

UNIVERSIDADE FEDERAL DE SÃO CARLOS
Centro de Ciências Exatas e de Tecnologia
Programa de Pós Graduação em Estatística
Departamento de Estatística

USO DO PROCESSO GAMA PARA DADOS DE
SOBREVIVÊNCIA

Shirley Tieko Uêda

Orientador: Prof. Dr. Jorge Alberto Achcar

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

São Carlos - SP
Janeiro - 2005

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

U22up

Uêda, Shirley Tieko.

Uso do processo gama para dados de sobrevivência /
Shirley Tieko Uêda. -- São Carlos : UFSCar, 2005.
115 p.

Dissertação (Mestrado) -- Universidade Federal de São
Carlos, 2005.

1. Estatística matemática. 2. Processo gama. 3. Análise
de sobrevivência (biometria). 4. Análise bayesiana
hierárquica. 5. Algoritmos MCMC. 6. Confiabilidade. I. Título.

CDD: 519.5 (20^a)

AMOR

Ainda que eu fale a língua dos homens e dos anjos, se não tiver Amor, serei como o bronze que soa ou como o címbalo que retine.

E ainda que eu distribuísse todos os meus bens entre os pobres, ainda que entregasse meu corpo para ser queimado, se não tiver Amor nada disso me adiantaria.

O Amor é paciente e bondoso.

O Amor não é ciumento, nem orgulhoso, nem vaidoso, nem egoísta. Não se alegra com a injustiça, mas se alegra com a verdade.

Tudo sofre, tudo crê, tudo espera, tudo suporta.

O Amor jamais acaba; havendo profecias, desaparecerão; havendo línguas cessarão; havendo ciências, passará.

Agora, pois permanecem a fé, a esperança e o Amor.

Porém o maior destes é o Amor.

1 CORÍNTIOS 13: 1-3-4-6-7-8-13

*Dedico este trabalho aos meus pais Lourenço Seiko Uêda e
Luzia Pimentel Uêda por estarem sempre ao meu lado em todos
os momentos difíceis e aos meus irmãos Eduardo pelo apoio e
incentivo e a Herbert (in memoriam) que tornou-se
luz em minha vida e da minha família.*

Agradecimentos

Primeiramente a DEUS e a Nossa Senhora Aparecida que sempre guiaram e iluminaram o meu caminho, dando-me paciência, sabedoria e força para superar as dificuldades encontradas.

Aos meus pais, Lourenço e Luzia, que nunca deixaram de me apoiar com vosso carinho e estímulo, mesmo nos momentos mais difíceis de nossas vidas, estando eles sempre ao meu lado dando condições físicas, financeiras e psicológicas para mais uma conquista.

Aos meus irmãos Eduardo por sempre estar ao meu lado e a Herbert (*in memoriam*) pelo incentivo ao meu ingresso na Pós-Graduação.

Ao meu orientador, Prof. Dr. Jorge Alberto Achcar, pela dedicação de transmitir os ensinamentos para realização deste trabalho, pela amizade, pelo convívio durante este período que tive a oportunidade de ter, pelo incentivo a pesquisa e principalmente por seu exemplo de dedicação e disciplina ao trabalho.

A Profa. Dra. Teresa Cristina Martins Dias, pela colaboração, pelo apoio computacional, pelo trabalho em equipe, pela amizade, pelas críticas e sugestões para o desenvolvimento deste trabalho.

Ao Prof. Dr. Francisco Louzada Neto e Prof. Dr. Francisco Rojas Rojas pelas críticas e sugestões dadas no exame de qualificação.

Ao Prof. Dr. Josmar Mazucheli (UEM) pelas discussões e sugestões sobre o tema em questão.

Aos professores do Departamento de Estatística da UFSCar, em especial, Prof. Luis Aparecido Milan, Prof. José Galvão Leite, Prof. Carlos Diniz e Prof. Josemar Rodrigues.

Ao secretário da pós-graduação, Robson, pela paciência e disposição em sempre nos ajudar.

Aos amigos do PPG-DES pelos momentos inesquecíveis que passamos juntos em nosso ambiente de trabalho.

Aos amigos de São Carlos pela amizade, apoio e carinho, em especial a: Carol, Cesar, Claudinei, Eliza, Heverton, Marcelo Ikeda, Marcel, Márcia, Miriam, Nivalda, Silvia e Zé Paulo.

À Coordenação de Aperfeiçoamento Pessoal de Nível Superior - CAPES pelo auxílio financeiro.

Enfim, agradeço a todos aqueles que direta ou indiretamente acompanharam o desenvolvimento deste trabalho.

Resumo

Nesta dissertação, apresentamos uma análise clássica e uma abordagem Bayesiana para obter inferências dos parâmetros de interesse, considerando as distribuições exponencial e de Weibull para os tempos de sobrevivência. Assumimos um processo gama para as taxas individuais e a presença de covariadas relacionadas com os tempos de sobrevivência com censuras do tipo II.

Alguns conceitos e resultados de análise estatística de testes de vida acelerados são apresentados, em particular, um estudo sobre o Modelo de Lei de Potência Inversa, considerando que os tempos de sobrevivência são ajustados por uma distribuição exponencial.

A metodologia proposta neste trabalho está ilustrada a três conjuntos de dados, onde dois são referentes à análise de sobrevivência com dados médicos e o outro a dados de confiabilidade.

Abstract

In this dissertation, we introduce classical and Bayesian approaches to get inferences on the parameters of interest, considering exponential and Weibull distributions for the lifetimes. For a Bayesian analysis, we assume a gamma process for the individual rates considering type II censoring data and the presence of covariates. We also consider accelerated life tests assuming an inverse power law model and an exponential distribution for the lifetimes.

The proposed methodology is illustrated in three examples.

Sumário

1	Introdução	1
2	Uma Abordagem Clássica do Modelo de Regressão Exponencial para Dados de Sobrevivência	4
2.1	Modelo Probabilístico	5
2.2	Modelagem de Dados de Sobrevivência na Presença de Covariáveis	7
2.2.1	Modelo Linear	8
2.2.2	Modelo de Riscos Proporcionais de Cox	10
2.3	Ajustando o Modelo	13
2.3.1	Função de Verossimilhança	13
2.3.2	Estimação dos Parâmetros do Modelo	16
2.3.3	Precisão das Estimativas e Intervalos de Confiança	17
3	Uso do Processo Gama e Modelos de Fragilidade para a Função de Risco do Modelo Semi-Paramétrico de Cox	19
3.1	O Processo Gama	20
3.1.1	Uso do Processo Gama Sobre o Risco Acumulado	20
3.1.2	Uso do Processo Gama para a Verossimilhança com Dados-Agrupados	21
3.1.3	Relação com a Verossimilhança Parcial	24
3.1.4	Uso do Processo Gama Sobre o Risco de Base	27
3.2	Modelos de Fragilidade	29

4	Uma Abordagem Bayesiana do Modelo de Regressão Exponencial Para Dados de Sobrevida	32
4.1	Modelo Probabilístico	33
4.2	Formulação Bayesiana Usando Métodos MCMC	35
4.3	Distribuições Condicionais <i>a Posteriori</i>	40
4.4	Seleção de Modelos	41
5	Uso do Processo Gama Para Dados de Sobrevida Assumindo uma Distribuição Weibull: Uma Abordagem Bayesiana	43
5.1	Modelo Estatístico	44
5.2	Formulação Bayesiana Usando Métodos MCMC	45
5.3	Distribuições Condicionais <i>a Posteriori</i>	48
6	Aplicações	50
6.1	Aplicação com Dados Não-Censurados	51
6.1.1	Método Clássico para o caso exponencial	52
6.1.2	Método Bayesiano para o caso exponencial	54
6.1.3	Método Bayesiano para o caso Weibull	65
6.2	Aplicação com Dados Censurados	69
6.2.1	Método Clássico para o caso exponencial	72
6.2.2	Método Bayesiano para o caso exponencial	74
7	Conceitos Básicos de Testes Acelerados em Controle de Qualidade	79
7.1	Variáveis de Estresse e Tipos de testes acelerados	80
7.2	Formas de aceleração de testes	81
7.3	Formas de Aplicação da Carga	83
7.4	Considerações Importantes para a Realização de um Teste Acelerado	86
8	Relações Estresse-Resposta: Arrhenius, Eyring e Potência Inversa	90
8.1	Relação de Arrhenius	91

8.2	Relação de Eyring	94
8.3	Relação de Potência Inversa	96
9	Uma Análise Bayesiana do Modelo de Potência Inversa Assumindo uma Distribuição Exponencial	98
9.1	Modelo Estatístico	99
9.2	Abordagem Bayesiana do Modelo	100
9.3	Distribuições Condicionais <i>a Posteriori</i> para λ, β_0 e β_1	102
9.4	Introdução de Efeitos Aleatórios no Modelo	102
9.5	Distribuições Condicionais <i>a Posteriori</i> para $\boldsymbol{\lambda}, \mathbf{w}, \beta_0, \beta_1$ e σ_w^2	104
9.6	Aplicação	106
10	Considerações Futuras	109
	Referências Bibliográficas	111

Capítulo 2

Uma Abordagem Clássica do Modelo de Regressão Exponencial para Dados de Sobrevivência

Neste capítulo desenvolvemos uma abordagem clássica do modelo de regressão exponencial com uma variável que descreve o tempo de vida, e assumimos a presença de observações censuradas e não-censuradas. Para realização deste estudo consideramos técnicas usuais relacionadas aos modelos paramétricos e semi-paramétricos (Cox, 1972; Green e Yandell, 1985), onde utilizamos o método de máxima verossimilhança para estimar os parâmetros do modelo.

Através do método de máxima verossimilhança é possível obtermos os intervalos de confiança assintóticos ou exatos para os estimadores pontuais com o objetivo de obtermos inferências por intervalo para os parâmetros do modelo.

As técnicas que apresentamos aqui serão aplicadas a dois conjuntos de dados reais, onde um dos conjuntos assume a presença de observações censuradas e o outro observações não-censuradas.

2.1 Modelo Probabilístico

Em análise de dados de sobrevivência e confiabilidade um dos modelos probabilísticos que tem sido utilizado intensivamente na literatura pela sua eficiência em descrever variáveis como a vida útil de um determinado equipamento, tempos de falha e tempos de vida é o *modelo exponencial*, um caso especial do *modelo Weibull* quando o parâmetro de forma $p = 1$.

Seja T uma variável aleatória não-negativa e contínua representando os verdadeiros tempos de sobrevivência, que provém de uma distribuição exponencial com parâmetro $\mu \geq 0$. Então, a função densidade de probabilidade (*f.d.p.*) para modelar os tempos de sobrevivência é dada na forma:

$$f(t_i|\mu_i) = \frac{1}{\mu_i} \exp\left(-\frac{t_i}{\mu_i}\right), \quad (2.1)$$

onde:

$t > 0$;

$i = 1, \dots, n$;

n : é o tamanho da amostra e

μ_i : é o parâmetro desconhecido a ser estimado.

O parâmetro μ_i representa a média do tempo de vida do i – *ésimo* indivíduo, $E(T_i)$, e deve ter a mesma unidade de medida do tempo de falha t_i , isto é, se o tempo de falha tiver uma unidade de medida em segundos, minutos ou horas, então, o parâmetro μ_i terá a unidade de medida equivalente a do tempo t_i . A variância do tempo de vida do i – *ésimo* indivíduo é representada por $Var(T_i) = \mu_i^2$. Usualmente se considera mesmo valor para μ_i para todos indivíduos.

A função de sobrevivência para o i – *ésimo* indivíduo é dada por (ver por exemplo, Miller, 1981; Lawless, 1982; Louzada-Neto *et al.*, 2002):

$$S(t_i|\mu_i) = P(T \geq t_i) = 1 - F(t_i), \quad (2.2)$$

onde $F(t_i|\mu_i)$ representa a função densidade de probabilidade acumulada para o i – *ésimo* indivíduo.

No caso exponencial temos,

$$F(t_i|\mu_i) = 1 - \exp\left(-\frac{t_i}{\mu_i}\right).$$

Neste caso a função de sobrevivência para o i – *ésimo* indivíduo é dada por (ver por exemplo, Miller, 1981; Lawless, 1982; Collet, 1994; Louzada-Neto *et al.*, 2002):

$$S(t_i|\mu_i) = \exp\left(-\frac{t_i}{\mu_i}\right). \quad (2.3)$$

A função de sobrevivência do i – *ésimo* indivíduo $S(t_i|\mu_i)$, também é conhecida na literatura como razão de sobrevivência acumulada. Esta função define a probabilidade do i – *ésimo* indivíduo sobreviver mais do que um determinado tempo t_i , podendo ser representada através de gráficos denominados *curvas de sobrevivência*, sendo estes muito utilizados para a comparação de dois ou mais tratamentos em estudo.

O modelo exponencial é apropriado em situações onde os indivíduos apresentam funções de risco constante que podem depender de variáveis concomitantes, essa é uma característica muito importante da distribuição exponencial. Esta função tem uma relação matemática com a função de sobrevivência para o i – *ésimo* indivíduo que é expressa por (ver por exemplo, Miller, 1981; Lawless, 1982):

$$h(t_i|\mu_i) = \frac{f(t_i|\mu_i)}{S(t_i|\mu_i)}, \quad (2.4)$$

onde $f(t_i|\mu_i)$ é a função densidade de probabilidade de T correspondente ao i – *ésimo* indivíduo dada em (2.1) e $S(t_i|\mu_i)$ é a função de sobrevivência para o i – *ésimo* indivíduo dada em (2.3).

Se considerarmos que a função taxa de falha para a distribuição exponencial é representada por:

$$h(t_i|\mu_i) = \lambda_i,$$

então:

$$\begin{aligned}\lambda_i &= \frac{f(t_i|\mu_i)}{S(t_i|\mu_i)} \\ &= \frac{\frac{1}{\mu_i} \exp\left(-\frac{t_i}{\mu_i}\right)}{\exp\left(-\frac{t_i}{\mu_i}\right)} \\ &= \frac{1}{\mu_i}.\end{aligned}$$

Logo,

$$\lambda_i = \frac{1}{\mu_i}, \tag{2.5}$$

isto quer dizer que, a função de risco não depende do tempo t_i , ou seja, uma unidade velha em uso tem a mesma probabilidade de falhar em um tempo futuro do que uma unidade ainda não utilizada. Por exemplo, uma lâmpada com 1000 horas de uso tem o mesmo risco de falhar que uma lâmpada recém fabricada ainda não utilizada. Esta propriedade é conhecida como *Falta de Memória* da distribuição exponencial. Também é importante ressaltar que somente esta distribuição contínua apresenta tal característica (ver por exemplo, Magalhães e Lima, 2002; Magalhães, 2004).

2.2 Modelagem de Dados de Sobrevivência na Presença de Covariáveis

É muito comum em estudos da área médica a presença de variáveis explicativas que apresentam alguma relação com os tempos de sobrevivência. Para a realização de uma análise eficiente do efeito das variáveis explicativas associadas ao tempo de sobrevivência

é recomendável a utilização de um modelo de regressão que seja adequado para dados censurados.

A literatura oferece duas classes de modelos que foram propostos para a análise de sobrevivência na presença de variáveis explicativas: *modelos paramétricos* e *semi-paramétricos*.

Os *modelos paramétricos* que também são conhecidos na literatura como modelos de tempos de vida acelerados são mais eficientes, mas não apresentam grande flexibilidade se comparado aos *modelos semi-paramétricos*. Um caso especial é dado pelo *modelo de regressão de Cox*, (ver por exemplo, Cox, 1972) intensivamente utilizado em análises de dados clínicos devido a sua grande flexibilidade e facilidade de incorporar covariáveis dependentes do tempo.

2.2.1 Modelo Linear

O modelo linear é muito utilizado para explorar a relação entre as variáveis explicativas e o tempo até a ocorrência de um determinado evento de interesse o qual denominamos de variável resposta, cujo modelo estatístico é expresso na forma:

$$y_i = \beta_0 + \beta_i x_i + \varepsilon, \quad i = 1, \dots, n. \quad (2.6)$$

onde:

y : é a variável resposta;

x : é uma variável explicativa;

β_0 e β_i : são os parâmetros desconhecidos de regressão e

ε : é um erro aleatório.

Na análise de sobrevivência a situação usual de interesse é uma resposta que envolve um tempo até a ocorrência de um determinado evento e a presença de censuras, onde podemos utilizar um modelo de regressão para estudar a relação entre as variáveis. Este

tipo de resposta associada ao comportamento de variáveis com a presença de observações censuradas usualmente não permite a utilização direta do modelo linear sob normalidade do erro. Este problema pode ser contornado em duas formas equivalentes, sendo apresentadas a seguir (ver por exemplo, Colosimo, 2001):

1– podemos transformar a resposta com a finalidade de obtermos um modelo linear usual ou

2– utilizarmos um componente sistemático não-linear nos parâmetros e uma distribuição assimétrica para o componente estocástico.

Quando utilizamos um modelo linear para a transformação logarítmica da resposta, isto é equivalente a usarmos o seguinte componente sistemático:

$$\exp \{ \beta_0 + \beta_1 x \}, \tag{2.7}$$

e uma distribuição log-normal para o erro. No entanto, existem outras distribuições assimétricas para o erro que não permitem a inversão para o modelo linear.

A utilização da distribuição exponencial para o erro e um componente sistemático da forma (2.7) é o modelo de regressão mais simples e mais utilizado na literatura de análise de sobrevivência.

A partir da combinação de um componente sistemático e uma distribuição exponencial para o erro com média unitária ($f(\varepsilon) = e^{-\varepsilon}$) podemos considerar o seguinte modelo:

$$T = \exp \{ \beta_0 + \beta_1 x \} \times \varepsilon, \tag{2.8}$$

que é um modelo de regressão exponencial com uma variável que descreve o tempo de vida.

O modelo (2.8) admite uma relação não-linear entre T e x no seu componente determinístico e um erro com distribuição assimétrica. Na literatura que refere-se aos modelos lineares generalizados (ver por exemplo, McCullagh e Nelder, 1989), temos uma função de ligação logarítmica e a resposta com a distribuição exponencial. Portanto, o modelo

(2.8) é linearizável e tomando o logaritmo de T obtemos:

$$Y = \log(T) = \beta_0 + \beta_1 x + z, \quad (2.9)$$

onde $z = \log(\varepsilon)$. Podemos observar que o modelo (2.9) é semelhante ao modelo linear exceto para a distribuição dos erros que não segue uma distribuição normal.

O erro z usualmente tem uma distribuição valor extremo padrão expressa por (ver por exemplo, Lawless, 1982):

$$f(z) = \exp\{z - e^z\}. \quad (2.10)$$

Esta distribuição é bastante utilizada em análise de sobrevivência, pois caracteriza de forma adequada a distribuição do logaritmo de certos tempos de vida e está relacionada com a distribuição exponencial de T .

2.2.2 Modelo de Riscos Proporcionais de Cox

O modelo de riscos proporcionais foi proposto por Cox (1972), sendo um dos modelos mais utilizados devido a sua flexibilidade, ou seja, não requer a escolha de uma distribuição de probabilidade para os tempos de falha, além de levar em consideração as covariáveis relacionadas com esse tempo.

Este modelo supõe independência entre os tempos observados e modela a função de risco através da função (ver por exemplo, Parmar e Machin, 1995):

$$\lambda(t) = \lambda_0(t) \exp\{-x_i' \boldsymbol{\beta}\}, \quad (2.11)$$

onde:

$\lambda_0(t)$: é a função de risco de base (taxa de risco para $x_i = 0$);

$\boldsymbol{\beta}$: é o vetor de parâmetros desconhecidos;

x_i : é o vetor de covariáveis observadas para o i – *ésimo* indivíduo.

O modelo de riscos proporcionais é assim denominado pelo fato da razão entre as taxas de falhas para dois indivíduos ser constante no tempo, ou seja, tomando a razão de riscos para dois indivíduos i e j e aplicando ao modelo de riscos proporcionais, obtemos o resultado abaixo:

$$\begin{aligned} \frac{\lambda(t_i)}{\lambda(t_j)} &= \frac{\lambda_0(t) \exp\{-x_i'\boldsymbol{\beta}\}}{\lambda_0(t) \exp\{-x_j'\boldsymbol{\beta}\}} \\ &= \exp\{x_j'\boldsymbol{\beta} - x_i'\boldsymbol{\beta}\}, \quad i \neq j, \quad i, j = 1, \dots, n, \end{aligned} \quad (2.12)$$

onde observamos que a razão de riscos não depende do tempo.

Para estimar os parâmetros de tal modelo, o método de máxima verossimilhança usual não é apropriado, pois contém o componente não-paramétrico $\lambda_0(t)$. Desse modo, Cox (1975) formalizou o método de máxima verossimilhança parcial que consiste em condicionar a função de verossimilhança nos tempos de ocorrência do evento de modo a eliminar a função de perturbação $\lambda_0(t)$.

Suponha que os dados consistem de uma amostra de tamanho n , onde temos $r \leq n$ falhas distintas nos tempos de sobrevivência $t_1 \leq t_2 \leq \dots \leq t_r$. Segundo Cox (1975), a probabilidade condicional da j –ésima observação falhar no tempo t_j , conhecendo as observações que estão sob risco em t_j é expressa por (ver por exemplo, Collet, 1994):

$$\frac{\lambda_j(t_j)}{\sum_{l \in R(t_j)} \lambda_l(t_j)}, \quad (2.13)$$

onde $R(t_j)$ representa o conjunto de indivíduos que estão sob risco no tempo t_j .

Substituindo (2.11) na expressão (2.13) obtemos:

$$\begin{aligned} \frac{\lambda_j(t_j)}{\sum_{l \in R(t_j)} \lambda_l(t_j)} &= \frac{\lambda_0(t) \exp\{-x_j'\boldsymbol{\beta}\}}{\sum_{l \in R(t_j)} \lambda_0(t) \exp\{-x_l'\boldsymbol{\beta}\}} \\ &= \frac{\lambda_0(t) \exp\{-x_j'\boldsymbol{\beta}\}}{\lambda_0(t) \sum_{l \in R(t_j)} \exp\{-x_l'\boldsymbol{\beta}\}} \end{aligned}$$

Logo,

$$\frac{\lambda_j(t_j)}{\sum_{l \in R(t_j)} \lambda_l(t_j)} = \frac{\exp\{-x_j'\boldsymbol{\beta}\}}{\sum_{l \in R(t_j)} \exp\{-x_l'\boldsymbol{\beta}\}}. \quad (2.14)$$

Observe que ao substituirmos (2.11) em (2.13), a função de perturbação $\lambda_0(t)$ é eliminada do modelo.

Logo, a função de verossimilhança proposta por Cox (1975) para fazer inferências sob o modelo é dada por:

$$L(\boldsymbol{\beta}|\mathbf{x}, \mathbf{t}) = \prod_{j=1}^r \frac{\exp\{x_j'\boldsymbol{\beta}\}}{\sum_{l \in R(t_j)} \exp\{x_l'\boldsymbol{\beta}\}}, \quad (2.15)$$

onde x_j é o vetor de covariadas para os indivíduos que não sobrevivem até o j –ésimo tempo de sobrevivência t_j . O termo no denominador da expressão (2.15) é a soma dos valores de $\exp(x'\boldsymbol{\beta})$ para todos os indivíduos que estão em risco no tempo t_j .

Seja $t_1^0, t_2^0, \dots, t_n^0$ os verdadeiros tempos de sobrevivência e suponha também que temos n tempos de sobrevivência observados denotados por t_1, t_2, \dots, t_n e que δ_i representa uma variável indicadora de censura na forma:

$$\delta_i = \begin{cases} 1, & \text{se } t_i = t_i^0 \quad (\text{uma morte ou falha observada de alguma espécie}) \text{ e} \\ 0, & \text{se } t_i < t_i^0 \quad (\text{uma observação censurada}), \end{cases} \quad i = 1, \dots, n.$$

Então, a função de verossimilhança dada em (2.15) pode ser reescrita na forma:

$$L(\boldsymbol{\beta}|\mathbf{x}, \mathbf{t}) = \prod_{i=1}^n \left[\frac{\exp\{x_i'\boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} \exp\{x_j'\boldsymbol{\beta}\}} \right]^{\delta_i}, \quad (2.16)$$

onde $R(t_i)$ é o conjunto de indivíduos em risco no tempo t_i .

A partir da função de verossimilhança dada em (2.16), podemos obter a correspondente função log-verossimilhança que é dada por:

$$l(\boldsymbol{\beta}|\mathbf{x}, \mathbf{t}) = \log(L(\boldsymbol{\beta}|\mathbf{x}, \mathbf{t})) = \sum_{i=1}^n \delta_i \left[x_i' \boldsymbol{\beta} - \log \left(\sum_{j \in R(t_i)} \exp \{x_j' \boldsymbol{\beta}\} \right) \right]. \quad (2.17)$$

Os estimadores de máxima verossimilhança dos parâmetros do modelo de riscos proporcionais podem ser obtidos maximizando (2.17) a partir de métodos numéricos como por exemplo o Método de Newton-Raphson (ver por exemplo, Kalbfleisch e Prentice, 1980; Miller, 1981; Hougaard, 2000).

2.3 Ajustando o Modelo

Os modelos paramétricos são caracterizados por quantidades desconhecidas que são denominados parâmetros, os quais podem ser estimados por vários métodos estatísticos.

O método mais utilizado em modelos de regressão linear é o de Mínimos Quadrados. Quando o estudo envolve tempos de falha, em geral, o método de mínimos quadrados não é apropriado para esta situação, pelo fato de não ser possível incorporar as censuras no processo de estimação.

O ajuste de um modelo paramétrico para um conjunto de dados de sobrevivência que assume a presença de censuras é realizado utilizando o método de máxima verossimilhança para estimar os parâmetros do modelo, tal método é eficiente pelo fato de incorporar as observações censuradas e também possuir algumas propriedades para o estudo de grandes amostras.

2.3.1 Função de Verossimilhança

A função de verossimilhança para os parâmetros é a função densidade de probabilidade conjunta das variáveis aleatórias sob o qual os dados constituem pontos observados, tal

que os dados sejam avaliados nestes pontos. Sendo a função de verossimilhança composta por parâmetros desconhecidos, podemos utilizar o método de máxima verossimilhança para estimar as quantidades de interesse.

Vamos primeiramente considerar a situação em que temos os tempos de sobrevivência completos para n indivíduos observados, onde não existe a presença de observações censuradas.

Se a função densidade de probabilidade da variável aleatória associada com o tempo de sobrevivência é representada por $f(t|\boldsymbol{\theta})$ e $\boldsymbol{\theta}$ representando o vetor de parâmetros a serem estimados, então, a função de verossimilhança para as n observações independentes t_1, t_2, \dots, t_n é simplesmente o produto das *f.d.p.*'s representado por:

$$L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{t}) = \prod_{i=1}^n f(t_i). \quad (2.18)$$

Agora se considerarmos a situação mais comum da análise de sobrevivência, onde temos a presença de uma ou mais observações censuradas associada ao tempo de sobrevivência, a função de verossimilhança deve incorporar as censuras.

Considere que as observações possam ser divididas em dois conjuntos de dados com as r primeiras observações completas dos indivíduos que são denotadas por $t_{(1)} < t_{(2)} < \dots < t_{(r)}$ (dados não-censurados) e um conjunto com as $(n - r)$ observações restantes $t_1^*, t_2^*, \dots, t_{(n-r)}^*$ censuradas à direita. A contribuição das r primeiras observações não-censuradas é dada por:

$$\prod_{j=1}^r f(t_{(j)}|\boldsymbol{\theta}), \quad (2.19)$$

onde cada observação não-censurada no termo (2.19) representa a sua respectiva função densidade de probabilidade.

Naturalmente, não podemos ignorar as $(n - r)$ observações restantes dos tempos de sobrevivência, onde estas assumem a presença de censuras. Se os tempos de sobrevivência são censurados no tempo t^* e supomos que o tempo de vida de um indivíduo é pelo menos

t^* , então a probabilidade deste evento é:

$$\begin{aligned} P(T \geq t^*) &= 1 - P(T < t^*) \\ &= 1 - F(t^*) \\ &= S(t^*). \end{aligned} \tag{2.20}$$

Assim, a contribuição das $(n - r)$ observações censuradas é dada pela função de sobrevivência, e estas observações informam que o tempo de falha é maior que o tempo de censura observado. Observar que r é suposto fixo e conhecido (censuras de tipo II).

Portanto, a função de verossimilhança total para dados de sobrevivência que assumem a presença de censuras de tipo II é expressa por:

$$L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{t}) = \prod_{j=1}^r f(t_{(j)}) \prod_{l=1}^{n-r} S(t_l^*), \tag{2.21}$$

onde o primeiro produto representa as r observações completas e o segundo produto as $(n - r)$ observações restantes que são censuradas.

Supor agora que os dados são considerados como pares de n observações, onde o par para o i -ésimo indivíduo é (t_i, δ_i) , $i = 1, 2, \dots, n$. A notação δ_i denota uma variável indicadora de censura que assume o valor *zero* se os tempos de sobrevivência são censurados e 1 se os tempos de sobrevivência são não-censurados. Logo, podemos reescrever a função de verossimilhança na forma:

$$L(\boldsymbol{\theta}|\mathbf{x}, \mathbf{t}) = \prod_{i=1}^n [f(t_i)]^{\delta_i} [S(t_i^*)]^{1-\delta_i}. \tag{2.22}$$

A função dada em (2.22) é equivalente a (2.21) e podemos maximizar (2.22) com respeito aos parâmetros nas funções densidade de probabilidade e sobrevivência.

A forma de verossimilhança (2.22) é dada para dados com censuras de tipo I onde $\delta_i = 1$ se $t_i \leq c_i$ onde c_i é um tempo de censura fixo e $\delta_i = 0$ se $t_i > c_i$ para $i = 1, 2, \dots, n$. Em (2.22), c_i é igual a t_i^* .

2.3.2 Estimação dos Parâmetros do Modelo

Mencionamos anteriormente que os modelos probabilísticos são caracterizados por quantidades desconhecidas que conferem uma forma geral aos modelos, e podemos estimar estas quantidades através do método de máxima verossimilhança, onde tais quantidades quando estimadas por este método são conhecidas como *estimativas de máxima verossimilhança* (*e.m.v.*).

Considerando que $T_1^0, T_2^0, \dots, T_n^0$ são os verdadeiros tempos de sobrevivência de uma amostra de tamanho n e assumindo que as variáveis são independentes e identicamente distribuídas com função de sobrevivência $S_0(t|\lambda) = P(T_i^0 > t)$, com observações arbitrariamente censuradas à direita, temos $t_i = \min(T_i^0, c_i)$ onde c_i é o tempo de censura.

Sob o modelo exponencial (2.1), a função de verossimilhança dada em (2.22) é escrita na forma:

$$L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x}) = \prod_{i=1}^n f^{\delta_i}(t_i|\lambda_i) S^{1-\delta_i}(t_i|\lambda_i), \quad (2.23)$$

onde:

$f(t_i|\lambda_i) = \lambda_i \exp(-\lambda_i t_i)$ com diferentes λ_i para cada indivíduo e

$S(t_i|\lambda_i) = \exp(-\lambda_i t_i)$ é a função de sobrevivência para $i = 1, \dots, n$.

Logo, a função de verossimilhança para $\boldsymbol{\lambda}' = (\lambda_1, \dots, \lambda_n)$ é dada por:

$$L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x}) = \prod_{i=1}^n \lambda_i^{\delta_i} \exp(-\lambda_i t_i). \quad (2.24)$$

Sem alterar o resultado podemos considerar o log da função de verossimilhança sendo computacionalmente mais fácil e prático. Vamos usar a notação $l(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x})$ para representar o logaritmo da verossimilhança de λ que é expresso por:

$$l(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x}) = \log(L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x})) = \sum_{i=1}^n (\delta_i \log \lambda_i - \lambda_i t_i). \quad (2.25)$$

Considerando mesmo valor λ_i para todos indivíduos, isto é, $\lambda_i = \lambda$, os estimadores de máxima verossimilhança podem ser obtidos através da resolução da equação $\frac{\partial l(\lambda)}{\partial \lambda} = 0$. Existem situações onde é necessário a utilização do procedimento de Newton-Raphson (ver por exemplo, Kalbfleisch e Prentice, 1980; Miller, 1981; Hougaard, 2000) ou algum outro método iterativo para obter as *e.m.v.* do parâmetro de interesse, especialmente na presença de covariáveis.

2.3.3 Precisão das Estimativas e Intervalos de Confiança

Utilizamos o método de máxima verossimilhança com o intuito de obter os estimadores do modelo, o qual são popularmente conhecidos na literatura como *estimadores pontuais*. Através deste método também é possível construir os intervalos de confiança para os parâmetros. A construção destes intervalos é realizada a partir de propriedades para grandes amostras dos estimadores de interesse.

O resultado mais importante para grandes amostras, trata-se da precisão dos estimadores de máxima verossimilhança, estabelecendo a seguinte relação (ver por exemplo, Gamerman e Migon, 1993):

$$Var(\hat{\theta}) \approx \{E[I(\theta)]\}^{-1},$$

isto é, a matriz de covariância dos estimadores de máxima verossimilhança é aproximadamente o inverso da esperança de $I(\theta)$, onde esta é conhecida como *matriz de informação de Fisher* sendo expressa por:

$$I(\theta) = \begin{bmatrix} -\frac{\partial^2 l}{\partial \theta_1^2} & -\frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} & \cdots & -\frac{\partial^2 l}{\partial \theta_1 \partial \theta_n} \\ -\frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} & -\frac{\partial^2 l}{\partial \theta_2^2} & \cdots & -\frac{\partial^2 l}{\partial \theta_2 \partial \theta_n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\partial^2 l}{\partial \theta_1 \partial \theta_n} & -\frac{\partial^2 l}{\partial \theta_2 \partial \theta_n} & \cdots & -\frac{\partial^2 l}{\partial \theta_n^2} \end{bmatrix}.$$

Existem situações onde há dificuldade em se calcular a esperança; uma forma de contornar esta dificuldade é simplesmente utilizar a informação observada em lugar da informação esperada, isto é,

$$Var(\hat{\theta}) \approx [I(\hat{\theta})]^{-1}.$$

Assim o inverso da matriz de informação observada $I(\hat{\theta})$ é formada pelas variâncias e covariâncias dos estimadores de máxima verossimilhança $\hat{\theta}$, onde $\hat{\theta} \stackrel{a}{\sim} N\left(\theta; [I(\hat{\theta})]^{-1}\right)$ (normalidade assintótica) (ver por exemplo, Gamerman e Migon, 1993).

Um intervalo de confiança aproximado $(1 - \alpha)$ 100% para cada componente θ_i do vetor θ é dado por:

$$\hat{\theta}_i \pm z_{\alpha/2} \sqrt{Var(\hat{\theta}_i)}, \quad i = 1, \dots, n,$$

onde $\sqrt{Var(\hat{\theta}_i)}$ é a estimativa do erro-padrão obtida através da matriz de variância-covariância, denotada por $Var(\hat{\theta})$ e $z_{\alpha/2}$ é um percentil da distribuição normal $N(0, 1)$.

Uma outra propriedade importante do estimador de máxima verossimilhança é a possibilidade de estimar uma função dos parâmetros $\eta = g(\theta)$. O estimador de máxima verossimilhança de η é $\hat{\eta} = g(\hat{\theta})$, isto quer dizer que, para estimar η é necessário apenas substituir θ pelo seu respectivo estimador de máxima verossimilhança, e se existir interesse em construir um intervalo de confiança para η para verificar a precisão dos estimadores é necessário encontrar uma estimativa para o erro-padrão de η , esta estimativa é obtida utilizando o método delta (ver por exemplo, Louzada-Neto, 1991; Dias, 1992; Hosmer e Lemeshow, 1999).

Este resultado é importante em análise de sobrevivência, pois em geral o pesquisador está interessado em estimar a função de sobrevivência ou outras funções dos parâmetros, como o tempo de sobrevivência mediano.

Capítulo 3

Uso do Processo Gama e Modelos de Fragilidade para a Função de Risco do Modelo Semi-Paramétrico de Cox

Neste capítulo, apresentamos uma *priori* muito comum utilizada para modelar o componente determinístico do modelo semi-paramétrico de Cox (1972). Esta *priori* é popularmente conhecida na literatura como Processo Gama e denotada por $\mathcal{GP}[a, b]$.

Os métodos Bayesianos para modelos não-paramétricos e semi-paramétricos em análise de sobrevivência tem se tornado comuns, devido ao grande avanço tecnológico que permitiu o aprimoramento computacional das técnicas estatísticas, como por exemplo, o desenvolvimento de algoritmos computacionais eficientes, conhecidos como métodos MCMC (*Monte Carlo Markov chain*). A grande utilização dos métodos não-paramétricos é devido a modelagem que é especificada por suposições fracas. Os modelos semi-paramétricos misturam métodos paramétricos e não-paramétricos como o Modelo de Cox (Cox, 1972).

Considerando *a priori* um processo gama para o modelo semi-paramétrico de Cox (1972), construímos a função de verossimilhança, a distribuição *a posteriori* e discutimos as técnicas que podem ser utilizadas a partir de uma *priori* $\mathcal{GP}[a, b]$ para obter as inferências de interesse.

3.1 O Processo Gama

O processo gama é derivado de uma família de distribuições com densidade gama e pode ser descrito da seguinte forma: seja $\Gamma[a, b]$ denotando uma distribuição gama com parâmetro de forma $a > 0$ e parâmetro de escala $b > 0$. Considere $a(t)$, para $t > 0$ uma função crescente e contínua tal que $a(0) = 0$ e seja $Z(t)$, $t \geq 0$ um processo estocástico com as seguintes propriedades:

- (a) $Z(0) = 0$;
- (b) $Z(t)$ é independente em intervalos disjuntos e
- (c) para $t > s$, $Z(t) - Z(s) \sim \Gamma[c(a(t) - a(s)), c]$.

Portanto, o processo $\{Z(t) : t \geq 0\}$ é chamado um processo gama e é denotado por (ver por exemplo, Ibrahim *et al.*, 2001):

$$Z(t) \sim \mathcal{GP}[ca(t), c],$$

onde $a(t)$ é a média do processo gama e c é uma influência ou parâmetro de confiança aproximado da média.

3.1.1 Uso do Processo Gama Sobre o Risco Acumulado

Sob o modelo de Cox (1972), a probabilidade de sobrevivência de n indivíduos dada uma matriz de covariáveis X é dada por (ver por exemplo, Ibrahim *et al.*, 2001):

$$P(Y > y | \beta, X, H_0) = \exp \left[- \sum_{i=1}^n \exp \left\{ \exp(x'_i \beta) H_0(y_i) \right\} \right]. \quad (3.1)$$

O processo gama é frequentemente utilizado como uma *priori* para a função de risco base acumulada $H_0(y)$.

Neste caso,

$$H_0 \sim \mathcal{GP}(c_0 H^*, c_0), \quad (3.2)$$

onde $H^*(y)$ é uma função crescente com $H^*(0) = 0$.

A função H^* é frequentemente suposta como uma função paramétrica conhecida com vetor de hiperparâmetro γ_0 . Por exemplo, se H^* corresponde a distribuição exponencial, então, $H^*(y) = \gamma_0 y$, onde γ_0 é um hiperparâmetro especificado. Se H^* corresponde a uma distribuição de Weibull, então, $H^*(y) = \eta_0 y^{k_0}$, onde $\gamma_0 = (\eta_0, k_0)'$ é um vetor especificado de hiperparâmetros.

A função de sobrevivência marginal é dada por:

$$P(Y > y | \beta, X, \gamma_0, c_0) = \prod_{i=1}^n [\phi(jV_i)]^{c_0 [H^*(y_{(i)}) - H^*(y_{(i-1)})]}, \quad (3.3)$$

onde:

$$V_i = \sum_{l \in R} \exp(x_l' \beta);$$

R_i é o conjunto em risco no tempo $y_{(i)}$ e

$y_{(1)} < y_{(2)} < \dots < y_{(n)}$ são tempos ordenados distintos.

Para dados contínuos, onde os tempos de sobrevivência ordenados são todos distintos, a verossimilhança de (β, γ_0, c_0) pode ser obtida pela diferenciação de (3.3). Esta verossimilhança é definida apenas quando os tempos de sobrevivência observados são distintos (ver por exemplo, Kalbfleisch, 1978; Clayton, 1991).

3.1.2 Uso do Processo Gama para a Verossimilhança com Dados-Agrupados

Considerando uma partição finita dos tempos, $0 < s_1 < s_2 < \dots < s_J$, onde $s_J > y_i$ para todo $i = 1, 2, \dots, n$, temos os J intervalos disjuntos $(0, s_1], (s_1, s_2], \dots, (s_{J-1}, s_J]$ e seja $I_j = (s_{j-1}, s_j]$. Os dados observados D são assumidos serem avaliados como grupos dentro destes intervalos, tal que:

$$D = (X, R_j, D_j : j = 1, 2, \dots, J),$$

onde:

R_j é o conjunto em risco e

D_j é o conjunto de falhas do j – *ésimo* intervalo I_j .

Seja h_j denotando o incremento no risco de base acumulado no j – *ésimo* intervalo, que é:

$$h_j = H_0(s_j) - H_0(s_{j-1}), \quad j = 1, 2, \dots, J.$$

A *priori* processo gama, (3.2), implica que os h'_j s são independentes e:

$$h_j \sim \mathcal{GP}(\alpha_{0j} - a_{0,j-1}, c_0), \quad (3.4)$$

onde $\alpha_{0j} = c_0 H^*(s_j)$, H^* e c_0 estão definidos na seção anterior. Assim, os hiperparâmetros (H^*, c_0) para h_j consiste de uma função paramétrica do risco acumulado especificada $H^*(y)$ avaliada nos pontos finais dos intervalos do tempo, e um escalar positivo c_0 quantifica os graus da *priori* de confiança em $H^*(y)$. Se $H_0 \sim \mathcal{GP}(c_0 H^*, c_0)$, isto implica que todo incremento disjunto em H_0 tem uma *priori* dada por (3.4). Assim, a representação dos dados agrupados pode ser obtida como:

$$P(y_i \in I_j | h) = \exp \left\{ - \exp(x'_i \beta) \sum_{k=1}^{j-1} h_k \right\} \left[1 - \exp \left\{ -h_j \exp(x'_i \beta) \right\} \right],$$

onde $h = (h_1, h_2, \dots, h_J)'$.

Logo, a função de verossimilhança para dados agrupados é dada por:

$$L(\beta, h | D) \propto \prod_{j=1}^J G_j, \quad (3.5)$$

onde:

$$G_j = \exp \left\{ -h_j \sum_{k \in R_j - D_j} \exp(x'_k \beta) \right\} \prod_{l \in D_j} \left[1 - \exp \left\{ -h_j \exp(x'_l \beta) \right\} \right]. \quad (3.6)$$

Observe que a verossimilhança para dados agrupados na expressão em (3.6) é muito geral e não-limitada para o caso em que os h'_j s são realizações de um processo gama sobre H_0 . Visto que, a função risco de base H_0 entra na verossimilhança em (3.6) através dos h'_j s, temos que os parâmetros da verossimilhança são (β, h) e assim necessitamos de uma distribuição *a priori* conjunta para (β, h) . Um caso importante é quando consideramos um risco de base constante em intervalos com $h_j = \Delta_j \lambda_j$ e $\Delta_j = s_j - s_{j-1}$. Neste caso, observamos uma grande similiaridade entre as verossimilhanças (3.3) e (3.6). Na ausência de covariadas, (3.6) é reduzido a:

$$G_j = \exp \{ -h_j (r_j - d_j) \} \{ 1 - \exp(-h_j) \}^{d_j},$$

onde r_j e d_j representam o número de indivíduos no conjunto R_j e D_j , respectivamente.

Uma *priori* típica para β é uma distribuição $N(\mu_0, \Sigma_0)$. Assim, a distribuição *a posteriori* conjunta para (β, h) pode ser escrita como:

$$\begin{aligned} \pi(\beta, h|D) &\propto \prod_{j=1}^J \left[G_j h_j^{(\alpha_{0j} - \alpha_{0,j-1}) - 1} \exp(-c_0 h_j) \right] \\ &\times \exp \left(-\frac{1}{2} (\beta - \mu_0) \Sigma_0^{-1} (\beta - \mu_0) \right). \end{aligned} \quad (3.7)$$

Para gerar uma amostra da distribuição *a posteriori* conjunta de (β, h) , podemos mostrar que $(\beta|h, D)$ é log-côncava em β e assim utilizar o algoritmo de rejeição adaptativo, usado eficientemente para gerar amostras do componente β . Além disso, $[h|\beta, D]$ também é log-côncava nos componentes de h . Podemos assim utilizar os recursos MCMC (*Monte Carlo Markov chain*).

As distribuições condicionais *a posteriori* necessárias para o algoritmo *Gibbs Sampling* são dadas por:

$$(a) \pi(\beta_j | \beta_{(j)}, h, D) \propto \prod_{j=1}^J G_j \exp \left\{ -\frac{1}{2} (\beta - \mu_0) \Sigma_0^{-1} (\beta - \mu_0) \right\} \quad (3.8)$$

$$(b) \pi(h_j | h_{(j)}, \beta, D) \propto h_j^{(\alpha_{0j} - \alpha_{0,j-1}) - 1} \exp \left\{ -h_j \left(\sum_{k \in R_j - D_j} \exp(x'_k \beta) + c_0 \right) \right\},$$

onde $h_{(j)}$ denota o vetor h com exceção do j -ésimo componente.

Observe que a distribuição condicional de $h_j | h_{(j)}, \beta, D$ é dada por uma uma distribuição gama,

$$h_j | h_{(j)}, \beta, D \sim \Gamma \left[\alpha_{0j} - \alpha_{0,j-1}, \sum_{k \in R_j - D_j} \exp(x'_k \beta) + c_0 \right]. \quad (3.9)$$

3.1.3 Relação com a Verossimilhança Parcial

A verossimilhança parcial definida por Cox (1975) pode ser obtida como um caso limitado da *posteriori* marginal de β no modelo de Cox sobre uma *priori* com um processo gama para o risco de base acumulado. Para este objetivo, denotamos os intervalos de tempo como $(0, s_1], (s_1, s_2], \dots, (s_{j-1}, s_j]$ e supomos $H_0 \sim \mathcal{GP}(c_0 H^*, c_0)$.

Seja $y_{(1)} < y_{(2)} < \dots < y_{(n)}$ denotando as falhas ordenadas ou tempos de censura, e $h_j = H_0(y_{(j)}) - H_0(y_{(j-1)})$, então:

$$h_j \sim \Gamma(c_0 h_{0j}, c_0), \quad (3.10)$$

onde $h_{0j} = H^*(y_{(j)}) - H^*(y_{(j-1)})$.

Seja $\mathcal{A}_j = \sum_{l \in R} \exp(x'_l \beta)$ e $E_{\mathcal{GP}}$ a esperança com respeito a *priori* processo gama.

Então, temos:

$$\begin{aligned}
P(Y > y|X, \beta, H_0) &= \exp \left\{ - \sum_{j=1}^n \exp(x'_j \beta) H_0(y_{(j)}) \right\} \\
&= \exp \left\{ - \sum_{j=1}^n h_j \sum_{l \in R} \exp(x'_l \beta) \right\}
\end{aligned}$$

e

$$\begin{aligned}
E_{\mathcal{GP}} &= [P(Y > y|X, \beta, H_0) | H^*] = \prod_{j=1}^n \left(\frac{c_0}{c_0 + \mathcal{A}_j} \right)^{c_0 h_{0j}} \\
&= \prod_{j=1}^n \exp \left\{ c_0 H^*(y_{(j)}) \log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right) \right\}.
\end{aligned}$$

Agora seja $\theta = (\beta', h_0, c_0)'$, onde $h_0(y) = \frac{d}{dy} H^*(y)$. Então, podemos escrever a função de verossimilhança como:

$$\begin{aligned}
L(\theta|\mathcal{D}) &= \prod_{j=1}^n \exp \left\{ c_0 H^*(y_{(j)}) \log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right) \right\} \\
&\quad \times \left\{ -c_0 \frac{d}{dy_{(j)}} H^*(y_{(j)}) \left(\log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right) \right) \right\}^{\nu_i} \\
&= \prod_{j=1}^n \exp \left\{ H^*(y_{(j)}) \log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right)^{c_0} \right\} \\
&\quad \times \left\{ -c_0 h_0(y_{(j)}) \left(\log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right) \right) \right\}^{\nu_i}.
\end{aligned}$$

Seja $d = \sum_{i=1}^n \nu_i$ e $h^* = \prod_{j=1}^n [h_0(y_{(j)})]^{\nu_i}$. Logo,

$$\lim_{c_0 \rightarrow 0} \exp \left\{ H^*(y_{(j)}) \log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right)^{c_0} \right\} = 0$$

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

$$\begin{aligned} \lim_{c_0 \rightarrow 0} \log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right) &= \log \left(1 - \frac{\exp(x'_j \beta)}{\mathcal{A}_j} \right) \\ &\approx \frac{\exp(x'_j \beta)}{\mathcal{A}_j} \end{aligned}$$

para $j = 1, 2, \dots, n - 1$. Assim, temos:

$$\lim_{c_0 \rightarrow 0} \frac{L(\theta|\mathcal{D})}{c_0^d \log(c_0) h^*} \approx \prod_{i=1}^n \left[\frac{\exp(x'_i \beta)}{\mathcal{A}_i} \right]^{\nu_i}. \quad (3.11)$$

Observe que o lado direito de (3.11) é precisamente a verossimilhança parcial de Cox.

Se $c_0 \rightarrow \infty$, encontramos a função de verossimilhança baseada sobre (β, h_0) . Observamos também que:

$$\begin{aligned} &\lim_{c_0 \rightarrow 0} \left[\exp \left\{ H^*(y_{(j)}) \log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right)^{c_0} \right\} \right. \\ &\quad \left. \times \left\{ -c_0 h_0(y_{(j)}) \left(\log \left(1 - \frac{\exp(x'_j \beta)}{c_0 + \mathcal{A}_j} \right) \right) \right\}^{\nu_i} \right] \\ &= \exp \left\{ -H^*(y_{(j)}) \exp(x'_j \beta) \right\} \left\{ h_0(y_{(j)}) \exp(x'_j \beta) \right\}^{\nu_i}, \end{aligned}$$

e portanto,

$$\lim_{c_0 \rightarrow 0} L(\beta, c_0, h_0) = \prod_{j=1}^n \left[\exp \left\{ -H^*(y_{(j)}) \exp(x'_j \beta) \right\} \right] \left\{ h_0(y_{(j)}) \exp(x'_j \beta) \right\}^{\nu_i}. \quad (3.12)$$

Assim, vimos que (3.12) é a função de verossimilhança de (β, h_0) baseada no modelo de riscos proporcionais.

3.1.4 Uso do Processo Gama Sobre o Risco de Base

Uma alternativa para descrever detalhadamente o modelo semi-paramétrico de Cox (1972) é especificar uma *priori* com um processo gama sobre a taxa de risco (ver por exemplo, Dykstra e Laud, 1981). Consideramos uma aproximação discreta do processo gama estendido. Especificamente, construímos a verossimilhança usando um modelo de risco de base constante e usamos a única informação de que os tempos de falha se encontram dentro dos intervalos. Seja $0 = s_0 < s_1 < \dots < s_j$ uma partição finita dos tempos e

$$\delta_j = h_0(s_j) - h_0(s_{j-1})$$

denotando o incremento no risco de base no intervalo $(s_{j-1}, s_j]$, $j = 1, 2, \dots, J$ e $\delta = (\delta_1, \delta_2, \dots, \delta_J)'$. Para construir a função de verossimilhança aproximada de (β, δ) , utilizamos a metodologia proposta por Ibrahim, Chen e MacEachern (1999). Para um indivíduo arbitrário na população, a função de sobrevivência para o modelo de Cox no tempo y é dado por:

$$\begin{aligned} S(y|x) &= \exp \left\{ -\eta \int_0^y h_0(u) du \right\} \\ &\approx \exp \left\{ -\eta \left(\sum_{i=1}^J \delta_i (y - s_{i-1})^+ \right) \right\}, \end{aligned} \quad (3.13)$$

onde:

$$\begin{aligned} h_0(0) &= 0, \\ (u)^+ &= \begin{cases} u & \text{se } u > 0 \\ 0 & \text{outro caso} \end{cases} \text{ e} \\ \eta &= \exp(x'\beta). \end{aligned}$$

Obtemos esta primeira aproximação, visto que, a especificação δ não detalha a taxa de risco completamente, mas simplesmente o δ_j . Para o propósito de aproximação, tomamos

o incremento na taxa de risco, δ_j , para ocorrer imediatamente o s_{j-1} . Seja p_j denotando a probabilidade de uma falha no intervalo $(s_{j-1}, s_j]$, $j = 1, 2, \dots, J$.

Usando (3.13), temos:

$$\begin{aligned} p_j &= S(s_{j-1}) - S(s_j) \\ &\approx \exp \left\{ -\eta \sum_{l=1}^{j-1} \delta_l (s_{j-1} - s_{l-1}) \right\} \left[1 - \exp \left\{ -\eta (s_j - s_{j-1}) \sum_{l=1}^j \delta_l \right\} \right]. \end{aligned}$$

Assim, no j -ésimo intervalo $(s_{j-1}, s_j]$, a contribuição para a função de verossimilhança para uma falha é p_j , e $S(s_j)$ para uma observação censurada à direita. Para $j = 1, 2, \dots, J$, seja d_j o número de falhas, D_j o conjunto exposto as falhas, c_j o número de observações censuradas à direita e C_j o conjunto de sujeitos que são censurados. E, seja $\mathcal{D} = (\eta, y, X, \nu)$ denotando os dados.

A função de verossimilhança para dados agrupados é dada por:

$$L(\beta, \delta | \mathcal{D}) = \prod_{j=1}^J \left\{ \exp \{ -\delta_j (a_j + b_j) \} \prod_{k \in D_j} [1 - \exp \{ -\eta_k T_j \}] \right\}, \quad (3.14)$$

onde:

$$\eta_k = \exp \left(x'_k \beta \right),$$

$$a_j = \sum_{l=j+1}^J \sum_{k \in D_j} \eta_k (s_{l-1} - s_{j-1}),$$

$$b_j = \sum_{l=j}^M \sum_{k \in C_l} \eta_k (s_l - s_{j-1}) \quad \text{e} \quad (3.15)$$

$$T_j = (s_j - s_{j-1}) \sum_{l=1}^j \delta_l. \quad (3.16)$$

Observe que esta verossimilhança envolve uma segunda aproximação. Em vez de condicionar sobre o evento de tempos exatos, condicionamos sobre o conjunto de falhas

e o conjunto de eventos censurados à direita em cada intervalo, e assim aproximamos os dados contínuos censurados à direita por dados agrupados.

3.2 Modelos de Fragilidade

Particularmente em estudos que envolvem dados clínicos, podemos ter diversos fatores das covariadas que afetam significativamente a distribuição do tempo de sobrevivência, e quando isto ocorre, denominamos este fato de heterogeneidade dos dados.

Este assunto tem sido discutido por vários autores, como por exemplo, Vaupel, Manton e Stallard (1979) que utilizaram o conceito de fragilidade para descrever o tempo de sobrevivência relacionado a diversos fatores entre indivíduos aparentemente semelhantes, Hourgaard (1995) apresenta um estudo sobre os modelos de fragilidade para dados de sobrevivência, Klein e Moeschberger (1997) discutem métodos baseados em incorporar uma fragilidade em um modelo de riscos proporcionais, discutindo detalhes técnicos.

A noção básica de fragilidade é a introdução de efeitos aleatórios em modelos para dados de sobrevivência, quando a heterogeneidade dos dados não é observada. Em sua forma mais simples, uma fragilidade é um fator de proporcionalidade aleatório, onde este embora não seja observado modifica a função de risco de um indivíduo ou de indivíduos que estão relacionados.

Os modelos de fragilidade são extensões dos modelos de riscos proporcionais que são mais conhecido na literatura como Modelo de Cox (ver por exemplo, Cox, 1972), onde este é o modelo mais popular em análise de sobrevivência. Geralmente, a maioria das aplicações com dados clínicos em análise de sobrevivência assume uma população homogênea a ser estudada implicitamente, isto significa que, todos os indivíduos em estudo estão sujeitos em princípio ao mesmo fator de risco, como por exemplo, risco de morte, risco da doença retornar, entre outros.

Em muitas aplicações, o estudo populacional pode assumir não ser homogêneo, mas deve ser considerado como uma amostra heterogênea, ou seja, uma mistura de indivíduos

com riscos diferentes. Existem muitos casos em que é impossível medir todas as covariadas que estão relacionadas à doença de interesse, muitas vezes o motivo de não medir todas as covariáveis é devido a razões econômicas ou devido a importância de algumas covariadas ainda ser desconhecida.

A aproximação de fragilidade é um conceito de modelagem estatística, que permite assumir que a amostra apresente heterogeneidade sobre os riscos dos indivíduos, causado por covariadas não-mensuráveis. Estatisticamente, um modelo de fragilidade é um modelo onde introduzimos um efeito aleatório para os dados de sobrevivência, onde o efeito aleatório ou fragilidade, apresenta ter um efeito multiplicativo na função risco de base.

Podemos distinguir duas classes de modelos de fragilidade, que são citadas a seguir:

- (1) modelos com tempo de sobrevivência para o caso univariado e
- (2) modelos que descrevem tempo de sobrevivência para o caso multivariado, como por exemplo, riscos competitivos, eventos recorrentes, ocorrência de uma doença em parentes, etc.

No primeiro caso, o tempo de vida univariado independente é utilizado para descrever a influência das covariadas não-observadas em um modelo de riscos proporcionais. A variabilidade dos dados de sobrevivência é separada em duas partes, onde uma parte depende dos fatores de risco, e é teoricamente previsível, e a outra parte é inicialmente imprevisível, mesmo quando toda a informação pertinente é conhecida.

A separação destas duas fontes de variabilidade tem a vantagem de que a heterogeneidade poderá explicar alguns resultados inesperados ou dar uma interpretação alternativa de alguns resultados, como por exemplo, a convergência das funções de riscos de dois tratamentos diferentes.

Porém o segundo caso, onde temos os tempos de sobrevivência multivariados é mais interessante, pois os tempos são dependentes.

Uma maneira de modelar a dependência dos tempos de evento agrupados é através de um efeito aleatório agrupado - fragilidade. Este efeito aleatório explica a dependência até o

ponto em que conhecemos a fragilidade, logo os eventos seriam independentes. Em outras palavras, os tempos de vida são condicionalmente independentes, desde que determinada a fragilidade. Esta aproximação pode ser usada para tempos de sobrevivência de indivíduos relacionados ou observações periódicas do mesmo indivíduo.

Capítulo 4

Uma Abordagem Bayesiana do Modelo de Regressão Exponencial Para Dados de Sobrevida

Neste capítulo, consideramos um processo gama para as taxas individuais para o desenvolvimento de uma abordagem Bayesiana hierárquica para um modelo de regressão exponencial, assumindo dados de sobrevivência com a presença de covariadas e observações censuradas e não-censuradas.

A metodologia que apresentamos aqui é uma alternativa para o modelo exponencial usual com uma taxa determinística para ajustar dados de sobrevivência na presença de covariadas e censuras. Realizamos o desenvolvimento da análise Bayesiana hierárquica para o modelo usando métodos MCMC (*Monte Carlo Markov chain*) e ilustramos esta metodologia a dois conjuntos de dados reais, que também foram utilizados para ilustrar a metodologia do Capítulo 2.

Vamos considerar neste capítulo a mesma modelagem apresentada no capítulo anterior, onde vamos considerar a distribuição exponencial para os tempos de sobrevivência assumindo um processo gama para as taxas individuais.

4.1 Modelo Probabilístico

Um modelo de sobrevivência usualmente é composto por dois componentes: *probabilístico* e *determinístico*. O componente *probabilístico* é representado por alguma distribuição de sobrevivência, como por exemplo a Weibull, exponencial, log-normal, entre outras conhecidas na literatura. O componente *determinístico* do modelo de sobrevivência geralmente é uma função que relata a média de sobrevivência ou então uma função deste parâmetro com os níveis de covariadas.

Vamos considerar como componente probabilístico a distribuição exponencial que é uma das distribuições mais utilizadas na literatura para modelar tempos de sobrevivência, onde a *f.d.p.* desta distribuição é dada em (2.1). Se considerarmos que a taxa de falha é dada como em (2.5), então a *f.d.p.* é expressa na forma:

$$f(t_i|\lambda_i) = \lambda_i \exp(-\lambda_i t_i), \quad (4.1)$$

onde:

$$t > 0;$$

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

n : é o tamanho da amostra e

λ_i : é o parâmetro desconhecido que representa a taxa de falha constante do modelo (4.1) para o i – *ésimo* indivíduo.

Podemos assumir que o componente determinístico do modelo de sobrevivência é representado pela taxa de falha dada por:

$$\lambda_i = e^{-x'_i \boldsymbol{\beta}}, \quad (4.2)$$

onde:

$$x'_i \boldsymbol{\beta} = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki};$$

$x'_i = (1, x_{1i}, \dots, x_{ki})$: é o vetor de covariáveis e

$\boldsymbol{\beta}' = (\beta_0, \beta_1, \dots, \beta_k)$: é o vetor dos parâmetros de regressão.

Para uma análise Bayesiana, podemos assumir uma família de distribuições gama para as taxas de falha λ_i dada por:

$$\lambda_i \sim \Gamma \left[\lambda_0(t_i), e^{x_i' \boldsymbol{\beta}} \right], \quad (4.3)$$

isto é, $\lambda_1, \lambda_2, \dots, \lambda_n$ são consideradas realizações de um processo gama, onde $\Gamma[a, b]$ genericamente denota uma distribuição gama com média $\frac{a}{b}$ e variância $\frac{a}{b^2}$.

Assim, temos de (4.3) que,

$$E(\lambda_i) = \lambda_0(t_i) e^{-x_i' \boldsymbol{\beta}}, \quad (4.4)$$

onde podemos identificar (4.4) como um modelo de riscos proporcionais (ver por exemplo, Cox, 1972). Podemos considerar escolhas diferentes para a função de base $\lambda_0(t_i)$ na qual é notada uma grande flexibilidade de ajuste para os dados de sobrevivência.

Como um caso especial, assumir $\lambda_0(t_i) = \exp(\alpha_i)$; logo temos,

$$\begin{aligned} E(\lambda_i) &= \lambda_0(t_i) e^{-x_i' \boldsymbol{\beta}} \\ &= \exp(\alpha_i) e^{-x_i' \boldsymbol{\beta}} \\ &= \exp(\alpha_i - x_i' \boldsymbol{\beta}). \end{aligned} \quad (4.5)$$

Generalizações deste modelo podem ser consideradas com outras famílias de distribuições paramétricas, por exemplo, as distribuições normal, logística ou log-gama (ver por exemplo, Kalbfleisch e Prentice, 1980).

É interessante apontar que nas últimas três décadas, estão ocorrendo muitos avanços nos métodos para análises estatísticas de dados de sobrevivência na presença de covariadas, entre eles destaca-se o modelo de riscos proporcionais que é um modelo semi-paramétrico introduzido por Cox (1972) dado em (4.4).

Nas áreas de epidemiologia e medicina usualmente há uma tendência para agrupar a

ocorrência de uma doença em famílias, devido a exposição da ação ambiental ou devido a predisposição genética. Neste caso, há uma necessidade de introduzir um componente aleatório extra dentro do modelo de riscos proporcionais. Estes modelos são classificados como modelos de “fragilidade” (ver Clayton, 1991; Ibrahim, Chen e Sinha, 2001). Podemos observar que o termo α_i em (4.5) pode ser tratado como uma “fragilidade” do componente individual.

Assumindo a presença de dados de sobrevivência censurados, consideramos uma análise Bayesiana do modelo de regressão exponencial baseada em métodos MCMC (*Monte Carlo Markov chain*), como o algoritmo *Gibbs Sampling* (ver por exemplo, Gelfand e Smith, 1990) com *Metropolis-Hastings* (ver por exemplo, Smith e Roberts, 1993).

4.2 Formulação Bayesiana Usando Métodos MCMC

Seja $T_1^0, T_2^0, \dots, T_n^0$ os verdadeiros tempos de sobrevivência de uma amostra de tamanho n , onde assumimos que as variáveis aleatórias são independentes e identicamente distribuídas com função de sobrevivência expressa por:

$$S^0(t_i|\lambda_i) = P(T_i^0 > t_i) = \exp(-\lambda_i t_i), \quad i = 1, 2, \dots, n, \quad (4.6)$$

com observações sujeitas a censuras arbitrárias à direita.

Se o período de observação para o i – *ésimo* indivíduo é limitado por um valor c_i (*fixo*), então, o tempo de sobrevivência observado para o i – *ésimo* indivíduo é dado por:

$$T_i = \min(T_i^0, c_i).$$

Defina uma variável indicadora de censura δ_i dada por,

$$\delta_i = \begin{cases} 1 & \text{se } T_i = T_i^0 \text{ (uma morte ou falha observada de alguma espécie) e} \\ 0 & \text{se } T_i < T_i^0 \text{ (uma observação censurada)} \end{cases}. \quad (4.7)$$

Assim, temos que sob o modelo (4.1), a função de verossimilhança para $\boldsymbol{\lambda}' = (\lambda_1, \dots, \lambda_n)$ é expressa por:

$$L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x}) = \prod_{i=1}^n [f(t_i|\lambda_i)]^{\delta_i} [S(t_i|\lambda_i)]^{1-\delta_i}, \quad (4.8)$$

onde:

$f(t_i|\lambda_i)$: é a função densidade de probabilidade dada em (4.1) e
 $S(t_i|\lambda_i)$: é a função de sobrevivência dada em (4.6).

Então temos:

$$\begin{aligned} L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x}) &= \prod_{i=1}^n [\lambda_i \exp(-\lambda_i t_i)]^{\delta_i} [\exp(-\lambda_i t_i)]^{1-\delta_i} \\ &= \prod_{i=1}^n \lambda_i^{\delta_i} \exp(-\delta_i \lambda_i t_i) \frac{\exp(-\lambda_i t_i)}{\exp(-\delta_i \lambda_i t_i)} \\ &= \prod_{i=1}^n \lambda_i^{\delta_i} \exp(-\lambda_i t_i). \end{aligned} \quad (4.9)$$

Portanto, a função de verossimilhança para $\boldsymbol{\lambda}$ é dada por:

$$L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{x}) = \prod_{i=1}^n \lambda_i^{\delta_i} \exp(-\lambda_i t_i), \quad (4.10)$$

onde $\boldsymbol{\lambda}' = (\lambda_1, \dots, \lambda_n)$.

A partir da equação dada em (4.10), considerando que $\lambda_0(t_i) = \exp(\alpha_i)$, temos um problema para a obtenção dos estimadores de máxima verossimilhança, pois o número de parâmetros da função de verossimilhança é maior do que o tamanho amostral n e isto conduz a não-identificabilidade (ver por exemplo, Titterington *et al.*, 1985, Gelman e Rubin, 1992) dos parâmetros. Portanto, a abordagem Bayesiana hierárquica para este caso é mais recomendável do que uma análise clássica para o modelo proposto.

Para uma análise Bayesiana hierárquica do modelo, consideramos (4.3) como a densidade *a priori* para λ_i , $i = 1, \dots, n$, isto é,

$$\lambda_i \sim \Gamma \left[e^{\alpha_i}, e^{x'_i \boldsymbol{\beta}} \right]. \quad (4.11)$$

Para o segundo estágio da análise Bayesiana hierárquica consideramos as seguintes densidades *a priori* para $\boldsymbol{\alpha}$, β_l e σ_α^2 :

$$(a) \quad \alpha_i \sim N(0; \sigma_\alpha^2) \quad (4.12)$$

$$(b) \quad \sigma_\alpha^2 \sim IG(a; b)$$

$$(c) \quad \beta_l \sim N(c_l; d_l^2),$$

onde:

$$l = 0, 1, 2, \dots, k;$$

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

a, b, c e d : são hiperparâmetros conhecidos;

$IG(a; b)$: denota uma distribuição gama inversa com média $\frac{b}{(a-1)}$ e variância $\frac{b^2}{[(a-1)^2(a-2)]}$ e

$N(c; d^2)$: denota uma distribuição normal com média c e variância d^2 .

Assim as densidades *a priori* em (4.11) e (4.12) são dadas por,

$$\pi(\lambda_i | \mathbf{t}, \mathbf{x}, \delta) = \frac{[e^{x'_i \boldsymbol{\beta}}]^{\exp(\alpha_i)}}{\Gamma[\exp(\alpha_i)]} \lambda_i^{\exp(\alpha_i)-1} \exp(-\lambda_i e^{x'_i \boldsymbol{\beta}}), \quad i = 1, 2, \dots, n. \quad (4.13)$$

$$\pi(\alpha_i | \mathbf{t}, \mathbf{x}, \delta) = \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right), \quad i = 1, 2, \dots, n. \quad (4.14)$$

$$\pi(\beta_l | \mathbf{t}, \mathbf{x}, \delta) = \frac{1}{\sqrt{2\pi d_l^2}} \exp\left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right), \quad l = 0, 1, \dots, k. \quad (4.15)$$

$$\pi(\sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \delta) = (\sigma_\alpha^2)^{-(\alpha+1)} \exp(-b/\sigma_\alpha^2), \quad \sigma_\alpha^2 > 0. \quad (4.16)$$

Assumindo independência *a priori* sobre os parâmetros, a distribuição *a posteriori* conjunta para $\boldsymbol{\lambda}' = (\lambda_1, \dots, \lambda_n)$, $\boldsymbol{\alpha}' = (\alpha_1, \dots, \alpha_k)$, $\boldsymbol{\beta}' = (\beta_1, \dots, \beta_k)$ e σ_α^2 é dada por:

$$\begin{aligned} \pi(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \delta) &\propto \left[\prod_{i=1}^n \lambda_i^{\delta_i} e^{(-\lambda_i t_i)} \right] \left[\prod_{i=1}^n \frac{[e^{x'_i \boldsymbol{\beta}}]^{\exp(\alpha_i)}}{\Gamma[\exp(\alpha_i)]} \lambda_i^{\exp(\alpha_i)-1} e^{(-\lambda_i e^{x'_i \boldsymbol{\beta}})} \right] \quad (4.17) \\ &\times \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \right] (\sigma_\alpha^2)^{-(\alpha+1)} \exp(-b/\sigma_\alpha^2) \\ &\times \left[\prod_{l=0}^k \exp\left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right) \right]. \end{aligned}$$

Podemos reescrever a distribuição *a posteriori* dada em (4.17) na forma,

$$\begin{aligned} \pi(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \delta) &\propto \prod_{i=1}^n \frac{[e^{x'_i \boldsymbol{\beta}}]^{\exp(\alpha_i)}}{\Gamma[\exp(\alpha_i)]} \prod_{i=1}^n \lambda_i^{\exp(\alpha_i)+\delta_i-1} e^{\{-(t_i + e^{x'_i \boldsymbol{\beta}})\lambda_i\}} \quad (4.18) \\ &\times \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \right] (\sigma_\alpha^2)^{-(\alpha+1)} \exp(-b/\sigma_\alpha^2) \\ &\times \left[\prod_{l=0}^k \exp\left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right) \right] \end{aligned}$$

Considerando que $\lambda_0(t_i) = e^{\alpha_i}$, temos:

$$\begin{aligned} \pi(\boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \delta) &\propto \prod_{i=1}^n \frac{[e^{x'_i \boldsymbol{\beta}}]^{\lambda_0(t_i)}}{\Gamma[\lambda_0(t_i)]} \prod_{i=1}^n \lambda_i^{\lambda_0(t_i)+\delta_i-1} \exp\left\{-\left(t_i + e^{x'_i \boldsymbol{\beta}}\right)\lambda_i\right\} \quad (4.19) \\ &\times \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \right] (\sigma_\alpha^2)^{-(\alpha+1)} \exp(-b/\sigma_\alpha^2) \\ &\times \left[\prod_{l=0}^k \exp\left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right) \right] \end{aligned}$$

Observamos que:

$$\prod_{i=1}^n \int_0^\infty \lambda_i^{\lambda_0(t_i)+\delta_i-1} \exp\left\{-\left(t_i + e^{x'_i \boldsymbol{\beta}}\right)\lambda_i\right\} d\lambda_i = \prod_{i=1}^n \frac{\Gamma[\lambda_0(t_i) + \delta_i]}{(t_i + e^{x'_i \boldsymbol{\beta}})^{\lambda_0(t_i)+\delta_i}}.$$

Integrando (4.19) em relação a λ_i , a distribuição *a posteriori* conjunta para $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ e σ_α^2 é dada por:

$$\begin{aligned} \pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) &\propto \frac{\exp\left\{\sum_{i=1}^n \lambda_0(t_i) x'_i \boldsymbol{\beta}\right\}}{\prod_{i=1}^n \Gamma[\lambda_0(t_i)]} \prod_{i=1}^n \frac{\Gamma[\lambda_0(t_i) + \delta_i]}{(t_i + e^{x'_i \boldsymbol{\beta}})^{\lambda_0(t_i) + \delta_i}} \quad (4.20) \\ &\times \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \right] \\ &\times (\sigma_\alpha^2)^{-(\alpha+1)} \exp(-b/\sigma_\alpha^2) \\ &\times \left[\prod_{l=0}^k \exp\left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right) \right]. \end{aligned}$$

Visto que,

$$\Gamma[\lambda_0(t_i) + \delta_i] = \begin{cases} \lambda_0(t_i) \Gamma[\lambda_0(t_i)] & \text{se } \delta_i = 1 \\ \Gamma[\lambda_0(t_i)] & \text{se } \delta_i = 0 \end{cases}$$

temos,

$$\Gamma[\lambda_0(t_i) + \delta_i] = \lambda_0^{\delta_i}(t_i) \Gamma[\lambda_0(t_i)].$$

Como $\lambda_0(t_i) = e^{\alpha_i}$, a distribuição *a posteriori* conjunta para $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ e σ_α^2 é dada por:

$$\begin{aligned} \pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) &\propto \exp\left\{\sum_{i=1}^n e^{\alpha_i} x'_i \boldsymbol{\beta}\right\} \left\{ \frac{\prod_{i=1}^n e^{\alpha_i \delta_i}}{\prod_{i=1}^n (t_i + e^{x'_i \boldsymbol{\beta}})^{\exp(\alpha_i) + \delta_i}} \right\} \quad (4.21) \\ &\times \left[\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left(-\frac{1}{2\sigma_\alpha^2} \alpha_i^2\right) \right] (\sigma_\alpha^2)^{-(\alpha+1)} \exp(-b/\sigma_\alpha^2) \\ &\times \left[\prod_{l=0}^k \exp\left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right) \right]. \end{aligned}$$

4.3 Distribuições Condicionais *a Posteriori*

Dada a distribuição *a posteriori* em (4.12) as distribuições condicionais necessárias para o algoritmo *Gibbs Sampling* são dadas por:

$$(a) \pi(\alpha_i | \boldsymbol{\theta}_{(\alpha_i)}, \mathbf{t}, \mathbf{x}, \delta) \propto \exp\left\{-\frac{\alpha_i^2}{2\sigma_\alpha^2}\right\} \Psi_1(\theta, \mathbf{t}, \mathbf{x}, \delta), \quad (4.22)$$

onde:

$$\Psi_1(\theta, \mathbf{t}, \mathbf{x}, \delta) = \exp\left\{\alpha_i \delta_i + e^{\alpha_i} x'_i \boldsymbol{\beta} - \exp(\alpha_i) \ln(t_i + e^{x'_i \boldsymbol{\beta}})\right\}, \quad i = 1, \dots, n \text{ e}$$

$\boldsymbol{\theta}_{(\alpha_i)}$: é o vetor de todos os parâmetros exceto α_i ;

$$(b) \pi(\beta_0 | \boldsymbol{\theta}_{(\beta_0)}, \mathbf{t}, \mathbf{x}, \delta) \propto \exp\left\{-\frac{1}{2d_0^2} (\beta_0 - c_0)^2\right\} \Psi_2(\theta, \mathbf{t}, \mathbf{x}, \delta), \quad (4.23)$$

onde:

$$\Psi_2(\theta, \mathbf{t}, \mathbf{x}, \delta) = \exp\left\{\beta_0 \sum_{i=1}^n e^{\alpha_i} - \sum_{i=1}^n [\exp(\alpha_i) + \delta_i] \ln(t_i + e^{x'_i \boldsymbol{\beta}})\right\};$$

$$(c) \pi(\beta_l | \boldsymbol{\theta}_{(\beta_l)}, \mathbf{t}, \mathbf{x}, \delta) \propto \exp\left\{-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right\} \Psi_3(\theta, \mathbf{t}, \mathbf{x}, \delta), \quad (4.24)$$

onde:

$$\Psi_3(\theta, \mathbf{t}, \mathbf{x}, \delta) = \exp\left\{\beta_l \sum_{i=1}^n x_{li} e^{\alpha_i} - \sum_{i=1}^n [\exp(\alpha_i) + \delta_i] \ln(t_i + e^{x'_i \boldsymbol{\beta}})\right\}; \quad l = 1, 2, \dots, k;$$

$$(d) \sigma_\alpha^2 | \boldsymbol{\theta}_{(\sigma_\alpha^2)}, \mathbf{t}, \mathbf{x}, \delta \sim IG\left\{\frac{a+n}{2}, b + \frac{\sum_{i=1}^n \alpha_i^2}{2}\right\}. \quad (4.25)$$

Observar que $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_k$ devem ser gerados usando o algoritmo *Metropolis-Hastings* (Metropolis *et al.*, 1953; Hastings, 1970) e σ_α^2 pode ser gerado usando o algoritmo *Gibbs-Sampling* (Geman e Geman, 1984; Gelfand e Smith, 1990).

4.4 Seleção de Modelos

Para verificação da adequabilidade de um modelo podemos usar a densidade preditiva para t_i dado $t_{(i)} = (t_i, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$.

O uso da densidade preditiva em análise de dados de sobrevivência tem sido muito divulgado na literatura de estatística aplicada.

A densidade preditiva para t_i dado $t_{(i)}$ é dada por:

$$\begin{aligned} c_i &= f(t_i|t_{(i)}) \\ &= \int f(t_i|\lambda_i) \pi(\lambda_i|t_{(i)}) d\lambda_i. \end{aligned} \quad (4.26)$$

Visto que,

$$\pi(\lambda_i|t_{(i)}) = \int \int \int \pi(\lambda_i|\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2) \pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2|t_{(i)}, x_{(i)}, \delta_{(i)}) d\boldsymbol{\alpha} d\boldsymbol{\beta} d\sigma_\alpha^2,$$

temos:

$$\begin{aligned} c_i &= \int \lambda_i \exp(-\lambda_i t_i) \\ &\quad \times \left\{ \int \int \int \pi(\lambda_i|\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2) \pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2|t_{(i)}, x_{(i)}, \delta_{(i)}) d\boldsymbol{\alpha} d\boldsymbol{\beta} d\sigma_\alpha^2 \right\} d\lambda_i. \end{aligned} \quad (4.27)$$

Mudando a ordem de integração em (4.27), temos:

$$\begin{aligned} c_i &= \int \int \int \pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2|t_{(i)}, x_{(i)}, \delta_{(i)}) d\boldsymbol{\alpha} d\boldsymbol{\beta} d\sigma_\alpha^2 \\ &\quad \times \int_0^\infty \lambda_i e^{-\lambda_i t_i} \frac{[e^{x_i' \boldsymbol{\beta}}]^{\lambda_0(t_i)}}{\Gamma[\lambda_0(t_i)]} \lambda_i^{\lambda_0(t_i)-1} e^{-\lambda_i e^{x_i' \boldsymbol{\beta}}} d\lambda_i. \end{aligned}$$

Observamos que,

$$\begin{aligned}
& \int_0^\infty \lambda_i e^{-\lambda_i t_i} \frac{[e^{x'_i \boldsymbol{\beta}}]^{\lambda_0(t_i)}}{\Gamma[\lambda_0(t_i)]} \lambda_i^{\lambda_0(t_i)-1} e^{-\lambda_i e^{x'_i \boldsymbol{\beta}}} d\lambda_i \\
&= \frac{[e^{x'_i \boldsymbol{\beta}}]^{\lambda_0(t_i)}}{\Gamma[\lambda_0(t_i)]} \int_0^\infty \lambda_i^{\lambda_0(t_i)+1-1} \exp\left\{-\left(t_i + e^{x'_i \boldsymbol{\beta}}\right) \lambda_i\right\} d\lambda_i \\
&= \frac{[e^{x'_i \boldsymbol{\beta}}]^{\lambda_0(t_i)}}{\Gamma[\lambda_0(t_i)]} \frac{\Gamma[\lambda_0(t_i) + 1]}{(t_i + e^{x'_i \boldsymbol{\beta}})^{\lambda_0(t_i)+1}} \\
&= \frac{[e^{x'_i \boldsymbol{\beta}}]^{\lambda_0(t_i)}}{\Gamma[\lambda_0(t_i)]} \frac{\lambda_0(t_i) \Gamma[\lambda_0(t_i)]}{(t_i + e^{x'_i \boldsymbol{\beta}})^{\lambda_0(t_i)+1}} \\
&= \frac{\lambda_0(t_i) \exp\{\lambda_0(t_i) x'_i \boldsymbol{\beta}\}}{(t_i + e^{x'_i \boldsymbol{\beta}})^{\lambda_0(t_i)+1}},
\end{aligned}$$

onde $\lambda_0(t_i) = e^{\alpha_i}$.

Isto é,

$$c_i = \int \int \int \frac{e^{\alpha_i} \exp\{e^{\alpha_i} x'_i \boldsymbol{\beta}\}}{(t_i + e^{x'_i \boldsymbol{\beta}})^{\exp(\alpha_i)+1}} \pi(\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma_\alpha^2 | t_{(i)}, x_{(i)}, \delta_{(i)}) d\alpha d\boldsymbol{\beta} d\sigma_\alpha^2, \quad (4.28)$$

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

Podemos aproximar (4.28) pelas suas estimativas de Monte Carlo, obtidas pelas S amostras geradas de *Gibbs* dada por:

$$c_i = \hat{f}(t_i | t_{(i)}) = \frac{1}{S} \sum_{s=1}^S \frac{e^{\alpha_i^{(S)}} \exp\{e^{\alpha_i^{(S)}} x'_i \boldsymbol{\beta}^{(S)}\}}{(t_i + e^{x'_i \boldsymbol{\beta}^{(S)}})^{1+\exp(\alpha_i^{(S)})}}$$

onde $\alpha_i^{(S)}, \boldsymbol{\beta}^{(S)}$ são as amostras *Gibbs* geradas, $s = 1, 2, \dots, S$.

Podemos usar $c_i = \hat{f}(t_i | t_{(i)})$ na seleção de modelos. Consideramos para cada observação não-censurada t_i , plots de c_i versus i ($i = 1, \dots, n$) para diferentes modelos.

Para valores grandes de c_i (em média) temos indícios de que este indica o melhor modelo. Também podemos escolher o modelo que $c_l = \prod_{i=1}^n c_i(l)$ é o máximo, onde l indexa os diferentes modelos.

Capítulo 5

Uso do Processo Gama Para Dados de Sobrevivência Assumindo uma Distribuição Weibull: Uma Abordagem Bayesiana

No capítulo anterior desenvolvemos um procedimento Bayesiano para a análise dos tempos de sobrevivência considerando uma distribuição exponencial. Neste capítulo apresentamos a metodologia utilizada para o modelo exponencial considerando um outro modelo muito eficiente, que tem sido intensivamente utilizado em análise de dados de sobrevivência e confiabilidade para descrever variáveis como a vida útil de um determinado componente/equipamentos, tempos de falha ou tempos de sobrevivência conhecido como modelo *Weibull*.

Realizamos neste capítulo uma análise Bayesiana hierárquica para o modelo Weibull considerando observações não-censuradas, assumindo um processo gama para o parâmetro de escala μ . A metodologia utilizada foi baseada nos métodos MCMC (*Monte Carlo Markov chain*), especificamente o algoritmo *Gibbs Sampling* (ver por exemplo, Gelfand e Smith, 1990) com *Metropolis-Hastings* (ver por exemplo, Smith e Roberts, 1993).

5.1 Modelo Estatístico

Um modelo usual para descrever dados de sobrevivência é dada pela distribuição Weibull com *f.d.p.*, $f(t; \mu, p)$, expressa por,

$$f(t_i|\mu_i, p) = \frac{p}{\mu_i} \left(\frac{t_i}{\mu_i}\right)^{p-1} \exp\left\{-\left(\frac{t_i}{\mu_i}\right)^p\right\}, \quad t > 0, \quad (5.1)$$

onde $p, \mu_i > 0$ são os parâmetros de forma e escala, respectivamente e $i = 1, \dots, n$.

Quando $p = 1$ em (5.1), obtemos a distribuição exponencial como caso particular.

Na literatura é comum encontrarmos a distribuição Weibull escrita sob diferentes parametrizações, mas a parametrização dada em (5.1) permite uma interpretação direta do parâmetro de escala.

A função de risco para (5.1) é dada por,

$$h(t_i|\mu_i, p) = \frac{p}{\mu_i} \left(\frac{t_i}{\mu_i}\right)^{p-1}, \quad i = 1, \dots, n \quad (5.2)$$

e a função de sobrevivência é escrita na forma,

$$S(t_i|\mu_i, p) = \exp\left\{-\left(\frac{t_i}{\mu_i}\right)^p\right\}, \quad i = 1, \dots, n. \quad (5.3)$$

Assumimos uma família de distribuições gama para $\lambda_i = \frac{1}{\mu_i}$ da densidade Weibull (5.1), isto é,

$$\lambda_i \sim \Gamma\left[\lambda_0(t_i), e^{x_i'\beta}\right], \quad (5.4)$$

onde $\Gamma[a, b]$ denota uma distribuição gama com média $\frac{a}{b}$ e variância $\frac{a}{b^2}$.

Portanto, temos para (5.4) que:

$$E(\lambda_i) = \lambda_0(t_i) e^{-x_i'\beta},$$

é um *modelo de risco proporcional*. Uma variedade de formas podem ser consideradas para $\lambda_0(t_i)$, onde existe uma notável flexibilidade do ajuste para os dados de sobrevivência.

Para o modelo Weibull consideramos a mesma função utilizada para o modelo exponencial, onde a função de base é dada por:

$$\lambda_0(t_i) = \exp(\alpha_i).$$

Logo,

$$\begin{aligned} E(\lambda_i) &= \lambda_0(t_i) e^{-x_i' \boldsymbol{\beta}} \\ &= e^{\alpha_i} e^{-x_i' \boldsymbol{\beta}} \\ &= \exp\{\alpha_i - x_i' \boldsymbol{\beta}\}, \end{aligned} \tag{5.5}$$

onde $x_i' \boldsymbol{\beta} = \beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}$, para $i = 1, 2, \dots, n$.

Consideramos uma análise Bayesiana hierárquica do modelo Weibull com uma taxa determinística para ajustar dados de sobrevivência na presença de covariadas e observações censuradas e não-censuradas, baseada nos métodos MCMC (*Monte Carlo Markov Chain*), como o algoritmo *Gibbs Sampling* com *Metropolis-Hastings*.

5.2 Formulação Bayesiana Usando Métodos MCMC

Seja $T_1^0, T_2^0, \dots, T_n^0$ os verdadeiros tempos de sobrevivência de uma amostra de tamanho n . Assumimos que as variáveis aleatórias são independentes e identicamente distribuídas com função de sobrevivência $S^0(t) = P(T_i^0 > t)$. Com observações sujeitas arbitrariamente censuradas à direita, e o período seguinte para o i –ésimo indivíduo é limitado para um valor c_i .

O tempo de sobrevivência observado para o i –ésimo indivíduo é dado por $T_i = \min(T_i^0, c_i)$. Definir $\delta_i = 1$ se $T_i = T_i^0$ (uma morte ou falha observada de alguma espécie) e $\delta_i = 0$ se $T_i < T_i^0$ (uma observação censurada).

Assim, sob o modelo (5.1), a função de verossimilhança para $\boldsymbol{\mu}' = (\mu_1, \mu_2, \dots, \mu_n)$ e p é dada por,

$$L(\boldsymbol{\mu}, p | \mathbf{t}, \mathbf{x}, \delta) = \prod_{i=1}^n f^{\delta_i}(t_i | \mu_i, p) S^{1-\delta_i}(t_i | \mu_i, p) \quad (5.6)$$

onde:

$f(t_i | \mu_i, p)$ é a função densidade de probabilidade dada em (5.1) e

$S(t_i | \mu_i, p)$ é a função de sobrevivência para $i = 1, \dots, n$ dada em (5.3).

Então,

$$\begin{aligned} L(\boldsymbol{\mu}, p | \mathbf{t}, \mathbf{x}, \delta) &= \prod_{i=1}^n \left(\frac{p}{\mu_i}\right)^{\delta_i} \left(\frac{t_i}{\mu_i}\right)^{\delta_i(p-1)} \exp\left\{-\delta_i \left(\frac{t_i}{\mu_i}\right)^p\right\} \exp\left\{-(1-\delta_i) \left(\frac{t_i}{\mu_i}\right)^p\right\} \\ &= \prod_{i=1}^n \left(\frac{p}{\mu_i}\right)^{\delta_i} \left(\frac{t_i}{\mu_i}\right)^{\delta_i(p-1)} \exp\left\{-\left(\frac{t_i}{\mu_i}\right)^p\right\} \\ &= p^{\sum_{i=1}^n \delta_i} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\mu_i}\right)^p\right\} \prod_{i=1}^n \frac{t_i^{\delta_i(p-1)}}{\mu_i^{\delta_i p}} \\ &= p^s \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\mu_i}\right)^p\right\} \prod_{i=1}^n \frac{t_i^{\delta_i(p-1)}}{\mu_i^{\delta_i p}} \end{aligned} \quad (5.7)$$

onde $s = \sum_{i=1}^n \delta_i$ representa o número de falhas ou mortes.

Como mencionado no capítulo anterior, considerando que $\lambda = \frac{1}{\mu}$ e assumindo $\lambda_0(t_i) = \exp(\alpha_i)$, temos um problema para a obtenção dos estimadores de máxima verossimilhança, onde notamos que o número de parâmetros da função de verossimilhança é maior do que o tamanho amostral n , o que leva a não-identificabilidade (ver por exemplo, Titterington *et al.*, 1985, Gelman e Rubin, 1992) dos parâmetros. A utilização da metodologia Bayesiana para este caso é mais recomendável do que uma análise clássica para o modelo proposto devido superar este tipo de problema.

Para o primeiro estágio da análise Bayesiana do modelo Weibull assumimos as seguintes densidades *a priori* para as densidades de λ_i e p ,

$$\lambda_i \sim \Gamma\left[e^{\alpha_i}; e^{x_i \beta}\right], \quad \text{para } i = 1, \dots, n \quad (5.8)$$

e

$$p \sim \Gamma [c; d], \quad c, d \text{ conhecidos.}$$

Assim temos:

$$\pi (\lambda_i | \mathbf{t}, \mathbf{x}, \delta) = \frac{[e^{x'_i \boldsymbol{\beta}}]^{\exp(\alpha_i)}}{\Gamma [\exp(\alpha_i)]} \lambda_i^{\exp(\alpha_i)-1} \exp \left(-\lambda_i e^{x'_i \boldsymbol{\beta}} \right), \quad i = 1, 2, \dots, n \quad (5.9)$$

e

$$\pi (p | \mathbf{t}, \mathbf{x}, \delta) = \frac{b^a}{\Gamma [a]} p^{a-1} \exp (-pb). \quad (5.10)$$

Para o segundo estágio da análise Bayesiana hierárquica consideramos as seguintes densidades *a priori* para $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)$, σ_α^2 e β_l ,

$$\begin{aligned} (a) \quad \alpha_i &\sim N(0; \sigma_\alpha^2) \\ (b) \quad \sigma_\alpha^2 &\sim IG[a; b] \\ (c) \quad \beta_l &\sim N(c_l; d_l), \end{aligned} \quad (5.11)$$

onde a , b , c_l e d_l são os hiperparâmetros conhecidos, $i = 1, \dots, n$ e $l = 0, 1, \dots, k$.

Assumindo independência *a priori* sobre os parâmetros, a distribuição *a posteriori* conjunta para $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, p e σ_α^2 é expressa por,

$$\begin{aligned} \pi (\boldsymbol{\mu}, \boldsymbol{\alpha}, \boldsymbol{\beta}, p, \sigma_\alpha^2 | \mathbf{t}, \mathbf{x}, \delta) &\propto \left\{ \prod_{i=1}^n \frac{(e^{x'_i \boldsymbol{\beta}})^{e^{\alpha_i}}}{\Gamma [e^{\alpha_i}]} \mu_i^{-(e^{\alpha_i}+1)} \exp \left(-\frac{e^{x'_i \boldsymbol{\beta}}}{\mu_i} \right) \right\} \\ &\times \left\{ \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp \left(-\frac{\alpha_i^2}{2\sigma_\alpha^2} \right) \right\} (\sigma_\alpha^2)^{-(a+1)} \exp \left(\frac{b}{\sigma_\alpha^2} \right) \\ &\times \left\{ \prod_{l=0}^k \exp \left(-\frac{1}{2d_l^2} (\beta_l - c_l)^2 \right) \right\} p^{c-1} e^{-dp} p^s \\ &\times \left\{ \prod_{i=1}^n \frac{t_i^{\delta_i(p-1)}}{\mu_i^{\delta_i p}} \right\} \left\{ \exp \left(-\sum_{i=1}^n \left(\frac{t_i}{\mu_i} \right)^p \right) \right\}. \end{aligned} \quad (5.12)$$

5.3 Distribuições Condicionais *a Posteriori*

Dada a distribuição *a posteriori* em (5.12), as distribuições condicionais para o algoritmo *Gibbs Sampling* são dadas por,

$$(a) \pi(\mu_i | \boldsymbol{\theta}_{(\mu_i)}, \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) \propto \mu_i^{-(e^{\alpha_i+1})} \mu_i^{-\delta_i p} \exp\left\{-\frac{e^{x'_i \boldsymbol{\beta}}}{\mu_i}\right\} \exp\left\{-\left(\frac{t_i}{\mu_i}\right)^p\right\}, \quad (5.13)$$

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

$$(b) \pi(\alpha_i | \boldsymbol{\theta}_{(\alpha_i)}, \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) \propto \frac{(e^{x'_i \boldsymbol{\beta}})^{\alpha_i}}{\Gamma[e^{\alpha_i}]} \mu_i^{-(e^{\alpha_i+1})} \exp\left\{-\frac{\alpha_i^2}{2\sigma_\alpha^2}\right\}, \quad (5.14)$$

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

$$(c) \pi(\beta_l | \boldsymbol{\theta}_{(\beta_l)}, \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) \propto \exp\left\{-\frac{1}{2d_l^2} (\beta_l - c_l)^2\right\} \prod_{i=1}^n (e^{x'_i \boldsymbol{\beta}})^{e^{\alpha_i}} \exp\left\{-\frac{e^{x'_i \boldsymbol{\beta}}}{\mu_i}\right\}, \quad (5.15)$$

para $l = 0, 1, 2, \dots, k$.

$$(d) \pi(p | \boldsymbol{\theta}_{(p)}, \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) \propto p^{c-1} e^{-dp} p^s \left[\prod_{i=1}^n \frac{t_i^{\delta_i(p-1)}}{\mu_i^{\delta_i p}} \right] \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\mu_i}\right)^p\right\}. \quad (5.16)$$

$$(e) \pi(\sigma_\alpha^2 | \boldsymbol{\theta}_{(\sigma_\alpha^2)}, \mathbf{t}, \mathbf{x}, \boldsymbol{\delta}) \propto (\sigma_\alpha^2)^{-(a+1)} \exp\left\{\frac{-b}{\sigma_\alpha^2}\right\} \quad (5.17)$$

$$\times (\sigma_\alpha^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma_\alpha^2} \sum_{i=1}^n \alpha_i^2\right\}$$

$$\propto (\sigma_\alpha^2)^{-\left(\frac{n}{2}+a+1\right)} \exp\left\{-\frac{1}{\sigma_\alpha^2} \left(b + \frac{1}{2} \sum_{i=1}^n \alpha_i^2\right)\right\}.$$

Logo,

$$\sigma_{\alpha}^2 | \boldsymbol{\theta}_{(\sigma_{\alpha}^2)}, \mathbf{t}, \mathbf{x}, \boldsymbol{\delta} \sim IG \left[\frac{n}{2} + a, b + \frac{1}{2} \sum_{i=1}^n \alpha_i^2 \right]. \quad (5.18)$$

Observamos que a distribuição condicional *a posteriori* para σ_{α}^2 é representada pela distribuição gama inversa e que as distribuições condicionais *a posteriori* para $\boldsymbol{\mu}$, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$ e p não apresentam formas conhecidas, esta dificuldade é superada utilizando o algoritmo de *Metropolis-Hastings*, que permite a geração de observações que provém de distribuições desconhecidas.

Como uma alternativa para analisar o modelo proposto, utilizamos o *software* BUGS (*Bayesian Inference Using Gibbs Sampling*) que foi desenvolvido por Spiegelhalter, Thomas, Gilks e Best, 1994.

Capítulo 6

Aplicações

Neste capítulo, ilustramos a metodologia deste trabalho a dois conjuntos de dados reais, considerando dados de sobrevivência com a presença de covariadas, observações censuradas e não-censuradas.

A primeira aplicação é referente a um conjunto de dados com a presença de observações não-censuradas, que foi introduzido por Feigl e Zelen (1965), onde foram observados os tempos de sobrevivência em semanas e a contagem do número de glóbulos brancos de 17 pacientes com leucemia.

O segundo conjunto a ser analisado foi introduzido por Hosmer e Lemeshow (1999) e apresenta os tempos de sobrevivência de 100 pacientes com HIV, onde temos a presença de observações censuradas e duas covariadas (idade e história de droga empregada) relacionadas com os tempos de sobrevivência.

Para o desenvolvimento de uma análise clássica de ambos conjuntos, encontramos dificuldades para estimar os parâmetros do modelo quando assumimos um caso especial para a função de base do modelo de riscos proporcionais de Cox (ver por exemplo, Cox, 1972); neste caso, utilizamos métodos Bayesianos que contornam o problema de não-identificabilidade dos parâmetros através dos recursos MCMC (*Monte Carlo Markov chain*).

6.1 Aplicação com Dados Não-Censurados

No conjunto de dados da Tabela 6.1 introduzido por Feigl e Zelen (1965), temos os tempos de sobrevivência observados de 17 pacientes (grupo AG positivo) com leucemia. Os tempos de sobrevivência t_i são dados em semanas desde o diagnóstico. Uma única covariada x está associada a cada paciente; esta variável independente x_i é dada pelo logaritmo da contagem de glóbulos brancos (*CGB*) no momento do diagnóstico.

Tabela 6.1- Dados de sobrevivência para 17 pacientes com leucemia.

t_i	65	156	100	134	16	108	121	4	39
<i>CGB</i>	2300	750	4300	2600	6000	10000	10000	17000	5400
x_i	3.36	2.88	3.63	3.41	3.78	4.02	4.00	4.23	3.73
t_i	143	56	26	22	1	1	5	65	
<i>CGB</i>	7000	9400	32000	35000	100000	100000	52000	100000	
x_i	3.85	3.97	4.51	4.54	5.00	5.00	4.72	5.00	

Utilizamos este conjunto de dados para realizar uma análise clássica e uma abordagem Bayesiana hierárquica do modelo exponencial e de Weibull, assumindo a presença de uma única covariada e observações não-censuradas.

Para a análise Bayesiana consideramos duas situações, onde para a primeira situação assumimos que a taxa de falha λ_i é dada por (4.2) e a segunda situação a taxa de falha λ_i é dada por (4.3).

Na ordem realizamos uma comparação das duas situações para o método Bayesiano, onde utilizamos a densidade preditiva $c_i = \hat{f}(t_i|t_{(i)})$ para a seleção de modelos, na qual considerando para cada observação não-censurada t_i plots de c_i versus i ($i = 1, 2, \dots, 17$) para o modelo tradicional e para o modelo com um processo gama, considerando as distribuições exponencial e de Weibull.

6.1.1 Método Clássico para o caso exponencial

Utilizamos a distribuição exponencial dada em (2.1) para modelar os tempos de sobrevivência dada na Tabela 6.1.

Consideramos uma reparametrização da forma (ver por exemplo, Lawless, 1982),

$$y_i = \ln(t_i) = \theta_i + z_i, \quad (6.1)$$

onde:

$\theta_i = \ln(\mu_i)$, onde μ_i é a média da distribuição de T_i ;

$\mu_i = e^{\theta_i}$ e

z_i é uma variável aleatória com distribuição valor extremo padrão com *f.d.p.* dada em (2.10).

Na presença de um vetor de covariadas x_i conhecidas assumimos,

$$y_i = \ln(t_i) = x_i' \boldsymbol{\beta} + z_i, \quad (6.2)$$

onde:

$x_i' \boldsymbol{\beta} = \beta_0 + \beta_1 x_{1i}$;

$x_i' = (1, x_{1i})$: é o vetor da covariável x e

$\boldsymbol{\beta}' = (\beta_0, \beta_1)$: é o vetor dos parâmetros de regressão.

Logo, temos:

$$\begin{aligned} \lambda_i &= \exp \{ \beta_0 + \beta_1 (x_i - \bar{x}) \} \\ &= \exp \{ \beta_0 + \beta_1 x_i^* \}, \end{aligned} \quad (6.3)$$

onde \bar{x} é a média de x para o grupo de 17 pacientes com leucemia e $\lambda_i = \frac{1}{\theta_i}$.

As estimativas de máxima verossimilhança para $\hat{\beta}_0$ e $\hat{\beta}_1$ são dadas pela solução das

equações a seguir:

$$\sum_{i=1}^n t_i x_i^* e^{-\beta_1 x_i^*} = 0 \quad (6.4)$$

$$\beta_0 = \log \left(\frac{1}{n} \sum_{i=1}^n t_i e^{-\beta_1 x_i^*} \right).$$

Para a resolução do sistema dado em (6.4), utilizamos o método de Newton-Raphson (ver por exemplo, Kalbfleisch e Prentice, 1980; Miller, 1981; Hougaard, 2000) através de recursos computacionais dados nos *softwares SAS* (ver por exemplo, SAS Institute Inc., 1997) e *R* (ver por exemplo, Venables e Ripley, 1994). Através destes *softwares* também é possível calcular as matrizes de informação observada e esperada .

Considerando a derivada segunda de (6.4), obtemos a matriz de informação observada que é dada por:

$$\begin{bmatrix} 16.9933 & -0.0003 \\ -0.0003 & 5.8464 \end{bmatrix}. \quad (6.5)$$

Portanto, o inverso da matriz de informação é dada por:

$$\begin{bmatrix} 0.058847 & -0.000003 \\ -0.000003 & 0.171047 \end{bmatrix}. \quad (6.6)$$

Logo, as inferências a respeito dos parâmetros de interesse são dadas na Tabela 6.2.

Tabela 6.2- Inferências clássicas dos parâmetros do modelo log-linear.

Parâmetro	e.m.v.	desvio padrão	IC 95%
β_0	3.9339	0.2426	(3.4585; 4.4094)
β_1	-1.1093	0.4136	(-1.91995; -0.2987)

A partir dos resultados da Tabela 6.2 observamos que considerando um modelo log-linear, o fator de risco x_1 apresenta um efeito importante para os tempos de so-

brevivência dos pacientes, pois o *zero* não está incluído no intervalo de confiança.

Como mencionado nos capítulos anteriores, observamos que se considerarmos $\lambda_0(t_i) = \exp(\alpha_i)$, teremos um problema de não-identificabilidade, ou seja, o número de parâmetros será maior do que o número de observações.

Neste caso, os métodos Bayesianos superam tal dificuldade através dos recursos MCMC (*Monte Carlo Markov chain*); logo a seguir, apresentamos uma solução Bayesiana considerando um modelo log-linear e um modelo com um processo gama para as taxas de falha individuais λ_i para as distribuições exponencial e de Weibull.

6.1.2 Método Bayesiano para o caso exponencial

Como na seção anterior, para a análise Bayesiana também consideramos a distribuição exponencial para modelar os tempos de sobrevivência e assumimos o modelo log-linear dado em (6.2),

$$y_i = \ln(t_i) = \beta_0 + \beta_1 x_i^* + z_i, \quad (6.7)$$

onde z_i é uma variável aleatória com distribuição valor extremo padrão (2.10).

Logo, a função de verossimilhança para β_0 e β_1 é dada por:

$$L(\beta_0, \beta_1 | \mathbf{t}, \mathbf{x}, \delta) = \exp \left\{ \sum_{i=1}^n z_i - \sum_{i=1}^n e^{z_i} \right\} \quad (6.8)$$

onde $z_i = y_i - \beta_0 - \beta_1 x_i^*$.

Assumindo densidades *a priori* normais para β_l , $l = 0, 1$ dada em (4.12), as distribuições condicionais *a posteriori* necessárias para o algoritmo *Gibbs Sampling* para esta aplicação são dadas por:

$$(a) \pi(\beta_0 | \beta_1, \mathbf{t}, \mathbf{x}) \propto \exp \left\{ -\frac{1}{2d_0^2} (\beta_0 - c_0)^2 \right\} \Psi_1(\beta_0, \beta_1, \mathbf{t}, \mathbf{x}) \quad (6.9)$$

onde:

$$\Psi_1(\beta_0, \beta_1, \mathbf{t}, \mathbf{x}) = \exp \left\{ -n\beta_0 - e^{-\beta_0} \sum_{i=1}^n t_i e^{-\beta_1 x_i^*} \right\};$$

$$(b) \pi(\beta_1 | \beta_0, \mathbf{t}, \mathbf{x}) \propto \exp \left\{ -\frac{1}{2d_1^2} (\beta_1 - c_1)^2 \right\} \Psi_2(\beta_0, \beta_1, \mathbf{t}, \mathbf{x}) \quad (6.10)$$

onde:

$$\Psi_2(\beta_0, \beta_1, \mathbf{t}, \mathbf{x}) = \exp \left\{ -\beta_1 \sum_{i=1}^n x_i^* - e^{-\beta_0} \sum_{i=1}^n t_i e^{-\beta_1 x_i^*} \right\}.$$

Assumindo que $\lambda_i = \exp(\beta_0 + \beta_1 x_i^*)$, com densidades *a priori* $N(c_l, d_l)$ para $l = 0, 1$ como dado em (6.9) e (6.10), com $c_0 = 1$, $d_0 = (0.001)^{-1}$, $c_1 = 1.5$ e $d_1 = (0.001)^{-1}$, obtemos as inferências Bayesianas tradicionais para o modelo log-linear dado em (6.2) usando o algoritmo *Gibbs* (ver por exemplo, Casella e George, 1992). Observe que ao considerarmos estes valores para os hiperparâmetros, temos densidades *a priori* não-informativas para os parâmetros β_0 e β_1 .

A partir das distribuições condicionais *a posteriori* para β_0 e β_1 geramos 31000 amostras *Gibbs* para cada cadeia, descartando as 1000 primeiras amostras (*burn-in samples*), os resultados obtidos são apresentados na Tabela 6.3.

Os estimadores de máxima verossimilhança e seus respectivos desvios-padrão e intervalos assintóticos para β_0 e β_1 sobre o modelo log-linear (6.2) são dados na Tabela 6.2. É importante salientar que os resultados clássicos foram obtidos via resultados assintóticos do estimador de máxima verossimilhança, onde esses resultados podem não ser muito precisos quando o tamanho amostral não é muito grande ou existe uma grande proporção de dados censurados.

Podemos observar que os resultados Bayesianos dados pela Tabela 6.3 são semelhantes aos resultados obtidos pelo método clássico (ver Tabela 6.2), visto que, utilizamos *prioris* não-informativas para os parâmetros de interesse, pois $E(\beta_0) = 1$, $Var(\beta_0) = 1000$ e $E(\beta_1) = 1.5$, $Var(\beta_1) = 1000$.

Tabela 6.3- Resumo *a posteriori* dos parâmetros do modelo log-linear assumindo uma distribuição exponencial.

Parâmetro	média <i>a posteriori</i>	desvio padrão	IC 95%
β_0	3.994	0.2497	(3.5340; 4.5100)
β_1	-1.102	0.4195	(-1.9160; -0.2653)

Agora, considerando o modelo de regressão exponencial (2.1) com λ_i dada pela distribuição gama em (4.3) com $\lambda_0(t_i) = \exp(\alpha_i)$, $x_i^* = x_i - \bar{x}$, $i = 1, \dots, 17$ e as densidades *a priori* dadas em (4.12) com $a = 10$, $b = 0.1$, $c_0 = 1$, $d_0 = (0.001)^{-1}$, $c_1 = 1.5$ e $d_1 = (0.001)^{-1}$ obtemos as inferências Bayesianas usando o algoritmo *Gibbs* (ver por exemplo, Casella e George, 1992) com *Metropolis-Hastings* (ver por exemplo, Chib e Greenberg, 1995).

Observe que considerando esses valores para os hiperparâmetros das densidades *a priori* temos: $E(\sigma_\alpha^2) = 100$, $Var(\sigma_\alpha^2) = 1000$; $E(\beta_0) = 1$, $Var(\beta_0) = 1000$ e $E(\beta_1) = 1.5$, $Var(\beta_1) = 1000$, assim, para este caso observamos que *as prioris* utilizadas são não-informativas como no método Bayesiano tradicional.

A partir das distribuições condicionais *a posteriori* para $\alpha_1, \dots, \alpha_{17}, \beta_0, \beta_1$ e σ_α^2 geramos 102000 amostras *Gibbs*, e desta cadeia, descartamos as 2000 primeiras amostras (*burn-in samples*).

Na Tabela 6.3 temos os resumos *a posteriori* de interesse do modelo log-linear, considerando um processo gama para as taxas de falha do modelo exponencial.

Tabela 6.4- Resumos *a posteriori* do modelo log-linear assumindo uma distribuição exponencial e um processo gama para as taxas de falha.

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
α_1	0.2302	0.6917	(-1.1500; 1.5750)
α_2	0.2245	0.6988	(-1.1490; 1.5990)
α_3	0.0349	0.6428	(-1.258; 1.2700)
α_4	0.0495	0.6506	(-1.2800; 1.2940)

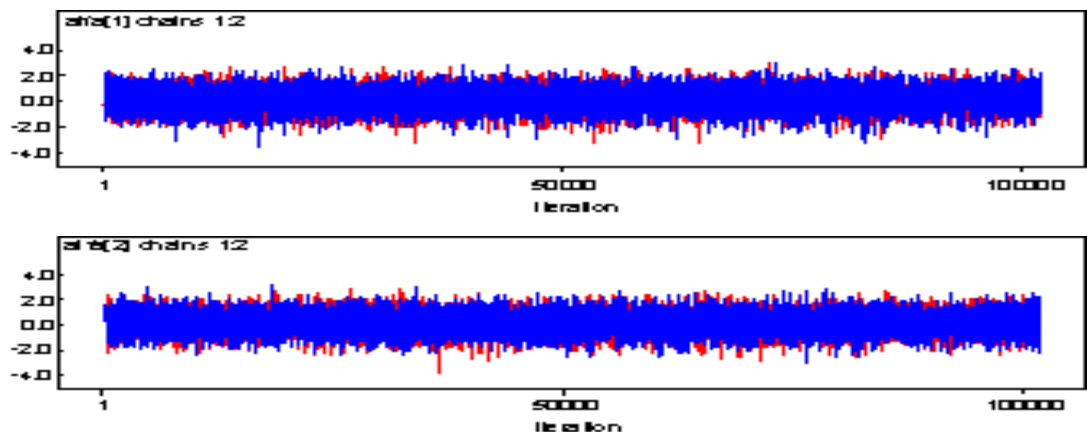
Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
α_5	0.3765	0.7214	(-1.0300; 1.8320)
α_6	-0.1484	0.6115	(-1.4260; 0.9816)
α_7	-0.1634	0.6073	(-1.4260; 0.9594)
α_8	0.4647	0.7431	(-0.9661; 1.9890)
α_9	0.2116	0.6867	(-1.1670; 1.5490)
α_{10}	-0.1441	0.6026	(-1.3910; 0.9822)
α_{11}	0.0307	0.6350	(-1.2690; 1.2230)
α_{12}	-0.0014	0.6360	(-1.3040; 1.1980)
α_{13}	0.0236	0.6421	(-1.2980; 1.2440)
α_{14}	0.4576	0.7582	(-0.9975; 2.0240)
α_{15}	0.4564	0.7547	(-1.0040; 1.9810)
α_{16}	0.2885	0.6996	(-1.0870; 1.6640)
α_{17}	-0.4021	0.5882	(-1.6390; 0.6774)
β_0	3.362	0.4008	(2.5740; 4.1510)
β_1	-1.926	0.6862	(-3.2740; -0.5789)
λ_1	0.0126	0.0105	(0.0010; 0.0098)
λ_2	0.0055	0.0048	(0.4033; 0.0041)
λ_3	0.0118	0.0092	(0.0010; 0.0095)
λ_4	0.0085	0.0067	(0.7275; 0.0069)
λ_5	0.0350	0.0298	(0.0029; 0.0270)
λ_6	0.0131	0.0101	(0.0011; 0.0107)
λ_7	0.0118	0.0091	(0.9749; 0.0096)
λ_8	0.0908	0.0811	(0.0074; 0.0682)
λ_9	0.0225	0.0181	(0.0019; 0.0179)
λ_{10}	0.0098	0.0076	(0.0008; 0.0080)
λ_{11}	0.0211	0.0164	(0.0018; 0.0171)
λ_{12}	0.0477	0.0369	(0.0041; 0.0386)

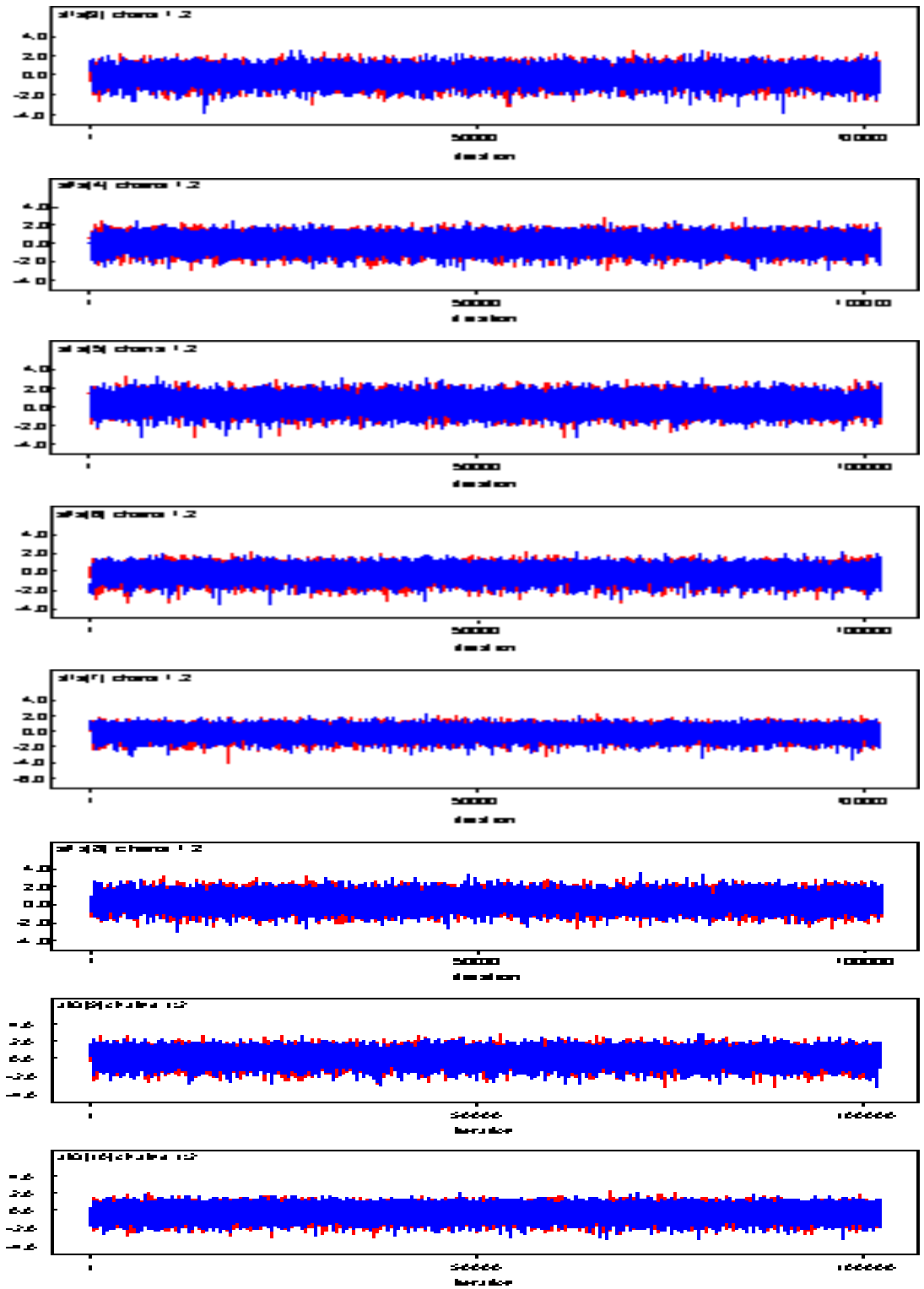
Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
λ_{13}	0.0545	0.0428	(0.0046; 0.0440)
λ_{14}	0.3695	0.3811	(0.0225; 0.2518)
λ_{15}	0.3677	0.3770	(0.0226; 0.2517)
λ_{16}	0.1476	0.1256	(0.0119; 0.1133)
λ_{17}	0.0237	0.0187	(0.0018; 0.0191)
σ_α^2	0.7879	0.1644	(0.5352; 1.1750)

A partir dos resultados da Tabela 6.4, observamos que os resultados Bayesianos para o modelo log-linear assumindo uma distribuição exponencial e um processo gama para as taxas de falha são diferentes, especialmente para os parâmetros β_0 e β_1 .

É importante salientar que os métodos assintóticos baseados na normalidade dos estimadores de máxima verossimilhança podem não ser muito precisos e dependem em geral de uma boa reparametrização. Isso pode levar a intervalos mais curtos, mas não realísticos.

Observamos a convergência das cadeias para todos os parâmetros de interesse, usando procedimentos de diagnósticos no pacote CODA (*Convergence Diagnostic e Output Analysis*) e consideramos para as distribuições condicionais *a posteriori* de λ , α , β_0 , β_1 e σ_α^2 duas cadeias com diferentes valores iniciais.





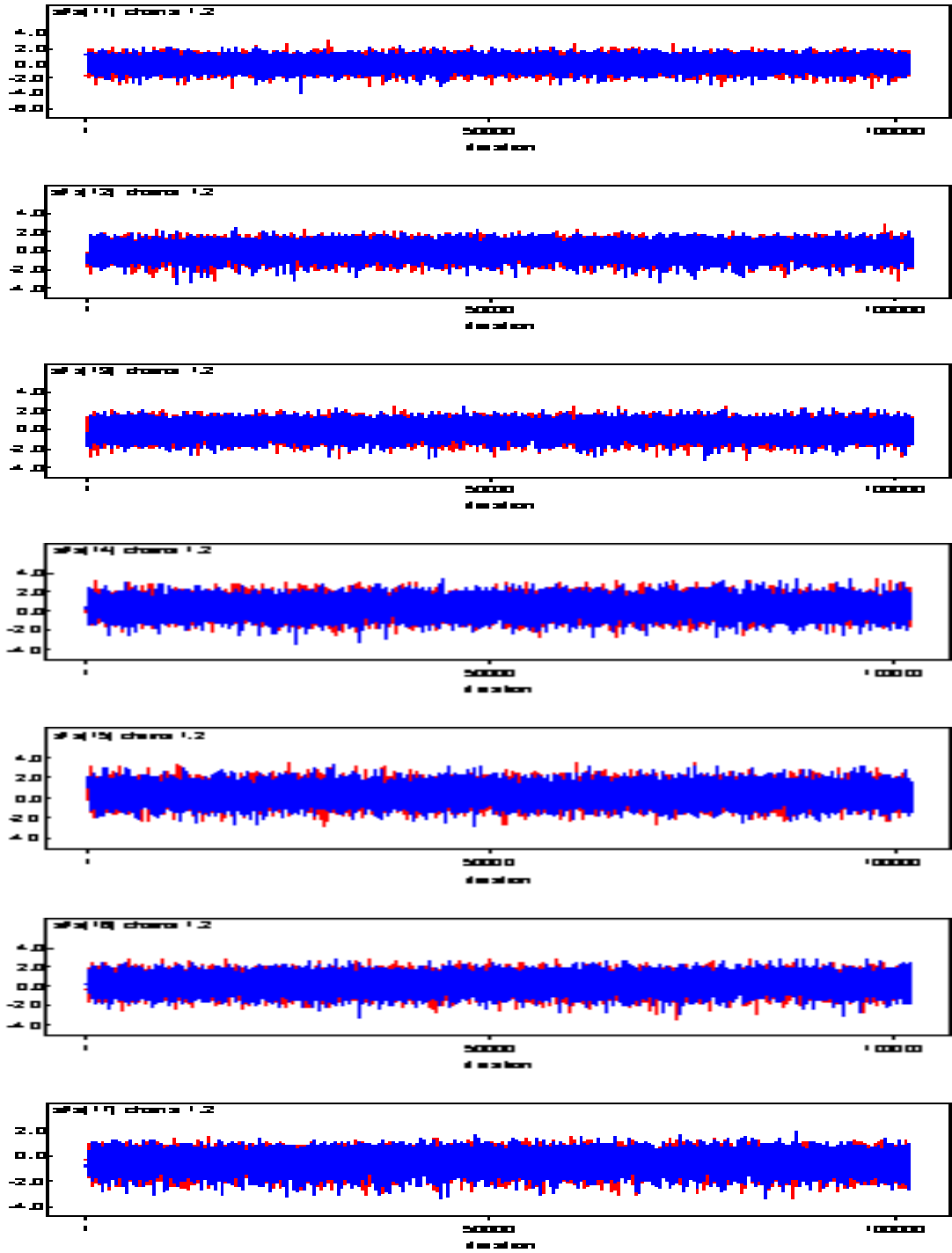


Figura 6.1- Diagnóstico de convergência para o vetor de parâmetros α .

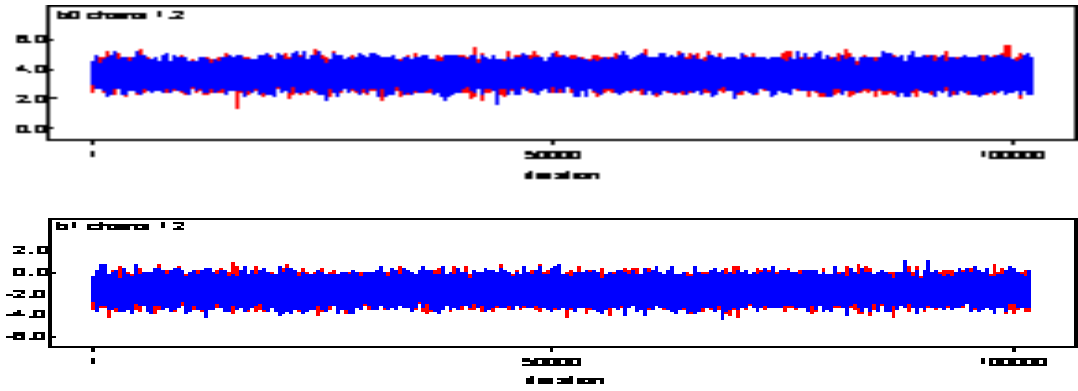
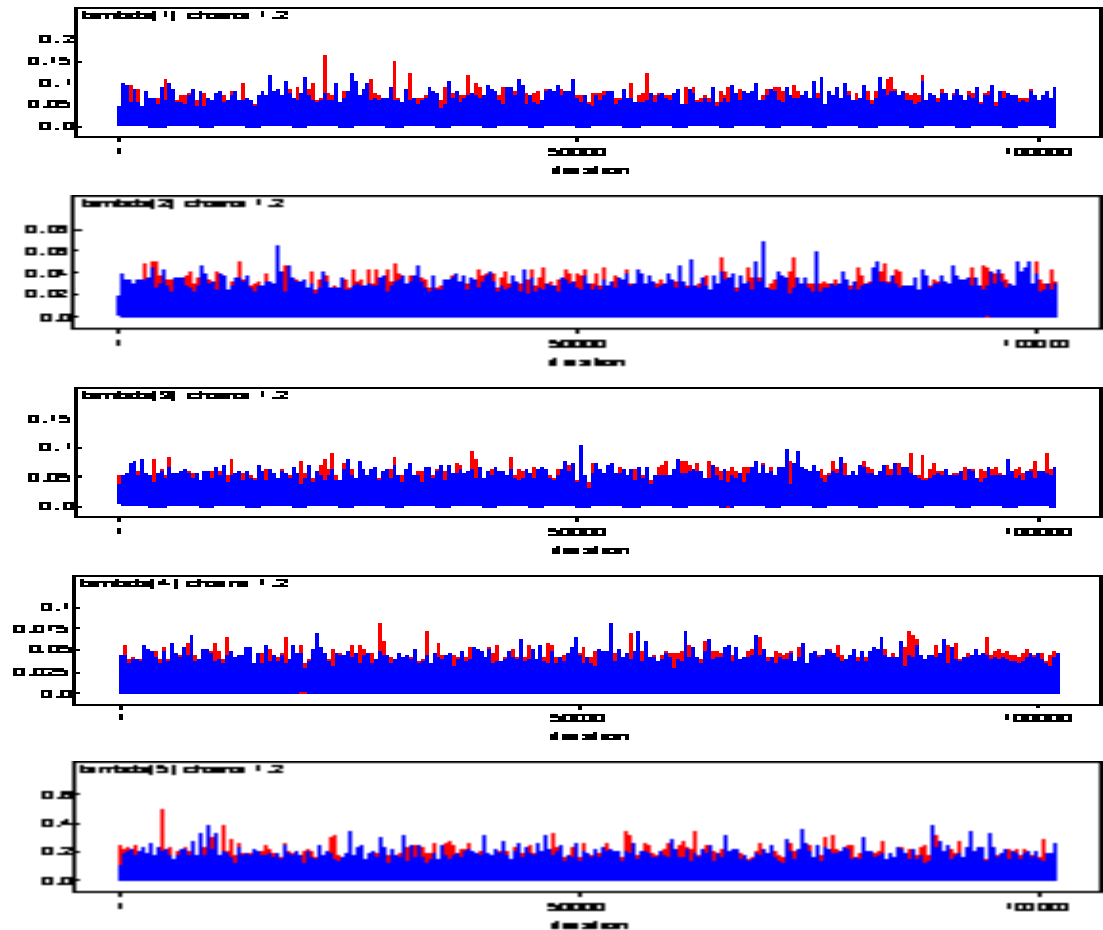
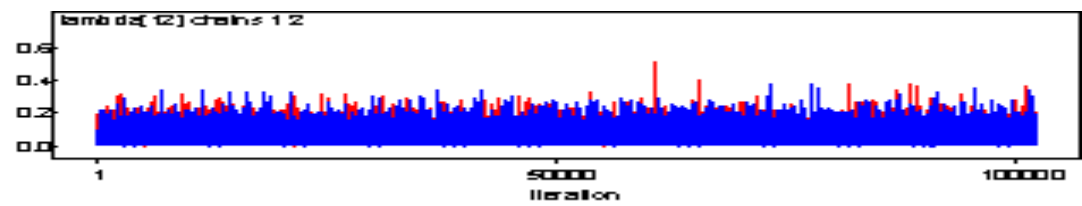
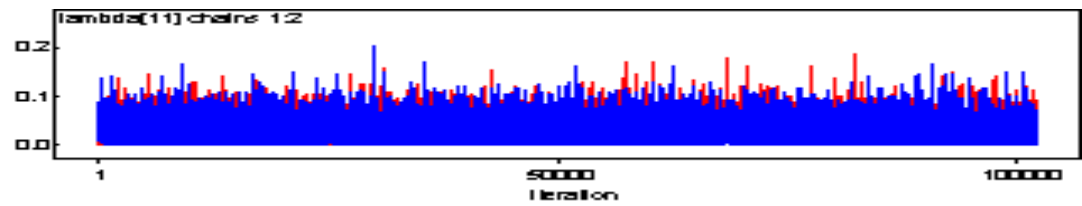
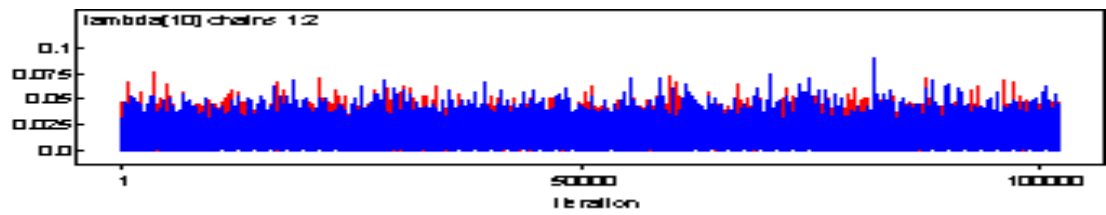
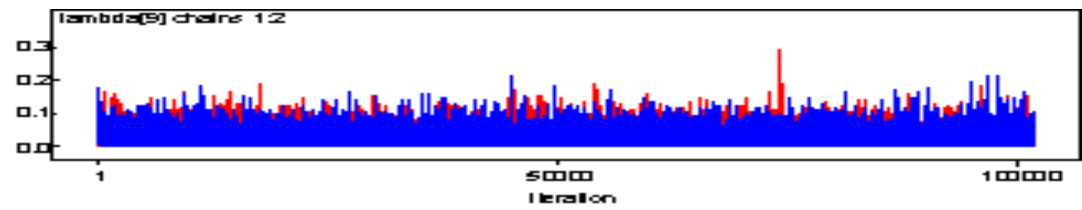
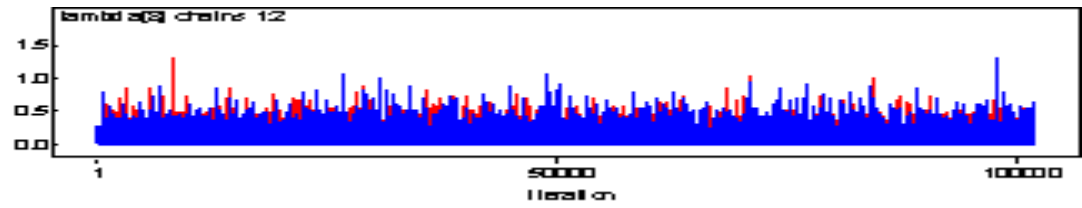
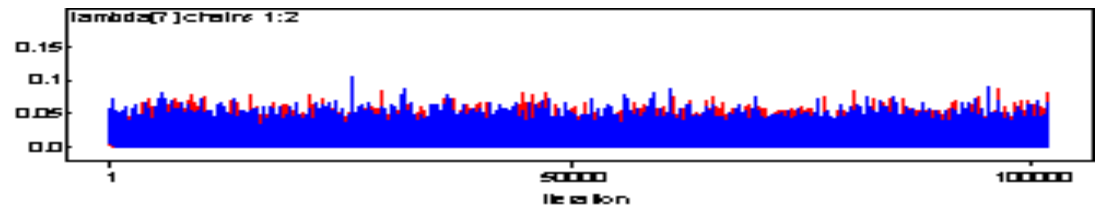
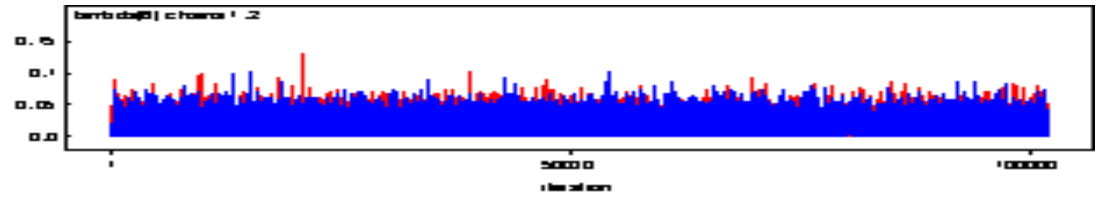


Figura 6.2- Diagnóstico de convergência para o vetor de parâmetros β .





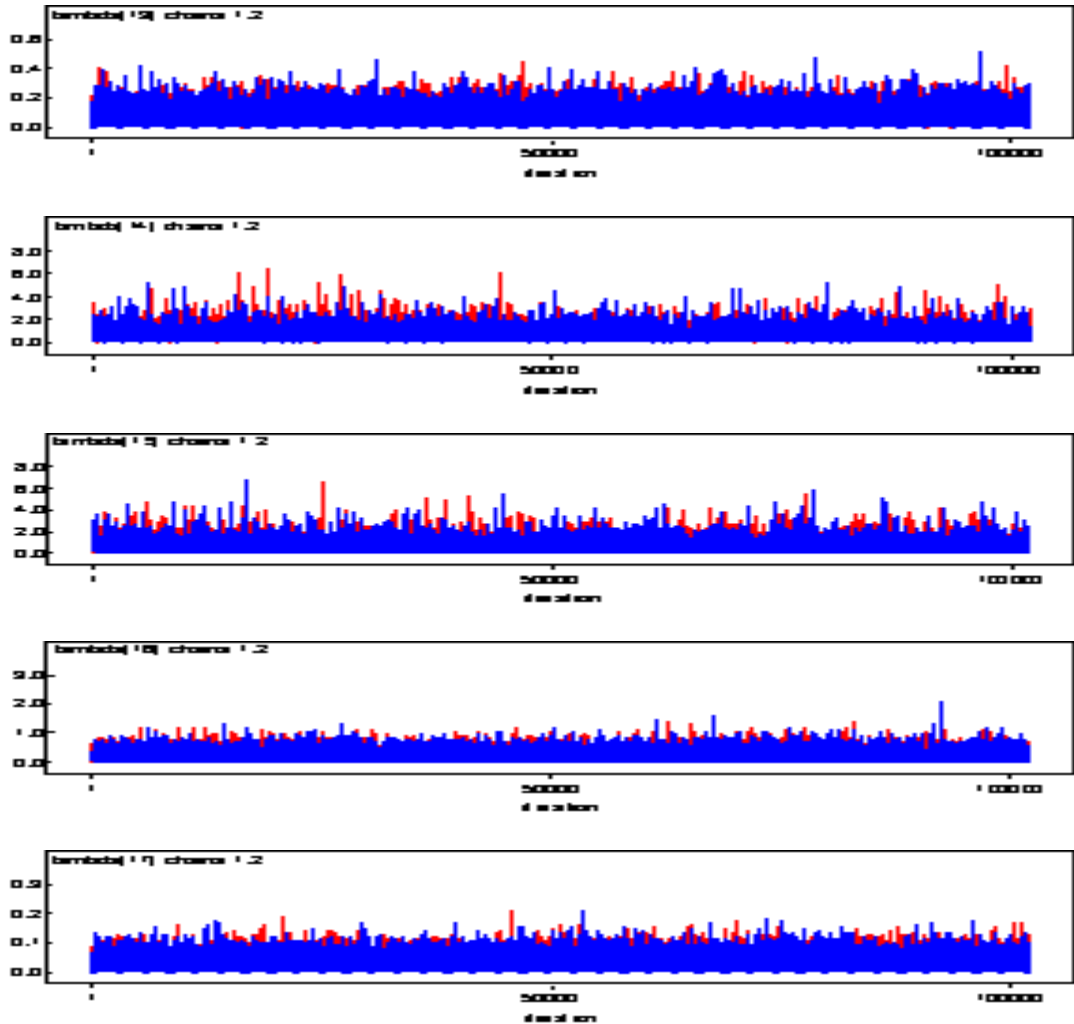


Figura 6.3- Diagnóstico de convergência para o vetor de parâmetros λ .

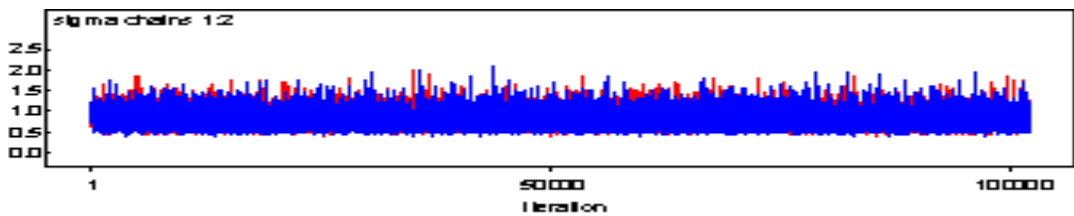


Figura 6.4- Diagnóstico de convergência para o parâmetro σ_{α}^2 .

Para a seleção de modelos, consideramos a abordagem descrita na seção 4.4 sobre as densidades preditivas. A Figura 6.5 mostra os plots de c_i versus i , considerando o modelo (4.1) com λ_i dado por (4.2), representado por (\bullet) e também consideramos o modelo (4.1) com λ_i dado por (4.3) representado por ($*$).

Os resultados obtidos para quase todas as observações estão próximas para ambos modelos para este caso, mas notamos que existe para algumas observações um melhor ajuste quando assumimos o modelo de regressão exponencial com um processo gama para as taxas de falha, e também observamos que $\sum_{i=1}^{17} \hat{c}_i = 0.4472$ para o modelo log-linear tradicional dado em (6.2) e $\sum_{i=1}^{17} \hat{c}_i = 0.4955$ para o modelo exponencial com um processo gama para as taxas de falha λ_i , logo, existem indícios de que o uso do processo gama para dados de sobrevivência apresenta inferências melhores do que utilizando os métodos Bayesianos tradicionais.

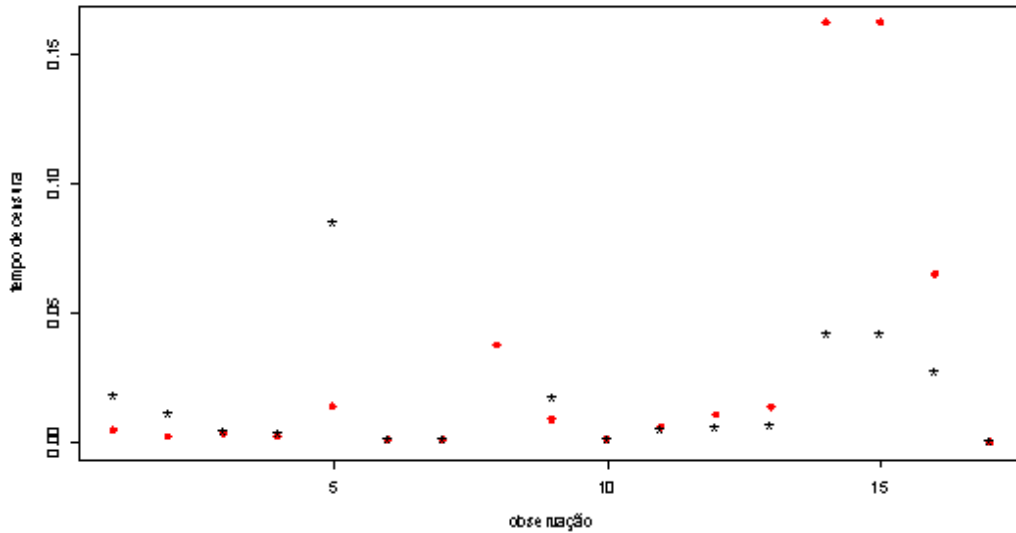


Figura 6.5- Valores das densidades preditivas considerando o modelo de regressão exponencial.

(\bullet): modelo log-linear; $*$: uso do processo gama).

6.1.3 Método Bayesiano para o caso Weibull

Antes de apresentarmos a análise Bayesiana tradicional, é importante observar que as estimativas de máxima verossimilhança para β_0 , β_1 e p obtidas pela iteração de Newton-Raphson (ver por exemplo, Kalbfleisch e Prentice, 1980; Miller, 1981; Hougaard, 2000) são dadas por:

Tabela 6.5- Inferências clássicas dos parâmetros do modelo log-linear assumindo uma distribuição de Weibull.

Parâmetro	e.m.v.	desvio padrão	IC 95%
β_0	3.9425	0.2506	(3.4514; 4.4362)
β_1	-1.0982	0.4175	(-1.9166; -0.2799)
p	1.0218	0.2059	(0.6181; 1.4255)

Para a análise Bayesiana tradicional, consideramos o modelo de Weibull (5.1) e assumimos que $\lambda_i = \exp(-\beta_0 - \beta_1 x_i^*)$ para $i = 1, \dots, 17$ com densidades *a priori* normais não-informativas para β_0 e β_1 , pois com $c_0 = 4$, $d_0 = (0.001)^{-1}$, $c_1 = -2$ e $d_1 = (0.001)^{-1}$ temos: $E(\beta_0) = 4$, $Var(\beta_0) = 1000$ e $E(\beta_1) = -2$, $Var(\beta_1) = 1000$.

Obtemos para β_0 , β_1 e p as inferências Bayesianas tradicionais usando o algoritmo *Gibbs* (ver por exemplo, Casella e George, 1992). Os resultados obtidos estão apresentados na Tabela 6.6, onde temos os resumos *a posteriori* de interesse considerando o modelo log-linear dado em (6.3).

Tabela 6.6- Resumo *a posteriori* dos parâmetros do modelo log-linear assumindo uma distribuição de Weibull.

Parâmetro	média <i>a posteriori</i>	desvio padrão	IC 95%
β_0	4.0690	0.9175	(2.4580; 6.0530)
β_1	-1.1010	0.4352	(-1.9470; -0.2419)
p	1.0160	0.2035	(0.6595; 1.4560)

Assumindo o modelo Weibull (5.1) com λ_i dada pela distribuição gama (5.4) com $\lambda_0(t_i) = \exp(\alpha_i)$ e as densidades *a priori* (5.8) e (5.9) com $a = 1, b = 1, c = 1, d = 1, c_0 = 4, d_0 = 1, c_1 = -2$ e $d_1 = 1$, obtemos as inferências Bayesianas usando o algoritmo *Gibbs* com *Metropolis-Hastings*.

Podemos observar que, neste caso, utilizamos densidades *a priori* totalmente informativas para os parâmetros de interesse, visto que: $E(\sigma_\alpha^2) = 1, Var(\sigma_\alpha^2) = 1; E(p) = 1, Var(p) = 1; E(\beta_0) = 4, Var(\beta_0) = 1$ e $E(\beta_1) = -2, Var(\beta_1) = 1$.

A partir das distribuições condicionais *a posteriori* dos parâmetros de interesse dada por (5.13), (5.14), (5.15), (5.16) e (5.17) geramos 101500 amostras *Gibbs*, descartando as 1500 primeiras amostras das cadeias (*burn-in samples*). Observamos a convergência das cadeias para todos os parâmetros de interesse, usando procedimentos de diagnósticos no pacote CODA (*Convergence Diagnostic e Output Analysis*) e consideramos para as distribuições condicionais *a posteriori* de μ, α, β, p e σ_α^2 duas cadeias com diferentes valores iniciais.

Na Tabela 6.7, temos os resumos *a posteriori* para os parâmetros de interesse, considerando a distribuição de Weibull para os tempos de sobrevivência.

Tabela 6.7- Resumo *a posteriori* do modelo log-linear assumindo uma distribuição de Weibull e um processo gama para as taxas de falha.

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
α_1	0.2310	0.6998	(-1.1580; 1.6350)
α_2	0.2367	0.7066	(-1.1560; 1.6680)
α_3	-0.0644	0.6315	(-1.3790; 1.1360)
α_4	-0.0340	0.6460	(-1.3790; 1.1900)
α_5	0.3960	0.7494	(-1.0060; 1.9620)
α_6	-0.3129	0.6014	(-1.632; 0.7617)
α_7	-0.3389	0.5993	(-1.639; 0.7383)
α_8	0.5001	0.7914	(-0.9558; 2.201)

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
α_9	0.1739	0.6731	(-1.1890; 1.5030)
α_{10}	-0.3094	0.6044	(-1.6090; 0.7807)
α_{11}	-0.0913	0.6262	(-1.4170; 1.0810)
α_{12}	-0.1611	0.6236	(-1.4990; 0.9947)
α_{13}	-0.1342	0.6221	(-1.4520; 1.0330)
α_{14}	0.4775	0.7968	(-0.9599; 2.1910)
α_{15}	0.4844	0.7956	(-0.9672; 2.221)
α_{16}	0.2347	0.6983	(-1.1410; 1.6460)
α_{17}	-0.6370	0.5917	(-1.9470; 0.3938)
β_0	4.6220	0.7032	(3.2950; 6.0540)
β_1	-2.5940	0.6999	(-3.9820; -1.2320)
σ_α	0.7894	0.2486	(0.4364; 6.0540)
p	1.3810	0.2150	(0.9853; 1.8280)
λ_1	0.0036	0.0051	(1.255E - 4; 0.01702)
λ_2	0.0012	0.0022	(2.617E - 5; 0.006627)
λ_3	0.0033	0.0049	(9.462E - 5; 0.01627)
λ_4	0.0022	0.0034	(5.69E - 5; 0.01118)
λ_5	0.0131	0.0159	(6.694E - 4; 0.0549)
λ_6	0.0037	0.0055	(8.685E - 5; 0.01831)
λ_7	0.0032	0.0049	(7.041E - 5; 0.01621)
λ_8	0.0474	0.0523	(0.00277; 0.188)
λ_9	0.0075	0.0093	(3.291E - 4; 0.03248)
λ_{10}	0.0026	0.0042	(5.391E - 5; 0.01339)
λ_{11}	0.0068	0.0088	(2.556E - 4; 0.03047)
λ_{12}	0.0192	0.0219	(9.029E - 4; 0.07804)
λ_{13}	0.0228	0.0251	(0.001151; 0.09033)

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
λ_{14}	0.3453	0.3840	(0.0180; 1.3980)
λ_{15}	0.3474	0.3850	(0.0183; 1.3940)
λ_{16}	0.0940	0.0894	(0.0063; 0.3339)
λ_{17}	0.0072	0.0107	($1.719E - 4$; 0.0352)

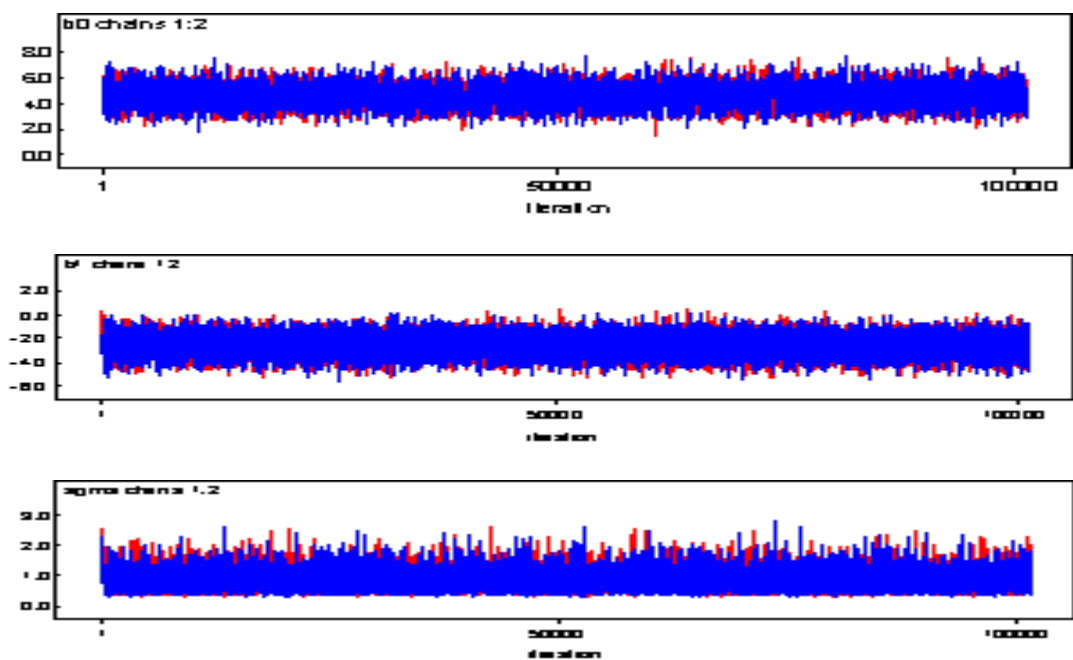


Figura 6.6- Diagnóstico de convergência dos parâmetros das cadeias geradas.

A partir da Tabela 6.7, observamos que considerando um modelo Weibull com um processo gama para λ_i , o fator de risco x_1 mostra um efeito importante para os tempos de sobrevivência dos pacientes.

A convergência das cadeias pode ser analisada a partir da Figura 6.6, através dos gráficos dos parâmetros β_0 , β_1 e σ ; observamos que as duas cadeias geradas apresentam indicações de que os valores gerados para ambas cadeias convergiram.

A Figura 6.7 mostra os plots de c_i versus i considerando o modelo (4.1) com λ_i dado por (4.2), representado por (\bullet) e também consideramos o modelo (4.1) com λ_i dado por

(4.3) representado por (*).

Observamos que o uso do processo gama para as taxas de falha, melhora o ajuste do modelo de regressão de Weibull e os valores preditivos são grandes em geral do que usando o modelo log-linear. Também observamos que $\sum_{i=1}^n \hat{c}_i = 0.2928$ para o modelo log-linear tradicional dado em (6.2) e $\sum_{i=1}^n \hat{c}_i = 1.0818$ para o modelo de Weibull com um processo gama para as taxas λ_i .

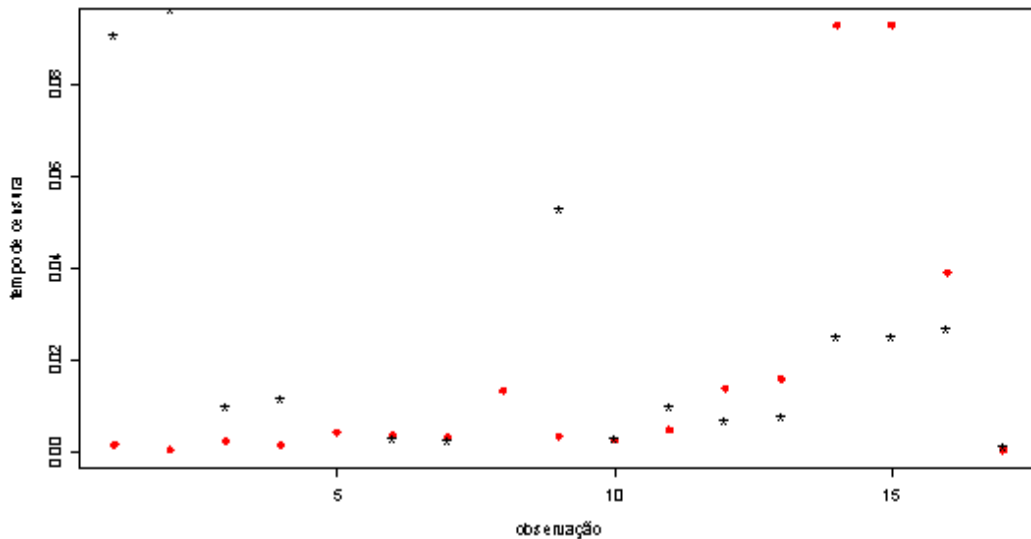


Figura 6.7- Valores das densidades preditivas considerando o modelo de regressão de Weibull.

(●: modelo log-linear; *: uso do processo gama).

6.2 Aplicação com Dados Censurados

No conjunto de dados da Tabela 6.8 temos os tempos de sobrevivência de 100 pacientes com HIV positivo introduzido por Hosmer e Lemeshow (1999). O estudo destes dados foi realizado e coletado por uma grande Organização de Apoio a Saúde - HMO.

As pessoas foram registradas no estudo de janeiro de 1989 à dezembro de 1991 depois de confirmado o diagnóstico de HIV positivo. O estudo foi finalizado em dezembro de 1995. Depois de um diagnóstico confirmado de HIV, os pacientes foram acompanhados vivos até a morte devido a AIDS ou complicações relativas a doença. Também há pacientes que foram acompanhados até o final do estudo ou até que abandonassem o tratamento (dados censurados).

A variável de interesse é o tempo de sobrevivência em meses depois de um diagnóstico confirmado de HIV e associadas ao tempo de sobrevivência temos as variáveis preditoras: idade da pessoa no início do estudo em anos, história *a priori* de droga empregada (1 se sim e 0 se não) e estágio no final do estudo (1 se a pessoa saiu do estudo devido a AIDS e 0 por algum outro motivo não relacionado com a doença).

Tabela 6.8 - Tempos de sobrevivência de pacientes com HIV positivo.

<i>id</i>	<i>t_i</i>	<i>idade</i>	<i>droga</i>	δ_i	<i>id</i>	<i>t_i</i>	<i>idade</i>	<i>droga</i>	δ_i
1	5	46	0	1	15	34	36	0	1
2	6	35	1	0	16	1	36	0	1
3	8	30	1	1	17	4	54	0	1
4	3	30	1	1	18	19	35	0	0
5	22	36	0	1	19	3	44	1	0
6	1	32	1	0	20	2	38	0	1
7	7	36	1	1	21	2	40	0	0
8	9	31	1	1	22	6	34	1	1
9	3	48	0	1	23	60	25	0	0
10	12	47	0	1	24	11	32	0	1
11	2	28	1	0	25	2	42	1	0
12	12	34	0	1	26	5	47	0	1
13	1	44	1	1	27	4	30	0	0
14	15	32	1	1	28	1	47	1	1

<i>id</i>	<i>t_i</i>	<i>idade</i>	<i>droga</i>	δ_i	<i>id</i>	<i>t_i</i>	<i>idade</i>	<i>droga</i>	δ_i
29	13	41	0	1	55	7	36	1	1
30	3	40	1	1	56	3	30	1	1
31	2	43	0	1	57	3	42	1	1
32	1	41	0	1	58	2	32	1	1
33	30	30	0	1	59	32	34	0	1
34	7	37	0	1	60	3	38	1	1
35	4	42	1	1	61	10	33	0	0
36	8	31	1	1	62	11	39	1	1
37	5	39	1	1	63	3	39	1	1
38	10	32	0	1	64	7	33	1	1
39	2	51	0	1	65	5	34	1	1
40	9	36	0	1	66	31	34	0	1
41	36	43	0	1	67	5	46	1	1
42	3	39	0	1	68	58	22	0	1
43	9	33	0	1	69	1	44	1	1
44	3	45	1	1	70	3	37	0	0
45	35	33	0	1	71	43	25	0	1
46	8	28	0	1	72	1	38	0	1
47	11	31	0	1	73	6	32	0	1
48	56	20	1	0	74	53	34	0	1
49	2	44	0	0	75	14	29	0	1
50	3	39	1	1	76	4	36	1	1
51	15	30	0	1	77	54	21	0	1
52	1	31	0	1	78	1	26	1	1
53	10	33	0	1	79	1	32	1	1
54	1	50	1	1	80	8	42	0	1

<i>id</i>	<i>t_i</i>	<i>idade</i>	<i>droga</i>	δ_i	<i>id</i>	<i>t_i</i>	<i>idade</i>	<i>droga</i>	δ_i
81	5	40	1	1	91	4	35	0	1
82	1	37	1	1	92	57	36	0	1
83	1	47	0	1	93	1	41	1	1
84	2	32	1	1	94	12	36	1	0
85	7	41	1	0	95	7	35	1	1
86	1	46	1	0	96	1	34	1	1
87	10	26	1	1	97	5	28	0	1
88	7	32	1	1	98	60	29	0	0
89	24	30	0	0	99	2	35	1	0
90	12	31	1	0	100	1	34	1	1

6.2.1 Método Clássico para o caso exponencial

Vamos utilizar a modelagem da seção anterior, mas assumindo a presença de observações censuradas. Assim, consideramos o modelo log-linear,

$$y_i = \ln(t_i) = x_i' \boldsymbol{\beta} + z_i, \quad (6.11)$$

onde:

$$x_i' \boldsymbol{\beta} = \beta_0 + \beta_1 \text{idade}_i + \beta_2 \text{droga}_i;$$

$x_i' = (1, \text{idade}_i^*, \text{droga}_i)$: é o vetor das variáveis preditoras e

$\boldsymbol{\beta}' = (\beta_0, \beta_1, \beta_2)$: é o vetor dos parâmetros de regressão.

Logo,

$$\lambda_i = \exp \{ \beta_0 + \beta_1 (\text{idade}_i - \overline{\text{idade}}) + \beta_2 \text{droga}_i \}, \quad (6.12)$$

onde $\overline{\text{idade}}$ é a média das idades das 100 pessoas com HIV.

As estimativas de máxima verossimilhança para $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ são dadas pela solução

das equações a seguir:

$$\begin{aligned}
& -\sum_{i=1}^n \delta_i + \sum_{i=1}^n e^{y_i - \beta_0 - \beta_1 idade_i^* - \beta_2 droga_i} = 0 \\
& -\sum_{i=1}^n \delta_i idade_i + \sum_{i=1}^n idade_i e^{y_i - \beta_0 - \beta_1 idade_i^* - \beta_2 droga_i} = 0 \\
& -\sum_{i=1}^n \delta_i droga_i + \sum_{i=1}^n droga_i e^{y_i - \beta_0 - \beta_1 idade_i^* - \beta_2 droga_i} = 0.
\end{aligned} \tag{6.13}$$

Neste caso, é necessário utilizar um método numérico para obter as estimativas de máxima verossimilhança como na seção anterior. Utilizamos o método de Newton-Raphson (ver por exemplo, Kalbfleisch e Prentice, 1980; Miller, 1981; Hougaard, 2000) através de recursos computacionais do *software R*.

A matriz de informação observada é dada por:

$$\begin{bmatrix} 79.9775 & 27.6053 & 37.9927 \\ 27.6053 & 3854.7867 & 21.8219 \\ 37.9927 & 21.8219 & 37.9962 \end{bmatrix}. \tag{6.14}$$

O inverso da matriz de informação de Fisher é dada por:

$$\begin{bmatrix} 0.023821 & -0.000036 & -0.023798 \\ -0.000036 & 0.000260 & -0.000114 \\ -0.023798 & -0.000114 & 0.050180 \end{bmatrix}. \tag{6.15}$$

As inferências a respeito dos parâmetros de interesse são dados na Tabela 6.9 (estimadores de máxima verossimilhança e intervalos de confiança assintóticos).

Tabela 6.9- Inferências clássicas dos parâmetros do modelo log-linear.

Parâmetro	e.m.v.	desvio padrão	IC 95%
β_0	2.8287	0.1543	(2.5262; 3.1312)
β_1	-0.0916	0.0161	(-0.1232; -0.0599)
β_2	-1.0062	0.2240	(-1.4453; -0.5672)

A partir dos resultados da Tabela 6.9 observamos que considerando o modelo log-linear (6.11), a idade e a história da droga apresentam efeitos importantes para os tempos de sobrevivência das pessoas com HIV positivo.

6.2.2 Método Bayesiano para o caso exponencial

Assumir o modelo log-linear dado em (6.11) que é dado por:

$$y_i = \ln(t_i) = \beta_0 + \beta_1 idade^* + \beta_2 droga + z_i, \quad (6.16)$$

onde z_i é uma variável aleatória com distribuição valor extremo padrão (2.10).

A função de verossimilhança para $\beta' = (\beta_0, \beta_1, \beta_2)$ é dada por:

$$L(\beta) = \exp \left\{ \sum_{i=1}^n \delta_i z_i - \sum_{i=1}^n e^{z_i} \right\} \quad (6.17)$$

onde $z_i = y_i - \beta_0 - \beta_1 idade^* - \beta_2 droga$.

Assim,

$$\begin{aligned} L(\beta) &= \exp \left\{ \sum_{i=1}^n \delta_i (y_i - \beta_0 - \beta_1 idade^* - \beta_2 droga) \right\} \\ &\times \exp \left\{ - \sum_{i=1}^n e^{y_i - \beta_0 - \beta_1 idade^* - \beta_2 droga} \right\}. \end{aligned} \quad (6.18)$$

Assumindo as densidades *a priori* normais para β_l , $l = 0, 1$ dadas em (4.12), as distribuições condicionais *a posteriori* necessárias para o algoritmo *Gibbs Sampling* para esta aplicação são dadas por:

$$(a) \pi(\beta_0 | \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) \propto \exp \left\{ -\frac{1}{2d_0^2} (\beta_0 - c_0)^2 \right\} \Psi_1(\beta_0, \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) \quad (6.19)$$

onde:

$$\Psi_1(\beta_0, \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) = \exp \left\{ -\beta_0 \sum_{i=1}^n \delta_i - e^{-\beta_0} \sum_{i=1}^n t_i e^{-\beta_1 idade^* - \beta_2 droga} \right\};$$

$$(b) \pi(\beta_1 | \beta_0, \beta_2, \mathbf{t}, \mathbf{x}) \propto \exp \left\{ -\frac{1}{2d_l^2} (\beta_l - c_l)^2 \right\} \Psi_2(\beta_0, \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) \quad (6.20)$$

onde:

$$\Psi_2(\beta_0, \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) = \exp \left\{ -\beta_1 \sum_{i=1}^n idade_i^* - e^{-\beta_0} \sum_{i=1}^n t_i e^{-\beta_1 idade_i^* - \beta_2 droga_i} \right\}.$$

$$(c) \pi(\beta_2 | \beta_0, \beta_1, \mathbf{t}, \mathbf{x}) \propto \exp \left\{ -\frac{1}{2d_l^2} (\beta_l - c_l)^2 \right\} \Psi_3(\beta_0, \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) \quad (6.21)$$

onde:

$$\Psi_3(\beta_0, \beta_1, \beta_2, \mathbf{t}, \mathbf{x}) = \exp \left\{ -\beta_2 \sum_{i=1}^n droga_i - e^{-\beta_0} \sum_{i=1}^n t_i e^{-\beta_1 idade_i^* - \beta_2 droga_i} \right\}.$$

Para a análise Bayesiana tradicional do modelo de regressão exponencial, onde $\lambda_i = \exp(-\beta_0 - \beta_1 idade_i^* - \beta_2 droga_i)$, assumimos densidades *a priori* normais $N(c_l, d_l^2)$ para β_l , $l = 0, 1, 2$.

Assumindo $c_l = 1$ e $d_l = 0.001$, logo temos densidades *a priori* não-informativas para os parâmetros de interesse, pois $E(\beta_l) = 1$ e $Var(\beta_l) = 1000$ para $l = 0, 1, 2$.

Os resumos *a posteriori* de interesse foram obtidos a partir de 31000 amostras *Gibbs* geradas das distribuições condicionais *a posteriori*, onde descartamos as 1000 primeiras amostras (*burn-in samples*), os resultados são apresentados na Tabela 6.10.

Tabela 6.10- Resumos *a posteriori* do modelo log-linear assumindo uma distribuição exponencial.

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
β_0	2.6500	0.1413	(2.3820; 2.9350)
β_1	-0.0842	0.0145	(-0.0843; -0.0557)
β_2	-1.0690	0.2010	(-1.4620; -0.6717)

Assumindo o modelo de regressão exponencial (2.1), onde λ_i tem uma distribuição gama (4.3) com $\lambda_0(t_i) = \exp(\alpha_i)$, $idade_i^* = idade_i - \overline{idade}$, para $i = 1, \dots, 100$ e as

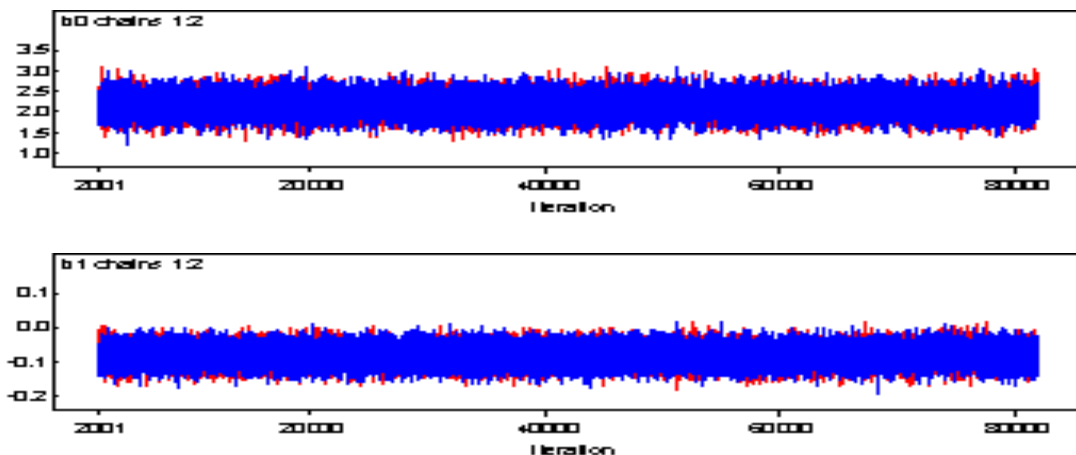
densidades *a priori* (4.12) com $a = 10$, $b = 0.1$, $c_l = 1$, $d_l = 0.001$, $l = 0, 1, 2$. Obtemos as inferências Bayesianas usando o algoritmo *Gibbs* (ver por exemplo, Casella e George, 1992) com *Metropolis-Hastings* (ver por exemplo, Chib e Greenberg, 1995).

Os resumos *a posteriori* para β_0 , β_1 , β_2 e σ_α^2 são apresentados na Tabela 6.11. Consideramos duas cadeias distintas com diferentes valores iniciais, cada uma com 80000 iterações da qual descartamos as 2000 primeiras amostras *Gibbs* geradas (*burn-in samples*).

Os valores das médias *a posteriori* para β_0 , β_1 , β_2 estão próximos aos valores obtidos pelo método clássico.

Tabela 6.11- Resumos *a posteriori* do modelo log-linear assumindo uma distribuição de exponencial e um processo gama para as taxas de falha.

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
β_0	2.1780	0.2210	(1.7420; 2.6100)
β_1	-0.0819	0.0233	(-0.1272; -0.0363)
β_2	-0.8931	0.3127	(-1.5030; -0.2767)
σ_α^2	0.1024	0.0165	(0.0759; 0.1401)



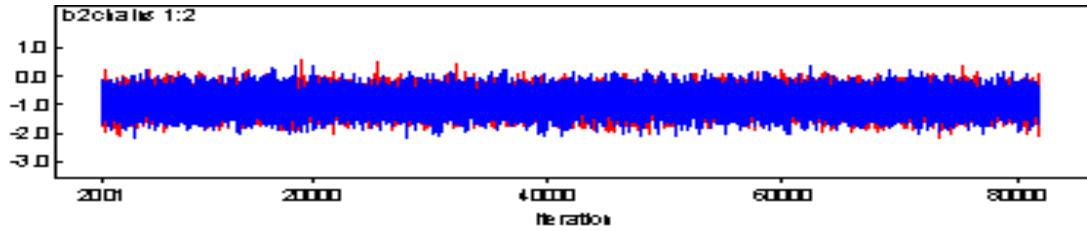


Figura 6.8- Diagnóstico de convergência para o vetor de parâmetros β .

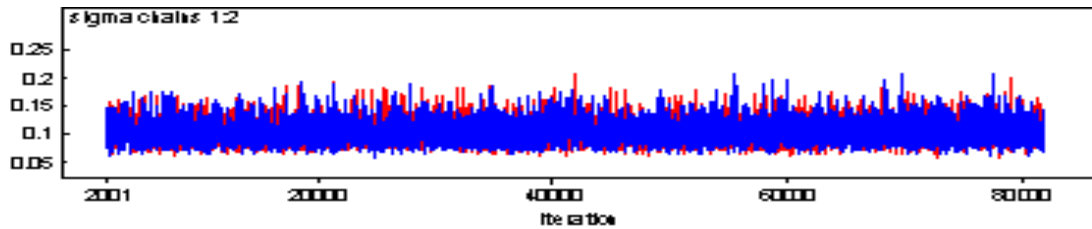


Figura 6.9- Diagnóstico de convergência para o parâmetro σ_α^2 .

Observamos a convergência das cadeias para todos os parâmetros de interesse, usando procedimentos de diagnósticos no pacote CODA (*Convergence Diagnostic e Output Analysis*).

A Figura 6.5 mostra os plots de c_i versus i , onde c_i refere-se as densidades preditivas descrita na seção 4.4. Consideramos o modelo (4.1) com λ_i dado por (4.2), representado por (\bullet) e também consideramos o modelo (4.1) com λ_i dado por (4.3) representado por ($*$).

Observamos que o uso do processo gama para as taxas de falha, melhora o ajuste do modelo de regressão exponencial, desde que em média. Os valores preditivos são grandes em geral do que usando o modelo log-linear.

Visto que $\sum_{i=1}^n \hat{c}_i = 6.59$ para o modelo log-linear tradicional dado em (6.2) e

$\sum_{i=1}^n \hat{c}_i = 7.6633$ para o modelo exponencial com um processo gama para as taxas de falha λ_i , podemos concluir que existem indícios de que o uso do processo gama para dados de sobrevivência apresenta inferências melhores do que utilizando os métodos Bayesianos tradicionais.

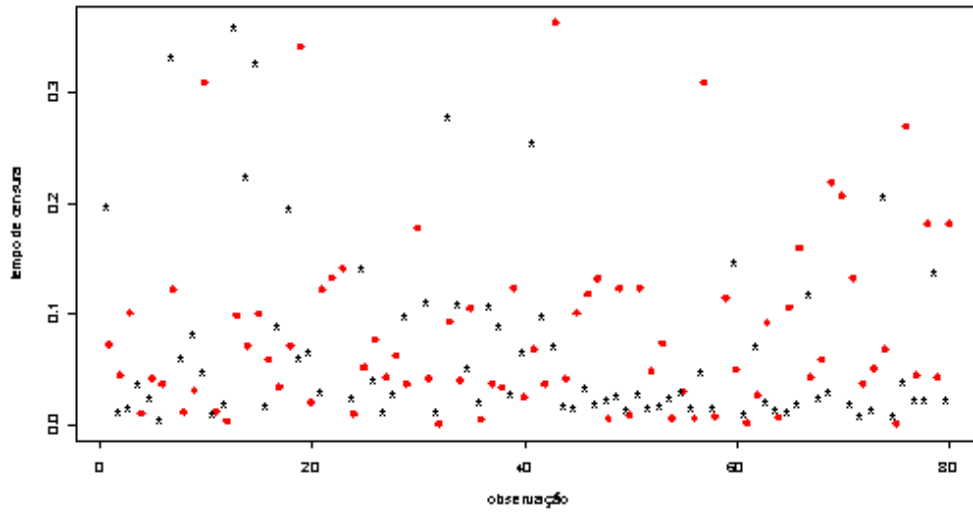


Figura 6.10- Valores das densidades preditivas considerando o modelo de regressão exponencial.

(●: modelo log-linear; *: uso do processo gama).

Capítulo 7

Conceitos Básicos de Testes Acelerados em Controle de Qualidade

Neste capítulo, apresentamos a teoria básica de testes de vida acelerados que engloba conceitos como falha, tipos de censura e conceitos específicos somente utilizados nestes tipos de testes.

Os testes acelerados consistem na indução da falha do componente e/ou equipamento por meio de um planejamento adequado, pois a análise da confiabilidade de um produto em situações normais de operação, pode ser cara e demorada.

O procedimento é baseado submetendo as unidades do produto de interesse a níveis de estresse mais altos ou mais baixos que aqueles em condições normais de uso (ver por exemplo, Nelson, 1990).

O principal interesse na área de confiabilidade é fazer inferências sobre a taxa de falha em condições normais de uso baseadas em dados obtidos em condições aceleradas. Na prática, isto significa que as informações referentes a confiabilidade dos produtos necessitam ser obtidas em um período curto de tempo para serem utilizadas em novos projetos e na melhoria dos existentes.

7.1 Variáveis de Estresse e Tipos de testes acelerados

Variável de estresse é aquela que utilizada num teste em níveis diferentes daqueles considerados em *condições normais de uso* encurta o tempo até a falha do produto ou material em um teste acelerado.

Os testes acelerados podem ser divididos em dois grupos, sendo esta classificação feita basicamente em função do tipo de dados obtidos a partir destes testes (ver por exemplo, Nelson, 1990, Borges *et al.*, 1996):

- testes de *vida* acelerados: são aqueles onde a resposta de interesse é o “tempo até a ocorrência da falha”. O objetivo neste caso é estimar o tempo médio até a falha (MTTF) e o tempo médio entre falhas (MTBF), a função de confiabilidade, etc. Através da modelagem de tais dados, podemos com certo cuidado extrapolar os resultados obtidos nos testes para as condições normais de uso.

- testes de *degradação* acelerados: neste caso, a resposta de interesse é alguma medida da performance do produto ou componente (resistência à tração, oxidação, etc.) medida ao longo do tempo. O objetivo aqui é estudar a degradação da performance do produto ao longo do tempo e utilizar esta informação para estimar a distribuição do tempo de vida do produto.

Considerando um tipo especial de produto como as memórias de computador do tipo RAM “*random access memory*”, uma das medidas da performance desse produto é o *tempo de acesso* e a “falha” neste caso poderia ser definida como o “tempo de acesso superior a um valor pré-determinado”. Portanto, num teste de vida acelerado para este produto a resposta de interesse seria o “tempo necessário para que o tempo de acesso da memória se tornasse superior ao valor pré-estabelecido”. Estes dados seriam coletados sob alguma condição estressante como por exemplo temperatura mais alta.

No caso dos testes de *degradação* estaria sendo observado a *degradação* desta característica do produto ao longo do tempo. Em outras palavras, o interesse é medir periodicamente o tempo de acesso da memória do tipo RAM sob alguma condição estressante.

Para este dispositivo a tensão nominal de uso é de 5 volts, e quando utilizada a níveis mais elevados, a tensão acelera a degradação do produto. Portanto, *tensão* é uma *variável de estresse* para estas memórias. Para o caso de um produto alimentício, em geral, temperatura e umidade são variáveis de estresse.

7.2 Formas de aceleração de testes

Basicamente acelera-se um teste submetendo as unidades a níveis mais altos de estresse do que aqueles encontrados nas condições normais de uso. As formas de aceleração podem ser divididas de acordo com dois tipos de variáveis de estresse (ver por exemplo, Nelson, 1990):

- *aceleração por alta taxa de uso*: neste caso, coloca-se o produto em operação com uma taxa de uso mais elevada do que a normal. Isto pode ser determinado de duas formas:

- 1) velocidade mais alta: colocamos uma unidade em funcionamento a uma velocidade mais alta do que ela funcionaria normalmente;

- 2) redução do tempo de descanso: o caso do mecanismo automático de acionamento de vidro é um exemplo típico. Deixamos o mecanismo funcionando por um certo número de ciclos bem maior do que o número diário esperado. Um outro exemplo é uma máquina de lavar roupas, pois em uma resistência a condição normal de uso é que esta fique em funcionamento uma ou duas horas por dia, o teste poderá ser submetido colocando a resistência da máquina de lavar para funcionar 24 horas seguidas.

- *aceleração por altos níveis de estresse*: consiste em colocar o produto em uso submetido a níveis altos das variáveis de estresse com o objetivo de encurtar o tempo de vida ou degradar a performance do produto de maneira mais rápida. Neste caso, trata-se de variáveis de estresse como temperatura, voltagem, vibração, entre outras.

Para alguns tipos de materiais e produtos temos definidas as variáveis de estresse mais comuns. Alguns exemplos são apresentados nas Tabelas 7.1 e 7.2 a seguir.

Tabela 7.1. Exemplos de Materiais; Medidas de Performance e Variáveis de Estresse.

Materiais		
Tipo	Medida de Performance	Variável de Estresse
1. metais	trinca, corrosão, oxidação	temperatura, umidade, sal
2. dielétricos e isolamentos	tempo até a falha, alongamento	temperatura, voltagem, vibração
3. alimentos e drogas	tempo de estocagem, pH reações químicas específicas	temperatura, umidade, radiação solar
4. plásticos	propriedades mecânicas, firmeza da cor	temperatura, vibração, choque

Tabela 7.2. Exemplos de Produtos; Medidas de Performance e Variáveis de Estresse.

Produtos		
Tipo	Medida de Performance	Variável de Estresse
1. semicondutores e componentes microeletrônicos	trinca, corrosão, oxidação	temperatura, corrente, voltagem, umidade, pressão
2. capacitores	tempo até a falha, alongamento	temperatura, voltagem, vibração
3. resistores	tempo de estocagem, pH reações químicas específicas	temperatura, umidade, radiação solar
4. contatos elétricos	propriedades mecânicas, firmeza da cor	temperatura, vibração, choque
5. lâmpadas	tempo até a falha, eficiência, luminosidade	voltagem, temperatura, choque (elétrico ou mecânico)

7.3 Formas de Aplicação da Carga

A carga de estresse num teste acelerado pode ser aplicada de várias formas, mas os níveis de stress só se aplicam para o segundo tipo de variável de estresse discutido na seção anterior.

A discussão será feita em termos de uma variável de estresse, mas os conceitos podem ser estendidos para o caso da utilização de combinação de variáveis (ver por exemplo, Nelson, 1990, Borges *et al.*, 1996).

· (A) *Constante*: é o tipo mais comum, onde cada unidade é submetida a um nível de estresse constante. Para fazer alguma modelagem dos dados obtidos é necessário utilizar vários níveis de estresse. No entanto, uma vez fixado o nível de estresse para uma determinada unidade, esta ficará submetida a este nível até o final do teste. As vantagens desta forma de aplicação de estresse vão desde a simplicidade na realização do teste até o uso de modelos simples na análise dos resultados. Os modelos com base em estresse constante estão bem definidos na literatura e empiricamente verificados para alguns materiais e produtos, que em uso, estão submetidos a níveis de estresse constante, estas razões levam a recomendar esta forma de aplicação da carga de estresse.

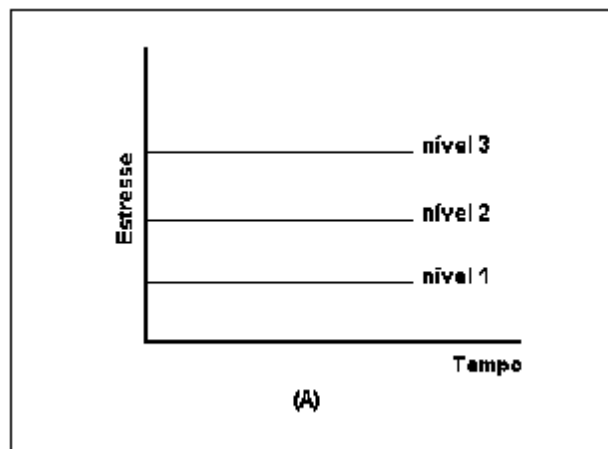


Figura 7.1- Forma de aplicação da carga de estresse do tipo constante.

· (B) *Escada (step stress)*: cada unidade é submetida a um nível de estresse por um período de tempo. Se a unidade não falhar, o nível de estresse é elevado para um novo patamar e o procedimento se repete. Portanto, nesta forma de aplicação da carga o nível de estresse é aumentado em pequenos “degraus” até que ocorra a falha. A principal vantagem do *step stress* é levar rapidamente à ocorrência de falhas e a maior desvantagem está na estimação das medidas de confiabilidade. A maioria dos produtos em condições normais de uso é submetida a uma carga constante e não do tipo *escada*. Portanto, o modelo deve de maneira adequada levar em consideração o efeito cumulativo da exposição a estresses sucessivos. O modelo a ser utilizado é mais complexo do que aquele para estresse constante. Uma outra desvantagem do *step stress* é que os modos de falha que ocorrem nos patamares mais altos, em geral, diferem daqueles que ocorrem em condições normais de uso.

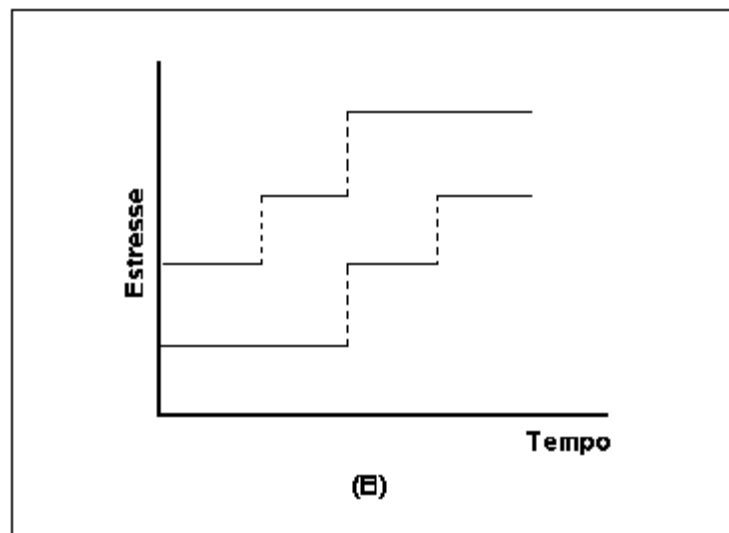


Figura 7.2- Forma de aplicação da carga de estresse do tipo escada.

· (C) *Progressivo*: cada unidade é submetida a um nível crescente de estresse, porém esse aumento não é feito em patamares, mas progressivamente. O uso de testes com estresse progressivo apresenta as mesmas vantagens e desvantagens do tipo *escada*. Além

disso, pode ser difícil em um teste acelerado controlar o aumento progressivo do nível de estresse.

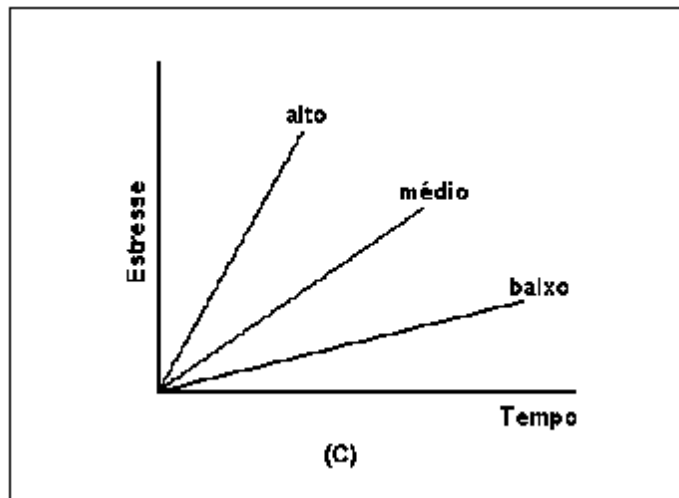


Figura 7.3- Forma de aplicação da carga de estresse do tipo progressivo.

· (D) *Cíclico*: cada unidade é submetida a estresse de níveis alto e baixo de maneira cíclica.

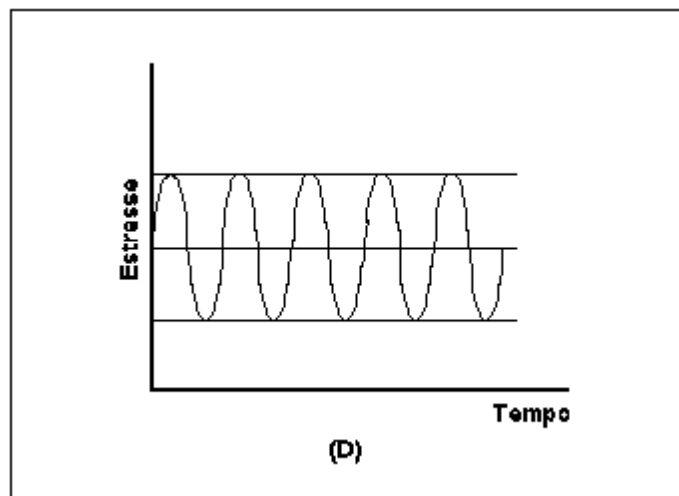


Figura 7.4- Forma de aplicação da carga de estresse do tipo cíclico.

· *Aleatório*: alguns produtos quando em uso, estão sujeitos a níveis de estresse que se alteram de maneira aleatória. Exemplos incluem componentes de pontes e elevados, e componentes estruturais de aviões. Num teste acelerado, os níveis de estresse serão aplicados seguindo uma distribuição o mais próxima possível da que ocorre na realidade, porém a níveis mais elevados.

A motivação principal para a utilização de estresses do tipo *step* ou *progressivo* está na possibilidade de estressar ainda mais as unidades sobreviventes e fazer com que todas ou quase todas falhem mais rapidamente. Em particular, quando o modo de falha e o efeito de aceleração da variável de estresse são bem entendidos, e quando se deseja fazer inferências para um ambiente cujo o estresse é aproximadamente constante, utilizar testes com outra forma de aplicação de estresse que não seja a forma de aplicação de carga do tipo constante, significa complicar desnecessariamente a situação. Entretanto, em outros casos é necessário realizar testes com aplicação de *step-stress* para identificar e fazer inferências sobre as constantes de um modelo cinético.

7.4 Considerações Importantes para a Realização de um Teste Acelerado

A validade e precisão das informações provenientes de um teste acelerado dependem do cuidado dispensado ao planejamento e a realização do mesmo. Este trabalho deve ser resolvido conjuntamente por Engenheiros e Estatísticos.

A seguir temos alguns elementos importantes que bem definidos garantem a validade e precisão das informações. Enfatizamos aqui os testes acelerados para estimação das medidas de confiabilidade do produto em condições de projeto.

· **Objetivo**: testes acelerados são realizados com vários objetivos. Os mais comuns incluem (ver por exemplo, Nelson, 1990, Borges *et al.*, 1996):

- identificar falhas de projeto: eliminá-las ou reduzi-las através de redundâncias, melhores componentes, etc;

- *burn in*: uma etapa do processo de manufatura que visa eliminar falhas que ocorrem logo no início do período de vida útil do produto;

- obter medidas de confiabilidade: para determinar se um projeto deve ser liberado para a fabricação ou se o produto deve ser liberado para um cliente. Estimar custos de garantia e serviço, taxas de falha, tempo médio até a falha (MTTF), taxas de degradação, etc;

- controle de qualidade: monitorar a confiabilidade do produto e tomar ações corretivas conforme necessário, por exemplo, quando um novo modo de falha aparece ou quando a vida ou performance se degradam.

Um teste pode ter um ou mais destes objetivos, que dizem respeito a projeto, fabricação, controle de qualidade, marketing, serviço de assistência técnica, etc. O importante é definir de maneira clara os objetivos do teste antes deste ser realizado, uma vez que ele determinará os cuidados a serem seguidos durante o planejamento e sua realização.

· **Escolha da medida de performance:** para muitos produtos existem normas técnicas que definem quais são as medidas de performance a serem utilizadas e como medir estes produtos (veja Tabelas 7.1 e 7.2). Em terminologia estatística, a medida de performance é chamada de *variável resposta*.

· **Definição de “falha”:** É importante que fabricante e consumidor concordem antecipadamente com a definição da *falha* para evitar problemas futuros.

· **Determinação de Condições Realistas de Teste:** o ideal é que um teste acelerado tente ao máximo simular as condições reais de uso, exceto pelos níveis utilizados da variável de estresse. Apesar dos esforços, muitos testes diferem das condições reais de uso. Os engenheiros são então levados a assumir que um produto cuja performance é boa nos testes, terá boa performance em uso.

· **Definição das Variáveis de Estresse:** aqui deparamos com o fato de que para muitos produtos, existem padrões para os testes e para os estresses de aceleração, ver

exemplos de variáveis de estresse nas Tabelas 7.1 e 7.2. Para outros produtos pode não haver padrão estabelecido, neste caso, é preciso determinar através de experimentação, os estresses apropriados. Há casos ainda em que pode ser conveniente e necessário utilizar mais do que uma variável de estresse, o importante é reconhecer qual estresse de fato acelera a vida do produto.

- **Definição da Forma de Aplicação da Carga de Estresse:** a escolha da forma de aplicação da carga de estresse (*constante*, tipo *escada*, etc) depende de como o produto é sobrecarregado nas condições de uso e de limitações práticas e teóricas.

- **Escolha do Plano Experimental:** um plano experimental é formado pelos três elementos que serão citados a seguir. Portanto, *escolher um plano experimental* significa:

- **determinar os níveis de estresse:** os níveis de estresse que serão utilizados não devem ser tão altos que produzam outros modos de falha que raramente ocorrem nas condições de projeto.

- **determinar o tamanho da amostra:** quantas unidades *ao todo* serão submetidas a teste.

- **determinar a proporção de alocação:** dada uma amostra de n unidades e o número de níveis de estresse, quantas unidades deverão ser submetidas a teste em cada nível.

A determinação destes 3 elementos e conseqüentemente do plano experimental deve ser feita em função da precisão desejada para os estimadores das quantidades de interesse.

- **Outras Variáveis:** além da variável de estresse, existem outras variáveis que afetam a vida do produto. O primeiro tipo consiste das *variáveis experimentais*, estas variáveis são controláveis e são investigadas em vários níveis com o objetivo de estudar seus efeitos. O outro tipo consiste das variáveis *constantes*, estas são mantidas constante durante o teste como por exemplo, manter o mesmo lote, mesmo operador, etc. Existem também as variáveis *não controláveis* mas *observáveis*. Seus valores são anotados para que se aprenda a respeito do seu efeito. Finalmente, o último tipo consiste das variáveis *não controláveis* e *não observáveis*. A aleatorização evita que o efeito destas variáveis se confunda com o

efeito do estresse.

· **Análise Estatística:** a análise estatística envolve a escolha de um modelo matemático apropriado para a realização de inferências a respeito da população em estudo. A escolha de um modelo deve ser baseada na teoria e experiência da engenharia. Um modelo para um teste de vida acelerado consiste na escolha de uma distribuição probabilística para o tempo de falha e de uma relação estresse-resposta. Existem alguns candidatos para a distribuição probabilística e para a relação estresse-resposta que são muito utilizados e conhecidos na literatura. Após a escolha de um possível modelo, os dados experimentais são utilizados inicialmente para verificar sua adequação. A validade do modelo deve ser verificada através de técnicas estatísticas. Caso necessário, o modelo deve ser modificado até que um adequado seja obtido. As conclusões são extraídas do modelo final, ajustado para responder aos objetivos do teste. Estas conclusões envolvem estimação pontual e por intervalos, além de testes de hipóteses.

Capítulo 8

Relações Estresse-Resposta: Arrhenius, Eyring e Potência Inversa

Neste capítulo, apresentamos as relações estresse-resposta mais utilizadas e conhecidas na literatura que foram desenvolvidas por estatísticos, matemáticos e engenheiros. Estas relações permitem que o analista utilize uma função densidade de probabilidade associada a uma função determinística que permite extrapolar as condições usuais de teste.

Em geral, a distribuição de sobrevivência de um determinado componente e/ou equipamento descreve matematicamente sua duração, onde existe a presença de diversas variáveis físicas que podem ser responsáveis por sua falha.

Considerar todas as possíveis variáveis físicas responsáveis pela falha é uma escolha totalmente inviável, mas este problema é facilmente solucionado com a utilização das relações estresse-resposta.

As relações estresse-resposta são baseadas em conceitos de mecânica estatística, como energia de ativação e teoria cinética. As relações mais comuns e utilizadas na literatura são: os modelos de Arrhenius e Eyring, que são muito aplicados em componentes eletrônicos. E também destaca-se o Modelo de Lei de Potência Inversa também conhecido como Modelo de Potência Inversa, onde este é muito comum na indústria e na engenharia em aplicações de capacitores elétricos (ver por exemplo, Lawless, 1982).

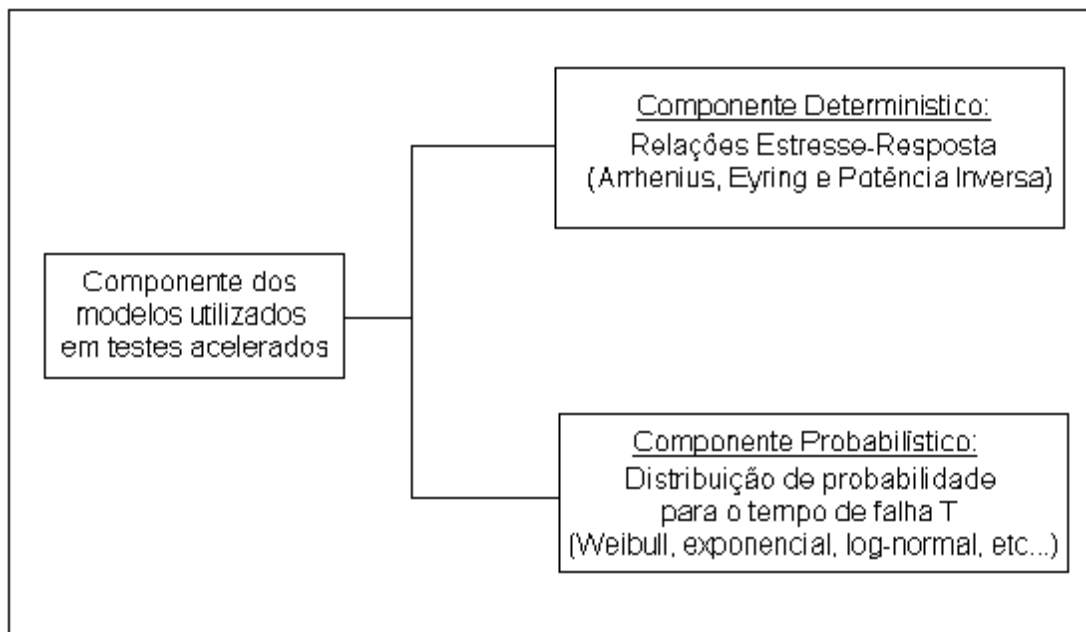


Figura 8.1- Componentes dos modelos utilizados em análise de testes de vida acelerados.

8.1 Relação de Arrhenius

O modelo de Arrhenius provavelmente é a relação estresse-resposta mais conhecida em testes de vida acelerados e intensivamente utilizado quando a variável de aceleração é do tipo térmica, ou seja, a variável de estresse é a temperatura

Esta relação é baseada na Lei de Arrhenius, proposta pelo físico-químico sueco Svandte Arrhenius em 1887. De acordo com esta lei, a taxa de uma reação química depende da temperatura segundo a seguinte expressão (ver por exemplo, Borges *et al.*, 1996):

$$R(T) = A \exp \left\{ \frac{-E_A}{kT} \right\}, \quad (8.1)$$

onde:

R : é a velocidade de reação;

A : é uma constante que é característica do mecanismo de falha do produto e das condições de teste;

E_A : é a energia de ativação, normalmente em elétron-volts;

k : é a constante de Boltzman: 8.617385×10^{-5} por K (Kelvin) e

τ : é a temperatura absoluta Kelvin: $273.16 + C^0$.

A energia de ativação é a energia que uma molécula tem que ter para participar na reação, ou seja, a energia de ativação é uma medida do efeito que a temperatura tem sobre a reação.

Assim podemos explicar a falha devido a uma reação química ou difusão através de uma relação simples. O produto falha quando uma quantidade crítica do produto químico reage, logo uma maneira de descrever esta situação é dada por (ver por exemplo, Borges *et al.*, 1996):

$$\text{Quantidade crítica} = \text{Taxa} \times \text{Tempo até a falha} \quad (8.2)$$

ou equivalentemente,

$$\text{Tempo até a falha} = \frac{\text{Quantidade crítica}}{\text{Taxa}}. \quad (8.3)$$

Isto significa que o Modelo Estresse-Resposta de Arrhenius é formulado assumindo que o tempo até a falha é inversamente proporcional à taxa dada em (8.3).

Substituindo (8.1) em (8.3) obtemos a Relação de Arrhenius dada por:

$$\begin{aligned} T &= \frac{\text{Quantidade crítica}}{A' \exp \left\{ -\frac{E}{k\tau} \right\}} \\ &= A \exp \left\{ \frac{E}{k\tau} \right\}, \end{aligned} \quad (8.4)$$

onde $A = \frac{\text{Quantidade crítica}}{A'}$.

A forma linearizada da relação de Arrhenius pode ser obtida aplicando o logaritmo em ambos lados da expressão dada em (8.4). Logo,

$$\log T = \log A + \frac{E}{k\tau}. \quad (8.5)$$

De uma forma equivalente vamos reescrever a expressão dada em (8.5) como,

$$\log T = \beta_0 + \beta_1 x, \quad (8.6)$$

onde:

$$\beta_0 = \log A;$$

$$\beta_1 = \frac{E}{k} \text{ e}$$

$$x = \frac{1}{\tau}.$$

Notemos que a variável x é o inverso da temperatura absoluta. Geralmente utilizamos $x = \frac{1000}{\tau}$ com a finalidade de evitar problemas computacionais provenientes do uso de números muito pequenos.

Portanto, podemos concluir que a relação estresse-resposta de Arrhenius especifica que o logaritmo do tempo de falha, é uma função linear do inverso da temperatura absoluta.

Muitos especialistas utilizam o termo fator de aceleração característica para descrever a taxa de vida ou aceleração característica entre o nível usual e um nível de estresse mais elevado ou,

$$\mathcal{A}_c = \frac{T}{T'},$$

da relação de Arrhenius dado em (8.4), podemos derivar esta função em relação a τ para o tempo de falha T e em relação a τ' para o tempo de falha T' .

Logo, o fator de aceleração para a relação de Arrhenius é dado por (ver por exemplo, Borges *et al.*, 1996):

$$\mathcal{A}_c = \exp \left\{ \frac{E}{k} \left[\frac{1}{\tau} - \frac{1}{\tau'} \right] \right\}. \quad (8.7)$$

8.2 Relação de Eyring

O modelo de Eyring foi formulado nos princípios de mecânica quântica, que é frequentemente usado quando a tensão térmica (temperatura) é a variável de aceleração. Porém, a relação de Eyring também é usada frequentemente para outros tipos de variáveis de tensão, como por exemplo umidade.

A relação de Eyring descrita aqui em uma forma simplificada envolve somente uma variável de estresse temperatura. Mas, esta também pode ser descrita em uma forma generalizada, onde a taxa de falha λ , é expressa em função de duas variáveis de estresse (ver por exemplo, Mann, Schaffer e Singpurwalla, 1974).

Considerando apenas uma variável de estresse que expressa a variação da taxa de falha como uma função da temperatura de operação, a relação de Eyring é determinada por (ver por exemplo, Achcar e Louzada-Neto, 1993):

$$\lambda_i = V_i \exp \left\{ \alpha - \frac{\beta}{V_i} \right\}, \quad (8.8)$$

onde:

$$-\infty < \alpha, \beta < \infty;$$

V : é o i -ésimo nível da variável de estresse temperatura V (a temperatura é avaliada em unidades absolutas, isto é, graus Kelvin ou graus Rankine) e

α e β são os parâmetros desconhecidos a serem estimados.

A relação de Eyring é semelhante a relação de Arrhenius, onde podemos observar esta semelhança com maior facilidade se reescrevermos a expressão dada em (8.8) na forma:

$$\begin{aligned}
\lambda_i &= V_i \exp \left\{ \alpha - \frac{\beta}{V_i} \right\} \\
&= V_i e^\alpha \exp \left\{ -\frac{\beta}{V_i} \right\} \\
&= V_i A \exp \left\{ -\frac{\beta}{V_i} \right\},
\end{aligned} \tag{8.9}$$

onde A é uma constante.

A relação de Arrhenius é dada por:

$$R(T) = A \exp \left\{ \frac{-E_A}{k\tau} \right\}.$$

Logo, comparando a expressão em (8.9) com a relação de Arrhenius, podemos ver que a única diferença é o termo que representa a variável de estresse, em geral, utilizando ambas as relações obtemos resultados semelhantes.

Para o modelo de Eyring o fator de aceleração para descrever a taxa de vida é determinado por:

$$\mathcal{A}_c = \frac{V}{V'},$$

logo da relação de Eyring dada em (8.8) temos:

$$\begin{aligned}
\mathcal{A}_c &= \frac{V \exp \left\{ \alpha - \frac{\beta}{V} \right\}}{V' \exp \left\{ \alpha - \frac{\beta}{V'} \right\}} \\
&= \frac{V}{V'} \exp \left\{ a - \frac{\beta}{V} \right\} \exp \left\{ - \left(\alpha - \frac{\beta}{V'} \right) \right\} \\
&= \frac{V}{V'} \exp \left\{ -\frac{\beta}{V} + \frac{\beta}{V'} \right\}.
\end{aligned} \tag{8.10}$$

Logo, o fator de aceleração é dado por:

$$\mathcal{A}_c = \frac{V}{V'} \exp \left\{ \beta \left(\frac{1}{V'} - \frac{1}{V} \right) \right\}. \quad (8.11)$$

8.3 Relação de Potência Inversa

O modelo de Lei de Potência Inversa mais conhecido na literatura como modelo de Potência Inversa é utilizado para modelar o tempo de falha em função de qualquer tipo de variável de estresse. Geralmente este modelo é usado para tensões aceleradas não-térmicas, como por exemplo em aplicações que incluem isolantes, dielétricos, fadiga de metais, lâmpadas incandescentes, etc.

Suponha que a variável de estresse V seja sempre positiva, o que acontece na maior parte das situações. Então, a relação de potência inversa é determinada por:

$$T = \frac{A}{V^{\beta_1}} \quad (8.12)$$

onde:

T : é o tempo de falha;

A e β_1 : são parâmetros característicos do produto, unidade, geometria, fabricação, método de teste, etc.

A forma linearizada do modelo de potência inversa dado na expressão (8.12) é dada por:

$$\begin{aligned} \log(T) &= \beta_0 + \beta_1 [-\log(V)] \\ &= \beta_0 + \beta_1 x. \end{aligned}$$

Logo, a forma linearizada do modelo é dada por:

$$\log(T) = \beta_0 + \beta_1 x. \quad (8.13)$$

O fator de aceleração é obtido derivando a expressão dada em (8.6) em relação ao tempo de falha T no nível de estresse V e o tempo de falha T' no nível de estresse de referência V' , ou seja,

$$\mathcal{A}_c = \left(\frac{V'}{V} \right)^{\beta_1}.$$

Das expressões (8.6) e (8.13), podemos derivar uma forma geral para a relação estresse-resposta dada por (ver por exemplo, Borges *et al.*, 1996):

$$\log(T) = \beta_0 + \beta_1 x, \quad (8.14)$$

onde:

se $x = \frac{1}{\tau}$, onde τ é a temperatura absoluta, então temos a relação de Arrhenius e

se $x = -\log(V)$, onde V é uma variável de estresse qualquer, então temos a relação de Potência Inversa.

A relação de Arrhenius é indicada quando a variável de estresse é a temperatura. Entretanto, na prática os exemplos de utilização desta relação não se limitam somente a este tipo de variável de estresse.

Capítulo 9

Uma Análise Bayesiana do Modelo de Potência Inversa Assumindo uma Distribuição Exponencial

Neste capítulo, apresentamos uma análise Bayesiana para dados de sobrevivência que provém de uma distribuição exponencial relacionada através do parâmetro λ , onde consideramos o Modelo de Lei de Potência Inversa como o componente determinístico para o modelo proposto.

Para a realização da abordagem Bayesiana consideramos a teoria dos testes de vida acelerados, assumindo um processo gama para as taxas de falha λ_i do modelo proposto. Um efeito aleatório extra foi introduzido para um grupo de unidades em cada nível de estresse, este efeito aleatório também conhecido na literatura como “fragilidade” é utilizado com o intuito de melhorar o ajuste do modelo para os dados de vida acelerados.

Desenvolvemos uma análise Bayesiana hierárquica para o modelo usando os métodos MCMC (*Monte Carlo Markov chain*), especificamente o algoritmo *Gibbs Sampling* (ver por exemplo, Gelfand e Smith, 1990) com *Metropolis-Hastings* (ver por exemplo, Smith e Roberts, 1993). A metodologia proposta neste capítulo está ilustrada a um conjunto de dados reais.

9.1 Modelo Estatístico

Considere um teste de vida acelerado de um experimento onde são aplicados nos itens k níveis de estresse V_i , para $i = 1, 2, \dots, k$, e que o comportamento de falhas dos componentes é especificado pela distribuição exponencial.

Seja T a variável aleatória que representa os tempos de sobrevivência com (*f.d.p.*) dada por:

$$f(t_i|\lambda_i) = \lambda_i \exp(-\lambda_i T_i), \quad (9.1)$$

onde $t > 0$, $\lambda > 0$ assume uma variável de estresse X que causa mudança na taxa de falha λ .

Visto que os dados classificam-se na maior parte como um processo aleatório, quando consideramos uma falha como sendo uma interação dados/stress, é natural considerar as falhas estabelecendo um processo aleatório (Processo de Poisson).

Nesta passagem, assumimos uma família de distribuições gama para as taxas de falha λ_i considerando (9.1), isto é,

$$\lambda_i \sim \Gamma[\lambda_0(i), e^{\beta_1 X_i}], \quad (9.2)$$

ou seja, $\lambda_1, \lambda_2, \dots, \lambda_n$ são realizações de um processo gama, onde $\Gamma[a, b]$ genericamente denota uma distribuição gama com média $\frac{a}{b}$ e variância $\frac{a}{b^2}$.

Assim, temos de (9.2) que,

$$E(\lambda_i) = \lambda_0(i) e^{-\beta_1 X_i}.$$

Podemos considerar escolhas diferentes para a função de base $\lambda_0(i)$, na qual é notada uma grande flexibilidade do ajuste para os dados de sobrevivência.

Como um caso especial, assumimos $\lambda_0(i) = \exp\{-(Z_i + \beta_0)\}$.

Portanto,

$$\begin{aligned}
 E(\lambda_i) &= \lambda_0(i) e^{-\beta_1 X_i} \\
 &= \exp\{-(Z_i + \beta_0)\} e^{-\beta_1 X_i} \\
 &= \exp\{-(Z_i + \beta_0 + \beta_1 X_i)\},
 \end{aligned} \tag{9.3}$$

onde $i = 1, 2, \dots, k$.

Para esta escolha de $\lambda_0(i)$, temos o modelo estresse-resposta geral introduzido por Klein e Basu (1981), onde β_0 e β_1 são os parâmetros desconhecidos a serem estimados do modelo proposto.

Considerando os dados com k níveis de estresse V_i , para $i = 1, 2, \dots, k$, assumindo aleatoriedade e o mecanismo de censura do tipo II, ou seja, o experimento termina quando observamos r_i falhas para cada nível de stress V_i , para $i = 1, 2, \dots, k$.

Assim, com n_i unidades no início de cada teste com estresse V_i , temos as observações não-censuradas ordenadas dada por $t_{1i}, t_{2i}, \dots, t_{r_i i}$ e $(n_i - r_i)$ observações censuradas igual a $t_{r_i i}$, para $i = 1, 2, \dots, k$.

Logo, a função de verossimilhança para $\boldsymbol{\lambda}$ é dada por (ver por exemplo, Achcar e Dias, 2003):

$$L(\boldsymbol{\lambda}|\mathbf{t}, \mathbf{v}) = \prod_{i=1}^k \lambda_i \exp(-\lambda_i T_i), \tag{9.4}$$

onde $\boldsymbol{\lambda}' = (\lambda_1, \lambda_2, \dots, \lambda_k)$ e $T_i = \sum_{j=1}^{r_i} t_{ij} + (n_i - r_i) t_{r_i i}$, para $i = 1, 2, \dots, k$.

9.2 Abordagem Bayesiana do Modelo

Para o Modelo de Potência Inversa assumimos $Z_i = 0$, $\beta_0 = \log \alpha$, $\beta_1 = \beta$ e $X_i = -\log V_i$ para a expressão em (9.3). Logo,

$$E(\lambda_i) = \frac{V_i^\beta}{\alpha}.$$

Assim temos de (9.2) que,

$$\lambda_i \sim \Gamma [e^{-\beta_0}, e^{\beta_1 X_i}], \quad (9.5)$$

para $i = 1, 2, \dots, k$.

Para o segundo estágio da análise Bayesiana, assumimos as seguintes distribuições *a priori* para β_0 e β_1 :

$$(a) \beta_0 \sim N(c_0; d_0^2) \quad (9.6)$$

$$(b) \beta_1 \sim N(c_1; d_1^2) \quad (9.7)$$

onde:

c_0, d_0, c_1 e d_1 são os hiperparâmetros conhecidos e

$N(c, d^2)$ denota uma distribuição normal com média c e variância d^2 .

Assumindo independência *a priori* entre os parâmetros β_0 e β_1 , obtemos a distribuição *a posteriori* conjunta para $(\boldsymbol{\lambda}, \beta_0, \beta_1)$ dada por:

$$\begin{aligned} \pi(\boldsymbol{\lambda}, \beta_0, \beta_1 | \mathbf{t}, \mathbf{v}) &\propto \left\{ \prod_{i=1}^k \lambda_i^{r_i} \exp(-\lambda_i T_i) \right\} \\ &\times \left\{ \prod_{i=1}^k \frac{(e^{\beta_1 X_i})^{e^{-\beta_0}}}{\Gamma[e^{-\beta_0}]} \lambda_i^{\exp(\beta_0)-1} \exp(-\lambda_i e^{\beta_1 X_i}) \right\} \\ &\times \exp \left\{ -\frac{1}{2d_0^2} (\beta_0 - c_0)^2 \right\} \exp \left\{ -\frac{1}{2d_1^2} (\beta_1 - c_1)^2 \right\}. \end{aligned} \quad (9.8)$$

Os sumários *a posteriori* podem ser obtidos utilizando os métodos MCMC (*Monte Carlo Markov chain*) para a distribuição *a posteriori* conjunta dada em (9.8).

9.3 Distribuições Condicionais *a Posteriori* para λ, β_0 e β_1

Dada a distribuição *a posteriori* em (9.8), as distribuições condicionais para o algoritmo *Gibbs Sampling* são dadas por:

$$(a) \lambda_i | \theta_{(\lambda_i)}, \mathbf{t}, \mathbf{v} \sim \Gamma \left[r_i + e^{-\beta_0}; T_i + V_i^{-\beta_1} \right], \quad (9.9)$$

onde $i = 1, 2, \dots, k$ e $\theta_{(\lambda_i)}$ é o vetor de todos os parâmetros exceto λ_i .

$$(b) \pi(\beta_0 | \theta_{(\beta_0)}, \mathbf{t}, \mathbf{v}) \propto \exp \left\{ -\frac{1}{2d_0^2} (\beta_0 - c_0)^2 \right\} \psi_1(\theta), \quad (9.10)$$

onde $\psi_1(\theta) = \exp \left\{ -2 \log \Gamma(e^{-\beta_0}) + \beta_1 \exp \left(-\beta_0 \sum_{i=1}^k X_i \right) + \exp \left(-\beta_0 \sum_{i=1}^k \log \lambda_i \right) \right\}$.

$$(c) \pi(\beta_1 | \theta_{(\beta_1)}, \mathbf{t}, \mathbf{v}) \propto \exp \left\{ -\frac{1}{2d_1^2} (\beta_1 - c_1)^2 \right\} \psi_2(\theta), \quad (9.11)$$

onde $\psi_2(\theta) = \exp \left\{ -\beta_1 \exp \left(-\beta_0 \sum_{i=1}^k X_i \right) - \sum_{i=1}^k \lambda_i e^{-\beta_1 X_i} \right\}$.

Observamos que a distribuição condicional *a posteriori* para o parâmetro $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ é representada pela distribuição gama e que as distribuições condicionais *a posteriori* para os parâmetros β_0 e β_1 não apresentam formas conhecidas, esta dificuldade é superada utilizando o algoritmo de *Metropolis-Hastings*, que permite a geração de observações que provém de distribuições desconhecidas.

9.4 Introdução de Efeitos Aleatórios no Modelo

Em aplicações industriais de testes acelerados, usualmente temos uma amostra aleatória de alguns grupos de unidades manufaturadas a ser testada em um nível de estresse especí-

fico. Neste caso, podemos ter ocorrência de falhas diferentes para grupos com unidades em adição ao efeito da variável estresse.

Neste caso, necessitamos introduzir um componente aleatório extra w_i , $i = 1, 2, \dots, k$ no modelo (9.1) com λ_i dado por:

$$\lambda_i \sim \Gamma [e^{w_i - Z_i - \beta_0}; e^{\beta_1 X_i}], \quad (9.12)$$

para $i = 1, 2, \dots, k$.

Tais modelos são conhecidos na literatura como modelos de “fragilidade” e são muito utilizados em aplicações médicas (ver por exemplo, Clayton, 1991).

Considerando o caso do Modelo de Potência Inversa em que $Z_i = 0$, $\beta_0 = \log \alpha$, $\beta_1 = \beta$ e $X_i = -\log V_i$, temos:

$$\lambda_i \sim \Gamma \left[\frac{e^{w_i}}{\alpha}; \frac{1}{V_i^\beta} \right], \quad (9.13)$$

para $i = 1, 2, \dots, k$.

Isto é,

$$E(\lambda_i) = \frac{e^{w_i} V_i^\beta}{\alpha}.$$

Para o segundo estágio da análise Bayesiana hierárquica, consideramos as seguintes distribuições *a priori* para β_0 , β_1 , w_i e σ_w^2 :

$$(a) \beta_0 \sim N(c_0; d_0^2), \quad (9.14)$$

$$(b) \beta_1 \sim N(c_1; d_1^2), \quad (9.15)$$

$$(c) w_i \sim N(0, \sigma_w^2), \quad (9.16)$$

$$(d) \sigma_w^2 \sim IG[a; b], \quad (9.17)$$

onde:

$$i = 1, 2, \dots, k;$$

a, b, c_0, d_0, c_1 e d_1 são os hiperparâmetros conhecidos;

$IG(a; b)$: denota uma distribuição gama inversa com média $\frac{b}{(a-1)}$ e variância $\frac{b^2}{[(a-1)^2(a-2)]}$ e

$N(c, d^2)$ denota uma distribuição normal com média c e variância d^2 .

Assumindo independência *a priori* sobre os parâmetros, a distribuição *a posteriori* conjunta para $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_k)$, $\mathbf{w} = (w_1, w_2, \dots, w_k)$, β_0, β_1 e σ_w^2 é dada por:

$$\begin{aligned} \pi(\boldsymbol{\lambda}, \mathbf{w}, \beta_0, \beta_1, \sigma_w^2 | \mathbf{t}, \mathbf{v}) &\propto \left\{ \prod_{i=1}^k \lambda_i e^{-\lambda_i} \right\} & (9.18) \\ &\times \left\{ \prod_{i=1}^k \frac{(e^{\beta_1 X_i})^{e^{-\beta_0 + w_i - 1}}}{\Gamma[e^{-\beta_0 + w_i}]} \lambda_i^{e^{-\beta_0 + w_i - 1}} \exp(-\lambda_i e^{\beta_1 X_i}) \right\} \\ &\times \prod_{i=1}^k \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left\{-\frac{w_i^2}{2\sigma_w^2}\right\} (\sigma_w^2)^{a-1} \exp\left\{-\frac{d}{\sigma_w^2}\right\} \\ &\times \exp\left\{-\frac{1}{2d_0^2}(\beta_0 - c_0)^2\right\} \exp\left\{-\frac{1}{2d_1^2}(\beta_1 - c_1)^2\right\}. \end{aligned}$$

onde $\boldsymbol{\lambda}' = (\lambda_1, \lambda_2, \dots, \lambda_k)$ e $\mathbf{w}' = (w_1, w_2, \dots, w_k)$.

9.5 Distribuições Condicionais *a Posteriori* para $\boldsymbol{\lambda}$,

$\mathbf{w}, \beta_0, \beta_1$ e σ_w^2

Dada a distribuição *a posteriori* em (9.18), as distribuições condicionais *a posteriori* necessárias para o algoritmo *Gibbs Sampling* são dadas por:

$$(a) \lambda_i | \theta_{(\lambda_i)}, \mathbf{t}, \mathbf{v} \sim \Gamma \left[r_i + e^{-\beta_0 + w_i}; T_i + V_i^{-\beta_1 X_i} \right], \quad (9.19)$$

onde $i = 1, 2, \dots, k$ e $\theta_{(\lambda_i)}$ é o vetor de todos os parâmetros exceto λ_i .

$$(b) \pi(w_i | \theta_{(w_i)}, \mathbf{t}, \mathbf{v}) \propto \exp \left\{ -\frac{w_i^2}{2\sigma_w^2} \right\} \psi_3(\theta), \quad (9.20)$$

onde $\psi_3(\theta) = \exp \left\{ \beta_1 \sum_{i=1}^k X_i \exp(-\beta_0 + w_i) + \sum_{i=1}^k \exp(-\beta_0 + w_i) \log \lambda_i - \log \Gamma [e^{-\beta_0 + w_i}] \right\}$, para $i = 1, 2, \dots, k$.

$$(c) \pi(\beta_0 | \theta_{(\beta_0)}, \mathbf{t}, \mathbf{v}) \propto \exp \left\{ -\frac{1}{2d_0^2} (\beta_0 - c_0)^2 \right\} \psi_4(\theta), \quad (9.21)$$

onde $\psi_4(\theta) = \exp \left\{ \beta_1 \sum_{i=1}^k X_i \exp(-\beta_0 + w_i) + \sum_{i=1}^k \exp(-\beta_0 + w_i) - \sum_{i=1}^k \log \Gamma [e^{-\beta_0 + w_i}] \right\}$.

$$(d) \pi(\beta_1 | \theta_{(\beta_1)}, \mathbf{t}, \mathbf{v}) \propto \exp \left\{ -\frac{1}{2d_1^2} (\beta_1 - c_1)^2 \right\} \psi_5(\theta), \quad (9.22)$$

onde $\psi_5(\theta) = \exp \left\{ \beta_1 \sum_{i=1}^k X_i \exp(-\beta_0 + w_i) - \sum_{i=1}^k \lambda_i e^{\beta_1 X_i} \right\}$.

$$(e) \sigma_w^2 | \theta_{(\sigma_w^2)}, \mathbf{t}, \mathbf{v} \sim IG \left[a + \frac{k}{2}; b + \frac{1}{2} \sum_{i=1}^k w_i^2 \right]. \quad (9.23)$$

Observe que as distribuições condicionais *a posteriori* para os parâmetros $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_k)$ e σ_w^2 apresentam formas conhecidas, sendo representadas pela distribuição gama e gama inversa, respectivamente. E as distribuições condicionais *a posteriori* para os parâmetros $\mathbf{w} = (w_1, w_2, \dots, w_k)$, β_0 e β_1 não apresentam formas conhecidas, onde tal dificuldade pode ser superada utilizando o algoritmo de *Metropolis-Hastings* (ver por exemplo, Metropolis *et al.*, 1953; Hastings, 1970).

9.6 Aplicação

Na Tabela 9.1, temos os tempos de sobrevivência em horas de unidades de Kevlar/expoxy 49, submetidos a um teste de vida acelerado com dois níveis de estresse de pressão (psi) V (conjunto de dados introduzido por Achcar e Dias, 2003). No primeiro nível de estresse $V_1 = 3700$, temos $r_1 = 18$ observações não-censuradas e $n_1 - r_1 = 6$ observações censuradas do tipo II. Para o segundo nível de estresse $V_2 = 4000$, temos 24 observações não-censuradas.

Tabela 9.1- Tempos de sobrevivência em horas de unidades de Kevlar/expoxy.

i	V_i	n_i	r_i	<i>tempos de sobrevivência em horas</i>																							
1	3700	24	18	225.2,	503.6,	1087.7,	1134.3,	1824.3,	1920.1,	2383.0,	2442.5,	3708.9,	3708.9,	4908.9,	5556.0,	6271.1,	7332.0,	7918.7,	7996.0,	9240.3,	9973.0						
2	4000	24	24	19.1,	24.3,	69.8,	71.2,	136.0,	199.1,	403.7,	432.2,	453.2,	514.1,	514.2,	541.6,	544.9,	554.2,	664.5,	694.1,	876.7,	930.4,	1254.9,	1275.6,	1536.8,	1755.5,	2046.2,	6177.5

Considerando o modelo exponencial dado em (9.1) para os tempos de sobrevivência e assumindo o modelo estresse-resposta geral $\lambda_i = \exp \{-(Z_i + \beta_0 + \beta_1 X_i)\}$ para $i = 1, 2$, onde $Z_i = 0$, $\beta_0 = \log(\alpha)$, $\beta_1 = \beta$ e $X_i = \log(V_i)$, $i = 1, 2$. Logo, temos o modelo de lei de potência inversa.

Assumimos para o primeiro estágio da análise Bayesiana do conjunto de dados da Tabela 9.1 as seguintes densidades *a priori* para β_0 e β_1 :

$$\beta_0 \sim N(174, (0.001)^{-1}) \quad (9.24)$$

e

$$\beta_1 \sim N(20, (0.001)^{-1}). \quad (9.25)$$

Observe que assumimos variâncias com valores grandes para as distribuições *a priori*, logo, temos densidades *a priori* não-informativas, pois $E(\beta_0) = 174$, $Var(\beta_0) = 1000$, $E(\beta_1) = 20$ e $Var(\beta_1) = 1000$.

A partir das distribuições condicionais *a posteriori* para β_0 e β_1 necessárias para o algoritmo *Gibbs Sampling*, foram geradas duas amostras de tamanho 53000 e desta amostra foram descartadas as 3000 primeiras observações (*burn-in sample*), onde obtemos das duas cadeias geradas uma amostra de tamanho 100000. Na Tabela 9.2, temos os resumos *a posteriori* dos parâmetros de interesse, considerando uma abordagem Bayesiana tradicional do modelo estresse-resposta geral $\lambda_i = \exp\{- (Z_i + \beta_0 + \beta_1 X_i)\}$.

Tabela 9.2- Resumos *a posteriori* do modelo estresse-resposta assumindo uma distribuição exponencial.

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
β_0	152.5	9.928	(128.9; 170.7)
β_1	19.47	1.208	(16.6; 21.69)

Agora, considerando o modelo exponencial dado em (9.1) e assumindo uma família de distribuições gama para as taxas de falha λ_i do modelo proposto, com as densidades *a priori* para β_0 e β_1 dadas em (9.24) e (9.25) realizamos uma análise Bayesiana hierárquica dos dados da Tabela 9.1.

A partir das distribuições condicionais *a posteriori* para β_0 e β_1 geramos duas cadeias de tamanho 150000, da qual descartamos as 3000 primeiras observações (*burn-in sample*). Os resultados obtidos são apresentados na Tabela 9.3.

Tabela 9.3- Resumos *a posteriori* do modelo estresse-resposta assumindo uma distribuição exponencial e um processo gama para as taxas de falha.

Parâmetro	média <i>a posteriori</i>	desvio-padrão	IC 95%
β_0	173.9	9.562	(155.0; 192.0)
β_1	20.09	1.169	(17.81; 22.34)
λ_1	0.000231	0.000053	(0.000139; 0.000345)
λ_2	0.001105	0.000220	(0.000717; 0.001579)

Podemos observar que os resumos *a posteriori* do modelo exponencial assumindo um processo gama para as taxas de falha apresenta resultados mais precisos do que os resultados obtidos pela análise Bayesiana tradicional.

Capítulo 10

Considerações Futuras

Neste trabalho de pesquisa, considerando que os dados de sobrevivência são ajustados por uma distribuição exponencial assumindo a presença de censuras do tipo II, desenvolvemos análises clássica e Bayesiana na presença de um vetor de covariáveis.

Na análise Bayesiana, considerando um processo gama como *priori* para as taxas individuais, λ_i , determinamos a densidade *a posteriori* conjunta para os parâmetros de interesse.

Uma análise Bayesiana para os tempos de sobrevivência com uma distribuição de Weibull também foi desenvolvida, na presença de covariáveis e dados não-censurados.

A partir dos resultados desta pesquisa, concluímos que o uso de um processo gama para modelos de regressão exponencial ou Weibull levam à inferências melhores do que obtidas usando os métodos Bayesianos tradicionais.

O uso dos *softwares* *WinBUGS* e *R* simplificam a obtenção dos sumários *a posteriori* considerando a análise Bayesiana hierárquica.

Alguns itens desta pesquisa podem ser destacados para serem estudados em um trabalho futuro:

- (1) outras escolhas para a função de base $\lambda_0(t)$ podem ser consideradas, o que leva a uma grande flexibilidade de ajuste aos dados (análise Bayesiana empírica);
- (2) diferentes distribuições *a priori* podem ser consideradas, incorporando opiniões

técnicas e informações de análises preliminares dos dados;

(3) implementar os métodos Bayesianos em outros *softwares* como por exemplo, *Ox* e *SAS* e comparar com os resultados dos *softwares WinBugs* e *R* que foram utilizados neste trabalho;

(4) considerar outros modelos de testes de vida acelerados, como por exemplo, os modelos de Arrhenius e de Eyring;

(5) assumir que os tempos de sobrevivência dos componentes seguem uma distribuição de Weibull.

Bibliografia

- [1] ACHCAR, J. A.; LOUZADA-NETO, F.; MAZUCHELI, J.; DIAS, T. C. M. (2003) - Uso do Processo Gama para Modelar Dados de Sobrevivência na Presença de Covariáveis: Uma Abordagem Bayesiana Hierárquica - (submetido).
- [2] ACHCAR, J. A.; DIAS, T. C. M. (2003) - Accelerated Life Tests Assuming a Gamma Process for the Failure Rates - (submetido).
- [3] ACHCAR, J. A.; LOUZADA-NETO, F. (1993) - Uso de Dados de Sobrevivência no Controle de Qualidade de Produtos Industriais assumindo uma Distribuição Exponencial e um Modelo de Estresse-Resposta Geral - *Estadística*, 45, 144, 145, 81-106.
- [4] BORGES, W. S.; COLOSIMO, E. A.; FREITAS, M. A. (1996) - Métodos Estatísticos e Melhoria da Qualidade: Construindo Confiabilidade em Produtos - *12º SINAPE* - Caxambu.
- [5] CASELLA, G.; GEORGE, E. (1992) - Explaining the Gibbs Sampler - *The American Statistician*, 46, 167-174.
- [6] CHIB, S.; GREENBERG, E. (1995) - Understanding the Metropolis-Hastings Algorithm - *The American Statistical Association*, 49, 327-335.
- [7] CLAYTON, D. G. (1991) - A Monte Carlo Method for Bayesian Inference in Frailty Models - *Biometrics*, 47, 467-485.

- [8] COLLET, D. (1994) - Modelling Survival Data in Medical Research - Chapman & Hall.
- [9] COLOSIMO, E. A. (2001) - Análise de Sobrevivência Aplicada - 46^a RBRAS e 9^o SEAGRO - ESALQ/USP- Piracicaba.
- [10] COX, D. R. (1972) - Regression Models and Life Tables (with discussion) - *Journal of the Royal Statistical Society*, B, 34, 187-220.
- [11] COX, D. R. (1975) - Partial Likelihood - *Biometrika*, 62, 269-276.
- [12] DIAS, T. C. M. (1992) - Testes Acelerados Assumindo um Modelo de Potência Inversa e Dados com Censuras do Tipo II - ICMSC-USP - Tese de Mestrado.
- [13] DYKSTRA, R. L.; LAUD, P. W. (1981) - A Bayesian nonparametric approach to reliability - *The Annals of Statistics*, 9, 356-367
- [14] FEIGL, P.; ZELEN, M. (1965) - Estimation of Exponential Survival Probabilities with Concomitant Information - *Biometrics*, 21, 826-838.
- [15] GAMERMAN, D.; MIGON, H. S. (1993) - Inferência Estatística: Uma Abordagem Integrada - UFRJ.
- [16] GELFAND, A. E.; SMITH, A. F. M. (1990) - Sampling-Based Approaches to Calculating Marginal Densities - *Journal of the American Statistical Association*, 85, 398-409.
- [17] GELMAN, A. E.; RUBIN, D. (1992) - Inference from Iterative Simulation using Multiple Sequences - *Statistical Science*, 7, 457-472.
- [18] 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.

- [19] GREEN, P. J.; YANDELL, B. S. (1985) - Semi-Parametric Generalized Linear Models - Springer-Verlag.
- [20] HASTINGS, W. K. (1970) - Monte Carlo Sampling Methods Using Markov Chains and their Applications - *Biometrika*, 57, 97-109.
- [21] HOSMER, D. W.; LEMESHOW, S. (1999) - Applied Survival Analysis: Regression Modeling of Time to Event Data - John Wiley & Sons.
- [22] HOUGAARD, P. (2000) - Analysis of Multivariate Survival Data - Springer-Verlag.
- [23] HOUGAARD, P. (1995) - Frailty Models for Survival Data - *Lifetime Data Analysis* 1, 255-273.
- [24] IBRAHIN, J. G.; CHEN, M-H.; SINHA, D. (2001) - Bayesian Survival Analysis - Springer-Verlag.
- [25] IBRAHIN, J. G.; CHEN, M-H.; MacEACHERN, S. N. (1999) - Bayesian Variable Selection for Proportional Hazards Models - *The Canadian Journal of Statistics*, 27, 701-717.
- [26] KALBFLEISCH, J. D. (1978) - Nonparametric Bayesian Analysis of Survival Time Data - *Journal of the Royal Statistical Society*, Series B 40, 214-221.
- [27] KALBFLEISCH, J. D.; PRENTICE, R. L. (1980) - The Statistical Analysis of Failure Time Data - John Wiley & Sons.
- [28] KLEIN, J. P.; BASU, A. P. (1981) - Weibull accelerated life tests when there are competing causes of failure, *Communications in Statistics-theory and methods*, A10, 20, 2073-2100.
- [29] KLEIN, J. P.; MOESCHBERGER, M. L. (1997) - Survival Analysis - Springer-Verlag.

- [30] LAWLESS, J. F. (1982) - Statistical Models and Methods for Lifetime Data - John Wiley & Sons.
- [31] LOUZADA-NETO, F. (1991) - Testes de Sobrevivência Acelerados: Uma Análise Bayesiana do Modelo de Eyring - ICMSC-USP - Tese de Mestrado.
- [32] LOUZADA-NETO, F.; MAZUCHELI, J.; ACHCAR, J. A. (2002) - Análise de Sobrevivência e Confiabilidade - IMCA.
- [33] MAGALHÃES, M. N.; LIMA, A. C. P. (2002) - Noções de Probabilidade e Estatística - USP.
- [34] MAGALHÃES, M. N. (2004) - Probabilidade e Variáveis Aleatórias - USP.
- [35] MANN, N. R.; SCHAFFER, R. E., SINGPURWALLA, N. D. (1974) - Methods for Statistical Analysis of Reliability and Life Test Data - John Wiley & Sons.
- [36] McCULLAGH, P.; NELDER, J. (1989) - Generalized Linear Models - Chapman & Hall.
- [37] METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N., TELLER, A. H., TELLER, E. (1953) - Equations of State Calculations by Fast Computing Machines - *Journal of Chemical Physics*, 21, 1087-1092.
- [38] MILLER, R. G. (1981) - Survival Analysis - John Wiley & Sons.
- [39] NELSON, W. (1982) - Applied Life Data Analysis - John Wiley & Sons.
- [40] NELSON, W. (1990) - Accelerated Testing - Statistical Models, Tests Plans and Data Analysis - John Wiley & Sons.
- [41] PARMAR, M. K. B.; MACHIN, D. (1995) - Survival Analysis: A Practical Approach - John Wiley & Sons.

- [42] SAS Institute Inc.(1997), Basic Statistics Using SAS Software Course Notes, Cary, NC: SAS Institute Inc.
- [43] SAS Institute Inc.(1997), SAS/STAT Software: Charges and Enhancements Through Release 6.12, Cary, NC: SAS Institute Inc.
- [44] SMITH, A. F. M.; ROBERTS, G. (1993) - Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods (with discussion) - *Journal of the Royal Statistical Society*, B, 55, 3-23.
- [45] SPIEGELHALTER, D.; THOMAS, A.; GILKS, W.; BEST, N. (1994) - BUGS - Bayesian Inference Using Gibbs Sampling - MRC Biostatistics Unit.
- [46] TITTERINGTON, D.; SMITH, A.; MAKOV, U. (1985) - Statistical of Finite Mixture Distributions - John Wiley & Sons.
- [47] VAUPEL, J. M.; MANTON, K.G.; STALLARD, E. (1979) - The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16, 439-454.
- [48] VENABLES, W. N.; RIPLEY, B. D. (1994) - Modern Applied Statistics with S-Plus - Springer-Verlag.