

Análise de referência Bayesiana para o modelo Weibull na aplicação de riscos competitivos

Camila Bertini Martins

Análise de referência Bayesiana para o modelo Weibull na aplicação de riscos competitivos

Camila Bertini Martins

Orientadora: Prof^ª. Dr^ª. Vera Lucia D. Tomazella

Co-orientador: Prof. Dr. Adriano Polpo

Dissertação apresentada ao Departamento de Estatística da Universidade Federal de São Carlos - DEs/UFSCar, para obtenção do título de Mestre em Estatística.

UFSCar - São Carlos

Abril/2009

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

M386ar

Martins, Camila Bertini.

Análise de referência Bayesiana para o modelo Weibull na aplicação de riscos competitivos / Camila Bertini Martins. -- São Carlos : UFSCar, 2009. 109 f.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2009.

1. Estatística. 2. Análise de referência bayesiana. 3. Distribuição Weibull. 4. Confiabilidade. I. Título.

CDD: 519.5 (20^a)

*“ A razão cardeal de toda a superioridade humana
é sem dúvida a vontade. O poder nasce do querer.*

*Sempre que o homem aplique a veemência e
perseverante energia de sua alma a um fim, ele
vencerá os obstáculos, e se não atingir o alvo, fará
pelo menos coisas admiráveis. ”*

José de Alencar

Aos meus pais Sonia Bertini Martins e Antônio José Martins

À minha irmã Ana Carolina e à minha avó Maria Aparecida

Agradecimentos

A Deus, por me dar forças durante esses anos de estudos.

À minha orientadora, Prof^a. Dr^a. Vera L. D. Tomazella, por acreditar em mim, apoiando-me e incentivando-me sempre, pela sua compreensão, dedicação e amizade.

Ao meu orientador, Prof. Dr. Adriano Polpo de Campos, pelo apoio, compreensão e, principalmente, pela amizade.

Ao Prof. Dr. Josemar Rodrigues e ao Prof. Dr. Mário de Castro pelas valiosas sugestões e comentários dados na realização do meu exame de qualificação.

Ao meu namorado, Mateus V. Polizeli, pelo carinho e dedicação.

À minha amiga, Caroline C. Biazolli, pela amizade e pelo auxílio na correção deste texto.

A todos os meus amigos, pelo companheirismo, pela ajuda nas horas difíceis e pelos momentos de descontração.

A todos os professores e funcionários do Departamento de Estatística da UFSCar que contribuíram para minha formação.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), pelo apoio financeiro (Processo 2007/01770-3).

Resumo

Há situações em que existem diversos fatores de risco de falha presentes ao mesmo tempo na vida de um sistema. Por essa razão, dizemos que esses fatores estão competindo para provocar a falha do mesmo. Entretanto, apenas um desses competidores é o responsável por determinada falha. O comportamento dessa falha é, na maioria das vezes, representado pela sua taxa de falha, que pode ser crescente, decrescente, constante ou fruto de suas combinações ao longo do tempo. Assim, é desejável o uso de um modelo probabilístico que represente cada uma dessas situações, apenas com mudanças nos valores dos seus parâmetros. Neste trabalho, estudamos, sob a perspectiva de análise de referência Bayesiana, a aplicação de riscos competitivos a partir do modelo Weibull, considerado bastante flexível. A análise de referência Bayesiana é um método de produzir afirmações inferenciais que dependem apenas do modelo assumido e dos dados observados (Bernardo, 1979). O objetivo é encontrar uma específica função a priori de referência conjunta para os parâmetros desconhecidos do modelo Weibull na aplicação de riscos competitivos e uma distribuição a posteriori de referência marginal para os parâmetros de interesse, a qual será dominada pelos dados observados. As distribuições a posteriori de referência são obtidas através do uso formal do teorema de Bayes com a função a priori de referência, podendo ser utilizadas para estimativas pontuais e testes de hipóteses, proporcionando um conjunto unificado de soluções Bayesianas objetivas para o nosso problema.

Abstract

There are situations where various risk factors of failure are present, in the same time, in the life of system. For this reason we say that these factors are competing to cause the system failure. However, only one of these competitors is responsible for the system failure. The failure behavior of one system is, in most times, represented for its failure rate, which may be increasing, decreasing, remain constant or be combinations of these over time. Therefore it is desirable to use a probabilistic model that only with changes in the values of the parameters representing each of these situations. In this work we studied from the perspective of Bayesian reference analysis to the application of competitive risks under the Weibull model due to high flexibility of this model. The reference analysis is a method to produce Bayesian inferential statements which only depend on the assumed model and the available data (Bernardo, 1979). The goal is to find a specific joint reference prior function for all the unknown parameters of Weibull model in the application of competitive risks and a marginal reference posterior to the parameters of interest, which is always dominated by the observed data. The reference posterior distributions are obtained through the use of the Bayes theorem with the reference prior function that can be used to point estimates and tests of hypotheses, providing a unified set of Bayesian objective solutions for our problem.

Sumário

1	Introdução	1
1.1	Introdução à análise de confiabilidade	2
1.2	Modelo Weibull	4
1.3	Objetivos do trabalho	8
2	Análise de referência	10
2.1	Construção da função a priori de referência	14
2.1.1	Caso de um parâmetro	14
2.1.2	Caso de um parâmetro perturbador	16
2.1.3	Caso multiparamétrico	20
2.2	Teste de hipóteses Bayesiano	23
2.2.1	Teste de hipóteses como um problema de decisão	24
2.2.2	Critério de Referência Bayesiano (BRC)	25
2.3	Análise de referência para o modelo Weibull	29
2.3.1	Teste de hipótese para o parâmetro de forma (β) do modelo Weibull	36
2.3.2	Inferência para a função de confiabilidade do modelo Weibull	38
3	Riscos competitivos	45
3.1	Introdução	45

3.1.1	Definição de riscos competitivos	46
3.2	Função de verossimilhança	48
3.3	O modelo Weibull na aplicação de riscos competitivos	49
3.4	Análise de referência	50
3.5	Critério de referência Bayesiano para igualdade dos parâmetros	55
3.6	Análise de referência para a confiabilidade em modelos de riscos competitivos	56
4	Aplicações	63
4.1	Estudo de simulação 1	63
4.2	Estudo de simulação 2	71
4.3	Exemplo com dados reais	78
4.4	Estudo de simulação 3	85
5	Considerações Finais	97
	Referências Bibliográficas	99
A	Dados Simulados	102
A.1	Conjunto de dados do Exemplo 2.1	102
A.2	Dados do Estudo de simulação 1	102
A.3	Dados do Estudo de simulação 2	105
A.4	Dados do Estudo de simulação 3	108

Capítulo 1

Introdução

A teoria para dados de sobrevivência/confiabilidade tem sido bastante desenvolvida com o objetivo de estudar a função de risco/sobrevivência de um indivíduo ou sistema. Essa metodologia permite determinar quais variáveis afetam a forma da função de risco e obter estimativas dessas funções para cada indivíduo ou componente. Este estudo envolve o acompanhamento de componentes (indivíduos) até a ocorrência de algum evento de interesse, por exemplo, a falha (morte) do mesmo.

Uma situação comum é que indivíduos ou sistemas estão sujeitos à morte ou falha, provocada por diferentes fatores de risco que competem para provocar a falha dos mesmos. No entanto, supõe-se que apenas um desses competidores seja o responsável por determinada falha.

Em geral, o comportamento da falha (ou morte) de um componente é representado pela sua taxa de falha. Por sua vez, essa taxa pode ser crescente, decrescente, constante ou fruto dessas combinações ao longo do tempo, principalmente quando o sistema está exposto a diferentes fatores de risco. Assim, é desejável o uso de um modelo probabilístico que com apenas mudanças nos valores de seus parâmetros represente cada uma dessas situações. Por exemplo, enquanto a Weibull permite risco crescente, decrescente ou constante, a distribuição Exponencial indica risco constante e a distribuição Log-Logística admite risco unimodal. Como a família de distribuições Weibull pode representar

algumas das mais diversas situações, utilizamo-la no trabalho para descrever a confiabilidade de um sistema sob diferentes fatores de risco.

Neste trabalho, abordamos o problema de sistemas com componentes em série, também conhecido como o problema de riscos competitivos, para estimar a curva de confiabilidade de cada um dos componentes do sistema. A idéia principal é utilizar o modelo paramétrico Weibull para as curvas de confiabilidade e estimá-las através do método Bayesiano, utilizando-se funções a priori de referência.

Neste capítulo, Seção 1.1, apresentamos uma introdução à análise de confiabilidade. Como nos interessa o modelo Weibull, descrevemo-lo na Seção 1.2, e na Seção 1.3, apresentamos os objetivos e o conteúdo deste trabalho.

1.1 Introdução à análise de confiabilidade

A análise de confiabilidade/sobrevivência tem sido bastante desenvolvida nestas últimas décadas devido ao desenvolvimento e aprimoramento de técnicas estatísticas combinados com o avanço computacional.

A análise de sobrevivência reúne um conjunto de métodos estatísticos úteis na análise de dados relacionados ao tempo até a ocorrência de um evento de interesse. Esses tempos são chamados de tempos de sobrevivência, de falha ou simplesmente de tempos de vida. Usualmente, a análise de sobrevivência diz respeito a dados biomédicos, enquanto a análise de confiabilidade se refere à área industrial.

A principal característica de dados de sobrevivência/confiabilidade é a presença de censura, que se resume à observação incompleta da variável tempo, causada por algum fator externo. Isto é, o acompanhamento de algum elemento foi interrompido por uma causa diferente da esperada: se essa fosse a morte do paciente por câncer, ele poderia ter mudado de cidade, o estudo poderia ter terminado para a análise dos dados ou ele ainda poderia ter morrido devido a outra causa. Se uma observação é censurada, toda a informação referente à resposta reside no conhecimento de que o tempo de falha é superior àquele

observado (Colosimo e Giolo, 2006).

Os dados de sobrevivência podem reunir não só os tempos de vida, mas também um conjunto de variáveis observáveis que podem ou não estar relacionadas com esses tempos. Essas variáveis são conhecidas como covariáveis ou variáveis explicativas.

O comportamento da variável aleatória contínua e não negativa, tempo de sobrevivência, $X \geq 0$, pode ser definido pelas seguintes funções: função densidade de probabilidade ($f(x)$), função de confiabilidade ($R(x)$) ou sobrevivência ($S(x)$), e função de risco ($h(x)$).

Essas três funções podem ser utilizadas para descrever os diferentes aspectos apresentados pelo conjunto de dados de tempos de sobrevivência (Lawless, 1982).

A função densidade de probabilidade é definida como o limite da probabilidade de um indivíduo falhar (morrer) no intervalo de tempo $[x, x + \Delta x]$ por unidade de tempo, e é expressa por

$$f(x) = \lim_{\Delta x \rightarrow 0^+} \frac{P(x \leq X < x + \Delta x)}{\Delta x},$$

em que $f(x) \geq 0$, para todo x .

A função de sobrevivência $S(x)$ ou confiabilidade $R(x)$ é definida como a probabilidade de um indivíduo sobreviver mais que um determinado tempo x , ou seja, não falhar após um certo tempo, e é expressa por

$$R(x) = P(X > x) = 1 - F(x),$$

em que $F(x) = \int_0^x f(u) du$, representa a função de distribuição acumulada. Quando $x \rightarrow \infty$, $R(x) = 0$ e quando $x = 0$, $R(x) = 1$.

Através da curva de confiabilidade, podemos comparar, por exemplo, dois ou mais fatores de risco de um sistema. Assim, uma curva de confiabilidade com declive acentuado representa curto tempo de vida, enquanto uma curva com aclive representa longo tempo de vida.

A função de risco $h(x)$, também conhecida como taxa de falha, é definida como o limite da probabilidade de um indivíduo falhar no intervalo de tempo

$[x, x + \Delta x]$, dado que o mesmo tenha sobrevivido até o tempo x , e é expressa por

$$h(x) = \lim_{\Delta x \rightarrow 0^+} \frac{P(x \leq X < x + \Delta x | X \geq x)}{\Delta x} = \frac{f(x)}{S(x)}.$$

Essa função analisa a probabilidade de falha (risco) de um determinado indivíduo e descreve como a probabilidade instantânea de falha se modifica com o passar do tempo. Pode apresentar diferentes comportamentos, como: crescente, decrescente, constante ou combinada.

Como o modelo Weibull é bastante flexível, acomodando diferentes formas da função de risco, utilizamo-lo neste trabalho para descrever e estudar as funções de confiabilidade. Na próxima seção, descrevemos o modelo Weibull.

1.2 Modelo Weibull

A distribuição Weibull é amplamente utilizada em confiabilidade (Martz e Waller, 1982). Foi apresentada originalmente por Weibull (1939) e utilizada em confiabilidade posteriormente por Weibull (1951) e, desde então, tem sido amplamente utilizada como modelo em aplicações biométricas, industriais e em muitas outras áreas. Uma das características importantes da distribuição Weibull na modelagem de tempos de sobrevivência ou confiabilidade está relacionada à sua flexibilidade em acomodar diferentes formas na taxa de falha (função de risco), apenas variando os valores de seus parâmetros.

Para o tempo $x \geq 0$, dados os parâmetros de forma ($\beta > 0$) e escala ($\eta > 0$), as funções densidade, de confiabilidade e de risco da distribuição Weibull ($Wei(x|\beta, \eta)$) são dadas, respectivamente, por

$$f(x|\beta, \eta) = \frac{\beta}{\eta^\beta} x^{\beta-1} \exp \left[- \left(\frac{x}{\eta} \right)^\beta \right], \quad (1.1)$$

$$R(x|\beta, \eta) = 1 - F(x|\beta, \eta) = \exp \left[- \left(\frac{x}{\eta} \right)^\beta \right] \quad (1.2)$$

e

$$h(x|\beta, \eta) = \frac{\beta}{\eta^\beta} x^{\beta-1}. \quad (1.3)$$

A função de verossimilhança para n dados de sobrevivência utilizando o modelo Weibull é expressa por

$$L(\beta, \eta | \mathbf{x}) = \prod_{i=1}^n \frac{\beta}{\eta^\beta} x_i^{\beta-1} \exp \left[- \left(\frac{x_i}{\eta} \right)^\beta \right]. \quad (1.4)$$

A média e a variância da distribuição Weibull, são dadas por

$$E(x|\beta, \eta) = \frac{\Gamma(1 + \eta^{-1})}{\beta^{1/\eta}},$$

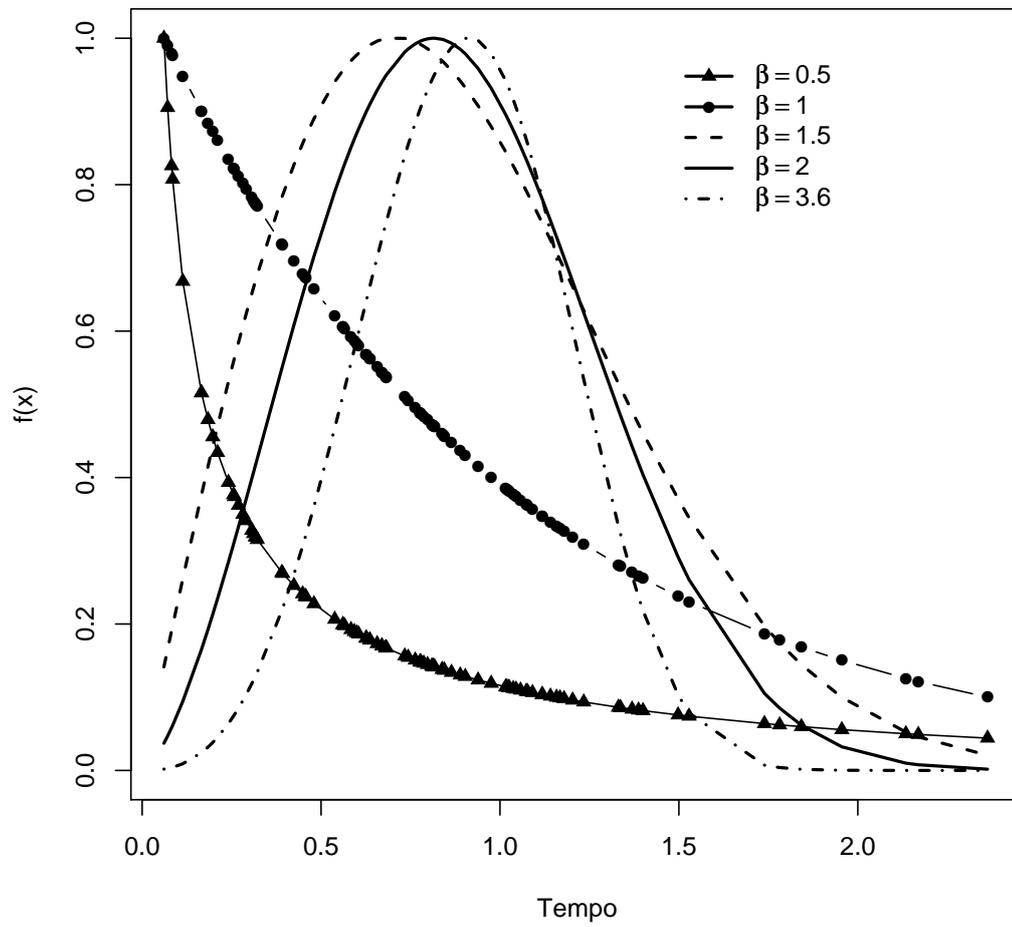
em que $\Gamma(\cdot)$ é a função gama, e

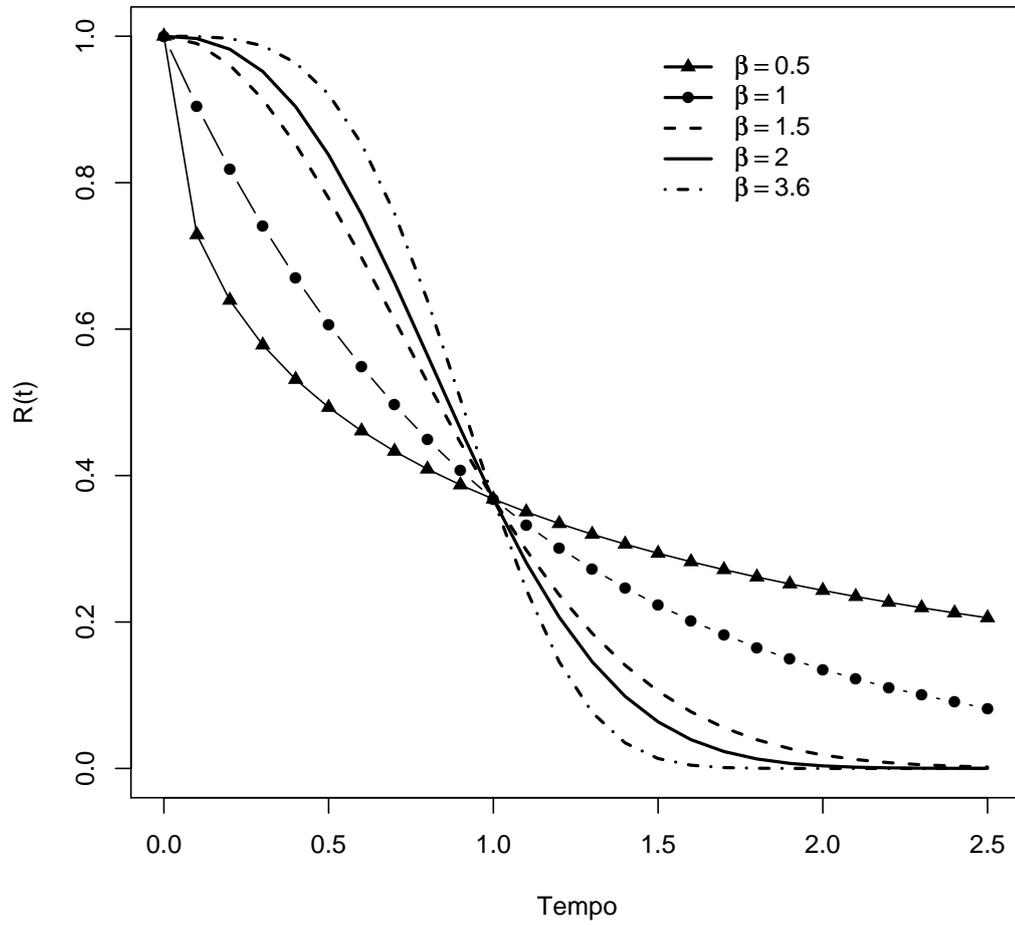
$$Var(x|\beta, \eta) = \frac{\Gamma(1 + 2\eta^{-1}) - \Gamma^2(1 + \eta^{-1})}{\beta^{2/\eta}}.$$

Para obtermos uma variedade de comportamentos dessa distribuição, basta alterar o parâmetro de forma β . Determinados valores desse parâmetro fornecem alguns casos particulares, por exemplo: $\beta = 1$ nos dá a distribuição Exponencial; $\beta = 2$, a distribuição Rayleigh, e obtemos uma aproximação das distribuições Log-Normal e Normal para $\beta = 2, 5$ e $\beta = 3, 6$, respectivamente.

Os comportamentos das funções de densidade, de confiabilidade e de risco do modelo Weibull estão representados nas Figuras 1.1, 1.2 e 1.3. Os casos particulares e aproximações do comportamento da distribuição Weibull, quando variamos β , estão apresentados na Figura 1.1.

Através da Figura 1.2, verificamos o comportamento da função de confiabilidade Weibull para diferentes valores de β . Vale ressaltar que, independente do valor de β , quando $x = \eta$, temos $R(x) = e^{-1}$. Observamos na Figura 1.3 as principais formas da taxa de falha para a distribuição Weibull, que podem ser descritas por $\beta < 1$ (taxa de falha decrescente), $\beta = 1$ (taxa de falha constante) e $\beta > 1$ (taxa de falha crescente).

FIGURA 1.1: Função densidade do modelo Weibull ($\eta = 1$).

FIGURA 1.2: Função de confiabilidade do modelo Weibull ($\eta = 1$).

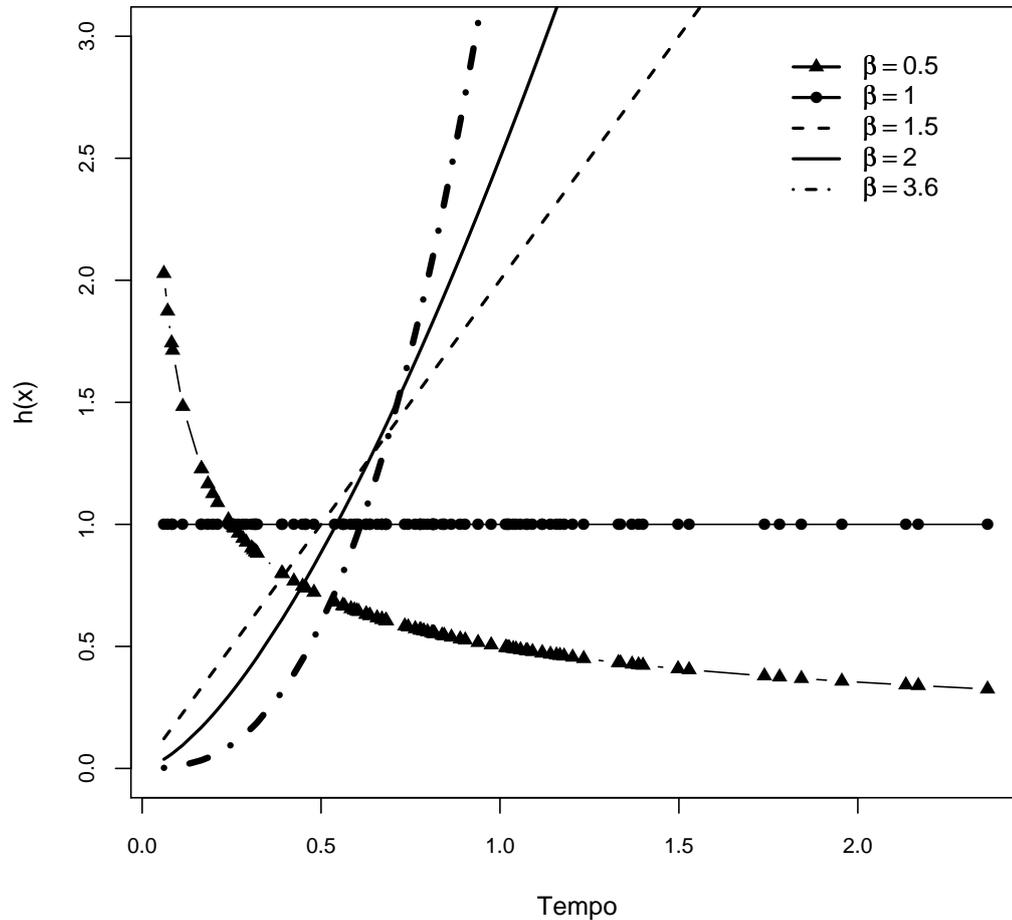


FIGURA 1.3: Função de risco do modelo Weibull ($\eta = 1$).

Na próxima seção apresentamos os objetivos deste trabalho.

1.3 Objetivos do trabalho

Os objetivos deste trabalho são: desenvolver os estimadores paramétricos das distribuições dos componentes em série (riscos competitivos), descrever o problema baseado na família de distribuições Weibull, utilizar a metodologia de análise de referência Bayesiana nos processos decisórios de estimação e testes de significância estatística, estruturação e definição dos estimadores do caso particular de duas potenciais causas de falha independentes, disponibilizando uma opção de estimação aos pesquisadores em geral e aplicações da teoria em casos práticos

reais que se encaixem em suas características.

Enfatizamos o objetivo de estudar, sob a perspectiva da análise de referência Bayesiana objetiva, a aplicação do modelo Weibull em problemas de confiabilidade sujeitos a fatores de risco competitivos.

Reconhecemos que, sob a perspectiva Bayesiana, a distribuição a posteriori para a quantidade de interesse representa a mais completa inferência que podemos fazer a respeito dessa quantidade. A distribuição a posteriori combina a informação contida nos dados com a informação prévia que possuíamos sobre a quantidade de interesse. O método de análise de referência Bayesiana foi inicialmente proposto por Bernardo (1979) e posteriormente aprimorado por Berger e Bernardo em uma série de trabalhos (Berger e Bernardo, 1989; Berger e Bernardo, 1992a; Berger e Bernardo, 1992b). O método de análise de referência generaliza muitos dos métodos alternativos propostos e inclui, como casos especiais, as soluções apropriadas que tais métodos apresentam em contextos específicos.

Como a operação Bayesiana muitas vezes é de difícil execução, utilizamos o método aproximado de simulação MCMC (Monte Carlo via Cadeias de Markov).

Este trabalho está organizado da seguinte maneira. No Capítulo 2, revisamos a metodologia de análise de referência introduzida por Bernardo (1979) e a aplicamos ao modelo Weibull. No Capítulo 3, revisamos o conceito de riscos competitivos e apresentamos a análise de referência para o modelo Weibull na aplicação de riscos competitivos. No Capítulo 4, apresentamos estudos de simulação e um exemplo com dados reais, utilizando a metodologia descrita ao longo do trabalho. E no Capítulo 5, se encontram as considerações finais.

Capítulo 2

Análise de referência

Neste capítulo, revisamos a metodologia de análise de referência introduzida por Bernardo (1979), descrevendo suas principais características e dando ênfase às propriedades de interesse para o trabalho. Mais detalhes se encontram em Bernardo (2005).

Os problemas de inferência estatística são normalmente resolvidos condicionados à suposição de que um determinado modelo estatístico é uma descrição adequada do mecanismo probabilístico que gerou os dados, bem como a escolha do modelo que naturalmente envolve um elemento de subjetividade. Isso se tornou uma prática habitual. No entanto, descrevemos como objetiva qualquer análise estatística que depende apenas do modelo assumido e dos dados observados. Nesse sentido, a análise de referência é um método de produzir inferência Bayesiana “objetiva”.

Argumentos fundacionais (Bernardo e Smith, 1994) ditam que os cientistas devem elicitar uma única distribuição a priori conjunta para todos os elementos desconhecidos do problema com base em informações disponíveis e usar o teorema de Bayes para combinar essa distribuição com as informações fornecidas pelos dados, encapsuladas na função de verossimilhança. Nesse contexto, a utilização ingênua de simples distribuições a priori “flats”, como sendo “não informativas”, frequentemente oculta suposições importantes que podem facilmente dominar ou mesmo invalidar a análise. O uso não crítico de tais

distribuições deve ser fortemente desestimulado, devendo se utilizar uma função a priori de referência adequada. Uma candidata natural nesses casos seria uma distribuição a priori “não informativa”, em que o conhecimento a priori deveria ser dominado pela informação fornecida pelos dados.

Os esforços se concentram em identificar, para cada problema particular de inferência, uma específica função a priori de referência conjunta para todos os elementos desconhecidos do problema que conduziria a uma distribuição a posteriori de referência (marginal) para a quantidade de interesse, a qual sempre será dominada pela informação fornecida pelos dados (Bernardo, 1979b).

Utiliza-se da teoria de informação estatística para fornecer um significado preciso para esta dominância exigida. Note que as funções a priori de referência não foram propostas como uma aproximação para as crenças pessoais dos cientistas, mas sim como uma coleção de consensos formais para as funções a priori que poderiam ser convenientemente utilizadas como padrões de comunicação científica.

As distribuições a posteriori de referência são obtidas através do uso formal do teorema de Bayes com a função a priori de referência. Se necessário, podem ser utilizadas para fornecer estimativas pontuais ou estimativas intervalares, testes de hipóteses, ou para prever o valor de futuras observações. Isso proporciona um conjunto unificado de soluções Bayesianas objetivas para os problemas convencionais de inferência científica.

Ao restringir a classe de candidatas a função a priori, a metodologia de referência torna possível incorporar à análise qualquer conhecimento a priori verdadeiro. Desse ponto de vista, a derivação de funções a priori de referência pode ser descrita como um novo e poderoso método para elicitación. Por outro lado, quando a informação a priori subjetiva é especificada, a correspondente distribuição a posteriori subjetiva pode ser comparada com a distribuição a posteriori de referência - daí o seu nome - para avaliar a importância relativa das opiniões iniciais na inferência final. A abordagem mais aceitável sobre análise de referência baseada na teoria de informação foi inicialmente proposta por Bernardo (1979) e aperfeiçoada posteriormente por Berger e Bernardo em uma série de

trabalhos (Berger e Bernardo, 1989, 1992a e 1992b). O método de análise de referência generaliza muitos dos métodos considerados alternativos e os inclui como casos especiais, como por exemplo, o método de Jeffreys.

O argumento de análise de referência é sempre definido em termos de algum modelo da forma geral $\mathcal{M} \equiv \{p(\mathbf{x}|\boldsymbol{\omega}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\omega} \in \Omega\}$, que descreve as condições das quais os dados foram gerados. Assume-se que os dados \mathbf{x} consistem de observações de um vetor aleatório $\mathbf{x} \in \mathcal{X}$, com função densidade de probabilidade $p(\mathbf{x}|\boldsymbol{\omega})$ para algum $\boldsymbol{\omega} \in \Omega$. Geralmente, os dados consistirão de uma amostra aleatória $\mathbf{x} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ de tamanho fixo n da densidade $p(\mathbf{y}|\boldsymbol{\omega})$, $\mathbf{y} \in \mathcal{Y}$.

Seja $\boldsymbol{\theta} = \boldsymbol{\theta}(\boldsymbol{\omega}) \in \Theta$ o vetor de interesse. O modelo \mathcal{M} assumido pode ser reparametrizado da forma

$$\mathcal{M} \equiv \{p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \in \Lambda\}, \quad (2.1)$$

em que $\boldsymbol{\lambda}$ é um vetor de parâmetros perturbadores; referido como modelo $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$. Condicional ao modelo assumido, toda afirmação inferencial Bayesiana sobre o valor de $\boldsymbol{\theta}$ é encapsulada na distribuição a posteriori $p(\boldsymbol{\theta}|\mathbf{x}) \propto \int_{\Lambda} p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})p(\boldsymbol{\theta}, \boldsymbol{\lambda})d\boldsymbol{\lambda}$, que combina a informação fornecida pelos dados \mathbf{x} com a informação a priori sobre $\boldsymbol{\theta}$. Intuitivamente, a função a priori de referência para $\boldsymbol{\theta}$, dado o modelo \mathcal{M} e a classe de distribuições a priori candidatas \mathcal{P} , é a distribuição conjunta $\pi(\boldsymbol{\theta}, \boldsymbol{\lambda}|\mathcal{M}, \mathcal{P})$ da qual podemos esperar ter efeito minimal sobre a inferência a posteriori para a quantidade de interesse $\boldsymbol{\theta}$.

A teoria de informação estatística é usada para definir a função a priori de referência como uma função matemática que descreve a situação em que os dados dominarão melhor o conhecimento a priori sobre a quantidade de interesse. Em termos gerais, o método pode informalmente ser descrito através da seguinte idéia: a quantidade de informação registrada sobre a característica de interesse $\boldsymbol{\theta}$ que esperamos observar com o experimento é obviamente uma questão de nosso conhecimento a priori, isto é, se temos um bom conhecimento a priori sobre $\boldsymbol{\theta}$, então esperamos não aprender muito com os dados observados do experimento; por outro lado, se o nosso conhecimento a priori sobre $\boldsymbol{\theta}$ é escasso, então aprenderemos muito com os dados. Assim, quanto maior a quantidade de informação

a priori, menor será a quantidade de informação fornecida pelos dados.

A idéia básica da função a priori de referência consiste em identificar a forma matemática de uma distribuição a priori “não informativa”, utilizar a teoria de informação para medir a quantidade de informação sobre θ e definir a função a priori de referência como aquela que maximiza a falta de informação sobre θ .

Pelo uso formal do teorema de Bayes, a função a priori de referência produz uma distribuição a posteriori marginal de referência para o vetor de interesse, dada por

$$\pi(\theta|\mathbf{x}, \mathcal{M}, \mathcal{P}) \propto \int_{\Lambda} p(\mathbf{x}|\theta, \lambda)\pi(\theta, \lambda|\mathcal{M}, \mathcal{P})d\lambda, \quad (2.2)$$

que poderia ser descrita como uma expressão matemática de um conteúdo inferencial dos dados \mathbf{x} com respeito ao valor de θ , que não adiciona conhecimento além do contido no modelo estatístico assumido \mathcal{M} e a classe de candidatas a priori \mathcal{P} .

Na busca de distribuições a posteriori objetivas, surgiram várias exigências que podem ser listadas como propriedades necessárias de uma solução proposta:

1. *Generalidade.* O procedimento deve ser completamente geral, isto é, aplicável a qualquer problema de inferência. Em particular, uma distribuição a posteriori objetiva $\pi(\theta|\mathbf{x})$ deve ser uma distribuição de probabilidade própria para qualquer conjunto de dados \mathbf{x} grande o suficiente para identificar os parâmetros desconhecidos.
2. *Invariância.* Para qualquer função um a um $\phi = \phi(\theta)$, a distribuição a posteriori $\pi(\phi|\mathbf{x})$ obtida do modelo reparametrizado $p(\mathbf{x}|\phi, \lambda)$ deve estar de acordo com a distribuição a posteriori $\pi(\theta|\mathbf{x})$ obtida do modelo original $p(\mathbf{x}|\theta, \lambda)$ no sentido que para qualquer conjunto de dados $\mathbf{x} \in \mathcal{X}$, $\pi(\phi|\mathbf{x}) = \pi(\theta|\mathbf{x})|d\theta/d\phi|$. Entretanto, se o modelo tem uma estatística suficiente $\mathbf{u} = \mathbf{u}(\mathbf{x})$, então a distribuição a posteriori $\pi(\theta|\mathbf{x})$ obtida do modelo completo $p(\mathbf{x}|\theta, \lambda)$ deve ser a mesma que a distribuição a posteriori $\pi(\theta|\mathbf{u})$ obtida do modelo equivalente $p(\mathbf{u}|\theta, \lambda)$.

3. *Marginalização consistente.* Se, para todo \mathbf{x} , a distribuição a posteriori $\pi_1(\boldsymbol{\theta}|\mathbf{x})$ obtida do modelo $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ é da forma $\pi_1(\boldsymbol{\theta}|\mathbf{x}) = \pi_1(\boldsymbol{\theta}|\mathbf{u})$ para alguma estatística $\mathbf{u} = \mathbf{u}(\mathbf{x})$ cuja distribuição amostral $p(\mathbf{u}|\boldsymbol{\theta}, \boldsymbol{\lambda}) = p(\mathbf{u}|\boldsymbol{\theta})$ depende apenas de $\boldsymbol{\theta}$, então a distribuição a posteriori $\pi_2(\boldsymbol{\theta}|\mathbf{u})$ obtida do modelo marginal $p(\mathbf{u}|\boldsymbol{\theta})$ deve ser a mesma que a distribuição a posteriori $\pi_1(\boldsymbol{\theta}|\mathbf{u})$ obtida do modelo original completo.
4. *Propriedades de amostragem consistente.* As propriedades sob repetição amostral da distribuição a posteriori devem ser consistentes com o modelo. Em particular, a família de distribuições a posteriori $\{\pi(\boldsymbol{\theta}|\mathbf{x}_j), \mathbf{x}_j \in \mathcal{X}\}$ que poderia ser obtida por replicação amostral de $p(\mathbf{x}_j|\boldsymbol{\theta}, \boldsymbol{\omega})$ deve se concentrar na região de Θ que contém o verdadeiro valor de $\boldsymbol{\theta}$.

De acordo com Bernardo (2005), a análise de referência parece ser o único método para derivar distribuições a posteriori objetivas que satisfaçam a todas essas propriedades. Neste capítulo, descrevemos os elementos básicos da análise de referência, suas propriedades e alguns conceitos necessários para a construção da função a priori de referência.

2.1 Construção da função a priori de referência

A função a priori de referência é objetiva no sentido de que é uma função matemática bem definida do parâmetro de interesse $\boldsymbol{\theta}$ e do modelo \mathcal{M} , e a classe de distribuições a priori candidatas \mathcal{P} não possui nenhum elemento subjetivo.

A seguir, revisamos a construção da função a priori de referência para o caso uniparamétrico.

2.1.1 Caso de um parâmetro

Seja $\theta \in \Theta \subset \mathbb{R}$ a quantidade de interesse, e seja \mathbf{x} os dados observados, que consistem de uma observação do modelo $\mathcal{M} \equiv \{p(\mathbf{x}|\theta), \mathbf{x} \in \mathcal{X}, \theta \in \Theta\}$.

Proposição 2.1 (Forma explícita da função a priori de referência)

Considere o modelo $\mathcal{M} \equiv \{p(\mathbf{x}|\theta), \mathbf{x} \in \mathcal{X}, \theta \in \Theta \subset \mathbb{R}\}$ e seja \mathcal{P} a classe de todas as distribuições a priori contínuas com suporte Θ . Seja $\mathbf{x}^k = (\mathbf{x}_1, \dots, \mathbf{x}_k)$ o resultado de k replicações independentes de \mathcal{M} tal que $p(\mathbf{x}^k|\theta) = \prod_{j=1}^k p(\mathbf{x}_j|\theta)$, e seja $\mathbf{u}_k = \mathbf{u}_k(\mathbf{x}^k)$ uma estatística suficiente assintoticamente. Seja $h(\theta)$ uma função contínua estritamente positiva tal que para k suficientemente grande, $\int_{\Theta} p(\mathbf{u}_k|\theta)h(\theta)d\theta < \infty$, onde $p(\mathbf{u}_k|\theta)$ é qualquer aproximação assintótica da distribuição a posteriori. Então, define-se

$$f_k(\theta) = \exp \left\{ \int p(\mathbf{u}_k|\theta) \log \left(\frac{p(\mathbf{u}_k|\theta)h(\theta)}{\int_{\Theta} p(\mathbf{u}_k|\theta)h(\theta)d\theta} \right) d\mathbf{u}_k \right\}, \quad (2.3)$$

$$f(\theta) = \lim_{k \rightarrow \infty} \frac{f_k(\theta)}{f_k(\theta_0)}, \quad (2.4)$$

sendo θ_0 qualquer ponto interior de Θ . Se $f(\theta)$ é uma função a priori permissível, então, para qualquer $c > 0$, $\pi(\theta, \mathcal{M}, \mathcal{P}) = cf(\theta)$ é uma função a priori de referência.

Uma prova formal para a proposição pode ser vista em Berger et al. (2005).

Definição 2.1 (Função a priori permissível)

Uma função positiva $f(\theta)$ é uma função a priori permissível para o modelo $\mathcal{M} \equiv \{p(\mathbf{x}|\theta), \mathbf{x} \in \mathcal{X}, \theta \in \Theta \subset \mathbb{R}\}$, se para todo $\mathbf{x} \in \mathcal{X}$ há alguma sequência crescente $\{\Theta_i\}_{i=1}^{\infty}$ de subconjuntos de Θ , tal que,

- $\lim_{i \rightarrow \infty} \Theta_i = \Theta$;
- $\int_{\Theta_i} \pi(\theta)d\theta < \infty$;
- $\lim_{i \rightarrow \infty} \int_{\mathcal{X}} p_i(x) \delta\{\pi_i(\theta|x), \pi(\theta|x)\} dx = 0$

em que $f(\theta)$ é a função a priori permissível para o modelo \mathcal{M} se sempre fornecer uma distribuição a posteriori própria. Todas as funções a priori próprias são permissíveis no sentido da definição acima, mas uma função a priori imprópria pode ou não ser permissível, embora a mesma pareça ser próxima de uma função a priori própria.

Proposição 2.2 (Priori de referência sob normalidade assintótica)

Considere o modelo probabilístico $\mathcal{M} \equiv \{p(\mathbf{x}|\theta), \mathbf{x} \in \mathcal{X}, \theta \in \Theta \subset \mathbb{R}\}$ e \mathcal{P} a classe de todas as distribuições a priori contínuas com suporte Θ . Se a distribuição a posteriori de θ , $\pi(\theta|\mathbf{x}_1, \dots, \mathbf{x}_k)$, é assintoticamente normal com desvio padrão $S(\hat{\theta})/\sqrt{k}$, em que $\hat{\theta}$ é um estimador consistente de θ , e $S(\theta)^{-1}$ é uma função a priori permissível, então uma função da forma

$$\pi(\theta) \propto \frac{1}{S(\theta)}$$

é a função a priori de referência. Sob condições de regularidade, ver Bernardo (2005), $\hat{\theta}$ é assintoticamente suficiente e consistente, a distribuição a posteriori de θ é assintoticamente normal com variância $i(\hat{\theta})^{-1}/k$, em que $\hat{\theta}$ é o estimador de máxima verossimilhança de θ e

$$i(\theta) = - \int_{\mathcal{X}} p(\mathbf{x}|\theta) \frac{\partial^2}{\partial \theta^2} \log p(\mathbf{x}|\theta) d\mathbf{x}$$

é um escalar da matriz de informação de Fisher. Se $i(\theta)^{1/2}$ é uma função a priori permissível, a função a priori de referência é a priori de Jeffreys, $\pi(\theta|\mathcal{M}, \mathcal{P}) \propto i(\theta)^{1/2}$.

Nesta subseção, revisamos a construção da função a priori de referência no caso uniparamétrico. A seguir, estendemos essa metodologia para o caso bidimensional, considerando um parâmetro de interesse e outro perturbador.

2.1.2 Caso de um parâmetro perturbador

Considere agora que o modelo \mathcal{M} contém um parâmetro perturbador. Então, $\mathcal{M} \equiv \{p(\mathbf{x}|\theta, \lambda), \mathbf{x} \in \mathcal{X}, \theta \in \Theta, \lambda \in \Lambda\}$, sendo que a quantidade de interesse é $\theta \in \Theta \subset \mathbb{R}$ e o parâmetro perturbador é $\lambda \in \Lambda \subset \mathbb{R}$. Para obtermos a distribuição a posteriori de referência requerida para θ , $\pi(\theta|\mathbf{x})$, uma função a priori de referência conjunta apropriada $\pi(\theta, \lambda)$ é necessária. Pelo teorema de Bayes, a correspondente distribuição a posteriori conjunta é $\pi(\theta, \lambda|\mathbf{x}) \propto p(\mathbf{x}|\theta, \lambda)\pi(\theta, \lambda)$ e, integrando em relação ao parâmetro perturbador, a distribuição a posteriori de referência marginal para o parâmetro de interesse é

$$\pi(\theta|\mathbf{x}) = \int_{\Lambda} \pi(\theta, \lambda|\mathbf{x}) d\lambda \propto \int_{\Lambda} p(\mathbf{x}|\theta, \lambda)\pi(\theta, \lambda) d\lambda.$$

A extensão da metodologia da construção da função a priori de referência para o caso de dois parâmetros segue o procedimento matemático usual do caso de um único parâmetro dado na Subseção 2.1.1. Então, o método de análise de referência procede combinando os resultados obtidos em duas aplicações sucessivas da solução uniparamétrica, conforme abaixo:

1. Condicional em θ , $p(\mathbf{x}|\theta, \lambda)$ depende apenas do parâmetro perturbador λ e, portanto, a metodologia do caso uniparamétrico pode ser utilizada para obter a função a priori de referência condicional $\pi(\lambda|\theta) = \pi(\lambda|\theta, \mathcal{M}, \mathcal{P})$.
2. Se $\pi(\lambda|\theta)$ tem integral finita em Λ , a função a priori de referência condicional $\pi(\lambda|\theta)$ pode ser utilizada para integrar em relação ao parâmetro perturbador e obter o modelo uniparamétrico fazendo

$$p(\mathbf{x}|\theta) = \int_{\Lambda} p(\mathbf{x}|\theta, \lambda)\pi(\lambda|\theta)d\lambda, \quad (2.5)$$

para o qual a metodologia uniparamétrica pode ser novamente aplicada para obter a função a priori de referência marginal $\pi(\theta) = \pi(\theta|\mathcal{M}, \mathcal{P})$.

3. A função a priori de referência para θ desejada é então $\pi(\theta, \lambda) = \pi(\lambda|\theta)\pi(\theta)$, e a distribuição a posteriori de referência requerida é

$$\pi(\theta|\mathbf{x}) \propto \int_{\Lambda} p(\mathbf{x}|\theta, \lambda)\pi(\lambda|\theta)d\lambda = p(\mathbf{x}|\theta)\pi(\theta). \quad (2.6)$$

A equação (2.5) sugere que a função a priori de referência condicional fornece um procedimento para eliminar os parâmetros perturbadores, um grande problema no paradigma frequentista.

Se a função a priori de referência $\pi(\lambda|\theta)$ não é própria, a equação (2.5) não define um modelo estatístico válido e, como consequência, uma aproximação é necessária para fornecer uma solução geral. Então, a integração pode ser feita em cada elemento de uma sequência crescente $\{\Lambda_i\}_{i=1}^{\infty}$ de subconjuntos de Λ convergindo para Λ tal que $\pi(\lambda|\theta)$ seja integrável. Assim, a equação (2.5) é substituída por

$$p_i(\mathbf{x}|\theta) = \int_{\Lambda_i} p(\mathbf{x}|\theta, \lambda)\pi_i(\lambda|\theta)d\lambda, \quad (2.7)$$

em que $\pi_i(\lambda|\theta)$ é a restrição própria renormalizada de $\pi(\lambda|\theta)$ para Λ_i da distribuição a posteriori de referência $\pi_i(\theta|\mathbf{x}) = \pi(\theta|\mathcal{M}_i, P)$. Para mais detalhes ver Bernardo (2005).

Como seria de se esperar, a função a priori de referência para θ não depende da escolha do parâmetro perturbador λ . Assim, para algum $\psi = \psi(\theta, \lambda)$ tal que (θ, ψ) é uma função um a um de (θ, λ) , a função a priori de referência para θ em termos de (θ, ψ) é simplesmente $\pi^\theta(\theta, \psi) = \pi^\theta(\theta, \lambda)/|\partial(\theta, \psi)/\partial(\theta, \lambda)|$, a transformação probabilística apropriada da função a priori de referência para θ em termos de (θ, λ) . Note entretanto que, como mencionado anteriormente, a função a priori de referência pode depender do parâmetro de interesse; assim, a função a priori de referência para θ pode diferir da função a priori de referência para ψ a menos que seja ψ uma transformação um a um de θ , ou que ψ seja assintoticamente independente de θ . Isso é uma consequência esperada do fato matemático de que a função a priori que maximiza a falta de informação sobre θ não é necessariamente a mesma que maximiza a falta de informação sobre a função $\psi = \psi(\theta, \lambda)$.

Por definição, a função a priori de referência deve ser uma função permissível, ver Bernardo (2005). Em particular, deve levar a distribuições a posteriori próprias para todo conjunto de dados suficientemente grande para identificar os parâmetros.

Para distribuições a posteriori sob normalidade assintótica, as funções a priori de referência são facilmente obtidas em termos da matriz de informação de Fisher.

Proposição 2.3 (Priori de referência sob binormalidade assintótica)

Seja $\mathbf{x} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\}$ consistindo de n observações condicionalmente independentes (dado θ) do modelo $\mathcal{M} = \{p(\mathbf{y}|\theta, \lambda), \mathbf{y} \in \mathbf{Y}, \theta \in \Theta, \lambda \in \Lambda\}$, $(\theta, \lambda) \in \Theta \times \Lambda \subseteq \mathbb{R} \times \mathbb{R}$ com dois parâmetros reais θ e λ , onde θ é a quantidade de interesse, e suponha que a distribuição a posteriori conjunta (θ, λ) é assintoticamente normal com matriz de dispersão $S(\hat{\theta}, \hat{\lambda})$, em que $\hat{\theta}$ e $\hat{\lambda}$ são estimadores consistentes de θ e λ e $S(\theta, \lambda) = H^{-1}(\theta, \lambda)$, sendo H a matriz de informação de

Fisher.

1. A função a priori de referência condicional de λ é

$$\pi(\lambda|\theta) \propto h_{2,2}(\theta, \lambda)^{1/2}, \quad \lambda \in \Lambda(\theta),$$

em que $h_{2,2}$ é o elemento (2,2) da matriz de informação de Fisher H .

2. Se $\pi(\lambda|\theta)$ não é própria, uma aproximação compacta $\{\Lambda_i(\theta), i = 1, 2, \dots\}$ para $\Lambda(\theta)$ é requerida e a função a priori de referência de λ dado θ é dada por

$$\pi_i(\lambda|\theta) = \frac{h_{2,2}(\theta, \lambda)^{1/2}}{\int_{\Lambda_i(\theta)} h_{2,2}(\theta, \lambda)^{1/2} d\lambda}, \quad \lambda \in \Lambda_i(\theta).$$

3. No interior de cada $\Lambda_i(\theta)$ a função a priori de referência marginal de θ é obtida da forma

$$\pi_i(\theta) \propto \exp \left\{ \int_{\Lambda_i(\theta)} \pi_i(\lambda|\theta) \log \left[s_{1,1}^{-1/2}(\theta, \lambda) \right] d\lambda \right\},$$

em que $s_{1,1}^{-1/2}(\theta, \lambda) = h_\theta(\theta, \lambda) = h_{11} - h_{12}h_{2,2}^{-1}h_{21}$.

4. A distribuição a posteriori de referência de θ dado \mathbf{x} é

$$\pi(\theta|\mathbf{x}) \propto \pi(\theta) \left\{ \int_{\Lambda(\theta)} \left\{ \prod_{i=1}^n p(x_i|\theta, \lambda) \right\} \pi(\lambda|\theta) d\lambda \right\}.$$

Sob condições de regularidade (ver Bernardo, 2005), a função a priori de referência pode ser reescrita como o produto de duas funções de parâmetros independentes, como estabelece o corolário abaixo.

Corolário 2.1 (Função a priori de referência sob fatoração)

Se o espaço de parâmetros perturbadores $\Lambda(\theta) = \Lambda$ é independente de θ , e as funções $s_{1,1}^{-1/2}(\theta, \lambda)$ e $h_{2,2}^{1/2}(\theta, \lambda)$ se fatoram na forma

$$\{s_{1,1}(\theta, \lambda)\}^{-1/2} = f_1(\theta) g_1(\lambda) \quad e \quad \{h_{2,2}(\theta, \lambda)\}^{1/2} = f_2(\theta) g_2(\lambda),$$

então,

$$\pi(\theta) \propto f_1(\theta) \quad e \quad \pi(\lambda|\theta) \propto g_2(\lambda).$$

A função a priori de referência relativa à ordenação (θ, λ) é definida por

$$\pi(\theta, \lambda) = f_1(\theta) g_2(\lambda).$$

Nesse caso, não existe a necessidade de uma aproximação compacta, mesmo que a função a priori de referência não seja própria (Bernardo e Smith, 1994). Vale ressaltar que este corolário e a proposição 2.3 são utilizados neste trabalho na derivação da função a priori de referência associada ao modelo Weibull.

Quando o vetor de parâmetros do modelo em questão possui mais que dois componentes, a idéia pode ser facilmente estendida (Bernardo e Smith, 1994). Essa abordagem é vista na subseção seguinte.

2.1.3 Caso multiparamétrico

Nesta subseção, revisamos uma extensão do caso de dois parâmetros, descrito anteriormente, para o caso multiparamétrico, com $(\theta, \lambda_1, \dots, \lambda_m)$, segundo Bernardo e Ramón (1998). A função a priori de referência é construída através de condicionamentos sucessivos, sempre aplicando a metodologia do caso uniparamétrico para a obtenção de cada componente da decomposição

$$\pi(\theta, \lambda_1, \dots, \lambda_m) = \pi(\lambda_m | \theta, \lambda_1, \dots, \lambda_{m-1}) \dots \pi(\lambda_1 | \theta) \pi(\theta).$$

Revisamos a metodologia para a construção da função a priori de referência multidimensional no caso em que as condições de regularidade são válidas. A abordagem apresentada é a que o vetor de parâmetros $(\theta, \lambda_1, \dots, \lambda_m)$ é dividido e ordenado em ν grupos. É importante observar que nessa abordagem ordenamos os ν diferentes grupos segundo sua maior importância inferencial, mas que, dentro de cada grupo, a ordenação dos parâmetros não é relevante (Berger e Bernardo, 1992b).

Considere $p(\mathbf{x} | \theta, \boldsymbol{\lambda})$ o modelo paramétrico, e $H(\theta, \boldsymbol{\lambda})$ a matriz de informação de Fisher $(m+1) \times (m+1)$. Então, a distribuição a posteriori de $(\theta, \boldsymbol{\lambda})$ é assintoticamente normal com média $(\hat{\theta}, \hat{\boldsymbol{\lambda}})$, os correspondentes estimadores de máxima verossimilhança, e matriz de covariâncias $S(\hat{\theta}, \hat{\boldsymbol{\lambda}})$, em que $S = H^{-1}$.

Disso, segue que, se S_j é a submatriz $j \times j$ superior de S , $j = 1, \dots, m+1$, $H_j = S_j^{-1}$ e $h_{j,j}$ é o elemento (j, j) de H_j , tal que $H_{m+1} = H$, então,

1. A distribuição marginal a posteriori de θ é assintoticamente normal com

desvio padrão

$$s_{1,1}(\hat{\theta}, \hat{\boldsymbol{\lambda}})^{1/2} = h_{1,1}(\hat{\theta}, \hat{\boldsymbol{\lambda}})^{-1/2};$$

2. A distribuição a posteriori condicional de λ_i dado $\theta, \lambda_1, \dots, \lambda_{i-1}$, é assintoticamente normal com desvio padrão

$$h_{i+1,i+1}(\theta, \lambda_1, \dots, \lambda_{i-1}, \hat{\lambda}_i, \dots, \hat{\lambda}_m)^{-1/2};$$

e uma possibilidade é usar sequencialmente a metodologia descrita no caso anterior para obter $\pi(\lambda_m|\theta, \lambda_1, \dots, \lambda_{m-1})$, $\pi(\lambda_{m-1}|\theta, \lambda_1, \dots, \lambda_{m-2})$, \dots , $\pi(\lambda_1|\theta)$ e $\pi(\theta)$, e produzir a distribuição a posteriori de referência desejada.

Proposição 2.4 (Normalidade assintótica - Caso multiparamétrico)

Seja $p(x|\theta, \boldsymbol{\lambda})$, $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_m)$ o modelo probabilístico com $m + 1$ parâmetros. Seja θ a quantidade de interesse e suponha que a distribuição conjunta de $(\theta, \boldsymbol{\lambda})$ é assintoticamente normal com matriz de covariâncias $S(\hat{\theta}, \hat{\boldsymbol{\lambda}})$. Então, se S_j é a submatriz $j \times j$ superior à esquerda de S , $H_j = S_j^{-1}$ e $h_{j,j}(\theta, \boldsymbol{\lambda})$ é o elemento (j, j) de H_j ,

1. As funções a priori de referência são

$$\pi(\lambda_m|\theta, \lambda_1, \dots, \lambda_{m-1}) \propto h_{m+1,m+1}^{1/2}(\theta, \boldsymbol{\lambda})$$

e

$$\pi(\lambda_i|\theta, \lambda_1, \dots, \lambda_{i-1}) \propto \exp \left\{ \int_{\Lambda_{i+1}} \dots \int_{\Lambda_m} \log h_{i,i}^{1/2}(\theta, \boldsymbol{\lambda}) \left\{ \prod_{j=i+1}^m \pi(\lambda_j|\theta, \lambda_1, \dots, \lambda_{j-1}) \right\} d\boldsymbol{\lambda}_{i+1} \right\},$$

em que $d\boldsymbol{\lambda}_j = d\lambda_j \times \dots \times d\lambda_m$, e $\pi(\lambda_i|\lambda_1, \dots, \lambda_{i-1})$, $i = 1, \dots, m$ são próprias. Se qualquer dessas funções a priori de referência condicionais não é própria, então uma aproximação compacta é necessária para as correspondentes integrais.

2. A função a priori de referência marginal de θ é

$$\pi(\theta) \propto \exp \left\{ \int_{\Lambda_1} \dots \int_{\Lambda_m} \log s_{1,1}^{-1/2}(\theta, \boldsymbol{\lambda}) \left\{ \prod_{j=1}^m \pi(\lambda_j|\theta, \lambda_1, \dots, \lambda_{j-1}) \right\} d\boldsymbol{\lambda}_1 \right\}.$$

Corolário 2.2 (Função a priori de referência sob fatoração)

Se os espaços dos parâmetros perturbadores $\Lambda_i(\theta, \lambda_1, \dots, \lambda_{i-1}) = \Lambda_i$ são independentes de ambos θ e λ'_i s, e as funções $s_{1,1}, h_{i,i}, \dots, h_{m,m}$ se fatoram da forma

$$s_{1,1}^{-1/2} = f_0(\theta)g_0(\lambda_1, \dots, \lambda_m) \text{ e}$$

$$h_{i+1,i+1}^{1/2} = f_i(\lambda_i)g_i(\theta, \lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_m), \quad i = 1, \dots, m,$$

então,

$$\pi(\theta) \propto f_0(\theta), \quad \pi(\lambda_i|\theta, \lambda_1, \dots, \lambda_{i-1}) \propto f_i(\lambda_i), \quad i = 1, \dots, m,$$

e não há necessidade para aproximações compactas, nem se $\pi(\lambda_i|\theta, \lambda_1, \dots, \lambda_{i-1})$ não for própria.

Berger e Bernardo recomendavam, antes de proporem a metodologia apresentada acima (Berger e Bernardo, 1992b), separar o vetor de parâmetros $(\theta, \lambda_1, \dots, \lambda_m)$ em apenas dois grupos, o grupo de parâmetros de interesse, e o grupo de parâmetros perturbadores (Berger e Bernardo, 1989).

Embora a abordagem com 2 grupos seja mais facilmente computável, os autores refinaram essa abordagem, uma vez que encontraram funções a priori de referência com comportamento insatisfatório (Berger e Bernardo, 1992b). Atualmente, a abordagem de ν grupos, no caso particular em que cada grupo é composto por um único parâmetro, é mais recomendada pelos autores (Bernardo e Ramón, 1998). Segundo Berger e Bernardo (1992b), devemos agrupar parâmetros somente quando houver algum motivo especial.

A abordagem com ν grupos tem se mostrado muito eficiente, parecendo não produzir resultados insatisfatórios. Esses resultados certamente não garantem que a função a priori com ν grupos sempre apresente bom comportamento, mas o fato é que o seu sucesso tem sido impressionante (Berger e Bernardo, 1992a). Entretanto, nessa abordagem, a função a priori de referência pode depender da ordenação do vetor de parâmetros. Então, um mesmo modelo paramétrico pode gerar diversas funções a priori de referência, cada uma associada a uma diferente ordenação do vetor de parâmetros.

Berger e Bernardo (1992b) recomendam que, para cada problema específico, a ordenação dos parâmetros deve ser de acordo com sua importância inferencial. Quando não é possível decidir a ordenação dos parâmetros perturbadores, os autores recomendam obter funções a priori de referência associadas a cada uma das possíveis ordenações dos mesmos e, então, escolher a função a priori que apresente o melhor comportamento inferencial. Todavia, Bernardo (1997) argumenta que na maioria das vezes a função a priori de referência é invariante com respeito à ordenação dos parâmetros de perturbação.

Neste trabalho (Capítulo 3), utilizamos a abordagem de ν grupos para o modelo Weibull na aplicação de riscos competitivos, devido a sua eficiência até agora verificada. Como também é de nosso interesse realizar testes de hipóteses para os parâmetros desse modelo, na próxima seção, revisamos o critério de referência Bayesiano (Bayesian Reference Criterion, BRC).

2.2 Teste de hipóteses Bayesiano

Testes de hipóteses têm sido assunto de polêmica desde sua formulação por Neyman e Pearson em 1930. Isso se deve principalmente ao fato de que sua formulação padrão frequentemente constitui uma simplificação grosseira do problema a ser resolvido. De fato, muitos dos problemas que tradicionalmente têm sido formulados em termos de testes de hipóteses são problemas realmente complexos de decisão em escolha de modelos, cuja solução apropriada naturalmente depende da estrutura do problema. Alguns desses importantes elementos estruturais são a motivação para escolher um modelo particular, a classe de modelos considerados, e a informação a priori disponível.

Uma estrutura matemática natural para a escolha do modelo é a teoria da decisão. Deve-se especificar o conjunto de modelos a ser considerado, para decidir se podemos assumir ou não que esse conjunto inclua o verdadeiro modelo para formular distribuições de probabilidade que descrevam a informação a priori de todos os elementos desconhecidos no problema, e para especificar uma função de perda medindo eventuais consequências de cada modelo escolhido. A melhor

alternativa dentro do conjunto dos modelos considerados é o modelo que minimiza a correspondente perda esperada a posteriori. Neste trabalho, revisamos os problemas de escolha de modelos, denominados Testes de Hipóteses, em que o modelo \mathcal{M} é provisoriamente aceito e desejamos testar se os dados avaliados são ou não compatíveis com um particular submodelo \mathcal{M}_0 . Mais detalhes se encontram em Bernardo e Rueda (2002).

2.2.1 Teste de hipóteses como um problema de decisão

Considere que o modelo $\mathcal{M} \equiv \{p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \in \Lambda\}$ fornece uma descrição apropriada do comportamento probabilístico dos dados observados $\mathbf{x} \in \mathcal{X}$ em termos do vetor de interesse $\boldsymbol{\theta} \in \Theta$ e do vetor de parâmetros perturbadores $\boldsymbol{\lambda} \in \Lambda$.

Em particular, a distribuição a posteriori marginal de $\boldsymbol{\theta}$ transmite informações sobre esse vetor de valores de interesse que podem ser compatíveis com as transmitidas pelos dados observados \mathbf{x} , com uma densidade de probabilidade relativamente alta. Nesses casos, um valor particular $\boldsymbol{\theta} = \boldsymbol{\theta}_0 \in \Theta$ da quantidade de interesse é sugerida na investigação como uma consideração especial, porque assumindo $\boldsymbol{\theta} = \boldsymbol{\theta}_0$ simplificaria muito o modelo, ou porque há argumentos adicionais que sugerem que $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Intuitivamente, a hipótese nula $H_0 \equiv \{\boldsymbol{\theta} = \boldsymbol{\theta}_0\}$ deve ser julgada compatível com os dados \mathbf{x} se $\boldsymbol{\theta}_0$ tiver uma distribuição a posteriori relativamente alta, entretanto, uma conclusão mais precisa é requerida.

Formalmente, o teste da hipótese $H_0 \equiv \{\boldsymbol{\theta} = \boldsymbol{\theta}_0\}$ é definido como um problema de decisão, em que o espaço das ações possui apenas dois elementos, aceitar (a_0) ou rejeitar (a_1) o uso do modelo restrito $\mathcal{M}_0 \equiv \{p(\mathbf{x}|\boldsymbol{\theta}_0, \boldsymbol{\lambda}), \boldsymbol{\lambda} \in \Lambda\}$ como uma aproximação conveniente do modelo assumido $\mathcal{M} \equiv \{p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \in \Lambda\}$. Para solucionar esse problema de decisão, é necessário especificar uma função de perda apropriada, $\{l[a_i, (\boldsymbol{\theta}, \boldsymbol{\lambda})], i = 0, 1\}$, medindo as consequências de aceitar ou rejeitar H_0 como uma função de valores $(\boldsymbol{\theta}, \boldsymbol{\lambda})$ dos parâmetros.

Dado \mathbf{x} , a ação ótima rejeitará H_0 se, e só se, a perda esperada a

posteriori de aceitação, $\int_{\Theta} \int_{\Lambda} l[a_0, \boldsymbol{\theta}, \boldsymbol{\lambda}] \pi(\boldsymbol{\theta}, \boldsymbol{\lambda} | \mathbf{x}) d\boldsymbol{\theta} d\boldsymbol{\lambda}$ for maior que a perda a posteriori esperada de rejeição, $\int_{\Theta} \int_{\Lambda} \{l[a_1, \boldsymbol{\theta}, \boldsymbol{\lambda}]\} d\boldsymbol{\theta} d\boldsymbol{\lambda}$, isto é, se

$$\int_{\Theta} \int_{\Lambda} \{l[a_0, \boldsymbol{\theta}, \boldsymbol{\lambda}] - l[a_1, \boldsymbol{\theta}, \boldsymbol{\lambda}]\} \pi(\boldsymbol{\theta}, \boldsymbol{\lambda} | \mathbf{x}) d\boldsymbol{\theta} d\boldsymbol{\lambda} > 0. \quad (2.8)$$

Portanto, apenas a diferença entre as perdas

$$\Delta l(H_0, \boldsymbol{\theta}, \boldsymbol{\lambda}) = l[a_0, \boldsymbol{\theta}, \boldsymbol{\lambda}] - l[a_1, \boldsymbol{\theta}, \boldsymbol{\lambda}], \quad (2.9)$$

que mede a vantagem de rejeitar H_0 como uma função de $(\boldsymbol{\theta}, \boldsymbol{\lambda})$, tem que ser especificada. Disso, segue que qualquer solução de Bayes para problemas de decisão deve ser da forma

$$\text{Rejeitar } H_0 \text{ se } \int_{\Theta} \int_{\Lambda} \Delta l(H_0, \boldsymbol{\theta}, \boldsymbol{\lambda}) \pi(\boldsymbol{\theta}, \boldsymbol{\lambda} | \mathbf{x}) d\boldsymbol{\theta} d\boldsymbol{\lambda} > 0 \quad (2.10)$$

para alguma função de diferença entre perdas $\Delta l(H_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$ e alguma função a priori $\pi(\boldsymbol{\theta}, \boldsymbol{\lambda})$.

Na próxima subseção, justificamos a escolha de uma particular função de perda diferença, a discrepância intrínseca. Isso, combinado com a análise de referência, propõe uma atrativa solução Bayesiana para problemas de teste de hipóteses, definidos como um problema de decidir se os dados disponíveis são ou não compatíveis com a hipótese que os parâmetros do modelo pertencem a algum subconjunto do espaço paramétrico (ver Bernardo e Rueda, 2002).

2.2.2 Critério de Referência Bayesiano (BRC)

Seja o modelo $\mathcal{M} \equiv \{p(\mathbf{x} | \boldsymbol{\theta}, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \in \Lambda\}$ uma descrição aceita do comportamento probabilístico dos dados $\mathbf{x} \in \mathcal{X}$, seja a_0 a decisão de trabalhar sob o modelo restrito $\mathcal{M}_0 \equiv \{p(\mathbf{x} | \boldsymbol{\theta}_0, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\lambda} \in \Lambda\}$, e seja a_1 a decisão para manter o modelo irrestrito \mathcal{M} . Nessa situação, $\Delta l(H_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$, que mede a vantagem de rejeitar H_0 como uma função de $(\boldsymbol{\theta}, \boldsymbol{\lambda})$, pode assumir a forma

$$\Delta l(H_0, \boldsymbol{\theta}, \boldsymbol{\lambda}) = \delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda}) - d^*, \quad d^* > 0, \quad (2.11)$$

em que

1. a função $\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$ é uma medida não-negativa da discrepância entre o modelo assumido $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ e a aproximação em $\{p(\mathbf{x}|\boldsymbol{\theta}_0, \boldsymbol{\lambda}), \boldsymbol{\lambda} \in \boldsymbol{\Lambda}\}$, tal que $\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0, \boldsymbol{\lambda}) = 0$, e
2. a constante $d^* > 0$ é um valor que mede a vantagem de trabalhar com o modelo simples quando esse é verdadeiro.

Escolhas de ambos, $\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$ e d^* , que podem ser apropriados para o uso geral, são discutidas.

Convenciona-se que a função de perda geralmente se foca na distância entre o valor verdadeiro e o nulo da quantidade de interesse, em vez da distância entre os modelos, que, usualmente, não são invariantes sob reparametrização. Perdas intrínsecas, no entanto, focam diretamente em quão diferente o modelo verdadeiro é do modelo nulo, e essas geralmente produzem soluções invariantes.

Definição 2.2 (Discrepância intrínseca)

A *discrepância intrínseca* $\delta\{p_1, p_2\}$ entre duas densidades de probabilidade $p_1(\mathbf{x})$ e $p_2(\mathbf{x})$ de uma quantidade aleatória $\mathbf{x} \in \boldsymbol{\chi}$ é

$$\delta\{p_1, p_2\} = \min \left\{ \int_{\boldsymbol{\chi}_1} p_1(\mathbf{x}) \log \frac{p_1(\mathbf{x})}{p_2(\mathbf{x})} d\mathbf{x}, \int_{\boldsymbol{\chi}_2} p_2(\mathbf{x}) \log \frac{p_2(\mathbf{x})}{p_1(\mathbf{x})} d\mathbf{x} \right\}, \quad (2.12)$$

em que $\kappa(p_2|p_1) = \int_{\boldsymbol{\chi}} p_1(\mathbf{x}) \log \frac{p_1(\mathbf{x})}{p_2(\mathbf{x})} d\mathbf{x}$ é a divergência de Kullback-Leiber entre $p_2(\mathbf{x})$ e $p_1(\mathbf{x})$ e que uma das integrais (ou soma) é finita. $\delta\{p_1, p_2\}$ é interpretada como o mínimo do valor esperado do logaritmo da razão das densidades de p_1 contra p_2 .

A *discrepância intrínseca* entre os dois modelos paramétricos para $\mathbf{x} \in \boldsymbol{\chi}$, $\mathcal{M}_1 = \{p_1(\mathbf{x}|\boldsymbol{\omega}), \mathbf{x} \in \boldsymbol{\chi}_1, \boldsymbol{\omega} \in \boldsymbol{\Omega}\}$ e $\mathcal{M}_2 = \{p_2(\mathbf{x}|\boldsymbol{\psi}), \mathbf{x} \in \boldsymbol{\chi}_2, \boldsymbol{\psi} \in \boldsymbol{\Psi}\}$, é a *discrepância intrínseca mínima* entre esses elementos,

$$\delta\{\mathcal{M}_1, \mathcal{M}_2\} = \min_{\boldsymbol{\omega} \in \boldsymbol{\Omega}, \boldsymbol{\psi} \in \boldsymbol{\Psi}} \delta\{p_1(\mathbf{x}|\boldsymbol{\omega}), p_2(\mathbf{x}|\boldsymbol{\psi})\}. \quad (2.13)$$

Segue da Denificação 2.1 que $\delta\{p_1(\mathbf{x}), p_2(\mathbf{x})\}$ fornece o mínimo esperado do logaritmo da razão de densidades $\log(p_i(\mathbf{x})/p_j(\mathbf{x}))$ em favor da verdadeira densidade que se obteria se os dados $\mathbf{x} \in \boldsymbol{\chi}$ fossem amostrados de $p_1(\mathbf{x})$ ou $p_2(\mathbf{x})$.

Em particular, se $p_1(\mathbf{x})$ e $p_2(\mathbf{x})$ são modelos de probabilidade alternativos para $\mathbf{x} \in \mathcal{X}$, e se assume que um deles é verdadeiro, então $\delta\{p_1(\mathbf{x}), p_2(\mathbf{x})\}$ é o mínimo da razão do logaritmo das verossimilhanças esperado para o modelo verdadeiro.

A discrepância intrínseca é simétrica, não-negativa e é igual a zero se, e somente se, $p_1(\mathbf{x}) = p_2(\mathbf{x})$. Além disso, é invariante sob transformação um a um de x , e é aditiva sob observações independentes; assim, se os dados observados \mathbf{x} consistem de uma amostra aleatória $\mathbf{x} = \{x_1, \dots, x_n\}$ de $p_1(\mathbf{x})$ ou $p_2(\mathbf{x})$, então $\delta\{p_1(\mathbf{x}), p_2(\mathbf{x})\} = n\delta\{p_1(x), p_2(x)\}$.

Definição 2.3 (Perda discrepante intrínseca)

A perda discrepante intrínseca $\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$ da substituição do modelo probabilístico $\mathcal{M} \equiv \{p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\theta} \in \Theta, \boldsymbol{\lambda} \in \Lambda\}$ pelo modelo com a restrição $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, $\mathcal{M}_0 \equiv \{p(\mathbf{x}|\boldsymbol{\theta}_0, \boldsymbol{\lambda}), \mathbf{x} \in \mathcal{X}, \boldsymbol{\lambda} \in \Lambda\}$, é a discrepância intrínseca entre a densidade de probabilidade $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ e a família de densidades de probabilidades $\{p(\mathbf{x}|\boldsymbol{\theta}_0, \boldsymbol{\lambda}), \boldsymbol{\lambda} \in \Lambda\}$, dada por

$$\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda}) = \min_{\boldsymbol{\lambda}_0 \in \Lambda} \delta\{p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda}), p(\mathbf{x}|\boldsymbol{\theta}_0, \boldsymbol{\lambda}_0)\}.$$

A discrepância intrínseca $\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$ entre $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ e \mathcal{M}_0 é a discrepância intrínseca entre a densidade de probabilidade assumida $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ e a aproximação com $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Note que $\delta(\boldsymbol{\theta}_0, \boldsymbol{\theta}, \boldsymbol{\lambda})$ é invariante sob reparametrização de $\boldsymbol{\theta}$ ou $\boldsymbol{\lambda}$.

Além disso, embora não explicitamente mostrado na notação, a função de discrepância intrínseca geralmente depende do tamanho amostral. Na verdade, se os dados $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^n$ consistem de uma amostra aleatória $\mathbf{x} = \{x_1, \dots, x_n\}$ de tamanho n de $p(x|\boldsymbol{\theta}_i, \boldsymbol{\lambda})$, então

$$\int_{\mathcal{X}} p(\mathbf{x}|\boldsymbol{\theta}_i, \boldsymbol{\lambda}) \log \frac{p(\mathbf{x}|\boldsymbol{\theta}_i, \boldsymbol{\lambda})}{p(\mathbf{x}|\boldsymbol{\theta}_j, \boldsymbol{\lambda}_j)} d\mathbf{x} = n \int_{\mathbb{R}} p(x|\boldsymbol{\theta}_i, \boldsymbol{\lambda}) \log \frac{p(x|\boldsymbol{\theta}_i, \boldsymbol{\lambda})}{p(x|\boldsymbol{\theta}_j, \boldsymbol{\lambda}_j)} dx, \quad (2.14)$$

tal que a discrepância intrínseca associada com o modelo completo $p(\mathbf{x}|\boldsymbol{\theta}, \boldsymbol{\lambda})$ é simplesmente n vezes a discrepância intrínseca associada ao modelo $p(x|\boldsymbol{\theta}, \boldsymbol{\lambda})$ que corresponde a uma só observação. No entanto, a definição pode ser usada em problemas onde \mathbf{x} não consiste de uma amostra aleatória.

De (2.10) e (2.11) segue que, com a função de perda discrepante intrínseca, a hipótese H_0 deve ser rejeitada se, e só se, a vantagem esperada a posteriori de

rejeitar θ_0 , dado o modelo \mathcal{M} e os dados \mathbf{x} , for suficientemente grande, tal que o critério de decisão se torna

$$\text{Rejeitar } H_0 \text{ se } d(\theta_0, \mathbf{x}) = \int_{\Theta} \int_{\Lambda} \delta(\theta_0, \theta, \lambda) \pi(\theta, \lambda | \mathbf{x}) d\theta d\lambda > d^*, \quad (2.15)$$

para $d^* > 0$, desde que $\delta(\theta_0, \theta, \lambda)$ e $d(\theta_0, \mathbf{x})$ sejam não-negativas.

A função $d(\theta_0, \mathbf{x})$ é uma medida contínua e não-negativa de quão inapropriado pode ser esperado ser a simplificação do modelo aceitando H_0 . Na verdade, $d(\theta_0, \mathbf{x})$ é uma medida precisa de uma quantidade de informação esperada (a posteriori) que poderá ser necessária para recuperar a densidade de probabilidade assumida $p(\mathbf{x} | \theta, \lambda)$ de uma aproximação em $\mathcal{M}_0 \equiv \{p(\mathbf{x} | \theta, \lambda), \lambda \in \Lambda\}$; é a medida de “poder de evidência” contra \mathcal{M}_0 dado $\mathcal{M} \equiv \{p(\mathbf{x} | \theta, \lambda), \theta \in \Theta, \lambda \in \Lambda\}$. Em linguagem tradicional, $d(\theta_0, \mathbf{x})$ é uma estatística de teste para H_0 , e a hipótese nula deve ser rejeitada se o valor de $d(\theta_0, \mathbf{x})$ ultrapassar o valor crítico d^* .

Naturalmente, ao implementar o critério de decisão, ambos, função a priori $\pi(\theta, \lambda)$ e a constante d^* , devem ser escolhidas. Essas duas questões importantes são sucessivamente tratadas, conduzindo para um critério de decisão geral de testes de hipóteses, o critério de referência Bayesiano.

Definição 2.4 (Critério de Referência Bayesiano (BRC))

Seja $\{p(\mathbf{x} | \theta, \lambda) | \theta \in \Theta, \lambda \in \Lambda\}$ o modelo estatístico que gerou os dados $\mathbf{x} \in \mathcal{X}$, e considere um valor preciso $\theta = \theta_0$ a ser testado depois de \mathbf{x} ter sido observado. Para decidir se aceita ou não o valor preciso θ_0 , pode ser usado como uma aproximação para o valor desconhecido de θ ,

1. Calcule a discrepância intrínseca $\delta(\theta_0, \theta, \lambda)$;
2. Derive a correspondente esperança a posteriori de referência $d(\theta_0, \mathbf{x}) = E[\delta(\theta_0, \theta, \lambda) | \mathbf{x}]$ e estipule esse número como uma medida de evidência contra a hipótese nula $H_0 \equiv \{\theta = \theta_0\}$.
3. Se uma decisão formal é requerida, rejeite a hipótese nula se, e só se, $d(\theta_0, \mathbf{x}) > d^*$, para algum d^* . Os valores $d^* \approx 1$ (não evidência contra H_0), $d^* \approx 2.5$ (leve evidência contra H_0) e $d^* > 5$ (evidência significativa contra H_0) podem ser convenientemente usados para a comunicação científica.

Na próxima seção, aplicamos a metodologia descrita neste capítulo para os parâmetros do modelo Weibull (1.1) e para sua função de confiabilidade (1.2).

2.3 Análise de referência para o modelo Weibull

Consideramos a metodologia descrita na Subseção 2.1.2 para o modelo Weibull (1.1), em que o parâmetro de interesse é η e o parâmetro perturbador é β .

O logaritmo da função de verossimilhança (1.4) é

$$l(\beta, \eta | \mathbf{x}) = n [\log(\beta) - \beta \log(\eta)] + (\beta - 1) \sum_{i=1}^n \log(x_i) - \sum_{i=1}^n \left(\frac{x_i}{\eta} \right)^\beta. \quad (2.16)$$

As derivadas primeiras e segundas de (2.16), que foram calculadas através do *software* Mathematica, são

$$\frac{\partial l(\beta, \eta | \mathbf{x})}{\partial \eta} = \beta \left(\frac{-n + \sum_{i=1}^n \left(\frac{x_i}{\eta} \right)^\beta}{\eta} \right),$$

$$\frac{\partial l(\beta, \eta | \mathbf{x})}{\partial \beta} = n \left(\frac{1}{\beta} - \log(\eta) \right) + \sum_{i=1}^n \log(x_i) - \sum_{i=1}^n \left(\frac{x_i}{\eta} \right)^\beta \log \left(\frac{x_i}{\eta} \right),$$

$$\frac{\partial^2 l(\beta, \eta | \mathbf{x})}{\partial \eta^2} = \frac{\beta}{\eta^2} \left(n - (1 + \beta) \sum_{i=1}^n \left(\frac{x_i}{\eta} \right)^\beta \right),$$

$$\frac{\partial^2 l(\beta, \eta | \mathbf{x})}{\partial \beta^2} = -\frac{n}{\beta^2} - \sum_{i=1}^n \left(\frac{x_i}{\eta} \right)^\beta \log \left(\frac{x_i}{\eta} \right)^2$$

e

$$\frac{\partial^2 l(\beta, \eta | \mathbf{x})}{\partial \eta \partial \beta} = -\frac{n - \eta \sum_{i=1}^n \left(\frac{\beta \log \left[\frac{x_i}{\eta} \right] x_i \left(\frac{x_i}{\eta} \right)^{\beta-1}}{\eta^2} + \left(\frac{x_i}{\eta} \right)^\beta \right)}{\eta}.$$

Para obtermos a matriz de informação de Fisher $H(\boldsymbol{\theta}) = E \left[-\frac{\partial^2 l(\boldsymbol{\theta} | \mathbf{x})}{\partial \theta_i \partial \theta_j} \right]$ para um vetor de parâmetros $\boldsymbol{\theta}$, primeiramente, devemos calcular $E \left[\frac{\partial^2 l(\boldsymbol{\theta} | \mathbf{x})}{\partial \theta_i \partial \theta_j} \right]$. Então,

$$E \left[\frac{\partial^2 l(\beta, \eta | \mathbf{x})}{\partial \eta^2} \right] = \frac{\beta}{\eta^2} \left[n - (1 + \beta) \sum_{i=1}^n E \left[\left(\frac{x_i}{\eta} \right)^\beta \right] \right], \text{ em que } E \left[\left(\frac{x_i}{\eta} \right)^\beta \right] = 1.$$

$$E \left[\frac{\partial^2 l(\beta, \eta | \mathbf{x})}{\partial \beta^2} \right] = -\frac{n}{\beta^2} - \sum_{i=1}^n E \left[\left(\frac{x_i}{\eta} \right)^\beta \log \left(\frac{x_i}{\eta} \right)^2 \right], \text{ em que}$$

$$E \left[\left(\frac{x_i}{\eta} \right)^\beta \log \left(\frac{x_i}{\eta} \right)^2 \right] = \frac{(-2 + \gamma)\gamma + \pi^2/6}{\beta^2}.$$

$$E \left[\frac{\partial^2 l(\beta, \eta | \mathbf{x})}{\partial \eta \partial \beta} \right] = -\frac{n - \eta \sum_{i=1}^n \left(\frac{\beta E \left[\log \left[\frac{x_i}{\eta} \right] x_i \left(\frac{x_i}{\eta} \right)^{\beta-1} \right] + E \left[\left(\frac{x_i}{\eta} \right)^\beta \right]}{\eta^2} \right)}{\eta}, \text{ em que}$$

$$E \left[\log \left[\frac{x_i}{\eta} \right] x_i \left(\frac{x_i}{\eta} \right)^{\beta-1} \right] = \frac{\eta \psi(2)}{\beta}.$$

Assim, a matriz de informação de Fisher e sua inversa são dadas por

$$H(\eta, \beta) = n \begin{bmatrix} \frac{\beta^2}{\eta^2} & \frac{-\psi(2)}{\eta} \\ \frac{-\psi(2)}{\eta} & \frac{\psi(2)^2 + \pi^2/6}{\beta^2} \end{bmatrix}$$

e

$$S(\eta, \beta) = \frac{6n}{\pi^2} \begin{bmatrix} \frac{(\psi(2)^2 + \pi^2/6)\eta^2}{\beta^2} & \psi(2)\eta \\ \psi(2)\eta & \beta^2 \end{bmatrix}.$$

em que $\psi(2) = 1 - \gamma$ e γ é a constante de Euler, cujo valor é aproximadamente 0,577216.

Temos que $\{s_{1,1}(\beta, \eta)\}^{-1/2}$ e $\{h_{2,2}(\beta, \eta)\}^{1/2}$ se fatoram na forma

$$\{s_{1,1}(\beta, \eta)\}^{-1/2} = \frac{1}{\eta} \sqrt{\frac{\beta^2}{(\psi(2)^2 + \frac{\pi^2}{6})}} = f_1(\eta)g_1(\beta) \quad \text{e}$$

$$\{h_{2,2}(\beta, \eta)\}^{1/2} = \frac{1}{\beta} \sqrt{\psi(2)^2 + \frac{\pi^2}{6}} = f_2(\eta)g_2(\beta).$$

Então, utilizando o Corolário 2.1, as funções a priori de referência para η e β dado η são dadas, respectivamente, por

$$\pi(\eta) \propto f_1(\eta) \propto \frac{1}{\eta} \quad \text{e} \quad \pi(\beta|\eta) \propto g_2(\beta) \propto \frac{1}{\beta}.$$

A função a priori de referência conjunta de (β, η) (Yang e Berger, 1996) é dada por

$$\pi(\eta, \beta) \propto \frac{1}{\eta\beta}. \tag{2.17}$$

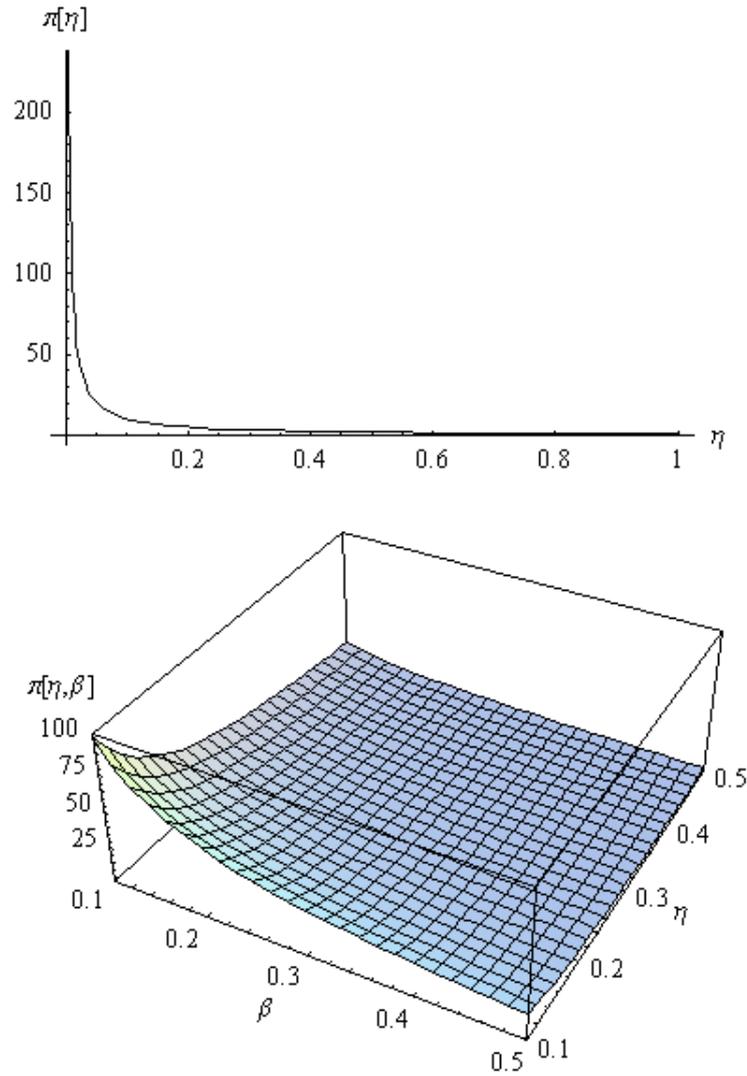


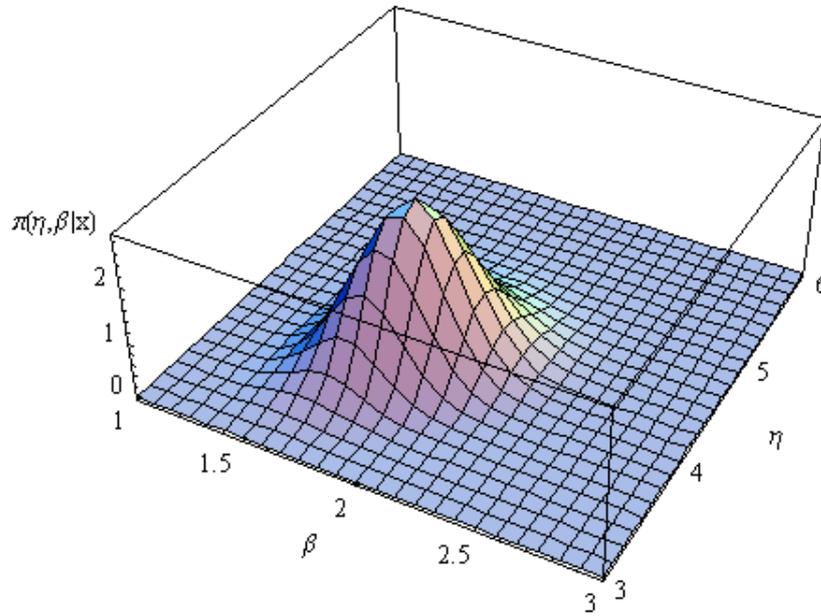
FIGURA 2.1: Funções densidade a priori para η e para (η, β) , respectivamente.

A Figura 2.1 representa a função a priori de referência para η e a função a priori conjunta para (η, β) .

Utilizando as equações (2.17) e (1.4), obtemos a distribuição a posteriori conjunta para η e β , expressa por

$$\pi(\eta, \beta | \mathbf{x}) \propto \frac{1}{\beta \eta} \prod_{i=1}^n \frac{\beta}{\eta^\beta} x_i^{\beta-1} \exp \left[- \left(\frac{x_i}{\eta} \right)^\beta \right]. \quad (2.18)$$

A Figura 2.2 representa a distribuição a posteriori conjunta de referência de (β, η) .

FIGURA 2.2: Distribuição a posteriori conjunta de (η, β) .

As distribuições a posteriori condicionais de β e η são dadas, respectivamente, por

$$\pi(\beta | \mathbf{x}, \eta) \propto \frac{\beta^{n-1}}{\eta^{n\beta+1}} \prod_{i=1}^n x_i^{\beta-1} \exp \left[- \left(\frac{x_i}{\eta} \right)^\beta \right] \quad (2.19)$$

e

$$\pi(\eta | \mathbf{x}, \beta) \propto \frac{1}{\eta^{n\beta+1}} \prod_{i=1}^n \exp \left[- \left(\frac{x_i}{\eta} \right)^\beta \right]. \quad (2.20)$$

Definidas a distribuição a posteriori e as condicionais, o próximo passo é a obtenção das medidas resumo a posteriori (média, desvio padrão (DP) e intervalo de credibilidade (Int. Cred. de 95%)) para os parâmetros de interesse. Como as distribuições condicionais não são conhecidas, devemos utilizar métodos de aproximação para a obtenção dessas medidas, como por exemplo, o método Metropolis-Hastings. Vale ressaltar que esse método é utilizado ao longo do trabalho devido ao fato das distribuições a posteriori não possuírem forma fechada. A seguir, exemplificamos a metodologia descrita nesta seção com um conjunto de dados de sobrevivência simulado. O *software* estatístico utilizado no trabalho foi o *R* e o conjunto de dados se encontra no Apêndice A. O computador utilizado para a execução dos mesmos foi um PC com processador AMD Turion 64 mobile (512 KB L2 cache, 1,58 GHz), com 960 MB de memória

RAM. Em todos os exemplos apresentados neste trabalho, a convergência das cadeias simuladas foi verificada através do critério de Gelman e Rubin (Gelman e Rubin, 1992). Alguns métodos gráficos, como superposição das densidades, gráfico das sequências e médias ergódicas também foram utilizados com a mesma finalidade. Neste trabalho, os gráficos das médias ergódicas foram construídos considerando apenas a primeira cadeia simulada dos parâmetros em questão.

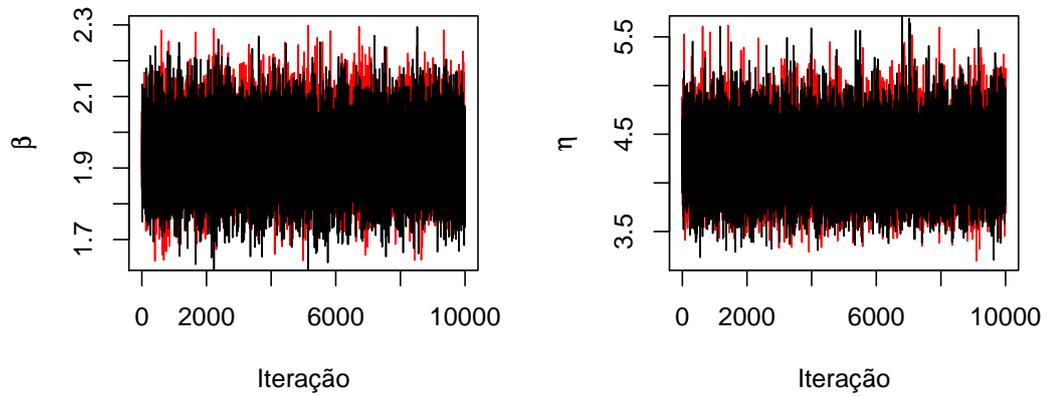
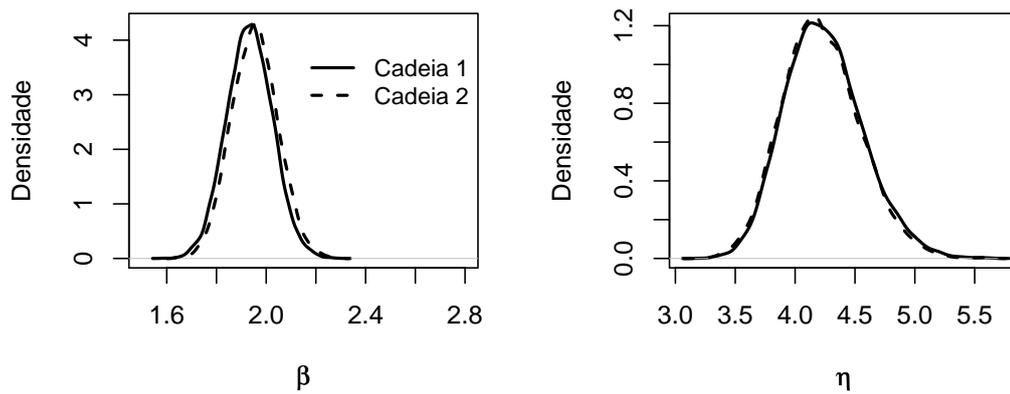
Exemplo 2.1 *Geramos um conjunto de dados de tamanho igual a 50 (Apêndice A), segundo o modelo Weibull (1.1), com parâmetros $\beta = 2$ e $\eta = 4$. Obtivemos duas amostras das distribuições a posteriori marginais de β e η , utilizando o algoritmo de Metropolis-Hastings (Hastings, 1970). As medidas resumo das amostras das cadeias 1 e 2, considerando aquecimento (burn in) de 15.000 com tamanho 10.000 e salto igual a 10, encontram-se na Tabela 2.1. O critério de Gelman e Rubin forneceu $\hat{R} = 1.00$ para ambas as cadeias dos parâmetros, indicando convergência das mesmas. O tempo de processamento para cada cadeia simulada, em segundos, é igual a 53.*

TABELA 2.1: Medidas resumo a posteriori para os parâmetros do modelo (1.1).

	Parâmetros	Média	DP	Int. Cred. de 95%
Cadeia 1	β	1,93	0,09	[1,75; 2,11]
	η	4,23	0,33	[3,65; 4,92]
Cadeia 2	β	1,93	0,09	[1,75; 2,12]
	η	4,22	0,32	[3,63; 4,88]

Verificamos na Tabela 2.1 que as médias a posteriori para ambos os parâmetros se encontram próximas do verdadeiro valor dos mesmos.

A Figura 2.3 mostra que há indícios de convergência para ambos os parâmetros, pois as cadeias se comportam dentro de uma “faixa”, ou seja, os valores gerados são próximos uns dos outros. Isso também pode ser visto na Figura 2.4, uma vez que as densidades estimadas dos parâmetros de interesse para as duas cadeias geradas estão próximas. Através da Figura 2.5, observamos que a média da cadeia converge para um determinado valor, indicando, após aproximadamente 15.000 iterações, a convergência das cadeias.

FIGURA 2.3: Traço das cadeias β e η .FIGURA 2.4: Funções densidade a posteriori estimadas de β e η .

A Figura 2.6 representa a curva de confiabilidade estimada considerando $\beta = 1,95$ e $\eta = 4,24$.

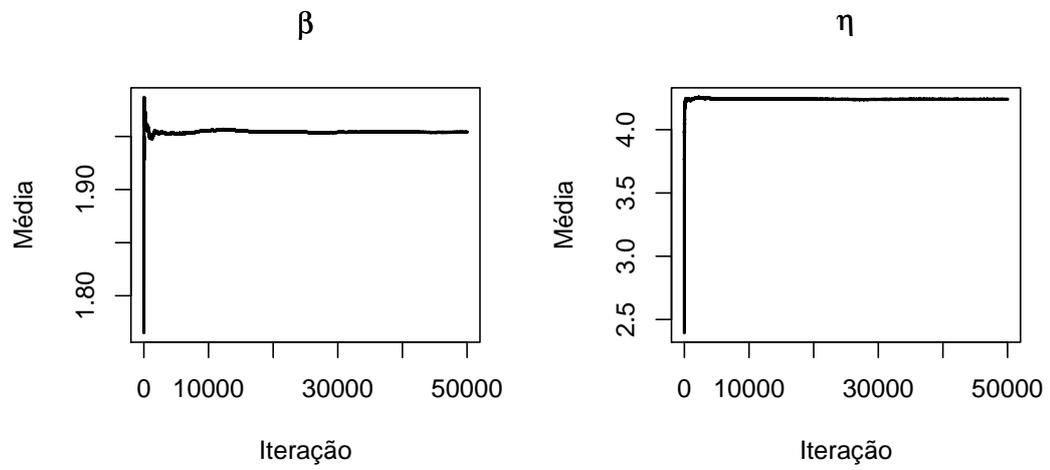
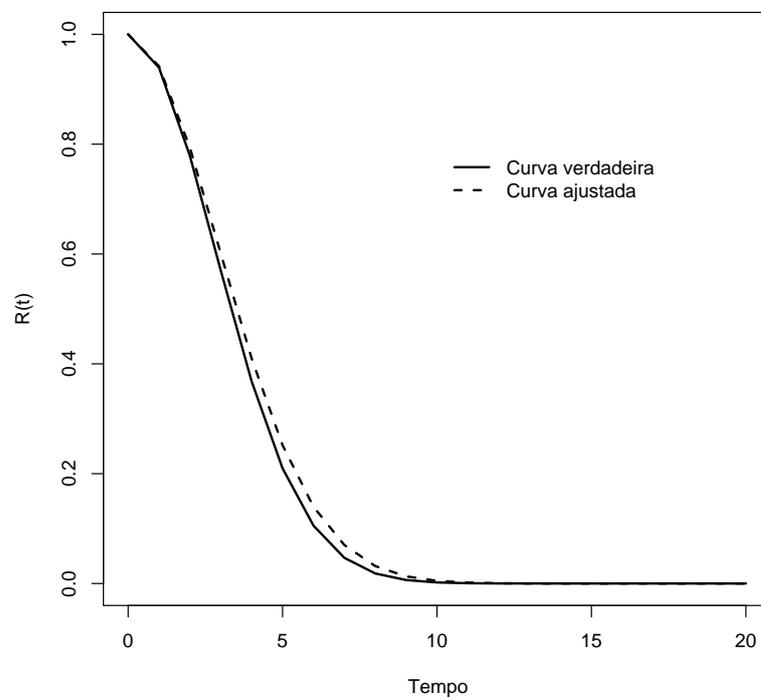
FIGURA 2.5: Médias ergódicas de β e η .

FIGURA 2.6: Função de confiabilidade estimada do modelo Weibull.

2.3.1 Teste de hipótese para o parâmetro de forma (β) do modelo Weibull

Nesta subseção, desenvolvemos o BRC descrito na Seção 2.2 para testarmos hipóteses envolvendo o parâmetro de forma do modelo Weibull.

Sejam os dados observados $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$ uma amostra aleatória do modelo (1.1). Considere η conhecido e o problema de testar se os dados são ou não compatíveis com a hipótese precisa $H_0 \equiv \{\beta = \beta_0\}$, sobre o valor do parâmetro de forma.

Dado η , a divergência logarítmica de Kullback-Leiber entre os modelos $Wei(\mathbf{x}|\beta_j, \eta)$ e $Wei(\mathbf{x}|\beta_i, \eta)$ é

$$\begin{aligned} \kappa(\beta_j|\beta_i) &= n \int Wei(\mathbf{x}|\beta_i, \eta) \log \left[\frac{Wei(\mathbf{x}|\beta_i, \eta)}{Wei(\mathbf{x}|\beta_j, \eta)} \right] dx \\ &= (-1 - \gamma) + \frac{\gamma\beta_j}{\beta_i} + \Gamma \left(\frac{\beta_i + \beta_j}{\beta_i} \right) + \log(\beta_i) - \log(\beta_j), \end{aligned} \quad (2.21)$$

onde γ é dada na página 30. É facilmente verificado que $\kappa(\beta_j|\beta_i) < \kappa(\beta_i|\beta_j)$ se, e somente se, $\beta_j < \beta_i$. Então, a discrepância intrínseca entre $p(\mathbf{x}|\beta_0, \eta)$ e $p(\mathbf{x}|\beta, \eta)$ é

$$\delta(\beta_0, \beta) = n \begin{cases} \kappa(\beta_0|\beta), & \text{se } \beta < \beta_0 \\ \kappa(\beta|\beta_0), & \text{se } \beta > \beta_0 \end{cases}.$$

A correspondente esperança a posteriori de referência $d(\beta_0, \mathbf{x})$ é dada por

$$d(\beta_0, \mathbf{x}) = \int_0^\infty \delta(\beta_0, \beta) \pi(\beta, \eta|\mathbf{x}) d\beta. \quad (2.22)$$

Porém, esse cálculo não pode ser determinado analiticamente. Assim, devemos utilizar um método numérico de solução de integrais para obtermos a estatística intrínseca $d(\beta_0, \mathbf{x})$.

Como o caso particular em que $\beta_0 = 1$ requer uma atenção especial, pois temos a hipótese de que os dados observados são compatíveis com uma distribuição Exponencial, construímos o BRC para esse caso.

Considere o modelo Weibull (1.1),

$$\mathcal{M} \equiv \left\{ p_1(\mathbf{x}|\boldsymbol{\beta}, \boldsymbol{\eta}) = \frac{\beta}{\eta^\beta} x^{\beta-1} \exp \left[- \left(\frac{x}{\eta} \right)^\beta \right], \beta, \eta > 0 \right\},$$

o modelo estatístico, o qual é assumido ter gerado os dados $\mathbf{x} \in \mathbb{R}^+$ e considere a hipótese nula a ser testada $H_0 \equiv \{\beta = 1\}$. Para decidirmos se aceitamos ou não a hipótese nula, utilizamos o BRC descrito na Seção 2.2. Para a implementação do mesmo, devemos:

1. Calcular a discrepância intrínseca ($\delta \{\mathcal{M}, \mathcal{M}_0\}$) entre os modelos \mathcal{M} e

$$\mathcal{M}_0 = \left\{ p_2(\mathbf{x}|\beta = 1, \eta) = \frac{1}{\eta} \exp \left[- \left(\frac{\mathbf{x}}{\eta} \right) \right], \mathbf{x} \in \mathbb{R}^+, \eta > 0 \right\}.$$

A discrepância intrínseca $\delta \{\mathcal{M}, \mathcal{M}_0\}$ é dada por

$$\delta(\mathcal{M}, \mathcal{M}_0) = \min \delta \{ \kappa(p_2|p_1), \kappa(p_1|p_2) \}.$$

Então, utilizando (2.21) temos que a discrepância intrínseca entre \mathcal{M} e \mathcal{M}_0 é dada por

$$\delta(\mathcal{M}_0, \mathcal{M}) = n \begin{cases} \kappa(\mathcal{M}_0|\mathcal{M}), & \text{se } \beta < 1 \\ \kappa(\mathcal{M}|\mathcal{M}_0), & \text{se } \beta > 1 \end{cases}.$$

2. Derivar a correspondente esperança a posteriori de referência

$$d(\mathcal{M}_0, \mathbf{x}) = E[\delta(\mathcal{M}_0, \mathcal{M})|\mathbf{x}] = \int_{\mathbb{R}^+} \delta(\mathcal{M}_0, \mathcal{M}) \pi(\beta|\mathbf{x}) d\beta,$$

e estipular este número como uma medida de evidência contra a hipótese nula H_0 .

Como esse cálculo não pode ser determinado analiticamente, utilizamos os valores dos parâmetros gerados pelo método de Metropolis-Hastings para obtermos a estatística intrínseca $d(\mathcal{M}_0, \mathbf{x})$.

Considerando os dados do Exemplo 2.1, a medida de evidência $d(\mathcal{M}_0, \mathbf{x})$ é igual a 48,20, indicando evidência significativa contra a hipótese nula.

Nesta seção, estudamos o modelo Weibull do ponto de vista da inferência Bayesiana objetiva, construímos a função a priori de referência para os parâmetros de interesse, calculamos a distribuição a posteriori para os mesmos e desenvolvemos o teste de hipótese Bayesiano para o parâmetro de forma do modelo (1.1). Exemplificamos a metodologia com dados gerados, utilizando MCMC para a obtenção das amostras a posteriori para os parâmetros. Seguimos essa mesma idéia para a função de confiabilidade do modelo Weibull, que é de nosso interesse.

2.3.2 Inferência para a função de confiabilidade do modelo Weibull

Considerando que em geral o interesse do pesquisador é a função de confiabilidade $R(x)$, apresentamos a análise de referência para a função de confiabilidade do modelo Weibull (1.3) (ver Moala et al. 2008). Para encontrarmos a função a priori de referência, primeiramente, consideramos a reparametrização $W = \beta$ e $R = \exp \left[- \left(\frac{t_0}{\eta} \right)^\beta \right]$ (t_0 fixo) em (1.1), que pode ser reescrita da forma

$$f(x|t_0, R, W) = W \log(1/R) \left(\frac{x^{W-1}}{t_0^W} \right) \exp \left\{ - \left(\frac{x}{t_0} \right)^W \log(1/R) \right\}. \quad (2.23)$$

Fazendo $Y = \frac{X}{t_0}$ e considerando o jacobiano da transformação $J = |X|$ em (2.23) obtemos

$$f(y|R, W) = W \log \left(\frac{1}{R} \right) y^{W-1} \exp \left\{ -y^W \log \left(\frac{1}{R} \right) \right\}, \quad (2.24)$$

ou seja, $y|R, W \sim Weibull \left(W, \frac{1}{(\log(\frac{1}{R}))^{1/W}} \right)$.

A função de verossimilhança para os n dados de sobrevivência considerando o modelo dado em (2.24) é

$$L(R, W|\mathbf{y}) = \left(\log \left(\frac{1}{R} \right) \right)^n W^n \prod_{i=1}^n y_i^{(W-1)} R^{y_i^W}. \quad (2.25)$$

Na Figura 2.7, observamos que a função de verossimilhança (2.25) é aparentemente bem comportada, possuindo um ponto de máximo e se assemelha à função de verossimilhança de um modelo Normal.

Para obtermos a função a priori de referência, consideramos a metodologia descrita na Subseção 2.1.2 para o modelo (2.24), em que R é o parâmetro de interesse e W o parâmetro perturbador.

O logaritmo da função de verossimilhança (2.25) é

$$\begin{aligned} l(R, W|\mathbf{y}) = n \left[\log \left(\log \left(\frac{1}{R} \right) \right) + \log(W) \right] + \log(R) \sum_{i=1}^n y_i^W \\ + (W - 1) \sum_{i=1}^n \log(y_i). \end{aligned} \quad (2.26)$$

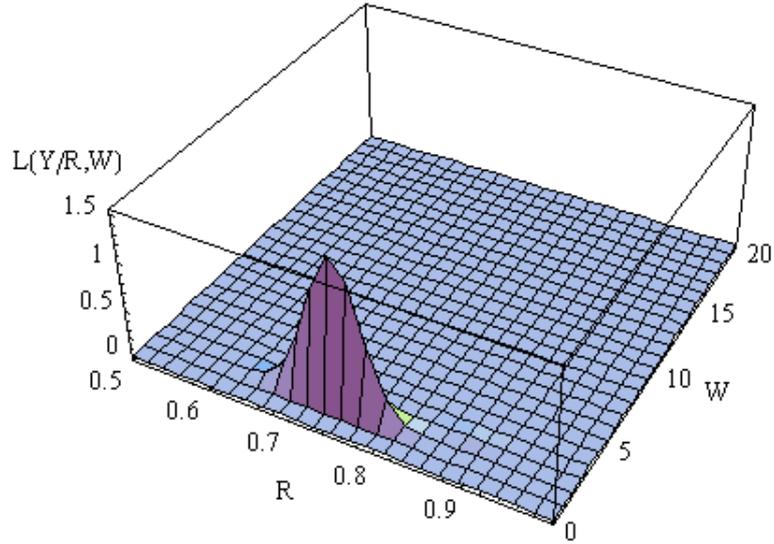


FIGURA 2.7: Função de verossimhança do modelo (2.25).

As derivadas primeiras e segundas de (2.26), que foram calculadas através do *software* Mathematica, são

$$\frac{\partial l(R, W|\mathbf{y})}{\partial R} = \frac{n}{R \log(R)} + \frac{\sum_{i=1}^n y_i^W}{R},$$

$$\frac{\partial l(R, W|\mathbf{y})}{\partial W} = \frac{n}{W} + \sum_{i=1}^n \log(y_i) + \log(R) \sum_{i=1}^n \log(y_i) y_i^W,$$

$$\frac{\partial^2 l(R, W|\mathbf{y})}{\partial R^2} = -\frac{n}{R^2 \log(R)^2} - \frac{n}{R^2 \log(R)} - \frac{\sum_{i=1}^n y_i^W}{R^2},$$

$$\frac{\partial^2 l(R, W|\mathbf{y})}{\partial W^2} = -\frac{n}{W^2} + \log(R) \sum_{i=1}^n \log(y_i)^2 y_i^W$$

e

$$\frac{\partial^2 l(R, W|\mathbf{y})}{\partial R \partial W} = \frac{\partial l(R, W)}{\partial W \partial R} = \frac{\sum_{i=1}^n \log(y_i) y_i^W}{R}.$$

As esperanças para o cálculo da matriz de informação de Fisher são

$$E \left[\frac{\partial^2 l(R, W|\mathbf{y})}{\partial R^2} \right] = -\frac{n}{R^2 \log(R)^2} - \frac{n}{R^2 \log(R)} - \frac{\sum_{i=1}^n E[y_i^W]}{R^2}, \text{ em que}$$

$$E[y_i^W] = \frac{1}{\log\left(\frac{1}{R}\right)}.$$

$$E \left[\frac{\partial^2 l(R, W|\mathbf{y})}{\partial W^2} \right] = -\frac{n}{W^2} + \log(R) \sum_{i=1}^n E[\log(y_i)^2 y_i^W], \text{ onde}$$

$$E [\log(y_i)^2 y_i^W] = \frac{(-2 + \gamma)\gamma + \pi^2/6 + \log \log \left(\frac{1}{R}\right) (-2 + 2\gamma + \log \log \left(\frac{1}{R}\right))}{W^2 \log \left(\frac{1}{R}\right)}.$$

$$E \left[\frac{\partial^2 l(R, W | \mathbf{y})}{\partial R \partial W} \right] = \frac{\sum_{i=1}^n E [\log(y_i) y_i^W]}{R}, \text{ em que}$$

$$E [\log(y_i) y_i^W] = \frac{-1 + \gamma + \log \log \left(\frac{1}{R}\right)}{W \log(R)}.$$

Assim, a matriz de informação de Fisher e sua inversa são dadas por

$$H(R, W) = n \begin{bmatrix} \frac{1}{R^2 \log \left(\frac{1}{R}\right)^2} & \frac{b_1(R)}{RW \log \left(\frac{1}{R}\right)} \\ \frac{b_1(R)}{RW \log \left(\frac{1}{R}\right)} & \frac{b_2(R)}{W^2} \end{bmatrix}$$

e

$$S(R, W) = \frac{6n}{\pi^2} \begin{bmatrix} R^2 \log \left(\frac{1}{R}\right)^2 b_2(R) & -RW \log \left(\frac{1}{R}\right) b_1(R) \\ -RW \log \left(\frac{1}{R}\right) b_1(R) & W^2 \end{bmatrix},$$

em que $b_1(R) = -\psi(2) + \log \left(\log \left(\frac{1}{R}\right)\right)$, $b_2(R) = \psi(2)^2 + \pi^2/6 + \log \left(\log \left(\frac{1}{R}\right)\right) (-2\psi(2) + \log \left(\log \left(\frac{1}{R}\right)\right))$, γ e $\psi(2)$ dados na página 30.

Temos que $\{s_{1,1}(R, W)\}^{-1/2}$ e $\{h_{2,2}(R, W)\}^{1/2}$ se fatoram na forma

$$\{s_{1,1}(R, W)\}^{-1/2} = \frac{1}{R \log \left(\frac{1}{R}\right) \sqrt{b_2(R)}} \cdot 1 = f_1(R) g_1(W) \text{ e}$$

$$\{h_{2,2}(R, W)\}^{1/2} = \frac{1}{W} \sqrt{b_2(R)} = f_2(R) g_2(W).$$

Então, utilizando o Corolário 2.1, as funções a priori de referência para R e para W dado R são dadas, respectivamente, por

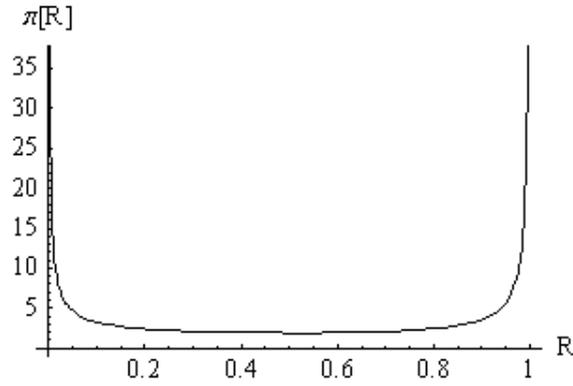
$$\pi(R) \propto f_1(R) \propto \frac{1}{R \log \left(\frac{1}{R}\right) \sqrt{b_2(R)}} \text{ e} \quad (2.27)$$

$$\pi(W|R) \propto g_2(W) \propto \frac{1}{W}. \quad (2.28)$$

A função a priori de referência conjunta de (R, W) é dada por

$$\pi(R, W) \propto \frac{1}{WR \log \left(\frac{1}{R}\right) \sqrt{b_2(R)}}. \quad (2.29)$$

A Figura 2.8 mostra a função a priori de referência para a confiabilidade R em (2.27).

FIGURA 2.8: Função a priori de referência para a confiabilidade R .

Para inferir sobre a função de confiabilidade do modelo Weibull, devemos encontrar sua distribuição a posteriori. Assim, considerando as equações (2.25) e (2.29), a distribuição a posteriori conjunta de R e W é dada por

$$\pi(R, W|y) \propto \frac{(\log(\frac{1}{R}))^{n-1} W^{n-1} \prod_{i=1}^n y_i^{(W-1)} R^{y_i^W - 1}}{\sqrt{\psi(2)^2 + \pi^2/6 + \log(\log(\frac{1}{R}))(-2\psi(2) + \log(\log(\frac{1}{R})))}} \quad (2.30)$$

e a distribuição a posteriori marginal para R é

$$\pi(R|y) \propto a \int_0^\infty W^{n-1} \prod_{i=1}^n y_i^{(W-1)} R^{y_i^W - 1} dW, \quad (2.31)$$

em que $a = \frac{(\log(\frac{1}{R}))^{n-1}}{\sqrt{b_2(R)}}$.

As distribuições a posteriori condicionais de R e W são dadas, respectivamente, por

$$\pi(R|y, W) \propto \frac{(\log(\frac{1}{R}))^{n-1} \prod_{i=1}^n R^{y_i^W - 1}}{\sqrt{\psi(2)^2 + \pi^2/6 + \log(\log(\frac{1}{R}))(-2\psi(2) + \log(\log(\frac{1}{R})))}} \quad (2.32)$$

e

$$\pi(W|y, R) \propto W^{n-1} \prod_{i=1}^n y_i^{(W-1)} R^{y_i^W - 1}. \quad (2.33)$$

A seguir, exemplificamos a metodologia desenvolvida nesta subseção com o conjunto de dados de sobrevivência simulado no Exemplo 2.1. Encontramos as estimativas da função de confiabilidade para vários valores de t_0 , construindo a curva de confiabilidade estimada.

Exemplo 2.2 Considerando o conjunto de dados gerado no Exemplo 2.1 e utilizando o algoritmo de Metropolis-Hastings na equação (2.30), encontramos duas amostras da distribuição a posteriori de R para $t_0 = 2$, cujo valor teórico de R é 0,778. As medidas resumo das amostras das cadeias 1 e 2, considerando aquecimento (burn in) de 20.000 com tamanho 10.000 e salto igual a 5, encontram-se na Tabela 2.2. O critério de Gelman e Rubin forneceu $\hat{R} = 1.00$ para ambas as cadeias dos parâmetros, indicando convergência das mesmas. O tempo de processamento para cada cadeia simulada, em segundos, é igual a 200,48.

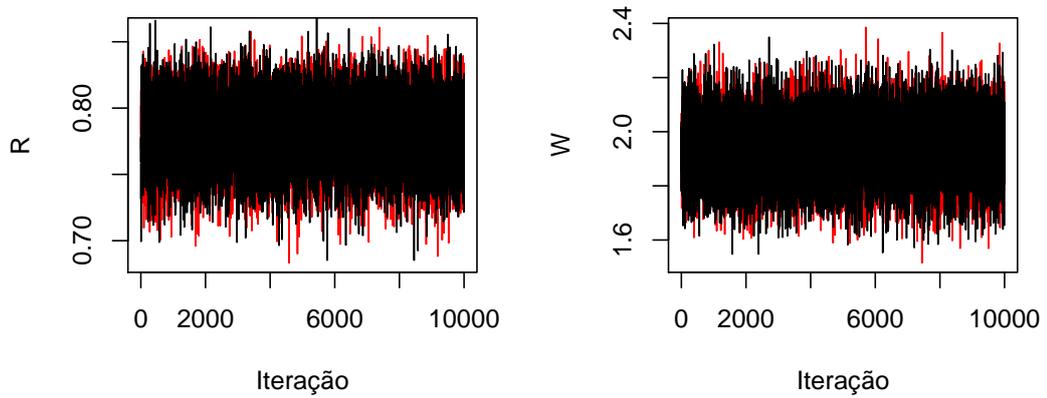
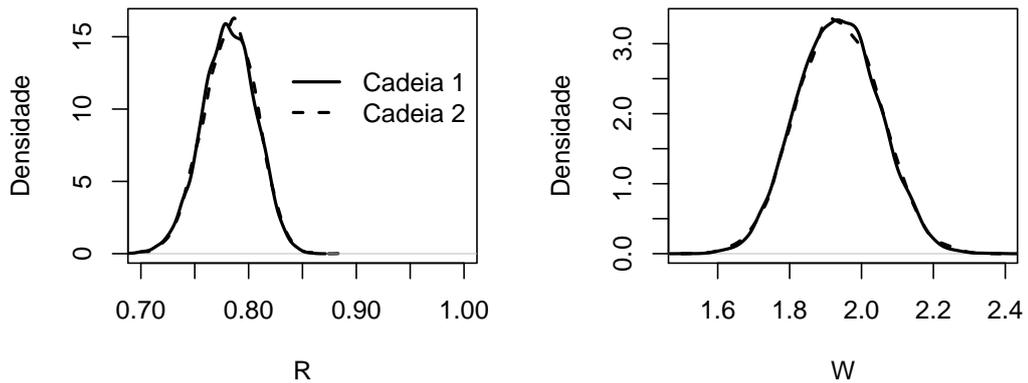
TABELA 2.2: Medidas resumo a posteriori para os parâmetros do modelo (2.25).

	Parâmetro	Média	DP	Int. Cred. de 95%
Cadeia 1	R	0,782	0,024	[0,732; 0,827]
	W	1,935	0,112	[1,723; 2,154]
Cadeia 2	R	0,783	0,024	[0,734; 0,828]
	W	1,936	0,114	[1,716; 2,158]

A Figura 2.9 mostra que há indícios de convergência para ambos os parâmetros, pois as cadeias se comportam de maneira homogênea. Isso também pode ser observado na Figura 2.10, uma vez que as densidades estimadas dos parâmetros de interesse para as duas cadeias geradas estão próximas. Através da Figura 2.11, verificamos que a média da cadeia converge para um determinado valor, indicando, após aproximadamente 20.000 iterações, a convergência das cadeias.

Para cada valor selecionado de t_0 , encontramos a respectiva distribuição a posteriori para R , verificamos a convergência de maneira análoga à apresentada para $t_0 = 2$. Estimamos a curva de confiabilidade para alguns valores de t_0 , tomando como estimador pontual a média a posteriori da cadeia 1. A Figura 2.12 apresenta a curva estimada para alguns valores de t_0 (0, 25, 0,5, 1, 2, 3, 4, 5, 6, 7, 8, 9 e 10), os intervalos de credibilidade de 95% e a curva teórica. Observamos que a curva de confiabilidade estimada está próxima da teórica e a mesma se encontra entre os limites de credibilidade.

Estudamos a análise de referência objetiva para os parâmetros do modelo

FIGURA 2.9: Traço das cadeias de R e W .FIGURA 2.10: Funções densidade a posteriori estimadas de R e W .

Weibull e para sua confiabilidade. Exemplificamos essa metodologia para um conjunto de dados de sobrevivência simulado e utilizamos o método de Metropolis-Hastings para obtenção das amostras das distribuições a posteriori para os parâmetros de interesse. Como o objetivo deste trabalho é utilizar a análise de referência em modelos de riscos competitivos, no capítulo seguinte, revisamos conceitos destes modelos.

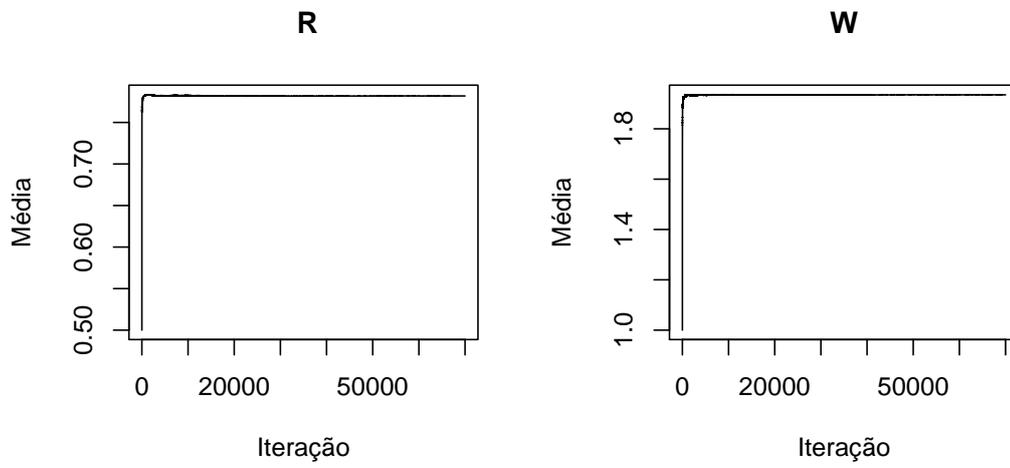
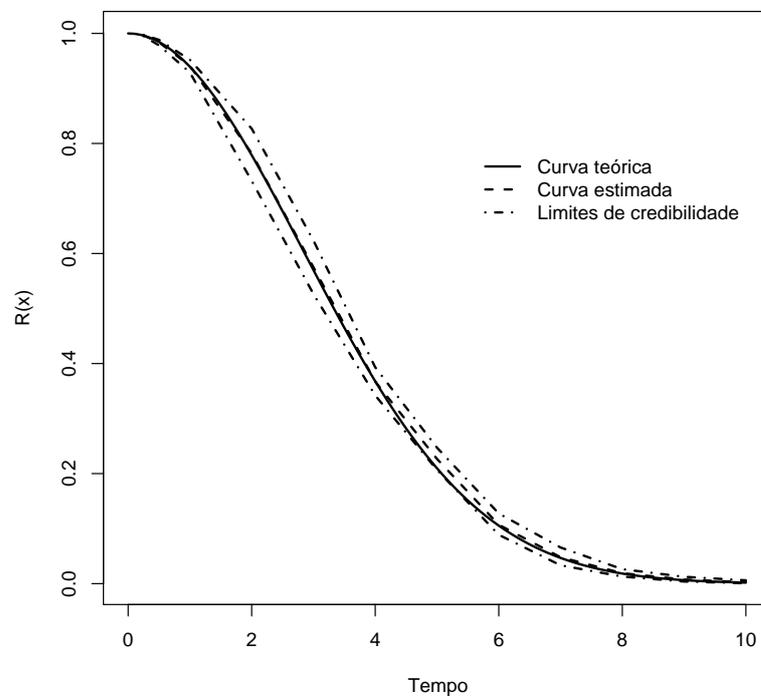
FIGURA 2.11: Médias ergódicas para R e W .

FIGURA 2.12: Curva de confiabilidade estimada.

Capítulo 3

Riscos competitivos

Neste capítulo, apresentamos as definições básicas e a notação utilizada na análise do tempo de falha em modelos de fatores de risco competitivos (ver Lawless, 1982; Coque, 2004; Gichangi e Vach, 2005; Lindqvist, 2006). A Seção 3.1 contém uma introdução sobre riscos competitivos; na Seção 3.2, discutiremos sobre a construção da função de verossimilhança. A Seção 3.3 apresenta o modelo Weibull na aplicação de riscos competitivos; na Seção 3.4, apresentamos a análise de referência para o modelo em questão. Na Seção 3.5, desenvolvemos o BRC para igualdade dos parâmetros dos fatores de risco e, por fim, na Seção 3.6, apresentamos a análise de referência para a função de confiabilidade do modelo.

3.1 Introdução

Muitos estudos em análise de sobrevivência estão centrados em apenas uma causa de morte (falha). Entretanto, causas que competem entre si para que o evento de interesse ocorra são também comuns. Os estudos que consideram diversas causas de falha são conhecidos na literatura como análise de dados de riscos competitivos. A teoria de riscos competitivos teve início com David Bernoulli, em 1760, numa tentativa de desvendar o risco de morte por varíola e por outras causas. Esse é um exemplo clássico de riscos competitivos, em que indivíduos estão expostos a múltiplas causas de morte. Aplicações similares ocorrem em demografia, ciências atuariais e confiabilidade.

Em confiabilidade, o termo riscos competitivos se refere a situações em que sistemas sob investigação estão expostos a mais de uma possibilidade de ocorrência do evento de interesse. As diferentes possibilidades são as causas e o acontecimento do primeiro evento de interesse, a falha.

Uma situação comum é que um sistema esteja sujeito a diversos fatores de falha, e que apenas um desses fatores causará sua falha, ou seja, a falha de qualquer componente resulta na falha de todo o sistema. Portanto, dizemos que esses fatores estão competindo para extinguir a vida desse sistema. O modelo para o tempo de vida sujeito a esses fatores é conhecido como modelo de fatores de risco competitivos ou, por apenas, modelo de riscos competitivos.

3.1.1 Definição de riscos competitivos

Suponha que para um particular sistema existam \mathcal{C} ($\mathcal{C} \geq 2$) potenciais causas independentes de falhas. Assim, cada sistema estará sujeito a \mathcal{C} fatores de falha, ou seja, estará sujeito à falha de algum de seus componentes, sendo chamados de fatores de risco competitivos. Cada componente está exposto a \mathcal{C} fatores de risco independentes, sendo que somente um causará a falha do sistema. Esse fator é conhecido como a causa da falha.

Suponha que $X_{i1}, X_{i2}, \dots, X_{i\mathcal{C}}$ sejam variáveis aleatórias independentes, contínuas e positivas denominadas tempos de vida para uma determinada unidade i ($i = 1, \dots, n$) exposta aos ℓ fatores ($\ell = 1, \dots, \mathcal{C}$). Ou seja, $X_{i\ell}$ representa o tempo de vida do i -ésimo sistema (indivíduo) sob o ℓ -ésimo fator de risco. Contudo, todos os \mathcal{C} fatores agem simultaneamente, e como somente um irá causar a falha do sistema, adotaremos outra variável aleatória $T_i = \min \{X_{i1}, X_{i2}, \dots, X_{i\mathcal{C}}\}$. Esse fato é ilustrado no diagrama da Figura 3.1. Além disso, observaremos uma variável indicadora $\delta_{i\ell}$, que representa a causa da falha da i -ésima unidade sob o ℓ -ésimo fator de risco, ou seja,

$$\delta_{i\ell} = \begin{cases} 1, & \text{se a unidade } i \text{ falhou por } \ell, \text{ ou seja, } T_i = X_{i\ell} \\ 0, & \text{se a unidade } i \text{ não falhou por } \ell, \text{ ou seja, } T_i \neq X_{i\ell} \end{cases}.$$

Considerando uma amostra de n observações nas condições anteriores, os

dados estão dispostos da seguinte forma: $(T_1, \delta_{11}, \delta_{12}, \dots, \delta_{1c}), \dots, (T_n, \delta_{n1}, \delta_{n2}, \dots, \delta_{nc})$. Na prática, os dados são frequentemente censurados à direita por algum fator de risco independente.

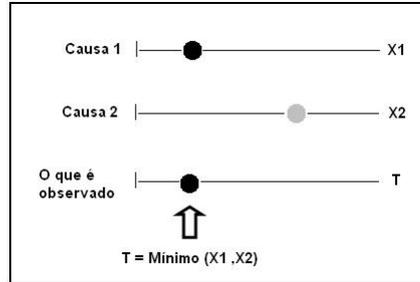


FIGURA 3.1: Riscos Competitivos. Diagrama de duas causas de falhas.

A seguir, apresentamos um exemplo real de dados referentes a riscos competitivos. Esses dados foram retirados de Coque (2004) e os mesmos se encontram na Tabela 3.1 .

Exemplo 3.1 *A bomba HDR-750CV fornece refrigeração aos mancais de uma prensa que está posicionada ao lado de um forno a 980°C. A bomba também fica próxima ao forno e, portanto, sujeita à alta temperatura, por isso, a degradação de alguns componentes da bomba é frequente. A prensa mencionada trava frequentemente, pois a bomba apresenta dois fatores de risco de falha:*

Fator A: Degradação da vedação do rotor e

Fator B: Degradação da lubrificação.

A empresa possui dados de falha da bomba em questão, em que algumas falhas são devidas à degradação da vedação do rotor (Fator A) e outras falhas, por degradação da lubrificação (Fator B). Os dados estão apresentados na Tabela 3.1. O objetivo da empresa é estimar os fatores de risco para a falha da bomba e determinar qual é o fator de risco mais crítico para a falha da bomba.

Os dados apresentados no Exemplo 3.1 são utilizados na Seção 3.3, onde é desenvolvida a abordagem clássica para o modelo Weibull na aplicação de riscos competitivos, e no Capítulo 4, onde apresentamos uma aplicação da metodologia apresentada no Capítulo 2.

TABELA 3.1: Histórico de falhas da bomba HDR-750CV.

Dias	Fator de Risco	Dias	Fator de Risco	Dias	Fator de Risco
310	B	690	A	1150	A
320	A	720	B	1230	A
360	B	730	B	1370	A
380	B	800	A	1380	A
450	A	810	A	1400	A
525	B	810	B	1420	B
530	B	840	A	1430	A
565	A	850	B	1490	B
620	A	870	A	1540	B
648	A	880	B	1660	B
665	B	930	A	1740	A
670	A	950	B	1850	B
675	A	1010	A	1930	B

3.2 Função de verossimilhança

Em análise de sobrevivência/confiabilidade os dados são frequentemente censurados à direita, isto é, os tempos de vida são conhecidos somente para uma parte das unidades. Quanto a outra parte, sabemos apenas que as unidades sobreviveram até determinado instante (observado) (Rodrigues et al., 2008).

Considere os dados da forma $(\mathbf{T}, \boldsymbol{\delta})$ para n unidades independentes observadas. Na prática, tais dados são frequentemente censurados à direita por alguma causa independente dos C riscos. Por exemplo, a censura poderia ser devido ao tempo limite do experimento. Se a i -ésima unidade é não-censurada, ou seja, falhou por ℓ , observamos o tempo de falha T_i e a indicadora de falha $\delta_{i\ell} = 1$. Por outro lado, se a i -ésima unidade é censurada à direita sabemos apenas que $X_{i\ell} > T_i$. Então, a função de verossimilhança é dada por (Lindqvist, 2006)

$$L(\theta_1, \dots, \theta_C | \mathbf{t}) = \prod_{\ell=1}^C \prod_{i=1}^n f(t_i | \theta_\ell)^{\delta_{i\ell}} R(t_i | \theta_\ell)^{1-\delta_{i\ell}} \quad (3.1)$$

$$= \prod_{\ell=1}^c L_{\ell}(\theta_{\ell}|\mathbf{t}). \quad (3.2)$$

Podemos escrever $L(\theta_1, \dots, \theta_c|\mathbf{t})$ como o produto de verossimilhanças $L_{\ell}(\theta_{\ell}|\mathbf{t})$, em que $L_{\ell}(\theta_{\ell}|\mathbf{t}) = \prod_{i=1}^n f_{\ell}(t_i)^{\delta_{i\ell}} R_{\ell}(t_i)^{1-\delta_{i\ell}}$ tem a forma da função de verossimilhança de uma amostra censurada (maiores detalhes ver Rodrigues et al., 2008) com a ℓ -ésima causa de falha em que todas as observações, onde $T_i \neq X_{i\ell}$, são tratadas como censura.

Note que, quando analisamos um determinado fator ℓ , o modelo de fatores de risco competitivos considera como falha os tempos em que as falhas ocorreram pelo fator ℓ e considera como censura, além dos tempos que não apresentaram falhas, os tempos de falha provocados por fatores diferentes de ℓ .

Para obtenção das estimativas dos parâmetros do modelo de riscos competitivos, maximiza-se $\log(L(\theta_1, \dots, \theta_c|\mathbf{t}))$ em (3.1).

3.3 O modelo Weibull na aplicação de riscos competitivos

Considere uma amostra de n tempos t_i ($i = 1, 2, \dots, n$) com os respectivos rótulos $\delta_{i\ell}$. Suponha que para uma população que siga o modelo Weibull existam C fatores de falha independentes.

As funções densidade, de confiabilidade e de risco sob o modelo Weibull na aplicação de riscos competitivos, para o indivíduo i , são dadas, respectivamente, por

$$f(t_i|\beta_{\ell}, \eta_{\ell}) = \frac{\beta_{\ell}}{\eta_{\ell}^{\beta_{\ell}}} t_i^{\beta_{\ell}-1} \exp \left[- \left(\frac{t_i}{\eta_{\ell}} \right)^{\beta_{\ell}} \right], \quad \ell = 1, 2, \dots, C, \quad (3.3)$$

$$R(t_i|\beta_{\ell}, \eta_{\ell}) = \exp \left[- \left(\frac{t_i}{\eta_{\ell}} \right)^{\beta_{\ell}} \right] \quad (3.4)$$

e

$$h(t_i|\beta_{\ell}, \eta_{\ell}) = \frac{\beta_{\ell}}{\eta_{\ell}^{\beta_{\ell}}} t_i^{\beta_{\ell}-1}. \quad (3.5)$$

Substituindo (3.3) e (3.4) em (3.1), a função de verossimilhança sob o modelo Weibull assume a forma

$$L(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t}) = \prod_{i=1}^n \prod_{\ell=1}^{\mathbf{C}} \left(\frac{\beta_{\ell}}{\eta_{\ell}^{\beta_{\ell}}} t_i^{\beta_{\ell}-1} \right)^{\delta_{i\ell}} \exp \left[- \left(\frac{t_i}{\eta_{\ell}} \right)^{\beta_{\ell}} \right]. \quad (3.6)$$

Neste trabalho, consideramos a abordagem de riscos competitivos na presença de apenas dois fatores de riscos, ou seja, $\mathbf{C} = 2$.

Consideramos os dados da Tabela 3.1 para obtenção das estimativas de máxima verossimilhança para os parâmetros do modelo de riscos competitivos (3.6). As estimativas foram obtidas minimizando $-\log(L(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\eta}))$ através da função *optim* contida no *software* estatístico *R*. A Tabela 3.2 mostra os resultados obtidos, onde se encontram as EMV, desvio padrão (DP) e intervalo de confiança (IC) de 95% para os parâmetros do modelo em questão, onde observamos que todos os parâmetros de interesse são significativos a um nível de 5%, pois o valor 0 não pertence ao intervalo de confiança de 95%.

TABELA 3.2: EMV, Desvio Padrão e IC de 95% para os parâmetros do modelo (3.6).

Parâmetro	EMV	DP	IC de 95%
β_A	2,42	0,42	[1,60 ; 3,25]
η_A	1457,87	144,33	[1174,99 ; 1740,76]
β_B	2,16	0,40	[1,40 ; 2,94]
η_B	1489,63	159,90	[1176,21 ; 1803,05]

3.4 Análise de referência

Nesta seção, desenvolvemos a construção da função a priori de referência para os parâmetros do modelo Weibull na aplicação de riscos competitivos (3.6), considerando a metodologia descrita na Subseção 2.1.3.

Para a obtenção da função a priori de referência fazemos uso da suposição de normalidade assintótica da distribuição a posteriori e consideramos que o

parâmetro de interesse é η_1 e os demais β_1, η_2 e β_2 são parâmetros perturbadores, uma vez que temos apenas dois fatores competitivos de risco.

O logaritmo da função de verossimilhança (3.6) com $\mathbf{C} = 2$, é

$$l(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\eta}) = \sum_{i=1}^n \sum_{\ell=1}^2 \left\{ \delta_{i\ell} [\log(\beta_\ell) - \beta_\ell \log(\eta_\ell) + (\beta_\ell - 1) \log(t_i)] - \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \right] \right\} \quad (3.7)$$

As derivadas de primeira e segunda ordem de (3.7) para $\ell = 1, 2$, obtidas através do *software* Mathematica, são dadas por

$$\begin{aligned} \frac{\partial l(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t})}{\partial \eta_\ell} &= \sum_{i=1}^n \frac{-\beta_\ell \delta_{i\ell}}{\eta_\ell} + \sum_{i=1}^n \frac{\beta_\ell t_i \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 1}}{\eta_\ell^2}, \\ \frac{\partial l(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t})}{\partial \beta_\ell} &= \sum_{i=1}^n \left(\frac{1}{\beta_\ell} - \log(\eta_\ell) + \log(t_i) \right) \delta_{i\ell} - \sum_{i=1}^n \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right), \\ \frac{\partial^2 l(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t})}{\partial^2 \eta_\ell} &= \sum_{i=1}^n \frac{\beta_\ell \delta_{i\ell}}{\eta_\ell^2} - \sum_{i=1}^n \left(\frac{(\beta_\ell - 1) \beta_\ell t_i^2 \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 2}}{\eta_\ell^4} + \frac{2 \beta_\ell t_i \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 1}}{\eta_\ell^3} \right), \\ \frac{\partial^2 l(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t})}{\partial \eta_\ell \partial \beta_\ell} &= \sum_{i=1}^n -\frac{\delta_{i\ell}}{\eta_\ell} + \sum_{i=1}^n \frac{\left(\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} + \beta_\ell \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right) \right)}{\eta_\ell} \end{aligned}$$

e

$$\frac{\partial^2 l(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t})}{\partial^2 \beta_\ell} = \sum_{i=1}^n -\frac{\delta_{i\ell}}{\beta_\ell^2} - \sum_{i=1}^n \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right)^2.$$

As esperanças para o cálculo da matriz de informação de Fisher são

$$E \left[\frac{\partial^2 l(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t})}{\partial^2 \eta_\ell} \right] = \sum_{i=1}^n \frac{\beta_\ell \delta_{i\ell}}{\eta_\ell^2} - \sum_{i=1}^n \left(\frac{(\beta_\ell - 1) \beta_\ell E \left[t_i^2 \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 2} \right]}{\eta_\ell^4} + \frac{2 \beta_\ell E \left[t_i \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 1} \right]}{\eta_\ell^3} \right), \text{ em que}$$

$$E \left[t_i^2 \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 2} \right] = \eta_\ell^2 \text{ e } E \left[t_i \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell - 1} \right] = \eta_\ell.$$

$$E \left[\frac{\partial^2 l(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\eta})}{\partial \eta_\ell \partial \beta_\ell} \right] = \sum_{i=1}^n -\frac{\delta_{i\ell}}{\eta_\ell} + \sum_{i=1}^n \frac{\left(E \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \right] + \beta_\ell E \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right) \right] \right)}{\eta_\ell},$$

$$\text{em que } E \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \right] = 1 \text{ e } E \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right) \right] = \frac{1 - \gamma}{\beta_\ell}.$$

$$E \left[\frac{\partial^2 l(\boldsymbol{\beta}, \boldsymbol{\eta} | \mathbf{t})}{\partial^2 \beta_\ell} \right] = \sum_{i=1}^n -\frac{\delta_{i\ell}}{\beta_\ell^2} - \sum_{i=1}^n E \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right)^2 \right],$$

$$\text{em que } E \left[\left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \log \left(\frac{t_i}{\eta_\ell} \right)^2 \right] = \frac{(-2 + \gamma)\gamma + \pi^2/6}{\beta_\ell^2}.$$

Assim, a matriz de informação de Fisher relativa à ordenação $(\eta_1, \beta_1, \eta_2, \beta_2)$ é dada por

$$H = n \begin{bmatrix} \frac{\beta_1(1+\beta_1+\bar{\delta}_1)}{\eta_1^2} & \frac{(2+\gamma+\bar{\delta}_1)}{\eta_1} & 0 & 0 \\ \frac{(-2+\gamma+\bar{\delta}_1)}{\eta_1} & \frac{(-2+\gamma)\gamma+\psi'(2)+1+\bar{\delta}_1}{\beta_1^2} & 0 & 0 \\ 0 & 0 & \frac{\beta_2(1+\beta_2+\bar{\delta}_2)}{\eta_2^2} & \frac{(2+\gamma+\bar{\delta}_2)}{\eta_2} \\ 0 & 0 & \frac{(-2+\gamma+\bar{\delta}_2)}{\eta_2} & \frac{(-2+\gamma)\gamma+\psi'(2)+1+\bar{\delta}_2}{\beta_2^2} \end{bmatrix},$$

em que $\bar{\delta}_\ell = \sum_{i=1}^n \frac{\delta_{i\ell}}{n}$, $\ell = 1, 2$, γ é a constante de Euler e $\psi'(2) = \frac{\pi^2}{6} - 1$.

A inversa da matriz de informação é dada por

$$S = \frac{1}{n} \begin{bmatrix} \frac{\eta_1^2 [b_3 + \bar{\delta}_1]}{\beta_1 b_4(\beta_1)} & -\frac{\beta_1 \eta_1 (-2 + \gamma + \bar{\delta}_1)}{b_4(\beta_1)} & 0 & 0 \\ -\frac{\beta_1 \eta_1 (-2 + \gamma + \bar{\delta}_1)}{b_4(\beta_1)} & \frac{\beta_1^2 [1 + \beta_1 + \bar{\delta}_1]}{b_4(\beta_1)} & 0 & 0 \\ 0 & 0 & \frac{\eta_2^2 [b_3 + \bar{\delta}_2]}{\beta_2 b_4(\beta_2)} & -\frac{\beta_2 \eta_2 (-2 + \gamma + \bar{\delta}_2)}{b_4(\beta_2)} \\ 0 & 0 & -\frac{\beta_2 \eta_2 (-2 + \gamma + \bar{\delta}_2)}{b_4(\beta_2)} & \frac{\beta_2^2 [1 + \beta_2 + \bar{\delta}_2]}{b_4(\beta_2)} \end{bmatrix},$$

em que $b_3 = (-2 + \gamma)\gamma + \pi^2/6$, $b_4(\beta_\ell) = [b_5 - b_6 \bar{\delta}_\ell - (1 + \beta_\ell) \bar{\delta}_\ell^2]$, $b_5 = \gamma^2 + \pi^2/6 + 2\gamma(-1 + \beta_\ell) + (-4 + \pi^2)\beta_\ell$ e $b_6 = -1 + \pi^2/6 - 5\beta_\ell + \gamma(-2 + \gamma + 2\beta_\ell)$ para $\ell = 1, 2$.

Para obtermos a função a priori de referência conjunta dos parâmetros η_1, β_1, η_2 e β_2 , consideramos o caso de quatro grupos com a ordenação

$$\pi(\eta_1, \beta_1, \eta_2, \beta_2) = \pi(\beta_2 | \eta_1, \beta_1, \eta_2) \pi(\eta_2 | \eta_1, \beta_1) \pi(\beta_1 | \eta_1) \pi(\eta_1),$$

seguindo o procedimento descrito na Subseção 2.1.3.

Considere H_4 igual a matriz de informação de Fisher H , e seja, $h_{4,4} = \frac{(-2+\gamma)\gamma + \psi'(2) + 1 + \bar{\delta}_2}{\beta_2^2}$ o elemento (4, 4) da matriz H_4 , de modo que $h_{4,4}^{1/2}$ pode ser fatorado da forma

$$h_{4,4}^{1/2} = \frac{1}{\beta_2} \left\{ (-2 + \gamma)\gamma + \psi'(2) + 1 + \bar{\delta}_2 \right\}^{1/2} = f_3(\beta_2)g_3(\eta_1, \beta_1, \eta_2).$$

Então,

$$\pi(\beta_2 | \eta_1, \beta_1, \eta_2) \propto f_3(\beta_2) = \frac{1}{\beta_2}. \quad (3.8)$$

Considere agora S_4 igual à inversa da matriz de informação de Fisher S , e S_3 a submatriz superior 3×3 de $S(\boldsymbol{\beta}, \boldsymbol{\eta})$. Devemos encontrar $h_{3,3}$, o elemento (3, 3) da matriz H_3 , e mostrar que esse se fatora como $h_{3,3}^{1/2} = f_2(\eta_2) g_2(\eta_1, \beta_1, \beta_2)$.

Temos que

$$S_3 = \frac{1}{n} \begin{bmatrix} \frac{\eta_1^2 [b_3 + \bar{\delta}_1]}{\beta_1 b_4(\beta_1)} & -\frac{\beta_1 \eta_1 (-2 + \gamma + \bar{\delta}_1)}{b_4(\beta_1)} & 0 \\ -\frac{\beta_1 \eta_1 (-2 + \gamma + \bar{\delta}_1)}{b_4(\beta_1)} & \frac{\beta_1^2 [1 + \beta_1 + \bar{\delta}_1]}{b_4(\beta_1)} & 0 \\ 0 & 0 & \frac{\eta_2^2 [b_3 + \bar{\delta}_2]}{\beta_2 b_4(\beta_2)} \end{bmatrix}.$$

Invertendo S_3 , temos $h_{3,3} = \frac{\beta_2 b_4(\beta_2)}{n \eta_2^2 [((-2 + \gamma)\gamma + \pi^2/6) + \bar{\delta}_2]}$ de modo que $h_{3,3}^{1/2}$ pode ser fatorado da forma

$$h_{3,3}^{1/2} = \frac{1}{\eta_2} \left\{ \frac{\beta_2 b_4(\beta_2)}{n [((-2 + \gamma)\gamma + \pi^2/6) + \bar{\delta}_2]} \right\}^{1/2} = f_2(\eta_2)g_2(\eta_1, \beta_1, \beta_2).$$

Então,

$$\pi(\eta_2 | \eta_1, \beta_1) \propto f_2(\eta_2) = \frac{1}{\eta_2}. \quad (3.9)$$

O próximo passo é encontrar $h_{2,2}$, o elemento (2, 2) da matriz H_2 , a partir de S_2 , a submatriz superior 2×2 de S , e mostrar que $h_{2,2}^{1/2} = f_1(\beta_1)g_1(\eta_1, \eta_2, \beta_2)$. Para tanto, precisamos primeiramente inverter a matriz S_2 , dada por

$$S_2 = \frac{1}{n} \begin{bmatrix} \frac{\eta_1^2 [b_3 + \bar{\delta}_1]}{\beta_1 b_4(\beta_1)} & -\frac{\beta_1 \eta_1 (-2 + \gamma + \bar{\delta}_1)}{b_4(\beta_1)} \\ -\frac{\beta_1 \eta_1 (-2 + \gamma + \bar{\delta}_1)}{b_4(\beta_1)} & \frac{\beta_1^2 [1 + \beta_1 + \bar{\delta}_1]}{b_4(\beta_1)} \end{bmatrix}.$$

Invertendo S_2 , temos $h_{2,2} = \frac{(-2+\gamma)\gamma + \psi'(2) + 1 + \bar{\delta}_1}{\beta_1^2}$ de modo que $h_{2,2}^{1/2}$ pode ser fatorado da forma

$$h_{2,2}^{1/2} = \frac{1}{\beta_1} \left\{ ((-2 + \gamma)\gamma + \psi'(2) + 1 + \bar{\delta}_1) \right\}^{1/2} = f_1(\beta_1)g_1(\eta_1, \eta_2, \beta_2).$$

Então,

$$\pi(\beta_1|\eta_1) \propto f_1(\beta_1) = \frac{1}{\beta_1}. \quad (3.10)$$

Finalmente, basta encontrarmos o elemento $s_{1,1}$ de S e mostrar que $s_{1,1}^{-1/2} = f_0(\eta_1)g_0(\beta_1, \eta_2, \beta_2)$. Para isso, precisamos inverter $s_{1,1} = \frac{\eta_1^2 [b_3 + \bar{\delta}_1]}{\beta_1 b_4(\beta_1)}$. Assim, temos que $s_{1,1}^{-1/2}$ pode ser fatorado da forma

$$s_{1,1}^{-1/2} = \frac{1}{\eta_1} \left\{ \frac{\beta_1 b_4(\beta_1)}{(-2 + \gamma)\gamma + \pi^2/6 + \bar{\delta}_1} \right\}^{1/2} = f_0(\eta_1)g_0(\beta_1, \eta_2, \beta_2).$$

Então,

$$\pi(\eta_1) \propto f_0(\eta_1) = \frac{1}{\eta_1}. \quad (3.11)$$

Portanto, a função a priori de referência relativa à ordenação $(\eta_1, \beta_1, \eta_2, \beta_2)$ é dada por

$$\pi(\eta_1, \beta_1, \eta_2, \beta_2) \propto \frac{1}{\eta_1 \beta_1 \eta_2 \beta_2}, \quad (3.12)$$

em que $\eta_1, \beta_1, \eta_2, \beta_2 \in \mathbb{R}^+$. Neste caso, vale ressaltar que a função a priori de referência coincide com a função a priori de Jeffreys.

Para inferir sobre os parâmetros do modelo Weibull, devemos encontrar a distribuição a posteriori dos parâmetros. Assim, levando em conta (3.6) e (3.12), a distribuição a posteriori conjunta de η_1, β_1, η_2 e β_2 é dada por

$$\pi(\boldsymbol{\eta}, \boldsymbol{\beta}|\mathbf{t}) \propto \prod_{i=1}^n \prod_{\ell=1}^2 \left(\left(\frac{\beta_\ell}{\eta_\ell} t_i^{\beta_\ell - 1} \right)^{\delta_{i\ell}} \exp \left[- \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \right] \frac{1}{\eta_\ell \beta_\ell} \right), \quad (3.13)$$

As distribuições a posteriori condicionais de β_ℓ e η_ℓ são dadas, respectivamente, por

$$\pi(\eta_\ell|\mathbf{t}, \boldsymbol{\beta}) \propto \prod_{i=1}^n \left(\left(\frac{1}{\eta_\ell} \right)^{\delta_{i\ell}} \exp \left[- \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \right] \frac{1}{\eta_\ell} \right) \quad (3.14)$$

e

$$\pi(\beta_\ell|\mathbf{t}, \boldsymbol{\eta}) \propto \prod_{i=1}^n \left(\left(\frac{\beta_\ell}{\eta_\ell} t_i^{\beta_\ell - 1} \right)^{\delta_{i\ell}} \exp \left[- \left(\frac{t_i}{\eta_\ell} \right)^{\beta_\ell} \right] \frac{1}{\beta_\ell} \right). \quad (3.15)$$

Definida a distribuição a posteriori, o próximo passo é a obtenção das medidas resumo para os parâmetros de interesse. No Capítulo 4, exemplificamos a metodologia desenvolvida até agora, através de dois estudos de simulação para o modelo em questão, e apresentamos uma aplicação considerando os dados do Exemplo 3.1.

Na próxima seção, descrevemos o critério de referência Bayesiano (Bernardo e Rueda, 2002), utilizado para testar a igualdade dos parâmetros das distribuições dos tempos de vida dos dois fatores de risco independentes.

3.5 Critério de referência Bayesiano para igualdade dos parâmetros

Considere dois fatores de risco, fator 1 e fator 2, ($\ell = 1, 2$), n tempos t_i ($i = 1, \dots, n$) observados com os respectivos rótulos $\delta_{i\ell}$. Seja

$$\mathcal{M} \equiv \left\{ p_1(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\eta}) = \prod_{\ell=1}^2 \left(\frac{\beta_\ell}{\eta_\ell^{\beta_\ell}} t^{\beta_\ell-1} \right)^{\delta_\ell} \exp \left[- \left(\frac{t}{\eta_\ell} \right)^{\beta_\ell} \right], \beta_\ell, \eta_\ell > 0, \delta_2 = 1 - \delta_1 \right\},$$

o modelo estatístico que gerou os dados $\mathbf{t} \in \mathbb{R}^+$ e considere a hipótese nula a ser testada $H_0 \equiv \{\beta_1 = \beta_2, \eta_1 = \eta_2\}$. Para decidirmos se aceitamos ou não a hipótese nula, utilizamos o BRC descrito no Capítulo 2. Para a implementação do mesmo devemos:

1. Calcular a discrepância intrínseca ($\delta \{\mathcal{M}, \mathcal{M}_0\}$) entre os modelos \mathcal{M} e

$$\mathcal{M}_0 \equiv \left\{ p_2(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\eta}) = \frac{\beta_1}{\eta_1^{\beta_1}} t^{\beta_1-1} \exp \left[-2 \left(\frac{t}{\eta_1} \right)^{\beta_1} \right], \beta_1, \eta_1 > 0 \right\}.$$

A discrepância intrínseca $\delta \{\mathcal{M}, \mathcal{M}_0\}$ é dada por

$$\delta \{\mathcal{M}, \mathcal{M}_0\} = \min \{ \kappa(p_2|p_1), \kappa(p_1|p_2) \}.$$

Assim,

$$\kappa(p_j|p_i) = -\frac{1}{2} + 2^{-\frac{\beta_i}{\beta_j}} \left(\frac{\eta_j}{\eta_i} \right)^{\beta_i} \Gamma \left[\frac{\beta_i + \beta_j}{\beta_j} \right] + \frac{(\beta_i - \beta_j)(\gamma + \log(2) - \beta_j \log(\eta_j))}{2\beta_j}$$

$$+\frac{1}{2}(-\log(\beta_i) + \log(\beta_j) + \beta_i \log(\eta_i) - \beta_j \log(\eta_j)), \quad i, j = 1, 2. \quad (3.16)$$

e $\delta \{\mathcal{M}, \mathcal{M}_0\}$ é encontrada numericamente.

2. Derivar a correspondente esperança a posteriori de referência

$$d(\mathcal{M}_0, \mathbf{t}) = E[\delta(\mathcal{M}_0, \mathcal{M})|\mathbf{t}] = \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \delta(\mathcal{M}_0, \mathcal{M}) \pi(\boldsymbol{\beta}, \boldsymbol{\eta}|\mathbf{t}) d\boldsymbol{\eta} d\boldsymbol{\beta},$$

e estipular esse número como uma medida de evidência contra a hipótese nula H_0 .

Porém, esse cálculo não pode ser efetuado analiticamente, por isso, utilizamos os valores dos parâmetros gerados pelo método de Metropolis-Hastings para obtermos a estatística intrínseca $d(\mathcal{M}_0, \mathbf{t})$.

Na próxima seção, desenvolvemos a análise de referência para o modelo de riscos competitivos, considerando a confiabilidade como um parâmetro do mesmo.

3.6 Análise de referência para a confiabilidade em modelos de riscos competitivos

Considerando que o interesse é a confiabilidade dos componentes de um sistema em série, nesta seção, construímos a priori de referência diretamente para as confiabilidades, considerando a densidade em Y dada em (2.24) e a metodologia descrita na Subseção 2.1.3.

Para a obtenção da função a priori de referência consideramos a seguinte reparametrização

$$W_\ell = \beta_\ell \text{ e } R_\ell = \exp \left[- \left(\frac{t_\ell}{\eta_\ell} \right)^{\beta_\ell} \right], \quad \ell = 1, 2,$$

em que t_ℓ é fixo. Observe que R_ℓ é função de t_ℓ , no entanto, utilizamos esta notação para simplificação.

Então, considere que cada fator de risco tenha a distribuição dada em (2.24) e suponha que Y_{i1} e Y_{i2} sejam variáveis aleatórias independentes, contínuas

e positivas denominadas tempos de vida para uma determinada unidade i ($i = 1, \dots, n$) exposta aos dois fatores de risco competitivos. Com base no cenário de riscos competitivos apresentado na Seção 3.1, temos a variável aleatória $T_i = \min \{Y_{i1}, Y_{i2}\}$ e que a variável indicadora $\delta_{i\ell}$ é dada por

$$\delta_{i\ell} = \begin{cases} 1, & \text{se a unidade } i \text{ falhou por } \ell, \text{ ou seja, } T_i = Y_{i\ell} \\ 0, & \text{se a unidade } i \text{ não falhou por } \ell, \text{ ou seja, } T_i \neq Y_{i\ell} \end{cases}.$$

O parâmetro de interesse escolhido para a construção da função a priori foi R_1 e os demais, R_2 , W_1 e W_2 , foram considerados como parâmetros perturbadores.

A função de verossimilhança para n dados de riscos competitivos considerando o modelo Weibull, $\mathbf{R} = (R_1, R_2)$ e $\mathbf{W} = (W_1, W_2)$, é dada por

$$L(\mathbf{R}, \mathbf{W}|\mathbf{t}) = \prod_{i=1}^n \prod_{\ell=1}^2 \left\{ \left(W_\ell \log \left(\frac{1}{R_\ell} \right) t_i^{W_\ell - 1} \right)^{\delta_{i\ell}} \exp \left[-t_i^{W_\ell} \log \left(\frac{1}{R_\ell} \right) \right] \right\}. \quad (3.17)$$

O logaritmo da função de verossimilhança (3.17) é

$$l(\mathbf{R}, \mathbf{W}|\mathbf{t}) = \sum_{i=1}^n \sum_{\ell=1}^2 \left\{ \delta_{i\ell} \left[\log(W_\ell) + \log \log \left(\frac{1}{R_\ell} \right) + (W_\ell - 1) \log(t_i) \right] - t_i^{W_\ell} \log \left(\frac{1}{R_\ell} \right) \right\}. \quad (3.18)$$

As derivadas de primeira e segunda ordem da equação (3.19) para $\ell = 1, 2$, obtidas através do *software* Mathematica, são dadas por

$$\begin{aligned} \frac{\partial l(\mathbf{R}, \mathbf{W}|\mathbf{t})}{\partial R_\ell} &= -\frac{\sum_{i=1}^n \delta_{i\ell}}{R_\ell \log \left(\frac{1}{R_\ell} \right)} + \frac{\sum_{i=1}^n t_i^{W_\ell}}{R_\ell}, \\ \frac{\partial l(\mathbf{R}, \mathbf{W}|\mathbf{t})}{\partial W_\ell} &= \sum_{i=1}^n \left(\frac{1}{W_\ell} + \log(t_i) \right) \delta_{i\ell} - \log \left(\frac{1}{R_\ell} \right) \sum_{i=1}^n t_i^{W_\ell} \log(t_i), \\ \frac{\partial^2 l(\mathbf{R}, \mathbf{W}|\mathbf{t})}{\partial R_\ell^2} &= \sum_{i=1}^n \left(-\frac{\delta_{i\ell}}{R_\ell^2 \log \left(\frac{1}{R_\ell} \right)^2} + \frac{\delta_{i\ell}}{R_\ell^2 \log \left(\frac{1}{R_\ell} \right)} \right) - \frac{\sum_{i=1}^n t_i^{W_\ell}}{R_\ell^2}, \\ \frac{\partial^2 l(\mathbf{R}, \mathbf{W}|\mathbf{t})}{\partial R_\ell \partial W_\ell} &= \frac{\sum_{i=1}^n \log(t_i) t_i^{W_\ell}}{R_\ell} \end{aligned}$$

e

$$\frac{\partial^2 l(\mathbf{R}, \mathbf{W} | \mathbf{t})}{\partial W_\ell^2} = -\frac{\sum_{i=1}^n \delta_{i\ell}}{W_\ell} - \log\left(\frac{1}{R_\ell}\right) \sum_{i=1}^n \log(t_i)^2 t_i^{W_\ell}$$

As esperanças para o cálculo da matriz de informação de Fisher são

$$E\left[\frac{\partial^2 l(\mathbf{R}, \mathbf{W} | \mathbf{t})}{\partial R_\ell^2}\right] = \sum_{i=1}^n \left(-\frac{\delta_{i\ell}}{R_\ell^2 \log\left(\frac{1}{R_\ell}\right)^2} + \frac{\delta_{i\ell}}{R_\ell^2 \log\left(\frac{1}{R_\ell}\right)} \right) - \frac{\sum_{i=1}^n E\left[t_i^{W_\ell}\right]}{R_\ell^2},$$

$$\text{em que } E\left[t_i^{W_\ell}\right] = \frac{1}{\log\left(\frac{1}{R_\ell}\right)}.$$

$$E\left[\frac{\partial^2 l(\mathbf{R}, \mathbf{W} | \mathbf{t})}{\partial R_\ell \partial W_\ell}\right] = \frac{\sum_{i=1}^n E\left[\log(t_i) t_i^{W_\ell}\right]}{R_\ell}, \text{ em que}$$

$$E\left[\log(t_i) t_i^{W_\ell}\right] = \frac{(-1 + \gamma) + \log \log\left(\frac{1}{R_\ell}\right)}{W_\ell \log(R_\ell)}.$$

$$E\left[\frac{\partial^2 l(\mathbf{R}, \mathbf{W} | \mathbf{t})}{\partial W_\ell^2}\right] = -\frac{\sum_{i=1}^n \delta_{i\ell}}{W_\ell} - \log\left(\frac{1}{R_\ell}\right) \sum_{i=1}^n E\left[\log(t_i)^2 t_i^{W_\ell}\right], \text{ em que}$$

$$E\left[\log(t_i)^2 t_i^{W_\ell}\right] = \frac{(-2 + \gamma)\gamma + \pi^2/6 + \log \log\left(\frac{1}{R_\ell}\right) \left(-2 + 2\gamma + \log \log\left(\frac{1}{R_\ell}\right)\right)}{W_\ell^2 \log\left(\frac{1}{R_\ell}\right)}.$$

Assim, a matriz de informação de Fisher é dada por

$$H(\mathbf{R}, \mathbf{W}) = n \begin{bmatrix} \frac{\bar{\delta}_1 + (1 - \bar{\delta}_1) \log\left(\frac{1}{R_1}\right)}{R_1^2 \log\left(\frac{1}{R_1}\right)^2} & 0 & \frac{q_1(R_1)}{R_1 W_1 \log\left(\frac{1}{R_1}\right)} & 0 \\ 0 & \frac{\bar{\delta}_2 + (1 - \bar{\delta}_2) \log\left(\frac{1}{R_2}\right)}{R_2^2 \log\left(\frac{1}{R_2}\right)^2} & 0 & \frac{q_1(R_2)}{R_2 W_2 \log\left(\frac{1}{R_2}\right)} \\ \frac{q_1(R_1)}{R_1 W_1 \log\left(\frac{1}{R_1}\right)} & 0 & \frac{q_2(R_1)}{W_1^2} & 0 \\ 0 & \frac{q_1(R_2)}{R_2 W_2 \log\left(\frac{1}{R_2}\right)} & 0 & \frac{q_2(R_2)}{W_2^2} \end{bmatrix},$$

em que $q_1(R_\ell) = -\psi(2) + \log \log\left(\frac{1}{R_\ell}\right)$, $q_2(R_\ell) = \bar{\delta}_\ell + (-2 + \gamma)\gamma + \frac{\pi^2}{6} + \log \log\left(\frac{1}{R_\ell}\right) \left(-2\psi(2) + \log \log\left(\frac{1}{R_\ell}\right)\right)$, $\bar{\delta}_\ell = \sum_{i=1}^n \frac{\delta_{i\ell}}{n}$, $\ell = 1, 2$, γ e $\psi(2)$ dados anteriormente.

Com inversa dada por

$$S(\mathbf{R}, \mathbf{W}) = \begin{bmatrix} \frac{R_1^2 \log\left(\frac{1}{R_1}\right)^2 q_1(R_1)}{q_3(R_1)} & 0 & \frac{W_1 q_4(R_1)}{q_3(R_1)} & 0 \\ 0 & \frac{R_2^2 \log\left(\frac{1}{R_2}\right)^2 q_1(R_2)}{q_3(R_2)} & 0 & \frac{W_2 q_4(R_2)}{q_3(R_2)} \\ \frac{W_1 q_4(R_1)}{q_3(R_1)} & 0 & \frac{W_1 q_5(R_1)}{q_3(R_1)} & 0 \\ 0 & \frac{W_2 q_4(R_2)}{q_3(R_2)} & 0 & \frac{W_2 q_5(R_2)}{q_3(R_2)} \end{bmatrix},$$

sendo $q_3(R_\ell) = (-1 + \bar{\delta}_\ell) \left[\bar{\delta}_\ell + (-1 + \gamma)^2 + \log \log \left(\frac{1}{R_\ell} \right) \left(-2\psi(2) + \log \log \left(\frac{1}{R_\ell} \right) \right) - \log \left(\frac{1}{R_\ell} \right) (\bar{\delta}_\ell + q_2(R_\ell)) \right] + \bar{\delta}_\ell \frac{\pi^2}{6}$, $q_4(R_\ell) = R_\ell \log \left(\frac{1}{R_\ell} \right) q_1(R_\ell)$ e $q_5(R_\ell) = -\bar{\delta}_\ell + (\bar{\delta}_\ell - 1) \log \left(\frac{1}{R_\ell} \right)$

Para obtermos a função a priori de referência conjunta dos parâmetros \mathbf{R} e \mathbf{W} , consideramos o caso de quatro grupos, com a ordenação

$$\pi(R_1, R_2, W_1, W_2) = \pi(W_2 | R_1, R_2, W_1) \pi(W_1 | R_1, R_2) \pi(R_2 | R_1) \pi(R_1),$$

seguindo o procedimento descrito na Subseção 2.1.3.

Considere H_4 igual à matriz de informação de Fisher $H(\mathbf{R}, \mathbf{W})$ e seja $h_{4,4} = \frac{q_2(R_2)}{W_2^2}$ o elemento (4, 4) da matriz H_4 , de modo que $h_{4,4}^{1/2}$ pode ser fatorado da forma

$$h_{4,4}^{1/2} = \frac{1}{W_2} [q_2(R_2)]^{1/2} = f_3(W_2) g_3(R_1, R_2, W_1).$$

Então,

$$\pi(W_2 | R_1, R_2, W_1) \propto f_3(W_2) = \frac{1}{W_2}. \quad (3.19)$$

Considere agora S_4 igual à inversa da matriz de informação de Fisher $S(\mathbf{R}, \mathbf{W})$, e S_3 a submatriz superior 3×3 de $S(\mathbf{R}, \mathbf{W})$. Devemos encontrar $h_{3,3}$, o elemento (3, 3) da matriz H_3 , e mostrar que esse se fatora como $h_{3,3}^{1/2} = f_2(W_1)$

$g_2(R_1, R_2, W_2)$. Temos que

$$S_3 = \begin{bmatrix} \frac{R_1^2 \log\left(\frac{1}{R_1}\right)^2 q_1(R_1)}{q_3(R_1)} & 0 & \frac{W_1 q_4(R_1)}{q_3(R_1)} \\ 0 & \frac{R_2^2 \log\left(\frac{1}{R_2}\right)^2 q_1(R_2)}{q_3(R_2)} & 0 \\ \frac{W_1 q_4(R_1)}{q_3(R_1)} & 0 & \frac{W_1 q_5(R_1)}{q_3(R_1)} \end{bmatrix}$$

Invertendo S_3 , temos $h_{3,3} = \frac{q_2(R_1)}{W_1^2}$ de modo que $h_{3,3}^{1/2}$ pode ser fatorado da forma

$$h_{3,3}^{1/2} = \frac{1}{W_1} [q_2(R_1)]^{1/2} = f_2(W_1)g_2(R_1, R_2, W_2).$$

Então,

$$\pi(W_1|R_1, R_2) \propto f_2(W_1) = \frac{1}{W_1}. \quad (3.20)$$

O próximo passo é encontrar $h_{2,2}$, o elemento (2, 2) da matriz H_2 , a partir de S_2 , a submatriz superior 2×2 de $S(\mathbf{R}, \mathbf{W})$, e mostrar que $h_{2,2}^{1/2} = f_1(R_2)g_1(R_1, W_1, W_2)$. Para tanto, precisamos primeiramente inverter a matriz S_2 dada por

$$S_2 = \frac{1}{n} \begin{bmatrix} \frac{R_1^2 \log\left(\frac{1}{R_1}\right)^2 q_1(R_1)}{q_3(R_1)} & 0 \\ 0 & \frac{R_2^2 \log\left(\frac{1}{R_2}\right)^2 q_1(R_2)}{q_3(R_2)} \end{bmatrix}.$$

Invertendo S_2 , temos $h_{2,2} = \frac{(-2+\gamma)\gamma+\psi'(2)+1+\bar{\delta}_1}{\beta_1^2}$ de modo que $h_{2,2}^{1/2}$ pode ser fatorado da forma

$$\begin{aligned} h_{2,2}^{1/2} &= \frac{1}{R_2 \log\left(\frac{1}{R_2}\right)} \left[(-1 + \bar{\delta}_2) \left(1 + \log\left(\frac{1}{R_2}\right) \right) + \frac{-1 + \bar{\delta}_2 + \pi^2/6}{q_2(R_2)} \right]^{1/2} \\ &= f_1(R_2)g_1(R_1, W_1, W_2). \end{aligned}$$

Então,

$$\begin{aligned} \pi(R_2|R_1) &\propto f_1(R_2) \\ &= \frac{1}{R_2 \log\left(\frac{1}{R_2}\right)} \left[(-1 + \bar{\delta}_2) \left(1 + \log\left(\frac{1}{R_2}\right) \right) + \frac{-1 + \bar{\delta}_2 + \pi^2/6}{q_2(R_2)} \right]^{1/2}. \quad (3.21) \end{aligned}$$

Finalmente, basta encontrarmos o elemento $s_{1,1}$ de $S(\mathbf{R}, \mathbf{W})$ e mostrar que $s_{1,1}^{-1/2} = f_0(R_1)g_0(R_2, W_1, W_2)$. Para isso, precisamos inverter

$$s_{1,1} = \frac{R_1^2 \log\left(\frac{1}{R_1}\right)^2 q_1(R_1)}{q_3(R_1)}.$$

Assim, temos que $s_{1,1}^{-1/2}$ pode ser fatorado da forma

$$\begin{aligned} s_{1,1}^{-1/2} &= \frac{1}{R_1 \log\left(\frac{1}{R_1}\right)} \left[(-1 + \bar{\delta}_1) \left(1 + \log\left(\frac{1}{R_1}\right) \right) + \frac{-1 + \bar{\delta}_1 + \pi^2/6}{q_2(R_1)} \right]^{1/2} \\ &= f_0(R_1)g_0(R_2, W_1, W_2). \end{aligned}$$

Então,

$$\begin{aligned} \pi(R_1) &\propto f_0(R_1) \\ &= \frac{1}{R_1 \log\left(\frac{1}{R_1}\right)} \left[(-1 + \bar{\delta}_1) \left(1 + \log\left(\frac{1}{R_1}\right) \right) + \frac{-1 + \bar{\delta}_1 + \pi^2/6}{q_2(R_1)} \right]^{1/2}. \end{aligned} \quad (3.22)$$

Portanto, a função a priori de referência relativa à ordenação (R_1, R_2, W_1, W_2) é dada por

$$\begin{aligned} \pi(R_1, R_2, W_1, W_2) &\propto \prod_{\ell=1}^2 \left\{ \frac{1}{R_\ell \log\left(\frac{1}{R_\ell}\right)} \left[(-1 + \bar{\delta}_\ell) \left(1 + \log\left(\frac{1}{R_\ell}\right) \right) \right. \right. \\ &\quad \left. \left. + \frac{-1 + \bar{\delta}_\ell + \pi^2/6}{q_2(R_\ell)} \right]^{1/2} \frac{1}{W_\ell} \right\}, \end{aligned} \quad (3.23)$$

em que $W_\ell \in \mathbb{R}^+$ e $0 < R_\ell < 1$.

Para inferir sobre os parâmetros do modelo Weibull devemos encontrar a distribuição a posteriori dos mesmos. Assim, levando em conta (3.17) e (3.23), a distribuição a posteriori conjunta de \mathbf{R} e \mathbf{W} é dada por

$$\begin{aligned} \pi(\mathbf{R}, \mathbf{W} | \mathbf{t}) &= \prod_{i=1}^n \prod_{\ell=1}^2 \left\{ \left(W_\ell \log\left(\frac{1}{R_\ell}\right) t_i^{W_\ell-1} \right)^{\delta_{i\ell}} \exp \left[-t_i^{W_\ell} \log\left(\frac{1}{R_\ell}\right) \right] \right. \\ &\quad \left. \otimes \frac{1}{R_\ell \log\left(\frac{1}{R_\ell}\right)} \left[(-1 + \bar{\delta}_\ell) \left(1 + \log\left(\frac{1}{R_\ell}\right) \right) + \frac{-1 + \bar{\delta}_\ell + \pi^2/6}{q_2(R_\ell)} \right]^{1/2} \frac{1}{W_\ell} \right\} \end{aligned} \quad (3.24)$$

As distribuições a posteriori condicionais de R_ℓ e W_ℓ são dadas, respectivamente, por

$$\pi(R_\ell | \mathbf{t}, W_\ell) = \prod_{i=1}^n \prod_{\ell=1}^2 \left\{ \left(\log\left(\frac{1}{R_\ell}\right) t_i^{W_\ell-1} \right)^{\delta_{i\ell}} \exp \left[-t_i^{W_\ell} \log\left(\frac{1}{R_\ell}\right) \right] \right\}$$

$$\otimes \frac{1}{R_\ell \log\left(\frac{1}{R_\ell}\right)} \left\{ \left[(-1 + \bar{\delta}_\ell) \left(1 + \log\left(\frac{1}{R_\ell}\right) \right) + \frac{-1 + \bar{\delta}_\ell + \pi^2/6}{q_2(R_\ell)} \right]^{1/2} \right\} \quad (3.25)$$

e

$$\pi(W_\ell | \mathbf{t}, R_\ell) = \prod_{i=1}^n \prod_{\ell=1}^2 \left\{ \left(W_\ell t_i^{W_\ell - 1} \right)^{\delta_{i\ell}} \exp \left[-t_i^{W_\ell} \log\left(\frac{1}{R_\ell}\right) \right] \frac{1}{W_\ell} \right\}. \quad (3.26)$$

Definida a distribuição a posteriori, o próximo passo é a obtenção das medidas resumo para os parâmetros de interesse. Considerando essa abordagem, é possível efetuar o cálculo das bandas de credibilidade para as funções de confiabilidade dos componentes de um sistema em série.

Capítulo 4

Aplicações

Neste capítulo, aplicamos a teoria descrita no Capítulo 3 através de três estudos de simulação e de um exemplo real. O estudo de simulação 1 foi desenvolvido com base em dados simulados segundo o modelo Weibull (3.3) e o estudo 2, com base em dados simulados considerando distribuição Gama para os fatores de risco a fim de verificarmos a adequabilidade do método inferencial proposto. O exemplo real foi feito com os dados do Exemplo 3.1 (Coque, 2004) e, por fim, exemplificamos a metodologia apresentada na Seção 3.6, discutindo as vantagens e desvantagens da mesma.

Todos os cálculos e estimativas apresentadas neste trabalho foram feitos utilizando o *software* estatístico *R*. Os dados simulados se encontram no Apêndice A.

4.1 Estudo de simulação 1

Geramos quatro conjuntos de dados com tamanhos amostrais (n) iguais a 10, 25, 50 e 100, segundo o modelo Weibull (3.3), para fins de exemplificação. Todos os conjuntos foram gerados tomando $\beta_1 = 2$, $\eta_1 = 200$, $\beta_2 = 1$ e $\eta_2 = 400$. Obtivemos duas amostras das distribuições a posteriori marginais de β_1, η_1, β_2 e η_2 para cada n , utilizando o algoritmo de Metropolis-Hastings. As medidas resumo das amostras da cadeia 1, adotando aquecimento (*burn in*) de 30.000

com tamanho inicial de 65.000 e salto igual a 5 para $n = 10$, aquecimento (*burn in*) de 25.000 com tamanho inicial de 65.000 e salto igual a 8 para $n = 100$, e para os demais tamanhos amostrais, aquecimento (*burn in*) de 15.000 e salto igual a 10, encontram-se nas Tabelas 4.1 e 4.2. Aplicando o teste de Gelman e Rubin, para verificação da convergência, obtivemos $\hat{R} = 1$ para todos os casos, o que traz grandes indícios de convergência para as distribuições a posteriori dos parâmetros. O tempo de processamento para cada cadeia simulada, em segundos, é igual a 209,59.

TABELA 4.1: Medidas resumo a posteriori para os parâmetros relacionados ao fator de risco 1.

n	β_1			η_1		
	Média	DP	Int. Cred. de 95%	Média	DP	Int. Cred. de 95%
10	2,01	0,082	[1,85 ; 2,17]	198,4	7,94	[183,38 ; 213,90]
25	2,02	0,080	[1,86 ; 2,18]	200,1	7,57	[185,44 ; 215,21]
50	2,05	0,080	[1,90 ; 2,21]	199,5	7,29	[185,38 ; 214,02]
100	2,06	0,077	[1,91 ; 2,21]	199,7	6,72	[186,97 ; 213,22]

TABELA 4.2: Medidas resumo a posteriori para os parâmetros relacionados ao fator de risco 2.

n	β_2			η_2		
	Média	DP	Int. Cred. de 95%	Média	DP	Int. Cred. de 95%
10	0,99	0,041	[0,92 ; 1,08]	397,2	7,22	[383,20 ; 411,27]
25	1,00	0,041	[0,92 ; 1,08]	397,5	7,22	[383,70 ; 411,75]
50	1,01	0,039	[0,91 ; 1,06]	397,2	7,19	[383,55 ; 411,53]
100	0,98	0,038	[0,90 ; 1,05]	397,0	7,16	[383,07 ; 411,09]

Analisando as Tabelas 4.1 e 4.2, observamos que as estimativas estão próximas do verdadeiro valor de cada parâmetro e os intervalos de credibilidade contêm o verdadeiro valor de cada parâmetro. À medida que o tamanho da amostra aumenta, a tendência do desvio padrão é diminuir, o que já era esperado.

As figuras apresentadas neste exemplo dizem respeito à amostra de tamanho 10, e para os demais casos o resultado é análogo. A Figura 4.1 mostra que

há indícios de convergência para ambos os parâmetros, pois os valores gerados variam em torno de uma média. Isso também pode ser observado na Figura 4.3, uma vez que as densidades estimadas dos parâmetros de interesse para as duas cadeias geradas estão próximas. Note que, para o fator de risco 2, as densidades estimadas do parâmetro η coincidem. Na Figura 4.2, verificamos que as médias das cadeias convergem para um determinado valor, indicando, após aproximadamente 30.000 iterações, a convergência das cadeias.

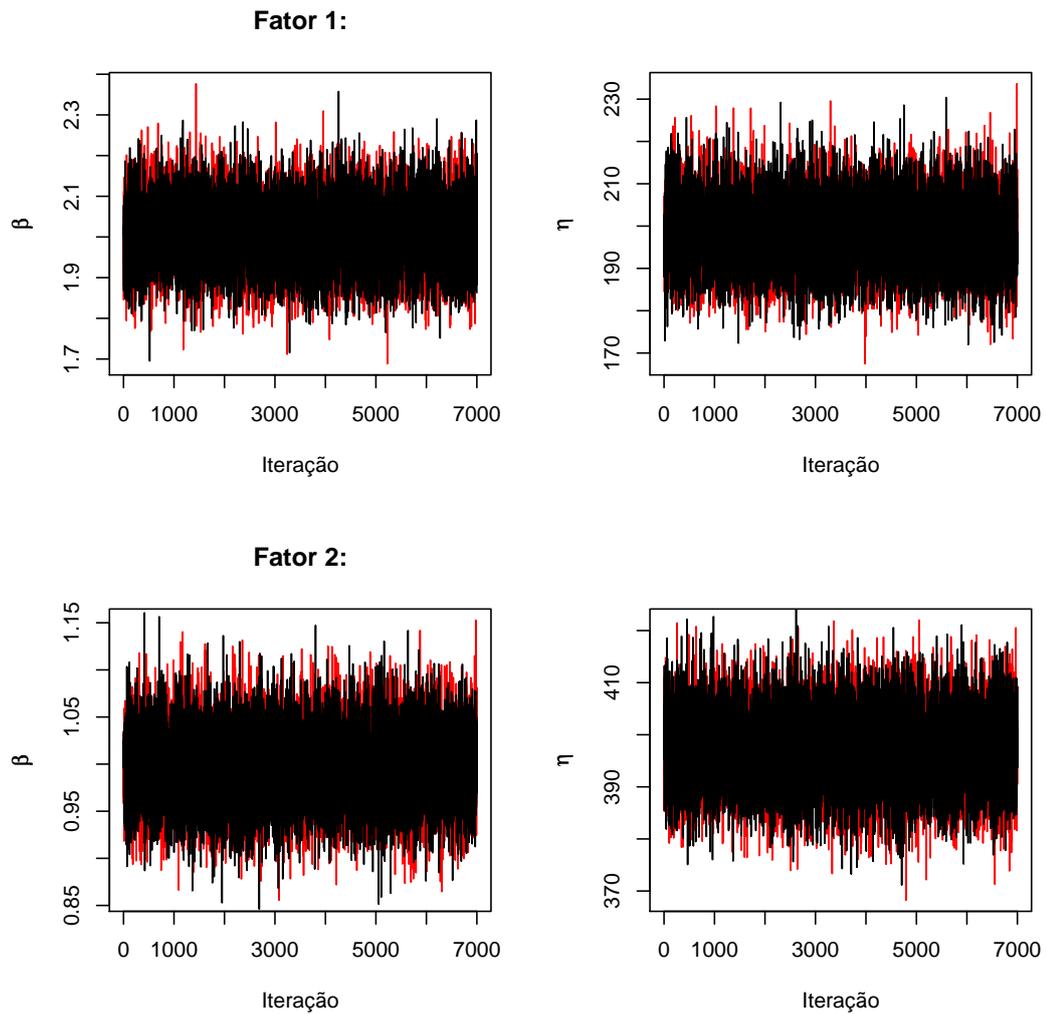


FIGURA 4.1: Traço das cadeias dos parâmetros dos fatores de risco 1 e 2 ($n = 10$).

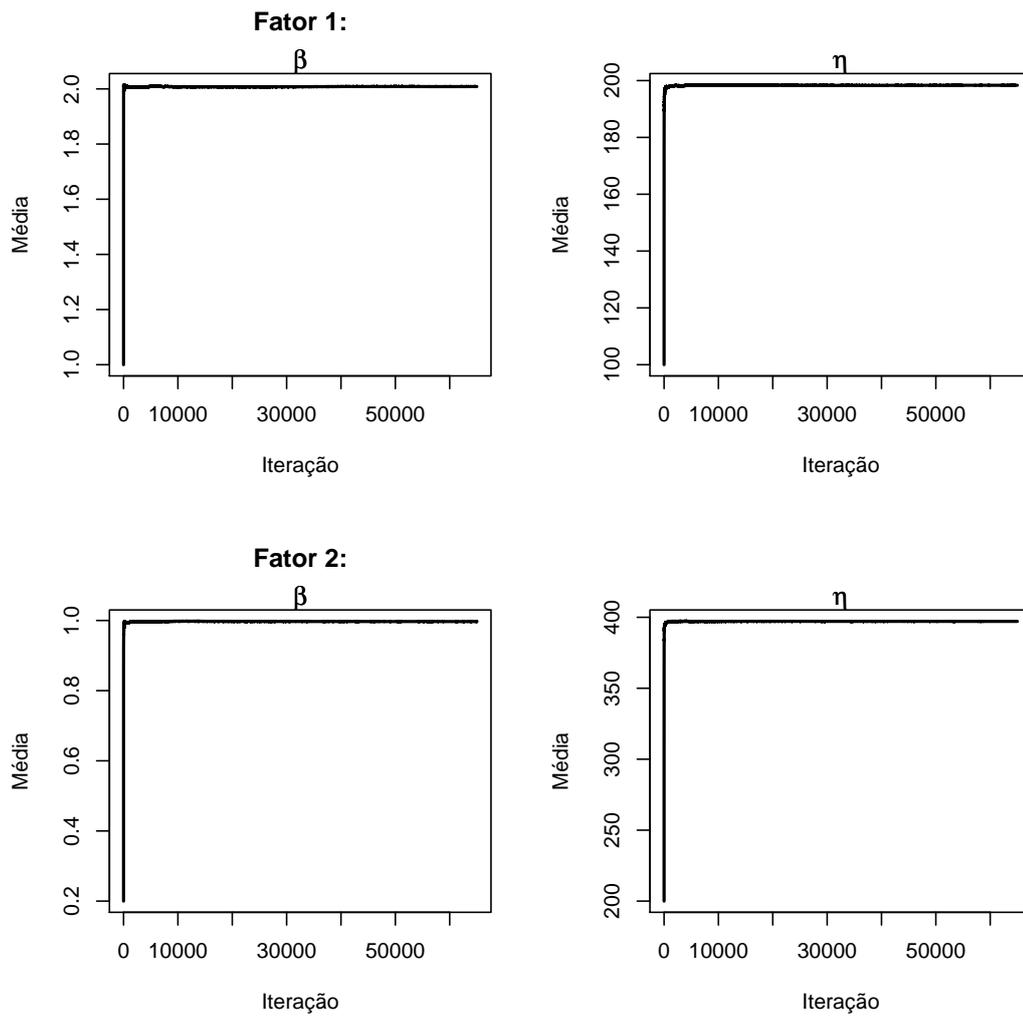


FIGURA 4.2: Médias ergódicas para os parâmetros dos fatores de risco 1 e 2 ($n = 10$).

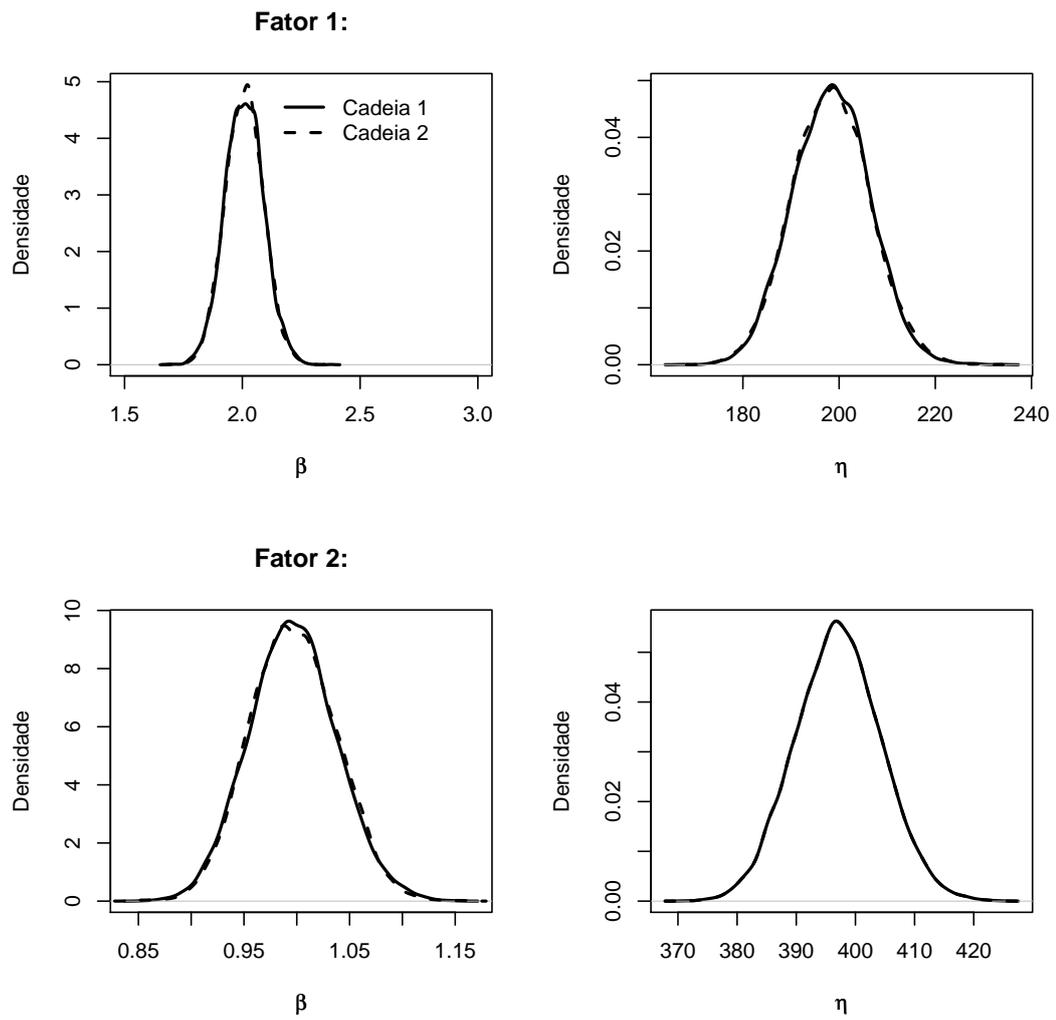


FIGURA 4.3: Funções densidade a posteriori estimadas dos parâmetros dos fatores de risco 1 e 2 ($n = 10$).

Para traçar os ajustes das curvas de confiabilidade foram utilizadas as estimativas encontradas nas Tabelas 4.1 e 4.2 para $n = 10$. Observando as Figuras 4.4, 4.5 e 4.6 verificamos que o ajuste do modelo Weibull é adequado aos dados em questão, indicando o esperado.

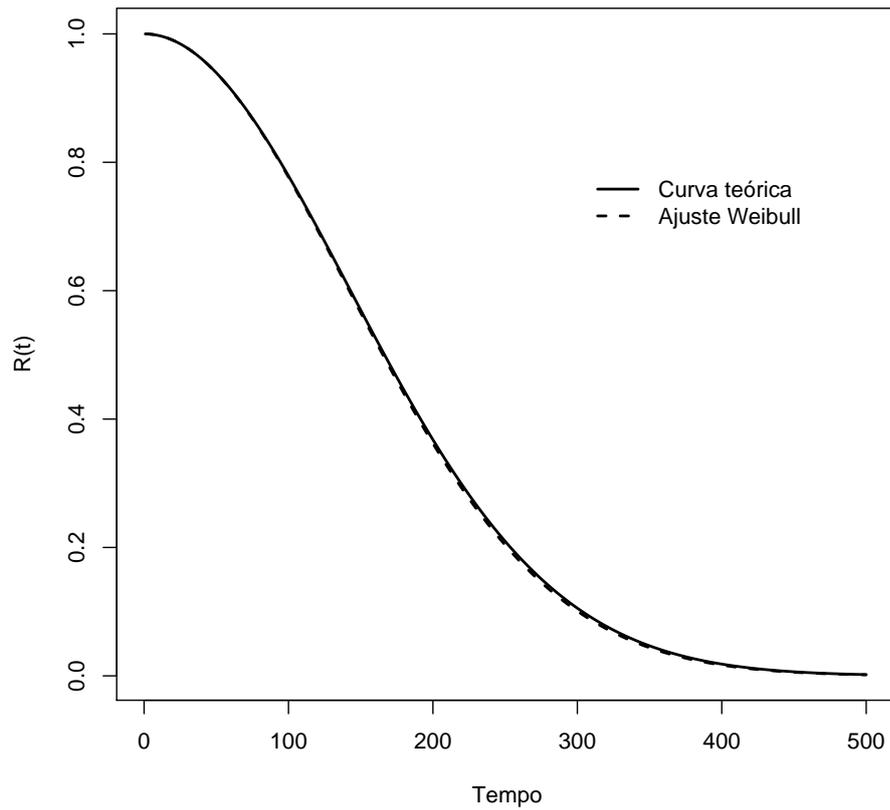


FIGURA 4.4: Ajuste da curva de confiabilidade do fator de risco 1.

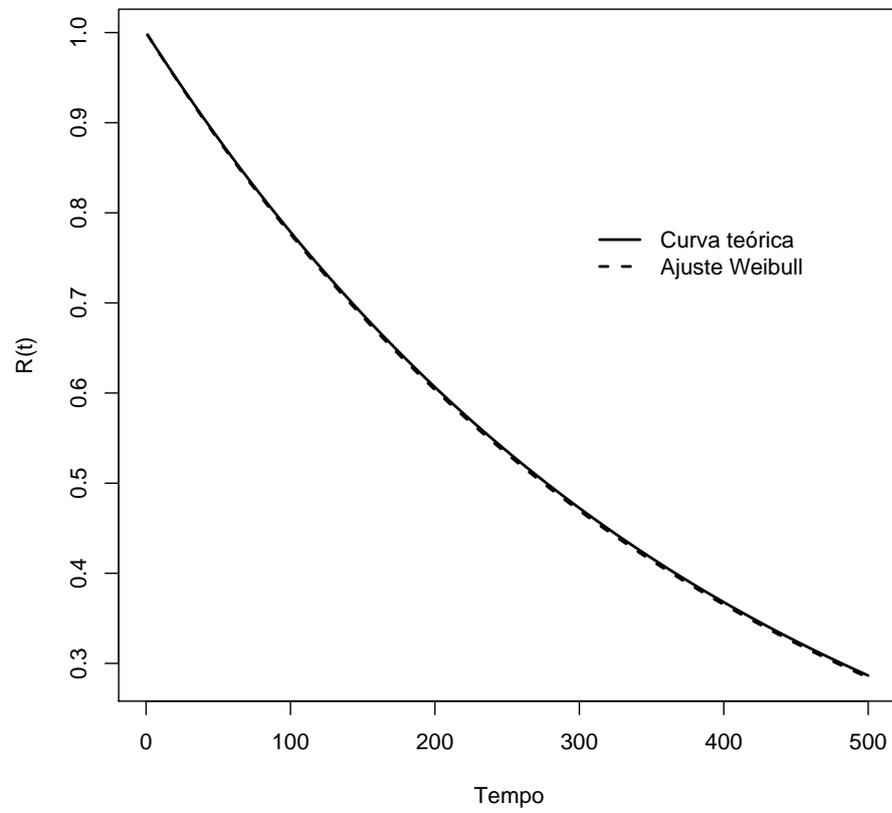


FIGURA 4.5: Ajuste da curva de confiabilidade do fator de risco 2.

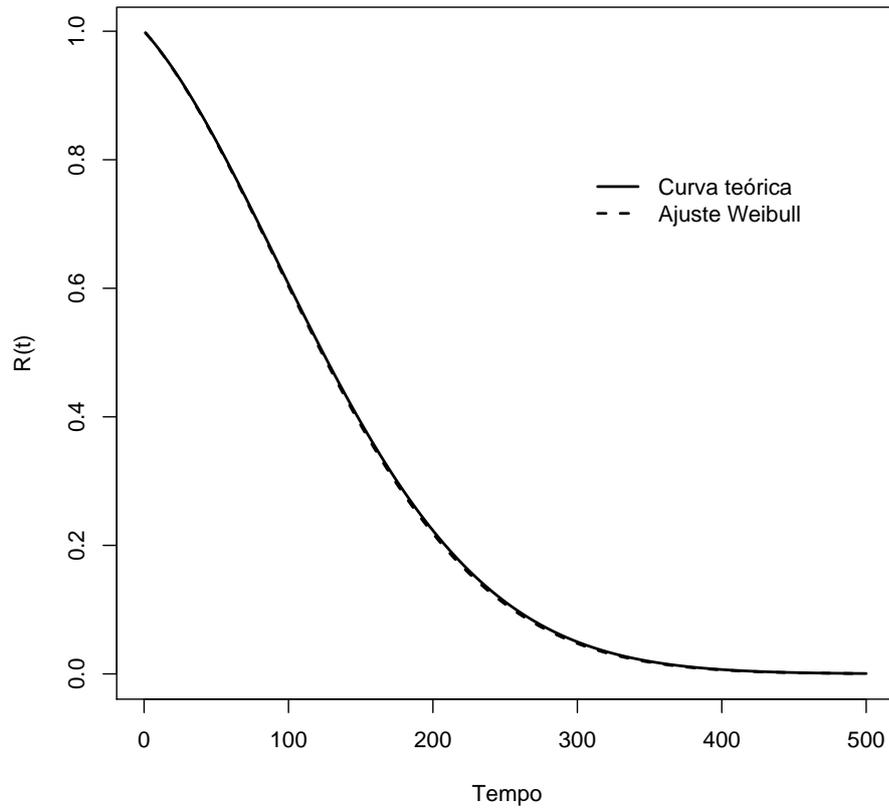


FIGURA 4.6: Ajuste da curva de confiabilidade do sistema.

Para verificar se há equivalência do comportamento de falha do fator 1 e do fator 2, utilizamos o critério de referência Bayesiano. Calculamos a medida de evidência $d(\mathcal{M}_0, \mathbf{t})$ para cada amostra, conforme descrito na Seção 3.5. Os resultados obtidos se encontram na Tabela 4.3. Nota-se que, para a amostra de tamanho 10, a medida fornece leve evidência contra a hipótese nula, porém, quando o tamanho da amostra aumenta, a mesma fornece evidência significativa para a rejeição de $H_0 \equiv \{\beta_1 = \beta_2, \eta_1 = \eta_2\}$.

TABELA 4.3: Medida de evidência contra a hipótese nula.

n	Medida de evidência	Decisão
10	1,92	Leve evidência
25	4,80	Evidência significativa
50	9,92	Evidência significativa
100	20,10	Evidência significativa

4.2 Estudo de simulação 2

Geramos quatro conjuntos de dados com tamanhos amostrais (n) iguais a 10, 25, 50 e 100, segundo um modelo $Gama(\alpha_\ell, \theta_\ell)$ com função densidade dada por

$$f(t_i|\alpha_\ell, \theta_\ell) = \frac{1}{\theta_\ell^{\alpha_\ell}} t_i^{(\alpha_\ell-1)} \exp\left\{-\frac{t_i}{\theta_\ell}\right\}.$$

Todos os conjuntos de dados foram gerados considerando $\alpha_1 = 12$, $\theta_1 = 0,5$, $\alpha_2 = 20$ e $\theta_2 = 1$. Obtivemos duas amostras das distribuições a posteriori marginais de β_1, η_1, β_2 e η_2 para cada n , utilizando o algoritmo de Metropolis-Hastings. As medidas resumo das amostras da cadeia 1, considerando aquecimento (*burn in*) de 20.000 com tamanho inicial 50.000 e salto igual a 5, se encontram nas Tabelas 4.4 e 4.5. Aplicando o teste de Gelman e Rubin para o estudo da convergência, obtivemos aproximadamente $\hat{R} = 1$ para todos os casos, o que traz grandes indícios de convergência para as distribuições a posteriori dos parâmetros. O tempo de processamento para cada cadeia simulada, em segundos, é igual a 159,71.

Analisando as Tabelas 4.4 e 4.5, observamos através dos intervalos de credibilidade que as estimativas são significativas. À medida que o tamanho da amostra aumenta, a tendência do desvio padrão é diminuir, o que é esperado.

As figuras apresentadas neste exemplo dizem respeito à amostra de tamanho 10, e para os demais casos o resultado é análogo. A Figura 4.7 mostra que há indícios de convergência para ambos os parâmetros, pois as cadeias se comportam dentro de uma “faixa”, ou seja, os valores gerados variam ao redor

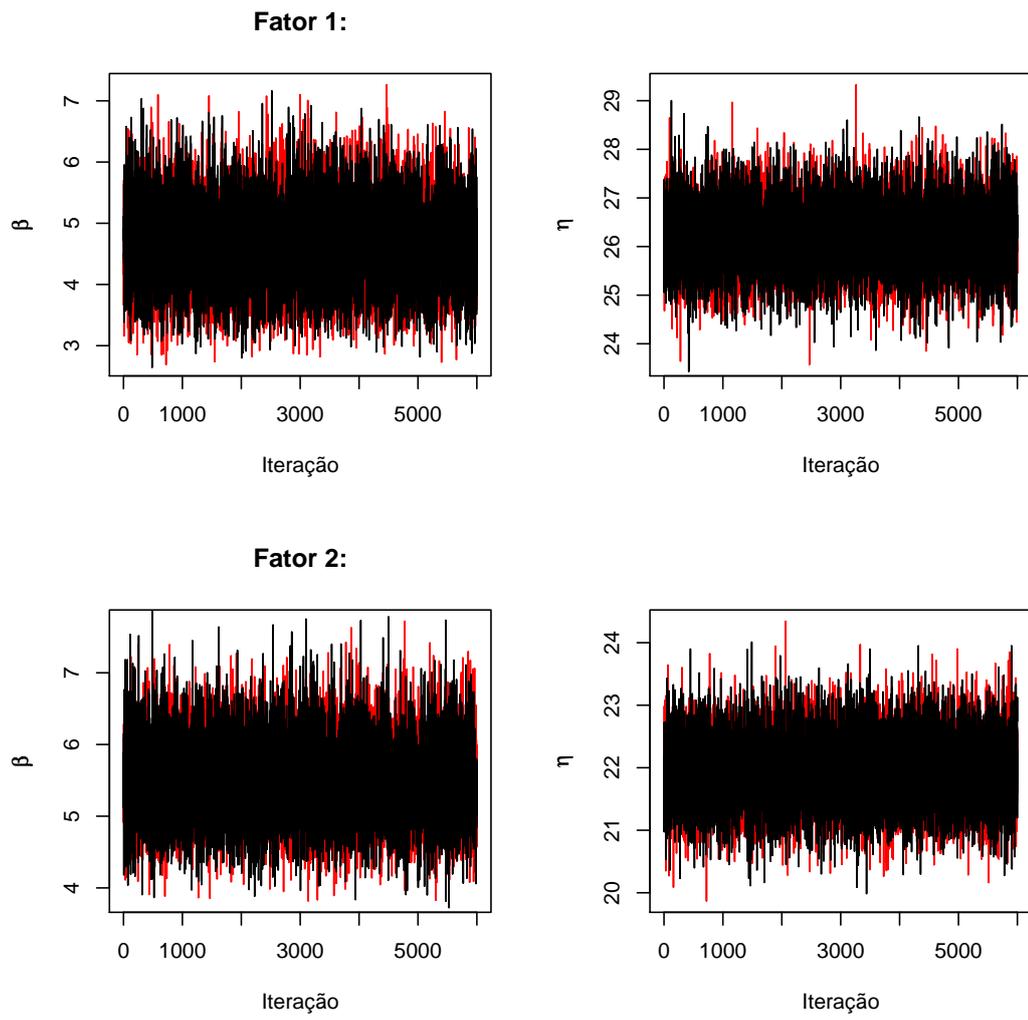
TABELA 4.4: Medidas resumo a posteriori para os parâmetros relacionados ao fator de risco 1.

n	β_1			η_1		
	Média	DP	Int. Cred. de 95%	Média	DP	Int. Cred. de 95%
10	4,617	0,672	[3,39 ; 6,04]	26,22	0,681	[24,92 ; 27,56]
25	4,240	0,556	[3,21 ; 5,40]	26,13	0,662	[24,88 ; 27,44]
50	4,039	0,468	[3,17 ; 5,00]	26,17	0,644	[24,95 ; 27,46]
100	4,495	0,438	[3,70 ; 5,40]	26,13	0,592	[25,01 ; 27,30]

TABELA 4.5: Medidas resumo a posteriori para os parâmetros relacionados ao fator de risco 2.

n	β_2			η_2		
	Média	DP	Int. Cred. de 95%	Média	DP	Int. Cred. de 95%
10	5,518	0,566	[4,45 ; 6,66]	21,93	0,556	[20,88 ; 23,03]
25	5,842	0,559	[4,80 ; 7,00]	22,24	0,498	[21,27 ; 23,24]
50	5,756	0,475	[4,88 ; 6,72]	22,39	0,442	[21,53 ; 23,30]
100	5,662	0,395	[4,90 ; 6,45]	22,08	0,377	[21,35 ; 22,82]

de uma média. Isso também pode ser observado na Figura 4.9, uma vez que as densidades estimadas dos parâmetros de interesse para as duas cadeias geradas estão próximas. Na Figura 4.8, verificamos que as médias das cadeias convergem para um determinado valor, indicando, após aproximadamente 20.000 iterações, a convergência das cadeias.

FIGURA 4.7: Traço das cadeias dos parâmetros dos fatores de risco 1 e 2 ($n = 10$).

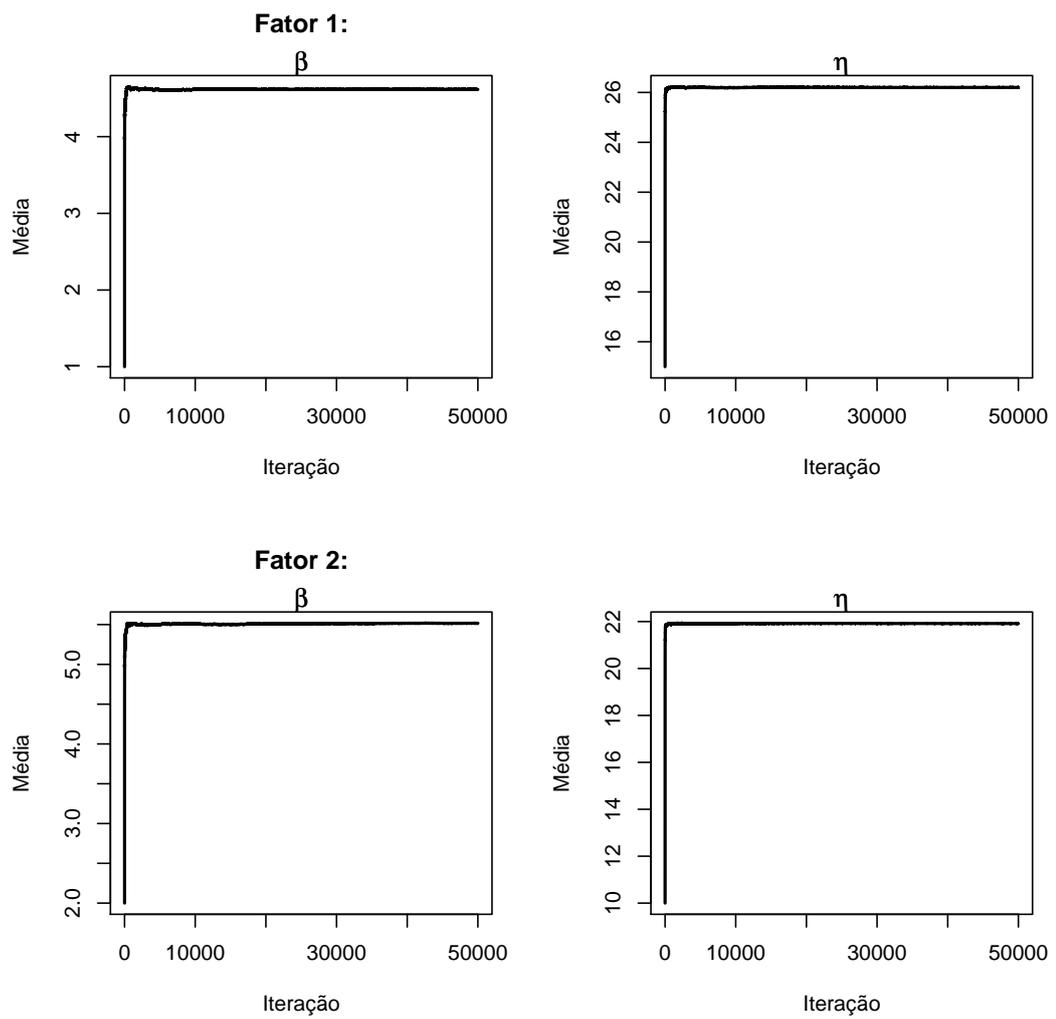


FIGURA 4.8: Médias ergódicas para os parâmetros dos fatores de risco 1 e 2 ($n = 10$).

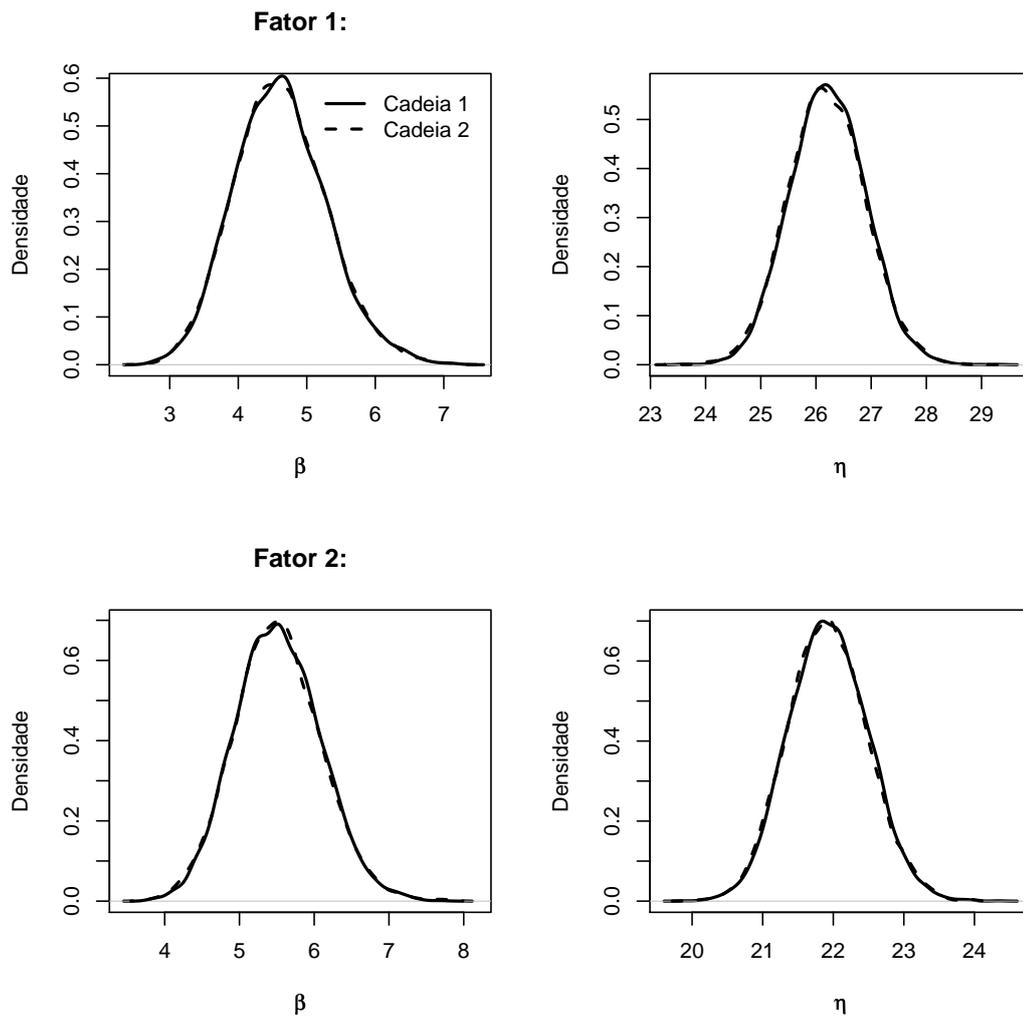


FIGURA 4.9: Funções densidade a posteriori estimadas dos parâmetros dos fatores de risco 1 e 2 ($n = 10$).

Analisando as Figuras 4.10, 4.11 e 4.12, as quais foram construídas considerando as estimativas encontradas nas Tabelas 4.4 e 4.5 e $n = 10$, verificamos que o ajuste do modelo Weibull é adequado aos dados em questão.

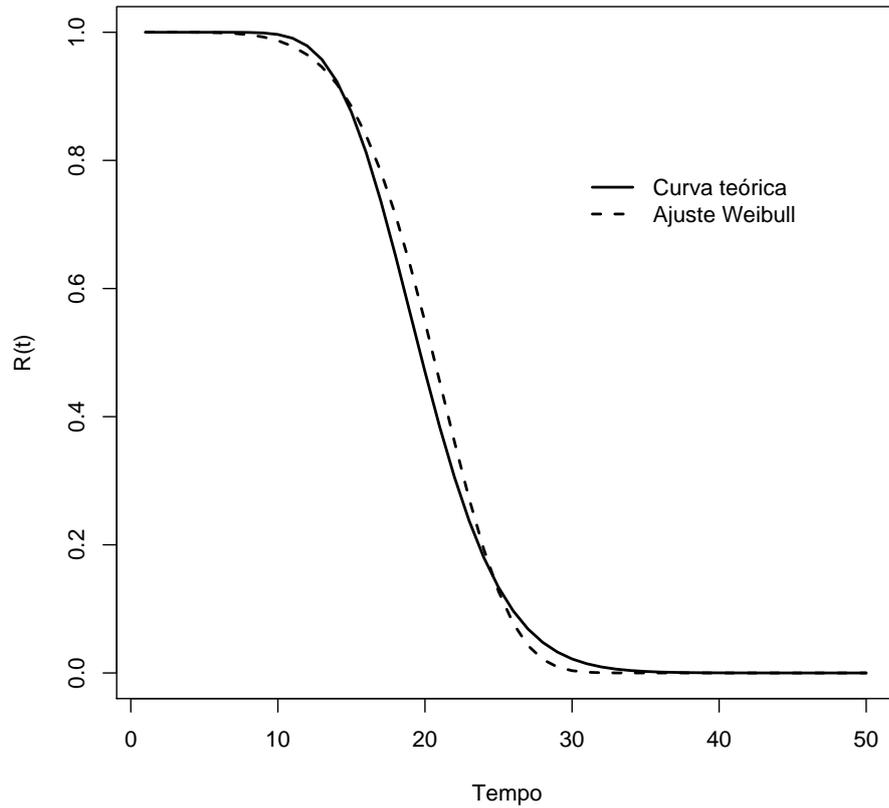


FIGURA 4.10: Ajuste da curva de confiabilidade do fator de risco 1.

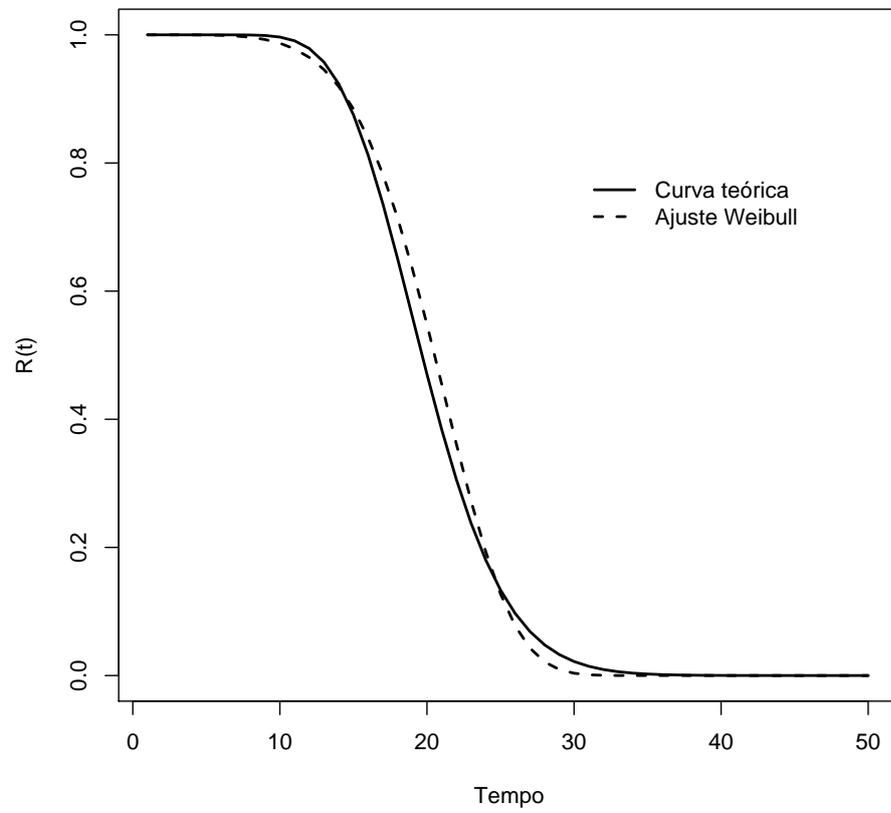


FIGURA 4.11: Ajuste da curva de confiabilidade do fator de risco 2.

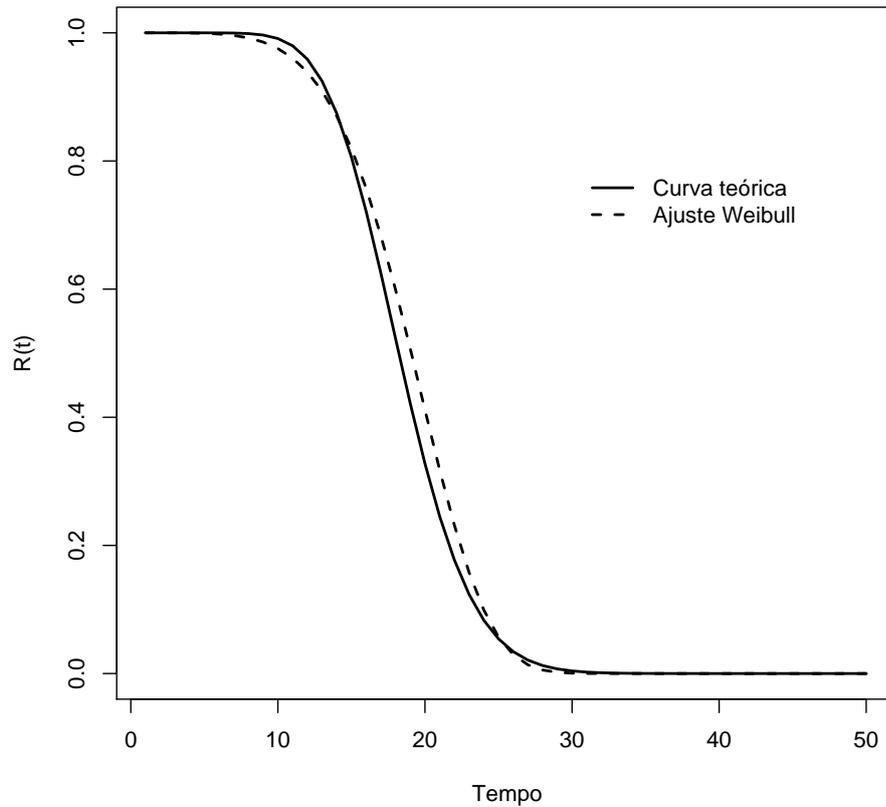


FIGURA 4.12: Ajuste da curva de confiabilidade do sistema.

Uma vez verificado que o método inferencial apresenta-se adequado, a seguir, apresentamos uma aplicação utilizando dados reais.

4.3 Exemplo com dados reais

Retornando aos dados do Exemplo 3.1 e utilizando o algoritmo de Metropolis-Hastings na equação (3.13), obtivemos duas amostras das distribuições a posteriori marginais de β_A, η_A, β_B e η_B . Agora, utilizamos os índices A e B ao invés de 1 e 2 para seguir a notação do Exemplo 3.1. As medidas resumo das amostras da cadeia 1, adotando aquecimento (*burn in*) de 25.000 com tamanho inicial 100.000 e salto igual a 10, se encontram na Tabela 4.6. O critério de Gelman e Rubin forneceu $\hat{R} = 1.00$ para ambas as cadeias dos parâmetros, indicando convergência das mesmas. O tempo de processamento para cada cadeia simulada,

em segundos, é igual a 1029,83.

TABELA 4.6: Medidas resumo a posteriori para os parâmetros do modelo (3.6).

Parâmetro	Média	Desvio Padrão	Intervalo de Credibilidade(95%)
β_A	2,29	0,27	[1,80 ; 2,85]
η_A	1399,00	72,66	[1261,63 ; 1544,98]
β_B	2,26	0,27	[1,74 ; 2,81]
η_B	1506,00	81,99	[1354,19 ; 1672,78]

Observamos na Tabela 4.6 que as estimativas dos parâmetros são significativas, uma vez que o valor zero não pertence ao intervalo de credibilidade de 95%, e que as mesmas estão próximas dos resultados assintóticos encontrados na Tabela 3.2. Comparando os resultados encontrados na Tabela 4.6 com os da Tabela 3.2, verificamos que o método Bayesiano forneceu desvios padrão menores e, conseqüentemente, intervalos com amplitude menor.

A Figura 4.13 sinaliza que há indícios de convergência para ambos os parâmetros, pois as cadeias se comportam dentro de uma “faixa”, ou seja, os valores gerados variam ao redor de um ponto médio. Isso também pode ser observado na Figura 4.15, uma vez que as densidades estimadas dos parâmetros de interesse para as duas cadeias geradas estão próximas. Analisando a Figura 4.14 verificamos que a média da cadeia converge para um determinado valor, indicando, após aproximadamente 25.000 iterações, a convergência das cadeias.

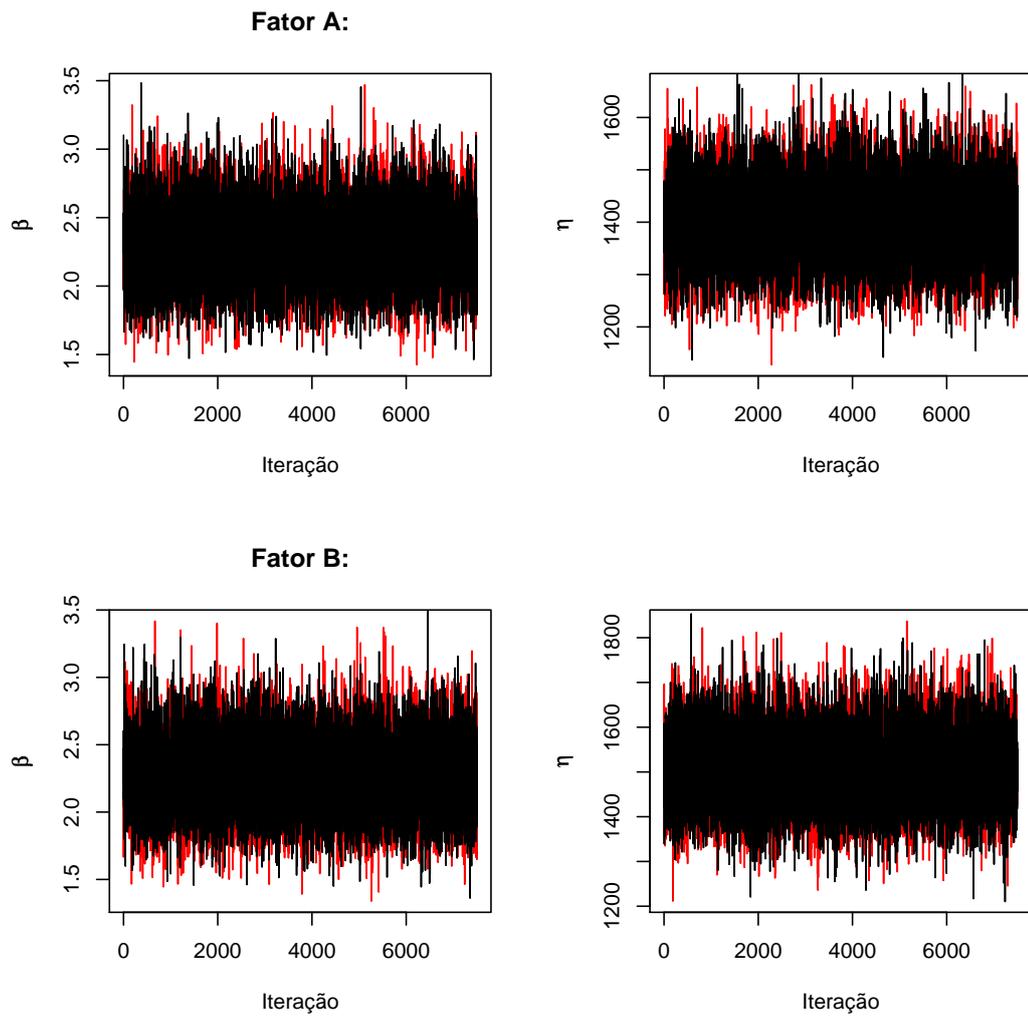


FIGURA 4.13: Traço das cadeias dos parâmetros dos fatores de risco A e B.

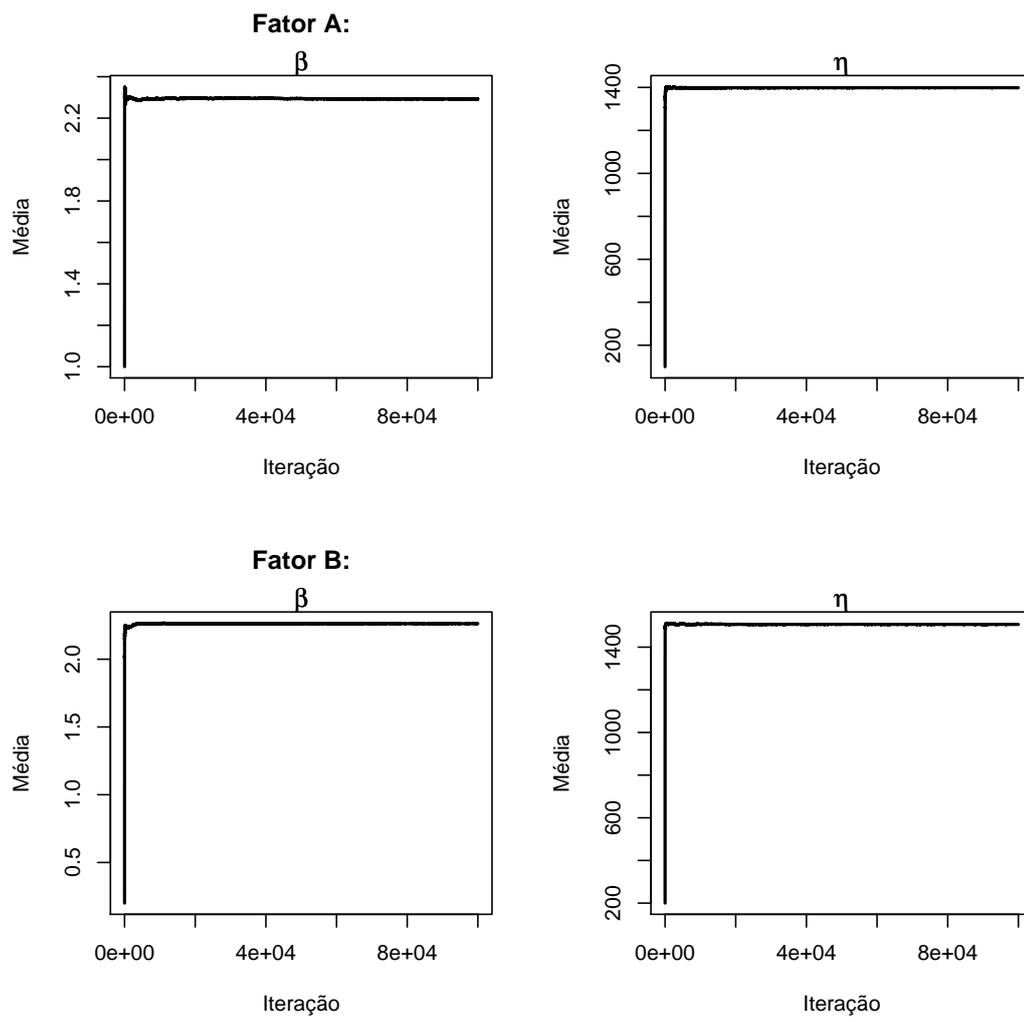


FIGURA 4.14: Médias ergódicas para os parâmetros dos fatores de risco A e B.

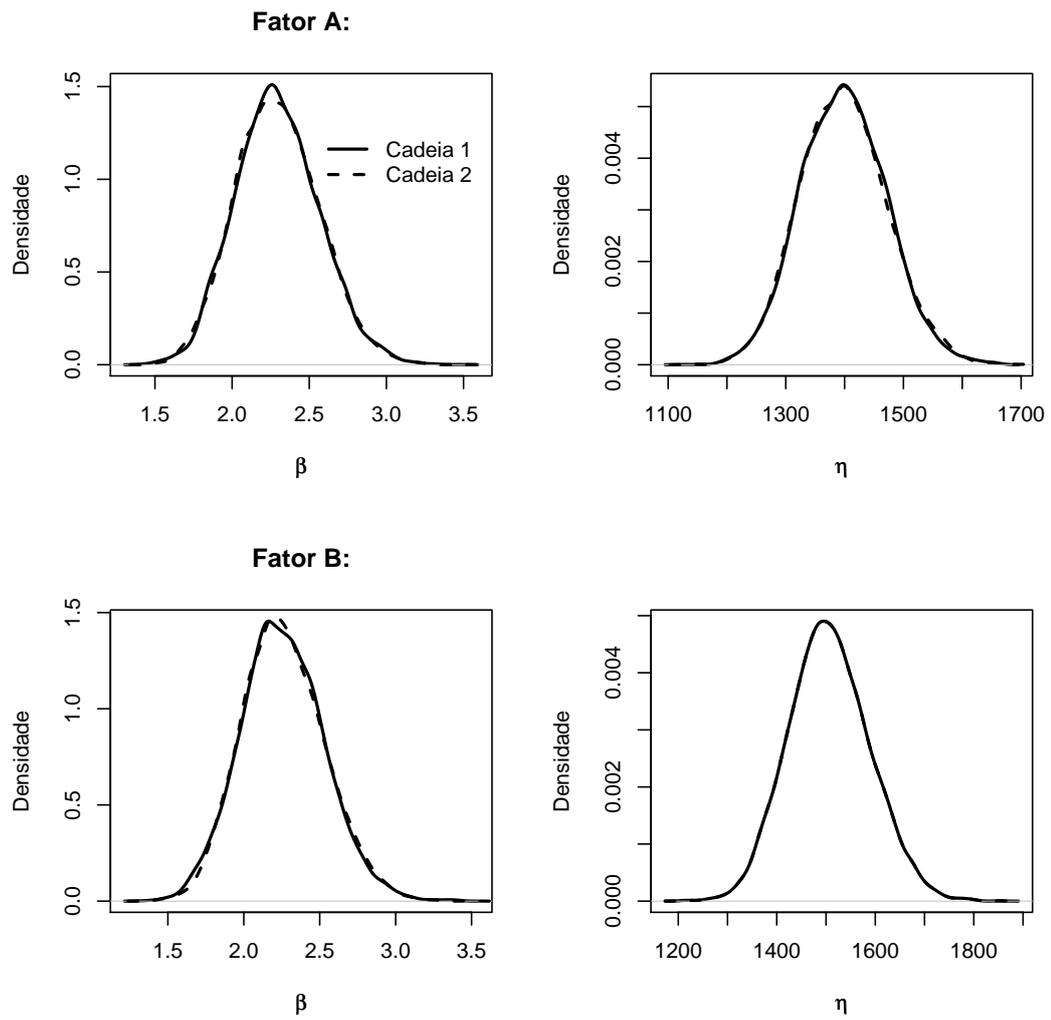


FIGURA 4.15: Funções densidade a posteriori estimadas dos parâmetros dos fatores de risco A e B.

Nas Figuras 4.16, 4.17 e 4.18, apresentamos as estimativas da confiabilidade obtidas com o ajuste Weibull juntamente com as estimativas e limites de confiança de Kaplan-Meier .

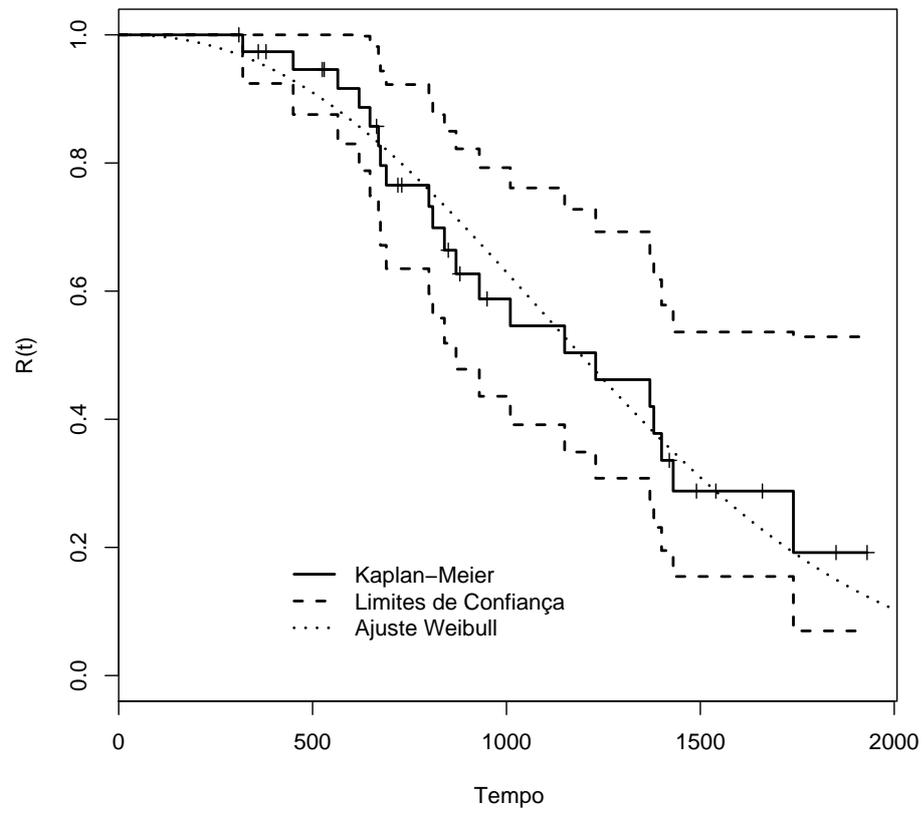


FIGURA 4.16: Ajuste da curva de confiabilidade do fator de risco A.

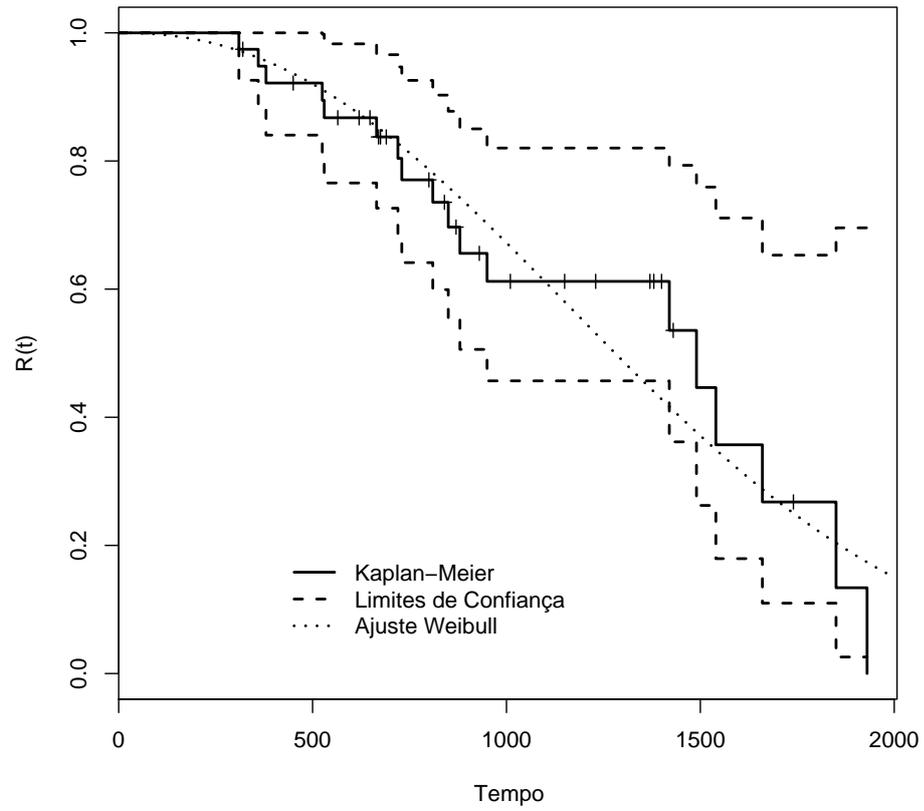


FIGURA 4.17: Ajuste da curva de confiabilidade do fator de risco B.

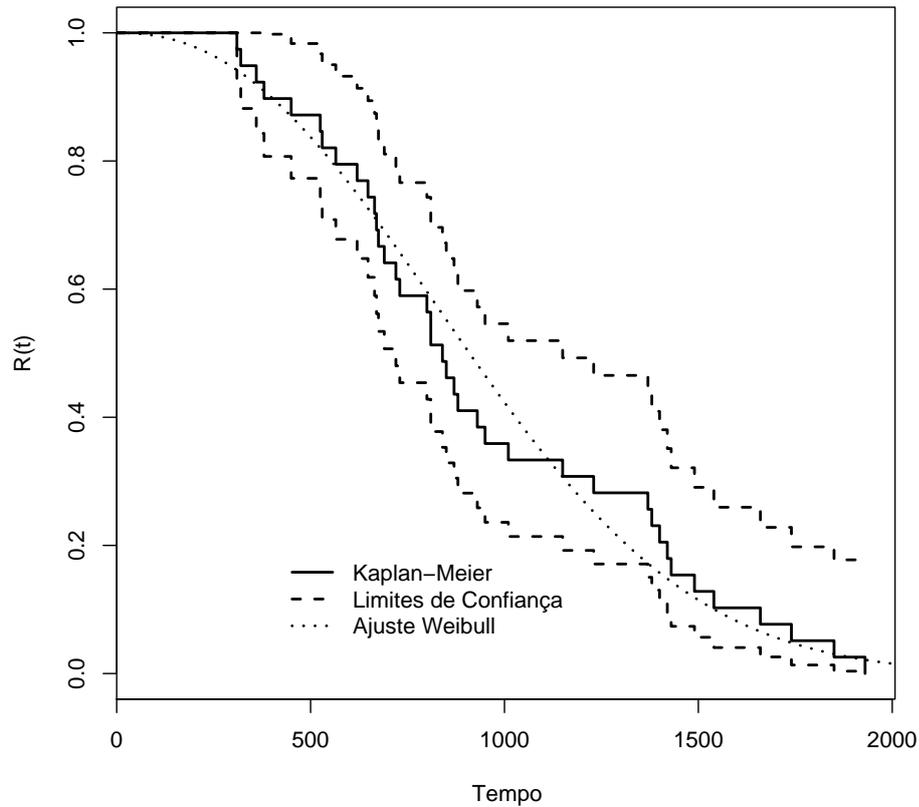


FIGURA 4.18: Ajuste da curva de confiabilidade do sistema.

Para verificar se há equivalência entre o comportamento de falha do fator A e do fator B, utilizamos o critério de referência Bayesiano. Calculamos a medida de evidência, conforme descrita na Seção 3.4, e obtivemos $d(\mathcal{M}_0, \mathbf{t}) = 0,89$, favorecendo a hipótese de igualdade dos parâmetros. Ou seja, as curvas de confiabilidade dos fatores de risco são estatisticamente iguais.

4.4 Estudo de simulação 3

Nesta seção, para fins de exemplificação, apresentamos um estudo de simulação considerando a abordagem descrita da Seção 3.4 e a apresentada na Seção 3.6, em que trabalhamos diretamente com as curvas de confiabilidade dos fatores de risco competitivos.

Geramos um conjunto de dados com tamanho amostral (n) igual a 100,

segundo o modelo Weibull (3.3), tomando $\beta_1 = 2$, $\eta_1 = 4$, $\beta_2 = 3$ e $\eta_2 = 5$. Obtivemos duas amostras das distribuições a posteriori marginais de R_1 , R_2 , β_1 , η_1 , β_2 e η_2 para cada método inferencial, utilizando o algoritmo de Metropolis-Hastings. As medidas resumo das amostras da cadeia 1, se encontram nas Tabelas 4.7, 4.8 (referentes ao método inferencial descrito na Seção 3.6) e 4.9 (referente ao método inferencial descrito na Seção 3.4). Aplicando o critério de Gelman e Rubin, obtivemos $\hat{R} \approx 1$, que traz grandes indícios de convergência. O tempo de processamento em média para cada cadeia simulada, em segundos, é igual a 951,25.

Considerando o método proposto na Seção 3.6, observamos nas Tabelas 4.7 e 4.8 que as estimativas estão próximas do verdadeiro valor de cada parâmetro, e que os intervalos de credibilidade contêm o verdadeiro valor de ambos os parâmetros. Vale ressaltar que esse método inferencial nos permite, mais facilmente, estimar os limites de credibilidade para as funções de confiabilidade dos fatores de risco e do sistema.

As figuras apresentadas neste exemplo dizem respeito a $t_1 = t_2 = 10$. A Figura 4.19 mostra que há indícios de convergência para ambos os parâmetros. Isso também pode ser observado na Figura 4.21, uma vez que as densidades estimadas dos parâmetros de interesse para as duas cadeias geradas estão próximas. Na Figura 4.20, verificamos que as médias das cadeias convergem para um determinado valor, indicando, após aproximadamente 10.000 iterações, a convergência das cadeias.

Considerando o método proposto na Seção 3.4, observamos pela Tabela 4.9 que as estimativas estão próximas do verdadeiro valor, e que os intervalos de credibilidade contêm o verdadeiro valor de cada parâmetro.

As Figuras 4.22 e Figura 4.24, mostram que há indícios de convergência para ambos os parâmetros. Na Figura 4.23, verificamos que as médias das cadeias convergem para um determinado valor, indicando, após aproximadamente 10.000 iterações, a convergência das cadeias.

TABELA 4.7: Medidas resumo a posteriori para confiabilidade do fator de risco

1.

t_1	Média	R verdadeiro	DP	Int. Cred. de 95%
1	0,9469	0,9394	0,0054	[0,936; 0,956]
2	0,8134	0,7788	0,2601	[0,759; 0,859]
3	0,5917	0,5697	0,0353	[0,523; 0,660]
4	0,3665	0,3678	0,0314	[0,305; 0,428]
5	0,2067	0,2096	0,0091	[0,189; 0,225]
6	0,1018	0,1054	0,0090	[0,085; 0,120]
7	0,0411	0,0468	0,0074	[0,028; 0,056]
8	0,0165	0,0183	0,0032	[0,011; 0,023]
9	0,0045	0,0063	0,0016	[0,002; 0,008]
10	0,0003	0,0019	0,0004	[0,000004; 0,0015]

TABELA 4.8: Medidas resumo a posteriori para confiabilidade do fator de risco

2.

t_2	Média	R verdadeiro	DP	Int. Cred. de 95%
1	0,9919	0,9920	0,0016	[0,988; 0,994]
2	0,9201	0,9380	0,0129	[0,891; 0,942]
3	0,7977	0,8057	0,0286	[0,741; 0,848]
4	0,6150	0,5992	0,0404	[0,530; 0,688]
5	0,3679	0,3679	0,0145	[0,341; 0,397]
6	0,1764	0,1776	0,0114	[0,155; 0,199]
7	0,0658	0,0643	0,0092	[0,049; 0,085]
8	0,0155	0,0166	0,0030	[0,010; 0,022]
9	0,0022	0,0029	0,0012	[0,0006; 0,005]
10	0,0004	0,0003	0,0004	[0,00001; 0,0015]

TABELA 4.9: Medidas resumo para os parâmetros do modelo (3.6).

Parâmetro	Média	DP	Int. de Cred. 95%
β_1	2,243	0,169	[1,921 ; 2,589]
η_1	4,002	0,181	[3,656 ; 4,388]
β_2	2,818	0,184	[2,468 ; 3,185]
η_2	5,080	0,238	[4,644 ; 5,574]

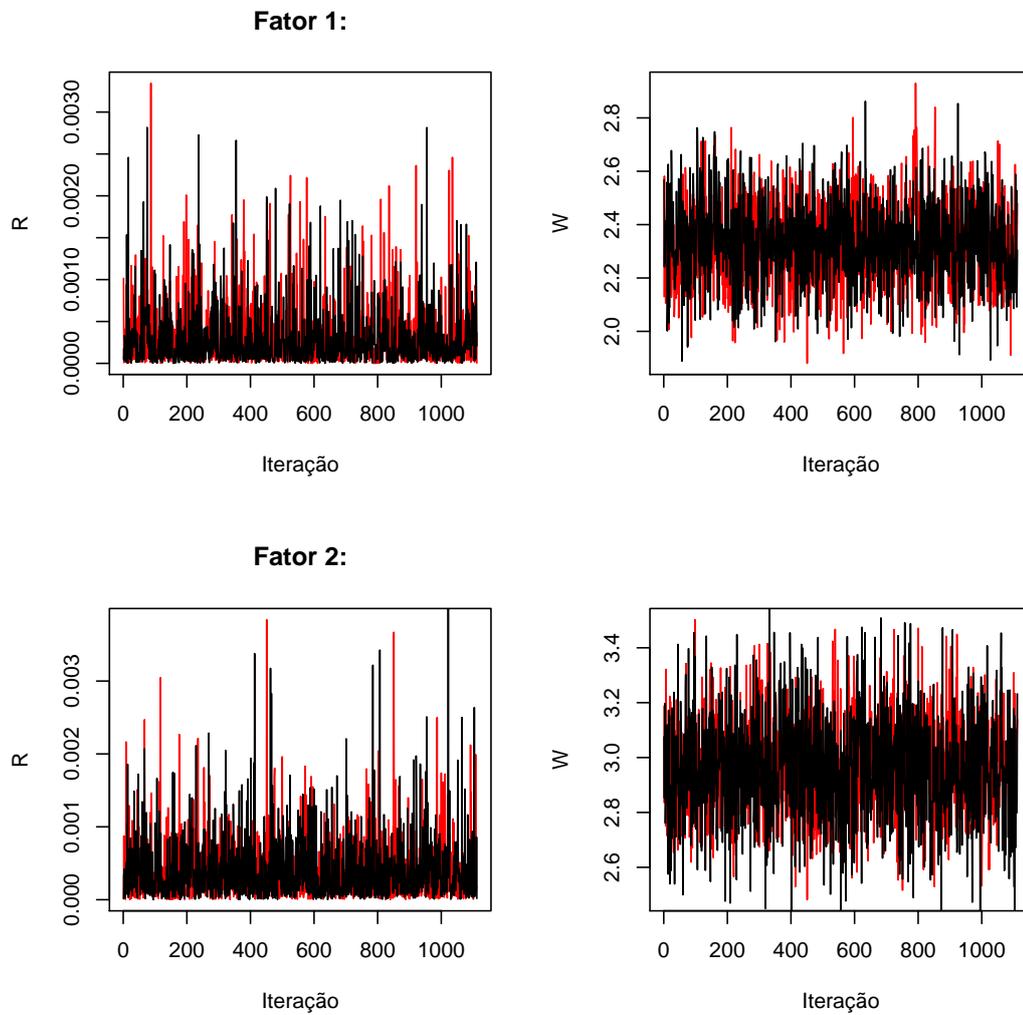


FIGURA 4.19: Traço das cadeias dos parâmetros dos fatores de risco 1 e 2.

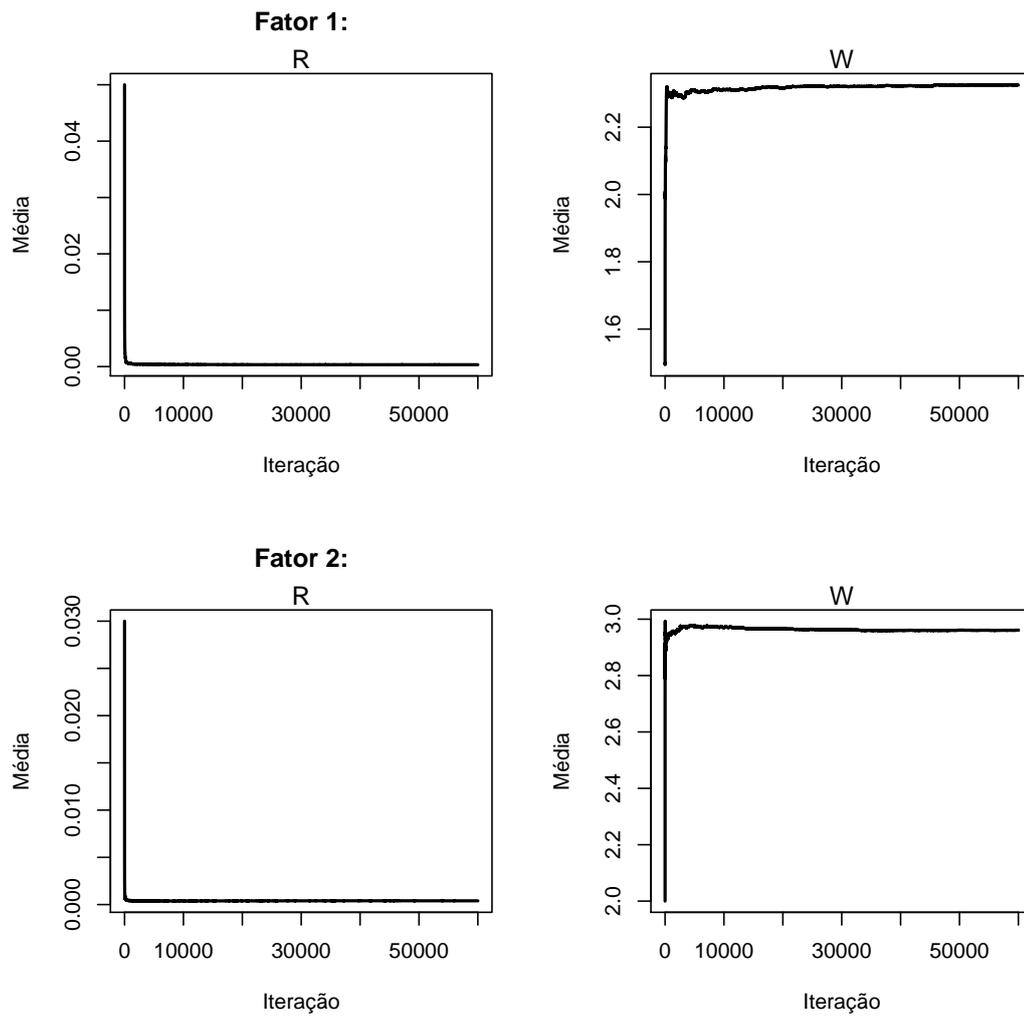


FIGURA 4.20: Médias ergódicas para os parâmetros dos fatores de risco 1 e 2.

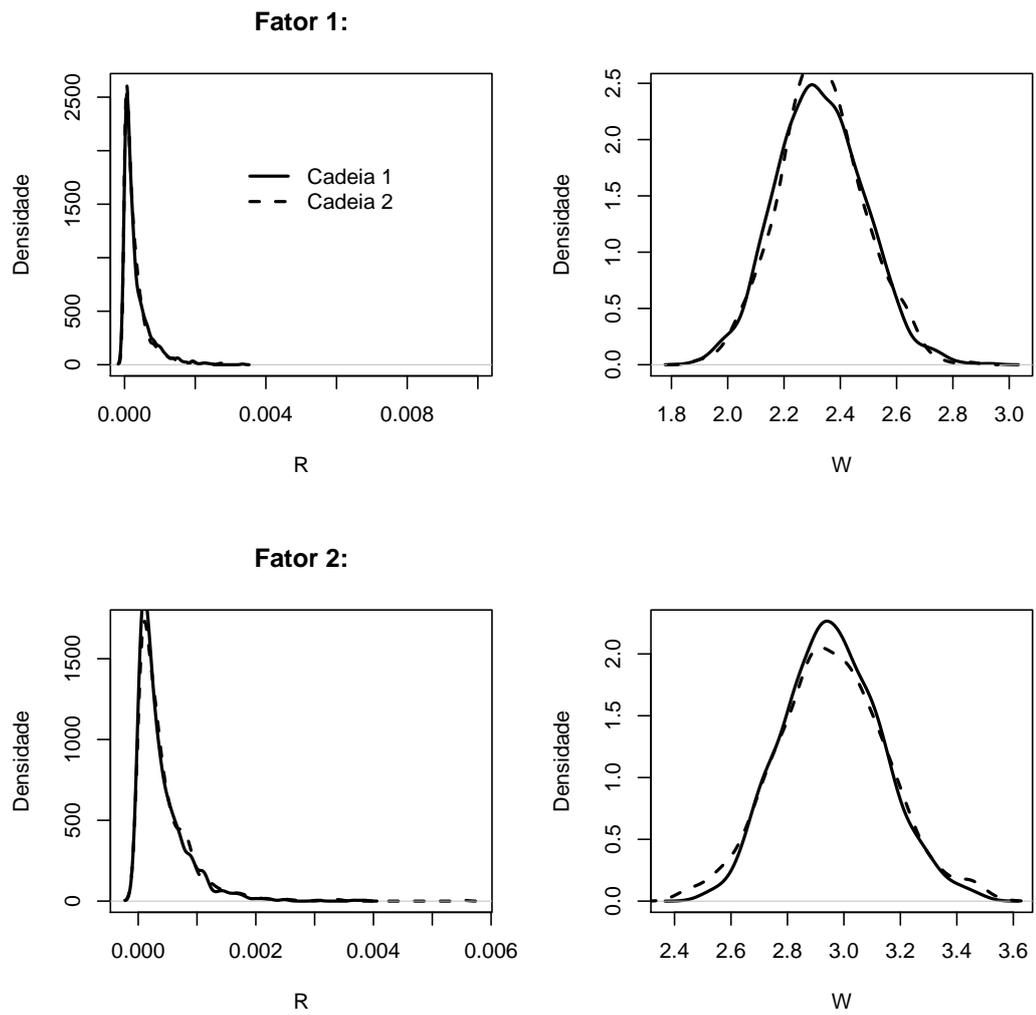


FIGURA 4.21: Funções densidade a posteriori estimadas dos parâmetros dos fatores de risco 1 e 2.

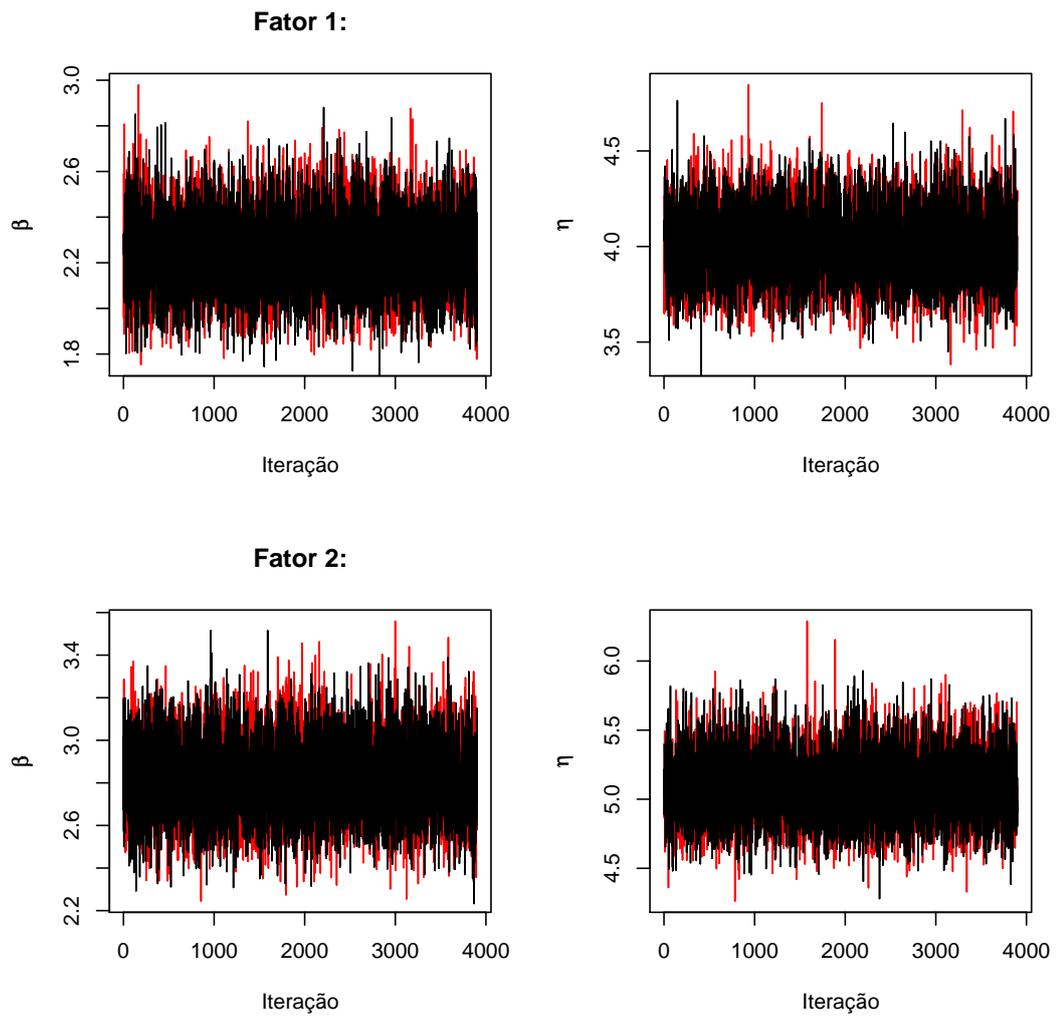


FIGURA 4.22: Traço das cadeias dos parâmetros dos fatores de risco 1 e 2.

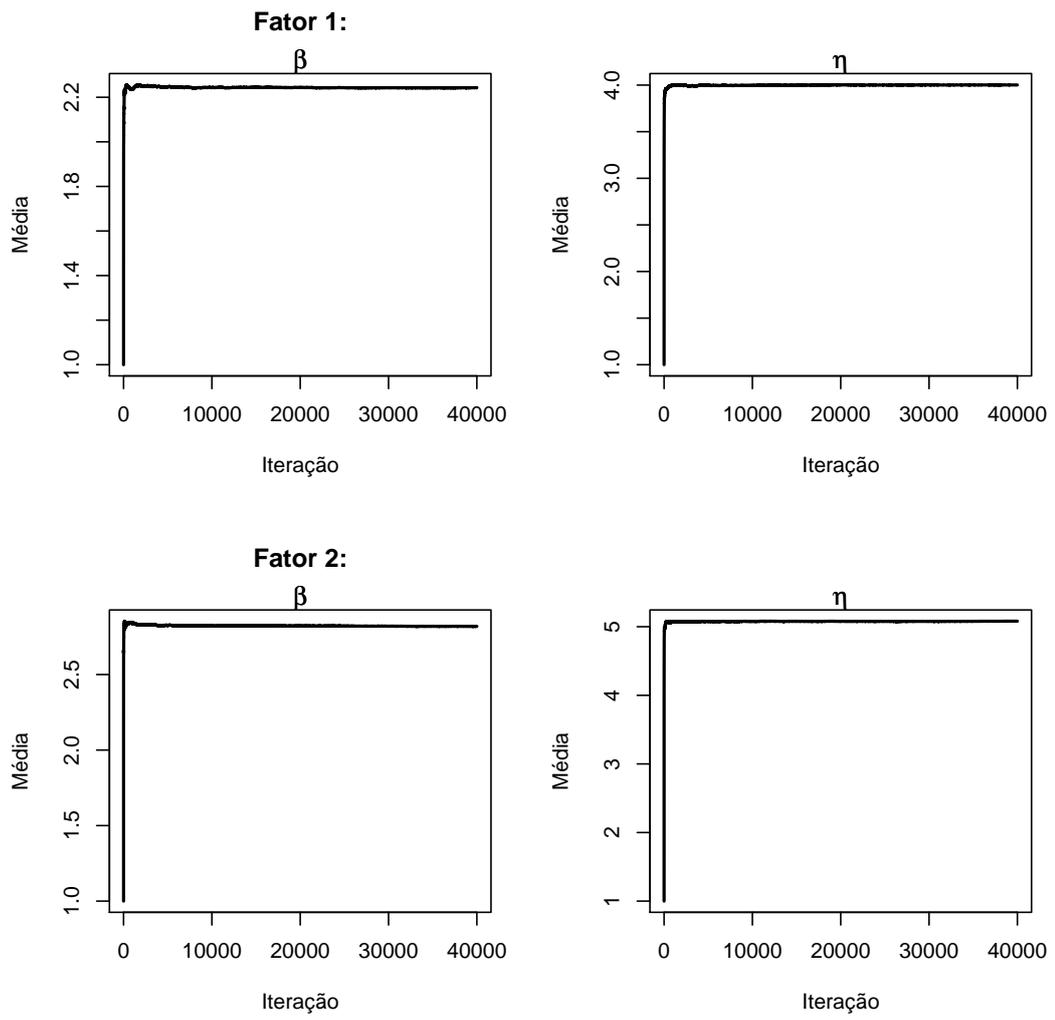


FIGURA 4.23: Médias ergódicas para os parâmetros dos fatores de risco 1 e 2.

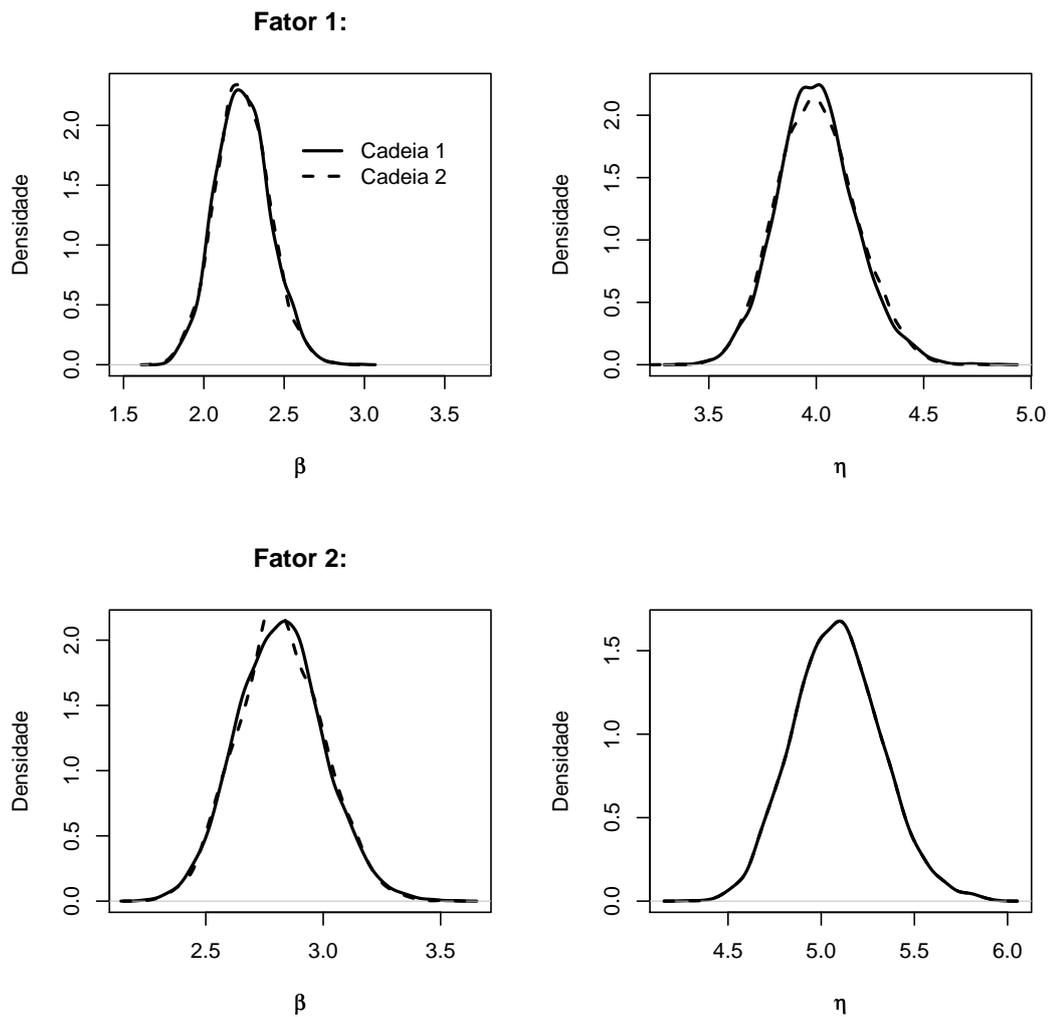


FIGURA 4.24: Funções densidade a posteriori estimadas dos parâmetros dos fatores de risco 1 e 2.

Nas Figuras 4.25, 4.26 e 4.27, apresentamos as estimativas da confiabilidade obtidas via métodos inferenciais descritos nas Seções 3.4 e 3.6 e a curva teórica do problema em questão. Note que ambos os métodos fornecem curvas próximas da teórica, indicando que ambos os métodos fornecem resultados satisfatórios. Para uma melhor visualização da comparação entre as curvas estimadas e a teórica, foi calculado o erro quadrático médio entre as mesmas (Tabela 4.10). Como os erros de ambos os métodos são similares e próximos de zero, não podemos concluir se há um método mais eficiente, apenas que ambos são adequados, com base neste conjunto de dados. Contudo, devemos fazer algumas observações em relação ao método apresentado na Seção 3.6. O cálculo das medidas resumo

demanda um esforço computacional maior, sendo que para valores de t_1 e t_2 que fornecem valores extremos na curva de confiabilidade apresentam complicação da estimação das medidas de interesse, porém o método permite o cálculo direto dos limites de credibilidade para a função de confiabilidade. Para uma melhor avaliação do método proposto, mais estudos devem ser realizados.

TABELA 4.10: Erro quadrático médio entre as curvas estimadas de confiabilidade e a teórica.

	Modelo (3.6)	Modelo (3.24)
Fator de risco 1	0.0002637938	0.0001634832
Fator de risco 2	0.0001420802	5.794376×10^{-05}
Sistema	8.337791×10^{-05}	5.145804×10^{-05}

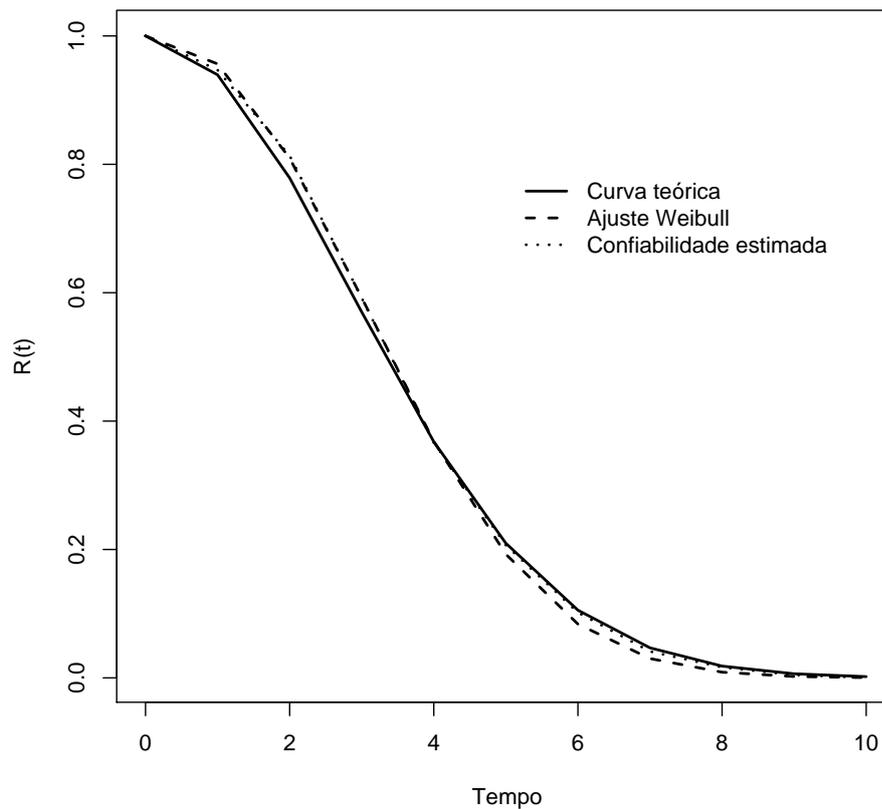


FIGURA 4.25: Ajuste da curva de confiabilidade do fator de risco 1.

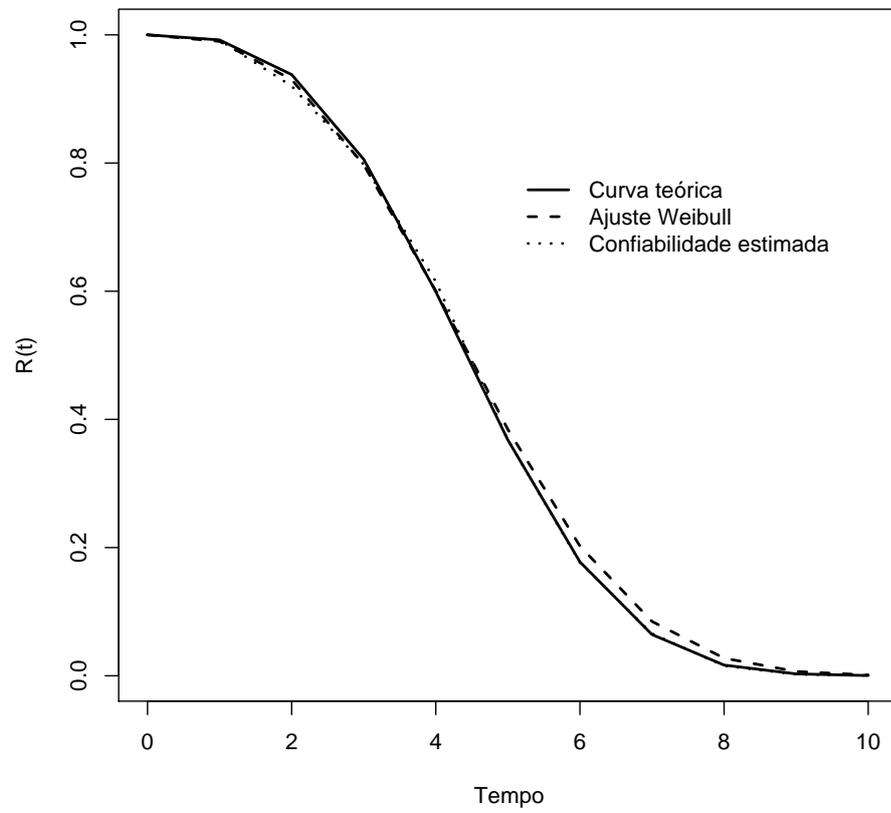


FIGURA 4.26: Ajuste da curva de confiabilidade do fator de risco 1

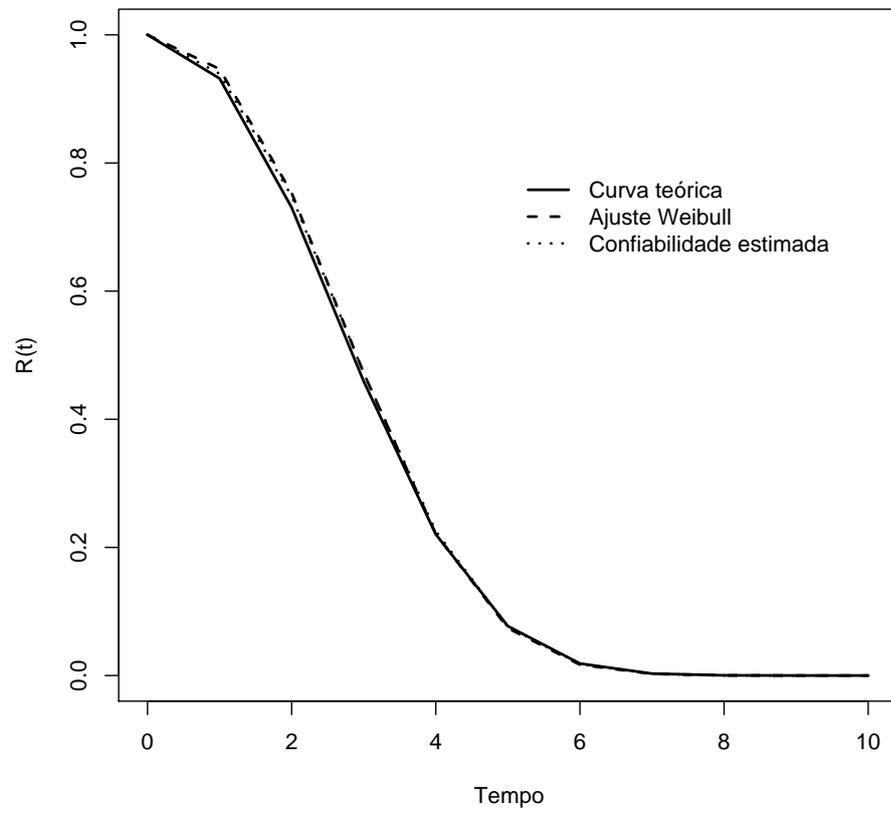


FIGURA 4.27: Ajuste da curva de confiabilidade do sistema.

Capítulo 5

Considerações Finais

Neste trabalho, estudamos, sob a perspectiva de análise de referência Bayesiana, a aplicação do modelo Weibull em problemas de confiabilidade sujeitos a fatores de risco competitivos. A utilização da distribuição Weibull foi um modelo adequado para representar os diferentes fatores de riscos, pois, devido apenas a mudanças nos valores de seus parâmetros, podemos representar diversas situações de falha, o que é comum na presença de diversos fatores de falha.

O modelo Weibull foi estudado com duas parametrizações diferentes: a primeira, considerando os parâmetros naturalmente (como na maioria dos textos); e a segunda, colocando a função de confiabilidade como um parâmetro do modelo, trabalhando, assim, diretamente com o interesse do estudo na modelagem. Os resultados obtidos nos levam a crer que os modelos em questão têm boa capacidade de adequação aos dados, segundo as duas parametrizações. A vantagem no segundo caso é que conseguimos construir limites de credibilidade mais facilmente do que no primeiro. Porém, os procedimentos computacionais para a estimação se tornam mais lentos neste caso.

Neste trabalho, destacam-se como tópico de pesquisas futuras, a necessidade de:

- Extensão dos estimadores para mais de dois fatores de risco competitivos;
- Estudar as propriedades do estimador desenvolvido na Seção 3.6 para melhor avaliação do método proposto, e

- Considerar o problema de sistemas em paralelo.

Referências Bibliográficas

- [1] BARLOW, J. O. E PROCHAN, F. *Statistical theory of reliability and life testing: probability models*. 1975.
- [2] BERGER, J. O. Bayesian analysis: A look at today and thoughts of tomorrow. *American Statistics Association* 95 (2000), 1269–1276.
- [3] BERGER, J. O. The case for objective bayesian analysis. *Bayesian Analysis, Rev.* 3 (2006), 385–402.
- [4] BERGER, J. O. E BERNARDO, J. M. Estimating a product of means: Bayesian analysis with reference priors. *American Statistics Association* 84 (1989), 200–2007.
- [5] BERGER, J. O. E BERNARDO, J. M. Ordered group reference priors with applications to a multinomial problem. *Biometrika* 79 (1992a), 25–27.
- [6] BERGER, J. O. E BERNARDO, J. M. Reference priors in a variance components problem. *Bayesian Analysis in Statistics an Econometrics* (1992b), 323–340.
- [7] BERGER, J. O. E BERNARDO, J. M. On the development of reference priors. *Bayesian Statistics* 4 (1992c), 35–60.
- [8] BERGER, J. O.; BERNARDO, J. M. E SUN, D. Reference priors from first principles: A general definition. Tech. rep., SAMSI, NC, USA, 2005.
- [9] BERGER, J. O.; BERNARDO, J. M. E SUN, D. The formal definition of reference priors. *Submitted of Annals of Statistics* 36 (2008).
- [10] BERNARDO, J. M. Expected information as expected utility. *Annals of Statistics* 7 (1979a), 686–690.
- [11] BERNARDO, J. M. Reference posterior distribution for bayesian inference. *Royal Statistics Society B* 41 (1979b), 113–147.
- [12] BERNARDO, J. M. Reference analysis. *Handbook of Statistics* 25 (2005), 17–90.
- [13] BERNARDO, J. M. E RAMÓN, J. M. An introduction to bayesian reference analysis: Inference on the ratio of multinomial parameters. *Royal Statistics Society D* 47 (1998), 101–135.

- [14] BERNARDO, J. M. E RUEDA, R. Bayesian hypothesis testing: A reference approach. *International Statistical Rev.* 70 (2002), 351–372.
- [15] BERNARDO, J. M. E SMITH, A. F. M. *Bayesian Theory*. 1994.
- [16] CHAIBUB NETO, E. Análise bayesiana de referência para modelos de calibração. Master's thesis, IME-USP, 2003.
- [17] COLOSIMO, E. A. E GIOLO, S. R. *Análise de sobrevivência aplicada*. 2006.
- [18] COQUE, M. A. Fatores competitivos sob o modelo weibull: Uma perspectiva bayesiana. Master's thesis, IME-USP, 2003.
- [19] GAMERMAN, D. E LOPES, H. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. 1997.
- [20] GELMAN, A. E RUBIN, D. R. A single series from the gibbs sampler provides a false sense of security. *Bayesian Statistics 4* (1992), 625–631.
- [21] GICHANGI, A. E VACH, W. The analysis of competing risks data: A guided tour. *Statistics Medicine* (2005).
- [22] HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika* 57 (1970), 97–109.
- [23] HOUGAARD, P. *Analysis of multivariate survival data*. 2000.
- [24] LAWLESS, J. F. *Statistical models and methods for lifetime data*. 1982.
- [25] LINDQVIST, H. Competing risks. Tech. rep., Norwegian University of Science and Tecnology, 2006.
- [26] LOUZADA NETO, F.; MAZUCHELI, J. E ACHCAR, J. A. *Introdução a análise de sobrevivência e confiabilidade*. 2002.
- [27] MAGALHÃES, M. N. *Probabilidade e variáveis aleatórias*. 2006.
- [28] MARTINS, C. B.; TOMAZELLA, V. L. D. E POLPO, A. Análise de referência bayesiana objetiva: Inferência para o modelo weibull. Tech. rep., UFSCar - DEs, 2008.
- [29] MARTZ, F. H., AND WALLER, A. R. *Bayesian reliability analysis*. 1982.
- [30] MOALA, F. A.; RODRIGUES, J. E TOMAZELLA, V. L. D. A note on the prior distribution for the weibull reliability function. *Communication in Statistics - Theory and Methods / Aceito para publicação* (2008).
- [31] NAKANO, E. Y. Uma extensão do modelo weibull bivariado de ryu: Uma aplicação bayesiana em riscos competitivos. Master's thesis, UFSCar, 2003.
- [32] PAULINO, C. D.; TURKMAN, M. A. A. E MURTEIRA, B. *Estatística Bayesiana*. 2003.
- [33] RODRIGUES, J.; CANCHO, V. G. E DE CASTRO, M. *Teoria unificada de análise de sobrevivência*. 2008.

-
- [34] WEIBULL, W. A statistical theory of strength of materials. *Ingeniors Vetenskaps Akademien 151-3* (1939), 45–55.
- [35] WEIBULL, W. A. A statistical distribution of wide applicability. *Journal of Applied Mechaniscs 18* (1951), 293–297.
- [36] YANG, R. E BERGER, J. O. A catalog of noninformative priors. Tech. rep., Perdue University, 1998.

Apêndice A

Dados Simulados

A.1 Conjunto de dados do Exemplo 2.1

TABELA A.1: Dados simulados segundo o modelo Weibull (1.1) com parâmetros $\beta = 2$ e $\eta = 4$.

2,29	1,46	2,09	1,39	3,54	5,36	4,24	3,29	2,26	0,41
7,34	5,49	2,22	10,42	3,88	3,51	3,89	5,25	0,89	3,56
4,23	0,75	2,35	2,65	3,88	2,40	3,12	4,87	3,40	1,93
9,05	5,17	2,48	3,99	4,03	1,50	1,27	2,78	5,67	1,98
3,68	1,10	2,03	4,64	4,26	6,71	7,08	6,81	2,74	3,82

A.2 Dados do Estudo de simulação 1

TABELA A.2: Dados simulados segundo o modelo Weibull (3.3) e $n = 10$.

Tempo	Fator de Risco	Tempo	Fator de Risco
33,16	2	126,17	1
40,30	2	141,80	1
94,04	1	166,29	1
107,77	2	221,72	1
118,43	1	254,47	1

TABELA A.3: Dados simulados segundo o modelo Weibull (3.3) e $n = 25$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
12, 54	2	126, 17	1	213, 63	1
33, 16	2	135, 91	1	221, 72	1
40, 30	2	141, 80	1	228, 03	1
50, 70	1	149, 78	2	232, 69	1
94, 04	1	166, 29	1	237, 50	2
107, 77	2	173, 69	1	254, 47	1
112, 43	1	193, 00	1	353, 90	1
113, 38	1	199, 75	1		
118, 43	1	204, 51	2		

TABELA A.4: Dados simulados segundo o modelo Weibull (3.3) e $n = 50$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
0, 36	2	113, 38	1	177, 8	1
2, 84	2	118, 43	1	189, 28	1
12, 54	2	119, 29	2	193	1
12, 97	2	126, 17	1	199, 75	1
33, 16	2	132, 68	1	200, 63	1
40, 3	2	132, 74	2	204, 51	2
50, 36	2	135, 65	1	213, 63	1
50, 7	1	135, 91	1	221, 72	1
54, 1	1	141, 8	1	225, 93	1
62, 06	2	141, 99	2	228, 03	1
67, 7	2	149, 78	2	232, 69	1
80, 97	2	160, 1	1	237, 5	2
94, 04	1	164, 87	1	254, 47	1
107, 77	2	166, 29	1	265, 07	1
108, 63	1	169, 76	1	279, 74	1
112, 2	1	173, 69	1	353, 9	1
112, 43	1	175, 41	1		

TABELA A.5: Dados simulados segundo o modelo Weibull (3.3) e $n = 100$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
0,27	2	108,72	2	173,69	1
0,36	2	112,2	1	175,41	1
0,62	2	112,43	1	177,8	1
2,84	2	113,38	1	178,47	1
6,31	2	115,57	1	185,71	1
12,54	2	116,42	1	186,19	2
12,97	2	118,43	1	189,28	1
22,08	2	119,26	1	189,37	1
33,16	2	119,29	2	193	1
37,53	2	119,86	2	196,9	2
39,47	2	125,54	2	199,75	1
39,52	1	126,17	1	200,63	1
40,3	2	128,05	1	204,51	2
50,36	2	129,29	1	213,63	1
50,7	1	129,37	2	221,72	1
54,1	1	131,65	2	224,42	1
55,16	1	132,68	1	225,93	1
59,91	2	132,74	2	228,03	1
60,91	1	133,37	1	232,69	1
62,06	2	135,65	1	237,5	2
62,72	1	135,91	1	250,61	1
64,95	2	141,8	1	254,47	1
67,7	2	141,99	2	258,65	1
69,29	1	147,35	1	262,71	1
70,41	2	149,78	2	265,07	1
71,3	1	149,8	2	266,97	1

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
79,24	1	153,41	1	272,2	2
80,03	1	156,85	1	279,68	1
80,97	2	160,1	1	279,74	1
94,04	1	164,87	1	289,36	2
99,88	1	165,82	1	349,57	1
100,29	1	166,29	1	353,9	1
107,77	2	166,41	1		
108,63	1	169,76	1		

A.3 Dados do Estudo de simulação 2

TABELA A.6: Dados simulados segundo um modelo Gama e $n = 10$.

Tempo	Fator de Risco	Tempo	Fator de Risco
18,06	2	19,61	1
19,04	2	18,86	2
15,48	2	14,46	2
18,49	2	17,42	2
26,18	2	18,68	2

TABELA A.7: Dados simulados segundo um modelo Gama e $n = 25$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
18.06	2	18.68	2	24.12	2
19.04	2	16.27	1	25.00	2
15.48	2	24.85	2	13.81	1
18.49	2	17.51	2	19.48	1
26.18	2	21.15	2	24.43	2
19.61	1	13.14	1	10.60	1
18.86	2	22.64	1	21.59	2
14.46	2	18.61	1		
17.42	2	20.50	2		

TABELA A.8: Dados simulados segundo um modelo Gama e $n = 50$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
10,6	1	17,51	2	20,57	1
11,21	1	17,59	1	21,15	2
13	1	18,06	2	21,47	2
13,14	1	18,1	2	21,59	2
13,81	1	18,2	2	21,63	2
14,12	1	18,49	2	22,64	1
14,33	1	18,61	1	23,61	1
14,46	2	18,68	2	24,12	2
15,17	2	18,86	2	24,43	2
15,43	2	18,89	2	24,49	2
15,48	2	19,04	2	24,85	2
15,53	2	19,48	1	24,99	1
15,77	2	19,61	1	25	2
16,27	1	19,67	2	25,13	2
16,41	2	19,99	2	26,18	2
17,04	2	20,23	1	29,95	2
17,42	2	20,5	2		

TABELA A.9: Dados simulados segundo um modelo Gama e $n = 100$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
10,33	2	17,95	1	20,33	2
10,6	1	17,99	1	20,36	2
11,21	1	18,06	2	20,38	2
11,4	2	18,1	2	20,5	2
12,02	2	18,17	1	20,57	1
12,27	2	18,2	2	20,58	1
13	1	18,45	2	20,98	1
13,14	1	18,47	2	21	1
13,63	1	18,49	2	21,15	2
13,81	1	18,52	1	21,47	2
13,86	1	18,59	1	21,59	2
14,12	1	18,61	1	21,63	2
14,33	1	18,63	2	21,84	2
14,46	2	18,65	2	22,1	2
14,6	2	18,68	2	22,64	1
14,64	2	18,86	2	22,77	1
15,17	2	18,89	2	23,5	2
15,43	2	19,04	2	23,61	1
15,48	2	19,04	2	24,01	2
15,53	2	19,15	2	24,06	2
15,71	2	19,27	2	24,12	2
15,77	2	19,36	2	24,16	2
16,19	2	19,46	2	24,43	2
16,19	0	19,48	1	24,49	2
16,27	1	19,61	1	24,85	2
16,41	2	19,67	2	24,99	1

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
16,59	2	19,88	2	25	2
16,67	2	19,99	2	25,13	2
17,04	2	20,02	1	25,68	2
17,23	2	20,06	2	26,18	2
17,41	1	20,09	2	26,47	2
17,42	2	20,19	1	29,95	2
17,51	2	20,23	1		
17,59	1	20,25	2		

A.4 Dados do Estudo de simulação 3

TABELA A.10: Dados simulados segundo o modelo Weibull (3.3) tomando $\beta_1 = 2$, $\eta_1 = 4$, $\beta_2 = 3$ e $\eta_2 = 5$.

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
2.83	1	2.84	2	3.12	1
2.65	2	5.55	1	2.99	2
3.47	2	4.99	2	2.80	1
2.41	1	2.11	2	3.65	1
5.35	1	2.04	1	3.13	1
5.64	2	2.23	1	1.09	2
5.22	2	3.06	2	1.67	1
2.02	1	4.33	1	1.35	1
2.57	2	0.81	1	1.03	1
1.29	1	2.86	1	4.49	1
2.49	1	1.32	1	0.91	1
2.40	1	2.21	1	1.57	2
1.01	1	1.55	2	7.39	2
6.49	1	1.46	1	3.22	2

Tempo	Fator de Risco	Tempo	Fator de Risco	Tempo	Fator de Risco
1.78	1	4.69	2	2.65	1
3.27	1	3.70	2	2.73	1
3.56	1	2.69	2	2.21	1
6.01	2	3.11	2	0.69	1
3.33	1	2.30	2	3.38	2
4.38	2	0.19	1	1.53	2
4.34	1	2.67	1	3.28	2
4.74	2	5.64	2	3.51	1
3.09	1	4.01	1	1.49	2
3.74	1	4.26	1	2.15	1
4.92	1	2.99	1	1.48	1
4.39	2	2.43	1	1.35	1
2.79	1	3.25	2	4.66	1
2.66	1	1.65	1	0.27	2
5.07	2	2.51	1	1.55	1
4.45	1	1.06	2	2.61	1
4.15	1	2.83	2	4.74	1
8.52	2	1.50	1	2.85	1
0.99	1	3.19	1		
0.53	1	4.60	1		
