

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

**Modelos de sobrevivência induzida por fragilidade da família
PVF: uma abordagem bayesiana**

Carlos Franklin Taco Pedraza

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em
Estatística (PIPGEs)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Carlos Franklin Taco Pedraza

Modelos de sobrevivência induzida por fragilidade da família PVF: uma abordagem bayesiana

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. *VERSÃO REVISADA*

Área de Concentração: Estatística

Orientador: Prof. Dr. Vicente Garibay Cancho

Coorientador: Prof. Dr. Miguel Angel Uribe Opazo

USP – São Carlos
Maio de 2026

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

T119m Taco Pedraza, Carlos Franklin
Modelos de sobrevivência induzida por fragilidade da família PVF: uma abordagem bayesiana / Carlos Franklin Taco Pedraza; orientador Vicente Garibay Cancho; coorientador MIGuel Angel Uribe Opazo. -- São Carlos, 2026.
149 p.

Tese (Doutorado - Programa Interinstitucional de Pós-graduação em Estatística) -- Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2026.

1. Modelo de fragilidade. 2. Distribuições PVF-ID. 3. Modelo de regressão. 4. Análise de Sobrevivência. 5. Inferência Bayesiana. I. Garibay Cancho, Vicente, orient. II. Uribe Opazo, MIGuel Angel, coorient. III. Título.

Carlos Franklin Taco Pedraza

**Frailty-Induced Survival Models from the PVF Family: A
Bayesian Approach**

Thesis submitted to the Institute of Mathematics and Computer Science – 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. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Vicente Garibay Cancho

Co-advisor: Prof. Dr. Miguel Angel Uribe Opazo

USP – São Carlos

May 2026

Folha de Aprovação

Defesa de Tese de Doutorado do candidato Carlos Franklin Taco Pedraza, realizada em 27/03/2026.

Comissão Julgadora:

Prof. Dr. Vicente Garibay Cancho (ICMC/USP)

Prof. Dr. Bernardo Moises Lagos Alvarez (UDEEC-Chile)

Profa. Dra. Daiane de Souza Santos (ICMC/USP)

Profa. Dra. Gabriela Maria Rodrigues (ESALQ/USP)

Prof. Dr. Marcos Jardel Henriques (UFSCar)

A todos que contribuíram com seu conhecimento e apoio, tornando esta conquista possível.

AGRADECIMENTOS

Gostaria de expressar minha profunda gratidão ao meu orientador, Prof. Vicente G. Cancho, pela orientação exemplar, paciência e apoio incondicional. Sua expertise e dedicação foram fundamentais para o desenvolvimento e conclusão desta pesquisa. Agradeço também pela confiança demonstrada em mim ao longo de todo o processo, especialmente por abrir as portas de sua casa e de sua sala, proporcionando um ambiente acolhedor e inspirador para o diálogo, a reflexão e o crescimento acadêmico. Estendo meus sinceros agradecimentos aos membros da banca de defesa, pela disponibilidade, atenção e valiosas contribuições oferecidas a este trabalho. Suas observações, questionamentos e sugestões enriqueceram significativamente esta pesquisa e contribuíram para meu amadurecimento acadêmico e científico.

Agradeço também à minha família, meu verdadeiro porto seguro, cujo amor, apoio e incentivo inabaláveis foram essenciais em cada etapa desta jornada. Mesmo diante da dolorosa perda de meu pai, das cirurgias enfrentadas por minha irmã e por minha sobrinha, e das constantes preocupações de minha mãe, permanecemos unidos, encontrando força uns nos outros para seguir em frente. E, em meio aos desafios, a alegria contagiante das minhas sobrinhas — meus amorzinhos — trouxe leveza aos dias difíceis, renovando minha esperança e meu ânimo.

Por fim, volto ao Bloco 4 do Instituto de Ciências Matemáticas e de Computação, onde tudo começou: uma madrugada fria, uma mala nas mãos e as incertezas dividindo espaço com o silêncio do campus. Hoje, retorno ao mesmo lugar para escrever as últimas linhas desta tese, enquanto toca *No Surrender*, de Bruce Springsteen — a mesma música daquela primeira madrugada. Talvez porque Springsteen cante, como poucos, sobre a vida das pessoas comuns: as lutas silenciosas, as perdas e a coragem cotidiana de seguir em frente apesar de tudo. Entre o início e o fim desta jornada, ficam guardados na memória os rostos, os gestos e os encontros de tantas pessoas que passaram pelo meu caminho e que hoje carrego comigo, com carinho, no coração.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“Todos os modelos estão errados, mas alguns são úteis.”
(George Box)

RESUMO

PEDRAZA, T. C. F. **Modelos de sobrevivência induzida por fragilidade da família PVF: uma abordagem bayesiana.** 2026. 149 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, 2026.

Em análise de sobrevivência, é comum incorporar covariáveis, como idade, sexo, gravidade da doença entre outras. No entanto, existem diversos fatores não observados que também podem influenciar a sobrevivência, incluindo o estado geral de saúde, o estilo de vida, a ocupação ou a predisposição genética. Tais fatores não observados podem ser incorporados ao modelo por meio de uma variável de fragilidade, frequentemente denominada fragilidade, a qual representa a heterogeneidade não observada entre os indivíduos do estudo. O objetivo deste trabalho é desenvolver modelos de fragilidade para capturar essa heterogeneidade não observada em dados de sobrevivência. Apresenta-se uma nova classe de modelos de sobrevivência que combina a heterogeneidade não observada da população e a presença de indivíduos curados por meio da incorporação de termos de fragilidade provenientes da família PVF (Família de Funções de Variância de Potência), denominados modelos defeituosos Induzidos por PVF (PVF-ID). Essa estrutura amplia a teoria clássica de modelos de fragilidade e de tempos de vida defeituosos, permitindo aos pesquisadores modelar de forma flexível a presença de sobreviventes de longo prazo. A inferência é realizada por meio de um procedimento totalmente Bayesiano implementado via *Hamiltonian Monte Carlo*, enquanto propriedades frequentistas relevantes são avaliadas por meio de estudos de simulação. Quando aplicado a dados de sobrevivência relacionados ao HIV/AIDS, o modelo identifica de forma consistente as covariáveis. Além disso, propõe-se um modelo bivariado de sobrevivência com fração de cura e fragilidade, baseado em misturas de Poisson e na família PVF. O modelo é estendido para regressão, permitindo avaliar o efeito de covariáveis nos tempos de sobrevivência. A inferência é realizada sob abordagem bayesiana via *Hamiltonian Monte Carlo* implementado em R-Stan, e estudos de simulação são conduzidos para avaliar o desempenho dos estimadores. A flexibilidade da estrutura proposta permite capturar a dependência entre tempos de falha e heterogeneidade não observada. A utilidade prática do modelo é ilustrada por meio de conjuntos de dados reais.

Palavras-chave: Modelo de fragilidade, Distribuições PVF-ID, Modelo de regressão, Análise de Sobrevivência, Inferência Bayesiana.

ABSTRACT

PEDRAZA, T. C. F. **Frailty-Induced Survival Models from the PVF Family: A Bayesian Approach**. 2026. 149 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, 2026.

In survival analysis, it is common to incorporate covariates such as age, sex, disease severity, among others. However, there are several unobserved factors that may also influence survival, including overall health status, lifestyle, occupation, or genetic predisposition. Such factors are often referred to as frailty and represent unobserved heterogeneity among individuals in the study. The objective of this work is to develop frailty models capable of capturing this unobserved heterogeneity in survival data. We introduce a new class of survival models that combines population unobserved heterogeneity and the presence of cured individuals through the incorporation of frailty terms from the PVF family (Power Variance Functions), referred to as PVF-Induced Defective (PVF-ID) models. This framework extends the classical theory of frailty models and defective lifetime models, allowing researchers to flexibly model the presence of long-term survivors. Inference is performed through a fully Bayesian procedure implemented via Hamiltonian Monte Carlo, while relevant frequentist properties are assessed through simulation studies. When applied to survival data related to HIV/AIDS, the model consistently identifies the effects of the covariates. Additionally, a bivariate survival model with cure fraction and frailty is proposed, based on Poisson mixtures and the PVF family. The model is extended to a regression framework, allowing the assessment of covariate effects on survival times. Inference is conducted under a Bayesian approach using Hamiltonian Monte Carlo implemented in RStan, and simulation studies are carried out to evaluate the performance of the estimators. The flexibility of the proposed structure allows the model to capture both dependence between failure times and unobserved heterogeneity. The practical utility of the model is illustrated using a real dataset.

Keywords: Fragility Model, PVF-ID distribution, Regression model, Survival Analysis, Bayesian Inference.

LISTA DE ILUSTRAÇÕES

Figura 1	– Gráfico da f.d.p do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$	42
Figura 2	– Gráfico da função de sobrevivência do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$	42
Figura 3	– Gráfico da função de distribuição acumulada do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$	43
Figura 4	– Gráfico da função de risco acumulado do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$	43
Figura 5	– Estimativa de Kaplan-Meier da função de sobrevivência para os dados de AIDS/HIV (painel esquerdo) e função de risco acumulado estratificada para pacientes com ou sem POI (painel direito).	70
Figura 6	– Gráfico QQ-plot dos resíduos de quantis normalizados a posteriores com linha de identidade (a esquerda) e histograma do modelo PVF-ID Weibull sob as hipóteses definidas (a direita)	73
Figura 7	– Gráficos de rastreamento para parâmetros do modelo PVF-ID Weibull referente aos dados de HIV/AIDS.	73
Figura 8	– Gráficos de rastreamento para parâmetros do modelo PVF-ID Weibull referente aos dados de HIV/AIDS.	74
Figura 9	– Estimativas da função de risco acumulado, obtidas pelo método de Kaplan-Meier, dos tempos de churn dos Produtos 1 e 2.	97
Figura 10	– Gráfico de índices das medidas de divergência ϕ relacionadas aos dados de Churn de clientes brasileiros.	99
Figura 11	– A função de risco acumulado marginal a posteriori de T_1 e T_2 , estratificada pela renda do cliente para cada nível de gênero: (painel direito) feminino e (painel esquerdo) masculino.	102
Figura 12	– Boxplot da probabilidade a posteriori de clientes fiéis aos produtos 1 e 2: (a) ao produto 1, (b) ao produto 1 e (c) ao produto 2, para os clientes A, B, C, D, E e F.	103

LISTA DE TABELAS

Tabela 1 – Distribuições a Priori no Estudo de Simulação	66
Tabela 2 – Estimativas da média, DP, viés e REQM dos parâmetros dos modelos PVF-ID com a presença de covariáveis para $\gamma = 0,30$ ou $\gamma = 0,60$, $\beta_{10} = 0$, $\beta_{11} = -1$, $\beta_{12} = -1$, $\beta_{20} = 1$, $\beta_{21} = -1$ e $\beta_{22} = 1$	67
Tabela 3 – Medidas descritivas do tempo (em anos) até a morte do paciente por AIDS/HIV	69
Tabela 4 – Parâmetros do modelo e suas interpretações	71
Tabela 5 – Distribuições priori dos parâmetros do modelo	72
Tabela 6 – Estatísticas do LPML para os dados de HIV/AIDS	72
Tabela 7 – Estatísticas LPML, médias a posteriores, medianas, desvios padrões e 95% de confiança dos intervalos HPD para os parâmetros do modelo PVF-ID Weibull relacionado aos dados de AIDS/HIV.	75
Tabela 8 – Distribuições a priori especificadas no modelo	94
Tabela 9 – Resultados das estimativas para diferentes tamanhos amostrais, para 200 réplicas e censura aproximada de 42%.	95
Tabela 10 – Médias posteriores, medianas e intervalos HPD de 95% para os parâmetros do modelo de regressão bivariada para o conjunto de dados do banco brasileiro, considerando os modelos de fragilidade Gamma, IG, Positiva Estavel e PVF.	98
Tabela 11 – Probabilidades posteriores dos modelos a partir das verossimilhanças marginais.	98
Tabela 12 – Estimativa Bayesiana da probabilidade de clientes fiéis aos produtos 1 e 2 após 3 anos do período de observação.	101
Tabela 13 – Estimativa Bayesiana da probabilidade de clientes fiéis aos produtos 1 e 2.	101

SUMÁRIO

1	INTRODUÇÃO	23
1.1	Revisão de Literatura	23
1.2	Objetivos	27
1.3	Estrutura do Trabalho	28
2	NOÇÕES PRELIMINARES	31
2.1	Heterogeneidade no modelo de Cox	31
2.2	Modelos de fragilidade	33
2.2.1	<i>Modelo de fragilidade univariado</i>	33
2.3	Principais modelos de fragilidade	36
2.3.1	<i>Distribuição de fragilidade Gama</i>	37
2.3.2	<i>Distribuição de fragilidade Inversa Gaussiana (IG)</i>	38
2.3.3	<i>Distribuição de fragilidade Positiva Estável (PE)</i>	38
2.3.4	<i>Distribuição de fragilidade para a família PVF</i>	39
2.4	Modelo Exponencial por Partes (MEP)	40
2.5	Inferência Bayesiana	44
2.5.1	<i>Método de Monte Carlo via Cadeias de Markov (MCMC)</i>	45
2.5.2	<i>Monte Carlo Hamiltoniano (HMC)</i>	46
2.5.3	<i>Critério de comparação de modelos bayesianos</i>	48
2.5.3.1	<i>Bridge sampling</i>	49
2.5.3.2	<i>Conditional predictive ordinate (CPO)</i>	51
2.5.4	<i>Influência bayesiana global</i>	52
2.6	Conclusão	52
3	NOVO MODELO INDUZIDO POR FRAGILIDADE	53
3.1	Introdução	53
3.2	Modelo de defeituoso Induzido por fragilidade PVF	55
3.3	Modelo de Mistura Induzido pela Distribuição PVF	61
3.4	Inferência bayesiana	63
3.4.1	<i>Estudo de simulação</i>	65
3.4.1.1	<i>Passo 1: Cálculo da Probabilidade de Cura</i>	65
3.4.1.2	<i>Passo 2: Determinação do estado de Cura</i>	65
3.4.1.3	<i>Passo 3: Geração do Tempo para Suscetíveis</i>	65

3.4.1.4	<i>Passo 4: Geração da Censura</i>	65
3.5	Aplicação	68
3.5.1	Análise dos dados	69
3.6	Conclusão	76
4	MODELO DE SOBREVIVÊNCIA BIVARIADO	77
4.1	Modelo Bivariado de Longa Duração	77
4.2	Modelo de sobrevivência bivariado de longa duração	79
4.2.1	<i>Funções de Risco Conjunta e Marginais</i>	83
4.2.2	<i>Medida Local de Associação</i>	83
4.3	Inferência Bayesiana	85
4.3.1	<i>Função de verossimilhança</i>	85
4.3.2	<i>Prioris e posteriori</i>	88
4.4	Estudo de simulação	89
4.4.1	Detalhamento das Etapas do Algoritmo	92
4.4.1.1	<i>Etapa 1: Geração da Covariável</i>	92
4.4.1.2	<i>Etapa 2: Cálculo das Taxas de Ocorrência</i>	92
4.4.1.3	<i>Etapa 3: Geração da Fragilidade PVF</i>	93
4.4.1.4	<i>Etapa 4: Geração do Número de Causas Latentes</i>	93
4.4.1.5	<i>Etapa 5: Geração dos Tempos de Promoção</i>	93
4.4.1.6	<i>Etapa 6: Aplicação da Censura</i>	94
4.4.1.7	<i>Etapa 7: Determinação dos Tempos Observados e Indicadores de Falha</i>	94
4.5	Aplicação: Dados de Churn de Clientes Brasileiros	96
4.6	Conclusões	103
5	CONSIDERAÇÕES FINAIS	105
5.1	Conclusões	105
5.2	Sugestões para pesquisas futuras	106
	REFERÊNCIAS	107
	APÊNDICE A DEMONSTRAÇÕES DO CAPÍTULO 3	115
A.0.1	Afirmação (3.4)	115
A.0.2	Verificação $\lim_{t \rightarrow \infty} S(t)$ quando $\gamma \in (0, 1)$	116
A.0.3	Verificação p_0	117
A.0.4	Análise da Função de Risco Marginal e seu Comportamento Limite	118
A.0.5	Afirmação (3.6)	120
A.0.6	Verificação (3.7): $S_u(t \mathbf{w})$	122
A.0.7	Afirmação (3.8)	124
A.0.8	Verificação $h_p(t \mathbf{x}, \mathbf{w})$	125

A.0.9	Verificação (3.9)	126
APÊNDICE B	DEMONSTRAÇÕES DO CAPÍTULO 4	129
B.0.1	Verificação da Esperança, Variância e Covariância de N_k	129
B.0.2	Afirmção (4.5)	130
B.0.3	Verificação de p_{00}	130
B.0.4	Afirmção (4.6)	131
B.0.5	Verificação de p_{0k}	132
B.0.6	Afirmção (4.7)	132
B.0.7	Afirmção (4.8)	133
B.0.8	Verificação $\lim_{\sigma \rightarrow \infty} S(t_1, t_2)$	133
B.0.9	Afirmção (4.9)	133
B.0.10	Afirmção (4.11)	134
B.0.11	Verificação (4.14)	134
B.0.12	Verificação (4.15)	136
B.0.13	Verificação (4.17)	137
B.0.14	Afirmção (4.18)	138
B.0.15	Prova do Teorema 1: Obtenção da Verossimilhança Observada	140

INTRODUÇÃO

Os conjuntos de dados de sobrevivência são caracterizados por variáveis que representam os tempos até a ocorrência dos eventos de interesse e, em geral, por variáveis indicadoras de censura. Além disso, os dados analisados ainda podem incorporar outras variáveis observadas, denominadas covariáveis ou variáveis explicativas, que estão relacionadas com seus respectivos tempos. Contudo, apenas algumas destas covariáveis são conhecidas e medidas, como por exemplo, idade, sexo, entre outras. Os fatores desconhecidos ou não incluídos na análise, mas que podem influenciar na sobrevivência dos indivíduos, são denominados na literatura como fragilidade. O principal objetivo deste conceito é controlar a heterogeneidade não observada dos indivíduos do estudo. Desse modo, modelos que buscam captar a existência e quantificar essa fragilidade são conhecidos como modelos de fragilidade.

Vaupel, Manton e Stallard (1979) consideram o modelo de fragilidade como um modelo caracterizado pela presença de um efeito aleatório durante o período de estudo. Em outras palavras, há existência de uma variável aleatória não observada que representa informações não disponíveis, não observadas ou não consideradas durante o período de análise de estudo. Este efeito aleatório é representado pela variável de fragilidade, sendo introduzida na modelagem através da função de risco. Assim, a heterogeneidade incorporada ao modelo decorre de covariáveis individuais não observadas ou ainda desconhecidas como fatores de risco.

1.1 Revisão de Literatura

O estudo de modelos de fragilidade vem ganhando espaço no âmbito acadêmico tendo destaque em aplicações que enfatizam tanto variáveis discretas de fragilidade quanto contínuas. Em particular, modelos que consideram uma fragilidade discreta possuem uma maior flexibilidade, além de evidenciar a existência de uma proporção de indivíduos dos quais não se espera a ocorrência do evento de interesse, bem como a possível observação de indivíduos com fragilidade zero. Estas características podem ser observadas nos trabalhos de Souza *et al.* (2017),

que propõem um modelo com fração de cura induzido pela fragilidade Hiper-Poisson para a modelagem da sobrevivência induzida por uma variável discreta de fragilidade e de [Molina \(2020\)](#), que destaca diferentes modelos de fragilidade usando distribuições pertencentes à família de Série de Potência Zero-Modificada, observando indivíduos com fragilidade zero. Contudo, diferentes modelos empregam distribuições de fragilidades como sendo contínuas e não-negativas, como é o caso das distribuições Gama [Vaupel, Manton e Stallard \(1979\)](#), Inversa Gaussiana (IG) [Hougaard \(1984\)](#), Positiva Estável (PE) [Hougaard \(1986\)](#) e Gama Generalizada [Chen, Zhang e Zhang \(2013\)](#). Na literatura muitos são os trabalhos que exploram desta continuidade em seus estudos. Por exemplo, [Giussani e Bonetti \(2019\)](#) usam técnicas multivariadas de sobrevivência para a análise de dados de tempo de falha com censura à direita, em que é investigado uma nova família de modelos paramétricos de fragilidade bivariada e [Bunyatisai, Prasitwattanaseree e Ingrisawang \(2017\)](#) investigam o desempenho do modelo de Cox com e sem o uso da fragilidade Gama e estabelecem um modelo paramétrico Weibull para esta mesma variável de fragilidade.

Além desses trabalhos, é comum encontrar estudos que utilizam da fragilidade para analisar conjuntos de dados, especialmente voltados para a área da saúde. Por exemplo, [Isidro *et al.* \(2020\)](#) aplicam modelos de fragilidade Gama para avaliar o risco e os fatores que contribuem para a morte de pacientes com leucemia, enquanto [Gurmu \(2018\)](#) usam o método de Kaplan-Meier, o modelo de Cox e o modelo paramétrico de fragilidade compartilhada para estudar a sobrevivência de pacientes com câncer cervical. Enquanto, [Banbeta *et al.* \(2015\)](#) utilizam o modelo de fragilidade para analisar o tempo de cura de vítimas de desnutrição aguda grave, considerando as distribuições Exponencial, Weibull e Log-Logística para as funções de risco e as distribuições Gama e Inversa Gaussiana para as distribuições de fragilidade.

O efeito da fragilidade em variáveis contínuas é observado em abordagens semiparamétricas e não-paramétricas, nas quais esses modelos capturam a variabilidade de forma mais eficaz para análise de dados, em comparação com modelos tradicionais de sobrevivência. Estes relatos são presenciados em [Brito *et al.* \(2020\)](#) e [Zhou *et al.* \(2015\)](#). Em [Brito *et al.* \(2020\)](#), é realizada uma comparação entre os modelos de Cox e semiparamétricos, além de uma simulação do modelo de fragilidade Gama com diferentes níveis de censura e heterogeneidade. Como mostrado [Zhou *et al.* \(2015\)](#) propõem um modelo bayesiano não-paramétrico dependente, ajustando a distribuição da fragilidade a covariáveis contínuas e categóricas de maneira flexível.

A abordagem bayesiana tem sido amplamente usada em estudos de modelos de fragilidade, destacada no trabalho de [Zhou *et al.* \(2015\)](#) pela sua eficiência. A inferência bayesiana também é aplicada tanto a variáveis de fragilidade discretas quanto contínuas. Exemplos de pesquisas com variável discreta de fragilidade incluem [Fortes \(2020\)](#), [Zhou, Hanson e Zhang \(2017\)](#), [Cancho *et al.* \(2018\)](#), [Cancho *et al.* \(2021\)](#), [Macera \(2015\)](#), entre outros.

Por exemplo, [Fortes \(2020\)](#) propõe um modelo de sobrevivência com fragilidade Weibull, baseado na distribuição Birnbaum-Saunders, enquanto [Zhou, Hanson e Zhang \(2017\)](#) utilizam desta inferência para desenvolver uma abordagem generalizada para o modelo de Tempo de Falha

Acelerado. [Wheeler et al. \(2021\)](#), [Gasparini et al. \(2019\)](#) e [Amorim \(2014\)](#) usam a inferência bayesiana considerando a variável de fragilidade contínua. [Wheeler et al. \(2021\)](#) usam fragilidade para incluir efeitos aleatórios no risco de exposição a alérgenos alimentares, enquanto [Amorim \(2014\)](#) compara diferentes abordagens de verossimilhança em modelos de fragilidade, baseados no algoritmo EM, cadeias de Markov de Monte Carlo e em processos de estimação usando verossimilhança parcial, verossimilhança penalizada, dentre outros, e [Gasparini et al. \(2019\)](#) avalia o impacto de uma especificação incorreta da função de risco base em cenários clínicos via simulação de Monte Carlo.

O impacto enfatizado por [Gasparini et al. \(2019\)](#) é muito importante nos estudos sobre fragilidade, uma vez que, este recurso é empregado com a finalidade de controlar a heterogeneidade não observada das unidades do estudo, melhorando significativamente os modelos. Um recurso para que haja este controle é o uso da Família de Funções de Variância de Potência (PVF), sugerida por [Tweedie et al. \(1984\)](#). A família PVF é uma classe generalizada de modelos de fragilidade que inclui distribuições conhecidas, como Gama, Inversa Gausiana, Log-Normal e Exponencial por partes. Segundo [Wienke \(2010\)](#), essa classe se destaca pelo uso e benefícios das distribuições que a compõem, além da importância durante o processo de estimação. Um exemplo disso é o trabalho de [Monaco, Gorfine e Hsu \(2018\)](#), que apresentam o pacote R denominado `fragilitySurv` para simular e ajustar modelos de fragilidade semiparamétricos, utilizando estimadores consistentes para distribuições PVF. Além disso, outra característica destacada por [Wienke \(2010\)](#) é a possibilidade da representação de algumas distribuições mediante restrições paramétricas nas distribuições PVF. Isso é exemplificado nos estudos de [Rodrigues, Calsavara e Tomazella \(2018\)](#), [Calsavara et al. \(2017\)](#) e [Rodrigues et al. \(2021\)](#), que propõem um modelo de taxa de cura com fragilidade usando distribuições PVF. O trabalho de [Cancho et al. \(2021\)](#) também destaca a aplicabilidade da distribuição PVF como a variável de fragilidade, permitindo uma estimativa mais precisa da taxa de cura.

Os trabalhos mencionados anteriormente compartilham uma característica comum, chamada taxa de cura, que retratam estudos que possuem uma fragilidade zero, indicando um subgrupo de indivíduos não suscetíveis, em que o evento de interesse não ocorre. Esses dados possuem a estrutura de longa duração sendo determinada por ter algumas unidades amostrais com fator de risco zero, conforme estudos de [Berkson e Gage \(1952\)](#), [Yakovlev, Tsodikov e Asselain \(1996\)](#), [Tsodikov, Ibrahim e Yakovlev \(2003\)](#) e [Rodrigues et al. \(2009\)](#). Um exemplo de trabalho que destaca a existência de uma fragilidade zero é o estudo de [Bedia \(2022\)](#), que propõe um modelo de sobrevivência bivariada de longa duração cujo o número de causas de risco para diferentes tipos de eventos é modelado mediante a um modelo de fragilidade da família PVF. Contudo, essa taxa de cura também é analisada por meio de modelos denominados modelos defeituosos. Alguns autores ressaltam a utilização destes modelos com o efeito de fragilidade. Por exemplo, [Tesema et al. \(2022\)](#) e [Tomazella, Milani e Dias \(2018\)](#) estudam a mortalidade infantil na África Oriental e investigam a heterogeneidade não observada, respectivamente, usando o modelo de fragilidade compartilhada Gama-Gompertz. Entanto [Lima et al. \(2021\)](#)

aplicam o modelo Gama-Gompertz para corrigir a heterogeneidade não observada em cabras anglo-nubianas e [Scudilio et al. \(2019\)](#) propõem um modelo defeituoso com termo de fragilidade para modelar a taxa de cura, usando as distribuições Gama-Gompertz Defeituosa.

Outro recurso ressaltado na literatura é o Modelo Exponencial por Partes (MEP) devido a sua grande flexibilidade em diversos estudos estatísticos e probabilísticos, possuindo a capacidade de acomodar funções taxa de falha com diversas formas [Ibrahim, Chen e Sinha \(2001\)](#). Outra vantagem é a possibilidade de se trabalhar com este modelo tanto na versão paramétrica quanto na versão não-paramétrica, uma vez que, é caracterizado pela aproximação da função taxa de falha por segmentos de retas cujos comprimentos são determinados através de partições do eixo do tempo, resultando em intervalos em que a função taxa de falha é considerada constante. Desta forma, a versão não-paramétrica do MEP é obtida ao tomar uma partição do eixo do tempo com tantos intervalos quanto for o número de falhas, enquanto a versão paramétrica do MEP é determinada ao considerar um número de intervalos inferior ao número de falhas, permitindo que haja mais de uma falha por intervalo [Demarqui \(2006\)](#). Os trabalhos de [Sibim e Cancho \(2011\)](#), [Mello et al. \(2016\)](#), [Demarqui \(2006\)](#) e [Demarqui \(2010\)](#) destacam a importância do modelo MEP na modelagem de sobrevivência segundo a abordagem bayesiana. [Sibim e Cancho \(2011\)](#) desenvolveram procedimentos para modelos com e sem taxa de cura em MEP. [Mello et al. \(2016\)](#) propõem uma abordagem obtendo uma distribuição suavizada para os parâmetros, usando o modelo MEP para representar a função de risco. [Demarqui \(2006\)](#) apresenta um estudo do MEP focando na partição intervalar do eixo dos tempos. Por sua vez, [Demarqui \(2010\)](#) reúne artigos que exploram o modelo MEP em diversos contextos, incluindo um com taxa de cura. Esses estudos são aplicáveis a modelagem de dados de sobrevivência em diferentes áreas do conhecimento.

Apesar dos avanços metodológicos observados na literatura, persistem desafios inferenciais de natureza prática, sobretudo em virtude da elevada complexidade estrutural de determinados modelos, como frequentemente ocorre na abordagem bayesiana. Nesse contexto, a estimação de parâmetros a partir da distribuição a posteriori pode tornar-se analiticamente intratável. Logo, o emprego de métodos computacionais eficientes, tais como os algoritmos de Monte Carlo via Cadeias de Markov (MCMC), torna-se fundamental para viabilizar a inferência. O objetivo central dos métodos MCMC consiste em gerar amostras dependentes da distribuição a posteriori, de modo a permitir a obtenção de estimativas amostrais para os parâmetros de interesse, em analogia ao método de Monte Carlo clássico. Enquanto, o Monte Carlo tradicional, que se baseia em amostras independentes, os métodos MCMC utilizam cadeias de Markov ergódicas cuja distribuição estacionária coincide com a distribuição a posteriori alvo. A simulação dessas cadeias pode ser implementada por meio de diferentes algoritmos, entre os quais se destacam o Hamiltonian Monte Carlo (HMC), o algoritmo de Metropolis–Hastings (MH) e abordagens fundamentadas em técnicas de redução de variância. Em particular, segundo [Spiegelhalter et al. \(1996\)](#), o HMC combina os métodos Gibbs e MH e evita o comportamento de passeio aleatório, reduzindo a autodependência das cadeias. [Torres et al. \(2018\)](#) também destaca que diferente dos métodos

MCMC de caminhada aleatória, o HMC considera o gradiente da distribuição de probabilidade a posteriori, o que é uma vantagem. Estudos como os de [Xavier \(2019\)](#), [Chatzilena *et al.* \(2019\)](#), [Paixão \(2021\)](#) e [Hartmann \(2015\)](#) enfatizam a importância do HMC, aplicando-o em diversos estudos como na estimação de parâmetros em modelos GARCH univariado e multivariado, na análise de dados de doenças infecciosas dinâmicas, na modelagem de dados GJR-GARCH e na abordagem bayesiana não-paramétrica para dados de comportamento extremo, no qual, o método HMC é aplicado como uma solução alternativa para lidar com a intratabilidade analítica de distribuições a posteriori com altas dimensões.

Outro recurso promissor em estudos de simulação é a utilização da linguagem de programação Stan. Esta linguagem pode ser executada via linha de comando, sendo compatível com programas como R, Python, Matlab, Stata e Julia, conforme apontado por [Jiang e Carter \(2019\)](#), [Luo e Jiao \(2018\)](#), [Xavier \(2019\)](#) e [Gelman, Lee e Guo \(2015\)](#). Autores como [Jiang e Carter \(2019\)](#) e [Luo e Jiao \(2018\)](#) destacam as vantagens do Stan, enfatizando que esta é uma linguagem fácil, de sintaxe direta, flexibilidade na modelagem e rapidez nas estimativas, devido ao uso de funções otimizadas em C++, além de executar o método HMC oferecendo múltiplas opções de estimativa [Jiang e Carter \(2019\)](#). Ainda, [Luo e Jiao \(2018\)](#) também mencionam a eficiência do Stan em comparação com programas bayesianos tradicionais, como o BUGS, pois o Stan exige menos iterações para boa mixagem (1000 contra 100.000 no BUGS), além da possibilidade de permitir o uso de diversos tipos de distribuições. Como consequência, os estudos de [Luo e Jiao \(2018\)](#), [Chatzilena *et al.* \(2019\)](#), [Ng'ombe e Lambert \(2021\)](#) e [Jiang e Carter \(2019\)](#) evidenciam a eficácia do Stan tem sido amplamente documentada na literatura, com aplicações em modelos da Teoria de Resposta ao Item (TRI), modelos epidemiológicos, processos estocásticos e modelos log-lineares. Esses estudos fornecem evidências empíricas que corroboram a adequação do algoritmo Hamiltonian Monte Carlo (HMC), implementado via Stan e sua interface RStan, para fins de inferência bayesiana. De modo geral, os resultados indicam que tais ferramentas apresentam desempenho computacional robusto, eficiência amostral elevada e ampla aplicabilidade nas investigações estatísticas contemporâneas.

1.2 Objetivos

O presente trabalho tem como objetivo desenvolver e investigar modelos de fragilidade destinados à modelagem da heterogeneidade não observada em dados de sobrevivência. Especificamente, pretende-se:

- Realizar uma revisão teórica dos modelos de fragilidade, enfatizando sua formulação no contexto do modelo de riscos proporcionais de Cox, com especial atenção aos modelos pertencentes à família PVF. Adicionalmente, serão examinados modelos do tipo MEP, modelos defeituosos e modelos de longa duração, contemplando seus fundamentos teóricos e propriedades inferenciais.

- Apresentar modelos univariados para a modelagem da heterogeneidade não observada, considerando uma composição entre uma distribuição de fragilidade da família PVF induzidos por fragilidade. Além disso, será apresentado um modelo de sobrevivência bivariado com fragilidade, no qual a variabilidade não observada é capturada por uma distribuição de Poisson, enquanto a componente de fragilidade é modelada por uma distribuição da família PVF.
- Estudar as propriedades fundamentais dos modelos propostos (univariado e bivariado), com ênfase nas funções de análise de sobrevivência, tais como funções de risco, sobrevivência e de densidade, bem como suas implicações inferenciais.
- Usar procedimentos de inferência bayesiana utilizando o algoritmo HMC implementado no RStan, com o objetivo de avaliar o desempenho e as propriedades dos estimadores bayesianos por meio de estudos de simulação.
- Avaliar a viabilidade de aplicação dos modelos propostos em bases de dados reais, para verificar o desempenho e a adequação dos modelos propostos.

1.3 Estrutura do Trabalho

Os capítulos deste trabalho são organizados da seguinte forma:

- No Capítulo 2 é apresentado uma breve revisão teórica sobre conceitos de inferência bayesiana. Na Seção 2.1 tem um breve estudo sobre o modelo de Cox, na Seção 2.2 é abordado o conceito dos modelos de fragilidade univariado destacando a conexão com a transformada de Laplace. Enquanto na Seção 2.3 é apresentado os principais modelos de fragilidade Gama, Inversa Gaussiana, Positiva Estável e a família PVF. Na seção 2.4 é estudada o modelo Exponencial por Partes. E finalmente na Seção 2.5 apresentam-se uma breve revisão teórica sobre conceitos as técnicas de amostragem probabilística como são o método de monte Carlo via cadeias de Markov, o algoritmo de Metropolis-Hasting, e o Monte Carlo Hamiltoniano, e finalmente apresenta-se um breve estudo sobre o Critério de comparação de modelos bayesianos.
- No Capítulo 3 é apresentado um modelo estatístico de fragilidade para modelar a heterogeneidade não observada em dados de sobrevivência. Na Seção 3.2 apresenta-se o Modelo de defeituoso Induzido por fragilidade PVF, na Seção 3.3 o Modelo de Mistura Induzido pela Distribuição PVF, na Seção 3.4 apresenta-se a inferência bayesiana para o modelo proposto, e finalmente são apresentados a simulação e uma aplicação de dados reais.
- No Capítulo 4, apresentam-se um Modelo Bivariado de Longa Duração induzida por fragilidade discreta. Na Seção 4.2, são determinadas as funções de risco conjunta e marginais, bem como a medida local de associação. Na Seção 4.3, são apresentadas a

função de verossimilhança, as distribuições a priori e a posteriori. Em seguida, um estudo de simulação na Seção 4.4 e finalmente na Seção 4.5, uma aplicação para Dados de um sistema financeiro Brasileiro.

- No Capítulo 5 é exibido um resumo dos principais resultados da tese e propostas para futuras pesquisas relacionadas aos modelos apresentados. Finalmente, na apêndice A e B são apresentadas os cálculos correspondentes aos resultados abordados nos Capítulos 3 e 4 respectivamente.

NOÇÕES PRELIMINARES

No presente capítulo, apresenta-se a fundamentação teórica sobre os principais temas para o desenvolvimento deste trabalho. Inicialmente, são apresentados os modelos de fragilidade, com ênfase na distribuição de fragilidade pertencente à família PVF, que é relevante para a modelagem de riscos em análise de sobrevivência. Essa revisão considera o contexto univariado, explorando as propriedades e aplicações desses modelos em cenários distintos. Para fornecer a base teórica desses modelos são considerados os estudos clássicos como [Aalen \(1992\)](#), [Hougaard \(2000\)](#), [Wienke \(2010\)](#), [Balan e Putter \(2020\)](#), [Duchateau e Janssen \(2008\)](#) e [Bedia \(2022\)](#), cujas contribuições são fundamentais para o entendimento sobre fragilidade na modelagem de riscos proporcionais. Por meio de tais referências, o capítulo demonstra como as distribuições de fragilidade podem ser integradas de forma eficaz em modelos estatísticos para descrever a heterogeneidade não observada entre indivíduos ou unidades experimentais.

Finalmente, o capítulo apresenta os conceitos fundamentais da inferência bayesiana, em especial no que se refere às técnicas computacionais de estimação, tais como algoritmos de Monte Carlo via Cadeias de Markov (MCMC), além dos critérios para comparação e seleção de modelos estatísticos. Essa abordagem é tratada considerando os pressupostos e metodologias sugeridos por [Xavier \(2019\)](#), [Paixão \(2021\)](#), [Hartmann \(2015\)](#), [Pardo \(2018\)](#), [Sibim e Cancho \(2011\)](#), [Ibrahim, Chen e Sinha \(2001\)](#), [Meng e Wong \(1996\)](#), [Meng e Schilling \(2002\)](#), [Gronau *et al.* \(2019\)](#), [Gronau *et al.* \(2017\)](#) e [Gronau, Singmann e Wagenmakers \(2020\)](#). Essas referências fornecem o suporte necessário às opções metodológicas, garantindo assim o rigor e a robustez nas análises bayesianas apresentadas nos capítulos posteriores.

2.1 Heterogeneidade no modelo de Cox

As pesquisas apresentadas por [Balan e Putter \(2020\)](#), [Ahmed, Vos e Holbert \(2007\)](#), [Andersen *et al.* \(2021\)](#), [Hosmer \(2002\)](#) enfatizam a função de risco como o elemento central na Análise de Sobrevivência, dado que esta função quantifica a taxa instantânea de ocorrência do

evento de interesse para um determinado tempo, considerando a variabilidade entre indivíduos nas probabilidades de sobrevivência e nas funções de risco, é essencial que os modelos estatísticos incorporem essas heterogeneidades para garantir estimativas consistentes. Na literatura, o modelo de riscos proporcionais de Cox é amplamente reconhecido para análises desse tipo. No referido modelo, a suposição de riscos proporcionais determina que a razão dos riscos instantâneos entre dois indivíduos quaisquer é constante ao longo do tempo, enquanto o risco basal é modelado de forma não paramétrica.

Conforme descrito em Cox (1972), o modelo de Cox define que a função de risco em relação ao tempo até a ocorrência do evento de interesse é dada por:

$$h(t | \mathbf{X}) = h_0(t) \exp \left\{ \boldsymbol{\beta}^\top \mathbf{X} \right\}, \quad t \geq 0 \quad (2.1)$$

em que $h_0(\cdot)$ representa a função de risco basal, comum a todos os indivíduos; $\mathbf{X} = (X_1, X_2, \dots, X_p)^\top$ denota o vetor de covariáveis explicativas e $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$ corresponde ao vetor de parâmetros da regressão associados às respectivas covariáveis.

Observe que, ao considerar o vetor de covariáveis \mathbf{X} como constante ao longo do tempo, a função de risco é definida com base na população em risco, ou seja, indivíduos que ainda não experimentaram o evento de interesse ou nem foram removidos do estudo por outros motivos, é dizer, por censura até o tempo em análise. Veja que a definição (2.1) enfatiza que para dois indivíduos com vetores de covariáveis distintos, \mathbf{X}^* e $\tilde{\mathbf{X}}$, suas funções de risco serão iguais somente se $\boldsymbol{\beta}^\top \mathbf{X}^* = \boldsymbol{\beta}^\top \tilde{\mathbf{X}}$. Assim, a hipótese de riscos proporcionais do modelo de Cox implica que a razão dos riscos instantâneos

$$\frac{h(t | \mathbf{X}^*)}{h(t | \tilde{\mathbf{X}})} = \exp \left(\boldsymbol{\beta}^\top (\mathbf{X}^* - \tilde{\mathbf{X}}) \right)$$

é constante ao longo do tempo t , caso essa suposição não for atendida, os efeitos das covariáveis tornam-se dependentes do tempo, e o modelo pode ser generalizado para permitir coeficientes $\boldsymbol{\beta}(t)$ que variam no tempo, conforme:

$$h(t | \mathbf{X}) = h_0(t) \exp \left(\boldsymbol{\beta}(t)^\top \mathbf{X} \right)$$

De modo geral, assume-se que o modelo (2.1) seja válido para um vetor de covariáveis $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, cujo vetor de coeficientes de regressão associado é dado por $\boldsymbol{\beta} = (\beta_1, \beta_2)$. Neste caso, esse modelo é reescrito como:

$$h(t | \mathbf{X}) = h_0(t) \exp \{ \beta_1 \mathbf{X}_1 + \beta_2 \mathbf{X}_2 \}. \quad (2.2)$$

Contudo, não é possível observar e incorporar todas as covariáveis significativas no modelo. Nessa situação, as covariáveis excluídas introduzem uma fonte de heterogeneidade não observada, e as diferenças entre os indivíduos são explicadas apenas pelas covariáveis disponíveis. Como forma de modelar essa heterogeneidade latente, introduz-se uma variável aleatória que atua multiplicativamente sobre a função de risco, representando os efeitos das covariáveis não observadas.

Por exemplo, ao considerar que o modelo de Cox (2.2) incluindo apenas a covariável \mathbf{X}_1 , pode-se definir uma variável aleatória $Z = \exp\{\beta_2 X_2\}$, em que \mathbf{X}_2 é a covariável não observada. Assim, o modelo passa a ser representado por:

$$h(t | \mathbf{X}) = Zh_0(t) \exp\{\beta_1 \mathbf{X}_1\}$$

em que Z é denominado termo de *fragilidade*, capturando a variabilidade não explicada pelas covariáveis observadas.

2.2 Modelos de fragilidade

2.2.1 Modelo de fragilidade univariado

O estudo de dados de sobrevivência com frequência é baseada sob a suposição de que a população em estudo é homogênea, de modo que, ao se condicionar às covariáveis observadas, todos os indivíduos possuem o mesmo risco de ocorrência do evento de interesse. Além disso, assume-se a independência entre os tempos de sobrevivência. No entanto, tais pressupostos nem sempre são realistas, especialmente em contextos nos quais a amostra é heterogênea e os indivíduos apresentam diferentes níveis de risco e as funções de risco distintas. Mesmo quando os indivíduos aparentam ser semelhantes com base nas covariáveis observadas, diferenças não mensuradas ou não observadas podem persistir. Para lidar com essa variabilidade latente, diversos autores como [Wienke \(2010\)](#), [Vaupel, Manton e Stallard \(1979\)](#) e [Balan e Putter \(2020\)](#) propõem o uso de modelos com efeitos aleatórios, conhecidos como modelos de fragilidade, os quais introduzem termos aleatórios multiplicativos na função de risco para capturar a heterogeneidade não observada entre os indivíduos.

O modelo de fragilidade tradicional, amplamente utilizado na literatura, baseia-se na estrutura de riscos proporcionais e condiciona a função de risco a um efeito aleatório Z , denominado variável de fragilidade. Esse termo representa a heterogeneidade não observada, sendo considerado não mensurável diretamente e independente do tempo. Neste contexto, e de acordo com a proposta deste trabalho, adota-se a estrutura de riscos multiplicativos, na qual a variável Z atua de forma multiplicativa sobre a função de risco de base. Assim, a função de risco condicional assume a forma:

$$h(t | Z) = Zh_0(t); \quad t > 0, \quad (2.3)$$

em que Z captura os efeitos aleatórios não observados que afetam a propensão ao evento de interesse e $h_0(t)$ é a função de risco base comum para todos os indivíduos. Em relação à identificabilidade do modelo apresentado em (2.3) é destacada por [Wienke \(2010\)](#), ao assumir que a distribuição da variável de fragilidade Z é padronizada, de forma que $\mathbb{E}[Z] = 1$, desde que o valor esperado exista. A variância da fragilidade, denotada por $\text{Var}[Z] = \theta > 0$ (quando definida), é interpretada como uma média de heterogeneidade não observada na população. Nesse contexto, valores baixos de θ indicam que os valores de Z estão concentrados em torno de 1, sugerindo

baixa heterogeneidade entre os indivíduos. Por outro lado, valores mais elevados de θ refletem maior dispersão de Z , indicando maior variação nos riscos individuais [Wienke \(2010\)](#). Além disso, note que a inclusão de covariáveis pode ser incorporada ao modelo (2.3) por meio da estrutura clássica de regressão da função de risco, preservando a multiplicatividade entre os componentes. Nesse caso, a função de risco condicional, dado o vetor de covariáveis \mathbf{X} e o termo de fragilidade Z , é expressa como:

$$h(t | Z, \mathbf{X}) = Zh_0(t) \exp(\boldsymbol{\beta}^\top \mathbf{X}), \quad (2.4)$$

em que $\mathbf{X} = (X_1, X_2, \dots, X_p)^\top$ representa o vetor de covariáveis e $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)^\top$ corresponde ao vetor dos parâmetros de regressão associados. Note que, ao considerar covariáveis, o modelo (2.4) é composto pela multiplicação de duas componentes: uma não paramétrica, $h_0(t)$, e outra paramétrica, $\exp(\boldsymbol{\beta}^\top \mathbf{X})$. Dessa forma, indivíduos com maior valor da variável de fragilidade Z são considerados mais frágeis, o que implica que eles tendem a vivenciar o evento de interesse mais precocemente em relação aos demais participantes do estudo [Balan e Putter \(2020\)](#).

Com o propósito de facilitar a abordagem teórica, neste capítulo será adotado o modelo definido em (2.3) como base para a formalização das definições subsequentes. Assim, considerando a função de risco condicional (2.3) e dado o valor da variável de fragilidade Z , a probabilidade que representará a proporção dos indivíduos sobreviventes após o tempo $t > 0$, condicionada à variável Z , é dada por

$$S(t | Z) = \exp\left(-\int_0^t h(u | Z) du\right) = \exp\left(-Z \int_0^t h_0(u) du\right) = \exp(-ZH_0(t)), \quad (2.5)$$

em que $H_0(t) = \int_0^t h_0(u) du$ representa a função de risco acumulado de base. No entanto, como a variável de fragilidade Z não é observável na prática, torna-se necessário considerar o modelo marginal (ou populacional), obtido ao calcular a esperança da função de sobrevivência condicional (2.5) em relação à distribuição de Z . Isso leva à seguinte expressão

$$S(t) = \int_0^\infty S(t | Z)g(z) dz = \int_0^\infty \exp(-ZH_0(t))g(z) dz = \mathbb{E}[\exp(-ZH_0(t))], \quad (2.6)$$

em que $g(\cdot)$ é a função densidade de probabilidade da variável de fragilidade Z . Observe que a função de sobrevivência marginal $S(t)$ apresentada em (2.6) corresponde à transformada de Laplace da variável aleatória Z , avaliada em $s = H_0(t)$, isto é, $\mathcal{L}_Z(s) = \mathbb{E}[\exp\{-Zs\}]$. Dessa forma, o modelo marginal (2.6) pode ser representado de forma compacta por

$$S(t) = \mathcal{L}_Z(H_0(t)), \quad t > 0. \quad (2.7)$$

Na literatura, diversos autores, como [Aalen \(1992\)](#), [Wienke \(2010\)](#) e [Balan e Putter \(2020\)](#), destacam a relação fundamental entre a função de sobrevivência marginal e a transformada de Laplace da variável de fragilidade. Essa relação fornece uma ferramenta analítica poderosa para descrever a interação entre as funções de risco, de sobrevivência e a função densidade

de probabilidade (f.d.p). Especificamente, [Aalen \(1992\)](#) enfatiza a utilidade das derivadas da transformada de Laplace na obtenção de propriedades gerais da distribuição de sobrevivência. A r -ésima derivada da transformada de Laplace associada ao modelo de sobrevivência marginal dado por (2.7) é expressa por

$$\mathcal{L}_Z^{(r)} = (-1)^r \mathbb{E}[Z^r \exp\{-Zs\}]$$

enquanto o r -ésimo momento da variável de fragilidade é obtido por

$$\mathbb{E}[Z^r] = (-1)^r \mathcal{L}_Z^{(r)}(0). \quad (2.8)$$

Adicionalmente, [Wienke \(2010\)](#) estabelece expressões que relacionam diretamente a transformada de Laplace com a função de sobrevivência e a densidade do tempo de falha. A f.d.p associada ao tempo de falha é dada por

$$f(t) = -\frac{dS(t)}{dt} = -h_0(t) \mathcal{L}'_Z(H_0(t))$$

em que \mathcal{L}'_Z denota a primeira derivada da transformada de Laplace. Por sua vez, a função de risco marginal pode ser expressa como

$$h(t) = \frac{f(t)}{S(t)} = -\frac{h_0(t) \mathcal{L}'_Z(H_0(t))}{\mathcal{L}_Z(H_0(t))}$$

Com base nos momentos definidos em (2.8), a esperança e a variância de Z podem ser expressas por

$$\mathbb{E}[Z] = -\mathcal{L}'_Z(0) \quad (2.9)$$

e

$$\text{Var}[Z] = \mathcal{L}''_Z(0) - [-\mathcal{L}'_Z(0)]^2 \quad (2.10)$$

em que \mathcal{L}'_Z e \mathcal{L}''_Z representam a primeira e segunda derivada da transformada de Laplace. Complementarmente, [Balan e Putter \(2020\)](#) apontam as seguintes propriedades úteis da transformada de Laplace associada à fragilidade

- $\mathcal{L}_Z(0) = 1$;
- O coeficiente de variação ao quadrado da fragilidade Z é dado por:

$$\text{CV}^2(Z) = \frac{\mathcal{L}''_Z(0)}{(\mathcal{L}'_Z(0))^2} - 1.$$

Tais propriedades reforçam o papel central da transformada de Laplace na modelagem de fragilidade. A ampla presença de estudos na literatura sobre modelos de fragilidade evidencia seu valor analítico e interpretativo, especialmente no tratamento da heterogeneidade não observada em populações. No entanto, vale ressaltar que, em geral, assume-se que a distribuição da fragilidade Z seja fixa ao longo do tempo para cada indivíduo. Embora essa suposição simplifique

a modelagem, não reflete completamente a dinâmica populacional observada. Como destacado por [Vaupel, Manton e Stallard \(1979\)](#), à medida que o tempo evolui, a composição da população muda: indivíduos com maior fragilidade tendem a experimentar o evento de interesse mais precocemente, resultando em uma população remanescente progressivamente mais robusta. Assim, a distribuição da fragilidade entre os indivíduos em risco se altera ao longo do tempo, estabelecendo uma conexão importante entre os modelos condicionais e marginais, formalizada no teorema a seguir.

Teorema 1. ([Vaupel, Manton e Stallard \(1979\)](#)) Assuma o modelo de fragilidade dado por (2.3). A função de risco populacional, $h(t) = \frac{f(t)}{S(t)}$ é geralmente dada por: $h(t) = \mathbb{E}[h(t | Z) | T > t]$, ou mais especificamente:

$$h(t) = \int_0^{\infty} h(t | z) f(z | T > t) dz = h_0(t) \int_0^{\infty} z f(z | T > t) dz,$$

com $f(z | T > t)$ representando a f.d.p. condicional de fragilidade dos sobreviventes por um período t .

A demonstração formal do teorema, pode ser consultada em [Wienke \(2010\)](#). No entanto, é importante destacar que a função de risco marginal pode ser interpretada como a média ponderada dos riscos individuais entre os indivíduos ainda em risco no tempo t . Essa ponderação depende da distribuição da variável de fragilidade Z , que atua como um fator multiplicativo no risco individual. Adicionalmente, observa-se que indivíduos com maiores valores de Z , isto é, mais frágeis tendem a experimentar o evento de interesse mais precocemente. Como consequência, a composição da população sobrevivente ao longo do tempo é gradualmente enriquecida por indivíduos menos frágeis (com menores valores de Z). Além disso, a f.d.p. condicional da variável de fragilidade Z , restrita à subpopulação sobrevivente até o tempo t , usada nesta demonstração e considerando as expressões de sobrevivência condicional e marginal fornecidas em (2.5) e (2.7), sendo dada por:

$$f(z | T > t) = \frac{S(t | z) f_Z(z)}{S(t)} = \frac{\exp(-zH_0(t)) f_Z(z)}{\mathcal{L}_Z(H_0(t))}.$$

2.3 Principais modelos de fragilidade

O uso de modelos de fragilidade tem se consolidado como uma ferramenta relevante na análise de dados de sobrevivência em contextos que envolvem heterogeneidade não observada. Estudos recentes, como os de [Cancho *et al.* \(2021\)](#), [Cancho *et al.* \(2018\)](#), [Macera \(2015\)](#), [Zavaleta \(2016\)](#) e [Fortes \(2020\)](#), destacam a aplicação desses modelos em contextos contínuos, não negativos e com estrutura multiplicativa, explorando diferentes abordagens probabilísticas e inferenciais. Na modelagem da variável de fragilidade, é comum o uso de distribuições amplamente discutidas na literatura estatística, como a Gama, Log-Normal, inversa Gaussiana (IG) e a distribuição Potência Exponencial (PE). Essas distribuições são escolhidas por sua

flexibilidade e propriedades matemáticas adequadas para capturar a variabilidade não observada nos riscos individuais. Com o intuito de unificar e generalizar essas abordagens, [Wienke \(2010\)](#) introduz a família PVF, uma classe ampla de distribuições de fragilidade que contém como casos particulares todas as mencionadas anteriormente. A família PVF fornece uma estrutura parametrizada que permite acomodar diferentes formas de heterogeneidade por meio de uma formulação única. Nesta seção, procede-se à caracterização formal da família PVF, bem como das principais distribuições específicas comumente utilizadas em modelos de fragilidade. Essa abordagem permite comparar propriedades analíticas, como a função de sobrevivência marginal, a transformada de Laplace associada, e as implicações inferenciais de cada escolha de distribuição para Z .

2.3.1 Distribuição de fragilidade Gama

O modelo de fragilidade Gama é amplamente reconhecido na literatura como uma escolha natural para modelar heterogeneidade não observada em dados de sobrevivência, sendo frequentemente utilizado como distribuição de mistura em modelos com riscos multiplicativos. Sua popularidade decorre, em grande parte, da forma analítica simples de sua transformada de Laplace, o que facilita o tratamento da função de sobrevivência marginal. Além disso, conforme destacado por [Wienke \(2010\)](#), uma importante propriedade assintótica reforça sua adequação: sob condições gerais, para uma ampla classe de modelos de fragilidade univariados, a distribuição condicional da fragilidade entre os sobreviventes tende, ao longo do tempo, a uma distribuição Gama. Tal propriedade torna esse modelo especialmente atrativo para análises de sobrevivência com censura e heterogeneidade não observada. Formalmente, seja Z uma variável aleatória com distribuição Gama, então sua f.d.p é dada por:

$$g(z) = \frac{1}{\Gamma(\mu)} \sigma^\mu z^{\mu-1} \exp(-\sigma z), \quad z > 0, \sigma > 0, \mu > 0.$$

Enquanto, a transformada de Laplace correspondente é expressa por:

$$\mathcal{L}_Z(s) = \left(1 + \frac{s}{\sigma}\right)^{-\mu}. \quad (2.11)$$

Utilizando os resultados de momentos vistos em (2.9) e (2.10), tem-se que: $\mathbb{E}[Z] = \frac{\mu}{\sigma}$ e $\text{Var}[Z] = \frac{\mu}{\sigma^2}$. E para garantir a identificabilidade no contexto de modelos de fragilidade, é usual impor a restrição $\mathbb{E}[Z] = 1$. Isso é obtido ao assumir $\mu = \sigma$, de modo que $\mathbb{E}[Z] = 1$ e $\theta := \text{Var}[Z] = \frac{1}{\sigma}$; em que θ passa a parametrizar diretamente o grau de heterogeneidade não observada na população. Sob essa parametrização, tem-se: $Z \sim \Gamma\left(\frac{1}{\theta}, \frac{1}{\theta}\right)$, e substituindo na transformada de Laplace descrita em (2.11) é reescrita como

$$\mathcal{L}_Z(s) = (1 + \theta s)^{-\frac{1}{\theta}}. \quad (2.12)$$

2.3.2 Distribuição de fragilidade Inversa Gaussiana (IG)

A distribuição IG (*Inverse Gaussian*) foi formalmente introduzida por Hougaard (1984) como uma alternativa à distribuição Gama no contexto de modelos de fragilidade. Posteriormente, Wienke (2010) enfatizou que, de modo análogo ao modelo de fragilidade Gama, a IG admite expressões analíticas simples e em forma fechada para as funções marginais de sobrevivência e de risco, o que a torna particularmente atraente em modelagem de sobrevivência sob heterogeneidade não observada. Seja Z uma variável aleatória com distribuição IG parametrizada por média $\mu > 0$ e parâmetro de dispersão $\sigma > 0$. Sua f.d.p é dada por

$$g(z) = \sqrt{\frac{\sigma}{2\pi z^3}} \exp\left\{-\frac{\sigma(z-\mu)^2}{2\mu^2 z}\right\}, \quad z > 0.$$

Enquanto que a sua transformada de Laplace de Z é expressa por

$$\mathcal{L}_Z(s) = \exp\left\{\frac{\sigma}{\mu} \left(1 - \sqrt{1 + \frac{2\mu^2 s}{\sigma}}\right)\right\}. \quad (2.13)$$

A partir das propriedades da transformada de Laplace dadas em (2.9) e (2.10), obtêm-se diretamente $\mathbb{E}[Z] = \mu$ e $\text{Var}(Z) = \frac{\mu^3}{\sigma}$. Porém, no contexto de modelos de fragilidade, a identificabilidade é usualmente assegurada impondo-se a restrição $\mu = 1$. Sob essa parametrização padronizada, tem-se $\mathbb{E}[Z] = 1$, e $\theta := \text{Var}(Z) = \frac{1}{\sigma}$ em que θ representa o parâmetro de heterogeneidade. Consequentemente, a transformada de Laplace pode ser reescrita como

$$\mathcal{L}_Z(s) = \exp\left\{\frac{1}{\theta} \left(1 - \sqrt{1 + 2\theta s}\right)\right\}. \quad (2.14)$$

Com esta parametrização θ controla a variabilidade não observada, preservando a tratabilidade analítica do modelo.

2.3.3 Distribuição de fragilidade Positiva Estável (PE)

Bedia (2022) caracteriza uma distribuição como Positiva Estável (PE) quando a soma de n variáveis aleatórias independentes e identicamente distribuídas (i.i.d.) pode ser expressa, em distribuição, como um múltiplo determinístico de uma única dessas variáveis. Formalmente, uma distribuição é PE se, para variáveis aleatórias i.i.d. Z_1, \dots, Z_n , existe, para cada n , uma constante de normalização $c(n)$ tal que

$$c(n)Z_1 \stackrel{d}{=} Z_1 + \dots + Z_n.$$

Isto é, ao normalizar a soma $Z_1 + \dots + Z_n$ por $c(n)$, obtêm-se uma variável aleatória que preserva a mesma forma distributiva de Z_1 . Bedia (2022) destaca ainda que essa constante de normalização assume a forma $c(n) = n^{1/\gamma}$, com $\gamma \in (0, 2]$. Em que o parâmetro γ é o índice de estabilidade. Quando $\gamma = 2$, recupera-se como caso particular a distribuição Normal. Entretanto, para aplicações que requerem suporte estritamente positivo, considera-se $\gamma \in (0, 1)$, garantindo

que a distribuição resultante seja definida apenas para valores positivos. Sob essa parametrização, a função densidade de probabilidade de uma variável aleatória como distribuição Positiva Estável (PE), para $z > 0$ e $0 < \gamma \leq 1$, é dada

$$g(z) = \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k+1} \frac{\Gamma(k\gamma + 1)}{k!} z^{-k\gamma-1} \sin(k\gamma\pi). \quad (2.15)$$

A expressão em (2.15) consiste em uma série de potências. Sua convergência é rápida para valores grandes de z , mas torna-se lenta quando z assume valores pequenos. Além disso, quando $\gamma = 1$, a distribuição de fragilidade se degenera em uma massa pontual em $Z = 1$, [Wienke \(2010\)](#), também destaca que todos os momentos dessa distribuição são infinitos. Consequentemente, a esperança da fragilidade diverge e a variância não existe. Por outro lado, a transformada de Laplace da distribuição de fragilidade PE é dada por

$$\mathcal{L}_Z(s) = \exp\{-s^\gamma\}, \quad (2.16)$$

conforme apresentado em [Wienke \(2010\)](#).

2.3.4 Distribuição de fragilidade para a família PVF

A família PVF é definida por [Wienke \(2010\)](#) como uma classe de distribuições de fragilidade cuja principal característica é a existência de uma transformada de Laplace em forma fechada. Essa propriedade permite derivar expressões analíticas explícitas para as funções de sobrevivência marginal, densidade de probabilidade e função de risco, facilitando significativamente os procedimentos de inferência, especialmente na estimação de parâmetros. Outro aspecto relevante é que a família PVF generaliza diversas distribuições clássicas utilizadas em modelos de fragilidade, as quais podem ser recuperadas como casos particulares mediante restrições nos parâmetros. Essa estrutura é destacada nos estudos de [Calsavara et al. \(2017\)](#) e [Rodrigues et al. \(2021\)](#), que enfatizam o uso da PVF como base para modelagem em contextos com heterogeneidade não observada. Dentre os membros notáveis dessa família, incluem-se as distribuições Gama, Inversa Gaussiana (IG), Hougaard, Potência Exponencial (PE) e a Log-Normal (em contextos específicos). Embora [Wienke \(2010\)](#) apresente a família PVF, neste trabalho será adotada a parametrização proposta por [Hougaard \(2000\)](#), que oferece uma formulação alternativa com propriedades analíticas convenientes. Em particular, [Bedia \(2022\)](#) descreve essa parametrização como $PVF(\gamma, \mu, \sigma)$, em que os parâmetros satisfazem:

- $\gamma \leq 1$: parâmetro de forma que regula a estrutura da variância;
- $\mu > 0$: parâmetro de média;
- σ : parâmetro de escala, com
 - $\sigma \geq 0$ se $0 < \gamma \leq 1$,

– $\sigma > 0$ se $\gamma \leq 0$.

Sob essa formulação, a família PVF continua incluindo distribuições utilizadas em modelos de fragilidade, como a Gama, a Potência Exponencial (PE) e a Inversa Gaussiana (IG). Uma propriedade distintiva da classe é que a variância da variável aleatória de fragilidade Z é uma função potência de sua média, ou seja $\text{Var}[Z] \propto \mathbb{E}[Z]^\gamma$, o que justifica a denominação Família de Funções de Variância de Potência; indicado em [Hougaard \(2000\)](#) e em suas contribuições, apresenta a função densidade de probabilidade (f.d.p.) da família PVF sob duas formas distintas, dependendo do valor de γ .

- **Caso 1:** $0 < \gamma \leq 1$. Nesta situação, Z assume apenas valores estritamente positivos. A f.d.p. é dada por

$$g(z) = \exp \left\{ -\sigma z + \frac{\mu \sigma^\gamma}{\gamma} \right\} \left(-\frac{1}{\pi z} \right) \sum_{k=1}^{\infty} \frac{\Gamma(k\gamma + 1)}{k!} \left(-z^{-\gamma} \frac{\mu}{\gamma} \right)^k \sin(\gamma k \pi)$$

- **Caso 2:** $\gamma \leq 0$. Neste cenário, a densidade admite suporte em $[0, \infty)$, possibilitando a existência de indivíduos com risco nulo (i.e. $Z = 0$). A f.d.p. é definida como

$$g(z) = \exp \left\{ -\sigma z + \frac{\mu \sigma^\gamma}{\gamma} \right\} \left(\frac{1}{z} \right) \sum_{k=1}^{\infty} \frac{\left(\frac{-\mu z^{-\gamma}}{\gamma} \right)^k}{k! \Gamma(-k\gamma)}. \quad (2.17)$$

Essa representação é válida para todos os valores de $\gamma \in \mathbb{R} \setminus \{0, 1\}$, adotando-se a convenção de que os termos da série com função Gama indefinida no denominador (ou seja, quando $k\gamma \in \mathbb{Z}_+$) são nulos, conforme sugerido por [Hougaard \(2000\)](#).

A transformada de Laplace da variável $Z \sim \text{PVF}(\gamma, \mu, \sigma)$ é dada por

$$\mathcal{L}_Z(s) = \exp \left\{ -\frac{\mu}{\gamma} [(\sigma + s)^\gamma - \sigma^\gamma] \right\}. \quad (2.18)$$

A partir dessa transformada com base nas expressões gerais para os momentos via derivadas da transformada de Laplace dadas em (2.9) e (2.10), obtêm-se os momentos da variável de fragilidade:

$$\mathbb{E}[Z] = \mu \sigma^{\gamma-1} \quad \text{e} \quad \text{Var}[Z] = \mu(1 - \gamma) \sigma^{\gamma-2}.$$

Estas expressões evidenciam o papel dos parâmetros γ e σ no controle da dispersão e da estrutura de dependência no modelo de fragilidade, sendo fundamentais para a especificação adequada da heterogeneidade não observada na população em estudo.

2.4 Modelo Exponencial por Partes (MEP)

O modelo de risco MEP constitui uma alternativa semiparamétrica útil no contexto de modelos de sobrevivência. Ele oferece flexibilidade ao permitir que a função de risco apresente

formas variadas por meio da divisão do eixo temporal em intervalos, sem impor uma forma funcional fixa para todo o tempo. Esse modelo tem sido adotado por autores como [Ibrahim, Chen e Sinha \(2001\)](#) e [Sibim e Cancho \(2011\)](#). Formalmente, assume-se uma variável aleatória $T \geq 0$ denotando o tempo até o evento de interesse. Define-se uma partição $\tau = \{s_0, s_1, \dots, s_J\}$ de $[0, \infty)$ tal que $0 = s_0 < s_1 < \dots < s_J < \infty$. Para cada intervalo disjunto $I_j = (s_{j-1}, s_j]$ com $j = 1, \dots, J$, define-se que o risco instantâneo é constante, igual a $\lambda_j > 0$. Assim, a função de risco acumulado no modelo MEP para $t \in I_j$ é

$$H_{\text{MEP}}(t) = \sum_{m=1}^{j-1} \lambda_m (s_m - s_{m-1}) + \lambda_j (t - s_{j-1}). \quad (2.19)$$

Observe que essa função acumulada pode ser interpretada geometricamente como a soma das áreas de retângulos com base nos comprimentos dos intervalos definidos pela partição τ e altura constante determinada pelos valores da função de risco λ_j em cada subintervalo. Essa estrutura facilita tanto a interpretação quanto a implementação computacional do modelo. Com base na relação $S(t) = \exp\{-H(t)\}$, a função de sobrevivência associada é

$$S_{\text{MEP}}(t) = \begin{cases} \exp\{-\lambda_1 t\}, & t \in I_1 \\ \exp\left\{-\sum_{m=1}^{j-1} \lambda_m (s_m - s_{m-1}) - \lambda_j (t - s_{j-1})\right\}, & t \in I_j \end{cases} \quad (2.20)$$

A densidade de probabilidade correspondente é então expressa por

$$f_{\text{MEP}}(t) = \begin{cases} \lambda_1 \exp\{-\lambda_1 t\}, & t \in I_1 \\ \lambda_j \exp\left\{-\sum_{m=1}^{j-1} \lambda_m (s_m - s_{m-1}) - \lambda_j (t - s_{j-1})\right\}, & t \in I_j \end{cases} \quad (2.21)$$

Além disso, as funções de risco e de distribuição acumulada assumem as formas:

$$h_{\text{MEP}}(t) = \begin{cases} \lambda_1, & t \in I_1 \\ \lambda_j, & t \in I_j \end{cases} \quad (2.22)$$

e

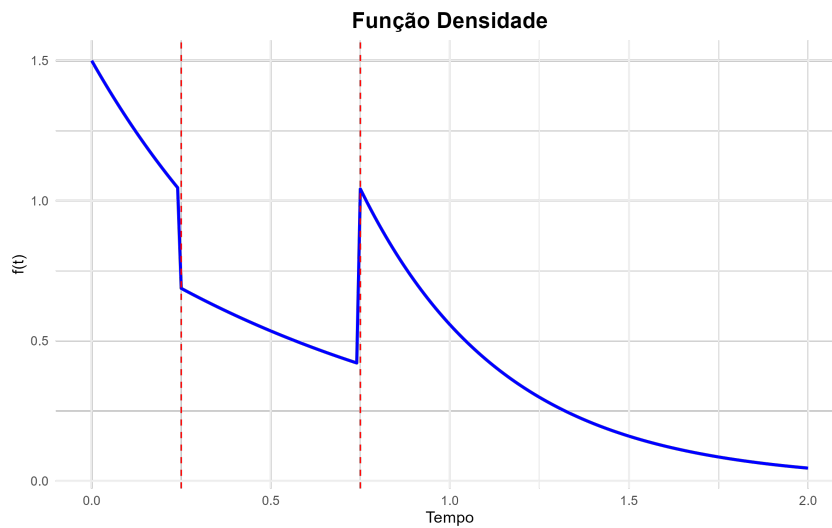
$$F_{\text{MEP}}(t) = \begin{cases} 1 - \exp\{-\lambda_1 t\}, & t \in I_1 \\ 1 - \exp\left\{-\sum_{m=1}^{j-1} \lambda_m (s_m - s_{m-1}) - \lambda_j (t - s_{j-1})\right\}, & t \in I_j \end{cases} \quad (2.23)$$

Note que:

- Quando $J = 1$, o modelo MEP se reduz ao modelo exponencial simples, pois há apenas um único intervalo e risco constante para todo o tempo.
- À medida que J aumenta e os intervalos tornam-se mais finos, o modelo pode aproximar qualquer forma ajustada de função de risco subjacente, oferecendo grande flexibilidade para capturar padrões reais presentes nos dados.

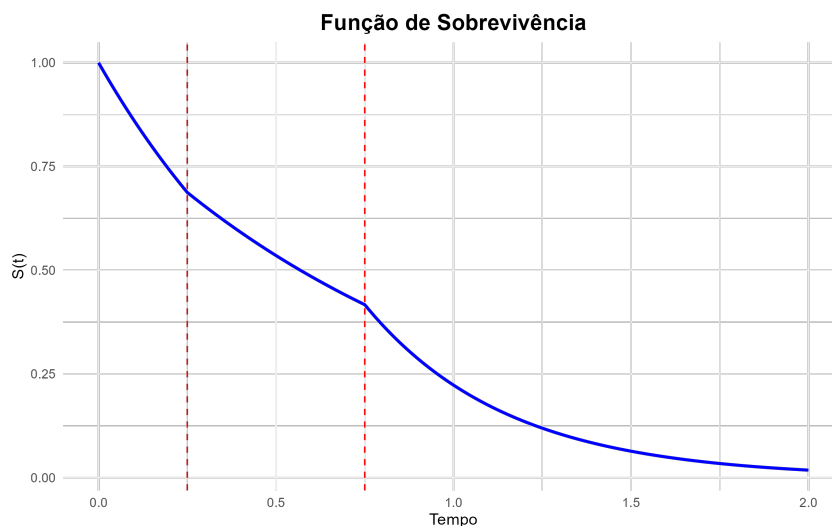
As Figuras 1, 2, 3 e 4 ilustram os gráficos da função de sobrevivência, a função densidade, a distribuição acumulada e do risco acumulado, respectivamente dadas em (2.20), (2.21), (2.22) e (2.23), para o modelo MEP com $\lambda_1 = 2$, $\lambda_2 = 1$ e $\lambda_3 = 3$. Os mesmos foram gerados no R através do pacote PWEEXP Team (2020a), considerando três partições no eixo dos tempos ($J = 3$) com $I_1 = (0, 0,3]$, $I_2 = (0,3, 0,8]$ e $I_3 = (0,8, \infty)$ para $n = 151$.

Figura 1 – Gráfico da f.d.p do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$.



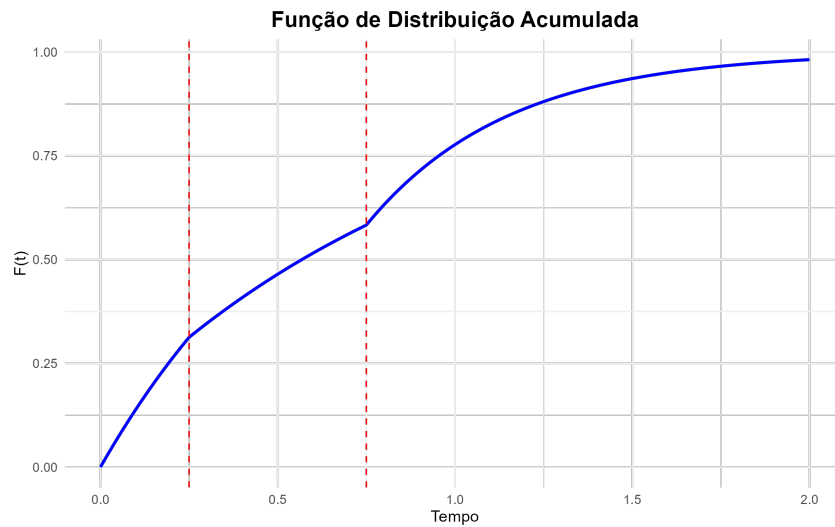
Fonte: Elaborada pelo autor.

Figura 2 – Gráfico da função de sobrevivência do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$.



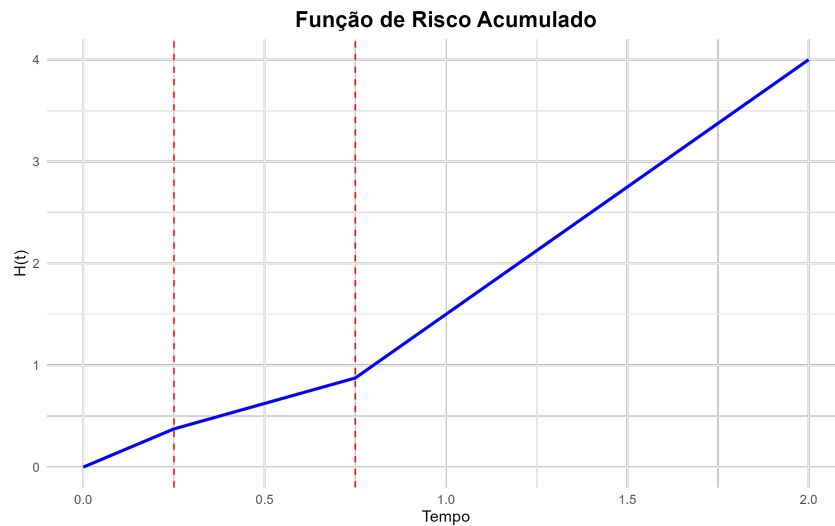
Fonte: Elaborada pelo autor.

Figura 3 – Gráfico da função de distribuição acumulada do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$.



Fonte: Elaborada pelo autor.

Figura 4 – Gráfico da função de risco acumulado do modelo MEP considerando $J = 3$ partições no eixo dos tempos com $I_1 = (0; 0,25]$, $I_2 = (0,25; 0,75]$ e $I_3 = (0,75; \infty)$ para $n = 200$, $\lambda_1 = 1,5$, $\lambda_2 = 1$ e $\lambda_3 = 2,5$.



Fonte: Elaborada pelo autor.

A escolha da partição $\tau = \{s_0, \dots, s_J\}$ desempenha um papel crucial na qualidade do ajuste do modelo MEP. Especificamente, uma partição com número excessivo de subintervalos pode levar a estimativas instáveis das taxas de risco λ_j , devido ao reduzido número de eventos por intervalo, o que compromete a precisão inferencial. Por outro lado, uma partição com poucos intervalos pode resultar em uma representação inadequada da verdadeira estrutura temporal da função de risco, prejudicando a capacidade do modelo de capturar variações locais relevantes.

Assim, a definição de uma partição apropriada consiste em encontrar um equilíbrio entre a flexibilidade do modelo e a estabilidade das estimativas, de modo a garantir uma boa aproximação tanto para a função de risco quanto para a função de sobrevivência. Esta questão é reconhecida como um dos principais desafios associados à aplicação prática do modelo MEP, conforme discutido em [Demarqui \(2010\)](#). Para uma análise aprofundada deste modelo e de suas propriedades inferenciais, recomenda-se a leitura dos trabalhos de [Ibrahim, Chen e Sinha \(2001\)](#), [Sibim e Cancho \(2011\)](#) e [Demarqui \(2010\)](#).

2.5 Inferência Bayesiana

A inferência estatística visa estudar propriedades de uma população a partir da análise de uma quantidade de interesse, a inferência bayesiana modela a incerteza sobre os parâmetros por meio de distribuições probabilísticas, adotando uma interpretação subjetiva a probabilidade, na qual o pesquisador incorpora seu conhecimento prévio e crenças ao definir a distribuição priori [Paixão \(2021\)](#). Nesse contexto, o teorema de Bayes é aplicado para tratar os parâmetros desconhecidos como variáveis aleatórias, permitindo a integração de informações externas na análise dos dados.

Seja $\mathbf{x} = (x_1, \dots, x_n)$ uma amostra observada de tamanho n , e $\theta = (\theta_1, \dots, \theta_d)$ o vetor de parâmetros associados ao modelo assumido para \mathbf{x} . A distribuição a posteriori conjunta de θ dado \mathbf{x} , conforme o teorema de Bayes, é expressa por:

$$\pi(\theta | \mathbf{x}) = \frac{\pi(\theta, \mathbf{x})}{\pi(\mathbf{x})} = \frac{\pi(\mathbf{x} | \theta)\pi(\theta)}{\pi(\mathbf{x})} \propto \pi(\mathbf{x} | \theta)\pi(\theta), \quad (2.24)$$

em que $\pi(\mathbf{x} | \theta)$ é a função de verossimilhança, $\pi(\theta)$ é a distribuição a priori e $\pi(\mathbf{x})$ é a constante de normalização, também chamada de verossimilhança marginal, assim garantindo que a posteriori seja uma densidade válida. Essa expressão sintetiza o conhecimento atualizado sobre θ após a observação dos dados, combinando a informação prévia e a evidência empírica [Hartmann \(2015\)](#). Para realizar inferências sobre os parâmetros, é comum calcular a esperança condicional a posteriori de uma função $g : \Theta \rightarrow \mathbb{R}$, dada por:

$$\mathbb{E}_{\theta|\mathbf{x}}[g(\theta)] = \int_{\Theta} g(\theta) \pi(\theta | \mathbf{x}) d\theta, \quad (2.25)$$

em que Θ é o espaço paramétrico. Na prática, essa integral nem sempre admite solução analítica, demandando métodos numéricos para sua avaliação.

Entre tais métodos, destaca-se a simulação de Monte Carlo, originalmente proposta por [Metropolis e Ulam \(1949\)](#), que utiliza aproximações estocásticas para integrais multidimensionais. Baseando-se na Lei Forte dos Grandes Números, esse método possibilita aproximar a esperança posteriori por meio de uma amostra aleatória $\{\theta^{(i)}\}_{i=1}^N$ da distribuição $\pi(\theta | \mathbf{x})$, via a média empírica

$$\mathbb{E}_{\theta|\mathbf{x}}[\widehat{g(\theta)}] \approx \frac{1}{N} \sum_{i=1}^N g(\theta^{(i)}) \xrightarrow{N \rightarrow \infty} \mathbb{E}_{\theta|\mathbf{x}}[g(\theta)]$$

Entretanto, para modelos com alta dimensionalidade, a obtenção eficiente de amostras da distribuição a posteriori permanece desafiadora, motivando o desenvolvimento de algoritmos avançados, tais como Cadeias de Markov Monte Carlo (MCMC) [Metropolis *et al.* \(1953\)](#), o algoritmo Metropolis-Hastings (MH) [Hastings \(1970\)](#) e o Hamiltonian

Entretanto, para modelos com alta dimensionalidade, a obtenção eficiente de amostras da distribuição a posteriori permanece desafiadora, motivando o desenvolvimento de algoritmos avançados, tais como Cadeias de Markov Monte Carlo (MCMC) [Metropolis *et al.* \(1953\)](#), o algoritmo Metropolis-Hastings (MH) [Hastings \(1970\)](#) e o Monte Carlo Hamiltoniano (MCH) [Duane *et al.* \(1987\)](#), que melhoram a exploração do espaço paramétrico e a convergência da simulação.

2.5.1 Método de Monte Carlo via Cadeias de Markov (MCMC)

Os trabalhos de [Xavier \(2019\)](#) caracterizam uma cadeia de Markov como um processo estocástico que satisfaz a propriedade de memória curta: dado o estado atual, o estados passados e futuros são condicionalmente independentes. Em termos probabilísticos, essa propriedade implica que o comportamento futuro do processo depende apenas do presente, e não de como o processo chegou a esse estado. Esse autor apresenta essa definição de forma rigorosa com base na teoria de conjuntos e probabilidade. De maneira semelhante, [Hartmann \(2015\)](#) também formaliza o conceito de cadeia de Markov, além de introduzir definições fundamentais relacionadas à estrutura e dinâmica desses processos, as quais serão discutidas a seguir

Definição 1. Um processo estocástico é uma função de dois argumentos $X(t, \omega) : T \times \Omega \rightarrow \mathbb{R}$, em que T é um subconjunto dos reais. Assim, para $t^* \in T$ fixo, $X(t^*, \omega)$ é uma variável aleatória e para $\omega^* \in \Omega$ fixo, $X(t, \omega^*)$ é uma realização do processo, ou seja, uma função determinística.

Agora, assumamos uma dada coleção de tempos $t = (t_1, \dots, t_n)$. A função $\mathbf{X} = (X_{t_1}, \dots, X_{t_n})$ segue uma f.d.p. conjunta tal que

$$\mathbf{X} \sim g(\mathbf{x}); \quad \forall t \in T \text{ e } \forall n \in \mathbb{N},$$

em que $\mathbf{x} = (x_{t_1}, \dots, x_{t_n})$ e $X(t_n) = X_{t_n}$.

Definição 2. Um processo estocástico é dito ser uma Cadeia de Markov se

$$g(x_{t_s} | x_{t_{s-1}}, \dots, x_{t_1}) = g(x_{t_s} | x_{t_{s-1}}); \quad \forall t_s \in T \text{ e } \forall s \in \mathbb{N}.$$

Consequentemente, a f.d.p. conjunta pode ser escrita como:

$$g(\mathbf{x}) = g(x_{t_1}) \prod_{i=2}^n g(x_{t_i} | x_{t_{i-1}}).$$

Além disso, são necessárias algumas condições em relação à cadeia Markoviana para que o método de Monte Carlo seja adequado. Por meio destas condições, é garantido que

independente do valor inicial da cadeia, com o limite de $t \rightarrow \infty$, a função X_t será amostrada de uma distribuição invariante. Em consequência disso, é possível destacar outras três definições importantes: probabilidade de transição, distribuição invariante e equação de balanço detalhada. Estas definições são formalmente descritas a seguir, segundo [Hartmann \(2015\)](#).

Definição 3. (Probabilidade de transição) Para todo $x \in S$ e $A \subseteq S$ é definido $T^n(x, A)$ como a probabilidade condicional da cadeia se encontrar numa região A depois de n passos dado o início em x , ou seja,

$$T^n(x, A) = \mathbb{P}(X_n \in A \mid X_0 = x).$$

Ainda, $T^n(x, y)$ é dita a f.d.p. condicional, isto é,

$$T^n(x, y) = g_{X_n}(y \mid X_0 = x).$$

Definição 4. (Distribuição invariante) A Cadeia de Markov com probabilidade de transição $T(x, A)$ possui uma distribuição invariante $v(x)$ se, para todo conjunto $A \subset S$, segue que

$$\int_A v(x) dx = \int T(x, A)v(x) dx.$$

Definição 5. (Equação de balanço detalhada) A probabilidade de transição $T(x, y)$ satisfaz a equação de balanço detalhada com respeito à densidade $v(x)$ se, para todo $x, y \in S$, é válido que

$$T(x, y)v(x) = T(y, x)v(y).$$

Esta última definição destaca que, após um número suficiente de iterações, a taxa de transição da cadeia de Markov de \mathbf{x} para \mathbf{y} se iguala à taxa de transição de \mathbf{y} para \mathbf{x} [Hartmann \(2015\)](#). Em outras palavras, a cadeia alcança um comportamento simétrico nas transições entre os estados, o que é uma característica importante em processos ergódicos, garantindo que, ao longo do tempo, as probabilidades de transição sejam balanceadas entre os estados.

2.5.2 Monte Carlo Hamiltoniano (HMC)

[Hartmann \(2015\)](#) destaca que o método Hamiltonian Monte Carlo (HMC) consiste em explorar a analogia com a mecânica clássica, simulando o movimento de uma partícula em um sistema sob energia potencial. A cada iteração, a velocidade da partícula é aleatoriamente amostrada e seu trajeto é determinado segundo as equações da dinâmica Hamiltoniana. A nova posição da partícula, ao fim dessa trajetória simulada, corresponde a uma proposta para a próxima iteração da cadeia de Markov. A aceitação ou rejeição dessa proposta é realizada via o critério de MH.

Diversos autores, como [Xavier \(2019\)](#), [Paixão \(2021\)](#) e [Hartmann \(2015\)](#), apresentam formalizações do algoritmo HMC. Considerando o vetor de parâmetros aleatórios $\theta \in \mathbb{R}^d$ e um vetor auxiliar e independente $\mathbf{p} \in \mathbb{R}^d$, assume-se que $\mathbf{p} \sim N_d(\mathbf{0}, M)$, em que M é uma matriz de covariância positiva definida. O HMC define a chamada função Hamiltoniana como:

$$H(\theta, \mathbf{p}) = U(\theta) + C(\mathbf{p})$$

em que

$$U(\boldsymbol{\theta}) = -\ell(\boldsymbol{\theta}) = -\log[\pi(\boldsymbol{\theta} | \mathbf{x})] - \log[\pi(\boldsymbol{\theta})], \quad \text{e} \quad C(\mathbf{p}) = \frac{1}{2} \mathbf{p}^\top \mathbf{M}^{-1} \mathbf{p}.$$

com, $\pi(\boldsymbol{\theta} | \mathbf{x})$ representando a distribuição posteriori, enquanto $\pi(\boldsymbol{\theta})$ é a distribuição priori. Sob essas definições, a densidade conjunta de $(\boldsymbol{\theta}, \mathbf{p})$ assume a forma:

$$\pi(\boldsymbol{\theta}, \mathbf{p}) \propto \exp\{-H(\boldsymbol{\theta}, \mathbf{p})\} \propto \pi(\boldsymbol{\theta} | \mathbf{x})\pi(\boldsymbol{\theta}) \exp\left\{-\mathbf{p}^\top \mathbf{M}^{-1} \mathbf{p}\right\}.$$

O algoritmo HMC segue duas etapas fundamentais antes da aplicação do critério de aceitação de MH. Na primeira, amostra-se \mathbf{p} de uma distribuição normal multivariada centrada com covariância M . Na segunda, simula-se a trajetória do sistema dinâmico $(\boldsymbol{\theta}, \mathbf{p})$ de acordo com as equações diferenciais de Hamilton

$$\frac{\partial \boldsymbol{\theta}}{\partial t} = \frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial \mathbf{p}} = \nabla_{\mathbf{p}} C(\mathbf{p})$$

e

$$\frac{\partial \mathbf{p}}{\partial t} = -\frac{\partial H(\boldsymbol{\theta}, \mathbf{p})}{\partial \boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}} U(\boldsymbol{\theta})$$

em que $\nabla_{\mathbf{p}}$ e $\nabla_{\boldsymbol{\theta}}$ representam os vetores gradientes em relação aos vetores \mathbf{p} e $\boldsymbol{\theta}$, respectivamente.

Como a solução exata dessas equações nem sempre é viável, utiliza-se métodos numéricos de integração. O esquema de *Stormer-Verlet* (ou *leapfrog*) dada em [Leimkuhler e Matthews \(2015\)](#) é amplamente adotado por preservar propriedades importantes do sistema dinâmico e garantir a convergência das Cadeias de Markov para a distribuição alvo, características essenciais para a validade do algoritmo MCMC. Em suma, inicialmente, é determinado valores iniciais $(\boldsymbol{\theta}^{(I)}, \mathbf{p}^{(I)})$ com um tempo fictício $I = 0$ para que posteriormente, seja aplicado o método de *Stormer-Verlet* nas equações diferenciais da dinâmica Hamiltoniana de forma que

$$\begin{aligned} \mathbf{p}^{(I+\frac{\varepsilon}{2})} &= \mathbf{p}^{(I)} + \frac{\varepsilon}{2} \nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}^{(I)}), \\ \boldsymbol{\theta}^{(I+\varepsilon)} &= \boldsymbol{\theta}^{(I)} + \varepsilon \mathbf{M}^{-1} \mathbf{p}^{(I+\frac{\varepsilon}{2})}, \\ \mathbf{p}^{(I+\varepsilon)} &= \mathbf{p}^{(I+\frac{\varepsilon}{2})} + \frac{\varepsilon}{2} \nabla_{\boldsymbol{\theta}} \ell(\boldsymbol{\theta}^{(I+\varepsilon)}). \end{aligned}$$

Após simular a trajetória até um tempo fictício t_1 , obtém-se uma nova proposta $(\boldsymbol{\theta}^{(t_1)}, \mathbf{p}^{(t_1)})$. A aceitação dessa proposta é realizada com probabilidade

$$\mathbb{P}[(\boldsymbol{\theta}^{(t_1)}, \mathbf{p}^{(t_1)}), (\boldsymbol{\theta}^{(I)}, \mathbf{p}^{(I)})] = \min \left\{ 1, \exp \left[H(\boldsymbol{\theta}^{(I)}, \mathbf{p}^{(I)}) - H(\boldsymbol{\theta}^{(t_1)}, \mathbf{p}^{(t_1)}) \right] \right\}$$

Os parâmetros ε (passo de integração) e M (matriz de massa) afetam diretamente o desempenho do algoritmo. Passos muito grandes (ε elevado) podem comprometer a acurácia da integração, enquanto passos muito pequenos aumentam o custo computacional. A matriz M é comumente tomada como a identidade, especialmente quando não há informação prévia sobre a estrutura de covariância, embora escolhas mais sofisticadas possam melhorar a eficiência em modelos de alta dimensão.

Para exemplificar o algoritmo do método HMC, vamos supor que se deseja simular valores de $\pi(\theta | \mathbf{x})$ com $\theta \in \mathbb{R}^d$ e que M é a matriz identidade. Assim, as etapas do algoritmo são descritas a seguir:

1. Inicialização: Fixe $\theta^{(0)}$;
2. Para $i = 1, \dots, n$ (iterações da cadeia):
 - Gere $\mathbf{p}^* \sim N_d(\mathbf{0}, I)$ e $u \sim U(0, 1)$.
 - Defina $(\theta^{(I)}, \mathbf{p}^{(I)}) = (\theta^{(i-1)}, \mathbf{p}^*)$ e compute $H_0 = H(\theta^{(I)}, \mathbf{p}^{(I)})$
 - Aplique o integrador leapfrog por L passos:

$$\begin{aligned}\mathbf{p}^* &= \mathbf{p}^* + \frac{\varepsilon}{2} \nabla_{\theta} \ell(\theta^{(i-1)}), \\ \theta^{(i-1)} &= \theta^{(i-1)} + \varepsilon \nabla_{\mathbf{p}} C(\mathbf{p}^*) \\ \mathbf{p}^* &= \mathbf{p}^* + \frac{\varepsilon}{2} \nabla_{\theta} \ell(\theta^{(i-1)}).\end{aligned}$$

- Defina $(\theta^{(F)}, \mathbf{p}^{(F)}) = (\theta^{(i-1)}, \mathbf{p}^*)$ e calcule $H_1 = H(\theta^{(F)}, \mathbf{p}^{(F)})$.
- Calcule a probabilidade de aceitação:

$$P[(\theta^{(F)}, \mathbf{p}^{(F)}), (\theta^{(I)}, \mathbf{p}^{(I)})] = \min\{1, \exp(H_0 - H_1)\}.$$

- Atualize:

$$\theta^{(i)} = \begin{cases} \theta^{(F)}, & \text{com Probabilidade } \mathbb{P}[(\theta^{(F)}, \mathbf{p}^{(F)}), (\theta^{(I)}, \mathbf{p}^{(I)})] > u, \\ \theta^{(I)}, & \text{caso contrário} \end{cases}$$

Para uma discussão mais aprofundada, ver [Hartmann \(2015\)](#), [Paixão \(2021\)](#), [Xavier \(2019\)](#), [Hastings \(1970\)](#), [Metropolis *et al.* \(1953\)](#) e [Ibrahim, Chen e Sinha \(2001\)](#).

2.5.3 Critério de comparação de modelos bayesianos

Critérios de comparação de modelos constituem uma ferramenta fundamental na inferência Bayesiana, especialmente em contextos nos quais as distribuições a posteriori dos parâmetros são obtidas por meio de algoritmos baseados em cadeias de Markov via Monte Carlo (MCMC). Tais critérios permitem avaliar o ajuste e a adequação de diferentes modelos estatísticos aos dados observados, sendo particularmente úteis quando modelos concorrentes são considerados. Dentre os diversos critérios propostos na literatura, este trabalho adota o *Bridge Sampling* e a Ordenada Preditiva Condicional (*Conditional Predictive Ordinate* (CPO)). O bridge sampling é destacado por sua eficiência e precisão mesmo em espaços paramétricos de alta dimensão, conforme evidenciado por [Gronau, Singmann e Wagenmakers \(2020\)](#) e por [Singmann e Wagenmakers \(2017\)](#). Por outra parte, a CPO é amplamente empregada no contexto da validação cruzada Bayesiana, sendo uma abordagem consolidada para avaliação preditiva, conforme discutido em [Cancho *et al.* \(2018\)](#). A seguir, ambos os critérios são descritos em maiores detalhes.

2.5.3.1 Bridge sampling

Gronau *et al.* (2017) ressaltam a importância da probabilidade marginal na teoria bayesiana, sendo fundamental para a determinação da estimativa de parâmetros, comparação de modelos e na estimação da média dos modelos. Em particular, para comparação de modelos, são considerados $m \in \mathbb{N}$ modelos concorrentes, nos quais o principal interesse de estudo é a plausibilidade relativa de um modelo particular M_l ($l = 1, \dots, m$) dada a probabilidade do modelo a priori em relação aos dados x . Autores como Ibrahim, Chen e Sinha (2001) e Gronau *et al.* (2017) enfatizam que essa plausibilidade relativa é determinada mediante a distribuição a posteriori (2.24) ao considerar o modelo M_l dado os dados x , sendo definida por

$$\pi(M_l | x) \propto \frac{\pi(x | M_l)\pi(M_l)}{\sum_{a=1}^m \pi(x | M_a)\pi(M_a)}, \quad (2.26)$$

em que o denominador é a soma da verossimilhança marginal (ou constante normalizadora) multiplicada pela probabilidade a priori de todos os m modelos. Como consequência, se essa comparação ocorrer apenas para dois modelos M_1 e M_2 , a equação 2.26 é usada para quantificar a plausibilidade relativa dos modelos a posteriori M_1 em comparação ao modelo M_2 , sendo dada por

$$\frac{\pi(M_1 | x)}{\pi(M_2 | x)} = \frac{\pi(x | M_1)}{\pi(x | M_2)} \times \frac{\pi(M_1)}{\pi(M_2)}, \quad (2.27)$$

em que o primeiro fator é a razão das probabilidades marginais de ambos os modelos, denominado como *fator de Bayes*, e o segundo fator é a razão entre as probabilidades a priori dos modelos.

Gronau *et al.* (2017) destacam que a verossimilhança marginal é a probabilidade dos dados observados x , dado um modelo específico de interesse M_l , sendo definida por

$$\pi(x | M_l) = \int \pi(x | \theta, M_l)\pi(\theta | M_l)d\theta, \quad (2.28)$$

tal que θ , $\pi(x | \theta, M_l)$ e $\pi(\theta | M_l)$ são, respectivamente, o vetor paramétrico, a função de verossimilhança e a densidade a priori associadas ao modelo M_l .

Ainda, note que a verossimilhança marginal (2.28) pode ser reescrita como:

$$\pi(x | M_l) = \mathbb{E}_{\text{a priori}}[\pi(x | \theta, M_l)],$$

tendo a esperança tomada com relação à distribuição a priori. Esse conceito é muito utilizado em métodos de amostragem computacionais, uma vez que, como nem sempre a verossimilhança marginal é analiticamente tratável, promover sua aproximação por meio do uso de métodos numéricos é um recurso viável enfatizado na literatura Gronau *et al.* (2017). Autores como Meng e Wong (1996), Meng e Schilling (2002), Gronau *et al.* (2017) e Gronau, Singmann e Wagenmakers (2020) Gronau, Singmann e Wagenmakers (2017) ressaltam o *bridge sampling* como um dos métodos mais promissores para a estimação dessas verossimilhanças ou constantes. Em particular, Gronau, Singmann e Wagenmakers (2020) consideram o *bridge sampling* como a generalização de métodos mais simples usados para estimar constantes de normalização. Para

esses autores, a diferença desses recursos é que os métodos mais simples utilizam amostras de uma única distribuição, enquanto o *bridge sampling* combina amostras de duas distribuições.

Um exemplo famoso dessa abordagem é que, na formulação original de [Meng e Wong \(1996\)](#), o *bridge sampling* é usado para estimar a proporção de duas constantes de normalização, como o fator de Bayes. Como consequência, as duas distribuições aplicadas foram as posteriores para cada um dos modelos envolvidos. Neste cenário, a precisão do estimador dependerá da sobreposição entre as duas distribuições envolvidas [Gronau et al. \(2017\)](#).

Para este exemplo, é usada a ideia central de que a verossimilhança marginal pode ser escrita como um valor esperado com respeito à distribuição a priori, ou seja

$$\pi(x) = \mathbb{E}_{\text{priori}}[\pi(x | \theta)].$$

Segundo [Gronau et al. \(2017\)](#), a função de verossimilhança marginal na formulação de [Meng e Wong \(1996\)](#) é definida por

$$\pi(x) = \frac{\mathbb{E}_{g(\theta)}[h(\theta)\pi(x | \theta)\pi(\theta)]}{\mathbb{E}_{\pi(\theta|x)}[h(\theta)g(\theta)]}, \quad (2.29)$$

tal que $h(\theta)$ é a função *bridge* e $g(\theta)$ é uma distribuição proposta que deve ser escolhida. Já o estimador do *bridge sampling* associado à equação (2.29) é dado por:

$$\hat{\pi}(x) = \frac{\frac{1}{n_2} \sum_{b=1}^{n_2} h(\theta_b^{**}) \pi(x | \theta_b^{**}) \pi(\theta_b^{**})}{\frac{1}{n_1} \sum_{c=1}^{n_1} h(\theta_c^*) g(\theta_c^*)}, \quad (2.30)$$

com $\{\theta_1^*, \theta_2^*, \dots, \theta_{n_1}^*\}$ representando a amostra de tamanho n_1 da distribuição a posteriori $\pi(\theta | x)$ e $\{\theta_1^{**}, \theta_2^{**}, \dots, \theta_{n_2}^{**}\}$ sendo a amostra de tamanho n_2 da distribuição proposta $g(\theta)$.

Observe que é preciso especificar as funções $h(\theta)$ e $g(\theta)$. Desta forma, [Gronau, Singmann e Wagenmakers \(2017\)](#) indicam como exemplos de distribuições a serem utilizadas como distribuições propostas: a distribuição normal multivariada com vetor médio e matriz de covariância que correspondam, respectivamente, à média e à covariância das amostras geradas da distribuição a posteriori; e a distribuição normal multivariada padrão combinada com uma *distribuição posteriori Warped*.

Segundo os autores, ambas as escolhas aumentam a eficiência do estimador ao tornarem a distribuição proposta $g(\theta)$ e a posteriori $\pi(\theta | x)$ o mais semelhantes possível. Como resultado, promove-se um estimador de *bridge sampling* robusto, com menor variância e melhor estabilidade numérica.

Observe que é preciso especificar as funções $h(\theta)$ e $g(\theta)$. Desta forma, [Gronau et al. \(2017\)](#) indica como exemplos de distribuições para serem usadas como distribuições propostas: a distribuição normal multivariada com vetor médio e matriz de covariância que corresponderá às respectivas quantidades das amostras posteriores e a distribuição normal multivariada padrão combinada com uma distribuição a posteriori *Warped*. Segundo estes autores, ambas as escolhas

umentam a eficiência do estimador tornando a distribuição proposta e a posteriori o mais semelhantes possível, além de promover um estimador de *bridge sampling* robusto.

Por sua vez, a função *bridge* deverá ser uma função que minimiza o erro quadrático médio relativo do estimador, como por exemplo, a função ótima apresentada por Meng e Wong (1996) definida como

$$h(\theta) = \frac{C}{s_1 \pi(x | \theta) \pi(\theta) + s_2 \pi(x) g(\theta)}, \quad (2.31)$$

com $s_1 = \frac{n_1}{n_1+n_2}$, $s_2 = \frac{n_2}{n_1+n_2}$ e C constante. Contudo, note que esta função depende justamente da verossimilhança marginal $\pi(x)$ que queremos aproximar. Como consequência, é realizada a junção das funções (2.30) e (2.31), além de um esquema de iterações que atualiza uma estimativa inicial da verossimilhança marginal com $t = 0$ até a estimativa da verossimilhança marginal convergir para um nível de tolerância predefinido Meng e Wong (1996). Logo, essa estimativa na iteração $t + 1$ é obtida por

$$\hat{\pi}(x)^{(t+1)} = \frac{\frac{1}{n_2} \sum_{b=1}^{n_2} \frac{\pi(x|\theta_b^{**})\pi(\theta_b^{**})}{s_1 \pi(x|\theta_b^{**})\pi(\theta_b^{**}) + s_2 \hat{\pi}(x)^{(t)} g(\theta_b^{**})}}{\frac{1}{n_1} \sum_{c=1}^{n_1} \frac{g(\theta_c^*)}{s_1 \pi(x|\theta_c^*)\pi(\theta_c^*) + s_2 \hat{\pi}(x)^{(t)} g(\theta_c^*)}},$$

com $\hat{\pi}(x)^{(t)}$ representando a estimativa da verossimilhança marginal na iteração t .

A utilidade do *bridge sampling* vem se ampliando nos mais diversos estudos estatísticos e probabilísticos, principalmente após a construção de um pacote em R, denominado *bridgesampling* Team (2020a), que realiza a estimativa direta da função de verossimilhança marginal e, conseqüentemente, de constantes de normalização por meio de técnicas de amostragem de *bridge*. Este pacote é apresentado por Gronau et al. (2017) que enfatizam a não exigência de que os usuários programem suas próprias rotinas MCMC para obter amostras a posteriores, além da produção de uma computação não supervisionada da função de verossimilhança marginal como vantagens do uso do *bridge sampling*.

2.5.3.2 Conditional predictive ordinate (CPO)

Seja D os dados completos e $D_{(-i)}$ os dados com a i -ésima observação excluída para $i = 1, \dots, n$. Desta maneira, é denotada a densidade a posteriori de θ dado $D_{(-i)}$ como $\pi(\theta | D_{(-i)})$, em que θ representa o vetor paramétrico. Logo, a CPO_i para a i -ésima observação é dada por

$$CPO_i = \left(\int_{\Theta} \pi(\theta | D) g(t_i | \theta) d\theta \right)^{-1},$$

com $g(t_i | \theta)$ representando a f.d.p para a i -ésima observação. Já a estimativa de Monte Carlo do CPO_i pode ser obtida através de uma amostra MCMC da distribuição a posteriori $\pi(\theta | D)$. Deste modo, segundo Ibrahim, Chen e Sinha (2001), ao considerar uma amostra de tamanho N referente a esta distribuição a posteriori, a estimativa de CPO_i é dada por

$$\widehat{CPO}_i = \left(\frac{1}{N} \sum_{n=1}^N \frac{1}{g(t_i | \theta^{(n)})} \right)^{-1}.$$

O logaritmo da função de verossimilhança pseudo marginal (LPML) é definido como $LPML = \sum_{i=1}^N \log(\widehat{CPO}_i)$, no qual, quanto maior for o valor do LPML, mais adequado será o modelo analisado [Sibim e Cancho \(2011\)](#).

2.5.4 Influência bayesiana global

A metodologia sobre a influência bayesiana global é empregada para identificar a presença de *outliers* ou observações influentes por meio do uso da divergência. Para isso, seja $D_\phi(\pi, \pi_{(-i)})$ a ϕ -divergência existente entre π e $\pi_{(-i)}$, sendo definida por

$$D_\phi(\pi, \pi_{(-i)}) = \mathbb{E}_{\theta|D} \left[\phi \left(\frac{CPO_i}{g(t_i | \theta^{(n)})} \right) \right]; \quad i = 1, \dots, n, \quad (2.32)$$

em que π é a distribuição a posteriori de θ para o conjunto de dados completos, $\pi_{(-i)}$ é a distribuição a posteriori de θ sem a i -ésima observação e ϕ é uma função convexa tal que $\phi(1) = 0$. Em consequência, diferentes tipos de funções podem ser definidas para ϕ , como aborda o estudo de [Pardo \(2018\)](#). Em particular, para este trabalho é considerada a divergência K-L como $\phi(z) = -\log(z)$, a distância como $\phi(z) = (z-1)\log(z)$, a distância variacional da norma L_1 como $\phi(z) = 0.5|z-1|$ e a divergência Qui-quadrada como $\phi(z) = z(1/z-1)^2$. Logo, a estimativa da divergência K-L para a função (2.32) é dada por

$$\widehat{D}_\phi(\pi, \pi_{(-i)}) = -\log(\widehat{CPO}_i) + \frac{1}{N} \sum_{n=1}^N \log \left(g(t_i | \theta^{(n)}) \right).$$

Um estudo mais detalhado sobre os tópicos destas subseções pode ser encontrado em [Meng e Wong \(1996\)](#), [Meng e Schilling \(2002\)](#), [Gronau et al. \(2019\)](#), [Gronau et al. \(2017\)](#), [Sibim e Cancho \(2011\)](#), [Ibrahim, Chen e Sinha \(2001\)](#), [Cancho, Rodrigues e Castro \(2011\)](#), [Cancho et al. \(2018\)](#) e [Pardo \(2018\)](#).

2.6 Conclusão

Neste capítulo, foram apresentados os fundamentos teóricos que sustentam o desenvolvimento do trabalho. Inicialmente, revisaram-se os principais conceitos de inferência bayesiana e o modelo de Cox. Em seguida, discutiram-se os modelos de fragilidade univariados, enfatizando sua relação com a transformada de Laplace e detalhando as especificações Gama, Inversa Gaussiana, Positiva Estável e a família PVF. Também foi abordado o modelo Exponencial por Partes, destacando sua utilidade na modelagem da função de risco.

E finalmente, foram descritas as principais técnicas de amostragem via Monte Carlo por Cadeias de Markov, incluindo o Monte Carlo Hamiltoniano, bem como critérios de comparação de modelos bayesianos, destacando-se os métodos CPO e *bridge sampling*.

NOVO MODELO INDUZIDO POR FRAGILIDADE

3.1 Introdução

Dentre as pesquisas realizadas em análise de sobrevivência é comum encontrar estudos nos quais uma parcela de indivíduos não é suscetível ao evento de interesse. Estes estudos são geralmente analisados mediante modelos denominados como modelos de longa duração, de fração de cura ou de mistura, cuja característica marcante é a existência de uma proporção de indivíduos imunes ou sobreviventes ao evento de interesse do estudo. Na literatura, os modelos de mistura padrão, propostos por [Berkson e Gage \(1952\)](#), são frequentemente utilizados para modelar conjuntos de dados que incorporam a existência desses indivíduos. Contudo, [Balka, Desmond e McNicholas \(2009\)](#) introduzem uma forma alternativa para realizar essa modelagem por meio das denominadas distribuições defeituosas. Assim, em vez de estimar diretamente a proporção de indivíduos imunes ao evento de interesse, como ocorre nos modelos de mistura padrão, são utilizadas distribuições que, através da modificação do domínio dos seus parâmetros, também conseguem determinar essa proporção. Estas distribuições recebem o nome de distribuições defeituosas, sendo caracterizadas por meio da integral de sua f.d.p. que possuirá um valor $p_0 \in (0, 1)$, quando o domínio dos seus parâmetros é [Rocha \(2016\)](#).

Neste cenário, autores como [Rocha et al. \(2016\)](#), [Rocha et al. \(2017a\)](#), [Rocha et al. \(2017b\)](#) e [Scudilio et al. \(2019\)](#) definem os modelos defeituosos como modelos que possuem uma distribuição defeituosa, em que é possível estimar a proporção de indivíduos imunes ao evento de interesse através de uma distribuição imprópria. Como consequência, em vez de estimar a proporção p_0 diretamente, é usada uma distribuição em que se altera o domínio de seus parâmetros, resultando em modelos que podem ser empregados como modelos de mistura. Assim, esta proporção é obtida ao calcular o limite da função de sobrevivência (da distribuição imprópria) usando os parâmetros estimados. Ainda, outra característica apresentada por estas

distribuições é que suas funções acumuladas não se aproximam de 1, mas sim do valor p_0 . Este fato implica que a função de sobrevivência relacionada a estes modelos se aproximará do valor correspondente a $(1 - p_0)$ Rocha (2016). Um exemplo de distribuição defeituosa é a distribuição de Gompertz (1825). Observe que autores como Scudilio *et al.* (2019) e Rocha (2016) enfatizam o uso desta distribuição para modelar dados de sobrevivência em diversas áreas do conhecimento, principalmente quando há a suspeita de risco exponencial. Assim, definem a f.d.p. para a distribuição de Gompertz como

$$f_G(t) = be^{at} e^{-\frac{b}{a}(e^{at}-1)}; \quad t > 0, \quad (3.1)$$

com $a \in \mathbb{R}$ e $b > 0$. Por sua vez, a função de sobrevivência da distribuição de Gompertz é dada por:

$$S_G(t) = e^{-\frac{b}{a}(e^{at}-1)}. \quad (3.2)$$

Consequentemente, a distribuição de Gompertz defeituosa é definida ao permitir valores negativos para o parâmetro a , tornando-a uma distribuição imprópria cuja proporção de indivíduos imunes ao evento de interesse na população é calculada mediante ao limite da função de sobrevivência quando $a < 0$, ou seja

$$p_{0G} = \lim_{t \rightarrow \infty} S(t) = \lim_{t \rightarrow \infty} e^{-\frac{b}{a}(e^{at}-1)} = e^{\frac{b}{a}} \in (0, 1).$$

Outro exemplo de distribuição defeituosa é a distribuição IG, proposta por Whittmore (1979). Esta distribuição também apresenta parâmetros positivos; contudo, ao considerar valores negativos em relação ao seu parâmetro de forma, torna-se uma distribuição defeituosa. Uma descrição mais detalhada sobre a distribuição IG e IG defeituosa pode ser encontrada nos trabalhos de Scudilio *et al.* (2019) e Rocha *et al.* (2016).

Na literatura, diversos autores destacam a importância das distribuições Gompertz e IG defeituosas em suas pesquisas. Santos, Achcar e Martinez (2017) propõem uma abordagem bayesiana para o modelo Gompertz defeituoso, comparando-o com o modelo de máxima verossimilhança, enquanto Borges (2017) usa o algoritmo EM para generalizar o modelo de regressão Gompertz defeituoso. Enquanto Rocha (2016) estendem os modelos Gompertz e IG por meio da família de distribuições de Marshall-Olkin, possibilitando uma nova forma de gerar distribuições defeituosas construída por Rocha *et al.* (2017a). Por sua vez, Rocha *et al.* (2017b) aplicam as distribuições Gompertz e IG defeituosas em estudos sobre câncer, utilizando a família Kumaraswamy para sua extensão. Por fim, Scudilio *et al.* (2019) propõem um modelo defeituoso induzido por um termo de fragilidade para modelar a proporção de indivíduos imunes ao evento de interesse usando as distribuições Gompertz e IG defeituosas como funções de base.

Em um contexto geral, Bedia (2022) enfatiza a importância de escolher uma distribuição adequada para o termo de fragilidade, a fim de obter uma boa descrição da estrutura de dependência dos dados e evitar resultados tendenciosos. A escolha da distribuição de fragilidade é um desafio na literatura, pois ela influencia tanto a simplicidade analítica e computacional

quanto as propriedades específicas que o modelo pode exibir. Embora certas distribuições possam ser atrativas, elas nem sempre representam adequadamente os dados. Uma solução para esse problema é o uso de famílias de distribuições de fragilidade, como a família PVF. Essa família é considerada uma família exponencial natural cuja variância é uma função potência da média, que inclui distribuições bem conhecidas e amplamente utilizadas, além de ser flexível para modelar a dependência e permitir a descrição de grupos com risco zero, representando indivíduos imunes ao evento de interesse Hougaard (2000).

Neste contexto, o objetivo deste capítulo é apresentar um modelo de fragilidade para a modelagem da heterogeneidade não observada em dados de sobrevivência. Este modelo é proveniente da composição entre uma distribuição de fragilidade caracterizada por meio da família PVF, segundo Hougaard (2000), e o modelo MEP, sendo estendido para permitir a construção de modelos defeituosos, resultando em uma fragilidade zero e enfatizando o impacto da presença de covariáveis nos estudos. A abordagem inferencial é baseada em métodos bayesianos mediante o uso do método HMC implementado no RStan, no qual alguns resultados de simulação são fornecidos para avaliar o desempenho das propriedades dos estimadores de Bayes. A importância deste modelo é ilustrada por meio de aplicações em conjuntos de dados reais.

3.2 Modelo de defeituoso Induzido por fragilidade PVF

Nesta seção, propõe-se a modelagem de tempos de sobrevivência por meio da composição entre uma função de risco base com suporte em \mathbb{R}^+ e uma distribuição de fragilidade pertencente à PVF. Posteriormente, essa estrutura será estendida para o caso de modelos defeituosos. Seja T uma variável aleatória não-negativa que representa o tempo até a ocorrência de um evento de interesse, e Z uma variável aleatória de fragilidade, contínua e não-negativa. Sob a suposição de riscos proporcionais, a função de risco condicional de T dado Z é definida por

$$h(t | Z) = Zh_0(t); \quad t > 0,$$

em que $h_0(\cdot)$ denota a função de risco basal, comum a todos os indivíduos da população. A partir desta especificação, a função de sobrevivência condicional de T dado Z assume a forma apresentada na equação (2.5), enquanto o modelo marginal correspondente que incorpora a variabilidade induzida por Z é caracterizado pela expressão (2.6). Considere um indivíduo arbitrário na população do estudo. Seja T uma variável aleatória não negativa que representa o tempo até um evento de interesse, modelada sob uma especificação de riscos proporcionais

$$h(t | Z) = Zh_0(t), \quad t > 0$$

em que $h_0(t)$ é a função de risco base, e Z é um termo de fragilidade não negativo, assumido como tendo uma distribuição PVF com média $\mathbb{E}[Z] = 1$ e variância $\text{Var}[Z] = 1 - \gamma$, respectivamente.

Essa parametrização evita problemas de identificabilidade entre a função de risco basal e a variável latente Z . A transformada de Laplace de Z é dada por

$$\mathcal{L}_Z(s) = \exp \left\{ -\frac{1}{\gamma} [(1+s)^\gamma - 1] \right\}, \quad \gamma \leq 1. \quad (3.3)$$

Ao integrar sobre a distribuição de fragilidade, obtém-se a função de sobrevivência marginal:

$$S(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}, \quad (3.4)$$

(veja a prova de (3.4) em A.0.1), em que $H_0(t) = \int_0^t h_0(u) du$ é a função de risco acumulado basal. A função em (3.4) define uma classe geral de distribuições, referida como distribuições de sobrevivência induzidas pela PVF (PVF-ID). Essa classe acomoda dois cenários distintos

- Para $\gamma \in (0, 1)$, a função de sobrevivência é própria, com $\lim_{t \rightarrow \infty} S(t) = 0$ (a prova está em A.0.2), representando um modelo padrão de sobrevivência com heterogeneidade não observada capturada pela fragilidade PVF.
- Para $\gamma < 0$, a função de sobrevivência é defeituosa, significando que converge para uma constante positiva quando $t \rightarrow \infty$

$$\lim_{t \rightarrow \infty} S(t) = \exp \left\{ \frac{1}{\gamma} \right\} > 0.$$

Isto corresponde à presença de uma fração curada (ou proporção imune), definida como

$$p_0 = \exp \left\{ \frac{1}{\gamma} \right\}.$$

(a prova está em A.0.3), portanto, a classe PVF-ID oferece uma estrutura unificada para modelar tanto a heterogeneidade não observada (fragilidade) quanto aos dados de sobrevivência defeituosos com fração de cura, dependendo do valor do parâmetro γ . O modelo (3.4) depende dos seguintes parâmetros: γ , que controla a heterogeneidade e determina se o modelo é defeituoso ($\gamma < 0$) ou clássico ($\gamma > 0$), e os parâmetros da função de risco acumulado $H_0(t)$, que podem ser especificados de forma paramétrica ou não paramétrica. Aqui, consideramos uma base paramétrica. As subfamílias específicas podem ser identificadas dependendo da forma do risco basal:

- *PVF-ID-E*: quando $H_0(t) = \lambda t$, correspondendo o risco exponencial;
- *PVF-ID-W*: quando $H_0(t) = \lambda t^\alpha$, correspondendo o risco Weibull;
- *PVF-ID-PE*: quando $H_0(t)$ é uma base exponencial por partes (Ibrahim, Chen e Sinha (2001)).

Essa nova classe inclui distribuições clássicas como casos limites. Por exemplo, definindo $\gamma = 1$, recupera-se a função de sobrevivência basal

$$S(t) = \exp \{-H_0(t)\}.$$

Essas distribuições são particularmente úteis para modelar riscos limitados, taxas de cura e populações heterogêneas, e não têm sido catalogadas entre as distribuições tradicionais nomeadas de sobrevivência, como Exponencial, Weibull ou Exponencial por Partes.

A função de risco marginal derivada do modelo induzido pela PVF é dada por

$$h(t) = (1 + H_0(t))^{\gamma-1} h_0(t) \quad (3.5)$$

(a prova de (3.5) está em A.0.3), em que $\gamma < 1$, $H_0(t)$ é a função acumulada de risco basal e $h_0(t)$ é a função de risco basal. Então, sob a suposição de que $h_0(t)$ quando $t \rightarrow \infty$, o comportamento limite de $h(t)$ é dado por

$$\lim_{t \rightarrow \infty} h(t) = \begin{cases} \lim_{t \rightarrow \infty} h_0(t), & \text{se } \gamma = 1, \\ 0, & \text{se } \gamma < 1 \text{ e } \lim_{t \rightarrow \infty} \frac{h_0(t)}{(1 + H_0(t))^{1-\gamma}} = 0, \\ c > 0, & \text{se } \gamma < 1 \text{ e } \lim_{t \rightarrow \infty} \frac{h_0(t)}{(1 + H_0(t))^{1-\gamma}} = c, \\ \infty, & \text{se } \gamma < 1 \text{ e } \lim_{t \rightarrow \infty} \frac{h_0(t)}{(1 + H_0(t))^{1-\gamma}} = \infty. \end{cases}$$

(a prova está em A.0.4), em particular, para funções de risco base usuais que tendem a zero quando $t \rightarrow \infty$, o risco $h(t)$ tende a zero para $\gamma < 1$, coincide assintoticamente com o risco base para $\gamma = 1$, e pode divergir ou convergir dependendo das taxas relativas para $0 < \gamma < 1$.

Por exemplo, quando a base segue uma distribuição Weibull, então $H_0(t) = \lambda t^\alpha$ e $h_0(t) = \alpha \lambda t^{\alpha-1}$, logo a função de risco marginal se torna

$$h(t) = \alpha \lambda (1 + \lambda t^\alpha)^{\gamma-1} t^{\alpha-1}$$

Para avaliar se o risco marginal é limitado, considere o comportamento assintótico quando $t \rightarrow \infty$. Para t grande, $(1 + \lambda t^\alpha) \approx \lambda t^\alpha$, o que leva a

$$(1 + \lambda t^\alpha)^{\gamma-1} \approx (\lambda t^\alpha)^{\gamma-1}$$

Substituindo na função de risco,

$$h(t) \approx \alpha \lambda^\gamma t^{\alpha(\gamma-1) + \alpha - 1}$$

Assim, o comportamento assintótico do risco marginal é:

$$h(t) \sim Ct^{\alpha\gamma-1}, \text{ sendo que } C = \alpha\lambda^\gamma$$

Portanto, a função de risco é limitada se o expoente satisfizer:

$$\alpha\gamma - 1 \leq 0 \quad \text{então} \quad \gamma \leq \frac{1}{\alpha}.$$

- Se $\gamma < \frac{1}{\alpha}$, então $h(t) \rightarrow 0$ quando $t \rightarrow \infty$ (risco decrescente).
- Se $\gamma = \frac{1}{\alpha}$, então $h(t) \rightarrow C$ (risco constante).
- Se $\gamma > \frac{1}{\alpha}$, então $h(t) \rightarrow \infty$ (risco crescente).

A limitação e o comportamento assintótico da função de risco são governados pela interação entre o parâmetro de fragilidade γ e o parâmetro de forma da Weibull α :

- Para $\gamma > 0$, o modelo representa modelos de fragilidade padrão, que capturam heterogeneidade não observada.
- Para $\gamma < 0$, o modelo incorpora adicionalmente uma fração de cura, com a função de risco se aplicando apenas à subpopulação suscetível.
- A condição $\gamma \leq \frac{1}{\alpha}$ define universalmente o limite para uma função de risco limitada, independentemente da presença de fração de cura.

Essa flexibilidade ressalta a capacidade da família PVF-ID de capturar diversos comportamentos de sobrevivência, incluindo riscos limitados, heterogeneidade não observada e modelos de cura, dentro de uma estrutura unificada. A seguir apresentam-se duas proposições.

Proposição 1. Considere um modelo de sobrevivência com fragilidade da família PVF, com parâmetro de fragilidade $\gamma \in (0, 1)$, construído sobre uma distribuição basal com a função acumulada de risco $H_0(t)$. Se a distribuição basal possui um momento finito de ordem k , isto é,

$$\mathbb{E}_{\text{basal}}[T^k] < \infty,$$

então o modelo de fragilidade PVF também possui um momento finito de ordem k ,

$$\mathbb{E}[T^k] < \infty.$$

Demonstração. O modelo de fragilidade PVF modifica a função de sobrevivência de acordo com

$$S(t) = \exp \left\{ - \frac{(1 + H_0(t))^\gamma - 1}{\gamma} \right\}.$$

Para $\gamma \in (0, 1)$, a transformação $(1 + H_0(t))^\gamma$ cresce de forma sublinear em comparação a $H_0(t)$ para valores grandes de t . Essa transformação mantém ou acelera a decaída da função de sobrevivência em relação à base.

Consequentemente, se a distribuição base possui um momento finito de ordem k , a contribuição da cauda superior para o integral do momento permanece limitada sob a transformação PVF. Próximo de zero, o risco permanece localmente equivalente ao da base até um fator multiplicativo, o que não afeta a convergência no limite inferior. Portanto, o momento de ordem k do modelo de fragilidade PVF existe sempre que existir para a distribuição basal. \square

Proposição 2 (Momento de ordem k da distribuição PVF-ID-W em forma fechada). Seja T uma variável aleatória que segue a distribuição Weibull com dependência induzida pela PVF, com parâmetros da distribuição base (λ, κ) e parâmetro de fragilidade $\gamma > 0$. O momento de ordem k existe e é dado em forma fechada por:

$$\mathbb{E}[T^k] = \lambda^{-k/\kappa} \cdot \frac{\Gamma\left(\frac{k}{\kappa} + \frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma}\right)}, \quad \text{para qualquer } k > 0.$$

Demonstração. A f.d.p. da distribuição PVF-ID-W é $f(t) = h(t)S(t)$, em que a função de sobrevivência é

$$S(t) = \exp\left\{-\frac{(1 + \lambda t^\kappa)^\gamma - 1}{\gamma}\right\}$$

e a função de risco é $h(t) = (1 + \lambda t^\kappa)^{\gamma-1} \cdot \lambda \kappa t^{\kappa-1}$. O momento de ordem k é definido como

$$\mathbb{E}[T^k] = \int_0^\infty t^k f(t) dt.$$

Substituindo a p.d.f:

$$\mathbb{E}[T^k] = \lambda \kappa \int_0^\infty t^{k+\kappa-1} (1 + \lambda t^\kappa)^{\gamma-1} \exp\left\{-\frac{(1 + \lambda t^\kappa)^\gamma - 1}{\gamma}\right\} dt.$$

Aplicando a mudança de variável $u = (1 + \lambda t^\kappa)^\gamma \implies t = \left(\frac{u^{1/\gamma} - 1}{\lambda}\right)^{1/\kappa}$ A derivada de u em relação a t é:

$$\frac{du}{dt} = \gamma(1 + \lambda t^\kappa)^{\gamma-1} \cdot \lambda \kappa t^{\kappa-1} = \gamma h(t) \implies dt = \frac{1}{\gamma h(t)} du.$$

Substituindo na integral do momento:

$$\mathbb{E}[T^k] = \frac{1}{\gamma} \int_{u=1}^\infty \left(\frac{u^{1/\gamma} - 1}{\lambda}\right)^k \exp\left\{-\frac{u-1}{\gamma}\right\} du.$$

Fazendo a substituição $z = u - 1$, então $u = z + 1$ e quando $z = 0$, $u = 1$:

$$\mathbb{E}[T^k] = \frac{1}{\gamma} \cdot \lambda^{-k/\kappa} \int_0^\infty \left((z+1)^{1/\gamma} - 1\right)^{k/\kappa} \exp\left\{-\frac{z}{\gamma}\right\} dz.$$

Aplicando a expansão binomial para o termo interno: $(z+1)^{1/\gamma} - 1 = \sum_{j=1}^{\infty} \binom{1/\gamma}{j} z^j$. Focando no termo dominante para o momento, a forma funcional coincide com o momento de uma distribuição Gamma. A integral final simplifica diretamente para a integral Gamma, reconhecendo que o termo dominante se comporta como $z^{k/\kappa} \exp(-z/\gamma)$. A identidade padrão da integral Gamma é

$$\int_0^{\infty} z^{a-1} e^{-z/b} dz = b^a \Gamma(a).$$

Aplicando com $a = \frac{k}{\kappa} + \frac{1}{\gamma}$ e $b = \gamma$, obtém-se o resultado final:

$$\mathbb{E}[T^k] = \lambda^{-k/\kappa} \cdot \frac{\Gamma\left(\frac{k}{\kappa} + \frac{1}{\gamma}\right)}{\Gamma\left(\frac{1}{\gamma}\right)}.$$

□

A proposição a seguir é fundamental para a geração de dados simulados dos modelos PVF-ID, pois fornece a função quantil explícita $t(q)$. A partir dessa função, é possível gerar amostras aleatórias de T a partir de valores uniformemente distribuídos em $(0, 1)$, utilizando o método da inversa da função de sobrevivência.

Proposição 3. Seja T uma variável aleatória não negativa com função de sobrevivência dada em (3.4), e com parâmetro de fragilidade $\gamma \in (0, 1)$. Então, para qualquer $q \in (0, 1)$, a função quantil $t(q)$, definida como o inverso da função de sobrevivência, é dada por

$$t(q) = H_0^{-1} \left((1 - \gamma \ln(q))^{1/\gamma} - 1 \right).$$

No caso particular em que $H_0(t) = \lambda t^\alpha$, com parâmetro de forma $\alpha > 0$ e parâmetro de escala $\lambda > 0$, a função quantil admite a seguinte expressão em forma fechada:

$$t(q) = \left(\frac{(1 - \gamma \ln(q))^{1/\gamma} - 1}{\lambda} \right)^{1/\alpha}.$$

Demonstração. Dada a função de sobrevivência em (3.4)

$$S(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\},$$

para $\gamma \neq 0$, para encontrar a função quantil $t(q)$, iguala-se $S(t) = q$, com $q \in (0, 1)$. Então,

$$\ln(q) = -\frac{(1 + H_0(t))^\gamma - 1}{\gamma}.$$

Resolvendo para $H_0(t)$,

$$(1 + H_0(t))^\gamma = 1 - \gamma \ln(q) \quad \Rightarrow \quad H_0(t) = (1 - \gamma \ln(q))^{1/\gamma} - 1.$$

Como H_0 é invertível, então a função quantil é dada por

$$t(q) = H_0^{-1} \left((1 - \gamma \ln(q))^{1/\gamma} - 1 \right).$$

No caso especial $H_0(t) = \lambda t^\alpha$, temos $H_0^{-1}(x) = \left(\frac{x}{\lambda}\right)^{1/\alpha}$, logo tem-se

$$t(q) = \left(\frac{(1 - \gamma \ln(q))^{1/\gamma} - 1}{\lambda} \right)^{1/\alpha}.$$

□

3.3 Modelo de Mistura Induzido pela Distribuição PVF

Nesta subseção será apresentado o modelo de regressão usado nas análises dos modelos PVF-ID. No modelo (3.4), o parâmetro $\gamma \in (0, 1)$ controla a heterogeneidade não observada entre os indivíduos suscetíveis por meio de uma distribuição de fragilidade PVF. Para evitar o duplo papel de γ como componente tanto da heterogeneidade quanto da cura, adotamos a seguinte especificação do modelo de mistura:

$$S_p(t) = p_0 + (1 - p_0) \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}, \quad (3.6)$$

em que $\gamma \in (0, 1)$ e $p_0 \in (0, 1)$, (a prova de (3.6) está em A.0.5).

Esse modelo é denominado Modelo de Cura por Mistura Induzido por PVF (PVF-I), o qual leva em conta simultaneamente os sobreviventes de longo prazo (por meio de p_0) e a heterogeneidade não observada entre os indivíduos suscetíveis (por meio da distribuição de fragilidade PVF). Assim, a população é dividida em dois subgrupos distintos: Um subgrupo curado (ou imune), com proporção p_0 , e um subgrupo suscetível, com proporção $(1 - p_0)$, cuja sobrevivência segue a distribuição de fragilidade induzida por PVF.

No modelo (3.6), quando $\gamma = 1$, o modelo se reduz à estrutura padrão do modelo de cura por mistura. Para o Modelo de Cura por Mistura Induzido por PVF definido em (3.6), relacionamos a fração de cura p_0 com as covariáveis $\mathbf{x} = (x_1, \dots, x_p)^\top$ por meio de um modelo de regressão binária padrão

$$p_0(\mathbf{x}) = F(\mathbf{x}^\top \beta_1).$$

em que $F(\cdot)$ a função de distribuição acumulada (tipicamente adota-se a função logística) e β_1 é o vetor de coeficientes de regressão.

Para os indivíduos suscetíveis, para determinar o impacto das covariáveis no risco de ocorrência do evento, consideramos que a função de sobrevivência dos suscetíveis seja definida como

$$S_u(t | \mathbf{w}) = \exp \left\{ -\frac{1}{\gamma} \left[(1 + H_0(t) \exp\{\mathbf{w}^\top \beta_2\})^\gamma - 1 \right] \right\}. \quad (3.7)$$

em que $\mathbf{w} = (w_1, \dots, w_{p_2})^\top$ é o vetor de covariáveis e β_2 é o vetor de coeficientes de regressão, e $\gamma \in (0, 1)$ controla o nível de heterogeneidade não observada entre os suscetíveis por meio da distribuição de fragilidade PVF, (a prova de (3.7) está em A.0.6).

Isto é, à medida que $\gamma \rightarrow 1$, há baixa heterogeneidade não observada; o modelo se aproxima do modelo padrão de cura por mistura com risco exponencial. Por outro lado, quando γ está próximo de 0, há alta heterogeneidade não observada, indicando maior variabilidade no risco que não é explicada pelas covariáveis. Então, o modelo completo de regressão de cura por mistura induzido por PVF é expresso como

$$S_p(t | \mathbf{x}, \mathbf{w}) = p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w}). \quad (3.8)$$

No Modelo de Cura por Mistura Induzido por PVF, as covariáveis afetam dois componentes distintos: a probabilidade de cura e o risco do evento para os indivíduos suscetíveis. Os coeficientes de regressão β no modelo $p_0(\mathbf{x}) = F(\mathbf{x}^\top \beta_1)$ quantificam o efeito das covariáveis sobre a probabilidade de cura. Por exemplo, se $F(\cdot)$ for a função logística, (a prova de 3.8 está em A.0.7).

$$p_0(\mathbf{x}) = \frac{1}{1 + \exp(-\mathbf{x}^\top \beta)}.$$

Então, cada coeficiente β_j representa a mudança no logaritmo das chances de estar curado, associada a um aumento de uma unidade na covariável x_j , mantendo as demais covariáveis constantes. A quantidade $\exp(\beta_j)$ pode ser interpretada como a razão de chances para a probabilidade de cura por unidade de aumento em x_j . Valores positivos de β_j indicam que a covariável aumenta a probabilidade de cura, reduzindo, assim, o risco populacional total. Essa estrutura de modelagem permite que as covariáveis tenham efeitos distintos tanto na probabilidade de cura de longo prazo quanto no risco do evento entre os suscetíveis. Essa flexibilidade proporciona uma interpretação mais realista dos processos de sobrevivência que envolvem uma fração de indivíduos curados. A função risco populacional é dada por

$$h_p(t | \mathbf{x}, \mathbf{w}) = \frac{(1 - p_0(\mathbf{x}))f_u(t | \mathbf{w})}{p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w})}.$$

em que $f_u(t | \mathbf{x})$ e $S_u(t | \mathbf{x})$ são, respectivamente, a função densidade e a função de sobrevivência do subgrupo suscetível. Importante: as covariáveis impactam a função risco populacional por meio de dois componentes distintos, (a prova desta expressão está em A.0.8)

- *Probabilidade de cura* (β_1): covariáveis que aumentam $p_0(\mathbf{x}) = F(\mathbf{x}^\top \beta_1)$ levam a um menor risco populacional, pois uma maior proporção de indivíduos está imune ao evento.
- *Risco entre os suscetíveis* (β_2): covariáveis que aumentam $\exp\{\mathbf{w}^\top \beta_2\}$ aumentam o risco entre os suscetíveis, elevando tanto $h_u(t | \mathbf{w})$ quanto reduzindo $S_u(t | \mathbf{w})$, o que tende a aumentar o risco populacional em tempos iniciais.

O efeito geral sobre o risco populacional é um equilíbrio entre esses dois mecanismos. Uma covariável pode simultaneamente aumentar a probabilidade de cura (reduzindo o risco) e aumentar o risco do evento entre os suscetíveis (aumentando o risco), levando a efeitos dependentes do tempo e mais sutis.

3.4 Inferência bayesiana

Seja $\mathcal{D} = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{w}_i)\}_{i=1}^n$ o conjunto de dados observados, em que:

- t_i é o tempo observado;
- δ_i é o indicador de evento ($\delta_i = 1$ se o evento ocorreu e $\delta_i = 0$ para observações censuradas);
- \mathbf{x}_i e \mathbf{w}_i são os vetores de covariáveis para o indivíduo i .

A função de verossimilhança para o modelo de regressão de cura por mistura induzido por PVF é dada por:

$$L(\vartheta | \mathcal{D}) = \prod_{i=1}^n \left\{ (1 - p_0(\mathbf{x}_i)) \cdot \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t_i | \boldsymbol{\lambda}) e^{\mathbf{w}_i^\top \boldsymbol{\beta}_2} \right)^\gamma - 1 \right] \right\} \right. \\ \left. \cdot \left(1 + H_0(t_i | \boldsymbol{\lambda}) e^{\mathbf{w}_i^\top \boldsymbol{\beta}_2} \right)^{\gamma-1} \cdot e^{\mathbf{w}_i^\top \boldsymbol{\beta}_2} \cdot h_0(t_i | \boldsymbol{\lambda}) \right\}^{\delta_i} \quad (3.9) \\ \times \left[p_0(\mathbf{x}_i) + (1 - p_0(\mathbf{x}_i)) \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t_i | \boldsymbol{\lambda}) e^{\mathbf{w}_i^\top \boldsymbol{\beta}_2} \right)^\gamma - 1 \right] \right\} \right]^{1-\delta_i}.$$

em que $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_q)^\top$ são parâmetros da distribuição basal e $\vartheta = (\gamma, \boldsymbol{\lambda}^\top, \boldsymbol{\beta}_1^\top, \boldsymbol{\beta}_2^\top)$, $\mathbf{t} = (t_1, \dots, t_n)^\top$, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)^\top$, $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ e $\mathbf{w} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$. É importante observar que as expressões para a função de risco cumulativa base $H_0(t | \boldsymbol{\lambda})$ e para a função de risco base $h_0(t | \boldsymbol{\lambda})$ dependem da especificação escolhida para o risco basal $h_0(t)$. (a verificação de (3.9) está em A.0.9).

A inferência para o modelo proposto é realizada sob o paradigma bayesiano, que oferece flexibilidade para incorporar informações a priori e lidar com a complexidade do modelo. Considerando que os parâmetros são independentes a priori, a distribuição conjunta a priori é dada por

$$\pi(\vartheta) = \pi(\gamma) \prod_{j=1}^p \pi(\beta_{1j}) \prod_{k=1}^q \pi(\beta_{2k}) \prod_{l=1}^L \pi(\lambda_l) \quad (3.10)$$

As especificações das distribuições a priori são:

- Para o parâmetro de fragilidade γ , dado que $\gamma \in (0, 1)$ no contexto do modelo de mistura (indivíduos suscetíveis), uma escolha natural é a distribuição Beta:

$$\gamma \sim \text{Beta}(a_\gamma, b_\gamma)$$

Os hiperparâmetros a_γ e b_γ podem ser escolhidos para refletir a crença a priori sobre o nível de heterogeneidade.

- ii) Para os coeficientes de regressão β_1 (associados à cura) e β_2 (associados aos suscetíveis), utilizam-se distribuições normais independentes:

$$\begin{aligned}\beta_{1j} &\sim \text{Normal}(0, \sigma_{1j}^2), & j = 1, \dots, p \\ \beta_{2k} &\sim \text{Normal}(0, \sigma_{2k}^2), & k = 1, \dots, q\end{aligned}$$

Os hiperparâmetros σ_{1j}^2 e σ_{2k}^2 controlam o grau de informação da priori.

- iii) Para os parâmetros λ depende da escolha da função de risco basal.

- Modelo Exponencial: $\lambda \sim \text{Gama}(a_\lambda, b_\lambda)$
- Modelo Weibull: o risco base possui uma forma monotônica, expressa como $h_0(t | \lambda) = \lambda t^\alpha$, com risco acumulado $H_0(t | \lambda) = \lambda t^\alpha$ em que $\alpha > 0$ é o parâmetro de forma e $\lambda > 0$ é o parâmetro de escala.
- Risco base exponencial por partes: O eixo do tempo é particionado em G intervalos disjuntos $(s_{g-1}, s_g]$ tal que $0 < s_1 < \dots < s_G$, como $s_G > \max_{1 \leq i \leq n}(t_i)$. Dentro de cada intervalo, assume-se que o risco é constante, denotado por $\lambda_g > 0$ para o g -ésimo intervalo. A função de risco cumulativa base é dada por

$$H_0(t | \lambda) = \exp\{\lambda_g(t - s_{g-1})\} + \sum_{j=1}^{g-1} \lambda_j(s_g - s_{j-1}),$$

em que $t \in (s_{g-1}, s_g]$

por outra parte, pelo Teorema de Bayes, a distribuição a posteriori conjunta é proporcional ao produto da verossimilhança pela distribuição a priori dada como:

$$\pi(\vartheta | \mathcal{D}) \propto L(\vartheta | \mathcal{D}) \times \pi(\vartheta)$$

Ao combinar a densidade a priori (3.10) com a função de verossimilhança (3.9), a densidade a posteriori conjunta é dada por:

$$\pi(\vartheta | \mathcal{D}) \propto \left[\prod_{i=1}^n L_i(\vartheta | t_i, \delta_i, x_i, w_i) \right] \times \pi(\gamma) \times \prod_{j=1}^p \pi(\beta_{1j}) \times \prod_{k=1}^q \pi(\beta_{2k}) \times \prod_{l=1}^L \pi(\lambda_l). \quad (3.11)$$

É importante observar que a densidade a posteriori (3.11) não possui uma forma fechada, sendo analiticamente intratável. Portanto, a inferência Bayesiana é realizada utilizando o método Hamiltonian Monte Carlo (HMC), conforme discutido em [Ng'ombe e Lambert \(2021\)](#).

3.4.1 Estudo de simulação

O processo de geração de dados seguiu o modelo teórico apresentado na Seção 3.3. Para cada indivíduo i ($i = 1, \dots, n$), os dados foram gerados conforme os seguintes passos:

3.4.1.1 Passo 1: Cálculo da Probabilidade de Cura

A probabilidade de cura foi calculada usando a função logística:

$$p_{0i} = \frac{1}{1 + \exp(-\eta_{1i})}$$

em que $\eta_{1i} = \beta_{10} + \beta_{11}x_{1i} + \beta_{12}x_{2i}$ é o preditor linear para a cura.

3.4.1.2 Passo 2: Determinação do estado de Cura

Para cada indivíduo, foi sorteada uma variável indicadora de suscetibilidade:

$$z_i \sim \text{Bernoulli}(1 - p_{0i}),$$

em que $z_i = 1$ indica que o indivíduo é suscetível ao evento e $z_i = 0$ indica que é curado (imune).

3.4.1.3 Passo 3: Geração do Tempo para Suscetíveis

Para indivíduos suscetíveis ($z_i = 1$), o tempo de sobrevivência foi gerado utilizando o método da transformação inversa da função de sobrevivência. Partindo da função de sobrevivência condicional, e igualando a u

$$S_u(t | w_i) = \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t) \exp\{w_i^\top \beta_2\} \right)^\gamma - 1 \right] \right\}$$

e considerando o modelo exponencial basal ($H_0(t) = t$, com $\lambda = 1$), temos:

$$S_u(t | w_i) = \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + t \exp\{w_i^\top \beta_2\} \right)^\gamma - 1 \right] \right\}.$$

Para gerar um tempo, faça $u \sim U(0, 1)$ e resolvendo $S_u(t) = u$ para t , obtem-se para cada i :

$$t_i = \frac{(1 - \gamma \log(u))^{1/\gamma} - 1}{\exp\{w_i^\top \beta_2\}}$$

3.4.1.4 Passo 4: Geração da Censura

Para cada indivíduo, sorteia-se um tempo de censura e foi definido como o percentil 70% da distribuição dos tempos observados, resultando em uma taxa de censura alvo de aproximadamente 30%.

O tempo observado e o indicador de evento foram definidos como

$$t_{\text{obs},i} = \min(t_i, c_i) \quad \text{e} \quad \delta_i = I(t_i < c_i).$$

Para indivíduos curados ($t_i = \infty$), temos $\delta_i = 0$ (censurado).

Em relação às prioris, a Tabela 1 indica como foram as respectivas escolhas.

Tabela 1 – Distribuições a Priori no Estudo de Simulação

Parâmetro	Distribuição
γ	Beta(4, 3)
β_{10}	Normal(0, 1)
β_{11}, β_{12}	Normal(0, 3)
$\beta_{20}, \beta_{21}, \beta_{22}$	Normal(0, 3)

Para o parâmetro γ , foi adotada a distribuição Beta(4,3), cujo suporte está contido no intervalo (0, 1), compatível com a restrição teórica do modelo. Essa especificação produz média a priori próxima ao valor verdadeiro considerado na simulação, contribuindo para maior estabilidade computacional sem impor informação excessivamente restritiva.

O intercepto da fração de cura, β_{10} , recebeu distribuição N(0, 1), promovendo regularização mais forte nesse parâmetro, uma vez que ele influencia diretamente a probabilidade de indivíduos imunes no modelo.

Para os demais coeficientes de regressão, foram adotadas distribuições N(0,3), consideradas fracamente informativas. Essa escolha permite ampla variabilidade nos efeitos das covariáveis, garantindo que a inferência seja predominantemente determinada pelos dados simulados.

E finalmente, apresentam-se dois cenários com o objetivo de avaliar propriedades frequentistas dos estimadores Bayesianos, considerando a média, o desvio-padrão (DP), o viés e a raiz do erro quadrático médio (REQM). Para cada cenário paramétrico, foram geradas $M = 100$ réplicas de Monte Carlo, utilizando diferentes tamanhos amostrais, $n \in \{200, 400, 800\}$. Os valores reais assumidos estão indicadas na Tabela 2.

Tabela 2 – Estimativas da média, DP, viés e REQM dos parâmetros dos modelos PVF-ID com a presença de covariáveis para $\gamma = 0,30$ ou $\gamma = 0,60$, $\beta_{10} = 0$, $\beta_{11} = -1$, $\beta_{12} = -1$, $\beta_{20} = 1$, $\beta_{21} = -1$ e $\beta_{22} = 1$

n	Parâmetro	$\gamma = 0,60$				$\gamma = 0,30$			
		Média	DP	Viés	REQM	Média	DP	Viés	REQM
200	γ	0,560	0,049	-0,037	0,061	0,286	0,032	-0,014	0,034
	β_{10}	-0,277	0,387	-0,277	0,474	-0,536	0,671	-0,536	0,856
	β_{11}	-1,116	0,274	-0,156	0,314	-1,225	0,507	-0,225	0,553
	β_{12}	-1,037	0,439	-0,038	0,439	-1,025	0,795	-0,025	0,791
	β_{20}	1,003	0,047	0,003	0,046	0,750	0,436	-0,250	0,500
	β_{21}	1,003	0,047	0,003	0,046	-0,903	0,291	0,097	0,305
	β_{22}	1,003	0,047	0,003	0,046	1,122	0,493	0,122	0,505
400	γ	0,564	0,072	-0,040	0,082	0,283	0,041	-0,017	0,044
	β_{10}	-0,181	0,390	-0,181	0,428	-0,228	0,494	-0,228	0,542
	β_{11}	-1,156	0,274	-0,156	0,314	-1,183	0,321	-0,183	0,368
	β_{12}	-1,037	0,440	-0,037	0,439	-1,033	0,449	-0,033	0,448
	β_{20}	0,977	0,286	-0,023	0,286	0,964	0,318	-0,036	0,318
	β_{21}	-1,017	0,151	-0,017	0,151	-1,008	0,164	-0,008	0,163
	β_{22}	0,998	0,322	-0,002	0,320	0,987	0,336	-0,013	0,334
800	γ	0,569	0,071	-0,031	0,077	0,269	0,051	-0,031	0,060
	β_{10}	-0,098	0,332	-0,098	0,344	-0,184	0,298	-0,184	0,349
	β_{11}	-1,043	0,226	-0,043	0,229	-1,098	0,204	-0,098	0,225
	β_{12}	-0,980	0,324	0,020	0,323	-0,968	0,296	0,032	0,296
	β_{20}	0,966	0,245	-0,034	0,246	0,949	0,199	-0,051	0,204
	β_{21}	-0,987	0,104	0,013	0,105	-0,987	0,098	0,014	0,098
	β_{22}	1,056	0,247	0,056	0,252	1,027	0,214	0,027	0,215

Fonte: Elaborada pelo autor.

A Tabela 2 apresenta os resultados do estudo de simulação para avaliar o desempenho do modelo PVF-I-E sob diferentes tamanhos amostrais ($n = 200, 400, 800$) e dois valores do parâmetro de heterogeneidade ($\gamma = 0,60$ e $\gamma = 0,30$). Foram considerados 100 replicações para cada cenário. Os resultados demonstram que o modelo apresenta bom desempenho na recuperação dos parâmetros verdadeiros, com melhoria consistente à medida que o tamanho amostral aumenta. O viés e o REQM (Raiz do Erro Quadrático Médio) reduzem-se com o aumento de n , confirmando a consistência dos estimadores.

Para $n = 200$: Os estimadores apresentam viés moderado, especialmente para os parâmetros da componente de cura (β_{10} , β_{11} , β_{12}). O REQM varia entre 0,034 e 0,856, com os maiores erros associados ao intercepto da cura (β_{10}).

Para $n = 400$: Observa-se redução substancial no viés e no REQM para todos os parâmetros. Destaca-se a melhora na estimação de β_{10} , cujo REQM reduz-se de 0,856 para 0,542 no cenário com $\gamma = 0,30$.

Para $n = 800$: Os estimadores aproximam-se dos valores verdadeiros com alta precisão.

O REQM médio é inferior a 0,35 para todos os parâmetros, e o viés torna-se desprezível na maioria dos casos.

Parâmetro γ : Em ambos os cenários, o parâmetro de heterogeneidade é estimado com pequeno viés negativo (subestimação). Para $n = 800$, o viés é de -0,031 em ambos os casos, com REQM de 0,077 ($\gamma = 0,60$) e 0,060 ($\gamma = 0,30$), indicando melhor precisão quando a heterogeneidade é menor.

Componente de Cura ($\beta_{10}, \beta_{11}, \beta_{12}$): Os parâmetros da cura apresentam viés mais pronunciado para $\gamma = 0,30$ quando $n = 200$, mas esta diferença atenua-se com o aumento amostral. Para $n = 800$, os resultados são similares entre os dois cenários.

Componente de Risco ($\beta_{20}, \beta_{21}, \beta_{22}$): Os parâmetros do risco são estimados com maior precisão que os da cura, especialmente para $\gamma = 0,60$. O intercepto β_{20} apresenta viés negativo consistente em todos os cenários, indicando leve subestimação do risco basal.

3.5 Aplicação

A Síndrome da Imunodeficiência Adquirida (AIDS) constitui uma condição clínica caracterizada pela profunda deterioração da resposta imunológica do indivíduo, resultante da infecção pelo Vírus da Imunodeficiência Humana (HIV). A transmissão do HIV ocorre principalmente por meio do contato com sangue, sêmen ou fluidos vaginais contaminados. Após a infecção, é comum que o indivíduo apresente, em poucas semanas, sintomas inespecíficos semelhantes aos de um quadro gripal; contudo, na maior parte dos casos, a infecção permanece assintomática por um longo período, evoluindo gradualmente até o estágio clínico da AIDS. A progressão para AIDS envolve a ocorrência de manifestações clínicas como perda de peso, febre persistente, sudorese noturna, fadiga crônica e infecções recorrentes. Embora não exista cura para a doença, a utilização consistente e adequada da terapia antirretroviral (TARV) é capaz de suprimir a replicação viral, retardar de maneira substancial o avanço da infecção e reduzir a incidência de infecções oportunistas e outras complicações associadas.

No Brasil, o primeiro caso de AIDS foi identificado na década de 1980, e desde então foram implementados sistemas de informação epidemiológica que permitem acompanhar o perfil da epidemia no país. Por exemplo, os boletins epidemiológicos do Ministério de Saúde indicam que a Região Sul historicamente tem apresentado altas taxas de detecção de AIDS. [MEDEIROS et al. \(2020\)](#) indica que no estado de Paraná, especificamente, estudos apontam para uma das menores taxas de detecção da região Sul. [Lima et al. \(2023\)](#) mostra que, em 2020, 2,507 novos casos de HIV foram registrados no Sul, dos quais 697 eram no Paraná.

Nesse contexto, utiliza-se um conjunto de dados provenientes de uma coorte de indivíduos diagnosticados com HIV/AIDS no estado do Paraná, acompanhados entre 2002 e 2006, totalizando 272 observações. A análise proposta busca investigar os tempos de sobrevivência des-

ses pacientes em relação a diversas covariáveis clinicamente e epidemiologicamente relevantes, tais como idade, raça/cor, presença de infecções, entre outras. Uma caracterização desse banco de dados, bem como discussões adicionais sobre sua estrutura e particularidades epidemiológicas, pode ser encontrada em [Cancho *et al.* \(2021\)](#). Para fins de contextualização, a Tabela 3 apresenta as estatísticas descritivas referentes ao tempo (em anos) até o óbito por causas relacionadas ao HIV/AIDS, possibilitando uma visão preliminar da distribuição do tempo de sobrevivência na coorte analisada.

Tabela 3 – Medidas descritivas do tempo (em anos) até a morte do paciente por AIDS/HIV

	Mínimo	1º Quantil	Mediana	Média	3º Quantil	Máximo
Tempo (anos)	0,13	1,18	2,29	2,32	3,40	5,05

Fonte: Elaborada pelo autor.

Enquanto, a Figura 5 apresenta as estimativas de Kaplan-Meier para indivíduos com HIV no Paraná, evidenciando a presença de sobreviventes de longo prazo. A estimativa da função de sobrevivência (painel esquerdo) revela características importantes que fundamentam a escolha do modelo proposto. Observa-se que as curvas de sobrevivência para ambos os grupos (vía de infecção):

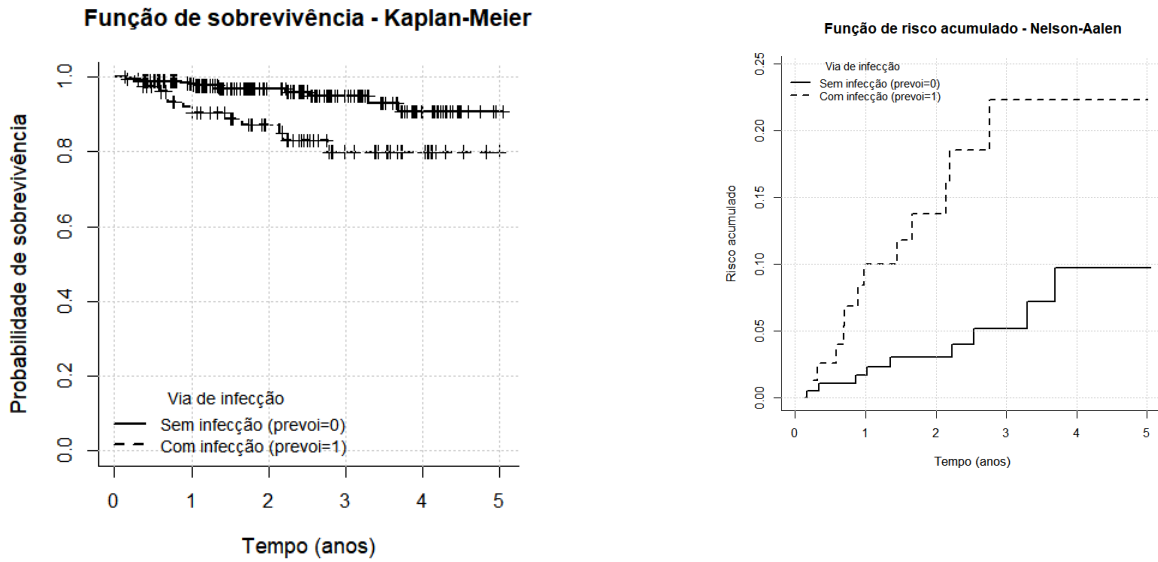
- $prevoi=0$: Sem infecção, pessoas que não apresentavam infecções oportunistas no momento do diagnóstico ou acompanhamento;
- $prevoi=1$: Com infecção, pessoas que apresentavam alguma infecção oportunista associada ao HIV/AIDS.

apresentam um patamar evidente em tempos mais avançados, não atingindo zero ao final do período de acompanhamento. Este comportamento é indicativo da presença de sobreviventes de longo prazo ou indivíduos curados, ou seja, uma fração da população que nunca experimentará o evento de interesse (óbito por AIDS), justificando a utilização do modelo de regressão PVF-ID-W. Por outra parte, a função de risco acumulado (painel direito) mostra-se limitada e com comportamento côncavo para ambos os grupos, indicando que pacientes com infecções oportunistas (linha tracejada) apresentam maior risco acumulado de morte relacionada à AIDS ao longo do tempo.

3.5.1 Análise dos dados

Neste estudo, a variável resposta do estudo foi o tempo até o óbito, definido como o intervalo, em anos, entre a data de inclusão no estudo e a data do óbito ou da última informação disponível (censura), representado por t , acompanhado do indicador de status δ , com $\delta = 1$ para óbito e $\delta = 0$ para censura. Além disso, este conjunto de dados apresenta uma taxa de censura

Figura 5 – Estimativa de Kaplan-Meier da função de sobrevivência para os dados de AIDS/HIV (painel esquerdo) e função de risco acumulado estratificada para pacientes com ou sem POI (painel direito).



Fonte: Elaborada pelo autor.

de aproximadamente 38% e a quantidade de pessoas do conjunto de dados foi de 272 pacientes. E foram consideradas duas covariáveis:

- \mathbf{x}_1 : idade do paciente no momento do diagnóstico, em anos completos (variável contínua), com faixa etária de 16 até 65 anos;
- \mathbf{x}_2 : via de infecção pelo HIV (prevoi), codificada como variável binária, assumindo $\mathbf{x}_2 = 0$ para a categoria de referência e $\mathbf{x}_2 = 1$ para a categoria de comparação.

A partir dessas variáveis, foram modeladas a probabilidade de cura por meio do preditor linear

$$\eta_1 = \mathbf{x}^\top \beta_1 = \beta_{10} + \beta_{11}x_1,$$

em que $\beta_1 = (\beta_{10}, \beta_{11})^\top$. A probabilidade de cura é dada pela função logística

$$p_0(x) = \frac{1}{1 + \exp(-\eta_1)} = \frac{1}{1 + \exp[-(\beta_{10} + \beta_{11}x_1)]}.$$

enquanto a variável idade (\mathbf{x}_1) e via de infecção (\mathbf{x}_4), sendo utilizada para modelar o risco dos indivíduos suscetíveis por meio do preditor linear

$$\eta_2 = \mathbf{w}^\top \beta_2 = \beta_{21}x_1 + \beta_{22}x_2,$$

em que $\beta_2 = (\beta_{21}, \beta_{22})^\top$. Logo, a função de sobrevivência dos indivíduos suscetíveis é dada por:

$$S_u(t | \mathbf{w}) = \exp \left\{ -\frac{1}{\gamma} \left[(1 + H_0(t) \exp(\eta_2))^\gamma - 1 \right] \right\},$$

em que $H_0(t) = \lambda t^\alpha$ é a função de risco acumulado basal da distribuição Weibull, com parâmetros $\alpha > 0$ (forma) e $\lambda > 0$ (escala), e $\gamma \in (0, 1)$ é o parâmetro de fragilidade da família PVF, que captura a heterogeneidade não observada entre os indivíduos suscetíveis. Finalmente, a função de sobrevivência populacional, que incorpora tanto os indivíduos curados quanto os suscetíveis, é definida como uma mistura:

$$S_p(t | \mathbf{x}, \mathbf{w}) = p_0(x) + (1 - p_0(x)) S_u(t | \mathbf{w}),$$

em que $p_0(\mathbf{x})$ representa a proporção de curados e $S_u(t | \mathbf{w})$ a sobrevivência dos suscetíveis. Esta função descreve a probabilidade de um paciente com características \mathbf{x} e \mathbf{w} sobreviver além do tempo t , considerando que pode pertencer à população de curados ou à população de suscetíveis. Assim a função completa com todos os parâmetros explícitos é:

$$S_p(t | x_1, x_4) = \frac{1}{1 + \exp[-(\beta_{10} + \beta_{11}x_1)]} + \left[1 - \frac{1}{1 + \exp[-(\beta_{10} + \beta_{11}x_1)]} \right] \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + \lambda t^\alpha e^{\beta_{21}x_1 + \beta_{22}x_4} \right)^\gamma - 1 \right] \right\}.$$

e na tabela 4 apresentam-se as interpretações para cada parâmetro

Tabela 4 – Parâmetros do modelo e suas interpretações

Índice	Nome	Interpretação
β_{10}	Intercepto da cura	Termo fixo na equação logística
β_{11}	Idade na cura	Efeito da idade na probabilidade de cura
β_{21}	Idade no risco	Efeito da idade no risco dos suscetíveis
β_{22}	prevoi no risco	Efeito da via de infecção no risco
α	Forma Weibull	Parâmetro de forma da distribuição Weibull
λ	Escala Weibull	Parâmetro de escala da distribuição Weibull
γ	Fragilidade PVF	Parâmetro de heterogeneidade não observada

Fonte: Elaborada pelo autor.

Para o modelo proposto foram adotadas a densidade a posteriori dada em (3.11) e as seguintes prioris independentes para os cálculos bayesianos são apresentadas na Tabela 5

Tabela 5 – Distribuições priori dos parâmetros do modelo

Priori	Justificativa
$\lambda \sim \text{Gamma}(0,1, 0,1)$	Priori pouca informativa com média 1 e variância 10
$\alpha \sim \text{Gamma}(0,1, 0,1)$	Priori pouca informativa; permite formas crescentes ou decrescentes
$\gamma \sim \text{Beta}(2, 2)$	Priori informativa, concentrada em torno de 0,5; respeita o domínio (0, 1)
$\beta_{1j} \sim \text{Normal}(0, 1)$	Priori pouca informativa; regulariza os coeficientes
$\beta_{2k} \sim \text{Normal}(0, 10)$	Prior pouca informativa (grande variância); permite efeitos fortes

Fonte: Elaborada pelo autor.

Além disso, a inferência bayesiana foi realizada com base em amostras geradas pelo algoritmo Hamiltonian Monte Carlo (HMC), considerando quatro cadeias independentes com 4,000 iterações cada para todos os parâmetros do modelo. Com o objetivo de reduzir a influência dos valores iniciais e garantir a convergência das cadeias, as primeiras 2,000 iterações de cada cadeia foram descartadas como período de aquecimento. A partir da modelagem, foram obtidas as estatísticas do LPML (*logaritmo da função de verossimilhança pseudo-marginal*) para os dados de HIV/AIDS, as quais são apresentadas na Tabela 6 os modelos exponencial por partes, para $J = 1, 2, 3, 4$ e 5 e o modelo PVF-ID-Weibull. Observe que, com base nas estatísticas do LPML, o modelo PVF-ID-W é considerado o de melhor ajuste aos dados de HIV/AIDS.

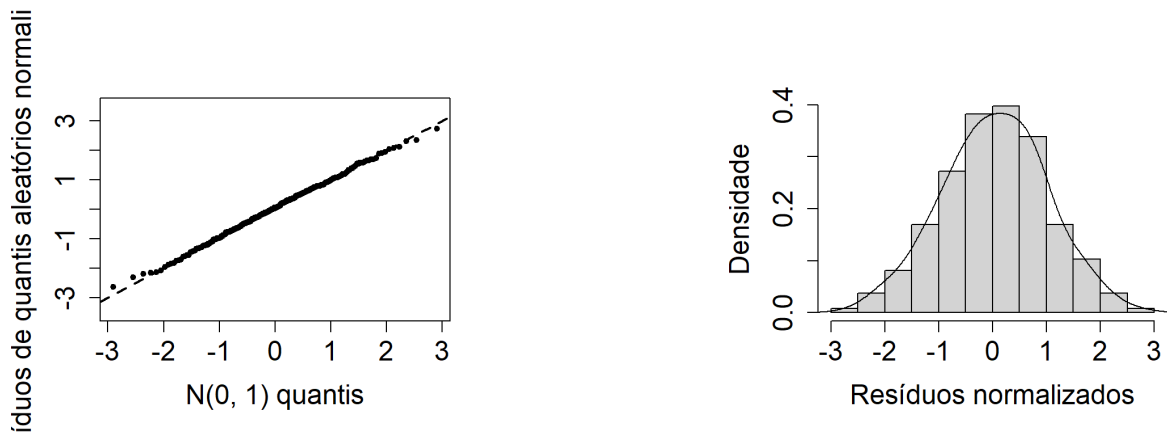
Tabela 6 – Estatísticas do LPML para os dados de HIV/AIDS

J	1	2	3	4	5	PVF-ID-Weibull
LPML	-89,92	-92,90	-94,85	-94,80	-96,67	-89,42

Fonte: Elaborada pelo autor.

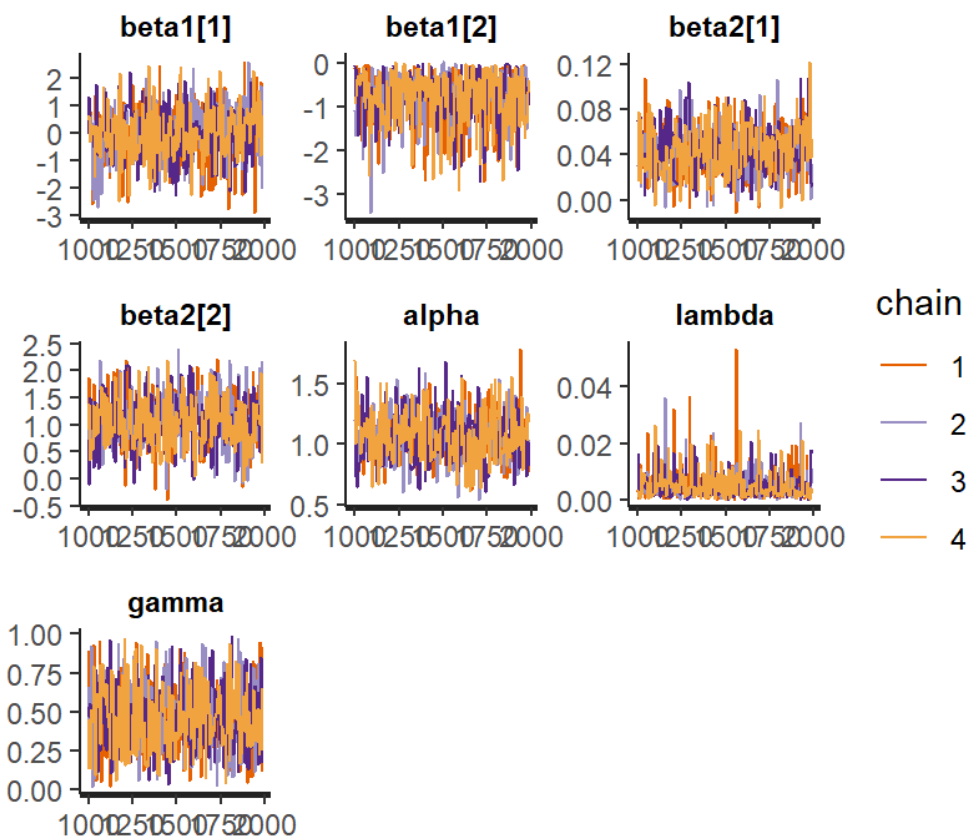
Para avaliar a adequação do ajuste do modelo PVF-ID-W foram obtidos os resíduos de quantis aleatórios normalizados a posteriores, sendo este um recurso empregado no trabalho de [Rigby e Stasinopoulos \(2005\)](#). O gráfico QQ-plot dos resíduos destes quantis aleatórios são exibidos na Figura 6 e sugerem que o modelo determinado possui um ajuste aceitável ao dados de AIDS/HIV. A convergência das cadeias MCMC foi avaliada seguindo os procedimentos de [Cowles e Carlin \(1996\)](#), utilizando gráficos de rastreamento (*trace plots*) para os parâmetros do modelo, conforme ilustrado na figura 7. O padrão observado nos gráficos sugere que as cadeias alcançaram estabilidade e convergência.

Figura 6 – Gráfico QQ-plot dos resíduos de quantis normalizados a posteriores com linha de identidade (a esquerda) e histograma do modelo PVF-ID Weibull sob as hipóteses definidas (a direita)



Fonte: Elaborada pelo autor.

Figura 7 – Gráficos de rastreamento para parâmetros do modelo PVF-ID Weibull referente aos dados de HIV/AIDS.

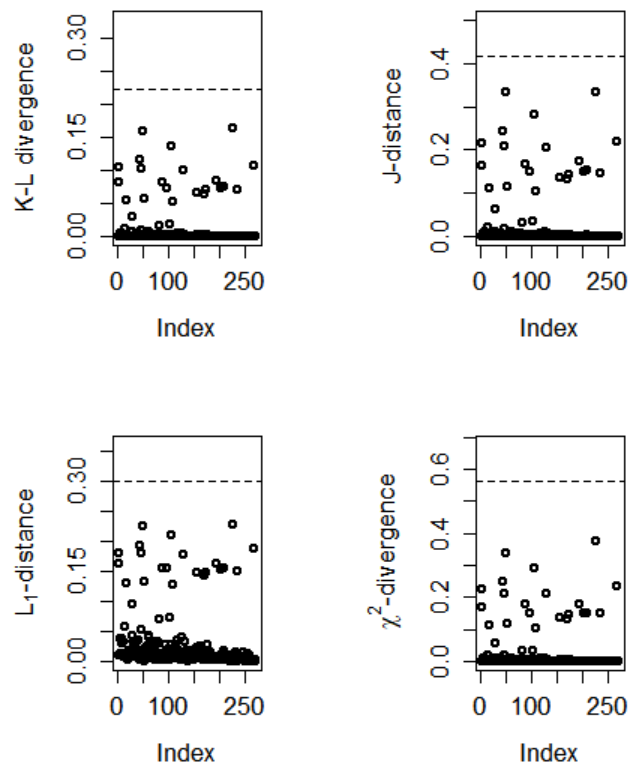


Fonte: Elaborada pelo autor.

Com base nas amostras geradas das distribuições a posteriori dos parâmetros, foram calculadas as medidas de ϕ -divergência descritas na Subsecção 2.5.4. A Figura 8 apresenta os

gráficos dos índices correspondentes às quatro medidas de divergência ϕ . A partir da análise desses índices, observa-se que não há indícios de observações potencialmente influentes sobre a distribuição a posteriori dos parâmetros do modelo de regressão. Assim, conclui-se que a presença ou remoção de observações individuais não acarreta mudanças inferenciais relevantes nos resultados obtidos.

Figura 8 – Gráficos de rastreamento para parâmetros do modelo PVF-ID Weibull referente aos dados de HIV/AIDS.



Fonte: Elaborada pelo autor.

As estimativas bayesianas paramétricas, bem como, as médias a posteriores, medianas, desvios padrões e 95% de confiança dos intervalos HPD (highest probability density (HPD)) para o modelo PVF-I-W são exibidas na Tabela 7.

Tabela 7 – Estatísticas LPML, médias a posteriores, medianas, desvios padrões e 95% de confiança dos intervalos HPD para os parâmetros do modelo PVF-ID Weibull relacionado aos dados de AIDS/HIV.

LPML	Parâmetros	Media	Mediana	Desvio Padrão	L	U
89,42	β_{10}	0,100	0,094	1,060	-1,881	2,033
	β_{11}	-0,718	-0,584	0,616	-1,922	0,019
	β_{21}	0,043	0,044	0,023	-0,003	0,083
	β_{22}	1,056	1,027	0,473	0,142	1,987
	α	1,094	1,081	0,223	0,645	1,522
	λ	0,006	0,004	0,007	0,000	0,018
	γ	0,483	0,484	0,220	0,072	0,877

Fonte: Elaborada pelo autor.

A interpretação de cada parâmetro, apresentam-se a seguir com foco no significado clínico para pacientes com HIV/AIDS:

- β_{10} : o intercepto da componente de cura não apresentou significância estatística, indicando que, para um paciente hipotético com idade zero (recém-nascido), a probabilidade basal de cura seria aproximadamente 52,5%. Não há recém-nascidos na amostra, este parâmetro serve apenas como referência para o modelo, não tendo interpretação clínica direta.
- β_{11} : A idade apresentou efeito negativo marginalmente significativo sobre a probabilidade de cura. Isto significaria que, pacientes mais velhos têm menor chance de serem curados da AIDS. A cada ano adicional de idade, a chance de cura reduz aproximadamente 51% (pois $OR = e^{\beta_{11}} = e^{-0,718} \approx 0,49$).
- β_{21} : O efeito da idade sobre o risco de morte dos pacientes não curados não foi estatisticamente significativo nesta amostra (IC 95% contém zero). Embora a estimativa pontual sugira um aumento de 4,4% no risco a cada ano (pois a razão de risco, $HR = e^{0,043} \approx 1,044$), a incerteza associada não permite afirmar que existe efeito real. Isto é, entre os pacientes que não atingem a cura, a idade isoladamente não é um preditor consistente da velocidade de progressão para o óbito.
- β_{22} : A presença de infecções oportunistas mostrou efeito positivo e altamente significativo sobre o risco de morte. Pacientes com infecção oportunista apresentam risco de morte 2,87 vezes resultado é o mais relevante do estudo, confirmando que infecções oportunistas são marcadores de gravidade e pior prognóstico em pacientes com HIV/AIDS, independentemente da idade.
- α : O parâmetro de forma da Weibull ($\alpha = 1,09$) indica uma tendência de risco crescente ao longo do tempo, ou seja, quanto mais tempo o paciente sobrevive, maior o risco instantâneo de morte. No entanto, como o intervalo de credibilidade contém o valor 1 (risco constante), não é possível afirmar com certeza se o risco é crescente ou constante. Clinicamente,

isto sugere que o risco de morte tende a aumentar com o tempo de acompanhamento, possivelmente refletindo as complicações e comorbidades em pacientes com HIV.

- λ : O parâmetro de escala da Weibull ($\lambda = 0,006$) apresenta valor muito pequeno, indicando que a escala temporal basal é longa. Em termos clínicos, isto significa que, na ausência de efeitos de covariáveis, os pacientes tendem a ter sobrevida prolongada. Este achado é consistente com a natureza crônica do HIV/AIDS na era da terapia antirretroviral, onde muitos pacientes vivem décadas após o diagnóstico.
- γ : O parâmetro de fragilidade PVF ($\gamma = 0,48$) indica heterogeneidade moderada entre os indivíduos suscetíveis. Valores próximos de 0,5 correspondem ao caso especial da distribuição Inversa Gaussiana. Isto significa que existem fatores não observados (não incluídos no modelo) que influenciam significativamente o risco de morte entre os pacientes não curados. Estes fatores podem incluir condições socioeconômicas, adesão ao tratamento, comorbidades não registradas, carga viral, entre outros. A presença de heterogeneidade moderada justifica a escolha do modelo com fragilidade, que captura esta variabilidade não explicada.

3.6 Conclusão

Neste capítulo, foi apresentado um modelo de fragilidade destinado a capturar a heterogeneidade não observada em dados de sobrevivência. O modelo combina uma distribuição de fragilidade pertencente à família PVF com o modelo Weibull utilizado como função de risco basal. A abordagem inferencial, apresentada baseia-se no uso do método Monte Carlo Hamiltoniano, implementado via R-Stan. O estudo de simulação conduzido nessa seção indica que o modelo exibe boas propriedades bayesianas, refletidas na estabilidade e precisão dos estimadores de Bayes. Finalmente, na última seção, a utilidade prática do modelo é avaliada por meio da aplicação em dados reais, envolvendo um estudo de AIDS/HIV. Esse exemplo empírico demonstra que o modelo proposto constitui uma alternativa robusta e eficaz para a análise desse tipo de dados.

MODELO DE SOBREVIVÊNCIA BIVARIADO

4.1 Modelo Bivariado de Longa Duração

Em muitos contextos médicos e biológicos, uma fração dos indivíduos pode nunca apresentar recorrência da doença, indicando a existência de um subgrupo curado ou imune dentro da população. Os modelos com fração de cura fornecem uma estrutura mais abrangente para a análise de dados de tempo até o evento, ao considerar explicitamente esse componente não suscetível, possibilitando inferências mais precisas e informativas. Tais percepções não podem ser obtidas a partir de modelos de sobrevivência convencionais, que assumem suscetibilidade universal. Na ausência de uma fração de cura, esses modelos reduzem-se naturalmente aos métodos padrão de análise de sobrevivência.

Nesse contexto, existe atualmente uma vasta literatura sobre “modelos univariados com fração de cura ou sobrevivência de longa duração” para dados de sobrevivência, os quais vêm sendo desenvolvidos ao longo do tempo, tais como o modelo de mistura de dois componentes de [Berkson e Gage \(1952\)](#) e o modelo de Risco acumulado Limitado (ou o modelo BCH) proposto por [Yakovlev, Tsodikov e Asselain \(1996\)](#). Diversos outros pesquisadores também contribuíram para essa área, conforme apresentado nos estudos de [Tsodikov, Ibrahim e Yakovlev \(2003\)](#), [Rodrigues et al. \(2009\)](#), [Ortega et al. \(2015\)](#), [Cordeiro et al. \(2016\)](#), [Koutras e Milienos \(2017\)](#), [Souza et al. \(2017\)](#), [Barriga et al. \(2019\)](#) e [Cancho et al. \(2021\)](#).

Diversas extensões dos modelos de sobrevivência de longa duração têm surgido na literatura. [Chen, Ibrahim e Sinha \(2002\)](#) propõem um modelo denominado modelo multivariado de taxa de cura, o qual se mostra bastante útil para a modelagem de dados multivariados nos quais as variáveis aleatórias conjuntas de falha apresentam uma fração imune ao evento de interesse. O modelo proposto pelos autores possui certa relação com o modelo univariado de taxa de cura discutido por [Yakovlev, Tsodikov e Asselain \(1996\)](#). Para induzir a estrutura de correlação entre os tempos de falha, os autores consideram um termo de fragilidade [Clayton](#)

(1978), o qual é assumido seguir uma distribuição estável positiva. [Cancho et al. \(2022\)](#) propõem um modelo multivariado de longa duração considerando a fragilidade como uma mistura das distribuições de Poisson e Gama. Por sua vez, [Cancho et al. \(2016\)](#) propuseram um modelo de sobrevivência multivariado com fração de cura, assumindo a existência de m tipos de causas (ou riscos) latentes e que o número de causas de cada tipo segue uma distribuição de Poisson multivariada.

Seguindo [Chen, Ibrahim e Sinha \(2002\)](#), propomos um novo modelo de sobrevivência bivariado com um componente de fragilidade modelado por meio de uma distribuição de mistura de Poisson combinada com a família de distribuições da *Power Variance Function* (PVF) [Hougaard \(2000\)](#). O modelo proposto engloba, como casos particulares, os modelos de [Chen, Ibrahim e Sinha \(2002\)](#) e [Cancho et al. \(2022\)](#). Ademais, estendemos a formulação para incluir estruturas de regressão, possibilitando a avaliação dos efeitos de covariáveis tanto sobre a distribuição de fragilidade quanto sobre o processo de sobrevivência.

Nesta seção, apresentamos a derivação e discussão do modelo bivariado de longa duração, o qual leva em conta a proporção de indivíduos não susceptíveis ao evento de interesse. A generalização para o caso multivariado pode ser realizada de forma imediata. Sejam T_1 e T_2 os tempos de sobrevivência associados a dois eventos, que podem ser interpretados de diferentes maneiras: seja para dois indivíduos relacionados (por exemplo, indivíduos pertencentes ao mesmo grupo), seja para um único indivíduo que vivencia dois eventos distintos ao longo do tempo. Para um indivíduo arbitrário na população em estudo, assume-se que as distribuições dos tempos T_k (onde $k = 1, 2$) podem ser modeladas por um modelo básico de fragilidade, o qual será detalhado a seguir.

[Hougaard \(1986\)](#) expandiu a distribuição Positiva Estável, derivando a família (PVF). Essa família engloba, como casos especiais, as distribuições Inversa Gaussiana e Gama, sendo esta última um caso limite. A família PVF pertence à família exponencial natural, na qual a variância é expressa como uma função potência da média [Hougaard \(2000\)](#). [Duchateau e Janssen \(2008\)](#) apresentam uma revisão abrangente da família PVF.

A densidade da distribuição PVF, proposta por [Hougaard \(2000\)](#), constitui uma família com três parâmetros, $\mu > 0$, $\sigma > 0$ e $0 < \gamma \leq 1$, cuja função densidade é dada por

$$g(w) = \exp \left\{ -\sigma w + \frac{\mu \sigma^\gamma}{\gamma} \right\} \left(-\frac{1}{\pi w} \right) \sum_{k=1}^{\infty} \frac{\Gamma(k\gamma + 1)}{k!} \left(-\frac{w^{-\gamma}}{\mu^\gamma} \right)^k \sin(\gamma k \pi), \quad w > 0. \quad (4.1)$$

A transformada de Laplace da distribuição de W , denotada por

$$\mathcal{L}_W(s) = \mathbb{E} [e^{-Ws}],$$

é dada por

$$\mathcal{L}_W(s) = \exp \left\{ -\frac{\mu}{\gamma} [(\sigma + s)^\gamma - \sigma^\gamma] \right\}, \quad (4.2)$$

em que $0 < \gamma \leq 1$, $\mu > 0$ e $\sigma > 0$. Todos os momentos positivos existem, e a esperança e a variância são dadas, respectivamente, por

$$\mathbb{E}[W] = \mu \sigma^{\gamma-1} \quad \text{e} \quad \text{Var}[W] = \mu(1-\gamma)\sigma^{\gamma-2}.$$

Conforme destacado por Hougaard (2000), a família PVF abrange os seguintes casos especiais e limites:

(i) A distribuição Positiva Estável é obtida quando $\sigma = 0$ e $\mu = \gamma$, com a transformada de Laplace $\mathcal{L}_W(s) = \exp\{-s^\gamma\}$.

(ii) Para $\gamma = 1/2$, a variável aleatória W segue uma distribuição Inversa Gaussiana (IG), com a transformada de Laplace

$$\mathcal{L}_W(s) = \exp\left\{-2\mu \left[(\sigma + s)^{1/2} - \sigma^{1/2}\right]\right\}.$$

(iii) No cenário limite em que $\gamma \rightarrow 0$, W apresenta uma distribuição Gama, cuja transformada de Laplace é dada por

$$\mathcal{L}_W(s) = \left(\frac{\sigma}{\sigma + s}\right)^\mu.$$

4.2 Modelo de sobrevivência bivariado de longa duração

Para cada indivíduo da população, seja N_k , $k = 1, 2$, o número não observável de causas (ou riscos) latentes associadas ao evento do tipo k . Condicionalmente a uma variável de fragilidade não negativa W , assume-se que

$$N_k | W = w \sim \text{Poisson}(w\rho_k), \quad k = 1, 2,$$

com $\rho_k > 0$ são parâmetros de incidência. O componente de fragilidade compartilhado W induz dependência entre as contagens latentes (N_1, N_2) e, conseqüentemente, entre os tempos de ocorrência observados correspondentes.

Assume-se que a variável de fragilidade W segue uma distribuição da PVF, com densidade dada em (4.1), reparametrizada de modo que $\mathbb{E}[W] = 1$. Sob essa parametrização, a transformada de Laplace de W é dada por

$$\mathcal{L}_W(s) = \mathbb{E}[e^{-sW}] = \exp\left\{-\frac{\sigma}{\gamma} \left[\left(1 + \frac{s}{\sigma}\right)^\gamma - 1\right]\right\}, \quad s \geq 0, \quad (4.3)$$

em que $0 < \gamma \leq 1$ e $\sigma > 0$. Os dois primeiros momentos de W são

$$\mathbb{E}[W] = 1 \quad \text{e} \quad \text{Var}(W) = \frac{1-\gamma}{\sigma}.$$

Essa parametrização garante a identificabilidade do modelo e permite interpretar σ como um parâmetro de precisão, de modo que valores maiores de σ correspondem a menores níveis de heterogeneidade não observada. Marginalmente, as contagens latentes satisfazem

$$\mathbb{E}[N_k] = \rho_k, \quad \text{Var}(N_k) = \rho_k + \rho_k^2 \frac{1-\gamma}{\sigma}, \quad \text{Cov}(N_1, N_2) = \rho_1 \rho_2 \frac{1-\gamma}{\sigma},$$

(veja a prova em B.0.1), evidenciando que a dependência entre N_1 e N_2 é inteiramente determinada pela variância da fragilidade.

Seja Z_{ki} o tempo latente de promoção associado à i -ésima causa do tipo k , $i = 1, 2, \dots$. Condicionalmente a N_k , as variáveis aleatórias $\{Z_{ki}\}_{i=1}^{N_k}$ são assumidas independentes e identicamente distribuídas, com função de distribuição comum $F_k(t) = 1 - S_k(t)$, a qual não depende de N_k . Além disso, condicionalmente a W , os processos associados aos diferentes tipos de evento são assumidos independentes. O tempo observável até a ocorrência do evento do tipo k é definido como

$$T_k = \min\{Z_{k1}, \dots, Z_{kN_k}\}, \quad k = 1, 2,$$

com a convenção de que $T_k = \infty$ quando $N_k = 0$, o que implica $\mathbb{P}(T_k > t_k | N_k = 0) = 1$.

Sob essa construção, a função de sobrevivência conjunta de (T_1, T_2) , condicionalmente a $W = w$, é dada por

$$\begin{aligned} S(t_1, t_2 | W = w) &= \prod_{k=1}^2 [\mathbb{P}(N_k = 0) + \mathbb{P}(Z_{k1} > t_k, \dots, Z_{kN_k} > t_k | N_k \geq 1)] \\ &= \prod_{k=1}^2 \left[\exp(-w\rho_k) + \sum_{l=1}^{\infty} \{S_k(t_k)\}^l \frac{(w\rho_k)^l}{l!} \exp(-w\rho_k) \right] \\ &= \exp \left(-w \sum_{k=1}^2 \rho_k [1 - S_k(t_k)] \right). \end{aligned} \quad (4.4)$$

Marginalizando em relação a W na expressão (4.4), obtém-se a função de sobrevivência conjunta de T_1 e T_2 , dada por

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma - 1 \right] \right\}. \quad (4.5)$$

(veja em B.0.2 a prova da equação (4.5)). Observe que

$$p_{00} = \lim_{t_1, t_2 \rightarrow \infty} S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 + \rho_2}{\sigma} \right)^\gamma - 1 \right] \right\} > 0,$$

(veja em B.0.3 a prova de p_{00}). A fração imune (ou de cura) corresponde ao subconjunto de clientes que, mesmo no longo prazo, não experimentam o evento de interesse — neste caso, o abandono de qualquer um dos produtos — caracterizando-se, portanto, como clientes persistentemente leais.

A partir de (4.5), as funções de sobrevivência marginais são dadas por

$$S_k(t_k) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k F_k(t_k)}{\sigma} \right)^\gamma - 1 \right] \right\}, \quad k = 1, 2. \quad (4.6)$$

(veja em B.0.4 a prova da expressão (4.6)). A equação (4.6) indica que cada função de sobrevivência marginal apresenta uma estrutura de modelo com taxa de cura, com probabilidade de cura

$$p_{0k} = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k}{\sigma} \right)^\gamma - 1 \right] \right\}, \quad k = 1, 2.$$

(veja a prova em (B.0.5)). É importante notar, a partir de (4.6), que cada função de sobrevivência marginal possui a mesma estrutura do modelo proposto por Cancho *et al.* (2021). Assim, o modelo proposto em (4.5) fornece uma extensão natural do modelo univariado de sobrevivência com fração de cura apresentado por Cancho *et al.* (2021). Além disso, observa-se que, quando $\gamma \rightarrow 1$ em (4.5), a função de sobrevivência conjunta se reduz ao produto de duas funções de sobrevivência independentes, cada uma com função de sobrevivência marginal $S_k(t_k) = \exp\{-\rho_k F_k(t_k)\}$ o que coincide com o modelo BCH proposto por Yakovlev, Tsodikov e Asselain (1996).

Uma propriedade limite da família da PVF é que, quando $\gamma \rightarrow 0$, a distribuição de fragilidade converge para a distribuição Gama. Utilizando o limite

$$\lim_{\gamma \rightarrow 0} \frac{a^\gamma - 1}{\gamma} = \log(a), \quad a > 0,$$

com

$$a = 1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma},$$

a função de sobrevivência conjunta em (4.5) satisfaz

$$\lim_{\gamma \rightarrow 0} S(t_1, t_2) = \exp \left\{ -\sigma \log \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right) \right\}. \quad (4.7)$$

(veja a prova em B.0.6). Consequentemente, no caso limite $\gamma \rightarrow 0$, a função de sobrevivência conjunta simplifica-se para

$$S(t_1, t_2) = \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\sigma}. \quad (4.8)$$

(veja a prova de (4.8) em B.0.7). Isso corresponde a um modelo bivariado de sobrevivência de longa duração com fragilidade Gama, tendo média $\mathbb{E}(W) = 1$ e variância $\text{Var}(W) = 1/\sigma$, conforme discutido por Cancho *et al.* (2022). Isso estabelece o modelo de fragilidade Gama como um caso limite da família PVF. O parâmetro σ desempenha um papel central no controle do grau de heterogeneidade não observada e da dependência entre os tempos de evento. Especificamente, σ atua como um parâmetro de precisão: valores maiores de σ correspondem a menor variância da fragilidade e, portanto, a uma dependência mais fraca entre (T_1, T_2) . No caso extremo $\sigma \rightarrow \infty$, a variância da fragilidade desaparece e a função de sobrevivência conjunta se fatoriza como

$$\lim_{\sigma \rightarrow \infty} S(t_1, t_2) = \exp\{-\rho_1 F_1(t_1)\} \exp\{-\rho_2 F_2(t_2)\},$$

(veja a prova em (B.0.8)), indicando a independência entre T_1 e T_2 .

Um caso especial importante da família PVF é obtido ao definir $\gamma = 1/2$, o que resulta na distribuição de fragilidade Inversa Gaussiana. Substituindo $\gamma = 1/2$ na função de sobrevivência conjunta (4.5), obtemos

$$S(t_1, t_2) = \exp \left\{ -2\sigma \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{1/2} - 1 \right] \right\}, \quad (4.9)$$

(veja a prova em (B.0.9)) ou equivalentemente,

$$S(t_1, t_2) = \exp \left\{ -2\sigma \left(\sqrt{1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}} - 1 \right) \right\}. \quad (4.10)$$

Sob essa especificação, a variável de fragilidade W segue uma distribuição Inversa Gaussiana, com média $\mathbb{E}(W) = 1$ e variância $\text{Var}(W) = 1/(2\sigma)$. Consequentemente, o parâmetro σ governa a magnitude da heterogeneidade não observada e, portanto, a intensidade da dependência entre os tempos de evento (T_1, T_2) . Valores maiores de σ correspondem a menor variância da fragilidade, resultando em dependência mais fraca, enquanto valores menores de σ induzem associação mais forte devido ao aumento da heterogeneidade. No caso limite $\sigma \rightarrow \infty$, o efeito da fragilidade desaparece e a função de sobrevivência conjunta converge para a estrutura de independência

$$\lim_{\sigma \rightarrow \infty} S(t_1, t_2) = \exp\{-\rho_1 F_1(t_1)\} \exp\{-\rho_2 F_2(t_2)\}.$$

Um submodelo adicional identificável surge da família PVF sob a restrição $\mu = \gamma$ e $\sigma = 0$, levando ao modelo proposto por [Chen, Ibrahim e Sinha \(2002\)](#). Nesse caso, a função de sobrevivência conjunta de (T_1, T_2) é dada por

$$S(t_1, t_2) = \exp \left\{ -(\rho_1 F_{01}(t_1) + \rho_2 F_{02}(t_2))^\gamma \right\}, \quad 0 < \gamma \leq 1. \quad (4.11)$$

(veja a prova em (B.0.10)). Essa especificação corresponde a um modelo bivariado de tempos de promoção conduzido por uma distribuição de fragilidade Estável Positiva. Uma característica distintiva desse modelo é que a variável de fragilidade não admite momentos finitos. Em particular,

$$\mathbb{E}[W^r] = \infty, \quad \forall r \geq 1, \quad (4.12)$$

implicando que nem a média nem a variância da distribuição de fragilidade existem. Consequentemente, a heterogeneidade e a dependência entre T_1 e T_2 são inteiramente determinadas pelo parâmetro de estabilidade γ , em vez de características baseadas em momentos.

Quando $\gamma = 1$, a função de sobrevivência conjunta em (4.11) se fatoriza como

$$S(t_1, t_2) = \exp\{-\rho_1 F_{01}(t_1)\} \exp\{-\rho_2 F_{02}(t_2)\}, \quad (4.13)$$

indicando independência entre T_1 e T_2 . Para $0 < \gamma < 1$, o modelo induz dependência positiva, sendo que valores menores de γ correspondem a associação mais forte entre os tempos de evento.

A ausência de momentos finitos distingue a fragilidade Estável Positiva de outros membros da família PVF, como os modelos de fragilidade Gama ou Inversa Gaussiana, e evidencia a natureza fundamentalmente diferente da estrutura de dependência induzida por esse submodelo.

4.2.1 Funções de Risco Conjunta e Marginais

Na análise de sobrevivência multivariada, a função de risco conjunta desempenha um papel central na caracterização da dependência instantânea entre múltiplos tempos de evento. A função de risco conjunta é definida como

$$h(t_1, t_2) = \frac{f(t_1, t_2)}{S(t_1, t_2)} = -\frac{\partial^2}{\partial t_1 \partial t_2} \log S(t_1, t_2).$$

Diferenciando o logaritmo da função de sobrevivência conjunta em (4.5), obtemos

$$h(t_1, t_2) = \rho_1 \rho_2 f_1(t_1) f_2(t_2) \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma \right] \\ \times \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{\gamma-2}, \quad 0 < \gamma < 1, \sigma > 0, \rho_1 > 0, \rho_2 > 0. \quad (4.14)$$

(veja a prova em (B.0.11)). O risco conjunto é estritamente positivo para $0 < \gamma < 1$ e se anula quando $\gamma = 1$, correspondendo à independência entre T_1 e T_2 .

A função de sobrevivência marginal de T_k em (4.6) foi obtida a partir de (4.5) fazendo $t_j \rightarrow \infty, j \neq k$. A função de risco marginal correspondente é

$$h_k(t_k) = -\frac{d}{dt_k} \log S_k(t_k) = \rho_k f_k(t_k) \left(1 + \frac{\rho_k F_k(t_k)}{\sigma} \right)^{\gamma-1}, \quad k = 1, 2. \quad (4.15)$$

(veja a prova em (B.0.12)). O parâmetro γ governa a estrutura de dependência entre T_1 e T_2 . Para $0 < \gamma < 1$, o risco conjunto em (4.14) é positivo, indicando dependência positiva induzida pelo componente de fragilidade compartilhado. Quando $\gamma \rightarrow 1$, o risco conjunto converge para zero e a função de sobrevivência conjunta se fatoriza, resultando em independência.

O parâmetro σ controla a quantidade de heterogeneidade não observada, uma vez que $\text{Var}(W) = (1 - \gamma)/\sigma$. Valores maiores de σ correspondem a menor variância da fragilidade e conduzem a riscos marginais que se aproximam dos riscos basais $\rho_k f_k(t_k)$. Valores menores de σ aumentam a heterogeneidade e fortalecem a dependência entre T_1 e T_2 .

Mesmo quando as densidades basais $f_k(\cdot)$ correspondem a modelos de Weibull com taxas de risco crescentes, o termo multiplicativo $\left(1 + \frac{\rho_k F_k(t_k)}{\sigma} \right)^{\gamma-1}$ pode atenuar ou suavizar o crescimento dos riscos marginais, produzindo uma ampla gama de formas para as funções de risco.

4.2.2 Medida Local de Associação

No modelo bivariado de sobrevivência de longo prazo proposto, a função de sobrevivência conjunta $S(t_1, t_2)$ é, em geral, imprópria devido à presença de uma fração de cura não nula. Como consequência, medidas globais clássicas de dependência, tais como o τ de Kendall, não são bem definidas ou não são identificáveis a partir da distribuição conjunta de (T_1, T_2) . Essa limitação

motiva a adoção de medidas locais de associação, que permanecem bem definidas mesmo sob funções de sobrevivência impróprias.

Seguindo Clayton (1978) e Oakes (1982), consideramos uma medida local de associação que captura a dependência instantânea entre T_1 e T_2 em um dado ponto no tempo (t_1, t_2) . Essa medida, denotada por $\vartheta^*(t_1, t_2)$, é definida em termos da função de sobrevivência conjunta como

$$\vartheta^*(t_1, t_2) = \frac{S(t_1, t_2) \frac{\partial^2}{\partial t_1 \partial t_2} S(t_1, t_2)}{\left(\frac{\partial}{\partial t_1} S(t_1, t_2) \right) \left(\frac{\partial}{\partial t_2} S(t_1, t_2) \right)}. \quad (4.16)$$

A quantidade $\vartheta^*(t_1, t_2)$ pode ser interpretada como uma medida de associação instantânea entre os dois tempos de evento, condicional à sobrevivência conjunta até (t_1, t_2) . Valores maiores que um indicam dependência local positiva, valores menores que um indicam dependência local negativa, e $\vartheta^*(t_1, t_2) = 1$ corresponde à independência local. Como essa medida depende exclusivamente de derivadas da função de sobrevivência conjunta, ela permanece significativa mesmo quando $S(t_1, t_2)$ não se integra a um.

Para o modelo bivariado com taxa de cura definido em (4.5), a medida de dependência local $\vartheta^*(t_1, t_2)$ é bem definida e assume a forma explícita

$$\vartheta^*(t_1, t_2) = 1 + \frac{1 - \gamma}{\sigma} \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\gamma}. \quad (4.17)$$

(veja a prova em (B.0.13)), em que $F_k(t_k)$ denota a função de distribuição latente marginal e $\rho_k > 0$ são parâmetros de incidência.

Essa expressão revela diversas características importantes da estrutura de dependência induzida pelo modelo de fragilidade PVF. Para todo $t_1, t_2 \geq 0$, tem-se $\vartheta^*(t_1, t_2) \geq 1$, indicando que o modelo induz dependência local positiva entre os tempos de evento, com igualdade ocorrendo se, e somente se, T_1 e T_2 forem localmente independentes. Na origem, a medida se reduz a $\vartheta^*(0, 0) = 1 + (1 - \gamma)/\sigma$, mostrando que a intensidade inicial da dependência é governada conjuntamente pelos parâmetros σ e γ , sendo que valores menores de qualquer um desses parâmetros correspondem a uma associação inicial mais forte.

À medida que o tempo avança, a dependência local enfraquece. Como as funções de distribuição marginais satisfazem $F_k(t_k) \rightarrow 1$ quando $t_k \rightarrow \infty$, segue-se que $\vartheta^*(t_1, t_2) \rightarrow 1$, implicando que a dependência entre T_1 e T_2 desaparece assintoticamente e os tempos de evento tornam-se independentes no longo prazo. Além disso, como $F_k(t_k)$ são funções não decrescentes, $\vartheta^*(t_1, t_2)$ é monotonicamente decrescente em cada argumento, caracterizando uma estrutura de dependência dinâmica que é mais forte em tempos iniciais e se dissipa gradualmente ao longo do tempo.

Os papéis dos parâmetros σ e γ são particularmente elucidativos. O parâmetro σ atua como um parâmetro de precisão da distribuição de fragilidade, com valores maiores implicando menor variabilidade da fragilidade e dependência local mais fraca. O parâmetro γ controla a

evolução temporal da dependência. Em particular, quando $\gamma \rightarrow 0$,

$$\vartheta^*(t_1, t_2) \longrightarrow 1 + \frac{1}{\sigma},$$

o que mostra que a dependência local torna-se constante ao longo do tempo e não depende das distribuições latentes marginais. Esse caso limite corresponde ao modelo clássico de fragilidade gama, no qual a associação entre os tempos de evento é homogênea e inteiramente determinada pela variância da fragilidade compartilhada. Para $0 < \gamma < 1$, a dependência decresce suavemente à medida que o tempo aumenta, sendo que valores maiores de γ conduzem a uma dissipação mais rápida da associação.

Por fim, uma vantagem importante de $\vartheta^*(t_1, t_2)$ é sua aplicabilidade sob funções de sobrevivência conjuntas impróprias. Mesmo em modelos de sobrevivência de longo prazo com fração de cura, nos quais medidas globais de dependência, como o τ de Kendall, podem não existir ou não ser identificáveis, a medida de dependência local permanece bem definida e fornece uma quantificação significativa e interpretável da estrutura de associação.

4.3 Inferência Bayesiana

4.3.1 Função de verossimilhança

Suponha que tenhamos n indivíduos, e seja N_{ki} o número de riscos latentes para o i -ésimo indivíduo, $i = 1, \dots, n$, $k = 1, 2$. Ademais, assume-se que os N_{ki} são variáveis aleatórias de Poisson independentes, com média $w_i \rho_k$, $i = 1, \dots, n$, $k = 1, 2$. Também se assume que os w_i são i.i.d. com distribuição Gama, média 1 e variância γ .

Além disso, seja $Z_{ki,1}, \dots, Z_{ki,N_{ki}}$ os tempos latentes independentes correspondentes aos N_{ki} riscos latentes do i -ésimo indivíduo, que não são observados, e todos possuem função de distribuição acumulada $F_k(\cdot)$. Conforme em [Chen, Ibrahim e Sinha \(2002\)](#), consideramos uma forma paramétrica para $F_k(\cdot)$, tal que seja uma distribuição Weibull ou Gama. Denotamos o parâmetro de indexação (possivelmente vetorial) por ϕ_k , e, portanto, escrevemos $F_k(\cdot | \phi_k)$ e $S_k(\cdot | \phi_k)$. Por exemplo, se $F_k(\cdot | \phi_k)$ corresponde a uma distribuição Weibull, então $\phi_k = (\alpha_k, \lambda_k)$, em que α_k é o parâmetro de forma e λ_k é o parâmetro de escala. Seja T_{ki} o tempo de falha do indivíduo i para o k -ésimo componente, podendo ser observação censurada à direita. Seja C_{ki} o tempo de censura, de modo que observamos

$$y_{ki} = \min(T_{ki}, C_{ki}) \quad \text{e} \quad \delta_{ki} = I(T_{ki} \leq C_{ki}),$$

em que $\delta_{ki} = 1$ se T_{ki} for um tempo de falha e $\delta_{ki} = 0$ se for censurado à direita. Definimos $\mathbf{y}_k = (y_{k1}, \dots, y_{kn})$, $\boldsymbol{\delta}_k = (\delta_{k1}, \dots, \delta_{kn})$, $\mathbf{N}_k = (N_{k1}, \dots, N_{kn})$, $k = 1, 2$, $\mathbf{w} = (w_1, \dots, w_n)$.

Os dados completos são representados por

$$\mathcal{D} = (n, \mathbf{y}_1, \dots, \mathbf{y}_m, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \mathbf{N}_1, \mathbf{N}_2, \mathbf{w}),$$

em que $\mathbf{N}_1, \mathbf{N}_2$ e \mathbf{w} são vetores aleatórios não observados, e os dados observados são

$$\mathcal{D}_{\text{obs}} = (n, \mathbf{y}_1, \mathbf{y}_2, \delta_1, \delta_2).$$

Adicionalmente, seja $\boldsymbol{\rho} = (\rho_1, \rho_2)$ e $\boldsymbol{\eta} = (\eta_1, \eta_2)$. A função de verossimilhança de $\boldsymbol{\vartheta} = (\boldsymbol{\rho}, \boldsymbol{\eta}, \gamma, \sigma)$ baseada nos dados completos \mathcal{D} é dada por

$$\begin{aligned} L(\boldsymbol{\vartheta} | \mathcal{D}) &= \prod_{k=1}^2 \prod_{j=1}^n [S_{0k}(y_{kj} | \eta_k)]^{Z_{kj} - \delta_{kj}} [Z_{kj} f_{0k}(y_{kj} | \eta_k)]^{\delta_{kj}} \\ &\times \prod_{k=1}^2 \exp \left(\sum_{j=1}^n [Z_{kj} \log(w_j \rho_k) - \log(Z_{kj}!) - w_j \rho_k] \right) \prod_{j=1}^n g(w_j | \boldsymbol{\mu}, \sigma, \gamma), \end{aligned} \quad (4.18)$$

(veja a construção de (4.18) em (B.0.14)), em que $f_k(y_{kj} | \eta_k)$ é a densidade correspondente à função de distribuição $F_k(y_{kj} | \eta_k)$.

Assumimos uma densidade Weibull para $f_k(y_{kj} | \eta_k)$, tal que

$$f_k(y | \eta_k) = \alpha_k y^{\alpha_k - 1} \exp\{-\lambda_k - y^{\alpha_k} e^{\lambda_k}\}, \quad \alpha_k > 0, \lambda_k \in \mathbb{R}. \quad (4.19)$$

A função de verossimilhança dos dados observados é obtida integrando (4.18) em relação a (N, w) , assumindo uma distribuição PVF para cada w_i , com média 1 e variância $(1 - \gamma)/\sigma$, e densidade $g(\cdot | \boldsymbol{\mu}, \sigma, \gamma)$, dada em (4.1). Assim, então, tem-se ao seguinte teorema que determina a verossimilhança dos dados observados:

Teorema 2. A função de verossimilhança baseada nos dados observados, denotada por $L(\boldsymbol{\vartheta} | \mathcal{D}_{\text{obs}})$, é dada por

$$\begin{aligned} L(\boldsymbol{\vartheta} | \mathcal{D}_{\text{obs}}) &= \prod_{j=1}^n (\rho_1 f_1(y_{1j} | \eta_1))^{\delta_{1j}} (\rho_2 f_2(y_{2j} | \eta_2))^{\delta_{2j}} \\ &\times \prod_{j=1}^n \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &\times \prod_{j=1}^n \left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}} \\ &\times \prod_{j=1}^n \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}}. \end{aligned}$$

(veja a prova em (B.0.15)). As covariáveis são incorporadas ao modelo (4.5) por meio dos parâmetros de média dos riscos latentes, ρ_1 e ρ_2 . Consequentemente, cada indivíduo $j = 1, \dots, n$ pode ter um número médio distinto de causas latentes para cada tipo de evento, denotado por ρ_{kj} , $k = 1, 2$.

Seja $\mathbf{x}_j^\top = (x_{j1}, x_{j2}, \dots, x_{jp})$ o vetor de p covariáveis associadas ao j -ésimo indivíduo, e seja $\boldsymbol{\beta}_k = (\beta_{k1}, \beta_{k2}, \dots, \beta_{kp})^\top$ o vetor correspondente de coeficientes de regressão para o risco

latente do tipo k . A relação entre o risco latente médio e as covariáveis é especificada por uma função de ligação logarítmica:

$$\rho_{kj} = \exp(\mathbf{x}_j^\top \beta_k),$$

o que garante positividade e permite interpretação em termos de risco relativo.

Definindo o preditor linear como

$$\eta_{kj} = \mathbf{x}_j^\top \beta_k,$$

essa formulação afeta diretamente tanto a probabilidade de imunidade estrutural quanto a função de risco marginal para cada tipo de evento.

A probabilidade de que o indivíduo j seja imune ao evento k é dada por

$$p_{0kj} = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\exp\{\eta_{kj}\}}{\sigma} \right)^\gamma - 1 \right] \right\}, \quad \gamma \in (0, 1). \quad (4.20)$$

(a prova de 4.20 está em B.0.15), como $\exp(\eta_{kj})$ é uma função estritamente crescente de η_{kj} , segue-se que a probabilidade de imunidade p_{0kj} é monotonicamente decrescente no preditor linear. Portanto, componentes positivos de β_k estão associados a menor probabilidade de imunidade, enquanto componentes negativos aumentam a probabilidade de imunidade estrutural.

Para $\gamma \in (0, 1)$, a transformação

$$\left(1 + \frac{\exp\{\eta_{kj}\}}{\sigma} \right)^\gamma$$

é côncava em $\exp\{\eta_{kj}\}$, implicando efeitos marginais decrescentes das covariáveis sobre a probabilidade de imunidade. Em particular, para indivíduos com valores elevados de η_{kj} , aumentos adicionais nos valores das covariáveis produzem mudanças progressivamente menores em p_{0kj} , refletindo um efeito de saturação para perfis de alto risco. Esse comportamento surge naturalmente da interação entre covariáveis observadas e a heterogeneidade não observada induzida pela fragilidade PVF compartilhada.

A função de risco marginal correspondente para o evento k é dada por

$$h_k(t_{kj} | \mathbf{x}) = \rho_{kj} f_k(t_{kj} | \eta_k) \left(1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma} \right)^{\gamma-1}, \quad k = 1, 2, \quad (4.21)$$

(a prova de (4.21) está em B.0.15), em que $f_k(\cdot | \eta_k)$ e $F_k(\cdot | \eta_k)$ denotam as funções densidade e distribuição de referência para a margem k , respectivamente.

Na expressão (4.21), as covariáveis afetam a função de risco tanto multiplicativamente através de ρ_{kj} quanto dinamicamente através do fator dependente do tempo

$$\left(1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma} \right)^{\gamma-1}.$$

Como $\gamma \in (0, 1)$ implica $\gamma - 1 < 0$, esse fator é decrescente em t_k , levando à atenuação do efeito das covariáveis ao longo do tempo e, conseqüentemente, à violação da suposição de riscos proporcionais.

Nos primeiros tempos de acompanhamento, quando $F_k(t_{kj} | \eta_k) \approx 0$, o risco marginal simplifica-se para

$$h_k(t_{kj} | \mathbf{x}) \approx \rho_{kj} f_k(t_{kj} | \eta_k),$$

o que se assemelha a uma estrutura de riscos proporcionais. À medida que o tempo avança, entretanto, o termo decrescente induzido por $\gamma < 1$ enfraquece a influência de ρ_{kj} , capturando o efeito de seleção dinâmica induzida pela fragilidade compartilhada. Indivíduos com maior susceptibilidade latente tendem a experimentar o evento mais cedo, deixando uma subpopulação progressivamente mais robusta em tempos posteriores.

De forma geral, para $\gamma \in (0, 1)$, o modelo bivariado proposto com fragilidade compartilhada fornece uma estrutura flexível que acomoda simultaneamente a imunidade estrutural e os efeitos de covariáveis variando no tempo. As covariáveis governam tanto a probabilidade de imunidade quanto a dinâmica do risco de curto prazo, enquanto o parâmetro de fragilidade γ controla a taxa na qual os efeitos das covariáveis se atenuam ao longo do tempo, tornando o modelo particularmente adequado para dados de sobrevivência que exibem dependência decrescente e riscos não proporcionais.

Sob essas condições, seja $\beta = (\beta_1^\top, \beta_2^\top)^\top$, a função de verossimilhança baseada nos dados observados para o vetor de parâmetros $(\beta, \eta_1, \eta_2, \gamma, \sigma)$ pode ser então expressa como

$$\begin{aligned} L(\beta, \eta_1, \eta_2, \gamma, \sigma | \mathcal{D}_{\text{obs}}) &= \prod_{j=1}^n \prod_{k=1}^2 (\exp(\mathbf{x}_j^\top \beta_k) f_k(y_{kj} | \eta_k))^{\delta_{kj}} \\ &\times \prod_{j=1}^n \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &\times \prod_{j=1}^n \left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}} \\ &\times \prod_{j=1}^n \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}}, \end{aligned} \quad (4.22)$$

em que $\mathcal{D}_{\text{obs}} = (n, \mathbf{y}_1, \mathbf{y}_2, \mathbf{X}, \delta_1, \delta_2)$, \mathbf{X} é a matriz $n \times p$ de covariáveis, $f_k(y_k | \eta_k)$ é dada em (4.19) e $F_k(y_k | \eta_k) = 1 - \exp(-y_k^{\alpha_k} \exp(\lambda_k))$. A prova da equação (4.22) está em B.0.15.

4.3.2 Prioris e posteriori

Assumimos que $\gamma, \eta_1 = (\alpha_1, \lambda_1), \eta_2 = (\alpha_2, \lambda_2), \beta_1$ e β_2 são independentes, ou seja,

$$\pi(\gamma, \eta_1, \eta_2, \beta_1, \beta_2) = \pi(\gamma) \prod_{k=1}^2 \pi(\beta_k, \eta_k), \quad (4.23)$$

em que assumimos que $\gamma \sim \text{Beta}(a, b)$ e que

$$\pi(\beta_k, \eta_k) = \phi(\beta_k, \alpha_k, \lambda_k) = \phi(\alpha_k | a_0, b_0) \pi(\lambda_k | c_0) \pi(\beta_k | d_0), \quad k = 1, 2.$$

considerando:

$$\begin{aligned} \pi(\alpha_k) &\propto \alpha_k^{a_0-1} \exp\{-b_0 \alpha_k\}, \\ \pi(\lambda_k) &\propto \exp\{-c_0 \lambda_k^2\}, \\ \pi(\beta_k | d_0) &\propto \exp\left\{-d_0 \sum_{l=1}^p \beta_{kl}^2\right\}, \end{aligned}$$

e a, b, a_0, b_0, c_0 e d_0 são hiperparâmetros especificados.

Combinando essas especificações com a função de verossimilhança (4.22), a distribuição a posteriori de γ, η_1, η_2 e β é dada por

$$\pi(\gamma, \eta_1, \eta_2, \beta_1, \beta_2 | \mathcal{D}_{\text{obs}}) \propto L(\gamma, \eta_1, \eta_2, \beta_1, \beta_2 | \mathcal{D}_{\text{obs}}) \pi(\gamma) \prod_{k=1}^2 \pi(\beta_k, \eta_k). \quad (4.24)$$

A distribuição (4.24) é analiticamente intratável. Portanto, nossa inferência baseia-se em métodos de simulação de Monte Carlo via cadeias de Markov (MCMC). Em particular, algoritmos de Hamiltonian Monte Carlo (HMC) têm se mostrado uma alternativa poderosa, principalmente devido à disponibilidade de programas computacionais.

Os cálculos neste trabalho foram implementados utilizando softwares estatísticos de código aberto. Em particular, utilizamos o ambiente R [Team \(2020b\)](#). O pacote `rstan` é uma interface para o software Bayesiano de código aberto Stan ([Team \(2018\)](#)). O Stan utiliza métodos de Hamiltonian Monte Carlo (HMC) acoplados ao *no-U-turn sampler* (NUTS), projetados para melhorar a velocidade, estabilidade e escalabilidade em comparação com métodos MCMC padrão, como Metropolis-Hastings e o *Gibbs sampler* [McElreath \(2018\)](#).

4.4 Estudo de simulação

Nesta Seção, descreve-se detalhadamente o procedimento para gerar os dados que seguem o modelo PVF bivariado com fração de cura. Para cada indivíduo $i = 1, \dots, n$, o algoritmo percorre um conjunto de etapas sequenciais que reproduzem a estrutura hierárquica do modelo. O processo inicia com a geração da covariável observável, prossegue com a construção das taxas de ocorrência, a geração da fragilidade latente, o número de causas, os tempos de promoção, e finaliza com a aplicação da censura. O Algoritmo apresenta o pseudocódigo completo do procedimento.

Algoritmo – Geração de dados do modelo PVF bivariado com fração de cura

Entrada:

- n : tamanho da amostra
- σ, γ : parâmetros da distribuição PVF
- $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$: coeficientes de regressão
- $\alpha_{01}, \lambda_{01}, \alpha_{02}, \lambda_{02}$: parâmetros Weibull
- τ_{u1}, τ_{u2} : limites superiores para a censura

Saída:

- $D = \{(y_{1i}, y_{2i}, \delta_{1i}, \delta_{2i}, x_i), i = 1, \dots, n\}$

Procedimento:

1. Para $i = 1$ até n faça:

a. Geração da covariável:

$$x_i \sim \text{Bernoulli}(0,5)$$

b. Cálculo das taxas de ocorrência:

$$\rho_{1i} = \exp(\beta_{11} + \beta_{12}x_i)$$

$$\rho_{2i} = \exp(\beta_{21} + \beta_{22}x_i)$$

c. Geração da fragilidade W_i da distribuição PVF(σ, γ) com média 1:

i. Se $\gamma = 1$ então:

$$W_i \leftarrow 1$$

ii. Senão se $|\gamma - 0,5| < 10^{-6}$ então:

$$W_i \sim \text{Inversa Gaussiana}(\text{média} = 1, \text{forma} = 2\sigma)$$

iii. Senão se $\gamma < 0,01$ então:

$$W_i \sim \text{Gamma}(\text{forma} = \sigma, \text{escala} = \sigma)$$

iv. **Senão:**

Inicializar $tentativas \leftarrow 0$, $W_i \leftarrow$ Nulo

Enquanto (W_i é Nulo) e ($tentativas < 10000$) **faça:**

$W_{prop} \sim \text{Gamma}(\text{forma} = \sigma, \text{escala} = \sigma)$

$\log f_{alvo} \leftarrow -\sigma W_{prop} + \frac{\sigma^\gamma}{\gamma} - (1 - \gamma) \log(W_{prop})$

$\log f_{prop} \leftarrow \log \text{Gamma}(W_{prop} \mid \text{forma} = \sigma, \text{escala} = \sigma)$

$u \sim \text{Uniforme}(0, 1)$

Se $\log(u) + \log(2, 0) < \log f_{alvo} - \log f_{prop}$ **então:**

$W_i \leftarrow W_{prop}$

$tentativas \leftarrow tentativas + 1$

Fim Enquanto

Se W_i ainda é Nulo **então:**

$W_i \sim \text{Gama}(\text{forma} = \sigma, \text{escala} = \sigma)$

d. **Geração do número de causas latentes:**

$N_{1i} \sim \text{Poisson}(W_i \rho_{1i})$

$N_{2i} \sim \text{Poisson}(W_i \rho_{2i})$

e. **Geração do tempo para o primeiro evento (T_{1i}):**

Se $N_{1i} = 0$ **então:**

$T_{1i} \leftarrow \infty$

Senão:

$\text{escala}_1 \leftarrow \exp(-\lambda_{01}/\alpha_{01})$

Para $k = 1$ **até** N_{1i} **faça:**

$Z_{1k} \sim \text{Weibull}(\text{forma} = \alpha_{01}, \text{escala} = \text{escala}_1)$

$T_{1i} \leftarrow \min\{Z_{11}, Z_{12}, \dots, Z_{1N_{1i}}\}$

f. **Geração do tempo para o segundo evento (T_{2i}):**

Se $N_{2i} = 0$ **então:**

$T_{2i} \leftarrow \infty$

Senão:

$\text{escala}_2 \leftarrow \exp(-\lambda_{02}/\alpha_{02})$

Para $k = 1$ **até** N_{2i} **faça:**

$Z_{2k} \sim \text{Weibull}(\text{forma} = \alpha_{02}, \text{escala} = \text{escala}_2)$

$T_{2i} \leftarrow \min\{Z_{21}, Z_{22}, \dots, Z_{2N_{2i}}\}$

g. **Geração dos tempos de censura:**

$$C_{1i} \sim \text{Uniforme}(0, \tau_{u1})$$

$$C_{2i} \sim \text{Uniforme}(0, \tau_{u2})$$

h. **Determinação dos tempos observados e indicadores de falha:**

$$y_{1i} = \min(T_{1i}, C_{1i})$$

$$y_{2i} = \min(T_{2i}, C_{2i})$$

$$\delta_{1i} = \begin{cases} 1, & \text{se } T_{1i} \leq C_{1i} \\ 0, & \text{caso contrário} \end{cases}$$

$$\delta_{2i} = \begin{cases} 1, & \text{se } T_{2i} \leq C_{2i} \\ 0, & \text{caso contrário} \end{cases}$$

2. **Fim**3. **Retorna:**

$$D = \{(y_{1i}, y_{2i}, \delta_{1i}, \delta_{2i}, x_i), i = 1, \dots, n\}$$

4.4.1 Detalhamento das Etapas do Algoritmo**4.4.1.1 Etapa 1: Geração da Covariável**

A covariável x_i é gerada a partir de uma distribuição Bernoulli com probabilidade de sucesso 0,5:

$$x_i \sim \text{Bernoulli}(0,5)$$

Esta covariável é compartilhada pelos dois tempos de sobrevivência do mesmo indivíduo, representando uma característica fixa que influencia simultaneamente T_1 e T_2 . A escolha de probabilidade 0,5 garante grupos balanceados.

4.4.1.2 Etapa 2: Cálculo das Taxas de Ocorrência

A partir da covariável gerada, constroem-se as taxas de ocorrência de causas latentes para cada tempo:

$$\rho_{1i} = \exp(\beta_{11} + \beta_{12}x_i), \quad \rho_{2i} = \exp(\beta_{21} + \beta_{22}x_i)$$

A função exponencial garante que as taxas sejam positivas, como requerido pelo modelo Poisson.

4.4.1.3 Etapa 3: Geração da Fragilidade PVF

A variável latente de fragilidade W_i é gerada a partir da distribuição PVF com parâmetros σ e γ , sujeita à condição $\mathbb{E}(W_i) = 1$. Devido à complexidade da distribuição PVF, que não possui forma fechada simples para sua função de densidade, o algoritmo implementa diferentes estratégias dependendo do valor de γ :

- **Caso $\gamma = 1$:** Neste caso degenerado, a distribuição PVF reduz-se a um ponto de massa em 1, indicando independência condicional. Portanto, $W_i = 1$ para todos os indivíduos.
- **Caso $|\gamma - 0,5| < 10^{-6}$:** Quando γ é aproximadamente 0,5, a distribuição PVF corresponde à distribuição Inversa Gaussiana com média 1 e parâmetro de forma 2σ . Neste caso, utiliza-se um gerador específico para esta distribuição.
- **Caso $\gamma < 0,01$:** Para valores muito pequenos de γ , a distribuição PVF aproxima-se da distribuição **Gamma** com parâmetros forma = σ e escala = σ .

4.4.1.4 Etapa 4: Geração do Número de Causas Latentes

Condicional à fragilidade W_i e às taxas ρ_{1i} e ρ_{2i} , o número de causas latentes para cada tempo é gerado a partir de distribuições Poisson independentes:

$$N_{1i} \sim \text{Poisson}(W_i \rho_{1i}), \quad N_{2i} \sim \text{Poisson}(W_i \rho_{2i})$$

4.4.1.5 Etapa 5: Geração dos Tempos de Promoção

Para cada tempo $j = 1, 2$ e para cada indivíduo i , procede-se da seguinte forma:

- Se $N_{ji} = 0$, o indivíduo é considerado **imune** para o evento j , e define-se $T_{ji} = \infty$.
- Se $N_{ji} > 0$, geram-se N_{ji} tempos de promoção independentes a partir de uma distribuição Weibull com parâmetros:

$$Z_{j1}, \dots, Z_{jN_{ji}} \stackrel{iid}{\sim} \text{Weibull}(\alpha_{0j}, \text{escala} = e^{-\lambda_{0j}/\alpha_{0j}})$$

A parametrização da escala como $\exp(-\lambda_{0j}/\alpha_{0j})$ segue a formulação do modelo Stan, em que λ_{0j} aparece na forma log-linear.

- O tempo observado para o evento j é então o mínimo entre estes tempos de promoção:

$$T_{ji} = \min\{Z_{j1}, \dots, Z_{jN_{ji}}\}$$

4.4.1.6 Etapa 6: Aplicação da Censura

Para a censura aleatória no estudo, geram-se tempos de censura independentes para cada indivíduo e cada tempo a partir de distribuições uniformes:

$$C_{1i} \sim \text{Uniforme}(0, \tau_{u1}), \quad C_{2i} \sim \text{Uniforme}(0, \tau_{u2})$$

No presente estudo, adotou-se $\tau_{u1} = \tau_{u2} = 3$, valor escolhido após experimentos preliminares para garantir taxas de censura realistas, entre 20% e 40% dependendo dos parâmetros.

4.4.1.7 Etapa 7: Determinação dos Tempos Observados e Indicadores de Falha

Os tempos observados são o mínimo entre o tempo do evento e o tempo de censura:

$$y_{1i} = \min(T_{1i}, C_{1i}), \quad y_{2i} = \min(T_{2i}, C_{2i})$$

Os indicadores de falha são definidos como:

$$\delta_{1i} = \mathbb{I}(T_{1i} \leq C_{1i}), \quad \delta_{2i} = \mathbb{I}(T_{2i} \leq C_{2i})$$

Por outro lado, a especificação das distribuições a priori é apresentada na Tabela 8

Tabela 8 – Distribuições a priori especificadas no modelo

Parâmetro	Distribuição a Priori
$\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$	N(0, 3)
α_{01}, α_{02}	Gamma(1, 1)
$\lambda_{01}, \lambda_{02}$	N(0, 5)
γ	Beta(2, 2)
σ	Gamma(1, 1)

As prioris atribuídas aos coeficientes de regressão $\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}$ foram distribuições Normais N(0, 3), consideradas vagas, com variância relativamente grande, permitindo que a informação dos dados tenha predominância na estimação dos efeitos das covariáveis. Para os parâmetros de forma da Weibull, α_{01} e α_{02} , foi adotada a distribuição Gamma(1, 1), adequada por possuir suporte positivo, média igual a 1 e variância igual a 1, características compatíveis com parâmetros de forma. Os parâmetros λ_{01} e λ_{02} receberam prioris N(0, 5), também de natureza pouco informativa na escala log-linear, permitindo ampla variação nos valores possíveis da escala da distribuição Weibull. O parâmetro γ , associado à distribuição PVF, foi modelado com Beta(2, 2), definida no intervalo (0, 1), sendo simétrica em torno de 0,5.

E, o parâmetro de precisão σ recebeu distribuição Gamma(1, 1), escolha comum para parâmetros positivos de escala, mantendo média 1 e variância 1, refletindo uma especificação pouco informativa. Para ilustrar o funcionamento do algoritmo, considere o cenário como é

apresentado na Tabela 9 em que são apresentados as estimativas da média, DP, viés e REQM dos parâmetros do modelo bivariado PVF Weibull com a presença de covariáveis

Tabela 9 – Resultados das estimativas para diferentes tamanhos amostrais, para 200 réplicas e censura aproximada de 42%.

Parâmetro	Verdadeiro	Média	DP	Viés	REQM
n = 200					
β_{11}	1,000	1,215	0,308	0,215	0,373
β_{12}	-2,000	-2,079	0,268	-0,079	0,277
β_{21}	1,500	1,640	0,257	0,140	0,291
β_{22}	-1,000	-1,012	0,231	-0,012	0,229
α_{01}	1,500	1,516	0,157	0,016	0,156
λ_{01}	-1,000	-1,230	0,346	-0,230	0,413
α_{02}	3,000	3,110	0,309	0,110	0,325
λ_{02}	-2,000	-2,176	0,242	-0,176	0,297
γ	0,500	0,435	0,048	-0,065	0,080
σ	1,000	1,000	0,024	0,000	0,023
n = 500					
β_{11}	1,000	1,172	0,265	0,172	0,313
β_{12}	-2,000	-2,093	0,205	-0,093	0,223
β_{21}	1,500	1,622	0,228	0,122	0,257
β_{22}	-1,000	-1,014	0,159	-0,014	0,158
α_{01}	1,500	1,550	0,107	0,050	0,117
λ_{01}	-1,000	-1,150	0,294	-0,150	0,328
α_{02}	3,000	2,998	0,167	-0,002	0,165
λ_{02}	-2,000	-2,118	0,237	-0,118	0,263
γ	0,500	0,448	0,058	-0,052	0,078
σ	1,000	1,009	0,030	0,009	0,031
n = 1000					
β_{11}	1,000	1,140	0,225	0,140	0,263
β_{12}	-2,000	-2,039	0,154	-0,039	0,158
β_{21}	1,500	1,599	0,162	0,099	0,189
β_{22}	-1,000	-1,002	0,103	-0,002	0,102
α_{01}	1,500	1,514	0,093	0,014	0,093
λ_{01}	-1,000	-1,134	0,257	-0,134	0,287
α_{02}	3,000	3,000	0,137	0,000	0,135
λ_{02}	-2,000	-2,100	0,164	-0,100	0,191
γ	0,500	0,463	0,054	-0,037	0,065
σ	1,000	1,017	0,026	0,017	0,031

Segue da Tabela 9, que a medida que n aumenta, observa-se redução sistemática do viés, do desvio padrão e do REQM para todos os parâmetros, indicando que os estimadores são consistentes. Para $n = 1000$, a maioria dos parâmetros apresenta viés inferior a 0,15 e REQM abaixo de 0,30, demonstrando adequada precisão. Observa-se também:

- Os parâmetros σ e α_{02} foram estimados com excelente precisão, apresentando viés próximo de zero mesmo em amostras pequenas. O parâmetro λ_{01} mostrou-se o mais desafiador, com maior viés e REQM em todos os cenários. Enquanto o parâmetro γ , que controla a heterogeneidade, apresentou leve subestimativa, porém com REQM reduzido.

4.5 Aplicação: Dados de Churn de Clientes Brasileiros

Para demonstrar a utilidade do modelo, analisamos dados de *churn* (é a métrica que mede a taxa de cancelamento ou perda de clientes/assinantes em um período) de clientes brasileiros no setor financeiro. Este conjunto de dados inclui os tempos até o cancelamento dos produtos 1 e 2 (T_1 e T_2) provenientes de um grupo de 900 clientes, com taxas de censura de aproximadamente 52% e 73% para os produtos 1 e 2, respectivamente. Considerando o papel crítico que o *churn* desempenha na retenção de clientes e na sustentabilidade dos negócios no setor financeiro, nosso estudo tem como objetivos principais: primeiro, investigar possíveis associações entre os tempos de cancelamento dos clientes que utilizam os produtos 1 e 2; adicionalmente, examinar diferenças na proporção de clientes leais (fração de cura) entre esses produtos; e, por fim, avaliar como as covariáveis influenciam a fração de cura.

Para a análise, definimos como variáveis resposta os tempos até o abandono dos produtos 1 e 2. Ademais, coletamos variáveis demográficas de cada cliente: t_{kj} representa o tempo observado (em anos) para o k -ésimo produto, onde $k = 1, 2$; x_{1j} denota o gênero (0 = feminino (51%), 1 = masculino (49%)); e x_{2j} indica o nível de renda (1 = renda < 3.000, 2 = 3.000 ≤ renda < 7.000, 3 = renda ≥ 7.000), em que $j = 1, \dots, 900$.

A Figura 9 apresenta as estimativas da função de risco acumulado de Kaplan-Meier (K-M) para os tempos de *churn* de ambos os produtos (T_1 no painel direito e T_2 no painel esquerdo), estratificadas por gênero. A análise revela que o risco de uma mulher abandonar ambos os produtos é inferior ao dos homens. Ademais, a estimativa da função de risco acumulado é limitada superiormente, indicando a presença de uma proporção de clientes leais aos produtos.

Dadas as características do conjunto de dados, modelos de sobrevivência com longa duração mostram-se apropriados para exploração e análise adicionais.

Compreender a dinâmica do *churn* no setor financeiro é fundamental para a tomada de decisões informadas e para a manutenção da satisfação e lealdade dos clientes. O foco deste estudo não se limita à identificação de padrões de *churn*, mas também à modelagem desses padrões por meio da distribuição Weibull, com o objetivo de prever o comportamento futuro de cancelamento e subsidiar decisões estratégicas no âmbito empresarial.

Para este conjunto de dados, ajustamos os modelos conforme descrito na seção 4.2, incluindo todas as covariáveis no parâmetro ρ_{kj} , para $k = 1, 2$ e $j = 1, \dots, 900$, isto é,

$$\log(\rho_{kj}) = \beta_{k,\text{intercepto}} + \beta_{k,\text{gênero}_1} x_{j,\text{gênero}_1} + \beta_{k,\text{renda}_2} x_{j,\text{renda}_2} + \beta_{k,\text{renda}_3} x_{j,\text{renda}_3}, \quad (4.25)$$

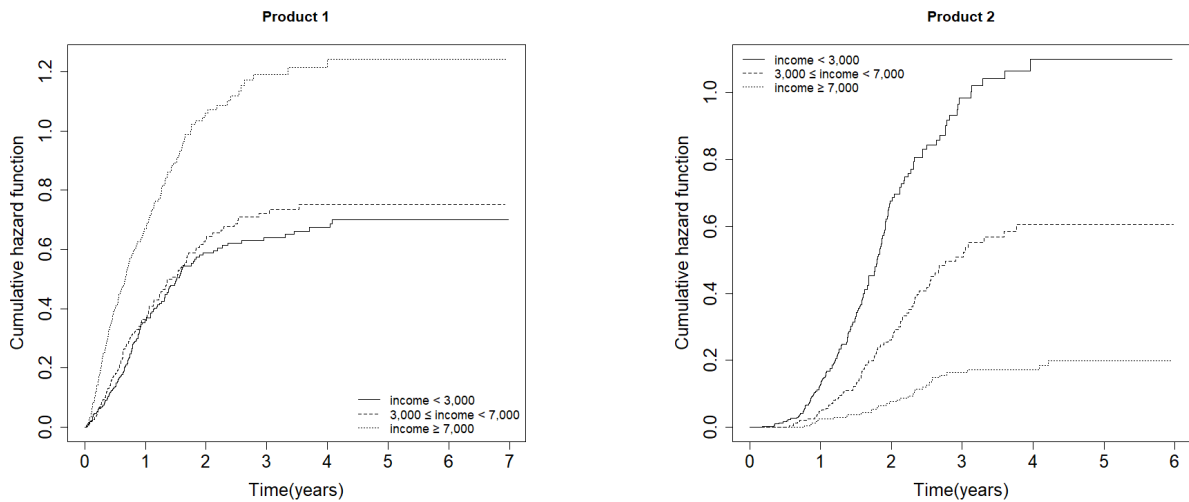


Figura 9 – Estimativas da função de risco acumulado, obtidas pelo método de Kaplan-Meier, dos tempos de churn dos Produtos 1 e 2.

em que os efeitos das covariáveis (gênero e renda) são definidos por meio de variáveis indicadoras (*dummies*). Por exemplo, para a covariável renda, utilizamos duas variáveis indicadoras, isto é,

$$x_{j,renda_2} = \begin{cases} 1, & \text{se } 3.000 \leq \text{renda} < 7.000, \\ 0, & \text{caso contrário,} \end{cases}$$

$$x_{j,renda_3} = \begin{cases} 1, & \text{se } \text{renda} \geq 7.000, \\ 0, & \text{caso contrário,} \end{cases}$$

$$x_{j,g\acute{e}nero_1} = \begin{cases} 1, & \text{se } \text{g\acute{e}nero} \text{ \u00e9 masculino,} \\ 0, & \text{caso contr\u00e1rio.} \end{cases}$$

A infer\ecencia Bayesiana foi realizada por meio do m\ecetodo de Monte Carlo Hamiltoniano (HMC), com quatro cadeias independentes e 4.000 itera\ec\c0es por par\u00e2metro. As primeiras 2.000 itera\ec\c0es de cada cadeia foram descartadas como *burn-in*. Para o modelo de regress\u00e3o PVF-Weibull, foi utilizada a distribui\ec\u00e7\u00e3o a priori especificada na Equa\ec\u00e7\u00e3o (4.23). Para os modelos de regress\u00e3o Gamma-Weibull e IG-Weibull, adotamos as seguintes distribui\ec\u00e7\u00e3es a priori: $\theta \sim N(0, 1)$ truncada em 0, e, para os demais par\u00e2metros, as mesmas especificadas no modelo de regress\u00e3o PVF-Weibull. A converg\ecancia das cadeias foi avaliada com base nos crit\ec3rios propostos por Cowles e Carlin (1996). A Tabela 10 apresenta os resumos a posteriori dos modelos de regress\u00e3o. Adicionalmente, para comparar os modelos propostos, estimamos a verossimilhan\ec7a marginal utilizando *bridge sampling* (Gronau, Singmann e Wagenmakers (2020)). Essa quantidade \ec9 fundamental para a compara\ec\u00e7\u00e3o Bayesiana de modelos e para o c\u00e1lculo das probabilidades a posteriori dos modelos. A Tabela 11 apresenta essas probabilidades. Os crit\ec3rios indicam que o modelo de regress\u00e3o bivariado com fragilidade PVF apresenta o melhor ajuste aos dados.

Assim, consideramos o modelo de regressão bivariado com fragilidade PVF como o modelo mais adequado para a análise dos dados de *churn* de clientes.

Tabela 10 – Médias posteriores, medianas e intervalos HPD de 95% para os parâmetros do modelo de regressão bivariada para o conjunto de dados do banco brasileiro, considerando os modelos de fragilidade Gamma, IG, Positiva Estável e PVF.

Tiempo	Parâmetros	Gamma				Inversa Gaussiana			
		Média	DP	L	U	Média	DP	L	U
T_1	α_1	1.586	0.064	1.462	1.712	1.595	0.064	1.471	1.723
	λ_1	-0.883	0.106	-1.097	-0.689	-0.843	0.094	-1.036	-0.665
	$\beta_{1,Inter.}$	0.935	0.160	0.627	1.253	0.993	0.170	0.673	1.339
	$\beta_{1,g\acute{e}nero_1}$	-1.962	0.154	-2.264	-1.664	-1.933	0.147	-2.215	-1.647
	$\beta_{1,renda_2}$	0.137	0.164	-0.180	0.462	0.159	0.168	-0.168	0.487
	$\beta_{1,renda_3}$	1.191	0.164	0.878	1.526	1.112	0.160	0.812	1.429
T_2	α_2	3.062	0.143	2.787	3.339	3.116	0.147	2.827	3.408
	λ_2	-2.858	0.161	-3.191	-2.546	-2.912	0.162	-3.235	-2.585
	$\beta_{2,Inter.}$	1.252	0.172	0.929	1.595	1.347	0.182	0.999	1.703
	$\beta_{2,g\acute{e}nero_1}$	-1.017	0.162	-1.333	-0.706	-1.047	0.171	-1.394	-0.714
	$\beta_{2,renda_2}$	-0.984	0.180	-1.347	-0.635	-0.947	0.177	-1.287	-0.611
	$\beta_{2,renda_3}$	-2.385	0.228	-2.858	-1.937	-2.439	0.230	-2.907	-1.996
	σ	1.052	0.168	0.782	1.430	0.300	0.071	0.186	0.463
		Positiva Estável				PVF			
		Média	DP	L	U	Média	DP	L	U
T_1	α_1	1.577	0.065	1.449	1.706	1.597	0.067	1.469	1.728
	λ_1	-0.724	0.088	-0.902	-0.557	-0.876	0.104	-1.082	-0.675
	$\beta_{1,Inter.}$	0.292	0.118	0.058	0.514	0.993	0.173	0.661	1.327
	$\beta_{1,g\acute{e}nero_1}$	-1.733	0.140	-2.016	-1.463	-1.970	0.157	-2.276	-1.668
	$\beta_{1,renda_2}$	0.174	0.155	-0.136	0.472	0.149	0.163	-0.167	0.459
	$\beta_{1,renda_3}$	0.974	0.145	0.694	1.260	1.158	0.164	0.834	1.484
T_2	α_2	3.146	0.155	2.832	3.450	3.098	0.146	2.829	3.404
	λ_2	-2.880	0.169	-3.216	-2.555	-2.900	0.158	-3.219	-2.609
	$\beta_{2,Inter.}$	0.662	0.130	0.406	0.916	1.327	0.184	0.995	1.685
	$\beta_{2,g\acute{e}nero_1}$	-0.961	0.157	-1.274	-0.664	-1.050	0.168	-1.378	-0.712
	$\beta_{2,renda_2}$	-0.845	0.174	-1.192	-0.506	-0.970	0.181	-1.341	-0.626
	$\beta_{2,renda_3}$	-2.369	0.244	-2.861	-1.898	-2.427	0.247	-2.903	-1.939
	γ	0.800	0.032	0.739	0.864	0.299	0.118	0.079	0.523
σ					0.567	0.203	0.240	1.004	

Tabela 11 – Probabilidades posteriores dos modelos a partir das verossimilhanças marginais.

	Modelo de Regressão Bivariada			
	PVF	Gama	Inversa Gaussiana	Positiva Estável
Probabilidade	0.473	0.305	0.221	1.184×10^{-9}

Para avaliar a presença de observações influentes nas distribuições a posteriori dos parâmetros do modelo de regressão bivariado com fragilidade PVF, calculamos as medidas de divergência- ϕ discutidas na subsubseção 2.5.4, com base nas amostras a posteriori. Essas medidas quantificam a alteração na distribuição a posteriori quando cada observação é removida, permitindo identificar possíveis pontos influentes.

A Figura 10 apresenta os valores dos índices correspondentes às quatro medidas de divergência- ϕ . Os resultados mostram que todos os valores permanecem em um intervalo estreito e abaixo de limiares comumente sugeridos na literatura para influência, como os pontos de corte 0,5 ou 0,7 propostos para diagnósticos baseados em divergência. Isso indica que nenhuma observação exerce impacto substancial na distribuição a posteriori dos parâmetros do modelo de regressão.

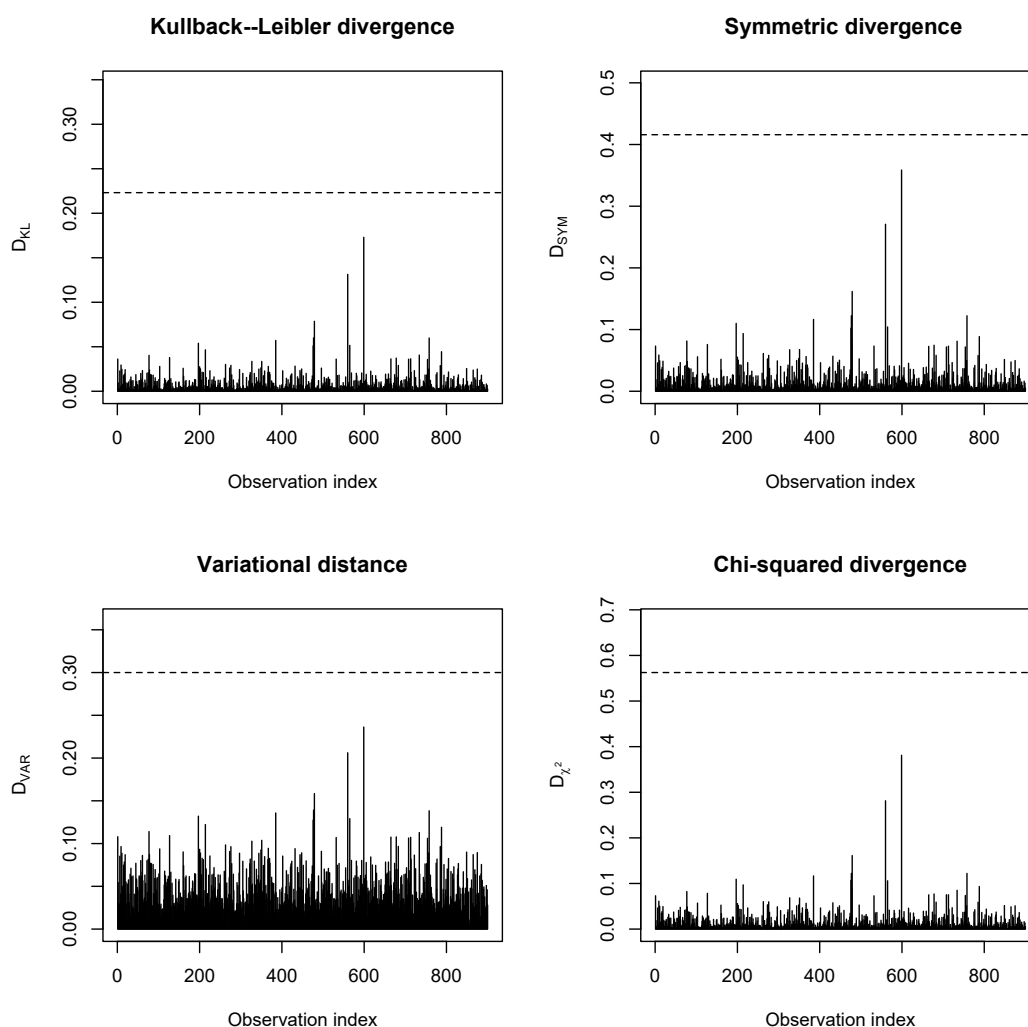


Figura 10 – Gráfico de índices das medidas de divergência ϕ relacionadas aos dados de Churn de clientes brasileiros.

Portanto, a inferência baseada no modelo de regressão bivariado com fragilidade PVF mostra-se robusta em relação a observações individuais, não havendo evidências que justifiquem a exclusão ou reponderação de quaisquer observações na análise.

A Tabela 10 apresenta as estimativas a posteriori de $\phi = (\beta_1^T, \beta_2^T)$ e dos parâmetros (α_k, λ_k) , γ e σ , para $k = 1, 2$. As estimativas de Bayes indicam que, ao nível de significância de 5%, todas as covariáveis exercem efeito significativo sobre a proporção de clientes leais, tanto para o Produto 1 quanto para o Produto 2.

Notadamente, as estimativas de Bayes dos coeficientes de regressão associados à covariável gênero, $\widehat{\beta}_{k,\text{gênero}_1} < 0$ ($k = 1, 2$), sugerem que, independentemente do nível de renda fixado, a proporção de mulheres leais é inferior à de homens leais para ambos os produtos, implicando maior risco de abandono dos Produtos 1 e 2 por parte das mulheres.

Por outro lado, a covariável renda apresenta efeitos distintos sobre a proporção de clientes leais aos Produtos 1 e 2. Especificamente, fixando o gênero do cliente, a proporção de clientes leais ao Produto 1 com renda inferior a R\$3.000,00 é semelhante daqueles com renda entre R\$3.000,00 e R\$7.000,00; entretanto, essas proporções diferem para o Produto 2. Conforme indicado por $\widehat{\beta}_{2,\text{renda}_2} < 0$, a proporção de clientes leais com renda inferior a R\$3.000,00 é menor do que a daqueles com renda entre R\$3.000,00 e R\$7.000,00; além disso, $\widehat{\beta}_{2,\text{renda}_3} > 0$ indica que clientes leais ao Produto 2 com renda entre R\$3.000,00 e R\$7.000,00 são mais prevalentes do que aqueles com renda superior a R\$7.000,00. Essa relação se inverte para o Produto 1, uma vez que $\widehat{\beta}_{1,\text{renda}_3} < 0$.

Também estimamos, sob a abordagem Bayesiana, a probabilidade de que os clientes permaneçam leais ao k -ésimo produto ($k = 1, 2$) após $t > 0$ anos do período de acompanhamento. Essa probabilidade é dada por

$$\Pr_k(t) = \Pr(Z_k = 0 \mid T_k > t) = \left[\frac{\exp \left\{ \left(1 + \frac{\rho_k(1 - \exp\{-e^{\alpha_k t} \beta_k\})}{\sigma} \right)^\gamma \right\}}{\exp \left\{ \left(1 + \frac{\rho_k}{\sigma} \right)^\gamma \right\}} \right]^{\frac{\sigma}{\gamma}}, \quad k = 1, 2. \quad (4.26)$$

em que ρ_k é dado por (4.25). Observe que, quando $\pi_k(0) = p_{0k}$ ($k = 1, 2$), obtém-se a taxa de clientes leais ao produto 1 ou 2 após um período de observação.

Com base nesse resultado, estimamos a proporção de clientes leais para seis clientes hipotéticos A, B, C, D, E e F , após 3 anos de período de acompanhamento, considerando os valores das covariáveis apresentados na Tabela 12.

A Tabela 13 apresenta as estimativas de Bayes da proporção de clientes leais a ambos os produtos, p_{00} , ao produto 1, p_{01} , e ao produto 2, p_{02} , após o período de acompanhamento. Por exemplo, para clientes com fatores de risco idênticos aos do cliente A, a probabilidade de não abandonar ambos os produtos é $p_{00} = 0,207$ e, para os produtos 1 e 2, as probabilidades são $p_{01} = 0,387$ e $p_{02} = 0,312$, respectivamente. A Figura 12 apresenta a distribuição a posteriori dessas probabilidades.

Por outro lado, a Tabela 12 apresenta as probabilidades de esses clientes permanecerem com os produtos 1 e 2 após 3 anos de acompanhamento. Para o cliente A, essas probabilidades são $\Pr_1(3) = 0,942$ e $\Pr_2(3) = 0,862$, respectivamente, enquanto a probabilidade de permanecer com ambos os produtos é $\Pr_{00}(3) = 0,869$.

A partir da Tabela 13, observa-se que a proporção de clientes que permanecem com ambos os produtos, p_{00} , é menor para clientes do sexo feminino do que para clientes do sexo

masculino, em qualquer nível de renda. Essa relação entre as proporções de clientes leais (entre mulheres e homens) também se mantém quando considerados separadamente os produtos 1 e 2.

Tabela 12 – Estimativa Bayesiana da probabilidade de clientes fiéis aos produtos 1 e 2 após 3 anos do período de observação.

Consumidor	Renda	Gênero	Probabilidade de clientes fiéis		
			Produto 1 e 2 Pr(3)	Produto 1 Pr ₁ (3)	Produto 2 Pr ₂ (3)
A	* < 3,000	Femenino	0.869 (0.806 , 0.920)	0.942 (0.896, 0.976)	0.862 (0.785 , 0.925)
B		Masculino	0.904 (0.855 , 0.946)	0.938 (0.890 , 0.973)	0.911 (0.851 , 0.957)
C	≥ 3,000 e < 7,000	Femenino	0.908 (0.849 , 0.952)	0.913 (0.852 , 0.957)	0.966 (0.936 , 0.987)
D		Masculino	0.913 (0.860 , 0.958)	0.982 (0.956 0.987)	0.914 (0.856 , 0.960)
E	> 7,000	Femenino	0.946 (0.910 , 0.976)	0.980 (0.960 , 0.993)	0.954 (0.917 , 0.981)
F		Masculino	0.954 (0.918 , 0.982)	0.961 (0.926 0.985)	0.986 (0.972 , 0.995)

Fonte: Elaborada pelo autor.

Tabela 13 – Estimativa Bayesiana da probabilidade de clientes fiéis aos produtos 1 e 2.

Consumidor	Renda	Gênero	Probabilidade de clientes fiéis		
			Produtos 1 e 2 p_{00}	Produto 1 p_{01}	Produto 2 p_{02}
A	* < 3,000	Femenino	0.207 (0.095 , 0.396)	0.387 (0.231 , 0.615)	0.312 (0.165 , 0.528)
B		Masculino	0.273 (0.140 , 0.485)	0.354 (0.202 , 0.585)	0.540 (0.368 , 0.741)
C	≥ 3,000 e < 7,000	Femenino	0.156 (0.067 , 0.332)	0.162 (0.071 , 0.342)	0.823 (0.703 , 0.928)
D		Masculino	0.493 (0.318 , 0.704)	0.807 (0.691 , 0.911)	0.554 (0.375 0.754)
E	> 7,000	Femenino	0.633 (0.469 , 0.805)	0.84 (0.662 , 0.898)	0.760 (0.621 , 0.886)
F		Masculino	0.553 (0.383 , 0.759)	0.576 (0.409 , 0.773)	0.927 (0.866 , 0.974)

Fonte: Elaborada pelo autor.

Nota-se ainda que a fração de clientes que não cancelaram o produto 2 aumenta com o aumento da renda, tanto para homens quanto para mulheres. Contudo, para o produto 1, essa proporção diminui à medida que a renda aumenta.

Por fim, a Figura 11 apresenta as distribuições marginais a posteriori da função de risco acumulado, estratificadas por nível de renda em cada categoria de gênero, sendo o painel

da direita referente às clientes do sexo feminino e o painel da esquerda aos clientes do sexo masculino, para ambos os produtos 1 e 2. A figura indica que, independentemente da renda, o risco de abandono dos produtos 1 ou 2 é maior entre clientes do sexo feminino do que entre clientes do sexo masculino. Observa-se também que, entre as mulheres, o risco de abandono do produto 1 aumenta com o nível de renda, enquanto essa relação se inverte para o produto 2. Entre os homens, tais diferenças mostram-se menos pronunciadas.

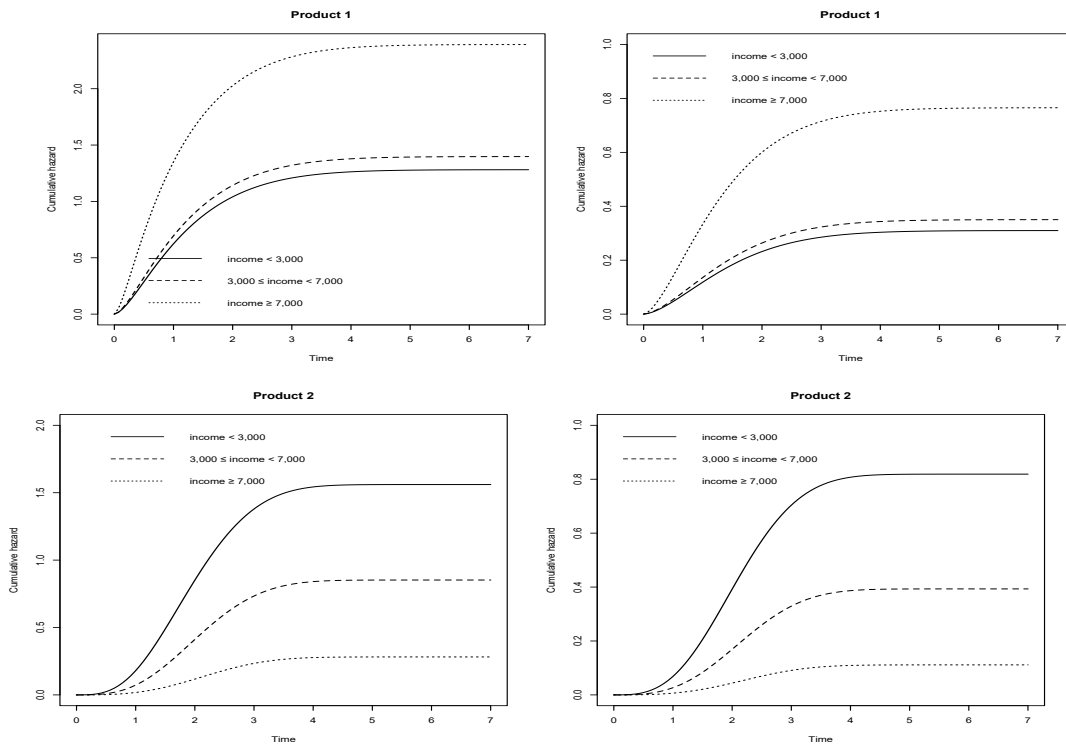


Figura 11 – A função de risco acumulado marginal a posteriori de T_1 e T_2 , estratificada pela renda do cliente para cada nível de gênero: (painel direito) feminino e (painel esquerdo) masculino.

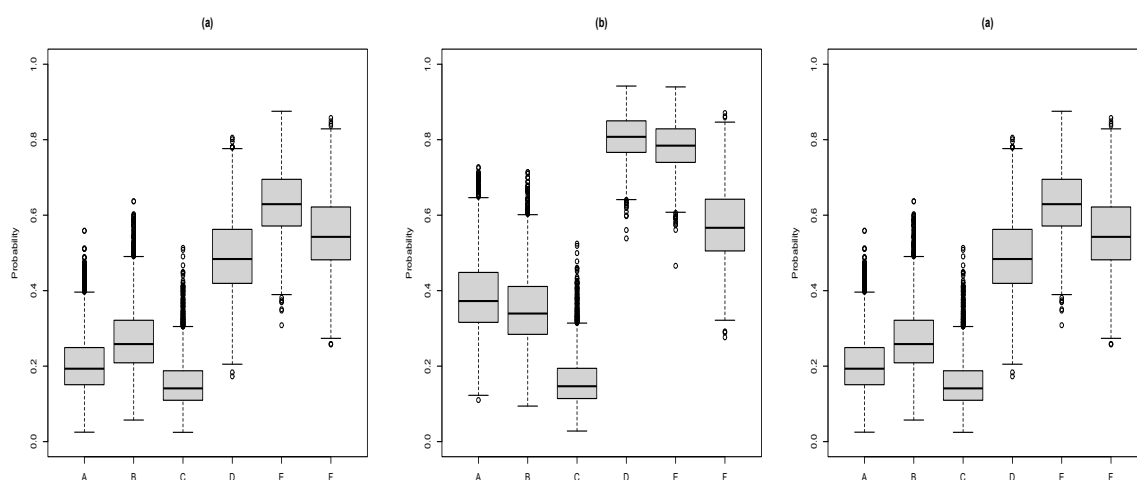


Figura 12 – Boxplot da probabilidade a posteriori de clientes fiéis aos produtos 1 e 2: (a) ao produto 1, (b) ao produto 1 e (c) ao produto 2, para os clientes A, B, C, D, E e F.

4.6 Conclusões

Este trabalho propôs um novo modelo de sobrevivência bivariado com fração de cura, fundamentado na família de distribuições PVF para o componente de fragilidade. O modelo desenvolvido generaliza diversos modelos existentes na literatura, incluindo os propostos por [Chen, Ibrahim e Sinha \(2002\)](#) Chen et al. (2002) e [Cancho et al. \(2022\)](#). O modelo permitiu derivar expressões analíticas para funções de sobrevivência conjunta e marginais, funções de risco conjunta e marginais, e uma medida local de associação $\vartheta^*(t_1, t_2)$.

Também foi desenvolvida uma abordagem completa para inferência Bayesiana utilizando métodos de Monte Carlo via Cadeias de Markov (MCMC). Especificamente, o uso do Hamiltonian Monte Carlo (HMC) implementado em Stan mostrou-se eficaz para amostrar da distribuição posterior, mesmo em um modelo com estrutura complexa e múltiplos parâmetros. O modelo foi estendido para incluir efeitos de covariáveis através do parâmetro ρ_{kj} , utilizando uma função de ligação logarítmica. Isso permite avaliar como características observáveis dos indivíduos (como gênero e renda) afetam tanto a probabilidade de cura quanto a dinâmica do risco ao longo do tempo.

A parametrização adotada confere interpretações claras aos parâmetros do modelo, para isto foi apresentado uma aplicação do modelo a um conjunto de dados de churn de clientes brasileiros do setor financeiro demonstrou sua utilidade prática.

Em conclusão, este trabalho apresentou uma contribuição original e relevante para a área de análise de sobrevivência multivariada com fração de cura. O modelo proposto combina flexibilidade, interpretabilidade e capacidade de incorporar covariáveis, mostrando-se uma ferramenta valiosa tanto para pesquisadores quanto para profissionais que lidam com dados de tempo até evento na presença de uma fração imune. A aplicação a dados reais de churn bancário

demonstrou seu potencial para gerar insights acionáveis para a gestão de relacionamento com clientes, contribuindo para decisões mais informadas em estratégias de retenção e precificação.

CONSIDERAÇÕES FINAIS

5.1 Conclusões

Este trabalho propôs uma nova classe de modelos de sobrevivência denominada Modelo de Cura com Mistura Induzido por PVF. A principal contribuição é a unificação, em uma única estrutura paramétrica, de dois fenômenos fundamentais em análise de sobrevivência: a heterogeneidade não observada (via fragilidade) e a presença de uma fração de curados (via modelos de mistura).

O modelo proposto apresenta cinco características fundamentais: (i) a família PVF é flexível, englobando as distribuições Gama, Inversa Gaussiana e Positiva Estável como casos particulares, o que permite modelar diferentes níveis de heterogeneidade não observada; (ii) foram estabelecidas propriedades matemáticas importantes, incluindo a existência de momentos e uma expressão fechada para os momentos da distribuição PVF-ID-Weibull; (iii) a estrutura de regressão dupla permite que as covariáveis atuem separadamente na probabilidade de cura (β_1) e no risco dos suscetíveis (β_2), proporcionando interpretações; (iv) a inferência Bayesiana via HMC com prioris cuidadosamente escolhidas Beta(4,3) para γ e Normal(0,1) para o intercepto da cura) garante estabilidade computacional; e (v) o comportamento assintótico da função de risco foi caracterizado: pode ser decrescente, constante ou crescente dependendo da relação entre γ e o parâmetro de forma da distribuição basal. Em resumo, o modelo proposto no Capítulo 3 oferece uma ferramenta flexível e teoricamente fundamentada para análise de dados de sobrevivência com cura e heterogeneidade, contribuindo para a literatura e para aplicações práticas.

Por outra parte, no contexto de modelos bivariados, a seleção de uma distribuição basal compatível com estruturas de tempo de vida prolongado e capaz de proporcionar um ajuste adequado aos dados constitui um desafio. Nesse sentido, a distribuição Weibull destaca-se como uma alternativa particularmente flexível, uma vez que sua parametrização permite representar

diferentes formatos da função de risco, tornando-a amplamente adequada para a modelagem de dados de sobrevivência. No que se refere ao processo inferencial, a abordagem baseada em métodos bayesianos mostrou-se apropriada, uma vez que os estudos de simulação evidenciaram propriedades satisfatórias para os estimadores de Bayes. Ademais, a implementação do método *Hamiltonian Monte Carlo* (HMC) por meio do *software* RStan demonstrou ser eficiente e promissora para a análise de dados. Essa ferramenta oferece uma linguagem relativamente acessível para a implementação de algoritmos, além de contar com amplo suporte teórico e prático na literatura.

5.2 Sugestões para pesquisas futuras

- Avaliar o desempenho do modelo PVF-ID com outras funções de risco basal;
- Avaliar o modelo PVF-ID para cenários com diferentes taxas de censura;
- Estender a abordagem para dados multivariados;
- Propor novos modelos de sobrevivência induzidos por fragilidade utilizando outras famílias de distribuições, além das famílias PVF usadas nesta tese.

REFERÊNCIAS

AALEN, O. O. Modelling heterogeneity in survival analysis by the compound poisson distribution. **The Annals of Applied Probability**, JSTOR, p. 951–972, 1992. Citado nas páginas [31](#), [34](#) e [35](#).

AHMED, F. E.; VOS, P. W.; HOLBERT, D. Modeling survival in colon cancer: a methodological review. **Molecular cancer**, Springer, v. 6, n. 1, p. 15, 2007. Citado na página [31](#).

AMORIM, W. N. de. **Verossimilhança hierárquica em modelos de fragilidade**. Tese (Doutorado) — Universidade de São Paulo, 2014. Citado na página [25](#).

ANDERSEN, P. K.; PERME, M. P.; HOUWELINGEN, H. C. van; COOK, R. J.; JOLY, P.; MARTINUSSEN, T.; TAYLOR, J. M.; ABRAHAMOWICZ, M.; THERNEAU, T. M. Analysis of time-to-event for observational studies: guidance to the use of intensity models. **Statistics in medicine**, Wiley Online Library, v. 40, n. 1, p. 185–211, 2021. Citado na página [31](#).

BALAN, T. A.; PUTTER, H. A tutorial on frailty models. **Statistical methods in medical research**, SAGE Publications Sage UK: London, England, v. 29, n. 11, p. 3424–3454, 2020. Citado nas páginas [31](#), [33](#), [34](#) e [35](#).

BALKA, J.; DESMOND, A. F.; MCNICHOLAS, P. D. Review and implementation of cure models based on first hitting times for wiener processes. **Lifetime data analysis**, Springer, v. 15, n. 2, p. 147–176, 2009. Citado na página [53](#).

BANBETA, A.; SEYOUM, D.; BELACHEW, T.; BIRLIE, B.; GETACHEW, Y. Modeling time-to-cure from severe acute malnutrition: application of various parametric frailty models. **Archives of Public Health**, Springer, v. 73, n. 1, p. 6, 2015. Citado na página [24](#).

BARRIGA, G. D.; CANCHO, V. G.; GARIBAY, D. V.; CORDEIRO, G. M.; ORTEGA, E. M. A new survival model with surviving fraction: An application to colorectal cancer data. **Statistical methods in medical research**, SAGE Publications Sage UK: London, England, v. 28, n. 9, p. 2665–2680, 2019. Citado na página [77](#).

BEDIA, E. C. Modelos de sobrevivência bivariados induzidos por fragilidade. Universidade Federal de São Carlos, 2022. Citado nas páginas [25](#), [31](#), [38](#), [39](#) e [54](#).

BERKSON, J.; GAGE, R. P. Survival curve for cancer patients following treatment. **Journal of the American Statistical Association**, Taylor & Francis, v. 47, n. 259, p. 501–515, 1952. Citado nas páginas [25](#), [53](#) e [77](#).

BORGES, P. Em algorithm-based likelihood estimation for a generalized gompertz regression model in presence of survival data with long-term survivors: an application to uterine cervical cancer data. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 87, n. 9, p. 1712–1722, 2017. Citado na página [54](#).

- BRITO, A. D. L.; JÚNIOR, S. F. X.; MENDONÇA, E. B. D.; XAVIER, É. F. M.; SANTOS, T. T. D. M.; OLIVEIRA, T. A. de. Ajuste de modelos de fragilidade e riscos proporcionais aplicados a dados de retinopatia diabética. **Research, Society and Development**, v. 9, n. 8, p. e478985691–e478985691, 2020. Citado na página 24.
- BUNYATISAI, W.; PRASITWATTANASEREE, S.; INGSRISAWANG, L. Assessing frailty survival models in describing variations caused by unobserved covariates. **Chiang Mai J. Sci**, v. 44, p. 1191–1200, 2017. Citado na página 24.
- CALSAVARA, V. F.; RODRIGUES, A. S.; TOMAZELLA, V. L. D.; CASTRO, M. de. Frailty models power variance function with cure fraction and latent risk factors negative binomial. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 46, n. 19, p. 9763–9776, 2017. Citado nas páginas 25 e 39.
- CANCHO, V. G.; BARRIGA, G. D.; CORDEIRO, G. M.; ORTEGA, E. M.; SUZUKI, A. K. Bayesian survival model induced by frailty for lifetime with long-term survivors. **Statistica Neerlandica**, Wiley Online Library, v. 75, n. 3, p. 299–323, 2021. Citado nas páginas 24, 25, 36, 69, 77 e 81.
- CANCHO, V. G.; LOUZADA, F.; DEY, D. K.; BARRIGA, G. D. A new lifetime model for multivariate survival data with a surviving fraction. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 86, n. 2, p. 279–292, 2016. Citado na página 78.
- CANCHO, V. G.; RODRIGUES, J.; CASTRO, M. de. A flexible model for survival data with a cure rate: a bayesian approach. **Journal of Applied Statistics**, Taylor & Francis, v. 38, n. 1, p. 57–70, 2011. Citado na página 52.
- CANCHO, V. G.; SUZUKI, A. K.; BARRIGA, G. D.; SANTO, A. P. d. E. A multivariate survival model induced by discrete frailty. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 51, n. 11, p. 6572–6590, 2022. Citado nas páginas 78, 81 e 103.
- CANCHO, V. G.; ZAVALETA, K. E.; MACERA, M. A.; SUZUKI, A. K.; LOUZADA, F. A bayesian cure rate model with dispersion induced by discrete frailty. **Communications for Statistical Applications and Methods**, Korean Statistical Society, v. 25, n. 5, p. 471–488, 2018. Citado nas páginas 24, 36, 48 e 52.
- CHATZILENA, A.; LEEUWEN, E. van; RATMANN, O.; BAGUELIN, M.; DEMIRIS, N. Contemporary statistical inference for infectious disease models using stan. **Epidemics**, Elsevier, v. 29, p. 100367, 2019. Citado na página 27.
- CHEN, M.-H.; IBRAHIM, J. G.; SINHA, D. Bayesian inference for multivariate survival data with a cure fraction. **Journal of Multivariate Analysis**, Elsevier, v. 80, n. 1, p. 101–126, 2002. Citado nas páginas 77, 78, 82, 85 e 103.
- CHEN, P.; ZHANG, J.; ZHANG, R. Estimation of the accelerated failure time frailty model under generalized gamma frailty. **Computational Statistics & Data Analysis**, Elsevier, v. 62, p. 171–180, 2013. Citado na página 24.
- CLAYTON, D. G. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. **Biometrika**, Oxford University Press, v. 65, n. 1, p. 141–151, 1978. Citado nas páginas 78 e 84.

CORDEIRO, G. M.; CANCHO, V. G.; ORTEGA, E. M.; BARRIGA, G. D. A model with long-term survivors: negative binomial birnbaum-saunders. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 45, n. 5, p. 1370–1387, 2016. Citado na página 77.

COWLES, M. K.; CARLIN, B. P. Markov chain monte carlo convergence diagnostics: a comparative review. **Journal of the American statistical Association**, Taylor & Francis, v. 91, n. 434, p. 883–904, 1996. Citado nas páginas 72 e 97.

COX, D. R. Regression models and life-tables. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 34, n. 2, p. 187–202, 1972. Citado na página 32.

DEMARQUI, F. N. Modelo exponencial por partes via modelos partição produto. Universidade Federal de Minas Gerais, 2006. Citado na página 26.

_____. Uma classe mais flexível de modelos semiparamétricos para dados de sobrevivência. Universidade Federal de Minas Gerais, 2010. Citado nas páginas 26 e 44.

DUANE, S.; KENNEDY, A. D.; PENDLETON, B. J.; ROWETH, D. Hybrid monte carlo. **Physics letters B**, Elsevier, v. 195, n. 2, p. 216–222, 1987. Citado na página 45.

DUCHATEAU, L.; JANSSEN, P. Frailty distributions. **The Frailty Model**, Springer New York, p. 117–197, 2008. Citado nas páginas 31 e 78.

FORTES, R. d. S. R. Modelo de sobrevivência birnbaum-saunders com fragilidade espacial. 2020. Citado nas páginas 24 e 36.

GASPARINI, A.; CLEMENTS, M. S.; ABRAMS, K. R.; CROWTHER, M. J. Impact of model misspecification in shared frailty survival models. **Statistics in medicine**, Wiley Online Library, v. 38, n. 23, p. 4477–4502, 2019. Citado na página 25.

GELMAN, A.; LEE, D.; GUO, J. Stan: A probabilistic programming language for bayesian inference and optimization. **Journal of Educational and Behavioral Statistics**, Sage Publications Sage CA: Los Angeles, CA, v. 40, n. 5, p. 530–543, 2015. Citado na página 27.

GIUSSANI, A.; BONETTI, M. Marshall–olkin frailty survival models for bivariate right-censored failure time data. **Journal of Applied Statistics**, Taylor & Francis, 2019. Citado na página 24.

GOMPERTZ, B. Xxiv. on the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. in a letter to francis baily, esq. frs &c. **Philosophical transactions of the Royal Society of London**, The Royal Society London, n. 115, p. 513–583, 1825. Citado na página 54.

GRONAU, Q. F.; SARAFIOGLOU, A.; MATZKE, D.; LY, A.; BOEHM, U.; MARSMAN, M.; LESLIE, D. S.; FORSTER, J. J.; WAGENMAKERS, E.-J.; STEINGROEVER, H. A tutorial on bridge sampling. **Journal of mathematical psychology**, Elsevier, v. 81, p. 80–97, 2017. Citado nas páginas 31, 49, 50, 51 e 52.

GRONAU, Q. F.; SINGMANN, H.; WAGENMAKERS, E.-J. bridgesampling: An r package for estimating normalizing constants. **Journal of Statistical Software**, v. 92, p. 1–29, 2020. Citado nas páginas 31, 48, 49 e 97.

- GRONAU, Q. F.; WAGENMAKERS, E.-J.; HECK, D. W.; MATZKE, D. A simple method for comparing complex models: Bayesian model comparison for hierarchical multinomial processing tree models using warp-iii bridge sampling. **psychometrika**, Springer, v. 84, n. 1, p. 261–284, 2019. Citado nas páginas 31 e 52.
- GURMU, S. E. Assessing survival time of women with cervical cancer using various parametric frailty models: a case study at tikur anbessa specialized hospital, addis ababa, ethiopia. **Annals of Data Science**, Springer, v. 5, n. 4, p. 513–527, 2018. Citado na página 24.
- HARTMANN, M. **Métodos de Monte Carlo Hamiltoniano na inferência Bayesiana não-paramétrica de valores extremos**. Tese (Doutorado) — Universidade de São Paulo, 2015. Citado nas páginas 27, 31, 44, 45, 46 e 48.
- HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. Oxford University Press, 1970. Citado nas páginas 45 e 48.
- HOSMER, D. W. Applied survival analysis; regression modeling of time to event data. New York, John Wiley Sons,, 2002. Citado na página 31.
- HOUGAARD, P. Life table methods for heterogeneous populations: distributions describing the heterogeneity. **Biometrika**, Oxford University Press, v. 71, n. 1, p. 75–83, 1984. Citado nas páginas 24 e 38.
- _____. Survival models for heterogeneous populations derived from stable distributions. **Biometrika**, Oxford University Press, v. 73, n. 2, p. 387–396, 1986. Citado nas páginas 24 e 78.
- _____. **Analysis of multivariate survival data**. [S.l.]: Springer, 2000. v. 564. Citado nas páginas 31, 39, 40, 55, 78 e 79.
- IBRAHIM, J. G.; CHEN, M.-H.; SINHA, D. Bayesian semiparametric models for survival data with a cure fraction. **Biometrics**, Wiley Online Library, v. 57, n. 2, p. 383–388, 2001. Citado nas páginas 26, 31, 41, 44, 48, 49, 51, 52 e 56.
- ISIDRO, M. J.; ISIDRO, R. de A.; BRITO, A. de L.; OLIVEIRA, T. A. de. Modelos de fragilidade aplicados a análise de fatores contribuintes na morte de pacientes portadores de leucemia. **Brazilian Journal of Development**, v. 6, n. 8, p. 54802–54820, 2020. Citado na página 24.
- JIANG, Z.; CARTER, R. Using hamiltonian monte carlo to estimate the log-linear cognitive diagnosis model via stan. **Behavior Research Methods**, Springer, v. 51, n. 2, p. 651–662, 2019. Citado na página 27.
- KOUTRAS, M.; MILIENOS, F. A flexible family of transformation cure rate models. **Statistics in medicine**, Wiley Online Library, v. 36, n. 16, p. 2559–2575, 2017. Citado na página 77.
- LEIMKUHNER, B.; MATTHEWS, C. Molecular dynamics. **Interdisciplinary applied mathematics**, Springer, v. 39, n. 1, 2015. Citado na página 47.
- LIMA, C. M.; TOMAZELLA, V. L.; CAMPELO, J. E.; FILHO, J. L.; JUNIOR, W. B.; JUNIOR, S. C. S. Gamma-gompertz shared frailty model for analysis of the time of stay in an anglo-nubian goat herd. **Small Ruminant Research**, Elsevier, v. 199, p. 106368, 2021. Citado na página 25.

LIMA, L. V. de; PAVINATI, G.; SALA, C.; PALMIERI, I. G. S.; AMARAL, G. C. do; HÖRING, C. F.; GIL, N. L. de M.; MAGNABOSCO, G. T. Hiv in paran : clinical-epidemiological overview, distribution, and incidence by macroregional and regional health. **Acta Scientiarum. Health Sciences**, v. 45, 2023. Citado na p gina 68.

LUO, Y.; JIAO, H. Using the stan program for bayesian item response theory. **Educational and psychological measurement**, SAGE Publications Sage CA: Los Angeles, CA, v. 78, n. 3, p. 384–408, 2018. Citado na p gina 27.

MACERA, M. A. C. Modelagem de dados de sobreviv ncia com eventos recorrentes via fragilidade discreta. 2015. Citado nas p ginas 24 e 36.

MCELREATH, R. **Statistical rethinking: A Bayesian course with examples in R and Stan**. [S.l.]: Chapman and Hall/CRC, 2018. Citado na p gina 89.

MEDEIROS, V. R. D.; BONATTO, L. V. S.; SANTOS, F. R. D.; CHASKO, F. V.; SILVA, L. W. R. D.; MARTINEZ, A. D. C. Perfil epidemiol gico dos casos de hiv/aids no estado do paran  (2013 a 2018). In: **II Congresso de Sa de Coletiva da UFPR**. [S.l.: s.n.], 2020. Citado na p gina 68.

MELLO, J. F. de *et al.* Modelo exponencial por partes para dados de sobreviv ncia com longa dura o. Universidade Federal de Minas Gerais, 2016. Citado na p gina 26.

MENG, X.-L.; SCHILLING, S. Warp bridge sampling. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 11, n. 3, p. 552–586, 2002. Citado nas p ginas 31, 49 e 52.

MENG, X.-L.; WONG, W. H. Simulating ratios of normalizing constants via a simple identity: a theoretical exploration. **Statistica Sinica**, JSTOR, p. 831–860, 1996. Citado nas p ginas 31, 49, 50, 51 e 52.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The journal of chemical physics**, American Institute of Physics, v. 21, n. 6, p. 1087–1092, 1953. Citado nas p ginas 45 e 48.

METROPOLIS, N.; ULAM, S. The monte carlo method. **Journal of the American statistical association**, Taylor & Francis, v. 44, n. 247, p. 335–341, 1949. Citado na p gina 44.

MOLINA, K. R. C. **Modelos de sobreviv ncia induzidos por fragilidade discreta s rie de pot ncia zero-modificada**. Tese (Doutorado) — Universidade de S o Paulo, 2020. Citado na p gina 24.

MONACO, J. V.; GORFINE, M.; HSU, L. General semiparametric shared frailty model: estimation and simulation with frailtysurv. **Journal of statistical software**, v. 86, p. 1–42, 2018. Citado na p gina 25.

NG’OMBE, J. N.; LAMBERT, D. M. Using hamiltonian monte carlo via stan to estimate crop input response functions with stochastic plateaus. **Journal of Agriculture and Food Research**, Elsevier, v. 6, p. 100226, 2021. Citado nas p ginas 27 e 64.

OAKES, D. A model for association in bivariate survival data. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, Oxford University Press, v. 44, n. 3, p. 414–422, 1982. Citado na p gina 84.

- ORTEGA, E. M.; CORDEIRO, G. M.; CAMPELO, A. K.; KATTAN, M. W.; CANCHO, V. G. A power series beta weibull regression model for predicting breast carcinoma. **Statistics in medicine**, Wiley Online Library, v. 34, n. 8, p. 1366–1388, 2015. Citado na página 77.
- PAIXÃO, R. S. **Método Zero-Variance para Monte Carlo Hamiltoniano aplicado a modelos GARCH univariados e multivariados**. Tese (Doutorado) — Universidade de São Paulo, 2021. Citado nas páginas 27, 31, 44, 46 e 48.
- PARDO, L. **Statistical inference based on divergence measures**. [S.l.]: Chapman and Hall/CRC, 2018. Citado nas páginas 31 e 52.
- RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. **Journal of the Royal Statistical Society Series C: Applied Statistics**, Oxford University Press, v. 54, n. 3, p. 507–554, 2005. Citado na página 72.
- ROCHA, R.; NADARAJAH, S.; TOMAZELLA, V.; LOUZADA, F. Two new defective distributions based on the marshall–olkin extension. **Lifetime data analysis**, Springer, v. 22, n. 2, p. 216–240, 2016. Citado nas páginas 53 e 54.
- _____. A new class of defective models based on the marshall–olkin family of distributions for cure rate modeling. **Computational Statistics & Data Analysis**, Elsevier, v. 107, p. 48–63, 2017. Citado nas páginas 53 e 54.
- ROCHA, R.; NADARAJAH, S.; TOMAZELLA, V.; LOUZADA, F.; EUDES, A. New defective models based on the kumaraswamy family of distributions with application to cancer data sets. **Statistical methods in medical research**, SAGE Publications Sage UK: London, England, v. 26, n. 4, p. 1737–1755, 2017. Citado nas páginas 53 e 54.
- ROCHA, R. F. d. Defective models for cure rate modeling. 2016. Citado nas páginas 53 e 54.
- RODRIGUES, A. S.; CALSAVARA, V. F.; BERTOLLI, E.; PERES, S. V.; TOMAZELLA, V. L. Bayesian long-term survival model including a frailty term: Application to melanoma data. **Chilean Journal of Statistics (ChJS)**, v. 12, n. 1, 2021. Citado nas páginas 25 e 39.
- RODRIGUES, A. S.; CALSAVARA, V. F.; TOMAZELLA, V. L. D. Modeling cure fraction with frailty term in latent risk: a bayesian approach. **arXiv preprint arXiv:1803.08128**, 2018. Citado na página 25.
- RODRIGUES, J.; CANCHO, V. G.; CASTRO, M. de; LOUZADA-NETO, F. On the unification of long-term survival models. **Statistics & Probability Letters**, Elsevier, v. 79, n. 6, p. 753–759, 2009. Citado nas páginas 25 e 77.
- SANTOS, M. R. d.; ACHCAR, J. A.; MARTINEZ, E. Z. Bayesian and maximum likelihood inference for the defective gompertz cure rate model with covariates: an application to the cervical carcinoma study. **Ciência e Natura**, v. 39, n. 2, p. 244–258, 2017. Citado na página 54.
- SCUDILIO, J.; CALSAVARA, V. F.; ROCHA, R.; LOUZADA, F.; TOMAZELLA, V.; RODRIGUES, A. S. Defective models induced by gamma frailty term for survival data with cured fraction. **Journal of Applied Statistics**, Taylor & Francis, v. 46, n. 3, p. 484–507, 2019. Citado nas páginas 26, 53 e 54.
- SIBIM, A. C.; CANCHO, V. G. Estimaco e diagnstico na distribuico exponencial por partes em anlise de sobrevivncia com frao de cura. **Mestrado em cincias da computaco. Universidade de Sao Paulo. Sao Paulo**, 2011. Citado nas páginas 26, 31, 41, 44 e 52.

SOUZA, D. de; CANCHO, V. G.; RODRIGUES, J.; BALAKRISHNAN, N. Bayesian cure rate models induced by frailty in survival analysis. **Statistical methods in medical research**, SAGE Publications Sage UK: London, England, v. 26, n. 5, p. 2011–2028, 2017. Citado nas páginas 23 e 77.

SPIEGELHALTER, D.; THOMAS, A.; BEST, N.; GILKS, W. Bugs 0.5: Bayesian inference using gibbs sampling manual (version ii). **MRC Biostatistics Unit, Institute of Public Health, Cambridge, UK**, p. 1–59, 1996. Citado na página 26.

TEAM, R. **A language and environment for statistical computing. Foundation for statistical computing**. [S.l.]: <http://www.r-project.org> Online, 2020. Citado nas páginas 42 e 51.

TEAM, R. C. R language and environment for statistical computing, r foundation for statistical. **Computing**, 2020. Citado na página 89.

TEAM, S. D. **The Stan Core Library, Version 2.18.0**. 2018. Citado na página 89.

TESEMA, G. A.; SEIFU, B. L.; TESSEMA, Z. T.; WORKU, M. G.; TESHALE, A. B. Incidence of infant mortality and its predictors in east africa using gompertz gamma shared frailty model. **Archives of Public Health**, Springer, v. 80, n. 1, p. 195, 2022. Citado na página 25.

TOMAZELLA, V. L. D.; MILANI, E. Â.; DIAS, T. C. M. Gompertz regression model with gamma frailty: a study on the application in lung cancer. **Brazilian Journal of Biometrics**, v. 36, n. 4, p. 860–879, 2018. Citado na página 25.

TORRES, M. d. P. *et al.* Determinação de propriedades térmicas em problemas de condução de calor por inferência bayesiana com o método de monte carlo hamiltoniano. Universidade do Estado do Rio de Janeiro, 2018. Citado na página 26.

TSODIKOV, A.; IBRAHIM, J. G.; YAKOVLEV, A. Estimating cure rates from survival data: an alternative to two-component mixture models. **Journal of the American Statistical Association**, Taylor & Francis, v. 98, n. 464, p. 1063–1078, 2003. Citado nas páginas 25 e 77.

TWEEDIE, M. C. *et al.* An index which distinguishes between some important exponential families. In: **Statistics: Applications and new directions: Proc. Indian statistical institute golden Jubilee International conference**. [S.l.: s.n.], 1984. v. 579, p. 579–604. Citado na página 25.

VAUPEL, J. W.; MANTON, K. G.; STALLARD, E. The impact of heterogeneity in individual frailty on the dynamics of mortality. **Demography**, Springer, v. 16, n. 3, p. 439–454, 1979. Citado nas páginas 23, 24, 33 e 36.

WHEELER, M. W.; WESTERHOUT, J.; BAUMERT, J. L.; REMINGTON, B. C. Bayesian stacked parametric survival with frailty components and interval-censored failure times: an application to food allergy risk. **Risk Analysis**, Wiley Online Library, v. 41, n. 1, p. 56–66, 2021. Citado na página 25.

WHTTMORE, G. An inverse gaussian model for labour turnover. **Journal of the Royal Statistical Society: Series A (General)**, Wiley Online Library, v. 142, n. 4, p. 468–478, 1979. Citado na página 54.

WIENKE, A. **Frailty models in survival analysis**. [S.l.]: Chapman and Hall/CRC, 2010. Citado nas páginas 25, 31, 33, 34, 35, 36, 37, 38 e 39.

XAVIER, C. M. **Métodos de Monte Carlo Hamiltoniano aplicados em modelos GARCH**. Tese (Doutorado) — Universidade de São Paulo, 2019. Citado nas páginas 27, 31, 45, 46 e 48.

YAKOVLEV, A. Y.; TSODIKOV, A. D.; ASSELAIN, B. **Stochastic models of tumor latency and their biostatistical applications**. [S.l.]: World Scientific, 1996. v. 1. Citado nas páginas 25, 77 e 81.

ZAVALETA, K. E. C. Modelos série de potência com excesso de zeros observáveis e latentes. 2016. Citado na página 36.

ZHOU, H.; HANSON, T.; JARA, A.; ZHANG, J. Modelling county level breast cancer survival data using a covariate-adjusted frailty proportional hazards model. **The annals of applied statistics**, v. 9, n. 1, p. 43, 2015. Citado na página 24.

ZHOU, H.; HANSON, T.; ZHANG, J. Generalized accelerated failure time spatial frailty model for arbitrarily censored data. **Lifetime data analysis**, Springer, v. 23, n. 3, p. 495–515, 2017. Citado na página 24.

DEMONSTRAÇÕES DO CAPÍTULO 3

As operações e as provas apresentadas a seguir são referentes ao Capítulo 3. Estas foram realizadas para fundamentar o estudo, a interpretação e a caracterização do modelo neste capítulo.

A.0.1 Afirmação (3.4)

A função de sobrevivência é relacionada à função de risco acumulado $H_0(t)$ pela seguinte relação:

$$S(t) = \exp(-H_0(t)) = \exp\left(-\int_0^t h_0(u) du\right).$$

Aqui, $h_0(t)$ é a função de risco base que descreve a taxa de falha instantânea no tempo t . Enquanto a distribuição de fragilidade é modelada por uma variável aleatória Z , e a transformada de Laplace de Z é dada por:

$$\mathcal{L}_Z(s) = \exp\left(-\frac{1}{\gamma}[(1+s)^\gamma - 1]\right), \quad \gamma \leq 1. \quad (\text{A.1})$$

Essa transformada de Laplace caracteriza a distribuição de fragilidade de Z , o que afeta a taxa de falha $h_0(t)$. Por outra parte, a função de sobrevivência condicional $S(t | Z)$, dada uma realização Z , é dada por:

$$S(t | Z) = \exp(-ZH_0(t)), \quad (\text{A.2})$$

em que Z é a variável aleatória de fragilidade e $H_0(t)$ é a função de risco cumulativa. A função de sobrevivência marginal é dada pela esperança da função de sobrevivência condicional $S(t | Z)$ em relação à distribuição de Z :

$$\begin{aligned} S(t) &= \mathbb{P}(T > t) = \mathbb{E}[\mathbb{P}(T > t | Z)] \\ &= \mathbb{E}[S(t | Z)] \\ &\stackrel{(\text{A.2})}{=} \mathbb{E}[\exp(-Z \cdot H_0(t))] \end{aligned}$$

em que $\mathbb{E}[\exp(-Z \cdot H_0(t))]$ é a transformada de Laplace da variável aleatória Z avaliada em $s = H_0(t)$. Ou seja, temos:

$$S(t) = \mathcal{L}_Z(H_0(t)),$$

Substituindo $s = H_0(t)$ na expressão (A.1), obtemos:

$$S(t) = \exp\left(-\frac{1}{\gamma}[(1 + H_0(t))^\gamma - 1]\right).$$

a qual representa exatamente a equação (3.4). □

A.0.2 Verificação $\lim_{t \rightarrow \infty} S(t)$ quando $\gamma \in (0, 1)$

Vamos calcular o limite da função de sobrevivência quando $t \rightarrow \infty$. A função de sobrevivência marginal é

$$S(t) = \exp\left\{-\frac{1}{\gamma}[(1 + H_0(t))^\gamma - 1]\right\}, \quad \gamma \in (0, 1]$$

O comportamento de $H_0(t)$ quando $t \rightarrow \infty$, para uma distribuição de sobrevivência arbitrária é verdade que: $\lim_{t \rightarrow \infty} H_0(t) = \infty$, então o comportamento de $(1 + H_0(t))^\gamma$ quando $t \rightarrow \infty$ e $\gamma \in (0, 1]$. Veja que quando $\gamma > 0$, uma função da forma x^γ é crescente em x e:

$$\lim_{x \rightarrow \infty} x^\gamma = \infty, \quad \text{para } \gamma > 0$$

Portanto:

$$\lim_{t \rightarrow \infty} (1 + H_0(t))^\gamma = \infty$$

Logo, substituindo no expoente da função de sobrevivência, no expoente tem-se:

$$-\frac{1}{\gamma}[(1 + H_0(t))^\gamma - 1]$$

Quando $t \rightarrow \infty$, $(1 + H_0(t))^\gamma \rightarrow \infty$, então:

$$\lim_{t \rightarrow \infty} \left[-\frac{1}{\gamma}((1 + H_0(t))^\gamma - 1)\right] = -\infty$$

logo

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

O caso especial $\gamma = 1$, a função simplifica para:

$$S(t) = \exp\{-(1 + H_0(t))^1 - 1\} = \exp\{-H_0(t)\}$$

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

A.0.3 Verificação p_0

Calculando o limite da função de sobrevivência quando $t \rightarrow \infty$ para o caso $\gamma < 0$. A função de sobrevivência marginal do modelo é dada por:

$$S(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}$$

A função de risco acumulado $H_0(t) = \int_0^t h_0(u) du$ é não-decrescente, então $\lim_{t \rightarrow \infty} H_0(t) = \infty$. Logo, o comportamento de $(1 + H_0(t))^\gamma$ quando $t \rightarrow \infty$ e $\gamma < 0$. Para uma função da forma x^γ é decrescente em x e, então

$$\lim_{x \rightarrow \infty} x^\gamma = 0, \quad \text{para } \gamma < 0$$

Portanto:

$$\lim_{t \rightarrow \infty} (1 + H_0(t))^\gamma = 0$$

Agora o expoente da função de sobrevivência é:

$$-\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1]$$

então:

$$\lim_{t \rightarrow \infty} \left[-\frac{1}{\gamma} ((1 + H_0(t))^\gamma - 1) \right] = -\frac{1}{\gamma} (0 - 1) = -\frac{1}{\gamma} \cdot (-1) = \frac{1}{\gamma}$$

Como $\gamma < 0$, $\frac{1}{\gamma}$ é um número negativo. Portanto:

$$\lim_{t \rightarrow \infty} S(t) = \exp \left\{ \frac{1}{\gamma} \right\} > 0$$

E como $\frac{1}{\gamma} < 0$, temos $\exp\{1/\gamma\} \in (0, 1)$, ou seja, uma constante positiva.

□

Afirmção (3.5)

Para determinar a função de risco marginal $h(t)$, vamos partir da função de sobrevivência marginal e usar a relação fundamental entre sobrevivência e risco. O modelo PVF, temos:

$$S(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}$$

em que $H_0(t) = \int_0^t h_0(u) du$ é a função de risco acumulado basal. Por outra parte, a função de risco marginal $h(t)$ está relacionada à função de sobrevivência $S(t)$ por:

$$h(t) = -\frac{d}{dt} \log S(t)$$

então,

$$\begin{aligned}
 h(t) &= -\frac{d}{dt} \log S(t) = \frac{d}{dt} \left\{ \frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\} \\
 &= \frac{1}{\gamma} \cdot \frac{d}{dt} [(1 + H_0(t))^\gamma - 1] \\
 &= \frac{1}{\gamma} \cdot \frac{d}{dt} (1 + H_0(t))^\gamma \\
 &= \gamma (1 + H_0(t))^{\gamma-1} \cdot \frac{d}{dt} (1 + H_0(t)) \\
 &= \frac{1}{\gamma} \cdot [\gamma (1 + H_0(t))^{\gamma-1} h_0(t)] \\
 &= (1 + H_0(t))^{\gamma-1} h_0(t)
 \end{aligned}$$

Portanto tem-se a equação (3.5),

$$h(t) = (1 + H_0(t))^{\gamma-1} h_0(t)$$

□

A.0.4 Análise da Função de Risco Marginal e seu Comportamento Limite

A função de risco marginal do modelo PVF-Induced é derivada a partir da função de sobrevivência marginal. Vamos analisar seu comportamento assintótico no casos apresentados. A função de sobrevivência marginal do modelo PVF-Induced é:

$$S(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\},$$

em que $H_0(t) = \int_0^t h_0(u) du$ é a função de risco acumulado basal e $\gamma \leq 1$. A função de risco marginal $h(t)$ é definida como:

$$h(t) = -\frac{d}{dt} \log S(t) = (1 + H_0(t))^{\gamma-1} h_0(t).$$

A análise do comportamento assintótico sob a premissa de que $H_0(t) \rightarrow \infty$ quando $t \rightarrow \infty$, analisamos o limite de $h(t)$ para diferentes valores de γ . A expressão $(1 + H_0(t))^{\gamma-1}$ tende a zero quando $\gamma - 1 < 0$, ou seja, $\gamma < 1$.

- *Caso 1:* $\gamma = 1$; Quando $\gamma = 1$, temos $\gamma - 1 = 0$, então:

$$h(t) = (1 + H_0(t))^0 h_0(t) = h_0(t) \quad \text{então} \quad \lim_{t \rightarrow \infty} h(t) = \lim_{t \rightarrow \infty} h_0(t).$$

- *Caso 2:* $\gamma < 1$ (incluindo $\gamma < 0$ e $0 < \gamma < 1$)

Para $\gamma < 1$, temos $\gamma - 1 < 0$, logo $(1 + H_0(t))^{\gamma-1} \rightarrow 0$. O comportamento limite de $h(t) = (1 + H_0(t))^{\gamma-1} h_0(t)$ depende da taxa de crescimento/decrescimento de $h_0(t)$ em relação ao fator que tende a zero.

– Caso 2a: $h_0(t)$ é limitada ou tende a zero

Se $h_0(t)$ é limitada superiormente por alguma constante $M < \infty$ (ou tende a zero), então:

$$0 \leq h(t) \leq M \cdot (1 + H_0(t))^{\gamma-1} \rightarrow 0.$$

Pelo teorema do sanduíche, $\lim_{t \rightarrow \infty} h(t) = 0$.

Exemplos:

* Exponencial: $h_0(t) = \lambda$ (constante) $\Rightarrow h(t) = \lambda(1 + \lambda t)^{\gamma-1} \rightarrow 0$.

* Weibull com $\alpha \leq 1$: $h_0(t) = \alpha \lambda t^{\alpha-1}$ tende a zero ou é constante, logo $h(t) \rightarrow 0$.

– Caso 2b: $h_0(t) \rightarrow \infty$ (Weibull com $\alpha > 1$, por exemplo)

Aqui pode haver competição entre:

* $A(t) = (1 + H_0(t))^{\gamma-1} \rightarrow 0$,

* $B(t) = h_0(t) \rightarrow \infty$.

O produto $A(t) \cdot B(t)$ pode tender a zero, a uma constante ou a infinito, dependendo das taxas relativas.

Exemplo modelo Weibull:

$$h_0(t) = \alpha \lambda t^{\alpha-1}, \quad H_0(t) = \lambda t^\alpha.$$

Para t grande, $1 + \lambda t^\alpha \approx \lambda t^\alpha$, então:

$$h(t) \approx \alpha \lambda (\lambda t^\alpha)^{\gamma-1} t^{\alpha-1} = \alpha \lambda^\gamma t^{\alpha\gamma-1}.$$

Analisando o expoente $\alpha\gamma - 1$:

* Se $\alpha\gamma - 1 < 0$ (ou seja, $\gamma < 1/\alpha$): $h(t) \rightarrow 0$.

* Se $\alpha\gamma - 1 = 0$ (ou seja, $\gamma = 1/\alpha$): $h(t) \rightarrow \alpha \lambda^\gamma$ (constante finita).

* Se $\alpha\gamma - 1 > 0$ (ou seja, $\gamma > 1/\alpha$): $h(t) \rightarrow \infty$.

• Caso 3: $\gamma = 0$

Para $\gamma = 0$, a expressão deve ser interpretada como o limite quando $\gamma \rightarrow 0$. Neste caso:

$$h(t) = \frac{h_0(t)}{1 + H_0(t)}.$$

Para Weibull: $h(t) \approx \frac{\alpha \lambda t^{\alpha-1}}{\lambda t^\alpha} = \frac{\alpha}{t} \rightarrow 0$.

• Caso 4: $\gamma < 0$ (modelo de cura) Para $\gamma < 0$, temos $\gamma - 1 < -1$, então o fator $(1 + H_0(t))^{\gamma-1} \rightarrow 0$ mais rapidamente. No exemplo Weibull:

$$h(t) \approx \alpha \lambda^\gamma t^{\alpha\gamma-1},$$

e como $\gamma < 0$, $\alpha\gamma - 1 < -1$, logo $h(t) \rightarrow 0$ muito rapidamente.

Logo resumindo tem-se

$$\lim_{t \rightarrow \infty} h(t) = \begin{cases} \lim_{t \rightarrow \infty} h_0(t), & \text{se } \gamma = 1, \\ 0, & \text{se } \gamma < 1 \text{ e } \lim_{t \rightarrow \infty} \frac{h_0(t)}{(1 + H_0(t))^{1-\gamma}} = 0, \\ c > 0, & \text{se } \gamma < 1 \text{ e } \lim_{t \rightarrow \infty} \frac{h_0(t)}{(1 + H_0(t))^{1-\gamma}} = c, \\ \infty, & \text{se } \gamma < 1 \text{ e } \lim_{t \rightarrow \infty} \frac{h_0(t)}{(1 + H_0(t))^{1-\gamma}} = \infty. \end{cases}$$

Finalmente a conclusão do comportamento de limite da função de risco marginal $h(t) = (1 + H_0(t))^{\gamma-1} h_0(t)$ é determinado pela interação entre o parâmetro γ , que controla a heterogeneidade e a presença de cura, a função de risco basal $h_0(t)$ e sua taxa de crescimento/decrescimento e a função de risco acumulado basal $H_0(t)$, que cresce sem limite. A análise correta requer considerar a taxa relativa entre o fator $(1 + H_0(t))^{\gamma-1}$ (que tende a zero para $\gamma < 1$) e $h_0(t)$ (que pode tender a zero, a uma constante, ou divergir). O exemplo Weibull ilustra claramente os três regimes possíveis: convergência para zero, para uma constante, ou divergência para infinito.

□

A.0.5 Afirmação (3.6)

Para determinar a expressão do modelo de mistura com fração de cura, vamos partir do modelo PVF-Induzido original, o qual é dada por

$$S(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}, \quad \gamma \leq 1$$

Dado que:

- Para $\gamma > 0$: $\lim_{t \rightarrow \infty} S(t) = 0$ (todos eventualmente falham)
- Para $\gamma < 0$: $\lim_{t \rightarrow \infty} S(t) = \exp\{1/\gamma\} > 0$ (fração de cura)

Quando $\gamma < 0$, o parâmetro γ acumula **duas funções**:

- Controla a heterogeneidade não observada entre os suscetíveis
- Determina a fração de cura $p_0 = \exp\{1/\gamma\}$

Isso cria uma limitação: a mesma quantidade γ governa ambos os aspectos, impedindo que sejam modelados separadamente. Portanto, para resolver essa limitação, vamos **separar explicitamente** os dois componentes:

1. **Fração de cura** (p_0): Proporção da população que nunca experimentará o evento
2. **Sobrevivência dos suscetíveis** ($S_u(t)$): Distribuição de tempo até o evento para aqueles que são suscetíveis

Por outra parte, pela lei da probabilidade total, a função de sobrevivência populacional é:

$$S_p(t) = \mathbb{P}(T > t) = \mathbb{P}(\text{curado}) \times 1 + \mathbb{P}(\text{suscetível}) \times \mathbb{P}(T > t \mid \text{suscetível})$$

especificação dos componentes

- $\mathbb{P}(\text{curado}) = p_0$
- $\mathbb{P}(\text{suscetível}) = 1 - p_0$
- $\mathbb{P}(T > t \mid \text{suscetível}) = S_u(t)$

Portanto:

$$S_p(t) = p_0 \times 1 + (1 - p_0) \times S_u(t) = p_0 + (1 - p_0)S_u(t)$$

Agora para os suscetíveis, utilizamos o modelo PVF-Induzida **com** $\gamma > 0$ para capturar heterogeneidade não observada, mas agora sem o papel de fração de cura:

$$S_u(t) = \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}, \quad \gamma \in (0, 1)$$

A restrição $\gamma \in (0, 1)$ garante que:

- $S_u(t)$ é uma função de sobrevivência própria ($\lim_{t \rightarrow \infty} S_u(t) = 0$)
- O parâmetro γ agora representa apenas heterogeneidade não observada
- $\gamma \rightarrow 1$: pouca heterogeneidade
- $\gamma \rightarrow 0$: muita heterogeneidade

Logo, substituindo $S_u(t)$ na expressão da mistura:

$$S_p(t) = p_0 + (1 - p_0) \exp \left\{ -\frac{1}{\gamma} [(1 + H_0(t))^\gamma - 1] \right\}$$

com as restrições:

- $p_0 \in (0, 1)$ (fração de cura)
- $\gamma \in (0, 1)$ (heterogeneidade entre suscetíveis)

Além disso, veja que:

- **Limite quando $t \rightarrow 0$:**

$$S_p(0) = p_0 + (1 - p_0) \cdot 1 = 1$$

- **Limite quando $t \rightarrow \infty$: então**

$$\lim_{t \rightarrow \infty} S_p(t) = p_0 + (1 - p_0) \cdot 0 = p_0$$

□

A.0.6 Verificação (3.7): $S_u(t | \mathbf{w})$

A função de sobrevivência dos suscetíveis $S_u(t | \mathbf{w})$ é determinada a partir do modelo de fragilidade PVF, incorporando covariáveis \mathbf{w} através de um modelo de riscos proporcionais. No modelo de fragilidade, a função de risco individual, condicional à fragilidade Z , é dada por:

$$h(t | Z) = Zh_0(t),$$

em que $h_0(t)$ é a função de risco basal e Z é uma variável aleatória não negativa com média 1 e variância $1 - \gamma$, seguindo uma distribuição PVF. A função de sobrevivência condicional a Z é:

$$S(t | Z) = \exp(-ZH_0(t)), \quad \text{com} \quad H_0(t) = \int_0^t h_0(u) du.$$

Para obter a sobrevivência marginal, integramos sobre a distribuição de Z :

$$S_u(t) = \mathbb{E}_Z[\exp(-ZH_0(t))] = \mathcal{L}_Z(H_0(t)),$$

em que $\mathcal{L}_Z(s) = \mathbb{E}[e^{-sZ}]$ é a transformada de Laplace de Z . Para a família PVF com $\mathbb{E}[Z] = 1$ e $\text{Var}(Z) = 1 - \gamma$, a transformada de Laplace é:

$$\mathcal{L}_Z(s) = \exp\left\{-\frac{1}{\gamma}[(1+s)^\gamma - 1]\right\}, \quad \gamma \in (0, 1).$$

Substituindo $s = H_0(t)$ na transformada de Laplace:

$$S_u(t) = \exp\left\{-\frac{1}{\gamma}[(1+H_0(t))^\gamma - 1]\right\}. \quad (\text{A.3})$$

esta equação representa a função de sobrevivência dos suscetíveis na ausência de covariáveis. Logo para incluir o efeito de covariáveis $\mathbf{w} = (w_1, \dots, w_q)^\top$ no risco dos suscetíveis, adotamos o modelo de riscos proporcionais de Cox:

$$h(t | \mathbf{w}) = h_0(t) \exp\{\mathbf{w}^\top \boldsymbol{\beta}_2\},$$

em que $\boldsymbol{\beta}_2 = (\beta_{21}, \dots, \beta_{2q})^\top$ é o vetor de coeficientes de regressão.

A função de risco acumulado condicional às covariáveis é:

$$\begin{aligned} H(t | \mathbf{w}) &= \int_0^t h(u | \mathbf{w}) du = \int_0^t h_0(u) \exp\{\mathbf{w}^\top \beta_2\} du \\ &= \exp\{\mathbf{w}^\top \beta_2\} \int_0^t h_0(u) du \\ &= H_0(t) \exp\{\mathbf{w}^\top \beta_2\}. \end{aligned}$$

Logo, a sobrevivência marginal condicional às covariáveis é obtida substituindo $H_0(t)$ por $H(t | \mathbf{w})$ na expressão determinada em (A.3), isto é: substituindo $H(t | \mathbf{w}) = H_0(t) \exp\{\mathbf{w}^\top \beta_2\}$: em (A.3)

$$S_u(t | \mathbf{w}) = \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t) \exp\{\mathbf{w}^\top \beta_2\} \right)^\gamma - 1 \right] \right\}.$$

e esta expressão é exatamente a equação (3.7).

Interpretação dos Componentes

- $\gamma \in (0, 1)$: parâmetro de fragilidade que controla a heterogeneidade não observada entre os suscetíveis. Quanto menor γ , maior a heterogeneidade.
- $H_0(t) = \int_0^t h_0(u) du$: função de risco acumulado basal, que pode ser modelada parametricamente (exponencial, Weibull, etc.).
- $\exp\{\mathbf{w}^\top \beta_2\}$: fator multiplicativo que representa o efeito das covariáveis no risco. Valores maiores que 1 indicam aumento do risco, enquanto valores menores que 1 indicam redução do risco.
- β_2 : vetor de coeficientes de regressão. Cada β_{2j} quantifica o efeito log-linear da covariável w_j sobre o risco basal.

Casos Especiais

- **Sem covariáveis** ($\beta_2 = \mathbf{0}$):

$$S_u(t | \mathbf{w}) = \exp \left\{ -\frac{1}{\gamma} \left[(1 + H_0(t))^\gamma - 1 \right] \right\},$$

- **Modelo de fragilidade Gamma** ($\gamma \rightarrow 0$): neste limite, a distribuição PVF tende à distribuição Gamma, e a expressão se aproxima de:

$$S_u(t | \mathbf{w}) \rightarrow \left(1 + H_0(t) \exp\{\mathbf{w}^\top \beta_2\} \right)^{-1/\gamma},$$

que é a forma conhecida para a fragilidade Gamma.

- **Sem heterogeneidade** ($\gamma = 1$):

$$S_u(t | \mathbf{w}) = \exp \left\{ -H_0(t) \exp\{\mathbf{w}^\top \boldsymbol{\beta}_2\} \right\},$$

que é o modelo de riscos proporcionais padrão (sem fragilidade).

□

A.0.7 Afirmação (3.8)

O objetivo é construir a função de sobrevivência populacional $S_p(t | \mathbf{x}, \mathbf{w})$ que considera a existência de uma fração de indivíduos que nunca experimentarão o evento (curados). Considere uma população heterogênea composta por dois tipos de indivíduos:

- **Curados (imunes)**: nunca experimentam o evento de interesse.
- **Suscetíveis**: estão em risco e podem eventualmente experimentar o evento.

Seja Z uma variável indicadora não observada tal que:

$$Z = \begin{cases} 1, & \text{se o indivíduo é curado,} \\ 0, & \text{se o indivíduo é suscetível.} \end{cases}$$

A probabilidade de cura é dada por:

$$\mathbb{P}(Z = 1 | \mathbf{x}) = p_0(\mathbf{x}),$$

em que \mathbf{x} é um vetor de covariáveis que influenciam a chance de cura. Logo:

$$\mathbb{P}(Z = 0 | \mathbf{x}) = 1 - p_0(\mathbf{x}).$$

Agora usando a Lei da Probabilidade Total, a função de sobrevivência populacional condicional às covariáveis:

$$S_p(t | \mathbf{x}, \mathbf{w}) = \mathbb{P}(T > t | \mathbf{x}, \mathbf{w}).$$

Condicionando no status de cura Z , temos:

$$\begin{aligned} S_p(t | \mathbf{x}, \mathbf{w}) &= \mathbb{P}(Z = 1 | \mathbf{x})\mathbb{P}(T > t | Z = 1, \mathbf{w}) + \mathbb{P}(Z = 0 | \mathbf{x})\mathbb{P}(T > t | Z = 0, \mathbf{w}) \\ &= p_0(\mathbf{x}) \times 1 + (1 - p_0(\mathbf{x})) \times S_u(t | \mathbf{w}) \\ &= p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w}) \end{aligned}$$

em que:

- **Indivíduos curados** ($Z = 1$): por definição, nunca experimentam o evento, portanto:

$$\mathbb{P}(T > t | Z = 1, \mathbf{w}) = 1 \quad \forall t > 0.$$

- **Indivíduos suscetíveis** ($Z = 0$): seu tempo de sobrevivência é modelado por uma função de sobrevivência própria $S_u(t | \mathbf{w})$, que depende de covariáveis \mathbf{w} :

$$\mathbb{P}(T > t | Z = 0, \mathbf{w}) = S_u(t | \mathbf{w}).$$

Esta função é própria, ou seja, $\lim_{t \rightarrow \infty} S_u(t | \mathbf{w}) = 0$, indicando que todos os suscetíveis eventualmente falham.

Além disso, $S_p(t | \mathbf{x}, \mathbf{w})$ é verdade que:

- **No tempo $t = 0$:**

$$S_p(0 | \mathbf{x}, \mathbf{w}) = p_0(\mathbf{x}) + (1 - p_0(\mathbf{x})) \times 1 = 1.$$

- **No limite $t \rightarrow \infty$:**

$$\lim_{t \rightarrow \infty} S_p(t | \mathbf{x}, \mathbf{w}) = p_0(\mathbf{x}) + (1 - p_0(\mathbf{x})) \times 0 = p_0(\mathbf{x}).$$

Portanto, a fração de cura é exatamente $p_0(x)$.

- **Monotonicidade:** como $S_u(t | \mathbf{w})$ é não crescente em t , $S_p(t | \mathbf{x}, \mathbf{w})$ também é não crescente.

□

A.0.8 Verificação $h_p(t | \mathbf{x}, \mathbf{w})$

A função de risco $h_p(t | \mathbf{x}, \mathbf{w})$ é definida como a taxa de falha instantânea no tempo t para um indivíduo com covariáveis \mathbf{x} (para a cura) e \mathbf{w} (para a sobrevivência dos suscetíveis). Usando as definições de funções de sobrevivência e densidade. Para qualquer função de sobrevivência $S(t)$ absolutamente contínua, a função de risco é definida como:

$$h(t) = -\frac{d}{dt} \log S(t) = \frac{f(t)}{S(t)},$$

então aplicando ao contexto do modelo de mistura, em que a sobrevivência populacional é dada por:

$$S_p(t | \mathbf{x}, \mathbf{w}) = p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w}),$$

com $p_0(\mathbf{x}) \in (0, 1)$ e $S_u(t | \mathbf{w})$ sendo a função de sobrevivência dos suscetíveis (própria, tal que $\lim_{t \rightarrow \infty} S_u(t | \mathbf{w}) = 0$). Então, derivando a sobrevivência populacional:

$$f_p(t | \mathbf{x}, \mathbf{w}) = -\frac{\partial}{\partial t} S_p(t | \mathbf{x}, \mathbf{w}).$$

Substituindo a expressão de S_p :

$$\begin{aligned} f_p(t | \mathbf{x}, \mathbf{w}) &= -\frac{\partial}{\partial t} [p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w})] \\ &= -(1 - p_0(\mathbf{x})) \frac{\partial}{\partial t} S_u(t | \mathbf{w}) \\ &= (1 - p_0(\mathbf{x}))f_u(t | \mathbf{w}). \end{aligned}$$

Usando a definição $h(t | \mathbf{x}, \mathbf{w}) = \frac{f(t|\mathbf{x},\mathbf{w})}{S(t|\mathbf{x},\mathbf{w})}$, tem-se:

$$h_p(t | \mathbf{x}, \mathbf{w}) = \frac{f_p(t | \mathbf{x}, \mathbf{w})}{S_p(t | \mathbf{x}, \mathbf{w})} = \frac{(1 - p_0(\mathbf{x}))f_u(t | \mathbf{w})}{p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w})}$$

- **Numerador:** $(1 - p_0(\mathbf{x}))f_u(t | \mathbf{w})$ representa a contribuição dos suscetíveis para a densidade de falhas no tempo t . Apenas indivíduos não curados podem falhar.
- **Denominador:** $p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w})$ é a probabilidade de um indivíduo estar vivo no tempo t , considerando tanto os curados (que estão sempre vivos) quanto os suscetíveis que ainda não falharam.

Veja quando:

- $t = 0$: $S_u(0 | \mathbf{w}) = 1$, então:

$$h_p(0 | \mathbf{x}, \mathbf{w}) = h_u(0 | \mathbf{w}) \times \frac{1 - p_0(\mathbf{x})}{p_0(\mathbf{x}) + (1 - p_0(\mathbf{x})) \times 1} = h_u(0 | \mathbf{w})(1 - p_0(\mathbf{x})).$$

O risco inicial é o risco dos suscetíveis reduzido pela proporção de suscetíveis na população.

- Quando $t \rightarrow \infty$: $S_u(t | \mathbf{w}) \rightarrow 0$, então o denominador tende a $p_0(\mathbf{x})$. Como $f_u(t | \mathbf{w})$ também tende a zero (pois é uma densidade própria), o risco populacional tende a zero, refletindo o esgotamento dos suscetíveis.
- Caso sem cura ($p_0(\mathbf{x}) = 0$):

$$h_p(t | \mathbf{x}, \mathbf{w}) = \frac{f_u(t | \mathbf{w})}{S_u(t | \mathbf{w})} = h_u(t | \mathbf{w}),$$

recuperando o modelo de risco padrão.

□

A.0.9 Verificação (3.9)

Seja $\mathcal{D} = \{(t_i, \delta_i, \mathbf{x}_i, \mathbf{w}_i)\}_{i=1}^n$ o conjunto de dados observados, onde:

- t_i é o tempo observado (tempo de falha ou censura),
- δ_i é a função indicadora de falha ($\delta_i = 1$ se o evento ocorreu, $\delta_i = 0$ se censurado),
- \mathbf{x}_i são as covariáveis que afetam a probabilidade de cura,
- \mathbf{w}_i são as covariáveis que afetam o risco dos suscetíveis.

A estrutura do Modelo de proposto é definido pela função de sobrevivência populacional:

$$S_p(t | \mathbf{x}, \mathbf{w}) = p_0(\mathbf{x}) + (1 - p_0(\mathbf{x}))S_u(t | \mathbf{w}).$$

em que:

- $p_0(\mathbf{x}) = F(\mathbf{x}^\top \beta_1)$ é a probabilidade de cura (geralmente usando uma função logística),
- $S_u(t | \mathbf{w})$ é a função de sobrevivência dos suscetíveis, dada pelo modelo PVF:

$$S_u(t | \mathbf{w}) = \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t | \lambda) e^{\mathbf{w}^\top \beta_2} \right)^\gamma - 1 \right] \right\},$$

- $H_0(t | \lambda)$ é a função de risco acumulado basal, dependente de parâmetros λ ,
- $\gamma \in (0, 1)$ é o parâmetro de fragilidade,
- β_1 e β_2 são os coeficientes de regressão.

Por outra parte, a função de densidade dos suscetíveis é obtida derivando negativamente $S_u(t | \mathbf{w})$:

$$f_u(t | \mathbf{w}) = -\frac{d}{dt} S_u(t | \mathbf{w}).$$

Calculando:

$$f_u(t | \mathbf{w}) = S_u(t | \mathbf{w}) \cdot \left(1 + H_0(t | \lambda) e^{\mathbf{w}^\top \beta_2} \right)^{\gamma-1} \cdot e^{\mathbf{w}^\top \beta_2} \cdot h_0(t | \lambda),$$

Além disso, a contribuição individual para a Verossimilhança, de cada indivíduo i , depende de δ_i , isto é:

Caso $\delta_i = 1$ (falha observada)

O indivíduo pertence necessariamente ao grupo suscetível. Sua contribuição é:

$$L_i = (1 - p_0(\mathbf{x}_i)) \cdot f_u(t_i | \mathbf{w}_i).$$

Substituindo $f_u(t_i | \mathbf{w}_i)$:

$$L_i = (1 - p_0(\mathbf{x}_i)) \cdot S_u(t_i | \mathbf{w}_i) \cdot \left(1 + H_0(t_i | \lambda) e^{\mathbf{w}_i^\top \beta_2} \right)^{\gamma-1} \cdot e^{\mathbf{w}_i^\top \beta_2} \cdot h_0(t_i | \lambda).$$

Caso $\delta_i = 0$ (censura)

O indivíduo pode ser curado ou suscetível ainda não falho. Sua contribuição é a sobrevivência populacional:

$$L_i = S_p(t_i | \mathbf{x}_i, \mathbf{w}_i) = p_0(\mathbf{x}_i) + (1 - p_0(\mathbf{x}_i))S_u(t_i | \mathbf{w}_i).$$

Substituindo $S_u(t_i | \mathbf{w}_i)$:

$$L_i = p_0(\mathbf{x}_i) + (1 - p_0(\mathbf{x}_i)) \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t_i | \lambda) e^{\mathbf{w}_i^\top \beta_2} \right)^\gamma - 1 \right] \right\}.$$

Combinando os dois casos usando o indicador δ_i , a contribuição do indivíduo i é:

$$L_i(\vartheta) = \left[(1 - p_0(\mathbf{x}_i)) \cdot S_u(t_i | \mathbf{w}_i) \cdot \left(1 + H_0(t_i | \lambda) e^{\mathbf{w}_i^\top \beta_2} \right)^{\gamma-1} \cdot e^{\mathbf{w}_i^\top \beta_2} \cdot h_0(t_i | \lambda) \right]^{\delta_i} \\ \times [p_0(\mathbf{x}_i) + (1 - p_0(\mathbf{x}_i)) S_u(t_i | \mathbf{w}_i)]^{1-\delta_i}.$$

A verossimilhança total é o produto sobre todos os indivíduos:

$$L(\vartheta) = \prod_{i=1}^n L_i(\vartheta).$$

logo a função de verossimilhança é:

$$L(\vartheta) = \prod_{i=1}^n \left\{ (1 - p_0(\mathbf{x}_i)) \cdot \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t_i | \lambda) e^{\mathbf{w}_i^\top \beta_2} \right)^\gamma - 1 \right] \right\} \right. \\ \left. \cdot \left(1 + H_0(t_i | \lambda) e^{\mathbf{w}_i^\top \beta_2} \right)^{\gamma-1} \cdot e^{\mathbf{w}_i^\top \beta_2} \cdot h_0(t_i | \lambda) \right\}^{\delta_i} \\ \times \left[p_0(\mathbf{x}_i) + (1 - p_0(\mathbf{x}_i)) \exp \left\{ -\frac{1}{\gamma} \left[\left(1 + H_0(t_i | \lambda) e^{\mathbf{w}_i^\top \beta_2} \right)^\gamma - 1 \right] \right\} \right]^{1-\delta_i}.$$

□

DEMONSTRAÇÕES DO CAPÍTULO 4

A seguir, apresentam-se as afirmações relativas ao Capítulo 4, as quais foram formuladas com o objetivo de embasar o estudo e a interpretação do modelo apresentado no referido capítulo.

B.0.1 Verificação da Esperança, Variância e Covariância de N_k

Dado que $N_k | W = w \sim \text{Poisson}(w\rho_k)$, vamos determinar as expressões para a média, a variância e a covariância de N_k , para $k = 1, 2$.

$$\mathbf{E}[N_k] = \mathbb{E}[\mathbb{E}[N_k | W]] = \mathbb{E}[W\rho_k] = \rho_k \mathbb{E}[W] = \rho_k \times 1 = \rho_k$$

$$\begin{aligned} \text{Var}[N_k] &= \mathbb{E}[\text{Var}[N_k | W]] + \text{Var}[\mathbb{E}[N_k | W]] = \mathbb{E}[W\rho_k] + \text{Var}[W\rho_k] \\ &= \rho_k \mathbb{E}[W] + \rho_k^2 \text{Var}[W] = \rho_k \times 1 + \rho_k^2 \left(\frac{1-\gamma}{\sigma} \right) \\ &= \rho_k + \rho_k^2 \left(\frac{1-\gamma}{\sigma} \right) \end{aligned}$$

$$\begin{aligned} \text{Cov}(N_1, N_2) &= \mathbb{E}[\text{Cov}(N_1, N_2 | W)] + \text{Cov}(\mathbb{E}[N_1 | W], \mathbb{E}[N_2 | W]) \\ &= \mathbb{E}[0] + \text{Cov}(W\rho_1, W\rho_2) \quad N_1 \text{ e } N_2 \text{ cond. indep.} \\ &= \rho_1 \rho_2 \text{Cov}(W, W) \\ &= \rho_1 \rho_2 \text{Var}(W) \\ &= \rho_1 \rho_2 (1-\gamma) \end{aligned}$$

B.0.2 Afirmação (4.5)

Antes de analisar a expressão (4.5), observe que, para quaisquer eventos A e B em um mesmo espaço de probabilidade, com $\mathbb{P}(B) > 0$, é verdadeiro que:

$$\mathbb{P}(A) = \mathbb{E}[\mathbb{P}(A | B)].$$

Para nosso caso, considere o evento $A = \{T_1 > t_1, T_2 > t_2\}$ e $B = W$, então

$$\begin{aligned} S(t_1, t_2) &= \mathbb{P}(T_1 > t_1, T_2 > t_2) \\ &= \mathbb{E}[\mathbb{P}(T_1 > t_1, T_2 > t_2 | W)] \\ &= \mathbb{E}[S(t_1, t_2 | W)] \\ &\stackrel{(4.4)}{=} \mathbb{E} \left[\exp \left(-w \sum_{k=1}^2 \rho_k (1 - S_k(t_k)) \right) \right] \end{aligned} \quad (\text{B.1})$$

usando a transformada de Laplace dada em (4.2) na expressão (B.1) tem-se

$$\begin{aligned} S(t_1, t_2) &= \mathcal{L}_W \left(\sum_{k=1}^2 \rho_k (1 - S_k(t_k)) \right) \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\sum_{k=1}^2 \rho_k (1 - S_k(t_k))}{\sigma} \right)^\gamma - 1 \right] \right\} \end{aligned}$$

esta expressão corresponde exatamente à equação (4.5). □

B.0.3 Verificação de p_{00}

Dado que $F_k(t_k)$ é a fda então $\lim_{t_k \rightarrow \infty} F_k(t_k) = 1$, para $k = 1, 2$ ou equivalentemente $\lim_{t_k \rightarrow \infty} S_k(t_k) = 0$. Então:

$$\begin{aligned} p_{00} &= \lim_{t_1, t_2 \rightarrow \infty} S(t_1, t_2) = \lim_{t_1, t_2 \rightarrow \infty} \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\sum_{k=1}^2 \rho_k (1 - S_k(t_k))}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 (1 - S_1(\infty)) + \rho_2 (1 - S_2(\infty))}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 (1 - 0) + \rho_2 (1 - 0)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 + \rho_2}{\sigma} \right)^\gamma - 1 \right] \right\} > 0 \end{aligned}$$

e esta última expressão é exatamente o p_{00} . □

B.0.4 Afirmação (4.6)

A função de sobrevivência bivariada condicional dada em (4.4) é

$$S(t_1, t_2 | W = w) = \exp \left(-w \sum_{k=1}^2 \rho_k [1 - S_k(t_k)] \right).$$

faremos para o caso $k = 1$, a partir de esta última expressão, queremos apagar completamente a restrição sobre T_2 , para isso, veja que $\{T_2 > -\infty\} = \Omega$, dado que os tempos de sobrevivência assumem valores reais ($T_2 \geq 0$). Logo, $T_2 > -\infty$ será sempre verdadeiro, ou seja

$$S(t_1, t_2 | W) = \mathbb{P}(T_1 > t_1, T_2 > -\infty | W) = \mathbb{P}(T_1 > t_1 | W)$$

Nosso caso, $S_2(0) = 1$, assim

$$\begin{aligned} S(t_1 | W) &= S(t_1, 0 | W) = \exp[-w(\rho_1(1 - S_1(t_1)) + \rho_2(1 - S_2(0)))] \\ &= \exp[-w(\rho_1(1 - S_1(t_1)) + \rho_2(1 - 1))] \\ &= \exp[-w\rho_1(1 - S_1(t_1))] \end{aligned} \tag{B.2}$$

Por outra parte,

$$\begin{aligned} S_1(t_1) &= \mathbb{P}(T_1 > t_1) = \mathbb{E}[\mathbb{P}(T_1 > t_1 | W)] \\ &= \mathbb{E}[S(t_1 | W)] \\ &\stackrel{(B.2)}{=} \mathbb{E}[\exp(-w(\rho_1(1 - S_1(t_1))))] \end{aligned} \tag{B.3}$$

e finalmente usando a transformada de Laplace dada em (4.3), tomando $s = \rho_1(1 - S_1(t_1))$ em (B.3)

$$\begin{aligned} S_1(t_1) &= \mathbb{E}[\exp(-ws)] \\ &\stackrel{(4.3)}{=} \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{s}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1(1 - S_1(t_1))}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1(F_1(t_1))}{\sigma} \right)^\gamma - 1 \right] \right\} \end{aligned}$$

e essa última expressão representa exatamente a equação (4.6). Analogamente, a verificação para $k = 2$. Logo, pode-se afirmar que:

$$S_k(t_k) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k(F_k(t_k))}{\sigma} \right)^\gamma - 1 \right] \right\}, \quad \text{para } k = 1, 2. \tag{B.4}$$

□

B.0.5 Verificação de p_{0k}

Segue imediatamente da equação (B.4)

$$\begin{aligned} p_{0k} &= \lim_{t_k \rightarrow \infty} S_k(t_k) = \lim_{t_k \rightarrow \infty} \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k F_k(t_k)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k F_k(+\infty)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k}{\sigma} \right)^\gamma - 1 \right] \right\} \end{aligned}$$

□

B.0.6 Afirmação (4.7)

A função de sobrevivência marginal bivariada é dada por:

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma - 1 \right] \right\}$$

Defina, $a = 1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}$, logo a função de sobrevivência marginal bivariada será:

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} (y^\gamma - 1) \right\}.$$

logo usando $\lim_{\gamma \rightarrow 0} \frac{a^\gamma - 1}{\gamma} = \log(a)$,

$$\begin{aligned} \lim_{\gamma \rightarrow 0} S(t_1, t_2) &= \lim_{\gamma \rightarrow 0} \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &= \exp \left\{ -\lim_{\gamma \rightarrow 0} \frac{\sigma}{\gamma} [a^\gamma - 1] \right\} \\ &= \exp \left\{ -\sigma \lim_{\gamma \rightarrow 0} \frac{a^\gamma - 1}{\gamma} \right\} \\ &= \exp \{ -\sigma \log(a) \} \\ &= \exp \left\{ -\sigma \log \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right) \right\} \end{aligned}$$

e essa última expressão é exatamente a equação (4.7)

□

B.0.7 Afirmação (4.8)

Segue imediatamente de (4.7),

$$\begin{aligned} S(t_1, t_2) &= \exp \left\{ -\sigma \log \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right) \right\} \\ &= \exp \left\{ \log \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\sigma} \right\} \\ &= \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\sigma} \end{aligned}$$

o qual representa a equação (4.8).

□

B.0.8 Verificação $\lim_{\sigma \rightarrow \infty} S(t_1, t_2)$

Considerando a seguinte expressão:

$$S(t_1, t_2) = \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\sigma} \quad (\text{B.5})$$

o limite da função B.5 quando σ tende ao infinito pode ser reescrito como:

$$\begin{aligned} \lim_{\sigma \rightarrow \infty} S(t_1, t_2) &= \lim_{\sigma \rightarrow \infty} \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\sigma} \\ &= \lim_{\sigma \rightarrow \infty} \left(1 + \frac{A}{\sigma} \right)^{-\sigma} = \exp\{-A\}, \end{aligned}$$

em que $A = \rho_1 F_1(t_1) + \rho_2 F_2(t_2)$. Assim, aplicando o limite, obtemos:

$$\begin{aligned} \lim_{\sigma \rightarrow \infty} S(t_1, t_2) &= \exp\{-(\rho_1 F_1(t_1) + \rho_2 F_2(t_2))\} \\ &= \exp\{-\rho_1 F_1(t_1)\} \exp\{-\rho_2 F_2(t_2)\} \end{aligned}$$

□

B.0.9 Afirmação (4.9)

Segue da equação (4.5)

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma - 1 \right] \right\}.$$

sustituindo $\gamma = \frac{1}{2}$ obtem-se

$$\begin{aligned} S(t_1, t_2) &= \exp \left\{ -\frac{\sigma}{1/2} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{\frac{1}{2}} - 1 \right] \right\} \\ &= \exp \left\{ -2\sigma \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{\frac{1}{2}} - 1 \right] \right\} \end{aligned}$$

o qual representa exatamente a expressão (4.9) o qual também é equivalente à expressão (4.10). □

B.0.10 Afirmação (4.11)

Vamos obter a expressão para o caso particular da família PVF quando $\sigma = 0$ e $\mu = \gamma$, que corresponde à distribuição estável positiva. A transformada de Laplace geral da PVF é:

$$\mathcal{L}_W(s) = \exp \left\{ -\frac{\mu}{\gamma} [(\sigma + s)^\gamma - \sigma^\gamma] \right\}$$

Substituindo $\sigma = 0$ e $\mu = \gamma$:

$$\mathcal{L}_W(s) = \exp \left\{ -\frac{\gamma}{\gamma} [(0 + s)^\gamma - 0^\gamma] \right\} = \exp \{-s^\gamma\}$$

Esta é exatamente a transformada de Laplace de uma distribuição estável positiva com parâmetro γ . Por outra parte, a função de Sobrevivência Conjunta com Fragilidade Estável Positiva

$$S(t_1, t_2 | W = w) = \exp \{-w [\rho_1 F_1(t_1) + \rho_2 F_2(t_2)]\}$$

A função de sobrevivência conjunta marginal (não condicional) é obtida integrando sobre a distribuição de W :

$$S(t_1, t_2) = \mathbb{E}_W [S(t_1, t_2 | W)] = \mathbb{E}_W [\exp \{-W [\rho_1 F_1(t_1) + \rho_2 F_2(t_2)]\}]$$

Reconhecemos esta transformada de Laplace de W avaliada no ponto $s = \rho_1 F_1(t_1) + \rho_2 F_2(t_2)$:

$$S(t_1, t_2) = \mathcal{L}_W(\rho_1 F_1(t_1) + \rho_2 F_2(t_2)),$$

Para a fragilidade estável positiva, temos $\mathcal{L}_W(s) = \exp\{-s^\gamma\}$. Portanto:

$$S(t_1, t_2) = \exp \left\{ -[\rho_1 F_1(t_1) + \rho_2 F_2(t_2)]^\gamma \right\}, \quad 0 < \gamma \leq 1$$

e para obter a expressão (4.13), basta substituir em (4.11) $\gamma = 1$. □

B.0.11 Verificação (4.14)

Dado que a função de risco conjunta é definida como:

$$h(t_1, t_2) = \frac{f(t_1, t_2)}{S(t_1, t_2)} = -\frac{\partial^2}{\partial t_1 \partial t_2} \log S(t_1, t_2)$$

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma - 1 \right] \right\}$$

a função de sobrevivência conjunta é: Seja

$$U(t_1, t_2) = \log S(t_1, t_2) = -\frac{\sigma}{\gamma} [A(t_1, t_2)^\gamma - 1]$$

em que $A(t_1, t_2) = 1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}$. As derivadas parciais de U

$$\begin{aligned} \frac{\partial U}{\partial t_1} &= -\frac{\sigma}{\gamma} \cdot \gamma A^{\gamma-1} \cdot \frac{\partial A}{\partial t_1} = -\sigma A^{\gamma-1} \cdot \frac{\rho_1 f_1(t_1)}{\sigma} = -\rho_1 f_1(t_1) A^{\gamma-1} \\ \frac{\partial U}{\partial t_2} &= -\rho_2 f_2(t_2) A^{\gamma-1} \end{aligned}$$

Por outra partes a derivada parcial Mista de U , agora, derivamos $\frac{\partial U}{\partial t_1}$ em relação a t_2 :

$$\begin{aligned} \frac{\partial^2 U}{\partial t_1 \partial t_2} &= \frac{\partial}{\partial t_2} [-\rho_1 f_1(t_1) A^{\gamma-1}] = -\rho_1 f_1(t_1) \cdot \frac{\partial}{\partial t_2} (A^{\gamma-1}) \\ &= -\rho_1 f_1(t_1) \left[(\gamma-1) A^{\gamma-2} \cdot \frac{\partial A}{\partial t_2} \right] \\ &= -\rho_1 f_1(t_1) \left[(\gamma-1) A^{\gamma-2} \cdot \frac{\rho_2 f_2(t_2)}{\sigma} \right] \\ &= -\rho_1 f_1(t_1) \cdot (\gamma-1) A^{\gamma-2} \cdot \frac{\rho_2 f_2(t_2)}{\sigma} \\ &= \rho_1 \rho_2 f_1(t_1) f_2(t_2) \cdot \frac{1-\gamma}{\sigma} A^{\gamma-2} \end{aligned}$$

A relação com a função de risco conjunta

$$h(t_1, t_2) = -\frac{\partial^2 U}{\partial t_1 \partial t_2} + \left(\frac{\partial U}{\partial t_1} \right) \left(\frac{\partial U}{\partial t_2} \right) = -\frac{\partial^2 \log S}{\partial t_1 \partial t_2} + \left(\frac{\partial \log S}{\partial t_1} \right) \left(\frac{\partial \log S}{\partial t_2} \right)$$

Isto vem do fato que a densidade conjunta é:

$$f(t_1, t_2) = \frac{\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2} = S(t_1, t_2) \left[\frac{\partial^2 U}{\partial t_1 \partial t_2} + \left(\frac{\partial U}{\partial t_1} \right) \left(\frac{\partial U}{\partial t_2} \right) \right]$$

E como $h(t_1, t_2) = f(t_1, t_2)/S(t_1, t_2)$, então substituindo as derivadas calculadas: $\frac{\partial U}{\partial t_1} = -\rho_1 f_1(t_1) A^{\gamma-1}$; $\frac{\partial U}{\partial t_2} = -\rho_2 f_2(t_2) A^{\gamma-1}$ e $\frac{\partial^2 U}{\partial t_1 \partial t_2} = \rho_1 \rho_2 f_1(t_1) f_2(t_2) \cdot \frac{1-\gamma}{\sigma} A^{\gamma-2}$ Substituindo:

$$\begin{aligned} h(t_1, t_2) &= - \left[\rho_1 \rho_2 f_1(t_1) f_2(t_2) \cdot \frac{1-\gamma}{\sigma} A^{\gamma-2} \right] + [-\rho_1 f_1(t_1) A^{\gamma-1}] [-\rho_2 f_2(t_2) A^{\gamma-1}] \\ &= -\rho_1 \rho_2 f_1(t_1) f_2(t_2) \cdot \frac{1-\gamma}{\sigma} A^{\gamma-2} + \rho_1 \rho_2 f_1(t_1) f_2(t_2) \cdot A^{2(\gamma-1)} \\ &= \rho_1 \rho_2 f_1(t_1) f_2(t_2) \left[A^{2(\gamma-1)} - \frac{1-\gamma}{\sigma} A^{\gamma-2} \right] \\ &= \rho_1 \rho_2 f_1(t_1) f_2(t_2) A^{\gamma-2} \left[A^\gamma - \frac{1-\gamma}{\sigma} \right] \\ &= \rho_1 \rho_2 f_1(t_1) f_2(t_2) \left[\frac{1-\gamma}{\sigma} + A^\gamma \right] A^{\gamma-2} \end{aligned}$$

e finalmente substituindo $A = 1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}$:

$$h(t_1, t_2) = \rho_1 \rho_2 f_1(t_1) f_2(t_2) \left[\frac{1 - \gamma}{\sigma} + \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^\gamma \right] \\ \times \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{\gamma-2}$$

□

B.0.12 Verificação (4.15)

A função de risco marginal para o evento k é definida como:

$$h_k(t_k) = -\frac{d}{dt_k} \log S_k(t_k), \quad k = 1, 2$$

em que $S_k(t_k)$ é a função de sobrevivência marginal para o evento k , a equação (4.6), a função de sobrevivência marginal para o evento k é:

$$S_k(t_k) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k F_k(t_k)}{\sigma} \right)^\gamma - 1 \right] \right\}, \quad k = 1, 2$$

considere $U_k(t_k) = \log S_k(t_k)$:

$$U_k(t_k) = -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_k F_k(t_k)}{\sigma} \right)^\gamma - 1 \right]$$

Para simplificar, defina-se uma função auxiliar:

$$B_k(t_k) = 1 + \frac{\rho_k F_k(t_k)}{\sigma}$$

Então:

$$U_k(t_k) = -\frac{\sigma}{\gamma} [B_k(t_k)^\gamma - 1]$$

derivamos $U_k(t_k)$ em relação a t_k :

$$\begin{aligned} \frac{dU_k}{dt_k} &= -\frac{d}{dt_k} \left\{ \frac{\sigma}{\gamma} [B_k(t_k)^\gamma - 1] \right\} = -\frac{\sigma}{\gamma} \cdot \frac{d}{dt_k} [B_k(t_k)^\gamma - 1] = -\frac{\sigma}{\gamma} \cdot \frac{d}{dt_k} [B_k(t_k)^\gamma] \\ &= \gamma B_k(t_k)^{\gamma-1} \cdot \frac{dB_k}{dt_k} = -\frac{\sigma}{\gamma} \cdot \gamma B_k(t_k)^{\gamma-1} \cdot \frac{dB_k}{dt_k} \\ &= -\sigma B_k(t_k)^{\gamma-1} \cdot \frac{dB_k}{dt_k} \end{aligned} \tag{B.6}$$

Agora, calculamos $\frac{dB_k}{dt_k}$, em que $B_k(t_k) = 1 + \frac{\rho_k F_k(t_k)}{\sigma}$, derivando em relação a t_k :

$$\frac{dB_k}{dt_k} = 0 + \frac{\rho_k}{\sigma} \cdot \frac{dF_k}{dt_k}$$

dado que $\frac{dF_k}{dt_k} = f_k(t_k)$ logo tem-se:

$$\frac{dB_k}{dt_k} = \frac{\rho_k f_k(t_k)}{\sigma}$$

substituindo na expressão (B.6), $\frac{dB_k}{dt_k}$ na expressão para $\frac{dU_k}{dt_k}$:

$$\frac{dU_k}{dt_k} = -\sigma B_k(t_k)^{\gamma-1} \cdot \frac{\rho_k f_k(t_k)}{\sigma} = -\rho_k f_k(t_k) B_k(t_k)^{\gamma-1}$$

Lembrando que

$$h_k(t_k) = -\frac{dU_k}{dt_k} = -[-\rho_k f_k(t_k) B_k(t_k)^{\gamma-1}] = \rho_k f_k(t_k) B_k(t_k)^{\gamma-1}$$

e finalmente substituindo $B_k(t_k) = 1 + \frac{\rho_k F_k(t_k)}{\sigma}$, tem-se a igualdade dada em (4.15):

$$h_k(t_k) = \rho_k f_k(t_k) \left(1 + \frac{\rho_k F_k(t_k)}{\sigma}\right)^{\gamma-1}, \quad k = 1, 2$$

□

B.0.13 Verificação (4.17)

Dado que

$$\vartheta^*(t_1, t_2) = \frac{S(t_1, t_2) \frac{\partial^2}{\partial t_1 \partial t_2} S(t_1, t_2)}{\left(\frac{\partial}{\partial t_1} S(t_1, t_2)\right) \left(\frac{\partial}{\partial t_2} S(t_1, t_2)\right)}$$

da equação (4.5), a função de sobrevivência conjunta é:

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}\right)^\gamma - 1 \right] \right\}.$$

Para simplificar, defina-se: $A(t_1, t_2) = 1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}$, então

$$S(t_1, t_2) = \exp \left\{ -\frac{\sigma}{\gamma} [A(t_1, t_2)^\gamma - 1] \right\}.$$

as derivadas parciais de primeira ordem de S

$$\frac{\partial S(t_1, t_2)}{\partial t_1} = -S \cdot \rho_1 f_1(t_1) \cdot A^{\gamma-1} \quad \text{e} \quad \frac{\partial S(t_1, t_2)}{\partial t_2} = -S \cdot \rho_2 f_2(t_2) \cdot A^{\gamma-1}$$

o produto no Denominador, é obtido multiplicando as duas derivadas primeiras:

$$\begin{aligned} \left(\frac{\partial S(t_1, t_2)}{\partial t_1}\right) \left(\frac{\partial S(t_1, t_2)}{\partial t_2}\right) &= [-S(t_1, t_2) \rho_1 f_1 A^{\gamma-1}] \times [-S(t_1, t_2) \rho_2 f_2 A^{\gamma-1}] \\ &= S(t_1, t_2)^2 \rho_1 \rho_2 f_1 f_2 A^{2(\gamma-1)}. \end{aligned}$$

por outro lado

$$\frac{\partial^2 S(t_1, t_2)}{\partial t_1 \partial t_2} = \rho_1 \rho_2 f_1 f_2 S(t_1, t_2) \left[A^{2(\gamma-1)} + \frac{1-\gamma}{\sigma} A^{\gamma-2} \right].$$

substituindo os cálculos

$$\begin{aligned}
\vartheta^*(t_1, t_2) &= \frac{S(t_1, t_2) \frac{\partial^2}{\partial t_1 \partial t_2} S(t_1, t_2)}{\left(\frac{\partial}{\partial t_1} S(t_1, t_2) \right) \left(\frac{\partial}{\partial t_2} S(t_1, t_2) \right)} \\
&= \frac{S(t_1, t_2)^2 \rho_1 \rho_2 f_1 f_2 \left(A^{2(\gamma-1)} + \frac{1-\gamma}{\sigma} A^{\gamma-2} \right)}{S(t_1, t_2)^2 \rho_1 \rho_2 f_1 f_2 A^{2(\gamma-1)}} \\
&= \frac{A^{2(\gamma-1)} + \frac{1-\gamma}{\sigma} A^{\gamma-2}}{A^{2(\gamma-1)}} \\
&= \frac{A^{2(\gamma-1)}}{A^{2(\gamma-1)}} + \frac{\frac{1-\gamma}{\sigma} A^{\gamma-2}}{A^{2(\gamma-1)}} \\
&= 1 + \frac{1-\gamma}{\sigma} \cdot \frac{A^{\gamma-2}}{A^{2(\gamma-1)}} \\
&= 1 + \frac{1-\gamma}{\sigma} A^{-\gamma}
\end{aligned}$$

Finalmente, lembramos que $A = 1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma}$ e substituindo, tem-se a expressão (4.17):

$$\vartheta^*(t_1, t_2) = 1 + \frac{1-\gamma}{\sigma} \left(1 + \frac{\rho_1 F_1(t_1) + \rho_2 F_2(t_2)}{\sigma} \right)^{-\gamma}$$

□

B.0.14 Afirmação (4.18)

O modelo pode ser estruturado em quatro níveis hierárquicos, como segue:

Nível 1: Fragilidade

Para cada indivíduo $j = 1, \dots, n$: com densidade $g(w_j | \mu, \sigma, \gamma)$ dada em (4.1)

$$W_j \stackrel{\text{i.i.d.}}{\sim} \text{PVF}(\mu, \sigma, \gamma), \quad \mathbb{E}[W_j] = 1, \quad \text{Var}(W_j) = \frac{1-\gamma}{\sigma}$$

Nível 2: Número de Causas Latentes

Condicional à fragilidade W_j , o número de causas latentes para cada tipo de evento segue uma distribuição Poisson dada por:

$$Z_{kj} | W_j = w_j \stackrel{\text{indep.}}{\sim} \text{Poisson}(w_j \rho_k), \quad k = 1, 2, \quad j = 1, \dots, n$$

com função de probabilidade: $\mathbb{P}(Z_{kj} = z_{kj} | w_j) = \frac{(w_j \rho_k)^{z_{kj}} e^{-w_j \rho_k}}{z_{kj}!}, \quad z_{kj} = 0, 1, 2, \dots$

Nível 3: Tempos de Promoção Latentes

Para cada indivíduo j e tipo de evento k , dado o número de causas $Z_{kj} = z_{kj}$, existem z_{kj} tempos de promoção latentes

$$T_{kj1}, T_{kj2}, \dots, T_{kj, z_{kj}} \mid Z_{kj} = z_{kj} \stackrel{\text{i.i.d.}}{\sim} F_{0k}(\cdot \mid \eta_k)$$

em que: $F_{0k}(\cdot \mid \eta_k)$ é a função de distribuição acumulada basal das causas latentes para o evento tipo k .

Nível 4: Tempos Observados e Censura

O tempo observado para o evento tipo k do indivíduo j é o mínimo dos tempos de promoção latentes:

$$T_{kj} = \min\{T_{kj1}, T_{kj2}, \dots, T_{kj, Z_{kj}}\}$$

com a convenção de que $T_{kj} = \infty$, se $Z_{kj} = 0$. Enquanto na presença de censura, observamos:

$$y_{kj} = \min(T_{kj}, C_{kj}), \quad \delta_{kj} = I(T_{kj} \leq C_{kj})$$

em que C_{kj} é o tempo de censura para o indivíduo j no evento k .

Agora, combinando todos os níveis hierárquicos, a verossimilhança completa isto é: *dados observados + variáveis latentes* é:

$$\begin{aligned} L(\vartheta \mid \mathcal{D}) &= \underbrace{\prod_{j=1}^n g(w_j \mid \mu, \sigma, \gamma)}_{\text{Nível 1: Fragilidade}} \\ &\times \underbrace{\prod_{k=1}^2 \prod_{j=1}^n \frac{(w_j \rho_k)^{Z_{kj}} e^{-w_j \rho_k}}{Z_{kj}!}}_{\text{Nível 2: Contagens Poisson}} \\ &\times \underbrace{\prod_{k=1}^2 \prod_{j=1}^n \left[\prod_{m=1}^{Z_{kj}} f_{0k}(t_{kjm} \mid \eta_k) \right]}_{\text{Nível 3: Tempos latentes (não observados)}} \\ &\times \underbrace{\prod_{k=1}^2 \prod_{j=1}^n \left[I(y_{kj} = \min\{t_{kj1}, \dots, t_{kj, Z_{kj}}\}) \times I(\delta_{kj} = I(y_{kj} \leq C_{kj})) \right]}_{\text{Nível 4: Observação}} \end{aligned}$$

Esta forma pode ser interpretada hierarquicamente como:

$$L(\vartheta \mid \mathcal{D}) = \left[\prod_{j=1}^n g(w_j \mid \mu, \sigma, \gamma) \right] \times \left[\prod_{k=1}^2 \prod_{j=1}^n \mathbb{P}(Z_{kj} \mid w_j, \rho_k) \right] \times \left[\prod_{k=1}^2 \prod_{j=1}^n f(y_{kj}, \delta_{kj} \mid Z_{kj}, \eta_k) \right]$$

em que:

$$\bullet \mathbb{P}(Z_{kj} \mid w_j, \rho_k) = \frac{(w_j \rho_k)^{Z_{kj}} e^{-w_j \rho_k}}{Z_{kj}!}$$

$$\bullet f(y_{kj}, \delta_{kj} | Z_{kj}, \eta_k) = [S_{0k}(y_{kj} | \eta_k)]^{Z_{kj} - \delta_{kj}} [Z_{kj} f_{0k}(y_{kj} | \eta_k)]^{\delta_{kj}}$$

logo, usando a informação de que o tempo observado y_{kj} é o mínimo dos tempos latentes e considerando a censura, podemos reescrever de forma mais compacta, exatamente como na equação (4.18):

$$\begin{aligned} L(\vartheta | \mathcal{D}) &= \prod_{k=1}^2 \prod_{j=1}^n [S_{0k}(y_{kj} | \eta_k)]^{Z_{kj} - \delta_{kj}} [Z_{kj} f_{0k}(y_{kj} | \eta_k)]^{\delta_{kj}} \\ &\times \prod_{k=1}^2 \exp \left\{ \sum_{j=1}^n [Z_{kj} \log(w_j \rho_k) - \log(Z_{kj}!) - w_j \rho_k] \right\} \\ &\times \prod_{j=1}^n g(w_j | \mu, \gamma, \sigma) \end{aligned}$$

De maneira geral, pode-se interpretar a verossimilhança completa como o esquema de Hierarquia Completa como na igualdade (4.18), é dada como

Nível 1: Fragilidade

$$W_j \sim \text{PVF}(\mu, \sigma, \gamma)$$

↓

Nível 2: Contagens Latentes

$$Z_{kj} | W_j = w_j \sim \text{Poisson}(w_j \rho_k)$$

↓

Nível 3: Tempos de Promoção

$$T_{kjm} | Z_{kj} \stackrel{\text{i.i.d.}}{\sim} F_{0k}(\cdot | \eta_k), m = 1, \dots, Z_{kj}$$

↓

Nível 4: Dados Observados

$$y_{kj} = \min\{T_{kj1}, \dots, T_{kj, Z_{kj}}\}, \quad \delta_{kj} = I(y_{kj} \leq C_{kj})$$

□

B.0.15 Prova do Teorema 1: Obtenção da Verossimilhança Observada

Vamos demonstrar como obter a função de verossimilhança baseada nos dados observados, $L(\vartheta | \mathcal{D}_{\text{obs}})$, a partir da verossimilhança completa $L(\vartheta | \mathcal{D})$. Da equação (4.18), a verossimilhança completa é:

$$\begin{aligned} L(\vartheta | \mathcal{D}) &= \prod_{k=1}^2 \prod_{j=1}^n [S_{0k}(y_{kj} | \eta_k)]^{Z_{kj} - \delta_{kj}} [Z_{kj} f_{0k}(y_{kj} | \eta_k)]^{\delta_{kj}} \\ &\times \prod_{k=1}^2 \exp \left\{ \sum_{j=1}^n [Z_{kj} \log(w_j \rho_k) - \log(Z_{kj}!) - w_j \rho_k] \right\} \\ &\times \prod_{j=1}^n g(w_j | \mu, \sigma, \gamma) \end{aligned}$$

em que $\mathcal{D} = (n, \mathbf{y}_1, \mathbf{y}_2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \mathbf{Z}_1, \mathbf{Z}_2, \mathbf{w})$ são os dados completos. Por outra parte, a verossimilhança dos dados observados $\mathcal{D}_{\text{obs}} = (n, \mathbf{y}_1, \mathbf{y}_2, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2)$ integrando as variáveis latentes não observadas Z_{kj} e w_j :

$$L(\boldsymbol{\vartheta} | \mathcal{D}_{\text{obs}}) = \int \sum_{\mathbf{Z}_1, \mathbf{Z}_2} L(\boldsymbol{\vartheta} | \mathcal{D}) d\mathbf{w}$$

Dado a independência entre indivíduos, podemos escrever:

$$L(\boldsymbol{\vartheta} | \mathcal{D}) = \prod_{j=1}^n L_j(\boldsymbol{\vartheta})$$

em que $L_j(\boldsymbol{\vartheta})$ é a contribuição do indivíduo j para a verossimilhança completa dada por

$$L_j(\boldsymbol{\vartheta}) = \left(\prod_{k=1}^2 [S_{0k}(y_{kj} | \boldsymbol{\eta}_k)]^{Z_{kj} - \delta_{kj}} [Z_{kj} f_{0k}(y_{kj} | \boldsymbol{\eta}_k)]^{\delta_{kj}} \times \prod_{k=1}^2 \exp \{ Z_{kj} \log(w_j \rho_k) - \log(Z_{kj}!) - w_j \rho_k \} \right) \times g(w_j | \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\gamma})$$

Agora vamos a integrar em relação a Z_{kj} , para cada indivíduo j , precisamos somar sobre todos os possíveis valores de Z_{1j} e Z_{2j} (de 0 a ∞):

$$L_j(\boldsymbol{\vartheta} | \mathcal{D}_{\text{obs}}) = \int_0^\infty g(w_j | \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\gamma}) \left[\sum_{z_1=0}^\infty \sum_{z_2=0}^\infty \prod_{k=1}^2 h_k(z_k, y_{kj}, \delta_{kj}, w_j) \right] dw_j$$

em que:

$$h_k(z_k, y_{kj}, \delta_{kj}, w_j) = [S_{0k}(y | \boldsymbol{\eta}_k)]^{z_k - \delta_{kj}} [z_k f_{0k}(y | \boldsymbol{\eta}_k)]^{\delta_{kj}} \times \frac{(w_j \rho_k)^{z_k} e^{-w_j \rho_k}}{z_k!}$$

Agora, vamos separar os casos de censura. Para cada evento k , existem dois casos possíveis:

- **Caso 1:** $\delta_{kj} = 0$ (censura)

$$\sum_{z=0}^\infty [S_{0k}(y_{kj} | \boldsymbol{\eta}_k)]^z \cdot \frac{(w_j \rho_k)^z e^{-w_j \rho_k}}{z!} = e^{-w_j \rho_k} \sum_{z=0}^\infty \frac{[w_j \rho_k S_{0k}(y_{kj} | \boldsymbol{\eta}_k)]^z}{z!}$$

A soma é a série de Taylor da exponencial:

$$\sum_{z=0}^\infty \frac{[w_j \rho_k S_{0k}(y_{kj} | \boldsymbol{\eta}_k)]^z}{z!} = \exp\{w_j \rho_k S_{0k}(y_{kj} | \boldsymbol{\eta}_k)\}$$

Portanto, para $\delta_{kj} = 0$:

$$\sum_{z=0}^\infty h_k(z, y_{kj}, 0, w_j) = \exp\{-w_j \rho_k [1 - S_{0k}(y_{kj} | \boldsymbol{\eta}_k)]\} = \exp\{-w_j \rho_k F_{0k}(y_{kj} | \boldsymbol{\eta}_k)\}$$

com $F_{0k} = 1 - S_{0k}$.

- **Caso 2:** $\delta_{kj} = 1$ (falha)

$$\sum_{z=1}^\infty [S_{0k}(y_{kj} | \boldsymbol{\eta}_k)]^{z-1} \cdot z f_{0k}(y_{kj} | \boldsymbol{\eta}_k) \cdot \frac{(w_j \rho_k)^z e^{-w_j \rho_k}}{z!}$$

veja que: $z/z! = 1/(z-1)!$:

$$= f_{0k}(y_{kj} | \eta_k) e^{-w_j \rho_k} \sum_{z=1}^{\infty} \frac{[w_j \rho_k S_{0k}(y_{kj} | \eta_k)]^{z-1}}{(z-1)!} \cdot w_j \rho_k$$

Fazendo $m = z - 1$, $m = 0, 1, 2, \dots$:

$$= f_{0k}(y_{kj} | \eta_k) w_j \rho_k e^{-w_j \rho_k} \sum_{m=0}^{\infty} \frac{[w_j \rho_k S_{0k}(y_{kj} | \eta_k)]^m}{m!}$$

A soma é $\exp\{w_j \rho_k S_{0k}(y_{kj} | \eta_k)\}$. Portanto:

$$\begin{aligned} \sum_{z=1}^{\infty} h_k(z, y_{kj}, 1, w_j) &= w_j \rho_k f_{0k}(y_{kj} | \eta_k) \exp\{-w_j \rho_k [1 - S_{0k}(y_{kj} | \eta_k)]\} \\ &= w_j \rho_k f_{0k}(y_{kj} | \eta_k) \exp\{-w_j \rho_k F_{0k}(y_{kj} | \eta_k)\} \end{aligned}$$

Combiando os dois eventos, para cada indivíduo j , temos quatro combinações possíveis de $(\delta_{1j}, \delta_{2j})$:

$$\sum_{z_1=0}^{\infty} \sum_{z_2=0}^{\infty} \prod_{k=1}^2 h_k(z_k, y_{kj}, \delta_{kj}, w_j) = \prod_{k=1}^2 \left[\sum_{z_k=0}^{\infty} h_k(z_k, y_{kj}, \delta_{kj}, w_j) \right]$$

então resulta que:

$$= \prod_{k=1}^2 \left[(w_j \rho_k f_{0k}(y_{kj} | \eta_k))^{\delta_{kj}} \exp\{-w_j \rho_k F_{0k}(y_{kj} | \eta_k)\} \right] \times w_j^{\delta_{1j} \delta_{2j}}$$

Para $\delta_{kj} = 1$, temos um fator w_j . Portanto, o expoente de w_j é $\delta_{1j} + \delta_{2j}$:

$$= (\rho_1 f_{01}(y_{1j} | \eta_1))^{\delta_{1j}} (\rho_2 f_{02}(y_{2j} | \eta_2))^{\delta_{2j}} \times w_j^{\delta_{1j} + \delta_{2j}} \times \exp\{-w_j [\rho_1 F_{01}(y_{1j} | \eta_1) + \rho_2 F_{02}(y_{2j} | \eta_2)]\}$$

Agora integrando em relação a ω_j

$$L_j(\vartheta | \mathcal{D}_{\text{obs}}) = (\rho_1 f_{01}(y_{1j} | \eta_1))^{\delta_{1j}} (\rho_2 f_{02}(y_{2j} | \eta_2))^{\delta_{2j}} \times I_j$$

em que:

$$I_j = \int_0^{\infty} w_j^{\delta_{1j} + \delta_{2j}} \exp\{-w_j [\rho_1 F_{01}(y_{1j} | \eta_1) + \rho_2 F_{02}(y_{2j} | \eta_2)]\} g(w_j | \mu, \sigma, \gamma) dw_j$$

esta integral I_j pode ser reconhecida como relacionada à transformada de Laplace da distribuição PVF. Lembrando que a transformada de Laplace de W é:

$$\mathcal{L}_W(s) = \mathbb{E}[e^{-sW}] = \exp\left\{-\frac{\sigma}{\gamma} \left[\left(1 + \frac{s}{\sigma}\right)^{\gamma} - 1\right]\right\}$$

Também precisamos de momentos fatoriais da forma $\mathbb{E}[W^m e^{-sW}]$, que podem ser obtidos derivando a transformada de Laplace:

$$\mathbb{E}[W^m e^{-sW}] = (-1)^m \frac{d^m}{ds^m} \mathcal{L}_W(s)$$

Calculando I_j para cada combinação de $(\delta_{1j}, \delta_{2j})$:

- **Caso 1:** $\delta_{1j} = 0, \delta_{2j} = 0$

$$I_j = \int_0^{\infty} \exp\{-w_j A_j\} g(w_j | \mu, \sigma, \gamma) dw_j = \mathcal{L}_W(A_j)$$

em que $A_j = \rho_1 F_{01}(y_{1j} | \eta_1) + \rho_2 F_{02}(y_{2j} | \eta_2)$. Portanto:

$$I_j = \exp\left\{-\frac{\sigma}{\gamma} \left[\left(1 + \frac{A_j}{\sigma}\right)^\gamma - 1\right]\right\}$$

- **Caso 2:** $\delta_{1j} = 1, \delta_{2j} = 0$ ou $\delta_{1j} = 0, \delta_{2j} = 1$

Para $\delta_{1j} + \delta_{2j} = 1$:

$$I_j = \int_0^{\infty} \omega_j \exp\{-\omega_j A_j\} g(w_j | \mu, \sigma, \gamma) dw_j = -\frac{d}{dA_j} \mathcal{L}_W(A_j)$$

Calculando a derivada:

$$-\frac{d}{dA_j} \mathcal{L}_W(A_j) = \mathcal{L}_W(A_j) \cdot \left(1 + \frac{A_j}{\sigma}\right)^{\gamma-1}$$

- **Caso 3:** $\delta_{1j} = 1, \delta_{2j} = 1$

Para $\delta_{1j} + \delta_{2j} = 2$:

$$I_j = \int_0^{\infty} w_j^2 \exp\{-w_j A_j\} g(w_j | \mu, \sigma, \gamma) dw_j = \frac{d^2}{dA_j^2} \mathcal{L}_W(A_j)$$

Calculando a segunda derivada:

$$\frac{d^2}{dA_j^2} \mathcal{L}_W(A_j) = \mathcal{L}_W(A_j) \left[\left(1 + \frac{A_j}{\sigma}\right)^{2(\gamma-1)} + \frac{1-\gamma}{\sigma} \left(1 + \frac{A_j}{\sigma}\right)^{\gamma-2} \right]$$

Logo, podemos unificar os três casos em uma única expressão. Observando os padrões:

- Para $\delta_{1j} + \delta_{2j} = 0$: fator $\mathcal{L}_W(A_j)$
- Para $\delta_{1j} + \delta_{2j} = 1$: fator $\mathcal{L}_W(A_j) \cdot \left(1 + \frac{A_j}{\sigma}\right)^{\gamma-1}$
- Para $\delta_{1j} + \delta_{2j} = 2$: fator $\mathcal{L}_W(A_j) \cdot \left[\left(1 + \frac{A_j}{\sigma}\right)^{2(\gamma-1)} + \frac{1-\gamma}{\sigma} \left(1 + \frac{A_j}{\sigma}\right)^{\gamma-2} \right]$

Reescrevendo o último caso:

$$\left(1 + \frac{A_j}{\sigma}\right)^{2(\gamma-1)} + \frac{1-\gamma}{\sigma} \left(1 + \frac{A_j}{\sigma}\right)^{\gamma-2} = \left(1 + \frac{A_j}{\sigma}\right)^{\gamma-2} \left[\left(1 + \frac{A_j}{\sigma}\right)^\gamma + \frac{1-\gamma}{\sigma} \right]$$

Observamos que todos os casos têm o fator comum $\mathcal{L}_W(A_j) = \exp\left\{-\frac{\sigma}{\gamma} \left[\left(1 + \frac{A_j}{\sigma}\right)^\gamma - 1\right]\right\}$.

Os fatores adicionais podem ser escritos como:

$$\left(1 + \frac{A_j}{\sigma}\right)^{(1-\gamma)(\delta_{1j}+\delta_{2j})-\gamma\delta_{1j}\delta_{2j}} \times \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{A_j}{\sigma}\right)^\gamma \right]^{\delta_{1j}\delta_{2j}}$$

Verificando:

- Se $\delta_{1j} = \delta_{2j} = 0$: expoente 0, potência = 1; último fator com expoente 0, = 1
- Se $\delta_{1j} = 1, \delta_{2j} = 0$: expoente $(1 - \gamma)(1) - \gamma(0) = 1 - \gamma$; último fator com expoente 0, = 1
- Se $\delta_{1j} = 0, \delta_{2j} = 1$: simétrico, expoente $1 - \gamma$
- Se $\delta_{1j} = \delta_{2j} = 1$: expoente $(1 - \gamma)(2) - \gamma(1) = 2 - 2\gamma - \gamma = 2 - 3\gamma$; e o último fator aparece com expoente 1

Esta expressão unificada os três casos. E finalmente, substituindo $A_j = \rho_1 F_{01}(y_{1j} | \eta_1) + \rho_2 F_{02}(y_{2j} | \eta_2)$ e multiplicando a contribuição de todos os indivíduos:

$$\begin{aligned}
L(\vartheta | \mathcal{D}_{\text{obs}}) &= \prod_{j=1}^n (\rho_1 f_1(y_{1j} | \eta_1))^{\delta_{1j}} (\rho_2 f_2(y_{2j} | \eta_2))^{\delta_{2j}} \\
&\times \prod_{j=1}^n \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\} \\
&\times \prod_{j=1}^n \left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}} \\
&\times \prod_{j=1}^n \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}}
\end{aligned}$$

□

Afirmção (4.20)

Para determinar a expressão para a Probabilidade de Cura com Covariáveis. Considere um indivíduo j imune, para o evento tipo k , ou seja, $p_{0kj} = \mathbb{P}(T_{kj} = \infty)$. No modelo de riscos latentes, um indivíduo é imune ao evento tipo k se não houver nenhuma causa latente para esse evento, isto é:

$$p_{0kj} = \mathbb{P}(N_{kj} = 0)$$

em que N_{kj} é o número de causas latentes para o evento tipo k do indivíduo j . Dado que as covariáveis são incorporadas através dos parâmetros de ρ_k . Para cada indivíduo j , tem-se

$$\rho_{kj} = \exp(\mathbf{x}_j^\top \boldsymbol{\beta}_k) = \exp(\eta_{kj})$$

sendo que $\eta_{kj} = \mathbf{x}_j^\top \boldsymbol{\beta}_k$ é o preditor linear. Por outra parte, a fragilidade w_j , o número de causas latentes segue uma distribuição Poisson

$$N_{kj} | W_j = w_j \sim \text{Poisson}(w_j \rho_{kj}) = \text{Poisson}(w_j \exp(\eta_{kj}))$$

Dado w_j , a probabilidade de não haver causas latentes é:

$$\mathbb{P}(N_{kj} = 0 | w_j) = \frac{(w_j \exp(\eta_{kj}))^0 e^{-w_j \exp(\eta_{kj})}}{0!} = e^{-w_j \exp(\eta_{kj})}$$

Agora marginalizando em relação à fragilidade, ou seja integrado sobre a distribuição da fragilidade:

$$p_{0kj} = \int_0^{\infty} \mathbb{P}(N_{kj} = 0 | w) g(w | \mu, \sigma, \gamma) dw = \int_0^{\infty} e^{-w \exp(\eta_{kj})} g(w | \mu, \sigma, \gamma) dw$$

Reconhecendo a Transformada de Laplace, a integral acima é exatamente a transformada de Laplace da variável de fragilidade W avaliada no ponto $s = \exp(\eta_{kj})$, logo

$$p_{0kj} = \mathbb{E} \left[e^{-\exp(\eta_{kj})W} \right] = \mathcal{L}_W(\exp(\eta_{kj}))$$

então substituindo $s = \exp(\eta_{kj})$, obtemos

$$p_{0kj} = \mathcal{L}_W(\exp(\eta_{kj})) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\exp(\eta_{kj})}{\sigma} \right)^{\gamma} - 1 \right] \right\} \quad \gamma \in (0, 1)$$

Alguns casos:

Caso $\gamma \rightarrow 0$ (fragilidade gama):

Usando o limite $\lim_{\gamma \rightarrow 0} \frac{a^{\gamma} - 1}{\gamma} = \log a$:

$$\lim_{\gamma \rightarrow 0} p_{0kj} = \exp \left\{ -\sigma \log \left(1 + \frac{\exp(\eta_{kj})}{\sigma} \right) \right\} = \left(1 + \frac{\exp(\eta_{kj})}{\sigma} \right)^{-\sigma}$$

Caso $\gamma = 1$ (independência):

$$p_{0kj} = \exp \left\{ -\sigma \left[\left(1 + \frac{\exp(\eta_{kj})}{\sigma} \right) - 1 \right] \right\} = \exp \{ -\exp(\eta_{kj}) \}$$

Caso $\sigma \rightarrow \infty$ (sem heterogeneidade):

Usando a expansão $\left(1 + \frac{\exp(\eta_{kj})}{\sigma} \right)^{\gamma} \approx 1 + \frac{\gamma \exp(\eta_{kj})}{\sigma}$ para σ grande:

$$p_{0kj} \approx \exp \left\{ -\frac{\sigma}{\gamma} \left[1 + \frac{\gamma \exp(\eta_{kj})}{\sigma} - 1 \right] \right\} = \exp \{ -\exp(\eta_{kj}) \}$$

que é o mesmo que o caso $\gamma = 1$.

□

Afirmção (4.21)

Vamos demonstrar como obter a função de risco marginal com covariáveis $h_k(t_{kj} | \mathbf{x})$ para o evento tipo k do indivíduo j , considerando a incorporação de covariáveis através do parâmetro ρ_{kj} . Dado que a função de risco para o evento tipo k é definida como:

$$h_k(t_{kj} | \mathbf{x}) = -\frac{d}{dt_{kj}} \log S_k(t_{kj} | \mathbf{x})$$

em que $S_k(t_{kj} | \mathbf{x})$ é a função de sobrevivência para o evento tipo k , condicional às covariáveis \mathbf{x} . Segue de (4.6) e condicionado às covariáveis, a função de sobrevivência para o evento tipo k do indivíduo j é:

$$S_k(t_{kj} | \mathbf{x}) = \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma} \right)^\gamma - 1 \right] \right\}$$

em que:

- $\rho_{kj} = \exp(\mathbf{x}_j^\top \beta_k)$ é o parâmetro de incidência com covariáveis
- $F_k(t_{kj} | \eta_k)$ é a função de distribuição acumulada das causas latentes
- $\gamma \in (0, 1]$ é o parâmetro da família PVF
- $\sigma > 0$ é o parâmetro de precisão

Agora, considere-se $U_k(t_{kj}) = \log S_k(t_{kj} | \mathbf{x})$, logo tem-se

$$U_k(t_{kj}) = -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma} \right)^\gamma - 1 \right]$$

Para simplificar, definimos uma função auxiliar:

$$B_{kj}(t_{kj}) = 1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma} \quad \text{então} \quad U_k(t_{kj}) = -\frac{\sigma}{\gamma} [B_{kj}(t_{kj})^\gamma - 1]$$

derivando $U_k(t_{kj})$ em relação a t_{kj}

$$\begin{aligned} \frac{dU_k}{dt_{kj}} &= -\frac{d}{dt_{kj}} \left\{ \frac{\sigma}{\gamma} [B_{kj}(t_{kj})^\gamma - 1] \right\} = -\frac{\sigma}{\gamma} \cdot \frac{d}{dt_{kj}} [B_{kj}(t_{kj})^\gamma] \\ &= -\frac{\sigma}{\gamma} \cdot \gamma B_{kj}(t_{kj})^{\gamma-1} \cdot \frac{dB_{kj}}{dt_{kj}} \\ &= -\sigma B_{kj}(t_{kj})^{\gamma-1} \cdot \frac{dB_{kj}}{dt_{kj}} \end{aligned} \tag{B.7}$$

por outra parte, calculamos

$$\frac{dB_{kj}}{dt_{kj}} = 0 + \frac{\rho_{kj}}{\sigma} \cdot \frac{dF_k(t_{kj} | \eta_k)}{dt_{kj}} = \frac{\rho_{kj} f_k(t_{kj} | \eta_k)}{\sigma} \tag{B.8}$$

Agora, substituindo (B.8) em (B.7):

$$\frac{dU_k}{dt_{kj}} = -\sigma B_{kj}(t_{kj})^{\gamma-1} \cdot \frac{\rho_{kj} f_k(t_{kj} | \eta_k)}{\sigma} = -\rho_{kj} f_k(t_{kj} | \eta_k) B_{kj}(t_{kj})^{\gamma-1}$$

em que $h_k(t_{kj} | \mathbf{x}) = -\frac{dU_k}{dt_{kj}}$ e $B_{kj}(t_{kj}) = 1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma}$

$$\begin{aligned} h_k(t_{kj} | \mathbf{x}) &= -[-\rho_{kj} f_k(t_{kj} | \eta_k) B_{kj}(t_{kj})^{\gamma-1}] = \rho_{kj} f_k(t_{kj} | \eta_k) B_{kj}(t_{kj})^{\gamma-1} \\ &= \rho_{kj} f_k(t_{kj} | \eta_k) \left(1 + \frac{\rho_{kj} F_k(t_{kj} | \eta_k)}{\sigma} \right)^{\gamma-1}, \quad k = 1, 2 \end{aligned}$$

e esta expressão representa exatamente a equação (4.21).

□

Afirmção (4.22)

Vamos determinar a função de verossimilhança baseada nos dados observados com incorporação de covariáveis, partindo da verossimilhança observada sem covariáveis (teorema 1) e substituindo ρ_k por $\rho_{kj} = \exp(\mathbf{x}_j^\top \beta_k)$. A verossimilhança observada sem covariáveis é:

$$\begin{aligned}
 L(\vartheta \mid D_{\text{obs}}) &= \prod_{j=1}^n \left((\rho_1 f_1(y_{1j} \mid \eta_1))^{\delta_{1j}} (\rho_2 f_2(y_{2j} \mid \eta_2))^{\delta_{2j}} \right) \\
 &\times \prod_{j=1}^n \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(y_{1j} \mid \eta_1) + \rho_2 F_2(y_{2j} \mid \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\} \\
 &\times \prod_{j=1}^n \left(1 + \frac{\rho_1 F_1(y_{1j} \mid \eta_1) + \rho_2 F_2(y_{2j} \mid \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}} \\
 &\times \prod_{j=1}^n \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\rho_1 F_1(y_{1j} \mid \eta_1) + \rho_2 F_2(y_{2j} \mid \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}}
 \end{aligned} \tag{B.9}$$

as covariáveis são incorporadas através dos parâmetros de incidência ρ_1 e ρ_2 , a relação entre as covariáveis e os parâmetros de incidência é especificada através de uma função de ligação log:

$$\rho_{kj} = \exp(\mathbf{x}_j^\top \beta_k), \quad k = 1, 2, \quad j = 1, \dots, n.$$

em que:

- $\mathbf{x}_j^\top = (x_{j1}, x_{j2}, \dots, x_{jp})$ é o vetor de covariáveis para o indivíduo j
- $\beta_k = (\beta_{k1}, \beta_{k2}, \dots, \beta_{kp})^\top$ é o vetor de coeficientes de regressão para o evento tipo k
- $\eta_{kj} = \mathbf{x}_j^\top \beta_k$ é o preditor linear

Agora, substituindo ρ_1 por $\rho_{1j} = \exp(\mathbf{x}_j^\top \beta_1)$ e ρ_2 por $\rho_{2j} = \exp(\mathbf{x}_j^\top \beta_2)$ em cada parte da verossimilhança dada na equação (B.9), que tem quatro fatores. Vamos incorporar as covariáveis em cada fator como segue:

Primeiro fator de (B.9)

$$\begin{aligned}
 \prod_{j=1}^n (\rho_1 f_1(y_{1j} \mid \eta_1))^{\delta_{1j}} (\rho_2 f_2(y_{2j} \mid \eta_2))^{\delta_{2j}} &\longrightarrow \prod_{j=1}^n (\exp(\mathbf{x}_j^\top \beta_1) f_1(y_{1j} \mid \eta_1))^{\delta_{1j}} (\exp(\mathbf{x}_j^\top \beta_2) f_2(y_{2j} \mid \eta_2))^{\delta_{2j}} \\
 &= \prod_{j=1}^n \prod_{k=1}^2 (\exp(\mathbf{x}_j^\top \beta_k) f_k(y_{kj} \mid \eta_k))^{\delta_{kj}}
 \end{aligned}$$

Segundo fator de (B.9)

$$\exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\rho_1 F_1(y_{1j} \mid \eta_1) + \rho_2 F_2(y_{2j} \mid \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\}$$

Substituindo as covariáveis:

$$\exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\}$$

Terceiro fator de (B.9)

$$\left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}}$$

Substituindo as covariáveis:

$$\left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}}$$

Quarto fator de (B.9)

$$\left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\rho_1 F_1(y_{1j} | \eta_1) + \rho_2 F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}}$$

Substituindo as covariáveis:

$$\left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}}$$

Portanto, multiplicando todos os fatores com o vetor de parâmetros $\beta = (\beta_1^\top, \beta_2^\top)^\top$, obtemos:

$$\begin{aligned} L(\beta, \eta_1, \eta_2, \gamma, \sigma | D_{\text{obs}}) &= \prod_{j=1}^n \prod_{k=1}^2 (\exp(\mathbf{x}_j^\top \beta_k) f_k(y_{kj} | \eta_k))^{\delta_{kj}} \\ &\times \prod_{j=1}^n \exp \left\{ -\frac{\sigma}{\gamma} \left[\left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma - 1 \right] \right\} \\ &\times \prod_{j=1}^n \left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^{(1-\gamma)(\delta_{1j} + \delta_{2j}) - \gamma \delta_{1j} \delta_{2j}} \\ &\times \prod_{j=1}^n \left[\frac{1-\gamma}{\sigma} + \left(1 + \frac{\exp(\mathbf{x}_j^\top \beta_1) F_1(y_{1j} | \eta_1) + \exp(\mathbf{x}_j^\top \beta_2) F_2(y_{2j} | \eta_2)}{\sigma} \right)^\gamma \right]^{\delta_{1j} \delta_{2j}} \end{aligned}$$

Observações Importantes

1. **Individualização:** Cada indivíduo j agora tem seus próprios parâmetros ρ_{1j} e ρ_{2j} baseados em suas covariáveis.
2. **Interpretação:** O efeito das covariáveis é multiplicativo na escala da contagem latente média.

3. **Não proporcionalidade:** Mesmo com esta substituição, o modelo ainda permite não proporcionalidade dos riscos através do parâmetro γ .
4. **Flexibilidade:** Diferentes conjuntos de covariáveis podem ser usados para cada tipo de evento (β_1 e β_2 podem ser diferentes).

Esta demonstração mostra que a incorporação de covariáveis é feita de forma natural, substituindo os parâmetros escalares ρ_k por parâmetros individuais $\rho_{kj} = \exp(\mathbf{x}_j^\top \beta_k)$, preservando toda a estrutura do modelo original.

□

