



Universidade de Brasília  
Departamento de Estatística

Distribuição Weibull discreta exponenciada para dados com presença de  
censura: uma abordagem clássica e bayesiana

Marcílio Ramos Pereira Cardial

Brasília  
2017



Marcílio Ramos Pereira Cardial

**Distribuição Weibull discreta exponenciada para dados com presença de censura: uma abordagem clássica e bayesiana**

Orientador:

Prof. Dr. **Eduardo Yoshio Nakano**

Co-orientadora:

Profa. Dra. **Juliana Betini Fachini-Gomes**

Dissertação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas da Universidade de Brasília como requisito para obtenção do título de Mestre em Estatística.

**Brasília**  
**2017**



## Agradecimentos

Primeiramente, a Deus, por todas as conquistas que já me fez alcançar não deixando em nenhum momento me faltar saúde, paz e disposição para realizar as minhas tarefas diárias. Certamente sem sua presença em minha vida nada seria possível, muito menos a execução deste trabalho.

Agradeço a minha família, aos meus pais Domingos e Maria, que sempre me incentivaram a estudar e não desistir diante do primeiro obstáculo sempre me apoiando em minhas decisões. Também agradeço a Sirlene pela compreensão e companherismo.

Aos professores Eduardo Yoshio Nakano e Juliana Betinni Fachini-Gomes pela ótima orientação e por dedicar bastante tempo compartilhando seus conhecimentos comigo, sempre com muita paciência e disposição. Agradeço também ao professor Cleber Giuglioli Carrasco da Universidade Estadual de Goiás (UEG), que me incentivou a cursar uma Pós-Graduação em Estatística.

Aos meus amigos e colegas de Pós-graduação Alex e Luciana e de graduação Vando e Samuel pela amizade e companherismo indispensáveis.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa de estudos, que me auxiliou muito nesta jornada de estudos.

Por fim, à todos que direta ou indiretamente contribuíram para que esse objetivo fosse alcançado.



## Resumo

Distribuição Weibull discreta exponenciada para dados com presença de censura: uma abordagem clássica e bayesiana

Este trabalho faz um estudo inferencial da distribuição Weibull Discreta Exponenciada para dados de sobrevivência discretos. A distribuição em estudo é uma extensão da distribuição Weibull Discreta de Nakagawa e Osaki. Esta generalização da distribuição Weibull Discreta fornece um ajuste paramétrico em que se permite trabalhar com taxas de falha crescente, decrescente, constante, forma de banheira e unimodal. A formulação do modelo foi realizada no contexto de Análise de Sobrevivência, que considera a existência de dados censurados, e a inferência foi realizada nas abordagens clássica e bayesiana. Na abordagem clássica, a partir do método de máxima verossimilhança, foram utilizados métodos computacionais de otimização para encontrar as estimativas pontuais e intervalares e utilizou-se o Teste de Razão de Verossimilhanças (TRV) para realização de testes de hipóteses como forma de seleção de modelos. Na abordagem bayesiana foram utilizadas técnicas de MCMC (*Markov Chain Monte Carlo*) para obtenção das estimativas pontuais dos parâmetros do modelo e seus respectivos intervalos de credibilidade HPD (*Highest Posterior Density*) e o teste de significância genuinamente bayesiano (FBST - *Full Bayesian Significance Test*). A metodologia apresentada foi aplicada em dados simulados e ilustrada em dois conjuntos de dados reais: o primeiro refere-se ao tempo de sobrevivência de pacientes diagnosticados com câncer de pescoço e cabeça e o segundo considera o tempo até o alívio da dor em pacientes com dor lombar crônica não específica. Todas as simulações e estimativas foram obtidas por meio do *software* livre R.

**Palavras-chave:** Distribuição exponenciada; Distribuição discreta, Análise de sobrevivência; Dados censurados.



## Abstract

Exponentiated discrete Weibull distribution for censored data: a classical and bayesian approach

This work presented an inferential study of the Exponentiated Discrete Weibull distribution for discrete survival data. This distribution is an extension of the Nakagawa e Osaki's Discrete Weibull distribution. This generalization of the Discrete Weibull distribution provides a parametric adjustment that allows to work with increasing, decreasing, constant, bathtub shape and unimodal hazard rates. The model was formulated in the context of Survival Analysis, which considers the existence of censored data, and inferences were made in the classical and bayesian approaches. In the classical approach, from the maximum likelihood method, computational optimization methods were used to find the point and interval estimates and the Likelihood Ratio Test was used to perform hypothesis tests as a form of model selection. In addition, in the Bayesian approach, the use of the MCMC (Markov Chain Monte Carlo) techniques was used to obtain the point estimates of the model's parameters and their respective HPD (Highest Posterior Density) credible intervals and the FBST (Full Bayesian Significance Test). The methodology presented was applied in simulated data and illustrated in two real data sets: the first refers to the survival time of patients diagnosed with head and neck cancer and the second performed an analysis of the time to relieving pain in patients with chronic non-specific back pain. All simulations and estimates were obtained by the free software R.

**Keywords:** Exponentiated distribution; Discrete distribution; Survival analysis; Censored data.



---

## Lista de Abreviaturas e Siglas

EQM: Erro Quadrático Médio.

$f_w$ : Função densidade de probabilidade da distribuição Weibull.

$F_w$ : Função distribuição acumulada da Weibull.

$F_{wd}$ : Função distribuição acumulada da Weibull Discreta.

$F_{we}$ : Função distribuição acumulada da Weibull Exponenciada.

$F_{wde}$ : Função distribuição acumulada da Weibull Discreta Exponenciada.

FBST: *Full Bayesian Significance Test* (Teste de Significância Genuinamente Bayesiano).

GE: Geométrico(a) Exponenciado(a).

$h_w$ : Taxa de falha da distribuição Weibull.

$h_{wd}$ : Taxa de falha da distribuição Weibull Discreta.

$h_{wde}$ : Taxa de falha da distribuição Weibull Discreta Exponenciada.

HPD: *Highest Posterior Density* (Densidade a Posteriori Máxima).

K-M: Kaplan-Meier

MCMC: *Markov Chain Monte Carlo* (Monte Carlo via Cadeias de Markov).

$p_{wd}$ : Distribuição de probabilidades da Weibull Discreta.

$p_{wde}$ : Distribuição de probabilidades da Weibull Discreta Exponenciada.

$p_{wed}$ : Distribuição de probabilidades da Weibull Exponenciada Discreta.

$\hat{S}_{km}$ : Função de sobrevivência estimada via Kaplan-Meier

$S_w$ : Função de sobrevivência da distribuição Weibull.

$S_{wd}$ : Função de sobrevivência da distribuição Weibull Discreta.

$S_{wde}$ : Função de sobrevivência da distribuição Weibull Discreta Exponenciada.

TRV: Teste da Razão de Verossimilhanças.

WD: Weibull Discreto(a).

WE: Weibull Exponenciado(a).

WDE: Weibull Discreto(a) Exponenciado(a).

WED: Weibull Exponenciado(a) Discreto(a).



## Lista de Figuras

|    |   |    |
|----|---|----|
| 1  | Função densidade de probabilidade da distribuição Weibull . . . . .   | 27 |
| 2  | Função de sobrevivência da distribuição Weibull . . . . .   | 28 |
| 3  | Taxa de falha da distribuição Weibull . . . . .   | 29 |
| 4  | Distribuição de probabilidade da Weibull Discreta . . . . .   | 30 |
| 5  | Função de sobrevivência da distribuição Weibull Discreta . . . . .  | 31 |
| 6  | Taxa de falha da distribuição Weibull Discreta . . . . .  | 32 |
| 7  | Distribuição de probabilidade da Weibull Discreta Exponenciada . . . . .  | 33 |
| 8  | Função de sobrevivência da distribuição Weibull Discreta Exponenciada . .   | 34 |
| 9  | Taxa de falha da distribuição Weibull Discreta Exponenciada . . . . .   | 35 |
| 10 | Distribuições a posteriori marginais com diferentes distribuições a priori de $\gamma$ para o Cenário 1 . . . . .   | 47 |
| 11 | Função de sobrevivência do estimador K-M dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .                           | 49 |
| 12 | TTT plot dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .   | 50 |
| 13 | Ajuste dos modelos WDE, WD, GE e Geométrico aos dados (em meses) de pacientes diagnosticados com câncer de pescoço e cabeça . . . . .                           | 51 |
| 14 | Ajuste dos modelos WDE, WD, GE e Geométrico aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .                        | 53 |
| 15 | Ajuste do modelo WDE com intervalo de confiança 95% aos tempos de sobrevivência (em meses) de pacientes diagnosticados com câncer de pescoço e cabeça . . . . . | 54 |
| 16 | Taxa de falha ajustada pelo modelo WDE referentes aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .                  | 54 |
| 17 | Ajuste do modelo WDE aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça segundo a abordagem bayesiana . . .                     | 56 |
| 18 | Função de sobrevivência do estimador K-M dos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica                  | 57 |
| 19 | TTT plot referente aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .                              | 58 |
| 20 | Ajuste dos modelos WDE, WD, GE e Geométrico aos dados (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .                       | 59 |
| 21 | Ajuste dos modelos WDE, WD, GE e Geométrico aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .     | 60 |

|    |  |    |
|----|--|----|
| 22 | Ajuste dos modelos WDE, WD e GE com intervalo de confiança 95% aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . . | 62 |
| 23 | Taxa de falha ajustada pelos modelos WDE,WD e GE referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .    | 62 |
| 24 | Ajuste dos modelos WDE, WD e GE aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica segundo a abordagem bayesiana . . . . .  | 64 |

## Lista de Tabelas

|    |  |    |
|----|--|----|
| 1  | Cenários utilizados na simulação . . . . .   | 40 |
| 2  | Estimativas do vício e EQM do Cenário 1 . . . . .  | 42 |
| 3  | Estimativas do vício e EQM do Cenário 2 . . . . .  | 42 |
| 4  | Estimativas do vício e EQM do Cenário 3 . . . . .  | 43 |
| 5  | Estimativas do vício e EQM do Cenário 4 . . . . .  | 44 |
| 6  | Influência da escolha dos hiperparâmetros $a_2$ e $b_2$ da priori de $\gamma$ na inferência dos parâmetros da distribuição WDE para o Cenário 1. . . . .                     | 45 |
| 7  | Tempos de sobrevivência em dias de pacientes com câncer de pescoço e cabeça . . . . .  | 48 |
| 8  | Tempos de sobrevivência em meses de pacientes com câncer de pescoço e cabeça . . . . .   | 48 |
| 9  | Estimativas do modelo WDE referentes aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .  | 50 |
| 10 | Estimativas da função de sobrevivência dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .  | 51 |
| 11 | TRV para os modelos referentes aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .  | 53 |
| 12 | Inferência bayesiana dos parâmetros do modelo WDE dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça . . . . .                               | 55 |
| 13 | Tempos de sobrevivência referentes a Aplicação 2 . . . . .   | 57 |
| 14 | Estimativas dos modelos WDE,WD e GE referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .               | 58 |
| 15 | Estimativas da função de sobrevivência dos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .                       | 59 |
| 16 | TRV para os modelos referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .                               | 61 |
| 17 | Inferência bayesiana dos parâmetros dos modelos WDE, WD e GE dos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . . | 63 |
| 18 | FBST para os modelos referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica . . . . .                              | 63 |



# Sumário

|   |    |
|---|----|
| <b>1 Introdução</b> . . . . .                                 | 6  |
| <b>2 Revisão de Literatura</b> . . . . .                      | 8  |
| 2.1 Análise de Sobrevivência . . . . .                        | 8  |
| 2.1.1 Função de Probabilidade . . . . .                       | 9  |
| 2.1.2 Função de Sobrevivência . . . . .                       | 10 |
| 2.1.3 Taxa de Falha . . . . .                                 | 11 |
| 2.1.4 Estimador de Kaplan-Meier . . . . .                     | 11 |
| 2.2 Discretização de Distribuições Contínuas . . . . .        | 12 |
| 2.3 Inferência Clássica . . . . .                             | 13 |
| 2.3.1 Método de Máxima Verossimilhança . . . . .              | 13 |
| 2.3.2 Intervalo de Confiança . . . . .                        | 14 |
| 2.3.3 Teste de Hipóteses . . . . .                            | 15 |
| 2.4 Inferência Bayesiana . . . . .                            | 16 |
| 2.4.1 Teorema de Bayes . . . . .                              | 17 |
| 2.4.2 Prioris . . . . .                                       | 18 |
| 2.4.3 Estimativa Pontual . . . . .                            | 19 |
| 2.4.4 Estimativas Intervalares . . . . .                      | 20 |
| 2.4.5 Teste de Hipóteses Bayesianos . . . . .                 | 20 |
| 2.4.6 <i>Full Bayesian Significance Test</i> (FBST) . . . . . | 21 |
| 2.5 Classe de Distribuições Exponenciadas . . . . .           | 23 |
| <b>3 Metodologia</b> . . . . .                                | 26 |
| 3.1 Distribuição Weibull . . . . .                            | 26 |
| 3.1.1 Caso Contínuo . . . . .                                 | 27 |
| 3.1.2 Caso Discreto . . . . .                                 | 29 |
| 3.2 Distribuição Weibull Discreta Exponenciada . . . . .      | 32 |
| 3.2.1 Outra Formulação da WDE . . . . .                       | 36 |
| 3.3 Inferência Clássica . . . . .                             | 37 |
| 3.4 Inferência Bayesiana . . . . .                            | 38 |
| <b>4 Simulações Computacionais</b> . . . . .                  | 40 |
| 4.1 Vício e Erro Quadrático Médio (EQM) . . . . .             | 40 |
| 4.2 Estudo de Sensibilidade da Escolha da Priori . . . . .    | 44 |
| <b>5 Aplicações em dados reais</b> . . . . .                  | 48 |
| 5.1 Aplicação 1 . . . . .                                     | 48 |
| 5.1.1 Abordagem Clássica . . . . .                            | 50 |
| 5.1.2 Abordagem Bayesiana . . . . .                           | 55 |
| 5.2 Aplicação 2 . . . . .                                     | 56 |
| 5.2.1 Abordagem Clássica . . . . .                            | 58 |

---

|  |           |
|--|-----------|
| 5.2.2 Abordagem Bayesiana . . . . .  | 62        |
| <b>6 Considerações Finais . . . . .</b>                                    | <b>66</b> |
| <b>Referências . . . . .</b>   | <b>72</b> |
| <b>Apêndices . . . . .</b>   | <b>74</b> |
| A.1 Equações de Verossimilhança da WDE . . . . .                           | 74        |
| A.2 Intervalos de Confiança para os Parâmetros . . . . .                   | 75        |
| A.3 Logaritmo da Função de Verossimilhança Utilizada Computacionalmente. . | 78        |
| A.4 Scripts Desenvolvidos . . . . .  | 79        |

# 1 Introdução

A estatística é uma ciência baseada na teoria das probabilidades e que tem como objetivo apresentar conclusões em situações de incerteza a partir da amostra. Essa ciência entre outros tópicos envolve o planejamento do experimento a ser realizado, a coleta qualificada dos dados, a inferência, o processamento, a análise e a disseminação das informações. Uma das áreas da estatística é a análise de sobrevivência ou confiabilidade.

Segundo Nakano e Carrasco (2006) a análise de sobrevivência ou confiabilidade é um conjunto de técnicas e modelos estatísticos na análise de experimentos cuja variável resposta é o tempo até a ocorrência de um evento de interesse. Em análise de sobrevivência o evento de interesse em diversas situações é atribuído à morte de um indivíduo que também identifica-se como falha.

Uma das características da análise de sobrevivência é a presença de dados censurados, que são observações incompletas ou parciais da variável resposta. A omissão dos dados censurados pode acarretar conclusões viciadas.

Nesse contexto, uma das distribuições que mais vem se destacando na análise de dados de sobrevivência é a distribuição Weibull, devido à flexibilidade do ajuste de dados que é proporcionado por essa distribuição (Ryan, 2009). No entanto, conforme Carrasco *et al.* (2008), a distribuição Weibull não fornece um ajuste paramétrico razoável para fenômenos de modelagem com taxa de falhas não-monótona, como taxas de falha em forma de banheira e unimodal, que são comuns em análise de sobrevivência.

Um exemplo de taxa de falha em forma de banheira é a experiência da mortalidade humana, com uma alta taxa de mortalidade infantil que decai rapidamente, permanecendo neste nível durante alguns anos antes de crescer novamente. A taxa de falha unimodal pode ser observada no decurso de uma doença cuja mortalidade atinge um pico depois de algum período finito diminuindo gradualmente em seguida.

Nos últimos anos, novas classes de distribuições para dados contínuos foram propostas com base em modificações da distribuição Weibull para lidar com taxas de falha em forma de banheira e unimodal. Entre elas, estão a Weibull Exponenciada (WE) introduzida por Mudholkar e Srivastava (1993), a distribuição Weibull Estendida apresentada por Xie *et al.* (2002), a distribuição Weibull Modificada proposta por Lai *et al.* (2003), entre outras.

A distribuição Weibull e suas generalizações vêm sendo usadas ao longo dos anos em análise de sobrevivência quando os dados são contínuos. Entretanto, em alguns casos, esses são discretos. Com base em estudos apresentados por Nakano e Carrasco (2006), nem sempre é aceitável a utilização de um modelo contínuo para análise de dados discretos,

pois em alguns casos pode-se observar um resultado pouco satisfatório.

Nakagawa e Osaki (1975) propuseram a distribuição Weibull Discreta (WD) para corresponder à distribuição Weibull contínua. Ao considerar uma generalização da distribuição Weibull para dados discretos, Nekoukhou e Bidram (2015) propuseram a distribuição Weibull Discreta Exponenciada (WDE). No entanto, as inferências dos parâmetros do modelo apresentado pelos mesmos não consideraram uma possível presença de censuras.

Neste trabalho será considerada a distribuição Weibull Discreta Exponenciada, que é uma extensão da distribuição Weibull discreta. A distribuição em estudo será considerada no contexto de análise de sobrevivência permitindo a presença de dados censurados no processo inferencial, sendo esse realizado nas abordagens clássica e bayesiana.

Este trabalho está dividido em seis capítulos. No Capítulo 2, é apresentada uma revisão de literatura com os conceitos estatísticos básicos que serão utilizados posteriormente no Capítulo 3, que refere-se a metodologia adotada para a formulação do modelo WDE segundo as abordagens clássica e bayesiana no contexto de análise de sobrevivência. As simulações e seus resultados são apresentados no Capítulo 4, enquanto que o Capítulo 5 ilustra duas aplicações do modelo proposto em conjuntos de dados reais. Por fim, no Capítulo 6 são apresentadas as considerações finais do trabalho.

## 2 Revisão de Literatura

No presente capítulo, será feita uma revisão bibliográfica. Na Seção 2.1, são apresentados os principais conceitos de análise de sobrevivência, as funções que a caracterizam considerando os casos contínuo e discreto, bem como o estimador de Kaplan-Meier. A Seção 2.2 apresenta a discretização de distribuições contínuas. A Seção 2.3 refere-se à análise de sobrevivência a partir de uma abordagem clássica (frequentista) e a Seção 2.4 a partir de uma abordagem bayesiana. Por fim, na Seção 2.5 mostra-se como é obtida uma distribuição exponenciada a partir de uma distribuição base. Nas Seções 2.1 e 2.3 utilizou-se como referência principal: Colosimo e Giolo (2006).

### 2.1 Análise de Sobrevivência

A análise de sobrevivência pode ser definida como um conjunto de procedimentos estatísticos para análise de dados relacionados a variável resposta, tempo até a ocorrência de um determinado evento de interesse. A análise de sobrevivência é muito utilizada na área médica, na qual indivíduos submetidos a um determinado evento de interesse são colocados em estudo a fim de estimar o tempo de vida dos mesmos. Porém, essa técnica pode ser utilizada em outras áreas como a engenharia, na qual são realizados estudos em que os produtos são colocados sob testes para estimar características relacionadas aos seus tempos de vida. Nesse caso, é conhecida como análise de confiabilidade.

O estudo em análise de sobrevivência deve ser feito de forma cuidadosa pelo fato da possível presença de dados censurados, que é a existência de observações incompletas ou parciais, ou seja, quando não se observa o evento de interesse. É importante ressaltar que as censuras ocorrem, pois nem sempre o evento de interesse é observado para todos os elementos de um determinado estudo.

Sendo assim, há a necessidade da introdução de uma variável que indique se o tempo de vida foi ou não observado. Essa variável é definida na literatura como variável indicadora de censura ou falha, e será igual a um se o tempo de sobrevivência for observado e zero caso contrário. A variável indicadora de censura denotada por  $\delta$ , é definida por:

$$\delta = \begin{cases} 0, & \text{se o tempo é censurado} \\ 1, & \text{se o tempo é observado} \end{cases}$$

É válido lembrar que, mesmo censurados, todos os resultados provenientes de um estudo devem ser usados na análise estatística, pois a omissão da censura nos cálculos estatísticos poderá acarretar em estimativas viciadas.

Quanto à classificação dos tipos de censura, pode-se dividi-las em três grupos que são a censura à direita, à esquerda e intervalar. Na censura à direita, o tempo de interesse está à direita do tempo registrado, enquanto que na censura à esquerda o evento de interesse está à esquerda do tempo registrado. Na censura intervalar, quando os indivíduos são acompanhados em visitas periódicas, sabe-se apenas que o evento de interesse ocorreu em algum momento entre duas visitas.

Ainda, pode-se dividir a censura à direita em três subgrupos:

- Censura do Tipo I: o estudo termina após um período pré-estabelecido de tempo.
- Censura do Tipo II: o estudo termina após ter ocorrido o evento de interesse em um número pré-estabelecido de indivíduos estudados.
- Censura do Tipo Aleatório: ocorre quando um indivíduo é retirado no decorrer do estudo sem ter ocorrido a falha. Também ocorre, por exemplo, se o indivíduo falhar por uma razão diferente da estudada.

No presente trabalho, será considerada apenas a censura à direita aleatória, que é a mais utilizada em trabalhos de análise de sobrevivência.

### 2.1.1 Função de Probabilidade

Para uma variável aleatória contínua  $T$ , define-se  $f(\cdot)$  como uma função densidade de probabilidade se a mesma satisfizer as seguintes condições (Magalhães e Lima, 2013):

i)  $f(t) \geq 0$ , para todo  $t \in \mathbb{R}$ ;

ii)  $\int_{-\infty}^{\infty} f(t)dt = 1$ .

Sendo assim,  $f$  é uma função não negativa utilizada para representar a distribuição de probabilidade de uma variável aleatória contínua  $T$ . Para calcular a probabilidade da variável aleatória contínua  $T$  estar no intervalo  $[a, b]$ , basta calcular a seguinte integral:

$$P(a \leq T \leq b) = \int_a^b f(t)dt.$$

A variável aleatória contínua  $T$  tem um número infinito não enumerável de valores possíveis e a probabilidade dessa assumir um determinado valor  $t$  é zero. Por esta razão, a função densidade de probabilidade não é utilizada em valores pontuais e sim em intervalos.

Ao contrário de uma variável aleatória contínua, uma variável aleatória discreta

possui um conjunto enumerável de valores possíveis, sendo em muitos casos, definida para números inteiros positivos, ou seja,  $t = 0, 1, 2, \dots$ .

Segundo Magalhães e Lima (2013), ao considerar uma variável aleatória discreta  $T$ , a função que atribui a cada valor da variável aleatória a sua probabilidade é denominada de função discreta de probabilidade ou distribuição de probabilidade, que é representada por  $p(t)$ . Sendo assim,  $p(t) = P(T = t)$  deve satisfazer as seguintes condições:

$$i) 0 \leq p(t) \leq 1, \quad t = 0, 1, 2, \dots;$$

$$ii) \sum_{t=0}^{\infty} p(t) = 1.$$

### 2.1.2 Função de Sobrevivência

A função de sobrevivência no tempo  $t$  é a capacidade de um indivíduo sobreviver até o tempo  $t$ . Em outras palavras, é a probabilidade do indivíduo não falhar no intervalo de  $[0, t]$ .

Seja  $T$  uma variável aleatória contínua positiva ( $T > 0$ ). Define-se  $S(t)$  como a função de sobrevivência de um indivíduo e pode-se escrever  $S(t)$  em termos probabilísticos como:

$$\begin{aligned} S(t) &= P(T > t) \\ &= \int_t^{\infty} f(u) du, \quad t > 0. \end{aligned}$$

Aqui,  $f(\cdot)$  é a função densidade de probabilidade de  $T$ .

Conseqüentemente, a função de distribuição acumulada é a probabilidade do evento complementar do indivíduo sobreviver até o tempo  $t$ . Em termos probabilísticos, tem-se  $F(t) = 1 - S(t)$ .

No caso em que  $T$  é uma variável aleatória discreta não negativa, que assume apenas valores inteiros ( $t = 0, 1, 2, \dots$ ), tem-se que a função de sobrevivência pode ser definida como (Fernandes, 2013):

$$\begin{aligned} S_d(t) &= P(T > t) \\ &= \sum_{k=t+1}^{\infty} P(T = k), \quad t = 0, 1, 2, \dots \end{aligned}$$

### 2.1.3 Taxa de Falha

A taxa de falha, também conhecida como função de risco, é representada por  $h(t)$  e especifica a taxa de falha instantânea de um indivíduo, ou seja, a probabilidade do mesmo vir a falhar instantaneamente no tempo  $t$ , dado que este sobreviveu até o tempo  $t$ . A taxa de falha é dada por:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}. \quad (1)$$

A taxa de falha é utilizada para representar o comportamento da variável tempo de sobrevivência. Essa função pode assumir forma crescente, decrescente ou constante, ou ainda, vir a assumir forma não-monótona tendo a forma de banheira ou unimodal.

Pode-se utilizar uma relação matemática que descreve a taxa de falha (1), relacionando a função densidade de probabilidade com a função de sobrevivência. Para distribuições contínuas, a relação é expressa por:

$$h(t) = \frac{f(t)}{S(t)}.$$

Segundo Fernandes (2013), no caso em que as distribuições de probabilidade são discretas, a função de risco é igual a zero, exceto nos pontos em que pode ocorrer uma falha. A função de risco para dados discretos, denotada por  $h_d(t)$  e definida no intervalo  $0 < h_d(t) < 1$ , pode ser expressa como:

$$\begin{aligned} h_d(t) &= P(T = t | T \geq t) \\ &= \frac{P(T = t)}{P(T \geq t)} \\ &= \frac{P(T = t)}{P(T > t) + P(T = t)} \\ &= \frac{p(t)}{S_d(t) + p(t)}, \quad t = 0, 1, 2, \dots \end{aligned}$$

### 2.1.4 Estimador de Kaplan-Meier

O estimador de Kaplan-Meier é um estimador não paramétrico para a função de sobrevivência, é também chamado de estimador limite-produto. Segundo Kaplan e Meier (1958), suponha que existam “ $n$ ” indivíduos sob teste e “ $k$ ” falhas distintas nos tempos  $t_1 < t_2 < \dots < t_k$ , para  $k \leq n$ , podendo ocorrer mais de uma falha num mesmo tempo

(empate). O estimador de Kaplan-Meier da função de sobrevivência é dado por:

$$\hat{S}_{km}(t) = \prod_{i:t_i \leq t} \left(1 - \frac{d_i}{n_i}\right), \quad (2)$$

em que  $d_i$  é o número de falhas no tempo  $t_i$  e  $n_i$  é o número de observações sob risco no tempo  $t_i$ ,  $i = 1, 2, \dots, k$ .

O estimador de Kaplan-Meier baseia-se no fato de que o número de intervalos é igual ao número de tempos de falhas distintas e, por esse motivo, o gráfico vem a ser uma função escada com degraus nos tempos observados de falha.

O estimador de Kaplan-Meier é mais eficiente que os outros estimadores não paramétricos por ser um estimador não viciado para a função de sobrevivência, tanto para grandes como para pequenas amostras (Kaplan e Meier, 1958). Além disso, esse estimador é fracamente consistente, converge assintoticamente para um processo gaussiano e é estimador de máxima verossimilhança da função de sobrevivência.

## 2.2 Discretização de Distribuições Contínuas

Uma variável aleatória é definida como sendo uma quantidade associada a cada possível resultado do espaço amostral. Caso uma variável aleatória assuma valores em um conjunto enumerável essa é denominada variável aleatória discreta. Entretanto, se o conjunto de valores é qualquer intervalo dos números reais, que seria um conjunto não enumerável, a quantidade é denominada variável aleatória contínua.

A variável aleatória proposta é o tempo até a ocorrência de um evento de interesse. Em geral, esse tempo é tratado na escala contínua. No entanto, há situações em que esse tempo é registrado em uma escala discreta, por exemplo, quando o tempo é medido em ciclos ou número de sessões de um tratamento (contagem).

Modelos de variáveis discretas podem ser obtidas a partir de todos os modelos de variáveis aleatórias contínuas, agrupando os tempos em intervalos unitários.

Seja  $X$  uma variável aleatória contínua. A variável aleatória discreta é dada por  $T = [X]$ , em que  $[X]$  representa “a parte inteira de  $X$ ”. Se  $F_X(x)$  é a função de distribuição acumulada de  $X$ , a distribuição de probabilidade de  $T$ , representada por  $p(t)$ , pode ser escrita como (Nakano e Carrasco, 2006):

$$\begin{aligned} p(t) &= P(T = t) = P(t \leq X < t + 1), \\ &= F_X(t + 1) - F_X(t) \quad t = 0, 1, 2, \dots \end{aligned} \quad (3)$$

## 2.3 Inferência Clássica

O objetivo da inferência estatística é inferir conclusões sobre alguns aspectos da população a partir de uma amostra representativa. Normalmente, a inferência preocupa-se com o valor de um ou mais parâmetros<sup>1</sup> desconhecidos. Na inferência estatística, existem duas abordagens: a inferência clássica (frequentista), descrita nesta seção, e a inferência bayesiana descrita na seção subsequente.

O processo inferencial segundo a abordagem frequentista é realizado a partir da função de verossimilhança. Entretanto, no contexto de análise de sobrevivência, deve-se considerar uma função de verossimilhança que permita a presença de dados censurados. Desta forma, nesta seção são apresentados métodos para encontrar a estimativa pontual obtida a partir do método de máxima verossimilhança, estimativas intervalares obtidas a partir do intervalo de confiança e o teste de hipóteses para testar hipóteses em relação aos parâmetros. O processo inferencial que será descrito nas próximas seções é válido para dados contínuos e discretos. Como neste trabalho serão abordados dados discretos, a notação utilizada será  $p(t)$  para a distribuição de probabilidade e  $S_d(t)$  para a função de sobrevivência.

### 2.3.1 Método de Máxima Verossimilhança

Seja  $t_1, t_2, \dots, t_n$  uma amostra aleatória observada de uma variável aleatória discreta  $T$ ;  $p(t; \boldsymbol{\theta})$  representa a função de probabilidade da população e  $\boldsymbol{\theta}$  é o vetor de parâmetros da função de probabilidade. A função de verossimilhança para  $\boldsymbol{\theta}$  é dada por:

$$L(\mathbf{t}|\boldsymbol{\theta}) = \prod_{i=1}^n p(t_i; \boldsymbol{\theta}). \quad (4)$$

A expressão (4) é utilizada considerando que todos os elementos da amostra experimentaram o evento de interesse. Entretanto, em análise de sobrevivência possivelmente alguns desses elementos não sofreram o evento de interesse. Desta forma, os dados censurados, como são tecnicamente chamados, devem ser separados dos elementos que sofreram o evento de interesse frequentemente chamados de dados não-censurados.

Dessa forma, as observações são divididas em dois conjuntos, as  $r$  primeiras são as não-censuradas ( $1, 2, \dots, r$ ) e sua contribuição para a função de verossimilhança é dada por  $p(t_i; \boldsymbol{\theta})$ , e as  $n - r$  seguintes são as censuradas ( $r + 1, r + 2, \dots, n$ ), sendo sua contribuição para função de verossimilhança dada pela função de sobrevivência  $S_d(t_i; \boldsymbol{\theta})$ . Assim, a

---

<sup>1</sup>funções de valores populacionais

função de verossimilhança assume a seguinte forma:

$$L(\mathbf{t}|\boldsymbol{\theta}) \propto \prod_{i=1}^r p(t_i; \boldsymbol{\theta}) \prod_{i=r+1}^n S_d(t_i; \boldsymbol{\theta}). \quad (5)$$

A expressão (5) é equivalente a:

$$L(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta}) \propto \prod_{i=1}^n [p(t_i; \boldsymbol{\theta})]^{\delta_i} [S_d(t_i; \boldsymbol{\theta})]^{1-\delta_i}, \quad (6)$$

em que  $\delta_i$  é a variável indicadora de falha, que assume valor 1 se o tempo  $t_i$  for de falha e 0 se for censura à direita;  $p(t_i; \boldsymbol{\theta})$  e  $S_d(t_i; \boldsymbol{\theta})$  são, respectivamente, a distribuição de probabilidade e função de sobrevivência do modelo considerado.

Ao aplicar o logaritmo na função de verossimilhança na equação (6), tem-se:

$$\ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta}) = \sum_{i=1}^n [\delta_i \log [p(t_i; \boldsymbol{\theta})] + (1 - \delta_i) \log [S_d(t_i; \boldsymbol{\theta})]] + c. \quad (7)$$

em que  $c$  é uma constante que não depende de  $\boldsymbol{\theta}$ .

Os estimadores de máxima verossimilhança são os valores que maximizam  $L(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})$  ou equivalentemente  $\ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})$ , normalmente representado por  $\hat{\boldsymbol{\theta}}$ , e são obtidos resolvendo o sistema de equações:

$$\frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0}. \quad (8)$$

### 2.3.2 Intervalo de Confiança

A construção de intervalos de confiança é feita a partir da matriz informação de Fisher definida por:

$$I_f(\boldsymbol{\theta}) = E \left[ \left( \frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^2 \right] = - \frac{\partial^2 \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}.$$

Sob certas condições de regularidade tem-se que:

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

em que  $I_f(\hat{\boldsymbol{\theta}})$  é a informação de Fisher observada da amostra.

Para obter intervalos de confiança utiliza-se então, a distribuição assintótica do estimador de máxima verossimilhança. Para grandes amostras, essa propriedade estabelece

que:

$$\hat{\boldsymbol{\theta}} \stackrel{a}{\sim} N_k \left( \boldsymbol{\theta}, \text{Var}(\hat{\boldsymbol{\theta}}) \right),$$

ou seja,  $\hat{\boldsymbol{\theta}}$  converge assintoticamente para uma distribuição normal multivariada com média  $\boldsymbol{\theta}$  e matriz de variância e covariância  $\text{Var}(\hat{\boldsymbol{\theta}})$ , em que  $k$  é a dimensão de  $\hat{\boldsymbol{\theta}}$ .

Assim, um intervalo aproximado de  $(1 - \alpha) \times 100\%$  de confiança para o parâmetro  $\theta_j$ ,  $j = 1, 2, \dots, k$ , é dado por:

$$\left[ \hat{\theta}_j - Z_{(1-\alpha/2)} \sqrt{\text{Var}(\hat{\theta}_j)}; \hat{\theta}_j + Z_{(1-\alpha/2)} \sqrt{\text{Var}(\hat{\theta}_j)} \right], \quad j = 1, 2, \dots, k, \quad (9)$$

em que  $\hat{\theta}_j$  é o estimador de máxima verossimilhança de  $\theta_j$ ,  $Z_{(1-\alpha/2)}$  é o quantil  $(1 - \alpha/2)$  de uma distribuição normal padrão e  $\text{Var}(\hat{\theta}_j)$  é a estimativa da variância de  $\theta_j$ .

De forma análoga, um intervalo de confiança para uma função real 1:1 (inversível)  $g(\theta)$  definida em  $\Theta$ , pode ser obtido por meio do *método delta* que estabelece que:

$$\widehat{g(\theta)} \stackrel{a}{\sim} N \left( g(\theta), \left[ g'(\hat{\theta}) \right]^2 \text{Var}(\hat{\theta}) \right), \quad (10)$$

em que  $\widehat{g(\theta)} = g(\hat{\theta})$  é o estimador de máxima verossimilhança de  $g(\theta)$  e  $g'(\hat{\theta})$  é a derivada de primeira ordem de  $g(\hat{\theta})$ .

A expressão (10) é utilizada quando se há o interesse em calcular intervalos de confiança para funções dos parâmetros de uma determinada distribuição.

### 2.3.3 Teste de Hipóteses

Para um modelo com um vetor  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)'$  de parâmetros, em algumas situações deseja-se testar hipóteses relacionadas a este vetor ou a um subconjunto dele.

O teste da razão de verossimilhanças (TRV) é baseado na função de verossimilhança e envolve a comparação dos valores do logaritmo da função de verossimilhança maximizada e sob  $H_0$ . A estatística do teste é dada por:

$$TRV = -2 \log \left[ \frac{L(\mathbf{t}, \boldsymbol{\delta} | \hat{\boldsymbol{\theta}}_0)}{L(\mathbf{t}, \boldsymbol{\delta} | \hat{\boldsymbol{\theta}})} \right] = 2[\log L(\mathbf{t}, \boldsymbol{\delta} | \hat{\boldsymbol{\theta}}) - \log L(\mathbf{t}, \boldsymbol{\delta} | \hat{\boldsymbol{\theta}}_0)], \quad (11)$$

que, sob  $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ , segue aproximadamente uma distribuição qui-quadrado com  $\rho$  graus de liberdade que representa a diferença do número de parâmetros dos modelos (modelo completo e modelo sob  $H_0$ ). Para amostras grandes,  $H_0$  é rejeitada, a um nível de  $\alpha \times 100\%$

de significância, se  $TRV > \chi_{\rho,1-\alpha}^2$ . Aqui,  $\chi_{\rho,1-\alpha}^2$  é o quantil  $(1 - \alpha)$  de uma distribuição qui-quadrado com  $\rho$  graus de liberdade.

## 2.4 Inferência Bayesiana

A essência da estatística do ponto de vista clássico (frequentista) está em pensar no processo aleatório que produziu os dados observados, de forma a reduzir o desconhecimento em relação a uma quantidade de interesse, que é atribuída ao parâmetro  $\theta$ , podendo este desconhecimento assumir diferentes graus. Já a estatística do ponto de vista bayesiano, os graus de incerteza são representados por modelos probabilísticos para  $\theta$ , sendo assim, as quantidades observáveis e os parâmetros de um modelo estatístico segundo esta abordagem são todas aleatórias.

Para fixar as principais diferenças entre as duas metodologias, pode-se tomar como exemplo, o lançamento de um dado honesto, que tem como resultados possíveis: 1, 2, 3, 4, 5 e 6. Na inferência clássica, o processo aleatório é repetido um número infinito de vezes sob as mesmas condições, baseando-se na frequência relativa da ocorrência de eventos, quando são realizadas várias repetições do evento de interesse. Assim, para o exemplo do dado, as observações consistem do número de ocorrência da face um, representado por  $y$ , em  $n$  lançamentos. A proporção da face um é  $y/n$ . Os frequentistas formulam o conceito de probabilidade de face um,  $\theta$ , como o valor de  $y/n$ , com  $n$  tendendo ao infinito. Conseqüentemente, o parâmetro  $\theta$  pode ser estimado a partir das observações visualizadas. Perceba que desta forma o parâmetro  $\theta$  se torna fixo.

Na inferência bayesiana, a informação sobre  $\theta$  é aleatória, sendo que esta é obtida através de informações que se tem em relação aos dados ou pela informação de um especialista. Esta informação é resumida probabilisticamente sendo representada por  $\pi(\theta)$ , conhecida como distribuição a priori de  $\theta$ .

Segundo Dobson e Barnett (2008), pode-se diferenciar as inferências clássica e bayesiana através de duas estatísticas bastante conhecidas: valor-p e o intervalo de confiança. Essas diferenças são apresentadas a seguir:

### Inferência Clássica

- O valor-p é a probabilidade de observar dados mais extremos do que aqueles observados na amostra dado que a hipótese nula é verdadeira.
- O intervalo de confiança de  $(1 - \alpha) \times 100\%$  é um intervalo que contém o verdadeiro valor em  $(1 - \alpha) \times 100\%$  das ocasiões, se o processo aleatório pudesse ser repetido várias vezes. Por definição, um intervalo de confiança de  $(1 - \alpha) \times 100\%$  não inclui o verdadeiro valor em  $\alpha \times 100\%$  das ocasiões.

As definições anteriores dependem do conceito de múltiplas repetições de processos aleatórios que gerem os dados. Alternativamente, seriam mais naturais, definições que dependam apenas dos dados já observados, conforme apresentado a seguir:

### Inferência Bayesiana

- A decisão é tomada a partir da probabilidade estimada da hipótese nula ser verdadeira, dado os dados observados.
- O intervalo de credibilidade de  $(1-\alpha) \times 100\%$  é um intervalo que contém o verdadeiro valor com probabilidade de  $(1-\alpha)$ .

Sendo assim, conforme Dobson e Barnett (2008), as definições ideais são possíveis a partir de um paradigma estatístico alternativo conhecido como inferência bayesiana.

Carlin e Louis (2000) listam algumas vantagens da utilização de inferência bayesiana:

- Métodos bayesianos fornecem ao usuário a capacidade de incorporar formalmente informação prévia;
- As inferências são condicionais sobre os dados observados;
- A razão para parar a experimentação não afeta a inferência bayesiana;
- Respostas bayesianas são mais facilmente interpretadas por não especialistas;
- Toda a análise bayesiana segue diretamente a partir da distribuição a posteriori; não são necessárias quaisquer teorias distintas de avaliação, testes, comparações múltiplas, etc;
- Qualquer pergunta pode ser respondida diretamente por meio da análise bayesiana;
- procedimentos bayesianos e bayesianos empíricos possuem automaticamente várias propriedades de otimização.

A seguir são apresentados o Teorema de Bayes, o conceito de prioris além de métodos para encontrar: a estimativa pontual, estimativas intervalares e teste de hipóteses bayesianos obtidos a partir da distribuição a posteriori.

#### 2.4.1 Teorema de Bayes

Considere uma quantidade de interesse desconhecida  $\theta$ . A informação a priori que se dispõe sobre  $\theta$ , denotada aqui por  $\pi(\theta)$ , pode ser aumentada observando uma quantidade aleatória  $T$  relacionada com  $\theta$ . A distribuição amostral (verossimilhança)  $L(t|\theta)$  define

esta relação. A ideia de que, após observar  $T = t$ , a quantidade de informação sobre  $\theta$  aumenta é bastante intuitiva e a fórmula de Bayes é a regra de atualização utilizada para quantificar esse aumento de informação. Assumindo que  $\theta$  é um parâmetro contínuo, tem-se que a fórmula de Bayes pode ser descrita como:

$$\pi(\theta|t) = \frac{\pi(t, \theta)}{\pi(t)} = \frac{L(t|\theta)\pi(\theta)}{\pi(t)} = \frac{L(t|\theta)\pi(\theta)}{\int \pi(\theta, t)d\theta}. \quad (12)$$

Note que  $1/\pi(t)$ , que não depende de  $\theta$ , funciona como uma constante normalizadora de  $\pi(\theta|t)$  (Ehlers, 2011).

Sendo assim, a regra de atualização é formada pela função de verossimilhança e a distribuição a priori de  $\theta$ . Segundo Dobson e Barnett (2008), uma vez que a constante normalizadora é a mesma para cada valor do numerador, pode-se simplificar a expressão (12) para:

$$\pi(\theta|t) \propto L(t|\theta)\pi(\theta), \quad (13)$$

em que  $\propto$  significa “proporcional a”. A equação (13) é chamada distribuição a posteriori de  $\theta$ , sendo assim, pode-se descrever esta expressão como:

$$\text{distribuição a posteriori} \propto \text{verossimilhança} \times \text{distribuição a priori.}$$

Em resumo, segundo Paulino *et al.* (2003), para os bayesianos a distribuição a posteriori incorpora, via do Teorema de Bayes, toda informação disponível sobre o parâmetro: informação inicial (priori) + informação da experiência ou da amostra (verossimilhança).

Conforme Gamerman e Lopes (2006), o conceito de priori e posteriori estão sempre relacionados à observação considerada em um dado momento. É possível que, após observar  $t$  e obtendo uma distribuição a posteriori, uma nova observação  $y$  também relacionada com  $\theta$  através de uma eventualmente diferente função de verossimilhança possa ser realizada. Neste caso a posteriori relativa a  $t$  é a priori relativa a  $y$  e a nova posteriori pode ser obtida por meio de uma nova aplicação do teorema de Bayes.

É importante ressaltar que a distribuição a posteriori é de suma importância na inferência bayesiana, pois toda a análise é feita a partir da mesma.

## 2.4.2 Prioris

Certamente, uma das principais vantagens da abordagem bayesiana é a incorporação de informação prévia ou externa à experimentação. A utilização de tal informação em inferência bayesiana requer a especificação de uma distribuição a priori para a quantidade

de interesse  $\theta$ .

O conhecimento adquirido previamente sobre  $\theta$ , por sua vez, é representado por uma função densidade de um modelo paramétrico, cujos parâmetros devem ser informados de acordo com este conhecimento. Sendo assim, os parâmetros indexadores das funções densidade são chamados de hiperparâmetros.

Segundo Ehlers (2011), a ideia é que as distribuições a priori e a posteriori pertençam a mesma classe de distribuições e assim a atualização do conhecimento que se tem de  $\theta$  envolve apenas uma mudança nos hiperparâmetros, desta forma, o aspecto sequencial da abordagem bayesiana pode ser explorada definindo apenas a regra de atualização dos hiperparâmetros uma vez que as distribuições permanecem as mesmas.

Esta ideia, conforme Gamerman e Lopes (2006), traz vantagens para o procedimento inferencial resultante produzido devido à simplificação da análise que se torna restrita a um subconjunto de todas as distribuições. Entretanto, deve-se ter cuidado com a utilização indiscriminada de prioris conjugadas pelo fato dessas nem sempre serem a representação adequada da incerteza a priori. Sua utilização está muitas vezes associada à tratabilidade analítica decorrente.

### 2.4.3 Estimativa Pontual

Enquanto que na abordagem frequentista a estimativa pontual  $\hat{\theta}$  é usualmente obtida através do método de máxima verossimilhança da quantidade desconhecida, porém fixa  $\theta$ , na abordagem bayesiana a estimativa pontual  $\tilde{\theta}$  é encontrada a partir algum ponto da posteriori  $\pi(\theta|t)$ . Em geral este ponto é uma medida de resumo da mesma, tal como a média, mediana ou moda, que está diretamente ligada a uma função de perda.

Para cada possível valor de  $\theta$  e cada possível estimativa  $a \in \Theta$ , associa-se uma perda  $\mathcal{L}(a, \theta)$  de modo que quanto maior a distância entre  $a$  e  $\theta$  maior o valor da perda. Neste caso a perda esperada a posteriori é dada por:

$$E[\mathcal{L}(a, \theta|t)] = \int \mathcal{L}(a, \theta)\pi(\theta|t)d\theta,$$

e a regra de Bayes consiste em escolher o valor de  $a$  que minimiza esta perda esperada (Ehlers, 2011).

Ao considerar, a função de perda quadrática

$$\mathcal{L}(a, \theta) = (a - \theta)^2,$$

o estimador de Bayes de  $\theta$  é a média a posteriori, expressa por:

$$\tilde{\theta} = E(\theta|t).$$

#### 2.4.4 Estimativas Intervalares

Quando se é realizada a estimativa pontual no ponto de vista bayesiano, toda informação da distribuição posteriori está abreviada em um ou mais números dependendo da quantidade de parâmetros. Devido à necessidade de informar o quão precisas são estas estimativas, é estudado o intervalo de credibilidade ou intervalo de confiança bayesiano.

O intervalo de credibilidade fornece a probabilidade do parâmetro pertencer ao intervalo, assim, pode-se especificar um intervalo de probabilidade  $1 - \alpha$ , a partir da distribuição a posteriori.

Diz-se que  $[\tau_1; \tau_2]$  é um intervalo de credibilidade para  $\theta$ , com coeficiente de confiança  $1 - \alpha$  se (Bolfarine e Sandoval, 2010):

$$\int_{\tau_1}^{\tau_2} \pi(\theta|t)d\theta = 1 - \alpha. \quad (14)$$

Note que existem infinitos intervalos  $[\tau_1; \tau_2]$  que satisfazem (14). No entanto, pode ser desejável que o comprimento deste intervalo seja o mínimo possível. Quando a distribuição a posteriori é simétrica, geralmente os intervalos simétricos são os de menor comprimento. O intervalo de confiança bayesiano de menor comprimento é conhecido como o intervalo de densidade a posteriori máxima “*highest posterior density*” (HPD). Métodos computacionais são, em geral, necessários para obtenção do intervalo HPD.

É pacífico afirmar, conforme Paulino *et al.* (2003), que o intervalo de credibilidade HPD não coincide em geral com o intervalo de confiança clássico de comprimento mínimo.

#### 2.4.5 Teste de Hipóteses Bayesianos

O problema de testar  $H_0 : \theta \in \Theta_0$  versus  $H_1 : \theta \in \Theta_1 = \Theta - \Theta_0$  utilizando a abordagem bayesiana é conceitualmente mais simples do que na abordagem clássica.

Atendendo à interpretação direta das hipóteses em confronto, é suficiente calcular as respectivas probabilidades à posteriori.

Caso pretenda optar por uma das hipóteses em função de uma grandeza relativa,

pode-se calcular o *Odds* (Chance) à posteriori a favor de  $H_0$ :

$$O(H_0|t) = \frac{P(H_0|t)}{P(H_1|t)} = \frac{P(H_0|t)}{1 - P(H_0|t)}. \quad (15)$$

É possível obter o *Odds* a priori de maneira análoga:

$$O(H_0) = \frac{P(H_0)}{P(H_1)} = \frac{P(H_0)}{1 - P(H_0)}. \quad (16)$$

Com o objetivo de medir a influência dos dados  $t$  na alteração da credibilidade relativa  $H_0$  e  $H_1$  calcula-se o Fator de Bayes que é a razão entre o *Odds* a posteriori (15) e o *Odds* a priori (16), expressa por:

$$B_0(t) = \frac{O(H_0|t)}{O(H_0)} = \frac{\frac{P(H_0|t)}{1 - P(H_0|t)}}{\frac{P(H_0)}{1 - P(H_0)}}. \quad (17)$$

Aqui  $B_0(t)$  é denominado como o Fator de Bayes à favor de  $H_0$  (ou contra  $H_1$ ). Aplicando o logaritmo no Fator de Bayes, obtém-se uma relação aditiva entre as quantidades transformadas:

$$\ln B_0(t) = \ln O(H_0|t) - \ln O(H_0), \quad (18)$$

essas quantidades são chamadas de pesos de evidência. Assim,  $\ln B_0(t)$  é visto como o peso da evidência à posteriori descontado do correspondente peso da evidência à priori.

Conforme Paulino *et al.* (2003), o fator de Bayes ou seu logaritmo representam um peso relativo da evidência contida nos dados a favor de uma ou outra das hipóteses em confronto. Um fator de Bayes muito grande ou muito pequeno relativamente a 1 representa uma tendência bastante forte nos dados a favor de uma hipótese, contra a outra, entendida no sentido de que uma hipótese é muito mais ou muito menos provável do que era a priori. A correspondente desigualdade entre as duas chances conduz que a hipótese favorecida tenha uma probabilidade a posteriori superior à da sua alternativa, a não ser que se verifique uma relação fortemente contrária entre as probabilidades a priori.

#### 2.4.6 Full Bayesian Significance Test (FBST)

A abordagem frequentista, tendo como base a distribuição amostral, utiliza como medida de evidência em testes de hipóteses o nível descritivo ou valor-p, com a finalidade de mensurar a evidência trazida pelos dados em favor da hipótese nula  $H_0$ . Sabe-se que o valor-p é a probabilidade de observar dados mais extremos (se o processo aleatório for

repetido) do que aqueles observados na amostra, dado que a hipótese nula é verdadeira. Sendo assim, baseado nos dados amostrais que não favorecem a hipótese nula, o valor  $p$  indicará valores menores que o nível de significância, levando a decisão de rejeitar a hipótese nula e não rejeitar a hipótese alternativa.

É observado que a inferência realizada pela abordagem clássica calcula o nível descritivo levando em consideração a informação dos dados que poderiam ter sido observados, mas ainda não o foram, violando o Princípio da Verossimilhança, segundo o qual todo processo de decisão deve ser feito nos dados devidamente observados. Na abordagem bayesiana, a medida de evidência é calculada com base na função de verossimilhança dos dados observados e na distribuição a priori para a quantidade desconhecida, obedecendo ao Princípio da Verossimilhança e levando em consideração a hipótese alternativa (Fernandes, 2013).

Com o objetivo de criar uma medida bayesiana de evidência para hipóteses precisas Pereira e Stern (1999) apresentaram o teste de significância genuinamente bayesiano “ *Full Bayesian Significance Test* ” (FBST), que necessita apenas da distribuição a posteriori dos parâmetros. É um teste baseado no valor- $e$ , que representa a evidência da hipótese nula  $H_0$ .

O FBST é um teste cujo objetivo é testar hipóteses precisas baseadas no cálculo da probabilidade a posteriori da região HPD, que é tangente ao conjunto que define a hipótese nula. Este é um teste intuitivo, com fácil caracterização geométrica e pode ser implementado a partir de técnicas de otimização e integração numérica.

Considere uma variável aleatória  $T$  e um modelo estatístico paramétrico a ela associado  $(\Upsilon, A, F, \Theta, \pi)$ , em que  $\Upsilon$  é o espaço amostral, conjunto dos possíveis valores de  $t$ ,  $A$  é uma sigma-álgebra conveniente de subconjuntos de  $\Upsilon$ ,  $F$  é uma classe de distribuições de probabilidade em  $A$ , indexadas no espaço paramétrico  $\Theta$  e  $\pi$  é uma densidade a priori em  $\Theta$ . Considere também que, após a observação de  $t$ , obtém-se a densidade a posteriori de  $\theta$ ,  $\pi(\theta|t)$ , restringindo essa a uma função densidade de probabilidade.

Inicialmente, define-se  $\xi_\phi$  como um subconjunto do espaço paramétrico em que a densidade a posteriori  $\pi(\theta|t)$  é maior que  $\phi$  (Pereira e Stern, 1999):

$$\xi_\phi = \{\theta \in \Theta | \pi(\theta|t) > \phi\}.$$

A credibilidade de  $\xi_\phi$  é a sua probabilidade a posteriori:

$$K(\xi_\phi) = \int_{\xi_\phi} \pi(\theta|t) d\theta.$$

Definindo  $\theta^*$  como o argumento no qual a densidade a posteriori atinge o valor

máximo sob a hipótese  $H_0$  e  $\pi^*$  como o valor dessa densidade:

$$\theta^* \in \arg \max_{\theta \in \Theta} \pi(\theta|t)$$

e

$$\pi^* = \pi(\theta^*|t).$$

E definindo também  $\xi^*$  como o “conjunto tangente” da hipótese  $H_0$ :

$$\xi^* = \{\theta \in \Theta | \pi(\theta|t) > \pi^*\}.$$

Assim, a credibilidade do “conjunto tangente” à hipótese nula é definida por:

$$K^* = \pi(\xi^*|t) = \int_{\xi^*} \pi(\theta|t) d\theta.$$

Logo, a evidência a favor da hipótese  $H_0$ , decorrente da observação dos dados é definida por:

$$Ev(H_0) = \text{valor-e} = 1 - K^*.$$

Se a probabilidade de  $\xi^*$  é “grande”, isto significa que o conjunto de valores da hipótese  $H_0$  pertence a uma região de baixa probabilidade e a evidência trazida pelos dados é contra a hipótese  $H_0$ . Por outro lado, se a probabilidade de  $\xi^*$  é “pequena”, então o conjunto de valores da hipótese  $H_0$  está em uma região de alta probabilidade e a evidência trazida pelos dados é em favor da hipótese  $H_0$ . A definição anterior é bastante geral, uma vez que foi criada com o objetivo de testar hipóteses precisas, ou seja, uma hipótese nula para a qual a sua dimensão seja menor que a dimensão do espaço paramétrico, isto é,  $\dim(\Theta_0) < \dim(\Theta)$  (Pereira e Stern, 1999).

## 2.5 Classe de Distribuições Exponenciadas

Diversos autores já trabalharam e vem trabalhando com essa nova classe de distribuições buscando encontrar uma distribuição que acomode o máximo ou todas as formas de taxa de falha existentes na literatura, entre eles estão Mudholkar e Srivastava (1993) que propuseram a WE, Nekoukhou e Bidram (2015) que apresentaram a WDE.

A ideia da classe de distribuições exponenciadas é adicionar um parâmetro de forma  $\gamma > 0$  na função distribuição base. Seja  $T$  uma variável aleatória e  $G(t)$  a função distri-

buição acumulada base, a nova distribuição EXP- $G(t)$  terá função distribuição acumulada expressa da seguinte forma:

$$F(t) = [G(t)]^\gamma. \quad (19)$$

Ao considerar que  $T$  tem distribuição exponenciada como definida na equação (19), a distribuição EXP- $G(t)$  tem como caso especial a função distribuição base, quando  $\gamma = 1$ .

Caso a distribuição seja contínua, para  $t \geq 0$  a função densidade de probabilidade é dada por:

$$f(t) = \gamma[G(t)]^{\gamma-1}g(t), \quad (20)$$

em que  $g(t)$  é a derivada de primeira ordem de  $G(t)$ .

Caso a distribuição seja discreta, para  $t = 0, 1, 2, \dots$  a distribuição de probabilidade é dada por:

$$p(t) = \begin{cases} (G(0))^\gamma, & \text{se } t = 0 \\ (G(t))^\gamma - (G(t-1))^\gamma, & \text{se } t = 1, 2, \dots \end{cases} \quad (21)$$

O próximo capítulo detalha uma classe de distribuição exponenciada aplicando a teoria descrita nesta seção utilizando a distribuição Weibull contínua e discreta. Esse procedimento, a partir da inclusão de um novo parâmetro, aumenta a flexibilidade do ajuste de dados. Também realiza-se inferências da distribuição Weibull Discreta Exponenciada, tanto na abordagem clássica como na bayesiana.



### 3 Metodologia

Quando a variável resposta é o tempo até a ocorrência de um evento de interesse, em geral, esse tempo é modelado pelas distribuições Exponencial, Weibull, Gama, Log-normal e Log-logística. Contudo, há um crescimento de trabalhos para acomodar as diversas situações em estudo. O objetivo deste capítulo é propor a distribuição Weibull Discreta Exponenciada (WDE) para analisar dados de sobrevivência discretos nas abordagens clássica e bayesiana. Na Seção 3.1 é apresentada a distribuição Weibull considerando os casos contínuo e discreto, posteriormente na Seção 3.2 apresenta-se a distribuição WDE. Por fim, tem-se inferências da distribuição WDE dentro do contexto de análise de sobrevivência considerando a abordagem clássica (Seção 3.3) e a abordagem bayesiana (Seção 3.4).

#### 3.1 Distribuição Weibull

A distribuição Weibull recebeu esse nome após um engenheiro e matemático suéco chamado Waloddi Weibull propor, em 1939, uma distribuição apropriada para a resistência de ruptura dos materiais (Ryan, 2009). Essa distribuição tem sido o modelo probabilístico mais aceito e utilizado na modelagem de dados de sobrevivência devido à flexibilidade que o mesmo apresenta. Essa versatilidade se deve aos dois parâmetros que a distribuição apresenta: o parâmetro de escala e o parâmetro de forma, que proporcionam uma variedade de formas. A distribuição Weibull acomoda taxa de falha crescente, decrescente e constante. Entretanto, a mesma não permite que se trabalhe com taxas de falhas não-monótonas tais como taxa de falha em forma de banheira e unimodal, que usualmente vem sendo utilizadas em trabalhos de análise de sobrevivência.

Com o intuito de sanar essa deficiência, desde 1970, muitas extensões da distribuição Weibull têm sido propostas para melhorar a sua capacidade para se ajustar a diversos dados de tempo de vida. Ao longo dos últimos 40 anos o número de artigos relacionados com as várias extensões da distribuição Weibull e suas aplicações ultrapassam a marca de 4000 (Lai, 2014).

Um fator comum entre as distribuições generalizadas é que a distribuição Weibull é um caso especial das mesmas. A mudança realizada para obter tais distribuições é usualmente realizada na função distribuição acumulada ou na taxa de falha. No presente trabalho a modificação ocorrerá na função distribuição acumulada conforme visto na Seção 2.5.

Nesta seção será descrita a distribuição Weibull para os casos contínuo e discreto, apresentando a função de probabilidade, função de sobrevivência, função de risco e os

seus respectivos comportamentos.

### 3.1.1 Caso Contínuo

Uma variável aleatória contínua  $T$  com distribuição Weibull, tem a função densidade de probabilidade dada por:

$$f_w(t) = \frac{\beta}{\eta^\beta} t^{\beta-1} \exp \left\{ - \left( \frac{t}{\eta} \right)^\beta \right\}, \quad t \geq 0, \quad (22)$$

em que  $\beta > 0$  é o parâmetro de forma, e  $\eta > 0$  é o parâmetro de escala. O parâmetro  $\eta$  tem a mesma unidade de medida de  $t$  e  $\beta$  não tem unidade de medida.

A Figura 1 apresenta o gráfico da função densidade de probabilidade para diferentes valores de  $\eta$  e  $\beta$ . Nessa figura, pode-se observar que a distribuição Weibull pode ser utilizada tanto para acomodar tempos simétricos como tempos com assimetria à direita.

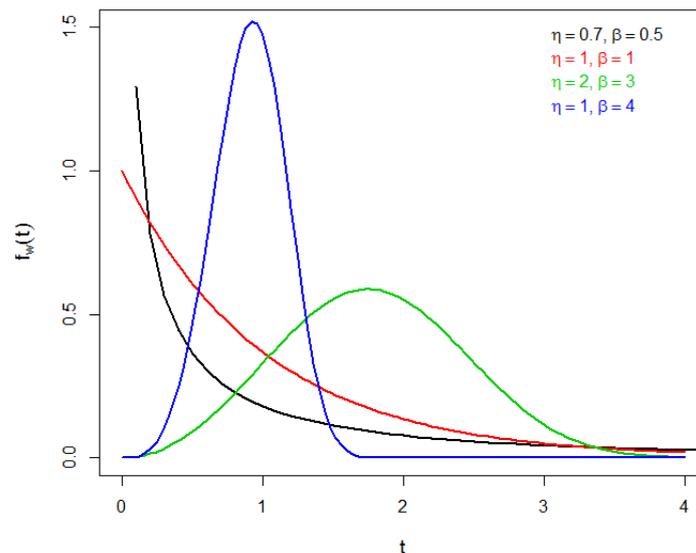


Figura 1: Função densidade de probabilidade da distribuição Weibull

A função de sobrevivência da distribuição Weibull é dada por:

$$S_w(t) = \exp \left\{ - \left( \frac{t}{\eta} \right)^\beta \right\}, \quad t \geq 0. \quad (23)$$

A Figura 2 trata-se do gráfico da função de sobrevivência para diferentes valores de  $\eta$  e  $\beta$ .

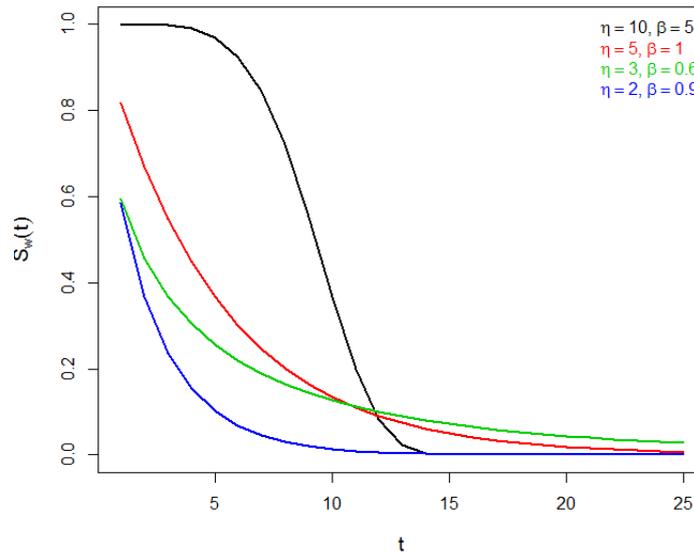


Figura 2: Função de sobrevivência da distribuição Weibull

Consequentemente tem-se a função distribuição acumulada, definida por:

$$F_w(t) = 1 - \exp \left\{ - \left( \frac{t}{\eta} \right)^\beta \right\}, \quad t \geq 0. \quad (24)$$

A taxa de falha do modelo Weibull é dada por:

$$h_w(t) = \frac{\beta}{\eta^\beta} t^{\beta-1}, \quad t \geq 0. \quad (25)$$

A Figura 3 refere-se a taxa de falha do modelo Weibull com o parâmetro de escala  $\eta = 1$  e diferentes valores do parâmetro de forma ( $\beta$ ).

Note que a taxa de falha da distribuição Weibull apresenta diferentes formas que estão ligadas diretamente ao parâmetro de forma  $\beta$ , isto é:

- Quando  $\beta > 1$ , a taxa de falha é estritamente crescente;
- Quando  $\beta = 1$ , a taxa de falha é constante (a distribuição Weibull se reduz a distribuição Exponencial);
- Quando  $\beta < 1$ , a taxa de falha é estritamente decrescente.

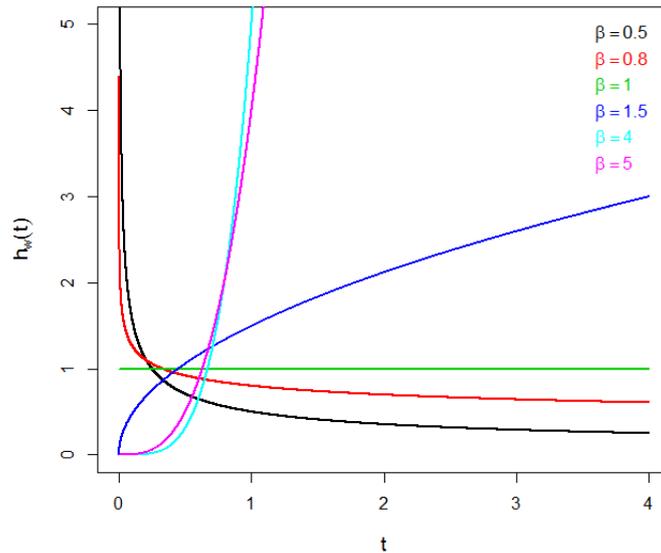


Figura 3: Taxa de falha da distribuição Weibull

### 3.1.2 Caso Discreto

Conforme visto na Seção (2.2), ao obter a variável discreta  $T = [X]$ , em que  $[X]$  representa a parte inteira de  $X$ , a distribuição de probabilidade da da distribuição Weibull Discreta (WD) proposta por Nakagawa e Osaki (1975) é expressa por:

$$\begin{aligned}
 p_{wd}(t) &= P[T = t] \\
 &= P[t \leq X < t + 1] \\
 &= F_w(t + 1) - F_w(t) \\
 &= \left[ 1 - \exp \left\{ - \left( \frac{t + 1}{\eta} \right)^\beta \right\} \right] - \left[ 1 - \exp \left\{ - \left( \frac{t}{\eta} \right)^\beta \right\} \right] \\
 &= q^{t^\beta} - q^{(t+1)^\beta} \quad t = 0, 1, 2, \dots,
 \end{aligned} \tag{26}$$

em que  $q = \exp \left\{ \frac{-1}{\eta^\beta} \right\}$ . Aqui  $0 < q < 1$  e  $\beta > 0$  é o parâmetro de forma da distribuição.

A Figura 4 apresenta o gráfico da função de probabilidade dessa distribuição para diferentes valores de  $q$  e  $\beta$ . Pode-se observar que a distribuição WD pode ser utilizada tanto para acomodar tempos discretos simétricos como tempos discretos com assimetria à direita.

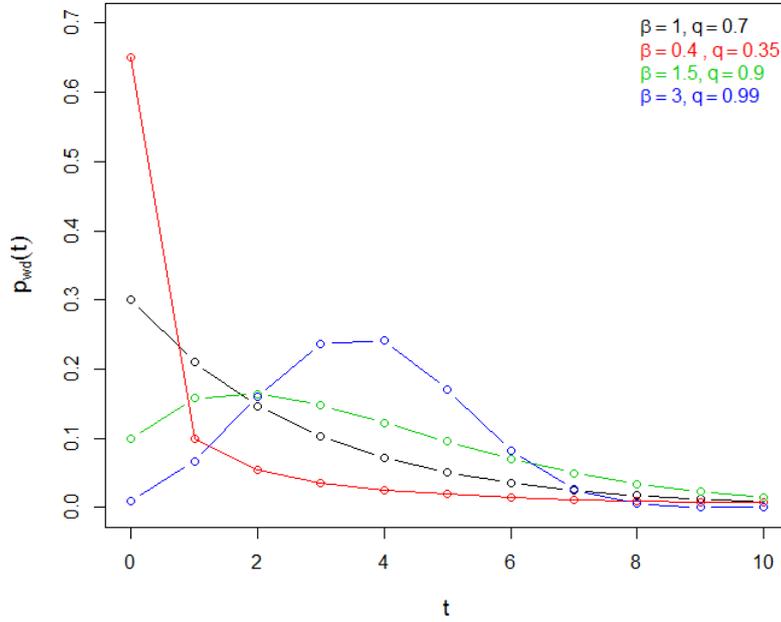


Figura 4: Distribuição de probabilidade da Weibull Discreta

A função de sobrevivência e a função distribuição acumulada da distribuição WD são expressas, respectivamente, por:

$$\begin{aligned}
 S_{wd}(t) &= \sum_{k=t+1}^{\infty} P(T = k) \\
 &= \sum_{k=t+1}^{\infty} q^{k^\beta} - q^{(t+1)^\beta} \\
 &= \left( q^{t^\beta} - q^{(t+1)^\beta} \right) + \left( q^{(t+1)^\beta} - q^{(t+2)^\beta} \right) + \left( q^{(t+2)^\beta} - q^{(t+3)^\beta} \right) + \dots \\
 &= q^{(t+1)^\beta}, \quad t = 0, 1, 2, \dots
 \end{aligned} \tag{27}$$

e

$$F_{wd}(t) = 1 - q^{(t+1)^\beta}, \quad t = 0, 1, 2, \dots \tag{28}$$

A Figura 5 trata-se do gráfico da função de sobrevivência da WD para diferentes valores de  $q$  e  $\beta$ . Note que, devido ao fato de os tempos serem discretos (assumindo apenas valores inteiros não negativos) o gráfico vem a ser uma função escada com degraus nos tempos observados de falha.

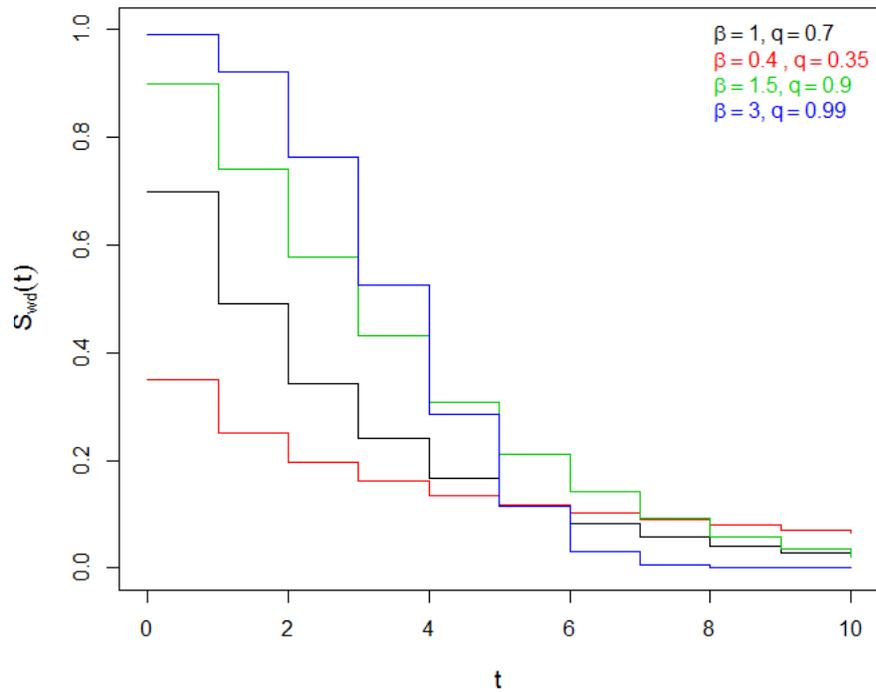


Figura 5: Função de sobrevivência da distribuição Weibull Discreta

A taxa de falha da distribuição WD é dada por:

$$h_{wd}(t) = \frac{q^{t^\beta} - q^{(t+1)^\beta}}{q^{t^\beta}}, \quad t = 0, 1, 2, \dots \quad (29)$$

A Figura 6 refere-se ao gráfico da taxa de falha da WD para diferentes valores de  $q$  e  $\beta$ .

Note que a taxa de falha da distribuição WD apresenta diferentes formas que estão ligadas diretamente ao seu parâmetro de forma  $\beta$ , isto é:

- Quando  $\beta > 1$ , a taxa de falha é estritamente crescente;
- Quando  $\beta = 1$ , a taxa de falha é constante (a distribuição WD se reduz a distribuição Geométrica);
- Quando  $\beta < 1$ , a taxa de falha é estritamente decrescente.

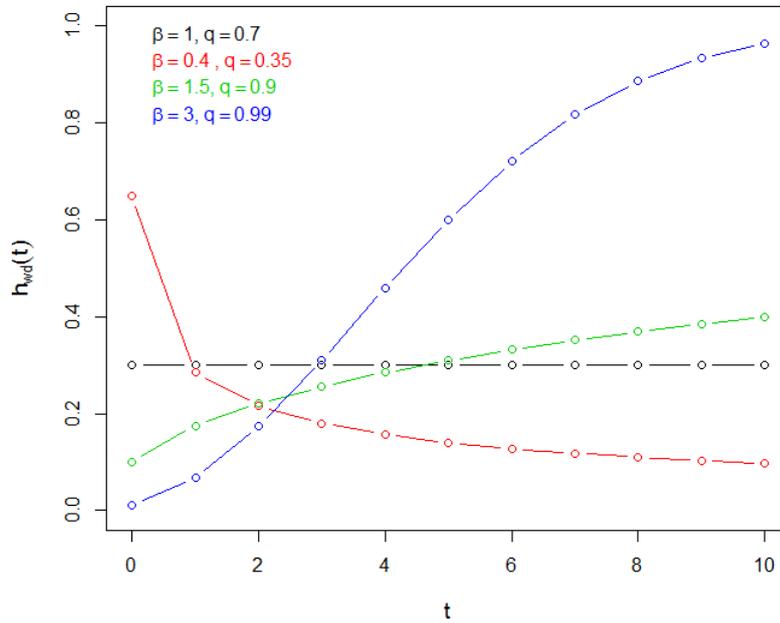


Figura 6: Taxa de falha da distribuição Weibull Discreta

### 3.2 Distribuição Weibull Discreta Exponenciada

Na Seção 2.5 foi definido como encontrar a distribuição exponenciada a partir de uma distribuição base. Dessa forma, ao considerar a definição dada pela equação (19) e a função distribuição acumulada dada pela equação (28), tem-se que a função distribuição acumulada da Weibull Discreta Exponenciada (WDE) pode ser expressa por:

$$\begin{aligned} F_{wde}(t) &= [F_{wd}(t)]^\gamma \\ &= \left[1 - q^{(t+1)^\beta}\right]^\gamma, \quad t = 0, 1, 2, \dots \end{aligned} \quad (30)$$

Aqui,  $\gamma > 0$  é o parâmetro de forma que também pode ser chamado de parâmetro de resiliência<sup>2</sup> (Nekoukhou e Bidram, 2015).

Conseqüentemente, conforme visto na Seção 2.5, a WDE tem distribuição de probabilidade dada por:

$$\begin{aligned} p_{wde}(t) &= P[T = t] \\ &= \left[1 - q^{(t+1)^\beta}\right]^\gamma - \left[1 - q^{t^\beta}\right]^\gamma, \quad t = 0, 1, 2, \dots \end{aligned} \quad (31)$$

<sup>2</sup>termo provindo do latim “silie” (saltar) acrescido do prefixo “re” (novamente) sendo entendido como voltar ao estado em que se encontrava anteriormente.

em que  $\gamma > 0$ ,  $\beta > 0$  e  $0 < q < 1$ .

A Figura 7 apresenta o gráfico da função de probabilidade dessa distribuição para diferentes valores de  $q$ ,  $\beta$  e  $\gamma$ . Pode-se observar que a distribuição WDE também pode ser utilizada tanto para acomodar tempos discretos simétricos como tempos discretos com assimetria à direita.

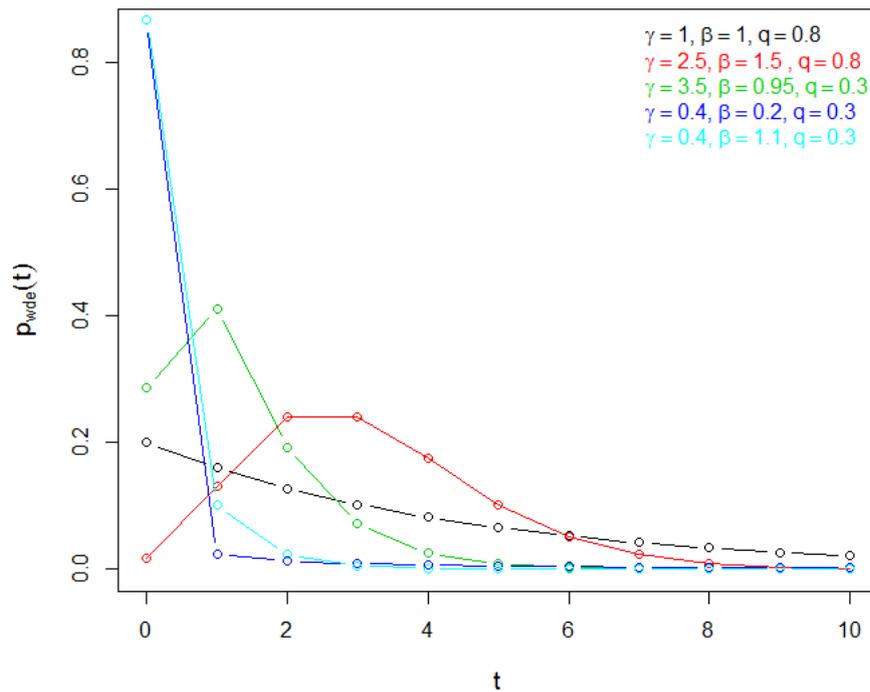


Figura 7: Distribuição de probabilidade da Weibull Discreta Exponenciada

A função de sobrevivência da distribuição WDE é expressa por:

$$S_{wde}(t) = 1 - \left[1 - q^{(t+1)^\beta}\right]^\gamma, \quad t = 0, 1, 2, \dots \quad (32)$$

A Figura 8 trata-se do gráfico da função de sobrevivência da WDE para diferentes valores de  $q$ ,  $\beta$  e  $\gamma$ .

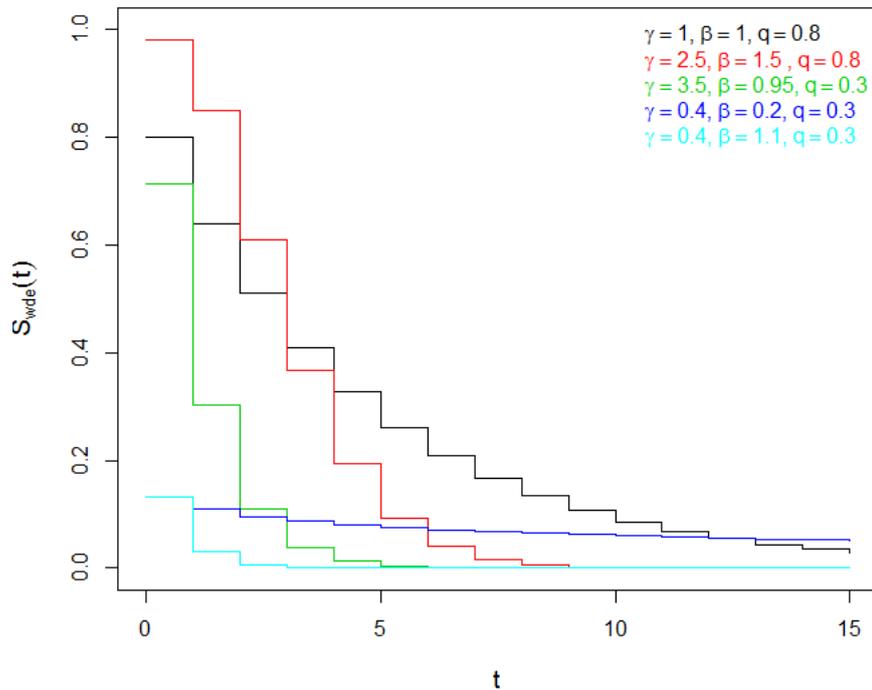


Figura 8: Função de sobrevivência da distribuição Weibull Discreta Exponenciada

Ademais, a função taxa de falha da distribuição WDE é definida por:

$$\begin{aligned}
 h_{wde}(t) &= \frac{p_{wde}(t)}{S_{wde}(t) + p_{wde}(t)} \\
 &= \frac{[1 - q^{(t+1)\beta}]^\gamma - [1 - q^{t\beta}]^\gamma}{1 - [1 - q^{t\beta}]^\gamma}, \quad t = 0, 1, 2, \dots
 \end{aligned} \tag{33}$$

A Figura 9 refere-se ao gráfico da taxa de falha da WDE para diferentes valores de  $q$ ,  $\beta$  e  $\gamma$ . Certamente uma das vantagens de se trabalhar com essa generalização da WD é o fato dela fornecer um ajuste paramétrico em que se permite trabalhar com vários tipos de taxas de falha existentes na literatura. A partir da Figura 9 nota-se que dependendo da combinação dos parâmetros a função taxa de falha assume comportamento crescente, decrescente, forma de banheira e unimodal.

Assim, conforme Nekoukhou e Bidram (2015), a distribuição WDE mostra-se como uma distribuição mais flexível para o ajuste dos dados do que a distribuição WD que acomoda apenas taxas de falhas monótonas e constantes.

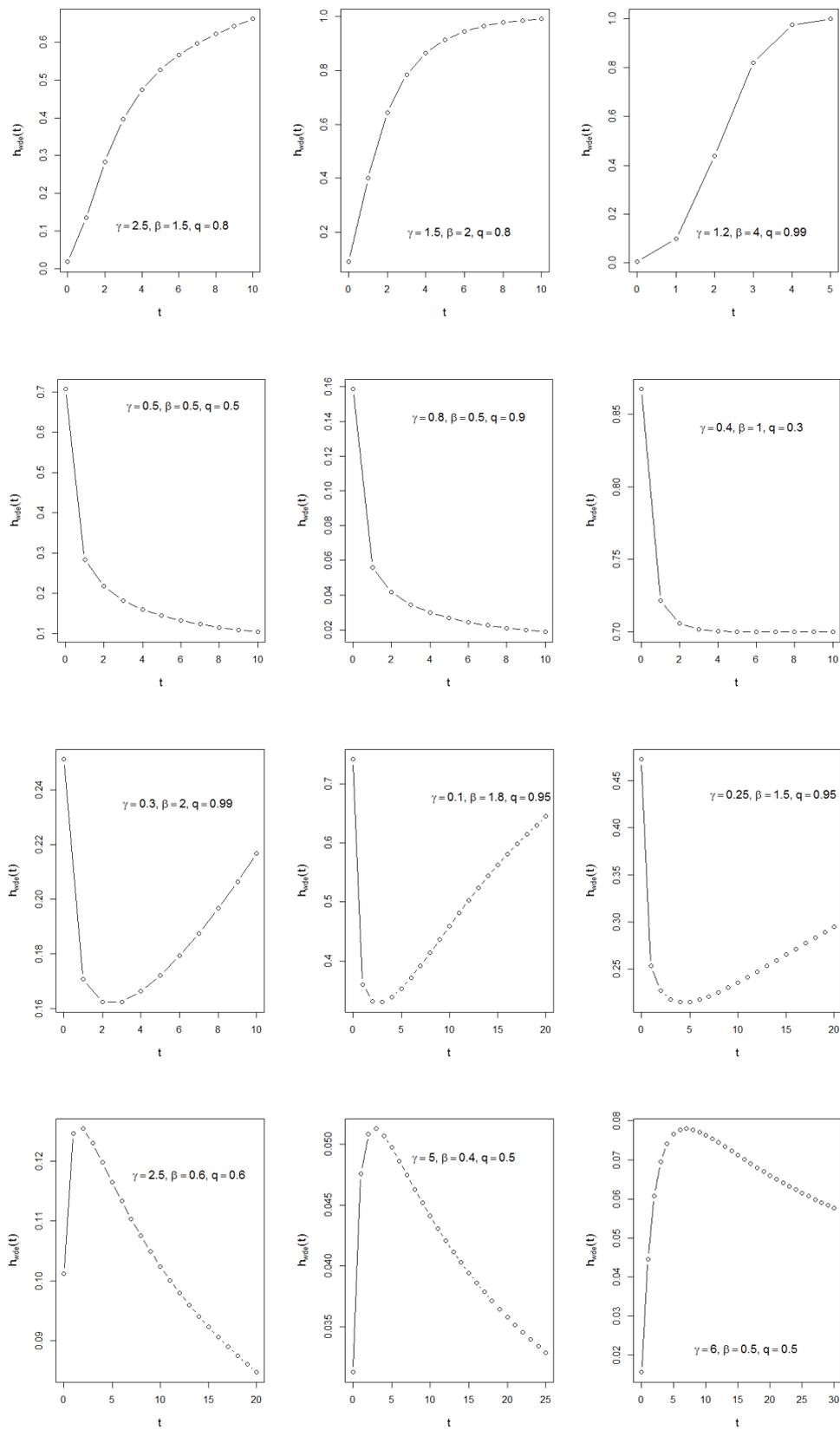


Figura 9: Taxa de falha da distribuição Weibull Discreta Exponenciada

A distribuição WDE assume alguns sub-casos especiais dependendo dos valores dos seus parâmetros. A WDE se reduz a:

- Distribuição WD, se  $\gamma = 1$ ;
- Distribuição Geométrica Exponenciada, se  $\beta = 1$ ;
- Distribuição Geométrica, se  $\gamma = 1$  e  $\beta = 1$ ;
- Distribuição Rayleigh Discreta Exponenciada, se  $\beta = 2$ ;
- Distribuição Rayleigh Discreta, se  $\gamma = 1$  e  $\beta = 2$  (Nekoukhou e Bidram, 2015).

### 3.2.1 Outra Formulação da WDE

Uma forma alternativa de encontrar a WDE, é primeiramente exponenciar a distribuição Weibull contínua, obtendo-se a distribuição Weibull Exponenciada (WE) proposta por Mudholkar e Srivastava (1993). Em seguida discretizar a WE contínua, obtendo-se a Weibull Exponenciada Discreta (WED) que é equivalente a WDE, como pode ser mostrado a seguir.

Ao exponenciar a função definida na equação (24), obtém-se a função distribuição acumulada da WE contínua:

$$\begin{aligned} F_{we}(t) &= [F_w(t)]^\gamma \\ &= \left[ 1 - \exp \left\{ - \left( \frac{t}{\eta} \right)^\beta \right\} \right]^\gamma, \quad t \geq 0. \end{aligned} \quad (34)$$

O próximo passo é discretizar a WE definida na equação (34). Ou seja,

$$\begin{aligned} p_{wed}(t) &= P[T = t] \\ &= P[t \leq F_{we}(t) < t + 1] \\ &= F_{we}(t + 1) - F_{we}(t) \\ &= \left[ 1 - \exp \left\{ - \left( \frac{(t + 1)}{\eta} \right)^\beta \right\} \right]^\gamma - \left[ 1 - \exp \left\{ - \left( \frac{t}{\eta} \right)^\beta \right\} \right]^\gamma \\ &= \left[ 1 - q^{(t+1)^\beta} \right]^\gamma - \left[ 1 - q^{t^\beta} \right]^\gamma, \quad t = 0, 1, 2, \dots, \end{aligned} \quad (35)$$

em que  $q = \exp \left\{ \frac{-1}{\eta^\beta} \right\}$ ,  $0 < q < 1$ ;  $\beta > 0$  e  $\gamma > 0$ .

Pelas equações (31) e (35), tem-se que:

$$p_{wde}(t) = p_{wed}(t).$$

Desta forma, conclui-se que os dois métodos de formulação se equivalem. Logo, a distribuição WDE é igual a distribuição WED.

### 3.3 Inferência Clássica

Seja uma amostra aleatória observada  $t_1, t_2, \dots, t_n$  de uma população com distribuição de probabilidade e função de sobrevivência da WDE definidas na equações (31) e (32). O vetor de parâmetros  $\boldsymbol{\theta}$  é definido por  $\boldsymbol{\theta} = (q, \beta, \gamma)$  e  $\delta_i$  denota a variável indicadora de censura, no qual é igual a 1 se o tempo  $t_i$  for de falha e 0 se for censura à direita. A função de verossimilhança da WDE é dada por:

$$\begin{aligned} L(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta}) &\propto \prod_{i=1}^n [p_{wde}(t_i)]^{\delta_i} [S_{wde}(t_i)]^{(1-\delta_i)} \\ &\propto \prod_{i=1}^n \left\{ \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[ 1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)}, \end{aligned} \quad (36)$$

em que  $\gamma > 0$ ,  $\beta > 0$  e  $0 < q < 1$ .

Aplicando o logaritmo na função de verossimilhança (36), tem-se:

$$\begin{aligned} \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \delta_i \log \left\{ \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[ 1 - q^{t_i^\beta} \right]^\gamma \right\} \right\} + \\ &+ \sum_{i=1}^n \left\{ (1 - \delta_i) \log \left\{ 1 - \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma \right\} \right\} + c. \end{aligned} \quad (37)$$

em que  $c$  é uma constante que não depende de  $\boldsymbol{\theta}$ .

Derivando a expressão (37) em relação a  $q$ ,  $\beta$  e  $\gamma$  e igualando a zero obtém-se as equações de verossimilhança (disponíveis no Apêndice A.1) representadas por:

$$\frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})}{\partial q} = 0, \quad \frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})}{\partial \beta} = 0, \quad \frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta})}{\partial \gamma} = 0. \quad (38)$$

Os valores de  $\hat{q}$ ,  $\hat{\beta}$  e  $\hat{\gamma}$  que satisfazem as equações (38), são os estimadores de máxima verossimilhança da distribuição WDE. Contudo, esses cálculos são extensos e por este motivo é recomendado utilizar métodos computacionais de otimização como o algoritmo de Newton-Raphson, que utiliza a matriz de derivadas segundas do logaritmo da função de verossimilhança (37).

Como os parâmetros da distribuição WDE tem limitação no espaço paramétrico, ou seja,  $\gamma > 0$ ,  $\beta > 0$  e  $0 < q < 1$  é necessário fazer uma transformação para deixar os parâmetros irrestritos, e então construir intervalos de confiança em que os limites são

definidos considerando o espaço paramétrico de cada parâmetro. Assim, para o parâmetro  $q$  foi considerada a transformação log-log e para os parâmetros  $\beta$  e  $\gamma$  foi considerada a transformação logarítmica (Apêndice A.2).

Desta forma, os intervalos de confiança para os parâmetros  $q$ ;  $\beta$  e  $\gamma$ , são dados respectivamente por:

$$\left[ (\hat{q})e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}}; (\hat{q})e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} \right], \quad (39)$$

em que  $\hat{u} = \log[-\log \hat{q}]$ ;

$$\left[ \hat{\beta}e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}}; \hat{\beta}e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} \right], \quad (40)$$

em que  $\hat{v} = \log \hat{\beta}$ , e

$$\left[ \hat{\gamma}e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}}; \hat{\gamma}e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} \right], \quad (41)$$

em que  $\hat{w} = \log \hat{\gamma}$  e  $Z_{1-\alpha/2}$  é o quantil  $(1 - \alpha/2)$  de uma distribuição normal padrão.

Como definido na Seção 2.3.2 os valores de  $\widehat{Var}(\hat{u})$ ,  $\widehat{Var}(\hat{v})$  e  $\widehat{Var}(\hat{w})$ , que também são obtidos numericamente calculados via *método delta*.

### 3.4 Inferência Bayesiana

Ao considerar a função de verossimilhança (36) e  $\boldsymbol{\theta} = (q, \beta, \gamma)$ , como  $\beta > 0$ ,  $\gamma > 0$  e  $0 < q < 1$ , foram consideradas as seguintes distribuições a priori para os parâmetros:  $\beta \sim Gama(a_1, b_1)$ ,  $\gamma \sim Gama(a_2, b_2)$  e  $q \sim Beta(a_3, b_3)$ , em que  $a_1$ ,  $a_2$ ,  $a_3$ ,  $b_1$ ,  $b_2$  e  $b_3$  são hiperparâmetros positivos conhecidos. Supondo a independência dos parâmetros a priori, a distribuição a priori pode ser descrita da seguinte maneira:

$$\pi(\boldsymbol{\theta}) \propto (\beta^{a_1-1}e^{-b_1\beta}) \times (\gamma^{a_2-1}e^{-b_2\gamma}) \times (q^{a_3-1}(1-q)^{b_3-1}).$$

Assim, a distribuição a posteriori dos parâmetros é proporcional a:

$$\begin{aligned} \pi(\boldsymbol{\theta}|\mathbf{t}, \boldsymbol{\delta}) &\propto \pi(\boldsymbol{\theta})L(\mathbf{t}, \boldsymbol{\delta}|\boldsymbol{\theta}) \\ \pi(\boldsymbol{\theta}|\mathbf{t}, \boldsymbol{\delta}) &\propto \beta^{a_1-1}e^{-b_1\beta}\gamma^{a_2-1}e^{-b_2\gamma}q^{a_3-1}(1-q)^{b_3-1} \\ &\quad \prod_{i=1}^n \left\{ \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[ 1 - q^{t_i^\beta} \right]^\gamma \right\}^{\delta_i} \left\{ 1 - \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma \right\}^{(1-\delta_i)}. \end{aligned} \quad (42)$$

E as distribuições condicionais a posteriori são dadas por:

$$\pi(\beta|\gamma, q, \mathbf{t}, \boldsymbol{\delta}) \propto \beta^{a_1-1} e^{-b_1\beta} \prod_{i=1}^n \left\{ [1-q^{(t_i+1)\beta}]^\gamma - [1-q^{t_i^\beta}]^\gamma \right\}^{\delta_i} \left\{ 1 - [1-q^{(t_i+1)\beta}]^\gamma \right\}^{(1-\delta_i)},$$

$$\pi(\gamma|\beta, q, \mathbf{t}, \boldsymbol{\delta}) \propto \gamma^{a_2-1} e^{-b_2\gamma} \prod_{i=1}^n \left\{ [1-q^{(t_i+1)\beta}]^\gamma - [1-q^{t_i^\beta}]^\gamma \right\}^{\delta_i} \left\{ 1 - [1-q^{(t_i+1)\beta}]^\gamma \right\}^{(1-\delta_i)}$$

e

$$\pi(q|\gamma, \beta, \mathbf{t}, \boldsymbol{\delta}) \propto q^{a_3-1} (1-q)^{b_3-1} \prod_{i=1}^n \left\{ [1-q^{(t_i+1)\beta}]^\gamma - [1-q^{t_i^\beta}]^\gamma \right\}^{\delta_i} \left\{ 1 - [1-q^{(t_i+1)\beta}]^\gamma \right\}^{(1-\delta_i)}.$$

A distribuição a posteriori (42) não pode ser obtida analiticamente, mas amostras da mesma podem ser obtidas numericamente através dos métodos MCMC (*Markov Chain Monte Carlo*) com passos do algoritmo Metropolis-Hastings (Metropolis *et al.* (1953); Hastings (1970)).

## 4 Simulações Computacionais

Neste capítulo serão apresentadas as simulações computacionais e os seus resultados, que foram obtidos através do *software* livre R. As simulações tem por objetivo gerar tempos de sobrevivência da distribuição WDE pelo método da inversão e considerando o mecanismo de censura a direita aleatório. A censura foi incorporada nas amostras independentemente do tempo de sobrevivência através da variável indicadora de censura gerada por uma distribuição Bernoulli, com os percentuais de censura que estão especificados nas seções 4.1 e 4.2.

Foram considerados diferentes cenários para simulação, levando em conta o comportamento da taxa de falha da distribuição em estudo, conforme a Figura 9. Dessa forma, na Tabela 1 estão descritos os cenários escolhidos para as simulações.

Tabela 1: Cenários utilizados na simulação

| Cenário | $q$  | $\beta$ | $\gamma$ | Taxa de falha     |
|---------|------|---------|----------|-------------------|
| 1       | 0,8  | 1,5     | 2,5      | Crescente         |
| 2       | 0,9  | 0,5     | 0,8      | Decrescente       |
| 3       | 0,99 | 2       | 0,3      | Forma de Banheira |
| 4       | 0,5  | 0,5     | 6        | Unimodal          |

Na Seção 4.1 são calculados o vício e o Erro quadrático médio (EQM) dos estimadores do modelo a partir dos cenários descritos na Tabela 1 e na Seção 4.2 é feito um estudo de sensibilidade da escolha da priori considerando o Cenário 1.

### 4.1 Vício e Erro Quadrático Médio (EQM)

Segundo Bolfarine e Sandoval (2010), o Erro Quadrático Médio (EQM) de um estimador  $\hat{\theta}$  do parâmetro  $\theta$  é dado por:

$$EQM[\hat{\theta}] = E[(\hat{\theta} - \theta)^2].$$

Tem-se ainda que:

$$EQM[\hat{\theta}] = Var[\hat{\theta}] + B^2(\hat{\theta}),$$

em que

$$B(\hat{\theta}) = E[\hat{\theta}] - \theta,$$

é denominado o vício do estimador  $\hat{\theta}$ . Um estimador  $\hat{\theta}$  é não viciado para  $\theta$  se:

$$E[\hat{\theta}] = \theta,$$

ou seja  $B(\hat{\theta}) = 0$  para todo  $\theta \in \Theta$ . E ainda se  $\lim_{n \rightarrow \infty} B(\hat{\theta}) = 0$ , pode-se dizer que o estimador  $\hat{\theta}$  é assintoticamente não viciado para  $\theta$ . No caso em que  $\hat{\theta}$  é um estimador não viciado para  $\theta$  tem-se que:

$$EQM[\hat{\theta}] = Var[\hat{\theta}],$$

ou seja o EQM de  $\hat{\theta}$  se reduz à sua variância.

Naturalmente duas propriedades desejáveis do estimador é que este seja assintoticamente não viciado e que tenha o mínimo EQM, ou seja:

$$\begin{aligned} i) \lim_{n \rightarrow \infty} B(\hat{\theta}) &= 0 \\ ii) \lim_{n \rightarrow \infty} EQM[\hat{\theta}] &= 0 \end{aligned}$$

Sendo assim, uma vez que, estas propriedades para o estimador ou para um vetor de estimadores são verificadas pode-se comprovar a acurácia do modelo proposto.

O vício e o EQM dos estimadores do modelo Weibull Discreto Exponenciado foram calculados a partir de 10.000 réplicas de Monte Carlo considerando amostras de tamanho  $n = 30; 80; 100; 200$  e 500 e percentuais de censura iguais a 0%; 10%; 20% e 30%.

A Tabela 2 refere-se as estimativas do vício e EQM para cada parâmetro conforme o Cenário 1 para diferentes tamanhos de amostra e percentuais de censura.

Nota-se pelos resultados obtidos na Tabela 2 a acurácia do modelo proposto, uma vez que as estimativas do vício e EQM são relativamente pequenas mesmo para amostras consideravelmente pequenas ( $n = 30$ ), e que as estimativas do vício tendem a valores mais próximos de zero quando o tamanho da amostra aumenta.

Tabela 2: Estimativas do vício e EQM do Cenário 1

|     |           | Cenário 1                               |        |         |        |         |        |         |        |
|-----|-----------|---|--------|---------|--------|---------|--------|---------|--------|
|     |           | $q = 0, 8; \beta = 1, 5; \gamma = 2, 5$ |        |         |        |         |        |         |        |
| $n$ | Parâmetro | Censuras                                |        |         |        |         |        |         |        |
|     |           | 0%                                      |        | 10%     |        | 20%     |        | 30%     |        |
|     |           | Vício                                   | EQM    | Vício   | EQM    | Vício   | EQM    | Vício   | EQM    |
| 30  | $q$       | -0,0041                                 | 0,0017 | -0,0004 | 0,0022 | 0,0035  | 0,0028 | 0,0084  | 0,0035 |
|     | $\beta$   | 0,0206                                  | 0,0284 | -0,0218 | 0,0338 | -0,0643 | 0,0441 | -0,1065 | 0,0585 |
|     | $\gamma$  | 0,1303                                  | 0,0877 | 0,1419  | 0,0782 | 0,1529  | 0,0695 | 0,1763  | 0,0710 |
| 80  | $q$       | -0,0048                                 | 0,0007 | -0,0043 | 0,0012 | -0,0021 | 0,0016 | 0,0022  | 0,0019 |
|     | $\beta$   | 0,0003                                  | 0,0104 | -0,0501 | 0,0157 | -0,0971 | 0,0272 | -0,1410 | 0,0418 |
|     | $\gamma$  | 0,0959                                  | 0,0545 | 0,1354  | 0,0739 | 0,1647  | 0,0599 | 0,1848  | 0,0538 |
| 100 | $q$       | -0,0055                                 | 0,0007 | -0,0049 | 0,0010 | -0,0031 | 0,0014 | 0,0008  | 0,0016 |
|     | $\beta$   | -0,0037                                 | 0,0086 | -0,0531 | 0,0141 | -0,1009 | 0,0252 | -0,1476 | 0,0405 |
|     | $\gamma$  | 0,0925                                  | 0,0573 | 0,1317  | 0,0648 | 0,1659  | 0,0538 | 0,1884  | 0,0545 |
| 200 | $q$       | -0,0049                                 | 0,0004 | -0,0064 | 0,0008 | -0,0046 | 0,0011 | -0,0014 | 0,0011 |
|     | $\beta$   | -0,0060                                 | 0,0051 | -0,0616 | 0,0106 | -0,1076 | 0,0221 | -0,1582 | 0,0362 |
|     | $\gamma$  | 0,0692                                  | 0,0461 | 0,1309  | 0,0780 | 0,1770  | 0,0655 | 0,1946  | 0,0543 |
| 500 | $q$       | -0,0046                                 | 0,0004 | -0,0071 | 0,0006 | -0,0058 | 0,0008 | -0,0009 | 0,0007 |
|     | $\beta$   | -0,0076                                 | 0,0032 | -0,0666 | 0,0083 | -0,1143 | 0,0197 | -0,1611 | 0,0319 |
|     | $\gamma$  | 0,0541                                  | 0,0428 | 0,1246  | 0,0729 | 0,1751  | 0,0627 | 0,1846  | 0,0617 |

Na Tabela 3 são apresentadas as estimativas do vício e EQM para cada parâmetro referentes ao Cenário 2 para diferentes tamanhos de amostra e percentuais de censura.

Tabela 3: Estimativas do vício e EQM do Cenário 2

|     |           | Cenário 2                               |        |         |        |         |        |         |        |
|-----|-----------|---|--------|---------|--------|---------|--------|---------|--------|
|     |           | $q = 0, 9; \beta = 0, 5; \gamma = 0, 8$ |        |         |        |         |        |         |        |
| $n$ | Parâmetro | Censuras                                |        |         |        |         |        |         |        |
|     |           | 0%                                      |        | 10%     |        | 20%     |        | 30%     |        |
|     |           | Vício                                   | EQM    | Vício   | EQM    | Vício   | EQM    | Vício   | EQM    |
| 30  | $q$       | -0,0042                                 | 0,0011 | 0,0052  | 0,0009 | 0,0149  | 0,0011 | 0,0236  | 0,0013 |
|     | $\beta$   | 0,0092                                  | 0,0035 | 0,0079  | 0,0039 | 0,0072  | 0,0047 | 0,0039  | 0,0057 |
|     | $\gamma$  | 0,0289                                  | 0,0157 | 0,0394  | 0,0107 | 0,0461  | 0,0094 | 0,0560  | 0,0125 |
| 80  | $q$       | -0,0043                                 | 0,0005 | 0,0037  | 0,0005 | 0,0129  | 0,0007 | 0,0245  | 0,0011 |
|     | $\beta$   | 0,0001                                  | 0,0013 | -0,0032 | 0,0015 | -0,0052 | 0,0019 | -0,0026 | 0,0026 |
|     | $\gamma$  | 0,0246                                  | 0,0096 | 0,0389  | 0,0095 | 0,0479  | 0,0079 | 0,0500  | 0,0069 |
| 100 | $q$       | -0,0045                                 | 0,0005 | 0,0030  | 0,0005 | 0,0127  | 0,0006 | 0,0240  | 0,0011 |
|     | $\beta$   | -0,0014                                 | 0,0011 | -0,0046 | 0,0014 | -0,0065 | 0,0017 | -0,0042 | 0,0024 |
|     | $\gamma$  | 0,0241                                  | 0,0095 | 0,0388  | 0,0102 | 0,0472  | 0,0083 | 0,0504  | 0,0077 |
| 200 | $q$       | -0,0041                                 | 0,0004 | 0,0034  | 0,0005 | 0,0129  | 0,0005 | 0,0245  | 0,0009 |
|     | $\beta$   | -0,0025                                 | 0,0008 | -0,0059 | 0,0010 | -0,0082 | 0,0011 | -0,0057 | 0,0015 |
|     | $\gamma$  | 0,0184                                  | 0,0070 | 0,0363  | 0,0095 | 0,0483  | 0,0085 | 0,0490  | 0,0061 |
| 500 | $q$       | -0,0035                                 | 0,0002 | 0,0041  | 0,0003 | 0,0134  | 0,0004 | 0,0260  | 0,0008 |
|     | $\beta$   | -0,0031                                 | 0,0005 | -0,0061 | 0,0007 | -0,0084 | 0,0007 | -0,0049 | 0,0008 |
|     | $\gamma$  | 0,0141                                  | 0,0042 | 0,0331  | 0,0059 | 0,0472  | 0,0053 | 0,0448  | 0,0042 |

Observa-se na Tabela 3, que as estimativas do vício e EQM são maiores quando a amostra é pequena e diminuem gradativamente à medida que a amostra aumenta, o que atesta a acurácia do modelo WDE.

A Tabela 4 refere-se as estimativas do vício e EQM para cada parâmetro de acordo com o Cenário 3 para diferentes tamanhos de amostra e percentuais de censura.

Tabela 4: Estimativas do vício e EQM do Cenário 3

|     |          | Cenário 3                           |                    |         |                    |         |                    |        |                    |
|-----|----------|-------------------------------------|--------------------|---------|--------------------|---------|--------------------|--------|--------------------|
|     |          | $q = 0,99; \beta = 2; \gamma = 0,3$ |                    |         |                    |         |                    |        |                    |
| $n$ | Par.     | Censuras                            |                    |         |                    |         |                    |        |                    |
|     |          | 0%                                  |                    | 10%     |                    | 20%     |                    | 30%    |                    |
|     |          | Vício                               | EQM                | Vício   | EQM                | Vício   | EQM                | Vício  | EQM                |
| 30  | $q$      | 0,0002                              | $1 \times 10^{-5}$ | 0,0010  | $1 \times 10^{-5}$ | 0,0019  | $1 \times 10^{-5}$ | 0,0029 | $2 \times 10^{-4}$ |
|     | $\beta$  | 0,0571                              | 0,0079             | 0,0357  | 0,0047             | 0,0195  | 0,0029             | 0,0079 | 0,0021             |
|     | $\gamma$ | 0,0157                              | 0,0059             | 0,0377  | 0,0087             | 0,0615  | 0,0127             | 0,0862 | 0,0180             |
| 80  | $q$      | 0,0006                              | $2 \times 10^{-6}$ | 0,0014  | $5 \times 10^{-6}$ | 0,0026  | $1 \times 10^{-5}$ | 0,0038 | $2 \times 10^{-4}$ |
|     | $\beta$  | 0,0378                              | 0,0041             | 0,0184  | 0,0015             | 0,0126  | 0,0011             | 0,0136 | 0,0016             |
|     | $\gamma$ | 0,0029                              | 0,0019             | 0,0205  | 0,0029             | 0,0397  | 0,0047             | 0,0622 | 0,0081             |
| 100 | $q$      | 0,0006                              | $2 \times 10^{-6}$ | 0,0014  | $4 \times 10^{-6}$ | 0,0027  | $1 \times 10^{-5}$ | 0,0039 | $2 \times 10^{-4}$ |
|     | $\beta$  | 0,0337                              | 0,0034             | 0,01482 | 0,0011             | 0,0111  | 0,0008             | 0,0162 | 0,0019             |
|     | $\gamma$ | 0,0015                              | 0,0016             | 0,0187  | 0,0023             | 0,0366  | 0,0038             | 0,0572 | 0,0068             |
| 200 | $q$      | 0,0005                              | $1 \times 10^{-6}$ | 0,0014  | $3 \times 10^{-6}$ | 0,0028  | $9 \times 10^{-6}$ | 0,0043 | $2 \times 10^{-5}$ |
|     | $\beta$  | 0,0244                              | 0,0017             | 0,0071  | 0,0004             | 0,0089  | 0,0006             | 0,0240 | 0,0028             |
|     | $\gamma$ | -0,0007                             | 0,0008             | 0,0169  | 0,0013             | 0,0317  | 0,0023             | 0,0465 | 0,0041             |
| 500 | $q$      | 0,0004                              | $7 \times 10^{-7}$ | 0,0015  | $2 \times 10^{-6}$ | 0,0029  | $9 \times 10^{-6}$ | 0,0044 | $2 \times 10^{-5}$ |
|     | $\beta$  | 0,0161                              | 0,0008             | 0,0039  | 0,0001             | 0,0088  | 0,0005             | 0,0260 | 0,0024             |
|     | $\gamma$ | -0,0018                             | 0,0004             | 0,0152  | 0,0006             | 0,02787 | 0,0013             | 0,0401 | 0,0024             |

Para o Cenário 3 que considera a taxa de falha em forma de banheira, as estimativas do vício e EQM apresentados na Tabela 4, foram muito próximas de zero independente do tamanho da amostra, fato que comprova a acurácia do modelo WDE para este cenário.

Na Tabela 5 tem-se as estimativas do vício e EQM para cada parâmetro em concordância com o Cenário 4 para diferentes tamanhos de amostra e percentuais de censura.

As estimativas do vício e EQM apresentadas Tabela 5 são valores relativamente pequenos e que tendem a diminuir quando o tamanho da amostra ( $n$ ) aumenta, o que confirma a acurácia do modelo em estudo.

Tabela 5: Estimativas do vício e EQM do Cenário 4

|     |           | Cenário 4                            |        |         |        |         |        |         |        |
|-----|-----------|--------------------------------------|--------|---------|--------|---------|--------|---------|--------|
|     |           | $q = 0, 5; \beta = 0, 5; \gamma = 6$ |        |         |        |         |        |         |        |
| $n$ | Parâmetro | Censuras                             |        |         |        |         |        |         |        |
|     |           | 0%                                   |        | 10%     |        | 20%     |        | 30%     |        |
|     |           | Vício                                | EQM    | Vício   | EQM    | Vício   | EQM    | Vício   | EQM    |
| 30  | $q$       | 0,0019                               | 0,0042 | 0,0068  | 0,0046 | 0,0102  | 0,0051 | 0,0144  | 0,0057 |
|     | $\beta$   | 0,0134                               | 0,0053 | -0,0002 | 0,0055 | -0,0159 | 0,0061 | -0,0326 | 0,0073 |
|     | $\gamma$  | 0,2590                               | 0,1357 | 0,3222  | 0,1746 | 0,3836  | 0,2173 | 0,4396  | 0,2587 |
| 80  | $q$       | -0,0049                              | 0,0015 | -0,0033 | 0,0016 | -0,0017 | 0,0019 | 0,0017  | 0,0022 |
|     | $\beta$   | 0,0016                               | 0,0018 | -0,0139 | 0,0021 | -0,0302 | 0,0029 | -0,0467 | 0,0045 |
|     | $\gamma$  | 0,2339                               | 0,1020 | 0,3188  | 0,1562 | 0,4172  | 0,2273 | 0,4996  | 0,2887 |
| 100 | $q$       | -0,0063                              | 0,0012 | -0,0051 | 0,0013 | -0,0039 | 0,0015 | -0,0009 | 0,0017 |
|     | $\beta$   | -0,0002                              | 0,0014 | -0,0156 | 0,0018 | -0,0324 | 0,0027 | -0,0492 | 0,0042 |
|     | $\gamma$  | 0,2314                               | 0,1014 | 0,3179  | 0,1557 | 0,4257  | 0,2300 | 0,5083  | 0,2917 |
| 200 | $q$       | -0,0074                              | 0,0007 | -0,0082 | 0,0007 | -0,0082 | 0,0009 | -0,0040 | 0,0009 |
|     | $\beta$   | -0,0023                              | 0,0007 | -0,0193 | 0,0011 | -0,0365 | 0,0022 | -0,0525 | 0,0036 |
|     | $\gamma$  | 0,2021                               | 0,0937 | 0,3304  | 0,1667 | 0,4711  | 0,2594 | 0,5357  | 0,3054 |
| 500 | $q$       | -0,0075                              | 0,0004 | -0,0105 | 0,0005 | -0,0119 | 0,0006 | -0,0055 | 0,0005 |
|     | $\beta$   | -0,0033                              | 0,0003 | -0,0216 | 0,0008 | -0,0394 | 0,0019 | -0,0542 | 0,0033 |
|     | $\gamma$  | 0,1608                               | 0,0889 | 0,3589  | 0,1926 | 0,5293  | 0,3113 | 0,5535  | 0,3198 |

De maneira geral, as estimativas do vício e EQM diminuem a medida que aumenta-se o tamanho da amostra na situação sem censura, uma vez que, a função de verossimilhança conta apenas com a distribuição de probabilidade. Quando se tem censura, sempre haverá vício, pois os “verdadeiros” valores dos parâmetros não consideram a censura, sendo assim, é esperado que o vício apareça. Naturalmente observa-se, que as estimativas do vício e EQM tornam-se maiores quando o percentual de censura é maior pois, essas estimativas são naturalmente viesadas pelo fato da função de verossimilhança na presença de censuras contar com a contribuição da função de sobrevivência.

A acurácia do modelo WDE foi comprovada em todos os Cenários. Dessa forma, o modelo pode ser utilizado quando se tem taxa de falha crescente (Cenário 1) e decrescente (Cenário 2). Além disso, tem-se a vantagem deste modelo ser utilizado quando a função de risco tem forma de banheira (Cenário 3) e unimodal (Cenário 4).

## 4.2 Estudo de Sensibilidade da Escolha da Priori

Nesta seção será feito um estudo de sensibilidade da escolha dos hiperparâmetros da distribuição a priori de  $\gamma$ . A distribuição a priori de  $\gamma$  foi escolhida pois quando este parâmetro for igual a um o modelo WDE se reduz ao modelo WD. Como visto na Seção 3.4, considerou-se a priori que  $\gamma \sim Gama(a_2, b_2)$  e pretende-se variar os valores dos hiperparâmetros  $a_2, b_2$ . O objetivo é observar o impacto desses hiperparâmetros nos

resultados da inferência.

Para este estudo foi utilizado o Cenário 1 ( $q = 0,8; \beta = 1,5; \gamma = 2,5$ ). As estimativas para o cenário foram baseadas em uma cadeia de tamanho 10.000, considerando um *burn-in* de 1.000. A Tabela 6 apresenta um estudo de sensibilidade da escolha dos hiperparâmetros da distribuição a priori de  $\gamma$  com as estimativas pontuais dos parâmetros (média a posteriori) da distribuição WDE, seus respectivos intervalos HPD de 95% de credibilidade e também o valor-e da hipótese  $H_{01}$ . Este estudo foi realizado considerando uma amostra de tamanho  $n=200$  e 0% de censura. Foi decidido aqui adotar uma amostra sem censuras para proporcionar um maior controle do verdadeiro valor dos parâmetros nas amostras.

Tabela 6: Influência da escolha dos hiperparâmetros  $a_2$  e  $b_2$  da priori de  $\gamma$  na inferência dos parâmetros da distribuição WDE para o Cenário 1.

| $a_2$ | $b_2$ | Parâmetros | Estimativas<br>(média a posteriori) | Intervalo HPD<br>95% | valor-e:<br>$H_{01} : \gamma = 1$ |
|-------|-------|------------|-------------------------------------|----------------------|-----------------------------------|
| 0,001 | 0,001 | $q$        | 0,7271                              | [0,6291; 0,8176]     | 0,0274                            |
|       |       | $\beta$    | 1,3971                              | [1,1965; 1,5808]     |                                   |
|       |       | $\gamma$   | 3,1249                              | [2,1831; 4,1952]     |                                   |
| 0,01  | 0,01  | $q$        | 0,7303                              | [0,6352; 0,8177]     | 0,0314                            |
|       |       | $\beta$    | 1,4036                              | [1,2242; 1,6048]     |                                   |
|       |       | $\gamma$   | 3,0869                              | [2,2014; 4,1546]     |                                   |
| 1     | 1     | $q$        | 0,7435                              | [0,6535; 0,8282]     | 0,0677                            |
|       |       | $\beta$    | 1,4303                              | [1,2360; 1,6050]     |                                   |
|       |       | $\gamma$   | 2,9467                              | [2,1469; 3,9197]     |                                   |
| 1     | 2     | $q$        | 0,7595                              | [0,6668; 0,8385]     | 0,1442                            |
|       |       | $\beta$    | 1,4649                              | [1,2583; 1,6461]     |                                   |
|       |       | $\gamma$   | 2,7866                              | [1,9944; 3,7073]     |                                   |
| 1     | 3     | $q$        | 0,7752                              | [0,6952; 0,8495]     | 0,2115                            |
|       |       | $\beta$    | 1,5002                              | [1,3026; 1,6839]     |                                   |
|       |       | $\gamma$   | 2,6279                              | [1,9229; 3,3796]     |                                   |
| 1     | 5     | $q$        | 0,8009                              | [0,7399; 0,8659]     | 0,3093                            |
|       |       | $\beta$    | 1,5634                              | [1,3949; 1,7398]     |                                   |
|       |       | $\gamma$   | 2,3861                              | [1,8138; 2,9559]     |                                   |
| 2     | 1     | $q$        | 0,7357                              | [0,6418; 0,8214]     | 0,0192                            |
|       |       | $\beta$    | 1,4135                              | [1,2156; 1,5926]     |                                   |
|       |       | $\gamma$   | 3,0288                              | [2,1453; 4,0137]     |                                   |
| 3     | 1     | $q$        | 0,7296                              | [0,6363; 0,8209]     | 0,0099                            |
|       |       | $\beta$    | 1,402                               | [1,2080; 1,5929]     |                                   |
|       |       | $\gamma$   | 3,0963                              | [2,1644; 4,1045]     |                                   |
| 5     | 1     | $q$        | 0,7156                              | [0,6200; 0,8089]     | $2 \times 10^{-4}$                |
|       |       | $\beta$    | 1,3739                              | [1,1869; 1,5578]     |                                   |
|       |       | $\gamma$   | 3,2418                              | [2,2769; 4,3217]     |                                   |

É importante destacar que, neste estudo de sensibilidade, os valores dos hiperparâmetros da distribuição a priori de  $\beta$  foram mantidos iguais a 0,001 e o valores dos hiperparâmetros da distribuição a priori de  $q$  iguais a 1. Seria equivalente dizer que as prioris utilizadas nos parâmetros  $q$  e  $\beta$  foram não informativas (difusas).

A partir da Tabela 6 é possível notar que a inferência é robusta quanto à média da priori de  $\gamma$  quando a variância à priori não é pequena. Note que as estimativas dos parâmetros e seus intervalos HPD não variaram muito quando a variância à priori de  $\gamma$  é maior do que 1, mesmo variando os valores da média a priori. Lembre que, se  $\gamma \sim Gama(a_2, b_2)$  então a média e a variância de  $\gamma$  são, respectivamente  $a_2/b_2$  e  $a_2/b_2^2$ . Entretanto, nos casos em que a variância a priori é pequena (prioris informativas) e, conseqüentemente, a média a priori menor do que 1, nota-se menores valores da média a posteriori de  $\gamma$ . Nestes casos é possível notar a influência da distribuição a priori nas estimativas à posteriori dos parâmetros.

O FBST foi realizado para cada par de hiperparâmetros da Tabela 6 considerando o limite 0,05 para o valor-e. Observa-se que no Cenário 1, o verdadeiro valor do parâmetro adotado nas simulações ( $\gamma=2,5$ ) é diferente de 1. Desta forma, para as prioris não informativas (prioris com variância grande), a inferência é dominada pela amostra e, conseqüentemente, a hipótese  $H_{01} : \gamma = 1$  é rejeitada (valor-e menor do que 0,05). Além disso, observa-se que ao passo que  $a_2$  aumenta (aumenta a média a priori de  $\gamma$ ), a média a posteriori de  $\gamma$  tende a aumentar, conseqüentemente o valor-e passa a se tornar cada vez menor, evidenciando mais a rejeição de  $H_{01}$ . Em contrapartida, quando a distribuição a priori passa a ser informativa (variância pequena) e centrada em valores menores do que 1, a estimativa de  $\gamma$  tende a diminuir (se aproximar mais de  $H_{01}$ ), aumentando a evidência de que a hipótese  $H_{01}$  possa ser verdadeira.

A Figura 10 apresenta as distribuições a posteriori para os parâmetros  $q$ ,  $\beta$  e  $\gamma$ , considerando os diversos valores para os hiperparâmetros da distribuição a priori de  $\gamma$ , com  $a_2 = 0,001$ ,  $b_2 = 0,001$  (—);  $a_2 = 0,01$ ,  $b_2 = 0,01$  (—);  $a_2 = 1$ ,  $b_2 = 1$  (—);  $a_2 = 1$ ,  $b_2 = 2$  (—);  $a_2 = 1$ ,  $b_2 = 3$  (—);  $a_2 = 1$ ,  $b_2 = 5$ (—);  $a_2 = 2$ ,  $b_2 = 1$  (—);  $a_2 = 3$ ,  $b_2 = 1$  (—) e  $a_2 = 5$ ,  $b_2 = 1$  (—).

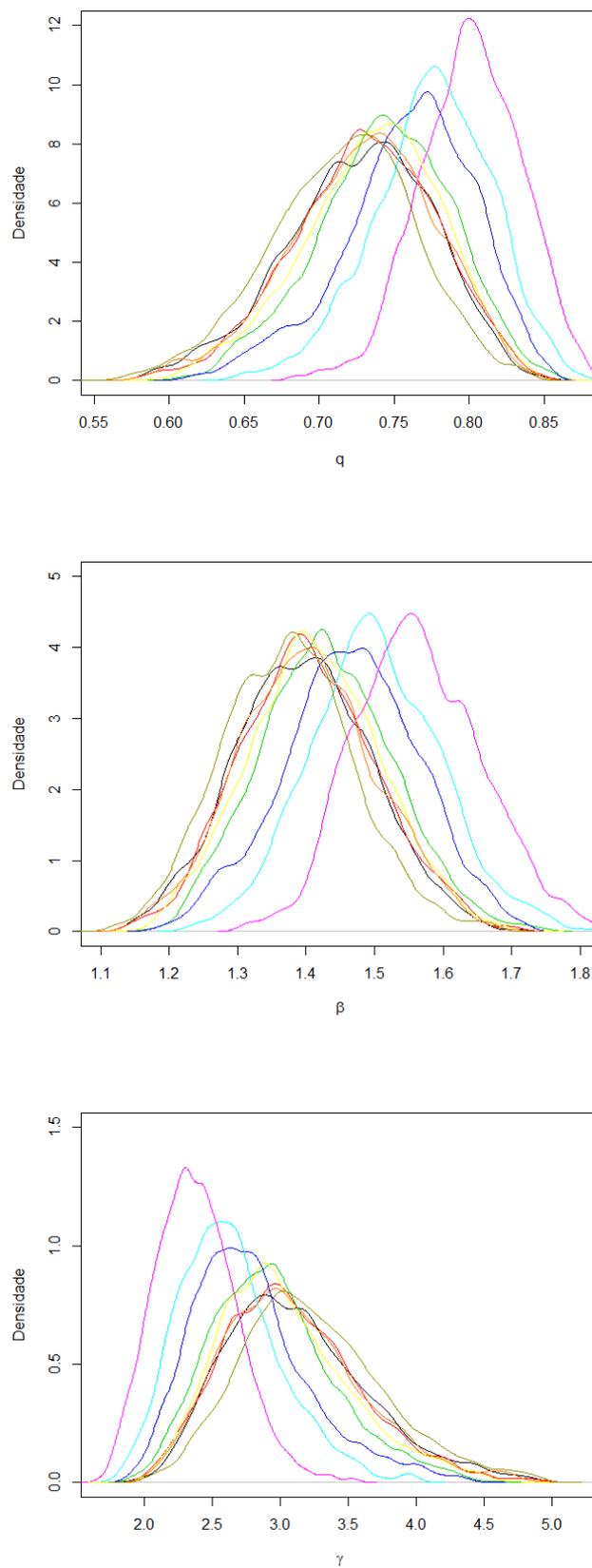


Figura 10: Distribuições a posteriori marginais com diferentes distribuições a priori de  $\gamma$  para o Cenário 1

## 5 Aplicações em dados reais

Neste capítulo são apresentadas duas aplicações em que o modelo WDE é utilizado. Na primeira aplicação os dados possuem taxa de falha unimodal, sendo assim, nesta aplicação espera-se o modelo em estudo tenha um bom ajuste aos dados. Já a segunda aplicação tem função de risco decrescente, dessa forma, espera-se que o modelo proposto obtenha um bom ajuste aos dados e que possivelmente um modelo mais simples (WD e Geométrico Exponenciado - GE) também.

### 5.1 Aplicação 1

Os dados desta aplicação são provenientes do trabalho de Efron (1988). Trata-se de um estudo com 51 pacientes com câncer de pescoço e cabeça realizado pelo Grupo de Oncologia do Norte da Califórnia. Os tempos de sobrevivência em dias destes pacientes são apresentados na Tabela 7.

Tabela 7: Tempos de sobrevivência em dias de pacientes com câncer de pescoço e cabeça

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

Nota: “+” indica as observações censuradas.

Fonte: Efron (1988)

Assim como em Efron (1988), os dados (em dias) da Tabela 7 foram transformados em meses considerando que 1 mês = 30,438 dias e em seguida discretizados. Os dados (em meses) considerando a transformação descrita são apresentados na Tabela 8.

Tabela 8: Tempos de sobrevivência em meses de pacientes com câncer de pescoço e cabeça

| Tempos de sobrevivência em meses  |
|---|
| 0; 1; 1; 2; 2; 2+; 2; 2; 2; 3; 3; 4; 4; 4;<br>4; 4; 4; 4; 4; 5; 5; 5; 5; 5; 5; 5; 5; 6+; 7;<br>7; 7; 8; 8; 9; 9+; 9; 10+; 13; 13; 13;<br>14; 17; 17+; 19; 19; 36; 36+; 37; 40+;<br>44+; 46+; 46 |

Nota: “+” indica as observações censuradas.

Dessa forma, o evento de interesse é o óbito do indivíduo, e a variável resposta  $T$  é o

tempo, em meses completos, do início do estudo até o óbito do paciente ou censura. Aqui  $T = 0$  indica que o paciente morreu antes de completar um mês de diagnóstico positivo.

O banco de dados é composto por 51 observações dentre as quais 09 são censuradas, que corresponde a aproximadamente 18% do conjunto de dados. A Figura 11 mostra a função de sobrevivência estimada pelo método de Kaplan-Meier (K-M) para os pacientes com câncer de pescoço e cabeça.

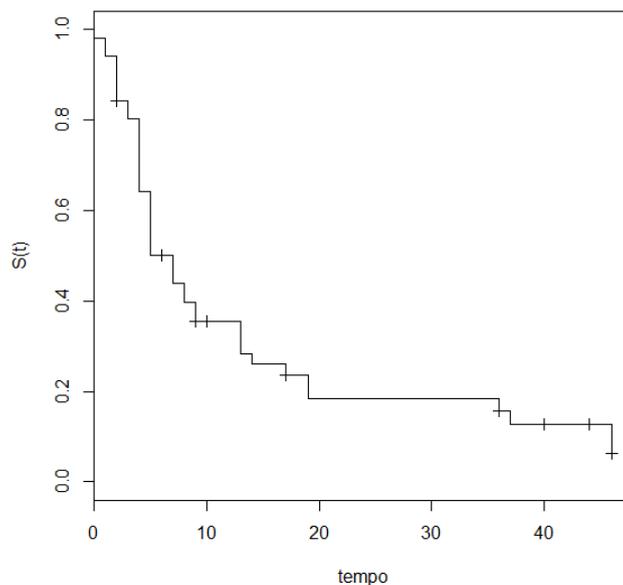


Figura 11: Função de sobrevivência do estimador K-M dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

Como existem várias formas que o gráfico da função de taxa de falha da variável  $T$  pode assumir, é importante utilizar algum método que possa identificar o modelo mais apropriado para esta variável. O gráfico do tempo total em teste (curva TTT), proposto por Aarset (1987), trata-se de um recurso gráfico que auxilia na escolha do modelo probabilístico a ser utilizado com base na função de risco. A Figura 12 apresenta o gráfico da curva TTT para este conjunto de dados.

Observa-se pela Figura 12 que a curva TTT apresenta comportamento côncavo e depois convexo, o que indica que a função taxa de falha é unimodal. Dessa forma, o modelo mais apropriado para a variável em estudo, seria uma distribuição que acomode taxa de falha unimodal. Sendo assim, há evidências que a distribuição WDE se adequa bem aos dados.

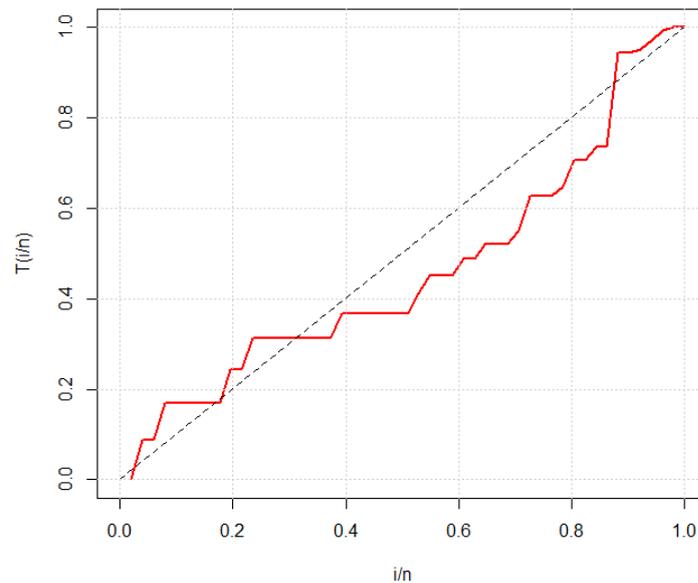


Figura 12: TTT plot dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

### 5.1.1 Abordagem Clássica

A análise descritiva realizada na seção anterior apontou haver evidências que a distribuição WDE é adequada aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça.

Assim, foram calculadas as estimativas pontuais e intervalares dos parâmetros para o modelo WDE aplicado aos dados que são apresentadas na Tabela 9. O nível de confiança considerado para o intervalo de confiança foi de 95%.

Tabela 9: Estimativas do modelo WDE referentes aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

| Parâmetros | Estimativas | Intervalos de Confiança |
|------------|-------------|-------------------------|
| $q$        | 0,16473     | [0,1375; 0,1942]        |
| $\beta$    | 0,2893      | [0,2519; 0,3321]        |
| $\gamma$   | 18,813      | [14,229; 24,873]        |

A seguir foram calculadas as estimativas da função de sobrevivência dos modelos WDE, WD, Geométrico Exponenciado (GE) e Geométrico apresentados em paralelo com as função de sobrevivência estimada pelo método de K-M. Esses resultados são mostrados na Tabela 10.

Tabela 10: Estimativas da função de sobrevivência dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

| Tempo (meses) | Quantidade de falhas | K-M    | WDE    | WD      | GE     | Geométrico |
|---------------|----------------------|--------|--------|---------|--------|------------|
| 0             | 1                    | 0,9804 | 0,9662 | 0,9066  | 0,9354 | 0,9349     |
| 1             | 2                    | 0,9412 | 0,8892 | 0,8342  | 0,8771 | 0,8740     |
| 2             | 5                    | 0,8431 | 0,8077 | 0,7714  | 0,8231 | 0,8171     |
| 3             | 2                    | 0,8030 | 0,7324 | 0,7154  | 0,7727 | 0,7639     |
| 4             | 8                    | 0,6424 | 0,6655 | 0,6649  | 0,7256 | 0,7142     |
| 5             | 7                    | 0,5019 | 0,6068 | 0,61908 | 0,6815 | 0,6677     |
| 7             | 3                    | 0,4391 | 0,5101 | 0,5387  | 0,6016 | 0,5836     |
| 8             | 2                    | 0,3973 | 0,4702 | 0,5032  | 0,5653 | 0,5456     |
| 9             | 2                    | 0,3555 | 0,4349 | 0,4705  | 0,5313 | 0,5101     |
| 13            | 3                    | 0,2844 | 0,3276 | 0,3622  | 0,4147 | 0,3897     |
| 14            | 1                    | 0,2607 | 0,3070 | 0,3397  | 0,3898 | 0,3643     |
| 17            | 1                    | 0,2370 | 0,2559 | 0,2812  | 0,3239 | 0,2977     |
| 19            | 2                    | 0,1843 | 0,2287 | 0,2484  | 0,2864 | 0,2602     |
| 36            | 1                    | 0,1580 | 0,1061 | 0,0906  | 0,1007 | 0,0829     |
| 37            | 1                    | 0,1264 | 0,1022 | 0,0855  | 0,0947 | 0,0775     |
| 46            | 1                    | 0,0632 | 0,0747 | 0,0514  | 0,0545 | 0,0423     |

Ao fazer uma análise entre as diferenças das estimativas dos modelos apresentados e as estimativas empíricas (K-M) obtidos na Tabela 10, nota-se que o modelo WDE possui as estimativas mais próximas ao modelo empírico em quase todos os casos, o que indica um melhor ajuste desse modelo ao conjunto de dados. As estimativas dos modelos GE e Geométrico apresentam as maiores diferenças em relação as estimativas do modelo empírico na maioria dos casos, dando fortes indícios que esses modelos não se ajustam bem aos dados tal como o modelo WD que possui diferenças razoáveis (Figura 13).

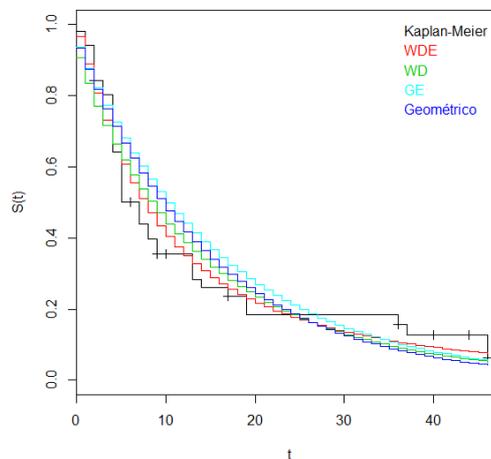


Figura 13: Ajuste dos modelos WDE, WD, GE e Geométrico aos dados (em meses) de pacientes diagnosticados com câncer de pescoço e cabeça

Segundo Nakano e Carrasco (2006), define-se o erro máximo cometido na estimação como um teste estatístico de ajuste de modelos baseados nas diferenças entre as esitmativas do modelo em estudo e as estimativas empíricas (K-M), expresso por:

$$\epsilon = \max|\hat{S}(t) - \hat{S}_{km}(t)|. \quad (43)$$

Ao considerar os modelos da Tabela 10. São apresentados os erros máximos cometidos na estimação dos modelos WDE; WD; GE e Geométrico respectivamente:

$$\begin{aligned} \epsilon_{wde} &= \max|\hat{S}_{wde}(t) - \hat{S}_{km}(t)| \\ &= 0,1049; \end{aligned}$$

$$\begin{aligned} \epsilon_{wd} &= \max|\hat{S}_{wd}(t) - \hat{S}_{km}(t)| \\ &= 0,1171; \end{aligned}$$

$$\begin{aligned} \epsilon_{ge} &= \max|\hat{S}_{ge}(t) - \hat{S}_{km}(t)| \\ &= 0,1796; \end{aligned}$$

e

$$\begin{aligned} \epsilon_g &= \max|\hat{S}_g(t) - \hat{S}_{km}(t)| \\ &= 0,1658. \end{aligned}$$

Tem-se que,  $\epsilon_{wde} < \epsilon_{wd} < \epsilon_g < \epsilon_{ge}$ . Indicando novamente um melhor ajuste do modelo WDE para este conjunto de dados.

O que foi descrito acima, pode ser melhor visualizado na Figura 14.

Para completar o estudo em relação a qual modelo é o mais adequado ao conjunto de dados, a fim de verificar se um modelo mais simples (menos parâmetros) pode ser utilizado, o teste da razão de verossimilhanças (TRV) é realizado para os modelos WD, GE e Geométrico. Considerando o nível de significância  $\alpha = 0,05$  e as seguintes hipóteses:

- $H_{01}$ :  $\gamma = 1$ : O modelo WD mostra-se adequado;
- $H_{02}$ :  $\beta = 1$ : O modelo GE mostra-se adequado;
- $H_{03}$ :  $\gamma = 1$  e  $\beta = 1$ : O modelo Geométrico mostra-se adequado.

Os resultados dos testes são mostrados na Tabela 11.

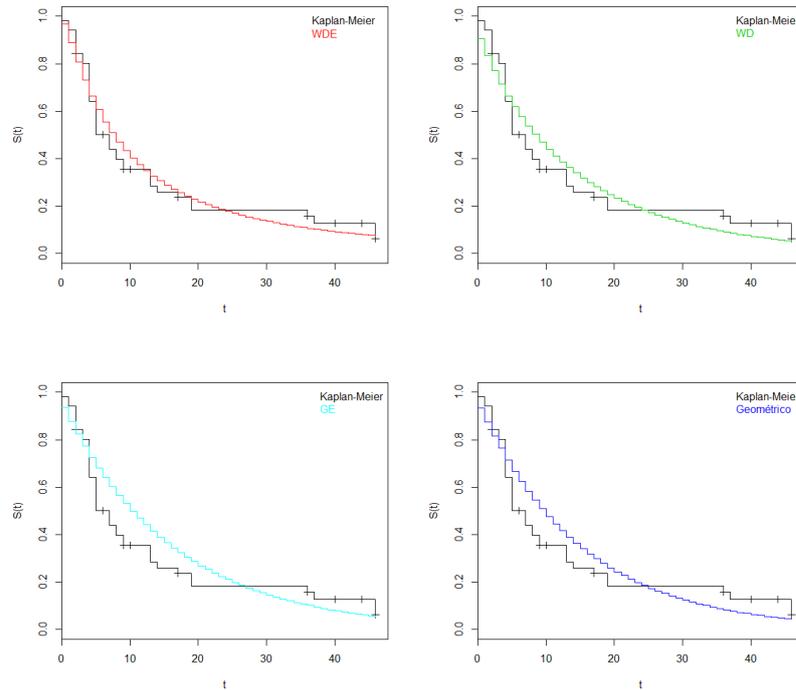


Figura 14: Ajuste dos modelos WDE, WD, GE e Geométrico aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

Tabela 11: TRV para os modelos referentes aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

| Hipótese | Modelo     | Estatística do teste | Graus de Liberdade | valor-p  |
|----------|------------|----------------------|--------------------|----------|
| –        | WDE        | Modelo Completo      |                    |          |
| $H_{01}$ | WD         | 10,39533             | 1                  | 0,001263 |
| $H_{02}$ | GE         | 10,99681             | 1                  | 0,000913 |
| $H_{03}$ | Geométrico | 10,57452             | 2                  | 0,005056 |

A partir da Tabela 11 comprova-se que o modelo WDE mostra-se mais adequado para este banco de dados, uma vez que, o valor-p é menor que o nível de significância ( $\alpha = 0,05$ ) para os três modelos em teste, sendo assim, as hipóteses  $H_{01}$ ,  $H_{02}$  e  $H_{03}$  são fortemente rejeitadas. Conclui-se assim, que modelos mais simples como os modelos WD e GE que acomodam taxas de falhas crescente, decrescente e constante e o modelo Geométrico que acomodam taxas de falha constante não seriam modelos adequados para este banco de dados.

A Figura 15 apresenta graficamente as estimativas da função de sobrevivência do estimador de Kaplan-Meier e do modelo WDE com o intervalo de confiança de 95% para a função de sobrevivência do modelo WDE calculados via *método delta*.

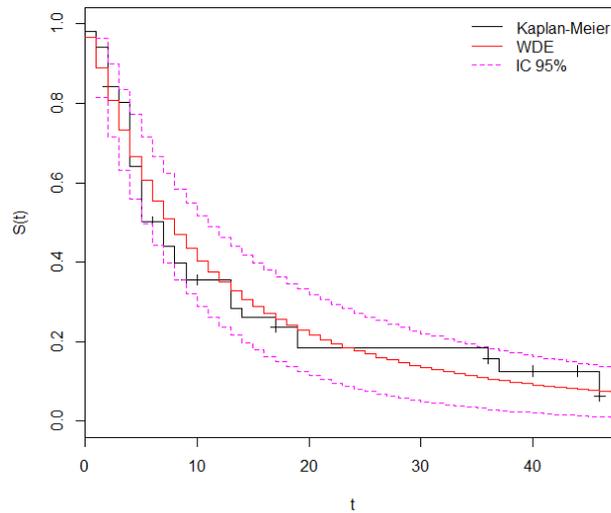


Figura 15: Ajuste do modelo WDE com intervalo de confiança 95% aos tempos de sobrevivência (em meses) de pacientes diagnosticados com câncer de pescoço e cabeça

A partir da Figura 15 pode-se observar o bom ajuste do modelo WDE aos dados, o grande número de observações censuradas no fim do estudo, ocasionou um ajuste razoável do modelo aos dados ao final da curva. Entretanto, todas as estimativas empíricas encontram-se dentro do intervalo de confiança de 95% para a função de sobrevivência do modelo, o que evidencia ainda mais o bom ajuste.

A Figura 16 refere-se a taxa de falha ajustada pelo modelo WDE aplicado aos dados de pacientes com câncer de pescoço e cabeça.

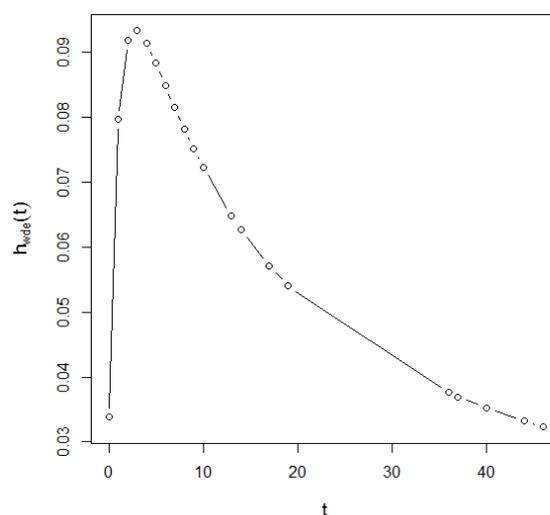


Figura 16: Taxa de falha ajustada pelo modelo WDE referentes aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

Como esperado, a taxa de falha para esta aplicação é unimodal. Pode-se observar no compartimento da curva, que a taxa de falha atinge um pico, que se encontra no terceiro mês ( $t= 3$  meses) com probabilidade aproximadamente de 0,0932, diminuindo gradualmente em seguida.

### 5.1.2 Abordagem Bayesiana

Para a aplicação 1, foram adotadas como prioris não informativas,  $q \sim Beta(1; 1)$ ,  $\beta \sim Gama(0, 001; 0, 001)$  e  $\gamma \sim Gama(0, 001; 0, 001)$ , uma vez que,  $0 < q < 1$ ,  $\beta > 0$  e  $\gamma > 0$ . Toda inferência dos parâmetros foi realizada via *Markov Chain Monte Carlo* (MCMC) através do pacote *MCMCPack* do *software* “R”, que utiliza como núcleo de transição uma cadeia de passeio aleatório.

Na Tabela 12 são apresentadas as estimativas pontuais (médias a posteriori) do modelo WDE com os respectivos intervalos de credibilidade HPD de 95% e o FBST para os parâmetros, levando em consideração as seguintes hipóteses:

- $H_{01}$ :  $\gamma = 1$ : O modelo WD mostra-se adequado;
- $H_{02}$ :  $\beta = 1$ : O modelo GE mostra-se adequado;
- $H_{03}$ :  $\gamma = 1$  e  $\beta = 1$ : O modelo Geométrico mostra-se adequado.

Tabela 12: Inferência bayesiana dos parâmetros do modelo WDE dos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça

| Parâmetros | Estimativas<br>(médias a posteriori) | Intervalo<br>HPD<br>95% | valor-e               |                      |                                       |
|------------|--------------------------------------|-------------------------|-----------------------|----------------------|---------------------------------------|
|            |                                      |                         | $H_{01} : \gamma = 1$ | $H_{02} : \beta = 1$ | $H_{03} : \gamma = 1;$<br>$\beta = 1$ |
| $q$        | 0,1935                               | [0,1579; 0,2452]        |                       |                      |                                       |
| $\beta$    | 0,3185                               | [0,2888; 0,3542]        | 0,0208                | 0,0177               | 0,0178                                |
| $\gamma$   | 17,425                               | [13,142; 19,997]        |                       |                      |                                       |

As estimativas pontuais (médias a posteriori) dos parâmetros segundo abordagem bayesiana são próximas as estimativas pontuais segundo a abordagem clássica. Entretanto, as estimativas intervalares estimadas na inferência bayesiana (intervalo HPD 95%) para os parâmetros  $\beta$  e  $\gamma$ , tem amplitude menor que as estimativas intervalares estimadas na inferência clássica. O FBST, rejeita as hipóteses  $H_{01}$ ,  $H_{02}$  e  $H_{03}$  como esperado pelo fato da taxa de falha ser unimodal, pois o valor-e é menor que 0,05. Sendo assim, os modelos WD, GE e Geométrico não se mostram adequados para o conjunto de dados em estudo.

Para verificação do ajuste do modelo WDE considerando o processo inferencial bayesiano, na Figura 17 são apresentadas a curva de sobrevivência estimada pelo estimador de

K-M com a curva de sobrevivência estimada pelo modelo WDE sobreposta, acompanhada do intervalo HPD de 95% para a função de sobrevivência do modelo.

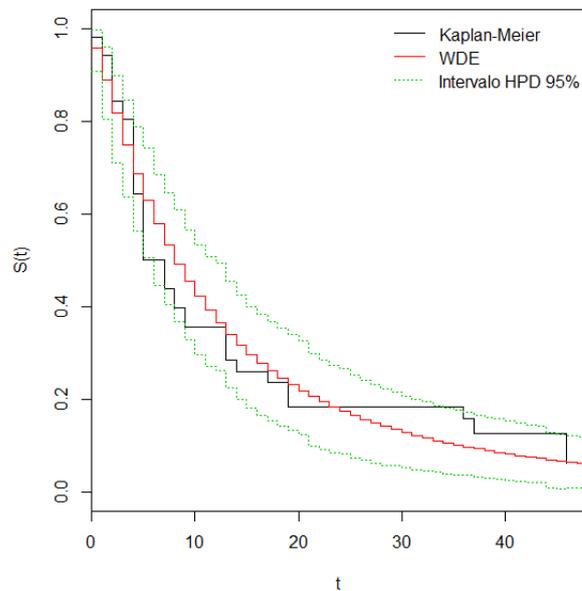


Figura 17: Ajuste do modelo WDE aos tempos de sobrevivência (em meses) de pacientes com câncer de pescoço e cabeça segundo a abordagem bayesiana

Nota-se pela Figura 17, tal como na abordagem clássica, que o modelo WDE na abordagem bayesiana obteve um bom ajuste aos dados. Sendo que, praticamente todas as estimativas da função de sobrevivência estimadas pelo método de K-M encontram-se dentro do intervalo HPD de 95% para a função de sobrevivência do modelo WDE.

## 5.2 Aplicação 2

Os dados para a aplicação 2 são referentes dos trabalhos de Corrêa *et al.* (2016) e Silva *et al.* (2017), que mediram o tempo até o alívio da dor em pacientes com dor lombar crônica não específica submetidos a sessões de 30 minutos de estimulação com uma corrente interferencial. Cada um dos pacientes realizaram 3 sessões por semana em dias alternados, por 4 semanas, totalizando 12 sessões. Neste trabalho, o evento de interesse (alívio da dor) foi definido como o número de sessões que cada paciente foi submetido até ser observado o evento de interesse, que é a diminuição ou alívio da intensidade da dor. Os tempos de sobrevivência em número de sessões são apresentados na Tabela 13 a seguir.

Tabela 13: Tempos de sobrevivência referentes a Aplicação 2

| Tempos de sobrevivência em número de sessões       |
|--|
| 1;10;2;1;1;1;2;1;1;1;5;1;12+;2;2;2;1;1;4;2;1;2;5;  |
| 2;1;1;3;1;1;1;2;1;5;4;1;1;2;1;1;1;1;1;1;1;6;1;     |
| 1;8;2;1;1;1;1;2;1;4;5;6;1;1;1;6;1;1;1;1;1;2;1;1;1; |
| 1;4;2;2;1;1;1;1;3;3;1;1;3;3;1;3;1;2.               |

Nota: “+” indica as observações censuradas.

O alívio da dor foi caracterizado como a redução em no mínimo 50% da escala numérica verbal de dor, em relação ao valor observado no início do tratamento. Aqui, a variável  $T$  foi definida como o número de sessões até o alívio da dor menos 1, isto é,  $T = 0$  indica que a dor foi aliviada em apenas uma sessão ( $T = 0; 1; 2; \dots; 11$ ). A variável  $T$  foi definida dessa maneira de forma que o valor mínimo fosse zero (visto que o número mínimo de sessões é igual a um) ajustando-se ao suporte da distribuição em estudo.

O banco de dados composto por 100 observações dentre as quais apenas uma é censurada, que corresponde a 1% do conjunto de dados que se encontra no tempo final do estudo. A Figura 18 mostra a função de sobrevivência estimada pelo método de K-M para os pacientes submetidos a sessões para alívio da dor lombar crônica não específica que evidencia o que foi descrito acima.

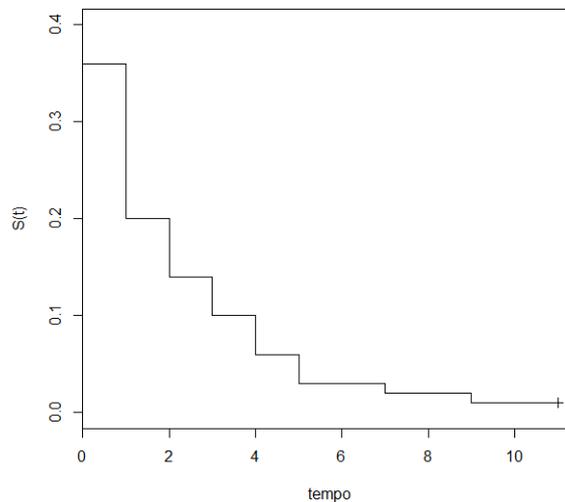


Figura 18: Função de sobrevivência do estimador K-M dos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

A Figura 19 apresenta o gráfico do tempo total em teste (curva TTT) para este conjunto de dados.

Nota-se pela Figura 19 que a curva TTT apresenta comportamento convexo, o que indica que a função taxa de falha é monotonicamente decrescente.

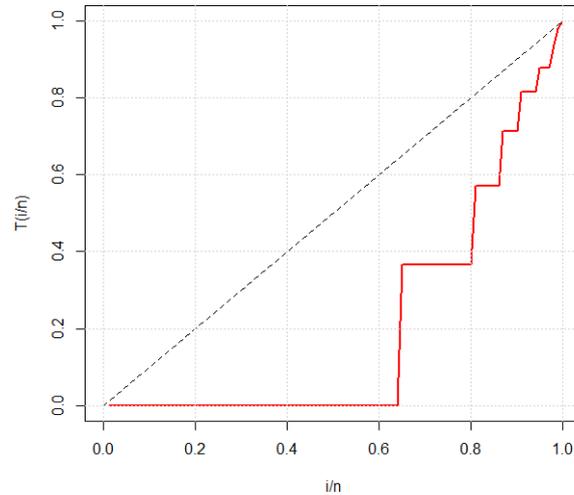


Figura 19: TTT plot referente aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

Dessa forma, o modelo mais apropriado para a variável em estudo, seria uma distribuição que acomode taxa de falha decrescente. Sendo assim, supostamente a distribuição WDE se adequa bem aos dados. Além disso há evidências que modelos mais simples com apenas dois parâmetros, como os modelos WD e GE sejam adequados, uma vez que, estes acomodam taxas de falhas crescente, decrescente e constante.

### 5.2.1 Abordagem Clássica

A análise descritiva realizada na seção anterior apontou haver evidências que a distribuição WDE, WD e GE são adequadas aos tempo de sobrevivência referentes a aplicação 2. Assim, foram calculadas as estimativas pontuais e intervalares dos parâmetros para os modelos WDE, WD e GE aplicados aos dados de pacientes com dor lombar crônica não específica que são apresentadas na Tabela 14. O nível de confiança considerado para o intervalo de confiança foi de 95%.

Tabela 14: Estimativas dos modelos WDE,WD e GE referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

| Modelo | Parâmetros | Estimativas | Intervalos de Confiança |
|--------|------------|-------------|-------------------------|
| WDE    | $q$        | 0,5528      | [0,4627; 0,6338]        |
|        | $\beta$    | 0,8306      | [0,6900; 0,9997]        |
|        | $\gamma$   | 0,5417      | [0,3906; 0,7511]        |
| WD     | $q$        | 0,3399      | [0,2678; 0,4132]        |
|        | $\beta$    | 0,6073      | [0,4925; 0,7490]        |
| GE     | $q$        | 0,6898      | [0,6014; 0,7624]        |
|        | $\gamma$   | 0,3693      | [0,2663; 0,5121]        |

Em seguida, foram calculadas as estimativas da função de sobrevivência dos modelos WDE, WD, GE e Geométrico apresentados em paralelo com as função de sobrevivência estimada pelo método de K-M. Esses resultados são mostrados na Tabela 15.

Tabela 15: Estimativas da função de sobrevivência dos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

| Tempo (sessões) | Quantidade de falhas | K-M  | WDE    | WD     | GE     | Geométrico |
|-----------------|----------------------|------|--------|--------|--------|------------|
| 0               | 64                   | 0,36 | 0,3533 | 0,3399 | 0,3509 | 0,5099     |
| 1               | 16                   | 0,20 | 0,2071 | 0,1933 | 0,2122 | 0,2600     |
| 2               | 6                    | 0,14 | 0,1311 | 0,1221 | 0,1366 | 0,1326     |
| 3               | 4                    | 0,10 | 0,0862 | 0,0817 | 0,0904 | 0,0676     |
| 4               | 4                    | 0,06 | 0,0582 | 0,0568 | 0,0608 | 0,0345     |
| 5               | 3                    | 0,03 | 0,0399 | 0,0406 | 0,0412 | 0,0176     |
| 7               | 1                    | 0,02 | 0,0195 | 0,0220 | 0,0192 | 0,0046     |
| 9               | 1                    | 0,01 | 0,0098 | 0,0127 | 0,0091 | 0,0012     |

Assim, tem-se o erro máximo cometido na estimação para os modelos em estudo dados por:

$$0,0122 = \epsilon_{ge} < 0,0138 = \epsilon_{wde} < 0,0183 = \epsilon_{wd} < 0,1499 = \epsilon_g.$$

Ao verificar as diferenças entre as estimativas dos modelos apresentados, as estimativas empíricas (K-M) obtidos na Tabela 15 e o erro máximo cometido na estimação, nota-se que os modelos WDE e GE possuem as estimativas mais próximas ao modelo empírico em quase todos os casos, o que indica um melhor ajuste destes modelos ao conjunto de dados (Figura 20)

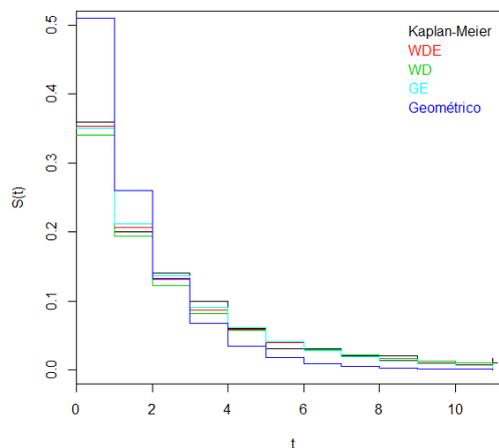


Figura 20: Ajuste dos modelos WDE, WD, GE e Geométrico aos dados (em número de sessões) de pacientes com dor lombar crônica não específica

Nota-se também, que as estimativas do modelo WD mantêm-se sempre próximas das estimativas de K-M, confirmando as evidências de que este modelo também possa ser utilizado para este banco de dados. Em contrapartida, o modelo Geométrico não possui todas as estimativas próximas ao modelo empírico como nos outros casos e o erro máximo cometido na estimação é bem maior em relação aos outros modelos. Sendo assim, há evidências que este modelo não seja adequado para os dados referentes a aplicação 2. Os resultados obtidos que foram descritos podem ser melhor visualizados na Figura 21.

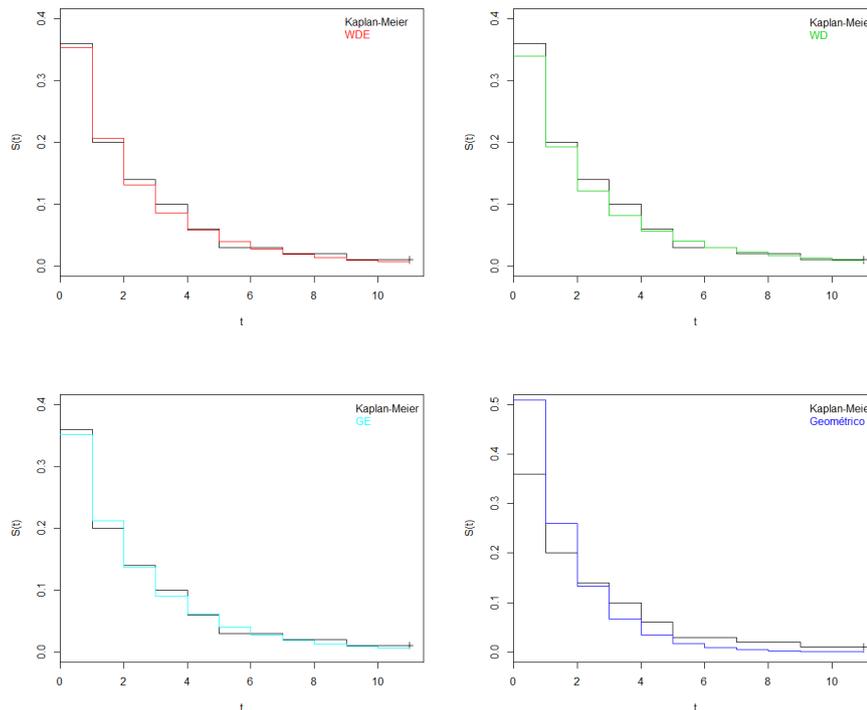


Figura 21: Ajuste dos modelos WDE, WD, GE e Geométrico aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

Para verificar se de fato os modelos WD e GE seriam modelos ideais para o conjunto de dados em estudo e se a evidência de que o modelo Geométrico não seja adequado para os dados da aplicação 2, o TRV é realizado para os três modelos. Considerando o nível de significância  $\alpha = 0,05$  e as seguintes hipóteses:

- $H_{01}$ :  $\gamma = 1$ : O modelo WD mostra-se adequado;
- $H_{02}$ :  $\beta = 1$ : O modelo GE mostra-se adequado;
- $H_{03}$ :  $\gamma = 1$  e  $\beta = 1$ : O modelo Geométrico mostra-se adequado.

A Tabela 16 refere-se aos resultados dos testes.

Tabela 16: TRV para os modelos referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

| Hipótese | Modelo     | Estatística do teste | Graus de Liberdade | valor-p  |
|----------|------------|----------------------|--------------------|----------|
| –        | WDE        |                      | Modelo Completo    |          |
| $H_{01}$ | WD         | 0,046687             | 1                  | 0,828932 |
| $H_{02}$ | GE         | 0,302798             | 1                  | 0,582134 |
| $H_{03}$ | Geométrico | 17,66773             | 2                  | 0,000146 |

Os resultados obtidos na Tabela 16 comprovam que os modelos WD e GE também pode ser utilizados, uma vez que, o valor-p é maior que o nível de significância ( $\alpha = 0,05$ ). Sendo assim, as hipóteses  $H_{01}$  e  $H_{02}$  não são rejeitadas para os modelos em teste. Porém, para o modelo Geométrico, a hipótese  $H_{03}$  é rejeitada pois o valor-p é menor que o nível de significância. Conclui-se assim, que o modelos mais simples como os modelos WD e GE que acomodam taxas de falhas crescente, decrescente e constante também são modelos adequados para os dados da aplicação 2. Já o modelo Geométrico que acomoda apenas taxas de falha constante não se mostra adequado.

Enfatiza-se que, apesar do modelo WD também poder ser utilizado por este acomodar taxas de falha crescente, decrescente e constante, o modelo WDE mostra-se como um modelo melhor para o ajuste dos dados, pois apresenta a curva de sobrevivência mais próxima a curva de sobrevivência empírica e as estimativas mais próximas as estimativas de K-M em relação ao modelo supracitado. O modelo WDE também apresenta vantagens em relação ao modelo GE, pois o modelo em estudo é mais versátil para o ajuste de dados em comparação a este modelo.

Assim, a Figura 22 apresenta graficamente as estimativas da função de sobrevivência do estimador de Kaplan-Meier e dos modelos WDE, WD e GE, com o intervalo de confiança de 95% para a função de sobrevivência dos modelos WDE, WD e GE calculados via *método delta*.

Na Figura 22, observa-se o bom ajuste dos modelos WDE, WD e GE aos dados. Evidenciando que modelos mais simples (menos parâmetros) como os modelos WD e GE também podem ser utilizados para os dados desta aplicação. Nota-se também que, todas as estimativas empíricas encontram-se dentro do intervalo de confiança de 95% para a função de sobrevivência dos modelos WDE, WD e GE, o que evidencia ainda mais o bom ajuste.

A Figura 23 refere-se a taxa de falha ajustada pelos modelos WDE, WD e GE aplicados aos dados de pacientes com dor lombar crônica não específica.

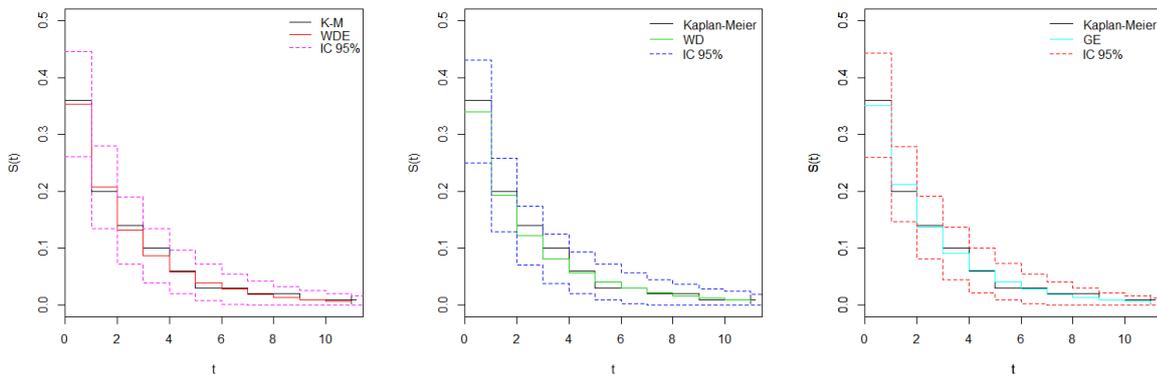


Figura 22: Ajuste dos modelos WDE, WD e GE com intervalo de confiança 95% aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

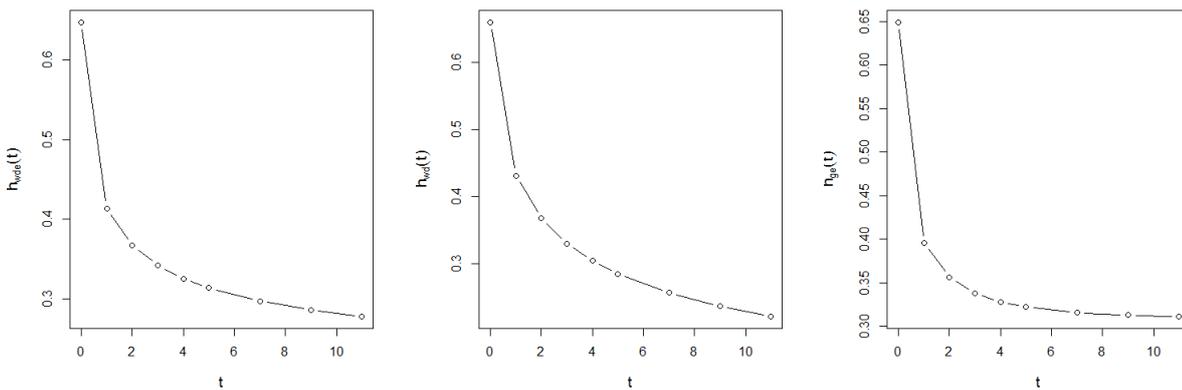


Figura 23: Taxa de falha ajustada pelos modelos WDE, WD e GE referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

A taxa de falha para esta aplicação como esperado é monótona decrescente. Pode-se observar no compartimento das curvas, que a taxa de falha tem um ponto de máximo nos três casos em  $t = 0$ , que representa uma sessão com probabilidade 0,6467 para o modelo WDE; 0,6601 para o modelo WD e 0,6490 para o modelo GE. Também verifica-se, que as probabilidades citadas vão diminuindo a medida que o tempo aumenta.

### 5.2.2 Abordagem Bayesiana

Para a aplicação 2, novamente foram adotadas como prioris não informativas,  $q \sim \text{Beta}(1; 1)$ ,  $\beta \sim \text{Gama}(0,001; 0,001)$  e  $\gamma \sim \text{Gama}(0,001; 0,001)$ , uma vez que,  $0 < q < 1$ ,  $\beta > 0$  e  $\gamma > 0$ . Toda inferência dos parâmetros foi realizada via *Markov Chain Monte Carlo* (MCMC) através do pacote *MCMCPack* do *software* “R”, que utiliza como núcleo de transição uma cadeia de passeio aleatório.

Na Tabela 17 são apresentadas as estimativas pontuais (médias a posteriori) dos modelos WDE, WD e GE com os respectivos intervalos de credibilidade HPD de 95%.

Tabela 17: Inferência bayesiana dos parâmetros dos modelos WDE, WD e GE dos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

| Modelo | Parâmetros | Estimativas<br>(médias a posteriori) | Intervalo HPD<br>95% |
|--------|------------|--------------------------------------|----------------------|
| WDE    | $q$        | 0,5525                               | [0,2660; 0,8965]     |
|        | $\beta$    | 0,8393                               | [0,4179; 1,3583]     |
|        | $\gamma$   | 0,6194                               | [0,1486; 1,2291]     |
| WD     | $q$        | 0,3608                               | [0,2743; 0,4572]     |
|        | $\beta$    | 0,6211                               | [0,4664; 0,7848]     |
| GE     | $q$        | 0,6917                               | [0,5704; 0,7973]     |
|        | $\gamma$   | 0,3831                               | [0,2174; 0,5761]     |

Assim como na aplicação 1, na aplicação 2 as estimativas pontuais clássicas estão próximas às estimativas pontuais bayesianas. As estimativas intervalares bayesianas neste caso tem amplitude maior que as estimativas intervalares clássicas, sendo que o intervalo HPD de 95% para o parâmetro  $\gamma$  inclui o valor 1 (distribuição WDE se reduz a distribuição WD) e para o parâmetro  $\beta$  também inclui o valor 1 (distribuição WDE se reduz a distribuição GE).

Como critério de seleção de modelos o FBST é apresentado na Tabela 18, levando em consideração as seguintes hipóteses:

- $H_{01}$ :  $\gamma = 1$ : O modelo WD mostra-se adequado;
- $H_{02}$ :  $\beta = 1$ : O modelo GE mostra-se adequado;
- $H_{03}$ :  $\gamma = 1$  e  $\beta = 1$ : O modelo Geométrico mostra-se adequado.

Tabela 18: FBST para os modelos referentes aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica

| Hipótese | Modelo     | valor-e |
|----------|------------|---------|
| $H_{01}$ | WD         | 0,2795  |
| $H_{02}$ | GE         | 0,3331  |
| $H_{03}$ | Geométrico | 0,0000  |

A partir da Tabela 18 nota-se que, o FBST não rejeita as hipóteses  $H_{01}$  e  $H_{02}$ , pois nesses casos o valor-e é maior que 0,05 pelo fato da taxa de falha ser decrescente, sendo assim, os modelos WD e GE também mostram-se adequados para o conjunto de dados da aplicação 2. Em contrapartida, na hipótese  $H_{03}$  o valor-e estimado é menor que 0,05,

sendo assim, o modelo Geométrico não é adequado para modelar os tempos (em número de sessões) até o alívio da dor em pacientes com dor lombar crônica não específica

Para verificação do ajuste dos modelos aos dados considerando o processo inferencial bayesiano na Figura 24 é apresentada a curva de sobrevivência estimada pelo estimador de K-M com a curva de sobrevivência estimada pelos modelos WDE, WD e GE acompanhada dos respectivos intervalos HPD de 95% para a função de sobrevivência destes modelos.

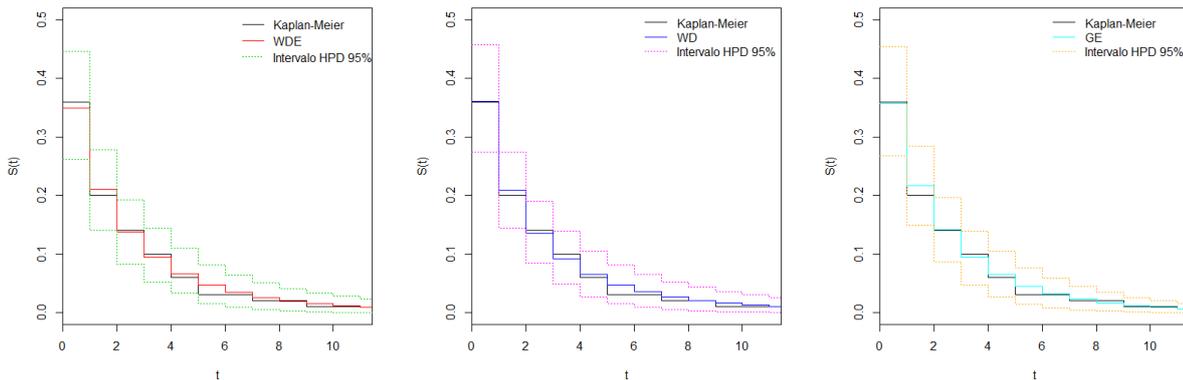


Figura 24: Ajuste dos modelos WDE, WD e GE aos tempos de sobrevivência (em número de sessões) de pacientes com dor lombar crônica não específica segundo a abordagem bayesiana

Na Figura 24 observa-se o bom ajuste dos modelos WDE, WD e GE aos dados e que a curva da função de sobrevivência estimada pelo método de K-M encontra-se dentro do intervalo HPD de 95% para as funções de sobrevivência dos modelos WDE, WD e GE evidenciando ainda mais o bom ajuste destes modelos aos dados.



## 6 Considerações Finais

A distribuição WDE é uma distribuição bastante versátil, uma vez que, acomoda taxas de falha monotónas (crescente, decrescente), constante e não-monotónas (forma de banheira e unimodal). Sendo assim, uma das vantagens de se trabalhar com essa distribuição é o fato dela fornecer um ajuste paramétrico em que se permite trabalhar com vários tipos de taxas de falha existentes na literatura.

Essa versatilidade se deve ao fato da distribuição WDE apresentar 3 parâmetros. Sendo que, o parâmetro de forma  $\gamma > 0$  introduzido como expoente na função distribuição acumulada base em conjunto com o parâmetro de forma  $\beta > 0$ , permite essa flexibilidade. Além disso a distribuição em estudo conta com o parâmetro  $0 < q < 1$  que representa uma probabilidade. Alguns sub-casos especiais dessa distribuição são possíveis quando são fixados certos valores para  $\beta$  e  $\gamma$ .

A distribuição WDE foi proposta por Nekoukhou e Bidram (2015). No entanto, as inferências dos parâmetros do modelo apresentado pelos mesmos não consideraram uma possível presença de censuras, que é uma característica primordial em análise de sobrevivência. Sendo que, o processo inferencial considerado por esses autores foi realizado apenas seguindo a abordagem clássica.

A motivação deste trabalho surgiu do interesse de utilizar essa distribuição que contém propriedades atrativas no contexto de análise de sobrevivência. Dessa forma, neste trabalho considerou-se um processo inferencial que permitisse a presença de dados censurados, utilizando as duas abordagens de Inferência Estatística existentes na literatura: a abordagem clássica e a abordagem bayesiana.

A formulação do modelo WDE, ao considerar o processo inferencial clássico foi desenvolvido a partir do método de máxima verossimilhança utilizando métodos computacionais de otimização para a obtenção das estimativas. Já o processo inferencial bayesiano, apesar da distribuição a posteriori conjunta não ser conhecida, amostras da mesma foram facilmente obtidas através dos métodos MCMC (*Markov Chain Monte Carlo*). Como forma de seleção de modelos, testes paramétricos de hipóteses foram realizados. Na abordagem clássica considerou-se o Teste da Razão de Verossimilhanças (TRV) e na abordagem bayesiana considerou-se o FBST (*Full Bayesian Significance Test*).

No Capítulo 4, foram descritas as simulações computacionais realizadas via *software* R, os tempos de sobrevivência foram gerados considerando quatro cenários definidos a partir da taxa de falha da distribuição dependendo da combinação dos parâmetros. O vício e o EQM para cada parâmetro da distribuição segundo o processo inferencial clássico para os quatro cenários foram calculados para diferentes tamanhos de amostra e percentuais de censura. De modo geral, todos os cenários apresentaram bons resultados. As

características dos dados, tais como tamanho de amostra e percentual de censura influenciaram nas estimativas dos parâmetros. Entretanto, esse impacto foi pequeno, visto que em todos os casos analisados, nos quatro cenários, as medidas das estimativas estavam muito próximas dos verdadeiros valores dos parâmetros e os valores do EQM eram bem pequenos.

Além disso, foi realizado um estudo de análise de sensibilidade da escolha dos hiperparâmetros da distribuição a priori de  $\gamma$ , em que pôde-se observar que o processo inferencial bayesiano adotado é robusto quanto à média da priori de  $\gamma$  quando a variância à priori não é pequena.

As aplicações apresentadas no Capítulo 5 permitiram ilustrar o uso do modelo WDE em dados reais nas abordagens clássica e bayesiana. Para a primeira aplicação referente ao tempo de sobrevivência em meses de pacientes diagnosticados com câncer de pescoço e cabeça que tinha taxa de falha unimodal, apenas o modelo WDE mostrava-se adequado para ajustar os dados nas duas abordagens em detrimento dos outros modelos testados (WD, GE e Geométrico). Fato que, foi comprovado com a realização de testes de hipóteses, o TRV e o FBST, para as abordagens clássica e bayesiana respectivamente. Esses testes não apresentaram resultados significativos para os modelos sob  $H_0$ . Além do mais, através da curva de sobrevivência empírica (K-M) e as curvas de sobrevivência estimadas, o modelo WDE apresentou um melhor ajuste aos dados, tendo as estimativas mais próximas ao modelo empírico em quase todos os casos.

A segunda aplicação considera o tempo em número de sessões até o alívio da dor em pacientes com dor lombar crônica não específica. Essa aplicação que tinha taxa de falha decrescente, além do modelo WDE, os modelos WD e GE mostravam-se adequados para ajustar os dados nas duas abordagens, fato contrário ao modelo Geométrico. Para comprovação de que os modelos testados realmente tinham um bom ajuste, testes de hipóteses foram realizados na abordagem clássica (TRV) e na abordagem bayesiana (FBST). Sendo que, para os modelos WD e GE os testes apresentaram resultados significativos e o modelo Geométrico não apresentou. Além disto, a curva de sobrevivência empírica e as curvas de sobrevivência estimadas revelaram de fato que os modelos WDE, WD e GE possuíam bons ajustes aos dados. No entanto, o modelo WDE apresentou-se de forma mais atrativa pela flexibilidade do ajuste de dados que é proporcionado por este modelo.

De maneira geral, ao considerar as aplicações, a formulação do modelo WDE no contexto de Análise de Sobrevivência considerando as abordagens clássica e bayesiana, com exceção das estimativas intervalares para a aplicação 2, não revelou grandes diferenças entre as duas abordagens em relação as estimativas apresentadas, mantendo-se sempre relativamente próximas.

Desta forma, o modelo WDE nas duas abordagens em estudo, mostra-se como um

modelo mais flexível para o ajuste dos dados de sobrevivência discretos, do que outros modelos para o mesmo fim. Como os modelos WD e GE que acomodam apenas taxas de falha monótonas e constante e o modelo Geométrico que acomoda apenas taxas de falha constante.

Como propostas para trabalhos futuros sugere-se:

- A inclusão de covariáveis no modelo WDE a fim de considerar um modelo de regressão;
- Inferências do modelo WDE considerando a metodologia da fração de cura.



## Referências

- M. V. Aarset. How to identify a bathtub hazard rate. *IEEE Transactions on Reliability*, 36(1):106–108, 1987.
- H. Bolfarine e M. C. Sandoval. *Introdução a Inferência Estatística*. SBM, Rio de Janeiro, 2010.
- G. H. V. Brunello e E. Y. Nakano. Inferência Bayesiana no modelo Weibull discreto em dados com presença de censura. *TEMA Tend. Mat. Apl. Comput.*, 16(2):97–110, 2015.
- B. P. Carlin e T. A. Louis. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall/CRC, 2000.
- J. M. F. Carrasco, E. M. M. Ortega, e G. M. Cordeiro. A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics and Data Analysis*, 53(2): 450–462, 2008.
- E. A. Colosimo e S. R. Giolo. *Análise de Sobrevida Aplicada*. Edgard Blucher, São Paulo, 2006. ABE - Projeto Fisher.
- J. B. Corrêa, L. O. Costa, N. T. Oliveira, W. P. Lima, K. A. Sluka, e R. E. Liebano. Effects of the carrier frequency of interferential current on pain modulation and central hypersensitivity in people with chronic nonspecific low back pain: A randomized placebo-controlled trial. *European Journal of Pain*, 20(10):1653–1666, 2016.
- A. J. Dobson e A. Barnett. *An Introduction to Generalized Linear Models*. CRC Press, 2008.
- B. Efron. Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of The American Statistical Association*, 83:414–425, 1988. doi: 10.1080/01621459.1988.10478612.
- R. S. Ehlers. Inferência Bayesiana. *Departamento de Matemática Aplicada e Estatística, ICMC-USP*, 2011.
- L. M. Fernandes. Inferência Bayesiana em modelos discretos com fração de cura. Dissertação (Mestrado em Estatística) - Departamento de Estatística, Universidade de Brasília, Brasília., 2013.
- D. Gamerman e H. F. Lopes. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. CRC Press, 2006.

- W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- E. L. Kaplan e P. Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.
- C. D. Lai. Generalized Weibull Distributions. In *Generalized Weibull Distributions*, pages 23–75. Springer, 2014.
- C. D. Lai, M. Xie, e D. N. P. Murthy. A modified Weibull distribution. *IEEE Transactions on reliability*, 52(1):33–37, 2003.
- M. N. Magalhães e C. P. D. Lima. *Noções de Probabilidade e Estatística*. EDUSP, São Paulo, 2013. 7<sup>a</sup> ed. rev.
- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, e E. Teller. Equations of state calculations by fast computing machines. *The journal of Chemical Physics*, 21(6):1087–1092, 1953.
- G. S. Mudholkar e D. K. Srivastava. Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2):299–302, 1993.
- T. Nakagawa e S. Osaki. The discrete weibull distribution. *IEEE Transactions on Reliability*, 24(5):300–301, 1975.
- E. Y. Nakano e C. G. Carrasco. Uma avaliação do uso de um modelo contínuo na análise de dados discretos de sobrevivência. *TEMA Tend. Mat. Apl. Comput.*, 7(1):91–100, 2006.
- V. Nekoukhrou e H. Bidram. The exponentiated discrete Weibull distribution. *SORT*, 39:127–146, 2015.
- C. D. Paulino, M. A. A. Turkman, e B. Murteira. *Estatística Bayesiana*. Fundação Calouste Gulbekian, Lisboa, 2003.
- C. A. B. Pereira e J. M. Stern. Evidence and credibility: full bayesian significance test of precise hypothesis. *Entropy*, 1(4):99–110, 1999.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2015. URL <https://www.R-project.org/>.
- T. Ryan. *Estatística Moderna para engenharia*. Elsevier Brasil, 2009.
- J. F. Silva, R. E. Liebano, J. B. Corrêa, R. Y. Matsushita, e E. Y. Nakano. Análise do tempo até o alívio da intensidade da dor em pacientes com dor lombar crônica não

específica via modelo de riscos proporcionais de Cox. *Ciência e Natura*, 39(2):(no prelo), 2017.

M. Xie, Y. Tang, e T. N. Goh. A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering and System Safety*, 76(3):279–285, 2002.



## Apêndices

### A.1 Equações de Verossimilhança da WDE

$$\frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta})}{\partial q} = 0$$

$$+ \sum_{i=1}^n \left\{ \delta_i \left\{ \frac{\left[ \hat{\gamma} (1 - \hat{q}^{(t_i+1)^{\hat{\beta}}})^{\hat{\gamma}-1} (t_i + 1)^{\hat{\beta}} \hat{q}^{(t_i+1)^{\hat{\beta}}-1} \right] - \left[ \hat{\gamma} (1 - \hat{q}^{t_i^{\hat{\beta}}})^{\hat{\gamma}-1} t_i^{\hat{\beta}} \hat{q}^{t_i^{\hat{\beta}}-1} \right]}{\left[ 1 - \hat{q}^{(t_i+1)^{\hat{\beta}}} \right]^{\hat{\gamma}} - \left[ 1 - \hat{q}^{t_i^{\hat{\beta}}} \right]^{\hat{\gamma}}} \right\} \right\}$$

$$+ \sum_{i=1}^n \left\{ (1 - \delta_i) \left\{ \frac{\left[ -\hat{\gamma} (1 - \hat{q}^{(t_i+1)^{\hat{\beta}}})^{\hat{\gamma}-1} (t_i + 1)^{\hat{\beta}} \hat{q}^{(t_i+1)^{\hat{\beta}}-1} \right]}{1 - \left[ 1 - \hat{q}^{(t_i+1)^{\hat{\beta}}} \right]^{\hat{\gamma}}} \right\} \right\} = 0;$$

$$\frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta})}{\partial \beta} = 0$$

$$+ \sum_{i=1}^n \left\{ \delta_i \left\{ \frac{\left[ \hat{\gamma} (1 - \hat{q}^{(t_i+1)^{\hat{\beta}}})^{\hat{\gamma}-1} (t_i + 1)^{\hat{\beta}} \log(t_i + 1) \hat{q}^{(t_i+1)^{\hat{\beta}}} \log(\hat{q}) \right]}{\left[ 1 - \hat{q}^{(t_i+1)^{\hat{\beta}}} \right]^{\hat{\gamma}} - \left[ 1 - \hat{q}^{t_i^{\hat{\beta}}} \right]^{\hat{\gamma}}} \right\} \right\}$$

$$- \sum_{i=1}^n \left\{ \delta_i \left\{ \frac{\left[ \hat{\gamma} (1 - \hat{q}^{t_i^{\hat{\beta}}})^{\hat{\gamma}-1} t_i^{\hat{\beta}} \log(t_i) \hat{q}^{t_i^{\hat{\beta}}} \log(\hat{q}) \right]}{\left[ 1 - \hat{q}^{(t_i+1)^{\hat{\beta}}} \right]^{\hat{\gamma}} - \left[ 1 - \hat{q}^{t_i^{\hat{\beta}}} \right]^{\hat{\gamma}}} \right\} \right\}$$

$$+ \sum_{i=1}^n \left\{ (1 - \delta_i) \left\{ \frac{\left[ -\hat{\gamma} (1 - \hat{q}^{(t_i+1)^{\hat{\beta}}})^{\hat{\gamma}-1} (t_i + 1)^{\hat{\beta}} \log(t_i + 1) \hat{q}^{(t_i+1)^{\hat{\beta}}} \log(\hat{q}) \right]}{1 - \left[ 1 - \hat{q}^{(t_i+1)^{\hat{\beta}}} \right]^{\hat{\gamma}}} \right\} \right\} = 0$$

e

$$\frac{\partial \ell(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta})}{\partial \gamma} = 0.$$

$$\begin{aligned} & \sum_{i=1}^n \left\{ \delta_i \left[ \frac{\left(1 - \hat{q}^{(t_i+1)\hat{\beta}}\right)^{\hat{\gamma}} \log\left(1 - \hat{q}^{(t_i+1)\hat{\beta}}\right) \left(1 - \hat{q}^{t_i\hat{\beta}}\right)^{\hat{\gamma}} \log\left(1 - \hat{q}^{t_i\hat{\beta}}\right)}{\left(1 - \hat{q}^{(t_i+1)\hat{\beta}}\right)^{\hat{\gamma}} - \left(1 - \hat{q}^{t_i\hat{\beta}}\right)^{\hat{\gamma}}} \right] \right\} \\ + & \sum_{i=1}^n \left\{ (1 - \delta_i) \left[ \frac{-\left(1 - \hat{q}^{(t_i+1)\hat{\beta}}\right)^{\hat{\gamma}} \log\left(1 - \hat{q}^{(t_i+1)\hat{\beta}}\right)}{1 - \left(1 - \hat{q}^{(t_i+1)\hat{\beta}}\right)^{\hat{\gamma}}} \right] \right\} = 0 \end{aligned}$$

## A.2 Intervalos de Confiança para os Paramêtros

### Parâmetro $q$

Ao considerar a seguinte transformação

$$u = \log[-\log q] \quad \implies \quad q = e^{-e^u},$$

e que:

$$\hat{u} = \log[-\log \hat{q}] \quad \implies \quad \hat{q} = e^{-e^{\hat{u}}}.$$

Assim:

$$\begin{aligned} P \left( -Z_{1-\alpha/2} < \frac{\hat{u} - u}{\sqrt{\widehat{Var}(\hat{u})}} < Z_{1-\alpha/2} \right) &= 1 - \alpha \\ -Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} < \hat{u} - u < Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} \\ -\hat{u} - Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} < -u < -\hat{u} + Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} \\ \hat{u} + Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} > u > \hat{u} - Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} \\ \hat{u} - Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})} < u < \hat{u} + Z_{1-\alpha/2} \sqrt{\widehat{Var}(\hat{u})}. \end{aligned}$$

Aplicando  $e^{-e}$ , na desigualdade, fica:

$$e^{-e^{\hat{u}-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}}} < \underbrace{e^{-e^u}}_q < e^{-e^{\hat{u}+Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}}}.$$

Tem-se:

$$\begin{aligned} e^{-e^{\hat{u}+Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}}} &< q < e^{-e^{\hat{u}-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}}} \\ e^{-e^{\hat{u}}}e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} &< q < e^{-e^{\hat{u}}}e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} \\ \left[ e^{-e^{\hat{u}}} \right] e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} &< q < \left[ e^{-e^{\hat{u}}} \right] e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} \\ (\hat{q})e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} &< q < (\hat{q})e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}}. \end{aligned}$$

Logo, o intervalo de confiança para o parâmetro  $q$  é dado por:

$$\left[ (\hat{q})e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} ; (\hat{q})e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{u})}} \right].$$

## Parâmetro $\beta$

Ao considerar a seguinte transformação

$$v = \log \beta \quad \Longrightarrow \quad \beta = e^v,$$

e que:

$$\hat{v} = \log \hat{\beta} \quad \Longrightarrow \quad \hat{\beta} = e^{\hat{v}}.$$

Assim:

$$\begin{aligned} P \left( -Z_{1-\alpha/2} < \frac{\hat{v} - v}{\sqrt{\widehat{Var}(\hat{v})}} < Z_{1-\alpha/2} \right) &= 1 - \alpha \\ -Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} < \hat{v} - v < Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} \\ -\hat{v} - Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} < -v < -\hat{v} + Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} \\ \hat{v} + Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} > v > \hat{v} - Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} \\ \hat{v} - Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})} < v < \hat{v} + Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}. \end{aligned}$$

Aplicando  $e$ , na desigualdade, fica:

$$e^{\hat{v}-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} < \underbrace{e^{\hat{v}}}_{\beta} < e^{\hat{v}+Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}}.$$

Tem-se:

$$\begin{aligned} e^{\hat{v}-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} &< \beta < e^{\hat{v}+Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} \\ e^{\hat{v}} e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} &< \beta < e^{\hat{v}} e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} \\ [e^{\hat{v}}] e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} &< \beta < [e^{\hat{v}}] e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} \\ \hat{\beta} e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} &< \beta < \hat{\beta} e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}}. \end{aligned}$$

Logo, o intervalo de confiança para o parâmetro  $\beta$  é dado por:

$$\left[ \hat{\beta} e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}}; \hat{\beta} e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{v})}} \right].$$

### Parâmetro $\gamma$

Ao considerar a seguinte transformação

$$w = \log \gamma \quad \Longrightarrow \quad \gamma = e^w,$$

e que:

$$\hat{w} = \log \hat{\gamma} \quad \Longrightarrow \quad \hat{\gamma} = e^{\hat{w}}.$$

Assim:

$$P \left( -Z_{1-\alpha/2} < \frac{\hat{w} - w}{\sqrt{\widehat{Var}(\hat{w})}} < Z_{1-\alpha/2} \right) = 1 - \alpha$$

$$\begin{aligned} -Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} &< \hat{w} - w < Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} \\ -\hat{w} - Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} &< -w < -\hat{w} + Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} \\ \hat{w} + Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} &> w > \hat{w} - Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} \\ \hat{w} - Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})} &< w < \hat{w} + Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}. \end{aligned}$$

Aplicando  $e$ , na desigualdade, fica:

$$e^{\hat{w}-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} < \underbrace{e^{\hat{w}}}_{\gamma} < e^{\hat{w}+Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}}.$$

Tem-se:

$$\begin{aligned} e^{\hat{w}-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} &< \gamma < e^{\hat{w}+Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} \\ e^{\hat{w}}e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} &< \gamma < e^{\hat{w}}e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} \\ [e^{\hat{w}}]e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} &< \gamma < [e^{\hat{w}}]e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} \\ \hat{\gamma}e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} &< \gamma < \hat{\gamma}e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}}. \end{aligned}$$

Logo, o intervalo de confiança para o parâmetro  $\gamma$  é dado por:

$$\left[ \hat{\gamma}e^{-Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}}, \hat{\gamma}e^{Z_{1-\alpha/2}\sqrt{\widehat{Var}(\hat{w})}} \right].$$

### A.3 Logaritmo da Função de Verossimilhança Utilizada Computacionalmente

Ao considerar o logaritmo da função de verossimilhança (que permita a presença de dados cenurados) da distribuição WDE, tem-se:

$$\begin{aligned} \ell(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \delta_i \log \left\{ \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma - \left[ 1 - q^{t_i^\beta} \right]^\gamma \right\} \right\} + \\ &+ \sum_{i=1}^n \left\{ (1 - \delta_i) \log \left\{ 1 - \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma \right\} \right\}. \end{aligned} \quad (44)$$

Sejam  $d$  e  $e \in \mathbb{R}$ . Considere também a seguinte relação:

$$(d - e) = d \left( 1 - \frac{e}{d} \right) \implies \log(d - e) = \log d + \log \left( 1 - \frac{e}{d} \right), \quad (45)$$

Ao utilizar a relação (45) na primeira parte do logaritmo da função de verossimilhança (44) da distribuição WDE, tem-se:

$$\begin{aligned} \ell(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \delta_i \left[ \log \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma + \log \left[ 1 - \left( \frac{1 - q^{t_i^\beta}}{1 - q^{(t_i+1)^\beta}} \right)^\gamma \right] \right] \right\} + \\ &+ \sum_{i=1}^n \left\{ (1 - \delta_i) \log \left\{ 1 - \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma \right\} \right\}. \end{aligned}$$

Fica:

$$\begin{aligned} \ell(\mathbf{t}, \boldsymbol{\delta} | \boldsymbol{\theta}) &= \sum_{i=1}^n \left\{ \delta_i \left[ \gamma \log \left[ 1 - q^{(t_i+1)^\beta} \right] + \log \left[ 1 - \left( \frac{1 - q^{t_i^\beta}}{1 - q^{(t_i+1)^\beta}} \right)^\gamma \right] \right] \right\} + \\ &+ \sum_{i=1}^n \left\{ (1 - \delta_i) \log \left\{ 1 - \left[ 1 - q^{(t_i+1)^\beta} \right]^\gamma \right\} \right\}. \end{aligned} \quad (46)$$

Decidiu-se então, adotar a função *log1p* do *software R* que calcula  $\log(1 + x)$ ,  $x \in \mathbb{R}$ , com mais precisão. Para,  $\beta = \beta$ ,  $\gamma = \gamma$ ,  $\delta = \delta$ , o logaritmo da função de verossimilhança, fica assim representado no R:

```
vero<-function(theta,t,censura){
  q<-theta[1]
  beta<-theta[2]
  a<-theta[3]
  if ((q > 0) && (q < 1) && (beta > 0) && (a > 0))
  return (-1*sum(censura*((a*log1p(-q^((t+1)^beta)))
+(log1p(-((1-q^(t)^beta))/(1-q^((t+1)^beta)))^a)))+(1-censura)*
(log1p(-(1-q^((t+1)^beta) )^a)))/10^15)
  else return(-Inf)
}
```

Devido a problemas computacionais nas estimativas dos parâmetros, pelo fato do logaritmo da função de verossimilhança retornar valores “muito grandes” dependendo dos tempos observados ou da combinação dos parâmetros, decidiu-se dividir o logaritmo função de verossimilhança por  $10^{15}$ . Dessa forma,  $10^{15}$  funciona como uma constante normalizadora para o logaritmo da função de verossimilhança.

## A.4 Scripts Desenvolvidos

```
##### Geração dos tempos de sobrevivência da distribuição WDE
### q --- parâmetro do modelo
### beta --- parâmetro do modelo
### a --- parâmetro do modelo (gamma)
set.seed(2016)
redweibull<-function (n, q, beta,a, zero = FALSE){
  u <- runif(n)
  if (zero) {
s <- floor(((log(1-exp((log(u))/(a))))/(log(q)))^(1/beta)) - 1
  }
}
```

```

    else {
s <- floor(((log(1-exp((log(u))/(a))))/(log(q)))^(1/beta))
    }
s
}
n<-50, 80, 100, 200, 500 ## tamanho da amostra
t<-redweibull(n,q,beta,a, zero = FALSE)
p.censura<- 0.0, 0.1, 0.2, 0.3   ### percentual censura
censura<-rbinom(n,1,1 - p.censura)

##### Modelo WDE #####
##### Distribuição de probabilidades
DEW<-function(q,beta,a,t){
  (1-q^((t+1)^beta) )^a - (1-q^((t)^beta) )^a
}

##### Função de Sobrevivência
SDEW<-function(q,beta,a,t){
  1-(1-q^((t+1)^beta) )^a
}

##### Inferência Clássica #####
##### Função de log-verossimilhança
vero<-function(theta,t,censura){
  q<-theta[1]
  beta<-theta[2]
  a<-theta[3]
  if ((q > 0) && (q < 1) && (beta > 0) && (a > 0))
  return (-1*sum(censura*((a*log1p(-q^((t+1)^beta)))
+(log1p(-((1-q^((t)^beta))/(1-q^((t+1)^beta)))^a)))+(1-censura)*
(log1p(-((1-q^((t+1)^beta) )^a)))/10^15)
  else return(-Inf)
}

##### Vício e EQM
M<-10000   ## replicas de monte carlo

q.est<-beta.est<-a.est<-matrix(0,M,2)

```

```

for (i in 1:M){
  t<-redweibull(n,q,beta,a, zero = FALSE)
  censura<-rbinom(n,1,1-p.censura)
  v2 <- optim(c(q,beta,a),t=t,censura=censura,vero)
  q.est[i,1]<-v2$par[1]
  q.est[i,2]<-v2$convergence

  beta.est[i,1]<-v2$par[2]
  beta.est[i,2]<-v2$convergence

  a.est[i,1]<-v2$par[3]
  a.est[i,2]<-v2$convergence

  cat(i,"\n")
}

vicio.q<-mean(q.est[,1]-q)    ### vício
EQM.q<-mean((q.est[,1]-q)^2)  ##EQM

vicio.beta<-mean(beta.est[,1]-beta)    ### vício
EQM.beta<-mean((beta.est[,1]-beta)^2)  ##EQM

vicio.a<-mean(a.est[,1]-a)    ### vício
EQM.a<-mean((a.est[,1]-a)^2)  ##EQM

##### Estimativas pontuais (Estimativas de máx. vero.)
v2 <- optim(c(q,.beta,a),t=t,censura=censura,vero=NULL,hessian=T)
q.est<- v2$par[1]            ## estimativa do parâmetro q
beta.est<- v2$par[1]         ## estimativa do parâmetro beta
a.est<- v2$par[1]           ## estimativa do parâmetro gamma
v2$hessian                   ## hessiana

##### Intervalo de confiança para os parâmetros
require(emdbook)
#### parâmetro q
var.u<-deltavar(fun = log(-log(q)),Sigma=(10^-15)*solve(v2$hessian[1,1]),
meanval=c(q = q.est))
LI.q<-(v2$par[1])^(exp(qnorm(0.975)*sqrt(var.u)))  ##limite inferior

```

```

LS.q<-(v2$par[1])^(exp(-qnorm(0.975)*sqrt(var.u))) ##limite superior

#### parâmetro beta
var.v<-deltavar(fun = log(beta),Sigma=(10^-15)*solve(v2$hessian[2,2]),
meanval=c(beta = beta.est))
LI.beta<-(v2$par[2])*exp((-qnorm(0.975)*sqrt(var.v))) ##limite inferior
LS.beta<-(v2$par[2])*exp((qnorm(0.975)*sqrt(var.v))) ##limite superior

#### parâmetro gamma
var.w<-deltavar(fun = log(a),Sigma=(10^-15)*solve(v2$hessian[3,3]),
meanval=c(a = a.est))
LI.a<-(v2$par[3])*exp((-qnorm(0.975)*sqrt(var.w))) ##limite inferior
LS.a<-(v2$par[3])*exp((qnorm(0.975)*sqrt(var.w))) ##limite superior

##### TRV
##### H01: gamma=1
##### Função de Verossimilhança WD
vero1<-function(theta,t,censura){
  q<-theta[1]
  beta<-theta[2]
  a<-1
  if ((q > 0) && (q < 1)&& (beta > 0))
  return (-1*sum(censura*((a*log1p(-q^((t+1)^beta)))
+(log1p(-((1-q^((t)^beta))/(1-q^((t+1)^beta))))^(a))))+(1-censura)*
(log1p(-((1-q^((t+1)^beta) )^a)))/10^15)
  else return(-Inf)
}

v3 <- optim(c(q.est,beta.est),t=t,censura=censura,vero1,NULL,hessian=T)
lwde01<-v3$value*10^15
lwde<-v2$value*10^15 ## modelo completo

TRV01<- -2*(lwde-lwde01) ## estatística do teste
1-pchisq(TRV01,1) ## valor-p

##### H02: beta=1
##### Função de Verossimilhança GE
vero2<-function(theta,t,censura){
  q<-theta[1]

```

```

    beta<-1
    a<-theta[2]
    if ((q > 0) && (q < 1)&& (a > 0)&& (a < 10))
    return (-1*sum(censura*((a*log1p(-q^((t+1)^beta)))
    +(log1p(-((1-q^((t)^beta))/(1-q^((t+1)^beta)))^a))))+(1-censura)*
    (log1p(-((1-q^((t+1)^beta) )^a)))/10^15)
    else return(-Inf)
}

v4 <- optim(c(v3$par[1],v3$par[2]),t=t,censura=censura,vero2,NULL,hessian=T)
lwde02<-v4$value*10^15
lwde<-v2$value*10^15      ##modelo completo

TRV02<- -2*(lwde-lwde02)    ## estatística do teste
1-pchisq(TRV02,2)          ## valor-p

##### H03: gamma=1; beta=1
##### Função de Verossimilhança da distribuição Geométrica
vero3<-function(theta,t,censura){
  q<-theta[1]
  beta<-1
  a<-1
  if ((q > 0) && (q < 1))
  return (-1*sum(censura*((a*log1p(-q^((t+1)^beta)))
  +(log1p(-((1-q^((t)^beta))/(1-q^((t+1)^beta)))^a))))+(1-censura)*
  (log1p(-((1-q^((t+1)^beta) )^a)))/10^15)
  else return(-Inf)
}

v5 <- optim(c(v3$par[1]),t=t,censura=censura,vero3,NULL,hessian=T)
lwde03<-v5$value*10^15
lwde<-v2$value*10^15      ##modelo completo

TRV03<- -2*(lwde-lwde03)    ## estatística do teste
1-pchisq(TRV03,2)          ## valor-p

##### Ajuste dos modelos WDE,WD, GE e Gemométrico aos dados
require(survival)

```

```

km<-survfit(Surv(t,censura)~1)
plot(km,conf.int=F,xlab="t",ylab="S(t)",ylim=c(0,1)) # K-M

tt<-seq(0,max(t))
s.est<-SDEW(v2$par[1],v2$par[2],v2$par[3],tt)
points(tt,s.est,type="s",col=2) # WDE

s.est1<-SDEW(v3$par[1],v3$par[2],1,tt)
points(tt,s.est1,type="s",col=3) # WD

s.est2<-SDEW(v4$par[1],1,v4$par[2],tt)
points(tt,s.est2,type="s",col=5) # GE

s.est3<-SDEW(v5$par[1],1,1,tt)
points(tt,s.est3,type="s",col=4) # Geométrico

##### Ajuste do modelo WDE aos dados com IC 95%
##### para a função de sobrevivência do modelo
km<-survfit(Surv(t,censura)~1)
plot(km,conf.int=F,xlab="t",ylab="S(t)",ylim=c(0,1))

tt<-seq(0,max(t))
s.est1<-SDEW(v2$par[1],v2$par[2],v2$par[3],tt)
points(tt,s.est1,type="s",col=2)

##### IC 95% da função se sob.
t<-(0:1000)
inv<-solve(v2$hessian)
var<-diag(inv)
Sigma=inv
s.weibull<-LI.sob<-SDEW(v2$par[1],v2$par[2],v2$par[3],t=t)
Sigma=inv
require(emdbook)
##### calculo da variância de S(t)
var.sob<-numeric(length(t))
for (i in 1:length(t)) {
  x<-t[i]
  var.sob[i]<-deltavar(fun = 1-(1-q^((x+1)^beta) )^a ,Sigma=(10^-15)*Sigma,
meanval=c(q = v2$par[1],beta = v2$par[2],a = v2$par[3]))}

```

```

#### Construindo IC 95% da função se sob.
LI.sob<-s.weibull - qnorm(0.975)*var.sob^.5
LS.sob<-s.weibull + qnorm(0.975)*var.sob^.5

### Intervalo simétrico (truncando em 0 e 1)
LI.sob[which(LI.sob<0)]<-0
LS.sob[which(LS.sob>1)]<-1

points(t,LI.sob,type="s",lty=2,col=6)
points(t,LS.sob,type="s",lty=2,col=6)

legend('topright', xpd=TRUE,c("K-M","WDE",
"IC 95%"),bty="n",cex=1,col=c(1,2,6),lty=c(1,1,2))

##### Inferência Bayesiana #####
#### logaritmo da distribuição a posteriori
log.post<- function(x,censura){
  q<-x[1]
  beta<-x[2]
  a<-x[3]
  if ((q > 0) && (q < 1) && (beta > 0) && (a > 0 ))
  return(((10^-3)-1)*log(beta)-(10^-3)*beta + ((10^-3)-1)*log(a)-(10^-3)*a +
  sum(censura*(log(DEW(q,beta,a,t)))+(1-censura)*(log(SDEW(q,beta,a,t)))) )
  else return( -Inf)}

##### Estimativas pontuais (médias a posteriori)
require(MCMCpack)
theta<- MCMCmetrop1R(log.post,c(q,.beta,a),burnin=1000,mcmc=10000,
logfun = TRUE,censura=censura)
q.est<-mean(theta[,1])          ## estimativa do parâmetro q
beta.est<-mean(theta[,2])      ## estimativa do parâmetro beta
a.est<-mean(theta[,3])         ## estimativa do parâmetro gamma

##### Intervalo HPD para os parâmetros
library("TeachingDemos")
emp.hpd(theta[,1])
emp.hpd(theta[,2])
emp.hpd(theta[,3])

```

```
##### Ajuste do modelo WDE aos dados com intervalo HPD de 95%
#### para a função de sobrevivência do modelo
km<-survfit(Surv(t,censura)~1)
plot(km,conf.int=F,xlab="t",ylab="S(t)",ylim=c(0,1)) # K-M

tempo<-(0:1000)
est.theta<-numeric(length(tempo))
LI<-numeric()
LS<-numeric()
for (i in 0:length(tempo)){
x<-tempo[i]
ec= SDEW(theta[,1],theta[,2],theta[,3],x)
est.theta[i]<-mean(ec)
LI[i]<-emp.hpd(ec)[1]
LS[i]<-emp.hpd(ec)[2]}
points(tempo, est.theta, type="s",
xlab="t",ylab="S(t)",col=2) #modelo WDE
points(tempo,LI,lty=3,type="s",col=3)
points(tempo,LS,lty=3,type="s",col=3)
legend('topright', xpd=TRUE,c("Kaplan-Meier","WDE",
"Intervalo HPD 95%"),bty="n",cex=1,col=c(1,2,3),lty=c(1,1,3))

##### FBST
l.post<-numeric()
for (i in 1:length(theta[,1])){
  l.post[i]<-log.post(theta[i,],censura)}

##### H01: gamma=1
### max do log.post sob H01
log.post.H01<- function(x,censura){
q<-x[1]
beta<-x[2]
a<-1
if ((q > 0) && (q < 1) && (beta > 0) )
return(-1*(((10^-3)-1)*log(beta)-(10^-3)*beta +
((10^-3)-1)*log(a)-(10^-3)*a +
sum(censura*(log(DEW(q,beta,a,t)))+(1-censura)*(log(SDEW(q,beta,a,t)))))) )
else return( -Inf)}
```

```

aa<-optim(c(q.est,beta.est),log.post.H01,censura=censura)
max.l.post<- -1*aa$value      ### maximo da posteriori sob H01: gamma=1

evalor<- sum(l.post<max.l.post)/length(theta)  ### valor-e

##### H02: beta=1
### max do log.post sob H02
log.post.H02<- function(x,censura){
  q<-x[1]
  beta<-1
  a<-x[2]
  if ((q > 0) && (q < 1) && (a > 0) )
  return(-1*(((10^-3)-1)*log(beta)-(10^-3)*beta +
  ((10^-3)-1)*log(a)-(10^-3)*a +
  sum(censura*(log(DEW(q,beta,a,t)))+(1-censura)*(log(SDEW(q,beta,a,t)))))) )
  else return( -Inf)}
aa1<-optim(c(aa$par[1],aa$par[2]),log.post.H02,censura=censura)
max.l.post<- -1*aa1$value    ### maximo da posteriori sob H02: beta=1

evalor<- sum(l.post<max.l.post)/length(theta)  ### valor-e

H03: gamma=1; beta=1
### max do log.post sob H03
log.post.H03<- function(x,censura){
  q<-x[1]
  beta<-1
  a<-1
  if ((q > 0) && (q < 1) )
  return(-1*(((10^-3)-1)*log(beta)-(10^-3)*beta +
  ((10^-3)-1)*log(a)-(10^-3)*a +
  sum(censura*(log(DEW(q,beta,a,t)))+(1-censura)*(log(SDEW(q,beta,a,t)))))) )
  else return( -Inf)}
aa2<-optim(aa$par[1],log.post.H03,censura=censura)
max.l.post<- -1*aa2$value    ### maximo da posteriori sob H03: gamma=1 e beta =1

evalor<- sum(l.post<max.l.post)/length(theta)  ### valor-e

```