

Modelo de Mistura Padrão de Longa Duração com Censura Uniforme-Exponencial

Josenildo de Souza Chaves

Modelo de Mistura Padrão de Longa Duração com Censura Uniforme-Exponencial

Josenildo de Souza Chaves

Orientador: Prof. Dr. Josemar Rodrigues

Tese apresentada ao Departamento de Estatística da Universidade Federal de São Carlos – DEs – UFSCar, como parte dos requisitos para obtenção do título de Doutor em Estatística.

UFSCar – São Carlos – SP

Março de 2010

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

C512mm

Chaves, Josenildo de Souza.

Modelo de mistura padrão de longa duração com censura uniforme-exponencial / Josenildo de Souza Chaves. -- São Carlos : UFSCar, 2010.

92 f.

Tese (Doutorado) -- Universidade Federal de São Carlos, 2010.

1. Análise de sobrevivência. 2. Mistura de distribuições. 3. Inferência bayesiana. 4. Simulação estocástica. I. Título.

CDD: 519.9 (20^a)

Josenildo de Souza Chaves

Modelo de Mistura Padrão de Longa Duração com Censura Uniforme-Exponencial

Tese apresentada à Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Doutor em Estatística.

Aprovada em 25 de março de 2010.

BANCA EXAMINADORA

Presidente



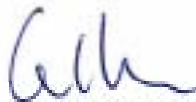
Prof. Dr. Josemar Rodrigues (DEs-UFSCar/Orientador)

1º Examinador



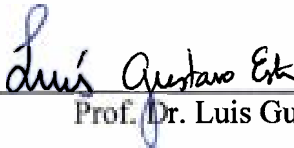
Prof. Dr. Antonio Carlos Pedroso de Lima (IME-USP)

2º Examinador



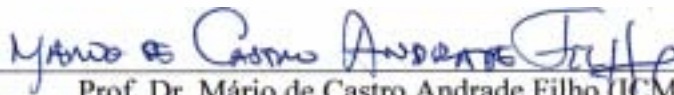
Prof. Dr. Enrico Antonio Colosimo (UFMG)

3º Examinador



Prof. Dr. Luis Gustavo Esteves (IME-USP)

4º Examinador



Prof. Dr. Mário de Castro Andrade Filho (ICMC-USP)

Aos meus pais José e Branca e a
minha esposa Eleni.

Agradecimentos

Agradeço a Deus que me conduziu firmemente durante todo o período de trabalho. Desejo também agradecer a muitas pessoas que foram importantes para o sucesso deste trabalho. Ao Prof. Dr. Josemar Rodrigues pela orientação segura e dedicada, pela ativa participação e motivação durante toda a pesquisa. Aos professores da Pós-Graduação, especialmente ao Prof. Dr. Francisco Louzada Neto e ao Prof. Dr. Jorge Achcar que conjuntamente com o Prof. Dr. Josemar Rodrigues, contribuíram significativamente para a minha formação acadêmica com os cursos por eles ministrados. Ao Prof. Dr. Mário de Castro e ao Prof. Dr. Heleno Bolfarine pelas várias sugestões e comentários decorrentes do exame de qualificação. À Prof.^a Dr.^a Vera Tomazella e família pelo incentivo e apoio à minha vinda e permanência na cidade de São Carlos. Aos amigos do Curso com os quais nestes quatro anos de estudo pude compartilhar ideias e principalmente pela amizade. Aos funcionários do DEs-UFSCar e Biblioteca Comunitária da UFSCar pela atenção e serviços prestados. À minha esposa Eleni Andrade pela paciência, carinho e motivação em todos os momentos. Aos meus pais José e Branca e aos meus irmãos pelas palavras de apoio e incentivo neste período de muitas ausências. Aos meus colegas professores e funcionários do Departamento de Matemática da UFMA pelo constante incentivo que tenho recebido. À CAPES pelo suporte financeiro concedido. Agradeço também a outras pessoas e instituições que não estão citadas e que de alguma forma contribuíram para a realização deste trabalho.

Muito obrigado,
Josenildo Chaves.

“Não devemos esquecer-nos do velho axioma de que, quando todas as outras hipóteses falham, a que resta, não importa quão improvável seja, deve traduzir a verdade”.

Sir Arthur Conan Doyle – Os Planos do Submarino
Bruce-Partington.

Resumo

Na análise de dados de sobrevivência é frequente a ocorrência de um grande número de indivíduos censurados à direita. Este fato pode ser a indicação de que para uma fração de indivíduos no estudo o evento de interesse nunca vai ocorrer, ou seja, uma fração de indivíduos da população é de curados ou imunes. Este caso não é admitido pela teoria de sobrevivência usual, que em geral considera que todos os indivíduos em risco não terão cura durante o período de acompanhamento. Por isso, os modelos de sobrevivência com fração de cura, ou de longa duração, têm recebido muita atenção em anos recentes. Utilizamos a distribuição exponencial para o tempo de sobrevivência dos indivíduos em risco e a uniforme-exponencial para o tempo de censura. Em muitas situações é evidente que o mecanismo de censura é informativo. Lagakos & Williams (1978) propuseram uma classe de modelos em que o papel do mecanismo de censura em análise de sobrevivência é avaliado e Lagakos (1979) apresentou várias situações em que a suposição de censura não-informativa é violada. Este trabalho tem como objetivo principal verificar o impacto da censura informativa uniforme-exponencial na análise de dados de sobrevivência sob o modelo de mistura padrão.

Abstract

In survival data analysis it is common the occurrence of a large number of individuals to the right. This fact can indicate that, in a fraction of the individuals the event of interest will never happen, in other words, a fraction of individuals of the population is cured or immune. This case is not usually taken into account by the usual survival theory that, in general, considers that the individuals at risk will not achieve cure during the follow-up period. Therefore, the survival models with cure fraction, or long-term survival models, have received a lot of attention in recent years. We consider the exponential distribution for the survival time of individuals at risk and the uniform-exponential distribution for the censoring time. In many situations, it is evident that the censoring mechanism is informative. Lagakos & Williams (1978) proposed a class of models where the acting of the censoring mechanism in the survival time is evaluated and Lagakos (1979) presented several situations in which the assumption of noninformative censoring is violated. The main purpose of this work is to verify the impact of informative uniform-exponential censoring in the survival data analysis under the standard mixture model.

Sumário

Lista de Figuras	iv
Lista de Tabelas	v
1 Introdução	1
1.1 Modelo Uniforme-Exponencial	3
1.2 Objetivos	7
1.3 Organização do Trabalho	8
2 Modelo com Fração de Cura Unificado	9
2.1 Função de Sobrevivência de Longa Duração	10
2.2 Modelo com Fração de Cura Unificado	12
2.3 Função de Probabilidade de Cura	13
2.4 Identificabilidade do Modelo de Mistura Padrão	15
3 Censura Não-Informativa e Censura Informativa	18
3.1 Censura Não-Informativa	20
3.1.1 Função de Verossimilhança e Matriz de Informação de Fisher	20
3.1.2 Relação Funcional Entre os Parâmetros das Distribuições de T e X	22
3.1.3 Censura Não-Informativa Uniforme-Exponencial	23

3.1.4	Função de Verossimilhança com Dados Ampliados	28
3.2	Censura Informativa	30
3.2.1	Função de Verossimilhança e Matriz de Informação de Fisher .	30
3.2.2	Censura Informativa Uniforme-Exponencial	30
4	Estudo de Simulação e Aplicação com Dados Reais	35
4.1	Algoritmo Para Simulação do Modelo de Mistura Padrão	36
4.2	Estudo de Simulação	36
4.2.1	Simulação com p_0 Não-Fixado	37
4.2.2	Simulação com T_0 Não-Fixado	41
4.3	Aplicação com Dados Reais	42
4.3.1	Inferência Clássica com Dados de Leucemia	42
5	Inferência Bayesiana com Priori de Jeffreys e Censura UExp . .	47
5.1	Distribuição <i>a Priori</i> de Jeffreys	47
5.2	Densidades <i>a Posteriori</i> Condicionais	48
5.3	Exemplo com Dados de Leucemia	49
6	Conclusões e Propostas de Trabalhos Futuros	53
A	Provas de Teoremas do Capítulo 3	55
B	Distribuição Exponencial Ponderada	58
B.1	Distribuições Ponderadas	58
B.2	Distribuição Exponencial Ponderada	59
B.3	Comparação Entre X e X^w	60
B.4	Exemplos de Distribuições Ponderadas	61
C	Programas em R para o Modelo de Mistura Padrão	63

C.1 Programa para Estimação Clássica	63
C.2 Programa para Estimação Bayesiana	70
Referências	88

Lista de Figuras

1.1	Estimativas paramétricas e estimativas de Kaplan-Meier (EKM) da função de sobrevivência de longa duração $S_p(t)$ com dados simulados com fração de cura e função de sobrevivência sem a presença de imunes $S(t)$	2
1.2	Funções densidade uniforme-exponencial. (a) $T_0 = 10$ e diferentes valores de λ ; (b) $\lambda = 0,25$ e diferentes valores de T_0	6
1.3	Funções de sobrevivência uniforme-exponencial para diferentes valores de λ e $T_0 = 20$	6
1.4	Funções de risco uniforme-exponencial para diferentes valores de λ e $T_0 = 50$	7
2.1	Reprodução baseada em Peng & Zhang (2008). (a) Curvas de sobrevivência com fração de cura; (b) curvas de sobrevivência de indivíduos em risco.	17
3.1	Probabilidade de um indivíduo ser censurado dado que está em risco (κ^*, κ) versus p_0 para $\lambda = 0,25$ e $p_c = 0,50$. (a) $T_0 = 200$ e $p_{0sup} = 0,4898$; (b) $T_0 = 10$ e $p_{0sup} = 0,2099$	25
4.1	Probabilidades de cobertura com $n = 50, \lambda = 0,25, T_0 = 50$ e $p_c = 0,50$	38
4.2	Comprimento de intervalos de confiança (IC) de 95% com $n = 50, \lambda = 0,25, T_0 = 50$ e $p_c = 0,50$	38
4.3	Estimativas de máxima verossimilhança com $n = 50, \lambda = 0,25, T_0 = 50$ e $p_c = 0,50$	39
4.4	Probabilidades de cobertura com $n = 400, \lambda = 0,25, T_0 = 50$ e $p_c = 0,50$	39
4.5	Comprimento de intervalos de confiança (IC) de 95% com $n = 400, \lambda = 0,25, T_0 = 50$ e $p_c = 0,50$	40

4.6	Estimativas de máxima verossimilhança com $n = 400, \lambda = 0, 25, T_0 = 50$ e $p_c = 0, 50$.	40
4.7	Estimativas de máxima verossimilhança, probabilidades de cobertura e comprimentos de intervalos de confiança 95% de λ com $n = 50, \lambda = 0, 25, p_0 = 0, 30$ e $p_c = 0, 50$.	41
4.8	Estimativas de máxima verossimilhança, probabilidades de cobertura e comprimentos de intervalos de confiança 95% de p_0 com $n = 50, \lambda = 0, 25, p_0 = 0, 30$ e $p_c = 0, 50$.	41
4.9	Estimativas de máxima verossimilhança (EMV) e estimativas de Kaplan-Meier (EKM) da função de sobrevivência $S_p(t)$ com os dados do Exemplo 4.1.	45
4.10	Estimativas de máxima verossimilhança (EMV) e estimativas de Kaplan-Meier (EKM) da função de probabilidade de cura $p(t)$ com os dados do Exemplo 4.1.	45
5.1	Densidades marginais <i>a posteriori</i> de λ, p_0 e κ – dados do Exemplo 4.1.	51
5.2	Histórico das simulações – dados do Exemplo 4.1.	51
5.3	Autocorrelações – dados do Exemplo 4.1.	52
5.4	Densidade marginal <i>a posteriori</i> de κ^* , histórico das simulações e autocorrelações – dados do Exemplo 4.1.	52
B.1	Dispersão da distribuição uniforme-exponencial em relação a exponencial.	62

Lista de Tabelas

4.1	Inferência clássica com censura não-informativa com os dados do Exemplo 4.1.	43
4.2	Inferência clássica com censura informativa $UExp(\lambda_c, T_0 = 5, 1)$ com os dados do Exemplo 4.1.	43
4.3	Inferência clássica com censura informativa $UExp(\lambda_c, T_0 = 10)$ com os dados do Exemplo 4.1.	44
4.4	Inferência clássica com censura informativa $UExp(\lambda_c, T_0 \rightarrow \infty)$ com os dados do Exemplo 4.1.	44
5.1	Inferência Bayesiana com censura informativa $UExp(\lambda_c, T_0 = 5, 1)$ com os dados do Exemplo 4.1.	51

Capítulo 1

Introdução

Em muitos estudos de sobrevivência com dados clínicos, por exemplo em estudos de vários tipos de câncer, verifica-se que uma parte dos indivíduos pode ser considerada curada. Esta característica não é admitida pela teoria de sobrevivência usual, que em geral não considera a presença de indivíduos curados durante a terapia ou período de acompanhamento do estudo. Neste caso, são recomendados, os modelos de sobrevivência com fração de cura, ou de longa duração e o uso destes é muito frequente em anos recentes, como pode ser visto por exemplo em Maller & Zhou (1996), Yakovlev & Tsodikov (1996), Sy & Taylor (2000), Ibrahim *et al.* (2001), Yau & Ng (2001), Tsodikov *et al.* (2003), Chen *et al.* (2004), Rodrigues *et al.* (2009), entre outros. Os modelos com fração de cura foram desenvolvidos originalmente para analisar dados de sobrevivência de indivíduos portadores de câncer por Boag (1949) e Berkson & Gage (1952), que propuseram um modelo de mistura em que há uma proporção de indivíduos curados na população. Iremos nos referir ao modelo de Berkson & Gage (1952) como modelo de mistura padrão. Chen *et al.* (1999) e Tsodikov *et al.* (2003) sugerem modelos alternativos a esse, baseados no número de células suscetíveis a desenvolver metástase. Rodrigues *et al.* (2009) propõem uma teoria unificada da qual o modelo de mistura padrão, que será definido no Capítulo 2, e o modelo de Chen *et al.* (1999) são casos particulares.

Dados de tempo de vida nos quais está caracterizada a presença de unidades amostrais de longa duração ou indivíduos curados possuem curva de sobrevivência,

$S_p(t) = P(T > t)$, como a apresentada na Figura 1.1. Nestas situações, em que o estimador de Kaplan-Meier (vide por exemplo, Kaplan & Meier, 1958; Colosimo & Giolo, 2006) da função de sobrevivência $S_p(t)$ converge para um valor $p_0 > 0$ quando $t \rightarrow \infty$, fica caracterizada a possibilidade de cura ou a presença de imunes na população.

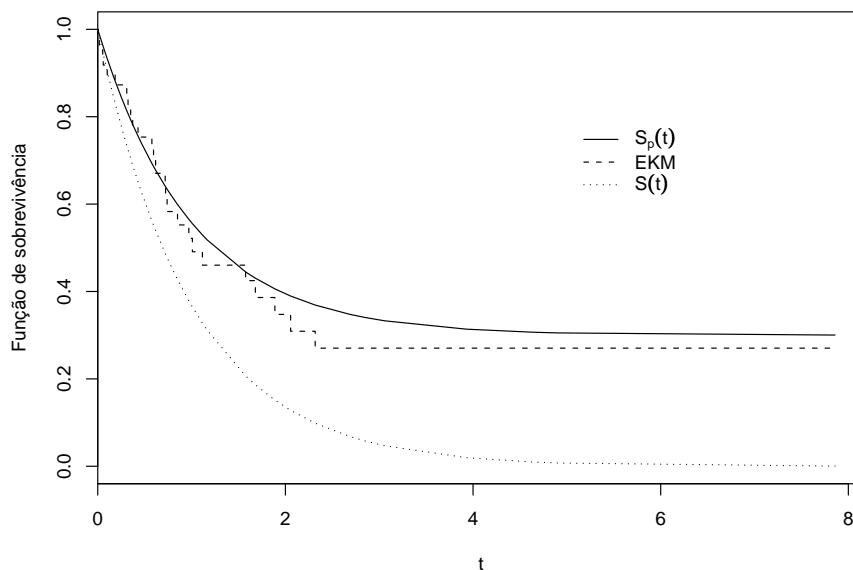


FIGURA 1.1: Estimativas paramétricas e estimativas de Kaplan-Meier (EKM) da função de sobrevivência de longa duração $S_p(t)$ com dados simulados com fração de cura e função de sobrevivência sem a presença de imunes $S(t)$.

Os métodos usualmente encontrados na literatura para analisar dados de sobrevivência com censuras e fração de cura são desenvolvidos sob suposição de censura não-informativa. Lagakos & Williams (1978) propuseram uma classe de modelos em que o papel do mecanismo de censura é avaliado. Lagakos (1979) apresentou várias situações em que a suposição de censura não-informativa é violada.

Neste trabalho desenvolvemos estudos clássicos (vide Capítulos 3 e 4) e Bayesiano (vide Capítulo 5), com um mecanismo de censura informativo mais flexível que os usuais, encontrados na literatura. Este mecanismo será expresso através do modelo uniforme-exponencial.

Definimos na Seção 1.1 o modelo uniforme-exponencial que utilizamos para

representar a distribuição de probabilidade do tempo de censura. Algumas propriedades relacionadas com as distribuições uniforme e exponencial são determinadas. Apresentamos na Seção 1.2 os objetivos do trabalho e na Seção 1.3 a organização dos capítulos seguintes.

1.1 Modelo Uniforme-Exponencial

Considere uma variável aleatória T com função de sobrevivência $S(t)$. Um modelo uniforme-exponencial híbrido das distribuições uniforme ($T \sim U(0, T_0)$) e exponencial ($T \sim Exp(\lambda)$) pode ser formulado pelas equações

$$\frac{-dS(t)}{dt} = \begin{cases} \frac{1}{T_0}, & \text{se } T \sim U(0, T_0), \\ \lambda S(t), & \text{se } T \sim Exp(\lambda). \end{cases}$$

De acordo com Piantadosi & Crowley (1995), combinando estas duas equações diferenciais, um modelo uniforme-exponencial como uma função implícita do tempo é definido por: $\log[S(t)] + \lambda T_0 S(t) = \lambda(T_0 - t)$. Eles utilizaram este modelo para ajustar dados de sobrevivência de pacientes com câncer de pulmão que receberam quimioterapia e terapia de irradiação. O modelo uniforme-exponencial com a função de sobrevivência especificada como uma função explícita do tempo, foi utilizado por Patterson & Smith (1985) e Ghitany (1993). Em análise de sobrevivência tradicional, Patterson & Smith (1985) utilizaram as distribuições Weibull, exponencial e exponencial mista para o tempo de vida, e uniforme e uniforme-exponencial para o tempo de censura. Em análise de sobrevivência de longa duração Ghitany (1993) utilizou para o tempo de vida uma distribuição exponencial e para o tempo de censura as distribuições uniforme-exponencial em $(0, T_0)$, uniforme em $(0, T_0)$ e uma distribuição degenerada em T_0 . Embora a formulação de Piantadosi & Crowley (1995), seja atraente em originalidade, foi pouco difundida em análise de sobrevivência. Para os nossos objetivos atuais, optamos pelo modelo uniforme-exponencial utilizado por Patterson & Smith (1985) e Ghitany (1993).

Suponha que há interesse em modelar uma situação que pode ser representada por uma função de risco híbrida das distribuições uniforme e exponen-

cial. Definimos, como uma função explícita do tempo, uma distribuição uniforme-exponencial para uma v.a. T , denotada por $T \sim UExp(\lambda, T_0)$, como segue.

Definição 1.1 Uma variável aleatória T tem distribuição uniforme-exponencial (UExp) com parâmetros λ e T_0 , o tempo de duração do estudo (conhecido), se sua função de distribuição de probabilidade, G , for dada por

$$G(t; \lambda) = 1 - \left(1 - \frac{t}{T_0}\right) e^{-\lambda t}, \quad (1.1)$$

em que $0 < t < T_0$ e $\lambda > 0$.

Uma nova formulação da distribuição uniforme-exponencial e algumas propriedades de uma v.a. $T \sim UExp(\lambda, T_0)$ são dadas nas seguintes proposições:

Proposição 1.1 Suponha que $T_1 \sim Exp(\lambda)$ e $T_2 \sim U(0, T_0)$, T_0 conhecido, são variáveis aleatórias independentes com funções de sobrevivência $S_1(t)$ e $S_2(t)$, respectivamente. Então, a v.a. $T = \min\{T_1, T_2\}$ tem função de distribuição de probabilidade dada por (1.1).

Prova. Seja $G(t) = P(T \leq t)$. Então,

$$G(t) = P(\min\{T_1, T_2\} \leq t) = 1 - P(T_1 > t, T_2 > t) = 1 - S_1(t)S_2(t).$$

Logo $G(t)$ é dada por (1.1). ■

Esta formulação sugere a construção de novas distribuições de probabilidade pela substituição de T_1 na Proposição 1.1 por outras distribuições a fim de generalizar o modelo uniforme-exponencial. Além disso, outras distribuições de probabilidade envolvendo T_0 podem ser propostas para ajustar dados de tempo de vida.

Proposição 1.2 A função densidade de probabilidade (fdp) g e a função de risco h de uma v.a. $T \sim UExp(\lambda, T_0)$ são dadas respectivamente por

$$(a) \quad g(t; \lambda) = \left[\frac{1}{T_0} + \lambda \left(1 - \frac{t}{T_0}\right) \right] e^{-\lambda t}, \quad 0 < t < T_0, \quad \lambda > 0 \quad (1.2)$$

e

$$(b) \quad h(t; \lambda) = \lambda + \frac{1}{T_0 - t}, \quad 0 < t < T_0, \quad \lambda > 0. \quad (1.3)$$

Prova. (a) Basta usar a equação $g(t; \lambda) = G'(t; \lambda)$. (b) Pela definição da função de risco, temos que $h(t; \lambda) = g(t; \lambda)/S(t; \lambda)$, em que $g(t; \lambda)$ e $S(t; \lambda) = P(T > t)$ representam, a fdp e a função de sobrevivência da v.a. T , respectivamente. Desse modo, os resultados (1.2) e (1.3) são determinados trivialmente. ■

Note que, $\lim_{t \rightarrow 0} g(t; \lambda) = \frac{1}{T_0} + \lambda$ e $\lim_{t \rightarrow T_0} h(t; \lambda) = \infty$, nas equações (1.2) e (1.3), respectivamente.

Proposição 1.3 Se $T \sim UExp(\lambda, T_0)$, então

$$(a) \lim_{\lambda \rightarrow 0} G(t; \lambda) = G_U(t; T_0) \quad e \quad (b) \lim_{T_0 \rightarrow \infty} G(t; \lambda) = G_{Exp}(t; \lambda), \quad (1.4)$$

em que $G_U(\cdot; T_0)$ e $G_{Exp}(\cdot; \lambda)$ são as funções de distribuições acumuladas das variáveis aleatórias uniforme em $(0, T_0)$ e exponencial com parâmetro λ , respectivamente.

Prova. Os resultados (a) e (b) são obtidos ao tomarmos o limite da expressão (1.1) quando $\lambda \rightarrow 0$ e $T_0 \rightarrow \infty$, respectivamente. ■

Proposição 1.4 O valor esperado e a variância de uma v.a. $T \sim UExp(\lambda, T_0)$ são dados respectivamente por

$$(a) E(T) = \frac{1}{\lambda} \left[1 - \frac{1}{\lambda T_0} (1 - e^{-\lambda T_0}) \right] \quad (1.5)$$

e

$$(b) Var(T) = \frac{2}{\lambda^2} \left[1 - \frac{2}{\lambda T_0} + \left(1 + \frac{2}{\lambda T_0}\right) e^{-\lambda T_0} \right] - \frac{1}{\lambda^2} \left[1 - \frac{1}{\lambda T_0} (1 - e^{-\lambda T_0}) \right]^2. \quad (1.6)$$

Prova. (a) Da definição de valor esperado de uma v.a. T com valores possíveis em \mathbb{R}_+ , temos que

$$E(T) = \int_0^\infty [1 - G(t; \lambda)] dt = \int_0^\infty \left(1 - \frac{t}{T_0} \right) e^{-\lambda t} dt.$$

Após uma integração por partes obtém-se o resultado desejado. (b) Basta calcular $E(T^2)$, em seguida determinar $Var(T)$. ■

Apresentamos, na Figura 1.2, gráficos da fdp da v.a. $T \sim UExp(\lambda, T_0)$ para alguns valores de λ e T_0 . Os gráficos foram construídos a fim de ilustrar a Proposição

1.2. Podemos observar tanto o comportamento exponencial quanto o uniforme da v.a. T .

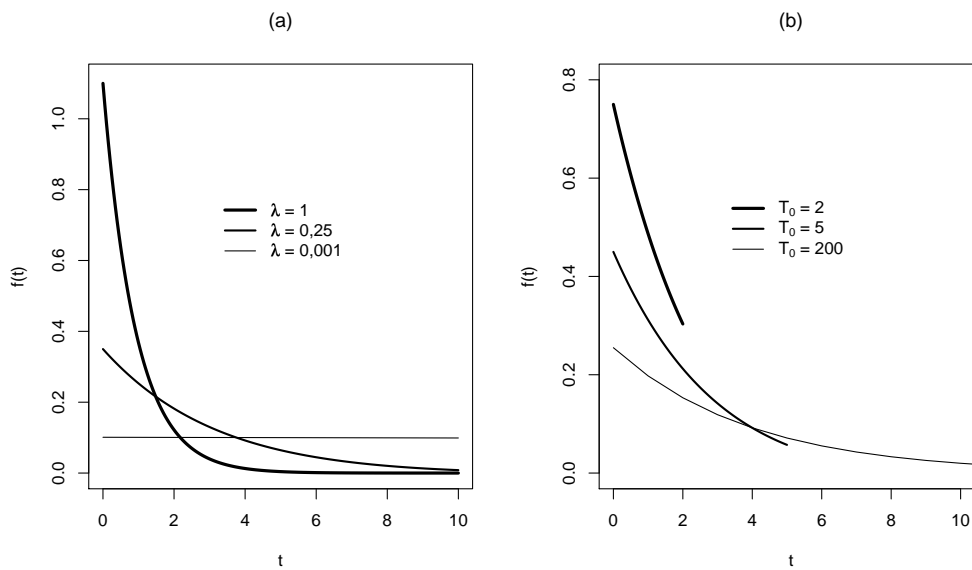


FIGURA 1.2: Funções densidade uniforme-exponencial. (a) $T_0 = 10$ e diferentes valores de λ ; (b) $\lambda = 0,25$ e diferentes valores de T_0 .

Podemos observar diferentes exemplos da função de sobrevivência de uma v.a. T uniforme-exponencial na Figura 1.3. As curvas foram traçadas para $T_0 = 20$ e $\lambda = 1; 0,25; 0,1; 0,001$.

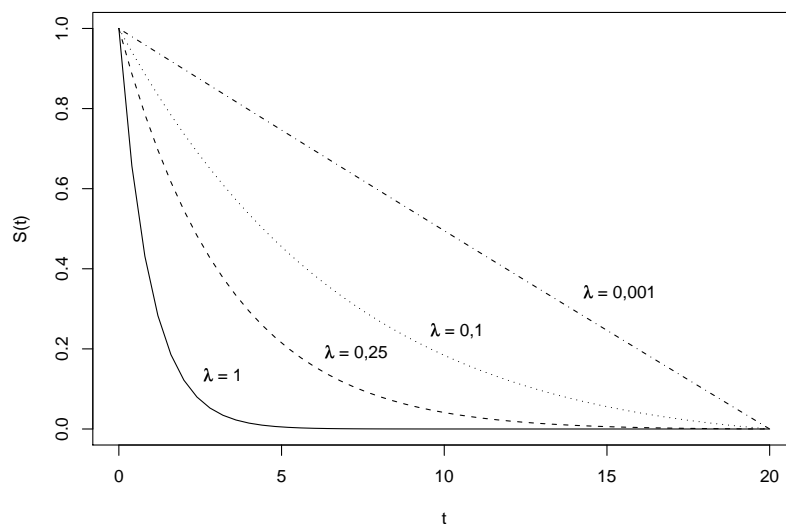


FIGURA 1.3: Funções de sobrevivência uniforme-exponencial para diferentes valores de λ e $T_0 = 20$.

Podemos observar diferentes exemplos da função de risco de uma v.a. T uniforme-exponencial na Figura 1.4. As curvas foram traçadas para $T_0 = 50$ e $\lambda = 4; 2; 1; 0,25$.

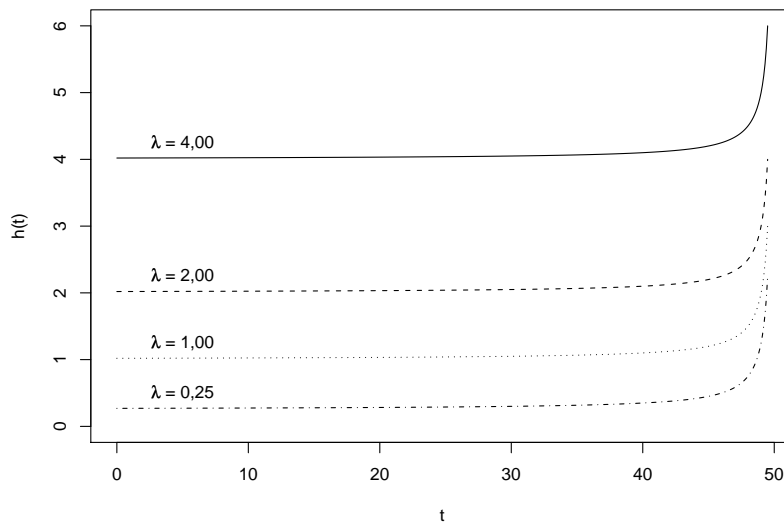


FIGURA 1.4: Funções de risco uniforme-exponencial para diferentes valores de λ e $T_0 = 50$.

1.2 Objetivos

Este trabalho tem como objetivo principal avaliar, com o uso de dados simulados, o impacto da censura informativa uniforme-exponencial em inferências dos parâmetros do modelo de mistura padrão. Especificamente, destacamos os seguintes objetivos:

- Obter a matriz de informação de Fisher para os parâmetros do modelo de mistura padrão com censura informativa uniforme-exponencial e censura não-informativa.
- Desenvolver um algoritmo e programa computacional para simular dados de acordo com o modelo de mistura padrão com censura uniforme-exponencial.
- Construir intervalos de confiança assintóticos para os parâmetros do modelo de mistura padrão com censura uniforme-exponencial.

- Avaliar, utilizando dados simulados, as propriedades assintóticas dos estimadores de máxima verossimilhança e intervalos de confiança dos parâmetros do modelo de mistura padrão com censura não-informativa e censura informativa uniforme-exponencial.
- Aplicar a metodologia desenvolvida em conjuntos de dados reais.

1.3 Organização do Trabalho

Este trabalho está organizado em seis capítulos e três apêndices. Apresentamos no Capítulo 2 a função de sobrevivência de longa duração, o modelo com fração de cura unificado e, como caso particular deste, o modelo de mistura padrão. Definimos uma função de probabilidade de cura sob o enfoque do modelo com fração de cura unificado, e analisamos a identificabilidade do modelo de mistura padrão. No Capítulo 3 introduzimos o modelo de mistura padrão com censura informativa com distribuição uniforme-exponencial e com censura não-informativa. As matrizes de informação de Fisher são obtidas analiticamente para os dois casos utilizando os teoremas construídos no Apêndice A. A função de verossimilhança com censura não-informativa com dados ampliados é também introduzida. No Capítulo 4 propomos um algoritmo para simulação de amostras do modelo de mistura padrão. Além disso, desenvolvemos uma inferência clássica para analisar conjuntos de dados simulados e um conjunto de dados reais com mecanismo de censura não-informativo e informativo uniforme-exponencial. No Capítulo 5 desenvolvemos uma inferência Bayesiana objetiva para o modelo de mistura padrão sob distribuições *a priori* de Jeffreys. O conjunto de dados reais utilizado no Capítulo 4 é analisado. No Capítulo 6 apresentamos as conclusões e propostas de trabalhos futuros. Introduzimos no Apêndice B as distribuições ponderadas com interesse específico na distribuição exponencial ponderada que generaliza o modelo exponencial ordinário. Além disso, o modelo uniforme-exponencial é obtido como um caso especial. Dois programas computacionais que desenvolvemos em R (R Development Core Team, 2009) estão apresentados no Apêndice C, o primeiro para obtenção da inferência clássica do Capítulo 4 e o segundo para a inferência Bayesiana do Capítulo 5.

Capítulo 2

Modelo com Fração de Cura Unificado

Em estudos clínicos de sobrevivência ou em análise de dados de tempos de falha, o evento de interesse associado a uma determinada causa pode ser a morte de um indivíduo, a falha de um sistema, a remissão de uma doença, a perda de um cliente por uma instituição financeira ou a falência de uma empresa, entre outros. Em muitos desses estudos ou experimentos o evento de interesse especificado poderá nunca ocorrer. Isto significa que uma parcela de unidades da população é de curados ou imunes e a função de sobrevivência associada é imprópria. Diferentemente dos modelos usuais de sobrevivência é assumido que $\lim_{t \rightarrow \infty} S_p(t) = p_0$, em que p_0 é a proporção de imunes ou fração de cura e a proporção complementar $1 - p_0$ é de não-curados. Na metodologia que apresentamos, a classificação individual de curados ou não-curados numa amostra não é possível. Entretanto, se o gráfico do estimador de Kaplan & Meier (1958) da função de sobrevivência apresentar cauda longa, num nível visivelmente acima de zero, teremos uma forte indicação de que as observações censuradas irão determinar uma fração de cura maior do que zero.

Neste capítulo sumarizamos o modelo com fração de cura unificado proposto por Rodrigues *et al.* (2009), que considera a presença de indivíduos imunes ou curados no período de acompanhamento do estudo. O modelo de Berkson & Gage (1952) também denominado de modelo de mistura padrão, utilizado por vários

autores, vide por exemplo, Farewell (1982), Yakovlev & Tsodikov (1996), Maller & Zhou (1996), Ibrahim *et al.* (2001), Chen *et al.* (2002) e Tsodikov *et al.* (2003), é obtido como caso particular do modelo com fração de cura unificado. Na Seção 2.1 apresentamos a função de sobrevivência de longa duração, na Seção 2.2 o modelo com fração de cura unificado, na Seção 2.3 uma função de probabilidade de cura e na Seção 2.4, a identificabilidade do modelo de mistura padrão é discutida.

2.1 Função de Sobrevivência de Longa Duração

Nesta seção destacamos dois teoremas apresentados em Rodrigues *et al.* (2009). De acordo com Feller (1968), temos a seguinte definição:

Definição 2.1 Seja $\{a_m\}$, $m = 0, 1, 2, \dots$, uma sequência de números reais. Se para $s \in [0, 1]$,

$$A(s) = a_0 + a_1s + a_2s^2 + \dots \quad (2.1)$$

converge, então $A(s)$ é definida como função geradora da sequência $\{a_m\}$.

Seja M uma variável aleatória representando o número de causas ou riscos da ocorrência de um particular evento de interesse com distribuição

$$p_m = P(M = m), \quad m = 0, 1, 2, \dots \quad (2.2)$$

Dado $M = m$ considere $Z_i, i = 1, \dots, m$, variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) representando o tempo de ocorrência do evento de interesse devido à i -ésima causa ou risco. Então, a função de sobrevivência de um indivíduo da população em risco associado à i -ésima causa é dada por

$$S(t) = P(Z_i > t), \quad i = 1, \dots, m, \quad t \geq 0, \quad (2.3)$$

em que $S(t)$ é uma função de sobrevivência própria, isto é, $\lim_{t \rightarrow \infty} S(t) = 0$.

O tempo de ocorrência do evento de interesse para um indivíduo da população é dado por $T = \min\{Z_0, Z_1, \dots, Z_M\}$, sendo que $P(Z_0 = \infty) = 1$. Esta suposição admite a possibilidade de uma proporção p_0 da população não apresentar a

ocorrência do evento de interesse. Fração de cura, proporção de curados e proporção de imunes da população são denominações equivalentes de p_0 . As variáveis aleatórias Z_i e M são variáveis latentes, enquanto a variável aleatória T é uma variável observável com função de sobrevivência dada por

$$S_p(t) = P(T > t). \quad (2.4)$$

Teorema 2.1 Dada uma função de sobrevivência própria $S(t)$, para os indivíduos em risco da população, temos que a função de sobrevivência de longa duração da variável aleatória T é dada por

$$S_p(t) = A(S(t)) = \sum_{m=0}^{\infty} p_m (S(t))^m, \quad (2.5)$$

em que $A(\cdot)$ é a função geradora da sequência $\{p_m\}$, que é convergente para $0 \leq S(t) \leq 1$.

Prova. Pode ser vista em Rodrigues *et al.* (2009), mas devido o uso da mesma para provar o Teorema 2.4, uma reprodução é dada como segue. Note que

$$S_p(t) = P(T > t) = P(\min \{Z_0, Z_1, \dots, Z_M\} > t).$$

Então,

$$\begin{aligned} S_p(t) &= P(Z_0 > t, M = 0) + P(\min \{Z_1, Z_2, \dots, Z_M\} > t, M \geq 1) \\ &= P(Z_0 > t, M = 0) + P(Z_1 > t, Z_2 > t, \dots, Z_M > t, \cup_{i=1}^{\infty} \{M = i\}) \\ &= P(Z_0 > t, M = 0) + P(Z_1 > t, M = 1) \\ &\quad + P(Z_1 > t, Z_2 > t, M = 2) + \dots + P(Z_1 > t, \dots, Z_m > t, M = m) + \dots \\ &= P(Z_0 > t | M = 0)P(M = 0) + P(Z_1 > t | M = 1)P(M = 1) \\ &\quad + P(Z_1 > t, Z_2 > t | M = 2)P(M = 2) \\ &\quad + P(Z_1 > t, Z_2 > t, \dots, Z_m > t | M = m)P(M = m) + \dots \\ &= p_0 + S(t)p_1 + (S(t))^2 p_2 + \dots + (S(t))^m p_m + \dots \\ &= p_0 + \sum_{m=1}^{\infty} p_m (S(t))^m = A(S(t)). \blacksquare \end{aligned}$$

Podemos observar que $S_p(t)$ não é uma função de sobrevivência própria, isto é, $\lim_{t \rightarrow \infty} S_p(t) > 0$. A fração de cura da população é dada pelo seguinte teorema:

Teorema 2.2 Dada uma função de sobrevivência própria $S(t)$, para os elementos da população em risco, temos que

$$\lim_{t \rightarrow \infty} S_p(t) = p_0 = P(M = 0), \quad (2.6)$$

em que p_0 é a proporção de indivíduos curados ou imunes da população.

Prova. Vide Rodrigues *et al.* (2009). ■

2.2 Modelo com Fração de Cura Unificado

Com a finalidade de formular o modelo de mistura padrão, considere de acordo com Rodrigues *et al.* (2009) o seguinte teorema:

Teorema 2.3 O modelo com fração de cura unificado é dado por

$$S_p(t) = p_0 + (1 - p_0)S^*(t), \quad (2.7)$$

sendo que

$$S^*(t) = \sum_{m=1}^{\infty} p_m^* (S(t))^m, \quad p_m^* = \frac{p_m}{1 - p_0}.$$

Prova. Escrevendo $S_p(t)$ como

$$S_p(t) = p_0 + \sum_{m=1}^{\infty} p_m (S(t))^m = p_0 + (1 - p_0) \frac{\sum_{m=1}^{\infty} p_m (S(t))^m}{1 - p_0},$$

a equação (2.7) é obtida. ■

Devemos observar que $S^*(t)$ é uma função de sobrevivência própria associada aos indivíduos da população em risco. Um caso particular do modelo de longa duração unificado dado pela equação (2.7), assumindo que o número de causas M segue uma distribuição de Bernoulli com parâmetro $1 - p_0$, é o modelo de modelo de mistura padrão, dado por

$$S_p(t) = p_0 + (1 - p_0)S(t), \quad (2.8)$$

em que $S(t) = P(T > t | M = 1)$ é uma função de sobrevivência própria, isto é,

$$\lim_{t \rightarrow \infty} S(t) = 0.$$

De acordo com a equação (2.5) a derivada da função de sobrevivência de longa duração é dada por

$$S'_p(t) = -f(t) \sum_{m=1}^{\infty} mp_m(S(t))^{m-1}, \quad (2.9)$$

sendo que $f(t)$ é a função densidade própria associada a $S(t)$. Da equação (2.9), a subfunção densidade associada à função de sobrevivência de longa duração é dada por

$$f_p(t) = f(t) \frac{dA(s)}{ds} \Big|_{s=S(t)} = -S'_p(t). \quad (2.10)$$

A subfunção de risco de longa duração é dada por

$$h_p(t) = \frac{f_p(t)}{S_p(t)} = f(t) \frac{\frac{dA(s)}{ds} \Big|_{s=S(t)}}{S_p(t)}. \quad (2.11)$$

Decorre da equação (2.7) que a função densidade própria da população em risco e sua correspondente função de risco são dadas por

$$f^*(t) = -\frac{dS^*(t)}{dt} = f(t) \frac{\frac{dA(s)}{ds} \Big|_{s=S(t)}}{1 - p_0} \quad (2.12)$$

e

$$h^*(t) = \frac{f^*(t)}{S^*(t)} = \frac{S_p(t)h_p(t)}{(1 - p_0)S^*(t)} = \frac{S_p(t)}{S_p(t) - p_0} h_p(t). \quad (2.13)$$

Como $\frac{S_p(t)}{S_p(t) - p_0} > 1$, então $h^*(t) > h_p(t)$, $\forall t$. Devemos observar que $h^*(t)$ em (2.13) não possui a propriedade de riscos proporcionais (Cox, 1972). Quando variáveis explicativas são ligadas a p_0 em (2.11), o modelo com fração de cura unificado, (2.7), tem a propriedade de riscos proporcionais de Cox se, e somente se, a distribuição de probabilidade de M for Poisson com parâmetro θ . Além desta, outras propriedades do modelo com fração de cura unificado são demonstradas em Rodrigues *et al.* (2009).

2.3 Função de Probabilidade de Cura

Nesta seção apresentamos uma questão de bastante interesse que ocorre durante o tempo de acompanhamento em análise de sobrevivência na presença de

indivíduos curados, conforme analisada em Gordon (1990), Maller & Zhou (1996) e Lambert *et al.* (2010).

Dado que um indivíduo tenha sido acompanhado até o tempo $t > 0$ e o evento falha não tenha sido observado, desejamos determinar a probabilidade de que este indivíduo esteja curado. Sob o modelo de mistura padrão, Maller & Zhou (1996) calcularam esta probabilidade e apresentaram estimativas não-paramétricas em dados de dois grupos de pacientes com leucemia.

Utilizando um modelo de mistura de duas distribuições de Gompertz, uma para associar a morte por câncer de mama e outra para associar a morte por outras causas, Gordon (1990) apresentou uma definição equivalente para calcular a proporção de curados dado o tempo t após a cirurgia. Uma observação necessária é que neste modelo a proporção de pacientes curados da doença morre de outras causas, isto é, a função de sobrevivência é própria.

Baseada em Maller & Zhou (1996), definimos a função de probabilidade de cura do seguinte modo:

Definição 2.2 Dado que um indivíduo tenha sobrevivido até o tempo $t > 0$, a probabilidade de que este indivíduo esteja curado é determinada pela função de probabilidade de cura

$$p(t) = P(M = 0|T > t). \quad (2.14)$$

A função $p(t)$, de acordo com o modelo com fração de cura unificado, é obtida pelo seguinte teorema.

Teorema 2.4 Dado que um indivíduo tenha sobrevivido até o tempo $t > 0$, a função de probabilidade de cura, baseada no modelo com fração de cura unificado, é dada por

$$p(t) = \frac{p_0}{A(S(t))} = \frac{p_0}{S_p(t)}. \quad (2.15)$$

Prova. Pela Definição 2.2 temos que $p(t) = P(M = 0|T > t)$ e pela fórmula de Bayes,

$$p(t) = \frac{P(M = 0, T > t)}{P(T > t)} = \frac{P(T > t|M = 0)P(M = 0)}{P(T > t)}. \quad (2.16)$$

Note que $P(T > t | M = 0) = 1, \forall t$ e $P(M = 0) = p_0$. Pela prova do Teorema 2.1 $S_p(t) = P(T > t)$. Logo, $p(t)$ é dada por (2.15). ■

Observar que $p(t)$ é não-decrescente. A conexão desta função com as funções $S_p(t)$ e $h_p(t)$ pode ser explorada na busca de outras propriedades do modelo com fração de cura unificado.

2.4 Identificabilidade do Modelo de Mistura Padrão

Em análise de dados de sobrevivência sem a presença de fração de cura, o problema de identificabilidade tem sido amplamente estudado, vide por exemplo, Tsiatis (1975), Emoto & Matthews (1990), Kalbfleisch & Prentice (2002), Dunson & Dinse (2002), Ebrahimi *et al.* (2003) e Oller *et al.* (2007). Sob a presença de fração de cura Farewell (1986), Gordon (1990), Li *et al.* (2001) e Peng & Zhang (2008) investigaram a identificabilidade considerando o modelo de mistura padrão e outros modelos de sobrevivência com fração de cura. Em geral, uma família paramétrica de densidades $f(t_j; \boldsymbol{\theta})$ é identificável se distintos valores de $\boldsymbol{\theta}$ determinam distintos membros da família de densidades $\{f(t_j; \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$ (McLachlan & Peel, 2000).

Para determinar a identificabilidade do modelo de mistura padrão seguimos a abordagem dada aos modelos de mistura por Yakowitz & Spragins (1968), Titterington *et al.* (1985) e a abordagem dada aos modelos de sobrevivência de longa duração por Li *et al.* (2001). Com este propósito, seja $\mathcal{F} = \{F(t; \boldsymbol{\theta} | M = 1) : t > 0, \boldsymbol{\theta} \in \Theta\}$ a classe de distribuições condicionais próprias para o tempo de vida T dos indivíduos em risco (suscetíveis), $t < T^+$, para excluir distribuições degeneradas em $t = +\infty$. Então, a classe \mathcal{H} de todos os modelos de mistura padrão é dada por

$$\begin{aligned} \mathcal{H} &= \{F_p(t; \boldsymbol{\theta}, p_0) : F_p(t; \boldsymbol{\theta}, p_0) = (1 - p_0)F(t; \boldsymbol{\theta} | M = 1), \\ &\quad F(t; \boldsymbol{\theta} | M = 1) \in \mathcal{F}, p_0 \in \Theta\}. \end{aligned} \quad (2.17)$$

A definição seguinte, baseada em Li *et al.* (2001), estabelece a identificabilidade de \mathcal{H} . Note que é necessário que todos os membros de \mathcal{H} sejam distintos.

Definição 2.3 Considere dois membros da classe de modelos de mistura padrão \mathcal{H} ,

dados por $F_p(t; \boldsymbol{\theta}, p_0) = (1 - p_0)F(t; \boldsymbol{\theta} | M = 1)$ e $F_p^*(t; \boldsymbol{\theta}^*, p_0^*) = (1 - p_0^*)F^*(t; \boldsymbol{\theta}^* | M = 1)$. Então, \mathcal{H} é uma classe identificável se $F_p(t; \boldsymbol{\theta}, p_0) = F_p^*(t; \boldsymbol{\theta}^*, p_0^*)$ se, e somente se, $p_0 = p_0^*$ e $F(t; \boldsymbol{\theta} | M = 1) = F^*(t; \boldsymbol{\theta}^* | M = 1)$, para quase todo t em $(0, T^+)$.

Considerando a Definição 2.3 a identificabilidade do modelo de mistura padrão é formalizada pelo seguinte teorema:

Teorema 2.5 (Li *et al.*, 2001). Suponha que a função de sobrevivência condicional própria, $S(t; \boldsymbol{\theta} | M = 1)$, seja completamente especificada por um modelo paramétrico. Então, a função de sobrevivência $S_p(t; \boldsymbol{\theta}, p_0) = p_0 + (1 - p_0)S(t; \boldsymbol{\theta} | M = 1)$, $t < T^+$ é identificável.

Prova. Pela Definição 2.3 suponha que $S_p(t; \boldsymbol{\theta}, p_0) = S_p^*(t; \boldsymbol{\theta}^*, p_0^*)$. Então $(1 - p_0)(1 - S(t; \boldsymbol{\theta} | M = 1)) = (1 - p_0^*)(1 - S^*(t; \boldsymbol{\theta}^* | M = 1))$. Segue que a razão,

$$\frac{1 - p_0}{1 - p_0^*} = \frac{1 - S^*(t; \boldsymbol{\theta}^* | M = 1)}{1 - S(t; \boldsymbol{\theta} | M = 1)} = c,$$

sendo que c é uma constante positiva independente de t . Portanto, $S^*(t; \boldsymbol{\theta}^* | M = 1)$ não pertence à família \mathcal{F} da qual $S(t; \boldsymbol{\theta} | M = 1)$ pertence se $c \neq 1$, a menos que $\boldsymbol{\theta}^* = \boldsymbol{\theta}$. ■

Para ilustrar graficamente o problema de identificabilidade do modelo de mistura padrão, apresentamos o seguinte exemplo.

Exemplo 2.1 (Peng & Zhang, 2008). Considere dois modelos de mistura padrão. O modelo $S_{p1}(t)$ com fração de cura 0,3 e função de sobrevivência dos indivíduos em risco $S_1(t)$ seguindo uma distribuição de Weibull com parâmetros de forma e escala iguais a 0,5. O outro modelo $S_{p2}(t)$ com fração de cura 0,2 e função de sobrevivência dos indivíduos em risco $S_2(t)$ seguindo uma mistura de uma distribuição degenerada em ∞ e uma distribuição de Weibull com parâmetros de forma e escala iguais a 0,5 e pesos 0,125 e 0,875, respectivamente. A Figura 2.1 (a) mostra duas curvas de sobrevivência $S_p(t)$, em que $S_{p1}(t) = S_{p2}(t)$, enquanto a Figura 2.1 (b) mostra duas curvas de sobrevivência dos indivíduos em risco $S(t)$, em que $S_1(t)$ e $S_2(t)$ são totalmente diferentes com o aumento de t . Neste caso, os parâmetros de interesse do modelo $S_p(t) = p_0 + (1 - p_0)S(t)$ não são identificáveis.

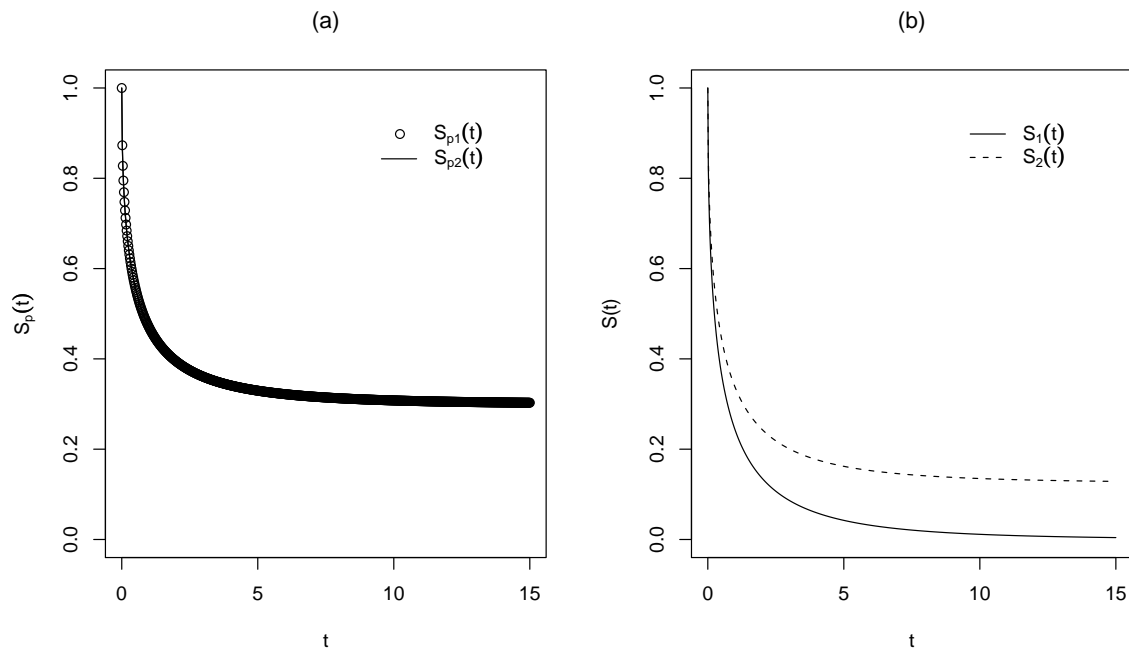


FIGURA 2.1: Reprodução baseada em Peng & Zhang (2008). (a) Curvas de sobrevivência com fração de cura; (b) curvas de sobrevivência de indivíduos em risco.

Neste capítulo foi apresentado, sob o enfoque de função geradora de probabilidades, o modelo com fração de cura unificado e o modelo de mistura padrão. Apresentamos também a função de probabilidade de cura e o problema de identificabilidade do modelo de mistura padrão. No próximo capítulo este modelo será utilizado para destacar o impacto da censura informativa em tempos de vida com distribuição exponencial e tempos de censura com distribuição uniforme-exponencial.

Capítulo 3

Censura Não-Informativa e Censura Informativa

Na análise de dados de sobrevivência a maioria dos estudos utiliza a suposição de que a função de sobrevivência não sofre interferência do mecanismo de censura, sendo admitido que uma determinada observação é censurada simplesmente porque o correspondente evento de interesse, falha ou morte, não ocorreu no período de acompanhamento ou tempo de estudo. Entretanto, existem muitas situações, como mostram Williams & Lagakos (1977), Lagakos (1979), Link (1989) e Siannis *et al.* (2005) em que o tempo de sobrevivência é claramente perturbado pelo mecanismo de censura. Por exemplo, em experimentos clínicos nos quais o paciente é removido do estudo devido ao alto grau de toxicidade da terapia. Por outro lado, pacientes que tiveram melhora significativa ou foram curados durante a terapia podem ser removidos do estudo causando um aumento do tempo de vida esperado. Conseqüentemente, não é razoável ignorar o efeito que o mecanismo de censura produz no tempo de sobrevivência. Lagakos & Williams (1978) propuseram uma classe de modelos em que um parâmetro é especificado para medir o efeito do mecanismo de censura no tempo de sobrevivência. Siannis (2004) e Siannis *et al.* (2005) utilizaram um modelo no qual a dependência entre o tempo de sobrevivência e o mecanismo de censura é avaliado por meio de um parâmetro que é coeficiente de uma função de vício introduzida na distribuição de censura.

Neste capítulo o desenvolvimento teórico que apresentamos revela o impacto da censura informativa na matriz de informação de Fisher, calculada analiticamente para o modelo de mistura padrão, sugerida por Maller & Zhou (1996). Com este propósito, determinamos uma importante relação funcional entre os parâmetros da distribuição do tempo de sobrevivência dos indivíduos em risco e a distribuição do tempo de censura. Apresentamos um estudo com um mecanismo de censura informativo mais flexível que os usuais encontrados na literatura. Este mecanismo está expresso pelo modelo uniforme-exponencial que pode ser encontrado em Patterson & Smith (1985) e Ghitany (1993). As matrizes de informação de Fisher são obtidas com o uso da função de verossimilhança com censura não-informativa e com censura informativa. Para calcular os elementos da matriz de informação de Fisher utilizamos os resultados teóricos que desenvolvemos no Apêndice A. Em particular destacamos os teoremas associados às derivadas parciais de segunda ordem em relação à fração de cura p_0 . A função de verossimilhança com censura não-informativa com dados ampliados é também apresentada.

Com a finalidade de representar o modelo de mistura padrão completamente parametrizado considere a seguinte definição.

Definição 3.1 Seja T uma variável aleatória representando o tempo de ocorrência do evento de interesse e seja M uma variável aleatória Bernoulli representando o número de causas do evento de interesse, em que $M = 0$ se o indivíduo não está em risco (indivíduo curado ou imune) e $M = 1$ caso contrário. Seja $P(M = 0) = p_0 \in (0, 1)$ a proporção de indivíduos curados que nunca falharão. O modelo de mistura padrão é dado por

$$S_p(t; \boldsymbol{\theta}, p_0) = p_0 + (1 - p_0)S(t; \boldsymbol{\theta} | M = 1), \quad (3.1)$$

sendo que $\boldsymbol{\theta} = (\theta_1, \dots, \theta_k)'$ e p_0 pertencem ao espaço paramétrico $\Theta = \mathbb{R}^k \times (0, 1)$ e $S(t; \boldsymbol{\theta} | M = 1)$ é uma função de sobrevivência própria associada aos indivíduos em risco da população.

Para caracterizar os dois tipos de censura que iremos trabalhar, censura não-informativa e censura informativa, introduzimos, motivada por Lawless (1982), Kalbfleisch & Prentice (2002) e Siannis *et al.* (2005), a seguinte definição.

Definição 3.2 Sob o modelo (3.1) o mecanismo de censura é informativo se existir uma relação funcional entre os parâmetros da distribuição de probabilidade da v.a. T e os parâmetros da distribuição de probabilidade da v.a. X , o tempo de censura. Caso contrário, a censura é não-informativa.

3.1 Censura Não-Informativa

3.1.1 Função de Verossimilhança e Matriz de Informação de Fisher

O conjunto de dados \mathcal{D} que temos interesse em analisar consiste de n observações (Y_i, δ_i) das variáveis aleatórias (Y, δ) . Seja T o tempo de vida dos indivíduos em risco e seja X o tempo de censura variáveis aleatórias independentes com funções de distribuição F e G , respectivamente. A variável aleatória Y_i é definida por

$$Y_i = \min \{T_i, X_i\}, \quad i = 1, \dots, n, \quad (3.2)$$

sendo que T_1, \dots, T_n são variáveis aleatórias i.i.d. com função de distribuição F , independentes de X_1, \dots, X_n , as quais são i.i.d. com função de distribuição G . A variável indicadora δ_i é tal que $\delta_i = 0$, se a i -ésima observação é censurada e $\delta_i = 1$, se a i -ésima observação está associada a uma falha. Note que $\delta_i, i = 1, \dots, n$, pode ser representada do seguinte modo

$$\delta_i = \begin{cases} 1, & \text{se } T_i \leq X_i, \\ 0, & \text{se } T_i > X_i. \end{cases} \quad (3.3)$$

No modelo (3.1) existem dois mecanismos de censura atuando simultaneamente, um dado pela variável aleatória δ_i definida em (3.3) e outro pela variável aleatória

$$M_i = \begin{cases} 1, & \text{se o indivíduo } i \text{ está em risco,} \\ 0, & \text{se o indivíduo } i \text{ não está em risco (curado),} \end{cases} \quad (3.4)$$

em que, de acordo com a Definição 3.1, $M_i \sim \text{Bernoulli}(1 - p_0)$ i.i.d., $i = 1, \dots, n$.

Suponha que a censura seja não-informativa, isto é, a função de distribuição G da variável aleatória X_i não envolve $\boldsymbol{\theta}$ e p_0 . Representamos a função de verossimilhança com censura não-informativa $L(\boldsymbol{\theta}, p_0; \mathcal{D})$, de acordo com a Definição 3.2, Maller & Zhou (1996) e de Castro *et al.* (2009) pela expressão

$$L(\boldsymbol{\theta}, p_0; \mathcal{D}) \propto \prod_{i=1}^n [f_p(y_i; \boldsymbol{\theta}, p_0)]^{\delta_i} [S_p(y_i; \boldsymbol{\theta}, p_0)]^{1-\delta_i}, \quad (3.5)$$

sendo que $f_p(\cdot) = -S'_p(\cdot)$.

Na equação (3.5) o estimador de máxima verossimilhança (EMV) $\hat{\boldsymbol{\vartheta}}$ de $\boldsymbol{\vartheta} = (\theta_1, \dots, \theta_k, p_0)'$ é determinado pela maximização da função de log-verossimilhança, $l = l(\boldsymbol{\theta}, p_0; \mathcal{D}) = \log L(\boldsymbol{\theta}, p_0; \mathcal{D})$, com respeito a $\boldsymbol{\theta}$ e p_0 . Admitindo que $l(\boldsymbol{\theta}, p_0; \mathcal{D}) = l(\boldsymbol{\vartheta}; \mathcal{D})$ é contínua e diferenciável em Θ , o sistema formado pelas equações

$$\frac{\partial l(\boldsymbol{\vartheta}; \mathcal{D})}{\partial \vartheta_i} = 0, \quad i = 1, \dots, k + 1$$

é resolvido por meio de algoritmos computacionais de otimização não-linear, por exemplo, Newton-Raphson e quase-Newton (Nocedal & Wright, 1999).

A matriz de informação esperada ou matriz de informação de Fisher denotada por $\mathbf{I}(\boldsymbol{\vartheta})$ tem componentes dadas por

$$I_{ij} = -E \left\{ \frac{\partial^2 l(\boldsymbol{\vartheta}; \mathcal{D})}{\partial \vartheta_i \partial \vartheta_j} \right\}, \quad i, j = 1, \dots, k + 1. \quad (3.6)$$

Pela propriedade de normalidade assintótica do EMV de $\boldsymbol{\vartheta}$, a distribuição assintótica de $\hat{\boldsymbol{\vartheta}}$ é normal multivariada com valor esperado $\boldsymbol{\vartheta}$ e matriz de covariância $\mathbf{I}(\boldsymbol{\vartheta})^{-1}$, isto é, $\hat{\boldsymbol{\vartheta}} \approx \mathcal{N}(\boldsymbol{\vartheta}, \mathbf{I}(\boldsymbol{\vartheta})^{-1})$. Um estimador consistente de $\mathbf{I}(\boldsymbol{\vartheta})$ é a matriz de informação observada $\mathbf{I}_0(\boldsymbol{\vartheta})$ cujas componentes são dadas por

$$-\frac{\partial^2 l(\boldsymbol{\vartheta})}{\partial \vartheta_i \partial \vartheta_j} \Big|_{\boldsymbol{\vartheta}=\hat{\boldsymbol{\vartheta}}}, \quad i, j = 1, \dots, k + 1.$$

Um intervalo de confiança de $100(1 - \alpha)\%$ para um determinado parâmetro ϑ_j , baseado na equação (3.6), é dado por $\hat{\vartheta}_j \pm z_{\alpha/2} \sqrt{I_{jj}^{-1}}$, em que I_{jj}^{-1} é o elemento (j, j) da matriz $\mathbf{I}^{-1}(\boldsymbol{\vartheta})$ e $z_{\alpha/2}$ é o quantil $\alpha/2$ de uma distribuição normal padronizada. Para testar a hipótese $H_0 : \boldsymbol{\vartheta} = \boldsymbol{\vartheta}_0$, podemos utilizar a estatística da razão de verossimilhanças $\Lambda = -2[l(\boldsymbol{\vartheta}_0) - l(\hat{\boldsymbol{\vartheta}})]$, sendo que Λ tem distribuição assintótica $\chi^2_{(k+1)}$.

Podemos observar que o cálculo da matriz de informação de Fisher correspondente à função de verossimilhança (3.5), se for realizado pela aplicação dos Teoremas A.2, A.3 e A.4 do Apêndice A, é menos oneroso em relação ao cálculo baseado nas derivadas parciais de segunda ordem definido em (3.6), pelo menos para as componentes que envolvem p_0 .

Embora esta seção tenha o enfoque da censura não-informativa, devemos chamar atenção para a existência de uma relação funcional entre os parâmetros das distribuições T e X , que possui papel fundamental para caracterizar a censura informativa da Definição 3.2. Esta relação funcional é expressada pela probabilidade κ^* definida na próxima subseção.

3.1.2 Relação Funcional Entre os Parâmetros das Distribuições de T e X

Uma configuração de dependência entre os parâmetros das distribuições de T , X , M e δ_i , $i = 1, \dots, n$ segue do fato de que a probabilidade de censura do i -ésimo indivíduo, p_c , é dada por

$$\begin{aligned} p_c &= P(\delta_i = 0) \\ &= p_0 \overbrace{P(\delta_i = 0 | M_i = 0)}^1 + (1 - p_0) \overbrace{P(\delta_i = 0 | M_i = 1)}^{\kappa^*}, \quad i = 1, \dots, n. \end{aligned} \quad (3.7)$$

Logo,

$$p_c = p_0 + (1 - p_0)\kappa^*. \quad (3.8)$$

Nas equações (3.7) e (3.8), κ^* representa a probabilidade do i -ésimo indivíduo ser censurado dado que está em risco. Esta probabilidade determina uma relação funcional entre os parâmetros das distribuições de T e X caracterizada pelo seguinte teorema.

Teorema 3.1 Considere o modelo de mistura padrão dado em (3.1) e seja X uma v.a. representando o tempo de censura com função de distribuição $G(x; \psi)$ independente da v.a. T , o tempo de sobrevivência dos indivíduos em risco. Então, a

probabilidade κ^* do i -ésimo indivíduo ser censurado dado que está em risco, é dada por

$$\kappa^* = \int_0^{\infty} S(x; \boldsymbol{\theta}|M_i = 1) g(x; \boldsymbol{\psi}|M_i = 1) dx, \quad i = 1, 2, \dots, n, \quad (3.9)$$

em que $g(\cdot) = G'(\cdot)$.

Prova. Temos que

$$\begin{aligned} \kappa^* &= P(\delta_i = 0|M_i = 1) = P(T > X|M_i = 1) \\ &= \int \int_{\{(t,x): t>x, x>0\}} f(t, x; \boldsymbol{\theta}, \boldsymbol{\psi}|M_i = 1) dt dx, \\ &= \int \int_{\{(t,x): t>x, x>0\}} f(t; \boldsymbol{\theta}|M_i = 1) g(x; \boldsymbol{\psi}|M_i = 1) dt dx \\ &= \int_0^{\infty} \left\{ \int_x^{\infty} f(t; \boldsymbol{\theta}|M_i = 1) dt \right\} g(x; \boldsymbol{\psi}|M_i = 1) dx. \end{aligned}$$

Portanto, κ^* é dada pela equação (3.9). ■

3.1.3 Censura Não-Informativa Uniforme-Exponencial

Utilizamos nesta subseção uma distribuição uniforme-exponencial para o tempo de censura conforme a equação (1.1), com parâmetros λ_c e T_0 , isto é $X \sim UExp(\lambda_c, T_0)$. Assumimos no modelo de mistura padrão (3.1) e função de verossimilhança (3.5) que os tempos de vida dos indivíduos em risco seguem uma distribuição exponencial com parâmetro λ e função densidade de probabilidade (fdp) $f(t; \lambda) = f(t; \lambda|M = 1)$. Desse modo, na função de verossimilhança (3.5) $\boldsymbol{\theta} = \lambda$ e a função de sobrevivência de longa duração do i -ésimo indivíduo é dada por

$$S_p(y_i; \lambda, p_0) = p_0 + (1 - p_0)S(y_i; \lambda|M = 1), \quad (3.10)$$

sendo que $S(y_i; \lambda|M = 1) = e^{-\lambda y_i}$. Então, a função densidade imprópria é dada por

$$f_p(y_i; \lambda, p_0) = (1 - p_0)f(y_i; \lambda) = (1 - p_0)\lambda e^{-\lambda y_i}. \quad (3.11)$$

Para $T \sim Exp(\lambda)$ e $X \sim UExp(\lambda_c, T_0)$ em (3.9), após alguma álgebra obtemos

$$\begin{aligned} \kappa^*(\lambda, \lambda_c) &= \frac{1}{\lambda + \lambda_c} [1 - e^{-(\lambda + \lambda_c)T_0}] \left(\frac{\lambda}{T_0(\lambda + \lambda_c)} + \lambda_c \right) \\ &\quad + \lambda_c(\lambda + \lambda_c) e^{-(\lambda + \lambda_c)T_0}. \end{aligned} \quad (3.12)$$

Neste caso, conseguimos a determinação analítica de κ^* , mas em geral métodos de integração numérica são requeridos.

Assumindo que p_c e p_0 são fixos e desconhecidos, a relação funcional entre os parâmetros λ e λ_c expressa pela quantidade $\kappa^*(\lambda, \lambda_c)$ na equação (3.12) estabelece que o mecanismo de censura não pode ser ignorado, caracterizando a censura informativa, a qual será considerada na Seção 3.2. Para introduzir esta dependência na função de verossimilhança por meio da distribuição de censura, percebemos que o procedimento natural requer a determinação de λ_c em função de κ^* , λ e T_0 . Em várias tentativas de uma solução geral em λ_c não obtivemos sucesso. Felizmente, a solução sugerida quando $T_0 \rightarrow \infty$ fornece uma importante reparametrização em (3.12) do ponto de vista de interpretação estatística e objetivos de simulação. Esta reparametrização é dada por

$$\kappa = \frac{\lambda_c}{\lambda + \lambda_c}, \quad (3.13)$$

ou

$$\lambda_c = \frac{\lambda\kappa}{1 - \kappa}. \quad (3.14)$$

Note que, das equações (3.12) e (3.13),

$$\lim_{T_0 \rightarrow \infty} \kappa^*(\lambda, \lambda_c) = \kappa.$$

De acordo com a equação (3.13), κ representa a proporção da taxa de falha da distribuição de censura em relação à taxa de falha total quando as variáveis T e X são exponencialmente distribuídas. Outra propriedade de $\kappa^*(\lambda, \lambda_c)$ que faremos uso é dada por

$$\lim_{\lambda_c \rightarrow 0} \kappa^*(\lambda, \lambda_c) = \frac{1}{\lambda T_0} (1 - e^{-\lambda T_0}). \quad (3.15)$$

A equação (3.15) representa o quociente entre a probabilidade de falha no intervalo $(0, T_0)$ e a taxa de falha até o tempo T_0 dos indivíduos em risco. Esta equação representa ainda o limite inferior de $\kappa^*(\lambda, \lambda_c)$ quando $\lambda_c \rightarrow 0$, κ_{inf}^* . Logo, a partir das equações (3.12) e (3.15) temos que

$$\kappa_{inf}^* = \frac{1}{\lambda T_0} (1 - e^{-\lambda T_0}). \quad (3.16)$$

Com as equações (3.8) e (3.16) estabelecemos o limite superior de p_0 , representado por

$$p_{0sup} = \frac{p_c - \kappa_{inf}^*}{1 - \kappa_{inf}^*}.$$

Apresentamos na Figura 3.1 os gráficos de $\kappa^* = \kappa^*(\lambda, \lambda_c)$ e κ em função de p_0 .

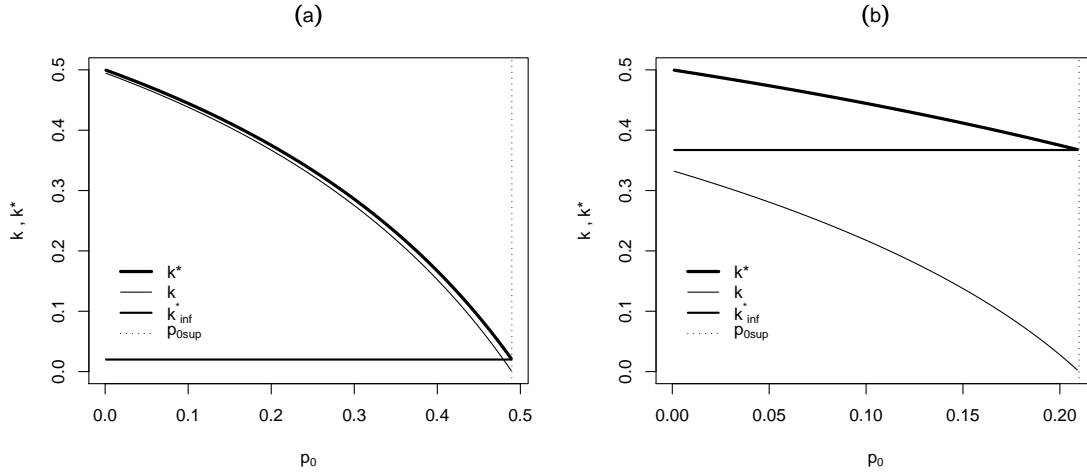


FIGURA 3.1: Probabilidade de um indivíduo ser censurado dado que está em risco (κ^* , κ) versus p_0 para $\lambda = 0,25$ e $p_c = 0,50$. (a) $T_0 = 200$ e $p_{0sup} = 0,4898$; (b) $T_0 = 10$ e $p_{0sup} = 0,2099$.

Utilizando as equações (3.10) e (3.11), a função de verossimilhança com censura não-informativa é dada por

$$L(\lambda, p_0; \mathcal{D}) \propto \prod_{i=1}^n [(1 - p_0)\lambda e^{-\lambda y_i}]^{\delta_i} [p_0 + (1 - p_0)e^{-\lambda y_i}]^{1 - \delta_i}. \quad (3.17)$$

A função de log-verossimilhança correspondente à equação (3.17) é dada por

$$l(\lambda, p_0; \mathcal{D}) \propto \sum_{i=1}^n \left\{ \delta_i [\log(1 - p_0) + \log \lambda - \lambda y_i] + (1 - \delta_i) [\log(p_0 + (1 - p_0)e^{-\lambda y_i})] \right\}.$$

As segundas derivadas parciais de $l(\lambda, p_0; \mathcal{D})$, com respeito a λ e p_0 são dadas por

$$\frac{\partial^2 l}{\partial \lambda^2} = \sum_{i=1}^n \left\{ -\frac{\delta_i}{\lambda^2} + (1 - \delta_i) \frac{p_0(1 - p_0)y_i^2 e^{-\lambda y_i}}{[p_0 + (1 - p_0)e^{-\lambda y_i}]^2} \right\}, \quad (3.18)$$

$$\frac{\partial^2 l}{\partial p_0 \partial \lambda} = \sum_{i=1}^n (1 - \delta_i) \frac{y_i e^{-\lambda y_i}}{[p_0 + (1 - p_0) e^{-\lambda y_i}]^2} \quad (3.19)$$

e

$$\frac{\partial^2 l}{\partial p_0^2} = \sum_{i=1}^n \left\{ \frac{-\delta_i}{(1 - p_0)^2} - (1 - \delta_i) \frac{(1 - e^{-\lambda y_i})^2}{[p_0 + (1 - p_0) e^{-\lambda y_i}]^2} \right\}. \quad (3.20)$$

A partir dos Teoremas A.2, A.3 e A.4 do Apêndice A e de (3.18) – (3.20) os valores esperados das segundas derivadas parciais de $-l(\lambda, p_0; \mathcal{D})$ com respeito a λ e p_0 são dados por

$$E \left\{ -\frac{\partial^2 l}{\partial \lambda^2} \right\} = (1 - p_0) n E \left\{ \frac{1 - e^{-\lambda X}}{\lambda^2} - \frac{p_0 X^2 e^{-\lambda X}}{p_0 + (1 - p_0) e^{-\lambda X}} \right\} = n I_{\lambda \lambda}, \quad (3.21)$$

$$E \left\{ -\frac{\partial^2 l}{\partial p_0 \partial \lambda} \right\} = -n E \left\{ \frac{X e^{-\lambda X}}{p_0 + (1 - p_0) e^{-\lambda X}} \right\} = n I_{p_0 \lambda} = n I_{\lambda p_0} \quad (3.22)$$

e

$$E \left\{ -\frac{\partial^2 l}{\partial p_0^2} \right\} = \frac{1}{1 - p_0} n E \left\{ \frac{1 - e^{-\lambda X}}{p_0 + (1 - p_0) e^{-\lambda X}} \right\} = n I_{p_0 p_0}. \quad (3.23)$$

Para a determinação analítica dos valores esperados nas equações (3.21) – (3.23) com respeito à variável aleatória $X \sim UExp(\lambda_c, T_0)$, definimos a transformação paramétrica

$$\varphi^*(\alpha, \nu, \beta) = \int_0^{T_0} \frac{x^{\nu-1} e^{-\alpha x}}{1 + \beta e^{-x}} dx, \quad (3.24)$$

com $\alpha > 0$, $\beta = \frac{1 - p_0}{p_0} > 0$ e $\nu = 1, 2, \dots$, a qual chamaremos de transformação integral φ^* -paramétrica.

A transformação integral φ^* -paramétrica, dada pela equação (3.24), possui várias propriedades relacionadas com as funções Lerch, Lerch-Zeta, Zeta de Riemann, dilogaritmo, exponencial-integral, entre outras, que podem ser encontradas em Abramowitz & Stegun (1964) e Gradshteyn & Ryzhik (2007). Algumas dessas propriedades são dadas por

- $\lim_{T_0 \rightarrow \infty} \varphi^*(\alpha, \nu, 1) = (\nu - 1)! \Phi(-1, \nu, \alpha)$, em que $\Phi(\cdot, \cdot, \cdot)$ é a função Lerch (Gradshteyn & Ryzhik, 2007) com representação integral dada por

$$\Phi(\zeta, \alpha, \nu) = \frac{1}{\Gamma(\alpha)} \int_0^\infty \frac{x^{\alpha-1} e^{-\nu x}}{1 - \zeta e^{-x}} dx, \quad \zeta \neq 1, \alpha > 0, \nu > 0.$$

- $\varphi^*(\alpha, \nu, \beta) \leq \frac{(\nu - 1)!}{\alpha^\nu} E\left(\frac{1}{1 + \beta e^{-X}}\right)$, se $X \sim \text{Gama}(\nu, \alpha)$ com função densidade de probabilidade dada por

$$f(x) = \frac{\alpha^\nu}{\Gamma(\nu)} x^{\nu-1} e^{-\alpha x}, \quad \alpha > 0, \quad \nu = 1, 2, \dots$$

- $\lim_{T_0 \rightarrow \infty} \varphi^*(1, \nu, \beta) = \frac{(\nu - 1)!}{-\beta} Li_\nu(-\beta)$, em que $Li_\nu(\cdot)$ é a função polilogaritmo (Cvijović & Klinowski, 1997) com representação em série dada por

$$Li_\nu(x) = \sum_{j=1}^{\infty} \frac{x^j}{j^\nu}, \quad \nu \geq 1, \quad |x| < 1.$$

Desenvolvemos algumas soluções analíticas e computacionais com representação em série da integral φ^* com uso de métodos tradicionais como o método de Simpson, trapezóides, pontos médios e combinação destes. Nenhum destes métodos de integração numérica que utilizamos produziu resultados satisfatórios em relação ao tempo computacional e magnitude do erro, principalmente para valores de p_0 próximos aos extremos do intervalo $(0, 1)$. Felizmente, para $0 < p_0 < 1$ observamos que as soluções por integração de Monte Carlo produziram resultados satisfatoriamente equivalentes às soluções obtidas com a função de integração numérica da linguagem R (R Development Core Team, 2009), que utiliza quadratura adaptativa. Portanto, para desenvolver a análise de sensibilidade da censura informativa motivada pela matriz de informação de Fisher, a transformação integral φ^* será obtida numericamente com a função “integrate” da linguagem R, implicando num custo computacional bastante inferior em relação às soluções por integração de Monte Carlo.

Para determinar a matriz de informação de Fisher utilizando as equações (3.21) – (3.24) considere o seguinte teorema.

Teorema 3.2 Sob o modelo de mistura padrão com fração de cura p_0 , sejam $T_i \sim \text{Exp}(\lambda)$ e $X_i \sim \text{UExp}(\lambda_c, T_0)$ o tempo de sobrevivência e o tempo de censura do i -ésimo indivíduo em risco variáveis aleatórias i.i.d., $i = 1, \dots, n$, respectivamente e suponha que a censura seja não-informativa. Então, a matriz de informação de Fisher, baseada na função de verossimilhança (3.17), é dada por

$$\mathbf{I}(\lambda, p_0) = n \begin{bmatrix} I_{\lambda\lambda} & I_{\lambda p_0} \\ I_{p_0\lambda} & I_{p_0 p_0} \end{bmatrix}, \quad (3.25)$$

em que

$$\begin{aligned} I_{\lambda\lambda} = & \frac{(1-p_0)}{\lambda^2} \left[1 - \left(\frac{(1-\kappa)^2}{\lambda T_0} + \kappa \right) (1 - \exp(-\frac{\lambda T_0}{1-\kappa})) - \frac{\lambda^2 \kappa^2}{(1-\kappa)^2} \exp(-\frac{\lambda T_0}{1-\kappa}) \right] \\ & - \frac{(1-p_0)}{\lambda^2} \left(\frac{1}{\lambda T_0} + \frac{\kappa}{1-\kappa} \right) \varphi^* \left(\frac{1}{1-\kappa}, 3, \frac{1-p_0}{p_0} \right) \\ & + \frac{1-p_0}{\lambda^3 T_0} \frac{\kappa}{1-\kappa} \varphi^* \left(\frac{1}{1-\kappa}, 4, \frac{1-p_0}{p_0} \right), \end{aligned} \quad (3.26)$$

$$\begin{aligned} I_{\lambda p_0} = I_{p_0\lambda} = & - \left(\frac{1}{T_0} + \lambda \frac{\kappa}{1-\kappa} \right) \frac{1}{\lambda^2 p_0} \varphi^* \left(\frac{1}{1-\kappa}, 2, \frac{1-p_0}{p_0} \right) \\ & + \frac{\kappa}{1-\kappa} \frac{1}{\lambda^2 T_0} \varphi^* \left(\frac{1}{1-\kappa}, 3, \frac{1-p_0}{p_0} \right) \end{aligned} \quad (3.27)$$

e

$$\begin{aligned} I_{p_0 p_0} = & \frac{1}{\lambda p_0 (1-p_0)} \left(\frac{1}{T_0} + \lambda \frac{\kappa}{1-\kappa} \right) \varphi^* \left(\frac{\kappa}{1-\kappa}, 1, \frac{1-p_0}{p_0} \right) \\ & - \frac{1}{\lambda p_0 (1-p_0)} \frac{\kappa}{T_0 (1-\kappa)} \varphi^* \left(\frac{\kappa}{1-\kappa}, 2, \frac{1-p_0}{p_0} \right) \\ & - \frac{1}{\lambda p_0 (1-p_0)} \left(\frac{1}{T_0} + \lambda \frac{\kappa}{1-\kappa} \right) \varphi^* \left(\frac{\kappa}{1-\kappa}, 1, \frac{1-p_0}{p_0} \right) \\ & + \frac{1}{\lambda p_0 (1-p_0)} \frac{\kappa}{T_0 (1-\kappa)} \varphi^* \left(\frac{1}{1-\kappa}, 2, \frac{1-p_0}{p_0} \right). \end{aligned} \quad (3.28)$$

Prova. Após uma aplicação direta das equações (3.21) – (3.24) os elementos $I_{\lambda\lambda}$, $I_{\lambda p_0}$ e $I_{p_0 p_0}$ da matriz de informação $\mathbf{I}(\lambda, p_0)$ são obtidos. ■

Observe que a matriz $\mathbf{I}(\lambda, p_0)$ envolve a quantidade κ definida pela equação (3.13), caracterizando que os mecanismos de censura são informativos. Para resolver este problema consideramos a censura informativa na Seção 3.2.

3.1.4 Função de Verossimilhança com Dados Ampliados

Uma maneira mais atrativa de representar a função de verossimilhança (3.17) consiste na introdução de uma variável latente W no modelo, consequentemente aumentando o conjunto de dados. Note que $L(\lambda, p_0; \mathcal{D})$ pode ser escrita do

seguinte modo:

$$L(\lambda, p_0; \mathcal{D}) \propto (1 - p_0)^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n [f(y_i; \lambda)]^{1-\delta_i} \prod_{i=1}^n [p_0 + (1 - p_0)S(y_i; \lambda)]^{1-\delta_i}. \quad (3.29)$$

Observamos que a equação (3.29) pode ser expressa numa forma fatorada com a inclusão das variáveis latentes W_i , $i = 1, \dots, n$ independentes, seguindo uma distribuição de Bernoulli com parâmetro p_i dado por

$$p_i = \frac{p_0}{p_0 + (1 - p_0)S(y_i; \lambda)}, \quad i = 1, \dots, n. \quad (3.30)$$

Logo, a função de verossimilhança ampliada pelos dados artificiais $\{w_1, \dots, w_n\}$ é dada por

$$\begin{aligned} L_a(\lambda, p_0; \mathcal{D}) &\propto (1 - p_0)^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n [f(y_i; \lambda)]^{1-\delta_i} \prod_{i=1}^n [p_0 + (1 - p_0)S(y_i; \lambda)]^{1-\delta_i} \\ &\times \prod_{i=1}^n \left[\left(\frac{p_0}{p_0 + (1 - p_0)S(y_i; \lambda)} \right)^{w_i} \right]^{1-\delta_i} \\ &\times \prod_{i=1}^n \left[\left(1 - \frac{p_0}{p_0 + (1 - p_0)S(y_i; \lambda)} \right)^{1-w_i} \right]^{1-\delta_i}. \end{aligned} \quad (3.31)$$

Reescrevendo a equação (3.31) temos que

$$L_a(\lambda, p_0; \mathcal{D}) \propto p_0^{\sum_{i=1}^n (1-\delta_i)w_i} (1 - p_0)^{n - \sum_{i=1}^n (1-\delta_i)w_i} \prod_{i=1}^n [f(y_i; \lambda)]^{\delta_i} [S(y_i; \lambda)]^{(1-w_i)(1-\delta_i)}. \quad (3.32)$$

Logo, uma forma fatorada de (3.17) é dada por

$$L_a(\lambda, p_0; \mathcal{D}) \propto p_0^{\sum_{i=1}^n (1-\delta_i)w_i} (1 - p_0)^{n - \sum_{i=1}^n (1-\delta_i)w_i} \lambda^{\sum_{i=1}^n \delta_i} e^{-\lambda \sum_{i=1}^n y_i [1 - (1-\delta_i)w_i]}. \quad (3.33)$$

Podemos observar que a forma fatorada de $L_a(\lambda, p_0; \mathcal{D})$ em (3.32) simplifica a função de verossimilhança (3.29) permitindo a implementação de algoritmos computacionais mais simples nos contextos clássico e Bayesiano. Um estudo de simulação com inferência Bayesiana via métodos de Monte Carlo em cadeias de Markov (MCMC) e dados ampliados (Tanner, 1996) será objeto de pesquisas futuras (vide Capítulo 6).

3.2 Censura Informativa

Nesta seção iremos considerar o mesmo conjunto de dados e variáveis aleatórias definidas por (3.2) – (3.4) e distribuições de probabilidade $X \sim UExp(\lambda_c, T_0)$ e (3.10).

3.2.1 Função de Verossimilhança e Matriz de Informação de Fisher

A função de verossimilhança com censura informativa, motivada por Lagakos (1979), Lawless (1982) e Maller & Zhou (1996), é dada por

$$L^*(\boldsymbol{\theta}, p_0, \boldsymbol{\psi}; \mathcal{D}) = \prod_{i=1}^n f_p^{\delta_i}(y_i; \boldsymbol{\theta}, p_0) S_p^{1-\delta_i}(y_i; \boldsymbol{\theta}, p_0) \prod_{i=1}^n 1 - G(y_i; \boldsymbol{\psi})^{\delta_i} g^{1-\delta_i}(y_i; \boldsymbol{\psi}). \quad (3.34)$$

sendo que $g(\cdot) = G'(\cdot)$ e $\boldsymbol{\psi}$ é um vetor de parâmetros associado à distribuição de censura $G(\cdot)$. Em (3.34) estamos assumindo que o segundo produtório não pode ser ignorado porque depende de $\boldsymbol{\theta}$ (Lawless, 1982). Fica estabelecida dessa forma a função de verossimilhança com censura informativa que utilizamos neste trabalho, caracterizada pela relação funcional entre os vetores $\boldsymbol{\theta}$ e $\boldsymbol{\psi}$.

Podemos escrever $l^*(\boldsymbol{\theta}, p_0, \boldsymbol{\psi}; \mathcal{D}) = \log L^*(\boldsymbol{\theta}, p_0, \boldsymbol{\psi}; \mathcal{D})$, a função de log-verossimilhança com censura informativa, da seguinte maneira:

$$l^*(\boldsymbol{\theta}, p_0, \boldsymbol{\psi}; \mathcal{D}) = l(\boldsymbol{\theta}, p_0; \mathcal{D}) + \sum_{i=1}^n [\delta_i \log(1 - G(y_i; \boldsymbol{\psi})) + (1 - \delta_i) \log g(y_i; \boldsymbol{\psi})], \quad (3.35)$$

em que $l(\boldsymbol{\theta}, p_0; \mathcal{D})$ decorre de (3.17).

A matriz de informação de Fisher $\mathbf{I}^*(\boldsymbol{\theta}, p_0, \boldsymbol{\psi})$ baseada em (3.35) pode ser obtida pela aplicação dos Teoremas A.1, A.2 e A.3 do Apêndice A.

3.2.2 Censura Informativa Uniforme-Exponencial

Considere $\boldsymbol{\theta} = \lambda$ e $\boldsymbol{\psi} = \lambda_c$ na equação (3.34). Então,

$$L^*(\lambda, p_0, \lambda_c; \mathcal{D}) = \prod_{i=1}^n [f_p(y_i; \lambda, p_0)(1 - G(y_i; \lambda_c))]^{\delta_i} [S_p(y_i; \lambda, p_0)g(y_i; \lambda_c)]^{1-\delta_i}. \quad (3.36)$$

A função de log-verossimilhança correspondente à função (3.36) é dada por

$$l^*(\lambda, p_0, \lambda_c; \mathcal{D}) = \sum_{i=1}^n \delta_i \log[f_p(y_i; \lambda, p_0)(1 - G(y_i; \lambda_c))] + \sum_{i=1}^n (1 - \delta_i) \log[S_p(y_i; \lambda, p_0)g(y_i; \lambda_c)]. \quad (3.37)$$

Note que

$$l^*(\lambda, p_0, \lambda_c; \mathcal{D}) = l(\lambda, p_0; \mathcal{D}) + \sum_{i=1}^n [\delta_i \log(1 - G(y_i; \lambda_c)) + (1 - \delta_i) \log(g(y_i; \lambda_c))].$$

Das equações (1.1), (3.10) e (3.14), segue que

$$l^*(\lambda, p_0, \kappa; \mathcal{D}) = \sum_{i=1}^n \delta_i [\log(1 - p_0) + \log \lambda - \lambda y_i + \log(1 - \frac{y_i}{T_0})] + \sum_{i=1}^n (1 - \delta_i) [\log(p_0 + (1 - p_0)e^{-\lambda y_i})] + \sum_{i=1}^n \left\{ (1 - \delta_i) \log\left[\frac{1}{T_0} + \frac{\lambda \kappa}{1 - \kappa} \left(1 - \frac{y_i}{T_0}\right)\right] - \frac{\lambda \kappa}{1 - \kappa} y_i \right\}. \quad (3.38)$$

As segundas derivadas parciais de (3.38), com respeito a λ , p_0 e κ são dadas por

$$\begin{aligned} \frac{\partial^2 l^*}{\partial \lambda^2} &= \frac{\partial^2 l}{\partial \lambda^2} - \sum_{i=1}^n (1 - \delta_i) \frac{(\kappa/(1 - \kappa))^2 (1 - \frac{y_i}{T_0})^2}{\left[\frac{1}{T_0} + \frac{\lambda \kappa}{1 - \kappa} \left(1 - \frac{y_i}{T_0}\right)\right]^2}, \\ \frac{\partial^2 l^*}{\partial \kappa^2} &= \sum_{i=1}^n (1 - \delta_i) \left[\frac{(1 - \frac{y_i}{T_0}) 2\lambda}{(1 - \kappa)^3 \left[\frac{1}{T_0} + \frac{\lambda \kappa}{1 - \kappa} \left(1 - \frac{y_i}{T_0}\right)\right]} - \frac{(1 - \frac{y_i}{T_0})^2 \lambda^2}{(1 - \kappa)^4 \left[\frac{1}{T_0} + \frac{\lambda \kappa}{1 - \kappa} \left(1 - \frac{y_i}{T_0}\right)\right]^2} \right] \\ &\quad - \sum_{i=1}^n \left[(1 - \delta_i) \frac{2\lambda y_i}{(1 - \kappa)^3} + \delta_i \frac{2\lambda y_i}{(1 - \kappa)^3} \right], \\ \frac{\partial^2 l^*}{\partial \lambda \partial \kappa} &= \sum_{i=1}^n (1 - \delta_i) \left[\frac{1 - \frac{y_i}{T_0}}{(1 - \kappa)^2 \left[\frac{1}{T_0} + \frac{\lambda \kappa}{1 - \kappa} \left(1 - \frac{y_i}{T_0}\right)\right]} - \frac{(1 - \frac{y_i}{T_0})^2 \lambda \kappa}{(1 - \kappa)^3 \left[\frac{1}{T_0} + \frac{\lambda \kappa}{1 - \kappa} \left(1 - \frac{y_i}{T_0}\right)\right]^2} \right] \\ &\quad - \sum_{i=1}^n \left[(1 - \delta_i) \frac{y_i}{(1 - \kappa)^2} + \delta_i \frac{y_i}{(1 - \kappa)^2} \right], \\ \frac{\partial^2 l^*}{\partial \kappa \partial \lambda} &= \frac{\partial^2 l^*}{\partial \lambda \partial \kappa}, \quad \frac{\partial^2 l^*}{\partial p_0 \partial \lambda} = \frac{\partial^2 l^*}{\partial \lambda \partial p_0} = \frac{\partial^2 l}{\partial p_0 \partial \lambda}, \quad \frac{\partial^2 l^*}{\partial p_0 \partial \kappa} = \frac{\partial^2 l^*}{\partial \kappa \partial p_0} = 0 \text{ e } \frac{\partial^2 l^*}{\partial p_0^2} = \frac{\partial^2 l}{\partial p_0^2}. \end{aligned}$$

Observe que $\frac{\partial^2 l}{\partial \lambda^2}$, $\frac{\partial^2 l}{\partial p_0 \partial \lambda}$ e $\frac{\partial^2 l}{\partial p_0^2}$ são dadas pelas equações (3.18), (3.19) e (3.20), respectivamente.

Os valores esperados das segundas derivadas parciais de $-l^*(\lambda, p_0, \kappa; \mathcal{D})$, com respeito a λ , p_0 e κ , obtidos pela aplicação dos Teoremas A.1 – A.3 do Apêndice A e por (3.26) – (3.28) são dados por

$$E \left\{ -\frac{\partial^2 l^*}{\partial \lambda^2} \right\} = nI_{\lambda\lambda}^* = nI_{\lambda\lambda} + nI_{\lambda\lambda}^c, \quad (3.39)$$

em que $nI_{\lambda\lambda} = E \left\{ -\frac{\partial^2 l}{\partial \lambda^2} \right\}$,

$$I_{\lambda\lambda}^c = \frac{\kappa^2}{(1-\kappa)^2} E \left\{ S_p(X) \frac{(1 - \frac{X}{T_0})^2}{[\frac{1}{T_0} + \frac{\lambda\kappa}{1-\kappa}(1 - \frac{X}{T_0})]^2} \right\},$$

$$E \left\{ -\frac{\partial^2 l^*}{\partial \kappa^2} \right\} = nE \{ S_p(X) Q_1(X) + Q_2(X) \} = nI_{\kappa\kappa}^*, \quad (3.40)$$

com

$$Q_1(X) = \frac{-(1 - \frac{X}{T_0})2\lambda}{(1-\kappa)^3[\frac{1}{T_0} + \frac{\lambda\kappa}{1-\kappa}(1 - \frac{X}{T_0})]} + \frac{(1 - \frac{X}{T_0})^2\lambda^2}{(1-\kappa)^4[\frac{1}{T_0} + \frac{\lambda\kappa}{1-\kappa}(1 - \frac{X}{T_0})]^2} + \frac{2\lambda X}{(1-\kappa)^3}$$

e

$$Q_2(X) = \frac{(1-p_0)2\lambda}{(1-\kappa)^3} \left[\frac{1}{\lambda}(1 - e^{-\lambda X}) - X e^{-\lambda X} \right],$$

$$E \left\{ -\frac{\partial^2 l^*}{\partial \lambda \partial \kappa} \right\} = nE \{ S_p(X) Q_3(X) + Q_4(X) \} = nI_{\lambda\kappa}^*, \quad (3.41)$$

com

$$Q_3(X) = \frac{-(1 - \frac{X}{T_0})}{(1-\kappa)^2[\frac{1}{T_0} + \frac{\lambda\kappa}{1-\kappa}(1 - \frac{X}{T_0})]} + \frac{(1 - \frac{X}{T_0})^2\lambda\kappa}{(1-\kappa)^3[\frac{1}{T_0} + \frac{\lambda\kappa}{1-\kappa}(1 - \frac{X}{T_0})]^2} + \frac{X}{(1-\kappa)^2}$$

e

$$Q_4(X) = \frac{(1-p_0)}{(1-\kappa)^2} \left[\frac{1}{\lambda}(1 - e^{-\lambda X}) - X e^{-\lambda X} \right],$$

$$E \left\{ -\frac{\partial^2 l^*}{\partial \lambda \partial p_0} \right\} = E \left\{ -\frac{\partial^2 l}{\partial \lambda \partial p_0} \right\} = nI_{\lambda p_0} = nI_{\lambda p_0}^*, \quad (3.42)$$

$$E \left\{ -\frac{\partial^2 l^*}{\partial p_0^2} \right\} = E \left\{ -\frac{\partial^2 l}{\partial p_0^2} \right\} = nI_{p_0 p_0} = nI_{p_0 p_0}^* \quad (3.43)$$

e

$$E \left\{ -\frac{\partial^2 l^*}{\partial \kappa \partial p_0} \right\} = nI_{\kappa p_0}^* = 0. \quad (3.44)$$

Portanto, segue das equações (3.39) – (3.44) que a matriz de informação de Fisher é dada por

$$\mathbf{I}^*(\lambda, p_0, \kappa) = n \begin{bmatrix} I_{\lambda\lambda}^* & I_{\lambda p_0}^* & I_{\lambda\kappa}^* \\ I_{\lambda p_0}^* & I_{p_0 p_0}^* & 0 \\ I_{\lambda\kappa}^* & 0 & I_{\kappa\kappa}^* \end{bmatrix}, \quad (3.45)$$

em que $I_{\lambda p_0}^* = I_{\lambda p_0}$, $I_{p_0 p_0}^* = I_{p_0 p_0}$ e $I_{\lambda\lambda}^* = I_{\lambda\lambda} + I_{\lambda\lambda}^c$. Calculamos $I_{\lambda\lambda}^c$, $I_{\lambda\kappa}^*$ e $I_{\kappa\kappa}^*$ aplicando o Teorema A.1 do Apêndice A. Note que $I_{\lambda\lambda}^c$ é determinado analiticamente do seguinte modo:

$$I_{\lambda\lambda}^c = \frac{\kappa^2}{(1-\kappa)^2 T_0} \int_0^{T_0} \frac{p_0(T_0-x)^2 \exp\left(\frac{-\lambda\kappa x}{1-\kappa}\right) + (1-p_0)(T_0-x)^2 \exp\left(-\frac{\lambda x}{1-\kappa}\right)}{1 + \frac{\lambda\kappa}{1-\kappa}(T_0-x)} dx. \quad (3.46)$$

Para continuar a determinação de $I_{\lambda\lambda}^c$, definimos a transformação paramétrica

$$\varphi_1^*(\alpha, \nu, \lambda_c) = \int_0^{T_0} \frac{x^{\nu-1} e^{-\alpha x}}{1 + \lambda_c(T_0-x)} dx, \quad (3.47)$$

em que $\alpha > 0$, λ_c é dado por (3.14), $T_0 > 0$ e $\nu = 1, 2, \dots$, denominada de transformação integral φ_1^* -paramétrica.

Utilizando as equações (3.46) e (3.47), após algumas simplificações algébricas temos que

$$\begin{aligned} I_{\lambda\lambda}^c &= \frac{\kappa^2}{(1-\kappa)^2 T_0} \left\{ T_0^2 \varphi_1^* \left(\frac{\lambda\kappa}{1-\kappa}, 1, \frac{\lambda\kappa}{1-\kappa} \right) - 2T_0 \varphi_1^* \left(\frac{\lambda\kappa}{1-\kappa}, 2, \frac{\lambda\kappa}{1-\kappa} \right) \right\} \\ &+ \frac{\kappa^2(1-p_0)}{(1-\kappa)^2 T_0} \left\{ T_0^2 \varphi_1^* \left(\frac{\lambda}{1-\kappa}, 1, \frac{\lambda\kappa}{1-\kappa} \right) - 2T_0 \varphi_1^* \left(\frac{\lambda}{1-\kappa}, 2, \frac{\lambda\kappa}{1-\kappa} \right) \right\} \\ &+ \frac{\kappa^2}{(1-\kappa)^2 T_0} \varphi_1^* \left(\frac{\lambda\kappa}{1-\kappa}, 3, \frac{\lambda\kappa}{1-\kappa} \right) + \frac{\kappa^2(1-p_0)}{(1-\kappa)^2 T_0} \varphi_1^* \left(\frac{\lambda}{1-\kappa}, 3, \frac{\lambda\kappa}{1-\kappa} \right), \end{aligned} \quad (3.48)$$

em que κ é dado pela equação (3.13). Expressões analíticas de $I_{\lambda\kappa}^*$ e $I_{\kappa\kappa}^*$ podem ser derivadas com o uso da equação (3.47).

Note que $I_{\lambda\lambda}^* \geq I_{\lambda\lambda}$, o que resulta em não podermos ignorar o mecanismo de censura na estimação de λ e consequentemente na estimação de p_0 .

Neste capítulo destacamos a função de verossimilhança com censura não-informativa e com censura informativa uniforme-exponencial e a estimação dos parâmetros de interesse sob o ponto de vista clássico. Utilizamos estas funções de verossimilhanças no próximo capítulo em um estudo de simulação e em um exemplo com dados reais.

Capítulo 4

Estudo de Simulação e Aplicação com Dados Reais

Neste capítulo apresentamos um estudo de simulação do modelo de mistura padrão definido no Capítulo 3 e analisamos um conjunto de dados reais. Para simular amostras do modelo utilizamos o algoritmo que propomos na Seção 4.1 com implementação em linguagem R (vide Apêndice C). Usamos o método aceitação-rejeição (AR) com proposta exponencial para obter amostras aleatórias do modelo uniforme-exponencial. Uma apresentação detalhada do método AR pode ser encontrada, por exemplo, em Rubinstein & Kroese (2008). A obtenção das estimativas de máxima verossimilhança, intervalos de confiança assintóticos e probabilidades de cobertura, é dada pelo programa que desenvolvemos em R, Seção C.1 do Apêndice C. O método de Newton-Raphson (NR) introduzido inicialmente para obter as EMV conduziu o processo de otimização em determinadas regiões do espaço paramétrico a matrizes de informações singulares. Para evitar matrizes singulares no método NR, uma alternativa é o uso de uma reparametrização apropriada, como propõem Maller & Zhou (1996). Optamos por um esquema de maximização condicionada que usa uma rotina de maximização da linguagem R.

Os intervalos de confiança e probabilidades de cobertura foram estimados tendo como base as propriedades assintóticas dos EMV (Brown *et al.*, 2001; Berger, 2006).

4.1 Algoritmo Para Simulação do Modelo de Mistura Padrão

Para atender aos objetivos do estudo de simulação do modelo de mistura padrão com fração de cura p_0 , o algoritmo seguinte permite obter D réplicas de amostras aleatórias de tamanho n das variáveis aleatórias tempo de vida dos indivíduos em risco (T), tempo de censura (X), número de causas (M), indicador de censura (δ) e $Y = \min \{T, X\}$.

Algoritmo 4.1

(1) Amostra aleatória das variáveis aleatórias M , T e X .

(1.1) Gerar (m_1, \dots, m_n) de $M \sim \text{Bernoulli}(1 - p_0)$.

(1.2) Gerar (t_1, \dots, t_n) de f , a fdp de T .

(1.3) Gerar (x_1, \dots, x_n) de g , a fdp de X .

(2) Amostra aleatória das variáveis $Y = \min \{T, X\}$ e δ . Para $i = 1, \dots, n$,

(2.1) Fazer $t_i = \infty$, se $m_i = 0$. Caso contrário, fazer $t_i = t_i$.

(2.2) Fazer $y_i = \min \{t_i, x_i\}$.

(2.3) Fazer $\delta_i = 1$, se $t_i \leq x_i$. Caso contrário, fazer $\delta_i = 0$.

(3) Repetir os passos (1) e (2) até obter D réplicas.

Este algoritmo é semelhante ao algoritmo desenvolvido por Mizoi (2004) em linguagem Ox (Doornik, 2006) para o número de causas competindo para o evento de interesse com distribuição de Poisson.

4.2 Estudo de Simulação

Para verificar o impacto da censura informativa nas probabilidades de cobertura e comprimentos de intervalos de confiança da fração de cura e taxa de falha,

utilizamos conjuntos de dados gerados pelo Algoritmo 4.1 com $M \sim \text{Bernoulli}(1 - p_0)$, $T \sim \text{Exp}(\lambda)$ e $X \sim \text{UExp}(\lambda_c, T_0)$. Consideramos funções de verossimilhança com censura não-informativa e com censura informativa uniforme-exponencial.

Os intervalos de confiança assintóticos para λ e p_0 foram obtidos do seguinte modo:

$$\hat{\lambda} \pm z_{\alpha/2} \sqrt{\hat{I}_{11}^{-1}} \quad \text{e} \quad \hat{p}_0 \pm z_{\alpha/2} \sqrt{\hat{I}_{22}^{-1}} \quad ,$$

em que \hat{I}_{jj}^{-1} , $j = 1, 2$, denota o j -ésimo elemento da diagonal da inversa da matriz de informação, estimado conforme a censura seja não-informativa com matriz de informação observada ou informativa com matriz de informação de Fisher (3.45) e $z_{\alpha/2}$ é o quantil $\alpha/2$ de uma distribuição normal padrão. Considerando $\epsilon = 10^{-5}$, os EMV de $\boldsymbol{\theta} = (\lambda, p_0)'$ com censura não-informativa e os EMV de $\boldsymbol{\theta} = (\lambda, p_0, \kappa)'$ com censura informativa, foram obtidos quando na i -ésima iteração temos satisfeita a condição $|\hat{\boldsymbol{\theta}}^{(i)} - \hat{\boldsymbol{\theta}}^{(i-1)}| \leq \epsilon$.

4.2.1 Simulação com p_0 Não-Fixado

No estudo de simulação desenvolvemos uma inferência clássica baseada em 10000 réplicas de amostras aleatórias para $n = 50$ e 400 , $T_0 = 50$, $p_c = 0,50$, $\lambda = 0,25$ e $p_0 = 0,02, 0,04, \dots, 0,44$, $p_0 \leq p_{0sup} = 0,4565$. Para verificar o impacto da censura informativa uniforme-exponencial nas estimativas pontuais e intervalares, apresentamos as Figuras 4.1 – 4.6. Podemos observar as seguintes estimativas: probabilidades de cobertura, comprimentos de intervalos de confiança de 95% e EMV de λ e p_0 .

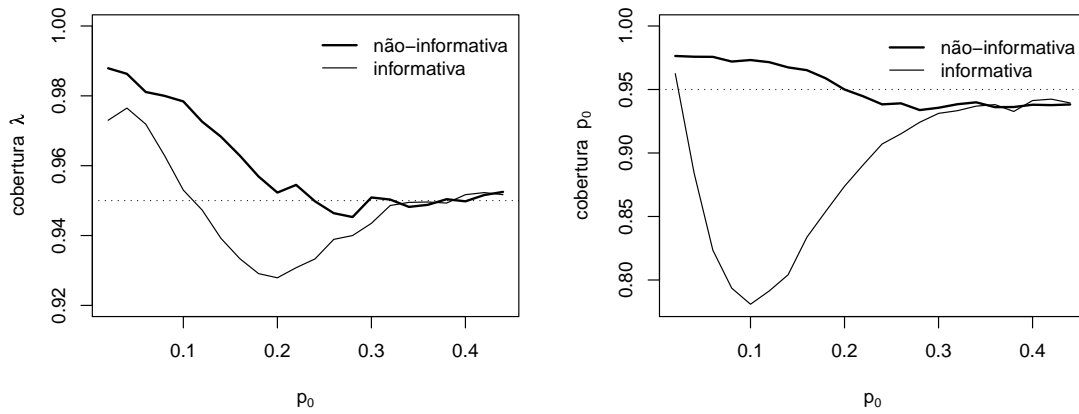


FIGURA 4.1: Probabilidades de cobertura com $n = 50$, $\lambda = 0,25$, $T_0 = 50$ e $p_c = 0,50$.

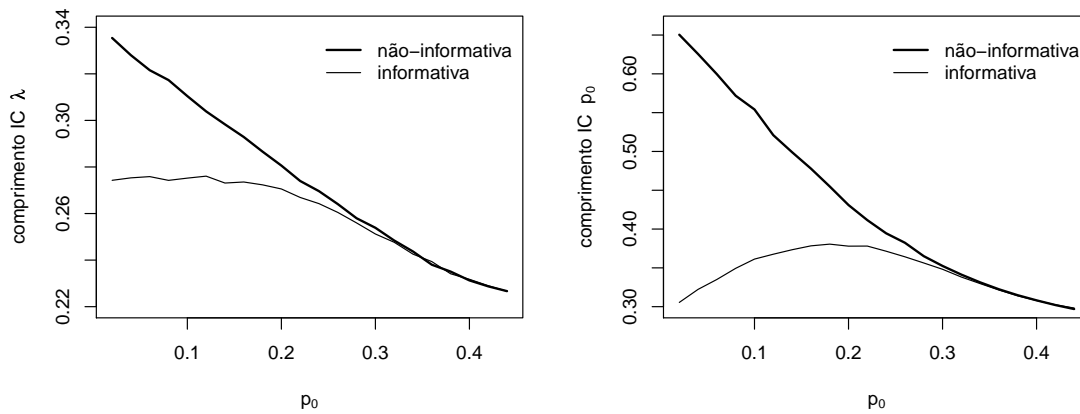


FIGURA 4.2: Comprimento de intervalos de confiança (IC) de 95% com $n = 50$, $\lambda = 0,25$, $T_0 = 50$ e $p_c = 0,50$.

As Figuras 4.1 e 4.2 para valores de $p_0 \leq 0,2$ revelam uma forte instabilidade nas estimativas das probabilidades de cobertura e comprimentos dos intervalos de confiança assintóticos. Além disso, podemos observar que houve forte redução dos comprimentos dos intervalos de confiança com censura informativa em relação aos comprimentos dos intervalos de confiança com censura não-informativa.

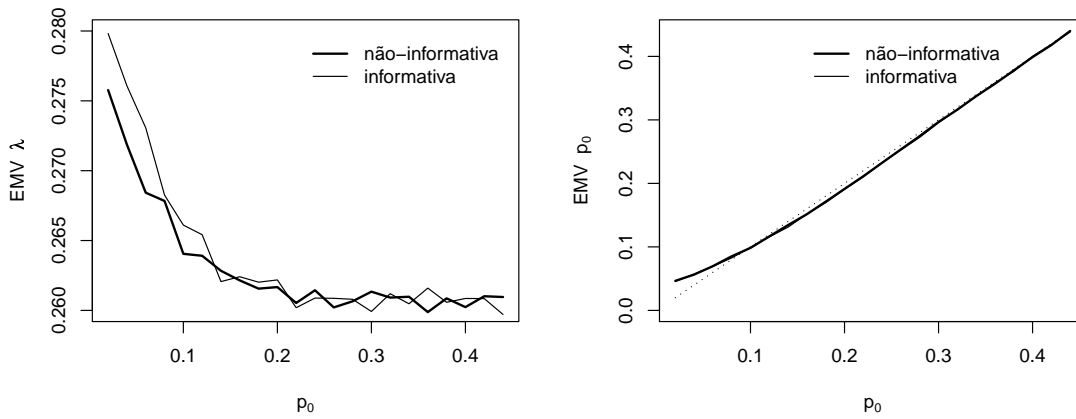


FIGURA 4.3: Estimativas de máxima verossimilhança com $n = 50$, $\lambda = 0,25$, $T_0 = 50$ e $p_c = 0,50$.

Note que as Figuras 4.4, 4.5 e 4.6 a seguir estão construídas para $n = 400$. Com o aumento do tamanho da amostra de $n = 50$ para $n = 400$ temos forte redução dos comprimentos dos intervalos de confiança assintóticos de λ e p_0 na Figura 4.5.

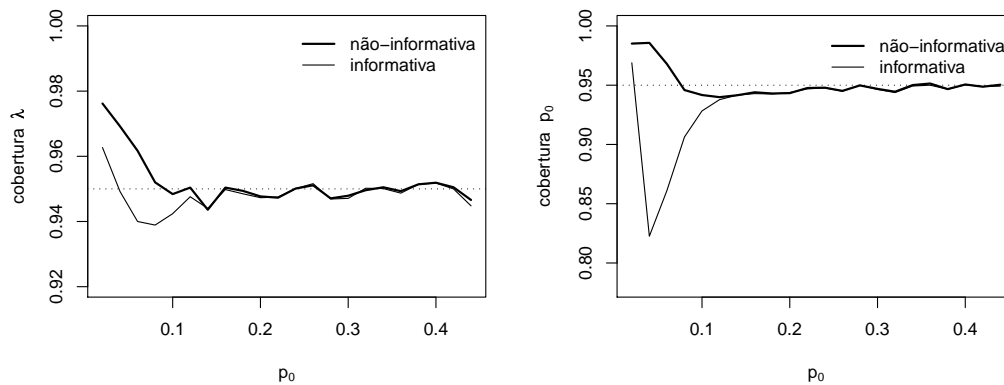


FIGURA 4.4: Probabilidades de cobertura com $n = 400$, $\lambda = 0,25$, $T_0 = 50$ e $p_c = 0,50$.

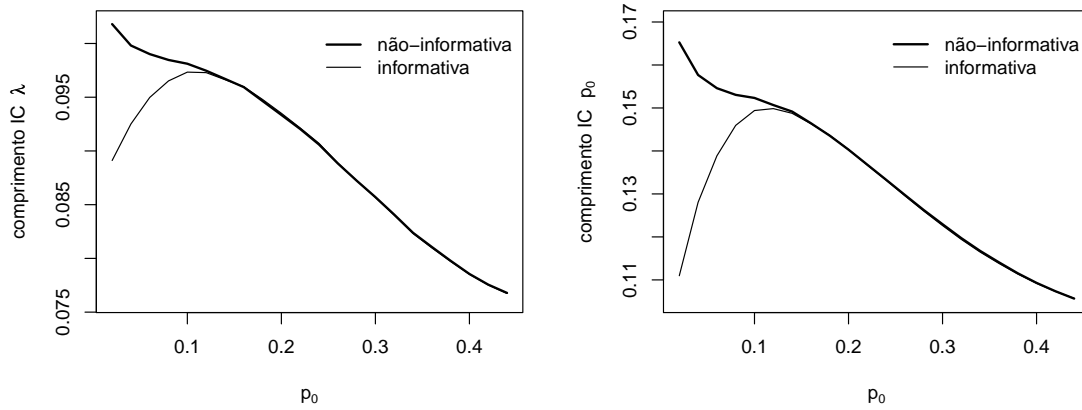


FIGURA 4.5: Comprimento de intervalos de confiança (IC) de 95% com $n = 400$, $\lambda = 0,25$, $T_0 = 50$ e $p_c = 0,50$.

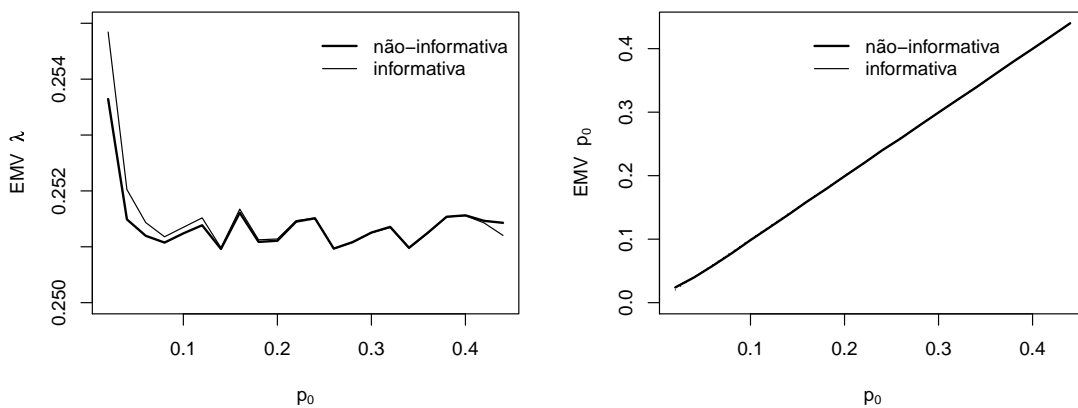


FIGURA 4.6: Estimativas de máxima verossimilhança com $n = 400$, $\lambda = 0,25$, $T_0 = 50$ e $p_c = 0,50$.

De acordo com as Figuras 4.1, 4.2, 4.4 e 4.5, observamos que as estimativas dos comprimentos dos intervalos de confiança assintóticos do parâmetro λ , principalmente na região $0 < p_0 \leq 0.1$, apresentam forte redução quando levamos em conta que a censura é informativa. O vício das estimativas de máxima verossimilhança de λ tende a diminuir quando p_0 se aproxima de p_c , conforme as Figuras 4.3 e 4.6. Quando $p_0 \downarrow 0$, as probabilidades de cobertura sofrem uma forte influência da variável $X \sim UExp(\lambda, T_0)$. Por outro lado, se $p_0 \uparrow p_c$ as probabilidades de cobertura

sufrem uma forte influência da variável $M_i \sim \text{Bernoulli}(1 - p_0)$, $i = 1, \dots, n$.

4.2.2 Simulação com T_0 Não-Fixado

Para analisar o impacto da escolha de T_0 nas EMV com censura informativa uniforme-exponencial, simulamos 10000 réplicas de amostras aleatórias para $n = 50$, $\lambda = 0, 25$, $p_0 = 0, 30$, $p_c = 0, 50$ e $T_0 = 20, 30, \dots, 200$. As Figuras 4.7 e 4.8 revelam um impacto maior nas probabilidades de cobertura e comprimentos de intervalos de confiança 95% de λ e p_0 para $T_0 \leq 50$.

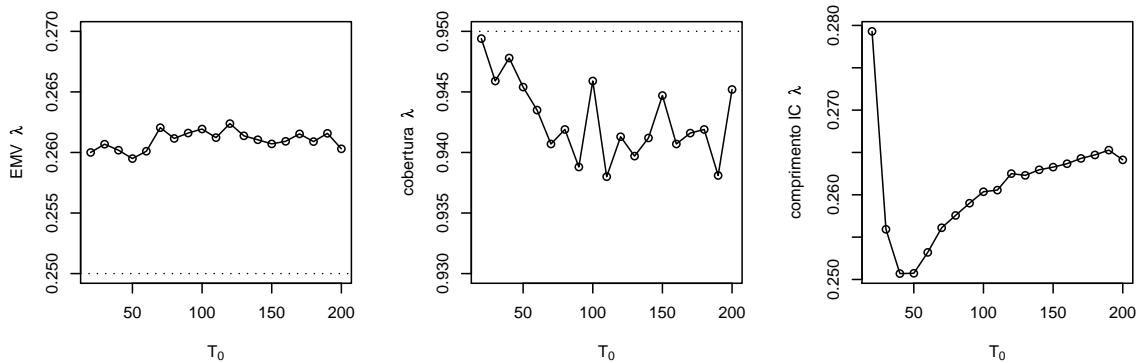


FIGURA 4.7: Estimativas de máxima verossimilhança, probabilidades de cobertura e comprimentos de intervalos de confiança 95% de λ com $n = 50$, $\lambda = 0, 25$, $p_0 = 0, 30$ e $p_c = 0, 50$.

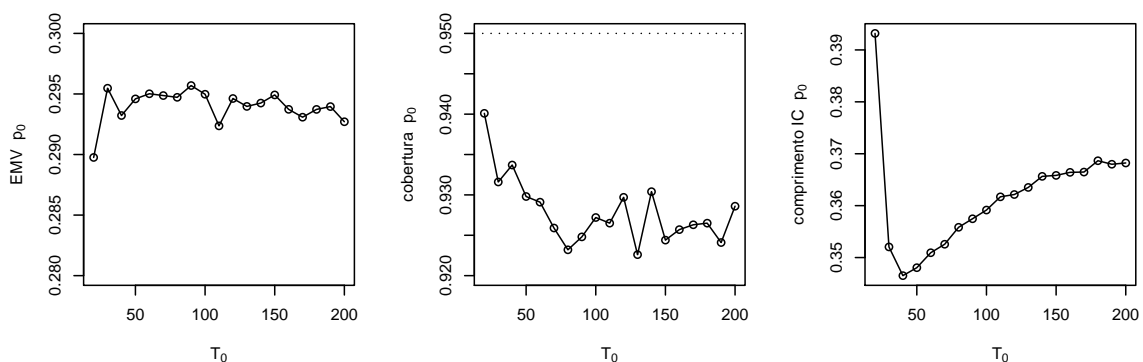


FIGURA 4.8: Estimativas de máxima verossimilhança, probabilidades de cobertura e comprimentos de intervalos de confiança 95% de p_0 com $n = 50$, $\lambda = 0, 25$, $p_0 = 0, 30$ e $p_c = 0, 50$.

4.3 Aplicação com Dados Reais

Nesta seção analisamos dados de pacientes com leucemia descritos em (Kersey *et al.*, 1987). Determinamos estimativas de máxima verossimilhança (EMV) e intervalos de confiança para os parâmetros λ , p_0 , κ e κ^* definidos no Capítulo 3. Estimativas de Kaplan-Meier (EKM) da função de sobrevivência e da função de probabilidade de cura são também calculadas.

As EKM da função de probabilidade de cura $p(t)$ foram obtidas, de acordo com Maller & Zhou (1996), do seguinte modo:

$$\hat{p}(t) = \frac{\hat{p}_n}{\hat{S}_n(t)}, \quad (4.1)$$

em que, \hat{p}_n é dado pelo valor mínimo do EKM de $S_p(t)$ e $\hat{S}_n(t)$ é dado pelo EKM de $S_p(t)$.

Os intervalos de confiança assintóticos estão construídos com censura não-informativa com matriz de informação observada e com censura informativa com matriz de informação de Fisher (3.45). Para construir intervalos de confiança de κ^* utilizamos o método delta que pode ser encontrado em Migon & Gamerman (1999). Para estimar $Var(\hat{\kappa}^*)$, é necessário escrever κ^* dada pela equação (3.12) na forma

$$\kappa^* = \left[\frac{(1 - \kappa)^2}{\lambda T_0} + \kappa \right] \left(1 - e^{-\frac{\lambda T_0}{1 - \kappa}} \right) + \frac{\kappa \lambda^2 e^{-\frac{\lambda T_0}{1 - \kappa}}}{(1 - \kappa)^2}. \quad (4.2)$$

4.3.1 Inferência Clássica com Dados de Leucemia

Exemplo 4.1 Analisamos neste exemplo dados descritos em Kersey *et al.* (1987) relativos ao tempo de recorrência T (em anos) e indicador de censura δ de 46 pacientes do Grupo 1 (transplante alogênico). Este conjunto de dados foi estudado por Maller & Zhou (1996) e ainda por Rodrigues *et al.* (2008). Dentre as medidas resumo da amostra temos que $p_c = 28,3\%$, tempo médio de sobrevivência = 1,52 ano, desvio padrão = 1,63 ano, tempo mínimo = 0,03 ano e tempo máximo = 5,00 anos.

De acordo com as Tabelas 4.1 – 4.4 verificamos que as EMV de λ e da fração de cura (não sofrer recorrência da leucemia após o transplante alogênico) p_0

calculadas com o uso da função de log-verossimilhança com censura não-informativa correspondente à equação (3.17) são aproximadamente iguais às EMV calculadas com o uso da função de log-verossimilhança com censura informativa (3.38), não dependendo de T_0 . Entretanto, as alterações no erro padrão das EMV de λ e p_0 , podem provocar diferenças importantes nos comprimentos dos intervalos de confiança assintóticos de acordo com as Tabelas 4.1 – 4.4. Notamos que o valor de κ^* pode ser considerado pequeno, por isso, os resultados da Tabela 4.1 obtidos com censura não-informativa e matriz de informação observada não podem ser descartados. Situações como esta, em que é possível considerar que o mecanismo de censura seja não-informativo, já estão configuradas no estudo de simulação deste capítulo. Entretanto, é necessário esclarecer de acordo com (3.12) e (3.39) que a relação de dependência entre os parâmetros da distribuição de sobrevivência e distribuição de censura estabelecida por κ^* em (3.9) não permite admitir que o mecanismo de censura informativo seja ignorado.

TABELA 4.1: Inferência clássica com censura não-informativa com os dados do Exemplo 4.1.

Parâmetro	EMV	Erro padrão	Intervalo de confiança de 95%	Comprimento do intervalo
λ	1,4332	0,2756	(0,8931; 1,9733)	1,0802
p_0	0,2711	0,0676	(0,1385; 0,4037)	0,2652

TABELA 4.2: Inferência clássica com censura informativa $UExp(\lambda_c, T_0 = 5, 1)$ com os dados do Exemplo 4.1.

Parâmetro	EMV	Erro padrão	Intervalo de confiança de 95%	Comprimento do intervalo
λ	1,4330	0,3165	(0,8127; 2,0533)	1,2406
p_0	0,2711	0,0778	(0,1185; 0,4236)	0,3051
κ	0,0001	0,0494	(-0,0968; 0,0970)	0,1938
κ^*	0,1368	0,0469	(0,0449; 0,2287)	0,1838

TABELA 4.3: Inferência clássica com censura informativa $UExp(\lambda_c, T_0 = 10)$ com os dados do Exemplo 4.1.

Parâmetro	Erro		Intervalo	Comprimento do intervalo
	EMV	padrão	de confiança de 95%	
λ	1,4332	0,2812	(0,8822; 1,9843)	1,1021
p_0	0,2711	0,0717	(0,1304; 0,4117)	0,2813
κ	0,0185	0,0308	(-0,0418; 0,0788)	0,1206
κ^*	0,0857	0,0310	(0,0250; 0,1465)	0,1215

TABELA 4.4: Inferência clássica com censura informativa $UExp(\lambda_c, T_0 \rightarrow \infty)$ com os dados do Exemplo 4.1.

Parâmetro	Erro		Intervalo	Comprimento do intervalo
	EMV	padrão	de confiança de 95%	
λ	1,4331	0,2901	(0,8646; 2,0017)	1,1371
p_0	0,2711	0,0741	(0,1259; 0,4163)	0,2904
κ	0,1150	0,0325	(0,0512; 0,1788)	0,1276
κ^*	0,1150	0,0325	(0,0512; 0,1788)	0,1276

Os intervalos de confiança de κ com o limite inferior negativo nas Tabelas 4.2 e 4.3 podem ser evitados, como mostra a inferência Bayesiana que desenvolvemos no Capítulo 5.

Sobre a escolha de T_0 fazemos o seguinte comentário, adicional ao que foi feito na Subseção 4.2.2. Numa análise teórica, usando a função de verossimilhança com censura não-informativa, fixando valores para λ e p_0 para investigar a escolha de T_0 Ghitany (1993) concluiu que para um esquema de censura fixada num tempo T_0 , as variâncias de $\hat{\lambda}$ e \hat{p}_0 são minimizadas quando $T_0 \rightarrow \infty$. Entretanto, nos cenários com dados reais como este do Exemplo 4.1 é bastante plausível que se admita T_0 o mais próximo possível, superiormente, do valor máximo amostral, ou fazer $T_0 \rightarrow \infty$ e trabalhar com a censura exponencial, uma vez que o valor real de T_0 não está disponível.

As EMV e EKM da função de sobrevivência $S_p(t)$ e da função de probabilidade de cura $p(t)$ para $T_0 = 10$ estão representadas nas Figuras 4.9 e 4.10, respectivamente. Note que um paciente com 3.5 anos de acompanhamento tem estimativas de probabilidade de cura aproximadamente igual a 1.

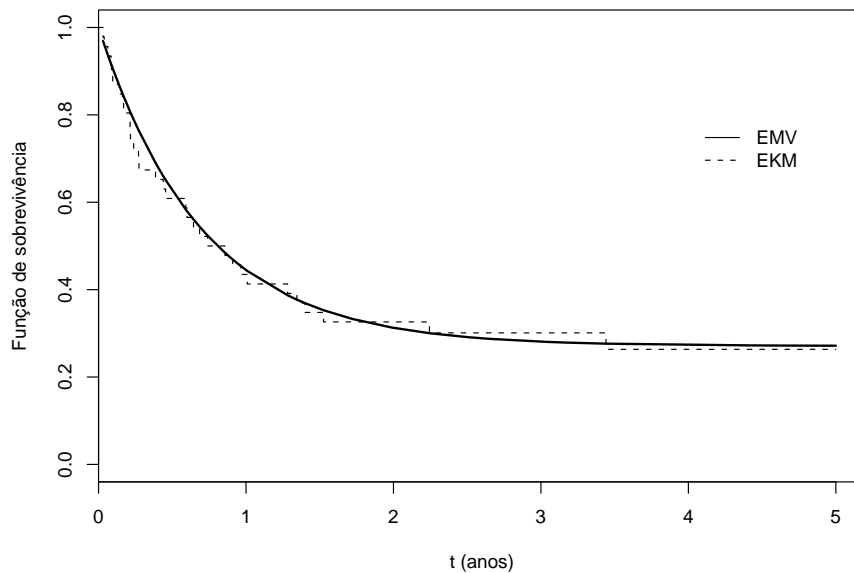


FIGURA 4.9: Estimativas de máxima verossimilhança (EMV) e estimativas de Kaplan-Meier (EKM) da função de sobrevivência $S_p(t)$ com os dados do Exemplo 4.1.

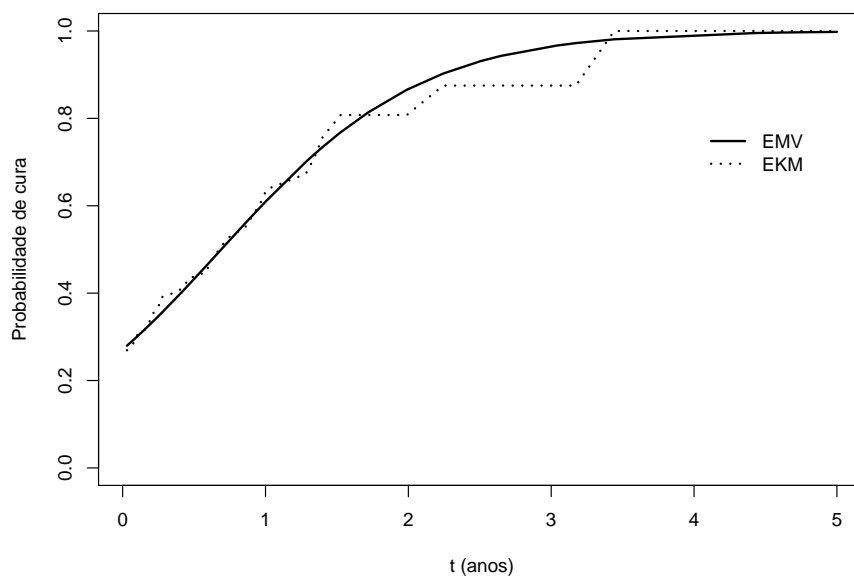


FIGURA 4.10: Estimativas de máxima verossimilhança (EMV) e estimativas de Kaplan-Meier (EKM) da função de probabilidade de cura $p(t)$ com os dados do Exemplo 4.1.

Desenvolvemos neste capítulo um estudo de simulação e aplicações com dados reais sob a inferência clássica. A análise Bayesiana com dados simulados enfrentou algumas dificuldades quanto ao tempo computacional e convergência do algoritmo MCMC e até aqui não foi possível a obtenção de resultados satisfatórios. Para ilustrar a inferência Bayesiana utilizamos os dados do Exemplo 4.1 e apresentamos os resultados no Capítulo 5.

Capítulo 5

Inferência Bayesiana com Priori de Jeffreys e Censura UExp

Neste capítulo propomos uma inferência Bayesiana objetiva (IBO) para avaliar as probabilidades de cobertura e os comprimentos dos intervalos de credibilidade dos parâmetros de interesse. Apresentamos um procedimento Bayesiano com distribuição *a priori* de Jeffreys (Jeffreys, 1961). O desenvolvimento teórico está elaborado sob o mecanismo de censura informativo uniforme-exponencial. O procedimento Bayesiano para o caso de censura não-informativa pode ser deduzido de modo semelhante ao que está descrito nas Seções 5.1 e 5.2. Na seção 5.3 analisamos um conjunto de dados reais.

5.1 Distribuição *a Priori* de Jeffreys

Nesta seção formulamos a distribuição *a priori* objetiva de Jeffreys para os parâmetros do modelo de mistura padrão dado pela equação (3.1). De acordo com a terminologia empregada por Berger (2006) o termo *objetiva* representa um estado em que não se dispõe de informação subjetiva. Também denominado estado em que a informação *a priori* não é significativa em relação à informação amostral ou estado de conhecimento vago ou difuso (Paulino *et al.*, 2003), ou estado em que mesmo havendo informação disponível não se deseja utilizá-la, segundo Migon &

Gamerman (1999). Outros termos empregados são não-informativa, não-subjetiva, referência, *default* e convencional.

Para atingir os propósitos desta seção utilizamos a seguinte definição:

Definição 5.1 A distribuição *a priori* não-informativa de Jeffreys para $\boldsymbol{\theta}$ e p_0 do modelo de mistura padrão e κ , com censura informativa, é dada por

$$\pi(\boldsymbol{\theta}, p_0, \kappa) \propto \{\det(\mathbf{I}^*(\boldsymbol{\theta}, p_0, \kappa))\}^{1/2}, \quad (5.1)$$

em que $\mathbf{I}^*(\boldsymbol{\theta}, p_0, \kappa)$ é a matriz de informação de Fisher.

Consequentemente, escrevemos a distribuição *a priori* de Jeffreys para $\boldsymbol{\theta}$, p_0 e κ , baseada na função de verossimilhança (3.36) com censura informativa uniforme-exponencial, como

$$\pi(\boldsymbol{\theta}, p_0, \kappa) \propto [I_{\kappa\kappa}^*(I_{\lambda\lambda}^* I_{p_0 p_0} - I_{\lambda p_0}^2) - I_{p_0 p_0} (I_{\lambda\kappa}^*)^2]^{1/2}, \quad (5.2)$$

em que $I_{\lambda\lambda}^*$, $I_{\kappa\kappa}^*$, $I_{\lambda\kappa}^*$, $I_{\lambda p_0}$ e $I_{p_0 p_0}$ estão dados em (3.39) – (3.43).

5.2 Densidades *a Posteriori* Condicionais

Reescrevendo a função de verossimilhança de acordo com (3.36), temos que

$$\begin{aligned} L^*(\lambda, p_0, \kappa | \mathcal{D}) &\propto \lambda^{\sum_{i=1}^n \delta_i} e^{-\lambda \sum_{i=1}^n y_i (\delta_i + \frac{\kappa}{1-\kappa})} \prod_{i=1}^n [p_0 + (1-p_0)e^{-\lambda y_i}]^{1-\delta_i} \\ &\times \prod_{i=1}^n \left[\left(\frac{1}{T_0} + \frac{\lambda k}{1-k} \left(1 - \frac{y_i}{T_0}\right) \right) \right]^{1-\delta_i}. \end{aligned} \quad (5.3)$$

De (5.2) e (5.3), a função densidade *a posteriori* conjunta das quantidades λ , p_0 e κ é dada por

$$\begin{aligned} \pi(\lambda, p_0, \kappa | \mathcal{D}) &\propto \lambda^{\sum_{i=1}^n \delta_i} e^{-\lambda \sum_{i=1}^n y_i (\delta_i + \frac{\kappa}{1-\kappa})} \prod_{i=1}^n [p_0 + (1-p_0)e^{-\lambda y_i}]^{1-\delta_i} \\ &\times \prod_{i=1}^n \left[\left(\frac{1}{T_0} + \frac{\lambda k}{1-k} \left(1 - \frac{y_i}{T_0}\right) \right) \right]^{1-\delta_i} \\ &\times [I_{\kappa\kappa}^*(I_{\lambda\lambda}^* I_{p_0 p_0} - I_{\lambda p_0}^2) - I_{p_0 p_0} (I_{\lambda\kappa}^*)^2]^{1/2}, \end{aligned} \quad (5.4)$$

sendo que T_0 é conhecido.

As funções densidade *a posteriori* condicionais dos parâmetros λ e p_0 e κ são dadas, respectivamente por

$$\begin{aligned} \pi(\lambda|p_0, \kappa, \mathcal{D}) &\propto \lambda^{\sum_{i=1}^n \delta_i} e^{-\lambda \sum_{i=1}^n y_i (\delta_i + \frac{\kappa}{1-\kappa})} \prod_{i=1}^n [p_0 + (1-p_0)e^{-\lambda y_i}]^{1-\delta_i} \\ &\times \prod_{i=1}^n \left[\left(\frac{1}{T_0} + \frac{\lambda k}{1-k} \left(1 - \frac{y_i}{T_0}\right) \right) \right]^{1-\delta_i} \\ &\times [I_{\kappa\kappa}^* (I_{\lambda\lambda}^* I_{p_0 p_0} - I_{\lambda p_0}^2) - I_{p_0 p_0} (I_{\lambda\kappa}^*)^2]^{1/2}, \end{aligned} \quad (5.5)$$

$$\begin{aligned} \pi(p_0|\lambda, \kappa, \mathcal{D}) &\propto p_0^{n-\sum_{i=1}^n \delta_i} (1-p_0)^{\sum_{i=1}^n \delta_i} \prod_{i=1}^n \left[\left(1 + \frac{1-p_0}{p_0} e^{-\lambda y_i}\right) \right]^{1-\delta_i} \\ &\times [I_{\kappa\kappa}^* (I_{\lambda\lambda}^* I_{p_0 p_0} - I_{\lambda p_0}^2) - I_{p_0 p_0} (I_{\lambda\kappa}^*)^2]^{1/2} \end{aligned} \quad (5.6)$$

e

$$\begin{aligned} \pi(\kappa|\lambda, p_0, \mathcal{D}) &\propto e^{-\frac{\lambda\kappa}{1-\kappa} \sum_{i=1}^n y_i} \prod_{i=1}^n \left[\left(\frac{1}{T_0} + \frac{\lambda k}{1-k} \left(1 - \frac{y_i}{T_0}\right) \right) \right]^{1-\delta_i} \\ &\times [I_{\kappa\kappa}^* (I_{\lambda\lambda}^* I_{p_0 p_0} - I_{\lambda p_0}^2) - I_{p_0 p_0} (I_{\lambda\kappa}^*)^2]^{1/2}. \end{aligned} \quad (5.7)$$

Podemos obter amostras aleatórias das quantidades λ , p_0 e κ das equações (5.5), (5.6) e (5.7) utilizando o algoritmo Metropolis-Hastings (MH), vide por exemplo, Hastings (1970). Uma apresentação mais recente do algoritmo MH pode ser vista em Chib & Greenberg (1995) ou em Gamerman & Lopes (2006).

5.3 Exemplo com Dados de Leucemia

A estimação Bayesiana das quantidades de interesse, vide programa em R da Seção C.2 do Apêndice C, foi obtida por meio de simulação Monte Carlo em Cadeias de Markov (MCMC) usando o algoritmo MH. Determinamos estimativas pontuais e intervalos de credibilidade com distribuição *a priori* de Jeffreys dos parâmetros λ , p_0 , κ e κ^* .

Exemplo 5.1 – Inferência Bayesiana com censura informativa com os dados do Exemplo 4.1. Desenvolvemos uma inferência Bayesiana com censura informativa uniforme-exponencial e com distribuição *a priori* de Jeffreys utilizando o algoritmo MH a partir de 60000 iterações. As 10000 primeiras iterações foram descartadas para diminuir a influência dos pontos iniciais. Para minimizar o problema de autocorrelação consideramos saltos de tamanho 10 para formar uma amostra final de tamanho 5000 das distribuições *a posteriori* de λ , p_0 , κ e κ^* . As densidades marginais *a posteriori* de λ , p_0 , κ e κ^* foram estimadas com o uso das equações (5.5), (5.6) e (5.7). Verificamos a indicação de convergência das cadeias pelo exame dos gráficos dos históricos das simulações das Figuras 5.2 e 5.4 e pelo método de Geweke (vide por exemplo, Paulino *et al.*, 2003; Gamerman & Lopes, 2006). Utilizamos este método, por meio do pacote CODA da linguagem R. Considerando a partição de 10% dos valores iniciais e a partição de 50% dos valores finais, observamos que os valores $z_G(\lambda) = -0,0034$, $z_G(p_0) = 1,1378$, $z_G(\kappa) = 1,1199$ e $z_G(\kappa^*) = 0,3078$ pertencem ao intervalo $(-1,96; 1,96)$ indicando convergência das cadeias sob a suposição que a estatística Z_G do teste de Geweke segue uma distribuição assintoticamente normal padrão. Outros aspectos computacionais podem ser observados na Seção C.2 do Apêndice C.

A Tabela 5.1 apresenta um resumo das densidades *a posteriori* de λ , p_0 , κ e κ^* . Podemos observar que para $T_0 = 5,1$ houve uma forte redução no comprimento de intervalos de credibilidade em relação aos intervalos de confiança da Tabela 4.2 e que a EMV e a estimativa Bayesiana de κ são bastante diferentes. O intervalo de credibilidade para este parâmetro preserva a suposição de $\kappa \geq 0$, diferentemente do intervalo de confiança assintótico.

TABELA 5.1: Inferência Bayesiana com censura informativa $UExp(\lambda_c, T_0 = 5, 1)$ com os dados do Exemplo 4.1.

Parâmetro	Média	Desvio padrão	Intervalo de credibilidade de 95%	Comprimento do intervalo
λ	1,4406	0,2766	(0,9416; 2,0281)	1,0865
p_0	0,2758	0,0662	(0,1546; 0,4114)	0,2568
κ	0,0130	0,0127	(0,0003; 0,0461)	0,0458
κ^*	0,1503	0,0301	(0,1015; 0,2185)	0,1170

As Figuras 5.1 – 5.4 apresentam as densidades marginais *a posteriori*, os históricos das cadeias e as autocorrelações.

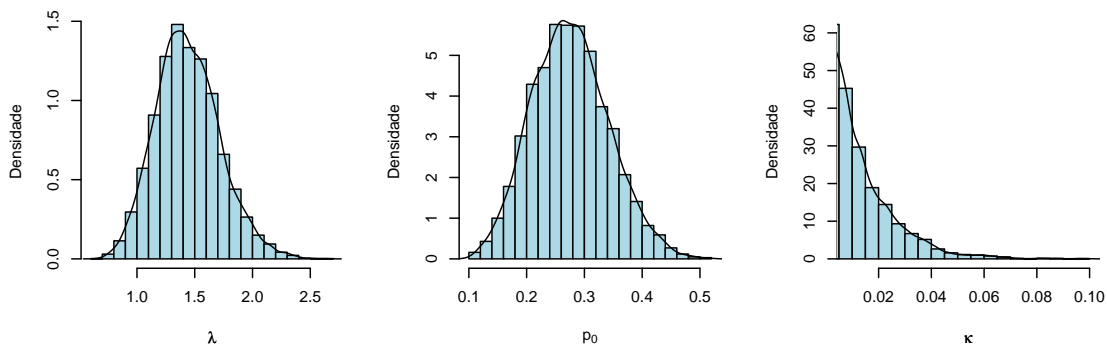
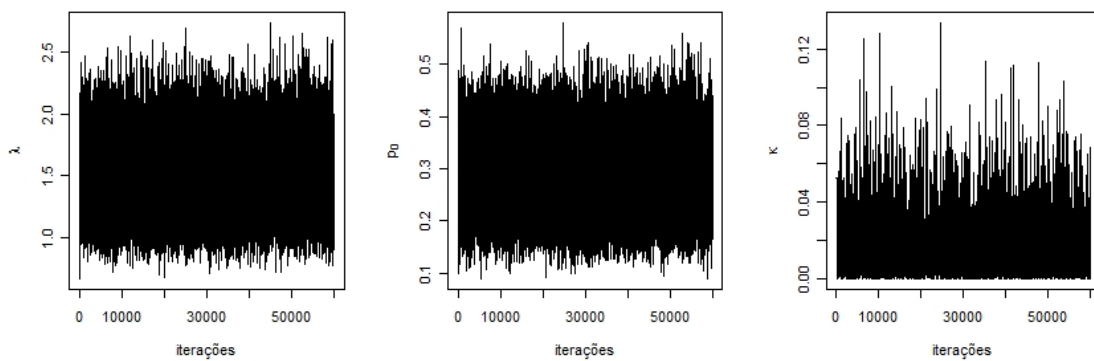
FIGURA 5.1: Densidades marginais *a posteriori* de λ , p_0 e κ – dados do Exemplo 4.1.

FIGURA 5.2: Histórico das simulações – dados do Exemplo 4.1.

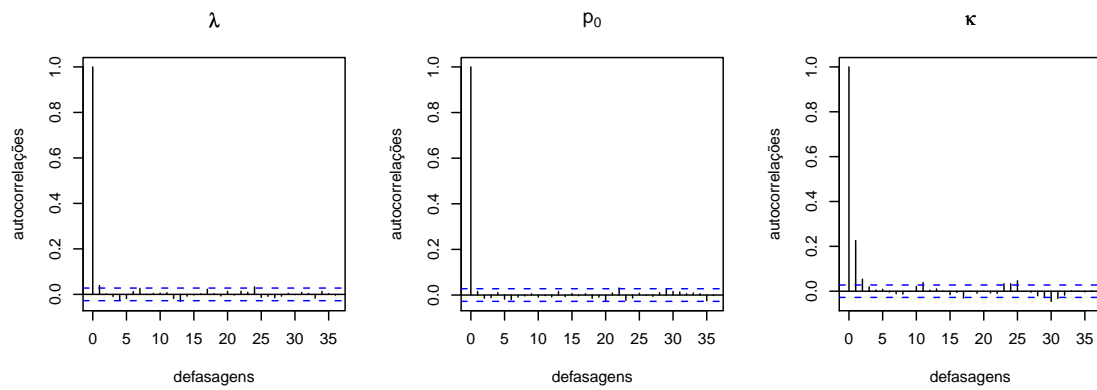
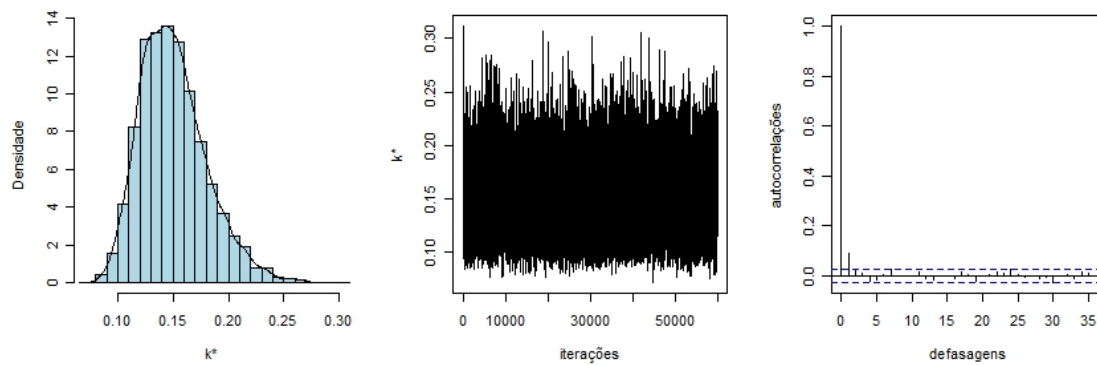


FIGURA 5.3: Autocorrelações – dados do Exemplo 4.1.

FIGURA 5.4: Densidade marginal *a posteriori* de κ^* , histórico das simulações e autocorrelações – dados do Exemplo 4.1.

Capítulo 6

Conclusões e Propostas de Trabalhos Futuros

O desenvolvimento teórico do Capítulo 3 com os resultados do estudo de simulação e aplicação com dados reais do Capítulo 4 são suficientes para concluir que a distribuição de censura não deve ser ignorada na função de verossimilhança para o modelo de mistura padrão. A dependência entre tempo de sobrevivência e mecanismo de censura é objeto de pesquisa de vários autores e acreditamos que este tema continuará sendo amplamente explorado. É importante mencionar que os modelos envolvidos, exponencial, uniforme-exponencial e Bernoulli não foram analisados na presença de covariáveis neste trabalho. Portanto, a inclusão de covariáveis no modelo de mistura padrão com censura informativa faz parte das nossas pesquisas futuras. Utilizando a matriz de informação de Fisher, verificamos que as probabilidades de cobertura e comprimentos dos intervalos de confiança de λ e de p_0 foram fortemente reduzidos com o uso da censura informativa à medida que p_0 se afasta de p_c .

No Capítulo 4 os procedimentos de estimação que desenvolvemos para inferência clássica produziram resultados satisfatórios nas aplicações da metodologia. Entretanto, em relação ao método Bayesiano objetivo do Capítulo 5, com o uso de dados simulados encontramos algumas instabilidades quanto à convergência do algoritmo MCMC no intervalo $0 < p_0 \leq 0,1$ e tempo de processamento. Uma sugestão para tentar reduzir o tempo de processamento com dados simulados em

estudos de probabilidade de cobertura é o uso de um esquema de computação paralela, combinada ou não com o uso da linguagem Ox.

Como principais contribuições deste trabalho, destacamos a relação de dependência entre os parâmetros das distribuições de probabilidade de T e X expressada pela equação (3.9), a determinação analítica da matriz de informação de Fisher com censura não-informativa e com censura informativa uniforme-exponencial, os Teoremas (A.2) – (A.4) do Apêndice A e o Algoritmo 4.1 para simulação de amostras aleatórias do modelo de mistura padrão. Finalmente, mencionamos a evidente redução obtida em estimativas dos comprimentos dos intervalos de confiança assintóticos da taxa de falha e da fração de cura quando não se ignora a censura informativa e os programas computacionais desenvolvidos em R para o ajuste dos modelos.

Como pesquisas futuras enfatizamos os seguintes tópicos:

- Estudar para pequenas e grandes amostras as propriedades frequentistas dos intervalos de credibilidade dos parâmetros do modelo de mistura padrão com distribuição *a priori* de Jeffreys (Berger, 2006).
- Estudar a sensibilidade de covariáveis na probabilidade de cobertura dos intervalos de confiança e de credibilidade da fração de cura e da taxa de falha.
- Estimar as quantidades de interesse utilizando a função de verossimilhança com dados ampliados e implementar os algoritmos MCMC em linguagem Ox.
- Aplicar a metodologia desenvolvida com a distribuição exponencial ponderada introduzida no Apêndice B e verificar o ganho de informação em relação à distribuição exponencial.
- Aplicar a metodologia desenvolvida para dados reais com covariáveis com o uso do pacote GAMLSS (de Castro *et al.*, 2010).
- Utilizar a distribuição $UExp(\lambda, T_0)$, entre outras, para os tempos de sobrevivência e de censura, e investigar soluções analíticas em κ^* para os parâmetros da distribuição de censura.

Apêndice A

Provas de Teoremas do Capítulo 3

Determinamos a matriz de informação de Fisher considerando o seguinte teorema, formulado a partir do Lema 2 de Ghitany *et al.* (1994) e do Lema 2.1 de Zhou & Maller (1995).

Teorema A.1 Suponha que T_1, \dots, T_n sejam tempos de sobrevivência independentes com distribuição acumulada $(1 - p_i)F_i(t_i)$, $p_i \in (0, 1)$ e X_1, \dots, X_n tempos de censura independentes, $Y_i = \min \{T_i, X_i\}$ e T_i e X_i variáveis aleatórias independentes, $i = 1, \dots, n$. Então, para qualquer função mensurável positiva, $Q : \mathbb{R} \rightarrow \mathbb{R}_+$,

$$E(\delta_i Q(Y_i)) = (1 - p_i) E \left\{ \int_{[0, X_i]} Q(y) dF_i(y) \right\} \quad (\text{A.1})$$

e

$$E((1 - \delta_i)Q(Y_i)) = E[\{1 - (1 - p_i)F_i(X_i)\} Q(X_i)]. \quad (\text{A.2})$$

Prova. Considerar o Lema 2 de Ghitany *et al.* (1994) e o Lema 2.1 de Zhou & Maller (1995) para obter (A.1) e (A.2). ■

Desenvolvemos a seguir três resultados que proporcionam melhor desempenho computacional no cálculo da matriz de informação de Fisher com censura não-informativa, pois as derivadas parciais de primeira e segunda ordem da função de log-verossimilhança em relação à θ e p_0 em (3.5) não são requeridas. Estes resultados estão representados pelos seguintes teoremas.

Teorema A.2 Seja $Y_i = \min \{T_i, X_i\}$ sob um modelo de mistura padrão com fração de cura $p_0 \in (0, 1)$, em que o tempo de censura X_i e o tempo de sobrevivência do i -ésimo indivíduo em risco T_i são variáveis aleatórias i.i.d., $i = 1, \dots, n$. Suponha que as funções de distribuições $F(t; \boldsymbol{\theta} | M = 1) = 1 - S(t; \boldsymbol{\theta} | M = 1)$ e G das variáveis aleatórias T_i e X_i não envolvam p_0 e $S_p(Y_i; \boldsymbol{\theta}, p_0)$ denote a função de sobrevivência de longa duração. Então, o elemento $E \{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D}) / \partial p_0^2\}$ da matriz de informação de Fisher é dado por

$$E \left\{ \frac{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D})}{\partial p_0^2} \right\} = nE \left\{ \frac{1 - S(X; \boldsymbol{\theta} | M = 1)}{(1 - p_0)S_p(X; \boldsymbol{\theta}, p_0)} \right\}. \quad (\text{A.3})$$

Prova. A prova será dada utilizando a função de verossimilhança com censura não-informativa (3.5), mas o resultado (A.3) continua válido no caso de função de verossimilhança com censura informativa, definida na Seção 3.2. Considerando a função de verossimilhança (3.5) e a equação (3.1), a função de log-verossimilhança $l(\boldsymbol{\theta}, p_0; \mathcal{D})$ pode ser escrita na forma

$$l(\boldsymbol{\theta}, p_0; \mathcal{D}) \propto \sum_{i=1}^n [\delta_i \log(1 - p_0) + \delta_i \log f(y_i; \boldsymbol{\theta} | M = 1) + (1 - \delta_i) \log S_p(y_i; \boldsymbol{\theta}, p_0)]. \quad (\text{A.4})$$

Então,

$$\frac{\partial^2 l}{\partial p_0^2} = \sum_{i=1}^n \left[\delta_i \frac{-1}{(1 - p_0)^2} - (1 - \delta_i) \frac{(1 - S(y_i; \boldsymbol{\theta} | M = 1))^2}{S_p^2(y_i; \boldsymbol{\theta}, p_0)} \right]. \quad (\text{A.5})$$

Aplicando (A.1) e (A.2) em (A.5), temos que

$$E \left\{ \delta_i \frac{1}{(1 - p_0)^2} \right\} = E \left\{ \frac{1 - S(X; \boldsymbol{\theta} | M = 1)}{1 - p_0} \right\} \quad (\text{A.6})$$

e

$$E \left\{ (1 - \delta_i) \frac{(1 - S(Y_i; \boldsymbol{\theta} | M = 1))^2}{[p_0 + (1 - p_0)S(Y_i; \boldsymbol{\theta} | M = 1)]^2} \right\} = E \left\{ \frac{(1 - S(X; \boldsymbol{\theta} | M = 1))^2}{S_p(X; \boldsymbol{\theta}, p_0)} \right\}. \quad (\text{A.7})$$

Combinando as equações (A.6) e (A.7), concluímos que $E\{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D}) / \partial p_0^2\}$ é dado pela equação (A.3). ■

Teorema A.3 Considere as condições do Teorema A.2. Então, o elemento $E\{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D}) / \partial p_0 \partial \theta_j\}$ da matriz de informação de Fisher é dado por

$$E \left\{ \frac{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D})}{\partial p_0 \partial \theta_j} \right\} = nE \left\{ \frac{\partial S(X; \boldsymbol{\theta} | M = 1) / \partial \theta_j}{S_p(X; \boldsymbol{\theta}, p_0)} \right\}, \quad j = 1, \dots, k. \quad (\text{A.8})$$

Prova. Basta determinar $\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D})/\partial p_0 \partial \theta_j$ em (A.4) e utilizar procedimento semelhante à prova do Teorema A.2 para obter o resultado (A.8). ■

Teorema A.4 Considere as condições do Teorema A.2 e suponha que o mecanismo de censura seja não-informativo para o modelo de mistura padrão com fração de cura p_0 . Então, o elemento $E \{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D})/\partial \theta_j \partial \theta_r\}$ da matriz de informação de Fisher é dado por

$$\begin{aligned}
E \left\{ \frac{-\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D})}{\partial \theta_j \partial \theta_r} \right\} &= n(1-p_0)E \left\{ \int_0^X \frac{[\partial f(y; \boldsymbol{\theta})/\partial \theta_j][\partial f(y; \boldsymbol{\theta})/\partial \theta_r]}{f(y; \boldsymbol{\theta})} dy \right\} \\
&\quad - n(1-p_0)E \left\{ \int_0^X [\partial^2 f(y; \boldsymbol{\theta})/\partial \theta_j \partial \theta_r] dy \right\} \\
&\quad + nE \left\{ \frac{\partial S_p(X; \boldsymbol{\theta}, p_0)}{\partial \theta_j} \frac{\partial S_p(X; \boldsymbol{\theta}, p_0)}{\partial \theta_r} \frac{1}{S_p(X; \boldsymbol{\theta}, p_0)} \right\} \\
&\quad - nE \left\{ \frac{\partial^2 S_p(X; \boldsymbol{\theta}, p_0)}{\partial \theta_j \partial \theta_r} \right\}, \quad j, r = 1, \dots, k,
\end{aligned} \tag{A.9}$$

em que $f(y; \boldsymbol{\theta}) = f(y; \boldsymbol{\theta} | M = 1)$.

Prova. Basta determinar $\partial^2 l(\boldsymbol{\theta}, p_0; \mathcal{D})/\partial \theta_j \partial \theta_r$ em (A.4) e utilizar procedimento semelhante à prova do Teorema 5.2 para obter o resultado (A.9). ■

Apêndice B

Distribuição Exponencial Ponderada

Neste apêndice introduzimos uma distribuição exponencial ponderada que pertence à classe de distribuições ponderadas estudada por Patil *et al.* (1986), Gupta & Kirmani (1990), Gelfand & Dalal (1990), Bayarri & DeGroot (1992), Navarro *et al.* (2001) e Castillo & Pérez-Casany (2005), entre outros. Destacamos o modelo exponencial ponderado como um elemento fundamental, no caso de experimentos com distribuição do tempo de sobrevivência X exponencial, $X \sim Exp(\cdot)$ e outros possíveis modelos relacionados, $X^w \sim f^w(\cdot)$, que podem diminuir o vício dos estimadores dos parâmetros de interesse e melhor explicar o conjunto de dados em estudo. O modelo uniforme-exponencial que utilizamos para o tempo de censura é obtido como caso particular.

B.1 Distribuições Ponderadas

A idealização de distribuições ponderadas foi proposta inicialmente por Fisher (1934). Posteriormente, o conceito foi estabelecido por Rao (1965) e possui aplicações em várias áreas tais como medicina, confiabilidade, meio ambiente e finanças. Em um estudo de doenças crônicas, Zelen & Feinleib (1969) chamaram atenção para o fato de que previsões do tempo de sobrevivência de pacientes que participaram de um programa de prevenção com diagnóstico pré-clínico, quando comparadas às de um grupo controle com diagnóstico no estado clínico, requerem

uma metodologia baseada em amostras viciadas do tipo “length-biased sampling”, que são casos particulares de amostras de distribuições ponderadas. Gupta & Kirmani (1990) utilizaram distribuições ponderadas em um estudo de confiabilidade de produtos e componentes, e várias relações entre as v.a.’s X e X^w são desenvolvidas considerando a função de sobrevivência, a função de risco e o tempo médio de vida residual. Fisher (1934) e Rao (1965) sugeriram flexibilizar o modelo proposto multiplicando a probabilidade (função densidade) de que um particular valor faça parte da amostra por uma função não-negativa tal que os dados sejam uma amostra da distribuição ponderada. A ideia central consiste em transformar o modelo e não os dados. Com efeito, considere a seguinte definição:

Definição B.1 Seja X uma v.a. contínua não-negativa com função densidade $f(x; \theta)$ e seja $w(x; \theta, \phi)$ uma função não-negativa. Então, a v.a. X^w com função densidade

$$f^w(x; \theta, \phi) = \frac{w(x; \theta, \phi)f(x; \theta)}{E(w(X; \theta, \phi))}, \quad 0 < E(w(X; \theta, \phi)) < \infty, \quad (\text{B.1})$$

é chamada de v.a. ponderada correspondente à X .

Note que a função densidade original é multiplicada pela função de peso $w(\cdot)$ e dividida pela constante normalizadora $E(w(\cdot))$. Para uma v.a. discreta X a equação (B.1) continua válida após a substituição das fdp’s pelas correspondentes distribuições de probabilidade $p(x; \theta)$ e $p^w(x; \theta, \phi)$.

Em geral pode ser conhecido somente que $w(\cdot)$ é limitado por duas funções de peso especificadas (Bayarri & Berger, 1998). Uma classe especial deste tipo de modelo surge quando a amostra é selecionada apenas em uma determinada região do espaço amostral associado à v.a. X (Bayarri & Berger, 1998).

B.2 Distribuição Exponencial Ponderada

Como caso particular de distribuição ponderada, motivada por Castillo & Pérez-Casany (2005) e Kokonendji *et al.* (2008), introduzimos a seguinte definição:

Definição B.2 Seja X uma v.a. com fdp $f(x; \theta) = \theta \exp(-\theta x)$, $x > 0$, $\theta > 0$. Então X^w tem distribuição exponencial ponderada se a sua fdp é expressa na forma

$$f^w(x; \theta, \phi) = \frac{w(x; \theta, \phi) \theta \exp(-\theta x)}{E(w(X; \theta, \phi))}, \quad (\text{B.2})$$

em que $w(x; \theta, \phi) = \exp\{rt(x; \theta, \phi)\}$ e r é uma quantidade conhecida.

Considerando a função de peso $w(x; \theta, \phi) = w(x; \phi)$ em (B.1) Patil *et al.* (1986) estabeleceram vários resultados para comparar as v.a.'s X e X^w discretas. Mais especificamente, Castillo & Pérez-Casany (2005) e Kokonendji *et al.* (2008) usando uma v.a. X de Poisson estabeleceram vários resultados para caracterizar sub ou super-dispersão nas correspondentes versões ponderadas X^w . Neste contexto, para comparar as distribuições X e X^w definidas em (B.2) alguns resultados podem ser determinados, por exemplo, analisando a função $t(x; \theta, \phi)$ quanto à convexidade e monotonicidade. Na próxima seção introduzimos três definições e dois teoremas que dão início a uma metodologia para comparar as distribuições X e X^w .

B.3 Comparação Entre X e X^w

Definição B.3 (Croquet, 1960). A v.a. X é mais dispersa do que a v.a. Y , representada por $X \succ Y$, quando para qualquer função convexa g , $E(g(X)) \geq E(g(Y))$.

Definição B.4 (Castillo & Pérez-Casany, 2005). Sejam X e Y variáveis aleatórias. Então, X é estocasticamente maior que Y ($X \succ_{st} Y$) se, e somente se, $P(X > a) \geq P(Y > a)$, $a \in \mathbb{R}$.

Teorema B.1 (Ross, 1983). Sejam X e Y variáveis aleatórias. Então, $X \succ_{st} Y$ se, e somente se, $E(f(X)) \geq E(f(Y))$, para qualquer função $f(\cdot)$ crescente.

Teorema B.2 Seja X^w uma v.a. com distribuição exponencial ponderada de acordo com a equação (B.2) e seja $w(x; \theta, \phi) = w(x; \phi)$. Então, $X^w \succ_{st} X$ ($Y \prec_{st} X$) se $t(x; \phi)$ é monótona crescente (monótona decrescente), com $r > 0$.

Prova. A prova segue diretamente do Teorema 1 de Patil *et al.* (1986). ■

Definição B.5 Uma v.a. X^w com distribuição exponencial ponderada é mais dispersa (ou menos dispersa) do que uma v.a. X com distribuição exponencial com taxa de falha θ se o coeficiente de variação (CV) de X^w for tal que

$$CV(X^w) = \frac{[Var(X^w)]^{1/2}}{E(X^w)} > 1 \quad (CV(X) < 1). \quad (\text{B.3})$$

Resta determinar as restrições que devemos impor sobre $t(x; \theta, \phi)$ e r para que $CV(X^w) < 1$ (> 1) e $Var(X^w) \leq \frac{1}{\theta^2}$ ($\leq \frac{1}{\theta^2}$).

B.4 Exemplos de Distribuições Ponderadas

Exemplo B.1 *Distribuição de equilíbrio.* A distribuição de equilíbrio surge naturalmente em processos de renovação com componente de vida $X \sim F(x)$ (Navarro *et al.*, 2001). A v.a. X^w correspondente tem fdp dada por

$$f^w(x) = \frac{1 - F(x)}{E(X)} = \frac{S(x)}{E(X)}, \quad x > 0. \quad (\text{B.4})$$

Note que, neste caso a função de peso, $w(x) = S(x)/f(x)$ é o inverso da função de risco $h(x) = f(x)/S(x)$. A distribuição de equilíbrio como exponencial ponderada é dada por

$$f^w(x; \theta) = \frac{\exp\{-t(x; \theta)\} \theta e^{-\theta x}}{\theta E(X)}, \quad (\text{B.5})$$

em que $t(x; \theta) = H(x; \theta) - \theta x$ e $H(x; \theta) = \int_x^0 h(u) du = -\log S(x)$ é a função de risco acumulada.

Exemplo B.2 *Distribuição de Weibull.* Considere X^w uma v.a. com distribuição de Weibull cuja fdp é dada por

$$f^w(x; \theta, c) = c\theta x^{c-1} \exp(-\theta x^c) = w(x; \theta, c)\theta \exp(-\theta x), \quad (\text{B.6})$$

em que $c > 0$, $\theta > 0$ e $x > 0$. De (B.2), $w(x; \theta, c) = \exp(t(x; \theta, c))$, $r = 1$ e $t(x; \theta, c) = \log c + (c - 1) \log x - \theta x^c + \theta x$.

Exemplo B.3 *Distribuição uniforme-exponencial.* Considere X^w uma v.a. com distribuição uniforme-exponencial cuja fdp é dada por

$$f^w(x; \lambda, T_0) = \left[\frac{1}{T_0} + \lambda \left(1 - \frac{x}{T_0}\right) \right] \exp(-\lambda x), \quad (\text{B.7})$$

em que $0 < x < T_0$, $\lambda > 0$. Note que, de (B.2), $r = 1$ e $t(x; \lambda, T_0) = \log\left[\frac{1}{\lambda T_0} + \left(1 - \frac{x}{T_0}\right)\right]$.

A Figura B.1 apresenta a razão de variâncias, $Var(X^w)/Var(X)$, para $0 < \lambda < 5$ e $T_0 = 2, 10, 50$.

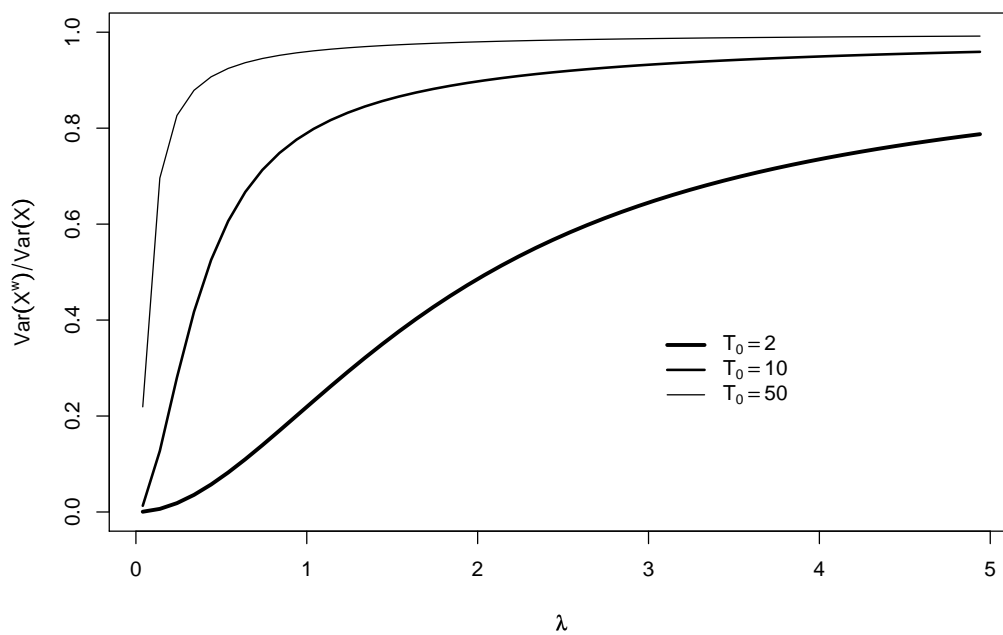


FIGURA B.1: Dispersão da distribuição uniforme-exponencial em relação a exponencial.

Apêndice C

Programas em R para o Modelo de Mistura Padrão

C.1 Programa para Estimação Clássica

Este programa determina estimativas de probabilidades de cobertura, comprimento de intervalos de confiança e EMV de λ e p_0 para o modelo de mistura padrão com censura informativa UExp do estudo de simulação do Capítulo 4.

```
# Verossimilhança com censura informativa UExp
rm(list=ls(all=TRUE))
set.seed(21341)
ptm <- proc.time()
D=10000; n=400
theta<-numeric();alfa=numeric();alpha<-numeric();alph=numeric();ka=numeric()
kaa=numeric(); alphacfinal=numeric()
teta=numeric(); thet=numeric();alph2=numeric();thet2=numeric(); kaa2=numeric()
Pc<-numeric();thetafinal<-numeric()
alphafinal<-numeric();ICalpha1<-numeric();ICalpha2<-numeric()
ICtheta1<-numeric();ICtheta2<-numeric()
Atalpha<-numeric();Attheta<-numeric() ; iteracao<-numeric()
d11<-numeric();d12<-numeric();d22<-numeric(); d13=numeric(); d33=numeric()
Iesperada11=numeric(); Iesperada12=numeric();Iesperada22=numeric()
I11<-numeric();I12<-numeric(); I22<-numeric(); I13=numeric() ;A11<-numeric()
```

```

IF<-numeric();INF11=numeric();INF12=numeric();INF22=numeric()
  V11=numeric();V12=numeric();V22=numeric(); pcure=numeric()
In11<-numeric();In12<-numeric();In13<-numeric()
In21<-numeric();In22<-numeric();In23<-numeric()
In31<-numeric();In32<-numeric();In33<-numeric()
Tn=numeric();Tzero=numeric();Tt=numeric(); ksfinal=numeric()
stat=numeric();p=numeric();af=numeric();ka=numeric();ks=numeric()
kst=numeric();kie=numeric(); ncalphae=numeric();ncalphad=numeric()
ncthetae=numeric();ncthetad=numeric();ncalpha=numeric();nctheta=numeric()
liminfalpha=numeric();limsupalpha=numeric();liminftheta=numeric()
limsuptheta=numeric(); mediaalpha=numeric();mediatheta=numeric()
mediaka=numeric();mediaks=numeric() ; meanPc=numeric(); meaniter=numeric()
y1=numeric();y<-numeric();t<-numeric(); z=numeric();c0=numeric();c=numeric()
alfac=numeric();alfacexp=numeric();alfacest=numeric(); kafinal=numeric()
Pcens=0.50; alfa0=0.25; T0=50 # verificar se T0 > t0min
rz=uniroot(function (t0) (1/(alfa0*t0))*(1-exp(-alfa0*t0))-Pcens ,
lower = 0.001, upper = 10000, tol = 1e-20)
t0min=rz$root
M0=10000*n
ksi=(1/(alfa0*T0))*(1-exp(-alfa0*T0))
p0max=(Pcens-ksi)/(1-ksi)
p0<-seq(0.02,0.99*p0max,0.02) # fração de cura a ser estimada
for (i0 in 1:length(p0)){
ks[i0]=(Pcens-p0[i0])/(1-p0[i0])
fks= uniroot(function (alfac) (1/(alfa0+alfac))*(1-exp(-(alfa0+
alfac)*T0))*((1/T0)*(alfa0/(alfa0+alfac))+ alfac)+
+alfac*(alfa0+alfac)*exp(-(alfa0+alfac)*T0)-ks[i0], lower = 0,
upper = 10000, tol = 1e-20)
alfac[i0]=fks$root
for (k in 1:D){
# geracao de c
c0=(1/T0) + alfac[i0]
j0=0
for (i1 in 1:M0) {
u=runif(1, 0, 1)
z[i1]=rexp(1, rate = alfac[i0])
fz=((1/T0) + alfac[i0]*(1-(z[i1]/T0)))*exp(-alfac[i0]*z[i1])
hz=alfac[i0]*exp(-alfac[i0]*z[i1])

```

```

if ((u <= fz/(c0*hz)) & (z[i1]<=T0)) {c[j0]=z[i1]
j0=j0+1}
if (j0==n+1) break
}
ac=n/(i1-1)
y0<-rep(10^20,n)
y1<-rexp(n,alfa0)
N<-rbinom(n,1,1-p0[i0])
delta <- rep(0,n)
for (i in 1:n) {
  if (N[i]==0) y[i]<-y0[i]
  else y[i]=y1[i] # or else y[i]<-min(y0[i],y1[i])
  t[i]=min(y[i],c[i])
  if (y[i]<=c[i]) delta[i]=1
}
Pc[k]<-(n-sum(delta))/n #porcentagem de censura
epsilon=0.00001 # verificar magnitude dos parâmetros
alpha[1]=1/mean(t); theta[1]=0.5*Pc[k]
ka[1]= 0.5*Pc[k]
M1=100
u1=function(alpha1){
sum(delta*(log(1-theta[1])+log(alpha1) -alpha1*t)+(1-delta)*(log(theta[1]+
(1-theta[1])*exp(-alpha1*t))))+
sum(delta*(log(1-(t/T0)))+(1-delta)*(log((1/T0)+(alpha1*ka[1]/(1-ka[1]))*
(1-(t/T0))))-(alpha1*ka[1]/(1-ka[1]))*t)
}
alph = optimize(u1,c(0,1000000), maximum = T)$maximum
alpha[2]=alph
u2 = function(theta1){
sum(delta*(log(1-theta1)+log(alpha[2]) -alpha[2]*t)+(1-delta)*(log(theta1+
(1-theta1)*exp(-alpha[2]*t))))+
sum(delta*(log(1-(t/T0)))+(1-delta)*(log((1/T0)+(alpha[2]*ka[1]/(1-ka[1]))*
(1-(t/T0))))-(alpha[2]*ka[1]/(1-ka[1]))*t)
}
thet = optimize(u2,c(0,1), maximum = T)$maximum
theta[2]=thet
u3=function(ka1){
sum(delta*(log(1-theta[2])+log(alpha[2]) -alpha[2]*t)+(1-delta)*(log(theta[2]+

```



```

(1-theta[2])*exp(-alpha[2]*t))))+
sum(delta*(log(1-(t/T0)))+(1-delta)*(log((1/T0)+(alpha[2]*ka1/(1-ka1))*
(1-(t/T0))))-(alpha[2]*ka1/(1-ka1))*t)
}
kaa = optimize(u3,c(0,1), maximum = T)$maximum
ka[2] = kaa
iteracao[k]=0
for (j in 3:M1) {
if ((abs(alpha[j-1]-alpha[j-2])>epsilon)|(abs(theta[j-1]-
theta[j-2])>epsilon)|(abs(ka[j-1]-ka[j-2])>epsilon)) {
u1=function(alpha1){
sum(delta*(log(1-theta[j-1])+log(alpha1) -alpha1*t)+(1-delta)*(log(theta[j-1]+
(1-theta[j-1])*exp(-alpha1*t))))+
sum(delta*(log(1-(t/T0)))+(1-delta)*(log((1/T0)+(alpha1*ka[j-1]/(1-ka[j-1]))*
(1-(t/T0))))-(alpha1*ka[j-1]/(1-ka[j-1]))*t)
}
alph2=optimize(u1,c(0,1000000), maximum = T)$maximum
alpha[j]=alph2
u2=function(theta1){
sum(delta*(log(1-theta1)+log(alpha[j]) -alpha[j]*t)+(1-delta)*
(log(theta1+(1-theta1)*exp(-alpha[j]*t))))+
sum(delta*(log(1-(t/T0)))+(1-delta)*(log((1/T0)+(alpha[j]*ka[j-1]/(1-ka[j-1]))*
(1-(t/T0))))-(alpha[j]*ka[j-1]/(1-ka[j-1]))*t)
}
thet2=optimize(u2,c(0,1), maximum = T)$maximum
theta[j]=thet2
u3=function(ka1){
sum(delta*(log(1-theta[j])+log(alpha[j]) -alpha[j]*t)+(1-delta)*(log(theta[j]+
(1-theta[j])*exp(-alpha[j]*t))))+
sum(delta*(log(1-(t/T0)))+(1-delta)*(log((1/T0)+(alpha[j]*ka1/(1-ka1))*
(1-(t/T0))))-(alpha[j]*ka1/(1-ka1))*t)
}
kaa2 = optimize(u3,c(0,1), maximum = T)$maximum
ka[j] = kaa2
}
else break
}
alphafinal[k]=alpha[j-1]

```

```

thetafinal[k]=theta[j-1]
kafinal[k]=ka[j-1]
alphacfina[k]=(alphafinal[k]*kafinal[k]/(1-kafinal[k]))
alphaci=alphacfina[k]
p0i=thetafinal[k]
kai=kafinal[k]
alphai=alphafinal[k]
iteracao[k]=j
ksfinal[k]=(1/(alphafinal[k]+alphaci))*(1-exp(-(alphafinal[k]+alphaci)*T0))*
((1/T0)*(alphafinal[k]/(alphafinal[k]+alphaci))+
alphaci) + alphaci*(alphafinal[k]+alphaci)*exp(-(alphafinal[k]+alphaci)*T0)
# Expected values ##### # # # # # # #
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(1-p0i)*((1/(alphai^2))*(1-exp(-alphai*u)) - p0i*(u^2)*
(exp(-alphai*u))/Sp)*fu +
Sp*(((kai/(1-kai))*(1-u/T0)/((1/T0) + alphaci*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[k]=I11
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
-(u*exp(-alphai*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[k]=I12
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
Su=exp(-alphai*u)
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)
I22=as.numeric(I22[1])

```

```

d22[k]=I22
rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphi*u)
Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*(alphi*kai/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=((1-p0i)/((1-kai)^2))*((1/alphi)*(1-exp(-alphi*u))-u*exp(-alphi*u))
fu(((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[k]=I13
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu(((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
Q1(((1-u/T0)*(-2*alphi/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphi^2)/(((1-kai)^4)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphi*u/((1-kai)^3))
Q2(((1-p0i)*2*alphi/((1-kai)^3))*((1/alphi)*(1-exp(-alphi*u))-u*
exp(-alphi*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[k]=I33
Iesperada=matrix(c(n*d11[k],n*d12[k],n*d13[k],n*d12[k],n*d22[k],0,
n*d13[k],0,n*d33[k]),ncol=3)
Iinv=solve(Iesperada)
In11[k]<-Iinv[1,1]; In12[k]<-Iinv[1,2]; In13[k]<-Iinv[1,3]
In21[k]<-Iinv[2,1]; In22[k]<-Iinv[2,2]; In23[k]<-Iinv[2,3]
In31[k]<-Iinv[3,1]; In32[k]<-Iinv[3,2]; In33[k]<-Iinv[3,3]
# limites e comprimentos de intervalos
ICalpha1[k]=c(alphafinal[k]-qnorm(0.975)*sqrt(Iinv[1,1]))
ICalpha2[k]=c(alphafinal[k]+qnorm(0.975)*sqrt(Iinv[1,1]))
ICtheta1[k]=c(thetafinal[k]-qnorm(0.975)*sqrt(Iinv[2,2]))
ICtheta2[k]=c(thetafinal[k]+qnorm(0.975)*sqrt(Iinv[2,2]))
Atalpha[k]=ICalpha2[k]-ICalpha1[k]

```

```

Attheta[k]=ICtheta2[k]-ICtheta1[k]
}
n0=0;n1=0;k0=0;k1=0
for(i in 1:k){
      if (ICalpha1[i]>alfa0)    n0=n0+1
      if (ICalpha2[i]<alfa0)    n1=n1+1
      if (ICtheta1[i]>p0[i0])   k0=k0+1
      if (ICtheta2[i]<p0[i0])   k1=k1+1
}
ncalphae[i0]=n0/k;ncalphad[i0]=n1/k
ncthetae[i0]=k0/k;ncthetad[i0]=k1/k
ncalphi[i0]=ncalphae[i0]+ncalphad[i0]
nctheti[i0]=ncthetae[i0]+ncthetad[i0]
liminfalpha[i0]=mean(ICalpha1); limsupalpha[i0]=mean(ICalpha2)
liminftheta[i0]=mean(ICtheta1); limsuptheta[i0]=mean(ICtheta2)
mediaalpha[i0]<-mean(alfafinal); mediatheta[i0]<-mean(thetafinal)
mediaka[i0]<-mean(kafinal); mediaks[i0]=mean(ksfinal)
meanPc[i0]=mean(Pc); meaniter[i0]=mean(iteracao)
INF11[i0]=mean(Iesperada11)
INF12[i0]=mean(Iesperada12); INF22[i0]=mean(Iesperada22)
V11[i0]=mean(In11); V12[i0]=mean(In12); V22[i0]=mean(In22)
}
Ialpha=limsupalpha-liminfalpha
Itheta=limsuptheta-liminftheta # amplitude
cobert1=matrix(c(p0,ks,mediaalpha,ncalphae,ncalphad,ncalphi,Ialpha,
mediatheta,ncthetae,ncthetad,nctheti,Itheta),length(p0),12)
IFoIC=matrix(c(meaniter,INF11,INF12,INF22,liminfalpha,limsupalpha,
liminftheta,limsuptheta),length(p0),8)
cobert= matrix(c(p0,ks,mediaalpha,ncalphae,ncalphad,ncalphi,Ialpha,
mediatheta,ncthetae,ncthetad,nctheti,Itheta,INF11,INF12,INF22,
V11,V12,V22,liminfalpha,limsupalpha,liminftheta,limsuptheta),
length(p0),22)
pcure= mediatheta[i0]/(mediatheta[i0]+(1-mediatheta[i0])*exp(-mediaalpha[i0]*t))
timeh=(proc.time() - ptm)/3600
# RESULTADOS
timeh
c(mediaalpha,mediatheta,mediaka,mediaks)
par(mfrow=c(2,4))

```

```

plot(p0, 1-ncalpha, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
abline(h=0.95, lty=3)
plot(p0, 1-nctheta, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
abline(h=0.95, lty=3)
plot(p0, Ialpha, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
plot(p0, Itheta, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
plot(p0, mediaalpha, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
abline(h=alfa0, lty=3)
plot(p0, mediatheta, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
lines(p0, p0, lty=3)
plot(sort(t), sort(pcure), xlab=expression(t), ylab=expression(p(t)),
     type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
plot(1-ncalpha,1- nctheta, type="o", pch=21, bg=par("bg"),
     col = "blue", cex=1)
# END

```

C.2 Programa para Estimação Bayesiana

Este programa determina estimativas Bayesianas de λ , p_0 , κ e κ^* e comprimento de intervalos de credibilidade do modelo de mistura padrão com censura informativa UExp e dados do Exemplo 4.1.

```

# Estimação Bayesiana com censura informativa U-Exp
rm(list=ls(all=TRUE))
ptm=proc.time()
set.seed(21341)
D=1
Pc=numeric();y=numeric();t=numeric()
theta<-numeric();alpha<-numeric();thetafinal<-numeric();ICalpha1<-numeric()
ICalpha2<-numeric();ICtheta1<-numeric();ICtheta2<-numeric();ICalphha<-numeric()

```

```

ICtheta<-numeric();alphaf<-numeric();thetaf<-numeric();af=numeric()
ka=numeric();ks=numeric();kst=numeric();kav=numeric() ; ka1=numeric()
ncalphae=numeric();ncalphad=numeric();ncthetae=numeric();ncthetad=numeric()
ncalpha=numeric();nctheta=numeric();liminfalpha=numeric();limsupalpha=numeric()
liminftheta=numeric();limsuptheta=numeric();Ialpha=numeric();Itheta=numeric()
mediaalpha=numeric();mediatheta=numeric(); meanPc=numeric(); mediaka=numeric()
INF11=numeric();INF12=numeric();INF22=numeric(); V11=numeric();V12=numeric()
V22=numeric();alpha1=numeric();theta1=numeric() ; k11=numeric(); ki=numeric()
iteralpha=numeric();itertheta=numeric();sti1=numeric();sti2=numeric()
r1a=numeric();r2a=numeric();lra=numeric();Ualfga=numeric();Ualfaa=numeric()
I1111<-numeric(); I1<-numeric(); I2<-numeric() ; I3<-numeric(); I4<-numeric()
I5<-numeric(); I6<-numeric(); I7<-numeric(); I8<-numeric();I9<-numeric()
I10<-numeric(); I11<-numeric(); I12<-numeric(); I13<-numeric(); I14<-numeric()
I1a<-numeric(); I2a<-numeric();d111<-numeric() ; Sp=numeric(); fu=numeric()
I121<-numeric(); I11<-numeric();I12<-numeric();I22<-numeric();d11<-numeric()
I221<-numeric();I221<-numeric();d22<-numeric();IF11<-numeric();IF21<-numeric()
IF1<-numeric();I112<-numeric(); d12<-numeric();I33<-numeric();d112<-numeric()
I122<-numeric();I12122<-numeric();d13<-numeric();I222<-numeric()
I2222<-numeric();d222<-numeric(); d33<-numeric();IF12<-numeric()
IF22<-numeric();IF2<-numeric();J1=numeric();J2=numeric();J3=numeric()
J4=numeric(); J5=numeric();J6=numeric();sr1=numeric(); sr2=numeric()
r1t=numeric();r2t=numeric();lrt=numeric() ;r1k=numeric(); r2k=numeric()
lrk=numeric();Ualfat=numeric() ; Iesperada=numeric(); J10=numeric()
alfac=numeric();alfacexp=numeric();alfacest=numeric()
af=numeric();ka=numeric(); kie=numeric();z=numeric();c0=numeric()
c=numeric();alfac=numeric();alfacexp=numeric();alfacest=numeric()
nb=numeric(); taxa1=numeric();taxa2=numeric() ;taxa3=numeric()
alphaest=numeric(); thetaest=numeric() ; cobert=numeric(); kaest=numeric()
delta=numeric(); thet=numeric(); ICka1=numeric();ICka2=numeric(); kse=numeric()
ICka1=numeric(); ICka2=numeric(); ICks1=numeric(); ICks2=numeric()
T0=5.1
for (i0 in 1:1){

for (k in 1:D){

t=c(0.0301, 0.0384, 0.0630, 0.0849, 0.0877, 0.0959, 0.1397, 0.1616, 0.1699,
0.2137, 0.2137, 0.2164, 0.2384, 0.2712, 0.2740, 0.3863, 0.4384, 0.4548, 0.5918,
0.6000, 0.6438, 0.6849, 0.7397, 0.8575, 0.9096, 0.9644, 1.0082, 1.2822, 1.3452,

```

```

1.4000, 1.5260, 1.7205, 1.9890, 2.2438, 2.5068, 2.6466, 3.0384, 3.1726, 3.4411,
4.4219, 4.4356, 4.5863, 4.6904, 4.7808 ,4.9863, 5.0000)
delta=c(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0)
n=length(t)

Pc[k]<-(n-sum(delta))/n #Porcentagem de censura

## ---- Verificar dados simulados -----

M=60000
alpha[1]=1/mean(t) # valor inicial de theta
theta[1]=0.5*Pc[k] # valor inicial de alpha
ks[1]=(Pc[k]-theta[1])/(1-theta[1]) # valor inicial de k*
kse[1]=ks[1]
ka[1]=0.5*ks[1] # valor inicial de kappa
cont1=0; cont2=0; cont3=0
Ualfga[1]=1
Ualfaa[1]=0.01 # Ualfaa inicial deve ser sempre menor que U inicial
J1[1]=.9 ; J2[1]=0.5 ; J3[1]=.9 ; J4[1]=0.5
theta1[1]=0.5
Ualfgt=1
Ualfat[1]=0.1 # Ualfat inicial deve ser sempre menor que G inicial
Ualfga[1]=1
#thet=1
r1a[1]=0.5; r2a[1]=0.2 ; r1t[1]=0.3; r2t[1]=0.2
lra[1]=r1a[1]-r2a[1] ; lrt[1]=r1t[1]-r2t[1]
alpha1[1]=alpha[1];theta1[1]=theta[1] ; ka1[1]=ka[1]

require(survival) ; ekm<- survfit(Surv(t,delta) ~ 1) ; sv=ekm$surv
sv0=ekm$surv
km=sv0[length(sv0)]
nb[1]=0

for (j in 2:M) {
#### GERACÃO DE ALPHA
# NEW (lambda)

```

```

a=sum(delta)+1; S=sum(t*(delta+(ka[j-1]/(1-ka[j-1]))))
#a=sum(delta)+1; S=sum(delta*t)

#alpha1[j]=rexp(1,sum(t)/n)
alpha1[j]=rgamma(1,a,S)
# Jefreys
# Expected values ##### # ## # # ##
p0i=theta[j-1]; alphai=alpha1[j]; alphaci=(alpha1[j]*ka[j-1]/(1-ka[j-1]))
kai=ka[j-1]

# Exata para d11 = dlambd,lambd
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(1-p0i)*((1/(alphai^2))*(1-exp(-alphai*u))- p0i*(u^2)*(exp(-alphai*u))/Sp)*fu +
Sp*(((kai/(1-kai))*(1-u/T0)/((1/T0) + alphaci*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[j]=I11

# Exata para d12 = dlambd,p0
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
-(u*exp(-alphai*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[j]=I12

# Exata para d22 = dpp
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
Su=exp(-alphai*u)
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)

```



```

I22=as.numeric(I22[1])
d22[j]=I22

# Exata para d13 = dlambdak
rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphi*u)
Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*(alphi*kai/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=(((1-p0i)/((1-kai)^2))*((1/alphi)*(1-exp(-alphi*u))-u*exp(-alphi*u)))
fu=(((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[j]=I13

# Exata para d33 = dkk
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu=(((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
Q1=((1-u/T0)*(-2*alphi/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphi^2)/(((1-kai)^4)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphi*u/((1-kai)^3))
Q2=(((1-p0i)*2*alphi/((1-kai)^3))*((1/alphi)*(1-exp(-alphi*u))-
u*exp(-alphi*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[j]=I33

Iesperada=matrix(c(d11[j],d12[j],d13[j],d12[j],d22[j],0,d13[j],0,d33[j]),ncol=3)
J1[j]=sqrt(det(Iesperada))
##

r1a[j]=sum((1-delta)*(log((theta[j-1]+(1-theta[j-1])*exp(-alpha1[j]*t)))+
log((1/T0)+(alpha1[j]*ka[j-1]/(1-ka[j-1]))*(1-t/T0)))) + log(J1[j])

```

```

#
# ##### OLD (lambda)
alphai=alpha[j-1]
##
# Exata para d11 = dlambdai,lambda
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(1-p0i)*((1/(alphai^2))*(1-exp(-alphai*u))- p0i*(u^2)*(exp(-alphai*u))/Sp)*fu +
Sp*(((kai/(1-kai))*(1-u/T0)/((1/T0) + alphaci*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[j]=I11

# Exata para d12 = dlambdai,p0
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
-(u*exp(-alphai*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[j]=I12

# Exata para d22 = dppi
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
Su=exp(-alphai*u)
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)
I22=as.numeric(I22[1])
d22[j]=I22

# Exata para d13 = dlambdai,k
rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphai*u)

```

```

Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*(alphi*kai/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=((((1-p0i)/((1-kai)^2))*((1/alphi)*(1-exp(-alphi*u))-u*exp(-alphi*u)))
fu(((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[j]=I13

# Exata para d33 = dkk
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu(((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
Q1(((1-u/T0)*(-2*alphi/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphi^2)/(((1-kai)^4)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphi*u/((1-kai)^3))
Q2(((1-p0i)*2*alphi/((1-kai)^3))*((1/alphi)*(1-exp(-alphi*u))-
u*exp(-alphi*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[j]=I33

Iesperada=matrix(c(d11[j],d12[j],d13[j],d12[j],d22[j],0,d13[j],0,d33[j]),ncol=3)
J2[j]=sqrt(det(Iesperada))
#
r2a[j]=sum((1-delta)*(log((theta[j-1]+(1-theta[j-1])*exp(-alpha[j-1]*t)))+
log((1/T0)+(alpha[j-1]*ka[j-1]/(1-ka[j-1]))*(1-t/T0)))) + log(J2[j])

lra[j]=r1a[j]-r2a[j]
Ualfaa[j]=min(1,exp(lra[j]))
if (runif(1)<=Ualfaa[j]) {alpha[j]=alpha1[j]
cont1=cont1+1
}
else alpha[j]=alpha[j-1]

```

```

#### Geração de theta

#### NEW (p0)
theta1[j]=rbeta(1,n-sum(delta)+1,sum(delta)+1)
p0i=theta1[j]; alphai=alpha[j]
alphaci=(alpha[j]*ka[j-1]/(1-ka[j-1])); kai=ka[j-1]

##
# Exata para d11 = dlambd,lambd
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(1-p0i)*((1/(alphai^2))*(1-exp(-alphai*u))- p0i*(u^2)*(exp(-alphai*u))/Sp)*fu +
Sp*(((kai/(1-kai))*(1-u/T0)/((1/T0) + alphaci*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[j]=I11

# Exata para d12 = dlambd,p0
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
-(u*exp(-alphai*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[j]=I12

# Exata para d22 = dpp
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
Su=exp(-alphai*u)
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)
I22=as.numeric(I22[1])

```

```

d22[j]=I22

# Exata para d13 = dlambdak
rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphi*u)
Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*(alphi*kai/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=(((1-p0i)/((1-kai)^2))*((1/alphi)*(1-exp(-alphi*u))-u*exp(-alphi*u)))
fu=(((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[j]=I13

# Exata para d33 = dkk
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu=(((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
Q1=((1-u/T0)*(-2*alphi/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphi^2)/(((1-kai)^4)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphi*u/((1-kai)^3))
Q2=(((1-p0i)*2*alphi/((1-kai)^3))*((1/alphi)*(1-exp(-alphi*u))-
u*exp(-alphi*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[j]=I33

Iesperada=matrix(c(d11[j],d12[j],d13[j],d12[j],d22[j],0,d13[j],0,d33[j]),ncol=3)
J3[j]=sqrt(det(Iesperada))
##
r1t[j]=sum((1-delta)*log(1+((1-theta1[j])/theta1[j])*
exp(-alpha[j]*t))) + log(J3[j])
#

```

```

## OLD (p0)
p0i= theta[j-1]
##
# Exata para d11 = dlambd,lambd
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu((((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
(1-p0i)*((1/(alphi^2))*(1-exp(-alphi*u))- p0i*(u^2)*(exp(-alphi*u))/Sp)*fu +
Sp*((kai/(1-kai))*(1-u/T0)/((1/T0) + alphi*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[j]=I11

# Exata para d12 = dlambd,p0
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu((((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
-(u*exp(-alphi*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[j]=I12

# Exata para d22 = dpp
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
Su=exp(-alphi*u)
fu((((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u)))
((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)
I22=as.numeric(I22[1])
d22[j]=I22

# Exata para d13 = dlambd,k
rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphi*u)
Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+

```

```

((1-u/T0)^2)*(alphai*kai/(((1-kai)^3)*((1/T0)+(alphai*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=(((1-p0i)/((1-kai)^2))*((1/alphai)*(1-exp(-alphai*u))-u*exp(-alphai*u)))
fu=(((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[j]=I13

# Exata para d33 = dkk
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu=(((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
Q1=(((1-u/T0)*(-2*alphai/(((1-kai)^3)*((1/T0)+(alphai*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphai^2)/(((1-kai)^4)*((1/T0)+(alphai*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphai*u/((1-kai)^3))
Q2=(((1-p0i)*2*alphai/((1-kai)^3))*((1/alphai)*
(1-exp(-alphai*u))-u*exp(-alphai*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[j]=I33

Iesperada=matrix(c(d11[j],d12[j],d13[j],d12[j],d22[j],0,d13[j],0,d33[j]),ncol=3)
J4[j]=sqrt(det(Iesperada))
##

r2t[j]=sum((1-delta)*log(1+((1-theta[j-1])/theta[j-1])*
exp(-alpha[j]*t))) + log(J4[j]) # 1 original
lrt[j]=r1t[j]-r2t[j]

if (runif(1)<=min(1,exp(lrt[j]))) {theta[j]=theta1[j]
#if ((U2<=Ualfat[j])& (theta1[j]<= Pc[k])) {theta[j]=theta1[j]
cont2=cont2+1
}
else theta[j]=theta[j-1]

```

```

#### GERAÇÃO DE k (kappa)

## NEW (k)
ks[j]=(1/(alpha[j]+alpha[j]*ka[j-1]/(1-ka[j-1])))*(1-exp(-(alpha[j]+alpha[j]*
ka[j-1]/(1-ka[j-1]))*T0))*((1/T0)*(alpha[j]/(alpha[j]+alpha[j]*
ka[j-1]/(1-ka[j-1])))+ alpha[j]*ka[j-1]/(1-ka[j-1])) + alpha[j]*
ka[j-1]/(1-ka[j-1])*(alpha[j]+alpha[j]*ka[j-1]/(1-ka[j-1]))*
exp(-(alpha[j]+alpha[j]*ka[j-1]/(1-ka[j-1]))*T0)

ka1[j]=runif(1, 0, ks[j])
p0i=theta[j]; alphai=alpha[j]; alphaci=(alpha[j]*ka1[j]/(1-ka1[j])); kai=ka1[j]
##
# Exata para d11 = dlambd,lambd
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(1-p0i)*((1/(alphai^2))*(1-exp(-alphai*u))- p0i*(u^2)*(exp(-alphai*u))/Sp)*fu +
Sp*(((kai/(1-kai))*(1-u/T0)/((1/T0) + alphaci*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[j]=I11

# Exata para d12 = dlambd,p0
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
-(u*exp(-alphai*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[j]=I12

# Exata para d22 = dpp
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphai*u))
Su=exp(-alphai*u)
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))

```



```

((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)
I22=as.numeric(I22[1])
d22[j]=I22

# Exata para d13 = dlambdak
rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphi*u)
Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*(alphi*kai/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=((((1-p0i)/((1-kai)^2))*(1/alphi)*(1-exp(-alphi*u))-u*exp(-alphi*u)))
fu=((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[j]=I13

# Exata para d33 = dkk
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu=((1/T0) + alphi*(1-u/T0))*(exp(-(alphi)*u))
Q1=((1-u/T0)*(-2*alphi/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphi^2)/(((1-kai)^4)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphi*u/((1-kai)^3))
Q2=((((1-p0i)*2*alphi/((1-kai)^3))*(1/alphi)*
(1-exp(-alphi*u))-u*exp(-alphi*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[j]=I33

Iesperada=matrix(c(d11[j],d12[j],d13[j],d12[j],d22[j],0,d13[j],0,d33[j]),ncol=3)
J5[j]=sqrt(det(Iesperada))
##

```

```

r1k[j]= -log(ks[j]) -(alpha[j]*ka1[j]/(1-ka1[j]))*sum(t) + sum((1-delta)*
log((1/T0)+(alpha[j]*ka1[j]/(1-ka1[j]))*(1-t/T0))) + log(J5[j])
#
## OLD (k)

alphaci=(alpha[j]*ka[j-1]/(1-ka[j-1])) ;kai= ka[j-1]
##
rmc1<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(1-p0i)*((1/(alphi^2))*(1-exp(-alphi*u))- p0i*(u^2)*(exp(-alphi*u))/Sp)*fu +
Sp*(((kai/(1-kai))*(1-u/T0)/((1/T0) + alphaci*(1-u/T0)))^2)*fu
}
I11<-integrate(rmc1, lower = 0, upper = T0)
I11=as.numeric(I11[1])
d11[j]=I11

# Exata para d12 = dlambdap0
rmc2<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
-(u*exp(-alphi*u)/Sp)*fu
}
I12<-integrate(rmc2, lower = 0, upper = T0)
I12=as.numeric(I12[1])
d12[j]=I12

# Exata para d22 = dpp
rmc3<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
Su=exp(-alphi*u)
fu((((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
((1-Su)/((1-p0i)*Sp))*fu
}
I22<-integrate(rmc3, lower = 0, upper = T0)
I22=as.numeric(I22[1])
d22[j]=I22

# Exata para d13 = dlambdak

```

```

rmc4<-function(u){
Sp=p0i+(1-p0i)*exp(-alphi*u)
Q3=((1-u/T0)*(-1/(((1-kai)^2)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*(alphi*kai/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + u/((1-kai)^2))
Q4=(((1-p0i)/((1-kai)^2))*((1/alphi)*(1-exp(-alphi*u))-u*exp(-alphi*u)))
fu=(((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
(Sp*Q3 + Q4)*fu
}
I13<-integrate(rmc4, lower = 0, upper = T0)
I13=as.numeric(I13[1])
d13[j]=I13

# Exata para d33 = dkk
rmc5<-function(u){
Sp=(p0i+(1-p0i)*exp(-alphi*u))
fu=(((1/T0) + alphaci*(1-u/T0))*(exp(-(alphaci)*u)))
Q1=((1-u/T0)*(-2*alphi/(((1-kai)^3)*((1/T0)+(alphi*kai/(1-kai))*(1-u/T0))))+
((1-u/T0)^2)*((alphi^2)/(((1-kai)^4)*((1/T0)+(alphi*kai/(1-kai))*
(1-u/T0))^2))) + 2*alphi*u/((1-kai)^3))
Q2=(((1-p0i)*2*alphi/((1-kai)^3))*((1/alphi)*
(1-exp(-alphi*u))-u*exp(-alphi*u)))
(Sp*Q1 + Q2)*fu
}
I33<-integrate(rmc5, lower = 0, upper = T0)
I33=as.numeric(I33[1])
d33[j]=I33

Iesperada=matrix(c(d11[j],d12[j],d13[j],d12[j],d22[j],0,d13[j],0,d33[j]),ncol=3)
J6[j]=sqrt(det(Iesperada))
##

r2k[j]= -log(ks[j-1]) -(alpha[j]*ka[j-1]/(1-ka[j-1]))*sum(t) + sum((1-delta)*
log((1/T0)+(alpha[j]*ka[j-1]/(1-ka[j-1]))*(1-t/T0))) + log(J6[j])
lrk[j]=r1k[j]-r2k[j]

if (runif(1)<=min(1,exp(lrk[j]))) {ka[j]=ka1[j]
cont3=cont3+1

```

```

}
else ka[j]=ka[j-1]
kse[j]=(1/(alpha[j]+alpha[j]*ka[j]/(1-ka[j])))*(1-exp(-(alpha[j]+alpha[j]*
ka[j]/(1-ka[j]))*T0))*((1/T0)*(alpha[j]/(alpha[j]+alpha[j]*ka[j]/(1-ka[j])))+
alpha[j]*ka[j]/(1-ka[j])) + alpha[j]*ka[j]/(1-ka[j])*(alpha[j]+alpha[j]*
ka[j]/(1-ka[j]))*exp(-(alpha[j]+alpha[j]*ka[j]/(1-ka[j]))*T0)
# kse = k* estimado

} # chave M

d0=10000
salto=10
alphaf1=alpha[d0+1:M]
alphaf=alphaf1[seq((1), M-d0,salto)]
thetaf1=theta[d0+1:M]
thetaf=thetaf1[seq((1), M-d0,salto)]
kaf1=ka[d0+1:M]
kaf=kaf1[seq((1),M-d0,salto)]
ksf1=kse[d0+1:M]
ksf=ksf1[seq((1),M-d0,salto)]

ICalpha1[k]=quantile(alphaf,c(0.025))
ICalpha2[k]=quantile(alphaf,c(0.975))
ICtheta1[k]=quantile(thetaf,c(0.025))
ICtheta2[k]=quantile(thetaf,c(0.975))
ICka1[k]=quantile(kaf,c(0.025))
ICka2[k]=quantile(kaf,c(0.975))
ICks1[k]=quantile(ksf,c(0.025))
ICks2[k]=quantile(ksf,c(0.975))

taxa1[k]=cont1/M ; taxa2[k]=cont2/M; taxa3[k]=cont3/M
alphaest[k]=mean(alphaf); thetaest[k]=mean(thetaf); kaest[k]=mean(kaf)

} # chave do D

} # chave do i0

```

```
# Amostras
P1=matrix(c(alphaf,thetaf, kaf, ksf),length(alphaf),4)
write.table(P1, file = "kersey_bayes_P1.txt", quote = TRUE,
  sep = " ", dec = ".", row.names = FALSE, col.names = TRUE)
P2=matrix(c(alpha,theta, ka, ks),length(alpha),4)
write.table(P2, file = "kersey_bayes_P2.txt", quote = TRUE,
  sep = " ", dec = ".", row.names = FALSE, col.names = TRUE)

library(coda)
s1c=mcmc(alphaf)
s2c=mcmc(thetaf)
s3c=mcmc(kaf)
s4c=mcmc(ksf)
c(geweke.diag(s1c),geweke.diag(s2c), geweke.diag(s3c), geweke.diag(s4c))

timeh=(proc.time() - ptm)/3600
# Resultados
#dados iniciais
timeh
D; n; T0; i0
c(taxa1, taxa2, taxa3, taxa3)

#####
# END
par(mfrow=c(1,3))
hist(alphaf,col="light blue",probability = TRUE,nclass=20,
xlab=expression(lambda),ylab="Frequencia",main="")
hist(thetaf,col="light blue",probability = TRUE,nclass=20,
xlab= expression(p[0]),ylab="Frequencia",main="")
hist(kaf,col="light blue",probability = TRUE,nclass=20,
xlab=expression(kappa),ylab="Frequencia",main="")

par(mfrow=c(1,3))
plot(alpha, type='l', xlab='iterações', ylab=expression(lambda))
plot(theta, type='l', xlab='iterações', ylab= expression(p[0]))
plot(ka, type='l', xlab='iterações', ylab=expression(kappa))

par(mfrow=c(1,3))
```

```
acf (alphaf, main='', xlab='defasagens', ylab='autocorrelações')
acf (thetaf, main='', xlab='defasagens', ylab='autocorrelações')
acf (kaf, main='', xlab='defasagens', ylab='autocorrelações')

par(mfrow=c(1,3))
hist(ksf,col="light blue",probability = TRUE,nclass=20,xlab='k*',
ylab="Frequencia",main="")
plot(ks, type='l', xlab='iterações', ylab='k*')
acf (ksf, main='', xlab='defasagens', ylab='autocorrelações')

c(taxa1, taxa2, taxa3, taxa3)
c(cont1,cont2,cont3, cont3)

Ictheta=c(ICtheta1[k],ICtheta2[k])
c(mean(alphaf),sd(alphaf),ICalpha1[k],ICalpha2[k],ICalpha2[k]-ICalpha1[k])
c(mean(thetaf),sd(thetaf),ICtheta1[k],ICtheta2[k],ICtheta2[k]-ICtheta1[k])
c(mean(kaf),sd(kaf),ICka1[k],ICka2[k],ICka2[k]-ICka1[k])
c(mean(ksf),sd(ksf),ICks1,ICks2,ICks2-ICks1)
# END
```

Referências

- Abramowitz, M. & Stegun, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, D.C.
- Bayarri, M. J. & Berger, J. (1998). Robust Bayesian analysis of selection models. *The Annals of Statistics*, **26**, 645–659.
- Bayarri, M. J. & DeGroot, M. (1992). A BAD view of weighted distributions and selection models. *Proceedings of the Fourth Valencia International Meeting – Bayesian Statistics 4*. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds. Oxford University Press, 17–33.
- Berger, J. (2006). The case for objective Bayesian analysis. *Bayesian Analysis*, **1**, 385–402.
- Berkson, J. & Gage, R. (1952). Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, **47**, 501–515.
- Boag, J. W. (1949). Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society B*, **11**, 15–53.
- Brown, L. D., Cai, T. T. & DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, **16**, 101–133.
- Castillo, J. & Pérez-Casany, M. (2005). Overdispersed and underdispersed Poisson generalizations. *Journal of Statistical Planning and Inference*, **134**, 486–500.
- Chen, M.-H., Ibrahim, J. G. & Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**, 909–919.
- Chen, M.-H., Harrington, D. P. & Ibrahim, J. G. (2002). Bayesian cure rate model for malignant melanoma: a case-study of Eastern Cooperative Oncology Group trial E1690. *Applied Statistics*, **51**, 135–150.
- Chen, M.-H., Ibrahim, J. G. & Sinha, D. (2004). A new joint model for longitudinal and survival data with a cure fraction. *Journal of Multivariate Analysis*, **91**, 18–34.
- Chib, S. & Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *The American Statistician*, **49**, 327–335.

- Colosimo, E. A. & Giolo, S. R. (2006). *Análise de Sobrevivência Aplicada*. Edgard Blücher, São Paulo.
- Cox, D. R. (1972). Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society B*, **34**, 187–220.
- Croquet, G. (1960). Le thorme representation integrales dans les cones convexes. *Annales de l' Institut Fourier (Grenoble)*, **10**, 333–334.
- Cvijović, D. & Klinowski, J. (1997). Continued-fraction expansions for the Riemann zeta function and polylogarithms. *Proceedings of the American Mathematical Society*, **125**, 2543–2550.
- de Castro, M., Cancho, V. G. & Rodrigues, J. (2009). A Bayesian long-term survival model parametrized in the cured fraction. *Biometrical Journal*, **51**, 443–445.
- de Castro, M., Cancho, V. G. & Rodrigues, J. (2010). A hands-on approach for fitting long-term survival models under the GAMLSS framework. *Computer Methods and Programs in Biomedicine*, **97**, 168–177.
- Doornik, J. A. (2006). *Ox: An Object-Oriented Matrix Language*. Timberlake Consultants Press, London, fourth edition.
- Dunson, D. B. & Dinse, G. E. (2002). Bayesian models for multivariate current status data with informative censoring. *Biometrics*, **58**, 79–88.
- Ebrahimi, N., Molefe, D. & Ying, Z. (2003). Identifiability and censored data. *Biometrika*, **90**, 724–727.
- Emoto, S. E. & Matthews, P. C. (1990). A Weibull model for dependent censoring. *The Annals of Statistics*, **18**, 1556–1577.
- Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, **38**, 1041–1046.
- Farewell, V. T. (1986). Mixture models in survival analysis: are they worth the risk? *The Canadian Journal of Statistics*, **14**, 257–262.
- Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, vol. 1. Wiley, New York, third edition.
- Fisher, R. A. (1934). The effect of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, **6**, 13–25.
- Gamerman, D. & Lopes, H. F. (2006). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman & Hall/CRC, London, second edition.
- Gelfand, A. E. & Dalal, S. R. (1990). A note on overdispersed exponential families. *Biometrika*, **77**, 55–64.

- Ghitany, M. E. (1993). On the information matrix of exponential mixture models with long-term survivors. *Biometrical Journal*, **35**, 15–27.
- Ghitany, M. E., Maller, R. A. & Zhou, S. (1994). Exponential mixture models with long-term survivors and covariates. *Journal of Multivariate Analysis*, **49**, 218–241.
- Gordon, N. H. (1990). Application of the theory of finite mixtures for the estimation of ‘cure’ rates of treated cancer patients. *Statistics in Medicine*, **9**, 397–407.
- Gradshteyn, I. S. & Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products*. Edited by A. Jeffrey and D. Zwillinger. Academic Press, New York, 7th edition.
- Gupta, R. C. & Kirmani, S. N. U. A. (1990). The role of weighted distribution in stochastic modeling. *Communications in Statistics – Theory and Methods*, **19**, 3147–3162.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**, 97–109.
- Ibrahim, J. G., Chen, M.-H. & Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York.
- Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press, London, third edition.
- Kalbfleisch, J. D. & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. Wiley, New York, second edition.
- Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, **53**, 457–481.
- Kersey, J. H., Weisdorf, D., Nesbit, M. E., Lebien, T. W., Woods, W. G., Mcglave, P. B., Kim, T., Vallera, D. A., Goldman, A. I., Bostrom, B., Hurd, D. & Ramsay, N. K. C. (1987). Comparison of autologous and allogeneic bone marrow transplantation for treatment of high-risk refractory acute lymphoblastic leukemia. *New England Journal of Medicine*, **317**, 461–467.
- Kokonendji, C. C., Mizère, D. & Balakrishnan, N. (2008). Connections of the Poisson weight function to overdispersion and underdispersion. *Journal of Statistical Planning and Inference*, **138**, 1287–1296.
- Lagakos, S. W. (1979). General right censoring and its impact on the analysis of survival data. *Biometrics*, **35**, 139–156.
- Lagakos, S. W. & Williams, J. S. (1978). Models for censored survival analysis: A cone class of variable-sum models. *Biometrika*, **65**, 181–189.
- Lambert, P. C., Dickman, P. W. & Weston, C. L. (2010). Estimating the cure fraction in population-based cancer studies by using finite mixture models. *Journal of the Royal Statistical Society C (Applied Statistics)*, **59**, 35–55.

- Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. Wiley, New York.
- Li, C.-S., Taylor, J. M. G. & Sy, J. P. (2001). Identifiability of cure models. *Statistics and Probability Letters*, **54**, 389–395.
- Link, W. A. (1989). A model for informative censoring. *Journal of the American Statistical Association*, **84**, 749–752.
- Maller, R. A. & Zhou, S. (1996). *Survival Analysis with Long-Term Survivors*. Wiley, New York.
- McLachlan, G. & Peel, D. (2000). *Finite Mixture Models*. Wiley, New York.
- Migon, H. S. & Gamerman, D. (1999). *Statistical Inference: An Integrated Approach*. Arnold, London.
- Mizoi, M. F. (2004). *Influência local em modelos de sobrevivência com fração de cura*. Tese de doutorado, IME – Universidade de São Paulo, São Paulo, Brasil.
- Navarro, J., Ruiz, J. M. & Del Aguila, Y. (2001). Parametric estimation from weighted samples. *Biometrical Journal*, **43**, 297–311.
- Nocedal, J. & Wright, S. J. (1999). *Numerical Optimization*. Springer, New York.
- Oller, R., Gómez, G. & Calle, M. L. (2007). Interval censoring: identifiability and the constant-sum property. *Biometrika*, **94**, 61–70.
- Patil, G. P., Rao, C. R. & Ratnaparkhi, M. V. (1986). On discrete weighted distributions and their use in model choice for observed data. *Communications in Statistics – Theory and Methods*, **15**, 907–918.
- Patterson, B. H. & Smith, P. J. (1985). Asymptotic and finite sample behavior of the time on test estimator under random censorship when lifetimes are not exponential. *Communications in Statistics – Theory and Methods*, **14**, 1643–1658.
- Paulino, C. D., Turkman, M. A. A. & Murteira, B. (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.
- Peng, Y. & Zhang, J. (2008). Identifiability of a mixture cure frailty model. *Statistics and Probability Letters*, **78**, 2604–2608.
- Piantadosi, S. & Crowley, J. (1995). An implicitly defined parametric model for censored survival data and covariates. *Biometrics*, **51**, 249–258.
- R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment. *Pergamon Press and Statistical Publishing Society*. G. P. Patil, ed., 320–332.

- Rodrigues, J., Cancho, V. G. & de Castro, M. (2008). *Unified Theory of Survival Analysis*. Brazilian Statistical Association, São Paulo. (in Portuguese).
- Rodrigues, J., Cancho, V. G., de Castro, M. & Louzada-Neto, F. (2009). On the unification of the long-term survival models. *Statistics and Probability Letters*, **79**, 753–759.
- Ross, S. M. (1983). *Stochastic Processes*. Wiley, New York.
- Rubinstein, R. Y. & Kroese, D. P. (2008). *Simulation and the Monte Carlo Method*. Wiley, Hoboken, second edition.
- Siannis, F. (2004). Applications of a parametric model for informative censoring. *Biometrics*, **60**, 704–714.
- Siannis, F., Copas, J. & Lu, G. (2005). Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics*, **6**, 77–91.
- Sy, J. P. & Taylor, J. M. G. (2000). Estimation in a Cox proportional hazards cure model. *Biometrics*, **56**, 227–236.
- Tanner, M. A. (1996). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*. Springer, New York.
- Titterton, D. M., Smith, A. F. M. & Makov, U. E. (1985). *Statistical Analysis of Finite Mixture Distributions*. Wiley, New York.
- Tsiatis, A. (1975). A nonidentifiability aspect of the problem of competing risks. *Proceedings of the National Academy of Sciences of the USA*, **72**, 20–22.
- Tsodikov, A. D., Ibrahim, J. G. & Yakovlev, A. Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. *Journal of the American Statistical Association*, **98**, 1063–1078.
- Williams, J. S. & Lagakos, S. W. (1977). Models for censored survival analysis: constant-sum and variable-sum models. *Biometrika*, **64**, 215–224.
- Yakovlev, A. Y. & Tsodikov, A. D. (1996). *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, Singapore.
- Yakowitz, S. J. & Spragins, J. D. (1968). On the identifiability of finite mixtures. *Annals of Mathematical Statistics*, **39**, 209–214.
- Yau, K. K. W. & Ng, A. S. K. (2001). Long-term survivor mixture model with random effects: application to a multi-centre clinical trial of carcinoma. *Statistics in Medicine*, **20**, 1591–1607.
- Zelen, M. & Feinleib, M. (1969). On the theory of screening for chronic diseases. *Biometrika*, **56**, 601–614.
- Zhou, S. & Maller, R. A. (1995). The likelihood ratio test for the presence of immunes in a censored sample. *Statistics*, **27**, 181–201.