

MODELOS PARA DADOS DE SOBREVIVÊNCIA MULTIVARIADOS COM CENSURA INFORMATIVA

Silvana Schneider

Departamento de Estatística - ICEX - UFMG

Julho de 2017

MODELOS PARA DADOS DE SOBREVIVÊNCIA MULTIVARIADOS COM CENSURA INFORMATIVA

Silvana Schneider

Orientador: Prof. Dr. Fábio Nogueira Demarqui
Coorientador: Prof. Dr. Enrico Antônio Colosimo

Tese submetida ao Programa de Pós-graduação em Estatística da Universidade Federal de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Doutor em Estatística.

Departamento de Estatística
Instituto de Ciências Exatas
Universidade Federal de Minas Gerais

Belo Horizonte, MG - Brasil
Julho de 2017

Aos meus pais, Irene e Orlando.

“... é caminhando que se faz o caminho...”

Titãs

Agradecimentos

A Deus, obrigada Pai amado, pela minha vida, por todas oportunidades que tem me proporcionado, pelas pessoas que tem colocado em meu caminho, por demonstrar o Seu amor e cuidado todos os dias.

Aos meus pais, Orlando e Irene Müller Schneider, pelo amor incondicional, por não medir esforços para nos ajudar e pelos exemplos. À minha irmã Roberta, por estar sempre presente, se preocupando, incentivando e dando força.

Ao professor Fábio Demarqui, por TODA orientação, pelo conhecimento transmitido, pela paciência, dedicação e preocupação, pelos conselhos profissionais e pelas instruções sobre concursos. Ao professor Enrico Colosimo, pela co-orientação deste trabalho e experiência transmitida.

Aos membros da banca examinadora, professores Mário de Castro, Leonardo S. Bastos, Wagner B. de Souza, Lourdes C. Montenegro e ao Cristiano C. Santos pelas contribuições, correções e sugestões. Ao professor Vinícius D. Mayrink, pela colaboração.

À Michele Morales, amiga de todas as horas, que mesmo com a distância se fez presente em todos os momentos, principalmente nos mais difíceis.

Aos amigos da PIB-BH e da ABU-BH, Osana, Medson, Antunes, Obadias, Mery, André e Tolentino, pelos passeios e trilhas por MG. Ao pr. Pedro e à Vivi, pelos almoços de domingo, em especial à Vivi, pelas conversas estatísticas. Às amigas Van, Sara e Angélica, por todo suporte e pelos passeios culturais.

À Dani, amiga nos momentos felizes e tristes e à sua família, pelos pratos mineiros. À Paula, amiga de todas as horas, inclusive as passadas na federal, pelo incentivo e dicas sobre concursos. Ao Allyson, por propor, e também por aceitar os desafios ciclísticos, faça sol ou chuva, calor ou “frio”.

À Lola Flores, por todo suporte nestes últimos dois anos, por me ensinar ver o outro lado da moeda. Ao Bernardo Limberger, pelas correções e pelo suporte linguístico. Também agradeço à Natália e à Maíra, amigas e companheiras de república, por compreenderem meu estresse nos últimos tempos.

Aos amigos da estatística, por todas as trocas de conhecimento, pelas ajudas e horas de estudo. Em especial à Jussiane, à Juliana, ao Paulo, ao Estevão, ao Zé Luis, ao Cristiano, por TODA ajuda computacional e teórica, também por TODAS conversas; à Nívea, ao Francisco e ao Márcio, por toda ajuda, estudos e partidas de boliche.

Aos professores do departamento de Estatística, por todo conhecimento transmitido e dedicação à ciência. Aos funcionários, em especial à Maísa, Cristina, Rogéria, Rose, auxiliar de serviços gerais Maria e aos bibliotecários Célio e Erlaine.

À FAPEMIG e à CAPES, pelas bolsas de estudo e pelo apoio financeiro prestado em diversas ocasiões.

Muito obrigada!
Silvana Schneider

Resumo

Na utilização da grande parte dos métodos apresentados na literatura para analisar dados de sobrevivência, é necessário fazer a suposição de que o mecanismo gerador da censura é não informativo, ou seja, assumir que a distribuição dos tempos de censura não depende de nenhum parâmetro da distribuição dos tempos de falha. Porém, em muitas situações, essa suposição pode ser inadequada e levar a inferências enganosas.

Com o objetivo de acomodar a dependência entre os tempos de falha e os tempos de censura informativa, consideramos a abordagem fornecida pelos modelos de fragilidade, especificamente a abordagem proposta por Huang & Wolfe (2002). Primeiramente propomos uma versão completamente paramétrica ao modelo de Huang & Wolfe (2002), utilizando a distribuição Weibull para ajustar os tempos de falha e os tempos de censura informativa. Em seguida, também propomos uma versão semiparamétrica, considerando a distribuição Exponencial por Partes para modelar os tempos de falha e de censura. Para esses dois modelos mostramos as etapas da estimação via abordagens de máxima verossimilhança e bayesiana. Posteriormente, propomos uma versão bayesiana do modelo de Huang & Wolfe (2002).

A censura informativa também pode influenciar as estimativas obtidas pelos modelos para analisar dados de sobrevivência com fração de cura. Para tanto, propomos um modelo paramétrico capaz de captar a dependência entre os tempos de promoção e os tempos de censura informativa, em que empregamos a distribuição Weibull para ajustar os tempos de promoção e os tempos de censura informativa. Além desse modelo, também propomos um modelo semiparamétrico para os dados de fração de cura com censura informativa, utilizando a distribuição Exponencial por Partes para modelar os tempos de promoção e tempos de censura. Apresentamos as etapas para estimação através das abordagens de máxima verossimilhança e bayesiana.

Realizamos estudos de simulação Monte Carlo, considerando três cenários: dados gerados com correlações positiva, negativa e nula entre os tempos de falha e censura, para dados com e sem fração de cura. Apresentamos uma aplicação em dados sobre mortalidade em centros de diálise renal dos Estados Unidos, cedidos pelo estudo chamado *Dialysis Outcomes and Practice Patterns Study* (DOPPS), utilizando os modelos propostos com censura informativa. Por último, mostramos uma aplicação em dados sobre câncer de melanoma, fornecidos pelo Programa *Surveillance, Epidemiology, and End Results* (SEER), empregando os modelos propostos com censura informativa e fração de cura.

Palavras-chave: Fração de cura, modelo Exponencial por Partes, distribuição Weibull.

Abstract

To use the most of the methods presented in the literature to analyze survival data, it is necessary to make the assumption that the mechanism generating the censoring is noninformative, that is, assume that the distribution of the censoring times does not depend on any parameter of the lifetime distribution. However, in many situations, this assumption can be inadequate and lead to misleading inferences.

With the objective to accommodate the dependence between lifetime times and informative censoring, we consider the approach provided by the frailty models, specifically the approach proposed by Huang & Wolfe (2002). First, we propose a fully parametric version of the Huang & Wolfe (2002)'s model, using the Weibull distribution to adjust the lifetime times and informative censoring times. Next, we also propose a semiparametric version, considering the Piecewise Exponential distribution to model the lifetime and censoring times. For these two models, we show the estimation steps via maximum likelihood and Bayesian approaches. Later we propose a Bayesian version of the Huang & Wolfe (2002)'s model.

Informative censoring can also influence the estimates obtained by the models to analyze multivariate survival data with cure rate. In order to do so, we propose a fully parametric model capable of capturing the dependence between promotion times and informative censoring times, in which we use the Weibull distribution to adjust promotion times and informative censoring times. In addition to this model, we also propose a semiparametric model for cure rate data with informative censoring, using the Piecewise Exponential distribution to model the promotion times and the censoring times. We present the steps for estimation through the maximum likelihood and Bayesian approaches.

Monte Carlo simulation studies were performed considering three scenarios: data generated with a positive, negative and null correlation between lifetime and censoring times for data with and without cure rate. An application is carried out on mortality data in renal dialysis centers in the United States, provided by the study called Dialysis Outcomes and Practice Patterns Study (DOPPS), using the models proposed with informative censoring. Finally, an application is presented in data on melanoma cancer, provided by the Surveillance, Epidemiology, and End Results (SEER) program, using the proposed models with informative censoring and cure rate.

Keywords: Cure Rate, Piecewise Exponential Model, Weibull distribution.

Índice

1	Introdução	1
1.1	Motivação e objetivos	6
1.2	Organização da tese	8
2	Definições e revisão da literatura	10
2.1	Introdução	10
2.2	Modelo Exponencial por Partes	11
2.2.1	Definições e modelo	13
2.3	Modelo de Cox	14
2.4	Modelos com fragilidade	15
2.4.1	Fragilidade em um contexto multivariado	16
2.5	Modelos de sobrevivência com fração de cura	19
2.5.1	Definições dos modelos de tempos de promoção	21
3	Modelos de fragilidade para censura informativa	24
3.1	Construção da função de verossimilhança	24
3.2	Abordagem proposta por Huang & Wolfe (2002)	27
3.3	Modelos propostos para censura informativa	33
3.3.1	Modelo Weibull	34

3.3.1.1	Abordagem frequentista	34
3.3.1.2	Abordagem bayesiana	35
3.3.2	Modelo Exponencial por Partes	38
3.3.2.1	Abordagem frequentista	39
3.3.2.2	Abordagem bayesiana	40
3.3.3	Versão bayesiana para o modelo de Huang & Wolfe (2002)	42
3.4	Modelos propostos para censura informativa com fração de cura	45
3.4.1	Modelo de fração de cura Weibull	47
3.4.1.1	Abordagem frequentista	47
3.4.1.2	Abordagem bayesiana	48
3.4.2	Modelo de fração de cura Exponencial por Partes	50
3.4.2.1	Abordagem frequentista	51
3.4.2.2	Abordagem bayesiana	52
4	Estudo de simulação Monte Carlo	54
4.1	Estudos Monte Carlo para modelos com censura informativa sem fração de cura	55
4.1.1	Avaliação dos estimadores de máxima verossimilhança	55
4.1.2	Avaliação dos estimadores bayesianos	67
4.1.3	Comparações entre os estimadores	70
4.2	Estudos Monte Carlo para modelos com censura informativa e fração de cura	84
4.2.1	Avaliação dos estimadores de máxima verossimilhança	85
4.2.2	Abordagem bayesiana	89
4.2.3	Comparação entre os estimadores	90
5	Aplicações em dados reais	100
5.1	Análise de dados de mortalidade em centros de diálise renal dos Estados Unidos	100

5.2	Análise de dados de câncer de melanoma	104
6	Considerações finais	110
A	Uma aproximação para a matriz de informação de Fisher para os mo- delos com censura informativa	113
A.1	Modelo de Huang & Wolfe (2002)	113
A.2	Modelo Weibull	116
A.3	Modelo Exponencial por Partes	118
A.4	Modelo de fração de cura Weibull	120
A.5	Modelo de fração de cura exponencial por partes	122
B	Estudos com dados simulados para os modelos com censura informativa	125
C	Seleção de modelos	132
C.1	Critério de informação da desviância	132
C.2	Estatística logaritmo da pseudoverossimilhança marginal	133
C.3	Critério de informação de Watanabe (2010)	134
	Bibliografia	136

Capítulo 1

Introdução

Análise de Sobrevivência e Confiabilidade englobam métodos para dados em que o objeto de estudo é o tempo até a ocorrência de um evento de interesse, que pode ser, por exemplo, o tempo até o óbito de um paciente (Análise de Sobrevivência) ou tempo até a falha de um equipamento (Confiabilidade). Os índices de sobrevivência, juntamente com a mortalidade, incidência e prevalência, são importantes indicadores para avaliar a eficiência de um sistema de saúde no controle de doenças (Danieli *et al.*, 2012). Lawless (2011) e Colosimo & Giolo (2006), entre outros autores, apresentaram uma relação de exemplos de dados de sobrevivência que surgem em diversas áreas do conhecimento, como Medicina, Engenharia e Ciências Sociais, e demonstraram como esses métodos podem ser aplicados.

Na Análise de Sobrevivência, além de acomodarmos a informação sobre o tempo até o evento de interesse, denominado por tempo de falha ou tempo de vida, podemos incluir as observações incompletas, nomeadas por censuras ou observações censuradas. A censura pode ocorrer quando o acompanhamento do indivíduo é interrompido, devido à ocorrência de óbito por motivos diferentes da causa do evento de interesse ou quando o indivíduo deixa de participar do estudo por questões pessoais, comumente chamada de censura por abandono ou *dropout*. A censura também pode ocorrer quando o indivíduo não experimenta o evento de interesse até o final do estudo, neste caso chamada de censura administrativa.

As censuras podem ser denominadas como:

- i*) Censura à direita, quando há perda de acompanhamento ou não ocorrência do evento de interesse durante o período de observação, ou antes do término do estudo. Dentro dessa categoria, podemos ter:
 - Censura do tipo I, quando a duração do estudo é pré-estabelecida. Observamos que, ao término do mesmo, ainda pode haver indivíduos que não apresentaram o evento de interesse. Neste caso, o número de falhas é aleatório (Lawless, 2011).

- Censura do tipo II, quando o término do estudo ocorre após um número pré-determinado de indivíduos experimentarem o evento de interesse. Ao final, os indivíduos que não apresentaram o evento de interesse são censurados. Neste caso, o tempo de acompanhamento é aleatório.
- Censura do tipo aleatória, quando o indivíduo deixa de ser acompanhado ao longo do estudo, por algum motivo não relacionado com a causa do evento de interesse.
- ii*) Censura à esquerda, quando o indivíduo estava sob risco de experimentar o evento de interesse antes de entrar no estudo, ou seja, quando o tempo registrado é maior que o tempo de falha.
- iii*) Censura intervalar, quando sabemos apenas que o evento de interesse ocorreu entre dois tempos observados.

A relação entre o tempo de falha e a censura pode ser diferenciada de acordo com as causas geradoras da censura. Por exemplo, em estudos clínicos sobre câncer de mama, algumas causas de observações incompletas podem ser: a) término do estudo; b) óbito devido a causas externas; c) remoção do estudo se houver evidências clínicas da ineficiência do tratamento; d) abandono do estudo devido aos efeitos colaterais da terapia que a paciente está recebendo. Em a) e b) a censura ocorre independentemente da falha, ou seja, ela não traz informação alguma sobre o tempo da falha, nesses casos dizemos que a censura é não informativa. Em c) e d) as observações censuradas parecem prever algo sobre o tempo de falha, ou seja, elas estão relacionadas com a falha, nesses casos dizemos que a censura é informativa (Lagakos, 1979). Segundo Braekers & Veraverbeke (2005), dados que podem conter os dois tipos de censura são denominados de dados com censura parcialmente informativa.

Para a aplicação da maioria dos métodos apresentados na literatura, é necessário fazer a suposição de que os tempos de falha e censura são independentes, ou seja, que o tempo de sobrevivência não é influenciado pelas causas geradoras de censura. Porém, em muitos casos essa suposição pode ser inadequada. Fisher & Kanarek (1974) apontaram que em muitos estudos médicos pode haver duas causas de censura: alguns indivíduos são censurados por abandonarem o estudo (nestes casos a suposição de independência é questionável, pois a saída pode estar relacionada com a causa do evento de interesse) e outros são censurados pois não sofreram o evento de interesse até o final do estudo. Williams & Lagakos (1977) comentaram que as observações censuradas devido ao agravamento ou à melhora da doença podem fornecer alguma informação sobre o tempo de sobrevivência, podendo não ser adequado ignorá-las no processo inferencial. Outros exemplos de dados de sobrevivência sujeitos à censura informativa podem ser encontrados em Lagakos (1979).

Leung *et al.* (1997) discutiram algumas suposições feitas sobre o mecanismo da censura e compararam o efeito da censura informativa em alguns métodos. Huang *et al.* (2004) apresentaram um método para testar a correlação entre os tempos de falha e censura baseado na correlação estimada entre os resíduos *martingale* do modelo para os tempos de falha e censura, respectivamente.

Com o objetivo de acomodar a censura informativa, Fisher & Kanarek (1974) propuseram um modelo que considera as censuras informativas e administrativas baseado na suposição de que as censuras por abandono durante o estudo sempre ocorrem no mesmo tempo. Além do mais, incluíram na função de sobrevivência um parâmetro que capta a associação entre os tempos de falha e censura, resultando em um aumento ou em uma diminuição do valor da função de sobrevivência. Lagakos & Williams (1978) apresentaram uma classe de modelos denominada por classe cone, em que é incorporado um parâmetro na função densidade de probabilidade dos tempos de falha, sob a restrição de que a soma entre as probabilidades de falha e censura seja um. Esse parâmetro reflete o grau que a censura afeta o tempo de sobrevivência. Assumindo que a censura somente ocorre em uma subpopulação definida pela distribuição da fragilidade, Link (1989) propôs estimador de Kaplan-Meier modificado.

Posteriormente, Emoto *et al.* (1990) sugeriram um modelo Weibull bivariado para acomodar a dependência entre os tempos de falha e censura. Scharfstein & Robins (2002) publicaram um método utilizando uma abordagem baseada na suposição de riscos proporcionais estratificado. Dunson & Dinse (2002) propuseram uma abordagem bayesiana para ajustar dados com censura intervalar informativa e com diversas causas que podem levar à ocorrência do evento de interesse. Nesse modelo de Dunson & Dinse (2002), para garantir identificabilidade, as diferentes causas foram assumidas como independentes, quando condicionadas em uma variável latente. Zhang & Rao (2004) propuseram um estimador de máxima verossimilhança generalizada (EMVG) para o modelo de riscos proporcionais com censura parcialmente informativa.

A correlação entre os tempos de falha e censura também foi estudada por Siannis (2004) e Siannis *et al.* (2005). Eles apresentaram um modelo paramétrico que permite incluir a dependência entre os tempos de falha e censura em termos de um parâmetro que expressa a dependência entre esses tempos, denotado por δ , e uma função viés, denotada por $B(t, \theta)$, em que θ é o parâmetro desconhecido de interesse.

Em Siannis (2004) e Siannis *et al.* (2005), foi assumido que a distribuição condicional do tempo observável de censura, C , dado o tempo observável de falha, T , é semelhante com a distribuição marginal de T ; porém, com a inclusão de um parâmetro que depende dos tempos de falha, δ . Se $\delta = 0$, os tempos de falha e censura são independentes. Os autores

mencionaram que o parâmetro δ pode ser pensado como uma medida da dependência entre T e C , e a função viés $B(t, \theta)$ como uma medida do padrão da dependência, ambos pré-especificados.

Siannis *et al.* (2005) apresentaram uma análise de sensibilidade para o modelo paramétrico exposto em (1.1), considerando o parâmetro δ fixo e assumindo valores pequenos, e distribuição Exponencial para ajustar os tempos de falha e censura. Siannis (2004) explorou o uso de modelos paramétricos definidos no trabalho de Siannis *et al.* (2005), apresentando o modelo para permitir a presença de censura informativa e não informativa, considerando a distribuição Weibull para ajustar os tempos de falha e censura. Posteriormente, Siannis (2011) modificou o modelo proposto em Siannis (2004) utilizando abordagem não paramétrica e o modelo de riscos proporcionais de Cox(1972) para ajustar os tempos de falha e censura, resultando em um modelo mais flexível. Porém, concluiu que este é mais intensivo computacionalmente do que os modelos apresentados em Siannis (2004) e Siannis *et al.* (2005).

Staplin (2012) propôs a utilização do Modelo Exponencial por Partes (MEP) para modelar os tempos na abordagem proposta por Siannis (2004) e Siannis *et al.* (2005). Staplin (2012) constatou, através de uma análise de sensibilidade, que o MEP preserva a simplicidade computacional dos modelos paramétricos de Siannis (2004) e Siannis *et al.* (2005) e proporciona a flexibilidade que a abordagem de Siannis (2011) possibilita.

Outra solução alternativa para modelar a densidade conjunta dos tempos de falha e censura é através das funções cópula, como podemos ver por exemplo em Jiang *et al.* (2005), que propuseram um modelo em que a estrutura da dependência entre os tempos de eventos recorrentes é formulada com uma função cópula, com distribuições marginais não especificadas, considerando o caso em que um evento primário censura um evento secundário. Huang & Zhang (2008) apresentaram uma análise de sensibilidade para modelos de regressão em sobrevivência, assumindo uma função cópula para ajustar a dependência entre os tempos de falha e a censura, com o modelo de riscos proporcionais de Cox (1972) e o estimador proposto por Breslow (1972) para ajustar as distribuições marginais. Ding (2010) forneceu condições de identificabilidade para um modelo bivariado com covariáveis em dados com censura informativa utilizando a função cópula para gerar a estrutura de dependência entre os tempos.

Os modelos baseados em funções cópula contêm um parâmetro que carrega a informação sobre o tipo de dependência entre os tempos de falha e censura, podendo ser negativa, positiva ou nula, como podemos ver em Suzuki (2012). Do mesmo modo que os modelos propostos por Siannis (2004) e Siannis *et al.* (2005), os modelos formados a partir de funções cópula também requerem uma análise de sensibilidade, e o parâmetro que induz

a dependência entre falha e censura precisa ser pré-definido.

Outra forma para modelar a densidade conjunta dos tempos de falha e censura é por meio dos modelos de fragilidade. Huang & Wolfe (2002) propuseram um modelo de fragilidade para censura informativa, capaz de acomodar a correlação entre os tempos de falha e a censura, para dados de sobrevivência agrupados. O modelo permite que algumas causas de censura possam ser analisadas como informativas e outras como não informativas. Liu *et al.* (2004) estenderam o modelo proposto por Huang & Wolfe (2002) para problemas com eventos recorrentes, considerando a fragilidade em nível do indivíduo. Chiang & Wang (2004) também apresentaram um modelo para eventos recorrentes, porém, com censura informativa, no qual a dependência entre os eventos recorrentes e os tempos de censura foi modelada mediante uma variável latente. Zhang *et al.* (2007) apresentaram um modelo de regressão para dados com censura intervalar e informativa, usando o modelo de fragilidade para captar a estrutura de dependência.

Santos Junior (2012) considerou que os tempos de falha e de censura são condicionalmente independentes, dado uma fragilidade no contexto univariado, assumindo distribuição Gama para a fragilidade e distribuições Exponencial e Weibull para os tempos de falha e censura. Souza (2015) apresentou dois modelos paramétricos para modelar dados de sobrevivência univariados com censura informativa, baseado no modelo proposto por Huang & Wolfe (2002), porém utilizando uma abordagem bayesiana em que as funções taxa de falha para os tempos são modeladas pela distribuição Weibull.

A censura informativa também pode afetar a função de sobrevivência em dados com fração de cura. Li *et al.* (2007) propuseram um modelo considerando uma função cópula de Arquimedes para fornecer a densidade conjunta dos tempos de falha e censura, com modelo de mistura padrão para fração de cura nos tempos de falha. Othus *et al.* (2009) propuseram um modelo semiparamétrico para dados com fração de cura, incluindo a censura informativa, sem fazer suposição paramétrica sobre a relação da dependência. Chaves & Rodrigues (2011) também apresentaram um modelo de fração de cura com censura informativa, em que os tempos de falha são ajustados pela distribuição Exponencial considerando o modelo de mistura padrão, e os tempos de censura pela distribuição Exponencial-uniforme. Freitas & Rodrigues (2013) verificaram por intermédio de um estudo de simulação o impacto causado pela censura informativa na convergência e no tamanho dos intervalos assintóticos; considerando o modelo proposto por Chaves & Rodrigues (2011).

1.1 Motivação e objetivos

Os dados de sobrevivência multivariados com censura informativa são muito comuns na prática real. Por exemplo, os dados do estudo sobre memória no envelhecimento analisados em Breitner *et al.* (1999), Bandeen-Roche & Liang (2002) e Cheng *et al.* (2010); o conjunto de dados sobre infecção pelo HIV em uma coorte de prostitutas femininas no Senegal (Naskar *et al.*, 2005); dados sobre o estudo da insuficiência renal crônica (Shou *et al.*, 2017), entre outros. Para exemplificar este tipo de dado consideramos o conjunto de dados sobre mortalidade em centros de diálise fornecido pelo *Dialysis Outcomes and Practice Patterns Study* (DOPPS). Em 1996, o DOPPS iniciou um estudo para avaliar os diferentes padrões e práticas utilizados nas instalações de diálise renal dos Estados Unidos. Atualmente, a coorte contém 108.788 pacientes de 711 instalações de diálise. De 1996 a 2016, 2.1951 (20,2%) pacientes morreram, 2.664 (2,4%) alteraram a modalidade de diálise, 2.931 (2,7%) se retiraram da diálise, 3.866 (3,6%) receberam transplantes e o restante permaneceu no estudo e na diálise, ou foram transferidos para outras clínicas de diálise, ou eram pacientes transitórios. De acordo com Huang & Wolfe (2002), a retirada, o transplante e a alteração da modalidade de diálise provavelmente são informativos. Como apontou Huang *et al.* (2004), os motivos exatos de retirada são desconhecidos, mas é suspeito que ela esteja relacionada uma piora do estado de saúde.

Como vimos na revisão da literatura apresentada na seção anterior, as abordagens frequentemente utilizadas para analisar dados de sobrevivência com censura informativa são os modelos baseados em funções cópula, os modelos baseados na abordagem proposta por Siannis (2004) e Siannis *et al.* (2005) e os modelos de fragilidade. Desses possíveis métodos, escolhemos o modelo de fragilidade, uma vez que ele fornece uma maneira simples para obter a função densidade de probabilidade conjunta dos tempos de falha e censura. Essa densidade conjunta pode ser obtida assumindo-se que os tempos de falha e censura são condicionalmente independentes dada uma fragilidade. De acordo com Huang *et al.* (2004), os modelos de fragilidade são frequentemente utilizados para analisar dados de sobrevivência multivariados. Além do mais, os modelos propostos por Siannis (2004) e os modelos formados a partir funções cópula requerem uma análise de sensibilidade, pois o parâmetro que induz a dependência entre os tempos de falha e censura precisa ser pré-especificado. Segundo Schwarz *et al.* (2013), os modelos baseados na função cópula são identificáveis se uma das distribuições marginais for conhecida ou se o parâmetro que mede a dependência entre os tempos de falha e censura for conhecido.

Com o objetivo de analisar dados de sobrevivência multivariados com censura informativa, propomos neste trabalho alguns modelos que consideram a combinação dos tópicos

“dados agrupados” e “censura informativa” no mesmo conjunto de dados. Utilizamos a abordagem proposta por Huang & Wolfe (2002) para fornecer a estrutura de dependência entre os tempos de falha e os tempos de censura. Propomos uma versão completamente paramétrica para o modelo de Huang & Wolfe (2002), empregando a distribuição Weibull para ajustar os tempos de falha e censura. De acordo com Wienke (2011), a conveniência do modelo Weibull é devido à sua flexibilidade e simplicidade das funções taxa de falha e de sobrevivência, e por poder ser aplicada em diversos conjunto de dados complexos (Murthy *et al.*, 2004). Conforme Lawless (2011), a distribuição Weibull é provavelmente a mais usada para ajustar tempos de vida, apresenta uma distribuição paramétrica com apenas dois parâmetros, capaz de acomodar funções taxa de falha estritamente crescentes, decrescentes e constantes.

Também propomos uma versão semiparamétrica, utilizando a distribuição Exponencial por Partes (EP) para modelar os tempos de falha e e os tempos de censura. Optamos por utilizar a distribuição Exponencial por Partes pois ela proporciona uma abordagem simples de modelagem, uma vez que dada uma partição do eixo do tempo, as funções taxa de falha são consideradas constantes em cada intervalo, permitindo ajustar funções taxa de falha de variadas formas (Demarqui *et al.*, 2008). O modelo Exponencial por Partes, embora simples e paramétrico, pode ser pensado como não paramétrico se o número de intervalos for grande ou não especificado. Portanto, o MEP serve como referência para comparações com outros modelos semiparamétricos ou com modelos completamente paramétricos (Ibrahim *et al.*, 2010). O modelo Exponencial por Partes preserva as características de um modelo paramétrico e é menos intensivo computacionalmente do que os modelos não paramétricos. A distribuição EP possui função de sobrevivência contínua, ao passo que o estimador de Kaplan-Meier tem função de sobrevivência discreta com saltos determinados pelos tempos de falha observados (Kim & Proschan, 1991). Para esses modelos Weibull e Exponencial por Partes com censura informativa, propostos neste trabalho, apresentamos os passos para realizar a estimação via abordagens de máxima verossimilhança e bayesiana.

Juntamente, propomos uma versão bayesiana para o modelo de Huang & Wolfe (2002), estendendo a abordagem apresentada por Clayton (1991) para modelos de fragilidade. A prova das propriedades assintóticas do método proposto por Huang & Wolfe (2002) ainda não foram apresentadas, mas eles comentaram que os resultados da simulação parecem apontar para a validade assintótica de seu método. De acordo com Mostafa & Ghorbal (2011), a abordagem bayesiana tem vantagem de lidar com amostras pequenas de dados censurados em relação à abordagem frequentista. Além do mais, a abordagem bayesiana permite a inclusão de informações fornecidas, por exemplo, de estudos anteriores,

por meio da distribuição *a priori*. Conforme Murthy *et al.* (2004), o método bayesiano é mais apropriado que qualquer outro método, quando o pesquisador tiver alguma informação subjetiva sobre os parâmetros. No entanto, quando não existe informação sobre os parâmetros, distribuições *a priori* não informativas podem ser utilizadas.

Ademais, com o propósito de analisar dados de sobrevivência multivariados com censura informativa e fração de cura, propomos estender a abordagem de Huang & Wolfe (2002) para incluir a proporção de indivíduos curados. Nesse sentido, propomos um modelo completamente paramétrico, considerando a distribuição Weibull para modelar os tempos de promoção e os tempos de censura. Por fim, propomos um modelo semi-paramétrico, utilizando a distribuição Exponencial por Partes para ajustar os tempos de promoção e de censura. Para ambos os modelos, Weibull e MEP, com censura informativa e fração de cura, apresentamos a estimação via abordagens de máxima verossimilhança e bayesiana.

1.2 Organização da tese

No Capítulo 2, apresentado a seguir, exibimos um resumo dos conceitos básicos empregados em Análise de Sobrevivência. Na Seção 2.2 mostramos os conceitos do Modelo Exponencial por Partes. Na Seção 2.3 apresentamos uma breve revisão sobre o modelo de Cox. Na Seção 2.4 exibimos uma revisão sobre os modelos de fragilidade no contexto multivariado, bem como uma relação das principais distribuições utilizadas no ajuste da fragilidade. Ademais, na Seção 2.5 apresentamos as definições dos modelos de fração de cura, especificamente os modelos de tempos de promoção.

No Capítulo 3 definimos as notações utilizadas nos modelos de sobrevivência com censura informativa, inicialmente para o modelo proposto por Huang & Wolfe (2002), e em seguida para os modelos propostos neste trabalho. Na Seção 3.1 mostramos a construção das função de verossimilhança para problemas com censura informativa. Na Seção 3.2 descrevemos detalhadamente o modelo proposto por Huang & Wolfe (2002), também exibimos os passos do algoritmo EM Monte Carlo.

Na Seção 3.3 apresentamos os modelos propostos nesta tese para dados de sobrevivência multivariados com censura informativa. Inicialmente, propomos uma extensão completamente paramétrica para o modelo de Huang & Wolfe (2002), utilizando a distribuição Weibull para ajustar os tempos de falha e os tempos de censura. Para este modelo, mostramos as etapas para estimação via abordagens frequentista e bayesiana. Também propomos uma versão semiparamétrica do modelo de Huang & Wolfe (2002),

considerando a distribuição Exponencial por Partes para ajustar os tempos de falha e de censura, apresentamos as equações para estimação via abordagens frequentista e bayesiana. Posteriormente, propomos a versão bayesiana para o modelo de Huang & Wolfe (2002), empregando processos de contagem.

Na Seção 3.4 mostramos os modelos propomos neste trabalho para analisar dados de sobrevivência com fração de cura e censura informativa. Primeiramente apresentamos um modelo paramétrico, em que usamos a distribuição Weibull para modelar os tempos de promoção e os tempos de censura, também consideramos a abordagem proposta por Huang & Wolfe (2002) para fornecer a estrutura de dependência entre os tempos de promoção e os tempos de censura informativa. Em seguida, exibimos um modelo semiparamétrico com distribuição Exponencial por Partes para ajustar os tempos de promoção e os tempos de censura informativa, similarmente utilizamos a abordagem de Huang & Wolfe (2002) para prover a estrutura de dependência entre os tempos. Para ambos os modelos apresentamos as etapas de estimação através das abordagens frequentista e bayesiana.

No Capítulo 4 exibimos as simulações de Monte Carlo para avaliar a consistência dos estimadores, para os modelos com censura informativa (Seção 4.1) e modelos para dados com fração de cura e censura informativa (Seção 4.2). No Capítulo 5 apresentamos aplicações em dois conjuntos de dados reais, sendo um deles, um estudo sobre mortalidade em centros de diálise renal dos Estados Unidos e o outro, um estudo sobre câncer de melanoma em pacientes diagnosticados no período de 2004 a 2013.

Por fim, no Capítulo 6 apresentamos os comentários finais sobre os modelos propostos neste trabalho. Também listamos algumas sugestões de trabalhos futuros.

Capítulo 2

Definições e revisão da literatura

2.1 Introdução

Neste trabalho utilizamos uma variável aleatória contínua não-negativa, denotada por T , para representar o tempo de falha. A função densidade de probabilidade associada à variável aleatória T será denotada por $f(t)$ com função de distribuição

$$F(t) = P(T \leq t) = \int_0^t f(u)du, \quad t > 0.$$

A probabilidade de um indivíduo não falhar até um certo tempo t , ou seja, a probabilidade de um indivíduo sobreviver até o tempo t , é dada pela função de sobrevivência

$$S(t) = P(T > t) = 1 - F(t), \quad t > 0. \quad (2.1)$$

Portanto, como $F(t)$ é uma função monótona não decrescente, com $\lim_{t \rightarrow \infty} F(t) = 1$, a função de sobrevivência $S(t)$ é uma função monótona não crescente, com $S(0) = 1$ e $S(\infty) = \lim_{t \rightarrow \infty} S(t) = 0$. Como os tempos são positivos, todas as funções são definidas no intervalo $[0, \infty)$.

Outra função importante em análise de sobrevivência é a função taxa de falha, denotada por $h(t)$. Segundo Colosimo & Giolo (2006), a função taxa de falha é mais informativa do que a função de sobrevivência, pois funções de sobrevivência semelhantes podem ter funções taxa de falha completamente diferentes. Essa função descreve a taxa instantânea de falha no tempo t é definida por

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t < T \leq t + \Delta t | T > t)}{\Delta t} = \frac{f(t)}{S(t)}. \quad (2.2)$$

A probabilidade aproximada da falha ocorrer em $(t, t + \Delta t]$, dado que o indivíduo sobreviveu até o tempo t , é expressa por $h(t)\Delta t$.

A função taxa de falha acumulada é definida como $H(t) = \int_0^t h(u)du$. Como consequência das equações (2.1) e (2.2), podemos derivar a função de sobrevivência, ou seja, sabendo que $h(t) = f(t)/S(t)$ e que $F(t) = 1 - S(t)$, portanto

$$F'(t) = -S'(t) \Leftrightarrow f(t) = -S'(t).$$

Logo,

$$h(t) = \frac{-S'(t)}{S(t)} \quad e \quad \frac{-S'(t)}{S(t)} = \frac{d}{dt} (-\log(S(t))).$$

Então,

$$H(t) = \int_0^t h(u)du = -\log \{S(t)\}.$$

Equivalentemente, temos

$$S(t) = \exp \{-H(t)\} = \exp \left\{ - \int_0^t h(u)du \right\}.$$

A função densidade de probabilidade para os tempos pode ser obtida por

$$f(t) = h(t) \exp \left\{ - \int_0^t h(u)du \right\}.$$

As relações acima nos mostram que se conhecemos uma das funções, podemos obter as demais. Mais detalhes podem ser vistos em Lawless (2011), Colosimo & Giolo (2006), Ibrahim *et al.* (2010), entre outros.

2.2 Modelo Exponencial por Partes

Breslow (1972) sugeriu uma abordagem de estimação alternativa àquela apresentada por Cox (1972), para estimar os coeficientes de regressão β e as taxas de falha λ_0 , e também uma estimativa mais simples para a função de sobrevivência. Breslow (1972) considerou a estimação simultânea de β e λ_0 através da função de verossimilhança conjunta, considerando a função taxa de falha constante entre os distintos tempos de falha, enquanto que Cox (1972) utiliza a função de verossimilhança condicional no conjunto de indivíduos sob risco, vide contribuição apresentada na discussão do trabalho de Cox (1972). Kalbfleisch & Prentice (1973) apresentaram a função de verossimilhança marginal para os parâmetros de regressão do modelo proposto por Cox (1972). Posteriormente,

Breslow (1974) apresentou uma aplicação em dados de leucemia, utilizando a proposta exposta em Breslow (1972). Essa abordagem introduzida por Norman Breslow em 1972, hoje é chamada de Modelo Exponencial por Partes (MEP).

Friedman (1982) apresentou propriedades assintóticas dos estimadores de máxima verossimilhança para os parâmetros do modelo proposto por Breslow (1972). Kim & Proschan (1991) utilizaram o MEP para estimar a função de sobrevivência e realizaram comparações entre o MEP e o estimador de Kaplan-Meier, concluindo que a função de sobrevivência fornecida pelo MEP é uma função contínua, ao passo que o estimador de Kaplan-Meier fornece uma função de sobrevivência discreta com descontinuidades nos tempos de falha distintos observados. Gamerman (1994) propôs uma abordagem bayesiana para a estimação dos parâmetros do MEP, fornecendo uma suavização nas estimativas da função de sobrevivência e da função taxa de falha.

Em Arjas & Gasbarra (1994), a função taxa de falha basal do MEP foi modelada não parametricamente com a utilização de um processo de saltos com estrutura *martingal* para a distribuição *a priori*, fornecendo uma maneira para se obter uma quantidade aleatória de intervalos, porém sem considerar covariáveis no modelo. O MEP também foi utilizado no ajuste da função taxa de falha do modelo de fragilidade de Qiou *et al.* (1999), no qual a fragilidade foi modelada por meio da distribuição Estável Positiva e a distribuição correlacionada proposta por Arjas & Gasbarra (1994) foi utilizada para ajustar as funções taxa de falha do MEP. A distribuição Exponencial por Partes foi utilizada para modelar a taxa de falha em testes de vida acelerados sob uma abordagem frequentista em Barbosa *et al.* (1996). Chen & Ibrahim (2001) propuseram um método de máxima verossimilhança para a estimação dos parâmetros em uma classe de modelos de sobrevivência semiparamétricos com fração de cura, em que os tempos são modelados pela distribuição Exponencial por Partes considerando o número de intervalos fixo.

Kim *et al.* (2007) propuseram uma classe de modelos semiparamétricos para dados de sobrevivência com fração de cura, baseada em um processo *martingal* para modelar a taxa de falha, proporcionando flexibilidade para a função sobrevivência. A estrutura dinâmica proposta permite que o número de intervalos e o tempo final de cada intervalo sejam aleatórios. Paralelamente, Demarqui *et al.* (2008) apresentaram uma modelagem bayesiana considerando o número de intervalos e o tempo final de cada intervalo como quantidades aleatórias, utilizando a estrutura do modelo partição produto (MPP) introduzido por Barry & Hartigan (1992). Posteriormente, Demarqui *et al.* (2012) estenderam o modelo proposto por Gamerman (1991), incluindo a estrutura do MPP para gerar os intervalos nesse modelo dinâmico.

2.2.1 Definições e modelo

Para definirmos o MEP precisamos especificar uma partição finita do eixo do tempo, $\{s_0, s_1, \dots, s_b\}$, em que $0 = s_0 < s_1 < s_2 < \dots < s_b < \infty$, $b \in \mathbb{N}$, com $s_b > t$, sendo t o tempo observado. Essa partição divide o eixo do tempo em b intervalos disjuntos, denotados por $I_1 = (s_0, s_1]$, $I_2 = (s_1, s_2]$, \dots , $I_b = (s_{b-1}, s_b]$. Dessa forma, temos a grade $\rho = \{s_0, s_1, \dots, s_b\}$, que particiona o eixo do tempo em b intervalos. As definições desta Seção foram extraídas de Demarqui (2010).

A função taxa de falha base é aproximada por segmentos de retas definidos pelos intervalos, $I_j = (s_{j-1}, s_j]$, $j = 1, \dots, b$. Portanto, a função taxa de falha é assumida como sendo constante em cada intervalo, e é dada por

$$h(t) = \lambda_j, \quad t \in I_j = (s_{j-1}, s_j], \quad j = 1, \dots, b,$$

em que λ_j são valores positivos para $j = 1, \dots, b$.

A função taxa de falha acumulada para o MEP, denotada por $H(t)$, é dada pela soma das áreas dos retângulos em que as bases são dadas pelos comprimentos dos intervalos I_j e as alturas pelas respectivas funções taxa de falha, λ_j . Logo, a função taxa de falha acumulada pode ser expressa como

$$H(t|\boldsymbol{\lambda}, \rho) = \sum_{j=1}^b \lambda_j (t_j - s_{j-1}),$$

em que t_j é definido, convenientemente, da seguinte forma

$$t_j = \begin{cases} s_{j-1}, & \text{se } t < s_{j-1}, \\ t, & \text{se } t \in I_j, \\ s_j, & \text{se } t > s_j, \end{cases}$$

para $j = 1, 2, \dots, b$, sendo t o tempo observado.

Uma vez que $S(t) = \exp\{-H(t)\}$, a função de sobrevivência para o MEP será dada por

$$S(t|\boldsymbol{\lambda}, \rho) = \exp \left\{ - \sum_{j=1}^b \lambda_j (t_j - s_{j-1}) \right\}. \quad (2.3)$$

A função densidade para a variável aleatória T é obtida tomando-se menos a derivada da função sobrevivência, dada em (2.3), em relação a t . Portanto, dizemos que a variável aleatória T segue uma distribuição Exponencial por Partes (EP) com grade ρ e vetor de

parâmetros $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_b)$, denotada por $T \sim EP(\rho, \boldsymbol{\lambda})$, se sua função densidade de probabilidade é dada por

$$f(t|\boldsymbol{\lambda}, \rho) = \lambda_j \exp \left\{ - \sum_{j=1}^b \lambda_j (t_j - s_{j-1}) \right\},$$

para $t \in I_j = (s_{j-1}, s_j]$ e $\lambda_j > 0, j = 1, 2, \dots, b$.

O modelo MEP serve como ponto de referência para fazer comparações com outros modelos semiparamétricos ou com modelos totalmente paramétricos para dados de sobrevivência (Ibrahim *et al.*, 2010). Este modelo é bastante geral pois pode acomodar várias formas para a função taxa de falha base. Além disso, podemos notar que se $b = 1$, temos o modelo Exponencial como caso particular.

2.3 Modelo de Cox

Com o objetivo de avaliar o efeito que uma ou mais covariáveis exercem sobre o tempo de falha, Cox (1972) propôs uma extensão ao método apresentado por Kaplan & Meier (1958), incluindo uma função não negativa das covariáveis, de forma multiplicativa na função taxa de falha.

O modelo proposto por Cox (1972) considera que a distribuição do tempo de falha, $T > 0$, pode ser ajustada através da função taxa de falha de T , condicional nas covariáveis $\boldsymbol{x} = (x_1, \dots, x_p)'$, denotada aqui por

$$h(t|\boldsymbol{x}) = h_0(t) \exp(\boldsymbol{\beta}'\boldsymbol{x}), \tag{2.4}$$

em que $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ denota o vetor de coeficientes de regressão, e $h_0(t)$ representa uma função não negativa, geralmente chamada de função taxa de falha base, pois $h_0(t)$ não depende de \boldsymbol{x} . A função taxa de falha de base $h_0(t)$ pode ser modelada parametricamente ou não parametricamente. O intercepto β_0 não é incluído na expressão acima, pois o componente $h_0(t)$ absorve esse termo constante.

Segundo Kalbfleisch (1978), o modelo apresentado em (2.4) é apropriado para variáveis aleatórias contínuas, sendo denominado por modelo de taxas de falhas proporcionais, pois tem a propriedade de garantir que a razão entre as taxas de dois indivíduos seja constante no tempo, desde que o efeito das covariáveis seja constante no tempo. Ou seja, a razão

das funções taxa de falha entre os indivíduos i e j é dada por

$$\frac{h_i(t|\mathbf{x}_i)}{h_j(t|\mathbf{x}_j)} = \frac{h_0(t) \exp(\boldsymbol{\beta}'\mathbf{x}_i)}{h_0(t) \exp(\boldsymbol{\beta}'\mathbf{x}_j)} = \exp(\boldsymbol{\beta}'\mathbf{x}_i - \boldsymbol{\beta}'\mathbf{x}_j) = \exp(\boldsymbol{\beta}'(\mathbf{x}_i - \mathbf{x}_j)).$$

As covariáveis podem aumentar ou diminuir o risco da ocorrência do evento de interesse, porém talvez existam fatores não mensuráveis ou desconhecidos que influenciam a ocorrência do evento de interesse. Nestes casos, pode ser inserido um efeito aleatório para modelar esses fatores, como veremos na Seção 2.4.

Alternativamente, a expressão (2.4), temos a função de sobrevivência

$$P(T \geq t|\mathbf{x}) = \exp\{-H_0(t) \exp(\boldsymbol{\beta}'\mathbf{x})\}, \quad (2.5)$$

em que $\exp\{-H_0(t)\} = P(T \geq t|\mathbf{x} = \mathbf{0})$ é a função de sobrevivência base, não especificada num contexto semiparamétrico. De acordo com Kalbfleisch (1978), a expressão (2.5) tem a vantagem de acomodar distribuições contínuas e discretas para os tempos de falha.

2.4 Modelos com fragilidade

A associação entre os tempos de falha pode ser encontrada quando temos, por exemplo, vários tempos observados para um mesmo indivíduo (eventos recorrentes), ou tempos de indivíduos vindos da mesma família (correlacionados geneticamente), ou ainda tempos de pacientes vindos do mesmo bloco cirúrgico. Inicialmente, Clayton (1978) explorou a interpretação de riscos relativos que compartilham um termo aleatório comum. Para captar essa associação entre os indivíduos, Vaupel *et al.* (1979), no estudo sobre o impacto da heterogeneidade sobre a dinâmica da mortalidade, definiram a fragilidade como sendo uma variável aleatória multiplicativa na taxa de mortalidade. Posteriormente, Clayton & Cuzick (1985) propuseram uma generalização multivariada do modelo de taxas de falha proporcionais de Cox(1972), incluindo o efeito aleatório, para representar a heterogeneidade, e as covariáveis representando os efeitos fixos.

Clayton (1991) apresentou uma abordagem bayesiana semiparamétrica para analisar dados de sobrevivência multivariados, utilizando o modelo de taxas de falha proporcionais, condicional na fragilidade. Este modelo é uma generalização semiparamétrica do trabalho de Clayton & Cuzick (1985). Em paralelo à abordagem bayesiana, Nielsen *et al.* (1992) propuseram uma abordagem com processo de contagem e uma variável subjacente (fragilidade), que aumenta a susceptibilidade ao evento de interesse.

2.4.1 Fragilidade em um contexto multivariado

Segundo Colosimo & Giolo (2006), uma das abordagens utilizadas para análise de dados de sobrevivência multivariados é a abordagem condicional, em que condicionado em uma variável aleatória, as demais são independentes. A fragilidade pode ser usada para modelar a associação entre os tempos de eventos, nesse caso, o modelo de fragilidade assume que todos os eventos em um grupo são independentes, dado as variáveis de fragilidade (Wienke, 2011). Hanagal (2011) também argumentou que os modelos de fragilidade podem ser aplicados a dados de sobrevivência multivariados para captar a correlação entre os tempos dos indivíduos de um mesmo grupo ou para captar a correlação presente nos eventos recorrentes para um mesmo indivíduo.

O modelo de fragilidade aplicado a estes casos é conhecido por *modelo de fragilidade compartilhada*, pois os indivíduos que estão no mesmo grupo compartilham o mesmo efeito aleatório, ou seja, eles apresentam a mesma fragilidade. O modelo é formulado estendendo-se o modelo proposto por Cox (1972), no qual o efeito aleatório é introduzido para captar a associação dos tempos em nível de grupo, ou nos casos de eventos recorrentes para captar a associação entre os diferentes tempos associados a um mesmo indivíduo.

A função taxa de falha para o i -ésimo indivíduo no k -ésimo grupo, condicional ao valor latente (fragilidade) z_k e a um vetor de dimensão p de covariáveis, que podem ser em nível de grupo e em nível de indivíduo, $\mathbf{x}_{i,k}$, é expressa por

$$h_{i,k}(t|z_k, \mathbf{x}_{i,k}) = h_0(t)z_k \exp(\boldsymbol{\beta}'\mathbf{x}_{i,k}), \quad (2.6)$$

em que, $t_{i,k}$ denota o tempo de sobrevivência para o i -ésimo sujeito no k -ésimo grupo, sendo $i = 1, 2, \dots, n_k$, em que n_k representa o número de indivíduos no k -ésimo grupo, e $k = 1, 2, \dots, m$, m representa o número total de grupos; $\boldsymbol{\beta}$ denota vetor $p \times 1$ de coeficientes de regressão desconhecidos; $h_0(t)$ denota função taxa de falha basal desconhecida e z_k denota a fragilidade (não observável) associada ao k -ésimo grupo.

O modelo apresentado em (2.6) pode ser reescrito como

$$h_{i,k}(t|w_k, \mathbf{x}_{i,k}) = h_0(t_{i,k}) \exp(\boldsymbol{\beta}'\mathbf{x}_{i,k} + w_k),$$

em que z_k em (2.6) equivale a $\exp(w_k)$. Normalmente assume-se que os efeitos aleatórios, w_k , tem média zero e variância desconhecida (Colosimo & Giolo, 2006).

A razão das funções taxa de falha entre o indivíduo i e indivíduo j , neste caso, depende tanto da diferença entre as covariáveis quanto do valor da fragilidade, como podemos ver

na expressão

$$\frac{h_{i,k}(t|\mathbf{x}_{i,k}, z_k)}{h_{j,k'}(t|\mathbf{x}_{j,k'}, z_{k'})} = \frac{h_0(t)z_k \exp(\mathbf{x}'_{i,k}\boldsymbol{\beta})}{h_0(t)z_{k'} \exp(\boldsymbol{\beta}'\mathbf{x}_{j,k'})} = \frac{z_k}{z_{k'}} \exp(\boldsymbol{\beta}'(\mathbf{x}_{i,k} - \mathbf{x}_{j,k'})). \quad (2.7)$$

A partir da expressão (2.7), podemos ter as seguintes interpretações distintas para o vetor $\boldsymbol{\beta}$ (Colosimo & Giolo, 2006):

- Quando estamos comparando dois indivíduos, i e j , de um mesmo grupo, ou seja $k = k'$, o valor da variável fragilidade será o mesmo. Neste caso, a interpretação do vetor de coeficientes $\boldsymbol{\beta}$ é semelhante ao caso com dados independentes, como podemos ver na razão dada por

$$\frac{h_{i,k}(t|\mathbf{x}_{i,k}, z_k)}{h_{j,k}(t|\mathbf{x}_{j,k}, z_k)} = \frac{h_0(t)z_k \exp(\boldsymbol{\beta}'\mathbf{x}_{i,k})}{h_0(t)z_k \exp(\boldsymbol{\beta}'\mathbf{x}_{j,k})} = \exp(\boldsymbol{\beta}'(\mathbf{x}_{i,k} - \mathbf{x}_{j,k})).$$

- Quando estamos comparando dois indivíduos, i e j , com os mesmos valores das covariáveis, porém de grupos diferentes. Neste caso a razão entre as taxas de falha será a razão entre as fragilidade, obtida por

$$\frac{h_{i,k}(t|\mathbf{x}, z_k)}{h_{j,k'}(t|\mathbf{x}, z_{k'})} = \frac{h_0(t)z_k \exp(\boldsymbol{\beta}'\mathbf{x})}{h_0(t)z_{k'} \exp(\boldsymbol{\beta}'\mathbf{x})} = \frac{z_k}{z_{k'}}.$$

- Finalmente, quando estamos comparando dois indivíduos pertencentes a grupos distintos e com covariáveis diferentes, temos a equação apresentada em (2.7).

Frequentemente, utiliza-se distribuições paramétricas para o termo de fragilidade z_k , sendo mais populares as distribuições com média finita tal como a distribuição Gama e distribuição Log-normal (Ibrahim *et al.*, 2010). A distribuição Gama é utilizada na maioria das aplicações, pois é possível encontrar a distribuição marginal analiticamente (Wienke, 2011). Clayton & Cuzick (1985) utilizaram a distribuição Gama para modelar a distribuição da fragilidade, pois ela mantém a distribuição a *posteriori* com formas simples. Shih & Louis (1995) mostraram que diferentes distribuições de fragilidade podem levar a estruturas de correlação bem diferentes. Aslanidou *et al.* (1998) utilizaram a distribuição Gama(κ, κ) para modelar a distribuição da fragilidade, em que a média igual a um faz com que os parâmetros do modelo de fragilidade sejam identificáveis.

Hougaard (1986) propôs o uso de uma família de distribuições com três parâmetros, denominada Estável Positiva. Essa família inclui distribuições definidas nos números positivos, tais como, distribuição Gama, distribuições degeneradas e distribuição Normal Inversa como casos particulares, as propriedades dessa família de distribuições podem ser

vistas em Hougaard (1986). Qiou *et al.* (1999) descreveram uma modelagem bayesiana para dados de sobrevivência multivariados com dependência, usando a distribuição Estável Positiva, $S_\alpha(\sigma, \beta, \epsilon)$, para a fragilidade.

Quando o modelo de fragilidade está escrito na forma da equação (2.6), em geral assume-se que as fragilidades são independente e identicamente distribuídas com densidade de probabilidade conhecida, com valores nos reais positivos. De acordo com Wienke (2011), as distribuições Log-normal e Gama são as mais importantes e frequentemente utilizadas para a distribuição da fragilidade. Além disso, os modelos de fragilidade Log-normal são úteis para modelar a dependência em problemas multivariados, sendo a principal vantagem devido à flexibilidade nas estruturas de dependência por meio da distribuição Normal Multivariada (Wienke, 2011). Hougaard (1984) compararam os modelos de fragilidade utilizando as distribuições Gama e Normal Inversa para a fragilidade. Segundo Ibrahim *et al.* (2010), se utilizarmos as distribuições Gama e Log-normal para ajustar o efeito aleatório inserido no modelo de Cox, o modelo não preserva a propriedade de riscos proporcionais, quando consideramos abordagem marginal.

Assumimos distribuição Normal para a variável aleatória W , denota por $W \sim N(\mu; \sigma^2)$, para gerar fragilidades na forma $Z = \exp(W)$. A relação entre os parâmetros das distribuições para as variáveis aleatórias W e Z é expressa por (vide, por exemplo, Wienke (2011))

$$\begin{aligned} m &= \mathbf{E}(Z) = e^{\mu + \frac{\sigma^2}{2}}, \\ s^2 &= \mathbf{V}(Z) = e^{2\mu + \sigma^2(e^{\sigma^2} - 1)}. \end{aligned} \tag{2.8}$$

Os modelos de fragilidade Log-normal podem ser escritos de duas formas: *i*) se assumirmos a restrição que $m = 1$ em (2.8), então temos a definição padrão introduzida originalmente por Vaupel *et al.* (1979), implicando que a fragilidade média em uma população é igual a um (no início do seguimento); *ii*) se assumirmos a restrição $\mu = 0$ em (2.8), isso significa que o logaritmo da fragilidade tem média zero, sendo esta a forma mais comum na literatura e nos pacotes estatísticos por causa da relação direta com os modelos lineares generalizados. Portanto, é natural assumir que a variável aleatória transformada $W = \log(Z)$ tem distribuição Normal com média zero (Wienke, 2011).

Logo, podemos utilizar a distribuição Log-normal(0, τ), em que $\tau = 1/\sigma^2$, para modelar a distribuição do efeito aleatório, ou seja, $Z \sim LN(0, \tau)$, em que τ é o parâmetro de precisão, com função densidade dada por

$$f(z) = \frac{\sqrt{\tau}}{\sqrt{2\pi z}} \exp \left\{ \frac{-\tau(\log(z))^2}{2} \right\}.$$

Outra vantagem do modelo do modelo de fragilidade Log-normal, além da flexibili-

dade para modelar diversas estruturas de dependência, é a sua interpretação, tornando esta classe modelo mais conveniente para aplicação por pesquisadores de diversas áreas do conhecimento (Wienke, 2011). Portanto, neste trabalho utilizaremos o modelo de fragilidade com distribuição Log-normal para o efeito aleatório.

2.5 Modelos de sobrevivência com fração de cura

Modelos de sobrevivência com fração de cura, também denominados por modelos de sobrevivência com longa duração, são modelos pertencentes a uma classe em que algumas unidades experimentais não irão experimentar o evento de interesse, pois são não-susceptíveis. Esses modelos são aplicados em larga escala na área médica para dados de sobrevivência associados a doenças oncológicas, pois devido ao avanço científico, a eficácia dos tratamentos tem aumentado, resultando em um maior percentual de pacientes curados. Exemplos em que estes modelos foram utilizados podem ser vistos nos livros Maller & Zhou (1996), Frederick *et al.* (2002), Ibrahim *et al.* (2010), Klein *et al.* (2016), entre outros. No entanto, os modelos com fração de cura também podem ser utilizados nas áreas da engenharia, Nelson (2005) e Meeker & LuValle (1995); em dados de criminalidade, Partanen (1969), Maltz & McCleary (1977) e Broadhurst & Mailer (1990); em dados financeiros, Anscombe (1961); são alguns dos exemplos em que estes modelos foram empregados.

Os modelos encontrados na literatura para ajustar dados de sobrevivência com longa duração são classificados em duas classes: *i*) os modelos derivados de uma abordagem de mistura, inicialmente propostos por Boag (1949) e Berkson & Gage (1952), denominados de Modelos de Mistura; *ii*) os modelos provenientes de uma abordagem proposta por Yakovlev *et al.* (1993), chamados de Modelos de Tempos de Promoção.

Boag (1949) propôs um modelo formado por uma mistura de distribuições, composta pela distribuição Log-normal para representar os tempos de falha, pela distribuição degenerada para representar os tempos de sobrevivência dos indivíduos imunes e um componente para representar a probabilidade dos indivíduos sofrerem a recidiva. Berkson & Gage (1952) apresentaram um modelo de mistura, semelhante ao modelo de Boag (1949), porém com distribuição Exponencial para ajustar os tempos de falha e uma distribuição degenerada para representar a sobrevivência dos indivíduos curados, dado por

$$S_{\text{pop}}(t) = \pi + (1 - \pi)S^*(t), \quad t > 0, \quad (2.9)$$

em que $\pi \in (0, 1)$ representa a fração dos indivíduos curados na população, $S^*(t)$ denota

a função de sobrevivência para o grupo dos indivíduos não curados na população, vide Ibrahim *et al.* (2010).

Farewell (1977, 1982) sugeriram um modelo em que a função logística fosse empregada para modelar a probabilidade de incidência. Kuk & Chen (1992) e Taylor (1995) consideraram estimação não paramétrica para os tempos de falha. Calsavara *et al.* (2013) incluíram um efeito aleatório no modelo de mistura proposto por Boag (1949) e Berkson & Gage (1952), para captar a heterogeneidade das duas subpopulações, modelando a fragilidade através da distribuição Gama. Gonzales *et al.* (2013) incorporaram covariáveis no modelo de Calsavara *et al.* (2013), tanto na função de sobrevivência dos indivíduos suscetíveis quanto na proporção de curados.

Uma abordagem alternativa aos modelos de mistura foi proposta por Yakovlev *et al.* (1993, 1994) e Yakovlev & Tsodikov (1996), baseada no número de células tumorais e no tempo que cada uma dessas células leva para desenvolver o tumor metastático detectável. Em Yakovlev *et al.* (1994), os autores comentaram sobre a dificuldade em interpretar o parâmetro que representa a proporção de indivíduos curados, de forma biologicamente plausível nos modelos de mistura padrão, alegando que nestes modelos todos os pacientes com pelos menos uma célula maligna são considerados como uma população homogênea, todos tendo a mesma probabilidade de sobreviver.

Para incluir a informação do tamanho do tumor, Yakovlev *et al.* (1993) propuseram um modelo em que o número de células tumorais ativas foi pressuposto como sendo uma variável aleatória latente, N . Neste modelo, indivíduos com $N = 0$ células metastáticas ativas são considerados indivíduos curados, enquanto que os indivíduos com $N > 0$ células tumorais correspondem a fração de indivíduos suscetíveis na população. Associado a cada célula metastática existe uma variável aleatória não observável $Z_k, k = 1, \dots, N$, para denotar o tempo que estas células levam até desenvolver o câncer detectável, denominados tempos de promoção.

Chen *et al.* (1999) apresentaram uma abordagem bayesiana para o modelo proposto por Yakovlev *et al.* (1993), considerando distribuição Weibull para os tempos de promoção e mostraram que o modelo preserva a estrutura de riscos proporcionais, se é assumido que o número de células tem distribuição Poisson. Ibrahim *et al.* (2001) publicaram uma versão semiparamétrica do modelo bayesiano de Chen *et al.* (1999), em que os tempos de promoção são ajustados pela distribuição Exponencial por Partes. Chen *et al.* (2002b) também estudaram o modelo semiparamétrico com tempos de promoção seguindo uma distribuição EP e propuseram uma *priori* informativa para as taxas de falha. Sinha *et al.* (2003) apresentaram uma extensão do modelo de Chen *et al.* (1999) para dados multivariados, em que os efeitos aleatórios são ajustados pela distribuição Estável Positiva.

Yin & Ibrahim (2005) sugeriram uma estrutura aditiva para a função taxa de falha, ao invés do modelo multiplicativo proposto por Chen *et al.* (1999). Yin (2005) propôs uma generalização para o trabalho de Chen *et al.* (1999), incluindo um efeito aleatório para modelar a heterogeneidade das células tumorais de diferentes indivíduos.

Kim *et al.* (2007) propuseram um modelo dinâmico para ajustar as funções taxa de falha do modelo de Ibrahim *et al.* (2001), considerando uma grade aleatória para o eixo do tempo e variação do efeito das covariáveis no tempo. Demarqui *et al.* (2014) apresentaram uma generalização do trabalho de Ibrahim *et al.* (2001), em que a grade do MEP é considerada uma quantidade aleatória, obtida através do modelo Partição Produto. Rodrigues *et al.* (2009) estenderam o modelo proposto por Chen *et al.* (1999), e forneceram um mecanismo para a construção de modelos de fração de cura de tempos de promoção, podendo considerar diferentes distribuições para variável latente N . Castro *et al.* (2009) e Cancho *et al.* (2011) aplicaram a abordagem de Rodrigues *et al.* (2009), considerando distribuição Binomial Negativa para o número de causas competitivas e distribuição Weibull e Exponencial por Partes para os tempos.

2.5.1 Definições dos modelos de tempos de promoção

Para incluir a informação sobre o tumor nos modelos de fração de cura, Yakovlev & Tsodikov (1996) apresentaram um modelo no qual o número de células malignas sobreviventes é incorporado, bem como o tempo que essas células levam até desenvolver um tumor detectável. O número de células tumorais foi ajustado por uma distribuição Poisson, com média θ , $\theta > 0$. Esse modelo pode ser derivado como apresentado em Ibrahim *et al.* (2010), de onde retiramos as expressões expostas nesta seção.

Alternativamente, aos modelos de mistura, Yakovlev & Tsodikov (1996) apresentaram um modelo baseado na estrutura de riscos competitivos. Seja N uma variável aleatória não observável para denotar o número de causas competindo para a ocorrência do evento de interesse. Assuma que N tem distribuição Poisson com média θ . Seja Z_i uma variável aleatória que denota o tempo de ocorrência do evento de interesse devido à i -ésima causa, definido como tempo de promoção. Dado $N = n$, as variáveis aleatórias não observáveis $Z_i, i = 1, \dots, n$ são assumidas independentes e identicamente distribuídas, com função de distribuição comum dada por $F(z) = 1 - S(z)$, que não depende de N .

Assim, o tempo para a ocorrência do evento de interesse é dada por $T = \min\{Z_i, 0 \leq i \leq N\}$, em que $P(Z_0 = \infty) = 1$.

A função de sobrevivência populacional é dada por

$$\begin{aligned}
S_{pop}(t) &= P(\text{n\~ao ocorrer o evento de interesse at\~e o tempo } t) \\
&= P(N = 0) + P(Z_1 > t, \dots, Z_N > t, N \geq 1) \\
&= P(N = 0) + P(Z_1 > t, \dots, Z_N > t | N \geq 1) \times P(N \geq 1), \quad t > 0.
\end{aligned} \tag{2.10}$$

Como é assumido $N \sim Poisson(\theta)$, a função de sobrevivência resulta em

$$\begin{aligned}
S_{pop}(t) &= \exp(-\theta) + \sum_{k=1}^{\infty} S(t)^k \frac{\theta^k}{k!} \exp(-\theta) \\
&= \exp(-\theta + \theta S(t)) = \exp(-\theta F(t)), \quad t > 0.
\end{aligned} \tag{2.11}$$

Podemos notar que $S_{pop}(\infty) = \exp(-\theta) > 0$, portanto (2.11) corresponde a uma função de sobrevivência imprópria.

Yakovlev *et al.* (1994) salientou que a fórmula apresentada em (2.10) mostra explicitamente a contribuição de duas características no crescimento do tumor: o número inicial de células malignas sobreviventes e a probabilidade da progressão dessas células.

Podemos notar que a probabilidade de cura é dada por

$$S_{pop}(\infty) \equiv P(N = 0) = \exp(-\theta). \tag{2.12}$$

Se $\theta \rightarrow \infty$, a fração de cura tende a 0, se $\theta \rightarrow 0$, a fração de cura tende a 1.

A partir de (2.12) as funções densidade de probabilidade e função taxa de falha, ambas impróprias, podem ser derivadas, resultando em

$$f_{pop}(t) = \theta f(t) \exp(-\theta F(t))$$

e

$$h_{pop}(t) = \theta f(t). \tag{2.13}$$

Note que, $h_{pop}(t) \rightarrow 0$ à medida que $t \rightarrow \infty$, e $\int_0^{\infty} h_{pop}(t) dt < \infty$. Também podemos notar que a função taxa de falha $h_{pop}(t)$ é obtida de forma multiplicativa de θ e $f(t)$, e portanto preserva a estrutura de riscos proporcionais, quando as covariáveis são incluídas através de θ .

A função de sobrevivência da população suscetível é dada por

$$S^*(t) = P(T > t | N \geq 1) = \frac{\exp(-\theta F(t)) - \exp(-\theta)}{1 - \exp(-\theta)}, \quad t > 0, \tag{2.14}$$

em que $S^*(t)$ é uma função de sobrevivência própria.

De acordo com Chen *et al.* (1999), a relação matemática entre o modelo de mistura (2.9) e o modelo de tempos de promoção (2.11) pode ser obtida quando escrevemos a função de sobrevivência populacional do modelo (2.11) da seguinte forma

$$S_{pop}(t) = \exp(-\theta) + \{1 - \exp(-\theta)\} S^*(t), \quad t > 0,$$

em que $S^*(t)$ é dada por (2.14). Logo, a função de sobrevivência $S_{pop}(t)$ do modelo de mistura é obtida quando $\pi = \exp(-\theta)$ e a função de sobrevivência associada à população de não curados é dada por $S^*(t)$. Consequentemente, todo modelo definido como em (2.11) pode ser escrito como um modelo de mistura padrão, e todo modelo definido como (2.9) corresponde de alguma forma a um modelo de tempos de promoção.

Li *et al.* (2001) mostraram que os modelos de tempos de promoção, dados por $S_{pop}(t) = \exp(-\theta F(t))$, somente são identificáveis se os tempos de promoção são ajustados por uma distribuição paramétrica. Chen *et al.* (2002b) apontaram que a estimação do parâmetro da fração de cura, θ , é altamente afetada pela forma não paramétrica de $F(t)$, sugerindo a utilização de um número pequeno de intervalos para a distribuição Exponencial por Partes; outro fato que pode deixar o modelo não identificável ou quase não identificável, é quando o tempo de acompanhamento for insuficiente e ocorrerem poucos eventos, vide Li *et al.* (2001).

Capítulo 3

Modelos de fragilidade para censura informativa

A presença de censuras nos dados de sobrevivência faz com que seja necessário adotar algumas suposições sobre a distribuição conjunta entre os tempos de falha e censura, quando o interesse for, por exemplo, estimar os parâmetros associados à distribuição dos tempos até a falha. Segundo Staplin (2012), se utilizarmos os métodos padrão, que supõem independência entre os tempos de falha e censura, a função de sobrevivência poderá ser sobrestimada quando a associação entre os tempos de falha e censura for positiva, e subestimada quando a associação for negativa. Além do mais, de acordo com Staplin (2012), uma possível causa da censura informativa pode ser atribuída àqueles fatores associados ao tempo de censura, que também estão relacionados com o tempo até a falha.

3.1 Construção da função de verossimilhança

Seja Y uma variável aleatória não negativa que representa o tempo observável, sendo $Y = \min(T, C, A)$, em que T denota a variável aleatória tempo até a falha, C denota a variável aleatória tempo até a censura informativa e A denota a variável aleatória tempo até a censura administrativa. Denotaremos por $\delta^{(T)}$ a variável aleatória indicadora de falha e por $\delta^{(C)}$ variável aleatória indicadora de censura informativa. Então, temos que

$$\delta^{(T)} = \begin{cases} 1, & \text{se } T \leq \min(C, A), \\ 0, & \text{caso contrário,} \end{cases} \quad \text{e} \quad \delta^{(C)} = \begin{cases} 1, & \text{se } C \leq \min(T, A), \\ 0, & \text{caso contrário.} \end{cases}$$

Para indicar que ocorreu uma censura administrativa utilizamos $\delta^{(T)} = 0$ e $\delta^{(C)} = 0$. As

definições apresentadas nesta seção foram extraídas de Lawless (2011).

Seja $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(T)}, \boldsymbol{\theta}^{(C)}, \boldsymbol{\theta}^{(A)})$ o vetor de parâmetros, em que $\boldsymbol{\theta}^{(T)}$, $\boldsymbol{\theta}^{(C)}$ e $\boldsymbol{\theta}^{(A)}$ são os vetores de parâmetros associados às distribuições dos tempos de falha, censura informativa e censura administrativa, respectivamente.

Para os tempos de falha, censura informativa e censura administrativa, respectivamente, denotaremos:

$f^{(T)}(y_i|\boldsymbol{\theta}^{(T)})$, $f^{(C)}(y_i|\boldsymbol{\theta}^{(C)})$ e $f^{(A)}(y_i|\boldsymbol{\theta}^{(A)})$ como funções densidade,
 $S^{(T)}(y_i|\boldsymbol{\theta}^{(T)})$, $S^{(C)}(y_i|\boldsymbol{\theta}^{(C)})$ e $S^{(A)}(y_i|\boldsymbol{\theta}^{(A)})$ como funções de sobrevivência e
 $h^{(T)}(y_i|\boldsymbol{\theta}^{(T)})$, $h^{(C)}(y_i|\boldsymbol{\theta}^{(C)})$ e $h^{(A)}(y_i|\boldsymbol{\theta}^{(A)})$ como funções taxa de falha.

Se os tempos de falha T são dependentes dos tempos de censura C , precisamos encontrar a distribuição conjunta de (T, C) . Assumindo que todos os tempos de falha T e censura C são mutuamente independentes, e assumindo que a distribuição dos tempos de censura C não depende de nenhum parâmetro da distribuição dos tempos de falha T , o cálculo da função de verossimilhança fica simples algebricamente Lawless (2011). Nesse caso, precisamos identificar as distribuições dos tempos de falha e censura, sem levar em conta a relação entre esses tempos.

Assumindo o mecanismo de censura não informativa, isto é, que distribuição dos tempos censurados $f^{(C)}(y_i)$ não depende dos parâmetros associados à distribuição dos tempos de falha $f^{(T)}(y_i)$ (Kalbfleisch & Prentice, 2011), as seguintes probabilidades podem ser obtidas pelo simples produto entre as funções dos tempos de falha e censura:

$$\begin{aligned} P(T \in (y_i, y_i + \Delta y_i], C > y_i) &= P(T \in (y_i, y_i + \Delta y_i])S^{(C)}(y_i) \\ &= f^{(T)}(y_i)S^{(C)}(y_i), \end{aligned}$$

$$\begin{aligned} P(C \in (y_i, y_i + \Delta y_i], T > y_i) &= P(C \in (y_i, y_i + \Delta y_i])S^{(T)}(y_i) \\ &= f^{(C)}(y_i)S^{(T)}(y_i), \end{aligned}$$

e

$$P(T > y_i, C > y_i) = S^{(T)}(y_i)S^{(C)}(y_i).$$

Considerando o conjunto de dados observados, $D = \{(y_i, \delta_i^{(T)}, \delta_i^{(C)})\}$, $i = 1, \dots, n$, a função de verossimilhança é dada por

$$\begin{aligned} L \propto & \prod_{i=1}^n [f^{(T)}(y_i)S^{(C)}(y_i)]^{\delta_i^{(T)}} \\ & \times [f^{(C)}(y_i)S^{(T)}(y_i)]^{\delta_i^{(C)}} \\ & \times [S^{(T)}(y_i)S^{(C)}(y_i)]^{(1-\delta_i^{(T)})(1-\delta_i^{(C)})}. \end{aligned}$$

Aplicando a relação de equivalência apresentada em (2.2), temos

$$L \propto \prod_{i=1}^n [h^{(T)}(y_i)]^{\delta_i^{(T)}} S^{(T)}(y_i) \\ \times [h^{(C)}(y_i)]^{\delta_i^{(C)}} S^{(C)}(y_i).$$

Ao assumir que a censura não é informativa, declaramos que a distribuição dos tempos de censura C não depende dos parâmetros $\boldsymbol{\theta}^{(T)}$ associados à distribuição dos tempos de falha T (Kalbfleisch & Prentice, 2011). Então, se a censura é não informativa, a função de verossimilhança resulta na expressão

$$L(\boldsymbol{\theta}^{(T)}) \propto \prod_{i=1}^n h^{(T)}(y_i | \boldsymbol{\theta}^{(T)})^{\delta_i^{(T)}} S^{(T)}(y_i | \boldsymbol{\theta}^{(T)}).$$

Assumir erroneamente que o mecanismo de censura é não informativo, quando na verdade é informativo, pode levar a graves vieses, conforme discutido em Leung *et al.* (1997). Portanto, quando a i -ésima observação é uma falha ($\delta_i^{(T)} = 1$, $\delta_i^{(C)} = 0$), a contribuição para a verossimilhança é dada por

$$P(y_i, \delta_i^{(T)} = 1, \delta_i^{(C)} = 0) = P(T \in (y_i, y_i + \Delta y_i], C > y_i, A > y_i).$$

De forma análoga, quando a i -ésima observação corresponde a uma censura informativa ($\delta_i^{(T)} = 0$, $\delta_i^{(C)} = 1$), a contribuição para a verossimilhança é dada por

$$P(y_i, \delta_i^{(T)} = 0, \delta_i^{(C)} = 1) = P(C \in (y_i, y_i + \Delta y_i], T > y_i, A > y_i).$$

Finalmente, quando a i -ésima observação equivale a uma censura administrativa ($\delta_i^{(T)} = 0$, $\delta_i^{(C)} = 0$), a contribuição para a verossimilhança é dada por

$$P(y_i, \delta_i^{(T)} = 0, \delta_i^{(C)} = 0) = P(A \in (y_i, y_i + \Delta y_i], T > y_i, C > y_i).$$

Supondo que a variável aleatória A é independente das variáveis aleatórias T e C , porém as variáveis T e C não são independentes, ou seja, $(T, C) \perp A$. Então, temos

$$P(T \in (y_i, y_i + \Delta y_i], C > y_i, A > y_i) = P(T \in (y_i, y_i + \Delta y_i], C > y_i) S^{(A)}(y_i | \boldsymbol{\theta}^{(A)}), \quad (3.1)$$

$$P(C \in (y_i, y_i + \Delta y_i], T > y_i, A > y_i) = P(C \in (y_i, y_i + \Delta y_i], T > y_i) S^{(A)}(y_i | \boldsymbol{\theta}^{(A)}), \quad (3.2)$$

$$P(A \in (y_i, y_i + \Delta y_i], T > y_i, C > y_i) = f^{(A)}(y_i | \boldsymbol{\theta}^{(A)}) P(T > y_i, C > y_i). \quad (3.3)$$

As probabilidades em (3.1), (3.2) e (3.3) são utilizadas para a obtenção da função de verossimilhança para uma amostra de n observações. Considerando o vetor de parâmetros dado por $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(T)}, \boldsymbol{\theta}^{(C)}, \boldsymbol{\theta}^{(A)})$, a função de verossimilhança é dada por

$$\begin{aligned} L(\boldsymbol{\theta}|D) &= \prod_{i=1}^n \left[P(T \in (y_i, y_i + \Delta y_i], C > y_i | \boldsymbol{\theta}) \right]^{S^{(A)}(y_i | \boldsymbol{\theta}^{(A)}) \delta_i^{(T)}} \\ &\quad \times \left[P(C \in (y_i, y_i + \Delta y_i], T > y_i | \boldsymbol{\theta}) \right]^{S^{(A)}(y_i | \boldsymbol{\theta}^{(A)}) \delta_i^{(C)}} \\ &\quad \times \left[f^{(A)}(y_i | \boldsymbol{\theta}^{(A)}) P(T > y_i, C > y_i) \right]^{(1-\delta_i^{(T)})(1-\delta_i^{(C)})}. \end{aligned}$$

Como A é independente de (T, C) , temos

$$\begin{aligned} L(\boldsymbol{\theta}|D) &\propto \prod_{i=1}^n \left[P(T \in (y_i, y_i + \Delta y_i], C > y_i | \boldsymbol{\theta}) \right]^{\delta_i^{(T)}} \\ &\quad \times \left[P(C \in (y_i, y_i + \Delta y_i], T > y_i | \boldsymbol{\theta}) \right]^{\delta_i^{(C)}} \\ &\quad \times \left[P(T > y_i, C > y_i) \right]^{(1-\delta_i^{(T)})(1-\delta_i^{(C)})}. \end{aligned}$$

Uma das maneiras de obter essa função densidade conjunta entre os tempos de falha e censura é assumir que esses tempos são condicionalmente independentes, ou seja, dado um efeito aleatório podemos assumir independência entre os tempos de falha T e censura C .

Suponha que existam m grupos com n_k indivíduos no k -ésimo grupo, sendo $N = \sum_{k=1}^m n_k$ indivíduos no total. Assuma que os indivíduos no k -ésimo grupo compartilham a mesma fragilidade não observável, denotada por W_k , introduzida na modelagem para captar a correlação entre os tempos associados aos indivíduos pertencentes ao mesmo grupo.

Vamos denotar por

$\boldsymbol{\theta} = \left(\boldsymbol{\beta}^{(T)'} , \boldsymbol{\beta}^{(C)'} , \alpha , \sigma^2 , h_0^{(T)}(\cdot) , h_0^{(C)}(\cdot) \right)'$, o vetor de parâmetros;

$\mathbf{d}_{ik} = \left(y_{ik} , \delta_{ik}^{(T)} , \delta_{ik}^{(C)} , x_{ik}^{(T)} , x_{ik}^{(C)} \right)'$, o vetor de dados observados para o i -ésimo indivíduo no k -ésimo grupo;

$\mathbf{d}_k = (\mathbf{d}_{k1}, \dots, \mathbf{d}_{kn_k})'$, o conjunto de dados observados no grupo k ;

$\mathbf{d} = (\mathbf{d}'_1, \dots, \mathbf{d}'_m)'$, o conjunto de dados observados.

3.2 Abordagem proposta por Huang & Wolfe (2002)

No modelo proposto por Huang & Wolfe (2002) as funções taxa de falha são expressas

como

$$h_{ik}^{(T)}(y|\mathbf{X}_{ik}^{(T)}, \mathbf{X}_{ik}^{(C)}, W_k, C_{ik}) = h_0^{(T)}(y) \exp\left(\boldsymbol{\beta}^{(T)'} \mathbf{X}_{ik}^{(T)} + W_k\right) \quad (3.4)$$

e

$$h_{ik}^{(C)}(y|\mathbf{X}_{ik}^{(T)}, \mathbf{X}_{ik}^{(C)}, W_k, T_{ik}) = h_0^{(C)}(y) \exp\left(\boldsymbol{\beta}^{(C)'} \mathbf{X}_{ik}^{(C)} + \alpha W_k\right), \quad (3.5)$$

em que $W_k \sim N(0, \sigma^2)$, $\mathbf{X}_{ik}^{(T)}$ e $\mathbf{X}_{ik}^{(C)}$ são os vetores de covariáveis associadas aos tempos de falha e de censura, com dimensões $p \times 1$ e $q \times 1$, respectivamente. As funções $h_0^{(T)}(y)$ e $h_0^{(C)}(y)$ denotam as funções taxa de falha de base para os tempos de falha e censura, respectivamente. Essas funções $h_0^{(T)}(y)$ e $h_0^{(C)}(y)$ podem ser modeladas de forma paramétrica ou não paramétrica. Condicional nas covariáveis e na fragilidade, os tempos de falha e censura são independentes. A função taxa de falha para os tempos de censura administrativa, A , não precisa ser especificada, pois estes tempos são independentes dos tempos de falha e censura informativa.

O parâmetro α , $\alpha \in \mathbb{R}$, quantifica a dependência entre os tempos de falha e tempos de censura. Se $\alpha = 0$, a fragilidade permanece apenas nos tempos da falha, captando a associação somente entre os indivíduos que estão no mesmo grupo. Nesse caso, a fragilidade não afeta a censura e, portanto, a censura informativa não precisa ser considerada no modelo, pois ela não afeta a inferência sobre os parâmetros $\boldsymbol{\beta}^{(T)}$ e σ^2 .

Por outro lado, se $\alpha \neq 0$ a fragilidade é comum aos tempos de falha e tempos de censura, indicando a dependência entre ambos. Se $\alpha > 0$, quanto maior a fragilidade, mais rápido a censura tenderá a ocorrer. Se $\alpha < 0$, quanto maior a fragilidade, maior a probabilidade da falha ocorrer antes da censura.

A abordagem proposta por Huang & Wolfe (2002) baseia-se em dois pressupostos: os indivíduos no mesmo grupo compartilham uma fragilidade comum; dentro de cada grupo, a censura é independente da falha. Esses pressupostos são comuns nos modelos de fragilidade com censura não informativa. Porém, a abordagem de Huang & Wolfe (2002) acrescenta uma nova característica que permite uma associação entre as funções taxa de falha dos tempos de falha e as funções taxa de falha dos tempos de censura, de modo que a censura possa ser informativa para a falha.

No modelo de Huang & Wolfe (2002), a dependência marginal entre os tempos de falha T e os tempos de censura informativa C é apenas devido à fragilidade, W_k , que afeta tanto a falha quanto a censura. Além disso, a função taxa de falha dos tempos censurados para um indivíduo com fragilidade e^{W_k} é afetada por $(e^{W_k})^\alpha$. A correlação entre $\log(T)$ e $\log(C)$ introduzida por este modelo não tem forma fechada (Huang & Wolfe, 2002). Para ilustrar a relação entre essa correlação e α realizamos um estudo de simulação, em que utilizamos $h_0^{(T)}(y) = h_0^{(C)}(y) = 0, 1$, sem considerar covariáveis e α assumindo valores no

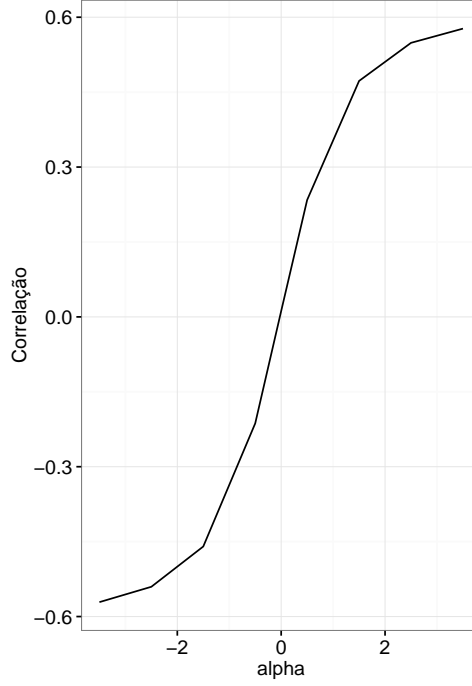


Figura 3.1: Relação entre α e a correlação entre $\log(T)$ e $\log(C)$ quando as funções taxa de falha para T e C são $h^{(T)}(y) = 0, 1 \exp(W)$ e $h^{(C)}(y) = 0, 1 \exp(\alpha W)$, respectivamente.

intervalo de $-3,0$ a $3,0$ em (3.4) e (3.5). Os resultados dessa relação são apresentados na Figura 3.1. Observamos que a correlação entre $\log(T)$ e $\log(C)$ tem o mesmo sinal que α e que ela aumenta a medida que $|\alpha|$ aumenta.

A função de verossimilhança condicional nas fragilidades para o i -ésimo indivíduo no k -ésimo grupo é dada por

$$\begin{aligned}
 L_{ik}(\mathbf{d}_{ik}; \boldsymbol{\theta}, w_k) \propto & \left[h_0^{(T)}(y_{ik}) \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right]^{\delta_{ik}^{(T)}} \\
 & \times \exp \left\{ -H_0^{(T)}(y_{ik}) \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right\} \\
 & \times \left[h_0^{(C)}(y_{ik}) \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right]^{\delta_{ik}^{(C)}} \\
 & \times \exp \left\{ -H_0^{(C)}(y_{ik}) \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right\}.
 \end{aligned} \tag{3.6}$$

Huang & Wolfe (2002) realizaram a estimação dos coeficientes de regressão $\boldsymbol{\beta}^{(T)'}$ e $\boldsymbol{\beta}^{(C)'}$, do parâmetro α e da variância do efeito aleatório, por meio do método de máxima verossimilhança. Para tanto, consideraram a função de verossimilhança marginal para o k -ésimo grupo, dada por

$$P_k(\mathbf{d}_k; \boldsymbol{\theta}) = \int_{-\infty}^{\infty} f(w_k; \sigma^2) \prod_{i=1}^{n_k} L_{ik}(\mathbf{d}_{ik}; \boldsymbol{\theta}, w_k) dw_k, \tag{3.7}$$

em que $f(\cdot; \sigma^2)$ é a função de densidade da distribuição Normal(0, σ^2) e $L_{ik}(\mathbf{d}_{ik}; \boldsymbol{\theta}, w_k)$ é a função de verossimilhança condicional dada em (3.6).

Para obter os estimadores de máxima verossimilhança é necessário maximizar a função de verossimilhança marginal, obtida pelo produto de (3.7) para todos os grupos $k = 1, \dots, m$. Conforme apontado por Huang & Wolfe (2002), se os efeitos aleatórios w_k fossem conhecidos, então poderíamos ajustar (3.4) e (3.5) separadamente. Desta forma, o algoritmo EM proposto por Dempster *et al.* (1977) surge como uma alternativa muito atraente para a estimação dos parâmetros, onde a variável latente é a fragilidade não observada.

No passo E do algoritmo EM é calculada a esperança do logaritmo da função de verossimilhança dos dados aumentados, condicional nos dados observados e nos valores atuais do vetor de parâmetros $\boldsymbol{\theta}$. Portanto,

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = E \left[\log (L(\boldsymbol{\theta}; \mathbf{w}, \mathbf{d})) | \mathbf{d}, \boldsymbol{\theta}^{(r)} \right], \quad (3.8)$$

em que $\boldsymbol{\theta}^{(r)}$ corresponde ao vetor de valores que maximizaram a função $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r-1)})$ no passo r , e a função de verossimilhança dos dados aumentados $L(\boldsymbol{\theta}; \mathbf{w}, \mathbf{d})$ é dada por

$$L(\boldsymbol{\theta}; \mathbf{w}, \mathbf{d}) = \prod_{k=1}^m L_k(\boldsymbol{\theta}; w_k, \mathbf{d}_k) f_{W_k|\mathbf{d}_k}(w_k; \boldsymbol{\theta}), \quad (3.9)$$

com $L_k(\boldsymbol{\theta}; w_k, \mathbf{d}_k) = \prod_{i=1}^{n_k} L_{ik}(\boldsymbol{\theta}; w_k, d_{ik})$.

A função densidade de probabilidade das fragilidades W_k condicional nos dados observados $f_{W_k|\mathbf{d}_k}(w_k; \boldsymbol{\theta})$, utilizada em (3.9) é dada por

$$f_{W_k|\mathbf{d}_k}(w_k; \boldsymbol{\theta}) = \frac{f(w_k; \sigma^2) \prod_{i=1}^{n_k} L_{ik}(\mathbf{d}_{ik}; \boldsymbol{\theta}, w_k)}{P_k(\mathbf{d}_k; \boldsymbol{\theta})}. \quad (3.10)$$

Como a integral em (3.7) não tem forma fechada, torna-se muito difícil maximizar a função de verossimilhança marginal diretamente, e o valor esperado em (3.8) também não apresenta expressões analíticas conhecidas. Uma possível solução é utilizarmos um estimador de Monte Carlo para estimarmos $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$.

Dada uma amostra dos efeitos aleatórios na iteração $r + 1$, $w_k^{(r+1,1)}, \dots, w_k^{(r+1,L)}$, da distribuição $f_{W_k|\mathbf{d}_k}(w_k; \boldsymbol{\theta}^{(r)})$, gerada pelo algoritmo de Metropolis-Hastings (Metropolis *et al.*, 1953) e (Hastings, 1970), podemos estimar a esperança em (3.8) por meio da

aproximação

$$Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = \frac{1}{S} \sum_{s=1}^S \log(L(\boldsymbol{\theta}; \mathbf{w}^{(r+1,s)}, \mathbf{d})). \quad (3.11)$$

Esse tipo de aproximação foi proposto por Wei & Tanner (1990) e, nesses casos, o algoritmo passa a ser chamado de algoritmo EM Monte Carlo (EMMC).

Portanto, considerando a função de verossimilhança dos dados aumentados dada em (3.9), no cálculo da esperança (3.11) temos

$$\begin{aligned} Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = & \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\log(h_0^{(T)}(y_{ik})) + \boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + \hat{E}(w_k|\mathbf{d}_k) \right] \right. \\ & - H_0^{(T)}(y_{ik}) \exp\left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}\right) \hat{E}(\exp(w_k)|\mathbf{d}_k) \\ & + \delta_{ik}^{(C)} \left[\log(h_0^{(C)}(y_{ik})) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha \hat{E}(w_k|\mathbf{d}_k) \right] \\ & \left. - H_0^{(C)}(y_{ik}) \exp\left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}\right) \hat{E}(\exp(\alpha w_k)|\mathbf{d}_k) \right\} \\ & - \frac{1}{2} [\log(2\pi) - \log(\sigma^2)] - \frac{1}{2\sigma^2} \hat{E}(w_k^2|\mathbf{d}_k), \end{aligned} \quad (3.12)$$

em que as esperanças $\hat{E}(w_k|\mathbf{d}_k)$, $\hat{E}(\exp(w_k)|\mathbf{d}_k)$ e $\hat{E}(\exp(\alpha w_k)|\mathbf{d}_k)$ são calculadas numericamente a partir de amostras dos efeitos aleatórios w_k , de tamanho S , geradas da função densidade de probabilidade condicional (3.10). Dessa forma, obtêm-se uma estimativa para o valor esperado de funções dos efeitos aleatórios da seguinte forma,

$$\begin{aligned} \hat{E}(w_k|\mathbf{d}_k) &= \frac{1}{S} \sum_{s=1}^S w_k^{(s)}, \\ \hat{E}(\exp(w_k)|\mathbf{d}_k) &= \frac{1}{S} \sum_{s=1}^S \exp\left(w_k^{(s)}\right), \\ \hat{E}(\exp(\alpha w_k)|\mathbf{d}_k) &= \frac{1}{S} \sum_{s=1}^S \exp\left(\alpha w_k^{(s)}\right), \\ \hat{E}(w_k \exp(\alpha w_k)|\mathbf{d}_k) &= \frac{1}{S} \sum_{s=1}^S w_k^{(s)} \exp\left(\alpha w_k^{(s)}\right) \end{aligned} \quad (3.13)$$

e

$$\hat{E}(w_k^2|\mathbf{d}_k) = \frac{1}{S} \sum_{s=1}^S \left(w_k^{(s)}\right)^2.$$

No passo M, obtêm-se as estimativas para os parâmetros $\boldsymbol{\theta}$ que maximizam a função $Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$. Para maximizar essa função $Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ aplica-se o método da verossimilhança perfilada, vide por exemplo Vaida *et al.* (2000), que apresentaram um modelo com efeito

aleatório e censura não informativa. Conforme Huang & Wolfe (2002), para valores fixos de $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$ e α , as estimativas $\hat{h}_0^{(T)}(\cdot)$ e $\hat{h}_0^{(C)}(\cdot)$ que maximizam $Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ são dados por

$$\hat{h}_0^{(T)}(y) = \frac{\delta_{ik}^{(T)}}{\sum_{t_{jl} \geq t_{ik}} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{jl}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k)}$$

e

$$\hat{h}_0^{(C)}(y) = \frac{\delta_{ik}^{(C)}}{\sum_{t_{jl} \geq t_{ik}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{jl}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k)}. \quad (3.14)$$

que correspondem às estimativas das funções taxa de falha de base propostas por Breslow (1974). A partir de (3.14) obtêm-se as estimativas para $\hat{H}_0^{(T)}(\cdot)$ e $\hat{H}_0^{(C)}(\cdot)$ substituindo essas funções em $Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ e derivando com respeito aos parâmetros $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α , σ^2 , encontram-se as seguintes funções escore:

$$\begin{aligned} U(\boldsymbol{\beta}^{(T)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(T)} \left[\delta_{ik}^{(T)} - \hat{H}_0^{(T)}(y_{ik}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k) \right] \right\}, \\ U(\boldsymbol{\beta}^{(C)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(C)} \left[\delta_{ik}^{(C)} - \hat{H}_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right] \right\}, \\ U(\alpha) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \hat{E}(w_k | \mathbf{d}_k) - \hat{H}_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(w_k \exp(\alpha w_k) | \mathbf{d}_k) \right\} \end{aligned}$$

e

$$U(\sigma^2) = \sum_{i=1}^{n_k} \left\{ \frac{1}{2(\sigma^2)^2} \hat{E}(w_k^2 | \mathbf{d}_k) - \frac{1}{2\sigma^2} \right\}. \quad (3.15)$$

As derivadas parciais de primeira e segunda ordem da função de verossimilhança, para esse modelo e para os demais modelos propostos nesta tese, encontram-se no Apêndice A.

Para obter as estimativas de $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α , e σ^2 a função multirroot do pacote root-Solve (Soetaert, 2016) pode ser utilizada. Tendo essas estimativas de $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α , e σ^2 , atualizam-se as estimativas das funções de taxa falha base, até obter convergência. Esse método de estimação para modelos de fragilidade com censura informativa, Equações (3.7) - (3.13), foi apresentado por Huang & Wolfe (2002).

Segundo Vaida *et al.* (2000), um bom ponto de partida (chute inicial) para $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, $h_0^{(T)}(\cdot)$ e $h_0^{(C)}(\cdot)$ é dado pelo ajuste da regressão de Cox usual sem considerar a fragilidade no ajuste. A convergência da log-verossimilhança é apenas aproximadamente monótona, devido às aproximações no passo E (Vaida *et al.*, 2000).

No cálculo da matriz de informação de Fisher precisamos derivar a função log-verossimilhança calculada com respeito aos dados observados. Porém, quando utilizamos o algoritmo EM, temos somente função de verossimilhança conjunta dos dados incompletos (observados) e efeitos aleatórios (dados não observados). Para solucionar esse problema, Louis (1982) demonstrou que a matriz de informação de Fisher pode ser aproximada pela log-verossimilhança calculada com respeito aos dados pseudo-completos (dados observados e dados não observados).

Murphy (1994) e Murphy (1995) demonstraram a consistência e normalidade assintótica para o modelo de fragilidade gama sem covariáveis. Posteriormente, Parner *et al.* (1998) estendeu estes resultados para modelo de fragilidade gama com covariáveis, mostrando a consistência dos estimadores e a variância assintótica. Vaida *et al.* (2000) apresentou a matriz de informação de Fisher aproximada, a partir do método proposto por Louis (1982), para os modelos de fragilidade, dada por

$$I(\theta) \approx E[-l''(\theta; \mathbf{d}, \mathbf{w})|\mathbf{d}, \theta] - E[s(\theta; \mathbf{d}, \mathbf{w})s(\theta; \mathbf{d}, \mathbf{w})'|\mathbf{d}, \theta],$$

em que $l''(\theta; \mathbf{d}, \mathbf{w})$ é a matriz de derivadas parciais de segunda ordem de $\log(L(\theta; \mathbf{d}, \mathbf{w}))$ e $s(\theta; \mathbf{d}, \mathbf{w})$ é o vetor de derivadas parciais de primeira ordem de $\log(L(\theta; \mathbf{d}, \mathbf{w}))$ com respeito a θ .

O cálculo do primeiro termo de $I(\theta)$ é baseado na esperança do tipo $E[h(\mathbf{w}_k)|\mathbf{d}, \hat{\theta}]$, similar ao passo E do algoritmo EM Monte Carlo. Para o segundo termo de $I(\theta)$ calcula-se a matriz $s(\theta; \mathbf{d}, \mathbf{w})s(\theta; \mathbf{d}, \mathbf{w})'$ para cada \mathbf{w} simulado de $p(\mathbf{w}|\mathbf{d}, \hat{\theta})$, então faz-se a esperança da matriz. Portanto, por Parner *et al.* (1998), $\hat{\theta}$ é consistente para θ e satisfaz a normalidade assintótica

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, I(\theta)^{-1}),$$

quando $n \rightarrow \infty$.

Os resultados assintóticos de Parner *et al.* (1998) foram demonstrado para modelos de fragilidade com censura não informativa. Porém, Huang & Wolfe (2002) comentaram que essas demonstrações precisam ser modificadas para modelos de fragilidade com censura informativa e que os resultados de simulação apresentados em Huang & Wolfe (2002) apontaram para a validade da teoria assintótica do modelo por eles proposto.

3.3 Modelos propostos para censura informativa

Nesta seção vamos propor modelos para dados de sobrevivência multivariados com

censura à direita e informativa, sem fração de cura, considerando a abordagem condicional proposta por Huang & Wolfe (2002).

3.3.1 Modelo Weibull

Esta seção apresenta uma versão completamente paramétrica do modelo proposto por Huang & Wolfe (2002) assumindo que os tempos de falha e censura seguem distribuições Weibull.

A distribuição Weibull foi apresentada inicialmente no trabalho de Fisher & Tippett (1928), porém Weibull (1939) mostrou que essa distribuição poderia ser utilizada em conjuntos de dados com diferentes características. De acordo com Lawless (2011), o modelo Weibull provavelmente é o modelo mais utilizado para ajustar dados de sobrevivência. O modelo Weibull possui apenas dois parâmetros, e é capaz de acomodar funções taxa de falha estritamente crescentes, decrescentes e constantes.

Se os tempos de falha e censura, T e C , têm distribuição Weibull, com parâmetros de forma denotados por $\kappa^{(T)}$ e $\kappa^{(C)}$, respectivamente, e escala denotados por $\gamma^{(T)}$ e $\gamma^{(C)}$, respectivamente, com funções de taxa de falha base dadas por

$$h_0^{(T)}(y) = \kappa^{(T)} y^{\kappa^{(T)}-1} \gamma^{(T)}$$

e

$$h_0^{(C)}(y) = \kappa^{(C)} y^{\kappa^{(C)}-1} \gamma^{(C)}.$$

Logo, incluindo as covariáveis e os efeitos aleatórios, a função de verossimilhança condicional nos efeitos aleatórios, para o i -ésimo indivíduo no k -ésimo grupo, é dada por

$$\begin{aligned} L_{ik}(\boldsymbol{\theta}; w_k, \mathbf{d}_k) &\propto \left[\kappa^{(T)} y_i^{\kappa^{(T)}-1} \gamma^{(T)} \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right]^{\delta_{ik}^{(T)}} \\ &\times \exp \left\{ -y_i^{\kappa^{(T)}} \gamma^{(T)} \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right\} \\ &\times \left[\kappa^{(C)} y_i^{\kappa^{(C)}-1} \gamma^{(C)} \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right]^{\delta_{ik}^{(C)}} \\ &\times \exp \left\{ -y_i^{\kappa^{(C)}} \gamma^{(C)} \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right\}. \end{aligned} \quad (3.16)$$

3.3.1.1 Abordagem frequentista

Considerando a função de verossimilhança condicional (3.16) e a função de densidade dos efeitos aleatórios no cálculo da esperança (3.11), temos

$$\begin{aligned}
Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)}) = & \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\log(\kappa^{(T)}) + (\kappa^{(T)} - 1) \log(y_{ik}) + \log(\gamma^{(T)}) \right. \right. \\
& + \boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + \hat{E}(w_k|\mathbf{d}_k) \left. \right] - y_{ik}^{\kappa^{(T)}} \gamma^{(T)} \exp\left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}\right) \hat{E}(\exp(w_k)|\mathbf{d}_k) \\
& + \delta_{ik}^{(C)} \left[\log(\kappa^{(C)}) + (\kappa^{(C)} - 1) \log(y_{ik}) + \log(\gamma^{(C)}) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} \right. \\
& + \alpha \hat{E}(w_k|\mathbf{d}_k) \left. \right] - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp\left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}\right) \hat{E}(\exp(\alpha w_k)|\mathbf{d}_k) \left. \right\} \\
& - \frac{1}{2} [\log(2\pi) - \log(\sigma^2)] - \frac{1}{2\sigma^2} \hat{E}(w_k^2|\mathbf{d}_k).
\end{aligned} \tag{3.17}$$

Derivando $Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ dada em (3.17) com respeito aos parâmetros $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α , σ^2 , $\kappa^{(T)}$, $\gamma^{(T)}$, $\kappa^{(C)}$, $\gamma^{(C)}$, encontramos as respectivas funções escore. Novamente, as funções $U(\boldsymbol{\beta}^{(T)})$, $U(\boldsymbol{\beta}^{(C)})$, $U(\alpha)$, $U(\sigma^2)$ são obtidas em (3.15).

$$\begin{aligned}
U(\kappa^{(T)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\kappa^{(T)}} + \log(y_{ik}) \right] \right. \\
&\quad \left. - y_{ik}^{\kappa^{(T)}} \log(y_{ik}) \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k)|\mathbf{d}_k) \right\}, \\
U(\gamma^{(T)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\gamma^{(T)}} \right] - y_{ik}^{\kappa^{(T)}} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k)|\mathbf{d}_k) \right\}, \\
U(\kappa^{(C)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\kappa^{(C)}} + \log(y_{ik}) \right] \right. \\
&\quad \left. - y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k)|\mathbf{d}_k) \right\}
\end{aligned}$$

e

$$U(\gamma^{(C)}) = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\gamma^{(C)}} \right] - y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k)|\mathbf{d}_k) \right\}.$$

As funções escore $U(\kappa^{(T)})$, $U(\gamma^{(T)})$, $U(\kappa^{(C)})$ e $U(\gamma^{(C)})$ não possuem solução analítica, portanto, para estimar esses parâmetros torna-se necessário a utilização de métodos numéricos. Uma das maneiras para encontrar essas estimativas pode ser feita pelo *software* R Core Team (2017), particularmente, por intermédio da função `multroot` do pacote `rootSolve` (Soetaert, 2016).

3.3.1.2 Abordagem bayesiana

Nos modelos propostos neste trabalho iremos assumir as seguintes especificações para

as distribuições *a priori*:

$$\begin{aligned}\tau &\sim \text{Gama}(a_\tau, b_\tau), \\ \beta_p^{(T)} &\sim \text{Normal}(0, \tau_\beta), \quad p = 1, \dots, P, \\ \beta_q^{(C)} &\sim \text{Normal}(0, \tau_\beta), \quad q = 1, \dots, Q\end{aligned}$$

e

$$\alpha \sim \text{Normal}(0, \tau_\alpha),$$

em que τ_β e τ_α denotam os parâmetros de precisão da distribuição. Os parâmetros τ , $\beta^{(T)}$, $\beta^{(C)}$ e α são considerados como sendo independentes.

Os métodos de geração conhecidos como Monte Carlo via cadeias de Markov (MCMC) podem ser utilizados para obter amostras das distribuições condicionais *a posteriori*. Detalhes destes métodos podem ser encontrados em Gamerman & Lopes (2006), Brooks *et al.* (2011), entre outros. Um dos procedimentos MCMC que pode ser utilizado é o amostrador de Gibbs, proposto por Geman & Geman (1984) para reconstrução de imagens, baseado no conhecimento das distribuições condicionais completas, e difundido no meio da inferência estatística por Gelfand & Smith (1990). Nos casos em que essas distribuições não apresentam forma conhecida, um dos métodos utilizados pode ser o algoritmo conhecido como Metropolis-Hastings. Esse algoritmo foi inicialmente proposto por Metropolis *et al.* (1953) e generalizado por Hastings (1970).

Apresentamos na sequência a versão bayesiana para o modelo Weibull proposto na Seção 3.3.1.1. Se os tempos de falha e censura, T e C , têm distribuição Weibull, com parâmetros de forma denotados por $\kappa^{(T)}$ e $\kappa^{(C)}$, e escala denotados por $\gamma^{(T)}$ e $\gamma^{(C)}$, temos a função de verossimilhança condicional apresentada em (3.16).

Para a especificação *a priori* dos parâmetros utilizamos as seguintes distribuições:

$$\begin{aligned}\kappa^{(T)} &\sim \text{Gama}(a_\kappa^{(T)}, b_\kappa^{(T)}), \\ \kappa^{(C)} &\sim \text{Gama}(a_\kappa^{(C)}, b_\kappa^{(C)}), \\ \gamma^{(T)} &\sim \text{Gama}(a_\gamma^{(T)}, b_\gamma^{(T)})\end{aligned}$$

e

$$\gamma^{(C)} \sim \text{Gama}(a_\gamma^{(C)}, b_\gamma^{(C)}).$$

em que os parâmetros $\kappa^{(T)}$, $\kappa^{(C)}$, $\gamma^{(T)}$ e $\gamma^{(C)}$ são assumidos independentes *a priori*.

Consequentemente, precisamos obter a distribuição *a posteriori* do vetor $\theta = (\beta^{(T)}, \beta^{(C)}, \alpha, \tau, \kappa^{(T)}, \kappa^{(C)}, \gamma^{(T)}, \gamma^{(C)}, \mathbf{W})$. Assumimos que os parâmetros são independentes *a priori*, ou seja, $\pi(\theta) = \pi(\beta^{(T)})\pi(\beta^{(C)})\pi(\alpha)\pi(\tau)\pi(\kappa^{(T)})\pi(\kappa^{(C)})\pi(\gamma^{(T)})\pi(\gamma^{(C)})$.

As distribuições condicionais completas para os parâmetros e para os efeitos aleatórios são dadas por

$$\begin{aligned} \pi(\beta_p^{(T)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \exp\left(\delta_{ik}^{(T)} \beta_p^{(T)} x_{ik}^{(T)} - y_{ik}^{\kappa^{(T)}} \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k)\right) \\ &\quad \times \exp\left(\frac{\tau_\beta (\beta_p^{(T)})^2}{2}\right), \end{aligned}$$

$$\begin{aligned} \pi(\beta_q^{(C)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \exp\left(\delta_{ik}^{(C)} \beta_q^{(C)} x_{ik}^{(C)} - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k)\right) \\ &\quad \times \exp\left(-\frac{\tau_\beta (\beta_q^{(C)})^2}{2}\right), \end{aligned}$$

$$\begin{aligned} \pi(\alpha | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \exp\left(\delta_{ik}^{(C)} \alpha w_k - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k)\right) \\ &\quad \times \exp\left(-\frac{\tau_\alpha \alpha^2}{2}\right), \end{aligned}$$

$$\begin{aligned} \pi(\kappa^{(T)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \left\{ \left[\kappa^{(T)} y_{ik}^{\kappa^{(T)}} \right]^{\delta_{ik}^{(T)}} \exp\left(-y_{ik}^{\kappa^{(T)}} \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k)\right) \right\} \\ &\quad \times (\kappa^{(T)})^{a_\kappa^{(T)}} \exp(-\kappa^{(T)} b_\kappa^{(T)}), \end{aligned}$$

$$\begin{aligned} \pi(\kappa^{(C)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \left\{ \left[\kappa^{(C)} y_{ik}^{\kappa^{(C)}} \right]^{\delta_{ik}^{(C)}} \exp\left(-y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k)\right) \right\} \\ &\quad \times (\kappa^{(C)})^{a_\kappa^{(C)}} \exp(-\kappa^{(C)} b_\kappa^{(C)}), \end{aligned}$$

$$(\gamma^{(T)} | \dots) \sim \text{Gama}\left(a_\gamma^{(T)} + \sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ik}^{(T)}, b_\gamma^{(T)} + \sum_{k=1}^m \sum_{i=1}^{n_k} y_{ik}^{\kappa^{(T)}} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k)\right),$$

$$(\gamma^{(C)} | \dots) \sim \text{Gama}\left(a_\gamma^{(C)} + \sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ik}^{(C)}, b_\gamma^{(C)} + \sum_{k=1}^m \sum_{i=1}^{n_k} y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k)\right),$$

$$(\tau | \dots) \sim \text{Gama}\left(a_\tau + \frac{m}{2}, b_\tau + \frac{\sum_{k=1}^m w_k^2}{2}\right)$$

e

$$\begin{aligned} \pi(w_k | \dots) \propto & \prod_{i=1}^{n_k} \left\{ \exp\left(\delta_{ik}^{(T)} w_k - y_{ik}^{\kappa^{(T)}} \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k)\right) \right. \\ & \times \exp\left(\delta_{ik}^{(C)} \alpha w_k - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_{ik})\right) \left. \right\} \\ & \times \exp\left(-\frac{\tau w_k^2}{2}\right). \end{aligned}$$

As distribuições condicionais completas para os parâmetros $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α , $\kappa^{(T)}$ e $\kappa^{(C)}$ e para as fragilidades \mathbf{w} são log-côncavas. Dessa forma, um dos métodos MCMC que podem ser utilizados é o método da rejeição adaptativa proposto por Gilks *et al.* (1995).

3.3.2 Modelo Exponencial por Partes

Nesta seção propomos um modelo para dados de sobrevivência agrupados com censura informativa, em que os tempos de falha e censura são ajustados pelo modelo Exponencial por Partes, e a estrutura de dependência é obtida por meio da abordagem condicional do modelo de fragilidade, proposto por Huang & Wolfe (2002). Dessa maneira, o modelo proposto acomoda o efeito aleatório, fornece a dependência entre os tempos de falha e censura, é paramétrico, porém permanece flexível para ajustar qualquer função taxa de falha basal.

Seja $\rho^{(T)}$ a grade de tempos que particiona o eixo dos tempos de falha em b intervalos, sendo a partição finita do eixo do tempo da falha T dada por $0 < s_1 < s_2 < \dots < s_b < \infty$, com $s_b > y$, em que y é o tempo de falha observado, com b intervalos dados por $(0, s_1], (s_1, s_2], \dots, (s_{b-1}, s_b]$. De forma análoga, seja $\rho^{(C)}$ a grade dos tempos que particiona o eixo do tempo até a censura em d intervalos, sendo a partição finita do eixo do tempo de censura C dada por $0 < c_1 < c_2 < \dots < c_d < \infty$, com $c_d > y$, em que y é o tempo de censura observado, com d intervalos dados por $(0, c_1], (c_1, c_2], \dots, (c_{d-1}, c_d]$.

Assumindo que os tempos T e C têm distribuição Exponencial por Partes, as funções de taxa de falha de base para os tempos de falha e censura são dadas, respectivamente, por

$$h_0^{(T)}(y) = \lambda_j^{(T)}, \quad y \in I_j^{(T)} = (s_{j-1}, s_j], \quad j = 1, \dots, b$$

e

$$h_0^{(C)}(y) = \lambda_l^{(C)}, \quad y \in I_l^{(C)} = (c_{l-1}, c_l], \quad l = 1, \dots, d.$$

Portanto, incluindo as covariáveis e os efeitos aleatórios, a função de verossimilhança

condicional nos efeitos aleatórios, para o i -ésimo indivíduo no k -ésimo grupo, é dada por

$$\begin{aligned}
L_{ik}(\boldsymbol{\theta}; w_k, \mathbf{d}_k) &\propto \prod_{j=1}^b \left[\lambda_j^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right]^{\delta_{ikj}^{(T)}} \\
&\times \exp \left\{ - \sum_{j=1}^b \lambda_j^{(T)} (y_{ikj} - s_{j-1}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right\} \\
&\times \prod_{l=1}^d \left[\lambda_l^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right]^{\delta_{ikl}^{(C)}} \\
&\times \exp \left\{ - \sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\},
\end{aligned} \tag{3.18}$$

em que o tempo y_{ikj} é definido como na Seção 2.2.1.

3.3.2.1 Abordagem frequentista

Considerando a função de verossimilhança condicional (3.18) na obtenção da função de verossimilhança dos dados completos (3.9), e esta no cálculo da esperança (3.11), temos

$$\begin{aligned}
Q_S(\boldsymbol{\theta} | \boldsymbol{\theta}^{(r)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \sum_{j=1}^b \left\{ \delta_{ikj}^{(T)} \left[\log(\lambda_j^{(T)}) + \boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + \hat{E}(w_k | \mathbf{d}_k) \right] \right\} \right. \\
&- \sum_{j=1}^b \left\{ \lambda_j^{(T)} (y_{ikj} - s_{j-1}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k) \right\} \\
&+ \sum_{l=1}^d \left\{ \delta_{ikl}^{(C)} \left[\log(\lambda_l^{(C)}) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha \hat{E}(w_k | \mathbf{d}_k) \right] \right\} \\
&- \sum_{l=1}^d \left\{ \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right\} \left. \right\} \\
&- \frac{1}{2} [\log(2\pi) - \log(\sigma^2)] - \frac{1}{2\sigma^2} \hat{E}(w_k^2 | \mathbf{d}_k).
\end{aligned} \tag{3.19}$$

Derivando (3.19) com respeito aos parâmetros $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α , σ^2 , $\boldsymbol{\lambda}^{(T)}$, $\boldsymbol{\lambda}^{(C)}$, encontramos as respectivas funções escore. As funções escore $U(\boldsymbol{\beta}^{(T)})$, $U(\boldsymbol{\beta}^{(C)})$, $U(\alpha)$ e $U(\sigma^2)$ são calculadas como em (3.2), e as funções escore para $\boldsymbol{\lambda}^{(T)}$ e $\boldsymbol{\lambda}^{(C)}$ são dadas por

$$U(\lambda_j^{(T)}) = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \frac{\delta_{ikj}^{(T)}}{\lambda_j^{(T)}} - (y_{ikj} - s_{j-1}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k) \right\}, \quad j = 1, \dots, b$$

e

$$U\left(\lambda_l^{(C)}\right) = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \frac{\delta_{ikl}^{(C)}}{\lambda_l^{(C)}} - (y_{ikl} - c_{l-1}) \exp\left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}\right) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right\}, \quad l = 1, \dots, d.$$

Assim, as estimativas para as taxas são obtidas por

$$\hat{\lambda}_j^{(T)} = \frac{\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikj}^{(T)}}{\sum_{k=1}^m \sum_{i=1}^{n_k} (y_{ikj} - s_{j-1}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k)}, \quad j = 1, \dots, b$$

e

$$\hat{\lambda}_l^{(C)} = \frac{\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikl}^{(C)}}{\sum_{k=1}^m \sum_{i=1}^{n_k} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k)}, \quad l = 1, \dots, d.$$

3.3.2.2 Abordagem bayesiana

Nesta seção é apresentada uma abordagem bayesiana para o modelo de fragilidade em que a distribuição Exponencial por Partes é assumida para os tempos de falha e censura. Portanto, considerando a função de verossimilhança condicional apresentada em (3.18), com o objetivo de especificar o modelo, precisamos definir a distribuição *a priori* conjunta de $(\boldsymbol{\beta}^{(T)}, \boldsymbol{\beta}^{(C)}, \boldsymbol{\lambda}^{(T)}, \boldsymbol{\lambda}^{(C)}, \alpha, \tau)$, $\tau = 1/\sigma^2$. Isso pode ser feito de forma independente, como segue

$$\begin{aligned} \pi\left(\boldsymbol{\beta}^{(T)}, \boldsymbol{\lambda}^{(T)} | \rho^{(T)}\right) &= \pi\left(\boldsymbol{\beta}^{(T)}\right) \pi\left(\boldsymbol{\lambda}^{(T)} | \rho^{(T)}\right), \\ \pi\left(\boldsymbol{\beta}^{(C)}, \boldsymbol{\lambda}^{(C)} | \rho^{(C)}\right) &= \pi\left(\boldsymbol{\beta}^{(C)}\right) \pi\left(\boldsymbol{\lambda}^{(C)} | \rho^{(C)}\right). \\ \pi\left(\boldsymbol{\lambda}^{(T)} | \rho^{(T)}\right) &= \prod_{j=1}^b \pi\left(\lambda_j^{(T)}\right) \quad e \quad \pi\left(\boldsymbol{\lambda}^{(C)} | \rho^{(C)}\right) = \prod_{l=1}^d \pi\left(\lambda_l^{(C)}\right). \end{aligned}$$

Uma distribuição *a priori* comumente utilizada para a taxa de falha base $\boldsymbol{\lambda}$ é distribuição *a priori* independente Gama $\lambda_j \sim G(a_{\lambda_j}, b_{\lambda_j})$ para $j = 1, 2, \dots, b$, em que a_{λ_j} e b_{λ_j} são hiperparâmetros. Portanto, vamos assumir as seguintes distribuições *a priori*

$$\lambda_j^{(T)} \sim \text{Gama}\left(a_{\lambda_j}^{(T)}, b_{\lambda_j}^{(T)}\right), \quad \forall j = 1, \dots, b,$$

e

$$\lambda_l^{(C)} \sim \text{Gama} \left(a_{\lambda_l}^{(C)}, b_{\lambda_l}^{(C)} \right), \forall l = 1, \dots, d.$$

A suposição acima é equivalente a assumir que, dado a grade de tempos, as taxas de falha têm distribuição *a priori* independentes. A família de distribuições Gama tem sido amplamente utilizada como distribuição *a priori* para ajustar a taxa de falha do MEP, pelo fato de esta distribuição corresponder à distribuição conjugada para a distribuição Exponencial por Partes, facilitando assim o processo de inferência em relação ao MEP.

As distribuições condicionais completas para o parâmetros e para efeito aleatório são dadas por

$$\begin{aligned} \pi \left(\beta_p^{(T)} | \dots \right) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{j=1}^b \exp \left(\delta_{ikj}^{(T)} \beta_p^{(T)} x_{ik}^{(T)} - \lambda_j^{(T)} (y_{ikj} - s_{j-1}) \right. \\ &\quad \left. \times \exp(\beta^{(T)' \mathbf{x}_{ik}^{(T)} + w_k) \right) \exp \left(-\frac{\tau_\beta (\beta_p^{(T)})^2}{2} \right), \end{aligned}$$

$$\begin{aligned} \pi \left(\beta_q^{(C)} | \dots \right) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{l=1}^d \exp \left(\delta_{ikl}^{(C)} \beta_q^{(C)} x_{ik}^{(C)} - \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \right. \\ &\quad \left. \times \exp(\beta^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \exp \left(-\frac{\tau_\beta (\beta_q^{(C)})^2}{2} \right), \end{aligned}$$

$$\left(\lambda_j^{(T)} | \dots \right) \sim \text{Gama} \left(\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikj}^{(T)} + a_{\lambda_j}^{(T)}, \sum_{k=1}^m \sum_{i=1}^{n_k} (y_{ikj} - s_{j-1}) \exp(\beta^{(T)' \mathbf{x}_{ik}^{(T)} + w_k) + b_{\lambda_j}^{(T)} \right),$$

$$\left(\lambda_l^{(C)} | \dots \right) \sim \text{Gama} \left(\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikl}^{(C)} + a_{\lambda_l}^{(C)}, \sum_{k=1}^m \sum_{i=1}^{n_k} (y_{ikl} - c_{l-1}) \exp(\beta^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k) + b_{\lambda_l}^{(C)} \right),$$

$$\begin{aligned} \pi \left(\alpha | \dots \right) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{l=1}^d \exp \left(\delta_{ikl}^{(C)} \alpha w_k - \lambda_l^{(C)} (y_{ikl} - s_{l-1}) \exp(\beta^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \\ &\quad \times \exp \left(-\frac{\tau_\alpha \alpha^2}{2} \right), \end{aligned}$$

$$\left(\tau | \dots \right) \sim \text{Gama} \left(a_\tau + \frac{m}{2}, b_\tau + \frac{\sum_{k=1}^m w_k^2}{2} \right)$$

e

$$\begin{aligned} \pi(w_k | \dots) \propto & \prod_{i=1}^{n_k} \left\{ \prod_{j=1}^b \exp\left(\delta_{ikj}^{(T)} w_k - \lambda_j^{(T)} (y_{ikj} - s_{j-1}) \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)} + w_k)\right) \right. \\ & \times \left. \prod_{l=1}^d \exp\left(\delta_{ikl}^{(C)} \alpha w_k - \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k)\right) \right\} \\ & \times \exp\left(-\frac{\tau w_k^2}{2}\right). \end{aligned}$$

3.3.3 Versão bayesiana para o modelo de Huang & Wolfe (2002)

Nesta seção propomos uma extensão da abordagem apresentada por Clayton (1991) para modelos de fragilidade, acrescentando a dependência entre os tempos de falha e censura como proposto por Huang & Wolfe (2002).

Andersen & Gill (1982) consideraram o modelo de Cox (1972) como sendo um caso particular do processo de contagem proposto por Aalen (1978), e demonstraram as propriedades assintóticas para os coeficientes de regressão. Clayton (1991) inseriu o efeito aleatório na abordagem proposta por Aalen (1978) e apresentou a estimação sob enfoque bayesiano.

Seja $N_{ik}^{(T)}(y)$ o processo que conta o número de falhas do i -ésimo indivíduo no k -ésimo grupo no intervalo $[y, y + dy)$. De forma análoga, seja $N_{ik}^{(C)}(y)$ o processo que conta o número de censuras do i -ésimo indivíduo no k -ésimo grupo no intervalo $[y, y + dy)$.

Seja $I_{ik}^{(T)}(y)$ a intensidade associada ao processo aleatório $N_{ik}^{(T)}(y)$, e $I_{ik}^{(C)}(y)$ a intensidade associada ao processo $N_{ik}^{(C)}(y)$. De acordo com Clayton (1991), essas intensidades podem ser dadas pelos produtos

$$I_{ik}^{(T)}(y) = V_{ik}^{(T)}(y) h_{0_{ik}}^{(T)}(y) \exp\left(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)} + w_k\right)$$

e

$$I_{ik}^{(C)}(y) = V_{ik}^{(C)}(y) h_{0_{ik}}^{(C)}(y) \exp\left(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k\right),$$

em que $h_{0_{ik}}^{(T)}(y)$ equivale à função taxa de falha base para o i -ésimo indivíduo no k -ésimo grupo no tempo y , e $h_{0_{ik}}^{(C)}(y)$ equivale a função taxa de falha base para os tempos de censura y para o indivíduo i no grupo k . A variável indicadora $V_{ik}^{(T)}(y)$ denota o processo observado para os tempos de falha, em que $V_{ik}^{(T)}(y) = 1$ indica se o indivíduo está sob risco em y e $V_{ik}^{(T)}(y) = 0$ caso contrário. A variável indicadora $V_{ik}^{(C)}(y)$ denota o processo observado para os tempos de censura, em que $V_{ik}^{(C)}(y) = 1$ indica se o indivíduo está sob

risco em y e $V_{ik}^{(C)}(y) = 0$ caso contrário.

Seja $dN_{ik}^{(T)}(y)$ o incremento de $N_{ik}^{(T)}(y)$ no intervalo $[y, y + dy)$, com intensidade $I_{ik}^{(T)}(y)$. Portanto, $dN_{ik}^{(T)}(y) \sim \text{Poisson} \left(V_{ik}^{(T)}(y) h_{0_{ik}}^{(T)}(y) \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right)$. De forma análoga, seja $dN_{ik}^{(C)}(y)$ o incremento de $N_{ik}^{(C)}(y)$ no intervalo $[y, y + dy)$, com intensidade $I_{ik}^{(C)}(y)$. Logo, $dN_{ik}^{(C)}(y) \sim \text{Poisson} \left(V_{ik}^{(C)}(y) h_{0_{ik}}^{(C)}(y) \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right)$, mais detalhes em Clayton (1991) e Mostafa & Ghorbal (2011).

Os incrementos seguem um processo de contagem, com $dN_{ik}^{(T)}(y) = 1$ quando a falha é observada no intervalo $[y, y + dy)$, e $dN_{ik}^{(T)}(y) = 0$ caso contrário. Semelhantemente, $dN_{ik}^{(C)}(y) = 1$ quando uma censura é observada em $[y, y + dy)$, e $dN_{ik}^{(C)}(y) = 0$ caso contrário.

Considere os tempos de falha $j = 1, \dots, J$, com intervalos $[y_j^{(T)}, y_{j+1}^{(T)})$. Portanto, a contribuição do i -ésimo indivíduo no k -ésimo grupo, se for falha, para a função de verossimilhança é dada por

$$L_{ik}^{(T)} \left(\boldsymbol{\theta}^{(T)} | \mathbf{d}_{ik}, w_k \right) = \prod_{j=1}^J \left[V_{ik}^{(T)}(y_j^{(T)}) h_{0_{ik}}^{(T)}(y_j^{(T)}) \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right]^{dN_{ik}^{(T)}(y_j^{(T)})} \\ \times \exp \left\{ -V_{ik}^{(T)}(y_j^{(T)}) h_{0_{ik}}^{(T)}(y_j^{(T)}) \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right) \right\},$$

em que $\boldsymbol{\theta}^{(T)} = \left(\boldsymbol{\beta}^{(T)'}, \mathbf{h}_0^{(T)}(\cdot) \right)'$ e $\mathbf{h}_0^{(T)}(\cdot)$ é o vetor de taxas de falha base.

Considere os tempos de censura $l = 1, \dots, L$, com intervalos $[y_l^{(C)}, y_{l+1}^{(C)})$. Logo, a contribuição do i -ésimo indivíduo no k -ésimo grupo, sendo censura informativa, para a função de verossimilhança é dada por

$$L_{ik}^{(C)} \left(\boldsymbol{\theta}^{(C)} | \mathbf{d}_{ik}, w_k \right) = \prod_{l=1}^L \left[V_{ik}^{(C)}(y_l^{(C)}) h_{0_{ik}}^{(C)}(y_l^{(C)}) \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right]^{dN_{ik}^{(C)}(y_l^{(C)})} \\ \times \exp \left\{ -V_{ik}^{(C)}(y_l^{(C)}) h_{0_{ik}}^{(C)}(y_l^{(C)}) \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right) \right\},$$

em que $\boldsymbol{\theta}^{(C)'} = \left(\alpha, \boldsymbol{\beta}^{(C)'}, \mathbf{h}_0^{(C)}(\cdot) \right)'$ e $\mathbf{h}_0^{(C)}(\cdot)$ é o vetor de taxas de falha de base para os tempos de censura.

Denote $L_{ik}^{(T)} = L_{ik}^{(T)} \left(\boldsymbol{\beta}^{(T)}, \mathbf{h}_0^{(T)}(\cdot) | \mathbf{d}_{ik}, w_k \right)$ e $L_{ik}^{(C)} = L_{ik}^{(C)} \left(\alpha, \boldsymbol{\beta}^{(C)}, \mathbf{h}_0^{(C)}(\cdot) | \mathbf{d}_{ik}, w_k \right)$. Portanto, a função de verossimilhança é dada por

$$L \left(\alpha, \boldsymbol{\beta}^{(T)}, \boldsymbol{\beta}^{(C)}, \mathbf{h}_0^{(T)}(\cdot), \mathbf{h}_0^{(C)}(\cdot) | \mathbf{d}, \mathbf{w} \right) = \prod_{k=1}^m \prod_{i=1}^{n_k} L_{ik}^{(T)} L_{ik}^{(C)}.$$

Sejam $\mathbf{H}_0^{(T)}(\cdot)$ a função taxa de falha acumulada e $dH_0^{(T)}(y_j^{(T)}) = H_0^{(T)}(y_{j+1}^{(T)}) -$

$H_0^{(T)}(y_j^{(T)})$, o incremento da função taxa de falha acumulada para os tempos de falha no intervalos $[y_j^{(T)}, y_{j+1}^{(T)})$. Similarmente, sejam $\mathbf{H}_0^{(C)}(\cdot)$ a função taxa de falha acumulada para os tempos de censura e $dH_0^{(C)}(y_l^{(C)}) = H_0^{(C)}(y_{l+1}^{(C)}) - H_0^{(C)}(y_l^{(C)})$, o incremento da função taxa de falha acumulada para os tempos de censura em $[y_l^{(C)}, y_{l+1}^{(C)})$.

Assumimos processos gamas com incrementos independentes para modelar $\mathbf{H}_0^{(T)}(\cdot)$ e $\mathbf{H}_0^{(C)}(\cdot)$, dados por

$$dH_0^{(T)}(y_j^{(T)}) \sim \text{Gama} \left(c \left[H_*^{(T)}(y_{j+1}^{(T)}) - H_*^{(T)}(y_j^{(T)}) \right], c \right), \quad j = 1, \dots, J,$$

e

$$dH_0^{(C)}(y_l^{(C)}) \sim \text{Gama} \left(r \left[H_*^{(C)}(y_{l+1}^{(C)}) - H_*^{(C)}(y_l^{(C)}) \right], r \right), \quad l = 1, \dots, L,$$

em que $c > 0$ e $r > 0$ são constantes conhecidas e $H_*^{(T)}(\cdot)$ e $H_*^{(C)}(\cdot)$ são funções de taxas de falha acumuladas conhecidas.

Para completarmos as especificações do modelo, assumimos independência *a priori* entre os objetos aleatórios $(\boldsymbol{\beta}^{(T)}, \boldsymbol{\beta}^{(C)}, \alpha, \tau, \mathbf{H}_0^{(T)}(\cdot), \mathbf{H}_0^{(C)}(\cdot))$. Como a variável aleatória W_k é latente e, considerando a técnica de aumento de dados, também deve ser estimada. Portanto, é necessário obtermos as distribuições do vetor $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^{(T)}, \boldsymbol{\beta}^{(C)}, \alpha, \tau, \mathbf{H}_0^{(T)}(\cdot), \mathbf{H}_0^{(C)}(\cdot), \mathbf{W} \right)$.

A seguir apresentamos as distribuições condicionais completas (dcc) para os componentes de $\boldsymbol{\theta}$.

Para os incrementos na taxa de falha acumulada dos tempos de falha, as dcc são dada por

$$dH_0^{(T)}(y_j^{(T)} | \dots) \sim \text{Gama} \left(c \left[H_*^{(T)}(y_{j+1}^{(T)}) - H_*^{(T)}(y_j^{(T)}) \right] + dN^{(T)}(y_j^{(T)}), c + R_+^{(T)} \right),$$

em que $R_{ik}^{(T)} = V_{ik}^{(T)}(y_j^{(T)}) \exp \left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right)$ e $R_+^{(T)}$ é a soma de $R_{ik}^{(T)}$ para todos indivíduos com tempos maiores que do i -ésimo indivíduo.

Para os incrementos na taxa de falha acumulada dos tempos de censura, as dcc são dada por

$$dH_0^{(C)}(y_l^{(C)} | \dots) \sim \text{Gama} \left(r \left[H_*^{(C)}(y_{l+1}^{(C)}) - H_*^{(C)}(y_l^{(C)}) \right] + dN^{(C)}(y_l^{(C)}), r + R_+^{(C)} \right),$$

em que $R_{ik}^{(C)} = V_{ik}^{(C)}(y_l^{(C)}) \exp \left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right)$ e $R_+^{(C)}$ é a soma de $R_{ik}^{(C)}$ para todos indivíduos com tempos maiores que do i -ésimo indivíduo.

A distribuição condicional completa para $\beta^{(T)}$ é dada por

$$\pi(\beta_p^{(T)} | \dots) \propto \exp\left(-\frac{\tau_\beta}{2}(\beta_p^{(T)})^2\right) \prod_{k=1}^m \prod_{i=1}^{n_k} \exp\left(dN_{ik}^{(T)}(y_j^{(T)})\beta_p^{(T)}x_{ik}^{(T)} - V_{ik}^{(T)}(y_j^{(T)})h_{0_{ik}}^{(T)}(y_j^{(T)}) \exp\left(\beta^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k\right)\right).$$

De forma análoga, a distribuição condicional completa para $\beta^{(C)}$ é dada por

$$\pi(\beta_q^{(C)} | \dots) \propto \exp\left(-\frac{\tau_\beta}{2}(\beta_q^{(C)})^2\right) \prod_{k=1}^m \prod_{i=1}^{n_k} \exp\left(dN_{ik}^{(C)}(y_l^{(C)})\beta_q^{(C)}x_{ik}^{(C)} - V_{ik}^{(C)}(y_l^{(C)})h_{0_{ik}}^{(C)}(y_l^{(C)}) \exp\left(\beta^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k\right)\right).$$

As demais distribuições condicionais completas são dadas por

$$\pi(\alpha | \dots) \propto \exp\left(-\frac{\tau_\alpha(\alpha)^2}{2}\right) \prod_{k=1}^m \prod_{i=1}^{n_k} \exp\left(dN_{ik}^{(C)}(y_l^{(C)})\alpha w_k - V_{ik}^{(C)}(y_l^{(C)})h_{0_{ik}}^{(C)}(y_l^{(C)}) \exp\left(\beta^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k\right)\right).$$

$$(\tau | \dots) \sim \text{Gama}\left(a_\tau + \frac{m}{2}, b_\tau + \frac{\sum_{k=1}^m w_k^2}{2}\right)$$

e

$$\begin{aligned} \pi(w_k | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \left\{ \exp\left(dN_{ik}^{(T)}(y_j^{(T)})w_k - V_{ik}^{(T)}(y_j^{(T)})h_{0_{ik}}^{(T)}(y_j^{(T)}) \exp\left(\beta^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k\right)\right) \right. \\ &\quad \times \exp\left(dN_{ik}^{(C)}(y_l^{(C)})\alpha w_k - V_{ik}^{(C)}(y_l^{(C)})h_{0_{ik}}^{(C)}(y_l^{(C)}) \exp\left(\beta^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k\right)\right) \left. \right\} \\ &\quad \times \exp\left(-\frac{\tau w_k^2}{2}\right). \end{aligned}$$

3.4 Modelos propostos para censura informativa com fração de cura

Neste trabalho também propomos incluir fração de cura no modelo para censura informativa e considerar que os tempos de promoção e censura são independentes quando condicionados na fragilidade e nas covariáveis.

Seja N o número de causas competindo para a ocorrência do evento de interesse. Seja Z_i a variável aleatória que denota o tempo de ocorrência do evento de interesse devido à i -ésima causa, conhecido como tempo de promoção. Seja $T = \min\{Z_1, \dots, Z_N\}$ e $Y = \min\{T, C, A\}$, no qual C é uma variável aleatória que representa o tempo observável até a censura informativa e A denota o tempo observável até a censura administrativa. Consideramos que a função de verossimilhança para o i -ésimo indivíduo no k -ésimo grupo, condicional nos efeitos aleatórios, é dada por

$$L_{ik}(\boldsymbol{\theta}; \mathbf{d}_k, w_k) \propto \left[h_{pop}^{(T)}(y_{ik} | \boldsymbol{\theta}^{(T)}, w_k) \right]^{\delta_{ik}^{(T)}} S_{pop}^{(T)}(y_{ik} | \boldsymbol{\theta}^{(T)}, w_k) \times \left[h^{(C)}(y_{ik} | \boldsymbol{\theta}^{(C)}, w_k) \right]^{\delta_{ik}^{(C)}} S^{(C)}(y_{ik} | \boldsymbol{\theta}^{(C)}, w_k), \quad (3.20)$$

em que $S_{pop}^{(T)}(y_{ik} | \boldsymbol{\theta}^{(T)}, w_k)$ e $h_{pop}^{(T)}(y_{ik} | \boldsymbol{\theta}^{(T)}, w_k)$ são dadas em (2.12) e (2.13), respectivamente.

Substituindo $h_{pop}^{(T)}(y_{ik} | \boldsymbol{\theta}^{(T)}, w_k)$ e $S_{pop}^{(T)}(y_{ik} | \boldsymbol{\theta}^{(T)}, w_k)$ pelas equações equivalentes, temos a seguinte função de verossimilhança condicional:

$$L_{ik}(\boldsymbol{\theta}; \mathbf{d}_k, w_k) \propto \left[\boldsymbol{\xi}_{ik} f^{(Z)}(y_{ik} | \boldsymbol{\theta}^{(Z)}) \right]^{\delta_{ik}^{(T)}} \exp \left\{ -\boldsymbol{\xi}_{ik} F^{(Z)}(y_{ik} | \boldsymbol{\theta}^{(Z)}) \right\} \times \left[h^{(C)}(y_{ik} | \boldsymbol{\theta}^{(C)}, w_k) \right]^{\delta_{ik}^{(C)}} S^{(C)}(y_{ik} | \boldsymbol{\theta}^{(C)}, w_k),$$

em que $\boldsymbol{\xi}_{ik} = \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k)$.

Portanto, as funções taxa de falha são dadas por

$$h_{pop}^{(T)}(y_{ik}; \delta_{ik}^{(T)}, \mathbf{x}_{ik}^{(T)}, w_k) = \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) f^{(Z)}(y_{ik}) \quad (3.21)$$

e

$$h^{(C)}(y_{ik}; \delta_{ik}^{(C)}, \mathbf{x}_{ik}^{(C)}, w_k) = h_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k). \quad (3.22)$$

Conseqüentemente, substituindo (3.21) e (3.22) na função de verossimilhança (3.20) temos a seguinte função de verossimilhança condicional nos efeitos aleatórios:

$$L_{ik}(\boldsymbol{\theta}; \mathbf{d}_k, w_k) \propto \left[\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) f^{(Z)}(y_{ik}) \right]^{\delta_{ik}^{(T)}} \times \exp \left\{ -\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) [F^{(Z)}(y_{ik})] \right\} \times \left[h_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right]^{\delta_{ik}^{(C)}} \times \exp \left\{ -H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}. \quad (3.23)$$

3.4.1 Modelo de fração de cura Weibull

Nesta seção vamos propor um modelo para ajustar dados de sobrevivência com fração de cura e censura informativa, considerando a distribuição Weibull($\kappa^{(Z)}$, $\gamma^{(Z)}$), para ajustar os tempos de promoção, e a distribuição Weibull($\kappa^{(C)}$, $\gamma^{(C)}$) para modelar os tempos de censura. As funções taxa de falha de base são obtidas por

$$h^{(Z)}(y) = \kappa^{(Z)} y^{\kappa^{(Z)}-1} \gamma^{(Z)} \quad e \quad h_0^{(C)}(y) = \kappa^{(C)} y^{\kappa^{(C)}-1} \gamma^{(C)}.$$

Conseqüentemente, substituindo as funções taxa de falha base em (3.23) temos a seguinte função de verossimilhança condicional nos efeitos aleatórios, para o i -ésimo indivíduo no k -ésimo grupo,

$$\begin{aligned} L_{ik}(\boldsymbol{\theta}; \mathbf{d}_{ik}, w_k) \propto & \left[\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \kappa^{(Z)} y_{ik}^{\kappa^{(Z)}-1} \gamma^{(Z)} \exp\left(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}\right) \right]^{\delta_{ik}^{(T)}} \\ & \times \exp\left\{ -\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \left[1 - \exp\left(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}\right) \right] \right\} \\ & \times \left[\kappa^{(C)} y_{ik}^{\kappa^{(C)}-1} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right]^{\delta_{ik}^{(C)}} \\ & \times \exp\left\{ -y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}. \end{aligned} \quad (3.24)$$

3.4.1.1 Abordagem frequentista

Assim como no caso do modelo sem fração de cura, vamos utilizar o algoritmo EM Monte Carlo para realizar a estimação dos parâmetros. As derivadas parciais de primeira e segunda ordem da função de verossimilhança para os modelos de fração de cura com censura informativa encontram-se no Apêndice A.

Do mesmo modo que nos modelos sem fração de cura, considerando a função de verossimilhança condicional (3.24) no cálculo da função de verossimilhança dos dados aumentados (3.9), e esta, por sua vez, no cálculo da esperança (3.11), resultando em

$$\begin{aligned} Q_S(\boldsymbol{\theta} | \boldsymbol{\theta}^{(r)}) = & \sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ik}^{(T)} \left[\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + \hat{E}(w_k | \mathbf{d}_k) + \log(\kappa^{(Z)}) + (\kappa^{(Z)} - 1) \log(y_{ik}) \right. \\ & \left. + \log(\gamma^{(Z)}) - y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)} \right] - \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + \hat{E}(w_k | \mathbf{d}_k)) \\ & \times \left\{ 1 - \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \right\} \\ & + \delta_{ik}^{(C)} \left[\log(\kappa^{(C)}) + (\kappa^{(C)} - 1) \log(y_{ik}) + \log(\gamma^{(C)}) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha \hat{E}(w_k | \mathbf{d}_k) \right] \\ & - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha \hat{E}(w_k | \mathbf{d}_k)) \\ & - \frac{1}{2} \left[\log(2\pi) + \log(\sigma^2) \right] - \frac{1}{2\sigma^2} \hat{E}(w_k^2 | \mathbf{d}_k). \end{aligned}$$

Derivando $Q_S(\boldsymbol{\theta}|\boldsymbol{\theta}^{(r)})$ com respeito aos parâmetros $\boldsymbol{\beta}^{(T)}, \kappa^{(Z)}, \kappa^{(C)}, \boldsymbol{\beta}^{(C)}, \gamma^{(Z)}, \gamma^{(C)}, \alpha$ e σ^2 , encontramos as seguintes funções escore:

$$\begin{aligned}
U(\boldsymbol{\beta}^{(T)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(T)} \left[\delta_{ik}^{(T)} - \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(w_k | \mathbf{d}_k) (1 - \exp(-H^{(Z)}(y_{ik}))) \right] \right\}, \\
U(\boldsymbol{\beta}^{(C)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(C)} \left[\delta_{ik}^{(C)} - H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right] \right\}, \\
U(\alpha) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \hat{E}(w_k | \mathbf{d}_k) - H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(w_k \exp(\alpha w_k) | \mathbf{d}_k) \right\}, \\
U(\sigma^2) &= \sum_{i=1}^{n_k} \left\{ \frac{1}{2\sigma^4} \hat{E}(w_k^2 | \mathbf{d}_k) - \frac{1}{2\sigma^2} \right\},
\end{aligned} \tag{3.25}$$

$$\begin{aligned}
U(\kappa^{(Z)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\kappa^{(Z)}} + \log(y_{ik}) - y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \gamma^{(Z)} \right] \right. \\
&\quad \left. + \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k) \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \left(-y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \gamma^{(Z)} \right) \right\},
\end{aligned}$$

$$\begin{aligned}
U(\gamma^{(Z)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\gamma^{(Z)}} - y_{ik}^{\kappa^{(Z)}} \right] + \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(\exp(w_k) | \mathbf{d}_k) \right. \\
&\quad \left. \times \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \left(-y_{ik}^{\kappa^{(Z)}} \right) \right\},
\end{aligned}$$

$$\begin{aligned}
U(\kappa^{(C)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\kappa^{(C)}} + \log(y_{ik}) \right] - y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \right. \\
&\quad \left. \times \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right\}
\end{aligned}$$

e

$$U(\gamma^{(C)}) = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\gamma^{(C)}} \right] - y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right\}.$$

3.4.1.2 Abordagem bayesiana

Consideramos a distribuição Weibull denotada por $Z \sim \text{Weibull}(\kappa^{(Z)}, \gamma^{(Z)})$ para ajustar os tempos de promoção, e a distribuição Weibull denotada por $C \sim \text{Weibull}(\kappa^{(C)}, \gamma^{(C)})$ para ajustar os tempos de censura. A função de verossimilhança condicional nos efeitos aleatórios para o i -ésimo indivíduo no k -ésimo grupo, para esse modelo, se encontra na equação (3.24).

Vamos utilizar as seguintes distribuições *a priori*:

$$\begin{aligned}\kappa^{(Z)} &\sim \text{Gama} \left(a_{\kappa}^{(Z)}, b_{\kappa}^{(Z)} \right), \\ \kappa^{(C)} &\sim \text{Gama} \left(a_{\kappa}^{(C)}, b_{\kappa}^{(C)} \right), \\ \gamma^{(Z)} &\sim \text{Gama} \left(a_{\gamma}^{(Z)}, b_{\gamma}^{(Z)} \right)\end{aligned}$$

e

$$\gamma^{(C)} \sim \text{Gama} \left(a_{\gamma}^{(C)}, b_{\gamma}^{(C)} \right),$$

em que os parâmetros $\kappa^{(Z)}$, $\kappa^{(C)}$, $\gamma^{(Z)}$ e $\gamma^{(C)}$ são assumidos independentes *a priori*.

Portanto, as distribuições condicionais completas para os parâmetros e para os efeitos aleatórios são dadas por

$$\begin{aligned}\pi(\beta_p^{(T)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \exp \left\{ \delta_{ik}^{(T)} \beta_p^{(T)} x_{ik}^{(T)} - \exp(\beta^{(T)' \mathbf{x}_{ik}^{(T)} + w_k) \right. \\ &\quad \left. \times \left[1 - \exp \left(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)} \right) \right] \right\} \exp \left(\frac{\tau_{\beta} (\beta_p^{(T)})^2}{2} \right),\end{aligned}$$

$$\begin{aligned}\pi(\beta_q^{(C)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \exp \left(\delta_{ik}^{(C)} \beta_q^{(C)} x_{ik}^{(C)} - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\beta^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \\ &\quad \times \exp \left(-\frac{\tau_{\beta} (\beta_q^{(C)})^2}{2} \right),\end{aligned}$$

$$\pi(\alpha | \dots) \propto \prod_{k=1}^m \prod_{i=1}^{n_k} \exp \left(\delta_{ik}^{(C)} \alpha w_k - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\beta^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \exp \left(-\frac{\tau_{\alpha} \alpha^2}{2} \right),$$

$$\begin{aligned}\pi(\kappa^{(Z)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \left\{ \left[\kappa^{(Z)} y_{ik}^{\kappa^{(Z)}} \right]^{\delta_{ik}^{(T)}} \exp \left(\exp(\beta^{(T)' \mathbf{x}_{ik}^{(T)} + w_k) \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \right) \right\} \\ &\quad \times (\kappa^{(Z)})^{a_{\kappa}^{(Z)}} \exp(-\kappa^{(Z)} b_{\kappa}^{(Z)}),\end{aligned}$$

$$\begin{aligned}\pi(\kappa^{(C)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \left\{ \left[\kappa^{(C)} y_{ik}^{\kappa^{(C)}} \right]^{\delta_{ik}^{(C)}} \exp \left(-y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\beta^{(C)' \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \right\} \\ &\quad \times (\kappa^{(C)})^{a_{\kappa}^{(C)}} \exp(-\kappa^{(C)} b_{\kappa}^{(C)}),\end{aligned}$$

$$\begin{aligned} \pi(\gamma^{(Z)}|\dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \left\{ \left[\gamma^{(Z)} \exp(-y_{ik}^{\kappa(Z)} \gamma^{(Z)}) \right]^{\delta_{ik}^{(T)}} \exp \left(\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right. \right. \\ &\quad \left. \left. \times \exp(-y_{ik}^{\kappa(Z)} \gamma^{(Z)}) \right) \right\} (\gamma^{(Z)})^{a_\gamma^{(Z)}} \exp(-\gamma^{(Z)} b_\gamma^{(Z)}), \\ (\gamma^{(C)}|\dots) &\sim \text{Gama} \left(a_\gamma^{(C)} + \sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ik}^{(C)}, b_\gamma^{(C)} + \sum_{k=1}^m \sum_{i=1}^{n_k} y_{ik}^{\kappa(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right), \\ (\tau|\dots) &\sim \text{Gama} \left(a_\tau + \frac{m}{2}, b_\tau + \frac{\sum_{k=1}^m w_k^2}{2} \right) \end{aligned}$$

e

$$\begin{aligned} \pi(w_k|\dots) &\propto \prod_{i=1}^{n_k} \left\{ \exp \left(\delta_{ik}^{(T)} w_k - \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \left[1 - \exp(-y_{ik}^{\kappa(Z)} \gamma^{(Z)}) \right] \right) \right. \\ &\quad \left. \times \exp \left(\delta_{ik}^{(C)} \alpha w_k - y_{ik}^{\kappa(C)} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \right\} \\ &\quad \times \exp \left(-\frac{\tau w_k^2}{2} \right). \end{aligned}$$

3.4.2 Modelo de fração de cura Exponencial por Partes

Nesta seção consideramos a distribuição Exponencial por Partes para ajustar os tempos de promoção, denota por $Z \sim EP(\rho^{(Z)}, \boldsymbol{\lambda}^{(Z)})$ com $\boldsymbol{\lambda}^{(Z)} = (\lambda_1^{(Z)}, \lambda_2^{(Z)}, \dots, \lambda_b^{(Z)})'$, e também para ajustar os tempos de censura, denotada por $C \sim EP(\rho^{(C)}, \boldsymbol{\lambda}^{(C)})$ em que $\boldsymbol{\lambda}^{(C)} = (\lambda_1^{(C)}, \lambda_2^{(C)}, \dots, \lambda_d^{(C)})'$.

Seja $\rho^{(Z)}$ a grade que divide o eixo dos tempos de promoção em intervalos e $\rho^{(C)}$ a grade que divide o eixo dos tempos de censura. Conseqüentemente, as funções taxa de falha de base para os tempos de promoção Z e os tempos de censura C , são dadas por

$$h^{(Z)}(y|\boldsymbol{\lambda}^{(Z)}) = \lambda_j^{(Z)}, y \in I_j^{(Z)} = (s_{j-1}, s_j] \quad e \quad h_0^{(C)}(y|\boldsymbol{\lambda}^{(C)}) = \lambda_l^{(C)}, y \in I_l^{(C)} = (c_{l-1}, c_l]. \quad (3.26)$$

Substituindo as funções taxa de falha base (3.26) na função de verossimilhança (3.23), temos a função de verossimilhança condicional nos efeitos aleatórios, para o i -ésimo indi-

víduo no k -ésimo grupo dada por

$$\begin{aligned}
L_{ik}(\boldsymbol{\theta}; \mathbf{d}_{ik}, w_k) &= \prod_{j=1}^b \left\{ \left(\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \lambda_j^{(T)} \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right)^{\delta_{ikj}^{(T)}} \right. \\
&\quad \times \exp \left[- \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \left(1 - \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right) \right] \left. \right\} \\
&\quad \times \prod_{l=1}^d \left[\lambda_l^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right]^{\delta_{ikl}^{(C)}} \\
&\quad \times \exp \left\{ - \sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}.
\end{aligned} \tag{3.27}$$

3.4.2.1 Abordagem frequentista

Empregando a função de verossimilhança condicional (3.27) no cálculo da esperança utilizada no algoritmo EMMC, dada na equação (3.11), temos a seguinte esperança:

$$\begin{aligned}
Q_S(\boldsymbol{\theta} | \boldsymbol{\theta}^{(r)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \sum_{j=1}^b \left\{ \delta_{ikj}^{(T)} \left[\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + \hat{E}(w_k | \mathbf{d}_k) + \log(\lambda_j^{(Z)}) - \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right] \right. \\
&\quad \left. - \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \left[1 - \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right] \right\} \\
&\quad + \sum_{l=1}^d \delta_{ikl}^{(C)} \left[\log(\lambda_l^{(C)}) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha \hat{E}(w_k | \mathbf{d}_k) \right] \\
&\quad - \sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \\
&\quad - \frac{1}{2} [\log(2\pi) + \log(\sigma^2)] - \frac{\hat{E}(w_k^2 | \mathbf{d}_k)}{2\sigma^2}.
\end{aligned} \tag{3.28}$$

Derivando (3.28) com respeito aos parâmetros $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\lambda}^{(Z)}$, $\boldsymbol{\lambda}^{(C)}$, $\boldsymbol{\beta}^{(C)}$, α e σ^2 , encontramos as funções escore. As funções $U(\boldsymbol{\beta}^{(T)})$, $U(\boldsymbol{\beta}^{(C)})$, $U(\alpha)$ e $U(\sigma^2)$ são dadas em (3.25), as demais são apresentadas a seguir:

$$\begin{aligned}
U(\lambda_j^{(Z)}) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ikj}^{(T)} \left[\frac{1}{\lambda_j^{(Z)}} + (y_{ikj} - s_{j-1}) \right] \right. \\
&\quad \left. + \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}) \hat{E}(w_k | \mathbf{d}_k) \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) (-y_{ikj} - s_{j-1}) \right\}
\end{aligned}$$

e

$$U(\lambda_l^{(C)}) = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ikl}^{(C)} \left[\frac{1}{\lambda_l^{(C)}} \right] - (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}) \hat{E}(\exp(\alpha w_k) | \mathbf{d}_k) \right\}.$$

3.4.2.2 Abordagem bayesiana

Com o objetivo de especificar o modelo, precisamos definir a distribuição *a priori* conjunta de $(\boldsymbol{\beta}^{(T)}, \boldsymbol{\beta}^{(C)}, \boldsymbol{\lambda}^{(Z)}, \boldsymbol{\lambda}^{(C)}, \alpha, \tau)$, $\tau = 1/\sigma^2$. Isso pode ser feito de forma independente, como realizado na Seção 3.3.2.2. Aqui também consideramos distribuições *a priori* Gama para as taxas de falha, dadas por

$$\lambda_j^{(Z)} \sim \text{Gama} \left(a_{\lambda_j}^{(Z)}, b_{\lambda_j}^{(Z)} \right), \forall j = 1, \dots, b,$$

e

$$\lambda_l^{(C)} \sim \text{Gama} \left(a_{\lambda_l}^{(C)}, b_{\lambda_l}^{(C)} \right), \forall l = 1, \dots, d.$$

Para fazer inferência sobre os parâmetros de interesse precisamos da distribuição conjunta *a posteriori* destes parâmetros. A seguir apresentamos as distribuição condicionais completas.

$$\begin{aligned} \pi(\beta_p^{(T)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{j=1}^b \exp \left\{ \delta_{ikj}^{(T)} \beta_p^{(T)} x_{ik}^{(T)} - \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right. \\ &\quad \left. \times \left[1 - \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right] \right\} \exp \left(- \frac{\tau_\beta (\beta_p^{(T)})^2}{2} \right), \end{aligned}$$

$$\begin{aligned} \pi(\beta_q^{(C)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{l=1}^d \exp \left(\delta_{ikl}^{(C)} \beta_q^{(C)} x_{ik}^{(C)} - \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \right. \\ &\quad \left. \times \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \exp \left(- \frac{\tau_\beta (\beta_q^{(C)})^2}{2} \right), \end{aligned}$$

$$\begin{aligned} \pi(\lambda_j^{(Z)} | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{j=1}^b \left[\lambda_j^{(Z)} \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right]^{\delta_{ikj}^{(T)}} \\ &\quad \times \exp \left(\exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \exp \left(- \sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right) \\ &\quad \times (\lambda_j^{(Z)})^{a_{\lambda_j}^{(Z)}} \exp \left(- \lambda_j^{(Z)} b_{\lambda_j}^{(Z)} \right), \end{aligned}$$

$$(\lambda_l^{(C)} | \dots) \sim \text{Gama} \left(\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikl}^{(C)} + a_\lambda^{(C)}, \sum_{k=1}^m \sum_{i=1}^{n_k} (y_{ikl} - c_{l-1}) \exp(\mathbf{x}_{ik}^{(C)' } \boldsymbol{\beta}^{(C)} + \alpha w_k) + b_\lambda^{(C)} \right),$$

$$\begin{aligned} \pi(\alpha | \dots) &\propto \prod_{k=1}^m \prod_{i=1}^{n_k} \prod_{l=1}^d \exp \left(\delta_{ikl}^{(C)} \alpha w_k - \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \\ &\quad \times \exp \left(-\frac{\tau_\alpha \alpha^2}{2} \right), \end{aligned}$$

$$(\tau | \alpha, \boldsymbol{\theta}^{(T)}, \boldsymbol{\theta}^{(C)}) \sim \text{Gama} \left(a_\tau + \frac{m}{2}, b_\tau + \frac{\sum_{k=1}^m w_k^2}{2} \right)$$

e

$$\begin{aligned} \pi(w_k | \mathbf{w}_{(-k)}, \boldsymbol{\theta}, \alpha) &\propto \prod_{i=1}^{n_k} \prod_{j=1}^b \left\{ \exp \left(\delta_{ikj}^{(T)} w_k - \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \right) \right. \\ &\quad \times \left[1 - \exp \left(\sum_{j=1}^b \lambda_j^{(Z)} (y_{ikj} - s_{j-1}) \right) \right] \left. \right\} \\ &\quad \times \prod_{l=1}^d \exp \left(\delta_{ikl}^{(C)} \alpha w_k - \lambda_l^{(C)} (y_{ikl} - c_{l-1}) \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right) \left. \right\} \\ &\quad \times \exp \left(-\frac{\tau w_k^2}{2} \right). \end{aligned}$$

Capítulo 4

Estudo de simulação Monte Carlo

Neste capítulo são apresentados estudos Monte Carlo para avaliar a consistência dos estimadores de máxima verossimilhança, obtidos pelo algoritmo EMMC para os modelos propostos no Capítulo 3, e compará-los àqueles obtidos usando o modelo proposto por Huang & Wolfe (2002). Também são apresentados estudos Monte Carlo para avaliar os estimadores bayesianos média, mediana e moda *a posteriori* obtidos utilizando os modelos propostos anteriormente.

Com o objetivo de comparar os diferentes modelos, consideramos três cenários em ambas as análises: dados gerados com correlação positiva ($\alpha = 1$) entre os tempos de falha e censura; dados gerados supondo independência ($\alpha = 0$); e conjunto de dados gerados com correlação negativa ($\alpha = -1$). Para cada cenário são ajustados os modelos que não levam em conta a censura informativa, e os modelos que consideram a censura informativa. Em cada cenário geramos 500 conjuntos de dados.

Na obtenção dos resultados foi utilizado o *software* R Core Team (2017), particularmente, a função `multroot` do pacote `rootSolve` (Soetaert, 2016) para resolver equações não lineares, e a função `ARMS` do pacote `dln` (Petris, 2015) para gerar os efeitos aleatórios. A geração das amostras *a posteriori* foi realizada através do *software* JAGS (Plummer, 2015), por meio do pacote `R2jags` (Su & Yajima, 2012).

A apresentação dos resultados está dividida em duas partes: modelos de censura informativa sem fração de cura (Seção 4.1); modelos de censura informativa com fração de cura (Seção 4.2).

4.1 Estudos Monte Carlo para modelos com censura informativa sem fração de cura

Esta seção apresenta os resultados obtidos através dos modelos de censura informativa para dados de sobrevivência multivariados sem fração de cura, sob enfoques frequentista e bayesiano.

Na geração dos dados assumimos que os tempos de falha e censura seguem distribuições Weibull. Portanto, temos as seguintes equações para gerar os tempos de falha, t_{ik} , e censura, c_{ik} , respectivamente:

$$t_{ik} = \left(\frac{-\log(u_i)}{\gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik} + w_k)} \right)^{\frac{1}{\kappa^{(T)}}}$$

e

$$c_{ik} = \left(\frac{-\log(v_i)}{\gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik} + \alpha w_k)} \right)^{\frac{1}{\kappa^{(C)}}},$$

em que $u_i \sim U[0; 1]$, $v_i \sim U[0, 1]$, $\kappa^{(T)} = 2$, $\kappa^{(C)} = 2$, $\gamma^{(T)} = 0, 1$, $\gamma^{(C)} = 0, 02$, as covariáveis foram consideradas como sendo $X_1^{(T)} = X_1^{(C)} = X_1 \sim U[-10, 10]$ e $X_2^{(T)} = X_2^{(C)} = X_2 \sim \text{Ber}(0, 5)$, com coeficientes de regressão $\boldsymbol{\beta}^{(T)} = (0, 1, -1, 4)$, $\boldsymbol{\beta}^{(C)} = (0, 2, 1, 2)$. Os tempos de censura administrativa foram gerados de uma distribuição Uniforme $U[0, 10]$. Os efeitos aleatórios foram gerados de uma distribuição Normal, $w_k \sim N(0, 1)$. Consideramos um tamanho amostral de $n=200$ indivíduos, distribuídos aleatoriamente em 40 grupos, ou seja, 5 indivíduos por grupo.

Vamos denotar os modelos por

M0: Modelo Weibull;

M1: Modelo MEP com 10 intervalos;

M2: Modelo MEP com 20 intervalos;

M3: Modelo MEP com número de intervalos igual ao número de falhas e censuras distintas;

M4: Modelo com ajuste não paramétrico para as funções taxa de falha.

4.1.1 Avaliação dos estimadores de máxima verossimilhança

Nesta seção apresentamos os resultados para as simulações de Monte Carlo conside-

rando a abordagem frequentista na estimação dos parâmetros. Primeiramente, analisamos o efeito do número de indivíduos por grupo quando utilizamos o modelo de Huang & Wolfe (2002). Para tanto, na Tabela 4.1 apresentamos a média, mediana, erro padrão empírico, probabilidade de cobertura e vício relativo para 500 conjuntos de dados, considerando que o número de observações por grupo variasse de 1 a 10. Podemos perceber que quando consideramos dois ou mais indivíduos por grupo, o vício relativo diminui consideravelmente, em relação ao caso em que consideramos um indivíduo por grupo. Essa característica também foi observada por Huang & Wolfe (2002). Huang & Wolfe (2002) comentaram que, se o tamanho dos grupos for pequeno, as estimativas da fragilidade e dos demais parâmetros são instáveis. Com o objetivo de comparar os resultados obtidos por meio dos modelos propostos neste trabalho com aqueles fornecidos pelo modelo de Huang & Wolfe (2002), optamos por considerar cinco indivíduos em cada grupo.

Para escolha do número de réplicas para o algoritmo EM Monte Carlo nos baseamos no trabalho de Wei & Tanner (1990). Estes autores comentaram que o algoritmo EM Monte Carlo pode ser ineficiente se for inicializado com um número de réplicas muito grande para o passo E, pois as estimativas podem ficar longe dos valores verdadeiros. Wei & Tanner (1990) propõem começar gerando amostras pequenas, ir aumentando gradativamente e observar a convergência.

Com o objetivo de escolher o número máximo de iterações, aplicamos os modelos propostos em alguns conjuntos de dados, notamos que os valores convergem rapidamente, como podemos ver nos gráficos do traço das estimativas que encontram-se nas Figuras B.1 e B.2 (Apêndice B). Para correção do erro Monte Carlo utilizamos a média das últimas três estimativas, após a convergência. Para o número de réplicas do efeito aleatório w_k em cada iteração, S , consideramos a seguinte sequência: $S = 10$ para as iterações 10 - 20, $S = 25$ para as iterações 21 - 50, $S = 50$ para as iterações 51 - 100, $S = 75$ para as iterações 101 - 150, $S = 100$ para as iterações 151 - 200, $S = 125$ para as iterações 201 - 250, $S = 150$ para as iterações 251 - 300, $S = 200$ para as iterações 301 - 350, $S = 250$ para as iterações 351 - 400.

Para definirmos o critério de parada utilizamos a abordagem proposta por Booth & Hobert (1999), dada pela diferença relativa das estimativas nas iterações r e $r + 1$, obtida por

$$\max_i \left(\frac{|\theta_i^{(r+1)} - \theta_i^{(r)}|}{|\theta_i^{(r)}| + \epsilon_1} \right) < \epsilon_2,$$

em que $\epsilon_1 = 0,001$ e $\epsilon_2 = 0,0001$ são constantes especificadas, $\theta_i^{(r)}$ é a estimativa de θ_i na r -ésima iteração, sendo $\boldsymbol{\theta} = (\boldsymbol{\beta}^{(T)}, \boldsymbol{\beta}^{(C)}, \alpha, \tau)$ o vetor de parâmetros de interesse.

Tabela 4.1: Parâmetros estimados pelo modelo de Huang & Wolfe(2002), supondo censura informativa. Dados com 40% falha, 32% censura informativa e 18% censura administrativa.

	Verdadeiro	Média	Mediana	EPE	LI	LS	PC	Vício (%)
1 indivíduo por grupo								
$\beta_1^{(T)}$	0,1	0,080	0,0785	0,031	0,019	0,1407	0,902	-19,857
$\beta_2^{(T)}$	-1,4	-1,419	-1,418	0,350	-2,104	-0,733	0,956	-1,336
$\beta_1^{(C)}$	0,2	0,178	0,174	0,037	0,105	0,251	0,920	-11,121
$\beta_2^{(C)}$	1,2	1,286	1,281	0,391	0,520	2,053	0,944	7,204
α	1	0,623	0,669	1,032	-1,400	2,645	0,928	-37,732
σ^2	1	0,505	0,342	0,444	0,000	1,376	0,864	-49,490
2 indivíduos por grupo								
$\beta_1^{(T)}$	0,1	0,099	0,097	0,030	0,039	0,159	0,956	-1,126
$\beta_2^{(T)}$	-1,4	-1,421	-1,406	0,342	-2,091	-0,750	0,956	-1,464
$\beta_1^{(C)}$	0,2	0,200	0,196	0,039	0,124	0,277	0,958	0,225
$\beta_2^{(C)}$	1,2	1,233	1,232	0,391	0,466	2,000	0,946	2,773
α	1	1,050	0,972	0,617	-0,159	2,260	0,966	5,036
σ^2	1	0,978	0,882	0,560	0,000	2,075	0,968	-2,229
3 indivíduos por grupo								
$\beta_1^{(T)}$	0,1	0,100	0,098	0,031	0,039	0,160	0,948	-0,297
$\beta_2^{(T)}$	-1,4	-1,433	-1,391	0,347	-2,113	-0,753	0,944	-2,373
$\beta_1^{(C)}$	0,2	0,199	0,196	0,034	0,131	0,266	0,960	-0,620
$\beta_2^{(C)}$	1,2	1,261	1,244	0,348	0,579	1,943	0,942	5,096
α	1	1,062	0,965	0,478	0,125	1,998	0,962	6,153
σ^2	1	0,988	0,945	0,502	0,004	1,973	0,962	-1,153
4 indivíduos por grupo								
$\beta_1^{(T)}$	0,1	0,100	0,099	0,028	0,046	0,154	0,948	-0,021
$\beta_2^{(T)}$	-1,4	-1,412	-1,386	0,328	-2,056	-0,769	0,954	-0,890
$\beta_1^{(C)}$	0,2	0,202	0,201	0,037	0,131	0,274	0,948	1,221
$\beta_2^{(C)}$	1,2	1,216	1,220	0,371	0,489	1,943	0,952	1,323
α	1	1,040	0,984	0,374	0,307	1,774	0,944	4,039
σ^2	1	0,991	0,963	0,422	0,165	1,817	0,940	-0,924
5 indivíduos por grupo								
$\beta_1^{(T)}$	0,1	0,1021	0,1003	0,0278	0,0476	0,1565	0,956	2,057
$\beta_2^{(T)}$	-1,4	-1,413	-1,409	0,301	-2,003	-0,824	0,958	-0,940
$\beta_1^{(C)}$	0,2	0,202	0,200	0,034	0,136	0,269	0,952	1,237
$\beta_2^{(C)}$	1,2	1,217	1,215	0,358	0,516	1,918	0,958	1,393
α	1	1,036	0,984	0,383	0,286	1,786	0,964	3,604
σ^2	1	1,026	0,932	0,486	0,072	1,980	0,944	2,598
10 indivíduos por grupo								
$\beta_1^{(T)}$	0,1	0,102	0,100	0,027	0,050	0,155	0,938	2,241
$\beta_2^{(T)}$	-1,4	-1,400	-1,388	0,297	-1,983	-0,818	0,96	-0,014
$\beta_1^{(C)}$	0,2	0,202	0,201	0,032	0,139	0,265	0,944	1,091
$\beta_2^{(C)}$	1,2	1,222	1,210	0,354	0,528	1,917	0,946	1,860
α	1	1,037	1,001	0,313	0,422	1,651	0,956	3,660
σ^2	1	0,973	0,880	0,504	0,000	1,961	0,96	-2,711

¹EPE: erro padrão empírico calculado a partir das 500 estimativas dos parâmetros, LI: limite inferior do intervalo de confiança empírico calculado a partir do EPE, LS: limite superior do intervalo de confiança empírico calculado a partir do EPE, PC: probabilidade de cobertura, Vício: calculado a partir do valor estimado pela média.

Com o objetivo de avaliar o ajuste em modelos mistos com efeito aleatório, Chen *et al.* (2002a) consideraram a seguinte aproximação para a função log-verossimilhança

$$\log L(\hat{\boldsymbol{\theta}}; \mathbf{d}) = \sum_{k=1}^m \log \left\{ \frac{1}{S} \sum_{s=1}^S L(\mathbf{d}_k | w_k^{(s)}, \hat{\boldsymbol{\theta}}) \right\}, \quad (4.1)$$

em que S é o número de réplicas do efeito aleatório.

Seja D_θ a dimensão do vetor de parâmetros $\boldsymbol{\theta}$ e $N = \sum_{k=1}^m n_k$ o número total de observações. Chen *et al.* (2002a) utilizaram $-\log L(\hat{\boldsymbol{\theta}}; \mathbf{d}) + a_P$, se $a_P = D_\theta$, tem-se o critério de Akaike, AIC; se $a_P = (1/2) \log(N) D_\theta$ tem-se o critério de Schwarz, BIC; e se $a_P = \log(\log(N)) D_\theta$ tem-se o critério de Hannan-Quin, HQ. Modelos com valores desses critérios menores são preferíveis. Dessa forma, além de avaliar o vício relativo, intervalo de confiança e probabilidade de cobertura, também vamos comparar a qualidade dos ajustes por meio dos critérios AIC, BIC e HQ.

A seguir apresentamos o ajuste empregando os cinco modelos apresentados na Seção 3.1. As Tabelas 4.2 e 4.3 apresentam as médias, erros padrões, probabilidade de cobertura e critérios para o cenário com correlação positiva. Da Tabela 4.2 nota-se que o melhor modelo segundo os critérios é o modelo Weibull (M0), lembrando que os dados foram gerados de distribuições Weibull. Em seguida, o modelo M1 teve o menor valor para os critérios, ou seja, considerar 10 intervalos para os tempos de falha e censura é melhor que considerar 20 intervalos. O modelo M4 apresentou menor vício relativo, porém com altos valores para todos os critérios. Um fato interessante ocorre ao compararmos os modelos M3 e M4, o primeiro considera número de intervalos para o modelo MEP igual ao número de falhas e censuras distintas, o segundo considera o estimador proposto por Breslow (1974), portanto ambos possuem o mesmo número de parâmetros, porém o valor dos critérios para M4 é aproximadamente o dobro do valor dos critérios para M3, vide Tabela 4.2.

Ainda considerando o cenário com correlação positiva, ao compararmos os ajustes supondo censura informativa (Tabela 4.2) com os ajustes supondo censura não informativa (Tabela 4.3) pode-se observar, de maneira geral, que o vício relativo é aproximadamente três ou mais vezes maior para os modelos que não consideram a censura informativa. A probabilidade de cobertura fica próxima de 95% para quase todos os parâmetros, exceto para τ , quando ajustamos os modelos supondo censura informativa (Tabela 4.2); ao passo que, quando ajustamos os modelos supondo censura não informativa, essa probabilidade apresenta-se menor na maioria dos parâmetros (Tabela 4.3). Além do mais, os critérios apontam para os modelos com censura informativa.

As Tabelas 4.4 e 4.5 apresentam as estimativas pontuais e intervalares para o cenário

com correlação negativa entre os tempos. De acordo com a Tabela 4.4, nota-se que o melhor ajuste apontado pelos critérios foi o modelo M0, seguido do modelo M1. Porém, dentre todos os modelos, de uma forma geral, os modelos M3 e M4 apresentaram estimativas mais próximas dos valores reais que as estimativas fornecidas pelos modelos M0 e M1. Ao comparar o ajuste supondo censura informativa (Tabela 4.4) com o ajuste que considera censura não informativa (Tabela 4.5), novamente podemos constatar um vício relativo menor para os modelos que levam consideração a censura informativa.

O ajuste para o cenário em que os tempos de falha e censura foram gerados supondo independência encontra-se nas Tabelas 4.6 e 4.7. A partir da Tabela 4.6 percebe-se que o modelo M4 apresenta vícios menores, porém com valores para os critérios sendo aproximadamente o dobro dos demais modelos. Ao observamos os valores dos critérios para os modelos com suposição de censura informativa (Tabela 4.6) com os valores obtidos para o modelos que consideram a censura não informativa (Tabela 4.7), percebe-se que os valores para os três critérios são muito próximos, indicando que os modelos com ambas suposições podem ser aplicados. Os vícios relativos também são semelhantes para as duas abordagens.

Para avaliar os critérios clássicos AIC, BIC e HQ calculamos a razão entre os valores desses critérios. Esta razão foi calculada da seguinte forma

$$\text{Razão} = \frac{\text{Valor do critério para o modelo com ajuste incorreto}}{\text{Valor do critério para o modelo com ajuste correto}}.$$

Portanto, o valor da razão sendo maior que 1, indica que o modelo sob ajuste incorreto possui o maior valor, logo o modelo sob ajuste correto é melhor. As distribuições das razões para os cenários com correlação negativa, nula e positiva estão representadas na Figura 4.1.

A Tabela 4.8 apresenta a proporção do número de vezes que o critério indica o modelo com ajuste correto. Observa-se que os critérios não distinguem muito bem entre aplicação com censura informativa e censura não informativa. Porém, por exemplo, para o cenário com correlação negativa, entre os modelos M0, M1, M2, M3 e M4 em 495 (99%) conjuntos de dados o critério AIC aponta para o modelo M0, e entre os modelo M1, M2, M3 e M4 o critério AIC indica para o modelo M1 em torno de 90,8% das vezes.

Tabela 4.2: Ajuste com abordagem frequentista e censura informativa, dados com correlação positiva ($\alpha = 1$).

	Verdadeiro	Média (VR%)	EPE	EPL	IC	PC	
M0							
$\beta_1^{(T)}$	0,1	0,101 (1,2)	0,026	0,027	[0,051; 0,152]	0,954	$\overline{AIC} = 318,238$
$\beta_2^{(T)}$	-1,4	-1,469 (-4,9)	0,315	0,313	[-2,087; -0,851]	0,942	$\overline{BIC} = 334,729$
$\beta_1^{(C)}$	0,2	0,208 (4,2)	0,032	0,032	[0,146; 0,271]	0,940	$\overline{HQ} = 324,912$
$\beta_2^{(C)}$	1,2	1,252 (4,4)	0,379	0,345	[0,509; 1,996]	0,960	
α	1	1,050 (5,0)	0,331	0,327	[0,402; 1,699]	0,938	
σ^2	1	1,049 (4,9)	0,448	0,481	[0,172; 1,927]	0,952	
τ	1	1,160 (16,0)	0,478	0,576	[0,002; 2,318]	0,936	
M1							
$\beta_1^{(T)}$	0,1	0,092 (-7,7)	0,025	0,027	[0,043; 0,141]	0,958	$\overline{AIC} = 335,464$
$\beta_2^{(T)}$	-1,4	-1,434 (-2,4)	0,308	0,312	[-2,038; -0,829]	0,948	$\overline{BIC} = 378,342$
$\beta_1^{(C)}$	0,2	0,198 (-0,8)	0,031	0,033	[0,138; 0,259]	0,942	$\overline{HQ} = 352,816$
$\beta_2^{(C)}$	1,2	1,322 (10,1)	0,375	0,350	[0,588; 2,056]	0,956	
α	1	1,052 (5,2)	0,372	0,370	[0,328; 1,776]	0,950	
σ^2	1	0,888 (-11,2)	0,402	0,469	[0,101; 1,675]	0,960	
τ	1	1,430 (43,0)	0,651	0,634	[0,001; 3,174]	0,926	
M2							
$\beta_1^{(T)}$	0,1	0,094 (-6,2)	0,026	0,026	[0,044; 0,144]	0,958	$\overline{AIC} = 343,044$
$\beta_2^{(T)}$	-1,4	-1,437 (-2,6)	0,311	0,316	[-2,047; -0,827]	0,946	$\overline{BIC} = 418,906$
$\beta_1^{(C)}$	0,2	0,201 (0,3)	0,032	0,033	[0,139; 0,263]	0,942	$\overline{HQ} = 373,744$
$\beta_2^{(C)}$	1,2	1,322 (10,2)	0,377	0,351	[0,583; 2,062]	0,950	
α	1	1,057 (5,7)	0,373	0,373	[0,326; 1,789]	0,942	
σ^2	1	0,914 (-8,6)	0,417	0,490	[0,097; 1,730]	0,960	
τ	1	1,392 (39,2)	0,634	0,736	[0,001; 3,054]	0,928	
M3							
$\beta_1^{(T)}$	0,1	0,095 (-5,1)	0,026	0,027	[0,045; 0,145]	0,958	$\overline{AIC} = 387,001$
$\beta_2^{(T)}$	-1,4	-1,447 (-3,3)	0,312	0,314	[-2,058; -0,835]	0,948	$\overline{BIC} = 641,505$
$\beta_1^{(C)}$	0,2	0,201 (0,7)	0,032	0,033	[0,139; 0,264]	0,944	$\overline{HQ} = 489,995$
$\beta_2^{(C)}$	1,2	1,303 (8,6)	0,353	0,349	[0,610; 1,996]	0,940	
α	1	1,052 (5,2)	0,367	0,366	[0,334; 1,771]	0,948	
σ^2	1	0,932 (-6,8)	0,421	0,490	[0,106; 1,758]	0,968	
τ	1	1,352 (35,2)	0,616	0,756	[0,001; 2,879]	0,926	
M4							
$\beta_1^{(T)}$	0,1	0,099 (-1,1)	0,026	0,028	[0,048; 0,15]	0,956	$\overline{AIC} = 859,397$
$\beta_2^{(T)}$	-1,4	-1,454 (-3,9)	0,314	0,333	[-2,07; -0,839]	0,952	$\overline{BIC} = 1113,901$
$\beta_1^{(C)}$	0,2	0,205 (2,7)	0,032	0,036	[0,142; 0,269]	0,942	$\overline{HQ} = 962,391$
$\beta_2^{(C)}$	1,2	1,241 (3,4)	0,375	0,336	[0,506; 1,976]	0,962	
α	1	1,048 (4,8)	0,341	0,409	[0,381; 1,716]	0,942	
σ^2	1	0,998 (-0,2)	0,444	0,488	[0,127; 1,869]	0,966	
τ	1	1,246 (24,6)	0,547	0,537	[0,001; 2,578]	0,928	

¹VR: vício relativo, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.3: Ajuste com abordagem frequentista e censura não informativa, dados com correlação positiva ($\alpha = 1$).

	Verdadeiro	Média (VR%)	EPE	EPL	IC	PC	
M0							
$\beta_1^{(T)}$	0,1	0,089 (-11,1)	0,027	0,027	[0,036; 0,142]	0,942	$\overline{AIC} = 318,851$
$\beta_2^{(T)}$	-1,4	-1,58 (-12,8)	0,333	0,324	[-2,232; -0,927]	0,926	$\overline{BIC} = 335,342$
σ^2	1	0,936 (-6,4)	0,454	0,459	[0,045; 1,827]	0,962	$\overline{HQ} = 325,525$
τ	1	1,458 (45,8)	0,685	0,801	[0,001; 4,26]	0,972	
M1							
$\beta_1^{(T)}$	0,1	0,081 (-19,1)	0,026	0,027	[0,03; 0,132]	0,892	$\overline{AIC} = 337,314$
$\beta_2^{(T)}$	-1,4	-1,513 (-8,1)	0,321	0,324	[-2,143; -0,884]	0,936	$\overline{BIC} = 380,193$
σ^2	1	0,762 (-23,8)	0,401	0,462	[-0,024; 1,548]	0,920	$\overline{HQ} = 354,667$
τ	1	2,02 (102)	1,148	1,137	[0,001; 7,103]	0,966	
M2							
$\beta_1^{(T)}$	0,1	0,082 (-18,1)	0,026	0,027	[0,03; 0,134]	0,902	$\overline{AIC} = 344,889$
$\beta_2^{(T)}$	-1,4	-1,52 (-8,6)	0,326	0,323	[-2,159; -0,882]	0,932	$\overline{BIC} = 420,751$
σ^2	1	0,786 (-21,4)	0,415	0,464	[-0,028; 1,599]	0,932	$\overline{HQ} = 375,589$
τ	1	2,018 (101,8)	1,115	1,145	[0,001; 7,743]	0,972	
M3							
$\beta_1^{(T)}$	0,1	0,083 (-17,4)	0,027	0,027	[0,031; 0,135]	0,904	$\overline{AIC} = 388,778$
$\beta_2^{(T)}$	-1,4	-1,531 (-9,3)	0,326	0,323	[-2,169; -0,892]	0,928	$\overline{BIC} = 643,283$
σ^2	1	0,798 (-20,2)	0,418	0,459	[0,001; 1,618]	0,936	$\overline{HQ} = 491,772$
τ	1	1,892 (89,2)	1,075	1,072	[-2,296; 6,081]	0,960	
M4							
$\beta_1^{(T)}$	0,1	0,086 (-14,2)	0,027	0,026	[0,033; 0,138]	0,924	$\overline{AIC} = 860,123$
$\beta_2^{(T)}$	-1,4	-1,546 (-10,4)	0,327	0,355	[-2,186; -0,905]	0,932	$\overline{BIC} = 1114,628$
σ^2	1	0,847 (-15,3)	0,439	0,373	[0,001; 1,707]	0,952	$\overline{HQ} = 963,117$
τ	1	1,702 (70,2)	0,929	0,810	[0,001; 4,862]	0,958	

¹VR: vício relativo, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.4: Ajuste com abordagem frequentista e censura informativa, dados com correlação negativa ($\alpha = -1$).

	Verdadeiro	Média (VR%)	EPE	EPL	IC	PC	
M0							
$\beta_1^{(T)}$	0,1	0,103 (3,1)	0,026	0,025	[0,053; 0,154]	0,946	$\overline{AIC} = 310,682$
$\beta_2^{(T)}$	-1,4	-1,431 (-2,2)	0,315	0,304	[-2,05; -0,813]	0,948	$\overline{BIC} = 327,173$
$\beta_1^{(C)}$	0,2	0,206 (3,0)	0,033	0,031	[0,142; 0,271]	0,95	$\overline{HQ} = 317,356$
$\beta_2^{(C)}$	1,2	1,246 (3,8)	0,326	0,321	[0,606; 1,886]	0,94	
α	-1	-1,074 (-7,4)	0,414	0,378	[-1,886; -0,262]	0,944	
σ^2	1	1,014 (1,4)	0,485	0,499	[0,063; 1,966]	0,946	
τ	1	1,308 (30,8)	0,641	0,683	[0,001; 3,221]	0,946	
M1							
$\beta_1^{(T)}$	0,1	0,097 (-2,8)	0,025	0,025	[0,048; 0,146]	0,942	$\overline{AIC} = 326,648$
$\beta_2^{(T)}$	-1,4	-1,385 (1,1)	0,288	0,298	[-1,949; -0,819]	0,952	$\overline{BIC} = 369,527$
$\beta_1^{(C)}$	0,2	0,203 (1,4)	0,031	0,032	[0,141; 0,264]	0,948	$\overline{HQ} = 344,001$
$\beta_2^{(C)}$	1,2	1,252 (4,3)	0,349	0,326	[0,567; 1,936]	0,94	
α	-1	-1,099 (-9,9)	0,358	0,393	[-1,799; -0,398]	0,948	
σ^2	1	0,948 (-5,2)	0,425	0,504	[0,114; 1,782]	0,958	
τ	1	1,342 (34,2)	0,576	0,735	[0,001; 3,14]	0,956	
M2							
$\beta_1^{(T)}$	0,1	0,098 (-1,9)	0,025	0,026	[0,048; 0,148]	0,948	$\overline{AIC} = 334,369$
$\beta_2^{(T)}$	-1,4	-1,392 (0,5)	0,29	0,296	[-1,961; -0,824]	0,95	$\overline{BIC} = 410,231$
$\beta_1^{(C)}$	0,2	0,205 (2,4)	0,032	0,032	[0,142; 0,267]	0,948	$\overline{HQ} = 365,069$
$\beta_2^{(C)}$	1,2	1,259 (4,9)	0,351	0,329	[0,572; 1,946]	0,940	
α	-1	-1,102 (-10,2)	0,362	0,393	[-1,811; -0,393]	0,942	
σ^2	1	0,966 (-3,4)	0,436	0,503	[0,112; 1,82]	0,948	
τ	1	1,319 (31,9)	0,573	0,724	[0,001; 3,066]	0,952	
M3							
$\beta_1^{(T)}$	0,1	0,099 (-1,0)	0,026	0,026	[0,049; 0,149]	0,946	$\overline{AIC} = 382,023$
$\beta_2^{(T)}$	-1,4	-1,404 (-0,3)	0,294	0,302	[-1,98; -0,827]	0,946	$\overline{BIC} = 653,897$
$\beta_1^{(C)}$	0,2	0,205 (2,7)	0,032	0,033	[0,143; 0,268]	0,942	$\overline{HQ} = 492,046$
$\beta_2^{(C)}$	1,2	1,255 (4,5)	0,348	0,327	[0,573; 1,936]	0,94	
α	-1	-1,099 (-9,9)	0,36	0,396	[-1,805; -0,393]	0,948	
σ^2	1	0,972 (-2,8)	0,442	0,514	[0,106; 1,838]	0,952	
τ	1	1,31 (31,0)	0,574	0,742	[0,001; 2,997]	0,952	
M4							
$\beta_1^{(T)}$	0,1	0,101 (1,4)	0,026	0,028	[0,051; 0,152]	0,946	$\overline{AIC} = 909,909$
$\beta_2^{(T)}$	-1,4	-1,418 (-1,3)	0,300	0,416	[-2,006; -0,829]	0,948	$\overline{BIC} = 1181,783$
$\beta_1^{(C)}$	0,2	0,208 (3,9)	0,032	0,031	[0,144; 0,271]	0,942	$\overline{HQ} = 1019,932$
$\beta_2^{(C)}$	1,2	1,218 (1,5)	0,342	0,215	[0,547; 1,888]	0,938	
α	-1	-1,088 (-8,8)	0,381	0,209	[-1,834; -0,341]	0,954	
σ^2	1	0,969 (-3,1)	0,445	0,353	[0,098; 1,840]	0,948	
τ	1	1,351 (35,1)	0,598	0,471	[0,001; 3,552]	0,968	

¹VR: vício relativo, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.5: Ajuste com abordagem frequentista e censura não informativa, dados com correlação negativa ($\alpha = -1$).

	Verdadeiro	Média (VR%)	EPE	EPL	IC	PC	
M0							
$\beta_1^{(T)}$	0,1	0,112 (11,7)	0,027	0,026	[0,059; 0,164]	0,924	$\overline{AIC} = 310,368$
$\beta_2^{(T)}$	-1,4	-1,343 (4,0)	0,324	0,310	[-1,978; -0,709]	0,940	$\overline{BIC} = 326,859$
σ^2	1	0,865 (-13,5)	0,434	0,421	[0,014; 1,715]	0,950	$\overline{HQ} = 317,042$
τ	1	1,738 (73,8)	0,866	0,954	[0,001; 8,116]	0,984	
M1							
$\beta_1^{(T)}$	0,1	0,106 (5,9)	0,025	0,026	[0,056; 0,155]	0,940	$\overline{AIC} = 327,254$
$\beta_2^{(T)}$	-1,4	-1,303 (6,9)	0,292	0,313	[-1,876; -0,729]	0,946	$\overline{BIC} = 370,132$
σ^2	1	0,790 (-20,9)	0,389	0,444	[0,026; 1,554]	0,946	$\overline{HQ} = 344,606$
τ	1	1,912 (91,2)	0,831	1,011	[-5,83; 9,654]	0,990	
M2							
$\beta_1^{(T)}$	0,1	0,107 (7,0)	0,026	0,026	[0,057; 0,157]	0,944	$\overline{AIC} = 334,768$
$\beta_2^{(T)}$	-1,4	-1,311 (6,3)	0,295	0,312	[-1,888; -0,734]	0,946	$\overline{BIC} = 410,629$
σ^2	1	0,805 (-19,5)	0,399	0,447	[0,022; 1,588]	0,952	$\overline{HQ} = 365,467$
τ	1	1,872 (87,2)	0,823	1,004	[0,001; 9,915]	0,992	
M3							
$\beta_1^{(T)}$	0,1	0,108 (8,0)	0,026	0,026	[0,057; 0,159]	0,940	$\overline{AIC} = 382,307$
$\beta_2^{(T)}$	-1,4	-1,325 (5,4)	0,299	0,311	[-1,909; -0,739]	0,948	$\overline{BIC} = 654,180$
σ^2	1	0,817 (-18,3)	0,405	0,422	[0,023; 1,607]	0,956	$\overline{HQ} = 492,329$
τ	1	1,813 (81,3)	0,803	0,885	[0,001; 8,231]	0,990	
M4							
$\beta_1^{(T)}$	0,1	0,111 (10,8)	0,026	0,028	[0,059; 0,162]	0,934	$\overline{AIC} = 911,521$
$\beta_2^{(T)}$	-1,4	-1,342 (4,2)	0,305	0,378	[-1,939; -0,743]	0,952	$\overline{BIC} = 1183,395$
σ^2	1	0,819 (-18,0)	0,408	0,392	[0,021; 1,619]	0,958	$\overline{HQ} = 1021,544$
τ	1	1,857 (85,7)	0,790	0,858	[0,001; 10,735]	0,994	

¹VR: vício relativo, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.6: Ajuste com abordagem frequentista e censura informativa, dados com correlação nula ($\alpha = 0$).

	Verdadeiro	Média(VR%)	EPE	EPL	IC	PC	
M0							
$\beta_1^{(T)}$	0,1	0,102 (2,1)	0,027	0,026	[0,049; 0,154]	0,958	$\overline{AIC} = 318,649$
$\beta_2^{(T)}$	-1,4	-1,429 (-2,1)	0,325	0,314	[-2,066; -0,793]	0,944	$\overline{BIC} = 335,140$
$\beta_1^{(C)}$	0,2	0,208 (4,1)	0,032	0,029	[0,146; 0,271]	0,928	$\overline{HQ} = 325,323$
$\beta_2^{(C)}$	1,2	1,245 (3,7)	0,339	0,315	[0,579; 1,911]	0,944	
α	0	-0,003 (-0,3)	0,258	0,229	[-0,509; 0,504]	0,944	
σ^2	1	1,003 (0,3)	0,462	0,463	[0,097; 1,909]	0,946	
τ	1	1,292 (29,2)	0,589	0,650	[0,001; 3,045]	0,952	
M1							
$\beta_1^{(T)}$	0,1	0,094 (-6,4)	0,026	0,027	[0,042; 0,145]	0,946	$\overline{AIC} = 336,183$
$\beta_2^{(T)}$	-1,4	-1,388 (0,8)	0,312	0,312	[-1,999; -0,777]	0,946	$\overline{BIC} = 379,061$
$\beta_1^{(C)}$	0,2	0,203 (1,7)	0,030	0,03	[0,145; 0,262]	0,936	$\overline{HQ} = 353,535$
$\beta_2^{(C)}$	1,2	1,284 (7,0)	0,332	0,322	[0,633; 1,935]	0,930	
α	0	-0,087 (-8,7)	0,304	0,248	[-0,684; 0,509]	0,952	
σ^2	1	0,914 (-8,6)	0,422	0,474	[0,087; 1,741]	0,956	
τ	1	1,515 (51,5)	0,651	0,862	[0,001; 6,693]	0,996	
M2							
$\beta_1^{(T)}$	0,1	0,095 (-5,3)	0,026	0,027	[0,043; 0,147]	0,946	$\overline{AIC} = 343,518$
$\beta_2^{(T)}$	-1,4	-1,396 (0,3)	0,315	0,312	[-2,014; -0,778]	0,952	$\overline{BIC} = 419,379$
$\beta_1^{(C)}$	0,2	0,205 (2,3)	0,031	0,03	[0,145; 0,264]	0,936	$\overline{HQ} = 374,218$
$\beta_2^{(C)}$	1,2	1,288 (7,3)	0,334	0,322	[0,633; 1,943]	0,928	
α	0	-0,081 (-8,1)	0,302	0,247	[-0,672; 0,511]	0,954	
σ^2	1	0,935 (-6,5)	0,431	0,483	[0,089; 1,781]	0,956	
τ	1	1,475 (47,5)	0,639	0,773	[0,001; 5,884]	0,990	
M3							
$\beta_1^{(T)}$	0,1	0,096 (-4,4)	0,027	0,027	[0,043; 0,148]	0,944	$\overline{AIC} = 391,064$
$\beta_2^{(T)}$	-1,4	-1,405 (-0,4)	0,316	0,317	[-2,023; -0,787]	0,950	$\overline{BIC} = 660,016$
$\beta_1^{(C)}$	0,2	0,205 (2,7)	0,030	0,030	[0,146; 0,265]	0,932	$\overline{HQ} = 499,905$
$\beta_2^{(C)}$	1,2	1,281 (6,7)	0,336	0,321	[0,623; 1,938]	0,928	
α	0	-0,075 (-7,4)	0,297	0,247	[-0,658; 0,508]	0,958	
σ^2	1	0,944 (-5,6)	0,434	0,502	[0,094; 1,794]	0,962	
τ	1	1,455 (45,5)	0,622	0,766	[0,001; 5,918]	0,992	
M4							
$\beta_1^{(T)}$	0,1	0,099 (-0,9)	0,027	0,028	[0,046; 0,152]	0,938	$\overline{AIC} = 908,049$
$\beta_2^{(T)}$	-1,4	-1,416 (-1,1)	0,318	0,354	[-2,04; -0,792]	0,948	$\overline{BIC} = 1177,001$
$\beta_1^{(C)}$	0,2	0,209 (4,3)	0,031	0,031	[0,148; 0,269]	0,926	$\overline{HQ} = 1016,890$
$\beta_2^{(C)}$	1,2	1,226 (2,1)	0,331	0,343	[0,577; 1,874]	0,950	
α	0	-0,025 (-2,5)	0,288	0,233	[-0,589; 0,539]	0,956	
σ^2	1	0,969 (-3,1)	0,441	0,433	[0,104; 1,834]	0,956	
τ	1	1,395 (39,5)	0,592	0,608	[0,001; 5,116]	0,984	

¹VR: vício relativo, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.7: Ajuste com abordagem frequentista e censura não informativa, dados com correlação nula ($\alpha = 0$).

	Verdadeiro	Média(VR%)	EPE	EPM	IC	PC	
M0							
$\beta_1^{(T)}$	0,1	0,102 (2,1)	0,027	0,026	[0,049; 0,155]	0,956	$\overline{AIC} = 318,569$
$\beta_2^{(T)}$	-1,4	-1,428 (-1,9)	0,324	0,312	[-2,062; -0,794]	0,952	$\overline{BIC} = 335,061$
σ^2	1	0,999 (-0,1)	0,462	0,454	[0,093; 1,905]	0,946	$\overline{HQ} = 325,243$
τ	1	1,298 (29,8)	0,605	0,659	[0,001; 3,071]	0,952	
M1							
$\beta_1^{(T)}$	0,1	0,094 (-5,9)	0,026	0,026	[0,043; 0,145]	0,944	$\overline{AIC} = 336,138$
$\beta_2^{(T)}$	-1,4	-1,383 (1,2)	0,310	0,308	[-1,991; -0,775]	0,954	$\overline{BIC} = 379,016$
σ^2	1	0,905 (-9,5)	0,421	0,457	[0,079; 1,731]	0,962	$\overline{HQ} = 353,490$
τ	1	1,446 (44,6)	0,676	0,771	[-0,826; 3,719]	0,960	
M2							
$\beta_1^{(T)}$	0,1	0,095 (-4,8)	0,026	0,027	[0,043; 0,147]	0,950	$\overline{AIC} = 343,378$
$\beta_2^{(T)}$	-1,4	-1,391 (0,6)	0,314	0,313	[-2,006; -0,776]	0,958	$\overline{BIC} = 419,239$
σ^2	1	0,926 (-7,4)	0,429	0,478	[0,085; 1,768]	0,958	$\overline{HQ} = 374,078$
τ	1	1,418 (41,8)	0,662	0,791	[0; 3,722]	0,962	
M3							
$\beta_1^{(T)}$	0,1	0,096 (-3,9)	0,027	0,027	[0,044; 0,148]	0,944	$\overline{AIC} = 390,968$
$\beta_2^{(T)}$	-1,4	-1,402 (-0,1)	0,314	0,311	[-2,018; -0,786]	0,956	$\overline{BIC} = 659,919$
σ^2	1	0,938 (-6,2)	0,432	0,474	[0,092; 1,784]	0,966	$\overline{HQ} = 499,808$
τ	1	1,381 (38,1)	0,646	0,763	[0,001; 3,258]	0,946	
M4							
$\beta_1^{(T)}$	0,1	0,099 (-0,9)	0,027	0,028	[0,047; 0,151]	0,942	$\overline{AIC} = 907,977$
$\beta_2^{(T)}$	-1,4	-1,414 (-1,0)	0,317	0,359	[-2,036; -0,793]	0,954	$\overline{BIC} = 1176,929$
σ^2	1	0,962 (-3,8)	0,441	0,417	[0,097; 1,827]	0,958	$\overline{HQ} = 1016,818$
τ	1	1,336 (33,6)	0,610	0,636	[0,001; 3,069]	0,948	

¹VR: vício relativo, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

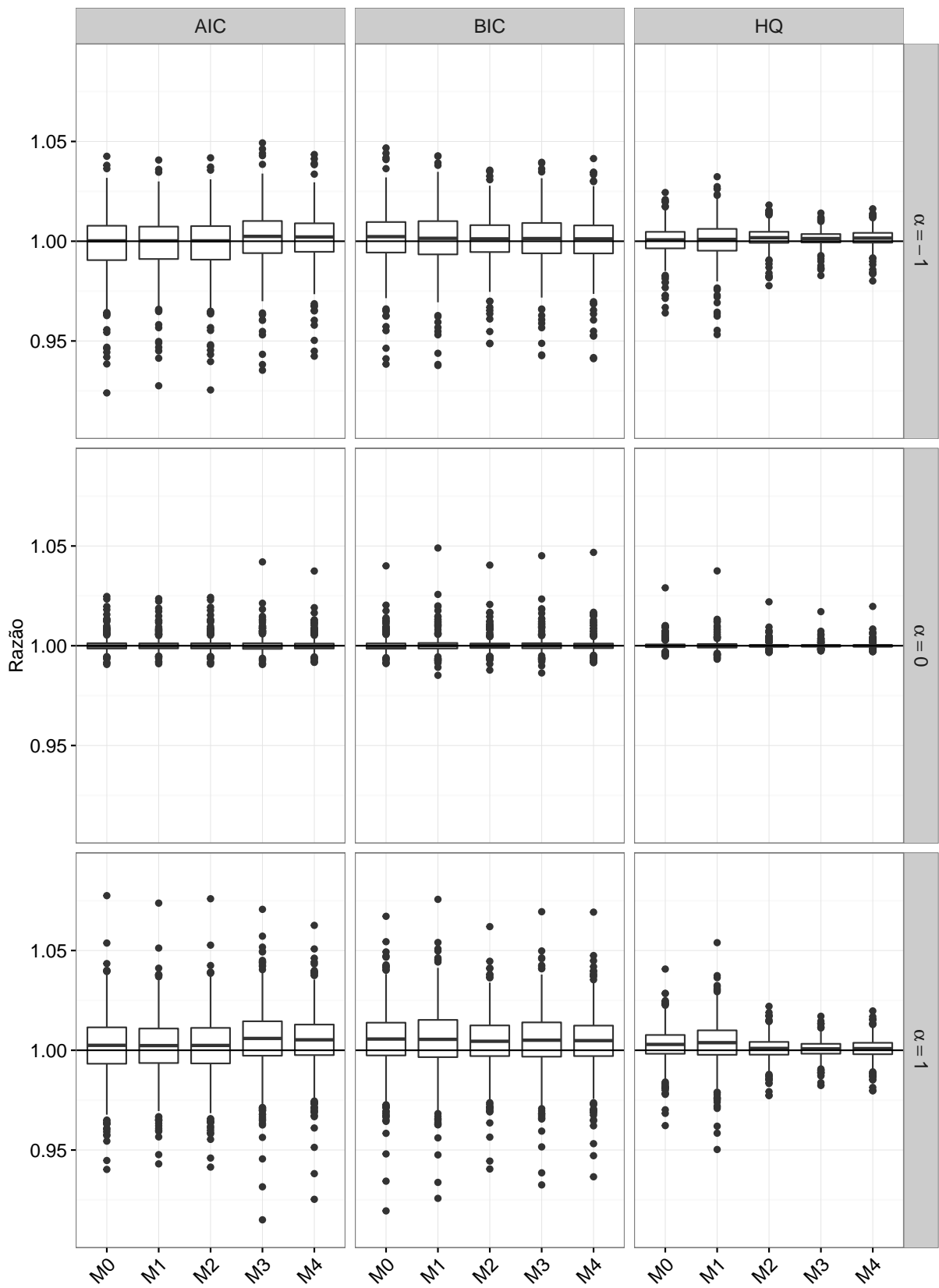


Figura 4.1: Razão entre o valor dos critérios sob ajuste incorreto pelo valor dos critérios sob ajuste correto para ajuste com abordagem frequentista.

Tabela 4.8: Proporção do número de vezes em que o critério indica para o modelo com ajuste correto.

	AIC	BIC	HQ	AIC	BIC	HQ	AIC	BIC	HQ
	$\alpha = -1$			$\alpha = 0$			$\alpha = 1$		
M0	0,506	0,506	0,506	0,466	0,466	0,466	0,576	0,576	0,576
M1	0,572	0,572	0,572	0,456	0,456	0,456	0,668	0,668	0,668
M2	0,558	0,558	0,558	0,524	0,524	0,524	0,640	0,640	0,640
M3	0,550	0,550	0,550	0,496	0,496	0,496	0,654	0,654	0,654
M4	0,678	0,678	0,678	0,497	0,497	0,497	0,576	0,576	0,576
Proporção do número de vezes em que os critérios indicam M1 entre M1, M2, M3 e M4.									
	AIC	BIC	HQ	AIC	BIC	HQ	AIC	BIC	HQ
	$\alpha = -1$			$\alpha = 0$			$\alpha = 1$		
M1	0,908	1	0,998	0,890	1,000	0,998	0,898	1,000	1,000
Proporção do número de vezes em que os critérios indicam M2 entre M1, M2, M3 e M4.									
	AIC	BIC	HQ	AIC	BIC	HQ	AIC	BIC	HQ
	$\alpha = -1$			$\alpha = 0$			$\alpha = 1$		
M2	0,092	0,000	0,002	0,110	0,000	0,002	0,102	0,000	0,000

4.1.2 Avaliação dos estimadores bayesianos

Nesta seção são apresentados os resultados obtidos através da abordagem bayesiana. Na geração das amostras *a posteriori* foi considerada uma cadeia de tamanho 155.000, sendo as primeiras 5.000 iterações descartadas para eliminar o efeito dos valores iniciais, consideramos um espaçamento de tamanho 30 para reduzir a autocorrelação, resultando em uma amostra efetiva de 5.000 observações.

Em todos os modelos consideramos as seguintes especificações para as distribuições *a priori*:

$$\begin{aligned}\beta_p^{(T)} &\sim \text{Normal}(0; 0, 01), \quad p = 1, \dots, P, \\ \beta_q^{(C)} &\sim \text{Normal}(0; 0, 01), \quad q = 1, \dots, Q, \\ \alpha &\sim \text{Normal}(0; 0, 01)\end{aligned}$$

e

$$\tau \sim \text{Exponencial}(1).$$

Para escolha da distribuição *a priori* para o parâmetro de precisão τ , realizamos um estudo Monte Carlo considerando as distribuições Gama(10; 10), Gama(1; 1) e Gama(0, 1; 0, 1) para τ , aplicamos o MEP com 10 intervalos para os tempos de falha e censura, em 500 conjuntos de dados com correlação positiva, negativa e nula entre os tempos T e C , e ajustamos o modelo supondo censura informativa e não informativa. Os resultados utilizando as distribuições Gama(1; 1) e Gama(10; 10) ficaram bem próximos, como podemos ver nas

Tabelas B.3, B.4 e B.5 (Apêndice B). Apesar de a distribuição Gama(10; 10) apresentar um desempenho levemente melhor, optamos por utilizar Gama(1; 1) por ter uma variância maior que Gama(10; 10).

No ajuste com o modelo não paramétrico (M4) consideramos as seguintes especificações *a priori*:

$$dH_0^{(T)} \left(y_j^{(T)} \right) \sim \text{Gama} \left(c \left[H_*^{(T)} \left(y_{j+1}^{(T)} \right) - H_*^{(T)} \left(y_j^{(T)} \right) \right], c \right), j = 1, \dots, J,$$

e

$$dH_0^{(C)} \left(y_l^{(C)} \right) \sim \text{Gama} \left(r \left[H_*^{(C)} \left(y_{l+1}^{(C)} \right) - H_*^{(C)} \left(y_l^{(C)} \right) \right], r \right), l = 1, \dots, L,$$

em que $r = 0,01$ e $c = 0,01$, constantes fixas, especificadas previamente e escolhidas mediante uma análise de sensibilidade. Consideramos $\left[H_*^{(T)} \left(y_{j+1}^{(T)} \right) - H_*^{(T)} \left(y_j^{(T)} \right) \right] = k \times \left[y_{j+1}^{(T)} - y_j^{(T)} \right]$ com $k = 1$ e intervalo $\left[y_{j+1}^{(T)} - y_j^{(T)} \right]$ fornecido pelos tempos de falha dos dados. Similarmente, $\left[H_*^{(C)} \left(y_{l+1}^{(C)} \right) - H_*^{(C)} \left(y_l^{(C)} \right) \right] = k \times \left[y_{l+1}^{(C)} - y_l^{(C)} \right]$ com $k = 1$ e intervalo $\left[y_{l+1}^{(C)} - y_l^{(C)} \right]$ extraído dos tempos de censura dos dados.

Para o MEP consideramos as seguintes distribuições *a priori*:

$$\lambda_j^{(T)} \sim \text{Gama}(0,01; 0,01), j = 1, \dots, b,$$

e

$$\lambda_l^{(C)} \sim \text{Gama}(0,01; 0,01), l = 1, \dots, d.$$

Realizamos uma análise de sensibilidade para avaliar qual seria o melhor número de intervalos para os tempos de falha e censura, dentre o conjunto de intervalos (1, 2, 4, 6, 8, 10, 20, 30, \dots , 100). Também avaliamos o caso em que o número de intervalos para os tempos de falha e censura foram, respectivamente, iguais ao número de falhas e censuras distintas. Para essa análise foram gerados dois bancos de dados com tamanho amostral igual a 400, para garantir que houvesse pelo menos uma falha em cada intervalo, com as mesmas especificações dos dados gerados com tamanho amostral igual a 200. Então, aplicamos o modelo aumentando o número de intervalos, e avaliamos os seguintes critérios de comparação de modelos: o Logaritmo da Pseudo Verossimilhança Marginal (LPML), o Critério de Informação da Desviância (DIC) e o critério de informação proposto por Watanabe (2010) (WAIC); mais detalhes sobre os critérios podem ser vistos no Apêndice C. Sendo o LPML, quanto maior melhor; DIC, quanto menor, melhor e WAIC, quanto maior, melhor o ajuste.

Para facilitar a comparação, apresentamos os três critérios na mesma escala, escre-

vendo os critérios LPML e WAIC na escala do DIC, ou seja, $LPML = -2LPML$ e $WAIC = -2WAIC$. Dessa forma os três critérios são avaliados da forma “quanto menor o valor do critério, melhor o ajuste do modelo”. Os resultados da comparação entre os critérios LPML, DIC e WAIC, para dois conjuntos de dados encontram-se na Tabela B.2 (Apêndice B). Podemos observar que os melhores ajustes foram com 10 intervalos para o primeiro conjunto de dados (Conjunto 1) e 20 intervalos para o segundo conjunto de dados (Conjunto 2). Portanto, nas simulações posteriores consideramos 10, 20 intervalos e número de intervalos igual ao número de falhas e censuras distintas.

No ajuste com o modelo Weibull consideramos as seguintes distribuições *a priori*:

$$\begin{aligned}\kappa^{(T)} &\sim \text{Gama}(1; 1), \\ \kappa^{(C)} &\sim \text{Gama}(1; 1), \\ \gamma^{(T)} &\sim \text{Gama}(0, 01; 0, 1)\end{aligned}$$

e

$$\gamma^{(C)} \sim \text{Gama}(0, 01; 0, 1).$$

A distribuição $\text{Gama}(0, 01; 0, 1)$ apresenta média=0,1 e variância=1, e apresentou um desempenho melhor que a distribuição Exponencial(1), para os parâmetros $\gamma^{(T)}$ e $\gamma^{(C)}$.

As Tabelas 4.9 e 4.10 apresentam os valores médios das 500 médias *a posteriori*, das 500 medianas *a posteriori*, das 500 modas *a posteriori* e dos 500 intervalos de 95% de credibilidade de mais alta densidade (HPD, *highest probability density*), bem como a probabilidade de cobertura e o valor médio dos critérios DIC, CPO e WAIC para os dados gerados com correlação positiva. Da Tabela 4.9 observa-se que o vício relativo foi menor quando consideramos a moda *a posteriori*, fato que também ocorreu nos demais cenários. O melhor modelo apontado pelos critérios é o modelo M0, seguido do modelo M1. De uma forma geral, os vícios relativos do modelo M4 são menores. A partir da Tabela 4.10 nota-se que o ajuste considerando censura não informativa, em geral, resulta em vícios maiores e probabilidade de cobertura menor, comparado aos resultados da Tabela 4.9. Os critérios também apontam para o ajuste com suposição de censura informativa.

Para os dados com correlação negativa, as estatísticas média, mediana, moda e intervalos HPD são apresentados nas Tabelas 4.11 e 4.12. Com base na Tabela 4.11, percebe-se que, ao analisarmos a moda *a posteriori* para os coeficientes associados aos tempos de falha, o modelo M1 apresentou vícios menores que os demais modelos. Os critérios indicam que o modelo M0 teve o melhor ajuste, seguido pelo modelo M1. Assim como no cenário com correlação positiva, as estimativas geradas pelos modelos supondo censura não informativa (Tabela 4.12) são mais viesadas que as estimativas produzidas pelos modelos que

supõem censura informativa (Tabela 4.11).

As Tabelas 4.13 e 4.14 apresentam as estimativas *a posteriori* para os dados com correlação nula. Da Tabela 4.13 observa-se que considerando a moda *a posteriori*, os vícios relativos do modelo M3 quase sempre foram menores do que os vícios dos demais modelos. Ao comparar o ajuste dos modelos com censura informativa (Tabela 4.13) com o ajuste dos modelos sem a suposição de censura informativa (Tabela 4.14), nota-se que os critérios apontam para os modelos com suposição de censura não informativa, porém em muitos cenários os valores dos critérios são bem próximos dos valores para o modelo com suposição de censura informativa.

Também calculamos a razão entre os valores dos critérios DIC, LPML e WAIC para os modelos que consideram incorretamente a suposição sobre a censura pelos valores dos critérios para os modelos que supõem corretamente o mecanismo de censura. Na Figura 4.2 exibe-se a distribuição dessa razão entre os critérios. Observa-se que nos cenários com correlação positiva e negativa, os critérios apontam para o modelo com ajuste correto, pois a razão sendo maior do que o valor um, mostra que o critério para o modelo com ajuste incorreto é maior.

A Tabela 4.15 apresenta a proporção do número de vezes em que os critérios apontam para o modelo com ajuste correto. A partir dessa tabela percebe-se que, em geral, em 80% das vezes ou mais, para os cenários com dependência entre os tempos de falha e censura, os critérios indicam que o modelo com censura informativa teve um ajuste melhor, ao passo que para o cenário com correlação nula, os modelos apresentam uma proporção menor, corroborando com as informações expostas na Figura 4.2.

4.1.3 Comparações entre os estimadores

Para comparar o ajuste dos modelos utilizando abordagens frequentista e bayesiana, calculamos a razão dos vícios relativos dos modelos com suposição incorreta para a censura pelos vícios relativos dos modelos com suposição correta, para ambas abordagens. Portanto, se a razão for maior que um, indica que o vício do modelo incorreto é maior. Na Figura 4.3 podemos analisar o comportamento dessa razão para os modelos sob abordagem frequentista (cinza claro) e sob abordagem bayesiana (cinza escuro).

Da Figura 4.3 observa-se que a razão é sempre maior que um para os coeficientes $\beta_1^{(T)}$ e $\beta_2^{(T)}$, nos cenários com correlação positiva e negativa, indicando que os modelos que consideram a censura informativa apresentam vício relativo, em geral, menor que os modelos que consideram censura não informativa. Para o parâmetro de precisão τ essa razão tende a ser menor que um, esse fato pode ser devido ao parâmetro α , que pode

Tabela 4.9: Ajuste utilizando modelos com abordagem bayesiana supondo censura informativa em dados gerados com correlação positiva ($\alpha = 1$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC
M0							
$\beta_1^{(T)}$	0,1	0,099 (-1,3)	0,098 (-1,6)	0,098 (-2,0)	0,026	[0,048; 0,150]	0,968 $\overline{LPML} = 345,086$
$\beta_2^{(T)}$	-1,4	-1,488 (-6,3)	-1,481 (-5,8)	-1,468 (-4,9)	0,314	[-2,107; -0,883]	0,944 $\overline{DIC} = 367,432$
$\beta_1^{(C)}$	0,2	0,210 (5,1)	0,209 (4,7)	0,208 (3,9)	0,033	[0,147; 0,274]	0,954 $\overline{WAIC} = 343,059$
$\beta_2^{(C)}$	1,2	1,300 (8,3)	1,290 (7,5)	1,278 (6,5)	0,354	[0,62; 1,999]	0,940
α	1	1,136 (13,6)	1,095 (9,5)	1,023 (2,3)	0,346	[0,527; 1,833]	0,960
τ	1	1,254 (25,4)	1,135 (13,5)	0,954 (-4,6)	0,570	[0,403; 2,378]	0,978
M1							
$\beta_1^{(T)}$	0,1	0,095 (-4,6)	0,095 (-4,8)	0,095 (-5,4)	0,026	[0,044; 0,147]	0,958 $\overline{LPML} = 360,856$
$\beta_2^{(T)}$	-1,4	-1,491 (-6,5)	-1,483 (-5,9)	-1,468 (-4,9)	0,316	[-2,112; -0,881]	0,942 $\overline{DIC} = 381,908$
$\beta_1^{(C)}$	0,2	0,207 (3,5)	0,206 (3,0)	0,204 (2,2)	0,033	[0,143; 0,273]	0,950 $\overline{WAIC} = 358,669$
$\beta_2^{(C)}$	1,2	1,305 (8,8)	1,295 (7,8)	1,276 (6,3)	0,354	[0,623; 2,002]	0,944
α	1	1,133 (13,3)	1,089 (8,9)	1,017 (1,7)	0,363	[0,492; 1,865]	0,970
τ	1	1,328 (32,6)	1,196 (19,6)	0,988 (-1,2)	0,62	[0,411; 2,548]	0,984
M2							
$\beta_1^{(T)}$	0,1	0,097 (-3,0)	0,097 (-3,3)	0,096 (-3,8)	0,026	[0,046; 0,149]	0,958 $\overline{LPML} = 373,733$
$\beta_2^{(T)}$	-1,4	-1,498 (-7,0)	-1,489 (-6,4)	-1,472 (-5,1)	0,317	[-2,123; -0,889]	0,942 $\overline{DIC} = 391,996$
$\beta_1^{(C)}$	0,2	0,209 (4,6)	0,208 (4,2)	0,207 (3,4)	0,034	[0,144; 0,276]	0,948 $\overline{WAIC} = 370,678$
$\beta_2^{(C)}$	1,2	1,299 (8,3)	1,289 (7,5)	1,273 (6,1)	0,354	[0,62; 1,997]	0,950
α	1	1,140 (14,0)	1,096 (9,6)	1,021 (2,1)	0,366	[0,496; 1,875]	0,972
τ	1	1,297 (29,7)	1,168 (16,8)	0,965 (-3,5)	0,606	[0,402; 2,491]	0,978
M3							
$\beta_1^{(T)}$	0,1	0,096 (-4,2)	0,096 (-4,5)	0,095 (-4,9)	0,026	[0,045; 0,148]	0,962 $\overline{LPML} = 611,816$
$\beta_2^{(T)}$	-1,4	-1,475 (-5,3)	-1,467 (-4,8)	-1,448 (-3,5)	0,316	[-2,096; -0,865]	0,94 $\overline{DIC} = 465,496$
$\beta_1^{(C)}$	0,2	0,209 (4,4)	0,208 (3,9)	0,206 (3,1)	0,034	[0,143; 0,275]	0,948 $\overline{WAIC} = 499,633$
$\beta_2^{(C)}$	1,2	1,291 (7,6)	1,282 (6,8)	1,261 (5,1)	0,349	[0,619; 1,98]	0,952
α	1	1,147 (14,7)	1,102 (10,2)	1,022 (2,2)	0,369	[0,496; 1,889]	0,968
τ	1	1,318 (31,8)	1,187 (18,7)	0,978 (-2,2)	0,617	[0,406; 2,536]	0,982
M4							
$\beta_1^{(T)}$	0,1	0,099 (-0,4)	0,099 (-0,6)	0,099 (-1,0)	0,026	[0,051; 0,149]	0,931 $\overline{LPML} = 1048,707$
$\beta_2^{(T)}$	-1,4	-1,47 (-5,0)	-1,463 (-4,5)	-1,452 (-3,7)	0,307	[-2,049; -0,908]	0,907 $\overline{DIC} = 926,208$
$\beta_1^{(C)}$	0,2	0,211 (5,7)	0,21 (5,1)	0,208 (4,2)	0,034	[0,149; 0,276]	0,919 $\overline{WAIC} = 962,288$
$\beta_2^{(C)}$	1,2	1,239 (3,3)	1,235 (2,9)	1,228 (2,3)	0,341	[0,602; 1,873]	0,889
α	1	1,149 (14,9)	1,107 (10,7)	1,041 (4,1)	0,356	[0,53; 1,867]	0,953
τ	1	1,256 (25,6)	1,135 (13,5)	0,974 (-2,6)	0,578	[0,413; 2,414]	0,976

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.10: Ajuste utilizando modelos com abordagem bayesiana supondo censura não informativa em dados gerados com correlação positiva ($\alpha = 1$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC	
M0								
$\beta_1^{(T)}$	0,1	0,087 (-12,5)	0,087 (-12,8)	0,087 (-13,2)	0,026	[0,036; 0,139]	0,926	$\overline{LPML} = 354,004$
$\beta_2^{(T)}$	-1,4	-1,624 (-1)	-1,615 (-15,4)	-1,596 (-14)	0,326	[-2,265; -0,995]	0,916	$\overline{DIC} = 378,284$
τ	1	1,307 (30,7)	1,175 (17,5)	0,967 (-3,3)	0,616	[0,402; 2,52]	0,974	$\overline{WAIC} = 350,183$
M1								
$\beta_1^{(T)}$	0,1	0,086 (-13,7)	0,086 (-13,9)	0,086 (-14,3)	0,027	[0,034; 0,139]	0,926	$\overline{LPML} = 369,097$
$\beta_2^{(T)}$	-1,4	-1,601 (-14,4)	-1,593 (-13,7)	-1,576 (-12,7)	0,328	[-2,247; -0,969]	0,928	$\overline{DIC} = 391,245$
τ	1	1,335 (33,5)	1,195 (19,6)	0,976 (-2,4)	0,644	[0,397; 2,601]	0,982	$\overline{WAIC} = 365,256$
M2								
$\beta_1^{(T)}$	0,1	0,088 (-12,4)	0,087 (-12,8)	0,087 (-13,3)	0,027	[0,035; 0,14]	0,936	$\overline{LPML} = 382,213$
$\beta_2^{(T)}$	-1,4	-1,611 (-15,1)	-1,603 (-14,5)	-1,585 (-13,2)	0,329	[-2,26; -0,976]	0,912	$\overline{DIC} = 401,617$
τ	1	1,306 (30,6)	1,169 (16,9)	0,960 (-3,9)	0,632	[0,386; 2,549]	0,978	$\overline{WAIC} = 377,343$
M3								
$\beta_1^{(T)}$	0,1	0,087 (-13,5)	0,086 (-13,8)	0,086 (-14,2)	0,027	[0,034; 0,14]	0,926	$\overline{LPML} = 619,546$
$\beta_2^{(T)}$	-1,4	-1,587 (-13,3)	-1,578 (-12,7)	-1,564 (-11,7)	0,328	[-2,234; -0,955]	0,924	$\overline{DIC} = 475,195$
τ	1	1,323 (32,3)	1,183 (18,3)	0,967 (-3,3)	0,642	[0,389; 2,587]	0,980	$\overline{WAIC} = 506,405$
M4								
$\beta_1^{(T)}$	0,1	0,089 (-10,7)	0,089 (-11,0)	0,088 (-11,5)	0,027	[0,036; 0,143]	0,938	$\overline{LPML} = 1084,849$
$\beta_2^{(T)}$	-1,4	-1,599 (-14,2)	-1,589 (-13,6)	-1,573 (-12,4)	0,333	[-2,254; -0,958]	0,916	$\overline{DIC} = 937,286$
τ	1	1,282 (28,2)	1,146 (14,6)	0,936 (-6,4)	0,625	[0,376; 2,513]	0,972	$\overline{WAIC} = 969,016$

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.11: Ajuste utilizando modelos com abordagem bayesiana supondo censura informativa em dados gerados com correlação negativa ($\alpha = -1$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC
M0							
$\beta_1^{(T)}$	0,1	0,102 (1,8)	0,102 (1,6)	0,101 (1,4)	0,025	[0,053; 0,151]	0,948 $\overline{LPML} = 336,492$
$\beta_2^{(T)}$	-1,4	-1,437 (-2,7)	-1,432 (-2,3)	-1,417 (-1,2)	0,297	[-2,021; -0,863]	0,952 $\overline{DIC} = 355,524$
$\beta_1^{(C)}$	0,2	0,211 (5,6)	0,210 (5,2)	0,209 (4,6)	0,032	[0,15; 0,273]	0,946 $\overline{WAIC} = 334,036$
$\beta_2^{(C)}$	1,2	1,268 (5,6)	1,260 (5,0)	1,247 (4,0)	0,327	[0,638; 1,912]	0,930
α	-1	-1,180 (-18,0)	-1,132 (-13,2)	-1,048 (-4,8)	0,377	[-1,936; -0,523]	0,962
τ	1	1,264 (26,4)	1,146 (14,6)	0,961 (-3,9)	0,575	[0,397; 2,397]	0,980
M1							
$\beta_1^{(T)}$	0,1	0,101 (0,7)	0,100 (0,5)	0,100 (0,1)	0,026	[0,051; 0,151]	0,950 $\overline{LPML} = 350,816$
$\beta_2^{(T)}$	-1,4	-1,419 (-1,4)	-1,413 (-0,9)	-1,404 (-0,3)	0,301	[-2,013; -0,838]	0,948 $\overline{DIC} = 368,386$
$\beta_1^{(C)}$	0,2	0,205 (2,7)	0,205 (2,4)	0,204 (1,8)	0,032	[0,144; 0,269]	0,940 $\overline{WAIC} = 348,254$
$\beta_2^{(C)}$	1,2	1,266 (5,5)	1,259 (4,9)	1,246 (3,8)	0,326	[0,637; 1,908]	0,940
α	-1	-1,166 (-16,6)	-1,116 (-11,6)	-1,027 (-2,7)	0,379	[-1,93; -0,505]	0,964
τ	1	1,316 (31,6)	1,192 (19,2)	1,001 (0,1)	0,605	[0,408; 2,508]	0,976
M2							
$\beta_1^{(T)}$	0,1	0,102 (1,7)	0,101 (1,5)	0,101 (1,1)	0,026	[0,052; 0,152]	0,946 $\overline{LPML} = 363,789$
$\beta_2^{(T)}$	-1,4	-1,431 (-2,2)	-1,424 (-1,7)	-1,41 (-0,7)	0,303	[-2,026; -0,845]	0,948 $\overline{DIC} = 379,559$
$\beta_1^{(C)}$	0,2	0,207 (3,7)	0,207 (3,3)	0,205 (2,7)	0,032	[0,145; 0,271]	0,932 $\overline{WAIC} = 360,411$
$\beta_2^{(C)}$	1,2	1,269 (5,8)	1,262 (5,1)	1,247 (3,9)	0,327	[0,638; 1,915]	0,928
α	-1	-1,127 (-12,7)	-1,082 (-8,2)	-1,037 (-3,7)	0,382	[-1,934; -0,504]	0,956
τ	1	1,324 (32,4)	1,199 (19,9)	0,999 (-0,1)	0,611	[0,401; 2,481]	0,970
M3							
$\beta_1^{(T)}$	0,1	0,101 (0,7)	0,100 (0,5)	0,100 (0,1)	0,026	[0,051; 0,151]	0,948 $\overline{LPML} = 617,058$
$\beta_2^{(T)}$	-1,4	-1,408 (-0,6)	-1,402 (-0,1)	-1,391 (0,7)	0,302	[-2,003; -0,825]	0,952 $\overline{DIC} = 457,462$
$\beta_1^{(C)}$	0,2	0,207 (3,5)	0,206 (3,2)	0,205 (2,5)	0,032	[0,145; 0,271]	0,942 $\overline{WAIC} = 498,024$
$\beta_2^{(C)}$	1,2	1,255 (4,6)	1,248 (4)	1,235 (2,9)	0,325	[0,629; 1,895]	0,938
α	-1	-1,134 (-13,4)	-1,086 (-8,6)	-1,001 (-0,1)	0,381	[-1,929; -0,502]	0,964
τ	1	1,315 (31,5)	1,190 (19,0)	0,992 (-0,8)	0,603	[0,397; 2,487]	0,968
M4							
$\beta_1^{(T)}$	0,1	0,102 (2,3)	0,102 (2,1)	0,102 (1,7)	0,026	[0,052; 0,153]	0,944 $\overline{LPML} = 1146,787$
$\beta_2^{(T)}$	-1,4	-1,433 (-2,4)	-1,426 (-1,9)	-1,413 (-0,9)	0,303	[-2,029; -0,848]	0,952 $\overline{DIC} = 983,574$
$\beta_1^{(C)}$	0,2	0,214 (6,9)	0,213 (6,5)	0,212 (5,8)	0,033	[0,149; 0,28]	0,952 $\overline{WAIC} = 1024,408$
$\beta_2^{(C)}$	1,2	1,262 (5,2)	1,255 (4,6)	1,241 (3,4)	0,334	[0,619; 1,92]	0,922
α	-1	-1,179 (-17,9)	-1,131 (-13,1)	-1,042 (-4,2)	0,400	[-1,977; -0,481]	0,968
τ	1	1,302 (30,2)	1,176 (17,6)	0,977 (-2,3)	0,605	[0,398; 2,495]	0,978

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.12: Ajuste utilizando modelos com abordagem bayesiana supondo censura não informativa em dados gerados com correlação negativa ($\alpha = -1$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC	
M0								
$\beta_1^{(T)}$	0,1	0,109 (9,3)	0,109 (9,1)	0,109 (8,8)	0,026	[0,059; 0,16]	0,940	$\overline{LPML} = 344,038$
$\beta_2^{(T)}$	-1,4	-1,401 (0,0)	-1,394 (0,4)	-1,376 (1,7)	0,304	[-1,999; -0,814]	0,956	$\overline{DIC} = 363,867$
τ	1	1,351 (35,1)	1,225 (22,5)	1,019 (1,9)	0,614	[0,428; 2,563]	0,982	$\overline{WAIC} = 340,017$
M1								
$\beta_1^{(T)}$	0,1	0,111 (10,7)	0,110 (10,5)	0,110 (10,1)	0,027	[0,059; 0,163]	0,932	$\overline{LPML} = 358,496$
$\beta_2^{(T)}$	-1,4	-1,373 (1,9)	-1,366 (2,4)	-1,353 (3,4)	0,311	[-1,983; -0,772]	0,948	$\overline{DIC} = 376,359$
τ	1	1,346 (34,5)	1,217 (21,7)	1,010 (1,0)	0,621	[0,417; 2,571]	0,984	$\overline{WAIC} = 354,276$
M2								
$\beta_1^{(T)}$	0,1	0,112 (12,1)	0,112 (11,8)	0,111 (11,4)	0,027	[0,06; 0,165]	0,930	$\overline{LPML} = 371,070$
$\beta_2^{(T)}$	-1,4	-1,388 (0,9)	-1,381 (1,4)	-1,367 (2,3)	0,313	[-2,003; -0,784]	0,942	$\overline{DIC} = 386,146$
τ	1	1,319 (31,9)	1,192 (19,2)	0,992 (-0,8)	0,610	[0,408; 2,519]	0,976	$\overline{WAIC} = 365,939$
M3								
$\beta_1^{(T)}$	0,1	0,110 (10,5)	0,110 (10,3)	0,109 (9,7)	0,027	[0,058; 0,163]	0,936	$\overline{LPML} = 624,138$
$\beta_2^{(T)}$	-1,4	-1,354 (3,3)	-1,349 (3,7)	-1,338 (4,6)	0,311	[-1,967; -0,756]	0,942	$\overline{DIC} = 465,284$
τ	1	1,327 (32,7)	1,195 (19,9)	0,999 (0,0)	0,618	[0,405; 2,536]	0,978	$\overline{WAIC} = 503,588$
M4								
$\beta_1^{(T)}$	0,1	0,113 (13,1)	0,113 (12,8)	0,112 (12,4)	0,027	[0,061; 0,166]	0,936	$\overline{LPML} = 1154,062$
$\beta_2^{(T)}$	-1,4	-1,399 (0,0)	-1,393 (0,5)	-1,379 (1,5)	0,313	[-2,016; -0,796]	0,948	$\overline{DIC} = 989,274$
τ	1	1,305 (30,5)	1,179 (17,9)	0,979 (-2,1)	0,607	[0,4; 2,502]	0,976	$\overline{WAIC} = 1029,700$

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.13: Ajuste utilizando modelos com abordagem bayesiana supondo censura informativa em dados gerados com correlação nula ($\alpha = 0$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC
M0							
$\beta_1^{(T)}$	0,1	0,099 (-0,6)	0,099 (-0,8)	0,099 (-1,2)	0,026	[0,049; 0,151]	0,950 $\overline{LPML} = 355,913$
$\beta_2^{(T)}$	-1,4	-1,448 (-3,4)	-1,441 (-2,9)	-1,433 (-2,3)	0,309	[-2,058; -0,852]	0,940 $\overline{DIC} = 381,727$
$\beta_1^{(C)}$	0,2	0,209 (4,9)	0,209 (4,6)	0,209 (4,3)	0,029	[0,154; 0,267]	0,944 $\overline{WAIC} = 351,983$
$\beta_2^{(C)}$	1,2	1,271 (5,9)	1,262 (5,2)	1,246 (3,9)	0,324	[0,646; 1,908]	0,938
α	0	-0,039 (-3,9)	-0,035 (-3,5)	-0,032 (-3,2)	0,256	[-0,549; 0,464]	0,954
τ	1	1,257 (25,7)	1,136 (13,6)	0,952 (-4,8)	0,580	[0,397; 2,401]	0,978
M1							
$\beta_1^{(T)}$	0,1	0,098 (-1,5)	0,098 (-1,8)	0,098 (-2,2)	0,027	[0,047; 0,151]	0,962 $\overline{LPML} = 370,589$
$\beta_2^{(T)}$	-1,4	-1,427 (-1,9)	-1,419 (-1,4)	-1,404 (-0,3)	0,312	[-2,042; -0,825]	0,944 $\overline{DIC} = 395,402$
$\beta_1^{(C)}$	0,2	0,208 (3,8)	0,207 (3,5)	0,206 (3,1)	0,030	[0,149; 0,267]	0,908 $\overline{WAIC} = 366,824$
$\beta_2^{(C)}$	1,2	1,282 (6,8)	1,273 (6,1)	1,258 (4,8)	0,324	[0,658; 1,92]	0,940
α	0	-0,041 (-4,1)	-0,037 (-3,7)	-0,037 (-3,7)	0,269	[-0,576; 0,49]	0,954
τ	1	1,351 (35,1)	1,216 (21,6)	1,001 (0,1)	0,636	[0,413; 2,603]	0,982
M2							
$\beta_1^{(T)}$	0,1	0,099 (-0,1)	0,099 (-0,4)	0,099 (-0,7)	0,027	[0,048; 0,152]	0,960 $\overline{LPML} = 383,434$
$\beta_2^{(T)}$	-1,4	-1,439 (-2,8)	-1,431 (-2,2)	-1,414 (-1,3)	0,314	[-2,057; -0,833]	0,948 $\overline{DIC} = 405,583$
$\beta_1^{(C)}$	0,2	0,209 (4,5)	0,208 (4,2)	0,207 (3,7)	0,031	[0,15; 0,269]	0,902 $\overline{WAIC} = 378,714$
$\beta_2^{(C)}$	1,2	1,275 (6,2)	1,266 (5,5)	1,248 (4,5)	0,324	[0,651; 1,913]	0,936
α	0	-0,029 (-2,9)	-0,026 (-2,6)	-0,021 (-2,1)	0,268	[-0,562; 0,498]	0,956
τ	1	1,319 (31,9)	1,187 (18,7)	0,983 (-1,7)	0,623	[0,405; 2,55]	0,984
M3							
$\beta_1^{(T)}$	0,1	0,099 (-1,5)	0,098 (-1,8)	0,098 (-2,3)	0,027	[0,046; 0,151]	0,958 $\overline{LPML} = 635,303$
$\beta_2^{(T)}$	-1,4	-1,412 (-0,9)	-1,405 (-0,4)	-1,392 (0,3)	0,312	[-2,025; -0,809]	0,952 $\overline{DIC} = 483,382$
$\beta_1^{(C)}$	0,2	0,209 (4,3)	0,208 (3,9)	0,207 (3,5)	0,031	[0,149; 0,269]	0,916 $\overline{WAIC} = 515,127$
$\beta_2^{(C)}$	1,2	1,266 (5,5)	1,257 (4,8)	1,243 (3,6)	0,32	[0,65; 1,898]	0,940
α	0	-0,032 (-3,2)	-0,029 (-2,9)	-0,026 (-2,6)	0,271	[-0,572; 0,5]	0,952
τ	1	1,339 (33,9)	1,203 (20,3)	0,992 (-0,8)	0,635	[0,405; 2,59]	0,986
M4							
$\beta_1^{(T)}$	0,1	0,100 (0,1)	0,099 (-0,2)	0,099 (-0,7)	0,027	[0,048; 0,153]	0,942 $\overline{LPML} = 1148,777$
$\beta_2^{(T)}$	-1,4	-1,437 (-2,6)	-1,429 (-2,1)	-1,419 (-1,4)	0,314	[-2,056; -0,831]	0,942 $\overline{DIC} = 994,518$
$\beta_1^{(C)}$	0,2	0,212 (6,1)	0,212 (5,8)	0,210 (5,2)	0,031	[0,152; 0,274]	0,934 $\overline{WAIC} = 1026,804$
$\beta_2^{(C)}$	1,2	1,257 (4,7)	1,248 (4,0)	1,232 (2,6)	0,329	[0,622; 1,905]	0,942
α	0	-0,023 (-2,3)	-0,02 (-2,0)	-0,018 (-1,8)	0,269	[-0,559; 0,507]	0,962
τ	1	1,276 (27,6)	1,147 (14,7)	0,939 (-6,0)	0,607	[0,386; 2,471]	0,978

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.14: Ajuste utilizando modelos com abordagem bayesiana supondo censura não informativa em dados gerados com correlação nula ($\alpha = 0$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC	
M0								
$\beta_1^{(T)}$	0,1	0,099 (-1,2)	0,099 (-1,4)	0,098 (-1,7)	0,026	[0,048; 0,15]	0,958	$\overline{LPML} = 355,289$
$\beta_2^{(T)}$	-1,4	-1,479 (-5,7)	-1,472 (-5,2)	-1,461 (-4,3)	0,309	[-2,089; -0,883]	0,940	$\overline{DIC} = 377,009$
τ	1	1,191 (19,1)	1,082 (8,2)	0,908 (-9,2)	0,532	[0,389; 2,239]	0,974	$\overline{WAIC} = 351,109$
M1								
$\beta_1^{(T)}$	0,1	0,099 (0,0)	0,099 (-0,2)	0,099 (-0,6)	0,027	[0,048; 0,152]	0,956	$\overline{LPML} = 370,144$
$\beta_2^{(T)}$	-1,4	-1,447 (-3,3)	-1,439 (-2,8)	-1,425 (-1,5)	0,314	[-2,065; -0,842]	0,952	$\overline{DIC} = 389,727$
τ	1	1,228 (22,8)	1,111 (11,1)	0,927 (-6,3)	0,563	[0,388; 2,334]	0,972	$\overline{WAIC} = 365,862$
M2								
$\beta_1^{(T)}$	0,1	0,101 (1,3)	0,101 (1,1)	0,101 (0,6)	0,027	[0,049; 0,154]	0,958	$\overline{LPML} = 383,001$
$\beta_2^{(T)}$	-1,4	-1,459 (-4,3)	-1,452 (-3,7)	-1,442 (-3,1)	0,315	[-2,081; -0,851]	0,942	$\overline{DIC} = 399,777$
τ	1	1,199 (19,9)	1,084 (8,4)	0,907 (-9,3)	0,551	[0,381; 2,285]	0,972	$\overline{WAIC} = 377,775$
M3								
$\beta_1^{(T)}$	0,1	0,099 (0,0)	0,099 (-0,3)	0,099 (-0,8)	0,027	[0,048; 0,153]	0,958	$\overline{LPML} = 633,294$
$\beta_2^{(T)}$	-1,4	-1,432 (-2,3)	-1,425 (-1,8)	-1,414 (-0,9)	0,314	[-2,049; -0,826]	0,950	$\overline{DIC} = 478,170$
τ	1	1,214 (21,4)	1,097 (9,7)	0,913 (-8,7)	0,561	[0,378; 2,317]	0,968	$\overline{WAIC} = 514,138$
M4								
$\beta_1^{(T)}$	0,1	0,102 (1,6)	0,101 (1,3)	0,101 (0,8)	0,027	[0,049; 0,155]	0,942	$\overline{LPML} = 1147,956$
$\beta_2^{(T)}$	-1,4	-1,462 (-4,4)	-1,455 (-3,9)	-1,446 (-3,3)	0,317	[-2,084; -0,85]	0,938	$\overline{DIC} = 988,194$
τ	1	1,159 (15,9)	1,047 (4,7)	0,878 (-12,2)	0,535	[0,365; 2,211]	0,964	$\overline{WAIC} = 1025,720$

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.15: Proporção do número de vezes em que o critério aponta para o modelo com ajuste correto.

	DIC	LPML	WAIC	DIC	LPML	WAIC	DIC	LPML	WAIC
	$\alpha = -1$			$\alpha = 0$			$\alpha = 1$		
M0	0,826	0,958	0,908	0,892	0,676	0,792	0,822	0,960	0,920
M1	0,812	0,970	0,926	0,940	0,658	0,830	0,820	0,952	0,904
M2	0,800	0,958	0,908	0,93	0,622	0,816	0,826	0,952	0,910
M3	0,794	0,702	0,902	0,822	0,560	0,730	0,824	0,702	0,886
M4	0,766	0,912	0,900	0,930	0,768	0,794	0,830	0,968	0,960

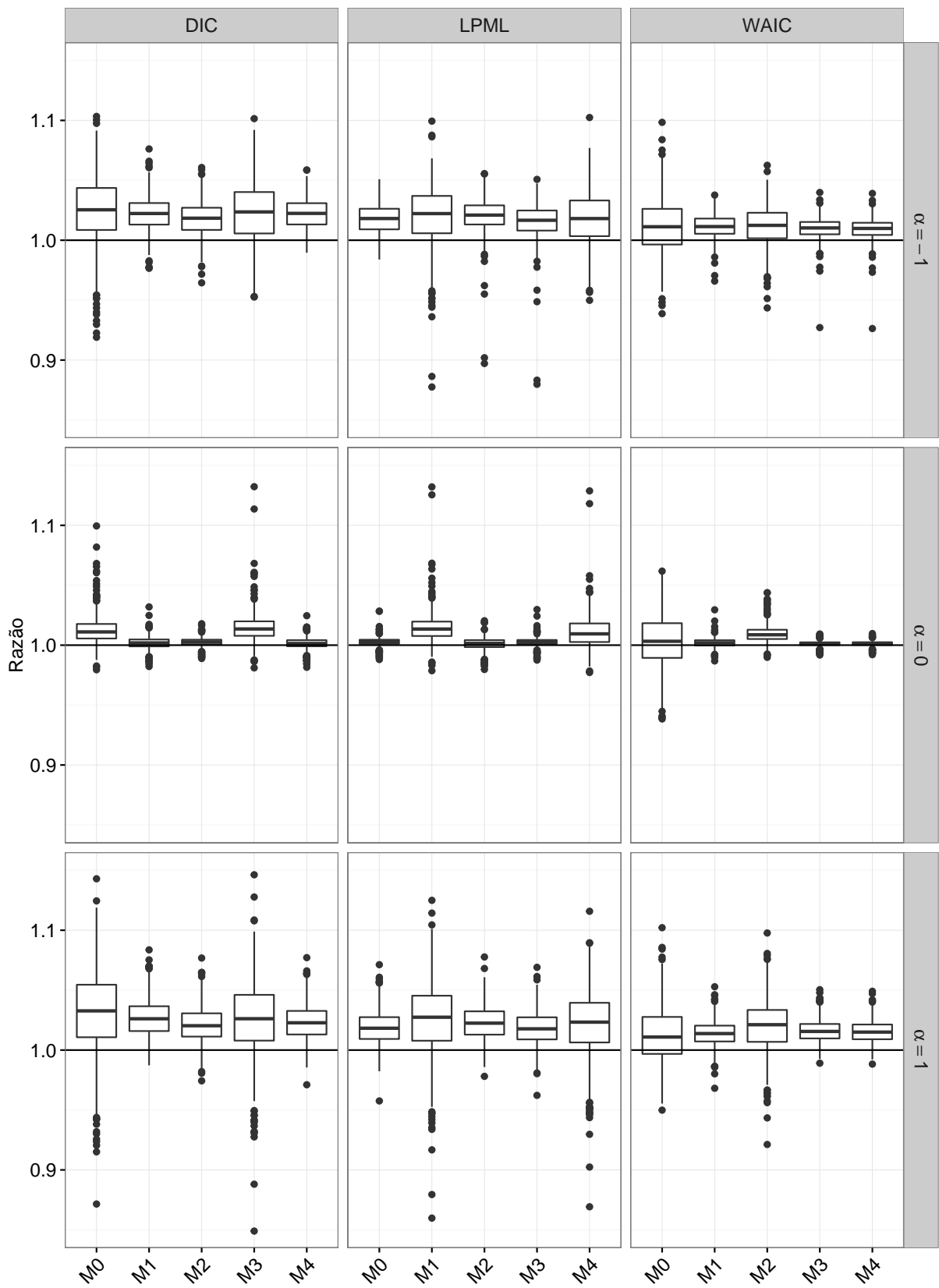


Figura 4.2: Razão entre o valor dos critérios sob ajuste incorreto pelo valor dos critérios sob ajuste correto em ajuste com abordagem bayesiana.

atrapalhar a estimação de τ .

A Tabela 4.16 apresenta a média e a mediana da razão entre o vício relativo do ajuste pelo modelo supondo incorretamente o mecanismo de censura pelo vício relativo do modelo supondo corretamente o mecanismo de censura. Nota-se que, mesmo observando a mediana, essa razão é maior que o valor 1 em todos os casos, exceto para o parâmetro de precisão τ e para o coeficiente de regressão $\beta_1^{(T)}$ pelo modelo M1 no cenário com correlação negativa. Por exemplo, se para o parâmetro $\beta_1^{(T)}$ no cenário com correlação positiva, pelo ajuste com o modelo M0, a razão é 1,209, indicando que o vício relativo do modelo que supõem censura não informativa é aproximadamente 21% maior que o vício para o modelo que supõem censura informativa, em metade das vezes. Percebe-se que os valores médios são bem maiores que 1, na maioria das vezes. A Tabela 4.17 apresenta a proporção do número de vezes que o vício relativo do modelo que supõem corretamente o mecanismo de censura é menor que o vício relativo do modelo que supõem incorretamente o mecanismo de censura.

Para comparar os modelos sob abordagens frequentista e bayesiana, também calculamos a razão entre o vício relativo dos modelos clássicos pelo vício relativo dos modelos bayesianos. No cálculo do vício relativo para os modelos sob abordagem bayesiana foi considerada a moda *a posteriori*. A Figura 4.4 apresenta a distribuição dessas razões para os três cenários. Nota-se que, de uma forma geral, os modelos clássicos fornecem vício relativo maior do que os modelos sob abordagem bayesiana.

A Tabela 4.18 apresenta a média e a mediana da razão entre os vícios dos modelos clássicos pelos vícios relativos do modelos bayesianos. A partir da Tabela 4.18 observa-se que na maioria das vezes essa razão é maior do que 1, indicando que o modelo sob abordagem bayesiana produz vício menor. A proporção do número de vezes em que o vício relativo é menor para o modelo bayesiano se encontra na Tabela 4.19.

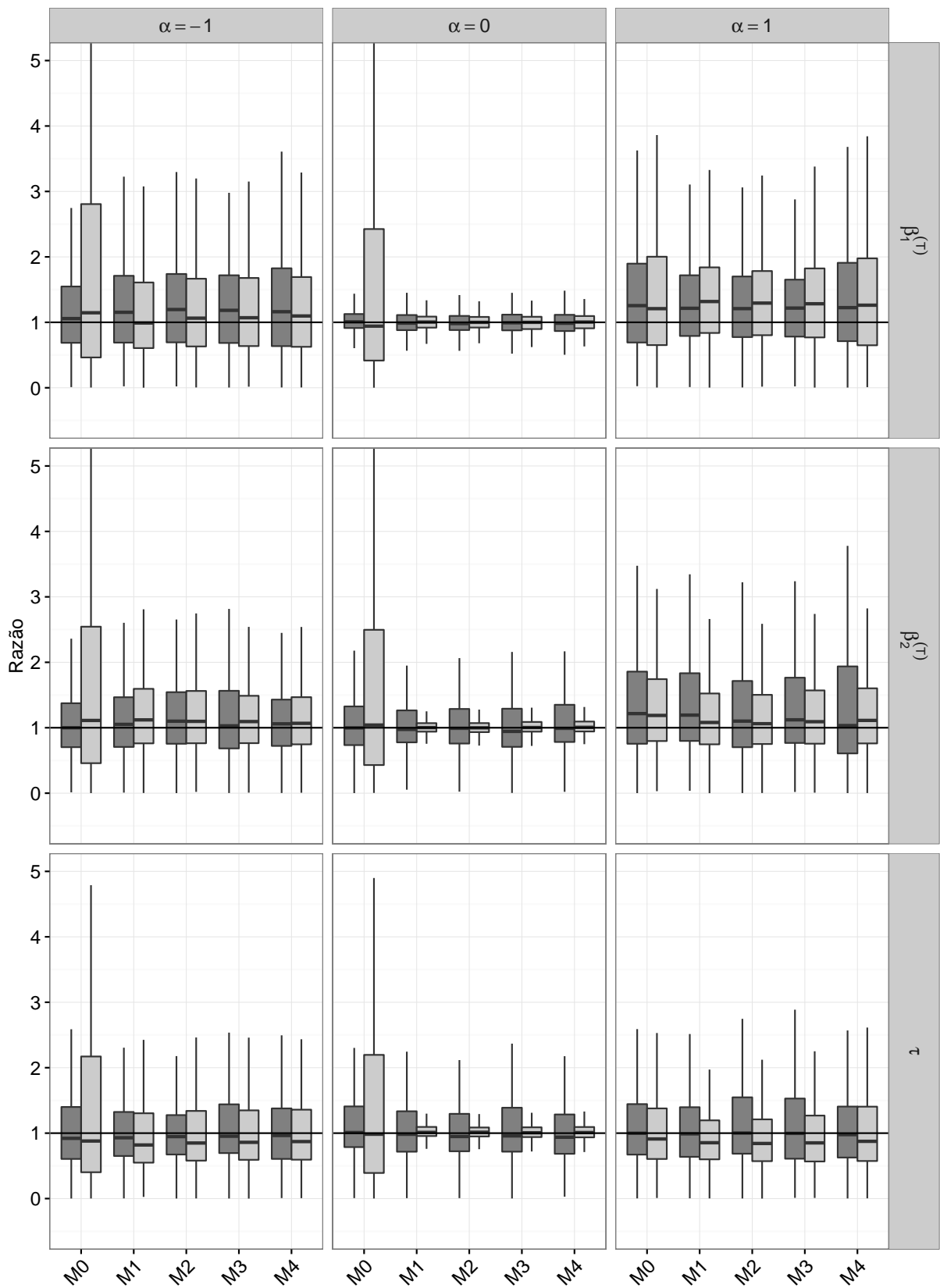


Figura 4.3: Razão entre o vício relativo absoluto do modelo incorreto (suposição incorreta sobre o mecanismo de censura) pelo vício relativo absoluto do modelo correto, para modelos sob abordagem bayesiana (cinza escuro) e abordagem frequentista (cinza claro).

Tabela 4.16: Razão entre o vício relativo do modelo supondo mecanismo de censura incorreto pelo vício relativo do modelo supondo mecanismo de censura corretamente.

	$\beta_1^{(T)}$	$\beta_2^{(T)}$	τ	$\beta_1^{(T)}$	$\beta_2^{(T)}$	τ						
Abordagem Bayesiana						Abordagem Frequentista						
	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana
$\alpha = -1$												
M0	2,126	1,057	1,534	1,001	1,925	0,921	3,324	1,121	2,223	1,083	5,598	1,095
M1	11,850	1,153	1,583	1,052	3,792	0,929	2,320	0,989	3,562	1,120	1,930	1,220
M2	2,258	1,195	5,486	1,099	6,011	0,9474	2,507	1,064	2,310	1,097	2,983	1,177
M3	2,188	1,183	2,478	1,029	1,858	0,953	3,023	1,070	1,822	1,094	89,763	1,163
M4	5,684	1,161	4,651	1,059	1,874	0,9643	3,262	1,097	7,345	1,067	2,394	1,147
$\alpha = 0$												
M0	1,300	1,007	1,339	0,998	2,905	1,009	1,124	1,001	1,198	1,003	1,131	0,994
M1	3,262	0,987	1,309	0,974	2,272	0,983	1,167	1,004	1,310	1,002	1,081	0,988
M2	5,980	0,981	1,865	0,992	2,422	0,948	1,604	1,001	1,093	0,997	1,105	0,989
M3	1,296	0,993	1,737	0,944	3,385	0,957	1,040	0,999	2,049	1,003	1,167	0,994
M4	1,330	0,986	2,060	0,991	2,106	0,937	1,087	1,004	1,119	1,006	1,066	0,990
$\alpha = 1$												
M0	3,695	1,255	4,869	1,216	2,567	0,998	6,215	1,209	2,333	1,188	2,211	1,097
M1	4,170	1,214	37,476	1,193	3,413	0,989	3,302	1,318	2,092	1,080	3,499	1,171
M2	3,249	1,209	3,272	1,101	1,863	1,002	2,616	1,294	2,877	1,061	7,690	1,187
M3	2,357	1,218	3,016	1,121	2,153	0,996	6,173	1,242	3,853	1,076	4,129	1,142
M4	2,982	1,224	2,612	1,033	2,056	0,977	2,826	1,262	3,450	1,112	3,630	1,142

¹Para abordagem frequentista, Média: média das 500 médias estimadas e Mediana: mediana das 500 médias estimadas. Para abordagem bayesiana, Média: média das 500 modas *a posteriori* e Mediana: mediana das 500 modas *a posteriori*.

Tabela 4.17: Proporção do número de vezes em que o vício relativo absoluto do modelo correto (suposição correta sobre o mecanismo de censura) é menor do que o vício relativo do modelo incorreto (suposição incorreta sobre o mecanismo de censura).

	$\beta_1^{(T)}$	$\beta_2^{(T)}$	τ	$\beta_1^{(T)}$	$\beta_2^{(T)}$	τ
	Abordagem Bayesiana			Abordagem Frequentista		
$\alpha = -1$						
M0	0,540	0,500	0,424	0,570	0,544	0,556
M1	0,558	0,548	0,426	0,496	0,582	0,658
M2	0,580	0,588	0,428	0,518	0,570	0,622
M3	0,566	0,532	0,468	0,532	0,572	0,622
M4	0,580	0,548	0,458	0,548	0,542	0,602
$\alpha = 0$						
M0	0,518	0,498	0,514	0,508	0,514	0,450
M1	0,466	0,476	0,474	0,530	0,512	0,452
M2	0,462	0,492	0,436	0,508	0,494	0,470
M3	0,476	0,438	0,424	0,500	0,514	0,464
M4	0,470	0,488	0,418	0,522	0,534	0,474
$\alpha = 1$						
M0	0,618	0,624	0,496	0,594	0,630	0,568
M1	0,648	0,628	0,482	0,704	0,554	0,640
M2	0,632	0,572	0,502	0,680	0,546	0,642
M3	0,646	0,570	0,498	0,606	0,538	0,568
M4	0,614	0,524	0,486	0,622	0,582	0,600

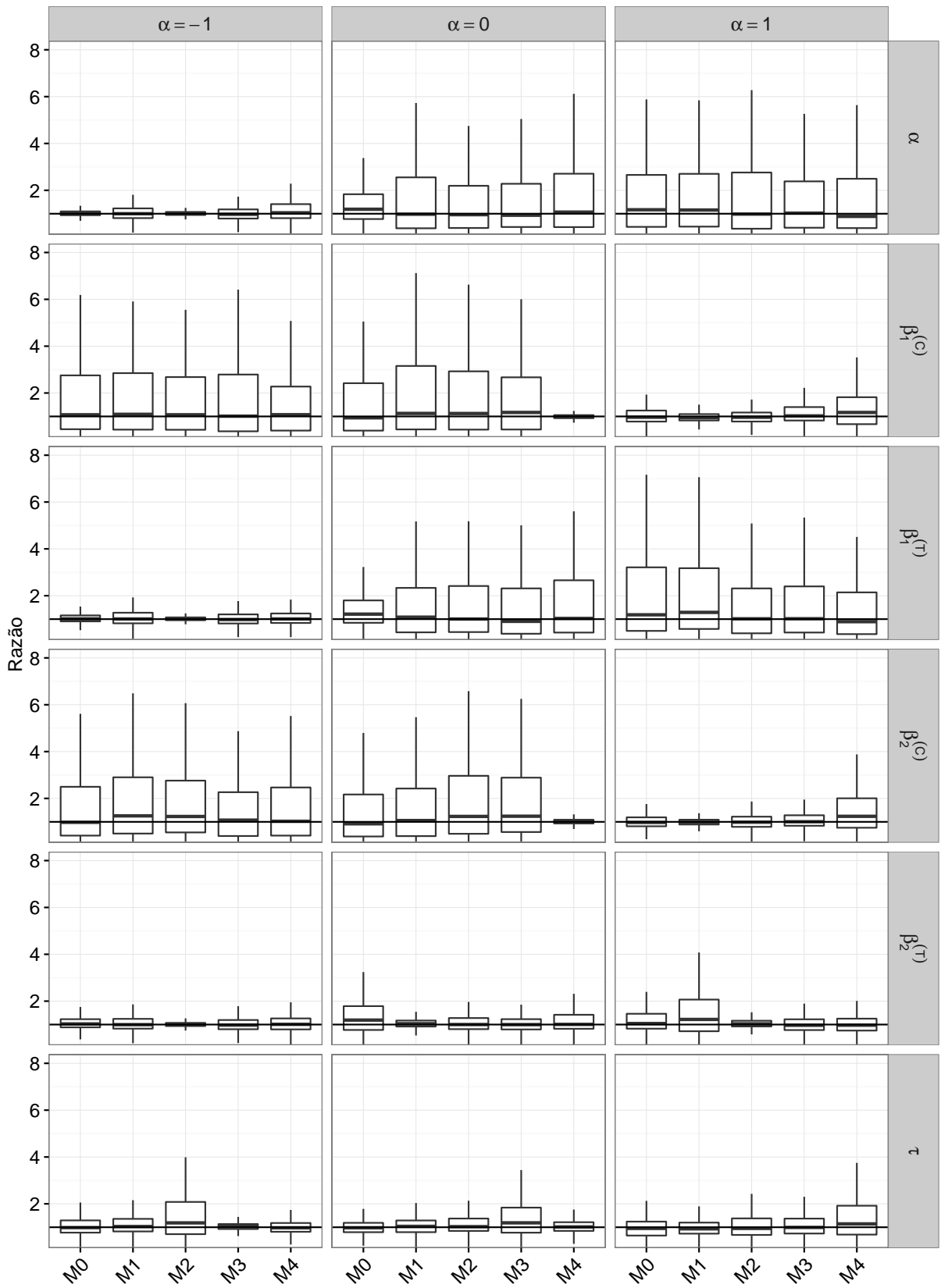


Figura 4.4: Razão entre vício relativo absoluto dos modelos com abordagem frequentista pelo vício relativo absoluto dos modelos com abordagem bayesiana, para censura informativa.

Tabela 4.18: Razão entre o vício relativo do modelo com abordagem frequentista pelo vício relativo do modelo com abordagem bayesiana, ambos supondo censura informativa.

$\beta_1^{(T)}$		$\beta_2^{(T)}$		$\beta_1^{(C)}$		$\beta_2^{(C)}$		α		τ		
Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	
$\alpha = -1$												
M0	4,699	1,078	4,970	1,019	4,726	0,994	3,927	0,912	5,313	1,042	4,644	1,288
M1	12,507	0,978	3,013	0,959	5,172	0,936	11,532	1,069	5,429	1,169	4,249	1,158
M2	4,594	0,983	12,927	1,024	5,709	0,886	135,741	1,071	4,636	1,090	4,842	1,071
M3	3,286	1,012	5,128	1,074	3,686	0,941	5,113	1,128	8,316	1,125	3,693	1,169
M4	1,544	0,995	1,901	0,965	2,002	0,945	2,176	0,965	1,871	1,023	2,480	1,169
$\alpha = 0$												
M0	3,041	1,112	18,347	1,045	4,768	1,147	5,444	0,969	4,098	0,916	4,595	1,223
M1	5,369	1,084	4,357	1,005	5,841	0,905	2,881	1,031	7,244	1,185	6,978	1,289
M2	2,946	1,035	3,451	1,012	2,904	0,898	11,458	0,988	60,348	1,288	3,612	1,230
M3	2,873	1,073	8,904	1,023	7,392	0,917	5,085	1,056	9,768	1,231	14,668	1,240
M4	1,179	1,006	1,844	0,974	1,171	0,984	1,355	0,984	2,931	1,005	3,416	1,236
$\alpha = 1$												
M0	1,733	1,026	1,445	0,995	1,197	1,001	1,331	0,977	4,574	1,009	2,705	1,190
M1	2,046	1,036	4,715	1,001	2,711	0,996	3,178	1,007	6,855	1,044	3,790	1,218
M2	1,409	1,031	1,613	0,968	1,598	0,973	2,595	0,984	1,951	1,031	3,037	1,187
M3	6,542	1,030	2,981	0,998	2,274	1,029	46,253	1,034	8,213	1,012	4,844	1,263
M4	1,609	1,014	1,549	0,950	1,913	0,933	2,887	0,950	2,418	0,993	2,174	1,144

Tabela 4.19: Proporção do número de vezes em que o vício relativo do modelo bayesiano é menor do que o vício relativo do modelo clássico, ambos com censura informativa.

	$\beta_1^{(T)}$	$\beta_2^{(T)}$	$\beta_1^{(C)}$	$\beta_2^{(C)}$	α	τ
$\alpha = -1$						
M0	0,528	0,510	0,498	0,478	0,516	0,572
M1	0,496	0,486	0,476	0,518	0,542	0,560
M2	0,498	0,504	0,472	0,524	0,520	0,528
M3	0,502	0,524	0,482	0,538	0,544	0,556
M4	0,478	0,460	0,402	0,458	0,548	0,592
$\alpha = 0$						
M0	0,538	0,508	0,538	0,494	0,476	0,548
M1	0,532	0,504	0,476	0,510	0,572	0,588
M2	0,518	0,502	0,466	0,498	0,570	0,562
M3	0,528	0,508	0,468	0,518	0,55	0,586
M4	0,520	0,466	0,450	0,482	0,506	0,614
$\alpha = 1$						
M0	0,532	0,492	0,508	0,464	0,518	0,624
M1	0,602	0,502	0,498	0,508	0,560	0,604
M2	0,594	0,462	0,488	0,484	0,536	0,604
M3	0,532	0,498	0,516	0,524	0,506	0,580
M4	0,542	0,438	0,426	0,452	0,490	0,578

4.2 Estudos Monte Carlo para modelos com censura informativa e fração de cura

Nesta seção apresentamos os resultados obtidos para o estudo de Monte Carlo, utilizando estimação sob abordagem frequentista e bayesiana, para dados de sobrevivência com fração de cura. Também apresentamos uma aplicação em dados sobre câncer de melanoma.

Para avaliar a eficiência dos estimadores para os modelos com fração de cura geramos 500 bancos de dados, cada um com tamanho amostral igual a 400, distribuídos aleatoriamente em 40 grupos com tamanhos iguais. Consideramos as distribuições Weibull($\kappa^{(T)}, \gamma^{(T)}$), para gerar os tempos de falha, e Weibull($\kappa^{(C)}, \gamma^{(C)}$) para gerar os tempos de censura, sendo $\kappa^{(T)} = 1$, $\kappa^{(C)} = 2$, $\gamma^{(T)} = 1$, $\gamma^{(C)} = 0,135$. Consideramos fração de cura para os tempos de falha, portanto utilizamos a relação $S_{pop}(t) = \exp(-\xi F(t))$, com $\xi = \exp(\beta^{(T)'X^{(T)} + W)$, consideramos uma covariável binária dada por $\text{Ber}(0,5)$. Logo, $\xi_i = \exp(\beta_0^{(T)} + \beta_1^{(T)}X_{ik}^{(T)} + W_k)$, os coeficientes de regressão como sendo $\beta^{(T)} = (-0,6, 0,4)$. A escolha para os valores dos parâmetros foi baseada em Cancho *et al.* (2011).

Consideramos duas covariáveis para os tempos de censura, uma contínua e outra binária, dadas por $N(0,1)$ e $\text{Ber}(0,5)$, respectivamente, com coeficientes de regressão $\beta^{(C)} = (0,1, -0,1)$. Também consideramos três cenários: dados gerados com correlação positiva, $\alpha = 1$, dados gerados com correlação negativa, $\alpha = -1$ e dados gerados supondo independência entre os tempos de falha e censura, $\alpha = 0$.

Para gerar os tempos observados e as variáveis indicadoras de falha e censura, seguimos os seguintes passos:

1. Gere $u_i, v_i \sim U[0,1]$;
2. Gere $s_i \sim U[0,10]$;
3. Seja $\gamma^{(C)*} \leftarrow \gamma^{(C)} \exp(\beta^{(C)'X^{(C)}}) \exp(\alpha w)$;
4. $\xi_i \leftarrow \exp(\beta_0^{(T)} + \beta_1^{(T)}X_{ik}^{(T)}) \exp(w_k)$;
5. Se $u_i \leftarrow \exp(-\xi_i)$, então $t_i \leftarrow \inf$;
 Se $u_i > \exp(-\xi_i)$, então $t_i \leftarrow \left\{ -\log \left(\left(\frac{\log(u_i)}{\xi_i} \right) + 1 \right) / \gamma^{(T)} \right\}^{1/\kappa^{(T)}}$;
6. $c_i \leftarrow \left\{ \frac{-\log(v_i)}{\gamma^{(C)*}} \right\}^{1/\kappa^{(C)}}$;
7. $y_i \leftarrow \min(t_i, c_i, s_i)$;
8. Se $y_i = t_i$, então $\delta_i^{(T)} = 1$, caso contrário $\delta_i^{(T)} = 0$;
 Se $y_i = c_i$, então $\delta_i^{(C)} = 1$, caso contrário $\delta_i^{(C)} = 0$.

Em cada conjunto de dados foi considerado aproximadamente 40% de falhas, 35% de censura informativa e 25% de censura administrativa. Também foi considerado por volta de 50% de indivíduos curados.

Vamos denotar os modelos por

M5: Modelo Weibull com fração de cura;

M6: Modelo MEP com 5 intervalos e fração de cura;

M7: Modelo MEP com 10 intervalos e fração de cura.

4.2.1 Avaliação dos estimadores de máxima verossimilhança

Nesta seção são expostos os resultados obtidos através dos ajustes considerando abordagem frequentista na estimação. O número de réplicas de Monte Carlo, bem como o critério de parada adotado, são análogos aos utilizados na Seção 4.1.1.

A Tabela 4.23 apresenta a média, erro padrão, probabilidade de cobertura e valores médios dos critérios para o cenário com correlação positiva. Observa-se que, em geral, o vício relativo é menor quando utilizamos o modelo *M5*, uma vez que os dados são gerados a partir de um modelo Weibull. Também nota-se que, as estimativas obtidas pelo modelo *M6* são mais próximas dos verdadeiros valores para os coeficientes $\beta_0^{(T)}$ e $\beta_1^{(T)}$ que as estimativas obtidas pelo modelo *M7*. Ao comparar os resultados dos ajustes para os modelos que não consideram a censura como informativa, percebe-se que o vício relativo para o intercepto $\beta_0^{(T)}$ é maior que o vício para o modelos que supõem censura não informativa. A probabilidade de cobertura foi menor ao se considerar suposição de independência entre os tempos de falha e censura.

O ajuste para o cenário com correlação negativa é apresentado na Tabela 4.24, considerando a suposição de censura informativa e não informativa. Nesta tabela observa-se que o vício relativo para os parâmetros $\beta_0^{(T)}$ e $\beta_1^{(T)}$ foi menor para o modelo *M6*. Também observa-se que o vício para o intercepto $\beta_0^{(T)}$ nos modelos que não supõem censura informativa é o valor do vício relativo é bem maior que para os modelos que levam em consideração a censura informativa. A probabilidade de cobertura para $\beta_0^{(T)}$ também diminui ao se utilizar modelos sem a suposição de censura informativa.

A Tabela 4.25 apresenta o ajuste para os dados com correlação nula. Nota-se que, de forma geral, o valor dos vícios relativos é menor nos modelos que supõem censura não informativa. A probabilidade de cobertura, para quase todos os parâmetros, fica próximo de 95%, tanto no ajuste supondo censura informativa, quanto no ajuste com a suposição de censura não informativa.

Tabela 4.20: Ajuste com abordagem frequentista para os dados com correlação positiva ($\alpha = 1$).

	Verdadeiro	Média(VR%)	EPE	EPL	IC	PC	
Censura Informativa							
M5							
$\beta_0^{(T)}$	-0,6	-0,601 (-0,1)	0,274	0,214	[-1,139; -0,063]	0,974	$\overline{AIC} = 457,325$
$\beta_1^{(T)}$	0,4	0,402 (0,4)	0,167	0,164	[0,074; 0,729]	0,948	$\overline{BIC} = 477,282$
$\beta_1^{(C)}$	0,1	0,101 (1,2)	0,089	0,090	[-0,074; 0,276]	0,950	$\overline{HQ} = 465,228$
$\beta_2^{(C)}$	-0,1	-0,097 (2,8)	0,179	0,178	[-0,449; 0,255]	0,950	
α	1	1,014 (1,4)	0,213	0,209	[0,597; 1,432]	0,956	
σ^2	1	0,994 (-0,6)	0,328	0,342	[0,351; 1,638]	0,952	
M6							
$\beta_0^{(T)}$	-0,6	-0,607 (-1,2)	0,228	0,239	[-1,055; -0,159]	0,952	$\overline{AIC} = 459,630$
$\beta_1^{(T)}$	0,4	0,402 (0,6)	0,168	0,164	[0,074; 0,731]	0,948	$\overline{BIC} = 491,562$
$\beta_1^{(C)}$	0,1	0,095 (-4,8)	0,085	0,089	[-0,071; 0,262]	0,950	$\overline{HQ} = 472,276$
$\beta_2^{(C)}$	-0,1	-0,091 (9,1)	0,172	0,176	[-0,427; 0,245]	0,940	
α	1	0,878 (-12,2)	0,192	0,181	[0,502; 1,254]	0,930	
σ^2	1	1,010 (1,0)	0,340	0,334	[0,343; 1,678]	0,950	
M7							
$\beta_0^{(T)}$	-0,6	-0,579 (3,5)	0,319	0,220	[-1,206; 0,047]	0,983	$\overline{AIC} = 559,627$
$\beta_1^{(T)}$	0,4	0,405 (1,4)	0,166	0,163	[0,079; 0,732]	0,949	$\overline{BIC} = 611,517$
$\beta_1^{(C)}$	0,1	0,099 (-1,4)	0,088	0,089	[-0,074; 0,271]	0,946	$\overline{HQ} = 580,176$
$\beta_2^{(C)}$	-0,1	-0,095 (5,1)	0,177	0,177	[-0,441; 0,251]	0,946	
α	1	0,948 (-5,2)	0,202	0,186	[0,553; 1,343]	0,952	
σ^2	1	1,008 (0,8)	0,356	0,322	[0,31; 1,705]	0,965	
Censura Não Informativa							
M5							
$\beta_0^{(T)}$	-0,6	-0,651 (-8,5)	0,219	0,235	[-1,080; -0,222]	0,932	$\overline{AIC} = 455,625$
$\beta_1^{(T)}$	0,4	0,404 (0,9)	0,169	0,168	[0,071; 0,737]	0,944	$\overline{BIC} = 475,582$
σ^2	1	0,972 (-2,8)	0,339	0,344	[0,308; 1,637]	0,950	$\overline{HQ} = 463,528$
M6							
$\beta_0^{(T)}$	-0,6	-0,644 (-7,4)	0,222	0,221	[-1,08; -0,208]	0,938	$\overline{AIC} = 468,661$
$\beta_1^{(T)}$	0,4	0,403 (0,9)	0,169	0,166	[0,072; 0,735]	0,942	$\overline{BIC} = 500,593$
σ^2	1	0,967 (-3,3)	0,336	0,316	[0,309; 1,625]	0,956	$\overline{HQ} = 481,306$
M7							
$\beta_0^{(T)}$	-0,6	-0,631 (-5,2)	0,229	0,212	[-1,08; -0,182]	0,941	$\overline{AIC} = 555,011$
$\beta_1^{(T)}$	0,4	0,404 (1,1)	0,169	0,166	[0,072; 0,736]	0,941	$\overline{BIC} = 606,900$
σ^2	1	0,966 (-3,4)	0,335	0,316	[0,309; 1,623]	0,953	$\overline{HQ} = 575,559$

¹VR: vício relativo, EPE: erro padrão empírico calculado a partir das 500 estimativas, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.21: Ajuste com abordagem frequentista para os dados com correlação negativa ($\alpha = -1$).

	Verdadeiro	Média (VR%)	EPE	EPL	IC	PC	
Censura Informativa							
M5							
$\beta_0^{(T)}$	-0,6	-0,589 (1,8)	0,321	0,229	[-1,218; 0,039]	0,986	$\overline{AIC} = 455,243$
$\beta_1^{(T)}$	0,4	0,402 (0,5)	0,168	0,163	[0,072; 0,731]	0,940	$\overline{BIC} = 475,199$
$\beta_1^{(C)}$	0,1	0,100 (0,2)	0,087	0,085	[-0,071; 0,271]	0,950	$\overline{HQ} = 463,146$
$\beta_2^{(C)}$	-0,1	-0,098 (1,7)	0,167	0,168	[-0,427; 0,230]	0,938	
α	-1	-1,012 (-1,2)	0,209	0,215	[-1,423; -0,601]	0,950	
σ^2	1	0,987 (-1,3)	0,337	0,335	[0,326; 1,647]	0,950	
M6							
$\beta_0^{(T)}$	-0,6	-0,595 (0,8)	0,229	0,227	[-1,044; -0,146]	0,940	$\overline{AIC} = 455,957$
$\beta_1^{(T)}$	0,4	0,401 (0,2)	0,168	0,161	[0,072; 0,73]	0,940	$\overline{BIC} = 487,889$
$\beta_1^{(C)}$	0,1	0,095 (-4,6)	0,083	0,084	[-0,067; 0,258]	0,956	$\overline{HQ} = 468,603$
$\beta_2^{(C)}$	-0,1	-0,093 (7,0)	0,159	0,166	[-0,406; 0,22]	0,950	
α	-1	-0,959 (4,1)	0,195	0,220	[-1,34; -0,577]	0,954	
σ^2	1	0,991 (-0,9)	0,334	0,342	[0,336; 1,647]	0,950	
M7							
$\beta_0^{(T)}$	-0,6	-0,571 (4,9)	0,273	0,229	[-1,106; -0,036]	0,972	$\overline{AIC} = 524,411$
$\beta_1^{(T)}$	0,4	0,399 (-0,3)	0,166	0,161	[0,074; 0,723]	0,945	$\overline{BIC} = 576,300$
$\beta_1^{(C)}$	0,1	0,099 (-0,6)	0,084	0,085	[-0,065; 0,264]	0,949	$\overline{HQ} = 544,959$
$\beta_2^{(C)}$	-0,1	-0,098 (2,1)	0,162	0,168	[-0,416; 0,22]	0,941	
α	-1	-0,993 (0,7)	0,204	0,223	[-1,392; -0,593]	0,951	
σ^2	1	0,981 (-1,9)	0,331	0,344	[0,332; 1,63]	0,953	
Censura Não Informativa							
M5							
$\beta_0^{(T)}$	-0,6	-0,547 (8,9)	0,220	0,234	[-0,978; -0,115]	0,930	$\overline{AIC} = 451,786$
$\beta_1^{(T)}$	0,4	0,403 (0,7)	0,169	0,165	[0,073; 0,733]	0,940	$\overline{BIC} = 471,744$
σ^2	1	0,934 (-6,6)	0,333	0,323	[0,281; 1,587]	0,960	$\overline{HQ} = 459,689$
M6							
$\beta_0^{(T)}$	-0,6	-0,529 (11,8)	0,233	0,209	[-0,986; -0,073]	0,928	$\overline{AIC} = 455,192$
$\beta_1^{(T)}$	0,4	0,402 (0,4)	0,169	0,164	[0,071; 0,732]	0,944	$\overline{BIC} = 487,124$
σ^2	1	0,926 (-7,4)	0,328	0,306	[0,283; 1,569]	0,948	$\overline{HQ} = 467,838$
M7							
$\beta_0^{(T)}$	-0,6	-0,515 (14,2)	0,224	0,265	[-0,953; -0,076]	0,936	$\overline{AIC} = 644,210$
$\beta_1^{(T)}$	0,4	0,398 (-0,5)	0,165	0,164	[0,077; 0,723]	0,945	$\overline{BIC} = 696,099$
σ^2	1	0,932 (-6,8)	0,324	0,438	[0,297; 1,567]	0,959	$\overline{HQ} = 664,759$

¹VR: vício relativo, EPE: erro padrão empírico calculado a partir das 500 estimativas, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

Tabela 4.22: Ajuste com abordagem frequentista para os dados com correlação nula ($\alpha = 0$).

	Verdadeiro	Média (VR%)	EPE	EPL	IC	PC	
Censura Informativa							
M5							
$\beta_0^{(T)}$	-0,6	-0,589 (1,8)	0,383	0,240	[-1,339; 0,161]	0,996	$\overline{AIC} = 460,774$
$\beta_1^{(T)}$	0,4	0,405 (1,3)	0,167	0,165	[0,078; 0,733]	0,946	$\overline{BIC} = 480,732$
$\beta_1^{(C)}$	0,1	0,102 (1,9)	0,080	0,080	[-0,055; 0,259]	0,956	$\overline{HQ} = 468,678$
$\beta_2^{(C)}$	-0,1	-0,096 (3,6)	0,162	0,159	[-0,414; 0,221]	0,960	
α	0	0,003 (0,3)	0,127	0,128	[-0,246; 0,251]	0,942	
σ^2	1	0,984 (-1,6)	0,339	0,331	[0,32; 1,649]	0,956	
M6							
$\beta_0^{(T)}$	-0,6	-0,592 (1,3)	0,232	0,237	[-1,047; -0,137]	0,938	$\overline{AIC} = 463,930$
$\beta_1^{(T)}$	0,4	0,404 (1,0)	0,167	0,164	[0,077; 0,731]	0,950	$\overline{BIC} = 495,862$
$\beta_1^{(C)}$	0,1	0,098 (-1,7)	0,078	0,081	[-0,055; 0,251]	0,954	$\overline{HQ} = 476,576$
$\beta_2^{(C)}$	-0,1	-0,093 (6,7)	0,156	0,159	[-0,4; 0,213]	0,956	
α	0	-0,033 (-3,3)	0,114	0,125	[-0,257; 0,192]	0,938	
σ^2	1	0,979 (-2,1)	0,335	0,328	[0,323; 1,635]	0,956	
M7							
$\beta_0^{(T)}$	-0,6	-0,555 (7,6)	0,418	0,228	[-1,373; 0,264]	0,995	$\overline{AIC} = 535,665$
$\beta_1^{(T)}$	0,4	0,404 (0,9)	0,166	0,164	[0,078; 0,729]	0,948	$\overline{BIC} = 587,554$
$\beta_1^{(C)}$	0,1	0,102 (1,7)	0,079	0,081	[-0,054; 0,257]	0,952	$\overline{HQ} = 556,214$
$\beta_2^{(C)}$	-0,1	-0,097 (2,5)	0,159	0,159	[-0,408; 0,213]	0,954	
α	0	-0,018 (-1,8)	0,119	0,127	[-0,253; 0,217]	0,943	
σ^2	1	0,977 (-2,3)	0,333	0,319	[0,325; 1,629]	0,959	
Censura Não Informativa							
M5							
$\beta_0^{(T)}$	-0,6	-0,602 (-0,4)	0,235	0,234	[-1,063; -0,142]	0,948	$\overline{AIC} = 460,604$
$\beta_1^{(T)}$	0,4	0,405 (1,2)	0,167	0,165	[0,078; 0,732]	0,946	$\overline{BIC} = 480,561$
σ^2	1	0,983 (-1,7)	0,337	0,341	[0,322; 1,644]	0,954	$\overline{HQ} = 468,507$
M6							
$\beta_0^{(T)}$	-0,6	-0,590 (1,7)	0,233	0,221	[-1,047; -0,133]	0,934	$\overline{AIC} = 473,551$
$\beta_1^{(T)}$	0,4	0,404 (1,0)	0,166	0,164	[0,078; 0,73]	0,950	$\overline{BIC} = 505,4828$
σ^2	1	0,977 (-2,3)	0,333	0,320	[0,324; 1,63]	0,954	$\overline{HQ} = 486,196$
M7							
$\beta_0^{(T)}$	-0,6	-0,563 (6,2)	0,233	0,267	[-1,32; 0,194]	0,998	$\overline{AIC} = 741,948$
$\beta_1^{(T)}$	0,4	0,404 (1,1)	0,166	0,165	[0,079; 0,729]	0,950	$\overline{BIC} = 793,837$
σ^2	1	0,977 (-2,3)	2,294	0,501	[0,321; 1,633]	0,955	$\overline{HQ} = 762,496$

¹VR: vício relativo, EPE: erro padrão empírico calculado a partir das 500 estimativas, EPL: média dos 500 erros padrão assintóticos, IC: intervalo de confiança empírico calculado a partir do EPE, \overline{AIC} , \overline{BIC} e \overline{HQ} : média dos 500 AIC, BIC e HQ, respectivamente.

4.2.2 Abordagem bayesiana

Nesta seção são apresentados os ajustes para os modelos M5, M6 e M7 utilizando abordagem bayesiana na estimação. Para a escolha das especificações para as distribuições *a priori* para os coeficientes de regressão, para o parâmetro de dependência e para o parâmetro de precisão, consideramos

$$\beta_p^{(T)} \sim N(0; 0, 01), \quad p = 1, \dots, P,$$

$$\beta_q^{(C)} \sim N(0; 0, 01), \quad q = 1, \dots, Q,$$

$$\alpha \sim N(0; 0, 01)$$

e

$$\tau \sim \text{Exponencial}(1).$$

Para o modelo Weibull realizamos uma análise de sensibilidade e optamos por utilizar as seguintes distribuições

$$\kappa^{(T)} \sim \text{Gama}(1, 1),$$

$$\kappa^{(C)} \sim \text{Gama}(1, 1),$$

$$\gamma^{(T)} \sim \text{Gama}(1, 1),$$

$$\gamma^{(C)} \sim \text{Gama}(1, 1).$$

As especificações *a priori* utilizadas para o Modelo Exponencial por Partes foram

$$\lambda_j^{(T)} \sim \text{Gama}(0, 1; 0, 1), \quad j = 1, \dots, b,$$

e

$$\lambda_l^{(C)} \sim \text{Gama}(0, 1; 0, 1), \quad l = 1, \dots, d.$$

A Tabela 4.26 apresenta a média, mediana, moda e desvio padrão *a posteriori*, bem como o intervalo HPD e o valor médio dos critérios. Observa-se que, mais uma vez o moda *a posteriori* apresenta menor vício relativo, fato que também ocorre nos demais cenários. Portanto, considerando-se a moda *a posteriori*, nota-se que o vício relativo para o intercepto $\beta_0^{(T)}$ foi menor quando utilizamos o modelo M6, porém este modelo (M6) não estimou bem o parâmetro da dependência α . Os critérios de comparação de modelos indicam que o modelo M5 apresentou um ajuste melhor. Ao compararmos o ajuste supondo censura informativa em relação ao ajuste supondo censura não informativa,

percebe-se que os critérios DIC, LPML e WAIC apontam para os modelos com censura informativa.

O ajuste para os dados gerados com correlação negativa é apresentado na Tabela 4.27. Nesta tabela percebe-se que, em geral, o modelo M5 resulta em vícios menores, os critérios também apontam que o melhor modelo é M5, como esperado, uma vez que os dados foram gerados de uma modelo Weibull. Comparando-se os ajustes com censura informativa em relação aos ajustes com censura não-informativa, novamente, os vícios relativos são maiores quando se supõem independência entre os tempos de falha e censura. Os critérios DIC, LPML e WAIC também visam os modelos com censura informativa.

A Tabela 4.28 apresenta o ajuste para os dados gerados com correlação nula. Desta tabela percebe-se que os critérios apontam para os modelos que supõem censura não informativa. Os vícios relativos obtidos em ambas as suposições, com e sem censura informativa, são bem próximos.

A Tabela 4.29 apresenta a proporção do número de vezes que o critério aponta para o modelo com ajuste correto. Percebe-se que, nos cenários em que são consideradas correlação positiva ou negativa na geração dos dados, os critérios acertam em pelo menos 80% das vezes. Ao passo que no cenário com correlação nula essa proporção diminui para os critérios LPML e WAIC. A distribuição dessas razões pode ser vista na Figura 4.5.

4.2.3 Comparação entre os estimadores

Da mesma forma como na comparação entre as abordagens frequentista e bayesiana para os modelos sem fração de cura, calculamos a razão entre o vício relativo dos modelos supondo ajuste incorreto pelo vício relativo dos modelos supondo ajuste correto. A distribuição desses valores encontra-se na Figura 4.6.

Da Figura 4.6 observa-se que, exceto para o modelo M7, os valores da razão variam menos para a abordagem frequentista (em cinza claro). De maneira geral, a razão para os modelos sob abordagem bayesiana (em cinza escuro) mostra-se similar nos três cenários.

A Tabela 4.30 mostra as proporções do número de vezes que o vício relativo é menor para o modelo que supõem ajuste correto. A Tabela 4.31 apresenta a média e a mediana para razão entre os vícios relativos, para as abordagens bayesiana e frequentista. Observa-se que a média desses valores é sempre maior que 1, porém quando analisamos a mediana percebemos que, em geral, esses valores são próximos de 1. Da Tabela 4.30, percebe-se que as proporções estão próximas de 50%, porém da Tabela 4.31, nota-se que os vícios relativos podem ser altos, quando utilizamos os modelos com suposição incorreta sobre a censura.

Tabela 4.23: Ajuste utilizando modelos com abordagem bayesiana em dados gerados com correlação positiva ($\alpha = 1$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	DP	HPD	PC	
Censura Informativa								
M5								
$\beta_0^{(T)}$	-0,6	-0,566 (5,7)	-0,567 (5,5)	-0,568 (5,4)	0,228	[-1,011; -0,122]	0,944	DIC = 496,704
$\beta_1^{(T)}$	0,4	0,403 (0,7)	0,402 (0,5)	0,402 (0,6)	0,163	[0,085; 0,722]	0,942	\overline{LPML} = 484,162
$\beta_1^{(C)}$	0,1	0,101 (0,8)	0,101 (0,8)	0,101 (0,1)	0,089	[-0,074; 0,275]	0,952	\overline{WAIC} = 483,769
$\beta_2^{(C)}$	-0,1	-0,110 (-10,3)	-0,109 (-9,8)	-0,108 (-8,2)	0,177	[-0,456; 0,234]	0,954	
α	1	1,034 (3,4)	1,021 (2,1)	0,996 (-0,4)	0,193	[0,673; 1,417]	0,940	
τ	1	1,118 (11,8)	1,069 (6,9)	0,979 (-2,1)	0,355	[0,513; 1,825]	0,948	
M6								
$\beta_0^{(T)}$	-0,6	-0,534 (11)	-0,566 (5,7)	-0,579 (3,6)	0,319	[-1,065; 0,074]	0,966	\overline{DIC} = 502,426
$\beta_1^{(T)}$	0,4	0,404 (1,1)	0,404 (0,9)	0,403 (0,6)	0,164	[0,084; 0,724]	0,944	\overline{LPML} = 489,154
$\beta_1^{(C)}$	0,1	0,096 (-4,2)	0,096 (-4,3)	0,096 (-4,2)	0,089	[-0,077; 0,268]	0,962	\overline{WAIC} = 488,705
$\beta_2^{(C)}$	-0,1	-0,095 (5,2)	-0,094 (5,6)	-0,093 (7,0)	0,175	[-0,438; 0,247]	0,958	
α	1	0,903 (-9,7)	0,892 (-10,8)	0,869 (-13,1)	0,175	[0,576; 1,253]	0,852	
τ	1	1,102 (10,2)	1,051 (5,1)	0,962 (-3,8)	0,352	[0,503; 1,805]	0,938	
M7								
$\beta_0^{(T)}$	-0,6	-0,43 (28,4)	-0,499 (16,8)	-0,555 (7,4)	0,398	[-1,064; 0,425]	0,966	\overline{DIC} = 506,557
$\beta_1^{(T)}$	0,4	0,403 (0,8)	0,403 (0,7)	0,402 (0,5)	0,164	[0,085; 0,723]	0,944	\overline{LPML} = 493,624
$\beta_1^{(C)}$	0,1	0,099 (-0,9)	0,099 (-0,9)	0,099 (-0,8)	0,09	[-0,075; 0,274]	0,950	\overline{WAIC} = 493,171
$\beta_2^{(C)}$	-0,1	-0,101 (0,1)	-0,101 (0,3)	-0,099 (0,7)	0,177	[-0,445; 0,244]	0,942	
α	1	0,978 (-2,2)	0,966 (-3,4)	0,945 (-5,5)	0,186	[0,63; 1,348]	0,912	
τ	1	1,110 (11,0)	1,060 (6,0)	0,971 (-2,9)	0,353	[0,507; 1,813]	0,946	
Censura Não Informativa								
M5								
$\beta_0^{(T)}$	-0,6	-0,639 (-6,6)	-0,638 (-6,4)	-0,633 (-5,5)	0,224	[-1,078; -0,203]	0,940	\overline{DIC} = 506,794
$\beta_1^{(T)}$	0,4	0,407 (1,6)	0,406 (1,5)	0,405 (1,3)	0,166	[0,083; 0,731]	0,938	\overline{LPML} = 496,406
τ	1	1,115 (11,5)	1,063 (6,3)	0,975 (-2,5)	0,363	[0,499; 1,838]	0,950	\overline{WAIC} = 495,562
M6								
$\beta_0^{(T)}$	-0,6	-0,610 (-1,7)	-0,622 (-3,6)	-0,624 (-4,1)	0,266	[-1,094; -0,108]	0,960	\overline{DIC} = 509,979
$\beta_1^{(T)}$	0,4	0,407 (1,7)	0,406 (1,6)	0,403 (0,6)	0,167	[0,082; 0,732]	0,940	\overline{LPML} = 499,672
τ	1	1,113 (11,3)	1,060 (6,0)	0,968 (-3,2)	0,363	[0,494; 1,833]	0,948	\overline{WAIC} = 498,813
M7								
$\beta_0^{(T)}$	-0,6	-0,528 (11,9)	-0,579 (3,5)	-0,609 (-1,6)	0,357	[-1,118; 0,192]	0,972	\overline{DIC} = 515,267
$\beta_1^{(T)}$	0,4	0,406 (1,6)	0,405 (1,5)	0,404 (0,9)	0,166	[0,082; 0,731]	0,938	\overline{LPML} = 505,115
τ	1	1,112 (11,2)	1,059 (5,9)	0,973 (-2,7)	0,362	[0,496; 1,835]	0,948	\overline{WAIC} = 504,218

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.24: Ajuste utilizando modelos com abordagem bayesiana em dados gerados com correlação negativa ($\alpha = -1$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	EP	HPD	PC	
Censura Informativa								
M5								
$\beta_0^{(T)}$	-0,6	-0,605 (-0,8)	-0,603 (-0,5)	-0,596 (0,7)	0,229	[-1,054; -0,161]	0,946	DIC = 492,615
$\beta_1^{(T)}$	0,4	0,403 (0,9)	0,403 (0,7)	0,401 (0,2)	0,162	[0,09; 0,72]	0,934	\overline{LPML} = 480,605
$\beta_1^{(C)}$	0,1	0,099 (-0,3)	0,099 (-0,3)	0,099 (-0,3)	0,084	[-0,065; 0,264]	0,946	\overline{WAIC} = 480,037
$\beta_2^{(C)}$	-0,1	-0,109 (-9,1)	-0,109 (-8,9)	-0,106 (-6,3)	0,167	[-0,433; 0,217]	0,946	
α	-1	-1,039 (-3,9)	-1,021 (-2,1)	-0,989 (1,0)	0,211	[-1,46; -0,65]	0,954	
τ	1	1,161 (16,1)	1,107 (10,7)	1,013 (1,3)	0,373	[0,525; 1,903]	0,950	
M6								
$\beta_0^{(T)}$	-0,6	-0,522 (13)	-0,554 (7,7)	-0,568 (5,3)	0,322	[-1,06; 0,102]	0,950	\overline{DIC} = 494,614
$\beta_1^{(T)}$	0,4	0,403 (0,8)	0,403 (0,7)	0,401 (0,3)	0,162	[0,087; 0,72]	0,934	\overline{LPML} = 483,519
$\beta_1^{(C)}$	0,1	0,096 (-4,0)	0,096 (-4)	0,096 (-3,7)	0,084	[-0,068; 0,26]	0,952	\overline{WAIC} = 482,909
$\beta_2^{(C)}$	-0,1	-0,097 (2,5)	-0,097 (2,6)	-0,1 (0,1)	0,166	[-0,423; 0,227]	0,958	
α	-1	-1,010 (-1,0)	-0,992 (0,8)	-0,957 (4,3)	0,207	[-1,423; -0,63]	0,946	
τ	1	1,135 (13,5)	1,082 (8,2)	0,994 (-0,6)	0,366	[0,511; 1,864]	0,956	
M7								
$\beta_0^{(T)}$	-0,6	-0,427 (29,2)	-0,504 (16,2)	-0,554 (8,0)	0,423	[-1,088; 0,462]	0,970	\overline{DIC} = 500,565
$\beta_1^{(T)}$	0,4	0,403 (0,9)	0,403 (0,7)	0,402 (0,1)	0,162	[0,088; 0,72]	0,942	\overline{LPML} = 488,916
$\beta_1^{(C)}$	0,1	0,099 (-0,8)	0,099 (-0,8)	0,098 (-0,4)	0,085	[-0,066; 0,264]	0,948	\overline{WAIC} = 488,290
$\beta_2^{(C)}$	-0,1	-0,103 (-3,5)	-0,103 (-3,3)	-0,104 (-2,8)	0,167	[-0,429; 0,222]	0,944	
α	-1	-1,045 (-4,5)	-1,026 (-2,6)	-0,992 (0,5)	0,215	[-1,472; -0,65]	0,952	
τ	1	1,145 (14,5)	1,091 (9,1)	1,002 (0,2)	0,370	[0,515; 1,881]	0,952	
Censura Não Informativa								
M5								
$\beta_0^{(T)}$	-0,6	-0,526 (12,3)	-0,528 (12,0)	-0,531 (11,4)	0,239	[-0,991; -0,061]	0,948	\overline{DIC} = 502,119
$\beta_1^{(T)}$	0,4	0,406 (1,5)	0,405 (1,3)	0,405 (1,2)	0,164	[0,087; 0,727]	0,948	\overline{LPML} = 492,086
τ	1	1,169 (16,9)	1,113 (11,3)	1,014 (1,4)	0,384	[0,517; 1,934]	0,958	\overline{WAIC} = 491,106
M6								
$\beta_0^{(T)}$	-0,6	-0,446 (25,7)	-0,489 (18,4)	-0,512 (14,7)	0,355	[-1,022; 0,244]	0,958	\overline{DIC} = 506,072
$\beta_1^{(T)}$	0,4	0,406 (1,6)	0,406 (1,5)	0,403 (0,9)	0,164	[0,086; 0,728]	0,936	\overline{LPML} = 495,663
τ	1	1,173 (17,3)	1,116 (11,6)	1,022 (2,2)	0,386	[0,515; 1,939]	0,958	\overline{WAIC} = 494,659
M7								
$\beta_0^{(T)}$	-0,6	-0,324 (46,3)	-0,422 (29,6)	-0,486 (19,4)	0,479	[-1,043; 0,695]	0,956	\overline{DIC} = 510,674
$\beta_1^{(T)}$	0,4	0,406 (1,5)	0,406 (1,4)	0,407 (1,5)	0,164	[0,086; 0,727]	0,938	\overline{LPML} = 500,711
τ	1	1,175 (17,6)	1,119 (11,9)	1,021 (2,2)	0,386	[0,517; 1,943]	0,960	\overline{WAIC} = 499,662

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

Tabela 4.25: Ajuste utilizando modelos com abordagem bayesiana em dados gerados com correlação nula ($\alpha = 0$).

	Verdadeiro	Média(VR%)	Mediana(VR%)	Moda(VR%)	EP	HPD	PC	
Censura Informativa								
M5								
$\beta_0^{(T)}$	-0,6	-0,579 (3,5)	-0,581 (3,2)	-0,583 (2,8)	0,233	[-1,033; -0,126]	0,950	DIC = 513,963
$\beta_1^{(T)}$	0,4	0,406 (1,6)	0,406 (1,5)	0,407 (1,7)	0,164	[0,088; 0,727]	0,946	LPML = 501,792
$\beta_1^{(C)}$	0,1	0,101 (1,1)	0,101 (1,1)	0,101 (1,4)	0,08	[-0,055; 0,258]	0,960	WAIC = 500,905
$\beta_2^{(C)}$	-0,1	-0,106 (-6,3)	-0,106 (-6,0)	-0,105 (-4,6)	0,159	[-0,417; 0,203]	0,952	
α	0	0,004 (0,4)	0,005 (0,5)	0,007 (0,7)	0,133	[-0,258; 0,265]	0,958	
τ	1	1,144 (14,4)	1,091 (9,1)	0,994 (-0,6)	0,371	[0,512; 1,883]	0,950	
M6								
$\beta_0^{(T)}$	-0,6	-0,475 (20,9)	-0,534 (11,0)	-0,562 (6,3)	0,395	[-1,072; 0,299]	0,960	\overline{DIC} = 517,448
$\beta_1^{(T)}$	0,4	0,407 (1,7)	0,406 (1,5)	0,405 (1,3)	0,164	[0,087; 0,728]	0,946	\overline{LPML} = 505,308
$\beta_1^{(C)}$	0,1	0,098 (-1,6)	0,098 (-1,6)	0,099 (-0,9)	0,081	[-0,059; 0,255]	0,962	\overline{WAIC} = 504,398
$\beta_2^{(C)}$	-0,1	-0,096 (3,6)	-0,096 (3,8)	-0,095 (5,0)	0,159	[-0,409; 0,214]	0,954	
α	0	-0,035 (-3,5)	-0,033 (-3,3)	-0,031 (-3,1)	0,130	[-0,293; 0,219]	0,960	
τ	1	1,144 (14,4)	1,090 (9,0)	0,997 (-0,3)	0,371	[0,511; 1,881]	0,952	
M7								
$\beta_0^{(T)}$	-0,6	-0,345 (42,4)	-0,45 (25,2)	-0,536 (11,6)	0,490	[-1,084; 0,78]	0,964	\overline{DIC} = 522,331
$\beta_1^{(T)}$	0,4	0,409 (1,5)	0,408 (1,4)	0,408 (1,2)	0,164	[0,089; 0,729]	0,948	\overline{LPML} = 510,279
$\beta_1^{(C)}$	0,1	0,101 (1,0)	0,101 (1,0)	0,101 (-1,0)	0,081	[-0,057; 0,258]	0,954	\overline{WAIC} = 509,353
$\beta_2^{(C)}$	-0,1	-0,101 (-1,6)	-0,101 (-1,5)	-0,101 (-3,0)	0,159	[-0,418; 0,206]	0,944	
α	0	-0,013 (-1,3)	-0,012 (-1,2)	-0,009 (-1,1)	0,133	[-0,277; 0,247]	0,962	
τ	1	1,158 (15,8)	1,104 (10,4)	1,001 (1,0)	0,376	[0,518; 1,907]	0,954	
Censura Não Informativa								
M5								
$\beta_0^{(T)}$	-0,6	-0,586 (2,4)	-0,587 (2,1)	-0,586 (2,4)	0,236	[-1,044; -0,125]	0,948	\overline{DIC} = 511,276
$\beta_1^{(T)}$	0,4	0,408 (2,0)	0,408 (1,8)	0,409 (2,3)	0,164	[0,088; 0,728]	0,942	\overline{LPML} = 501,528
τ	1	1,103 (10,3)	1,053 (5,3)	0,964 (-3,6)	0,356	[0,495; 1,814]	0,948	\overline{WAIC} = 500,595
M6								
$\beta_0^{(T)}$	-0,6	-0,472 (21,4)	-0,536 (10,6)	-0,569 (5,1)	0,407	[-1,082; 0,327]	0,954	\overline{DIC} = 514,870
$\beta_1^{(T)}$	0,4	0,408 (1,9)	0,407 (1,9)	0,405 (1,3)	0,165	[0,087; 0,729]	0,948	\overline{LPML} = 505,093
τ	1	1,104 (10,4)	1,053 (5,3)	0,964 (-3,6)	0,356	[0,494; 1,813]	0,952	\overline{WAIC} = 504,149
M7								
$\beta_0^{(T)}$	-0,6	-0,343 (41,9)	-0,456 (23,8)	-0,536 (10,7)	0,499	[-1,085; 0,711]	0,974	\overline{DIC} = 519,729
$\beta_1^{(T)}$	0,4	0,407 (1,8)	0,407 (1,7)	0,405 (1,1)	0,164	[0,088; 0,729]	0,944	\overline{LPML} = 510,067
τ	1	1,106 (10,6)	1,055 (5,5)	0,962 (-3,8)	0,357	[0,497; 1,817]	0,956	\overline{WAIC} = 509,094

¹VR: vício relativo, Média: média das 500 médias *a posteriori*, Mediana: média das 500 medianas *a posteriori*, Moda: média das 500 modas *a posteriori*, DP: média dos 500 desvios padrão *a posteriori*, PC: probabilidade de cobertura a partir do 500 HPD, \overline{LPML} , \overline{DIC} e \overline{WAIC} : média dos 500 LPML, DIC e WAIC, respectivamente.

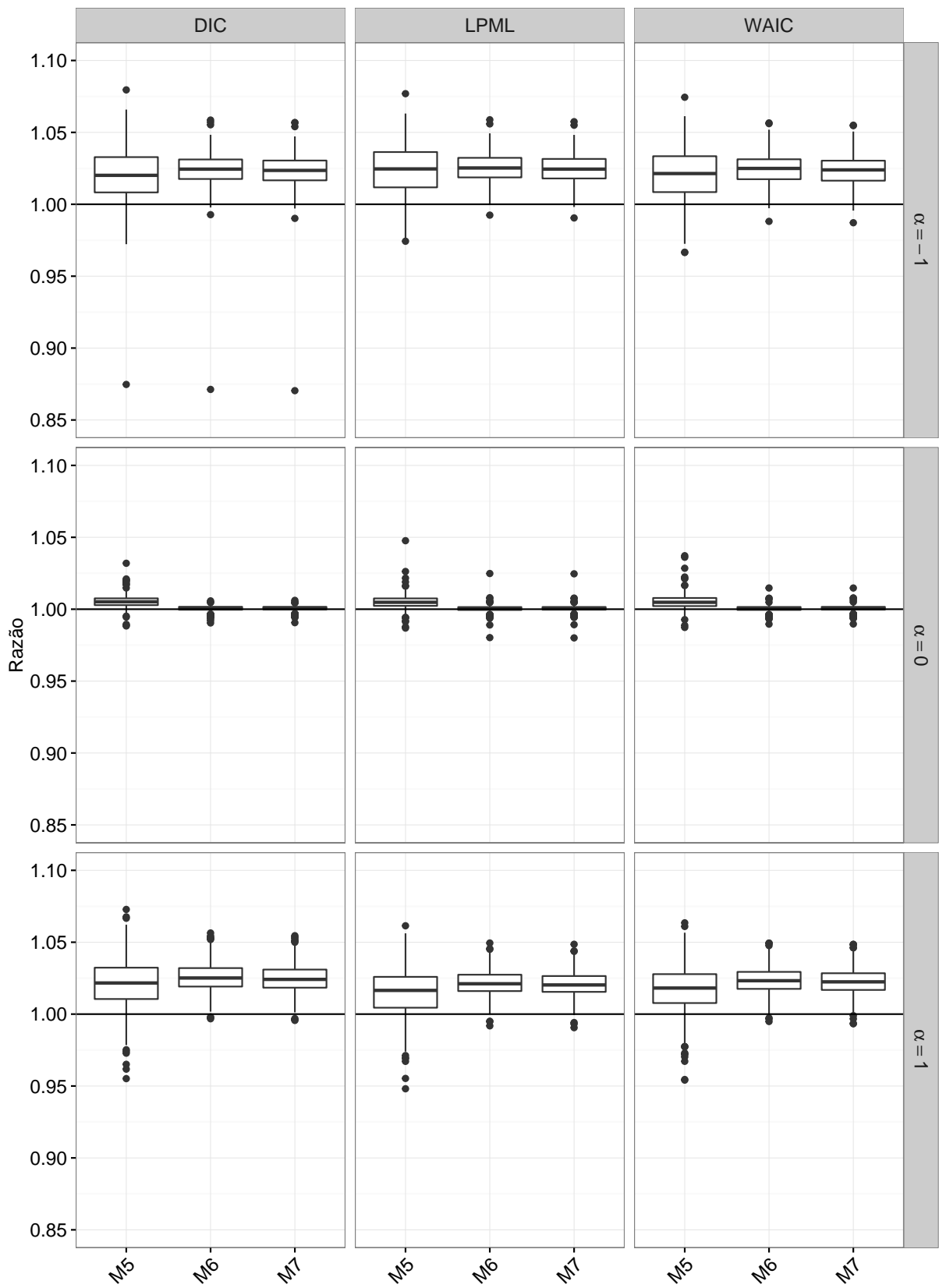


Figura 4.5: Razão entre valor dos critérios para os modelos incorretos pelos valores para os modelos corretos.

Tabela 4.26: Proporção do número de vezes que o critério aponta para o modelo correto.

		DIC	LPML	WAIC
$\alpha = -1$	M5	0,856	0,990	0,986
	M6	0,894	0,994	0,994
	M7	0,860	0,994	0,992
$\alpha = 0$	M5	0,916	0,676	0,708
	M6	0,900	0,666	0,692
	M7	0,890	0,658	0,682
$\alpha = 1$	M5	0,878	0,994	0,994
	M6	0,818	0,992	0,988
	M7	0,868	0,992	0,990

A Figura 4.7 exibe a razão entre o vício relativo dos modelos sob abordagem frequentista pelo vício relativo dos modelos sob abordagem bayesiana. Observa-se que os valores dessa razão são próximos de 1, mostrando equivalência entre as abordagens. Exceto para o modelo M7, que em alguns casos mostra ajuste melhor para os modelos sob abordagem bayesiana.

De forma geral, o modelo M7 apresentou estimativas com viés alto, isso pode ser devido ao número de intervalos para o eixo dos tempos. Chen *et al.* (2002b) comentaram que a estimação da fração de cura pode ser altamente afetada pela natureza não paramétrica da distribuição $F(t)$. Esses autores recomendaram escolher uma quantidade pequena ou moderada para o número de intervalos (J) e fazer uma análise de sensibilidade. Chen *et al.* (2002b) utilizaram os seguintes números de intervalos, $J = (1, 5, 10)$, para o eixo dos tempos no ajuste do MEP. Os autores Cancho *et al.* (2011) apresentaram um modelo flexível capaz de ajustar diversas formas para a função taxa de falha, utilizando número de intervalos igual a $J = (1, 2, 3, 4, 5, 6)$ para o MEP.

Tabela 4.27: Proporção do número de vezes que o vício relativo do modelo correto é menor que o vício relativo do modelo incorreto em dados com fração de cura.

	$\beta_0^{(T)}$	$\beta_1^{(T)}$	τ	$\beta_0^{(T)}$	$\beta_0^{(T)}$	τ
	Abordagem Bayesiana			Abordagem Frequentista		
$\alpha = -1$						
M5	0,548	0,510	0,484	0,540	0,530	0,588
M6	0,552	0,500	0,444	0,568	0,522	0,596
M7	0,566	0,526	0,514	0,416	0,434	0,420
$\alpha = 0$						
M5	0,508	0,506	0,44	0,518	0,512	0,470
M6	0,508	0,540	0,506	0,484	0,508	0,498
M7	0,516	0,504	0,496	0,384	0,388	0,370
$\alpha = 1$						
M5	0,494	0,518	0,536	0,540	0,536	0,586
M6	0,484	0,506	0,510	0,516	0,528	0,580
M7	0,494	0,518	0,536	0,466	0,470	0,468

Tabela 4.28: Razão entre vício relativo de modelo incorreto (com suposição incorreta sobre o mecanismo gerador de censura) pelo vício relativo do modelo correto (com suposição correta sobre o mecanismo gerador de censura em dados com fração de cura.

	$\beta_0^{(T)}$	$\beta_1^{(T)}$	τ	$\beta_0^{(T)}$	$\beta_1^{(T)}$	τ
	Abordagem Bayesiana			Abordagem Frequentista		
	Média	Mediana		Média	Mediana	
$\alpha = -1$						
M5	5,950	1,093	1,742 1,0126	3,088	0,991	1,843 1,079 2,452 1,016 1,835 1,081
M6	2,168	1,064	2,327 1,001	2,516	0,956	2,440 1,125 1,760 1,012 1,536 1,089
M7	2,109	1,088	24,318 1,032	2,050	1,014	7,779 1,055 4,265 1,075 41,184 1,055
$\alpha = 0$						
M5	3,320	1,004	1,980 1,004	1,420	0,957	1,248 1,005 1,288 1,002 1,146 0,994
M6	1,398	1,011	1,886 1,024	1,994	1,004	3,678 0,998 1,099 1,001 1,124 0,999
M7	1,676	1,007	6,304 1,003	2,474	0,994	4,019 1,002 4,376 1,003 5,609 0,987
$\alpha = 1$						
M5	1,899	0,991	2,249 1,036	1,826	1,031	1,713 1,035 1,728 1,028 1,946 1,059
M6	1,767	0,980	3,054 1,007	7,463	1,009	1,462 1,011 3,005 1,011 1,705 1,052
M7	1,710	0,965	2,168 1,033	2,002	1,014	2,883 1,072 5,033 1,081 3,929 1,073

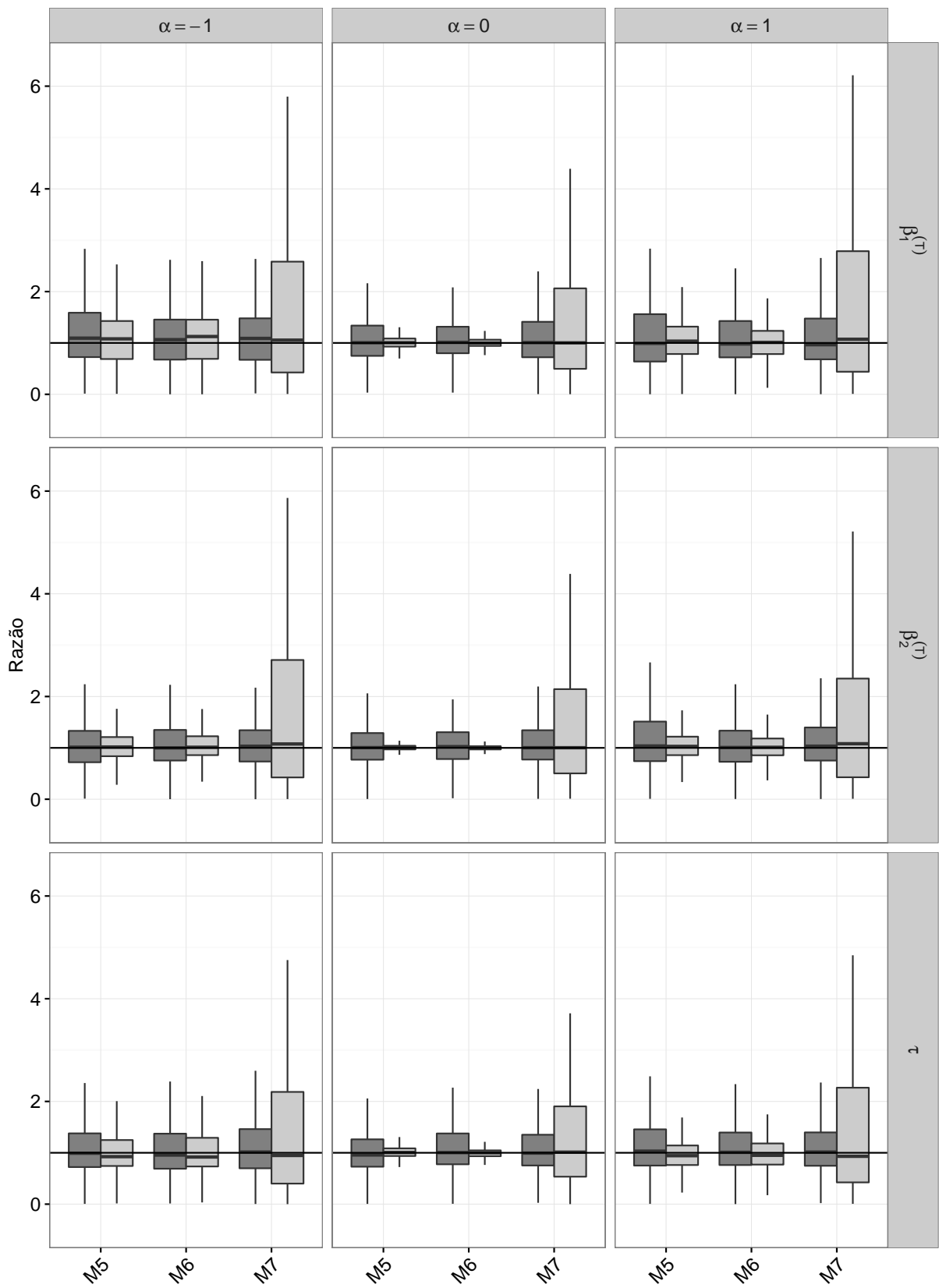


Figura 4.6: Razão entre o vício relativo dos modelos incorretos pelo vício relativo dos modelos corretos, para abordagem frequentista (cinza claro) e abordagem bayesiana (cinza escuro).

Tabela 4.29: Proporção do número de vezes que o vício relativo do modelo com abordagem bayesiana e censura informativa é menor que o vício relativa dos modelos com abordagem frequentista e censura informativa, em dados com fração de cura.

	$\beta_0^{(T)}$	$\beta_1^{(T)}$	$\beta_1^{(C)}$	$\beta_2^{(C)}$	α	τ
$\alpha = -1$						
M5	0,522	0,458	0,470	0,502	0,520	0,564
M6	0,444	0,402	0,478	0,446	0,516	0,540
M7	0,352	0,360	0,396	0,390	0,404	0,468
$\alpha = 0$						
M5	0,492	0,472	0,486	0,502	0,530	0,578
M6	0,470	0,444	0,458	0,426	0,484	0,602
M7	0,398	0,410	0,386	0,370	0,340	0,460
$\alpha = 1$						
M5	0,454	0,486	0,502	0,480	0,568	0,580
M6	0,460	0,472	0,462	0,456	0,478	0,562
M7	0,428	0,438	0,436	0,436	0,446	0,530

Tabela 4.30: Razão entre vício relativo de modelo com abordagem frequentista e censura informativa pelo vício relativo do modelo com abordagem bayesiana e censura informativa em dados com fração de cura.

	$\beta_0^{(T)}$		$\beta_1^{(T)}$		$\beta_1^{(C)}$		$\beta_2^{(C)}$		α		τ	
	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana	Média	Mediana
$\alpha = -1$												
M5	2,473	1,018	1,347	0,985	1,780	0,984	1,733	1,001	1,598	1,014	4,197	1,078
M6	1,468	0,973	1,779	0,947	1,207	0,987	1,288	0,971	2,229	1,006	2,219	1,073
M7	1,534	0,957	2,046	0,975	1,415	0,995	1,998	0,991	1,936	0,999	2,062	1,147
$\alpha = 0$												
M5	1,681	0,986	1,543	0,986	1,134	0,988	18,876	1,003	2,243	1,021	6,895	1,118
M6	1,522	0,976	1,320	0,967	1,159	0,982	1,436	0,961	1,279	0,992	2,771	1,145
M7	1,850	1,020	1,443	1,022	3,833	1,007	1,692	0,993	1,415	0,971	2,206	1,166
$\alpha = 1$												
M5	1,468	0,956	1,533	0,988	2,043	1,001	2,412	0,983	1,541	1,041	2,925	1,104
M6	1,385	0,976	2,795	0,985	2,461	0,982	1,459	0,972	1,210	0,994	3,925	1,107
M7	1,637	0,986	1,776	0,988	1,662	0,991	3,790	0,992	1,389	1,012	2,132	1,101

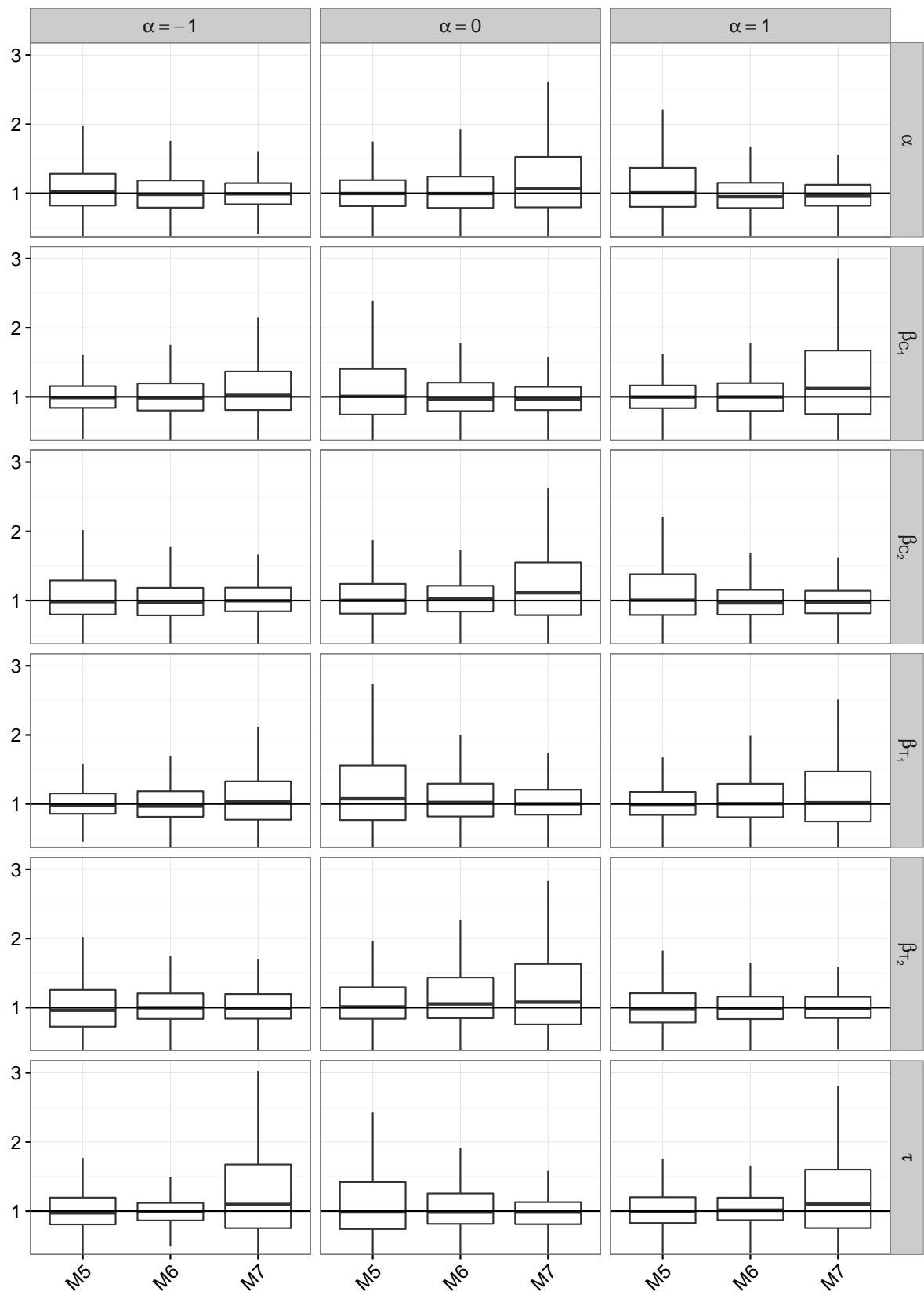


Figura 4.7: Razão entre o vício relativo dos modelos com abordagem frequentista pelo vício relativo dos modelos com abordagem bayesiana.

Capítulo 5

Aplicações em dados reais

5.1 Análise de dados de mortalidade em centros de diálise renal dos Estados Unidos

Nesta seção apresentamos o ajuste dos modelos, sob abordagem frequentista e bayesiana, para uma amostra do estudo chamado *Dialysis Outcomes and Practice Patterns Study* (DOPPS), sobre mortalidade em centros de diálise dos Estados Unidos. Antes de discorrermos sobre os resultados obtidos pelos modelos apresentados anteriormente, vamos compreender um pouco sobre o objetivo do estudo DOPPS.

A hemodiálise é um tratamento feito em pacientes com insuficiência renal crônica. Esse tratamento é um procedimento feito por uma máquina que limpa e filtra o sangue, liberando os resíduos prejudiciais à saúde, ajudando no controle da pressão arterial e na manutenção das substâncias como sódio, potássio, ureia e creatinina em equilíbrio. Na hemodiálise uma máquina recebe o sangue do doente por meio de um acesso vascular (tubo ou fístula arteriovenosa) e depois é enviado para o filtro de diálise. No filtro, o sangue é exposto a um fluido, posteriormente ocorre a retirada desse fluido e das toxinas em excesso, e então é feita a devolução do sangue limpo para o paciente, pelo acesso vascular; vide <http://sbn.org.br/publico/tratamentos/hemodialise>.

O procedimento de hemodiálise geralmente leva de três a quatro horas, realizado três ou quatro vezes por semana, ou, dependendo da situação do paciente, diariamente. O volume total do fluido utilizado no tratamento depende do peso dos pacientes. As características do processo, que incluem o tempo de duração diária, o número de sessões semanais, a velocidade de fluxo do fluido e o tipo de máquina de diálise, variam muito

de uma clínica para outra. Essas características são nomeadas de Padrões de Prática, (Huang & Wolfe, 2002). Como os Padrões de Prática variam muito de uma clínica para outra, em 1996 foi criado o estudo DOPPS para identificar as melhores práticas no tratamento de pacientes com insuficiência renal crônica. Atualmente, o DOPPS acompanha mais de 70.000 pacientes em hemodiálise, em mais de vinte países, observando os pacientes ao longo do tempo, correlacionando práticas e resultados, ajudando os pesquisadores a identificar as melhores características que melhoram a vida dos pacientes; vide <http://dopps.org/ProgramHome.aspx>.

Os dados investigados nesta seção foram fornecidos pelo programa *Dialysis Outcomes and Practice Patterns Study*. O DOPPS é coordenado pela *Arbor Research Collaborative for Health, Ann Arbor, MI, EUA*. As interpretações e conclusões desses dados são de responsabilidade desta autora, e não devem ser vistas como uma posição oficial do DOPPS ou da *Arbor Research Collaborative for Health*.

O período de coleta dos dados analisados foi de julho de 1996 até junho de 2015. Consideramos uma amostra de 2.526 pacientes distribuídos em 71 clínicas de diálise, com número de pacientes por clínica variando de 6 a 123. Durante o estudo, dos 2.526 pacientes, 646 (25,6%) faleceram; 139 (5,5%) tiveram alteração da modalidade de diálise ou retiraram-se do estudo; 130 (5,1%) receberam transplante; e o restante, 1.611 (63,8%) permaneceram no estudo durante o acompanhamento do DOPPS ou possuíam motivo de retirada desconhecido.

Classificamos os pacientes que faleceram como falha; os indivíduos doentes que tiveram alteração da modalidade de diálise e os que deixaram o estudo como censura informativa, pois segundo Huang & Wolfe (2002), suspeita-se que o motivo mais comum para os pacientes deixarem o tratamento é a piora do estado de saúde; o transplante também pode ser informativo; e os demais tempos foram classificados como censura não informativa.

A idade dos pacientes variou de 18 a 98 anos, sendo a idade média de 62,83 com desvio padrão de 15,38. As distribuições das demais variáveis consideradas na análise podem ser vistas na Tabela 4.20.

Neste ajuste temos quatro variáveis indicadoras, sendo $\delta^{(T)} = 1$ se o indivíduo falhou e $\delta^{(T)} = 0$ caso contrário; $\delta^{(C)} = 1$ se o indivíduo teve censura informativa e $\delta^{(C)} = 0$ caso contrário; $\delta^{(R)} = 1$ se o indivíduo recebeu transplante e $\delta^{(R)} = 0$ caso contrário; e $\delta^{(A)} = 1$ se o indivíduo teve censura administrativa. Dessa forma, temos as seguintes funções taxa de falha:

$$h_{ik}^{(T)}(y|\mathbf{X}_{ik}^{(T)}, \mathbf{X}_{ik}^{(C)}, W_k, C_{ik}, R_{ik}) = h_0^{(T)}(y) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{X}_{ik}^{(T)} + W_k),$$

$$h_{ik}^{(C)}(y|\mathbf{X}_{ik}^{(T)}, \mathbf{X}_{ik}^{(C)}, W_k, T_{ik}, R_{ik}) = h_0^{(C)}(y) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{X}_{ik}^{(C)} + \alpha_2 W_k),$$

Tabela 5.1: Distribuição dos eventos de falha, censura informativa e transplante.

Covariáveis	Categorias	Total (%)	Falha (%)	Cens. Inf. (%)	Transp. (%)
Idade	≥ 50 anos	2048 (81,1)	581 (89,9)	114 (82,0)	77 (59,2)
	< 50 anos	473 (18,7)	65 (10,1)	24 (17,3)	53 (40,8)
Sexo	Masculino	1388 (54,9)	348 (53,9)	67 (48,2)	80 (61,5)
	Feminino	1136 (45,0)	297 (46,0)	71 (51,1)	50 (38,5)
Raça	Negro	811 (32,1)	162 (25,1)	37 (26,6)	34 (26,2)
	Outro	1715 (67,9)	484 (74,9)	102 (73,4)	96 (73,8)
Diabetes	Sim	1198 (47,4)	336 (52,0)	61 (43,9)	39 (30,0)
	Não	1309 (51,8)	300 (46,4)	77 (55,4)	90 (69,2)

¹Cens.Inf. denota Censura Informativa, Transp. denota transplante.

$$h_{ik}^{(R)}(y|\mathbf{X}_{ik}^{(R)}, \mathbf{X}_{ik}^{(R)}, W_k, T_{ik}, C_{ik}) = h_0^{(R)}(y) \exp(\boldsymbol{\beta}'_R \mathbf{X}_{ik}^{(R)} + \alpha_3 W_k),$$

em que $h_{ik}^{(R)}(\cdot)$ denota a função taxa de falha dos indivíduos que receberam transplante.

No ajuste dos modelos sob abordagem bayesiana, consideramos um descarte de 10.000, um espaçamento de 10 e 5.000 amostras *a posteriori*. A Tabela 4.21 apresenta o valor para os critérios de comparação de modelos. Observa-se nesta tabela que o melhor ajuste apontado pelos critérios DIC, LPML e WAIC, para os modelos com abordagem bayesiana, foi com o modelo MEP com cinco intervalos. Para abordagem frequentista o modelo MEP com cinco intervalos apresentou menor valor pelo critério BIC, e o ajuste MEP com 10 intervalos apresentou menor valor pelos critérios AIC e HQ. Com o objetivo de manter a comparabilidade entre as abordagens, apresentamos na Tabela 4.22 o ajuste utilizando cinco intervalos para o eixo dos tempos do MEP, para abordagens frequentista e bayesiana.

Tabela 5.2: Análise dos critérios de informação para os modelos.

	Weibull	MEP (5 int)	MEP (10 int)	MEP (15 int)
Abordagem Frequentista				
AIC	3147,91	3086,13	3063,77	3227,12
BIC	3188,68	3144,37	3151,13	3343,61
HQ	3162,72	3107,27	3095,48	3269,41
Abordagem Bayesiana				
DIC	3214,80	3209,61	3210,87	3219,13
LPML	3167,06	3165,65	3170,95	3177,05
WAIC	3167,01	3165,61	3170,90	3177,00

A Tabela 4.22 apresenta o ajuste utilizando o modelo Exponencial por Partes com cinco intervalos sob abordagens frequentista e bayesiana. Para abordagem frequentista, apresentamos a estimativa pontual, intervalar e erro padrão estimado. Para abordagem bayesiana, apresentamos a moda, erro padrão e intervalo HPD *a posteriori*.

Ao analisar o ajuste sob abordagem frequentista e censura informativa na Tabela 4.22,

Tabela 5.3: Resultados para dados DOPPS com modelo MEP considerando cinco intervalos, ajuste sob enfoques frequentista e bayesiano.

		Abordagem Frequentista						
		Censura Informativa			Censura Não Informativa			Vies(%)
		Estimativa	EP	IC	Estimativa	EP	IC	
Falha	Idade	0,4325	0,047	[0,340; 0,525]	0,4436	0,048	[0,350; 0,538]	2,56
	Sexo	0,0111	0,081	[-0,148; 0,170]	0,0205	0,081	[-0,138; 0,179]	84,68
	Raça	-0,4291	0,099	[-0,623; -0,235]	-0,4562	0,103	[-0,658; -0,254]	6,31
	Diabetes	0,2148	0,081	[0,056; 0,374]	0,2136	0,082	[0,053; 0,374]	-0,55
	σ^2	0,1048	0,033	[0,040; 0,169]	0,1264	0,064	[0,001; 0,252]	20,61
Cens. Infor.	Idade	0,2542	0,097	[0,064; 0,444]				
	Sexo	-0,2512	0,179	[-0,602; 0,100]				
	Raça	-0,4116	0,215	[-0,833; 0,010]				
	Diabetes	-0,1025	0,182	[-0,459; 0,254]				
	α_2	2,2218	0,563	[1,118; 3,325]				
Transplante	Idade	-0,7208	0,084	[-0,885; -0,556]				
	Sexo	0,2244	0,182	[-0,132; 0,581]				
	Raça	-0,8208	0,205	[-1,223; -0,419]				
	Diabetes	-0,5235	0,197	[-0,910; -0,137]				
	α_3	0,8206	0,403	[0,031; 1,610]				
		Abordagem Bayesiana						
		Censura Informativa			Censura Não Informativa			Viés(%)
		Moda	DP	HPD	Moda	DP	HPD	
Falha	Idade	0,4382	0,048	[0,345; 0,532]	0,4503	0,047	[0,357; 0,542]	2,76
	Sexo	-0,0018	0,082	[-0,147; 0,165]	0,0222	0,082	[-0,150; 0,173]	-1333,33
	Raça	-0,4499	0,1	[-0,641; -0,249]	-0,4686	0,104	[-0,677; -0,269]	4,16
	Diabetes	0,2066	0,079	[0,054; 0,369]	0,2168	0,081	[0,060; 0,373]	4,94
	τ	5,0192	1,564	[3,067; 8,830]	4,4919	1,3897	[2,530; 7,647]	11,75
Cens. Infor.	Idade	0,2441	0,096	[0,061; 0,433]				
	Sexo	-0,2944	0,175	[-0,637; 0,054]				
	Raça	-0,4679	0,21	[-0,881; -0,073]				
	Diabetes	-0,1578	0,174	[-0,455; 0,219]				
	α_2	1,6189	0,478	[0,915; 2,746]				
Transplante	Idade	-0,7438	0,084	[-0,881; -0,555]				
	Sexo	0,2106	0,188	[-0,144; 0,586]				
	Raça	-0,8938	0,208	[-1,235; -0,422]				
	Diabetes	-0,4526	0,2	[-0,923; -0,131]				
	α_3	0,6628	0,348	[0,090; 1,452]				

¹Estimativa: estimativa de máxima verossimilhança, EP: erro padrão, IC: intervalo de confiança calculado a partir do EP, Viés: vício relativo entre a estimativa obtida pelo modelo com censura não informativa e a estimativa obtida pelo modelo com censura informativa (considerado padrão ouro).

percebe-se que o parâmetro que mede a dependência entre os tempos de falha e censura informativa, α_2 , apresenta estimativa $\alpha_2 = 2,2218$ [1,1118; 3,325]. Logo, podemos concluir que a associação entre os tempos de falha e censura é significativa no modelo. O parâmetro que representa a dependência entre tempos de falha e transplante, α_3 , apresentou estimativa $\alpha_3 = 0,8206$ [0,031; 1,610], sendo portanto, significativo no modelo a um nível de 5% de confiança. Nota-se que as covariáveis idade, raça e diabetes são significativas tanto para ajustar o tempo de falha, quanto para ajustar o tempo de espera até o transplante. Já para o tempo até a censura informativa apenas a covariável idade apresentou-se significativa. Também observa-se que o coeficiente da variável raça teve um aumento de 6,31% ao se considerar censura informativa, sendo em torno de -0,456 quando é utilizado o modelo supondo censura não informativa e -0,429 quando é empregado o modelo supondo censura informativa, sendo em ambos os casos significativa.

Observando-se o ajuste sob abordagem bayesiana e censura informativa na Tabela 4.22, podemos notar que o parâmetro α_2 apresentou uma estimativa menor que o valor estimado pela abordagem frequentista, podendo ser em consequência da especificação *a priori*. No entanto, também foi significativo, sendo $\alpha_2 = 1,6189$ [0,915; 2,746]. Da mesma forma, a estimativa para o parâmetro α_3 foi menor, $\alpha_3 = 0,6628$ [0,090; 1,452], porém, significativa. As covariáveis idade, raça e diabetes apresentaram-se significativas no ajuste dos tempos de falha e transplante, para os tempos de censura informativa as covariáveis idade e raça foram significativas. Ao comparar o ajuste considerando a suposição de censura informativa com o ajuste supondo censura não informativa, para abordagem bayesiana, verifica-se que o coeficiente da covariável raça apresentou um aumento de 4,16% ao considerar censura informativa no modelo.

5.2 Análise de dados de câncer de melanoma

Nesta seção apresentamos uma análise em dados sobre câncer de melanoma fornecidos pelo Programa *Surveillance, Epidemiology, and End Results* (SEER), utilizando os modelos propostos na Seção 3.2 para dados de sobrevivência com fração de cura e censura informativa.

De acordo com Dimatos *et al.* (2009), o câncer de pele é o tipo de câncer mais incidente em ambos os sexos no Brasil, com as maiores taxas estimadas para a região Sul, sendo a neoplasia maligna que vem apresentando maior crescimento na última década, tornando-se um problema de saúde pública. Carvalho *et al.* (2004) constataram que, dos indivíduos com diagnóstico de melanoma admitidos no Hospital de Clínicas de Porto Alegre entre

1999 e 2000, aproximadamente 10% dos casos foram atribuídos a mutações em genes de predisposição, ou seja, hereditários. Criado *et al.* (1999) realizaram um estudo com pacientes diagnosticados com melanoma maligno cutâneo no Hospital do Servidor Público Estadual (FMO) - SP, no período de 1963 a 1997, no qual também evidenciaram um aumento da ocorrência deste melanoma ao longo dos anos.

O aumento do número de casos diagnosticados com melanoma cutâneo também ocorreu nos países desenvolvidos (Criado *et al.*, 1999). McHenry *et al.* (1992), realizaram um estudo com 1430 pacientes identificados com melanoma na Escócia durante 1979 a 1989, e notaram que a incidência do câncer dobrou nesse período. Fato que também foi observado no Canadá, em um estudo retrospectivo de 1972 a 1981 (Carmichael & Wilson, 1992). De acordo com Carvalho *et al.* (2004), a incidência da neoplasia maligna originada a partir dos melanócitos tem aumentado drasticamente nas últimas quatro décadas em diversos países industrializados. Nos Estados Unidos, 44.200 novos casos foram registrados em 1999, resultando em 7.300 mortes; além do mais, entre 1973 e 1998 a incidência e a mortalidade por melanoma neste país aumentaram continuamente (Carvalho *et al.*, 2004).

Segundo o Consórcio para Genética do Melanoma (*Melanoma Genetics Consortium* - GenoMEL), o melanoma cutâneo tornou-se cada vez mais comum nos últimos anos, podendo ser causado pela exposição ao sol em pessoas de pele clara no período de férias, feriados, ou durante a prática de esportes, <http://genomel.org/info-for-patients/research-in-leeds-2/>.

O Programa de Vigilância, Epidemiologia e Resultados Finais (SEER) do Instituto Nacional do Câncer fornece informações sobre estatísticas de câncer dos Estados Unidos, com o objetivo de reduzir a incidência de câncer nesta população. O Programa SEER abrange, atualmente, aproximadamente 28 % da população dos Estados Unidos, cobrindo 26% dos afro-americanos, 38% dos hispânicos, 44% dos ameríndios e dos nativos do Alasca, 50% dos asiáticos e 67% dos havaianos. São coletados rotineiramente dados demográficos, local do tumor primário, morfologia e estadiamento no momento do diagnóstico do tumor, primeiro tratamento e acompanhamento do estado vital, desde 01 de janeiro de 1973, <https://seer.cancer.gov> acessado em 05.03.2017.

Os registros do SEER contém informações de 13 estados dos Estados Unidos, sendo eles Connecticut, Iowa, Novo México, Utah, as áreas metropolitanas de San-Francisco-Oakland, San-Jose-Monterey, o condado de Los Angeles, Califórnia, Atlanta, Geórgia, Detroit, Michigan, Seattle-Puget Sound, Washington, 10 condados da Geórgia Rural e os registros de tumor nos nativos do Alasca (Lachiewicz *et al.*, 2008).

Extraímos um amostra de dados sobre câncer de melanoma do Programa *Surveillance, Epidemiology, and End Results*, da base de dados “SEER*Stat: Incidence - SEER 18 Regs

Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2015 Sub (2000-2013) <Katrina/Rita Population Adjustment>” em <https://seer.cancer.gov>. Seleccionamos pacientes com diagnóstico entre 2004 e 2013, pois a partir de 2004 as variáveis tamanho e extensão do tumor, número de linfonodos, entre outras características, passaram a ser incluídas nos registros do SEER (Johnson, 2004).

Portanto, de todos os pacientes diagnosticados com melanoma cutâneo no período de 2004 a 2013, excluímos os indivíduos que sem informação sobre o tamanho do tumor, ou sobre o número de linfonodos ou ainda, sobre o estadiamento da doença. Destes, sorteamos aleatoriamente 5%, resultando em uma amostra de 2.280 indivíduos. Dos 2.280 registros, 218 (9,6%) apresentaram óbito devido ao câncer, 178 (7,8%) óbito devido a outros fatores, considerados nesta análise como censura informativa, e 1.884 (82,6%) estavam vivos até o final da coleta, considerados como censura administrativa. O estado em que o paciente foi diagnosticado foi utilizado como variável identificadora de grupo, logo, nesta amostra o número de indivíduos por grupo variou de 39 (1,71%) a 1141 (50,04%).

De acordo com Seiffert *et al.* (1993), os estágios do tumor: Localizado, Regional e Distante, também podem ser denominados de Estadiamento I até IV. O Comitê Americano sobre Câncer (*American Joint Committee on Cancer's* (AJCC)) apresentou a classificação do tumor, levando em consideração o tumor primário (T), os nódulos linfáticos regionais (N) e a presença de Metástases distante (M). Sendo T1, T2, T2, T4, tamanho e/ou extensão crescente do primeiro tumor; N0 sem linfonodo regional metastático, N1, N2, N3 número crescente de nódulos linfáticos; M0 sem metástases distantes, M1 metástase distante. Os estágio I - IV são apresentados na Tabela 4.34.

Tabela 5.4: Classificação do tumor segundo Comitê Americano sobre Câncer (AJCC).

Estadiamento do tumor (AJCC)	T	N	M
I	T1	N0	M0
II	T2	N0	M0
III	T3	N0	M0
IV	T4	Qualquer N	M0
	Qualquer T	N1	M0
	Qualquer T	Qualquer N	M1

¹Tabela extraída do Guia comparativo de estadiamento para o câncer, (Seiffert *et al.*, 1993).

Consideramos as variáveis sexo, idade no diagnóstico, tamanho aproximado do tumor em milímetros, número exato de linfonodos regionais examinados pelo patologista e encontrados com metástases, presença de linfonodos envolvidos no câncer no momento do diagnóstico (positivo/negativo) e estadiamento do tumor (I - IV) (Johnson, 2004). As variáveis contínuas foram padronizadas para análise, sexo feminino foi considerado como

categoria de referência. As demais informações descritivas utilizadas nas análises deste estudo se encontram na Tabela 4.35.

Tabela 5.5: Análise descritiva dos 2.280 pacientes diagnosticados com melanoma cutâneo.

Covariáveis		Total	Falha	Cens. Inf.
		Média (DP)	Média (DP)	Média (DP)
Idade		58,77 (16,66)	63,21 (16,35)	75,89 (12,77)
Número nódulos positivos		60,59 (47,42)	50,15 (48,07)	70,58 (43,94)
Tamanho do Tumor		21,32 (67,12)	33,21 (79,54)	18,28 (19,69)
	Categorias	Freq (%)	Freq (%)	Freq (%)
Sexo	Feminino	955 (41,88)	59 (27,06)	62 (34,83)
	Masculino	1325 (58,11)	159 (72,93)	116 (65,17)
Linfonodos	Positivo	239 (10,48)	82 (37,61)	14 (7,9)
	Negativo	2041 (89,52)	136 (62,38)	164 (92,1)
Estadiamento	I	1643 (72,06)	38 (17,43)	108 (60,67)
	II	356 (15,61)	67 (30,74)	52 (29,21)
	III	198 (8,7)	55 (25,23)	11 (6,2)
	IV	83 (3,6)	58 (26,60)	7 (3,9)

O modelo que forneceu um ajuste melhor, segundo os critérios de comparação de modelos sob abordagem bayesiana, foi o modelo Weibull com censura informativa, como podemos ver na Tabela 4.36. Portanto, na Tabela 4.37 apresentamos o ajuste utilizando o modelo Weibull com e sem censura informativa, para as abordagens frequentista e bayesiana.

Analisando o ajuste com o modelo Weibull supondo censura informativa e utilizando abordagem frequentista na Tabela 4.37, observa-se que o parâmetro que ajusta a dependência entre os tempos de promoção e censura informativa, α , apresentou uma estimativa $\hat{\alpha} = 0,545$, porém não sendo significativo. As covariáveis idade, sexo, tamanho do tumor, estadiamento I, II e III apresentaram-se significativas para a proporção de cura. O coeficiente associado a covariável presença de linfonodos apresentou uma diminuição de 49,5% quando foi considerado a suposição de censura não informativa ao invés de censura informativa no ajuste do modelo.

Observando o ajuste com censura informativa e abordagem bayesiana, também apresentado na Tabela 4.37, notamos que a estimativa para o parâmetro α foi um pouco menor do que a estimativa obtida pela abordagem frequentista, $\hat{\alpha} = 0,381$, também não sendo significativo. As covariáveis idade, sexo, tamanho do tumor, estadiamento I, II e II também foram significativas no ajuste da proporção de cura. O coeficiente da covariável presença de linfonodos apresentou uma diminuição de 115,9% no ajuste supondo censura não informativa, comparado com o valor obtido pelo ajuste supondo censura informativa.

Tabela 5.6: Valor do critério de comparação aplicado ao conjunto de dados sobre câncer de melanoma

	Weibull	MEP 5 int	MEP 10int	MEP 15 int
Censura Informativa				
LPML	1531,849	1537,817	1542,829	1549,205
DIC	1527,968	1534,753	1540,476	1546,957
WAIC	1531,423	1537,489	1542,273	1548,759
Censura Não Informativa				
LPML	1532,962	1538,365	1542,641	1550,182
DIC	1528,295	1534,903	1538,156	1546,408
WAIC	1532,429	1537,986	1542,186	1549,697

Tabela 5.7: Ajuste para dados sobre câncer de melanoma (SEER) com o modelo Weibull sob abordagens frequentista e bayesiana.

		Abordagem Frequentista							
		Censura Informativa			Censura Não Informativa				
		Estimativa	EP	IC	Estimativa	EP	IC	Viés	
Prop. cura	β_0	2,128	0,240	[1,658; 2,599]	2,150	0,223	[1,713; 2,587]	1,038	
	β_{idade}	0,319	0,079	[0,164; 0,473]	0,317	0,079	[0,162; 0,472]	-0,565	
	β_{sexo}	-0,345	0,154	[-0,647; -0,044]	-0,343	0,155	[-0,647; -0,039]	0,695	
	β_{tumor}	0,186	0,042	[0,103; 0,268]	0,187	0,042	[0,105; 0,269]	0,643	
	$\beta_{linfon.}$	-0,081	0,269	[-0,609; 0,446]	-0,122	0,271	[-0,653; 0,409]	-49,509	
	$\beta_{estad.I}$	-4,767	0,253	[-5,263; -4,271]	-4,788	0,254	[-5,286; -4,290]	-0,441	
	$\beta_{estad.II}$	-2,567	0,228	[-3,013; -2,120]	-2,581	0,229	[-3,030; -2,132]	-0,553	
	$\beta_{estad.III}$	-1,926	0,241	[-2,398; -1,454]	-1,896	0,241	[-2,368; -1,424]	1,558	
	α	0,545	0,466	[-0,368; 1,458]					
	Cens. Inf.	β_{idade}	1,535	0,111	[1,317; 1,752]				
β_{sexo}		-0,205	0,158	[-0,515; 0,105]					
β_{tumor}		-0,041	0,102	[-0,241; 0,159]					
$\beta_{linfon.}$		0,031	1,595	[-3,095; 3,157]					
$\beta_{estad.I}$		-0,434	0,588	[-1,586; 0,719]					
$\beta_{estad.II}$		0,173	0,567	[-0,939; 1,284]					
$\beta_{estad.III}$		-0,301	1,02	[-2,301; 1,698]					
σ^2		0,066	0,055	[-0,042; 0,174]	0,0630	0,049	[0,001; 0,159]	-4,83	
		Abordagem Bayesiana							
		Censura Informativa			Censura Não Informativa				
		Moda	DP	HPD	Moda	DP	HPD	Viés	
Prop. cura	β_0	2,191	0,427	[1,518; 3,154]	2,182	0,628	[1,453; 3,848]	-0,411	
	β_{idade}	0,329	0,079	[0,174; 0,484]	0,334	0,082	[0,167; 0,484]	1,519	
	β_{sexo}	-0,342	0,153	[-0,652; -0,053]	-0,309	0,158	[-0,653; -0,038]	9,649	
	β_{tumor}	0,185	0,045	[0,082; 0,256]	0,182	0,046	[0,083; 0,260]	-1,622	
	$\beta_{linfon.}$	-0,069	0,280	[-0,670; 0,420]	-0,149	0,279	[-0,761; 0,342]	-115,942	
	$\beta_{estad.I}$	-4,861	0,264	[-5,331; -4,287]	-4,770	0,266	[-5,372; -4,318]	1,872	
	$\beta_{estad.II}$	-2,684	0,240	[-3,089; -2,156]	-2,708	0,238	[-3,096; -2,160]	-0,894	
	$\beta_{estad.III}$	-1,866	0,249	[-2,376; -1,42]	-1,823	0,248	[-2,341; -1,375]	2,304	
	α	0,381	0,345	[-0,240; 1,117]					
	Cens. Inf.	β_{idade}	1,555	0,111	[1,348; 1,780]				
β_{sexo}		-0,207	0,160	[-0,540; 0,081]					
β_{tumor}		-0,026	0,138	[-0,368; 0,147]					
$\beta_{linfon.}$		-0,704	0,798	[-2,178; 0,912]					
$\beta_{estad.I}$		-2,378	0,495	[-3,062; -1,161]					
$\beta_{estad.II}$		-1,645	0,500	[-2,541; -0,571]					
$\beta_{estad.III}$		-1,249	0,707	[-2,542; 0,157]					
τ		3,374	1,526	[0,909; 6,581]	2,5740	1,555	[0,980; 6,615]	-23,711	

¹Estimativa: estimativa de máxima verossimilhança, EP: erro padrão, IC: intervalo de confiança calculado a partir do EP, Viés: vício relativo entre a estimativa obtida pelo modelo com censura não informativa e a estimativa obtida pelo modelo com censura informativa (considerado padrão ouro).

Capítulo 6

Considerações finais

O objetivo desta tese residiu em construir modelos gerais para dados de sobrevivência com censura informativa, que fossem capazes de capturar a dependência entre os tempos de falha e os tempos de censura. Para tanto, consideramos os modelos de fragilidade, particularmente o modelo proposto por Huang & Wolfe (2002), para analisar dados de sobrevivência multivariados com censura informativa.

Neste trabalho, apresentamos as seguintes extensões para o modelo de Huang & Wolfe (2002): Propomos um modelo completamente paramétrico, utilizando a distribuição Weibull para ajustar os tempos de falha e censura; um modelo semiparamétrico empregando o MEP para modelar os tempos de falha e censura; realizamos a inferência sob enfoques frequentista e bayesiano para estes modelos, Weibull e MEP, com censura informativa; apresentamos uma abordagem bayesiana para o modelo de Huang & Wolfe (2002).

Com o objetivo de analisar dados de sobrevivência multivariados com censura informativa e fração de cura, mostramos as seguintes extensões: Um modelo de fragilidade completamente paramétrico, com distribuição Weibull para ajustar os tempos de promoção e os tempos de censura informativa; um modelo semiparamétrico, com distribuição Exponencial por Partes para ajustar os tempos de promoção e censura informativa; apresentamos as inferências via abordagens frequentista e bayesiana para ambos os modelos, Weibull e MEP, para fração de cura com censura informativa.

Para avaliar e comparar os modelos propostos, realizamos estudos de simulação, tanto para dados gerados sem fração de cura quanto para dados gerados com fração de cura, considerando três cenários: dados gerados com correlações negativa, nula e positiva entre os tempos de falha e censura. Os novos modelos de censura informativa foram então aplicados a bancos de dados reais, dados sobre pacientes em diálise e dados sobre câncer de melanoma.

Os critérios de seleção de modelos AIC, BIC e HQ (abordagem frequentista) e os critérios DIC, LPML e WAIC (abordagem bayesiana) foram explorados na avaliação da qualidade de ajuste dos modelos. Também avaliamos estimativas médias, erro padrão empírico, probabilidade de cobertura e razão entre os vícios relativos para os ajustes sob abordagem frequentista. Apresentamos estimativas média, mediana, moda, erro padrão e probabilidade de cobertura *a posteriori* para os ajustes sob enfoque bayesiano. Enquanto o erro padrão apresenta avaliações para cada parâmetro, os critérios são medidas globais da qualidade do ajuste. Apresentamos uma comparação entre as estimativas obtidas considerando censura informativa e não informativa, bem como uma comparação entre estimativas obtidas pelas abordagens frequentista e bayesiana.

De forma geral, no cenário com censura informativa e sem fração de cura, quando os dados são gerados com correlação positiva ou negativa, o ajuste com o MEP e estimação via abordagem frequentista apresentou erro padrão empírico menor do que o modelo proposto por Huang & Wolfe (2002). Porém, o vício relativo, em geral, é menor para o modelo não paramétrico. Os critérios indicam que o modelo Weibull (M0) fornece o melhor ajuste, em seguida pelo modelo MEP com 10 intervalos (M1).

Ainda avaliando o cenário sem fração de cura, analisando as estimativas obtidas pela abordagem bayesiana, percebemos que a moda *a posteriori* mostrou-se o estimador com menor viés em todos os cenários. Portanto, comparando a moda *a posteriori* dos diversos modelos, notamos que, quando temos dependência positiva entre os tempos, o modelo não paramétrico (M4) apresentou, quase sempre, vícios menores do que os demais modelos. No entanto, quando observamos os dados gerados com correlação negativa, o modelo MEP com 10 intervalos (M1) apresentou vícios menores para quase todos os parâmetros. Os critérios bayesianos apontaram para o ajuste com o modelo Weibull (M0), e o segundo menor valor foi para o modelo MEP com 10 intervalos (M1).

Assim como no ajuste dos modelos sem fração de cura, o desempenho dos modelos com censura informativa e fração de cura, apontado pelos critérios clássicos, foi melhor quando utilizamos modelos com menos parâmetros. Ao analisarmos o erro padrão empírico, percebemos que o modelo MEP com cinco intervalos (M6) apresentou valores menores em todos os cenários, para quase todos os parâmetros.

Para os dados gerados com fração de cura, as estimativas obtidas através da moda *a posteriori* também apresentaram vícios menores. Logo, comparando a moda *a posteriori* entre os diferentes modelos ajustados, percebemos que o modelo Weibull (M5) apresentou menor vício relativo. De forma geral, o modelo MEP com 10 intervalos (M7) apresentou vícios maiores, tanto via abordagens frequentista e bayesiana, podendo ser devido à natureza semiparamétrica do MEP para ajustar os tempos de promoção.

Assim como Huang & Wolfe (2002), que concluíram que o método proposto, considerando correlação entre os tempos de falha e censura, resultou em uma redução do viés, comparando com o método padrão que considera apenas censura não informativa. Também observamos que os modelos com censura informativa apresentaram vícios menores do que os modelos com censura não informativa.

Diante das conclusões expostas podemos enxergar como próximos passos deste trabalho:

- Realizar mais simulações comparando diferentes distribuições *a priori* para as taxas de falha do modelo Exponencial por Partes, por exemplo, utilizar a distribuição Gama(α, β) com $\alpha \rightarrow 0$ e $\beta \rightarrow 0$, pois segundo Gelman *et al.* (2014a), no caso limite essa distribuição é não informativa. Também incluir a estrutura dinâmica proposta por Gamerman (1991) a fim de eliminar a dificuldade encontrada para se determinar a distribuição *a priori* para a taxa de falha;
- Estender os modelos para captar também a correlação espacial entre os indivíduos, utilizando o modelo CAR. Um dos trabalhos que pode ser utilizado como base é o modelo proposto por Banerjee *et al.* (2003);
- Utilizar a distribuição Estável Positiva proposta por Hougaard (1986) para modelar a fragilidade, pois ela preserva a estrutura de riscos proporcionais. Essa distribuição também foi empregada por Qiou *et al.* (1999) para ajustar a fragilidade em uma modelagem bayesiana para dados de sobrevivência multivariados;
- Nos modelos de fração de cura, para controlar o grau de parametrização na cauda direta da função de sobrevivência, pode-se assumir uma estrutura hierárquica. Considerando distribuições *a priori* Gama independentes para as taxas de falha λ_j , com média e variância das por

$$\mu_j = \frac{H_0(a_j) - H_0(a_{j-1})}{a_j - a_{j-1}} \quad \text{e} \quad \sigma_j^2 = \mu_j \kappa^j,$$

em que $H_0(\cdot)$ é a função taxa de falha acumulada, e o parâmetro κ , $0 < \kappa < 1$ é especificado previamente, vide Ibrahim *et al.* (2001) e Demarqui *et al.* (2014);

- Estender a abordagem bayesiana do modelo não paramétrico para incluir fração de cura;
- Estender a modelagem proposta para acomodar dados de sobrevivência com censura intervalar. Também pode-se realizar extensões para dados com eventos recorrentes.

Apêndice A

Uma aproximação para a matriz de informação de Fisher para os modelos com censura informativa

Louis (1982) demonstrou como obter uma matriz de informação de Fisher aproximada para o problema de dados incompletos.

Seja \mathbf{d} o conjunto de dados observados, e $\mathbf{d}_{comp} = (\mathbf{d}, \mathbf{w})$ o conjunto de dados completos, que incluem o conjunto de dados observados e o vetor de fragilidades não observadas, \mathbf{w} . Seja $\boldsymbol{\theta} = (\boldsymbol{\beta}^{(T)'}, \boldsymbol{\beta}^{(C)'}, \alpha, \sigma^2)$ o vetor de parâmetros estimados. Portanto, de acordo com Louis (1982), a matriz de informação de Fisher é dada por

$$I(\boldsymbol{\theta}) \approx -E \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \middle| \mathbf{d}, \boldsymbol{\theta} \right] - E \left[\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left[\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^t \middle| \mathbf{d}, \boldsymbol{\theta} \right],$$

em que as esperanças são calculadas com respeito ao vetor de fragilidades \mathbf{w} . O primeiro termo denota a esperança da matriz de derivadas parciais de segunda ordem, o segundo termo a esperança da matriz resultante do produto entre o vetor gradiente, $\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, e o vetor gradiente transposto, $\left[\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^t$.

A.1 Modelo de Huang & Wolfe (2002)

No caso do modelo proposto por Huang & Wolfe (2002), o vetor de taxas de falha de base é denotado por $\mathbf{h}^{(T)}(\cdot)$, com tamanho igual ao número de falhas distintas; e o

vetor de taxas de falha para os tempos de censura, denotado por $\mathbf{h}^{(C)}(\cdot)$, com tamanho igual ao número de censuras distintas. Assim, o vetor de parâmetros estimados é dado por $\boldsymbol{\theta} = (\boldsymbol{\beta}^{(T)'}, \boldsymbol{\beta}^{(C)'}, \alpha, \mathbf{h}^{(T)}(\cdot), \mathbf{h}^{(C)}(\cdot), \sigma^2)$.

Se os valores dos efeitos aleatórios $W_k, k = 1, \dots, m$ são conhecidos, então o logaritmo da verossimilhança para os dados completos é dado por

$$\begin{aligned} \log(L(\mathbf{d}; \boldsymbol{\theta}, w_k)) = & \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\log(h_0^{(T)}(y_{ik})) + \boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right] \right. \\ & - H_0^{(T)}(y_{ik}) \exp\left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}\right) \exp(w_k) \\ & + \delta_{ik}^{(C)} \left[\log(h_0^{(C)}(y_{ik})) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right] \\ & \left. - H_0^{(C)}(y_{ik}) \exp\left(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)}\right) \exp(\alpha w_k) \right\} \\ & - \frac{1}{2} [\log(2\pi) + \log(\sigma^2)] - \frac{1}{2\sigma^2} w_k^2. \end{aligned} \quad (\text{A.1})$$

As derivadas parciais de primeira ordem, são dadas por

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(T)} \left[\delta_{ik}^{(T)} - H_0^{(T)}(y_{ik}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right] \right\}, \quad (\text{A.2})$$

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(C)} \left[\delta_{ik}^{(C)} - H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right] \right\}, \quad (\text{A.3})$$

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ w_k \left[\delta_{ik}^{(C)} - H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right] \right\} \quad (\text{A.4})$$

e

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \sigma^2} = \sum_{i=1}^{n_k} \left(\frac{1}{2\sigma^4} w_k^2 - \frac{1}{2\sigma^2} \right). \quad (\text{A.5})$$

Para derivar as taxas de falha de base temos que lembrar que

$$H_0^{(T)}(y_{ik}) = \sum_{y_{jk} \leq y_{ik}} h_0^{(T)}(y_{jk}) \quad \text{e} \quad H_0^{(C)}(y_{ik}) = \sum_{y_{jk} \leq y_{ik}} h_0^{(C)}(y_{jk}).$$

Conseqüentemente,

$$\sum_{i=1}^{n_k} \sum_{y_{jk} \leq y_{ik}} h_0^{(T)}(y_{jk}) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) = \sum_{i=1}^{n_k} h_0^{(T)}(y_{ik}) \sum_{y_{jk} \geq y_{ik}} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{jk}^{(T)} + w_k),$$

e

$$\sum_{i=1}^{n_k} \sum_{y_{jk} \leq y_{ik}} h_0^{(C)}(y_{jk}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) = \sum_{i=1}^{n_k} h_0^{(C)}(y_{ik}) \sum_{y_{jk} \geq y_{ik}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{jk}^{(C)} + \alpha w_k).$$

Seja $R(y_{ik})$ o conjunto de indivíduos que estão sob risco em y_i , então $y_{jk} \geq y_{ik}$ equivale a $j \in R(y_{ik})$. Substituindo as expressões acima na equação (A.1), temos a log-verossimilhança,

$$\begin{aligned} \log(L(\mathbf{d}; \boldsymbol{\theta}, w_k)) &= \sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ik}^{(T)} \left[\log(h_0^{(T)}(y_{ik})) + \boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k \right] \\ &\quad - h_0^{(T)}(y_{ik}) \sum_{y_{jk} \geq y_{ik}} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{jk}^{(T)} + w_k) \\ &\quad + \delta_{ik}^{(C)} \left[\log(h_0^{(C)}(y_{ik})) + \boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k \right] \\ &\quad - h_0^{(C)}(y_{ik}) \sum_{y_{jk} \geq y_{ik}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{jk}^{(C)} + \alpha w_k) \\ &\quad - \frac{1}{2} [\log(2\pi) + \log(\sigma^2)] - \frac{1}{2\sigma^2} w_k^2. \end{aligned} \tag{A.6}$$

Derivando (A.6), temos

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial h_0^{(T)}(y_{ik})} = \frac{\delta_{ik}^{(T)}}{h_0^{(T)}(y_{ik})} - \sum_{y_{jk} \geq y_{ik}} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{jk}^{(T)} + w_k)$$

e

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial h_0^{(C)}(y_{ik})} = \frac{\delta_{ik}^{(C)}}{h_0^{(C)}(y_{ik})} - \sum_{y_{jk} \geq y_{ik}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{jk}^{(C)} + \alpha w_k).$$

As derivadas parciais de segunda ordem, são dadas por

$$\begin{aligned} \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \beta^{(T)} \partial \beta^{(T)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -H_0^{(T)}(y_{ik}) \exp\left(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)}\right) \mathbf{x}_{ik}^{(T)} \mathbf{x}_{ik}^{(T)'} \exp(w_k) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \left\{ h_0^{(T)}(y_{ik}) \right\}^2} &= - \frac{\delta_{ik}^{(T)}}{\left\{ h_0^{(T)}(y_{ik}) \right\}^2}, \end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial h_0^{(T)}(y_{ik})} &= - \sum_{y_{jk} \geq y_{ik}} x_{jk}^{(T)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{jk}^{(T)} + w_k), \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \boldsymbol{\beta}^{(C)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}) \mathbf{x}_{ik}^{(C)} \mathbf{x}_{ik}^{(C)' \exp(\alpha w_k)} \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \alpha} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}) \exp(\alpha w_k) w_k^2 \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \left\{ h_0^{(C)}(y_{ik}) \right\}^2} &= - \frac{\delta_{ik}^{(C)}}{\left\{ h_0^{(C)}(y_{ik}) \right\}^2}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \alpha} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -H_0^{(C)}(y_{ik}) \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}) \mathbf{x}_{ik}^{(C)} \exp(\alpha w_k) w_k \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial h_0^{(C)}(y_{ik})} &= - \sum_{y_{jk} \geq y_{ik}} x_{jk}^{(C)} \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{jk}^{(C)} + \alpha w_k), \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial h_0^{(C)}(y_{ik})} &= - \sum_{y_{jk} \geq y_{ik}} w_k \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{jk}^{(C)} + \alpha w_k), \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \sigma^2 \partial \sigma^2} &= \sum_{k=1}^m \left\{ -\frac{w_k^2}{(\sigma^2)^3} + \frac{1}{2(\sigma^2)^2} \right\}.
\end{aligned}$$

Os demais elementos da matriz são nulos.

A.2 Modelo Weibull

Consideramos a distribuição Weibull para ajustar os tempos de falha e censura, com funções taxa de falha dadas por

$$h_0^{(T)}(y) = \kappa^{(T)} y^{\kappa^{(T)} - 1} \gamma^{(T)} \quad \text{e} \quad h_0^{(C)}(y) = \kappa^{(C)} y^{\kappa^{(C)} - 1} \gamma^{(C)}.$$

O vetor de parâmetros estimados é dado por $\boldsymbol{\theta} = (\boldsymbol{\beta}^{(T)'}, \boldsymbol{\beta}^{(C)'}, \alpha, \kappa^{(T)}, \gamma^{(T)}, \kappa^{(C)}, \gamma^{(C)}, \sigma^2)$. As derivadas parciais de primeira ordem para os parâmetros $\boldsymbol{\beta}^{(T)}$, $\boldsymbol{\beta}^{(C)}$, α e σ^2 são dadas como nas equações (A.2) - (A.5).

As derivadas parciais de primeira ordem em relação aos parâmetros $\kappa^{(T)}$, $\gamma^{(T)}$, $\kappa^{(C)}$ e $\gamma^{(C)}$ são obtidas por

$$\begin{aligned}\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\kappa^{(T)}} + \log(y_{ik}) \right] - y_{ik}^{\kappa^{(T)}} \log(y_{ik}) \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) \right\}, \\ \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \gamma^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\gamma^{(T)}} \right] - y_{ik}^{\kappa^{(T)}} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) \right\}, \\ \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\kappa^{(C)}} + \log(y_{ik}) \right] - y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(C)}} + \alpha w_k) \right\}\end{aligned}$$

e

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \gamma^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\gamma^{(C)}} \right] - y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}} + \alpha w_k) \right\}.$$

As derivadas parciais de segunda ordem são dadas por

$$\begin{aligned}\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \boldsymbol{\beta}^{(T)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(T)}} \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) x_{ik}^{(T)} x_{ik}^{(T)'} \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(T)} \partial \kappa^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[-\frac{1}{(\kappa^{(T)})^2} \right] - y_{ik}^{\kappa^{(T)}} \log(y_{ik}) \log(y_{ik}) \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \kappa^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(T)}} \log(y_{ik}) \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) x_{ik}^{(T)} \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\gamma^{(T)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[-\frac{1}{(\gamma^{(T)})^2} \right] \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(T)} \partial \gamma^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(T)}} \log(y_{ik}) \gamma^{(T)} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \gamma^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(T)}} \exp(\boldsymbol{\beta}^{(T)' \mathbf{x}_{ik}^{(T)}} + w_k) x_{ik}^{(T)} \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \boldsymbol{\beta}^{(C)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}} + \alpha w_k) x_{ik}^{(C)} x_{ik}^{(C)'} \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \alpha} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}} + \alpha w_k) x_{ik}^{(C)} w_k \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\kappa^{(C)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[-\frac{1}{(\kappa^{(C)})^2} \right] - y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}} + \alpha w_k) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \kappa^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' \mathbf{x}_{ik}^{(C)}} + \alpha w_k) x_{ik}^{(C)} \right\},\end{aligned}$$

$$\begin{aligned} \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \kappa^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) w_k \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\gamma^{(C)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[-\frac{1}{(\gamma^{(C)})^2} \right] \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \gamma^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) x_{ik}^{(C)} \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \gamma^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) w_k \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(C)} \partial \gamma^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\} \end{aligned}$$

e

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\sigma^2)^2} = \sum_{k=1}^m \left\{ -\frac{w_k^2}{(\sigma^2)^3} + \frac{1}{2(\sigma^2)^2} \right\}.$$

Os demais elementos da matriz são nulos.

A.3 Modelo Exponencial por Partes

O vetor de taxas de falha de base é denotado por $\boldsymbol{\lambda}^{(T)} = (\lambda_1^{(T)}, \lambda_2^{(T)}, \dots, \lambda_j^{(T)}, \dots, \lambda_b^{(T)})'$, com tamanho igual bs ; e o vetor de taxas de falha para os tempos de censura, denotado por $\boldsymbol{\lambda}^{(C)} = (\lambda_1^{(C)}, \lambda_2^{(C)}, \dots, \lambda_l^{(C)}, \dots, \lambda_d^{(C)})'$, com tamanho igual d . Assim, o vetor de parâmetros estimados é dado por $\boldsymbol{\theta} = (\boldsymbol{\beta}^{(T)'}, \boldsymbol{\beta}^{(C)'}, \alpha, \boldsymbol{\lambda}^{(T)}, \boldsymbol{\lambda}^{(C)}, \sigma^2)$.

Se os valores dos efeitos aleatórios $W_k, k = 1, \dots, m$ são conhecidos, então o logaritmo da verossimilhança para os dados completos é dada por (3.18), em que $\mathbf{h}_0^{(T)}(\cdot) = \boldsymbol{\lambda}^{(T)}$ e $\mathbf{h}_0^{(C)}(\cdot) = \boldsymbol{\lambda}^{(C)}$. Conseqüentemente, as derivadas parciais de primeira ordem para os parâmetros $\beta^{(T)}, \beta^{(C)}, \alpha$ e σ^2 são dadas como nas equações (A.2) - (A.5).

Como as funções taxa de falha acumulada são dadas por

$$H_0^{(T)}(y_{ik}) = \sum_{j=1}^b \lambda_j^{(T)} (y_{ikj} - s_j) \quad \text{e} \quad H_0^{(C)}(y_{ik}) = \sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l),$$

Temos as seguintes derivadas de primeira ordem, com respeito as taxas de falhas $\lambda_j^{(T)}$ e $\lambda_l^{(C)}$,

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \lambda_j^{(T)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\frac{\delta_{ikj}^{(T)}}{\lambda_j^{(T)}} - (y_{ikj} - s_j) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right\}$$

e

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \lambda_l^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\frac{\delta_{ikl}^{(C)}}{\lambda_l^{(C)}} - (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}.$$

As derivadas parciais de segunda ordem são dadas por

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \boldsymbol{\beta}^{(T)'}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{j=1}^b \lambda_j^{(T)} (y_{ikj} - s_j) \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \mathbf{x}_{ik}^{(T)} \mathbf{x}_{ik}^{(T)'} \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \lambda_j^{(T)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -(y_{ikj} - s_j) x_{jk}^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \left\{ \lambda_j^{(T)} \right\}^2} = -\frac{\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikj}^{(T)}}{\left\{ \lambda_j^{(T)} \right\}^2},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \boldsymbol{\beta}^{(C)'}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} \mathbf{x}_{ik}^{(C)'} \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \alpha} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} w_k \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \lambda_j^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -(y_{ikl} - c_l) x_{jk}^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha^2} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) w_k^2 \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \lambda_j^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -(y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) w_k \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \left\{ \lambda_l^{(C)} \right\}^2} = -\frac{\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikl}^{(C)}}{\left\{ \lambda_l^{(C)} \right\}^2}$$

e

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\sigma^2)^2} = \sum_{k=1}^m \left\{ -\frac{w_k^2}{(\sigma^2)^3} + \frac{1}{2(\sigma^2)^2} \right\}.$$

Os demais elementos da matriz são nulos.

A.4 Modelo de fração de cura Weibull

Considerando o modelo Weibull com fração de cura proposto na Seção 3.4.1, temos o vetor de parâmetros denotado por $\boldsymbol{\theta} = (\boldsymbol{\beta}^{(T)'}, \boldsymbol{\beta}^{(C)'}, \alpha, \sigma^2, \kappa^{(Z)}, \kappa^{(C)}, \gamma^{(Z)}, \gamma^{(C)})$.

Seja \mathbf{d} o conjunto de dados observados, e $\mathbf{d}_{comp} = (\mathbf{d}, \mathbf{w})$ o conjunto de dados completos, que incluem o conjunto de dados observados e o vetor de fragilidades não observadas, \mathbf{w} .

As derivadas parciais de primeira ordem são dadas por

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(T)} \left[\delta_{ik}^{(T)} - \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \left(1 - \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \right) \right] \right\},$$

$$\begin{aligned} \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(Z)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\kappa^{(Z)}} + \log(y_{ik}) - y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \gamma^{(Z)} \right] \right. \\ &\quad \left. + \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \left(-y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \gamma^{(Z)} \right) \right\}, \end{aligned}$$

$$\begin{aligned} \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \gamma^{(Z)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\gamma^{(Z)}} - y_{ik}^{\kappa^{(Z)}} \right] \right. \\ &\quad \left. + \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \left(-y_{ik}^{\kappa^{(Z)}} \right) \right\}, \end{aligned}$$

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(C)} \left[\delta_{ik}^{(C)} - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right] \right\},$$

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ w_k \left[\delta_{ik}^{(C)} - y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right] \right\},$$

$$\begin{aligned} \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\kappa^{(C)}} + \log(y_{ik}) \right] \right. \\ &\quad \left. - y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}, \end{aligned}$$

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \gamma^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[\frac{1}{\gamma^{(C)}} \right] - y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}$$

e

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \sigma^2} = \sum_{i=1}^{n_k} \left\{ \frac{w_k^2}{2\sigma^4} - \frac{1}{2\sigma^2} \right\}.$$

As derivadas parciais de segunda ordem são dadas por

$$\begin{aligned}
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \boldsymbol{\beta}^{(T)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \mathbf{x}_{ik}^{(T)} \mathbf{x}_{ik}^{(T)'} \left[1 - \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \right] \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \kappa^{(Z)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \mathbf{x}_{ik}^{(T)} \gamma^{(Z)} y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \gamma^{(Z)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \mathbf{x}_{ik}^{(T)} y_{ik}^{\kappa^{(Z)}} \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\kappa^{(Z)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[\frac{1}{\kappa^2} - y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \log(y_{ik}) \gamma^{(Z)} \right] \right. \\
&\quad \left. - \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \gamma^{(Z)} \log(y_{ik}) y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \right. \\
&\quad \left. \times \left[1 - y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)} \right] \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(Z)} \partial \gamma^{(Z)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[-y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \right] - \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) \right. \\
&\quad \left. \times y_{ik}^{\kappa^{(Z)}} \log(y_{ik}) \left[1 - y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)} \right] \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\gamma^{(Z)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(T)} \left[-\frac{1}{(\gamma^{(Z)})^2} \right] + \exp(\boldsymbol{\beta}^{(T)' } \mathbf{x}_{ik}^{(T)} + w_k) \right. \\
&\quad \left. \times \exp(-y_{ik}^{\kappa^{(Z)}} \gamma^{(Z)}) y_{ik}^{2\kappa^{(Z)}} \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \boldsymbol{\beta}^{(C)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} \mathbf{x}_{ik}^{(C)'} \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \alpha} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} w_k \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\kappa^{(C)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[-\frac{1}{(\kappa^{(C)})^2} \right] - y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \log(y_{ik}) \gamma^{(C)} \right. \\
&\quad \left. \times \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \kappa^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} \right\}, \\
\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \kappa^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \gamma^{(C)} \exp(\boldsymbol{\beta}^{(C)' } \mathbf{x}_{ik}^{(C)} + \alpha w_k) w_k \right\},
\end{aligned}$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial(\gamma^{(C)})^2} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ik}^{(C)} \left[-\frac{1}{(\gamma^{(C)})^2} \right] \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \gamma^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \gamma^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \exp(x_{ik}^{(C)} \beta^{(C)} + \alpha w_k) w_k \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \kappa^{(C)} \partial \gamma^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -y_{ik}^{\kappa^{(C)}} \log(y_{ik}) \exp(x_{ik}^{(C)} \beta^{(C)} + \alpha w_k) \right\}$$

e

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial(\sigma^2)^2} = \sum_{k=1}^m \left\{ -\frac{w_k^2}{(\sigma^2)^3} + \frac{1}{2(\sigma^2)^2} \right\}.$$

Os demais elementos da matriz são nulos.

A.5 Modelo de fração de cura exponencial por partes

Nesta seção apresentamos as derivadas parciais necessárias para o cálculo da matriz de informação de Fisher, referentes ao modelo de fração de cura com MEP proposto na Seção 3.4.2.

Derivando a função log-verossimilhança condicional (3.27) com respeito aos parâmetros $\boldsymbol{\beta}^{(T)}$, $\lambda_j^{(Z)}$, $\lambda_i^{(C)}$, $\boldsymbol{\beta}^{(C)}$, α , σ^2 , encontramos as seguintes derivadas parciais de primeira ordem:

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(T)} \left[\delta_{ik}^{(T)} - \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right. \right. \\ \left. \left. \times \left(1 - \exp\left(-\sum_{j=1}^b \lambda_j^{(Z)} (y_{ij} - s_{j-1})\right) \right) \right] \right\},$$

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \lambda_j^{(Z)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ikj}^{(T)} \left[\frac{1}{\lambda_j^{(Z)}} + (y_{ij} - s_{j-1}) \right] \right. \\ \left. + \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \exp\left(-\sum_{j=1}^b \lambda_j^{(Z)} (y_{ij} - s_{j-1})\right) (-(y_{ij} - s_{j-1})) \right\},$$

$$\begin{aligned}\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(C)} \left[\delta_{ik}^{(C)} - \sum_{l=1}^d \lambda_l^{(C)} (y_{il} - s_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right] \right\}, \\ \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ w_k \left[\delta_{ik}^{(C)} - \sum_{l=1}^d \lambda_l^{(C)} (y_{il} - s_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right] \right\}, \\ \frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \lambda_l^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ikl}^{(C)} \left[\frac{1}{\lambda_l^{(C)}} \right] - (y_{il} - s_{l-1}) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \right\}\end{aligned}$$

e

$$\frac{\partial l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \sigma^2} = \sum_{i=1}^{n_k} \left\{ \frac{1}{2\sigma^4} (w_k^2) - \frac{1}{2\sigma^2} \right\}.$$

A seguir apresentamos as derivadas parciais de segunda ordem:

$$\begin{aligned}\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \boldsymbol{\beta}^{(T)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -x_{ik}^{(T)} x_{ik}^{(T)'} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \right. \\ &\quad \left. \times \left(1 - \exp\left(-\sum_{j=1}^b \lambda_j^{(Z)} (y_{ij} - s_{j-1})\right) \right) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(T)} \partial \lambda_j^{(T)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ x_{ik}^{(T)} \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \exp\left(-\sum_{j=1}^b \lambda_j^{(Z)} (y_{ij} - s_{j-1})\right) \right. \\ &\quad \left. \times (-(y_{ij} - s_{j-1})) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\lambda_j^{(Z)})^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ \delta_{ikj}^{(T)} \left[\frac{-1}{(\lambda_j^{(Z)})^2} \right] \right. \\ &\quad \left. + \exp(\boldsymbol{\beta}^{(T)'} \mathbf{x}_{ik}^{(T)} + w_k) \exp\left(-\sum_{j=1}^b \lambda_j^{(Z)} (y_{ij} - s_{j-1})\right) (-(y_{ij} - s_{j-1}))^2 \right\} \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \boldsymbol{\beta}^{(C)'}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) \mathbf{x}_{ik}^{(C)} \mathbf{x}_{ik}^{(C)'} \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \alpha} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) x_{ik}^{(C)} w_k \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \boldsymbol{\beta}^{(C)} \partial \lambda_j^{(C)}} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -(y_{ikl} - c_l) \mathbf{x}_{jk}^{(C)} \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{jk}^{(C)} + \alpha w_k) \right\}, \\ \frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha^2} &= \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -\sum_{l=1}^d \lambda_l^{(C)} (y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{ik}^{(C)} + \alpha w_k) w_k^2 \right\},\end{aligned}$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \alpha \partial \lambda_j^{(C)}} = \sum_{k=1}^m \sum_{i=1}^{n_k} \left\{ -(y_{ikl} - c_l) \exp(\boldsymbol{\beta}^{(C)'} \mathbf{x}_{jk}^{(C)} + \alpha w_k) w_k \right\},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial \left\{ \lambda_l^{(C)} \right\}^2} = - \frac{\sum_{k=1}^m \sum_{i=1}^{n_k} \delta_{ikl}^{(C)}}{\left\{ \lambda_l^{(C)} \right\}^2},$$

$$\frac{\partial^2 l(\mathbf{d}; \boldsymbol{\theta}, w_k)}{\partial (\sigma^2)^2} = \sum_{k=1}^m \left\{ -\frac{w_k^2}{(\sigma^2)^3} + \frac{1}{2(\sigma^2)^2} \right\}.$$

Os demais elementos da matriz são nulos.

Apêndice B

Estudos com dados simulados para os modelos com censura informativa

Neste apêndice apresentamos alguns resultados complementares sobre os modelos com censura informativa expostos na Seção 3.1. Primeiramente avaliamos a convergência dos estimadores de máxima verossimilhança, vide Figura B.1 para o modelo proposto por Huang & Wolfe (2002) e Figura B.2 para o modelo MEP.

Na Tabela B.1 relatamos a frequência do número de NA's (observações que não apresentaram estimativas por causa de problemas numéricos) obtidos no cálculo do erro padrão assintótico proposto por Louis (1982). Essas amostras não foram descartadas, portanto, no cálculo das demais estimativas são considerados 500 conjuntos de dados.

Posteriormente mostramos a comparação dos critérios bayesianos LPML, DIC e WAIC, para dois conjuntos de dados, vide Tabela B.2. Podemos observar que os melhores ajustes foram com 10 intervalos (Conjunto 1) e 20 intervalos (Conjunto 2). Portanto, nas simulações consideramos 10, 20 e número de intervalos igual ao número de falhas e censuras distintas, para o modelo MEP.

Para escolha da distribuição *a priori* do parâmetro de precisão τ , realizamos um estudo de sensibilidade utilizando as distribuições Gama(10; 10), Gama(1; 1) e Gama(0, 1; 0, 1) nos cenários: correlação positiva, vide Tabela B.3; correlação negativa, vide Tabela B.4, e correlação nula, vide Tabela B.5.

Como podemos analisar nas Tabelas B.3, B.4 e B.5 os ajustes considerando as distribuições $G(1; 1)$ e $G(10; 10)$ foram bem similares. Os valores dos critérios apontaram que o ajuste com $G(10; 10)$ apresentou um desempenho levemente melhor. No entanto, optamos por utilizar a distribuição $G(1; 1)$.

Tabela B.1: Frequência do número de NA's encontrados no cálculo do erro padrão estimado proposto por Louis, considerando 500 réplicas.

	CI	CNI	CI	CNI	CI	CNI
	$\alpha = 1$		$\alpha = -1$		$\alpha = 0$	
M0						
$\beta_1^{(T)}$	6	1	1	1	1	0
$\beta_2^{(T)}$	1	1	3	4	3	0
$\beta_1^{(C)}$	2		1		1	
$\beta_2^{(C)}$	1		1		1	
α	14		28		4	
σ^2	14	5	28	8	6	2
M1						
$\beta_1^{(T)}$	15	18	10	3	8	3
$\beta_2^{(T)}$	6	10	19	9	9	2
$\beta_1^{(C)}$	4		8		2	
$\beta_2^{(C)}$	5		9		5	
α	76		129		16	
σ^2	79	56	131	35	31	28
M2						
$\beta_1^{(T)}$	18	14	8	1	5	3
$\beta_2^{(T)}$	6	11	17	12	8	6
$\beta_1^{(C)}$	7		5		1	
$\beta_2^{(C)}$	11		7		2	
α	81		116		12	
σ^2	84	53	119	36	35	21
M3						
$\beta_1^{(T)}$	14	9	8	5	5	5
$\beta_2^{(T)}$	6	7	18	14	8	4
$\beta_1^{(C)}$	2		5		2	
$\beta_2^{(C)}$	4		4		1	
α	64		112		15	
σ^2	69	57	115	32	37	23
M4						
$\beta_1^{(T)}$	14	14	13	19	13	21
$\beta_2^{(T)}$	17	46	68	91	44	61
$\beta_1^{(C)}$	40		18		5	
$\beta_2^{(C)}$	201		66		50	
α	141		141		25	
σ^2	82	18	68	22	25	16

¹CI: censura informativa, CNI: censura não informativa.

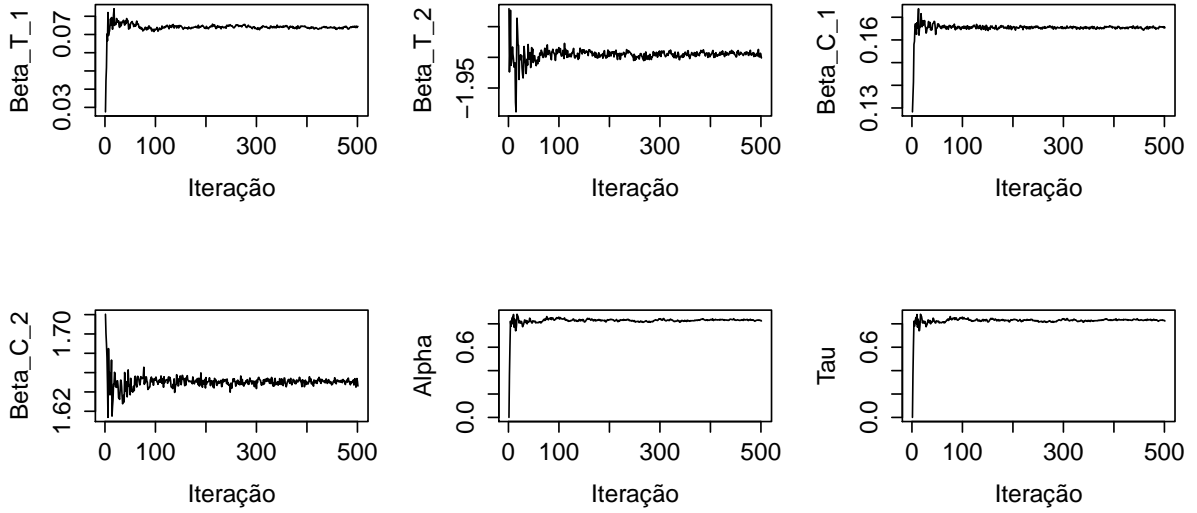


Figura B.1: Convergência para os parâmetros do modelo de Huang & -Wolfe (2002).

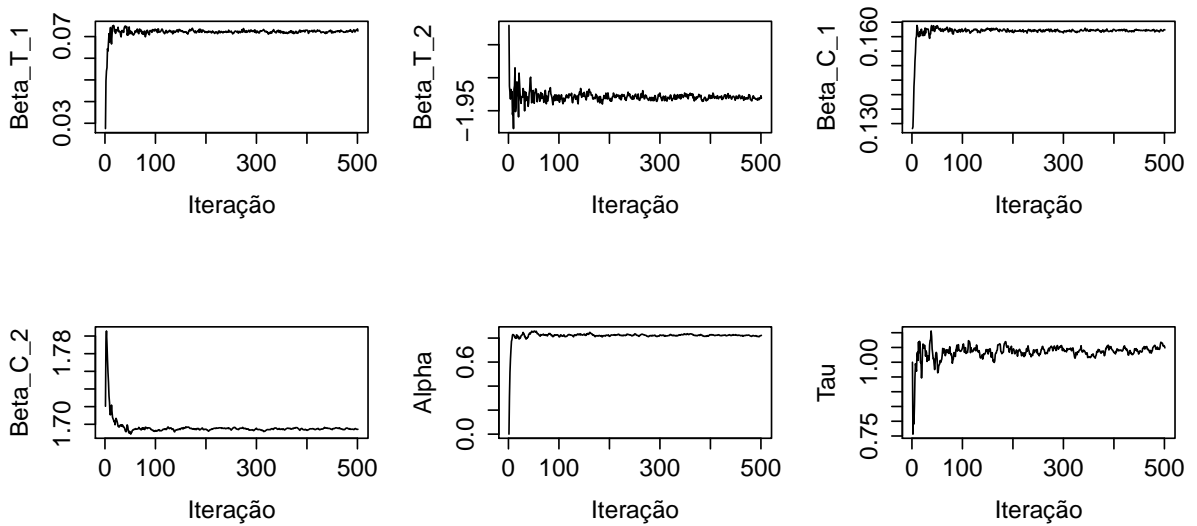


Figura B.2: Convergência para os parâmetros do modelo de fragilidade com distribuição Exponencial por Partes para os tempos.

Tabela B.2: Avaliação dos critérios de informação para o modelo de fragilidade com distribuição Exponencial por Partes para os tempos.

Nº Intervalos	LPML	DIC	WAIC	LPML	DIC	WAIC
	Conjunto 1			Conjunto 2		
1	812,150	830,709	811,938	825,335	836,830	825,189
2	773,037	800,902	772,484	783,029	798,485	782,660
4	751,729	784,762	751,038	757,459	772,631	756,570
6	746,319	780,698	745,490	750,387	764,575	749,506
8	742,691	774,410	741,614	753,268	766,850	752,107
10	741,718	771,214	740,685	758,132	774,458	756,602
20	747,511	781,371	746,301	749,049	762,142	747,601
30	766,521	794,853	764,512	769,323	778,969	768,008
40	776,243	802,914	773,177	779,779	788,158	777,623
50	788,737	812,704	784,315	791,68	794,085	788,376
60	819,499	827,391	810,725	814,067	812,989	808,439
70	843,938	842,918	831,454	839,713	828,048	829,277
80	854,351	847,036	834,904	862,267	830,967	847,325
90	906,741	859,185	864,897	903,87	854,134	873,965
100	970,479	864,104	890,811	963,599	874,466	905,031
Nº falhas dist,	1278,097	983,012	1067,547	1338,618	964,317	1087,261

¹Nº falhas dist.: número de intervalos igual ao número de falhas e censuras distintas.

Tabela B.3: Comparação entre as distribuições *a priori* para τ considerando o Modelo MEP* em dados com correlação positiva.

	Média	Moda	Mediana	HPD		Viés rel.(%)	LPML	DIC	WAIC	
Distribuição <i>a priori</i> Gama(0,1;0,1) para τ										
$\beta_1^{(T)}$	0,1	0,094	0,093	0,094	0,043	0,146	-6,826	361,514	388,863	359,437
$\beta_2^{(T)}$	-1,4	-1,485	-1,465	-1,477	-2,104	-0,877	-4,619			
$\beta_1^{(C)}$	0,2	0,207	0,204	0,206	0,143	0,273	2,242			
$\beta_2^{(C)}$	1,2	1,311	1,283	1,301	0,630	2,011	6,878			
α	1	1,252	1,048	1,168	0,460	2,235	4,834			
τ	1	1,854	1,038	1,445	0,363	4,447	3,843			
Distribuição <i>a priori</i> Gama(1;1) para τ										
$\beta_1^{(T)}$	0,1	0,095	0,095	0,095	0,044	0,147	-5,319	360,828	381,911	358,653
$\beta_2^{(T)}$	-1,4	-1,491	-1,470	-1,483	-2,114	-0,884	-5,012			
$\beta_1^{(C)}$	0,2	0,207	0,205	0,206	0,143	0,273	2,286			
$\beta_2^{(C)}$	1,2	1,305	1,275	1,295	0,626	2,002	6,236			
α	1	1,133	1,011	1,089	0,493	1,863	1,113			
τ	1	1,327	0,990	1,196	0,414	2,546	-0,966			
Distribuição <i>a priori</i> Gama(10;10) para τ										
$\beta_1^{(T)}$	0,1	0,096	0,096	0,096	0,046	0,147	-3,950	360,147	374,964	358,014
$\beta_2^{(T)}$	-1,4	-1,495	-1,470	-1,487	-2,114	-0,889	-5,007			
$\beta_1^{(C)}$	0,2	0,207	0,204	0,206	0,143	0,272	2,161			
$\beta_2^{(C)}$	1,2	1,299	1,266	1,289	0,622	1,993	5,511			
α	1	1,050	0,989	1,028	0,516	1,628	-1,121			
τ	1	1,054	0,968	1,025	0,566	1,594	-3,154			
Censura Não Iformativa										
Distribuição <i>a priori</i> Gama(0,1;0,1) para τ										
$\beta_1^{(T)}$	0,1	0,085	0,084	0,085	0,033	0,137	-15,914	369,803	402,644	366,104
$\beta_2^{(T)}$	-1,4	-1,589	-1,559	-1,580	-2,235	-0,959	-11,382			
τ	1	1,960	1,019	1,473	0,338	4,915	1,939			
Distribuição <i>a priori</i> Gama(1;1) para τ										
$\beta_1^{(T)}$	0,1	0,086	0,086	0,086	0,034	0,139	-14,445	369,176	391,309	365,257
$\beta_2^{(T)}$	-1,4	-1,601	-1,576	-1,593	-2,247	-0,969	-12,602			
τ	1	1,335	0,987	1,195	0,397	2,601	-1,312			
Distribuição <i>a priori</i> Gama(10;10) para τ										
$\beta_1^{(T)}$	0,1	0,087	0,087	0,087	0,036	0,139	-13,262	368,697	381,038	364,659
$\beta_2^{(T)}$	-1,4	-1,609	-1,582	-1,601	-2,252	-0,979	-13,023			
τ	1	1,052	0,973	1,024	0,560	1,599	-2,733			

¹* número de intervalos = 10.

Tabela B.4: Comparação entre as distribuições *a priori* para τ considerando o Modelo MEP* em dados com correlação negativa.

	Média	Moda	Mediana	HPD		Viés rel.(%)	LPML	DIC	WAIC	
Distribuição <i>a priori</i> Gama(0,1;0,1) para τ										
$\beta_1^{(T)}$	0,1	0,100	0,099	0,099	0,050	0,149	-1,039	352,411	375,584	349,952
$\beta_2^{(T)}$	-1,4	-1,410	-1,390	-1,404	-2,000	-0,831	0,699			
$\beta_1^{(C)}$	0,2	0,210	0,208	0,209	0,147	0,274	3,867			
$\beta_2^{(C)}$	1,2	1,256	1,235	1,249	0,626	1,902	2,893			
α	-1	-1,292	-1,085	-1,208	-2,271	-0,501	-8,479			
τ	1	1,750	1,031	1,398	0,355	4,070	3,107			
Distribuição <i>a priori</i> Gama(1;1) para τ										
$\beta_1^{(T)}$	0,1	0,100	0,099	0,099	0,050	0,149	-0,858	351,858	369,816	349,266
$\beta_2^{(T)}$	-1,4	-1,425	-1,410	-1,419	-2,013	-0,848	-0,682			
$\beta_1^{(C)}$	0,2	0,209	0,208	0,209	0,147	0,273	3,779			
$\beta_2^{(C)}$	1,2	1,257	1,236	1,249	0,626	1,902	2,977			
α	-1	-1,184	-1,048	-1,135	-1,951	-0,518	-4,820			
τ	1	1,292	0,975	1,169	0,401	2,461	-2,510			
Distribuição <i>a priori</i> Gama(10;10) para τ										
$\beta_1^{(T)}$	0,1	0,100	0,099	0,100	0,050	0,150	-0,636	351,354	363,774	348,737
$\beta_2^{(T)}$	-1,4	-1,435	-1,418	-1,429	-2,012	-0,869	-1,254			
$\beta_1^{(C)}$	0,2	0,209	0,207	0,208	0,146	0,273	3,451			
$\beta_2^{(C)}$	1,2	1,256	1,231	1,248	0,624	1,898	2,608			
α	-1	-1,108	-1,030	-1,079	-1,725	-0,550	-3,016			
τ	1	1,046	0,967	1,018	0,560	1,583	-3,324			
Censura Não Informativa										
Distribuição <i>a priori</i> Gama(0,1;0,1) para τ										
$\beta_1^{(T)}$	0,1	0,110	0,109	0,110	0,058	0,162	9,449	359,036	385,914	355,036
$\beta_2^{(T)}$	-1,4	-1,355	-1,331	-1,347	-1,967	-0,754	4,911			
τ	1	1,909	1,070	1,487	0,363	4,580	6,983			
Distribuição <i>a priori</i> Gama(1;1) para τ										
$\beta_1^{(T)}$	0,1	0,111	0,110	0,110	0,059	0,163	10,103	358,586	376,539	354,303
$\beta_2^{(T)}$	-1,4	-1,373	-1,353	-1,366	-1,983	-0,772	3,358			
τ	1	1,346	1,012	1,217	0,417	2,571	1,243			
Distribuição <i>a priori</i> Gama(10;10) para τ										
$\beta_1^{(T)}$	0,1	0,1112	0,1106	0,1109	0,0592	0,1633	10,633	358,0854	368,036	353,738
$\beta_2^{(T)}$	-1,4	-1,388	-1,376	-1,3827	-1,987	-0,798	1,685			
τ	1	1,0601	0,9787	1,0319	0,5716	1,6029	-2,129			

¹* número de intervalos = 10.

Tabela B.5: Comparação entre as distribuições *a priori* para τ considerando o Modelo MEP* em dados com correlação nula.

	Média	Moda	Mediana	HPD		Viés rel,(%)	LPML	DIC	WAIC	
Distribuição <i>a priori</i> Gama(0,1;0,1) para τ										
$\beta_1^{(T)}$	0,1	0,096	0,095	0,096	0,045	0,148	-4,552	372,267	410,388	368,563
$\beta_2^{(T)}$	-1,4	-1,425	-1,403	-1,417	-2,037	-0,825	-0,219			
$\beta_1^{(C)}$	0,2	0,209	0,208	0,209	0,150	0,269	3,787			
$\beta_2^{(C)}$	1,2	1,276	1,248	1,267	0,648	1,917	4,015			
α	0	-0,071	-0,052	-0,063	-0,769	0,605	-5,238			
τ	1	1,989	1,024	1,478	0,345	5,036	2,366			
Distribuição <i>a priori</i> Gama(1;1) para τ										
$\beta_1^{(T)}$	0,1	0,097	0,096	0,097	0,045	0,149	-3,804	371,365	396,699	367,411
$\beta_2^{(T)}$	-1,4	-1,439	-1,423	-1,432	-2,048	-0,839	-1,642			
$\beta_1^{(C)}$	0,2	0,209	0,207	0,208	0,150	0,269	3,496			
$\beta_2^{(C)}$	1,2	1,274	1,248	1,265	0,646	1,912	3,988			
α	0	-0,061	-0,049	-0,056	-0,593	0,459	-4,859			
τ	1	1,310	0,978	1,179	0,401	2,523	-2,154			
Distribuição <i>a priori</i> Gama(10;10) para τ										
$\beta_1^{(T)}$	0,1	0,097	0,097	0,097	0,046	0,149	-3,001	370,681	385,656	366,665
$\beta_2^{(T)}$	-1,4	-1,448	-1,431	-1,442	-2,050	-0,857	-2,247			
$\beta_1^{(C)}$	0,2	0,208	0,207	0,208	0,150	0,268	3,382			
$\beta_2^{(C)}$	1,2	1,272	1,245	1,264	0,647	1,911	3,738			
α	0	-0,056	-0,045	-0,052	-0,514	0,397	-4,464			
τ	1	1,045	0,962	1,017	0,560	1,581	-3,768			
Censura Não Informativa										
Distribuição <i>a priori</i> Gama(0,1;0,1) para τ										
$\beta_1^{(T)}$	0,1	0,099	0,099	0,099	0,047	0,152	-1,389	370,624	398,16	366,541
$\beta_2^{(T)}$	-1,4	-1,436	-1,417	-1,429	-2,054	-0,828	-1,243			
τ	1	1,631	0,963	1,293	0,337	3,803	-3,679			
Distribuição <i>a priori</i> Gama(1;1) para τ										
$\beta_1^{(T)}$	0,1	0,100	0,099	0,100	0,048	0,152	-0,027	370,215	389,752	365,89
$\beta_2^{(T)}$	-1,4	-1,447	-1,425	-1,440	-2,065	-0,842	-3,323			
τ	1	1,228	0,927	1,111	0,388	2,334	22,764			
Distribuição <i>a priori</i> Gama(10;10) para τ										
$\beta_1^{(T)}$	0,1	0,1002	0,0997	0,1	0,0486	0,1522	-0,320	370,112	388,453	366,100
$\beta_2^{(T)}$	-1,4	-1,450	-1,437	-1,444	-2,059	-0,854	-2,628			
τ	1	1,0251	0,9442	0,9972	0,5508	1,553	-5,577			

¹* número de intervalos = 10.

Tabela B.6: Estatísticas sumário.

Distribuição	Mínimo	1° Quartil	Mediana	Média	3° Quartil	Máximo
G(10;10)	0,233	0,759	0,969	1,002	1,203	2,530
G(1;1)	0,001	0,289	0,681	1,049	1,455	9,947
G(0,1;0,1)	0,000	0,000	0,007	1,054	0,341	28,210

Apêndice C

Seleção de modelos

George E. P. Box, um dos estatísticos mais brilhantes do final do século XX, publicou em seu livro *Empirical Model-Building and Response Surfaces* (Box & Draper, 1987), a famosa frase “Todos os modelos são errados, mas alguns são úteis”. Dentre os modelos possíveis, precisamos mostrar qual(ou quais) é(são) o(s) modelo(s) que melhor se ajusta(m) ao problema.

Existem vários critérios para escolher o modelo com o melhor ajuste, destes, utilizamos os critérios bayesianos. A abordagem bayesiana padrão para seleção de modelos é baseada na distribuição preditiva, pois esta permite comparações de modelos arbitrários, ao passo que nas comparações utilizando distribuições *a posteriori* os modelos precisam ser “encaixados” (Gelfand *et al.*, 1992).

Nas próximas Seções apresentamos um breve resumo sobre os três critérios empregados para avaliar os modelos: estatística logaritmo da pseudo verossimilhança marginal (LPML), critério de informação do desvio (DIC) e critério de informação de Watanabe (WAIC).

C.1 Critério de informação da desviância

$$D(\boldsymbol{\theta}) = -2 \sum_{i=1}^n \log[f(y_i|\boldsymbol{\theta})],$$

e

$$\bar{D} = \frac{1}{R} \sum_{r=1}^R D(\theta_r).$$

em que r é o r -ésimo elemento de uma amostra de tamanho R . O AIC é obtido por

$$AIC = \bar{D} + 2q,$$

em que q é o número de parâmetros do modelo.

Em muitos modelos complexos, tais como modelos hierárquicos, é complicado saber o número de parâmetros. Então Spiegelhalter *et al.* (2002) propuseram uma generalização do AIC, substituindo a medida do número de parâmetros pela estimativa do número efetivo de parâmetros, utilizando média *a posteriori* do desvio. Ou seja,

$$\rho_D = E[D(\boldsymbol{\theta})|\mathbf{y}] - D[E(\boldsymbol{\theta})|\mathbf{y}].$$

Assim, o critério da informação do desvio é dado por

$$DIC = \bar{D} + \hat{\rho}_D.$$

Quanto menor o valor de DIC melhor o ajuste do modelo.

C.2 Estatística logaritmo da pseudoverossimilhança marginal

Outro critério de comparação de modelos muito utilizado na comparação de modelos é estatística logaritmo da pseudo verossimilhança marginal, proposto por Gelfand *et al.* (1992). Este critério é baseado na distribuição preditiva.

Sejam J modelos propostos com M_j denotado por $f(\mathbf{Y}|\theta_j; X, M_j) \times \pi(\theta_j)$. Se w_j denota a probabilidade *a priori* para M_j , então a probabilidade de M_j é

$$p(M_j|\mathbf{Y}) = \frac{f(\mathbf{Y}|M_j) \times w_j}{\sum_{j=1}^J f(\mathbf{Y}|M_j) \times w_j},$$

em que $f(\mathbf{Y}|M_j)$ é a distribuição preditiva ou distribuição marginal conjunta dos dados, sob o modelo M_j . Para \mathbf{y} o modelo que produz maior $p(M_j|\mathbf{y})$ será o modelo selecionado.

Gelfand *et al.* (1992) propõem validar as distribuições preditivas condicionais resultante da omissão de um valor das respostas observadas. Ou seja, a distribuição preditiva dada por

$$f(Y_i|Y_{(-i)}) = \int f(Y_i|Y_{(-i)}, \theta, X) \pi(\theta|Y_{(-i)}) d\theta,$$

em que Y_i é a resposta do i -ésimo indivíduo e $Y_{(-i)}$ é o vetor $n - 1 \times 1$ de respostas.

Os autores, Gelfand *et al.* (1992), comentam que amostrar de $f(Y_r|Y_{(-r)}, \theta, X)$ não tem

problema na maioria das vezes, porém amostrar de $\pi(\theta|y_{(-i)})$ pode ser complicado. Nesse caso, propõem utilização do amostrador de Gibbs para produzir a *posteriori* conjunta $\pi(\theta|\mathbf{y})$.

A ordenada da distribuição preditiva condicional para a i -ésima observação CPO_i pode ser interpretada como o peso da densidade marginal para esta observação. A CPO_i pode ser calculado pela aproximação,

$$C\hat{P}O_i = \left[\frac{1}{R} \sum_{r=1}^R \frac{1}{f(y_i|\theta_r)} \right]^{-1},$$

sendo $\theta_1, \dots, \theta_R$ uma amostra de tamanho R da distribuição *a posteriori* $\pi(\theta|\mathbf{y})$, mais detalhes podem ser vistos em (Chen *et al.*, 2012).

A partir do CPO_i podemos calcular a estatística chamada logaritmo da pseudo verossimilhança marginal (LPML), vide Ghosh & Gönen (2008), dada por

$$LPML = \sum_{i=1}^n \log CPO_i,$$

em que $CPO_i = f(y_i|y_{(-i)})$, $y_{(-i)} = (y_1, y_2, \dots, y_{i-1}, y_{i+1}, \dots, y_n)$. Quanto maior o valor de LPML, melhor é o ajuste.

C.3 Critério de informação de Watanabe (2010)

Segundo Watanabe (2010), muitos modelos estatísticos são singulares, isto é a matriz de Informação de Fisher não é positiva definida. Por exemplos mistura de normais, mistura de distribuições de probabilidade, modelos ocultos de Markov, são singulares. Em modelos singulares a estimador de máxima verossimilhança não satisfaz a normalidade assintótica, logo o AIC não é igual a esperança do erro.

Portanto, o autor Watanabe (2010) propôs um critério que levado em consideração a variância do logaritmo da distribuição preditiva para cada observação. Em que a distribuição preditiva de Bayes é definida por

$$f^*(x) = E_{\theta} [f(x|\theta)].$$

A função perda definida por

$$B_tL(n) = -\frac{1}{n} \sum_{i=1}^n \log f^*(X_i).$$

A função da variância definida por

$$V(n) = \sum_{i=1}^n \{E_{\theta} [(\log f(X_i|\theta))^2] - E_{\theta} [\log f(X_i|\theta)]^2\},$$

que mostra a flutuação a distribuição *a posteriori*. Logo, o critério proposto por Watanabe (2010) resulta em

$$WAIC(n) = B_t L(n) + \frac{\beta}{n} V(n),$$

em que $0 < \beta < \infty$ e quando $\beta = 1$ é o caso mais importante, pois corresponde a estimação de Bayes.

Em Gelman *et al.* (2014b) podemos encontrar uma comparação entre os critérios AIC, DIC e WAIC. No qual os autores (Gelman *et al.*, 2014b), concluem que os resultados do WAIC são similares aos resultados do DIC, porém a correção para o número efetivo de parâmetros é mais consistente que o comportamento assintótico do WAIC somente quando número de parâmetros for igual ao número de observações.

Bibliografia

- Aalen, Odd. 1978. Nonparametric inference for a family of counting processes. *The Annals of Statistics*, 701–726.
- Andersen, Per Kragh, & Gill, Richard David. 1982. Cox's regression model for counting processes: a large sample study. *The Annals of Statistics*, **10**(4), 1100–1120.
- Anscombe, Francis John. 1961. Estimating a mixed-exponential response law. *Journal of the American Statistical Association*, **56**(295), 493–502.
- Arjas, Elja, & Gasbarra, Dario. 1994. Nonparametric Bayesian inference from right censored survival data, using the Gibbs sampler. *Statistica Sinica*, **4**, 505–524.
- Aslanidou, Helen, Dey, Dipak K, & Sinha, Debajyoti. 1998. Bayesian analysis of multivariate survival data using Monte Carlo methods. *Canadian Journal of Statistics*, **26**(1), 33–48.
- Bandeem-Roche, Karen, & Liang, Kung-Yee. 2002. Modelling multivariate failure time associations in the presence of a competing risk. *Biometrika*, **89**(2), 299–314.
- Banerjee, Sudipto, Wall, Melanie M, & Carlin, Bradley P. 2003. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. *Biostatistics*, **4**(1), 123–142.
- Barbosa, Emanuel P, Colosimo, Enrico A, & Louzada, Francisco. 1996. Accelerated life tests analyzed by a piecewise exponential distribution via generalized linear models. *IEEE Transactions on Reliability*, **45**(4), 619–623.
- Barry, Daniel, & Hartigan, John A. 1992. Product partition models for change point problems. *The Annals of Statistics*, **20**(1), 260–279.

- Berkson, Joseph, & Gage, Robert P. 1952. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, **47**(259), 501–515.
- Boag, John W. 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. *Journal of the Royal Statistical Society*, **B11**(1), 15–53.
- Booth, James G, & Hobert, James P. 1999. Maximizing generalized linear mixed model likelihoods with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society*, **B61**(1), 265–285.
- Box, George EP, & Draper, Norman Richard. 1987. *Empirical model-building and response surfaces*. Vol. 424. Wiley New York.
- Braekers, Roel, & Veraverbeke, Noël. 2005. Cox’s regression model under partially informative censoring. *Communications in Statistics Theory and Methods*, **34**(8), 1793–1811.
- Breitner, JCS, Wyse, BW, Anthony, JC, Welsh-Bohmer, KA, Steffens, DC, Norton, Maria C, Tschanz, JT, Plassman, BL, Meyer, MR, Skoog, I, *et al.* 1999. APOE- ϵ 4 count predicts age when prevalence of AD increases, then declines The Cache County Study. *Neurology*, **53**(2), 321–321.
- Breslow, Norman. 1972. Contribution to the discussion on the Cox (1972) paper. *D. R. Cox (1972) paper*.
- Breslow, Norman. 1974. Covariance analysis of censored survival data. *Biometrics*, **30**(1), 89–99.
- Broadhurst, Roderic G, & Mailer, Ross A. 1990. The recidivism of prisoners released for the first time: Reconsidering the effectiveness question. *Australian & New Zealand Journal of Criminology*, **23**(2), 88–104.
- Brooks, Steve, Gelman, Andrew, Jones, Galin, & Meng, Xiao-Li. 2011. *Handbook of markov chain monte carlo*. CRC press.
- Calsavara, Vinicius F, Tomazella, Vera LD, & Fogo, José C. 2013. The effect of frailty term in the standard mixture model. *Chilean Journal of Statistics*, **4**(2), 95–109.
- Cancho, Vicente G, Rodrigues, Josemar, & Castro, Mário de. 2011. A flexible model for survival data with a cure rate: a Bayesian approach. *Journal of Applied Statistics*, **38**(1), 57–70.

- Carmichael, Vicki E, & Wilson, Kenneth S. 1992. Primary cutaneous malignant melanoma: experience of the British Columbia Cancer Agency from 1972 to 1981. *Canadian Journal of Surgery. Journal Canadien de Chirurgie*, **35**(6), 589–597.
- Carvalho, Charles André, Giugliani, Roberto, Ashton-Prolla, Patrícia, da Cunha, Mauricio Estrela, & Bakos, Lucio. 2004. Melanoma hereditário: prevalência de fatores de risco em um grupo de pacientes no Sul do Brasil. *Anais Brasileiros de Dermatologia*, **79**(1), 53–60.
- Castro, Mário de, Cancho, Vicente G, & Rodrigues, Josemar. 2009. A Bayesian long-term survival model parametrized in the cured fraction. *Biometrical Journal*, **51**(3), 443–455.
- Chaves, Josenildo de Souza, & Rodrigues, Josemar. 2011. Standard Exponential Cure Rate Model with Noninformative or Informative Uniform-Exponential Censoring. *Communications in Statistics—Simulation and Computation*, **40**(3), 364–382.
- Chen, Junliang, Zhang, Daowen, & Davidian, Marie. 2002a. A Monte Carlo EM algorithm for generalized linear mixed models with flexible random effects distribution. *Biostatistics*, **3**(3), 347–360.
- Chen, Ming-Hui, & Ibrahim, Joseph G. 2001. Maximum likelihood methods for cure rate models with missing covariates. *Biometrics*, **57**(1), 43–52.
- Chen, Ming-Hui, Ibrahim, Joseph G, & Sinha, Debajyoti. 1999. A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, **94**(447), 909–919.
- Chen, Ming-Hui, Harrington, David P, & Ibrahim, Joseph G. 2002b. Bayesian cure rate models for malignant melanoma: a case-study of Eastern Cooperative Oncology Group trial E1690. *Journal of the Royal Statistical Society*, **C51**(2), 135–150.
- Chen, Ming-Hui, Shao, Qi-Man, & Ibrahim, Joseph G. 2012. *Monte Carlo methods in Bayesian computation*. Springer Series in Statistics.
- Cheng, Yu, Fine, Jason P, & Bandeen-Roche, Karen. 2010. Association analyses of clustered competing risks data via cross hazard ratio. *Biostatistics*, **11**(1), 82–92.
- Chiang, Chin-Tsang, & Wang, Mei-Cheng. 2004. Smoothing estimation of rate function for recurrent event data with informative censoring. *Annals of the Institute of Statistical Mathematics*, **56**(1), 87–100.

- Clayton, David, & Cuzick, Jack. 1985. Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society*, **A148**(2), 82–117.
- Clayton, David G. 1978. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, **65**(1), 141–151.
- Clayton, David G. 1991. A Monte Carlo method for Bayesian inference in frailty models. *Biometrics*, **47**(2), 467–485.
- Colosimo, Enrico Antônio, & Giolo, Suely Ruiz. 2006. *Análise de sobrevivência aplicada*. 1rd edn. São Paulo: Blücher.
- Cox, David R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society*, **B34**(2), 187–220.
- Criado, PR, Vasconcellos, C, Sittart, JAS, Valente, NYS, Moura, BPS, Barbosa, GL, & Ichihara, C. 1999. Melanoma maligno cutâneo primário: estudo retrospectivo de 1963 a 1997 no Hospital do Servidor Público Estadual de São Paulo. *Revista da Associação Médica Brasileira*, **45**(2), 157–62.
- Danieli, Coraline, Remontet, Laurent, Bossard, Nadine, Roche, Laurent, & Belot, Aurélien. 2012. Estimating net survival: the importance of allowing for informative censoring. *Statistics in Medicine*, **31**(8), 775–786.
- Demarqui, Fabio N, Loschi, Rosangela H, & Colosimo, Enrico A. 2008. Estimating the grid of time-points for the piecewise exponential model. *Lifetime Data Analysis*, **14**(3), 333–356.
- Demarqui, Fabio N, Loschi, Rosangela H, Dey, Dipak K, & Colosimo, Enrico A. 2012. A class of dynamic piecewise exponential models with random time grid. *Journal of Statistical Planning and Inference*, **142**(3), 728–742.
- Demarqui, Fabio N, Dey, Dipak K, Loschi, Rosangela H, & Colosimo, Enrico A. 2014. Fully semiparametric Bayesian approach for modeling survival data with cure fraction. *Biometrical Journal*, **56**(2), 198–218.
- Demarqui, Fábio Nogueira. 2010. *Uma classe mais flexível de modelos semiparamétricos para dados de sobrevivência*. Ph.D. thesis, Universidade Federal de Minas Gerais.

- Dempster, Arthur P, Laird, Nan M, & Rubin, Donald B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society*, **B39**(1), 1–38.
- Dimatos, Dimitri C., Duarte, Felipe O., S., Machado Robson, Vieira, Vilberto J., A., Vasconcellos Zulmar, & Neves, Rodrigo D. 2009. Melanoma cutâneo no Brasil. *Arquivos Catarinenses de Medicina*, **38**(Suplemento 01), 14–19.
- Ding, A Adam. 2010. Identifiability conditions for covariate effects model on survival times under informative censoring. *Statistics & Probability Letters*, **80**(11), 911–915.
- Dunson, David B, & Dinse, Gregg E. 2002. Bayesian models for multivariate current status data with informative censoring. *Biometrics*, **58**(1), 79–88.
- Emoto, Sherrie E, Matthews, Peter C, *et al.* 1990. A Weibull model for dependent censoring. *The Annals of Statistics*, **18**(4), 1556–1577.
- Farewell, Vern T. 1982. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics*, **38**(4), 1041–1046.
- Farewell, Vernon T. 1977. A model for a binary variable with time-censored observations. *Biometrika*, **64**(1), 43–46.
- Fisher, Lloyd, & Kanarek, Paula. 1974. Presenting censored survival data when censoring and survival times may not be independent. *Reliability and Biometry*, 303–326.
- Fisher, Ronald Aylmer, & Tippett, Leonard Henry Caleb. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Pages 180–190 of: Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24. Cambridge Univ Press.
- Frederick, L, Page, David L, Fleming, Irvin D, Fritz, April G, Balch, Charles M, Haller, Daniel G, Morrow, Monica, *et al.* 2002. *AJCC cancer staging manual*. Vol. 1. Springer Science & Business Media.
- Freitas, Luiz Antonio, & Rodrigues, Josemar. 2013. Standard exponential cure rate model with informative censoring. *Communications in Statistics-Simulation and Computation*, **42**(1), 8–23.
- Friedman, Michael. 1982. Piecewise exponential models for survival data with covariates. *The Annals of Statistics*, **10**(1), 101–113.

- Gamerman, Dani. 1991. Dynamic Bayesian models for survival data. *Applied Statistics*, **40**(1), 63–79.
- Gamerman, Dani. 1994. Bayes estimation of the piece-wise exponential distribution. *IEEE Transactions on Reliability*, **43**(1), 128–131.
- Gamerman, Dani, & Lopes, Hedibert F. 2006. *Markov chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman & Hall/CRC.
- Gelfand, Alan E, & Smith, Adrian FM. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, **85**(410), 398–409.
- Gelfand, Alan E, Dey, Dipak K, & Chang, Hong. 1992. *Model determination using predictive distributions with implementation via sampling-based methods*. Tech. rept. Department of Statistics, Stanford University, California.
- Gelman, Andrew, Carlin, John B, Stern, Hal S, & Rubin, Donald B. 2014a. *Bayesian data analysis*. 2nd edn. Chapman & Hall/CRC.
- Gelman, Andrew, Hwang, Jessica, & Vehtari, Aki. 2014b. Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, **24**(6), 997–1016.
- Geman, Stuart, & Geman, Donald. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, 721–741.
- Ghosh, Pulak, & Gönen, Mithat. 2008. Bayesian modeling of multivariate average bioequivalence. *Statistics in Medicine*, **27**(13), 2402–2419.
- Gilks, Wally R, Best, NG, & Tan, KKC. 1995. Adaptive rejection Metropolis sampling within Gibbs sampling. *Applied Statistics*, **44**(4), 455–472.
- Gonzales, Jhon Franky Bernedo, Tomazella, Vera Lucia Damasceno, & Taconelli, Joao Paulo. 2013. Estimaco paramétrica do modelo de mistura com fragilidade gama na presena de covariáveis. *Revista Brasileira de Biometria*, **31**(2), 233–247.
- Hanagal, David D. 2011. *Modeling survival data using frailty models*. Chapman & Hall/CRC Boca Raton.
- Hastings, W Keith. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.

- Hougaard, Philip. 1984. Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, **71**(1), 75–83.
- Hougaard, Philip. 1986. Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, **73**(2), 387–396.
- Huang, Xuelin, & Wolfe, Robert A. 2002. A frailty model for informative censoring. *Biometrics*, **58**(3), 510–520.
- Huang, Xuelin, & Zhang, Nan. 2008. Regression survival analysis with an assumed copula for dependent censoring: a sensitivity analysis approach. *Biometrics*, **64**(4), 1090–1099.
- Huang, Xuelin, Wolfe, Robert A, & Hu, Chengcheng. 2004. A test for informative censoring in clustered survival data. *Statistics in Medicine*, **23**(13), 2089–2107.
- Ibrahim, Joseph G, Chen, Ming-Hui, & Sinha, Debajyoti. 2001. Bayesian semiparametric models for survival data with a cure fraction. *Biometrics*, **57**(2), 383–388.
- Ibrahim, Joseph G, Chen, Ming-Hui, & Sinha, Debajyoti. 2010. *Bayesian survival analysis*. Springer Series in Statistics.
- Jiang, Hongyu, Fine, Jason P, Kosorok, Michael R, & Chappell, Rick. 2005. Pseudo self-consistent estimation of a copula model with informative censoring. *Scandinavian Journal of Statistics*, **32**(1), 1–20.
- Johnson, Carol Hahn. 2004. *The SEER Program Coding and Staging Manual 2004*. Cancer Statistics Branch, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, National Institutes of Health, Public Health Service, US Department of Health and Human Services.
- Kalbfleisch, John D. 1978. Non-parametric Bayesian analysis of survival time data. *Journal of the Royal Statistical Society*, **B40**(2), 214–221.
- Kalbfleisch, John D, & Prentice, Ross L. 1973. Marginal likelihoods based on Cox’s regression and life model. *Biometrika*, **60**(2), 267–278.
- Kalbfleisch, John D, & Prentice, Ross L. 2011. *The statistical analysis of failure time data*. 2nd edn. John Wiley & Sons.
- Kaplan, Edward L, & Meier, Paul. 1958. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, **53**(282), 457–481.

- Kim, Jee Soo, & Proschan, Frank. 1991. Piecewise exponential estimator of the survivor function. *IEEE Transactions on Reliability*, **40**(2), 134–139.
- Kim, Sungduk, Chen, Ming-Hui, Dey, Dipak K, & Gamerman, Dani. 2007. Bayesian dynamic models for survival data with a cure fraction. *Lifetime Data Analysis*, **13**(1), 17–35.
- Klein, John P, Van Houwelingen, Hans C, Ibrahim, Joseph G, & Scheike, Thomas H. 2016. *Handbook of survival analysis*. Chapman & Hall/CRC.
- Kuk, Anthony YC, & Chen, Chen-Hsin. 1992. A mixture model combining logistic regression with proportional hazards regression. *Biometrika*, **79**(3), 531–541.
- Lachiewicz, Anne M, Berwick, Marianne, Wiggins, Charles L, & Thomas, Nancy E. 2008. Survival differences between patients with scalp or neck melanoma and those with melanoma of other sites in the Surveillance, Epidemiology, and End Results (SEER) program. *Archives of Dermatology*, **144**(4), 515–521.
- Lagakos, Stephan W, & Williams, JS. 1978. Models for censored survival analysis: A cone class of variable-sum models. *Biometrika*, **65**(1), 181–189.
- Lagakos, SW. 1979. General right censoring and its impact on the analysis of survival data. *Biometrics*, 139–156.
- Lawless, Jerald F. 2011. *Statistical models and methods for lifetime data*. 2nd edn. John Wiley & Sons.
- Leung, Kwan-Moon, Elashoff, Robert M, & Affi, Abdelmonem A. 1997. Censoring issues in survival analysis. *Annual Review of Public Health*, **18**(1), 83–104.
- Li, Chin-Shang, Taylor, Jeremy MG, & Sy, Judy P. 2001. Identifiability of cure models. *Statistics & Probability Letters*, **54**(4), 389–395.
- Li, Yi, Tiwari, Ram C, & Guha, Subharup. 2007. Mixture cure survival models with dependent censoring. *Journal of the Royal Statistical Society*, **B69**(3), 285–306.
- Link, William A. 1989. A model for informative censoring. *Journal of the American Statistical Association*, **84**(407), 749–752.
- Liu, Lei, Wolfe, Robert A, & Huang, Xuelin. 2004. Shared frailty models for recurrent events and a terminal event. *Biometrics*, **60**(3), 747–756.

- Louis, Thomas A. 1982. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society*, **B44**(2), 226–233.
- Maller, Ross, & Zhou, Xian. 1996. *Survival analysis with long-term survivors*. Wiley New York.
- Maltz, Michael D, & McCleary, Richard. 1977. The mathematics of behavioral change recidivism and construct validity. *Evaluation Review*, **1**(3), 421–438.
- McHenry, Pamela M, Hole, David J, & MacKie, Rona M. 1992. Melanoma in people aged 65 and over in Scotland, 1979-89. *Bmj*, **304**(6829), 746–749.
- Meeker, William Q, & LuValle, Michael J. 1995. An accelerated life test model based on reliability kinetics. *Technometrics*, **37**(2), 133–146.
- Metropolis, Nicholas, Rosenbluth, Arianna W, Rosenbluth, Marshall N, Teller, Augusta H, & Teller, Edward. 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, **21**(6), 1087–1092.
- Mostafa, Ayman A, & Ghorbal, Anis Ben. 2011. Using WinBUGS to Cox model with changing from the baseline hazard function. *Applied Mathematical Sciences*, **5**(45), 2217–2240.
- Murphy, Susan A. 1994. Consistency in a proportional hazards model incorporating a random effect. *The Annals of Statistics*, **22**(2), 712–731.
- Murphy, Susan A. 1995. Asymptotic theory for the frailty model. *The Annals of Statistics*, **23**(1), 182–198.
- Murthy, DN Prabhakar, Xie, Min, & Jiang, Renyan. 2004. *Weibull models*. John Wiley & Sons.
- Naskar, Malay, Das, Kalyan, & Ibrahim, Joseph G. 2005. A semiparametric mixture model for analyzing clustered competing risks data. *Biometrics*, **61**(3), 729–737.
- Nelson, Wayne B. 2005. *Applied life data analysis*. John Wiley & Sons.
- Nielsen, Gert G, Gill, Richard D, Andersen, Per Kragh, & Sørensen, Thorkild IA. 1992. A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian journal of Statistics*, **19**(1), 25–43.

- Othus, Megan, Li, Yi, & Tiwari, Ram C. 2009. A class of semiparametric mixture cure survival models with dependent censoring. *Journal of the American Statistical Association*, **104**(487), 1241–1250.
- Parner, Erik, *et al.* 1998. Asymptotic theory for the correlated gamma-frailty model. *The Annals of Statistics*, **26**(1), 183–214.
- Partanen, Juha. 1969. On waiting time distributions. *Acta Sociologica*, **12**(3), 132–143.
- Petris, Giovanni. 2015. dlm: an R package for bayesian analysis of dynamic linear models. *R package version 1.1-4*.
- Plummer, Martyn. 2015. JAGS version 3.3. 0 user manual. *JAGS Version 3.3.0*.
- Qiou, Zuqiang, Ravishanker, Nalini, & Dey, Dipak K. 1999. Multivariate survival analysis with positive stable frailties. *Biometrics*, **55**(2), 637–644.
- R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rodrigues, Josemar, Cancho, Vicente G, de Castro, Mário, & Louzada-Neto, Francisco. 2009. On the unification of long-term survival models. *Statistics & Probability Letters*, **79**(6), 753–759.
- Santos Junior, Paulo Cerqueira dos. 2012. *Análise de sobrevivência na presença de censura informativa*. Ph.D. thesis, Universidade Federal de Minas Gerais, <http://hdl.handle.net/1843/ICED-8TQGD8>.
- Scharfstein, Daniel O, & Robins, James M. 2002. Estimation of the failure time distribution in the presence of informative censoring. *Biometrika*, **89**(3), 617–634.
- Schwarz, Maik, Jongbloed, Geurt, & Van Keilegom, Ingrid. 2013. On the identifiability of copulas in bivariate competing risks models. *Canadian Journal of Statistics*, **41**(2), 291–303.
- Seiffert, JE, Shambaugh, EM, Weiss, MA, *et al.* 1993. SEER program: comparative staging guide for cancer. *Bethesda, MD: Cancer Statistics Branch Surveillance Program, Division of Cancer Prevention and Control (DCPC). National Cancer Institute (US) Bethesda MD*.
- Shih, Joanna H, & Louis, Thomas A. 1995. Assessing gamma frailty models for clustered failure time data. *Lifetime Data Analysis*, **1**(2), 205–220.

- Shou, Haochang, Hsu, Jesse Y, Xie, Dawei, Yang, Wei, Roy, Jason, Anderson, Amanda H, Landis, J Richard, Feldman, Harold I, Parsa, Afshin, Jepson, Christopher, *et al.* 2017. Analytic Considerations for Repeated Measures of eGFR in Cohort Studies of CKD. *Clinical Journal of the American Society of Nephrology*, **12**, 1–9.
- Siannis, Fotios. 2004. Applications of a parametric model for informative censoring. *Biometrics*, **60**(3), 704–714.
- Siannis, Fotios. 2011. Sensitivity analysis for multiple right censoring processes: Investigating mortality in psoriatic arthritis. *Statistics in Medicine*, **30**(4), 356–367.
- Siannis, Fotios, Copas, John, & Lu, Guobing. 2005. Sensitivity analysis for informative censoring in parametric survival models. *Biostatistics*, **6**(1), 77–91.
- Sinha, Debajyoti, Chen, Ming-Hui, & Ibrahim, Joseph G. 2003. Bayesian Inference for Survival Data with a Surviving Fraction. *Lecture Notes-Monograph Series*, **43**, 117–138.
- Soetaert, Karline. 2016. rootSolve: Nonlinear root finding, equilibrium and steady-state analysis of ordinary differential equations. *R package version 1.7*.
- Souza, Renata Camila de. 2015. *Análise de sobrevivência na presença de censura informativa: uma abordagem bayesiana*. Ph.D. thesis, Universidade Federal de Minas Gerais, <http://hdl.handle.net/1843/BUOS-A3ZFFU>.
- Spiegelhalter, David J, Best, Nicola G, Carlin, Bradley P, & Van Der Linde, Angelika. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, **B64**(4), 583–639.
- Staplin, Natalie. 2012. *Informative censoring in transplantation statistics*. Ph.D. thesis, University of Southampton, <https://eprints.soton.ac.uk/341789>.
- Su, Yu-Sung, & Yajima, Masanao. 2012. R2jags: A Package for Running jags from R. *R package version 0.5-7*.
- Suzuki, Adriano Kamimura. 2012. *Modelos de sobrevivência bivariados baseados na cópula FGM: uma abordagem bayesiana*. Ph.D. thesis, Universidade Federal de São Carlos, <https://repositorio.ufscar.br/bitstream/handle/ufscar/4487/4292.pdf>.
- Taylor, Jeremy MG. 1995. Semi-parametric estimation in failure time mixture models. *Biometrics*, **51**(3), 899–907.
- Vaida, Florin, Xu, Ronghui, *et al.* 2000. Proportional hazards model with random effects. *Statistics in Medicine*, **19**(24), 3309–3324.

- Vaupel, James W, Manton, Kenneth G, & Stallard, Eric. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, **16**(3), 439–454.
- Watanabe, Sumio. 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *The Journal of Machine Learning Research*, **11**, 3571–3594.
- Wei, Greg CG, & Tanner, Martin A. 1990. A Monte Carlo implementation of the EM algorithm and the poor man’s data augmentation algorithms. *Journal of the American Statistical Association*, **85**(411), 699–704.
- Weibull, W. 1939. A statistical theory of the strength of materials. *Royal Swedish Institute for Engineering Research*, **151**, 45p.
- Wienke, Andreas. 2011. *Frailty models in survival analysis*. Chapman & Hall/CRC Biostatistics Series.
- Williams, JS, & Lagakos, SW. 1977. Models for censored survival analysis: constant-sum and variable-sum models. *Biometrika*, **64**(2), 215–224.
- Yakovlev, A Yu, & Tsodikov, Alexander D. 1996. *Stochastic models of tumor latency and their biostatistical applications*. Vol. 1. Series in Mathematical Biology and Medicine.
- Yakovlev, A Yu, Asselain, B, Bardou, VJ, Fourquet, A, Hoang, T, Rochefediere, A, & Tsodikov, AD. 1993. A simple stochastic model of tumor recurrence and its application to data on premenopausal breast cancer. *Biometrie et analyse de donnees spatio-temporelles*, **12**, 66–82.
- Yakovlev, Andrej Yu, Cantor, Alan B, & Shuster, Jonathan J. 1994. Parametric versus non-parametric methods for estimating cure rates based on censored survival data. *Statistics in Medicine*, **13**(9), 983–986.
- Yin, Guosheng. 2005. Bayesian cure rate frailty models with application to a root canal therapy study. *Biometrics*, **61**(2), 552–558.
- Yin, Guosheng, & Ibrahim, Joseph G. 2005. A general class of Bayesian survival models with zero and nonzero cure fractions. *Biometrics*, **61**(2), 403–412.
- Zhang, Haimeng, & Rao, M Bhaskara. 2004. On generalized maximum likelihood estimation in the proportional hazards model with partially informative censoring. *Metrika*, **59**(2), 125–136.

Zhang, Zhigang, Sun, Liuquan, Sun, Jianguo, & Finkelstein, Dianne M. 2007. Regression analysis of failure time data with informative interval censoring. *Statistics in Medicine*, **26**(12), 2533–2546.