risco condicional,

$$h(t|\mathbf{x}) = \gamma(t)^{\gamma-1} \exp(\mathbf{x}'_i \boldsymbol{\beta}), \quad t \geq 0. \quad (2.18)$$

Temos o efeito das covariáveis, na função de risco do modelo (2.18), como multiplicativo; mas é comum considerarmos uma transformação na variável aleatória T , tal que um modelo de locação-escala é obtido.

A transformação usada é dada por

$$Y = \ln(T) = \alpha + \gamma Z, \quad (2.19)$$

com $\sigma = \gamma^{-1}$, $\mu = -\log(\exp(\mathbf{x}'_i \boldsymbol{\beta})) = -\mathbf{x}'_i \boldsymbol{\beta}$ e Z uma variável aleatória com distribuição valor-extremo padrão, cuja densidade $f(z) = \exp(z - e^z)$, $-\infty < z < \infty$.

Na distribuição padronizada, temos a função densidade como sendo,

$$f(y) = \frac{1}{\sigma} \exp \left[\frac{y - \mu}{\sigma} - \exp \left(\frac{y - \mu}{\sigma} \right) \right], \quad (2.20)$$

em que σ e μ são os parâmetros de locação e escala, respectivamente, da distribuição.

2.3.3 As Distribuições Gama e Gama Generalizada

A distribuição gama tem sido usada frequentemente, em problemas de confiabilidade industrial e sobrevivência humana. Essa distribuição é caracterizada por dois parâmetros: γ e v . Quando $0 < \gamma < 1$, a taxa de risco decresce monotonicamente de infinito para v , com o tempo crescendo de 0 para infinito. Quando $\gamma > 1$ a taxa de risco cresce monotonicamente de 0 para v , com o tempo crescendo de 0 para infinito. Quando $\gamma = 1$, a taxa de risco é igual a v , constante (caso da exponencial).

Seja W_i uma variável aleatória com uma distribuição Gama(γ, v), com densidade,

$$f(w_i|\gamma, v) = \frac{v^\gamma}{\Gamma(\gamma)} w_i^{\gamma-1} \exp\{-vw_i\}, \quad (2.21)$$

com $v > 0$, $\gamma > 0$ e $w_i > 0$ para $i = 1, 2, \dots, n$.

Da equação (2.21) observamos que,

$$\int_0^{\infty} w_i^{\gamma-1} \exp\{-vw_i\} dw_i = \frac{\Gamma(\gamma)}{v^\gamma}. \quad (2.22)$$

A função de sobrevivência é dada por

$$S(w) = \int_w^{\infty} \frac{v^\gamma}{\Gamma(\gamma)} x^{\gamma-1} \exp\{-vx\} dx. \quad (2.23)$$

A esperança, a variância, a moda, o r-ésimo momento e a função geradora de momentos desta distribuição, são dadas respectivamente, por:

$$E(W) = \frac{\gamma}{v}, \quad (2.24)$$

$$Var(W) = \frac{\gamma-1}{v^2}, \quad \gamma \geq 1, \quad (2.25)$$

$$Mo(W) = \frac{\gamma-1}{v}, \quad \gamma \geq 1, \quad (2.26)$$

$$E(W^r) = \left(\frac{1}{v}\right)^r \frac{\Gamma(\gamma+1/v)}{\Gamma(\gamma)}, \quad (2.27)$$

$$E(e^{tw}) = (1-w/v)^\gamma, \quad w < v. \quad (2.28)$$

A figura 2.3 representa as funções de densidade da distribuição gama, para alguns valores de v .

Outros resultados para a distribuição gama podem ser encontrados no apêndice A.

A distribuição gama generalizada com parâmetros γ , v e ζ tem densidade dada por,

$$f(w_i|\gamma, v, \zeta) = \frac{\zeta}{\Gamma(\gamma)} v^{\zeta\gamma} w_i^{\zeta\gamma-1} \exp\{-[vw_i]^\zeta\}, \quad (2.29)$$

em que $w_i > 0$, $i = 1, 2, \dots, n$.

A esperança, a variância, a moda e o r-ésimo momento desta distribuição, são dadas respectivamente, por:

$$E(W) = \frac{\Gamma(\gamma+1/\zeta)}{v\Gamma(\gamma)}, \quad \gamma > -1/\zeta, \quad (2.30)$$

$$Var(W) = \frac{1}{v^2} \left\{ \frac{\Gamma(\gamma+2/\zeta)}{\Gamma(\gamma)} - \left[\frac{\Gamma(\gamma+1/\zeta)}{\Gamma(\gamma)} \right]^2 \right\}, \quad \gamma > -2/\zeta, \quad (2.31)$$

$$Mo(W) = \frac{(\gamma-1/\zeta)^{1/\zeta}}{v}, \quad \gamma > 1/\zeta, \quad (2.32)$$

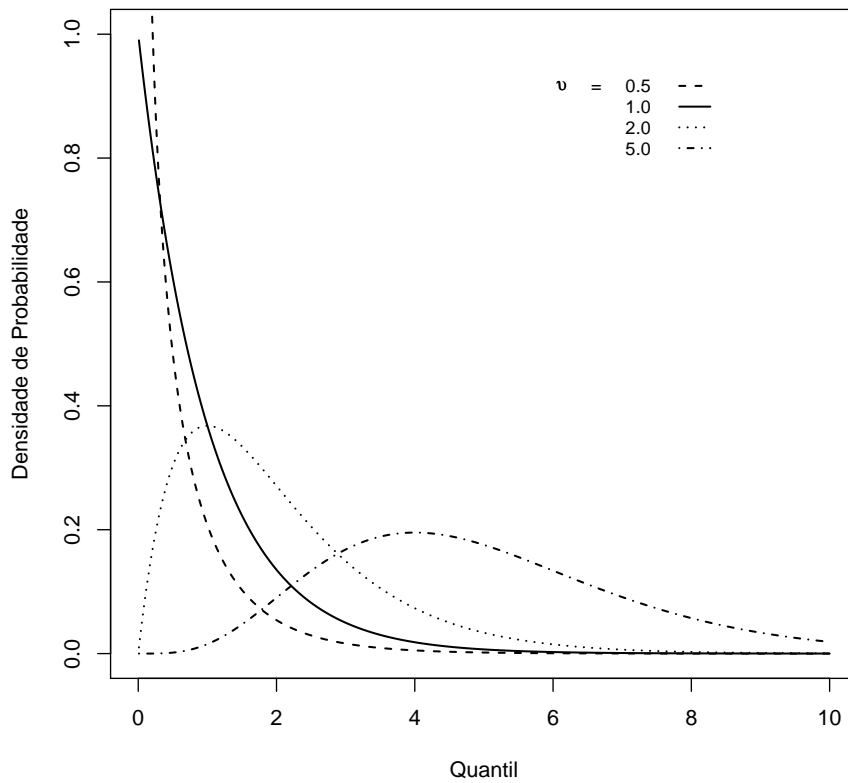


Figura 2.3: Funções de densidade da distribuição gama para diferentes valores de ν .

$$E(W^r) = \frac{1}{\nu^r} \frac{\Gamma(\gamma + r/\zeta)}{\Gamma(\gamma)}, \quad \gamma > -r/\zeta. \quad (2.33)$$

É fácil ver que a distribuição gama generalizada se reduz para a distribuição exponencial se $\zeta = 1$ e $\gamma = 1$; à distribuição Weibull se $\gamma = 1$; e à distribuição gama se $\zeta = 1$.

Além das distribuições exponencial, Weibull, gama e gama generalizada que são muito utilizadas na modelagem de dados de sobrevivência, temos outras distribuições consideradas na literatura; dentre elas podemos citar a log-normal, log-logística, Erlang e a F generalizada (ver maiores detalhes em Kalbfleisch & Prentice, 2002).

2.4 Dados de Eventos Recorrentes

Em muitas áreas de estudos, tratamos de situações onde alguns eventos de interesse não são terminais e podem ocorrer mais que uma vez para o mesmo indivíduo, temos então os chamados eventos recorrentes. Os dados para eventos recorrentes surgem em várias áreas, como biomedicina, criminologia, demografia e confiabilidade industrial. Por exemplo, um infrator pode reincidir no crime algumas vezes, vários tumores podem ser observados no mesmo paciente (recidivas), episódios de pneumonias podem recorrer em pacientes com síndrome de imunodeficiência humana e um equipamento pode falhar em repetidas experimentações.

Diferentemente do caso mais simples, em que se observa um único evento de interesse, e a partir do momento em que o mesmo acontece, o indivíduo deixa de ser observado. Para o caso de eventos recorrentes, o paciente continua sendo observado mesmo depois da ocorrência de cada evento.

Como as técnicas usuais não são adequadas para analisar este tipo de dados, muitos pesquisadores consideram como período da observação, apenas o tempo até o instante da primeira ocorrência do evento. Deste modo acabam desperdiçando dados que seriam obtidos durante todo acompanhamento, pois utilizam somente parte desta informação.

Este tipo de dados surge naturalmente em estudos longitudinais envolvendo múltiplos tempos de sobrevivência. Os dados sobre o i -ésimo indivíduo consiste no número total m_i de eventos (tempos de sobrevivência), observados no período de tempo $[0, \tau_i]$ e a época ordenada desses tempos de sobrevivência, no tempo $0 \leq t_{i1} < t_{i2} < \dots < t_{im_i} \leq T_i$. Além disso, podemos ter informações das covariáveis para cada paciente definido por um vetor \mathbf{x} , e um vetor dos indicadores de censuras. Interessa nestes estudos caracterizar o processo de ocorrência dos eventos para os indivíduos. Adicionalmente podemos estar interessados na comparação de tratamentos, com base no tempo distinto de cada indivíduo. A idéia é explicar a natureza da variação entre pacientes em termos do tratamento, das covariáveis e de outros fatores que podem não serem observáveis (ver por exemplo, Tomazella, 2003).

Existe uma extensa literatura sobre dados de eventos recorrentes. Se o interesse está na avaliação dos tratamentos ou nas covariáveis fixas, métodos marginais que focalizam a sobrevivência marginal média e/ou função da razão para eventos recorrentes são utilizados. Wei *et al.* (1989) propuseram modelos de riscos proporcionais separados para o tempo total de vida, T_j , com $j = 1, \dots, m_i$ (ou seja, tempo do início do experimento para o j -ésimo evento sendo independente do número de eventos). Pepe & Cai (1993) consideraram métodos gráficos para analisar os dados e o custo de sistemas de reparos,

onde a idéia principal foi propor um estimador não-paramétrico para a função da média cumulativa. Usando idéia similar, Lawless & Nadeau (1995) utilizaram Processos de Poisson para desenvolver modelos que enfocam o número esperado de eventos ocorridos em um determinado intervalo de tempo. Cook (1995) sugeriu modificações e discutiu os modelos de Poisson fazendo uma série de considerações partindo do ponto de vista clínico. Com a mesma abordagem, Cook & Lawless (1997) consideraram situações em que, além dos eventos recorrentes, existiam eventos terminais que impediam a ocorrência dos demais eventos e, portanto, induziam a censuras dependentes (ver por exemplo, Tomazella, 2003).

Entender como eventos ocorrem com o tempo é uma questão importante e que pode surgir quando temos dados de eventos recorrentes. Além disso, esses métodos não-paramétricos podem perder precisão relativa usando uma análise baseada em modelos completamente especificados, particularmente se a correlação entre os tempos para os eventos é grande (Oakes, 1986). Uma comparação de métodos marginais pode ser encontrado em Lawless & Nadeau (1995).

Outras abordagens alternativas têm sido desenvolvidas e propostas para tratarem eventos recorrentes (ver por exemplo, Hougaard, 2000; Ng & Cook, 1997; Turnbull *et al.*, 1997; Doganaksoy & Nelson, 1998). Apesar disso, os métodos estatísticos para a análise desses tipos de dados não estão completamente disseminados, e o assunto permite muitas discussões. Neste trabalho, os exemplos de aplicação são de estudos na área médica.

No primeiro exemplo, consideramos um conjunto de dados de sobrevivência introduzido por McGilchrist & Aisbett (1991) relacionado a infecção no rim, onde há recorrência da infecção de 38 rins de pacientes, usando máquinas portáteis de diálise, são registradas. Infecções podem ocorrer no local da inserção do cateter. O tempo anotado, chamado tempo de infecção, é o tempo de sobrevivência (em dias) do paciente até ocorrer a infecção e o cateter ter sido removido, ou tempo censurado, onde o cateter foi removido por outras razões. O cateter é reinsertado depois de algum tempo e o segundo tempo de infecção é novamente observado ou censurado (conjunto de dados na tabela 2.1).

Um segundo exemplo explorado nesta tese é sobre transplantes de medula óssea como um tratamento padrão para leucemia aguda. Prognósticos para a recuperação podem depender de fatores de risco conhecidos no tempo do transplante, tais como paciente e/ou idade e sexo do doador, o estágio inicial da doença, o tempo de prognóstico para o transplante, etc.

Neste estudo, foram observados 137 pacientes com leucemia mielóide aguda (LMA) e leucemia linfoblástica aguda; também conhecida como leucemia linfóide aguda (ou LLA). Os dados foram obtidos em 4 hospitais, onde os pacientes receberam tratamento (conjunto de dados introduzidos por Klein & Moeschberger, 1997, página 464).

Tabela 2.1: Tempos de recorrência das infecções de rins em 38 pacientes.

Paciente	Primeiro tempo	Segundo tempo	1 ^a censura	2 ^a censura	Sexo
1	8	16	1	1	1
2	23	13	1	0	2
3	22	28	1	1	1
4	447	318	1	1	2
5	30	12	1	1	1
6	24	245	1	1	2
7	7	9	1	1	1
8	511	30	1	1	2
9	53	196	1	1	2
10	15	154	1	1	1
11	7	333	1	1	2
12	141	8	1	0	2
13	96	38	1	1	2
14	149	70	0	0	2
15	536	25	1	0	2
16	17	4	1	0	1
17	185	117	1	1	2
18	292	114	1	1	2
19	22	159	0	0	2
20	15	108	1	0	2
21	152	562	1	1	1
22	402	24	1	0	2
23	13	66	1	1	2
24	39	46	1	0	2
25	12	40	1	1	1
26	113	201	0	1	2
27	132	156	1	1	2
28	34	30	1	1	2
29	2	25	1	1	1
30	130	26	1	1	2
31	27	58	1	1	2
32	5	43	0	1	2
33	152	30	1	1	2
34	190	5	1	0	2
35	119	8	1	1	2
36	54	16	0	0	2
37	6	78	0	1	2
38	63	8	1	0	1

[Censura (0); ocorrência da infecção (1); masculino (1); feminino (2)]

Para analisar esses dados podemos considerar diferentes modelos de sobrevivência multivariados introduzidos na literatura para análise de dados de tempos de vida dependentes, na presença de um vetor de covariáveis e observações censuradas; e propor novas formas de modelagem estatística, que será o objetivo desta tese.

Para capturar a correlação entre dois ou mais tempos de sobrevivência, podemos considerar a introdução de “fragilidades” ou variáveis latentes (veja por exemplo, Clayton & Cuzick, 1985; Oakes, 1986, 1989; Shih & Louis, 1992), assumindo modelos de riscos proporcionais.

2.5 Modelos de Regressão

É comum, a análise de sobrevivência estar centrada na relação entre o tempo de sobrevivência e variáveis explicativas de interesse no modelo; onde podemos estar interessados em investigar a influência das covariáveis nos tempos de sobrevivência. Uma forma de fazer isso é através de modelos de regressão, onde a dependência do tempo de sobrevivência e a covariável é explicitamente identificada. Nesses casos temos na literatura muitas técnicas para analisar estes dados de regressão, como por exemplo, técnicas não-paramétricas (ver por exemplo, Miller, 1976; Buckley & James, 1979); e técnicas paramétricas, na presença de dados censurados (ver por exemplo, Feigl & Zelen, 1965; Lawless, 1982).

A investigação da influência de covariáveis nos tempos de sobrevivência, em populações heterogêneas, é feita usualmente por meio da função de risco.

Um popular modelo de regressão semi-paramétrico para análise de dados de sobrevivência foi introduzido por Cox (1972) assumindo riscos proporcionais (veja também, Cox & Oakes, 1984). Nesse modelo, é assumido que os tempos de sobrevivência são independentes, ou seja, os indivíduos não são correlacionados.

Em muitas situações práticas, especialmente em estudos médicos, é comum termos tempos de sobrevivência dependentes, quando os indivíduos são relacionados um com o outro (mesma família, medidas repetidas no mesmo indivíduo ou duas ou mais medidas no mesmo paciente). Esses dados de sobrevivência são denominados de dados multivariados de sobrevivência (Kalbfleisch & Prentice, 2002; Hougaard, 2000).

Na literatura, normalmente esses tipos de dados são ajustados e analisados através de modelos de regressão (Fleming & Lin, 2000; Hougaard, 2000). Análises estatísticas para tempos de sobrevivência multivariados, usando uma distribuição multivariada específica é

feita usando estimação por máxima verossimilhança, que é uma extensão da metodologia paramétrica para tempos de falha independentes.

Do ponto de vista paramétrico, os modelos de sobrevivência são construídos a partir de dois componentes: aleatório e determinístico. O componente aleatório é representado por uma distribuição de probabilidade vinculada ao comportamento do tempo de sobrevivência, enquanto o componente determinístico é representado pelo relacionamento entre os parâmetros desta distribuição e as covariáveis.

Os modelos de regressão multiplicativos são definidos por

$$h(t_i; \mathbf{x}_i) = h_0(t_i) G(\mathbf{x}_i, \boldsymbol{\beta}) \quad (2.34)$$

onde $G(\mathbf{x}_i, \boldsymbol{\beta})$ representa o componente determinístico na formulação do modelo, $h_0(t)$ representa o componente aleatório. O termo $h_0(\cdot)$ é uma função positiva que descreve a dependência do risco sobre o tempo, é chamada função de risco de base, sendo comum a todos os indivíduos. O subscrito 0 reflete o risco para valores zero da covariável (é um risco basal, sendo uma referência para comparação com os riscos observados na presença das covariáveis), $\mathbf{x}'_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ é um vetor $(1 \times p)$ de covariáveis e $\boldsymbol{\beta}$ é um vetor p -dimensional de parâmetros desconhecidos de efeitos fixos associados a \mathbf{x}_i .

Para diferentes valores do vetor \mathbf{x} , as funções de risco são assumidas serem proporcionais, já que estes modelos pertencem à classe dos modelos de riscos proporcionais de Cox; onde (2.34) é o modelo multiplicativo mais conhecido.

Quando a função risco de base $h_0(t)$ é indexada por um parâmetro, digamos θ , estes modelos são considerados modelos de regressão paramétricos.

Contemplamos várias formas para a função $G(\cdot)$ na equação (2.34). Considerando, como exemplo, $G(\cdot) = \exp(\cdot)$ em (2.34), e a definição (2.5), podemos verificar que para a i -ésima unidade, a função de risco é

$$h(t_i; \mathbf{x}_i) = \frac{f(t_i; \mathbf{x}_i)}{S(t_i; \mathbf{x}_i)} = h_0(t_i) \exp\{\mathbf{x}'_i \boldsymbol{\beta}\} \quad (2.35)$$

onde $\mathbf{x}'_i \boldsymbol{\beta} = \eta_i$ é o preditor linear.

O modelo (2.35) implica que a razão de riscos para dois indivíduos, em qualquer tempo, depende das diferenças entre os preditores lineares. Integrando esta expressão nos segundo e terceiro termos, temos

$$\int_0^{t_i} \frac{f(u_i)}{1 - F(u_i)} du_i = \left(\int_0^{t_i} h_0(u_i) du_i \right) \exp\{\mathbf{x}'_i \boldsymbol{\beta}\}$$

aplicando o logaritmo, obtemos

$$-\log(1 - F(t_i)) = -H_0(t_i) \exp\{\mathbf{x}'_i \boldsymbol{\beta}\} \quad (2.36)$$

onde $H_0(t_i) = \int_0^{t_i} h_0(u_i) du_i$ é a função de risco acumulada.

Aplicando exponencial na expressão acima, encontramos a função de sobrevivência do i -ésimo indivíduo, dada por

$$S(t_i; \mathbf{x}_i) = \exp \{ \log(1 - F(t_i)) \} = \exp \{ -H_0(t_i) \exp(\mathbf{x}'_i \boldsymbol{\beta}) \} \quad (2.37)$$

$i = 1, 2, \dots, n$.

A aplicação do modelo de Cox tem levado, quase que exclusivamente a considerações de riscos proporcionais; que é relevante em muitas situações reais, mas em muitos casos temos interesse onde o risco não é proporcional. Neste caso, uma possível alternativa é a classe de modelos de regressão em análise de sobrevivência, conhecida por modelos aditivos.

Desta forma, quando o efeito das covariáveis é expresso aditivamente na função de risco, os modelos de regressão aditivos são definidos pela seguinte função de risco na unidade i de \mathcal{D} (conjunto de dados de sobrevivência, com n unidades independentes e p covariáveis observadas) no instante $t \in T$,

$$h(t_i; \mathbf{x}_i, \boldsymbol{\beta}) = h_0(t_i) + G(\mathbf{x}_i(\mathbf{t}), \boldsymbol{\beta}) \quad (2.38)$$

onde $h_0(\cdot)$ é a função de risco de base, $G(\cdot)$ é uma função positiva escolhida usualmente como $G(x) = x$, e $\boldsymbol{\beta}$ é um vetor p -dimensional de parâmetros desconhecidos de efeitos fixos associados a \mathbf{x}_i .

Este modelo é o mais conhecido dos modelos aditivos de Alen (1980), onde mesmo com as covariáveis que não são observadas ao longo do tempo, ele permite averiguar a influência dessas covariadas variando com o tempo.

Temos modelos de regressão em análise de sobrevivência, apresentados numa formulação mais geral, tanto para modelos multiplicativos e aditivos, (ver por exemplo, Lin & Ying, 1995).

2.6 Modelando as Fragilidades

Usualmente em análise de sobrevivência, supomos que os efeitos dos tratamentos ou dos riscos nos indivíduos são condicionalmente independentes, dadas as covariáveis observadas. Normalmente essa suposição é bem plausível, dependendo de como o estudo

é planejado, mas quando temos eventos múltiplos ou repetidos dentro de um mesmo indivíduo, essa suposição de independência passa a ser questionável e devemos levar isso em consideração na estimação do efeito do tratamento.

Em problemas de eventos recorrentes entretanto, é razoável assumir que exista dependência entre as observações (Cox, 1972). Motivado por situações onde existe tal dependência entre indivíduos ou grupos de indivíduos, neste trabalho utilizaremos uma das possíveis maneiras de solucionar o problema, que é através de modelos que incorporam *fragilidade* (*frailty models*), introduzido por Vaupel *et al.* (1979). Clayton (1978) e Oakes (1989) consideraram o primeiro modelo de fragilidade para dados multivariados. Posteriormente alguns outros autores usaram esse modelo, onde podemos citar, Self & Prentice (1986); Ng & Cook (1997) e Andersen *et al.* (1993) com uma primeira abordagem Bayesiana para esses tipos de modelos.

Como extensão do modelo semi-paramétrico de Cox (1972), destacam-se as abordagens de *fragilidade* e variância robusta, que permitem introduzir as distribuições marginais dos tempos de falha de formas arbitrária, individuais ou de grupos. O modelo de *fragilidade* é caracterizado pela utilização de um efeito aleatório, ou seja, de uma variável aleatória não-observável, que representa as informações que não podem ou não foram observadas, como por exemplo em dados longitudinais (elementos abandonaram o estudo por mudarem de localidade), levantamentos amostrais ou em estudos de epidemiologia. No modelo com variância robusta, esta correlação não é modelada especificamente, mas costuma-se levar em consideração a estimação robusta da matriz de variância dos estimadores do vetor que contém os parâmetros do modelo.

Os modelos com *fragilidade* podem ser considerados como modelos com efeitos aleatórios, com duas fontes de variação: a variabilidade natural explicada pela função de risco e a variabilidade comum para os indivíduos pertencentes ao mesmo grupo ou variabilidade comum para alguns eventos de um dado indivíduo, que é explicada pelas *fragilidades*.

Para dados multivariados, estes modelos também podem explicar a dependência entre as unidades ou observações pertencentes ao mesmo grupo.

Diante de um conjunto de dados de sobrevivência com n unidades e p covariáveis observadas, podemos encontrar uma heterogeneidade em cada unidade (individual) causada pela não observação de alguns fatores de riscos. Por exemplo, Hougaard (1986b), ao analisar doentes após um enfarte, observou que o risco de morte era inicialmente bastante alto, diminuindo posteriormente. A interpretação biológica é que, passada a crise, uma parte do músculo do coração morre devido à falta de oxigênio e o órgão se ajusta à capacidade reduzida. Alternativamente, o risco de morte decrescente pode ser explicado pela heterogeneidade não observada (*fragilidade*). A introdução da *fragilidade* é feita na função de risco de cada unidade, incorporando um termo de *fragilidade* w_i que é responsável pela

heterogeneidade não observada da unidade i ; $i = 1, 2, \dots, n$.

As *fragilidades* podem ser introduzidas de forma multiplicativa (ver por exemplo, Clayton, 1978); ou de forma aditiva na função de risco (ver por exemplo, da Silva, 2001; Tomazella, 2003), visando assim responder as diferentes formas de avaliar a influência da heterogeneidade entre as unidades na função de risco, ou intensidade dos processos de contagem.

Em geral, as *fragilidades* são assumidas independentes e identicamente distribuídas (iid) para cada grupo ou indivíduo. É comum na literatura assumir a distribuição gama para os efeitos aleatórios não observados (ver por exemplo, Oakes, 1989; Clayton, 1991). Isto normalmente ocorre por conveniência matemática, uma vez que para assegurar a identificabilidade do modelo, as *fragilidades* têm média 1.

Considerando o modelo com a presença de covariáveis fixas e de efeitos aleatórios não observados, temos usando (2.34) que

$$h(t_{ij}|x_{ij}, w_i) = h_0(t_{ij}) \exp\{\mathbf{x}'_{ij}\boldsymbol{\beta}\}w_i,$$

para $i = 1, 2, \dots, l$ e $j = 1, 2, \dots, n_i$. Temos que $h_0(t_{ij})$ é a função risco de base e w_i representa a *fragilidade*.

Aslanidou *et al.* (2004) consideraram *fragilidades* univariadas para os grupos com $W_i \sim \text{Gama}(1/k, 1/k)$, onde k é a variância dos W_i (desconhecida) e a distribuição a priori também gama. Sinha (1998) considera $W_i \sim \text{Gama}(k^{-1}, k^{-1})$, com W_i representando a fragilidade da i -ésima unidade. Em ambos os casos, valores altos de k , representam uma correlação positiva entre as unidades do mesmo grupo e maior heterogeneidade entre as de grupos diferentes.

Outros autores usaram diferentes distribuições para as *fragilidades*, como Hougaard (1986a) e Hougaard (1986b) que usaram uma distribuição estável positiva. Esta distribuição tem a propriedade de preservar a proporcionalidade entre as funções de risco dos modelos multiplicativos. Para estudar as influências genéticas e ambientais na morte prematura de adultos adotivos, Dias (2002) utilizou um modelo com *fragilidade* aditiva correlacionada, que considerou a heterogeneidade e dependência entre unidades ou grupos e permitiu modelar conjuntos de dados com vários graus de associação. Andersen *et al.* (1993) e Hougaard (1995) apresentaram uma revisão dos modelos de *fragilidade* multiplicativos numa perspectiva frequentista, enquanto Shih & Louis (1992) elaboraram um revisão completa desses modelos sob um enfoque Bayesiano.

Várias propostas têm sido feitas para introduzir efeitos aleatórios nos modelos para análise de dados multivariados de sobrevivência (ver por exemplo, Pickles & Crouchley, 1995; Hougaard, 2000), como o modelo de *fragilidade* compartilhada (*shared frailty model*) discutidos em da Silva Matuda (2005), em que é focalizado a análise de eventos agrupados.

Clayton (1991) usa um processo de Levy (Kalbfleisch, 1978) como um modelo Bayesiano não-paramétrico para o risco base, aplicado a dados contínuos, ou seja, dados com nenhuma restrição.

Capítulo 3

Métodos Bayesianos

Apresentamos neste capítulo alguns conceitos gerais de análise Bayesiana, que é uma alternativa para a modelagem de dados em situações geralmente mais complexas.

Na inferência clássica um dos objetivos costuma ser em determinar que generalizações podemos fazer (se algumas são possíveis) sobre a população de interesse à partir da amostra colhida da mesma. Desejamos estimar o parâmetro θ , com $\theta \in \Theta$ um escalar ou vetor desconhecido, mas fixo. No modelo Bayesiano o parâmetro θ é um escalar ou vetor aleatório (não observável), onde toda incerteza deve ser quantificada em termos de probabilidade. Então, a inferência Bayesiana se baseia em probabilidades subjetivas ou credibilidades a posteriori associadas com diferentes valores dos parâmetros θ e condicionadas pelo particular valor de y observado.

A quantidade θ serve como indexador da família de possíveis distribuições para as observações y , representando características de interesse que se deseja conhecer para ter uma descrição completa do processo.

A estimação paramétrica na inferência Bayesiana pode encontrar problemas com modelos mais complexos, que podem resultar em distribuições a posteriori de difícil tratamento analítico, como por exemplo, a presença de um grande número de parâmetros a ser estimados, a dificuldade na obtenção das densidades marginais de forma analítica, ou ainda distribuições a priori e a posteriori que não são conjugadas. Nesses casos, os métodos analíticos de aproximação não são indicados, sendo necessários métodos de aproximação numérica para estimação dos parâmetros de interesse. Os métodos analíticos de aproximação são mais eficientes computacionalmente, porém são baseados em resultados assintóticos e na suposição de normalidade. Também, são mais difíceis para programar à

medida que o número de parâmetros aumenta.

Os métodos de aproximação mais utilizados são os métodos de aproximação de Laplace (Tierney *et al.*, 1989), a aproximação pela distribuição normal e o uso de quadraturas Gaussianas.

Nas últimas décadas, vários métodos numéricos têm sido desenvolvidos para resolver as questões mencionadas acima. Posteriormente falaremos sobre alguns desses métodos mais utilizados, em especial métodos de simulação de amostras da distribuição a posteriori de interesse.

O procedimento Bayesiano é baseado no teorema de Bayes, ou seja, dados os tempos de falha $\mathbf{y}' = (y_1, y_2, \dots, y_n)$ obtidos de um modelo paramétrico $f(\mathbf{y}|\boldsymbol{\theta})$, $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_k)$, temos

$$p(\mathbf{y}|\boldsymbol{\theta}) = \frac{f(\boldsymbol{\theta}|\mathbf{y}) \pi(\boldsymbol{\theta})}{\int f(\boldsymbol{\theta}|\mathbf{y}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}}, \quad (3.1)$$

em que $p(\boldsymbol{\theta}|\mathbf{y})$ é a distribuição a posteriori de $\boldsymbol{\theta}$ quando $\mathbf{Y} = \mathbf{y}$, combinando a função de verossimilhança $f(\boldsymbol{\theta}|\mathbf{y})$ com a priori $\pi(\boldsymbol{\theta})$.

Se nos interessa algum componente de $\boldsymbol{\theta}$, digamos θ_i ($i = 1, \dots, k$), obtemos a distribuição marginal integrando a distribuição conjunta $p(\boldsymbol{\theta}|\mathbf{y})$, ou seja,

$$p(y|\theta_i) = \int_{\Theta_{-i}} f(\boldsymbol{\theta}|\mathbf{y}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}_{-i}, \quad (3.2)$$

onde o subscrito $(-i)$ implica a integração de todos os componente exceto θ_i .

Os cálculos de (3.2) nem sempre são fáceis, havendo muitas vezes necessidade de recorrer a procedimentos numéricos. Esses procedimentos se baseiam em amostragem que englobem métodos de integração simples de Monte Carlo, métodos de reamostragem por importância e métodos de Monte Carlo via Cadeias de Markov (MCMC). Estes são mais simples para implementação e não apresentam restrições quanto ao número de parâmetros a serem estimados.

Entre os métodos MCMC, os mais utilizados são o amostrador de Gibbs (*Gibbs Sampling*) e Metropolis-Hastings (ver por exemplo, Gamerman, 1997). É importante observar que o uso destes algoritmos e métodos MCMC em geral, são necessários se a geração não-iterativa da distribuição da qual se deseja obter uma amostra for muito complicada ou custosa.

3.1 Monte Carlo via Cadeias de Markov (MCMC)

A idéia básica por trás desse método é a de transformar o problema estático que temos, num problema de natureza dinâmica, construindo para isso um processo estocástico temporal artificial, que seja fácil de simular e que convirja, fracamente, para a distribuição original. Este processo temporal, em geral, é uma cadeia de Markov homogênea cuja distribuição de equilíbrio é a distribuição que pretendemos simular. Para os Bayesianos, é a densidade à posteriori $p(\boldsymbol{\theta}|\mathbf{x})$.

Existem dois métodos para gerar cadeias de Markov com distribuição estacionária especificada. O primeiro, que tem sido usado por muitos anos em física estatística é o algoritmo de Metropolis (Metropolis *et al.*, 1953), que mais tarde foi generalizado por Hastings (1970). O outro método é o *Gibbs Sampling* (ver por exemplo, Geman & Geman, 1984), que foi trazido dentro da influência da literatura estatística por Gelfand & Smith (1990).

A seguir, descrevemos alguns desses algoritmos comumente utilizados.

3.1.1 Amostrador de *Gibbs* (*Gibbs Sampling*)

O amostrador de Gibbs foi desenvolvido por Geman & Geman (1984), como uma ferramenta para modelos de processamento de imagens; é uma técnica usada para gerar variáveis aleatórias de uma distribuição, sem utilizar diretamente sua densidade de probabilidade. Este é um esquema Markoviano de atualização no processo de amostragem da distribuição conjunta dos parâmetros, via amostragem iterativa das distribuições condicionais completas. Gelfand & Smith (1990) apresentaram aplicações interessantes deste procedimento em uma grande variedade de problemas estatísticos. Casella & George (1992) ilustraram a teoria com dois casos: um bivariado e outro com dimensão maior do que 2.

Como descrição geral do método, que é essencialmente um processo iterativo de amostragem de uma cadeia de Markov com núcleo de transição formado pelas condicionais completas, considere que a densidade a posteriori conjunta, suprimindo a dependência dos dados, seja denotada por $\pi(\theta_1, \theta_2, \dots, \theta_k)$ e as distribuições condicionais completas para que os k parâmetros estejam disponíveis (ver por exemplo, Tomazella, 2003). Essas distribuições são representadas por

$$\pi(\theta_1|\theta_2, \theta_3, \dots, \theta_{k-1}, \theta_k), \pi(\theta_2|\theta_1, \theta_3, \dots, \theta_k), \dots, \pi(\theta_k|\theta_1, \theta_2, \dots, \theta_{k-1})$$

Assim, o algoritmo (amostrador de Gibbs) é dado por:

i- Escolha valores iniciais $\theta_2^{(0)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}$

ii- Inicialize o contador da cadeia $r = 1$ e gere um valor $\theta_1^{(1)}$ da densidade condicional

$$\pi(\theta_1|\theta_2^{(0)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}).$$

iii- Similarmente, gere um valor $\theta_2^{(1)}$ da densidade condicional

$$\pi(\theta_2|\theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_k^{(0)}).$$

iv- Proceda sucessivamente até o valor $\theta_k^{(1)}$ da densidade condicional

$$\pi(\theta_k|\theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{k-1}^{(1)})$$

Então com essa nova realização $\theta^{(1)}$ para o vetor $\boldsymbol{\theta}$, substituir os valores iniciais no esquema iterativo. Atualize o contador r para $r + 1$ e repita este procedimento por B vezes. Sob condições de regularidade, como apresentado por Geman & Geman (1984), temos que $\theta_i^{(r)}$ tende em distribuição a $p(\theta_i)$, isto é, para r grande assume-se que a convergência é atingida em uma iteração cuja distribuição esteja arbitrariamente próxima da distribuição de equilíbrio. Portanto $\theta_i^{(r)}$ pode ser considerado (aproximadamente), como uma observação simulada da distribuição marginal de θ_i , $\pi(\theta_i)$, $i = 1, 2, \dots, k$. Replicações paralelas independentes de todo o algoritmo acima N vezes produz N conjuntos de vetores dos parâmetros e então, para cada elemento de $\boldsymbol{\theta}$, obtemos uma amostra simulada de tamanho N da distribuição a posteriori conjunta.

3.1.2 Algoritmo de Metropolis-Hastings

O algoritmo de Metropolis (Metropolis *et al.*, 1953) foi desenvolvido com o objetivo de estudar as propriedades de equilíbrio de grandes sistemas de partículas, tais como elétrons em um átomo. Hastings (1970) sugeriu uma generalização deste algoritmo, ilustrando o seu uso na simulação de desvios Normais e Poisson, e de matrizes aleatórias ortogonais.

Este algoritmo pode ser utilizado em alguma etapa do amostrador de Gibbs, quando as distribuições condicionais a posteriori não são facilmente identificadas como possuidoras de uma forma padrão (por exemplo Normal, gama, etc.), que impossibilita a geração direta a partir destas distribuições.

Como descrição geral para o caso de θ ser um escalar, considere uma densidade de transição (ou densidade proposta), $q(\theta, \theta^*)$. Esta densidade de transição não possui necessariamente probabilidade de equilíbrio π , porém deve representar uma regra de passagem que defina uma cadeia. Considere também a probabilidade de aceitação $p(\theta^{(r-1)}, \theta^*)$ definida abaixo. Uma amostra de $\pi(\theta)$ é obtida assim:

Passo 0: escolha um valor inicial $\theta = \theta^{(0)}$ e inicialize o contador da cadeia $r = 1$;

Passo 1: amostre θ^* da densidade proposta $q(\theta^{(r-1)}, \theta^*)$;

Passo 2: amostre u de uma distribuição Uniforme $(0, 1)$. Aceite o valor θ^* gerado se

$$u \leq p(\theta^{(r-1)}, \theta^*) = \min \left\{ 1, \frac{\pi(\theta^*) q(\theta^*, \theta^{(r-1)})}{\pi(\theta^{(r-1)}) q(\theta^{(r-1)}, \theta^*)} \right\}$$

Caso contrário, fique com $\theta^{(r)} = \theta^{(r-1)}$;

Passo 3: atualize o contador r para $r + 1$ e volte ao *passo 1*, até a convergência ser alcançada.

Para r suficientemente grande, $\theta^{(r)}$ é uma amostra da distribuição a posteriori $\pi(\theta)$.

Para o caso vetorial $\boldsymbol{\theta}' = (\theta_1, \theta_2, \dots, \theta_p)$, teremos a densidade de transição dada por $q(\boldsymbol{\theta}, \boldsymbol{\theta}^*)$ e a probabilidade de aceitação dada por $p(\boldsymbol{\theta}^{(r-1)}, \boldsymbol{\theta}^*)$. Depois, procedemos analogamente ao algoritmo de Metropolis-Hastings.

3.2 Critérios de Comparação de Modelos

Um problema importante, consiste na avaliação e escolha do modelo que melhor represente a situação em estudo. Como uma situação especial, podemos usar um super modelo como a distribuição gama generalizada (ver seção 2.3.3). Neste caso, se o intervalo de credibilidade para o parâmetro γ inclui o valor um, isto é uma indicação que o uso da distribuição Weibull, na presença de “fragilidade”, oferece um bom ajuste para os dados de sobrevivência.

Nesta tese utilizamos o critério DIC, que é um critério especificamente usado para a

seleção de modelos na metodologia Bayesiana, onde as amostras da distribuição a posteriori para os parâmetros do modelo, são obtidas usando os métodos MCMC.

A *deviance* é definida por,

$$D(\boldsymbol{\theta}) = -2 \log L(\boldsymbol{\theta}) + c, \quad (3.3)$$

com $\boldsymbol{\theta}$ é um vetor de parâmetros desconhecidos do modelo, $L(\boldsymbol{\theta})$ a função de verossimilhança do modelo e c é uma constante que não necessita ser conhecida quando é feita a comparação entre modelos.

O critério DIC definido por Spiegelhalter *et al.* (2002) é dado por,

$$DIC = D(\hat{\boldsymbol{\theta}}) + 2n_D, \quad (3.4)$$

em que $D(\hat{\boldsymbol{\theta}})$ é a deviance avaliada na média a posteriori $\hat{\boldsymbol{\theta}} = E(\boldsymbol{\theta}|\text{dados})$ e n_D é o número efetivo de parâmetros do modelo, dado por $n_D = \bar{D} - D(\hat{\boldsymbol{\theta}})$, onde $\bar{D} = E(D(\boldsymbol{\theta})|\text{dados})$ é a *deviance* a posteriori medindo a qualidade dos dados fixados para o modelo. Pequenos valores de DIC indicam melhores modelos. Note que esses valores podem ser negativos.

Vale ressaltar que estes métodos têm como objetivo englobar a complexidade do modelo no critério de seleção, pois são critérios que “penalizam” a verossimilhança (Paulino *et al.*, 2003).

Também podemos usar a preditiva condicional ordinal (CPO) (ver por exemplo, Carlin & Louis, 1997) dada por,

$$c_i(l) = \int f^{(l)}(t_{1i}, t_{2i}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}|\mathbf{t}_1, \mathbf{t}_2)d\boldsymbol{\theta}, \quad (3.5)$$

em que $f^{(l)}(t_{1i}, t_{2i}|\boldsymbol{\theta})$ é a função de verossimilhança para o vetor de parâmetros $\boldsymbol{\theta}$ associado ao modelo utilizado considerando a observação (t_{1i}, t_{2i}) , $i = 1, 2, \dots, n$ e $\pi(\boldsymbol{\theta}|\mathbf{t}_1, \mathbf{t}_2)$ é a distribuição a posteriori conjunta para $\boldsymbol{\theta}$. Observar que estamos considerando o caso especial de dados bivariados, mas os resultados são válidos para qualquer dimensão.

Um estimador de Monte Carlo para (3.5) baseado em S amostras geradas por Gibbs é dada por,

$$\hat{c}_i(l) = \frac{1}{S} \sum_{s=1}^S f^{(l)}(t_{1i}, t_{2i}|\boldsymbol{\theta}^{(s)}). \quad (3.6)$$

Grandes valores de $\hat{c}_i(l)$ (em média) indicam melhores modelos. Também podemos considerar a soma dos logaritmos de $\hat{c}_i(l)$, ou seja, $\hat{c}_i(l) = \sum_{i=1}^n \log(\hat{c}_i(l))$ como um critério para a escolha do melhor modelo.

3.3 Modelos Hierárquicos

Em um modelo de Bayesiano hierárquico, ao invés de especificar a distribuição a *priori* como uma função simples, ela é especificada em uma hierarquia (ver por exemplo, Paulino *et al.*, 2003).

Normalmente, o modelo de Bayes, de forma geral é dado por

$$\begin{aligned} X|\theta &\sim f(x|\theta) \\ \Theta|\xi &\sim \pi(\theta|\xi); \end{aligned} \tag{3.7}$$

então, condicionalmente em θ , X tem densidade amostral $f(x|\theta)$; e condicionalmente em ξ , Θ tem densidade a priori $\pi(\theta|\xi)$.

Do modelo (3.7), calculamos a distribuição a posteriori $\pi(\theta|x, \xi)$, usando a fórmula de Bayes.

No modelo hierárquico especificamos a distribuição a *priori* de forma hierárquica, colocando outro nível (ou estágio) no modelo; então temos,

$$\begin{aligned} X|\theta &\sim f(x|\theta) \\ \Theta|\xi &\sim \pi(\theta|\xi) \\ \xi &\sim \psi(\xi), \end{aligned} \tag{3.8}$$

onde assumimos que $\psi(\cdot)$ é conhecido e não depende de nenhum outro hiperparâmetro desconhecido.

A metodologia Bayesiana hierárquica modela a incerteza nos hiperparâmetros através de novas distribuições, induzindo assim, uma decomposição da distribuição a *priori* em dois ou mais níveis.

O modelo Bayesiano considerado em (3.8) é definido através de uma hierarquia de dois níveis, traduzindo o primeiro nível da distribuição amostral dos dados e o segundo a distribuição a *priori* para o parâmetro θ . Existem situações onde surgem de forma natural estruturas hierárquicas para quantidades observáveis, correspondendo a uma decomposição do modelo amostral em dois ou mais níveis.

Formalmente, o modelo Bayesiano hierárquico é um modelo Bayesiano $\{f(x|\theta), \pi(\theta)\}$, onde $\pi(\theta)$ pode ser decomposto nas condicionais $\pi_1(\theta|\alpha_1), \pi_2(\alpha_1|\alpha_2), \dots, \pi_{l-1}(\alpha_{l-2}|\alpha_{l-1})$

e na distribuição marginal $\pi_1(\alpha_{l-1})$, de modo que,

$$c(\theta) = \int_l \pi(\theta | \alpha_1), \pi_2(\alpha_1 | \alpha_2), \dots, \pi_{l-1}(\alpha_{l-2} | \alpha_{l-1}) d\alpha_1 \dots d\alpha_{l-1}, \quad (3.9)$$

em que α_i denota os hiperparâmetros de nível $i = 2, 3, \dots$ e a integral é estendida ao conjunto dos valores possíveis para $(\alpha_1, \alpha_2, \dots, \alpha_{l-1})$.

Esta decomposição da distribuição a priori é, geralmente justificada pela dificuldade de se quantificar exatamente a informação dessa distribuição e pelo interesse em incorporar a incerteza decorrente sobre os hiperparâmetros.

Em geral, a distribuição a *priori* do primeiro nível é uma distribuição conjugada natural do modelo amostral por motivos de flexibilidade computacional. A dificuldade natural em especificar os hiperparâmetros, que é maior quanto mais elevado o nível de hierarquia, torna comum adotar para estas hierarquias distribuições não informativas em níveis mais elevados, já que raramente é útil na prática estender a hierarquia a mais do que dois níveis. Deve-se todavia acrescentar que más especificações nas distribuições além do primeiro nível têm muito menor impacto inferencial do que se tal ocorrer na distribuição a *priori* padrão.

Uma aplicação interessante do conceito de hierarquia é quando a informação a *priori* disponível só pode ser convenientemente resumida através de uma mistura de distribuições. Isto implica em considerar uma distribuição discreta para

$$\pi(\theta | \xi) = \sum_{i=1}^k \pi(\theta | \xi_i) \pi(\xi_i). \quad (3.10)$$

Verificamos que a distribuição a posteriori de θ é também uma mistura pois,

$$\pi(\theta | x) = \frac{\pi(\theta | \xi) f(x | \theta)}{\int \pi(\theta | \xi) f(x | \theta) d\theta} = \frac{\sum_{i=1}^k f(x | \theta) \pi(\theta | \xi_i) \pi(\xi_i)}{\sum_{i=1}^k \pi(\xi_i) \int f(x | \theta) \pi(\theta | \xi_i) d\theta}. \quad (3.11)$$

Note que a distribuição a posteriori condicional de θ dado ξ_i é obtida como sendo

$$\pi(\theta | x, \xi_i) = \frac{f(x | \theta) \pi(\theta | \xi)}{\int f(x | \theta) \pi(\theta | \xi) d\theta} = \frac{f(x | \theta) \pi(\theta | \xi_i)}{m(\theta | \xi_i)}, \quad (3.12)$$

e, portanto, $f(x | \theta) \pi(\theta | \xi) = \pi(\theta | x, \xi_i) m(\theta | \xi_i)$. Ainda temos que $m(\theta | \xi_i) \stackrel{\mathcal{D}}{\simeq} \int f(x | \theta) \pi(\theta | \xi) d\theta$.

Assim, podemos escrever a posteriori de θ como

$$\pi(\theta|x) = \frac{\sum_{i=1}^k \pi(\theta|x, \xi_i) m(\theta|\xi_i) \pi(\xi_i)}{\sum_{i=1}^k m(\theta|\xi_i) \pi(\xi_i)} = \sum_{i=1}^k \pi(\theta|x, \xi_i) \pi(\xi_i|x). \quad (3.13)$$

Note que $\pi(\xi_i|x) = \frac{m(\theta|\xi_i) \pi(\xi_i)}{\sum_{i=1}^k m(\theta|\xi_i) \pi(\xi_i)}$, isto é, a distribuição preditiva é uma mistura de preditivas condicionais.

3.4 Diagnósticos de Convergência

Neste campo temos uma extensa investigação, onde tem sido propostas várias estratégias para abordar o problema da avaliação da convergência do método de Gibbs (ver por exemplo, Cowles & Carlin, 1996).

Os problemas teóricos associados à avaliação da convergência do método de Gibbs, num contexto geral, têm a ver com a convergência dos métodos MCMC. Esses métodos são ótimas ferramentas para resolver diversos problemas reais, mas algumas dificuldades surgem para utilizá-los, tais como: número necessário de iterações para se obter a convergência, as iterações iniciais da amostra podem ser influenciadas pelos valores iniciais dos parâmetros e as sequências de valores podem apresentar correlação entre os parâmetros. Não existe uma técnica geral para resolver essas questões, porém, existem métodos de verificação de convergência que são baseados nas propriedades da cadeia de Markov, para indicar ou não, a convergência da amostra simulada para a distribuição marginal.

Uma avaliação da convergência pode ser feita preliminarmente, analisando os gráficos ou as medidas descritivas, obtidas a partir dos valores simulados para os parâmetros de interesse. Os gráficos mais frequentes são aqueles que usam a quantidade de interesse estimada ao longo das iterações e um gráfico da estimativa da distribuição marginal a posteriori deste parâmetro (ver por exemplo, Dias, 2002).

As medidas descritivas são a média, o desvio-padrão e os quantis. Uma outra avaliação da convergência é feita usando-se algumas técnicas de diagnóstico desta convergência. As técnicas de diagnóstico mais populares são descritas por Gelman & Rubin (1992), Raftery & Lewis (1992) e Heidelberger (1983). Estas técnicas estão implementadas na biblioteca CODA (*Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output*) (Best, N. and Cowles, M. and Vines, K., 1997), do *software WinBUGS* (ver por exemplo,

Spiegelhalter, D. J.; Thomas, A.; Best, N. G.; Gilks, W. R., 2003).

Geralmente, esses métodos são utilizados conjuntamente a fim de obtermos uma indicação de convergência, uma vez que, nenhum deles é considerado melhor do que o outro. Os autores do CODA dizem que Cowles & Carlin (1996) testaram alguns dos diagnósticos desta biblioteca e encontraram alguns exemplos que falharam na detecção da convergência. Por isso, recomenda-se que o uso desses diagnósticos seja combinado também com uma análise gráfica e medidas descritivas.

Gamerman (1991) revisou algumas técnicas de identificação e monitoração informal e de verificação de convergência. Brooks & Roberts (1995) revisaram alguns dos métodos de convergência e discutiram a implementação destes e possíveis extensões; os autores também utilizaram os critérios de interpretação e aplicação para compará-los.

3.4.1 Método de Gelman e Rubin

Gelman & Rubin (1992) sugerem uma abordagem da convergência do método Gibbs, usando componentes da variância de sequências multiplas da cadeia, simuladas a partir de diferentes pontos iniciais dispersos. Eles comparam a variância dentro e entre as cadeias, usadas para estimar o fator pelo qual o parâmetro de escala da distribuição a posteriori marginal pode ser reduzido, quando o número de iterações é grande. O método funciona da seguinte forma (ver por exemplo, Paulino *et al.*, 2003)

- (a) Simulam-se $m \geq 2$ sequências, cada uma de comprimento $2n$, a partir dos valores iniciais simulados de uma distribuição com dispersão similar à dispersão da distribuição de equilíbrio. Usam-se apenas as n últimas iterações de cada sequência.
- (b) Seja U a quantidade de interesse que desejamos estimar que é uma função do parâmetro. Chamando de u_{ij} o valor que U assume na j -ésima iteração da cadeia, obtemos

$$\bar{u}_i = \frac{1}{n} \sum_{j=1}^n u_{ij}$$

$$s_i^2 = \frac{1}{n-1} \sum_{j=1}^n (u_{ij} - \bar{u}_i)^2$$

a média e a variância, respectivamente, de U para cada sequência $i = 1, 2, \dots, m$.

- (c) Calculam-se os seguintes componentes da variância,

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

a média das m variâncias dentro das sequências, cada uma baseada em $n - 1$ graus de liberdade, e

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

a variância entre as médias das m sequências, \bar{u}_i , cada uma baseada em n valores de U .

- (d) Estima-se a média de U como uma média amostral de todos os mn valores simulados de U , isto é,

$$\hat{\mu} = \bar{u}_{..} = \frac{1}{m} \sum_{i=1}^m \bar{u}_i.$$

- (e) Estima-se a variância de U como uma média ponderada de K e B

$$\hat{\sigma}^2 = \frac{n - 1}{n} K + \frac{1}{n} B$$

$\hat{\sigma}^2$ super-estima σ^2 se a distribuição inicial for super-dispersa e não-viciada no equilíbrio.

- (f) Cria-se uma distribuição t-Student conservativa, ou seja, com poucos g.l. (graus de liberdade), para U com relação a $\hat{\mu}$; dispersão $\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^2 + \frac{B}{mn}}$ e g.l. $\nu = 2\hat{V}^2 / \text{Var}(\hat{V})$, onde

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

As variâncias e covariâncias são estimadas a partir dos m valores amostrais de s_i^2 , \bar{u}_i e \bar{u}_i^2 .

- (g) Estima-se um indicador de convergência (fator de redução que sofreria o parâmetro de escala da distribuição t-Student para U), se as simulações continuassem como,

$$\sqrt{\hat{R}} = \sqrt{\frac{\hat{V}}{K} \frac{\tau}{\tau - 2}}$$

que se aproxima de um quando o número de iterações aumenta.

À medida que n cresce, ambos os estimadores acabarão convergindo para \hat{V} , e R convergirá para 1. Assim, \hat{R} pode ser usado como indicador de convergência pela avaliação de sua proximidade a 1. Gelman *et al.* (1995) sugerem aceitar como garantia de convergência valores de $\hat{R} \leq 1.1$.

Capítulo 4

Alguns Modelos de Regressão para Dados de Sobrevivência Multivariados Obtidos pela Inclusão de Fatores Aleatórios

Neste capítulo, apresentamos os modelos de regressão Weibull e gama generalizado, considerando os dados de sobrevivência na forma multivariada.

4.1 Modelo de Regressão Weibull

Seja T_{ji} uma variável aleatória denotando o tempo de sobrevivência do i -ésimo indivíduo ($i = 1, 2, \dots, n$) na j -ésima medida repetida para o mesmo indivíduo ($j = 1, 2, \dots, k$) com uma distribuição Weibull com densidade,

$$f(t_{ji}|\nu_j, \lambda_j(i)) = \nu_j \lambda_j(i) t_{ji}^{\nu_j-1} \exp\{-\lambda_j(i) t_{ji}^{\nu_j}\}, \quad (4.1)$$

tal que $t_{ji} > 0$; $\nu_j > 0$ é o parâmetro de forma e $\lambda_j(i)$ é o parâmetro de escala.

Para capturar a correlação entre as medidas repetidas $T_{1i}, T_{2i}, \dots, T_{ki}$ para o mesmo indivíduo, introduzimos uma variável latente ou de “fragilidade” W_i , para cada indivíduo e $i = 1, 2, \dots, n$ com distribuição normal, ou seja,

$$W_i \stackrel{\text{iid}}{\sim} N(0, \sigma_w^2). \quad (4.2)$$

Na presença de um vetor de covariáveis $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$ e a variável latente W_i , assumimos o modelo de regressão na equação (4.1), dado por,

$$\lambda_j(i) = \exp\{w_i + \beta_j' \mathbf{x}_i\}, \quad (4.3)$$

em que $\beta_j' = (\beta_{j1}, \beta_{j2}, \dots, \beta_{jp})$ é o vetor de parâmetros da regressão, $j = 1, 2, \dots, k$.

A função de risco é dada por,

$$h_j(t_{ji} | \mathbf{x}_i, w_i) = \nu_j t_{ji}^{\nu_j - 1} \exp\{w_i + \beta_j' \mathbf{x}_i\}. \quad (4.4)$$

A função de sobrevivência para um dado t_{ji} é dada por,

$$S(t_{ji} | \mathbf{x}_i, w_i) = \exp\{-t_{ji}^{\nu_j} e^{w_i + \beta_j' \mathbf{x}_i}\}, \quad (4.5)$$

para $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, k$.

Denotamos o modelo definido pelas equações (4.1)-(4.5) como “modelo 1”.

De (4.4), observamos que podemos ter risco constante, decrescente ou crescente, assumindo, respectivamente, $\nu_j = 1$, $\nu_j < 1$ or $\nu_j > 1$.

A média e a variância condicional para T_{ji} dado \mathbf{x}_i e w_i são dadas, respectivamente por

$$E(T_{ji} | \mathbf{x}_i, w_i) = \frac{\Gamma(1 + 1/\nu_j)}{\exp\left\{\frac{1}{\nu_j} (w_i + \beta_j' \mathbf{x}_i)\right\}}, \quad (4.6)$$

e

$$\text{Var}(T_{ji} | \mathbf{x}_i, w_i) = \frac{1}{\exp\left\{\frac{2}{\nu_j} (w_i + \beta_j' \mathbf{x}_i)\right\}} \left\{ \Gamma\left(1 + \frac{2}{\nu_j}\right) - \Gamma^2\left(1 + \frac{1}{\nu_j}\right) \right\}, \quad (4.7)$$

para $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$.

A média não-condicional para T_{ji} é obtida do resultado, $E(T_{ji} | \mathbf{x}_i) = E[E(T_{ji} | \mathbf{x}_i, w_i)]$, onde

$$E(T_{ji}|\mathbf{x}_i) = \frac{\Gamma(1 + 1/\nu_j)}{\exp\left(\frac{\beta'_j \mathbf{x}_i}{\nu_j}\right)} E\{e^{-W_i/\nu_j}\}. \quad (4.8)$$

Observe, já que $W_i \sim N(0, \sigma_w^2)$, temos que $g(w_i) = e^{-w_i/\nu_j} \stackrel{a}{\sim} N\{g(0); [g'(0)]^2 \sigma_w^2\}$ (“método delta”), ou seja,

$$e^{-W_i/\nu_j} \stackrel{a}{\sim} N\left[1; \frac{\sigma_w^2}{\nu_j^2}\right]. \quad (4.9)$$

Então, a média não-condicional para T_{ji} dado \mathbf{x}_i é,

$$E(T_{ji}|\mathbf{x}_i) \approx \frac{\Gamma(1 + 1/\nu_j)}{\exp\left(\frac{\beta'_j \mathbf{x}_i}{\nu_j}\right)}, \quad (4.10)$$

para $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$.

A variância não-condicional para T_{ji} é obtida da $\text{Var}(T_{ji}|\mathbf{x}_i) = \text{Var}\{E(T_{ji}|\mathbf{x}_i, \mathbf{w}_i)\} + E\{\text{Var}(T_{ji}|\mathbf{x}_i, w_i)\}$, ou seja,

$$\begin{aligned} \text{Var}(T_{ji}|\mathbf{x}_i) &\approx \frac{\Gamma^2(1 + 1/\nu_j)}{\exp\left(\frac{2\beta'_j \mathbf{x}_i}{\nu_j}\right)} \text{Var}(e^{-W_i/\nu_j}) + \\ &+ \frac{[\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)]}{\exp\left(\frac{2\beta'_j \mathbf{x}_i}{\nu_j}\right)} E(e^{-2W_i/\nu_j}). \end{aligned} \quad (4.11)$$

Também, usando o “método delta”, observamos que $g(W_i) = e^{-2W_i/\nu_j} \stackrel{a}{\sim} N\left[1; \frac{4\sigma_w^2}{\nu_j^2}\right]$, ou seja,

$$\begin{aligned} \text{Var}(T_{ji}|\mathbf{x}_i) &\approx \frac{\sigma_w^2 \Gamma^2(1 + 1/\nu_j)}{\nu_j^2 \exp\left(\frac{2\beta'_j \mathbf{x}_i}{\nu_j}\right)} + \\ &+ \frac{1}{\exp\left(\frac{2\beta'_j \mathbf{x}_i}{\nu_j}\right)} \{\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)\}. \end{aligned} \quad (4.12)$$

Observe que não considerando a presença da “fragilidade” W_i , a variância para T_{ji} dado \mathbf{x}_i é

$$\text{Var}(T_{ji}|\mathbf{x}_i) \approx \frac{\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)}{\exp\left(\frac{2\beta_j'\mathbf{x}_i}{\nu_j}\right)}. \quad (4.13)$$

Da equação (4.12) e (4.13), observamos que a variabilidade extra-Weibull na presença da “fragilidade” W_i com distribuição normal (4.2), é dada por

$$\frac{\sigma_w^2 \Gamma^2(1 + 1/\nu_j)}{\nu_j^2 \exp\left(\frac{2\beta_j'\mathbf{x}_i}{\nu_j}\right)}, \quad (4.14)$$

para $j = 1, 2, \dots, k; i = 1, 2, \dots, n$.

Um modelo diferente pode ser considerado substituindo (4.3) por

$$\lambda_j(i) = w_i e^{\beta_j'\mathbf{x}_i}, \quad (4.15)$$

com

$$W_i \stackrel{\text{iid}}{\sim} \text{Gama}(\phi^{-1}, \phi^{-1}). \quad (4.16)$$

De (4.16), observe que $E(W_i) = 1$ e $\text{Var}(W_i) = \phi$.

Assumamos o modelo definido por (4.1) e (4.15) como “modelo 2”.

Do “modelo 2”, a média e variância condicionais para T_{ji} dado \mathbf{x}_i e w_i , são dadas, respectivamente por

$$E(T_{ji}|\mathbf{x}_i, w_i) = \frac{\Gamma(1 + 1/\nu_j)}{w_i^{1/\nu_j} e^{\beta_j'\mathbf{x}_i/\nu_j}}, \quad (4.17)$$

e

$$\text{Var}(T_{ji}|\mathbf{x}_i, w_i) = \frac{\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)}{w_i^{2/\nu_j} e^{2\beta_j'\mathbf{x}_i/\nu_j}}, \quad (4.18)$$

para $i = 1, 2, \dots, n; j = 1, 2, \dots, k$.

Seguindo os mesmos argumentos usados na determinação da média e da variância não-condicional para T_{ji} assumindo o “modelo 1”, e observando que a “fragilidade” W_i tem uma distribuição $\text{Gamma}(\phi^{-1}, \phi^{-1})$, a média não-condicional para T_{ji} assumindo o “modelo 2” (ver apêndice A), é dada por

$$E(T_{ji}|\mathbf{x}_i) = \frac{\Gamma(1 + 1/\nu_j)(\phi^{-1})^{1/\nu_j}\Gamma(\phi^{-1} - \nu_j^{-1})}{\exp\{\boldsymbol{\beta}'_j\mathbf{x}_i/\nu_j\}\Gamma(\phi^{-1})}, \quad (4.19)$$

para $\phi^{-1} > \nu_j^{-1}$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$.

A variância não-condicional para T_{ji} é dada por

$$\begin{aligned} \text{Var}(T_{ji}|\mathbf{x}_i) &= \frac{(\phi^{-1})^{2/\nu_j}}{\exp(2\boldsymbol{\beta}'_j\mathbf{x}_i/\nu_j)} \times \\ &\times \left\{ \frac{\Gamma(1 + 2/\nu_j)\Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} - \left[\frac{\Gamma(1 + 1/\nu_j)\Gamma(\phi^{-1} - 1/\nu_j)}{\Gamma(\phi^{-1})} \right]^2 \right\}, \end{aligned} \quad (4.20)$$

para $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$.

Generalizações do “modelo 1” e “modelo 2” podem ser consideradas assumindo que a covariável \mathbf{x}_i também afeta o parâmetro de forma ν_j , ou seja, assumindo o modelo de regressão $\nu_j(i) = \exp\{\boldsymbol{\alpha}'\mathbf{x}_i\}$. Denotamos estes modelos como “modelo 3” e “modelo 4”, respectivamente.

Observe que esses modelos introduzem novas perspectivas para analisar dados de sobrevivência multivariados, isto é, são novos modelos introduzidos nesta tese.

4.1.1 Análise Bayesiana para o Modelo

Assumindo tempos de vida na presença de observações censuradas e vetor de covariável $\mathbf{x} = (x_1, x_2, \dots, x_p)'$, definimos (como dado na subseção 2.1.1) uma variável indicadora para as observações censuradas e não censuradas, por

$$\delta_{ji} = \begin{cases} 1 & \text{para o tempo de vida observado} \\ 0 & \text{para a observação censurada} \end{cases} \quad (4.21)$$

Assumindo o “modelo 1” definido por (4.1), (4.2) e (4.3), a função de verossimilhança é dada por,

$$L(\mathbf{t}|\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{w}) = \prod_{i=1}^n \prod_{j=1}^k [f(t_{ji}|\mathbf{x}_i, w_i)]^{\delta_{ji}} [S(t_{ji}|\mathbf{x}_i, w_i)]^{1-\delta_{ji}}, \quad (4.22)$$

em que $S(t_{ji}|\mathbf{x}_i, w_i)$ é a função de sobrevivência definida em (4.5), $\boldsymbol{\beta}' = (\beta_1, \beta_2, \dots, \beta_k)$,

$\beta'_j = (\beta_{j1}, \beta_{j2}, \dots, \beta_{jp})$, $j = 1, 2, \dots, k$; $\nu' = (\nu_1, \nu_2, \dots, \nu_k)$, $\mathbf{w}' = (w_1, w_2, \dots, w_n)$,
 $\mathbf{x}'_i = (x_{i1}, x_{i2}, \dots, x_{ip})$; $i = 1, 2, \dots, n$.

Ou seja,

$$L(\mathbf{t}|\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{w}) = \prod_{i=1}^n \prod_{j=1}^k \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j-1)} \exp\{\delta_{ji}[w_i + \beta'_j \mathbf{x}_i]\} \exp\{-t_{ji}^{\nu_j} e^{w_i + \beta'_j \mathbf{x}_i}\}. \quad (4.23)$$

Assumindo o “modelo 2”, a função de verossimilhança é (das equações(4.15) e (4.22))
dada por

$$L(\mathbf{t}|\mathbf{x}, \boldsymbol{\beta}, \boldsymbol{\nu}, \mathbf{w}) = \prod_{i=1}^n \prod_{j=1}^k \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j-1)} w_i^{\delta_{ji}} e^{\delta_{ji} \beta'_j \mathbf{x}_i} \exp\{-t_{ji}^{\nu_j} w_i e^{\beta'_j \mathbf{x}_i}\} \quad (4.24)$$

Para uma análise Bayesiana hierárquica do “modelo 1”, assumimos no primeiro estágio,
as seguintes distribuições a priori para os parâmetros:

$$\nu_j \sim \text{Gama}(a_j, b_j); \quad (4.25)$$

$$\beta_{jl} \sim N(0; c_{jl}^2),$$

onde $j = 1, 2, \dots, k$; $l = 1, 2, \dots, p$; a_j, b_j, c_{jl} são hiperparâmetros conhecidos e $\text{Gama}(a, b)$
denota uma distribuição gama com média a/b e variância a/b^2 .

No segundo estágio da análise Bayesiana hierárquica, assumimos uma distribuição a
priori gama para σ_w^2 , ou seja,

$$\sigma_w^2 \sim \text{Gama}(d, e); \quad (4.26)$$

tal que d e e são hiperparâmetros conhecidos. Assumimos também independência entre
os parâmetros.

Combinando-se (4.23), (4.25) e (4.26), obtemos a distribuição a posteriori conjunta
para $\mathbf{w}, \boldsymbol{\nu}, \boldsymbol{\beta}$ e σ_w^2 dada por,

$$\begin{aligned} \pi(\boldsymbol{\nu}, \boldsymbol{\beta}, \mathbf{w}, \sigma_w^2 | \mathbf{x}, \mathbf{t}) &\propto \left\{ \prod_{i=1}^n \exp\left(-\frac{w_i^2}{2\sigma_w^2}\right) \right\} \left\{ \prod_{j=1}^k \prod_{l=1}^p \exp\left(-\frac{\beta_{jl}^2}{2c_{jl}^2}\right) \right\} \times \quad (4.27) \\ &\times (\sigma_w^2)^{d-1} \exp(-e\sigma_w^2) \left(\prod_{j=1}^k \nu_j^{a_j-1} e^{-b_j \nu_j} \right) \prod_{i=1}^n \prod_{j=1}^k \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j-1)} \times \\ &\times \exp\{\delta_{ji}(w_i + \beta'_j \mathbf{x}_i)\} \exp\{-t_{ji}^{\nu_j} e^{w_i + \beta'_j \mathbf{x}_i}\}. \end{aligned}$$

As distribuições condicionais necessárias para o amostrador de Gibbs são dadas por,

$$\pi(\nu_j | \beta_{jl}, w_i, \sigma_w^2, \mathbf{x}, \mathbf{t}) \propto \nu_j^{a_j-1} \exp(-b_j \nu_j) \prod_{i=1}^n \left[\nu_j^{\delta_{ji}} t_{ji}^{\alpha_{ji}(\nu_j-1)} \exp\left(-t_{ji}^{\nu_j} e^{w_i + \sum_{l=1}^p \beta_{jl} x_{il}}\right) \right],$$

$$j = 1, 2, \dots, k.$$

$$\pi(\beta_{jl} | \nu_j, w_i, \sigma_w^2, \mathbf{x}, \mathbf{t}) \propto \exp\left[-\frac{\beta_{jl}}{2c_j^2} + \sum_{i=1}^n \left(\delta_{ji} \beta_{jl} x_i - t_{ji}^{\nu_j} e^{w_i + \sum_{l=1}^p \beta_{jl} x_{il}}\right)\right],$$

$$j = 1, 2, \dots, k; l = 1, 2, \dots, p. \quad (4.28)$$

$$\pi(w_i | \nu_j, \beta_{jl}, \sigma_w^2, \mathbf{x}, \mathbf{t}) \propto \exp\left(-\frac{w_i^2}{2\sigma_w^2} + \sum_{j=1}^k \delta_{ji} w_i - \sum_{j=1}^k t_{ji}^{\nu_j} e^{w_i + \sum_{l=1}^p \beta_{jl} x_{il}}\right),$$

$$i = 1, 2, \dots, n.$$

$$\pi(\sigma_w^2 | \nu_j, \beta_{jl}, w_i, \mathbf{x}, \mathbf{t}) \propto \exp\left[-\sum_{i=1}^n \frac{w_i^2}{2\sigma_w^2} - e\sigma_w^2\right] (\sigma_w^2)^{d-1}$$

Para obter os resumos a posteriori de interesse, simulamos amostras da distribuição a posteriori conjunta (4.27) usando o método MCMC como o popular algoritmo amostrador de Gibbs, ou o algoritmo Metropolis-Hastings.

Uma grande simplificação na simulação das amostras para a distribuição a posteriori conjunta, é dado pelo *software WinBUGS* (Spiegelhalter, D. J.; Thomas, A.; Best, N. G.; Gilks, W. R., 2003), que requer somente a especificação da distribuição conjunta para os dados e a distribuição a priori para os parâmetros.

Assumindo o “modelo 2”, onde $W_i \sim \text{Gama}(1/\phi, 1/\phi)$ e considerando as mesmas prioris (4.25) para ν_j e β_{jl} ; e uma distribuição a priori uniforme para ϕ , ou seja,

$$\phi \sim U(0, f) \quad (4.29)$$

onde $U(a, b)$ denota uma distribuição uniforme no intervalo (a, b) e f é um hiperparâmetro conhecido.

A distribuição a posteriori conjunta para $\mathbf{w}, \boldsymbol{\nu}, \boldsymbol{\beta}$ e ϕ é dada por,

$$\begin{aligned}
 \pi(\boldsymbol{\nu}, \boldsymbol{\beta}, \mathbf{w}, \phi | \mathbf{x}, \mathbf{t}) &\propto \left\{ \prod_{i=1}^n w_i^{\phi^{-1}-1} e^{-\phi^{-1}w_i} \right\} \phi^{f-1} e^{-f\phi} \times \\
 &\times \left\{ \prod_{j=1}^k \prod_{l=1}^p \exp\left(-\frac{\beta_{jl}^2}{2c_j^2}\right) \right\} \left\{ \prod_{j=1}^k \nu_j^{a_j-1} e^{-b_j\nu_j} \right\} \times \\
 &\times \prod_{i=1}^n \prod_{j=1}^k \nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j-1)} w_i^{\delta_{ji}} \exp\{\delta_{ji}\beta'_{j}\mathbf{x}_i\} \exp\{-t_{ji}^{\nu_j} w_i e^{\beta'_{j}\mathbf{x}_i}\}.
 \end{aligned} \tag{4.30}$$

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

$$\begin{aligned}
 \pi(\phi | \nu_j, \beta_{jl}, w_i, \mathbf{x}, \mathbf{t}) &\propto \left(\prod_{i=1}^n w_i^{\frac{1}{\phi}-1} \right) \phi^{f-1} \exp\left[-\sum_{i=1}^n \frac{w_i}{\phi} - f\phi\right] \\
 \pi(\nu_j | \beta_{jl}, w_i, \phi, \mathbf{x}, \mathbf{t}) &\propto \left(\nu_j^{a_j-1} \right) \prod_{i=1}^n \left[\nu_j^{\delta_{ji}} t_{ji}^{\delta_{ji}(\nu_j-1)} \right] \times \\
 &\times \exp\left\{-b_j\nu_j - \sum_{i=1}^n t_{ji}^{\nu_j} w_i e^{\sum_{l=1}^p \beta_{jl}x_{il}}\right\}, \\
 &j = 1, 2, \dots, k.
 \end{aligned} \tag{4.31}$$

$$\begin{aligned}
 \pi(\beta_{jl} | \nu_j, w_i, \phi, \mathbf{x}, \mathbf{t}) &\propto \exp\left[-\frac{\beta_{jl}^2}{2c_j^2} + \sum_{i=1}^n \delta_{ji}\beta'_{j}\mathbf{x}_i - \sum_{i=1}^n t_{ji}^{\nu_j} w_i e^{\sum_{l=1}^p \beta_{jl}x_{il}}\right], \\
 &j = 1, 2, \dots, k; l = 1, 2, \dots, p.
 \end{aligned}$$

$$\begin{aligned}
 \pi(w_i | \nu_j, \beta_{jl}, \phi, \mathbf{x}, \mathbf{t}) &\propto \exp\left[-\frac{w_i}{\phi} - \sum_{j=1}^k t_{ji}^{\nu_j} w_i e^{\sum_{l=1}^p \beta_{jl}x_{il}}\right] \left(w_i^{\frac{1}{\phi}-1}\right) \left(\prod_{j=1}^k w_i^{\delta_{ji}}\right), \\
 &i = 1, 2, \dots, n.
 \end{aligned}$$

4.2 Modelo de Regressão Gama Generalizada

Nesta seção, assumimos que os tempos de vida T_{ji} tem uma distribuição gama generalizada com densidade,

$$f(t_{ji}|\nu_j, \mu_j(i), \theta_j) = \frac{\theta_j}{\Gamma(\nu_j)} [\mu_j(i)]^{\theta_j \nu_j} t_{ji}^{\theta_j \nu_j - 1} \exp \{ -[\mu_j(i) t_{ji}]^{\theta_j} \} \quad (4.32)$$

onde $t_{ji} > 0$, $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$; $\theta_j > 0$; $\nu_j > 0$ e $\mu_j(i) > 0$.

A distribuição gama generalizada apresenta uma família bastante flexível de distribuições, que inclui como caso especial, as distribuições exponencial ($\theta_j = \nu_j = 1$), Weibull ($\nu_j = 1$) e gama ($\theta_j = 1$). A distribuição log-normal também surge como uma forma limite de (4.32), ou seja, o modelo gama generalizado inclui como casos especiais, as mais comumente distribuições usadas para tempos de vida. Isto é útil para a discriminação entre esses e outros modelos.

A função de sobrevivência para um dado valor de T_{ji} , é dada por,

$$S(t_{ji}|\nu_j, \mu_j(i), \theta_j) = P(T_{ji} > t_{ji}) = \frac{\theta_j}{\Gamma(\nu_j)} [\mu_j(i)]^{\theta_j \nu_j} \int_{t_{ji}}^{\infty} z^{\theta_j \nu_j - 1} e^{-[\mu_j(i) z]^{\theta_j}} dz \quad (4.33)$$

Para capturar a correlação entre essas medidas repetidas $T_{1i}, T_{2i}, \dots, T_{ki}$ para o mesmo indivíduo, introduzimos “fragilidades” W_i , $i = 1, 2, \dots, n$. Na presença do vetor de covariáveis $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$, assumimos o modelo de regressão,

$$\mu_j(i) = \exp\{w_i + \beta_j' \mathbf{x}_i\} \quad (4.34)$$

onde W_i tem uma distribuição normal (4.2) $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$, denotado como “modelo 5”, ou,

$$\mu_j(i) = w_i e^{\beta_j' \mathbf{x}_i} \quad (4.35)$$

onde W_i tem uma distribuição gama (4.16), denotado como “modelo 6”.

Assumindo o “modelo 5”, consideramos as seguintes distribuições a priori no primeiro estágio da análise Bayesiana hierárquica:

$$\nu_j \sim \text{Gama}(a_j, b_j), \quad (4.36)$$

$$\theta_j \sim \text{Gama}(c_j, d_j),$$

$$\beta_{jl} \sim N(0, e_{jl}^2),$$

em que $j = 1, 2, \dots, k$; $l = 1, 2, \dots, p$; a_j, b_j, c_j, d_j e e_{jl} são hiperparâmetros conhecidos.

No segundo estágio da análise Bayesiana hierárquica, assumimos a priori gama (4.26) para σ_w^2 .

Assumindo o “modelo 6”, consideramos as mesmas prioris (4.36) para ν_j, θ_j e β_{jl} , e uma priori gama (4.26) para ϕ .

Para desenvolver uma análise Bayesiana para a distribuição gama generalizada de dados de sobrevivência multivariados na presença de covariáveis e observações censuradas, necessitamos de distribuições a priori informativas para obter a convergência do algoritmo amostrador de Gibbs. Observe que, usando a distribuição gama generalizada, usualmente temos grandes dificuldades para obter inferências clássicas de interesse (ver por exemplo, Hager & Bain, 1970; Parr & Webster, 1965).

Amostras da distribuição a posteriori conjunta para os parâmetros do “modelo 3” ou “modelo 4”, são obtidas também usando o método MCMC.

4.3 Aplicação aos Dados de McGilchrist e Aisbett

Apresentamos a seguir, os resultados obtidos na análise do conjunto de dados de sobrevivência introduzido por McGilchrist & Aisbett (1991) relacionado a infecção nos rins, onde há recorrência da infecção de 38 rins de pacientes usando máquinas de diálise portáteis são registradas. Infecções podem ocorrer no local da inserção do cateter. O tempo anotado, chamado tempo de infecção, é o tempo de sobrevivência (em dias) do paciente até ocorrer a infecção e o cateter ter sido removido, ou o tempo é censurado e o cateter foi removido por outras razões. O cateter é reinserido depois de algum tempo e o segundo tempo de infecção é novamente observado ou censurado (veja tabela 2.1, no capítulo 2).

4.3.1 Análise Descritiva dos Dados

Inicialmente vamos fazer uma análise descritiva dos dados usando o estimador não paramétrico de Kaplan-Meier, para cada tempo até a primeira e segunda infecção, respectivamente.

Para o tempo da primeira infecção temos 32 valores não-censurados e 6 com censura à direita. A média é 136,785 dias e a mediana é 63 dias, considerando as observações completas. A Tabela 4.1 apresenta o estimador Kaplan-Meier dessa variável.

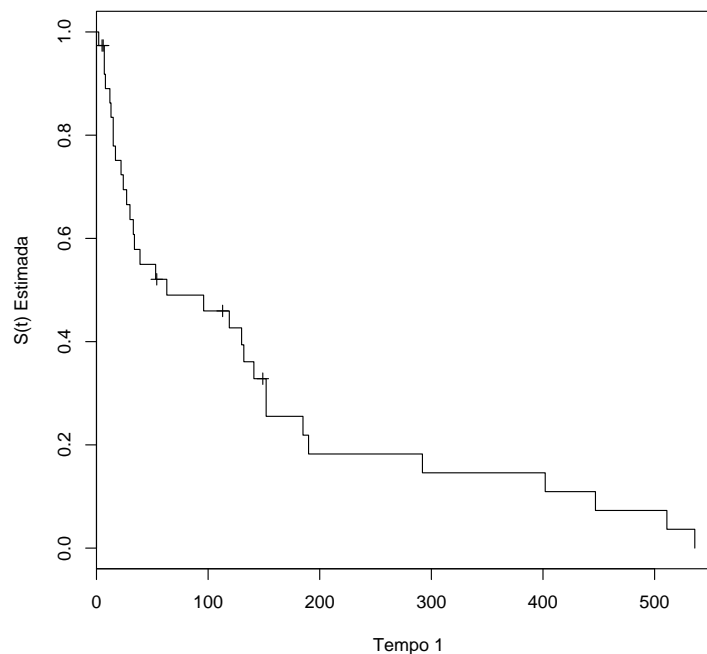


Figura 4.1: Estimador não-paramétrico de Kaplan-Meier para a função de sobrevivência dos tempos da primeira infecção.

No caso do tempo da segunda infecção, temos a média 133,503 dias e a mediana 78 dias, considerando os dados completos. Para essa variável temos 26 valores não-censurados e 12 com censura à direita. Na Tabela 4.2 temos o estimador Kaplan-Meier para essa variável.

Encontramos os estimadores de máxima verossimilhança para os parâmetros do modelo Weibull, usando métodos iterativos e posteriormente obtivemos as estimativas das funções de sobrevivência para os tempos da primeira infecção, usando o modelo Weibull e também o método de Kaplan-Meier, cujos resultados estão na Tabela 4.1 e podem ser visualizados através da Figura 4.1.

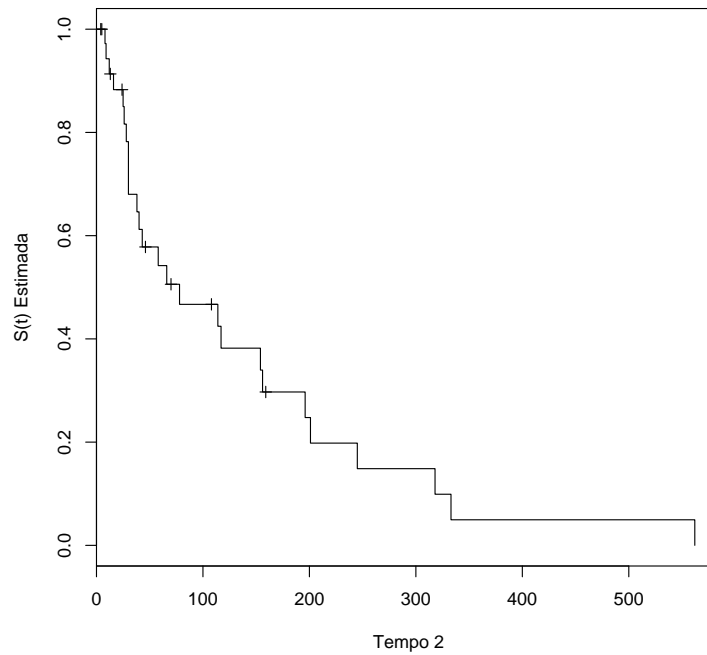


Figura 4.2: Estimador não-paramétrico de Kaplan-Meier para a função de sobrevivência dos tempos da segunda infecção.

Analogamente fazemos o mesmo para o tempo da segunda infecção e apresentamos os resultados na Tabela 4.2 e na Figura 4.2.

Tabela 4.1: Estimativas não-paramétricas da função de sobrevivência para os tempos até 1ª infecção, usando o estimador produto-limite de Kaplan-Meier e a função de sobrevivência do modelo Weibull estimado pelo método de máxima verossimilhança.

tempo	Kaplan-Meier	Weibull
2	0,97368421	0,96545687
5	0,97368421	0,92851604
6	0,97368421	0,91755192
7	0,91804511	0,90704635
8	0,89022556	0,89693170
12	0,86240602	0,85953839
13	0,83458647	0,85081522
15	0,77894737	0,83398473
17	0,75112782	0,81788694
22	0,72330827	0,78033858
24	0,69437594	0,76624748
27	0,66544361	0,74597120
30	0,63651128	0,72662915
33	0,60757895	0,70813050
34	0,57864662	0,70213840
39	0,54971429	0,67337604
53	0,52078195	0,60185971
54	0,52078195	0,59718818
63	0,49014772	0,55737564
96	0,45951349	0,43873919
113	0,45951349	0,39027772
119	0,42669110	0,37478329
130	0,39386871	0,34829206
132	0,36104631	0,34372395
141	0,32822392	0,32404296
149	0,32822392	0,30767636
152	0,25528527	0,30179474
185	0,21881595	0,24512195
190	0,18234662	0,23766690
292	0,14587730	0,13011582
402	0,10940797	0,07092264
447	0,07293865	0,05584797
511	0,03646932	0,04005613
536	0,00000111	0,03525452

Tabela 4.2: Estimativas não-paramétricas da função de sobrevivência para os tempos até 2ª infecção, usando o estimador produto-limite de Kaplan-Meier e a função de sobrevivência do modelo Weibull estimada pelo método de máxima verossimilhança.

tempo	Kaplan-Meier	Weibull
4	1,00000000	0,96987450
5	1,00000000	0,96249832
8	0,97222222	0,94071634
9	0,94276094	0,93356874
12	0,91329966	0,91245665
13	0,91329966	0,90552776
16	0,88285634	0,88505965
24	0,88285634	0,83273568
25	0,85015796	0,82641852
26	0,81615164	0,82014966
28	0,78214532	0,80775529
30	0,68012637	0,79554952
38	0,64612005	0,74855391
40	0,61211373	0,73724771
43	0,57810741	0,72060933
46	0,57810741	0,70434812
58	0,54197570	0,64290446
66	0,50584399	0,60495934
70	0,50584399	0,58683756
78	0,46693291	0,55221051
108	0,46693291	0,43963111
114	0,42448446	0,42003991
117	0,38203602	0,41057473
154	0,33958757	0,30998546
156	0,29713912	0,30531303
159	0,29713912	0,29843627
196	0,24761594	0,22534627
201	0,19809275	0,21695344
245	0,14856956	0,15535802
318	0,09904637	0,08928832
333	0,04952319	0,07968559
562	0,00000321	0,01403871

A partir dos resultados apresentados nas Tabelas 4.1 e 4.2, fizemos os gráficos das funções de sobrevivências estimadas por Kaplan-Meier e das funções de sobrevivências estimadas através do modelo Weibull, que são apresentados na Figura 4.3. Notamos que o ajuste gráfico para o modelo Weibull parece ser adequado, para ambos os tempos de infecção.

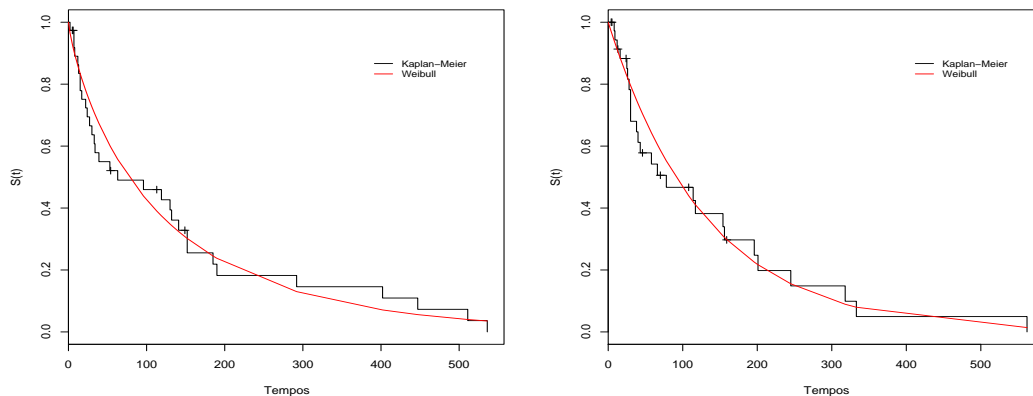


Figura 4.3: Funções de sobrevivências estimadas pelo modelo Weibull e funções de sobrevivências estimadas por Kaplan-Meier, para os tempo da primeira e da segunda infecção, respectivamente.

Na Figura 4.4, temos gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelo modelo Weibull.

A partir da Figura 4.4, verificamos que existe aproximadamente uma relação linear, indicando que este modelo é possivelmente adequado para os dados sob estudo.

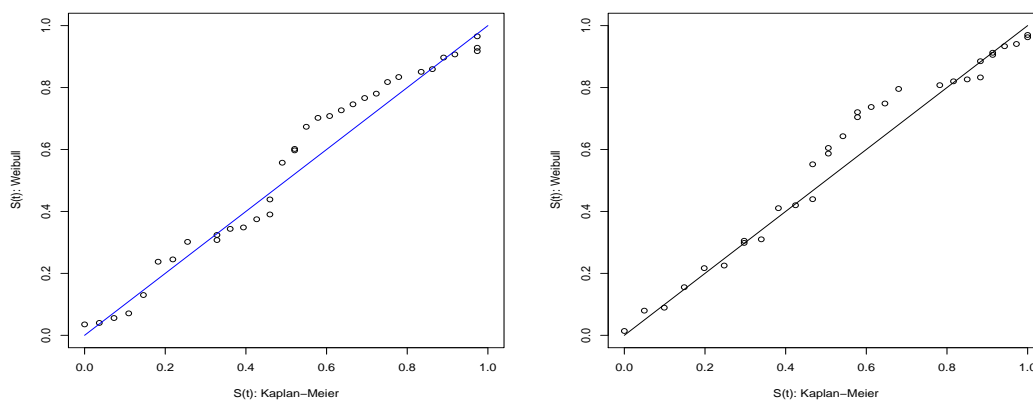


Figura 4.4: Gráfico das sobrevivências estimadas por Kaplan-Meier e das sobrevivências estimadas pelo modelo Weibull, para os tempo da primeira e da segunda infecção.

Para confirmar os resultados obtidos anteriormente, foram construídos os gráficos linearizados para o modelo Weibull, tomando os dois tempos de infecção. A Figura 4.5

mostra estes gráficos para o modelo Weibull, onde notamos que não há um afastamento marcante de uma reta (Q-Q plots da distribuição Weibull).

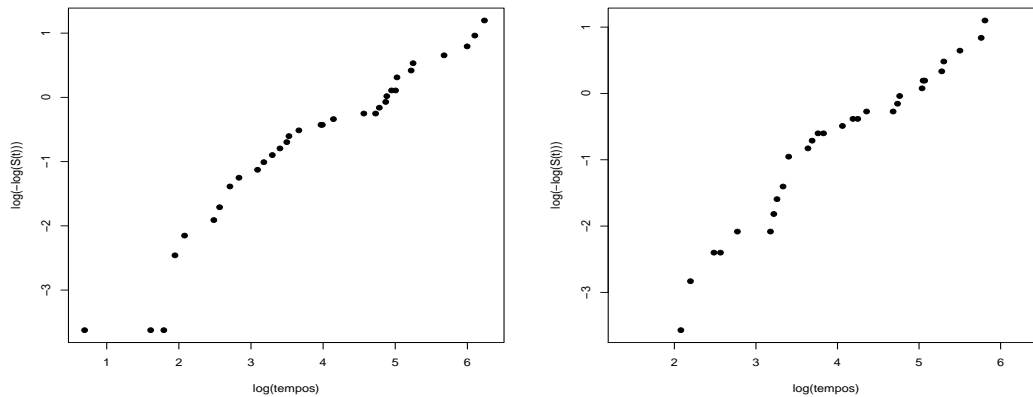


Figura 4.5: Gráfico de $\log(t)$ e $\log(-\log(S(t)))$, para os tempo da primeira e da segunda infecção.

Portanto, pelos procedimentos gráficos apresentados anteriormente, juntamente com os resultados obtidos, podemos concluir que o ajuste do modelo Weibull proposto é satisfatório.

4.3.2 Análise Bayesiana dos dados

Para analisar os tempos de recorrência das infecções (ver Tabela 2.1, no capítulo 2), assumimos um modelo de regressão Weibull (“modelo 1”), na presença de uma “fragilidade” W_i com uma distribuição normal (4.2).

Neste caso, temos somente uma covariável x_i (sexo: $x = 2$ para masculino; $x = 1$ para feminino) e $k = 2$ tempos de recorrência.

Da equação (4.3), temos o modelo de regressão (j -ésima medição do i -ésimo indivíduo)

$$\lambda_j(i) = \exp\{\beta_{1j} + \beta_{2j}x_i + w_i\}, \quad (4.37)$$

$i = 1, 2, \dots, 38; j = 1, 2$.

Para uma análise Bayesiana do “modelo 1”, assumimos as distribuições a priori (4.25) e (4.26) com $a_1 = b_1 = a_2 = b_2 = 1$; $c_{11} = c_{21} = c_{12} = c_{22} = 10$ e $d = e = 0.1$.

Buscando eliminar a correlação entre os valores gerados, escolhemos a cada 10 amostras simuladas por Gibbs (*lag 10*). Usando o *software WinBUGS*, descartamos as primeiras

5000 amostras simuladas por Gibbs (“*burn-in-sample*”) para eliminar os efeitos dos valores iniciais nos parâmetros do modelo. Obtemos os resumos a posteriori de interesse (veja Tabela 4.3), para uma amostra final de tamanho 2000. A Convergência para o algoritmo amostrador de Gibbs foi monitorado usando métodos gráficos, como o de séries temporais e o Gelman & Rubin (1992) para as amostras simuladas.

Este mesmo procedimento de simulação foi considerado para os outros modelos apresentados nesta seção.

Na Tabela 4.3, também temos as estimativas de Monte Carlo para a média a posteriori do tempo mediano de sobrevivência em cada tempo de recorrência. Observe que o tempo mediano de sobrevivência, que não inclui a covariável x_i , é dado por $Med_j = [(\log 2) e^{-\beta_{1j}}]^{1/\nu_j}$, $j = 1, 2$.

Tabela 4.3: Resumos a posteriori (“modelo 1”)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
β_{21}	1,9820	0,5694	(0,8885; 3,1550)
β_{22}	0,7594	0,5570	(-0,3279; 1,8510)
β_{11}	-5,5490	0,8571	(-7,3400; -3,9880)
β_{12}	-5,6110	0,8805	(-7,4620; -3,9940)
med 1	120,40	31,780	(69,350; 194,50)
med 2	106,60	29,520	(62,720; 173,10)
ν_1	1,0880	0,1610	(0,8012; 1,4270)
ν_2	1,1310	0,1742	(0,8113; 1,5100)
$1/\sigma_w^2$	3,1350	3,1060	(0,7188; 11,270)

Na Figura 4.6, temos gráficos de séries temporais para amostras simuladas por Gibbs. Por esses gráficos, observamos a convergência do algoritmo em todos os casos.

Na Figura 4.7 observamos, usando Gelman & Rubin (1992), a convergência do algoritmo considerando diferentes valores iniciais para os parâmetros.

Para o “modelo 2”, que é definido pela densidade Weibull (4.1), onde $\lambda_j(i)$ é dado por (4.15), temos,

$$\lambda_j(i) = w_i \exp\{\beta_{1j} + \beta_{2j}x_i\}, \quad (4.38)$$

$i = 1, 2, \dots, 38$; $j = 1, 2$. Assumimos as distribuições a priori (4.25) e (4.29) com $a_1 = b_1 = a_2 = b_2 = 1$; $c_{11} = c_{21} = c_{12} = c_{22} = 10$ e $f = 5$.

Seguindo os mesmos passos considerados na simulação para a geração das amostras da distribuição a posteriori conjunta dos parâmetros do “modelo 1”, temos na Tabela 4.4, os resumos a posteriori de interesse assumindo o tamanho de amostra de Gibbs final 2000.

Na Figura 4.8, temos os gráficos para as amostras simuladas por Gibbs.

Na Figura 4.9, temos gráficos para os valores de Gelman & Rubin (1992), assumindo

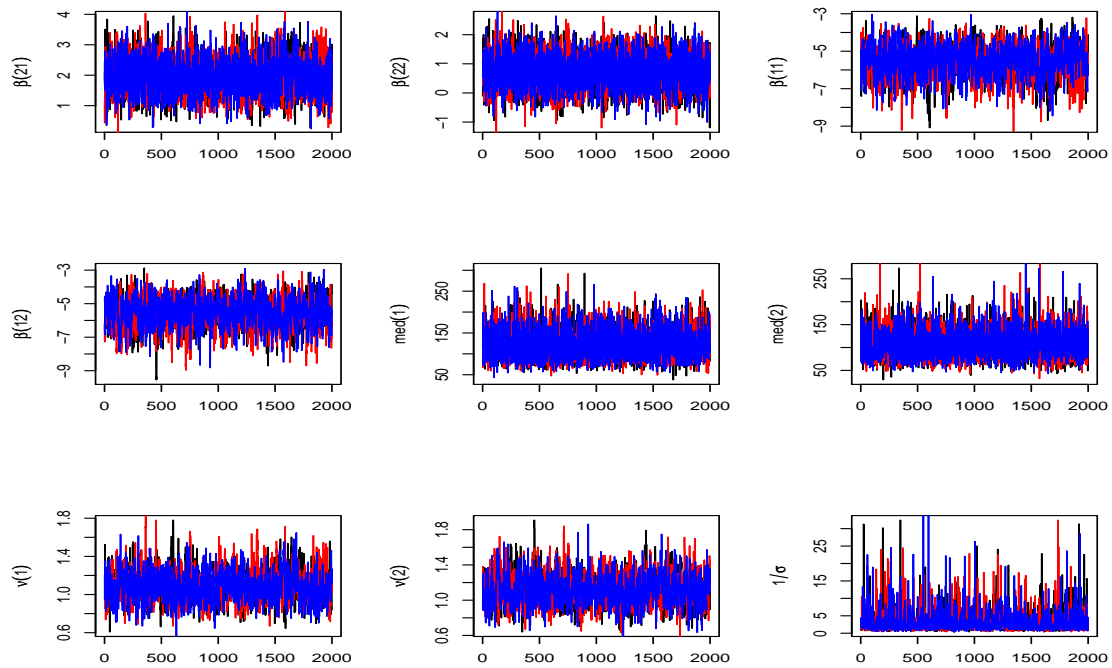


Figura 4.6: Amostras simuladas por Gibbs (“modelo 1”)

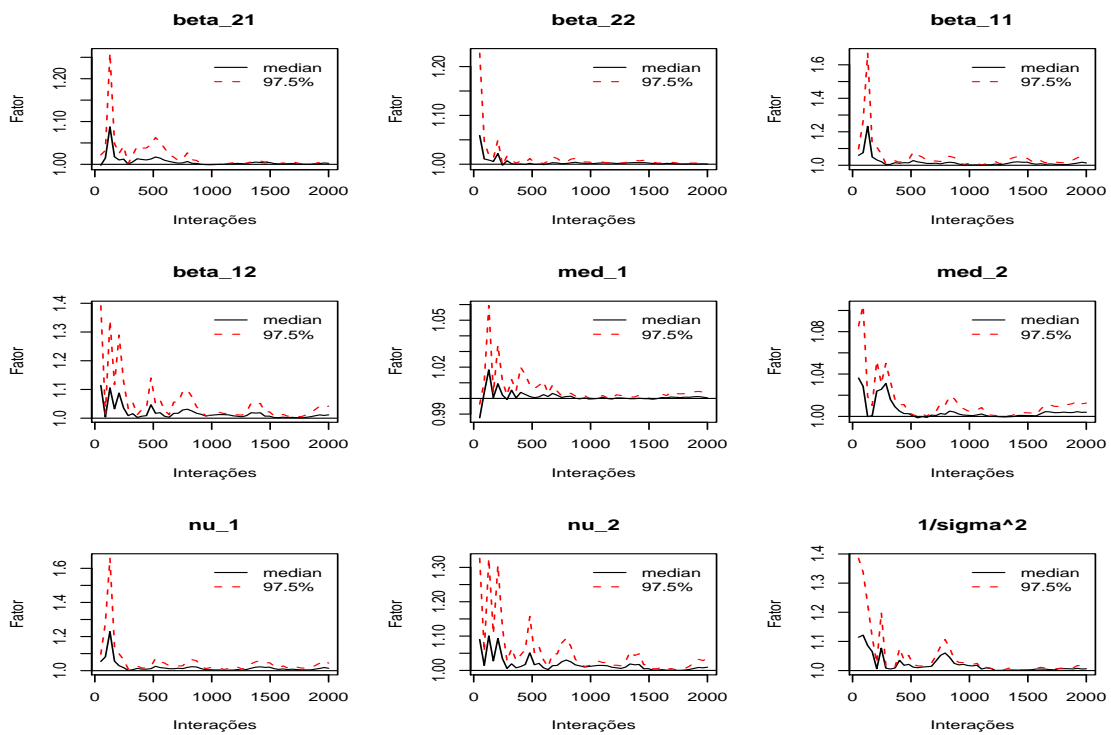


Figura 4.7: Valores de Gelman e Rubin (“modelo 1”)

Tabela 4.4: Resumos a Posteriori (“modelo 2”).

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
β_{21}	2,2370	0,6192	(1,0770; 3,5080)
β_{22}	1,0180	0,6234	(-0,1979; 2,2800)
β_{11}	-5,6340	0,8506	(-7,3790; -4,1350)
β_{12}	-5,6690	0,8794	(-7,5240; -4,0420)
med 1	98,590	26,260	(54,860; 157,90)
med 2	87,340	23,140	(50,320; 141,90)
ν_1	1,1560	0,1719	(0,8508; 1,5130)
ν_2	1,1950	0,1863	(0,8604; 1,5900)
$1/\phi$	2,4670	2,6360	(0,7685; 7,7670)

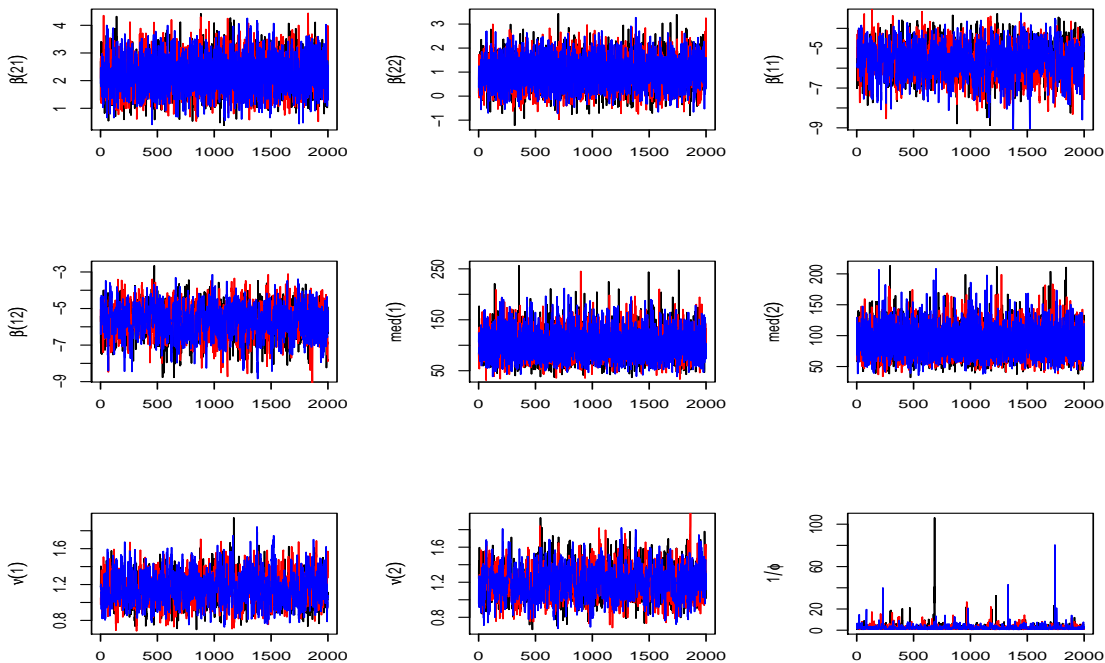


Figura 4.8: Amostras simuladas por Gibbs (“modelo 2”)

diferentes valores iniciais para as cadeias através da qual, observamos a convergência para o algoritmo.

Dos resultados das Tabelas (4.3) e (4.4), observamos resultados similares considerando o “modelo 1” e “modelo 2”. Observamos que em ambos os modelos, temos um efeito significativo da covariável sexo para os primeiros tempos de recorrência, já que o zero não está incluso no intervalo de credibilidade de 95% para β_{21} . Do mesmo modo, observamos que sexo não tem efeito significativo no segundo tempo de recorrência, já que o zero está incluído no intervalo de credibilidade de 95% para β_{22} .

Para discriminação de modelos, podemos utilizar uma medida Bayesiana existente,

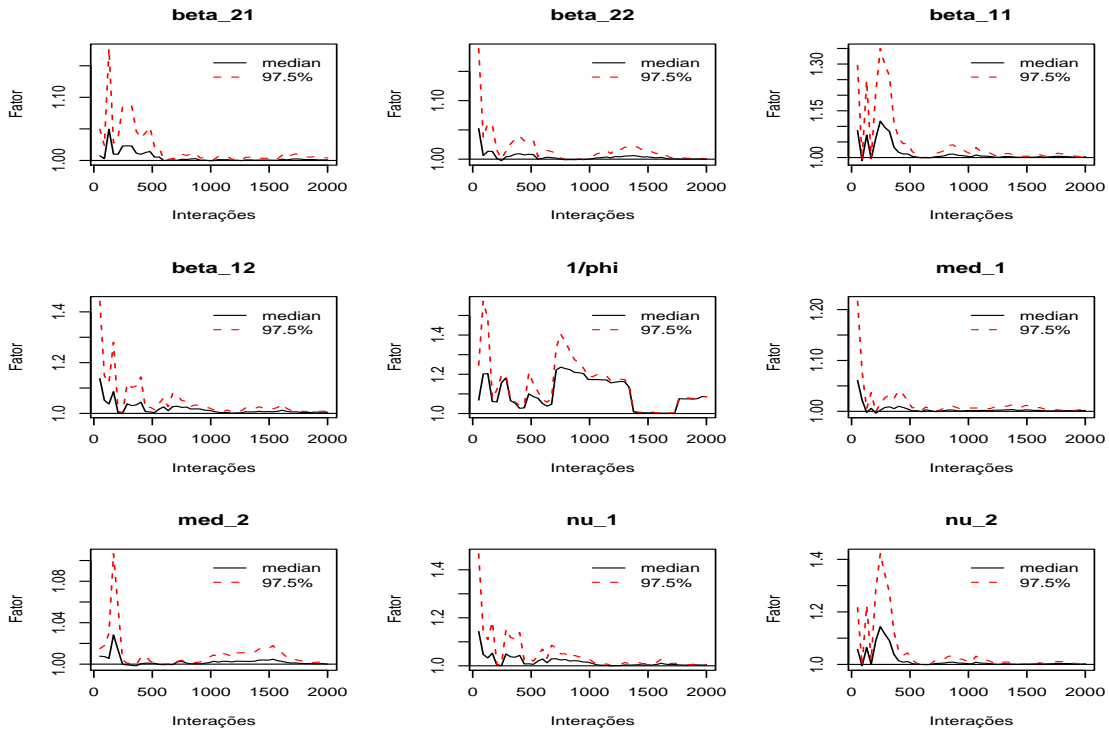


Figura 4.9: Valores de Gelman e Rubin (“modelo 2”)

como o DIC - *Deviance Information Criterion* (Spiegelhalter *et al.*, 2002). Um estimador de Monte Carlo para o DIC, baseado nas 2000 amostras de Gibbs simuladas, considerando o “modelo 1” é dado por $DIC = 667,07$. Considerando o “modelo 2”, temos $DIC = 662,86$, ou seja, já que temos uma pequena diminuição no valor de DIC assumindo o “modelo 2”, podemos concluir que o “modelo 2” é mais adequado para os tempos de recorrência da infecção para os rins dos pacientes. É importante salientar que outros métodos de discriminação de modelos podem ser utilizados para decidir qual é o melhor modelo.

Mais adiante, uma modificação pode ser assumida para o “modelo 1” e “modelo 2”, introduzindo o efeito da covariável sexo (x_i) no parâmetro de forma ν_j , $j = 1, 2$.

Desta maneira, assumimos para o “modelo 1” e “modelo 2”, o modelo de regressão para o parâmetro de forma dado por,

$$\nu_j(i) = \exp\{\alpha_{1j} + \alpha_{2j}x_i\}, \quad (4.39)$$

$i = 1, 2, \dots, 38$; $j = 1, 2$.

Denotamos esses modelos como “modelo 3” e “modelo 4”.

Para o “modelo 3” e “modelo 4”, assumimos distribuições a priori informativas normais para β_{1j} e β_{2j} , considerando médias fechadas para as médias a posteriori obtidas para β_{1j} e β_{2j} assumindo “modelo 1” e “modelo 2”, respectivamente. Também assumimos priors

normais para α_{1j} e α_{2j} , $j = 1, 2$, considerando variâncias pequenas.

Na Tabela 4.5, temos os resumos à posteriori obtidos das 2000 amostras de Gibbs simuladas para a distribuição à posteriori conjunta de interesse.

Tabela 4.5: Resumos à posteriori (“modelo 3” e “modelo 4”)

Modelo	Parâmetro	Média	D.P.	Intervalo Credibilidade de 95%
“modelo 3” DIC = 660,87	β_{21}	2,0050	0,2932	(1,4410; 2,5880)
	β_{22}	1,9470	0,3007	(1,3610; 2,5370)
	β_{11}	-5,9650	0,2970	(-6,5460; -5,3840)
	β_{12}	-6,0650	0,2937	(-6,6580; -5,4970)
	α_{21}	0,0603	0,1186	(-0,1818; 0,2850)
	α_{22}	-0,1855	0,1180	(-0,4252; 0,0356)
	α_{11}	0,1395	0,0662	(0,0064; 0,2670)
	α_{12}	0,1899	0,0697	(0,0452; 0,3159)
	$1/\sigma_w^2$	1,7700	0,8131	(0,6983; 3,8310)
“modelo 4” DIC = 658,06	β_{21}	2,0170	0,3044	(1,4210; 2,5960)
	β_{22}	1,9510	0,3013	(1,3610; 2,5440)
	β_{11}	-5,9410	0,2938	(-6,5280; -5,3700)
	β_{12}	-6,0510	0,2921	(-6,6460; -5,4680)
	α_{21}	0,1142	0,1242	(-0,1362; 0,3476)
	α_{22}	-0,1263	0,1281	(-0,3848; 0,1203)
	α_{11}	0,1772	0,0696	(0,0430; 0,3143)
	α_{12}	0,2267	0,0691	(0,0838; 0,3593)
	$1/\phi$	2,1050	1,9220	(0,7696; 5,5800)

Nas Figuras 4.10 e 4.11, temos os gráficos para as amostras simuladas por Gibbs considerando o “modelo 3” e o “modelo 4”.

Nas Figuras (4.12) e (4.13) temos gráficos para os valores de Gelman & Rubin (1992) as quais, observamos a convergência para os parâmetros via algoritmo *Gibbs sampling*.

Dos resultados da Tabela 4.5, observamos que o “modelo 3” e o “modelo 4” apresentam inferências similares. Observamos que a covariável x_i (sexo) não tem um efeito significativo no parâmetro de forma da distribuição Weibull para os tempos de recorrência, já que o zero está incluso no intervalo de credibilidade de 95% para α_{21} e α_{22} assumindo ambos os modelos.

Também observamos que, o “modelo 4” apresenta um valor menor para o DIC (658,06) se comparado aos modelos 1, 2 e 3.

Outro caminho para checar se o modelo de regressão Weibull é mais adequado para os dados é assumir a distribuição gama generalizada.

Observar que a inclusão das covariáveis nos parâmetros de forma (ver (4.39)) pode ser importante para a obtenção de melhores inferências Bayesianas. Isso não é, em geral, considerado na literatura.

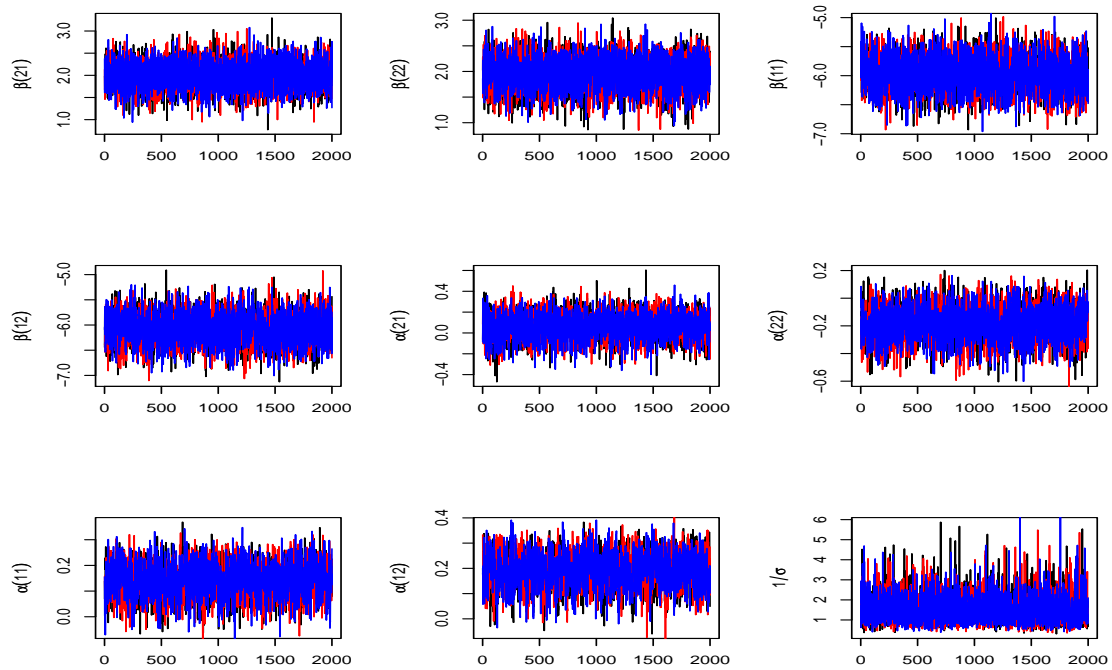


Figura 4.10: Amostras simuladas por Gibbs (“modelo 3”)

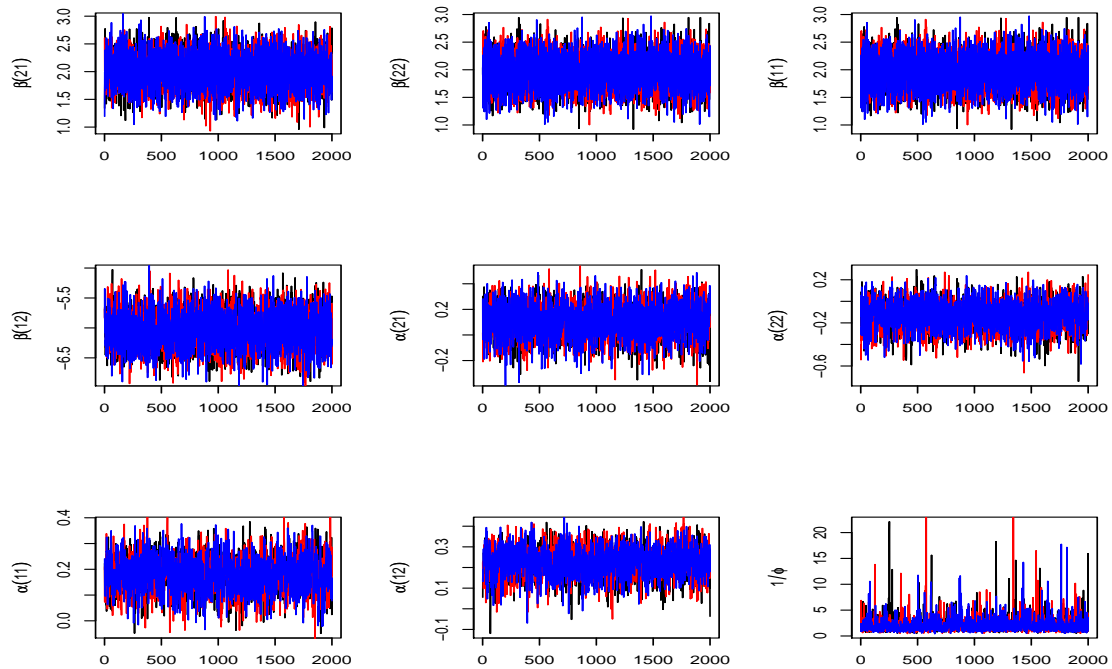


Figura 4.11: Amostras simuladas por Gibbs (“modelo 4”)

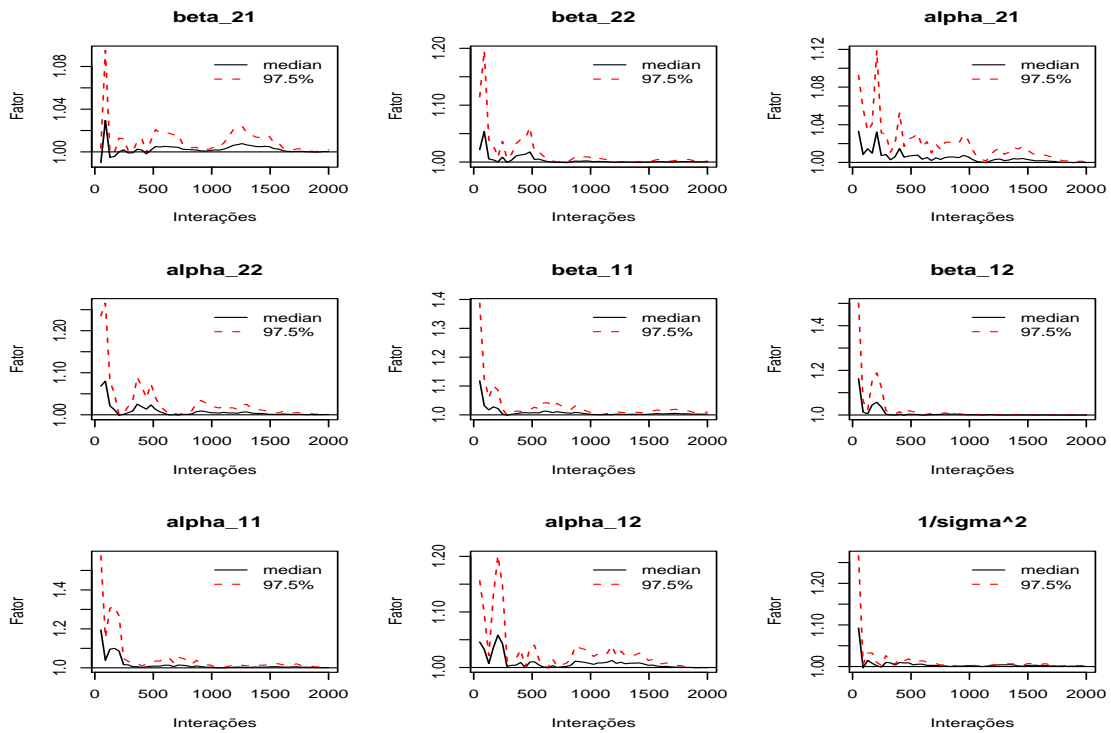


Figura 4.12: Valores de Gelman e Rubin (“modelo 3”)

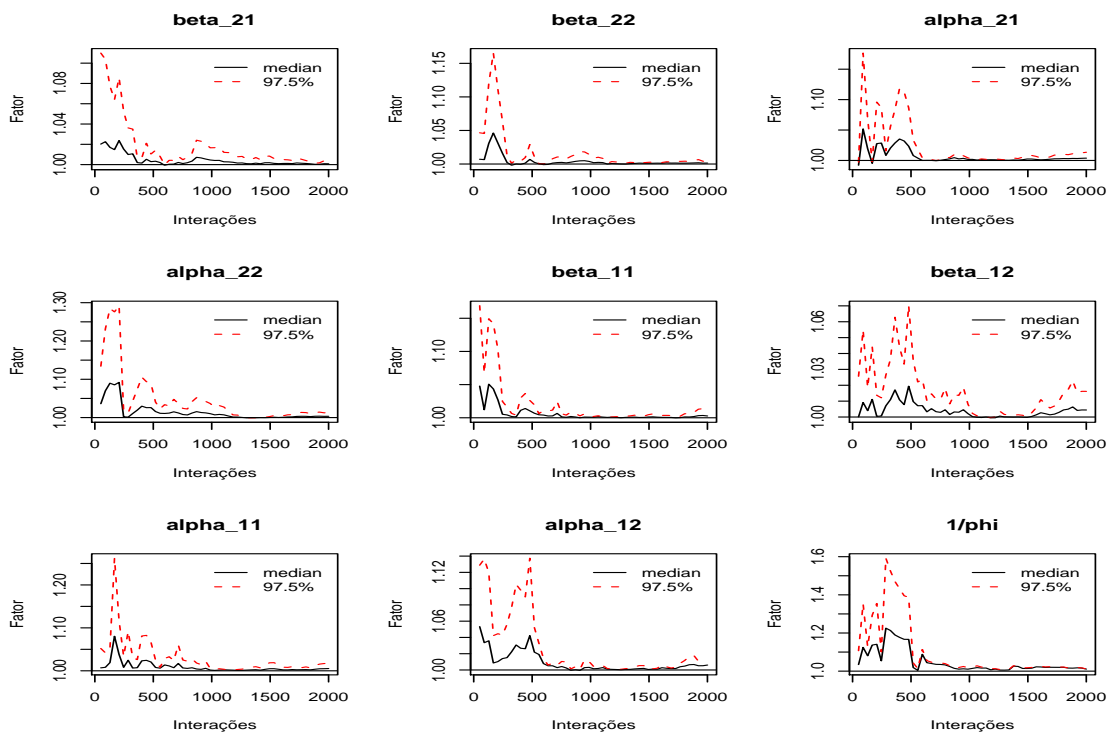


Figura 4.13: Valores de Gelman e Rubin (“modelo 4”)

Considerando o “modelo 5” com uma densidade gama generalizada (4.32) e modelo de regressão,

$$\mu_j(i) = \exp\{\beta_{1j} + \beta_{2j}x_i + w_i\}, \tag{4.40}$$

em que a “fragilidade” W_i tem uma distribuição normal (4.2), assumamos as prioris (4.36) e (4.26) para os parâmetros do modelo com os valores dos hiperparâmetros $a_1 = a_2 = b_1 = b_2 = c_1 = c_2 = d_1 = d_2 = 1$ e distribuições normais para β_{11} , β_{12} , β_{21} e β_{22} com variâncias iguais a um.

Assumindo o “modelo 6”, com uma densidade gama generalizada (4.32) e um modelo de regressão,

$$\mu_j(i) = w_i \exp\{\beta_{1j} + \beta_{2j}x_i\}, \tag{4.41}$$

em que a “fragilidade” W_i tem uma distribuição gama (4.16), assumamos as mesmas distribuições a priori consideradas para o “modelo 5”, no primeiro estágio da análise Bayesiana hierárquica e uma priori Gama(1, 1) para o parâmetro ϕ .

Na Tabela 4.6, temos os resumos a posteriori de interesse considerando o “modelo 5” e o “modelo 6”.

Tabela 4.6: Resumos a posteriori (“modelo 5” e “modelo 6”)

Modelo	Parâmetro	Média	D.P.	Intervalo Credibilidade de 95%
“modelo 5”	β_{21}	1, 8270	0, 4312	(0, 9458; 2, 6560)
	β_{22}	0, 8333	0, 4353	(-0, 0450; 1, 6790)
	β_{11}	-5, 1650	0, 6007	(-6, 1090; -3, 7340)
	β_{12}	-4, 7660	0, 5579	(-5, 7620; -3, 5470)
	θ_1	1, 4920	0, 7905	(0, 6186; 3, 7110)
	θ_2	1, 3930	0, 6994	(0, 6453; 3, 3830)
	ν_1	0, 9544	0, 5290	(0, 2398; 2, 2420)
	ν_2	1, 2580	0, 5967	(0, 3553; 2, 6830)
	$1/\sigma_w^2$	1, 9060	0, 7860	(0, 8513; 3, 9330)
“modelo 6”	β_{21}	1, 9010	0, 4240	(1, 0380; 2, 6750)
	β_{22}	0, 9467	0, 4284	(0, 1055; 1, 7830)
	β_{11}	-4, 8950	0, 6403	(-5, 8470; -3, 3000)
	β_{12}	-4, 6930	0, 5428	(-5, 6470; -3, 4300)
	θ_1	1, 3900	0, 7163	(0, 6086; 3, 3030)
	θ_2	1, 4630	0, 6724	(0, 6761; 3, 2120)
	ν_1	1, 0330	0, 5737	(0, 2642; 2, 4950)
	ν_2	1, 1450	0, 5517	(0, 3419; 2, 5530)
	$1/\phi$	2, 8710	2, 1210	(1, 1220; 7, 3010)

Observar que os intervalos de credibilidade de ν_1 e ν_2 incluem o valor 1, considerando

os modelos 5 e 6, um indicativo de que o modelo de regressão de Weibull é satisfatório para os dados.

Os resultados deste capítulo mostram que a inclusão de efeitos aleatórios com diferentes estruturas, podem ser uma boa alternativa para analisar dados de sobrevivência bivariados, na presença de dados censurados e covaráveis. Esses resultados podem ser utilizados para dados de sobrevivência multivariados de qualquer dimensão. Essa nova metodologia é de grande interesse prático.

Capítulo 5

Modelos Paramétricos Exponenciais Bivariados para Dados de Sobrevivência sem Censuras

Em muitas aplicações de testes de vida, usualmente temos dois tempos de vida T_1 e T_2 associados para cada unidade. Nestas aplicações, os modelos mais populares e mais aplicados em tempos de vida são dados pelas distribuições exponenciais bivariadas. Dentre essas distribuições exponenciais bivariadas, alguns modelos estão sendo extensivamente usados por engenheiros de confiabilidade e pesquisadores médicos: o modelo exponencial bivariado Block & Basu (1974); o modelo exponencial bivariado Farlie-Gumbel-Morgenstern (Farlie, 1960); o modelo exponencial bivariado Gumbel (1960); o modelo exponencial bivariado Freund (1961) e o modelo exponencial bivariado Marshall & Olkin (1967a,b). Outras distribuições paramétricas exponenciais bivariadas são introduzidas na literatura (ver por exemplo, Hougaard, 1986b; Downton, 1970; Arnold & Strauss, 1988; Sarkar, 1987).

Para analisar dados de tempo de vida correlacionados, também podemos considerar modelos Bayesianos hierárquicos com distribuições exponenciais para os tempos de vida T_1 e T_2 , na presença de diferentes estruturas de “fragilidades” para capturar a correlação entre os tempos de vida (ver por exemplo, Clayton & Cuzick, 1985; Oakes, 1986, 1989; Shih & Louis, 1992)(Ver resultados do Capítulo 4).

A análise Bayesiana para modelos de tempos de vida bivariados tem sido considerada como uma boa alternativa para obter as inferências de interesse, quando comparada com

a inferência clássica usando métodos de máxima verossimilhança, em termos de precisão para estimar os parâmetros de interesse. Neste caso, o uso de métodos MCMC (Markov Chain Monte Carlo) (Gelfand & Smith, 1990; Chib & Greenberg, 1995), leva a uma grande simplificação para obter os resumos a posteriori de interesse.

Como uma motivação e exemplo de dados de tempos de vida bivariados para este capítulo, consideramos o conjunto de dados obtido de Meintanis (2007) e introduzido na Tabela 5.1.

Os dados representam partidas de futebol onde ao menos um gol foi feito pelo time da casa, sendo que esse gol foi marcado diretamente de uma cobrança de penalti, falta ou qualquer outro chute direto. Neste sentido, T_1 representa o tempo, em minutos, do primeiro gol anotado por algum dos times e T_2 representa o primeiro gol anotado pelo time da casa, de qualquer uma das formas citadas anteriormente. Neste caso, todas as possibilidades estão em aberto, por exemplo, $T_1 < T_2$, ou $T_1 > T_2$ ou $T_1 = T_2 = Y$. Os jogos sem gols não foram considerados e não há censuras nos dados.

Tabela 5.1: Dados da Liga dos Campeões da UEFA

2005 – 2006	T_1	T_2	2004 – 2005	T_1	T_2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man.United-Fenerbahce	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kiev-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villareal-Benfica	72	72	Bayern-M. Tel Aviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. Tel Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

Na Figura 5.1, temos o gráfico de T_1 versus T_2 , onde notamos de fato, alguma correlação entre essas variáveis. A correlação de Pearson entre T_1 e T_2 é dada por $r = 0.465$.

Para analisar os dados de tempos de vida bivariados introduzidos na Tabela 5.1, assumimos diferentes distribuições exponenciais bivariadas, sob o paradigma Bayesiano e usando o *software WinBUGS* para obter os resumos a posteriori de interesse.

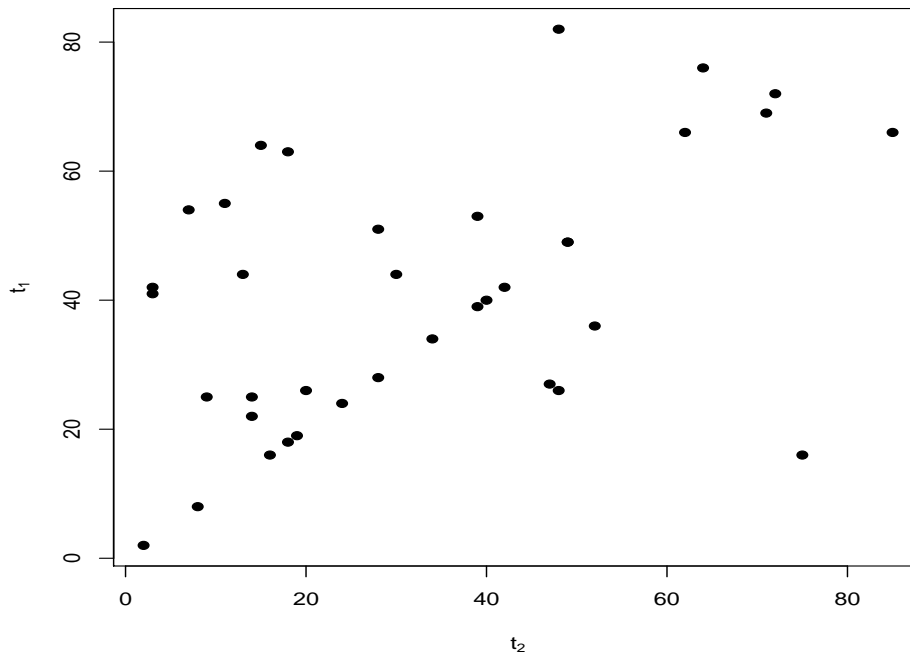


Figura 5.1: Scatterplot de t_2 vs t_1

5.1 Algumas Distribuições Exponenciais Bivariadas, num contexto Bayesiano

5.1.1 A Distribuição Exponencial Bivariada de Block & Basu

A distribuição exponencial bivariada de Block & Basu (1974) com parâmetros λ_1 , λ_2 e λ_3 para os tempos de vida T_1 e T_2 tem uma função densidade conjunta dada por,

$$f(t_1, t_2) = \begin{cases} f_1(t_1, t_2) = \frac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}} \exp\{-\lambda_1 t_1 - \lambda_{23} t_2\}, & \text{se } t_1 < t_2 \\ f_2(t_1, t_2) = \frac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}} \exp\{-\lambda_{13} t_1 - \lambda_2 t_2\}, & \text{se } t_1 \geq t_2, \end{cases} \quad (5.1)$$

sendo $\lambda_{12} = \lambda_1 + \lambda_2$, $\lambda_{13} = \lambda_1 + \lambda_3$, $\lambda_{23} = \lambda_2 + \lambda_3$ e $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ e $\lambda_3 \geq 0$.

As médias e as variâncias para T_1 e T_2 são dadas respectivamente por,

$$\mu_1 = E(T_1) = \frac{1}{\lambda_{13}} + \frac{\lambda_2 \lambda_3}{\lambda \lambda_{12} \lambda_{13}}, \tag{5.2}$$

$$\mu_2 = E(T_2) = \frac{1}{\lambda_{23}} + \frac{\lambda_1 \lambda_3}{\lambda \lambda_{12} \lambda_{23}},$$

$$\sigma_1^2 = Var(T_1) = \frac{1}{\lambda_{13}^2} + \frac{\lambda_2 \lambda_3 (2\lambda_1 \lambda + \lambda_2 \lambda_3)}{\lambda^2 \lambda_{12}^2 \lambda_{13}^2},$$

$$\sigma_2^2 = Var(T_2) = \frac{1}{\lambda_{23}^2} + \frac{\lambda_1 \lambda_3 (2\lambda_2 \lambda + \lambda_1 \lambda_3)}{\lambda^2 \lambda_{12}^2 \lambda_{23}^2}.$$

O coeficiente de correlação para T_1 e T_2 é dado por,

$$\rho_{12} = \frac{\lambda_{12} [(\lambda_1^2 + \lambda_2^2) \lambda \lambda_3 + \lambda_1 \lambda_2 \lambda_3^2]}{[\lambda^2 \lambda_{12}^2 + \lambda_2 \lambda_3 (2\lambda \lambda_1 + \lambda_2 \lambda_3)]^{1/2} [\lambda^2 \lambda_{12}^2 + \lambda_1 \lambda_3 (2\lambda \lambda_2 + \lambda_1 \lambda_3)]^{1/2}}. \tag{5.3}$$

A covariância entre T_1 e T_2 é dada por

$$Cov(T_1, T_2) = \frac{(\lambda_1^2 + \lambda_2^2) \lambda_3 \lambda + \lambda_1 \lambda_2 \lambda_3^2}{\lambda^2 \lambda_{12} \lambda_{13} \lambda_{23}}. \tag{5.4}$$

O coeficiente de correlação ρ_{12} é tal que $0 \leq \rho_{12} \leq 1$ e $\rho_{12} = 0$ somente para o caso $\lambda_3 = 0$ ou $\lambda_1 = \lambda_2 = 0$.

Considerando uma amostra aleatória de tamanho n , dada por $(T_{11}, T_{21}), (T_{12}, T_{22}), \dots, (T_{1n}, T_{2n})$, a função de verossimilhança para λ_1, λ_2 e λ_3 é dada por,

$$L(\lambda_1, \lambda_2, \lambda_3) = \prod_{i=1}^n f_1^{\delta_i}(t_{1i}, t_{2i}) f_2^{1-\delta_i}(t_{1i}, t_{2i}), \tag{5.5}$$

em que $\delta_i = 1$ se $T_{1i} < T_{2i}$ e $\delta_i = 0$ se $T_{1i} \geq T_{2i}$.

Ou seja,

$$L(\lambda_1, \lambda_2, \lambda_3) = \frac{\lambda_1^r \lambda_2^n \lambda_3^{n-r} \lambda_{23}^r \lambda_{13}^{n-r}}{\lambda_{12}^n} \exp \{-\lambda_1 n \bar{t}_1 - \lambda_2 n \bar{t}_2 - \lambda_3 R\}, \tag{5.6}$$

no qual $n \bar{t}_1 = \sum_{i=1}^n t_{1i}$, $n \bar{t}_2 = \sum_{i=1}^n t_{2i}$, $r = \sum_{i=1}^n \delta_i$ e $R = \sum_{i=1}^n [\delta_i t_{2i} + (1 - \delta_i) t_{1i}]$.

A introdução de variáveis latentes para analisar dados de sobrevivência bivariados usando o paradigma Bayesiano, tem sido considerada por muitos autores (veja por exem-

plo, Achcar & Leandro, 1998; Lee *et al.*, 2007).

Para uma análise Bayesiana do modelo Block & Basu, considerando a introdução de uma variável latente N_1 representando o número de observações tais que $T_{1i} < T_{2i}$, assumimos as seguintes densidades a priori para N_1 , λ_1 , λ_2 e λ_3 :

$$\begin{aligned} N_1 &\sim b\left(n, \frac{\lambda_1}{\lambda_{12}}\right), \\ \lambda_1 &\sim \text{Gama}(a_1, b_1), \\ \lambda_2 &\sim \text{Gama}(a_2, b_2), \\ \lambda_3 &\sim \text{Gama}(a_3, b_3), \end{aligned} \tag{5.7}$$

em que a_1 , b_1 , a_2 , b_2 , a_3 e b_3 são hiperparâmetros conhecidos, $b\left(n, \frac{\lambda_1}{\lambda_{12}}\right)$ denota uma distribuição binomial com média $\lambda_1/\lambda_{12} = P(T_1 < T_2)$; $\text{Gama}(a, b)$ denota uma distribuição gama com média a/b e variância a/b^2 . Além disso, assumimos independência entre os parâmetros N_1 , λ_1 , λ_2 e λ_3 .

A densidade a posteriori conjunta é dada por,

$$\begin{aligned} \pi(N, \lambda_1, \lambda_2, \lambda_3 | \mathbf{t}_1, \mathbf{t}_2) &\propto \binom{n}{N_1} \left(\frac{\lambda_1}{\lambda_{12}}\right)^{N_1} \left(\frac{\lambda_2}{\lambda_{12}}\right)^{n-N_1} \frac{\lambda^n}{\lambda_{12}^n} \times \\ &\times \lambda_{23}^r \lambda_{13}^{n-r} \lambda_1^{r+a_1-1} \lambda_2^{n-r+a_2-1} \lambda_3^{a_3-1} \times \\ &\times \exp\{-(n\bar{t}_1 + b_1)\lambda_1 - (n\bar{t}_2 + b_2)\lambda_2 - (R + b_3)\lambda_3\}, \end{aligned} \tag{5.8}$$

no qual $n\bar{t}_1 = \sum_{i=1}^n t_{1i}$, $n\bar{t}_2 = \sum_{i=1}^n t_{2i}$, $r = \sum_{i=1}^n \delta_i$, $R = \sum_{i=1}^n [\delta_i t_{2i} + (1 - \delta_i)t_{1i}]$, $\mathbf{t}_1 = (t_{11}, t_{12}, \dots, t_{1n})'$ e $\mathbf{t}_2 = (t_{21}, t_{22}, \dots, t_{2n})'$.

Amostras simuladas da distribuição a posteriori conjunta (5.8) podem ser obtidas usando métodos MCMC, como o amostrador de Gibbs e o algoritmo Metropolis-Hastings (ver, Achcar & Leandro, 1998).

5.1.2 A Distribuição Exponencial Correlacionada da Família Farlie-Gumbel-Morgenstern

A densidade bivariada de Farlie-Gumbel-Morgenstern (Farlie, 1960) tem uma forma geral dada por,

$$f(t_1, t_2) = f_1(t_1)f_2(t_2) + \rho f_1(t_1)f_2(t_2)[1 - 2F_1(t_1)][1 - 2F_2(t_2)], \quad (5.9)$$

em que $f_1(t_1)$ e $f_2(t_2)$ são respectivamente, as densidades marginais para T_1 e T_2 ; $F_1(t_1)$ e $F_2(t_2)$ são as funções de distribuição correspondentes a f_1 e f_2 , respectivamente; ρ é o coeficiente de correlação entre T_1 e T_2 onde $-1 \leq \rho \leq 1$.

Assumindo densidades marginais exponenciais para T_1 e T_2 , ou seja, $f_j(t_j) = \lambda_j e^{-\lambda_j t_j}$, $j = 1, 2$, a função de densidade conjunta para T_1 e T_2 é (de (5.9)) dada por,

$$f(t_1, t_2) = \lambda_1 \lambda_2 e^{-\lambda_1 t_1 - \lambda_2 t_2} + \rho \lambda_1 \lambda_2 e^{-\lambda_1 t_1 - \lambda_2 t_2} (2e^{-\lambda_1 t_1} - 1)(2e^{-\lambda_2 t_2} - 1), \quad (5.10)$$

no qual $\lambda_1 \geq 0$, $\lambda_2 \geq 0$ e $-1 \leq \rho \leq 1$.

Considerando uma amostra aleatória de tamanho n para (T_1, T_2) , a função de verossimilhança para ρ , λ_1 e λ_2 é dada por,

$$\begin{aligned} L(\rho, \lambda_1, \lambda_2) &= \prod_{i=1}^n f(t_{1i}, t_{2i}) \\ &= \prod_{i=1}^n [\lambda_1 \lambda_2 e^{-\lambda_1 t_{1i} - \lambda_2 t_{2i}} + \rho \lambda_1 \lambda_2 e^{-\lambda_1 t_{1i} - \lambda_2 t_{2i}} (2e^{-\lambda_1 t_{1i}} - 1)(2e^{-\lambda_2 t_{2i}} - 1)]. \end{aligned} \quad (5.11)$$

Para uma análise Bayesiana da distribuição exponencial bivariada de Farlie-Gumbel-Morgenstern, considerando a introdução da variável latente U_i com distribuição de Bernoulli, ou seja,

$$U_i \sim \text{Bernoulli}(\theta_i), \quad (5.12)$$

em que θ_i é dada por,

$$\theta_i = \frac{\lambda_1 \lambda_2 e^{-\lambda_1 t_{1i} - \lambda_2 t_{2i}}}{\lambda_1 \lambda_2 e^{-\lambda_1 t_{1i} - \lambda_2 t_{2i}} + \rho \lambda_1 \lambda_2 e^{-\lambda_1 t_{1i} - \lambda_2 t_{2i}} (2e^{-\lambda_1 t_{1i}} - 1)(2e^{-\lambda_2 t_{2i}} - 1)},$$

para $i = 1, 2, \dots, n$, assumimos as seguintes distribuições a priori para λ_1 , λ_2 e ρ :

$$\begin{aligned} \lambda_1 &\sim \text{Gama}(a_1, b_1), \\ \lambda_2 &\sim \text{Gama}(a_2, b_2), \\ \rho &\sim U[-1, 1], \end{aligned} \quad (5.13)$$

no qual $U[-1, 1]$ denota uma distribuição uniforme no intervalo $[-1, 1]$; a_1 , b_1 , a_2 e b_2 são

hiperparâmetros conhecidos. Também assumimos independência entre λ_1 , λ_2 e ρ .

A distribuição a posteriori conjunta para λ_1 , λ_2 e ρ dado os dados observados $\mathbf{t}_1 = (t_{11}, t_{12}, \dots, t_{1n})$, $\mathbf{t}_2 = (t_{21}, t_{22}, \dots, t_{2n})$ e dados não observados $\mathbf{u} = (u_1, u_2, \dots, u_n)$ é dada por,

$$\pi(\lambda_1, \lambda_2, \rho | \mathbf{t}_1, \mathbf{t}_2, \mathbf{u}) \propto L(\rho, \lambda_1, \lambda_2) \times \left\{ \prod_{i=1}^n \theta_i^{u_i} (1 - \theta_i)^{1-u_i} \right\} \lambda_1^{a_1-1} e^{-b_1 \lambda_1} \lambda_2^{a_2-1} e^{-b_2 \lambda_2}, \quad (5.14)$$

no qual $L(\rho, \lambda_1, \lambda_2)$ é a função de verossimilhança (5.11) e θ_i é dada por (5.12).

Ou seja,

$$\begin{aligned} \pi(\lambda_1, \lambda_2, \rho | \mathbf{t}_1, \mathbf{t}_2, \mathbf{u}) &\propto \lambda_1^{n+a_1-1} \lambda_2^{n+a_2-1} \rho^{n-v} \exp \{ -\lambda_1(b_1 + n\bar{t}_1) - \lambda_2(b_2 + n\bar{t}_2) \} \\ &\times \exp \left\{ \sum_{i=1}^n (1 - u_i) \log(2e^{-\lambda_1 t_{1i}} - 1) + \right. \\ &\quad \left. \sum_{i=1}^n (1 - u_i) \log(2e^{-\lambda_2 t_{2i}} - 1) \right\}, \end{aligned} \quad (5.15)$$

sendo $v = \sum_{i=1}^n u_i$.

Para as amostras simuladas da distribuição a posteriori conjunta de, λ_1 , λ_2 e ρ , dados os valores iniciais $\lambda_1^{(0)}$, $\lambda_2^{(0)}$ e $\rho^{(0)}$, seguimos o algoritmo:

- i Gerar U_i de uma distribuição de Bernoulli com probabilidade de sucesso θ_i , $i = 1, 2, \dots, n$ dada em (5.12);
- ii Gerar $\lambda_1^{(0)}$, $\lambda_2^{(0)}$ e $\rho^{(0)}$ das distribuições a posteriori condicionais $\pi(\lambda_1 | \lambda_2, \rho, \mathbf{t}_1, \mathbf{t}_2, \mathbf{u})$, $\pi(\lambda_2 | \lambda_1, \rho, \mathbf{t}_1, \mathbf{t}_2, \mathbf{u})$ e $\pi(\rho | \lambda_1, \lambda_2, \mathbf{t}_1, \mathbf{t}_2, \mathbf{u})$, respectivamente.

5.1.3 Distribuição Exponencial Bivariada de Gumbel

A distribuição exponencial bivariada de Gumbel (1960) tem a função de sobrevivência conjunta e a função de probabilidade conjunta dadas respectivamente por,

$$S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2) = \exp \{ -(\alpha t_1 + \beta t_2 + \theta \alpha \beta t_1 t_2) \}, \quad (5.16)$$

e

$$f(t_1, t_2) = \{(1 - \theta)\alpha\beta + \theta\alpha^2\beta t_1 + \theta\alpha\beta^2 t_2 + \theta^2\alpha^2\beta^2 t_1 t_2\} S(t_1, t_2),$$

com $\alpha \geq 0$, $\beta \geq 0$ e $0 < \theta < 1$.

A função de verossimilhança para α , β e θ baseada em uma amostra aleatória de tamanho n é dada por,

$$L(\alpha, \beta, \theta) = \prod_{i=1}^n \{(1 - \theta)\alpha\beta + \theta\alpha^2\beta t_{1i} + \theta\alpha\beta^2 t_{2i} + \theta^2\alpha^2\beta^2 t_{1i} t_{2i}\} \quad (5.17)$$

$$\times \exp \left\{ -\alpha \sum_{i=1}^n t_{1i} - \beta \sum_{i=1}^n t_{2i} + \theta\alpha\beta \sum_{i=1}^n t_{1i} t_{2i} \right\}.$$

O coeficiente de correlação entre T_1 e T_2 é dado por,

$$\rho = 1 - \frac{1}{\theta} \exp\left(\frac{1}{\theta}\right) Ei\left(\frac{1}{\theta}\right), \quad (5.18)$$

no qual $Ei(x)$ é a função exponencial integral dada por,

$$Ei(x) = - \int_{-x}^{\infty} \frac{e^{-t}}{t} dt. \quad (5.19)$$

Vamos assumir as seguintes distribuições a priori para os parâmetros α , β e θ :

$$\begin{aligned} \alpha &\sim \text{Gama}(a_1, b_1), \\ \beta &\sim \text{Gama}(a_2, b_2), \\ \theta &\sim U[0, 1], \end{aligned} \quad (5.20)$$

em que a_1 , b_1 , a_2 e b_2 são hiperparâmetros conhecidos. Também assumimos independência a priori para os parâmetros α , β e θ .

A distribuição a posteriori conjunta para α , β e θ é dada combinando a função de verossimilhança (5.17) com a priori (5.20), ou seja,

$$\begin{aligned} \pi(\alpha, \beta, \theta | \mathbf{t}_1, \mathbf{t}_2) &\propto \prod_{i=1}^n \{(1 - \theta)\alpha\beta + \theta\alpha^2\beta t_{1i} + \theta\alpha\beta^2 t_{2i} + \theta^2\alpha^2\beta^2 t_{1i} t_{2i}\} \quad (5.21) \\ &\times \exp \left\{ -(b_1 + n\bar{t}_1)\alpha - (b_2 + n\bar{t}_2)\beta + \theta\alpha\beta \sum_{i=1}^n t_{1i} t_{2i} + a_1 \log \alpha \right\} \\ &\times \exp \{a_2 \log \beta\}. \end{aligned}$$

Amostras da distribuição a posteriori (5.21) serão simuladas via métodos MCMC.

5.1.4 Distribuição Exponencial Bivariada de Freund

A distribuição exponencial bivariada de Freund (1961) tem a densidade conjunta dada por,

$$f(t_1, t_2) = \begin{cases} f_1(t_1, t_2) = \alpha_1 \beta_2 \exp \{-\beta_2 t_2 - (\alpha_1 + \alpha_2 - \beta_2)t_1\}, & \text{se } t_1 < t_2 \\ f_2(t_1, t_2) = \alpha_2 \beta_1 \exp \{-\beta_1 t_1 - (\alpha_1 + \alpha_2 - \beta_1)t_2\}, & \text{se } t_1 > t_2, \end{cases} \quad (5.22)$$

com $\alpha_1 > 0$, $\alpha_2 > 0$, $\beta_1 > 0$ e $\beta_2 > 0$.

Observar que T_1 e T_2 tem densidades marginais, dadas por misturas de exponenciais, isto é,

$$f_1(t_1) = \frac{1}{(\alpha_1 + \alpha_2 - \beta_1)} [(\alpha_1 - \beta_1)(\alpha_2 + \beta_2)e^{(\alpha_1 + \alpha_2)t_1} + \beta_1 \alpha_2 e^{-\beta_1 t_1}], \quad (5.23)$$

e

$$f_2(t_2) = \frac{1}{(\alpha_1 + \alpha_2 - \beta_2)} [(\alpha_2 - \beta_2)(\alpha_1 + \beta_1)e^{(\alpha_1 + \alpha_2)t_2} + \beta_2 \alpha_1 e^{-\beta_2 t_2}].$$

Com médias,

$$\begin{aligned} E(T_1) &= \frac{\beta_1 + \alpha_2}{\beta_1(\alpha_1 + \alpha_2)}, \\ E(T_2) &= \frac{\beta_2 + \alpha_1}{\beta_2(\alpha_1 + \alpha_2)}. \end{aligned} \quad (5.24)$$

Coefficiente de correlação dado por,

$$\rho = \frac{\beta_1 \beta_2 - \alpha_1 \alpha_2}{\sqrt{\beta_1^2 + 2\alpha_1 \alpha_2 + \alpha_2^2} \sqrt{\beta_2^2 + 2\alpha_1 \alpha_2 + \alpha_1^2}},$$

sendo $-1/3 < \rho < 1$.

A função de verossimilhança para $\alpha_1, \beta_1, \alpha_2$ e β_2 baseada em uma amostra aleatória de tamanho n , é dada por,

$$L(\alpha_1, \alpha_2, \beta_1, \beta_2) = \prod_{i=1}^n f_1^{\delta_i}(t_{1i}, t_{2i}) f_2^{1-\delta_i}(t_{1i}, t_{2i}), \quad (5.25)$$

em que $\delta_i = 1$ se $t_{1i} \leq t_{2i}$ e $\delta_i = 0$ se $t_{1i} > t_{2i}$, $i = 1, 2, \dots, n$.

Ou seja,

$$\begin{aligned} L(\alpha_1, \alpha_2, \beta_1, \beta_2) &= \alpha_1^r \beta_2^r \alpha_2^{n-r} \beta_1^{n-r} \exp \left\{ -\beta_1 \sum_{i=1}^n (1 - \delta_i) t_{1i} - \beta_2 \sum_{i=1}^n \delta_i t_{2i} \right\} \\ &\times \exp \left\{ (\alpha_1 + \alpha_2 - \beta_2) \sum_{i=1}^n \delta_i t_{1i} - (\alpha_1 + \alpha_2 - \beta_1) \sum_{i=1}^n (1 - \delta_i) t_{2i} \right\}, \end{aligned} \quad (5.26)$$

com $r = \sum_{i=1}^n \delta_i$

Assumamos distribuições a priori gama para $\alpha_1, \alpha_2, \beta_1$ e β_2 , ou seja,

$$\begin{aligned} \alpha_1 &\sim \text{Gama}(a_1, b_1), \\ \alpha_2 &\sim \text{Gama}(a_2, b_2), \\ \beta_1 &\sim \text{Gama}(c_1, d_1), \\ \beta_2 &\sim \text{Gama}(c_2, d_2), \end{aligned} \quad (5.27)$$

em que $a_1, b_1, a_2, b_2, c_1, d_1, c_2$ e d_2 são hiperparâmetros conhecidos. Também assumir independência a priori entre os parâmetros.

A distribuição a posteriori conjunta para $\alpha_1, \alpha_2, \beta_1$ e β_2 é dada por,

$$\begin{aligned} L(\alpha_1, \alpha_2, \beta_1, \beta_2 | \mathbf{t}_1, \mathbf{t}_2) &\propto \alpha_1^{r+a_1-1} \alpha_2^{n-r+a_2-1} \beta_1^{n-r+c_1-1} \beta_2^{r+c_2-1} \\ &\times \exp \left\{ - \left(b_1 + \sum_{i=1}^n \delta_i t_{1i} + \sum_{i=1}^n (1 - \delta_i) t_{2i} \right) \alpha_1 - \right. \\ &- \left(b_2 + \sum_{i=1}^n \delta_i t_{1i} + \sum_{i=1}^n (1 - \delta_i) t_{2i} \right) \alpha_2 - \\ &- \left(d_1 + \sum_{i=1}^n (1 - \delta_i) t_{1i} - \sum_{i=1}^n (1 - \delta_i) t_{2i} \right) \beta_1 - \\ &\left. - \left(d_2 + \sum_{i=1}^n \delta_i t_{2i} - \sum_{i=1}^n \delta_i t_{1i} \right) \beta_2 \right\}. \end{aligned} \quad (5.28)$$

As distribuições condicionais necessárias para o amostrador de Gibbs são dadas por,

$$\begin{aligned} \alpha_1 | \alpha_2, \beta_1, \beta_2, t_1, t_2 &\sim \text{Gama} \left(r + a_1, b_1 + \sum_{i=1}^n \delta_i t_{1i} + \sum_{i=1}^n (1 - \delta_i) t_{2i} \right), \\ \alpha_2 | \alpha_1, \beta_1, \beta_2, t_1, t_2 &\sim \text{Gama} \left(n - r + a_2, b_2 + \sum_{i=1}^n \delta_i t_{1i} + \sum_{i=1}^n (1 - \delta_i) t_{2i} \right), \\ \beta_1 | \alpha_1, \alpha_2, \beta_2, t_1, t_2 &\sim \text{Gama} \left(n - r + c_1, d_1 + \sum_{i=1}^n (1 - \delta_i) t_{1i} - \sum_{i=1}^n (1 - \delta_i) t_{2i} \right), \\ \beta_2 | \alpha_1, \alpha_2, \beta_1, t_1, t_2 &\sim \text{Gama} \left(r + c_2, d_2 + \sum_{i=1}^n \delta_i t_{2i} - \sum_{i=1}^n \delta_i t_{1i} \right). \end{aligned} \quad (5.29)$$

5.1.5 Distribuição Bivariada de Marshall & Olkin

A distribuição bivariada de Marshall & Olkin (1967a,b) tem densidade conjunta dada por,

$$f(t_1, t_2) = \begin{cases} f_1(t_1, t_2) = \theta_1 \theta_{23} \exp \{-\theta_1 t_1 - (\theta_{23} t_2)\}, & \text{se } t_1 < t_2 \\ f_2(t_1, t_2) = \theta_2 \theta_{13} \exp \{-\theta_2 t_2 - (\theta_{13} t_1)\}, & \text{se } t_1 > t_2 \\ f_3(t_1, t_2) = \theta_3 \exp \{-\theta_{123} t_2\}, & \text{se } t_1 = t_2, \end{cases} \quad (5.30)$$

no qual $\theta_1 > 0$, $\theta_2 > 0$, $\theta_3 > 0$, $\theta_{23} = \theta_2 + \theta_3$, $\theta_{13} = \theta_1 + \theta_3$ e $\theta_{123} = \theta_1 + \theta_2 + \theta_3$.

As médias para T_1 e T_2 e o coeficiente de correlação são dados, respectivamente por,

$$\begin{aligned} E(T_1) &= \frac{1}{\theta_{13}}, \\ E(T_2) &= \frac{1}{\theta_{23}}, \\ \rho &= \frac{\theta_3}{\theta_{123}}. \end{aligned} \quad (5.31)$$

A função de verossimilhança para θ_1 , θ_2 e θ_3 baseada em uma amostra aleatória de tamanho n é dada por,

$$L(\theta_1, \theta_2, \theta_3) = \prod_{i=1}^n f_1^{\delta_{1i}}(t_{1i}, t_{2i}) f_2^{\delta_{2i}}(t_{1i}, t_{2i}) f_3^{\delta_{3i}}(t_{1i}, t_{2i}), \quad (5.32)$$

em que $\delta_{1i} = 1$ se $t_{1i} \leq t_{2i}$, $\delta_{1i} = 0$ caso contrário; $\delta_{2i} = 1$ se $t_{1i} > t_{2i}$, $\delta_{2i} = 0$ em outros casos; $\delta_{3i} = 1$ se $t_{1i} = t_{2i}$ e $\delta_{3i} = 0$ em outros casos.

Ou seja,

$$\begin{aligned} L(\theta_1, \theta_2, \theta_3) &= \theta_1^{r_1} \theta_{23}^{r_1} \theta_2^{r_2} \theta_{13}^{r_2} \theta_3^{r_3} \exp \left\{ -\theta_1 \sum_{i=1}^n \delta_{1i} t_{1i} - \theta_{23} \sum_{i=1}^n \delta_i t_{2i} \right\} \\ &\times \exp \left\{ -\theta_2 \sum_{i=1}^n \delta_{2i} t_{2i} - \theta_{13} \sum_{i=1}^n \delta_{2i} t_{1i} - \theta_{123} \sum_{i=1}^n \delta_{3i} t_{2i} \right\}, \end{aligned} \quad (5.33)$$

sendo $r_1 = \sum_{i=1}^n \delta_{1i}$, $r_2 = \sum_{i=1}^n \delta_{2i}$ e $r_3 = \sum_{i=1}^n \delta_{3i}$.

Assumimos distribuições a priori gama independentes para θ_1 , θ_2 e θ_3 dadas por,

$$\begin{aligned} \theta_1 &\sim \text{Gama}(a_{11}, b_{11}), \\ \theta_2 &\sim \text{Gama}(a_{21}, b_{21}), \\ \theta_3 &\sim \text{Gama}(a_{31}, b_{31}), \end{aligned} \quad (5.34)$$

em que a_{11} , b_{11} , a_{21} , b_{21} , a_{31} e b_{31} são hiperparâmetros conhecidos. Assumir independência a priori entre θ_1 , θ_2 e θ_3 .

A distribuição a posteriori conjunta para θ_1 , θ_2 e θ_3 é dada por,

$$\begin{aligned} \pi(\theta_1, \theta_2, \theta_3 | \mathbf{t}_1, \mathbf{t}_2) &\propto \theta_1^{a_{11}+r_1-1} \theta_2^{a_{21}+r_2-1} \theta_3^{a_{31}+r_3-1} \theta_{23}^{r_1} \theta_{13}^{r_2} \\ &\times \exp \left\{ - \left(b_{11} + \sum_{i=1}^n \delta_{1i} t_{1i} \right) \theta_1 - \left(b_{21} + \sum_{i=1}^n \delta_{2i} t_{2i} \right) \theta_2 - \right. \\ &\left. - b_{31} \theta_3 - \theta_{23} \sum_{i=1}^n \delta_i t_{2i} - \theta_{13} \sum_{i=1}^n \delta_{2i} t_{1i} - \theta_{123} \sum_{i=1}^n \delta_{3i} t_{2i} \right\}. \end{aligned} \quad (5.35)$$

Também usamos o amostrador de Gibbs ou o algoritmo de Metropolis-Hastings para simular amostras da distribuição a posteriori (5.35).

5.1.6 Presença de “Fragilidades” para Capturar a Correlação entre Distribuições Exponenciais

Para capturar a correlação entre dois tempos de vida, também podemos considerar a introdução de “fragilidades” ou variáveis latentes (ver por exemplo, Clayton & Cuzick, 1985; Oakes, 1986, 1989; Shih & Louis, 1992) assumindo densidades marginais exponenciais dadas por,

$$f_1(t_{1i}) = \lambda_{1i} \exp(-\lambda_{1i}t_{1i}), \tag{5.36}$$

$$f_2(t_{2i}) = \lambda_{2i} \exp(-\lambda_{2i}t_{2i}),$$

para $i = 1, 2, \dots, n$; sendo

$$\lambda_{1i} = \alpha_1 \exp(-w_i),$$

$$\lambda_{2i} = \alpha_2 \exp(-w_i).$$

Assumindo uma distribuição normal para as “fragilidades” ou variáveis latentes W_i , dada por,

$$W_i \sim N(0, \sigma_w^2). \tag{5.37}$$

Observar que esse modelo é obtido dos resultados do Capítulo 4, considerando os parâmetros de forma das distribuições de Weibull iguais a 1.

Para uma análise Bayesiana hierárquica deste modelo, assumimos no primeiro estágio, as seguintes distribuições a priori:

$$\alpha_1 \sim \text{Gama}(a_1, b_1), \tag{5.38}$$

$$\alpha_2 \sim \text{Gama}(a_2, b_2),$$

sendo a_1, b_1, a_2 e b_2 hiperparâmetros conhecidos. Assumir independência a priori entre α_1 e α_2 .

Para o segundo estágio da análise Bayesiana hierárquica, assumimos uma distribuição gama para $\xi_w = 1/\sigma_w^2$, ou seja,

$$\xi_w \sim \text{Gama}(c, d), \tag{5.39}$$

em que c e d são hiperparâmetros conhecidos. Observe que σ_w^2 tem uma distribuição gama

inversa,

$$\sigma_w^2 \sim \text{GI}(c, d). \quad (5.40)$$

Baseado em uma amostra aleatória de tamanho n , dado $w_i, t_{1i}, t_{2i}, i = 1, 2, \dots, n$, a função de verossimilhança para α_1 e α_2 é dada por,

$$L(\alpha_1, \alpha_2) = \alpha_1^n \alpha_2^n \exp \left\{ -2 \sum_{i=1}^n w_i - \alpha_1 \sum_{i=1}^n t_{1i} e^{-w_i} - \alpha_2 \sum_{i=1}^n t_{2i} e^{-w_i} \right\}. \quad (5.41)$$

A distribuição a posteriori conjunta para α_1, α_2 e σ_w^2 é dada por,

$$\begin{aligned} \pi(\alpha_1, \alpha_2, \sigma_w^2 | \mathbf{t}_1, \mathbf{t}_2, \mathbf{w}) &\propto \alpha_1^{a_1+n-1} \alpha_2^{a_2+n-1} (\sigma_w^2)^{-\left(\frac{n}{2}+e+1\right)} \\ &\times \exp \left\{ -\frac{1}{\sigma_w^2} \left(f + \frac{1}{2} \sum_{i=1}^n w_i^2 \right) \right\} \exp \left\{ -2 \sum_{i=1}^n w_i - \right. \\ &\left. - \left(b_1 + \sum_{i=1}^n t_{1i} e^{-w_i} \right) \alpha_1 - \left(b_2 + \sum_{i=1}^n t_{2i} e^{-w_i} \right) \alpha_2 \right\}, \end{aligned} \quad (5.42)$$

com $\alpha_1 > 0, \alpha_2 > 0$ e $\sigma_w^2 > 0$.

Para simular as amostras da distribuição a posteriori conjunta (5.42) seguimos o algoritmo,

i Gerar a variável latente não observável W_i pela distribuição condicional

$$\pi(w_i | \alpha_1, \alpha_2, \sigma_w^2, \mathbf{t}_1, \mathbf{t}_2) \propto \exp \left\{ -\frac{w_i^2}{2\sigma_w^2} \right\} \exp \left\{ -2w_i - \alpha_1 t_{1i} e^{-w_i} - \alpha_2 t_{2i} e^{-w_i} \right\},$$

para $i = 1, 2, \dots, n$;

ii Gerar α_1 de,

$$\alpha_1 | \alpha_2, \sigma_w^2, \mathbf{t}_1, \mathbf{t}_2, \mathbf{w} \sim \text{Gama} \left(n + a_1, b_1 + \sum_{i=1}^n t_{1i} e^{-w_i} \right); \quad (5.43)$$

iii Gerar α_2 de,

$$\alpha_2 | \alpha_1, \sigma_w^2, \mathbf{t}_1, \mathbf{t}_2, \mathbf{w} \sim \text{Gama} \left(n + a_2, b_2 + \sum_{i=1}^n t_{2i} e^{-w_i} \right);$$

iv Gerar σ_w^2 de,

$$\sigma_w^2 | \alpha_1, \alpha_2, \mathbf{t}_1, \mathbf{t}_2, \mathbf{w} \sim \text{IG} \left(c + \frac{n}{2}, d + \frac{1}{2} \sum_{i=1}^n w_i^2 \right).$$

5.2 Análise dos dados da Liga dos Campeões da UEFA

Para analisar os dados da liga dos campeões da UEFA introduzido na Tabela 5.1, vamos assumir as diferentes distribuições exponenciais bivariadas introduzidas neste capítulo.

Assumindo a distribuição exponencial bivariada de Block & Basu (5.1) com valores dos hiperparâmetros $a_1 = a_2 = a_3 = b_1 = b_2 = b_3 = 0,01$ para as distribuições a priori de λ_1, λ_2 e λ_3 apresentadas em (5.7) (priors aproximadamente não-informativas), usamos o *software WinBUGS* para gerar amostras da distribuição a posteriori para λ_1, λ_2 e λ_3 . Geramos 55000 amostras, onde as primeiras 5000 amostras foram descartadas (“*burn-in sample*”) para eliminar o efeito dos valores iniciais dos parâmetros do modelo. Escolhendo cada 10^a amostra simulada por Gibbs, obtemos uma amostra final de tamanho 5000 para os sumários a posteriori de interesse (veja Tabela 5.2).

A convergência do algoritmo amostrador de Gibbs foi monitorada usando métodos existentes, como gráficos de séries temporais e o de Gelman & Rubin para as amostras simuladas por Gibbs.

Tabela 5.2: Resumos a posteriori (distribuição de Block & Basu)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
λ_1	0,00015	0,00037	(6,13E - 13; 0,0012)
λ_2	0,00074	0,00151	(3,69E - 12; 0,0054)
λ_3	0,04442	0,00530	(0,0344; 0,0552)

Nas Figuras 5.2 e 5.3 observamos através dos gráficos de séries temporais e o de Gelman & Rubin (1992), a convergência do algoritmo considerando diferentes valores iniciais para os parâmetros.

Na Tabela 5.9, temos os valores obtidos para o DIC e a soma dos logaritmos das CPO's, baseado nas 5000 amostras geradas por Gibbs.

Assumindo a distribuição exponencial bivariada de Farlie-Gumbel-Morgenstern (5.10) e as distribuições a priori (5.13) com valores dos hiperparâmetros $a_1 = a_2 = b_1 = b_2 = 0,1$,

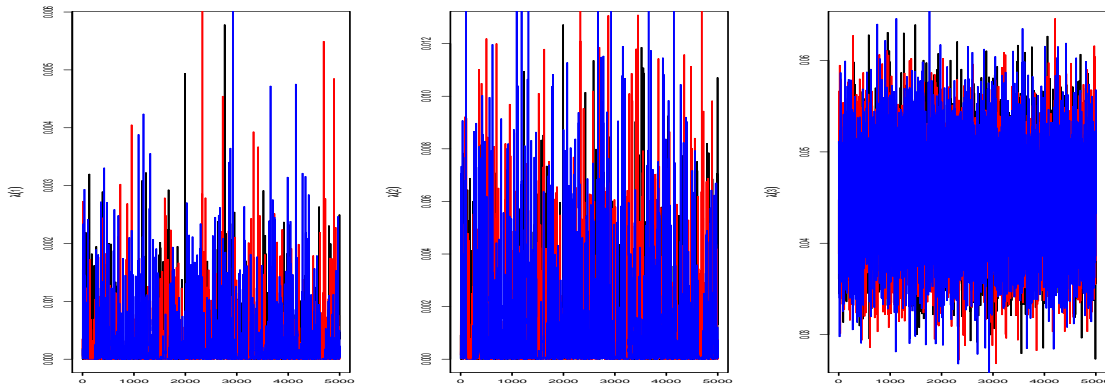


Figura 5.2: Amostras simuladas por Gibbs (distribuição de Block & Basu)

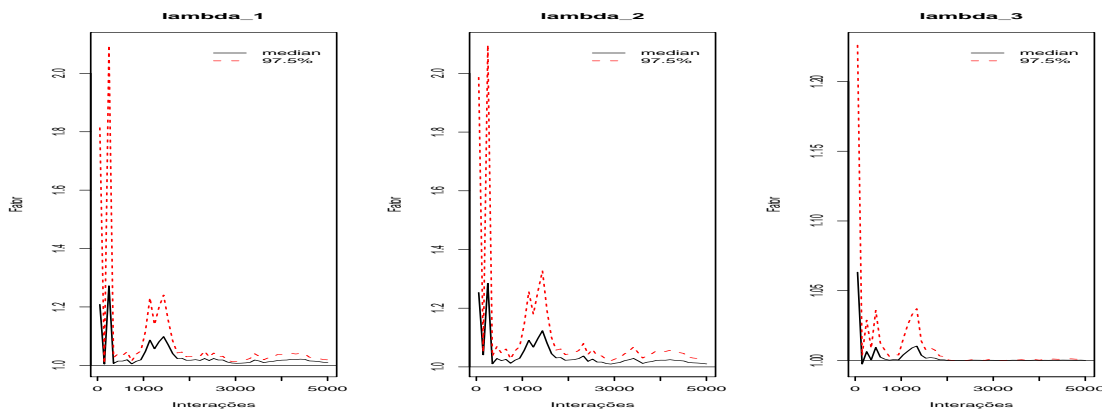


Figura 5.3: Valores de Gelman e Rubin (distribuição de Block & Basu)

e usando os mesmos passos MCMC usados para o modelo de Block & Basu, temos na Tabela 5.3 os resumos a posteriori considerando as 5000 amostras geradas por Gibbs obtidas usando o *software WinBUGS*.

Tabela 5.3: Resumos a posteriori (distribuição de Farlie-Gumbel-Morgenstern)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
λ_1	0,02551	0,00414	(0,0181; 0,0340)
λ_2	0,03272	0,00533	(0,0229; 0,0438)
ρ	0,66870	0,29291	(-0,0902; 0,9900)

Nas Figuras 5.4 e 5.5 observamos a convergência para os parâmetros do modelo de Farlie-Gumbel-Morgenstern.

Na Tabela 5.9, temos estimadores de Monte Carlo para o DIC e a soma dos logaritmos das CPO's.

Assumindo a distribuição exponencial bivariada de Gumbel (5.16) com valores dos hiperparâmetros $a_1 = a_2 = b_1 = b_2 = 0,01$ para as distribuições a priori (5.20) (prioris

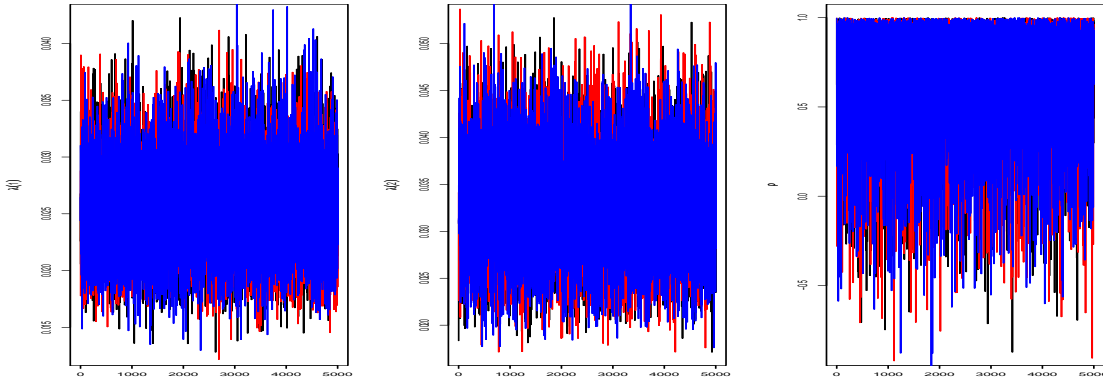


Figura 5.4: Amostras simuladas por Gibbs (distribuição de Farlie-Gumbel-Morgenstern)

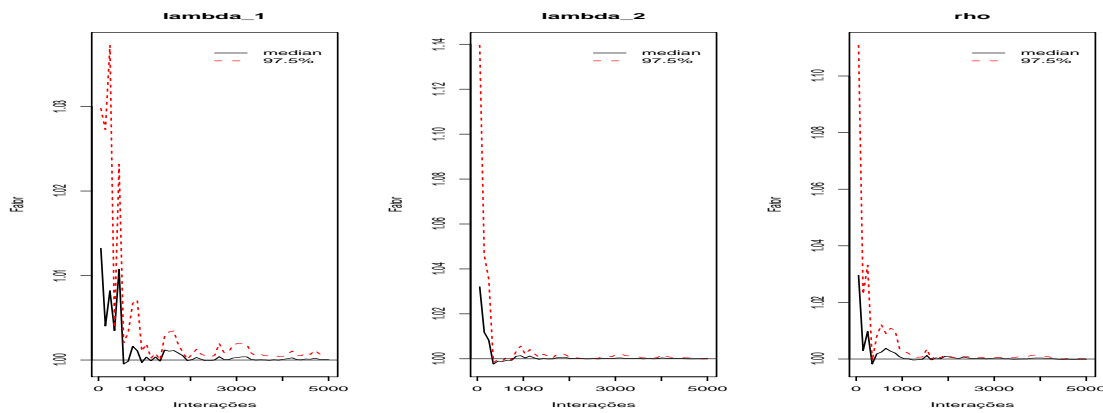


Figura 5.5: Valores de Gelman e Rubin (distribuição de Farlie-Gumbel-Morgenstern)

não-informativas), temos na Tabela 5.4, os resumos a posteriori de interesse baseado nas 5000 amostras geradas por Gibbs.

Tabela 5.4: Resumos a posteriori (distribuição de Gumbel)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
α	0,02383	0,00403	(0,0167; 0,0236)
β	0,02934	0,00501	(0,0203; 0,0291)
θ	0,17190	0,15822	(0,0045; 0,1242)

Nas Figuras 5.6 e 5.7 observamos a convergência para os parâmetros do modelo de Gumbel.

Assumindo a distribuição exponencial bivariada de Freund (5.22) com valores do hiperparâmetros $a_1 = a_2 = b_1 = b_2 = c_1 = c_2 = d_1 = d_2 = 0,1$ para as distribuições a priori (5.27), temos na Tabela 5.5 os resumos a posteriori de interesse baseado nas 5000 amostras geradas por Gibbs.

Nas Figuras 5.8 e 5.9 observamos a convergência para os parâmetros do modelo de

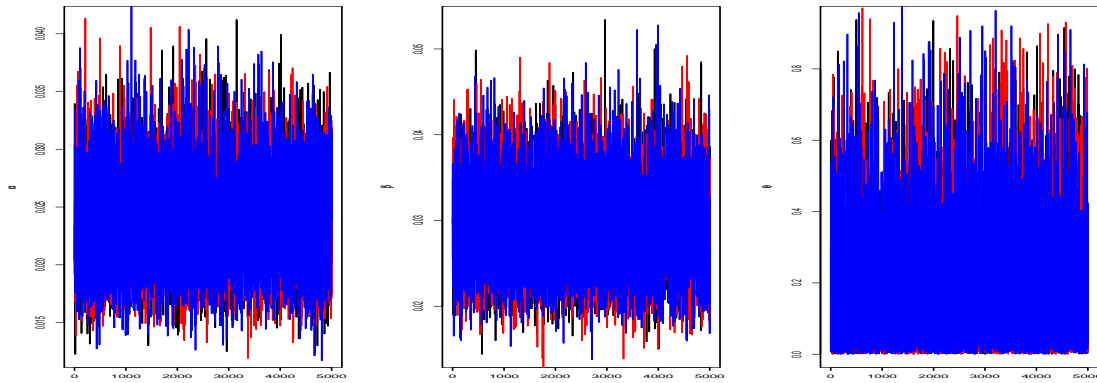


Figura 5.6: Amostras simuladas por Gibbs (distribuição de Gumbel)

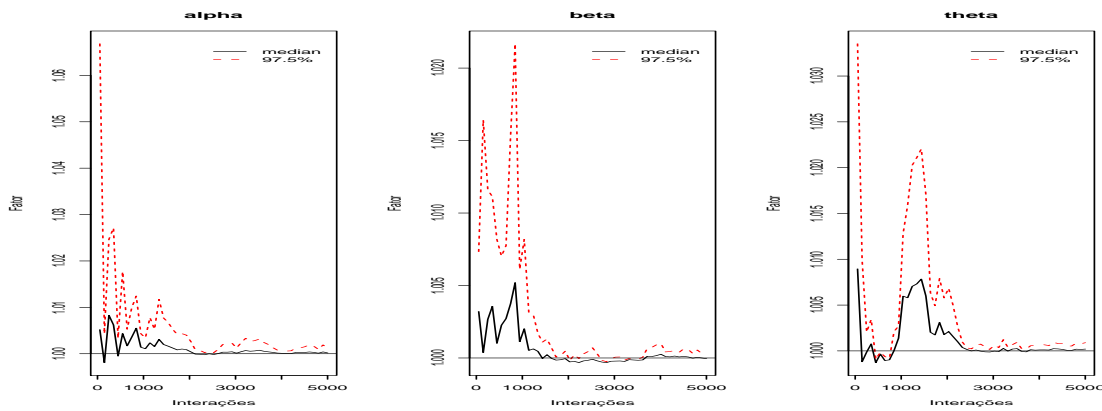


Figura 5.7: Valores de Gelman e Rubin (distribuição de Gumbel)

Tabela 5.5: Resumos a posteriori (distribuição de Freund)

Parâmetros	Média	D.P.	Intervalo de Credibilidade de 95%
α_1	0,00561	0,00235	(0,0021; 0,0108)
α_2	0,02890	0,00521	(0,0196; 0,0402)
β_1	0,07162	0,01282	(0,0487; 0,0982)
β_2	0,04424	0,01796	(0,0164; 0,0856)

Freund.

Assumindo a distribuição exponencial bivariada de Marshall & Olkin (5.30), com valores dos hiperparâmetros $a_{11} = b_{11} = a_{21} = b_{21} = 0, 1$ e $a_{31} = b_{31} = 1$ para as distribuições a priori (5.34), temos na Tabela 5.6 os resumos a posteriori de interesse baseado nas 5000 amostras geradas por Gibbs.

Nas Figuras 5.10 e 5.11 observamos a convergência para os parâmetros do modelo de Marshall & Olkin.

Assumindo a presença de “fragilidade” normal para capturar a correlação entre duas distribuições exponenciais (5.36) e (5.37) com valores dos hiperparâmetros $a_1 = b_1 =$

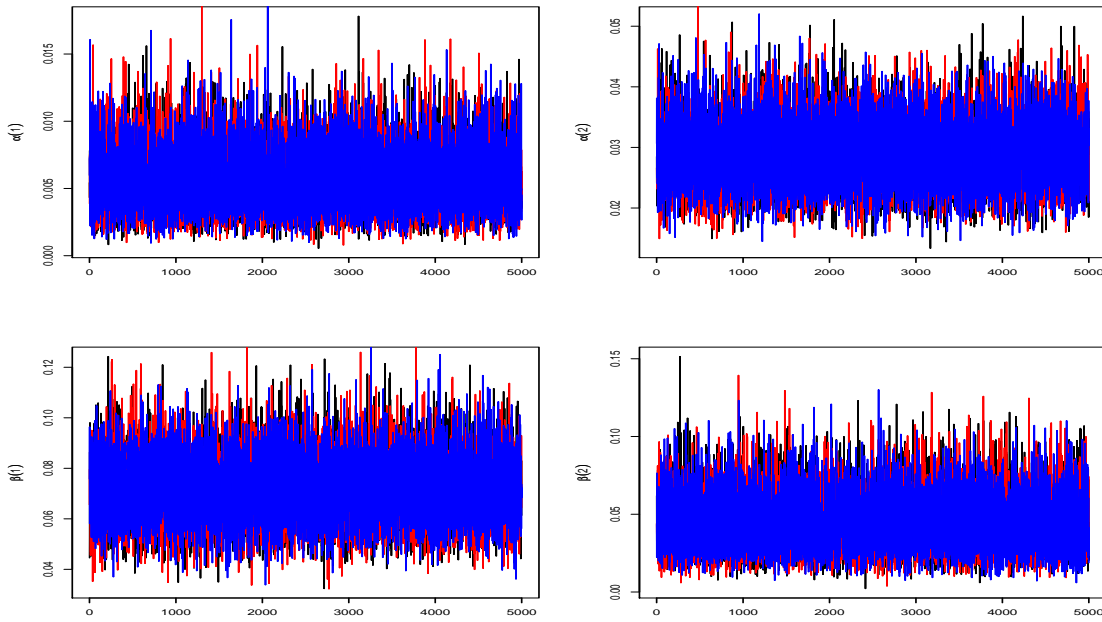


Figura 5.8: Amostras simuladas por Gibbs (distribuição de Freund)

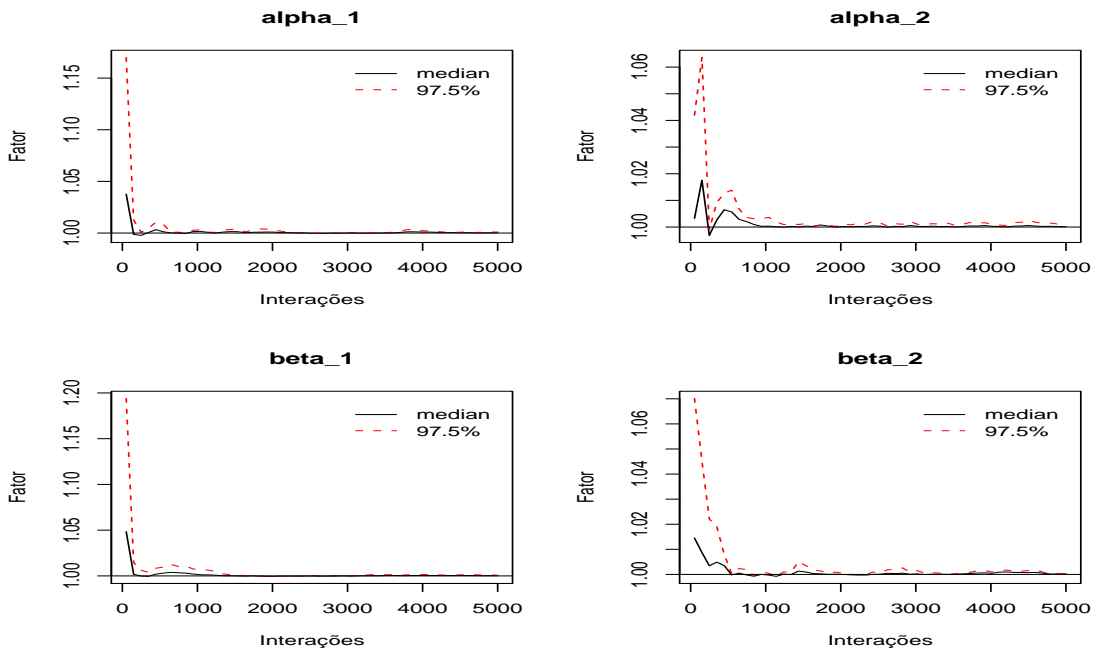


Figura 5.9: Valores de Gelman e Rubin (distribuição de Freund)

$a_2 = b_2 = 0,01$ (priors não-informativas) para as distribuições a priori (5.38) e valores dos hiperparâmetros $c = d = 1$ para a distribuição a priori (5.40), nós temos na Tabela 5.7 os resumos a posteriori de interesse baseados nas 5000 amostras geradas por Gibbs.

Nas Figuras 5.12 e 5.13 observamos a convergência para os parâmetros do modelo com

Tabela 5.6: Resumos a posteriori (distribuição de Marshall & Olkin)

Parâmetros	Média	D.P.	Intervalo de Credibilidade de 95%
θ_1	0,02143	0,00470	(0,0120; 0,0309)
θ_2	0,03012	0,00514	(0,0209; 0,0407)
θ_3	0,00393	0,00331	(0,0001; 0,0122)

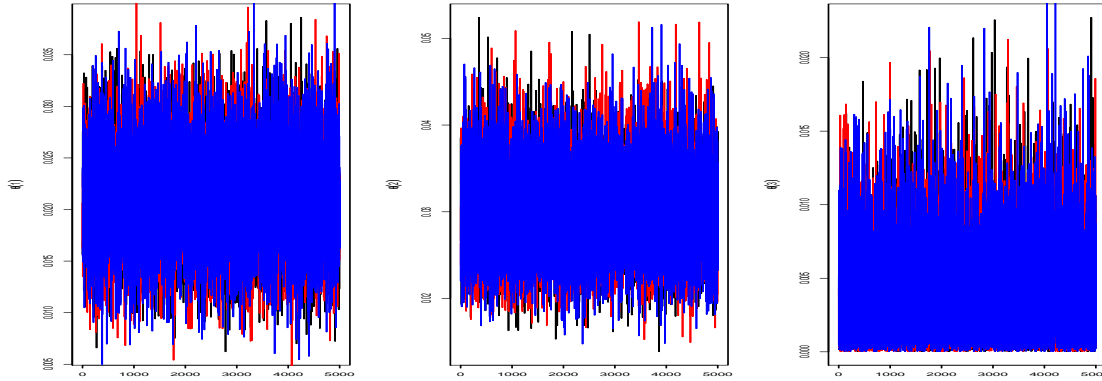


Figura 5.10: Amostras simuladas por Gibbs (distribuição de Marshall & Olkin)

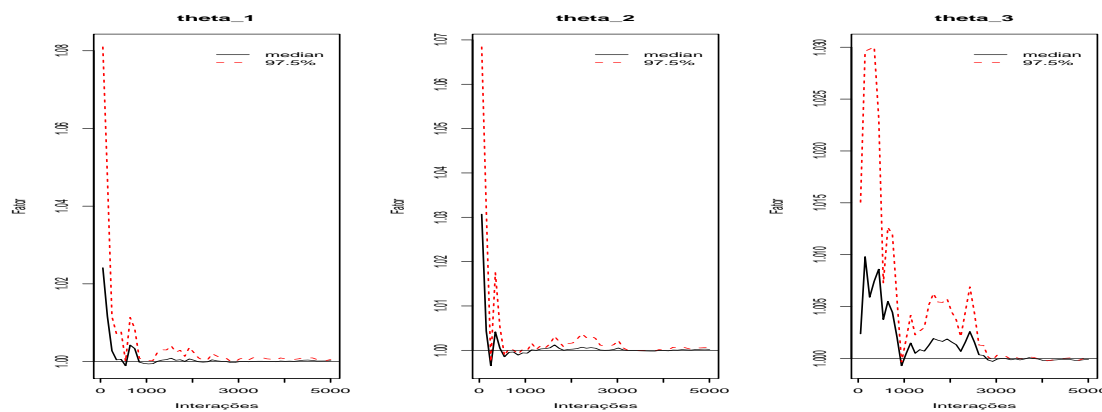


Figura 5.11: Valores de Gelman e Rubin (distribuição de Marshall & Olkin)

Tabela 5.7: Resumos a posteriori (presença de “fragilidade”)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
α_1	0,02403	0,00452	(0,0160; 0,0337)
α_2	0,03122	0,00601	(0,0207; 0,0442)
σ_w^2	0,26210	0,10550	(0,1210; 0,5242)
$\tau = 1/\sigma_w^2$	4,39703	1,65704	(1,9090; 8,2980)

presença de “fragilidade”.

Na Tabela 5.8 temos o resumo a posteriori de interesse assumindo distribuições exponenciais independentes, com densidades $f(t_j) = \lambda_j \exp\{-\lambda_j t_j\}$, $j = 1, 2$ e distribuições a priori gama independentes (Gama(0, 1; 0, 1)) para λ_1 e λ_2 , respectivamente, também con-

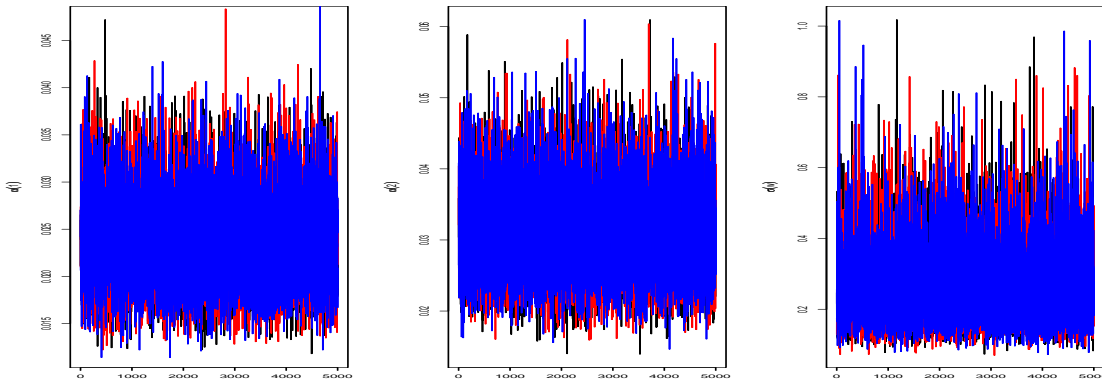


Figura 5.12: Amostras simuladas por Gibbs (modelo de “fragilidade”)

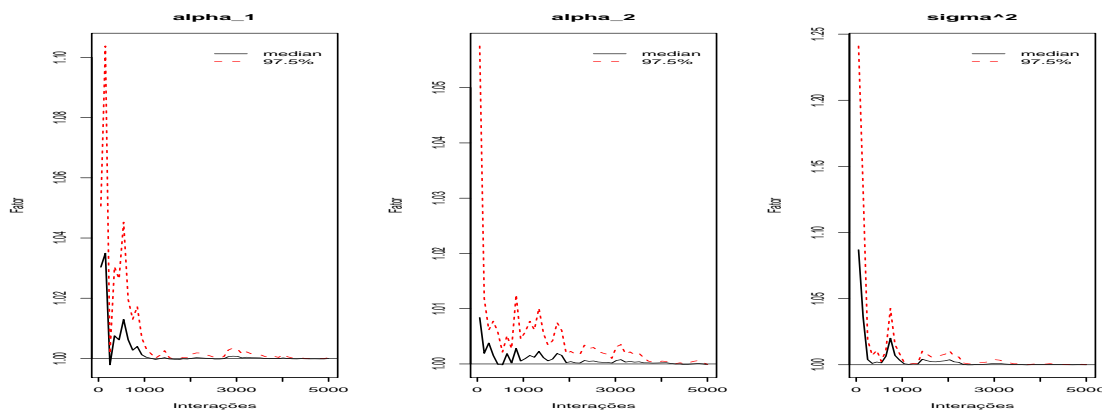


Figura 5.13: Valores de Gelman e Rubin (modelo de “fragilidade”)

siderando 5000 amostras geradas por Gibbs. Este modelo foi considerado com finalidade comparativa com os outros modelos apresentados.

Tabela 5.8: Resumos a posteriori (distribuições exponenciais independentes)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
λ_1	0,02464	0,00402	(0,0173; 0,0332)
λ_2	0,03051	0,00503	(0,0214; 0,0410)

A partir dos resultados da Tabela 5.9 concluímos que o modelo de Freund é melhor ajustado pelos dados, quando consideramos o critério DIC; considerando a soma dos logaritmos das CPO's observamos que o modelo com presença de “fragilidade” normal apresenta maior valor, isto é, uma indicação que este modelo é melhor ajustado pelos dados. Observar que na prática, poderíamos ter a presença de dados censurados e covariáveis. Isso em geral pode acarretar grandes dificuldades inferenciais, considerando cada um dos modelos propostos assumindo métodos clássicos de inferência.

O uso de métodos Bayesianos, especialmente considerando o *software WinBUGS* abre

Tabela 5.9: Estimadores de Monte Carlo para o DIC soma dos logaritmos das CPO's

Modelo	DIC	Soma dos logaritmos das CPO's
Block & Basu	644,749	-99,167
Farlie, Morgenstern & Gumbell	680,998	-125,435
Gumbel	686,951	-100,228
Freund	639,775	-107,974
Marshall & Olkin	687,102	-115,039
Presença de "fragilidade" normal	696,838	-97,988
Distribuições exponenciais independentes	685,053	--

novos horizontes na análise de dados de sobrevivência multivariados.

Este estudo ilustra o uso de diferentes modelos candidatos para analisar dados de sobrevivência bivariados. É importante salientar que para cada conjunto de dados, um dos modelos propostos pode se mostrar superior aos outros. Além disso, o custo computacional pode variar muito para os diferentes modelos.

Estudos de simulação para os modelos propostos poderiam ser úteis para avaliar as propriedades dos estimadores Bayesianos obtidos. Também seria interessante avaliar os efeitos de reparametrizações para cada modelo.

Capítulo 6

Análise Bayesiana da Distribuição Exponencial Bivariada de Block & Basu com Censuras e Covariáveis

Para analisar tempos de vida bivariados podemos assumir diferentes distribuições paramétricas introduzidas na literatura (Freund, 1961; Marshall & Olkin, 1967a; Sarkar, 1987; Downton, 1970; Gumbel, 1960; Hawkes, 1972; Hougaard, 1986b; Arnold & Strauss, 1988) (ver Capítulo 5). Dentre todas essas distribuições de tempos de vida bivariados, um modelo tem sido muito bem explorado na literatura: a distribuição exponencial de Block & Basu (1974). Neste capítulo apresentamos essa distribuição e generalizamos os resultados obtidos por Achcar & Leandro (1998) considerando a presença de dados censurados e usando o *software WinBUGS*. Essa distribuição se destaca entre as várias distribuições introduzidas no Capítulo 5, pois as marginais tem distribuições dadas por misturas de exponenciais, a distribuição conjunta é absolutamente contínua e o coeficiente de correlação é positivo. Esta distribuição foi derivada originalmente por excluir a parte singular da distribuição de Marshall & Olkin (1967b). A distribuição de Marshall & Olkin não é absolutamente contínua, mas satisfaz a propriedade da falta de memória.

A distribuição exponencial bivariada de Block & Basu com parâmetros λ_1 , λ_2 e λ_3 , introduzidas em (5.1) para os tempos de sobrevivência T_1 e T_2 tem uma função densidade conjunta dada por,

$$f(t_1, t_2) = \begin{cases} f_1(t_1, t_2) = \frac{\lambda\lambda_1\lambda_{23}}{\lambda_{12}} \exp\{-\lambda_1 t_1 - \lambda_{23} t_2\}, & \text{se } t_1 < t_2 \\ f_2(t_1, t_2) = \frac{\lambda\lambda_2\lambda_{13}}{\lambda_{12}} \exp\{-\lambda_{13} t_1 - \lambda_2 t_2\}, & \text{se } t_1 \geq t_2, \end{cases} \quad (6.1)$$

com $\lambda_{12} = \lambda_1 + \lambda_2$, $\lambda_{13} = \lambda_1 + \lambda_3$, $\lambda_{23} = \lambda_2 + \lambda_3$ e $\lambda = \lambda_1 + \lambda_2 + \lambda_3$.

A função de sobrevivência conjunta para a distribuição de Block & Basu é dada por,

$$S(t_1, t_2) = P(T_1 > t_1, T_2 > t_2) = \begin{cases} S_1(t_1, t_2) & \text{se } t_1 < t_2 \\ S_2(t_1, t_2) & \text{se } t_1 \geq t_2, \end{cases} \quad (6.2)$$

em que

$$S_1(t_1, t_2) = \frac{\lambda}{\lambda_{12}} \exp\{-\lambda_1 t_1 - \lambda_{23} t_2\} - \frac{\lambda_3}{\lambda_{12}} \exp\{-\lambda t_2\},$$

e

$$S_2(t_1, t_2) = \frac{\lambda}{\lambda_{12}} \exp\{-\lambda_{13} t_1 - \lambda_2 t_2\} - \frac{\lambda_3}{\lambda_{12}} \exp\{-\lambda t_1\}.$$

A distribuição de Block & Basu com densidade (6.1) tem distribuições marginais dadas por misturas de exponenciais para T_1 e T_2 (veja Block & Basu, 1974), e função geradora conjunta dada por,

$$m(s_1, s_2) = \frac{\lambda}{\lambda_{12}(\lambda - s_1 - s_2)} \left\{ \frac{\lambda_1 \lambda_{23}}{\lambda_{23} - s_2} + \frac{\lambda_2 \lambda_{13}}{\lambda_{13} - s_1} \right\}. \quad (6.3)$$

De (6.3), encontramos os momentos de interesse para T_1 e T_2 ; dessa forma as médias e as variâncias para T_1 e T_2 são dadas por (5.2) e a covariância entre T_1 e T_2 é dada por (5.4).

Usamos métodos MCMC para obter os resumos à posteriori de interesse. Achcar & Leandro (1998) introduzem uma análise Bayesiana da distribuição de Block & Basu, usando métodos MCMC não considerando a presença de observações censuradas. Neste capítulo generalizamos esses resultados incluindo a presença de dados censurados e covariáveis.

6.1 Uma Análise Bayesiana na Presença de Dados Censurados

Suponha que ambos, T_1 ou T_2 podem ser censurados e que a censura é independente dos tempos de sobrevivência. Vamos sub-dividir as n observações em quatro classes:

C_1 : ambos t_{1i} e t_{2i} são tempos de sobrevivência observados;

C_2 : t_{1i} é um tempo de sobrevivência e t_{2i} é um tempo de censura (ou seja, sabemos somente que $T_{2i} \geq t_{2i}$);

C_3 : t_{1i} é um tempo de censura e t_{2i} é um de sobrevivência;

C_4 : ambos t_{1i} e t_{2i} são tempos de censura.

A função de verossimilhança para um modelo contínuo (ver por exemplo, Lawless, 1982, página 479) é dada por,

$$L = \prod_{i \in C_1} f(t_{1i}, t_{2i}) \prod_{i \in C_2} \left(-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} \right) \prod_{i \in C_3} \left(-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} \right) \prod_{i \in C_4} S(t_{1i}, t_{2i}), \quad (6.4)$$

em que $f(t_{1i}, t_{2i})$ é dado por (6.1); $S(t_{1i}, t_{2i})$ é dado por (6.2) e $-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}}$ e $-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}}$ são dados, respectivamente, por,

$$-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} = \begin{cases} S'_{1t_1}(t_{1i}, t_{2i}) & \text{se } t_{1i} < t_{2i} \\ S'_{2t_1}(t_{1i}, t_{2i}) & \text{se } t_{1i} \geq t_{2i}, \end{cases} \quad (6.5)$$

com

$$S'_{1t_1}(t_{1i}, t_{2i}) = \frac{\lambda\lambda_1}{\lambda_{12}} \exp \{-\lambda_1 t_{1i} - \lambda_{23} t_{2i}\},$$

e

$$S'_{2t_1}(t_{1i}, t_{2i}) = \frac{\lambda\lambda_{13}}{\lambda_{12}} \exp \{-\lambda_{13} t_{1i} - \lambda_2 t_{2i}\} - \frac{\lambda\lambda_3}{\lambda_{12}} \exp \{-\lambda t_{1i}\}.$$

E

$$-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} = \begin{cases} S'_{1t_2}(t_{1i}, t_{2i}) & \text{se } t_{1i} < t_{2i} \\ S'_{2t_2}(t_{1i}, t_{2i}) & \text{se } t_{1i} \geq t_{2i}, \end{cases} \quad (6.6)$$

com

$$S'_{1t_2}(t_{1i}, t_{2i}) = \frac{\lambda\lambda_{23}}{\lambda_{12}} \exp\{-\lambda_1 t_{1i} - \lambda_{23} t_{2i}\} - \frac{\lambda\lambda_3}{\lambda_{12}} \exp\{-\lambda t_{2i}\},$$

e

$$S'_{2t_2}(t_{1i}, t_{2i}) = \frac{\lambda\lambda_2}{\lambda_{12}} \exp\{-\lambda_{13} t_{1i} - \lambda_2 t_{2i}\}.$$

Definamos as variáveis indicadoras δ_{1i} , δ_{2i} e v_i , por,

$$\delta_{ji} = \begin{cases} 1 & \text{se } t_{ji} \text{ é um tempo de sobrevivência observado} \\ 0 & \text{se } t_{ji} \text{ é uma observação censurada,} \end{cases} \quad (6.7)$$

para $j = 1, 2$; $i = 1, 2, \dots, n$, e,

$$v_i = \begin{cases} 1 & \text{se } t_{1i} < t_{2i} \\ 0 & \text{se } t_{1i} \geq t_{2i}. \end{cases} \quad (6.8)$$

Neste sentido, o logaritmo da função de verossimilhança para λ_1 , λ_2 e λ_3 (ver (6.4)) é dado por,

$$\begin{aligned} l(\lambda_1, \lambda_2, \lambda_3) &= \log L(\lambda_1, \lambda_2, \lambda_3) = \exp \left\{ \sum_{i=1}^n v_i \delta_{1i} \delta_{2i} \log f_1(t_{1i}, t_{2i}) + \right. \\ &+ \sum_{i=1}^n (1 - v_i) \delta_{1i} \delta_{2i} \log f_2(t_{1i}, t_{2i}) + \sum_{i=1}^n v_i \delta_{1i} (1 - \delta_{2i}) \times \\ &\times \log S'_{1t_1}(t_{1i}, t_{2i}) + \sum_{i=1}^n (1 - v_i) \delta_{1i} (1 - \delta_{2i}) \log S'_{2t_1}(t_{1i}, t_{2i}) + \\ &+ \sum_{i=1}^n v_i (1 - \delta_{1i}) \delta_{2i} \log S'_{1t_2}(t_{1i}, t_{2i}) + \sum_{i=1}^n (1 - v_i) (1 - \delta_{1i}) \delta_{2i} \times \\ &\times \log S'_{2t_2}(t_{1i}, t_{2i}) + \sum_{i=1}^n v_i (1 - \delta_{1i}) (1 - \delta_{2i}) \log S_1(t_{1i}, t_{2i}) + \\ &\left. + \sum_{i=1}^n (1 - v_i) (1 - \delta_{1i}) (1 - \delta_{2i}) \log S_2(t_{1i}, t_{2i}) \right\}, \end{aligned} \quad (6.9)$$

em que $f_1(t_{1i}, t_{2i})$ e $f_2(t_{1i}, t_{2i})$ são dados por (6.1); $S_1(t_{1i}, t_{2i})$ e $S_2(t_{1i}, t_{2i})$ são dados por (6.2); $S'_{1t_1}(t_{1i}, t_{2i})$ e $S'_{2t_1}(t_{1i}, t_{2i})$ são dados por (6.5) e $S'_{1t_2}(t_{1i}, t_{2i})$ e $S'_{2t_2}(t_{1i}, t_{2i})$ são dados por (6.6).

Para uma análise Bayesiana da distribuição de Block & Basu na presença de observações censuradas, assumimos distribuições gama independentes para os parâmetros λ_k , ou seja,

$$\lambda_k \sim \text{Gama}(a_k, b_k), \tag{6.10}$$

para $k = 1, 2$ e 3 ; a_k e b_k são hiperparâmetros conhecidos; $\text{Gama}(a_k, b_k)$ denota uma distribuição gama com média a_k/b_k e variância a_k/b_k^2 .

Observe que estamos usando distribuições a priori gama (6.10) por ser os parâmetros λ_k positivos. Também podemos assumir hiperparâmetros desconhecidos em (6.10) e considerar uma análise Bayesiana hierárquica. Outras distribuições a priori podem ser usadas assumindo dependência entre os parâmetros, mas os resultados são similares, especialmente considerando grandes tamanhos de amostras.

A distribuição a posteriori conjunta para λ_1, λ_2 e λ_3 é dada por,

$$\pi(\lambda_1, \lambda_2, \lambda_3 | \mathbf{t}_1, \mathbf{t}_2) \propto \left\{ \prod_{k=1}^3 \lambda_k^{a_k-1} e^{-b_k \lambda_k} \right\} L(\lambda_1, \lambda_2, \lambda_3), \tag{6.11}$$

no qual $L(\lambda_1, \lambda_2, \lambda_3)$ é a função de verossimilhança e \mathbf{t}_1 e \mathbf{t}_2 são vetores das observações.

Para simular amostras para a distribuição a posteriori conjunta (6.11), usamos métodos MCMC. Neste caso, podemos simular amostras para a distribuição a posteriori conjunta (6.11) das distribuições condicionais $\pi(\lambda_1 | \lambda_2, \lambda_3, \mathbf{t}_1, \mathbf{t}_2)$, $\pi(\lambda_2 | \lambda_1, \lambda_3, \mathbf{t}_1, \mathbf{t}_2)$ e $\pi(\lambda_3 | \lambda_1, \lambda_2, \mathbf{t}_1, \mathbf{t}_2)$ usando o amostrador de Gibbs ou o algoritmo de Metropolis-Hastings.

6.2 Presença de Covariáveis e Observações Censuradas

Agora, vamos assumir que temos um vetor de covariáveis $\mathbf{x} = (x_1, x_2, \dots, x_p)'$ associado a cada tempo de vida bivariado (t_1, t_2) com uma distribuição exponencial de Block & Basu, com densidade (6.1). A presença de covariáveis para dados de sobrevivência bivariados tem sido discutidas por alguns autores (ver por exemplo, Hanagal, 2008; Huster *et al.*, 1989).

Na presença de um vetor de covariável \mathbf{x} , considere o seguinte modelo de regressão:

$$\lambda_{1i} = \alpha_1 \exp \left\{ \boldsymbol{\beta}'_1 \mathbf{x}_i \right\}, \tag{6.12}$$

$$\lambda_{2i} = \alpha_2 \exp \left\{ \boldsymbol{\beta}'_2 \mathbf{x}_i \right\},$$

com $\boldsymbol{\beta}'_j = (\beta_{j1}, \beta_{j2}, \dots, \beta_{jp})'$; $j = 1, 2$ sendo o vetor dos parâmetros da regressão e $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})$, $i = 1, 2, \dots, n$.

Também considerando a presença de observações censuradas, o logaritmo da função de verossimilhança para $\alpha_1, \alpha_2, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ e λ_3 é dado por (6.9), onde λ_1 e λ_2 são substituídos por λ_{1i} e λ_{2i} dados por (6.12).

Para uma análise Bayesiana do modelo, assumimos as seguintes distribuições a priori para $\alpha_1, \alpha_2, \beta_{1l}, \beta_{2l}$ e λ_3 :

$$\begin{aligned} \alpha_k &\sim \text{Gamma}(c_k, d_k), \\ \lambda_3 &\sim \text{Gamma}(e, f), \\ \beta_{kl} &\sim N(0, \sigma_{kl}^2), \end{aligned} \tag{6.13}$$

para $k = 1, 2$; $l = 1, 2, \dots, p$; c_k, d_k, e, f e σ_{kl}^2 são hiperparâmetros conhecidos e $N(0, \sigma_k^2)$ denota uma distribuição normal com média igual a zero e variância igual a σ_{kl}^2 . O uso de uma distribuição a priori normal para o parâmetro de regressão β_{kl} é justificável, já que este parâmetro é definido para todos os valores reais. Além disso assumimos independência a priori entre os parâmetros.

Amostras da distribuição a posteriori conjunta para os parâmetros do modelo são simuladas usando métodos MCMC. Como estaremos usando o *software WinBUGS* na geração dessas amostras, não apresentaremos as distribuições condicionais necessárias para o amostrador de Gibbs.

6.3 Exemplos

6.3.1 Tempos de Recorrência para a Infecção de rins de Pacientes

Como uma primeira aplicação da metodologia proposta, novamente vamos considerar o conjunto de dados de sobrevivência introduzido por McGilchrist & Aisbett (1991), relacionado a infecção de rins, onde há recorrência da infecção dos rins de 38 pacientes. Esses valores são registrados usando máquinas de diálise portáteis. As infecções podem ocorrer no local de inserção do cateter. O tempo registrado, chamado tempo de infecção, é o tempo de sobrevivência (em dias) do paciente até ocorrer uma infecção e o cateter ter sido removido, ou o tempo censurado, onde o cateter foi removido por outras razões. O cateter é reinscrito depois de algum tempo e o segundo tempo de infecção é novamente observado ou censurado (conjunto de dados na Tabela 2.1).

Para analisar os tempos de recorrência das infecções, vamos assumir agora a distribuição exponencial bivariada de Block & Basu com densidade (6.1).

Como uma primeira análise, não consideramos a presença da covariável sexo. Denotamos esse modelo como “modelo 1”.

Para uma análise Bayesiana do “modelo 1”, assumamos a distribuição a priori gama (6.10) para λ_1 , λ_2 e λ_3 , com valores dos hiperparâmetros $a_1 = a_2 = a_3 = 1$ e $b_1 = b_2 = b_3 = 100$, isto é, distribuições a priori aproximadamente não-informativas.

Usando o *software WinBUGS*, descartamos as primeiras 5000 amostras simuladas por Gibbs (“burn-in-samples”) para eliminar os efeitos dos valores iniciais para os parâmetros do modelo. Escolhendo cada 10^a amostra simulada por Gibbs, obtemos uma amostra final de tamanho 5000 para termos os resumos a posteriori de interesse. Convergência do algoritmo amostrador de Gibbs foi monitorada usando diferentes métodos, como os gráficos em séries temporais e o de Gelman-Rubin para as amostras simuladas.

Na Tabela 6.1, temos as médias a posteriori, os desvios padrões a posteriori e os intervalos de credibilidade de 95% para os parâmetros λ_1 , λ_2 e λ_3 . Também temos na Tabela 6.1, os resumos a posteriori para as médias, variâncias e desvios padrões para os tempos de sobrevivência T_1 e T_2 (ver (5.2)); e a covariância e o coeficiente de correlação (ver (5.3) e (5.4)) para os tempos de sobrevivência T_1 e T_2 .

É interessante observar que baseado nas observações não-censuradas, temos a média e o desvio-padrão amostrais para o 1º tempo (32 observações não-censuradas) e o 2º tempo

Tabela 6.1: Resumos a posteriori (“modelo 1”)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
λ_1	0,00561	0,00193	(0,0019; 0,0091)
λ_2	0,00532	0,00212	(0,0015; 0,0094)
λ_3	0,00314	0,00220	(0,0001; 0,0083)
μ_1	131,200	21,8801	(94,590; 129,00)
μ_2	137,500	25,1802	(95,260; 193,00)
σ_1	125,800	20,6204	(91,780; 172,00)
σ_2	130,400	22,7000	(92,980; 180,90)
σ_1^2	16240	5564	(8423; 29580)
σ_2^2	17520	6351	(8646; 32710)
Cov ₁₂	15,5400	7,45401	(1,0730; 16,810)
ρ_{12}	0,00101	0,00054	(0,0001; 0,0018)

(26 observações não-censuradas), dadas por $\bar{T}_1 = 121,80$; $\bar{T}_2 = 113,50$; $sd_{T_1} = 153,30$ e $sd_{T_2} = 131,40$. Comparando esses resultados com os da Tabela 6.1 (Monte Carlo estimado para as médias a posteriori e intervalo de credibilidade de 95% para μ_1, μ_2, σ_1 e σ_2), temos uma indicação de bom ajuste para a distribuição de Block & Basu com densidade dada por (6.1) para os tempos de recorrência de infecção dos rins dos pacientes apresentados na Tabela 2.1.

Nas Figuras 6.1 e 6.2, temos os gráficos em séries temporais e o de Gelman-Rubin para as amostras simuladas por Gibbs. Através desses gráficos, observamos a convergência do algoritmo em todos os casos.

Com finalidade comparativa, assumir agora T_1 e T_2 variáveis aleatórias exponenciais independentes, com parâmetros λ_1 e λ_2 , respectivamente, e distribuições a priori Gama(1,100) para λ_1 e λ_2 , temos na Tabela 6.2, os resumos a posteriori para $\lambda_1, \lambda_2, \mu_1 = 1/\lambda_1$ e $\mu_2 = 1/\lambda_2$ (μ_1 e μ_2 são as médias para T_1 e T_2) baseadas nas amostras simuladas por Gibbs de tamanho 5000. Observamos resultados mais precisos, especialmente para os parâmetros μ_1 e μ_2 (menores intervalos de credibilidade) considerando a distribuição exponencial bivariada de Block & Basu (veja Tabela 6.1)

Tabela 6.2: Resumos a posteriori (T_1 e T_2 independentes)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
λ_1	0,00764	0,00130	(0,0053; 0,0103)
λ_2	0,00791	0,00150	(0,0053; 0,0111)
μ_1	135,500	24,0802	(96,77; 189,20)
μ_2	130,701	25,7504	(90,20; 189,80)

Observe que usando métodos de inferência clássica, baseados na normalidade assintótica dos estimadores de máxima verossimilhança para λ_1, λ_2 e λ_3 , podemos ter grandes dificuldades computacionais e resultados das inferências não muito exatos para os

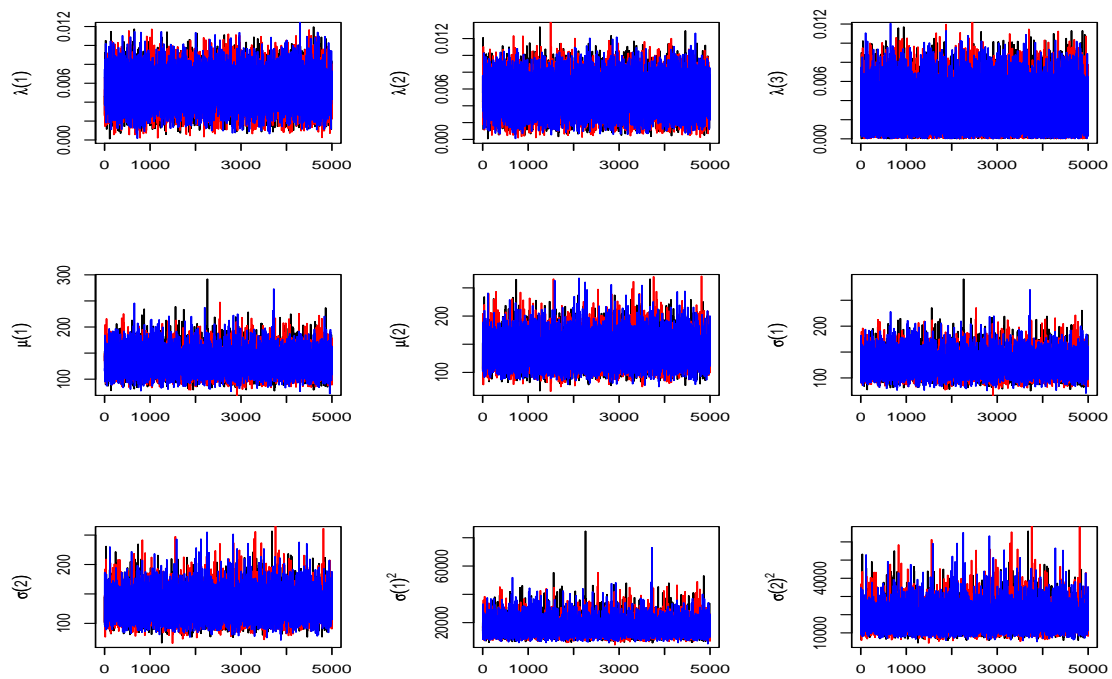


Figura 6.1: Amostras simuladas por Gibbs (“modelo 1”)

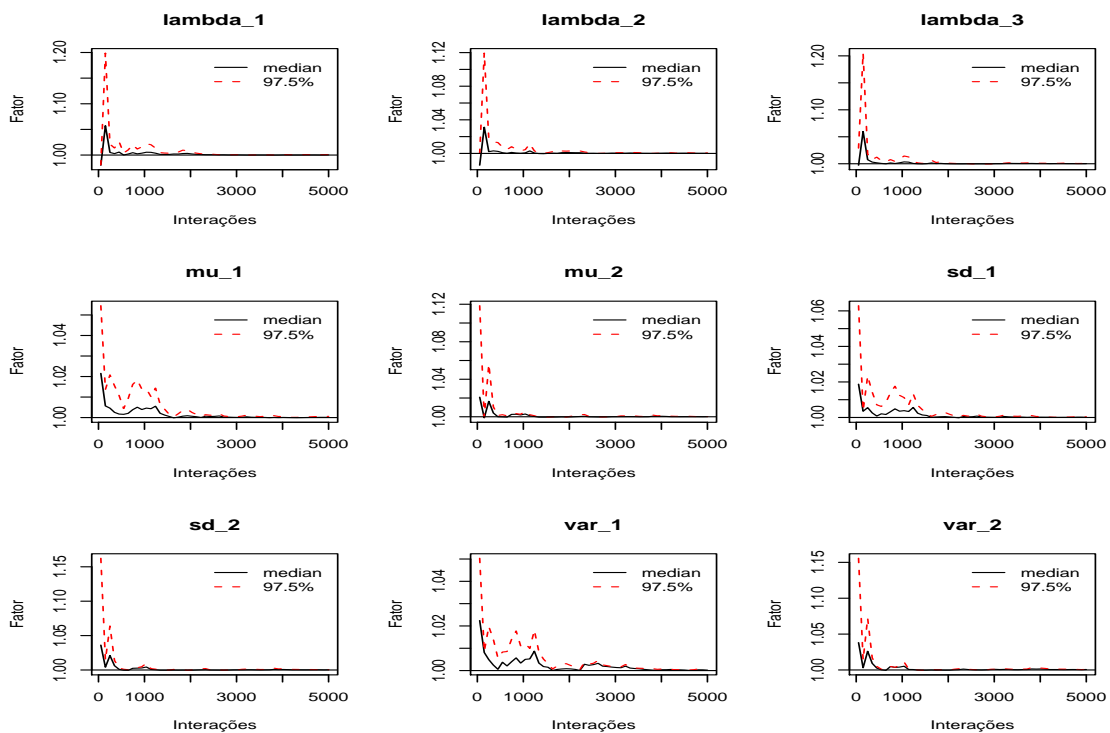


Figura 6.2: Valores de Gelman e Rubin (“modelo 1”)

parâmetros da distribuição exponencial bivariada de Block & Basu.

Na presença da covariável sexo, denotado por x , assumimos o modelo de regressão introduzido na Seção 6.2 para os parâmetros λ_1 e λ_2 , ou seja,

$$\begin{aligned} \lambda_{1i} &= \alpha_1 \exp \{ \beta_1 x_i \}, \\ \lambda_{2i} &= \alpha_2 \exp \{ \beta_2 x_i \}, \end{aligned} \tag{6.14}$$

com $i = 1, 2, \dots, 38$; $x_i = 2$ (masculino) e $x_i = 1$ (feminino). Denotamos esse modelo como “modelo 2”.

Assumindo as distribuições a priori (6.13) com os valores dos hiperparâmetros $c_1 = c_2 = c_3 = 1$; $d_1 = d_2 = d_3 = 100$; $e = f = 1$; $\sigma_{11}^2 = \sigma_{21}^2 = 1$ (aqui $l = 1$, uma vez que temos somente uma covariável) e usando o *software WinBUGS* e também considerando um “*burn-in-sample*” de tamanho 5000 e tomando a cada 10^a amostra, temos na Tabela 6.3, os resumos a posteriori de interesse.

Tabela 6.3: Resumos a posteriori (“modelo 2”)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
α_1	0,00434	0,00150	(0,0016; 0,0074)
α_2	0,00521	0,00192	(0,0019; 0,0092)
β_1	1,63400	0,46104	(0,7324; 2,5460)
β_2	0,09363	0,54662	(-1,1240; 1,0720)
λ_3	0,00302	0,00203	(0,0001; 0,0073)

Nas Figuras 6.3 e 6.4, temos os gráficos de séries temporais e o de Gelman-Rubin para as amostras simuladas por Gibbs. Observamos a convergência para todos os parâmetros do “modelo 2”.

Dos resultados da Tabela 6.3, observamos que o zero está incluído no intervalo de credibilidade de 95% para β_2 , mas não para β_1 , ou seja, o segundo tempo T_2 não é afetado pela covariável sexo, mas o primeiro tempo T_1 é afetado pelo sexo.

Como uma comparação para os dois modelos propostos, podemos usar um critério de discriminação Bayesiano, como por exemplo o DIC (*Deviance Information Criterion*) introduzido por Spiegelhalter *et al.* (2002) e é dado automaticamente pelo *software WinBUGS*. Assumindo o “modelo 1”, temos um estimador de Monte Carlo para o DIC dado pelo valor 687,118; assumindo o “modelo 2” na presença da covariável sexo, o valor de DIC é dado por 674,961. Ou seja, o “modelo 2” se ajusta melhor ao conjunto de dados da Tabela 2.1 (menor valor para o DIC).

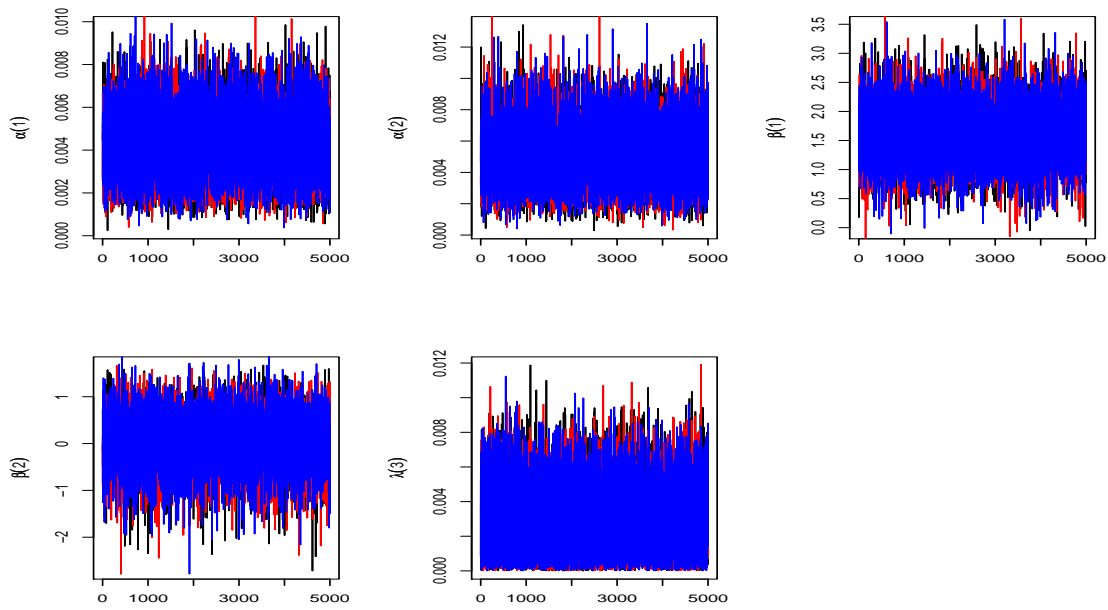


Figura 6.3: Amostras simuladas por Gibbs (“modelo 2”)

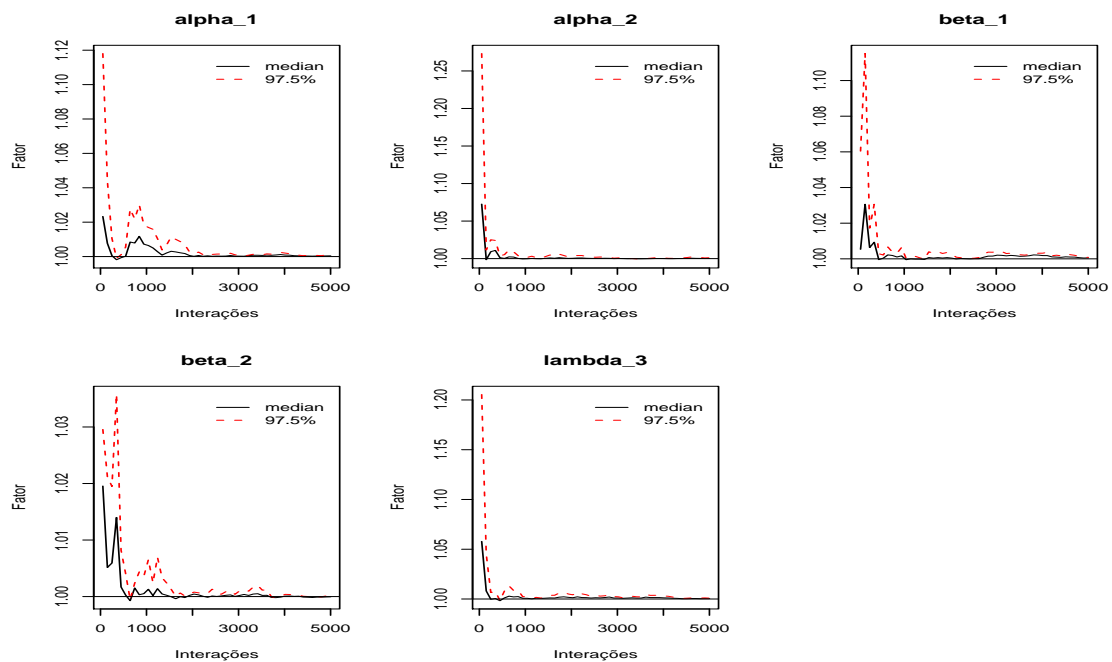


Figura 6.4: Valores de Gelman e Rubin (“modelo 2”)

6.3.2 Transplante de Medula Óssea para Leucemia

Transplantes de medula óssea são um tratamento padrão para leucemia aguda. Prognósticos para a recuperação podem depender de fatores de risco conhecidos no tempo

do transplante, tais como paciente e/ou idade e sexo do doador, o estágio inicial da doença, o tempo de prognóstico para transplante, etc. O prognóstico final pode mudar quando o desenvolvimento histórico pós-transplante leva a ocorrência de eventos, em tempos aleatórios, durante o processo de recuperação. Tais como, o desenvolvimento de doença aguda ou crônica transplantar-versus-hospedar (GVHD), retorno da contagem de plaquetas para níveis normais ou o desenvolvimento de infecções.

O transplante pode ser considerado um insucesso quando a leucemia retorna ao paciente (recidiva) ou quando ele/ela morre durante a remissão (tratamento leva a morte).

Neste estudo, 137 pacientes com leucemia mielóide aguda (LMA) e leucemia linfoblástica aguda (também conhecida como leucemia linfóide aguda (ou LLA)) receberam uma combinação de 16 mg/kg de Busulfan oral (BU) e 120 mg/kg de cyclophosphamide (Cy) (99 pacientes com LMA e 38 com LLA) e foram tratados em um dos quatro hospitais: 76 no da Universidade estadual de Ohio (OSU) em Columbus; 21 na Universidade Hahnemann (HU) na Philadelphia; 23 no Hospital St. Vincent's (SVH) em Sidney, Australia e 17 no Hospital Alfred (AH) em Melbourne, Australia (conjunto de dados introduzidos por Klein & Moeschberger, 1997, página 464).

Na análise considerada neste trabalho, assumimos como tempos de sobrevivência, os tempos (em dias) para a doença aguda transplantar-versus-hospedar (TA) com 111 observações censuradas e 26 não-censuradas e o tempo (em dias) da doença crônica transplantar-versus-hospedar (TC) com 76 observações censuradas e 61 não-censuradas e as seguintes covariáveis: idade dos pacientes em anos, idade do doador em anos, sexo do paciente, sexo do doador, CMV (cytomegalovirus immune status) do paciente, CMV do doador, tempo de espera do transplante em dias e os diferentes hospitais.

Como uma primeira análise, assumimos a distribuição exponencial bivariada de Block & Basu com densidade (6.1) na presença de dados censurados, mas não considerando a presença de covariáveis (“modelo 1”). Neste caso assumimos a distribuição a priori gama (6.10) para λ_k , $k = 1, 2, 3$, com hiperparâmetros $a_1 = a_2 = a_3 = 1$; $b_1 = b_2 = 100$ e $b_3 = 1000$. Usando o *software WinBUGS* com um “burn-in-sample” de tamanho 5000, simulamos uma amostra final de tamanho 5000, escolhendo a cada 50ª amostra simulada por Gibbs.

Convergência para o algoritmo amostrador de Gibbs foi verificado por gráficos de séries temporais e de Gelman & Rubin das amostras simuladas por Gibbs (veja Figura 6.5).

Na Tabela 6.4 temos os resumos a posteriori de interesse assumindo o “modelo 1”.

Nas Figuras 6.5 e 6.6, temos os gráficos de séries temporais e o de Gelman-Rubin para as amostras simuladas por Gibbs. Observamos a convergência para todos os parâmetros do “modelo 1”.

Assumindo o “modelo 1”, obtemos um estimador de Monte Carlo para o DIC baseado

Tabela 6.4: Resumos a posteriori (“modelo 1”)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
λ_1	0,00027	0,00005	(0,00018; 0,00038)
λ_2	0,00094	0,00013	(0,00070; 0,00120)
λ_3	0,00003	0,00003	(0,0000009; 0,000127)
μ_1	3500,00	689,200	(2405,00; 5098,00)
μ_2	1054,00	137,700	(821,500; 1361,00)
σ_1	3448,00	694,700	(2346,00; 5057,00)
σ_2	1053,00	136,900	(821,300; 1359,00)
Cov12	70,8200	60,6100	(2,18800; 226,000)
ρ_{12}	0,00002	0,00002	(0,0000006; 0,0000728)

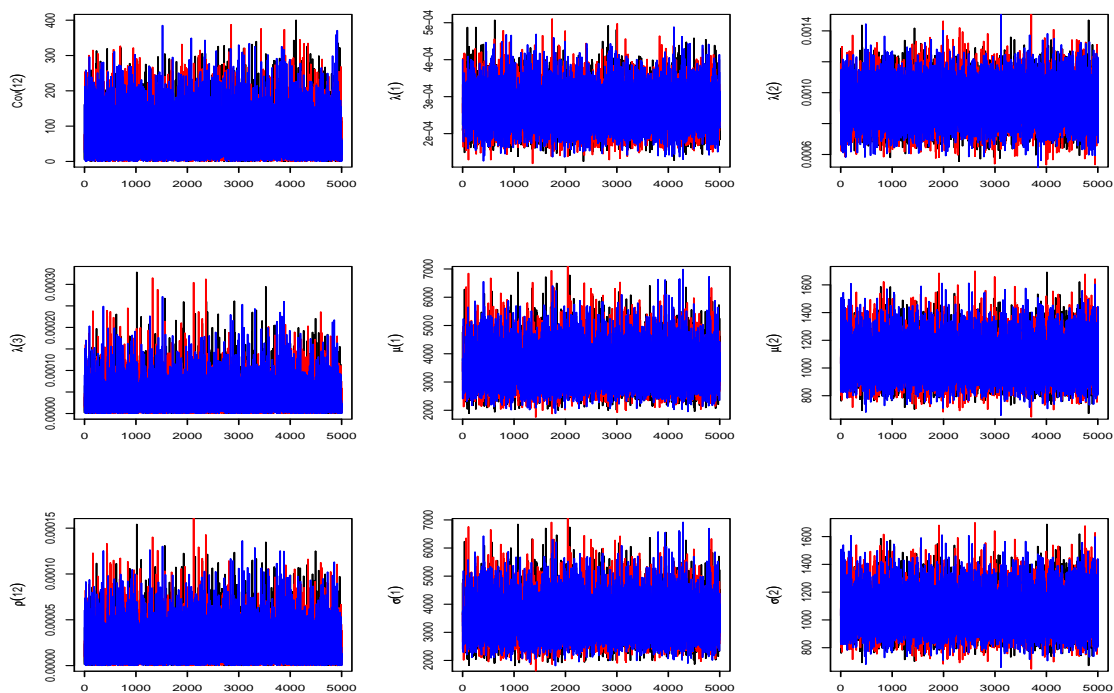


Figura 6.5: Amostras simuladas por Gibbs (“modelo 1”)

nas 5000 amostras simuladas por Gibbs dado por $DIC = 1456,87$.

Para verificar o efeito de outras distribuições a priori nos resumos a posteriori, também assumimos no “modelo 1” a reparametrização $\lambda_k = \exp(\gamma_k)$, $k = 1, 2$ e 3 , ou seja, $\gamma_k = \log(\lambda_k)$ e distribuições a priori uniforme aproximadamente não-informativas para γ_k no intervalo $(-20; -5)$. Dos resumos a posteriori obtidos (veja Tabela 6.5), observamos resultados de inferências similares quando consideramos distribuições a priori gama para λ_1, λ_2 e λ_3 (veja Tabela 6.4). Neste caso o DIC obtido é dado por $DIC = 1455,37$.

Na presença de dados censurados e covariáveis, assumimos o seguinte modelo de regressão (“modelo 2”):

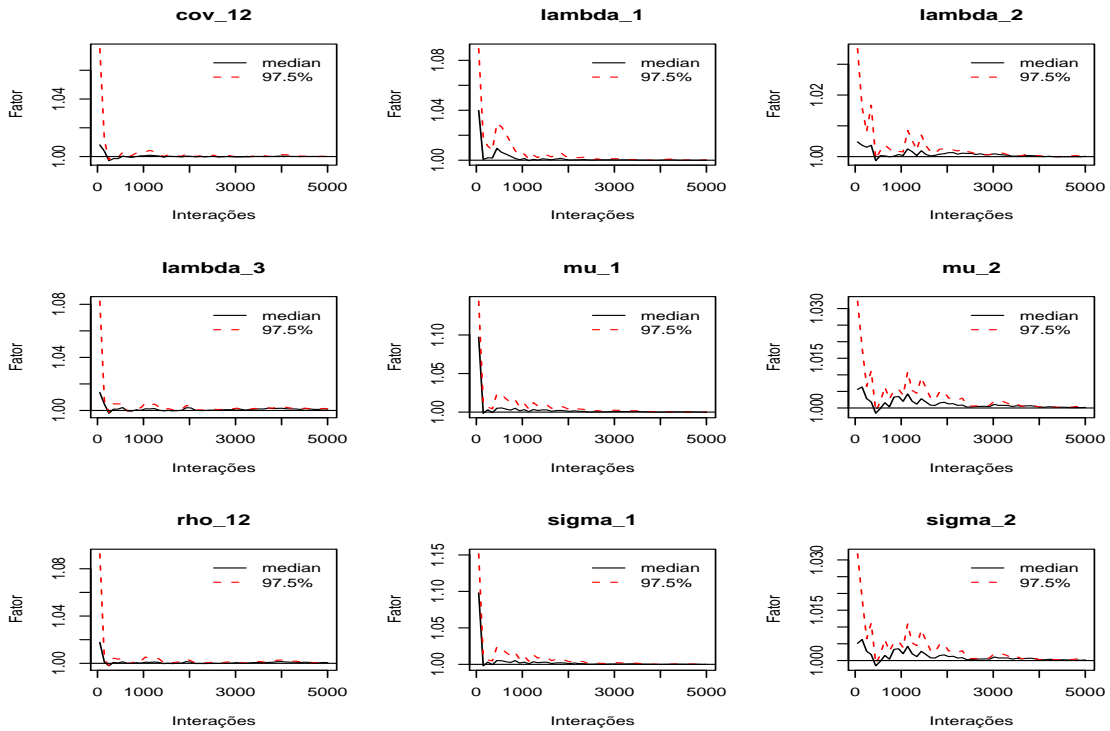


Figura 6.6: Valores de Gelman e Rubin (“modelo 1”)

Tabela 6.5: Resumos a posteriori (“modelo 1” com prioris uniformes para $\gamma_k, k = 1, 2, 3$)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
λ_1	0,000264	0,000052	(0,00017; 0,00037)
λ_2	0,000949	0,000122	(0,00072; 0,00121)
λ_3	0,000004	0,000012	(0,000000003; 0,000035)
μ_1	3898,000	797,0000	(2631,00; 5715,00)
μ_2	1068,000	139,2000	(827,500; 1373,00)
σ_1	3891,000	797,8000	(2621,00; 5710,00)
σ_2	1067,000	139,1000	(827,400; 1372,00)
Cov12	9,460000	25,00000	(0,00661; 84,3100)
ρ_{12}	0,0000025	0,000007	(0,000000001; 0,0000225)

$$\lambda_{1i} = \alpha_1 \exp \{ \beta_{11} (Z_{1i} - \bar{Z}_1) + \beta_{12} (Z_{2i} - \bar{Z}_2) + \beta_{13} Z_{3i} + \beta_{14} Z_{4i} + \beta_{15} Z_{5i} + \beta_{16} Z_{6i} + \beta_{17} (Z_{7i} - \bar{Z}_7) + \beta_{18} Z_{8i} + \beta_{19} Z_{9i} + \beta_{1,10} Z_{10i} \}, \quad (6.15)$$

e

$$\begin{aligned} \lambda_{2i} = & \alpha_2 \exp \{ \beta_{21} (Z_{1i} - \bar{Z}_1) + \beta_{22} (Z_{2i} - \bar{Z}_2) + \beta_{23} Z_{3i} + \beta_{24} Z_{4i} + \beta_{25} Z_{5i} + \\ & + \beta_{26} Z_{6i} + \beta_{27} (Z_{7i} - \bar{Z}_7) + \beta_{28} Z_{8i} + \beta_{29} Z_{9i} + \beta_{2,10} Z_{10i} \}, \end{aligned} \quad (6.16)$$

em que Z_{1i} é a idade do paciente; Z_{2i} é a idade do doador; Z_{3i} é o sexo do paciente; Z_{4i} é o sexo do doador; Z_{5i} é o CMV do paciente; Z_{6i} é o CMV do doador; Z_{7i} é o tempo de espera para o transplante, em dias; Z_{8i} , Z_{9i} e $Z_{10,i}$ são variáveis “dummy” relacionadas aos diferentes hospitais: ($Z_{8i} = 1$ para o OSU; 0 para os outros hospitais; $Z_{9i} = 1$ para o HU; 0 para os outros hospitais; $Z_{10,i} = 1$ para o SVH; 0 para os outros hospitais); \bar{Z}_1 , \bar{Z}_2 e \bar{Z}_7 são médias amostrais para as covariáveis Z_{1i} , Z_{2i} e Z_{7i} , respectivamente, $i = 1, 2, \dots, 137$.

Para uma análise Bayesiana considerando a distribuição exponencial bivariada de Block & Basu, assumimos as seguintes distribuições a priori (veja 6.13): $u_k \sim U(-20, 0)$ onde $\alpha_k = \log(u_k)$, para $k = 1, 2$; $\lambda_3 = \log(u_3)$, com $u_3 \sim U(-20, 0)$ e distribuições a priori uniformes $U(-a_{kl}, a_{kl})$ para os β_{kl} onde a_{jl} são hiperparâmetros com valores pequenos, $k = 1, 2$; $l = 1, 2, \dots, 10$. Também usamos distribuições a priori normais para os parâmetros da regressão β_{kl} , mas a convergência do algoritmo amostrador de Gibbs usando o *software WinBUGS* somente foi obtida tomando variâncias muito pequenas para as distribuições a priori normais com médias iguais a zero. A escolha dos valores dos hiperparâmetros foi feita usando informação a priori, como a informação prévia da análise Bayesiana não considerando a presença do covariáveis (uso de métodos Bayesianos empíricos).

Na Tabela 6.6, temos os resumos a posteriori obtidos usando o *software WinBUGS* com um “burn-in-sample” de tamanho 5000 e uma amostra final simulada por Gibbs de tamanho 5000, obtidas escolhendo a cada 50ª amostra simulada. O DIC obtido automaticamente do *WinBUGS* é dado por $DIC = 1443, 17$.

Já que o valor de DIC para este modelo é menor que o valor de DIC do modelo não incluindo a presença de covariáveis, temos uma indicação que o “modelo 2” pode ser melhor para este conjunto de dados.

Dos resultados da Tabela 6.6, observamos que o valor zero está incluído nos intervalos de credibilidade de 80% para muitos parâmetros da regressão β_{kl} , $k = 1, 2$ and $l = 1, 2, \dots, 10$; uma indicação de que a correspondente covariável não tem um efeito significativo nos tempos de sobrevivência TA e TC. Para alguns casos, como os parâmetros de regressão β_{11} , β_{12} e β_{22} observamos que o zero não esta incluído no intervalo de credibilidade. Isto é um indicativo de que a covariável Z_1 (idade do paciente) e Z_2 (idade do doador) tem efeito significativo na doença aguda (TA) e doença crônica (TC), respectivamente.

Tabela 6.6: Resumos a posteriori (“modelo 2”)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 80%
α_1	-0,92433	0,58021	(-0,94146; -0,90671)
α_2	-0,83696	0,74208	(-0,85144; -0,82217)
β_{11}	0,04239	0,03089	(0,00266; 0,08168)
β_{12}	0,04167	0,03063	(0,00275; 0,08083)
β_{13}	-0,01707	0,11360	(-0,16600; 0,14970)
β_{14}	-0,02402	0,11170	(-0,16840; 0,13950)
β_{15}	-0,00297	0,11310	(-0,15880; 0,15370)
β_{16}	0,05314	0,10610	(-0,10970; 0,17810)
β_{17}	0,00029	0,00050	(-0,00035; 0,00090)
β_{18}	-0,00644	0,11410	(-0,16280; 0,15330)
β_{19}	0,02379	0,11330	(-0,14300; 0,17050)
$\beta_{1,10}$	-0,08430	0,09216	(-0,18460; 0,05396)
β_{21}	-0,02651	0,02465	(-0,05838; 0,00488)
β_{22}	0,05332	0,02375	(0,02268; 0,08398)
β_{23}	-0,05284	0,10370	(-0,17730; 0,10290)
β_{24}	-0,06561	0,10050	(-0,18130; 0,08773)
β_{25}	0,08135	0,09398	(-0,06053; 0,18430)
β_{26}	-0,04558	0,10560	(-0,17500; 0,1125)
β_{27}	0,00059	0,00034	(0,00014; 0,00102)
β_{28}	0,03651	0,10840	(-0,12590; 0,17240)
β_{29}	0,04243	0,11020	(-0,12490; 0,17600)
$\beta_{2,10}$	-0,00018	0,11530	(-0,16070; 0,15890)
λ_3	-1,25164	0,62284	(-1,37218; -1,08207)

Também observamos alguma indicação de efeito para a covariável Z_7 (tempo de espera para o transplante) para o tempo da doença crônica TC (chronic graft-versus-host disease) já que o zero não está incluído no intervalo de credibilidade de 80% para β_{27} (0,00014; 0,00102).

A partir dos resultados deste capítulo, observamos que o uso de métodos Bayesianos para analisar dados de sobrevivência bivariados assumindo uma distribuição exponencial bivariada de Block & Basu na presença de dados censurados e covariáveis, pode ser uma boa alternativa para a análise de dados médicos de sobrevivência bivariados.

Os resultados desse capítulo generalizam os resultados obtidos por Achcar & Leandro (1998), quando temos censuras e covariáveis. O uso do *software WinBUGS* pode ser de grande utilidade, pois não requer muito esforço computacional na geração de amostras da distribuição a posteriori de interesse.

Observar que apesar desse modelo ter sido introduzido em 1974 (Block & Basu, 1974), ele não foi muito utilizado na análise de dados médicos, especialmente antes do advento das novas técnicas computacionais, dada as grandes dificuldades na obtenção das inferências clássicas precisas, mesmo na presença de amostras com tamanho amostral grandes, e na

presença de censuras e covariáveis.

Capítulo 7

Modelos Bivariados de Sobrevivência Derivados de Funções Cópulas

Como visto em capítulos anteriores, para analisar tempos de sobrevivência bivariados, podemos assumir diferentes distribuições paramétricas introduzidas na literatura (veja por exemplo, Freund, 1961; Marshall & Olkin, 1967a; Sarkar, 1987; Downton, 1970; Gumbel, 1960; Hawkes, 1972; Hougaard, 1986b; Arnold & Strauss, 1988; Block & Basu, 1974).

Assumindo somente um tempo de sobrevivência, uma distribuição popular é dada pela família Weibull (veja por exemplo, Lawless, 1982), já que este modelo apresenta uma grande flexibilidade: função de risco constante, crescente ou decrescente $h(t) = f(t)/S(t)$, onde $f(t)$ é a função densidade de probabilidade para o tempo de vida T e $S(t)$ é a função de sobrevivência, ou seja, $S(t) = P(T > t)$ para um valor fixado t .

Diferentes formulações da distribuição Weibull para tempos de sobrevivência bivariados, podem ser consideradas usando funções de cópulas (veja por exemplo, Nelsen, 1999; Trivedi & Zimmer, 2005a,b).

Inferências para esses diferentes modelos Weibull bivariados podem apresentar grandes dificuldades usando métodos de inferência clássica baseados nos métodos de máxima verossimilhança, especialmente na presença de dados censurados e covariáveis, uma situação comum em aplicações médicas.

Neste caso, consideramos o uso de métodos Bayesianos onde as amostras para a distribuição a posteriori conjunta de interesse são simuladas usando métodos MCMC.

7.1 Funções Cópulas

A seguir, apresentamos alguns conceitos básicos sobre funções cópulas.

Definição 1 *Uma função cópula é uma função que interliga marginais univariadas com sua distribuição multivariada completa. Para m variáveis aleatórias uniformes U_1, U_2, \dots, U_m a distribuição conjunta C é definida por,*

$$C(u_1, u_2, \dots, u_m, \rho) = Pr(U_1 \leq u_1, U_2 \leq u_2, \dots, U_m \leq u_m). \quad (7.1)$$

Funções cópulas podem ser usadas para interligar marginais com uma distribuição conjunta. Para dadas funções distribuição marginais univariadas $F_1(t_1), F_2(t_2), \dots, F_m(t_m)$, a função,

$$C(F_1(t_1), F_2(t_2), \dots, F_m(t_m)) = F(t_1, t_2, \dots, t_m), \quad (7.2)$$

que é definida usando uma função cópula C , resulta em uma função distribuição multivariada, com distribuições marginais univariadas com especificado $F_1(t_1), F_2(t_2), \dots, F_m(t_m)$.

Demonstração:

$$\begin{aligned} C(F_1(t_1), F_2(t_2), \dots, F_m(t_m), \rho) &= Pr[U_1 \leq F_1(t_1), U_2 \leq F_2(t_2), \dots, U_m \leq F_m(t_m)] \\ &= Pr[F_1^{-1}(u_1) \leq t_1, F_2^{-1}(u_2) \leq t_2, \dots, F_m^{-1}(u_m) \leq t_m] \\ &= Pr[T_1 \leq t_1, T_2 \leq t_2, \dots, T_m \leq t_m] \\ &= F(t_1, t_2, \dots, t_m). \end{aligned} \quad (7.3)$$

É importante chamar a atenção que cada função distribuição multivariada F pode ser escrita na forma de uma função cópula (Sklar, 1959); ou seja, se $F(t_1, t_2, \dots, t_m)$ é uma função distribuição multivariada conjunta com funções distribuições marginais univariadas $F_1(t_1), F_2(t_2), \dots, F_m(t_m)$, então existe uma função cópula $C(u_1, u_2, \dots, u_m)$, tal que,

$$F(t_1, t_2, \dots, t_m) = C(F_1(t_1), F_2(t_2), \dots, F_m(t_m)). \quad (7.4)$$

Se cada F_i é contínua, então C é única.

Para o caso especial de distribuições bivariadas, temos $m = 2$.

A aproximação para formular uma distribuição multivariada usando uma cópula é baseada na idéia que uma simples transformação pode ser feita de cada variável marginal,

onde cada variável marginal transformada tem uma distribuição uniforme. Uma vez feito isso, a estrutura de dependência pode ser expressa como uma distribuição multivariada sobre as uniformes obtidas, e uma cópula é precisamente uma distribuição multivariada a partir de variáveis aleatórias marginalmente uniformes.

Neste caso, há muitas famílias de cópulas que diferem no detalhe da dependência que elas representam.

No caso bivariado, temos T_1 e T_2 duas variáveis aleatórias com funções distribuições contínuas F_1 and F_2 .

A transformação de probabilidade integral pode ser aplicada separadamente para as duas variáveis aleatórias definidas como $U = F_1(t_1)$ and $V = F_2(t_2)$, onde U e V tem distribuições uniformes $(0, 1)$, mas são usualmente dependentes se T_1 e T_2 são dependentes (T_1 e T_2 independentes, implicam que U e V são independentes). Especificar a dependência entre T_1 e T_2 é o mesmo que especificar a dependência entre U e V .

Com as variáveis aleatórias uniformes U e V , o problema reduz a especificar uma distribuição bivariada entre duas uniformes, ou seja, uma cópula.

As funções cópulas satisfazem as seguintes condições: (1) O domínio de C é $F_1(t_1) \times F_2(t_2)$, onde $F_1(t_1)$ e $F_2(t_2)$ são subconjuntos de $I = [0, 1]$. (2) Uma cópula é reduzida: $C[F_1(t_1), F_2(t_2)] = 0$ se $F_i(t_i) = 0$ para $i = 1, 2$; e bi-crescente: $d^2C/dt_1t_2 \geq 0$ se a segunda derivada existe. (3) $C[F(t_i), 1] = F(t_i)$ para $i = 1, 2$.

Para todas as cópulas temos os limites de Fréchet-Hoeffding, dado por,

$$\max(0, u + v - 1) \leq C(u, v) \leq \min(u, v). \quad (7.5)$$

Diferentes famílias de cópulas são introduzidas na literatura. Uma delas é a cópula Gaussiana, que é construída da distribuição normal bivariada e é dada por,

$$C_\rho(u, v) = \Phi_\rho[\Phi^{-1}(u), \Phi^{-1}(v)], \quad (7.6)$$

em que Φ_ρ é a função distribuição bivariada de uma distribuição normal padrão com coeficiente de correlação ρ , ou seja,

$$\begin{aligned} \Phi_\rho(x, y) &= P(X \leq x, Y \leq y) = \\ &= \int_{-\infty}^x \int_{-\infty}^y \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)} [z^2 + w^2 - 2\rho zw]\right\} dzdw. \end{aligned} \quad (7.7)$$

7.2 Algumas Distribuições Bivariadas Weibull Derivadas de Funções Cópulas

Flores (2009) usa marginais Weibull para construir seis distribuições bivariadas Weibull apropriadas a análise de sobrevivência e testa suas performances através de simulações. Mais especificamente, esse trabalho compara e testa as performances das cópulas Farlie-Gumbel-Morgenstern, Ali-Mikhail-Haq, Gumbel-Hougaard, Gumbel-Barnett e Clayton com as geradas da distribuição bivariada Weibull.

Nesta tese consideremos uma primeira cópula (ver Flores, 2008), dada por,

$$C_1(u, v) = u + v - 1 + (1 - u)(1 - v) \exp \{-\theta \ln(1 - u) \ln(1 - v)\}, \quad (7.8)$$

com $u = F_1(t_1)$ e $v = F_2(t_2)$. Os tempos t_1 e t_2 pertencentes a reta real.

Também assumimos as distribuições Weibull marginais dadas por,

$$u = F_1(t_1) = 1 - \exp \left\{ - \left(\frac{t_1}{\lambda_1} \right)^{p_1} \right\}, \quad (7.9)$$

e

$$v = F_2(t_2) = 1 - \exp \left\{ - \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right\},$$

em que p_1 e p_2 são parâmetros de forma das distribuições de Weibull.

Assim,

$$\begin{aligned} H_1(t_1, t_2) &= C_1(F_1(t_1), F_2(t_2)) = F_1(t_1, t_2 | \lambda_1, \lambda_2, p_1, p_2, \theta) = \\ &= F_1(t_1) + F_2(t_2) - 1 + (1 - F_1(t_1))(1 - F_2(t_2)) \times \\ &\times \exp \{-\theta \ln(1 - F_1(t_1)) \ln(1 - F_2(t_2))\}. \end{aligned} \quad (7.10)$$

De (7.9) temos,

$$\begin{aligned} F_1(t_1, t_2 | \lambda_1, \lambda_2, p_1, p_2, \theta) &= 1 - \exp \left\{ - \left(\frac{t_1}{\lambda_1} \right)^{p_1} \right\} - \exp \left\{ - \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right\} + \\ &+ \exp \left\{ - \left(\frac{t_1}{\lambda_1} \right)^{p_1} - \left(\frac{t_2}{\lambda_2} \right)^{p_2} - \theta \left(\frac{t_1}{\lambda_1} \right)^{p_1} \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right\}, \end{aligned} \quad (7.11)$$

para $t_1 > 0$ e $t_2 > 0$.

Observe que se $\theta = 0$, temos variáveis aleatórias independentes. Também observe que a função de sobrevivência bivariada conjunta é dada por,

$$S_1(t_1, t_2) = 1 - F_1(t_1) - F_2(t_2) + F_1(t_1, t_2), \quad (7.12)$$

em que $F_1(t_1)$ e $F_2(t_2)$ são dados por (7.9) e $F_1(t_1, t_2)$ é dado por (7.11).

Ou seja,

$$S_1(t_1, t_2) = \exp \left\{ - \left(\frac{t_1}{\lambda_1} \right)^{p_1} - \left(\frac{t_2}{\lambda_2} \right)^{p_2} - \theta \left(\frac{t_1}{\lambda_1} \right)^{p_1} \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right\}. \quad (7.13)$$

Denotemos o modelo com a função de distribuição conjunta (7.11) e a de sobrevivência conjunta (7.13) como “modelo 1”.

Uma segunda função cópula assumida é a cópula de FGM (veja por exemplo, Farlie, 1960; Morgenstern, 1956) dada por,

$$C_2(u, v) = [1 - e^{\ln(1-u)}] [1 - e^{\ln(1-v)}] \{1 + \theta \exp [\ln(1-u) + \ln(1-v)]\}. \quad (7.14)$$

Ou seja,

$$C_2(u, v) = uv [1 + \theta(1-u)(1-v)], \quad (7.15)$$

com $-1 \leq \theta \leq 1$. O parâmetro θ estima a interdependência entre os tempos t_1 e t_2 .

De (7.9), temos,

$$\begin{aligned} F_2(t_1, t_2 | \lambda_1, \lambda_2, p_1, p_2, \theta) &= \left\{ 1 - \exp \left[- \left(\frac{t_1}{\lambda_1} \right)^{p_1} \right] \right\} \left\{ 1 - \exp \left[- \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right] \right\} \times \\ &\times \left\{ 1 + \theta \exp \left[- \left(\frac{t_1}{\lambda_1} \right)^{p_1} - \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right] \right\}. \end{aligned} \quad (7.16)$$

De (7.12) e (7.16), a função de sobrevivência conjunta é dada por,

$$\begin{aligned} S_2(t_1, t_2) &= \exp \left\{ - \left(\frac{t_1}{\lambda_1} \right)^{p_1} - \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right\} \left\{ 1 + \theta \left[1 - \exp \left(- \left(\frac{t_1}{\lambda_1} \right)^{p_1} \right) \right] \right\} \times \\ &\times \left[1 - \exp \left(- \left(\frac{t_2}{\lambda_2} \right)^{p_2} \right) \right] \}. \end{aligned} \quad (7.17)$$

Denotemos o modelo definido por (7.16) e (7.17) como “modelo 2”.

Outras funções cópulas também podem ser usadas; um caso especial é dado pela cópula Ali-Mikhail-Haq (veja por exemplo, Trivedi & Zimmer, 2005a), dada por,

$$C_3(u, v) = \frac{uv}{1 - \theta(1-u)(1-v)}, \quad (7.18)$$

para $-1 \leq \theta \leq 1$.

De (7.9) e (7.18), temos uma terceira distribuição bivariada Weibull com função de distribuição dada por,

$$F_3(t_1, t_2 | \lambda_1, \lambda_2, p_1, p_2, \theta) = \frac{\left\{1 - \exp\left[-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right]\right\} \left\{1 - \exp\left[-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right]\right\}}{\left\{1 - \theta \exp\left[-\left(\frac{t_1}{\lambda_1}\right)^{p_1} - \left(\frac{t_2}{\lambda_2}\right)^{p_2}\right]\right\}}. \quad (7.19)$$

A função de sobrevivência conjunta obtida de (7.12) é dada por,

$$S_3(t_1, t_2) = \exp\left[-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right] + \exp\left[-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right] - 1 + \frac{\left\{1 - \exp\left[-\left(\frac{t_1}{\lambda_1}\right)^{p_1}\right]\right\} \left\{1 - \exp\left[-\left(\frac{t_2}{\lambda_2}\right)^{p_2}\right]\right\}}{\left\{1 - \theta \exp\left[-\left(\frac{t_1}{\lambda_1}\right)^{p_1} - \left(\frac{t_2}{\lambda_2}\right)^{p_2}\right]\right\}}. \quad (7.20)$$

7.3 Uma Análise Bayesiana para o “Modelo 1” e “Modelo 2” na Presença de Dados Censurados

Suponha, como no capítulo 6, que ambos T_1 ou T_2 podem ser censurados e que a censura é independente dos tempos de sobrevivência. Vamos sub-dividir as n observações em quatro classes:

C_1 : ambos t_{1i} e t_{2i} são tempos de sobrevivência observados;

C_2 : t_{1i} é um tempo de sobrevivência e t_{2i} é um tempo de censura (ou seja, sabemos somente que $T_{2i} \geq t_{2i}$);

C_3 : t_{1i} é um tempo de censura e t_{2i} é um de sobrevivência;

C_4 : ambos t_{1i} e t_{2i} são tempos censurados.

A função de verossimilhança para um modelo contínuo (veja por exemplo, Lawless, 1982, page 479) é dada por,

$$L = \prod_{i \in C_1} f(t_{1i}, t_{2i}) \prod_{i \in C_2} \left(-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} \right) \prod_{i \in C_3} \left(-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} \right) \prod_{i \in C_4} S(t_{1i}, t_{2i}), \quad (7.21)$$

em que $f(t_{1i}, t_{2i})$ é a função densidade de probabilidade conjunta para T_{1i} e T_{2i} ; $S(t_{1i}, t_{2i})$ é a função de sobrevivência conjunta; $\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}}$ e $\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}}$ são as derivadas parciais de $S(t_{1i}, t_{2i})$, com respeito a t_{1i} e t_{2i} respectivamente.

Definamos as variáveis indicadoras δ_{1i} e δ_{2i} , por,

$$\delta_{ji} = \begin{cases} 1 & \text{se } t_{ji} \text{ é um tempo de sobrevivência observado} \\ 0 & \text{se } t_{ji} \text{ é uma observação censurada,} \end{cases} \quad (7.22)$$

para $j = 1, 2$; $i = 1, 2, \dots, n$.

Desta maneira, reescrevemos a função de verossimilhança (7.21) como,

$$\begin{aligned} L &= \prod_{i=1}^n [f(t_{1i}, t_{2i})]^{\delta_{1i}\delta_{2i}} \prod_{i=1}^n \left[-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}} \right]^{\delta_{1i}(1-\delta_{2i})} \prod_{i=1}^n \left[-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}} \right]^{(1-\delta_{1i})\delta_{2i}} \times \\ &\times \prod_{i=1}^n [S(t_{1i}, t_{2i})]^{(1-\delta_{1i})(1-\delta_{2i})}. \end{aligned} \quad (7.23)$$

7.3.1 “Modelo 1”

Assumindo o “modelo 1” definido por (7.11), substituir na função de verossimilhança (7.23), $S(t_{1i}, t_{2i})$ por (7.13) e $f(t_{1i}, t_{2i})$, $-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}}$ e $-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}}$, respectivamente, por,

$$\begin{aligned} f_1(t_{1i}, t_{2i}) &= \frac{\partial^2 F_1(t_{1i}, t_{2i})}{\partial t_{1i} \partial t_{2i}} = \frac{p_1 p_2 t_{1i}^{p_1-1} t_{2i}^{p_2-1}}{\lambda_1^{p_1} \lambda_2^{p_2}} \left[\left(1 + \frac{\theta t_{1i}^{p_1}}{\lambda_1^{p_1}} \right) \left(1 + \frac{\theta t_{2i}^{p_2}}{\lambda_2^{p_2}} \right) - \theta \right] \times \\ &\times \exp \left\{ - \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} - \theta \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right\}, \end{aligned} \quad (7.24)$$

$$\begin{aligned}
-\frac{\partial S_1(t_{1i}, t_{2i})}{\partial t_{1i}} &= \frac{p_1 t_{1i}^{p_1-1}}{\lambda_1^{p_1}} \left[1 + \theta \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] \times \\
&\times \exp \left\{ - \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} - \theta \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right\}, \quad (7.25)
\end{aligned}$$

e

$$\begin{aligned}
-\frac{\partial S_1(t_{1i}, t_{2i})}{\partial t_{2i}} &= \frac{p_2 t_{2i}^{p_2-1}}{\lambda_2^{p_2}} \left[1 + \theta \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \right] \times \\
&\times \exp \left\{ - \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} - \theta \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right\}. \quad (7.26)
\end{aligned}$$

7.3.2 “Modelo 2”

Assumindo “modelo 2” definido por (7.16), temos em (7.23), $S(t_{1i}, t_{2i})$ substituído por (7.17) e $f(t_{1i}, t_{2i})$, $-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{1i}}$ e $-\frac{\partial S(t_{1i}, t_{2i})}{\partial t_{2i}}$, substituído, respectivamente por,

$$\begin{aligned}
f_2(t_{1i}, t_{2i}) &= \frac{p_1 p_2 t_{1i}^{p_1-1} t_{2i}^{p_2-1}}{\lambda_1^{p_1} \lambda_2^{p_2}} \exp \left\{ - \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right\} \left\{ 1 + \theta - 2\theta \exp \left[- \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \right] \right. \\
&- \left. 2\theta \exp \left[- \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] + 4\theta \exp \left[- \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] \right\}, \quad (7.27)
\end{aligned}$$

$$\begin{aligned}
-\frac{\partial S_2(t_{1i}, t_{2i})}{\partial t_{1i}} &= \frac{p_1 t_{1i}^{p_1-1}}{\lambda_1^{p_1}} \exp \left\{ - \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right\} \left\{ 1 + \theta - 2\theta \exp \left[- \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \right] \right. \\
&- \left. \theta \exp \left[- \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] + 2\theta \exp \left[- \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] \right\}, \quad (7.28)
\end{aligned}$$

e

$$\begin{aligned}
-\frac{\partial S_2(t_{1i}, t_{2i})}{\partial t_{2i}} &= \frac{p_2 t_{2i}^{p_2-1}}{\lambda_2^{p_2}} \exp \left\{ - \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right\} \left\{ 1 + \theta - \theta \exp \left[- \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} \right] \right. \\
&\quad \left. - 2\theta \exp \left[- \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] + 2\theta \exp \left[- \left(\frac{t_{1i}}{\lambda_1} \right)^{p_1} - \left(\frac{t_{2i}}{\lambda_2} \right)^{p_2} \right] \right\}. \quad (7.29)
\end{aligned}$$

Assumindo o “modelo 1” ou o “modelo 2”, podemos considerar as seguintes distribuições a priori para λ_1 , λ_2 , p_1 , p_2 e θ :

$$\begin{aligned}
\lambda_j &\sim U(a_j, b_j), \\
p_j &\sim U(c_j, d_j), \\
\theta &\sim U(e, f),
\end{aligned} \quad (7.30)$$

para $j = 1, 2$; $U[a, b]$ denotando uma distribuição uniforme no intervalo $[a, b]$; a_j , b_j , c_j , d_j , e e f são hiperparâmetros conhecidos. Além disso, assumimos independência a priori entre os parâmetros.

Outras distribuições a priori também podem ser consideradas, como distribuições gama.

A distribuição a posteriori conjunta para $\boldsymbol{\theta} = (\lambda_1, \lambda_2, p_1, p_2, \theta)'$ é dada por,

$$\pi(\boldsymbol{\theta}|\mathbf{t}) \propto \pi(\boldsymbol{\theta})L(\boldsymbol{\theta}), \quad (7.31)$$

no qual $\pi(\boldsymbol{\theta})$ é a distribuição a priori conjunta para $\boldsymbol{\theta}$; $L(\boldsymbol{\theta})$ é a função de verossimilhança (7.23) assumindo “modelo 1” ou “modelo 2” e $\mathbf{t} = (t_1, t_2, \dots, t_n)'$ é o vetor de dados observados.

Para obtermos os resumos a posteriori de interesse, simulamos amostras para a distribuição a posteriori conjunta (7.31) usando métodos MCMC.

Deste modo, podemos simular amostras para a distribuição a posteriori conjunta das distribuições condicionais $\pi(\lambda_1|\lambda_2, p_1, p_2, \theta, \mathbf{t})$, $\pi(\lambda_2|\lambda_1, p_1, p_2, \theta, \mathbf{t})$, $\pi(p_1|\lambda_1, \lambda_2, p_2, \theta, \mathbf{t})$, $\pi(p_2|\lambda_1, \lambda_2, p_1, \theta, \mathbf{t})$ and $\pi(\theta|\lambda_1, \lambda_2, p_1, p_2, \mathbf{t})$ usando os algoritmos amostrador de Gibbs ou o Metropolis-Hastings quando as distribuições condicionais não são identificadas como distribuições conhecidas que são fáceis de simular.

7.4 Presença de Covariáveis e Dados Censurados

Agora assumamos a presença de um vetor $\mathbf{x} = (x_1, x_2, \dots, x_p)'$ de covariáveis associada a cada tempo de sobrevivência bivariado (t_1, t_2) assumindo o “modelo 1” ou o “modelo 2”.

Na presença de um vetor de covariáveis \mathbf{x} , consideremos o seguinte modelo de regressão:

$$\lambda_{1i} = \alpha_1 \exp \{ \beta_1' \mathbf{x}_i \}, \quad (7.32)$$

$$\lambda_{2i} = \alpha_2 \exp \{ \beta_2' \mathbf{x}_i \},$$

em que $\beta_j = (\beta_{j1}, \beta_{j2}, \dots, \beta_{jp})'$, $j = 1, 2$ é o vetor de parâmetros da regressão $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})'$, $i = 1, 2, \dots, n$. Também assumimos a presença de observações censuradas.

Para uma análise Bayesiana, assumimos as seguintes distribuições a priori para α_j , p_j , β_{jl} e θ :

$$\begin{aligned} \alpha_j &\sim U(a_j, b_j), \\ p_j &\sim U(c_j, d_j), \\ \theta &\sim U(e, f), \\ \beta_{jl} &\sim N(0, g^2), \end{aligned} \quad (7.33)$$

para $j = 1, 2$; $l = 1, 2, \dots, p$; a_j , b_j , c_j , d_j , e , f e g são hiperparâmetros conhecidos e $N(0, g^2)$ denota uma distribuição normal com média zero e variância igual a g^2 . Além disso assumimos independência a priori entre os parâmetros.

7.5 Exemplos

7.5.1 Tempos de Recorrência para a Infecção de rins de Pacientes

Como um primeiro exemplo, reconsiderar o conjunto de dados de sobrevivência introduzido por McGilchrist & Aisbett (1991), relacionado a infecção de rins, onde há recorrência da infecção dos rins de 38 (dados introduzidos na Tabela 2.1).

Para analisar os tempos de recorrência das infecções, assumir as duas distribuições Weibull bivariadas obtidas de funções cópulas, (“modelo 1” e “modelo 2”) introduzidos na Seção 7.2.

Como uma primeira análise, assumamos o “modelo 1” definido pela função distribuição conjunta (7.11) não considerando a presença da covariável sexo.

Para uma análise Bayesiana do “modelo 1” não considerando a presença da covariável sexo, assumamos distribuições a priori uniformes (veja (7.30)) com valores dos hiperparâmetros $a_1 = a_2 = 0$; $b_1 = b_2 = 500$, $c_1 = c_2 = 0$; $d_1 = d_2 = 2$; $e = 0$ e $f = 0,2$ para $\lambda_1, \lambda_2, p_1, p_2$ e θ .

Para simular amostras da distribuição a posteriori conjunta para $\lambda_1, \lambda_2, p_1, p_2$ e θ , usamos o *software WinBUGS*. No procedimento de simulação, descartamos as primeiras 5000 amostras simuladas por Gibbs (“*burn-in-samples*”) para eliminar o efeito dos valores iniciais para os parâmetros do modelo. Escolhendo a cada 20^a amostra simulada por Gibbs, obtemos uma amostra final de tamanho 5000 para gerar os resumos a posteriori de interesse.

Convergência do algoritmo amostrador de Gibbs foi monitorado usando métodos padrões existentes, como gráficos de séries temporais para as amostras simuladas.

Na Tabela 7.1, temos a média a posteriori, desvios-padrão a posteriori e intervalos de credibilidade de 95% para os parâmetros do modelo.

Nas Figuras 7.1 e 7.2, temos os gráficos de séries temporais e o de Gelman-Rubin para as amostras simuladas por Gibbs. Observamos a convergência para todos os parâmetros do “modelo 1” não considerando a covariável sexo.

Observe que usando métodos de inferência clássicos baseados na normalidade assintótica dos estimadores de máxima verossimilhança para $\lambda_1, \lambda_2, p_1, p_2$ e θ , podemos ter grandes dificuldades computacionais para a estimação.

Na presença da covariável sexo, denotada por x , assumimos o modelo de regressão

Tabela 7.1: Resumos a posteriori (“modelo 1” não considerando a covariável sexo)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
λ_1	135,3000	32,96000	(82,8200; 210,500)
λ_2	144,6000	31,66000	(94,2600; 217,400)
p_1	0,805500	0,109900	(0,59980; 1,03100)
p_2	0,989900	0,141300	(0,72670; 1,28200)
θ	0,076850	0,054840	(0,00283; 0,18940)
$1/\lambda_1$	0,007820	0,001865	(0,00475; 0,01209)
$1/\lambda_2$	0,007234	0,001526	(0,00460; 0,01061)

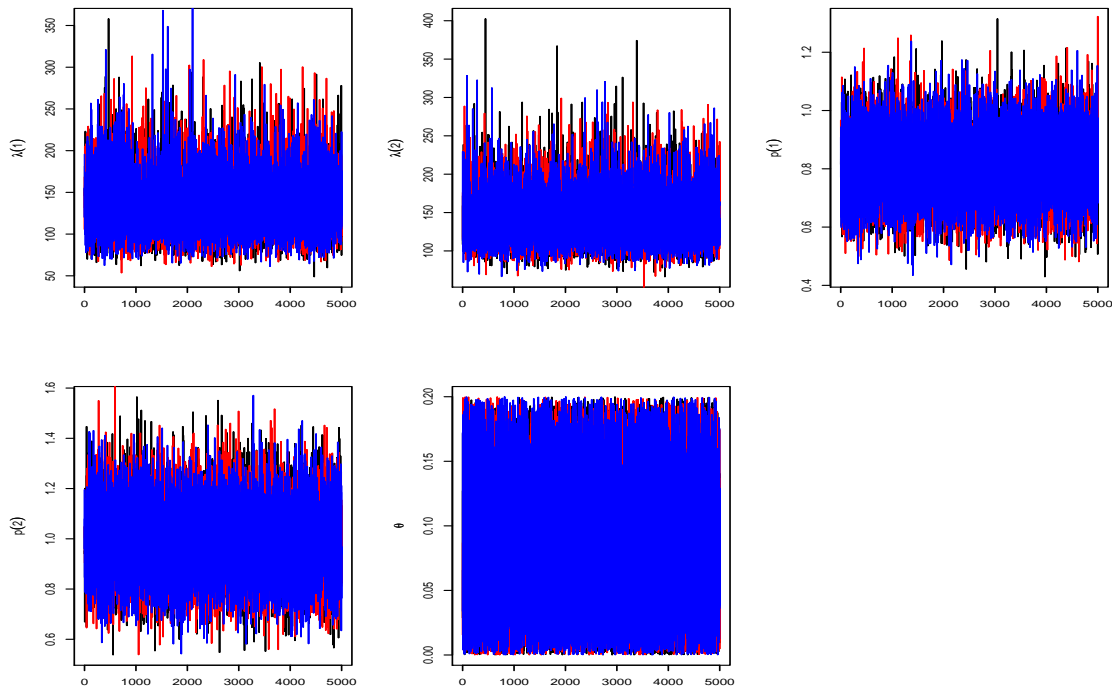


Figura 7.1: Amostras simuladas por Gibbs (“modelo 1” não considerando a covariável sexo)

introduzido na Seção 7.4, ou seja,

$$\lambda_{1i} = \alpha_1 \exp \{ \beta_1 \mathbf{x}_i \} \quad (7.34)$$

$$\lambda_{2i} = \alpha_2 \exp \{ \beta_2 \mathbf{x}_i \}$$

onde $i = 1, 2, \dots, 38$; $x_i = 1$ (masculino) e $x_i = 0$ (feminino), e “modelo 1” definido por (7.11).

Assumindo a distribuição a priori (7.33) com valores dos hiperparâmetros $a_1 = a_2 = 0$;

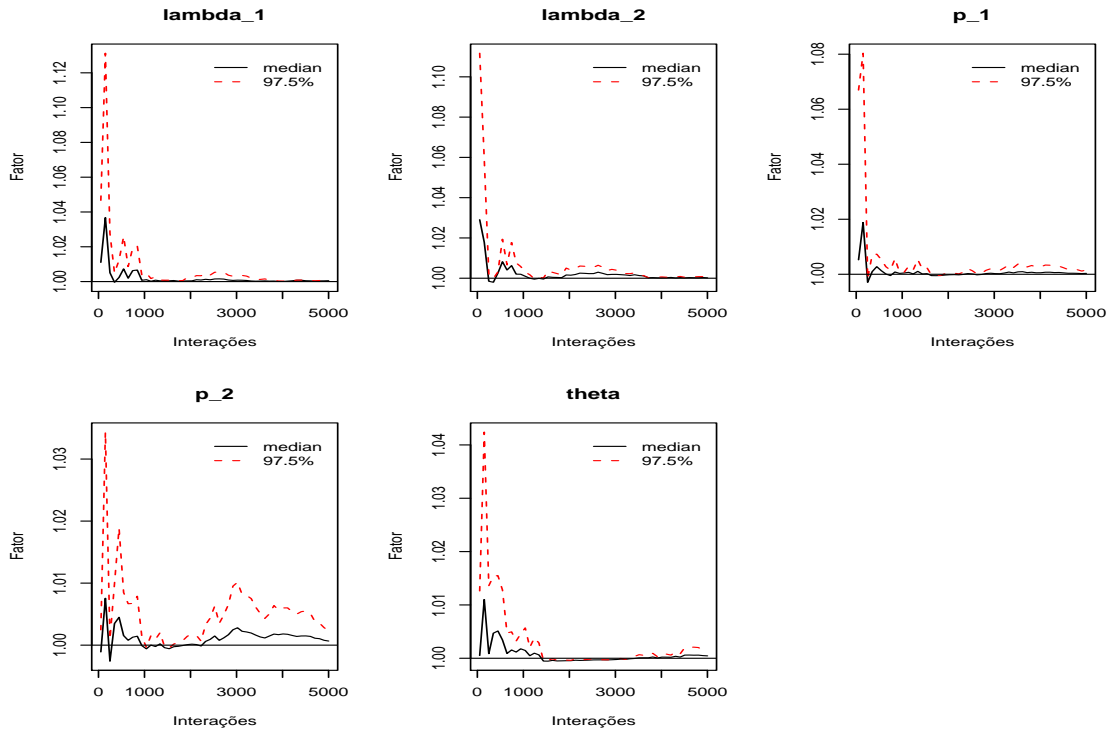


Figura 7.2: Valores de Gelman e Rubin (“modelo 1” não considerando a covariável sexo)

$b_1 = b_2 = 500$, $c_1 = c_2 = 0$; $d_1 = d_2 = 2$; $e = 0$; $f = 0,2$ e $g = 1$, e usando o *software WinBUGS* e também considerando um “burn-in-sample” de tamanho 5000 e uma amostra de Gibbs final de tamanho 5000, escolhendo cada 20^a amostra, temos na Tabela 7.2 os resumos a posteriori de interesse.

Dos resultados da Tabela 7.2, observamos que o valor zero está incluído no intervalo de credibilidade de 95% para β_2 , mas não para β_1 , ou seja, o tempo de sobrevivência T_2 não é afetado pelo sexo, mas T_1 é afetado pelo sexo.

Tabela 7.2: Resumos a posteriori (“modelo 1” na presença da covariável sexo)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
α_1	180,000	43,6600	(111,100; 281,40)
α_2	160,300	43,9400	(97,0500; 267,20)
β_1	-1,3990	0,42430	(-2,2100; -0,527)
β_2	-0,2329	0,44380	(-1,1150; 0,6322)
p_1	0,91900	0,12440	(0,68590; 1,1760)
p_2	0,94990	0,13910	(0,68990; 1,2400)
θ	0,06043	0,04984	(0,00162; 0,1787)

Nas Figuras 7.3 e 7.4, temos os gráficos de séries temporais e o de Gelman-Rubin para as amostras simuladas por Gibbs. Observamos a convergência para todos os parâmetros do “modelo 1” não considerando a covariável sexo.

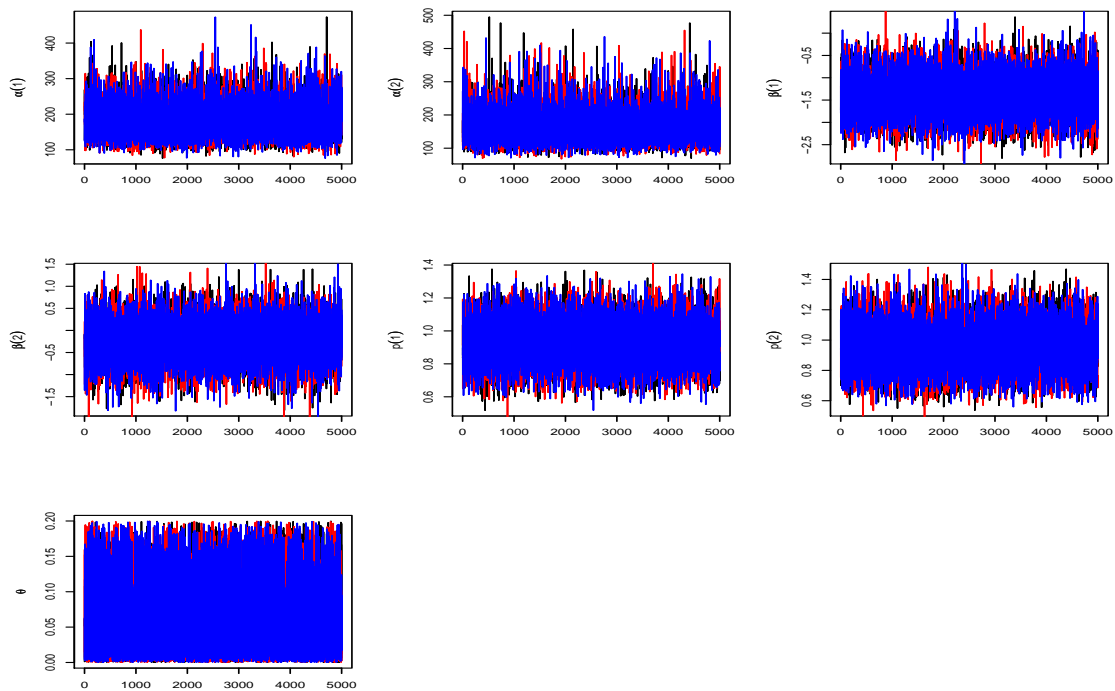


Figura 7.3: Amostras simuladas por Gibbs (“modelo 1”, na presença da covariável sexo)

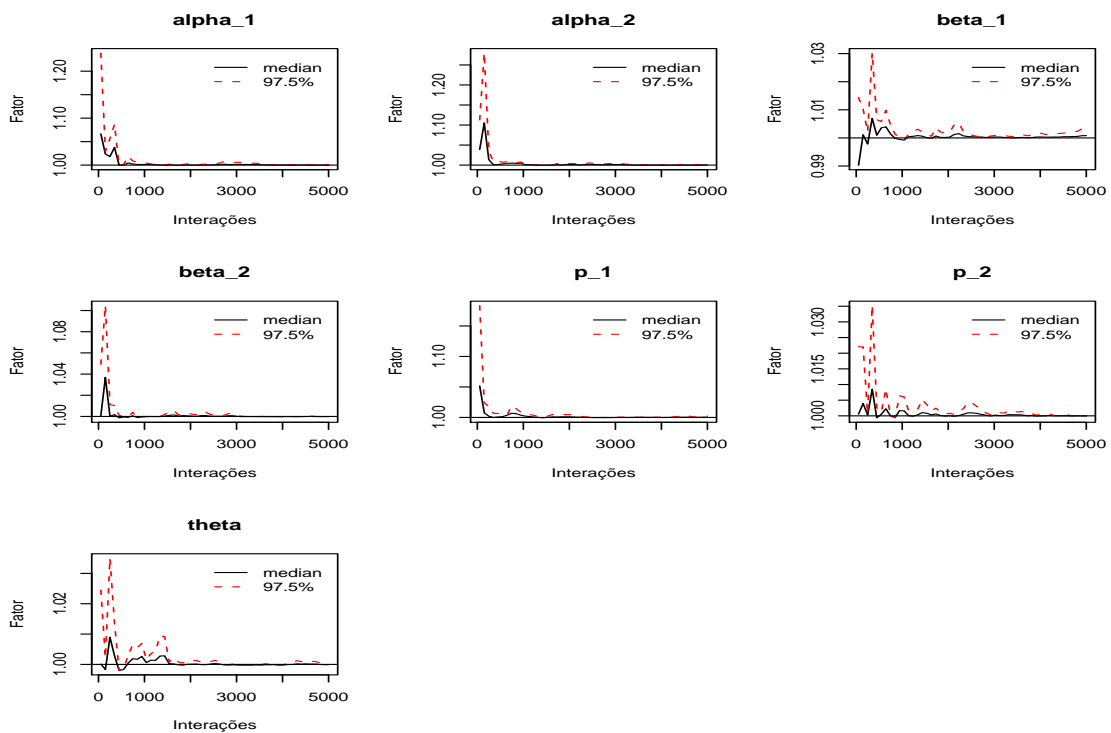


Figura 7.4: Valores de Gelman e Rubin (“modelo 1”, na presença da covariável sexo)

Para comparar os modelos propostos, usamos o critério de discriminação de modelo existente, como por exemplo o DIC (*Deviance information criterion*) introduzido por Spiegelhalter *et al.* (2002), e é dado automaticamente pelo *software WinBUGS*. Assumindo o “modelo 1” não considerando a presença da covariável sexo, temos um estimador de Monte Carlo para o DIC dado pelo valor 688,961; assumindo o “modelo 1” na presença da covariável sexo, o valor de DIC é dado por 680,427. Ou seja, o “modelo 1”, na presença da covariável sexo apresenta melhor ajuste para os dados da Tabela 2.1 (menor valor para o DIC).

Como segunda análise, assumir o “modelo 2” definido pela função de distribuição conjunta (7.16).

Para uma análise Bayesiana do “modelo 2” não considerando a presença da covariável sexo, assumimos as distribuições a priori uniformes dadas em (7.30), com os mesmos valores dos hiperparâmetros assumidos considerando o “modelo 1”.

Seguindo o mesmo procedimento de simulação considerado para o “modelo 1” e usando o *software WinBUGS*, temos na Tabela 7.3 os resumos a posteriori de interesse. Convergência do algoritmo amostrador de Gibbs foi monitorado, usando gráficos de séries temporais e o Gelman-Rubin das amostras simuladas.

Dos resultados da Tabela 7.3 considerando o “modelo 2”, observamos resultados similares aos obtidos considerando o “modelo 1” (veja Tabela 7.1).

Tabela 7.3: Resumos a posteriori (“modelo 2” não considerando a covariável sexo)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
λ_1	136,100	32,8900	(81,770; 208,20)
λ_2	143,200	30,2200	(93,010; 210,10)
p_1	0,82740	0,01088	(0,6311; 1,1036)
p_2	0,99690	0,01427	(0,7412; 1,3303)
θ	0,09880	0,05739	(0,0071; 0,2008)
$1/\lambda_1$	0,00776	0,00193	(0,0049; 0,0142)
$1/\lambda_2$	0,00721	0,00152	(0,0048; 0,0117)

Assumindo o “modelo 2” na presença da covariável sexo, o mesmo modelo de regressão (7.34) e as mesmas distribuições a priori assumidas para o “modelo 1” na presença da covariável sexo, temos na Tabela 7.4, os resumos a posteriori de interesse obtidos de uma amostra final simulada por Gibbs de tamanho 5000, usando o *software WinBUGS*.

Dos resultados da Tabela 7.4, observamos resultados similares aos do “modelo 1” (veja Tabela 7.2). A covariável sexo tem um efeito significativo no tempo de sobrevivência T_1 , uma vez que o valor zero não está incluído no intervalo de credibilidade de 95% para β_1 .

Estimadores de Monte Carlo para o DIC baseados nas 5000 amostras simuladas por Gibbs são dados por 687,356 (“modelo 2” não considerando a covariável sexo) e 676,995

Tabela 7.4: Resumos a posteriori (“modelo 2” na presença da covariável sexo)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
α_1	180,600	42,8300	(112,90; 279,40)
α_2	160,500	42,4100	(96,580; 262,60)
β_1	-1,4562	0,41222	(-2,261; -0,608)
β_2	-0,2891	0,43023	(-1,113; 0,5876)
p_1	0,94074	0,12761	(0,6944; 1,1860)
p_2	0,95883	0,13564	(0,6993; 1,2390)
θ	0,10233	0,05293	(0,0061; 0,1976)

(“modelo 2” na presença da covariável sexo). Ou seja, o “modelo 2” na presença da covariável sexo, dá um melhor ajuste para os tempos de sobrevivência bivariados da Tabela 2.1. Além de tudo, o “modelo 2” na presença da covariável sexo é o melhor, quando comparado com todos os outros modelos (“modelo 1” não considerando a covariável sexo; “modelo 1” na presença da covariável sexo; “modelo 2” não considerando a covariável sexo e “modelo 2” na presença da covariável sexo).

7.5.2 Transplante de Medula Óssea para Leucemia

Vamos considerar novamente os dados de transplante de medula óssea analisados no Capítulo 6. Transplantes de medula óssea são um tratamento padrão para leucemia aguda. Prognósticos para a recuperação podem depender de fatores de risco conhecidos no tempo do transplante, tais como paciente e/ou idade e sexo do doador, o estágio inicial da doença, o tempo de prognóstico para transplante entre muitos outros. O prognóstico final pode mudar quando o desenvolvimento histórico pós-transplante com ocorrência de eventos em tempos aleatórios durante o processo de recuperação, tais como o desenvolvimento de doença aguda ou crônica transplantar-versus-hospedar (GVHD), retornar a contagem de plaquetas para níveis normais, ou o desenvolvimento de infecções.

Neste estudo, 137 pacientes com leucemia receberam uma combinação de 16 mg/kg de Busulfan oral (BU) e 120 mg/kg de cyclophosphamide intravenoso (Cy) e foram tratados em um dos quatro hospitais: 76 no da Universidade estadual de Ohio (OSU) em Columbus; 21 na Universidade Hahnemann (HU) na Philadelphia; 23 no Hospital St. Vincent’s (SVH) em Sidney, Austrália e 17 no Hospital Alfred (AH) em Melbourne, Austrália (conjunto de dados introduzidos por Klein & Moeschberger, 1997, página 464).

Na análise considerada nesta seção, assumimos como tempos de sobrevivência, os tempos (em dias) a doença aguda transplantar-versus-hospedar (TA) com 111 observações

censuradas e 26 não-censuradas e o tempo (em dias) da doença crônica transplantar-versus-hospedar (TC) com 76 observações censuradas e 61 não-censuradas e as seguintes covariáveis: idade dos pacientes em anos, idade do doador em anos, sexo do paciente, sexo do doador, CMV (cytomegalovirus immune status) do paciente, CMV do doador, tempo de espera do transplante em dias e os diferentes hospitais.

Para analisar este conjunto de dados, assumimos distribuições Weibull bivariadas (“modelo 1” na presença ou não de covariáveis e “modelo 2” na presença ou não de covariáveis) obtidas de funções cópulas introduzidas na Seção 7.2.

Para todos os casos, assumimos o mesmo procedimento de simulação: simulação de 5000 amostras de Gibbs como “burn-in-sample” e uma amostra final de Gibbs também de tamanho 5000, pegando a cada 20ª amostra simulada e usando o *software WinBUGS*.

Assumindo o “modelo 1” com função de distribuição conjunta dada em (7.11) não considerando a presença de covariáveis, assumimos a reparametrização $\lambda_1 = \exp(\gamma_1)$, $\lambda_2 = \exp(\gamma_2)$ e as seguintes distribuições a priori para $\gamma_1, \gamma_2, p_1, p_2$ e θ : prioris uniformes $U(5; 100)$ para γ_1 e γ_2 ; prioris uniformes $U(0; 2)$ para p_1 e p_2 , e uma priori uniforme $U(0; 0, 5)$ para θ , ou seja, distribuições a priori aproximadamente não-informativas. A reparametrização para λ_1 e λ_2 foi necessária para melhorar a convergência do algoritmo amostrador de Gibbs.

Na Tabela 7.5, temos os resumos a posteriori de interesse.

Tabela 7.5: Resumos a posteriori (“modelo 1” não considerando a presença de covariáveis)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
γ_1	11,4000	1,13000	(9,6160; 13,990)
γ_2	7,13200	0,19974	(6,7720; 7,5520)
λ_1	260700	2827000	(15000; 1190000)
λ_2	1277,00	267,200	(873,10; 1904,0)
p_1	0,32601	0,05852	(0,2215; 0,4514)
p_2	0,69782	0,06971	(0,5679; 0,8388)
θ	0,09454	0,08603	(0,0003; 0,3234)
$1/\lambda_1$	0,00002	0,00002	(0,000001; 0,000067)
$1/\lambda_2$	0,00081	0,00016	(0,000525; 0,001145)

Nas Figuras 7.5 e 7.6, temos os gráficos de séries temporais e o de Gelman-Rubin para as amostras simuladas por Gibbs. Observamos a convergência para todos os parâmetros do “modelo 1” não considerando a covariável sexo.

Dos resultados da Tabela 7.5, observamos que o estimador de Monte Carlo para a média a posteriori de λ_1 baseado nas 5000 amostras simuladas por Gibbs é muito grande. De fato, temos 111 observações censuradas e somente 26 não-censuradas para os tempos de sobrevivência TA. Possivelmente uma distribuição a priori informativa usando informação a priori de especialistas poderia melhorar esse resultado.

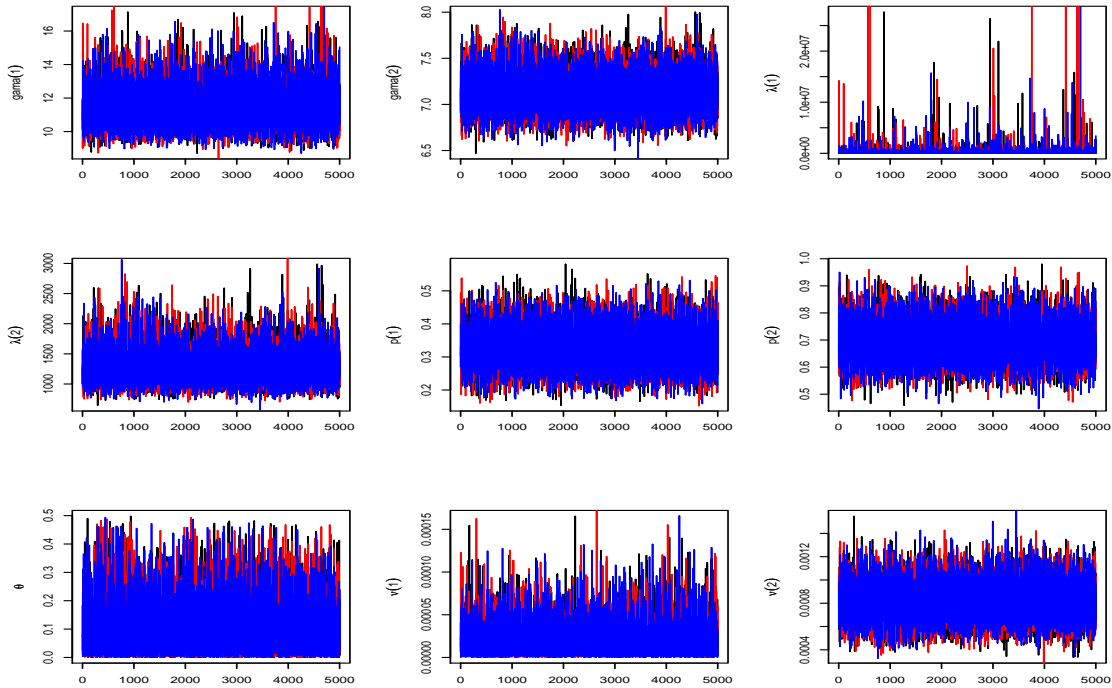


Figura 7.5: Amostras simuladas por Gibbs (“modelo 1”, não considerando a presença de covariáveis)

Assumindo o “modelo 1” na presença de covariáveis, assumamos o seguinte modelo de regressão:

$$\begin{aligned} \lambda_{1i} = & \alpha_1 \exp \{ \beta_{11} (Z_{1i} - \overline{Z_1}) + \beta_{12} (Z_{2i} - \overline{Z_2}) + \beta_{13} Z_{3i} + \beta_{14} Z_{4i} + \beta_{15} Z_{5i} + \beta_{16} Z_{6i} + \\ & + \beta_{17} (Z_{7i} - \overline{Z_7}) + \beta_{18} Z_{8i} + \beta_{19} Z_{9i} + \beta_{1,10} Z_{10i} \} \\ \text{e} & \end{aligned} \quad (7.35)$$

$$\begin{aligned} \lambda_{2i} = & \alpha_2 \exp \{ \beta_{21} (Z_{1i} - \overline{Z_1}) + \beta_{22} (Z_{2i} - \overline{Z_2}) + \beta_{23} Z_{3i} + \beta_{24} Z_{4i} + \beta_{25} Z_{5i} + \beta_{26} Z_{6i} + \\ & + \beta_{27} (Z_{7i} - \overline{Z_7}) + \beta_{28} Z_{8i} + \beta_{29} Z_{9i} + \beta_{2,10} Z_{10i} \} \end{aligned}$$

onde Z_{1i} é a idade do paciente; Z_{2i} é a idade do doador; Z_{3i} é o sexo do paciente; Z_{4i} é o sexo do doador; Z_{5i} é o CMV do paciente; Z_{6i} é o CMV do doador; Z_{7i} é o tempo de espera para o transplante em dias; Z_{8i} , Z_{9i} e Z_{10i} são variáveis “dummy” relacionadas aos diferentes hospitais, ($Z_{8i} = 1$ para o OSU; 0 para outros hospitais; $Z_{9i} = 1$ para o HU; 0 para outros hospitais; $Z_{10i} = 1$ para o SVH; 0 para outros hospitais); $\overline{Z_1}$, $\overline{Z_2}$ e $\overline{Z_7}$ são médias amostrais para as covariáveis Z_{1i} , Z_{2i} e Z_{7i} , respectivamente, $i = 1, 2, \dots, 137$.

Para o “modelo 1” na presença de covariáveis, assumimos as seguintes distribuições a priori (veja (7.33)): $\gamma_1 \sim U(5; 100)$ onde $\alpha_1 = \exp(\gamma_1)$; $\alpha_2 \sim U(0; 10000)$; $\theta \sim U(0; 0, 3)$;

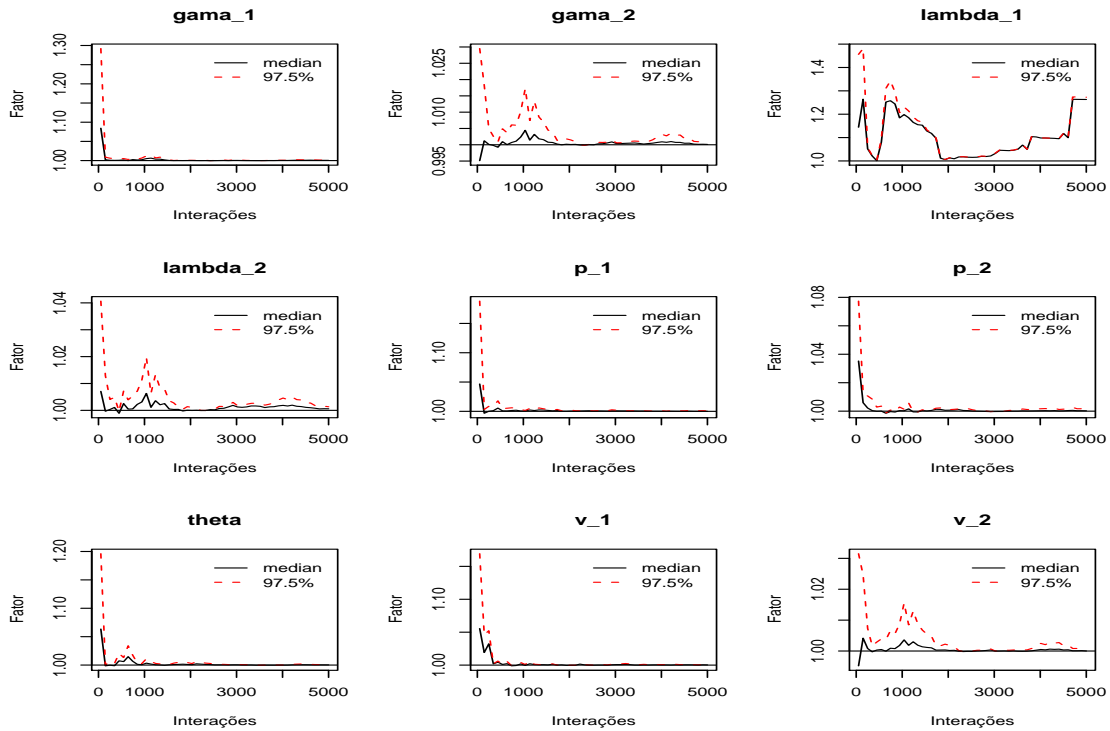


Figura 7.6: Valores de Gelman e Rubin (“modelo 1”, não considerando a presença de covariáveis)

$p_j \sim U(0; 2)$ para $j = 1, 2$ e distribuições a priori uniformes $U(-a_{jl}, a_{jl})$ para β_{jl} onde a_{jl} são hiperparâmetros conhecidos com pequenos valores, $j = 1, 2; l = 1, 2, \dots, 10$.

Também foi usado distribuições a priori normais para os parâmetros de regressão β_{jl} , mas a convergência do algoritmo amostrador de Gibbs usando o *software WinBUGS* somente foi obtida tomando variâncias muito pequenas para as distribuições a priori normais, com médias iguais a zero.

Observe também que foi utilizado distribuições a priori não-informativas para $\gamma_1, \alpha_2, \theta$ e $p_j, j = 1, 2$.

Na Tabela 7.6, temos os resumos a posteriori de interesse.

Dos resultados da Tabela 7.6, observamos que o zero está incluído no intervalo de credibilidade para muitos parâmetros da regressão $\beta_{jl}, j = 1, 2$ e $l = 1, 2, \dots, 10$; uma indicação que as correspondentes covariáveis não têm um efeito significativo nos tempos de sobrevivência TA e TC. Para alguns casos, como os parâmetros de regressão β_{11} e β_{22} , observamos que o zero não está incluído no intervalo de credibilidade. Isto é uma indicação que as covariáveis Z_1 (idade do paciente) e Z_2 (idade do doador) têm efeito significativo na doença aguda (TA) e na doença crônica (TC), respectivamente.

Também observamos alguma indicação de efeito para a covariável Z_7 (tempo de espera para o transplante) para o tempo de doença crônica, uma vez que o zero está muito

Tabela 7.6: Resumos a posteriori (“modelo 1” na presença de covariáveis)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 80%
α_1	427400	420300	(32240; 659310)
α_2	1250	325, 400	(879, 3; 1659, 0)
γ_1	11, 8100	1, 18900	(9, 465; 13, 67)
p_1	0, 32180	0, 05743	(0, 2487; 0, 3977)
p_2	0, 71110	0, 07438	(0, 6108; 0, 8084)
θ	0, 09120	0, 07312	(0, 0103; 0, 1935)
β_{11}	-0, 09516	0, 06528	(-0, 1737; -0, 0035)
β_{12}	-0, 07794	0, 06739	(-0, 1667; 0, 0157)
β_{13}	0, 00252	0, 11560	(-0, 1689; 0, 1634)
β_{14}	0, 00786	0, 11560	(-0, 1620; 0, 1679)
β_{15}	0, 00434	0, 11440	(-0, 1588; 0, 1572)
β_{16}	-0, 01465	0, 11460	(-0, 1196; 0, 1518)
β_{17}	-0, 00081	0, 00164	(-0, 0027; 0, 0013)
β_{18}	-0, 00098	0, 08657	(-0, 1228; 0, 1231)
β_{19}	-0, 00392	0, 11510	(-0, 1654; 0, 1504)
$\beta_{1,10}$	0, 05536	0, 09691	(-0, 0916; 0, 1686)
β_{21}	0, 02960	0, 03324	(-0, 0135; 0, 0687)
β_{22}	-0, 05187	0, 03173	(-0, 0901; -0, 0112)
β_{23}	0, 02353	0, 11160	(-0, 1420; 0, 1658)
β_{24}	0, 04961	0, 09866	(-0, 1021; 0, 1739)
β_{25}	-0, 03293	0, 08161	(-0, 1340; 0, 0918)
β_{26}	0, 02254	0, 11190	(-0, 1543; 0, 1684)
β_{27}	-0, 00058	0, 00048	(-0, 0011; -0, 0001)
β_{28}	-0, 01894	0, 08477	(-0, 1234; 0, 1107)
β_{29}	-0, 01350	0, 08605	(-0, 1241; 0, 1049)
$\beta_{2,10}$	-0, 00139	0, 11530	(-0, 1650; 0, 1586)

próximo do limite superior do intervalo de credibilidade de 80% para β_{27} (-0, 0011; 0, 00005).

O valor de DIC dado automaticamente pelo *software WinBUGS* é 1382,79 para o “modelo 1” não considerando a presença de covariáveis e 1379.41 para o “modelo 1” na presença de covariáveis, ou seja, uma indicação de que o “modelo 1” na presença de covariáveis melhor se ajusta aos dados.

Assumindo o “modelo 2” definido pela função de distribuição conjunta (7.16) não considerando a presença de covariáveis, também consideramos a reparametrização $\lambda_1 = \exp(\gamma_1)$ e $\lambda_2 = \exp(\gamma_2)$ e as mesmas distribuições a priori assumidas para o “modelo 1” não considerando a presença de covariáveis.

Na Tabela 7.7, temos os resumos a posteriori de interesse baseados nas 5000 amostras simuladas por Gibbs.

Dos resultados obtidos na Tabela 7.7, observamos inferências similares as obtidas considerando o “modelo 1” (veja tabela 7.5), exceto para o parâmetro θ que mede a associação

Tabela 7.7: Resumos a posteriori (“modelo 2” não considerando a presença de covariáveis)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
γ_1	11,3700	1,15100	(9,5160; 14,090)
γ_2	7,08201	0,19330	(6,7420; 7,5350)
λ_1	258700	1089000	(14000; 1310000)
λ_2	1232,00	249,300	(826,40; 1889,0)
p_1	0,32754	0,06101	(0,2015; 0,4611)
p_2	0,71032	0,07166	(0,5764; 0,8369)
θ	0,29584	0,12847	(0,0279; 0,4732)

entre os tempos de sobrevivência TA e TC (0,2958 assumindo o “modelo 2” e 0,0945 assumindo o “modelo 1”).

Assumindo o “modelo 2” na presença de covariáveis, o mesmo modelo de regressão (7.35) e as mesmas distribuições a priori consideradas para o “modelo 1” na presença de covariáveis, temos na Tabela 7.8, os resumos a posteriori de interesse obtidos de uma amostra simulada por Gibbs de tamanho 5000.

Dos resultados da Tabela 7.8, observamos resultados das inferências similares quando considerando o “modelo 1” (veja Tabela 7.6), exceto para o parâmetro θ .

Os valores de DIC para o “modelo 2” são dados por 1378,96 (“modelo 2” não considerando as covariáveis) e 1376,57 (“modelo 2” na presença de covariáveis), ou seja, temos indicação de que a presença de covariáveis melhora o ajuste do modelo para o conjunto de dados. Além de tudo, o “modelo 2” parece melhor quando comparado com o “modelo 1”, já que tem menor valor para o DIC.

Observar que o uso de funções cópulas pode abrir novas perspectivas para a construção de distribuições de sobrevivência multivariadas. Esses modelos são fundamentais para analisar dados de sobrevivência médicos multivariados, como observados nos dois exemplos considerados. Na prática existe muitas dificuldades na identificação de possíveis fatores de risco nos tempos de sobrevivência, como foi observado no exemplo de dados de transplante de medula óssea.

Esses resultados são de grande interesse nas aplicações médicas. O uso de métodos Bayesianos, via técnicas de simulação MCMC pode ser uma grande revolução na análise de dados de sobrevivência multivariados na presença de censuras e covariáveis.

Além disso, o uso do *software WinBUGS* leva a grandes simplificações na obtenção das inferências de interesse.

Observar que os resultados obtidos neste capítulo podem ser generalizados para vetores de dimensão maior do que dois e outras distribuições para os dados de sobrevivência.

Conclusões referentes as modelagens apresentadas anteriormente e novas perspectivas de modelagem, que serão objetos de futuras pesquisas, são apresentadas no próximo

Tabela 7.8: Resumos a posteriori (“modelo 2” na presença de covariáveis)

Parâmetro	Média	D.P.	Intervalo de credibilidade de 95%
α_1	382100	1237000	(31567; 685400)
α_2	1207,00	293,900	(870, 2; 1575)
γ_1	12,0320	1,21700	(10, 33; 13, 49)
p_1	0,32141	0,05863	(0, 2388; 0, 3967)
p_2	0,71830	0,07107	(0, 6157; 0, 8092)
θ	0,17118	0,08371	(0, 0447; 0, 2821)
β_{11}	-0,09054	0,06817	(-0, 1702; -0, 0006)
β_{12}	-0,07763	0,06482	(-0, 1633; 0, 0108)
β_{13}	0,00168	0,10997	(-0, 1564; 0, 1582)
β_{14}	0,01506	0,11430	(-0, 1484; 0, 1639)
β_{15}	0,00007	0,11415	(-0, 1592; 0, 1602)
β_{16}	-0,02133	0,11200	(-0, 1708; 0, 1438)
β_{17}	-0,00079	0,00156	(-0, 0029; 0, 0012)
β_{18}	0,00372	0,08934	(-0, 1201; 0, 1232)
β_{19}	-0,00712	0,11730	(-0, 1621; 0, 1607)
$\beta_{1,10}$	0,06023	0,09611	(-0, 0847; 0, 1752)
β_{21}	0,02894	0,03428	(-0, 0137; 0, 0698)
β_{22}	-0,05318	0,03095	(-0, 0932; -0, 0101)
β_{23}	0,02132	0,12311	(-0, 1392; 0, 1690)
β_{24}	0,04766	0,09415	(-0, 0941; 0, 1632)
β_{25}	-0,02877	0,08690	(-0, 1322; 0, 1083)
β_{26}	0,02531	0,11501	(-0, 1403; 0, 1645)
β_{27}	-0,00051	0,00022	(-0, 0011; 0, 0001)
β_{28}	-0,02340	0,08374	(-0, 1305; 0, 1027)
β_{29}	-0,01171	0,08295	(-0, 1238; 0, 1104)
$\beta_{2,10}$	-0,00166	0,11826	(-0, 1611; 0, 1496)

capítulo.

Capítulo 8

Conclusões e Perspectivas Futuras

Dados de sobrevivência longitudinais são comuns em muitos estudos em diversas áreas, como em medicina ou em engenharia. Usualmente, temos medidas repetidas para o mesmo paciente ou unidade. Nestes estudos, é comum a presença de covariáveis ou dados censurados. O uso de modelos Bayesianos hierárquicos, com “fragilidades” ou variáveis latentes assumindo diferentes estruturas, é um poderoso recurso para obtermos as inferências de interesse.

Observe que considerando tempos de sobrevivência independentes, com distribuição Weibull (4.1) e modelo de regressão (4.3) para a análise dos dados de sobrevivência introduzidos na Tabela 2.1, temos o valor de DIC dado por 678,82, considerando prioris não-informativas para os parâmetros do modelo e mesmos passos para o algoritmo de Gibbs, assumidos nos outros modelos propostos. Ou seja, já que o DIC é maior assumindo o modelo Weibull independente, temos uma grande indicação da presença de uma estrutura de correlação para os dados da tabela 2.1. Isso comprova a necessidade de usar modelos multivariados para dados de sobrevivência.

Na Tabela 8.1 temos os resumos a posteriori, assumindo modelos Weibull independentes. Nas Figuras 8.1 e 8.2, temos os gráficos para as amostras simuladas por Gibbs e o Gelman-Rubin, considerando os modelos Weibull independentes.

Nesta aplicação, já que temos somente uma covariável x_i (sexo; $x_i = 1$ para masculino e $x_i = 0$ para feminino), podemos comparar as médias e variâncias obtidas, assumindo a distribuição Weibull independente e o “modelo 1”, na presença de uma “fragilidade”. Observe que para o “modelo 1”, usamos as fórmulas aproximadas (4.10) e (4.12) para a média e variância não-condicionais dos tempos de sobrevivência (veja Tabela 8.2).

Tabela 8.1: Resumos a posteriori (modelos Weibull independentes)

Parâmetro	Média	D.P.	Intervalo de Credibilidade de 95%
β_{21}	1,55402	0,42712	(0,6873; 2,3750)
β_{22}	0,25361	0,43511	(-0,6191; 1,1000)
β_{11}	-4,87100	0,70781	(-6,3190; -3,5540)
β_{12}	-4,88802	0,72980	(-6,4030; -3,5710)
med 1	123,801	31,6400	(71,730; 193,80)
med 2	104,104	28,4700	(61,170; 171,00)
ν_1	0,93883	0,12383	(0,7112; 1,1960)
ν_2	0,97923	0,13924	(0,7279; 1,2720)

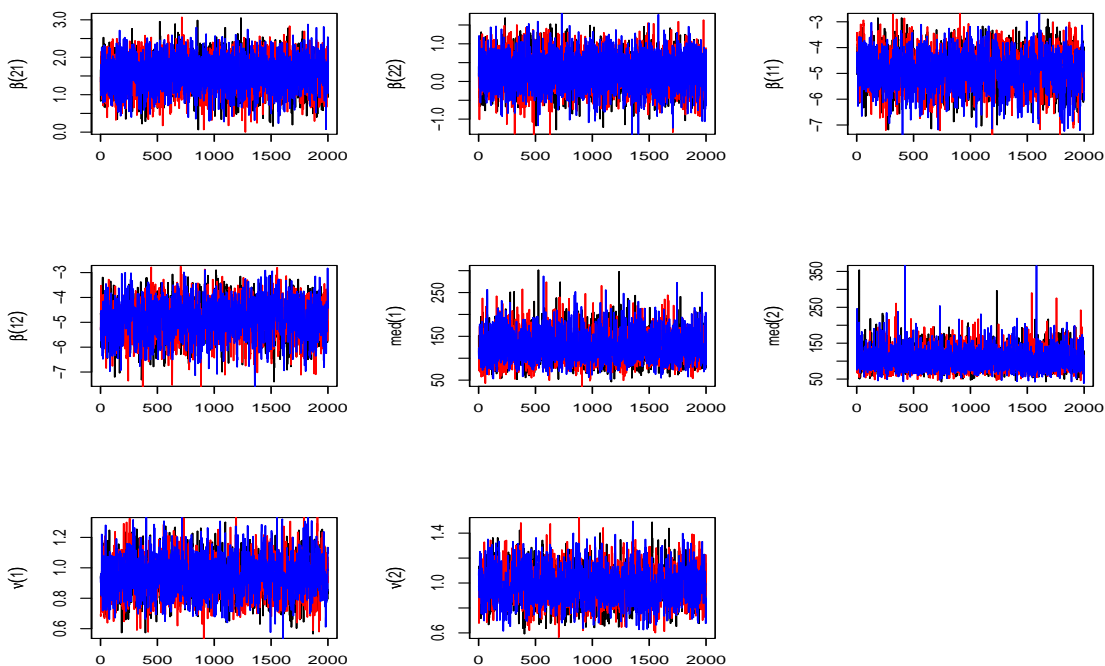


Figura 8.1: Amostras simuladas por Gibbs (modelos Weibull independentes)

Na Tabela 8.2, também temos a média e variância amostrais, para cada combinação da resposta x assumindo somente dados não-censurados.

Dos resultados da Tabela 8.2, observamos que as variâncias dos tempos de sobrevivência tem uma grande influência da presença da “fragilidade”. Outro ponto importante é que essas diferenças podem ser afetadas pelo tamanho das amostras, para cada uma das classes da covariável resposta sexo x .

Para os modelos apresentados anteriormente, verificamos que os modelos utilizando a distribuição gama generalizada apresentaram melhores resultados do que as modelagens, via distribuição Weibull. Isto comparando os modelos segundo os valores de DIC.

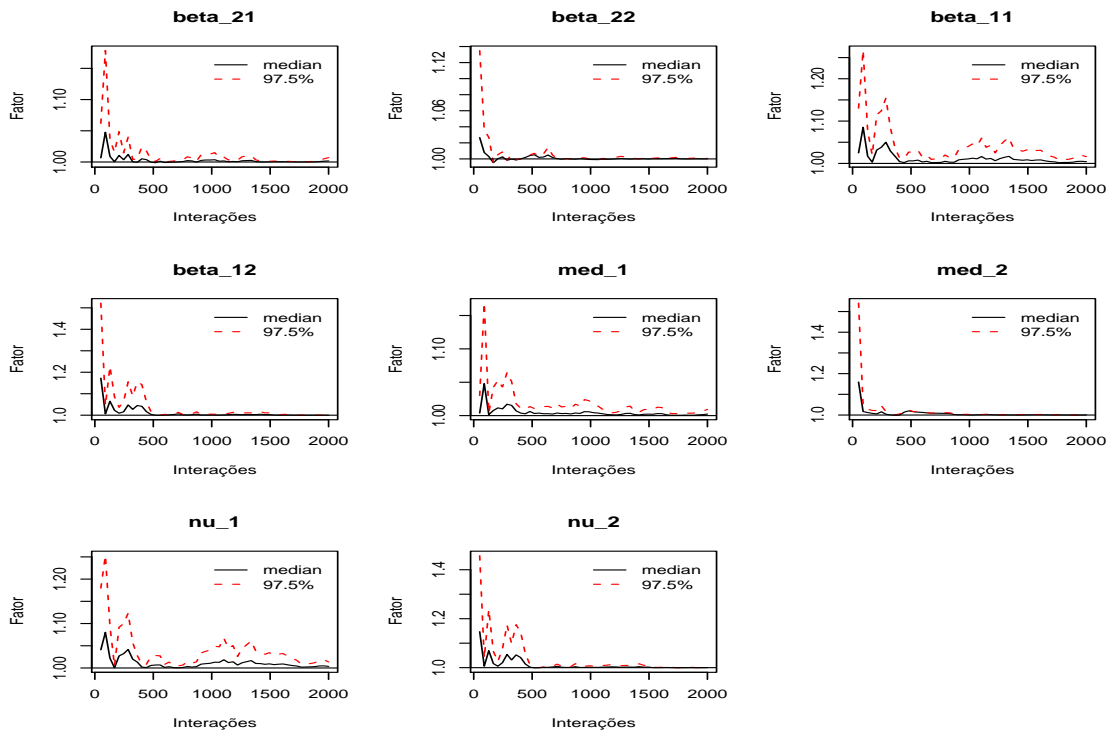


Figura 8.2: Convergência de Gelman-Rubin (modelos Weibull independentes)

Tabela 8.2: Médias e variâncias (“modelo 1” e distribuições Weibull independentes).

	dados sem censura		Weibull independentes	
	Média amostral	Variância amostral	Média	Variância
$(x = 1)$, resp 1	32,80	2052,09	34,360	1338,18
$(x = 0)$, resp 1	162,2	28358,6	186,95	39619,1
$(x = 1)$, resp 2	105,8	36214,1	117,29	14512,9
$(x = 0)$, resp 2	115,9	10609,0	148,36	23243,7
	“modelo 1”		“modelo 1”	
	Média nc	Variância nc	Média cond	Variância cond
$(x = 1)$, resp 1	25,680	565,550	32.240	1308.15
$(x = 0)$, resp 1	86,220	14307,6	179.45	36887.9
$(x = 1)$, resp 2	69,760	3836,03	108.22	13236.3
$(x = 0)$, resp 2	136,51	14658,7	131.98	20640.7

(resp=resposta; nc=não-condicional; cond=condicional; masculino ($x = 1$); feminino ($x = 0$))

No “modelo 5” utilizamos “fragilidade” aditiva enquanto que, para o “modelo 6” foi utilizado “fragilidade” multiplicativa, ambas para modelos Bayesianos hierárquicos.

Já que a literatura apresenta muitas distribuições exponenciais bivariadas diferentes, isto é de grande interesse para pesquisas aplicadas nas áreas médicas, engenharia ou outras áreas, para saber qual modelo é mais apropriado para ser usado nas aplicações. Usualmente, temos grande dificuldades e/ou resultados de inferência pobres, ou seja, não temos inferências exatas, usando inferência assintótica padrão baseada nos métodos de máxima

verossimilhança para essas distribuições bivariadas. O uso de métodos Bayesianos, especialmente considerando a metodologia MCMC e *softwares* existentes disponíveis como o *WinBUGS*, é uma boa alternativa para obtermos as inferências de interesse.

Para a análise Bayesiana do conjunto de dados introduzido na Tabela 5.1 (Liga dos campeões da UEFA), observamos que o modelo de Block & Basu é uma boa alternativa (menor DIC e maior soma dos logaritmos das CPO's). Considerando o menor valor para o DIC, observamos que o modelo de Freund é o melhor modelo para os dados apresentados no capítulo 5; e considerando a maior soma dos logaritmos das CPO's, observamos que o modelo na presença de "fragilidade" se apresenta melhor para os dados. É importante salientar que nas aplicações devemos ter cuidado por termos usado apenas uma técnica Bayesiana de discriminação de modelos, pois os resultados podem ser diferentes usando outros métodos de discriminação.

O uso de métodos MCMC para desenvolver uma análise Bayesiana para a distribuição exponencial bivariada de Block & Basu pode ser um caminho apropriado para analisar dados de sobrevivência bivariados, na presença de dados censurados.

Além disso na presença de dados censurados, usualmente podemos ter, especialmente em aplicações médicas, a presença de uma ou mais covariáveis, uma situação que pode nos trazer grandes dificuldades para estimarmos os parâmetros do modelo pelos métodos usuais de inferência clássica. Usualmente esses procedimentos de inferência clássica são baseados na normalidade assintótica dos estimadores de máxima verossimilhança (veja por exemplo, Lawless, 1982).

Também isso é importante para chamar a atenção ao uso de software Bayesianos disponíveis, tal como o *software WinBUGS*, que requer somente a especificação da distribuição para os dados e as distribuições a priori para os parâmetros do modelo, dando uma boa simplificação para termos os resumos a posteriori de interesse, onde isso foi observado nas aplicações consideradas nesta tese.

Também verificamos que o uso de diferentes funções cópulas pode ser uma boa alternativa para derivar distribuições dos tempos de sobrevivência bivariados. Outras distribuições de sobrevivência marginais também podem ser usadas, para obter distribuições de sobrevivência bivariadas derivadas de funções cópulas, ou mesmo multivariadas com mais de dois tempos de sobrevivência.

8.1 Estudos Futuros

Em estudos futuros estamos interessados em utilizar outras formas de distribuição para modelar esses dados. Uma alternativa será usarmos a distribuição bi-paramétrica exponencial generalizada (GE), que é dada por (ver, Kundu & Gupta, 2008)

$$f(x; \alpha, \lambda) = \alpha\lambda(1 - e^{-\lambda x})^{\alpha-1}e^{-\lambda x}; \quad x > 0 \quad (8.1)$$

onde $\alpha, \lambda > 0$ são os parâmetros de escala e de forma, respectivamente.

Sua função de distribuição é

$$F(x; \alpha, \lambda) = (1 - e^{-\lambda x})^\alpha; \quad \lambda, \alpha, x > 0, \quad (8.2)$$

a função de sobrevivência é

$$S(x; \alpha, \lambda) = 1 - (1 - e^{-\lambda x})^\alpha, \quad (8.3)$$

e a função de risco é

$$h(x; \alpha, \lambda) = \frac{\alpha\lambda(1 - e^{-\lambda x})^{\alpha-1}e^{-\lambda x}}{1 - (1 - e^{-\lambda x})^\alpha} \quad (8.4)$$

Denotamos esta distribuição por $GE(\alpha, \lambda)$, onde notamos que $GE(1, \lambda)$ representa a distribuição exponencial com parâmetro de escala λ . Observar que esta distribuição pode ser usada efetivamente para analisar muitos tipos de dados de sobrevivência, especialmente como uma alternativa para a distribuição gama. A distribuição $GE(\alpha, \lambda)$ pode ter taxas de falhas crescentes ou decrescentes, dependendo do parâmetro de forma, sendo de grande similaridade com a distribuição gama.

Embora esta distribuição tenha sido utilizada, em vários artigos, para casos de amostras completas (veja por exemplo, Gupta & Kundu, 1999), muito pouca atenção tem sido dada para casos de amostras censuradas. Um dos trabalhos que temos sobre dados censurados para a distribuição $GE(\alpha, \lambda)$ é o de Mitra & Kundu (2008).

Outra distribuição que pretendemos explorar é a distribuição BGE (Exponencial Beta Generalizada) (ver, Souza *et al.*, 2008), onde a distribuição GE é um caso especial. Sua densidade pode ser expressa como uma mistura de densidades exponenciais generalizadas, que é importante para obter as propriedades matemáticas desta distribuição Exponencial Beta Generalizada.

Sua função densidade de probabilidade é dada por

$$f(x) = \frac{\alpha\lambda}{B(a,b)} e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha a - 1} \{1 - (1 - e^{-\lambda x})^\alpha\}^{b-1}, \quad x > 0 \quad (8.5)$$

para $a > 0$, $b > 0$, $\lambda > 0$ e $\alpha > 0$. A função de risco é dada por

$$h(x) = \frac{\alpha\lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\alpha a - 1} \{1 - (1 - e^{-\lambda x})^\alpha\}^{b-1}}{B(a,b) I_{1-(1-e^{-\lambda x})^\alpha}(a,b)}, \quad x > 0 \quad (8.6)$$

A função de densidade (8.5) não envolve nenhuma função complicada e pode ser denotada por $X \sim \text{BGE}(a, b, \lambda, \alpha)$. A distribuição GE é um caso especial quando tomamos $a = b = 1$; se além disso consideramos $\alpha = 1$, obtemos uma distribuição exponencial com parâmetro λ .

Estudos de simulação para os modelos apresentados nesta tese também são objetos de futuras pesquisas, uma vez que eles podem dar mais consistência aos modelos apresentados. Outros métodos de discriminação de modelos Bayesianos podem ser implementados, dando novo suporte aos resultados já alcançados.

Referências Bibliográficas

- Achcar, J. A. & Leandro, R. A. (1998). Use of Markov Chain Monte Carlo methods in a Bayesian analysis of the Block and Basu bivariate exponential distribution. *Annals of the Institute of Statistical Mathematics*, **50**(3), 403–416.
- Achcar, J. A., Louzada-Neto, F. & Mazucheli, J. (2002). *Introdução à Análise de Sobre-vivência e Confiabilidade*. Minicurso: III Jornada Regional de Estatística e II Semana de Estatística, Maringá.
- Alen, O. O. (1980). A Model for Nonparametric Regression Analysis of Counting Process. *Lecture Notes in Statistics*.
- Andersen, P. K., Borgan, O., Gill, R. D. & Keiding, N. N. (1993). *Statistical Models Based on Counting Process*. Springer-Verlag, New York.
- Arnold, B. C. & Strauss, D. (1988). Bivariate distributions with exponential conditionals. *Journal of the American Statistical Association*, **83**(402), 522–527.
- Aslanidou, H., Dey, D. K. & Sinha, D. (2004). Bayesian analysis of multivariate survival data using Monte Carlo Methods. Technical report, Department of Statistics, University of Connecticut. Submitted to *Metron*.
- Best, N. and Cowles, M. and Vines, K. (1997). *CODA: Convergence Diagnosis and Output Analysis Software for Gibbs Sampling Output*. Cambridge, UK.
- Block, H. W. & Basu, A. P. (1974). A continuous bivariate exponential extension. *Journal of the American Statistical Association*, (69), 1031–1037.
- Brooks, S. & Roberts, G. (1995). Diagnosing convergence of Markov Chain Monte Carlo algorithms. Technical report, Department of Maths and Mathematical Statistics, University of Cambridge. Submitted to *Metron*.
- Buckley, T. & James, I. (1979). Linear regression with censored data. *Biometrika*, (66), 609–615.

- Carlin, B. P. & Louis, T. A. (1997). *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman Hall, London.
- Casella, G. & George, E. I. (1992). Explaining the Gibbs sampler. *The American Statistician*, (46), 167–174.
- Chib, S. & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, (49), 327–335.
- Clayton, D. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, (65), 141–151.
- Clayton, D. (1991). A Monte carlo methods for Bayesian inference in frailty models. *Biometrics*, (47), 467–485.
- Clayton, D. & Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society*, **A**(148), 82–117.
- Colosimo, E. A. & Giolo, S. R. (2006). *Análise de Sobrevida Aplicada*. Projeto Fisher - ABE. Edgard Blücher Ltda., São Paulo.
- Cook, R. (1995). The design and analysis of randomized trials with recurrent events. *Statistics in Medicine*, (14), 2081–2098.
- Cook, R. & Lawless, J. (1997). Marginal analysis of recurrence events and a terminating event. *Statistics in Medicine*, (16), 911–924.
- Cowles, M. K. & Carlin, B. P. (1996). Markov Chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association*, (91), 883–904.
- Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society*, **B**(34), 187–220.
- Cox, D. R. & Oakes, D. (1984). *Analysis of Survival Data*. Chapman & Hall, London.
- da Silva, G. L. (2001). *Análise Bayesiana de Modelos de Sobrevida com Fragilidade*. Tese de doutorado, Instituto Superior Técnico – Universidade Técnica de Lisboa, Lisboa.
- da Silva Matuda, N. (2005). *Fragilidade Gama e Variância Robusta: Extensões do Modelo Semiparamétrico de Cox*. Tese de doutorado, FSP – Universidade de São Paulo, São Paulo.

- Dias, T. C. M. (2002). *Análise Bayesiana para Dados de Sobrevivência em Modelos de Riscos Proporcionais*. Tese de doutorado, COPPE – Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- Doganaksoy, N. & Nelson, W. (1998). A method to compare two samples of recurrence data. *Lifetime Data Analysis*, (4), 51–63.
- Downton, F. (1970). Bivariate exponential distributions in reliability theory. *Journal of the Royal Statistical Society*, **B**(32), 408–417.
- Farlie, D. J. G. (1960). The performance of some correlation coefficients for a general bivariate distribution. *Biometrika*, (47), 307–323.
- Feigl, P. & Zelen, M. (1965). Estimation of exponential survival probabilities with concomitant information. *Biometrics*, (21), 826–837.
- Fleming, T. R. & Lin, D. Y. (2000). Survival analysis in clinical trials: past developments and future directions. *Biometrics*, (56), 971–983.
- Flores, A. Q. (2008). *Copula Functions and Bivariate Distributions for Survival Analysis: an Application to Political Survival*. manuscript, Wilf Family Department of Politics, New York University.
- Flores, A. Q. (2009). *Testing Copula Functions as a Method to Derive Bivariate Weibull Distributions*. manuscript, Wilf Family Department of Politics, New York University.
- Freund, J. E. (1961). A bivariate extension of the exponential distribution. *Journal of the American Statistical Association*, (56), 971–977.
- Gamerman, D. (1991). Dynamic Bayesian models for survival data. *Applied Statistics*, (40), 63–79.
- Gamerman, D. (1997). *Markov Chain Monte Carlo - Stochastic Simulation for Bayesian Inference*. Chapman & Hall, London.
- Gelfand, A. E. & Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, (85), 398–409.
- Gelman, A. & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, **7**(4), 457–472.
- Gelman, R., Carlin, J. B., Stern, H. S. & Rubin, D. B. (1995). *Bayesian Data Analysis*. Chapman & Hall, London.

- 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.
- Gumbel, E. J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association*, (55), 698–707.
- Gupta, R. D. & Kundu, D. (1999). Generalized Exponential Distributions. *Australian and New Zealand Journal Statistical*, **41**(2), 173–188.
- Hager, H. W. & Bain, L. J. (1970). Inferential procedures for the generalized gamma distribution. *Journal of the American Statistical Association*, (65), 1601–1609.
- Hanagal, D. D. (2008). Modeling heterogeneity for bivariate survival data by the Weibull distribution. *Statistical Papers*, (DO 10.1007/s 00362-008-0188-2).
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov Chains and their applications. *Biometrika*, (57), 97–109.
- Hawkes, A. G. (1972). A bivariate exponential distribution with applications to reliability. *Journal of the Royal Statistical Society*, **B**(34), 129–131.
- Heidelberger, W. (1983). Simulation run length control in the presence of an initial transient. *Operations Research*, (31), 1109–1144.
- Hougaard, P. (1986a). Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, **3**(73), 387–396.
- Hougaard, P. (1986b). A class of multivariate failure time distributions. *Biometrika*, **3**(73), 671–678.
- Hougaard, P. (1995). Frailty models for survival data. *Lifetime Data Analysis*, (1), 255–273.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer-Verlag, New York.
- Huster, W. J., Brookmeyer, R. & Self, S. G. (1989). Modeling paired survival data with covariates. *Biometrics*, (45), 145–156.
- Kalbfleisch, J. D. (1978). Nonparametric Bayesian analysis of survival time data. *Journal of the Royal Statistical Society*, **B**(40), 214–221.
- Kalbfleisch, J. D. & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley, New York.

- Klein, J. P. & Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York.
- Kundu, D. & Gupta, R. D. (2008). Generalized Exponential Distribution: Bayesian Estimations. Technical report, Department of Mathematics and Statistics, Indian Institute of Technology Kanpur. Submitted to *Metron*.
- Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley, New York.
- Lawless, J. F. & Nadeau, C. (1995). Some simple robust methods for the analysis of recurrent events. *Technometrics*, (37), 158–168.
- Lee, E. T. & Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis*. John Wiley and Sons, New Jersey.
- Lee, J., Kin, J. & Jung, S. H. (2007). Bayesian analysis of paired survival data using a bivariate exponential distribution. *Lifetime Data Analysis*, **1**(13), 119–137.
- Lin, D. Y. & Ying, Z. (1995). Semiparametric analysis of general additive-multiplicative hazards models for counting processes. *The Annals of Statistics*, (23), 1712–1734.
- Marshall, A. W. & Olkin, I. (1967a). A generalized bivariate exponential distribution. *Journal of Applied Probability*, (4), 291–302.
- Marshall, A. W. & Olkin, I. (1967b). A multivariate exponential distribution. *Journal of the American Statistical Association*, (62), 30–44.
- McGilchrist, C. A. & Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics*, (47), 461–466.
- Meintanis, S. G. (2007). Test of fit for Marshall-Olkin distributions with applications. *Journal of Statistical Planning and Inference*, (137), 3954–3963.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). Equations of state calculations by fast computing machine. *Journal of Chemical Physics*, (21), 1087–1091.
- Miller, R. G. (1976). Least squares regression with censored data. *Biometrika*, **63**, 449–464.
- Mitra, S. & Kundu, D. (2008). Analysis of left censored data from the generalized exponential distribution. *Journal of Statistical Computation and Simulation*, **78**(7), 669–679.

- Morgenstern, D. (1956). Einfache Beispiele Zweidimensionaler Verteilungen. *Mitteilungsblatt für Mathematische Statistik*, **8**, 234–253.
- Nelsen, R. B. (1999). *An Introduction to Copulas*. Springer-Verlag, New York.
- Ng, E. & Cook, R. (1997). Modeling two-state disease processes with random effects. *Lifetime Data Analysis*, **3**, 315–335.
- Oakes, D. (1986). Semiparametric inference in a model for association in bivariate survival data. *Biometrika*, **73**, 353–361.
- Oakes, D. (1989). Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, **84**, 487–493.
- Parr, V. B. & Webster, J. T. (1965). A method for discriminating between failure density functions used in reliability predictions. *Technometrics*, (7), 1–10.
- Paulino, C. D., Turkman, M. A. A. & Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.
- Pepe, M. S. & Cai, J. (1993). Some graphical displays and marginal regression analyses for recurrent failure times and time dependent covariates. *Journal of the American Statistical Association*, (88).
- Peto, R., Lee, P. N. & Paige, W. S. (1972). Statistical analysis of the biossay of continuous carcinogenes. *Breast Cancer*, (26), 258–261.
- Pickles, A. & Crouchley, R. (1995). A comparison of frailty models for multivariate survival data. *Statistic Medicine*, (14), 1447–1461.
- Raftery, A. L. & Lewis, S. (1992). How many iterations in the Gibbs sampler? *Bayesian Statistics*, (4), 763–774.
- Sarkar, S. K. (1987). A continuous bivariate exponential distribution. *Journal of the American Statistical Association*, (82), 667–675.
- Self, S. G. & Prentice, R. L. (1986). *Modern Statistical Methods in Chronic Disease Epidemiology*. John Wiley, New York.
- Shih, J. A. & Louis, T. A. (1992). Models and analysis for multivariate failure time data. Technical report, Division of Biostatistics, University of Minnesota. Submitted to *Metron*.

- Sinha, D. (1998). Semiparametric Bayesian analysis of multiple event time data. *Journal of the American Statistical Association*, (88), 979–983.
- Sklar, A. (1959). Fonctions de repartition à n-dimensions et leurs marges. *Inst. Stat. University Paris*, (8), 229–231.
- Souza, B. W., Santos, A. H. S. & Cordeiro, G. M. (2008). The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*, (0), 1–15.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society*, **B**(64), 583–639.
- Spiegelhalter, D. J.; Thomas, A.; Best, N. G.; Gilks, W. R. (2003). *WinBUGS User Manual (version 1.4)*. MRC Biostatistics Unit, Cambridge, UK.
- Therneau, T. M. & Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York.
- Tierney, L., Kass, R. E. & Kadane, J. B. (1989). Aproximate marginal densities for nonlinear functions. *Biometrika*, (76), 425–433.
- Tomazella, V. L. D. (2003). *Modelagem de Dados de Eventos Recorrentes via Processo de Poisson com Termo de Fragilidade*. Tese de doutorado, ICMC – Universidade de São Paulo, São Carlos.
- Trivedi, P. K. & Zimmer, D. M. (2005a). *Copula Modelling*. New Publishers, New York.
- Trivedi, P. K. & Zimmer, D. M. (2005b). Copula modelling: An introduction to practitioners. *Foundations and trends in econometrics*, **1**(1), 1–111.
- Turnbull, B., Jiang, W. & Clark, L. (1997). Regression model for recurrent event data: Parametric random effects models with measurement error. *Statistics in Medicine*, (16), 853–864.
- Vaupel, J. W., Manton, K. G. & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, (16), 439–454.
- Wei, L., Lin, D. & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, **84**(408), 1065–1073.
- Weibull, W. (1951). A Statistical distribution function of wide applicability. *Journal of Applied Mechanics*, pages 292–297.

Whittmore, A. & Altschuler, B. (1976). Lung cancer incidence in cigarette smokers: Further analysis of Doll and Hill's data for British physicians. *Biometrics*, (32), 805–816.

Apêndice A

Medidas condicionais e não-condicionais para a distribuição Gama

Seja W_i uma variável aleatória com uma distribuição Gama(γ, ν), com densidade,

$$f(w_i|\gamma, \nu) = \frac{\nu^\gamma}{\Gamma(\gamma)} w_i^{\gamma-1} e^{-\nu w_i} \quad (\text{A.1})$$

onde $w_i > 0$, $i = 1, 2, \dots, n$.

Da equação (A.1) observamos que,

$$\int_0^\infty w_i^{\gamma-1} e^{-\nu w_i} dw_i = \frac{\Gamma(\gamma)}{\nu^\gamma} \quad (\text{A.2})$$

A função de sobrevivência é dada por

$$S(w) = \int_w^\infty \frac{\nu^\gamma}{\Gamma(\gamma)} x^{\gamma-1} e^{-\nu x} dx \quad (\text{A.3})$$

Também observe que,

$$E(W_i^{-k}) = \int_0^\infty w_i^{-k} \frac{\nu^\gamma}{\Gamma(\gamma)} w_i^{\gamma-1} e^{-\nu w_i} dw_i = \frac{\nu^\gamma}{\Gamma(\gamma)} \int_0^\infty w_i^{(\gamma-k)-1} e^{-\nu w_i} dw_i$$

De (A.2), temos:

$$E(W_i^{-k}) = \frac{\nu^k \Gamma(\gamma - k)}{\Gamma(\gamma)} \quad (\text{A.4})$$

para $\gamma > k$.

Assumindo $\gamma = v = \phi^{-1}$, temos:

(i) Com $k = 1/\nu_j$,

$$E(W_i^{-1/\nu_j}) = \frac{(\phi^{-1})^{\nu_j^{-1}} \Gamma(\phi^{-1} - \nu_j^{-1})}{\Gamma(\phi^{-1})} \quad (\text{A.5})$$

para $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$;

(ii) Com $k = 2/\nu_j$,

$$E(W_i^{-2/\nu_j}) = \frac{(\phi^{-1})^{2/\nu_j} \Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} \quad (\text{A.6})$$

para $i = 1, 2, \dots, n$; $j = 1, 2, \dots, k$.

Temos ainda que

$$W_i \stackrel{\text{iid}}{\sim} \text{Gama}(\phi^{-1}, \phi^{-1}) \quad (\text{A.7})$$

De (A.7), temos

$$E(T_{ji}|\mathbf{x}_i) = E[E(T_{ji}|\mathbf{x}_i, w_i)] = \frac{\Gamma(1 + 1/\nu_j)}{\exp\{\boldsymbol{\beta}'_j \mathbf{x}_i / \nu_j\}} E(W_i^{-1/\nu_j})$$

Então, de (A.5), encontramos a média não-condicional para T_{ji} dada por,

$$E(T_{ji}|\mathbf{x}_i) = \frac{\Gamma(1 + 1/\nu_j)(\phi^{-1})^{1/\nu_j} \Gamma(\phi^{-1} - 1/\nu_j)}{\Gamma(\phi^{-1}) \exp\{\boldsymbol{\beta}'_j \mathbf{x}_i / \nu_j\}} \quad (\text{A.8})$$

De (A.7) e (A.8), e usando o resultado $\text{Var}(T_{ji}|\mathbf{x}_i) = E\{\text{Var}(T_{ji}|\mathbf{x}_i, w_i)\} + \text{Var}\{E(T_{ji}|\mathbf{x}_i, w_i)\}$, temos:

$$\begin{aligned} \text{Var}(T_{ji}|\mathbf{x}_i) &= \frac{[\Gamma(1 + 2/\nu_j) - \Gamma^2(1 + 1/\nu_j)]}{\exp\{2\boldsymbol{\beta}'_j \mathbf{x}_i / \nu_j\}} E(W_i^{-2/\nu_j}) + \\ &\quad + \frac{\Gamma^2(1 + 1/\nu_j)}{\exp\{2\boldsymbol{\beta}'_j \mathbf{x}_i / \nu_j\}} \text{Var}(W_i^{-1/\nu_j}) \end{aligned}$$

Observe que, $\text{Var}(W_i^{-1/\nu_j}) = E(W_i^{-2/\nu_j}) - [E(W_i^{-1/\nu_j})]^2$, ou seja, de (A.5) e (A.6),

$$\text{Var}(W_i^{-1/\nu_j}) = \frac{(\phi^{-1})^{2/\nu_j} \Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} - \frac{(\phi^{-1})^{2/\nu_j} \Gamma^2(\phi^{-1} - 1/\nu_j)}{\Gamma^2(\phi^{-1})} \quad (\text{A.9})$$

Também de (A.6) e (A.8), encontramos a variância não-condicional para T_{ji} dada por,

$$\begin{aligned} \text{Var}(T_{ji}|\mathbf{x}_i) &= \frac{(\phi^{-1})^{2/\nu_j}}{\exp(2\boldsymbol{\beta}'_j \mathbf{x}_i / \nu_j)} \times \\ &\times \left\{ \frac{\Gamma(1 + 2/\nu_j) \Gamma(\phi^{-1} - 2/\nu_j)}{\Gamma(\phi^{-1})} - \left[\frac{\Gamma(1 + 1/\nu_j) \Gamma(\phi^{-1} - 1/\nu_j)}{\Gamma(\phi^{-1})} \right]^2 \right\} \end{aligned}$$

Apêndice B

Alguns Programas Utilizados

Alguns dos programas, in *WinBUGS* (Spiegelhalter, D. J.; Thomas, A.; Best, N. G.; Gilks, W. R., 2003), apresentados na tese.

B 1. Modelo 1 - Distribuição Weibull, com dependência.

```
model {
  for(i in 1 : M) {
    for(j in 1 : N) {
      t[i, j] ~ dweib(r[i], mu[i,j])I(t.cen[i, j],)
      mu[i,j] <- exp(beta[i]+alpha[i]*x[j]+w[j])
    }
    alpha[i] ~ dnorm(0.0, 0.1)
    beta[i] ~ dnorm(0.0, 0.1)
    median[i] <- pow(log(2) * exp(-beta[i]), 1/r[i])
  }

  for(j in 1 : N) {
    w[j] ~ dnorm(0, tau)
  }

  for(i in 1 : M) {
    r [i] ~ dgamma(1, 1)
  }

  tau ~ dgamma(0.1, 0.1)
```



```

    }
Dados

list(t = structure(.Data =
  c(8, 23, 22, 447, 30, 24, 7, 511, 53, 15, 7,
    141, 96, NA, 536, 17, 185, 292, NA, 15, 152,
    402, 13, 39, 12, NA, 132, 34, 2, 130, 27, NA,
    152, 190, 119, NA, NA, 63, 16, NA, 28, 318, 12,
    245, 9, 30, 196, 154, 333, NA, 38, NA, NA, NA,
    117, 114, NA, NA, 562, NA, 66, NA, 40, 201, 156,
    30, 25, 26, 58, 43, 30, NA, 8, NA, 78, NA),
  .Dim = c(2, 38)),
t.cen = structure(.Data =
  c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 149,
    0, 0, 0, 0, 22, 0, 0, 0, 0, 0, 0, 113, 0, 0,
    0, 0, 0, 5, 0, 0, 0, 54, 6, 0, 0, 13, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 8, 0, 70, 25, 4, 0, 0,
    159, 108, 0, 24, 0, 46, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 5, 0, 16, 0, 8), .Dim = c(2, 38)), M = 2,
N = 38, x=c(1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0,
  0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0,
  0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1))

```

B 2. Modelo 4 - Distribuição Weibull (Covariável afetando o parâmetro de forma).

```

model {
  for(i in 1 : M) {
    for(j in 1 : N) {
      t[i, j] ~ dweib(r[i, j], mu[i, j])I(t.cen[i, j],)
      mu[i, j] <- w[j]*exp(beta1[i] + alpha1[i]*x[j])
      r[i, j] <- exp(beta2[i] + alpha2[i]*x[j])
      median[i, j] <- pow(log(2) * exp(-beta1[i]), 1/r[i, j])
    }
    alpha1[i] ~ dnorm(2, 10)
    alpha2[i] ~ dnorm(0, 10)
  }
}

```

```

        beta1[i] ~ dnorm(-6, 10)
        beta2[i] ~ dnorm(0, 10)
    }

    for(j in 1 : N) {
        w[j] ~ dgamma(eta, eta)
    }

    eta <- 1/phi
    phi ~ dunif(0, 10)
}

Dados

list(t = structure(.Data =
    c(8, 23, 22, 447, 30, 24, 7, 511, 53, 15, 7,
    141, 96, NA, 536, 17, 185, 292, NA, 15, 152,
    402, 13, 39, 12, NA, 132, 34, 2, 130, 27, NA,
    152, 190, 119, NA, NA, 63, 16, NA, 28, 318, 12,
    245, 9, 30, 196, 154, 333, NA, 38, NA, NA, NA,
    117, 114, NA, NA, 562, NA, 66, NA, 40, 201, 156,
    30, 25, 26, 58, 43, 30, NA, 8, NA, 78, NA),
    .Dim = c(2, 38)),
t.cen = structure(.Data =
    c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 149,
    0, 0, 0, 0, 22, 0, 0, 0, 0, 0, 0, 0, 113, 0, 0,
    0, 0, 0, 5, 0, 0, 0, 54, 6, 0, 0, 13, 0, 0,
    0, 0, 0, 0, 0, 0, 0, 8, 0, 70, 25, 4, 0, 0,
    159, 108, 0, 24, 0, 46, 0, 0, 0, 0, 0, 0, 0,
    0, 0, 5, 0, 16, 0, 8), .Dim = c(2, 38)), M = 2,
N = 38, x=c(1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0,
0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0,
0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1))

```

B 3. Modelo 5 - Distribuição Gama Generalizada.

```

model {
    for(i in 1 : M) {

```

```

for(j in 1 : N) {
t[i, j] ~ gen.gamma(r[i], mu[i, j], delta[i])I(t.cen[i, j],)
mu[i, j] <- exp(beta[i] + alpha[i]*x[j] + w[j])
      }
      beta[i] ~ dnorm(-5, 1)
      delta[i] ~ dgamma(1, 1)
      alpha[i] ~ dnorm(1, 1)
      r[i] ~ dgamma(1, 1)
}

      for(j in 1 : N) {
      w[j] ~ dnorm(0, tau)
      }
      tau ~ dgamma(1, 1)
}

Dados

list(t = structure(.Data =
      c(8, 23, 22, 447, 30, 24, 7, 511, 53, 15, 7,
      141, 96, NA, 536, 17, 185, 292, NA, 15, 152,
      402, 13, 39, 12, NA, 132, 34, 2, 130, 27, NA,
      152, 190, 119, NA, NA, 63, 16, NA, 28, 318, 12,
      245, 9, 30, 196, 154, 333, NA, 38, NA, NA, NA,
      117, 114, NA, NA, 562, NA, 66, NA, 40, 201, 156,
      30, 25, 26, 58, 43, 30, NA, 8, NA, 78, NA),
      .Dim = c(2, 38)),
t.cen = structure(.Data =
      c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 149,
      0, 0, 0, 0, 22, 0, 0, 0, 0, 0, 0, 113, 0, 0,
      0, 0, 0, 5, 0, 0, 0, 54, 6, 0, 0, 13, 0, 0,
      0, 0, 0, 0, 0, 0, 0, 8, 0, 70, 25, 4, 0, 0,
      159, 108, 0, 24, 0, 46, 0, 0, 0, 0, 0, 0, 0,
      0, 0, 5, 0, 16, 0, 8), .Dim = c(2, 38)), M = 2,
      N = 38, x=c(1, 0, 1, 0, 1, 0, 1, 0, 0, 1, 0,
      0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0,
      0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1))

```

B 4. Distribuição de Block & Basu.

```

model{
  for(i in 1 : n) {
    zeros[i] <- 0
    phi[i] <- -log(L[i])
    zeros[i] ~ dpois(phi[i])
    a1[i]<- (((lambda1+lambda2+lambda3)*lambda1*(lambda2+lambda3))/
(lambda1+lambda2))*exp(-lambda1*t1[i]-(lambda2+lambda3)*t2[i])
    a2[i]<- pow(a1[i],delta[i])
    b1[i]<-(((lambda1+lambda2+lambda3)*lambda2*(lambda1+lambda3))/
(lambda1+lambda2))*exp(-(lambda1+lambda3)*t1[i]-lambda2*t2[i])
    b2[i]<- pow(b1[i],1-delta[i])
    L[i]<- exp(log(a2[i]*b2[i]))
  }
  lambda1 ~ dgamma(0.1, 0.1)
  lambda2 ~ dgamma(0.1, 0.1)
  lambda3 ~ dgamma(0.1, 0.1)
}

Data

list(t1=c(26,63,19,66,40,49,8,69,39,82,72,66,25,41,16,18,22,42,36,
34,53,54,51,76,64,26,16,44,25,55,49,24,44,42,27,28,2),
t2=c(20,18,19,85,40,49,8,71,39,48,72,62,9,3,75,18,14,42,52,
34,39,7,28,64,15,48,16,13,14,11,49,24,30,3,47,28,2),
delta=c(0,0,0,1,0,0,0,1,0,0,0,0,0,0,1,0,0,0,1,
0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,1,0,0), n = 37)

```