**Confiabilidade em sistemas coerentes:
um modelo bayesiano Weibull**

Felipe Lunardi Bhering

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

Programa: Estatística

Orientador: Prof. Dr. Carlos Alberto de Bragança Pereira / IME - USP

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq

São Paulo, maio de 2013

Confiabilidade em sistemas coerentes: um modelo bayesiano Weibull

Esta dissertação trata-se da versão original
do aluno (Felipe Lunardi Bhering).

Confiabilidade em sistemas coerentes: um modelo bayesiano Weibull

Comissão Julgadora:

- Prof. Dr. Carlos Alberto de Bragança Pereira (Presidente) / IME - USP
- Prof. Dr. Adriano Polpo de Campos / UFSCar
- Prof. Dr. Enrico Antônio Colosimo / UFMG

Agradecimentos

Agradeço a todas pessoas e coisas que passaram de alguma forma na minha vida fizeram grandes contribuições para a realização desse trabalho. Priorizarei os últimos anos que tonaram isso possível e que concretizaram na escrita dessa dissertação. Agradeço a natureza, a sua força, magnitude e principalmente a benção de sua vida. Aos meus pais Denise e Nilton por terem me criado e dado apoio desde e durante a minha gestação. Agradeço a minha família, avós, tios e primos. Agradeço a Danila por todos ensinamentos, companheirismo, crescimento pessoal. Agradeço ao Professor Carlinhos e o Prof. Adriano pelas sugestões e dicas de caminhos que eles já trilharam e com excelência, pela liberdade que me deram, a tranquilidade que possuem e que geram nas pessoas. A escrita desse trabalho foi imensamente dada pela contribuição de sugestões dadas por eles como sugestões práticas da real necessidade do que deve ser escrito. O Prof. Adriano merece e muito grande significância e menção por tudo aqui criado e construído. Agradeço sua paciência e simplicidade e tudo que me ensinou durante todo esse tempo. Agradeço à Danila, Natália e Denise por lerem o texto inicial e darem sugestões. Grandes marcos na formação do mestrado foram dados pelo Prof. Luis Gustavo Esteves, sua forma simples de passar o conteúdo que já desde a graduação foi fundamental. A forma de pensar e ampliação de conceitos dada pelo Prof. Carlos Pereira e Prof. Sergio Wechsler nas disciplinas do mestrado. A formação de probabilidade dada pelo Prof. Adilson na disciplina de Probabilidade Avançada I e da Prof. Vera Fava na disciplina de Econometria da FEA que ampliou mais ainda a parte prática de conhecimentos estatísticos. Aos conhecimentos computacionais passados pela Prof. Florencia Leonardi essenciais nos dias de hoje a estatística. As discussões produtivas feitas com Manuel González, Gabriel Peixoto, Nathália Demétrico, Gustavo Miranda, Alexandre Patriota, Mel Oliveira, Andrés Rosso, Daniel de Brito Reis, William Duran, Andrés Arias Rodríguez, Eliardo nas disciplinas do mestrado. A ligação, congressos, vivências e tudo que o grupo que iniciou junto comigo o mestrado gerou Brian Melo, Raony Cesar, Bruno de Castro, Agatha Rodrigues, Carlos Relvas, Renata e Soane Mota. Agradeço, também e cito outros que tiveram presentes nesse ciclo Leandro Correia, Lina Thomas, Douglas Pinto, Márcio Diniz, Victor Cokin, Thiago Feitosa, Iesus Diniz, Victor Fossaluzza, Tiago Magalhães e Aníbal Emiliano. A contribuição grande a minha formação continuada e amizade de Evandro Makiyama. Não fosse pelo aprendizado tido com Gisela Tunes da Silva e Antonio Carlos Pedroso de Lima nas áreas de sobrevivência e seus ensinamentos de forma direta e bem formulada esse trabalho teria sido muito mais complicado de se realizar. O congresso EBEB Bayesiano e o comitê organizador e executivo que foi fundamental para a formação que tive nesses últimos tempos. A todos funcionários do IME, da secretaria, da biblioteca, faxineiros e guardas que possibilitam o funcionamento do Instituto. A amigos próximos nesse período Carlos Kozera e Marcela Ribeiro. Aos mestres presentes durante todo esse processo. A todos que fizeram parte de alguma forma e que não mencionei em meus agradecimentos.

Resumo

O principal objetivo desse trabalho é introduzir um modelo geral bayesiano Weibull hierárquico para dados censurados que estima a função de confiabilidade de cada componente para sistemas de confiabilidade coerentes. São introduzidos formas de estimação mais sólidas, sem a inserção de estimativas médias nas funções de confiabilidade (estimador plug-in). Através desse modelo, são expostos e solucionados exemplos na área de confiabilidade como sistemas em série, sistemas em paralelo, sistemas k-de-n, sistemas bridge e um estudo clínico com dados censurados intervalares. As soluções consideram que as componentes tem diferentes distribuições, e nesse caso, o sistema bridge ainda não havia solução na literatura. O modelo construído é geral e pode ser utilizado para qualquer sistema coerente e não apenas para dados da área de confiabilidade, como também na área de sobrevivência, dentre outros. Diversas simulações com componentes com diferentes proporções de censura, distintas médias, três tipos de distribuições e tamanhos de amostra foram feitas em todos os sistemas para avaliar a eficácia do modelo.

Palavras-chave: Estatística Bayesiana, Modelos Hierárquicos, Dados Censurados, Confiabilidade, Análise de Sobrevivência, Programação em R, Simulações, Métodos de Monte Carlo, MCMC, sistemas em série, sistemas em paralelo, censura à direita, censura à esquerda, censura intervalar, algoritmo EM, sistemas bridge, sistemas redundante, distribuição Weibull.

Abstract

The main purpose of this work is to introduce a general bayesian Weibull hierarchical model for censored data which estimates each reliability component's function from coherent systems. It's introduced estimation procedures which do not consider plug-in estimators. Also, it's exposed and solved with this model examples in reliability area such as series systems, parallel systems, k-out-of-n systems, bridge systems and a clinical study with interval censoring data. The problem of bridge system hadn't a solution before for the case of each component with different distribution. Actually, this model is general and can be used to analyse any kind of coherent system and censored data, not only reliability ones, but also survival data and others. Several components simulations with different censored proportions, distinct means, three kinds of distributions and sample size were made in all systems to evaluate model efficiency.

Keywords: Bayesian Statistics, Hierarchical Models, Censored Data, Reliability, Survival Analysis, R programming, Simulation, Monte Carlo Methods, MCMC, series systems, parallel systems, right censoring, left censoring, interval censoring, EM algorithm, bridge systems, redundant systems, Weibull distribution.

Sumário

Lista de Abreviaturas	ix
Lista de Símbolos	xi
Lista de Figuras	xiii
Lista de Tabelas	xv
1 Introdução	1
1.1 Objetivos e contribuições	1
1.2 Organização do Trabalho	1
2 Problemática, a censura e exemplos	3
2.1 Dados censurados	3
2.2 Confiabilidade e exemplos	4
3 Resolução e simulações	7
3.1 Sistemas em Série e em Paralelo	7
3.1.1 A modelagem	7
3.1.2 Simulações	12
3.2 Sistemas k-de-n	16
3.2.1 Modelo bayesiano hierárquico Weibull com censura intervalar	16
3.2.2 Posteriori própria	17
3.2.3 Simulações	20
3.3 Sistemas Ponte (Bridge)	21
3.3.1 Simulações	22
3.4 Tempo de retratação de acordo com dois tratamentos	25
3.4.1 Resolução	26
4 Conclusões	29
4.1 Considerações Finais e pesquisas futuras	29

Lista de Abreviaturas

<i>MLE</i>	estimador de máxima-verossimilhança
<i>BPE</i>	estimador plug-in bayesiano
<i>HPD</i>	intervalo Highest posterior density - intervalo de credibilidade
<i>MCMC</i>	procedimento de simulação Markov Chain Monte Carlo
<i>EM</i>	procedimento Expectation-maximization (esperança-maximização) com maior densidade de probabilidade
<i>EQM</i>	Erro Quadrático Médio
<i>SiPS</i>	Sistema paralelo-série
<i>SiSP</i>	Sistema série-paralelo

Lista de Símbolos

$k - de - n$	sistema que funciona apenas se pelo menos k componentes dos n estiverem funcionando
$2 - de - 3$	sistema que funciona apenas se pelo menos 2 componentes dos 3 estiverem funcionando
$(X_j)_{j \in \{1, \dots, k\}}$	equivalente à $(X_1, \dots, X_k)$
X_j	tempo de falha da componente j do sistema
$\min(X_1, \dots, X_k)$	menor valor dentre $(X_1, \dots, X_k)$
$\max(X_1, \dots, X_k)$	menor valor dentre $(X_1, \dots, X_k)$
$(\delta_j)_{j \in \{1, \dots, k\}}$	vetor equivalente à $(\delta_1, \dots, \delta_k)$
(δ_j)	indica qual componente gerou a falha $\delta = j$ se $T = X_j$ e 0 caso contrário
$f()$	função densidade
$R()$	função confiabilidade
θ	parâmetros do modelo a serem estimados
$\pi(\theta t, \delta)$	distribuição a posteriori
$\pi(\theta)$	distribuição a priori
(T, δ)	vetor de tempos de falha e vetor que indica de qual componente gerou a falha
$\mathbf{t}$	vetor $(t_1, \dots, t_n)$ de tempos observados
δ	vetor de variáveis que indicam qual componente gerou a falha
I_A	indicadora que tem o valor 1 caso ocorra o evento A e 0 caso contrário
$\theta = (\theta_1, \dots, \theta_k)$	vetor de parâmetros theta de todas componentes
$\theta_j = (\beta_j, \eta_j)$	parâmetros do modelo Weibull para a componente j
$\Gamma()$	função matemática gamma
m_β	média da distribuição do parâmetro β
m_η	média da distribuição do parâmetro η
v_β	variância da distribuição do parâmetro β
v_η	variância da distribuição do parâmetro η
$\propto$	tem distribuição (a menos de constante) do valor colocado a direita do símbolo
$\mathbf{m}_\beta$	vetor de média de beta respectivamente, de todas componentes
$\mathbf{m}_\eta$	vetor de média de eta de todas componentes
$\mathbf{v}_\beta$	vetor de variâncias de beta de todas componentes
$\mathbf{m}_\eta$	vetor de variâncias de eta de todas componentes

$\sim$	segue a distribuição indicada à direita do símbolo
n_p	tamanho da amostra
$E()$	função de esperança matemática
$Y()$	amostra da posteriori da função de confiabilidade da componente 1 do sistema
$\widehat{m}_\beta$	média estimada da distribuição do parâmetro β
$\widehat{m}_\eta$	média estimada da distribuição do parâmetro η
$\pi_j^b()$	classe de prioris não-informativas
$x+$	dado censurado à direita de x
$(x; y]$	intervalo aberto em x fechado em y

Lista de Figuras

2.1	Representação gráfica de um sistema de n componentes em série.	4
2.2	Representação gráfica de um sistema de n componentes em paralelo.	5
2.3	Representação gráfica de um sistema com estrutura 2-de-3.	5
2.4	Representação gráfica de Ponte de Wheatstone (bridge).	6
3.1	Representação gráfica da distribuição Weibull com diferentes parâmetros.	8
3.2	Funções de confiabilidade das componentes do sistema em série: (a) Componente 1; (b) Componente 2; (c) Componente 3.	13
3.3	Funções de confiabilidade das componentes do sistema em paralelo: (a) Componente 1; (b) Componente 2; (c) Componente 3.	14
3.4	Representação gráfica de um sistema com estrutura 2-de-3. (a) Sistema paralelo-série (SiPS); (b) Sistema série-paralelo (SiSP).	16
3.5	Funções de confiabilidade das componentes do sistema 2-de-3: (a) Componente 1; (b) Componente 2; (c) Componente 3.	21
3.6	Funções de confiabilidade das componentes 1, 2 e 3 do sistema ponte (bridge): (a) Componente 1; (b) Componente 2; (c) Componente 3.	23
3.7	Funções de confiabilidade das componentes 4 e 5 do sistema ponte (bridge): (a) Componente 4; (b) Componente 5.	23
3.8	Funções de confiabilidade do estudo para os dois tratamentos: (a) Radioterapia; (b) Radioterapia e Quimioterapia.	27

Lista de Tabelas

3.1	Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades da simulação 1 para os sistemas em série e paralelo.	13
3.2	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 1 que possui distribuição Weibull e com censura à direita (sistema em série - simulação 2).	14
3.3	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 2 que possui distribuição gamma e com censura à direita (sistema em série - simulação 2).	15
3.4	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 3 que possui distribuição log-normal e com censura à direita (sistema em série - simulação 2).	15
3.5	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 1 que possui distribuição Weibull e com censura à esquerda (sistema em paralelo - simulação 2).	15
3.6	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 2 que possui distribuição gamma e com censura à esquerda (sistema em paralelo - simulação 2).	15
3.7	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 3 que possui distribuição log-normal e com censura à esquerda (sistema em paralelo - simulação 2).	16
3.8	Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades da simulação 1 para o sistema 2-de-3.	21
3.9	Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades da simulação 1 para o sistema ponte (bridge).	22
3.10	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 1 que possui distribuição Weibull (simulação 2 -sistemas k-de-n e ponte (bridge)).	24
3.11	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 2 que possui distribuição gamma (simulação 2 -sistemas k-de-n e ponte (bridge)).	24
3.12	Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 3 que possui distribuição log-normal (simulação 2 -sistemas k-de-n e ponte (bridge)).	25

3.13	Intervalo de retratação para indivíduos com quadro recente de câncer de mama tratados com radioterapia apenas vs. radioterapia e quimioterapia.	26
3.14	Proporções de censura para o sistema em estudo com dois tratamentos.	26
3.15	Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades o estudo com dois tratamentos.	26

Capítulo 1

Introdução

“Man starts over again everyday, in spite of all he knows, against all he knows.” (Emil Cioran)

1.1 Objetivos e contribuições

A literatura que trata da estimação da confiabilidade de sistemas em série ou riscos competitivos é antiga e longa. O trabalho mais citado é o de ?, na qual foi desenvolvido um estimador não-paramétrico frequentista. Para uma visão bayesiana podem ser citados os artigos de ? e ?. Tudo indica que o artigo ? foi o primeiro a tratar a estimação bayesiana não-paramétrica de confiabilidade para sistemas em paralelo. Através de uma estrutura computacional mais complexa, o artigo ? utiliza prioris não-informativas e apresenta a estimação bayesiana de confiabilidade com modelos Weibull. A estimação foi feita através de modelos de Monte Carlos e o algoritmo apresentou dificuldades para convergência. Por isto, no presente trabalho utilizou-se o contexto de modelos hierárquicos com prioris não-informativas para contornar essa situação.

Dessa forma, este trabalho tem como objetivo introduzir o leitor um modelo geral bayesiano Weibull hierárquico para dados censurados. Esse modelo serve para estimar a sobrevivência em sistemas coerentes. Modelos gerais bayesianos e que considerem, todas formas de dados censurados, inclusive a censura intervalar não são encontrados na literatura. São introduzidos, também, formas de estimação mais sólidos, sem a utilização de aproximações frequentemente utilizadas como a inserção de estimativas médias nas funções de confiabilidade (estimador plug-in).

O artigo ? realizou uma simulação de três métodos distintos de estimação de confiabilidade. Comparou-se o estimador de ?, o estimador de máxima verossimilhança (MLE) e o estimador bayesiano Plug-in (BPE) para sistemas de confiabilidade com distribuição Weibull para todos os componentes. Os resultados indicam que MLE e BPE são similares e tem performance superior ao estimador Kaplan-Meier. No entanto, não foi construído intervalos de credibilidade neste artigo. Nesta dissertação, são construídas curvas de sobrevivência com intervalos de credibilidade HPD.

No presente trabalho, similarmente à ? e ? considerou-se um modelo com distribuição Weibull para a confiabilidade. Utilizou-se um modelo hierárquico com dois níveis, no primeiro nível com distribuições gamma independentes. Essas distribuições gammas foram reparametrizados para ter como parâmetros média e variância. Foram fixadas as variâncias e as médias definidas como prioris vagas (flat) para o segundo nível da hierarquia.

1.2 Organização do Trabalho

Para avaliação do modelo foram considerados alguns exemplos expressos no Capítulo 2. Os primeiros dois exemplos, introduzem a noção de sistemas em série e em paralelo. Sistemas em série são circuitos que estão funcionando se todos os componentes estão funcionando e sistemas em paralelo funcionam até que todos os componentes falhem. Outro exemplo exposto foi o de sistemas k-de-n, sistemas que funcionam se pelo menos k componentes estão funcionando. Um outro exemplo foi o de sistemas em ponte (bridge) e foi construída uma solução para o caso de variáveis com

diferentes distribuições, um problema que ainda não tinha uma solução na literatura. Finalmente, considerou-se, um exemplo em que simultaneamente ocorre censura intervalar, censura à direita e censura à esquerda. Todos os exemplos foram solucionados no Capítulo 3. Neste capítulo também descreve-se a modelagem utilizada para a resolução dos exemplos. Finalmente, no Capítulo 4 são dadas as considerações finais do trabalho e possíveis continuações.

Capítulo 2

Problemática, a censura e exemplos

“Onde existe vontade, existe realização.” (provérbio chinês)

Esse capítulo introduz a noção de dados censurados e apresenta exemplos de circuitos na área de confiabilidade que são bastante utilizados e importantes do ponto de vista tanto teórico como prático.

2.1 Dados censurados

Uma censura estatística pode ocorrer por diversos motivos. O truncamento ocorre quando escolhe-se um momento de parada. Considere uma experiência na qual verifica-se o tempo cronológico de funcionamento de uma lâmpada e esta experiência é finalizada após 10 dias. As lâmpadas que continuarem funcionando após esse tempo são consideradas como censura (à direita), e supõe-se que elas iriam quebrar em um dado momento se ela fosse observada por um grande período de tempo. Uma das lâmpadas já estava quebrada já no início do experimento, nesse caso temos uma censura à esquerda, ou seja, o evento já aconteceu quando a lâmpada foi observada. Ao mesmo tempo, se este experimento fosse observado em períodos de tempo partidos e não constantemente, e uma lâmpada parar de funcionar em um desses períodos, apenas temos a informação de que ela quebrou entre este momento de observação e o anterior. Isto corresponde à uma censura intervalar.

Dados perdidos são outro caso de censura. Ao analisar a sobrevivência de pacientes com uma doença, um deles não retornou ao tratamento e o hospital perdeu o seu contato, dessa forma, ele é censurado (à direita), ou seja, constata-se que ele, até a última data disponível estava vivo.

Outro motivo de censura ocorre quando o equipamento ou a forma como mede-se os dados não tem capacidade de medir o caso. Em um experimento que testa-se a resistência de um material, alguns dos produtos não chegaram a quebrar, pelo equipamento não ter criado a resistência necessária para a quebra do produto.

Dessa forma existem 3 tipos de censura: à esquerda, intervalar, e à direita. Por motivos financeiros, a medição de experimentos, muitas vezes é feita, de tempos em tempos, ou seja, em intervalos, e mesmo assim, a censura intervalar não é utilizada, e considera-se o tempo em que ocorreu o evento de interesse como a primeira medição feita em que verificou-se o evento.

A área da estatística é muito vasta e é aplicada em diversas áreas da ciência. Os dados censurados, também, são aplicados em diversas áreas e tem nomes distintos. A análise sobrevivência lida com o ramo de mortes de organismos biológicos ou falha em sistemas mecânicos, já na engenharia, a área da estatística que lida com dados censurados é a confiabilidade, na economia chama-se análise de duração, e, finalmente, na sociologia chama-se análise de eventos históricos. O enfoque deste trabalho e os exemplos, a partir desse momento, serão dados na área de confiabilidade.

2.2 Confiabilidade e exemplos

A área de confiabilidade lida com questões como a manutenção de máquinas e o controle de qualidade de produtos. Ela cresceu bastante pela demanda tecnológica por produtos de maior qualidade e com um padrão mais rigoroso. Também, através das guerras militares como a segunda guerra mundial no final da década de 1930. Na década de 50, cresceu o ramo dito análise de sobrevivência, ramo similar a confiabilidade, só que lida especificamente com a morte de organismos biológicos.

No presente trabalho, cada sistema e cada componente do sistema só tem duas possibilidades, ou está funcionando ou não está, já as variáveis que medem o tempo de funcionamento dessas variáveis binárias são contínuas, assumindo qualquer valor real no intervalo dado.

Nesse trabalho, todos os sistemas utilizados são ditos coerentes. Um sistema é dito coerente se cada um dos seus componentes é relevante e sua função de estrutura é monótona. Isso significa excluir a possibilidade de um conserto de um componente pode gerar falha de um sistema. Para mais detalhes sobre sistemas coerentes consultar o livro ?. Os sistemas k-de-n, os sistemas em paralelo e em série e o sistema bridge são casos particulares dos sistemas coerentes.

Dois conceitos chave na área de confiabilidade são sistemas em série e em paralelo e um importante resultado é que todo sistema coerente pode ser escrito como combinações de sistema em série e em paralelo.

Exemplo 2.2.1 *Um sistema em série com n componentes é um sistema que está funcionando apenas se todos os seus componentes estão funcionando, logo, a sua falha ocorre quando o primeiro dos componentes falhe. Sua representação gráfica é dada a seguir.*

Figura 2.1: Representação gráfica de um sistema de n componentes em série.

As redes de luzes de natal que piscam são um exemplo desse sistema, quando uma lâmpada delas queima, todas param de funcionar. Atualmente existem já algumas que não funcionam dessa forma. A informação que se tem é de que um componente falhou e os demais ainda não falharam, logo os demais componentes virão a falhar em um dado momento e possuem falha de valor maior do que o primeiro componente. Assim, ocorre a censura à direita para todos os componentes exceto o primeiro que falhou e gerou a quebra do sistema.

Exemplo 2.2.2 *Já um sistema em paralelo com n componentes é um sistema que é necessário que apenas um de seus componentes esteja funcionando para que toda a estrutura funcione, ou seja, a falha do sistema ocorre quando todos os componentes falham, visto de outra forma, até que o último componente falhe. Sua representação é dada a seguir.*

Figura 2.2: Representação gráfica de um sistema de n componentes em paralelo.

Um exemplo desse sistema em paralelo é quando uma casa possui um interruptor que liga mais de uma lâmpada. Assim, quando uma delas queima, as outras continuam a funcionar. Quando ocorre a quebra desse sistema, ou seja, todos componentes falham, em geral, temos apenas o tempo de falha do último componente (lâmpada). Todos outros componentes falharam em algum instante anterior a esse momento, ou seja ocorrendo censura à esquerda.

O conceito de redundância na engenharia tem intenção de aumentar a confiabilidade e segurança de um sistema. Ele é muito utilizado em sistemas de controle por computador e sistemas hidráulicos de aeronaves. Também, em sistemas de segurança como alarmes de incêndio e em sistemas de comando de submarinos.

Exemplo 2.2.3 Uma estrutura de um hardware (equipamento) que funciona caso 2 dos 3 componentes esteja funcionando é dito uma estrutura 2-de-3 ou redundante triplo-modular (TMR). A sua representação usual é dada por,

Figura 2.3: Representação gráfica de um sistema com estrutura 2-de-3.

A Figura 2.3 é meramente ilustrativa e para fins de análise, pois em cada sistema só existe uma única vez cada componente. Os computadores de sistema do Boeing 777 são desse tipo, dessa forma, existem 3 computadores, sendo que para funcionar pelo menos 2 necessitam estar funcionando. Cada um dos 3 computadores também é redundante triplo-modular, ou seja, cada um dos computadores tem canais interconectados utilizando 3 diferentes processadores (diferentes marcas), tais que apenas 2 processadores necessitam estar funcionando para que um computador funcione (?).

De um modo geral, um sistema com estrutura k -de- n (k e n inteiros com $k \leq n$) funciona se e somente se pelo menos k dos n componentes está funcionando.

Exemplo 2.2.4 Um circuito elétrico chamado de ponte (bridge) é um circuito onde existe uma ponte entre dois ramos da estrutura se interconectando. O exemplo, a seguir, e uma de suas representações, é dito Ponte de Wheatstone e tem esse nome devido à contribuição de Charles Wheatstone de popularizar o seu uso.

Figura 2.4: Representação gráfica de Ponte de Wheatstone (bridge).

As pontes de Wheatstone são muito utilizadas para medir resistências e voltagens. Já as pontes, de forma mais geral, possuem diversas aplicações como em filtro eletrônico e conversão de energia elétrica, dentre muitas outras. Algumas podem ser vistas em ?.

Exemplo 2.2.5 *Dados censurados intervalares são comuns em situações em que ocorre uma visita periódica na qual acompanha algum determinado evento do tipo sucesso-fracasso. Isso ocorre, por exemplo, na avaliação da eficácia de um tratamento para uma doença. Isso ocorre com frequência em consultas nas áreas biomédicas.*

Capítulo 3

Resolução e simulações

“Só percebemos o valor da água depois que a fonte seca.” (provérbio brasileiro)

Esse capítulo retoma os exemplos dados no capítulo anterior, resolvendo-os através do método proposto, um modelo bayesiano hierárquico geral para quaisquer dados censurados, à direita como à esquerda e intervalar, podendo inclusive possuir ao mesmo tempo diferentes censuras. Os detalhes para cada exemplos estão nas seções a seguir.

3.1 Sistemas em Série e em Paralelo

Uma versão inicial dessa seção foi publicada no artigo ?. Nessa seção serão discutidos e dadas formas de resolução para os sistemas em série e em paralelo, descritos no Exemplo 2.2.1 e o Exemplo 2.2.2. Para tal introduziremos conceitos matemáticos.

Desse modo, define-se a sequência de variáveis aleatórias $(X_j)_{j \in \{1, \dots, k\}} = (X_1, \dots, X_k)$, k inteiro. Cada componente (X_j) é uma variável aleatória independente que denota os tempo de falha (ou morte) da componente j de um sistema.

Nestes dois casos, conseguimos observar as funções $T = \min(X_1, \dots, X_k)$ para o sistema em série e $T = \max(X_1, \dots, X_k)$ para o sistema em paralelo. Um sistema com 3 componentes em série que falhou após 9 dias nos mostra que o primeiro componente, com valor menor (mínimo) de tempo de falha comparando-o com os demais, falhou em 9 dias. Já um sistema em paralelo com 3 componentes e tempo de falha de 10 dias indica que todos componentes falharam até 10 dias, sendo que a última, com maior tempo de falha (máximo), falhou em 10 dias.

Define-se, também, a sequência $\delta = j$ se $T = X_j$ e 0 caso contrário, para $j = 1, \dots, k$, assim, a sequência de $(\delta_j)_{j \in \{1, \dots, k\}}$ tem valor igual a j no componente com o mesmo valor de falha e 0 nos demais itens. No exemplo anterior se o componente 1 foi o primeiro componente a falhar em 9 dias, então essa sequência é dada por $(1, 0, 0)$. E se no segundo caso, caso o componente 2 tenha sido o último componente a falhar em 10 dias, a sequência é dada por $(0, 2, 0)$. Essa sequência indica de qual componente causou a falha do sistema.

3.1.1 A modelagem

A distribuição Weibull, que será apresentada, foi utilizada por sua adaptabilidade à diversos contextos e situações. Através de simulações com distribuições log-normal, gamma e weibull mostraremos essa qualidade dela. Essas distribuições são as mais utilizadas na área de confiabilidade e análise de sobrevivência.

Como dito no capítulo 2, os sistemas em série e em paralelo, tem, respectivamente, dados censurados à direita e à esquerda, desse modo, a nomenclatura utilizada a seguir será essa.

Dessa forma, essa seção apresentará um modelo de distribuição weibull para dados censurados à direita e à esquerda, bem como sua estimação e resultados.

3.1.1.1 Modelo Weibull

Por volta da década de 50, a distribuição mais utilizada nessa área era a exponencial, pela facilidade dos cálculos e qualidade dos seus resultados. Já no final dessa década, iniciou-se a utilização acentuada da distribuição Weibull, uma generalização da distribuição exponencial.

Desse modo, define-se para cada tempo de falha $X_j, j \in 1, \dots, k$, ter distribuição Weibull de parâmetro $\theta = (\beta, \eta)$ que possui função de densidade de probabilidade dada por

$$f(x, \beta, \eta) = \frac{\beta}{\eta} \left(\frac{x}{\eta} \right)^{\beta-1} \exp \left(-\frac{x}{\eta} \right)^{\beta}, x \geq 0, \beta > 0, \eta > 0 \quad (3.1)$$

Utilizaremos muito a função dita de confiabilidade ou sobrevivência da distribuição weibull, que significa a probabilidade da falha ou morte de um componente ser maior que um certo valor. A expressão é mostrada abaixo,

$$R_X(x) = P(X > x | (\beta, \eta)) = \exp \left(-\frac{x}{\eta} \right)^{\beta}, x \geq 0, \beta > 0, \eta > 0 \quad (3.2)$$

para parâmetros (β, η) .

Alterando os parâmetros da Weibull ela assume diversos comportamentos, a Figura 3.1 nos mostra diferentes parâmetros e representações gráficas da distribuição Weibull. Através dele, percebe-se que ela é bem flexível, dispondo-se de maneira bem diferente conforme muda-se os parâmetros. Quando o parâmetro dela $\beta = 1$ temos como caso particular a tão conhecida distribuição exponencial.

Figura 3.1: Representação gráfica da distribuição Weibull com diferentes parâmetros.

3.1.1.2 Função de Verossimilhança

A função de verossimilhança é uma função dos parâmetros de um modelo estatístico dado os dados observados existentes na amostra obtida. Através dela são feitas estimações dos parâmetros do modelo.

Considere um sistema com k componentes e seja $(X_j)_{j \in \{1, \dots, k\}}$ a sequência de variáveis aleatórias independentes que denotam os tempos de falha (ou morte) dos componentes. As funções $T = \min(X_1, \dots, X_k)$ para o sistema em série e $T = \max(X_1, \dots, X_k)$ para o sistema em paralelo são as funções observadas. Define-se que cada tempo de falha da variável aleatória X_j , com $j \in \{1, \dots, k\}$, como uma variável aleatória com distribuição Weibull de parâmetros $\theta_j = (\beta_j, \eta_j)$. Também, define-se a sequência $\delta = j$ se $T = X_j$ e 0 caso contrário, para $j = 1, \dots, k$, assim, a sequência de $(\delta_j)_{j \in \{1, \dots, k\}}$ tem valor igual a j no componente com o mesmo valor de falha e 0 nos demais itens. Além disso, temos uma amostra aleatória independente de n sistemas observados com a informação de qual foi o componente que gerou a falha do sistema, ou seja, $(T, \delta) = \{(T_i, \delta_i) : i = 1, \dots, n\}$.

Em um sistema em série temos a informação de que cada componente terá a possibilidade de: ou ter gerado a falha do sistema e dessa forma ter contribuído com o seu tempo de falha para a verossimilhança, ou termos a informação de que ele ainda não falhou e, portanto, o seu tempo de falha caso continuasse a funcionar é maior que o tempo de falha do sistema. Um sistema com 3 componentes em série que falhou após 9 dias nos mostra que o primeiro componente, com valor menor (mínimo) de tempo de falha comparando-o com os demais, falhou em 9 dias e os demais não falharam até 9 dias e se continuassem a funcionar falhariam após esse valor.

O vetor $\mathbf{t} = (t_1, \dots, t_n)$ é o vetor de observações independentes da amostra de n sistemas observado e $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$ ($\delta = j$ se $T = X_j$ e 0 caso contrário) indica qual foi o componente que gerou a falha do sistema. O vetor $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)$ denota o vetor dos parâmetros.

Assim, a função de verossimilhança para o sistema em série é dada por

$$\pi(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta}) \propto \prod_{j=1}^k \prod_{i=1}^n [f_j(t_i | \theta_j)]^{I_{\{\delta_i=j\}}} [R_j(t_i | \theta_j)]^{1-I_{\{\delta_i=j\}}} \quad (3.3)$$

$\delta_i = j$ indica a componente que gerou a falha do sistema e para essa componente temos a contribuição da densidade $f_j(t_i | \theta_j)$ e para as demais a contribuição do valor dado pela função de confiabilidade $R_j(t_i | \theta_j)$. A função I_A é a indicadora que vale 1 caso ocorra o evento A e 0 caso contrário.

Já um sistema em paralelo com 3 componentes e tempo de falha de 10 dias indica que todos componentes falharam em menos de 10 dias, sendo que a última, com maior tempo de falha, falhou em 10 dias.

Assim, a função de verossimilhança para o sistema em paralelo é dada por

$$\pi(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta}) \propto \prod_{j=1}^k \prod_{i=1}^n [f_j(t_i | \theta_j)]^{I_{\{\delta_i=j\}}} [1 - R_j(t_i | \theta_j)]^{1-I_{\{\delta_i=j\}}} \quad (3.4)$$

$\delta_i = j$ indica a componente que gerou a falha do sistema e para essa componente temos a contribuição da densidade $f_j(t_i | \theta_j)$ e para as demais a contribuição de $1 - R_j(t_i | \theta_j)$.

3.1.1.3 Modelo hierárquico

A área bayesiana cresceu significativamente após a década de 1970. Na década de 80, cresceu o uso dos modelos bayesianos hierárquicos. Da mesma forma que os parâmetros de uma distribuição são vistos como distribuições, isso pode ser feito em mais etapas. Os hiperparâmetros, ou seja, os parâmetros dos parâmetros das distribuições, são variáveis aleatórias com uma distribuição. Através desse método, a convergência de alguns modelos é mais rápida e os resultados convergem mais rapidamente para o valor exato. Os parâmetros, ou o primeiro nível de hierarquia, β_j e η_j

para $j = 1, \dots, n$ tiveram distribuições dadas pela distribuição gamma. A distribuição gamma de parâmetros (a, b) tem formulação matemática dada pela expressão abaixo:

$$f(x, a, b) = \frac{1}{(b^a \Gamma(a))} x^{a-1} \exp\left(-\frac{x}{b}\right), \quad x \geq 0 \quad (3.5)$$

com parâmetros forma (shape) $a > 0$ e escala (scale) $b > 0$.

A vantagem de usar a distribuição gamma é que a média dela é dada pelo valor ab e a variância (dispersão em torno da média) é dada por ab^2 . Desse modo, podemos reparametrizar a distribuição gamma com os parâmetros média e variância dela que são mais fáceis de tratar e interpretar. Através das equações 3.6 e 3.7 à seguir podemos reparametrizar a distribuição gamma em função dos parâmetros média e variância dela.

$$a = \frac{a^2 b^2}{ab^2} = \frac{\text{media}^2}{\text{variância}} \quad (3.6)$$

$$b = \frac{ab^2}{ab} = \frac{\text{variância}}{\text{media}} \quad (3.7)$$

Dessa forma, o segundo nível de hierarquia, os chamados hiperparâmetros, ou seja, parâmetros da distribuição dos parâmetros β_j e η_j podem ser reescritos pelos parâmetros média e variância de β_j e η_j denotados por $m_{\beta_j}, v_{\beta_j}, m_{\eta_j}$ e v_{η_j} . Foram dadas distribuições seguintes para os hiperparâmetros: $m_{\beta_j} \propto 1$ (priori vaga ou flat), $m_{\eta_j} \propto 1$ (priori vaga ou flat) e v_{β_j}, v_{η_j} valores constantes conhecidos e fixos.

Usaremos a notação de $\mathbf{m}_{\beta} = (m_{\beta_1}, \dots, m_{\beta_k})$, $\mathbf{v}_{\beta} = (v_{\beta_1}, \dots, v_{\beta_k})$, $\mathbf{m}_{\eta} = (m_{\eta_1}, \dots, m_{\eta_k})$, $\mathbf{v}_{\eta} = (v_{\eta_1}, \dots, v_{\eta_k})$ para denotar o vetor associado a essas distribuições. Todos esses valores são positivos.

Em um primeiro momento, o leitor pode achar estranho o porque usar uma priori vaga para média e fixar a variância e se isso gera alguma melhora para o algoritmo. Isso será melhor compreendido na seção seguinte que fala do algoritmo e da estimação. Esse método de hierarquia aumenta muito a velocidade como também a convergência do algoritmo que será descrito.

Em resumo, então, temos os seguinte níveis de hierarquia.

1o nível de hierarquia: $\beta_j \sim \text{gamma}(m_{\beta_j}, v_{\beta_j})$ e $\eta_j \sim \text{gamma}(m_{\eta_j}, v_{\eta_j})$

2o nível de hierarquia: $m_{\beta_j} \propto 1$, $m_{\eta_j} \propto 1$ e v_{β_j}, v_{η_j} valores conhecidos e fixos

para $j = 1, \dots, k$ distribuições independentes.

Considere $\zeta = (\mathbf{m}_{\beta}, \mathbf{v}_{\beta}, \mathbf{m}_{\eta}, \mathbf{v}_{\eta})$. A priori completa fica expressa por

$$\begin{aligned} \pi(\boldsymbol{\theta}, \zeta) &\propto \pi(\boldsymbol{\theta} | \mathbf{m}_{\beta}, \mathbf{v}_{\beta}, \mathbf{m}_{\eta}, \mathbf{v}_{\eta}) \pi(\mathbf{m}_{\beta}) \pi(\mathbf{v}_{\beta}) \pi(\mathbf{m}_{\eta}) \pi(\mathbf{v}_{\eta}) \\ &\propto \prod_{j=1}^k \pi(\theta_j | m_{\beta_j}, v_{\beta_j}, m_{\eta_j}, v_{\eta_j}) \pi(m_{\beta_j}) \pi(v_{\beta_j}) \pi(m_{\eta_j}) \pi(v_{\eta_j}) \end{aligned} \quad (3.8)$$

$$\propto \prod_{j=1}^k \frac{\beta_j^{m_{\beta_j}^2/v_{\beta_j}-1} \exp\{-m_{\beta_j} \beta_j / v_{\beta_j}\}}{(v_{\beta_j}/m_{\beta_j})^{m_{\beta_j}^2/v_{\beta_j}} \Gamma(m_{\beta_j}^2/v_{\beta_j})} \frac{\eta_j^{m_{\eta_j}^2/v_{\eta_j}-1} \exp\{-m_{\eta_j} \eta_j / v_{\eta_j}\}}{(v_{\eta_j}/m_{\eta_j})^{m_{\eta_j}^2/v_{\eta_j}} \Gamma(m_{\eta_j}^2/v_{\eta_j})} \quad (3.9)$$

Dessa forma, a posteriori para os sistemas em série e em paralelo, respectivamente, ficam dadas pelas expressões,

$$\pi(\boldsymbol{\theta}, \boldsymbol{\zeta} | \mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}, \boldsymbol{\zeta}) \prod_{j=1}^k \prod_{i=1}^n \left[\frac{t_i^{\beta_j-1} \beta_j}{\eta_j} \exp \left\{ - \left(\frac{t_i}{\eta_j} \right)^{\beta_j} \right\} \right]^{I_{\{\delta_i=j\}}} \left[\exp \left\{ - \left(\frac{t_i}{\eta_j} \right)^{\beta_j} \right\} \right]^{1-I_{\{\delta_i=j\}}}, \quad (3.10)$$

e

$$\pi(\boldsymbol{\theta}, \boldsymbol{\zeta} | \mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}, \boldsymbol{\zeta}) \prod_{j=1}^k \prod_{i=1}^n \left[\frac{t_i^{\beta_j-1} \beta_j}{\eta_j} \exp \left\{ - \left(\frac{t_i}{\eta_j} \right)^{\beta_j} \right\} \right]^{I_{\{\delta_i=j\}}} \left[1 - \exp \left\{ - \left(\frac{t_i}{\eta_j} \right)^{\beta_j} \right\} \right]^{1-I_{\{\delta_i=j\}}}, \quad (3.11)$$

em que $\mathbf{t} = (t_1, \dots, t_n)$ são os tempos de falha observados no sistema, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$ os indicadores de qual componente falhou no sistema, as outras componentes são as mesmas descritas anteriormente. Para chegar nas expressões foram aplicadas das fórmulas das distribuições as fórmulas de priori e posteriori.

Ressalta-se que a posteriori é própria. Esse resultado é essencial para a construção da posteriori. A versão mais geral desse teorema está na seção 3.2.2.

3.1.1.4 Algoritmo e estimação

Não existe forma fechada para a distribuição à posteriori acima, dessa forma é necessário usar um método de integração numérica ou algoritmos MCMC (Markov Chain Monte Carlo). Nesse trabalho, construiu-se amostras de posterioris de (β, η) através de algoritmos MCMC e utilizou-se os algoritmo EM (Expectation-Maximization). Esse procedimento, a seguir, é feito, para cada um dos componentes separadamente. Por isto, nessa seção os índices j não serão escritos. O procedimento é o mesmo para qualquer componente. A seguir, serão descritos os passos desse procedimento:

1. Define-se os valores iniciais dos hiperparâmetros fixos, as variâncias v_β e v_η . Sugere-se utilizar $v_\beta = v_\eta = 4$ Esses valores foram encontrados de forma heurística..
2. Define-se os valores iniciais para os parâmetros β e η e dos hiperparâmetros m_β , m_η a serem estimados.
3. Através dos valores iniciais fixos de m_β , m_η e do procedimento do MCMC gera-se uma amostra de tamanho n_p da distribuição a posteriori de β e η . Sugere-se usar $n_p = 1000$.
4. Passo E do algoritmo EM - Através da amostra a posteriori obtida no passo anterior 3, calcula-se a média a posteriori de β e η .
5. Passo M do algoritmo EM - Maximiza-se a amostra à posteriori obtida em 3 em relação à m_β e m_η , sendo os valores de β e η os obtidos em 4.
6. Atualiza-se os valores de m_β e m_η com os obtidos em 5 e os de β e η obtidos em 4 e retoma-se o processo em 1, desde que não haja convergência m_β e m_η .

O ciclo de iterações é realizado e repetido até que a convergência m_β e m_η ocorra. Quando isso ocorre, o algoritmo termina com uma amostra a posteriori de β e η obtida através desses novos valores de m_β e m_η . Sugere-se uma tolerância de 3 casas decimais entre o valor novo obtido e o

anterior de m_β e m_η como tolerância para finalizar a iteração e existir a convergência.

O próximo passo é estimar a função de confiabilidade das componentes. Define-se $\beta = (\beta_{11}, \beta_{12}, \dots, \beta_{1n_p})$ e $\eta = (\eta_{11}, \eta_{12}, \dots, \eta_{1n_p})$, a amostra à posteriori obtida via o procedimento EM e MCMC. Em muitos trabalhos, calcula-se a média dessa amostra obtida para β e η e este valor é colocado na função de confiabilidade. Esse procedimento chama-se plugin e é inferior ao que será explicado.

Considere para cada $t > 0$, $Y_\ell(t) = R(t | \beta_{1\ell}, \eta_{1\ell})$, para $\ell = 1, \dots, n_p$, na qual $R(t | \beta_{1\ell})$ função de confiabilidade Weibull. Podemos, assim, estimar a função de confiabilidade da componente 1 através da média a posteriori do cálculo funções de confiabilidade ou de distribuição das amostras mostrado na expressão a seguir,

$$\widehat{R(t)} = E[R(t) | dados] = \frac{1}{n_p} \sum_{\ell=1}^{n_p} Y_\ell(t) = \frac{1}{n_p} \sum_{\ell=1}^{n_p} R_\ell(t), \text{ para cada } t > 0.$$

Para cada $t > 0$ fixo, $Y(t) = (Y_1(t), \dots, Y_{n_p}(t))$ é uma amostra a posteriori da função de confiabilidade da componente 1, para obter o intervalo de credibilidade, basta obter o o intervalo de credibilidade bayesiano com maior densidade (HPD) ou calcular os quantis.

Para calcular a estimativa pontual da confiabilidade média é feito o seguinte,

$$E[T | data] = \frac{1}{n_p} \sum_{\ell=1}^{n_p} E[T | \beta_{1\ell}, \eta_{1\ell}], \ell = 1, \dots, n_p$$

em que $E[T | \beta_{1\ell}, \eta_{1\ell}] = \eta_{1\ell} \Gamma(1 + 1/\beta_{1\ell})$ é a média de uma variável aleatória com distribuição Weibull, e $\Gamma(\cdot)$ a função matemática gamma. Isso pode ser feito, também, para estimar a mediana e outras funções de interesse. O mesmo procedimento é feito para as outros componentes e será omitido.

3.1.2 Simulações

Foram realizadas duas simulações para avaliar a adaptabilidade da distribuição Weibull para sistemas em série e paralelo. Na primeira, mostra-se um cenário específico com figuras e propriedades. Já a segunda, mostra diversos cenários distintos possíveis com uma tabela de medidas-resumo.

3.1.2.1 Simulação 1

Essa simulação foi feita com 1 cenário com 3 componentes, distribuições diferentes, médias similares e variâncias distintas.

A primeira componente X_1 determinou-se ter distribuição Weibull com média 2 e variância 4, a segunda X_2 com distribuição gamma com média 2 e variância 0,667 e a terceira com distribuição log-normal X_3 com média 2,014 e variância 6,968. Foi gerada uma amostra de tamanho $n = 100$ dessas três componentes para o sistema em série e outra amostra de mesmo tamanho e com as mesmas componentes (distribuições com parâmetros iguais) para o sistema em paralelo.

Os dados simulados tiveram as seguintes características: (i) para o sistema em série, obteve-se, respectivamente, 64%, 80% e 56% de dados censurados para as componentes 1, 2 e 3; (ii) já para o sistema em paralelo observou-se 61%, 68% e 71% de dados censurados, respectivamente, para os mesmos três componentes. Isso significa que da amostra de 100 sistemas em série, em 64 desses casos a componente 1 foi censurada. O interesse primordial, desse exemplo, está na estimação das funções de confiabilidade das componentes. Dessa forma, esse exemplo, por possuir uma quantidade expressiva de dados censurados nas componentes o transforma num exemplo desafiador no sentido de avaliar a eficácia do método.

A estimação foi feita através via o método de MCMC e EM como comentado anteriormente. O procedimento foi feito da seguinte forma, gerou-se uma amostra de tamanho 20000 à posteriori, descartou-se os primeiros 10000 pontos amostrais (burn-in), visando atingir a medida estacionária.

Para estimação das funções de confiabilidade e a região de credibilidade utilizou-se uma amostra à posteriori de tamanho 1000. Foram considerados saltos (jumps) entre os pontos amostrais de tamanho 10, em que descarta-se uma amostra de tamanho 10 entre cada um dos pontos amostrais. Isto visa diminuir a auto-correlação entre esses pontos.

Determinou-se valores de $v_\beta = v_\eta = 4$ para todas componentes tanto para o sistema em paralelo como o em série. Os resultados encontrados na simulação foram os seguintes: $\widehat{\mathbf{m}}_\beta = (1.31, 4.12, 1.44)$ e $\widehat{\mathbf{m}}_\eta = (2.19, 2.03, 1.80)$ para o sistema em série e $\widehat{\mathbf{m}}_\beta = (1.28, 2.57, 0.92)$ e $\widehat{\mathbf{m}}_\eta = (2.67, 2.28, 2.11)$ para o sistema em paralelo.

A Tabela 3.1 mostra as estimativas de média e desvio padrão à posteriori de cada parâmetro do modelo para ambos o sistemas. Os desvios padrões são relativamente pequenos comparados com a grandeza da estimativa, além disso, a estimativa da do tempo de confiabilidade médio são próximos aos originais indicando uma boa performance do método.

Tabela 3.1: Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades da simulação 1 para os sistemas em série e paralelo.

	Estimativas do sistema em série		
	β	η	$E(T \beta, \eta)$
Componente 1	1,26 (0,17)	2,27 (0,41)	2,13 (0,45)
Componente 2	3,98 (0,53)	2,06 (0,12)	1,87 (0,11)
Componente 3	1,40 (0,17)	1,83 (0,22)	1,68 (0,22)
	Estimativas do sistema em paralelo		
	β	η	$E(T \beta, \eta)$
Componente 1	1,23 (0,15)	2,60 (0,27)	2,45 (0,22)
Componente 2	2,47 (0,32)	2,25 (0,14)	2,00 (0,13)
Componente 3	0,87 (0,11)	1,98 (0,33)	2,17 (0,29)

As Figuras 3.2 e 3.3 mostram a função de confiabilidade original de cada componente, comparando com a estimativa junto com o intervalo de credibilidade de 95% para os sistemas em série e em paralelo, respectivamente. Percebe-se que a estimativa ficou muito próxima da real. Nota-se, também, que em algumas das componentes a função de confiabilidade real está fora da fronteira do intervalo de credibilidade de 95%. Considerando que o a simulação foi feita com grande quantidade de dados censurados, sendo, assim, um exemplo desafiador no sentido de estimação, isso é esperado que aconteça.

Figura 3.2: Funções de confiabilidade das componentes do sistema em série: (a) Componente 1; (b) Componente 2; (c) Componente 3.

Figura 3.3: Funções de confiabilidade das componentes do sistema em paralelo: (a) Componente 1; (b) Componente 2; (c) Componente 3.

3.1.2.2 Simulação 2

Essa simulação foi feita para mostrar a qualidade do modelo bayesiano hierárquico Weibull proposto frente a diversas situações. Para isto, foram considerados 108 diferentes cenários com 1000 réplicas em cada um deles. Utilizou-se 3 diferentes tamanhos de amostra ($n = 30, 100$ e 1000), 3 diferentes proporções de dados censurados (0%, 20% e 40%), 2 diferentes médias (2 e 7), 3 componentes com diferentes distribuições (com distribuições Weibull, gamma e log-normal) e 2 tipos de sistema (série-censura à direita e paralelo-censura à esquerda).

A variância das distribuições geradoras foi fixada em 5. Para garantir as proporções de censurada desejadas, foi fixado um tempo para cada cenário no qual todos valores acima deste foram censurados à direita no sistema em série, e todos valores abaixo desse tempo foram censurados à esquerda no sistema em paralelo. Esse tipo de censura é não-aleatória. A simulação foi feita com uma amostra de tamanho 100 para cada cenário.

As Tabelas 3.2 até a 3.7 resumem alguns resultados da simulação. Nelas, estão descrito o viés e o Erro Quadrático Médio (EQM) para o tempo médio estimado de confiabilidade nos diversos cenários.

Em todos cenários, observa-se que o modelo foi bem estimado, pois o Viés e o EQM do tempo médio estimado encontrados são próximos à zero. Isso ocorreu mesmo em cenários com tamanho de amostra pequeno ($n = 30$) e com proporções de censura alta (40%). Os resultados demonstram boa performance tanto quando a distribuição da componente é Weibull, como, também, quando a distribuição é gamma e log-normal.

Tabela 3.2: Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 1 que possui distribuição Weibull e com censura à direita (sistema em série - simulação 2).

Cens.	Verd. $E(T)$	n = 30		n = 100		n = 1000	
		Viés	EQM	Viés	EQM	Viés	EQM
0%	2	-0,1300	0,2217	-0,0426	0,0572	-0,0057	0,0043
0%	7	0,0293	0,1523	0,0026	0,0474	-0,0095	0,0043
20%	2	-0,2813	0,4893	-0,1242	0,1275	-0,0200	0,0068
20%	7	-0,0522	0,2307	-0,0233	0,0565	-0,0148	0,0047
40%	2	-0,4293	0,7364	-0,2394	0,2925	-0,0239	0,0108
40%	7	-0,1773	0,3590	-0,0661	0,0874	-0,0162	0,0055

Tabela 3.3: *Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 2 que possui distribuição gamma e com censura à direita (sistema em série - simulação 2).*

Cens.	Verd. $E(T)$	n = 30		n = 100		n = 1000	
		Viés	EQM	Viés	EQM	Viés	EQM
0%	2	-0,1407	0,2286	-0,0191	0,0593	-0,0003	0,0055
0%	7	0,0006	0,1845	0,0081	0,0623	0,0098	0,0048
20%	2	-0,3829	0,6346	-0,1132	0,1234	-0,0560	0,0111
20%	7	0,0569	0,1721	0,1121	0,0669	0,1049	0,0153
40%	2	-0,5068	0,9424	-0,3084	0,3807	-0,1174	0,0299
40%	7	0,1389	0,2505	0,2372	0,1107	0,2336	0,0582

Tabela 3.4: *Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 3 que possui distribuição log-normal e com censura à direita (sistema em série - simulação 2).*

Cens.	Verd. $E(T)$	n = 30		n = 100		n = 1000	
		Viés	EQM	Viés	EQM	Viés	EQM
0%	2	-0,1160	0,1757	-0,0024	0,0396	-0,0089	0,0050
0%	7	-0,0348	0,1551	0,0462	0,0429	0,0220	0,0052
20%	2	0,0720	0,1410	0,2475	0,0779	0,2523	0,0655
20%	7	0,0786	0,1492	0,1848	0,0637	0,1620	0,0296
40%	2	0,2373	0,1795	0,4205	0,1981	0,4371	0,1928
40%	7	0,2190	0,2094	0,3438	0,1532	0,3337	0,1147

Tabela 3.5: *Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 1 que possui distribuição Weibull e com censura à esquerda (sistema em paralelo - simulação 2).*

Cens.	Verd. $E(T)$	n = 30		n = 100		n = 1000	
		Viés	EQM	Viés	EQM	Viés	EQM
0%	2	-0,1300	0,2217	-0,0426	0,0572	-0,0057	0,0043
0%	7	0,0293	0,1523	0,0026	0,0474	-0,0095	0,0043
20%	2	0,1030	0,1178	0,0034	0,0475	-0,0060	0,0043
20%	7	0,0387	0,1562	-0,0001	0,0490	-0,0094	0,0043
40%	2	-0,1813	0,2660	-0,0471	0,0579	-0,0065	0,0043
40%	7	0,0301	0,1813	0,0040	0,0597	-0,0099	0,0048

Tabela 3.6: *Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 2 que possui distribuição gamma e com censura à esquerda (sistema em paralelo - simulação 2).*

Cens.	Verd. $E(T)$	n = 30		n = 100		n = 1000	
		Viés	EQM	Viés	EQM	Viés	EQM
0%	2	-0,1407	0,2286	-0,0192	0,0594	-0,0003	0,0055
0%	7	0,0006	0,1845	0,0081	0,0623	0,0098	0,0048
20%	2	0,1029	0,1038	0,0448	0,0490	0,0008	0,0054
20%	7	0,1108	0,2143	0,0950	0,0739	0,0932	0,0139
40%	2	-0,1934	0,2463	-0,0245	0,0602	-0,0029	0,0055
40%	7	0,2185	0,2842	0,2105	0,1242	0,1850	0,0402

Tabela 3.7: Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 3 que possui distribuição log-normal e com censura à esquerda (sistema em paralelo - simulação 2).

Cens.	Verd. $E(T)$	n = 30		n = 100		n = 1000	
		Viés	EQM	Viés	EQM	Viés	EQM
0%	2	-0,1160	0,1757	-0,0024	0,0396	-0,0089	0,0050
0%	7	-0,0348	0,1551	0,0462	0,0429	0,0220	0,0052
20%	2	-0,0092	0,1253	0,0254	0,0405	0,0211	0,0052
20%	7	0,1003	0,1701	0,1757	0,0784	0,1375	0,0240
40%	2	-0,1165	0,1937	0,0575	0,0441	0,0624	0,0086
40%	7	0,2359	0,2727	0,3138	0,1548	0,2763	0,0825

3.2 Sistemas k-de-n

Os sistemas redundantes k-de-n foram explicitados no exemplo 2.2.3. No caso particular do sistema redundante 2-de-3, existem duas representações para esse sistema, uma em paralelo-série e outra em série-paralelo e que são mostradas na Figura 3.4. Nem sempre a representação é tão similar comparando paralelo-série e série-paralelo como nesse caso particular. Essas representações são esquemáticas e para fins de análise, já que nenhuma das componentes existe mais de uma vez ao mesmo tempo.

Figura 3.4: Representação gráfica de um sistema com estrutura 2-de-3. (a) Sistema paralelo-série (SiPS); (b) Sistema série-paralelo (SiSP).

A abordagem deste problema, feita com componentes independentes e com mesma distribuição, pode ser vista em ?. Neste trabalho, este tema será tratado com componentes com diferentes distribuições e resolvido através de censura intervalar.

Considera-se que no momento que ocorre a falha de um sistema anotam-se os dados de quais componentes estão ainda funcionando e quais componentes já falharam. Dessa forma, no exemplo mencionado, o equipamento pode falhar caso apenas os componentes 1 e 2 deixem de funcionar, e nesse caso o componente 3 é censurado à direita; pode quebrar caso apenas 1 e 3 deixem de funcionar e o componente 2 é censurado à direita e, finalmente, pode quebrar caso apenas 2 e 3 deixem de funcionar e o componente 1 é censurado à direita. Uma suposição é feita de que os tempos de falha não são coincidentes, ou seja, dois equipamentos não falham exatamente no mesmo instante de tempo e por isto não é possível que todos os componentes estejam quebrados no instante que o sistema falhar.

A seguir, é descrito um modelo utilizado para a resolução desse sistema e depois serão mostradas duas simulações feitas para a avaliação do modelo proposto.

3.2.1 Modelo bayesiano hierárquico Weibull com censura intervalar

A extensão do modelo anteriormente apresentado em 3.1.1 para acomodar casos intervalares é a seguinte. Seja $t_i = (t_{1i}, t_{2i})$ um vetor com observações do indivíduo i da amostra. Para todo

i tem-se $t_{1i} \leq t_{2i}$. Esse vetor indica o seguinte, caso se saiba o instante de falha, tem-se por convenção $t_{1i} = t_{2i}$, caso haja censura a direita $t_{2i} = \infty$, a esquerda $t_{1i} = 0$, nos demais casos existe a censura intervalar. Assim, de outra forma, podemos compactar, o vetor pela representação $t = \{(t_{11}, t_{21}), \dots, (t_{1n}, t_{2n})\}$, e ele detém todas as informações da amostra relativas a falha e censura.

Desse modo, a posteriori desse modelo que também agrega dados censurados intervalares é generalizada por

$$\pi(\boldsymbol{\theta}, \boldsymbol{\zeta} | \mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}, \boldsymbol{\zeta}) \prod_{j=1}^k \prod_{i=1}^n \left[\left[\frac{t_{1i}}{\eta_j} \right]^{\beta_j - 1} \frac{\beta_j}{\eta_j} \exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^{\beta_j} \right\} \right]^{I_{\{t_{1i}=t_{2i}\}}} \left[\exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^{\beta_j} \right\} - \exp \left\{ - \left(\frac{t_{2i}}{\eta_j} \right)^{\beta_j} \right\} \right]^{1 - I_{\{t_{1i}=t_{2i}\}}}, \quad (3.12)$$

O teorema que prova que a posteriori acima é própria está na próxima seção.

3.2.2 Posteriori própria

Essa seção é bem técnica, nela descreve-se e prova-se que a posteriori do modelo bayesiano hierárquico Weibull para dados censurados intervalares é própria para uma classe de distribuições à priori não-informativas. Esse resultado é essencial para a construção da posteriori.

A posteriori é própria se para todo componente vale $\int \pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) d\boldsymbol{\theta} < \infty$ onde $\boldsymbol{\theta} = (\beta, \eta)$. Desconsiderou-se os índices j , pois a prova vale para qualquer componente. A posteriori é dada por

$$\pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}) \prod_{i=1}^n \left[\left[\frac{t_{1i}}{\eta} \right]^{\beta - 1} \frac{\beta}{\eta} \exp \left\{ - \left(\frac{t_{1i}}{\eta} \right)^{\beta} \right\} \right]^{I_{\{t_{1i}=t_{2i}\}}} \left[\exp \left\{ - \left(\frac{t_{1i}}{\eta} \right)^{\beta} \right\} - \exp \left\{ - \left(\frac{t_{2i}}{\eta} \right)^{\beta} \right\} \right]^{1 - I_{\{t_{1i}=t_{2i}\}}}, \quad (3.13)$$

Recordando que a função densidade da Weibull de parâmetros (η, β) positivos é dada por:

$$\frac{\beta}{\eta} \left[\frac{t}{\eta} \right]^{\beta - 1} \exp \left[- \left(\frac{t}{\eta} \right)^{\beta} \right] \quad (3.14)$$

Considere agora uma classe geral de prioris não-informativas dadas por:

$$\pi^b(\theta, \beta) = \frac{1}{\eta \beta^b}, b \geq 0 \quad (3.15)$$

Teorema 1 *Seja uma amostra de dados em que nem todas censuras são à direita (e nesse caso ela já seria própria). Para $b \geq 0$ e $n = 1$ e existe falha, a posteriori acima é imprópria, e para $n \geq 1$ a posteriori acima é própria caso exista pelo menos dois valores distintos.*

Prova: Pela independência das componentes e por terem a mesma distribuição basta analisar a condição de que

$$\int \pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) d\boldsymbol{\theta} < \infty \quad (3.16)$$

Seja uma amostra qualquer tamanho n . Considere que $(1, \dots, n_f)$ são os i tais que $t_{1i} = t_{2i}$ e $(n_f+1, \dots, n)$ são os i tais que $t_{1i} \neq t_{2i}$.

$$\pi(\boldsymbol{\theta}|\mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}) \prod_{i=1}^{n_f} \left[\left[\frac{t_{1i}}{\eta} \right]^{\beta-1} \frac{\beta}{\eta} \exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^\beta \right\} \right] \quad (3.17)$$

$$\prod_{i=n_f+1}^n \left[\exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^\beta \right\} - \exp \left\{ - \left(\frac{t_{2i}}{\eta_j} \right)^\beta \right\} \right], \quad (3.18)$$

Seja $(a_1, \dots, a_{t+1})$ uma sequência de valores positivos e menores que 1 em que todo $a_{2n-1} > a_{2n}$ para $n \geq 1$ e $2n \leq t+1$. Para t inteiro, ímpar e $t \geq 4$, $(a_1 - a_2) * \dots * (a_t - a_{t+1}) < a_1 * a_2 * \dots * a_t$. Prova: Para tal considere o caso de $t = 3$, ou seja, duas parcelas. Temos que $(a_1 - a_2)(a_3 - a_4) < a_1(a_3 - a_4) < a_1a_3$. Considerando isso verdadeiro para $t = 2n - 1$, temos que $a_1a_3a_5 \dots a_{2n-1} > (a_1 - a_2) * \dots * (a_{2n-1} - a_{2n})$ e desse modo $(a_1a_3a_5 \dots a_{2n-1})(a_{2n+1} - a_{2n+2}) < (a_1a_3a_5 \dots a_{2n+1})$.

Na posteriori acima a parcela

$$\prod_{i=n_f+1}^n \left[\exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^\beta \right\} - \exp \left\{ - \left(\frac{t_{2i}}{\eta_j} \right)^\beta \right\} \right], \quad (3.19)$$

é do mesmo tipo que a prova acima. Isso pode ser visto, pois a função exponencial é estritamente crescente e para os casos com censura tem-se

$$t_{1i} < t_{2i} \Leftrightarrow \frac{t_{1i}}{\eta} < \frac{t_{2i}}{\eta} \Leftrightarrow -\frac{t_{1i}}{\eta} > -\frac{t_{2i}}{\eta} \quad (3.20)$$

$$\Leftrightarrow -\left(\frac{t_{1i}}{\eta} \right)^\beta > -\left(\frac{t_{2i}}{\eta} \right)^\beta \quad (3.21)$$

$$\exp \left\{ - \left(\frac{t_{1i}}{\eta} \right)^\beta \right\} > \exp \left\{ - \left(\frac{t_{2i}}{\eta} \right)^\beta \right\} \quad (3.22)$$

Para avaliar se a posteriori é própria, então, é necessário apenas avaliar que o majorante

$$\pi(\boldsymbol{\theta}|\mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}) \prod_{i=1}^{n_f} \left[\left[\frac{t_{1i}}{\eta} \right]^{\beta-1} \frac{\beta}{\eta} \exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^\beta \right\} \right] \quad (3.23)$$

$$\prod_{i=n_f+1}^n \left[\exp \left\{ - \left(\frac{t_{1i}}{\eta_j} \right)^\beta \right\} \right] < \infty \quad (3.24)$$

Para $n = 1$ e $t_{11} = t_{12}$,

$$\int_0^\infty \int_0^\infty \frac{1}{\eta^{\beta b}} \left[\left[\frac{t_{11}}{\eta} \right]^{\beta-1} \frac{\beta}{\eta} \exp \left\{ - \left(\frac{t_{11}}{\eta} \right)^\beta \right\} \right] d\eta d\beta = \quad (3.25)$$

$$\int_0^\infty \frac{1}{\beta^b} \int_0^\infty \left[\frac{\beta t_{11}^{\beta-1}}{\eta^{\beta+1}} \exp \left\{ - \left(\frac{t_{11}}{\eta} \right)^\beta \right\} \right] d\eta d\beta = \quad (3.26)$$

Recorda-se que uma gamma inversa com parâmetros α e β positivos tem função de densidade dada por:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} \exp \left\{ -\frac{\beta}{x} \right\} \quad (3.27)$$

agora fazendo a mudança de variável $\left(\frac{t_{11}}{\eta}\right)^\beta = \frac{\beta}{x}$, tem-se $t_{11}^\beta dx = \eta^{\beta-1} \beta^2 d\eta$, temos que,

$$\int_0^\infty \frac{1}{\beta^b} \int_0^\infty \left(\frac{\beta}{x}\right)^2 \frac{1}{\beta t_{11}} \exp \left\{ -\frac{\beta}{x} \right\} dx d\beta = \quad (3.28)$$

$$\int_0^\infty \frac{1}{\beta^b t_{11}} \int_0^\infty \frac{\beta}{x^2} \exp \left\{ -\frac{\beta}{x} \right\} dx d\beta = \quad (3.29)$$

assim, observa-se que a integral do meio compõe-se da distribuição de probabilidade de uma gamma invertida com parâmetros $(1, \beta)$,

$$\int_0^\infty \frac{1}{\beta^b} \frac{1}{t_{11}} d\beta = \infty \quad (3.30)$$

Agora analisemos o caso $n \geq 1$,

$$\pi(\boldsymbol{\theta} | \mathbf{t}, \boldsymbol{\delta}) \propto \pi(\boldsymbol{\theta}) \prod_{i=1}^{n_f} \left[\left[\frac{t_{1i}}{\eta} \right]^{\beta-1} \frac{\beta}{\eta} \right] \prod_{i=1}^n \exp \left\{ -\left(\frac{t_{1i}}{\eta} \right)^\beta \right\} \quad (3.31)$$

A integral fica o seguinte,

$$\int_0^\infty \int_0^\infty \frac{\beta^{n_f-b}}{\eta^{n_f \beta+1}} \exp \left\{ -\frac{\sum_{i=1}^n t_{1i}^\beta}{\eta^\beta} \right\} \prod_{i=1}^{n_f} t_{1i}^{\beta-1} d\eta d\beta = \quad (3.32)$$

fazendo a substituição de $\frac{\sum_{i=1}^n t_{1i}^\beta}{\eta^\beta} = \frac{\beta}{x}$, tem-se $\sum_{i=1}^n t_{1i}^\beta dx = \eta^{\beta-1} \beta^2 d\eta$, (aqui novamente temos uma na distribuição gamma invertida mas de parâmetros (n, β)) e seja $k_1 > \prod_{i=1}^{n_f} t_{1i}^{-1}$ real positivo. Considere também t_{1v} tal que $t_{1v} < \max(t_{11}, \dots, t_{1n_f})$, então para todo $b \geq 0$ temos que:

$$\int_0^\infty \frac{\beta^{-b}}{\left[\sum_{i=1}^n t_{1i}^\beta \right]^{n_f}} \prod_{i=1}^{n_f} t_{1i}^{\beta-1} \int_0^\infty \beta^{n_f-2} \left(\frac{\beta}{x} \right)^{n_f+1} \exp \left\{ -\frac{\beta}{x} \right\} dx d\beta = \quad (3.33)$$

$$\int_0^\infty \frac{\beta^{-b}}{\left[\sum_{i=1}^n t_{1i}^\beta \right]^{n_f}} \prod_{i=1}^{n_f} t_{1i}^{\beta-1} \int_0^\infty \beta^{2n_f-1} x^{-n_f-1} \exp \left\{ -\frac{\beta}{x} \right\} dx d\beta = \quad (3.34)$$

$$\int_0^\infty \frac{\beta^{n_f-1-b}}{\left[\sum_{i=1}^n t_{1i}^\beta \right]^{n_f}} \prod_{i=1}^{n_f} t_{1i}^{\beta-1} \int_0^\infty \beta^{n_f} x^{-n_f-1} \exp \left\{ -\frac{\beta}{x} \right\} dx d\beta = \quad (3.35)$$

$$\int_0^\infty \beta^{n_f-b-1} \Gamma(n_f) \frac{\prod_{i=1}^{n_f} t_{1i}^{\beta-1}}{\left[\sum_{i=1}^n t_{1i}^\beta\right]^{n_f}} d\beta < \int_0^\infty k_1 \beta^{n_f-b-1} \Gamma(n_f) \frac{\prod_{i=1}^{n_f} t_{1i}^\beta}{\left[\sum_{i=1}^n t_{1i}^\beta\right]^{n_f}} d\beta < (3.36)$$

$$\int_0^\infty k_1 \beta^{n_f-b-1} \Gamma(n_f) \frac{\prod_{i=1}^{n_f} t_{1i}^\beta}{\left[\sum_{i=1}^{n_f} t_{1i}^\beta\right]^{n_f}} d\beta < k_1 \Gamma(n_f) \int_0^\infty \beta^{n_f-b-1} \left[\frac{t_{1v}}{\max(t_{11}, \dots, t_{1n_f})} \right]^\beta d\beta < \infty (3.37)$$

Isso é válido, pois para um $a \in \mathbb{R}$,

$$\Gamma(a) = \int_0^\infty e^{-t} t^{a-1} dt < \infty (3.38)$$

Então, para como $h = \frac{t_{1v}}{\max(t_{11}, \dots, t_{1n_f})} < 1$, existe um $c \in \mathbb{R}$

$$\int_0^\infty \beta^{n_f-b-1} h^\beta d\beta = \int_0^\infty \beta^{n_f-b-1} e^{\ln h \beta} d\beta < (3.39)$$

e fazendo a substituição de $t = -\ln h \beta$, $c = -\left(\frac{1}{\ln h}\right)^{n_f-b-1}$,

$$< c \int_0^\infty e^{-t} t^{n_f-b-1} dt < c \int_0^\infty e^{-t} t^{z-1} dt < \infty (3.40)$$

sendo z o menor inteiro positivo maior que $n_f - b$.

3.2.3 Simulações

Foram realizada duas simulações para avaliar a adaptabilidade do método para o sistema k-de-n, especificamente para o sistema 2-de-3. A primeira mostra-se um cenário específico com figuras e propriedades. Já a segunda, mostra diversos cenários distintos possíveis com uma tabela de medidas-resumo.

3.2.3.1 Simulação 1

Foi realizada uma simulação de um caso do sistema k-de-n, o sistema 2-de-3. Esse sistema está em funcionamento, se e somente se, pelo menos 2 componentes das 3 esteja funcionando. O cenário utilizado para essa simulação foi um com componentes com diferentes distribuições, médias similares e variâncias distintas. As componentes escolhidas foram as mesmas da Simulação anterior dos Sistemas em Série e Paralelo.

Dessa forma, determinou-se a componente X_1 ter distribuição Weibull com média 2 e variância 4, a segunda componente, X_2 com distribuição gamma com média 2 e variância 0,667 e a terceira com distribuição log-normal X_3 com média 2,014 e variância 6,968. O tamanho da amostra para cada componente foi de $n = 100$.

A proporção de censura do sistema com esse tamanho amostral foi de 75% para a componente 1, 52% para a componente 2 e 73% para a componente 3. Novamente, esse é um sistema com uma grande porcentagem de dados censurados, o que torna o exemplo uma avaliação interessante da eficácia do método.

A estimação foi feita através dos algoritmos MCMC e EM. Foi gerada um amostra tamanho 20000 e descartou-se os 10000 valores (burn-in). Utilizou-se uma amostra à posteriori de tamanho 1000 para a estimação das funções de confiabilidade e a região de credibilidade. Considerou-se saltos (jumps) entre os pontos amostrais de tamanho 10, em que descarta-se uma amostra de tamanho 10 entre cada um dos pontos amostrais.

Determinou-se valores de $v_\beta = v_\eta = 4$ para todas componentes. Os resultados encontrados na simulação foram os seguintes: $\widehat{\mathbf{m}}_\beta = (1.11, 2.70, 1.02)$ e $\widehat{\mathbf{m}}_\eta = (2.43, 2.20, 2.03)$.

A Tabela 3.8 mostra as estimativas de média e desvio padrão à posteriori de cada parâmetro do modelo. Os desvios padrões são relativamente pequenos comparados com a grandeza da estimativa, além disso, a estimativa da do tempo de confiabilidade médio são próximos aos originais indicando uma boa performance do método.

Tabela 3.8: Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades da simulação 1 para o sistema 2-de-3.

	Estimativas do sistema 2-de-3		
	β	η	$E(T \beta, \eta)$
Componente 1	1,04 (0,17)	2,48 (0,35)	2,51 (0,46)
Componente 2	2,64 (0,24)	2,20 (0,11)	1,95 (0,10)
Componente 3	0,97 (0,16)	2,02 (0,27)	2,11 (0,32)

A Figura 3.5 mostra a função de confiabilidade original e a estimativa com o intervalo de credibilidade de 95% para cada componente. A estimativa ficou muito próxima da real e dentro dos intervalos de credibilidade.

Figura 3.5: Funções de confiabilidade das componentes do sistema 2-de-3: (a) Componente 1; (b) Componente 2; (c) Componente 3.

3.2.3.2 Simulação 2

Essa simulação foi feita para mostrar a qualidade do modelo bayesiano hierárquico Weibull proposto frente a diversas situações. Para tanto, englobou-se diferentes proporções de censura tanto à direita como à esquerda para componentes com distribuição Weibull, log-normal e gamma. Essa simulação, feita dessa forma, engloba tanto o casos como sistemas 2-de-3 aqui descritos, como também, os sistemas ponte descritos na seção a seguir. Portanto, esse exemplo será descrito no item 3.3.1.2.

3.3 Sistemas Ponte (Bridge)

Recordamos o exemplo 2.2.4 chamado de sistema ponte (bridge). Esse sistema até esse trabalho não havia solução para o caso de componentes com diferentes distribuições (não são i.i.d).

Podemos observar esse sistema e reescrevê-lo de uma forma que ele possa ser resolvido com o mesmo método anterior desenvolvido para o k-de-n. Observe que esse sistema falha caso, falhem conjuntamente algum dos seguintes conjuntos de componentes $\{1 \text{ e } 2\}$, $\{4 \text{ e } 5\}$, $\{1, 3 \text{ e } 5\}$ e $\{2, 3 \text{ e } 4\}$. Esses quatros conjuntos são chamadas de conjuntos de corte mínimo. Numa situação hipotética, considere que o sistema falhe por terem queimado os componentes 1 e 2 ,assim, verifica-se também a situação dos demais componentes 3, 4 e 5 e anota-se se estes estão funcionando no momento de quebra ou se também falharam. Suponha que 3 também estava quebrado e 4 e 5 não estavam.

Nesse caso o componente 3 tem dado censurado à esquerda e os componentes 4 e 5 censura à direita. Assim, caso o sistema quebre, as componentes podem ser tanto censuradas à direita como à esquerda (exceto a última componente a falhar) e pode existir ao mesmo tempo os dois casos dentro do mesmo sistema. Dessa forma, podemos resolver esse sistema da mesma forma que o anterior.

A seguir, serão descritas duas simulações para avaliar a performance desse modelo.

3.3.1 Simulações

Foram realizadas duas simulações para avaliar a adaptabilidade do método para o sistema bridge com 5 componentes descrito no capítulo 2. A primeira delas, mostra um cenário específico com figuras e propriedades. Já a segunda, mostra diversos cenários distintos possíveis com uma tabela de medidas-resumo.

3.3.1.1 Simulação 1

Foi realizada uma simulação para o exemplo de sistema ponte (bridge) dado nesse trabalho. Como dito anteriormente, esse sistema falha se conjuntamente algum dos seguintes conjuntos de componentes 1 e 2, 4 e 5, 1, 3 e 5 e 2, 3 e 4 falharem. O cenário utilizado para essa simulação foi um com componentes com diferentes distribuições, médias similares e variâncias distintas.

Determinou-se a componente X_1 ter distribuição Weibull com média 2 e variância 4, a segunda componente, X_2 com distribuição gamma com média 2 e variância 0,667, a terceira com distribuição log-normal X_3 com média 2,014 e variância 6,968, a quarta componente X_4 com distribuição gamma com média 2 e variância 3 e a quinta componente X_5 com distribuição log-normal com média 2,014 e variância 5. O tamanho da amostra para cada componente foi de $n = 100$.

A proporção de censura do sistema com esse tamanho amostral foi de 82% para a componente 1, 71% para a componente 2 e 89% para a componente 3, 84% para a componente 4 e, finalmente, 74% para a componente 5. Novamente, ressalta-se que esse é um sistema com uma grande porcentagem de dados censurados, o que torna o exemplo uma avaliação interessante da eficácia do método.

A estimação foi feita através dos algoritmos MCMC e EM. Foi gerada um amostra tamanho 20000 e descartou-se os 10000 valores (burn-in). Utilizou-se uma amostra à posteriori de tamanho 1000 para a estimação das funções de confiabilidade e a região de credibilidade. Considerou-se saltos (jumps) entre os pontos amostrais de tamanho 10, em que descarta-se uma amostra de tamanho 10 entre cada um dos pontos amostrais.

Determinou-se valores de $v_\beta = v_\eta = 4$ para todas componentes. Os resultados encontrados na simulação foram os seguintes: $\widehat{\mathbf{m}}_\beta = (1.08, 2.77, 1.19, 1.22, 1.14)$ e $\widehat{\mathbf{m}}_\eta = (2.28, 2.32, 2.11, 2.64, 2.02)$.

A Tabela 3.9 mostra as estimativas de média e desvio padrão à posteriori de cada parâmetro do modelo. Os desvios padrões são relativamente pequenos comparados com a grandeza da estimativa, além disso, a estimativa da do tempo de confiabilidade médio são próximos aos originais indicando uma boa performance do método.

Tabela 3.9: Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades da simulação 1 para o sistema ponte (bridge).

	Estimativas do sistema ponte (bridge)		
	β	η	$E(T \beta, \eta)$
Componente 1	1,00 (0,19)	2,30 (0,32)	2,40 (0,52)
Componente 2	2,69 (0,28)	2,33 (0,13)	2,07 (0,12)
Componente 3	1,09 (0,22)	2,14 (0,29)	2,16 (0,48)
Componente 4	1,13 (0,22)	2,73 (0,39)	2,69 (0,55)
Componente 5	1,07 (0,17)	2,03 (0,25)	2,02 (0,29)

As Figuras 3.6 e 3.7 mostram a função de confiabilidade original e a estimativa com o intervalo de credibilidade de 95% para cada componente. A estimativa ficou muito próxima da real e dentro dos intervalos de credibilidade.

Figura 3.6: Funções de confiabilidade das componentes 1, 2 e 3 do sistema ponte (bridge): (a) Componente 1; (b) Componente 2; (c) Componente 3.

Figura 3.7: Funções de confiabilidade das componentes 4 e 5 do sistema ponte (bridge): (a) Componente 4; (b) Componente 5.

3.3.1.2 Simulação 2

Essa simulação foi feita para mostrar a qualidade do modelo bayesiano hierárquico Weibull proposto frente a diversas situações. Foram considerados 216 diferentes cenários. Utilizou-se 3 diferentes tamanhos de amostra ($n = 30$, $n = 50$, 100 e 1000), 3 diferentes proporções de dados censurados à esquerda (0%, 20% e 40%), 3 diferentes proporções de dados censurados à direita (0%, 20% e 40%), 2 diferentes médias (2 e 7) e 3 componentes com diferentes distribuições (com distribuições Weibull, gamma e log-normal).

A variância das distribuições geradoras foi, novamente, fixada em 5. De forma a garantir as proporções de censura desejadas, foi fixado um tempo para cada cenário no qual todos valores acima deste foram censurados à direita, e todos valores abaixo desse tempo foram censurados à esquerda. Essa censura é não-aleatória. Utilizou-se para estimação os métodos de MCMC e EM com uma amostra de tamanho 100 para cada um dos cenários.

As Tabelas 3.10 até a 3.12 resumem alguns resultados dessa simulação. Nelas, estão descrito o viés e o Erro Quadrático Médio (EQM) para o tempo médio estimado de confiabilidade nos diversos cenários.

Em todos cenários, observa-se que o modelo foi bem estimado, pois o Viés e o EQM do tempo médio estimado encontrados são próximos à zero. Isso ocorreu mesmo em cenários com tamanho de amostra pequeno ($n = 30$) e com proporções de censura alta (40%). Ressalta-se que cenários em que existe 40% de censura à direita e 40% de censura à esquerda, no total, 80% dos dados estão

censurados. Os resultados demonstram boa performance tanto quando a distribuição da componente é Weibull, como, também, quando a distribuição é gamma e log-normal.

Tabela 3.10: *Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 1 que possui distribuição Weibull (simulação 2 -sistemas k-de-n e ponte (bridge)).*

Cens. (dir)	Cens. (esq)	Verd. $E(T)$	n = 30		n = 50		n = 100		n = 1000	
			Viés	EQM	Viés	EQM	Viés	EQM	Viés	EQM
0%	0%	2	-0.13	0.21	-0.07	0.11	-0.04	0.06	-0.01	0.00
0%	0%	7	0.03	0.15	0.02	0.08	0.00	0.05	-0.01	0.00
0%	20%	2	0.11	0.12	0.05	0.07	0.00	0.05	-0.01	0.00
0%	20%	7	0.04	0.16	0.02	0.09	0.00	0.05	-0.01	0.00
0%	40%	2	-0.18	0.26	-0.09	0.11	-0.05	0.06	-0.01	0.00
0%	40%	7	0.03	0.18	0.01	0.09	0.00	0.06	-0.01	0.00
20%	0%	2	-0.28	0.49	-0.24	0.33	-0.12	0.13	-0.02	0.01
20%	0%	7	-0.05	0.23	-0.04	0.10	-0.02	0.06	-0.01	0.00
20%	20%	2	-0.32	0.57	-0.28	0.41	-0.15	0.18	-0.02	0.01
20%	20%	7	-0.04	0.21	-0.04	0.11	-0.02	0.06	-0.01	0.00
20%	40%	2	-0.25	0.43	-0.27	0.41	-0.20	0.21	-0.02	0.01
20%	40%	7	-0.09	0.30	-0.05	0.11	-0.01	0.06	-0.01	0.01
40%	0%	2	-0.42	0.71	-0.43	0.71	-0.24	0.30	-0.02	0.01
40%	0%	7	-0.18	0.37	-0.13	0.19	-0.06	0.09	-0.02	0.01
40%	20%	2	-0.34	0.52	-0.39	0.61	-0.32	0.48	-0.04	0.02
40%	20%	7	-0.20	0.42	-0.19	0.27	-0.08	0.10	-0.02	0.01
40%	40%	2	-0.45	0.83	-0.41	0.80	-0.49	0.81	-0.05	0.02
40%	40%	7	-0.23	0.43	-0.21	0.29	-0.12	0.15	-0.02	0.01

Tabela 3.11: *Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 2 que possui distribuição gamma (simulação 2 -sistemas k-de-n e ponte (bridge)).*

Cens. (dir)	Cens. (esq)	Verd. $E(T)$	n = 30		n = 50		n = 100		n = 1000	
			Viés	EQM	Viés	EQM	Viés	EQM	Viés	EQM
0%	0%	2	-0.14	0.23	-0.10	0.12	-0.02	0.06	-0.00	0.01
0%	0%	7	0.00	0.18	0.00	0.11	0.01	0.06	0.01	0.00
0%	20%	2	0.13	0.10	0.09	0.07	0.05	0.05	0.00	0.01
0%	20%	7	0.11	0.21	0.09	0.13	0.10	0.07	0.09	0.01
0%	40%	2	-0.19	0.24	-0.11	0.12	-0.03	0.06	-0.00	0.01
0%	40%	7	0.22	0.28	0.20	0.18	0.21	0.12	0.18	0.04
20%	0%	2	-0.39	0.64	-0.25	0.31	-0.11	0.12	-0.06	0.01
20%	0%	7	0.06	0.17	0.09	0.11	0.11	0.07	0.11	0.02
20%	20%	2	-0.49	0.84	-0.25	0.31	-0.13	0.13	-0.05	0.01
20%	20%	7	0.08	0.21	0.13	0.13	0.15	0.08	0.15	0.03
20%	40%	2	-0.42	0.62	-0.33	0.46	-0.14	0.15	-0.04	0.01
20%	40%	7	0.13	0.26	0.19	0.17	0.22	0.12	0.19	0.04
40%	0%	2	-0.51	0.94	-0.39	0.64	-0.31	0.38	-0.12	0.03
40%	0%	7	0.14	0.25	0.19	0.16	0.24	0.11	0.23	0.06
40%	20%	2	-0.35	0.65	-0.37	0.60	-0.37	0.51	-0.09	0.02
40%	20%	7	0.04	0.37	0.13	0.21	0.22	0.12	0.23	0.06
40%	40%	2	-0.53	0.92	-0.40	0.66	-0.41	0.53	-0.09	0.03
40%	40%	7	0.03	0.39	0.07	0.30	0.17	0.15	0.23	0.06

Tabela 3.12: Viés e erro quadrático médio (EQM) da $E(T)$ estimada para os dados gerados da componente 3 que possui distribuição log-normal (simulação 2 -sistemas k-de-n e ponte (bridge)).

Cens. (dir)	Cens. (esq)	Verd. $E(T)$	n = 30		n = 50		n = 100		n = 1000	
			Viés	EQM	Viés	EQM	Viés	EQM	Viés	EQM
0%	0%	2	-0.12	0.17	-0.06	0.09	-0.00	0.04	-0.01	0.00
0%	0%	7	-0.03	0.16	0.01	0.09	0.05	0.04	0.02	0.01
0%	20%	2	-0.01	0.12	-0.02	0.08	0.02	0.04	0.02	0.01
0%	20%	7	0.10	0.17	0.14	0.13	0.17	0.08	0.14	0.02
0%	40%	2	-0.11	0.19	-0.02	0.09	0.06	0.04	0.06	0.01
0%	40%	7	0.24	0.27	0.28	0.22	0.31	0.15	0.28	0.08
20%	0%	2	0.07	0.14	0.17	0.10	0.25	0.08	0.25	0.07
20%	0%	7	0.08	0.15	0.14	0.10	0.18	0.06	0.16	0.03
20%	20%	2	-0.02	0.16	0.12	0.08	0.22	0.07	0.24	0.06
20%	20%	7	0.12	0.20	0.20	0.15	0.25	0.10	0.22	0.05
20%	40%	2	-0.14	0.32	0.02	0.16	0.19	0.06	0.23	0.05
20%	40%	7	0.17	0.28	0.26	0.21	0.32	0.15	0.28	0.09
40%	0%	2	0.21	0.23	0.33	0.18	0.42	0.20	0.44	0.19
40%	0%	7	0.22	0.21	0.28	0.18	0.34	0.15	0.33	0.11
40%	20%	2	0.06	0.33	0.10	0.26	0.32	0.15	0.38	0.15
40%	20%	7	0.14	0.26	0.23	0.21	0.33	0.16	0.33	0.11
40%	40%	2	-0.08	0.43	-0.05	0.33	0.16	0.14	0.33	0.11
40%	40%	7	0.10	0.38	0.17	0.27	0.32	0.16	0.34	0.12

3.4 Tempo de retratação de acordo com dois tratamentos

Recordamos o exemplo 2.2.5 em que descreve-se um exemplo de dados censurados intervalares na qual ocorre a visita periódica de um paciente com uma doença e avalia-se a eficácia do tratamento utilizado. Uma avaliação feita ao longo do tempo acompanhando o paciente e verificando os dados desde o início do tratamento é chamada de estudo prospectivo e uma avaliação baseada em dados de períodos passados já ocorridos chama-se retrospectiva.

Um paciente pode fazer diversas visitas ao longo de um tempo e em qualquer momento vir com um estado diferente. Para fins de análise utilizaremos o estudo descrito em ?, feito de modo retrospectivo. O estudo comentado nesse artigo compara pacientes com quadro recente câncer de mama que foram tratados com radioterapia e quimioterapia adjuvante com os tratados com apenas quimioterapia com respeito ao efeitos cosméticos do seu tratamento. O objectivo desse estudo foi comparar esses dois tratamentos para determinar se a quimioterapia afeta a razão de deterioração do estado cosmético. Nesse estudo, os pacientes fizeram visitas clínicas a cada período entre 4 a 6 meses. Esse tempo era frequentemente mais longo para pacientes que completavam a primeiro tratamento de irradiação e para os que viviam à grandes distâncias. São comparados o efeito do regime dos dois tratamentos no tempo até a deterioração cosmética, determinado pela aparência de retratação de mama. O estudo foi conduzido no Joint Center Radiation Therapy em Bostem de 1976 a 1980. A tabela 3.13 mostra os intervalos de tempo durante o resultado cosmético deteriorou-se. Por exemplo, se uma observação é igual a $(6, 14]$, então em 6 meses o paciente mostrou nenhuma deterioração no resultado cosmético inicial, mas em 14 meses, a retratação era presente. Esses dados visam comparar os dois tratamentos com respeito ao tempo até a deterioração cosmética. Valores com um sinal de + à direita são dados censurados à direita. Os dados intervalares com início no tempo 0 foram considerados dados censurados à esquerda.

Tabela 3.13: Intervalo de retratação para indivíduos com quadro recente de câncer de mama tratados com radioterapia apenas vs. radioterapia e quimioterapia.

Radioterapia			Radioterapia e Quimioterapia		
45+	(25,37]	37+	(8,12]	(0,5]	(30,34]
(6,10]	46+	(0,5]	(0,22]	(5,8]	13+
(0,7]	(26,40]	18+	(24,31]	(12,20]	(10,17]
46+	46+	24+	(17,27]	11+	(8,21]
46+	(27,34]	36+	(17,23]	(33,40]	(4,9]
(7,16]	(36,44]	(5,11]	(24,30]	31+	11+
17+	46+	(19,35]	(16,24]	(13,39]	(14,19]
(7,14]	(36,48]	(17,25]	13+	(19,32]	(4,8]
(37,44]	37+	24+	(11,13]	34+	34+
(0,8]	40+	32+	(16,20]	13+	(30,36]
(4,11]	(17,25]	33+	(18,25]	(16,24]	(18,24]
15+	46+	(19,26]	(17,26]	35+	(16,60]
(11,15]	(11,18]	37+	32+	(15,22]	(35,39]
22+	38+	34+	23+	(11,17]	21+
46+	(5,12]	36+	(44,48]	(22,32]	(11,20]
46+			(14,17]	(10,35]	48+

3.4.1 Resolução

A resolução do exemplo desta seção foi feita através do modelo bayesiano hierárquico weibul com censura intervalar generalizado em 3.2.1. Através desse modelo geral e utilizando o caso de um sistema com apenas uma componente calculamos separadamente os parâmetros e a curva de confiabilidade (ou retratação desse sistema). Esse sistema possui censura à direita, à esquerda e intervalar e estas proporções estão descritas na Tabela 3.14 a seguir. Observa-se que esse é um sistema com todos dados censurados e uma grande quantidade de dados com censura intervalar.

Tabela 3.14: Proporções de censura para o sistema em estudo com dois tratamentos.

	Proporções de censura		
	censura à direita	censura à esquerda	censura intervalar
Radio	54,35%	6,52%	39,13%
Radio e Químio	27,08%	4,17%	68,75%

Novamente, a estimação foi feita através dos algoritmos MCMC e EM. Foi gerada uma amostra tamanho 20000 e descartou-se os 10000 valores (burn-in). Utilizou-se uma amostra à posteriori de tamanho 1000 para a estimação das funções de confiabilidade e a região de credibilidade. Considerou-se saltos (jumps) entre os pontos amostrais de tamanho 10, em que descarta-se uma amostra de tamanho 10 entre cada um dos pontos amostrais. Determinou-se valores de $v_\beta = v_\eta = 4$ para todas componentes. Os resultados encontrados na simulação foram os seguintes: $\widehat{\mathbf{m}}_\beta = (1, 18; 2, 06)$ e $\widehat{\mathbf{m}}_\eta = (56, 70; 28, 30)$.

A Tabela 3.15 mostra as estimativas de média e desvio padrão à posteriori de cada parâmetro do modelo. Os desvios padrões são relativamente pequenos comparados com a grandeza das estimativas.

Tabela 3.15: Média à posteriori (desvio padrão-dp) de algumas estimativas de quantidades o estudo com dois tratamentos.

	Estimativas do estudo com dois tratamentos		
	β	η	$E(T \beta, \eta)$
Radio	1,12 (0,21)	56,76 (2,03)	56,02 (6,37)
Radio e Químio	2,00 (0,28)	28,40 (1,54)	25,20 (1,38)

As Figuras 3.8(a) e 3.8(b) mostram as estimativas com o intervalo de credibilidade de 95% para cada componente e também a curva de sobrevivência do estimador não-paramétrico de Turnbull. Esse estimador pode ser encontrado com mais detalhes em ?.

Figura 3.8: Funções de confiabilidade do estudo para os dois tratamentos: (a) Radioterapia; (b) Radioterapia e Quimioterapia.

Capítulo 4

Conclusões

“A satisfação está no esforço e não apenas na realização final.” (Mahatma Gandhi)

4.1 Considerações Finais e pesquisas futuras

Neste trabalho, foi introduzido um método de estimação weibull bayesiano hierárquico de sistemas coerentes. Foram considerados os três tipos de dados censurados, agregando tanto dados censurados à direita, à esquerda como também intervalares, inclusive podendo ter diferentes tipos de dados censurados no mesmo sistema. Os dados censurados são importantes e aplicados em diversas ciências, como medicina, economia, engenharia e sociologia. O enfoque nessa dissertação, foi de exemplos na área de confiabilidade, área da engenharia. Foram criadas simulações gerando posterioris através dos métodos de MCMC e EM para avaliar a eficácia do método sugerido. Também, houve a aplicação do método em um exemplo de dados censurados intervalares na área de sobrevivência.

As estimativas das funções de confiabilidade foram feitas através de funções à posteriori geradas, avaliando-as em diferentes tempos de confiabilidade. Essa forma de abordar é bem diferente da aproximação dita plug-in em que os valores estimados médios são simplesmente inseridos na fórmula de função de confiabilidade. Para sugestões futuras, pode-se ainda, aprimorar essa questão, considerando todas as curvas de confiabilidade possíveis através de técnicas da área dita Análise de Dados Funcionais.

Nos primeiros dois exemplos, foram avaliados dados censurados à direita e à esquerda. Estimou-se e avaliou-se diversos cenários possíveis com diferentes porcentagens de censura. Nesses sistemas permitia-se apenas um tipo censura nos sistemas, a censura à direita no sistema em série e a esquerda no sistema em paralelo. Sistemas em série são circuitos que estão funcionando se todos os componentes estão funcionando e sistemas em paralelo funcionam até que todos os componentes falhem. Fez-se uma simulação de um cenário desafiador com grande proporção de censura e uma simulação geral com diversas proporções de censura, distintas médias, três tipos de distribuições e tamanhos de amostra para avaliar a eficácia do modelo. Em todos casos, o modelo obteve ajustes com valores de Viés e EQM baixos comparados com os valores originais. Quanto maior o tamanho amostral, menor ficou a diferença de estimativas entre o estimado e o verdadeiro.

Um outro exemplo foi de sistemas k-de-n, sistemas que funcionam se pelo menos k componentes estão funcionando. Foi feita uma simulação para o sistema 2-de-3 e a curva de confiabilidade obtida foi bem próximo ao simulado. As simulações pelo método MCMC e EM feitas com modelos hierárquicos não apresentaram problemas de convergência como ocorria em contextos em que isso não era feito. Nesse sistema, existe a possibilidade do sistema ter um dado censurado à direita ou à esquerda. Mostrou-se um exemplo com os sistemas ponte (bridge), sistemas em que existe uma ponte entre dois ramos da estrutura. Essa forma cria diferentes possibilidades de falha do sistema. Anota-se a condição final de cada componente, ou seja, o seu funcionamento ou não-funcionamento e isso possibilita uma redução dessa estrutura mais complexa do bridge para um sistema em que cada componente é um dado censurado à direita ou à esquerda. Não havia ainda solução para

esse sistema para componentes com diferentes distribuições. Uma simulação avaliou a qualidade do modelo para um cenário com um sistema de ponte e obteve-se curvas de confiabilidade próximas as verdadeiras. Para avaliar de forma geral, criou-se uma simulação com diferentes proporções de censura variando de 0% à 80% divididos em censura à direita e à esquerda, diferentes médias, três tipos de distribuições e tamanhos de amostra. Os resultados foram resumidos através de quantidades como o Viés e o EQM, sendo esses valores todos próximos a 0, inclusive para altíssimas proporções de censura. Todos os dados foram resumidos em tabelas comparativas e gráficos.

Finalmente, introduziu-se um exemplo de um estudo de tempo de retratação de câncer de mama para dois tipos de tratamento. Esse exemplo conta com dados predominantemente intervalares e também com proporções de censura à direita e à esquerda, todos conjuntamente no mesmo exemplo. O modelo obteve estimativas com pequenos desvios padrões. Ressalta-se que um teorema importante da dissertação prova que a posteriori do modelo bayesiano hierárquico weibull para dados censurados intervalares, à direita e à esquerda é própria para uma classe de distribuições à priori não-informativas permitindo o uso desse modelo e estimação.

Dessa forma, conclui-se que o modelo com distribuição weibull comportou-se de forma altamente satisfatória à diferentes contextos e situações e demonstrou grande flexibilidade de adaptação a uma enorme gama de cenários.

Referências Bibliográficas

- A. Polpo e Pereira(2009)** M. Coque A. Polpo e C.A.B. Pereira. Statistical analysis for weibull distributions in presence of right and left censoring. *In The Proceedings of 8th International Conference on Reliability, Maintainability and Safety (ICRMS)*, 1:219–223. Citado na pág.
- A.S. Rodrigues e Polpo(2011)** M. Lauretto A.S. Rodrigues, T.C.M. Dias e A. Polpo. Reliability analysis in series systems: An empirical comparison between bayesian and classical estimators. *AIP Conference Proceedings*, 1443:214–221. Citado na pág.
- Bhattacharya e Samaniego(2010)** D. Bhattacharya e F. J. Samaniego. Estimating component characteristics from system failure-time data. *Wiley InterScience*. Citado na pág.
- Bhering et al.(2012)** F. L. Bhering, A. Polpo e C. Pereira. A hierarchical weibull bayesian model for right-censored and left-censored data. Em *XI BRAZILIAN MEETING ON BAYESIAN STATISTICS: EBEB 2012*, páginas 59–66. AIP. Citado na pág.
- Colosimo e Giolo(2006)** E. A. Colosimo e S. R. Giolo. *Análise de Sobrevida Aplicada*. Edgard Blucher. Citado na pág.
- Edwards et al.(2011)** C. Edwards, T. Lombaerts e H. Smaili. *Fault Tolerant Flight Control: A Benchmark Challenge*. Chapman and Hall/CRC. Citado na pág.
- Finkelstein e Wolfe(1985)** D. M. Finkelstein e R. A. Wolfe. A semiparametric model for regression analysis of interval-censored failure time data. *Biometrics*, 41. Citado na pág.
- Kaplan e Meier(1958)** E.L. Kaplan e P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481. Citado na pág.
- Polpo e Pereira(2009)** A. Polpo e C.A.B. Pereira. Reliability nonparametric bayesian estimation in parallel systems. *IEEE Transactions on Reliability*, 58(2):364–373. Citado na pág.
- Polpo e Sinha(2011)** A. Polpo e D. Sinha. Correction in bayesian nonparametric estimation in a series system or a competing-risks model. *Statistics and Probability Letters*, 81(12):1756–1759. Citado na pág.
- R.Barlow e Proschan(1981)** R.Barlow e F. Proschan. *Statistical Theory of Reliability and Life Testing*. To Begin with. Citado na pág.
- V. Salinas-Torres e Tiwari(2002)** C.A.B. Pereira V. Salinas-Torres e R. Tiwari. Bayesian nonparametric estimation in a series system or a competing risks model. *Journal of Nonparametric Statistics*, 14:449–458. Citado na pág.
- Williams(1990)** J. Williams. Bridge circuits marrying gain and balance. *Linear Technology Corp*, 43. Citado na pág.