

MÉTODOS DE MONTE CARLO EM
ANÁLISE DE SOBREVIVÊNCIA

Vicente Garibay Cancho

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

Área de Concentração: Estatística

Orientador: Prof. Dr. Heleno Bolfarine

*Durante a elaboração deste trabalho,
o autor recebeu apoio financeiro do CNPq*

– São Paulo, setembro de 1999 –

MÉTODOS DE MONTE CARLO EM ANÁLISE DE SOBREVIVÊNCIA

Este exemplar corresponde à redação final

da tese devidamente corrigida e

defendida por Vicente Garibay Cancho

e aprovada pela comissão julgadora.

São Paulo, 8 de setembro de 1999.

Banca examinadora:

- Prof. Dr. Heleno Bolfarine (Presidente) - IME - USP
- Prof. Dr. Francisco Louzada - Neto - UFSC.
- Prof. Dr. Helio Migon - UFRJ.
- Prof. Dr. Jorge Alberto Achcar - ICMC - USP.
- Prof. Dr. Reinaldo Boris Arellano Valle - PUCC-Chile.

A mis padres Rosendo e Eulogia (em memoria)

A mis hijos José e Daniel

A mi amada esposa Gladys

AGRADECIMENTOS

A Deus, por mais esta realização.

Ao meu orientador, Professor Heleno Bolfarine, pela oportunidade que tive de ter trabalhado mediante sua orientação segura e dedicada.

Aos Professores Jorge Alberto Achcar e Josemar Rodrigues pelo apoio que me brindaram.

Aos membros da comissão julgadora, em especial ao Prof. Francisco Louzada-Neto, pela observação de imprecisões e erros, na versão preliminar deste trabalho e por suas sugestões.

Aos professores do Departamento de Estatística do IME-USP, por minha formação.

Aos meus familiares, que sempre me incentivaram a continuar, principalmente nos momentos difíceis e especialmente a minha cunhada Elena.

A todos meus colegas que em forma direta ou indireta contribuíram para a finalização deste trabalho.

Ao CNPq, pelo apoio financeiro que permitiu realizar meus estudos.

Ao COSEAS, pela bolsa moradia.

E em especial, a minha esposa Gladys e meus filhos José Luis e Daniel Vincent.

Obrigados a todos.

RESUMO

O principal objetivo deste trabalho é desenvolver técnicas Monte Carlo via Cadeias de Markov para alguns modelos especiais considerados na Teoria da Confiabilidade e na Análise de Sobrevivência. Uma metodologia Bayesiana é desenvolvida para o modelo Weibull-exponenciada proposto por Mudholkar et al. (1995). Uma extensão deste modelo é também considerada adicionando covariáveis ao modelo originalmente proposto por Mudholkar et al. (1995). Reportamos um estudo comparativo do modelo proposto por Mudholkar et al. (1995) com várias famílias de distribuições usadas no ajuste de dados de tempos de vida com função de taxas de falha do tipo "bathtub". Finalmente desenvolvemos um procedimento de análise Bayesiana para o modelo de mistura exponencial com ou sem covariáveis e, comparamos tal análise com os procedimentos clássicos de análise. Consideramos a distribuição Weibull-exponenciada para o ajuste de dados de tempo de vida na presença de indivíduos imunes e estudamos o uso deste modelo no teste de adequação do modelo de mistura Weibull e exponencial considerando as abordagens clássica e Bayesiana.

ABSTRACT

The main object of this study is to develop Markov Chain Monte Carlo techniques in some special models typically considered in reliability theory and survival analysis. A Bayesian methodology is developed for the exponentiated-Weibull model proposed in Mudholkar et al. (1995). An extension of this model is also considered by adding covariates to the model originally proposed in Mudholkar et al. (1995). One of the special features of the exponentiated-Weibull model is that it allows fitting bathtub failure rate models to data sets. Numerical studies are reported comparing several methods for fitting bathtub models to data. Next, we consider a Bayesian analysis for the exponential and Weibull mixture models for incorporating long-term survival data to the model. Comparisons with the classical maximum likelihood methodology are reported and extensions considered by using the exponentiated-Weibull model in this situation.

Conteúdo

1	Introdução	1
1.1	Análise de Dados Tempo de Vida	1
1.2	Análise Bayesiana para os modelos de tempo de vida	3
1.2.1	Teorema de Bayes	3
1.2.2	Algoritmos de simulação	5
1.2.3	Algumas considerações sobre seleção de modelos	8
1.3	Objetivos	11
2	A familia de distribuições Weibull-exponenciada	12
2.1	Introdução	12
2.2	A distribuição Weibull-exponenciada	13
2.2.1	Momentos	13
2.2.2	Função de sobrevivência e função risco	14
2.3	Análise clássica do modelo Weibull-exponenciada	15
2.3.1	Estimação dos parâmetros por máxima verossimilhança	15
2.3.2	Inferência sobre os parametros α , θ e σ	17

2.3.3	Teste de bondade de ajuste	18
2.4	Análise Bayesiana do modelo Weibull-exponenciada	19
2.4.1	Densidade conjunta a posteriori para os parâmetros do modelo Weibull-exponenciada	19
2.4.2	Estimador de Bayes da função de sobrevivência e função risco . . .	21
2.4.3	Densidade conjunta a posteriori dos parâmetros do modelo Weibull	22
2.5	Exemplo de aplicação	23
3	Modelos de regressão Weibull-exponenciada	29
3.1	Introdução	29
3.2	O Modelo log-Weibull-exponenciada	30
3.3	O Modelo de regressão Weibull-exponenciada	30
3.4	Análise clássica para os modelos de regressão Weibull-exponenciada	32
3.4.1	Estimação dos parâmetros por máxima verossimilhança	32
3.4.2	Inferência sobre os parâmetros $\beta_1, \dots, \beta_p, \delta$ e θ	33
3.4.3	Teste para modelo de regressão Weibull	34
3.4.4	Teste para o modelo de regressão exponencial	35
3.5	Análise Bayesiana para o modelo de regressão Weibull-exponenciada	35
3.5.1	Densidade conjunta a posteriori dos parâmetros do modelo de regressão Weibull-exponenciada	36
3.5.2	Estimador de Bayes da função de sobrevivência no modelo de regressão Weibull-exponenciada	38

3.5.3	Densidade conjunta a posteriori dos parâmetros do modelo de regrssão Weibull	39
3.6	Exemplos de aplicação	40
3.6.1	Exemplo 1-Dados simulados	40
3.6.2	Exemplo 2-Teste de vida de Isolantes	44
4	Uso da distribuição Weibull-exponencial para dados com taxas de falha do tipo “bathtub”	53
4.1	Introdução	53
4.2	Modelos	54
4.2.1	A família de distribuições Weibull-exponenciada	55
4.2.2	A família de distribuições IDB	55
4.2.3	A família de distribuições exponencial-potência	56
4.2.4	Misturas de distribuições Gama	57
4.3	Inferência Bayesiana	57
4.3.1	O Modelo Weibull-exponenciada	58
4.3.2	O modelo IDB	59
4.3.3	O modelo exponencial-potência	60
4.3.4	Modelo de mistura distribuições de Gamas	61
4.4	Discriminação de modelos	62
4.5	Alguns exemplos	63
4.5.1	Exemplo 1	63

4.5.2	Exemplo 2	70
5	Comparação de dois modelos exponenciais com dados acelerados	78
5.1	Introdução	78
5.2	Formulação do modelo	79
5.3	Teste de hipóteses Bayesiano	80
5.4	A função de verossimilhança	82
5.5	Análise Bayesiana	83
5.5.1	DAP para os modelos M_1 e M_2	83
5.5.2	Priori de Jeffreys	88
5.5.3	Prioris Próprias	91
5.6	Aplicação	93
6	Modelo de mistura exponencial	103
6.1	Introdução	103
6.2	Análise clássica do modelo de mistura exponencial	104
6.2.1	Estimadores de máxima verossimilhança	105
6.2.2	Inferência sobre os parâmetros λ e p	106
6.2.3	Um estudo de simulação	108
6.2.4	Comparação de proporções de indivíduos imunes entre grupos	109
6.2.5	Estudo de simulação	110
6.3	Análise Bayesiana para modelos de mistura exponencial	111

6.3.1	Análise Bayesiana do modelo de mistura considerando a densidade a priori não informativa de Jeffreys	111
6.3.2	Estudo de simulação	115
6.4	Teste de hipóteses Bayesiano	117
6.4.1	Teste para $H_0 : p = 1$	117
6.4.2	Estudo de simulação	120
6.4.3	Comparação de proporções de sobreviventes	120
6.4.4	Teste para $H_0 : p_1 = \dots = p_G = 1$	125
6.5	Exemplo ilustrativo	127
6.5.1	Análise clássica	127
6.5.2	Análise Bayesiana	130
7	Modelos de mistura exponencial com covariáveis	137
7.1	Introdução	137
7.2	Análise clássica	138
7.2.1	Estimadores de máxima verossimilhança	140
7.2.2	Inferência sobre os parâmetros α e β	141
7.3	Análise Bayesiana	142
7.3.1	Análise Bayesiana do modelo de mistura exponencial com covariáveis usando o método MCMC	142
7.3.2	Estimador de Bayes do tempo médio de vida e para a proporção de sobreviventes dado um valor da covariável	144
7.4	Aplicação	144

7.4.1	Exemplo 1-Dados simulados	145
7.4.2	Exemplo 2-Dados de Kersey et al. 1987	149
8	Modelo de mistura Weibull-exponenciada	153
8.1	Introdução	153
8.2	Modelo	153
8.3	Uma análise clássica do modelo de mistura Weibull-exponenciada	154
8.3.1	Estimação dos parâmetros por máxima verossimilhança	155
8.3.2	Inferências sobre os parâmetros α , θ , λ e p	156
8.3.3	Teste do modelo de mistura Weibull	157
8.3.4	Teste para o modelo de mistura exponencial	158
8.4	Uma análise Bayesiana para o modelo de mistura Weibull-exponenciada . .	158
8.4.1	Densidade a priori independentes	159
8.4.2	Densidade conjunta a posteriori para os parâmetros do modelo de mistura Weibull-exponenciada	160
8.4.3	Densidade conjunta a posteriori para os parâmetros do modelo de mistura Weibull	162
8.4.4	Teste de hipóteses Bayesiano	163
8.5	Exemplos de aplicação	164
8.5.1	Exemplo 1-Dados Simulados	164
8.5.2	Exemplo 2-Dados de Kersey et al. 1987	171
Referências Bibliográficas		178

Apêndices	187
A Critério de convergência	187
B Alguns programas	189

Capítulo 1

Introdução

1.1 Análise de Dados Tempo de Vida

A análise estatística de dados de tempos de vida representa o principal interesse de duas áreas de estudo, a teoria de confiabilidade e, a análise de sobrevivência. Na confiabilidade, temos como exemplos os teste de vida onde os dados utilizados referem-se ao tempo de falha de um ítem, enquanto que, na análise de sobrevivência os dados representam em geral o tempo de vida ou tempo de remissão de uma doença, de pacientes ou animais sob experimentação. Esses dados de tempo de vida são frequentemente dados censuradas, isto é, para alguns elementos em estudo não sabemos seu tempo exato de vida, mas apenas que excede um certo valor, já que não é possível esperar que todas unidades experimentais em teste falhem e é necessário que o experimento se encerre num determinado tempo fixo, ou ainda, que ele se encerre depois de um certo número de falhas tenha ocorrido.

Na análise de sobrevivência e confiabilidade alguns conceitos são de suma importância, para isto considere uma variável aleatória contínua não negativa T , representando o tempo de vida de um indivíduo ou dispositivo de uma população homogênea. A distribuição de probabilidade de T pode ser representada por três formas equivalentes.

A função de densidade que é definida por,

$$f(t) = \lim_{\Delta t \rightarrow 0} p(t \leq T \leq t + \Delta t) = \frac{\partial F(t)}{\partial t}, \quad t > 0.$$

onde $F(t) = P(T \leq t)$ é a função de distribuição de T , a *função de sobrevivência* de T definida por,

$$S(t) = P(T \geq t) = \int_t^{\infty} f(x)dx,$$

onde esta função representa a probabilidade de um indivíduo (ou ítem) sobreviver pelo menos até o tempo t . Em alguns contextos, especialmente aqueles envolvendo tempos de vida de ítems essa função é chamada de *função de confiabilidade*. Pode-se mostrar que $S(t)$ é uma função monótona decrescente, contínua, com $S(0)=1$ e $\lim_{t \rightarrow \infty} S(t) = 0$.

Finalmente a *função risco* que especifica a taxa instantânea de falha ou morte no tempo t de um indivíduo (ou ítem), dado que ele sobreviveu até o tempo t , e que é definida por

$$h(t) = \lim_{\Delta t \rightarrow 0} = \frac{p(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}.$$

A densidade $f(t)$ e a função de sobrevivência $S(t)$ são representações comuns da distribuição de uma variável aleatória. Enquanto, a função risco $h(t)$ é uma caracterização mais especializada, mas é particularmente útil na modelagem de dados de sobrevivência. Frequentemente em aplicações, tem-se disponível a informação de como a taxa de falha muda com o tempo, e estas informações podem ser usadas na modelagem de $h(t)$. Na literatura existem alguns modelos que são usados com bastante frequência (veja por exemplo Lawles, 1982; Collet, 1994). Alguns destes modelos são os modelos exponencial e Weibull. Contudo, esses modelos não incluem o caso de funções de taxa de falha em forma de "bathtub" ou unimodal. Para tais situações foram propostos muitos modelos paramétricos entre eles estão as distribuições F generalizada de Printice (1975), gama generalizada de Stacy (1962) entre outros; não obstante a modelagem desses modelos é computacionalmente muito complicada na presença de dados censurados. Mudholkar et al. (1995) apresentam uma extenção da distribuição Weibull denominada Weibull-exponenciada que além de ter a característica de ter funções de risco unimodal e em forma de "bathtub", inclue

os modelos com funções risco monótonos. No Capítulo 2, apresentamos a distribuição Weibull-exponenciada e algumas de suas propriedades

1.2 Análise Bayesiana para os modelos de tempo de vida

Na análise de confiabilidade ou de sobrevivência, quando assumimos um modelo para os dados de tempos de vida, temos interesse em determinar o valor do parâmetro que caracteriza o modelo. Embora este parâmetro seja considerado fixo, não temos certeza do seu verdadeiro valor e assim procuramos uma forma de, através dos dados, expressar essa incerteza.

Uma alternativa é considerar uma distribuição para o parâmetro de interesse, que sob este ponto de vista é considerado uma variável aleatória (ver por exemplo Box e Tiao, 1973). Com esta formulação usamos o teorema de Bayes, que fornece um método para computar a distribuição do parâmetro, expressando nossa incerteza condicionada aos dado. Para análise de sobrevivência ou confiabilidade vários trabalhos utilizando abordagem Bayesiana já foram realizados (ver por exemplo, Achcar e Louzada-Neto, 1992; Rodrigues , Bolfarine e Louzada-Neto, 1993, entre outros).

1.2.1 Teorema de Bayes

Suponha que $t = (t_1, \dots, t_n)$ é um vetor de n observações onde a distribuição de probabilidade $\pi(t|\theta)$ depende de p parâmetros. A quantidade $\theta = (\theta_1, \dots, \theta_p)$ sob o ponto de vista Bayesiano, é considerado uma variável aleatória, representado características que se deseja conhecer. Se θ tem uma distribuição de probabilidade (denominda distribuição a priori) $\pi(\theta)$, então,

$$\pi(t|\theta)\pi(\theta) = \pi(t, \theta) = \pi(\theta|t)\pi(t).$$

A distribuição de θ dado as observações é dada por

$$\pi(\theta|t) = \frac{\pi(t|\theta)\pi(\theta)}{\pi(t)}, \quad (1.1)$$

onde

$$\pi(t) = c^{-1} = \int_{\Theta} \pi(t|\theta)\pi(\theta)d\theta.$$

Observe que $\pi(t)$ é a esperança matemática de $\pi(t|\theta)$ em relação a densidade a priori $\pi(\theta)$.

Se denotamos a função de verossimilhança $\pi(t|\theta)$ por $L(\theta)$, pelo teorema de Bayes podemos escrever (1.1) como

$$\pi(\theta|t) = c L(\theta)\pi(\theta),$$

onde c é a constante normalizadora de $\pi(\theta|t)$.

Portanto a densidade a posteriori (1.1) pode ser escrita como

$$\pi(\theta|t) \propto L(\theta)\pi(\theta). \quad (1.2)$$

O procedimento de inferência Bayesiana baseia-se na distribuição a posteriori das componentes de θ . Neste ponto surgem duas dificuldades, a primeira na obtenção da distribuição a posteriori marginal e a segunda no cálculo de momentos a posteriori de interesse. Em ambos casos é necessário a resolução de integrais que muitas vezes não apresentam uma solução analítica. Nesse caso, métodos numéricos (ver por exemplo, Naylor e Smith, 1982) ou métodos de aproximação de integrais, tal como o método de Laplace (ver por exemplo Tierney e Kadane, 1986) são necessários. Ocorre que para a aplicação deste último método é necessário determinar o valor que maximiza a função de densidade a posteriori e assim a medida que a dimensão do espaço paramétrico aumenta, aumentam as dificuldades para sua obtenção e a aplicação do método vai se tornando mais trabalhosa de ser implementada. Neste sentido, uma alternativa, é utilizar os algoritmos de simulação de Monte Carlo via Cadeias de Markov (MCMC) tais como os algoritmos de Gibbs (ver por exemplo, Gelfand e Smith, 1990) e Metropolis-Hastings (ver por exemplo,

Chib e Greenberg, 1995). Esse métodos permitem inclusive escolher densidades a priori sem prender a propriedade de conjugação.

Um outro elemento importante na inferência Bayesiana é a distribuição preditiva. Usualmente estamos interessados em fazer inferência sobre alguma observação futura, s , utilizando as observações passadas t . Tal inferência deve ser baseada na distribuição a posteriori de s , $\pi(s|t)$. Observamos que $\pi(s|t)$ pode ser obtida a partir da distribuição a posteriori conjunta de s e θ por integração, ou seja,

$$\pi(s|t) = \int \pi(s, \theta|t)d\theta = \int \pi(s|\theta, t)\pi(\theta|t)d\theta,$$

onde $\pi(s|\theta, t)$ em muitos modelos se reduz a $\pi(s|\theta)$.

1.2.2 Algoritmos de simulação

O algoritmo do amostrador de Gibbs (A.G.) e algoritmo de Metropolis-Hasting são um dos métodos de simulação de Monte Carlo via Cadeias de Markov (MCMC). O A.G. é uma técnica de MCMC para gerar amostras de uma densidade multivariada não normalizada. Os métodos de simulação MCMC atualmente têm sido utilizados na inferência Bayesiana (ver por exemplo, Gelfand e Smith, 1990 e Chib e Greenberg, 1995). Na inferência Bayesiana a distribuição estacionária desta Cadeia de Markov é a distribuição a posteriori. Geman e Geman (1984) apresentaram o algoritmo do A.G. em problemas de processamento de imagens. Entretanto o método já havia sido descrito por Metropolis (1953) e posteriormente por Hasting (1970). Gelfand e Smith (1990) popularizaram o A.G. na aplicação do método na área de inferência Bayesiana.

Descrevemos brevemente o procedimento do amostrador de Gibbs e em geral o algoritmo de Metropolis-Hasting. Suponha que desejamos estimar $\pi(\theta_1, \dots, \theta_p|D)$, densidade conjunta a posteriori do $(\theta_1, \dots, \theta_p)$ dado os dados. O algoritmo considera que as densidades condicionais a posteriori de cada parâmetro θ_i , dado todos os outros sejam conhecidas. Essas distribuições são denotados por $\pi(\theta_1|\theta_2, \dots, \theta_p, D)$, $\pi(\theta_2|\theta_1, \theta_3, \dots, \theta_p, D), \dots$, $\pi(\theta_p|\theta_1, \theta_2, \dots, \theta_{p-1}, D)$.

Com os valores iniciais $\theta_1^{(0)}, \dots, \theta_p^{(0)}$ geramos

$$\begin{aligned} \theta_1^{(1)} &\text{ de } \pi(\theta_1 | \theta_2^{(0)}, \dots, \theta_p^{(0)}, D), \\ \theta_2^{(1)} &\text{ de } \pi(\theta_2 | \theta_1^{(1)}, \theta_3^{(0)}, \dots, \theta_p^{(0)}, D), \\ &\vdots \quad \vdots \quad \vdots \\ \theta_p^{(1)} &\text{ de } \pi(\theta_p | \theta_1^{(1)}, \theta_2^{(1)}, \dots, \theta_{p-1}^{(1)}, D), \end{aligned} \tag{1.3}$$

com essa geração $\boldsymbol{\theta}^{(1)} = (\theta_1^{(1)}, \dots, \theta_p^{(1)})'$ substitui o valor inicial e continuamos iterando até a iteração k , onde se espera que a convergência da cadeia é alcançada. Sob certas condições de regularidade (veja, por exemplo Tierney, 1994), essa cadeia de Markov converge para distribuição estacionária para k suficientemente grande, isto é, $(\theta_1^{(k)}, \dots, \theta_p^{(k)})'$ tem uma distribuição aproximadamente igual a $\pi(\theta_1, \dots, \theta_p | D)$. Além disso, θ_i pode ser considerada como uma observação simulada de $\pi(\theta_i | D)$, a distribuição marginal de θ_i (veja, Geman e Geman, 1984). Se considerarmos s pontos iniciais independentes, podemos também replicar a iteração anterior r vezes. Seja $\boldsymbol{\theta}^{(i,s)} = (\theta_1^{(i,s)}, \dots, \theta_p^{(i,s)})'$ denota a realização de $\boldsymbol{\theta}$ obtida a partir da i -ésima iteração e da s -ésima réplica. Para estimar $\pi(\theta_1, \dots, \theta_p | D)$, seus momentos de interesse e intervalos de credibilidade, pode-se usar uma medida empírica atribuindo $1/r$ de peso a cada $(\theta_1^{(k,s)}, \dots, \theta_p^{(k,s)})$, $s = 1, \dots, r$. Alternativamente, Gelman e Rubin (1992) sugerem usar essa medida empírica para todas as variáveis considerando a segunda metade de todas as iterações e réplicas.

Para o uso da convergência e escolha prática de k e r , pode-se consultar Tanner e Wong (1987), Gelfand e Smith (1990), Casella e George (1991), e Gelman e Rubin (1992) e Cowles e Carlin (1996). Na prática, entretanto, ainda existe muita discussão sobre o monitoramento da convergência. Contudo, o método proposto por Gelman e Rubin (1992) o qual se baseia na técnica de análise de variância é a mais utilizada na inferência Bayesiana (veja por exemplo, Achcar e Perifera., 1998).

Quando as densidades condicionais não são facilmente identificáveis, ou se a geração não-iterativa for muito complicada o algoritmo de Metropolis-Hastings (M.H.) pode ser usado. Descrevemos a seguir o algoritmo de M.H. Suponha que desejamos uma amostra

da variável θ_1 em (1.3) onde sua densidade condicional a posteriori não é conhecida. Assim, consideramos a densidade genérica

$$\pi(\theta_1|\theta_2, \dots, \theta_p, D) = \frac{L(\theta_1, \dots, \theta_p)\pi(\theta_1, \dots, \theta_p)}{\int L(\theta_1, \dots, \theta_p)\pi(\theta_1, \dots, \theta_p)d\theta_1}. \quad (1.4)$$

Vamos denotar por $f(\theta_1)$ a densidade condicional em (1.4) suprimindo as variáveis condicionantes por brevidade. Assim define-se o núcleo de transição $q(\theta_1, X)$ que leva θ_1 a X . Se θ_1 é um valor real com suporte $(-\infty, \infty)$, construiremos $q(\cdot)$ de tal forma que $X = \theta_1 + \sigma Z$, sendo Z uma variável aleatória com distribuição normal padrão (ou t -student) e σ^2 tem que refletir a variância condicional de θ_1 em (1.4). Se θ_1 é uma variável limitada com suporte no intervalo (a, b) , podemos usar uma transformação, tal como $\theta' = \log\left(\frac{\theta_1-a}{b-\theta_1}\right)$, para levar (a, b) em $(-\infty, \infty)$ e então usamos o núcleo de transição anterior na densidade de θ' . Alternativamente, podemos aplicar um método simples sem a necessidade de considerar uma transformação. Construimos o núcleo de transição $q(\cdot)$ de modo que X seja distribuída uniformemente em (a, b) (veja, por exemplo Kuo et al. 1995). Outra forma de escolha, quando possível, é explorar a forma de $f(\theta_1)$ para obter o núcleo de transição $q(\cdot)$ (veja Chib e Greenberg, 1995). Por exemplo, se $f(\theta_1)$ pode ser dada por $f(\theta_1) \propto \Psi(\theta_1)h(\theta_1)$, onde $h(\theta_1)$ é uma densidade que pode ser amostrada e $\Psi(\theta_1)$ é uniformemente limitada, fazemos $q(\theta_1, X) = h(\theta_1)$. Outras possíveis escolhas podem ser encontradas em Chib e Greenberg (1995).

O algoritmo de M.H. pode ser descrito como segue:

- (i) inicialize θ_1 com um valor $\theta_1^{(0)}$, $j = 0$;
- (ii) gerar X de $q(\theta_1^{(j)}, \cdot)$ e u da distribuição uniforme $(0, 1)$;
- (iii) calcular $p = \min\left\{1, \frac{f(X)q(\theta_1^{(j)}, X)}{f(\theta_1^{(j)})q(X, \theta_1^{(j)})}\right\}$;
- (iv) se $u \leq p$; faça $\theta_1^{(j+1)} = X$, senão faça $\theta_1^{(j+1)} = \theta_1^{(j)}$;
- (v) repita (ii)- (iv) até que a distribuição estacionária seja atingida.

Observe que este algoritmo é definido usando a razão de dois valores obtidos de (1.4). Portanto, tem-se que conhecer a forma funcional da função de verossimilhança e a densidade a priori, mas, não a constante normalizador \mathcal{A} ,

1.2.3 Algumas considerações sobre seleção de modelos

A determinação do modelo se torna um problema fundamental da estatística. A literatura sobre a adequação ou checagem e seleção de modelos apresenta muitas abordagens, começando com a abordagem do fator Bayes. Várias modificações do fator de Bayes são apresentadas na literatura (veja por exemplo, Atkin, 1981; Berger e Perichi, 1992 ou Spiegelhalter e Smith, 1982). Geisser e Eddy (1979) sugerem uma abordagem preditiva baseada em métodos de validação cruzada para obter o pseudo-fator de Bayes.

Considere uma escolha entre dois modelos paramétricos cada um denotado por sua densidade conjunta $f(\underline{t}|\boldsymbol{\theta}_i, M_i)$ ou por sua função de verossimilhança $L(\boldsymbol{\theta}_i|\underline{t}, M_i)$, $i = 1, 2$. Suponha que w_i é a probabilidade de selecionar o modelo M_i , $i = 1, 2$ e $f(\underline{t}|M_i)$ é a distribuição preditiva (verossimilhança preditiva) do modelo M_i , que é definida por,

$$f(\underline{t}|M_i) = \int f(\underline{t}|\boldsymbol{\theta}_i, M_i)\pi(\boldsymbol{\theta}_i|M_i)d\boldsymbol{\theta}_i, \quad (1.5)$$

onde $\pi(\boldsymbol{\theta}_i|M_i)$ é a densidade a priori sobre o modelo M_i . Se \underline{t}_0 denota os dados observados, então escolhemos o modelo que tem maior valor $w_i f(\underline{t}_0|M_i)$.

Frequentemente considera-se $w_i = 0.5$, $i = 1, 2$ e o fator de Bayes de M_1 com respeito a M_2 é definido como

$$B_{12} = \frac{f(\underline{t}_0|M_1)}{f(\underline{t}_0|M_2)}. \quad (1.6)$$

A verossimilhança preditiva (1.5) pode ser aproximada por Monte Carlo usando S amostras geradas a partir da densidade a priori $\pi(\boldsymbol{\theta}_i|M_i)$, isto é,

$$\hat{f}(\underline{t}|M_i) = \frac{1}{S} \sum_{s=1}^S f(\underline{t}|\boldsymbol{\theta}_i^{(s)}, M_i), \quad (1.7)$$

onde $\boldsymbol{\theta}_i^{(s)}$ é a s -ésima amostra do vetor $\boldsymbol{\theta}_i$.

Podemos notar que, caso a distribuição a priori $\pi(\theta_i|M_i)$ seja imprópria, o que pode acontecer quando consideramos densidades a priori não informativas, $f(D|M_i)$ também será imprópria e consequentemente (1.6) não pode ser usada como critério de comparação, uma vez que não se tem uma razão de probabilidades. Esse fato pode restringir o uso do fator de Bayes à classe das prioris próprias, ou ainda a modelos razoavelmente simples, o que em geral não ocorre na prática, principalmente com relação a modelos de sobrevivência ou confiabilidade.

No sentido de superar restrições associadas ao fator de Bayes, muitos trabalhos recentemente publicados sugerem mudanças em (1.6) (veja por exemplo, Gelfand e Dey, 1994). O uso da distribuições preditivas em alguma forma tem sido reconhecida como a correta abordagem Bayesiana para a seleção de modelos. Particularmente, Box (1980) argumenta que densidades a posteriori devem ser usadas para a estimativa dos parâmetros condicional a adequabilidade do modelo, enquanto que densidades preditivas devem ser usadas para a análise da adequabilidade dos modelos na presença dos dados. Na avaliação de dois modelos quaisquer as densidades preditivas podem ser comparadas enquanto que as densidades a posteriori nem sempre, conforme visto acima.

A densidade preditiva que foi utilizada extensivamente em trabalhos recentes de Gelfand e Dey (1994) e Gelfand, Dey e Chang (1992) é a densidade de validação cruzada dada por,

$$f(t_r|D_{(r)}) = \frac{f(D)}{f(D_{(r)})} = \int f(t_r|\theta, D_{(r)})\pi(\theta|D_{(r)})d\theta, \quad (1.8)$$

onde $D_{(r)} = (t_1, \dots, t_{r-1}, t_{r+1}, \dots, t_n)$ e $\pi(\theta|D_{(r)})$ é a correspondente densidade a posteriori baseada em $D_{(r)}$. Essa densidade pode ser chamada de densidade preditiva condicional ordenada (CPO).

Na aproximação de validação cruzada, um único ponto é deletado, e este procedimento é uma rotina padrão em análise de regressão clássica e muitas técnicas de diagnósticos adotam esse ponto de vista. Isto também é estabelecido na literatura Bayesiana por Stone (1974).

Para obter uma estimativa da densidade preditiva de validação cruzada observe que

$$\begin{aligned} f(t_r|D_{(r)}) &= \frac{f(D)}{f(D_{(r)})} = \frac{\int \frac{f(D|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\boldsymbol{\theta}|D)f(D)} \cdot \pi(\boldsymbol{\theta}|D)d\boldsymbol{\theta}}{\int \frac{f(D_{(r)}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\pi(\boldsymbol{\theta}|D)f(D)} \cdot \pi(\boldsymbol{\theta}|D)d\boldsymbol{\theta}} \\ &= \frac{1}{\int \frac{1}{f(t_r|D_{(r)}, \boldsymbol{\theta})} \pi(\boldsymbol{\theta}|D)d\boldsymbol{\theta}}. \end{aligned} \quad (1.9)$$

Daí, por aproximação de Monte Carlo em (1.9) obtém-se a estimativa

$$\hat{f}(t_r|D_{(r)}) = B \left(\sum_{i=1}^B \frac{1}{f(t_r|D_{(r)}, \boldsymbol{\theta}_i)} \right)^{-1}. \quad (1.10)$$

onde B é o tamanho da amostra a posteriori de $\boldsymbol{\theta}$.

Se $t_r, r = 1, \dots, n$, são condicionalmente independentes dado $\boldsymbol{\theta}$, $f(t_r|D_{(r)}, \boldsymbol{\theta}_i)$ simplifica-se para $f(t_r|\boldsymbol{\theta}_i)$. Observe que esta quantidade pode ser considerada como a r -ésima componente da função de verossimilhança ou como a densidade condicional de t_r dado $\boldsymbol{\theta}_i$. Logo, (1.10) é a média harmônica das componentes da função verossimilhança ou média harmônica da densidade condicional avaliada na amostra a posteriori.

Podemos utilizar as estimativas $c_r(l) = \hat{f}(t_r|D_{(r)})$ na seleção dos modelos, considerando o gráfico dos valores $c_r(l)$ contra o número de observações r ($r = 1, \dots, n$) para os diferentes modelos; valores maiores (em média) indicam o melhor modelo. Também pode-se considerar para escolher o modelo que maximiza $c(l) = \prod_{r=1}^n c_r(l)$, onde l é o indexador do modelo.

Alternativamente, podemos considerar o gráfico da razão dos CPO's (ou log da razão dos CPO's) dos diferentes modelos contra o número de observações para selecionar o melhor modelo; valores da razão maiores que *um* (ou positivos) indicam a preferência do primeiro modelo.

Geisser e Eddy (1979) sugere que o produto das densidades preditivas $\prod_{r=1}^n f(t_r|D_{(r)}, M_l)$ pode ser usada como um indicador na seleção de modelos. Se temos dois modelos M_1 e

M_2 , temos que a razão,

$$\frac{\prod_{r=1}^n f(t_r | D_{(r)}, M_1)}{\prod_{r=1}^n f(t_r | D_{(r)}, M_2)} \quad (1.11)$$

pode ser usada como uma aproximação do fator de Bayes, a qual denota-se por $PSFB_{12}$ (pseudo-fator de Bayes,) escolhendo-se o modelo M_1 se $PSFB_{12} > 1$.

1.3 Objetivos

O principal objetivo deste trabalho é fazer análise Bayesiana para modelos de sobrevivência ou confiabilidade na ausência ou presença de indivíduos imunes utilizando os métodos de simulação via Cadeias de Markov com Monte Carlo.

Este trabalho está organizado em 8 Capítulos. O Capítulo 2 é dedicado à inferência Bayesiana considerando a distribuição Weibull-exponenciada para os tempos de sobrevivência e ao estudo de algumas propriedades dessa distribuição. No Capítulo 3 estendemos a família de distribuição Weibull-exponenciada para caso do modelos de regressão e apresentamos o procedimento de inferência clássica e Bayesiana para o modelo. No Capítulo 4 apresentamos um estudo comparativo da distribuição Weibull-exponenciada com as famílias de distribuições mais comumente usadas para modelar tempos de vida que apresentam função de taxa de falha do tipo "bathtub" desde uma perspectiva Bayesiana. No Capítulo 5 desenvolvemos uma análise Bayesiana considerando várias densidades a priori, para compararmos 2 populações sob condições aceleradas considerando a distribuição exponencial para os tempos de vida. No Capítulo 6 desenvolvemos o procedimento de inferência Bayesiana para os modelos de sobrevivência na presença de indivíduos imunes (modelos de mistura) considerando a distribuição exponencial para os tempos de vida. No Capítulo 7, apresentamos o procedimento de inferência Bayesiana para ^modelos de mistura exponencial na ~~pre~~ença de covariáveis. Finalmente no Capítulo 8 fazemos inferência clássica e Bayesiana para os modelos de mistura Weibull-exponenciada.

Capítulo 2

A família de distribuições

Weibull-exponenciada

2.1 Introdução

A família de distribuição Weibull é muito utilizada como um modelo para a distribuição de tempos de vida, tanto na área médica como na engenharia. Essa família de distribuições é adequada para modelar dados de tempo de vida com função risco monótonas ou constantes. Contudo, não é adequada para modelar situações onde a função de risco tem forma de banheira ou unimodal. Muitas famílias de distribuições foram introduzidas, para modelar as situações anteriores, como por exemplo, a família de distribuição F generalizada de Prentice (1975), a distribuição gama generalizada de Stacy (1962) ou duas famílias introduzidas por Slymen e Lachenbruch (1984). Porém, a modelagem de dados de tempo de vida com essas famílias de distribuições são computacionalmente muito complicadas na presença de censura.

Mudholkar et al.(1995) apresentam uma extensão da distribuição Weibull denominada família de distribuições Weibull-exponenciadas (WE) em aplicações de confiabilidade e sobrevivência. A nova família é obtida ao introduzir um segundo parâmetro de for-

ma θ adicional na função quantil da distribuição Weibull. Essa extensão, que contém distribuições com função de risco unimodal e em forma de "bathtub", também permite considerar uma ampla classe de modelos com funções risco monótonas.

Neste Capítulo apresentamos um estudo das propriedades da família de distribuição Weibull-exponenciada, bem como análises clássica e Bayesiana para tempos de vida com essa distribuição. O principal apporte deste capítulo é o uso de métodos Bayesianos para análise de dados de tempos de vida com distribuição Weibull-exponenciada.

2.2 A distribuição Weibull-exponenciada

A distribuição Weibull-exponenciada (WE) introduzida por Mudholkar et al. (1995), com parâmetros α , θ e σ para o tempo de vida T tem função de densidade dada por

$$f(t; \alpha, \theta, \sigma) = \frac{\alpha \theta}{\sigma} [1 - \exp(-(t/\sigma)^\alpha)]^{\theta-1} \times \exp(-(t/\sigma)^\alpha)(t/\sigma)^{\alpha-1}, \quad 0 < t < \infty \quad (2.1)$$

onde $\alpha > 0$, $\theta > 0$ são os parâmetros de forma e $\sigma > 0$ é o parâmetro de escala. A densidade (2.1) pode ser usada para especificar a distribuição exponencial como caso particular quando, $\alpha = 1$, e $\theta = 1$ e a distribuição Weibull no caso em que $\theta = 1$.

A função de distribuição de T , denotado por $F(t; \alpha, \theta, \sigma)$ para $t > 0$ é dada por

$$F(t; \alpha, \theta, \sigma) = [1 - \exp(-(t/\sigma)^\alpha)]^\theta. \quad (2.2)$$

2.2.1 Momentos

O k -ésimo momento da distribuição WE, para $k = 1, 2, \dots$ é dado por

$$\begin{aligned} \mu'_k &= E(T^k) = \int_0^\infty t^k f(t; \alpha, \theta, \sigma) dt \\ &= \int_0^\infty t^k \frac{\alpha \theta}{\sigma} [1 - \exp(-(t/\sigma)^\alpha)]^{\theta-1} \times \exp(-(t/\sigma)^\alpha)(t/\sigma)^{\alpha-1} dt \\ &= \theta \sigma^k \int_0^\infty y^{\frac{k}{\alpha}} (1 - \exp(-y))^{(\theta-1)} \exp(-y) dy. \end{aligned} \quad (2.3)$$

Em geral os momentos são intratáveis analiticamente, porém, podem ser estudadas numericamente. Mas, quando θ é um número enteiro (2.3), é dado por

$$\mu'_k = \sigma^k \theta \Gamma\left(\frac{k}{\alpha} + 1\right) \sum_{j=1}^{\theta-1} \binom{\theta-1}{j} \frac{1}{(j+1)^{k/\alpha+1}}. \quad (2.4)$$

Também, se $k/\alpha = r$ é um número $\overset{\circ}{\text{enteiro}}$ e positivo, (2.3) se reduz a

$$\mu'_k = \sigma^k \theta (-1)^r \left[\frac{d^r}{ds^r} B(s, \theta) \right]_{s=1}, \quad (2.5)$$

onde $B(s, \theta)$ denota a função beta (veja, por exemplo, Abramowitz e Stegun, 1965).

2.2.2 Função de sobrevivência e função risco

A função de sobrevivência de T que denotamos por $S(t; \alpha, \theta, \sigma) = P(T \geq t)$ é dada por

$$S(t; \alpha, \theta, \sigma) = 1 - [1 - \exp(-(t/\sigma)^\alpha)]^\theta. \quad (2.6)$$

A função risco de T , definida por $h(t; \alpha, \theta, \sigma) = f(t; \alpha\theta, \sigma)/S(t; \alpha, \theta, \sigma)$ é obtida apartir das expressões (2.1) e (2.6) e é dada por

$$h(t; \alpha, \theta, \sigma) = \frac{\alpha\theta[1 - \exp(-(t/\sigma)^\alpha)]^{\theta-1} \exp(-(t/\sigma)^\alpha)(t/\sigma)^{\alpha-1}}{\sigma[1 - (1 - \exp(-(t/\sigma)^\alpha))^\theta]}. \quad (2.7)$$

A grande flexibilidade deste modelo para o ajuste de dados de tempo de vida é dada pelas diferentes formas que a função de risco (2.7) pode apresentar, isto é,

- (i) monótona crescente se $\alpha \geq 1$ e $\alpha\theta \geq 1$,
- (ii) monótona decrescente se $\alpha \leq 1$ e $\alpha\theta \leq 1$,
- (iii) forma de banheira se $\alpha > 1$ e $\alpha\theta < 1$ e
- (iv) unimodal se $\alpha < 1$ e $\alpha\theta > 1$.

As monotonicidades em (i) e (ii) são estritas exceto para a distribuição exponencial correspondente a $\alpha = \theta = 1$.

A figura 2.1, mostra alguns casos especiais da função risco (2.7).

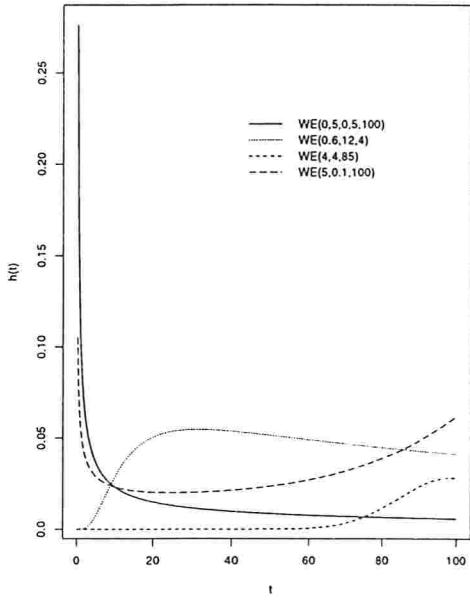


Figura 2.1: Alguma formas especiais da função risco da família de distribuição Weibull-exponenciada

2.3 Análise clássica do modelo Weibull-exponenciada

Nesta seção apresentaremos o procedimento de inferência clássica considerando a distribuição Weibull-exponenciada para os tempos de vida. Consideramos o método de máxima verossimilhança para estimação dos parâmetros e métodos assintóticos para a construção de intervalos de confiança e teste aproximados.

2.3.1 Estimação dos parâmetros por máxima verossimilhança

Sejam t_1, \dots, t_n observações de uma amostra onde t_i representa o tempo de vida ou tempo de censura de cada um dos n indivíduos em estudo. Considerando que os tempos de vida têm distribuição Weibull-exponenciada com função de densidade dada em (2.1), a função de verossimilhança para α , θ e σ é dada por

$$L(\alpha, \sigma, \theta) = \alpha^r \theta^r \sigma^{-r\alpha} \exp \left\{ - \sum_{i \in F} \left(\frac{t_i}{\sigma} \right)^\alpha \right\} \prod_{i \in F} t_i^{\alpha-1} \left(1 - \exp \left\{ - \left(\frac{t_i}{\sigma} \right)^\alpha \right\} \right)^{\theta-1}$$

$$\times \prod_{i \in C} \left(1 - \left(1 - \exp\left\{-\left(\frac{t_i}{\sigma}\right)^\alpha\right\}\right)^\theta\right), \quad (2.8)$$

onde r é número de falhas observadas, F denota o conjunto de observações não censuradas, C denota o conjunto de observações censuradas.

Os estimadores de máxima verossimilhança (EMV) de α , θ e σ são valores que maximizam a função de verossimilhança em (2.8), o que é equivalente a maximizar o logaritmo da função de verossimilhança, isto é, que maximiza: $\ell(\alpha, \theta, \sigma) = \log(L(\alpha, \sigma, \theta))$, ou seja

$$\begin{aligned} \ell(\alpha, \theta, \sigma) &= r \log\left(\frac{\alpha\theta}{\sigma}\right) + (\theta - 1) \sum_{i \in F} \log(g(t_i)) - \sum_{i \in F} \left(\frac{t_i}{\sigma}\right)^\alpha + (\alpha - 1) \sum_{i \in F} \log\left(\frac{t_i}{\sigma}\right) \\ &\quad + \sum_{i \in C} \log(S(t_i)), \end{aligned} \quad (2.9)$$

onde $g(t_i) = g(t_i; \alpha, \sigma) = 1 - \exp(-(t_i/\sigma)^\alpha)$ e $S(t_i) = S(t_i, \alpha, \theta, \sigma)$ é a função de sobrevivência dada em (2.6). Equivalentemente, os EMV resultam da solução do seguinte sistema de equações não lineares:

$$\frac{\partial \ell}{\partial \alpha} = \frac{r}{\alpha} + (\theta - 1) \sum_{i \in F} \frac{g_\alpha(t_i)}{g(t_i)} - \sum_{i \in F} \left(\frac{t_i}{\sigma}\right)^\alpha \log\left(\frac{t_i}{\sigma}\right) + \sum_{i \in F} \log\left(\frac{t_i}{\sigma}\right) + \sum_{i \in C} \frac{S_\alpha(t_i)}{S(t_i)} = 0 \quad (2.10)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{r}{\theta} + \sum_{i \in F} \log(g(t_i)) + \sum_{i \in C} \frac{S_\theta(t_i)}{S(t_i)} = 0, \quad (2.11)$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{r\alpha}{\sigma} + (\theta - 1) \sum_{i \in F} \underbrace{\frac{g'_\alpha(t_i)}{g(t_i)}}_{g_\theta} + \frac{\alpha}{\sigma} \sum_{i \in F} \left(\frac{t_i}{\sigma}\right)^\alpha + \sum_{i \in C} \frac{S_\sigma(t_i)}{S(t_i)} = 0, \quad (2.12)$$

onde $g_\alpha(t_i) = \exp\left(-\left(\frac{t_i}{\sigma}\right)^\alpha\right)\left(\frac{t_i}{\sigma}\right)^\alpha \log\left(\frac{t_i}{\sigma}\right)$, $g_\sigma(t_i) = -\frac{\alpha}{\sigma} \exp\left(-\left(\frac{t_i}{\sigma}\right)^\alpha\right)\left(\frac{t_i}{\sigma}\right)^\alpha$, $S_\alpha(t_i) = -\theta g^{\theta-1}(t_i)g_\alpha(t_i)$, $S_\theta(t_i) = -g^\theta(t_i) \log(g(t_i))$ e $S_\sigma(t_i) = -\theta g^{\theta-1}(t_i)g_\sigma(t_i)$.

Como não é possível encontrar uma forma analítica fechada para $\hat{\alpha}$, $\hat{\theta}$ e $\hat{\sigma}$ pode-se utilizar algum método numérico para resolver o sistema dado em (2.10)-(2.12). Um método mais utilizado é o método de Newton-Rapson. Esse método aparece como subrotina em alguns softwares, como por exemplo, no software Ox (veja Doornik, 1998).

2.3.2 Inferência sobre os parâmetros α , θ e σ

Inferência sobre os parâmetros são feitas usualmente através da normalidade assintótica dos estimadores de máxima verossimilhança (ver, por exemplo, Mudholkar et al. 1995). Assim, para amostras suficientemente grandes e sob certas condições de regularidade pode-se mostrar que

$$i^{1/2}(\alpha, \theta, \sigma) \begin{bmatrix} \hat{\alpha} \\ \hat{\theta} \\ \hat{\sigma} \end{bmatrix} - \begin{bmatrix} \alpha \\ \theta \\ \sigma \end{bmatrix} \xrightarrow{D} N(\mathbf{0}, I_3), \quad (2.13)$$

onde $i(\alpha, \theta, \sigma)^{1/2}$ é a matriz raiz quadrada da matriz de informação de Fisher e I_3 é a matriz identidade de ordem 3×3 . A distribuição assintótica dos EMV segue sendo válido se a matriz de informação de Fisher é substituída pela matriz de informação observada dada por

$$I(\hat{\alpha}, \hat{\theta}, \hat{\sigma}) = \begin{pmatrix} -\frac{\partial^2 \ell}{\partial \alpha^2} & -\frac{\partial^2 \ell}{\partial \alpha \partial \theta} & -\frac{\partial^2 \ell}{\partial \alpha \partial \sigma} \\ -\frac{\partial^2 \ell}{\partial \theta \partial \alpha} & -\frac{\partial^2 \ell}{\partial \theta^2} & -\frac{\partial^2 \ell}{\partial \theta \partial \sigma} \\ \text{simétrica} & -\frac{\partial^2 \ell}{\partial \sigma \partial \alpha} & -\frac{\partial^2 \ell}{\partial \sigma^2} \end{pmatrix} \Big|_{(\alpha, \theta, \sigma) = (\hat{\alpha}, \hat{\theta}, \hat{\sigma})}.$$

Considerando a normalidade assintótica dos estimadores de máxima verossimilhança $\hat{\alpha}$, $\hat{\theta}$ e $\hat{\sigma}$ podem-se calcular os intervalos de confiança para os parâmetros α , θ e σ e para a função de sobrevivência $S(t_0)$ para dado t_0 .

Representando a inversa da matriz de informação observada por

$$I^{-1}(\hat{\alpha}, \hat{\theta}, \hat{\sigma}) = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

e considerando um nível de confiança de $100(1 - \alpha)\%$, os intervalos de confiança aproximados para α , θ e σ são dados por:

$$\begin{aligned} ic(\alpha) &= [\hat{\alpha} - z_{\alpha/2}(b_{11})^{1/2}; \hat{\alpha} + z_{\alpha/2}(b_{11})^{1/2}], \\ ic(\theta) &= [\hat{\theta} - z_{\alpha/2}(b_{22})^{1/2}; \hat{\theta} + z_{\alpha/2}(b_{22})^{1/2}], \\ ic(\sigma) &= [\hat{\sigma} - z_{\alpha/2}(b_{11})^{1/2}; \hat{\sigma} + z_{\alpha/2}(b_{22})^{1/2}], \end{aligned} \quad (2.14)$$

onde $z_{\alpha/2}$ é o percentil correspondente da distribuição normal padronizada e b_{ii} é o i -ésimo elemento da diagonal da inversa da matriz de informação observada I avaliada nos EMV.

2.3.3 Teste de bondade de ajuste

O problema do teste de bondade de ajuste do modelo Weibull contra a classe de alternativas irrestrito é complexo. Contudo, restringindo as alternativas para a família Weibull-exponenciada podemos usar a estatística da razão de verossimilhanças (veja Rao, 1973) para testar a adequabilidade do submodelo Weibull. A hipótese nula, $H_{01} : \theta = 1$ correspondente ao submodelo Weibull e $H_{02} : \{\theta = 1, \alpha = 1\}$ correspondente ao submodelo exponencial (do modelo Weibull-exponenciada). Se uma destas hipóteses ou ambas são rejeitadas na classe da família Weibull-exponenciada então os submodelos são questionáveis. A estatística da razão de verossimilhanças para testar H_{0i} ($i = 1, 2$) são dadas por:

$$\Lambda_i = \sup_{R_{0i}} \{L(\alpha, \theta, \sigma)\} / \sup_R \{L(\alpha, \theta, \sigma)\}, \quad i = 1, 2,$$

onde R_{0i} é o espaço paramétrico correspondente a H_{0i} , $i = 1, 2$, e R é o espaço paramétrico irrestrito.

Em termos dos EMVs, a estatística da razão de verossimilhanças para testar a adequabilidade do modelo Weibull, isto é, $H_{01} : \theta = 1$ contra $H_{11} : \theta \neq 1$, se reduz a

$$\Lambda_1 = L(\hat{\alpha}_w, \theta = 1, \hat{\sigma}_w) / L(\tilde{\alpha}, \tilde{\theta}, \tilde{\sigma}), \quad (2.15)$$

onde $\hat{\alpha}_w$ e $\hat{\sigma}_w$ são EMV dos parâmetros do modelo Weibull e $\tilde{\alpha}$, $\tilde{\theta}$ e $\tilde{\sigma}$ são os EMV dos parâmetros do modelo Weibull-exponenciada. Sob a hipótese H_{01} , $-2 \log(\Lambda_1)$ tem distribuição aproximada qui-quadrado com *um* grau de liberdade para tamanhos amostrais suficientemente grandes (veja Rao, 1973). Similarmente para testar a adequabilidade do modelo exponencial a estatística da razão de verossimilhança é dada por

$$\Lambda_2 = L(\alpha = 1, \theta = 1, \hat{\sigma}_e) / L(\tilde{\alpha}, \tilde{\theta}, \tilde{\sigma}),$$

onde $\hat{\sigma}_e$ é o EMV do parâmetro no modelo exponencial. Sob a hipótese $H_{02} : \{\theta = 1, \alpha = 1\}$, $-2 \log(\Lambda_2)$ tem distribuição aproximada qui-quadrado com 2 graus de liberdade para tamanhos amostrais suficientemente grandes.

2.4 Análise Bayesiana do modelo Weibull-exponenciada

Na seção 2.3 foi considerada uma análise clássica do modelo Weibull-exponenciada para os tempos de falha, baseada em resultados assintóticos. Para amostras de tamanhos moderados ou pequenos esses resultados podem não ser confiáveis, tornando-se assim um problema sério uma vez que em geral, trabalha-se com amostras não muito grandes. Um procedimento alternativo que pode contornar esse problema é através da utilização de métodos Bayesianos. O uso de métodos Bayesianos, além de serem uma alternativa de análise, permitem ainda a incorporação de conhecimento a priori através de uma densidade a priori que seja informativa. Nesta seção apresenta-se uma análise Bayesiana do modelo Weibull-exponenciada usando o método de simulação via cadeias de Markov (veja, Gelfand e Smith, 1990).

2.4.1 Densidade conjunta a posteriori para os parâmetros do modelo Weibull-exponenciada

Para representar o grau de conhecimento sobre os parâmetros do modelo Weibull-exponenciada (2.1), consideramos para α , θ e σ as seguintes distribuições a priori:

$$\begin{aligned}\alpha &\sim \Gamma(a_1, b_1), \\ \theta &\sim \Gamma(a_2, b_2), \\ \sigma &\sim \Gamma(a_3, b_3),\end{aligned}\tag{2.16}$$

onde $\Gamma(a_i, b_i)$ denota a distribuição gama com média a_i/b_i e variância a_i/b_i^2 , e a_i e b_i , $i = 1, 2, 3$ são constantes conhecidas baseadas na informação a priori do especialista e da

análise preliminar dos dados.

Assumindo independência entre os parâmetros a densidade a priori conjunta de α , θ e σ é dada por:

$$\pi(\alpha, \theta, \sigma) \propto \alpha^{a_1-1} \theta^{a_2-1} \sigma^{a_3-1} \exp \{-\alpha b_1 - \theta b_2 - \sigma b_3\}, \quad (2.17)$$

onde $\alpha > 0$, $\theta > 0$ e $\sigma > 0$.

Observa-se que no caso em que $a_i = b_i = 0$, ($i = 1, 2, 3$) a densidade a priori conjunta de α , θ e σ dada em (2.17) é imprópria, ou seja,

$$\pi(\alpha, \theta, \sigma) \propto \frac{1}{\alpha \theta \sigma}.$$

Combinando (2.8)-(2.17) e o teorema de Bayes pode-se mostrar que a densidade conjunta a posteriori de α , θ e σ é dada por

$$\begin{aligned} \pi(\alpha, \theta, \sigma | D) &\propto \alpha^{n+a_1-1} \theta^{n+n_2-1} \sigma^{a_3-n\alpha-1} \exp \left\{ -\sum_{i \in F} \left(\frac{t_i}{\sigma} \right)^\alpha - b_1 \alpha - b_2 \theta - b_3 \sigma \right\} \\ &\times \prod_{i \in F} t_i^{\alpha-1} \left(1 - \exp \left\{ \left(\frac{t_i}{\sigma} \right)^\alpha \right\} \right)^{\theta-1} \prod_{i \in C} \left(1 - \left(1 - \exp \left\{ \left(\frac{t_i}{\sigma} \right)^\alpha \right\} \right)^\theta \right), \end{aligned} \quad (2.18)$$

onde $\alpha > 0$, $\theta > 0$, $\sigma > 0$ e D denota o conjunto de dados observados.

Notamos que a densidade a posteriori conjunta (2.18) não é uma densidade padrão, portanto só podemos avaliar as densidades a posteriori marginais por meio de métodos de aproximação, tais como o método de Laplace (Tierney , Kass e Kadane, 1986) ou usando métodos de simulação de Monte Carlo via Cadeias de Markov, tais como Gibbs Sampling (Casella e George, 1992) e Metropolis-Hasting (Chib e Greenberg, 1995). Nesta análise consideramos os métodos de simulação por serem de fácil implementação computacional.

Para obtermos uma amostra a posteriori das densidades marginais α , σ e θ , fazemos uso do algoritmo de Gibbs Sampling que baseia-se em sucessivas gerações das distribuições condicionais a posteriori de $\pi(\alpha|\theta, \sigma, D)$, $\pi(\theta|\alpha, \sigma, D)$ e $\pi(\sigma|\alpha, \theta, D)$.

Da densidade a posteriori conjunt dada em (2.18) pode-se mostrar que as densidades

condicionais a posteriori para o algoritmo de Gibbs são dadas por

$$\begin{aligned}\pi(\alpha|\theta, \sigma, D) &\propto \alpha^{n+a_1-1} \sigma^{-\alpha n} \exp\left\{-\sum_{i \in F} \left(\frac{t_i}{\sigma}\right)^\alpha - b_1 \alpha + \alpha \sum_{i \in F} \log(t_i) + A(\alpha, \sigma, \theta)\right\}, \\ \pi(\theta|\alpha, \sigma, D) &\propto \theta^{a_2+n-1} \exp\left\{-\theta b_2 + A_1(\alpha, \theta, \sigma)\right\}, \\ \pi(\sigma|\alpha, \theta, D) &\propto \sigma^{a_3-n\alpha-1} \exp\left\{-\sum_{i \in F} \left(\frac{t_i}{\sigma}\right)^\alpha - b_3 \theta + A(\alpha, \theta, \sigma)\right\},\end{aligned}\quad (2.19)$$

onde

$$A(\alpha, \theta, \sigma) = (\theta - 1) \sum_{i \in F} \log(1 - \exp\left\{-\left(\frac{t_i}{\sigma}\right)^\alpha\right\}) + \sum_{i \in C} \log(1 - (1 - \exp\left\{-\left(\frac{t_i}{\sigma}\right)^\alpha\right\})^\theta)$$

e

$$A_1(\alpha, \theta, \sigma) = \theta \sum_{i \in F} \log(1 - \exp\left\{-\left(\frac{t_i}{\sigma}\right)^\alpha\right\}) + \sum_{i \in C} \log(1 - (1 - \exp\left\{-\left(\frac{t_i}{\sigma}\right)^\alpha\right\})^\theta).$$

Observe que as densidades dadas em (2.19) não são conhecidas de modo que faremos uso do algoritmo de Metropolis-Hastings para gerar as variáveis α , θ e σ .

2.4.2 Estimador de Bayes da função de sobrevivência e função risco

Considerando as amostras Gibbs geradas em (2.19) de α , θ e σ , pode-se obter estimativas de Monte Carlo para os momentos de interesse. Um caso especial deste é dado pela função de sobrevivência (2.6) para um valor do tempo t_0 especificado. Um estimador de Bayes para $S(t_0)$ com respeito à função perda quadrática é dada por

$$E[S(t_0|D)] = \int \int \int \left(1 - [1 - \exp(-(t_0/\sigma)^\alpha)]^\theta\right) \pi(\alpha, \theta, \sigma|D) d\alpha d\theta d\sigma$$

a qual pode-se aproximar por sua estimativa de Monte Carlo por

$$\hat{S}(t_0) = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S \left[1 - \left[1 - \exp\left(-\left(\frac{t_0}{\sigma^{(r,s)}}\right)^{\alpha^{(r,s)}}\right) \right]^{\theta^{(r,s)}} \right]. \quad (2.20)$$

onde R e S são respectivamente, número total de iterações e simulações do amostrador de Gibbs. Similarmente, uma estimativa da função risco (2.7) para um valor do tempo

t_0 especificado com respeito à função de perda quadrática é dada por $E[h(t_0)|D]$. Uma estimativa de Monte Carlo é dada por

$$\hat{h}(t_0) = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S \frac{\alpha^{(r,s)} \theta^{(r,s)} [1 - \exp(-(\frac{t_0}{\sigma^{(r,s)}}))]^{\theta^{(r,s)}-1} \exp(-(\frac{t_0}{\sigma^{(r,s)}})^{\alpha^{(r,s)}}) (\frac{t_0}{\sigma^{(r,s)}})^{\alpha^{(r,s)}-1}}{\sigma^{(r,s)} [1 - (1 - \exp(-(\frac{t_0}{\sigma^{(r,s)}})^{\alpha^{(r,s)}}))^{\theta^{(r,s)}}]}, \quad (2.21)$$

onde $\alpha^{(r,s)}$, $\sigma^{(r,s)}$ e $\theta^{(r,s)}$ denotam as amostras de α , σ e θ na r -ésima iteração e s -ésima réplica.

2.4.3 Densidade conjunta a posteriori dos parâmetros do modelo Weibull

Consideramos as seguintes densidades a priori para α e σ para os parâmetros da distribuição de Weibull:

$$\alpha \sim \Gamma(a_1, b_1), \quad a_1 \text{ e } b_1 \text{ são conhecidos} \quad (2.22)$$

$$\sigma \sim \Gamma(a_2, b_2), \quad a_2 \text{ e } b_2 \text{ são conhecidos}, \quad (2.23)$$

onde $\Gamma(a_i, b_i)$ denota a distribuição Gama com média a_i/b_i , $i = 1, 2$.

Assumindo independência entre os parâmetros da distribuição Weibull e considerando a função de verossimilhança de (2.8) com $\theta = 1$, pode-se mostrar que a densidade conjunta a posteriori de α e σ é dada por

$$\pi(\alpha, \sigma | D) \propto \alpha^{a_1+r-1} \sigma^{a_2-r\alpha-1} \exp \left\{ - \sum_{i=1}^n \left(\frac{t_i}{\sigma} \right)^\alpha - (\alpha - 1) \sum_{i \in F} \log(t_i) - \alpha b_1 - \sigma b_2 \right\}, \quad (2.24)$$

onde $\alpha > 0$, $\sigma > 0$. Notemos que a distribuição a posteriori dada em (2.24) não tem uma forma padrão conhecida pelo que faremos uso do método MCMC.

Para obtermos uma amostra a posteriori das densidades marginais de α e σ , exploraremos o uso do método do amostrador de Gibbs que baseia-se em sucesivas gerações das distribuições marginais a posteriori $\pi(\alpha|\sigma, D)$ e $\pi(\sigma|\alpha, D)$.

De (2.24) pode-se mostrar que as densidades a posteriori condicionais para o algoritmo de Gibbs são dadas por

$$\begin{aligned}\pi(\alpha|\sigma, D) &\propto \alpha^{r+a_1-1} \sigma^{-r\alpha} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\sigma}\right)^\alpha - b_1\alpha + \alpha \sum_{i \in F} \log(t_i)\right\}, \\ \pi(\sigma|\alpha, D) &\propto \sigma^{a_1-r\alpha-1} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\sigma}\right)^\alpha - b_2\sigma\right\}.\end{aligned}\quad (2.25)$$

Observe que as densidades (2.25) não tem uma forma padrão conhecida e portanto deve-se usar o algoritmo de Metropolis-Hastings para gerar amostras de α e σ .

2.5 Exemplo de aplicação

Nesta seção ilustramos a metodologia apresentadas nas seções anteriores, através da análise do conjunto de dados analizado por Efron (1988) com reanálise em Mudholkar et al. (1995).

Na tabela 2.1, apresenta-se os tempos de sobrevivência em dias de 51 pacientes de cancer na região da cabeça-pescoço.

Tabela 2.1: Tempos de Sobrevivência de pacientes de cancer, com + indicando censura

7	34	42	63	64	74+	83	84	91	108	112
129	133	133	139	140	140	146	149	154	157	160
160	165	173	176	185+	218	225	241	248	273	277
279+	297	319+	405	417	420	440	523	523+	583	594
1101	1116+	1146	1226+	1349+	1412+	1417				

Análise clássica

Considerando a log-verossimilhança dada em (2.9) temos que o estimador de máxima verossimilhança do parâmetro $\eta = (\alpha, \theta, \sigma)$ é obtido pela resolução do sistema de equações não lineares dada em (2.10). Isto é feito através da rotina de Newton-Rapson do pacote Ox.

As estimativas de máxima verossimilhança são dadas por

$$\hat{\alpha} = 0,2943(0,13907), \quad \hat{\theta} = 18,036(27,410), \quad \hat{\sigma} = 0,14404(0,4995).$$

Cada valor entre parênteses ao lado das estimativas representa a raiz quadrada do correspondente elemento da diagonal da matriz $I^{-1}(\hat{\alpha}, \hat{\theta}, \hat{\sigma})$, onde $I(\alpha, \theta, \sigma)$ é a matriz de informação observada. Se admitimos a normalidade assintótica para os estimadores de máxima verossimilhança, $I^{-1}(\hat{\alpha}, \hat{\theta}, \hat{\sigma})$ corresponde à estimativa da matriz de covariâncias assintótica, e pode ser utilizada para fazer inferências sobre os parâmetros.

A estatística da razão de verosimilhanças para testar $H_{01} : \theta = 1$ dada em (2.15) é $\Lambda_1 = 7,7$, e indica a não aceitação do modelo de Weibull a um nível de significância 0,05 ($\chi^2_{(1,0,05)} = 3,8441$).

Análise Bayesiana

Para analisarmos os dados da tabela 2.1, desde uma perspectiva Bayesiana consideramos a densidade a priori para α, θ e σ dada em (2.17) com $a_1 = 8,66$, $b_1 = 29,43$, $a_2 = 31,27$, $b_2 = 1,734$, $a_3 = 2,5614$ e $b_3 = 17,78$ (a escolha desses valores para os parâmetros da densidade a priori, foram baseadas numa opinião experiente combinada com uma análise preliminar dos dados). A partir da densidades condicionais (2.19), geramos para cada um dos parâmetros do modelo Weibull-exponenciada 10 cadeias separadas de Gibbs cada uma com 5.000 iterações, e devido ao fato de todas as densidades condicionais marginais não apresentarem uma forma padrão conhecida utilizamos então o algoritmo de Metropolis-Hastings. Com o objetivo de diminuir o efeito de pontos iniciais desprezam-se as 2.500 primeiras iterações e a partir daí escolhem-se as iterações de 25 em 25, ou seja, para cada parâmetro considera-se a 2.525 – ésima, ..., 5.000 – ésima iterações; de modo que as 10 cadeias fornecem uma amostra Gibbs de tamanho 1000 para cada um dos parâmetros. A convergência das amostras geradas pelo algoritmo de Gibbs com Metropolis-Hastings foi monitorada utilizando o método proposto por Gelman e Rubin (1992) o qual se baseia na técnica de análise de variância.

Na tabela 2.2, apresenta-se um resumo das distribuições a posteriori marginais aproximadas e as estimativas dos fatores de redução de escala potencial \hat{R} (ver por exemplo, Gelman e Rubin, 1992). Observamos na tabela 2.2, que $\sqrt{\hat{R}}$ são bem menores que 1.1, o que nos indica que o algoritmo de Gibbs com Metropolis-Hasting gera para cada parâmetro amostras que convergem em distribuição para as distribuições condicionais marginais (2.19).

Tabela 2.2: Sumário a posteriori dos parâmetros do modelo Weibull-exponenciada

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α	0,3207	0,3197	0,0373	(0,2515 ; 0,3746)	1,000271
θ	18,3323	18,0802	4,1213	(11,1911;24,4389)	1,00014
σ	0,16057	0,1518	0,0703	(0,0550 ; 0,2705)	1,001939

Da tabela 2.2, note que os desvios padrões das estimativas obtidas pelo método MCMC são bem menores que os desvios padrões obtidos invertendo-se a matriz de informação observada. As médias a posteriori são próximas dos estimadores de máxima verossimilhança.

Na figura 2.2, representa-se as densidades a posteriori marginais aproximadas considerando os 1.000 pontos amostrais.

Com a finalidade de escolher entre os modelos Weibull e Weibull-exponenciada desde uma perspectiva Bayesiana ($H_{01} : \theta = 1$), utiliza-se a técnica das densidades preditivas condicionadas ordenadas (CPO) conforme descrito no capítulo 1. Para isso, consideramos a partir das estimativas dos CPOs a estimativa da verossimilhança preditiva para definir o pseudo-fator de Bayes. A estimativa da verossimilhança preditiva sob o modelo de Weibull-exponenciada com as 1.000 amostras de Gibbs resultou em $\hat{f}(t|M_1) = 7,584255 \times 10^{-68}$, e para o modelo de Weibull essa verossimilhança resultou em $\hat{f}(t|M_0) = 3,21158 \times 10^{-78}$. O pseudo-fator de Bayes do modelo Weibull com respeito ao modelo Weibull-exponenciada resultou em $PSFB_{01} = 4,23398 \times 10^{-11} < 1$ o qual indica que o modelo

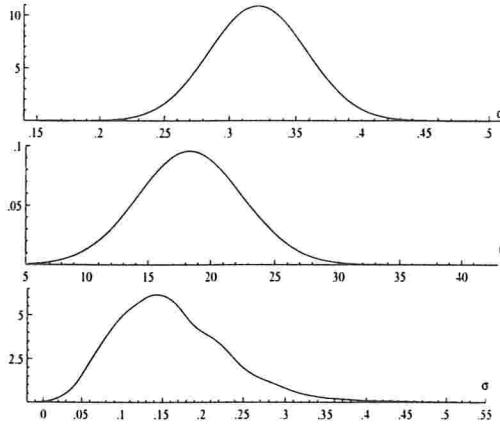


Figura 2.2: Densidades marginais a posteriori approximadas para os parâmetros do modelo Weibull-exponenciado, α , θ e σ

Weibull-exponenciada é mais adequado para ajustar os dados da tabela 2.1 e esse resultado coincide com o resultado do teste de hipóteses clássica. Contudo o enfoque Bayesiano apresenta uma indicação muito forte do que o enfoque clássico que o modelo Weibull-exponenciada é o mais adequado para o ajuste dos dados da tabela 2.1.

Na figura 2.3, plotamos o log a razão dos CPOs dos modelos Weibull e Weibull-exponenciada com respeito ao número de observações. Valores positivos do log da razão dos CPO indica a preferência do primeiro modelo. Notamos a partir da figura 2.3, o gráfico indica que 44 das 51 observações suportam ao modelo Weibull-exponenciada sobre o modelo Weibull. Esse método gráfico também confirma os resultados obtidos pelos métodos clássico e Bayesiano.

Na figura 2.4, temos densidade preditiva para um tempo de vida futuro, t , dada as observações D aproximada pelo método de Monte Carlo com as 1.000 amostras Gibbs dos parâmetros do modelo Weibull-exponenciada.

Na figura 2.5, representamos as função de sobrevivência a posteriori dada em (2.20)

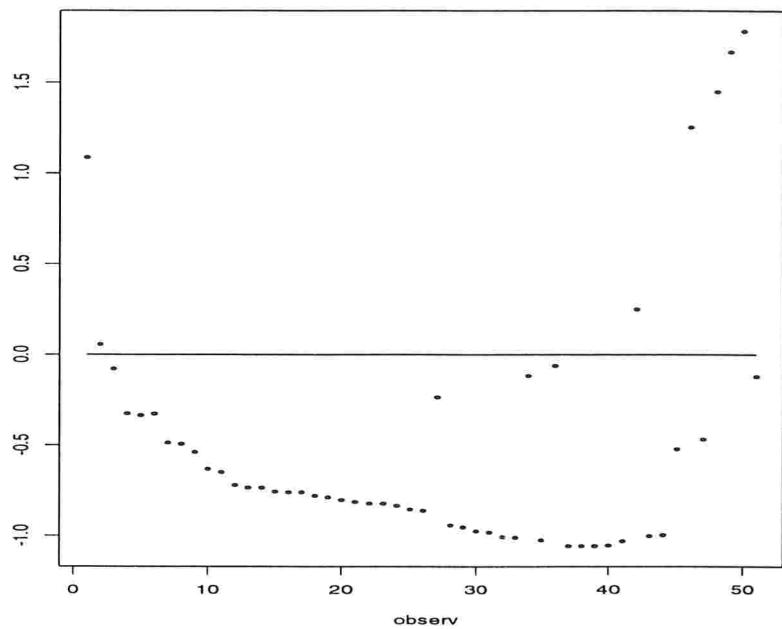


Figura 2.3: Gráfico do log da razão dos CPOs do modelo Weibull e o modelo Weibull-exponenciada

e a função risco a posteriori dada em (2.21), aproximadas pelo método de Monte Carlo com as 1.000 amostras Gibbs.

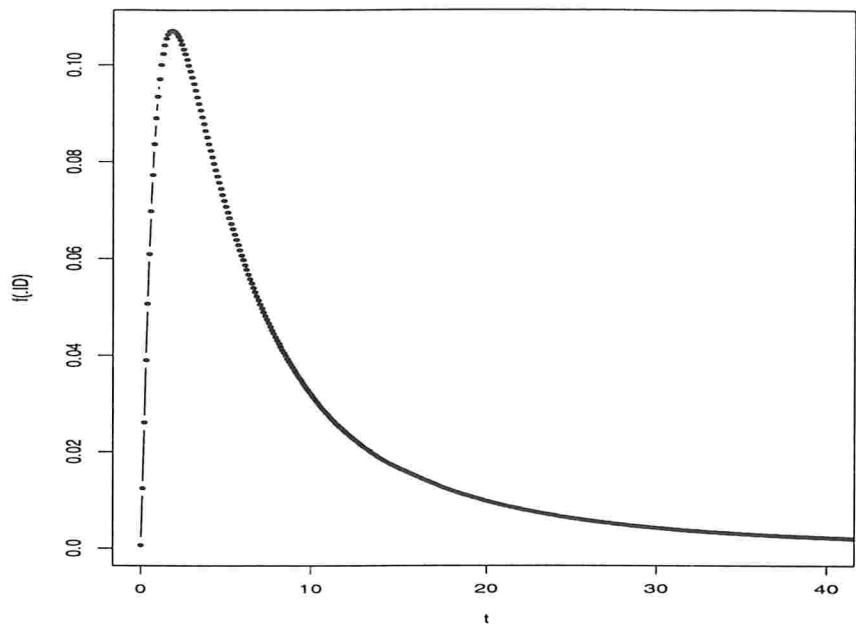


Figura 2.4: Densidade preditiva para tempo de vida futuro no modelo Weibull-exponenciada

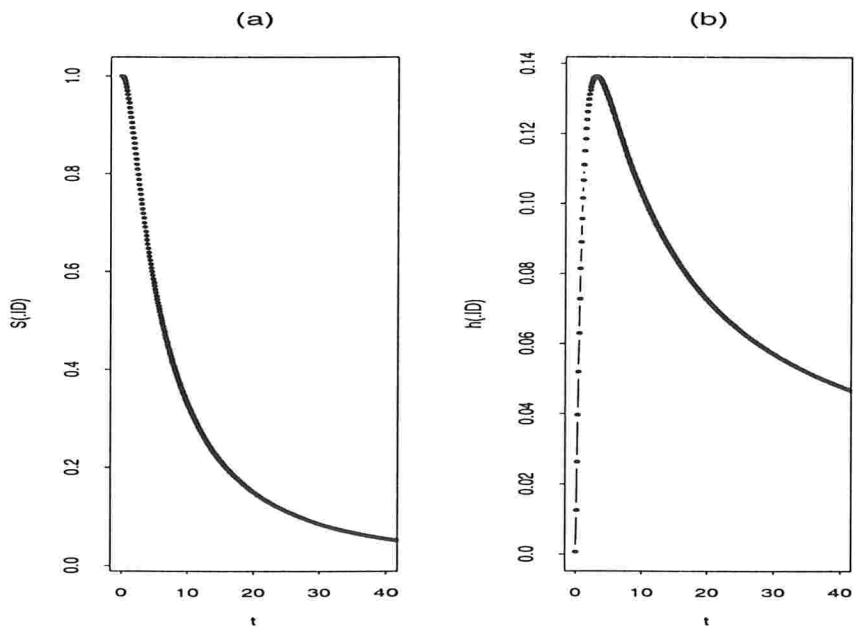


Figura 2.5: As funções : (a) sobrevivência e (b) risco a posteriori para o modelo Weibull-exponenciada

Capítulo 3

Modelos de regressão Weibull-exponenciada

3.1 Introdução

Em muitas situações práticas temos que o tempo de vida é influenciado por uma ou mais variáveis regressoras. Por exemplo, o tempo de vida de um determinado isolante é naturalmente afetado pelo nível de voltagem a que ele é sujeito e o tempo de sobrevivência de um paciente pode depender da idade, da quantidade de cigarros que ele fuma por dia, e de uma série de outros fatores. Uma maneira de determinar o relacionamento entre o tempo de vida e a covariável $\mathbf{x} = (x_1, \dots, x_p)$ é através de um modelo de regressão. Duas classes importantes de modelos de regressão são: modelos de riscos proporcionais para T e modelos de locação escala para o $\log T$. Abordaremos aqui apenas a segunda classe. Um descrição detalhada sobre modelos de riscos proporcionais podem ser obtidas em Cox e Oakes (1989), Kalbleisch e Prentice (1980), entre outros. Neste Capítulo, consideramos que o parâmetro de escala σ da distribuição Weibull-exponenciada dada em (2.1) depende de uma matriz de variáveis explicativas \mathbf{x} .

3.2 O Modelo log-Weibull-exponenciada

Considerando a transformação da variável aleatória $Y = \log T$ e a reparametrização $\sigma = \exp(\mu)$ e $\alpha = 1/\delta$, temos em (2.1), que a função de densidade de probabilidades de Y pode ser escrita como

$$f(y) = \frac{\theta}{\delta} \left[1 - \exp\left\{-\exp\left(\frac{y-\mu}{\delta}\right)\right\} \right]^{\theta-1} \exp\left\{\left(\frac{y-\mu}{\delta}\right) - \exp\left\{\left(\frac{y-\mu}{\delta}\right)\right\}\right\}, \quad (3.1)$$

onde $-\infty < y < \infty$, $\delta > 0$, $\theta > 0$ e $-\infty < \mu < \infty$. Observe que se $\theta = 1$, temos a densidade valor extremo para Y com parâmetro de locação μ e parâmetro de escala δ .

De (3.1), pode-se representar o modelo log-linear dado por

$$Y = \mu + \delta Z \quad (3.2)$$

onde Z tem densidade dada por

$$f(z) = \theta [1 - \exp\{-\exp(z)\}]^{\theta-1} \exp\{z - \exp(z)\}, \quad -\infty < z < \infty, \quad (3.3)$$

onde $\theta > 0$. Observe que quando $\theta = 1$, (3.3) é a densidade valor extremo padrão (ver por exemplo Lawless, 1982)

A figura 3.1, mostra algumas representações da densidade (3.3) para alguns valores dos parâmetros ($\theta = 0, 4; 1; 10; 100$).

A função de sobrevivência é dada por

$$S(y; \mu, \theta, \delta) = 1 - \left\{ 1 - \exp\left(-\exp\left(\frac{y-\mu}{\delta}\right)\right) \right\}^\theta. \quad (3.4)$$

3.3 O Modelo de regressão Weibul-exponenciada

Consideremos agora o modelo de regressão baseado na distribuição log-Weibull-exponenciada, dada em (3.1). Assumimos uma relação linear entre o log do tempo de vida, denotado

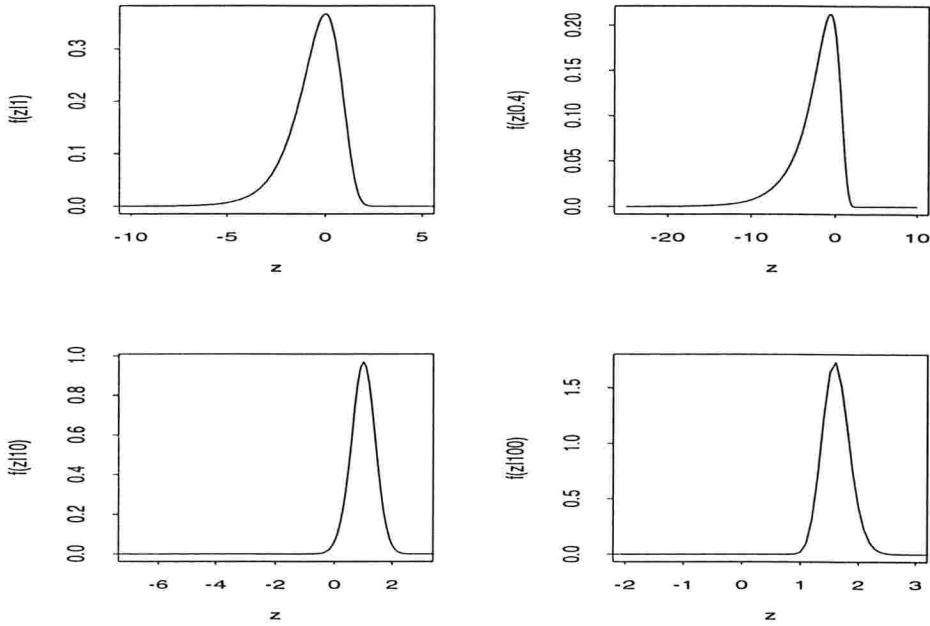


Figura 3.1: Algumas formas da densidade dada em (3.3).

por Y e o vetor de covariáveis $\mathbf{x} = (x_1, \dots, x_p)$. Y dado \mathbf{x} tem distribuição log-Weibul-exponenciada e a relação linear entre \mathbf{x} e Y pode ser representada da forma,

$$Y = \mathbf{x}\beta + \delta Z \quad (3.5)$$

onde $\beta = (\beta_1, \dots, \beta_p)'$ e $\delta > 0$, são parâmetros desconhecidos, e Z tem distribuição dada em (3.3).

A função de sobrevivência de Y dado \mathbf{x} , analogamente a (3.4), é representada por

$$S(y; \theta, \delta, \beta) = 1 - \{1 - \exp(-\exp(\frac{y - \mathbf{x}\beta}{\delta}))\}^\theta. \quad (3.6)$$

Como casos particulares do modelo dado por (3.5), temos os modelos de regressão exponencial para $\theta = \delta = 1$ e para $\theta = 1$ temos o modelo de regressão Weibull.

3.4 Análise clássica para os modelos de regressão Weibull-exponenciada

Nesta seção será apresentada uma análise clássica para a determinação de estimadores de máxima verossimilhança e intervalos de confiança para os parâmetros envolvidos no modelo de regressão dado em (3.5).

3.4.1 Estimação dos parâmetros por máxima verossimilhança

Sejam $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$, n observações independentes do modelo dado por (3.3), onde y_i representa o logarítmico do tempo de vida ou logarítmico do tempo censura, e $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})$ o vetor de covariáveis associadas ao i -ésimo indivíduo. Como antes, F e C denotam o conjunto de indivíduos (ítems) que resultaram em falha ou censura, respectivamente. A função log-verossimilhança para (β, δ, θ) é dada por

$$\begin{aligned} \ell(\beta, \theta, \delta) &= r \log(\theta) - r \log(\delta) + (\theta - 1) \sum_{i \in F} \log[1 - \exp\{-\exp\{\frac{y_i - \mathbf{x}_i \beta}{\delta}\}\}] + \\ &\quad + \sum_{i \in F} \left(\frac{y_i - \mathbf{x}_i \beta}{\delta} \right) - \sum_{i \in F} \exp\{\frac{y_i - \mathbf{x}_i \beta}{\delta}\} + \\ &\quad + \sum_{i \in C} \log[1 - (1 - \exp\{-\exp\{\frac{y_i - \mathbf{x}_i \beta}{\delta}\}\})^\theta], \end{aligned} \quad (3.7)$$

onde r é o número de observações não censuradas.

Os estimador de máxima verossimilhança de (β, δ, θ) é obtido maximizando-se a função log-verossimilhança (3.7), o que equivale a resolver o seguinte sistema de equações não lineares:

$$\frac{\partial \ell}{\partial \theta} = \frac{r}{\theta} + \sum_{i \in F} \log[1 - e^{-e^{z_i}}] - \sum_{i \in C} \frac{(1 - e^{-e^{z_i}})^\theta \log(1 - e^{-e^{z_i}})}{1 - (1 - e^{-e^{z_i}})^\theta} = 0, \quad (3.8)$$

$$\begin{aligned} \frac{\partial \ell}{\partial \delta} &= -\frac{r}{\delta} - \frac{1}{\delta} \sum_{i \in F} z_i + \sum_{i \in F} z_i e^{z_i} - \frac{(\theta - 1)}{\delta} \sum_{i \in F} \frac{z_i e^{z_i - e^{z_i}}}{1 - e^{-e^{z_i}}} - \\ &\quad \frac{\theta}{\delta} \sum_{i \in C} \frac{z_i (1 - e^{-e^{z_i}})^{\theta-1} \log(1 - e^{-e^{z_i}})}{1 - (1 - e^{-e^{z_i}})^\theta} = 0, \end{aligned} \quad (3.9)$$

$$\begin{aligned}\frac{\partial \ell}{\partial \beta_j} = & -\frac{1}{\delta} \sum_{i \in F} x_{ij} + \frac{1}{\delta} \sum_{i=1}^n x_{ij} e^{z_i} - \frac{(\theta-1)}{\delta} \sum_{i \in F} \frac{x_{ij} e^{z_i - \epsilon^{z_i}}}{1 - e^{-\epsilon^{z_i}}} - \\ & \frac{\theta-1}{\delta} \sum_{i \in C} \frac{x_{ij} (1 - e^{-\epsilon^{z_i}})^{\theta-1} \log(1 - e^{-\epsilon^{z_i}})}{1 - (1 - e^{-\epsilon^{z_i}})^\theta} = 0, \quad j = 1, \dots, p,\end{aligned}\tag{3.10}$$

onde $z_i = \frac{(y_i - \mathbf{x}_i \boldsymbol{\beta})}{\delta}$. Como não é possível obter de modo analítico os valores de $\hat{\beta}_1, \dots, \hat{\beta}_p$, $\hat{\delta}$ e $\hat{\theta}$ que satisfazem as equações acima, pode-se recorrer a métodos iterativos, como por exemplo o método de Newton-Rapson para calcular estimativas dos parâmetros.

3.4.2 Inferência sobre os parâmetros $\beta_1, \dots, \beta_p, \delta$ e θ

Para inferência sobre os parâmetros de um determinado modelo, usualmente, utiliza-se a normalidade assintótica dos estimadores de máxima verossimilhança. Assim para $\hat{\beta}_1, \dots, \hat{\beta}_p, \hat{\delta}$ e $\hat{\theta}$ EMVs de $\beta_1, \dots, \beta_p, \delta$ e θ sob certas condições de regularidade (veja, Cox e Hinkley, 1974) tem-se o seguinte resultado para $n \rightarrow \infty$:

$$i^{1/2}(\boldsymbol{\beta}, \delta, \theta) \begin{bmatrix} \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \\ \hat{\delta} \\ \hat{\theta} \end{bmatrix} - \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \\ \delta \\ \theta \end{bmatrix} \xrightarrow{D} N(\mathbf{0}, I_{p+2}) \tag{3.11}$$

onde $i^{1/2}(\boldsymbol{\beta}, \delta, \theta)$ é a matriz raiz quadrada da matriz de informação de Fisher e I_{p+2} é a matriz identidade de ordem $(p+2) \times (p+2)$. A normalidade assintótica (3.11) segue sendo válida se $i(\boldsymbol{\beta}, \delta, \theta)$ é substituído por seu estimador consistente $i(\hat{\boldsymbol{\beta}}, \hat{\delta}, \hat{\theta})$ ou simplesmente se considerarmos a matriz de informação avaliada nos estimadores de máxima verossimilhança dada por

$$I(\hat{\beta}, \hat{\delta}, \hat{\theta}) = \begin{bmatrix} -\frac{\partial^2 \ell}{\partial \beta_1} & \cdots & -\frac{\partial^2 \ell}{\partial \beta_1 \partial \beta_p} & -\frac{\partial^2 \ell}{\partial \beta_1 \partial \delta} & -\frac{\partial^2 \ell}{\partial \beta_p \partial \theta} \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ -\frac{\partial^2 \ell}{\partial \beta_p \partial \beta_1} & \cdots & -\frac{\partial^2 \ell}{\partial \beta_p^2} & -\frac{\partial^2 \ell}{\partial \beta_p \partial \delta} & -\frac{\partial^2 \ell}{\partial \beta_p \partial \theta} \\ & & -\frac{\partial^2 \ell}{\partial \delta^2} & -\frac{\partial^2 \ell}{\partial \delta \partial \theta} & -\frac{\partial^2 \ell}{\partial \theta^2} \\ \text{simétrica} & & & & \end{bmatrix} \Big|_{(\beta, \delta, \theta) = (\hat{\beta}, \hat{\delta}, \hat{\theta})}. \quad (3.12)$$

Considerando a distribuição normal assintótica (3.11) para os estimadores de máxima verossimilhança $\hat{\beta}$, $\hat{\delta}$ e $\hat{\theta}$ considerando um nível de confiança $100(1 - \alpha)\%$, intervalos de confiança são dados por:

$$ic(\beta_j) = (\hat{\beta}_j - z_{\alpha/2}(I^{(jj)})^{1/2}; \hat{\beta}_j + z_{\alpha/2}(I^{(jj)})^{1/2}), \quad j = 1, \dots, p \quad (3.13)$$

$$ic(\delta) = (\hat{\delta} - z_{\alpha/2}(I^{(p+1,p+1)})^{1/2}; \hat{\delta} + z_{\alpha/2}(I^{(p+1,p+1)})^{1/2}) \quad (3.14)$$

$$ic(\theta) = (\hat{\theta} - z_{\alpha/2}(I^{(p+2,p+2)})^{1/2}; \hat{\theta} + z_{\alpha/2}(I^{(p+2,p+2)})^{1/2}), \quad (3.15)$$

onde $I^{(jj)}$ é o j -ésimo elemento diagonal da matriz $I^{-1}(\hat{\beta}, \hat{\delta}, \hat{\theta})$, $j = 1, \dots, p$ e $z_{\alpha/2}$ é o percentil da distribuição normal padrão correspondente ao coeficiente de confiança estabelecido.

3.4.3 Teste para modelo de regressão Weibull

Se consideramos as seguintes hipóteses

$$H_0 : \theta = 1 \text{ versus } H_1 : \theta \neq 1, \quad (3.16)$$

estaremos testando a adequação do modelo de regressão Weibull para os tempos de vida.

Usamos especificamente a estatística

$$RV = -2 \log \left\{ \frac{L(1, \tilde{\delta}, \tilde{\beta})}{L(\hat{\theta}, \hat{\delta}, \hat{\beta})} \right\}, \quad (3.17)$$

onde $L(1, \tilde{\delta}, \tilde{\beta})$ corresponde ao máximo de $L(1, \delta, \beta)$ dado por

$$L(1, \delta, \beta) = \prod_{i \in F} \frac{1}{\delta} \exp\{z_i - e^{z_i}\} \prod_{i \in C} e^{-e^{z_i}}$$

sendo $z_i = (y_i - \mathbf{x}\beta)/\delta$, $i = 1, \dots, n$. A estatística (3.17) para grande amostras, tem distribuição qui-quadrado com *um* grau de liberdade sob a hipóteses H_0 (veja Rao, 1973).

3.4.4 Teste para o modelo de regressão exponencial

Considerar $H_0 : \theta = 1, \delta = 1$ versus, $H_1 : \theta \neq 1$, ou $\delta \neq 1$ equivale a testar se os dados admitem o modelo de regressão exponencial para T e a estatística usada para o teste é

$$RV = -2 \log \left\{ \frac{L(1, 1, \tilde{\beta})}{L(\hat{\theta}, \hat{\delta}, \tilde{\beta})} \right\}, \quad (3.18)$$

onde $\tilde{\beta}$ é o valor que maximiza a verossimilhança

$$L(1, 1, \beta) = \prod_{i \in F} \exp\{(y_i - \mathbf{x}\beta) - \exp\{y_i - \mathbf{x}_i\beta\}\} \prod_{i \in C} \exp\{-\exp(y_i - \mathbf{x}_i\beta)\}.$$

A estatística da razão de verossimilhança (3.18), sob H_0 tem distribuição assintótica qui-quadrado com 2 graus de liberdade (veja Rao, 1973).

3.5 Análise Bayesiana para o modelo de regressão Weibull-exponenciada

O procedimento clássico usual para análise de modelos de regressão Weibull-exponenciada consiste inicialmente, da utilização de alguns método iterativo para obtenção das estimativas dos parâmetros (estimativas de máxima verossimilhança). Como discutido na seção 3.4, as distribuições dos estimadores obtidos desta forma somente são conhecidas assintoticamente, o que pode resultar em um problema sério, uma vez que, em geral, trabalha-se com amostras não muito grandes. Desta forma, inferências baseadas nessas estatísticas podem não ser confiáveis, conforme mostra o exemplo 1 com dados simulados na seção 3.6. Um procedimento alternativo que pode contornar esse problema é através da utilização de métodos Bayesianos, que permite a incorporação de conhecimento a priori dos parâmetros do modelo através de uma densidade a priori.

Nesta seção apresentaremos uma análise Bayesiana do modelo de regressão Weibull-exponenciada usando métodos de simulação de Monte Carlo via Cadeias de Markov (MCMC), tais como Gibbs sampling (Casella e George, 1992) e Metropolis-Hastings (Chib e Greenberg, 1995).

3.5.1 Densidade conjunta a posteriori dos parâmetros do modelo de regressão Weibull-exponenciada

Para representar o grau de conhecimento sobre os parâmetros do modelo de regressão Weibull-exponenciada (3.5), consideramos as seguintes densidades à priori para θ , δ e β :

$$\begin{aligned}\theta &\sim \Gamma(c_1, d_1), \\ \delta &\sim \Gamma(c_2, d_2), \\ \beta_j &\sim N(\mu_{0j}, \sigma_{0j}^2), \quad j = 1, \dots, p,\end{aligned}\tag{3.19}$$

onde $\Gamma(c, d)$ denota a distribuição Gama com média c/d , $N(\mu, \sigma^2)$ denota a distribuição Normal com média μ e variância σ^2 , e $c_i, d_i, i = 1, 2$ e $\mu_{0j}, \sigma_{0j}, j = 1, \dots, p$ são constantes conhecidas baseadas na informação a priori do especialista e da análise preliminar dos dados.

Assumindo independência entre os parâmetros a densidade a priori conjunta é dada por:

$$\pi(\theta, \delta, \beta) \propto \theta^{c_1-1} \delta^{c_2-1} \exp \left\{ -d_1 \theta - d_2 \delta - \frac{1}{2} \sum_{j=1}^p \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}} \right)^2 \right\}, \tag{3.20}$$

onde $\theta > 0$, $\delta > 0$ e $-\infty < \beta_j < \infty$. Priori não informativa é obtida fazendo $c_i = d_i = 0$, $i = 1, 2$, $\mu_{0j} = 0$ e $\sigma_{0j}^2 \rightarrow \infty$, $j = 1, \dots, p$. Isto resulta em

$$\pi(\theta, \delta, \beta) \propto \frac{1}{\theta \delta}.$$

De (3.7) pode-se mostrar que a função verossimilhança para θ , δ e β é dada por $L(\theta, \delta, \beta) = \exp \{ \ell(\theta, \delta, \beta) \}$, isto é,

$$L(\theta, \delta, \beta) = \theta^r \delta^{-r} \exp \left\{ \sum_{i \in F} \left(\frac{y_i - \mathbf{x}_i \beta}{\delta} \right) - \sum_{i \in F} \exp \left\{ \frac{y_i - \mathbf{x}_i \beta}{\delta} \right\} \right\} \times \Psi_1(\theta, \delta, \beta), \tag{3.21}$$

onde

$$\begin{aligned}\Psi_1(\theta, \delta, \beta) &= \\ &\prod_{i \in F} [1 - \exp\{-\exp(\frac{y_i - \mathbf{x}_i \beta}{\delta})\}]^{\theta-1} \prod_{i \in C} [1 - [1 - \exp\{-\exp(\frac{y_i - \mathbf{x}_i \beta}{\delta})\}]^{\theta-1}]\end{aligned}$$

Combinando a densidade a priori conjunta (3.20) com a função verossimilhança (3.21) e o teorema de Bayes pode-se mostrar que a densidade a posteriori conjunta de θ , β e β é dada por

$$\begin{aligned}\pi(\theta, \delta, \beta | D) &\propto \theta^{r+c_1-1} \delta^{-r+c_2-1} \exp\left\{-d_1\theta - d_2\delta - \frac{1}{2} \sum_{j=1}^p \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}}\right)^2\right. \\ &\quad \left. + \sum_{i \in F} \left(\frac{y_i - \mathbf{x}_i \beta}{\delta}\right) - \sum_{i \in F} \exp\left\{\frac{y_i - \mathbf{x}_i \beta}{\delta}\right\}\right\} \times \Psi_1(\theta, \delta, \beta),\end{aligned}\quad (3.23)$$

onde $\theta > 0$, $\delta > 0$, $\beta_j \in \mathcal{R}$, $j = 1, \dots, p$ e D é o conjunto de dados. Conjecturamos que a densidade a posteriori conjunta para o caso da priori não informativa é própria embora ainda este seja um problema em aberto.

Notamos aqui que a densidade a posteriori conjunta não é uma densidade padrão. portanto só podemos avaliar as densidades a posteriori marginais por meio de métodos de aproximação, tais como o método de Laplace (Tierney, Kass e Kadane, 1986) ou usando métodos de simulação de Monte Carlo em Cadeias de Markov, tais como Gibbs Sampling (Casella e George, 1992) e Metropolis-Hasting (Chib E Greenberg, 1995). Nesta trabalho optamos pelos métodos de simulação por serem de fácil implementação computacional.

Para obtermos uma amostra a posteriori das densidades a posteriori marginais de $\theta > 0$, $\delta > 0$, $\beta_j \in \mathcal{R}$, $j = 1, \dots, p$, exploramos o uso do método do amostrador de Gibbs que baseia-se em sucessivas gerações das distribuições condicionais a posteriori $\pi(\theta|\delta, \beta, D)$, $\pi(\delta|\theta, \beta, D)$ e $\pi(\beta_j|\theta, \delta, \beta_{(-j)}, D)$ com $\beta_{(-j)} = (\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_p)$.

Pode-se mostrar que as densidades a posteriori condicionais para o algoritmo de Gibbs são dadas por

$$\pi(\theta|\delta, \beta, D) \propto \theta^{r+c_1-1} \exp\{\theta(B(\delta, \beta) - d_1)\}$$

$$\begin{aligned}
& + \sum_{i \in C} \log [1 - [1 - \exp \{-\exp(\frac{y_i - \mathbf{x}_i \beta}{\delta})\}]^\theta] \Big\}, \quad (3.24) \\
\pi(\delta|\theta, \beta, D) & \propto \delta^{-r+c_2-1} \exp \left\{ -d_2 \delta + \sum_{i \in F} \left(\frac{y_i - \mathbf{x}_i \beta}{\delta} \right) - \right. \\
& \quad \left. \sum_{i \in F} \exp \left\{ \frac{y_i - \mathbf{x}_i \beta}{\delta} \right\} + \log (\Psi_1(\theta, \delta, \beta)) \right\}, \\
\pi(\beta_j|\theta, \delta, \beta_{-j}, D) & \propto \exp \left\{ -\frac{1}{2} \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}} \right)^2 - \sum_{i \in F} \left(\frac{x_{ji} \beta_j}{\delta} \right) - \right. \\
& \quad \left. \sum_{i \in F} \exp \left\{ \frac{y_i - \mathbf{x}_i \beta}{\delta} \right\} + \log \Psi_1(\theta, \delta, \beta) \right\},
\end{aligned}$$

onde $B(\delta, \beta) = \prod_{i \in F} [1 - \exp \{-\exp(\frac{y_i - \mathbf{x}_i \beta}{\delta})\}]$.

Observe que as densidades em (3.24) não tem uma forma padrão conhecida, portanto devemos usar o algoritmo de Metropolis-Hastings dentro do ciclo do algoritmo de Gibbs para gerar θ , δ e β_j , $j = 1, \dots, p$.

3.5.2 Estimador de Bayes da função de sobrevivência no modelo de regressão Weibull-exponenciada

Considerando as amostras Gibbs geradas em (3.24) de θ , δ e β , pode-se obter estimativas de Monte Carlo para os momentos de interesse. Um caso especial deste é dado pela função de sobrevivência (3.4) para valores da covariáveis \mathbf{x}_h dada. Um estimador de Bayes para $S(Y_h)$ com respeito à função perda quadrática é dada por

$$E[S(Y_h|D)] = \int \left[1 - \left\{ 1 - \exp \left(-\exp \left(\frac{Y_h - \mathbf{x}_h \beta}{\delta} \right) \right) \right\}^{\theta-1} \right] \pi(\theta, \delta, \beta|D) d\theta d\delta d\beta.$$

Assim, essa integral pode ser aproximada por sua estimativa de Monte Carlo dada por

$$\hat{S}(Y_h|D) = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S \left[1 - \left\{ 1 - \exp \left(-\exp \left(\frac{y_h - x_{1h} \beta_1^{(r,s)} - \dots - x_{ph} \beta_p^{(r,s)}}{\delta^{(r,s)}} \right) \right) \right\}^{\theta^{(r,s)}-1} \right], \quad (3.25)$$

onde $\theta^{(r,s)}$, $\delta^{(r,s)}$, e $\beta_1^{(r,s)}, \dots, \beta_p^{(r,s)}$ denotam amostras de θ , δ e β_1, \dots, β_p , respectivamente na r -ésima iteração e na s -ésima replica, R e S são respectivamente, número total de

iterações e simulações no amostrador de Gibbs e \mathbf{x}_h é usualmente a mediana dos valores da covariável. Similarmente, pode-se determinar uma estimativa para a função risco.

3.5.3 Densidade conjunta a posteriori dos parâmetros do modelo de regressão Weibull

Consideramos as seguintes densidades a priori para δ e β :

$$\begin{aligned}\delta &\sim \Gamma(c_2, d_2), \\ \beta_j &\sim N(\mu_{0j}, \sigma_{0j}^2), \quad j = 1, \dots, p,\end{aligned}\tag{3.26}$$

onde $\Gamma(c, d)$ denota a distribuição Gama com média c/d , $N(\mu, \sigma^2)$ denota a distribuição Normal com média μ e variância σ^2 , e c_2 , d_2 , e μ_{0j} , σ_{0j} , $j = 1, \dots, p$ são constantes conhecidas baseadas na informação a priori do especialista e da análise preliminar dos dados.

Assumindo independência entre os parâmetros do modelo de regressão Weibull e considerando a função de verossimilhança (3.21) com $\theta = 1$, pode-se mostrar que a densidade a posteriori conjunta de δ e β é dada por

$$\begin{aligned}\pi(\delta, \beta | D) \propto & \delta^{-r+c_2-1} \exp \left\{ -d_2 \delta - \frac{1}{2} \sum_{j=1}^p \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}} \right)^2 + \sum_{i \in F} \left(\frac{y_i - \mathbf{x}_i \beta}{\delta} \right) - \right. \\ & \left. - \sum_{i=1}^n \exp \left\{ \frac{y_i - \mathbf{x}_i \beta}{\delta} \right\} \right\},\end{aligned}\tag{3.27}$$

onde $\delta > 0$, $\beta_j \in \mathcal{R}$, $j = 1, \dots, p$ e D é o conjunto de dados.

Para obtermos uma amostra a posteriori das densidades marginais a posteriori de $\delta > 0$, $\beta_j \in \mathcal{R}$, $j = 1, \dots, p$, exploramos o uso do algoritmo Gibbs que baseia-se em sucesivas gerações das distribuições condicionais a posteriori $\pi(\delta | \theta, \beta, D)$ e $\pi(\beta_j | \theta, \delta, \beta_{(-j)}, D)$ com $\beta_{(-j)} = (\beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_p)$.

Pode-se mostrar que as densidades a posteriori condicionais para o algoritmo de Gibbs

são dadas por

$$\begin{aligned}\pi(\delta|, \beta, D) &\propto \delta^{-r+c_2-1} \exp \left\{ -d_2 \delta + \sum_{i \in F} \left(\frac{y_i - \mathbf{x}_i \beta}{\delta} \right) - \sum_{i=1}^n \exp \left\{ \frac{y_i - \mathbf{x}_i \beta}{\delta} \right\} \right\}, \\ \pi(\beta_j | \delta, \beta_{-j}, D) &\propto \exp \left\{ -\frac{1}{2} \left(\frac{\beta_j - \mu_{0j}}{\sigma_{0j}} \right)^2 - \sum_{i \in F} \left(\frac{x_{ji} \beta_j}{\delta} \right) - \sum_{i=1}^n \exp \left\{ \frac{y_i - \mathbf{x}_i \beta}{\delta} \right\} \right\}.\end{aligned}\quad (3.28)$$

Observe que as densidades em (3.28) não tem uma forma padrão conhecida, portanto devemos usar o algoritmo de Metropolis-Hastings para gerar δ e β_j , $j = 1, \dots, p$.

3.6 Exemplos de aplicação

O objetivo desta seção é ilustrar a metodologia apresentada nas seções anteriores, através da análise de 2 conjuntos de dados

3.6.1 Exemplo 1-Dados simulados

Na tabela 3.1, temos 30 observações geradas da densidade (3.3) com $z = (y - \beta_0 - \beta_1 x)/\delta$ e parâmetros $\theta = 4$, $\delta = 2$, $\beta_0 = 4$ e $\beta_1 = 2$, não incluindo dados censurados.

Tabela 3.1: Dados gerados para modelo de regressão Weibull-exponenciada.

i	x_i	y_i	i	x_i	y_i	i	x_i	y_i
1	2.0	9.746	11	4.3	14.443	21	14	33.409
2	2.2	8.686	12	6.0	16.701	22	15	35.819
3	3.5	10.260	13	4.3	15.614	23	16	37.574
4	4.4	15.224	14	6.3	16.649	24	17	38.317
5	5.4	15.224	15	8.0	21.615	25	21	46.948
6	6.6	17.051	16	9	23.089	26	25	56.662
7	7.0	19.883	17	10	25.013	27	24	50.382
8	8.5	21.571	18	11	28.214	28	16	37.283
9	9.0	21.264	19	12	27.392	29	15	35.422
10	7.8	19.275	20	13	32.209	30	25	54.480

Análise clássica

Considerando a log-verossimilhança dada em (3.7) temos que o estimador de máxima verossimilhança do parâmetro $\eta = (\theta, \delta, \beta_0, \beta_1)$ é obtido pela resolução do sistema de equações não lineares dada em (3.8). Isto é feito através da rotina de Newton-Rapson do pacote **Ox**.

As estimativas de máxima verossimilhança são dadas por

$$\hat{\theta} = 3,566(5,3619), \quad \hat{\delta} = 1,781(2,4623), \quad \hat{\beta}_0 = 4,123(2,1447), \quad \text{e} \quad \hat{\beta}_1 = 1,99(0,0295).$$

Cada valor entre parênteses ao lado das estimativas representa a raiz quadrada do correspondente elemento da diagonal de $I^{-1}(\hat{\theta}, \hat{\delta}, \hat{\beta}_0, \hat{\beta}_1)$, onde $I(\theta, \delta, \beta_0, \beta_1)$ é a matriz de informação observada. Se admitimos a normalidade assintótica para a distribuição dos estimadores de máxima verossimilhança, $I^{-1}(\hat{\theta}, \hat{\delta}, \hat{\beta}_0, \hat{\beta}_1)$ corresponde à estimativa da matriz de covariâncias assintótica, e pode ser utilizada para computar inferências sobre os parâmetros.

A estatística de razão de razão de verossimilhanças para testar $H_0 : \theta = 1$ em (3.16) resultou em $RV = 1,06844$, e indica a aceitação do modelo Weibull a um nível de significância 0,05 ($\chi^2_{(1,0,05)} = 3,8441$).

Análise bayesiana

Para analisarmos os dados da tabela 3.1, desde uma perspectiva Bayesiana consideramos a densidade a priori para θ, δ, β_0 e β_1 dada em (3.19) com $c_1 = 4, d_1 = 1, c_2 = 32, d_2 = 16, \mu_{00} = 4, \sigma_{00}^2 = 3, \mu_{01} = 2$ e $\sigma_{01}^2 = 0,009$ (a escolha desses valores para os parâmetros da densidade a priori, foram baseadas numa opinião experiente combinada com uma análise preliminar dos dados). A partir das densidades condicionais (3.24), geramos para cada um dos parâmetros do modelo de regressão Weibull-exponenciada 10 cadeias separadas de Gibbs cada uma com 2.000 iterações, e devido ao fato de todas as densidades condicionais marginais não apresentarem uma forma padrão conhecida utilizamos então o algoritmo

de Metropolis-Hasting dentro do ciclo do amostrador de Gibbs. Com o objetivo de diminuir o efeito de pontos iniciais desprezam-se as 1.000 primeiras iterações e a partir de daí escolhem-se as iterações de 10 em 10 , ou seja, para cada parâmetro considera-se a 1.010 – ésima, . . . , 2.000 – ésima iterações; daí as 10 cadeias fornecem um amostra Gibbs de tamanho 1.000 para cada um dos parâmetros. A convergência das amostras geradas pelo algoritmo de Gibbs com Metropolis-Hasting foi monitorada utilizando-se o método proposto por Gelman e Rubin (1992) o qual se baseia na técnica de análise de variância.

Na tabela 3.2, apresenta-se um resumo das distribuições marginais a posteriori aproximadas e as estimativas dos fatores de redução de escala potencial \hat{R} (ver por exemplo, Gelman e Rubin, 1992). Observamos na tabela 3.2, que $\sqrt{\hat{R}}$ são bem menores que 1.1, o que nos indica que o algoritmo de Gibbs com Metropolis-Hasting geram para cada parâmetro amostras que convergem em distribuição para as distribuições condicionais marginais em (3.24).

Tabela 3.2: Sumário a posteriori dos parâmetros do modelo de Regressão Weibull-exponenciada

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
β_0	3,8506	3,8403	0,4469	(2,9932 ; 4,7162)	1,00667
β_1	1,9993	1,9998	0,0288	(1,9382 ; 2,0563)	1,00008
δ	2,0087	1,9862	0,2913	(1,4209 ; 2,6573)	1,00156
θ	3,9814	3,9459	0,7481	(1,0779 ; 5,6109)	1,00089

Na figura 3.2, representa-se as densidades a posteriori marginais aproximadas considerando os 1.000 pontos amostrais. Note que os desvios padrões obtidos com o enfoque Bayesiano são bem menores do que os obtidos com o método de máxima verossimilhança.

Com a finalidade de selecionar entre o modelo de regressão Weibull e o modelo de regressão Weibull-exponenciada o que se ajusta melhor aos dados da tabela (3.1) desde uma perspectiva Bayesiana (o teste de hipóteses dada em (3.16)), utilizamos a técnica

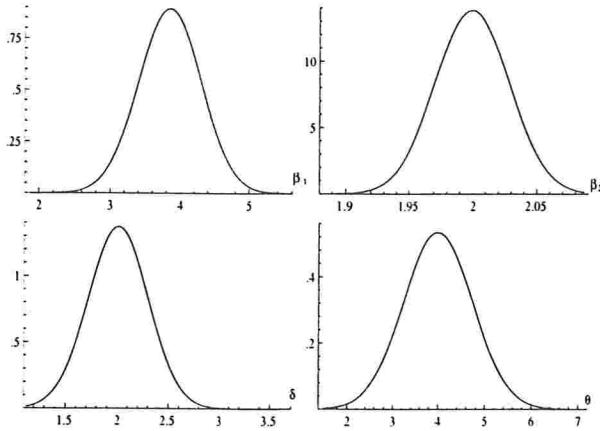


Figura 3.2: Densidades marginais a posteriori para os parâmetros do modelo de regressão Weibull-exponenciado, β_0 , β_1 , δ and θ

das densidades preditivas condicionadas ordenadas (CPO) (veja capítulo 1). Para isso, consideramos a partir das estimativas do CPO a estimativa da verossimilhança preditiva para definir o pseudo-fator de Bayes. A estimativa da verossimilhança preditiva do modelo de regressão Weibull-exponenciada resultou em $\hat{f}(\underline{y}|M_1) = 1.792 \times 10^{-20}$, e para o modelo de regressão Weibull essa verossimilhança resultou em $\hat{f}(\underline{y}|M_0) = 1.47 \times 10^{-28}$. O pseudo-fator de Bayes do modelo de regressão Weibull com respeito o modelo de regressão Weibull-exponenciada resultou em $PSFB_{01} = 2.8031 \times 10^{-9} < 1$ o qual indica que o modelo de regressão Weibull não é adequado para ajustar os dados da tabela 3.1 e esse resultado contraria o resultado do teste de hipóteses clássica (veja análise clássica acima). Este resultado nos mostra que parece ser temerário utilizar a distribuição assintótica do teste da razão de verossimilhança, pois ela está rejeitando uma hipótese verdadeira; resultados de simulação como o teste da razão parecem indicar que este apresenta níveis de significância bem maiores que os níveis teóricos que são estabelecidos.

O procedimento gráfico descrito no capítulo 1, é adotado para verificarmos o melhor modelo entre os modelos estudados. Considerando os dados da tabela 3.1, apresenta-se na tabela 3.3, os valores das estimativas da densidade preditivas condicionadas ordenadas

$c_i(l) = \hat{f}(y_i|D_{(i)})$, $i = 1, \dots, n$ aproximados por seu estimador de Monte Carlo baseadas nas amostras Gibbs, para cada modelo.

Na figura 3.3, apresenta-se os gráficos dos $c_i(l)$ contra os i para os modelos de regressão Weibull e Weibull-exponenciada. Da tabela 3.3 e figura 3.3, observamos que o melhor modelo é o modelo Weibull-exponenciada.

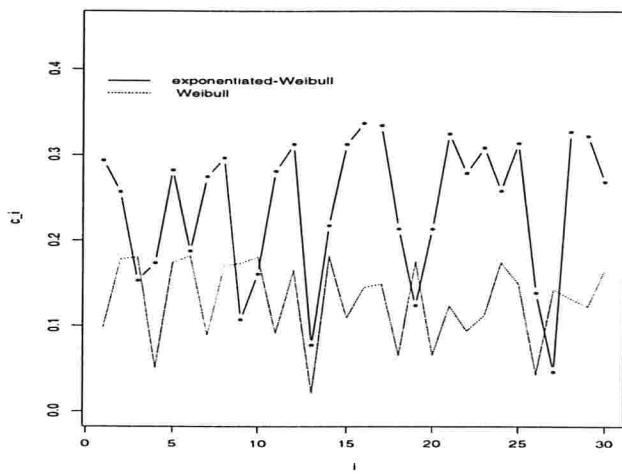


Figura 3.3: Plot de $c_i(l)$ versus i

3.6.2 Exemplo 2-Teste de vida de Isolantes

Os dados da tabela 3.4, são resultados de um teste acelerado de vida para isolantes de sistemas de uma nova classe **H**, e referentes ao tempo em horas que os isolantes levaram para tornar-se defeituosos. Os testes foram utilizadas em pequenos motores, para temperaturas elevadas. Dez motores foram colocadas para trabalhar a temperaturas de 190, 220, 240, e 260 °C, e inspecionados periodicamente para detectar a ocorrência de falha. Os Dados da tabela 3.4 são referentes à média entre o tempo em que a falha foi observada e o tempo de inspeção anterior. Os tempos entre checagem foram 7, 4, e 2 dias para as respectivas temperaturas. Estes dados foram analisados por Nelson (1990), que considerou o modelo de regressão log-normal para o ajuste dos tempos de vida.

i	Weibull ($c_i(1)$)	Weibull-exponenciada ($c_i(2)$)
1	0.0986040816	0.2935935436
2	0.1784192508	0.2569415188
3	0.1806369442	0.1532483601
4	0.0507562589	0.1736054647
5	0.1743275337	0.2823201432
6	0.1820799199	0.187552296
7	0.0889940721	0.2744424237
8	0.1705153204	0.2966126858
9	0.1735079874	0.1074380815
10	0.1808018307	0.1605345992
11	0.0918807093	0.2808851279
12	0.1657797775	0.312313336
13	0.0214895793	0.0774456402
14	0.1816987583	0.2175871757
15	0.1095415567	0.3124269618
16	0.1454376905	0.3370757668
17	0.1494318965	0.334371381
18	0.0652862838	0.2136894111
19	0.1757263783	0.1238903749
20	0.0656442795	0.2133202441
21	0.1238354895	0.3244610886
22	0.0940349458	0.2788427441
23	0.1123034737	0.3083304641
24	0.1745519103	0.2580601204
25	0.149597314	0.3134202149
26	0.0431443935	0.1390255184
27	0.1429360415	0.0460396558
28	0.1322505672	0.3270160164
29	0.1228709264	0.3220973127
30	0.1655935696	0.2685402203
$c(\ell)$	1.47×10^{-28}	1.792×10^{-20}

Tabela 3.3: Valores das estimativas dos $c_i(l)$, $i = 1 \dots, n$ para os modelos de regressão Weibull e Weibull-exponenciada

Tabela 3.4: Tempo de vida de isolantes (horas) para cada temperatura, em graus centígrados, com + indicando censura

Temperatura $^{\circ}C$			
190	220	240	260
7228	1764	1175	1128
7228	2436	1521	1464
7228	2436	1569	1512
8448	2436+	1617	1632
9167	2436	1665	1632+
9167	2436	1665	1632+
9167	3108	1713	1632+
9167	3108	1761	1632+
10511	3108	1881+	1632+
10511	3108	1953	1896

Análise clássica

Transformando a temperatura para assumir o relacionamento Arrhenius, conforme sugerido na literatura de engenharia, usamos o modelo

$$y_i = \beta_0 + \beta_1 x_i + \delta z_i, \quad i = 1, \dots, 40, \quad (3.29)$$

onde $y_i = \log(\text{tempo de vida do isolante } i)$, $x_i = 1000/(\text{temperatura } ^{\circ}C \text{ do } i\text{-ésimo motor} + 273.20)$, β_0 , β_1 e δ são parâmetros desconhecidos e Z_i tem distribuição dada em (3.3) com parâmetro θ . Para estimar os parâmetros, consideramos o mesmo procedimento dado anteriormente.

Como resultado as estimativas de máxima verossimilhança são dados por

$$\hat{\beta}_0 = -4,7261(1,5351), \quad \hat{\beta}_1 = 6,0491(0,4293), \quad \hat{\delta} = 0,6264(0,52895), \quad \hat{\theta} = 8,2099(19,21)$$

Para estudar a possibilidade de utilizar o modelo de regressão Weibull para os dados da tabela 3.4. Consideramos o teste de hipótese ($H_0 : \theta = 1$ vs $H_1 : \theta \neq 1$) e temos que a estatística da razão de verossimilhanças (3.17), resultou em, $RV = 1,38319$ e este resultado implica que o modelo Weibull é estatisticamente adequado para ajustar os dados ao nível de significância de 5%.

Análise Bayesiana para os tempos de vida de isolantes consideramos o modelo de regressão Weibull-exponenciada

Para analisarmos os dados da tabela 3.4, desde uma perspectiva Bayesiana consideramos o modelo de regressão Weibull-exponenciada com as densidades a priori para θ, δ, β_0 e β_1 dadas em (3.19) com $c_1 = 0,186$, $d_1 = 0,0227$, $c_2 = 1,422$, $d_2 = 2,293$, $\mu_{00} = -3,8415$, $\sigma_{00}^2 = 1,5$, $\mu_{01} = 5,95$ e $\sigma_{01}^2 = 10,5$. A partir da densidades condicionais (3.24) geramos uma amostra Gibbs através do algoritmo Gibbs com Metropolis-Hastings da seguinte forma: Geram-se 20 cadeias paralelas cada um com 2.000 iterações. Desprezam-se as 1.000 primeiras iterações, com objetivo de diminuir o efeito dos pontos iniciais e a partir de daí escolhem-se as iterações de 20 em 20, ou seja, para cada parâmetro em cada cadeia considera-se as iterações 1.020 – ésima, ..., 2.000 – ésima que resulta em uma amostra de Gibbs de tamanho 2.000. Para monitorar a convergência da cadeia utiliza-se a técnica de Gelman e Rubin (1992).

Tabela 3.5: Sumário a posteriori dos parâmetros do modelo de Regressão Weibull-exponenciada com os dados de tempos de vida de isolantes.

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
β_0	-4,6923	-4,6763	0,1815	(-5,0688 ; -4,3644)	1,0161108
β_1	6,0014	6,0017	0,0602	(5,8846 ; 6,1177)	1,000646
δ	0,69322	0,67411	0,16662	(0,4229 ; 1,04387)	1,013394
θ	8,2075	8,2023	0,2254	(7,7734 ; 8,6588)	1,0002416

Na tabela 3.5, apresenta-se o resumo da distribuição a posteriori do modelo de regressão Weibull-exponenciada e na figura 3.4, apresenta-se a densidade aproximada dos parâmetros do modelo de regressão Weibull-exponenciada. Também na tabela 3.5, tem-se o resultado das estimativas dos fatores de redução de escala potencial \hat{R} (veja, Gelman e Rubin, 1992) para todos os parâmetros. Os números de iterações considerados foram suficientes para se chegar à convergência pois $\sqrt{\hat{R}} < 1.1$ para todos os parâmetros.

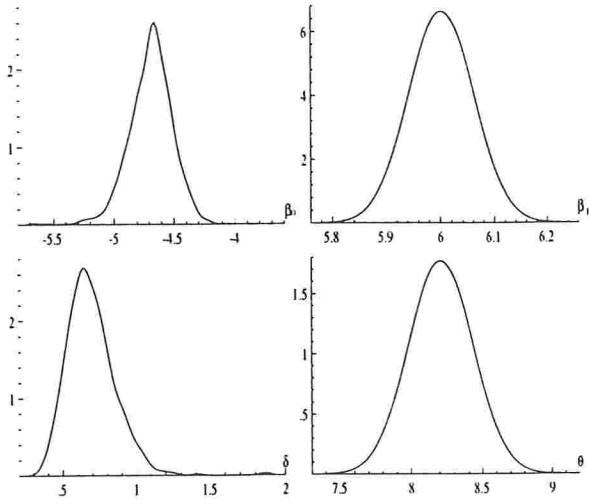


Figura 3.4: Densidades marginais a posteriori aproximados dos parâmetros do modelo de regressão Weibull-exponenciada, β_0 , β_1 , δ and θ com os dados de tempo de vida de isolantes

Análise Bayesiana para dos dados de tempos de vida de isolantes considerando o modelo de regressão Weibull

Para analisar os dados da tabela 3.4 agora, consideramos o modelo de regressão Weibull com as densidades a priori (3.26) para os parâmetros, com, $c_2 = 1,653$, $d_2 = 7,1875$, $\mu_{00} = -3,8415$, $\sigma_{00}^2 = 3$, $\mu_{01} = 5.9$, $\sigma_{01}^2 = 10$.

A partir das densidades condicionais marginais (3.28), geramos 20 cadeias separadas de Gibbs cada uma com 2.000 iterações e utilizamos o método proposto por Gelman Rubin (1992) para verificar a convergência das amostras. Para cada cadeia de Markov associada a distribuição de cada um dos parâmetro tomamos uma amostra com a metade das iterações, considerado as iterações 1.020 – éSIMA, ..., 2.000 – éSIMA, totalizando assim amostras Gibbs de 2.000 pontos amostrais.

As quantidades a posteriori de interesse obtidas a partir das amostras selecionadas são apresentadas na tabela 3.6, onde observamos que as estimativas dos fatores de redução

potencial são menores de 1.1 ($\sqrt{\hat{R}} < 1.1$), indicando a convergência das cadeias associadas as amostras geradas.

Tabela 3.6: Sumário a posteriori dos parâmetros do modelo de regressão Weibull

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
β_0	-3,8504	-3,8501	0,0480	(-3,94370;-3,75628)	1,000324
β_1	5,9612	5,9622	0,0263	(5,90635;6,00985)	1,003526
δ	0,23446	0,23336	0,00042	(0,186117;0,287874)	1,00232

Na figura 3.5, temos as densidades a posteriori marginais aproximadas para o modelo de regressão Weibull considerando os 2000 pontos amostrais.

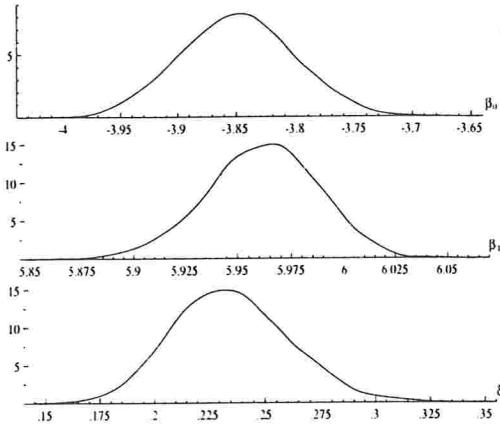


Figura 3.5: Densidades a posteriori aproximados para o modelo de regressão Weibull δ , β_0 e β_1

Usamos o procedimento Bayesiano para verificação da adequabilidade dos modelos de regressão Weibull e Weibull-exponenciada (teste de hipóteses $H_0 : \theta = 1$ (3.16)) para os dados da tabela 3.4, através das densidades preditivas CPO_i , $i = 1, \dots, 40$. Determinou-se a estimativas de Monte Carlo da verossimilhança preditiva com as amostras de Gibbs para os modelos de regressão Weibull e Weibull-exponenciada e essas quantidades são respectivamente $\hat{f}(y|M_0) = 3.617 \times 10^{-6}$ e $\hat{f}(y|M_1) = 2.403 \times 10^{-17}$. O pseudo-fator

de Bayes do modelo de regressão Weibull com respeito ao modelo de regressão Weibull-exponenciada resultou em, $PSFB_{01} = 1.2741 \times 10^{11} > 1$, o que indica que o modelo Weibull é bastante adequado para o ajuste dos dados da tabela 3.4, confirmando-se assim o resultado da inferência clássica. Cabe ressaltar que o suporte do enfoque Bayesiano ao modelo de regressão Weibull é bem maior que o apresentado pelo enfoque clássico.

Como na subseção anterior, é também adotado o método gráfico para verificarmos o melhor modelo entre os modelos estudados. Considerando os dados da tabela 3.4, apresenta-se na figura 3.6, o plot do log da razão dos CPO's dos modelos de regressão Weibull e Weibull-exponenciada com respeito ao numero de observações. Valores positivos do log da razão dos CPO indica a preferência do primeiro modelo. Observa-se, na figura 3.6 que 36 das 40 observações sustenta ao modelo de regressão Weibull sobre o modelo de regressão Weibull-exponenciada. Este método confirma o resultado obtida acima.

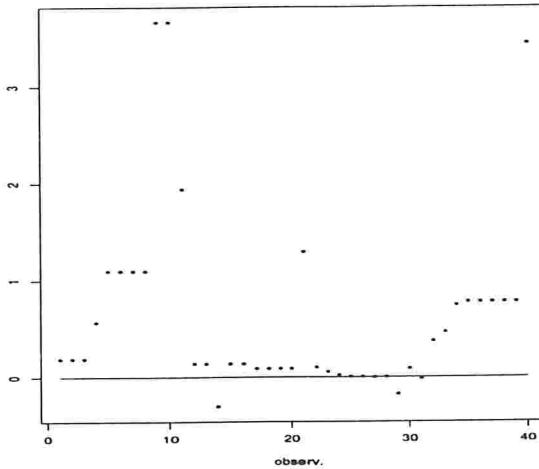


Figura 3.6: Gráfico do log da razão dos CPO's dos modelos de regressão Weibull e Weibull-exponenciada versus as observações

Novamente nota-se que o enfoque Bayesiano apresenta uma indicação muito mais forte do que o enfoque clássico de que o modelo Weibull é o mais adequado.

Na figura 3.7, temos a densidade preditiva para o log do tempo de vida futuro, y , dada as observações D para um valor da covariável $x_h = 190^0 C$ aproximada pelo

método de Monte Carlo com as 2.000 amostras Gibbs dos parâmetros do modelo Weibull-exponenciada.

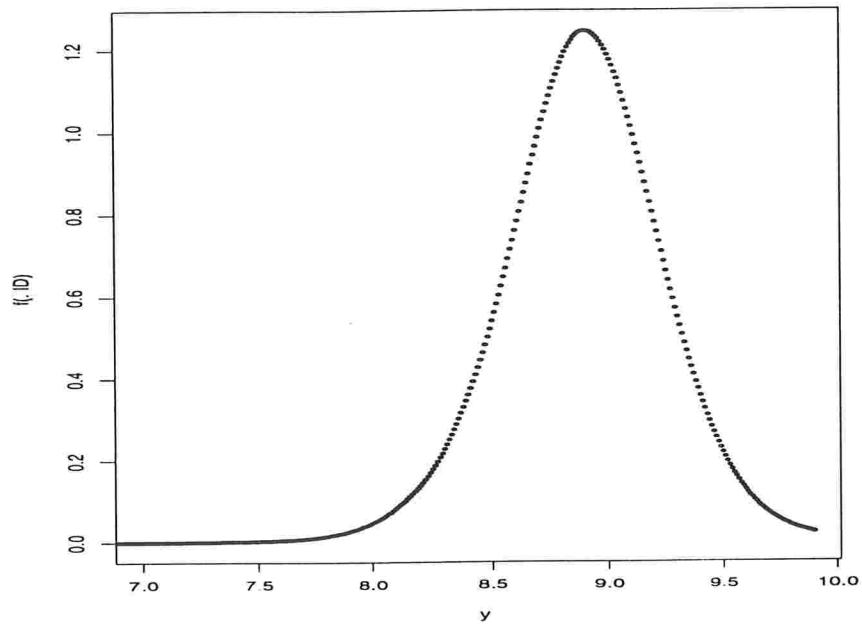


Figura 3.7: Densidade preditiva para o log do tempo de vida futuro no modelo de regressão Weibull-exponenciada

Na figura 3.8, representamos as funções de sobrevivência dada em (3.6) e de risco dada um valor da covariável $x_h = 190^0 C$, aproximadas pelo método de Monte Carlo com as 2.000 amostras Gibbs.

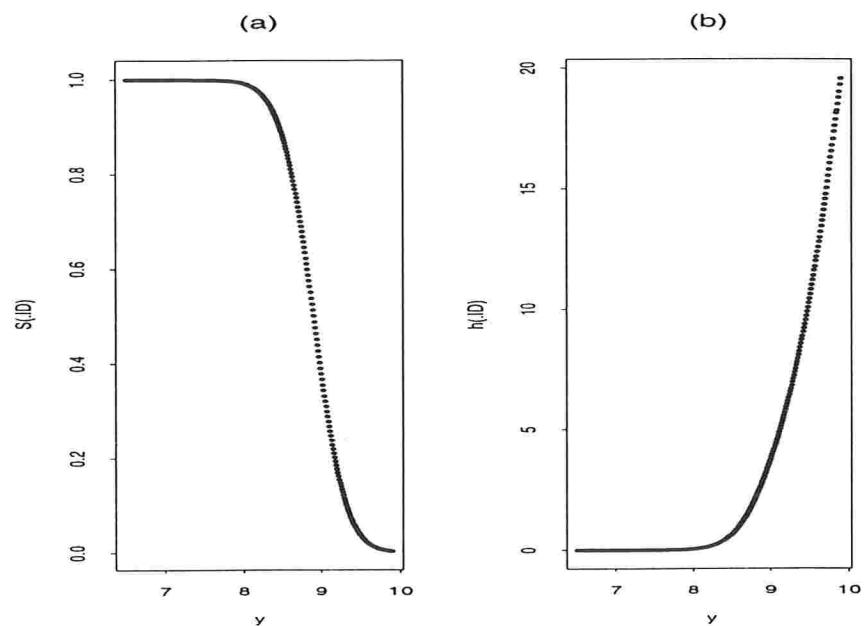


Figura 3.8: Funções de: (a) sobrevivência e (b) risco para o modelo de regressão Weibull-exponenciada dado um valor da covariável, $x_h = 190^0 C$

Capítulo 4

Uso da distribuição Weibull-exponencial para dados com taxas de falha do tipo “bathtub”

4.1 Introdução

Na análise de dados de tempos de vida especialmente no ciclo de vida de um produto pode ocorrer um primeiro período com taxas de falha inicial alta que decresce rapidamente com o tempo (mortalidade infantil) seguida, por um período onde as taxas de falha é aproximadamente constante e uma fase de envelhecimento onde as taxas de falha é crescente devido ao processo de desgaste natural do produto, indicado uma função de taxas de falha do tipo "bathtub". Algumas distribuições foram introduzidas para modelar esse tipo de dados, como a família de distribuições gama generalizadas proposta por Stacy(1962), a família de distribuições F generalizadas de Prentice (1975), a distribuição IDB proposta por Horth (1980), que descreve função de taxas de falha crescente (I), descrecente (D), constante e forma de "bathtub" (B), a qual motivo nome da família. e a família exponencial-potência formulada por Smith e Bain (1975). Uma razoável e comprensiva revisão para

modelos com função taxa de falha do tipo "bathtub" foram dadas por Rajarshi e Rajarshi (1988). O procedimento de inferência clássica padrão pode ter fortes dificuldades para esses modelos, especialmente na presença de dados censurados (veja por exemplo, Rajarshi e Rajarshi (1980); Lawless (1982) e Hjorth(1980)).

Conforme visto nos Capítulos anteriores Mudholkar et al.(1995) apresenta uma generalização simples da família de distribuições Weibull chamada família de distribuição Weibull-exponenciada que além de incluir distribuições com funções de taxas de falha do tipo "bathtub" e unimodal também admite uma ampla classe de distribuições com taxas de falha monótonas. Neste Capítulo comparamos diferentes modelos para análise de dados de tempos de vida com taxas de falha do tipo "bathtub". A distribuição Weibull-exponenciada é comparada com a família de distribuições IDB, exponencial-potência e a mistura de distribuições gama. A abordagem usada para o ajuste dos modelos é a abordagem Bayesiana baseada no método de simulação de Monte Carlo via cadeias de Markov (MCMC). A seleção do modelo é implementada usando o pseudo-fator de Bayes. A implementação de um estudo de simulação e a análise de um conjunto de dados reais parece indicar que o modelo Weibull-exponenciada é um dos que apresentam um melhor ajuste para esse tipo de dados.

4.2 Modelos

Nesta seção apresentamos algumas das famílias de distribuições mais comuns usadas para modelar tempos de vida com função de taxa de falha de tipo "bathtub". Não discutimos a abordagem de inferência clássica

4.2.1 A família de distribuições Weibull-exponenciada

Conforme visto na seção 2.2, a função de densidade de probabilidades para a distribuição Weibull-exponenciada é dada por

$$f(t) = \frac{\alpha\theta}{\sigma} [1 - \exp(-(t/\sigma)^\alpha)]^{\theta-1} \times \exp(-(t/\sigma)^\alpha)(t/\sigma)^{\alpha-1}, \quad 0 < t < \infty, \quad (4.1)$$

onde $\alpha > 0$, $\theta > 0$ são os parâmetros de forma e $\sigma > 0$ é o parâmetro de escala. A função de sobrevivência é dada por

$$S(t) = 1 - [1 - \exp(-(t/\sigma)^\alpha)]^\theta, \quad (4.2)$$

e a função risco de T é dada por

$$h(t) = \frac{\alpha\theta[1 - \exp(-(t/\sigma)^\alpha)]^{\theta-1} \exp(-(t/\sigma)^\alpha)(t/\sigma)^{\alpha-1}}{\sigma[1 - (1 - \exp(-(t/\sigma)^\alpha))^\theta]}. \quad (4.3)$$

A grande flexibilidade deste modelo para ajustar dados de sobrevivência é dado pelas diferentes formas da função risco (4.3) que pode tomar, isto é,

- (i) monótona crescente se $\alpha \geq 1$ e $\alpha\theta \geq 1$,
- (ii) monótona decrescente se $\alpha \leq 1$ e $\alpha\theta \leq 1$,
- (iii) forma de banheira se $\alpha > 1$ e $\alpha\theta < 1$ e
- (iv) unimodal se $\alpha < 1$ e $\alpha\theta > 1$,

4.2.2 A família de distribuições IDB

A função de sobrevivência da distribuição IBD é dada por

$$S(t) = \frac{\exp\{\delta t^2/2\}}{(1 + \beta t)^{\theta/\beta+1}} \quad (4.4)$$

onde $\delta > 0$, $\theta > 0$ e $\beta > 0$.

A função de densidade de probabilidades é

$$f(t) = \frac{(1 + \beta t)\delta t - \theta}{(1 + \beta t)^{\theta/\beta+1}} e^{\delta t^2/2}, \quad 0 < t < \infty. \quad (4.5)$$

A função de taxas de falha é dada por

$$h(t) = \delta t + \frac{\theta}{1 + \beta t}. \quad (4.6)$$

Casos especiais da distribuição IDB são dados por:

- (a) Se $\theta = 0$, temos a distribuição de Rayleigh;
- (b) Se $\delta = \beta = 0$, temos a distribuição exponencial;
- (c) Se $\delta = 0$, temos distribuições com taxas de falha decrescente;
- (d) Se $\delta \geq \theta\beta$ distribuições com taxas de falha crescente;
- (e) Se $0 < \delta < \theta\beta$ distribuições com taxas de falha do tipo "bathtub".

4.2.3 A família de distribuições exponencial-potência

Smith e Bain (1975,1976) apresentaram a família de distribuição exponencial-potência com dois parâmetros como uma distribuição alternativa para dados de tempos de vida com taxas de falha do tipo "bathtub". A função de densidade de probabilidade é

$$f(t) = \beta\alpha^{-\beta}t^{\beta-1} \exp\left(1 - e^{(\frac{t}{\alpha})^\beta} + \left(\frac{t}{\alpha}\right)^\beta\right), \quad 0 < t < \infty, \quad (4.7)$$

onde $\alpha > 0$ e $\beta > 0$;

A função de distribuição é dada por

$$F(t) = 1 - \exp\left(1 - e^{(\frac{t}{\alpha})^\beta}\right), \quad (4.8)$$

e a função risco é dada por

$$h(t) = \frac{f(t)}{1 - F(t)} = \alpha^{-\beta}\beta t^{\beta-1} \exp\left(\left(\frac{t}{\alpha}\right)^\beta\right) \quad (4.9)$$

A função risco (4.9) tem forma de banheira se $0 < \beta < 1$.

4.2.4 Misturas de distribuições Gama

Glasser (1980) apresentaram a mistura da distribuições gama com parâmetro de escala comum para os de tempos de vida com função de risco do tipo "bathtub". Neste modelo a função de densidade é dada por

$$f(t; \alpha_1, \alpha_2, \beta, p) = pf_1(t; \alpha_1, \beta) + (1 - p)f_2(t; \alpha_2, \beta) \quad (4.10)$$

onde $0 < p < 1$

$$f_j(t; \alpha_j, \beta) = \frac{\beta^{\alpha_j}}{\Gamma(\alpha_j)} t^{\alpha_j-1} e^{\beta t}, \quad \alpha_j > 0, \quad \beta > 0, \quad j = 1, 2.$$

A função de sobrevivência é dada por

$$S(t; \alpha_1, \alpha_2, \beta, p) = pS_1(t; \alpha_1, \beta) + (1 - p)S_2(t; \alpha_2, \beta), \quad t > 0.$$

onde $S_j(t; \alpha_j, \beta)$ é a função de sobrevivência do j -ésimo componente.

A função de taxas de falha é de tipo "bathtub" se:

- (a) $\alpha_1 > 1, \alpha_2 = 1$; f_1 estritamente crescente e f_2 uma exponencial
- (b) $\alpha_1 > 1, \alpha_2 < 1$.

4.3 Inferência Bayesiana

Assumimos que os tempos de vida são distribuídos independentemente e independentes do mecanismo de censura. Considerando que os dados de tempo de vida são censurados à direita, e que observamos $T_i = \min(T_i^0, C_i)$, onde T_i^0 é o tempo de vida do i -ésimo indivíduo e C_i é o tempo de censura do i -ésima indivíduo, $i = 1, \dots, n$. Neste caso a função de verossimilhança de θ , um vetor de parâmetros de dimensão p é dado por

$$L(\theta) = \prod_{i \in F} f(t_i; \theta) \prod_{i \in C} S(t_i; \theta), \quad (4.11)$$

onde F denota o conjunto de observações não censuradas, C denota o conjunto de observações censuradas, $f(\cdot)$ e $S(\cdot)$ é a função de densidade de probabilidades e a função de sobrevida, respectivamente.

Considerando uma densidade a priori $\pi(\theta)$ a densidade a posteriori de θ é dada por

$$\pi(\theta|D) = \frac{L(\theta)\pi(\theta)}{\int L(\theta)\pi(\theta)d\theta} \quad (4.12)$$

onde D denota o conjunto de dados observados.

Conforme visto nos capítulos anteriores, na abordagem Bayesiana, inferências são tipicamente baseadas nas densidades a posteriori marginais dos parâmetros envolvidos em (4.12), para os modelos descritos na seção anterior essas densidades marginais parecem não podermos ser obtidas analiticamente. Portanto, para esses casos, métodos de aproximação ou métodos de simulação de Monte Carlo via Cadeia de Markov (MCMC) são necessários. Neste Capítulo optamos pelo segundo método.

4.3.1 O Modelo Weibull-exponenciada

Considerando que os dados observados são não censurados, a função de verossimilhança (4.11) é dada por,

$$L(\alpha, \theta, \sigma) \propto \alpha^n \theta^n \sigma^{-n\alpha} \exp \left\{ - \sum_{i=1}^n \left(\frac{t_i}{\sigma} \right)^\alpha \right\} \prod_{i=1}^n t_i^{\alpha-1} (1 - \exp\{-(\frac{t_i}{\sigma})^\alpha\})^{\theta-1}. \quad (4.13)$$

Para análise Bayesiana, assumimos a seguinte densidade a priori para α , θ e σ :

$$\begin{aligned} \alpha &\sim \Gamma(a_1, b_1), \text{ com } a_1 \text{ e } b_1 \text{ conhecidos;} \\ \theta &\sim \Gamma(a_2, b_2), \text{ com } a_2 \text{ e } b_2 \text{ conhecidos;} \\ \sigma &\sim \Gamma(a_3, b_3), \text{ com } a_3 \text{ e } b_3 \text{ conhecidos.} \end{aligned} \quad (4.14)$$

onde $\Gamma(a_i, b_i)$ denota a distribuição gama com média $\frac{a_i}{b_i}$ e variância $\frac{a_i}{b_i^2}$. Além disso assumimos independência entre os parâmetros.

Combinando (4.13) e (4.14) e o teorema de Bayes pode-se mostrar que a densidade conjunta a posteriori de α , θ e σ é dada por,

$$\begin{aligned}\pi(\alpha, \theta, \sigma | D) &\propto \alpha^{n+a_1-1} \theta^{n+a_2-1} \sigma^{-n\alpha+a_3-1} \exp \left\{ -\sum_{i=1}^n \left(\frac{t_i}{\sigma} \right)^\alpha - b_1 \alpha - b_2 \theta - b_3 \sigma \right\} \\ &\quad \prod_{i=1}^n t_i^{\alpha-1} (1 - \exp\{-(\frac{t_i}{\sigma})^\alpha\})^{\theta-1},\end{aligned}\tag{4.15}$$

onde $\alpha > 0$, $\theta > 0$, $\sigma > 0$ e D denota o conjunto de dados observados sobre os tempos de falha.

Para obtermos uma amostra a posteriori das densidades marginais α , θ e σ , fazemos uso do algoritmo Gibbs Sampling que baseia-se em sucessivas gerações das distribuições condicionais a posteriori de $\pi(\alpha|\theta, \sigma, D)$, $\pi(\theta|\alpha, \sigma, D)$ e $\pi(\sigma|\alpha, \theta, D)$.

Da densidade a posteriori conjunta dada em (4.15) pode-se mostrar que as densidades condicionais a posteriori para o algoritmo de Gibbs são dadas por

$$\begin{aligned}\pi(\alpha|\theta, \sigma, D) &\propto \alpha^{n+a_1-1} \exp \left\{ -\sum_{i=1}^n \left(\frac{t_i}{\sigma} \right)^\alpha - b_1 \alpha \right\} \prod_{i=1}^n t_i^{\alpha-1} (1 - \exp\{-(\frac{t_i}{\sigma})^\alpha\})^{\theta-1}, \\ \pi(\theta|\alpha, \sigma, D) &\propto \theta^{n+a_2-1} e^{-b_2 \theta} \prod_{i=1}^n t_i^{\alpha-1} (1 - \exp\{-(\frac{t_i}{\sigma})^\alpha\})^{\theta-1}, \\ \pi(\sigma|\alpha, \theta, D) &\propto \theta^{-n\alpha+a_3-1} \exp \left\{ -\sum_{i=1}^n \left(\frac{t_i}{\sigma} \right)^\alpha - b_3 \sigma \right\} \prod_{i=1}^n t_i^{\alpha-1} (1 - \exp\{-(\frac{t_i}{\sigma})^\alpha\})^{\theta-1}.\end{aligned}\tag{4.16}$$

Observe que as densidades dadas em (4.16) não são conhecidas de modo que faremos uso do algoritmo de Metropolis-Hastings para gerar as variáveis α , θ e σ .

4.3.2 O modelo IDB

A função de verossimilhança de δ , β and θ é dada por,

$$L_n(\delta, \beta, \theta) = \exp \left\{ -\frac{\delta \sum_{i=1}^n t_i^2}{2} \right\} \prod_{i=1}^n \frac{(1 + \beta t_i) \delta t_i + \theta}{(1 + \beta t_i)^{\theta/\beta+1}}.\tag{4.17}$$

Assumindo independência entre os parâmetros, consideraremos a seguintes densidades a priori para δ , β e θ :

$$\delta \sim \Gamma(c_1, d_1); \text{ com } c_1 \text{ e } d_1 \text{ conhecidos};$$

$$\begin{aligned}\beta &\sim \Gamma(c_2, d_2); \text{ com } c_2 \text{ e } d_2 \text{ conhecidos;} \\ \theta &\sim \Gamma(c_3, d_3); \text{ com } c_3 \text{ e } d_3 \text{ conhecidos.}\end{aligned}\quad (4.18)$$

Considerando (4.17)-(4.18) e o teorema de Bayes pode-se mostrar que a densidade conjunta a posteriori de δ , β and θ é dada por,

$$\pi(\alpha, \theta, \sigma | D) \propto \delta^{c_1-1} \beta^{c_2-1} \theta^{c_3-1} \exp \left\{ -d_1 \delta - \frac{\delta \sum_{i=1}^n t_i^2}{2} - d_2 \beta - d_3 \theta \right\} \Psi_n(\delta, \beta, \theta), \quad (4.19)$$

$$\text{onde } \Psi_n(\delta, \beta, \theta) = \prod_{i=1}^n \frac{(1+\beta t_i)^{\delta t_i + \theta}}{(1+\beta t_i)^{\theta/\beta+1}}.$$

De (4.19) pode-se mostrar que as densidades condicionais requeridas para o algoritmo de Gibbs são dadas por

$$\begin{aligned}\pi(\delta | \beta, \theta, D) &\propto \delta^{c_1-1} \exp \left\{ -d_1 \delta - \frac{\delta \sum_{i=1}^n t_i^2}{2} \right\} \Psi_n(\delta, \beta, \theta), \\ \pi(\beta | \delta, \theta, D) &\propto \beta^{c_2-1} e^{-d_2 \beta} \Psi_n(\delta, \beta, \theta), \\ \pi(\theta | \delta, \beta, D) &\propto \theta^{c_3-1} e^{-d_3 \theta} \Psi_n(\delta, \beta, \theta).\end{aligned}\quad (4.20)$$

Observamos que as densidades condicionais dadas em (4.20) não são padrão, e portanto usamos o algoritmo de Metropolis-Hastings para gerar as variáveis δ , β e θ dentro do ciclo do amostrador de Gibbs.

4.3.3 O modelo exponencial-potência

De (4.7), a função de verossimilhança para α e β é dada por

$$L(\alpha, \beta) = \beta^n \alpha^{-n\beta} \prod_{i=1}^n t_i^{\beta-1} \exp \left\{ n + \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta - \sum_{i=1}^n e^{\left(\frac{t_i}{\alpha} \right)^\beta} \right\}. \quad (4.21)$$

Assumindo a densidades a priori,

$$\begin{aligned}\alpha &\sim \Gamma(e_1, f_1) \text{ com } e_1 \text{ e } f_1 \text{ conhecidos,} \\ \beta &\sim \Gamma(e_2, f_2) \text{ com } e_2 \text{ e } f_2 \text{ conhecidos.}\end{aligned}\quad (4.22)$$

De (4.21)-(4.22) e do teorema de Bayes pode-se mostrar que a densidade a posteriori de α e β é dada por,

$$\pi(\alpha, \beta | D) \propto \beta^{n+\epsilon_2-1} \alpha^{-n\beta+\epsilon_1-1} \prod_{i=1}^n t_i^{\beta-1} \exp \left\{ \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta - \sum_{i=1}^n e^{\left(\frac{t_i}{\alpha} \right)^\beta} - f_1 \alpha - f_2 \beta \right\}, \quad (4.23)$$

onde $\alpha > 0$ e $\beta > 0$.

De (4.23), obtemos a densidade condicional a posteriori para o algoritmo de Gibbs que é dada por

$$\begin{aligned} \pi(\alpha | \beta, D) &\propto \alpha^{-n\beta+\epsilon_1-1} \exp \left\{ \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta - \sum_{i=1}^n e^{\left(\frac{t_i}{\alpha} \right)^\beta} - f_1 \alpha \right\}, \\ \pi(\beta | \alpha, D) &\propto \beta^{n+\epsilon_2-1} \prod_{i=1}^n t_i^{\beta-1} \exp \left\{ \sum_{i=1}^n \left(\frac{t_i}{\alpha} \right)^\beta - \sum_{i=1}^n e^{\left(\frac{t_i}{\alpha} \right)^\beta} - f_2 \beta \right\}. \end{aligned} \quad (4.24)$$

De (4.24) observamos que as densidades condicionais a posteriori não são conhecidas pelo que faremos uso do algoritmo de Metropolis-Hastings para gerar as variáveis α e β dentro do ciclo do algoritmo de Gibbs.

4.3.4 Modelo de mistura distribuições de Gamas

A função de verossimilhança para mistura de distribuição de gamas em (4.10) (veja, Dempster et al. 1977) é dada por,

$$L(\alpha_1, \alpha_2, \beta, p) = \prod_{i=1}^n [pf(t_i; \alpha_1, \beta)]^{z_{i1}} [(1-p)f(t_i; \alpha_2, \beta)]^{1-z_{i1}}, \quad (4.25)$$

onde z_{i1} é uma função indicadora ($z_{i1} = 1(0)$ se t_i pertence à primeira (segunda) componente do modelo de mistura).

Assumindo as seguintes densidades a priori independentes:

$$\begin{aligned} p &\sim \text{Beta}(\gamma_1, \gamma_2), \text{ com } \gamma_1 \text{ e } \gamma_2 \text{ conhecidos,} \\ \alpha_j &\sim \Gamma(1, \lambda_j), \text{ com } \lambda_j \text{ conhecidos } j = 1, 2, \\ \beta &\sim \Gamma(\beta_1, \beta_2), \text{ } \beta_1 \text{ e } \beta_2 \text{ conhecidos,} \end{aligned} \quad (4.26)$$

onde $\text{Beta}(\delta_1, \delta_2)$ denota uma distribuição Beta com parâmetros δ_1 and δ_2 . Assim, a densidade conjunta a posteriori $\alpha_1, \alpha_2, \beta$ e p é dada por,

$$\begin{aligned}\pi(\alpha_1, \alpha_2, \beta, P|D, \underline{z}) &\propto p^{z_{.1} + \gamma_1 - 1} (1-p)^{n-z_{.1} + \gamma_2 - 1} \beta^{\alpha_1 z_{.1} + \alpha_2(n-z_{.1}) + \beta_1 - 1} \\ &\times \exp \left\{ -\beta \left(\sum_{i=1}^n t_i + \beta_2 \right) - \lambda_1 \alpha_1 - \lambda_2 \alpha_2 \right\} \frac{\left(\prod_{i=1}^n t_i^{z_{i1}} \right)^{\alpha_1}}{[\Gamma(\alpha_1)]^{z_{.1}}} \\ &\times \frac{\left(\prod_{i=1}^n t_i^{1-z_{i1}} \right)^{\alpha_2}}{[\Gamma(\alpha_1)]^{n-z_{.1}}},\end{aligned}\quad (4.27)$$

onde $z_{.1} = \sum_{i=1}^n z_{i1}$.

As densidades condicionais a posteriori para o algoritmo de Gibbs são dadas por:

$$\begin{aligned}\pi(z_{i1}|\alpha_1, \alpha_2, \beta, D) &\sim \text{beroulli} \left(\frac{pf(t_i; \alpha_1, \beta)}{pf(t_i; \alpha_1, \beta) + (1-p)f(t_i; \alpha_2, \beta)} \right); \\ \pi(p|\alpha_1, \alpha_2, \beta, \underline{z}, D) &\sim \text{Beta}(z_{.1} + \gamma_1, n - z_{.1} + \gamma_2), \\ \pi(\beta|\alpha_1, \alpha_2, p, \underline{z}, D) &\sim \Gamma \left(\alpha_1 z_{.1} + \alpha_2(n - z_{.1}) + \beta_1, \beta_2 + \sum_{i=1}^n t_i \right); \\ \pi(\alpha_1|\alpha_2, \beta, p, \underline{z}, D) &\propto \exp \{-\lambda_1 \alpha_1\} \frac{\beta^{\alpha_1 z_{.1}} \left(\prod_{i=1}^n t_i^{z_{i1}} \right)^{\alpha_1}}{[\Gamma(\alpha_1)]^{z_{.1}}}; \\ \pi(\alpha_2|\alpha_1, \beta, p, \underline{z}, D) &\propto \exp \{-\lambda_2 \alpha_2\} \frac{\beta^{\alpha_2(n-z_{.1})} \left(\prod_{i=1}^n t_i^{1-z_{i1}} \right)^{\alpha_2}}{[\Gamma(\alpha_2)]^{n-z_{.1}}}.\end{aligned}\quad (4.28)$$

Note que as duas últimas densidades condicionais requer do algoritmo Metropolis-Hasting no passo da implementação da metodologia MCMC.

4.4 Discriminação de modelos

Na seção 4.2 foram apresentados vários modelos para tempos de vida. Surge então a necessidade de determinar qual ou quais modelos são mais apropriadas para os tempos de vida com função de taxas de falha do tipo “bathtub”.

Os procedimentos descritos no capítulo 1, são adotados para verificarmos qual é o melhor modelo dentre os modelos estudados. Por exemplo, para o modelo Weibull-exponenciada a estimativa de Monte Carlo da densidade preditiva de t_r dado $D_{(r)}$ (CPO) com as amostras Gibbs é dada por

$$\hat{f}(t_r | D_{(r)})^{-1} = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S \frac{(\sigma^{(r,s)})^{\alpha^{(r,s)}}}{\alpha^{(r,s)} \theta^{(r,s)} [1 - \exp(-(\frac{t_r}{\sigma^{(r,s)}})^{\alpha^{(r,s)}})]^{\theta^{(r,s)}-1} \exp(-(\frac{t_r}{\sigma^{(r,s)}})^{\alpha^{(r,s)}})(t_r)^{\alpha^{(r,s)}-1}}. \quad (4.29)$$

sendo $D_{(r)} = (t_1, \dots, t_{r-1}, t_{r+1}, \dots, t_n)$, $r = 1, \dots, n$, $\alpha^{(r,s)}$, $\theta^{(r,s)}$ e $\sigma^{(r,s)}$ são gerados de (4.15) considerando S iterações para cada uma das R cadeias considerando diferentes pontos iniciais para α , θ e σ .

Da mesma forma, pode-se obter as estimativas dos CPO's para os modelos IDB, exponencial-potência e mistura de distribuições gammas.

Usaremos a estimativa do pseudo-fator de Bayes dada em (1.11), computada com as estimativas dos CPO's para discriminação entre os modelos considerados neste Capítulo.

4.5 Alguns exemplos

4.5.1 Exemplo 1

Para ilustrar a metodologia apresentada nas seções anteriores consideramos o conjunto de dados ilustradas em Arset (1987). Os dados descrevem os tempos de vida de 50 projetos industriais.

Modelo Weibull-exponenciada

Para análise dos dados da tabela (4.1), primeiramente assume-se a distribuição Weibull-exponenciada (4.1). Considerando a densidade a priori α , θ e σ dada em (4.14) com

Tabela 4.1: Tempos de sobrevivência de 50 projetos

0.1	0.2	1	1	1	1	1	2	3	6	7	11	12	18	18	18	18	21
32	36	40	45	46	47	50	55	60	63	63	67	67	67	72	75	79	82
82	83	84	84	84	85	85	85	85	85	86	86						

$a_1 = 8,6$, $b_1 = 1,83$, $a_2 = 7,58$, $b_2 = 51,98$. $a_3 = 26491,6$ and $b_3 = 290,251$ (a escolha desses valores para os parâmetros da densidade a priori, foram baseadas numa opinião experiente combinada com a análise preliminar dos dados), geramos 10 cadeias cada uma com 5.000 iterações e monitoramos a convergência das amostras Gibbs usando o método de Gelman e Rubin (1992) que utiliza a técnica de análise de variância para determinar se mais iterações são necessárias. Para cada parâmetro e em cada cadeia consideramos as iterações 2.525 – ésima, 2.550 – ésima, ..., 5.000 – ésima a qual produz uma amostra de Gibbs de tamanho 1.000. Na tabela 4.2, reportamos o resumo a posteriori dos parâmetros do modelo e, na figura 4.1, temos as densidades marginais a posteriori aproximadas obtidas considerando as 1.000 amostras de Gibbs. Também temos na tabela 4.2, o resultado da estimativa dos fatores de redução de escala potencial \hat{R} (veja, Gelmann e Rubin, 1992) para todos os parâmetros. Neste caso o número de iterações foi suficiente para a convergência ($\sqrt{\hat{R}} < 1.1$), para todos os parâmetros).

Tabela 4.2: Sumário a posteriori para o modelo Weibull-exponenciada

	Média	Mediana	D.P.	95% intervalo de credibilidade	\hat{R}
α	2,6444	2,6446	0,27758	(2,081139;3,16199)	1,002173
θ	0,20422	0,20432	0,05406	(0,10314;0,317079)	1,002107
σ	91,016	91,005	0,547375	(89,9555;92,0941)	1,0028216

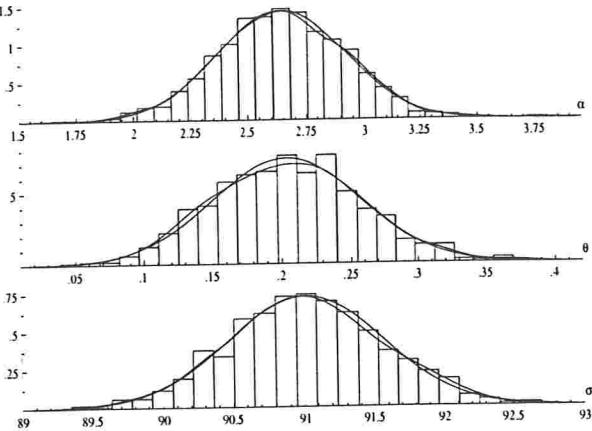


Figura 4.1: Densidades marginais a posteriori aproximadas de α , θ e σ parâmetros do modelo Weibull-exponenciada

Modelo IDB

Considerando o modelo IBD (4.4) e a densidade a priori dada em (4.19) com $c_1 = 33, 30$, $d_1 = 73669, 89$, $c_2 = 0, 8775$, $d_2 = 0, 2944$, $c_3 = 1, 4832$ and $d_3 = 7, 315$, geramos 20 cadeias separadas cada uma com 2.000 iterações e consideramos para cada parâmetro as iterações 1.020 – éssima, ..., 2.000 – éssima, de modo que as 20 cadeias fornecem uma amostra de tamanho 1.000. Na tabela 4.3, obtemos o resumo a posteriori dos parâmetros do modelo e, na figura 4.2, temos as densidades marginais a posteriori aproximados para δ , β e θ considerando as $S = 1.000$ amostras Gibbs. Também temos na tabela 4.3, a estimativa dos fatores de redução de escala potencial \hat{R} para todos os parâmetros, a qual indica que as cadeias convergiram.

Modelo exponencial-potência

Consideramos agora, o modelo exponencial-potência (4.7) e a densidade a priori (4.24) com $e_1 = 113.33$, $f_1 = 1.07$, $e_2 = 28.58$ e $f_2 = 53.4$, geramos 8 cadeias separadas

Tabela 4.3: Sumário a posteriori dos parâmetros do modelo IDB

	Média	Mediana	D.P.	95%Intervalo de credibilidade	\hat{R}
δ	0,00030	0,00030	0,00005	(0,000229;0,000407)	1,006153
β	0,72470	0,5024	0,8999	(0,0271;2.527)	1,0000402
θ	0,18837	0,16352	0,12462	(0,0247;0,48762)	1,0002546

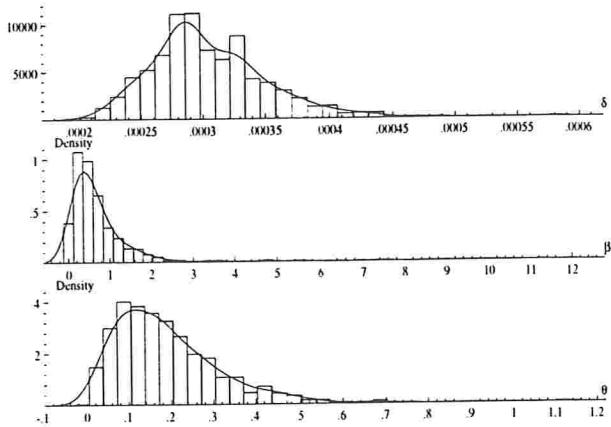


Figura 4.2: Densidades marginais a posteriori aproximadas de δ , β e θ parâmetros do modelo IDB

cada um com 50.000 iterações, foram consideradas as iterações 25.025 – ésima, 25.050 – ésima, . . . , 50.000 – ésima que produz uma amostra Gibbs de tamanho 8.000. Na tabela 4.4, obtemos o sumário a posteriori dos parâmetros do modelo e, na figura 4.3, temos a densidades marginais a posterioris aproximados considerando os $S = 8.000$ amostras Gibbs. Para todos os parâmetros, observamos (veja tabela 4.4), $\sqrt{\hat{R}} < 1,1$, indicando a convergência.

Tabela 4.4: Sumário a posteriori dos parâmetros do modelo exponencial-potência

	Média	Mediana	D.P.	95% Intervalo de credibilidade	\hat{R}
α	76,498	76,4919	2,991	(68,8193;81,3473)	1,000109
β	2,166	2,166	0,315	(1,18404;2,64272)	1,009575

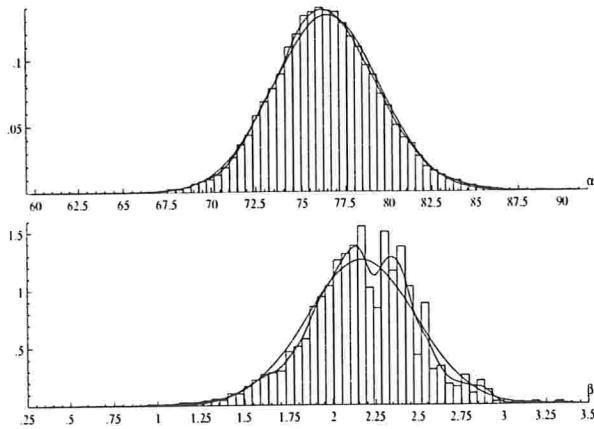


Figura 4.3: Densidades marginais a posteriori aproximadas de α e β parâmetros do modelo exponencial-potência

Modelo de Mistura

Finalmente considerando, a mistura de gamas dado em (4.10), e a densidade a priori (4.26) com $\gamma_1 = 0,5$, $\gamma = 0,5$, $\lambda_j = 0,000001$ ($j = 1, 2$), $\beta_1 = 1$ e $\beta_2 = 0,00002$, geramos 20 cadeias separadas de Gibbs cada uma com 5.000 iterações. Para cada parâmetro em cada cadeia consideramos as iterações 2.550 – ésima, 2.600 – ésima, ..., 5.000 – ésima produzindo uma amostra Gibbs de tamanho 1.000. Na tabela 4.5, reportamos as quantidades de interesse, média, mediana, desvio padrão (D.P.), intervalos de credibilidade de 95% e as estimativas dos fatores de redução de escala potêncial \hat{R} para cada um dos parâmetros.

Tabela 4.5: Sumário a posteriori dos parâmetros do modelo de mistura de distribuições gama

	Média	Mediana	D.P.	95% Intervalo de credibilidade	\hat{R}
α_1	14,542	12,494	10,405	(1,33710;39,6366)	1,000312
α_2	0,91499	0,89006	1,19612	(0,68565;1,32152)	1,000304
β	0,02148	0,02030	0,00655	(0,01271;0,04058)	1,085947
p	0,02672	0,00601	0,09164	(0,000012;0,0318169)	1,021205

Observamos na tabela 4.5, que $\sqrt{\hat{R}} < 1,1$ para todos os parâmetros o que nos indica a convergência das amostras geradas. Na figura 4.4, representamos a densidades marginais a posteriori aproximadas com as 1.000 amostras Gibbs para os parâmetros do modelo de mistura de distribuições gama.

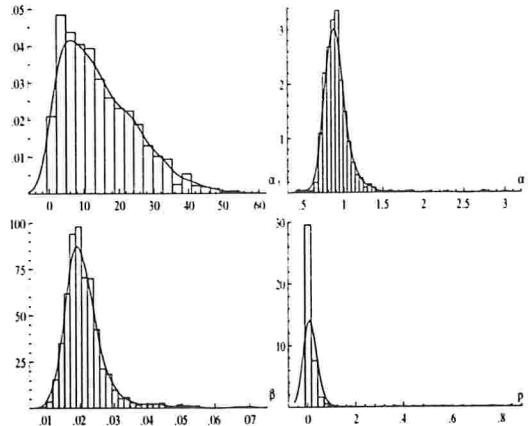


Figura 4.4: Densidades marginais a posteriori aproximados de α_1 , α_2 , β e p parâmetros do modelo de mistura de distribuições gama

Descriminação dos modelos

Para selecionar entre os modelos Weibull-exponenciada (4.1), IBD dada em (4.5), mistura de distribuições gama dada em (4.10) e exponencial-potência dada em (4.7), se ajusta melhor aos dados da tabela 4.1, utilizamos a técnica das densidades preditivas condicionais ordenadas (CPO) para o computo do pseudo-fator Bayes e o método gráfico de seleção de modelos (veja capítulo 1).

Com as amostras Gibbs geradas, temos na tabela 4.6, o pseudo-fator de Bayes (1.11) para a seleção do melhor modelo para os dados de tempo de vida dados na tabela 4.1.

Na figura 4.5, temos o gráfico da razão dos CPO para cada modelo com respeito aos outros versus as observações. Valores da razão de CPO maior que *um* indicam a preferência do primeiro modelo. Por exemplo, na figura 4.5, o “plot” do modelo Weibull-exponenciada versus o modelo exponencial-pontência indica que 28 das 50 observações sustentam o modelo Weibull-exponenciada sobre o modelo exponencial-potência. Similarmente, 38 observações sustentam o modelo Weibull-exponenciada sobre o modelo IDB, 33 observações sustentam ao modelo Weibull-exponenciada sobre o modelo de mistura de distribuições gamas, 38 observações sustentam o modelo IDB sobre o modelo exponencial-potência, 32 observações sustentam o modelo exponencial-potência sobre o modelo de mistura de gamas, e 39 observações sustentam ao modelo de mistura de gamas sobre o modelo IDB. Assim, em conclusão, esse critério e o método do pseudo-fator de Bayes mostram que o modelo Weibull-exponenciada é um dos que se ajustam melhor aos dados da tabela 4.1.

Tabela 4.6: Pseudo-Fator de Bayes

Modelos	PSFB
Weibull-exponenciada vs exponencial-potência	32,83
Weibull-exponenciadal vs IDB	17,61
Weibull-exponenciada vs mistura de gammas	620,75
exponencial-potência vs IDB	0,54
exponencial-potência vs mistura gamas	18,91
IDB vs mistura de gamas	35,26

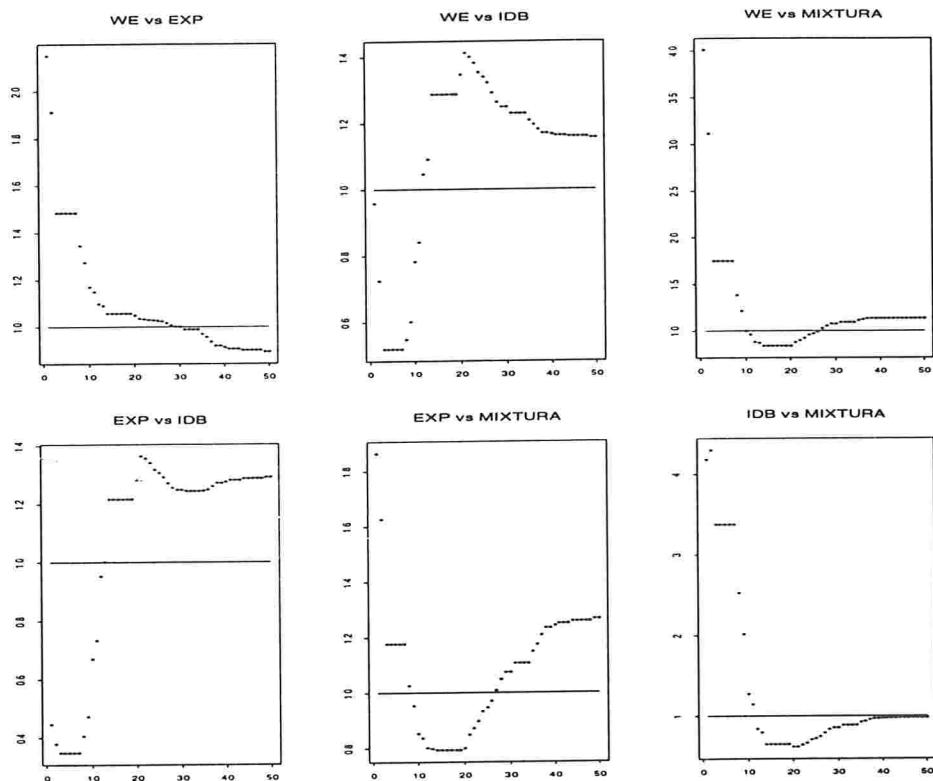


Figura 4.5: Gráfico da razão dos CPO para os diferentes modelos versus observações,

4.5.2 Exemplo 2

Na tabela 4.7, temos 50 observações geradas da densidade (5.1) com parâmetros $\alpha = 2$, $\theta = 0,3$ e $\sigma = 1$, não incluindo dados censurados.

Tabela 4.7: Dados gerados da densidade (4.1)

0,17087	0,67164	0,044448	1,0409	0,14450	0,44826	0,77785	0,0048173	0,0032595	0,27430
0,0015956	0,48105	0,66123	1,8675	0,22152	0,63347	0,012058	0,41785	1,9351	0,57678
0,67180	0,15366	1,0742	0,46588	0,038705	0,72372	0,17963	1,4175	0,48098	0,36706
0,64136	0,13995	1,4367	0,34256	0,29463	0,27305	0,060425	0,15204	1,0947	0,53263
1,2511	0,43501	0,30608	0,40391	0,025192	0,39053	0,42179	1,0422	0,0048170	1,3440

Modelo Weibull-exponenciada

Considerando o modelo Weibull-exponenciada (4.1), a a densidade a priori (4.14) com $a_1 = 0,4$, $b_1 = 0,2$, $a_2 = 0,6$, $b_2 = 0,18$, $a_3 = 0,001$ $b_3 = 0,001$, geramos 20 cadeias separadas cada uma com 2.000 iterações e, monitoramos a convergência usando o método de Gelman e Rubin (1992). Para cada parâmetro consideramos as 1.010 –ésima, 1.020 –ésima, ..., 2.000 –ésima as 20 cadeias produz um amostra de Gibbs de tamanho 2.000. Na tabela 4.8, reportamos a sumário a posterior dos parâmetros e, na figura 4.6 temos a densidade marginal a posteriori aproximado considerando as 2.000 amostras Gibbs. Também na tabela 4.8, temos a estimativa dos fatores de redução de escala \hat{R} (veja Gelman e Rubin, 1992) para todos prâmetros. Neste caso, o número de iterações consideradas foram suficientes para atingir a convergência uma vez que ($\sqrt{\hat{R}} < 1.1$) para todos parâmetros.

Tabela 4.8: Sumário a posteriori dos parâmetros do modelo Weibull-exponenciada

	Média	Mediana	D.P.	95% Intervalo de Credibilidade	\hat{R}
α	1,7279	1,7054	0,3570	(1,0938 ; 2,4497)	1,000382
θ	0,4098	0,3996	0,12686	(0,1822 ; 0,6866)	1,001075
σ	0,9944	0,9889	0,1753	(0,6783; 1,3637)	1,001018

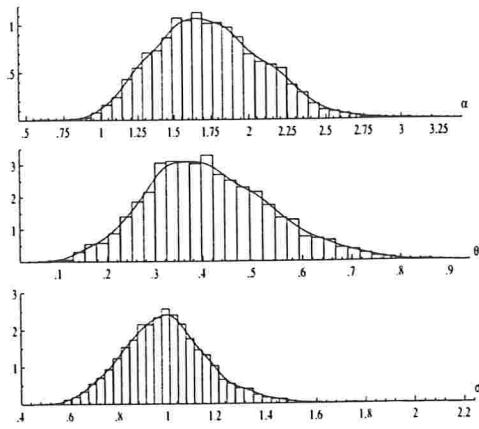


Figura 4.6: Densidades marginais a posteriori aproximado de α , θ e σ parâmetros do modelo Weibull-exponenciada

Modelo IBD

Considerando o modelo IBD (4.4) e a densidade a priori (4.19) com $c_1 = 7,76$, $d_1 = 3,94$, $c_2 = 34,68$, $d_2 = 2,63$, $c_3 = 6,75$ e $d_3 = 1,5$, geramos 20 cadeias separadas de Gibbs cada um com 2.000 iterações. Para cada parâmetro consideramos as iterações 1.010 – éssima, 1.020 – éssima, ..., 2.000 – éssima, totalizando assim uma amostra Gibbs de tamanho 2.000. Na tabela 4.9, apresentamos o sumário a posteriori para os parâmetros do modelo e, na figura 4.7, temos as densidades marginais a posteriori aproximadas considerando as 2.000 amostras Gibbs. Também temos na tabela 4.9, as estimativas dos fatores de redução de escala potencial \hat{R} para todos os parâmetros, o qual indica a convergência das cadeias.

Modelo exponencial-potência

Considerando agora, o modelo exponencial-potência (4.7) a densidade a priori (4.24) com $e_1 = 4,58$, $f_1 = 4,78$, $e_2 = 4,161$ e $f_2 = 6,451$, geramos 20 cadeias separadas de Gibbs cada uma com 5.000 iterações e consideramos as iterações 2.525 – éssima, 2.550 –

Tabela 4.9: Sumário a posteriori dos parâmetros do modelo IDB

	Média	Mediana	D.P.	95% Intervalo de Credibilidade	\hat{R}
δ	2,0098	1,9954	0,4110	(1,24363;2,8898)	1,000107
β	13,283	13,197	2,247	(9,3060 ; 18,2471)	1,0000402
θ	3,8369	3,7369	0,9027	(2,3219 ; 5,8996)	1,007290

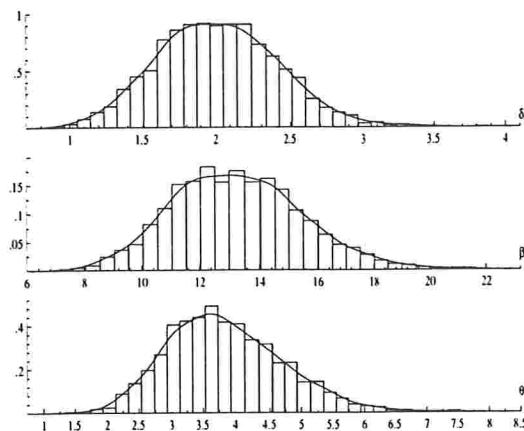


Figura 4.7: Densidades marginais a posteriori aproximada de δ , β e θ parâmetros do modelo IDB

ésima, . . . , 5.000 –ésima, a qual para as 20 cadeias produz uma amostra de tamanho 2.000. Na tabela 4.10 apresentamos o sumário a posteriori dos parâmetros do modelo e na figura 4.8 temos a densidade marginal a posteriori aproximado com as $S = 2.000$ amostras Gibbs. Para todos parâmetros observamos (ver tabela 4.10), $\sqrt{\hat{R}} < 1.1$, indicando a convergência.

Tabela 4.10: Sumário a posteriori dos parâmetros do modelo exponencial-potência

	Média	Mediana	D.P	95% Intervalo de Credibilidade	\hat{R}
α	0,9828	0,9674	0,16199	(0,7169 ; 1,3393)	1,0000081
β	0,6575	0,6419	0,20657	(0,2969 ; 1,1467)	1,0005173

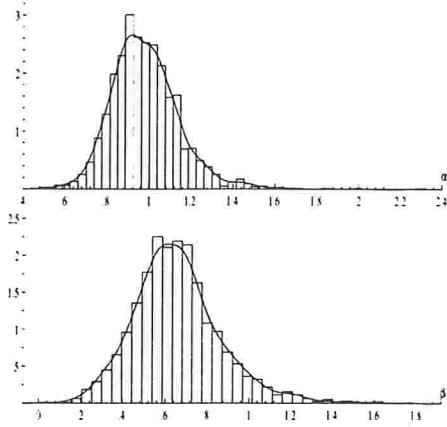


Figura 4.8: Densidades marginais a posteriori aproximadas de α e β do modelo exponencial-potência

Modelo de mistura de gamas

Finalmente considerando, a mistura de distribuições gamas dada em (4.10), e a densidade a priori (4.26) com $\gamma_1 = 1$, $\gamma_2 = 1$, $\lambda_j = 0,000001$ ($j = 1, 2$), $\beta_1 = 1$ $\beta_2 = 0,00001$, geramos 20 cadeias separadas de Gibbs cada um com 5.000 iterações. para cada parâmetro consideramos as iterações 2.550 – éSIMA, 2.600 – éSIMA, ..., 5.000 – éSIMA, as 20 cadeias produz uma amostra de tamanho 1.000. Na tabela 4.11, reportamos o sumário a posteriori dos parâmetros do modelo e, na figura 4.9, temos a densidades marginais a posteriori aproximadas pelas 1.000 amostras Gibbs. Também obtemos a convergência, já que, a estimativa dos fatores de redução de escala potencial $\hat{R} \approx 1$ para todos os parâmetros.

Tabela 4.11: Sumário a posteriori dos parâmetros do modelo de mistura de distribuições gamas

	Média	Mediana	D.P.	95% Intervalo de Credibilidade	\hat{R}
α_1	3,5629	2,9961	3,3662	(0,2050 ; 10,0671)	1,000312
α_2	0,4548	0,4416	0,0958	(0,3412 ; 0,6667)	1,000304
β	2,1423	2,2023	0,1023	(1,2383 ; 4,1721)	1,085947
p	0,0308	0,0054	0,10229	(0,000013 ; 0,4247)	1,021205

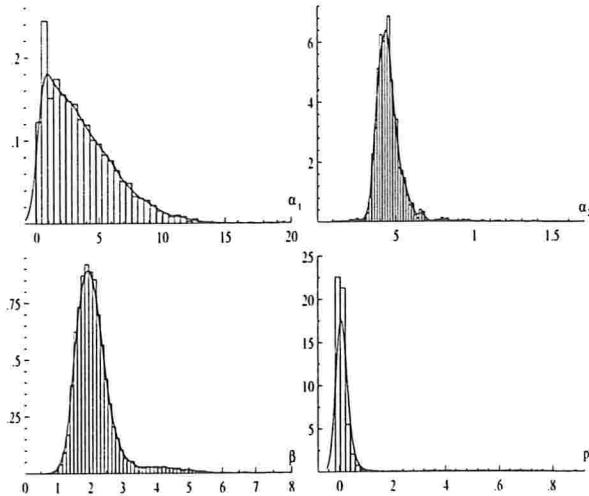


Figura 4.9: Densidades marginais a posteriori aproximados de α_1 , α_2 , β e p parâmetros do modelo de mistura de distribuições gamas

Discriminação de modelos

Como na subseção 5.1, na figura 4.10 temos o “plot” do log da razão CPO para os diferentes modelos com respeito aos outros contra o número de observações. Valores positivos do log da razão dos CPO indicam a preferência do primeiro modelo com respeito ao outro modelo. Na figura 4.10, modelo Weibull-exponenciada versus o modelo IDB o plot indica que

41 das 50 observações sustentam ao modelo Weibull-exponenciada sobre o modelo IDB. Similarmente, 47 observações sustentam ao modelo Weibull-exponenciada sobre o modelo exponencial-potência, 40 observações sustentam ao modelo Weibull-exponenciada sobre o modelo de mistura de distribuições gamas, 28 observações sustentam o modelo IDB sobre o modelo exponencial-potência, 35 observações sustentam o modelo exponencial-potência sobre o modelo de mistura de gamas, 36 observações sustentam o modelo de mistura de distribuições gamas sobre o modelo IDB. Este criterio mostra que o modelo Weibull-exponenciada é a que ajusta melhor aos dados da tabela 4.7.

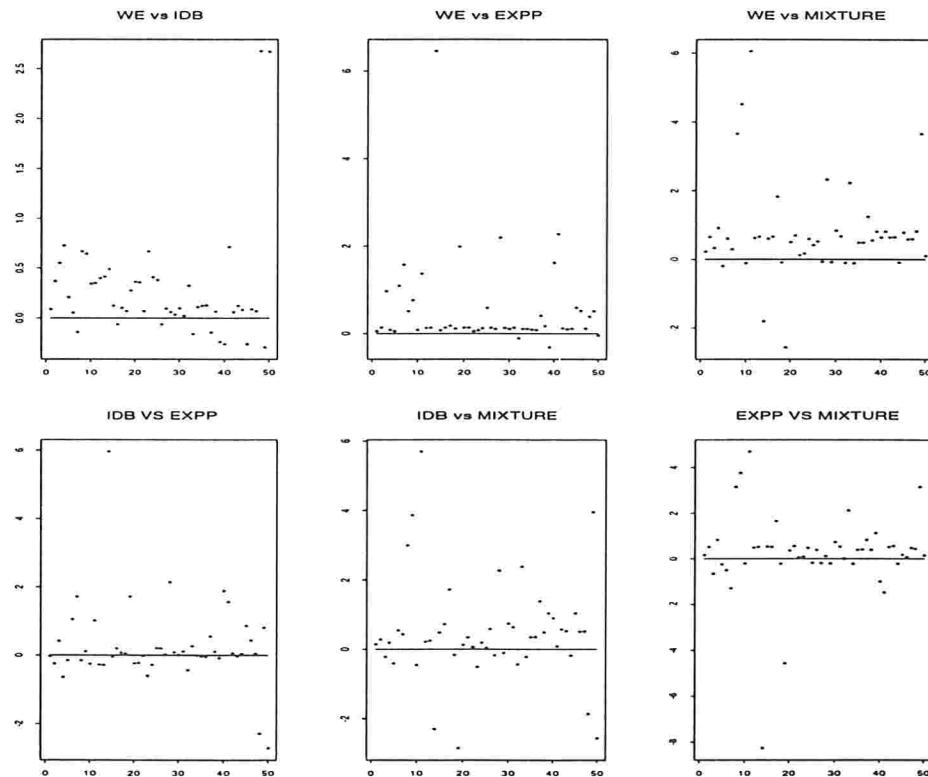


Figura 4.10: Gráfico do log da razão dos CPO para os diferentes modelos versus as observações

A mesma conclusões é também obtida pelo log do pseudo-fator de Bayes (1.11) (ver tabela 4.12).

Tabela 4.12: log do pseudo-fator de Bayes

Modelos	PSFB
Weibull-exponenciada vs IDB	13.9908
Weibull-exponenciada vs exponencial-potência	26.62271
Weibull-exponenciada vs mistura de gamas	38.0926
IDB vs exponencial-potência	12.63263
IDB vs mistura de gamas	24.10018
exponencial-potência vs mistura de gamas	11.46755

Capítulo 5

Comparação de dois modelos exponenciais com dados acelerados

5.1 Introdução

Um teste acelerado consiste basicamente de uma variedade de métodos para reduzir a vida de produtos ou provocar a degradação da performance deste rapidamente. O principal objetivo é obter rapidamente os dados a serem analisados e desta forma produzir informações sobre a vida de um produto ou sua performance sob condições normais de uso. Um abrangente estudo deste ramo da confiabilidade assim como numerosas referências são dadas em Nelson (1990). Contudo, há poucos artigos relacionados a testes de hipóteses. Louzada-Neto, Bolfarine e Rodrigues (1991) desenvolveram uma análise Bayesiana para comparar duas populações sob condições aceleradas considerando a distribuição exponencial, para os tempos de sobrevivência, assumindo dados sob esquema de censura tipo II e densidade a priori não informativa de Jeffreys, para os parâmetros do modelo obteve a densidade a posteriori para o quociente dos tempos médios de vida das duas populações. Rodrigues, Bolfarine, e Louzada-Neto (1993) estenderam o trabalho de Louzada et al. (1991) para o caso de p populações considerando para a comparação duas

abordagens; a primeira baseada na estatística da razão de verossimilhança e a segunda baseada no fator de Bayes a posteriori (Aitkin, 1991). Nas duas abordagens eles consideraram a parametrização ortogonal (Cox e Reid, 1987) nos parâmetros do modelo. Na abordagem Bayesiana consideram a densidade a priori não informativa de Jeffreys para os parâmetros do modelo sob cada uma das hipóteses.

Neste Capítulo apresentamos uma análise Bayesiana alternativa para o problema de comparar duas populações exponenciais com dados acelerados propostos por Louzada et al. (1991), análise é baseada em métodos de simulação de Monte Carlo via Cadeias de Markov.

5.2 Formulação do modelo

Suponha que temos disponível 2 conjunto de dados acelerados independentes, sob o esquema de censura tipo II, tiradas de populações com densidade exponencial sob o estresse V_i dado por

$$f(t; \theta_{ij}) = \frac{1}{\theta_{ij}} \exp\left\{-\frac{t}{\theta_{ij}}\right\} \quad (5.1)$$

onde θ_{ij} (parâmetro desconhecido) $i = 1, \dots, k$, $j=1,2$; é o tempo médio de falha da componente j sob o estresse V_i .

Neste trabalho consideremos o modelo de lei de potência (veja, por exemplo, Nelson, 1990) como a relação entre o parâmetro de escala e o estresse V_i , isto é,

$$\theta_{ij} = \frac{\alpha_{1j}}{V_i^{\alpha_{2j}}}, \quad (5.2)$$

onde $\alpha_{1j} > 0$ e $\alpha_{2j} \in \mathcal{R}$ são parâmetros desconhecidos $j = 1, 2$.

De (5.2), a função de ligação que especifica a relação entre o tempo médio de falha e o nível de estresse V_i é dado por

$$\log(\theta_{ij}) = X'_i \beta_j \quad (5.3)$$

onde $X'_i = (1, \log(V_i^{-1}))$ e o vetor de coeficientes de regressão é $\beta_j = (\beta_{1j}, \beta_{2j})$ dado pela seguinte parametrização

$$\beta_{1j} = \log(\alpha_{1j}), \quad \beta_{2j} = \alpha_{2j}. \quad (5.4)$$

O tempo médio de vida do j -ésimo componente sob o estresse V_i pode ser escrito como

$$\theta_{ij} = \exp\{X'_i \beta_j\}. \quad (5.5)$$

Suponha que tem-se interesse em testar as seguintes hipóteses

$$H_0 : \theta_{11} = \theta_{12} \quad \text{vs} \quad H_1 : \theta_{11} \neq \theta_{12}. \quad (5.6)$$

Isto é, temos interesse em comparar os tempos médios de vida de duas populações no estresse usual V_1 que tem distribuição exponencial. Observe que sob H_0 em (5.6), $\theta_{11} = \theta_{12}$ implica $X'_1 \beta_1 = X'_1 \beta_2$ (veja, (5.5)).

Neste trabalho testamos as hipóteses usando o fator de Bayes. De (5.6) consideramos o modelo M_1 sob a H_0 , ou seja, sob a suposição de que $\theta_{11} = \theta_{12} = \theta_1$ e o modelo M_2 sob H_1 , ou seja, considerando-se $\theta_{11} \neq \theta_{12}$. Um tratamento Bayesiano para testar as hipóteses em (5.6) usando o fator de Bayes a posteriori aproximado pelo método de Laplace é dado em Rodrigues et al. (1992).

5.3 Teste de hipóteses Bayesiano

Vamos descrever inicialmente alguns métodos alternativos para se obter aproximações através de métodos MCMC da verossimilhança preditiva sob o modelo M_j , $j = 1, 2$, definida em (1.5).

Conforme visto na seção 1.2.3 uma simples estimativa de Monte Carlo da função da verossimilhança preditiva, considerando amostra $\theta_1, \dots, \theta_n$ da densidade a priori $\pi(\theta_j | M_j)$

é dada por

$$\hat{f}_1(D|M_j) = \frac{1}{n} \sum_{i=1}^n L(\theta_i|M_j). \quad (5.7)$$

Raftery (1995) argumenta que esse estimador não funciona bem se houver muita discrepância entre a densidade a priori e a verossimilhança. Uma estimativa alternativa para (1.5) é realizar amostragem por importância usando uma função densidade de importância $g(\theta)$, definindo

$$\hat{f}_2(D|M_j) = \frac{\sum_{i=1}^n L(\theta_i|M_j)w_i}{\sum_{i=1}^n w_i},$$

onde $\theta_1, \dots, \theta_n$ é uma amostra da densidade de importância $g(\theta)$ e $w_i = \pi(\theta_i)/g(\theta_i)$.

Normalmente após um processo de simulação via cadeias de Markov, tem-se à disposição uma amostra $\theta_1, \dots, \theta_n$ de $\pi(\theta|D, M_j) = kL(\theta|M_j)\pi(\theta|M_j)$. Logo, esses valores podem ser aproveitados através do estimador \hat{f}_2 bastando para isso tomar $g(\theta) = \pi(\theta|D, M_j)$.

Nesse caso, o estimador se reduz a

$$\hat{f}_3(D|M_j) = \left[\frac{1}{n} \sum_{i=1}^n L(\theta_i|M_j)^{-1} \right]^{-1} \quad (5.8)$$

Este estimador é a média harmônica de valores da verossimilhança e foi originalmente proposto por Newton e Raftery (1994).

Uma outra aproximação para a verossimilhança preditiva (1.5) é obtida a partir do método de Laplace (veja por exemplo, Tierney e Kadane, 1986) a qual é dada por

$$\hat{f}_4(D|M_j) = (2\pi)^{d/2} |\hat{\Sigma}|^{1/2} L(\hat{\theta}|M_j) \pi(\hat{\theta}|M_j) \quad (5.9)$$

onde d é a dimensão do vetor de parâmetros θ , $\hat{\theta}$ é a moda da densidade a posteriori, isto é, o valor que maximiza $\ell(\theta) \propto \log(L(\theta|M_j)\pi(\theta|M_j))$ e $\hat{\Sigma}$ é a inversa da matriz de segundas derivadas de $\ell(\hat{\theta})$. Raftery(1995) sugere para o cálculo desta aproximação utilizar a amostra $\theta_1, \dots, \theta_n$ da densidade a posteriori. Uma estimativa da moda a posteriori é obtida calculando-se a posteriori para todos os pontos da amostra e escolhendo $\hat{\theta}$ de forma

que $\pi(\hat{\theta}) = \max_i \pi(\theta_i | D)$. Similarmente uma estimativa para a matriz de covariâncias a posteriori é dada pela matriz de covariâncias amostrais obtidas a partir da amostra gerada, ou seja,

$$\hat{\Sigma} = \frac{1}{n} \sum_{i=1}^n (\theta_i - \bar{\theta})(\theta_i - \bar{\theta})'$$

onde $\bar{\theta} = \frac{1}{n} \sum_{i=1}^n \theta_i$. Raftery denomina esse estimador de Laplace-Metropolis e discute formas alternativas para o cálculo de estimativas de θ e Σ .

Neste Cípitulo, o fator de Bayes dado em (1.6) do modelo sob H_0 em (5.6) com respeito ao modelo sob H_1 em (5.6) denotada por FB_{12} , é computado com as estimativas das verossimilhanças preditivas propostas por Newton e Raftery (1994) dada em (5.8) e Raftery (1995) dada em (5.9). Rejeita-se H_0 se $FB_{12} < 1$.

5.4 A função de verossimilhança

Nesta seção apresentamos a função de verossimilhança para os modelos M_1 e M_2 considerando dados sob o esquema de censura tipo II, isto é, o experimento termina quando r_i falhas são observadas entre as n_i unidades em testes no i -ésimo nível de estresse para $j = 1, 2$. Temos $t_{ij1}, t_{ij2}, \dots, t_{ijr_i}$ observações não-censuradas e $n_i - r_i$ observações censuradas iguais a t_{ijr_i} , para $i = 1, \dots, k$.

A função de verossimilhança para $\beta'_j = (\beta_{1j}, \beta_{2j})$, $j = 1, 2$, considerando k níveis de estresse e um esquema de censura tipo II, é dada por

$$L(\beta_j) = \prod_{i=1}^k e^{r_i X_i' \beta_j} \exp \left\{ -r_i \hat{\theta}_{ij} e^{-X_i' \beta_j} \right\}, \quad (5.10)$$

onde $\hat{\theta}_{ij} = \frac{A_{ij}}{r_i}$, $A_{ij} = \sum_{l=1}^k t_{ijl} + (n_i - r_i)t_{ijr_i}$

A função de verossimilhança para (β_1, β_2) sob o modelo M_2 é dada por

$$L(\beta_1, \beta_2 | M_2, D) = \prod_{i=1}^k e^{-r_i X_i' (\beta_1 + \beta_2)} \exp \left\{ -r_i (\hat{\theta}_{i1} e^{-X_i' \beta_1} + \hat{\theta}_{i2} e^{-X_i' \beta_2}) \right\}. \quad (5.11)$$

A função de verossimilhança para (β_1, β_2) sob o modelo M_1 é dada por

$$\begin{aligned} L(\beta_1, \beta_2 | M_1, D) &= \prod_{i=2}^k e^{-r_i X'_i (\beta_1 + \beta_2)} \exp \left\{ -r_i (\hat{\theta}_{i1} e^{-X'_i \beta_1} + \hat{\theta}_{i2} e^{-X'_i \beta_2}) \right\} \\ &\times e^{-2r_1 X'_1 \beta_1} \exp \left\{ -r_1 e^{X'_1 \beta_1} (\hat{\theta}_{11} + \hat{\theta}_{12}) \right\}. \end{aligned} \quad (5.12)$$

5.5 Análise Bayesiana

Um aspecto importante na análise Bayesiana é a especificação da densidade a priori $\pi(\beta_j)$ para $\beta_j, j = 1, \dots, k$. Com esse propósito, neste trabalho segue-se a abordagem de Bedrick et al.(1996) para formular uma densidade a priori para β_j chamada "Data Augmentation priors" (DAP). O Procedimento DAP consiste em determinar uma priori para β_j que tenha a forma funcional da função de verossimilhança. Também consideramos a obtenção da densidade a posteriori considerando densidades a priori não informativas e próprias, baseadas na distribuição normal e distribuição gama invertida.

5.5.1 DAP para os modelos M_1 e M_2

Para formular o DAP para os modelos M_1 e M_2 , consideramos os níveis de estresse que satisfazem $V_1 < \dots < V_k$ e, seja $\tilde{X}_i = (1, \log(V_i))'$, $i = 1, 2$ o nível de estresse ou covariável seleccionado tal que a matriz \tilde{X} de orden 2×2 seja não singular ($\tilde{X}_1 \neq \tilde{X}_2$). Daí temos que $\tilde{\theta}_{ij} = \exp\{\tilde{X}_i \beta_j\}$, implicando que

$$\tilde{\theta}_j = \exp\{\tilde{X} \beta_j\}, \quad j = 1, 2$$

onde $\tilde{\theta}'_j = (\tilde{\theta}_{1j}, \tilde{\theta}_{2j})$ e $\tilde{X} = \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \end{bmatrix}_{2 \times 2}$ é uma matriz não singular a qual implica

$$\beta_j = \tilde{X}^{-1} \log(\tilde{\theta}_j), \quad j = 1, 2.$$

Priori DAP para o modelo M_2

De (5.12) e seguindo o procedimento de Bedrick et al.(1996) para o modelo M_2 temos que a densidade priori é dada por

$$\pi(\beta_1, \beta_2 | M_2) \propto \prod_{i=1}^2 e^{-w_i \tilde{X}_i (\beta_1 + \beta_2)} \exp \left\{ -w_i (\hat{\theta}_{i1} e^{-\tilde{X}_i \beta_1} + \hat{\theta}_{i1} e^{-\tilde{X}_i \beta_1}) \right\}. \quad (5.13)$$

A priori média condicional para $(\tilde{\theta}_1, \tilde{\theta}_2)$ (veja, Bedrick, et al.,1996) induzido por (5.13) via mudança de variáveis são $GI(w_i, w_i \hat{\theta}_{ij})$ para, $i = 1, 2$ e $j = 1, 2$, independentes, com densidade conjunta

$$\begin{aligned} \pi(\tilde{\theta}_1, \tilde{\theta}_2 | M_2) &\propto \tilde{\theta}_{11}^{-(w_1+1)} \exp \left\{ -w_1 \frac{\tilde{\theta}_{11}}{\tilde{\theta}_{11}} \right\} \tilde{\theta}_{21}^{-w_2-1} \exp \left\{ -w_2 \frac{\tilde{\theta}_{21}}{\tilde{\theta}_{21}} \right\} \\ &\times \tilde{\theta}_{12}^{-(w_1+1)} \exp \left\{ -w_1 \frac{\tilde{\theta}_{12}}{\tilde{\theta}_{12}} \right\} \tilde{\theta}_{22}^{-w_2-1} \exp \left\{ -w_2 \frac{\tilde{\theta}_{22}}{\tilde{\theta}_{22}} \right\}. \end{aligned} \quad (5.14)$$

O procedimento para atribuir valores a priori para $(w_i, \hat{\theta}_{ij})$, é como segue: para uma escolha apropriada de \tilde{A}_{ij} e nível de estresse \tilde{X}_i fixado define-se para $i = 1, 2$, e $j = 1, 2$,

$$\tilde{p}_i = Pr[A_{ij} > \tilde{A}_{ij} | \tilde{X}_i, \tilde{\theta}_{ij}] = 1 - F\left(\frac{2\tilde{A}_{ij}}{\tilde{\theta}_{ij}}\right), \quad (5.15)$$

onde $F(.)$ é a distribuição qui-quadrado com $2r_i$ graus de liberdade. É razoável considerar $\hat{\theta}_{ij}$ como a informação a priori de $\tilde{\theta}_{ij}$ com o mesmo peso que $\hat{\theta}_{ij}$, isto é, $w_i = r_i$. Para um valor de τ_i , pode-se determinar γ_i de modo que

$$Pr(\tilde{p}_i > \gamma_i) = \tau_i.$$

Daí tem-se

$$\hat{\theta}_{ij} = \frac{\tilde{A}_{ij}}{r_i} U_{ij},$$

onde

$$U_{ij} = \frac{2G^{-1}(\tau_i)}{F^{-1}(1 - \gamma_i)},$$

e $G(.)$ é a distribuição $\Gamma(w_i, 1)$, para $i=1,2$.

Para asegurar a independência entre os parâmetros $\tilde{\theta}_{1j}$ e $\tilde{\theta}_{2j}$ escolhe-se os níveis de estresse seguinte (para detalhes da justificativa desta escolha veja Bedrick et al., 1996 e Rodrigues, 1998):

$$\begin{aligned}\tilde{X}_1 &= (1, \log(V_1))' \longrightarrow w_1 = r_1 \\ \tilde{X}_2 &= (1, \log(V_k))' \longrightarrow w_2 = r_k.\end{aligned}$$

Usando o procedimento de mudança de variáveis estabelecidos em (5.4), obtem-se a seguinte distribuição a posteriori para o modelo M_2 :

$$\begin{aligned}\pi(\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22} | M_2, D) &\propto (\alpha_{11}\alpha_{12})^{-(r+r_1+r_k+1)} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{r_1}) (V_k^{r_k}) \right]^{\alpha_{21}+\alpha_{22}} \\ &\quad \exp \left\{ - \sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\tau}_{ij} V_i^{\alpha_{2j}}}{\alpha_{1j}} \right\},\end{aligned}\tag{5.16}$$

onde

$$\begin{aligned}\hat{\tau}_{ij} &= \hat{\theta}_{ij}, \quad i = 2, \dots, k-1, \quad j = 1, 2 \\ &= \hat{\theta}_{1j} + \hat{\tilde{\theta}}_{1j}, \quad i = 1, \quad j = 1, 2 \\ &= \hat{\theta}_{kj} + \hat{\tilde{\theta}}_{2j}, \quad i = k, \quad j = 1, 2.\end{aligned}$$

De (5.16) pode-se mostrar que as densidades condicionais completas de α_{11} , α_{21} , α_{12} e α_{22} são dadas por

- (i) $\pi(\alpha_{11} | \alpha_{21}, \alpha_{12}, \alpha_{22}, M_2, D) \sim IG(r + r_1 + r_k; \sum_{i=1}^k r_i \hat{\tau}_{i1} V_i^{\alpha_{21}})$,
- (ii) $\pi(\alpha_{21} | \alpha_{11}, \alpha_{12}, \alpha_{22}, M_2, D) \propto \exp \left\{ -\frac{1}{\alpha_{11}} \sum_{i=1}^k r_i \hat{\tau}_{i1} V_i^{\alpha_{21}} \right\} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{r_1}) (V_k^{r_k}) \right]^{\alpha_{21}}$,
- (iii) $\pi(\alpha_{12} | \alpha_{11}, \alpha_{21}, \alpha_{22}, M_2, D) \sim IG(r + r_1 + r_k; \sum_{i=1}^k r_i \hat{\tau}_{i2} V_i^{\alpha_{22}})$,
- (iv) $\pi(\alpha_{22} | \alpha_{11}, \alpha_{21}, \alpha_{12}, M_2, D) \propto \exp \left\{ -\frac{1}{\alpha_{12}} \sum_{i=1}^k r_i \hat{\tau}_{i2} V_i^{\alpha_{22}} \right\} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{r_1}) (V_k^{r_k}) \right]^{\alpha_{22}}.$

Observe que (i) e (iii) são distribuições padrões e (ii) e (iv) são distribuições não conhecidas devendo-se portanto utilizar o algoritmo de Metropolis-Hastings dentro do ciclo do amostrador de Gibbs para obter as distribuições marginais a posteriori aproximadas.

Priori DAP para M_1

De (5.11) e seguindo o procedimento de Bedrick et al.(1996) para o modelo M_1 temos que a densidade a priori é dada por

$$\begin{aligned}\pi(\beta_1, \beta_2 | M_1) &\propto e^{-2w_1 \tilde{X}'_1 \beta_1} \exp \left\{ -2w_1 e^{-\tilde{X}'_1 \beta_1} \left(\frac{\hat{\theta}_{11} + \hat{\theta}_{12}}{2} \right) \right\} \\ &\times e^{-r_2 \tilde{X}'_2 (\beta_1 + \beta_2)} \exp \left\{ -r_2 (\hat{\theta}_{i1} e^{-\tilde{X}'_2 \beta_1} + \hat{\theta}_{i2} e^{-\tilde{X}'_2 \beta_2}) \right\},\end{aligned}\quad (5.17)$$

onde $\hat{\theta}_{i1}$ é a informação a priori de $\tilde{\theta}_{i1}$ e w_i é o peso de $\tilde{\theta}_{i1}$, $i = 1, 2$.

A priori média condicional para $(\tilde{\theta}_1, \tilde{\theta}_2)$ (veja, Bedrick, et al.,1996) induzida por (5.17) via a técnica de mudança de variáveis é dada por

$$\begin{aligned}\pi(\tilde{\theta}_1, \tilde{\theta}_2 | M_1) &\propto \tilde{\theta}_{11}^{-(2w_1+1+1)} \exp \left\{ -2w_1 \frac{\tilde{\theta}_1}{\tilde{\theta}_{11}} \right\} \tilde{\theta}_{21}^{-w_2-1} \exp \left\{ -w_2 \frac{\hat{\theta}_{21}}{\tilde{\theta}_{21}} \right\} \\ &\times \tilde{\theta}_{22}^{-w_2-1} \exp \left\{ -w_2 \frac{\hat{\theta}_{22}}{\tilde{\theta}_{22}} \right\},\end{aligned}\quad (5.18)$$

onde $\bar{\theta}_1 = \frac{\hat{\theta}_{11} + \hat{\theta}_{12}}{2}$. De (5.18) observa-se que as densidades a priori são independentes, com

$$\tilde{\theta}_{11} \sim \text{IG}(2w_1 + 1, 2w_1 \bar{\theta}_1), \quad (5.19)$$

$$\tilde{\theta}_{21} \sim \text{IG}(w_2, w_2 \hat{\theta}_{21}), \quad (5.20)$$

$$\tilde{\theta}_{22} \sim \text{IG}(w_2, w_2 \hat{\theta}_{22}), \quad (5.21)$$

onde $\text{IG}(a, b)$ denota a densidade gama invertida com média $\frac{b}{a-1}$ e variância $\frac{b^2}{(a-1)^2(a-2)}$.

O procedimento para atribuir os valores a priori para $\bar{\theta}_1$, é como segue: para uma escolha apropriada de \tilde{A}_1 e nível de estresse \tilde{X}_1 fixado acima define-se,

$$\tilde{p}_1 = Pr[A_1 > \tilde{A}_1 | \tilde{X}_1, \tilde{\theta}_{11}] = 1 - F\left(\frac{2\tilde{A}_1}{\tilde{\theta}_{11}}\right), \quad (5.22)$$

onde $F(\cdot)$ é a distribuição qui-quadrado com $4r_1$ graus de liberdade. Para um valor de τ_1 , pode-se determinar γ_1 de modo que

$$Pr(\tilde{p}_1 > \gamma_1) = \tau_1.$$

Daí tem-se

$$\bar{\tilde{\theta}}_1 = \frac{\tilde{A}_1}{2r_1} U_{11},$$

onde

$$U_{11} = \frac{2G^{-1}(\tau_1)}{F^{-1}(1 - \gamma_1)},$$

e $G(\cdot)$ é a distribuição $\Gamma(2r_1 + 1, 1)$, ou seja, a distribuição gama com parâmetros $2r_1 + 1$ e 1.

Segue-se o mesmo procedimento de (5.22) para se atribuir valores a priori para $\hat{\tilde{\theta}}_{21}$ e $\hat{\tilde{\theta}}_{22}$, obtendo-se

$$\hat{\tilde{\theta}}_{21} = \frac{\tilde{A}_{21}}{r_k} U_{21} \quad \text{e} \quad \hat{\tilde{\theta}}_{22} = \frac{\tilde{A}_{22}}{r_k} U_{22},$$

sendo

$$U_{21} = U_{22} = \frac{2G^{-1}(\tau_2)}{F^{-1}(1 - \gamma_2)},$$

onde $F(\cdot)$ é a distribuição de qui-quadrado com $2r_k$ graus de liberdade e $G(\cdot)$ a distribuição $\Gamma(r_k, 1)$.

Usando a técnica de mudança de variáveis em (5.4), obtem-se a seguinte distribuição a posteriori para o modelo M_1 :

$$\begin{aligned} \pi(\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22} | M_1, D) &\propto \alpha_{11}^{-(r+3r_1+r_k+1)} \alpha_{12}^{-(r-r_1+r_k+1)} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{3r_1}) (V_k^{r_k}) \right]^{\alpha_{21}} \\ &\quad \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{-r_1}) (V_k^{r_k}) \right]^{\alpha_{22}} \exp \left\{ - \sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\eta}_{ij} V_i^{\alpha_{2j}}}{\alpha_{1j}} \right\}, \end{aligned} \quad (5.23)$$

onde

$$\begin{aligned} \hat{\eta}_{i1} &= \hat{\theta}_{i1}, & i &= 2, \dots, k-1, \\ &= \hat{\theta}_{11} + \hat{\theta}_{12} + 2\bar{\tilde{\theta}}_1, & i &= 1, \\ &= \hat{\theta}_{k1} + \hat{\tilde{\theta}}_{21}, & i &= k \end{aligned}$$

e,

$$\hat{\eta}_{i2} = \hat{\theta}_{i2}, \quad i = 2, \dots, k-1,$$

$$= 0, \quad i = 1, \\ = \hat{\theta}_{k2} + \hat{\tilde{\theta}}_{22}, \quad i = k.$$

De (5.23) pode-se mostrar que as densidades condicionais completas para α_{11} , α_{21} , α_{12} e α_{22} são dadas por

- (i) $\pi(\alpha_{11}|\alpha_{21}, \alpha_{12}, \alpha_{22}, M_1, D) \sim IG(r + 3r_1 + r_k; \sum_{i=1}^k r_i \eta_{i1} V_i^{\alpha_{21}});$
- (ii) $\pi(\alpha_{21}|\alpha_{11}, \alpha_{12}, \alpha_{22}, M_1, D) \propto \exp\left\{-\frac{1}{\alpha_{11}} \sum_{i=1}^k r_i \hat{\eta}_{i1} V_i^{\alpha_{21}}\right\} \left[(\prod_{i=1}^k V_i^{r_i})(V_1^{3r_1})(V_k^{r_k})\right]^{\alpha_{21}};$
- (iii) $\pi(\alpha_{12}|\alpha_{11}, \alpha_{21}, \alpha_{22}, M_1, D) \sim IG(r - r_1 + r_k; \sum_{i=1}^k r_i \eta_{i2} V_i^{\alpha_{22}});$
- (iv) $\pi(\alpha_{22}|\alpha_{11}, \alpha_{21}, \alpha_{12}, M_1, D) \propto \exp\left\{-\frac{1}{\alpha_{12}} \sum_{i=1}^k r_i \hat{\eta}_{i2} V_i^{\alpha_{22}}\right\} \left[(\prod_{i=1}^k V_i^{r_i})(V_1^{-r_1})(V_k^{r_k})\right]^{\alpha_{22}},$

onde $r = \sum_{i=1}^k r_i$. Observe que (i) e (iii) são distribuições conhecidas e (ii) e (iv) são distribuições intratáveis de ponto de vista de simulação devendo-se portanto, utilizar o algoritmo de Metropolis-Hastings dentro do ciclo do amostrador de Gibbs para se obter as distribuições marginais a posteriori aproximadas.

5.5.2 Priori de Jeffreys

Considerando o modelo (5.1) e a lei de potência (5.2), a função de verossimilhança de $(\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22})$ sob a hipótese H_0 (modelo M_1) em (5.6) é dada por

$$L(\alpha_{11}\alpha_{21}, \alpha_{12}, \alpha_{22}|M_1) = \alpha_{11}^{-(r+r_1)} \alpha_{12}^{-(r-r_1)} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) V_1^{r_1} \right]^{\alpha_{21}} \left[\left(\prod_{i=1}^k V_i \right) V^{-r_1} \right]^{\alpha_{22}} \times \exp\left\{-\sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\eta}_{ij}^* V_i^{\alpha_{2j}}}{\alpha_{1j}}\right\}, \quad (5.24)$$

onde

$$\begin{aligned} \hat{\eta}_{i1}^* &= \hat{\theta}_{i1}, \quad i = 2, \dots, k \\ &= \hat{\theta}_{11} + \hat{\theta}_{12}, \quad i = 1. \end{aligned}$$

e

$$\begin{aligned} \hat{\eta}_{i2}^* &= \hat{\theta}_{i2}, \quad i = 2, \dots, k \\ &= 0, \quad i = 1, \end{aligned}$$

Pode-se mostrar que a matriz de informação de Fisher para o modelo M_1 é dada por

$$I_{M_1} = \begin{pmatrix} \frac{r+r_1}{\alpha_{11}^2} & -\frac{r_1 \log(V_1) + \sum_{i=1}^k r_i \log(V_i)}{\alpha_{11}} & 0 & 0 \\ -\frac{r_1 \log(V_1) + \sum_{i=1}^k r_i \log(V_i)}{\alpha_{11}} & r_1 \log^2(V_1) + \sum_{i=1}^k r_i \log^2(V_i) & 0 & 0 \\ 0 & 0 & \frac{r-r_1}{\alpha_{12}^2} & -\frac{\sum_{i=2}^k r_i \log(V_i)}{\alpha_{12}} \\ 0 & 0 & -\frac{\sum_{i=2}^k r_i \log(V_i)}{\alpha_{12}} & \sum_{i=2}^k r_i \log^2(V_i) \end{pmatrix} \quad (5.25)$$

A função de verossimilhança de $(\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22})$ sob a H_1 (modelo M_2) em (5.6) é dada por

$$L(\alpha_{11}\alpha_{21}, \alpha_{12}, \alpha_{22}|M_2) = (\alpha_{11}\alpha_{12})^{-r} \left[\prod_{i=1}^k V_i^{r_i} \right]^{\alpha_{21}+\alpha_{22}} \exp \left\{ - \sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\theta}_{ij} V_i^{\alpha_{2j}}}{\alpha_{1j}} \right\}. \quad (5.26)$$

Pode-se mostrar que a matriz de informação para o modelo M_2 é dada por

$$I_{M_2} = \begin{pmatrix} \frac{r}{\alpha_{11}^2} & -\frac{\sum_{i=1}^k r_i \log(V_i)}{\alpha_{11}} & 0 & 0 \\ -\frac{\sum_{i=1}^k r_i \log(V_i)}{\alpha_{11}} & \sum_{i=1}^k r_i \log^2(V_i) & 0 & 0 \\ 0 & 0 & \frac{r}{\alpha_{12}^2} & -\frac{\sum_{i=1}^k r_i \log(V_i)}{\alpha_{12}} \\ 0 & 0 & -\frac{\sum_{i=1}^k r_i \log(V_i)}{\alpha_{12}} & \sum_{i=2}^k r_i \log^2(V_i) \end{pmatrix} \quad (5.27)$$

De (5.25) e (5.27) pode-se mostrar que a priori não informativa de Jeffreys para os modelos M_1 e M_2 é dada por

$$\pi(\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22}|M_j) \propto \frac{1}{\alpha_{11}\alpha_{12}}. \quad (5.28)$$

Considerando (5.24) e (5.28) e o teorema de Bayes pode-se mostrar que a distribuição a posteriori para o modelo M_1 é dada por

$$\begin{aligned} \pi(\alpha_{11}, \alpha_{21}, \alpha_{12}, \alpha_{22} | M_1, D) &\propto \alpha_{11}^{-(r+r_1+1)} \alpha_{12}^{-(r-r_1+1)} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) V_1^{r_1} \right]^{\alpha_{21}} \\ &\quad \left[\left(\prod_{i=1}^k V_i^{r_i} \right) V_1^{-r_1} \right]^{\alpha_{22}} \exp \left\{ - \sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\eta}_{ij}^* V_i^{\alpha_{2j}}}{\alpha_{1j}} \right\}. \end{aligned} \quad (5.29)$$

As distribuições condicionais para o algoritmo de Gibbs para esse modelo são dadas por

- (i) $\pi(\alpha_{11} | \alpha_{21}, \alpha_{12}, \alpha_{22}, M_1, D) \sim IG(r + r_1; \sum_{i=1}^k r_i \hat{\eta}_{i1}^* V_i^{\alpha_{21}}),$
- (ii) $\pi(\alpha_{21} | \alpha_{11}, \alpha_{12}, \alpha_{22}, M_1, D) \propto \exp \left\{ - \frac{1}{\alpha_{11}} \sum_{i=1}^k r_i \hat{\eta}_{i1}^* V_i^{\alpha_{21}} \right\} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{r_1}) \right]^{\alpha_{21}},$
- (iii) $\pi(\alpha_{12} | \alpha_{11}, \alpha_{21}, \alpha_{22}, M_1, D) \sim IG(r - r_1; \sum_{i=1}^k r_i \hat{\eta}_{i2}^* V_i^{\alpha_{22}}),$
- (iv) $\pi(\alpha_{22} | \alpha_{11}, \alpha_{21}, \alpha_{12}, M_1, D) \propto \exp \left\{ - \frac{1}{\alpha_{12}} \sum_{i=1}^k r_i \hat{\eta}_{i2}^* V_i^{\alpha_{22}} \right\} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{-r_1}) \right]^{\alpha_{22}}.$

Utiliza-se o algoritmo de Metropolis-Hastings para obter amostras das densidades marginais em (ii) e (iv).

Considerando-se (5.26) e (5.28) e o teorema de Bayes pode-se mostrar que a distribuição a posteriori para o modelo M_2 é dada por

$$\pi(\alpha_{11}\alpha_{21}, \alpha_{12}, \alpha_{22} | M_2) \propto (\alpha_{11}\alpha_{12})^{-(r+1)} \left[\prod_{i=1}^k V_i^{r_i} \right]^{\alpha_{21}+\alpha_{22}} \exp \left\{ - \sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\theta}_{ij} V_i^{\alpha_{2j}}}{\alpha_{1j}} \right\}. \quad (5.30)$$

As distribuições condicionais para o algoritmo de Gibbs para esse modelo são dadas por

- (i) $\pi(\alpha_{11}|\alpha_{21}, \alpha_{12}, \alpha_{22}, M_2, D) \sim IG(r; \sum_{i=1}^k r_i \hat{\theta}_{i1} V_i^{\alpha_{21}}),$
- (ii) $\pi(\alpha_{21}|\alpha_{11}, \alpha_{12}, \alpha_{22}, M_2, D) \propto \exp\left\{-\frac{1}{\alpha_{11}} \sum_{i=1}^k r_i \hat{\theta}_{i1} V_i^{\alpha_{21}}\right\} \left[\left(\prod_{i=1}^k V_i^{r_i}\right)\right]^{\alpha_{21}},$
- (iii) $\pi(\alpha_{12}|\alpha_{11}, \alpha_{21}, \alpha_{22}, M_2, D) \sim IG(r; \sum_{i=1}^k r_i \hat{\theta}_{i2} V_i^{\alpha_{22}}),$
- (iv) $\pi(\alpha_{22}|\alpha_{11}, \alpha_{21}, \alpha_{12}, M_2, D) \propto \exp\left\{-\frac{1}{\alpha_{12}} \sum_{i=1}^k r_i \hat{\theta}_{i2} V_i^{\alpha_{22}}\right\} \left[\left(\prod_{i=1}^k V_i^{r_i}\right)\right]^{\alpha_{22}}.$

Nas situações (ii) e (iv), utiliza-se o algoritmo de Metropolis-Hasting para obter amostras das densidades marginais.

5.5.3 Prioris Próprias

Neste caso consideremos as seguintes densidades a priori independentes para ambos os modelos

$$\begin{aligned}\alpha_{1j} &\sim GI(a_{1j}, b_{1j}), \text{ onde } a_{1j} \text{ e } b_{1j}, j = 1, 2, \text{ são conhecidos,} \\ \alpha_{2j} &\sim N(\mu_{0j}, \sigma_{0j}^2), \text{ onde } \mu_{0j} \text{ e } \sigma_{0j}, j = 1, 2, \text{ são conhecidos,}\end{aligned}\quad (5.31)$$

onde $N(\mu, \sigma^2)$ denota a distribuição normal com média μ e variância σ^2 .

Considerando (5.24) e (5.31) e o teorema de Bayes pode-se mostrar que a distribuição a posteriori para o modelo M_1 é dada por

$$\begin{aligned}\pi(\alpha_{11}\alpha_{21}, \alpha_{12}, \alpha_{22}|M_1, D) &\propto \alpha_{11}^{-(r+r_1+a_{11}+1)} \alpha_{12}^{-(r-r_1+a_{12}+1)} \left[\left(\prod_{i=1}^k V_i^{r_i}\right) V^{r_1}\right]^{\alpha_{21}} \left[\left(\prod_{i=1}^k V_i\right) V_1^{-r_1}\right]^{\alpha_{22}} \\ &\exp\left\{-\sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\eta}_{ij}^* V_i^{\alpha_{2j}}}{\alpha_{1j}} - \frac{1}{2} \sum_{j=1}^2 \left(\frac{\alpha_{2j} - \mu_{0j}}{\sigma_{0j}}\right)^2 - \sum_{j=1}^2 \frac{b_{1j}}{\alpha_{1j}}\right\}.\end{aligned}\quad (5.32)$$

As distribuições condicionais para o algoritmo de Gibbs para esse modelo são dadas por

- (i) $\pi(\alpha_{11}|\alpha_{21}, \alpha_{12}, \alpha_{22}, M_1, D) \sim IG(r + r_1 + a_{11}; b_{11} + \sum_{i=1}^k r_i \hat{\eta}_{i1}^* V_i^{\alpha_{21}}),$
- (ii) $\pi(\alpha_{21}|\alpha_{11}, \alpha_{12}, \alpha_{22}, M_1, D) \propto \exp \left\{ -\frac{1}{\alpha_{11}} \sum_{i=1}^k r_i \hat{\eta}_{i1}^* V_i^{\alpha_{21}} - \frac{1}{2} \left(\frac{\alpha_{21} - \mu_{01}}{\sigma_{01}} \right)^2 \right\}$
 $\times \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{r_1}) \right]^{\alpha_{21}},$
- (iii) $\pi(\alpha_{12}|\alpha_{11}, \alpha_{21}, \alpha_{22}, M_1, D) \sim IG(r - r_1 + a_{12}; b_{12} + \sum_{i=1}^k r_i \hat{\eta}_{i2}^* V_i^{\alpha_{22}}),$
- (iv) $\pi(\alpha_{22}|\alpha_{11}, \alpha_{21}, \alpha_{12}, M_1, D) \propto \exp \left\{ -\frac{1}{\alpha_{12}} \sum_{i=1}^k r_i \hat{\eta}_{i2}^* V_i^{\alpha_{22}} - \frac{1}{2} \left(\frac{\alpha_{22} - \mu_{02}}{\sigma_{02}} \right)^2 \right\}$
 $\times \left[\left(\prod_{i=1}^k V_i^{r_i} \right) (V_1^{-r_1}) \right]^{\alpha_{22}}.$

Em (i) e (iii) as distribuições são conhecidas, em (ii) e (iv) as distribuições são de difícil tratamento e portanto, utiliza-se o algoritmo de Metropolis-Hastings para se obter amostras de suas densidades marginais.

Considerando (5.26) e (5.31) e o teorema de Bayes pode-se mostrar que a distribuição a posteriori para o modelo M_2 é dada por

$$\begin{aligned} \pi(\alpha_{11}\alpha_{21}, \alpha_{12}, \alpha_{22}|M_2) &\propto \alpha_{11}^{-(r+a_{11}+1)} \alpha_{12}^{-(r+a_{12}+1)} \left[\prod_{i=1}^k V_i^{r_i} \right]^{\alpha_{21}+\alpha_{22}} \\ &\times \exp \left\{ -\sum_{j=1}^2 \sum_{i=1}^k \frac{r_i \hat{\theta}_{ij} V_i^{\alpha_{2j}}}{\alpha_{1j}} - \frac{1}{2} \sum_{j=1}^2 \left(\frac{\alpha_{2j} - \mu_{0j}}{\sigma_{0j}} \right)^2 - \sum_{j=1}^2 \frac{b_{ij}}{\alpha_{1j}} \right\}. \end{aligned} \quad (5.33)$$

As distribuições condicionais para o algoritmo de Gibbs para esse modelo é dada por

- (i) $\pi(\alpha_{11}|\alpha_{21}, \alpha_{12}, \alpha_{22}, M_2, D) \sim IG(r + a_{11}; b_{11} + \sum_{i=1}^k r_i \hat{\theta}_{i1} V_i^{\alpha_{21}}),$
- (ii) $\pi(\alpha_{21}|\alpha_{11}, \alpha_{12}, \alpha_{22}, M_2, D) \propto \exp \left\{ -\frac{1}{\alpha_{11}} \sum_{i=1}^k r_i \hat{\theta}_{i1} V_i^{\alpha_{21}} - \frac{1}{2} \left(\frac{\alpha_{21} - \mu_{01}}{\sigma_{01}} \right)^2 \right\} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) \right]^{\alpha_{21}},$
- (iii) $\pi(\alpha_{12}|\alpha_{11}, \alpha_{21}, \alpha_{22}, M_2, D) \sim IG(r + a_{12}; b_{12} + \sum_{i=1}^k r_i \hat{\theta}_{i2} V_i^{\alpha_{22}}),$
- (iv) $\pi(\alpha_{22}|\alpha_{11}, \alpha_{21}, \alpha_{12}, M_2, D) \propto \exp \left\{ -\frac{1}{\alpha_{12}} \sum_{i=1}^k r_i \hat{\theta}_{i2} V_i^{\alpha_{22}} - \frac{1}{2} \left(\frac{\alpha_{22} - \mu_{02}}{\sigma_{02}} \right)^2 \right\} \left[\left(\prod_{i=1}^k V_i^{r_i} \right) \right]^{\alpha_{22}}.$

As distribuições de (i) e (iii) são conhecidas, em (ii) e (iv) as distribuições não são conhecidas e portanto utiliza-se o algoritmo de Metropolis-Hastings para obter amostras de suas densidades marginais.

5.6 Aplicação

Considere os dados da tabela 5.1, gerados assumindo-se uma distribuição exponencial e o modelo lei de potência com $\alpha_{1j} = 400$ e $\alpha_{2j} = 0.7$, $j=1,2$. Este dados são considerados em Rodrigues et al. (1993) e Louzada-Neto et al. (1991).

Tabela 5.1: Dados gerados com $\alpha_{1j} = 400$ e $\alpha_{2j} = 0.7$

i	V_i	n_i	r_i	θ_{ij}	Amostra 1	Amostra 2
1	5	10	4	129.65	16,30,41,63	14,27,42, 58
2	10	10	6	78.81	15,21,36,39,54,59	17,21,32,45,59,61
3	15	10	7	60.09	18,29,39,42,46,57,60	14,23,26,35,42,78
4	20	10	8	49.13	19,20,36,37,41,45,45,57	17,17,20,32,35,40,64,71
4	25	10	9	42.02	8,12,13,14,23,33,42,51,67	10,13,21,28,33,44,53,69,75

Considerando o modelo sob H_0 dado em (5.6) e a densidade a priori DAP para esse modelo com as seguintes informações a priori: $\tilde{A}_1 = 1017$, $\tilde{A}_{21} = 330$, $\tilde{A}_{22} = 425$, $\tau_1 = 0.9$, $\gamma_1 = 0,20$, $\tau_2 = 0,25$, $\gamma_2 = 0,70$, $r_1 = 4$ e $r_k = 9$, obtem-se $\bar{\theta}_1 = 161,44$, $\hat{\theta}_{21} = 34,73$ e $\hat{\theta}_{22} = 44,72$. Com essa escolha foram geradas 5 cadeias com 40.000 iterações cada uma e monitorou-se a convergência da amostra de Gibbs usando o método de Gelman e Rubin(1992), método este que utiliza a técnica de análise de variância para verificar se outras iterações são necessárias. Para cada parâmetro considerou-se as iterações $10.050^{th}, 10.100^{th}, \dots, 20.000^{th}$ de cada cadeia, produzindo-se assim uma amostra final de tamanho 2.000. Na tabela 5.2, apresenta-se o resumo das distribuições a posteriori e, na figura 5.1 tem-se as densidades a posteriori aproximadas obtidas considerando-se as 2.000 amostras de Gibbs. Também, na tabela 5.2, tem-se o resultado da redução potencial estimada de escala \hat{R} (veja, Gelman e Rubin,1992) para todos os parâmetros. Neste caso o número de iterações consideradas foram suficientes para obter a convergência (ou seja, $\sqrt{\hat{R}} < 1.1$), para todos os parâmetros, veja tabela 5.2).

Tabela 5.2: Sumário a posteriori para o modelo sob H_0 com DAP

	Média	Mediana	S.D	Intervalo de credibilidade 95%	\hat{R}
α_{11}	655,49	554,11	614,56	(212,86 ; 1.644,64)	1,008347
α_{21}	0,8326	0,8230	0,1986	(0,4607 ; 1,2252)	1,003009
α_{12}	1.050,70	366,30	2100,60	(21,60 ; 6.976,30)	1,012548
α_{22}	0,6282	0,6271	0,5005	(-0,3277 ; 1,5976)	1,0022077

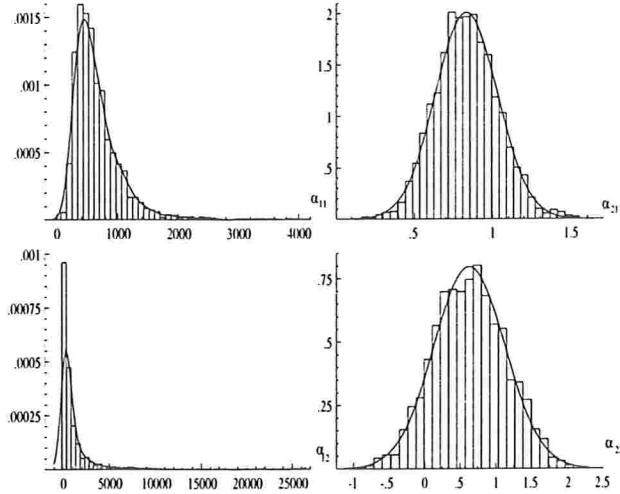


Figura 5.1: Densidades marginais a posteriori aproximados para α_{11} , α_{21} , α_{12} , α_{22} do modelo sob H_0

Para o modelo sob H_1 dado em (5.6) e DAP com as seguintes informações a priori: $\tilde{A}_{11} = 528$, $\tilde{A}_{21} = 330$, $\tilde{A}_{12} = 489$, $\tilde{A}_{22} = 425$, $\tau_1 = 0,90$, $\gamma_1 = 0,20$, $\tau_2 = 0,25$, $\gamma_2 = 0,70$, $r_1 = 4$ e $r_k = 9$, obtem-se $\hat{\theta}_{11} = 159,901$, $\hat{\theta}_{21} = 34,73$, $\hat{\theta}_{12} = 148,09$ e $\hat{\theta}_{22} = 44,72$. Com essa escolha foram geradas 5 cadeias com 40.000 iterações e monitorou-se a convergência da amostra Gibbs usando o método de Gelman e Rubin(1992). Para cada parâmetro considerou-se as iterações $20.050^{th}, 20.100^{th}, \dots, 40.000^{th}$ de cada cadeia, que produz uma amostra final de tamanho 2.000. Na tabela 3, apresenta-se o resumo da distribuição a

posteriori juntamente com o resultado da redução potencial estimada de escala \hat{R} (veja, Gelman e Rubin, 1992) para todos os parâmetros, com valores bastante próximos de um . Na figura 5.2, tem-se as densidades marginais a posteriori para α_{11} , α_{21} , α_{12} e α_{22} .

Tabela 5.3: Sumário a posteriori para o modelo sob H_1 com DAP

	Média	Mediana	S.D	Intervalo de credibilidade 95%	\hat{R}
α_{11}	754,40	582,40	608,70	(167,84 ; 2.254,66)	1,01388
α_{21}	0,8460	0,8392	0,2457	(0,3773 ; 1,3204)	1,015013
α_{12}	546,90	421,10	503,10	(104,37 ; 1.862,40)	1,012820
α_{22}	0,6721	0,6774	0,25947	(0,1741 ; 1,2099)	1,001652

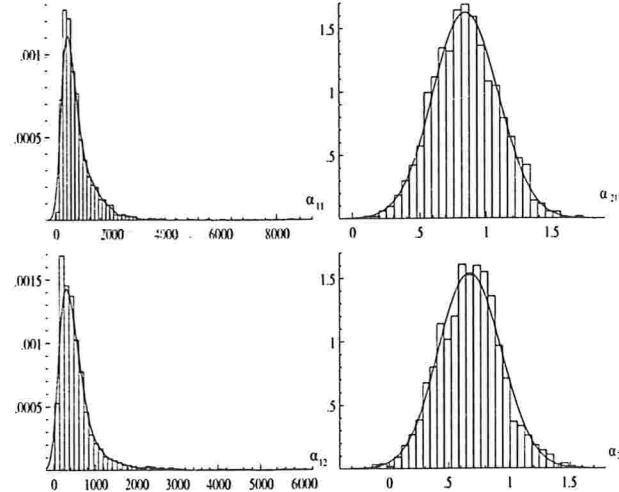


Figura 5.2: Densidades marginais a posteriori aproximados para α_{11} , α_{21} , α_{12} , α_{22} do modelo sob H_1 com DAP

Considerando o modelo sob H_0 e densidade a priori não informativa de Jeffreys dada em (5.28), foram geradas 10 cadeias cada uma com 20.000 iterações. Foram consideradas as iterações $10.100^{th}, 10.200^{th}, \dots, 20.000^{th}$ de cada cadeia, que produz uma amostra de tamanho 2.000. Na tabela 5.4, apresenta-se o resumo das distribuições a posteriori

juntamente com o resultado da redução potencial de escala para todos os parâmetros observamos que $\sqrt{\hat{R}} < 1.1$, indicando a convergência e, na figura 5.3, apresentamos as densidades marginais a posteriori aproximadas.

Tabela 5.4: Sumário a posteriori para o modelo sob H_0 com a densidade a priori não informativa

	Média	Mediana	S.D	Intervalo de credibilidade 95%	\hat{R}
α_{11}	540,30	412,76	427,97	(95,01 ; 1.757,29)	1.000945
α_{21}	0,7069	0,7177	0,2791	(0,1388 ; 1,2391)	1.0001597
α_{12}	1.825,00	524,00	4.952,00	(28,60 ; 11.468,50)	1.0611856
α_{22}	0,7440	0,7424	0,5364	(-0,2670 ; 1,8314)	1.0902412

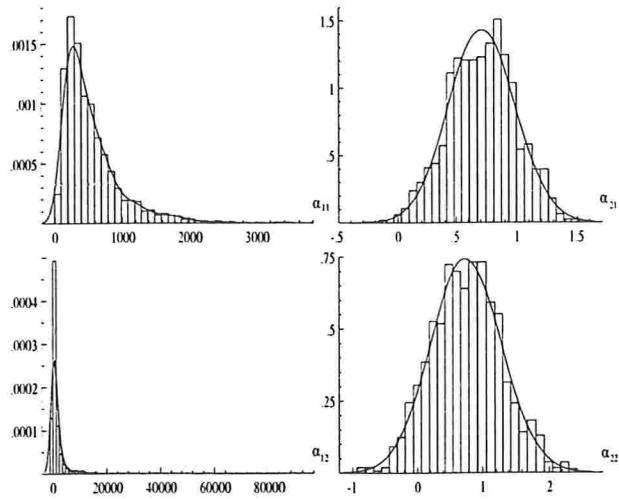


Figura 5.3: Densidades marginais a posteriori aproximados para α_{11} , α_{21} , α_{12} , α_{22} do modelo sob H_0 com priori não informativa

Considerando o modelo sob H_1 com a densidade a priori não informativa de Jeffreys, foram simuladas 10 cadeias cada um com 20.000 iterações, considerando-se as iterações $10.050^{th}, 10.100^{th}, \dots, 20.000^{th}$ de cada cadeia, produzindo uma amostra de tama-

nho 2.000. Na tabela 5.5, apresenta-se o resumo a da distribuição a posteriori juntamente com o resultado da redução potêncial de escala, para todos os parâmetros observamos que $\sqrt{\hat{R}} < 1.1$ indicando que convergência é atingida e, na figura 5.4, apresentamos as densidades marginais aproximadas a posteriori.

Tabela 5.5: Sumário a posteriori para o modelo sob H_1 com priori não informativa

	Média	Mediana	S.D	Intervalo de credibilidade 95%	\hat{R}
α_{11}	637,30	456,80	563,90	(92,46 ; 2.143,56)	1.03744
α_{21}	0,7454	0,7405	0,2889	(0,15123;1,3123)	1.006111
α_{12}	520,20	309,70	581,30	(55,81 ; 2.403,59)	1.05726
α_{22}	0,5848	0,5648	0,5828	(-0,0658 ; 1,2796)	1.00526

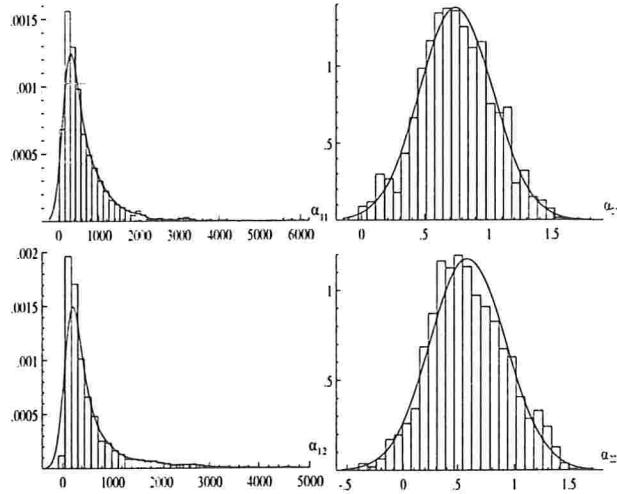


Figura 5.4: Densidades marginais a posteriori aproximados para α_{11} , α_{21} , α_{12} , α_{22} do modelo sob H_1 com priori não informativa

Considerando o modelo sob H_0 e densidade a priori informativa (5.31) com, $a_{1j} = 3$, $b_{1j} = 10.000$, $\mu_{0j} = 0$ e $\sigma_{0j} = 100$, $j = 1, 2$, foram simulados 10 cadeias cada um com 20,000 iterações, considerando-se as iterações $10.100^{th}, 10.200^{th}, \dots, 20.000^{th}$ de cada

cadeia, que produz uma amostra de tamanho 1.000. Na tabela 5.6, apresenta-se o resumo das distribuições a posteriori juntamente com o resultado da redução potêncial de escala. onde para todos os parâmetros observamos que $\sqrt{\hat{R}} < 1.1$ indicando convergência e, na figura 5.5, apresentamos as densidades marginais a posteriori aproximadas.

Tabela 5.6: Sumário a posteriori para o modelo sob H_0 com densidade a priori informativa

	Média	Mediana	S.D	Intervalo de credibilidade 95%	\hat{R}
α_{11}	337,28	325,76	88,18	(203,016;502,754)	1,000850
α_{21}	0,61848	0,61838	0,0850	(0,452688;0,757692)	1,0013037
α_{12}	403,55	338,64	235,14	(157,77;848,33)	1,001466
α_{22}	0,61733	0,60455	0,18073	(0,24091;0,95937)	1,0016882

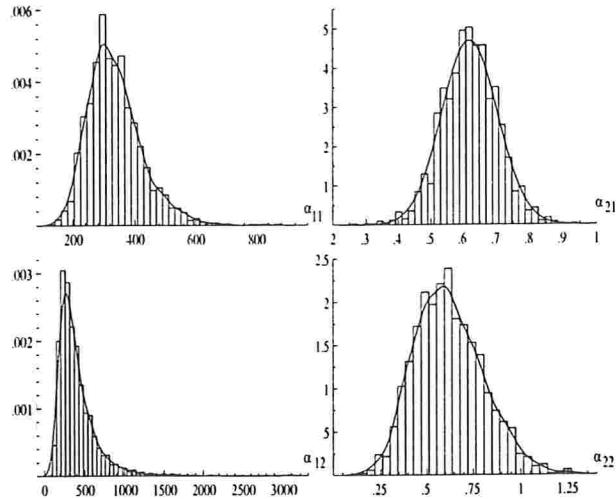


Figura 5.5: Densidades marginais a posteriores aproximados para α_{11} , α_{21} , α_{12} , α_{22} do modelo sob H_0 com priori informativa

Considerando o modelo sob H_1 com a densidade a priori informativa (5.31) com $a_{1j} = 3$, $b_{1j} = 10.000$, $\mu_{0j} = 0$ e $\sigma_{0j} = 100$, $j = 1, 2$, forma gerados 10 cadeias cada um

com 20.000 iterações, considerando-se as iterações $10.100^{th}, 10.200^{th}, \dots, 20.000^{th}$ de cada cadeia, que produz uma amostra de tamanho 1.000. Na tabela 5.7, apresenta-se o resumo das distribuições a posteriori juntamente com o resultado da redução potêncial de escala, para todos os parâmetros e observamos que $\sqrt{\hat{R}} < 1.1$, indicando a convergência e, na figura 5.6, apresentamos a densidades marginais a posteriori aproximadas.

Tabela 5.7: Sumário a posteriori para o modelo sob H_1 com densidade a priori informativa

	Média	Mediana	S.D	Intervalo de credibilidade 95%	\hat{R}
α_{11}	420,01	367,70	224,95	(209,15 ; 1.053,84)	1.002560
α_{21}	0,6760	0,66530	0,17052	(0,4437 ; 1,0715)	1.00532
α_{12}	362,81;	349,80	101,76	(233,86 ; 615,80)	1.01023
α_{22}	0,6386	0,6402	0,0668	(0,4877 ; 0,7960)	1.0070

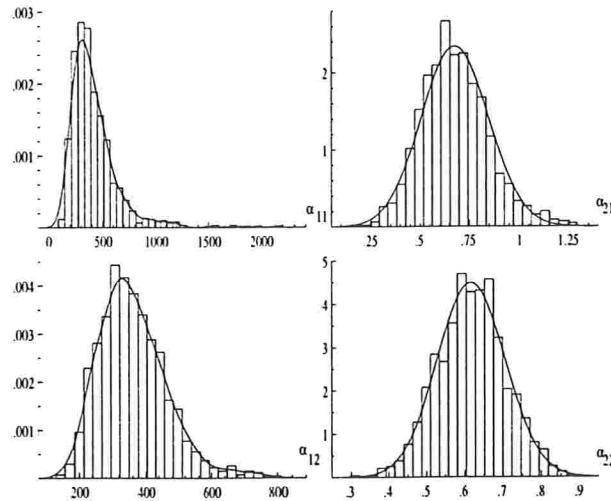


Figura 5.6: Densidades marginais a posteriori aproximados para α_{11} , α_{21} , α_{12} , α_{22} do modelo sob H_1 com priori própria

Com as amostras de Gibbs geradas sob os modelos decorrentes das hipóteses H_0 e H_1 ,

apresenta-se na tabela 5.8, o fator de Bayes com a estimativa da função de verossimilhança preditiva de Laplace-Metropolis, e a proposta por Newton e Raftery(1994). Daí, pode-se observar que o modelo sob H_0 é o que melhor se ajusta aos dados da tabela 5.1. O mesmo resultado foi obtido por Louzada-Neto et al. (1991).

Tabela 5.8: Fator de Bayes para descriminar o modelo sob H_0 vs o modelo sob H_1

Fator de Bayes	Priori DAP	Priori não informativa	Informativa
Laplace-Metropolis	6,3316	1.802,98	56,787
Newton-Raftery	308,8699	345,65	5,743

Nas figuras 5.7 , 5.8 e 5.9 apresenta-se as densidades marginais aproximados para θ_{11} , θ_{12} e $\theta_{11} - \theta_{12}$ do modelo sob H_0 com DAP, priori não informativa e priori informativa, nas tabelas 5.9 , 5.10 e 5.11 apresenta-se o resumo a posterior respectivo.

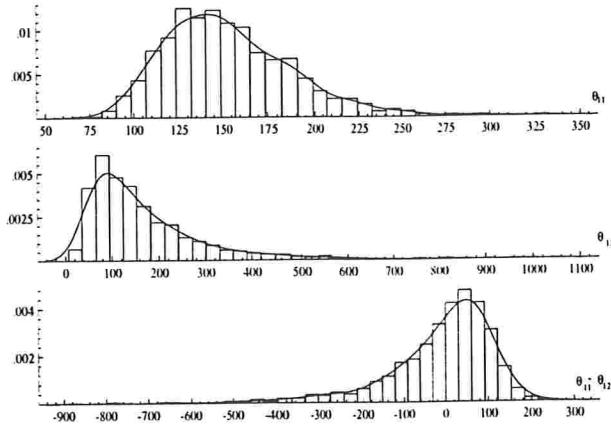


Figura 5.7: Densidades marginais a posterioris aproximados para θ_{11} , θ_{12} e $\theta_{11} - \theta_{12}$ do modelo sob H_0 com DAP

Tabela 5.9: Sumário a posteriori de (5.2) no nível de estresse usual, considerando DAP

	Média	Mediana	S.D	Intervalo de credibilidade 95%
θ_{11}	152,96	147,99	35,12	(97,363 ; 238,449)
θ_{12}	168,48	131,82	127,16	(37,75 ; 520,20)
$\theta_{11} - \theta_{12}$	-15,51	15.01	132,36	(-367,57 : 150.25)

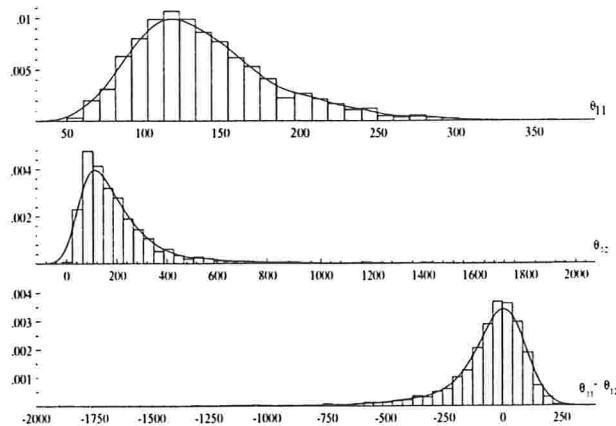


Figura 5.8: Densidades marginais a posteriori aproximados para θ_{11} , θ_{12} e $\theta_{11} - \theta_{12}$ do modelo sob H_0 com densidade a priori não informativa de Jeffreys

Tabela 5.10: Sumário a posteriori de (5.2) com o nível de estresse usual considerando densidade a priori não informativa de Jeffreys

	Média	Mediana	S.D	Intervalo de credibilidade 95%
θ_{11}	139,48	131,67	45,27	(71,60 ; 247,48)
θ_{12}	205,02	156,70	136,80	(43,08 ; 674,48)
$\theta_{11} - \theta_{12}$	-65,55	24,98	184,39	(-558,07 : 155,04)

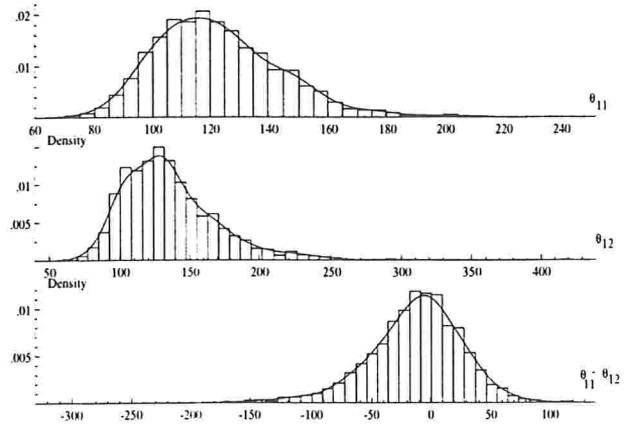


Figura 5.9: Densidades marginais a posteriori aproximados para θ_{11} , θ_{12} e $\theta_{11} - \theta_{12}$ do modelo sob H_0 com densidade a priori informativa

Tabela 5.11: Sumário a posteriori de (5.2) com o nível de estresse usual considerando densidade a priori informativa

	Média	Mediana	S.D	Intervalo de credibilidade 95%
θ_{11}	122,47	119,95	21,43	(87,728;159,778)
θ_{12}	135,58	130,15	33,79	(87,681;198,605)
$\theta_{11} - \theta_{12}$	-13,104	-9,763	39,509	(-104,022;44,671)

Capítulo 6

Modelo de mistura exponencial

6.1 Introdução

Os modelos paramétricos comumente usados na análise de sobrevivência e na análise de confiabilidade assumem alguma família de distribuições, por exemplo Weibull, exponencial, gama, etc, para os tempos de vida. Assim, todos os indivíduos morrem (itens falham) se observamos por um período de tempo suficientemente longo. Essa afirmação parece não ser verdadeira em muitos contextos, já que, não considera a possibilidade de que indivíduos sobreviventes (ou itens que não falham) por longo período sejam observados . Isto induz a considerar modelos de mistura que permitam incluir sobreviventes ou indivíduos que não falham. A idéia aqui é que a distribuição dos tempos de falha somente seja considerada para os indivíduos que eventualmente morrem. Isto é, a distribuição é definida condicionalmente dados os indivíduos que morrem e é, irrelevante para os indivíduos sobreviventes ou indivíduos curados (itens que não falham).

A vantagem dos modelos de mistura com sobreviventes por longo período em relação, aos modelos de sobrevivência comuns é que eles permitem um tipo de heterogeneidade limitada, mas, proveitosa, resultante do fato de que a subpopulação de sobreviventes por longo período tem taxa de falha zero.

Modelos de mistura com sobreviventes por longo período foram adequadamente ajustados para dados de cancer por Boag (1949), Berkson e Gage (1952) e Pocock (1992) com aplicações a dados sobre a curabilidade de cancer do seio e Goldman (1984) para a curabilidade de cancer testicular. Varias generalizações dos modelos propostos por Boag (1949) e Berkson e Gage (1952) foram consideradas por Farewell (1977, 1982 e 1986) para vários estudos biomédicos com alguma discussão geral do uso dos modelos de mistura com sobreviventes por longo periodo de tempo em experimentos de sobrevivência; Greenhouse e Wolfe (1984) utilizam a teoria de risco competitivos. Aplicações para dados de AIDS tem sido considerados por Struthers e Farewell (1989).

Outras aplicações interessantes fora do contexto biomédico têm aparecido em pesquisas de mercado (Ascombe (1961)), em estudos de criminologia (Maltz e McCleary (1977), Maltz (1984), Schimidt e White (1988) e Broadhurst e Maller (1990)) e em estudos de desemprego (Dunsmuir, Tweedie, Flack e Mengersen (1989) e Yamaguchi (1992)). Na teoria de confiabilidade temos aplicações deste modelo em Meeker (1987). Neste Capítulo o objetivo é desenvolver uma análise Bayesiana para o problema através de métodos computacionais, fazendo comparações com técnicas de inferência clássica.

6.2 Análise clássica do modelo de mistura exponencial

Seja b_i , $1 \leq i \leq n$, uma variável binária indicando que o i -ésimo indivíduo de uma população está em risco ou não com relação a algum tipo de falha ($b_i = 1$ se o indivíduo pertence ao grupo de risco e $b_i = 0$ caso contrario). Considera-se que os b_i são independentes e identicamente distribuidos tais que $P\{b_i = 1\} = p \in (0, 1)$ indica a proporção de indivíduos em risco e $P\{b_i = 0\} = 1 - p$ a proporção de indivíduos imunes ou indivíduos curados ou itens que não falham.

As variáveis b_i são não observadas mas, assume-se dado que os indivíduos pertencem

ao grupo de risco, isto é, $b_i = 1$, $1 \leq i \leq n$, os tempos de falha t_i^* são independentes e identicamente distribuidos como distribuição exponencial com média comum $1/\lambda$, $\lambda > 0$, especificado pela densidade

$$g(t; \lambda | b=1) = \lambda e^{-\lambda t}, \quad t \geq 0. \quad (6.1)$$

Associado com o i -ésimo indivíduo está o tempo de censura c_i , $i = 1, \dots, n$. Consideramos que a distribuição dos tempos de censura são não informativos, isto é, ela não envolve (λ, p) , os parâmetros de interesse. Consideramos que t_i^* e c_i são estocasticamente independentes. Quando associado ao indivíduo i , $b_i = 1$, observa-se tempo de censura ou falha $t_i = \min(t_i^*, c_i)$ e o indicador de censura $\delta_i = 1$ se ocorre falha antes do final do estudo e $\delta_i = 0$, caso contrário. Quando $b_i = 0$, $i = 1, \dots, n$ observa-se $t_i = c_i$ e $\delta_i = 0$, neste caso todos os tempos são censurados.

6.2.1 Estimadores de máxima verossimilhança

Seja $\theta = (\lambda, p)' \in \{\theta : \lambda > 0, 0 < p \leq 1\}$, o vetor de parâmetros a ser estimado. Um indivíduo que morreu (iten que falha) no tempo t_i contribui para função de verossimilhança com a parcela

$$P\{b_i = 1\}g(t_i; \lambda | b_i = 1) = p\lambda e^{-\lambda t_i},$$

enquanto, que um indivíduo curado o item que não falhou até o tempo t_i contribui para a função de verossimilhança com a parcela

$$P\{b_i = 0\} + P\{b_i = 1\}P(t_i^* > t_i; \lambda | b_i = 1) = 1 - p + pe^{-\lambda t_i},$$

que é probabilidade do indivíduo ser um sobrevivente por longo período mais a probabilidade de um indivíduo sobreviver pelo menos até o tempo t_i . Portanto, a contribuição do i -ésimo indivíduo para a função de verossimilhança é dada por

$$(p\lambda e^{-\lambda t_i})^{\delta_i} (1 - p + pe^{-\lambda t_i})^{1-\delta_i}.$$

Logo, dada uma amostra $t_1, \dots, t_n, \delta_1, \dots, \delta_n$, a função de verosimilhança para θ é dada por

$$L_n(\theta) = \prod_{i=1}^n \left(p\lambda e^{-\lambda t_i} \right)^{\delta_i} \left(1 - p + pe^{-\lambda t_i} \right)^{1-\delta_i}. \quad (6.2)$$

O estimador de máxima verossimilhança de θ é obtido maximizando-se (6.2), o que equivale a resolver o sistema de equações $S_n(\theta) = \sum_{i=1}^n s_i(\theta) = 0$, com $s_i(\theta) = (s_{i1}(\theta), s_{i2}(\theta))'$, onde

$$s_{i1}(\theta) = \frac{\partial \ell_i(\theta)}{\partial \lambda} = \frac{1}{\lambda} \left\{ \delta_i(1 - \lambda t_i) - (1 - \delta_i) \frac{p\lambda t_i e^{-\lambda t_i}}{1 - p + pe^{-\lambda t_i}} \right\}, \quad (6.3)$$

$$s_{i2}(\theta) = \frac{\partial \ell_i(\theta)}{\partial p} = \frac{1}{p} \left\{ \delta_i - (1 - \delta_i) \frac{p(1 - e^{-\lambda t_i})}{1 - p + pe^{-\lambda t_i}} \right\}, \quad (6.4)$$

sendo $\ell_i = \log(L_i(\theta))$ a i -ésima componente da função log-verossimilhança. Como não é possível obter de modo analítico os valores $\hat{\lambda}$ e \hat{p} que satisfazem as equações acima deve-se, usar algum método iterativo, como por exemplo, o de Newton-Raphson. Esse método iterativo aparece como uma rotina do pacote Ox (veja Doornik 1998).

6.2.2 Inferência sobre os parâmetros λ e p

A inferência sobre os parâmetros são feitas usualmente através da normalidade assintótica dos estimadores de máxima verossimilhança. Assim, para $\hat{\lambda}$ e \hat{p} , EMVs de λ e p , ponto interior do espaço paramétrico e sob certas condições de regularidade (veja por exemplo, Ghitany e Maller, 1992) tem-se o seguinte resultado, para $n \rightarrow \infty$:

$$I_n^{1/2}(\hat{\lambda}, \hat{p}) \left[\begin{pmatrix} \hat{\lambda} \\ \hat{p} \end{pmatrix} - \begin{pmatrix} \lambda \\ p \end{pmatrix} \right] \xrightarrow{D} N(\mathbf{0}, I_2), \quad (6.5)$$

onde $I_n^{1/2}(\lambda, p)$ é a matriz raiz quadrada da matriz de informação de Fisher e $N(\mathbf{0}, I_2)$ denota a distribuição normal bivariada padrão (I_2 é a matriz identidade de dimensão

2×2). Ghitany e Maller (1992) considerando os tempos de censura como sendo variáveis aleatórias independentes e identicamente distribuídas não degeneradas no ponto zero mostraram que os elementos da matriz de informação de Fisher são dados por:

$$\begin{aligned} I_n^{11} &= \frac{np}{\lambda^2} E \left[1 - e^{-\lambda c} - \frac{(1-p)(\lambda c)^2 e^{-\lambda c}}{1-p(1-e^{-\lambda c})} \right], \\ I_n^{22} &= \frac{n}{p} E \left[\frac{1-e^{-\lambda c}}{1-p(1-e^{-\lambda c})} \right], \\ I_n^{12} &= I_n^{21} = \frac{n}{\lambda} E \left[\frac{\lambda c e^{-\lambda c}}{1-p(1-e^{-\lambda c})} \right]. \end{aligned} \quad (6.6)$$

Observe em (6.6), que se as variáveis aleatórias dos tempo de censura convergem quase certamente a ∞ , os elementos da matriz de informação de Fisher se reduzem a

$$I_n^{11} = \frac{np}{\lambda^2}, \quad I_n^{22} = \frac{n}{p(1-p)}, \quad \text{e} \quad I_n^{12} = I_n^{21} = 0, \quad (6.7)$$

de modo que as variâncias assintóticas de EMV de λ e p são dadas por λ^2/np e $p(1-p)/n$, respectivamente.

Suponha que temos interesse em provar que a proporção de indivíduos imunes ou a proporção de indivíduos curados de certa enfermidade é nula, contra a alternativa que não é, isto é,

$$H_0 : p = 1. \quad (6.8)$$

$$H_1 : p < 1.$$

A estatística da razão de verossimilhança para testar H_0 é dada por

$$RV_n = 2 \left(\ell_n(\hat{\lambda}, \hat{p}) - \ell_n(\tilde{\lambda}, 1) \right), \quad (6.9)$$

onde $\hat{\lambda}$ e \hat{p} são EMV irrestritos e $\tilde{\lambda}$ é o EMV sob H_0 dada em (6.8). Sob certas condições de regularidade, Ghitany e Maller (1995) mostraram que a distribuição assintótica de (6.9) é uma mistura 50%-50% da distribuição qui-quadrado com *um* grau de liberdade e um ponto de massa em 0, isto é,

$$P(RV_n \leq x) = \frac{1}{2} + \frac{1}{2}P(\chi^2_{(1)} \leq x). \quad (6.10)$$

6.2.3 Um estudo de simulação

Nesta subseção apresenta-se um estudo de simulação para investigar o comportamento da distribuição assintótica da estatística do teste (6.10) para pequenas amostras. O estudo de simulação é baseado num modelo exponencial com média um para os tempo de vida e a variável c_i (independente dos tempo de vida t_i^*) correspondente ao mecanismo de censura, foi gerada como uma variável aleatória uniforme no intervalo $(0, 5)$. Várias combinações de tamanhos amostrais ($n = 20, 30, 50, 100$) foram consideradas no estudo. Mil réplicas foram geradas para cada simulação. Na tabela 6.1, apresenta-se os os níveis de significância empíricos para os diferentes tamanhos amostrais. Observa-se que os níveis de significância empíricos são quase o dobro dos níveis nominais teóricos o que indica que o teste não é preciso para pequenas amostras.

Tabela 6.1: Percentagem de rejeição de $H_0 : p = 1$ no modelo de sobrevivência com sobreviventes por longo período em 1000 replicações para diferentes tamanhos amostrais

n	10%	5%	1%
20	22.5	11.2	1.7
30	22.3	10.2	1.7
50	20.6	8.5	1.3
100	18.2	8.0	1.2

6.2.4 Comparação de proporções de indivíduos imunes entre grupos

Suponha que existem n indivíduos em um estudo, foram aplicados G tratamentos (grupos) diferentes, com n_g indivíduos para o tratamento g , $g = 1, \dots, G$ e, queremos verificar se existem diferenças nas proporções de indivíduos não imunes nos grupos. Isto é, queremos provar as hipóteses

$$H_0 : 0 < p_1 = \dots = p_g = p < 1, \quad (6.11)$$

$$H_1 : p_i \neq p, \quad \text{para algum, } i = 1, \dots, G,$$

onde $0 < p < 1$.

Githany, Maller e Zhou (1995) sob certas condições de regularidade mostraram que a estatística da razão de verossimilhança para testar (6.11),

$$RV_n = -2 \left[\ell_n(\hat{\theta}) - \ell_n(\tilde{\theta}) \right], \quad (6.12)$$

tem, quando $n \rightarrow \infty$, distribuição qui-quadrado com $G - 1$ graus de liberdade, onde $\hat{\theta}$ é o EMV irrestrito e $\tilde{\theta}$ é o EMV sob H_0 .

Githany et al. (1995) sugerem testar (6.11) em dois estágios, onde num primeiro estágio testa-se se a proporção de imunes nos grupos é nula, isto é,

$$H_0 : p_1 = \dots = p_G = 1, \quad (6.13)$$

$$H_1 : p_i \neq 1, \quad \text{para algum, } i = 1, \dots, G.$$

A estatística de teste para hipótese acima é dada por

$$d_n = 2 \sum_{g=1}^G \hat{\ell}_n(g) \xrightarrow{D} X_1 + \dots + X_G, \quad (6.14)$$

onde $\hat{\ell}_n(g) = \ell_n(\hat{\lambda}_{(g)}, \hat{p}_{(g)}) - \ell_n(\hat{\lambda}_{(g)}, 1)$ e X_g tem a distribuição dada em (6.10). Se aceitamos H_0 em (6.13) concluimos que não há imunes em nenhum grupo e a análise procede

como uma análise de sobrevivência ordinária. Se rejeita-se H_0 então consideramos a hipótese (6.11) como um segundo estágio. A distribuição assintótica da estatística em (6.13) é complicada e uma tabela é apresentada em Maller e Zhou (1996). O comportamento da estatística para n pequeno deixa bastante a desejar como vemos a seguir. Um problema clássico em aberto é testar alternativas restritas com relação aos p_i 's, como por exemplo, $H_1 : p_1 \leq \dots \leq p_G$ com pelo menos uma desigualdade estrita. Propomos uma solução parcial usando o enfoque bayesiano.

6.2.5 Estudo de simulação

Nesta subseção apresentamos um estudo de simulação de Monte Carlo com a finalidade de comparar os níveis de significância empíricos e teóricos para testar H_0 em (6.13). Neste estudo consideramos dois grupos ($G = 2$) com diferentes tamanhos amostrais ($n_1 = n_2 = 20, 30, 50, 100$). As amostras foram geradas a partir de H_0 (6.13); para os tempos de vida, consideramos um modelo exponencial com média um e para os tempos de censura independentemente dos tempos de vida consideramos a variável aleatória uniforme no intervalo (0, 5) para os dois grupos. Na tabela 6.2, apresenta-se os níveis de significância empíricos, onde observa-se que os níveis de significância são praticamente o dobro dos teóricos.

Tabela 6.2: Percentagem de rejeição de $H_0 : p_1 = p_2 = 1$ no modelo de sobrevivência por longo período com 1000 replicações

$n_1 = n_2$	10%	5%	1%
20	23.9	12.2	2.5
30	23.1	11.8	2.1
50	22.3	10.7	2.1
100	20.1	10.3	2.1

6.3 Análise Bayesiana para modelos de mistura exponencial

Como foi discutido na seção anterior a distribuição dos EMVs somente são conhecidas assintóticamente, tornando-se assim um problema sério, uma vez que, em geral, trabalha-se com amostras não muito grandes. Desta forma, inferências (intervalos de confiança, teste de hipóteses, etc.) baseadas nessas estatísticas podem não ser confiáveis, conforme mostram os estudos de simulação desenvolvidos anteriormente.

Um procedimento alternativo que pode contornar esse problema é através da utilização de métodos Bayesianos. Conforme vistos nos capítulos anteriores, o uso de métodos Bayesianos, além de ser uma alternativa de análise, permite a incorporação de conhecimento a priori através de uma densidade a priori que seja informativa. Caso não exista conhecimento a priori, ou seja difícil expressá-lo, considera-se densidades a priori não informativas.

6.3.1 Análise Bayesiana do modelo de mistura considerando a densidade a priori não informativa de Jeffreys

Considerando (6.7), pode-se mostrar que a densidade a priori não informativa de Jeffreys para os parâmetros λ e p é dada por,

$$\pi(\lambda, p) \propto \lambda^{-1}(1-p)^{-1/2}. \quad (6.15)$$

Considerando (6.15) e (6.2) e o teorema de Bayes pode-se mostrar que a densidade a posteriori conjunta para λ e p do modelo de mistura exponencial é dada por

$$\pi(\lambda, p|D) \propto p^r(1-p)^{-1/2}\lambda^{r-1}e^{-R\lambda} \prod_{i=1}^n (1-p+pe^{-\lambda t_i})^{1-\delta_i}, \quad (6.16)$$

onde D é o conjunto de dados observados, $r = \sum_{i=1}^n \delta_i$ e $R = \sum_{i=1}^n \delta_i t_i$.

O teorema 6.3.1 garante que a densidade conjunta a posteriori em (6.16) é própria.

Teorema 6.3.1 *Sob o modelo mistura exponencial, se a densidade a priori é dada por (6.15) então a densidade a posteriori (6.16) é própria.*

Prova: Desde que $(1 - p + pe^{-\lambda t_i}) \leq 1$ para todo p e $i = 1, \dots, n$, temos que

$$\begin{aligned} \int_0^1 \int_0^\infty \pi(\lambda, p|D) d\lambda dp &\propto \int_0^1 \int_0^\infty p^r (1-p)^{-1/2} \lambda^{r-1} e^{-R\lambda} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i} d\lambda dp \\ &\leq \int_0^1 \int_0^\infty p^r (1-p)^{-1/2} \lambda^{r-1} e^{-R\lambda} d\lambda dp \\ &= \frac{\Gamma(r+1)\Gamma(1/2)\Gamma(r)}{\Gamma(r+3/2)R^r} < \infty, \end{aligned}$$

o que mostra que a densidade a posteriori é própria.

Densidades a posteriori marginais pelo método de Laplace

Na inferência Bayesiana as inferências sobre os parâmetros são conduzidas com base nas densidades a posteriori marginais. As marginais para os parâmetros λ e p são obtidas por integração da densidade a posteriori conjunta (6.16) em relação a p e λ , respectivamente. Como em geral essas integrais não permitem soluções explícitas, utilizaremos o método de Laplace (veja, Tierney e Kadane, 1986).

A partir da densidade a posteriori conjunta dada em (6.16), a densidade marginal para λ pode ser escrita na forma:

$$\pi(\lambda|D) \propto \lambda^{r-1} e^{-\lambda R} \int_0^1 p^r (1-p)^{-1/2} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i} dp. \quad (6.17)$$

Portanto, utilizando a aproximação de Laplace em (6.17), obtem-se a densidade a posteriori marginal aproximada para λ dada por

$$\pi(\lambda|D) \propto \sigma \lambda^{r-1} \exp\{-\lambda R - nh(\lambda, \hat{p})\}, \quad (6.18)$$

onde

$$-nh(\lambda, p) = r \log p - \frac{1}{2} \log(1-p) + \sum_{i=1}^n (1-\delta_i) \log(1-p+pe^{-\lambda t_i}), \quad (6.19)$$

$\sigma = \{\frac{\partial^2 h(\lambda, \hat{p})}{\partial p^2}\}^{-1/2}$ e \hat{p} maximiza (6.19) para cada valor de λ fixo.

Pode-se mostrar que a densidade a posteriori marginal aproximada pelo método de Laplace para p é dada por

$$\pi(p|D) \propto \sigma^* p^r (1-p)^{-1/2} \exp\{-nh^*(\hat{\lambda}, p)\}, \quad (6.20)$$

onde

$$-nh^*(\lambda, p) = (r-1) \log \lambda - \lambda R + \sum_{i=1}^n (1-\delta_i) \log(1-p+pe^{-\lambda t_i}), \quad (6.21)$$

$\sigma^* = \{\frac{\partial^2 h^*(\hat{\lambda}, p)}{\partial p^2}\}^{-1/2}$ e $\hat{\lambda}$ maximiza (6.21) para cada valor de p fixo.

Análise Bayesiana usando método MCMC

Um método alternativo ao método de Laplace para a obtenção das densidades marginais de λ e p é dado pela utilização dos métodos de simulação via cadeia de Markov, como o Amostrador de Gibbs e o algoritmo de Metropolis-Hastings que são técnicas para gerar variáveis aleatórias correspondem as distribuições marginais indiretamente, isto é, estes métodos permitem gerar variáveis aleatórias correspondentes a uma distribuição marginal sem ter que calcular esta densidade. Assim, através da simulação de uma amostra da densidade a posteriori suficientemente grande pode-se calcular os momentos de interesse e outras características de interesse desta distribuição.

Considera-se as seguintes densidades a priori para λ e p :

$$\lambda \sim \Gamma(a_1, b_1), \quad \text{onde } a_1 \text{ e } b_1 \text{ são conhecidos,} \quad (6.22)$$

$$p \sim Beta(a_2, b_2), \quad \text{onde } a_2 \text{ e } b_2 \text{ são conhecidos,}$$

onde $\Gamma(a_1, b_1)$ denota a função de distribuição gama com média $\frac{a_1}{b_1}$ e variância $\frac{a_1}{b_1^2}$ e $Beta(a_2, b_2)$ denota a distribuição beta e a_1, b_1, a_2 e b_2 , são constantes conhecidas baseadas

das na informação a priori do especialista e da análise preliminar dos dados. Assumindo independência entre os parâmetros a densidade a priori conjunta é dada por:

$$\pi(\lambda, p) \propto p^{a_2-1} (1-p)^{b_2-1} \lambda^{a_1-1} \exp\{-b_1\lambda\} \quad (6.23)$$

onde $\lambda > 0$ e $0 < p < 1$. Observa-se que se $a_1 = b_1 = 0$, $a_2 = 0$ e $b_2 = 1/2$ obtem-se a densidade a priori não informativa de Jeffreys dada em (6.15). Para $a_1 = b_1 = 0$, $a_2 = b_2 = 1$, tem-se o caso da priori uniforme própria para p , mas, a densidade conjunta a priori dada em (6.23) se reduz a uma densidade a priori não informativa, ou seja,

$$\pi(\lambda, p) \propto \frac{1}{\lambda}, \quad \lambda > 0, \quad 0 \leq p \leq 1. \quad (6.24)$$

Considerando a densidade a priori dada em (6.23) e a função de verossimilhança (6.2), a densidade conjunta a posteriori de λ e p é dada por

$$\pi(\lambda, p|D) \propto p^{r+a_2-1} (1-p)^{b_2-1} \lambda^{r+a_1-1} e^{-(R+b_1)\lambda} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i}, \quad (6.25)$$

onde D denota o conjunto de dados observados.

Consideramos a seguir o caso da densidade a priori própria não informativa para p .

Teorema 6.3.2 *Se a densidade conjunta a priori é dada por (6.24) então a densidade conjunta a posteriori de λ e p , parâmetros do modelo de mistura exponencial, é dada por*

$$\pi(\lambda, p|D) \propto p^r \lambda^{r-1} e^{-R\lambda} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i}, \quad (6.26)$$

que é própria.

A demonstração deste teorema é similar a do teorema 6.3.1.

A partir da densidade conjunta a posteriori de λ e p em (6.25) podemos mostrar que as densidades a posteriori condicionais para o algoritmo de Gibbs são dadas por

$$\pi(p|\lambda, D) \propto p^{r+a_2-1} (1-p)^{b_2-1} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i}, \quad (6.27)$$

$$\pi(\lambda|p, D) \propto \lambda^{r+a_1-1} e^{-\lambda(R+b_1)} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i}. \quad (6.28)$$

Observe que (6.27) e (6.28) não são padrão pelo que faremos uso do algoritmo de Metropolis-Hasting para obter as amostras das densidades marginais de λ e p . O método de Laplace também pode ser usado para se obter densidades marginais aproximadas para λ e p a partir da densidade conjunta (6.25). +

6.3.2 Estudo de simulação

Com a finalidade de comparar a eficiência do estimador de máxima verossimilhança (EMV) e o estimador Bayesiano, um estudo de simulação foi realizado, baseado no modelo de sobrevivência por longo período. Considerou-se a variável aleatória b_i que indica se o indivíduo está em risco ou não, tem distribuição Bernoulli ($1, p$) com ($p = 0.2, 0.4, 0.6, 0.8$); que é a proporção de indivíduos em risco. Para os indivíduos em riscos, simulou-se os tempos de falha, t_i^* de uma distribuição exponencial com média um e para os tempos de censura simulou-se uma variável aleatória uniforme no intervalo (0, 5). O estimador Bayesiano considerado é a moda da densidade a posteriori conjunta de λ e p dada em (6.16) e da densidade a posteriori dada no teorema 6.3.2, denotadas estas modas por B e B_I , respectivamente. Várias combinações de tamanhos amostrais ($n = 30, 50, 100$) são consideradas. Mil réplicas foram geradas para cada simulação. Na tabela 6.3, apresenta-se as médias e o erro quadrático médio (EQM) dos estimadores. Observamos que para n pequeno o estimador de Bayes de λ é melhor, enquanto que os estimadores para p não diferem muito.

Tabela 6.3: Médias e erros quadráticos médios dos estimadores de máxima verossimilhança e dos estimadores Bayesianos, para 1.000 réplicas com diferentes tamanhos amostrais com $\lambda = 1$, do modelo de sobrevivência com sobreviventes por longo período.

			Média		EQM	
p	n		$\hat{\lambda}$	\hat{p}	$\hat{\lambda}$	\hat{p}
0.2	30	MV	2.0746	0.2189	7.0259	0.0068
		B	1.3976	0.2403	1.8663	0.0094
		B_I	1.5611	0.2284	3.6345	0.0081
	50	MV	1.4773	0.2041	0.7541	0.0044
		B	1.1264	0.2128	0.3611	0.0053
		B_I	1.1384	0.2169	0.3613	0.0053
	100	MV	1.1436	0.1993	0.1744	0.0021
		B	0.9945	0.2149	0.1574	0.0028
		B_I	0.9806	0.2118	0.1313	0.0028
0.4	30	MV	1.1453	0.4365	0.3376	0.0182
		B	1.0774	0.4497	0.2286	0.0162
		B_I	1.1469	0.4150	0.2887	0.0109
	50	MV	1.0789	0.4199	0.1432	0.0085
		B	0.9892	0.4349	0.1120	0.0086
		B_I	0.9898	0.4177	0.1170	0.0072
	100	MV	1.0572	0.4090	0.0733	0.0037
		B	0.9966	0.4102	0.0710	0.0044
		B_I	0.9699	0.4172	0.0644	0.0047
0.6	30	MV	1.2301	0.5944	0.1786	0.0108
		B	1.0496	0.6128	0.1208	0.0135
		B_I	1.0944	0.5971	0.1395	0.0113
	50	MV	1.0976	0.6033	0.0805	0.0068
		B	0.9931	0.6295	0.0746	0.0089
		B_I	0.9859	0.6088	0.0723	0.0070
	100	MV	1.0380	0.6036	0.0444	0.0039
		B	0.9897	0.6154	0.0149	0.0047
		B_I	0.9735	0.6110	0.0393	0.0044
0.8	30	MV	1.2375	0.7468	0.1294	0.0101
		B	1.0896	0.7858	0.0820	0.0079
		B_I	1.1406	0.7606	0.1028	0.0078
	50	MV	1.1248	0.7815	0.0572	0.0049
		B	1.0384	0.8080	0.0457	0.0052
		B_I	1.0437	0.7873	0.0463	0.0043
	100	MV	1.0474	0.7982	0.0273	0.0027
		B	1.0027	0.8121	0.0262	0.0032
		B_I	0.9913	0.8057	0.0233	0.0030

6.4 Teste de hipóteses Bayesiano

Na seção 6.2 considerou-se um teste de hipóteses para verificar se a proporção de indivíduos curados de certa enfermidade é nula ou não (veja as hipóteses em (6.8)) e verificou-se também se as proporções de indivíduos curados de alguma enfermidade sobmetidos aos diferentes tratamentos são as mesmas (veja as hipóteses em (6.11)). Nesta seção testa-se estas hipóteses considerando-se uma perspectiva Bayesiana.

Conforme vimos no capítulo 1, uma alternativa Bayesiana para testar as hipóteses dadas em (6.8), (6.11) e (6.13), é através da determinação do fator de Bayes do modelo sob a hipótese nula (modelo M_0) com respeito à hipótese alternativa (modelo M_1). Como foi sugerido por Geisser e Eddy (1979) e enfatizado no capítulo 1, neste trabalho utilizaremos o pseudo-fator de Bayes para testar essas hipóteses, considerando as verossimilhanças preditivas dos modelos M_0 e M_1 , e essas verossimilhanças são obtidas como o produto das densidades preditivas condicionais ordenadas (CPO) definidas por

$$f(t_i|D_{(r)}) = \int f(t_i|\theta, D_{(r)})\pi(\theta|D_{(r)}) \quad (6.29)$$

onde $D_{(r)} = (t_1, \dots, t_{i-1}, t_{i+1}, t_{i+2}, \dots, t_n)$. O pseudo-fator de Bayes do modelo M_0 com respeito ao modelo M_1 é definido por

$$PSFB_{01} = \frac{\prod_{i=1}^n f(t_i|D_{(r)}, M_0)}{\prod_{i=1}^n f(t_i|D_{(r)}, M_1)}, \quad (6.30)$$

onde $f(t_i|D_{(i)}, M_i)$, $i = 0, 1$ é a verossimilhança preditiva do modelo sob H_i , $i = 0, 1$. Como nos capítulos anteriores, se $PSFB_{01} < 1$ consideramos M_1 como o melhor modelo. Portanto, rejeita-se H_0 .

6.4.1 Teste para $H_0 : p = 1$

Nesta subseção testa-se as hipóteses (6.8) da seção 6.2, através de uma perspectiva Bayesiana, utilizando-se o pseudo-fator de Bayes. Para isto determina-se as verossimilhanças preditivas sob as hipóteses H_0 e H_1 dadas em (6.8).

Dado $D = (t_1, \dots, t_n)$, uma amostra aleatória, a função de verossimilhança para $\theta = (\lambda, 1)'$ sob H_0 em (6.8) é dada por

$$L_n(\lambda|M_0) = \lambda^r \exp\left\{-\lambda \sum_{i=1}^n t_i\right\}, \quad (6.31)$$

onde $r = \sum_{i=1}^n \delta_i$. Considera-se a seguinte densidade a priori para λ dada por

$$\lambda \sim \Gamma(a, b) \quad \text{onde } a \text{ e } b \text{ são conhecidos,} \quad (6.32)$$

e $\Gamma(a, b)$ representa a distribuição gama com média a/b e variância a/b^2 .

Usando (6.31), (6.32) e o teorema de Bayes pode-se mostrar que a densidade a posteriori dos parâmetros do modelo M_0 é dada por

$$\lambda|D, M_0 \sim \Gamma(r + a, \sum_{i=1}^n t_i + b). \quad (6.33)$$

A verossimilhança preditiva sob H_0 (modelo M_0) é dada por

$$f(D|M_0) = \frac{\Gamma(r + a)}{\left(\sum_{i=1}^n t_i + b\right)^{r+a}}, \quad (6.34)$$

onde $\Gamma(\cdot)$ é a função gamma.

A função de verossimilhança para $\theta = (\lambda, p)'$ sob H_1 (modelo M_1) é dada em (6.2) e a densidade a posteriori conjunta do modelo M_1 com densidades a priori independentes para λ e p dada em (6.22) é portanto dada por

$$\pi(\lambda, p|D, M_1) \propto p^{r+a_2-1} (1-p)^{b_2-1} \lambda^{r+a_1-1} e^{-(R+b_1)\lambda} \prod_{i=1}^n (1 - p + pe^{-\lambda t_i})^{1-\delta_i}. \quad (6.35)$$

Sob a hipótese H_1 não é possível obter uma expressão analítica para a verossimilhança preditiva como ocorre sob H_0 . Uma estimativa da verossimilhança preditiva sob H_1 (modelo M_1) pode ser obtida por integração de Monte Carlo com amostras a posteriori de λ e p . Isto é, podemos determinar a densidade preditiva de validação cruzada $f(t_r|D_{(r)}, M_1)$.

Vimos que uma estimativa desta densidade é a média harmônica das componentes da função de verossimilhança avaliadas na amostra a posteriori (veja, capítulo 1), isto é,

$$\hat{f}(t_r|D_{(r)}, M_1) = B \left(\sum_{i=1}^B \left[(p^{(i)} \lambda^{(i)} e^{-\lambda^{(i)} t_r})^{\delta_r} (1 - p^{(i)} + p^{(i)} e^{-\lambda^{(i)} t_r})^{1-\delta_r} \right]^{-1} \right)^{-1}, \quad (6.36)$$

onde B é o tamanho da amostra a posteriori de λ e p , obtida utilizando o algoritmo de Gibbs com Metropolis-Hastings. Logo, uma estimativa da verossimilhança preditiva é dada por $\hat{f}(D|M_1) = \prod_{r=1}^n \hat{f}(t_r|D_{(r)}, M_1)$.

Seguindo o procedimento proposto por Gelfand e Dey (1994), uma estimativa alternativa a (6.36) é obtida utilizando-se o método de Laplace. Considere a definição da densidade de validação cruzada dada em (6.29)

$$f(t_r|D_{(r)}, M_1) = \frac{\int L_n(\lambda, p|D, M_1) \pi(\lambda, p|M_1) d\lambda dp}{\int L_{n-1}(\lambda, p|D_{(r)}, M_1) \pi(\lambda, p|M_1) d\lambda dp}, \quad (6.37)$$

onde $L_n(\lambda, p|D, M_1)$ é função de verossimilhança dada em (6.2) para o modelo M_1 dadas n observações, $L_{n-1}(\lambda, p|D, M_1)$ é a função de verossimilhança dada em (6.2) para o modelo M_1 dadas $n-1$ observações (sem a observação t_r) e $\pi(\lambda, p|M_1)$ é a densidade a priori do modelo M_1 .

Usando a aproximação de Laplace no numerador e denominador em (6.37) obtem-se que

$$\hat{f}(t_r|D_{(r)}, M_1) = \left\{ \frac{\Sigma}{\Sigma^*} \right\}^{1/2} \exp\{-(nh^*(\hat{\lambda}, \hat{p}) - nh(\tilde{\lambda}, \tilde{p}))\} \quad (6.38)$$

onde $(\hat{\lambda}, \hat{p})$ é o valor que maximiza

$$-nh^*(\lambda, p) = \log L_n(\lambda, p|D, M_1) + \log \pi(\lambda, p|M_1),$$

$\Sigma^* = \det(nD^2(\hat{\lambda}, \hat{p}))$, sendo $D^2(\hat{\lambda}, \hat{p})$ a matriz hessiana de $h(\lambda, p)$ calculada em $(\hat{\lambda}, \hat{p})$ e $(\tilde{\lambda}, \tilde{p})$ o valor que maximiza

$$-nh(\lambda, p) = \log L_{n-1}(\lambda, p|D, M_1) + \log \pi(\lambda, p|M_1),$$

$\Sigma = \det(nD^2(\tilde{\lambda}, \tilde{p}))$, sendo $D^2(\tilde{\lambda}, \tilde{p})$ a matriz hessiana de $h(\lambda, p)$ avaliada em $(\tilde{\lambda}, \tilde{p})$.

6.4.2 Estudo de simulação

Nesta subseção apresenta-se um estudo de simulação para verificar o comportamento do pseudo fator de Bayes para discriminar entre os modelos sob H_0 e sob H_1 . As verossimilhanças preditivas foram approximadas utilizando-se o método de Laplace e o método MCMC. Para este último método considerou-se 50 amostras para cada tamanho amostral que levaram a rejeição pelo método Laplace. O estudo de simulação é baseado num modelo exponencial com média um para os tempo de vida e a variável aleatória dos tempos de censura c_i são independentes dos tempos de vida t_i^* , tendo distribuição uniforme no intervalo $(0, 5)$. Para o modelo sob H_1 com a priori dada em (6.22), tomamos $a_1 = b_1 = 0$ e $a_2 = b_2 = 1$, enquanto que sob H_0 , para a priori (6.32) tomamos $a = b = 0$. Várias combinações de tamanhos amostrais ($n = 20, 30, 50, 100$) foram consideradas no estudo. Mil amostras foram geradas para cada simulação. Na tabela 6.4, apresenta-se número de vezes que a hipóteses foi rejeitada considerando-se os pseudo-fator de Bayes do modelo sob H_0 com respeito o modelo sob H_1 . Consideramos que H_0 é rejeitado quando $PSBF_{01} < 1$. Por exemplo, com $n = 20$, de 50 amostras apresentando rejeição com o método de Laplace apenas 47 apresentam rejeição com o método MCMC.

Tabela 6.4: Número de vezes que $H_0 : p = 1$ foi rejeitada em 1000 amostras

PSFB	n			
	20	30	50	100
Laplace	142	141	124	98
MCMC	47(50)	46(50)	44(50)	42(50)

6.4.3 Comparação de proporções de sobreviventes

Suponhamos agora que G tratamentos diferentes foram aplicados a n indivíduos de modo que, cada um destes tratamentos foi aplicado a n_g indivíduos escolhidos ale-

toriamente, $g = 1, \dots, G$. Queremos verificar se as proporções de indivíduos que são curados pelos diferentes tratamentos é a mesma, isto é, testar as hipóteses dada em (6.11). Para isto, observou-se os tempos de falha ou censura $t_{11}, \dots, t_{n_11}, \dots, t_{1G}, \dots, t_{n_GG}$, $\delta_{11}, \dots, \delta_{n_11}, \dots, \delta_{1G}, \dots, \delta_{n_GG}$, sendo, $t_{ig} = \min(t_{ig}^*, c_{ig})$, onde, t_{ig}^* denota o tempo de falha do i -ésimo indivíduo em risco, devido ao g -ésimo tratamento (grupo) $g = 1, \dots, G$ e $i = 1, \dots, n_g$, com distribuição exponencial com média $\frac{1}{\lambda_g}$, c_{ig} é o tempo de censura do i -ésimo indivíduo no g -ésimo grupo e δ_{ig} é o indicador de censura ($\delta_{ig} = 1$, se ocorre falha no i -ésimo indivíduo e g -ésimo tratamento (grupo) e $\delta_{ig} = 0$ caso contrário).

Como na seção anterior utiliza-se o fator de Bayes para testar a hipótese dada em (6.11). Para isto, determina-se as estimativas das verossimilhanças preditivas sob a hipótese nula $H_0 : p_1 = \dots, p_G = p$ (modelo M_0) e sob $H_1 : p_j \neq p$ para algum $j = 1 \dots G$ (modelo M_1), utilizando-se amostras a posteriori de cada modelo.

Densidade conjunta a posteriori dos parâmetros do modelo M_1

A função de verossimilhança para $p_1, \dots, p_G, \lambda_1, \dots, \lambda_G$ sob o modelo M_1 considerando que as G amostras são independentes é dada por

$$L(\lambda_1, \dots, \lambda_G, p_1, \dots, p_G) = \prod_{j=1}^G (\lambda_j p_j)^{\delta_{.j}} \exp \left\{ - \sum_{j=1}^G \lambda_j T_j \right\} \times \prod_{j=1}^G \prod_{i=1}^{n_j} \left(1 - p_j + p_j e^{-\lambda_j t_{ij}} \right)^{1-\delta_{ij}}, \quad (6.39)$$

onde $\delta_{.j} = \sum_{i=1}^{n_j} \delta_{ij}$ e $T_j = \sum_{i=1}^{n_j} \delta_{ij} t_{ij}$.

Considere as seguintes densidades a priori independentes para p_j e λ_j

$$\lambda_j \sim \Gamma(c_j, d_j), \quad \text{onde } c_j \text{ e } d_j, j = 1, \dots, G \text{ são conhecidos,} \quad (6.40)$$

$$p_j \sim Beta(e_j, f_j), \quad \text{onde } e_j \text{ e } b_j, j = 1, \dots, G, \text{ conhecidos.}$$

Pelo teorema de Bayes com (6.39) e (6.40), pode-se mostrar que a densidade a posteriori para o modelo M_1 é dada por

$$\begin{aligned}\pi(\underline{\lambda}, \underline{P}|D) &\propto \prod_{j=1}^G \lambda_j^{c_j+\delta_{.j}-1} p_j^{e_j+\delta_{.j}-1} (1-p_j)^{f_j-1} \exp \left\{ -\sum_{j=1}^G \lambda_j(T_j + d_j) \right\} \\ &\times \prod_{j=1}^G \prod_{i=1}^{n_j} \left(1 - p_j + p_j e^{-\lambda_j t_{ij}} \right)^{1-\delta_{ij}}.\end{aligned}\quad (6.41)$$

onde $\underline{\lambda}' = (\lambda_1, \dots, \lambda_g)$ e $\underline{P}' = (p_1, \dots, p_G)$.

Teorema 6.4.1 Se a densidade a priori para $\theta' = (\underline{P}', \underline{\lambda}')$ para o modelo M_1 é dada por (6.40) então a densidade a posteriori conjunta do modelo M_1 é própria.

Prova: Desde que $(1 - p_j + p_j e^{-\lambda_j t_{ij}})^{1-\delta_{ij}} \leq 1$, para todo p_j , $j = 1, \dots, G$ e $i = 1, \dots, n_j$, temos então que

$$\begin{aligned}\int, \dots, \int \pi(\underline{\lambda}, \underline{P}|D) d\underline{\lambda} d\underline{P} &\leq \prod_{j=1}^G \left(\int_0^1 p_j^{e_j+\delta_{.j}-1} (1-p_j)^{f_j-1} dp_j \right) \\ &\times \left(\int_0^\infty \lambda_j^{c_j+\delta_{.j}-1} \exp \{-\lambda_j(T_j + d_j)\} d\lambda_j \right) \\ &= \prod_{i=1}^G \frac{\Gamma(e_j + \delta_{.j}) \Gamma(f_j)}{\Gamma(e_j + \delta_{.j} + f_j)} \times \frac{\Gamma(c_j + \delta_{.j})}{(T_j + d_j)^{c_j+\delta_{.j}}} < \infty.\end{aligned}$$

o que mostra que a densidade a posteriori do modelo M_1 é própria.

A partir da densidade a posteriori conjunta de θ sob o modelo M_1 dado em (6.41) pode-se mostrar que as condicionais completas para a implementação do algoritmo de Gibbs são dadas por

$$\begin{aligned}\pi(\lambda_j | \underline{\lambda}_{(j)}, \underline{P}, D) &\propto \lambda_j^{c_{.j}+\delta_{.j}-1} e^{-\lambda_j(T_j+d_j)} \Psi(\lambda_j, p_j) \\ \pi(p_j | \underline{\lambda}, \underline{P}_{(j)}, D) &\propto p_j^{e_j+\delta_{.j}-1} (1-p_j)^{f_j-1} \Psi(\lambda_j, p_j),\end{aligned}\quad (6.42)$$

onde $\underline{\lambda}_{(j)} = (\lambda_1, \dots, \lambda_{j-1}, \lambda_{j+1}, \dots, \lambda_G)$, $\underline{P}_{(j)} = (p_1, \dots, p_{j-1}, p_{j+1}, \dots, p_G)$ e

$$\Psi(\lambda_j, p_j) = \prod_{i=1}^{n_j} \left(1 - p_j + p_j e^{-\lambda_j t_{ij}} \right)^{1-\delta_{ij}}.$$

Observe que as densidades a posteriori condicionais em (6.42) não possuem uma distribuição padrão, pelo que faremos uso do algoritmo de Metropolis-Hastings com Gibbs para obter as amostras a posteriori.

A estimativa da verossimilhança preditiva sob o modelo M_1 será obtida como o produto das densidades preditivas de validação cruzada e, a estimativa dessa densidade com as amostras a posteriori dos parâmetros sob o modelo M_1 é dada por

$$\hat{f}(t_{rj}|D_{(rj)}, M_1) = \begin{cases} B \left(\sum_{k=1}^B [p_j^{(k)} \lambda_j^{(k)} \exp(-\lambda_j^{(k)} t_{rj})]^{-1} \right)^{-1}, & \text{se } \delta_{rj} = 1 \\ B \left(\sum_{k=1}^B [1 - p_j^{(k)} (1 + \exp(-\lambda_j^{(k)} t_{rj}))]^{-1} \right)^{-1}, & \text{se } \delta_{rj} = 0, \end{cases} \quad (6.43)$$

onde $(\lambda_j^{(k)}, p_j^{(k)})$, $k = 1, \dots, B$, é amostra a posteriori de λ_j ; p_j ; B é o tamanho da amostra no modelo M_1 e $D_{(rj)}$ denota o conjunto de observações sem a observação correspondente ao indivíduo r do grupo j , $j = 1, \dots, G$ e $r = 1, \dots, n_j$.

Densidade conjunta a posteriori para os parâmetros sob o modelo M_0

A função de verossimilhança para $(\lambda_1, \dots, \lambda_G, p)$ sob H_0 (modelo M_0) é dada por

$$L(\underline{\lambda}, p|D) \propto p^r \prod_{j=1}^G \lambda_j^{\delta_{rj}} \exp \left\{ - \sum_{i=1}^G \lambda_j T_i \right\} \prod_{j=1}^G \prod_{i=1}^{n_j} (1 - p + pe^{\lambda_j t_{ij}})^{1-\delta_{ij}}, \quad (6.44)$$

onde $r_1 = \sum_{j=1}^G \sum_{i=1}^{n_j} \delta_{ij}$. Para os parâmetros sob o modelo M_0 consideramos a seguintes densidades a priori independentes:

$$\lambda_j \sim \Gamma(g_j, h_j), \text{ onde } g_j \text{ e } h_j \text{ são conhecidos } j = 1, \dots, G; \quad (6.45)$$

$$p \sim \text{Beta}(e, w), \text{ onde } e \text{ e } w \text{ são conhecidos.}$$

Considerando (6.44) e (6.45) pode-se mostrar que a densidade conjunta a posteriori sob o modelo M_0 é dada por

$$\begin{aligned} \pi(\underline{\lambda}, p|D) &\propto \prod_{j=1}^G \lambda_j^{g_j + \delta_{rj} - 1} p^{e + r_1 - 1} (1 - p)^{w - 1} \exp \left\{ - \sum_{j=1}^G \lambda_j (T_j + h_j) \right\} \\ &\times \prod_{j=1}^G \prod_{i=1}^{n_j} (1 - p(1 - e^{\lambda_j t_{ij}}))^{1-\delta_{ij}}. \end{aligned} \quad (6.46)$$

Teorema 6.4.2 No modelo sob H_0 dada em (6.11) (modelo M_0), se a densidade a priori para $(\underline{\lambda}, p)$ é dada em (6.45) então a densidade a posteriori dada em (6.46) é própria.

Prova: Já que $(1 - p + pe^{\lambda_j t_{ij}})^{1-\delta_{ij}} \leq 1$ para todo $j = 1, \dots, G$ e $i = 1, \dots, n_j$, temos então que

$$\begin{aligned} \int \dots \int \pi(\underline{\lambda}, p | D) d\underline{\lambda} dp &\leq \int_0^1 p^{e+r_1-1} (1-p)^{w-1} dp \prod_{j=1}^G \left(\int_0^\infty \lambda_j^{g_j+\delta_j-1} e^{-\lambda_j(T_j+h_j)} d\lambda_j \right) \\ &= \frac{\Gamma(e+r_1)\Gamma(w)}{\Gamma(e+r_1+w)} \prod_{j=1}^G \frac{\Gamma(g_j+\delta_j)}{[T_j+h_j]^{g_j+\delta_j}} < \infty. \end{aligned}$$

O teorema 6.4.2 segue sendo válido no caso que a densidade a priori para o modelo M_0 é dada por $\pi(\underline{\lambda}, p) \propto 1 / \prod_{j=1}^G \lambda_j$. Isto é, densidades a priori não informativas independentes.

A partir da densidade conjunta a posteriori do modelo M_0 dada em (6.46), as condicionais completas para o algoritmo de Gibbs são dadas por

$$\begin{aligned} \pi(\lambda_j | \underline{\lambda}_{(j)}, p, D) &\propto \lambda_j^{g_j+\delta_j-1} \exp\{-\lambda_j(T_j+h_j)\} \prod_{i=1}^{n_j} (1 - p(1 - e^{-\lambda_j t_{ij}}))^{1-\delta_{ij}} \\ \pi(p | \underline{\lambda}, D) &\propto p^{e+r_1-1} (1-p)^{w-1} \prod_{j=1}^G \prod_{i=1}^{n_j} (1 - p(1 - e^{-\lambda_j t_{ij}}))^{1-\delta_{ij}}. \end{aligned} \quad (6.47)$$

Como as condicionais a posteriori dadas em (6.47) não são distribuições conhecidas utilizaremos a algoritmo de Metropolis-Hasting para obter as amostras a posteriori.

A estimativa da verossimilhança preditiva sob o modelo M_0 (modelo sob H_0 dada em (6.11)) será obtida como o produto da estimativas das densidades preditivas de validação cruzada ou CPO, dadas em (8.24). A estimativa dessas densidades com as amostras Gibbs dos parâmetros sob o modelo M_0 são dadas por

$$\hat{f}(t_{rj} | D_{(rj)}, M_0) = \begin{cases} B \left(\sum_{k=1}^B [p^{(k)} \lambda_j^{(k)} \exp(-\lambda_j^{(k)} t_{rj})]^{-1} \right)^{-1}, & \text{se } \delta_{rj} = 1 \\ B \left(\sum_{k=1}^B [1 - p^{(k)} (1 + \exp(-\lambda_j^{(k)} t_{rj}))]^{-1} \right)^{-1}, & \text{se } \delta_{rj} = 0, \end{cases} \quad (6.48)$$

onde $(\lambda_j^{(k)}, p^{(k)})$ é a $k-$ ésima amostra Gibbs de λ_j , p no modelo M_1 , $k = 1, \dots, B$, $j = 1, \dots, G$ e B é o tamanho da amostra.

Com (6.43) e (6.48) pode-se calcular o pseudo-fator de Bayes do modelo M_0 com respeito ao modelo M_1 , que denotaremos por $PSFB_{01}$ e rejeita-se H_0 se $PSFB_{01} < 1$.

6.4.4 Teste para $H_0 : p_1 = \dots = p_G = 1$

Nesta subseção testa-se as hipóteses dadas em (6.13) usando a perpectiva Bayesiana, isto é, queremos determinar o fator de Bayes do modelo sob H_0 (modelo M_0) com respeito ao modelo sob H_1 (modelo M_1).

A função de verossimilhança para $(\lambda_1, \dots, \lambda_G)$ sob o modelo M_0 é dada por

$$L(\lambda_1, \dots, \lambda_G) \propto \prod_{j=1}^G \lambda_j^{\delta_{.j}} \exp \left\{ - \sum_{j=1}^G \lambda_j t_{.j} \right\}, \quad (6.49)$$

onde $t_{.j} = \sum_{i=1}^{n_j} t_{ij}$. Observe que (6.49) é a função de verossimilhança correspondente a G amostras independentes do modelo de sobrevivência exponencial com média λ_j , $j = 1, \dots, G$.

Vamos a considerar as seguintes densidades a priori independentes para $\lambda_1, \dots, \lambda_G$:

$$\lambda_j \sim \Gamma(l_j, k_j), \text{ onde } l_j \text{ e } k_j \text{ são conhecidos.} \quad (6.50)$$

Considerando (6.49) e (6.50), pode-se mostrar que a densidade conjuta a posteriori para o modelo M_0 é dada por

$$\pi(\lambda_1, \dots, \lambda_G | D) \propto \prod_{j=1}^G \lambda_j^{l_j + \delta_{.j} - 1} \exp \left\{ - \sum_{j=1}^G \lambda_j (t_{.j} + k_j) \right\}. \quad (6.51)$$

Observe que a densidade a posteriori dada em (6.51) é o produto de distribuições gamas e daí pode-se mostrar que a verossimilhança preditiva sob o modelo M_0 é dada por

$$f(D|M_0) = \prod_{i=1}^G \frac{\Gamma(\delta_{.j} + c_j)}{(t_{.j} + d_j)^{\delta_{.j} + c_j}}, \quad (6.52)$$

onde $\Gamma(\cdot)$ denota a função gama.

A densidade a posteriori sob a hipóteses alternativa H_1 dada em (6.13) é similar à densidade dada em (6.41) se consideramos em nossa hipótese $H_1 : p_g \neq 1$, para $g = 1, \dots, G$. Logo, a verossimilhança preditiva para este modelo pode ser obtida como em (6.43).

Estudo de simulação

Nesta subseção apresenta-se um estudo de simulação para verificar o comportamento do pseudo fator de Bayes para discriminar entre os modelos sob $H_0 : p_1 = p_2 = 1$ (modelo M_0) e sob $H_1 : p_g \neq 1$ (modelo M_1), $g = 1, 2$, como dado em (6.13). A aproximação das verossimilhanças preditivas foi obtida utilizando-se os métodos de Laplace e MCMC. Para este último método considerou-se 50 réplicas para cada tamanho amostral que apresentaram rejeição pelo método de Laplace. O estudo de simulação é baseado em dois modelos exponenciais com média um para os tempos de vida e a variável aleatória dos tempos de censura c_i são independentes dos tempo de vida t_i^* tendo sido gerados como uma variável aleatória uniforme no intervalo $(0, 5)$. Considerou-se para os parâmetros da densidade a priori para o modelo M_1 dado em (6.40), com $c_j = d_j = 0$ e $e_j = f_j = 1$, $j = 1, 2$ e a densidade a priori para o modelo M_0 dado (6.50) com $l_j = k_j = 0$, $j = 1, 2$. Várias combinações de tamanhos amostrais ($n_1 = n_2 = n = 20, 30, 50, 100$) foram consideradas no estudo. Mil réplicas foram geradas para cada simulação. Na tabela 6.5, apresenta-se o número de vezes que a hipótese foi rejeitada. Rejeita-se H_0 se o pseudo-fator de Bayes do modelo M_0 (sob H_0) com respeito o modelo M_1 (sob H_1) é menor que um . Para $n = 20$, por exemplo, das 50 amostras onde se rejeita H_0 com o método de Laplace, rejeita-se apenas 45 com o método MCMC.

Tabela 6.5: Número de vezes que $H_0 : p_1 = p_2 = 1$ foi rejeitada em 1000 réplicas

	n			
	20	30	50	100
PSFB				
Laplace	159	142	138	102
MCMC	45(50)	44(50)	43(50)	47(50)

6.5 Exemplo ilustrativo

Nesta seção apresentamos um exemplo como aplicação dos procedimentos clássicos e Bayesianos, discutidos nas seções anteriores. O exemplo consiste em analisar os dados (veja tabela 6.6) obtidos por Kersey et al. (1987), os quais se referem aos tempos de recaída de pacientes de leucemia submetidos a dois tipos de transplante, "allogeneic" (Grupo 1) ou "autologous" (Grupo 2). O Grupo 1, consiste de observações de 46 pacientes, dos quais 13 correspondem a observações censuradas, enquanto, que o Grupo 2, consiste de observações de 44 pacientes dos quais 9 correspondem a observações censuradas.

6.5.1 Análise clássica

Nesta subseção apresenta-se uma aplicação da metodologia discutida na seção 2. Considerando-se apenas os dados do Grupo 1 da tabela 6.6, são apresentados na tabela 6.7 as estimativas de máxima verossimilhança e os intervalos de confiança de 95% para λ e p .

A seguir verifica-se a validade da hipótese dada em (6.8), isto é, testa-se se existem indivíduos que não apresentam recaída após o transplante "allogeneic". O estimador de máxima verossimilhança sob H_0 em (6.8) é $\tilde{\lambda} = 0,4728$, e a estatística da razão de verossimilhança dada em (6.9) resultou em $RV_n = 22,979$. O percentil de 99% denotado por $c_{0,99}$ da distribuição dada em (6.10) é tal que

$$\frac{1}{2} + \frac{1}{2}P(\chi^2_{(1)} \leq c_{0,99}) = 0,99,$$

Tabela 6.6: Tempos de recaída de pacientes com leucemia

Grupo 1				Grupo 2			
t_i	δ_i	t_i	δ_i	t_i	δ_i	t_i	δ_i
0.0301	1	0.9096	1	0.0575	1	0.3589	1
0.0484	1	0.9644	1	0.1096	1	0.4027	1
0.0630	1	1.0082	1	0.1370	1	0.4685	1
0.0849	1	1.2822	1	0.1452	1	0.4712	1
0.0877	1	1.3452	1	0.1479	1	0.4904	1
0.1397	1	1.4000	1	0.1534	1	0.5178	1
0.1616	1	1.7205	0	0.1753	1	0.5452	1
0.1699	1	1.9890	0	0.1836	1	0.5836	1
0.2137 ²	1	2.2438	1	0.2000	1	0.6110	1
0.2164	1	2.5068	0	0.2082	1	0.6137	1
0.2384	1	2.6466	0	0.2164	1	0.7589	1
0.2712	1	3.0384	0	0.2219	1	1.9836	0
0.2740	1	3.1726	0	0.2411	1	1.9973	0
0.3863	1	3.4411	1	0.2603	1	2.0110	1
0.4384	1	4.4219	0	0.2685 ²	1	2.8849	0
0.4548	1	4.4356	0	0.2712	1	2.9973	0
0.5918	1	4.5863	0	0.2849	1	3.2658	0
0.6000	1	4.6904	0	0.2877	1	4.0411	0
0.6348	1	4.7808	0	0.2904	1	4.2055 ²	0
0.6849	1	4.9863	0	0.3068	1	5.0548	0
0.7397	1	5.0000	0				
0.8575	1						

Tabela 6.7: Estimativas de máxima verossimilhança do modelo de mistura exponencial

	EMV	D.P	Intervalo de confiança 95%
$\hat{\lambda}$	1,4332	0,2756	(0,8931 ; 1,9733))
\hat{p}	0,7289	0,0675	(0,5966 ; 0,8613)

de onde tem-se que $c_{0,99} = 5,41$. Já que, $22,979 > 5,41$ rejeita-se H_0 ao nível de 1% de confiança, e portanto, há uma forte evidência de que $p < 1$.

Comparação de proporções de sobreviventes

Com os dados da tabela 6.6, verifica-se a validade das hipóteses dadas em (6.13), isto é, testa-se se existem indivíduos que não apresentam recaída após o transplante, nos 2 grupos. Os EMV sob H_0 em (6.13) são $\tilde{\lambda}_1 = 0,47284$ e $\tilde{\lambda}_2 = 0,8026$ e os EMV sob H_1 em (6.13) é $\hat{\lambda}_1 = 1,4332$, $\hat{p}_1 = 0,728893$, $\hat{\lambda}_2 = 2,681$ e $\hat{p}_2 = 0,79638$. A estatística da razão de verossimilhança dada em (6.14), resultou em $d_n = 62,604$. O percentil de 95% da distribuição assintótica (6.14) é $C_{0,95} = 4,23$. Como $62,604 > 4,23$, rejeita-se H_0 ao nível de confiança de 5% e portanto, há forte evidência de que pelos menos um $p_g < 1$, $g = 1,2$.

Como a hipótese H_0 dada em (6.13) foi rejeitada verifica-se a seguir se a porcentagem de indivíduos que apresentam recaída após os transplante é a mesma em ambos grupos, isto é, testa-se as hipóteses dada em (6.11). O EMV sob $H_0 : p_1 = p_2 = p$ são $\tilde{\lambda}_1 = 1,41$, $\tilde{\lambda}_2 = 2,684$ e $\check{p} = 0,76306$. A estatística da razão de verossimilhança dada em (6.12) resultou em $RV_n = 0,546$. O percentil de 95% da distribuição qui-quadrado com *um* grau de liberdade é $\chi^2_1 = 3,084$. Já que $0,546 < 3,084$ não rejeita-se H_0 ao nível de confiança de 5%.

6.5.2 Análise Bayesiana

Nesta subseção apresenta-se uma aplicação da metodologia Bayesiana apresentada na seção 6.3.

Densidade a posteriori marginal dos parâmetros do modelo de mistura pelo método de Laplace

Considerando os dados do Grupo 1 da tabela 6.6, apresenta-se na figura 6.1, as densidades a posteriori marginais dos parâmetros λ e p dadas em (6.18) e (6.20), aproximadas pelo método de Laplace para o modelo de sobrevivência com sobreviventes de longo período.

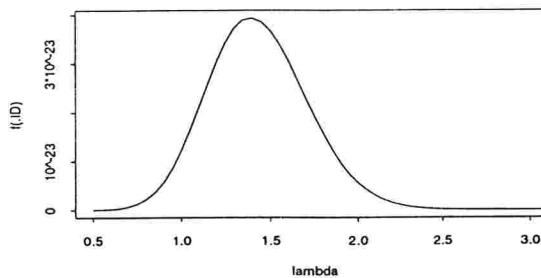
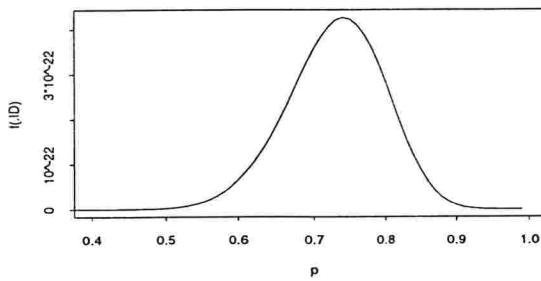


Figura 6.1: Densidades marginais a posteriori de p e λ para o modelo de mistura exponencial obtidas pelo método de Laplace

A média e desvio padrão a posteriori para p e λ aproximados pelo método de Laplace resultaram respectivamente em $E(p|D) = 0,73033$, $D.P.(p|D) = 0,0665$, $E(\lambda|D) = 1,4151$ e $D.P.(\lambda|D) = 0,2776$.

Método MCMC

Considerando os dados do Grupo 1 da tabela 6.6, para a inferência Bayesiana pelo método MCMC, consideramos as densidades a priori dada em (6.22) com $a_1 = 1$, $b_1 = 0,000002$ e $a_2 = b_2 = 0,5$. Observe que a densidade a priori é não informativa. Para se obter amostras da densidade a posteriori através do algoritmo de Gibbs com Metropolis-Hasting procede-se da seguinte forma. Geram-se 10 cadeias paralelas cada um com 20.000 iterações. Depreza-se as 10.000 primeiras iterações, com objetivo de diminuir o efeito dos pontos iniciais e a partir de daí escolhem-se as iterações de 10 em 10, ou seja, para cada parâmetro em cada cadeia considera-se as iterações 10.010 –ésima, ..., 20.000 –ésima o qual fornece uma amostra de tamanho 20000. Para monitorar a convergência da cadeia utiliza-se a técnica de Gelman e Rubin(1992), método este que utiliza a técnica de análise de variância para verificar se mais iterações são necessárias.

Na tabela 6.8, apresenta-se o resumo das distribuições a posteriori (6.25) e na figura 6.2, representa-se a densidade aproximada dos parâmetros do modelo de mistura exponencial. Também, na tabela 6.8, tem-se o resultado da redução de escala estimada da estatística \hat{R} (veja, Gelman e Rubin, 1992) para todos os parâmetros. O número de iterações considerado foi suficiente para se chegar a convergência, pois $\sqrt{\hat{R}} < 1,1$ para todos os parâmetros.

Tabela 6.8: Sumário a posteriori dos parâmetros do modelo de mistura exponencial

	Média	Mediana	D.P.	Intervalo de Credibilidade 95%	\hat{R}
λ	1,4328	1,4114	0,2780	(0,9312 ; 2,0222)	1,000044
p	0,7105	0,7129	0,06865	(0,5922;0,8393)	1,000186

Da análise Bayesiana anterior pode-se concluir que não há muita diferença entre os métodos Laplace e MCMC, pois, as densidades marginais são muitos similares mesmo que as densidades a priori para os dois métodos sejam diferentes. Observe também, que as estimativas pelo método de máxima verossimilhança são muito próximas a aquelas

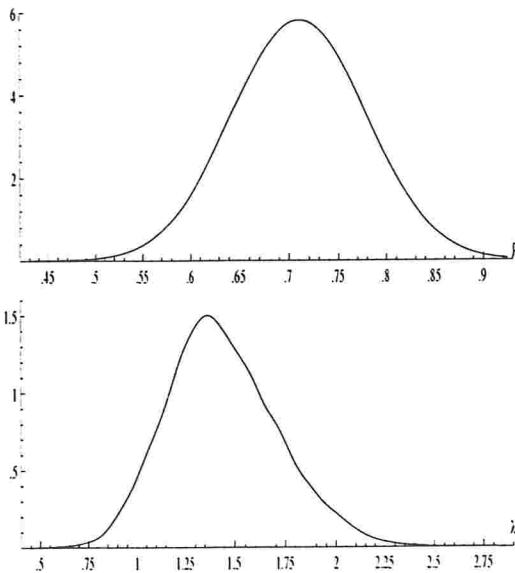


Figura 6.2: Densidades marginais a posteriori aproximadas de p e λ para o modelo de mistura exponencial pelo método MCMC

obtidas pela metodologia Bayesiana.

A seguir, vamos a testar as hipóteses dadas em (6.8), utilizando-se o pseudo-fator de Bayes. A verossimilhança preditiva sob H_0 (modelo M_0) dada em (6.34) considerando-se a densidade a priori (6.32) com $a = 1$ e $b = 0.00002$ é $f(D|M_0) = 1,776 \times 10^{-26}$. A estimativa da verossimilhança preditiva sob H_1 (modelo M_1) com as amostras a posteriori de λ e p , aproximado pela densidade de validação cruzada dada em (6.36) é $\hat{f}(D|M_1) = 8,637 \times 10^{-22}$. O pseudo-fator de Bayes para comparar os modelos sob M_0 e M_1 é $PSFB_{01} = 0,0000206$. Logo, para os dados do Grupo 1, da tabela 6.6 o modelo M_1 se ajusta melhor. Portanto, rejeita-se H_0 , confirmando a conclusão obtida pelo procedimento clássico. Observamos também que o fator de Bayes enfaticamente recomenda o modelo M_1 .

Comparação de proporções de sobreviventes

Nesta subseção verifica-se inicialmente as hipóteses dada em (6.13), utilizando-se o pseudo fator de Bayes do modelo sob H_0 (modelo M_0) com respeito ao modelo sob H_1 (modelo M_1). Se a hipótese H_0 em (6.13) é rejeitada, testa-se as hipóteses dadas em (6.11), isto é, que a proporção de indivíduos que apresentam recaída é a mesma nos dois grupos.

Para o modelo sob H_0 dado em (6.13) consideramos as densidades a priori dada em (6.50) com $l_g = 1$ e $k_g = 0,00001$, $g = 1, 2$. Neste caso a verossimilhança preditiva sob H_0 (modelo M_0) dada em (6.52) resultou em $f(D|M_0) = 1,739 \times 10^{-45}$.

Consideramos agora o modelo sob H_1 dada em (6.13) e as densidades a priori dadas em (6.40) com $c_g = 1$ e $d_g = 0,00001$, $g = 1, 2$. Geramos 20 cadeias separadas cada uma com 5.000 iterações. Deprezam-se as 2.500 primeiras iterações, e daí escolhem-se as iterações de 10 em 10, fornecendo uma amostra de Gibbs de tamanho 5.000. Na tabela 6.9, apresenta-se o resumo das distribuições a posteriori juntamente com a estimativa dos fatores de redução de escala potencial \hat{R} para todos os parâmetros que são praticamente iguais a um e na figura 6.3, representa-se as densidades marginais aproximadas para o modelo.

Tabela 6.9: Sumário a posteriori dos parâmetros do modelo sob H_1 dada em (6.13)

	Média	Mediana	D.P	Intervalo de Credibilidade 95%	\hat{R}
λ_1	1,5664	1,5176	0,3858	(0,98696 ;2,38594)	1,0081346
p_1	0,72420	0,72736	0,07031	(0,57904 ;0,854664)	1,0000234
λ_2	2,8132	2,7552	0,6077	(1,82204 ;4,18017)	1,0002838
p_2	0,79729	0,80109	0,05643	(0,676871; 0,893852)	1,0178026

A estimativa da verossimilhança preditiva sob a hipótese H_1 (Modelo M_1) dada em (6.13) resultou em, $\hat{f}(D|M_1) = 1,683 \times 10^{-35}$. O pseudo-fator de Bayes do modelo sob H_0 em (6.13) com respeito ao modelo sob H_1 em (6.13) é $PSFB_{01} = 1,033 \times 10^{-10} < 1$. Logo, rejeita-se H_0 . A mesma conclusão foi obtida com o procedimento clássico, embora

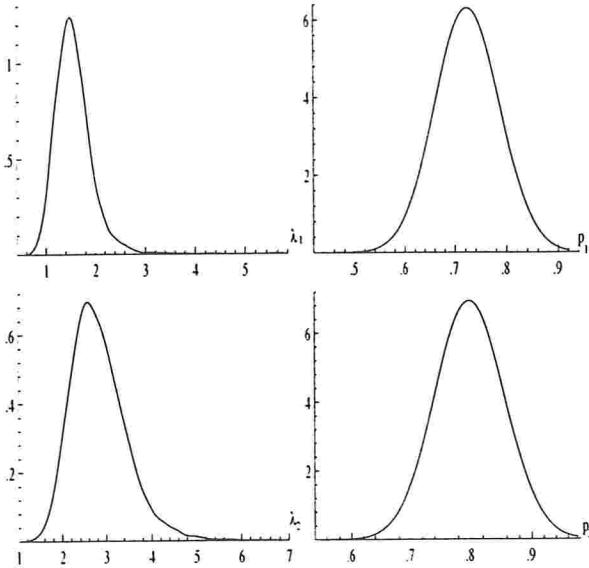


Figura 6.3: Densidades marginais a posteriori aproximadas de λ_1 , p_1 , λ_2 e p_2 do modelo sob H_1 em (6.13)

o pseudo-fator de Bayes parece mostrar que a diferença é mais acentuada.

A seguir vamos a testar as hipóteses dada em (6.11) de uma perspectiva Bayesiana. Considerando o modelo sob H_0 dada em (6.11) e as densidades a priori dadas em (6.45) com $g_j = 1$, $h_j = 0,00001$, $j = 1, 2$, e $e = f = 1$, geramos 20 cadeias separadas com 5.000 iterações cada uma para cada parâmetro. Desprezam-se as 2.500 primeiras iterações e a partir daí escolhem-se as iterações de 10 em 10, fornecendo uma amostra de tamanho 5.000. Na tabela 6.10, apresentamos o sumário da distribuição a posteriori juntamente com o resultado do fator de redução escala potencial estimada, para todos os parâmetros e observamos que $\sqrt{\hat{R}} < 1.1$ indicando a convergência e na figura 6.4, temos as densidades marginais a posteriori aproximadas.

Com as amostras de Gibbs geradas, as estimativas da verossimilhança preditiva para o modelo sob H_0 e H_1 em (6.11) são, $\hat{f}(D|M_0) = 3,363 \times 10^{-35}$ e $\hat{f}(D|M_1) = 1,683 \times 10^{-35}$. O pseudo-fator de Bayes do modelo sob H_0 com respeito ao modelo sob H_1 é $PSFB_{01} =$

Tabela 6.10: Sumário a posteriori dos parâmetros do modelo sob H_1 dada em (6.11)

	Média	Mediana	D.P	Intervalo de Credibilidade 95%	\hat{R}
λ_1	1,5463	1,5036	0.3791	(0,97685;2,36095)	1,0063249
λ_2	2,7600	2,7257	0.4874	(1,89705;3,78582)	1,0000291
p	0,75998	0,76184	0.0458	(0,663455;0,84437)	1,0032145

$1,9982 > 1$. Portanto, H_0 não é rejeitada, confirmando assim a conclusão obtida pelo método clássico.

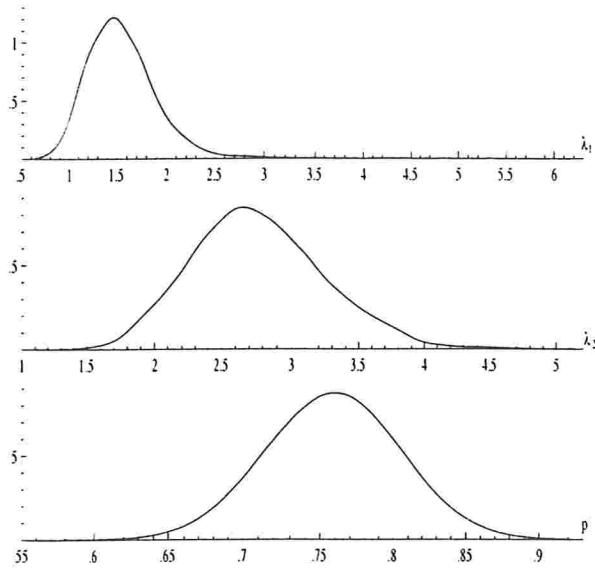


Figura 6.4: Densidades marginais a posteriori aproximadas de λ_1 , λ_2 e p do modelo sob H_0 em (6.11)

Teste para $H_0 : p_1 = p_2$ versus $H_1 : p_1 > p_2$

No caso de dois grupos ($G = 2$), muitas vezes tem-se interesse em verificar as hipóteses seguintes:

$$H_0 : p_1 = p_2 \text{ versus } H_1 : p_1 > p_2.$$

Enfatizamos que não existe procedimento clássico disponível para testar as hipóteses H_0 e H_1 . Um procedimento usual em neste caso, é obtermos inferência sobre $p_1 - p_2$ (ver por exemplo Box e Tiao, 1973). A função de densidade a posteriori de $p_1 - p_2$ pode ser aproximada utilizando as amostras Gibbs do modelo sob H_1 em (6.11) e essa densidade é representada na figura 6.5. A probabilidade a posteriori de que $p_1 - p_2 > 0$ é aproximadamente igual a 0,2039 dando um forte indício de que H_0 é verdadeira. Hsu (1982) considera este enfoque em problemas de deteção de pontos de mudança. Usando o amostrador de Gibbs fica fácil implementar este procedimento, pois temos disponível amostras independentes das duas distribuições. Quando temos mais de duas proporções para comparar, poderia ser estudada o comportamento destas duas a duas, e para conclusão final poderia ser utilizado um procedimento do tipo Bonferroni.

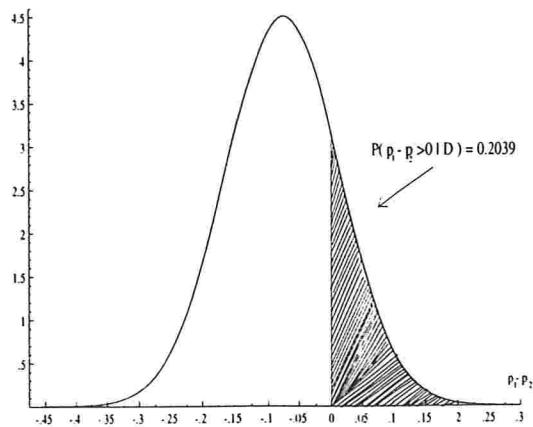


Figura 6.5: Densidade marginal a posteriori aproximado de $p_1 - p_2$

Capítulo 7

Modelos de mistura exponencial com covariáveis

7.1 Introdução

A análise de dados relacionados com os tempos de falha, onde a informação concomitante sobre os indivíduos está disponível na forma de vetores de covariáveis associadas com cada um, é de grande utilidade em muitas disciplinas tais como da medicina, engenharia e criminologia. Os métodos de covariáveis foram examinados extensivamente no contextos dos modelos de sobrevivência paramétrica para os quais a distribuição de tempos de sobrevivência depende do um vetor de covariáveis associado a cada indivíduo. Veja Feigel e Zelen (1965), Glasser (1967) e Maller (1988) para abordagens que acomodam censuras e covariáveis ao modelo de sobrevivência exponencial comum. Ghitany e Maller (1994) extenderam para o caso de covariáveis os modelos de mistura exponencial considerados por Ghitany e Maller (1992) em que a população em estudo tem uma distribuição de tempos de falha exponencial mas também contém uma proporção de imunes que são imortais no sentido de que eles não estão sujeitos a falhas. Atualmente, tais modelos de mistura com imunes e covariáveis estão em uso em muitas áreas, tais como medicina e criminologia.

Veja, por exemplo Farewell (1982-1986), Goldman (1984), Broadhurst e Maller (1990). Análise assintótica rigorosa não é geralmente tratada nesses estudos, mas se mencionando elas, por exemplo, frequentemente assumem tipos simples de censura e covariáveis limitadas. Estas restrições não são desejáveis, nem necessárias conforme mostram Ghitany e Maller (1994) utilizando a metodologia proposta por Fahrmeir e Kaufmann (1985) para modelos lineares generalizados. Contudo, certas condições de regularidade ainda são requeridas para a validade dos resultados assintóticos para os estimadores de máxima verossimilhança.

Na formulação considerada, as covariáveis são incorporadas a um modelo de mistura exponencial com sobreviventes por longo período, considerando que a proporção de indivíduos não imunes e a taxa de falhas dependem das covariáveis e de vetores de parâmetros desconhecidos, via funções logística e loglinear. Este tipo de parametrização é comum nos modelos lineares generalizados (veja por exemplo, McCullagh e Nelder, 1983) e também garante que os parâmetros pertençam ao espaço paramétrico de interesse.

O objetivo deste capítulo é apresentar uma análise Bayesiana para os modelos de mistura exponencial com covariáveis, utilizando métodos de simulação via cadeias de Markov para determinar as distribuições a posteriori dos parâmetros envolvidos. Esta metodologia não requer nenhuma suposição ou condição de regularidade sobre o modelo, o que é uma vantagem sobre o modelo clássico.

7.2 Análise clássica

Nesta seção apresenta-se o modelo de mistura exponencial com covariáveis e discutimos os estimadores de máxima verossimilhança (EMV) e teste de hipóteses dos parâmetros do modelo conforme apresentado por Githany e Maller (1994). Seja b_i , $1 \leq i \leq n$ uma sequência de variáveis aleatórias binárias, onde $b_i = 1$ indica que o i -ésimo indivíduo está em risco, sujeito a algum tipo de falha e $b_i = 0$ indica que o indivíduo não está em

risco. Nos não observamos os b_i de forma que não sabemos se um indivíduo é imune ou não, só podemos inferir a presença de imunes se muitas das observações são censuradas. Contudo, podemos ter vetores de covariáveis, $\mathbf{x}_i \in R^{k_1}$, $\mathbf{y}_i \in R^{k_2}$ contendo indicadores ou valores de quantidades como, por exemplo, idade, sexo, estado de saúde, disponível para cada indivíduo em consideração. Consideraremos que a covariável \mathbf{x}_i está associada à taxa de falhas (λ_i) e \mathbf{y}_i à proporção de indivíduos imunes, podendo inclusive serem as ser as mesmas, mas, geralmente elas diferem e podem ser consideradas com subvetores de um vetor geral de covariáveis.

Condisionalmente dado b_i , assumimos que os tempos de falha t_i^* são variáveis aleatórias independentes com função de densidade de probabilidades

$$f_i(t) = \lambda_i e^{-\lambda_i t}, \quad t \geq 0, \quad (7.1)$$

e a função de distribuição $F_i(t) = 1 - e^{-\lambda_i t}$. Assumimos que o parâmetro λ_i associado ao i -ésimo indivíduo está relacionado com as covariáveis \mathbf{x}_i através da relação

$$\lambda_i = \exp(\boldsymbol{\alpha}' \mathbf{x}_i), \quad (7.2)$$

onde $\boldsymbol{\alpha} \in R^{k_1}$ é um vetor k_1 -dimensional de parâmetros desconhecidos. Considera-se que distribuição de b_i está relacionado com a covariável \mathbf{y}_i através da relação

$$p_i = P(b_i = 1) = \frac{\exp\{\boldsymbol{\beta}' \mathbf{y}_i\}}{1 + \exp\{\boldsymbol{\beta}' \mathbf{y}_i\}} = 1 - P(b_i = 0), \quad (7.3)$$

onde $\boldsymbol{\beta} \in R^{k_2}$ é um vetor k_2 -dimensional de parâmetros desconhecidos.

A transformações em (7.2) e (7.3) garantem que os λ_i sejam positivos e restringe os p_i ao intervalo $(0, 1)$.

Usualmente, observamos o tempo de sobrevivência $t_i = \min(t_i^*, c_i)$, onde como antes os tempos de censura c_i são variáveis aleatórias independentes, os quais também são independentes de t_i^* e observamos $\delta_i = 1$ se $t_i^* \leq c_i$ e $\delta_i = 0$ caso contrário.

7.2.1 Estimadores de máxima verossimilhança

Seja $\theta = (\alpha', \beta')$ o vetor de parâmetros desconhecidos. A função log-verossimilhança para θ dado $(t_i, x_i, y_i, \delta_i)$, $1 \leq i \leq n$, pode ser escrita como

$$\ell_n(\theta) = \sum_{i=1}^n \{\delta_i(\log p_i + \log \lambda_i - \lambda_i t_i) + (1 - \delta_i)(1 - p_i + p_i e^{-\lambda_i t_i})\}. \quad (7.4)$$

O EMV de θ é obtido maximizando (7.4), o que equivale a resolver o sistema de equações $S_n(\theta) = \mathbf{0}$ dado a seguir.

De (7.1)-(7.2) podemos calcular o vetor das primeiras derivadas de $\partial \ell_n(\theta)/\partial \alpha$ e $\partial \ell_n(\theta)/\partial \beta$, que podemos escrever como

$$S_n(\theta) = \frac{\partial}{\partial \theta} \ell_n(\theta) = \begin{pmatrix} \frac{\partial}{\partial \alpha} \ell_n(\theta) \\ \frac{\partial}{\partial \beta} \ell_n(\theta) \end{pmatrix} = \sum_{i=1}^n \mathbf{X}_i s_i(\theta), \quad (7.5)$$

onde $s_i(\theta)$ são vetores de dimensão 2 com componentes

$$s_{i1}(\theta) = \delta_i(1 - \lambda_i t_i) - (1 - \delta_i) \frac{p_i \lambda_i t_i e^{-\lambda_i t_i}}{1 - p_i + p_i e^{-\lambda_i t_i}} \quad (7.6)$$

e

$$s_{i2}(\theta) = (1 - p_i) \left(\delta_i - \frac{(1 - \delta_i)p_i(1 - e^{-\lambda_i t_i})}{1 - p_i + p_i e^{-\lambda_i t_i}} \right), \quad (7.7)$$

e \mathbf{X}_i são matrizes não estocásticas de ordem $(k_1 + k_2) \times 2$ dadas por

$$\mathbf{X}_i = \begin{pmatrix} \mathbf{x}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{y}_i \end{pmatrix}, \quad (7.8)$$

$i = 1, \dots, n$. Como não é possível obter de modo analítico o estimador $\hat{\theta}$ que satifaz as equações acima, devemos utilizar algum método iterativo como, por exemplo, o método de Newton-Rapson para obter estimativas destes parâmetros. Esse método aparece como subrotina de alguns pacotes estatísticos tais como o Ox (veja, Doornik, 1998).

Pode-se mostrar que a matriz de segundas derivadas multiplicada por (-1) pode ser escrita como

$$\mathbf{F}_n(\theta) = -\frac{\partial^2}{\partial \theta^2} \ell_n(\theta) = - \begin{pmatrix} \frac{\partial^2}{\partial \alpha^2} \ell_n(\theta) & \frac{\partial^2}{\partial \alpha \partial \beta} \ell_n(\theta) \\ \frac{\partial^2}{\partial \beta \partial \alpha} \ell_n(\theta) & \frac{\partial^2}{\partial \beta^2} \ell_n(\theta) \end{pmatrix} \quad (7.9)$$

$$= \sum_{i=1}^n \mathbf{X}_i \mathcal{F}_i(\boldsymbol{\theta}) \mathbf{X}'_i, \quad (7.10)$$

onde $\mathcal{F}_i(\boldsymbol{\theta})$ é uma matriz aleatória simétrica de ordem 2×2 com elementos na diagonal dados por

$$f_i^{11}(\boldsymbol{\theta}) = \delta_i \lambda_i t_i + (1 - \delta_i) \frac{p_i \lambda_i t_i e^{-\lambda_i t_i}}{1 - p_i + p_i e^{-\lambda_i t_i}} \left\{ 1 - \frac{(1 - p_i) \lambda_i t_i}{1 - p_i + p_i e^{-\lambda_i t_i}} \right\}, \quad (7.11)$$

$$f_i^{22}(\boldsymbol{\theta}) = p_i (1 - p_i) \left\{ 1 - \frac{(1 - \delta_i) e^{-\lambda_i t_i}}{1 - p_i + p_i e^{-\lambda_i t_i}} \right\}, \quad (7.12)$$

e elementos fora da diagonal dados por

$$f_i^{12}(\boldsymbol{\theta}) = f_i^{21}(\boldsymbol{\theta}) = p_i (1 - p_i) (1 - \delta_i) \frac{\lambda_i t_i e^{-\lambda_i t_i}}{(1 - p_i + p_i e^{-\lambda_i t_i})^2}, \quad (7.13)$$

$i = 1, \dots, n$. Pode-se mostrar que a matriz de informação de Fisher, $\mathbf{D}_n = E(\mathbf{F}_n(\boldsymbol{\theta}))$, satisfaz $\mathbf{D}_n = E\{S_n(\boldsymbol{\theta})S_n(\boldsymbol{\theta})'\}$, de modo que de (7.9) pode-se escrever

$$\mathbf{D}_n = \sum_{i=1}^n \mathbf{X}_i \mathcal{D}_i(\boldsymbol{\theta}) \mathbf{X}'_i, \quad (7.14)$$

onde $\mathcal{D}_i = E\{\mathcal{F}_i(\boldsymbol{\theta})\}$, $1 \leq i \leq n$ são matrizes simétricas de ordem 2×2 , com elementos d_i^{rs} , $r, s = 1, 2$ dados por

$$\begin{aligned} d_i^{11} &= p_i E \left\{ 1 - e^{-\lambda_i c_i} - \frac{(1 - p_i) \lambda_i^2 c_i^2 e^{-\lambda_i c_i}}{1 - p_i + p_i e^{-\lambda_i c_i}} \right\}, \\ d_i^{22} &= p_i (1 - p_i)^2 E \left\{ \frac{1 - e^{-\lambda_i c_i}}{1 - p_i + p_i e^{-\lambda_i c_i}} \right\}, \\ d_i^{12} &= d_i^{21} = p_i (1 - p_i) E \left\{ \frac{\lambda_i c_i e^{-\lambda_i c_i}}{1 - p_i + p_i e^{-\lambda_i c_i}} \right\}, \end{aligned}$$

$i = 1, \dots, n$.

7.2.2 Inferência sobre os parâmetros α e β

A inferência sobre os parâmetros do modelo de mistura exponencial com covariáveis são feitas geralmente através da normalidade assintótica dos estimadores de máxima verosimilhança. Ghitany et al. (1994) mostraram que os EMV $\hat{\alpha}$ e $\hat{\beta}$ de α e β sob certas

condições de regularidade tem o seguinte comportamento, para $n \rightarrow \infty$:

$$\mathbf{D}_n^{1/2} \left[\begin{pmatrix} \hat{\boldsymbol{\alpha}} \\ \hat{\boldsymbol{\beta}} \end{pmatrix} - \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} \right] \xrightarrow{D} N(\mathbf{0}, I_{k_1+k_2}) \quad (7.15)$$

onde $\mathbf{D}_n^{1/2}$ é a matriz raíz quadrada da matriz de informação de Fisher dada em (7.14), $N(\mathbf{0}, I_{k_1+k_2})$ denota a distribuição normal multivariada padrão ($I_{k_1+k_2}$ é matriz identidade de dimensão $(k_1 + k_2) \times (k_1 + k_2)$).

7.3 Análise Bayesiana

Na seção anterior foi discutido que a distribuição dos estimadores de máxima verossimilhança somente é conhecida assintóticamente, o que pode trazer dificuldades, uma vez que em geral, trabalha-se com amostras pequenas e moderadas. Os estudos de simulação que apresentamos na seção 6.2 ilustram alguns dificuldades que podem ocorrer. Desta forma, inferências baseadas nessas estatísticas podem não ser confiáveis. Um procedimento alternativo que pode contornar esse problema é através de utilização de métodos Bayesianos que consideram os parâmetros do modelo como variáveis aleatórias e não como constantes desconhecidas. Nesta seção desenvolve-se esse método para o caso de modelos de mistura exponencial com sobreviventes de longo período e covariáveis através de métodos computacionais.

7.3.1 Análise Bayesiana do modelo de mistura exponencial com covariáveis usando o método MCMC

Dada $(t_i, \delta_i, \mathbf{x}_i, \mathbf{y}_i)$, $1 \leq i \leq n$, amostra aleatória de tamanho n , a função de verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta})$ é dada por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n (p_i \lambda_i \exp\{-\lambda_i t_i\})^{\delta_i} \left(1 - p_i + p_i e^{-\lambda_i t_i}\right)^{1-\delta_i}, \quad (7.16)$$

onde $p_i = \frac{\exp\{\boldsymbol{\beta}' \mathbf{y}_i\}}{1 + \exp\{\boldsymbol{\beta}' \mathbf{y}_i\}}$ e $\lambda_i = \exp(\boldsymbol{\alpha}' \mathbf{x}_i)$ dados em (7.2) e (7.3), respectivamente.

Para a inferência Bayesiana consideraremos as seguintes densidades a priori para α e β que consideramos independentes a priori:

$$\begin{aligned}\alpha_j &\sim N(\mu_{0j}, \sigma_{0j}^2), \quad \mu_{0j} \text{ e } \sigma_{0j}^2, j = 1, \dots, k_1 \text{ conhecidos,} \\ \beta_k &\sim N(\mu_{1k}, \sigma_{1k}^2), \quad \mu_{1k} \text{ e } \sigma_{1k}^2, k = 1, \dots, k_2 \text{ conhecidos,}\end{aligned}\quad (7.17)$$

onde $N(\mu, \sigma^2)$ denota a distribuição normal com média μ e variância σ^2 .

De (7.16)-(7.17) e do teorema de Bayes pode-se mostrar que a densidade a posteriori conjunta de α e β é dada por

$$\begin{aligned}\pi(\alpha, \beta | D) &\propto \exp \left\{ \sum_{i=1}^n \sum_{j=1}^{k_1} \delta_i x_{ij} \alpha_j - \sum_{i=1}^n \delta_i t_i e^{\alpha' \mathbf{x}_i} - \frac{1}{2} \sum_{j=1}^{k_1} \left(\frac{\alpha_j - \mu_{0j}}{\sigma_{0j}} \right)^2 - \frac{1}{2} \sum_{k=1}^{k_2} \left(\frac{\beta_k - \mu_{1k}}{\sigma_{1k}} \right)^2 \right\} \\ &\times \prod_{i=1}^n \left(\frac{\exp\{\beta' \mathbf{y}_i\}}{1 + \exp\{\beta' \mathbf{y}_i\}} \right)^{\delta_i} \left(1 - \frac{\exp\{\beta' \mathbf{y}_i\}}{1 + \exp\{\beta' \mathbf{y}_i\}} (1 - e^{-e^{\alpha' \mathbf{x}_i} t_i}) \right)^{1-\delta_i},\end{aligned}\quad (7.18)$$

onde D denota o conjunto de observações.

Da densidade a posteriori (7.18) pode-se mostrar que as densidades a posteriori condicionais para a aplicação do algoritmo de Gibbs são dadas por

$$\begin{aligned}\pi(\alpha_j | \boldsymbol{\alpha}_{(j)}, \boldsymbol{\beta}, D) &\propto \exp \left\{ \sum_{i=1}^n \delta_i x_{ij} \alpha_j - \sum_{i=1}^n \delta_i t_i e^{\alpha' \mathbf{x}_i} - \frac{1}{2} \left(\frac{\alpha_j - \mu_{0j}}{\sigma_{0j}} \right)^2 \right\} \Psi_1(\boldsymbol{\alpha}, \boldsymbol{\beta}), \\ \pi(\beta_k | \boldsymbol{\beta}_{(k)}, \boldsymbol{\alpha}, D) &\propto \exp \left\{ -\frac{1}{2} \left(\frac{\beta_k - \mu_{1k}}{\sigma_{1k}} \right)^2 \right\} \Psi_2(\boldsymbol{\alpha}, \boldsymbol{\beta}),\end{aligned}\quad (7.19)$$

onde $\boldsymbol{\alpha}_{(j)} = (\alpha_1, \dots, \alpha_{j-1}, \alpha_{j+1}, \dots, \alpha_{k_1})$, $\boldsymbol{\beta}_{(k)} = (\beta_1, \dots, \beta_{k-1}, \beta_{k+1}, \dots, \beta_{k_2})$,

$$\Psi_1(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^n \left(1 - \frac{\exp\{\beta' \mathbf{y}_i\}}{1 + \exp\{\beta' \mathbf{y}_i\}} (1 - e^{-e^{\alpha' \mathbf{x}_i} t_i}) \right)^{1-\delta_i},$$

$$\text{e} \quad \Psi_2(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^n \left(\frac{\exp\{\beta' \mathbf{y}_i\}}{1 + \exp\{\beta' \mathbf{y}_i\}} \right)^{\delta_i} \left(1 - \frac{\exp\{\beta' \mathbf{y}_i\}}{1 + \exp\{\beta' \mathbf{y}_i\}} (1 - e^{-e^{\alpha' \mathbf{x}_i} t_i}) \right)^{1-\delta_i},$$

As distribuições condicionais a posteriori dadas em (7.19) não possuem uma distribuição padrão, de modo que faremos uso do algoritmo de Metropolis-Hastings (veja, Chib e Greenberg, 1995) para gerar amostras a posteriori de $\boldsymbol{\alpha}$ e $\boldsymbol{\beta}$.

7.3.2 Estimador de Bayes do tempo médio de vida e para a proporção de sobreviventes dado um valor da covariável

Considerando as amostras Gibbs geradas em (7.19) de $\alpha_1, \dots, \alpha_{k_1}$ e $\beta_1, \dots, \beta_{k_2}$, podemos obter uma estimativa de Monte Carlo para os momentos de interesse. Um caso especial deste é dado pelo tempo médio de vida dado em (7.2) para um valor especificado da covariável \mathbf{x}_h . Um estimador de Bayes de λ_h com respeito à função de perda quadrática é dada por

$$E[\lambda_h|D] = \int \dots \int \exp \{ \alpha_1 x_{1h} + \dots + \alpha_{k_1 h} \} \pi(\alpha, \beta|D) d\alpha d\beta$$

a qual pode-se aproximar por sua estimativa de Monte Carlo

$$\hat{\lambda}_h = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S \exp \left\{ \alpha_1^{(r,s)} x_{1h} + \dots + \alpha_{k_1 h}^{(r,s)} \right\}, \quad (7.20)$$

onde $\alpha_1^{(r,s)}, \dots, \alpha_{k_1}^{(r,s)}$, denota os valores de $\alpha_1, \dots, \alpha_{k_1}$ respectivamente, obtidos na r -ésima iteração e a s - ésima replica, R e S são respectivamente, número total de iterações e simulações no amostrador de Gibbs e \mathbf{x}_h é um valor da covariável.

Similarmente uma estimativa da proporção de sobreviventes em (7.3) para um valor da covariável y_h com respeito à função perda quadrática é dada por $E[p_h|D]$. Um estimador de Monte Carlo é dado por

$$\hat{p}_h = \frac{2}{RS} \sum_{r=1}^R \sum_{s=\frac{S}{2}+1}^S = \frac{\exp \left\{ \beta_1^{(r,s)} y_{1h} + \dots + \beta_{k_2}^{(r,s)} y_{k_2 h} \right\}}{1 + \exp \left\{ \beta_1^{(r,s)} y_{1h} + \dots + \beta_{k_2}^{(r,s)} y_{k_2 h} \right\}} \quad (7.21)$$

7.4 Aplicação

Com a finalidade de ilustrar a metodologia apresentada nas seções 2 e 3, consideramos dois exemplos: o primeiro consiste de um conjunto de dados simulados e o segundo são os dados analizados por Kersey et al.(1980).

7.4.1 Exemplo 1-Dados simulados

A tabela 7.1, apresenta uma amostra de $n = 30$ observações de tempos de sobrevivência geradas do modelo de mistura exponencial com covariáveis. Considerou-se a variável b_i que indica se o indivíduo está em risco ou não, que tem distribuição de Bernoulli ($1, p_i$) com $p_i = \frac{\exp\{2+x_{i1}\}}{1+\exp\{2+x_{i1}\}}$. Para os indivíduos em risco, simulou-se os tempos de falha (morte), t_i^* a partir da distribuição exponencial com média dada em (7.2): $\lambda_i = \exp\{2,5 + 0,2x_{i1}\}$ e para as covariáveis gerou-se considerando a distribuição normal com média zero e variância 4 e para os tempos de censura, simulou-se uma variável aleatoria uniforme no intervalo $(0, 2)$.

Tabela 7.1: Amostra gerada do modelo de mistura exponencial com covariáveis.

i	t_i	δ_i	x_{i1}	i	t_i	δ_i	x_{i1}
1	0,06707	1	-1,60000	16	0,01053	0	0,32277
2	0,05666	1	2,06390	17	0,03374	1	0,22997
3	0,07656	1	-0,13747	18	0,02408	1	3,14980
4	0,40201	0	-2,60500	19	0,17374	1	-2,04450
5	0,03569	1	3,33550	20	0,02615	1	3,62700
6	0,13400	0	-1,79780	21	0,00416	1	1,41050
7	0,05527	1	1,10310	22	0,01355	1	0,15793
8	0,31485	1	0,90310	23	0,06404	1	3,79480
9	0,04123	1	1,73770	24	0,06758	1	1,80100
10	0,27833	0	-3,39280	25	0,02492	1	3,33170
11	0,02419	0	2,08630	26	0,02895	0	1,04340
12	0,01238	0	0,83944	27	0,05138	1	1,12030
13	0,13172	1	2,16810	28	0,41689	0	1,00790
14	0,13965	1	0,96796	29	0,06251	1	1,07880
15	0,03017	1	2,45200	30	0,25799	0	-2,74850

Análise clássica

Nesta subseção apresenta-se uma aplicação da metodologia discutida na seção 2 considerando os dados da tabela 7.1. Na tabela 7.2, apresenta-se as estimativas de máxima verossimilhança e os intervalos de confiança de 95% admitindo-se a normalidade assintótica

dos estimadores de máxima verossimilhança de α_0 , α_1 , β_0 e β_1 .

Tabela 7.2: Estimativas de máxima verossimilhança dos parâmetros do modelo de mistura exponencial com covariáveis

	EMV	D.P	Intervalo de confiança 95%
$\hat{\alpha}_0$	2,1323	0,4042	(1,1340 ; 2,9246))
$\hat{\alpha}_1$	0,30143	0,1798	(-0,0510 ; 0,6538)
$\hat{\beta}_0$	2,0295	1,0657	(-0,0393 ; 4,1183)
$\hat{\beta}_1$	0,9581	0,5432	(-0,1065 ; 2,0227)

Análise Bayesiana pelo método MCMC

Nesta subseção desenvolve-se a análise Bayesiana apresentada na seção 3 para o modelo de sobrevivência com sobreviventes de longo período com covariáveis utilizando-se o algoritmo de Gibbs com Metropolis-Hasting (G-MH) para os dados da tabela 7.1.

Para a inferência Bayesiana consideramos as densidades a priori dada em (7.15) com $\alpha_j \sim N(0, 10^5)$, $j = 0, 1$ e $\beta_k \sim N(0, 2 \times 10^8)$, $k = 0, 1$, quase não informativas. Para se obter amostras de Gibbs através do algoritmo G-MH, procede-se da seguinte forma: geram-se 20 cadeias paralelas cada um com 4.000 iterações. Depreza-se as 2.000 primeiras iterações, com objetivo de diminuir o efeito dos pontos iniciais, e a partir daí escolhem-se as iterações de 20 em 20, ou seja para cada parâmetro em cada cadeia considera-se a geração 1.020 – ésima, ..., 2.000 – ésima fornecendo uma amostra Gibbs de tamanho 2.000. Para monitorar a convergência utiliza-se a técnica de Gelman e Rubin (1992), método este que utiliza a técnica de análise de variância para verificar se mais iterações são necessárias. Na tabela 7.3, apresenta-se o resumo da distribuição a posteriori (7.18) e na figura 7.1 representa-se a densidade aproximada dos parâmetros do modelo de mistura exponencial com covariáveis. Também na tabela 7.3), tem-se o resultado do fator de redução de escala potencial estimada \hat{R} (veja, Gelman e Rubin, 1992) para todos os parâmetros. O número

de iterações considerado foi suficiente para se chegar à convergência, pois $\hat{R} < 1.1$ para todos os parâmetros.

Tabela 7.3: Sumário a posteriori dos parâmetros do modelo de mistura exponencial com covariáveis

	Medía	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α_0	2,3348	2,3388	0,2495	(1,8501 ; 2,8866)	1,08677
α_1	0,1858	0,1849	0,1469	(-0,0969 ; 0,4772)	1,00383
β_0	2,0976	2,0476	0,7348	(0,7889 ; 3,6790)	1,01045
β_1	1,0536	1,0574	0,3782	(0,3109 ; 1,7908)	1,00442

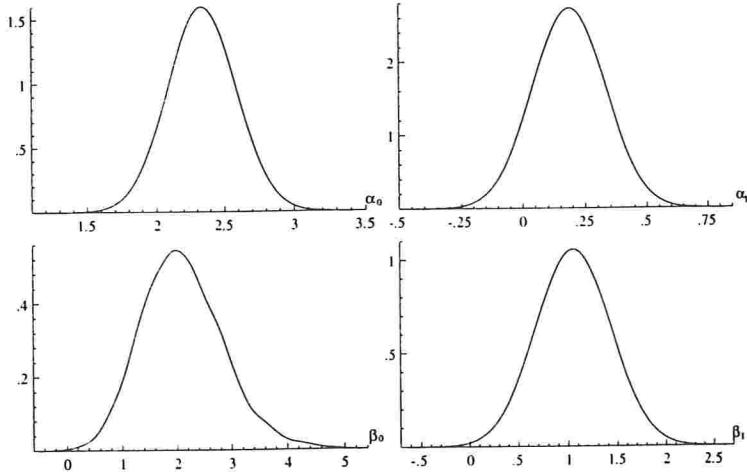


Figura 7.1: Densidades marginais a posteriori aproximadas dos parâmetros do modelo de mistura exponencial com covariáveis

Note que o desvio padrão do enfoque Bayesiano é bem menor que dos estimadores de máxima verossimilhança, mas, as médias a posteriori dos parâmetros do modelo são muito próximas aos estimadores de máxima verossimilhança. Contudo, é importante notar a redução na variância obtida com o procedimento Bayesiano, em alguns casos a redução

chega a quase 50%.

Na tabela 7.4, apresenta-se estimativas do tempo médio de vida e da proporção de sobreviventes dado um valor da covariável, $\mathbf{x}_h = \mathbf{y}_h = (1 \ 0, 2)$.

Tabela 7.4: Estimativas do tempo médio e da proporção de sobreviventes dado $\mathbf{x}_h = (1 \ 0, 2)$

	Valor real	EMV	G-MH
λ_h	12.679	8.958	11.033
p_h	0.900	0.850	0.892

Pode-se verificar através da tabela 7.4 resultados próximos para as estimativas de tempo de vida média e da proporção de sobreviventes obtidos pelo algoritmo de Gibbs com Metropolis-Hasting para a covariável $\mathbf{x}_h = \mathbf{y}_h = (1 \ 0, 2)$. Na figura 7.2, representa-se as densidades a posteriori do tempo médio λ_h e da proporção de sobreviventes, p_h para um valor da covariável.

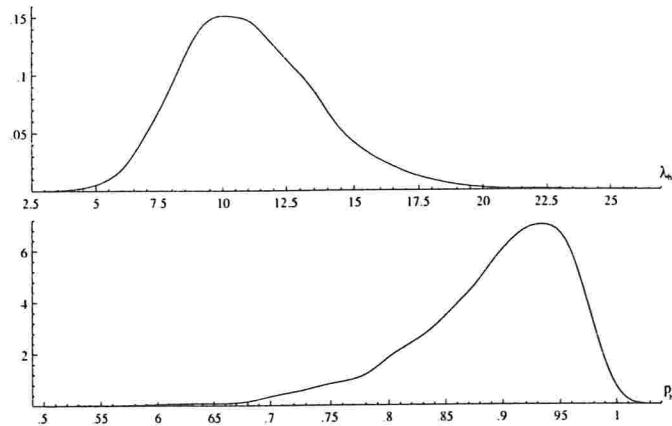


Figura 7.2: Densidades marginais a posteriori aproximados do tempo médio de vida e da proporção de sobreviventes para um valor dado da covariável $\mathbf{x}_h = (1 \ 0, 2)$

7.4.2 Exemplo 2-Dados de Kersey et al. 1987

Nesta subseção apresentamos um exemplo como aplicação dos procedimentos clássicos e bayesianos, discutidos nas seções anteriores. O exemplo consiste em analisar os dados (veja tabela 6.6) obtidos por Kersey et al. (1987), os quais se referem aos tempos de recaída de pacientes de leucemia submetidos a dois tipos de transplante, "allogeneic" (Grupo 1) ou "autologous" (Grupo 2). Neste caso consideraremos que o vetor de covariáveis \mathbf{x}_i associado ao i -ésimo indivíduo tem a seguinte forma:

$$\mathbf{x}_i = (1, x_i),$$

onde $x_i = 1$ para o Grupo 1 e $x_i = 2$ para o Grupo 2.

Análise clássica

Nesta subseção apresenta-se uma aplicação da metodologia discutida na seção 2 considerando os dados da tabela 6.6. Considerando a log-verossimilhança dada em (7.16) temos que o estimador de máxima verossimilhança do parâmetro $\theta = (\alpha_0, \alpha_1, \beta_0, \beta_1)$ é obtido pela resolução do sistema de equações não lineares $S_n(\theta) = 0$. Isto é feito através da subrotina de Newton-Rapson do pacote **Ox**.

As estimativas de máxima verossimilhança são dadas por

$$\hat{\alpha}_0 = -0,2664(0,4213), \quad \hat{\alpha}_1 = 0,6263(0,2581), \quad \hat{\beta}_0 = 0,6146(0,7809), \quad \text{e} \quad \hat{\beta}_1 = 0,3746(0,5081).$$

Cada valor entre parênteses ao lado das estimativas representa a raiz quadrada do correspondente elemento da diagonal de $F_n^{-1}(\hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta}_0, \hat{\beta}_1)$, onde $F_n(\alpha_0, \alpha_1, \beta_0, \beta_1)$ é a matriz de informação observada.

Análise Bayesiana pelo método MCMC

Nesta subseção desenvolve-se a análise bayesiana apresentada na seção 3 para o modelo de sobrevivência com sobreviventes de longo período com covariáveis utilizando-se o

algoritmo de Gibbs com Metropolis-Hasting (G-MH) para os dados da tabela 6.6.

Para a inferência Bayesiana consideramos as densidades a priori dada em (7.15) com $\alpha_j \sim N(0, 10^5)$, $j = 0, 1$ e $\beta_k \sim N(0, 2 \times 10^8)$, $k = 0, 1$, quase não informativas. Para se obter uma amostra de Gibbs através do algoritmo G-MH, procede-se da seguinte forma: geram-se 20 cadeias paralelas cada um com 3.000 iterações. Depreparam-se as 1.500 primeiras iterações, com objetivo de diminuir o efeito dos pontos iniciais, e a partir daí escolhem-se as iterações de 15 em 15, ou seja para cada parâmetro em cada cadeia considera-se as iterações 1.015 –ésima, ..., 2.000 –ésima que fornece uma amostra Gibbs de tamanho 1.500. Para monitorar a convergência utiliza-se a técnica de Gelman e Rubin (1992), método este que utiliza a técnica de análise de variância para verificar se mais iterações são necessárias. Na tabela 7.5, apresenta-se o resumo das distribuições a posteriori (7.18) e na figura 7.3, representa-se as densidades aproximadas dos parâmetros do modelo de mistura exponencial com covariáveis. Também na tabela 7.5, tem-se o resultado do fator de redução de escala potencial estimada \hat{R} (veja, Gelman e Rubin, 1992) para todos os parâmetros. O número de iterações considerado foi suficiente para se chegar à convergência, pois $\hat{R} < 1.1$ para todos os parâmetros.

Tabela 7.5: Sumário a posteriori dos parâmetros do modelo de mistura exponencial com covariáveis

	Medía	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α_0	-0,2578	-0,2588	0,0981	(-0,4381 ; -0,0782)	1,001057
α_1	0,6011	0,6017	0,0875	(0,4301 ; 0,7830)	1,001083
β_0	0,6170	0,6208	0,2682	(0,0967 ; 1,1374)	1,010045
β_1	0,3872	0,3868	0,1883	(0,0237 ; 0,7591)	1,00026

Na tabela 7.6, apresenta-se a estimativa do tempo médio de vida e a proporção de sobreviventes para os Grupos 1 e 2, e na figura 7.4, representa-se suas respectivas densidades

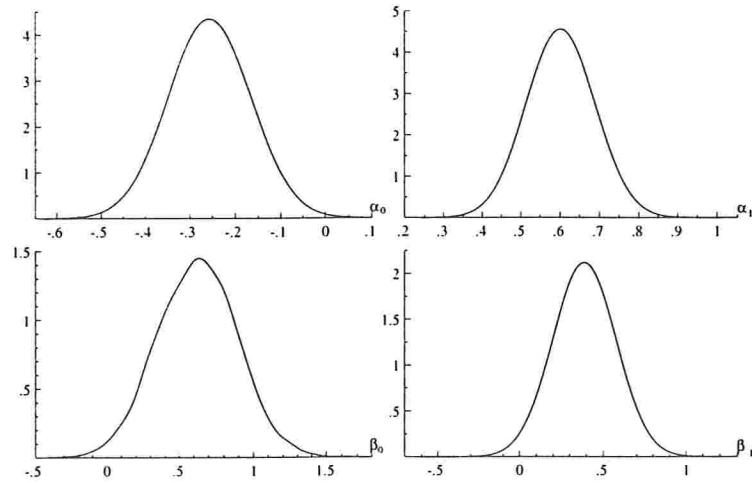


Figura 7.3: Densidades marginais a posteriori aproximadas dos parâmetros do modelo de mistura exponencial com covariáveis

a posteriori aproximadas.

Tabela 7.6: Estimativas do tempo médio e a proporção de sobreviventes para os Grupos 1 e 2

	Grupo 1		Grupo 2	
	MV	G-MH	MV	G-MH
λ	1,4332	1,5678	2,6810	2,7945
p	0,7289	0,7243	0,7964	0,7975

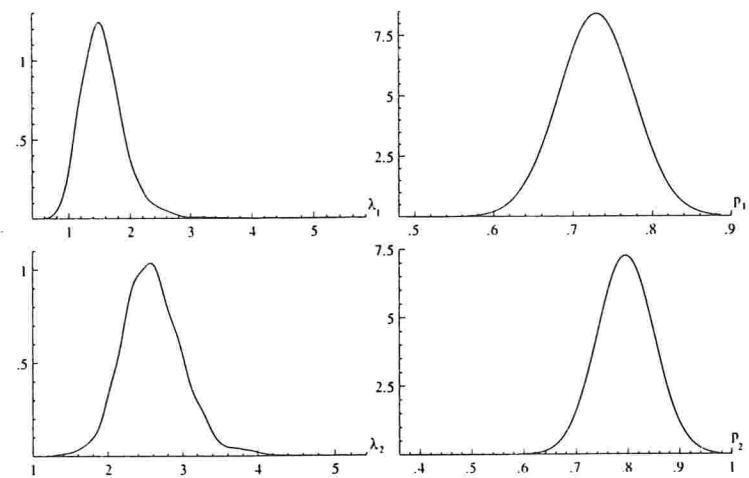


Figura 7.4: Densidades marginais a posteriori aproximadas do tempo médio de vida e a proporção de sobreviventes para os Grupos 1 e 2.

Capítulo 8

Modelo de mistura Weibull-exponenciada

8.1 Introdução

A distribuição Weibull-exponenciada (WE) foi introduzida por Mudholkar et al. (1995) e representa uma generalização das distribuições Weibull e exponencial. Neste Cípitulo consideramos a análise de sobrevivência com indivíduos imunes na população considerando a distribuição WE para os indivíduos que falham (modelo de mistura WE) e estuda-se o uso deste modelo no teste de adequação dos modelos Weibull e exponencial desde a perspectiva clássica e Bayesiana.

8.2 Modelo

Seguindo a formulação de Farewell (1977), considera-se b_i , $i = 1, \dots, n$, uma variável binária que indica i -ésimo indivíduo de uma população está em risco ou não com relação a algum tipo de falha ($b_i = 1$ se o indivíduo pertence ao grupo de risco e $b_i = 0$ caso

contrário). Considera-se que os b_i são independentes e identicamente distribuidos tais que $P(b_i = 1) = p \in (0, 1)$ indica a proporção de indivíduos em risco e $P(b_i = 0) = 1 - p$ a proporção de indivíduos imunes ou proporção de indivíduos curados de uma certa doença ou itens que nunca falham.

As variáveis b_i são não observadas mas, assume-se que dado que os indivíduos pertencem ao grupo de risco, isto é, $b_i = 1$, $i = 1, \dots, n$, os tempos de falha t_i^* , são independentes e identicamente distribuidos com distribuição Weibull-exponenciada. Isto é,

$$g(t; \alpha, \theta, \lambda | b = 1) = \alpha \theta \lambda^\alpha [1 - \exp\{-(\lambda t)^\alpha\}]^{\theta-1} \exp\{-(\lambda t)^\alpha\} t^{\alpha-1} \quad (8.1)$$

onde $\alpha > 0$ e $\theta > 0$ são os parâmetros de forma e $\lambda > 0$ o é parâmetro de escala. Observe que quando $\theta = 1$, (8.1) é função densidade da distribuição Weibull e quando $\alpha = 1$ e $\theta = 1$, (8.1) é a densidade da distribuição exponencial.

Associado com o i -ésimo indivíduo está um tempo de censura c_i . A distribuição dos tempos de censura são considerados não informativos, isto é, não envolve $(\alpha, \theta, \lambda, p)$. Consideramos que t_i^* e c_i , $i = 1, \dots, n$ são estocasticamente independentes. Para os indivíduos com $b_i = 1$, observamos tempo de censura ou falha (morte), $t_i = \min(t_i^*, c_i)$ e o indicador de censura $\delta_i = 1$ se ocorre falha e $\delta_i = 0$, caso contrário. Quando $b_i = 0$, observamos $t_i = c_i$ e, neste caso todos os tempos são censurados.

8.3 Uma análise clássica do modelo de mistura Weibull-exponenciada

Nesta seção tratamos de procedimentos inferenciais para o modelo de mistura Wibull-exponenciada. Fornecemos a seguir procedimentos de estimação dos parâmetros pelo método de máxima verossimilhança, assim como métodos assintóticos para a construção de testes e intervalos de confiança aproximados para os parâmetros do modelo.

8.3.1 Estimação dos parâmetros por máxima verossimilhança

Sejam t_1, \dots, t_n observações de uma amostra onde t_i representa o tempo de vida ou o tempo de censura de cada um dos n indivíduos em estudo. Seja $\boldsymbol{\eta} = (\alpha, \theta, \lambda, p)' \in \{\boldsymbol{\eta} : \alpha > 0, \theta > 0, \sigma > 0, 0 < p < 1\}$ o vetor de parâmetros a ser estimado.

Um indivíduo que falhou no tempo t_i , contribui para a função de verossimilhança com a parcela

$$P\{b_i = 1\}g(t; \alpha, \theta, \sigma | b_i = 1) = p\alpha\theta\lambda^\alpha [1 - \exp\{-(\lambda t_i)^\alpha\}]^{\theta-1} \exp\{-(\lambda t_i)^\alpha\}t_i^{\alpha-1},$$

enquanto, que um indivíduo que não falhou até o tempo t_i contribui para a função de verossimilhança com a parcela

$$P\{b_i = 0\} + P\{b_i = 1\}P(t_i^* > t_i; \alpha, \theta, \sigma | b_i = 1) = 1 - p[1 - \exp\{-(\lambda t_i)^\alpha\}]^\theta,$$

que é a probabilidade do indivíduo ser um sobrevivente por longo período mais a probabilidade de um indivíduo sobreviver pelo menos até o tempo t_i . Portanto, a contribuição do i -ésimo indivíduo para a função de verossimilhança é dada por

$$\left(p\alpha\theta\lambda^\alpha [1 - \exp\{-(\lambda t_i)^\alpha\}]^{\theta-1} \exp\{-(\lambda t_i)^\alpha\}t_i^{\alpha-1} \right)^{\delta_i} \left(1 - p[1 - \exp\{-(\lambda t_i)^\alpha\}]^\theta \right)^{1-\delta_i}.$$

Logo, dada uma amostra $t_1, \dots, t_n, \delta_1, \dots, \delta_n$, a função log-verossimilhança para $\boldsymbol{\eta}$ é dada por

$$\begin{aligned} \ell(\alpha, \theta, \lambda, p) &= r(\log(p\alpha\theta) + \alpha \log \lambda) - \lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha + (\theta - 1) \sum_{i=1}^n \delta_i \log(1 - \exp(-(\lambda t_i)^\alpha)) \\ &+ (\alpha - 1) \sum_{i=1}^n \delta_i \log t_i + \sum_{i=1}^n (1 - \delta_i) \log(1 - p[1 - \exp(-(\lambda t_i)^\alpha)]^\theta) \end{aligned} \quad (8.2)$$

O estimador de máxima verossimilhança de $\boldsymbol{\eta}$ é obtido maximizando (8.2), o que equivale a resolver o sistema de equações $S(\boldsymbol{\eta}) = \sum_{i=1}^n s_i(\boldsymbol{\eta}) = \mathbf{0}$, com $s_i(\boldsymbol{\eta}) = (s_{i1}(\boldsymbol{\eta}), s_{i2}(\boldsymbol{\eta}), s_{i3}(\boldsymbol{\eta}), s_{i4}(\boldsymbol{\eta}))'$, onde

$$s_{i1}(\boldsymbol{\eta}) = \frac{\partial \ell_i}{\partial \alpha} = \delta_i \left(\alpha^{-1} + \log(\lambda) + \frac{(\theta - 1)(\lambda t_i)^\alpha \log(\lambda t_i) e^{-(\lambda t_i)^\alpha}}{1 - e^{-(\lambda t_i)^\alpha}} - (\lambda t_i)^\alpha \log(\lambda t_i) + \log(t_i) \right) -$$

$$\frac{(1 - \delta_i) p (1 - e^{-(\lambda t_i)^\alpha})^\theta \theta (\lambda t_i)^\alpha \log(\lambda t_i) e^{-(\lambda t_i)^\alpha}}{(1 - e^{-(\lambda t)^\alpha}) (1 - p (1 - e^{-(\lambda t_i)^\alpha})^\theta)}, \quad (8.3)$$

$$s_{i2}(\boldsymbol{\eta}) = \frac{\partial \ell_i}{\partial \theta} = \delta_i \left(\theta^{-1} + \log(1 - e^{-(\lambda t_i)^\alpha}) \right) - \frac{(1 - \delta_i) p (1 - e^{-(\lambda t_i)^\alpha})^\theta \log(1 - e^{-(\lambda t_i)^\alpha})}{1 - p (1 - e^{-(\lambda t_i)^\alpha})^\theta}, \quad (8.4)$$

$$s_{i3}(\boldsymbol{\eta}) = \frac{\partial \ell_i}{\partial \lambda} = \delta_i \left(\frac{\alpha}{\lambda} + \frac{(\theta - 1) (\lambda t_i)^\alpha \alpha e^{-(\lambda t_i)^\alpha}}{\lambda (1 - e^{-(\lambda t_i)^\alpha})} - \frac{(\lambda t_i)^\alpha \alpha}{\lambda} \right) - \frac{(1 - \delta_i) p (1 - e^{-(\lambda t_i)^\alpha})^\theta \theta (\lambda t_i)^\alpha \alpha e^{-(\lambda t_i)^\alpha}}{\lambda (1 - e^{-(\lambda t_i)^\alpha}) (1 - p (1 - e^{-(\lambda t_i)^\alpha})^\theta)}, \quad (8.5)$$

$$s_{i4}(\boldsymbol{\eta}) = \frac{\partial \ell_i}{\partial p} = \frac{\delta_i}{p} - \frac{(1 - \delta_i) (1 - e^{-(\lambda t_i)^\alpha})^\theta}{1 - p (1 - e^{-(\lambda t_i)^\alpha})^\theta}, \quad (8.6)$$

sendo ℓ_i a i -ésima componente da função log-verossimilhança. Do sistema de equações não lineares dado acima não conseguimos obter analiticamente os estimadores de máxima verossimilhança, e precisamos de um método iterativo para solucionar o sistema. Este problema pode ser resolvido usando método de Newton-Raphson o qual aparece como rotina em alguns pacotes tais como **Ox** (veja, Doornik, 1998).

8.3.2 Inferências sobre os parâmetros α , θ , λ e p

Para fazermos inferências sobre os parâmetros α , θ , λ e p , utilizaremos a normalidade assintótica dos estimadores de máxima verossimilhança (EMV). Assim, para $\hat{\alpha}$, $\hat{\theta}$, $\hat{\lambda}$ e \hat{p} EMV's de α , θ , λ e p , ponto interior do espaço paramétrico e sob certas condições de regularidade (veja, Cox e Hinkley, 1974) pode-se mostrar, para $n \rightarrow \infty$, que

$$i^{1/2}(\alpha, \theta, \lambda, p) \begin{bmatrix} \hat{\alpha} \\ \hat{\theta} \\ \hat{\lambda} \\ \hat{p} \end{bmatrix} - \begin{bmatrix} \alpha \\ \theta \\ \lambda \\ p \end{bmatrix} \xrightarrow{D} N(\mathbf{0}, I_4) \quad (8.7)$$

onde $i^{1/2}(\alpha, \theta, \lambda, p)$ é a matriz raiz quadrada da matriz de informação de Fisher, $N(\mathbf{0}, I_4)$ denota a distribuição normal multivariada com vetor de médias igual $\mathbf{0}$ e matriz de

variâncias I_4 (I_4 denota a matriz indentidade de ordem 4×4). A normalidade assintótica (8.7) segue sendo válida se $i(\alpha, \theta, \lambda, p)$ é substituído por seu estimador consistente $i(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p})$ ou simplesmente considerarmos a matriz de informação avaliada nos estimadores de máxima verossimilhança denotada por $I(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p})$.

Considerando a distribuição assintótica normal dada em (8.7) para os estimadores de máxima verossimilhança $\hat{\alpha}$, $\hat{\theta}$, $\hat{\lambda}$ e \hat{p} , podemos construir intervalos para α , θ , λ e p , considerando um nível de confiança $100(1 - \gamma)\%$, dados por,

$$ic(\alpha) = (\hat{\alpha} - z_{\gamma/2}(I^{(11)})^{1/2}; \hat{\alpha} + z_{\gamma/2}(I^{(11)})^{1/2}), \quad (8.8)$$

$$ic(\theta) = (\hat{\theta} - z_{\gamma/2}(I^{(22)})^{1/2}; \hat{\theta} + z_{\gamma/2}(I^{(22)})^{1/2}), \quad (8.9)$$

$$ic(\lambda) = (\hat{\lambda} - z_{\gamma/2}(I^{(33)})^{1/2}; \hat{\lambda} + z_{\gamma/2}(I^{(33)})^{1/2}), \quad (8.10)$$

$$ic(p) = (\hat{p} - z_{\gamma/2}(I^{(44)})^{1/2}; \hat{p} + z_{\gamma/2}(I^{(44)})^{1/2}), \quad (8.11)$$

onde $I^{(jj)}$, $j = 1, 2, 3, 4$, é o j -ésimo elemento diagonal da matriz $i^{-1}(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p})$ e $z_{\gamma/2}$ é um percentil da distribuição normal padrão.

8.3.3 Teste do modelo de mistura Weibull

Considerando as hipóteses

$$H_0 : \theta = 1 \text{ versus } H_1 : \theta \neq 1, \quad (8.12)$$

estaremos testando a adequação do modelo de mistura Weibull para os tempos de vida. Usamos especificamente a estatística da razão de verossimilhança

$$\Lambda = -2 \left\{ \ell(\tilde{\alpha}, 1, \tilde{\lambda}, \tilde{p}) - \ell(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p}) \right\}, \quad (8.13)$$

onde $\ell(\tilde{\alpha}, 1, \tilde{\lambda})$ corresponde ao máximo de $\ell(\alpha, 1, \lambda)$ dado por

$$\ell(\alpha, 1, \lambda, p) = r(\log(p\alpha) + \alpha \log \lambda) - \lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha + (\alpha - 1) \sum_{i=1}^n \delta_i \log t_i + \sum_{i=1}^n (1 - \delta_i) \log(1 - p + p \exp(-(\lambda t_i)^\alpha))$$

A estatística (8.13) para grandes amostras, tem distribuição aproximada qui-quadrado com *um* grau de liberdade sob a hipóteses H_0 .

8.3.4 Teste para o modelo de mistura exponencial

Considerar

$$H_0 : \begin{pmatrix} \theta \\ \alpha \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ versus, } H_1 : \begin{pmatrix} \theta \\ \alpha \end{pmatrix} \neq \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (8.14)$$

equivale a testar que os dados admitem um modelo de mistura exponencial para os tempos de vida e a estatística da razão de verossimilhança usada para o teste é

$$\Lambda = -2 \left(\ell(1, 1, \tilde{\lambda}, \tilde{p}) - \ell(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p}) \right), \quad (8.15)$$

onde $\tilde{\lambda}$ é o valor que maximiza a log-verossimilhança

$$\ell(1, 1, \lambda, p) = r \log(p) + r \log \lambda - \lambda \sum_{i=1}^n \delta_i t_i + \sum_{i=1}^n (1 - \delta_i) \log(1 - p + p \exp(-\lambda t_i)).$$

A estatística da razão de verossimilhança (8.15), sob H_0 tem distribuição assintótica qui-quadrado com 2 graus de liberdade.

8.4 Uma análise Bayesiana para o modelo de mistura Weibull-exponenciada

Na seção anterior foi considerada uma análise clássica para o modelo de mistura Weibull-exponenciada baseada em resultados assintóticos que são válidos para amostras de tamanho grande. Já verificamos isto no caso exponencial. Para amostras de tamanhos moderados ou pequenos esses resultados assintóticos podem não ser confiáveis. Um procedimento alternativo que pode contornar esse problema é através da utilização de métodos Bayesianos que consideram os parâmetros como variáveis aleatórias.

Nesta seção, apresenta-se uma análise considerando-se a abordagem Bayesiana para o modelo de mistura Weibull-exponenciada. Neste ponto surgem duas dificuldades: a primeira se refere a obtenção da distribuição a posterior marginal e a segunda se refere ao cálculo de momentos de interesse. Em ambos casos é necessário a resolução de integrais que muitas vezes não apresentam solução analítica. Neste caso métodos de aproximação

ou métodos numéricos são necessários. Um método muito utilizado é o método de aproximação de Laplace (Tierney, Kass e Kadane, 1986). Para aplicação desse método é necessário maximizar funções que dependem de parâmetros, de modo que, a medida que a dimensão do espaço paramétrico aumenta, aumentam as dificuldades para sua obtenção e a aplicação do método vai se tornando mais trabalhosa e de difícil implementação. Para contornar esse problema neste trabalho utiliza-se os métodos de simulação de Monte Carlo via Cadeias de Markov, tais como o amostrador de Gibbs (Casella e George, 1992) e Metropolis-Hastings (Chib e Greenberg, 1995).

8.4.1 Densidade a priori independentes

Para representar o grau de conhecimento sobre os parâmetros do modelo de mistura Weibull-exponenciada, consideramos as seguintes densidades à priori para α , θ , λ e p :

$$\begin{aligned}\alpha &\sim \Gamma(\alpha_0, \alpha_1), \\ \theta &\sim \Gamma(\theta_0, \theta_1), \\ \lambda &\sim \Gamma(\lambda_0, \lambda_1), \\ p &\sim \text{Beta}(\gamma_0, \gamma_1),\end{aligned}\tag{8.16}$$

onde $\Gamma(a, b)$ denota a distribuição Gama com média a/b , $\text{Beta}(\gamma_0, \gamma_1)$ denota a distribuição Beta com média $\frac{\gamma_0}{(\gamma_0 + \gamma_1)}$, e α_i , θ_i , λ_i e γ_i , $i = 0, 1$, são constantes conhecidas baseadas na informação a priori do especialista e da análise preliminar dos dados. Assumindo independência entre os parâmetros a densidade a priori conjunta é dada por.

$$\pi(\alpha, \theta, \lambda, p) \propto p^{\gamma_0 - 1} (1 - p)^{\gamma_1 - 1} \alpha^{\alpha_0 - 1} \theta^{\theta_0 - 1} \lambda^{\lambda_0} \exp\{-\alpha_1 \alpha - \theta_1 \theta - \lambda_1 \lambda\},\tag{8.17}$$

onde $\alpha > 0$, $\theta > 0$, $\lambda > 0$ e $0 < p < 1$.

De (8.2) pode-se mostrar que a função verossimilhança para α , θ , λ e p é dada por $L(\alpha, \theta, \lambda, p) = \exp\{\ell(\alpha, \theta, \lambda, p)\}$, isto é,

$$L(\alpha, \theta, \lambda, p) = p^r \alpha^r \theta^r \lambda^{r\alpha} \exp\left\{-\lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha + (\alpha - 1) \sum_{i=1}^n \delta_i \log t_i + \right.\left.\dots\right\}\tag{8.18}$$

$$(\theta - 1) \sum_{i=1}^n \delta_i \log[1 - \exp\{-(\lambda t_i)^\alpha\}] + \sum_{i=1}^n (\delta_i - 1) \log \left(1 - p[1 - \exp\{-(\lambda t_i)^\alpha\}]^\theta\right)\},$$

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

8.4.2 Densidade conjunta a posteriori para os parâmetros do modelo de mistura Weibull-exponenciada

Considerando as densidade a priori (8.17) e a função de verossimilhança (8.18), a densidade a posteriori conjunta para α , θ , λ e p é dada por (veja, por exemplo, Box e Tiao, 1973)

$$\pi(\alpha, \theta, \lambda, p|D) \propto \pi(\alpha, \theta, \lambda, p)L(\alpha, \theta, \lambda, p),$$

isto é,

$$\begin{aligned} \pi(\alpha, \theta, \lambda, p|D) &\propto p^{\gamma_0+r-1}(1-p)^{\gamma_1-1}\alpha^{\alpha_0+r-1}\theta^{\theta_0+r-1}\lambda^{\lambda_0+r\alpha-1}\exp\left\{-\lambda^\alpha\sum_{i=1}^n\delta_i t_i^\alpha-\alpha_1\alpha-\right. \\ &\quad \theta_1\theta-\lambda_1\lambda+(\alpha-1)\sum_{i=1}^n\delta_i\log t_i+(\theta-1)\sum_{i=1}^n\delta_i\log[1-\exp\{-(\lambda t_i)^\alpha\}]+ \\ &\quad \left.\sum_{i=1}^n(\delta_i-1)\log\left(1-p[1-\exp\{-(\lambda t_i)^\alpha\}]^\theta\right)\right\}, \end{aligned} \quad (8.19)$$

onde D denota o conjunto dos dados observados.

No enfoque bayesiano, inferências são baseadas nas densidades a posteriori marginais dos parâmetros envolvidos no modelo. Pode-se observar que densidade a posteriori conjunta (8.19) não é uma densidade padrão, portanto, só podemos avaliar as densidades marginais por meio de métodos de aproximação tais como o método de Laplace ou através de métodos de simulação de Monte Carlo via cadeia de Markov (MCMC). Nesta seção optamos pelo método de simulação por serem de fácil implementação computacional. Para obtermos uma amostra das densidades a posteriori marginais de α , θ , λ e p , utiliza-se o amostrador de Gibbs que é uma técnica para gerar variáveis aleatórias de uma distribuição marginal independentemente, isto é, este método permite gerar variáveis aleatórias de uma distribuição marginal sem ter que calcular a forma da densidade.

De (8.19) pode-se mostrar que as densidades condicionais marginais $\pi(\alpha|\theta, \lambda, p, D)$, $\pi(\theta|\alpha, \lambda, p, D)$, $\pi(\lambda|\alpha, \theta, p, D)$ e $\pi(p|\alpha, \theta, \lambda, D)$ para o algoritmo de Gibbs são dadas por,

$$\begin{aligned}
(i) \quad \pi(\alpha|\theta, \lambda, p, D) &\propto \alpha^{\alpha_0+r-1} \lambda^{r\alpha} \exp \left\{ -\lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha - \alpha(\alpha_1 - \sum_{i=1}^n \delta_i \log t_i) + \Psi(\alpha|\theta, \lambda, p) \right\}, \\
(ii) \quad \pi(\theta|\alpha, \lambda, p, D) &\propto \theta^{\theta_0+r-1} \exp \left\{ \theta \left(\theta_0 - \sum_{i=1}^n \delta_i \log[1 - \exp\{-(\lambda t_i)^\alpha\}] \right) + \right. \\
&\quad \left. \sum_{i=1}^n (\delta_i - 1) \log \left(1 - p[1 - \exp\{-(\lambda t_i)^\alpha\}]^\theta \right) \right\}, \\
(iii) \quad \pi(\lambda|\alpha, \theta, p, D) &\propto \lambda^{\lambda_0+r\alpha-1} \exp \left\{ -\lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha - \lambda \lambda_1 + \Psi(\alpha|\theta, \lambda, p) \right\}, \\
(iv) \quad \pi(p|\alpha, \theta, \lambda, D) &\propto p^{\gamma_0+r-1} (1-p)^{\gamma_1-1} \exp \left\{ \sum_{i=1}^n (\delta_i - 1) \log \left(1 - p[1 - \exp\{-(\lambda t_i)^\alpha\}]^\theta \right) \right\},
\end{aligned} \tag{8.20}$$

onde

$$\Psi(\alpha|\theta, \lambda, p) = (\theta-1) \sum_{i=1}^n \delta_i \log[1 - \exp\{-(\lambda t_i)^\alpha\}] + \sum_{i=1}^n (\delta_i - 1) \log \left(1 - p[1 - \exp\{-(\lambda t_i)^\alpha\}]^\theta \right).$$

Observe em (i)-(iv) que as densidades são não conhecidas, portanto, utilizaremos o amostrador de Gibbs com algoritmo de Metropolis-Hasting para gerar as amostras a posteriori de α , θ , λ e p .

O procedimento para a aplicação do algoritmo de Metropolis-Hasting dentro do ciclo do amostrador de Gibbs é ilustrado brevemente para geração de α em (8.20), para isto denotamos por $f(\alpha)$ a densidade condicional (i) já que $\alpha > 0$ define-se a seguinte transformação $\alpha' = \log \alpha$ e consideramos o nucleo de transição $q(\alpha', X)$ que leva α a X ; por exemplo, $X = \alpha' + \sigma Z$, sendo Z uma variável aleatória normal padrão e σ^2 tem a variância de α' e computamos $f_{\alpha'}(\alpha') = f(\alpha)|d\alpha/d\alpha'|$, sendo a probabilidade de movimento da cadeia $p = \min\{1, \frac{f_{\alpha'}(X)}{f_{\alpha'}(\alpha')}\}$. Em cada passo do algoritmo de Gibbs transformamos α' ; $\alpha = \exp(\alpha')$. Similar procedimento pode ser aplicado para gerar θ , λ e para p o nucleo de transição $U(0, 1)$ (distribuição uniforme no intervalo $(0, 1)$).

8.4.3 Densidade conjunta a posteriori para os parâmetros do modelo de mistura Weibull

Considere as seguintes densidades a priori para α , λ e p parâmetros do modelo de mistura Weibull:

$$\begin{aligned}\alpha &\sim \Gamma(\alpha_0, \alpha_1), \text{ com } \alpha_0 \text{ e } \alpha_1 \text{ conhecidos} \\ \lambda &\sim \Gamma(\lambda_0, \lambda_1), \text{ com } \lambda_0 \text{ e } \lambda_1 \text{ conhecidos} \\ p &\sim \text{Beta}(\gamma_0, \gamma_1), \text{ com } \gamma_0 \text{ e } \gamma_1 \text{ conhecidos.}\end{aligned}\tag{8.21}$$

Assumindo independência entre α , λ e p e considerando a função de verossimilhança com $\theta = 1$ em (8.18), pode-se mostrar que a densidade conjunta a posteriori de α , λ e p é dada por

$$\begin{aligned}\pi(\alpha, \lambda, p|D) \propto & p^{r+\gamma_0-1}(1-p)^{\gamma_1-1} \alpha^{r+\alpha_0-1} \lambda^{r\alpha+\lambda_0-1} \exp\left\{-\lambda^\alpha \sum_{i=1}^n \delta_i t_i - \alpha\alpha_1 - \lambda\lambda_1\right\} \times \\ & \prod_{i=1}^n \left(t_i^{\alpha-1}\right)^{\delta_i} (1-p + p \exp\{-(\lambda t_i)^\alpha\})^{\delta_i-1},\end{aligned}\tag{8.22}$$

onde D é o conjunto dos dados observados.

De (8.22) pode-se mostrar que as densidades condicionais marginais $\pi(\alpha|\lambda, p, D)$, $\pi(\lambda|\alpha, p, D)$ e $\pi(p|\alpha, \lambda, D)$ para o algoritmo de Gibbs são dadas por,

$$\begin{aligned}(i) \quad \pi(\alpha|\lambda, p, D) &\propto \alpha^{r+\alpha_0-1} \lambda^{r\alpha} \exp\left\{-\lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha - \alpha(\alpha_1 + \sum_{i=1}^n \delta_i \log t_i)\right\} \Psi_1(\alpha, \lambda, p), \\ (ii) \quad \pi(\lambda|\alpha, p, D) &\propto \lambda^{r\alpha+\lambda_0-1} \exp\left\{-\lambda^\alpha \sum_{i=1}^n \delta_i t_i^\alpha - \lambda\lambda_1\right\} \Psi_1(\alpha, \lambda, p), \\ (iii) \quad \pi(p|\alpha, \lambda, D) &\propto p^{r+\lambda_0-1} (1-p)^{\lambda_1-1} \Psi_1(\alpha, \lambda, p),\end{aligned}\tag{8.23}$$

onde $\Psi_1(\alpha, \lambda, p) = \prod_{i=1}^n (1-p + p \exp\{-(\lambda t_i)^\alpha\})^{\delta_i-1}$.

Observe que as densidades a posteriori condicionais em (8.23) não possuem uma forma padrão, pelo que faremos uso do algoritmo de Metropolis-Hastings para gerar amostras a posteriori.

8.4.4 Teste de hipóteses Bayesiano

Na seção anterior considerou-se o modelo de mistura Weibull-exponenciada para verificar se certo conjunto de dados de tempos de vida podem ser ajustados pelo modelo de mistura Weibull ou pelo modelo de mistura exponencial. Nesta seção testa-se as hipóteses dadas em (8.12) e (8.14) a partir de uma perspectiva Bayesiana.

Como descrito no capítulo 1, consideramos a técnica das densidades preditivas condicionais ordenadas (CPO) definidas por

$$f(t_i|\underline{t}_{(i)}) = \int f(t_i|\theta, \underline{t}_{(i)})\pi(\theta|\underline{t}_{(i)})d\theta, \quad (8.24)$$

onde $\underline{t}_{(i)} = (t_1, \dots, t_{i-1}, t_{i+1}, \dots, t_n)$, para determinar a verossimilhança preditiva dos modelos sob as hipóteses nula e alternativa para testar a adequabilidade do modelo de mistura Weibull e do modelo de mistura exponencial, utilizando o pseudo-fator de Bayes (veja, por exemplo, Gelfand e Dey, 1994).

Usando as amostras geradas pelo algoritmo de Gibbs pode-se estimar a densidade preditiva (8.24) para o modelo de mistura Weibull (modelo sob H_0 dado em (8.12)), que denotaremos por, $f(t_i|\underline{t}_{(i)}, M_0)$. A estimativa de Monte Carlo desta densidade é dada por,

$$\hat{f}(t_i|\underline{t}_{(i)}, M_0) = \begin{cases} B \left(\sum_{k=1}^B [\alpha^{(k)} (\lambda^{(k)})^{\alpha^{(k)}} p^{(k)} \exp\{-(\lambda^{(k)})^{\alpha^{(k)}} t_i^{\alpha^{(k)}}\} t_i^{\alpha^{(k)}-1}]^{-1} \right)^{-1}, & \text{se } \delta_i = 1 \\ B \left(\sum_{k=1}^B [1 - p^{(k)}(1 + \exp(-(\lambda^{(k)} t_i)^{\alpha^{(k)}}))]^{-1} \right)^{-1}, & \text{se } \delta_i = 0, \end{cases} \quad (8.25)$$

onde B é o tamanho da amostra Gibbs, e $(\alpha^{(k)}, \lambda^{(k)}, p^{(k)})$ é a k -ésima amostra de Gibbs de α , λ e p . A estimativa da verossimilhança preditiva sob H_0 (8.12) é dada por $\hat{f}(\underline{t}|M_0) = \prod_{i=1}^n \hat{f}(t_i|\underline{t}_{(i)}, M_0)$.

Para aproximar a densidade preditiva para o modelo de mistura Weibull-exponenciada (modelo sob H_1 em (8.12)), que denotaremos por $\hat{f}(t_i|\underline{t}_{(i)}, M_1)$, uma estimativa de Monte

Carlo com as amostras Gibbs dos parâmetros do modelo é dada por

$$\hat{f}(t_i | \underline{t}_{(i)}, M_1) = \begin{cases} B \left(\sum_{k=1}^B [\alpha^{(k)} \theta^{(k)} (\lambda^{(k)})^{\alpha^{(k)}} p^{(k)} \exp\{-(\lambda^{(k)})^{\alpha^{(k)}} t_i^{\alpha^{(k)}}\} t_i^{\alpha^{(k)}-1} \times \right. \\ \left. (1 - \exp\{-(\lambda^{(k)} t_i)^{\alpha^{(k)}}\})^{\theta^{(k)}}]^{-1} \right)^{-1}, & \text{se } \delta_i = 1 \\ B \left(\sum_{k=1}^B [1 - p^{(k)} (1 + \exp(-(\lambda^{(k)} t_i)^{\alpha^{(k)}}))]^{-1} \right)^{-1}, & \text{se } \delta_i = 0, \end{cases} \quad (8.26)$$

onde $(\alpha^{(k)}, \theta^{(k)}, \lambda^{(k)}, p^{(k)})$ é a k -ésima amostra Gibbs de α , θ , λ e p . A estimativa da verossimilhança preditiva sob H_1 em (8.12) é dada por $\hat{f}(\underline{t} | M_1) = \prod_{i=1}^n \hat{f}(t_i | \underline{t}_{(i)}, M_1)$. Com (8.25) e (8.26) pode-se calcular o pseudo-fator de Bayes do modelo M_0 (modelo de mistura Weibull) com respeito ao modelo M_1 (modelo de mistura Weibull-exponenciada), que denotaremos por $PSFB_{01}$ e rejeita-se H_0 se $PSFB_{01} < 1$.

8.5 Exemplos de aplicação

Nesta seção apresentamos dois exemplos como aplicação dos procedimentos clássicos e Bayesianos, discutidos nas seções anteriores.

8.5.1 Exemplo 1-Dados Simulados

Na tabela 8.1, apresenta-se uma amostra de $n = 30$ observações de tempos de sobre-vivência geradas do modelo de mistura Weibull-exponenciada. Considerou-se que a variável b_i que indica se o indivíduo está em risco ou não, tem distribuição Bernoulli $(1, p)$ com $p = 0.7$. Para os indivíduos em riscos, simulou-se os tempos de falha (morte), t_i^* tendo distribuição Weibull-exponenciada com parâmetros $\alpha = 4$, $\theta = 1$ e $\lambda = 3$ e para os tempos de censura simulou-se uma variável aleatória exponencial com média 6.

Tabela 8.1: Dados gerados do modelo de mistura Weibull-exponenciada com + indicando censura

0,23242	0,073788	0,50877+	0,27088	6,3350+	0,27678	0,045127+	8,2059+	3,0524+	0,45935
0,29664	0,37975	0,14298	3,5607+	0,28250	0,41267	0,29316	0,27130	0,32717	0,19019
2,6033+	0,33151	0,30793	0,33978	0,19775	0,38698+	1,6313+	0,25243+	10,769+	0,25877

Análise clássica

Considerando a log-verossimilhança dada em (8.2) temos que o estimador de máxima verossimilhança do parâmetro $\boldsymbol{\eta} = (\alpha, \theta, \lambda, p)$ é obtido pela resolução do sistema de equações não lineares $\frac{\partial \ell}{\partial \alpha} = 0$, $\frac{\partial \ell}{\partial \theta} = 0$, $\frac{\partial \ell}{\partial \lambda} = 0$ e $\frac{\partial \ell}{\partial p} = 0$. Isto é feito através da rotina de Newton-Rapson do pacote **Ox**.

As estimativas de máxima verossimilhança são dadas por

$$\hat{\alpha} = 4,5415(2,4286), \quad \hat{\theta} = 0,6735(0,3225), \quad \hat{\lambda} = 2,9096(0,5183), \quad \hat{p} = 0,6826(0,0895).$$

Cada valor entre parênteses ao lado das estimativas representa a raiz quadrada do correspondente elemento da diagonal de $i^{-1}(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p})$, onde $i(\alpha, \theta, \lambda, p)$ é a matriz de informação observada. Se admitimos a normalidade assintótica para os dados da tabela 8.1, $i^{-1}(\hat{\alpha}, \hat{\theta}, \hat{\lambda}, \hat{p})$ corresponde à estimativa da matriz de covariâncias assintótica, e pode ser utilizada para computar inferências sobre os parâmetros.

Para estudar a possibilidade da utilização do modelo de mistura Weibull para os dados da tabela 8.1. A estatística da razão de verossimilhança para testar $H_0 : \theta = 1$ é $\Lambda = 0,21072$ e indica a aceitação do modelo Weibull a um nível de significância 0.05 ($\chi^2_{(1;0,05)} = 3,8441$). Testando o modelo de mistura exponencial ($H_0 : \alpha = 1 \theta = 1$) temos $\Lambda = 27,688$, o que indica que a um nível de significância de 5% ($\chi^2_{(2;0,05)} = 5,99$), o modelo de mistura exponencial não é adequado para ajuste dos dados.

Análise Bayesiana para o modelo de mistura Weibull-exponenciada

Para analisarmos os dados da tabela 8.1 desde uma perspectiva Bayesiana consideramos o modelo de mistura Weibull-exponenciada (WE) com as densidades a priori para α, θ, λ e p dada em (8.16) com $\alpha_0 = 3, 24$, $\alpha_1 = 0, 74$ $\theta_0 = 1, 403$, $\theta_1 = 2, 094$, $\lambda_0 = 32, 34$, $\lambda_1 = 11, 15$, $\gamma_0 = 146, 3$ e $\gamma_1 = 62, 7$ (a escolha desses valores para os parâmetros da densidade a priori, foram baseadas numa opinião experiente combinada com uma análise preliminar dos dados). A partir da densidades condicionais (8.20), geramos para cada um dos parâmetros do modelo de mistura Wiebull-exponenciada 20 cadeias separadas de Gibbs cada uma com 2.000 iterações, e devido ao fato de todas as densidades condicionais marginais (8.20) não apresentarem uma forma padrão conhecida utilizamos então o algoritmo de Metropolis-Hastings. Com o objetivo de diminuir o efeito de pontos iniciais desprezam-se as 1.000 primeiras iterações e a partir de daí escolhem-se as iterações de 20 em 20, ou seja, para cada parâmetro considera-se a 1.020 –ésima, ..., 2.000 –ésima iterações; daí as 20 cadeias fornecem um amostra Gibbs de tamanho 2.000 para cada um dos parâmetros. Essa seleção foi realizada de forma a garantir uma baixa correlação entre os elementos da amostra. Esse fato pode ser confirmado através da figura 8.1, que ilustra as funções de autocorrelação das amostras. A convergência das amostras geradas pelo algoritmo de Gibbs com Metropolis-Hastings foi monitorado utilizando o método proposto por Gelman e Rubin (1992) o qual se baseia na técnica de análise de variância (veja Gelman e Rubin, 1992).

Como pode-se observar na figura 8.1, todos os termos são próximos de zero, indicando que as amostras são não correlacionadas.

Na tabela 8.2, apresenta-se um resumo das distribuições a posteriori marginais aproximada conjuntamente com os estimadores dos fatores de redução de escala potencial \hat{R} (ver por exemplo, Gelman e Rubin, 1992). Observamos na tabela 8.2, que $\sqrt{\hat{R}} < 1.1$, o que nos indica que o algoritmo de Gibbs com Metropolis-Hastings gera para cada parâmetro amostras que convergem em distribuição para as distribuições condicionais marginais (8.20).

Na figura 8.2, representa-se as densidades a posteriori marginais aproximadas consi-

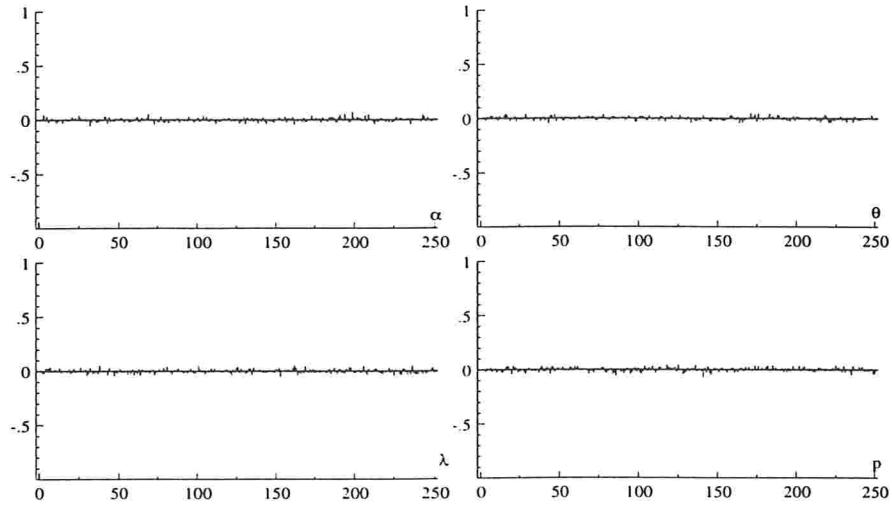


Figura 8.1: Autocorrelação das amostras geradas pelo algoritmo Metropolis-hasting com 2000 pontos

Tabela 8.2: Sumário a posteriori dos parâmetros do modelo de mistura Weibull-exponenciada

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α	3,9150	3,9054	0,2896	(3,3758 ; 4,5031)	1,01387
θ	0,96311	0,95632	0,11437	(0,7616 ; 1,1939)	1.00180
λ	3,0966	3,1053	0,2436	(2,5957 ; 3,5563)	1.00876
p	0,69977	0,70038	0,02929	(0,6434 ; 0,7568)	1.00019

derando os 1.000 pontos amostrais.

A partir dos resultados da tabela 8.2, podemos considerar o modelo de mistura Weibull para analisar os dados da tabela 8.1, visto que o estimador de Bayes para θ baseado na aproximação de Monte Carlo para a média a posteriori está bem próximo de um (observe também que o intervalo de credibilidade de 95% para θ inclui um), quando $\theta = 1$ o modelo Weibull-exponenciada (8.1) é o modelo de mistura Weibull.

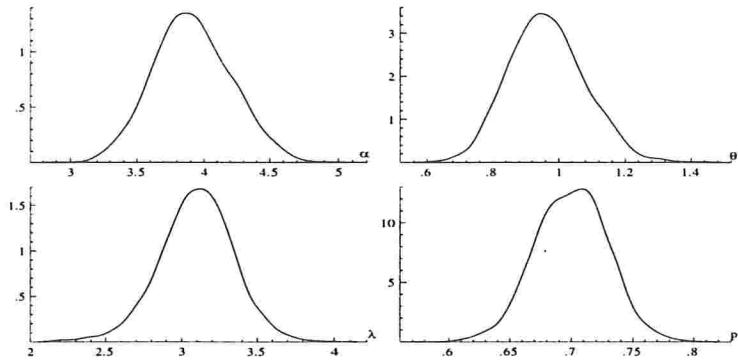


Figura 8.2: Densidades marginais a posteriori aproximados para os parâmetros do modelo de mistura Weibull-exponenciada α , θ , λ and p

Análise Bayesiana para o modelo de mistura Weibull

Para analisar os dados da tabela 8.1, consideramos agora o modelo de mistura Weibull com as densidades a priori (8.21) para os parâmetros, com, $\alpha_0 = 29, 17$, $\alpha_1 = 8, 33$, $\lambda_0 = 208, 91$, $\lambda_1 = 67, 4$ e $\gamma_0 = \gamma_1 = 1$

A partir das densidades condicionais marginais (8.23), geramos 20 cadeias separadas de Gibbs cada uma com 2.000 iterações e utilizamos o método proposto por Gelman e Rubin (1992) para verificar a convergência das cadeias. Para cada um dos parâmetro na cadeia consideramos as iterações 1.020 – ésima, ..., 2.000 – ésima totalizando assim uma amostra Gibbs de 2.000 pontos amostrais.

As quantidades a posteriori de interesse obtidas a partir das amostras selecionadas são dados na tabela 8.3, onde observamos que os fatores de redução potencial são menores de 1.1 ($\sqrt{\hat{R}} < 1.1$), indicando a convergência das amostras geradas.

Na figura 8.6, temos as densidades marginais a posteriori aproximadas considerando os 2.000 pontos amostrais.

Tabela 8.3: Sumário a posteriori dos parâmetros do modelo de mistura Weibull

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α	3,4980	3,4719	0,4595	(2,64616;4,44235)	1,09131
λ	3,1723	3,1782	0,1605	(2,64616;4,44235)	1,00671
p	0,67970	0,68115	0,03017	(0,617727;0,735598)	1,00038

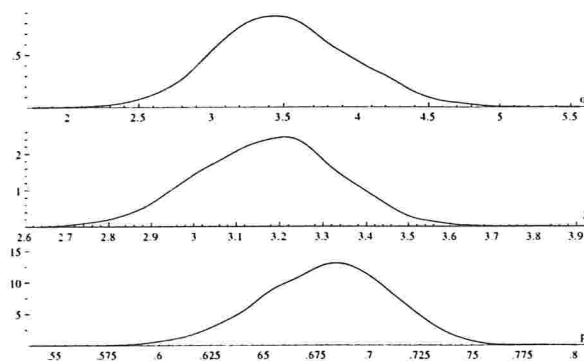


Figura 8.3: Densidades marginais a posteriori aproximados para o modelo de mistura Weibull α , λ and p

Teste de hipóteses Bayesiano

A estimativa de Monte Carlo da verossimilhança preditiva baseada nas amostras Gibbs dos parâmetros do modelo de mistura Weibul-exponenciada (modelo sob H_1 em (8.12)) resultou em $\hat{f}(D|M_1) = 0,328767$.

Para o modelo sob H_0 dado em (8.12)(modelo de mistura Weibull) uma estimativa da verossimilhança priditiva baseada nas amostras de Gibbs dos parâmetros do modelo Weibull resultou em $\hat{f}(D|M_0) = 1,13166$. O pseudo-fator de Bayes do modelo de mistura Weibull com respeito ao modelo Weibull-exponenciada é $PSFB_{01} = 3,464806$. Conclui-se que o modelo de mistura Weibull é adequado para o ajuste dos dados da tabela 8.1. A mesma conclusão foi obtida com o procedimento clássico.

Para testar se o modelo exponencial é adequado para o ajuste dos dados da tabela 8.1 desde a perspectiva Bayesiana, calculamos a verossimilhança preditiva que resultou em $\hat{f}(D|M_2) = 3,0809776 \times 10^{-6}$. O pseudo-fator de Bayes do modelo exponencial com respeito ao modelo Weibull-exponenciada é $PSFB_{21} = 9,37129 \times 10^{-6}$, indicando a rejeição do modelo exponencial para o ajuste dos dados da tabela 8.1, confirmando assim a conclusão obtidas com o método clássico.

Na figura 8.4, plotamos o log a razão dos CPO para os diferentes modelos com respeito ao número de observações. Valores positivos do log da razão dos CPO indica a preferência pelo primeiro modelo. Por exemplo, na figura 8.4, o gráfico para o modelo de mistura Weibull versus do modelo de mistura Weibull-exponenciada indica que 16 das 30 observações suportam o modelo de mistura Weibull sobre o modelo de mistura Weibull-exponenciada. Similarmente, 18 observações sustentam o modelo de mistura de Weibull-exponenciada sobre o modelo de mistura exponencial.

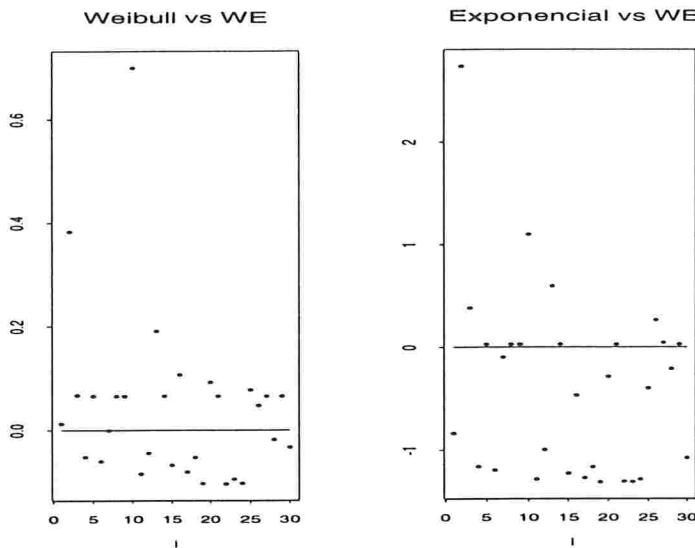


Figura 8.4: Gráfico da razão de CPO para comparar os diferentes modelos versus observações

8.5.2 Exemplo 2-Dados de Kersey et al. 1987

Nesta subseção apresentamos um exemplo como aplicação dos procedimentos clássicos e bayesianos, discutidos nas seções anteriores. O exemplo consiste em analisar os dados (veja tabela 6.6) obtidos por Kersey et al. (1987), os quais se referem aos tempos de recaída de pacientes de leucemia submetidos a dois tipos de transplante, "allogeneic" (Grupo 1) ou "autologous" (Grupo 2).

Análise clássica

Nesta subseção apresenta-se uma aplicação da metodologia discutida na seção 2. Considerando-se apenas os dados do Grupo 2 da tabela 6.6. Para estimar os parâmetros, consideramos o mesmo procedimento dado antes.

Como resultado as estimativas de máxima verossimilhança são dadas por

$$\hat{\alpha} = 0,42002(0,25013), \hat{\theta} = 44,452(20,56), \hat{\lambda} = 104.13(15,58), \hat{p} = 0.79629(0,0608).$$

Para estudar a possibilidade de utilizar o modelo de mistura Weibull para os dados grupo 2 da tabela 6.6, testa-se ($H_0 : \theta = 1$) e temos que a estatística da razão de verossimilhanças (8.13), resultou em, $\lambda = 11.083$ e este resultado implica que o modelo mistura Weibull não é estatisticamente adequada para ajustar os dados do grupo 2 da tabela 6.6 ao nível de significância de 5%, ($\chi^2_{1,0,05} = 3,8441$).

Testando o modelo de mistura exponencial ($H_0 : \alpha = 1 \theta = 1$) temos $\Lambda = 17,2112$, o que indica que a um nível de significância de 5% ($\chi^2_{(2;0,05)} = 5,99$), o modelo de mistura exponencial não é adequado para o ajuste dos dados.

Análise Bayesiana para o modelo de mistura Weibull-exponenciada

Para analisarmos os dados do Grupo 2 da tabela 6.6 desde uma perspectiva Bayesiana consideramos o modelo de mistura Weibull-exponenciada (WE) com as densidades a priori

para α, θ, λ e p dada em (8.16) com $\alpha_0 = 1, \alpha_1 = 0,00000002, \theta_0 = 1, \theta_1 = 0,00000002, \lambda_0 = 1, \lambda_1 = 0,0000002, \gamma_0 = 1$ e $\gamma_1 = 1$ (observe que a densidade a priori é não informativa). A partir da densidades condicionais (8.20), geramos para cada um dos parâmetros do modelo de mistura Weibull-exponenciada 20 cadeias separadas de Gibbs cada uma com 2.000 iterações, e devida ao fato de todas as densidades condicionais marginais (8.20) não apresentarem uma forma padrão conhecida utilizamos então o algoritmo de Metropolis-Hasting. Com o objetivo de diminuir o efeito de pontos iniciais desprezam-se as 1.000 primeiras iterações e a partir de daí escolhem-se as iterações de 20 em 20, ou seja, para cada parâmetro considera-se a 1020 – ésima, ..., 2000 – ésima iterações; daí as 20 cadeias fornecem um amostra Gibbs de tamanho 2.000 para cada um dos parâmetros. A convergência das amostras geradas pelo algoritmo de Gibbs com Metropolis-Hasting foi monitorada utilizando o método proposto por Gelman e Rubin (1992).

Na tabela 8.4, apresenta-se um resumo das distribuições a posteriori marginais aproximadas conjuntamente com as estimativas dos fatores de redução de escala potencial \hat{R} . Observamos na tabela 8.4, que $\sqrt{\hat{R}} < 1,1$, o que nos indica que o algoritmo de Gibbs com Metropolis-Hasting gera para cada parâmetro amostras que convergem em distribuição para as distribuições condicionais marginais (8.20).

Tabela 8.4: Sumário a posteriori dos parâmetros do modelo de mistura Weibull-exponenciada

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α	0,4191	0,4188	0,0129	(0,3939 ; 0,4455)	1,000387
θ	44,216	44,172	1,715	(40,919 ; 47,650)	1,00781
λ	104,33	104,29	2,216	(100,056 ; 108,866)	1,00626
p	0,7923	0,7931	0,02811	(0,7340 ; 0,8441)	1,00156

Na figura 8.5, representa-se as densidades marginais a posteriori aproximadas considerando os 1.000 pontos amostrais.

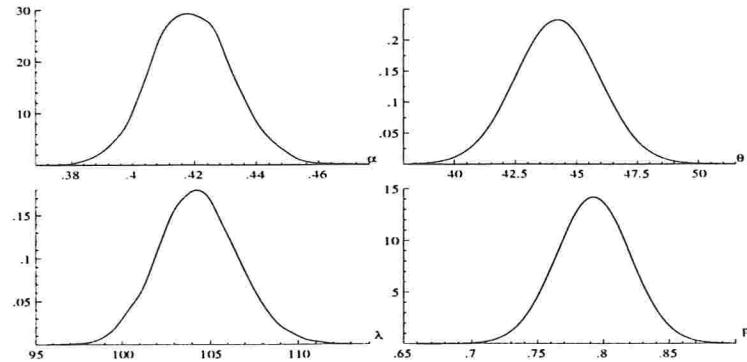


Figura 8.5: Densidades a posteriori margais aproximados para os parâmetros do modelo de mistura Weibull-exponenciada α , θ , λ and p

Análise Bayesiana para o modelo de mistura Weibull

Para analisar os dados da tabela 8.1, consideramos agora o modelo de mistura Weibull com as densidades a priori (8.21) para os parâmetros, com, $\alpha_0 = 1$, $\alpha_1 = 0,00000002$, $\lambda_0 = 1$, $\lambda_1 = 0,0000002$ e $\gamma_0 = \gamma_1 = 1$

A partir das densidades condicionais marginais (8.23), geramos 20 cadeias separadas de Gibbs cada uma com 2.000 iterações e utilizamos o

método proposto por Gelman Rubin (1992) par verificar a convergência das cadeias. Para cada um dos parâmetro na cadeia consideramos as iterações 1020 –ésima, . . . , 2000 –ésima totalizando assim uma amostra Gibbs de 2.000 pontos amostrais.

As quantidades a posteriori de interesse obtidas a partir das amostras selecionadas são dadas na tabela 8.5, onde observamos que os fatores de redução potencial são menores de 1.1 ($\sqrt{\hat{R}} < 1,1$), indicando a convergência das amostras geradas.

Na figura 8.6, temos as densidades marginais a posteriori aproximadas considerando os 2.000 pontos amostrais.

Tabela 8.5: Sumário a posteriori dos parâmetros do modelo de mistura Weibull

	Média	Mediana	D.P.	Intervalo de credibilidade 95%	\hat{R}
α	1,3516	1,3490	0,1368	(1,09102;1,62325)	1,000471
λ	2,4309	2,4230	0,2671	(1,9328;2,9882)	1,000914
p	0,7911	0,7917	0,02836	(0,7329 ; 0,8441)	1,0000173

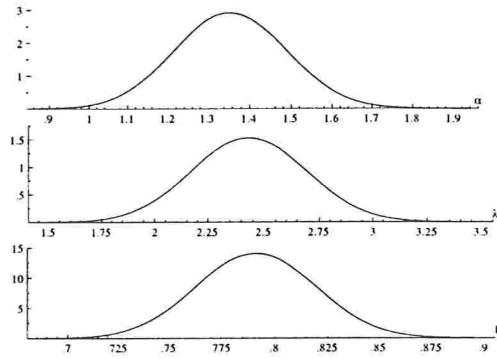


Figura 8.6: Densidades a posteriori aproximados para o modelo de mistura Weibull α , λ and p

Teste de hipóteses Bayesiano

A estimativa de Monte Carlo da verossimilhança preditiva baseada nas amostras Gibbs dos parâmetros do modelo de mistura Weibull-exponenciada (modelo sob H_1 em (8.12)) resultou em $\hat{f}(D|M_1) = 2,953 \times 10^{-7}$.

Para o modelo sob H_0 dado em (8.12)(modelo de mistura Weibull) uma estimativa da verossimilhança priditiva baseada nas amostras de Gibbs dos parâmetros do modelo Weibull resultou em $\hat{f}(D|M_0) = 1,681 \times 10^{-12}$. O pseudo-fator de Bayes do modelo de mistura Weibull com respeito ao modelo Weibull-exponenciada é $PSFB_{01} = 5,6925 \times 10^{-6}$. Conclui-se que o melhor modelo é a mistura Weibull-exponenciada para os dados do grupo 2 da tabela 6.6. A mesma conclusão foi obtida com o procedimento clássico.

Para testar se o modelo exponencial é adequado para o ajuste dos dados da tabela 6.6 desde a perspectiva Bayesiana, calculamos a verossimilhança preditiva que resultou em $\hat{f}(D|M_2) = 1,306 \times 10^{-10}$. O pseudo-fator de Bayes do modelo exponencial com respeito ao modelo Weibull-exponenciada é $PSFB_{21} = 0,00044$, indicando a rejeição do modelo exponencial para o ajuste dos dados do grupo 2 da tabela 6.6, confirmando assim as conclusões obtidas com o método clássico.

Na figura 8.7, plotamos o log a razão dos CPO para os diferentes modelos com respeito ao numero de observações. Valores positivos do log da razão dos CPOs indica a preferência do primeiro modelo. Por exemplo, na figura 8.7, o gráfico para o modelo de mistura Weibull versus do modelo de mistura Weibull-exponenciada indica que 26 das 44 observações suportam o modelo de mistura Weibull-exponenciada sobre o modelo de mistura Weibull. Similarmente, 38 observações sustentam ao modelo de mistura de Weibull-exponenciada sobre o modelo de mistura exponencial.

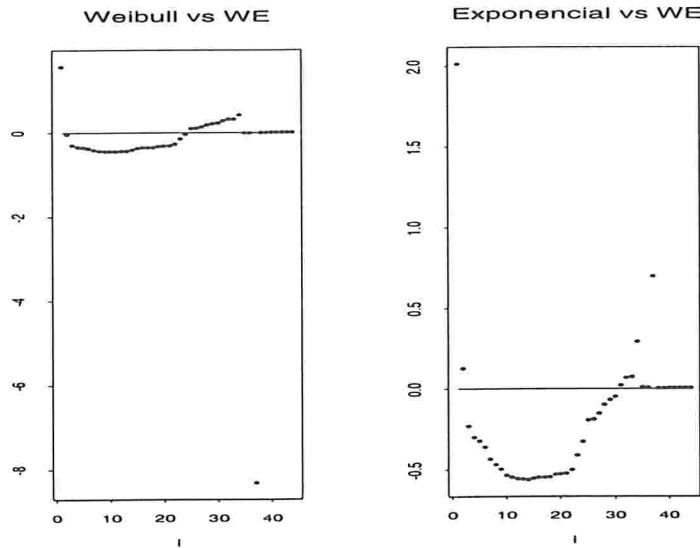


Figura 8.7: Gráfico da razão de CPO para comparar os diferentes modelos versus as observações

Conclusões

Finalizamos este trabalho, acreditando que a utilização de métodos Bayesianos aplicados à análise de dados de vida, baseados em métodos de simulação via Cadeias de Markov (MCMC), pode ser uma alternativa importante às técnicas de análise usuais.

Os algoritmos relacionados ao procedimento MCMC (algoritmo de Gibbs e Metropolis-Hasting) são de fácil implementação e não exigem um conhecimento computacional avançado, e, os resultados parecem ser mais precisos que os resultados do método de máxima verosimilhança. Portanto, podem ser usados de forma segura, além de facilitar os cálculos na inferência dos parâmetros do modelo.

Resumimos, a seguir, os principais resultados obtidos neste trabalho.

- (i) Desenvolvimento da metodologia Bayesiana para o modelo Weibull-exponenciada proposto por Mudholkar et al. (1995).
- (ii) Consideração do modelo de regressão Weibull-exponenciada. Este modelo estende o modelo proposto por Mudholkar para modelos com covariáveis.
- (iii) Desenvolvemos uma análise Bayesiana alternativa baseada em métodos MCMC, para o modelo proposto por Rodrigues, Bolfarine e Louzada-Neto (1993), para comparar populações exponenciais com dados acelerados.
- (iv) Desenvolvemos uma metodologia Bayesiana para comparar a distribuição Weibull-exponenciada com várias famílias de distribuições usadas para o ajuste de dados

vida com taxas de falha do tipo "bathtub".

- (v) Desenvolvemos uma metodologia Bayesiana para os modelos de mistura exponencial proposta por Githany et al. (1992).
- (vi) Desenvolvemos uma metodologia Bayesiana para o modelo em (v) com covariáveis associadas aos parâmetros do modelo de mistura exponencial.
- (vii) Propomos o modelo de mistura Weibull-exponenciada e utilizamos este modelo para estudar a adequabilidade dos modelos de mistura exponencial e Weibull.
- (viii) Desenvolvimento de uma metodologia para testar ordenação na comparação da proporções de sobreviventes quando vários tratamentos são considerados. Acreditamos que esta metodologia pode ser extendida para as prioris consideradas em Chen (1994).

Acreditamos que as técnicas desenvolvidas neste trabalho também podem ser aplicadas a outros modelos de sobrevivência ou confiabilidade, como os modelos de regressão Weibull-exponenciada com erros de medição nas covariadas e modelos de mistura Weibull-exponenciada com covariáveis, seja com ou sem erros de medição nas covariadas.

Bibliografia

- [1] Abramowitz, M. e Stegun, I.A. (1965). *Handbook of Mathematica functions*. New York: Dover
- [2] Achcar, J. A. and Bolfarine, H. (1986) The log-linear model with a generalized Gamma distribuition for the Error: A Bayesian approch. *Statistics and Probability Letters*, 4 325-332.
- [3] Achcar, J.L, and Louzada-Neto, F. (1992) A Bayesian approach for accelerated life tests considering the Weibull distribution, *Computational statistical*, 7, 355-369.
- [4] Achcar, J.A. and Pereira, G. A. (1998) Use of mixture of exponential power distributions for intervals-censored survival data in presence of covariates. *Notas do ICMSC*.
- [5] Aitkin, M. (1991). Posterior factor. *J. R. Stat. Soc., B*, 53, 111-142.
- [6] Anscombe, F. J. (1961). Estimating a mixed-exponential reposte law. *J. Amm. Statist. Ass.*, 56 493-502.
- [7] Arset, M. V. (1987). "How to identify bathtub hazard rate", *IEEE Trans. Reliability*, 36 106-108.
- [8] Berger, J. O. and Perechi, L. R. (1992). The intrinsic Bayes factor. Techneical report. Departament of statistics, Pardue University, Wes Lafayette.
- [9] Beerkson, J. and Gage, R.P. (1952). survival cure for cancer patients following tratment, *J. Amer. Statist. Ass.*, bf 47 501-515.

- [10] Bedrick, E.J., Christensen, R. and Johnson, W. (1996). *J. Amer. Statist. Ass.*, bf 91 1450-1460.
- [11] Boag, J.W. (1949) Maximum likelihood estimates of the proportion of patients cured by cancer therapy, *J. R. Stat. Soc., B*, 15-44.
- [12] Box, G. E. P. (1980). Sampling and Bayes inference in scientific modeling and Robustness. *J. R. Stat. Soc., A*, **143**, 383-430.
- [13] Box G. E. P. and Tiao, G. C. (1973) *Bayesian inference in statistical analysis.*, Addison-Wesley: Reading, Mass.
- [14] Broadhurst, R. G. and Maller, R. A. (1990) The recidivism of prisoners released for the first time: reconsidering the effectiveness questions. *Ast. N. Z. J. Crim.* **23**, 88-103.
- [15] Carlin, B. P. and Gelfand, A. E. (1993). Explaining the Gibbs sampler. *Amer. Statistician*, **46**, 167-174
- [16] Chen, M.-H. (1994). Importance-weighted marginal Bayesian posterior density estimation. *J. Amer. Statist. Assoc.* , **89**, 8818-824.
- [17] Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings. *Amer. Statistician*, **49**, 327-335
- [18] Collet, D. (1994). *Modelling survival data in medical research*. Chapman and Hall, London.
- [19] Cowles, M.K., and Carlin, B.P. (1996) Markov chain Monte Carlo convergence diagnostics: A comparative review. *J. Amm. Statist. Ass.*, bf 91 883-905.
- [20] Cox, D.R., and Oakes, D. (1984). *Analysis of Survival Data*. London: Chapman and Hall.

- [21] Dempster, A., Laird, N. and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Statist. Soc.*, **B**, 39, 1-38.
- [22] Devroye, L. (1986). *Non-Uniform Random Variate Generation*. Springer-Verlag. New York.
- [23] Dunsmuir, W., Tweedie, R., Flanck, L. and Mengersian, K. (1989). Modelling of transitions between employment sattes for young Australians, *austral. J. Statist.*, **31A**, 165-196.
- [24] Doornik, J. (1996). Ox: *An object-oriented matrix Programming language*. International Thomsom Business Press.
- [25] Doornik, J., Draisma, G. and Ooms, M. (1998). Ox tutorial
<http://www.nuff.ox.ac.uk/Users/Doornik>.
- [26] Fahrmeir, L. and Kaufmann, h. (1985). Consistency and asymptotic normality of the maximum likelihood estimator in generalized models, *Ann. Statist.*, **13**, 342-368.
- [27] Farewell, V. T. (1977a). A model for a binary variable with time censored observations, *Biometrika*, **64**, 43-46.
- [28] Farewell, V. T. (1977b). The commbined effect of breast cancer risk factors. *Cancer*, **40**, 931-936.
- [29] Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors, *Biometrics* **38**, 1041-1046.
- [30] Farewell, V. T. (1986). Mixture models in survival anlysis: are they worth the risk? *Canad. J. Statist.* **14**, 257-262.
- [31] Feigel, P. amd Zelen, M. (1965). Estimation of exponential survival probabilities with concomitant information. *Biometrics*, **21** 826-834.

- [32] Geisser, S., Eddy, W. (1979). A predictive approach to model selection. *J. Amer. Statist. Assoc.*, **79**, 153-160.
- [33] Gamerman, D. (1996). *Simulação estocástica via Cadeias de Markov*. 12º SINAPE-CAXAMBU-29/07 a 02/08/96.
- [34] Gelfand, A.E., Dey, D.K., and Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods (with discussion), in *Bayesian Statistics 4*, eds. J.M. Bernardo, J.O Berger, A. P Dawid, and A.F.M. Smith, Oxford:Oxford University Press, pp. 147-196.
- [35] Gelfand, A. E. and Dey, D. K. (1994). Bayesian model choice: Asymptotics and exact calculations. *J. Roy. Statist. Soc., Ser. B*, **56**, 501-514.
- [36] Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *J. Amer. Statist. Assoc.*, **85**, 398-409.
- [37] Gelfand, A. E., Smith, A. F. M., and Lee, T. M. (1992). Bayesian analysis of constrained parameter and truncated data problems using Gibbs sampling. *J. Amer. Statist. Assoc.*, **87**, 398-409.
- [38] Geman, S., and Geman, D (1984). Stochastics relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **6**, 681-889.
- [39] Gelman, A. and Rubin,, D. B. (1992). Inference from iterative simulating using multiple sequences (with discussion). *Statistical Science*, **7**, 457-511.
- [40] Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **6**, 721-741.
- [41] Ghitany, M. E. (1993) On the information matrix of exponential mixture models whit long-term survivors, *Biom. J.*, **35**, 15-27.

- [42] Ghitany, M. E. and Maller. R. A. (1992). Asymptotic results for exponential mixture models whit long-term survivors and covariates, *J. Mult. Analysis*, **49**, 218-241.
- [43] Ghitany, M. E., Maller, R. A. and Zhou, S. (1995). Exponential mixture models whit long-term survivors, *Statistics*, **23**, 231-336.
- [44] Ghitany, M. E. and Maller, R. A. (1995). Estimating the proportion of immunes in censored samples: a simulation study, *Statistics in Medicine* **14**, 39-49.
- [45] Glasser, R.E. (1980). Bathtub and realted failure rate characterization. *J. Amer. Statist. Assoc.*, **75** 667-672.
- [46] Glasser, M. (1967). Exponential survival with covariance. *J. Am. Statist. Assoc.*, **62** 651-568.
- [47] Goldman, A.I. (1984) Survivorship analysis when cure is a possibility: a Monte Carlo study, *Statistics in Medicine*, **3** 153-163.
- [48] Greenhouse, J. B. and Wolfe, R. A. (1984) A competing risks derivation of a mixture model for the analysis of survival data. *Comm. Statist.-Theor. Meth.*, **13** 3133-3154.
- [49] Kersey, J.H., Weisdorf, D.A., Nesbit, M.E., LeBien, T.W., T.W., Woods, W.G., McGlave, P.B., Kim, T.. Vallera, D.A., Goldaman, A.I.,Bostrom,B., Hurd, D. and Ramsay, N.K.C. (1987). Comparison of autologous and allogeneic bone marrow trans-paltation for treatment of high-risk refractory acute lymhoblastic leukaemia, *New England Journal of Medicine*, **317**, 461-467.
- [50] Hasting, W.K. (1970). Monte Carlo sampling methods using Markov chains and their aplications. *Biometrika*, **57**, 97-109.
- [51] Hjort, U. (1980). A realiability distributions with increasing, decreasing, constant and bathtub failure rates. *Technometrics*, **22**, 99-107.
- [52] Hsu, D. A. (1982). A Bayesian robust detection of shift in the risk structure of stock market returns. *J. Amer. Statist. Assoc.*, **77**, 29-39.

- [53] Laird, N. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, **38**, 936-974.
- [54] Lawles, J.F. (1982). *Statistical Models and Methods for Lifetime Data*, New York: John Wiley.
- [55] Leibelein. J. and Zelen, M. (1956). Statistical investigation of the fatigue life of deep groove ball bearing, *J. Res. Nat. Bur. Stand.*, **57** 273-316.
- [56] Louzada-neto, F., Bolfarine, H. e Rodrigues, J. (1991). Comparação de dois modelos exponenciais com dados acelerados: uma abordagem Bayesiana. *R. Bras. Estat.*, **52** 93-103.
- [57] Kass, R. E. and Raftery, A. E. (1995). Bayes factor. *J. Amer. Statist. Assoc.*, **90**, 773-795.
- [58] Kuo, L. e Yang, T. (1995). Bayesian Computation of software reliability. *Journal of computational and Graphical Statistics*, **4** 65-82.
- [59] Maltz, M. D. (1984). *Recidivism*. Academic Press, New York.
- [60] Maltz, M. D. And MacCleary, R. (1977). The mathematical of behavioral change: recidivism and construct validity, *Evaluation Quarterly*, **1**, 421-438.
- [61] Maller, R.A. (1988) On the exponential model for survival, *Biometrika*, **75**, 582-586.
- [62] Maller, R.A. and Zhou, S. (1992) The probability that the largest observation is censored, *J. Appl. Prob.*, **30**, 602-615.
- [63] Maller, R.A. and Zhou, S. (1993) Estimating the proportion of immunes in a censored sample, *Biometrika*, **79**, 731-739.
- [64] Maller, R.A. and Zhou, X. (1996). *Survival analysis with long-term survivors*, New York: John Wiley & Sons.

- [65] McCullagh, P., and Nelder, J. (1992). *Generalized Linear Models*, London; Chapman and Hall.
- [66] Meeker, Jr. W. Q. (1987). Limited failure population life test: applications to integrated circuit reliability. *Technometrics* **29**, 51-65.
- [67] Metropolis, N., Rosenbluth, A.W., Rosenblut, M.N., Teller, A.H., and Teller, E. (1953), Equation of state Calculations by Fast Computing Machines, *Journal of chemical physics*, **21**, 1087-1092.
- [68] Mudholkar, G.S., Srivastava, D.K., and Friemer,M. (1995) The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. *Technometrics* **37**, 436-445.
- [69] Nelson, W. B. (1990). *Accelerated Testing; statistical Models, test plans and data analysis*. New York: John Wiley.
- [70] Nelson, W.B. (1972). Graphical analysis of accelerated test data with the inverse power law, *IEE Trans. Reliab.*, **R21** 2-11.
- [71] Naylor, J. C. and Smith, A. F. M. (1982). Applications of a method for the efficient computation of posterior Distributions. *Applied Statistics*, **31**, 214-225.
- [72] Newton, M. A. and Raftery, A. E. (1994). Approximate Bayesian inference by the weighted likelihood bootstrap (with discussion). *J. Roy. Statist. Soc. Ser. B* **56**, 1-48.
- [73] Peto, R., Lee, P.N. and Paige, W.S. (1972). Statistical analysis of the bioassay of continuous carcinogens. *Br. J. Cancer*, **26** 258-261.
- [74] Pike, M.C. (1966). A method of analysis of a certain class of experiments in carcinogenesis. *Biometrics*, **35**, 142-161.
- [75] Pocock, S.J., Gore, S. M. and Kerr, G. (1982). Long-term survival analysis: the curability of breast cancer, *Statistics in Medicine*, **1**, 93-104.

- [76] Prentice, R.L. (1975). Discrimination among some parametric models. *Biometrika*, **61**, 539-544.
- [77] Raftery, A. E. (1987). Inference and prediction for a general order statistic model with unknown population size, *J. Amer. Statist. Assoc.*, **82**, 1163-1168.
- [78] Raftery, A. E. (1995). Hypothesis testing and selections. Em *Markov Chain Monte Carlo in practice* (Eds: W. R. Gilks, S. Richardson and D. J. Spiegelhalter), chap. 10. New York: Marcel Dekker.
- [79] Rajarshi, S. and Rajarshi, M.B (1988). Bathtub distributions: a review. *Commun. Statist.-Theory Meth.*, **17**(8), 2597-2621.
- [80] Rodrigues, J.; Bolfarine, H.; Neto, F.L. (1993). Comparing several accelerated life models. *Comm. Statist.-Theory Meth.* **22**(8), 2297-2308.
- [81] Rodrigues, J. (1998). Bayesian Analysis for the accelerated life tests with informative prior distributions obtained from fixed stresses. *Notas do ICMSC-USP-São Carlos*.
- [82] Stacy, E. W. (1962). A generalization of the gamma distributions. *Ann. Mathe. statist.* , **33** 1187-1192.
- [83] Smith, A.F.M. and Roberts, G.O. (1993). Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo Methods, *J. Roy. Statist. Soc., Ser. B*, **55**, 3-23.
- [84] Smith, R. M. and Bain, L. J. (1975) An exponential power life-testing distribution. *Commun. in Statist.*, **4**69-481.
- [85] Smith, R. M. and Bain, L. J. (1976). correlation-type goodness of fit sattistics with censored sampling. *Commun. in Statist.*, **4**69-481.
- [86] Spiegelhalter, D. J. and Smith, A. F. M. (1982). Bayes factor for linear and log-linear models with vague prior information, *J. R. Statist. Soc., Ser. B*, **44**, 377-387.

- [87] Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions (with discussion). *J. Roy. Stat. Soc. Ser B*, **36**, 111-147.
- [88] Struthers, C.A. and Farewell, V. T. (1989). A mixture model for time to AIDS data with left truncation and an uncertain origin, *Biometrika*. **76**, 814-817.
- [89] Tanner, M, and Wong, W. (1987). The calculating of posterior distributions by data augmentation *J. Amer. Statist. Assoc.*, **82**, 528-550.
- [90] Tierney, L. and Kandane, J.B. (1986). Accurate approximations for posterior moments and marginal densities, *J. Amer. Stat. Ass.* **81** 82-86.
- [91] Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Ann. Statist.*, **17** 1701-1762.
- [92] Yamaguchi, K. (1992). Accelerated failure time regression models with a regression model of surviving fraction: An application to the analysis of "permanent employment", *J. Amer. Statist. Assoc.*, **87**, 284-292.
- [93] Zhou, S. and Maller, R.A. (1995). The likelihood ratio test for the presence of im- munes in a censored sample, *Statistics*, **27** 181-201.

Apêndice A

Critério de convergência

Para verificarmos se as amostras geradas pelos algoritmos de Gibbs com Metropolis-Hasting estão realmente convergindo para uma distribuição estacionária, utilizamos neste trabalho o método, proposto por Gelman e Rubin (1992), o qual considera várias cadeias em paralelo partindo de valores iniciais distintos. Para cada parâmetro de interesse comparamos, a variabilidade dentro e entre as cadeias amostradas. Uma vez atingida a estacionariedade, digamos na i -ésima iteração, consideramos as realizações $(\theta_j, \theta_{j+h}, \dots, \theta_{j+nh})$, $j \geq i$ como uma amostra aleatória da distribuição desejada. Devemos considerar h razoavelmente grande de tal forma que 2 valores sucessivos de θ sejam aproximadamente independentes; assim teremos uma amostra independente e identicamente distribuído (iid) da distribuição de interesse.

A convergência é monitorada através do fator \hat{R} , introduzido por Gelman e Rubin (1992) e para tanto, consideremos $m \geq 2$ cadeias de comprimento $2n$. No caso de multimodal é recomendável inicializar pelo menos uma cadeia a partir de cada moda. Com as n , últimas observações, calculamos:

1. E : a variância entre as m cadeias:

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

onde $\bar{\theta}_{i..}$ é a média baseada nas n últimas iterações da cadeia i , $i = 1, \dots, m$ e $\bar{\theta}..$ é a média de $\bar{\theta}_{j..}, \dots, \bar{\theta}_{m..}$.

2. D : variância dentro das cadeias dadas por

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

sendo $s_i^2 = \frac{\sum_{j=1}^n (\theta_{ij} - \bar{\theta}_{i..})^2}{n-1}$.

3. Sob convergência, todos os nm valores são gerados da posteriori e a variância da distribuição a posteriori pode ser estimada de forma não viciada por

$$\hat{\sigma}^2 = \frac{1}{n} E + \frac{n-1}{n} D,$$

e sua média é estimada por $\hat{\mu} = \bar{\theta}..$

Resultado então que $\pi(\theta|D)$ tem uma distribuição aproximadamente t -student com centro em $\hat{\mu}$, escala $\sqrt{\hat{V}} = \sqrt{\hat{\sigma}^2 + E/mn}$ e graus de liberdade $df = \hat{V}^2/v\hat{a}r(\hat{V})$ onde

$$\begin{aligned} v\hat{a}r(\hat{V}) &= \left(\frac{n-1}{n}\right)^2 \frac{1}{m} v\hat{a}r(s_i^2) + \left(\frac{m+1}{mn}\right)^2 \frac{2}{m-1} E^2 + 2 \frac{(m+1)(n-1)}{mn^2} \\ &\quad \times \frac{n}{m} [c\hat{o}v(s_i^2, \bar{\theta}_{i..}^2) - 2\bar{\theta}.. c\hat{o}v(s_i^2, \bar{\theta}_{i..})] \end{aligned}$$

e onde a variância e covariância são estimadas a partir dos m valores amostrais de $\bar{\theta}_{i..}$ e s_i^2 ; $df \rightarrow \infty$ quando $n \rightarrow \infty$.

Um indicador da convergência da simulação iterativa é dada pelo fator redução de escala potencial dada por:

$$\sqrt{\hat{R}} = \sqrt{\frac{\hat{V}}{D} \frac{df}{df-2}},$$

esperando que esta quantidade convirja para 1 à medida que o número de iterações tende para o infinito ($n \rightarrow \infty$). Caso isto não ocorra, devemos considerar mais simulações para melhorar a inferência sobre a distribuição de interesse, ou seja, até que $\sqrt{\hat{R}} \approx 1$.

Apêndice B

Alguns programas

Neste apêndice, apresentamos alguns programas computacionais utilizados neste trabalho. Os programas foram desenvolvidos utilizando a *Procedure IML* (interactive matrix language) do software SAS versão 6.11.

1. Modulos para obtenção das amostras de α , θ e σ no modelo de Weibull-exponenciada, para os dados de Efron (1988).

```
PROC IML;
START INI;
USE EFRON;
READ ALL VAR{TEMPO CENSURA};
T=TEMPO/30.438;
C=CENSURA;
N=NROW(T);
R=C[+];
S0=SUM(C#((T/SIGMA)##ALPHA));
S1=SUM(C#LOG(T));
```

```

S2=SUM(C#LOG(1-EXP(-(T/SIGMA)##ALPHA)));
S3=SUM((1-C)#{1-(1-EXP(-(T/SIGMA)##ALPHA))##THETA});
L1=-R*ALPHA*LOG(SIGMA)+R*LOG(ALPHA)-S0+ALPHA*S1+(THETA-1)*S2+S3
L2=R*LOG(THETA)+ THETA*S2+S3;
L3=-R*ALPHA*LOG(SIGMA)-S0+(THETA-1)*S2+S3;
FINISH;

/* a1, a2 e a3 são os valores iniciais e cadeia e n1 é comprimento de cada cadeia. */

START NORM;

AALPHA=J(N1,1,.);TTTHETA=AALPHA;SSIGMA=TTTHETA;
AALPHA[1]=A1; TTTHETA[1]=A2; SSIGMA[1]=A3;
DO J=2 TO N1;
  ALPHA=AALPHA[J-1];
  THETA=TTTHETA[J-1];
  SIGMA=SSIGMA[J-1];
  RUNINI;
  DEN1=L1;
  ALP=RANGAM(199,ALP1)/ALP2;
  ALPHA=ALP;
  THETA=TTTHETA[J-1];
  SIGMA=SSIGMA[J-1];
  RUNINI;
  NUM1=L1;
  DEN2=L2;
  D1=NUM1-DEN1;

```

```
IF D1>=700 THEN D1=500;  
ELSE D1=D1;  
P1=MIN(D1,1);  
IF P1>=RANUNI(09) THEN AALPHA[J]=ALP;  
ELSE AALPHA[J]=AALPHA[J-1];  
THET=RANGAM(1134,THET1)/THET2;  
ALPHA=AALPHA[J];  
THETA=THET;  
SIGMA=SSIGMA[J-1];  
RUNINI;  
NUM2=L2;  
DEN3=L3;  
D2=NUM2-DEN2;  
IF D2>=700 THEN D2=500;  
ELSE D2=D2;  
P2=MIN(D2,1);  
IF P2>=RANUNI(10) THEN TTHETA[J]=THET;  
ELSE TTHETA[J]=TTHETA[J-1];  
SIG=RANGAM(2434,SIG1)/SIG2;  
ALPHA=AALPHA[J];  
THETA=TTHETA[J];  
SIGMA=SIG;  
RUNINI;  
NUM3=L3;
```

```

D3=NUM3-DEN3;
P3=MIN(D3,1);
IF P3>=RANUNI(102) THEN SSIGMA[J]=SIG;
ELSE SSIGMA[J]=SSIGMA[J-1];
END;
BB0=AALPHA[(N1/2+1):N1];
BB1=TTTHETA[(N1/2+1):N1];
BB2=SSIGMA[(N1/2+1):N1];
FINISH;

```

2. Módulo para obtenção de amostras de β_0 , β_1 , δ e θ no modelo de regressão Weibull-exponenciada utilizando o algoritmo de Gibbs com Metropolis-Hastings para os dados de Nelson (1990).

```

PROC IML;
STARTINI;
USE A;
READ ALL VAR{Y X CENSURA} ;
N=NROW(Y);
R=SUM(CENS);
Z=(Y-B0-B1*X)/DELTA;
F1=1-EXP(-EXP(Z));
S=SUM(CENS#LOG(F1));
S1=SUM((1-CENS)#LOG(1-F1##THETA));
L1=R*LOG(THETA)+THETA*S+S1;
L2=-R*LOG(DELTA)+SUM(CENS#Z)-SUM(CENS#EXP(Z))+(THETA-1)*S+S1;

```

```

L3=-R*B0/DELTA-SUM(CENS#EXP(Z))+(THETA-1)*S+S1 ;
L4= -B1*SUM(CENS#X/DELTA) -SUM(CENS#EXP(Z))+(THETA-1)*S+S1 ;
FINISH;

/* X1, X2 e X3 são os valores iniciais e N1 é o comprimento de cada cadeia. */

START NORM:

BB0=J(N1,1,.);BB1=BB0;DDELTA=BB1; TTHETA=J(N1,1,.);

BB0[1]=X1; BB1[1]=X2; DDELTA[1]=X3; TTHETA[1]=X4;

DO J=2 TO N1;

B0=BB0[J-1];
B1=BB1[J-1];
DELTA=DDELTA[J-1];
THETA=TTHETA[J-1];
RUNINI;
DEN1=L3; -----
BETA0=BB0[J-1]+SIG1*RANNOR(19);
B0=BETA0;
B1=BB1[J-1];
DELTA=DDELTA[J-1];
THETA=TTHETA[J-1];
RUNINI;
NUM1=L3;
DEN2=L4;
D1=NUM1-DEN1;
IF D1<=700 THEN D1=D1;

```

```

ELSE D1=500;

Q1=EXP(D1);

P1=MIN(Q1,1);

IF P1>=RANUNI(09) THEN BB0[J]=B0;

ELSE BB0[J]=BB0[J-1];

BETA1=BB[J-1]+SIG2*RANNOR(13);

B0=BB0[J];

B1=BETA1;

DELTA=DDELTA[J-1];

THETA=TTHETA[J-1];

RUNINI;

NUM2=L4;

DEN3=L2;

D2=NUM2-DEN2;

IF D2<=700 THEN D2=D2;

ELSE D2=500;

Q2=EXP(D2);

P2=MIN(Q2,1);

IF P2>=RANUNI(10) THEN BB1[J]=B1;

ELSE BB1[J]=BB1[J-1];

DEL=RANGAM(123,DEL1)/DEL2;

B0=BB0[J];

B1=BB1[J];

DELTA=DEL;

```

```

THETA=TTHETA[J-1];
RUNINI;
NUM3=L2;
DEN4=L1;
D3=NUM3-DEN3;
IF D3<=700 THEN D3=D3;
ELSE D3=500;
Q3=EXP(D3);
P3=MIN(Q3,1);
IF P3>=RANUNI(102) THEN DDELTA[J]=DEL;
ELSE DDELTA[J]=DDELTA[J-1];
THE=RANGAM(112,THE1)/THE2;
B0=BB0[J];
B1=BB1[J];
DELTA=DDELTA[J];
THETA=THE;
RUNINI;
NUM4=L1;
D4=NUM4-DEN4;
IF D4<=700 THEN D4=D4;
ELSE D4=500;
Q4=EXP(D4);
P4=MIN(Q4,1);
IF P4>=RANUNI(102) THEN TTHETA[J]=THE;

```

```

ELSE TTHETA[J]=TTHETA[J-1];
END;
BB0=BB0[(N1/2+1):N1];
BB1=BB1[(N1/2+1):N1];
DDELTA=DDELTA[(N1/2+1):N1];
TTHETA= TTHETA[(N1/2+1):N1];
FINISH;

```

3. Modulo para gerar amostras de α_{11} , α_{21} , α_{12} e α_{22} , sob H_1 na comparação duas amostras exponenciais com dados acelerados utilizando o algoritmo de Gibbs com Metropolis-Hasting.

```

PROC IML;
STARTINI;
R={4,6,7,8,9};
R1=R[+];
V={5,10,15,20,25};
A11={291.9,76.67,67.29,51.75,71.39};
A12={271.34,79.83,74.14,54.75,91.94};
S11=SUM(R#A11#(V##ALPHA21));
S12=SUM(R#A12#(V##ALPHA22));
L21=ALPHA21*(SUM(R#LOG(V))+4*LOG(5)+9*LOG(25))
-S11/ALPHA11+ALPHA21**2/2000;
L22=ALPHA22*(SUM(R#LOG(V))+4*LOG(5)+9*LOG(25))
-S12/ALPHA12+ALPHA22**2/2000;
FINISH;

```

```

/* X[1], X[2], X[3] e X[4] são os valores iniciais para cada cadeia e n1 é comprimento
de cada cadeia. */

START METROP;

X11=J(N1,1,..); X21=X11; X12=X21;X22=X12;
X11[1]=X[1]; X21[1]=X[2];X12[1]=X[3]; X22[1]=X[4];

DO J=2 TO N1;
    ALPHA11=X11[J-1];
    ALPHA21=X21[J-1];
    ALPHA12=X12[J-1];
    ALPHA22=X22[J-1];
    RUNINI;
    AL11=S11/RANGAM(4234,R1+13);
    ALPHA11=AL11;
    ALPHA21=X21[J-1];
    ALPHA12=X12[J-1];
    ALPHA22=X22[J-1];
    RUNINI;
    DEN21=L21;
    AL21=SIG1*RANNOR(1234)+X21[J-1];
    ALPHA11=AL11;
    ALPHA21=AL21;
    ALPHA12=X12[J-1];
    ALPHA22=X22[J-1];
    RUNINI;

```

```
NUM21=L21;  
D21=NUM21-DEN21;  
IF D21<=700 THEN D21=D21;  
ELSE D21=500;  
P21=MIN(1,EXP(D21));  
IF P21>=RANUNI(1234) THEN AL21=AL21;  
ELSE AL21=X21[J-1];  
ALPHA11=AL11;  
ALPHA21=AL21;  
ALPHA12=X12[J-1];  
ALPHA22=X22[J-1];  
RUNINI;  
AL12=S12/RANGAM(1234,R1+13);  
ALPHA11=AL11;  
ALPHA21=AL21;  
ALPHA12=AL12;  
ALPHA22=X22[J-1];  
RUNINI;  
DEN22=L22;  
AL22=SIG2*RANNOR(1234)+X22[j-1];  
ALPHA11=AL11;  
ALPHA21=AL21;  
ALPHA12=AL12;  
ALPHA22=AL22;
```

```

RUNINI;
NUM22=L22;
D22=NUM22-DEN22;
IF D22<=700 THEN D22=D22;
ELSE D22=500;
P22=MIN(1,EXP(D22));
IF P22>=RANUNI(999) THEN AL22=AL22;
ELSE AL22=X22[J-1];
X11[J]=AL11;
X21[J]=AL21;
X12[J]=AL12;
X22[J]=AL22;
END;
X11=X11[(N1/2+1):N1];
X21=X21[(N1/2+1):N1];
X12=X12[(N1/2+1):N1];
X22=X22[(N1/2+1):N1];
FINISH;

```

4. Modulo para a obtenção das amostras λ e p no modelo de mistura exponencial utilizando o algoritmo de Gibbs com Metropolis-Hasting.

```

PROC IML;
USE A;
READ ALL VART C;
STARTINI;

```

```

R=SUM(C);
P=EXP(BETA)/(1+EXP(BETA));
S=SUM((1-C)#LOG(1-P+P*EXP(-T*EXP(LAMBDA))));;
L1=(R+1)*LOG(P)-0.5*LOG(1-P)+S-LOG(1+EXP(BETA));
L2=(R+2)*LAMBDA-SUM(C#T)*EXP(LAMBDA)+S;
FINISH;

START METROP;
LAMB=J(N1,1,.); BBETA=J(N1,1,.);
LAMB[1]=X[1]; BBETA[1]=X[2];
DO J= 2 TO N1;
LAMBDA=LAMB[J-1];
BETA=BBETA[J-1];
RUNINI;
DEN1=L1;
BETA1=BBETA[J-1]+RANNOR(234)*SIG1;
LAMBDA=LAMB[J-1];
BETA=BETA1;
RUNINI;
NUM1=L1;
DEN2=L2;
D1=NUM1-DEN1;
IF D1>=700 THEN D1=500;
ELSE D1=D1;
Q1=EXP(D1);

```

```

P1=MIN(1,Q1);

IF P1>=RANUNI(123) THEN BBETA[J]=BETA1;
ELSE BBETA[J]=BBETA[J-1];
LAM1=LAMB[J-1]+RANNOR(13455)*SIG2;
LAMBDA=LAM1;
BETA=BBETA[J];
RUNINI;
NUM2=L2;
D2=NUM2-DEN2;
IF D2>=700 THEN D2=500;
ELSE D2=D2;
Q2=EXP(D2);
P2=MIN(1,Q2);
IF P2>=RANUNI(9901) THEN LAMB[J]=LAM1;
ELSE LAMB[J]=LAMB[J-1];
END;
B1=EXP(BBETA[(N1/2+1):N1])/(1+EXP(BBETA[(N1/2+1):N1]));
B2=EXP(LAMB[(N1/2+1):N1]);
FINISH;

```

5. Modulo para gerar amostras de α_0 , α_1 , β_1 e β_2 , no modelo de mistura exponencial com covariáveis.

```

PROCIML;
USEA;
READALLVAR{T C X};

```

```

STARTINI;
R=SUM(C);
LAMBDA=EXP(ALPHA0+ALPHA1#X);
P=EXP(BETA0+BETA1*X)/(1+EXP(BETA0+BETA1*X));
PSI1=SUM((1-C)#LOG(1-P+P#EXP(-LAMBDA#T)));
PSI2=SUM(C#LOG(P));
S1=SUM(C#T#LAMBDA);
L0=ALPHA0*R-S1-0.5*ALPHA0**2 /1000000+ PSI1;
L1=SUM(ALPHA1#X#C)-S1-0.5*ALPHA1**2/1000000+PSI1;
LL0=-0.5*BETA0**2/100000+PSI1+PSI2;
LL1=-0.5*BETA1**2/100000+PSI1+PSI2;
FINISH;

STARTMETROP;
AALPHA0=J(N1,1,.);AALPHA1=J(N1,1,.);BBETA0=J(N1,1,.);BBETA1=J(N1,1,.);
AALPHA0[1]=A1; AALPHA1[1]=A2; BBETA0[1]=B1;BBETA1[1]=B2;
DO J= 2 TO N1;
ALPHA0=AALPHA0[J-1];
ALPHA1=AALPHA1[J-1];
BETA0=BBETA0[J-1];
BETA1=BBETA1[J-1];
RUNINI;
DEN1=L0;
AL0=AALPHA0[J-1]+RANNOR(12323)*SIG1;
ALPHA0=AL0;

```

```

ALPHA1=AALPHA1[J-1];
BETA0=BBETA0[J-1];
BETA1=BBETA1[J-1];
RUN INI;
NUM1=L0;
DEN2=L1;
D1=NUM1-DEN1;
IF D1>=700 THEN D1=500;
ELSE D1=D1;
Q1=EXP(D1);
P1=MIN(1,Q1);
IF P1>=RANUNI(123) THEN AALPHA0[J]=AL0;
ELSE AALPHA0[J]=AALPHA0[J-1];
AL1=AALPHA1[J-1]+RANNOR(13455)*SIG2;
ALPHA0=AALPHA0[J];
ALPHA1=AL1;
BETA0=BBETA0[J-1];
BETA1=BBETA1[J-1];
RUN INI;
NUM2=L1;
DEN3=LL0;
D2=NUM2-DEN2;
IF D2>=700 THEN D2=500;
ELSE D2=D2;

```

```

Q2=EXP(D2);

P2=MIN(1,Q2);

IF P2>=RANUNI(9901) THEN AALPHA1[J]=AL1;

ELSE AALPHA1[J]=AALPHA1[J-1];

BE0=BBETA0[J-1]+RANNOR(656)*SIG2;

ALPHA0=AALPHA0[J];

ALPHA1=AALPHA1[J];

BETA0=BE0;

BETA1=BBETA1[J-1];

RUNINI;

NUM3=LL0;

DEN4=LL1;

D3=NUM3-DEN3;

IF D3>=700 THEN D3=500;

ELSE D3=D3;

Q3=EXP(D3);

P3=MIN(1,Q3);

IF P3>=RANUNI(1020) THEN BBETA0[J]=BE0;

ELSE BBETA0[J]=BBETA0[J-1];

BE1=BBETA1[J-1]+RANNOR(656)*SIG2;

ALPHA0=AALPHA0[J];

ALPHA1=AALPHA1[J];

BETA0=BBETA0[J];

BETA1=BE1;

```

```

RUN INI;
NUM4=LL1;
D4=NUM4-DEN4;
IF D4>700 THEN D4=500;
ELSE D4=D4;
Q4=EXP(D4);
P4=MIN(1,Q4);
IF P4>=RANUNI(1020) THEN BBETA1[J]=BE1;
ELSE BBETA1[J]=BBETA1[J-1];
END;
AA0=AALPHA0[(N1/2+1):N1];
AA1=AALPHA1[(N1/2+1):N1] ;
BB0=BBETA0[(N1/2+1):N1] ;
BB1=BBETA1[(N1/2+1):N1] ;
FINISH;

```

6. Programa desenvolvido no pacote Ox para obtenção das estimativas de maxima verossimilhança e o valor da estatística para testar $H_0 : p = 1$ versus $H_1 : p < 1$ no modelo de mistura exponencial.

```

#include<oxstd.h>
#include<oxdraw.h>
#include<oxfloat.h>
#include<maximize.h>
#include<simula.h>
#pragma link("maximize.oxo")

```

```

static decl g_mY;

log_vero(const vP,const adFunc, const avScore,const amHessian)
{
    decl n=rows(g_mY);

    decl cont,t,c;

    decl uns=ones(n,1);

    decl vero=zeros(1,n);

    decl a=abs(vP[0][0]);

    decl p=abs(vP[1][0]);

    for(cont=0;cont<n;++cont) {
        t=g_mY[cont][0];
        c=g_mY[cont][1];
        vero[0][cont]=c*(log(a)+log(p)-a*t)+(1-c)*log(1-p+p*exp(-1*a*t));
    }

    adFunc[0]=double(vero*uns);

    if(isnan(adFunc[0]))
        return 0;

    else
        return 1;
}

main()
{
    decl tempo=loadmat("a:tempo.dat",1);
    decl censura=loadmat("a:censura.dat",1);
}

```

```

g_mY= tempo|censura

decl lah0=sumc(g_mY[][1])/sumc(g_mY[][0]);

decl lvh0=sumc(g_mY[][1])*log(lah0)-lah0*sumc(g_mY[][0]);

print(rows(g_mY),lah0);

decl dfunc;

decl vP=< 1; 0.6 >; 

MaxControl(-1,20);

decl mhess=0.000001*unit(2);

decl ir;

ir=MaxBFGS(log_vero,&vP,&dfunc,&mhess,1);

Num2Derivative(log_vero,vP,&mhess);

if(ir==MAX_CONV||ir==MAX_WEAK_CONV)

{

print("Os EMV para modelo de mistura exponencial", vP);

decl hes=(-1/mhess));

print("A matriz de covariância ",hes)

print("O valor da estatística do teste",-2*(lvh0-dfunc));

}

}

```