
Modelagem conjunta de dados longitudinais e de sobrevivência para avaliação de desfechos clínicos do parto

Alexandre Cristovão Maiorano

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
UFSCar-USP

Alexandre Cristovão Maiorano

Modelagem conjunta de dados longitudinais e de sobrevivência para avaliação de desfechos clínicos do parto

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutor em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.

Orientador: Prof. Dr. Francisco Louzada Neto

São Carlos
Fevereiro de 2019

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
UFSCar-USP

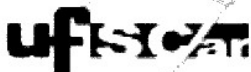
Alexandre Cristovão Maiorano

Joint modeling of longitudinal and survival data to evaluate
clinical outcomes of labor

Thesis submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DEs-UFSCar – in accordance with the requirements of the Statistics Interagency Graduate Program, for the degree of Doctor in Statistics.

Advisor: Prof. Dr. Francisco Louzada Neto

São Carlos
February 2019

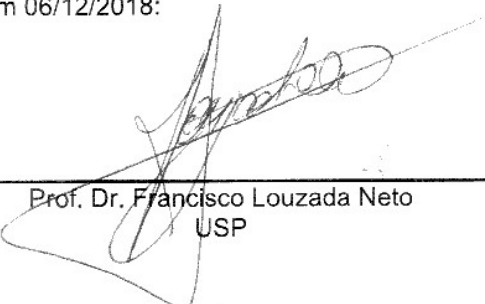


UNIVERSIDADE FEDERAL DE SÃO CARLOS

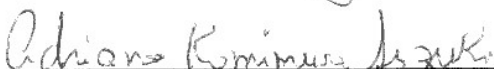
Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

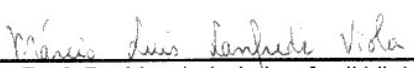
Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Alexandre Cristovão Maiorano, realizada em 06/12/2018:



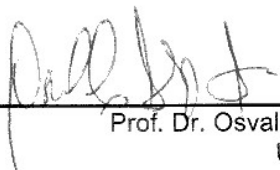
Prof. Dr. Francisco Louzada Neto
USP



Prof. Dr. Adriano Kaminura Suzuki
USP



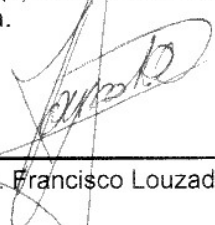
Prof. Dr. Marcio Luis Lanfredi Viola
UFSCar



Prof. Dr. Osvaldo Anacleto Junior
USP

Prof. Dr. Anderson Luiz Ara Souza
UFBA

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Anderson Luiz Ara Souza e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ão) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.



Prof. Dr. Francisco Louzada Neto

*Este trabalho é dedicado à minha família,
minha noiva Bárbara e aos meus amigos.
Em especial, aos alunos de doutorado
que passaram por essa etapa.*

AGRADECIMENTOS

Agradeço primeiramente a Deus pela trilha que segui e por ter tido fé, esperança e resiliência para ter superado este desafio.

Agradeço aos meus pais pela educação que me deram e ao apoio durante os anos de estudo. Especialmente minha mãe, que sempre me motivou a fazer o mestrado e, posteriormente o doutorado. À minha irmã, pelo companheirismo e amizade, bem como as discussões de pesquisa e dificuldade que passamos juntos durante os estudos.

Agradeço à minha noiva Bárbara, que desde quando passou no vestibular em Estatística está ao meu lado, compartilhando momentos inesquecíveis. Por ter me apoiado nos momentos difíceis e me motivado quando pensei em desistir. Pela compreensão aos finais de semana e feriados, em que eu precisava me dedicar aos estudos ao invés de desfrutar momentos de lazer.

Agradeço aos meus avós, que sempre se preocupavam e perguntavam como estavam os estudos. Aos meus colegas de doutorado pela união durante as matérias, compartilhando as dificuldades em momentos que jamais serão esquecidos.

A todos os professores que participaram de minha formação. Particularmente ao professor Louzada, que me proporcionou diversas experiências, fazendo com que minha formação fosse sólida e diversificada, melhorando ainda mais meu desenvolvimento. Agradeço também à professora Gleici pela orientação, experiência de trabalhar com os dados de parto e por ter me ajudado sempre que precisei, dando conselhos e revisando meu trabalho.

Agradeço ao Anderson Ara pela amizade que construímos, pelo conhecimento e confiança que tem me passado desde sempre. Também agradeço ao Pedro Ramos pelas discussões e sugestões feitas enquanto eu desenvolvia o estudo.

Agradeço à minha banca de defesa pela participação, contribuições e sugestões, melhorando ainda mais o trabalho realizado e, por fim, ao Itaú Unibanco pelas horas de estudo para que eu terminasse a pesquisa. Assim como os meus colegas de trabalho pela motivação e inspiração para finalizar a minha Tese.

RESUMO

MAIORANO, ALEXANDRE CRISTOVÃO. **Modelagem conjunta de dados longitudinais e de sobrevivência para avaliação de desfechos clínicos do parto**. 2019. 110 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Pelo fato da maioria das mortes e morbidades associadas à gravidez ocorrerem em torno do parto, a qualidade do cuidado nesse período é crucial para as mães e seus bebês. Para acompanhar as mulheres nessa etapa, o partograma tem sido a ferramenta mais utilizada nas últimas décadas e, devido à sua simplicidade, é frequentemente usado em países com baixa e média renda. No entanto, sua utilização é altamente questionada devido à ausência de evidências que justifiquem uma contribuição ao parto. Para melhorar a qualidade do parto nessas circunstâncias, o projeto BOLD tem sido desenvolvido com o intuito de reduzir a ocorrência de problemas indesejados e com a finalidade desenvolver uma ferramenta moderna, chamada de SELMA, que projeta-se como uma alternativa ao partograma. Com a finalidade de associar características fixas e dinâmicas avaliadas no parto e identificar quais elementos intra parto podem ser utilizados como gatilhos para realização de uma intervenção e, assim, prevenir um desfecho indesejado, propomos nesta tese a utilização de modelos de sobrevivência com covariáveis dependentes do tempo. Inicialmente, consideramos a modelagem de dados longitudinais e de sobrevivência utilizando funções de risco paramétricas flexíveis. Nesse caso, propomos a utilização de cinco generalizações da distribuição Weibull, da distribuição Nagakami e utilizamos um procedimento geral de seleção de modelos paramétricos usuais via distribuição Gamma generalizada, inédito na modelagem conjunta. Realizamos um extenso estudo de simulação para avaliar as estimativas de máxima verossimilhança e os critérios de discriminação. Além disso, a própria natureza do parto nos leva a um contexto de eventos múltiplos, nos remetendo à utilização dos modelos multiestados. Eles são definidos como modelos para um processo estocástico que a qualquer momento ocupa um conjunto discreto de estados. De uma forma geral, são os modelos mais comuns para descrever o desenvolvimento de dados de tempo de falha longitudinais e são frequentemente utilizados em aplicações médicas. Considerando o contexto de eventos múltiplos, propomos a inclusão de uma covariável dependente do tempo no modelo multiestados a partir de uma modificação dos dados, o que nos trouxe resultados satisfatórios e similares ao esperado na prática clínica.

Palavras-chave: covariáveis dependentes do tempo, função de risco flexível, modelo conjunto, modelo multiestados, modelagem do trabalho de parto.

ABSTRACT

MAIORANO, ALEXANDRE CRISTOVÃO. **Joint modeling of longitudinal and survival data to evaluate clinical outcomes of labor**. 2019. 110 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

As most pregnancy-related deaths and morbidities are clustered around the time of child birth, the quality of care during this period is crucial for mothers and their babies. To monitor the women at this stage, the partograph has been the central tool used in recent decades and, motivated by its simplicity, is frequently used in low-and middle-income countries. However, its use is highly questioned due to lack of evidence to justify a contribution to labor. To improve the quality of labor in these circumstances, the BOLD project has been developed in order to reduce the occurrence of pregnancy-related problems and in order to develop a modern tool, called SELMA, which is projected as an alternative to partograph. Aiming to associate fixed and dynamic characteristics evaluated in the delivery and to identify which elements can be used as triggers for performing an intervention, and thus preventing a bad outcome, this thesis proposes the use of survival models with time dependent covariates. Initially, we consider the joint modeling of survival and longitudinal data using flexible parametric hazard functions. In this sense, we propose the use of five generalizations of Weibull distribution, the Nagakami model and an inedited framework to discriminate usual parametric models via the generalized Gamma distribution, performing an extensive simulation study to evaluate the maximum likelihood estimations and the proposed discrimination criteria. Indeed, by its own nature, the birth leads us to a context of multiple events, referring to the use of multi-state models. These are models for a stochastic process which at any time occupies one of a few possible states. In general, they are the most common models to describe the development of longitudinal failure time data and are often used in medical applications. Considering this context, we proposed the inclusion of a time dependent covariate in the multi-state model using a modified version of the input data, which gave us satisfactory results similar to those expected in clinical practice.

Keywords: flexible hazard function, joint multi-state model, labor process, , multi-state model, time dependent covariates.

LISTA DE ILUSTRAÇÕES

Figura 1 – Ilustração do partograma.	20
Figura 2 – Dependência latente entre o processo de sobrevivência e longitudinal.	32
Figura 3 – Funções de densidade obtidas para os parâmetros da distribuição GG considerando a parametrização de Stacy.	53
Figura 4 – Funções de densidade obtidas para os parâmetros da distribuição GG considerando a parametrização de Prentice.	55
Figura 5 – Funções de densidade obtidas para os parâmetros da distribuição EW.	56
Figura 6 – Funções de densidade obtidas para os parâmetros da distribuição GW.	58
Figura 7 – Funções de densidade obtidas para os parâmetros da distribuição MOW.	60
Figura 8 – Funções de densidade obtidas para os parâmetros da distribuição GPW.	62
Figura 9 – Funções de densidade obtidas para os parâmetros da distribuição NK.	64
Figura 10 – Predições dinâmicas da probabilidade de sobrevivência para o modelo conjunto com função de risco GG considerando uma paciente sem dilatação na admissão, nulípara, com comorbidade ou complicação pré gravidez, não referenciada de outro hospital e com membrana aminótica intacta.	71
Figura 11 – Predições dinâmicas da probabilidade de sobrevivência para o modelo conjunto com função de risco log-normal considerando uma paciente sem dilatação na admissão, nulípara, com comorbidade ou complicação pré gravidez, não referenciada de outro hospital e com membrana aminótica intacta.	73
Figura 12 – Predições dinâmicas da probabilidade de sobrevivência para o modelo conjunto com função de risco NK considerando uma paciente sem dilatação na admissão, nulípara, com comorbidade ou complicação pré gravidez, não referenciada de outro hospital e com membrana aminótica intacta.	75
Figura 13 – Exemplo de modelo unidirecional.	79
Figura 14 – Exemplo de modelo progressivo.	79
Figura 15 – Exemplo de modelo bidirecional. Mudanças entre os estados 2 e 3 são possíveis.	80
Figura 16 – Exemplo de modelo recorrente.	80
Figura 17 – Modelo multiestados com 4 estados e 5 transições.	91
Figura 18 – Razão de risco para o modelo multiestados com efeitos fixos.	94
Figura 19 – Razão de risco para o modelo multiestados com covariável dependente do tempo.	96

LISTA DE TABELAS

Tabela 1 – Resultados da simulação para a distribuição GG considerando a parametrização de Stacy. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	51
Tabela 2 – Resultados da simulação para a distribuição GG considerando a parametrização de Prentice. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	52
Tabela 3 – Resultados da simulação para a distribuição EW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	54
Tabela 4 – Resultados da simulação para a distribuição GW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	57
Tabela 5 – Resultados da simulação para a distribuição MOW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	59
Tabela 6 – Resultados da simulação para a distribuição GPW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	61
Tabela 7 – Resultados da simulação para a distribuição NK. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).	63
Tabela 8 – Taxas de rejeição da hipótese nula ao nível de 5% de significância considerando o teste de razão de verossimilhanças via família GG.	66
Tabela 9 – Taxas de seleção dos modelos avaliados de acordo com o AIC.	66
Tabela 10 – Caracterização geral da amostra de acordo com as principais variáveis estudadas.	69
Tabela 11 – Resultados de log-verossimilhança($\log L(\cdot)$), AIC e BIC para todas as generalizações propostas da distribuição Weibull.	70
Tabela 12 – Estimativas de máxima verossimilhança, erro padrão e intervalo de confiança de 95% para o modelo conjunto com função de risco GG.	71
Tabela 13 – Resultados para discriminação de modelos via família GG.	72

Tabela 14 – Estimativas de máxima verossimilhança, erro padrão e intervalo de confiança de 95% para o modelo conjunto com função de risco log-normal.	72
Tabela 15 – Estimativas de máxima verossimilhança, erro padrão e intervalo de confiança de 95% para o modelo conjunto com função de risco NK.	74
Tabela 16 – Resultados de log-verossimilhança, AIC e BIC para diferentes funções paramétricas.	74
Tabela 17 – Exemplo de um banco de dados no formato “long” para multiestados.	83
Tabela 18 – Exemplo de Banco de dados com covariável dependente do tempo para o contexto usual de análise de sobrevivência.	89
Tabela 19 – Banco de dados no formato “long” e com a covariável dilatação como dependente do tempo para o modelo multiestados.	91
Tabela 20 – Frequência total para as transições do modelo multiestados.	91
Tabela 21 – Coeficientes estimados para o modelo multiestados com efeitos fixos.	93
Tabela 22 – Características dos pacientes A, B, C e D.	93
Tabela 23 – Probabilidades de transição (%) em $t = 6, 12, 18, 24$ e $30h$ considerando o estado inicial admissão e tempo igual a $t = 0$	94
Tabela 24 – Coeficientes estimados para o modelo multiestados com covariável dependente do tempo.	95

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Objetivos	24
1.1.1	<i>Geral</i>	24
1.1.2	<i>Específicos</i>	24
1.2	Estrutura da Tese	25
2	MODELAGEM DE DADOS LONGITUDINAIS E DE SOBREVIVÊNCIA	27
2.1	Análise do Tempo até o Evento de Interesse	27
2.2	Inclusão de Covariáveis Dependentes do Tempo	30
2.3	O Modelo Conjunto	32
2.4	Função de Verossimilhança Conjunta	33
2.5	Aproximação da Função de Verossimilhança Conjunta	35
2.6	Testes de Hipótese e Intervalo de Confiança para os Parâmetros	38
2.7	Predição Dinâmica da Probabilidade de Sobrevivência	40
2.8	Comentários Finais	41
3	MODELO CONJUNTO PARAMÉTRICO	43
3.1	Distribuição Gama Generalizada	44
3.2	Distribuição Weibull Exponenciada	45
3.3	Distribuição Weibull Generalizada	45
3.4	Distribuição Marshall-Olkin-Weibull	46
3.5	Distribuição Potência Weibull Generalizada	46
3.6	Distribuição Nakagami-m	46
3.7	Estudo de Simulação	47
3.7.1	<i>Especificações do Estudo</i>	47
3.7.2	<i>Estudo de Simulação para as Distribuições Propostas</i>	50
3.7.2.1	<i>Distribuição Gama Generalizada</i>	50
3.7.2.2	<i>Distribuição Weibull Exponenciada</i>	50
3.7.2.3	<i>Distribuição Weibull Generalizada</i>	55
3.7.2.4	<i>Distribuição Marshall-Olkin-Weibull</i>	56
3.7.2.5	<i>Distribuição Potência Weibull Generalizada</i>	58
3.7.2.6	<i>Distribuição Nakagami-m</i>	60

3.7.3	<i>Cr�terios para Sele�o de Modelos Param�tricos</i>	64
3.8	Modelo Conjunto no Contexto do Trabalho de Parto	67
3.8.1	<i>Modelagem a partir de Generaliza�es da Distribui�o Weibull</i>	68
3.8.2	<i>Sele�o de Modelos via Fam�lia Gama Generalizada</i>	70
3.8.3	<i>Modelo Nakagami-m</i>	73
3.9	Coment�rios Finais	75
4	MODELOS MULTIESTADOS	77
4.1	Processo Multiestados	77
4.2	Estrutura dos Modelos Multiestados	78
4.2.1	<i>Modelos Unidirecionais</i>	79
4.2.2	<i>Modelos Progressivos</i>	79
4.2.3	<i>Modelos Bidirecionais</i>	80
4.2.4	<i>Modelos Recorrentes</i>	80
4.3	Modelos de Regress�o Multiestados	80
4.4	Procedimentos de Estimac�o	81
4.4.1	<i>Intensidades de Transi�o: Modelos N�o Param�tricos</i>	83
4.4.2	<i>Intensidades de Transi�o: Modelos Semi Param�tricos</i>	84
4.4.3	<i>Estimac�o das Probabilidades de Transi�o e Erro Padr�o</i>	86
4.5	Modelo de Regress�o Multiestados - Covari�vel Dependente do Tempo	87
4.6	Modelo Multiestados no Contexto do Trabalho de Parto	89
4.6.1	<i>Modelo Multiestados com Efeitos Fixos</i>	90
4.6.2	<i>Modelo Multiestados com Covari�vel Dependente do Tempo</i>	92
4.7	Coment�rios Finais	97
5	CONSIDERA�ES FINAIS	99
5.1	Sugest�es para Trabalhos Futuros	104
	REFER�NCIAS	105

INTRODUÇÃO

Devido ao fato da maioria das mortes e morbidades associadas à gravidez ocorrerem em qualquer instante em torno do parto, a qualidade do cuidado durante esse período é crucial para as mães e seus bebês. Em um estudo recente, [Lawn et al. \(2014\)](#) mostram que entre o parto e a primeira semana pós parto ocorrem cerca de 73% de mortes de recém nascidos, 61% de mortes maternas e 44% de natimortos. Para fazer o monitoramento do parto e evitar possíveis problemas, tem-se frequentemente utilizado nas últimas décadas o partograma ([SOUZA et al., 2015](#); [OLADAPO et al., 2018](#)), ilustrado na Figura 1. Essa é uma ferramenta gráfica que considera a duração do trabalho de parto e a evolução da dilatação, sendo que suas principais características são as linhas de alerta e ação, que buscam funcionar como gatilhos para intervenções. Devido à sua simplicidade, o partograma é muito utilizado em países com baixa e média renda, visto que nesses locais as mulheres costumam ter seu parto sozinhas ou então com parteiras. Entretanto, sua utilização é muito questionada devido à ausência de evidências que realmente justifiquem sua contribuição no parto. Dentre os diversos pontos de discussão, temos as características de sua criação que, por exemplo, consideram a evolução da dilatação como de 1 cm/hora e a não inclusão de outras características da mulher nessa evolução ([SOUZA et al., 2015](#)).

Para melhorar a qualidade do parto, principalmente em países com poucos recursos, [Oladapo et al. \(2015\)](#), em parceria com a Organização Mundial da Saúde (OMS), desenvolvem atualmente o projeto *Better Outcomes in Labours Difficulty* (BOLD). O principal objetivo desse projeto é reduzir a mortalidade e morbidade materna e neonatal, abordando as barreiras críticas para o processo de cuidados de boa qualidade intra parto e reforçar a ligação entre os sistemas de saúde e comunidade. Além disso, os autores buscam desenvolver uma ferramenta moderna, chamada de *Simplified, Effective, Labour Monitoring-to-Action* (SELMA), a qual projeta-se como uma alternativa ao partograma. O projeto propôs a pesquisa em unidades de saúde da Nigéria e Uganda, sendo que foram coletadas informações relacionadas às características da mãe no momento da admissão na unidade de saúde, características do parto e os desfechos da gravidez.

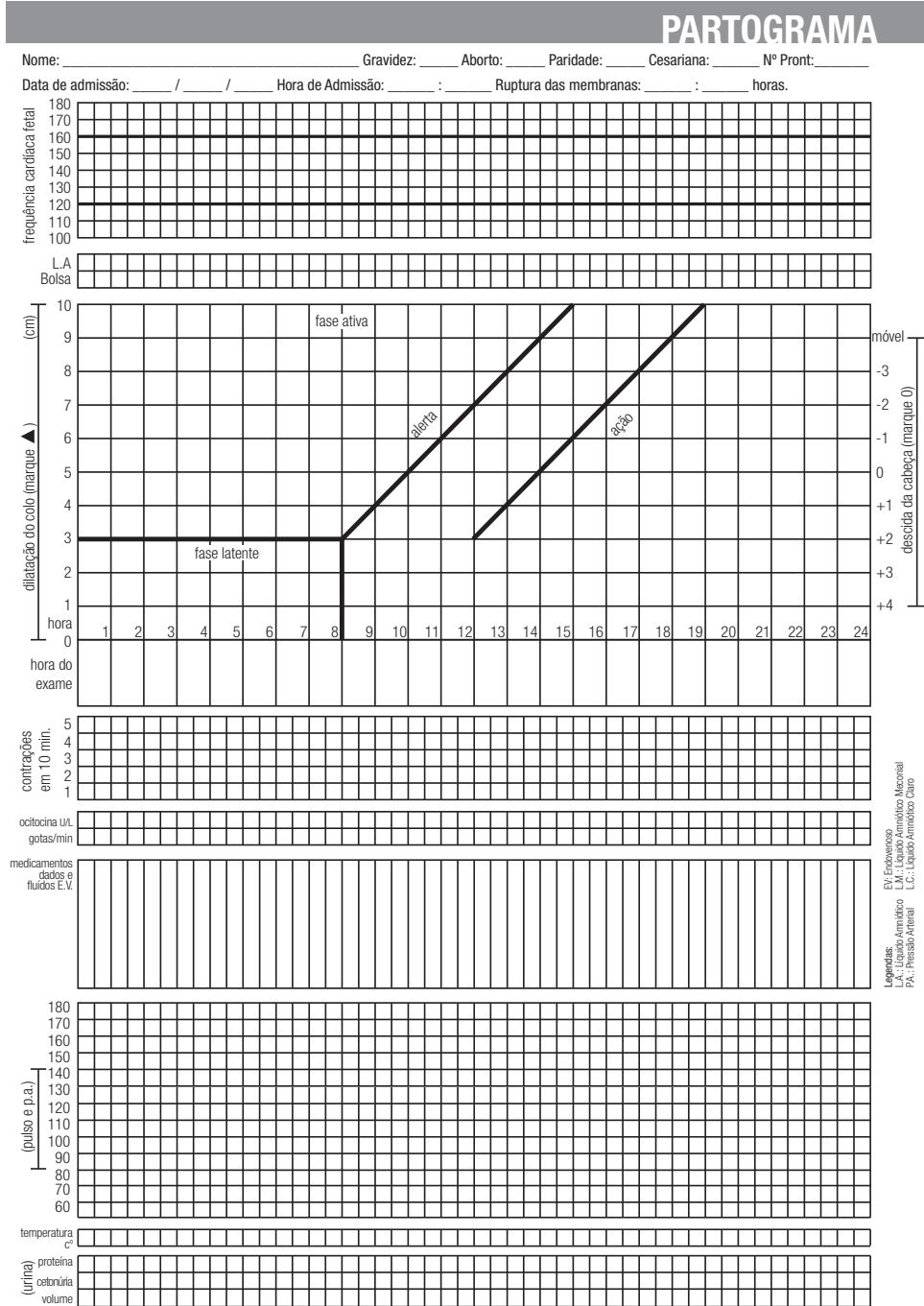


Figura 1 – Ilustração do partograma.

Fonte: SulAmérica (2016).

Como instrumento de avaliação, os autores consideraram informações estáticas, ou seja, que não se alteram durante o parto, e informações dinâmicas, que normalmente são incluídas como parte do partograma. Como exemplo de variáveis estáticas podemos citar características demográficas e do passado obstétrico da mãe, enquanto que, para as dinâmicas, temos frequência cardíaca fetal, dilatação cervical, contrações uterinas, uso de ocitocina e temperatura, por exemplo. Além disso, os autores também avaliaram informações dinâmicas que não estão associadas ao partograma como prestação de apoio social, posição e bem estar materno durante o parto. Por fim, avaliaram os dados relacionados aos desfechos da mãe e do bebê, tais como nascimento por cesárea ou parto normal, encaminhamento do bebê à UTI neonatal, morte da mãe ou do bebê, peso e sexo do bebê, etc.

Com a finalidade de associar essas informações e identificar quais elementos intra parto podem ser utilizados como gatilhos para realização de intervenção e, assim, prevenir um desfecho indesejado, podemos utilizar métodos estatísticos e verificar a possível associação desses desfechos com variáveis de interesse. Portanto, nesta Tese existe o interesse em responder diversas perguntas sobre o trabalho de parto, tais como:

- Verificar quais características maternas e da evolução do parto da mulher estão associadas ao parto vaginal ou por cesárea;
- Prever o risco de parto por cesárea dado que a mulher foi referenciada por outra unidade de saúde;
- Avaliar se a presença de comorbidades ou complicações pré gravidez podem induzir um parto por cesárea;
- Entender como a paridade prévia está relacionada com o desfecho do parto, ou seja, dado que a mulher já teve filho, o prognóstico do parto se diferencia do parto da mulher que nunca teve filho;
- Analisar a existência de um prognóstico diferenciado caso a membrana aminótica esteja rompida;
- Identificar qual o melhor momento para realizar uma intervenção diante da evolução do parto da mulher;
- Verificar a possibilidade de uma intervenção acelerar o parto vaginal;
- Compreender como a evolução dinâmica da dilatação pode afetar o desenvolvimento do parto.

Neste sentido, quando estamos interessados em avaliar a associação entre covariáveis dependentes do tempo dinamicamente em um contexto de análise de sobrevivência, podemos utilizar os modelos conjuntos. A ideia geral desse método é combinar um modelo de sobrevivência

com um modelo de medidas repetidas adequado para a variável longitudinal (RIZOPOULOS *et al.*, 2010). Em suma, combinando os dois processos, podemos prever o risco de um evento futuro baseando-se na evolução da variável observada repetidamente no tempo e, devido a essa característica, esses modelos têm ganhado atenção nos últimos anos. Segundo uma revisão sistemática da literatura publicada recentemente por Sudell, Kolamunnage-Dona e Tudur-Smith (2016), foram identificados 65 estudos voltados à área médica, com aplicação, em maior parte, à dados de câncer e AIDS/HIV tendo como finalidade estudar a conexão entre as duas componentes do modelo.

Geralmente, no modelo conjunto considera-se uma função de risco paramétrica que segue uma distribuição usual como, por exemplo, a distribuição Weibull, mais utilizada no contexto médico (SUDELL; KOLAMUNNAGE-DONA; TUDUR-SMITH, 2016), ou então métodos não paramétricos, que se baseiam em *splines* (BROWN; IBRAHIM; DEGRUTTOLA, 2005; CROWTHER; ABRAMS; LAMBERT, 2012; PHAM; PHAM, 2018). Do ponto de vista paramétrico, quando consideramos especificamente a distribuição Weibull (WEIBULL, 1951), temos que esta distribuição apresenta propriedades matemáticas convenientes e sua aplicação se dá em diversas áreas. No entanto, ela não pode ser usada para descrever funções de risco não monótonas, que possuem forma de banheira ou unimodal, por exemplo, sendo que para solucionar esse problema diversas generalizações têm sido propostas na literatura (PHAM; LAI, 2007).

Dessa forma, com a finalidade de utilizar uma abordagem paramétrica flexível para a função de risco do modelo conjunto, propomos nesta Tese a utilização de generalizações da distribuição Weibull com três parâmetros, que são capazes de lidar com funções de risco com formas complexas, tais como constante, crescente, decrescente, forma de banheira e unimodal. Nesse caso, avaliamos o uso das distribuições gama generalizada, dada pelas parametrizações de Stacy (1962) e Prentice (1974), Weibull exponenciada (MUDHOLKAR; SRIVASTAVA, 1993), Weibull generalizada (MUDHOLKAR; SRIVASTAVA; KOLLIA, 1996), Marshall-Olkin-Weibull (MARSHALL; OLKIN, 1997) e distribuição de potência Weibull generalizada (NIKULIN; HAGHIGHI, 2007). Nesse mesmo contexto, também avaliamos a utilização da distribuição Nakagami- m , que é um poderoso modelo estatístico frequentemente utilizado para modelar sinais de rádio. Proposta por NAKAGAMI (1960), essa distribuição se diferencia das anteriores por ser mais parcimoniosa, uma vez que tem apenas dois parâmetros.

Levando-se em conta que a família gama generalizada é capaz de generalizar diversas funções paramétricas usuais como exponencial, Weibull, log-normal e distribuição gama, ela pode ser usada para selecionar uma dessas famílias por meio de um procedimento geral via testes de razão de verossimilhanças. Nesse caso, estudamos a viabilidade de considerar a distribuição gama generalizada para fazer discriminação de modelos utilizando o teste de razão de verossimilhanças e também por outros critérios capazes de fazer discriminação de modelos, como o critério de informação de Akaike (AIC) e critério de informação Bayesiano (BIC). Destacamos que

tanto a proposta de utilizar a família gama generalizada, como o estudo de simulação para avaliar o desempenho dos critérios de informação sob a modelagem de dados longitudinais e de sobrevivência são inéditos na literatura dos modelos conjuntos.

No modelo conjunto geralmente analisamos apenas um desfecho. No entanto, durante o trabalho de parto podemos observar diversas situações como a realização de uma intervenção ou até mesmo o nascimento do bebê por parto vaginal ou cesárea, nos levando a uma situação de eventos múltiplos. Dado que temos múltiplos eventos e seus tempos associados, podemos utilizar os modelos multiestados, que são definidos como modelos para um processo estocástico que a qualquer momento ocupa um conjunto finito de estados e, em geral, são os modelos mais comuns para descrever a evolução longitudinal do tempo de falha (HOUGAARD, 1999). Esses tipos de modelos são usados em diversas aplicações médicas, em que os estados do modelo representam os níveis ou estados de uma doença. Modelos com essas características são úteis para responder diversas questões no contexto de análise de sobrevivência que não podem ser respondidas pelos modelos usuais, sendo que, tipicamente, nosso principal interesse é estimar intensidades e probabilidades de transição entre os estados (WREEDE; FIOCCO; PUTTER, 2011).

Em geral, ao introduzirmos covariáveis no modelo multiestados, consideramos que seus valores são constantes. Entretanto, particularmente na medicina, e ainda mais no trabalho de parto, essa suposição é clinicamente fraca, sendo mais relevante considerar que o risco da ocorrência de um particular evento está associado a covariáveis dependentes do tempo. Assim, estendemos a aplicação do modelo multiestados para considerar covariáveis dependentes do tempo considerando o modelo de Cox estendido (RIZOPOULOS, 2012b). No melhor do nosso conhecimento, não existe na literatura trabalhos que estendam o modelo de Cox no contexto do modelo multiestados. Apesar de assumir risco constante entre os tempos em que a variável longitudinal é avaliada, pode ser considerado como uma alternativa menos complexa para buscar possíveis *insights* no contexto multiestados.

Portanto, além de avaliar o uso de funções de risco paramétrica flexíveis na modelagem conjunta, também propomos a utilização do modelo multiestados com covariável dependente do tempo aos dados do trabalho de parto. Neste caso, realizamos diversas aplicações dos modelos estudados ao contexto do parto, auxiliando na compreensão e melhora do trabalho de parto. Vale destacar que para fazer a estimação dos parâmetros dos modelos estudados e os estudos de simulação apresentados na Tese foi utilizado o *software* R (R Core Team, 2018).

1.1 Objetivos

1.1.1 Geral

Propor o uso de modelos flexíveis que englobem características longitudinais e temporais no contexto do trabalho de parto. A partir de distribuições paramétricas que incorporam diferentes formas da função de risco, e da variação longitudinal da dilatação cervical, buscamos estudar o tempo de parto de forma a contribuir com a melhora da qualidade e cuidado intraparto.

1.1.2 Específicos

1. Propor o uso de modelos flexíveis que englobem características longitudinais e temporais no contexto do trabalho de parto;
2. Realizar estudo de distribuições paramétricas consideradas em análise de sobrevivência que incorporem funções de risco flexíveis. Privilegiaremos generalizações da distribuição Weibull, bem como da distribuição Nakagami- m , de modo que seja possível incorporar funções de risco crescente, decrescente, em forma de banheira e unimodal;
3. Realizar estudos de simulação para avaliar a eficiência do método de estimação dos parâmetros dos modelos propostos;
4. Fazer estudo de simulação de propriedades estatísticas e de testes de hipóteses, como teste de razão de verossimilhanças e critérios de informação AIC e BIC;
5. Avaliação de critérios de seleção de modelos no contexto de análise de sobrevivência e dados longitudinais, para selecionar os modelos paramétricos usuais, tais como Weibull, log-normal e gama, de modo que seja possível selecionar a distribuição que apresente melhor ajuste aos dados;
6. Comparar os resultados dos modelos propostos com outros já existentes;
7. Utilizar os modelos e métodos de seleção propostos para compreender o trabalho de parto, analisando-se o tempo até o nascimento;
8. Estudo e proposição de um modelo multiestados para a problemática do parto, de modo que seja possível incluir a dilatação cervical como covariável dependente do tempo em um contexto de eventos múltiplos;
9. Utilizar o modelo multiestados proposto para identificar potenciais preditores e mensurar as probabilidades e intensidades de transição entre as intervenções e desfechos do parto.

1.2 Estrutura da Tese

Esta Tese está organizada da seguinte maneira: no Capítulo 2 introduzimos os conceitos fundamentais para a análise de sobrevivência, a inclusão de covariáveis dependentes do tempo no modelo usual e no modelo conjunto, sendo que, para este último, detalhamos o método de estimação via máxima verossimilhança com aproximação via quadratura gaussiana, testes de hipótese para os parâmetros, intervalos de confiança usuais e via *bootstrap*, e a predição dinâmica da probabilidade de sobrevivência. No Capítulo 3 apresentamos as distribuições paramétricas propostas para o contexto do trabalho de parto, um estudo de simulação e uma aplicação que visa modelar o tempo até a cesárea considerando a dilatação como dependente do tempo. No Capítulo 4 introduzimos a teoria dos modelos multiestados, apresentamos suas principais características e as formas possíveis para associar covariáveis fixas e dinâmicas neste modelo. Ainda neste capítulo, exibimos um modelo para o trabalho de parto que leva em conta eventos múltiplos, caracterizados pela realização de uma intervenção, parto vaginal ou cesárea. Por fim, o Capítulo 5 apresenta as considerações finais sobre a Tese, bem como as sugestões para trabalhos futuros.

MODELAGEM DE DADOS LONGITUDINAIS E DE SOBREVIVÊNCIA

Durante o trabalho de parto, geralmente as mulheres são acompanhadas desde o momento de admissão à unidade de saúde até o nascimento do bebê. Durante esse período existe um monitoramento periódico por parte da equipe médica, conforme citado na utilização do partograma, o que gera diversos tipos de informações, tais como: sobre o estado de saúde da mulher e do bebê, a posição fetal, a realização de uma intervenção como indução, uso de ocitocina ou analgesia epidural. Além disso, ao final do parto temos informações sobre a forma de nascimento do bebê, se foi por parto vaginal ou por cesárea, se o bebê nasceu em boas condições ou se foi para UTI neonatal, assim como estado de saúde da mãe, caracterizando-se toda a evolução do parto até o momento de nascimento.

Dado que foram registradas algumas informações relacionadas ao tempo de ocorrência como, por exemplo, a data de admissão, o momento de realização da intervenção e a data de nascimento, temos a possibilidade de avaliar o tempo até os possíveis eventos de interesse. Nesse caso, sugerimos a utilização de métodos de análise de sobrevivência que, em sua essência, consideram o tempo até determinado evento ocorrer. Portanto, para contextualizar os métodos utilizados para a análise do trabalho de parto, apresentamos a seguir os conceitos fundamentais de análise de sobrevivência, bem como a estrutura de modelagem conjunta de dados longitudinais e de sobrevivência aplicada aos dados.

2.1 Análise do Tempo até o Evento de Interesse

Modelos de análise de sobrevivência consideram, basicamente, que a variável de interesse é o tempo até que um determinado evento ocorra, também conhecido como tempo de falha, tempo de sobrevivência ou tempo de evento (LAWLESS, 2011). Muito utilizado em estudos clínicos e epidemiológicos, esse método pode ser utilizado para avaliar o tempo até o óbito, tempo até

o surgimento de um tumor ou até mesmo o tempo até o desenvolvimento ou reincidência de uma doença, por exemplo. Uma outra característica inerente à análise de sobrevivência é a capacidade de incorporar informações de tempos incompletos, ou censurados, que são os tempos dos indivíduos para os quais o tempo de evento não foi completamente observado. Nesse caso, a análise de dados censurados depende da natureza do mecanismo de censura, que caracteriza como a censura ocorre. Em particular, podemos caracterizar o mecanismo de censura da seguinte maneira (LOUZADA-NETO; MAZUCHELI; ACHCAR, 2001):

- Censura à direita: o evento de interesse ocorre após um certo período de tempo. Esse tipo de censura ocorre, por exemplo, quando um indivíduo chega ao término do estudo sem que o evento de interesse tenha ocorrido (Tipo I), quando o estudo termina após o evento de interesse ocorrer em um número pré estabelecido de indivíduos (Tipo II) ou quando o paciente se muda da região ou centro do estudo (aleatória);
- Censura à esquerda: o evento de interesse ocorre antes de um certo período de tempo. Um exemplo típico é quando o indivíduo chega para o início do estudo já com a presença do evento de interesse;
- Censura intervalar: o evento de interesse ocorre em um intervalo de tempo, sem que se saiba exatamente quando ocorreu. Nesse caso, podemos citar a ocorrência do evento de interesse entre as idas do paciente ao médico ao longo do tempo.

Além disso, podemos classificá-lo de acordo com a dependência da probabilidade de censura e o processo de falha. Dizemos que a censura é informativa quando um sujeito se retira do estudo por razões diretamente relacionadas ao tempo esperado de falha. Por outro lado, a censura será não informativa quando não apresentar nenhum relacionamento com o processo de falha.

Definindo T^* como a variável aleatória contínua e não negativa que representa os tempos de falha em estudo, podemos expressar a probabilidade de um determinado evento ocorrer após um determinado tempo t como

$$S(t) = \Pr(T^* > t) = \int_t^{\infty} p(s) ds.$$

Nesse caso, essa é a função de probabilidade de sobrevivência, ou função de sobrevivência, sendo que $p(\cdot)$ denota a função de densidade de probabilidade de T^* . Essa função deve ser monótona não crescente a medida que t aumenta, em que $S(t = 0) = 1$ e $S(t = \infty) = 0$.

A função que descreve o risco instantâneo de um evento ocorrer em um intervalo de tempo $[t, t + dt)$, dado que o indivíduo sobreviveu até o tempo t , é a função de risco, definida por

$$h(t) = \lim_{dt \rightarrow 0} \frac{\Pr(t \leq T^* < t + dt \mid T^* \geq t)}{dt}, \quad t > 0.$$

As funções de sobrevivência e de risco são inversamente proporcionais, isto é, quando o risco aumenta, a probabilidade de sobrevivência diminui e vice-versa. Também, é possível escrever a função de sobrevivência em termos da função de risco

$$S(t) = \exp\{-H(t)\} = \exp\left\{-\int_0^t h(s)ds\right\},$$

em que $H(\cdot)$ é conhecida como função de risco acumulado, que descreve o risco acumulado até o tempo t . Essa função também pode ser interpretada como o número esperado de eventos a serem observados até o tempo t .

Quando estamos interessados em estimar essas funções, devemos levar em conta a informação da censura. Em particular, o tempo de evento observado para o indivíduo i , denotado por T_i , é definido como o valor mínimo entre o verdadeiro tempo de evento e o tempo de censura C_i . Para indicar a presença ou não de censura, usamos o indicador de evento $\delta_i = I(T_i^* \leq C_i)$, que assume valor 1 se o tempo de evento observado corresponde a um verdadeiro tempo de evento e 0 caso contrário. Dessa forma, avaliamos as características da distribuição de T_i^* usando somente as informações disponíveis, ou seja, utilizando o par ordenado $\{T_i, \delta_i\}$

A modelagem da função de sobrevivência pode acontecer de forma paramétrica ou não paramétrica. Os métodos não paramétricos geralmente são utilizados quando o processo de estimação é realizado sem fazer qualquer suposição acerca da distribuição de probabilidade do tempo de falha, ou seja, não impomos nenhum modelo teórico para as falhas e/ou censura. Para esse tipo de modelagem, o método mais conhecido é o estimador de Kaplan-Meier ([KAPLAN; MEIER, 1958](#))

$$\hat{S}_{KM}(t) = \prod_{i:t_i \leq t} \frac{r_i - d_i}{r_i},$$

em que

- t_1, \dots, t_k são os tempos únicos de falha ordenados observados na amostra;
- r_i denota o número de indivíduos que estão em risco no tempo t_i , $i = 1, \dots, k$;
- d_i é o número de eventos observados no tempo t_i .

Na abordagem paramétrica considera-se que a função de sobrevivência $S(t)$ possa ser modelada por alguma distribuição de probabilidades conhecida como, por exemplo, distribuição Weibull, distribuição log-normal ou distribuição exponencial, sendo que a estimação dos parâmetros geralmente é feita a partir do método de maximização da função de verossimilhança. Sendo assim, assumimos que o par ordenado $\{T_i, \delta_i\}$, $i = 1, \dots, n$, observados na amostra pertence a classe $\mathcal{P} = \{p(\cdot; \boldsymbol{\theta}), \boldsymbol{\theta} \in \mathbb{R}\}$, parametrizada por $\boldsymbol{\theta}$ e com função de distribuição associada $p(t; \boldsymbol{\theta})$.

Para considerar a informação da censura na construção da verossimilhança, levamos em conta que se observarmos o tempo de evento T_i em um sujeito i , indicaremos que ele

contribui para a verossimilhança com $p(T_i; \boldsymbol{\theta})$, enquanto se seu tempo de evento for censurado, isto é, $T_i^* > T_i = C_i$, indicaremos que tudo que sabemos é que ele sobreviveu até o respectivo tempo e, conseqüentemente, contribuirá com $S_i(T_i; \boldsymbol{\theta})$. Assim, temos a seguinte função de log-verossimilhança

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n [\delta_i \log p(T_i; \boldsymbol{\theta}) + (1 - \delta_i) \log S_i(T_i; \boldsymbol{\theta})].$$

Usando as relações $h(t) = p(t)/S(t)$ e $S(t) = \exp\{-H(t)\}$, também podemos reescrever a função de log-verossimilhança em função da função de risco

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^n \left[\delta_i \log h_i(T_i; \boldsymbol{\theta}) - \int_0^{T_i} h_i(s; \boldsymbol{\theta}) ds \right].$$

Dado que a função de log-verossimilhança é definida, podemos aplicar um algoritmo de maximização iterativo, tal como Newton-Raphson (LANGE, 2004), e encontrar as estimativas de $\hat{\boldsymbol{\theta}}$ cuja função é máxima.

Por fim, quando estamos interessados em avaliar o impacto de covariáveis nas estimativas de sobrevivência, podemos considerar os modelos de risco relativo, conhecidos também como modelo de Cox (COX, 1972). Esses modelos assumem que covariáveis possuem um efeito multiplicativo no risco de um evento e são formulados como

$$\begin{aligned} h_i(t | \mathbf{w}_i) &= \lim_{dt \rightarrow 0} \Pr(t \leq T^* < t + dt | T^* \geq t, \mathbf{w}_i) / dt \\ &= h_0(t) \exp(\boldsymbol{\gamma}^\top \mathbf{w}_i) \end{aligned} \quad (2.1)$$

em que $\mathbf{w}_i^\top = (w_{i1}, \dots, w_{ip})$ denota o vetor de covariáveis assumidas serem associadas à função de risco de cada indivíduo e $\boldsymbol{\gamma}$ é o vetor de coeficientes de regressão. A função $h_0(t)$ é chamada de risco basal, ou função de risco basal, e corresponde à função de risco associada a $\boldsymbol{\gamma}^\top \mathbf{w}_i = 0$.

2.2 Inclusão de Covariáveis Dependentes do Tempo

No modelo (2.1) assumimos que a função de risco depende somente de covariáveis cujos valores são constantes no tempo. No entanto, em alguns casos, é interessante avaliar se covariáveis dependentes do tempo estão associadas com o risco de ocorrência de um evento.

Conforme sugerido por Rizopoulos (2012b), essas covariáveis dependentes do tempo podem ser classificadas em duas categorias: variáveis exógenas e endógenas. A justificativa para essa distinção é que covariáveis endógenas necessitam de um tratamento especial. Assim, supondo que $y_i(t)$ seja o vetor da covariável no tempo t para o sujeito i e $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s < t\}$ a trajetória da covariável até o tempo t , a definição formal de uma covariável exógena é feita da seguinte maneira (KALBFLEISCH; PRENTICE, 2002)

$$\Pr(s \leq T_i^* < s + ds | T_i^* \geq s, \mathcal{Y}_i(s)) = \Pr(s \leq T_i^* < s + ds | T_i^* \geq s, \mathcal{Y}_i(t)) \quad (2.2)$$

para todo s, t tal que $0 < s \leq t$ e $ds \rightarrow 0$. Uma definição equivalente é

$$\Pr(\mathcal{Y}_i(t) | \mathcal{Y}_i(s), T_i^* \geq s) = \Pr(\mathcal{Y}_i(t) | \mathcal{Y}_i(s), T_i^* = s), \quad s \leq t \quad (2.3)$$

que formaliza a ideia de que $y_i(\cdot)$ está associado com a taxa de falha ao longo do tempo, mas o futuro do processo até um tempo qualquer $t > s$ não é afetado pela ocorrência de falha no tempo s . Segundo (RIZOPOULOS, 2012b), uma covariável exógena é um processo previsível, o que significa que seu valor é conhecido em qualquer momento s infinitesimalmente anterior a t . Por outro lado, uma covariável dependente do tempo endógena é aquela que não satisfaz as condições anteriores e, dessa forma, não é previsível. Fisher e Lin (1999) ressaltam que as covariáveis exógenas são aquelas que não estão diretamente relacionadas ao mecanismo de falha como, por exemplo, a idade de um indivíduo em um estudo de longo prazo, ou então, o nível de poluição do ar como um fator de risco para ataque de asma.

Sob as definições (2.2) e (2.3), podemos definir a função de sobrevivência condicionada à trajetória da covariável para covariáveis exógenas usando sua relação com a função de risco

$$\begin{aligned} S_i(t | \mathcal{Y}_i(t)) &= \Pr(T_i^* > t | \mathcal{Y}_i(t)) \\ &= \exp \left\{ - \int_0^t h_i(s | \mathcal{Y}_i(s)) ds \right\}. \end{aligned} \quad (2.4)$$

Com relação às covariáveis endógenas, elas geralmente são medidas dependentes do tempo avaliadas nos próprios indivíduos em estudo, ou seja, são geradas pelos indivíduos. Como exemplo, podemos citar a pressão sanguínea, peso relativo, histórico de procedimentos, marcadores genéticos, exposição a uma determinada doença, etc.

Para esses casos, existem características importantes que complicam a análise estatística. A primeira característica de uma covariável endógena é que, para sua existência, é necessário que o indivíduo em estudo não tenha experimentado a falha. Dessa forma, quando o evento de interesse é a morte, por exemplo, sua trajetória carrega a informação direta sobre o tempo de falha. Mais especificamente, dado que $y_i(t - ds)$ com $ds \rightarrow 0$ existe, a função de sobrevivência satisfaz a seguinte condição

$$S_i(t | \mathcal{Y}_i(t)) = \Pr(T_i^* > t | \mathcal{Y}_i(t)) = 1. \quad (2.5)$$

Por outro lado, a falha do indivíduo no tempo s corresponde à inexistência da covariável para $t \geq s$, o que tem implicação direta na violação da condição de endogeneidade (2.3). Além disso, uma consequência de (2.5) é que, ao contrário das covariáveis exógenas, a função de risco, antes definida como

$$h_i(t | \mathcal{Y}_i(t)) = \lim_{dt \rightarrow 0} \frac{\Pr(t \leq T_i^* < t + dt | T_i^* \geq t, \mathcal{Y}_i(t))}{dt}$$

não é diretamente relacionada à função de sobrevivência. Assim, as funções

$$S_i(t | \mathcal{Y}_i(t)) = \exp \left\{ - \int_0^t h_i(s | \mathcal{Y}_i(s)) ds \right\} \text{ e } p(t | \mathcal{Y}_i(t)) = h_i(t | \mathcal{Y}_i(t)) S_i(t | \mathcal{Y}_i(t))$$

não têm a interpretação das funções usuais. Devido à essa característica, a construção da função de verossimilhança baseada em $p(\cdot)$ e $S(\cdot)$ não tem sentido para variáveis endógenas. Outra característica das covariáveis endógenas é que geralmente são medidas com erro, e a última implicação sobre essas covariáveis é que a trajetória completa até algum tempo t não é completamente observada, pois as observações geralmente são feitas em momentos específicos como, por exemplo, uma ocasião em que o paciente visita o hospital para fazer avaliações. Nesse caso, a variável é medida apenas no momento da avaliação, implicando que os valores nos intervalos entre as avaliações são, possivelmente, desconhecidos.

2.3 O Modelo Conjunto

Quando estamos interessados em avaliar a associação entre variáveis dependentes do tempo, especificamente do tipo endógena, e modelagem de sobrevivência, é possível utilizar a modelagem conjunta entre a componente de sobrevivência e a longitudinal. A ideia é associar o modelo de sobrevivência, que é de principal interesse, com um modelo de medidas repetidas adequado para a variável longitudinal (RIZOPOULOS *et al.*, 2010) a partir de uma estrutura latente (Figura 2). Em suma, combinando os processos longitudinal e de sobrevivência, podemos prever o risco de um evento futuro baseando-se na evolução longitudinal da variável medida repetidamente no tempo e, devido a essa característica, o modelo conjunto tem ganhado atenção nos últimos anos (SUDELL; KOLAMUNNAGE-DONA; TUDUR-SMITH, 2016; PAGEDORGIU *et al.*, 2019). Para introduzir esse modelo, utilizaremos notação similar à seção anterior.

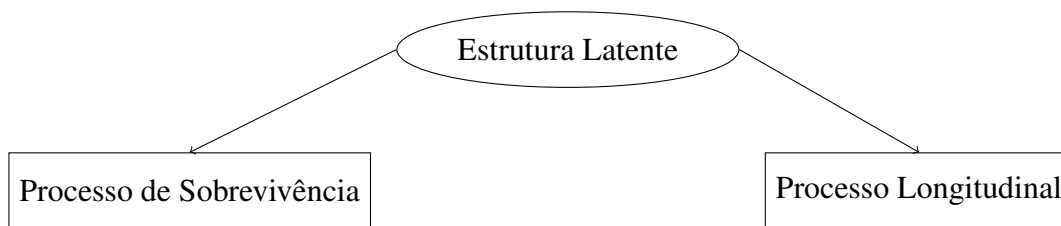


Figura 2 – Dependência latente entre o processo de sobrevivência e longitudinal.

Fonte: Elaborada pelo autor.

Definimos T_i^* como o tempo de evento verdadeiro para o i -ésimo sujeito, T_i o tempo de evento observado, definido como o mínimo entre o potencial tempo de censura C_i e T_i , isto é $T_i^* = \min(C_i, T_i)$, e $\delta_i = I(T_i^* \leq C_i)$ uma função indicadora de ocorrência de evento. Para a variável dependente do tempo, denotamos como $y_i(t)$ o valor observado no tempo t , resultando no conjunto de medidas $y_{ij} = \{y_i(t_{ij}), j = 1, \dots, n_i\}$, em que n_i representa o número de observações longitudinais para cada sujeito, do individuo i .

Para quantificar a evolução longitudinal no modelo de sobrevivência, consideramos que o risco relativo é relacionado a um conjunto de covariáveis \mathbf{w}_i e a verdadeira trajetória da variável

longitudinal, denotada por $m_i(t)$,

$$h_i(t | \mathcal{M}_i(t), \mathbf{w}_i) = h_0(t) \exp \left\{ \boldsymbol{\gamma}^\top \mathbf{w}_i + \alpha m_i(t) \right\},$$

em que $\mathcal{M}_i(t) = \{m_i(u), 0 \leq u < t\}$ denota a história do verdadeiro processo longitudinal não observável até o tempo t , $h_0(t)$ é a função de risco basal e α o parâmetro que mede a associação entre a componente longitudinal e de sobrevivência. Nesse caso, a função de sobrevivência pode ser obtida por

$$\begin{aligned} S_i(t | \mathcal{M}_i(t), \mathbf{w}_i; \boldsymbol{\beta}) &= \Pr(T_i^* > t | \mathcal{M}_i(t), \mathbf{w}_i) \\ &= \exp \left(- \int_0^t h_0(s) \exp \left\{ \boldsymbol{\gamma}^\top \mathbf{w}_i + \alpha m_i(s) \right\} ds \right), \end{aligned}$$

implicando que ela depende de toda a trajetória $\mathcal{M}_i(t)$.

Como a variável dependente do tempo t é, possivelmente, uma medida observada com erro, a modelagem conjunta postula um modelo linear misto longitudinal para descrever sua evolução

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= \mathbf{x}_i^\top(t) \boldsymbol{\beta} + \mathbf{z}_i^\top(t) \mathbf{b}_i + \varepsilon_i(t), \end{aligned}$$

sendo $m_i(t)$ a verdadeira, e não observável, trajetória e $\mathbf{x}_i(t)$ e $\mathbf{z}_i(t)$ observações associadas aos efeitos fixos, $\boldsymbol{\beta}$, e aleatórios, \mathbf{b}_i , respectivamente. Assumimos que os efeitos aleatórios \mathbf{b}_i seguem distribuição normal com média zero e matriz de covariâncias \mathbf{D} e que os termos independentes do erro $\varepsilon_i(t)$ seguem distribuição normal com média zero e variância σ^2 .

2.4 Função de Verossimilhança Conjunta

A estimação por máxima verossimilhança para o modelo conjunto se baseia na maximização da log-verossimilhança correspondente à distribuição conjunta do tempo até o evento e as medidas longitudinais $\{T_i, \delta_i, \mathbf{y}_i\}$, assumindo-se que o vetor de parâmetros associados aos efeitos aleatórios \mathbf{b}_i é subjacente às partes longitudinal e de sobrevivência.

Seja $\boldsymbol{\theta} = (\boldsymbol{\theta}_t^\top, \boldsymbol{\theta}_y^\top, \boldsymbol{\theta}_b^\top)$ o vetor de parâmetros, com $\boldsymbol{\theta}_t$ denotando o vetor de parâmetros associados ao tempo de evento, $\boldsymbol{\theta}_y$ o vetor de parâmetros associados ao desfecho longitudinal e $\boldsymbol{\theta}_b$ o vetor de parâmetros únicos da matriz de covariância dos efeitos aleatórios, \mathbf{y}_i o vetor de dimensão $n_i \times 1$ relacionado às observações para as respostas longitudinais do i -ésimo sujeito e $p(\cdot)$ uma função de densidade de probabilidade apropriada, podemos formular a contribuição do i -ésimo sujeito para a log-verossimilhança conjunta como

$$\begin{aligned} \log p(T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) &= \log \int p(T_i, \delta_i, \mathbf{y}_i, \mathbf{b}_i; \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \log \int p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}_t, \boldsymbol{\beta}) \left[\prod_j p\{y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}_y\} \right] p(\mathbf{b}_i; \boldsymbol{\theta}_b) d\mathbf{b}_i, \end{aligned} \quad (2.6)$$

em que a densidade condicional para o sub-modelo de sobrevivência $p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}_t, \boldsymbol{\beta})$ é dada por

$$\begin{aligned} p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}_t, \boldsymbol{\beta}) &= \{h_i(T_i | \mathcal{M}_i(T_i); \boldsymbol{\theta}_t, \boldsymbol{\beta})\}^{\delta_i} S_i(T_i | \mathcal{M}_i(T_i); \boldsymbol{\theta}_t, \boldsymbol{\beta}) \\ &= \left[h_0(T_i) \exp \left\{ \boldsymbol{\gamma}^\top \mathbf{w}_i + \alpha m_i(T_i) \right\} \right]^{\delta_i} \\ &\quad \times \exp \left(- \int_0^{T_i} h_0(s) \exp \left\{ \boldsymbol{\gamma}^\top \mathbf{w}_i + \alpha m_i(s) \right\} ds \right), \end{aligned} \quad (2.7)$$

sendo $h_0(\cdot)$ uma função positiva dependente do tempo. Além disso, a função densidade conjunta para as respostas longitudinais, juntamente com os efeitos aleatórios é dada por

$$\begin{aligned} p\{\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}\} p(\mathbf{b}_i; \boldsymbol{\theta}) &= \prod_j p\{y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}_y\} p(\mathbf{b}_i; \boldsymbol{\theta}_b) \\ &= (2\pi\sigma^2)^{-n_i/2} \exp \left\{ -\|\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta} - \mathbf{Z}_i\mathbf{b}_i\|^2 / 2\sigma^2 \right\} \\ &\quad \times (2\pi)^{-q_b/2} \det(\mathbf{D})^{-1/2} \exp \left(-\mathbf{b}_i^\top \mathbf{D}^{-1} \mathbf{b}_i / 2 \right) \end{aligned} \quad (2.8)$$

em que q_b representa a dimensionalidade do vetor de efeitos aleatórios e $\|\mathbf{x}\| = \{\sum_i x_i^2\}^{1/2}$ a norma euclidiana do vetor \mathbf{x} .

Os estimadores de máxima verossimilhança são geralmente obtidos a partir de algoritmos de maximização, tal como o algoritmo Expectation-Maximization (EM, [Dempster, Laird e Rubin \(1977\)](#)) ou Newton-Raphson ([Lange \(2004\)](#)). Nesse caso, estamos interessados em maximizar a função log-verossimilhança $\ell(\boldsymbol{\theta}) = \sum_i \log p(T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})$. No procedimento de estimação, o vetor escore associado ao logaritmo da função de verossimilhança (log-verossimilhança) é dado por

$$\begin{aligned} S(\boldsymbol{\theta}) &= \sum_i \frac{\partial}{\partial \boldsymbol{\theta}^\top} \log \int p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i; \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \sum_i \frac{1}{p(T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})} \frac{\partial}{\partial \boldsymbol{\theta}^\top} \int p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i; \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \sum_i \frac{1}{p(T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})} \int \frac{\partial}{\partial \boldsymbol{\theta}^\top} \{p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i; \boldsymbol{\theta})\} d\mathbf{b}_i \\ &= \sum_i \int \left[\frac{\partial}{\partial \boldsymbol{\theta}^\top} \log \{p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i; \boldsymbol{\theta})\} \right] \\ &\quad \times \frac{p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i; \boldsymbol{\theta})}{p(T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})} d\mathbf{b}_i \\ &= \sum_i \int A(\boldsymbol{\theta}, \mathbf{b}_i) p(\mathbf{b}_i | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) d\mathbf{b}_i, \end{aligned} \quad (2.9)$$

em que $A(\cdot)$ denota o vetor escore completo dos dados, dado por

$$A(\boldsymbol{\theta}, \mathbf{b}_i) = \partial \{ \log p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}) + \log p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) + \log p(\mathbf{b}_i; \boldsymbol{\theta}) \} / \partial \boldsymbol{\theta}^\top.$$

Note que o vetor escore observado dos dados é expresso como o valor esperado do vetor escore completo com relação à distribuição a posteriori dos efeitos aleatórios. Segundo [Rizopoulos \(2012b\)](#), isto implica que a equação (2.9) pode desempenhar um duplo papel. Em particular,

se as equações correspondentes forem resolvidas com respeito a $\boldsymbol{\theta}$, com $p(\mathbf{b}_i | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})$ fixo no valor de $\boldsymbol{\theta}$ da iteração anterior, esse procedimento corresponde ao algoritmo EM, enquanto que se as equações da função escore forem resolvidas com respeito a $\boldsymbol{\theta}$ considerando $p(\mathbf{b}_i | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})$, também como função de $\boldsymbol{\theta}$, o procedimento será correspondente à maximização direta da log-verossimilhança observada $\ell(\boldsymbol{\theta})$. Especificamente nesta Tese, para maximizar a verossimilhança dos modelos propostos utilizamos rotinas PORT (FOX; HALL; SCHRYER, 1978), disponíveis no pacote `nlmminb()` (GAY, 1990), que são alternativas ao método quase-Newton BFGS (NASH, 1990).

Por fim, com as estimativas obtidas, podemos calcular o erro padrão dos parâmetros baseando-se na matriz de informação estimada, tal que

$$\text{vâr}(\hat{\boldsymbol{\theta}}) = \left\{ \mathcal{J}(\hat{\boldsymbol{\theta}}) \right\}^{-1}, \quad \text{com} \quad \mathcal{J}(\hat{\boldsymbol{\theta}}) = - \sum_{i=1}^n \frac{\partial S_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}.$$

2.5 Aproximação da Função de Verossimilhança Conjunta

Em geral, as integrais envolvidas na especificação do vetor escore (2.9) não apresentam forma fechada, sendo necessário utilizar métodos numéricos para realizar seu cálculo. Podemos notar que duas integrais devem ser calculadas, uma para a função de sobrevivência e outra com respeito aos efeitos aleatórios. Em particular, a primeira é sempre unidimensional e pode ser aproximada usando quadraturas de Gauss-Kronrod (PRESS *et al.*, 2007), enquanto que a segunda depende da dimensionalidade dos efeitos aleatórios, podendo-se demandar alto esforço computacional (RIZOPOULOS, 2012b).

Para esse caso, uma escolha padrão é utilizar a quadratura de Gauss-Hermite, que aproxima uma integral a partir da soma ponderada do integrando em abcissas pré especificadas. Esse é um método de integração numérica frequentemente utilizado na estatística devido à sua relação com a distribuição normal (LIU; PIERCE, 1994). Em geral, o método se baseia em integrais do tipo

$$\int_{-\infty}^{\infty} f(x) \exp(-x^2) dx \quad (2.10)$$

sendo que, em muitas aplicações, o kernel da densidade normal é explicitamente observado no integrando. Nos casos em que esse fato não é visível, podemos multiplicar e dividir o integrando por $\exp(-x^2)$ ou por alguma outra função de densidade gaussiana. Nesse caso, a integral (2.10) é aproximada por

$$\int_{-\infty}^{\infty} f(x) \exp(-x^2) dx \approx \sum_{i=1}^m u_i f(x_i),$$

em que m denota o número de pontos, x_i são as raízes do polinômio de Hermite de ordem m e u_i são os pesos correspondentes. De maneira geral, queremos aplicar este tipo de quadratura para calcular uma integral expressa da seguinte forma

$$\int_{-\infty}^{\infty} g(t) dt,$$

em que $g(t) > 0$, para todo t , é a função de verossimilhança. O objetivo é aplicar uma transformação em t tal que $g(t)$ será avaliado em um intervalo adequado. Assim, podemos reescrever (2.10) de modo que

$$\int_{-\infty}^{\infty} f(t)\phi(t; \mu, \sigma)dt, \quad (2.11)$$

sendo $\phi(t; \mu, \sigma)$ a função de densidade da distribuição normal com média μ e variância σ^2 . Os nós e os pesos são modificados de maneira que $t_i = \mu + 2^{1/2}\sigma x_i$ e $u_i/\sqrt{\pi}$, tal que a escolha de μ e σ permitirão avaliar $g(t)$ em uma região adequada. Particularmente, assumimos $\hat{\mu}$ como a moda de $g(t)$ e $\hat{\sigma} = 1/\sqrt{\hat{j}}$, tal que

$$\hat{j} = -\frac{\partial^2}{\partial t^2} \log g(t) \Big|_{t=\hat{\mu}}.$$

Isto nos leva à densidade normal $\phi(t; \hat{\mu}, \hat{\sigma})$ ter a mesma derivada logarítmica parcial de segunda ordem, na moda, do integrando $g(t)$. Definindo

$$v(t) = \frac{g(t)}{\phi(t; \hat{\mu}, \hat{\sigma})}$$

podemos escrever

$$\int_{-\infty}^{\infty} g(t)dt = \int_{-\infty}^{\infty} v(t)\phi(t; \hat{\mu}, \hat{\sigma})dt.$$

Dessa forma, aplicando a quadratura na forma (2.11) usando $\phi(t; \hat{\mu}, \hat{\sigma})$, a função $v(t)$ e, conseqüentemente, $g(t)$, poderá ser avaliada no intervalo adequado, o que resulta em

$$\begin{aligned} \int_{-\infty}^{\infty} g(t)dt &\approx \sum_{i=1}^m \frac{u_i}{\sqrt{\pi}} v(\hat{\mu} + 2^{1/2}\hat{\sigma}x_i) \\ &= 2^{1/2}\hat{\sigma} \sum_{i=1}^m u_i^* g(\hat{\mu} + 2^{1/2}\hat{\sigma}x_i), \end{aligned}$$

com $u_i^* = u_i \exp(x_i^2)$.

No contexto do modelo conjunto, consideramos a quadratura de Gauss-Hermite para qualquer forma de $A(\cdot)$ como função dos efeitos aleatórios, permitindo-se que o vetor escore seja calculado, tal como

$$\begin{aligned} E \{A(\boldsymbol{\theta}, \mathbf{b}_i) \mid T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}\} &= \int A(\boldsymbol{\theta}, \mathbf{b}_i) p(\mathbf{b}_i \mid T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) d\mathbf{b}_i \\ &\approx 2^{q_b/2} \sum_{t_1 \dots t_q} \pi_t A(\boldsymbol{\theta}, \mathbf{b}_t \sqrt{2}) p(\mathbf{b}_t \sqrt{2} \mid T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) \exp(\|\mathbf{b}_t\|^2), \end{aligned}$$

em que $\sum_{t_1 \dots t_q}$ é a abreviação de $\sum_{t_1=1}^K \dots \sum_{t_q=1}^K$, sendo K o número de pontos de quadraturas e $\mathbf{b}_t^\top = (b_{t_1}, \dots, b_{t_q})$ as abcissas com pesos correspondentes iguais a π_t . A localização das abcissas e os pesos são utilizados para levar a aproximação exata da integral se o integrando for expresso da forma $\exp(-\mathbf{b}^\top \mathbf{b})l(\mathbf{b})$, com $l(\mathbf{b})$ denotando um polinômio de grau $2K - 1$, ou menor, que interpola as abcissas. Assim, a qualidade da aproximação da integral aumenta a medida que o

número de pontos K aumenta. No entanto, como a quadratura requer a avaliação do integrando sobre o produto cartesiano das abcissas para cada efeito aleatório, o esforço computacional aumenta exponencialmente com a dimensão q .

De acordo com Rizopoulos (2012b), um outro aspecto crítico que influencia bastante na aproximação é a localização dos pontos de quadratura com respeito à massa principal do integrando, ou seja, se $g(\mathbf{b}) = A(\boldsymbol{\theta}, \mathbf{b})p(\mathbf{b} | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})$ está concentrado longe do zero, ou se o espalhamento em $g(\mathbf{b})$ é diferente da função de peso $\exp(-\mathbf{b}^\top \mathbf{b})$, aplicar o procedimento padrão diretamente à $g(\mathbf{b})$ pode resultar em uma aproximação pobre, mesmo para grandes valores de K , pelo fato das abcissas não de serem alocadas corretamente onde as maiores massas de $g(\mathbf{b})$ estão. Para solucionar esse problema, sugere-se a utilização da regra adaptativa, que centraliza e escala apropriadamente o integrando em cada iteração do algoritmo de maximização. Mais especificamente,

$$E \{A(\boldsymbol{\theta}, \mathbf{b}_i) | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}\} \approx 2^{qb/2} |\hat{\mathbf{B}}_i|^{-1} \sum_{t_1 \dots t_q} \pi_t A(\boldsymbol{\theta}, \hat{\mathbf{r}}_t) p(\hat{\mathbf{r}}_t | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) \exp(\|\hat{\mathbf{r}}_t\|^2), \quad (2.12)$$

sendo que $\hat{\mathbf{r}}_t = \hat{\mathbf{b}}_i + \sqrt{2\hat{\mathbf{B}}_i^{-1}} \mathbf{b}_t$, $\hat{\mathbf{b}}_i = \arg \max_{\mathbf{b}} \{\log p(T_i, \delta_i, \mathbf{y}_i, \mathbf{b}; \boldsymbol{\theta})\}$ e $\hat{\mathbf{B}}_i$ denota o fator de Choleski de $\hat{\mathbf{H}}_i$, sendo $\hat{\mathbf{H}}_i = -\partial^2 \log p(T_i, \delta_i, \mathbf{y}_i, \mathbf{b}; \boldsymbol{\theta}) / \partial \mathbf{b} \partial \mathbf{b}^\top |_{\mathbf{b}=\hat{\mathbf{b}}_i}$. A partir dessa transformação, o integrando se aproximará do comportamento de uma distribuição normal $N(\mathbf{0}, 2^{-1}\mathbf{I})$ e, dado que o peso da quadratura de Gauss-Hermite é proporcional a função densidade dessa distribuição, teremos uma aproximação ótima. Assim, a regra adaptativa precisará de um menor número de quadraturas para obter um erro de mesma magnitude da regra padrão. No entanto, a necessidade de calcular a moda $\hat{\mathbf{b}}_i$ e a segunda derivada da matriz $\hat{\mathbf{H}}_i$ para cada indivíduo em cada iteração aumentará o esforço computacional.

Nesse caso, Rizopoulos (2012a) propôs a regra pseudo-adaptativa de Gauss-Hermite. A motivação do método se dá pelas propriedades da distribuição a posteriori dos efeitos aleatórios $p(\mathbf{b}_i | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})$. Escrevendo-a na escala logarítmica, podemos notar que essa densidade é proporcional a

$$\log p(\mathbf{b}_i | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) \propto \sum_{j=1}^{n_i} \log p\{y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}_y\} + \log p(\mathbf{b}_i; \boldsymbol{\theta}_b) + \log p(T_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}_t, \boldsymbol{\beta}).$$

Dessa forma, à medida que n_i cresce, o termo principal na função densidade log-posteriori é o logaritmo da função densidade do modelo linear misto, isto é, $\log p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}_y) = \sum_j \log p\{y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}_y\}$, que é quadrático em \mathbf{b}_i e apresenta forma similar à distribuição normal multivariada. A partir do Teorema do Limite Central Bayesiano (HINKLEY; COX, 1979) e sob as condições gerais de regularidade, à medida que $n_i \rightarrow \infty$,

$$p(\mathbf{b}_i | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) \xrightarrow{P} N(\tilde{\mathbf{b}}_i, \tilde{\mathbf{H}}_i^{-1}),$$

em que $\tilde{\mathbf{b}}_i = \arg \max_{\mathbf{b}} \{\log p(\mathbf{y}_i | \mathbf{b}; \boldsymbol{\theta}_y)\}$ e $\tilde{\mathbf{H}}_i = -\partial^2 \log p(\mathbf{y}_i | \mathbf{b}; \boldsymbol{\theta}_y) / \partial \mathbf{b} \partial \mathbf{b}^\top |_{\mathbf{b}=\tilde{\mathbf{b}}_i}$. Na prática, isso sugere que o aumento de n_i é suficiente para podermos recentralizar e reescalar o integrando para

cada sujeito usando somente informação proveniente do modelo com efeitos mistos da variável longitudinal. Logo, ao invés de usar a transformação padrão apresentada na quadratura adaptativa de Gauss-Hermite (2.12), os autores sugerem primeiro ajustar o modelo linear com efeitos mistos, extrair as estimativas empíricas de Bayes $\tilde{\mathbf{b}}_i = \arg \max_{\mathbf{b}} \{\log p(\mathbf{y}_i, \mathbf{b}; \tilde{\boldsymbol{\theta}}_y)\}$ e a respectiva matriz de covariância $\tilde{\mathbf{H}}_i^{-1}$, com $\tilde{\mathbf{H}}_i = -\partial^2 \log p(\mathbf{y}_i, \mathbf{b}; \tilde{\boldsymbol{\theta}}_y) / \partial \mathbf{b} \partial \mathbf{b}^\top |_{\mathbf{b}=\tilde{\mathbf{b}}_i}$, e usar a seguinte transformação

$$E \{A(\boldsymbol{\theta}, \mathbf{b}_i) | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}\} \approx 2^{q_b/2} |\tilde{\mathbf{B}}_i|^{-1} \sum_{t_1 \dots t_q} \pi_t A(\boldsymbol{\theta}, \tilde{\mathbf{r}}_t) p(\tilde{\mathbf{r}}_t | T_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) \exp(-\|\tilde{\mathbf{r}}_t\|^2),$$

em que $\tilde{\mathbf{r}}_t = \tilde{\mathbf{b}}_i + \sqrt{2\tilde{\mathbf{B}}_i^{-1}} \mathbf{b}_t$, $\tilde{\mathbf{B}}_i$ denota o fator de Choleski de $\tilde{\mathbf{H}}_i$ e $\tilde{\boldsymbol{\theta}}_y$ são as estimativas de máxima verossimilhança do modelo misto linear. De acordo com os autores, esse procedimento é similar ao adaptativo, mas o implementam apenas uma vez, no início da otimização, e não fazem atualização dos pontos da quadratura desde então. Com relação ao aspecto computacional, o método pseudo-adaptativo pode ser usado com menor número de pontos de quadratura do que o método padrão e, nesse caso, não é necessário fazer a atualização e o reescalonamento feito na regra adaptativa.

2.6 Testes de Hipótese e Intervalo de Confiança para os Parâmetros

Testes de hipóteses padrões para inferência via verossimilhança podem ser utilizados para o modelo conjunto (RIZOPOULOS, 2012b). Nesse caso, estamos interessados em testar hipóteses como:

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \quad \text{versus} \quad H_a : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0,$$

que, sob determinadas condições de regularidades (HINKLEY; COX, 1979), podem ser avaliadas pelos seguintes testes

- Teste de Razão de Verossimilhanças, definido pela estatística teste

$$TRV = -2 \left\{ l(\hat{\boldsymbol{\theta}}_0) - l(\hat{\boldsymbol{\theta}}) \right\},$$

em que $\hat{\boldsymbol{\theta}}_0$ e $\hat{\boldsymbol{\theta}}$ denotam as estimativas de máxima verossimilhança sob as hipóteses nula e alternativa, respetivamente;

- Teste de Escore, com estatística teste U definida por

$$U = S^\top(\hat{\boldsymbol{\theta}}_0) \left\{ \mathcal{I}(\hat{\boldsymbol{\theta}}_0) \right\}^{-1} S(\hat{\boldsymbol{\theta}}_0),$$

em que $S(\cdot)$ e $\mathcal{I}(\cdot)$ denotam a função escore e a matriz de informação de Fisher dos parâmetros do modelo sob a hipótese alternativa;

- Teste de Wald, cuja estatística teste W é dada por

$$W = (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_0)^\top \mathcal{I}(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{\theta}} - \hat{\boldsymbol{\theta}}_0).$$

Vale ressaltar que esses procedimentos requerem modelos encaixados e, levando em conta a teoria de verossimilhança, a distribuição assintótica de cada um desses testes segue uma distribuição Qui-Quadrado (χ^2) com grau de liberdade igual a diferença de dimensionalidade de $\hat{\theta}_0$ e $\hat{\theta}$, quando a hipótese nula é verdadeira (WILKS, 1938). Para um único parâmetro θ_j o teste de Wald é equivalente a $(\hat{\theta}_j - \hat{\theta}_{0j})/s.\hat{e}(\hat{\theta}_j)$, que sob a hipótese nula segue distribuição normal padrão.

Quando essa condição não é satisfeita é possível a utilização de testes ou critérios de informação, em que a principal ideia é se basear no valor da log-verossimilhança maximizada e aplicar uma penalização de acordo com o número de parâmetros utilizados. Os critérios mais comuns são o critério de informação de Akaike (AIC; (AKAIKE, 1974)) e critério de informação Bayesiano (BIC; (SCHWARZ *et al.*, 1978)), que são computados, respectivamente, por

$$AIC = -2l(\hat{\theta}) + 2n_{par} \quad \text{e} \quad BIC = -2l(\hat{\theta}) + n_{par} \log(n),$$

em que n_{par} denota o número de parâmetros no modelo, n o número de observações e $\hat{\theta}$ a estimativa de máxima verossimilhança de θ . Baseando-se nessas definições, quanto menor o valor dessas medidas, mais as estimativas dos parâmetros são adequadas para os dados considerados. Isso significa que quando estamos utilizando esses critérios para comparar dois modelos ajustados aos mesmos dados, decide-se pelo modelo que apresentar o menor valor do critério. Em geral, o AIC tende a selecionar modelos mais elaborados, enquanto que o BIC a modelos mais parcimoniosos.

Intervalos de confiança assintóticos de probabilidade podem ser obtidos para os parâmetros de interesse baseando-se na estatística de Wald sendo que, nesse caso, o intervalo com 95% de confiança é dado por $\hat{\theta} \pm 1,96s.\hat{e}(\hat{\theta})$, em que 1,96 é o percentil $z_{\alpha/2}$ ($\alpha = 5\%$) da distribuição normal.

Uma alternativa para se calcular intervalos de confiança para os parâmetros de interesse é a partir do método *bootstrap* (EFRON, 1979). Essa é uma técnica computacional que se baseia em reamostrar os dados, com reposição, calculando-se a estatística de interesse em cada nova amostra gerada. A partir desse procedimento podemos atribuir medidas de precisão, em termos de viés, variância, intervalos de confiança e previsão, entre outras, para as estimativas da amostra original.

Dentre as diferentes formas de *bootstrap*, utilizamos nessa Tese o método não paramétrico, que permite estimar a distribuição empírica dos estimadores sem fazer qualquer suposição sobre a distribuição dos mesmos. Sendo assim, para qualquer estatística $\hat{\theta} = s(\mathbf{x})$ de interesse, calculada em uma amostra $\mathbf{x} = (x_1, \dots, x_n)$, podemos fazer os seguintes passos:

1. Gerar amostras $\mathbf{x}^* = (x_1^*, \dots, x_n^*)$ com reposição;
2. Calcular a estatística de interesse $\hat{\theta}^* = s(\mathbf{x}^*)$;
3. Repetir B vezes os passos 1 e 2 para obter as estimativas de *bootstrap*.

A partir dos resultados obtidos, as estimativas de *bootstrap* com o respectivo erro padrão são dadas por

$$\hat{\theta}_B = \frac{\sum_{b=1}^B \hat{\theta}_b^*}{B} \quad \text{e} \quad s.\hat{e}.B = \sqrt{\frac{1}{B-1} \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta}_B)^2},$$

que podem ser utilizadas como estimativas para se calcular os intervalos de confiança.

No entanto, consideramos o método do percentil, que leva em conta os quantis empíricos das estimativas $\hat{\theta}^*$, sem que haja necessidade de se estimar o erro padrão. Assim, sendo $\hat{\theta}_{(1)}^*, \dots, \hat{\theta}_{(B)}^*$ as estimativas de *bootstrap* ordenadas, definimos os limites do intervalo de confiança como sendo os valores dos percentis $\hat{\theta}_{a/2}^*$ e $\hat{\theta}_{1-a/2}^*$ (EFRON, 1979).

2.7 Predição Dinâmica da Probabilidade de Sobrevida

Tendo como base o modelo conjunto, estamos interessados em prever as probabilidades de sobrevivência para um sujeito i dado o conjunto de medidas longitudinais $\mathcal{Y}_i(t) = \{y_i(s), 0 \leq s \leq t\}$ e um vetor de covariáveis fixas \mathbf{w}_i . Rizopoulos *et al.* (2010) sugere estimar a probabilidade condicional

$$\pi_i(u | t) = \Pr(T_i^* \geq u | T_i^* > t, \mathcal{Y}_i(t), \mathcal{D}_n), \quad (2.13)$$

sendo que $u > t$ e \mathcal{D}_n denota a amostra a qual o modelo foi ajustado.

Reescrevendo a equação (2.13)

$$\Pr(T_i^* \geq u | T_i^* > t, \mathcal{Y}_i(t), \mathcal{D}_n) = \int \Pr(T_i^* \geq u | T_i^* > t, \mathcal{Y}_i(t); \boldsymbol{\theta}) p(\boldsymbol{\theta} | \mathcal{D}_n) d\boldsymbol{\theta}$$

e usando a suposição de independência condicional entre o processo de sobrevivência e longitudinal, temos que

$$\begin{aligned} \Pr(T_i^* \geq u | T_i^* > t, \mathcal{Y}_i(t); \boldsymbol{\theta}) &= \int \Pr(T_i^* \geq u | T_i^* > t, \mathcal{Y}_i(t), \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i | T_i^* > t, \mathcal{Y}_i(t); \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \int \Pr(T_i^* \geq u | T_i^* > t, \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i | T_i^* > t, \mathcal{Y}_i(t); \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \int \frac{S_i\{u | \mathcal{M}_i(u, \mathbf{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}}{S_i\{t | \mathcal{M}_i(t, \mathbf{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}} p(\mathbf{b}_i | T_i^* > t, \mathcal{Y}_i(t); \boldsymbol{\theta}) d\mathbf{b}_i \end{aligned}$$

em que $S_i(\cdot)$ denota a função de sobrevivência, $\mathcal{M}_i(t)$ representa a história longitudinal, definida na Seção 2.3 e aproximada pelo modelo linear de efeito misto, e $\boldsymbol{\theta} = (\boldsymbol{\theta}_t^\top, \boldsymbol{\theta}_y^\top, \boldsymbol{\theta}_b^\top)^\top$ é o vetor de parâmetros do modelo conjunto.

Utilizando a estimativa de Bayes empírica para \mathbf{b}_i , Rizopoulos *et al.* (2010) propuseram o estimador de primeira ordem de $\pi_i(u | t)$, apresentada na equação (2.13), da seguinte maneira

$$\tilde{\pi}_i(u | t) = S_i\left\{u | \mathcal{M}_i(u, \hat{\mathbf{b}}_i^{(t)}, \hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}}\right\} / S_i\left\{t | \mathcal{M}_i(t, \hat{\mathbf{b}}_i^{(t)}, \hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}}\right\} + O([n_i(t)]^{-1}),$$

em que $\hat{\mathbf{b}}_i^{(t)}$ denota a moda a distribuição condicional, $\log p(\mathbf{b}_i | T_i^* > t, \mathcal{Y}_i(t), \mathbf{b}; \hat{\boldsymbol{\theta}})$, e $n_i(t)$ denota o número de observações longitudinais para o sujeito i no tempo t .

2.8 Comentários Finais

Neste capítulo apresentamos uma introdução ao contexto da modelagem de dados longitudinais e de sobrevivência. Exibimos como as covariáveis dinâmicas podem ser consideradas em um contexto de análise de sobrevivência e como elas são relacionadas com a função de risco e de sobrevivência. Apresentamos o modelo conjunto, a função de verossimilhança conjunta, aproximada por meio de quadraturas gaussianas, os de testes de hipótese e intervalo de confiança para o vetor de parâmetros $\boldsymbol{\theta}$ do modelo estudado. Por fim, exibimos como a predição dinâmica da probabilidade de sobrevivência pode ser feita considerando a modelagem conjunta de dados longitudinais e de sobrevivência.

O capítulo a seguir introduz as famílias de distribuições com função de risco flexível consideradas nesta Tese, bem como um estudo de simulação caracterizado por simulações que visam avaliar a eficiência do método da máxima verossimilhança para as distribuições propostas. Com a finalidade de selecionar modelos paramétricos usuais, avaliamos a eficiência da seleção de modelos via família gama generalizada e por critérios de informação como AIC e BIC. Por fim, os modelos estudados são ajustados ao contexto do trabalho de parto de modo que possibilite uma compreensão do mesmo levando-se em conta, principalmente, uma variável dinâmica avaliada longitudinalmente, a dilatação cervical.

MODELO CONJUNTO PARAMÉTRICO

Conforme citado no Capítulo 1, a função de distribuição Weibull é frequentemente utilizada em análise de sobrevivência, mas não pode ser usada para descrever tempos de falha com função de risco não monótona, como em forma de banheira ou unimodal. Sendo assim, propomos nesta Tese a utilização de generalizações da distribuição Weibull de modo que se obtenha um modelo mais flexível, capaz de lidar com diferentes formas de função de risco. Neste caso, consideramos as distribuições gama generalizada (STACY, 1962; PRENTICE, 1974), Weibull exponenciada (MUDHOLKAR; SRIVASTAVA, 1993), Weibull generalizada (MUDHOLKAR; SRIVASTAVA; KOLLIA, 1996), Marshall-Olkin-Weibull (MARSHALL; OLKIN, 1997) e função de potência Weibull generalizada (NIKULIN; HAGHIGHI, 2007). Uma vasta discussão sobre as generalizações recentes da distribuição Weibull é detalhada em Pham e Lai (2007).

Além do uso de generalizações da distribuição Weibull com função de risco flexível, também avaliamos a utilização da distribuição Nakagami- m (NK), proposta por NAKAGAMI (1960), que é um modelo estatístico poderoso utilizado frequentemente para modelar frequências de rádio (RAMOS; LOUZADA; RAMOS, 2018). Por possuir apenas dois parâmetros, pode ser considerada uma alternativa parcimoniosa em relação às generalizações da Weibull e outras distribuições paramétricas.

A seguir, apresentamos as distribuições flexíveis propostas, bem como um estudo de simulação para avaliar o método de estimação dos parâmetros e os critérios de discriminação de modelos paramétricos usuais. Também, apresentamos uma aplicação aos dados do trabalho de parto caracterizada pela modelagem de cesárea com dilatação cervical como dependente do tempo.

3.1 Distribuição Gama Generalizada

Introduzida por Stacy (1962), a distribuição gama-Weibull com três parâmetros é um modelo flexível para dados de confiabilidade devido à sua habilidade de acomodar várias formas da função de risco, como constante, crescente, decrescente, forma de banheira e unimodal. Como ela generaliza a distribuição gama com dois parâmetros, essa distribuição também é conhecida como gama generalizada (GG) e, para evitar confusão com a distribuição Weibull generalizada (GW), vamos nos referir a essa distribuição como GG.

Uma variável aleatória segue distribuição GG se sua função de densidade de probabilidades é dada por

$$f(t; a, b, k) = \frac{b}{\gamma(k)} a^{-bk} t^{bk-1} \exp\left\{-\left(\frac{t}{a}\right)^b\right\}, \quad t > 0$$

para $a, b, k > 0$, com função de sobrevivência e de risco iguais, respectivamente, a

$$S(t; a, b, k) = \frac{\gamma[k, (t/a)^b]}{\gamma(k)},$$

e

$$h(t; a, b, k) = \frac{b}{\gamma[k, (t/a)^b]} a^{-bk} t^{bk-1} \exp\left\{-\left(\frac{t}{a}\right)^b\right\}$$

em que $\gamma[\mu, \nu] = \int_{\nu}^{\infty} w^{\mu-1} e^{-w} dw$ é a função gama incompleta superior. Algumas distribuições importantes são casos especiais dessa distribuição, tais como a distribuição exponencial ($b = k = 1$), Weibull ($k = 1$), log-normal (caso limite quando $k \rightarrow \infty$) e gama ($b = 1$).

Para um comportamento mais estável quando essa distribuição tende à log-normal, uma parametrização alternativa foi proposta por Prentice (1974), que inclui uma nova classe de distribuições quando o parâmetro de forma k é negativo

$$f(t; \mu, \tau, q) = \frac{|q|}{\tau \gamma(q^{-2})} \left\{ q^{-2} (e^{-\mu t})^{q/\tau} \right\}^{q^{-2}} \exp\left\{-q^{-2} (e^{-\mu t})^{q/\tau}\right\}, \quad t > 0$$

com parâmetro de locação μ , escala $\tau > 0$ e forma q . As funções de sobrevivência e de risco para essa nova parametrização são dadas por

$$S(t; \mu, \tau, q) = \begin{cases} \frac{\gamma[q^{-2}, q^{-2} (e^{-\mu t})^{q/\tau}]}{\gamma(q^{-2})} & \text{if } q > 0 \\ 1 - \frac{\gamma[q^{-2}, q^{-2} (e^{-\mu t})^{q/\tau}]}{\gamma(q^{-2})} & \text{if } q < 0 \\ 1 - \Phi\left(\frac{\log(t) - \mu}{\tau}\right) & \text{if } q = 0, \end{cases}$$

em que $\Phi(\cdot)$ representa a função de distribuição acumulada da distribuição normal padrão, e

$$h(t; \mu, \tau, q) = \begin{cases} \frac{|q|}{\tau \gamma[q^{-2}, q^{-2} (e^{-\mu t})^{q/\tau}]} \left\{ q^{-2} (e^{-\mu t})^{q/\tau} \right\}^{q^{-2}} \exp\left\{-q^{-2} (e^{-\mu t})^{q/\tau}\right\} & \text{if } q > 0 \\ \frac{|q|}{\tau (\gamma(q^{-2}) - \gamma[q^{-2}, q^{-2} (e^{-\mu t})^{q/\tau}])} \left\{ q^{-2} (e^{-\mu t})^{q/\tau} \right\}^{q^{-2}} \exp\left\{-q^{-2} (e^{-\mu t})^{q/\tau}\right\} & \text{if } q < 0 \\ \frac{1}{\sqrt{2\pi\tau}} \exp\left\{-\frac{1}{2\tau^2} [\log(t) - \mu]^2\right\} / [1 - \Phi\left(\frac{\log(t) - \mu}{\tau}\right)] & \text{if } q = 0. \end{cases}$$

Sob essa parametrização, temos os casos especiais, como distribuição exponencial ($q = \tau = 1$), Weibull ($q = 1$), log-normal ($q = 0$) e gama ($q = \tau$). As propriedades, bem como aplicações na área de análise de sobrevivência, são discutidas em detalhes por [Cox et al. \(2007\)](#).

3.2 Distribuição Weibull Exponenciada

Introduzida por [Mudholkar e Srivastava \(1993\)](#), a distribuição Weibull exponenciada (EW) possui função de densidade de probabilidades dada por

$$f(t; a, b, \lambda) = a^{-1} b \lambda \left[1 - \exp \left\{ -(t/a)^b \right\} \right]^{\lambda-1} \exp \left\{ -(t/a)^b \right\} (t/a)^{b-1}, \quad t \geq 0$$

para $a, b, \lambda > 0$, com função de sobrevivência e de risco dadas, respectivamente, por

$$S(t; a, b, \lambda) = 1 - \left[1 - \exp \left\{ -(t/a)^b \right\} \right]^{\lambda},$$

e

$$h(t; a, b, \lambda) = \frac{a^{-1} b \lambda \left[1 - \exp \left\{ -(t/a)^b \right\} \right]^{\lambda-1} \exp \left\{ -(t/a)^b \right\} (t/a)^{b-1}}{1 - \left[1 - \exp \left\{ -(t/a)^b \right\} \right]^{\lambda}}.$$

A distribuição EW inclui os alguns casos especiais, como distribuição exponencial ($b = \lambda = 1$) e Weibull ($\lambda = 1$), com função de risco com formas análogas à GG.

3.3 Distribuição Weibull Generalizada

Proposta por [Mudholkar, Srivastava e Kollia \(1996\)](#), a distribuição Weibull generalizada (GW) possui a seguinte função de densidade de probabilidades

$$f(t; a, b, \lambda) = (ba)^{-1} (t/a)^{1/b-1} \left\{ 1 - \lambda (t/a)^{1/b} \right\}^{1/\lambda-1},$$

para $b \in \mathbb{R}$ and $\lambda, a > 0$. Nesse caso, as funções de sobrevivência e de risco são dadas, respectivamente, por

$$S(t; a, b, \lambda) = \left\{ 1 - \lambda (t/a)^{1/b} \right\}^{1/\lambda},$$

e

$$h(t; a, b, \lambda) = \frac{(ba)^{-1} (t/a)^{1/b-1}}{1 - \lambda (t/a)^{1/b}}.$$

Este também é um modelo muito flexível, apresentando função de risco constante, crescente, decrescente, em formato de banheira e unimodal. A distribuição GW inclui como casos especiais a distribuição exponencial ($\lambda = 0, b = 1$) e Weibull ($\lambda = 0$).

3.4 Distribuição Marshall-Olkin-Weibull

Marshall e Olkin (1997) propuseram um novo procedimento para introduzir um parâmetro adicional em uma família de distribuição. Nesse caso, os autores aplicaram o procedimento na distribuição Weibull, resultando na distribuição Marshall-Olkin-Weibull (MOW), com função de densidade dada por

$$f(t; a, b, \lambda) = \frac{\lambda a^{-1} b (t/a)^{b-1} \exp \{-(t/a)^b\}}{[1 - (1 - \lambda) \exp \{-(t/a)^b\}]^2},$$

para $a, b, \lambda > 0$, e com função de sobrevivência e de risco dadas, respectivamente, por

$$S(t; a, b, \lambda) = \frac{\lambda \exp \{-(t/a)^b\}}{1 - (1 - \lambda) \exp \{-(t/a)^b\}},$$

e

$$h(t; a, b, \lambda) = \frac{a^{-1} b (t/a)^{b-1}}{1 - (1 - \lambda) \exp \{-(t/a)^b\}}.$$

Além disso, Cordeiro e Lemonte (2013) derivaram diversas propriedades de estimação para a distribuição MOW. A função de risco resultante dessa distribuição apresenta diferentes formatos, assim como as citadas anteriormente.

3.5 Distribuição Potência Weibull Generalizada

A distribuição potência Weibull generalizada (GPW) é uma outra extensão da família Weibull que contém os diferentes formatos para a função de risco (NIKULIN; HAGHIGHI, 2007), ou seja, crescente, decrescente, unimodal e em forma de banheira. Sua função de densidade de probabilidades é dada por

$$f(t; a, b, \lambda) = \frac{b}{\lambda a^b} t^{b-1} \left\{ 1 + (t/a)^b \right\}^{1/\lambda - 1} \exp \left[1 - \left\{ 1 + (t/b)^a \right\}^{1/\lambda} \right],$$

para $a, b, \lambda > 0$, com função de sobrevivência e de riscos correspondentes, respectivamente, a

$$S(t; a, b, \lambda) = \exp \left[1 - \left\{ 1 + (t/b)^a \right\}^{1/\lambda} \right],$$

e

$$h(t; a, b, \lambda) = \frac{b}{\lambda a^b} t^{b-1} \left\{ 1 + (t/a)^b \right\}^{1/\lambda - 1}.$$

Os casos particulares da GPW são a distribuição Weibull ($\lambda = 1$) e exponencial ($\lambda = b = 1$).

3.6 Distribuição Nakagami-m

Introduzida por NAKAGAMI (1960), a distribuição NK é um modelo flexível para dados de confiabilidade devido à sua capacidade de acomodar diferentes formas da função de risco,

tais como crescente, decrescente, forma de banheira e unimodal. Uma variável aleatória com distribuição NK possui a seguinte função de densidade de probabilidades

$$f(t; a, b) = \frac{2}{\gamma(a)} \left(\frac{a}{b}\right)^a t^{2a-1} \exp\left(-\frac{a}{b}t^2\right), \quad t > 0$$

em que $a \geq 0.5$ e $b > 0$ são, respectivamente, os parâmetros de forma e escala, e $\gamma(\phi) = \int_0^\infty e^{-x} x^{\phi-1} dx$ é a função gama. Conforme apresentado em Ramos, Louzada e Ramos (2018), os momentos da NK são dados por

$$E(T^r | a, b) = \frac{\gamma(a+r/2)}{\gamma(a)} \left(\frac{b}{a}\right)^{r/2}$$

para $r \in \mathbb{N}$. Assim, a média e variância podem ser expressas como

$$\begin{aligned} E(T) &= \frac{\gamma(a+1/2)}{\gamma(a)} \left(\frac{b}{a}\right)^{1/2}, \\ \text{Var}(T) &= b \left(1 - \left(\frac{\gamma(a+1/2)}{\gamma(a)}\right)^2\right). \end{aligned}$$

A função de distribuição acumulada é dada por

$$F(t; a, b) = \int_0^{\frac{a}{b}t^2} \frac{1}{\gamma(a)} w^{a-1} e^{-w} dw = \frac{\gamma[a, \frac{a}{b}t^2]}{\gamma(a)},$$

em que $\gamma[u, v] = \int_0^v w^{u-1} e^{-w} dw$ é a função gama incompleta inferior. A função de sobrevivência é

$$S(t; a, b) = 1 - F(t; a, b) = \frac{\gamma[a, \frac{a}{b}t^2]}{\gamma(a)},$$

sendo que $\gamma[u, v] = \int_v^\infty w^{u-1} e^{-w} dw$ é a função gama incompleta superior, e a função de risco, dada por $h(t; a, b) = f(t; a, b)/S(t; a, b)$, é

$$h(t; a, b) = \frac{2}{\gamma[a, \frac{a}{b}t^2]} \left(\frac{a}{b}\right)^a t^{2a-1} \exp\left(-\frac{a}{b}t^2\right).$$

Algumas distribuições são casos especiais da NK, tais como Rayleigh ($a = 1$) e half-normal ($a = 1/2$). Além disso, esse modelo também está relacionado a distribuição gama. Seja $Y \sim \text{gama}(\mu, \Omega)$, então $T = \sqrt{Y}$ possui distribuição NK distribuição com $a = \mu$ e $b = \mu\Omega$. Dessa forma, devido à essa relação, o parâmetro a também pode assumir valores em $0 < a < 0.5$.

3.7 Estudo de Simulação

3.7.1 Especificações do Estudo

Com a finalidade de avaliar a eficiência da estimação por meio do método de máxima verossimilhança, realizamos um estudo de simulação com 500 réplicas (R) para um cenário

específico com diferentes tamanhos de amostra $n = 250, 500$ e 1000 . Baseando-se no estudo de simulação para modelos conjuntos proposto por Rizopoulos, Molenberghs e Lesaffre (2017), para cada sujeito $i = 1, \dots, n$, definimos o modelo longitudinal como

$$\begin{aligned} y_i(t) &= m_i(t) + \varepsilon_i(t) \\ &= \beta_1 Trt0_i + \beta_2 Trt1_i + \beta_3 \{t \times Trt0_i\} + \beta_4 \{t \times Trt1_i\} + b_{i0} + b_{i1}t + \varepsilon_i(t), \end{aligned}$$

em que $Trt0$ e $Trt1$ são variáveis dummy representando dois tratamentos, $\varepsilon_i(t) \sim N(0, \sigma^2)$ e $\mathbf{b}_i \sim N(\mathbf{0}, \mathbf{D})$. Para o modelo de sobrevivência, consideramos

$$h_i(t) = h_0(t) \exp\{\gamma Trt1_i + \alpha m_i(t)\}$$

sendo $h_0(t)$ a função de risco basal pertencente às diferentes distribuições do estudo: gama generalizada, extensões da distribuição Weibull e Nakagami- m .

Assumimos que os indivíduos foram avaliados longitudinalmente por um período de 20 anos e avaliamos as medidas longitudinais no início e ao longo de 16 períodos de tempo sorteados aleatoriamente. A distribuição da censura foi assumida como dependente do tratamento, ou seja, para o grupo $Trt0$ definimos os tempos de censura como seguindo distribuição Uniforme no intervalo $(0, 50)$, e para o grupo $Trt1$ como Uniforme no intervalo $(0, 60)$, resultando em uma censura média nos conjuntos de dados simulados em torno de 22,7%.

Para todas as amostras simuladas, os seguintes parâmetros foram definidos:

- Modelo longitudinal:

Efeitos fixos: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$ e $\beta_4 = 0,3718$; Matriz de covariâncias dos efeitos aleatórios: $\mathbf{D} = \begin{bmatrix} 0,8282 & -0,0196 \\ -0,0196 & 0,0328 \end{bmatrix}$; Desvio padrão do erro: $\sigma = 0,8902$.

- Modelo de sobrevivência:

Efeitos fixos: $\gamma = 0,4687$; Componente longitudinal: $\alpha = 0,4672$;

Função de risco basal: valores definidos especificamente para cada família de distribuição.

Para avaliar a eficiência dos estimadores, calculamos as médias, erro padrão, viés relativo (RBIAS) e raiz do erro quadrático médio (RMSE) das estimativas da seguinte maneira

$$\begin{aligned} \hat{\theta} &= \frac{1}{R} \sum_{r=1}^R \hat{\theta}_r, \quad s.e.(\hat{\theta}) = \sqrt{\frac{1}{R(R-1)} \sum_{r=1}^R (\hat{\theta}_r - \hat{\theta})^2}, \\ RBIAS(\hat{\theta}) &= \frac{\hat{\theta} - \theta}{\theta}, \quad RMSE(\hat{\theta}) = \sqrt{\frac{1}{R} \sum_{r=1}^R (\hat{\theta}_r - \theta)^2}, \end{aligned}$$

sendo θ o verdadeiro valor do parâmetro de interesse e R igual a 500. Também, calculamos a probabilidade de cobertura, cujos intervalos de 95% de confiança dos parâmetros foram calculados a partir de $B = 100$ réplicas de *bootstrap*, e exibimos em gráficos a função de densidade das estimativas dos parâmetros da função de risco basal para cada família de distribuição analisada.

Com relação aos valores iniciais do algoritmo de maximização da função de verossimilhança, consideramos as estimativas obtidas pelos modelos longitudinais e de sobrevivência ajustados separadamente aos dados e as estimativas das funções de distribuição ajustada aos dados sem considerar covariáveis, isto é, utilizando apenas o tempo e censura. Nesse caso, as estimativas do modelos longitudinal e de sobrevivência foram utilizadas nos valores iniciais dos efeitos fixos, matriz de covariâncias dos efeitos aleatórios e desvio padrão do erro, enquanto que as estimativas das funções de distribuição analisadas foram utilizadas na função de risco basal. Além disso, o valor inicial da componente longitudinal foi inicialmente definida como igual a um e a seguinte reparametrização foi considerada para a MOW: $a^* = (1/a)^b$.

No estudo para avaliar a discriminação de modelos via família GG e critérios de discriminação (AIC e BIC), simulamos 1000 réplicas para tamanhos de amostra iguais a $n = 250, 500$ e 1000. Foram gerados dados do modelo longitudinal com as seguintes distribuições:

- Gama: $f(t) = \frac{1}{b^a \Gamma(a)} t^{a-1} \exp(-t/b)$, com $a = 0.5128^{-2}$, $b = (0.5128^2/0.0214)$;
- Log-normal: $f(t) = \frac{1}{\sqrt{2\pi}bt} \exp\{-\{\log(t)-a\}^2/2b^2\}$, com $a = \log(1/0.0214) - 0.56$, $b = 0.5128$;
- Weibull: $f(t) = (a/b)(t/b)^{a-1} \exp\{-(t/b)^a\}$, com $a = 1/0.5128$, $b = 1/0.0214$;

e utilizados os dois critérios citados. Especificamente, para o teste de razão de verossimilhanças, avaliamos a taxa rejeição da hipótese nula considerando um nível de significância $\alpha = 0,05$ e para os critérios AIC e BIC a taxa de seleção quando o critério do respectivo modelo foi estritamente menor do que dos outros dois.

Vale destacar que para a análise dos resultados, em ambos os estudos de simulação, foram avaliados apenas os modelos que apresentaram convergência. Também, é importante mencionar que os parâmetros da matriz de covariâncias dos efeitos aleatórios são apresentados de acordo com a parametrização log-Cholesky (PINHEIRO; BATES, 1996) e, na implementação da aproximação da verossimilhança, consideramos 15 pontos na quadratura de Gauss-Kronrod, usada no cálculo da função de sobrevivência conjunta, e 5 pontos para a quadratura de Gauss-Hermite. Todo o procedimento de simulação foi desenvolvido no *software* R (R Core Team, 2018) e, devido à necessidade de computação intensiva para a maximização da log-verossimilhança do modelo conjunto, bem como as estimativas de intervalos de confiança via *bootstrap*, foram utilizados os recursos computacionais do cluster Euler, do Centro de Ciências Matemáticas Aplicadas à Indústria (CeMEAI).

3.7.2 Estudo de Simulação para as Distribuições Propostas

As seções a seguir apresentam os estudos de simulação realizados com a finalidade de verificar a eficiência dos estimadores dos parâmetros das distribuições propostas de acordo com especificação sugerida.

3.7.2.1 Distribuição Gama Generalizada

Antes de analisarmos os resultados da simulação em detalhes, é importante analisar a taxa de convergência das simulações realizadas de acordo com as duas parametrizações da distribuição gama generalizada. Em geral, observamos que, para os tamanhos de amostra de 250, 500 e 1000 houve uma taxa de convergência de 91,6, 95,0 e 95,6% para a parametrização de Stacy e 93,4, 93,4 e 94,8% para a de Prentice, respectivamente. Dessa forma, à medida que o tamanho da amostra aumenta, a taxa de convergência também aumenta. Além disso, a parametrização de Stacy apresenta uma taxa de convergência ligeiramente superior para amostra de 500 e 1000 observações.

A partir das Tabelas 1 e 2 observamos que, com exceção dos parâmetros da função de risco, houve boa aproximação nos valores dos demais parâmetros independentemente da parametrização considerada. Particularmente, os maiores valores do viés relativo e da raiz do erro médio quadrático, e também menores valores da probabilidade de cobertura, foram observados nos parâmetros da função de risco basal, indicando que existe alta variabilidade nas estimativas de máxima verossimilhança desses parâmetros. Além disso, também observamos baixa probabilidade de cobertura para o componente da matriz de covariâncias do modelo de efeitos mistos, D_{22} . Ao compararmos as duas parametrizações, constatamos que a parametrização de Stacy tende a apresentar aproximações ligeiramente melhores para os parâmetros das covariáveis, enquanto que a de Prentice apresenta melhores aproximações para os parâmetros da função de risco. Nesse caso, destacamos baixa aproximação nas estimativas dos parâmetros $\log(k)$ e q da parametrização de Stacy e Prentice, respectivamente. Em geral, esta última parametrização apresenta maiores probabilidades de cobertura independentemente dos parâmetros. Além disso, conforme observado nas Figuras 3 e 4, alguns parâmetros da função de risco apresentam função de densidade bimodal, confirmando a falta de identificabilidade da família gama generalizada.

3.7.2.2 Distribuição Weibull Exponenciada

Os resultados da simulação para a distribuição Weibull exponenciada estão apresentados na Tabela 3 e Figura 5. Para essa família, observamos alta taxa de convergência das simulações para os diferentes tamanhos de amostra, com valores de 99,0%, 98,8% e 97,6% para amostras de tamanho 250, 500 e 1000, respectivamente. Analisando os diferentes cenários, verificamos que os valores da raiz do erro médio quadrático se aproximaram de zero na maior parte dos casos, destacando valores maiores para $\log(b)$ e $\log(\lambda)$, que se sobressaíram em relação aos demais. Da mesma forma, notamos que, com exceção para $\log(b)$, o viés relativo se aproxima de zero

Tabela 1 – Resultados da simulação para a distribuição GG considerando a parametrização de Stacy. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8861	0,0043	-0,0018	0,0925	0,9410
	β_2	-0,6018	0,0042	-0,0052	0,0909	0,9389
	β_3	0,3423	0,0009	-0,0023	0,0194	0,9476
	β_4	0,3719	0,0009	0,0003	0,0189	0,9214
	$\log(\sigma)$	-0,1186	0,0008	0,0198	0,0173	0,9258
	γ	0,4368	0,0089	-0,0681	0,1929	0,9301
	α	0,4567	0,0020	-0,0224	0,0445	0,9454
	$\log(a)$	3,2926	0,0841	-0,1515	1,8911	0,9214
	$\log(b)$	1,7244	0,0939	1,5177	2,2613	0,9083
	$\log(k)$	-1,0306	0,0992	13,3718	2,3273	0,9192
	D_{11}	-0,1023	0,0029	0,0859	0,0636	0,9148
	D_{12}	-0,0195	0,0008	-0,0956	0,0168	0,9170
	D_{22}	-1,7433	0,0029	0,0161	0,0679	0,8559
	500	β_1	2,8847	0,0032	-0,0023	0,0702
β_2		-0,6034	0,0029	-0,0026	0,0640	0,9558
β_3		0,3431	0,0007	-0,0001	0,0148	0,9368
β_4		0,3716	0,0006	-0,0006	0,0133	0,9368
$\log(\sigma)$		-0,1163	0,0006	-0,0002	0,0123	0,9326
γ		0,4305	0,0063	-0,0815	0,1422	0,8947
α		0,4547	0,0014	-0,0268	0,0335	0,9221
$\log(a)$		3,5978	0,0424	-0,0729	0,9656	0,9242
$\log(b)$		1,7588	0,0896	1,5679	2,2259	0,8968
$\log(k)$		-1,1049	0,0927	14,4074	2,2672	0,9011
D_{11}		-0,0972	0,0020	0,0312	0,0435	0,8989
D_{12}		-0,0208	0,0005	-0,0321	0,0114	0,9200
D_{22}		-1,7333	0,0019	0,0102	0,0459	0,8947
1000		β_1	2,8926	0,0022	0,0005	0,0489
	β_2	-0,6025	0,0022	-0,0039	0,0477	0,9310
	β_3	0,3438	0,0005	0,0022	0,0106	0,9184
	β_4	0,3730	0,0004	0,0032	0,0094	0,9331
	$\log(\sigma)$	-0,1158	0,0004	-0,0045	0,0085	0,9351
	γ	0,4460	0,0043	-0,0485	0,0977	0,9414
	α	0,4572	0,0011	-0,0214	0,0254	0,9079
	$\log(a)$	3,7738	0,0150	-0,0275	0,3452	0,9477
	$\log(b)$	1,9088	0,1218	1,7870	2,9277	0,9059
	$\log(k)$	-1,2809	0,1229	16,8620	2,9445	0,9100
	D_{11}	-0,0978	0,0013	0,0376	0,0290	0,9393
	D_{12}	-0,0208	0,0004	-0,0320	0,0082	0,9268
	D_{22}	-1,7294	0,0014	0,0079	0,0334	0,8912

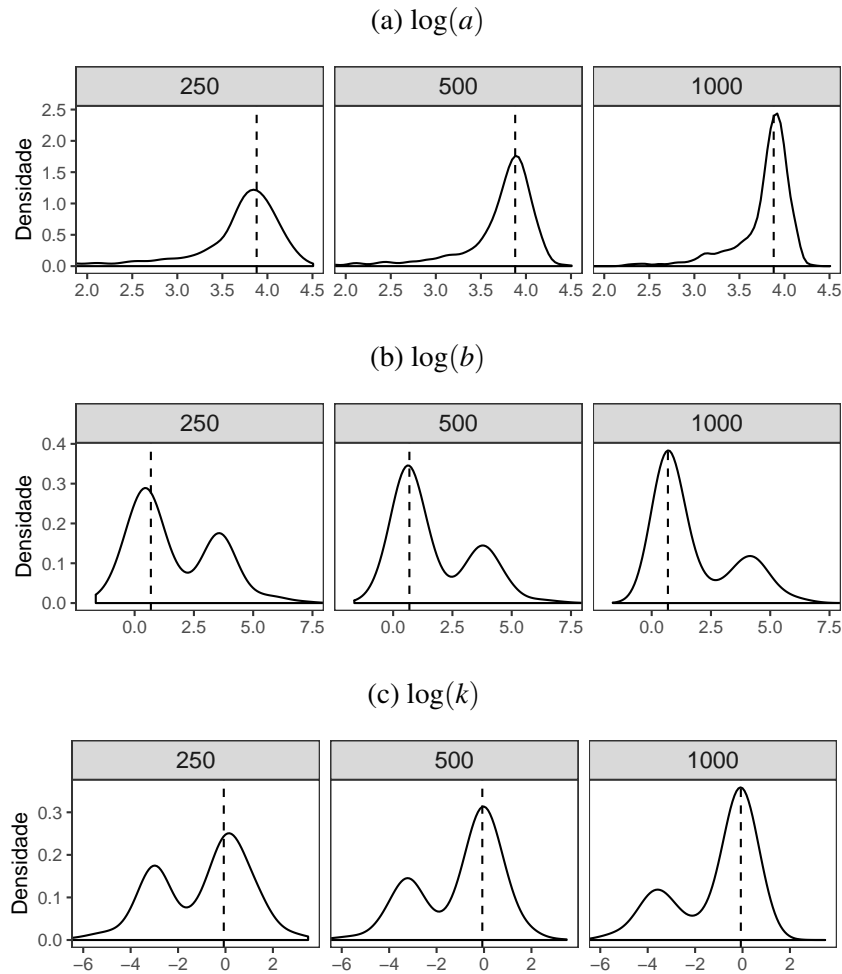
Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\log(a) = 3,8806$, $\log(b) = 0,6849$ e $\log(k) = -0,0717$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Tabela 2 – Resultados da simulação para a distribuição GG considerando a parametrização de Prentice. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8880	0,0043	-0,0011	0,0933	0,9443
	β_2	-0,5979	0,0043	-0,0116	0,0922	0,9400
	β_3	0,3393	0,0016	-0,0111	0,0351	0,9550
	β_4	0,3608	0,0035	-0,0297	0,0768	0,9636
	$\log(\sigma)$	-0,1026	0,0035	-0,1175	0,0776	0,9615
	γ	0,6161	0,0282	0,3146	0,6267	0,9272
	α	0,4555	0,0056	-0,0251	0,1220	0,9807
	μ	3,7424	0,0159	-0,0265	0,3587	0,9272
	$\log(\tau)$	-1,0818	0,0345	0,6665	0,8618	0,9036
	q	2,1035	0,1098	1,0294	2,5993	0,9036
	D_{11}	-0,1041	0,0030	0,1043	0,0653	0,9079
	D_{12}	-0,0203	0,0008	-0,0589	0,0179	0,9272
	D_{22}	-1,7498	0,0028	0,0198	0,0703	0,8223
500	β_1	2,8868	0,0033	-0,0016	0,0712	0,9143
	β_2	-0,6014	0,0030	-0,0057	0,0646	0,9422
	β_3	0,3432	0,0012	0,0004	0,0260	0,9507
	β_4	0,3691	0,0020	-0,0072	0,0442	0,9550
	$\log(\sigma)$	-0,1090	0,0023	-0,0628	0,0509	0,9443
	γ	0,5023	0,0166	0,0716	0,3607	0,9400
	α	0,4462	0,0038	-0,0449	0,0836	0,9443
	μ	3,8003	0,0086	-0,0115	0,1909	0,9208
	$\log(\tau)$	-1,1124	0,0340	0,7137	0,8679	0,9143
	q	2,2413	0,0992	1,1623	2,4567	0,9208
	D_{11}	-0,0962	0,0020	0,0207	0,0431	0,9229
	D_{12}	-0,0208	0,0005	-0,0336	0,0112	0,9550
	D_{22}	-1,7420	0,0019	0,0153	0,0491	0,8758
1000	β_1	2,8933	0,0023	0,0007	0,0504	0,9304
	β_2	-0,6026	0,0022	-0,0039	0,0480	0,9304
	β_3	0,3445	0,0005	0,0040	0,0114	0,9451
	β_4	0,3742	0,0005	0,0064	0,0104	0,9283
	$\log(\sigma)$	-0,1156	0,0004	-0,0058	0,0085	0,9367
	γ	0,4502	0,0045	-0,0396	0,1001	0,9346
	α	0,4598	0,0012	-0,0159	0,0272	0,9177
	μ	3,8150	0,0048	-0,0076	0,1086	0,9283
	$\log(\tau)$	-1,1214	0,0340	0,7275	0,8767	0,9177
	q	2,3176	0,1142	1,2360	2,7953	0,9283
	D_{11}	-0,0966	0,0014	0,0252	0,0304	0,9346
	D_{12}	-0,0210	0,0004	-0,0264	0,0081	0,9409
	D_{22}	-1,7390	0,0015	0,0136	0,0396	0,8291

Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\mu = 3,8444$, $\log(\tau) = -0,6491$ e $q = 1,0365$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Figura 3 – Funções de densidade obtidas para os parâmetros da distribuição GG considerando a parametrização de Stacy.



Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

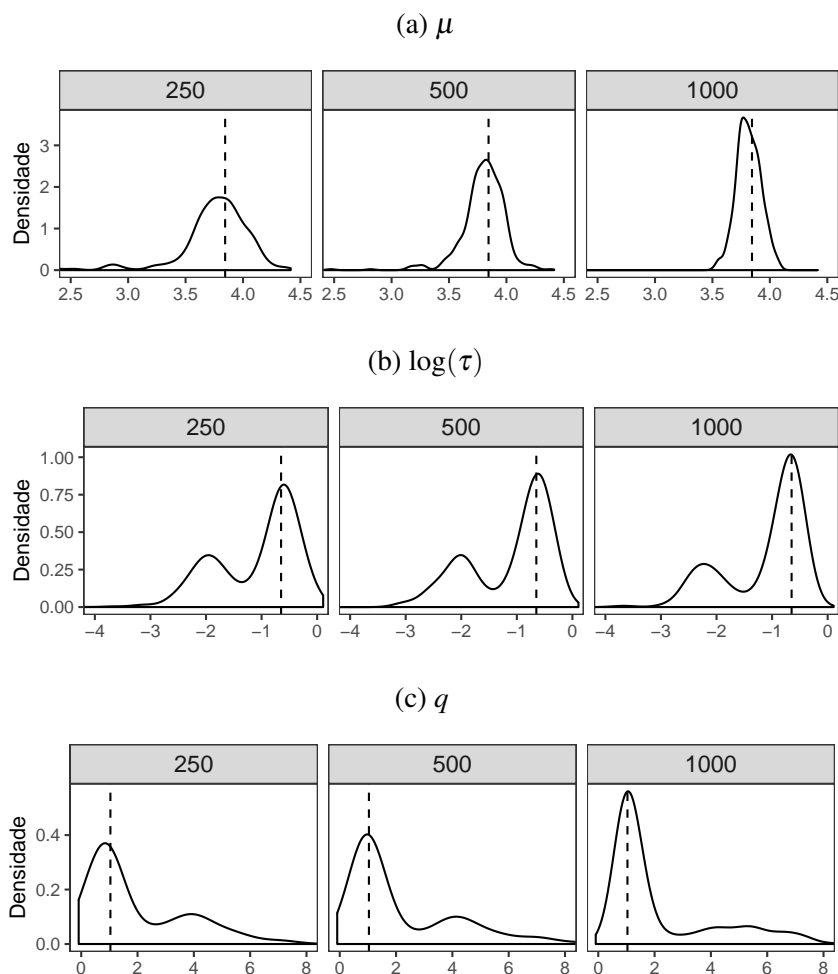
em todos os demais parâmetros, sendo menor à medida que n cresce. Nesse caso, vale ressaltar que, como o valor de λ simulado foi igual a um, não foi possível calcular o viés relativo desse parâmetro. Além disso, o viés relativo de $\log(b)$ se aproxima de um. Quanto às probabilidades de cobertura, notamos que tendem a ser menores para os parâmetros $\log(b)$ e $\log(\lambda)$, enquanto que, para os demais, os valores ficaram praticamente acima de 95%, com destaque para os maiores valores no maior tamanho de amostra ($n = 1000$). Por meio das funções de densidade das estimativas obtidas para os parâmetros da função de risco, constatamos que tanto $\log(b)$ quanto $\log(\lambda)$ apresentam funções de densidade bimodais, com suavização desse efeito para os maiores valores de amostra.

Tabela 3 – Resultados da simulação para a distribuição EW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8647	0,0087	-0,0092	0,1964	0,9677
	β_2	-0,5953	0,0076	-0,0159	0,1690	0,9617
	β_3	0,3414	0,0014	-0,0051	0,0313	0,9798
	β_4	0,3689	0,0013	-0,0079	0,0296	0,9738
	$\log(\sigma)$	-0,1176	0,0011	0,0108	0,0251	0,9718
	γ	0,3971	0,0114	-0,1528	0,2635	0,9395
	α	0,4610	0,0027	-0,0132	0,0602	0,9819
	$\log(a)$	3,5678	0,0385	-0,0770	0,9067	0,9476
	$\log(b)$	1,3187	0,0599	0,9314	1,4766	0,8851
	$\log(\lambda)$	-0,5001	0,0692	-	1,6197	0,8931
	D_{11}	-0,1065	0,0040	0,1303	0,0897	0,9617
	D_{12}	-0,0202	0,0011	-0,0615	0,0235	0,9778
	D_{22}	-1,7452	0,0042	0,0171	0,0988	0,9315
500	β_1	2,8748	0,0061	-0,0057	0,1360	0,9795
	β_2	-0,6023	0,0054	-0,0043	0,1189	0,9918
	β_3	0,3413	0,0013	-0,0051	0,0291	0,9877
	β_4	0,3699	0,0012	-0,0052	0,0273	0,9815
	$\log(\sigma)$	-0,1170	0,0011	0,0063	0,0251	0,9836
	γ	0,4145	0,0096	-0,1157	0,2177	0,9713
	α	0,4701	0,0029	0,0061	0,0646	0,9815
	$\log(a)$	3,7931	0,0173	-0,0188	0,3889	0,9733
	$\log(b)$	1,4275	0,0540	1,0907	1,4049	0,9076
	$\log(\lambda)$	-0,6743	0,0574	-	1,4347	0,9158
	D_{11}	-0,0996	0,0040	0,0564	0,0891	0,9815
	D_{12}	-0,0205	0,0017	-0,0465	0,0367	0,9795
	D_{22}	-1,7380	0,0036	0,0130	0,0830	0,9589
1000	β_1	2,8737	0,0095	-0,0061	0,2090	0,9835
	β_2	-0,6027	0,0102	-0,0036	0,2245	0,9794
	β_3	0,3405	0,0016	-0,0077	0,0343	0,9918
	β_4	0,3690	0,0020	-0,0074	0,0438	0,9876
	$\log(\sigma)$	-0,1158	0,0011	-0,0042	0,0249	0,9835
	γ	0,4290	0,0130	-0,0847	0,2890	0,9732
	α	0,4744	0,0047	0,0155	0,1041	0,9773
	$\log(a)$	3,8685	0,0165	0,0007	0,3627	0,9588
	$\log(b)$	1,5017	0,0533	1,1993	1,4311	0,9196
	$\log(\lambda)$	-0,7652	0,0549	-	1,4301	0,9175
	D_{11}	-0,0961	0,0044	0,0195	0,0960	0,9856
	D_{12}	-0,0209	0,0020	-0,0301	0,0445	0,9856
	D_{22}	-1,7111	0,0082	-0,0027	0,1797	0,9794

Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\log(a) = 3,8656$, $\log(b) = 0,6827$ e $\log(\lambda) = 0$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Figura 4 – Funções de densidade obtidas para os parâmetros da distribuição GG considerando a parametrização de Prentice.



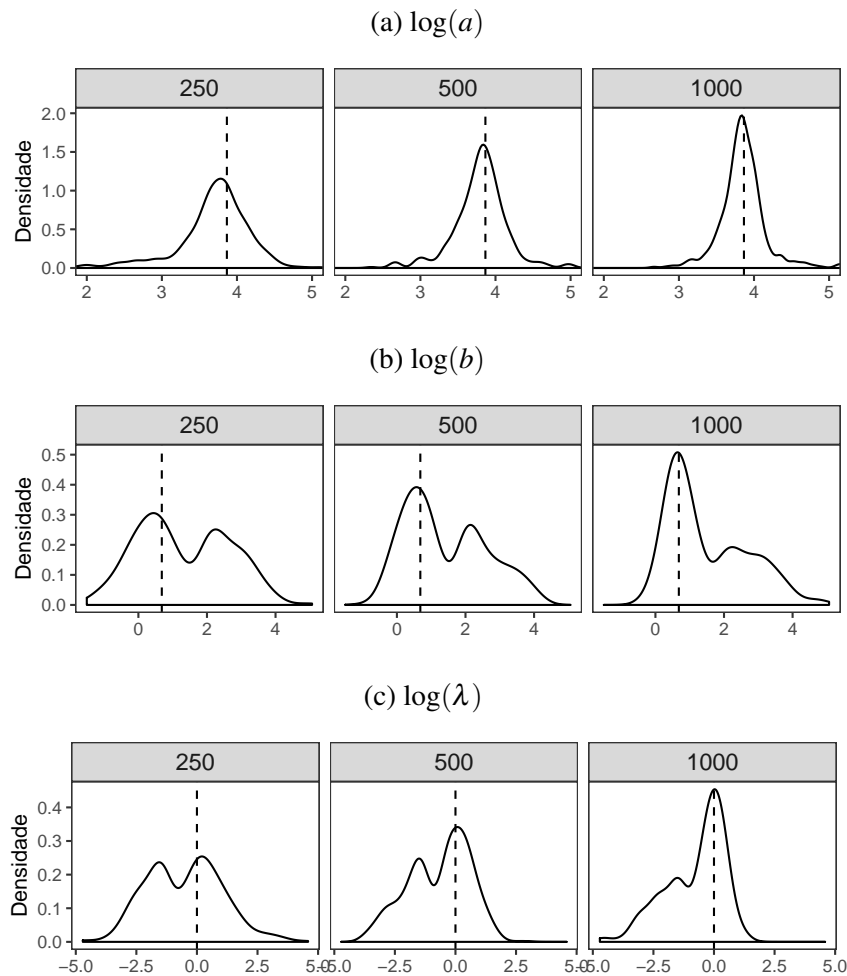
Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

3.7.2.3 Distribuição Weibull Generalizada

Os resultados do estudo de simulação feito para a distribuição Weibull generalizada são apresentados na Tabela 4 e Figura 6. Analisando os valores obtidos constatamos que, com exceção das medidas do parâmetro λ , para todos os cenários o viés relativo e raiz do erro médio quadrático dos parâmetros se aproximam de zero quando n cresce, sendo que para o referido parâmetro temos que ambas as medidas foram bastante altas. Nesse caso, existe uma suavização do efeito nas amostras de tamanhos maiores. Quanto às probabilidades de cobertura, observamos valores próximos a 95% para grande parte dos casos, sendo que D_{22} se destaca com valores bem abaixo do esperando, inclusive diminuindo com o tamanho da amostra. Por meio da Figura 6, temos que, ao contrário das distribuições anteriores, os parâmetros da Weibull generalizada apresentam funções de densidade unimodais, concentrando-se em torno do verdadeiro valor. Por fim, é importante citar que as taxas de convergência dos modelos ajustados aos dados foram

Figura 5 – Funções de densidade obtidas para os parâmetros da distribuição EW.



Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

inferiores aos outros modelos, sendo iguais a 81,2, 80,8 e 83,0% para n igual a 250, 500 e 1000.

3.7.2.4 Distribuição Marshall-Olkin-Weibull

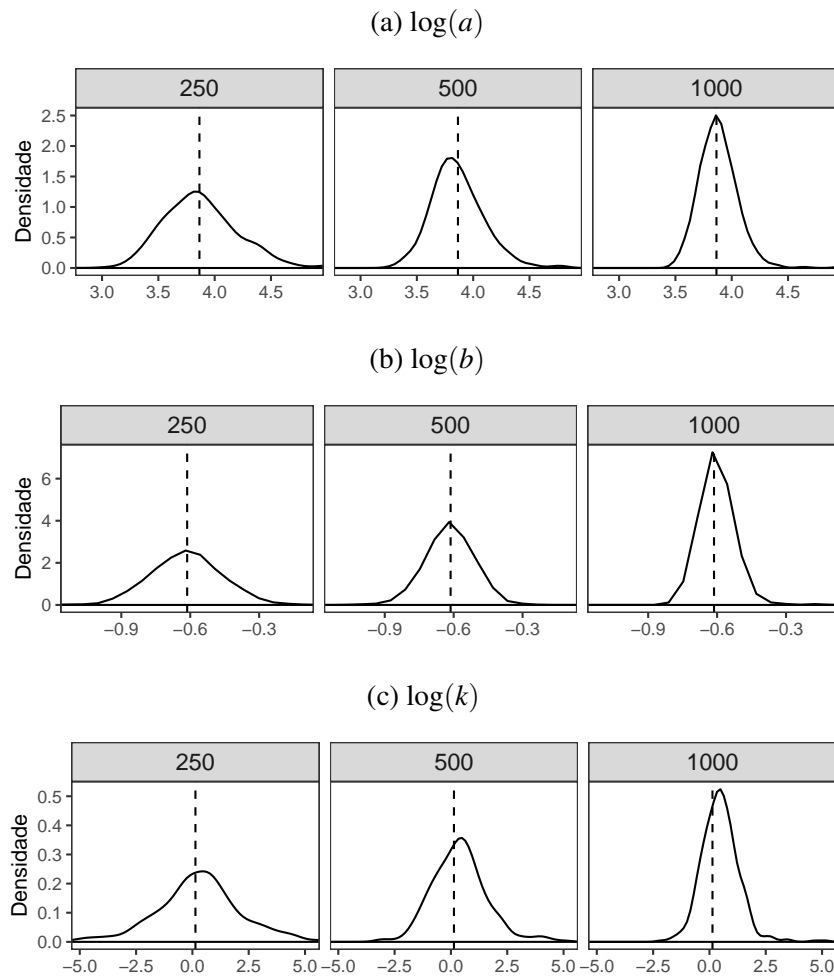
A Tabela 5 e Figura 7 apresentam os resultados para a distribuição Marshall-Olkin-Weibull. A partir dos modelos ajustados, observamos que existe alta taxa de convergência para o método de máxima verossimilhança, com valores iguais a 97,8, 99,4 e 99,4% para as amostras de 250, 500 e 1000 observações, respectivamente. Em geral, notamos boas estimativas para o viés relativo e raiz do erro médio quadrático, com valores próximos a zero na maior parte dos parâmetros analisados. Além disso, constatamos que os desvios em relação ao verdadeiro valor simulado são menores à medida que o tamanho da amostra cresce. Destacamos, também, que o parâmetro γ , bem como $\log(a^*)$ e $\log(\lambda)$, apresentaram os maiores vieses em comparação aos demais. Nesse caso, $\log(\lambda)$ e γ apresentam valores altos de viés relativo e raiz do erro médio quadrático principalmente para $n = 250$, decrescendo conforme n aumenta. No entanto, nota-se

Tabela 4 – Resultados da simulação para a distribuição GW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8907	0,0089	-0,0002	0,1793	0,9581
	β_2	-0,5894	0,0090	-0,0257	0,1813	0,9532
	β_3	0,3441	0,0026	0,0030	0,0514	0,9631
	β_4	0,3785	0,0027	0,0179	0,0555	0,9261
	$\log(\sigma)$	-0,1194	0,0052	0,0266	0,1042	0,9557
	γ	0,4173	0,0302	-0,1096	0,6109	0,9606
	α	0,4551	0,0066	-0,0258	0,1326	0,9655
	$\log(a)$	3,9683	0,0484	0,0268	0,9786	0,9483
	$\log(b)$	-0,5167	0,0338	-0,1576	0,6875	0,9409
	λ	1,4865	0,3694	10,0195	7,5558	0,9458
	D_{11}	-0,1006	0,0041	0,0669	0,0818	0,9360
	D_{12}	-0,0172	0,0023	-0,2020	0,0467	0,9507
	D_{22}	-1,7749	0,0067	0,0344	0,1478	0,7931
	500	β_1	2,8811	0,0182	-0,0035	0,3654
β_2		-0,5884	0,0181	-0,0272	0,3647	0,9455
β_3		0,3416	0,0041	-0,0044	0,0820	0,9455
β_4		0,3749	0,0044	0,0082	0,0887	0,9356
$\log(\sigma)$		-0,1172	0,0023	0,0075	0,0457	0,9307
γ		0,3224	0,0453	-0,3122	0,9217	0,9233
α		0,4195	0,0147	-0,1020	0,2994	0,9282
$\log(a)$		3,7757	0,0606	-0,0230	1,2192	0,9455
$\log(b)$		-0,5050	0,0747	-0,1767	1,5037	0,9381
λ		1,3497	0,3438	9,0050	7,0081	0,9035
D_{11}		-0,0801	0,0128	-0,1500	0,2565	0,9282
D_{12}		-0,0182	0,0042	-0,1549	0,0849	0,9480
D_{22}		-1,7627	0,0082	0,0273	0,1717	0,7772
1000		β_1	2,9002	0,0074	0,0031	0,1514
	β_2	-0,6024	0,0049	-0,0041	0,0988	0,9446
	β_3	0,3486	0,0016	0,0160	0,0334	0,9470
	β_4	0,3790	0,0019	0,0192	0,0389	0,9036
	$\log(\sigma)$	-0,1159	0,0007	-0,0036	0,0136	0,9518
	γ	0,4214	0,0187	-0,1009	0,3825	0,9349
	α	0,4490	0,0060	-0,0389	0,1241	0,9349
	$\log(a)$	3,8700	0,0249	0,0014	0,5062	0,9687
	$\log(b)$	-0,5658	0,0204	-0,0777	0,4177	0,9470
	λ	0,9529	0,2007	6,0638	4,1655	0,9084
	D_{11}	-0,0926	0,0032	-0,0171	0,0642	0,9542
	D_{12}	-0,0209	0,0008	-0,0307	0,0164	0,9446
	D_{22}	-1,7628	0,0058	0,0274	0,1272	0,6867

Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\log(a) = 3,8645$, $\log(b) = -0,6134$ e $\lambda = 0,1349$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Figura 6 – Funções de densidade obtidas para os parâmetros da distribuição GW.



Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

que para o maior tamanho de amostra o viés relativo é quase nulo em todos os parâmetros, indicando maior precisão das estimativas. Quanto às probabilidades de cobertura, temos que os valores ficaram acima de 90% para a maioria dos os casos e, segundo as funções de densidade obtidas, todos os parâmetros da função de risco proposta têm distribuição unimodal, com maior concentração em torno do verdadeiro valor à medida que o tamanho da amostra cresce.

3.7.2.5 Distribuição Potência Weibull Generalizada

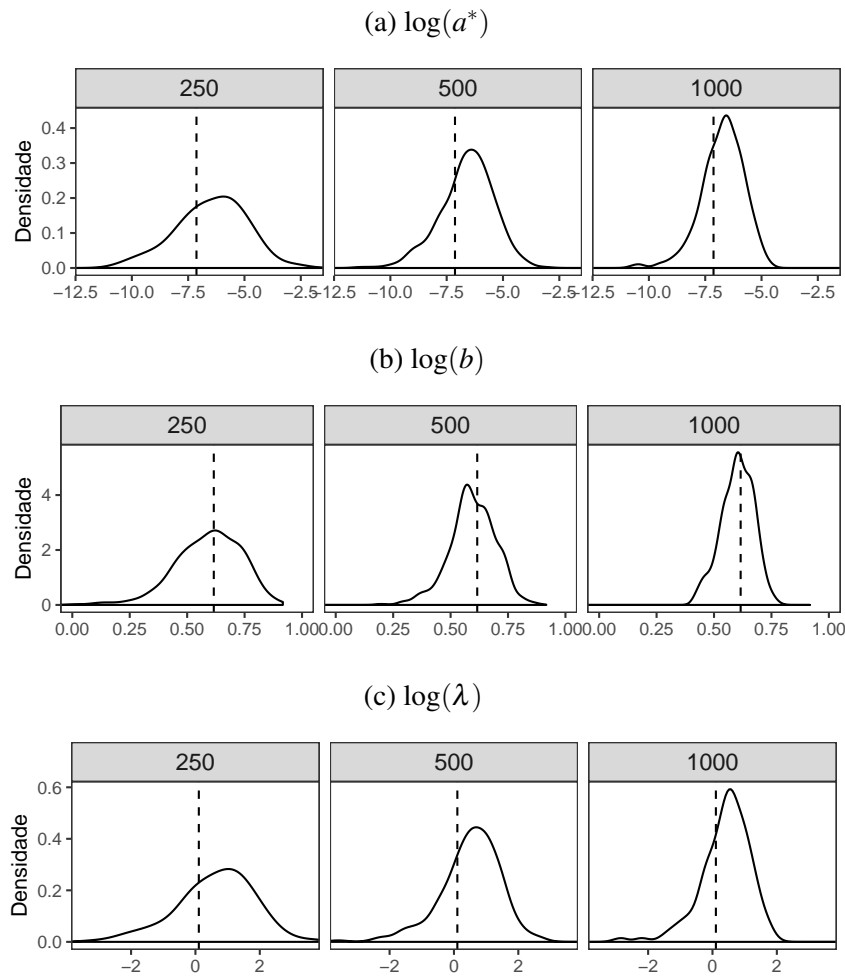
Apresentamos na Tabela 6 os resultados para o estudo de simulação da família de potências Weibull generalizada. Notamos que, com exceção dos parâmetros pertencentes à função de risco, $\log(a)$ e $\log(\gamma)$, que apresentaram valores do viés relativo e raiz do erro médio quadrático bem superiores aos demais, em todos os outros as estimativas de máxima verossimilhança foram bastante satisfatórias. Ainda, independentemente do tamanho da amostra, os valores de $\log(a)$ e $\log(\gamma)$ diferiram muito em relação ao verdadeiro valor, com raiz do erro

Tabela 5 – Resultados da simulação para a distribuição MOW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8835	0,0042	-0,0027	0,0924	0,9325
	β_2	-0,6012	0,0041	-0,0062	0,0906	0,9387
	β_3	0,3428	0,0009	-0,0009	0,0192	0,9346
	β_4	0,3709	0,0009	-0,0025	0,0189	0,9243
	$\log(\sigma)$	-0,1185	0,0008	0,0192	0,0171	0,9202
	γ	0,9930	0,5565	1,1187	12,3041	0,9223
	α	0,4592	0,0020	-0,0172	0,0456	0,9448
	$\log(a^*)$	-8,5641	0,2488	0,2010	5,6798	0,9039
	$\log(b)$	0,5756	0,0097	-0,0655	0,2176	0,9346
	$\log(\lambda)$	-1,3945	0,2380	-14,2145	5,4670	0,8855
	D_{11}	-0,1043	0,0028	0,1064	0,0634	0,9121
	D_{12}	-0,0200	0,0007	-0,0711	0,0164	0,9100
	D_{22}	-1,7359	0,0027	0,0117	0,0632	0,8896
500	β_1	2,8841	0,0031	-0,0025	0,0702	0,9195
	β_2	-0,6042	0,0029	-0,0011	0,0642	0,9577
	β_3	0,3432	0,0007	0,0003	0,0148	0,9356
	β_4	0,3708	0,0006	-0,0027	0,0136	0,9316
	$\log(\sigma)$	-0,1164	0,0006	0,0004	0,0123	0,9256
	γ	0,4355	0,0061	-0,0709	0,1405	0,9115
	α	0,4558	0,0015	-0,0244	0,0349	0,9336
	$\log(a^*)$	-7,4257	0,1603	0,0413	3,5818	0,9135
	$\log(b)$	0,5953	0,0045	-0,0335	0,1018	0,9296
	$\log(\lambda)$	-0,2943	0,1517	-3,7891	3,4031	0,9074
	D_{11}	-0,0959	0,0019	0,0179	0,0434	0,9095
	D_{12}	-0,0210	0,0005	-0,0260	0,0112	0,9316
	D_{22}	-1,7302	0,0019	0,0084	0,0448	0,9135
1000	β_1	2,8931	0,0022	0,0006	0,0487	0,9135
	β_2	-0,6028	0,0021	-0,0034	0,0473	0,9235
	β_3	0,3438	0,0005	0,0021	0,0103	0,9437
	β_4	0,3726	0,0004	0,0021	0,0090	0,9356
	$\log(\sigma)$	-0,1158	0,0004	-0,0042	0,0083	0,9276
	γ	0,4442	0,0042	-0,0523	0,0962	0,9477
	α	0,4572	0,0011	-0,0214	0,0256	0,9195
	$\log(a^*)$	-7,0442	0,0984	-0,0122	2,1930	0,9014
	$\log(b)$	0,6018	0,0032	-0,0230	0,0720	0,9396
	$\log(\lambda)$	0,1026	0,0932	-0,0274	2,0759	0,8893
	D_{11}	-0,0978	0,0013	0,0372	0,0286	0,9376
	D_{12}	-0,0205	0,0004	-0,0469	0,0082	0,9235
	D_{22}	-1,7271	0,0013	0,0066	0,0316	0,9135

Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\log(a^*) = -7,1308$, $\log(b) = 0,6159$ e $\log(\lambda) = 0,1055$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Figura 7 – Funções de densidade obtidas para os parâmetros da distribuição MOW.



Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

médio quadrático próximo a cinco para o parâmetro $\log(a)$ e viés relativo próximo a 400% para $\log(\gamma)$. Analisando as probabilidades de cobertura, constatamos que ficaram próximas do esperado, com valores maior que 90% em grande parte dos caso. Por meio da Figura 8, que apresenta as funções de densidade observadas nas estimativas das simulações realizadas, temos que o modelo proposto apresenta falta de identificabilidade com função de densidade bimodal tanto para $\log(a)$ quanto para $\log(\gamma)$.

3.7.2.6 Distribuição Nakagami- m

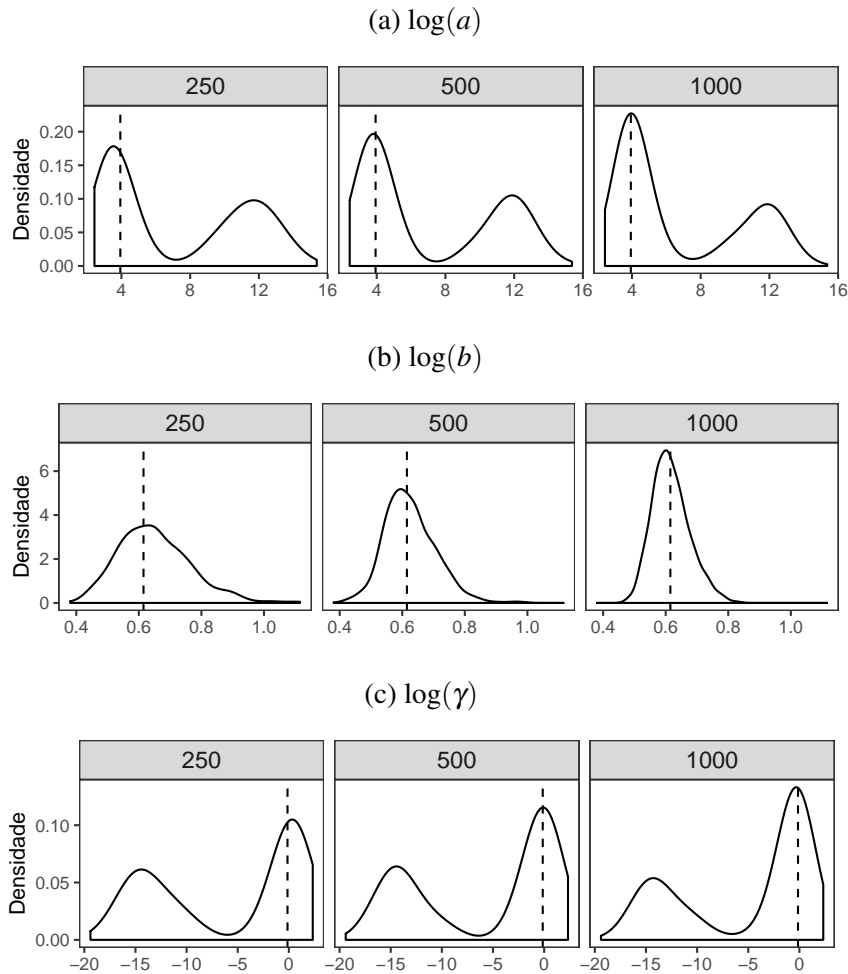
Analisando a Tabela 7, observamos que, em geral, o viés relativo e a raiz do erro quadrático dos parâmetros estimados se aproximam de zero nos três cenários analisados, com aproximação ainda maior à medida que o tamanho da amostra simulada cresce. Apesar de apresentarem valores satisfatórios, ressaltamos o valor para o viés relativo de $\log(a)$, que ficou entre 20 e 30%, e raiz do erro médio quadrático de $\log(b)$, que se sobressaiu em relação aos

Tabela 6 – Resultados da simulação para a distribuição GPW. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8857	0,0042	-0,0019	0,0926	0,9376
	β_2	-0,6006	0,0041	-0,0072	0,0918	0,9356
	β_3	0,3422	0,0009	-0,0027	0,0194	0,9457
	β_4	0,3712	0,0009	-0,0015	0,0190	0,9276
	$\log(\sigma)$	-0,1184	0,0008	0,0184	0,0175	0,9195
	γ	0,4330	0,0084	-0,0762	0,1902	0,9215
	α	0,4559	0,0019	-0,0242	0,0430	0,9537
	$\log(a)$	7,1405	0,1830	0,8148	5,1854	0,8934
	$\log(b)$	0,6504	0,0052	0,0577	0,1206	0,9115
	$\log(\gamma)$	-5,9724	0,3199	43,0848	9,2099	0,8753
	D_{11}	-0,1044	0,0029	0,1078	0,0644	0,9135
	D_{12}	-0,0194	0,0007	-0,0971	0,0167	0,9195
	D_{22}	-1,7385	0,0027	0,0133	0,0653	0,8753
	500	β_1	2,8844	0,0032	-0,0024	0,0705
β_2		-0,6045	0,0029	-0,0007	0,0646	0,9557
β_3		0,3430	0,0007	-0,0002	0,0150	0,9316
β_4		0,3709	0,0006	-0,0023	0,0133	0,9296
$\log(\sigma)$		-0,1165	0,0006	0,0014	0,0125	0,9195
γ		0,4310	0,0060	-0,0805	0,1389	0,8974
α		0,4543	0,0014	-0,0276	0,0340	0,9014
$\log(a)$		7,0693	0,1773	0,7967	5,0411	0,9195
$\log(b)$		0,6250	0,0036	0,0165	0,0819	0,9356
$\log(\gamma)$		-5,7829	0,3090	41,6857	8,9025	0,9034
D_{11}		-0,0969	0,0019	0,0285	0,0435	0,9014
D_{12}		-0,0206	0,0005	-0,0448	0,0115	0,9256
D_{22}		-1,7312	0,0019	0,0090	0,0454	0,9034
1000		β_1	2,8925	0,0022	0,0004	0,0487
	β_2	-0,6028	0,0021	-0,0035	0,0470	0,9333
	β_3	0,3440	0,0005	0,0025	0,0106	0,9232
	β_4	0,3725	0,0004	0,0018	0,0093	0,9253
	$\log(\sigma)$	-0,1159	0,0004	-0,0034	0,0085	0,9253
	γ	0,4419	0,0043	-0,0571	0,0993	0,9253
	α	0,4562	0,0011	-0,0234	0,0262	0,9030
	$\log(a)$	6,6950	0,1651	0,7016	4,5927	0,8990
	$\log(b)$	0,6168	0,0027	0,0031	0,0591	0,9455
	$\log(\gamma)$	-5,1142	0,2898	36,7498	8,1411	0,8788
	D_{11}	-0,0974	0,0013	0,0331	0,0289	0,9414
	D_{12}	-0,0207	0,0004	-0,0376	0,0081	0,9253
	D_{22}	-1,7270	0,0014	0,0065	0,0322	0,8929

Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\log(a) = 3,9346$, $\log(b) = 0,6149$ e $\log(\gamma) = -0,1354$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Figura 8 – Funções de densidade obtidas para os parâmetros da distribuição GPW.



Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

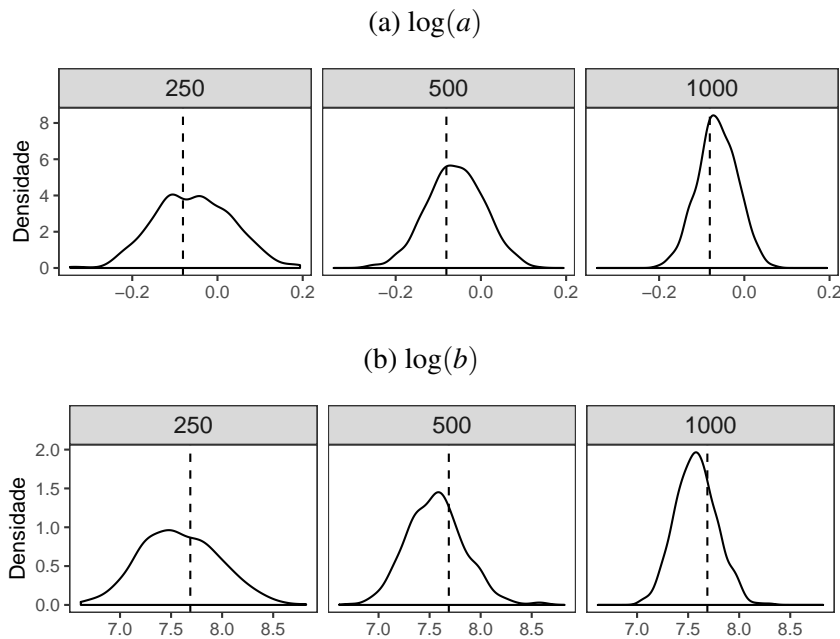
demais. Destacamos, também, que para n igual a 250 e 500 os valores da raiz do erro médio quadrático do parâmetro γ , relacionado à covariável no modelo de sobrevivência, foram iguais a 15 e 30%, respectivamente. Ao observarmos as probabilidades de cobertura, calculadas a partir de intervalos de 95% de confiança obtidos via *bootstrap*, temos que na maioria dos casos as probabilidades são superiores à 90%. A Figura 9 exibe as funções de densidade dos parâmetros da distribuição Nakagami- m . Como podemos observar, quanto menor o tamanho da amostra, maior a variabilidade das estimativas. Também, as distribuições de ambos os parâmetros são unimodais, sendo que a moda de $\log(a)$ se aproxima mais do verdadeiro valor do que a moda de $\log(b)$. É importante citar que todas os modelos simulados apresentaram convergência no método de estimação usado.

Tabela 7 – Resultados da simulação para a distribuição NK. As estatísticas calculadas são a média, erro padrão, viés relativo (RBIAS), raiz do erro quadrático médio (RMSE) e probabilidade de cobertura (CP).

n	Parâmetro	Média	Erro padrão	RBIAS	RMSE	CP 95%
250	β_1	2,8861	0,0041	-0,0018	0,0922	0,9480
	β_2	-0,6002	0,0041	-0,0079	0,0920	0,9340
	β_3	0,3420	0,0009	-0,0033	0,0194	0,9520
	β_4	0,3715	0,0009	-0,0007	0,0192	0,9240
	$\log(\sigma)$	-0,1185	0,0008	0,0185	0,0173	0,9220
	γ	0,4443	0,0084	-0,0520	0,1899	0,9340
	α	0,4586	0,0019	-0,0185	0,0439	0,9380
	$\log(a)$	-0,0560	0,0040	-0,3092	0,0929	0,9100
	$\log(b)$	7,5881	0,0170	-0,0130	0,3934	0,9340
	D_{11}	-0,1040	0,0028	0,1033	0,0639	0,9100
	D_{12}	-0,0196	0,0008	-0,0893	0,0169	0,9180
	D_{22}	-1,7418	0,0027	0,0152	0,0666	0,8640
500	β_1	2,8851	0,0031	-0,0021	0,0704	0,9200
	β_2	-0,6041	0,0029	-0,0013	0,0651	0,9580
	β_3	0,3431	0,0007	-0,0000	0,0148	0,9260
	β_4	0,3714	0,0006	-0,0010	0,0133	0,9300
	$\log(\sigma)$	-0,1165	0,0006	0,0013	0,0123	0,9180
	γ	0,4329	0,0061	-0,0763	0,1401	0,9060
	α	0,4539	0,0014	-0,0285	0,0342	0,9340
	$\log(a)$	-0,0598	0,0030	-0,2615	0,0696	0,8820
	$\log(b)$	7,5610	0,0126	-0,0166	0,3092	0,9040
	D_{11}	-0,0969	0,0019	0,0278	0,0434	0,9140
	D_{12}	-0,0207	0,0005	-0,0394	0,0114	0,9360
	D_{22}	-1,7333	0,0019	0,0102	0,0454	0,9040
1000	β_1	2,8931	0,0022	0,0006	0,0489	0,9280
	β_2	-0,6028	0,0021	-0,0035	0,0473	0,9260
	β_3	0,3437	0,0005	0,0017	0,0105	0,9280
	β_4	0,3730	0,0004	0,0033	0,0095	0,9380
	$\log(\sigma)$	-0,1160	0,0004	-0,0027	0,0085	0,9340
	γ	0,4455	0,0043	-0,0495	0,0987	0,9360
	α	0,4553	0,0010	-0,0255	0,0260	0,9080
	$\log(a)$	-0,0634	0,0021	-0,2170	0,0497	0,9000
	$\log(b)$	7,5774	0,0090	-0,0144	0,2289	0,9000
	D_{11}	-0,0972	0,0013	0,0311	0,0287	0,9400
	D_{12}	-0,0208	0,0004	-0,0344	0,0080	0,9280
	D_{22}	-1,7293	0,0014	0,0079	0,0338	0,8800

Nota – Valores simulados: $\beta_1 = 2,8913$, $\beta_2 = -0,6049$, $\beta_3 = 0,3431$, $\beta_4 = 0,3718$, $D_{11} = -0,0942$, $D_{12} = -0,0215$, $D_{22} = -1,7157$, $\log(\sigma) = -0,1163$, $\gamma = 0,4687$ e $\alpha = 0,4672$, $\log(a) = -0,0809$ e $\log(b) = 7,6882$. Os parâmetros D_{11} , D_{12} e D_{22} são apresentados com a transformação log-Cholesky.

Figura 9 – Funções de densidade obtidas para os parâmetros da distribuição NK.



Nota – A linha tracejada representa o verdadeiro valor do parâmetro.

Fonte: Elaborada pelo autor.

3.7.3 Critérios para Seleção de Modelos Paramétricos

Conforme citado anteriormente, a família de distribuições gama generalizada (GG) é bastante flexível e capaz de generalizar diversas distribuições paramétricas comumente usadas, como exponencial, Weibull, log-normal e gama. Entretanto, os resultados do estudo de simulação da Seção 3.7.2.1 mostraram que a distribuição proposta apresenta problemas de identificabilidade nos parâmetros, nos remetendo à situação onde diferentes valores para um mesmo parâmetro levam a um mesmo valor de função de verossimilhança, o que impacta no processo inferencial e testes de hipóteses (BREUSCH, 1986; FUKUMIZU *et al.*, 2003; CAO *et al.*, 2013).

Dessa forma, propomos nesta Tese a utilização da GG para selecionar uma distribuição mais simples que pertença à essa família de distribuições, possivelmente de um ou dois parâmetros, e que seja capaz de prover um ajuste adequado aos dados. Essa seleção pode ser feita a partir do teste de razão de verossimilhanças ou então por meio dos critérios de informação AIC e BIC. É importante destacar que tanto a utilização da seleção de modelos paramétricos usuais via família GG, quanto utilizando-se critérios de inovação, são temas nunca antes estudados no contexto de modelagem conjunta de componentes longitudinais e de sobrevivência.

Assim, sob a teoria do teste de razão de verossimilhanças, investigamos a performance em testar a hipótese nula de que o tempo de falha pode ser descrito por uma das distribuições candidatas versus a hipótese alternativa de que o tempo de falha pode ser descrito por uma outra família que não seja a da hipótese nula. Por exemplo, conforme apresentado na Seção 3.1, se

usarmos a parametrização de Stacy podemos avaliar na hipótese nula as distribuições exponencial ($b = k = 1$), Weibull ($k = 1$), log-normal (caso limite quando $k \rightarrow \infty$) e gama ($b = 1$), enquanto que para a de Prentice podemos avaliar a distribuição exponencial ($q = \tau = 1$), Weibull ($q = 1$), log-normal ($q = 0$) e gama ($q = \tau$). Dessa forma, consideramos um caso particular da gama na verossimilhança restrita versus a própria distribuição distribuição na verossimilhança irrestrita, isto é, a gama generalizada, podendo-se avaliar a possível escolha de um caso particular.

Os resultados do estudo de simulação realizado para avaliar a discriminação de modelos via teste de razão de verossimilhanças considerando a família gama generalizada, bem como por critérios de informação (AIC e BIC), estão apresentados nas Tabelas 8 e 9. Destacamos que o estudo foi desenvolvido seguindo os seguintes passos: geramos os dados a partir de uma das três distribuições (Weibull, log-normal e gama), ajustamos os três modelos separadamente e aplicamos o TRV considerando-se a função gama generalizada como valor de referência com um nível de significância de 5% para mensurar as taxas de seleção correta. Nesse caso, rejeitamos a hipótese de que os dados seguem a distribuição avaliada no TRV quando o p-valor do teste foi menor que 5%. Na seleção via critérios de informação comparamos o valor do AIC e BIC dos três casos particulares, sendo que o modelo adequado foi considerado como o que apresentou o menor valor entre os três.

A partir da Tabela 8, podemos observar que a utilização do teste de razão de verossimilhanças apresenta níveis de significância razoavelmente próximos do nível nominal assumido, isto é, com taxas de rejeição da hipótese nula dado que a respectiva hipótese é verdade entre 5% e 10%, principalmente para os modelos gama e Weibull. Quando os verdadeiros tempos de falha são gerados pelas distribuições gama ou Weibull, observamos que o teste apresenta alto poder de rejeição para a distribuição log-normal. No entanto, verificamos que há um confundimento em rejeitar a distribuição Weibull quando os tempos são gerados pela gama, e vice-versa, visto que a maior taxa de rejeição do modelo incorreto nessas situações é dada por valores em torno de 45%.

Quando simulamos os tempos de falha a partir da distribuição log-normal, notamos que a taxa de convergência do modelo com função de risco gama generalizada é baixa, inviabilizando a utilização desse critério para grande parte das amostras. Nesse caso, para as amostras em que foi possível realizar o teste notamos que existe alto grau de rejeição para o modelo Weibull e baixo para o modelo gama. Além disso, verificamos que a utilização da parametrização de Prentice tende a ser mais frequente quando os tempos são gerados pela Weibull.

No total, constatamos que para todos os casos o poder em rejeitar o modelo incorreto aumenta à medida que o tamanho de amostra cresce e, diferentemente do esperado, temos que a taxa de rejeição do modelo correto também aumenta com o tamanho de amostra para o modelo Weibull. Sendo assim, podemos considerar que o teste de razão de verossimilhanças é capaz de discriminar entre os modelos Weibull e log-normal com alto poder.

Analisando os resultados da Tabela 9 constatamos que o critério de seleção AIC apresenta resultados satisfatórios principalmente para o maior tamanho de amostra. Em geral, temos

melhores resultados para Weibull e log-normal sendo que, nesse caso, há boa discriminação entre os modelos mesmo com amostras menores. Apesar de existir uma chance razoável de selecionar o modelo gama em ambos os casos, as taxas de seleção da Weibull quando os tempos de falha foram gerados pela log-normal, e vice-versa, são praticamente nulas. Ainda, quando os verdadeiros tempos de falha são gerados pela gama, temos que o critério de seleção AIC apresenta performance moderada, com viés de seleção para Weibull. Também, notamos que a taxa de seleção do verdadeiro modelo aumenta e a do modelo incorreto diminui a medida que o tamanho da amostra cresce, independentemente da distribuição avaliada. Diferentemente do método de seleção anterior, esse método é robusto quando os tempos de falha seguem distribuição log-normal, sendo independente do ajuste do modelo GG.

É importante lembrar que, pelo fato das distribuições testadas terem as mesmas quantidades de parâmetros, os critérios AIC e BIC apresentam as mesmas decisões. Vale destacar que Yu (2017) obteve resultados de discriminação similares no contexto usual de análise de sobrevivência, isto é, sem a componente longitudinal.

Tabela 8 – Taxas de rejeição da hipótese nula ao nível de 5% de significância considerando o teste de razão de verossimilhanças via família GG.

Distribuição de Geração	n	Distribuição Ajustada				
		Gama	Log-normal	Weibull	Gama generalizada	
					Convergência	Parametrização de Stacy
Gama	250	0,084	0,420	0,059	0,826	0,913
	500	0,082	0,593	0,180	0,907	0,960
	1000	0,065	0,828	0,396	0,956	0,970
Log-normal	250	0,032	0,120	0,491	0,407	0,963
	500	0,010	0,124	0,906	0,406	0,995
	1000	0,037	0,136	1,000	0,329	1,000
Weibull	250	0,182	0,728	0,049	0,945	0,872
	500	0,309	0,936	0,075	0,969	0,889
	1000	0,468	0,998	0,103	0,975	0,892

Tabela 9 – Taxas de seleção dos modelos avaliados de acordo com o AIC.

Distribuição de Geração	n	Distribuição Ajustada			
		Gama	Log-normal	Weibull	Convergência
Gama	250	0,372	0,220	0,408	1,000
	500	0,522	0,126	0,351	0,997
	1000	0,673	0,058	0,269	0,985
Log-normal	250	0,273	0,685	0,042	1,000
	500	0,266	0,729	0,005	0,994
	1000	0,161	0,839	0,000	0,961
Weibull	250	0,224	0,059	0,717	1,000
	500	0,213	0,010	0,777	1,000
	1000	0,149	0,001	0,850	1,000

3.8 Modelo Conjunto no Contexto do Trabalho de Parto

Nas seções a seguir apresentamos uma série de aplicações dos modelos conjuntos com função de risco flexível no contexto do trabalho parto. No total, são analisadas informações de 6059 pacientes grávidas de apenas um bebê e que tiveram parto espontâneo. Além disso, são consideradas informações sobre dilatação cervical, membrana amniótica, paridade prévia, comorbidade ou complicação pré gravidez e se a mulher em trabalho de parto veio referenciada de outra unidade de saúde.

Conforme apresentado na Tabela 10, 34,5% das mulheres chegaram ao hospital com dilatação maior do que 4 cm, 78% com membrana amniótica intacta e 41,5% nunca tinham tido bebê anteriormente. Ainda, 14,3% das mulheres apresentavam comorbidades ou complicações pré gravidez e 4,8% foram referenciadas em trabalho de parto por outra unidade de saúde. Com relação às intervenções e desfechos do parto, temos que 31,8% recebeu aumento, 10,2% tiveram parto por cesárea e 89,8% vaginal. Analisando o tempo até os possíveis eventos de interesse, temos que o tempo médio até aumento, caracterizada como uma intervenção feita pelos médicos, é igual a 7,58 horas e tanto o tempo médio até cesárea, quanto até parto vaginal, são próximos a 8,9 horas.

Particularmente, consideramos um modelo para avaliar o tempo até cesárea, contemplando a seguinte estrutura: no modelo longitudinal, que avalia a evolução da dilatação cervical, consideramos uma tendência linear explicada pelo tempo e por paridade prévia, enquanto que o modelo de sobrevivência considera informação sobre paridade prévia, comorbidades ou complicações pré-gravidez, se a mulher em trabalho de parto foi referenciada por outra unidade de saúde e o status da membrana amniótica.

Como todas as mulheres que tiveram parto normal estão incluídas na análise, podemos levar em conta que o modelo para cesárea pode ser interpretado como um tipo de modelo de causa específica (PUTTER; FIOCCO; GESKUS, 2007), contrastando mulheres que tiveram parto por cesárea com aquelas que tiveram parto normal. Além disso, as mulheres com parto maiores do que 48 horas após a admissão foram censuradas e, na estimação dos parâmetros dos modelos analisados, consideramos o mesmo número de quadraturas apresentados no estudo de simulação, ou seja, 15 pontos na quadratura de Gauss-Kronrod e 5 pontos na de Gauss-Hermite. Com relação ao modelo de efeitos mistos, aplicamos a seguinte transformação na variável resposta dilatação: $\log(x + 1)$ e, para obtermos o erro padrão e intervalos de 95% de confiança para os parâmetros estimados, utilizamos 100 réplicas de *bootstrap*. Além disso, destacamos que tanto na formulação dos modelos, quanto na apresentação dos resultados, as classes da variável paridade prévia são representadas como “Para 1” e “Para 2+”, sendo que a primeira representa mulheres que já tiveram um filho anteriormente e a segunda dois ou mais filhos.

A seguir, apresentamos a estrutura geral dos modelos considerados para um paciente i observado no tempo j :

- Transformação na variável dilatação cervical:

$$- y_{ij} = \log(\text{Dilatação}_i(t_{ij}) + 1)$$

- Modelo longitudinal:

$$\begin{aligned} y_{ij} &= m_i(t_{ij}) + e_{ij} \\ &= \beta_0 + \beta_1 t_{ij} + \beta_2 \{\text{Para } 1_i\} + \beta_3 \{\text{Para } 2+i\} \\ &\quad + b_{i0} t_{ij} + b_{i1} t_{ij} + e_{ij} \end{aligned}$$

$$- b_i = (b_{i0}, b_{i1})^\top \sim \mathcal{N}(\mathbf{0}, \mathbf{D});$$

- D não estruturada;

$$- e_i = (e_{i1}, \dots, e_{in_i})^\top \sim \mathcal{N}(0, \sigma^2 I_{n_i}).$$

- Modelo de sobrevivência:

$$\begin{aligned} \log \{h_i(t)\} &= \log \{h_0(t)\} + \gamma_1 \{\text{Para } 1_i\} + \gamma_2 \{\text{Para } 2+i\} \\ &\quad + \gamma_3 \{\text{Comorbidades ou complicações pré gravidez}_i\} \\ &\quad + \gamma_4 \{\text{Mulher referenciada de outra unidade}_i\} \\ &\quad + \gamma_5 \{\text{Membrana amniótica rompida sem mecônio}_i\} \\ &\quad + \gamma_6 \{\text{Membrana amniótica rompida com mecônio}_i\} \\ &\quad + \alpha m_i(t). \end{aligned}$$

3.8.1 Modelagem a partir de Generalizações da Distribuição Weibull

Quanto à estimação dos parâmetros, é importante mencionar que função de risco GG considerando a proposta de [Prentice \(1974\)](#) apresentou instabilidade nos passos finais do algoritmo de otimização pelo fato de o parâmetro q ter convergido para zero. Quando q tende a zero, a quantidade $\gamma(q^{-2})$ tende ao infinito, causando instabilidade na verossimilhança. Portanto, para corrigir esse problema, aplicamos a seguinte transformação: $q^{-2} = q^*$. Com relação às demais distribuições, consideramos $a^* = 1/a$ para a distribuição EW e $a^* = (1/a)^b$ para a distribuição MOW.

A Tabela 11 apresenta os resultados da log-verossimilhança maximizada e critérios de informação para as distribuições GG, com a parametrização de Stacy e Prentice, EW, GW, MOW e GPW. Como podemos observar, a distribuição GG com parametrização de Prentice apresentou o maior valor em termos da função log-verossimilhança e menores AIC e BIC, indicando que, provavelmente, é a que mais se adéqua aos dados. Também, temos que a distribuição EW apresentou performance bastante similar. Esse resultado está de acordo com o apresentado por [Cox e Matheson \(2014\)](#) em uma comparação entre os modelos GG e EW sob o contexto usual de análise de sobrevivência.

Tabela 10 – Caracterização geral da amostra de acordo com as principais variáveis estudadas.

	Total N = 6059
Dilatação > 4 - n (%)	2088 (34,5)
Membrana amniótica - n (%)	
Intacta	4727 (78,0)
Rompida sem mecônio	1203 (19,9)
Rompida com mecônio	129 (2,1)
Paridade prévia - n (%)	
Não	2515 (41,5)
1	1550 (25,6)
+2	1994 (32,9)
Comorbidade ou complicação pré gravidez - n (%)	868 (14,3)
Mulher em trabalho de parto referenciada por outra unidade de saúde - n (%)	289 (4,8)
Realização de aumento - n (%)	1929 (31,8)
Tempo até aumento - média (desvio padrão)	7,58 (7,30)
Cesárea - n (%)	615 (10,2)
Tempo até cesárea - média (desvio padrão)	8,86 (7,63)
Parto vaginal - n (%)	5444 (89,8)
Tempo até parto vaginal - média (desvio padrão)	8,89 (7,65)

Para efeito de comparação, consideramos uma abordagem não paramétrica baseada em B-splines (RIZOPOULOS, 2012b), que expressa a função de risco como

$$h_0(t) = \sum_{k=1}^K \xi_k I(v_0 < v_1 < \dots < v_k),$$

em que k é o número de nós, $0 = v_0 < v_1 < \dots < v_K$ denota partições da escala de tempo, com v_K representando o maior valor observado, e denota o ξ_k o valor da função de risco no intervalo $(v_{k-1}, v_k]$. Nesse caso, à medida que o número de nós aumenta, a função de risco se torna mais flexível.

Comparando os resultados obtidos com uma abordagem que considera 5 nós internos, temos que apenas o modelo com função de risco MOW apresenta performance inferior ao não paramétrico, com diferenças ligeiramente inferiores. Além disso, vale ressaltar que o modelo baseado em B-splines possui 2 parâmetros a mais que os propostos, efeito que pode ser observado quando comparamos o BIC, cujo modelo não paramétrico tem a pior performance por ser menos parcimonioso.

As estimativas de máxima verossimilhança, bem como erro padrão e intervalo de 95% de confiança, calculados via *bootstrap*, para os parâmetros do modelo com maior performance estão dispostas na Tabela 12. Por meio dos resultados obtidos, temos que todas as variáveis são significantes para o modelo longitudinal de dilatação, com um impacto positivo de paridade prévia, indicando que mulheres que já tiveram filhos tendem a estar em um nível maior de dilatação do que as que nunca tiveram. Nesse caso, chegamos às conclusões de significância

para os parâmetros a partir dos intervalos de confiança, sendo que quando o zero está contido no intervalo, consideramos que o parâmetro não é significativamente diferente de zero. Analisando o modelo de sobrevivência, destacamos que somente paridade prévia, se a mulher foi referenciada por outra unidade de saúde e o status da membrana amniótica, especialmente com mecônio, são significativos para se ter um parto por cesárea. Em geral, mulheres múltiparas tendem a apresentar menor risco de cesárea, enquanto que mulheres transferidas de outra unidade de saúde e com membrana rompida com mecônio tendem a apresentar maior risco. Observando o valor do parâmetro α , concluímos que existe um impacto significativo da evolução da dilatação no evento de interesse. A razão de riscos para os efeitos fixos e evolução dinâmica da dilatação podem ser obtidas, respectivamente, por $\exp(\gamma_i)$ para $i = 1, \dots, 6$ e $\exp(\alpha)$. Como o limite inferior do intervalo de confiança para membrana rompida sem mecônio ficou bem próximo de zero, temos indícios que também deve haver um impacto dessa variável no parto por cesárea.

Uma das principais características do modelo conjunto é poder estimar a probabilidade de sobrevivência dinamicamente considerando a evolução da componente longitudinal. Dessa forma, a partir das equações apresentadas no Capítulo 2, estimamos as probabilidades de sobrevivência, que são equivalentes à mulher não ter cesárea, para um paciente com as seguintes características: sem dilatação na admissão, sem paridade prévia, com comorbidade ou complicação pré gravidez, membrana amniótica intacta e que não foi referenciada em trabalho de parto por outra unidade de saúde. Os resultados para quatro diferentes tempos, em horas, estão apresentados na Figura 10.

Tabela 11 – Resultados de log-verossimilhança(log L(.)), AIC e BIC para todas as generalizações propostas da distribuição Weibull.

Modelo	log L(.)	AIC	BIC
GG (Stacy)	-8780,38	17596,75	17717,52
GG (Prentice)	-8780,05	17596,09	17716,86
EW	-8780,16	17596,31	17717,08
GW	-8782,81	17601,62	17722,39
MOW	-8788,27	17612,55	17733,31
GPW	-8782,69	17601,38	17722,15
Modelo não paramétrico considerando B-Splines com 5 nós	-8783,85	17607,71	17741,89

3.8.2 Seleção de Modelos via Família Gama Generalizada

A Tabela 13 exibe os resultados da log-verossimilhança maximizada e os critérios de informação para as duas parametrizações da GG e os três modelos candidatos. Conforme podemos observar, a GG, independentemente da parametrização, bem como a log-normal apresentam os melhores resultados, devido aos maiores valores da função log-verossimilhança. Comparando as duas abordagens da GG, temos que a de Prentice apresenta o melhor resultado, justificando sua utilização de estabilidade para o caso limite em que q tende a zero. Portanto, consideramos essa parametrização como modelo de referência no testes de razão de verossimilhanças. Consi-

Tabela 12 – Estimativas de máxima verossimilhança, erro padrão e intervalo de confiança de 95% para o modelo conjunto com função de risco GG.

		Estimativa	Erro padrão	IC 95%	
Intercepto	β_0	1,4958	0,0080	1,4788	1,5097
Tempo	β_1	0,0858	0,0012	0,0838	0,0877
Para 1	β_2	0,1443	0,0124	0,1235	0,1632
Para 2+	β_3	0,1456	0,0106	0,1267	0,1662
	$\log(\sigma)$	-1,4993	0,0157	-1,5276	-1,4696
<i>Modelo de sobrevivência</i>					
Para 1	γ_1	-0,6174	0,1723	-0,8709	-0,3884
Para 2+	γ_2	-0,8891	0,2359	-1,1447	-0,6172
Comorbidade ou complicação pré gravidez	γ_3	0,0541	0,1086	-0,1767	0,2925
Mulher em trabalho de parto referenciada por outra unidade de saúde	γ_4	0,4982	0,1711	0,1619	0,7151
Membrana amniótica rompida sem mecônio	γ_5	0,2516	0,1261	-0,0647	0,4020
Membrana amniótica rompida com mecônio	γ_6	0,8261	0,3230	0,2330	1,2989
	α	1,4082	0,1521	1,0211	1,6758
	μ	3,5614	0,3235	3,0888	4,3011
	$\log(\tau)$	2,3669	0,8217	0,9799	3,1051
	$\log(q^*)$	1,7603	1,2808	1,0794	2,2808

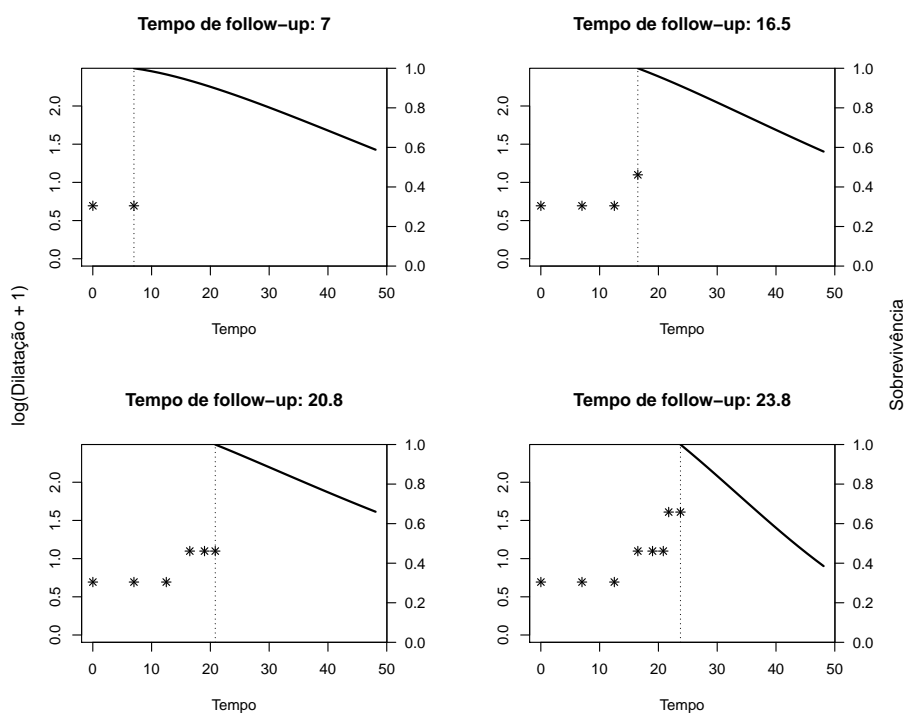


Figura 10 – Predições dinâmicas da probabilidade de sobrevivência para o modelo conjunto com função de risco GG considerando uma paciente sem dilatação na admissão, nulípara, com comorbidade ou complicação pré gravidez, não referenciada de outro hospital e com membrana aminótica intacta.

Fonte: Elaborada pelo autor.

derando o teste sugerido, temos claramente que a suposição dos modelos Weibull e gama foram rejeitadas ao nível de 5% de significância e que a função de risco log-normal se destaca pelo ajuste adequado aos dados. Ainda, nota-se que mesmo considerando os critérios de informação AIC e BIC, o modelo log-normal também apresenta bons resultados, com os menores valores entre os possíveis candidatos.

As estimativas de máxima verossimilhança para o modelo log-normal, bem como erro padrão e intervalo de 95% de confiança dos parâmetros estão dispostas na Tabela 14. Podemos observar que, com exceção dos parâmetros da função de risco, que dependem da parametrização utilizada, os demais valores foram praticamente iguais ao modelo GG. Podemos notar diferenças nos intervalos de confiança, principalmente para a variável relacionada à membrana amniótica. Nesse caso, o intervalo de confiança para o parâmetro associado à membrana amniótica rompida sem mecônio não contém o zero, sugerindo que existe um impacto positivo para o risco de cesárea quando a membrana está rompida.

Assim como feito para a GG na seção anterior, estimamos as probabilidades de sobrevivência dinâmicas de um paciente específico considerando a função de risco log-normal (Figura 11).

Tabela 13 – Resultados para discriminação de modelos via família GG.

Model	Log-verossimilhança	AIC	BIC	TRV	p-valor
Weibull	-8789,015	17612,03	17726,09	17,9374	<0,0001
Log-normal	-8780,174	17594,35	17708,40	0,2539	0,3072
Gama	-8786,566	17607,13	17721,19	13,0397	0,0003
GG (Stacy)	-8780,376	17596,75	17717,52	-	-
GG (Prentice)	-8780,047	17596,09	17716,86	(ref.)	-

Tabela 14 – Estimativas de máxima verossimilhança, erro padrão e intervalo de confiança de 95% para o modelo conjunto com função de risco log-normal.

		Estimativa	Erro padrão	IC 95%	
<i>Modelo longitudinal</i>					
Intercepto	β_0	1,4958	0,0076	1,4819	1,5102
Tempo	β_1	0,0858	0,0012	0,0836	0,0882
Para 1	β_2	0,1443	0,0114	0,1205	0,1614
Para 2+	β_3	0,1456	0,0097	0,1262	0,1665
	$\log(\sigma)$	-1,4992	0,0152	-1,5301	-1,4730
<i>Modelo de sobrevivência</i>					
Para 1	γ_1	-0,6179	0,1223	-0,8350	-0,4496
Para 2+	γ_2	-0,8896	0,1329	-1,1675	-0,6330
Comorbidade ou complicação pré gravidez	γ_3	0,0556	0,1129	-0,1819	0,2847
Mulher em trabalho de parto referenciada por outra unidade de saúde	γ_4	0,4986	0,1763	0,1482	0,8100
Membrana amniótica rompida sem mecônio	γ_5	0,2521	0,1054	0,0540	0,4736
Membrana amniótica rompida com mecônio	γ_6	0,8280	0,2218	0,4049	1,2765
	α	1,4144	0,1376	1,2346	1,7519
	μ	5,8241	0,2790	5,3869	6,4212
	$\log(\tau)$	0,3517	0,0448	0,2784	0,4388

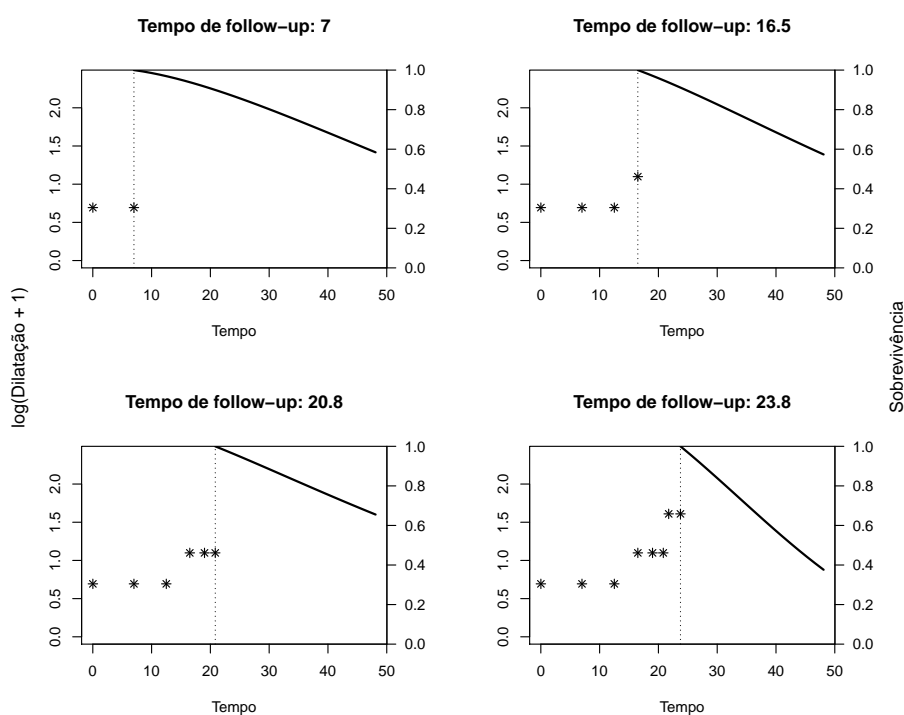


Figura 11 – Predições dinâmicas da probabilidade de sobrevivência para o modelo conjunto com função de risco log-normal considerando uma paciente sem dilatação na admissão, nulípara, com comorbidade ou complicação pré gravidez, não referenciada de outro hospital e com membrana aminótica intacta.

Fonte: Elaborada pelo autor.

3.8.3 Modelo Nakagami-m

Neste caso, consideramos o modelo conjunto com função de risco NK para o modelo de cesárea, cujos resultados estão apresentados na Tabela 15. Assim como nos modelos anteriores, observamos que todas as variáveis foram significativas no modelo longitudinal, com efeito de paridade prévia no aumento da dilatação. Para o modelo de sobrevivência, novamente apenas as variáveis paridade prévia, se a mulher em trabalho de parto foi referenciada por outra unidade de saúde e status da membrana amniótica foram significativas. Considerando esta última variável, constatamos comportamento similar ao modelo log-normal, em que membrana com e sem mecônio apresentaram intervalo de confiança de *bootstrap* sem inclusão do zero. Em geral, mulheres que já tiveram filho tendem a ter menor risco de cesárea e mulheres referenciadas por outra unidade de saúde ou então com membrana rompida tendem a ter maior risco.

Para efeito de comparação, apresentamos na Tabela 16 a log-verossimilhança e os critérios de informação AIC e BIC para diferentes famílias paramétricas com dois parâmetros. Notamos que a distribuição NK tem melhor performance que a Gompertz (GOMPERTZ, 1825) e resultados similares à Weibull. Diferentemente, o modelo proposto apresentou performance inferior à log-normal e gama.

Por fim, as estimativas de probabilidade de sobrevivência dinâmicas para a NK, segundo um perfil fixado de paciente, estão expostas na Figura 12.

Tabela 15 – Estimativas de máxima verossimilhança, erro padrão e intervalo de confiança de 95% para o modelo conjunto com função de risco NK.

		Estimativa	Erro padrão	IC 95%	
<i>Modelo longitudinal</i>					
Intercepto	β_0	1,4957	0,0077	1,4812	1,5121
Tempo	β_1	0,0858	0,0012	0,0833	0,0885
Para 1	β_2	0,1444	0,0106	0,1234	0,1641
Para 2+	β_3	0,1457	0,0095	0,1234	0,1603
	$\log(\sigma)$	-1,4990	0,0166	-1,5336	-1,4683
<i>Modelo de sobrevivência</i>					
Para 1	γ_1	-0,6309	0,1293	-0,8844	-0,3446
Para 2+	γ_2	-0,9025	0,1373	-1,1539	-0,6040
Comorbidade ou complicação pré gravidez	γ_3	0,0795	0,1074	-0,0858	0,2788
Mulher em trabalho de parto referenciada por outra unidade de saúde	γ_4	0,5077	0,1834	0,1400	0,8404
Membrana amniótica rompida sem mecônio	γ_5	0,2656	0,1024	0,0730	0,4367
Membrana amniótica rompida com mecônio	γ_6	0,8584	0,2368	0,3616	1,2610
	α	1,5003	0,1577	1,2369	1,8138
	$\log(a)$	-0,0934	0,0300	-0,1479	-0,0266
	$\log(b)$	10,3138	0,3908	9,6699	11,0512

Tabela 16 – Resultados de log-verossimilhança, AIC e BIC para diferentes funções paramétricas.

Critério	Weibull	Log-normal	Gama	Gompertz	NK
Log-verossimilhança	-8789,01	-8780,17	-8786,56	-8841,49	-8789,21
AIC	17612,03	17594,35	17607,13	17716,98	17612,44
BIC	17726,09	17708,40	17721,19	17831,04	17726,49

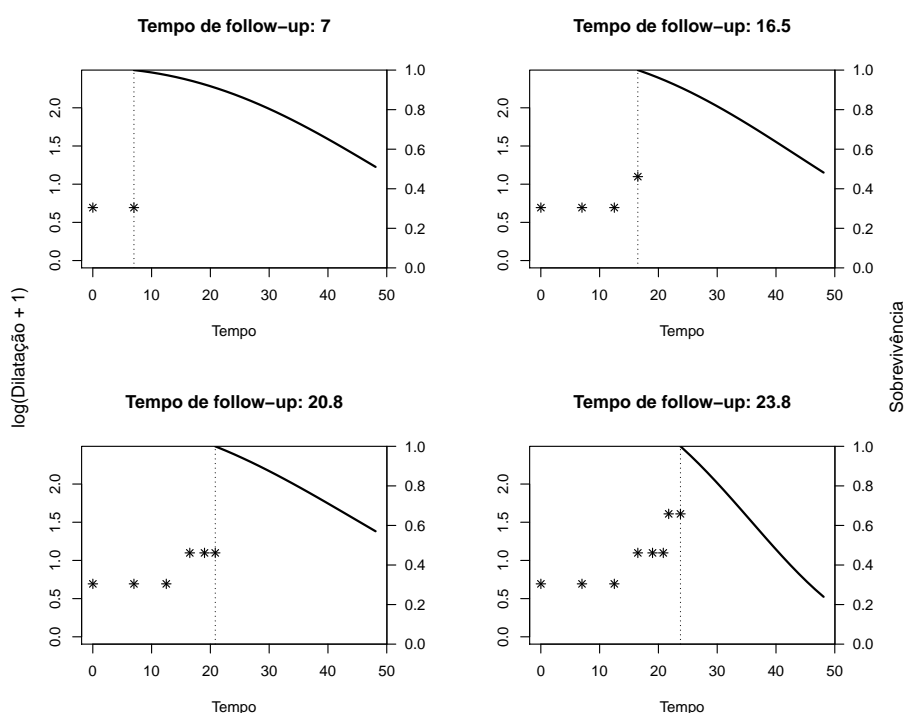


Figura 12 – Predições dinâmicas da probabilidade de sobrevivência para o modelo conjunto com função de risco NK considerando uma paciente sem dilatação na admissão, nulípara, com comorbidade ou complicação pré gravidez, não referenciada de outro hospital e com membrana aminótica intacta.

Fonte: Elaborada pelo autor.

3.9 Comentários Finais

Neste capítulo apresentamos as generalizações da distribuição Weibull com função de risco flexível propostas à modelagem conjunta de dados longitudinais e de sobrevivência. Dentre elas, apresentamos a distribuição gama generalizada, consideradas pela parametrizações de Stacy e Prentice, Weibull exponenciada, Weibull generalizada, Marshall-Olkin-Weibull e distribuição de potência Weibull generalizada. Também, introduzimos a distribuição Nakagami- m , que é uma família de dois parâmetros capaz de lidar com as diferentes formas da função de risco.

Um estudo de simulação foi realizado com finalidade de avaliar a performance do método de máxima verossimilhança e a seleção de modelos por teste de razão de verossimilhanças e critérios de informação. Em geral, observamos aproximações satisfatórias para os parâmetros. Entretanto, os maiores desvios em relação aos verdadeiros valores estão associados aos parâmetros da função de risco. Além disso, observamos problemas de identificabilidade nas distribuições gama generalizada, Weibull exponenciada e potência Weibull generalizada. Com relação à seleção de modelos paramétricos, verificamos maior poder de discriminação entre os modelos Weibull e log-normal, sendo que quando os tempos de falha são gerados pela log-normal observa-se problemas de convergência da gama generalizada, sugerindo a utilização dos critérios

de informação.

Por fim, os modelos estudados foram utilizados para avaliar o risco da ocorrência de parto por cesariana, considerando-se a dilatação como covariável no tempo, fato inédito no contexto do trabalho de parto. Nesse caso, propusemos a utilização de generalizações da distribuição Weibull, a seleção de modelos paramétricos usuais via distribuição gama generalizada e a aplicação do modelo conjunto com função de risco Nakagami- m . Os modelos utilizados foram úteis para compreender como as variáveis utilizadas se relacionam com um desfecho do parto por cesárea, considerando-se, principalmente, a evolução dinâmica da dilatação, visto que essa é uma das variáveis mais importantes do estudo. Os resultados obtidos permitem que médicos e enfermeiras tenham maior conhecimento sobre o complexo fenômeno do parto, permitindo-se que haja uma melhor qualidade e cuidado do mesmo, tomando-se decisões mais assertivas.

O capítulo a seguir introduz o conceito de modelos multiestados, as possíveis estruturas do modelo, as probabilidades e intensidades de transição e os métodos de estimação dos parâmetros. Também, exibimos como covariáveis dependentes do tempo podem ser utilizadas no contexto de eventos múltiplos, sendo que um modelo multiestados considerando covariáveis fixas e dinâmicas é ajustado ao trabalho de parto com três possíveis desfechos: realização de intervenção, parto vaginal e por cesárea.

MODELOS MULTIESTADOS

Nos capítulos anteriores consideramos modelos que avaliam apenas um desfecho no contexto de análise de sobrevivência. Entretanto, em muitos casos estão presentes múltiplos desfechos, sendo possível relacioná-los a partir de um modelo que leva em conta esse tipo de estrutura. Particularmente, no contexto do trabalho de parto, temos diversas situações que podem ser consideradas como desfechos, tais como a realização de uma intervenção pelo médico/enfermeira durante o trabalho de parto, o nascimento do bebê por parto vaginal ou cesárea, sendo interessante avaliá-los de forma conjunta. Dessa forma, no presente capítulo apresentamos os chamados modelos multiestados, que são os modelos mais utilizados quando esse tipo de característica está presente nos dados.

4.1 Processo Multiestados

Segundo [Meira-Machado *et al.* \(2008\)](#), um processo multiestados é um processo estocástico $(X(t), t \in \mathcal{T})$ que assume valores em um espaço de estados finito $\mathcal{S} = \{1, \dots, S\}$. Nesse caso, $\mathcal{T} = [0, \tau)$, $\tau < \infty$, é um intervalo de tempo e o valor do processo no tempo t é dado pelo estado ocupado no respectivo tempo. Com a evolução do processo, uma história \mathcal{H}_t^- é gerada pelas observações do processo ao longo do intervalo $[0, \tau)$, tais como os estados previamente visitados, os tempos de transição, etc. Matematicamente, \mathcal{H}_t^- é uma σ -álgebra de todos os eventos que podem acontecer antes de t ([COMMENGES; JACQMIN-GADDA, 2015](#)).

O processo multiestados é caracterizado pelas probabilidades de transição entre os estados g e h

$$p_{gh}(s, t) = \Pr(X(t) = h \mid X(s) = g, \mathcal{H}_t^-)$$

para $g, h \in \mathcal{S}$, $s, t \in \mathcal{T}$ e $s \leq t$. O processo multiestado também pode ser caracterizado pelas intensidades de transição

$$h_{gh}(t) = \lim_{dt \rightarrow 0} \frac{p_{gh}(t, t + dt)}{dt}$$

que representam o risco instantâneo de progressão do estado g para o estado h , sendo que tanto $p_{gh}(s, t)$ quanto $h_{gh}(t)$ dependem da história \mathcal{H}_t^- do processo.

Nesse caso, a função de risco acumulada é definida como $H_{gh}(t) = \int_0^t h_{gh}(u) du$ e assumimos que existem $K > 0$ transições possíveis entre os estados do modelo. Caso não seja possível transitar entre os estados g e h consideramos $H_{gh}(t) = 0$. Essas intensidades podem ser reproduzidas em uma matriz $\mathbf{H}(t)$ de dimensão $S \times S$ e com diagonal $H_{gg}(t) = -\sum_{h \neq g} H_{gh}(t)$. Esta última equação indica que indivíduos que não têm mais nenhuma transição permanecem no estado g .

Podemos classificar os estados do processo estocástico de duas maneiras: absorvente e transiente. Um estado $g \in \mathcal{S}$ é absorvente se para todo $t \in \mathcal{T}$, $h \in \mathcal{S}$, $j \neq h$, $h_{gh}(t) = 0$; caso contrário g é dito transiente (ANDERSEN; KEIDING, 2002). Dessa forma, os estados absorventes não permitem transições a partir deles, ou seja, uma vez que o indivíduo entra em um estado com essa característica ele permanece nele para sempre.

Levando em conta as suposições feitas sobre o modelo, podemos caracterizá-los como: modelo markoviano, semi-markoviano, oculto e não markoviano. Muitas vezes a utilização de determinado tipo de modelo é definida pela própria característica dos dados ou até mesmo pela avaliação da significância no risco de transição para uma variável que leva em conta o tempo gasto em um estado específico. Como exemplo, podemos citar o estudo apresentado por Andersen e Keiding (2002) que avaliou a violação da suposição markoviana em um modelo doença-morte com 3 estados.

No presente trabalho consideramos especificamente o caso markoviano. Nesse caso, o processo estocástico $(X(t), t \geq 0)$ é markoviano se para quaisquer s, t , com $0 \leq s < t$, e quaisquer $g, h \in \mathcal{S}$, tem-se

$$\Pr(X(t) = h \mid X(s) = g, \mathcal{H}_t^-) = \Pr(X(t) = h \mid X(s) = g),$$

ou seja, o futuro do processo após o tempo s depende somente do estado atual ocupado no tempo s .

Um caso especial de modelo markoviano é o modelo markoviano homogêneo, que assume que as intensidades de transição são funções constantes no tempo, ou seja, $h_{gh}(t) = h_{gh}$. No entanto, em algumas aplicações a hipótese de homogeneidade pode não ser realista, sendo necessária a utilização de um modelo não homogêneo.

4.2 Estrutura dos Modelos Multiestados

Dentre as diversas características dos modelos multiestados, podemos citar sua versatilidade para modelar diferentes tipos de estruturas. Essas estruturas são as formas com que os estados são relacionados e, a partir delas, podemos definir quais as transições entre os múltiplos

estados analisados. Como exemplo, temos os modelos unidirecionais, modelos progressivos, modelos bidirecionais e modelos recorrentes (TITMAN, 2008). A seguir, apresentamos as principais características de cada um deles.

4.2.1 Modelos Unidirecionais

Modelos unidirecionais consistem em uma simples cadeia de estados, em que os sujeitos iniciam no estado 1 e podem progredir através de estados subsequentes até um estado absorvente (Figura 13). No modelo de análise de sobrevivência temos, basicamente, dois estados: o estado 1, que indica que o sujeito está vivo, e o estado 2, em que o sujeito está morto. Portanto, temos que este é o caso mais simples de modelo unidirecional.

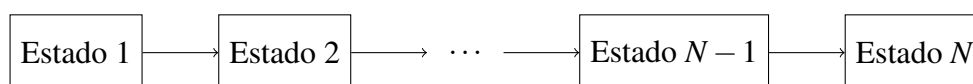


Figura 13 – Exemplo de modelo unidirecional.

Fonte: Elaborada pelo autor.

4.2.2 Modelos Progressivos

Estes modelos assumem a forma de um grafo acíclico direcionado e podem permitir a escolha de caminhos fora de um estado. No entanto, uma vez que um sujeito deixa um estado, ele não pode retornar. Um exemplo, comum desse tipo de modelo, é o modelo de doença crônica de 3 estados, em que os sujeitos iniciam no estado 1 (saudável) e podem progredir para os estados 2 (doente) ou 3 (morte). Uma vez que o sujeito está no estado 2 ele só pode progredir para o estado 3. Se considerarmos dois tipos de doença, teremos um modelo com 4 estados, conforme apresentado na Figura 14. Os modelos unidirecionais são exemplos da vasta classe de modelos progressivos.

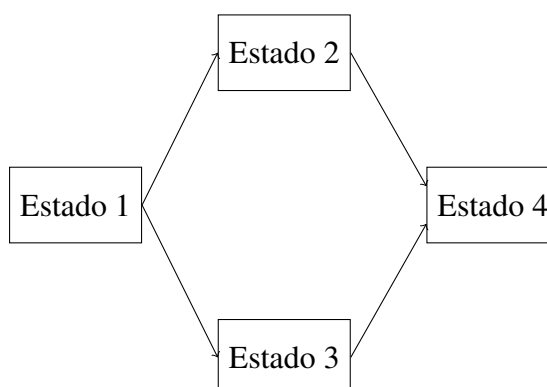


Figura 14 – Exemplo de modelo progressivo.

Fonte: Elaborada pelo autor.

4.2.3 Modelos Bidirecionais

Os modelos bidirecionais contém um estado absorvente, mas podem permitir transições em ambas as direções a partir de um estado transiente (Figura 15). Como exemplo, temos o modelo de doença de 3 estados em que os sujeitos podem se recuperar da doença. Alguns autores se referem à esses modelos como sendo reversíveis.

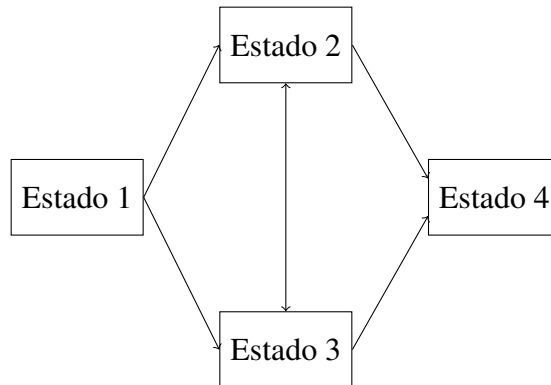


Figura 15 – Exemplo de modelo bidirecional. Mudanças entre os estados 2 e 3 são possíveis.

Fonte: Elaborada pelo autor.

4.2.4 Modelos Recorrentes

Modelos recorrentes não têm um estado absorvente e incluem estados em que, em algum momento, é eventualmente possível que o processo retorne a ele. O exemplo mais simples desse tipo de modelo é o modelo de doença-cura com 2 estados em que o estado 1 representa que o sujeito está saudável e o 2 que o sujeito está doente. Na Figura 16 apresentamos um exemplo desse modelo com 3 estados.

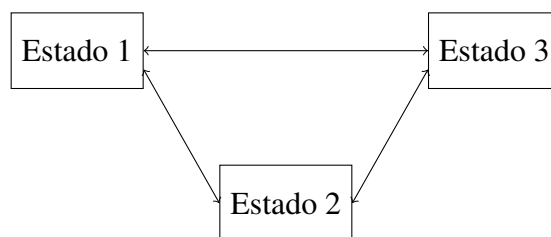


Figura 16 – Exemplo de modelo recorrente.

Fonte: Elaborada pelo autor.

4.3 Modelos de Regressão Multiestados

Definindo $X_i(t)$, $i = 1, \dots, n$, assumimos que as intensidades de transição são as mesmas para todos os indivíduos. No entanto, em situações práticas, pode ser interessante relacionar

características individuais às taxas de transição a partir de um vetor de covariáveis, \mathbf{w} , possivelmente dependente do tempo. Considerando um modelo de regressão geral, podemos escrever em termos de função de risco, tal como (MEIRA-MACHADO *et al.*, 2008)

$$h_{gh}(\cdot) = \varphi(h_{gh,0}(\cdot), \boldsymbol{\gamma}_{gh}^\top \mathbf{w}_i)$$

em que $h_{gh,0}(\cdot)$ é a função de risco basal entre os estados g e h , $\boldsymbol{\gamma}_{gh}$ é o vetor de parâmetros de regressão e \mathbf{w}_i é o vetor de covariáveis para o sujeito i .

Se tomarmos $\varphi(u(\cdot), v) = u(\cdot)e^v$ teremos o modelo de riscos proporcionais, isto é,

$$h_{gh}(t, \mathbf{w}) = h_{gh,0}(t) \exp(\boldsymbol{\gamma}_{gh}^\top \mathbf{w}).$$

Uma alternativa que permite coeficientes dependente do tempo, $\boldsymbol{\gamma}_{gh}(t)$, foi proposta por Aalen para dados de sobrevivida e depois utilizada em modelos multiestados. Essa abordagem pode ser obtida escolhendo $\varphi(u(\cdot), v) = u(\cdot) + v$, ou seja,

$$h_{gh}(t, \mathbf{w}) = h_{gh,0}(t) + \boldsymbol{\gamma}_{gh}^\top \mathbf{w}.$$

Para modelos markovianos, homogêneos e não homogêneos, geralmente assume-se o modelo de riscos proporcionais de Cox (COX, 1972), dado por,

$$h_{gh}(t, \mathbf{w}) = h_{gh,0}(t) \exp(\boldsymbol{\gamma}_{gh}^\top \mathbf{w}) \quad (4.1)$$

para relacionar as intensidades de transição $h_{gh,0}(t)$ às covariáveis \mathbf{w} .

Devido à sua frequente utilização em análise de sobrevivência e pelo efeito multiplicativo da função exponencial dos parâmetros associados à intensidade de transição, consideraremos esse tipo de modelo ao problema do trabalho de parto. A utilização de covariáveis nesses modelos permite prever as probabilidades para cada indivíduo, sendo que tanto nos modelos homogêneos quanto nos não homogêneos a inferência pode ser feita a partir de algumas suposições e manipulação dos dados.

A seguir exibimos os procedimentos de estimação para as probabilidades e intensidades de transição apresentados por Wreede, Fiocco e Putter (2010). Esses procedimentos são baseados em Andersen *et al.* (1993) e foram considerados na construção do pacote `mstate` (WREEDE; FIOCCO; PUTTER, 2011), disponível no *software* R (R Core Team, 2018).

4.4 Procedimentos de Estimação

Quando utilizamos modelos multiestados, a matriz de probabilidades de transição $\mathbf{P}(s, t)$ é a primeira quantidade de interesse. Seus elementos são da forma

$$P_{hj}(s, t) = \Pr(X(t) = j \mid X(s) = h),$$

que denotam a probabilidade de transição do estado h para o estado j no intervalo $(s, t]$. É importante ressaltar que um único elemento $P_{hj}(s, t)$ combina transições diretas e indiretas do estado h para o estado j . Em modelos markovianos em tempo contínuo e com espaço de estados finitos, dado $\mathbf{H}(t)$, a matriz de probabilidade de transição $\mathbf{P}(s, t)$ pode ser calculada pelo produto limite

$$\mathbf{P}(s, t) = \prod_{(s, t]} (\mathbf{I} + d\mathbf{H}(u)). \quad (4.2)$$

Em modelos clássicos de sobrevivência, em que consideramos somente a transição entre os estados vivo e morto, ou então entre algum estado pré determinado e a ocorrência do evento de interesse, podemos considerar um modelo de dois estados, definindo a função de sobrevivência, $S(t)$, como $S(t) = 1 - P_{12}(0, t) = P_{11}(0, t)$, em que 1 representa o estado inicial e 2 o evento de interesse.

Considere $N_{qi}(t)$ como o processo de contagem para eventos do tipo q para o sujeito i ($i = 1, \dots, n$), em que $q = 1, \dots, Q$ ($Q < K$). Dois ou mais eventos são considerados do mesmo “tipo” se eles possuírem risco basal comum. Dado que queremos analisar modelos de Cox em que as transições compartilham o mesmo risco basal, consideramos a teoria em termos dos tipos de transição ao invés das próprias transições. No caso geral, quando cada transição possui seu próprio risco basal ($Q = K$), cada tipo é equivalente a uma transição do estado g para o estado h . Nesses casos, a terminologia e notação adotadas retornam às usuais, ou seja, seguindo a equação (4.1), com a substituição de gh por q . No entanto, também é possível considerar modelos em que um “tipo” é definido por todas as transições a um certo estado provenientes de diferentes estados iniciais.

Na implementação da regressão de Cox, o “tipo” atuará como uma variável de estratificação e, na prática, podemos substituir o “tipo” por “estrato”. Esse fato não é relevante para análise, pois o estrato é definido pelos valores da covariável ou pelos grupos de transições, ao passo que cada estrato é determinado pelo seu próprio risco basal.

Nesse caso, definimos $Y_{qi}(t)$ como o processo em risco, sendo que $Y_{qi}(t) = 1$ se o indivíduo i está em risco para uma transição do tipo q no tempo t^- , tempo imediatamente anterior à t . Além disso, temos $N_q(t) = \sum_{i=1}^n N_{qi}(t)$ e $Y_q(t) = \sum_{i=1}^n Y_{qi}(t)$ como os processos que denotam respectivamente o número de eventos do tipo q até, inclusive, o tempo t e o tamanho do conjunto em risco para q no tempo t , que consiste de todos os sujeitos que podem experimentar uma transição q no tempo t .

Usualmente, os dados de estudos de sobrevivência são representados no formato de uma linha por sujeito, chamado de formato “wide”. Se este é o caso, os dados devem ser transformados para o formato “long”, em que cada linha do banco de dados representará uma transição, para que se adêquem ao cálculo dos estimadores dos modelos multiestados. Neste formato, cada sujeito terá número de linhas igual ao número de transições para os quais o sujeito está em risco. Um exemplo desse formato é apresentado na Tabela 17.

Tabela 17 – Exemplo de um banco de dados no formato “long” para multiestados.

id	from	to	trans	Tstart	Tstop	time	status	w	type	PR	w.1	w.2	w.3	age	T1
1	1	2	1	0	1	1	1	6	1	0	6	0	0	20	0
1	1	3	2	0	1	1	0	6	2	0	0	6	0	20	0
1	2	3	3	1	3	2	1	6	2	1	0	0	6	20	1
2	1	2	1	0	2	2	1	-2	1	0	-2	0	0	43	0
2	1	3	2	0	2	2	0	-2	2	0	0	-2	0	43	0
2	2	3	3	2	4	2	0	-2	2	1	0	0	-2	43	2
3	1	2	1	0	5	5	0	0	1	0	0	0	0	35	0
3	1	3	2	0	5	5	0	0	2	0	0	0	0	35	0

Fonte: [Wreede, Fiocco e Putter \(2010\)](#).

4.4.1 Intensidades de Transição: Modelos Não Paramétricos

Segundo [Wreede, Fiocco e Putter \(2010\)](#), em modelos com ausência de covariáveis, “tipo” e “transição” são equivalentes, pois o conjunto em risco para todas transições a partir do estado g são os mesmos e, para esse caso, escrevemos $Y_g(t)$ ao invés de $Y_{gh}(t)$. Como nesses modelos não temos os “parâmetros” associados às covariáveis, ou seja, $h_{gh}(t) = h_{gh0}(t)$, os chamamos de não paramétricos.

Nesses modelos o risco da transição $g \rightarrow h$ é estimado da seguinte maneira

$$d\hat{H}_{gh}(t) = \frac{dN_{gh}(t)}{Y_g(t)}, \quad (4.3)$$

em que $dN_{gh}(t)$ é definido como o número de transições do estado g para o estado h no tempo t . Somando todos os tempos de eventos até o tempo t obtemos o estimador de Nelson-Aalen $\hat{H}_{gh}(t)$ do risco acumulado para a transição $g \rightarrow h$.

Em analogia à análise de sobrevivência clássica, temos duas maneiras de calcular o erro padrão das estimativas do risco de transição: os estimadores de Aalen e Greenwood. O primeiro se baseia na suposição de que o verdadeiro risco acumulado é contínuo, enquanto que o segundo permite saltos.

A variância do estimador Aalen é calculada como

$$\widehat{\text{var}}(d\hat{H}_{gh}(t)) = \frac{dN_{gh}(t)}{Y_g^2(t)}, \quad \text{se } g \neq h$$

e

$$\widehat{\text{cov}}(d\hat{H}_{gh}(t), d\hat{H}_{g'h'}(t)) = 0, \quad \text{se } g \neq g' \text{ ou } g \neq h \neq h' \text{ se } g = g'.$$

Já o estimador Greenwood da variância é dado por

$$\widehat{\text{cov}}(d\hat{H}_{gh}(t), d\hat{H}_{g'h'}(t)) = \frac{(\delta_{hh'}Y_g(t) - dN_{gh}(t))dN_{g'h'}(t)}{Y_g^3(t)}, \quad \text{se } g \neq h, g \neq h,$$

em que $\delta_{hh'}$ é igual a 1 se $h = h'$ e 0 caso contrário, e

$$\widehat{\text{cov}}(d\hat{H}_{gh}(t), d\hat{H}_{g'h'}(t)) = 0, \quad \text{se } g \neq g'.$$

Nos dois casos as (co)variâncias remanescentes são definidas pela aditividade das (co)variâncias, sendo que a vantagem do estimador Greenwood é que produz erros padrão multinominais exatos quando não ocorre censura, por exemplo, $\widehat{\text{var}}P_{1j}(t) = P_{1j}(t)(1 - P_{1j}(t))/n$. No entanto, para grandes conjuntos de dados a diferença entre ambos é pequena.

4.4.2 Intensidades de Transição: Modelos Semi Paramétricos

O modelo de Cox é o modelo mais comum em análise de sobrevivência no caso em que covariáveis são consideradas (WREEDE; FIOCCO; PUTTER, 2010). Nesse caso, utilizamos uma variação tipo-específica do modelo de Cox comum (4.1), em que incluímos um vetor \mathbf{w} de covariáveis, é dado por,

$$h_q(t | \mathbf{w}) = h_{q0}(t) \exp(\boldsymbol{\gamma}^\top \mathbf{w}_q). \quad (4.4)$$

Nessa equação, $h_q(t | \mathbf{w})$ representa o risco para um evento do tipo q ($q = 1, \dots, Q$) de um indivíduo com vetor de covariáveis \mathbf{w} , h_0 é o risco basal para q , $\boldsymbol{\gamma}$ é o vetor de coeficientes de regressão para as covariáveis e \mathbf{w}_q um vetor de tamanho p das covariáveis tipo-específicas, derivado de \mathbf{w} . A maneira com que \mathbf{w}_q é definido a partir de \mathbf{w} depende especificamente do modelo estudado. A seguir, apresentaremos alguns exemplos.

A expressão apresentada em (4.4) é bem conhecida e conveniente, pois unifica vários tipos de modelos que são úteis para aplicações práticas e que são cobertos pela parte teórica e *softwares* disponíveis. Além disso, a formulação do modelo está intimamente relacionada ao formato de representação dos dados, formato *long*. Esse formato possui uma linha de cada indivíduo para cada transição ao qual ele está em risco, apresentando informações sobre a transição (a partir de qual estado, para qual estado e o número associado à transição), o tempo de entrada no estado atual e quando ele deixa de estar em risco no respectivo estado, e também uma variável indicadora de evento ou censura. Com relação às covariáveis, precisamos de uma coluna para cada transição.

Se considerarmos um modelo em que cada tipo corresponde a uma transição ($Q = K$), as covariáveis são fixas no baseline e tem o mesmo efeito para cada transição, o modelo (4.4) se torna igual a

$$h_q(t | \mathbf{w}) = h_{q0}(t) \exp(\boldsymbol{\gamma}^\top \mathbf{w}),$$

que é equivalente a $h_q(t | w) = h_{q0}(t) \exp(\gamma w)$ para apenas uma covariável. Nesse caso, a estimação se dá por meio do ajuste de um modelo de riscos proporcionais de Cox estratificado pela transição com efeito comum da covariável w .

Considerando o mesmo modelo, mas agora com covariáveis tipo-específicas, teremos

$$h_q(t | \mathbf{w}) = h_{q0}(t) \exp(\boldsymbol{\gamma}^\top \mathbf{w}_q). \quad (4.5)$$

Nesse caso, chamamos \mathbf{w}_q de covariável transição-específica. Além disso, esse tipo de covariável é um caso especial de covariáveis tipo-específicas e a teoria desenvolvida para esta última é utilizada para o caso transição-específica.

A equação apresentada em (4.4) descreve um modelo de Cox estratificado, em que cada estrato representa uma transição ou tipo. Embora o vetor de covariáveis \mathbf{w} não tenha efeito comum em todas as transições, as covariáveis transição-específicas \mathbf{w}_q possuem efeitos comuns, pois seus coeficientes $\boldsymbol{\gamma}$ são os mesmos para todas as transições.

Como exemplo, se considerarmos apenas um vetor básico \mathbf{w} e três possíveis transições, teremos três vetores de covariáveis transição-específica \mathbf{w}_1 , \mathbf{w}_2 e \mathbf{w}_3 , que são definidos como $\mathbf{w}_1 = (w, 0, 0)^\top$, $\mathbf{w}_2 = (0, w, 0)^\top$ e $\mathbf{w}_3 = (0, 0, w)^\top$, e um vetor de tamanho 3 para os coeficientes de regressão $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \gamma_3)^\top$. Por meio dessas definições e em conjunto com o modelo (4.5), temos as seguintes equações para os riscos de transição: $h_1(t | \mathbf{w}) = h_{10}(t) \exp(\gamma_1 \mathbf{w}_1)$, $h_2(t | \mathbf{w}) = h_{20}(t) \exp(\gamma_2 \mathbf{w}_2)$ e $h_3(t | \mathbf{w}) = h_{30}(t) \exp(\gamma_3 \mathbf{w}_3)$. Dessa forma, temos diferentes efeitos de \mathbf{w} nas transições 1, 2 e 3, com coeficientes de regressão iguais a γ_1 , γ_2 e γ_3 , respectivamente.

O modelo geral apresentado em (4.5) é equivalente a considerar um modelo de Cox para cada transição $g \rightarrow h$ separadamente, de acordo com a equação (4.6)

$$h_{gh}(t | \mathbf{w}) = h_{gh,0}(t) \exp(\boldsymbol{\gamma}_{gh}^\top \mathbf{w}). \quad (4.6)$$

em que $\boldsymbol{\gamma}_{gh}$ representa o vetor de coeficientes de regressão para a transição do estado g para o estado h .

Uma suposição adicional importante que é frequentemente feita, é que determinados riscos de base do modelo (4.4) são proporcionais, ou seja, as transições de $h \rightarrow j$ são proporcionais às transições de $k \rightarrow l$,

$$h_{gh,0}(t) = \tilde{\delta} h_{kl,0}(t),$$

que são os chamados modelos de riscos basais proporcionais. Nesse caso, as transições que compartilham o mesmo risco basal definem um tipo, o que significa que “transição” e “tipo” não são equivalentes nesse modelo e, portanto, teremos $Q < K$. Nesse caso, essa suposição reduz o número de parâmetros a serem estimados, diminuindo a complexidade do modelo multiestados.

Os coeficientes de regressão $\boldsymbol{\gamma}$ do modelo (4.4) são dados pelos valores que maximizam a função de log-verossimilhança parcial

$$\ell(\boldsymbol{\gamma}) = \sum_{q=1}^Q \sum_{i=1}^n \left\{ \int_0^\infty \boldsymbol{\gamma}^\top \mathbf{w}_{qi}(t) dN_{qi}(t) - \int_0^\infty \log S_q^{(0)}(\boldsymbol{\gamma}, t) dN_{qi}(t) \right\}$$

em que

$$S_q^{(0)}(\boldsymbol{\gamma}, t) = \sum_{i=1}^n \exp(\boldsymbol{\gamma}^\top \mathbf{w}_{qi}(t)) Y_{qi}(t).$$

O risco basal do evento tipo q , denotado por $h_{q0}(t, \boldsymbol{\gamma})$, é estimado por

$$d\hat{H}_{q0}(t, \hat{\boldsymbol{\gamma}}) = \frac{dN_q(t)}{S_q^{(0)}(\hat{\boldsymbol{\gamma}}, t)}, \quad (4.7)$$

sendo $\hat{\boldsymbol{\gamma}}$ o estimador de máxima verossimilhança de $\boldsymbol{\gamma}$. O estimador apresentado em (4.7) é chamado de estimador de Breslow e é uma versão ponderada do estimador não paramétrico apresentado em (4.3). Ainda, o estimador do risco basal acumulado $\hat{H}_{q0}(t, \hat{\boldsymbol{\gamma}})$ do evento tipo q é a soma nos tempos $u \leq t$ do tipo q das funções de risco estimadas na equação (4.7). A matriz de informação de Fisher $\mathcal{I}(\boldsymbol{\gamma})$ é dada por

$$\mathcal{I}(\boldsymbol{\gamma}) = \sum_{q=1}^Q \int_0^t V_q(\boldsymbol{\gamma}, t) dN_q(t),$$

com

$$V_q(\boldsymbol{\gamma}, t) = \frac{\mathbf{S}_q^{(2)}(\boldsymbol{\gamma}, t)}{\mathbf{S}_q^{(0)}(\boldsymbol{\gamma}, t)} - E_q(\boldsymbol{\gamma}, t) E_q^\top(\boldsymbol{\gamma}, t),$$

$$E_q(\boldsymbol{\gamma}, t) = \frac{\mathbf{S}_q^{(1)}(\boldsymbol{\gamma}, t)}{\mathbf{S}_q^{(0)}(\boldsymbol{\gamma}, t)},$$

$$\mathbf{S}_q^{(1)}(\boldsymbol{\gamma}, t) = \sum_{i=1}^n \mathbf{w}_{qi}(t) \exp(\boldsymbol{\gamma}^\top \mathbf{w}_{qi}(t)) Y_{qi}(t),$$

$$\mathbf{S}_q^{(2)}(\boldsymbol{\gamma}, t) = \sum_{i=1}^n \mathbf{w}_{qi}(t) \mathbf{w}_{qi}^\top(t) \exp(\boldsymbol{\gamma}^\top \mathbf{w}_{qi}(t)) Y_{qi}(t).$$

4.4.3 Estimação das Probabilidades de Transição e Erro Padrão

Assim como apresentado por [Wreede, Fiocco e Putter \(2010\)](#), para estimarmos as probabilidades de transição, dependentes ou não de covariáveis, e os respectivos erros padrão associados precisamos apenas conhecer as estimativas dos riscos de transição $\hat{\mathbf{H}}(t)$ e suas matrizes de variâncias-covariâncias $\widehat{\text{var}}(\hat{\mathbf{H}}(t))$. Para simplificar a apresentação dos procedimentos de estimação, consideraremos apenas a matriz de riscos $\hat{\mathbf{H}}$ ao invés de seus elementos H_{hj} individualmente.

Como $\hat{\mathbf{H}}$ é uma matriz, a matriz de covariâncias estimada é definida como $\widehat{\text{var}}(\text{vec}(\hat{\mathbf{H}}))$, em que $\text{vec}(\hat{\mathbf{H}})$ é definido como um vetor $(S^2 \times 1)$ composto pelas colunas de $\hat{\mathbf{H}}$ empilhadas, e a dimensão de $\widehat{\text{var}}(\hat{\mathbf{H}}(t))$ é igual a $S^2 \times S^2$.

Primeiramente, podemos estimar a matriz de probabilidades de transição a partir da seguinte equação

$$\hat{\mathbf{P}}(s, t) = \prod_{u \in (s, t]} (\mathbf{I} + d\hat{\mathbf{H}}(u)), \quad (4.8)$$

em que u indica todos os tempos em que os eventos ocorreram no intervalo $(s, t]$. Esse estimador é chamado de tipo-Aalen-Johanssen ([AALEN; JOHANSEN, 1978](#)). Note que a matriz de probabilidade de transição apresentada em (4.2) é calculada a partir do produto limite, enquanto que o estimador apresentado em (4.8) é baseado em um produto finito com alterações apenas nos tempos em que os eventos ocorreram.

As probabilidades de transição podem ser usadas para dois tipos de predição: “*forward*” e “*fixed horizon*”. A primeira fornece prognóstico futuro baseado no tempo atual e história do processo, isto é, em $\mathbf{P}(s, t)$ o tempo s se mantém fixo enquanto que o tempo t varia. Já na segunda, a previsão é feita a partir de vários pontos iniciais e apenas um ponto futuro, ou seja, t é fixo e s varia.

Os erros padrão das probabilidades estimadas podem ser usados para cálculo de intervalos de confiança das curvas de probabilidade. Nesse caso, as estimativas do erro padrão são calculados a partir da raiz quadrada dos elementos da diagonal da matriz de variância-covariância, dada por

$$\widehat{\text{var}}(\hat{\mathbf{P}}(s, t)), \quad (4.9)$$

sendo que os elementos $\widehat{\text{var}}(\hat{P}_{hj}, \hat{P}_{lm})$ dessa matriz $S^2 \times S^2$ possuem coordenadas $(S(j-1) + h, S(m-1) + l)$.

A matriz de variância-covariância apresentada em (4.8) é estimada pelo método delta, ao passo que o estimador (4.9) pode ser calculado de duas maneiras: diretamente ou por método recursivo. Além disso, existem dois possíveis métodos, Aalen e Greenwood, que dependem do estimador das (co)variâncias da função de risco acumulado utilizada. O tipo-Aalen pode ser calculado diretamente da seguinte maneira

$$\widehat{\text{var}}(\hat{\mathbf{P}}(s, t)) = \sum_{u \in (s, t]} \left\{ \hat{\mathbf{P}}(u, t)^\top \otimes \hat{\mathbf{P}}(s, u) \right\} \widehat{\text{var}}(d\hat{\mathbf{H}}(u)) \left\{ \hat{\mathbf{P}}(u, t) \otimes \hat{\mathbf{P}}(s, u)^\top \right\},$$

e o Greenwood calculado como

$$\widehat{\text{var}}(\hat{\mathbf{P}}(s, t)) = \sum_{u \in (s, t]} \left\{ \hat{\mathbf{P}}(u, t)^\top \otimes \hat{\mathbf{P}}(s, u-) \right\} \widehat{\text{var}}(d\hat{\mathbf{H}}(u)) \left\{ \hat{\mathbf{P}}(u, t) \otimes \hat{\mathbf{P}}(s, u-)^\top \right\}$$

em que \otimes representa o produto de Kronecker e $t-$ o último tempo de ocorrência de evento anterior a t . Os procedimentos para o cálculo recursivo, bem como as fórmulas recursivas para estimação nos métodos “*forward*” e “*fixed horizon*”, são apresentados com detalhes em [Wreede, Fiocco e Putter \(2010\)](#).

4.5 Modelo de Regressão Multiestados - Covariável Dependente do Tempo

Levando-se em conta o contexto usual de análise de sobrevivência, ou seja, com apenas um desfecho, o modelo de Cox (2.1) pode ser estendido para considerar covariáveis dependentes do tempo exógenas (apresentadas na Seção 2.2), usando a formulação de processo de contagem investigada por [Andersen e Gill \(1982\)](#) e apresentada em ([FLEMING; HARRINGTON, 1991](#)) e [Andersen et al. \(1993\)](#). A ideia da formulação desse modelo estendido é pensar na ocorrência dos eventos como realizações de um processo de Poisson muito lento. O modelo de Cox estendido, também conhecido como modelo de Andersen-Gill, é dado pela equação (4.10)

$$h_i(t | \mathcal{Y}_i(t), \mathbf{w}_i) = h_0(t) \exp \left\{ \boldsymbol{\gamma}^\top \mathbf{w}_i + \boldsymbol{\alpha} \mathbf{y}_i(t) \right\} \quad (4.10)$$

em que, assim como em (4.4), \mathbf{w}_i é o vetor de covariáveis de base e $\mathbf{y}_i(t)$ é o vetor de covariáveis dependentes do tempo, sendo que a interpretação de $\boldsymbol{\gamma}$ e $\boldsymbol{\alpha}$ com relação ao risco é semelhante.

Supondo que exista apenas uma covariável dependente do tempo no estudo, em um tempo qualquer t , temos que $\exp(\alpha)$ denota o aumento relativo no risco de um evento no tempo t dado pelo aumento de uma unidade em $y_i(t)$ no mesmo tempo t . Além disso, assim como apontado por (RIZOPOULOS, 2012b), dado que $y_i(t)$ é dependente do tempo o modelo (4.10) não assume que a razão de risco é constante no tempo.

Apesar do modelo de Cox estendido ser flexível, permitindo tanto covariáveis dependentes do tempo quanto truncamento à esquerda, múltiplas escalas de tempo, múltiplos eventos por sujeito e várias formas de modelos do tipo coorte (THERNEAU; GRAMBSCH, 2000), não é apropriado quando a covariável dependente do tempo é endógena (RIZOPOULOS, 2012b). Isso ocorre pelo motivo do modelo de Cox assumir que a covariável dependente do tempo é um processo previsível, medido sem erro e com trajetória completamente determinada. Especificamente, é assumido que os valores da covariável são alterados quando existe alguma avaliação do sujeito, sendo mantidos constantes entre os intervalos. Dessa forma, o modelo sugere que o risco de um evento, em um tempo t , é associado com o valor da covariável no mesmo tempo. No entanto, para muitas variáveis endógenas, assumir a suposição de função escada não caracteriza adequadamente a natureza da covariável. Por exemplo, podemos citar os casos em que biomarcadores são avaliados ao longo do tempo, em particular, quando se realiza uma análise desse tipo, utiliza-se a abordagem LVCF (*Last Value Carried Forward*), que considera o último valor observado da covariável que, segundo Prentice (1982), leva a estimativas e erros padrão severamente viesados.

Nesse contexto, Fisher e Lin (1999) enfatizam a importância do conhecimento da relação entre o desfecho e covariável analisada e também citam exemplos de como a forma da covariável pode variar ao longo do tempo. No geral, os exemplos apresentados avaliam a relação entre o fumo e morte, investigando se o hábito de fumar aumenta o risco de morte; a utilização de drogas redutoras do colesterol em indivíduos com altas taxas de gordura; a variação nos níveis do plasma sanguíneo em um período de 24 horas de uma droga que previne eventos vasculares e, por fim, transplante de medula óssea visando avaliar a sobrevivência dos pacientes após esse procedimento. Nesses casos, os autores citam a utilização de uma variável indicadora que varia de acordo com a característica estudada ao longo do tempo e que geralmente assume valor 1 quando uma determinada característica ocorre, como o hábito de fumar ou período de internação para tratamento do câncer, e 0 caso contrário.

A Tabela 18 apresenta, em termos práticos, a maneira mais comum de codificar covariáveis dependentes do tempo utilizando os tempos no formato $(T_{start}, T_{stop}]$. Destaca-se que as covariáveis fixas se mantêm constantes em todas as observações de um mesmo indivíduo, enquanto que os valores da covariável dependente do tempo é alterado a cada intervalo como, por exemplo, igual a 1,3 no intervalo $(0,15]$; 1,5 no intervalo $(15,45]$ e assim por diante. No

exemplo apresentado, o valor da creatinina no período inicial é 1,3; 1,5 no décimo quinto dia; 1,4 no quadragésimo sexto e 1,6 no septuagésimo terceiro dia. Nota-se que durante os intervalos é assumido que o valor da covariável se mantém constante, o que pode não faz sentido em casos específicos.

Tabela 18 – Exemplo de Banco de dados com covariável dependente do tempo para o contexto usual de análise de sobrevivência.

id	Tstart	Tstop	status	Idade	Creatinina
1	0	15	0	25	1,3
1	15	46	0	25	1,5
1	46	73	0	25	1,4
1	73	100	1	25	1,6

Fonte: [Therneau e Crowson \(2015\)](#).

Para a análise de uma covariável dependente do tempo em um contexto multiestados, propomos nesta Tese a extensão do banco de dados apresentado por [Therneau e Crowson \(2015\)](#) (Tabela 18) para o contexto de eventos múltiplos. Nesse caso, implementamos uma função que cria o banco de dados desejado conforme apresentado, já no contexto do parto, na Tabela 19. A partir desse novo banco de dados, seguimos com a estimação usual no modelo multiestados, conforme apresentado na Seção 4.4.1.

Vale destacar que o pacote `mstate` não realiza a expansão de covariáveis dependentes do tempo considerando o caso contínuo, justificando a contribuição da Tese.

4.6 Modelo Multiestados no Contexto do Trabalho de Parto

Nos modelos conjuntos que analisamos no Capítulo 3, consideramos apenas um desfecho, que é dado pelo parto por cesárea. No entanto, durante o trabalho de parto podemos observar diversas situações, tais como a realização de uma intervenção ou até mesmo o nascimento do bebê por parto vaginal ou cesárea, nos levando a uma situação de eventos múltiplos. Assim, dado que temos um caso com múltiplos eventos e seus tempos associados, podemos utilizar os modelos multiestados.

Nesta análise, construímos um modelo que leva em conta a realização de aumento, parto vaginal e por cesárea. A estrutura do modelo, bem como os possíveis estados foram definidos da seguinte maneira:

- Estado 1: A mulher permanece no estado inicial sem que seja feita nenhuma intervenção (Admissão);
- Estado 2: Realização de aumento (AOL);

- Estado 3: Realização de parto vaginal;
- Estado 4: Realização de parto por cesárea.

Devido à sua estrutura, o modelo multiestados é bastante útil nesse caso, pois pode nos dar um melhor entendimento sobre a relação entre os estados e o impacto entre eles para o desfecho do parto. Todos os pacientes iniciam o estudo no estado 1, que é definido pela chegada da mulher na unidade de saúde. A partir do estado inicial, é possível que a mulher mova para um estado transiente (estado 2), que representa a realização de aumento, ou então um estado absorvente (estados 3 e 4). Sendo um estado absorvente, quanto o paciente entra em um desses estados, não move para nenhum outro. Esse fato é determinado especificamente pelo nascimento do bebê e, conseqüentemente, esses estados têm essa característica. Em geral, o modelo possui cinco transições, conforme apresentado na Figura 17. Por meio da Tabela 20, que exhibe as frequências de transições a partir do estado inicial (admissão) e AOL, temos que, independentemente do estado, a maior frequência de transição é para o parto vaginal.

De maneira similar às análises anteriores, buscamos avaliar as transições do modelo incluindo informações como dilatação cervical, paridade prévia, status da membrana amniótica, comorbidades ou complicações pré gravidez e se a mulher em trabalho de parto foi referenciada de outra unidade de saúde. Além disso, o principal objetivo do estudo é incluir a dilatação cervical como covariável dependente do tempo, de modo que possamos avaliar a relação entre os estados com uma evolução dinâmica, de modo que se possa comparar os resultados esperados de acordo com evidências clínicas. Assim, analisamos em um primeiro modelo apenas se a dilatação na admissão é maior do que 4 cm, caracterizado como modelo de efeitos fixos via modelo de Cox padrão, e, em uma segunda abordagem, como dependente do tempo a partir de uma versão estendida do modelo de Cox para modelos multiestados, conforme apresentado na Seção 4.5.

4.6.1 Modelo Multiestados com Efeitos Fixos

Os resultados apresentados na Tabela 21 foram obtidos a partir do modelo que considera todas as variáveis como fixas, avaliadas na admissão. Particularmente, esse é o modelo que avalia a dilatação como maior ou menor ou igual a 4 cm. Observamos que em todas as transições pelo menos uma variável foi significativa ao nível de 5%. Considerando as estimativas de razão de risco, dadas por $\exp(\gamma)$, podemos destacar as seguintes influências em cada transição:

- Admissão → AOL: mulher com dilatação maior que 4 cm na admissão, membrana rompida, com ou sem mecônio, e referenciada por outra unidade de saúde tem maior risco de receber aumento, enquanto que mulher com paridade prévia, especialmente com 2 ou mais nascimentos anteriores, tem menor risco;

Tabela 19 – Banco de dados no formato “long” e com a covariável dilatação como dependente do tempo para o modelo multiestados.

id	from	to	trans	Tstart	Tstop	time	status	País	Dilatação
1	1	2	1	0,00	11,42	16,5	0	1	1
1	1	2	1	11,42	15,50	16,5	0	1	2
1	1	2	1	15,50	16,50	16,5	1	1	2
1	1	3	2	0,00	11,42	16,5	0	1	1
1	1	3	2	11,42	15,50	16,5	0	1	2
1	1	3	2	15,50	16,50	16,5	0	1	2
1	2	3	3	16,50	20,08	14,5	0	1	2
1	2	3	3	20,08	23,08	14,5	0	1	4
1	2	3	3	23,08	27,50	14,5	0	1	5
1	2	3	3	27,50	28,83	14,5	0	1	6
1	2	3	3	28,83	31,00	14,5	1	1	6
2	1	2	1	0,00	4,00	23,33	0	0	NA
2	1	2	1	4,00	12,67	23,33	0	0	3
2	1	2	1	12,67	16,67	23,33	0	0	3
2	1	2	1	16,67	20,67	23,33	0	0	4
2	1	2	1	20,67	23,33	23,33	1	0	5
2	1	3	2	0,00	4,00	23,33	0	0	NA
2	1	3	2	4,00	12,67	23,33	0	0	3
2	1	3	2	12,67	16,67	23,33	0	0	3
2	1	3	2	16,67	20,67	23,33	0	0	4
2	1	3	2	20,67	23,33	23,33	0	0	5

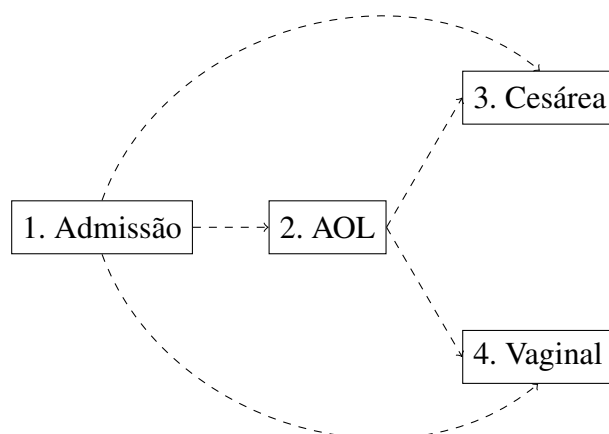


Figura 17 – Modelo multiestados com 4 estados e 5 transições.

Fonte: Elaborada pelo autor.

Tabela 20 – Frequência total para as transições do modelo multiestados.

	AOL	Cesárea	Vaginal
Admissão	1929 (31,8)	365 (6,0)	3765 (62,1)
AOL	-	250 (13,0)	1679 (87,0)

- Admissão → Cesárea: mulher com dilatação maior do que 4 cm, bem como com membrana amniótica rompida com mecônio apresenta maior risco de ter um bebê por cesárea diretamente da admissão. Por outro lado, multípara tem menor risco;
- Admissão → Vaginal: observamos que dilatação maior que 4 cm na admissão, membrana rompida sem mecônio e paridade prévia estão associados à maior risco de parto vaginal sem receber aumentação. Adicionalmente, presença de comorbidades ou complicações pré gravidez e referência da mulher em trabalho por outra unidade de saúde estão associados à menor risco de transição entre os estados;
- AOL → Cesárea: assumindo que a mulher recebeu aumentação, temos que dilatação maior que 4 cm na admissão está associado a um maior risco de parto por cesárea, enquanto que paridade prévia está associado a menor risco;
- AOL → Vaginal: para essa transição, observamos que dilatação maior que 4 cm na admissão, bem como paridade prévia, estão associados ao maior risco de transição. Adicionalmente, comorbidades ou complicações pré gravidez estão associados a menor risco de parto vaginal após realização de aumentação.

A partir da Figura 18 é possível visualizar as estimativas de razão de risco para as covariáveis em cada transição, incluindo o intervalo de confiança de 95% para cada estimativa. Nesse caso, valores de razão de risco maior que um indicam que há um maior risco de transição.

Apresentamos na Tabela 22 informações sobre quatro pacientes com diferentes características que podem ser usadas para fazer prognósticos do parto de acordo com o modelo ajustado, exibidas na Tabela 23. Nesse caso, a paciente A é uma paciente que chegou ao hospital com dilatação menor ou igual 4 cm, com membrana amniótica intacta, sem paridade prévia, sem comorbidades ou complicações pré gravidez e que não foi referenciada em trabalho de parto por outra unidade de saúde. A paciente B apresenta características similares à A, mas tem uma paridade prévia e dilatação maior que 4 cm na admissão. Temos que a paciente C representa um caso em que a mulher chega ao hospital com dilatação menor ou igual a 4 cm, com membrana rompida sem mecônio, uma paridade prévia, com comorbidades ou complicações pré gravidez e sem ser referenciada de outra unidade. Por fim, a paciente D é similar à C, mas sem comorbidades ou pré complicações e referenciada de outra unidade.

4.6.2 Modelo Multiestados com Covariável Dependente do Tempo

Na seção anterior consideramos o caso em que as covariáveis são fixas e avaliadas apenas no momento da admissão. No entanto, as mulheres foram avaliadas em diferentes períodos até o parto. Essas medidas foram feitas em períodos não fixos, avaliando-se também outras variáveis, tais como batimento cardíaco, número e duração das contrações, pressão arterial e posição

Tabela 21 – Coeficientes estimados para o modelo multiestados com efeitos fixos.

		γ	$\exp(\gamma)$	Erro padrão	p-valor
1. Admissão → AOL	Dilatação cervical > 4	0,385	1,470	0,053	<0,001
	Membrana amniótica rompida sem mecônio	0,663	1,940	0,052	<0,001
	Membrana amniótica rompida com mecônio	0,744	2,105	0,133	<0,001
	Para 1	-0,081	0,922	0,058	0,159
	Para 2+	-0,258	0,772	0,055	<0,001
	Comorbidade ou complicação pré gravidez	0,051	1,052	0,063	0,420
	Mulher em trabalho de parto referenciada por outra unidade de saúde	0,358	1,430	0,090	<0,001
2. Admissão → Cesárea	Dilatação cervical > 4	0,765	2,149	0,143	<0,001
	Membrana amniótica rompida sem mecônio	0,016	1,016	0,156	0,919
	Membrana amniótica rompida com mecônio	0,844	2,326	0,289	0,003
	Para 1	-0,433	0,648	0,141	0,002
	Para 2+	-0,714	0,490	0,143	<0,001
	Comorbidade ou complicação pré gravidez	0,135	1,144	0,147	0,359
	Mulher em trabalho de parto referenciada por outra unidade de saúde	0,401	1,494	0,221	0,069
3. Admissão → Vaginal	Dilatação cervical > 4	1,168	3,217	0,036	<0,001
	Membrana amniótica rompida sem mecônio	0,036	1,037	0,044	0,417
	Membrana amniótica rompida com mecônio	-0,264	0,768	0,135	0,051
	Para 1	0,448	1,565	0,042	<0,001
	Para 2+	0,497	1,644	0,039	<0,001
	Comorbidade ou complicação pré gravidez	-0,106	0,899	0,048	0,026
	Mulher em trabalho de parto referenciada por outra unidade de saúde	-0,256	0,774	0,089	0,004
4. AOL → Cesárea	Dilatação cervical > 4	0,496	1,643	0,172	0,004
	Membrana amniótica rompida sem mecônio	0,060	1,062	0,144	0,677
	Membrana amniótica rompida com mecônio	0,626	1,870	0,346	0,071
	Para 1	-0,556	0,573	0,184	0,002
	Para 2+	-0,659	0,518	0,198	<0,001
	Comorbidade ou complicação pré gravidez	-0,288	0,750	0,177	0,104
	Mulher em trabalho de parto referenciada por outra unidade de saúde	0,375	1,455	0,215	0,081
5. AOL → Vaginal	Dilatação cervical > 4	0,535	1,707	0,057	<0,001
	Membrana amniótica rompida sem mecônio	-0,042	0,958	0,056	0,445
	Membrana amniótica rompida com mecônio	0,160	1,174	0,144	0,268
	Para 1	0,347	1,415	0,061	<0,001
	Para 2+	0,512	1,668	0,058	<0,001
	Comorbidade ou complicação pré gravidez	-0,254	0,776	0,068	<0,001
	Mulher em trabalho de parto referenciada por outra unidade de saúde ^e	-0,073	0,929	0,100	0,460

Tabela 22 – Características dos pacientes A, B, C e D.

Covariável	Paciente A	Paciente B	Paciente C	Paciente D
Dilatação cervical > 4	Não	Sim	Não	Não
Membrana amniótica	Intacta	Intacta	Rompida sem mecônio	Rompida sem mecônio
Paridade prévia	Nulípara	Para 1	Para 1	Para 1
Comorbidade ou complicação pré gravidez	Não	Não	Sim	Não
Mulher em trabalho de parto referenciada por outra unidade de saúde	Não	Não	Não	Sim

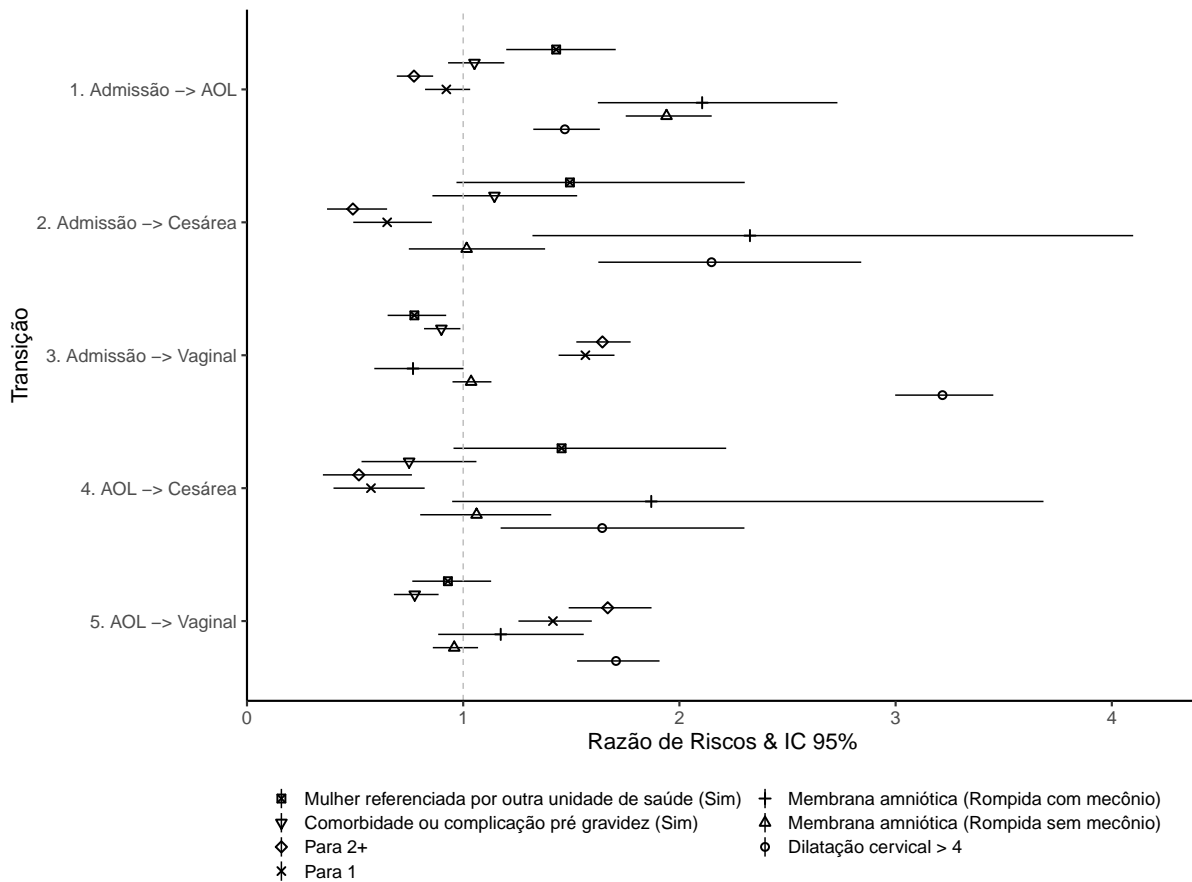


Figura 18 – Razão de risco para o modelo multiestados com efeitos fixos.

Fonte: Elaborada pelo autor.

Tabela 23 – Probabilidades de transição (%) em $t = 6, 12, 18, 24$ e $30h$ considerando o estado inicial admissão e tempo igual a $t = 0$.

	Paciente A					Paciente B				
	6h	12h	18h	24h	30h	6h	12h	18h	24h	30h
Permanece no estado inicial	59,6	31,4	16,7	9,2	5,0	19	1,9	0,2	0,0	0,0
Admissão → AOL	11,5	6,8	3,8	2,5	1,1	5,2	0,6	0,1	0,0	0,0
Admissão → Cesárea	2,1	7,8	12,2	14,8	16,7	1,6	2,8	3,0	3,0	3,1
Admissão → Vaginal	26,8	54,1	67,3	73,5	77,2	74,1	94,7	96,7	96,9	96,9
	Paciente C					Paciente D				
	6h	12h	18h	24h	30h	6h	12h	18h	24h	30h
Permanece no estado inicial	43,3	16,6	6,6	2,7	1,2	38,7	13,9	5,1	1,9	0,8
Admissão → AOL	18,5	9,7	4,8	2,7	1,2	21,2	8,7	3,5	1,7	0,6
Admissão → Cesárea	1,3	4,5	6,6	7,7	8,3	1,9	6,8	9,6	11	11,6
Admissão → Vaginal	36,9	69,3	82	86,9	89,4	38,3	70,7	81,7	85,4	87,0

fetal. Assim, apresentamos nesta seção o modelo que considera a dilatação como covariável dependente do tempo, conforme proposto na Seção 4.5.

Tabela 24 – Coeficientes estimados para o modelo multiestados com covariável dependente do tempo.

		γ	$\exp(\gamma)$	Erro padrão	p-valor
1. Admissão → AOL	Dilatação cervical	0,172	1,187	0,011	<0,001
	Membrana amniótica rompida sem mecônio	0,630	1,878	0,052	<0,001
	Membrana amniótica rompida com mecônio	0,687	1,989	0,133	<0,001
	* Para 1	-0,131	0,877	0,058	0,023
	Para 2+	-0,305	0,737	0,055	<0,001
	Comorbidade ou complicação pré gravidez	0,036	1,036	0,063	0,569
	Mulher em trabalho de parto referenciada por outra unidade de saúde	0,345	1,413	0,090	<0,001
2. Admissão → Cesárea	Dilatação cervical	0,300	1,350	0,022	<0,001
	Membrana amniótica rompida sem mecônio	0,011	1,011	0,156	0,943
	Membrana amniótica rompida com mecônio	0,791	2,205	0,285	0,006
	Para 1	-0,544	0,580	0,142	<0,001
	Para 2+	-0,770	0,463	0,143	<0,001
	Comorbidade ou complicação pré gravidez	0,098	1,103	0,147	0,505
	Mulher em trabalho de parto referenciada por outra unidade de saúde	0,351	1,421	0,220	0,111
3. Admissão → Vaginal	Dilatação cervical	0,745	2,107	0,008	<0,001
	Membrana amniótica rompida sem mecônio	-0,062	0,940	0,044	0,163
	* Membrana amniótica rompida com mecônio	-0,444	0,642	0,135	0,001
	† Para 1	0,066	1,068	0,042	0,122
	Para 2+	0,181	1,199	0,039	<0,001
	† Comorbidade ou complicação pré gravidez	-0,001	0,999	0,048	0,976
	† Mulher em trabalho de parto referenciada por outra unidade de saúde	-0,032	0,968	0,089	0,716
4. AOL → Cesárea	Dilatação cervical	0,097	1,102	0,032	0,002
	Membrana amniótica rompida sem mecônio	0,198	1,219	0,145	0,171
	* Membrana amniótica rompida com mecônio	0,690	1,994	0,347	0,046
	Para 1	-0,450	0,637	0,185	0,015
	Para 2+	-0,720	0,487	0,198	<0,001
	Comorbidade ou complicação pré gravidez	-0,178	0,837	0,177	0,314
	Mulher em trabalho de parto referenciada por outra unidade de saúde	0,397	1,488	0,213	0,063
5. AOL → Vaginal	Dilatação cervical	0,826	2,283	0,015	<0,001
	Membrana amniótica rompida sem mecônio	0,104	1,110	0,057	0,065
	Membrana amniótica rompida com mecônio	0,209	1,233	0,147	0,155
	Para 1	0,379	1,460	0,062	<0,001
	Para 2+	0,437	1,549	0,060	<0,001
	Comorbidade ou complicação pré gravidez	0,179	1,195	0,069	0,009
	Mulher em trabalho de parto referenciada por outra unidade de saúde	-0,186	0,831	0,101	0,066

Nota – Resultados diferentes em comparação com o modelo de efeitos fixos: * significante ao nível de 5%; † não significante ao nível de 5%.

Os resultados do modelo ajustado estão apresentados na Tabela 24 e Figura 19. A partir dos p-valores obtidos, observamos que, assim como no caso anterior, em todas as transições pelo menos uma variável foi significativa ao nível de 5%. Entretanto, analisando os coeficientes e tendo em conta a formulação do modelo, temos interpretações diferentes para o impacto da dilatação. Além disso, verificamos mudanças para paridade prévia, status da membrana amniótica, comorbidades e complicações pré gravidez e referência da mulher de outra unidade de saúde.

Notamos que a evolução da dilatação cervical é responsável por um aumento em todos os riscos de transição. Portanto, quanto maior a dilatação, maior o risco de transitar entre

os estados, demonstrando a importância da evolução cervical para o parto. Ainda, é possível destacar a diferença em relação ao modelo anterior para paridade prévia, com um nascimento anterior, na primeira transição, que agora é significativa, e para a mesma variável na transição da admissão para o parto vaginal, que deixou de ser significativa. Observamos, também, que comorbidades ou complicações pré gravidez e referência da mulher por outra unidade de saúde se tornaram significantes para o último caso citado. Além disso, verificamos que membrana amniótica rompida com mecônio se tornou significativa em duas transições, a partir da admissão para parto vaginal e de aumentação para cesárea.

Finalmente, em comparação com o modelo proposto na Seção 4.6.1, é importante destacar o impacto reverso de comorbidades ou complicações pré gravidez na transição de aumentação para parto vaginal. Nesse caso, a presença dessa característica causa um aumento no risco de transição.

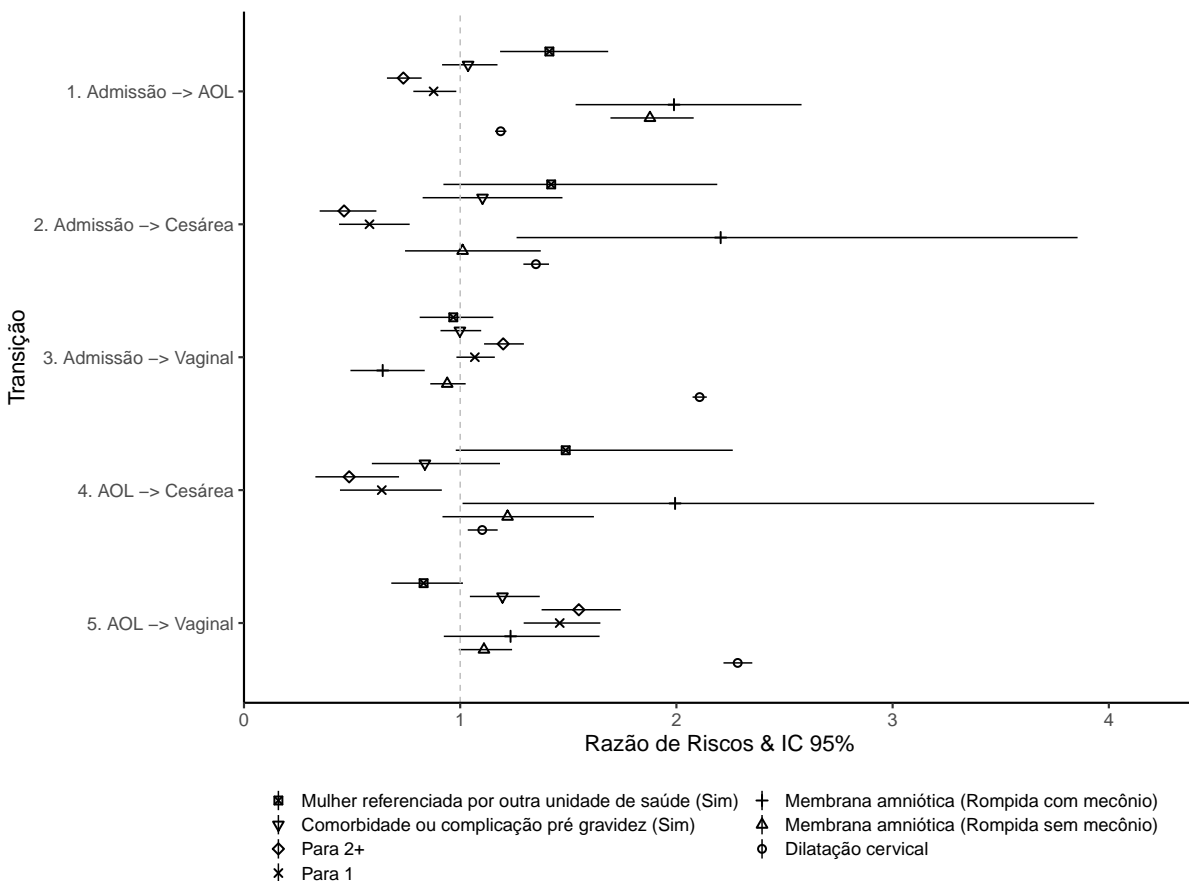


Figura 19 – Razão de risco para o modelo multiestados com covariável dependente do tempo.

Fonte: Elaborada pelo autor.

4.7 Comentários Finais

Nesse capítulo definimos o conceito de modelos multiestados, apresentamos as probabilidades e intensidades de transição, alguns exemplos de estruturas dos modelos, como as covariáveis podem ser consideradas e os procedimentos de estimação para o caso markoviano. Além disso, exibimos como covariáveis dependente do tempo podem ser consideradas a partir do modelo de Cox estendido para cada transição, apenas manipulando a forma de tratar os dados.

No contexto do trabalho de parto, estes resultados possibilitaram relacionar a realização de aumento (uso de oxitocina durante o trabalho de parto) com o tipo de parto, vaginal e cesariana, considerando a dilatação dependente do tempo em uma estrutura de eventos múltiplos. Nesse caso, fizemos uma comparação de um modelo de regressão multiestado considerando todas as covariáveis fixas no momento da admissão, com um modelo de regressão multiestado com uma covariável dependente do tempo (dilatação).

A seguir, apresentamos as considerações finais sobre a Tese, bem como algumas sugestões para trabalhos futuros.

CONSIDERAÇÕES FINAIS

O principal objetivo desta Tese, motivado pelo contexto do trabalho de parto, é propor um modelo que incorpore funções de risco complexas para a análise de dados de longitudinais e de sobrevivência. Geralmente, a modelagem conjunta utiliza funções paramétricas como Weibull, log-normal, gama e Gompertz, ou então métodos não paramétricos como splines. Baseando-se nas distribuições paramétricas, temos funções de risco constantes e monótonas crescentes ou decrescentes para a Weibull, Gompertz e gama, enquanto que a log-normal apresenta formas unimodais.

Considerando a Weibull, uma das distribuições mais usadas em análise de sobrevivência, por exemplo, temos que sua função de risco não é capaz de incorporar/acomodar função de risco não monótonas, como unimodal ou em formato de banheira. Sendo assim, nesta Tese propomos a avaliação da utilização de possíveis generalizações com três parâmetros, que são capazes incorporar as funções de risco citadas, tais como a distribuição gama generalizada (GG), com utilização de dois tipos de parametrizações, propostas por Stacy e Prentice, Weibull exponenciada (EW), Weibull generalizada (GW), Marshall-Olkin-Weibull (MOW) e distribuição de potência Weibull generalizada (GPW).

Neste mesmo contexto, propomos na modelagem conjunta a utilização da distribuição Nakagami- m , que apresenta função de risco crescente, decrescente, unimodal e em formato de banheira. Nesse caso, essa distribuição tem a vantagem de apresentar apenas dois parâmetros, tendo como casos particulares as distribuições Rayleigh e half-normal, além de estar relacionada com o modelo gama.

Como a distribuição gama generalizada é uma família paramétrica que tem como casos particulares as distribuições exponencial, Weibull, log-normal e gama, propomos um procedimento geral para selecionar um desses modelos considerando o teste de razão de verossimilhanças (TRV). Nesse caso, o teste busca comparar, em termos de discriminação, se o tempo de falha pode ser descrito por uma distribuição fixada na hipótese nula versus distribuições candidatas ao ajuste

na hipótese alternativa. Além disso, avaliamos a seleção por meio de critérios de informação, como AIC e BIC, destacando-se que tanto a seleção via família GG quanto a seleção por critérios de informação são inéditos na literatura da modelagem conjunta de dados longitudinais e de sobrevivência.

Para maximizar a logaritmo da função de verossimilhança dos modelos propostos, consideramos o método via regra pseudo-adaptativa de Gauss-Hermite, que se baseia em usar informações sobre o formato do integrando para o cálculo da quadratura ajustando-se um modelo misto separadamente para o componente longitudinal no início do algoritmo de estimação. Além disso, com a finalidade de verificar a eficiência desse método na estimação dos parâmetros das distribuições propostas, realizamos estudos de simulação para um cenário específico, variando o tamanho da amostra. Replicamos o estudo por 500 vezes de acordo com amostras de 250, 500 e 1000 observações e calculamos medidas de interesse como viés relativo, raiz do erro quadrático médio e probabilidades de cobertura utilizando intervalos de 95% de confiança via *bootstrap*. Por meio dos cenários simulados, observamos aproximações satisfatórias para as estimativas em todos os modelos estudados, sendo que, em geral, os maiores desvios em relação ao verdadeiro valor do parâmetro foram para os parâmetros vinculados à função de risco. Além disso, constatamos problema de identificabilidade nas distribuições GG, EW e GPW.

Avaliamos a performance dos critérios de seleção dos modelos paramétricos usuais a partir do teste de razão de verossimilhanças via família GG e critérios de informação como AIC e BIC. Nesse estudo, geramos 1000 amostras de tamanhos iguais a 250, 500 e 1000 para cada distribuição e comparamos o poder dos métodos em selecionar a verdadeira distribuição. Por meio dos resultados obtidos, destacamos que a utilização do teste de razão de verossimilhanças via família GG apresenta níveis de significância próximos do nível nominal, com valores entre 5% e 10% principalmente para os modelos Weibull e gama. Além disso, esse teste tem alto poder para rejeitar a distribuição log-normal quando os dados de falha são gerados pelas outras duas distribuições. No entanto, notamos que há um confundimento em discriminar as distribuições Weibull e gama e, quando geramos dados da log-normal, existe uma grande dificuldade de convergência para o modelo GG. Nesse caso, a própria dificuldade de convergência pode sugerir a utilização do modelo log-normal, que pode ser confirmado utilizando-se os critérios de informação, cujos resultados de discriminação, especialmente para grandes tamanhos de amostra, foram satisfatórios para as três distribuições usuais.

Os modelos estudados foram ajustados aos dados do trabalho de parto, sendo que um modelo conjunto para o tempo até cesárea com inclusão da evolução dinâmica da dilatação foi proposto. Além da dilatação, utilizamos informações sobre paridade prévia, status da membrana amniótica, presença de comorbidade ou complicações pré gravidez e se a mulher em trabalho de parto foi referenciada por outra unidade de saúde. A componente longitudinal, dada pela evolução dinâmica da dilatação, foi modelada a partir de uma tendência linear no tempo com informações sobre paridade prévia, enquanto que o modelo de sobrevivência incluiu todas as covariáveis

descritas no estudo. Como todas as mulheres que tiveram parto normal estão incluídas na análise, devemos lembrar que o modelo desenvolvido pode ser interpretado como um modelo de causa específica, contrastando mulheres que tiveram parto por cesárea com aquelas que tiveram parto normal.

Inicialmente, consideramos as generalizações da Weibull para a função de risco. Nesse caso, é importante citar que o algoritmo de maximização da log-verossimilhança apresentou instabilidade quando usamos a parametrização de Prentice na GG pelo fato de o parâmetro q ter convergido para zero. Quando q tende a zero, a quantidade $\gamma(q^{-2})$ tende ao infinito, resultando em instabilidade. Dessa forma, aplicamos a seguinte transformação $q^{-2} = q^*$ nos parâmetros. Por meio dos resultados obtidos, observamos que a distribuição GG com parametrização de Prentice apresentou os melhores resultados devido ao maior valor da log-verossimilhança e menores AIC e BIC. Também, notamos que a distribuição EW apresentou performance bastante similar, com resultados ligeiramente inferiores. Para avaliarmos a performance dos modelos propostos em comparação a um método não paramétrico, ajustamos um modelo com função de risco baseada em B-splines com 5 nós internos. Por meio dos resultados, observamos que maioria dos modelos propostos apresentaram melhor qualidade de ajuste principalmente se compararmos o BIC, que penaliza modelos mais complexos.

Utilizando a mesma base de dados e considerando-se a família de distribuição GG para fazer seleção de modelos via TRV e escolher entre as possíveis distribuições usuais (Weibull, log-normal e gama), observamos que a suposição dos modelos Weibull e gama foram rejeitadas ao nível de 5% de significância, enquanto que a função de risco log-normal se destacou pelo ajuste mais adequado aos dados. Além disso, destacamos que a função de risco log-normal também apresentou melhores resultados nos critérios de informação AIC e BIC, com menores valores em comparação com os outros modelos. Por fim, utilizamos a distribuição Nakagami- m , que apresentou performance superior à Gompertz, similar à Weibull e inferiores à log-normal e gama.

Os estudos condensados na Seção 3.8 mostraram que todas as covariáveis são significantes para o modelo longitudinal de dilatação, com um impacto positivo de paridade prévia, indicando que mulheres que já tiveram filhos tendem a estar em um nível maior de dilatação do que as que nunca tiveram filho. Analisando separadamente o modelo de sobrevivência, destacamos que a paridade prévia, o referenciamento por outra unidade de saúde e o status da membrana amniótica são significativas para o tempo até cesárea. Em geral, mulheres múltiparas tendem a apresentar menor risco de cesárea, enquanto que mulheres transferidas de outra unidade de saúde e com membrana rompida tendem a apresentar maior risco. Também, constatamos que existe um impacto significativo para a evolução da dilatação no evento de interesse, e que a única diferença com relação à interpretação das variáveis significativas foi para membrana rompida sem mecônio. Nesse caso, o limite inferior do intervalo de confiança para o modelo GG conteve o valor zero, enquanto que para log-normal e Nakagami não.

No modelo conjunto consideramos apenas um desfecho que é o parto por cesárea. No entanto, temos registros sobre desde a admissão no hospital até o nascimento bebê, incluindo informações sobre intervenções e nascimento por parto vaginal ou cesárea. Nesta situação temos a ocorrência de eventos múltiplos e, para associar esses eventos a seus tempos, utilizamos os modelos multiestados. Os modelos de múltiplos estados são definidos como modelos para um processo estocástico que a qualquer momento ocupa um conjunto finito de estados e, em geral, são os modelos mais comuns para descrever a evolução longitudinal do tempo de falha. Uma das grandes vantagens desses modelos é que são úteis para responder diversas questões no contexto de análise de sobrevivência que não podem ser respondidas pelos modelos usuais, como estimar intensidades e probabilidades de transição entre os estados, que neste caso seriam os tempos entre as intervenções.

Em geral, ao introduzirmos covariáveis no modelo multiestados, consideramos que seus valores são constantes. Entretanto, nosso objetivo nesta Tese é utilizar modelos que consideram covariáveis dependentes do tempo em sua construção. Assim, estendemos de maneira simplista a aplicação do modelo multiestados para considerar o modelo de Cox estendido em cada transição, manipulando a forma de tratar os dados e considerando os métodos usuais já disponíveis no pacote `mstate`. Com isso, propomos um modelo com quatro estados para a análise dos dados de trabalho de parto: admissão, realização de intervenção, parto vaginal e parto cesárea. Todos os pacientes iniciam o estudo no admissão, que é definido pela chegada da mulher na unidade de saúde. A partir do estado inicial, é possível que a mulher mova para um estado transiente, que representa a realização da amniotomia, que é caracterizada pelo uso de medicação para acelerar o trabalho de parto, ou então um estado absorvente, dado pelo parto vaginal ou cesárea.

Neste estudo, utilizamos as mesmas variáveis do modelo conjunto, com dois tratamentos para a dilatação: consideramos um caso de efeitos fixos, levando em conta se a dilatação era maior que 4 cm no momento da admissão, e um modelo com evolução dinâmica, como dependente do tempo. No primeiro caso, observamos que em todas as transições pelo menos uma variável foi significativa ao nível de 5% de significância. Assumindo que a mulher chegou com dilatação maior que 4 cm, com membrana rompida, com ou sem mecônio, e foi referenciada por outra unidade de saúde, teremos maior risco de realização de intervenção, enquanto que mulheres com paridade prévia, especialmente com dois ou mais nascimentos anteriores, tem menor risco. Para cesárea, sem amniotomia, notamos que o risco é maior para uma mulher com dilatação maior que 4 cm, bem como com membrana rompida com mecônio. Por outro lado, múltiparas tendem a apresentar menor risco. Levando em conta o parto vaginal, também sem amniotomia, notamos que risco cresce de acordo com dilatação maior que 4 cm na admissão, membrana rompida sem mecônio e paridade prévia, enquanto que a presença de comorbidades ou complicações pré gravidez, e se a mulher foi referenciada por outra unidade de saúde o diminuem. Tendo como base a cesárea, dado que foi realizada uma amniotomia, observamos que dilatação maior que 4 cm está associada a um maior risco de transição, enquanto que paridade prévia está relacionada a um menor risco. Relacionando-se parto vaginal e amniotomia, temos que dilatação maior que

4 cm, bem como paridade prévia estão associado a um maior risco, enquanto que comorbidades ou complicações pré gravidez diminuem o risco.

Assumindo a dilatação como dependente do tempo, notamos que quanto maior seu valor, maior o risco de transição entre os estados, mostrando sua importância para a evolução do parto principalmente no caso vaginal. Além disso, destacamos que paridade prévia com um nascimento prévio se tornou significativa na transição de amissão para aumento e membrana rompida com mecônio na transição de admissão para parto vaginal e também de aumento para cesárea. Por outro lado, paridade prévia com um nascimento anterior, comorbidade ou complicação pré gravidez e se a mulher foi transferida de unidade de saúde se tornaram insignificantes para a transição entre admissão e parto vaginal. Nesse caso, destaca-se o impacto reverso de comorbidade ou complicação prévia na transição de aumento para parto vaginal. Sob essas circunstâncias, a presença dessa característica causa um aumento no risco de transição

É notório que o modelo de Cox estendido para o caso multiestados mostrou melhores resultados em termos de evidências clínicas, alterando os coeficientes e significância de variáveis essenciais como, por exemplo, membrana amniótica rompida com presença de mecônio, que leva a ocorrência de cesárea. Assim como mencionado, esse fato ficou mais evidente quando usamos o modelo multiestados com covariável dependente do tempo proposto. Também, a transição de admissão para parto vaginal foi bastante afetada pela covariável dependente do tempo, evidenciando a importância da evolução da dilatação para o parto vaginal sem que haja aumento. Um outro ponto interessante, é a relação entre os resultados do modelo conjunto usando função de risco GG com o modelo multiestados, cujo valores do impacto de membrana rompida sem mecônio foram similares, indicando que esse fato não é tão significativo para cesárea como mostrado nos modelos log-normal e Nakagami.

Portanto, concluímos que é importante considerar modelos de sobrevivência com covariáveis dependente do tempo no contexto do trabalho de parto. Os modelos conjuntos paramétricos capazes de incorporar funções de risco com formas complexas, bem como o procedimento geral de seleção de modelos usuais via família GG, foram extremamente úteis para explicar como a evolução dinâmica da dilatação e outras variáveis fixas como, por exemplo, paridade prévia e status da membrana amniótica estão relacionadas ao tempo até cesárea, permitindo-se que médicos e enfermeiras possam atuar de maneira mais efetiva. Ainda, observamos que é possível utilizarmos modelos multiestados com covariável dependente do tempo para analisarmos no trabalho de parto de forma efetiva através da adequação do formato do banco de dados, sendo que esta modelagem apresentou resultados importantes para a problemática do trabalho de parto, com resultados similares ao esperado na prática clínica.

Finalmente, os modelos propostos, estudados e implementados nesta Tese podem ser utilizados em forma de algoritmo na construção de uma ferramenta de monitoração do parto para o fácil acesso em uma unidade de saúde, ou então disponibilizados como uma aplicação *mobile* para auxiliar médicos e enfermeiras de maneira eficiente, melhorando, assim, a qualidade e o

cuidado intra parto.

5.1 Sugestões para Trabalhos Futuros

Nesta Tese, utilizamos o modelo conjunto para apenas um desfecho, que foi a cesárea. No entanto, devido à observação dos dois tipos de parto, vaginal e por cesárea, esta estrutura pode ser considerada em modelos de riscos competitivos comparando o risco de ambos, conforme apresentado por [Elashoff, Li e Li \(2008\)](#), [Williamson *et al.* \(2008\)](#), [Hu, Li e Li \(2009\)](#), [Huang *et al.* \(2011\)](#). Neste estudo consideramos a dilatação como covariável dependente do tempo predita por um modelo com tendência linear. Entretanto é possível estender a utilização tanto para outros tipos de tendência como também considerar mais componentes longitudinais como a contração, pressão arterial, batimentos cardíacos do bebê, e outras variáveis registradas durante o trabalho de parto. Uma descrição sobre os trabalhos recentes com a utilização de múltiplos eventos longitudinais é apresentada por [Hickey *et al.* \(2016\)](#).

Outra possibilidade, é incluir o tipo de censura à esquerda ([PERICLEOUS, 2016](#); [KRÓL *et al.*, 2016](#)) ou então desenvolver um modelo conjunto com as mesmas distribuições propostas nesta Tese que leva em conta tanto indivíduos com medidas longitudinais quanto sem ([SOTO, 2014](#)). Podemos, também, investigar o efeito de distribuições defeituosas no contexto de fração de cura com covariável dependente do tempo ([ROCHA, 2016](#)) e utilizar um modelo longitudinal com erro de medida ([FULLER, 2009](#)). Nesse contexto, destacamos a forma de avaliação da dilatação, que pode ser subjetiva, indicando que o verdadeiro valor da dilatação não é medido corretamente no momento da avaliação.

Outro ponto a ser destacado é que, apesar da inclusão da covariável dependente do tempo ter sido essencial no modelo multiestados, o modelo de Cox estendido não é apropriado quando a variável é endógena ([RIZOPOULOS, 2012b](#)). Também, esse método usa a abordagem LVCF (*Last Value Carried Forward*), que considera o último valor observado para a covariável dependente do tempo. No entanto, quando a dilatação foi medida no mesmo tempo de evento, perdemos essa informação devido à essa abordagem. Além disso, se a mulher teve a primeira avaliação em um tempo após a admissão, os tempos iniciais não são considerados, pois se tornam valores faltantes, implicando em sua remoção.

Quando temos um processo longitudinal caracterizado por uma variável dependente do tempo em um processo multiestados, podemos usar os modelos conjuntos multiestados, que relacionam os dois processos segundo uma estrutura inerente a ambos ([TSIATIS; DAVIDIAN, 2004](#); [RIZOPOULOS, 2012b](#)). Sendo assim, trabalhos futuros devem considerar a relação entre o processo longitudinal e multiestados, conforme apresentado por [Ferrer *et al.* \(2016\)](#). Dado que o modelo multiestados conjunto foi pouco explorado, podemos utilizar as mesmas distribuições utilizadas nessa Tese para avaliar múltiplos desfechos.

REFERÊNCIAS

AALEN, O. O.; JOHANSEN, S. An empirical transition matrix for non-homogeneous markov chains based on censored observations. **Scandinavian Journal of Statistics**, JSTOR, p. 141–150, 1978. Citado na página [86](#).

AKAIKE, H. A new look at the statistical model identification. **IEEE transactions on automatic control**, Ieee, v. 19, n. 6, p. 716–723, 1974. Citado na página [39](#).

ANDERSEN, P. K.; BORGAN, O.; GILL, R. D.; KEIDING, N. **Statistical models based on counting processes**. [S.l.]: Springer, 1993. Citado nas páginas [81](#) e [87](#).

ANDERSEN, P. K.; GILL, R. D. Cox's regression model for counting processes: a large sample study. **The annals of statistics**, JSTOR, p. 1100–1120, 1982. Citado na página [87](#).

ANDERSEN, P. K.; KEIDING, N. Multi-state models for event history analysis. **Statistical methods in medical research**, SAGE Publications, v. 11, n. 2, p. 91–115, 2002. Citado na página [78](#).

BREUSCH, T. S. Hypothesis testing in unidentified models. **The Review of Economic Studies**, Wiley-Blackwell, v. 53, n. 4, p. 635–651, 1986. Citado na página [64](#).

BROWN, E. R.; IBRAHIM, J. G.; DEGRUTTOLA, V. A flexible b-spline model for multiple longitudinal biomarkers and survival. **Biometrics**, Wiley Online Library, v. 61, n. 1, p. 64–73, 2005. Citado na página [22](#).

CAO, G.; TODEM, D.; YANG, L.; FINE, J. P. Evaluating statistical hypotheses using weakly-identifiable estimating functions. **Scandinavian Journal of Statistics**, Wiley Online Library, v. 40, n. 2, p. 256–273, 2013. Citado na página [64](#).

COMMENGES, D.; JACQMIN-GADDA, H. **Dynamical Biostatistical Models**. [S.l.]: CRC Press, 2015. v. 86. Citado na página [77](#).

CORDEIRO, G. M.; LEMONTE, A. J. On the marshall–olkin extended weibull distribution. **Statistical papers**, Springer, v. 54, n. 2, p. 333–353, 2013. Citado na página [46](#).

COX, C.; CHU, H.; SCHNEIDER, M. F.; MUÑOZ, A. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. **Statistics in medicine**, Wiley Online Library, v. 26, n. 23, p. 4352–4374, 2007. Citado na página [45](#).

COX, C.; MATHESON, M. A comparison of the generalized gamma and exponentiated weibull distributions. **Statistics in medicine**, Wiley Online Library, v. 33, n. 21, p. 3772–3780, 2014. Citado na página [68](#).

COX, D. Regression models and life-tables. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 34, n. 2, p. 87–22, 1972. Citado nas páginas [30](#) e [81](#).

- CROWTHER, M. J.; ABRAMS, K. R.; LAMBERT, P. C. Flexible parametric joint modelling of longitudinal and survival data. **Statistics in Medicine**, Wiley Online Library, v. 31, n. 30, p. 4456–4471, 2012. Citado na página 22.
- DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. **Journal of the royal statistical society. Series B (methodological)**, JSTOR, p. 1–38, 1977. Citado na página 34.
- EFRON, B. Bootstrap methods: Another look at the jackknife. **Annals of Statistics**, p. 1–26, 1979. Citado nas páginas 39 e 40.
- ELASHOFF, R. M.; LI, G.; LI, N. A joint model for longitudinal measurements and survival data in the presence of multiple failure types. **Biometrics**, Wiley Online Library, v. 64, n. 3, p. 762–771, 2008. Citado na página 104.
- FERRER, L.; RONDEAU, V.; DIGNAM, J.; PICKLES, T.; JACQMIN-GADDA, H.; PROUST-LIMA, C. Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer. **Statistics in Medicine**, 2016. Citado na página 104.
- FISHER, L. D.; LIN, D. Y. Time-dependent covariates in the cox proportional-hazards regression model. **Annual review of public health**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 20, n. 1, p. 145–157, 1999. Citado nas páginas 31 e 88.
- FLEMING, T. R.; HARRINGTON, D. P. **Counting Processes and Survival Analysis**. [S.l.]: John Wiley & Sons, 1991. Citado na página 87.
- FOX, P. A.; HALL, A.; SCHRYER, N. L. The port mathematical subroutine library. **ACM Transactions on Mathematical Software (TOMS)**, ACM, v. 4, n. 2, p. 104–126, 1978. Citado na página 35.
- FUKUMIZU, K. *et al.* Likelihood ratio of unidentifiable models and multilayer neural networks. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 31, n. 3, p. 833–851, 2003. Citado na página 64.
- FULLER, W. A. **Measurement error models**. [S.l.]: John Wiley & Sons, 2009. v. 305. Citado na página 104.
- GAY, D. M. Usage summary for selected optimization routines. **Computing science technical report**, v. 153, p. 1–21, 1990. Citado na página 35.
- GOMPERTZ, B. On the nature of the function expressive of the law of human mortality and on the new mode of determining the value of life contingencies. **Philosophical transactions of the Royal Society of London**, The Royal Society, v. 115, p. 513–583, 1825. Citado na página 73.
- HICKEY, G. L.; PHILIPSON, P.; JORGENSEN, A.; KOLAMUNNAGE-DONA, R. Joint modelling of time-to-event and multivariate longitudinal outcomes: recent developments and issues. **BMC medical research methodology**, BioMed Central, v. 16, n. 1, p. 117, 2016. Citado na página 104.
- HINKLEY, D. V.; COX, D. **Theoretical statistics**. [S.l.]: Chapman and Hall/CRC, 1979. Citado nas páginas 37 e 38.
- HOUGAARD, P. Multi-state models: a review. **Lifetime data analysis**, Springer, v. 5, n. 3, p. 239–264, 1999. Citado na página 23.

- HU, W.; LI, G.; LI, N. A bayesian approach to joint analysis of longitudinal measurements and competing risks failure time data. **Statistics in medicine**, Wiley Online Library, v. 28, n. 11, p. 1601–1619, 2009. Citado na página 104.
- HUANG, X.; LI, G.; ELASHOFF, R. M.; PAN, J. A general joint model for longitudinal measurements and competing risks survival data with heterogeneous random effects. **Lifetime data analysis**, Springer, v. 17, n. 1, p. 80–100, 2011. Citado na página 104.
- KALBFLEISCH, J. D.; PRENTICE, R. L. **The Statistical Analysis of Failure Time Data**. 2nd ed. [S.l.]: John Wiley & Sons, 2002. Citado na página 30.
- KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American statistical association**, Taylor & Francis, v. 53, n. 282, p. 457–481, 1958. Citado na página 29.
- KRÓL, A.; FERRER, L.; PIGNON, J.-P.; PROUST-LIMA, C.; DUCREUX, M.; BOUCHÉ, O.; MICHIELS, S.; RONDEAU, V. Joint model for left-censored longitudinal data, recurrent events and terminal event: Predictive abilities of tumor burden for cancer evolution with application to the ffcd 2000–05 trial. **Biometrics**, Wiley Online Library, v. 72, n. 3, p. 907–916, 2016. Citado na página 104.
- LANGE, K. **Optimization**. [S.l.]: Springer, 2004. Citado nas páginas 30 e 34.
- LAWLESS, J. F. **Statistical models and methods for lifetime data**. [S.l.]: John Wiley & Sons, 2011. v. 362. Citado na página 27.
- LAWN, J. E.; BLENCOWE, H.; OZA, S.; YOU, D.; LEE, A. C.; WAISWA, P.; LALLI, M.; BHUTTA, Z.; BARROS, A. J.; CHRISTIAN, P. *et al.* Every newborn: progress, priorities, and potential beyond survival. **The Lancet**, Elsevier, v. 384, n. 9938, p. 189–205, 2014. Citado na página 19.
- LIU, Q.; PIERCE, D. A. A note on gauss-hermite quadrature. **Biometrika**, Oxford University Press, v. 81, n. 3, p. 624–629, 1994. Citado na página 35.
- LOUZADA-NETO, F.; MAZUCHELI, J.; ACHCAR, J. A. **Uma introdução à análise de sobrevivência e confiabilidade**. [S.l.]: Sociedad Chilena de Estadística, 2001. Citado na página 28.
- MARSHALL, A. W.; OLKIN, I. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. **Biometrika**, Oxford University Press, v. 84, n. 3, p. 641–652, 1997. Citado nas páginas 22, 43 e 46.
- MEIRA-MACHADO, L. F.; UÑA-ÁLVAREZ, J. de; CADARSO-SUÁREZ, C.; ANDERSEN, P. Multi-state models for the analysis of time-to-event data. **Statistical methods in medical research**, SAGE Publications, 2008. Citado nas páginas 77 e 81.
- MUDHOLKAR, G. S.; SRIVASTAVA, D. K. Exponentiated weibull family for analyzing bathtub failure-rate data. **IEEE transactions on reliability**, IEEE, v. 42, n. 2, p. 299–302, 1993. Citado nas páginas 22, 43 e 45.
- MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; KOLLIA, G. D. A generalization of the weibull distribution with application to the analysis of survival data. **Journal of the American Statistical Association**, Taylor & Francis, v. 91, n. 436, p. 1575–1583, 1996. Citado nas páginas 22, 43 e 45.

- NAKAGAMI, M. The m-distribution—a general formula of intensity distribution of rapid fading. **Statistical Methods in Radio Wave Propagation**, Pergamon Press, p. 3–36, 1960. Citado nas páginas 22, 43 e 46.
- NASH, J. C. **Compact numerical methods for computers: linear algebra and function minimisation**. [S.l.]: CRC press, 1990. Citado na página 35.
- NIKULIN, M.; HAGHIGHI, F. A chi-squared test for the generalized power weibull family for the head-and-neck cancer censored data. **Journal of Mathematical Sciences**, Springer, v. 142, n. 3, p. 2204–2204, 2007. Citado nas páginas 22, 43 e 46.
- OLADAPO, O. T.; SOUZA, J. P.; BOHREN, M. A.; TUNÇALP, Ö.; VOGEL, J. P.; FAWOLE, B.; MUGERWA, K.; GÜLMEZOĞLU, A. M. Who better outcomes in labour difficulty (bold) project: innovating to improve quality of care around the time of childbirth. **Reproductive health**, BioMed Central Ltd, v. 12, n. 1, p. 48, 2015. Citado na página 19.
- OLADAPO, O. T.; SOUZA, J. P.; FAWOLE, B.; MUGERWA, K.; PERDONÁ, G.; ALVES, D.; SOUZA, H.; REIS, R.; OLIVEIRA-CIABATI, L.; MAIORANO, A.; AKINTAN, A.; ALU, F. E.; OYENEYIN, L.; ADEBAYO, A.; BYAMUGISHA, J.; NAKALEMBE, M.; IDRIS, H. A.; OKIKE, O.; ALTHABE, F.; HUNDLEY, V.; DONNAY, F.; PATTINSON, R.; SANGHVI, H. C.; JARDINE, J. E.; TUNÇALP, ; VOGEL, J. P.; STANTON, M. E.; BOHREN, M.; ZHANG, J.; LAVENDER, T.; LILJESTRAND, J.; HOOPE-BENDER, P. ten; MATHAI, M.; BAHL, R.; GÜLMEZOĞLU, A. M. Progression of the first stage of spontaneous labour: A prospective cohort study in two sub-saharan african countries. **PLOS Medicine**, Public Library of Science, v. 15, n. 1, p. 1–30, 01 2018. Citado na página 19.
- PAPAGEORGIOU, G.; MAUFF, K.; TOMER, A.; RIZOPOULOS, D. An overview of joint modeling of time-to-event and longitudinal outcomes. **Annual Review of Statistics and Its Application**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, California 94303-0139, USA, n. 0, 2019. Citado na página 32.
- PERICLEOUS, P. **Parametric joint modelling for longitudinal and survival data**. Tese (Doutorado) — University of East Anglia, 2016. Citado na página 104.
- PHAM, H.; LAI, C.-D. On recent generalizations of the weibull distribution. **IEEE transactions on reliability**, IEEE, v. 56, n. 3, p. 454–458, 2007. Citado nas páginas 22 e 43.
- PHAM, H. T. T.; PHAM, H. Penalized spline joint models for longitudinal and time-to-event data. 2018. Citado na página 22.
- PINHEIRO, J. C.; BATES, D. M. Unconstrained parametrizations for variance-covariance matrices. **Statistics and computing**, Springer, v. 6, n. 3, p. 289–296, 1996. Citado na página 49.
- PRENTICE, R. Covariate measurement errors and parameter estimation in a failure time regression model. **Biometrika**, Biometrika Trust, v. 69, n. 2, p. 331–342, 1982. Citado na página 88.
- PRENTICE, R. L. A log gamma model and its maximum likelihood estimation. **Biometrika**, Oxford University Press, v. 61, n. 3, p. 539–544, 1974. Citado nas páginas 22, 43, 44 e 68.
- PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P. **Numerical recipes 3rd edition: The art of scientific computing**. [S.l.]: Cambridge university press, 2007. Citado na página 35.

PUTTER, H.; FIOCCO, M.; GESKUS, R. B. Tutorial in biostatistics: competing risks and multi-state models. **Statistics in medicine**, Wiley Online Library, v. 26, n. 11, p. 2389–2430, 2007. Citado na página 67.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Citado nas páginas 23, 49 e 81.

RAMOS, P. L.; LOUZADA, F.; RAMOS, E. Posterior properties of the nakagami-m distribution using noninformative priors and applications in reliability. **IEEE Transactions on reliability**, IEEE, v. 67, n. 1, p. 105–117, 2018. Citado nas páginas 43 e 47.

RIZOPOULOS, D. Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive gaussian quadrature rule. **Computational Statistics & Data Analysis**, Elsevier, v. 56, n. 3, p. 491–501, 2012. Citado na página 37.

_____. **Joint Models for Longitudinal and Time-to-Event Data: With Applications in R**. [S.l.]: Taylor & Francis, 2012. (Chapman & Hall/CRC Biostatistics Series). Citado nas páginas 23, 30, 31, 34, 35, 37, 38, 69, 88 e 104.

RIZOPOULOS, D.; MOLENBERGHS, G.; LESAFFRE, E. M. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. **Biometrical Journal**, Wiley Online Library, v. 59, n. 6, p. 1261–1276, 2017. Citado na página 48.

RIZOPOULOS, D. *et al.* Jm: An r package for the joint modelling of longitudinal and time-to-event data. **Journal of Statistical Software**, Citeseer, v. 35, n. 9, p. 1–33, 2010. Citado nas páginas 22, 32 e 40.

ROCHA, R. F. **Defective models for cure rate modeling**. Tese (Doutorado) — Universidade Federal de São Carlos, 2016. Citado na página 104.

SCHWARZ, G. *et al.* Estimating the dimension of a model. **The annals of statistics**, Institute of Mathematical Statistics, v. 6, n. 2, p. 461–464, 1978. Citado na página 39.

SOTO, D. C. F. **Modelos Birnbaum-Saunders para sobrevivência com dados longitudinais**. Tese (Doutorado) — Universidade de São Paulo, 2014. Citado na página 104.

SOUZA, J. P.; OLADAPO, O. T.; BOHREN, M. A.; MUGERWA, K.; FAWOLE, B.; MOSCOVICI, L.; ALVES, D.; PERDONA, G.; OLIVEIRA-CIABATI, L.; VOGEL, J. P. *et al.* The development of a simplified, effective, labour monitoring-to-action (selma) tool for better outcomes in labour difficulty (bold): study protocol. **Reprod Health**, 2015. Citado na página 19.

STACY, E. W. A generalization of the gamma distribution. **The Annals of mathematical statistics**, JSTOR, p. 1187–1192, 1962. Citado nas páginas 22, 43 e 44.

SUDELL, M.; KOLAMUNNAGE-DONA, R.; TUDUR-SMITH, C. Joint models for longitudinal and time-to-event data: a review of reporting quality with a view to meta-analysis. **BMC medical research methodology**, BioMed Central, v. 16, n. 1, p. 168, 2016. Citado nas páginas 22 e 32.

SULAMÉRICA. **Partograma - SulAmérica Saúde Ativa**. 2016. Disponível em: <<http://www.sulamerica.com.br/saude/projetopartoadequado/Partograma.pdf>>. Acesso em: 13/08/2016. Citado na página 20.

- THERNEAU, T.; CROWSON, C. Using time dependent covariates and time dependent coefficients in the cox model. **Survival Vignettes**, 2015. Citado na página 89.
- THERNEAU, T. M.; GRAMBSCH, P. M. **Modeling survival data: extending the Cox model**. [S.l.]: Springer Science & Business Media, 2000. Citado na página 88.
- TITMAN, A. C. **Model diagnostics in multi-state models of biological systems**. Tese (Doutorado) — University of Cambridge, 2008. Citado na página 79.
- TSIATIS, A. A.; DAVIDIAN, M. Joint modeling of longitudinal and time-to-event data: an overview. **Statistica Sinica**, JSTOR, p. 809–834, 2004. Citado na página 104.
- WEIBULL, W. A statistical distribution function of wide application. **J. Appl. Mech.**, v. 18, p. 287–293, 1951. Citado na página 22.
- WILKS, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. **The Annals of Mathematical Statistics**, JSTOR, v. 9, n. 1, p. 60–62, 1938. Citado na página 39.
- WILLIAMSON, P.; KOLAMUNNAGE-DONA, R.; PHILIPSON, P.; MARSON, A. Joint modelling of longitudinal and competing risks data. **Statistics in medicine**, Wiley Online Library, v. 27, n. 30, p. 6426–6438, 2008. Citado na página 104.
- WREEDE, L. C. de; FIOCCO, M.; PUTTER, H. The mstate package for estimation and prediction in non-and semi-parametric multi-state and competing risks models. **Computer methods and programs in biomedicine**, Elsevier, v. 99, n. 3, p. 261–274, 2010. Citado nas páginas 81, 83, 84, 86 e 87.
- WREEDE, L. de; FIOCCO, M.; PUTTER, H. mstate: An r package for the analysis of competing risks and multi-state models. **Journal of Statistical Software, Articles**, v. 38, n. 7, p. 1–30, 2011. Citado nas páginas 23 e 81.
- YU, H. **Use of Generalized Gamma Distribution in Modeling Lifetime Data**. Dissertação (Mestrado) — University of Texas at Arlington, 2017. Citado na página 66.