Universidade Estadual de Maringá
Programa de Pós-graduação em Bioestatística

MARCOS VINICIUS DE OLIVEIRA PERES

# Aplicações das Distribuições Weibull Modificada e Beta-Weibull na Presença de Frações de Cura sob o Enfoque Frequentista e Bayesiano.

Dissertação de Mestrado

Maringá - Paraná
2016

MARCOS VINICIUS DE OLIVEIRA PERES

# Aplicações das Distribuições Weibull Modificada e Beta-Weibull na Presença de Frações de Cura sob o Enfoque Frequentista e Bayesiano.

Dissertação apresentada ao Programa de Pós-Graduação em Bioestatística do Centro de Ciências Exatas da Universidade Estadual de Maringá como requisito parcial para obtenção do título de Mestre em Bioestatística.
Orientador: Prof Dr. Edson Z. Martinez
Coorientadora: Prof[a] Dra. Isolde Previdelli

Maringá - Paraná
2016

Dados Internacionais de Catalogação-na-Publicação (CIP)
(Biblioteca Central - UEM, Maringá – PR., Brasil)

P437a

Peres, Marcos Vinicius de Oliveira
Aplicações das distribuições Weibull modificada e Beta-Weibull na presença de frações de cura sob o enfoque Frequentista e Bayesiano / Marcos Vinicius de Oliveira Peres. -- Maringá, 2016.
93 f. : il. col., figs., tabs., + anexo e apêndice

Orientador: Prof. Dr. Edson Zangiacomi Martinez.
Co-orientadora: Prof.a Dr.a Isolde Terezinha Santos Previdelli.
Dissertação (mestrado) - Universidade Estadual de Maringá, Programa de Pós-Graduação em Bioestatística, 2016

1. Métodos quantitativos aplicados à saúde - Weibull modificada - Frequentista e Bayesiano - Comparação. 2. Métodos quantitativos aplicados à saúde - Beta-Weibull - Frequentista e Bayesiano - Comparação. 3. Análise de sobrevivência (Bioestatística) - Função de risco. 4. Intervalo de confiança perfilado. I. Martinez, Edson Zangiacomi, orient. II. Previdelli, Isolde Terezinha Santos, co-orient. III. Universidade Estadual de Maringá. Programa de Pós-Graduação em Bioestatística. IV. Título.

CDD 21.ed. 570.15195
MN-003314

*Dedico este trabalho a todos que me apoiaram nessa minha jornada.*

# Agradecimentos

Agradeço sobretudo a Deus, pois sem sua proteção não teria alcançado o fim dessa jornada.

A minha família, em especial a minha mãe e ao meu pai por sempre terem me incentivado a estudar em busca de uma vida melhor. Por sempre terem me apoiado e me ajudado nos momentos difíceis. Por terem batalhado durante suas vidas para que eu pudesse seguir o melhor caminho e por terem me ensinado valores que me fazem o que sou hoje.

Agradeço fervorosamente a meu orientador Edson Martinez por ter me acolhido como orientando, por seus conselhos, por compartilhar seus conhecimentos, por sua disposição e principalmente por sua paciência. Pelo valor e pela valorização de todas as atividades e trabalhos desenvolvidos. Sobretudo agradeço pelo grande exemplo de pessoa que é.

A professora Isolde Previdelli que desde do primeiro contato mostrou-se uma pessoa querida e dedicada. Por seus conselhos, broncas, pela paciência e por compartilhar sua experiência no momento do estágio.

Tenho muito a agradecer aos professores do departamento. A todos, mas principalmente aqueles que ministraram aulas por min assistidas e aos que me atenderam nos momentos de dúvidas. As professoras, Eniuce, Rosângela, Teresinha e Clédina, e aos professores Vanderly, Robson, Josmar, Eraldo e Diogo. Agradeço ainda a todos funcionários do departamento, aos professores, seguranças e aos zeladores.

Por todos momentos que passamos, não posso esquecer dos meus grandes amigos que conheci nesse mestrado e que fazem parte agora da minha vida. Agradeço muito a Rafaela, Guilherme, Marina, Matheus, Kelly, Omar, Emerson, Sérgio, Oilson, Marcos Jardel, Paulo, Ricardo, José André, Thiago, Márcia, Diego, e a todos que estiveram presente nessa jornada.

*"É melhor ter uma resposta aproximada para uma pergunta certa do que uma resposta exata para uma pergunta errada."*
*John Tukey*

# Resumo

As pesquisas médicas são de grande importância para o seguimento de pacientes com câncer especialmente após intervenções, como cirurgias de resseção, transplantes e quimio-oradioterapia, com o objetivo de compreender melhor o tratamento e melhorar a qualidade de vida desses indivíduos. É de grande importância a utilização de métodos adequados na modelagem desses dados. Dentre as ferramentas metodológicas existentes, uma de grande importância é a análise de sobrevivência. No contexto de análise de sobrevivência, o evento de interesse é muitas vezes relacionado com a morte ou recorrência de uma doença. Entretanto, ao final do estudo, é possível que uma parte da amostra não sofra o evento de interesse. Esses indivíduos podem ter sido curados ou são imunes ao evento de interesse, sendo dessa forma de grande importância estimar de forma adequada essa proporção de indivíduos não suscetíveis. Os modelos tradicionais, já muito conhecidos em análise de sobrevivência, em geral não são adequados para estimar essa proporção de imunes, sendo necessário modelos estatísticos mais complexos que incorporem esta informação. Atualmente existem várias técnicas para se estimar esta proporção de imunes, como os modelos de mistura com fração de cura e os modelos de não mistura. Foi utilizada uma análise baseada em duas distribuições não muito conhecidas no contexto prático, a distribuição Weibull modificada, uma distribuição de três parâmetros, e a distribuição beta-Weibull, com quatro parâmetros. Foram consideradas em ambas as distribuições a presença da fração de cura, dados censurados e covariáveis. Estimativas frequentistas (por máxima verossimilhança) e estimativas por inferência Bayesiana foram comparadas. Para verificar a adequação destes modelos na análise de dados reais, a Weibull modificada foi aplicada a dados de pacientes com adenocarcinoma gástrico, e para a beta-Weibull a dados de transplante de medula óssea. Ambos modelos considerados se adequaram de forma satisfatória aos dados e estimaram de forma adequada a proporção de cura. As estimativas Bayesianas e seus respectivos intervalos de alta densidade *a posteriori* (HPD) foram mais parcimoniosos do que os obtidos com o método de máxima verossimilhança.

**Palavras-chaves**: Análise de sobrevivência, câncer, intervalo de confiança perfilado, função de risco.

# Abstract

The medical research is of great importance for the monitoring of patients who have cancer - especially after medical interventions, such as surgical resections, organ transplants and chemoradiotherapy – aiming a better understanding in the treatment and a life quality improvement for the subjects. It is extremely important to use proper methods for the data modelling. Among the available tools, a very important one is the survival analysis. In the survival analysis context, the event of interest is often related to death or disease recurrence. However, in the study's conclusion it is possible for a part of the sample not to suffer from the event of interest. These patients can have been cured or be immune to the event of interest. Therefore, it is quite important to estimate in a proper way the proportion of not susceptible patients. The traditional models, which are well known in the survival analysis, usually are not adequate to estimate the immune proportion. It is necessary the use of complex statistic models to incorporate this information. Currently there are several methods to estimate the immune proportion, such as the mixing models with cure fraction and the not mixing models. In the present paper it has been used an analysis based on two not common distributions in the practical context, the Weibull modified distribution, a distribution composed by three parameters. Likewise, the beta-Weibull distribution, with four parameters. For both distributions, it was considered the cure fraction existence, censored data and the covariant. Frequentist estimates (by maximum likelihood) and Bayesian inference estimates were compared. To verify the adequacy of these models in the real data analysis the Weibull modified analysis was applied to the data of patients who have gastric adenocarcinoma. And the beta-Weibull analysis for the data of bone marrow transplant. Both considered models have adjusted in a satisfactory way to the data and properly estimated the cure proportion. The Bayesian estimates and their respective High Posterior Density intervals (HPD) were more parsimonious than the ones resulted from the method of maximum likelihood.

**Keywords**: Survival analysis, cancer, profile likelihood confidence interval, hazard function.

# Lista de ilustrações


Figura 1 – Representações de alguns TTT-plots . . . . . 23
Figura 2 – Intervalo de confiança perfilado para um parâmetro $\theta$. . . . . 28
Figura 3 – Gráficos da função densidade de probabilidade (a), da função de sobrevivência (b) e da função de risco (c) da distribuição Weibull modificada. . . 39
Figura 4 – Estimadores Kaplan-Meier para as função de sobrevivência dos dados de câncer gástrico dos pacientes em geral e para cada tratamento. . . . 49
Figura 5 – Gráfico TTT-plot dos dados de câncer gástrico dos pacientes em geral. . 51
Figura 6 – Gráficos dos estimadores de Kaplan–Meier para as funções de sobrevivência versus os respectivos valores preditos, das estimativas bayesianas do modelo com mistura, considerando a distribuição Weibull modificada e seus casos particulares. . . . . 54
Figura 7 – Funções de sobrevivência obtidas com estimativas Bayesianas do modelo de mistura, para WM e suas distribuições particulares. . . . . 55
Figura 8 – Funções de risco obtidas com estimativas Bayesianas do modelo de mistura. . . . . 56
Figura 9 – Gráfico TTT-plot dos dados de câncer gástrico dos pacientes para cada um dos tratamentos. . . . . 57
Figura 10 – Funções de sobrevivência obtidas com o modelo de mistura e não mistura, para a Weibull modificada na presença de covariável. . . . . 58
Figura 11 – Gráficos da função densidade de probabilidade (a), da função de sobrevivência (b) e da função de risco (c) da distribuição Beta-Weibull. . . . . 63
Figura 12 – Estimadores Kaplan-Meier para as função de sobrevivência dos dados de transplante de medula óssea dos pacientes em geral e por grupo. . . 68
Figura 13 – Funções de sobrevivência obtidas pelas estimativas Bayesianas (em vermelho) e por meio das estimativas frequentistas (em azul) do modelo de mistura, para BW, cada distribuição particular dela. . . . . 72

Figura 14 – Gráficos dos estimadores de Kaplan–Meier para as funções de sobrevivência versus os respectivos valores preditos, das estimativas Bayesianas (em vermelho) e frequentistas (em azul) do modelo com mistura, considerando a distribuição Beta-Weibull e seus casos particulares. . . . . 75
Figura 15 – Curvas ajustadas do modelo de mistura baseado na distribuição Beta-Weibull considerando os dados de transplantes de medula por grupo. . . 77

# Lista de tabelas


Tabela 1 – Estimativas frequentistas dos parâmetros por máxima verossimilhança assumindo um modelo de mistura, sem covariáveis. . . . . 51
Tabela 2 – Estimativas Bayesianas dos parâmetros assumindo um modelo de mistura, sem covariáveis. . . . . 52
Tabela 3 – Estimativas frequentistas dos parâmetros por máxima verossimilhança assumindo um modelo de não mistura, sem covariável. . . . . 53
Tabela 4 – Estimativas Bayesianas dos parâmetros assumindo um modelo de não mistura, sem covariável. . . . . 53
Tabela 5 – Estimativas frequentistas de máxima verossimilhança dos parâmetros assumindo os modelos de mistura e não mistura, baseados na distribuição WM, com a covariável incluída no proporção de cura $p$ e no parâmetro de forma $\beta$. . . . . 57
Tabela 6 – Estimativas Bayesianas dos parâmetros assumindo os modelos de mistura e não mistura, baseados na distribuição WM, com a covariável incluída no proporção de cura $p$ e no parâmetro de forma $\beta$. . . . . 58
Tabela 7 – Estimativas Bayesianas dos parâmetros, assumindo um modelo de mistura considerando os dados de transplante de medula óssea. . . . . 70
Tabela 8 – Estimativas frequentativas dos parâmetros, assumindo um modelo de mistura considerando os dados de transplante de medula óssea. . . . . 71
Tabela 9 – Estimativas Bayesianas dos parâmetros, assumindo um modelo de não mistura considerando os dados de transplante de medula óssea. . . . . 73
Tabela 10 – Estimativas frequentativas dos parâmetros, assumindo um modelo de não mistura considerando os dados de transplante de medula óssea. . 74
Tabela 11 – Estimativas Bayesianas dos parâmetros, assumindo um modelo de mistura com covariáveis, considerando os dados de transplante de medula óssea. . . . . 76
Tabela 12 – Dados Câncer Gástrico . . . . . 88

# Sumário


**1 Introdução . . . . . 13**
**2 Revisão Bibliográfica . . . . . 17**
2.1 Principais conceitos em Análise de Sobrevivência . . . . . 17
2.1.1 Tempo de Falha . . . . . 18
2.1.2 Censura . . . . . 18
2.1.3 Representação dos Dados de Sobrevivência . . . . . 19
2.1.4 Principais Funções em Análise de Sobrevivência . . . . . 20
2.1.4.1 Função de Sobrevivência . . . . . 20
2.1.4.2 Função de Risco . . . . . 21
2.1.4.3 Tempo Médio e Variância do Tempo Médio . . . . . 21
2.1.5 Estimador de Kaplan-Meier . . . . . 21
2.2 Gráfico TTT-plot . . . . . 22
2.3 Fração de Cura . . . . . 24
2.3.1 Modelos de fração de cura com mistura e não mistura . . . . . 24
2.4 Método da Máxima Verossimilhança . . . . . 25
2.4.1 Intervalos de Confiança . . . . . 26
2.4.1.1 Intervalo Assintótico . . . . . 26
2.4.1.2 Intervalo Perfilado . . . . . 27
2.4.2 Medida de Assimetria para Intervalo de Confiança Perfilado . . . . . 28
2.5 Funções de Ligação . . . . . 29
2.6 Método Delta . . . . . 30
2.7 Introdução à Inferência Bayesiana . . . . . 31
2.7.1 Teorema de Bayes . . . . . 31
2.7.2 Atualização da Incerteza . . . . . 32
2.7.3 Intervalo de Credibilidade e Intervalo de Alta Densidade *a Posteriori* 34
2.8 Critérios de Seleção . . . . . 34
**3 A Distribuição Weibull Modificada . . . . . 37**
3.1 Média e Variância . . . . . 38
3.2 Verossimilhança da Weibull Modificada . . . . . 40
3.2.1 Verossimilhança Weibull modificada na presença de dados censurados 42
3.2.2 Verossimilhança Weibull modificada na presença de dados censurados e fração de cura em um modelo de mistura . . . . . 43

3.2.3 Verossimilhança Weibull modificada na presença de dados censurados e fração de cura em um modelo de não-mistura . . . . . 45
3.3 Regressão com Ligação log-log Complementar no Parâmetro $p$ . . . . . 46
**4 Aplicação Weibull Modificada . . . . . 48**
4.1 Os Dados . . . . . 48
4.2 Métodos . . . . . 49
4.3 Resultados . . . . . 50
4.4 Discussão . . . . . 59
**5 A Distribuição Beta-Weibull . . . . . 61**
5.1 Média e Variância . . . . . 64
5.2 Função de Verossimilhança . . . . . 64
**6 Aplicações Beta-Weibull . . . . . 67**
6.1 Os Dados . . . . . 67
6.2 Métodos . . . . . 67
6.3 Resultados . . . . . 69
6.4 Discussão . . . . . 77
**7 Considerações Finais . . . . . 79**
**8 Estudos Futuros . . . . . 80**
**Referências . . . . . 81**
**Apêndice A Dados Câncer Gástrico . . . . . 88**
**Apêndice B Códigos Inferência Bayesiana . . . . . 89**
**Apêndice C Códigos Inferência Frequentista . . . . . 91**

# Capítulo 1

# Introdução

A análise de sobrevivência é um conjunto de técnicas e modelos estatísticos frequentemente utilizados nas áreas da saúde e nas engenharias (nesse caso, mais conhecida como análise de confiabilidade). A variável aleatória de interesse da análise de sobrevivência é o tempo até a ocorrência de algum evento definido, como por exemplo, o tempo até a ocorrência de um doença, o tempo até a morte do paciente, o tempo até a quebra de um aparelho eletrônico, etc.

Em análise de sobrevivência existem muitas técnicas consideradas padrões, incluindo o método Kaplan-Meier, o teste log-rank e o modelo de riscos proporcionais de Cox (LEE; GO, 1997; BRADBURN et al., 2003). Os modelos paramétricos de sobrevivência já são usados há décadas, motivados pela necessidade de se utilizar ferramentas mais refinadas para descrever estruturas de dados mais complexas, como a presença da fração de cura (LAMBERT et al., 2007). Uma vantagem dos modelos paramétricos em relação ao modelo de riscos proporcionais de Cox, é que são mais informativos. E ainda, o modelo de Cox assume que a razão entre as funções de risco entre dois diferentes grupos de covariável é constante e este fato pode não ocorrer em dados reais. Assim, modelos paramétricos podem ser mais flexíveis, principalmente quando não houver a proporcionalidade de riscos entre os grupos.

Uma importante característica do modelos paramétricos é que a função de sobrevivência assumida segue uma distribuição de probabilidade conhecida. Na década de 1950 surgiu a distribuição de probabilidade mais popular na análise de sobrevivência, a distribuição Weibull de dois parâmetros. Esta é aplicada em diversos campos, como engenharia de confiabilidade e prognóstico médico. De acordo com Lai (2014), a distribuição Weibull foi identificada primeiramente por Fréchet (1927) e usada posteriormente por Rosin e Rammler (1933) para descrever a distribuição do tamanho de partículas. Essa distribuição é assim chamada porque em 1951, o engenheiro e matemático sueco Ernst Hjalmar Waloddi Weibull

(1887-1979) (WEIBULL, 1951) a descreveu em detalhes.

Subsequentemente, Menon (1963) estudou o problema de se estimar os parâmetros de forma e escala da distribuição Weibull e Cohen (1965) utilizou o método de máxima verossimilhança para estimar os parâmetros considerando dados completos e censurados. Outros resultados foram obtidos por Thoman et al. (1969), tais como intervalos de confiança exatos e testes de hipóteses sobre os parâmetros e o poder do teste em relação ao parâmetro de forma. No campo da medicina, a distribuição Weibull é amplamente utilizada em pesquisas de câncer, devido à flexibilidade de sua função de risco e a facilidade para estimar os parâmetros (PETO et al., 1972). Em geral, a distribuição Weibull é adequada para modelar dados onde o risco é constante, crescente ou decrescente. Contudo, a distribuição Weibull não é apropriada para dados que apresentam função risco não monótona ou em forma de banheira.

Recentemente, tem-se intensificado o desenvolvimento de novas de distribuições mais flexíveis e um grande número de autores propuseram extensões para a distribuição Weibull de dois parâmetros tradicional. Isto deve-se à necessidade de encontrar distribuições que se ajustem de forma mais satisfatória aos conjuntos de dados reais.

Smith e Naylor (1987), por exemplo, descreveram uma distribuição Weibull de três parâmetros e desenvolveram estimadores de máxima verossimilhança e bayesianos para esta distribuição.

Mudholkar e Srivastava (1993) apresentaram uma nova forma para à distribuição Weibull, chamando esta nova distribuição de Weibull exponenciada. Este novo modelo inclui a possibilidade de se obter uma função de risco unimodal e de banheira, como também fornece uma classe maior para as funções de risco monótonas.

Mudholkar et al. (1996) adicionaram um parâmetro a distribuição Weibull, permitindo a ela um maior número de formas para a função de risco, como a unimodal e de banheira, e ampliando a classe de funções de forma monótona, chamando esta distribuição de Weibull generalizada.

Lai et al. (2003) proposeram outra modificação na distribuição Weibull, com a introdução de mais um parâmetro, a fim de acomodar não somente as formas monótonas, como também a forma de banheira para a função de risco. Esta distribuição é denominada distribuição Weibull modificada.

Uma generalização de quatro parâmetros da distribuição Weibull, também capaz de modelar uma função de taxa de risco em forma de banheira, foi proposta por Carrasco et al. (2008).

Ainda ampliando o número de extensões da distribuição Weibull, Famoye et al. (2005) introduziram um nova distribuição de quatro parâmetros, a beta-Weibull, que possui forma de banheira para a função de risco. E ainda mais recente, outras extensões da

distribuição Weibull foram introduzidas na literatura, como a beta Weibull modificada (SILVA et al., 2010) e a beta Weibull exponenciada (CORDEIRO et al., 2013a).

De acordo com Hjorth (1980), distribuições com um ou dois parâmetros apenas, como é o caso da distribuição Weibull, possuem limitações importantes como a impossibilidade de modelar dados que apresentam função de risco em forma de banheira. Contudo, as distribuições mais flexíveis e com maiores números de parâmetros, podem ter estimativas imprecisas, quando há um tamanho pequeno de amostra.

Por envolver muitos parâmetros, modelos de sobrevivência complexos em geral apresentam instabilidades numéricas na aplicação do método de estimação de máxima verossimilhança, principalmente quando o tamanho amostral é pequeno, dessa forma a abordagem Bayesiana para a estimação dos parâmetros pode ser mais adequada (BOLFARINE et al., 1991).

O desempenho dos estimadores de máxima verossimilhança e Bayesianos para a distribuição Weibull foi estudado por Canavos e Taokas (1973). Estes autores verificaram que o erro quadrático médio das estimativas Bayesianas foi consideravelmente menor do que o encontrado pelo método de máxima verossimilhança.

Segundo CANCHO et al. (2007), quando a amostra é relativamente pequena, o uso do métodos de máxima verossimilhança para alguns modelos pode apresentar resultados imprecisos.

Os parâmetros da distribuição Weibull modificada para dados sem censura foram estimados via método Bayesiano por Jiang et al. (2008), que concluíram que a construção de intervalo de credibilidade para esta distribuição é mais conveniente que quando comparado aos métodos de máxima verossimilhança.

Uma situação que podemos encontrar ao estudar dados médicos em análise de sobrevivência, em particular na investigação do câncer, ocorre quando espera-se que uma fração dos indivíduos não irá experimentar o evento de interesse. Esta fração é relacionada a uma "fração de cura", uma vez que é geralmente atribuída aos doentes curados da doença em questão. A presença da fração de cura em um conjunto de dados é geralmente sugerido por um gráfico de Kaplan-Meier da função de sobrevivência, que mostra um platô longo e estável com forte presença de dados censurados na extrema direita do gráfico (CORBIERE et al., 2009).

Modelos paramétricos são geralmente descritos por uma função de sobrevivência $S(t)$, de tal forma que o limite de $S(t)$ com $t$ tendendo ao infinito é zero. No entanto, isso não ocorre na presença de uma fração de cura, uma vez que a função de sobrevivência apresenta uma assíntota no valor correspondente à fração de pessoas "curadas". Para esses casos, há uma série de métodos estatísticos mais adequados, que consideram a população estudada como uma mistura de indivíduos suscetíveis que experimentam o evento de

interesse e indivíduos não-suscetíveis que supostamente nunca vão experimentá-lo. Estes métodos são descrito por diversos autores, como Boag (1949), Maller e Zhou (1996), Ng e McLachlan (1998), Angelis et al. (1999), Peng e Dear (2000),Tsodikov et al. (2003), Lambert et al. (2007) e muitos outros. Modelos Bayesianos para dados de sobrevivência com uma fração de cura são descritos, por exemplo, por Ibrahim et al. (2001), Chen et al. (2002), Achcar et al. (2012) e Martinez et al. (2013).

O objetivo do presente trabalho é apresentar uma comparação entre as estimativas de máxima verossimilhança e estimativas bayesianas para a distribuição Weibull modificada de três parâmetros, descrita por Lai et al. (2003) e para a distribuição beta-Weibull introduzida por Lambert et al. (2007), na presença de fração de cura, em dados censurados e com covariáveis.

O trabalho está organizado da seguinte ordem: no Capítulo 2, é apresentada uma breve revisão bibliográfica sobre a análise de sobrevivência e os métodos utilizados no presente trabalho. No Capítulo 3 apresentamos a distribuição Weibull modificada e algumas de suas propriedades. Em seguida, no Capítulo 4 é descrita uma aplicação da distribuição Weibull modificada na presença de fração de cura em dados de câncer gástrico. No Capítulo 4 abordada a distribuição Beta-Weibull e no Capítulo 5 uma aplicação dessa distribuição em um conjutno de dados da literatura. Por último são apresentadas possibilidades de estudos futuros e os anexos dos códigos computacionais usado no trabalho.

# Capítulo 2

# Revisão Bibliográfica

## 2.1 Principais conceitos em Análise de Sobrevivência

Segundo Colosimo e Giolo (2006), a análise de sobrevivência é uma das áreas da estatística que mais se desenvolveram nas últimas décadas, devido principalmente ao aprimoramento de técnicas estatísticas combinados com o avanço dos computadores. Pode-se dizer que ela engloba um conjunto de métodos e modelos, destinados à análise estatística de um tipo de dados que ocorrem quando, em um determinado grupo de indivíduos, é registrado o tempo decorrido de cada indivíduo, do instante inicial até a ocorrência de um evento de interesse (muitas vezes chamado de falha), que pode ser a morte do indivíduo, remissão do evento de interesse, tempo até a cura da doença. Na engenharia, em vez do termo "análise de sobrevivência", é comum o uso do termo "análise de confiabilidade", em que muitas vezes o evento de interesse é a falha de componentes mecânicos ou eletrônicos. Mais detalhes podem serem encontrados em Nelson (2003) e Cook e Lawless (2007).

Em estudos médicos, o tempo inicial pode corresponder frequentemente ao momento de recrutamento do indivíduo em um estudo experimental, em que muitas vezes o interesse é comparar dois ou mais tratamentos. O recrutamento do paciente, pode muitas vezes coincidir com o momento do diagnóstico de certa doença no indivíduo ou até mesmo a partir da realização de um procedimento cirúrgico ou a ocorrência de um evento adverso (como por exemplo um infarto). Dessa forma, a análise de sobrevivência é uma ferramenta importante nas áreas médicas (COLLETT, 2015), permitindo ampliar o entendimento da história natural de doenças específicas, avaliar o efeito de tratamentos sobre o tempo de sobrevida de pacientes, predizer o tempo de sobrevida de indivíduos com doenças crônicas, estudar a associação entre exposição e o tempo até a ocorrência do evento, dentre outras possibilidades.

A principal característica dos dados de sobrevivência, que a difere de outros tipos

de dados, é a presença de censuras, que são as observações incompletas ou parciais das respostas de interesse. Mesmo incompletas, essas informações devem ser incorporados na análise estatística desses dados, pois ainda que estas observações sejam incompletas, elas fornecem informações importantes sobre os tempos de sobrevida.

Dessa forma, para um melhor entendimento dos métodos trabalhados na análise de dados de sobrevivência, alguns conceitos básicos devem ser estudados, tais como, dados censurados, função de sobrevivência, função de risco e estimador de Kaplan-Meier.

### 2.1.1 Tempo de Falha

O tempo de falha é constituído por três elementos: o tempo inicial, a escala de medida e o evento de interesse.

- Tempo inicial: o início do estudo deve ser precisamente definido. Diversos critérios podem ser usados para definir o tempo inicial para cada indivíduo. Contudo, em um estudo, esse critério deve ser o mesmo a todos os indivíduos.
- Escala de medida: em geral, essa escala é o tempo real (calendário). Contudo, existem outras alternativas como: números de ciclos, distância percorrida, entre outros.
- Evento de interesse: em geral esses eventos são indesejados, e em geral chamados de falhas. É fundamental definir o evento de interesse de forma clara. Em certas situações, a falha define-se objetivamente, como a morte de um indivíduo, mas em outros casos esse falha não é tão evidente, como o tempo até um produto deixar de ser útil.

### 2.1.2 Censura

A maioria dos métodos de análises de sobrevivência considera a presença de censura. Essencialmente, a censura ocorre quando temos algumas informações sobre o tempo de sobrevivência individual, mas não sabemos o tempo de sobrevivência exatamente, isto é, o real tempo de sobrevivência pode ser superior ao tempo observado (KLEINBAUM; KLEIN, 2005). Censuras podem ocorrer por uma variedade de razões, dentre elas, a perda de acompanhamento do paciente no decorrer do tempo e a não ocorrência do evento de interesse até o término do experimento.

Colosimo e Giolo (2006) citam duas razões que justificam o uso dos dados censurados nas análises estatísticas: (I) Ainda que estas observações não sejam completas, elas fornecem informações sobre o tempo de vida do paciente e (II) a omissão das observações censuradas pode levar ao cálculo de estimativas viciadas.

Existem diferentes tipos de censuras, entre esse tipos vale ressaltar a censura à direita, à esquerda e intervalar. A censura à direita denota que o tempo de ocorrência do

evento de interesse é inferior ao tempo medido, sendo ainda caracterizado como: censura do tipo I, tipo II e aleatória.

- Censuras do tipo I ocorrem para os indivíduos que ao término do estudo, após um período pré-estabelecido de tempo, ainda não apresentaram o evento de interesse.
- Censuras do tipo II ocorrem para os indivíduos que ao término do estudo e após a ocorrência de um número pré-estabelecido de eventos de interesse, ainda não falharam.
- Censura aleatória, é aquela que mais ocorre na prática. Acontece quando um indivíduo é retirado durante o estudo por um motivo qualquer, sem que ainda tenha ocorrido a falha, ou se o indivíduo morre por uma razão diferente do desfecho de interesse.
- Censura à esquerda, não ocorre muito na prática. Acontece quando o tempo medido é maior do que o verdadeiro tempo de falha, ou seja, o evento de interesse ocorre antes daquele mensurado quando o indivíduo foi observado.
- Censura intervalar, é definida quando o evento de interesse $t$ ocorre entre dois instantes, isto é, $t \in [a, b]$. Neste tipo de censura não se sabe exatamente quando ocorreu o tempo da falha, mas apenas que o evento ocorreu num certo intervalo de tempo.

A censura aleatória mais comum de ocorrer, pode ser representada de forma simples usando duas variáveis aleatórias. Considere $T$ uma variável aleatória que representa o tempo de falha e $C$, outra variável aleatória independente de $T$, representando as observações censuradas. Portanto, para esse indivíduo, temos que seu tempo de falha é dado pela variável aleatória $t = \min(T, C)$, onde o indicador de censura é dado por

$$d = \begin{cases} 1, \text{ se } T \leq C \\ 0, \text{ se } T > C \end{cases} \tag{2.1}$$

Suponha que os pares $(T_i, C_i)$, para $i = 1, ..., n$ formam uma amostra aleatória de $n$ indivíduos. Podemos observar que, se todos apresentam $C_i = C$ (um valor fixado pelo pesquisador), temos nesse caso a censura do tipo I. Dessa forma, conclui-se que a censura do tipo I é um caso particular de censura aleatória. Mais detalhes sobre censuras podem ser encontradas em Lawless (2002) e Klein e Moeschberger (2005).

### 2.1.3 Representação dos Dados de Sobrevivência

Os dados de sobrevivência para um indivíduo $i$ $(i = 1, \ldots, n)$ em estudo são em geral representados por um par ordenado $(t_i, d_i)$, em que $t_i$ é o tempo de falha ou de

censura do indivíduo $i$ e $d_i$ é o indicador de falha ou de censura, isto é,

$$d_i = \begin{cases} 1, \text{ se } t_i \text{ é um tempo de falha} \\ 0, \text{ se } t_i \text{ é um tempo censurado} \end{cases}$$

Na presença de covariáveis medidas para o $i$-ésimo indivíduo, os dados de sobrevivência podem serem representados por $(t_i, d_i, \boldsymbol{x}_i)$, em que $\boldsymbol{x}$ é o vetor de covariáveis, que inclui por exemplo: idade, sexo e tipo de tratamento. No caso em que os dados de sobrevivência são intervalares, tem-se ainda a representação $(a_i, b_i, d_i, \boldsymbol{x}_i)$ em que $a_i$ e $b_i$ são, respectivamente, o limite inferior e o superior do intervalo observado para o $i$-ésimo individuo (COLOSIMO; GIOLO, 2006).

### 2.1.4 Principais Funções em Análise de Sobrevivência

Em análise de sobrevivência duas funções são consideradas de destaque: a função de sobrevivência (que representa o tempo de falha) e a função de risco. Outras duas funções de interesse são a expressão para estimar o tempo médio de falha e a que estima a variância do tempo médio de falha.

#### 2.1.4.1 Função de Sobrevivência

A função de sobrevivência é definida como a probabilidade de uma determinada observação não falhar até um certo tempo $t$, ou seja é a probabilidade de uma observação sobreviver até o tempo $t$. Em termos de probabilidade, podemos escrever,

$$S(t) = P(T \geq t) = \int_t^\infty f(t)\,\mathrm{d}t = 1 - F(t), \tag{2.2}$$

em que $f(t)$ é a função densidade de probabilidade e $F(t)$ é a função acumulada de probabilidade.

A função de sobrevivência deve satisfazer às seguintes propriedades:

- $S(t)$ é decrescente;
- $S(0) = 1$;
- $\lim_{t\to\infty} S(t) = 0$

A função de sobrevivência é muito utilizada para comparar ajustes de distribuições, pois ela pode ser representada graficamente por uma curva, chamada de curva de sobrevivência. Uma informação muito importante que é possível obter das curvas de sobrevivência é o percentual de indivíduos que ainda não falharam até determinado tempo.

#### 2.1.4.2 Função de Risco

A taxa de risco ou de falha em um intervalo é definida como a probabilidade de um indivíduo falhar durante um intervalo $(t, t + \Delta t]$, sabendo-se que esse indivíduo ainda não tinha falhado até o instante $t$. Em termos probabilísticos, temos então,

$$P(t < T \leq t + \Delta t | T > t) = \frac{P(t < T \leq t + \Delta t)}{P(T > t)} = \frac{S(t) - S(t + \Delta t)}{S(t)},$$

dividindo esta probabilidade por intervalo de tempo $\Delta t$ e considerando $\Delta t$ muito pequeno, temos a função de risco no instante $t$, ou seja,

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t},$$

resultando assim na relação

$$h(t) = \frac{f(t)}{S(t)}. \tag{2.3}$$

A função de risco $h(t)$, é muito útil para descrever a distribuição do tempo de falhas dos indivíduos. Ela descreve a forma com que a taxa de falha instantânea muda com o tempo. Interessante dizer que as forma das funções de risco são às vezes mais importantes de se modelar do que a própria função de sobrevivência, visto que diferentes funções de sobrevivência apresentam formas semelhantes, enquanto que as respectivas funções de risco podem diferir drasticamente. Formas comuns para a função de risco é a constante, decrescente, crescente, unimodal e de banheira (decrescente no início e crescente no fim).

#### 2.1.4.3 Tempo Médio e Variância do Tempo Médio

A média até o tempo de falha considerando uma distribuição de probabilidade, é definida como

$$\mu_T = \int_0^{+\infty} t f_0(t)\, \mathrm{d}t = \int_0^{+\infty} S_0(t)\, \mathrm{d}t. \tag{2.4}$$

Temos que a variância até o tempo de falha é dada por

$$\mathrm{Var}(T) = \int_0^{+\infty} t^2 f_0(t)\, \mathrm{d}t - \mu^2 = 2 \int_0^{+\infty} t S_0(t)\, \mathrm{d}t - \mu^2. \tag{2.5}$$

É muito comum essas expressões não terem formas analíticas fechadas, ficando suas soluções a cargo de se resolver numericamente a integral.

Outras funções e mais detalhes destas podem ser encontradas em Lawless (2002) e Klein e Moeschberger (2005).

### 2.1.5 Estimador de Kaplan-Meier

O estimador de Kaplan-Meier, é o mais utilizado para estimar a função de sobrevivência. Foi proposto por Kaplan e Meier (1958) e é também chamado de estimador

produto-limite. Esse estimador é uma adaptação da função de sobrevivência empírica que, na ausência de censuras, é definido como:

$$\hat{S}(t) = \frac{\text{nº de observações que não falharam até o tempo } t}{\text{nº total de observações no estudo}}.$$

Notemos que $\hat{S}(t)$ é uma "função escada" em que os degraus nos tempos observados de falhas são de tamanhos $1/n$, sendo $n$ o tamanho da amostra. Se houverem empates em qualquer tempo $t$, o tamanho do degrau é agora multiplicado pelo número de empates. O estimador de Kaplan-Meier considera tantos os intervalos de tempo quanto forem os números de falhas distintas, sendo que os limites dos intervalos de tempo são os tempos de falhas observados na amostra (COLOSIMO; GIOLO, 2006).

Para a expressão geral do estimador de Kaplan-Meier considere:

- $t_1 < t_2 < \cdots < t_k$, os $k$ tempos de falhas distintos e ordenados de falha,
- $d_j$ o número de falhas observadas em $t_j$, $j = 1, \ldots, k$, e
- $n_j$ o número de indivíduos sob risco no instante $t_j$ ou, em outras palavras, os indivíduos que não falharam e não foram censurados até o momento imediadamente anterior a $t_j$.

Dessa forma, o estimador de Kaplan-Meier é definido como:

$$\hat{S}(t) = \prod_{j:t_j<t} \left(1 - \frac{d_j}{n_j}\right). \tag{2.6}$$

A expressão para a variância assintótica do estimado de Kaplan-Meier é a seguinte:

$$\widehat{\text{Var}}\left(\hat{S}(t)\right) = \left[\hat{S}(t)\right]^2 \sum_{j:t_j<t} \frac{d_j}{n_j(n_j - d_j)}. \tag{2.7}$$

A prova de (2.7) e outros detalhes sobre o estimador de Kaplan-Meier podem ser encontrado em Kalbfleisch e Prentice (2011). A consistência e normalidade assintótica de $\hat{S}(t)$, sob certas condições de regularidade, foram demonstrado por Breslow et al. (1974) e no artigo original de Kaplan e Meier (1958) é provado que $\hat{S}(t)$ é o estimador de máxima verossimilhança de $S(t)$.

Outros estimadores para $S(t)$ são possíveis, detalhes complementares sobre eles podem serem encontrados em Tableman e Kim (2003) e Klein e Moeschberger (2005).

## 2.2 Gráfico TTT-plot

É comum em análise de sobrevivência desejar identificar previamente às análises a forma da função de risco. Um método visual que se destaca para estimar essa forma é

o gráfico conhecido como TTT-plot (gráfico do tempo total em teste) que foi proposto por Barlow e Campo (1975).

Aarset (1985) propôs uma forma empírica para determinar o comportamento da função de risco por meio da construção do gráfico TTT-plot, dada pela seguinte equação:

$$G\left(\frac{r}{n}\right) = \frac{\sum_{i=1}^{r} T_{[i]} + (n-r)T_{[r]}}{\sum_{i=1}^{r} T_{[i]}} \tag{2.8}$$

onde $r = 1, ..., n$ e $T_{[i]}$ com $i = 1, ..., n$ são as estatísticas de ordem da amostra.

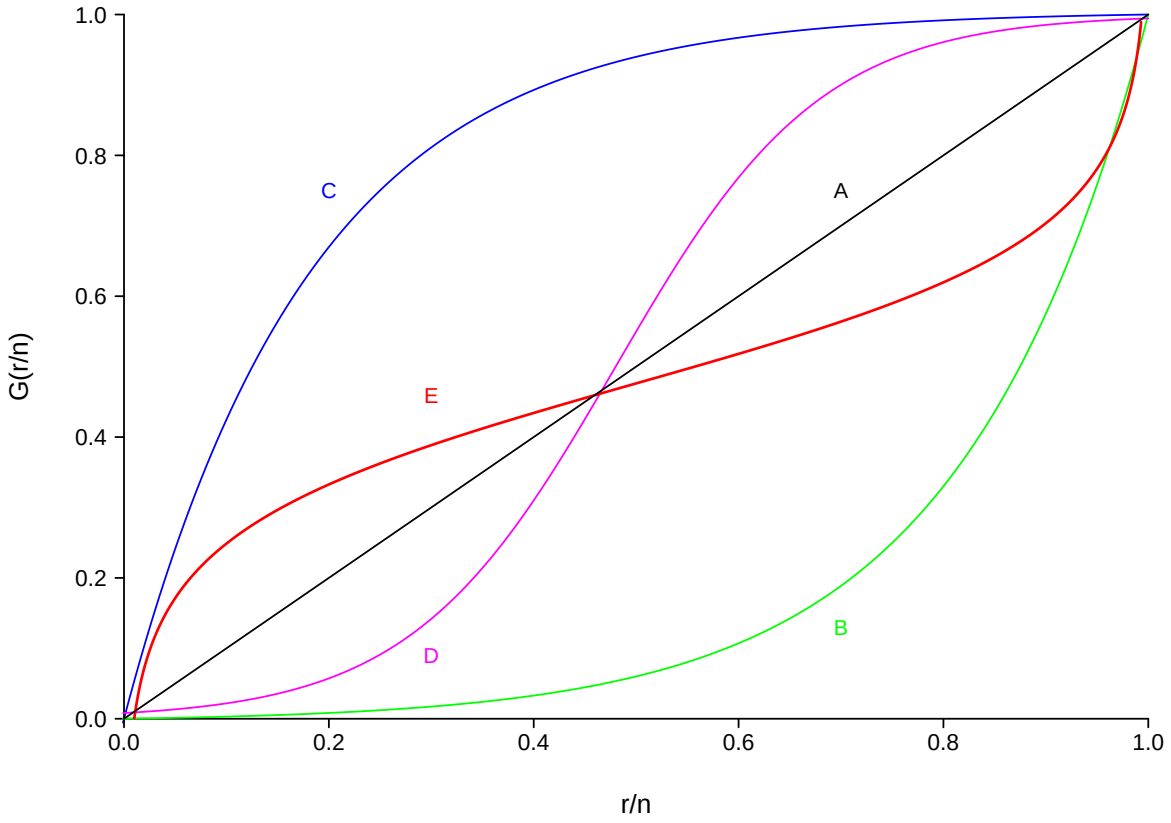


Figura 1 – Representações de alguns TTT-plots

A Figura 1 mostra alguns TTT-plots ilustrativos, onde é possível identificar as formas para a $h(t)$.

Temos assim, as seguintes identificações:

- Função de risco constante, quando o TTT-plot apresentar uma reta diagonal, representada em A,
- Função de risco crescente, quando TTT-plot descreve uma curva côncava, representada em B,
- Função de risco decrescente, curva convexa no TTT-plot, representada em C,

- Função de risco com forma de banheira, quando identificamos uma curva convexa e depois côncava no TTT-plot, representada em D,
- Função de risco unimodal, curva côncava e de depois convexa no TTT-plot, representada em E.

## 2.3 Fração de Cura

Com os tradicionais modelos de sobrevivência não é possível estimar qual é a proporção de cura de uma população. Dessa forma, é necessário a utilização de modelos estatísticos que são capazes de incorporar na sua modelagem estatística a fração de cura. Assim, vários modelos surgiram como uma maneira de superar as limitações dos modelos tradicionais, já que nessa nova classe de modelos admite-se que uma parte da população é imune ao evento de interesse.

### 2.3.1 Modelos de fração de cura com mistura e não mistura

Segundo Maller e Zhou (1996), a forma comum de assumir um modelo com fração de cura é considerar que a população em estudo é uma mistura de indivíduos suscetíveis ao evento de interesse e indivíduos não suscetíveis.

Há muitas formas de abordagens para modelar a fração de indivíduos que não estão suscetíveis ao evento de interesse, desde abordagens paramétricas a abordagens não paramétricas. Seja $T$ uma variável aleatória que representa o tempo até a ocorrência de um evento, e $t > 0$ uma observação de $T$. Uma função de sobrevivência do modelo de mistura com fração de cura padrão é dada por

$$S(t) = p + (1 - p)\, S_0(t), \tag{2.9}$$

onde $p$ é o parâmetro que representa a proporção de indivíduos que não estão suscetíveis ao evento de interesse, sendo que $(0 < p < 1)$, e $S_0(t)$ é a função de sobrevivência basal para os indivíduos que estão suscetíveis ao eventos, os não-curados.

Pode-se notar que, se $t$ tende ao infinito, então $S(t)$ tende a $p$, dado que $\lim_{t\to\infty} S_0(t) = 0$ . Dessa forma a função de sobrevivência do modelo com fração de cura padrão tem uma assíntota no valor $p$, que estima a proporção de curados. Para esse modelo a função de probabilidade para $T$ é

$$f(t) = \frac{dF(t)}{dt} = (1 - p) f_0(t), \tag{2.10}$$

onde $F(t) = 1 - S(t)$ e $f_0(t)$ é a função densidade basal de probabilidade para os indivíduos suscetíveis ao evento. Logo, para este modelo, a função de risco é dado por

$$h(t) = \frac{f(t)}{S(t)} = \frac{(1 - p) f_0(t)}{p + (1 - p) S_0(t)}. \tag{2.11}$$

Uma alternativa ao modelo de mistura é o modelo de não mistura, sugerido por alguns autores, como Tsodikov et al. (2003), Lambert et al. (2007) e Achcar et al. (2012). Este modelo define uma assíntota para o risco acumulado, e consequentemente para a fração de cura. A função de sobrevivência neste caso é dada por

$$S(t) = p^{F_0(t)} = \exp\left[\ln\left(p^{F_0(t)}\right)\right] = \exp[\ln(p)F_0(t)], \tag{2.12}$$

onde $F_0(t) = 1 - S_0(t)$. Considerando o modelo de não mistura temos que a função densidade de probabilidade para ele é dada por

$$f(t) = \frac{dF(t)}{dt} = -\ln(p)f_0(t)\exp[\ln(p)F_0(t)], \tag{2.13}$$

Assim, a função de risco para esse modelo é dada por

$$h(t) = \frac{f(t)}{S(t)} = -\ln(p)f_0(t). \tag{2.14}$$

## 2.4 Método da Máxima Verossimilhança

Esse é o método mais usual para a estimação dos parâmetros de uma distribuição de probabilidade, sendo utilizado pela primeira vez por Fisher (1912). O método de máxima verossimilhança busca os valores dos parâmetros que maximizam a função de verossimilhança. Seja $\boldsymbol{x} = (x_1, x_2, ..., x_n)$ uma amostra aleatória de tamanho $n$ da variável aleatória $X$ com função de probabilidade $f(x|\theta)$, com $\theta \in \Theta$, onde $\Theta$ é o espaço paramétrico. A função de verossimilhança de $\theta$ correspondente à distribuição da amostra aleatória observada e identicamente distribuída, é definida como

$$L(\theta; \boldsymbol{x}) = \prod_{i=1}^{n} f(x_i|\theta)$$

O estimador de máxima verossimilhança de $\theta$ é o valor $\hat{\theta} \in \Theta$ que maximiza a função de verossimilhança $L(\theta; \boldsymbol{x})$.

Muitas vezes é mais simples algebricamente e até mesmo computacionalmente encontrar o valor que maximiza o logaritmo da função de verossimilhança, $\ell(\theta; \boldsymbol{x})$. É fácil observar que maximizar $\ell(\theta; \boldsymbol{x})$ é o mesmo que maximizar $L(\theta; \mathbf{x})$, pois a função logaritmo é estritamente crescente. O estimador de máxima verossimilhança uniparamétrico pode ser encontrado como a raiz da equação de primeira derivada de $\ell(\theta; \boldsymbol{x})$, isto é

$$\ell'(\theta; \boldsymbol{x}) = \frac{\partial \ell(\theta; \boldsymbol{x})}{\partial \theta} = 0$$

Em alguns casos, a solução da equação de verossimilhança pode ser obtida explicitamente. Nas situações mais complicadas, a solução da equação será em geral obtida por

procedimentos numéricos. Para concluir que a solução da equação é ponto de máximo, é realizada a verificação da segunda derivada, tal que

$$\ell''(\theta; \boldsymbol{x}) = \frac{\partial^2 \ell(\theta; \boldsymbol{x})}{\partial \theta^2}|_{\theta=\hat{\theta}} < 0$$

Pode-se generalizar a função de verossimilhança para o caso multiparamétrico. Considere $\boldsymbol{\theta}$ um vetor de parâmetros de dimensão $n$, $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$, ou até mesmo uma matriz de parâmetros. Se $\boldsymbol{\theta}$ é um vetor de parâmetros, as equações de verossimilhança são dadas por

$$\ell'(\boldsymbol{\theta}; \boldsymbol{x}) = \frac{\partial \ell(\boldsymbol{\theta}; \boldsymbol{x})}{\partial \theta_j} = 0$$

com $j$ variando de $1$ a $j$, em que $j$ é o número de parâmetros em $\boldsymbol{\theta}$. Nesse caso, para garantir que a solução do sistema seja um ponto de máximo, a condição de segunda ordem é dada por uma análise na matriz Hessiana, que se refere à matriz de segunda ordem da função de verossimilhança. A condição é que

$$H = \frac{\partial^2 \ell(\boldsymbol{\theta}; \boldsymbol{x})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} < 0$$

seja negativa definida, onde cada elemento de $H$ é dado por

$$h_{ij} = \frac{\partial^2 \ell(\boldsymbol{\theta}; \boldsymbol{x})}{\partial \theta_i \partial \theta_j}.$$

Pode-se ter situações em que $\Theta$ é discreto ou em que o máximo de $\ell''(\boldsymbol{\theta}; \boldsymbol{x})$ ocorre na fronteira de $\Theta$. O estimador de máxima verossimilhança não pode então ser obtido a partir da solução da derivada de $\ell''(\boldsymbol{\theta}; \boldsymbol{x})$ . Nessas situações, o máximo é obtido a partir de uma análise da função de verossimilhança (BOLFARINE; SANDOVAL, 2001).

### 2.4.1 Intervalos de Confiança

Os intervalos de confiança dos parâmetros estimados por meio do método de máxima verossimilhança podem ser obtidos através de várias formas, como o método assintótico ou pelo perfil da verossimilhança. Intervalos de confiança para um parâmetro $\theta$ ou para um vetor de parâmetros $\boldsymbol{\theta}$ (nesse caso também chamado de região de confiança) podem ser construído usando-se qualquer estatística de teste (CORDEIRO; DEMÉTRIO, 2008).

#### 2.4.1.1 Intervalo Assintótico

O método assintótico é construído a partir da inversão do teste da Normal ou teste $Z$ (versão simplificado do teste de Wald), considerando que sobre certas condições $\hat{\theta} \sim N(0, 1)$, então o intervalo assintótico para um parâmetro $\theta_k$ com (1-$\alpha$)100% é dado por:

$$(\hat{\theta_k} - Z_{1-\alpha/2}\sqrt{\text{Var}(\hat{\theta_k})}, \hat{\theta_k} + Z_{1-\alpha/2}\sqrt{\text{Var}(\hat{\theta_k})}),$$

onde $Z_{1-\alpha/2}$ é o quantil de uma distribuição Normal padrão com coeficiente de confiança $\alpha$ e Var$(\hat{\theta}_k)$ é a variância de $\hat{\theta}_k$ que pode ser estimada por meio da matriz hessiana, onde

$$\text{Var}(\hat{\theta}) = -\left[E\left(H(\hat{\theta})\right)\right]^{-1} = \begin{bmatrix} \text{Var}(\theta_1) & \text{Cov}(\theta_1,\theta_2) & \cdots & \text{Cov}(\theta_1,\theta_k) \\ \text{Cov}(\theta_2,\theta_1) & \text{Var}(\theta_2) & \cdots & \text{Cov}(\theta_2,\theta_k) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(\theta_k,\theta_1) & \text{Cov}(\theta_k,\theta_2) & \ldots & \text{Var}(\theta_k) \end{bmatrix}.$$

#### 2.4.1.2 Intervalo Perfilado

Uma outra forma de construir intervalos de confiança é utilizando o teste das razões de verossimilhança. Nessa construção encontramos todos os valores de $\theta$ para o qual $l(\theta, \boldsymbol{x})$ está dentro de uma valor máximo tolerado de $l(\hat{\theta}, \boldsymbol{x})$. A teoria estatística nos diz que se $\boldsymbol{\theta}_0$ é o vetor dos verdadeiros valores dos parâmetros, então a estatística da razão de verossimilhança é

$$2\log\left(\frac{L(\hat{\boldsymbol{\theta}}, \boldsymbol{x})}{L(\boldsymbol{\theta}_0, \boldsymbol{x})}\right) = 2[l(\hat{\boldsymbol{\theta}}, \boldsymbol{x}) - l(\boldsymbol{\theta}_0, \boldsymbol{x})] \sim \chi^2_r,$$

onde $r$ são os graus de liberdade da distribuição qui-quadrado, sendo que $r$ é o número de parâmetros em restrição, isto é, o número do parâmetros em teste. Este é conhecido como teste da razão das verossimilhanças (LR, do inglês likelihood ratio), em que $H_0 : \hat{\boldsymbol{\theta}} = \boldsymbol{\theta}$ versus $H_1 : \hat{\boldsymbol{\theta}} \neq \boldsymbol{\theta}$ (MILLAR, 2011).

No caso particular em que $r = 1$, ao nível de significância de $\alpha = 5\%$, podemos rejeitar a hipótese nula se o valor da estatística for maior que $3,84$. Assim podemos usar o principio da razão de verossimilhança para construir intervalos de confiança, para $\theta$ com coeficiente $\alpha$.

Um intervalo aproximado de $100(1-\alpha)\%$ para $\theta$ consiste em todos os possíveis valores de $\theta$ para qual a hipótese nula não seja rejeitada. Então para um intervalo de $95\%$, consiste em todos os valores de $\theta$ para o qual

$$2[l(\boldsymbol{\theta}, \boldsymbol{x}) - l(\theta_0, \boldsymbol{x})] \leq 3,84,$$

ou

$$l(\theta_0, \boldsymbol{x}) \geq l(\hat{\theta}, \boldsymbol{x}) - 1,92.$$

Em outras palavras, o intervalo de 95% inclui todos os valores de $\theta$ para qual a função de log-verossimilhança não decai mais do que 1,92 unidades.

Quando o tamanho da amostra é grande ou quando a forma do perfil da verossimilhança é simétrica com curvatura não muito grande, o método LR tende a produzir intervalos muito semelhantes aos observados com base no método assintótico. Contudo, ao contrário dos intervalos assintóticos, os intervalos LR são escala-invariante, isto é, se encontrarmos

o intervalo LR para uma versão transformada do parâmetro, tal como $\phi = \log\left[\theta/\left(1-\theta\right)\right]$ e, em seguida, transformar os resultados de volta para a escala $\theta$, temos exatamente a mesma resposta que se aplicarmos o método LR diretamente na escala $\theta$. Esta pode ser considerada uma vantagem do método LR.

Na Figura 2 é mostrada a construção do intervalo perfilado de forma gráfica, o limite inferior e o superior de $\theta$ será onde a linha pontilhada corta o gráfico da função de log-verossimilhança.

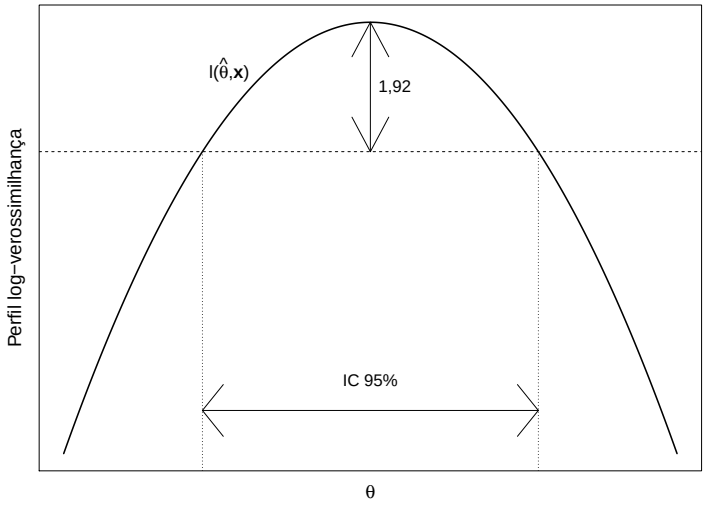


Figura 2 – Intervalo de confiança perfilado para um parâmetro $\theta$.

Outros detalhes e propriedades para o método assintótico podem ser obtido em Casella e George (1992) e Pawitan (2001).

### 2.4.2 Medida de Assimetria para Intervalo de Confiança Perfilado

Pode-se calcular uma medida de assimetria (A) para o intervalo de confiança perfilado de um parâmetro $\theta$ por meio da seguinte equação:

$$A = \frac{\left(L_s - \hat{\theta}\right) - \left(\hat{\theta} - L_i\right)}{L_s - L_i} \times 100, \tag{2.15}$$

onde $\hat{\theta}$ é o valor estimado do parâmetro, $L_i$ e $L_s$ são os limites inferiores e superiores do intervalo de confiança perfilado.

Se $A = 0$ o intervalo perfilado é simétrico, coincidindo com o assintótico, caso $A > 0$ o intervalo é assimétrico à direito e se $A < 0$ temos um intervalo assimétrico à esquerda. Quando maior o valor de $|A|$, maior assimetria apresenta o intervalo.

## 2.5 Funções de Ligação

Muitas vezes desejamos adicionar covariáveis ao modelo, onde essas covariáveis implicam nas estimativas de um ou mais parâmetros adicionais. A fim de adicionar essas covariáveis no modelo e relacioná-las aos parâmetros, é utilizado uma função de ligação, onde uma função monótona $g(\theta)$, relaciona o parâmetro $\theta$ a uma função linear (predito linear). Considera-se que um preditor linear é dado por

$$\eta = \boldsymbol{\varphi}\boldsymbol{x}^t = \varphi_0 + \varphi_1 x_1 \ ... \ \varphi_n x_n,$$

sendo que $\boldsymbol{\varphi} = (\varphi_0, \varphi_1 ... \varphi_n)$ é um vetor dos parâmetros do preditor linear, onde $\varphi_0$ é o intercepto do preditor e os outros parâmetros são respectivos a cada uma das covariáveis $x_1, ..., x_n$ do vetor de covariável $\boldsymbol{x}$. Na literatura é possível encontrar diversas funções de ligação, entretanto, temos que algumas das funções de ligação mais utilizadas para $g(\theta) = \eta$ e suas inversas, em que $\theta = g^{-1}(\eta)$, são

**(a)** Ligação Identidade

$$g(\theta) = \theta \text{ e } g^{-1}(\eta) = \eta$$

**(b)** Ligação Logarítmica

$$g(\theta) = \ln(\theta) \text{ e } g^{-1}(\eta) = \exp(\eta)$$

**(c)** Ligação Inversa

$$g(\theta) = \theta^{-1} \text{ e } g^{-1}(\eta) = \eta^{-1}$$

**(d)** Ligação Inversa Quadrática

$$g(\theta) = \theta^{-2} \text{ e } g^{-1}(\eta) = \eta^{-2}$$

**(e)** Ligação Raiz Quadrada

$$g(\theta) = \sqrt{\theta} \text{ e } g^{-1}(\eta) = \eta^2$$

**(f)** Ligação Logito

$$g(\theta) = \ln\left(\frac{\theta}{1-\theta}\right) \text{ e } g^{-1}(\eta) = \frac{exp(\eta)}{1+exp(\eta)}$$

**(g)** Ligação Probito

$$g(\theta) = \Phi^{-1}(\theta) \text{ e } g^{-1}(\eta) = \Phi(\eta)$$

em que $\Phi(.)$ é a função de distribuição acumulada de uma variável aleatória normal padrão, ou seja $\Phi(k) = \int_{-\infty}^{k} \sqrt{2\pi} \exp\left(-\frac{v^2}{2}\right) \mathbf{d}v$.

**(h)** Ligação Log-Log

$$g(\theta) = -\ln(-\ln(\theta)) \text{ e } g^{-1}(\eta) = \exp(-\exp(-\eta))$$

**(i)** Ligação Complementar Log-Log

$$g(\theta) = \ln(-\ln(1-\theta)) \text{ e } g^{-1}(\eta) = 1 - \exp(-\exp(\eta))$$

No momento de se escolher alguma dessas funções de ligação, vários fatores são considerados. É necessário conhecer o domínio do espaço paramétrico em questão, pois cada uma dessas funções possuem domínio e imagem diferentes. É possível perceber que entre essas, apenas **(a)** é uma função de ligação adequada para parâmetros em que o domínio deste é o conjunto $(-\infty, +\infty)$. Para os parâmetros em que o seu domínio são os números maiores que $0$, é adequado utilizar as seguintes funções de ligações: **(b)**, **(c)**, **(d)** e **(e)**. As outras funções de ligações (**(f)**, **(g)**, **(h)** e **(i)**), são adequadas para os parâmetros que pertencem ao intervalo $(0, 1)$.

## 2.6 Método Delta

O método delta é uma técnica comumente utilizada para obter intervalos de confiança para funções que possuem parâmetros estimados por máxima verossimilhança. A variância da função que se deseja obter o intervalo de confiança frequentemente é analiticamente difícil de se resolver. O método delta cria uma aproximação linear dessa função, expandindo-a até a primeira ordem de uma série de Taylor e obtendo uma boa aproximação para tamanhos de amostras grandes, graças à propriedade de invariância dos estimadores de máxima verossimilhança.

Sob condições de regularidade, que podem ser encontrado com mais detalhes em CORDEIRO (1992), se $\hat{\theta}$ é um vetor das estimativas de ML, temos que

$$\sqrt{n}\left(\hat{\theta} - \theta\right) \xrightarrow{d} N\left(0, \text{Var}\left(\hat{\theta}\right)\right).$$

Seja $f(\theta)$ um função de parâmetros $\hat{\theta}$. Expandindo $F(\hat{\theta})$ em uma série de Taylor em torno de $\theta$ e ignorando os termos superiores ao de primeira ordem, temos que

$$F(\hat{\theta}) = F(\theta) + (\hat{\theta} - \theta)' F'(\theta),$$

onde $F'(\theta)$ é a matriz de derivadas parciais de primeira ordem em relação a $\theta$. Assim temos que

$$\sqrt{n}\left[F\left(\hat{\theta}\right) - F\left(\theta\right)\right] = \sqrt{n}(\hat{\theta} - \theta)' \mathbf{F'}(\theta).$$

Dessa forma temos que

$$F(\hat{\theta}) \to N\left(F\left(\theta\right), \frac{\partial F(\theta)}{\partial \theta'} \text{Var}(\hat{\theta}) \frac{\partial F(\theta)}{\partial \theta}\right),$$

onde $Var(\hat{\theta})$ é a matriz de covariância das estimativas de $\theta$. Portanto, a estimativa da variância de $F\left(\hat{\theta}\right)$ é avaliada sobre as parciais das estimativas ML de $\theta$. Logo,

$$\text{Var}\left[F\left(\hat{\theta}\right)\right] = \frac{\partial F(\hat{\theta})}{\partial \hat{\theta}'}\text{Var}(\hat{\theta})\frac{\partial F(\hat{\theta})}{\partial \hat{\theta}}.$$

Uma forma de encontrar a matriz de covariância considera que, sob certas condições de regularidade,

$$\text{Var}\left(\hat{\theta}\right) \approx -\left[E\left(F'(\hat{\theta})\right)\right]^{-1}.$$

Dessa forma, pode-se dizer que a matriz de covariância dos estimadores de máxima verossimilhança é aproximadamente o negativo da inversa da esperança da matriz de segundas derivadas de $F(\theta)$. Nas situações em que é complicado ou impossível de calcular a esperança, podemos usar apenas $-\left[\left(F'(\hat{\theta})\right)\right]^{-1}$, que é um estimador consistente de $-\left[E\left(\ell(\hat{\theta})\right)\right]^{-1}$ (COLOSIMO; GIOLO, 2006). Nessa matriz os elementos da diagonal principal são as variâncias dos estimadores e os outros elementos são as covariâncias entre eles.

## 2.7 Introdução à Inferência Bayesiana

Os métodos Bayesianos têm se tornado uma ferramenta cada vez mais poderosa na análise de dados. Contudo, esta não é uma ideia recente. Em 1763, o reverendo e matemático inglês Thomas Bayes, em sua obra póstuma *An Essay Towards Solving a Problem in the Doctrine of Chances*, publicada por Bayes (1763), apresentou seus fundamentos, com base em probabilidades condicionais. Na Inferência Bayesiana, a informação que se tem sobre $\boldsymbol{\theta}$, assume diferentes graus de incerteza, estes são representados por meio de modelos probabilísticos para $\boldsymbol{\theta}$. Assim, é possível que os pesquisadores possuam diferentes graus de incerteza sobre $\boldsymbol{\theta}$, considerando modelos distintos para $\boldsymbol{\theta}$. Nesse método não existe distinção entre os dados observados e os parâmetros do modelo, todos são considerados como aleatórios (EHLERS, 2007) . De acordo com Gelman et al. (2014), a Inferência Bayesiana pode ser entendida como o processo de ajustar um modelo de probabilidade para um conjunto de dados, resumindo o resultado por uma distribuição de probabilidade nos parâmetros do modelo e em quantidades não observáveis como predição para novas observações.

### 2.7.1 Teorema de Bayes

A fundamentação da teoria de inferência Bayesiana é baseada no teorema de Bayes.

Sejam os eventos $\text{A}_1, \text{A}_2, ..., \text{A}_n$ uma sequência de eventos mutuamente exclusivos associados ao experimento aleatório, formando uma partição do espaço amostral $\Omega$ . Então,

para qualquer outro evento B, B $\subset \Omega$, e para todo $i = 1, 2, ..., n$, temos que

$$\mathsf{P}(\mathsf{A}_i|\mathsf{B}) = \frac{\mathsf{P}(\mathsf{B}|\mathsf{A}_i)\mathsf{P}(\mathsf{A}_i)}{\sum_{i}^{n} \mathsf{P}(\mathsf{B}|\mathsf{A}_i)\mathsf{P}(\mathsf{A}_i)}.$$

Pode-se interpretar esse teorema da seguinte forma: antes do conhecimento de qualquer informação sobre o evento $\mathsf{A}_i$, ao atribuir uma probabilidade a *priori* para $\mathsf{A}_i$, dada por $\mathsf{P}(\mathsf{A}_i)$, essa probabilidade é atualizada a partir da ocorrência do evento B . Essa probabilidade atualizada, ou probabilidade condicional do evento $\mathsf{A}_i$ dada a ocorrência do evento B, ou seja, a *posteriori* $\mathsf{P}(\mathsf{A}_i|\mathsf{B})$, é dada pelo teorema de Bayes.

### 2.7.2 Atualização da Incerteza

Considere uma quantidade de interesse desconhecida $\boldsymbol{\theta}$ representando os parâmetros de uma distribuição de probabilidade e um vetor de dados $\mathbf{x} = (x_1, x_2, ..., x_n)$, associado a uma variável aleatória $X_i$ com valores observados $x_i, i = 1, 2, ..., n$. A informação de que dispomos sobre $\boldsymbol{\theta}$, descrita probabilisticamente por meio de $\pi(\boldsymbol{\theta})$, pode ser aumentada observando-se um valor aleatório de $X_i$ relacionada com $\boldsymbol{\theta}$, na qual a distribuição amostral $p(x|\boldsymbol{\theta})$ é quem define essa relação. Uma vez observado $X_i = x_i$, a quantidade de informação sobre $\boldsymbol{\theta}$ aumenta e o teorema de Bayes é a regra de atualização aplicada para medir esse aumento de informação (EHLERS, 2007). Assumindo valores discretos de $\theta_1, \theta_2, ..., \theta_n$, temos que a distribuição *a posteriori* para $\theta_i$ dado $\boldsymbol{x}$, é dada por

$$\pi(\theta_i|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta_i)\pi(\theta_i)}{\sum_{i}^{n} p(\boldsymbol{x}|\theta_i)\pi(\theta_i)}. \tag{2.16}$$

Note que o parâmetro $\theta$ é considerado como um valor aleatório sob o ponto de vista Bayesiano.

Para o parâmetro $\boldsymbol{\theta}$ contínuo num dado intervalo, temos que a distribuição *posteriori* para $\theta_i$ dado $\boldsymbol{x}$, é dada por

$$\pi(\theta_i|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|\theta_i)\pi(\theta_i)}{\int_{\theta} p(\boldsymbol{x}|\theta)\pi(\theta_i)}. \tag{2.17}$$

Para um valor fixo de $x_i$, a função $p(\boldsymbol{x}|\theta) = L(\boldsymbol{\theta}; x_i)$ é a função de verossimilhança para cada um dos possíveis valores de $\boldsymbol{\theta}$, enquanto $\pi(\boldsymbol{\theta})$ é chamado de distribuição a *priori* de $\boldsymbol{\theta}$. Geralmente, utilizam-se distribuições a *priori* não informativas. Mais detalhes sobre distribuições a *priori* podem serem obtidas em Paulino et al. (2003), Jayaata et al. (2006) e Box e Tiao (2011). Estas duas fontes de informação, a distribuição *a priori* e a função de verossimilhança, são combinadas para obter a distribuição *posteriori* de $\boldsymbol{\theta}$, $\pi(\theta_i|\boldsymbol{x})$. Como nas expressões acima o denominador é igual a imagem do ponto $x_i$ pela função de probabilidade de $X$, o denominador das expressões são constantes. Assim a forma usual de descrever o teorema de Bayes para a atualização da incerteza é

$$\pi(\theta_i|\boldsymbol{x}) \propto L(\boldsymbol{\theta}; x_i)\pi(\theta_i)$$

Note que, ao omitir o denominador, temos uma proporcionalidade. Esta forma simplificada do teorema de Bayes é útil na estimação de parâmetros, pois o denominador é apenas uma constante normalizadora. Em outras situações, como seleção de modelos, este termo tem um papel importante (EHLERS, 2007; ROSSI, 2011). O objetivo da inferência Bayesiana é justamente formalizar a passagem de $\pi(\theta)$ para $\pi(\theta|\boldsymbol{x})$ (ROSSI, 2011).

Para calcular as inferências da distribuição *posteriori* conjunta deve-se que encontrar uma distribuição para um parâmetro específico de $\boldsymbol{\theta}$. Para encontrar essa distribuição, chamada de distribuição marginal de $\boldsymbol{\theta}$, é necessário integrar a distribuição *posteriori* conjunta em relação aos outros parâmetros do modelo, ou seja, temos que resolver a seguinte integral múltipla

$$\pi(\theta_i|\boldsymbol{x}) = \int ... \int \pi(\theta_i, \theta_{-i}|\boldsymbol{x})\, \boldsymbol{d}\theta_{-i}$$

em que $\theta_i$ é o parâmetro que esta sendo calculada a inferência e $\theta_{-i}$ é o conjunto complementar de parâmetros de $\boldsymbol{\theta}$, isto é, são os parâmetro do modelo excluindo o parâmetro $\theta_i$ (ROSA, 1998).

Muitas vezes, a forma analítica da distribuição marginal é multiparamétrica e complexa. Além disso, comumente não há uma forma analítica para ela, e para encontrar a solução dessa integral deve-se usar aproximações. Dessa forma, métodos numéricos iterativos, como o método de Monte Carlo via cadeias de Markov (MCMC) são utilizados para gerar valores de uma distribuição condicional a posteriori para cada um dos parâmetros (SORENSEN, 1996). Dentre os principais métodos de simulação que fazem o uso de cadeias de Markov, pode-se citar o amostrador de Gibbs (GELFAND et al., 1990) e o Metropolis-Hastings (CHIB; GREENBERG, 1995).

Segundo Zeger e Karim (1991), a motivação para o uso da amostragem de Gibbs é que a verdadeira distribuição *posteriori* pode ser aproximada por uma distribuição empírica de $B$ valores, de tal forma que $B$ seja grande o suficiente para que a amostragem de Gibbs atinja a convergência e a estacionariedade. A determinação da convergência e da estacionariedade do processo pode ser feita por técnicas gráficas e por métodos estatísticos. Dentre esses métodos estatísticos, podemos destacar Heidelberger e Welch (1983) Gelman e Rubin (1992) e Geweke et al. (1991). Outros métodos de verificar convergência e discussões sobre MCMC, podem ser encontradas em Cowles e Carlin (1996), Gilks (2005) e Liang et al. (2011). Casella e George (1992), mostram que para uma amostra suficientemente grande, $B \longrightarrow \infty$, por meio do amostrador de Gibbs, verifica-se que as condicionais completas para determinar as marginais, isto é, quando temos uma amostra suficientemente grande de um parâmetro, dados os demais parâmetros, é possível obter uma distribuição empírica que se aproxima suficientemente da distribuição marginal.

### 2.7.3 Intervalo de Credibilidade e Intervalo de Alta Densidade *a Posteriori*

Na inferência clássica tradicionalmente ao estimar um parâmetro é obtido um intervalo de confiança para o parâmetro, entretanto na inferência Bayesiana, toda informação que se tem a sob um parâmetro $\theta$ é determinado por meio de sua distribuição *a posteriori*. A partir dessa distribuição é possível estimar algumas medidas resumo do parâmetro $\theta$, como média, mediana e moda a *posteriori*, baseados nos valores observados da distribuição a *posteriori* de $\theta$. Dessa forma, toda a informação sobre o parâmetro ficaria restrita a uma estimação pontual. Logo, torna-se necessário utilizar uma medida de precisão de tais estimativas. Para isso, utiliza-se com o conceito de intervalo de credibilidade que permite medir a precisão com que estes valores foram estimados.

O intervalo de credibilidade de $100(1-\alpha)\%$ é definido como, um intervalo $C = [a, b]$ para $\theta$ se $P[\theta \in C] \geq 1 - \alpha$, sendo $\alpha$ uma probabilidade. Pode-se notar que a definição expressa de forma probabilística a pertinência de $\theta$ ou não ao intervalo. Assim, quanto menor for a amplitude do intervalo mais concentrada é a distribuição do parâmetro, ou seja a amplitude do intervalo informa a dispersão do conhecimento *a posteriori* de $\theta$ (EHLERS, 2007).

Outro intervalo Bayesiano frequentemente utilizado é chamado de intervalo de alta densidade (HPD, do inglês *highest posterior density*). É definido como, um intervalo $C = [a, b]$ com credibilidade $100(1-\alpha)\%$, onde para todo $\theta_1 \in C$ e para todo $\theta_2 \notin C$ : $P[\theta_1|\boldsymbol{x}] \geq P[\theta_2|\boldsymbol{x}]$. Assim, o intervalo de HPD contém os valores de $\theta$ que são *a posterior*i mais plausível, ou seja, $P[\theta|\boldsymbol{x}]$ é maior em todos os $\theta$'s dentro do intervalo de HPD do que para valores fora do intervalo. Graficamente, o intervalo HPD é construído determinando a intersecção de uma linha horizontal, com a densidade *posteriori*. As coordenadas $x$ dos dois pontos de intersecção a definir o intervalo HPD são encontradas ajustando a altura da linha horizontal até que a área sobre a curva seja igual a $(1-\alpha)$. Vale ressaltar que se a distribuição *a posterior* for simétrica e unimodal, ambos os intervalos serão muitos próximos (LESAFFRE; LAWSON, 2012).

## 2.8 Critérios de Seleção

É possível utilizar diversos critérios para a seleção de modelos. Em geral, esses critérios consideram a complexidade do modelo e penalizam a verossimilhança utilizando o número de parâmetros. Alguns critérios também consideram o tamanho da amostra. Esta penalização é feita subtraindo do valor da verossimilhança uma determinada quantidade, que depende de quantos parâmetros o modelo possui.

Para a estimação por máxima verossimilhança e para estimação Bayesiana, há diferentes critérios de seleção. Os critérios de seleção considerados para comparar estimação por máxima verossimilhança são: o $AIC$ (critério de informação Akaike), introduzido por

Akaike (1974), que utiliza a seguinte quantidade

$$AIC = -2\ell\left(\boldsymbol{\theta}\right) + 2k,$$

onde $\ell\left(\boldsymbol{\theta}\right)$ é a verossimilhança do modelo ajustado, $k$ é o número de parâmetros estimados pelo modelo. Outro critério utilizado é o $AICc$ (critério de informação Akaike corrigido), que segundo Sugiura (1978), é expresso por

$$AICc = AIC + \frac{2k(k+1)}{n-k-1},$$

onde $n$ é o tamanho da amostra. Outro critério de seleção, proposto por Schwarz et al. (1978) é o $BIC$ (critério de informação Bayesiano), dado por

$$BIC = -2\ell\left(\boldsymbol{\theta}\right) + k\ln n.$$

Para comparar os modelos estimados por inferência Bayesiana será utilizado o $DIC$ (critério de informação *deviance*) proposto por Spiegelhalter et al. (2002), que pode ser considerado uma generalização do $BIC$ sob o ponto de vista Bayesiano, é dado por

$$DIC = 2D(\overline{\boldsymbol{\theta}}) - D(\widehat{\boldsymbol{\theta}})$$

onde $D(\overline{\boldsymbol{\theta}})$ é a função *deviance* da média *a posteriori* e $D(\widehat{\boldsymbol{\theta}})$ é a medida *a posteriori* da qualidade do ajuste do modelo para os dados. Outro critério de ajuste utilizado é o $EAIC$ (critério de informação de Akaike estendido) proposto por Brooks (2002). Esse critério é definido como

$$EAIC = D(\overline{\boldsymbol{\theta}}) + 2k$$

onde $k$ é o número de parâmetros estimados pelo modelo. Será ainda utilizado o $EBIC$ (critério de informação de Bayesiano estendido), proposto por Carlin e Louis (2000), e expresso por

$$EBIC = D(\overline{\boldsymbol{\theta}}) + k\ln(n)$$

onde $n$ é o tamanho da amostra.

Em todos esses critérios quanto menor o seu valor, há um indicativo de melhor ajuste do modelo. Podemos especialmente dizer que comparando modelos, considera-se uma diferença significativa entre eles quando a diferença entre seus $DIC$'s é maior que 5 (BOX; TIAO, 2011).

Um outro critério utilizado para comparar modelos Bayesianos é usar a medida da log pseudo verossimilhança marginal (LMPL, do inglês *log pseudo marginal likelihood*) e o pseudo fator de Bayes. A LMPL é obtida das estatísticas ordenadas preditivas condicionais (CPO, em inglês *conditional predictive ordinates*) (GELFAND et al., 1992). Para a $i$-ésina observação a CPO$_i$ é dada por

$$f\left(\mathbf{D}_i|\boldsymbol{y}_{[i]}\right) = \int f\left(\mathbf{D}_i|\Theta\right) f\left(\Theta|\mathbf{D}_{[i]}\right) \mathrm{d}\Theta,$$

onde $\Theta$ é o vetor completo de parâmetros, $\mathbf{D}_i$ é cada ocorrência dos dados completos **D**, $\mathbf{D}_{[i]}$ é **D** sem a presente observação $i$ e $f\left(\Theta|\mathbf{D}_{[i]}\right)$ é a densidade *posteriori* de $\Theta$ dado $f\left(\Theta|\mathbf{D}_{[i]}\right),\ i = 1, \ldots, n$. É conhecido que uma aproximação MCMC para a CPO$_i$ é dada por

$$\widehat{CPO}_i = \left[\frac{1}{B}\sum_{b=1}^{B}\frac{1}{f\left(\mathbf{D}_i|\Theta_b\right)}\right]^{-1}, i = 1, \ldots, n,$$

onde $B$ é o números de iterações do modelo MCMC implementado depois do período de queima e $\Theta_b$ é o vetor de amostra obtido na $b$-ésima iteração (CHEN et al., 2012). Para um dado modelo, o valor estimado da LPML é dado por

$$\widehat{LPML} = \sum_{i=1}^{n}\ln\widehat{CPO}_i.$$

O maior valor da LPML, é considerado o melhor ajuste de modelo (GEISSER; EDDY, 1979). O valor da LPML se assemelha muito ao valor do $\ell(\hat{\boldsymbol{\theta}}, \boldsymbol{x})$. O correspondente pseudo fator de Bayes (PFB) comparando dois modelo, $m_1$ e $m_2$ é

$$PFB_{m_1m_2} = \exp\left(\widehat{LPML}_{m1} - \widehat{LPML}_{m2}\right).$$

Outra forma de observar o ajuste do modelo é por meio de gráficos. Dois gráficos são importantes nessa verificação. No primeiro, verifica-se a curva estimada se aproximada da curva empírica de Kapler-Meier. No segundo os valores observados são comparados com os valores preditos.

# Capítulo 3

# A Distribuição Weibull Modificada

A função densidade de probabilidade de uma variável aleatória $T$, que segue uma distribuição Weibull modificada (WM) é dada por

$$f_0(t) = \alpha t^{\beta-1}(\beta + \lambda t)\exp(\lambda t - \alpha t^{\beta} e^{\lambda t}), \tag{3.1}$$

em que $t > 0$, $\alpha > 0$, $\beta > 0$ e $\lambda > 0$. A respectiva função distribuição, $F_0(t) = P(T \leq t)$, é dada por

$$F_0(t) = 1 - \exp(-\alpha t^{\beta} e^{\lambda t}), \tag{3.2}$$

dessa forma a função de sobrevivência é

$$S_0(t) = 1 - F_0(t) = \exp(-\alpha t^{\beta} e^{\lambda t}), \tag{3.3}$$

com função de risco correspondente dada por

$$h_0(t) = \frac{f_0(t)}{S_0(t)} = \alpha t^{\beta-1}(\beta + \lambda t)\exp(\lambda t). \tag{3.4}$$

Observar que:

**(a)** Quando $\lambda = 0$, a expressão (3.1) se resume à função densidade de probabilidade de uma variável aleatória que segue uma distribuição Weibull com dois parâmetros, $\alpha$ e $\beta$.

**(b)** Quando $\lambda = 0$ e $\beta = 1$, a expressão (3.1) se resume à função densidade de probabilidade de uma variável aleatória que segue uma distribuição exponencial com um único parâmetro $\alpha$.

**(c)** Quando $\lambda = 0$ e $\beta = 2$, a expressão (3.1) se resume à função densidade de probabilidade de uma variável aleatória que segue uma distribuição de Rayleigh com um único parâmetro $\alpha$.

A WM é flexível nas formas das suas funções, tanto na forma da função densidade quanto na de sobrevivência e na de risco. Sobretudo, é interessante notar que a forma da sua função de risco, pode apresentar forma de banheira ou forma decrescente, conforme os valores dos parâmetros. Conforme Lai et al. (2003), a forma da função de risco, expressa em (3.4) depende do valor de $\beta$. A forma de (3.4) é definida em $\beta$ por causa do fator, $t^{\beta-1}$. Para definir sua forma, dois casos podem ocorrem, o caso onde $\beta > 1$ e onde $0 < \beta < 1$.

Caso 1, quando $\beta > 1$:

**(a)** É perceptível que $h_0(t)$ cresce conforme o valor de $t$ aumenta. Assim a expressão (3.4) é crescente;

**(b)** Temos que $h_0(0) = 0$ se $\beta > 1$ e $h_0(0) = \alpha$ se $\beta = 1$;

**(c)** Quando $t \longrightarrow \infty$, verifica-se que $h_0(t) \longrightarrow \infty$.

Caso 2, quando $0 < \beta < 1$:

**(a)** Pode-se notar que $h_0(t)$ cresce e depois decresce conforme o valor de $t$ aumenta, indicando que a expressão (3.4) assume forma de banheira;

**(b)** Quando $t \longrightarrow 0$, temos que $h_0(t) \longrightarrow \infty$, e quando $t \longrightarrow \infty$, também observamos $h_0(t) \longrightarrow \infty$;

**(c)** A derivada de $h_0(t)$ em relação a $t$, intercepta o eixo $t$ em

$$t' = \frac{\sqrt{\beta} - \beta}{\lambda}$$

para $t > 0$. Assim $h_0(t)$ é decrescente para $t < t'$ e é crescente para $t > t'$. É interessante notar que $t'$ é inversamente proporcional a $\lambda$, assim quando $\lambda$ cresce $t'$ decresce e vice-versa.

Notemos que a forma de banheira da $h_0(t)$ só é possível devido ao novo parâmetro de forma $\lambda$, que a diferencia da distribuição Weibull. Na Figura (3), tem-se os gráficos das expressões (3.1), (3.3) e (3.4) para alguns valores de parâmetros.

## 3.1 Média e Variância

A partir da Equação (2.4) considerando a distribuição WM, pode-se observar que o tempo médio até ocorrer a falha da distribuição WM é dado por

$$\mu = \int_0^{+\infty} \exp(-\alpha t^{\beta} e^{\lambda t})\, \mathrm{d}t. \tag{3.5}$$

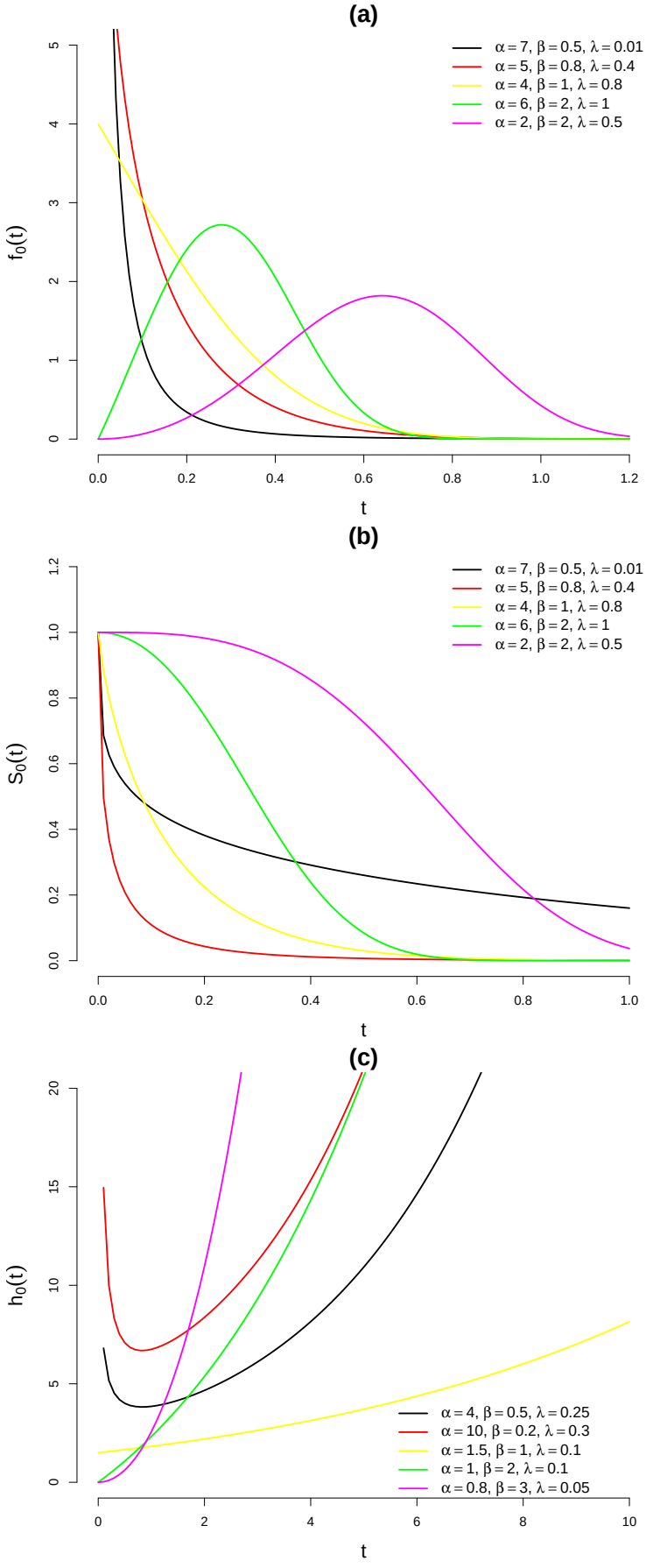


Figura 3 – Gráficos da função densidade de probabilidade (a), da função de sobrevivência (b) e da função de risco (c) da distribuição Weibull modificada.

Como já observado por Xie et al. (2002) esta integral não possui forma fechada, portanto sua solução deve ser encontrar por integração numérica. Analogamente por meio da Equação (2.5) pode-se verificar que a variância até o tempo de falha da distribuição WM pode ser expressa por

$$\mathrm{Var}(T) = 2\int_0^{+\infty} t\exp(-\alpha t^{\beta}e^{\lambda t})\,\mathrm{d}t - \mu^2. \tag{3.6}$$

Como a média (3.5), a variância até o tempo de falha também não possui forma fechada, sendo necessário uma integração numérica para encontrar sua solução.

Os momentos da distribuição WM podem serem encontrados em Nadarajah (2005).

## 3.2 Verossimilhança da Weibull Modificada

A função de verossimilhança considerando os parâmetros pertencentes ao vetor $\boldsymbol{\theta} = (\alpha, \beta, \lambda)^T$ é dada por

$$L(\boldsymbol{\theta}) = \alpha^n \prod_{i=1}^{n} t_i^{\beta-1}(\beta + \lambda t_i)\exp(\lambda t_i - \alpha t_i^{\beta}e^{\lambda t_i}).$$

Para uma amostra aleatória de tamanho $n$, o seu logaritmo é dado por

$$\ell(\boldsymbol{\theta}) = n\ln(\alpha) + (\beta-1)\sum_{i=1}^{n}\ln(t_i) + \sum_{i=1}^{n}\ln(\beta+\lambda t_i) + \lambda\sum_{i=1}^{n} t_i - \alpha\sum_{i=1}^{n} t_i^{\beta}e^{\lambda t_i},$$

onde $\ell(\boldsymbol{\theta})$ é $\ln\left(L(\boldsymbol{\theta})\right)$

A função escore para um parâmetro $\phi$ pertencente ao vetor $\boldsymbol{\theta}$ é definida por

$$U_{\phi}(\boldsymbol{\theta}) = \frac{\partial}{\partial\phi}\ell(\boldsymbol{\theta}).$$

Temos, portanto, as funções

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\ell(\boldsymbol{\theta}) &= \frac{n}{\alpha} - \sum_{i=1}^{n} e^{\lambda t_i}t_i^{\beta}, \\
\frac{\partial}{\partial\beta}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{1}{\beta+\lambda t_i} + \sum_{i=1}^{n}\ln(t_i) - \alpha\sum_{i=1}^{n} e^{\lambda t_i}t_i^{\beta}\ln(t_i), \\
\frac{\partial}{\partial\lambda}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{t_i}{\beta+\lambda t_i} + \sum_{i=1}^{n} t_i - \alpha\sum_{i=1}^{n} t_i^{\beta+1}e^{\lambda t_i}.
\end{aligned}$$

Dada uma amostra aleatória de tamanho $n$, as estimativas amostrais de máxima verossimilhança para $\alpha$, $\beta$ e $\lambda$ são dadas pelos valores de $\widehat{\alpha}$, $\widehat{\beta}$ e $\widehat{\lambda}$, respectivamente, que

satisfazem às expressões

$$\begin{aligned}
&\frac{n}{\widehat{\alpha}} - \sum_{i=1}^{n} e^{\widehat{\lambda} t_i} t_i^{\widehat{\beta}} = 0, \\
&\sum_{i=1}^{n} \frac{1}{\widehat{\beta} + \widehat{\lambda} t_i} + \sum_{i=1}^{n} \ln(t_i) - \widehat{\alpha} \sum_{i=1}^{n} e^{\widehat{\lambda} t_i} t_i^{\widehat{\beta}} \ln(t_i) = 0 \\
&\sum_{i=1}^{n} \frac{t_i}{\widehat{\beta} + \widehat{\lambda} t_i} + \sum_{i=1}^{n} t_i - \widehat{\alpha} \sum_{i=1}^{n} t_i^{\widehat{\beta}+1} e^{\widehat{\lambda} t_i} = 0.
\end{aligned}$$

Observar que, para encontrar estas estimativas de máxima verossimilhança é preciso utilizar métodos numéricos interativos. Será denotado por $U(\boldsymbol{\theta})$ um vetor de dimensão $3 \times 1$, dado por

$$U(\boldsymbol{\theta}) = \begin{bmatrix} U_\alpha(\boldsymbol{\theta}) \\ U_\beta(\boldsymbol{\theta}) \\ U_\lambda(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial \alpha}\ell \\ \frac{\partial}{\partial \beta}\ell \\ \frac{\partial}{\partial \lambda}\ell \end{bmatrix}. \tag{3.7}$$

As estimativas de máxima verossimilhança para $\alpha$, $\beta$ e $\lambda$ são obtidas por meio do método iterativo de Newton-Raphson, que se baseia na expansão da função escore $U(\boldsymbol{\theta})$ em série de Taylor em torno de um vetor de valores iniciais $\boldsymbol{\theta}^{(0)}$, de forma que

$$U(\boldsymbol{\theta}) \approx U(\boldsymbol{\theta}^{(0)}) + U'(\boldsymbol{\theta}^{(0)})(\boldsymbol{\theta} - \boldsymbol{\theta}^{(0)}),$$

em que $U'(\boldsymbol{\theta})$ é a derivada de segunda ordem de $\ell$ em relação a $\boldsymbol{\theta}^T$. Igualando $U(\boldsymbol{\theta})$ a zero, o método iterativo considera

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} + \left[-U'(\boldsymbol{\theta}^{(m)})\right]^{-1} U(\boldsymbol{\theta}^{(m)}),$$

para $m = 0, 1, 2, ...$ Observar que $\boldsymbol{\theta}^{(m)}$ é o vetor de valores encontrados para os parâmetros em $\boldsymbol{\theta}$ na $m$-ésima iteração. O método iterativo do escore de Fisher substitui $-U'(\boldsymbol{\theta})$ pela matriz de informação $-I(\boldsymbol{\theta})$, evitando a possibilidade de $-U'(\boldsymbol{\theta})$ não ser positiva definida. Assim, obtemos a aproximação

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} + \left[-I(\boldsymbol{\theta}^{(m)})\right]^{-1} U(\boldsymbol{\theta}^{(m)}).$$

A matriz de informação de Fisher $I(\boldsymbol{\theta})$ tem a forma

$$I(\boldsymbol{\theta}) = - \begin{bmatrix} \frac{\partial^2}{\partial \alpha^2}\ell(\boldsymbol{\theta}) & \frac{\partial^2}{\partial \alpha \partial \beta}\ell(\boldsymbol{\theta}) & \frac{\partial^2}{\partial \alpha \partial \lambda}\ell(\boldsymbol{\theta}) \\ \frac{\partial^2}{\partial \alpha \partial \beta}\ell(\boldsymbol{\theta}) & \frac{\partial^2}{\partial \beta^2}\ell(\boldsymbol{\theta}) & \frac{\partial}{\partial \beta \partial \lambda}\ell(\boldsymbol{\theta}) \\ \frac{\partial^2}{\partial \alpha \partial \lambda}\ell(\boldsymbol{\theta}) & \frac{\partial}{\partial \beta \partial \lambda}\ell(\boldsymbol{\theta}) & \frac{\partial^2}{\partial \lambda^2}\ell(\boldsymbol{\theta}) \end{bmatrix},$$

em que

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha^2}\ell(\boldsymbol{\theta}) &= -\frac{n}{\alpha^2}, \\
\frac{\partial^2}{\partial\beta^2}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n}\frac{1}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} e^{\lambda t_i} t_i^{\beta}\left(\ln t_i\right)^2, \\
\frac{\partial^2}{\partial\lambda^2}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n}\frac{t_i^2}{(\beta+\lambda(t_i))^2} - \alpha\sum_{i=1}^{n} t_i^{\beta+2} e^{\lambda t_i}, \\
\frac{\partial^2}{\partial\alpha\partial\beta}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n} e^{\lambda t_i} t_i^{\beta}\ln(t_i), \\
\frac{\partial^2}{\partial\alpha\partial\lambda}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n} t_i^{\beta+1} e^{\lambda t_i}, \\
\frac{\partial^2}{\partial\beta\partial\lambda}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n}\frac{t_i}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} t_i^{\beta+1} e^{\lambda t_i}\ln(t_i).
\end{aligned}$$

O erro padrão assintótico do estimador de máxima verossimilhança de um parâmetro $\theta$ é obtido da diagonal principal da inversa da matriz de informação de Fisher avaliada nos estimadores obtidos do método.

### 3.2.1 Verossimilhança Weibull modificada na presença de dados censurados

Seja $d_i$ uma observação de uma variável que indica se o tempo de sobrevida do indivíduo $i$ é censurada ou não, para $i = 1, ..., n$. Assim, para cada indivíduo, temos pares de observações $(t_i, d_i)$, tal que $d_i = 1$ indica que o evento de interesse foi observado e ocorreu no tempo $t_i$ e $d_i = 0$ indica que o indivíduo $i$ foi acompanhado até o tempo $t_i$ mas a informação sobre o instante em que o evento teria ocorrido é censurada. A função de verossimilhança é agora dada por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\left[f(t_i)\right]^{d_i}\left[S(t_i)\right]^{1-d_i}, \tag{3.8}$$

em que $S(t_i) = 1 - F(t_i)$ é a função de sobrevivência. Desprezando ainda a presença de fração de cura, das Equações (3.1) e (3.2), temos

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\left[\alpha t_i^{\beta-1}(\beta+\lambda t_i)\exp(\lambda t_i - \alpha t_i^{\beta} e^{\lambda t_i})\right]^{d_i}\left[1-\exp(-\alpha t_i^{\beta} e^{\lambda t_i})\right]^{1-d_i}.$$

O logaritmo desta expressão é dado por

$$\begin{aligned}
\ell(\boldsymbol{\theta}) &= n\ln(\alpha)\sum_{i=1}^{n} d_i + (\beta-1)\sum_{i=1}^{n} d_i\ln(t_i) + \sum_{i=1}^{n} d_i\ln(\beta+\lambda t_i) + \\
&\quad \sum_{i=1}^{n} d_i(\lambda t_i - \alpha t_i^{\beta} e^{\lambda t_i}) + \sum_{i=1}^{n}(1-d_i)\ln\left(1-\exp(-\alpha t_i^{\beta} e^{\lambda t_i})\right).
\end{aligned}$$

Derivando $\ell(\theta)$ em relação a $\alpha$, $\beta$ e $\lambda$, temos, respectivamente,

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\ell(\boldsymbol{\theta}) &= \frac{n}{\alpha}\sum_{i=1}^{n} d_i - \sum_{i=1}^{n} d_i e^{\lambda t_i} t_i^{\beta} + \sum_{i=1}^{n} \frac{(1-d_i)\, t_i^{\beta} e^{\lambda t_i} \exp\left(-\alpha e^{\lambda t_i} t_i^{\beta}\right)}{1-\exp\left(-\alpha e^{\lambda t_i} t_i^{\beta}\right)},\\
\frac{\partial}{\partial\beta}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \frac{d_i}{\beta+\lambda t_i} + \sum_{i=1}^{n} d_i \ln t_i + \alpha\sum_{i=1}^{n} \frac{(1-d_i)\exp\left(-\alpha e^{\lambda t_i} t_i^{\beta}\right) t_i^{\beta} e^{\lambda t_i}\ln(t_i)}{1-\exp\left(-\alpha e^{\lambda t_i} t_i^{\beta}\right)}\\
&\quad -\alpha\sum_{i=1}^{n} e^{\lambda t_i} d_i t_i^{\beta}\ln(t_i),\\
\frac{\partial}{\partial\lambda}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n} d_i\left(t_i - \alpha t_i e^{\lambda t_i} t_i^{\beta}\right) + \alpha\sum_{i=1}^{n} \frac{\exp\left((1-d_i) - \alpha e^{\lambda t_i} t_i^{\beta}\right) t_i^{\beta+1} e^{\lambda t_i}}{1-\exp\left(-\alpha e^{\lambda t_i} t_i^{\beta}\right)} + \sum_{i=1}^{n}\frac{d_i t_i}{\beta+\lambda t_i}.
\end{aligned}$$

As derivadas segundas, que compõem a matriz de informação de Fisher, são dadas por

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha^2}\ell(\boldsymbol{\theta}) &= -\frac{n}{\alpha^2}\sum_{i=1}^{n} d_i + \sum_{i=1}^{n}\frac{(1-d_i)\, t_i^{2\beta} A\, e^{2\lambda t_i}}{(1-A)^2},\\
\frac{\partial^2}{\partial\beta^2}\ell(\boldsymbol{\theta}) &= \alpha\sum_{i=1}^{n}\frac{(1-d_i)\,(\ln(t_i))^2 {t_i}^{\beta} e^{\lambda t_i} B}{(1-A)^2} - \sum_{i=1}^{n}\frac{d_i}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i e^{\lambda t_i} t_i^{\beta}\left(\ln(t_i)\right)^2,\\
\frac{\partial^2}{\partial\lambda^2}\ell(\boldsymbol{\theta}) &= \alpha\sum_{i=1}^{n}\frac{(1-d_i)\; e^{\lambda t_i} t_i^{\beta+2} B}{(1-A)^2} - \alpha\sum_{i=1}^{n} d_i e^{\lambda t_i} t_i^{\beta+2} - \sum_{i=1}^{n}\frac{d_i t_i^2}{(\beta+\lambda t_i)^2},\\
\frac{\partial^2}{\partial\alpha\partial\beta}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{(1-d_i)\; e^{\lambda t_i} t_i^{\beta}\ln(t_i) B}{(1-A)^2} - \sum_{i=1}^{n} d_i e^{\lambda t_i}\ln(t_i) t_i^{\beta},\\
\frac{\partial^2}{\partial\alpha\partial\lambda}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{(1-d_i)\; e^{\lambda t_i} t_i^{\beta+1} B}{(1-A)^2} - \sum_{i=1}^{n} d_i e^{\lambda t_i} t_i^{\beta+1},\\
\frac{\partial^2}{\partial\beta\partial\lambda}\ell(\boldsymbol{\theta}) &= \alpha\sum_{i=1}^{n}\frac{(1-d_i)\ln(t_i)\, {t_i}^{\beta+1} e^{\lambda t_i} B}{(1-A)^2} - \alpha\sum_{i=1}^{n} d_i e^{\lambda t_i} t_i^{\beta+1}\ln(t_i) - \sum_{i=1}^{n}\frac{d_i t_i}{(\beta+\lambda t_i)^2}.
\end{aligned}$$

onde $A = \exp\left(-\alpha e^{\lambda t_i} t_i^{\beta}\right)$ e $B = \left(1 - A e^{\lambda t_i} + A\alpha e^{\lambda t_i} {t_i}^{\beta}\right)$.

### 3.2.2 Verossimilhança Weibull modificada na presença de dados censurados e fração de cura em um modelo de mistura

Considerando a função de verossimilhança em (3.8), e substituindo nela as Equações (2.9) e (2.10), a função de verossimilhança, é dada agora por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n}\left[(1-p)\, f_0(t_i)\right]^{d_i}\left[p + (1-p)\, S_0(t_i)\right]^{1-d_i},$$

e seu logaritmo é dado por

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} d_i \ln\left[(1-p)\, f_0(t_i)\right] + \sum_{i=1}^{n}(1-d_i)\ln[p + (1-p)\, S_0(t_i)].$$

Considerando a distribuição Weibull modificada, temos

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= \ln(1-p)\sum_{i=1}^{n} d_i + (\ln\alpha)\sum_{i=1}^{n} d_i + (\beta-1)\sum_{i=1}^{n} d_i \ln t_i \\ &+ \sum_{i=1}^{n} d_i \ln(\beta+\lambda t_i) + \lambda\sum_{i=1}^{n} d_i t_i - \alpha\sum_{i=1}^{n} d_i t_i^{\beta} e^{\lambda t_i} \\ &+ \sum_{i=1}^{n}(1-d_i)\ln\left\{p+(1-p)\left[1-\exp(-\alpha t_i^{\beta}e^{\lambda t_i})\right]\right\}.\end{aligned}$$

Logo as derivadas de $\ell(\boldsymbol{\theta})$ em relação a cada parâmetro são:

$$\begin{aligned}\frac{\partial}{\partial\alpha}\ell(\boldsymbol{\theta}) &= \frac{1}{\alpha}\sum_{i=1}^{n} d_i - \sum_{i=1}^{n} d_i {t_i}^{\beta} e^{\lambda t_i} + \sum_{i=1}^{n}\frac{C{t_i}^{\beta}e^{\lambda t_i}A}{D}, \\ \frac{\partial}{\partial\beta}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{Ce^{\lambda t_i}\alpha {t_i}^{\beta}\ln(t_i)\,A}{D} + \sum_{i=1}^{n} d_i\ln(t_i) + \sum_{i=1}^{n}\frac{d_i}{\beta+\lambda t_i} - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta} e^{\lambda t_i}\ln(t_i) \\ \frac{\partial}{\partial\lambda}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{C\alpha {t_i}^{\beta+1}e^{\lambda t_i}A}{D} + \sum_{i=1}^{n}\frac{d_i t_i}{\beta+\lambda t_i} + \sum_{i=1}^{n} d_i t_i - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta+1} e^{\lambda t_i} \\ \frac{\partial}{\partial p}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{(1-d_i)\,A}{D} - \frac{1}{1-p}\sum_{i=1}^{n} d_i.\end{aligned}$$

Onde $C=(1-d_i)(1-p)$ e $D=p+(1-p)(1-A)$.

Assim as derivadas segundas são,

$$\begin{aligned}\frac{\partial^2}{\partial\alpha^2}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n}\frac{C{t_i}^{2\beta}e^{2\lambda t_i}A}{(1-A+pA)^2} - \frac{1}{\alpha^2}\sum_{i=1}^{n} d_i \\ \frac{\partial^2}{\partial\beta^2}\ell(\boldsymbol{\theta}) &= \alpha\sum_{i=1}^{n}\frac{C(\ln(t_i))^2 {t_i}^{\beta}e^{\lambda t_i}E}{(A-1+p)^2} - \sum_{i=1}^{n}\frac{d_i}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i(\ln(t_i))^2 {t_i}^{\beta}e^{\lambda t_i} \\ \frac{\partial^2}{\partial\lambda^2}\ell(\boldsymbol{\theta}) &= \alpha\sum_{i=1}^{n}\frac{C{t_i}^{\beta+2}e^{\lambda t_i}E}{(A-1+p)^2} - \sum_{i=1}^{n}\frac{d_i {t_i}^2}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta+2}e^{\lambda t_i} \\ \frac{\partial^2}{\partial p^2}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n}\frac{(1-d_i)\,A^2}{(1-A+pA)^2} - \frac{1}{(1-p)^2}\sum_{i=1}^{n} d_i \\ \frac{\partial^2}{\partial\alpha\partial\beta}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{C\ln(t_i)\,{t_i}^{\beta}e^{\lambda t_i}E}{(A-1+p)^2} - \sum_{i=1}^{n} d_i {t_i}^{\beta}\ln(t_i)\,e^{\lambda t_i}\end{aligned}$$

$$\begin{aligned}\frac{\partial^2}{\partial\alpha\partial\lambda}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\frac{Ce^{\lambda t_i}{t_i}^{\beta+1}E}{(A-1+p)^2} - \sum_{i=1}^{n} d_i {t_i}^{\beta+1}e^{\lambda t_i} \\ \frac{\partial^2}{\partial\alpha\partial p}\ell(\boldsymbol{\theta}) &= -\sum_{i=1}^{n}\frac{(1-d_i)\,{t_i}^{\beta}e^{\lambda t_i}A}{(1-A+pA)^2} \\ \frac{\partial^2}{\partial\beta\partial\lambda}\ell(\boldsymbol{\theta}) &= \alpha\sum_{i=1}^{n}\frac{C\ln(t_i)\,{t_i}^{\beta+1}e^{\lambda t_i}E}{(A-1+p)^2} - \sum_{i=1}^{n}\frac{d_i t_i}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta+1}\ln(t_i)\,e^{\lambda t_i} \\ \frac{\partial^2}{\partial\beta\partial p}\ell(\boldsymbol{\theta}) &= -\alpha\sum_{i=1}^{n}\frac{(1-d_i)\,{t_i}^{\beta}\ln(t_i)\,e^{\lambda t_i}A}{(1-A+pA)^2} \\ \frac{\partial^2}{\partial\lambda\partial p}\ell(\boldsymbol{\theta}) &= -\alpha\sum_{i=1}^{n}\frac{(1-d_i)\,{t_i}^{\beta+1}e^{\lambda t_i}A}{(1-A+pA)^2}\end{aligned}$$

Onde $E=A-1+p-Ae^{\lambda t_i}\alpha {t_i}^{\beta}$.

### 3.2.3 Verossimilhança Weibull modificada na presença de dados censurados e fração de cura em um modelo de não-mistura

Sendo que a função de máxima verossimilhança para dados censurados é dada em (3.8) e considerando as Equações (2.13) e (2.14), temos que a função de verossimilhança, nesse novo caso é dada por

$$\begin{aligned} L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} \left[h_0(t_i)\right]^{d_i} S_0(t_i) \\ &= \prod_{i=1}^{n} \left[-\ln(p) f_0(t_i) \exp(\ln(p) F_0(t_i))\right]^{d_i} \left[\exp[\ln(p) F_0(t_i)]\right]^{1-d_i} \\ &= \prod_{i=1}^{n} \left[-\ln(p) f_0(t_i)\right]^{d_i} \left[\exp[\ln(p) F_0(t_i)]\right] \end{aligned} \tag{3.9}$$

e seu logaritmo é dado por

$$\begin{aligned} \ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n} d_i \ln\left(h_0(t_i)\right) + \sum_{i=1}^{n} \ln\left(S_0(t_i)\right) \\ &= \ln\left(-\ln(p)\right) \sum_{i=1}^{n} d_i \ln\left(f_0(t_i)\right) + \ln(p) \sum_{i=1}^{n} F_0(t_i). \\ &= \ln\left(-\ln(p)\right) \sum_{i=1}^{n} d_i + \ln(\alpha) \sum_{i=1}^{n} d_i + (\beta - 1) \sum_{i=1}^{n} d_i \ln(t_i) \\ &\quad + \sum_{i=1}^{n} d_i \ln(\beta + \lambda t_i) + \lambda \sum_{i=1}^{n} d_i t_i - \alpha \sum_{i=1}^{n} d_i t_i^{\beta} e^{\lambda t_i} \\ &\quad + \ln(p) \sum_{i=1}^{n} \left(1 - \exp(-\alpha t_i^{\beta} e^{\lambda t_i})\right). \end{aligned}$$

As primeiras derivadas são:

$$\begin{aligned} \frac{\partial}{\partial \alpha}\ell(\boldsymbol{\theta}) &= \frac{1}{\alpha}\sum_{i=1}^{n} d_i - \sum_{i=1}^{n} d_i {t_i}^{\beta} e^{\lambda t_i} + \ln(p) \sum_{i=1}^{n} {t_i}^{\beta} e^{\lambda t_i} A \\ \frac{\partial}{\partial \beta}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n} d_i \ln(t_i) + \sum_{i=1}^{n} \frac{d_i}{\beta + \lambda t_i} - \alpha \sum_{i=1}^{n} d_i {t_i}^{\beta} e^{\lambda t_i} \ln(t_i) + \ln(p) \sum_{i=1}^{n} \alpha {t_i}^{\beta} e^{\lambda t_i} \ln(t_i) A \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial \lambda}\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \frac{d_i t_i}{\beta + \lambda t_i} + \sum_{i=1}^{n} d_i t_i - \alpha \sum_{i=1}^{n} d_i {t_i}^{\beta+1} e^{\lambda t_i} + \ln(p) \sum_{i=1}^{n} \alpha {t_i}^{\beta+1} e^{\lambda t_i} A \\ \frac{\partial}{\partial p}\ell(\boldsymbol{\theta}) &= \frac{1}{p \ln(p)} \sum_{i=1}^{n} d_i + \frac{1}{p}\left(n - \sum_{i=1}^{n} A\right). \end{aligned}$$

Por sua vez, as segundas derivadas são

$$\begin{aligned}
\frac{\partial^2}{\partial\alpha^2}\ell(\theta) &= -\ln(p)\sum_{i=1}^{n}\left({t_i}^{\beta}e^{\lambda t_i}\right)^2 A - \frac{1}{\alpha^2}\sum_{i=1}^{n} d_i \\
\frac{\partial^2}{\partial\beta^2}\ell(\theta) &= \alpha\ln(p)\sum_{i=1}^{n} A(\ln(t_i))^2 {t_i}^{\beta}e^{\lambda t_i}F - \sum_{i=1}^{n}\frac{d_i}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta}(\ln(t_i))^2 e^{\lambda t_i} \\
\frac{\partial^2}{\partial\lambda^2}\ell(\theta) &= \alpha\ln(p)\sum_{i=1}^{n} A{t_i}^{\beta+2}e^{\lambda t_i}F - \sum_{i=1}^{n}\frac{d_i {t_i}^2}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta}e^{\lambda t_i}{t_i}^2 \\
\frac{\partial^2}{\partial p^2}\ell(\theta) &= -\frac{1}{p}\sum_{i=1}^{n}\left({t_i}^{\beta}e^{\lambda t_i}\right)^2 A \\
\frac{\partial^2}{\partial\alpha\partial\beta}\ell(\theta) &= \ln(p)\sum_{i=1}^{n} A\ln(t_i)\,{t_i}^{\beta}e^{\lambda t_i}F - \sum_{i=1}^{n} d_i {t_i}^{\beta}e^{\lambda t_i}\ln(t_i) \\
\frac{\partial^2}{\partial\alpha\partial\lambda}\ell(\theta) &= \ln(p)\sum_{i=1}^{n} A{t_i}^{\beta+1}e^{\lambda t_i}F - \sum_{i=1}^{n} d_i {t_i}^{\beta+1}e^{\lambda t_i} \\
\frac{\partial^2}{\partial\alpha\partial p}\ell(\theta) &= \frac{1}{p}\sum_{i=1}^{n} {t_i}^{\beta}e^{\lambda t_i}A \\
\frac{\partial^2}{\partial\beta\partial\lambda}\ell(\theta) &= \alpha\ln(p)\sum_{i=1}^{n} A\ln(t_i)\,{t_i}^{\beta+1}e^{\lambda t_i}F - \sum_{i=1}^{n}\frac{d_i t_i}{(\beta+\lambda t_i)^2} - \alpha\sum_{i=1}^{n} d_i {t_i}^{\beta+1}\ln(t_i)\,e^{\lambda t_i} \\
\frac{\partial^2}{\partial\beta\partial p}\ell(\theta) &= \frac{1}{p}\sum_{i=1}^{n}\alpha {t_i}^{\beta}\ln(t_i)\,e^{\lambda t_i}A
\end{aligned}$$

e

$$\frac{\partial^2}{\partial\lambda\partial p}\ell(\theta) = \frac{1}{p}\sum_{i=1}^{n}\alpha {t_i}^{\beta+1}e^{\lambda t_i}A$$

Onde $A = \exp\left(-\alpha e^{\lambda t_i}t_i^{\beta}\right)$ e $F = \left(1-\alpha e^{\lambda t_i}{t_i}^{\beta}\right)$.

## 3.3 Regressão com Ligação log-log Complementar no Parâmetro $p$

Para introduzir uma regressão no parâmetro $p$, com o objetivo de comparar dois grupos, que podem ser, por exemplo, dois tratamentos, utiliza-se uma função de ligação. Nesse caso, temos que

$$\begin{cases} x = 1, \text{ para os tratados,} \\ x = 0, \text{ para os não tratados.} \end{cases}$$

Se uma função de ligação complementar log-log foi utilizada no parâmetro de cura, $p$, na aplicação da distribuição WM à dados reais, temos então que o parâmetro $p$ depende de outras variáveis ($p = g(\gamma_0, \gamma_1|x)$), portando para se obter uma estimativa do intervalo de

confiança para $p$ podemos utilizar o método delta. Assim,

$$\begin{aligned} g(\gamma_0, \gamma_1|x) &= 1 - \exp\left(-\exp(\gamma_0 + \gamma_1\, x)\right) \\ \frac{\partial}{\partial\gamma_0} g(\gamma_0, \gamma_1|x) &= \exp(\gamma_0 + \gamma_1\, x)\exp\left(-\exp(\gamma_0 + \gamma_1\, x)\right) \\ \frac{\partial}{\partial\gamma_1} g(\gamma_0, \gamma_1|x) &= x\exp(\gamma_0 + \gamma_1\, x)\exp\left(-\exp(\gamma_0 + \gamma_1\, x)\right) \end{aligned}$$

Logo,

$$\mathsf{Var}\left[g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right] \approx \left[\frac{\partial}{\partial\gamma_0} g(\gamma_0, \gamma_1|x) \;\; \frac{\partial}{\partial\gamma_1} g(\gamma_0, \gamma_1|x)\right] \widehat{\sum}\left[g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right] \begin{bmatrix} \frac{\partial}{\partial\gamma_0} g(\gamma_0, \gamma_1|x) \\ \frac{\partial}{\partial\gamma_1} g(\gamma_0, \gamma_1|x) \end{bmatrix}$$

onde

$$\widehat{\sum}\left[g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right] = \begin{bmatrix} \mathsf{Var}(\hat{\hat{\gamma_0}}) & \mathsf{cov}(\hat{\hat{\gamma_0}}, \hat{\gamma_1}) \\ \mathsf{cov}(\hat{\hat{\gamma_0}}, \hat{\hat{\gamma_1}}) & \mathsf{Var}(\hat{\hat{\gamma_1}}) \end{bmatrix}.$$

Notar então que,

$$\begin{aligned} p_0 &= g(\hat{\gamma_0}, \hat{\gamma_1}|x = 0) = 1 - \exp\left(-\exp(\gamma_0)\right) \\ p_1 &= g(\hat{\gamma_0}, \hat{\gamma_1}|x = 1) = 1 - \exp\left(-\exp(\gamma_0 + \gamma_1)\right). \end{aligned}$$

Nesse caso, $p_0$ e $p_1$ são respectivamente as frações de cura para os não tratados e para os tratados, $\gamma_1$ é interpretado como o efeito do tratamento sobre a fração de cura. Se a estimativa do parâmetro $\gamma_1$ foi significativa, pode-se dizer que o efeito do tratamento sobre a fração de cura é significativo.

É possível escrever o método delta na forma vetorial. Neste caso, temos que,

$$\begin{aligned} \mathsf{Var}\left[g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right] \approx\; & \mathsf{Var}(\hat{\gamma_0}) \left(\frac{\partial}{\partial\hat{\gamma_0}} g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right)^2 + \mathsf{Var}(\hat{\gamma_1}) \left(\frac{\partial}{\partial\hat{\gamma_1}} g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right)^2 \\ & + 2\mathsf{cov}(\hat{\gamma_0}, \hat{\gamma_1}) \left(\frac{\partial}{\partial\hat{\gamma_0}} g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right) \left(\frac{\partial}{\partial\hat{\gamma_1}} g(\hat{\gamma_0}, \hat{\gamma_1}|x)\right). \end{aligned}$$

No capítulo a seguir é feita uma aplicação da distribuição Weibull modificada a aplicada à dados reais de câncer gástrico na presença de fração de cura e censura.

# Capítulo 4

# Aplicação Weibull Modificada

## 4.1 Os Dados

É de conhecimento que o câncer gástrico é uma das principais causas de mortes relacionada a neoplasia (DICKEN et al., 2005) e a cirurgia de ressecção da mucosa é aceita como uma opção de tratamento para casos precoces da doença. Entre janeiro de 2002 e dezembro de 2007, Jácome et al. (2013) realizaram um estudo retrospectivo em pacientes com adenocarcinoma gástrico submetidos à resseção curativa com linfadenectomia D2 no Hospital de Câncer de Barretos (Barretos, São Paulo, Brasil). Foram acompanhados nesse estudo 201 pacientes. Dentre esses, 125 pacientes fizeram apenas a cirurgia de ressecção e os outros 76 pacientes, além da cirurgia, tiveram um tratamento complementar de quimioradioterapia (CRT). Desses pacientes 53,2% tiveram dados censurados, sendo, 57,9% tratados com cirurgia e CRT e 50,4% tratados apenas com cirurgia (MARTINEZ et al., 2013). A eficácia da linfadenectomia D2 para a cura em pacientes com início de câncer gástrico é discutida por Okamura et al. (1988).

As estimativas de Kaplan-Meier da função de sobrevivência para os dados de câncer gástrico são apresentado na Figura 4, onde é possível notar a presença de um platô que intercepta o eixo das ordenadas no valor de, aproximadamente, 0,5 no gráfico apresentado em (a), sugerindo assim que os modelos que não incluem a proporção $p$ de sobreviventes a longo prazo não são adequados para estes dados. O gráfico apresentado no painel (b) da Figura 4 descreve as funções de sobrevivência empíricas para cada um dos tipos de tratamento, onde ainda é possível observar a presença de planaltos na cauda direita da curva de sobrevivência estimada de Kaplan-Meier, o que sugere a adequação de um modelo que inclui um parâmetro que representa a fração de cura.

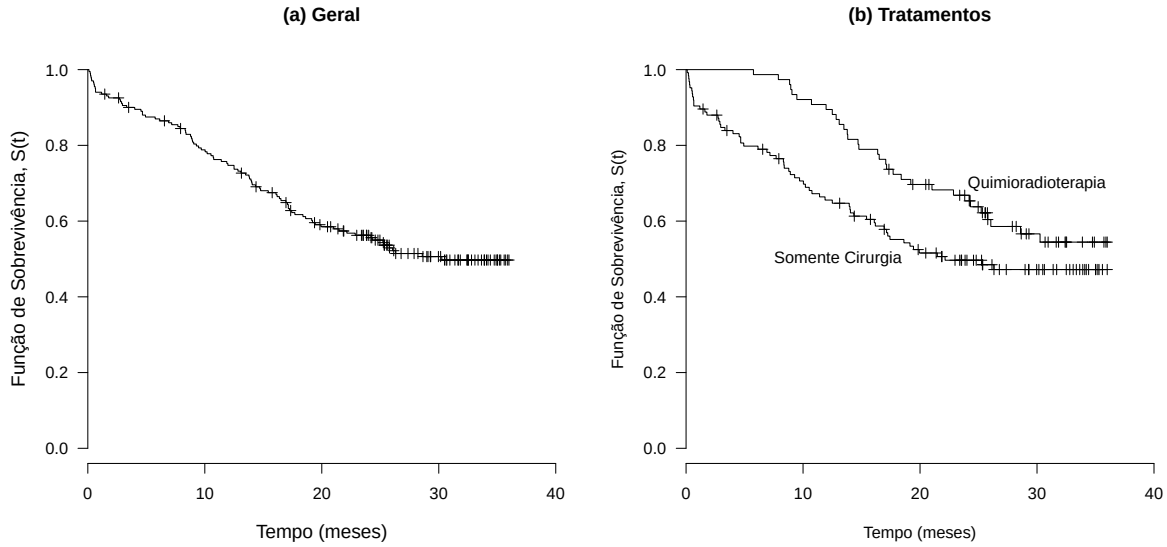


Figura 4 – Estimadores Kaplan-Meier para as função de sobrevivência dos dados de câncer gástrico dos pacientes em geral e para cada tratamento.

## 4.2 Métodos

Na aplicação da distribuição Weibull modificada para a estimação e inferência foram utilizado os *softwares*: *R* (R Core Team, 2014) e *OpenBUGS* (LUNN et al., 2000). A inferência clássica foi feita no *R* versão $3.1.2$, utilizando a função *MaxLik* do pacote de mesmo nome, descrita em Henningsen e Toomet (2011) e a inferência Bayesiana foi realizada no *OpenBUGS* versão $3.2.3$. Os diagnósticos de convergência do algoritmo MCMC das cadeias Bayesianas, foram obtidas no *R*.

O método numérico utilizado na função *maxLik* foi o de Newton-Rapshon, método esse utilizado com mais frequência para estimar os parâmetros por meio da máxima verossimilhança. A matriz de covariância, necessária para o uso do método delta, é obtida por meio da matriz Hessiana, que a própria função *maxLik* calcula.

Quando considerada a covariável no modelo, utilizou-se no parâmetro $\beta$, que é estritamente positivo, a função de ligação logarítmica, apresentada na Seção 2.5. Para o parâmetro $p$ a função de ligação utilizado foi a log-log complementar, devido esse parâmetro ser restrito ao intervalo $(0, 1)$.

Nesse trabalho foi considerado uma distribuição *a priori* gama, aos parâmetros $\alpha, \beta$ e$\lambda$, levando em consideração que eles são reais positivos. Entretanto, para o parâmetro de cura $p$, assumimos uma distribuição *a priori* beta, devido ele ser restrito ao intervalo $(0, 1)$. Então, temos que $\alpha \sim Gama(a_\alpha, b_\alpha)$ , $\beta \sim Gama(a_\beta, b_\beta)$ , $\lambda \sim Gama(a_\lambda, b_\lambda)$ e $p \sim Beta(c_p, d_p)$, sendo que $a_\alpha$ , $b_\alpha$ , $a_\beta$ , $b_\beta$ , $a_\lambda$ e $b_\lambda$ são os hiperparâmetros conhecidos da distribuição $Gama(a, b)$ com média $a/b$ e variância $a/b^2$ e $c_p$ e $d_p$ são os hiperparâmetros

da distribuição $Beta(c, d)$, com média $c/(c + d)$ e variância $cd/[(c + d)^2(c + d + 1)]$, onde consideramos $a_\alpha = b_\alpha = a_\beta = b_\beta = a_\lambda = b_\lambda = c_p = d_p = 1$, dessa forma tem-se *a priori* consideradas poucos informativas. Isso para a WM, para suas distribuições particulares, consideramos distribuições *a priori* apenas nos parâmetros não fixos.

Na presença de covariáveis, assumimos uma distribuição normal *a priori*, $N(e, f^2)$, com média $e$ e variância $f^2$, para cada parâmetro da função de ligação. Segundo a parametrização do *OpenBUGS*, considerou-se $e = 10^{-3}$ e $f^2 = 10^3$. Para todas as distribuições assumimos independência *a priori* entre os parâmetros incluídos no modelo. As distribuições *a posteriori* de interesse são obtidas a partir de amostras simuladas para a distribuição conjunta *a posteriori* usando procedimentos MCMC. Foi gerado para cada distribuição, 1.000.000 de amostras para cada parâmetro de interesse, sendo as primeiras 10.000 amostras descartadas a fim de minimizar os efeitos dos valores iniciais. Então as medidas *a posteriori* foram baseadas em 10.000 amostras, pois foi tomada uma amostra a cada 100 (salto de 100), obtendo amostras praticamente não correlacionadas. A convergência do algorítimo MCMC foi monitorada observando as séries temporais habituais para as amostras simuladas e as cadeias foram exportadas para o *R* onde a convergência foi verificada, por meio da função *heidel.diag(·)* e *geweke.diag(·)* do pacote *coda*, que aplica o teste de Heidelberger e Welch (1983) e o de Geweke et al. (1991), sendo a hipótese nula desses testes é a convergência da cadeia. Mais detalhes sobre o pacote podem ser encontrados em (PLUMMER et al., 2006) e Plummer et al. (2015).

## 4.3 Resultados

Considerando os dados de câncer gástrico introduzidos por Jácome et al. (2013) e disponibilizados em Martinez et al. (2013), a Tabela 1 apresenta as estimativas de máxima verossimilhança para os parâmetros de um modelo baseado na distribuição Weibull modificada (WM) e as distribuições que são casos especiais, sem considerar a presença de covariáveis. Estes modelos consideram uma função de sobrevivência baseada na equação (2.9) (modelo de mistura). Observa-se que o modelo baseado na distribuição WM apresentou o menor valor de $AIC, AICc$ e $BIC$ e ainda o maior valor do $\ell(\boldsymbol{\theta}, \boldsymbol{x})$, o que sugere uma melhor adequação aos dados. Além disso, a Figura 5, que mostra o gráfico TTT-plot para os dados gerais, indica que a função de risco para os dados é do tipo banheira (convexo e depois côncava), levando a considerar então que a distribuição WM nesse caso é a melhor para esses dados, devido ela comportar a forma banheira para a função de risco.

Em adição, as estimativas para a fração de cura $p$ considerando os modelos baseados nas distribuições Weibull e exponencial são, respectivamente $0,2788$ e $0,3622$, enquanto as estimativas obtidas pelos modelos baseados nas distribuições WM e Rayleigh são, respectivamente, $0,4965$ e $0,5$. Estas últimas estimativas parecem mais próximas

Tabela 1 – Estimativas frequentistas dos parâmetros por máxima verossimilhança assumindo um modelo de mistura, sem covariáveis.

| Modelos | | Estimativas (Erro Padrão) | Intervalos de Confiança $95\%$ | | | Critérios de Seleção | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Assintótico | Perfilado | A$^a$ | $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ | AIC | AICc | BIC |
| WM | $\alpha$ | 0,0772 (0,0219) | (0,0282;0,1261) | (0,0602;0,0965) | 6,36 | -441,8 | 891,6 | 891,8 | 904,8 |
| | $\beta$ | 0,5727 (0,1369) | (0,3043;0,8410) | (0,4894;0,6482) | -4,92 | | | | |
| | $\lambda$ | 0,0659 (0,0158) | (0,0349;0,0969) | (0,0527;0,0770) | -9,01 | | | | |
| | $p$ | 0,4965 (0,0383) | (0,4214;0,5716) | (0,4234;0,5691) | -0,32 | | | | |
| Weibull | $\alpha$ | 0,0502 (0,0158) | (0,0192;0,0812) | (0,0392;0,0641) | 11,74 | -447,1 | 900,2 | 900,3 | 910,1 |
| | $\beta$ | 0,9081 (0,1458) | (0,6223;1.1939) | (0,8328;0,9863) | 1.83 | | | | |
| | $p$ | 0,2788 (0,2254) | (-0,1629;0,7206) | (0,1747;0,3826) | -,.17 | | | | |
| Exponencial | $\alpha$ | 0,0472 (0,0118) | (0,0240;0,0703) | (0,0366;0,0610) | 13,14 | -447,3 | 898,6 | 898,7 | 905,2 |
| | $p$ | 0,3622 (0,0849) | (0,1958;0,5286) | (0,2749;0,4492) | -0,19 | | | | |
| Rayleigh | $\alpha$ | 0,0046 (0,0006) | (0,0034;0,0058) | (0,0004;0,017) | 49,70 | -473,4 | 950,8 | 950,9 | 957,4 |
| | $p$ | 0,5000 (0,0376) | (0,4263;0,5737) | (0,0678;0,918) | -1,69 | | | | |

$^a$ Medida de Assimetria para Intervalo de Confiança Perfilado.

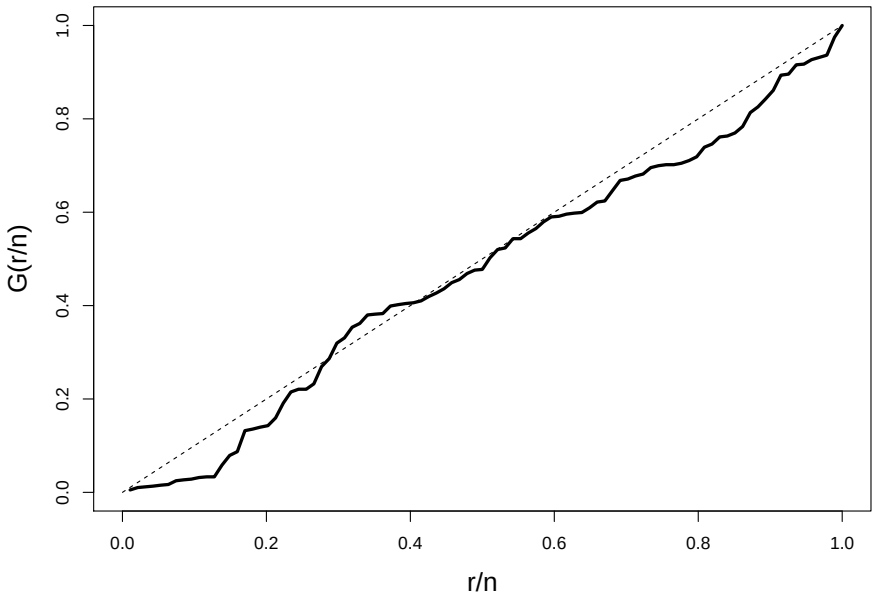


Figura 5 – Gráfico TTT-plot dos dados de câncer gástrico dos pacientes em geral.

daquelas sugeridas pelo gráfico da Figura 4 (painel esquerdo), em que a curva descreve um platô em um valor próximo de 0,5 para a função de sobrevivência. E ainda sobre a distribuição Weibull se for considerado um intervalo assintótico para a fração de cura, obtêm-se um limites negativos para o intervalo, não condizendo com o espaço paramétrico de $p$. Contudo, o intervalo de confiança perfilado é mais adequado, pois seu limite inferior está dentro das restrições do parâmetro.

Nota-se que as medidas de assimetria dos intervalos de confiança perfilados (A) é

em geral, muito altos, com valores maiores que 1. Assim, o intervalo assintótico não é o adequado a ser usado nesta aplicação. Observando-se ainda os intervalos de confiança, pode-se notar que eles não incluem os valores $\beta = 1,\ \beta = 2$ e $\lambda = 0$ na distribuição WM, fortalecendo a premissa que os casos particulares dessa distribuição não são adequados a estes dados. E além disso, os intervalos de confiança para $\beta$ incluem o valor de $\beta = 1$, no ajuste da Weibull, sugerindo que a distribuição exponencial seria uma boa opção, mas este o que não é o caso.

A Tabela 2 exibe as estimativas Bayesianas do modelo baseado na distribuição WM e seus casos específicos, também considerando que a função de sobrevivência tem uma forma de misturas. Observa-se que as estimativas Bayesianas são razoavelmente próximas às estimativas frequentistas, descritas na Tabela 1. Considerando os valores de $DIC,\ EAIC$ e LMPL, é novamente sugerido que o modelo baseado na distribuição WM é o mais adequado aos dados, dentre as distribuições consideradas. O modelo baseado na distribuição exponencial é o que apresenta menor valor de $EBIC$. Contudo, o valor estimado para $p$ não corresponde ao que é sugerido na Figura 4. Dessa forma, o modelo baseado na distribuição WM prevalece como o mais adequado, dado que ele parece estimar melhor o valor da fração de cura $p$ e sendo o único capaz de comportar a forma de banheira da função de risco.

Tabela 2 – Estimativas Bayesianas dos parâmetros assumindo um modelo de mistura, sem covariáveis.

| Modelos | Parâmetros | Médias posteriori | Intervalos HPD $95\%$ | Critérios de Seleção | | | | Diagnóstico de Convergencia | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | LMPL | DIC | EAIC | EBIC | p-valor HW | p-valor Geweke |
| WM | $\alpha$ | 0,0833 | (0,0423;0,1281 | -442,3 | 891,7 | 895,7 | 908,9 | 0,740 | 0,962 |
| | $\beta$ | 0,5658 | (0,3217;0,8382) | | | | | 0,542 | 0,400 |
| | $\lambda$ | 0,0635 | (0,0298;0,0946) | | | | | 0,559 | 0,776 |
| | $p$ | 0,4913 | (0,4099;0,5690) | | | | | 0,345 | 0,084 |
| Weibull | $\alpha$ | 0,0540 | (0,0271;0,0844) | -447,1 | 898,6 | 902,7 | 912,5 | 0,581 | 0,769 |
| | $\beta$ | 0,8943 | (0,6981;1,1130) | | | | | 0,868 | 0,434 |
| | $p$ | 0,2560 | (0.0002;0,4469) | | | | | 0,748 | 0,865 |
| Exponencial | $\alpha$ | 0,0469 | (0,0003;0.0430) | -447,4 | 898,2 | 900,5 | 907,1 | 0,501 | 0,693 |
| | $p$ | 0,3387 | (0,4186;0,5684) | | | | | 0,837 | 0,507 |
| Rayleigh | $\alpha$ | 0,0046 | (0,0028;0,0048) | -459,4 | 950,7 | 952,8 | 954,1 | 0,106 | 0,150 |
| | $p$ | 0,4992 | (0,0087;0,5379) | | | | | 0,142 | 0,654 |

Nas estimativas Bayesianas o intervalo HPD 95% para a fração de cura baseado na distribuição Weibull é mais adequado do que o apresentado na Tabela 1 se considerarmos o intervalo assintótico, pois o intervalo de confiança assintótico para o valor de $p$ extrapola seu espaço paramétrico (limite inferior negativo) ao contrario do intervalo HPD Bayesiano, que não extrapola. Isto evidencia que o intervalo Bayesiano é mais plausível para as inferências deste parâmetro, comparado aos intervalos de confiança assintóticos.

As Tabelas 3 e 4 exibem, respectivamente, estimativas frequentistas e Bayesianas de um modelo em que a função de sobrevivência tem a forma expressa pela equação (2.12), ou seja, modelos de não-mistura. Observa-se que as estimativas frequentistas (Tabela 3) são razoavelmente próximas das estimativas Bayesianas (Tabela 4). Pode-se dizer novamente, que o modelo baseado na distribuição WM é o que melhor se ajusta aos dados, considerando os critérios de seleção utilizados.

Tabela 3 – Estimativas frequentistas dos parâmetros por máxima verossimilhança assumindo um modelo de não mistura, sem covariável.

| Modelos | | Estimativas (Erro Padrão) | Intervalos de Confiança $95\%$ | | | Critérios de Seleção | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Assintótico | Perfilado | A[a] | $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ | AIC | AICc | BIC |
| WM | $\alpha$ | 0,0550 (0,0160) | (0,0236;0,0864) | (0,0602;0,0965) | 7,71 | -441,9 | 891,7 | 891,9 | 904,9 |
| | $\beta$ | 0,5653 (0,1372) | (0,2964;0,8342) | (0,4894;0,6482) | -4,62 | | | | |
| | $\lambda$ | 0,0746 (0,0166) | (0,0420;0,1072) | (0,0527;0,0770) | -9,59 | | | | |
| | $p$ | 0,4967 (0,0383) | (0,4216;0,5718) | (0,4234;0,5691) | 0,28 | | | | |
| Weibull | $\alpha$ | 0,0232 (0,0158) | (-0,0077;0,0542) | (0,0018;0,0298) | -52,86 | -447,4 | 900,8 | 900,9 | 910,7 |
| | $\beta$ | 0,9425 (0,1424) | (0,6632;1,2218) | (0,8578;1,0153) | -7,04 | | | | |
| | $p$ | 0,2242 (0,2251) | (-0,2170;0,6654) | (0,1568,0,2904) | -0,89 | | | | |
| Exponencial | $\alpha$ | 0,0236 (0,0119) | (0,0003;0,0430) | (0,0183;0,0303) | 12,47 | -447,6 | 899,2 | 899,3 | 905,8 |
| | $p$ | 0,2733 (0,1350) | (0,0087;0,5379) | (0,2202;0,3444) | 5,99 | | | | |
| Rayleigh | $\alpha$ | 0,0038 (0,0005) | (0,0028;0,0048) | (0,0004;0,017) | 53,14 | -471,8 | 947,5 | 947,6 | 954,1 |
| | $p$ | 0,4935 (0,0382) | (0,4186;0,5684) | (0,0678;0,918) | 1,99 | | | | |

[a] Medida de Assimetria para Intervalo de Confiança Perfilado.

Tabela 4 – Estimativas Bayesianas dos parâmetros assumindo um modelo de não mistura, sem covariável.

| Modelos | Parâmetros | Médias posteriori | Intervalos HPD $95\%$ | Critérios de Seleção | | | | Diagnóstico de Convergência | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | LMPL | DIC | EAIC | EBIC | p-valor HW | p-valor Geweke |
| WM | $\alpha$ | 0,0593 | (0,0298;0,0943) | -442,3 | 891,8 | 895,8 | 909,0 | 0,077 | 0,388 |
| | $\beta$ | 0,5599 | (0,3175;0,8437) | | | | | 0,324 | 0,419 |
| | $\lambda$ | 0,0717 | (0,0369;0,1065) | | | | | 0,429 | 0,196 |
| | $p$ | 0,4911 | (0,4084;0,5702) | | | | | 0,089 | 0,384 |
| Weibull | $\alpha$ | 0,0293 | (0,0080;0,0489) | -447,5 | 899,6 | 903,4 | 913,3 | 0,143 | 0,282 |
| | $\beta$ | 0,9333 | (0,7355;1,1510) | | | | | 0,133 | 0,344 |
| | $p$ | 0,2445 | (0.0170;0,4453) | | | | | 0,383 | 0,028 |
| Exponencial | $\alpha$ | 0,0269 | (0,0027;0.0485) | -447,7 | 898,5 | 901,0 | 907,6 | 0,931 | 0,949 |
| | $p$ | 0,2995 | (0,4181;0,5685) | | | | | 0,545 | 0,473 |
| Rayleigh | $\alpha$ | 0,0028 | (0,0027;0,0494) | -464,7 | 947,5 | 949,5 | 956,1 | 0,433 | 0,817 |
| | $p$ | 0,4929 | (0,4191;0,5660) | | | | | 0,295 | 0,954 |

Na Tabela 3, observa-se que os intervalos de confiança 95% para $\alpha$ e para $p$ considerando o modelo baseado na distribuição de Weibull apresentam valores negativos para os seus limites inferiores, se considerado o intervalo assintótico. Os intervalos de credibilidade

Bayesianos e os intervalos de confiança perfilado, que não se baseiam em propriedades assintóticas, são mais adequados. Observe-se ainda que a medida de assimetria mostra, que nesse caso, há em geral também uma assimetria muito grande nos intervalos de confiança perfilados.

Os gráficos da Figura 6 comparam os valores preditos dos modelos Bayesianos de mistura, baseados na distribuição WM, Weibull, exponencial e Rayleigh. Estas figuras novamente sugerem uma melhor adequação do modelo baseado na distribuição WM aos dados, dado que os valores preditos por este modelo são aqueles visualmente mais próximos às estimativas empíricas de Kaplan-Meier.

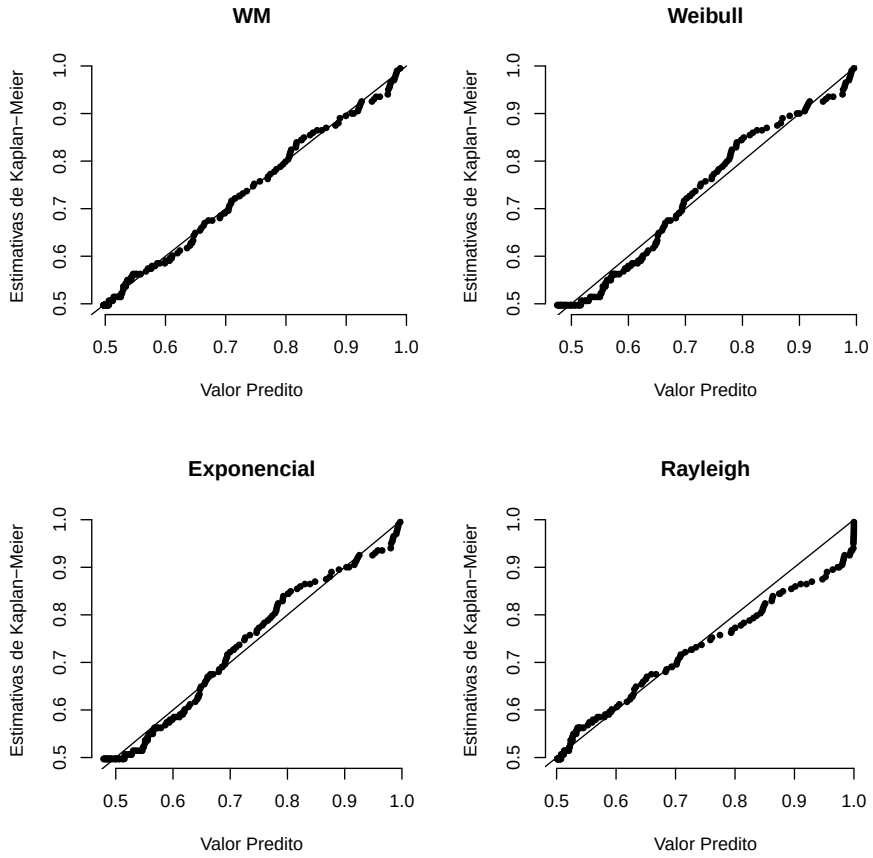


Figura 6 – Gráficos dos estimadores de Kaplan–Meier para as funções de sobrevivência versus os respectivos valores preditos, das estimativas bayesianas do modelo com mistura, considerando a distribuição Weibull modificada e seus casos particulares.

A Figura 7 apresenta as curvas de sobrevida obtidas por meio das estimativas do modelo Bayesiano com mistura, para cada uma das quatro distribuições estudadas, sobrepostas às curvas de Kaplan-Meier. Ainda que o parâmetro que descreve a fração de cura esteja presente em todos estes modelos, nota-se que as curvas obtidas dos modelos baseados nas distribuições de Weibull e exponencial não acompanham o platô descrito

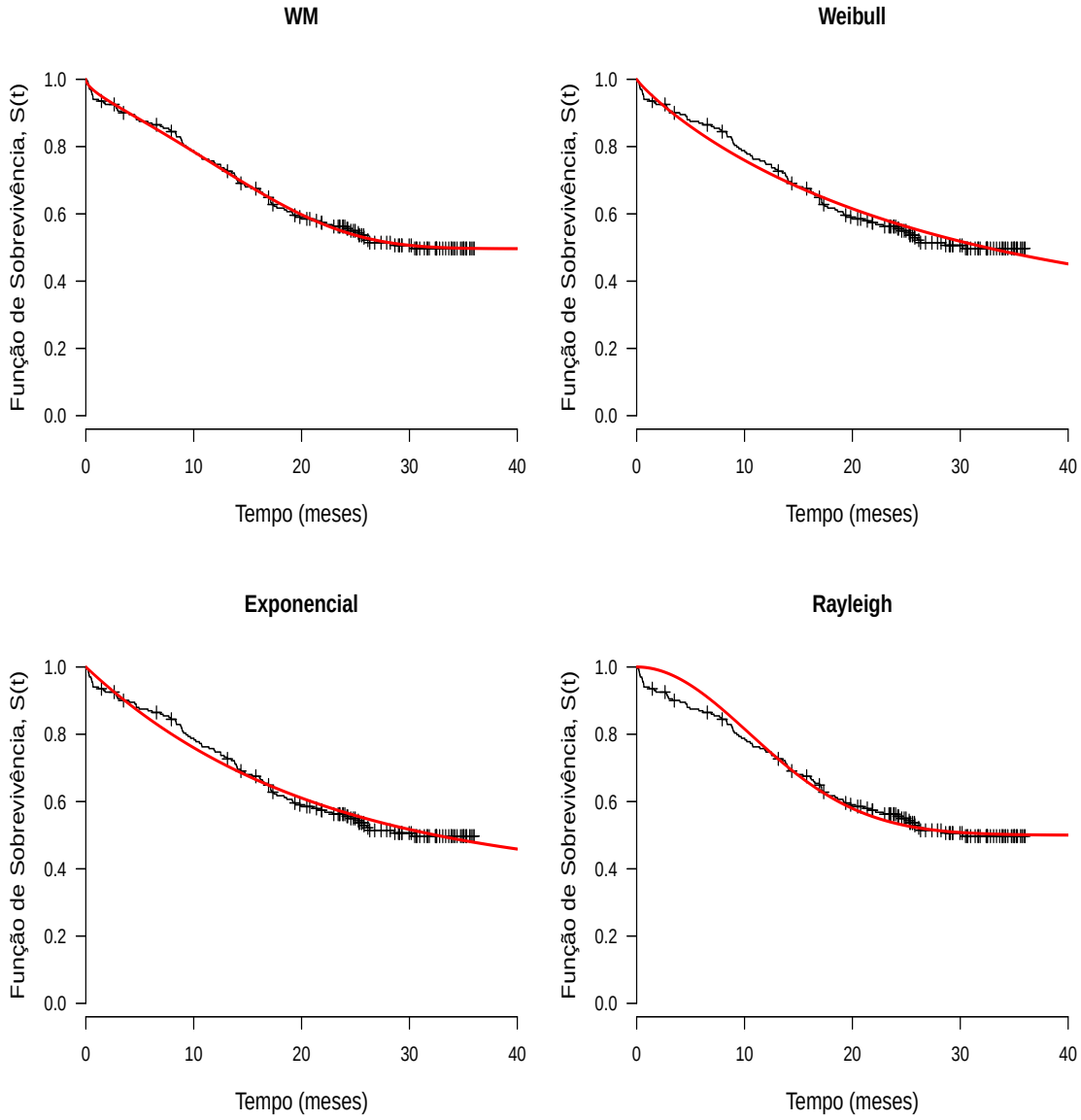


Figura 7 – Funções de sobrevivência obtidas com estimativas Bayesianas do modelo de mistura, para WM e suas distribuições particulares.

pelas curvas de Kaplan-Meier, e parecem aproximar-se de uma assíntota em um período posterior ao descrito pelos dados.

Os gráficos da Figura 8 descrevem as funções risco obtidas dos ajustes dos modelos considerando cada uma das distribuições estudadas. As curvas do gráfico do painel à esquerda **(a)** caracterizam todos os indivíduos, enquanto o gráfico à direita **(b)** descreve apenas os indivíduos suscetíveis (“não curados” da doença). Nota-se que, neste caso, a função risco de um modelo baseado na distribuição exponencial é constante. Para o modelo baseado na distribuição de Weibull, a curva é decrescente, para a Rayleigh a curva é crescente, enquanto para a WM tem a forma de banheira. Sugere-se então a adequabilidade da distribuição WM para esses dados, considerando a função de risco empírica apresentada

na Figura 5.

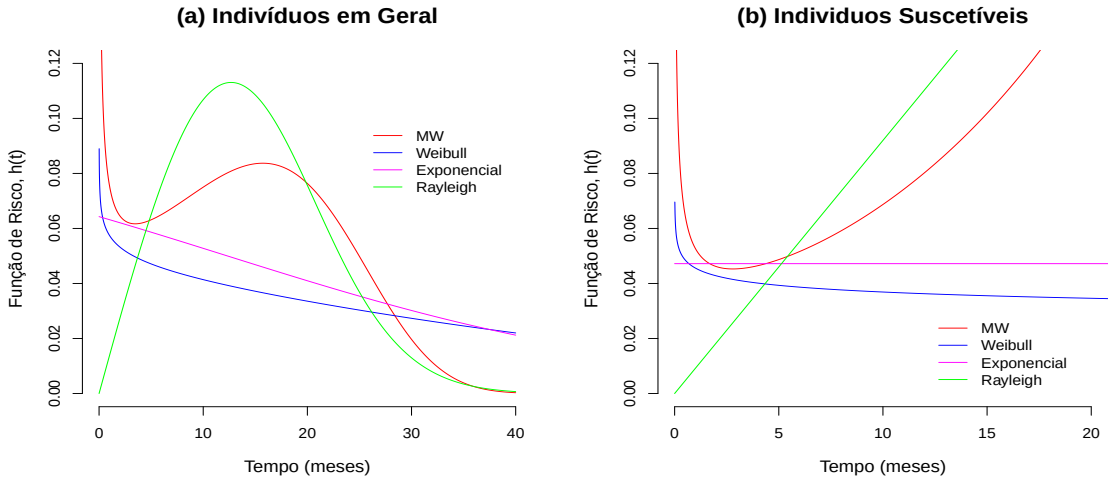


Figura 8 – Funções de risco obtidas com estimativas Bayesianas do modelo de mistura.

Embora as curvas de sobrevida ajustadas, apresentadas na Figura 7, sejam razoavelmente semelhantes para os modelos paramétricos baseados nas distribuições estudadas, nota-se na Figura 8 que as respectivas funções risco assumem formas radicalmente diferentes. Isto evidencia a importância da escolha de uma distribuição observando também o comportamento da função de risco. Nessa aplicação, dada a semelhança das estimativas entre os modelos de mistura e não mistura, não são exibidos gráficos com os ajustes obtidas das suas estimativas.

A Figura 9 apresenta os gráficos TTT-plot para a função de risco empírica dos diferentes tratamentos e as Tabelas 6 e 5 apresentam as estimativas dos parâmetros do modelo baseado na distribuição WM considerando a fração de cura e o tratamento como uma covariável, sendo comparados os modelos com mistura e sem mistura. A Tabela 5 apresenta as estimativas frequentistas de máxima verossimilhança, enquanto a Tabela 6 exibe as estimativas Bayesianas.

Os intervalos de confiança 95% para a fração de cura $p$ das estimativas de máxima verossimilhança foram obtidos pelo método delta, descrito na seção 2.6. Nota-se que, em todas as situações, os intervalos de confiança ou HPD para o parâmetro $\gamma_1$ contém o valor zero, enquanto os intervalos de confiança ou HPD para $\beta_1$ não contém o valor zero. Estes resultados sugerem que o tratamento com quimioradioterapia adjuvante não tem efeito sobre a proporção de óbitos dos pacientes, mas a quimioradioterapia pode trazer um tempo de sobrevida maior.

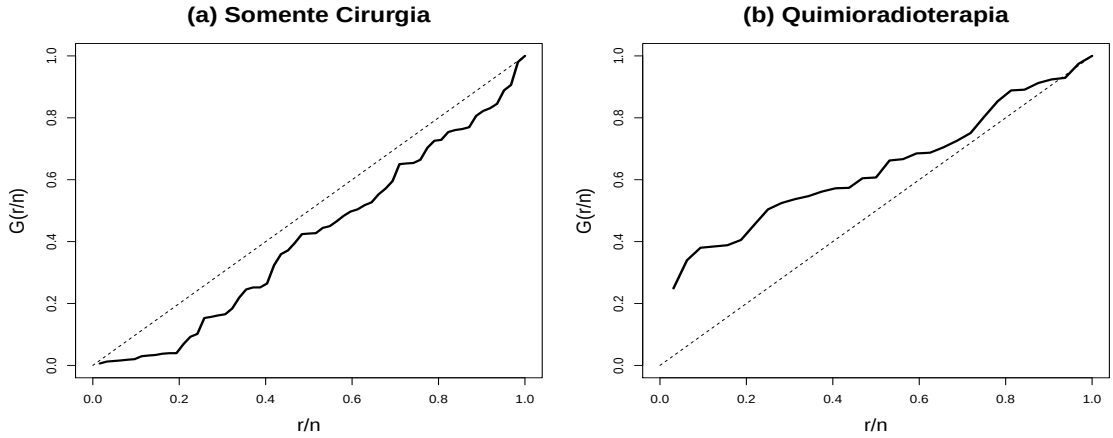


Figura 9 – Gráfico TTT-plot dos dados de câncer gástrico dos pacientes para cada um dos tratamentos.

Tabela 5 – Estimativas frequentistas de máxima verossimilhança dos parâmetros assumindo os modelos de mistura e não mistura, baseados na distribuição WM, com a covariável incluída no proporção de cura $p$ e no parâmetro de forma $\beta$.

| Modelos | | Estimativas (Erro Padrão) | Intervalos de Confiança $95\%$ | | | Critérios de Seleção | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Assintótico | Perfilado | A$^a$ | $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ | AIC | AICc | BIC |
| WM mistura | $\alpha$ | 0,0847 (0,0240) | (0,0376;0,1317) | (0,0667;0,1052) | 6,65 | -436,5 | 885,1 | 885,6 | 904,9 |
| | $\beta_0$ | -0,5897 (0,2459) | (-1,0717;-0,1076) | (-0,7711;-0,4472) | -12,01 | | | | |
| | $\beta_1$ | -0,6043 (0,3517) | (-1,2936;0,0851) | (-1,2496;-0,2701) | -31,76 | | | | |
| | $\lambda$ | 0.0829 (0,0168) | (0,0504;0,1153) | (0,0704;0,0938) | -6,45 | | | | |
| | $\gamma_0$ | -0,4466 (0,1389) | (-0,7190;-0.1792) | (-0,6652;-0,2429) | -3,54 | | | | |
| | $\gamma_1$ | 0.1982 (0,2254) | (-0,2436;0,6400) | (-0,1587;0,5158) | -5,83 | | | | |
| | $p_0$ | 0.4726 (0,0469) | (0,3806;0,5646) | * | | | | | |
| | $p_1$ | 0.5426 (0,0635) | (0,4171;0.6661) | * | | | | | |
| WM não mistura | $\alpha$ | 0,0597 (0,0174) | (0,0256;0.0938) | (0,0462;0,0755) | 7,97 | -436,1 | 884,1 | 884,6 | 903,6 |
| | $\beta_0$ | -0,5971 (0,2500) | (-1,0872;-0.1070) | (-0,7984;-0.4409) | -12,60 | | | | |
| | $\beta_1$ | -0,6343 (0,4023) | (-1,4228;0,1542) | (-1.3768;-0.2680) | -33,93 | | | | |
| | $\lambda$ | 0,0927 (0,0174) | (0,0586;0,1268) | (0,0790;0,1046) | -6,81 | | | | |
| | $\gamma_0$ | -0,4766 (0,1411) | (-0,7531;-0,2001) | (-0,6852;-0,2794) | -2,83 | | | | |
| | $\gamma_1$ | 0,2687 (0,2226) | (-0,1676;0,7050) | (-0,0590;0,5692) | -4,34 | | | | |
| | $p_0$ | 0.4625(0,0471) | (0,3703;0,5548) | * | | | | | |
| | $p_1$ | 0.5562 (0,0622) | (0,4343;0,6782) | * | | | | | |

* Não é obtido intervalos perfilados, pois estas estimativas são funções de parâmetros.
$^a$ Medida de Assimetria para Intervalo de Confiança Perfilado.

Se considerarmos as funções de risco apresentadas empíricas na Figura 9, onde no painel **(a)** aparentemente temos uma curva convexa, e em **(b)** um curva não muito bem definida, mas aparentemente parece ser côncava no inicio e convexa no final. Assim pode-se dizer que temos para os que receberam somente cirurgia uma função de risco decrescente, e para aqueles que receberam também a quimioradioterapia uma função de risco unimodal. Logo, como a distribuição WM não comporta essas formas ela pode não ser adequada a esses dados, quando considera-se as covariáveis tratamento no modelo.

Tabela 6 – Estimativas Bayesianas dos parâmetros assumindo os modelos de mistura e não mistura, baseados na distribuição WM, com a covariável incluída no proporção de cura $p$ e no parâmetro de forma $\beta$.

| Modelos | | Médias *posteriori* | Intervalos HPD 95% | Critérios de Seleção | | | | Diagnóstico de Convergência | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | LMPL | DIC | EAIC | EBIC | p-valor HW | p-valor Geweke |
| WM com mistura | $\alpha$ | 0,1021 | (0,0528;0,1509) | -437,5 | 883,6 | 896,6 | 923,0 | 0,115 | 0,161 |
| | $\beta_0$ | -0,7999 | (-1,3300;-0,3041) | | | | | 0,896 | 0,588 |
| | $\beta_1$ | -1,6970 | (-5,0730;-0,0402) | | | | | 0,851 | 0,881 |
| | $\lambda$ | 0,0893 | (0,0553;0,1202) | | | | | 0,623 | 0,812 |
| | $\gamma_0$ | -0,1132 | (-0,4763;0,2540) | | | | | 0,398 | 0,202 |
| | $\gamma_1$ | 0,2081 | (-0,5336;0,8800) | | | | | 0,589 | 0,498 |
| | $p_0$ | 0,4720 | (0,3831;0,5632) | | | | | 0,396 | 0,209 |
| | $p_1$ | 0,5235 | (0,3763;0,6637) | | | | | 0,792 | 0,478 |
| WM não mistura | $\alpha$ | 0,0715 | (0,0374;0,1080) | -436,9 | 882,6 | 894,8 | 921,2 | 0,304 | 0,126 |
| | $\beta_0$ | -0,7977 | (-1,3350;-0,3143) | | | | | 0,752 | 0,157 |
| | $\beta_1$ | -1,8320 | (-5,3080;-0,0016) | | | | | 0,335 | 0,280 |
| | $\lambda$ | 0,0990 | (0,0654;0,1313) | | | | | 0,856 | 0,595 |
| | $\gamma_0$ | -0,1453 | (-0,5254;0,2169) | | | | | 0,350 | 0,714 |
| | $\gamma_1$ | 0,2907 | (-0,3435;0,9363) | | | | | 0,490 | 0,424 |
| | $p_0$ | 0,4641 | (0,3716;0,5540) | | | | | 0,350 | 0,714 |
| | $p_1$ | 0,5357 | (0,4085;0,6673) | | | | | 0,884 | 0,541 |

A Figura 10 compara as curvas de sobrevida obtidas dos modelos bayesianos com e sem mistura, baseados na distribuição WM. Observa-se que as curvas obtidas do ajuste de um modelo frequentista (painel à esquerda) são semelhantes àquelas obtidas do modelo Bayesiano (painel à direita). É possível observar que as curvas estimadas e as de Kaplan-Meier, não são muito próximas, mostrando que a WM pode não ser adequada ao ajuste desses dados se comparamos os tratamentos. Entretanto é perceptível que os modelos em todos eles captaram de forma significativa os platôs formados pela fração de cura.

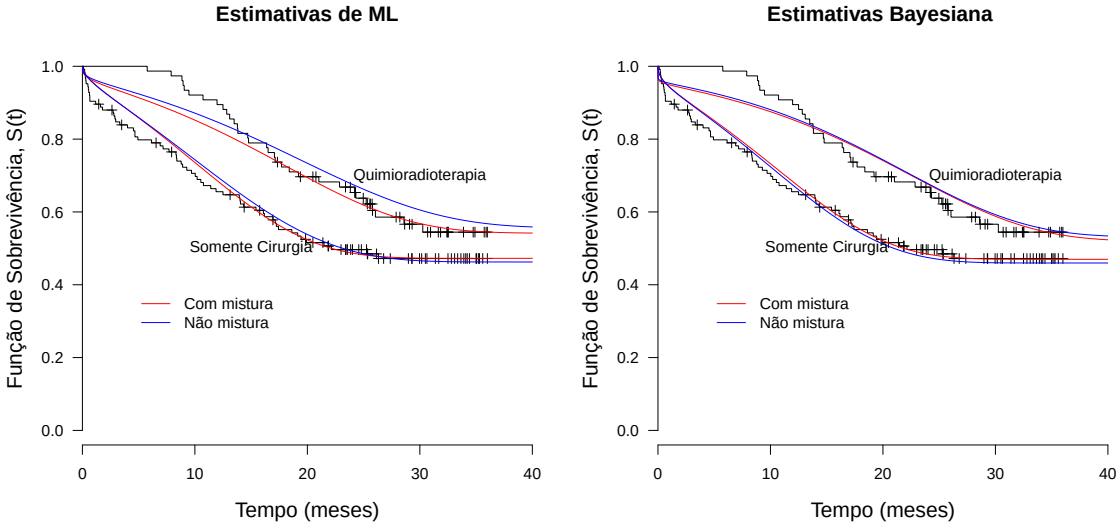


Figura 10 – Funções de sobrevivência obtidas com o modelo de mistura e não mistura, para a Weibull modificada na presença de covariável.

## 4.4 Discussão

A distribuição de Weibull é frequentemente usada em estudos em oncologia para a análise de dados de sobrevivência. Esta distribuição tem por vantagens a flexibilidade de sua função risco e a relativa facilidade encontrada na estimação de seus parâmetros. Entretanto, em alguns estudos médicos, pode se tornar necessário o uso de uma distribuição mais complexa para a análise dos dados. Esta complexidade pode estar na introdução de um parâmetro adicional na distribuição, de modo que ela comporte funções risco que não são sempre crescentes ou decrescentes, ou em introduzir uma regressão em dois ou mais parâmetros do modelo.

O presente estudo mostra que a distribuição WM é uma alternativa viável nesta situação, dado que ela permite uma função risco em forma de banheira, o que a torna mais adequada aos dados de câncer gástrico apresentados. Com os atuais avanços da pesquisa médica em novos medicamentos e tratamentos para o câncer, é esperado que os dados de sobrevivência em muitos estudos apresentem uma fração de indivíduos não suscetíveis ao evento de interesse. Este evento, nos estudos em câncer, pode ser a recidiva da doença ou o óbito. Dessa forma, os modelos com fração de cura tornam-se essenciais para as aplicações das ferramentas de análise de sobrevivência a dados reais.

Enquanto uma grande parte dos trabalhos publicados em revistas médicas utilizam ferramentas estatísticas mais simples, como o modelo de riscos proporcionais e o teste do log-rank, por serem talvez mais "fáceis" de serem realizados ou mesmo por falta de conhecimentos de outros métodos, e ainda por serem este os recursos presentes nos programas estatísticos comercias, os modelos paramétricos com fração de cura mostram-se versáteis em diferentes situações práticas. Como exemplo, enquanto as ferramentas usuais de comparações entre grupos baseiam-se em pressupostos de proporcionalidade de riscos, os modelos paramétricos são mais flexíveis e capazes de comportar diferentes formas para a função risco.

No presente estudo, os ajustes do modelo baseados no método da máxima verossimilhança mostraram-se adequados aos dados, mas observou-se que em algumas situações os intervalos assintóticos para os parâmetros podem não ser adequados, dado que seus limites podem extrapolar o espaço paramétrico. Uma possível forma de obter intervalos de confiança mais adequados é por meio de intervalos perfilados, principalmente em casos em que a função de verossimilhança do parâmetro não é simétrica. Um obstáculo ao uso dos intervalos perfilados é que na prática pode haver alguns problemas como por exemplo os perfis de verossimilhança serem monótonos, ou até mesmo a dificuldade em implementar uma função para calculá-lo. Dessa forma, os modelos Bayesianos podem trazer intervalos HPD mais factíveis.

Ambos os métodos aplicados nesse trabalho de mistura e não mistura para incorpo-

rar a fração de cura foram adequados, apresentando ajustes muitos parecidos entre eles. Em todos os casos que se considerou a distribuição WM na presença ou não de covariável, os dois métodos estimaram de forma conveniente a fração de cura.

Nota-se nessa aplicação que é fundamental ao tratar dados reais conhecer todos seus comportamentos, principalmente as características da função de risco, o que pode ser fundamental para identificar uma distribuição que se adeque de forma mais coerente aos dados.

# Capítulo 5

# A Distribuição Beta-Weibull

Se denotarmos por $G_0(t)$ a função de distribuição acumulada de uma variável aleatória $T$, a classe de distribuições Beta generalizada é definida por

$$F_0(t) = I_{G_0(t)}(\alpha, \beta) = \frac{B_{G_0(t)}(\alpha, \beta)}{B(\alpha, \beta)} = \frac{\int_0^{G_0(t)} w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w}{B(\alpha, \beta)}, \tag{5.1}$$

onde $\alpha > 0,\ \beta > 0,\ B(\alpha, \beta) = \Gamma(a)\Gamma(b)/\Gamma(\alpha+\beta)$ é a função beta, $\Gamma(a) = \int_0^\infty z^{\alpha-1}e^{-z}\,\mathrm{d}z$ é a função gama e $B_{G_0(t)}(\alpha, \beta)$ é a função beta incompleta.

Se $G_0(t)$ em (5.1) for a função de distribuição acumulada da distribuição normal com média $\mu$ e variância $\sigma^2$, temos a chamada distribuição Beta-Normal (EUGENE et al., 2002; FAMOYE et al., 2005).

Um modelo baseado na função distribuição acumulada da distribuição Weibull com parâmetro de forma $\gamma$ e parâmetro de escala $\lambda$ assume que

$$G_0(t) = 1 - exp\left[-\left(\frac{t}{\lambda}\right)\right],\ t > 0.$$

Substituindo esta expressão em (5.1) obtemos

$$F_0(t) = \frac{1}{B(\alpha, \beta)} \int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]} w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w. \tag{5.2}$$

A função de sobrevivência é dada por $S_0 = 1 - F_0(t)$. Note que essa função não possui forma fechada, sendo expressa em termos de integral. A função densidade de probabilidade da distribuição Beta-Weibull (BW) de quatro parâmetros é escrita como

$$f_0(t) = \frac{\gamma t^{\gamma-1}}{\lambda^\gamma B(\alpha, \beta)} \exp\left[-\beta\left(\frac{t}{\lambda}\right)^\gamma\right]\left\{1 - \exp\left[-\left(\frac{t}{\lambda}\right)^\gamma\right]\right\}^{\alpha-1},\ t > 0, \tag{5.3}$$

onde $\alpha > 0,\ \beta > 0,\ \gamma > 0$ e $\lambda > 0$. Com função de risco correspondente dada por

$$h_0(t) = \frac{f_0(t)}{S_0(t)} = \frac{\gamma t^{\gamma-1}\lambda^{-\gamma}\exp\left[-\beta\left(\frac{t}{\lambda}\right)^{\gamma}\right]\left\{1-\exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha-1}}{B(\alpha,\beta) - \int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]} w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w}. \tag{5.4}$$

Notemos que a distribuição BW é uma generalização de algumas distribuições já descritas na literatura, tal distribuições são:

1. **A distribuição exponencial Weibull (EW)**: A distribuição EW de três parâmetros (MUDHOLKAR; SRIVASTAVA, 1993; NASSAR; EISSA, 2003) é caso especial da BW quando $\beta = 1$ em (5.3).

2. **A distribuição Beta-exponencial (BE):** A distribuição BE de três parâmetros (NADARAJAH; KOTZ, 2006) é caso especial da BW quando $\gamma = 1$ em (5.3) .

3. **A distribuição Weibull:** Quando $\alpha = 1$ e $\beta = 1$ , a expressão (5.3) se reduz a função densidade de probabilidade da distribuição Weibull de dois parâmetros.

Cordeiro et al. (2011) encontaram formas fechadas para as expressões dos momentos da distribuição BW e ainda introduziram as estimativas de máxima verossimilhança para os parâmetros do modelo. Tem-se ainda que Cordeiro et al. (2013b) demonstraram que a densidade da BW pode ser expressa como uma mistura de densidades da Weibull e calcularam as expressões para seus momentos. A aplicabilidade da distribuição BW em dados reais foi investigada por Wahed et al. (2009) para modelar os dados de sobrevivência censurados de um estudo sobre câncer de mama.

É possível dizer que a BW é muito versátil nas formas de suas funções, sobretudo na função risco, como já notado por Lee et al. (2007). Ela pode apresentar as seguintes formas:

1. Constante: Quando $\alpha = \gamma = 1$, sendo que é contante no valor $\beta/\lambda$.

2. Decrescente: Quando $\alpha\gamma \leq 1$ e $\gamma \leq 1$.

3. Crescente: Quando $\alpha\gamma \geq 1$ e $\gamma \geq 1$.

4. Banheira: Quando $\alpha\gamma < 1$ e $\gamma > 1$.

5. Unimodal: Quando $\alpha\gamma > 1$ e $\gamma < 1$.

Na Figura (11), apresenta-se os gráficos das expressões (5.3) e (5.4) e para $S_0$ considerando alguns valores de parâmetros.

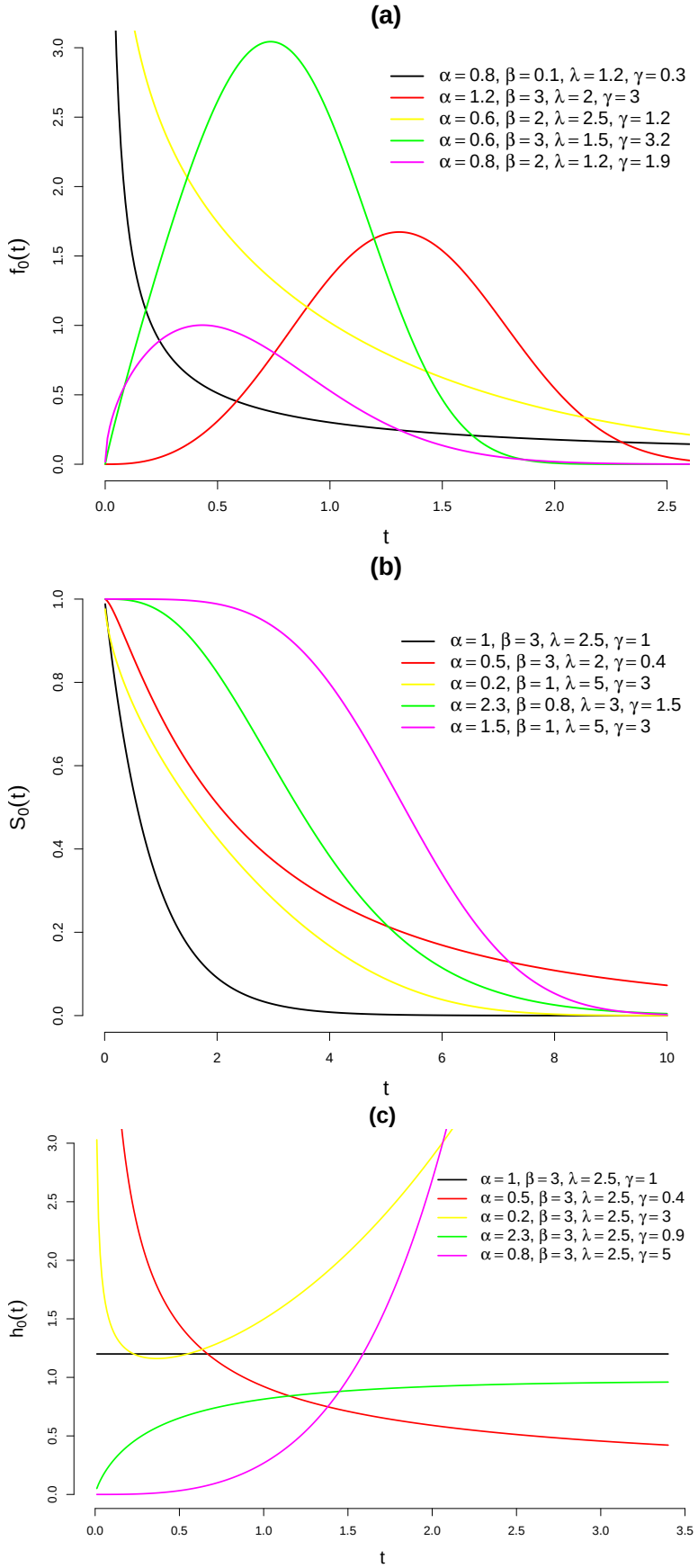


Figura 11 – Gráficos da função densidade de probabilidade (a), da função de sobrevivência (b) e da função de risco (c) da distribuição Beta-Weibull.

## 5.1 Média e Variância

De acordo com a Equação (2.4), temos que para a BW o tempo médio até a falha é dada por

$$\mu = \int_0^{+\infty} \frac{\gamma t^{\gamma-1}}{\lambda^\gamma B(\alpha,\beta)} \exp\left[-\beta\left(\frac{t}{\lambda}\right)^\gamma\right]\left\{1-\exp\left[-\left(\frac{t}{\lambda}\right)^\gamma\right]\right\}^{\alpha-1} \mathrm{d}t, \tag{5.5}$$

Esta integral não possui forma fechada, e sua solução deve ser encontrada por integração numérica. De forma análoga temos que a variância até o tempo de falha de acordo com a Equação (2.5) é dado por

$$\mathrm{Var}(T) = 2\int_0^{+\infty} t\left[1-\frac{1}{B(\alpha,\beta)}\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]} w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right]\mathrm{d}t - \mu^2. \tag{5.6}$$

Como a média até o tempo de falha a variância até o tempo de falha não apresenta forma fechada também, sendo necessário uma integração numérica para encontrar sua solução.

## 5.2 Função de Verossimilhança

A função de verossimilhança para a BW considerando os parâmetros pertencentes ao vetor $\boldsymbol{\theta} = (\alpha,\beta,\gamma,\lambda)^T$ é dada por

$$L(\boldsymbol{\theta}) = \left(\frac{\gamma}{\lambda^\gamma B(\alpha,\beta)}\right)^n \prod_{i=1}^n t_i^{\gamma-1}\exp\left[-\beta\left(\frac{t_i}{\lambda}\right)^\gamma\right]\left\{1-\exp\left[-\left(\frac{t_i}{\lambda}\right)^\gamma\right]\right\}^{\alpha-1}.$$

Para uma amostra aleatória de tamanho $n$, o seu logaritmo é dado por

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= n\ln(\gamma) - n\gamma\ln(\lambda) - \ln\left[B(\alpha,\beta)\right] + (\gamma-1)\sum_{i=1}^n \ln(t_i)\\ &+ \frac{\beta}{\lambda^\gamma}\sum_{i=1}^n (t_i)^\gamma + (\alpha-1)\sum_{i=1}^n \ln\left\{1-\exp\left[-\left(\frac{t_i}{\lambda}\right)^\gamma\right]\right\}\end{aligned}$$

onde $\ell(\boldsymbol{\theta})$ é $\ln\left(L(\boldsymbol{\theta})\right)$

Considerando agora a função de verossimilhança para dados censurados apresentada em (3.8) para a BW, temos as funções de verossimilhança e seu logaritmo, respectivamente, dadas por

$$\begin{aligned}L(\boldsymbol{\theta}) &= \prod_{i=1}^n \left[\frac{\gamma t^{\gamma-1}}{\lambda^\gamma B(\alpha,\beta)}\exp\left[-\beta\left(\frac{t}{\lambda}\right)^\gamma\right]\left\{1-\exp\left[-\left(\frac{t}{\lambda}\right)^\gamma\right]\right\}^{\alpha-1}\right]^{d_i}\\ &\times \prod_{i=1}^n \left[1-\frac{1}{B(\alpha,\beta)}\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]} w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right]^{1-d_i}\end{aligned}$$

e

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= \ln\left[\frac{\gamma t^{\gamma-1}}{\lambda^{\gamma}B(\alpha,\beta)}\right]\sum_{i=1}^{n}d_i+(\gamma-1)\sum_{i=1}^{n}d_i\ln(t_i)-\frac{\beta}{\lambda^{\gamma}}\sum_{i=1}^{n}d_i\,(t_i)^{\gamma}\\ &+ (\alpha-1)\sum_{i=1}^{n}d_i\ln\left\{1-\exp\left[-\left(\frac{t_i}{\lambda}\right)^{\gamma}\right]\right\}\\ &+ \sum_{i=1}^{n}(1-d_i)\ln\left[1-\frac{1}{B(\alpha,\beta)}\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]}w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right].\end{aligned}$$

Assumindo o modelo de mistura apresentado em (2.9) e (2.10), a função de verossimilhança na presença de dados censurados para a BW considerando os parâmetros pertencentes ao vetor $\boldsymbol{\theta}=(\alpha,\beta,\gamma,\lambda,p)^T$ é dada por

$$\begin{aligned}L(\boldsymbol{\theta}) &= \prod_{i=1}^{n}\left[\frac{(1-p)\gamma t^{\gamma-1}}{\lambda^{\gamma}B(\alpha,\beta)}\exp\left[-\beta\left(\frac{t}{\lambda}\right)^{\gamma}\right]\left\{1-\exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha-1}\right]^{d_i}\\ &\times \prod_{i=1}^{n}\left\{p+(1-p)\left[1-\frac{1}{B(\alpha,\beta)}\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]}w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right]\right\}^{1-d_i}\end{aligned}$$

sendo que seu respectivo logaritmo é dado por

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= \ln\left[\frac{(1-p)\gamma t^{\gamma-1}}{\lambda^{\gamma}B(\alpha,\beta)}\right]\sum_{i=1}^{n}d_i+(\gamma-1)\sum_{i=1}^{n}d_i\ln(t_i)-\frac{\beta}{\lambda^{\gamma}}\sum_{i=1}^{n}d_i\,(t_i)^{\gamma}\\ &+ (\alpha-1)\sum_{i=1}^{n}d_i\ln\left\{1-\exp\left[-\left(\frac{t_i}{\lambda}\right)^{\gamma}\right]\right\}\\ &+ \sum_{i=1}^{n}(1-d_i)\ln\left\{p+(1-p)\left[1-\frac{1}{B(\alpha,\beta)}\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]}w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right]\right\}.\end{aligned}\tag{5.7}$$

Além disso, considerando o modelo de não-mistura apresentado em (2.12) e (2.12) , a função de verossimilhança na presença de dados censurados considerando a Equação 3.10 para a BW considerando os parâmetros pertencentes ao vetor $\boldsymbol{\theta}=(\alpha,\beta,\gamma,\lambda,p)^T$, é dada por

$$\begin{aligned}L(\boldsymbol{\theta}) &= \prod_{i=1}^{n}\left[-\frac{\gamma(\ln p)t^{\gamma-1}}{\lambda^{\gamma}B(\alpha,\beta)}\exp\left[-\beta\left(\frac{t}{\lambda}\right)^{\gamma}\right]\left\{1-\exp\left[-\left(\frac{t}{\lambda}\right)^{\gamma}\right]\right\}^{\alpha-1}\right]^{d_i}\\ &\times \exp\left[\frac{lnp}{B(\alpha,\beta)}\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]}w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right]\end{aligned}$$

onde seu logaritmo é dado por

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= \ln\left[-\frac{\gamma(\ln p)t^{\gamma-1}}{\lambda^{\gamma}B(\alpha,\beta)}\right]\sum_{i=1}^{n}d_i+(\gamma-1)\sum_{i=1}^{n}d_i\ln(t_i)-\frac{\beta}{\lambda^{\gamma}}\sum_{i=1}^{n}d_i\,(t_i)^{\gamma}\\ &+ (\alpha-1)\sum_{i=1}^{n}d_i\ln\left\{1-\exp\left[-\left(\frac{t_i}{\lambda}\right)^{\gamma}\right]\right\}\\ &+ \frac{\ln p}{B(\alpha,\beta)}\sum_{i=1}^{n}\left[\int_0^{1-exp\left[-\left(\frac{t}{\lambda}\right)\right]}w^{\alpha-1}(1-w)^{\beta-1}\,\mathrm{d}w\right].\end{aligned}\tag{5.8}$$

Note que a derivada de todos as funções de log-verossimilhança apresentadas, são trabalhosas devido a sua complexidade e no caso dos modelos de mistura e não mistura estas funções não apresentam forma fechada.

# Capítulo 6

# Aplicações Beta-Weibull

## 6.1 Os Dados

Klein e Moeschberger (2005) apresentam um banco de dados com 137 pacientes submetidos a um transplante de medula óssea em quatro hospitais da América do Norte, sendo que 41% dos observações foram censuradas. Este dados foram obtidos originalmente de Copelan et al. (1991) e estão disponível no pacote *KMsurv* do *R*, com o nome de *bmt*. O transplante de medula óssea é o tratamento padrão para leucemia aguda. Um transplante é considerado falho se a leucemia retornar no paciente ou quando o paciente morre. Assim, o desfecho de interesse é o tempo em dias de sobrevivência livre de doença, ou seja, até à recaída ou até o fim do estudo.

As curvas de sobrevivência estimadas por Kaplan-Meier são apresentado na Figura 12, onde é possível notar a presença de um platô em $\hat{S}(t) = 0,35$ no gráfico apresentado em (a), sugerindo assim que um modelo de fração de cura pode adequar-se bem a estes dados. O gráfico apresentado no painel (b) da Figura 12 descreve as funções de sobrevivência empíricos para cada um dos grupos: leucemia linfoblástica aguda (ALL), leucemia mielóide aguda (AML) baixo risco e alto risco. Pode-se ainda notar nesse gráfico a presença de platos na cauda direita das curvas empíricas de sobrevivência, indicando que um modelo que inclui um parâmetro que representa a fração de cura é adequado a estes dados.

## 6.2 Métodos

Ao aplicar a distribuição beta-Weibull no conjunto de dados reais, foi dividido por 1000 os valores dos tempos de vida, a fim de facilitar a convergência do algoritmo de estimação. Para estimação e inferência dos parâmetros utilizamos apenas o *software R* versão $3.1.2$. O *OpenBUGS* não foi utilizado, pois a estrutura da função de verossimilhança

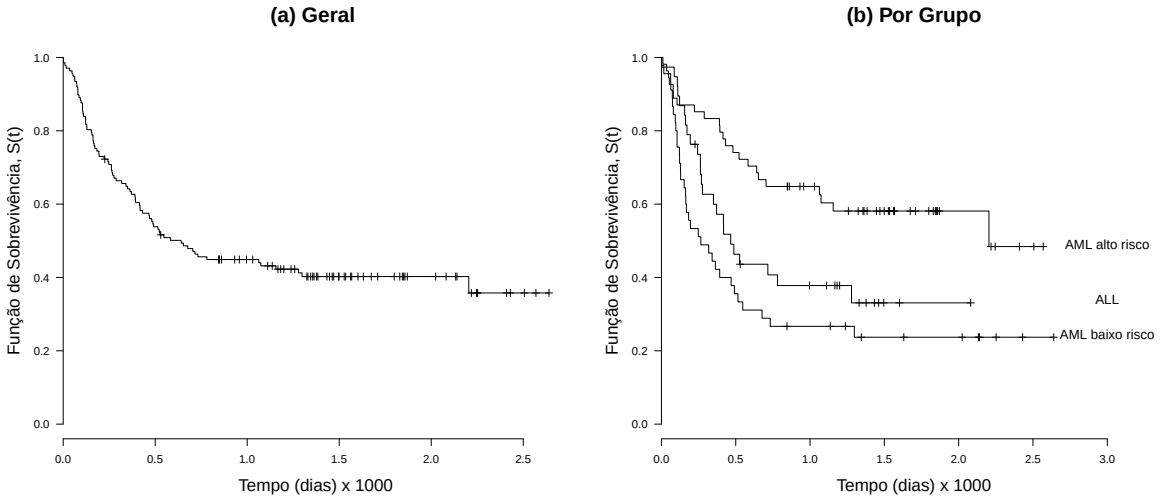


Figura 12 – Estimadores Kaplan-Meier para as função de sobrevivência dos dados de transplante de medula óssea dos pacientes em geral e por grupo.

apresenta a função Beta e a função Beta incompleta e o *OpenBUGS* não oferece suporte a essas funções. Para inferência clássica utilizou-se a função *maxLik* e para inferência Bayesiana foi utilizado a função *MCMCmetrop1R* do pacote *MCMCpack* (MARTIN et al., 2011).

A função *maxLik* disponibiliza vários métodos numéricos para estimação. Por padrão, ela utiliza o de Newton-Rapshon, método utilizado com maior frequência. Contudo, nessa aplicação, utilizou-se o método BFGSR que é uma variante do método Broyden-Fletcher-Goldfarb-Shanno (BFGS) (SHANNO, 1985) para o programa *R*, pois esse obteve melhores resultados. Utilizou-se a matriz de covariância, dada pela própria função *maxLik*, para calcular a correlação entre alguns parâmetros estimados.

Aplicou-se uma distribuição *a priori* gama, aos parâmetros $\alpha, \beta, \gamma$ e$\lambda$, levando em consideração que eles são reais positivos. Entretanto para o parâmetro de cura $p$, assumimos uma distribuição *a priori* beta, por ele estar restrito ao intervalo $(0, 1)$. Dessa forma, temos que $\alpha \sim Gama(a_\alpha, b_\alpha)$ , $\beta \sim Gama(a_\beta, b_\beta)$ ,$\gamma \sim Gama(a_\gamma, b_\gamma)$, $\lambda \sim Gama(a_\lambda, b_\lambda)$ e $p \sim Beta(c_p, d_p)$, sendo que $a_\alpha$ , $b_\alpha$ , $a_\beta$ , $b_\beta$ $a_\gamma$ , $b_\gamma$ , $a_\lambda$ e $b_\lambda$ são os hiperparâmetros conhecidos da distribuição $Gama(a, b)$ com média $a/b$ e variância $a/b^2$ e $c_p$ e $d_p$ são os hiperparâmetros da distribuição $Beta(c, d)$, com média $c/(c + d)$ e variância $cd/[(c + d)^2(c + d + 1)]$, onde consideramos $a_\alpha = b_\alpha = a_\beta = b_\beta = a_\gamma = b_\gamma = a_\lambda = b_\lambda = c_p = d_p = 1$, sendo assim tem-se *a priori* poucos informativas. Esses hiperparametros são considerados apenas para a beta-Weibull, para os casos particulares dela consideramos distribuições *a priori* aos parâmetros não fixados.

Ao considerar-se a variável grupo, definimos duas variáveis *dummy* $X_1$ e $X_2$, assumindo-se que para AML baixo risco tem-se que $X_1 = 0$ e $X_2 = 0$, para ALL tem-se que $X_1 = 1$ e $X_2 = 0$ e $X_1 = 0$ e $X_2 = 1$ para AML alto risco. Dessa forma, uma

regressão incluindo os esses niveis é aplicada considerando uma função de ligação *logito* para o parâmetro $p$ e uma função de ligação logarítmica para o parâmetro $\lambda$. Temos assim que

$$\ln\left(\frac{p_i(\mathbf{x})}{1-p_i(\mathbf{x})}\right) = \eta_0 + \eta_1 x_{1i} + \eta_2 x_{2i}$$

e

$$\ln\left(\lambda_i(\mathbf{x})\right) = \zeta_0 + \zeta_1 x_{1i} + \zeta_2 x_{2i},$$

sendo que $x_{1i}$ e $x_{2i}$ são as observações de $X_1$ e $X_2$ para o $i$-ésimo indivíduo. Assumimos uma distribuição normal *a priori*, $N(e, f^2)$, com média $e$ e variância $f^2$, para cada parâmetro de $\boldsymbol{\varphi}$, sendo $e = 0$ e $f^2 = 10$, tem-se uma *a priori* pouco informativo.

Para todas as distribuições assumimos independência *a priori* entre os parâmetros incluídos no modelo. As distribuições *a posteriori* de interesse são obtidas a partir de amostras simuladas para a distribuição conjunta *a posteriori* usando procedimentos MCMC. Para cada uma das distribuições, simulou-se 3.000.000 de amostras para cada parâmetro de interesse, sendo as primeiras 10.000 amostras descartadas a fim de minimizar os efeitos dos valores iniciais e ainda realizando salto de tamanho 200. Logo, as inferências são feitas sobre 15.000 amostras, obtendo-se dessa forma amostras sucessivas praticamente não correlacionadas. A convergência do algoritmo MCMC foi monitorado observando as séries temporais habituais para as amostras simuladas plotadas e as cadeias foram verificadas por meio da função *heidel.diag(·)* e *geweke.diag(·)* do pacote *coda*, sendo a hipótese nula desses testes é a convergência da cadeia. . Nessa aplicação, a fim de escolher o modelo que melhor se ajusta aos dados utilizou-se apenas a log pseudo verossimilhança marginal (LPML).

## 6.3 Resultados

Considerando os dados de transplante de medula óssea, a Tabela 7 apresenta as estimativas Bayesianas considerando a distribuição beta-Weibull com mistura e seus casos particulares. Estes modelos consideram a função de verossimilhança baseada na equação (5.7). Nota-se que os valores de LPML são muito próximos em todos os ajustes, podendo assim sugerir que todos os modelos se adequam de mesma forma aos dados. Note-se ainda que em todos os modelos, os valores estimados para $p$ são muito próximos, tanto nas médias *a posterior* (0,37) quanto no intervalo HPD, sendo que estas estimativas parecem adequadas ao valor sugerido pela Figura 12, painel (a).

A Tabela 8 apresenta as estimativas frequentistas, considerando também a função de verossimilhança da equação (5.7). É possível notar que os valores de $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ são bem próximos, e como nos resultados Bayesianos, sugere-se que todos os modelos se adequam

Tabela 7 – Estimativas Bayesianas dos parâmetros, assumindo um modelo de mistura considerando os dados de transplante de medula óssea.

| Modelos | | Médias posteriori | Intervalos HPD $95\%$ | Critério de Seleção | Diagnóstico de Convergência | |
|---|---|---|---|---|---|---|
| | | | | LMPL | p-valor HW | p-valor Geweke |
| BW | $\alpha$ | 1,4214 | (0,5112;2,9602) | -73,547 | 0,343 | 0,444 |
| | $\beta$ | 1,2179 | (0,0871;3,8578) | | 0,930 | 0,845 |
| | $\gamma$ | 0,7961 | (0,4335;1,1979) | | 0,418 | 0,456 |
| | $\lambda$ | 0,3249 | (0,0117,1,0890) | | 0,798 | 0,666 |
| | $p$ | 0,3748 | (0,2992;0,4652) | | 0,743 | 0,187 |
| EW | $\alpha$ | 1,3601 | (0,4878;2,6794) | -73,538 | 0,959 | 0,370 |
| | $\gamma$ | 0,8110 | (0,4689;1,2197) | | 0,530 | 0,048 |
| | $\lambda$ | 0,2342 | (0,0654;0,4395) | | 0,915 | 0,213 |
| | $p$ | 0,3743 | (0.2888;0,4603) | | 0,768 | 0,900 |
| BE | $\alpha$ | 0,9964 | (0,7033;1,3139) | -72,917 | 0,770 | 0,835 |
| | $\beta$ | 1,2271 | (0,4181;3,4995) | | 0,781 | 0,372 |
| | $\lambda$ | 0,3888 | (0,0132;1,0781) | | 0,992 | 0,628 |
| | $p$ | 0,3806 | (0,3033;0,4665) | | 0,681 | 0720 |
| Weibull | $\gamma$ | 0,9596 | (0,7808;1,1379) | -73,199 | 0,956 | 0,329 |
| | $\lambda$ | 0,3129 | (0,2330;0,4041) | | 0,521 | 0,996 |
| | $p$ | 0,3772 | (0,2914;0,4612) | | 0,401 | 0,882 |

aos dados. Em todos os modelos, os valores estimados para $p$ são muito próximos, assim como observado para os resultados Bayesianos, adequando-se o valor sugerido pela Figura 12, painel (a).

Entretanto, algumas estimativas frequentistas são consideravelmente diferentes das Bayesianas. Observando-se as estimativas para a BW, os parâmetros $\beta$ e $\lambda$ apresentados na Tabela 8 têm uma grande diferença em relação às apresentadas na Tabela 7. Uma possível explicação pode ser a correlação entre esses dois parâmetros. Sabendo-se que a correlação entre as estimativas de dois parâmetros, $\theta_1$ e $\theta_2$ pode ser calculado por $\text{cor}(\theta_1, \theta_2) = \text{cov}(\theta_1, \theta_2)/\sqrt{\text{Var}(\theta_1) \times \text{Var}(\theta_2)}$, utilizando as covariâncias obtidas por meio da matriz Hessiana, tem-se que $\text{cor}(\beta, \lambda) = 0,97$, e portanto uma correlação extremamente alta. Dessa forma, a alta correlação entre os dois parâmetros pode ser um dos motivos para a diferença entre as estimativas clássica e Bayesianas. Vale ressaltar que a Figura 10 apresentada os gráficos das curvas estimativas, apresentando no painel (a) que a curva estimativa pelo método Bayesiano e a estimada pelo método clássico são praticamente idênticas.

Caso parecido pode ser notado para a distribuição EW. Nesse caso, não há uma diferença muito grande entre as estimativas clássicas e Bayesianas. Contudo, nos pa-

Tabela 8 – Estimativas frequentativas dos parâmetros, assumindo um modelo de mistura considerando os dados de transplante de medula óssea.

| Modelos | | Estimativas (Erro Padrão) | IC $95\%$ | Critério de Seleção $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ |
|---|---|---|---|---|
| BW | $\alpha$ | 1,3696 (0,6794) | (0,0379;2,7012) | -86,5734 |
| | $\beta$ | 0,5081 (1,082) | (-1,6128;2,6291) | |
| | $\gamma$ | 0,8302 (0,2504) | (0,3394;1,3210) | |
| | $\lambda$ | 0,1441 (0,3195) | (-0,4820;0.7703) | |
| | $p$ | 0,3775 (0,0472) | (0,2852;0,4702) | |
| EW | $\alpha$ | 1,3792 (0,7354) | (-0,0622;2,8206) | -86,590 |
| | $\gamma$ | 0,8125 (0,2644) | (0,2943;1,3308) | |
| | $\lambda$ | 0,2965 (0.1627) | (-0,0224;0,6154) | |
| | $p$ | 0,3786 (0,0467) | (0,2871;0,4702) | |
| BE | $\alpha$ | 1,0275 (0,2052) | (0,6253;1,4297) | -86,849 |
| | $\beta$ | 0,2049 (0,0311) | (0,1440;0,2658) | |
| | $\lambda$ | 0,0807 (0,0007) | (0,0793;0,0821) | |
| | $p$ | 0,3844 (0,0443) | (0,2975;0,4712) | |
| Weibull | $\gamma$ | 0,9803 (0,0960) | (0,7921;1,1685) | -86,838 |
| | $\lambda$ | 0,3979 (0,0555) | (0,2892;0,5066) | |
| | $p$ | 0,3826 (0,0451) | (0,2943;0,4709) | |

râmetros $\alpha$ e $\lambda$ que possuem estimativas mais afastadas entre os métodos Bayesiano e clássico e ainda são os que apresentam intervalos de confiança inadequadas (limite inferior do intervalo é negativo), e possuem correlação negativa extremamente alta, isto é, $\text{cor}(\alpha, \lambda) = -0.96$. Isto pode explicar as estimativas inconstantes ou, como apresentado, intervalos inadequados. Mesmo assim, as estimativas de ambos os métodos geram gráficos praticamente idênticos, apresentados no painel (b) da Figura 10.

Algo diferente ocorre na distribuição BE, para a qual as estimativas dos parâmetros $\beta$ e $\lambda$ são diferentes entre as Bayesianas e as clássicas, mas obteve-se intervalos adequados e sem correlação alta entre os parâmetros. Pelo contrário, a correção entre todos os parâmetros é baixa. Além disso, o painel (c) da Figura 10 apresenta as curvas ajustadas para esses estimativas, mostrando curvas muito próximas. Pode-se então dizer que se trata de um típico problema de identificabilidade. A matriz de variância estimada para $\alpha,\ \beta,\ \lambda$ e $p$ é

$$\text{Var}(\alpha, \beta, \lambda, p) = \begin{bmatrix} 4,2 \times 10^{-2} & 2,9 \times 10^{-3} & -4,7 \times 10^{-8} & 6,0 \times 10^{-4} \\ 2,9 \times 10^{-3} & 9,67 \times 10^{-4} & 1,39 \times 10^{-6} & 2,66 \times 10^{-4} \\ -4,7 \times 10^{-8} & 1,39 \times 10^{-6} & 5,48 \times 10^{-7} & -4,89 \times 10^{-9} \\ 6,0 \times 10^{-4} & 2,66 \times 10^{-4} & -4,89 \times 10^{-9} & 1,96 \times 10^{-3} \end{bmatrix}.$$

O determinante dessa matriz é $\det(\text{Var}(\alpha, \beta, \lambda, p)) = 3,31 \times 10^{-14} \approx 0$. Dessa forma é

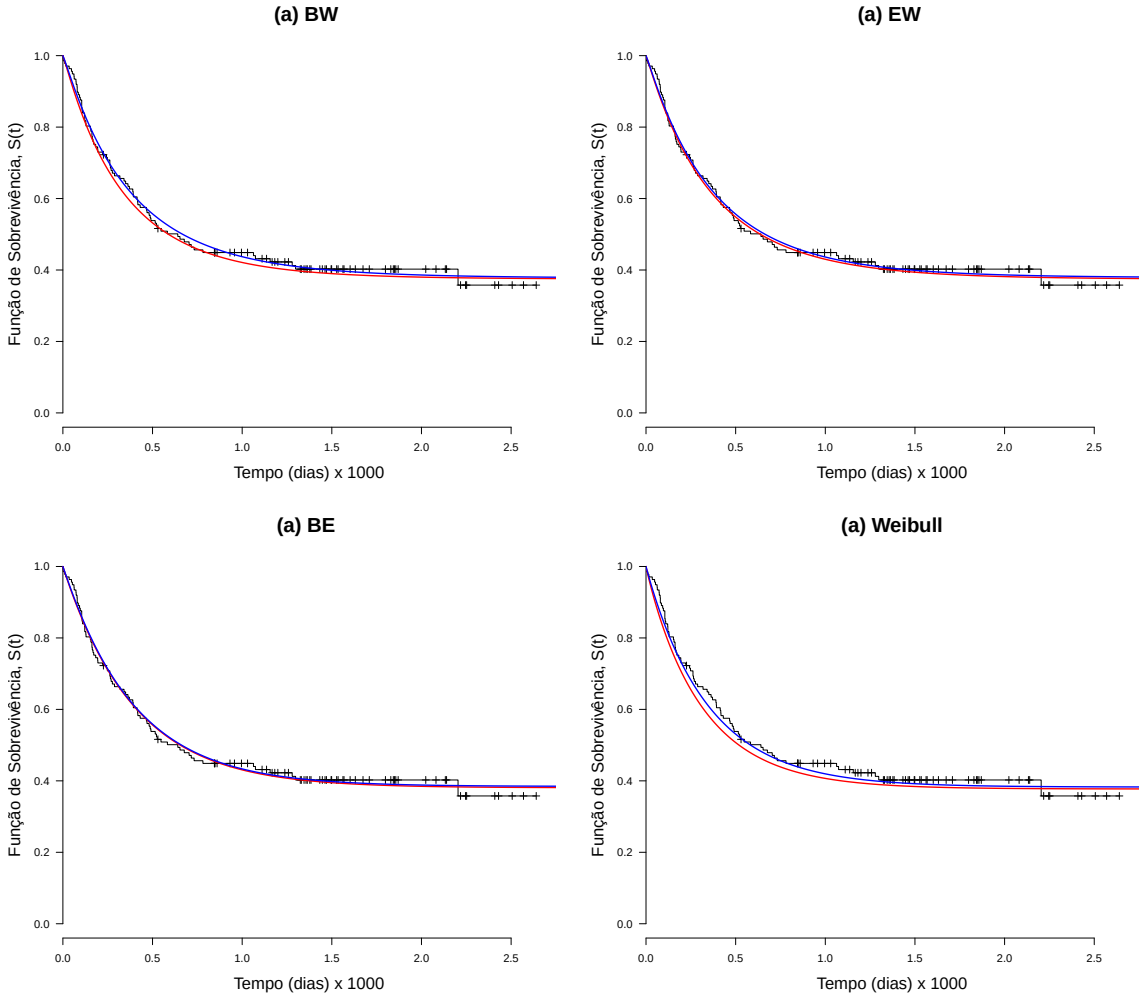


Figura 13 – Funções de sobrevivência obtidas pelas estimativas Bayesianas (em vermelho) e por meio das estimativas frequentistas (em azul) do modelo de mistura, para BW, cada distribuição particular dela.

possível verificar uma ortogonalidade entre as estimativas.

O ajuste do modelo baseado na distribuição Weibull não apresentou esses problemas. Estimativas para os parâmetros da distribuição Weibull foram muito próximos, comparando-se os métodos. Os intervalos de confiança apresentam resultados plausíveis. Isso se deve provavelmente, pelo fato de a distribuição ser mais simples que as outras, em relação ao número de parâmetros.

A Figura 13 apresenta os gráficos dos valores preditos pelas curvas de sobrevivência estimadas *versus* as estimativas de Kaplan-Meier. Nota-se quem em todas as distribuições, os pontos figuram próximos da linha diagonal, não havendo diferenças muito grandes, nem entre as distribuições e nem entre os métodos.

As estimativas Bayesianas do modelo de não mistura da distribuição beta-Weibull e seus casos particulares, considerando a equação de verossimilhança apresentada em (5.8), são descritas na Tabela 9. Observa-se que comparadas com as estimativas Bayesianas do

Tabela 9 – Estimativas Bayesianas dos parâmetros, assumindo um modelo de não mistura considerando os dados de transplante de medula óssea.

| Modelos | | Médias posteriori | Intervalos HPD $95\%$ | Critério de Seleção | Diagnóstico de Convergencia | |
|---|---|---|---|---|---|---|
| | | | | LMPL | p-valor HW | p-valor Geweke |
| BW | $\alpha$ | 1,3537 | (0,5301;2,7528) | -77,012 | 0,272 | 0,097 |
| | $\beta$ | 1,3789 | (0,0855;3,7454) | | 0,548 | 0,824 |
| | $\gamma$ | 0,8586 | (0,4536;1,3286) | | 0,372 | 0,503 |
| | $\lambda$ | 0,4519 | (0,0212,1,5307) | | 0,428 | 0,668 |
| | $p$ | 0,3682 | (0,2774;0,4601) | | 0,108 | 0,273 |
| EW | $\alpha$ | 1,3155 | (0,4941;2,6491) | -77,566 | 0,751 | 0,685 |
| | $\gamma$ | 0,8723 | (0,4605;1,3774) | | 0,676 | 0,882 |
| | $\lambda$ | 0,4519 | (0,0991;0,6430) | | 0,585 | 0,945 |
| | $p$ | 0,3674 | (0.2747;0,4564) | | 0,641 | 0,176 |
| BE | $\alpha$ | 1,1029 | (0,7934;1,4221) | -76,080 | 0,822 | 0,951 |
| | $\beta$ | 1,0600 | (0,0661;3,2384) | | 0,768 | 0,560 |
| | $\lambda$ | 0,4195 | (0,0219;1,2995) | | 0,895 | 0,726 |
| | $p$ | 0,3747 | (0,2900;0,4599) | | 0,191 | 0,801 |
| Weibull | $\gamma$ | 1,0343 | (0,8342;1,2218) | -76,654 | 0,852 | 0,442 |
| | $\lambda$ | 0,4289 | (0,3031;0,5923) | | 0,627 | 0,926 |
| | $p$ | 0,3703 | (0,2864;0,4601) | | 0,951 | 0,472 |

modelo de mistura da Tabela 7, as estimativas de não mistura Bayesianas da BW e de seus casos particulares são muito próximas, como esperado.

Os modelos de não mistura identificaram satisfatoriamente a fração de cura, em torno de 0,37. As curvas ajustadas do modelo de não mistura foram suprimidas devido a serem muito próximas daquelas obtidas do modelo de mistura. Contudo, os valores de LPML dos modelos de não mistura na Tabela 9 são, em todas as distribuições, inferiores se comparados com os do modelo de mistura da Tabela 7, sugerindo que os modelos de mistura têm melhor ajuste para esses dados.

A Tabela 10 apresenta as estimativas frequentistas considerando um modelo de não mistura com função de verossimilhança apresentada na equação (5.8), para a distribuição BW e seus casos particulares. Em todos os modelos, a fração de cura foi bem ajustada e as estimativas são muito próximas dos resultados considerando um modelo de mistura. Os valores de $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ são muito próximos do modelo frequentista com mistura, indicando que se considerar as estimativas frequentistas os dois modelos se adequam igualmente aos dados.

Entretanto, como no caso das estimativas do modelo de mistura, houve algumas

Tabela 10 – Estimativas frequentativas dos parâmetros, assumindo um modelo de não mistura considerando os dados de transplante de medula óssea.

| Modelos | | Estimativas (Erro Padrão) | IC $95\%$ | Critério de Seleção $\ell(\boldsymbol{\theta}, \boldsymbol{x})$ |
|---|---|---|---|---|
| BW | $\alpha$ | 1,3180 (0,3917) | (0,5501;2,0857) | -86,530 |
| | $\beta$ | 0,2207 (0,7469) | (-1,2431;1,6845) | |
| | $\gamma$ | 0,9122 (0,2193) | (0,4824;1,3420) | |
| | $\lambda$ | 0,0991 (0,3061) | (-0,5009;0.6990) | |
| | $p$ | 0,3678 (0,0501) | (0,2697;0,4658) | |
| EW | $\alpha$ | 1,4112 (0,8883) | (-0,3290;3,1529) | -86,467 |
| | $\gamma$ | 0,8426 (0,3582) | (0,1405;1,5446) | |
| | $\lambda$ | 0,4103 (0.2395 | (-0,0592;0,8798) | |
| | $p$ | 0,3719 (0,0502) | (0,2735;0,4702) | |
| BE | $\alpha$ | 1,0030 (0,1994) | (0,6118;1,3935) | -86,717 |
| | $\beta$ | 0,1209 (0,024) | (0,0741;0,1677) | |
| | $\lambda$ | 0,0686 (0,0001) | (0,0685;0,0687) | |
| | $p$ | 0,3709 (0,0477) | (0,2774;0,4643) | |
| Weibull | $\gamma$ | 1,0581 (0,1045) | (0,8532;1,2630) | -86,548 |
| | $\lambda$ | 0,5417 (0,0921) | (0,3613;0,7221) | |
| | $p$ | 0,3772 (0,0466) | (0,2859;0,4684) | |

diferenças entre as estimativas frequentistas referentes a Tabela 10 e as estimativas Bayesianas da Tabela 9. Essa diferença ocorre nas mesmas distribuições que o modelo de não mistura. Na distribuição BW novamente há uma divergência nos parâmetros $\beta$ e $\lambda$. Além das estimativas serem diferentes, existe o problema do intervalo de confiança ser inadequado ao parâmetro. Verifica-se que ocorre o mesmo problema de alta correlação entre esses dois parâmetros, sendo cor$(\beta, \lambda) = -0,89$.

No casos da distribuição EW encontra-se o mesmo obstáculo já verificado para os parâmetros $\alpha$ e $\lambda$, sendo altamente correlacionados, com cor$(\alpha, \lambda) = -0,95$. Já na distribuição BE, não é identificado problemas com os intervalos de confiança e nem para a correlação entre os parâmetros. Apenas identifica-se diferenças entre suas estimativas frequentistas e Bayesianas, e também determinando da matriz de variâncias próximo a 0. Contudo na distribuição Weibull, as estimativas são todas próximas, muito provavelmente devido essa distribuição sem mais simples que as outras consideradas..

Os resultados Bayesianos assumindo um modelo de mistura na presença de covariáveis da distribuição BW e seus casos especiais são apresentado na Tabela 11. Observando os intervalos HPD dos parâmetros das estimativas da BW, apenas $\eta_2$ não inclui em seu intervalo o valor 0, sugerindo que indivíduos no grupo AML baixo risco têm fração de cura diferente de indivíduos do grupo AML alto risco.

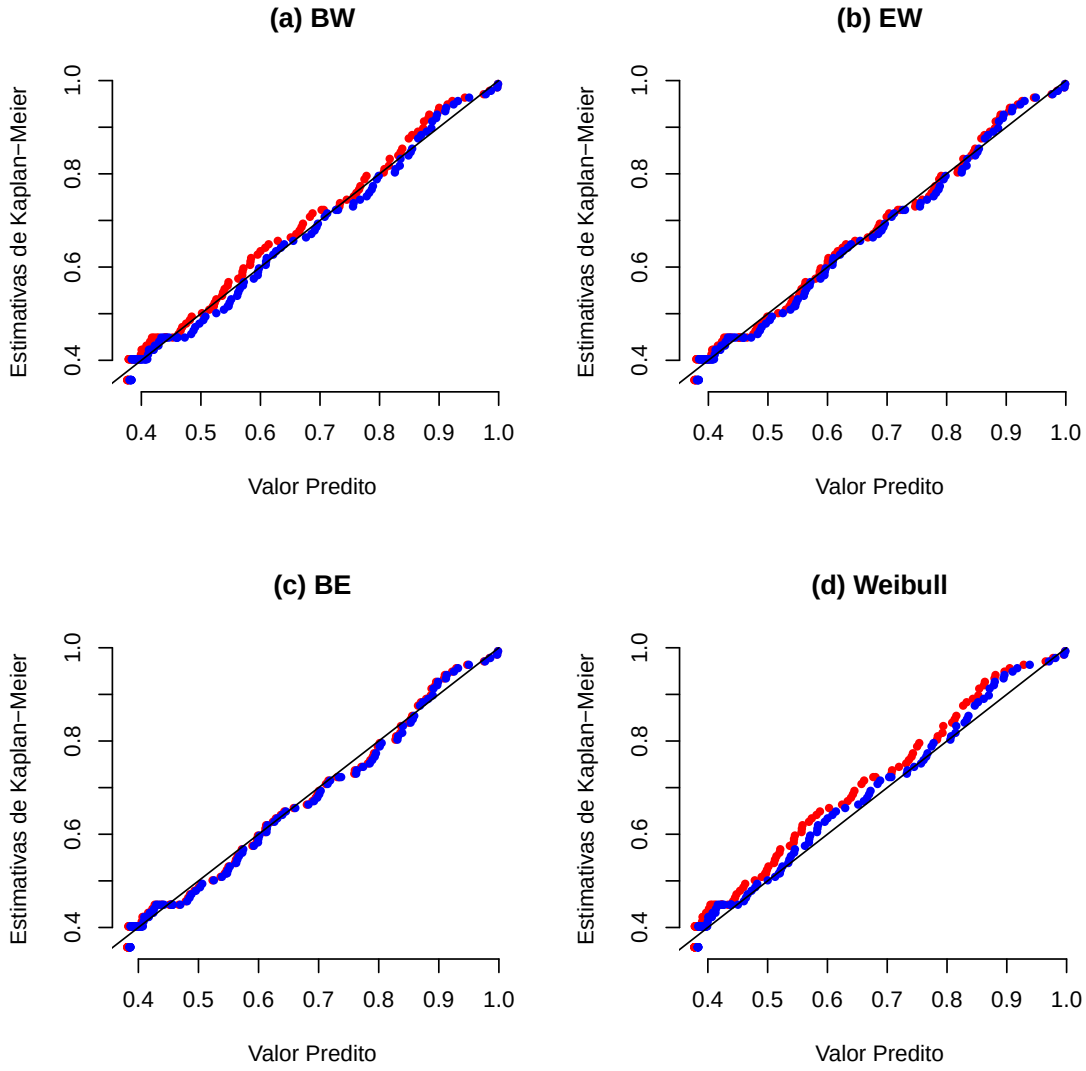


Figura 14 – Gráficos dos estimadores de Kaplan–Meier para as funções de sobrevivência versus os respectivos valores preditos, das estimativas Bayesianas (em vermelho) e frequentistas (em azul) do modelo com mistura, considerando a distribuição Beta-Weibull e seus casos particulares.

É possível obter as estimativas Bayesianas de fração de cura para cada grupo, considerando as amostra simuladas de $\eta_0,\ \eta_1$ e $\eta_2$ nas seguintes relações

$$p_{\text{AML baixo risco}} = \frac{exp(\eta_0)}{1+exp(\eta_0)},\quad p_{\text{ALL}} = \frac{exp(\eta_0+\eta_1)}{1+exp(\eta_0+\eta_1)} \quad \text{e} \quad p_{\text{AML alto risco}} = \frac{exp(\eta_0+\eta_2)}{1+exp(\eta_0+\eta_2)}.$$

Assim, as estimativas de fração de cura para os pacientes classificados no grupo como AML baixo risco, ALL e AML alto risco são respectivamente de 0,467 (intervalo HPD 95% de 0,320 até 0,602), 0,349 (intervalo HPD 95% de 0,197 até 0,503) e 0,259 (intervalo HPD 95% de 0,144 até 0,386).

A Figura 15 mostra as curvas de sobrevivências de Kaplan-Meier para cada grupo e com as respectivas curvas ajustas de acordo com as estimativas da distribuição beta-Weibull

com mistura. Note-se que as curvas obtidas com os modelos se assemelham as curvas empíricas estimadas pelo método de Kaplan-Meier, indicando assim que um ajuste do modelo para esses dados.

Tabela 11 – Estimativas Bayesianas dos parâmetros, assumindo um modelo de mistura com covariáveis, considerando os dados de transplante de medula óssea.

| Modelos | | Médias posteriori | Intervalos HPD $95\%$ | Critério de Seleção | Diagnóstico de Convergência | |
|---|---|---|---|---|---|---|
| | | | | LMPL | p-valor HW | p-valor Geweke |
| BW | $\alpha$ | 1,0129 | (0,3687;2,1743) | -66,403 | 0,453 | 0,729 |
| | $\beta$ | 1,2932 | (0,1107;3,5095) | | 0,223 | 0,255 |
| | $\gamma$ | 1,0321 | (0,5337;1,6854) | | 0,330 | 0,699 |
| | $\zeta_0$ | -0,3589 | (-1,7662;0,9063) | | 0,323 | 0,437 |
| | $\zeta_1$ | -0,6709 | (-1,3527;0,0187) | | 0,152 | 0,316 |
| | $\zeta_2$ | -0,9889 | (-1,6274;-0,4056) | | 0,146 | 0,410 |
| | $\eta_0$ | -0,1337 | (-0,7502;0,4156) | | 0,645 | 0,367 |
| | $\eta_1$ | -0,4843 | (-1,3096;0,4008) | | 0,619 | 0,710 |
| | $\eta_2$ | -0,9124 | (-1,7470;-0,0844) | | 0,475 | 0,149 |
| EW | $\alpha$ | 0,9920 | (0,3375;1,9698) | -66,339 | 0,606 | 0,843 |
| | $\gamma$ | 1,0456 | (0,5543;1,6848) | | 0,924 | 0,926 |
| | $\zeta_0$ | -0,5640 | (-1,4830;0,2432) | | 0,486 | 0,475 |
| | $\zeta_1$ | -0,6878 | (-1,3867;-0,0267) | | 0,209 | 0,480 |
| | $\zeta_2$ | -1,0025 | (-1,6199;-0,4032) | | 0,373 | 0,363 |
| | $\eta_0$ | -0,1461 | (-0,7618;0,4122) | | 0,923 | 0,760 |
| | $\eta_1$ | -0,4688 | (-1,3263;0,3899) | | 0,688 | 0,204 |
| | $\eta_2$ | -0,9032 | (-1,7279;-0,0601) | | 0,743 | 0,240 |
| BE | $\alpha$ | 1,0645 | (0,7532;1,3949) | -65,521 | 0,484 | 0,006 |
| | $\beta$ | 1,2022 | (0,1268;3,2823) | | 0,185 | 0,174 |
| | $\zeta_0$ | -0,4555 | (-1,8665;0,7963) | | 0,278 | 0,448 |
| | $\zeta_1$ | -0,6561 | (-1,3152;0,0242) | | 0,803 | 0,282 |
| | $\zeta_2$ | -0,9876 | (-1,6318;-0,4276) | | 0,088 | 0,312 |
| | $\eta_0$ | -0,1358 | (-0,7279;0,4218) | | 0,378 | 0,409 |
| | $\eta_1$ | -0,4753 | (-1,3250;0,3548) | | 0,958 | 0,535 |
| | $\eta_2$ | -0,9032 | (-1,7219;-0,0927) | | 0,937 | 0,614 |
| Weibull | $\gamma$ | 1,0372 | (0,8458;1,2476) | -65,557 | 0,100 | 0,265 |
| | $\zeta_0$ | -0,5784 | (-1,0565;-0,0332) | | 0,272 | 0,545 |
| | $\zeta_1$ | -0,6722 | (-1,3587;-0,0064) | | 0,960 | 0,942 |
| | $\zeta_2$ | -1,0041 | (-1,6542;-0,4332) | | 0,747 | 0,741 |
| | $\eta_0$ | -0,1378 | (-0,7671;0,4015) | | 0,278 | 0,567 |
| | $\eta_1$ | -0,4783 | (-1,3794;0,3407) | | 0,427 | 0,400 |
| | $\eta_2$ | -0,9051 | (-1,7219;-0,0499) | | 0,268 | 0,378 |

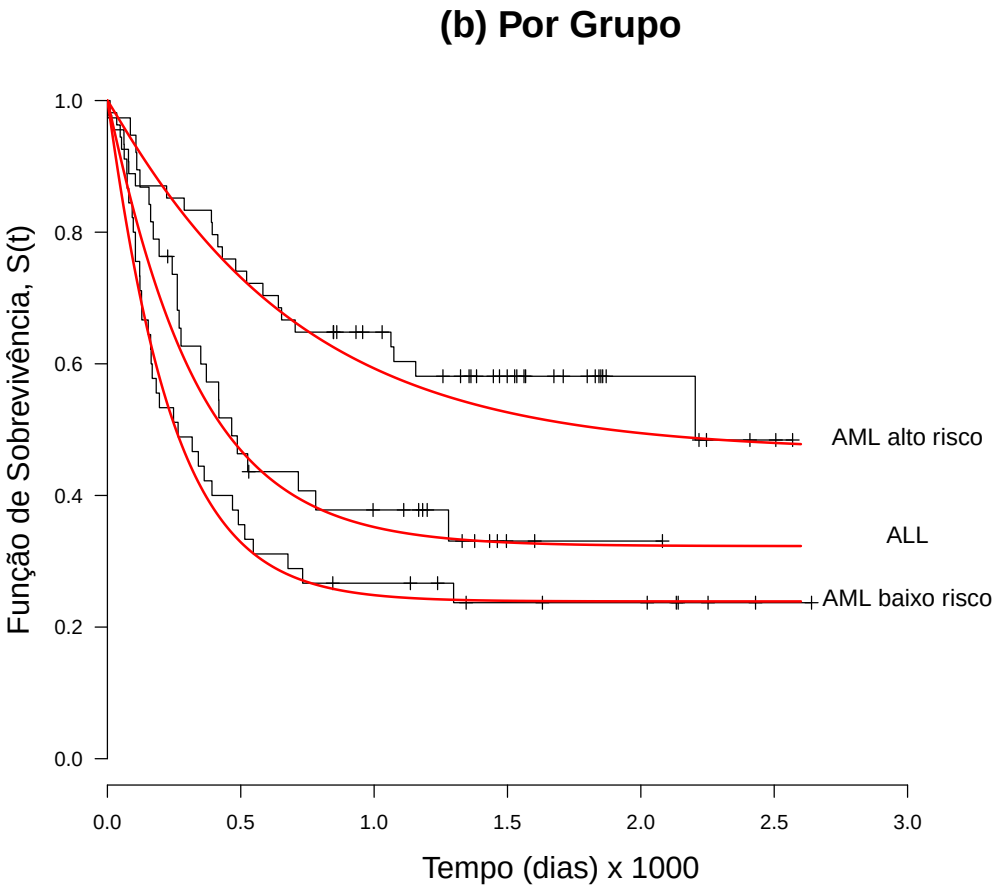


Figura 15 – Curvas ajustadas do modelo de mistura baseado na distribuição Beta-Weibull considerando os dados de transplantes de medula por grupo.

## 6.4 Discussão

É possível o uso da distribuição beta-Weibull na análise de dados médicos, contudo como já observado por Wahed et al. (2009) há obstáculos na aplicação do modelo baseado na BW, visto que sua função de sobrevivência não possui forma fechada, sendo necessário o uso de técnicas de integração numérica para estimar seus parâmetros. Este problema pode se tornar mais crítico, se for considerado covariáveis, aumentando assim a complexidade da função de verossimilhança.

A metodologia Bayesiana pode ser uma ferramenta mais parcimoniosa que a frequentistas, ao estimar os parâmetros da distribuição BW, e até mesmo de seus casos particulares. Quando o interesse é estimar a fração de cura, ambos estimaram de forma adequada, entretanto o método frequentista apresenta, em certos casos, intervalos de confiança assintóticos que não são plausíveis aos parâmetros. Dessa forma as estimativas Bayesianas são mais adequadas.

Pode-se perceber uma diferença mais clara entre as metodologias de estimação quando há presença de covariáveis, mesmo encontrando dificuldades na metodologia

Bayesiana (devido a presença das funções beta e beta incompleta). Diferente do casos da metodologia frequentista, em que não foi possível estimar valores ao inserir covariáveis. Isto pode ser um problema exclusivo a este banco de dados, ou da método de otimização utilizado.

Quanto ao detalhe da implementação Bayesiana, o uso do pacote *MCMCpack* mostrou-se eficiente na presença das complexas funções verossimilhanças. As estimativas foram obtidas com velocidade considerável e apresentaram resultados consistentes.

Portanto, o uso da inferência Bayesiana além da conhecida vantagem de incorporar uma conhecimento a *priori* dos parâmetros, principalmente na fração de cura, demonstrou vantagem na estimação dos parâmetros e obtenção de intervalos HPD fidedignos.

# Capítulo 7

# Considerações Finais

As comparações entre estimativas clássicas e Bayesianas em ambas as distribuições apresentadas no trabalho (Weibull modificada e Beta-Weibull), tanto com mistura e não mistura, mostrou que a fração de cura é bem estimada em todos os casos. Contudo, o método clássico pode apresentar maiores problemas numéricos do que o Bayesiano. Isto acontece principalmente quanto se observa a presença de covariáveis e, em alguns casos, o método clássico não apresentou intervalos de confiança assintóticos adequados. Uma solução mais adequada é a possível utilização de intervalos de confiança perfilados.

Nota-se também a importância de modelos mais complexo ao trabalho com dados médicos reais, visto que modelos mais simples podem não obter estimativas tão fidedignas. Isto evidencia a importância da informação de técnicas mais sofisticadas aos profissionais da pesquisa médica.

Um outro fator importante foi a necessidade de fazer a escolha e a verificação dos modelos de acordo com o comportamento empírico da curva de risco dos dados, evitando dessa forma o uso de distribuições que não comportem a forma mais adequada ao conjunto de dados que está sendo estudado.

Finalmente, observou-se a importância de diferentes métodos de estimação, sendo que o método Bayesiano apresentou no geral um melhor desempenho que o método frequentista. Mesmo nos casos em que as estimativas pontuais são semelhantes, o método Bayesiano não apresentou problema significativo de estimação e calculou intervalos HPD adequados aos casos estudados.

# Capítulo 8

# Estudos Futuros

Como objetivos futuros, espera-se atingir os seguintes pontos:

(a) Testar a adequação dos modelos apresentados considerando dados simulados. Neste caso, pode-se estudar a adequação dos modelos considerando diferentes tamanhos amostrais e valores para os parâmetros.

(b) Aprofundar o estudo da estrutura dos modelos, explorando outras funções de ligação (como a probito ou a logito) e a inserção de vetores de mais de uma covariável, utilizando dados reais como exemplos.

(c) Aprofundar o estudo dos modelos Bayesianos, considerando outras distribuições *a priori* para os parâmetros de interesse, a fim de fazer uma análise de sensibilidade para os hiperparâmetros escolhidos e explorar o uso de outros algoritmos computacionais.

(d) Ampliar a aplicação do modelo de mistura e não mistura na presença de fração de cura para outras distribuições.

(e) Construir um pacote no *software R* com o objetivo de estimar fração de cura para diversas distribuições.

# Referências

AARSET, M. V. The null distribution for a test of constant versus "bathtub" failure rate. **Scandinavian Journal of Statistics**, p. 55–61, 1985.

ACHCAR, J. A.; COELHO-BARROS, E. A.; MAZUCHELI, J. Cure fraction models using mixture and non-mixture models. **Tatra Mountains Mathematical Publications**, v. 51, n. 1, p. 1–9, 2012.

AKAIKE, H. A new look at the statistical model identification. **Automatic Control, IEEE Transactions on**, Ieee, v. 19, n. 6, p. 716–723, 1974.

ANGELIS, R. D.; CAPOCACCIA, R.; HAKULINEN, T.; SODERMAN, B.; VERDECCHIA, A. Mixture models for cancer survival analysis: application to population-based data with covariates. **Statistics in Medicine**, Wiley Online Library, v. 18, n. 4, p. 441–454, 1999.

BARLOW, R. E.; CAMPO, R. A. **Total Time on Test Processes and Applications to Failure Data Analysis.** [S.l.]: Defense Technical Information Center, 1975.

BAYES, M. P. M. An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, f. r. s. communicated by mr. price, in a letter to john canton, a. m. f. r. s. **Philosophical Transactions (1683-1775)**, The Royal Society, v. 53, p. 370–418, 1763. ISSN 02607085.

BOAG, J. W. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, v. 11, n. 1, p. 15–53, 1949.

BOLFARINE, H.; RODRIGUES, J.-F.; ACHCAR, J. **Análise de sobrevivência**. [S.l.]: ABE/IMUFRJ, 1991.

BOLFARINE, H.; SANDOVAL, M. C. **Introdução à inferência estatística**. [S.l.]: SBM, 2001.

BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. [S.l.]: John Wiley & Sons, 2011. v. 40.

BRADBURN, M.; CLARK, T.; LOVE, S.; ALTMAN, D. Survival analysis part II: multivariate data analysis–an introduction to concepts and methods. **British Journal of Cancer**, Nature Publishing Group, v. 89, n. 3, p. 431, 2003.

BRESLOW, N.; CROWLEY, J. et al. A large sample study of the life table and product limit estimates under random censorship. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 2, n. 3, p. 437–453, 1974.

BROOKS, S. **Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde**. [S.l.]: WILEY-BLACKWELL 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, 2002.

CANAVOS, G. C.; TAOKAS, C. P. Bayesian estimation of life parameters in the Weibull distribution. **Operations Research**, INFORMS, v. 21, n. 3, p. 755–763, 1973.

CANCHO, V. G.; ORTEGA, E. M. M.; BARRIGA, G. D. C. Comparison of modified Weibull models when bathtub-shaped failure rates. **Revista de Matemática e Estatística**, v. 25, n. 2, p. 111–136, 2007.

CARLIN, B. P.; LOUIS, T. A. **Bayes and empirical Bayes methods for data analysis**. [S.l.]: Chapman & Hall/CRC Boca Raton, 2000.

CARRASCO, J. M.; ORTEGA, E. M.; CORDEIRO, G. M. A generalized modified Weibull distribution for lifetime modeling. **Computational Statistics & Data Analysis**, Elsevier, v. 53, n. 2, p. 450–462, 2008.

CASELLA, G.; GEORGE, E. I. Explaining the Gibbs sampler. **The American Statistician**, Taylor & Francis, v. 46, n. 3, p. 167–174, 1992.

CHEN, M.-H.; IBRAHIM, J. G.; SINHA, D. Bayesian inference for multivariate survival data with a cure fraction. **Journal of Multivariate Analysis**, Elsevier, v. 80, n. 1, p. 101–126, 2002.

CHEN, M.-H.; SHAO, Q.-M.; IBRAHIM, J. G. **Monte Carlo methods in Bayesian computation**. [S.l.]: Springer Science & Business Media, 2012.

CHIB, S.; GREENBERG, E. Understanding the Metropolis-Hastings algorithm. **The american statistician**, Taylor & Francis Group, v. 49, n. 4, p. 327–335, 1995.

COHEN, A. C. Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples. **Technometrics**, Taylor & Francis Group, v. 7, n. 4, p. 579–588, 1965.

COLLETT, D. **Modelling survival data in medical research**. [S.l.]: CRC press, 2015.

COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. [S.l.]: Edgard Blücher, 2006. (ABE - Projeto Fisher).

COOK, R. J.; LAWLESS, J. F. **The statistical analysis of recurrent events**. [S.l.]: Springer Science & Business Media, 2007.

COPELAN, E. A.; BIGGS, J. C.; THOMPSON, J. M.; CRILLEY, P.; SZER, J.; KLEIN, J. P.; KAPOOR, N.; AVALOS, B. R.; CUNNINGHAM, I.; ATKINSON, K. Treatment for acute myelocytic leukemia with allogeneic bone marrow transplantation following preparation with bucy2. **Blood**, Am Soc Hematology, v. 78, n. 3, p. 838–843, 1991.

CORBIERE, F.; COMMENGES, D.; TAYLOR, J. M.; JOLY, P. A penalized likelihood approach for mixture cure models. **Statistics in medicine**, Wiley Online Library, v. 28, n. 3, p. 510–524, 2009.

CORDEIRO, G. Introdução à teoria de verossimilhança. UFRJ, ABE. **10º Simpósio Nacional de Probabilidade e Estatística, X SINAPE, Rio de Janeiro**, 1992.

CORDEIRO, G. M.; DEMÉTRIO, C. G. Modelos lineares generalizados e extensoes. **Recife: UFRPE**, 2008.

CORDEIRO, G. M.; GOMES, A. E.; SILVA, C. Q. da; ORTEGA, E. M. The beta exponentiated Weibull distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 83, n. 1, p. 114–138, 2013.

CORDEIRO, G. M.; NADARAJAH, S.; ORTEGA, E. M. General results for the beta Weibull distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 83, n. 6, p. 1082–1114, 2013.

CORDEIRO, G. M.; SIMAS, A. B.; STOŠIC, B. D. Closed form expressions for moments of the beta Weibull distribution. **Anais da Academia Brasileira de Ciências**, SciELO Brasil, v. 83, n. 2, p. 357–373, 2011.

COWLES, M. K.; CARLIN, B. P. Markov chain Monte Carlo convergence diagnostics: a comparative review. **Journal of the American Statistical Association**, Taylor & Francis, v. 91, n. 434, p. 883–904, 1996.

DICKEN, B. J.; BIGAM, D. L.; CASS, C.; MACKEY, J. R.; JOY, A. A.; HAMILTON, S. M. Gastric adenocarcinoma: review and considerations for future directions. **Annals of surgery**, Lippincott, Williams, and Wilkins, v. 241, n. 1, p. 27, 2005.

EHLERS, R. S. **Introdução à inferência bayesiana**. [S.l.: s.n.], 2007.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 31, n. 4, p. 497–512, 2002.

FAMOYE, F.; LEE, C.; OLUMOLADE, O. The beta-Weibull distribution. **Journal of Statistical Theory and Applications**, v. 4, n. 2, p. 121–136, 2005.

FISHER, R. A. On an absolute criterion for fitting frequency curves. Messenger of Mathematics, v. 41, p. 155–160, 1912.

FRÉCHET, M. Sur la loi de probabilité de l'écart maximum. In: **Annales de la societe Polonaise de Mathematique**. [S.l.: s.n.], 1927. v. 6, p. 93–116.

GEISSER, S.; EDDY, W. F. A predictive approach to model selection. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 74, n. 365, p. 153–160, 1979.

GELFAND, A. E.; DEY, D. K.; CHANG, H. **Model determination using predictive distributions with implementation via sampling-based methods**. [S.l.], 1992.

GELFAND, A. E.; HILLS, S. E.; RACINE-POON, A.; SMITH, A. F. Illustration of bayesian inference in normal data models using Gibbs sampling. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 85, n. 412, p. 972–985, 1990.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. [S.l.]: Taylor & Francis, 2014. v. 2.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical science**, JSTOR, p. 457–472, 1992.

GEWEKE, J. et al. **Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments**. [S.l.]: Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991. v. 196.

GILKS, W. R. **Markov chain monte carlo**. [S.l.]: Wiley Online Library, 2005.

HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, INFORMS, v. 31, n. 6, p. 1109–1144, 1983.

HENNINGSEN, A.; TOOMET, O. maxlik: A package for maximum likelihood estimation in *R*. **Computational Statistics**, v. 26, n. 3, p. 443–458, 2011.

HJORTH, U. A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. **Technometrics**, Taylor & Francis, v. 22, n. 1, p. 99–107, 1980.

IBRAHIM, J. G.; CHEN, M.-H.; SINHA, D. Bayesian semiparametric models for survival data with a cure fraction. **Biometrics**, JSTOR, p. 383–388, 2001.

JÁCOME, A. A.; WOHNRATH, D. R.; NETO, C. S.; FREGNANI, J. H. T.; QUINTO, A. L.; OLIVEIRA, A. T.; VAZQUEZ, V. L.; FAVA, G.; MARTINEZ, E. Z.; SANTOS, J. S. Effect of adjuvant chemoradiotherapy on overall survival of gastric cancer patients submitted to d2 lymphadenectomy. **Gastric Cancer**, Springer, v. 16, n. 2, p. 233–238, 2013.

JAYAATA, K.; MOHAN, D.; TAPAS, S. **An introduction to Bayesian analysis**. [S.l.]: Springer, 2006.

JIANG, H.; XIE, M.; TANG, L. Markov chain monte carlo methods for parameter estimation of the modified Weibull distribution. **Journal of Applied Statistics**, Taylor & Francis, v. 35, n. 6, p. 647–658, 2008.

KALBFLEISCH, J. D.; PRENTICE, R. L. **The statistical analysis of failure time data**. [S.l.]: John Wiley & Sons, 2011. v. 360.

KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American statistical association**, Taylor & Francis, v. 53, n. 282, p. 457–481, 1958.

KLEIN, J. P.; MOESCHBERGER, M. L. **Survival Analysis: Techniques for Censored and Truncated Data**. 2nd. ed. [S.l.]: Springer, 2005. (Statistics for Biology and Health).

KLEINBAUM, D. G.; KLEIN, M. **Survival Analysis A Self Learning Text**. 2nd. ed. [S.l.]: Springer, 2005. (Statistics for Biology and Health).

LAI, C.; XIE, M.; MURTHY, D. A modified Weibull distribution. **Reliability, IEEE Transactions on**, IEEE, v. 52, n. 1, p. 33–37, 2003.

LAI, C. D. **Generalized Weibull Distributions**. [S.l.]: Springer, 2014.

LAMBERT, P. C.; THOMPSON, J. R.; WESTON, C. L.; DICKMAN, P. W. Estimating and modeling the cure fraction in population-based cancer survival analysis. **Biostatistics**, Biometrika Trust, v. 8, n. 3, p. 576–594, 2007.

LAWLESS, J. F. **Statistical Models and Methods for Lifetime Data**. 2. ed. [S.l.]: Wiley-Interscience, 2002. (Wiley Series in Probability and Statistics).

LEE, C.; FAMOYE, F.; OLUMOLADE, O. Beta-Weibull distribution: some properties and applications to censored data. **Journal of modern applied statistical methods**, v. 6, n. 1, p. 17, 2007.

LEE, E. T.; GO, O. T. Survival analysis in public health research. **Annual review of public health**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 18, n. 1, p. 105–134, 1997.

LESAFFRE, E.; LAWSON, A. B. **Bayesian biostatistics**. [S.l.]: John Wiley & Sons, 2012.

LIANG, F.; LIU, C.; CARROLL, R. **Advanced Markov chain Monte Carlo methods: learning from past samples**. [S.l.]: John Wiley & Sons, 2011. v. 714.

LUNN, D. J.; THOMAS, A.; BEST, N.; SPIEGELHALTER, D. Winbugs-a Bayesian modelling framework: concepts, structure, and extensibility. **Statistics and computing**, Springer, v. 10, n. 4, p. 325–337, 2000.

MALLER, R. A.; ZHOU, X. **Survival analysis with long-term survivors**. [S.l.]: Wiley New York, 1996.

MARTIN, A. D.; QUINN, K. M.; PARK, J. H. MCMCpack: Markov chain Monte Carlo in R. **Journal of Statistical Software**, v. 42, n. 9, p. 22, 2011.

MARTINEZ, E. Z.; ACHCAR, J. A.; JÁCOME, A. A.; SANTOS, J. S. Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data. **Computer methods and programs in biomedicine**, Elsevier, v. 112, n. 3, p. 343–355, 2013.

MENON, M. Estimation of the shape and scale parameters of the Weibull distribution. **Technometrics**, Taylor & Francis Group, v. 5, n. 2, p. 175–182, 1963.

MILLAR, R. B. **Maximum likelihood estimation and inference: with examples in R, SAS and ADMB**. [S.l.]: John Wiley & Sons, 2011. v. 111.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K. Exponentiated Weibull family for analyzing bathtub failure-rate data. **IEEE Transactions on Reliability**, Institute of Electrical and Electronics Engineers, v. 42, n. 2, p. 299–302, 1993.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; KOLLIA, G. D. A generalization of the weibull distribution with application to the analysis of survival data. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 91, n. 436, p. 1575–1583, 1996.

NADARAJAH, S. On the moments of the modified Weibull distribution. **Reliability Engineering & System Safety**, Elsevier, v. 90, n. 1, p. 114–117, 2005.

NADARAJAH, S.; KOTZ, S. The beta exponential distribution. **Reliability engineering & system safety**, Elsevier, v. 91, n. 6, p. 689–697, 2006.

NASSAR, M. M.; EISSA, F. H. On the exponentiated Weibull distribution. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 32, n. 7, p. 1317–1336, 2003.

NELSON, W. B. **Recurrent events data analysis for product repairs, disease recurrences, and other applications**. [S.l.]: Society for Industrial and Applied Mathematics, 2003. (ASA-SIAM series on statistics and applied probability).

NG, S.; MCLACHLAN, G. On modifications to the long-term survival mixture model in the presence of competing risks. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 61, n. 1-2, p. 77–96, 1998.

OKAMURA, T.; TSUJITANI, S.; KORENAGA, D.; HARAGUCHI, M.; BABA, H.; HIRAMOTO, Y.; SUGIMACHI, K. Lymphadenectomy for cure in patients with early gastric cancer and lymph node metastasis. **The American journal of surgery**, Elsevier, v. 155, n. 3, p. 476–480, 1988.

PAULINO, C. D. M.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística bayesiana**. [S.l.: s.n.], 2003.

PAWITAN, Y. **In all likelihood: statistical modelling and inference using likelihood**. [S.l.]: Oxford University Press, 2001.

PENG, Y.; DEAR, K. B. A nonparametric mixture model for cure rate estimation. **Biometrics**, JSTOR, p. 237–243, 2000.

PETO, R.; LEE, P.; PAIGE, W. Statistical analysis of the bio-assay of continuous carcinogens. **British journal of cancer**, Nature Publishing Group, v. 26, n. 4, p. 258, 1972.

PLUMMER, M.; BEST, N.; COWLES, K.; VINES, K. Coda: Convergence diagnosis and output analysis for mcmc. **R News**, v. 6, n. 1, p. 7–11, 2006. Disponível em: <http://CRAN.R-project.org/doc/Rnews/>.

PLUMMER, M.; BEST, N.; COWLES, K.; VINES, K.; SARKAR, D.; BATES, D.; ALMOND, R.; PLUMMER, M. M. Package 'coda'. 2015.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2014. Disponível em: <http://www.R-project.org/>.

ROSA, G. J. M. **Análise Bayesiana de modelos lineares mistos robustos via Amostrador de Gibbs**. 57 f. Tese (Doutorado em Estatística e Experimentação Agronômica) — Escola Superior de Agricultura Luiz de Queiroz da Universidade de São Paulo, Piracicaba, 1998.

ROSIN, P.; RAMMLER, E. The laws governing the fineness of powdered coal. **Journal of the Institute of Fuelr**, v. 7, p. 29–36, 1933.

ROSSI, R. M. **Introdução aos métodos Bayesianos na análise de dados zootécnicos com uso do WinBUGS e R**. [S.l.]: Maringá: Eduem, 2011.

SCHWARZ, G. et al. Estimating the dimension of a model. **The annals of statistics**, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978.

SHANNO, D. F. On broyden-fletcher-goldfarb-shanno method. **Journal of Optimization Theory and Applications**, v. 46, n. 1, p. 87–94, 1985.

SILVA, G. O.; ORTEGA, E. M.; CORDEIRO, G. M. The beta modified Weibull distribution. **Lifetime Data Analysis**, Springer, v. 16, n. 3, p. 409–430, 2010.

SMITH, R. L.; NAYLOR, J. A comparison of maximum likelihood and Bayesian estimators for the three-parameter weibull distribution. **Applied Statistics**, JSTOR, p. 358–369, 1987.

SORENSEN, D. Gibbs sampling in quantitative genetics. **Intern Rapport. Statens Husdyrbrugsforsoeg (Denmark). no. 82.**, 1996.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit. **Journal of the Royal Statistical Society: Series B (Statistical Methodology)**, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002.

SUGIURA, N. Further analysts of the data by Akaike's information criterion and the finite corrections: Further analysts of the data by Akaike's. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 7, n. 1, p. 13–26, 1978.

TABLEMAN, M.; KIM, J. S. **Survival Analysis Using S: Analysis of Time-to-Event Data**. 1. ed. [S.l.]: Chapman and Hall/CRC, 2003. (Chapman & Hall/CRC Texts in Statistical Science).

THOMAN, D. R.; BAIN, L. J.; ANTLE, C. E. Inferences on the parameters of the Weibull distribution. **Technometrics**, Taylor & Francis, v. 11, n. 3, p. 445–460, 1969.

TSODIKOV, A.; IBRAHIM, J.; YAKOVLEV, A. Estimating cure rates from survival data. **Journal of the American Statistical Association**, v. 98, n. 464, 2003.

WAHED, A. S.; LUONG, T. M.; JEONG, J.-H. A new generalization of Weibull distribution with application to a breast cancer data set. **Statistics in medicine**, Wiley Online Library, v. 28, n. 16, p. 2077–2094, 2009.

WEIBULL, W. A statistical distribution function of wide applicability. **Journal of Applied Mechanics**, p. 293–297, 1951.

XIE, M.; TANG, Y.; GOH, T. N. A modified weibull extension with bathtub-shaped failure rate function. **Reliability Engineering & System Safety**, Elsevier, v. 76, n. 3, p. 279–285, 2002.

ZEGER, S. L.; KARIM, M. R. Generalized linear models with random effects; a gibbs sampling approach. **Journal of the American statistical association**, Taylor & Francis Group, v. 86, n. 413, p. 79–86, 1991.

# Apêndice A

# Dados Câncer Gástrico

Tabela 12 – Dados Câncer Gástrico

| tempo | d | x | tempo | d | x | tempo | d | x | tempo | d | x | tempo | d | x | tempo | d | x |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.63 | 1 | 0 | 1.61 | 1 | 0 | 20.76 | 0 | 1 | 22.86 | 1 | 1 | 4.01 | 1 | 0 | 7.89 | 1 | 1 |
| 35 | 0 | 0 | 17.7 | 1 | 1 | 24.28 | 0 | 1 | 36 | 0 | 0 | 14.77 | 1 | 1 | 25.56 | 0 | 1 |
| 10.72 | 1 | 1 | 15.43 | 1 | 0 | 26.15 | 0 | 0 | 8.95 | 1 | 1 | 9.7 | 1 | 0 | 8.36 | 1 | 0 |
| 30.69 | 0 | 1 | 24.21 | 1 | 1 | 11.41 | 1 | 0 | 36 | 0 | 1 | 17.43 | 1 | 0 | 29.97 | 0 | 0 |
| 24.97 | 0 | 1 | 35.72 | 0 | 1 | 25.79 | 0 | 1 | 36 | 0 | 1 | 27.89 | 0 | 1 | 6.91 | 1 | 0 |
| 19.84 | 0 | 0 | 35.59 | 0 | 0 | 16.94 | 0 | 0 | 36 | 0 | 0 | 23.55 | 0 | 0 | 15.76 | 0 | 0 |
| 32.5 | 0 | 0 | 19.93 | 1 | 0 | 25.59 | 0 | 1 | 25.79 | 1 | 1 | 13.78 | 1 | 1 | 10.76 | 1 | 0 |
| 31.38 | 0 | 0 | 36 | 0 | 0 | 21.88 | 0 | 0 | 26.32 | 0 | 0 | 20.49 | 0 | 0 | 7.7 | 1 | 0 |
| 0.2 | 1 | 0 | 32.53 | 0 | 1 | 19.38 | 0 | 1 | 25.33 | 0 | 0 | 18.39 | 1 | 1 | 14.38 | 0 | 0 |
| 29.08 | 0 | 1 | 28.98 | 0 | 0 | 8.39 | 1 | 0 | 14.05 | 1 | 0 | 25.36 | 0 | 0 | 35.3 | 0 | 0 |
| 22.14 | 1 | 0 | 23.52 | 0 | 0 | 36 | 0 | 1 | 7.93 | 0 | 0 | 20.49 | 0 | 1 | 9.05 | 1 | 1 |
| 31.64 | 0 | 1 | 24.57 | 0 | 0 | 3.49 | 0 | 0 | 21.84 | 0 | 0 | 17.14 | 1 | 0 | 16.38 | 1 | 1 |
| 0.26 | 1 | 0 | 21.48 | 1 | 0 | 36 | 0 | 1 | 17.34 | 0 | 1 | 36 | 0 | 1 | 24.8 | 0 | 0 |
| 25.76 | 0 | 1 | 11.88 | 1 | 0 | 1.45 | 0 | 0 | 4.93 | 1 | 0 | 29.31 | 0 | 0 | 32.47 | 0 | 1 |
| 23.85 | 0 | 0 | 32.4 | 0 | 1 | 18.62 | 1 | 0 | 30.49 | 0 | 0 | 12.83 | 1 | 1 | 34.08 | 0 | 0 |
| 24.21 | 0 | 1 | 31.84 | 0 | 1 | 24.31 | 1 | 1 | 2.96 | 1 | 0 | 25.26 | 0 | 0 | 2.73 | 1 | 0 |
| 36 | 0 | 1 | 34.41 | 0 | 0 | 13.13 | 0 | 0 | 31.68 | 0 | 0 | 19.14 | 1 | 0 | 16.18 | 1 | 0 |
| 34.74 | 0 | 1 | 13.95 | 1 | 0 | 4.54 | 1 | 0 | 29.31 | 0 | 1 | 23.82 | 0 | 1 | 30.63 | 0 | 0 |
| 33.36 | 0 | 0 | 26.32 | 1 | 0 | 6.15 | 1 | 0 | 0.53 | 1 | 0 | 29.28 | 0 | 0 | 33.09 | 0 | 0 |
| 4.67 | 1 | 0 | 6.55 | 0 | 0 | 36 | 0 | 0 | 8.78 | 1 | 0 | 2.89 | 1 | 0 | 16.09 | 1 | 0 |
| 35.2 | 0 | 0 | 33.88 | 0 | 1 | 28.59 | 1 | 1 | 13.09 | 1 | 1 | 0.66 | 1 | 0 | 21.02 | 1 | 1 |
| 33.65 | 0 | 0 | 10.03 | 1 | 0 | 36 | 0 | 1 | 0.1 | 1 | 0 | 0.3 | 1 | 0 | 3.32 | 1 | 0 |
| 22.99 | 0 | 0 | 28.65 | 0 | 1 | 36 | 0 | 0 | 0.49 | 1 | 0 | 21.38 | 0 | 0 | 10.53 | 1 | 0 |
| 34.21 | 0 | 0 | 35.03 | 0 | 0 | 36 | 0 | 1 | 14.01 | 1 | 0 | 0.33 | 1 | 0 | 16.51 | 1 | 1 |
| 28.22 | 0 | 1 | 13.82 | 1 | 1 | 2.63 | 0 | 0 | 8.85 | 1 | 1 | 12.5 | 1 | 0 | 36 | 0 | 0 |
| 17.24 | 1 | 0 | 19.44 | 1 | 0 | 30.95 | 0 | 1 | 36 | 0 | 0 | 36 | 0 | 0 | 11.97 | 1 | 1 |
| 26.78 | 0 | 0 | 12.5 | 1 | 1 | 19.21 | 1 | 1 | 0.56 | 1 | 0 | 35.89 | 0 | 1 | 7.17 | 1 | 0 |
| 23.39 | 0 | 1 | 30.16 | 0 | 0 | 36 | 0 | 0 | 26.05 | 1 | 1 | 14.7 | 1 | 1 | 2.8 | 1 | 0 |
| 25.33 | 0 | 1 | 23.39 | 0 | 0 | 17.07 | 1 | 1 | 36 | 0 | 0 | 0.23 | 1 | 0 | 1.78 | 1 | 0 |
| 5.76 | 1 | 1 | 4.67 | 1 | 0 | 17.14 | 1 | 1 | 36 | 0 | 1 | 9.28 | 1 | 0 | 8.91 | 1 | 0 |
| 25.3 | 1 | 0 | 10.2 | 1 | 0 | 36 | 0 | 0 | 30.26 | 1 | 1 | 36 | 0 | 1 | 14.34 | 1 | 0 |
| 33.91 | 0 | 0 | 8.32 | 1 | 0 | 0.66 | 1 | 0 | 27.37 | 0 | 0 | 32.8 | 0 | 0 | | | |
| 16.94 | 1 | 0 | 1.18 | 1 | 0 | 36 | 0 | 1 | 24.01 | 0 | 0 | 36 | 0 | 1 | | | |
| 34.9 | 0 | 1 | 13.49 | 1 | 1 | 36 | 0 | 1 | 9.47 | 1 | 1 | 25.23 | 1 | 1 | | | |

# Apêndice B

# Códigos Inferência Bayesiana

Para estimar os parâmetros da distribuição Weibull modificada com mistura no OpenBUGS foi utilizado o seguinte código:

```
model;
{
for (i in 1:m) {
f0t[i]  <-alpha*(pow(t[i],beta-1))*(beta+lambda*(t[i]))*exp(lambda*t[i])*
exp(-alpha*(pow(t[i],beta))*exp(lambda*t[i]))
s0t[i]   <- exp(-alpha*(pow(t[i],beta))*exp(lambda*t[i]))
L[i]     <- pow((1-pi)*f0t[i],d[i])*pow(pi+(1-pi)*s0t[i],1-d[i])
logL[i]  <- log(L[i])
zeros[i] <- 0
zeros[i] ~ dloglik(logL[i])
CPO[i]   <-1/L[i]
}
pi     ~ dbeta(1,1)
alpha  ~ dgamma(1,1)
beta   ~ dgamma(1,1)
lambda ~ dgamma(1,1)
}
```

Assumindo um modelo de não mistura deve-se apenas trocar a linha `L[i]` por

```
F0[i]  <- 1-S0[i]
h[i]   <- -(log(p))*f0[i]
L[i]   <- pow(h[i],d[i])*exp(F0[i]*log(p)).
```

Agora ao estimar os parametros da distribuição beta-Weibull no pacote MCMCpack utilizou-se o seguinte código

```
log.post<-function(t,d,theta)
{
alpha  <- theta[1]
Beta   <- theta[2]
gamma  <- theta[3]
lambda <- theta[4]
p      <- theta[5]
    if(theta[1]<=0)return(-Inf)
    if(theta[2]<=0)return(-Inf)
    if(theta[3]<=0)return(-Inf)
    if(theta[4]<=0)return(-Inf)
    if(theta[5]<=0)return(-Inf)
f0t   <-gamma/(lambda^gamma*beta(alpha,Beta))*t^(gamma-1)*
    exp(-Beta*(t/lambda)^gamma)*(1-exp(-(t/lambda)^gamma))^(alpha-1)
gt    <- 1-exp(-(t/lambda)^gamma)
S0t   <- 1-pbeta(gt,alpha,Beta)
log.like <-log(1-p)*sum(d)+sum(d*log(f0t))+sum((1-d)*
log(p+(1-p)*S0t))
priori  <-dbeta(p,1,1)*dgamma(alpha,1,1)*dgamma(Beta,1,1)*
dgamma(gamma,1,1)*dgamma(lambda,1,1)
log.priori<-log(priori)
L<-log.like+log.priori
if (is.na(L)==TRUE){return(-Inf)}else{return(L)}
}
```

Se um modelo de mistura for considerado basta trocar `log.lik` por

```
log.like <- log(-log(p))*sum(d)+sum(d*lof(f0t))+log(p)*sum(1-S0t).
```

# Apêndice C

# Códigos Inferência Frequentista

As estimativas frequentistas e intervalos de confiança assintóticos e perfilados para a distribuição Weibull modificada com mistura foram obtidas com o seguinte código:

```
######### Estimativas frequetistas e Intervalos de Confianças #########
WM   <- function(param) {
alpha  <- param[1]
beta   <- param[2]
lambda <- param[3]
fc     <- param[4]
if(param[1]<=0)return(-Inf)
if(param[2]<=0)return(-Inf)
if(param[3]<=0)return(-Inf)
if(param[4]<=0)return(-Inf)
S0t    <- exp(-alpha*(t^beta) * exp(lambda*t))
f0t    <- alpha*t^(beta-1)*(beta+lambda*t)*exp(lambda*t-alpha*t^beta*exp(lambda*t))
L    <- ((1-fc)*f0t)^d * (fc+(1-fc)*S0t)^(1-d)
logL <- sum(log(L))
if(is.na(logL)==TRUE){return(-Inf)}else{return(logL) }}

mle   <- maxLik(logLik = WM, start = c(0.1,0.5,0.05,0.5), method="NR")

summary(mle)

############### Intervalos de Confiança Assintótico 95% ################
k <- length(mle$estimate)
```

```
ic95 <- matrix(NA,nrow=k,ncol=2)

for (j in 1:k) {

ic95[j,1] <- mle$estimate[j] - qnorm(0.975)*sqrt(vcov(mle)[j,j])
ic95[j,2] <- mle$estimate[j] + qnorm(0.975)*sqrt(vcov(mle)[j,j])

}

ic95

############### Intervalos de Confiança Perfilado  95% ################

log.lik.WM  <- function(a,b,l,fc){
S0t    <- exp(-a*(t^b) * exp(l*t))
f0t    <- a*t^(b-1)*(b+l*t)*exp(l*t-a*t^b*exp(l*t))
L    <- ((1-fc)*f0t)^d * (fc+(1-p)*S0t)^(1-d)
logL <- sum(log(L))
return(logL)}

est   <- mle$estimate #estimativas

lcomp<-log.lik.WM(a=est[1],b=est[2],l=est[3],fc=est[4])

corte<-lcomp-1.92

s<-seq(0,2,0.0001)
z<-seq(0,1,0.0001)

f<-NA
g<-NA

########################## Intervalo Alpha ##########################
for(i in 1:length(s)){
f[i]<-log.lik.WM(s[i],b=est[2],l=est[3],fc=est[4])}

ff<-which(f>=corte)

ffmin<-min(ff)
```

```
ffmax<-max(ff)

alpha.inf<-s[ffmin]
alpha.sup<-s[ffmax]

########################## Intervalo Beta ########################
for(i in 1:length(s)){
f[i]<-log.lik.WM(a=est[1],s[i],l=est[3],fc=est[4])}

ff<-which(f>=corte)

ffmin<-min(ff)
ffmax<-max(ff)

beta.inf<-s[ffmin]
beta.sup<-s[ffmax]

########################## Intervalo Lambda ######################
for(i in 1:length(s)){
f[i]<-log.lik.WM(a=est[1],b=est[2],s[i],fc=est[4])}

ff<-which(f>=corte)

ffmin<-min(ff)
ffmax<-max(ff)

lambda.inf<-s[ffmin]
lambda.sup<-s[ffmax]

########################## Intervalo p ###########################
for(i in 1:length(z)){
g[i]<-log.lik.WM(a=est[1],b=est[2],l=est[3],z[i])}

gg<-which(g>=corte)

ggmin<-min(gg)
ggmax<-max(gg)

fc.inf<-s[ggmin]
```

```
fc.sup<-s[ggmax]

########################## Intervalos ############################

IC95<-matrix(c(alpha.inf,alpha.sup,beta.inf,beta.sup,lambda.inf,
lambda.sup,fc.inf,fc.sup),ncol=2,byrow = T,
dimnames = list(c("alpha","beta","lambda","fc"),
c("Inf","Sup")))
IC95
```

No caso de assumir um modelo de não mistura basta substituir as linhas `L` por

```
F0t  <- 1-S0t
L    <- ((-log(fc)*f0t)^d)*(exp(log(fc)*F0t)).
```