

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**Um novo modelo de sobrevivência com fração de cura:
inferência, diagnóstico e aplicações**

Renan de Oliveira da Cruz

Dissertação de Mestrado do Programa Interinstitucional de
Pós-Graduação em Estatística (PIPGEs)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Renan de Oliveira da Cruz

Um novo modelo de sobrevivência com fração de cura:
inferência, diagnóstico e aplicações

Dissertação 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 Mestre em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística. *VERSÃO REVISADA*

Área de Concentração: Estatística

Orientador: Prof. Dr. Adriano Kamimura Suzuki

USP – São Carlos
Dezembro de 2025

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

C957n Cruz, Renan de Oliveira da
Um novo modelo de sobrevivência com fração de
cura: inferência, diagnóstico e aplicações / Renan
de Oliveira da Cruz; orientador Adriano Kamimura
Suzuki. -- São Carlos, 2025.
148 p.

Dissertação (Mestrado - Programa
Interinstitucional de Pós-graduação em Estatística)
-- Instituto de Ciências Matemáticas e de
Computação, Universidade de São Paulo, 2025.

1. Modelos de Sobrevivência. 2. Fração de cura.
3. Esquema de ativação latente. 4. Distribuição
Bilal. 5. Inferência Bayesiana. I. Kamimura Suzuki,
Adriano, orient. II. Título.

Renan de Oliveira da Cruz

A new cure rate survival model: inference, diagnostics, and applications

Dissertation 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 Master in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Prof. Dr. Adriano Kamimura Suzuki

**USP – São Carlos
December 2025**



UNIVERSIDADE FEDERAL DE SÃO CARLOS

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

Folha de Aprovação

Defesa de Dissertação de Mestrado do candidato Renan de Oliveira da Cruz, realizada em 23/10/2025.

Comissão Julgadora:

Prof. Dr. Adriano Kamimura Suzuki (ICMC/USP)

Prof. Dr. Erlandson Ferreira Saraiva (UFMS)

Prof. Dr. Cynthia Arantes Vieira Tojeiro (UFG)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

*A todos que caminharam comigo – professores e colegas de graduação e pós-graduação –,
deixo meu sincero agradecimento.*

À Dona Edna e ao Seu Valdeci, meus pais, esta dedicatória é especialmente dirigida.

Obrigado!

AGRADECIMENTOS

Agradeço, antes de tudo, aos meus pais, Valdeci e Edna, pela paciência e pelo exemplo diário de trabalho e honestidade, sem os quais esta conquista não seria possível. Estendo meu reconhecimento aos meus irmãos pelo incentivo constante.

Aos amigos, pela torcida constante, pelas conversas que renovaram o ânimo e pelos cafés que se tornaram abrigo nos dias mais exigentes. A presença de vocês tornou o caminho mais leve.

Registro minha profunda gratidão ao meu orientador, Prof. Dr. Adriano Kamimura Suzuki, pela orientação firme e generosa, pela disponibilidade incansável e pelo rigor acadêmico que elevou este trabalho.

Aos professores de graduação do Instituto de Ciências Matemáticas e de Computação, pela base sólida de formação em Estatística, pelas aulas inspiradoras e pelo incentivo à curiosidade científica.

Aos professores da pós-graduação da USP e da UFSCar, pelos ensinamentos, debates e oportunidades que ampliaram horizontes e qualificaram esta pesquisa. A todos que contribuíram direta ou indiretamente, deixo meu sincero muito obrigado.

“Try again. Fail again. Fail better.”
(Samuel Beckett)

RESUMO

CRUZ, R. O. **Um novo modelo de sobrevivência com fração de cura: inferência, diagnóstico e aplicações**. 2025. 148 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2025.

Modelos de sobrevivência com fração de cura são essenciais quando parte dos indivíduos não apresenta o evento de interesse ao longo do acompanhamento. Em muitos cenários, pressupõe-se a existência de causas latentes não observáveis que, ao ativarem-se, podem desencadear o evento; esse mecanismo de ativação pode associar o desfecho à primeira causa, à última ou a uma selecionada aleatoriamente. Neste trabalho, estudam-se modelos de fração de cura que combinam tempos de falha Weibull com a distribuição discreta de Bilal para o número de causas latentes. A estimação é conduzida por abordagens clássica e Bayesiana, e sua avaliação baseia-se em estudos de simulação e na aplicação a dados reais. A análise inclui o uso de resíduos de Cox–Snell como ferramenta diagnóstica, destacando o papel da avaliação residual nesse contexto. No enquadramento Bayesiano, empregam-se critérios de comparação de modelos e análise de influência por deleção de casos. Em conjunto, os resultados evidenciam a utilidade da família proposta para capturar heterogeneidade latente e diferentes mecanismos de ativação, oferecendo uma alternativa flexível e interpretável para aplicações biomédicas e/ou industriais.

Palavras-chave: Modelos de Sobrevivência; Fração de cura; Esquema de ativação latente; Distribuição Bilal; Inferência Bayesiana.

ABSTRACT

CRUZ, R. O. **A new cure rate survival model: inference, diagnostics, and applications.** 2025. 148 p. Dissertação (Mestrado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2025.

Survival models with a cure rate are essential when a portion of individuals does not experience the event of interest during follow-up. In many settings, it is assumed that unobservable latent causes exist which, upon activation, may trigger the event; this activation mechanism can link the outcome to the first cause, the last cause, or one selected at random. In this work, we study cure rate models that combine Weibull failure times with the discrete Bilal distribution for the number of latent causes. Parameter estimation is conducted under both classical and Bayesian approaches, and performance is assessed through simulation studies and an application to real data. The analysis also employs Cox–Snell residuals as a diagnostic tool, underscoring the role of residual assessment in this context. Within the Bayesian framework, we use model comparison criteria and case-deletion influence analysis. Taken together, the results highlight the usefulness of the proposed family in capturing latent heterogeneity and distinct activation mechanisms, offering a flexible and interpretable alternative for biomedical and/or industrial applications.

Keywords: Survival Models; Cure Rate; Latent Activation Mechanism; Bilal Distribution; Bayesian Inference.

LISTA DE ILUSTRAÇÕES

Figura 1 – Função massa de probabilidade da distribuição Bilal para diferentes valores do parâmetro p	47
Figura 2 – Função de sobrevivência da distribuição Bilal para diferentes valores do parâmetro p	49
Figura 3 – Valor esperado, variância, assimetria e curtose da distribuição Bilal	52
Figura 4 – Função densidade de probabilidade da distribuição Weibull	54
Figura 5 – Função de sobrevivência da distribuição Weibull	56
Figura 6 – Média e variância pelos parâmetros da distribuição Weibull	58
Figura 7 – Função de risco da distribuição Weibull	59
Figura 8 – Resíduos de Cox-Snell do modelo bem especificado sem covariável	88
Figura 9 – Resíduos de Cox-Snell do modelo mal especificado sem covariável	89
Figura 10 – Divergência ψ para o conjunto de dados (a) no caso sem covariável	94
Figura 11 – Divergência ψ para o conjunto de dados (b) no caso sem covariável	94
Figura 12 – Divergência ψ para o conjunto de dados (e) no caso sem covariável	95
Figura 13 – Divergência ψ para o conjunto de dados (h) no caso sem covariável	95
Figura 14 – Estimador de Kaplan–Meier e curvas de sobrevivência dos modelos sem covariável ajustados por inferência clássica sob diferentes esquemas de ativação	112
Figura 15 – Resíduos do tipo Cox-Snell para os modelos ajustados sem covariável	113
Figura 16 – Divergência ψ para o conjunto de dados de melanoma	113
Figura 17 – Gráfico de traços	114
Figura 18 – Gráfico ACF	114
Figura 19 – Box plot para a distribuição <i>a posteriori</i> da proporções de curados	115
Figura 20 – Resíduos de Cox–Snell do modelo bem especificado com covariável	115
Figura 21 – Resíduos de Cox–Snell do modelo mal especificado com covariável	116
Figura 22 – Divergência ψ para o conjunto de dados (a) no caso com covariável	116
Figura 23 – Divergência ψ para o conjunto de dados (b) no caso com covariável	117
Figura 24 – Divergência ψ para o conjunto de dados (e) no caso com covariável	117
Figura 25 – Divergência ψ para o conjunto de dados (h) no caso com covariável	118
Figura 26 – Estimador de Kaplan–Meier e curvas de sobrevivência dos modelos com covariável ajustados por inferência clássica sob diferentes esquemas de ativação	119
Figura 27 – Resíduos do tipo Cox-Snell para os modelos ajustados com covariável	120
Figura 28 – Divergência ψ para o conjunto de dados de melanoma com covariável	120
Figura 29 – Gráfico de traços para o modelo com covariável	121

Figura 30 – Gráfico ACF para o modelo com covariável	121
Figura 31 – Box plot para a distribuição <i>a posteriori</i> da proporções de curados do modelo com covariável	122
Figura 32 – Divergência ψ para o conjunto de dados (c) no caso sem covariável	135
Figura 33 – Divergência ψ para o conjunto de dados (d) no caso sem covariável	136
Figura 34 – Divergência ψ para o conjunto de dados (f) no caso sem covariável	136
Figura 35 – Divergência ψ para o conjunto de dados (g) no caso sem covariável	137
Figura 36 – Divergência ψ para o conjunto de dados (c) no caso com covariável	137
Figura 37 – Divergência ψ para o conjunto de dados (d) no caso com covariável	138
Figura 38 – Divergência ψ para o conjunto de dados (f) no caso com covariável	138
Figura 39 – Divergência ψ para o conjunto de dados (g) no caso com covariável	139

LISTA DE TABELAS

Tabela 1 – Função de sobrevivência e função densidade para modelos de fração de cura sob diferentes mecanismos de ativação	65
Tabela 2 – Estudo de simulação abordagem clássica sem covariáveis sob o esquema de primeira ativação	86
Tabela 3 – Estudo de simulação abordagem clássica sem covariáveis sob o esquema de última ativação	86
Tabela 4 – Estudo de simulação abordagem clássica sem covariáveis sob o esquema de ativação aleatória	87
Tabela 5 – Estudo de simulação abordagem Bayesiana sem covariáveis sob o esquema de primeira ativação	90
Tabela 6 – Estudo de simulação abordagem Bayesiana sem covariáveis sob o esquema de última ativação	90
Tabela 7 – Estudo de simulação abordagem Bayesiana sem covariáveis sob o esquema de ativação aleatória	90
Tabela 8 – Média, DP e DR das estimativas dos parâmetros obtidas a partir do ajuste do modelo BWeibcr sem covariável sob o esquema de última ativação	92
Tabela 9 – Critérios Bayesianos de seleção de modelos no caso sem covariável	92
Tabela 10 – Medidas de divergência Bayesiana no caso sem covariável	93
Tabela 11 – Medidas resumo da variável de interesse	97
Tabela 12 – Estimativas de máxima verossimilhança do modelo sem covariáveis	97
Tabela 13 – Proporção de cura: $\pi = 2p^3 - 3p^2 + 1$	98
Tabela 14 – Métricas dos modelos ajustados sem a presença de covariáveis	98
Tabela 15 – Resumo a <i>posteriori</i> dos parâmetros do modelo BWeibcr sob o esquema de primeira ativação	99
Tabela 16 – Estudo de simulação abordagem clássica com covariáveis sob o esquema de primeira ativação	101
Tabela 17 – Estudo de simulação abordagem clássica com covariáveis sob o esquema de última ativação	102
Tabela 18 – Estudo de simulação abordagem clássica com covariáveis sob o esquema de ativação aleatória	102
Tabela 19 – Estudo de simulação abordagem Bayesiana com covariável sob o esquema de primeira ativação	104

Tabela 20 – Estudo de simulação abordagem Bayesiana com covariável sob o esquema de última ativação	104
Tabela 21 – Estudo de simulação abordagem Bayesiana com covariável sob o esquema de ativação aleatória	105
Tabela 22 – Média, DP e DR das estimativas dos parâmetros obtidas a partir do ajuste do modelo BWeibcr com covariável sob o esquema de última ativação	106
Tabela 23 – Critérios Bayesianos de seleção de modelos caso com covariável	106
Tabela 24 – Medidas de divergência Bayesiana no caso com covariável	107
Tabela 25 – Estimativas de máxima verossimilhança do modelo	108
Tabela 26 – Proporção de cura: $\pi(x) = 2p(x)^3 - 3p(x)^2 + 1$	109
Tabela 27 – Métricas dos modelos ajustados	109
Tabela 28 – Resumo <i>a posteriori</i> dos parâmetros do modelo BWeibcr sob o esquema de primeira ativação	111
Tabela 29 – Medidas de divergência para os dados de melanoma no modelo sem covariável	142
Tabela 30 – Medidas de divergência para os dados de melanoma no modelo com covariável	143

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Organização do Trabalho	24
2	ANÁLISE DE SOBREVIVÊNCIA: REVISÃO E CONCEITOS	25
2.1	Variável de interesse	26
2.2	Função de sobrevivência	26
2.3	Função de taxa de falha ou de risco	28
2.4	Função de Taxa de Falha Acumulada	29
2.5	Tempo Médio de Vida	30
2.6	Vida Média Residual	31
2.7	Resumo das Relações entre Funções	32
2.8	Estimadores para a Função de Sobrevivência	33
2.8.1	<i>Estimador de Nelson–Aalen</i>	33
2.8.2	<i>Estimador de Kaplan–Meier</i>	33
2.9	Dados Censurados	34
2.10	Modelos com Fração de Cura	36
2.11	Modelos com Mecanismo de Ativação	38
2.11.1	<i>Primeira ativação</i>	40
2.11.2	<i>Última ativação</i>	41
2.11.3	<i>Ativação aleatória</i>	42
3	DISTRIBUIÇÕES UTILIZADAS	45
3.1	Distribuição Bilal	45
3.1.1	<i>Função Massa de Probabilidade</i>	46
3.1.2	<i>Função de Distribuição Acumulada e Função de Sobrevivência</i>	48
3.1.3	<i>Função Geradora de Probabilidades</i>	49
3.1.4	<i>Momentos Estatísticos</i>	50
3.1.5	<i>Função de Risco</i>	51
3.2	Distribuição Weibull	52
3.2.1	<i>Função Densidade de Probabilidade</i>	53
3.2.2	<i>Função de Distribuição Acumulada e Função de Sobrevivência</i>	55
3.2.3	<i>Momentos Estatísticos</i>	57
3.2.4	<i>Função de Risco</i>	58

4	O MODELO BILAL WEIBULL COM FRAÇÃO DE CURA	61
4.1	Estimação - Abordagem Clássica	67
4.1.1	<i>Estimador de Máxima Verossimilhança</i>	68
4.1.2	<i>Informação de Fisher</i>	69
4.1.3	<i>Intervalos de Confiança via Informação de Fisher</i>	69
4.1.4	<i>Considerações Computacionais</i>	70
4.1.5	<i>Resíduos de Cox–Snell</i>	70
4.2	Estimação - Abordagem Bayesiana	78
4.2.1	<i>Distribuições a priori e a posteriori</i>	79
4.2.2	<i>Crítérios de comparação de modelos</i>	80
4.2.3	<i>Análise Bayesiana de Influência por Deleção de Casos</i>	82
5	RESULTADOS E APLICAÇÕES	85
5.1	Modelo Sem Covariável	85
5.1.1	<i>Estudo de simulação</i>	85
5.1.2	<i>Aplicação a dados reais</i>	93
5.2	Modelo Com Covariável	100
5.2.1	<i>Estudo de simulação</i>	101
5.2.2	<i>Aplicação a dados reais</i>	107
6	CONSIDERAÇÕES FINAIS	123
	REFERÊNCIAS	125
	APÊNDICE A	
	APÊNDICE	131
A.1	Verossimilhança e funções score sob cada mecanismo de ativação	131
A.1.1	<i>Primeira ativação</i>	132
A.1.1.0.1	Funções escore.	132
A.1.2	<i>Última ativação</i>	133
A.1.2.0.1	Funções escore.	133
A.1.3	<i>Ativação aleatória</i>	133
A.1.3.0.1	Funções escore.	134
	APÊNDICE B	
	APÊNDICE	135
	APÊNDICE C	
	APÊNDICE	141
	APÊNDICE D	
	APÊNDICE	145

INTRODUÇÃO

A Estatística, mais do que uma ciência de métodos e técnicas, configura-se como uma linguagem para compreender a incerteza presente na natureza e na sociedade. Desde os primórdios do pensamento humano, o desejo de quantificar o desconhecido e prever o comportamento de fenômenos incertos impulsionou o desenvolvimento de ideias que culminaram no surgimento da Estatística como disciplina formal. A interpretação probabilística do mundo, portanto, não é meramente um recurso matemático, mas um reflexo da condição humana frente à imprevisibilidade. Nas palavras de Karl Pearson, “a Estatística é a gramática da ciência”, pois traduz em números e modelos a complexidade dos fenômenos observados (PEARSON, 1892).

De acordo com Ronald Fisher, a Estatística pode ser entendida como a ciência da inferência, permitindo extrair conclusões válidas a partir de dados observacionais sujeitos ao acaso (FISHER, 1955). David Freedman e outros autores contemporâneos complementam essa visão ao destacar que a Estatística não só quantifica incertezas, mas também orienta decisões baseadas em evidências, consolidando-se como ferramenta essencial em diversas áreas do conhecimento (KAYE; FREEDMAN, 2011).

Dentro desse escopo, a análise de sobrevivência se insere como um dos ramos mais relevantes da Estatística aplicada, não se restringindo apenas ao tempo até a ocorrência de um evento de interesse, mas também se estendendo a outras variáveis contínuas, como distância, exposição acumulada ou até dimensões espaciais. Embora a ênfase histórica esteja na dimensão temporal – como nos estudos de mortalidade e falhas mecânicas –, autores como (LAWLESS, 2011) já destacavam que a estrutura conceitual da análise de sobrevivência poderia ser aplicada a qualquer domínio ordenado no qual se observe a ocorrência de eventos. Colosimo e Giolo (COLOSIMO; GIOLO, 2006) também coloca que a estrutura de sobrevivência pode ser adaptada para outras variáveis além do tempo, como número de ciclos ou distância até um evento. Kleinbaum e Klein (KLEINBAUM; KLEIN, 1996) definem a análise de sobrevivência como um conjunto de métodos estatísticos voltados à modelagem da ocorrência de eventos em função do

tempo, incorporando adequadamente a censura (disponibilidade apenas parcial de informações sobre a variável de interesse). Lawless (LAWLESS, 2011), por sua vez, reforça que sua principal força reside na capacidade de tratar dados incompletos. Assim, a análise de sobrevivência consolida-se como uma abordagem metodológica versátil e robusta, capaz de lidar com múltiplos tipos de variáveis e estruturas de censura.

Em termos probabilísticos, a análise de sobrevivência busca modelar a função de sobrevivência $S(t) = \mathbb{P}(T > t)$, que expressa a probabilidade do tempo até o evento ser superior a um determinado valor t . Dessa forma, a função de risco, definida como $h(t) = f(t)/S(t)$, quantifica a taxa instantânea de ocorrência do evento, sendo central em modelos como o de Cox ou em abordagens paramétricas clássicas (COLOSIMO; GIOLO, 2006).

A modelagem desses tempos de sobrevivência é realizada a partir de distribuições adequadas para variáveis não negativas, como a exponencial, a Weibull, a lognormal e a gama. Cada uma dessas distribuições oferece interpretações distintas sobre o comportamento do tempo até o evento, sendo selecionadas conforme a natureza do fenômeno analisado. Ademais, os modelos de regressão, como a regressão de Cox (semiparamétrica) e os modelos paramétricos baseados em Weibull, são amplamente utilizados para incorporar covariáveis e explorar fatores explicativos associados ao tempo de sobrevivência (COX, 1972).

A censura, conforme supramencionado, é um aspecto singular da análise de sobrevivência. Ela ocorre quando não se tem a observação completa do tempo até o evento, seja porque ele ainda não aconteceu no fim do estudo (censura à direita), já havia ocorrido antes do início da observação (censura à esquerda), ou se sabe apenas que ocorreu dentro de um intervalo (censura intervalar). Também se considera a censura aleatória, quando os tempos de censura e de falha são variáveis aleatórias independentes (COLOSIMO; GIOLO, 2006). O tratamento adequado desses tipos de censura é vital para evitar vieses na estimação dos parâmetros de interesse.

A literatura clássica apresenta diversos modelos de sobrevivência. O modelo exponencial é o mais simples, assumindo risco constante ao longo do tempo. A distribuição Weibull, uma generalização do modelo exponencial, permite uma função de risco crescente ou decrescente, sendo bastante utilizada na engenharia de confiabilidade e em estudos médicos. Outras distribuições, como a lognormal e a gama, foram progressivamente integradas ao repertório da análise de sobrevivência pela sua flexibilidade e pelo modo singular que cada uma se ajusta a dados de uma forma geral.

Neste contexto, a distribuição Weibull continua desempenhando papel central. Originada por Waloddi Weibull em 1951 (WEIBULL, 1951) para aplicações na engenharia, ela é caracterizada por sua função de risco parametrizável e flexível. Pode modelar processos com taxa de falha constante, crescente ou decrescente, sendo particularmente útil em contextos biomédicos e industriais, com ampla aceitação na modelagem de sobrevivência.

Nos últimos anos, especial atenção tem sido dada aos modelos com fração de cura,

uma estrutura particular dos modelos de longa duração (IBRAHIM *et al.*, 2001; CASTRO; CANCHO; RODRIGUES, 2009). Esses modelos admitem que uma parcela da população jamais experimentará o evento de interesse, sendo essencial em contextos como oncologia, onde parte dos pacientes tratados não apresenta recidiva da doença. A evidência empírica para tal fração pode ser observada em curvas de Kaplan-Meier com cauda longa.

Dentre os modelos com fração de cura, destacam-se os que incorporam variáveis latentes para representar causas ocultas de ativação do evento. Nesse contexto, assume-se que cada indivíduo possui um número latente de fatores de risco, cuja ocorrência conjunta leva ao evento. Esse paradigma foi amplamente explorado por autores como (TSODIKOV; IBRAHIM; YAKOVLEV, 2003) e (COONER; BANERJEE; MCBEAN, 2006), e posteriormente desenvolvido em diversas variantes com distribuições como a logarítmica (SUZUKI *et al.*, 2017), a Bell (GALLARDO; CASTRO; GÓMEZ, 2021) e a Poisson (SUZUKI; CANCHO; LOUZADA, 2016).

Nos modelos com causas latentes, a variável adjacente que determina o número dessas causas desempenha um papel fundamental. Ela representa a complexidade subjacente ao processo de falha e possibilita a incorporação de heterogeneidades não observáveis diretamente na população em estudo. A especificação adequada dessa variável é crucial tanto para a qualidade do ajuste do modelo quanto para a interpretação correta dos parâmetros estimados. Por esse motivo, destaca-se a importância de fomentar pesquisas que explorem diferentes distribuições para modelar o número de causas latentes, contribuindo para o desenvolvimento de uma literatura robusta e diversificada. Essa variedade de abordagens amplia o leque de modelos disponíveis, permitindo que pesquisadores escolham aquele que melhor se adapta às especificidades dos dados e à natureza do fenômeno analisado.

Dentre as distribuições discretas candidatas à modelagem da quantidade de variáveis adjacentes, destaca-se, recentemente, a distribuição Bilal. Proposta originalmente por Abd-Elrahman (ABD-ELRAHMAN, 2013) como uma distribuição contínua, sua versão discreta foi apresentada por Altun *et al.* (ALTUN; EL-MORSHEDY; ELIWA, 2022), evidenciando propriedades como cauda longa, assimetria positiva e elevada flexibilidade. Essa distribuição tem sido aplicada com sucesso em processos autorregressivos de contagem e mostra-se promissora para modelagem de número de causas latentes devido à sua capacidade de acomodar dispersão e curtose variadas.

Neste trabalho, propõe-se um novo modelo de sobrevivência de longa duração, o modelo Bilal Weibull com fração de cura (BWeibcr), no qual a variável latente que representa o número de causas é modelada por uma distribuição Bilal discreta e os tempos associados seguem distribuição Weibull. A motivação central para a proposta do modelo reside na necessidade de estruturas mais flexíveis para capturar mecanismos de ativação de causas latentes em cenários com fração de cura e censura. Embora modelos clássicos – como o Poisson–Weibull – sejam amplamente empregados, eles impõem estruturas rígidas à distribuição do número de causas latentes. A distribuição Poisson, por exemplo, assume média igual à variância, o que pode ser inadequado

em situações com sobredispersão ou subdispersão, limitando a capacidade de representar de forma adequada o comportamento observado em muitos conjuntos de dados reais. Ao adotar a distribuição Bilal, o BWeibcr introduz maior capacidade de acomodar heterogeneidades populacionais, ampliando o escopo de modelagem e contribuindo para a literatura ao disponibilizar uma nova alternativa para modelagem de dados.

As propriedades teóricas do modelo, sua estimação e aplicações práticas serão discutidas ao longo desta dissertação.

1.1 Organização do Trabalho

O Capítulo 2 sistematiza o arcabouço conceitual indispensável à análise de sobrevivência, contemplando as funções de sobrevivência, de risco e de risco acumulado, os principais tipos de censura, bem como os estimadores não paramétricos da função de sobrevivência. Nesse escopo, fundamenta-se a motivação para modelos de longa duração com fração de cura também.

O Capítulo 3 descreve o substrato probabilístico do modelo proposto, ancorado na distribuição Bilal discreta para o número de causas latentes M e na distribuição Weibull para os tempos de falha latentes.

O Capítulo 4 apresenta o modelo Bilal Weibull com fração de cura (BWeibcr), descrevendo seus fundamentos teóricos e os procedimentos de estimação nas abordagens clássica e Bayesiana. No enfoque clássico, discutem-se o estimador de máxima verossimilhança, a informação observada de Fisher, a construção dos intervalos de confiança e a análise dos resíduos de Cox–Snell. Em seguida, desenvolve-se a formulação Bayesiana, incluindo a especificação das distribuições *a priori*, a obtenção das distribuições *a posteriori* e os critérios de comparação de modelos. Por fim, apresenta-se a teoria da análise Bayesiana de influência por deleção de casos, utilizada para avaliar o impacto de observações individuais sobre as estimativas e a estabilidade do modelo.

O Capítulo 5 apresenta os resultados dos estudos de simulação e das aplicações a dados reais, tanto para o modelo sem covariável quanto para o modelo com covariável.

Por fim, o Capítulo 6 consolida as contribuições teórico-metodológicas e empíricas do estudo, discute implicações práticas do modelo proposto, explicita suas limitações e delinea perspectivas de pesquisa futura.

ANÁLISE DE SOBREVIVÊNCIA: REVISÃO E CONCEITOS

Conforme discutido anteriormente, a análise de sobrevivência é um ramo da Estatística que tem como um de seus objetivos investigar o tempo até a ocorrência de um evento de interesse, como a falha de um equipamento, o óbito de um paciente, a recidiva de uma doença ou até mesmo a desistência de um curso. Essa abordagem mostra-se especialmente útil quando há variabilidade nos tempos de ocorrência desses eventos, considerando não apenas o tempo de acompanhamento dos indivíduos, mas também o status de ocorrência do evento – isto é, se o indivíduo já o experimentou ou ainda está sob risco no momento da análise – além de possíveis covariáveis que possam influenciar esse tempo (KLEINBAUM; KLEIN, 1996).

A principal ferramenta analítica nesse contexto é a função de sobrevivência, que expressa a probabilidade de um indivíduo “sobreviver” além de um determinado instante de tempo. A partir dessa função, derivam-se outras medidas fundamentais para a compreensão da dinâmica do evento, como a função de risco (ou taxa de falha), que representa a taxa instantânea de ocorrência do evento em um dado momento, e a função de risco acumulado, que sintetiza a intensidade do risco ao longo do período observado. Destaca-se, ainda, a mediana de sobrevivência, definida como o tempo no qual 50% dos indivíduos já experimentaram o evento de interesse, sendo uma medida robusta frente à presença de dados censurados (COLOSIMO; GIOLO, 2006).

Essas ferramentas possibilitam análises aprofundadas sobre o comportamento dos indivíduos ao longo do tempo e oferecem subsídios relevantes para a tomada de decisão em diversas áreas, como medicina, engenharia, economia e ciências sociais. Além disso, o tratamento adequado de dados censurados – observações em que o evento de interesse não foi completamente observado – constitui uma característica distintiva da análise de sobrevivência, exigindo técnicas específicas para assegurar estimativas válidas.

Diante disso, este capítulo tem como objetivo apresentar e discutir de forma detalhada as

principais funções envolvidas na análise de sobrevivência, bem como os conceitos estatísticos fundamentais associados a esse campo. A maior parte do conteúdo aqui tratado tem como principal referência a obra *Análise de Sobrevivência Aplicada*, de Colosimo e Giolo (COLOSIMO; GIOLO, 2006), amplamente reconhecida no cenário nacional e frequentemente adotada em cursos de graduação, pós-graduação e em pesquisas acadêmicas voltadas ao tema.

2.1 Variável de interesse

O tempo de falha, também denominado tempo até o evento ou tempo de sobrevivência, refere-se à variável aleatória que representa a duração entre um ponto de referência inicial (como o início do acompanhamento, instalação de um equipamento ou início de um tratamento) e a ocorrência de um evento específico de interesse, como a falha de um sistema, o óbito de um paciente ou a desistência de um programa. Essa variável pode assumir apenas valores não negativos e, em contextos aplicados, frequentemente apresenta distribuição assimétrica e sujeita à presença de censura. A modelagem do tempo de falha é o cerne da análise de sobrevivência, sendo fundamental para a compreensão dos fatores que influenciam a ocorrência do evento e para a previsão de sua probabilidade ao longo do tempo (LAWLESS, 2011; KLEINBAUM; KLEIN, 1996).

Conforme definido por Colosimo e Giolo (2006), seja $T \subseteq \mathbb{R}_+$ uma variável aleatória com suporte não negativo, contínua ou discreta, representando a variável de interesse que, nesse contexto, geralmente corresponde ao tempo de falha ou ao tempo até a ocorrência de determinado evento.

É importante destacar que uma variável aleatória com suporte não negativo é aquela cuja função densidade ou função de probabilidade é estritamente positiva apenas em valores não negativos de seu domínio. Formalmente, o suporte de uma função f é definido por:

$$\text{supp}(f) = \{t \in T : f(t) > 0\}.$$

No contexto da análise de sobrevivência, as variáveis de interesse, como o tempo de falha ou tempo até a ocorrência de um evento, assumem apenas valores positivos ou nulos. Portanto, o suporte das distribuições consideradas neste cenário é dado por:

$$\text{supp}(f) = \{t \in \mathbb{R}_+\}.$$

2.2 Função de sobrevivência

A função de sobrevivência é um dos conceitos centrais na análise de tempo até evento, sendo amplamente empregada em áreas como confiabilidade de sistemas, estudos biomédicos,

atuária e engenharia. Esta função descreve a probabilidade de que um determinado evento, como a falha de um componente, a morte de um paciente ou o encerramento de um contrato, ainda não tenha ocorrido até um tempo específico t .

Formalmente, seja T uma variável aleatória não negativa que representa o tempo até a ocorrência do evento de interesse. A função de sobrevivência, denotada por $S(t)$, é definida como:

$$S(t) = \mathbb{P}(T > t), \quad t \geq 0. \quad (2.1)$$

Em outras palavras, $S(t)$ representa a probabilidade de um indivíduo sobreviver além de um determinado tempo t . Por exemplo, se $S(24) = 0,65$, isso indica que existe uma probabilidade de 65% de que a unidade analisada permaneça em operação, ou o indivíduo esteja vivo, ou o contrato esteja vigente, após 24 unidades de tempo.

A função de sobrevivência apresenta diversas propriedades matemáticas importantes. Primeiramente, trata-se de uma função não crescente, ou seja, $S(t_1) \geq S(t_2)$ sempre que $t_1 < t_2$, o que reflete o decaimento da probabilidade de sobrevivência ao longo do tempo. Para variáveis contínuas com suporte não negativo, tem-se que $S(0) = 1$, já que no instante inicial o evento ainda não ocorreu, e $\lim_{t \rightarrow \infty} S(t) = 0$, assumindo que o evento eventualmente ocorrerá com probabilidade um.

Além disso, existe uma relação direta entre a função de sobrevivência e a função de distribuição acumulada $F(t) = \mathbb{P}(T \leq t)$, dada por:

$$S(t) = 1 - F(t). \quad (2.2)$$

Quando a variável aleatória T é absolutamente contínua com função densidade de probabilidade $f(t)$, a função de sobrevivência pode ser expressa como:

$$S(t) = \int_t^{\infty} f(u) du. \quad (2.3)$$

De forma recíproca, a função densidade pode ser obtida a partir da derivada negativa da função de sobrevivência. De (2.2), tem-se $F(t) = 1 - S(t)$. Assim, ao derivar $F(t)$ em relação a t , obtém-se:

$$f(t) = \frac{d}{dt}F(t) = -\frac{d}{dt}S(t). \quad (2.4)$$

Outra propriedade relevante da função de sobrevivência está relacionada ao cálculo do valor esperado de T , isto é, a esperança matemática do tempo até o evento. Quando $\mathbb{E}(T)$ existe e o evento ocorre quase certamente, o valor esperado pode ser obtido diretamente a partir da

função de sobrevivência. De fato, a partir da definição de esperança para uma variável aleatória contínua $T \geq 0$ com densidade $f(t)$, tem-se:

$$\begin{aligned}
 \mathbb{E}(T) &= \int_0^{\infty} t f(t) dt = \int_0^{\infty} t (-S'(t)) dt \\
 &= \left[-tS(t) \right]_0^{\infty} + \int_0^{\infty} S(t) dt \\
 &= 0 + \int_0^{\infty} S(t) dt \quad (\text{pois } tS(t) \rightarrow 0 \text{ quando } t \rightarrow \infty) \\
 &= \int_0^{\infty} S(t) dt. \tag{2.5}
 \end{aligned}$$

Essa formulação é particularmente útil em aplicações práticas, pois permite o cálculo do tempo médio até o evento sem a necessidade explícita da função densidade. Todas essas propriedades tornam a função de sobrevivência uma ferramenta fundamental para descrever o comportamento temporal de fenômenos aleatórios, possibilitando o desenvolvimento de modelos estatísticos robustos.

2.3 Função de taxa de falha ou de risco

A função de taxa de falha, também conhecida como função de risco ou *hazard function*, é uma medida que descreve a propensão instantânea de ocorrência do evento de interesse em um determinado instante t , sob a condição de que tal evento ainda não tenha ocorrido até esse momento. Essa função desempenha um papel fundamental na análise de sobrevivência, especialmente em contextos onde se deseja compreender não apenas o tempo até o evento, mas a forma como o risco se distribui ao longo do tempo.

Conforme apresentado por [Colosimo e Giolo \(2006\)](#), a construção da função de taxa de falha pode ser motivada pela análise da probabilidade de ocorrência do evento no intervalo $[t_1, t_2)$, a qual é dada por:

$$\mathbb{P}(t_1 \leq T < t_2) = S(t_1) - S(t_2), \tag{2.6}$$

em que $S(t)$ representa a função de sobrevivência. Dado que o evento ainda não ocorreu até o instante t_1 , a probabilidade condicional de falha no intervalo $[t_1, t_2)$ é expressa por:

$$\frac{S(t_1) - S(t_2)}{S(t_1)}. \tag{2.7}$$

Dividindo essa quantidade pelo comprimento do intervalo $t_2 - t_1$, obtém-se a taxa de falha média no período considerado:

$$\frac{S(t_1) - S(t_2)}{(t_2 - t_1)S(t_1)}. \quad (2.8)$$

Ao fazer $t = t_1$ e $t_2 = t + \Delta t$, essa expressão assume a forma:

$$\frac{S(t) - S(t + \Delta t)}{\Delta t \cdot S(t)}. \quad (2.9)$$

Tomando o limite quando $\Delta t \rightarrow 0$, obtém-se a definição da função de taxa de falha instantânea:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t \mid T > t)}{\Delta t} = \frac{f(t)}{S(t)}, \quad (2.10)$$

em que $f(t)$ é a função densidade de probabilidade da variável aleatória T . A função $h(t)$ expressa, portanto, a taxa com que o evento ocorre exatamente no instante t , dado que não ocorreu anteriormente. Trata-se de uma medida local e intensiva, capaz de revelar nuances importantes sobre a dinâmica do processo de risco ao longo do tempo.

É importante observar que, embora diferentes campos do conhecimento adotem terminologias distintas para essa função – como taxa de falha, função de risco, hazard rate ou intensidade de falha – todas essas expressões se referem à mesma entidade matemática. A escolha da nomenclatura depende do contexto disciplinar, sendo comum a utilização de “função de risco” em estudos biomédicos e “taxa de falha” em aplicações de confiabilidade e engenharia.

2.4 Função de Taxa de Falha Acumulada

A função de taxa de falha acumulada, também chamada de função de risco acumulada, é uma medida que quantifica o risco total acumulado até um dado instante de tempo. Essa função representa a soma contínua das taxas de falha instantâneas ao longo do tempo, oferecendo uma visão integrada da intensidade com que ocorrem os eventos ao longo do período de observação. Seu uso é recorrente em contextos de confiabilidade, medicina, demografia e ciências atuariais, onde é fundamental avaliar como o risco se comporta cumulativamente até um determinado ponto no tempo.

A função de taxa de falha acumulada é especialmente útil quando se deseja compreender não apenas o risco pontual de falha, como faz a função $h(t)$, mas a quantidade total de risco enfrentado até o tempo t . Em termos práticos, ela fornece uma medida agregada da vulnerabilidade de um indivíduo ou sistema à ocorrência do evento de interesse desde o início do acompanhamento até o tempo t .

Seja $h(t)$ a função de taxa de falha instantânea de uma variável aleatória T , a função de taxa de falha acumulada, denotada por $H(t)$, é definida como:

$$H(t) = \int_0^t h(u) du. \quad (2.11)$$

Essa expressão representa a integral da função de risco instantâneo desde o tempo zero até o tempo t , agregando todas as contribuições de risco nesse intervalo. Diferente da função de sobrevivência, que mede a probabilidade de ainda não ter ocorrido o evento até certo tempo, a função $H(t)$ descreve o risco total acumulado nesse período. Embora não seja uma probabilidade no sentido clássico (isto é, não está limitada ao intervalo $[0, 1]$), sua interpretação está diretamente ligada à intensidade de ocorrência do evento.

A função de risco acumulada desempenha papel crucial na modelagem estatística. Em particular, ela permite expressar a função de sobrevivência por meio da relação:

$$S(t) = \exp(-H(t)), \quad (2.12)$$

o que evidencia a conexão entre a intensidade acumulada de risco e a probabilidade de sobrevivência. Assim, altos valores de $H(t)$ implicam baixa sobrevivência no tempo t , refletindo maior risco acumulado.

De modo geral, a análise da função $H(t)$ possibilita identificar períodos críticos de risco, comparar grupos de indivíduos ou componentes, e embasar decisões sobre intervenções, substituições ou monitoramentos mais intensivos. Sua interpretação clara e sua aplicabilidade transversal em diferentes disciplinas reforçam sua relevância como ferramenta estatística de primeira ordem na análise de tempo até evento.

2.5 Tempo Médio de Vida

O tempo médio de vida, também conhecido como valor esperado do tempo até o evento, é uma medida fundamental na análise de sobrevivência, pois resume em um único número a duração média até a ocorrência do evento de interesse. Essa medida permite quantificar, em termos esperados, o tempo que um indivíduo, sistema ou componente permanece em funcionamento ou livre do evento.

Seja T uma variável aleatória não negativa que representa o tempo até a ocorrência de um evento. O tempo médio de vida, denotado por $\mathbb{E}(T)$, pode ser expresso em termos da função de sobrevivência $S(t)$ como:

$$\mathbb{E}(T) = \int_0^{\infty} S(t) dt. \quad (2.13)$$

Do ponto de vista interpretativo, $\mathbb{E}(T)$ representa o tempo médio que se espera que um indivíduo ou sistema “sobreviva” antes da ocorrência do evento. Por exemplo, em estudos clínicos, pode indicar a duração média de vida após o diagnóstico de uma condição; em engenharia, o

tempo médio de operação até a falha de um componente. Trata-se, portanto, de uma medida de tendência central que condensa informações da distribuição de sobrevivência em um único valor sintético.

É importante notar que a existência de $\mathbb{E}(T)$ depende do comportamento assintótico de $S(t)$. Em particular, para que a integral que define o tempo médio de vida seja finita, a função de sobrevivência deve decrescer suficientemente rápido à medida que $t \rightarrow \infty$. Em alguns casos, como em distribuições com caudas pesadas, o tempo médio de vida pode ser infinito, mesmo que a função de sobrevivência seja bem definida para todo t .

Além disso, o tempo médio de vida serve como parâmetro de comparação entre diferentes grupos ou tratamentos, permitindo avaliar, por exemplo, qual intervenção médica resulta em maior ganho em termos de tempo de vida esperado.

2.6 Vida Média Residual

A vida média residual é uma medida que quantifica o tempo médio adicional de sobrevivência esperado para um indivíduo, sistema ou componente, dado que este já sobreviveu até um tempo t . Essa medida é especialmente relevante em situações em que a observação não se inicia no tempo zero, ou em contextos de atualização dinâmica da expectativa de vida, como em manutenção preventiva, monitoramento clínico contínuo e análise de contratos condicionais.

Matematicamente, conforme [Colosimo e Giolo \(2006\)](#), seja T uma variável aleatória contínua representando o tempo até a ocorrência do evento de interesse. A vida média residual no tempo t , denotada por $VMR(t)$, é definida como:

$$VMR(t) = \mathbb{E}[T - t \mid T > t] = \frac{\int_t^\infty (u - t)f(u) du}{S(t)} = \frac{\int_t^\infty S(u) du}{S(t)}, \quad (2.14)$$

em que $f(u)$ é a função densidade de T e $S(t) = \mathbb{P}(T > t)$ é a função de sobrevivência. A primeira igualdade corresponde à definição condicional do valor esperado, enquanto a segunda decorre da substituição por $f(u) = -\frac{d}{du}S(u)$ e da integração por partes.

Do ponto de vista interpretativo, $VMR(t)$ representa a esperança do tempo restante até a ocorrência do evento, condicionado à sobrevivência até o instante t . Por exemplo, se em um estudo clínico $VMR(6) = 4$, isso indica que, dado que o paciente sobreviveu até o sexto mês, espera-se que ele sobreviva, em média, mais quatro meses. Em engenharia, tal métrica pode ser utilizada para definir planos de manutenção baseada em condição ou para estimar o tempo de vida útil restante de um equipamento em operação.

A vida média residual também apresenta propriedades matemáticas interessantes. Primeiramente, quando $t = 0$, recupera-se o tempo médio de vida:

$$VMR(0) = \mathbb{E}(T) = \int_0^{\infty} S(u) du. \quad (2.15)$$

Além disso, em distribuições com a chamada propriedade de *falta de memória* (como a distribuição exponencial), a vida média residual é constante ao longo do tempo, ou seja, $VMR(t) = \mathbb{E}(T)$ para todo $t \geq 0$. Isso não ocorre em distribuições com taxa de falha dependente do tempo, como a Weibull ou a lognormal, nas quais a função $VMR(t)$ pode crescer, decrescer ou apresentar comportamento não monotônico.

A análise da vida média residual fornece, assim, uma perspectiva dinâmica da sobrevivência, complementando as informações fornecidas pela função de sobrevivência e pela função de taxa de falha.

2.7 Resumo das Relações entre Funções

Após a apresentação das principais funções associadas à análise de sobrevivência, é oportuno destacar as relações fundamentais que as conectam. Tais relações são essenciais para o desenvolvimento teórico e prático dos modelos utilizados.

Pelo Teorema Fundamental do Cálculo, a função densidade $f(t)$ relaciona-se com a função de distribuição acumulada $F(t)$ e com a função de sobrevivência $S(t)$ da seguinte forma:

$$f(t) = \frac{d}{dt}F(t), \quad (2.16)$$

$$f(t) = -\frac{d}{dt}S(t). \quad (2.17)$$

A função de risco, por sua vez, pode ser mostrada como sendo a razão entre a densidade e a função de sobrevivência, podendo também ser expressa em termos do logaritmo desta última:

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log S(t). \quad (2.18)$$

Já a função de sobrevivência pode ser expressa em função da função de risco acumulada $H(t)$:

$$S(t) = \exp(-H(t)) = \exp\left(-\int_0^t h(u) du\right), \quad (2.19)$$

$$H(t) = -\log S(t). \quad (2.20)$$

Por fim, a densidade também pode ser expressa diretamente em função da função de risco e da função de risco acumulada:

$$f(t) = h(t)S(t) = h(t) \exp(-H(t)). \quad (2.21)$$

2.8 Estimadores para a Função de Sobrevivência

Em situações práticas, a função de sobrevivência não é conhecida e, portanto, precisa ser estimada a partir de dados amostrais. A presença de dados censurados é uma característica comum em estudos de tempo até evento, o que exige métodos específicos para a estimação não paramétrica da função de sobrevivência. Neste contexto, destacam-se dois estimadores amplamente utilizados: o estimador de Nelson–Aalen e o estimador de Kaplan–Meier.

2.8.1 Estimador de Nelson–Aalen

O estimador de Nelson–Aalen é um estimador não paramétrico da função de risco acumulada $H(t)$, a partir da qual a função de sobrevivência pode ser estimada por meio da relação $\widehat{S}(t) = \exp(-\widehat{H}(t))$. Este método é especialmente útil por sua simplicidade e propriedades assintóticas bem comportadas, mesmo na presença de censura à direita.

Seja $t_1 < t_2 < \dots < t_k$ os tempos ordenados de falha observados, d_j o número de falhas em t_j , e r_j o número de indivíduos em risco imediatamente antes de t_j . O estimador de Nelson–Aalen da função de risco acumulada é dado por:

$$\widehat{H}(t) = \sum_{t_j \leq t} \frac{d_j}{r_j}. \quad (2.22)$$

Conseqüentemente, a estimativa da função de sobrevivência é obtida como:

$$\widehat{S}(t) = \exp\left(-\sum_{t_j \leq t} \frac{d_j}{r_j}\right). \quad (2.23)$$

O método foi proposto por Wayne Nelson em 1969 no contexto da confiabilidade de sistemas, e formalmente discutido em (NELSON, 1969), com importantes contribuições de Odd Aalen em 1978 no tratamento de processos de contagem e censura (AALEN, 1978).

2.8.2 Estimador de Kaplan–Meier

O estimador de Kaplan–Meier é o estimador não paramétrico mais amplamente utilizado para a função de sobrevivência em presença de censura à direita. Ele baseia-se no produto das probabilidades condicionais de sobrevivência em cada tempo de falha observado.

Seja novamente $t_1 < t_2 < \dots < t_k$ os tempos de falha distintos, com d_j falhas e r_j indivíduos em risco em t_j . O estimador de Kaplan–Meier é definido como:

$$\widehat{S}(t) = \prod_{t_j \leq t} \left(1 - \frac{d_j}{r_j}\right). \quad (2.24)$$

Esse estimador é uma função escada que decresce nos tempos de falha e permanece constante nos tempos de censura, capturando de forma eficiente o efeito da censura na estimação da sobrevivência.

O método foi introduzido por Edward L. Kaplan e Paul Meier em seu artigo de 1958 (KAPLAN; MEIER, 1958), sendo um marco na estatística aplicada à biomedicina e confiabilidade.

Ambos os estimadores têm interpretação intuitiva e sólida fundamentação teórica, sendo amplamente utilizados em softwares estatísticos e análises empíricas. A escolha entre eles pode depender do objetivo da análise: o estimador de Nelson–Aalen fornece diretamente a função de risco acumulado, enquanto o de Kaplan–Meier fornece diretamente a função de sobrevivência.

2.9 Dados Censurados

Em estudos de análise de sobrevivência, conforme antes mencionado, a variável de interesse é tipicamente o tempo até a ocorrência de um evento, como a falha de um equipamento, a recidiva de uma doença ou a morte de um paciente. No entanto, é comum que o tempo exato de ocorrência do evento não seja observado para todos os indivíduos da amostra, fenômeno conhecido como censura (KLEINBAUM; KLEIN, 1996; COLOSIMO; GIOLO, 2006; LOUZADA-NETO; MAZUCHELI; ACHCAR, 2001).

A censura pode ser definida como a limitação parcial da informação disponível sobre o tempo até o evento de interesse. Em outras palavras, para determinadas observações, conhece-se apenas uma desigualdade envolvendo o tempo do evento, ao invés de seu valor exato. A presença de censura não deve ser tratada como uma limitação, mas sim como uma característica intrínseca a muitos estudos prospectivos, especialmente em contextos médicos e industriais, e que deve ser incorporada corretamente à modelagem estatística.

Tipos de Censura

1. Censura à Direita: O tipo mais comum de censura é a censura à direita, que ocorre quando se sabe apenas que o evento de interesse ainda não ocorreu até certo instante. Formalmente, sejam T a variável aleatória que representa o tempo até o evento, e C a variável aleatória associada ao tempo de censura, tem-se:

$$t = \min(T, C), \quad \delta = \begin{cases} 1, & \text{se } T \leq C \quad (\text{evento observado}), \\ 0, & \text{se } T > C \quad (\text{censura à direita}). \end{cases} \quad (2.25)$$

Esse tipo de censura é típico quando um estudo é encerrado antes que todos os participantes tenham apresentado o evento de interesse.

2. Censura à Esquerda: Nesse caso, sabe-se que o evento de interesse ocorreu antes de certo ponto no tempo, mas o instante exato não é conhecido. Esse tipo de censura ocorre, por exemplo, em situações em que o evento já havia ocorrido antes do início da observação. Formalmente:

$$T \leq L, \quad \text{mas o valor exato de } T \text{ é desconhecido.} \quad (2.26)$$

Esse mecanismo é relevante em estudos retroativos ou quando o evento é detectado apenas indiretamente, como em exames laboratoriais que revelam que a infecção já está presente no momento da primeira observação.

3. Censura Intervalar: A censura intervalar ocorre quando se sabe que o evento ocorreu dentro de um intervalo de tempo, ou seja, $T \in (L, R)$, mas o tempo exato permanece desconhecido. Isso é comum em estudos com visitas periódicas, em que a ocorrência do evento é identificada apenas entre dois momentos de observação consecutivos:

$$L < T \leq R. \quad (2.27)$$

Um exemplo seria um ensaio clínico onde pacientes são examinados mensalmente, e o evento (por exemplo, recidiva tumoral) é detectado entre duas consultas.

Mecanismos de Geração da Censura

Além da classificação em termos do tipo de desigualdade, é importante considerar o mecanismo gerador da censura. Três formas principais são descritas na literatura:

- **Censura do Tipo I:** o estudo é encerrado em um tempo fixo pré-estabelecido.
- **Censura do Tipo II:** o estudo é encerrado após a observação de um número pré-determinado de eventos.
- **Censura Aleatória:** o tempo de censura varia aleatoriamente entre os indivíduos, sendo o caso mais comum na prática médica.

Esses mecanismos têm implicações distintas sobre a análise estatística, especialmente no contexto Bayesiano, onde a modelagem conjunta dos mecanismos de censura pode ser explicitamente incorporada à inferência (IBRAHIM *et al.*, 2001).

Em síntese, a censura é uma característica estrutural dos dados de sobrevivência e não uma falha de coleta de dados. Sua presença exige técnicas específicas para modelagem e

inferência, cuja incorporação adequada permite aproveitar ao máximo a informação disponível e garantir validade estatística às conclusões inferidas.

2.10 Modelos com Fração de Cura

Na análise de sobrevivência, os modelos com fração de cura – uma formulação particular dos modelos de longa duração – desempenham papel fundamental na investigação de dados em que o tempo até a ocorrência de um evento (como falha ou morte) é observado. Esses modelos são especialmente relevantes em contextos nos quais uma parcela dos indivíduos não experimenta o evento de interesse, mesmo após longos períodos de acompanhamento. Conforme discutido em (SOERJOMATARAM *et al.*, 2008), em pesquisas sobre câncer, um sobrevivente de longa duração é comumente definido como aquele que permanece vivo por pelo menos cinco anos após o diagnóstico. Ressalta-se, contudo, que esse intervalo de referência pode variar de acordo com as características da enfermidade e com outras variáveis clínicas ou sociodemográficas que influenciem o tempo de vida do paciente.

Segundo (BERKSON; GAGE, 1952), a concepção subjacente a esse termo decorre da ideia de que, diante de uma enfermidade letal como o câncer, os indivíduos que sobrevivem a um período suficientemente longo podem ser considerados “curados”. Todavia, é fundamental compreender que a taxa de sobrevivência em cinco anos não pode ser diretamente interpretada como a proporção de pacientes efetivamente curados. Não se pode assumir, por exemplo, que todos os pacientes não tratados fatalmente sucumbiriam nesse intervalo, nem tampouco que a totalidade dos indivíduos sem a doença estaria viva após cinco anos.

Nesse contexto, emerge o conceito de *fração de cura*, entendido como a proporção de indivíduos que nunca experimentarão o evento de interesse ao longo de um horizonte temporal suficientemente extenso. A incorporação desse componente nos modelos de sobrevivência permite distinguir, de maneira explícita, a parcela da população suscetível do subconjunto de indivíduos considerados imunes ao evento. Tal abordagem amplia a capacidade de interpretação dos resultados e fornece uma perspectiva mais abrangente sobre a dinâmica temporal dos fenômenos estudados, sendo particularmente útil em contextos biomédicos e industriais (IBRAHIM *et al.*, 2001).

A incorporação da fração de cura na modelagem de sobrevivência consiste em modificar a função de sobrevivência de modo a contemplar explicitamente a presença de indivíduos imunes ao evento de interesse. Nesse cenário, a função de sobrevivência populacional passa a ser representada como uma mistura entre a fração curada e a fração suscetível, sendo expressa, conforme abordado por Berkson e Gage (1952), da seguinte forma:

$$S_1(t) = \pi + (1 - \pi)S^*(t), \quad (2.28)$$

em que:

- $S_1(t)$: função de sobrevivência observada na população total;
- π : proporção da população considerada curada, ou seja, imune ao evento;
- $S^*(t)$: função de sobrevivência condicional ao subgrupo suscetível (não curado).

Ressalta-se que $S^*(t)$ é uma função de sobrevivência própria e, portanto, apresenta as propriedades típicas de uma função de sobrevivência já descritas anteriormente. Esse modelo pressupõe que a população é composta por dois subgrupos: um suscetível ao evento de interesse e outro que jamais o experimentará (grupo curado). Consequentemente, a função de sobrevivência da população total converge para a fração curada à medida que o tempo tende ao infinito. Matematicamente, essa relação pode ser expressa por:

$$\lim_{t \rightarrow \infty} S_1(t) = \pi. \quad (2.29)$$

em que $S_1(t) = \pi + (1 - \pi)S^*(t)$ é a função de sobrevivência da população total. Como $S^*(t) \rightarrow 0$ quando $t \rightarrow \infty$, o termo $(1 - \pi)S^*(t)$ desaparece assintoticamente, restando apenas a fração curada π .

A história dos modelos com fração de cura na análise de sobrevivência remonta à metade do século XX, quando os métodos tradicionais mostraram-se insuficientes diante da complexidade dos dados biomédicos. O trabalho pioneiro de Boag (1949) introduziu o conceito de modelo de mistura, representando a população como composta por dois subgrupos: um suscetível ao evento de interesse e outro considerado curado, ou seja, insuscetível à sua ocorrência. Poucos anos depois, Berkson e Gage (1952), em um estudo sobre tratamento de câncer, ampliaram essa formulação estatística ao propor explicitamente a incorporação da fração de cura em modelos de sobrevivência. Essa perspectiva rompeu com a hipótese de homogeneidade assumida nos modelos clássicos, ao reconhecer a heterogeneidade inerente às populações e permitir a descrição mais realista da dinâmica de ocorrência dos eventos, sobretudo em contextos epidemiológicos e clínicos.

Em trabalhos posteriores, diversos autores aprofundaram o estudo e a aplicação dos modelos com fração de cura, contribuindo para seu aperfeiçoamento e expansão metodológica. Farewell, em (FAREWELL, 1982) e (FAREWELL, 1986), propôs modelos baseados na distribuição Weibull, ampliando as possibilidades de adequação a diferentes contextos de dados. Ghitany e Maller (GHITANY; MALLER; ZHOU, 1994) também ofereceram uma contribuição relevante ao considerar a distribuição exponencial na formulação de modelos com fração de cura. Em um estudo aplicado, Sposto (SPOSTO, 2002) utilizou dados de crianças com câncer e empregou a análise de regressão de Cox, demonstrando a aplicabilidade do modelo em cenários clínicos reais. Já Perdona e Louzada-Neto (PERDONÁ; LOUZADA-NETO, 2011) propuseram um modelo

de risco geral que acomoda amplas famílias de modelos com taxa de cura, conferindo maior flexibilidade à abordagem. Além desses, diversos outros estudos têm explorado e expandido o tema, consolidando sua relevância na análise moderna de sobrevivência.

Essa percepção crucial da heterogeneidade populacional foi o catalisador que motivou a transição para modelos mais sofisticados. A necessidade emergente era a de desenvolver abordagens estatísticas que não apenas reconhecessem, mas também incorporassem essa complexidade na análise de sobrevivência. Assim, os modelos de fração de cura surgiram como uma resposta a essa demanda crescente por estruturas analíticas mais flexíveis, capazes de acomodar a diversidade de respostas temporais observadas na população em estudo.

2.11 Modelos com Mecanismo de Ativação

Com o avanço dos modelos de taxa de cura, diversos autores passaram a incorporar nas formulações a possibilidade de que o evento de interesse decorra de causas latentes e competitivas. Nesse contexto, a modelagem de dados de sobrevivência com fração de cura pode ser enriquecida por meio da incorporação de mecanismos de ativação, nos quais se assume que o número de causas que podem gerar o evento é uma variável aleatória, bem como os tempos de falha associados a cada uma delas. A ocorrência do evento é, assim, compreendida como consequência de um mecanismo de ativação entre essas causas latentes ou agentes ativadores, que seguem uma estrutura probabilística própria. Nessa abordagem, considera-se que o tempo até a ocorrência do evento não é diretamente modelado por uma única variável aleatória, mas sim pelo mínimo, máximo ou uma escolha aleatória de tempos associados a múltiplas causas latentes.

Os desenvolvimentos inaugurais nessa linha remontam a Yakovlev *et al.* (YAKOVLEV *et al.*, 1993) e a Yakovlev e Tsodikov (YAKOVLEV; TSODIKOV; ASSELAIN, 1996), que formalizaram modelos paramétricos de taxa de cura sob o paradigma do tempo de promoção, no qual o número de causas latentes é explicitamente modelado. Em sequência, Ibrahim *et al.* (IBRAHIM *et al.*, 2001) e Chen *et al.* (CHEN; IBRAHIM; SINHA, 1999; CHEN; IBRAHIM; SINHA, 2002) propuseram formulações Bayesianas e procedimentos de inferência, ampliando o arcabouço para cenários com covariáveis e estruturas hierárquicas. Avanços metodológicos adicionais foram introduzidos por Tsodikov *et al.* (TSODIKOV; IBRAHIM; YAKOVLEV, 2003), ao tratar explicitamente a ativação da primeira causa latente, e por Tournoud e Ecochard (TOURNOUD; ECOCHARD, 2007), ao considerar mecanismos de ativação mais flexíveis no contexto de doenças infecciosas. Em conjunto, tais contribuições consolidaram a fundamentação biológica e estatística dos modelos de cura baseados em causas latentes e estabeleceram bases robustas para aplicações biomédicas e epidemiológicas.

No contexto brasileiro, de Castro *et al.* (CASTRO; CANCHO; RODRIGUES, 2009; CASTRO; CANCHO; RODRIGUES, 2010), Ortega *et al.* (ORTEGA; CANCHO; PAULA, 2009)

e Louzada *et al.* ((LOUZADA; CANCHO; BARRIGA, 2012), (LOUZADA; CANCHO; YIQI, 2015)) aprofundaram o estudo da estrutura estocástica do número de causas latentes, propondo diferentes distribuições discretas para modelar a variável M , como a Poisson, binomial negativa, geométrica e séries de potências. Cancho *et al.* (CANCHO; RODRIGUES; CASTRO, 2011) e Kim *et al.* (KIM; CHEN; DEY, 2011) também exploraram essa vertente, com foco na análise Bayesiana e inferência via métodos de simulação.

A ideia central é considerar o número de causas latentes M – não observado – como uma variável discreta com suporte nos inteiros não negativos, cuja distribuição pode ser escolhida para refletir diferentes estruturas do processo gerador do evento. Condicionalmente a M , os tempos de promoção associados às causas, Z_j ($j = 1, \dots, M$), são assumidos independentes e identicamente distribuídos (i.i.d.) com função de distribuição $F(z) = 1 - S(z)$. O tempo observado T resulta da aplicação de um mecanismo de ativação sobre esses tempos ordenados, isto é, $T = \Phi(Z_{(1):(M)})$, onde $Z_{(1)} \leq \dots \leq Z_{(M)}$ denotam as estatísticas de ordem; exemplos incluem a primeira ativação $T = Z_{(1)}$, a última ativação $T = Z_{(M)}$ e a ativação aleatória $T = Z_R$ para algum R aleatório de $\{Z_j\}$, conforme discutido por Cooner *et al.* (COONER; BANERJEE; MCBEAN, 2006; COONER *et al.*, 2007). Note-se, ainda, que $M = 0$ caracteriza cura estrutural; por convenção, define-se $Z_0 = \infty$, isto é, Z_0 é uma variável aleatória degenerada em $+\infty$.

À luz do parágrafo precedente, em termos distintos, a estrutura do modelo em questão considera que o número de causas latentes M representa uma variável aleatória associada à quantidade de “agentes ativadores”, sendo cada agente caracterizado por um tempo de latência Z_i , para $i = 1, 2, \dots, M$. Esses tempos são assumidos como independentes e identicamente distribuídos segundo uma distribuição contínua com suporte não negativo. Nessa formulação, o tempo observado T é descrito como:

$$T = \begin{cases} \min\{Z_1, \dots, Z_M\}, & \text{primeira ativação,} \\ \max\{Z_1, \dots, Z_M\}, & \text{última ativação,} \\ Z_R, \quad R \sim \text{Unif}\{1, \dots, M\}, & \text{ativação aleatória,} \end{cases}$$

conforme discutido em (IBRAHIM *et al.*, 2001) e explorado por diversos autores posteriores, como Castro, Cancho e Rodrigues (2009), Suzuki, Cancho e Louzada (2016), Suzuki *et al.* (2017).

Diferentes distribuições têm sido propostas para modelar a variável M , cada uma resultando em distintas formas funcionais para a função de sobrevivência populacional. A distribuição de Poisson foi explorada em (SUZUKI; CANCHO; LOUZADA, 2016), enquanto a distribuição Logarítmica foi considerada em (SUZUKI *et al.*, 2017). A distribuição Bell foi recentemente incorporada nesse contexto, como apresentada em (CARREGARI, 2021), oferecendo maior flexibilidade no controle da variância de M . Adicionalmente, em (CASTRO; CANCHO; RODRIGUES, 2009), é proposta uma formulação em que o número de causas latentes M é modelado

por uma variável aleatória com distribuição binomial negativa.

O modelo aqui estudado combina dois blocos: (i) um componente discreto para o número de causas M , assumido com distribuição de Bilal (ALTUN; EL-MORSHEDY; ELIWA, 2022); e (ii) um componente contínuo para os tempos de falha, assumido com distribuição de Weibull (WEIBULL, 1951). Essa construção denomina-se, doravante, Bilal Weibull com fração de cura (BWeibcr).

No contexto da modelagem com causas latentes, conforme referido anteriormente, três esquemas de ativação têm sido considerados, cada um oferecendo uma representação distinta dos mecanismos subjacentes ao surgimento do evento de interesse. O primeiro deles, conhecido como *primeira ativação*, define o tempo até o evento como o menor dentre os tempos associados às causas latentes, isto é, $T = \min\{Z_1, \dots, Z_M\}$. Essa configuração é compatível com contextos clínicos, como o da recidiva de um câncer, nos quais a ocorrência do evento pode ser precipitada pela ativação de qualquer célula tumoral metastática residual, sendo suficiente que apenas uma delas se torne ativa para desencadear o desfecho. Em contraste, o esquema de *última ativação* assume que o evento ocorre apenas após a ativação de todas as causas latentes, formalmente $T = \max\{Z_1, \dots, Z_M\}$, como ocorre em sistemas de engenharia em que múltiplos componentes operam e a falha global só é observada quando todos os elementos deixam de funcionar. Já o modelo de *ativação aleatória* estabelece que o tempo até o evento é determinado por uma causa latente selecionada ao acaso entre as existentes, ou seja, $T = Z_R$ com $R \sim \text{Unif}(1, \dots, M)$, sendo apropriado para contextos em que não há conhecimento prévio sobre qual agente latente será responsável pela ocorrência do evento, como na análise da durabilidade de peças produzidas por diferentes fornecedores indistinguíveis do ponto de vista do risco. A escolha entre esses mecanismos deve refletir o conhecimento empírico sobre o fenômeno em estudo, pois ela afeta diretamente a formulação da função de sobrevivência populacional e, conseqüentemente, as inferências realizadas a partir do modelo estatístico.

2.11.1 Primeira ativação

No cenário de primeira ativação, assume-se que o evento de interesse ocorre assim que a primeira entre as causas latentes se torna ativa. Formalmente, define-se o tempo até o evento como:

$$T = \min\{Z_i; 1 \leq i \leq M\},$$

em que:

- T : denota a variável de interesse (tempo até o evento);
- M : representa o número de causas latentes associadas ao indivíduo;
- Z_i : é a variável aleatória que representa o tempo da i -ésima causa latente.

Assume-se que, condicionalmente a M , as variáveis Z_i são independentes e identicamente distribuídas (i.i.d.) e também independentes de M . Sob essas hipóteses, a função de sobrevivência populacional pode ser escrita em termos da função geradora de probabilidades (fgp) da variável M . De acordo com (RODRIGUES *et al.*, 2009), obtém-se:

$$S_{\text{pop}}(t) = \mathbb{P}(T > t) = \sum_{n=0}^{\infty} p_n \{S(t)\}^n = G_M(S(t)),$$

em que:

- $S(t) = \mathbb{P}(Z_i > t)$ é a função de sobrevivência comum a cada Z_i ;
- $p_n = \mathbb{P}(M = n)$ representa a função de massa de probabilidade de M ;
- $G_M(u) = \sum_{n=0}^{\infty} p_n u^n$ é a função geradora de probabilidades de M .

A demonstração da expressão acima é direta e decorre da definição da função de sobrevivência:

$$\begin{aligned} S_{\text{pop}}(t) &= \mathbb{P}(T > t) \\ &= \mathbb{P}(M = 0) + \mathbb{P}(Z_1 > t, \dots, Z_M > t \mid M \geq 1) \mathbb{P}(M \geq 1) \\ &= p_0 + \sum_{n=1}^{\infty} p_n \{\mathbb{P}(Z_i > t)\}^n \\ &= p_0 + \sum_{n=1}^{\infty} p_n S(t)^n \\ &= \sum_{n=0}^{\infty} p_n S(t)^n = G_M(S(t)). \end{aligned}$$

Essa formulação evidencia como a função de sobrevivência populacional resulta da composição entre a distribuição dos tempos latentes e a distribuição discreta do número de causas. No caso da primeira ativação, a função $S_{\text{pop}}(t)$ assume uma estrutura flexível que depende diretamente da fgp de M , permitindo incorporar diferentes escolhas de distribuições como Poisson, logarítmica, Bell ou Bilal.

2.11.2 Última ativação

Considerando o mecanismo de última ativação, o tempo até o evento de interesse é determinado pelo maior tempo entre as variáveis latentes associadas, isto é:

$$T = \max\{Z_i; 1 \leq i \leq M\},$$

em que Z_i representa, como anteriormente, o tempo da i -ésima causa latente e M o número total de causas latentes. Assume-se que as variáveis Z_i são independentes e identicamente distribuídas (i.i.d.) e também independentes da variável M .

Sob essas premissas, a função de sobrevivência populacional pode ser expressa como:

$$S_{\text{pop}}(t) = 1 + G_M(0) - G_M(F(t)),$$

em que:

- $F(t) = \mathbb{P}(Z_i \leq t)$ é a função de distribuição acumulada dos tempos latentes;
- $G_M(u) = \sum_{n=0}^{\infty} p_n u^n$ é a função geradora de probabilidades associada à variável discreta M ;
- $G_M(0) = \mathbb{P}(M = 0)$, de modo que $1 + G_M(0)$ assegura a normalização da função de sobrevivência.

Esse resultado decorre da caracterização do evento $T > t$ como equivalente à condição de que, se $M = m \geq 1$, ao menos uma das variáveis Z_i exceda t , o que é complementar ao caso em que todas as variáveis são menores ou iguais a t . Assim, a formulação considera a probabilidade acumulada até o tempo t sob a lógica de máxima latência entre os riscos.

2.11.3 Ativação aleatória

Por fim, no mecanismo de ativação aleatória, considera-se que o tempo até o evento de interesse é determinado pelo tempo de ativação de uma das causas latentes, escolhida aleatoriamente entre as M disponíveis. Mais precisamente, define-se:

$$T = Z_R, \quad \text{com } R \sim \text{Unif}\{1, 2, \dots, M\},$$

em que R é uma variável aleatória discreta, com distribuição uniforme condicional em $\{1, \dots, M\}$, e Z_i representa o tempo da i -ésima causa latente, conforme definido anteriormente.

Esse mecanismo modela situações em que todas as causas latentes têm igual probabilidade de se manifestar primeiro, independentemente de sua ordem ou natureza específica. Um exemplo seria o tempo até falha em um sistema composto por diversos subsistemas equivalentes, nos quais qualquer um pode falhar primeiro, e essa falha resulta na observação do evento. Trata-se de um modelo menos extremo do que os anteriores, pois não assume nem o cenário mais pessimista (mínimo) nem o mais conservador (máximo), mas sim uma escolha aleatória entre os riscos latentes.

Sob as hipóteses de independência entre M e os Z_i , bem como a suposição de que os Z_i são i.i.d., a função de sobrevivência populacional neste caso pode ser expressa da seguinte forma:

$$S_{\text{pop}}(t) = \mathbb{P}(M = 0) + \{1 - \mathbb{P}(M = 0)\} S(t),$$

em que $S(t) = \mathbb{P}(Z_i > t)$ representa a função de sobrevivência comum dos tempos latentes. Essa expressão mostra que a função de sobrevivência populacional resulta de uma média ponderada entre a probabilidade de cura $\mathbb{P}(M = 0)$ e a sobrevivência de um indivíduo suscetível cuja causa latente foi selecionada aleatoriamente. Assim, o modelo de ativação aleatória oferece uma representação flexível e intermediária entre os casos extremos de primeira e última ativação.

DISTRIBUIÇÕES UTILIZADAS

Adota-se a distribuição de Bilal para a variável M , que representa o número de causas latentes, e a distribuição Weibull para os tempos latentes Z_i . À luz dessas escolhas, este capítulo tem como objetivo apresentar, de forma sistemática, as principais propriedades dessas distribuições, bem como explicitar as definições fundamentais empregadas na formulação do modelo de sobrevivência desenvolvido nesta dissertação.

3.1 Distribuição Bilal

A modelagem de dados de contagem exige, frequentemente, distribuições discretas capazes de acomodar diferentes níveis de dispersão, assimetrias e formas de cauda. A distribuição de Poisson, embora clássica e amplamente utilizada, apresenta limitações importantes, sobretudo em contextos nos quais os dados observados exibem sobredispersão ou subdispersão. Diante disso, a literatura estatística tem investido no desenvolvimento de novas classes de distribuições discretas que ofereçam maior flexibilidade analítica e melhor capacidade de ajuste a dados reais.

Uma das estratégias recorrentes para essa finalidade consiste na discretização de distribuições originalmente contínuas, especialmente aquelas utilizadas na análise de sobrevivência. A partir dessa abordagem, diversas propostas foram introduzidas ao longo das últimas décadas, como as distribuições de Lindley Discreta, Weibull Discreta e Rayleigh Discreta, entre outras. Essas construções partem, em geral, da função de sobrevivência de uma variável contínua e aplicam transformações que garantem suporte discreto, mantendo ao mesmo tempo propriedades desejáveis como simetria, curtose controlável e formas explícitas para momentos e funções de risco.

Nesse cenário, a distribuição de Bilal surge como uma proposta contínua inovadora introduzida por [Abd-Elrahman \(2013\)](#), como caso particular de famílias de distribuições definidas pela mediana de amostras independentes, obtidas como polinômios finitos da função de sobrevi-

vência da distribuição original. Sua formulação proporciona flexibilidade adicional em relação a modelos tradicionais, permitindo representar uma ampla gama de comportamentos empíricos. Recentemente, [Altun, El-Morshedy e Eliwa \(2022\)](#) propuseram uma versão discretizada desse modelo, denominada distribuição de Bilal Discreta (DBL), derivada da função de sobrevivência da distribuição contínua original.

A distribuição DBL herda a flexibilidade da versão contínua, ao mesmo tempo em que oferece expressões fechadas para quantidades estatísticas relevantes, como média, variância, assimetria, curtose e função risco. Além disso, apresenta boa adaptabilidade a dados com caudas longas e estruturas de dispersão variáveis, sendo particularmente adequada para aplicações em séries temporais de contagem, processos autorregressivos inteiros e contextos de confiabilidade discreta.

A proposta de [Altun, El-Morshedy e Eliwa \(2022\)](#) contribui significativamente para o arcabouço de distribuições discretas modernas, reunindo estrutura matemática tratável e desempenho empírico satisfatório. As seções subsequentes apresentam a definição formal da distribuição de Bilal Discreta, suas propriedades estatísticas fundamentais que evidenciam sua utilidade na modelagem de dados de contagem.

3.1.1 Função Massa de Probabilidade

A construção do modelo discreto, conforme desenvolvido em ([ALTUN; EL-MORSHEDEY; ELIWA, 2022](#)), resulta em uma distribuição com suporte no conjunto dos inteiros não negativos, i.e., $\mathbb{N}_0 = \{0, 1, 2, \dots\}$, e dependente de um único parâmetro de escala $p \in (0, 1)$, relacionado à distribuição contínua original.

A função de massa de probabilidade (fmp) da distribuição DBL é expressa como:

$$f(x) = \mathbb{P}(X = x) = 2(p^3 - 1)p^{3x} - 3(p^2 - 1)p^{2x}, \quad x = 0, 1, 2, \dots \quad (3.1)$$

Essa expressão em forma fechada facilita a implementação computacional e a derivação de propriedades analíticas da distribuição. Nota-se que a fmp é composta por dois termos exponenciais dependentes de p , o que permite flexibilidade na forma da distribuição, sendo possível acomodar diferentes graus de assimetria e cauda.

Uma característica importante do modelo é que, para todo $p \in (0, 1)$, a função de massa de probabilidade é log-côncava, o que implica que a distribuição é unimodal e apresenta propriedades desejáveis para inferência estatística, como a existência de máximos globais na função de verossimilhança. Além disso, o comportamento da fmp sugere que a distribuição tende a ser assimétrica à direita e, para valores de p próximos de 1, tem-se como resultado caudas mais longas.

A estrutura da fmp torna a distribuição de Bilal Discreta uma alternativa promissora para

a modelagem de dados de contagem em que distribuições clássicas como Poisson ou binomial negativa não oferecem ajustes satisfatórios, especialmente em cenários de sobredispersão com assimetrias marcantes.

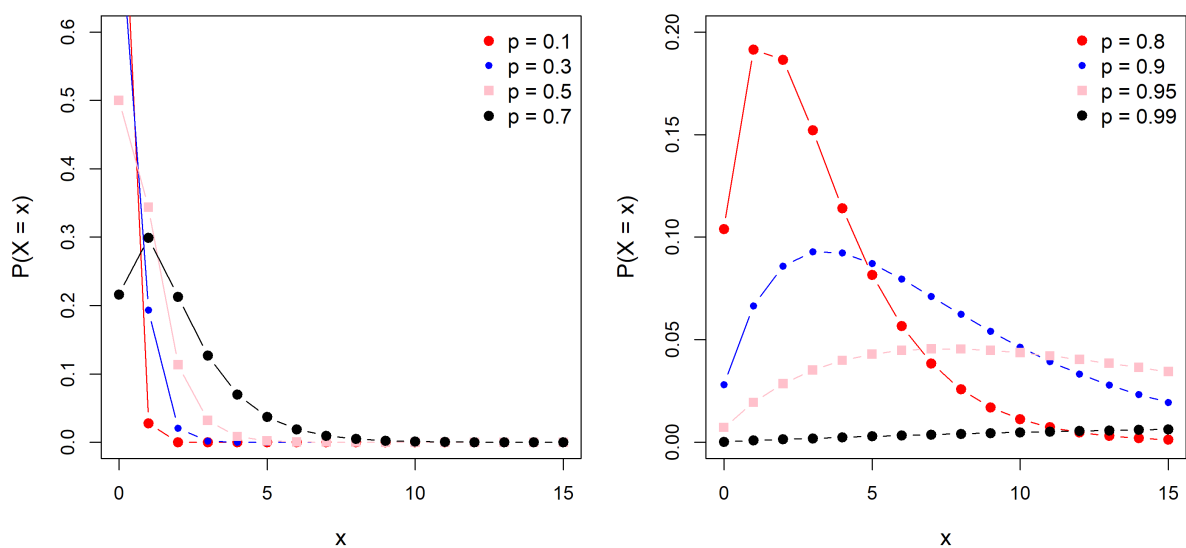
A Figura 1 apresenta a função de massa de probabilidade da distribuição de Bilal para diferentes valores do parâmetro p . O painel à esquerda exibe os comportamentos da distribuição para valores de p entre 0,1 e 0,7, enquanto o painel à direita contempla valores mais próximos de 1, variando entre 0,8 e 0,99.

Observa-se que, para valores menores de p , a distribuição apresenta uma concentração mais acentuada nos menores valores de m , evidenciando um decaimento exponencial rápido. Em particular, para $p = 0,1$, a massa de probabilidade está fortemente concentrada em $m = 0$, com valores subsequentes rapidamente se tornando desprezíveis. À medida que p aumenta, a cauda da distribuição se alonga e a variabilidade se eleva.

Já para valores de p mais próximos de 1, como ilustrado no painel da direita, a distribuição se torna mais dispersa e apresenta comportamentos não monotônicos. Para $p = 0,8$, por exemplo, a função de massa atinge seu pico em $m = 1$, indicando maior probabilidade de múltiplas causas latentes contribuírem para o evento. Quando $p = 0,95$ e $p = 0,99$, observa-se uma distribuição quase uniforme com pequenas probabilidades, sinalizando a presença de uma longa cauda e a possibilidade de ocorrência de um número maior de causas ativadoras.

Esses resultados demonstram a flexibilidade da distribuição de Bilal para modelar diferentes cenários de heterogeneidade latente, sendo possivelmente útil em contextos de análise de sobrevivência com taxa de cura e mecanismos de ativação competitiva.

Figura 1 – Função massa de probabilidade da distribuição Bilal para diferentes valores do parâmetro p



Fonte: Elaborada pelo autor.

3.1.2 Função de Distribuição Acumulada e Função de Sobrevivência

A função de distribuição acumulada (FDA) e a função de sobrevivência (FS) constituem elementos fundamentais na caracterização probabilística de variáveis aleatórias discretas. Para a distribuição de Bilal Discreta (DBL), essas funções são obtidas diretamente a partir da discretização da função de sobrevivência da versão contínua da distribuição de Bilal, conforme estabelecido em (ALTUN; EL-MORSHEDY; ELIWA, 2022).

A função de distribuição acumulada da DBL, denotada por $F(x) = \mathbb{P}(X \leq x)$, assume a seguinte forma:

$$F(x) = 1 - (3 - 2p^{x+1}) p^{2(x+1)}, \quad x \in \mathbb{N}_0. \quad (3.2)$$

A partir da expressão acima, define-se a função de sobrevivência, $S(x) = \mathbb{P}(X > x)$, como o complemento da FDA:

$$S(x) = (3 - 2p^{x+1}) p^{2(x+1)}, \quad x \in \mathbb{N}_0. \quad (3.3)$$

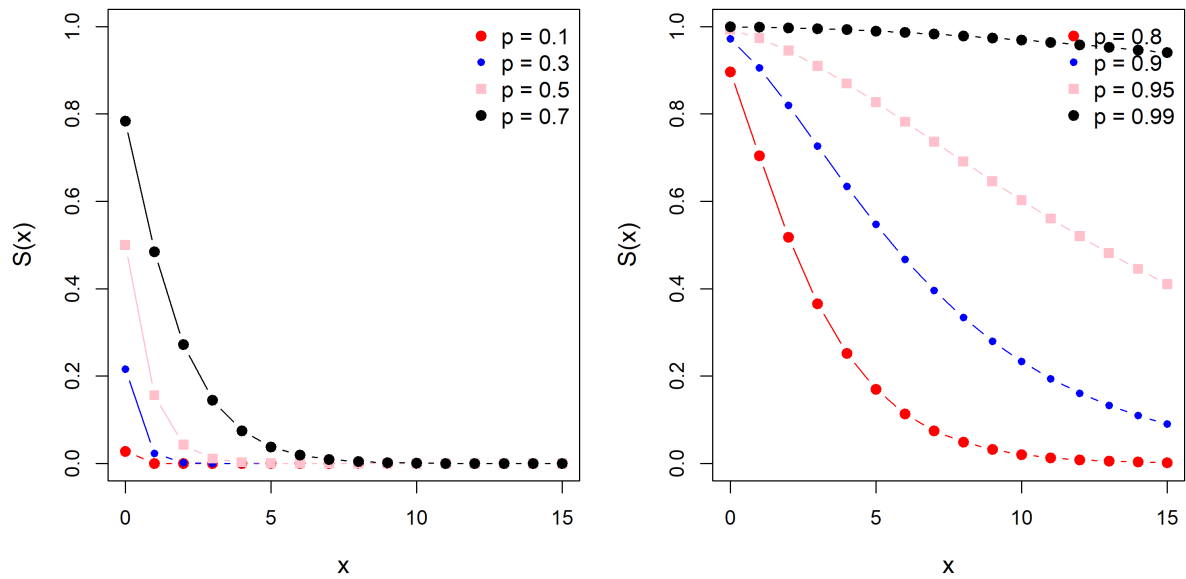
Essas expressões fornecem representações fechadas e computacionalmente eficientes para o cálculo das probabilidades acumuladas e de sobrevivência da DBL.

Além disso, a flexibilidade conferida pelo parâmetro p permite que a FDA da DBL assumam diferentes formatos, indo desde curvas abruptamente crescentes até comportamentos mais graduais, adaptando-se a uma ampla gama de distribuições empíricas observadas em dados de contagem.

A Figura 2 apresenta o comportamento da função de sobrevivência da distribuição de Bilal para diferentes valores do parâmetro p . O painel à esquerda mostra os resultados para valores de p relativamente baixos ($p = 0,1; 0,3; 0,5$ e $0,7$), enquanto o painel à direita contempla valores mais próximos de 1 ($p = 0,8; 0,9; 0,95$ e $0,99$).

Observa-se que, para valores menores de p , a função de sobrevivência decresce rapidamente à medida que m aumenta, refletindo o fato de que as maiores probabilidades estão concentradas nos menores valores de m . Esse comportamento indica uma baixa variabilidade no número de causas latentes ativadoras do evento de interesse, o que pode ser interpretado como um cenário de menor complexidade no mecanismo de ativação.

À medida que o parâmetro p se aproxima de 1, como mostrado no painel da direita, a função de sobrevivência se torna mais lenta em seu decaimento, apresentando caudas longas. Esse padrão evidencia a presença de maior heterogeneidade na quantidade de causas latentes — ou seja, há uma probabilidade não desprezível de que múltiplas causas estejam envolvidas antes da ocorrência do evento. Em particular, para $p = 0,99$, a função de sobrevivência permanece próxima de 1 por vários valores de m , o que caracteriza uma estrutura altamente dispersa, compatível com cenários onde o risco está diluído em muitos possíveis agentes ativadores.

Figura 2 – Função de sobrevivência da distribuição Bilal para diferentes valores do parâmetro p 

Fonte: Elaborada pelo autor.

3.1.3 Função Geradora de Probabilidades

A função geradora de probabilidades (FGP) é uma ferramenta fundamental na teoria das distribuições de probabilidade. Para uma variável aleatória discreta X , com suporte em \mathbb{N}_0 e função de massa de probabilidade $f(x)$, a FGP é definida como:

$$G_X(s) = \mathbb{E}[s^X] = \sum_{x=0}^{\infty} s^x f(x), \quad \text{para } s \in \mathbb{R} \text{ tal que a série converge.} \quad (3.4)$$

A FGP encapsula informações completas sobre a distribuição de X , permitindo caracterizar a variável aleatória. Além disso, a FGP é amplamente utilizada na análise de somas de variáveis aleatórias independentes, no estudo de processos estocásticos e na obtenção de propriedades de convolução.

No caso da distribuição de Bilal Discreta, a função geradora de probabilidades admite uma forma fechada, o que facilita sua aplicação em contextos analíticos e computacionais. A FGP da DBL é dada por:

$$G_X(s) = \frac{2(p^3 - 1)}{1 - p^3 s} - \frac{3(p^2 - 1)}{1 - p^2 s}. \quad (3.5)$$

A existência de uma FGP em forma fechada, como a apresentada acima, é uma vantagem considerável em termos de análise teórica e inferência estatística. Com ela, é possível derivar diretamente não apenas os momentos de primeira e segunda ordem (média e variância), mas também características como assimetria e curtose.

3.1.4 Momentos Estatísticos

Os momentos estatísticos de uma variável aleatória fornecem informações essenciais sobre o seu comportamento probabilístico. Em especial, os quatro primeiros momentos — média, variância, assimetria e curtose — são amplamente utilizados para descrever a tendência central, a dispersão, a simetria e a forma da cauda da distribuição.

Para a distribuição de Bilal Discreta (DBL), todos esses momentos admitem representações analíticas explícitas em função do parâmetro $p \in (0, 1)$, o que constitui uma importante vantagem em termos de inferência e interpretação.

A média, ou valor esperado, representa o centro de gravidade da distribuição, enquanto a variância mede o grau de dispersão dos valores em torno da média. Para distribuições discretas, esses momentos podem ser obtidos a partir da função geradora de momentos ou diretamente a partir da função geradora de probabilidades.

De acordo com [Altun, El-Morshedy e Eliwa \(2022\)](#), a média (valor esperado) da DBL é dada por:

$$\mathbb{E}(X) = \frac{p^2(p^2 + p + 3)}{(p^2 + p + 1)(1 - p^2)}, \quad (3.6)$$

enquanto a variância é expressa como:

$$\mathbb{V}\text{ar}(X) = \frac{p^2(3p^4 + 4p^3 - p^2 + 4p + 3)}{(p^2 + p + 1)^2(1 - p^2)^2}. \quad (3.7)$$

Ambas as medidas aumentam com o crescimento do parâmetro p , refletindo a maior dispersão e assimetria da distribuição à medida que a cauda se torna mais longa.

A assimetria mede o grau de simetria da distribuição em torno de sua média. Para a DBL, a assimetria é dada por:

$$\text{Assimetria}(X) = \frac{-3p^8 + 7p^7 - 3p^6 + 6p^5 + 44p^4 + 6p^3 - 3p^2 + 7p + 3}{p(3p^4 + 4p^3 - p^2 + 4p + 3)^{3/2}}. \quad (3.8)$$

Já a curtose, associada ao achatamento da distribuição, é expressa por:

$$\text{Curtose}(X) = \frac{3p^{12} + 10p^{11} + 19p^{10} + 72p^9 + 224p^8 + 206p^7 + 21p^6 + 206p^5 + 224p^4 + 72p^3 + 19p^2 + 10p + 3}{[p(3p^4 + 4p^3 - p^2 + 4p + 3)]^2} \quad (3.9)$$

Os valores da assimetria indicam que a distribuição é assimétrica à direita para todos os valores admissíveis de p , com o grau de assimetria diminuindo à medida que p se aproxima de 1.

A curtose, por sua vez, é sempre superior a 3, o que caracteriza a DBL como uma distribuição leptocúrtica – ou seja, com caudas mais pesadas e maior concentração em torno da média do que a distribuição Normal.

A Figura 3 ilustra o comportamento da média, variância, assimetria e curtose da distribuição de Bilal à medida que o parâmetro p varia no intervalo $(0, 1)$. A análise conjunta dessas métricas revela aspectos fundamentais sobre a forma, dispersão e cauda da distribuição.

No painel superior esquerdo, observa-se que a média da distribuição aumenta de forma acentuada à medida que $p \rightarrow 1$, indicando que valores mais elevados de p estão associados a maiores valores esperados da variável aleatória. Isso reflete uma maior tendência da distribuição a assumir valores mais afastados da origem.

De maneira semelhante, o painel superior direito mostra que a variância cresce rapidamente conforme p se aproxima de 1. Tal crescimento sugere que a dispersão da distribuição aumenta de forma significativa, caracterizando maior heterogeneidade e variabilidade na estrutura da distribuição quando p é elevada.

O painel inferior esquerdo apresenta a assimetria, que assume valores fortemente negativos para baixos valores de p , e tende rapidamente a zero à medida que p aumenta. Isso indica que, para p próximo de 0, a distribuição é altamente assimétrica positivamente, enquanto valores maiores de p resultam em distribuições mais simétricas.

Por fim, o painel inferior direito mostra a curtose, que é extremamente elevada para valores pequenos de p , sugerindo a presença de caudas pesadas e grande concentração de massa em torno da moda. À medida que p cresce, a curtose decresce abruptamente e se aproxima de valores baixos, indicando distribuições com caudas mais leves.

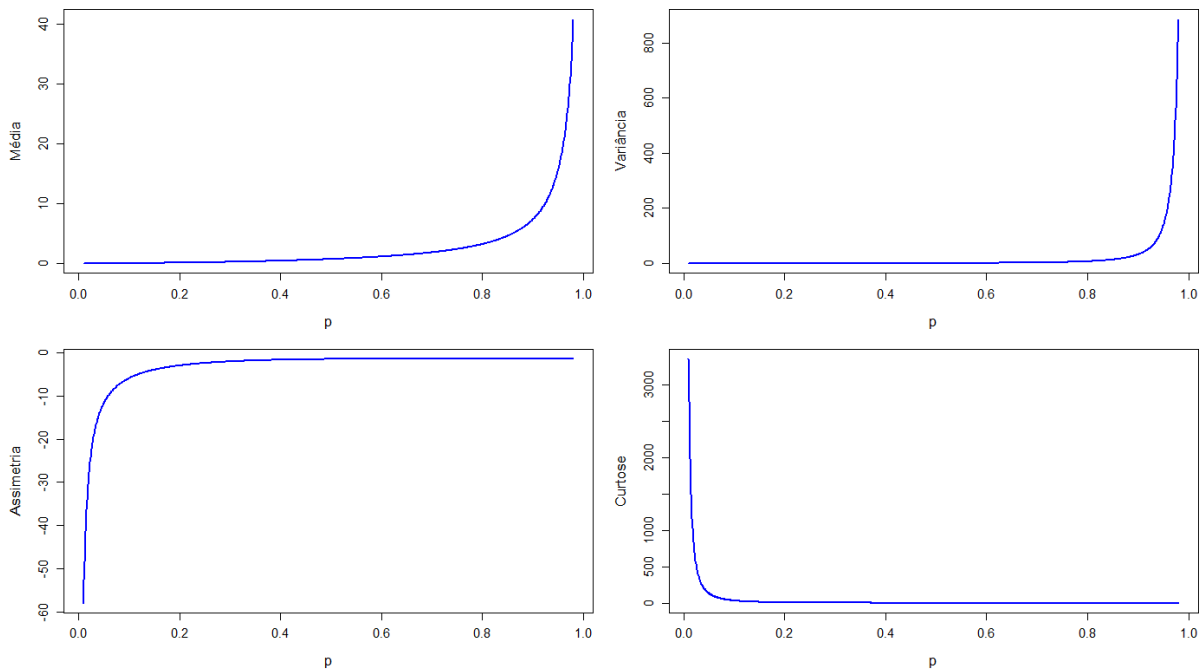
3.1.5 Função de Risco

Para a distribuição de Bilal Discreta, a função de risco é dada por:

$$h(x) = \frac{2(p^3 - 1)p^x - 3(p^2 - 1)}{3 - 2p^x}, \quad x \in \mathbb{N}_0. \quad (3.10)$$

Essa função descreve o comportamento da taxa de ocorrência do evento no instante x , condicionada à sobrevivência até esse ponto. No caso da DBL, o formato funcional de $h(x)$ permite que a distribuição represente diferentes padrões de risco, como crescimento inicial seguido de estabilização. Conforme discutido em (ALTUN; EL-MORSHEDY; ELIWA, 2022), para valores de p moderados ou altos, a função de risco apresenta comportamento crescente até um determinado ponto, após o qual tende à constância, o que é característico de fenômenos em que a probabilidade de falha aumenta rapidamente e depois se estabiliza. Tal flexibilidade torna a DBL adequada para aplicações em confiabilidade e modelagem de dados de falha discretos.

Figura 3 – Valor esperado, variância, assimetria e curtose da distribuição Bilal



Fonte: Elaborada pelo autor.

3.2 Distribuição Weibull

A distribuição de Weibull, originalmente proposta por Waloddi Weibull no artigo *A Statistical Distribution Function of Wide Applicability* publicado no *Journal of Applied Mechanics* em 1951 (WEIBULL, 1951), destaca-se por sua ampla aplicabilidade em diversas áreas das ciências aplicadas. Embora funções semelhantes já fossem conhecidas anteriormente, foi Weibull quem sistematizou sua formulação, interpretou-a sob uma ótica prática e defendeu sua versatilidade em aplicações empíricas. Em sua apresentação original, o autor destacou que a proposta não derivava de primeiros princípios probabilísticos, mas sim da observação empírica e da necessidade de uma função simples e flexível que se ajustasse bem a dados provenientes de contextos variados.

Historicamente, a distribuição de Weibull surgiu da análise de falhas mecânicas em materiais e estruturas, tendo sido utilizada inicialmente para modelar a resistência à tração de aço e de fibras naturais, a distribuição de tamanhos de partículas de cinzas volantes, bem como a vida útil de lâmpadas elétricas (WEIBULL, 1951). Com o tempo, a distribuição passou a ser adotada em um espectro mais amplo de aplicações, incluindo confiabilidade de sistemas, análise de sobrevivência, hidrologia, engenharia elétrica, ciências biomédicas e até mesmo no contexto de fenômenos naturais como a velocidade do vento (JOHNSON; KOTZ; BALAKRISHNAN, 1995).

Do ponto de vista funcional, a principal característica da distribuição de Weibull é sua capacidade de modelar diferentes tipos de comportamentos de risco, dependendo do valor de seus parâmetros. Isso a torna particularmente atrativa em estudos de confiabilidade e durabi-

lidade. Quando o parâmetro de forma é menor que 1, o modelo descreve uma taxa de falha decrescente, típica de falhas iniciais. Para valores iguais a 1, ela se reduz à distribuição exponencial, implicando uma taxa de falha constante. Já quando o parâmetro de forma é maior que 1, a taxa de falha é crescente, condizente com o envelhecimento de componentes ou sistemas. Essa versatilidade funcional é uma das razões centrais para sua popularização em ambientes industriais e acadêmicos.

Na literatura estatística moderna, a distribuição de Weibull é frequentemente abordada como um caso particular das distribuições do valor extremo, especialmente da distribuição de tipo III na classificação de Fisher-Tippett-Gnedenko (JOHNSON; KOTZ; BALAKRISHNAN, 1995). Modelos generalizados de valor extremo também incluem a Weibull como um subcaso (COLES *et al.*, 2001), reforçando seu papel no contexto da teoria de extremos.

Além disso, sua importância foi consolidada por meio de aplicações práticas em diversas frentes. Em confiabilidade, a Weibull é utilizada para estimar o tempo de vida de produtos, planejar manutenções preventivas e calcular garantias comerciais (MEEKER; ESCOBAR; PASCUAL, 2021). Em estatística médica, encontra aplicação na modelagem de tempos até a ocorrência de eventos como óbito ou recaída (LAWLESS, 2011). Em engenharia de software, é aplicada na modelagem do tempo entre falhas em sistemas computacionais (LYU *et al.*, 1996).

3.2.1 Função Densidade de Probabilidade

A distribuição de Weibull é frequentemente apresentada com diferentes parametrizações, refletindo tanto convenções adotadas em diferentes áreas quanto enfoques específicos de aplicação. As variações mais comuns dizem respeito à escolha da base da exponencial, à forma da função de escala, e à notação utilizada para os parâmetros de forma e escala. Essa flexibilidade, embora conceitualmente neutra, pode ocasionar confusões caso não se explicita a parametrização adotada.

Neste trabalho, utiliza-se a seguinte forma funcional para a função densidade de probabilidade (fdp) da distribuição de Weibull:

$$f(x) = \frac{\alpha}{\sigma} \left(\frac{x}{\sigma}\right)^{\alpha-1} \exp\left(-\left(\frac{x}{\sigma}\right)^{\alpha}\right), \quad (3.11)$$

em que $\alpha > 0$ representa o *parâmetro de forma*, responsável por controlar o comportamento da taxa de falha ao longo do tempo, e $\sigma > 0$ é o *parâmetro de escala*, que atua como um fator de dilatação horizontal da distribuição. A variável x assume valores positivos, isto é, $x > 0$.

A escolha desta parametrização segue a formulação clássica encontrada em diversas fontes da literatura estatística (JOHNSON; KOTZ; BALAKRISHNAN, 1995; MEEKER; ESCOBAR; PASCUAL, 2021), sendo especialmente comum em aplicações de confiabilidade e análise de sobrevivência. Nessa configuração, a fdp apresenta diferentes formatos dependendo do valor de α :

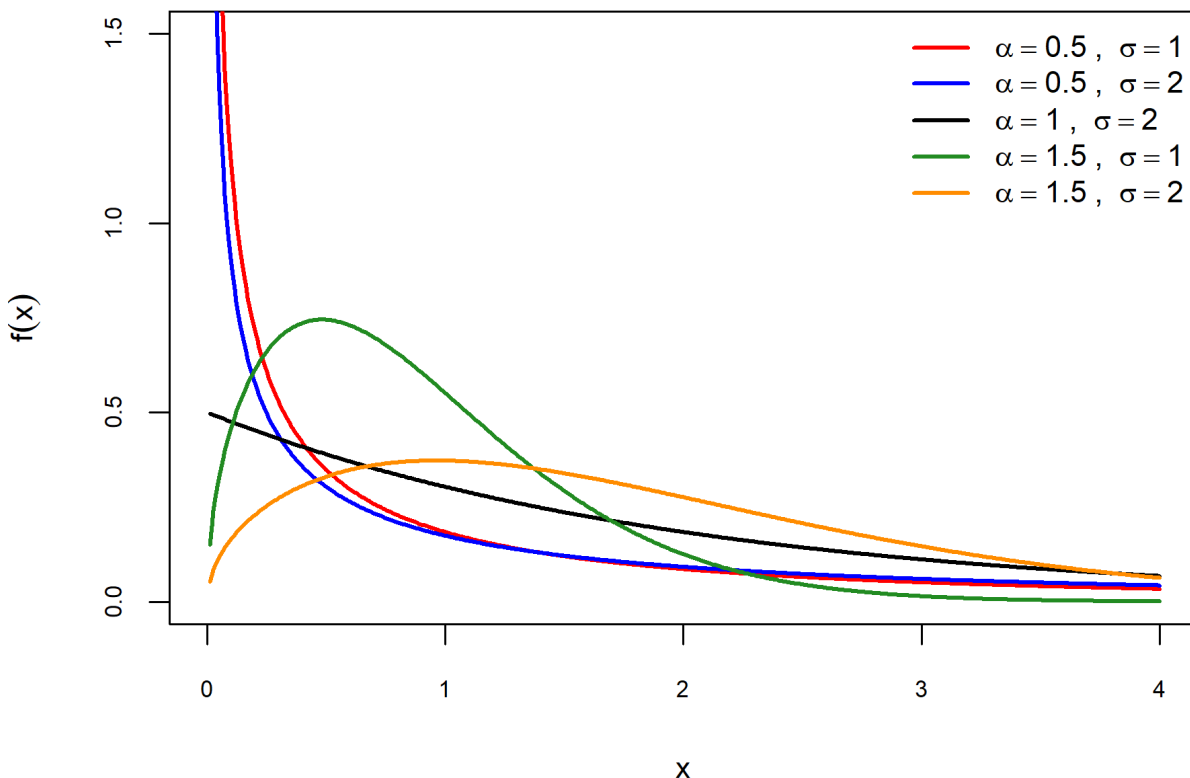
- Se $0 < \alpha < 1$, a função é decrescente, modelando uma taxa de falha decrescente com o tempo;
- Se $\alpha = 1$, a distribuição coincide com a exponencial, implicando uma taxa de falha constante;
- Se $\alpha > 1$, a fdp é unimodal e apresenta um máximo, refletindo uma taxa de falha crescente, típica de processos de envelhecimento.

Ainda, para $\alpha = 1$ tem-se:

$$f(x; \alpha = 1) = \frac{1}{\sigma} \left(\frac{x}{\sigma}\right)^{1-1} \exp\left(-\left(\frac{x}{\sigma}\right)^1\right) = \frac{1}{\sigma} \exp\left(-\left(\frac{x}{\sigma}\right)\right),$$

que denota a função massa de probabilidade da distribuição exponencial com o primeiro momento populacional igual a σ .

Figura 4 – Função densidade de probabilidade da distribuição Weibull



Fonte: Elaborada pelo autor.

A Figura 4 apresenta as curvas da função densidade de probabilidade da distribuição Weibull para diferentes combinações dos parâmetros de forma (α) e escala (σ). Observa-se que os diferentes pares de parâmetros resultam em comportamentos distintos da densidade, evidenciando a flexibilidade dessa distribuição.

Para os casos com $\alpha = 0,5$ (curvas vermelha e azul), a densidade é acentuadamente assimétrica à direita, com alta concentração de probabilidade próxima a $x = 0$ e rápida queda subsequente. Esse comportamento é característico de distribuições com $\alpha < 1$, onde a função densidade diverge em $x \rightarrow 0$ e a cauda é longa.

Quando $\alpha = 1$, representado pela curva preta, a distribuição Weibull coincide com a distribuição exponencial. Nesse cenário, a função densidade atinge seu valor máximo em $x = 0$ e decresce de forma monótona, sem apresentar ponto de máximo interno.

Para os casos com $\alpha = 1,5$ (curvas verde e laranja), a densidade assume uma forma unimodal com um pico deslocado para a direita. Esse comportamento indica uma distribuição com menor assimetria e maior curtose em relação aos casos anteriores. A presença de um ponto de máximo interno é típica de valores $\alpha > 1$, refletindo maior concentração em torno da moda.

Além disso, nota-se que o parâmetro de escala σ atua como um fator de dilatação horizontal: valores maiores de σ tornam a distribuição mais dispersa e achatada. Isso é evidente ao comparar pares de curvas com o mesmo valor de α e diferentes valores de σ — por exemplo, vermelho versus azul ($\alpha = 0,5$) e verde versus laranja ($\alpha = 1,5$).

3.2.2 Função de Distribuição Acumulada e Função de Sobrevivência

A partir da função densidade de probabilidade da distribuição de Weibull definida na equação (3.11), obtêm-se, por integração, as expressões para a função de distribuição acumulada (FDA) e a função de sobrevivência, que são fundamentais em aplicações envolvendo análise de confiabilidade e de sobrevivência.

A função de distribuição acumulada, denotada por $F(x)$, expressa a probabilidade de que a variável aleatória X assumira um valor inferior ou igual a x , ou seja, $F(x) = \mathbb{P}(X \leq x)$. Para a distribuição de Weibull com os parâmetros $\alpha > 0$ e $\sigma > 0$, tem-se:

$$F(x) = 1 - \exp\left(-\left(\frac{x}{\sigma}\right)^\alpha\right), \quad x > 0. \quad (3.12)$$

Essa expressão mostra que a distribuição acumulada da Weibull é contínua e estritamente crescente para $x > 0$. O termo $\left(\frac{x}{\sigma}\right)^\alpha$ governa a rapidez com que $F(x)$ se aproxima de 1, sendo diretamente influenciado pelos parâmetros de forma e escala.

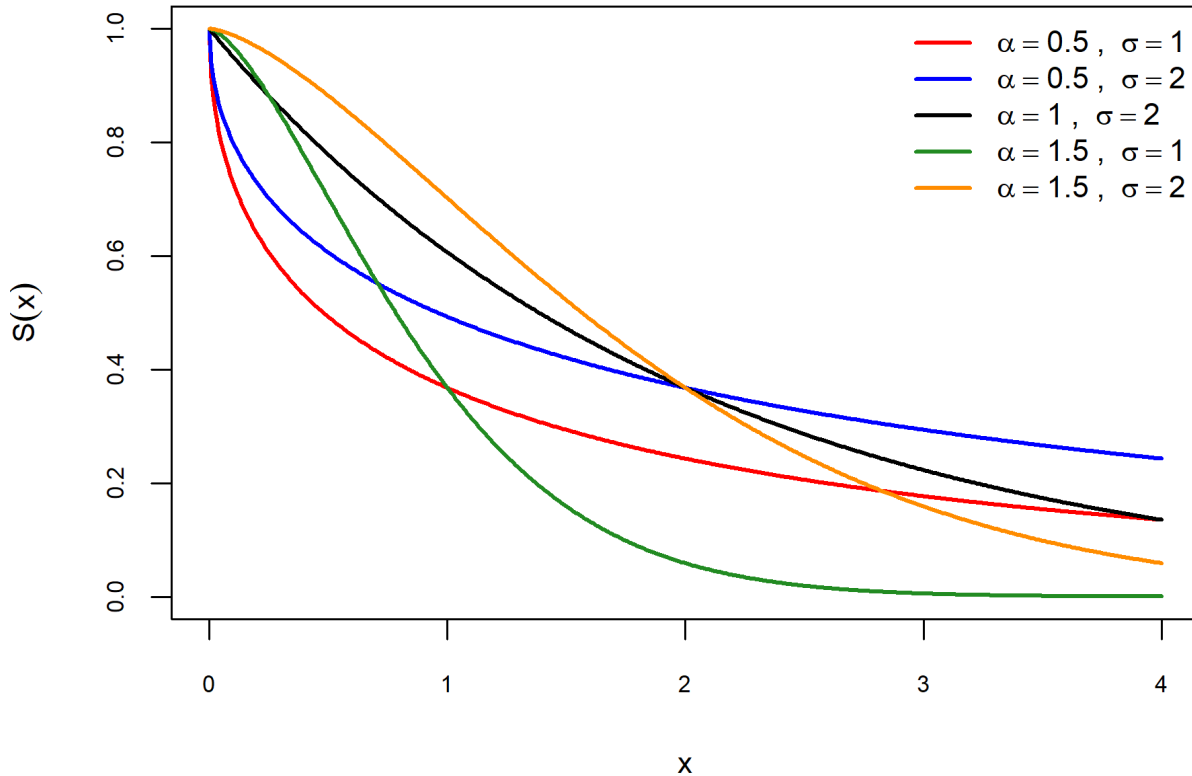
Complementar à função de distribuição acumulada está a função de sobrevivência, denotada por $S(x)$, que indica a probabilidade de sobrevivência além do tempo x , ou seja:

$$S(x) = \mathbb{P}(X > x) = 1 - F(x) = \exp\left(-\left(\frac{x}{\sigma}\right)^\alpha\right), \quad x > 0. \quad (3.13)$$

A função de sobrevivência é amplamente utilizada em contextos onde o interesse reside na permanência de um indivíduo, componente ou sistema além de um dado tempo. A forma

funcional de $S(x)$ na equação (3.13) evidencia a natureza decrescente da probabilidade de sobrevivência com o aumento de x , sendo essa diminuição modulada pelos parâmetros da distribuição.

Figura 5 – Função de sobrevivência da distribuição Weibull



Fonte: Elaborada pelo autor.

A Figura 5 apresenta as curvas da função de sobrevivência $S(x)$ da distribuição Weibull sob diferentes combinações dos parâmetros de forma (α) e escala (σ).

Observa-se que a taxa de decaimento da função de sobrevivência é fortemente influenciada pelo valor do parâmetro de forma α . Para $\alpha = 0,5$ (curvas vermelha e azul), as curvas decrescem mais lentamente para valores crescentes de x , caracterizando distribuições com cauda longa, ou seja, maior probabilidade de sobrevivência em longos períodos. Esse comportamento é típico de distribuições com risco decrescente ao longo do tempo.

Quando $\alpha = 1$ (curva preta), a função de sobrevivência assume a forma da distribuição exponencial, com decaimento aproximadamente constante, representando um risco constante ao longo do tempo.

Para $\alpha = 1,5$ (curvas verde e laranja), observa-se um decaimento mais acentuado da função de sobrevivência, indicando maior concentração de falhas nos primeiros períodos. Isso é coerente com o aumento do risco ao longo do tempo associado a valores de $\alpha > 1$.

O parâmetro de escala σ modula a dispersão da distribuição no eixo horizontal. Quando

σ aumenta (por exemplo, comparando as curvas verde e laranja ou vermelha e azul), a função de sobrevivência se torna mais “esticada”, indicando maior tempo de sobrevivência esperado. Em termos práticos, um maior σ retarda o decaimento da função $S(x)$, independentemente da forma da curva determinada por α .

3.2.3 Momentos Estatísticos

A obtenção dos momentos estatísticos de uma variável aleatória é fundamental para a caracterização de suas propriedades descritivas. No caso da distribuição de Weibull com função densidade de probabilidade dada pela equação (3.11), a média e a variância admitem expressões analíticas fechadas, as quais envolvem a função gama.

Seja X uma variável aleatória com distribuição Weibull parametrizada pelos parâmetros $\alpha > 0$ (forma) e $\sigma > 0$ (escala). Então, o primeiro momento (média) e o segundo momento central (variância) de X são dados por:

$$\mathbb{E}(X) = \sigma \Gamma\left(1 + \frac{1}{\alpha}\right), \quad (3.14)$$

$$\text{Var}(X) = \sigma^2 \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{1}{\alpha}\right)\right)^2 \right], \quad (3.15)$$

em que $\Gamma(\cdot)$ denota a função gama, definida por

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt, \quad z > 0.$$

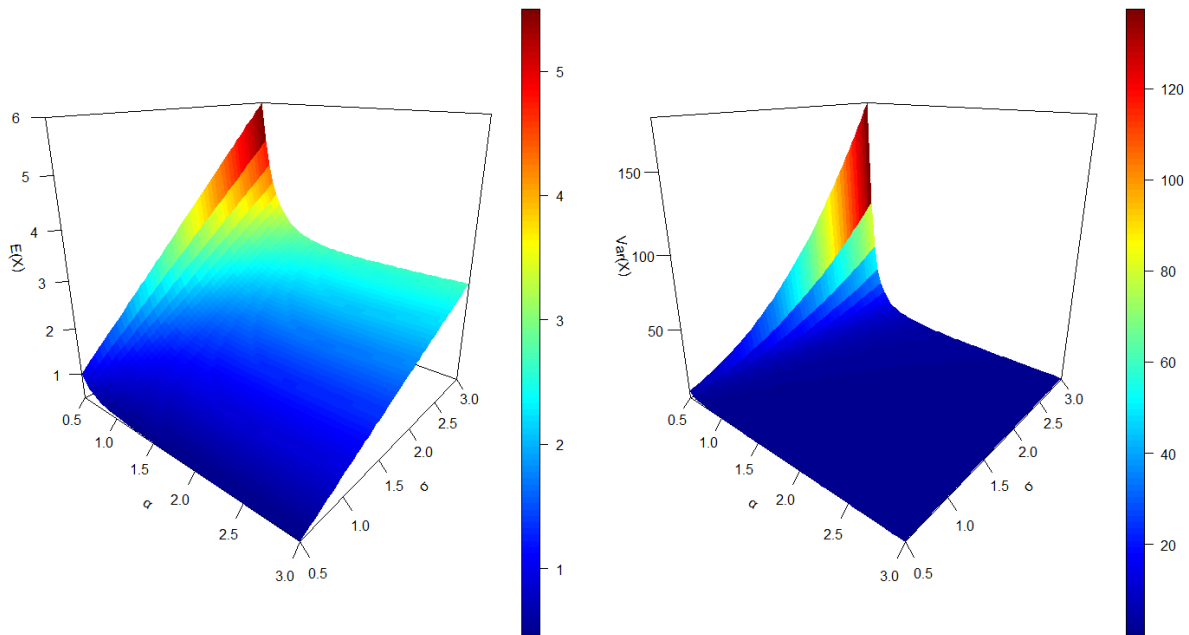
A equação (3.14) mostra que a média da distribuição é diretamente proporcional ao parâmetro de escala σ , enquanto a equação (3.15) evidencia a dependência da variância tanto do parâmetro de escala quanto da forma. À medida que α varia, os valores esperados e a dispersão dos dados se ajustam de maneira não linear, refletindo o comportamento assimétrico característico da distribuição de Weibull.

A Figura 6 apresenta, em painéis distintos, o comportamento da média (à esquerda) e da variância (à direita) da distribuição de Weibull em função dos parâmetros de forma α e de escala σ . Os gráficos tridimensionais permitem visualizar a sensibilidade de ambos os momentos em relação a variações nos parâmetros.

Observa-se que a média $\mathbb{E}(X)$ cresce monotonicamente com o aumento de σ , para qualquer valor fixado de α , o que é consistente com a forma funcional da equação (3.14). Adicionalmente, à medida que α aumenta, a taxa de crescimento da média tende a se suavizar.

No caso da variância $\text{Var}(X)$, nota-se um crescimento acentuado tanto com o aumento de σ quanto com a diminuição de α . Para baixos valores de α , a dispersão torna-se extremamente elevada, refletindo a sensibilidade do segundo momento à forma da distribuição.

Figura 6 – Média e variância pelos parâmetros da distribuição Weibull



Fonte: Elaborada pelo autor.

3.2.4 Função de Risco

A função de risco expressa a taxa instantânea de falha em um determinado tempo, dado que o indivíduo ou sistema sobreviveu até aquele instante. Formalmente, a função de risco pode ser definida como o quociente entre a função densidade de probabilidade $f(x)$ e a função de sobrevivência $S(x)$:

$$h(x) = \frac{f(x)}{S(x)}, \quad x > 0. \quad (3.16)$$

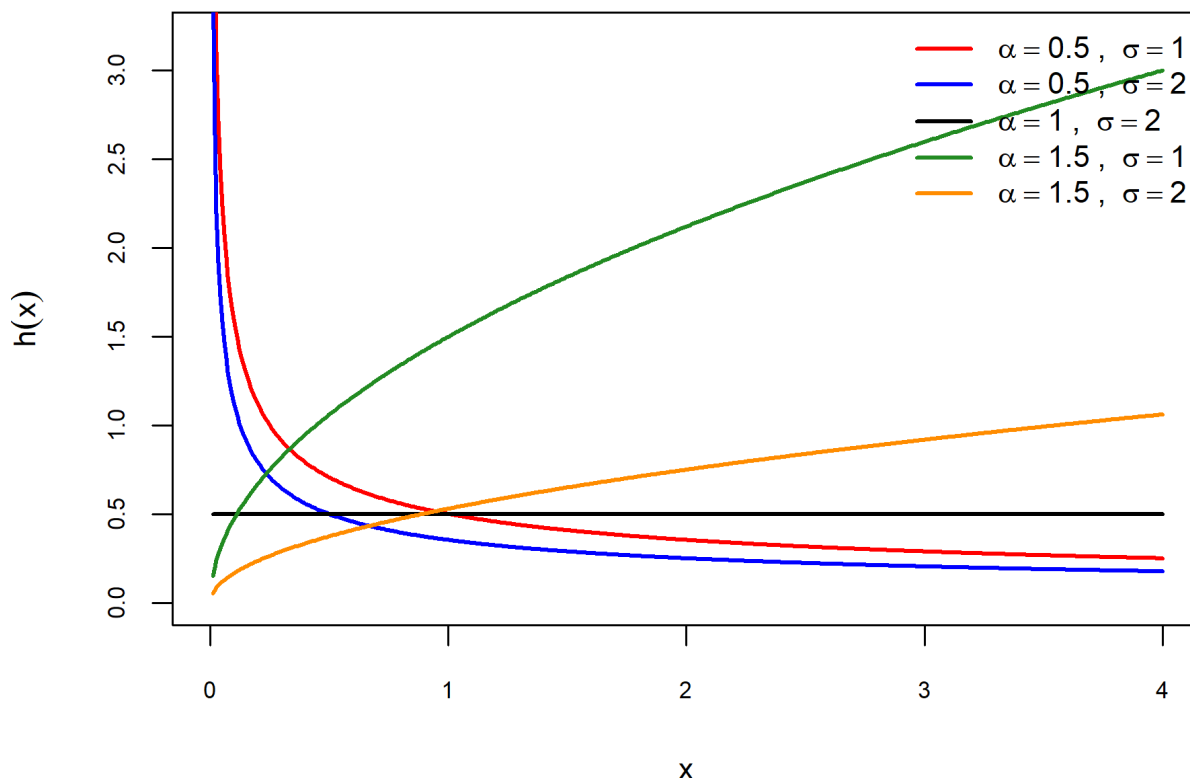
Substituindo as expressões das equações (3.11) e (3.13), obtém-se a função de risco da distribuição de Weibull:

$$h(x) = \frac{\alpha}{\sigma} \left(\frac{x}{\sigma} \right)^{\alpha-1}, \quad x > 0. \quad (3.17)$$

Essa expressão mostra que a função de risco da distribuição de Weibull possui um formato particularmente simples, dependendo apenas dos parâmetros α e σ . O comportamento qualitativo da função de risco é determinado essencialmente pelo parâmetro de forma α :

- Para $\alpha < 1$, a função de risco é decrescente;
- Para $\alpha = 1$, a função de risco é constante;
- Para $\alpha > 1$, a função de risco é crescente.

Figura 7 – Função de risco da distribuição Weibull



Fonte: Elaborada pelo autor.

A Figura 7 ilustra a função de risco $h(x)$ para diferentes combinações dos parâmetros de forma α e de escala σ da distribuição de Weibull, conforme a equação (3.17). Os perfis representados evidenciam a sensibilidade da taxa de falha ao valor do parâmetro α , confirmando os regimes teóricos de falha discutidos anteriormente.

Observa-se que, para $\alpha = 0,5$ (linhas vermelha e azul), a função de risco é decrescente. Esse padrão é típico de situações nas quais itens ou indivíduos apresentam maior propensão à falha nos momentos iniciais, com risco decrescendo à medida que o tempo avança. A diferença entre $\sigma = 1$ e $\sigma = 2$ afeta a escala, tornando a curva mais esticada no eixo horizontal, mas sem alterar a tendência decrescente.

Para o caso $\alpha = 1$ (linha preta), a função de risco é constante, correspondendo à distribuição exponencial. Esse cenário modela processos com ausência de memória, em que a chance de falha é a mesma em qualquer instante, independentemente da história anterior.

Nos casos em que $\alpha = 1,5$ (linhas verde e laranja), verifica-se uma função de risco crescente, o que é típico de sistemas que se deterioram com o tempo. Como esperado, o parâmetro σ atua novamente como fator de dilatação da curva, mantendo inalterada sua concavidade.

O MODELO BILAL WEIBULL COM FRAÇÃO DE CURA

A Seção 2.10 (Modelos com Fração de Cura) apresentou o panorama histórico e conceitual dos modelos com fração de cura, bem como a notação básica empregada ao longo deste trabalho. A Seção 2.11 (Modelos com Mecanismo de Ativação) detalhou a formulação com mecanismos de ativação, discutindo implicações técnicas do tempo observado sob diferentes regras. Para evitar repetições, adota-se aqui esse arcabouço como conhecido e passa-se diretamente à especificação utilizada nesta dissertação.

Considera-se um número latente de causas M associado à ocorrência do evento de interesse e tempos latentes Z_1, \dots, Z_M independentes e identicamente distribuídos, com função de distribuição $F(z; \theta)$ e função de sobrevivência $S(z; \theta) = 1 - F(z; \theta)$. Nesta aplicação, modela-se M por meio da distribuição Bilal (ALTUN; EL-MORSHEDY; ELIWA, 2022), e assume-se para os Z_i a distribuição contínua Weibull (WEIBULL, 1951). O tempo observado T será tratado segundo os mecanismos de ativação apresentados na Seção 2.11, a saber:

$$T = \min(Z_1, \dots, Z_M), \quad T = \max(Z_1, \dots, Z_M), \quad T = Z_R, \quad R \sim \text{Unif}\{1, \dots, M\}.$$

Com essa base, esta seção: (i) formaliza a especificação Bilal Weibull com fração de cura; (ii) deriva as expressões das funções de sobrevivência e densidade populacionais sob cada mecanismo de ativação; (iii) discussão a respeito do comportamento assintótico da função de sobrevivência populacional. As definições gerais, motivações históricas e exemplos ilustrativos permanecem concentrados nas Seções 2.10 e 2.11.

Cumprе ressaltar que, adiante, serão acionadas as fundamentações e formulações específicas das distribuições Bilal e Weibull, cujas expressões são tratadas de forma mais pormenorizada no Capítulo 3.

(1) *Esquema de primeira ativação:*

Na presente dissertação utiliza-se as terminologias e delineamento apresentados por Cooner *et al* em (COONER; BANERJEE; MCBEAN, 2006) e (COONER *et al.*, 2007). Assim sendo, sejam Z , M e T como as variáveis descritas anteriormente. Agora, considere-se M assumindo a distribuição Bilal descrita na Seção 3.1. Tendo a configuração de modelagem com o mínimo entre os tempos concorrentes, ou seja:

$$T = Z_{(1)} = \min \{Z_1, \dots, Z_M\}.$$

Como resultado, tem-se:

$$S_{\text{pop}}(t) = G_M(S(t)), \quad (4.1)$$

em que:

- S_{pop} : função de sobrevivência populacional;
- $G_M(\cdot)$: função geradora de probabilidades da distribuição da variável aleatória M ;
- $S(t)$: função de sobrevivência do tempo de ativação individual.

Considerando a distribuição Bilal para a variável aleatória M , temos que:

$$S_{\text{pop}}(t) = G_M(S(t)) = \frac{2(p^3 - 1)}{1 - p^3 S(t)} - \frac{3(p^2 - 1)}{1 - p^2 S(t)}. \quad (4.2)$$

Substituindo a função de sobrevivência da distribuição Weibull (3.13) em (4.2) tem-se:

$$S_{\text{pop}}(t) = \frac{2(p^3 - 1)}{1 - p^3 \exp\left(-\left(\frac{t}{\sigma}\right)^\alpha\right)} - \frac{3(p^2 - 1)}{1 - p^2 \exp\left(-\left(\frac{t}{\sigma}\right)^\alpha\right)}, \quad (4.3)$$

em que α e σ denotam, respectivamente, os parâmetros de forma e de escala da distribuição Weibull, ambos assumindo valores positivos ($\alpha, \sigma > 0$).

Pelo Teorema Fundamental do Cálculo: $f_{\text{pop}}(t) = -\frac{d}{dt}S_{\text{pop}}(t)$, ou seja:

$$\begin{aligned} f_{\text{pop}}(t) &= \frac{2 \cdot \alpha \cdot p^3 \cdot S(t) \cdot (y^{\alpha-1}) \cdot (p^3 - 1)}{\sigma^\alpha \cdot (1 - p^3 \cdot S(t))^2} - \frac{3 \cdot \alpha \cdot p^2 \cdot S(t) \cdot (y^{\alpha-1}) \cdot (p^2 - 1)}{\sigma^\alpha \cdot (1 - p^2 \cdot S(t))^2} \\ &= f(t) \left[\frac{2(p^3 - 1)p^3}{(1 - p^3 S(t))^2} - \frac{3(p^2 - 1)p^2}{(1 - p^2 S(t))^2} \right], \end{aligned} \quad (4.4)$$

em que $f(t)$ e $S(t)$ denotam, respectivamente, a função densidade de probabilidade e de sobrevivência especificadas em (3.11) e (3.13).

(2) Esquema de última ativação:

Nesta parte do trabalho, adota-se especificamente o mecanismo da última ativação, segundo o qual o evento ocorre apenas após a ativação do último dentre os M mecanismos latentes. Assim, se Z_1, Z_2, \dots, Z_M denotam os tempos individuais, o tempo relacionado ao evento de interesse é dado por:

$$T = Z_{(M)} = \max\{Z_1, Z_2, \dots, Z_M\}. \quad (4.5)$$

Conforme apresentado em (COONER *et al.*, 2007), a função de sobrevivência do modelo é dada por:

$$S_{pop}(t) = 1 + G_M(0) - G_M(F(t)), \quad (4.6)$$

em que $G_M(\cdot)$ é a função geradora de probabilidades de M e $F(t)$ é a função de distribuição acumulada do tempo de ativação individual; adota-se M com distribuição Bilal, cuja função geradora está descrita em 3.1.3, e consideram-se tempos individuais distribuídos segundo Weibull, com parâmetros de forma $\alpha > 0$ e de escala $\sigma > 0$ (ver 3.2); assim, a avaliação de G_M em $F(t)$ resulta em $G_M(F(t))$, que servirá de base para as expressões populacionais a seguir.

$$\begin{aligned} G_M(F(t)) &= \frac{2(p^3 - 1)}{1 - p^3 F(t)} - \frac{3(p^2 - 1)}{1 - p^2 F(t)} \\ &= \frac{2(p^3 - 1)}{1 - p^3(1 - e^{-(t/\sigma)^\alpha})} - \frac{3(p^2 - 1)}{1 - p^2(1 - e^{-(t/\sigma)^\alpha})}. \end{aligned} \quad (4.7)$$

Analogamente, a constante $G_M(0)$, isto é, a função geradora de probabilidades avaliada em zero, é dada por:

$$\begin{aligned} G_M(0) &= \frac{2(p^3 - 1)}{1 - 0} - \frac{3(p^2 - 1)}{1 - 0} \\ &= 2(p^3 - 1) - 3(p^2 - 1) \\ &= 2p^3 - 3p^2 + 1. \end{aligned} \quad (4.8)$$

Portanto, substituindo (4.7) e (4.8) em (4.6), a função de sobrevivência do modelo assume a forma:

$$\begin{aligned} S_{pop}(t) &= 1 + [2(p^3 - 1) - 3(p^2 - 1)] - \left[\frac{2(p^3 - 1)}{1 - p^3 F(t)} - \frac{3(p^2 - 1)}{1 - p^2 F(t)} \right] \\ &= 2 + 2p^3 - 3p^2 - \left[\frac{2(p^3 - 1)}{1 - p^3 F(t)} - \frac{3(p^2 - 1)}{1 - p^2 F(t)} \right]. \end{aligned} \quad (4.9)$$

A função densidade da população $f_{pop}(t)$ pode ser obtida pelo oposto da derivada da função de sobrevivência:

$$f_{pop}(t) = -\frac{d}{dt}S_{pop}(t). \quad (4.10)$$

Derivando a expressão de $S_{pop}(t)$, tem-se:

$$f_{pop}(t) = -\frac{d}{dt} \left\{ 2 + 2p^3 - 3p^2 - \left[\frac{2(p^3 - 1)}{1 - p^3F(t)} - \frac{3(p^2 - 1)}{1 - p^2F(t)} \right] \right\}.$$

Portanto, a densidade marginal é escrita como:

$$f_{pop}(t) = \frac{2(p^3 - 1)p^3 f(t)}{(1 - p^3F(t))^2} - \frac{3(p^2 - 1)p^2 f(t)}{(1 - p^2F(t))^2}. \quad (4.11)$$

Nas equações (4.9) e (4.11), $f(t)$ denota a função densidade de probabilidade *Weibull* com parâmetros α e σ , e $F(t)$ representa a respectiva função de distribuição acumulada com os mesmos parâmetros.

(3) Esquema de ativação aleatória:

Sejam Z_1, Z_2, \dots, Z_M os tempos latentes associados a cada uma das M causas. Sob o mecanismo de ativação aleatória, assume-se que o tempo observado T é obtido ao se selecionar aleatoriamente, com igual probabilidade, um dos Z_i , isto é,

$$T = Z_R \quad \text{com } R \sim \text{Uniforme}\{1, 2, \dots, M\}.$$

Assume-se ainda que os tempos Z_i são independentes e identicamente distribuídos segundo uma distribuição *Weibull*.

Assim como antes, a variável aleatória M , que representa o número de mecanismos latentes de ativação, segue a distribuição de Bilal (ALTUN; EL-MORSHEDY; ELIWA, 2022), cuja função massa de probabilidade é dada na Subseção 3.1.1.

Ainda,

$$\begin{aligned} \mathbb{P}(M = 0) &= 2(p^3 - 1)p^{3 \cdot 0} - 3(p^2 - 1)p^{2 \cdot 0} \\ &= 2p^3 - 2 - 3p^2 + 3 \\ &= 1 + 2p^3 - 3p^2. \end{aligned} \quad (4.12)$$

A função de sobrevivência para a população é definida a partir da mistura induzida pelo número de causas M e pela ativação aleatória. Sob ativação aleatória, a função de sobrevivência condicional ao número M é dada por:

$$\begin{aligned}
S_{pop}(t) &= \mathbb{P}(M = 0) + \{1 - \mathbb{P}(M = 0)\}S(t) \\
&= 1 + 2p^3 - 3p^2 + (3p^2 - 2p^3)S(t) \\
&= 1 + 2p^3 - 3p^2 + (3p^2 - 2p^3)\exp\left\{-\left(\frac{t}{\sigma}\right)^\alpha\right\}.
\end{aligned} \tag{4.13}$$

A função de densidade populacional, derivada da função de sobrevivência, é dada por:

$$\begin{aligned}
f_{pop}(t) &= -\frac{d}{dt}S_{pop}(t) \\
&= -(3p^2 - 2p^3)S'(t) \\
&= (3p^2 - 2p^3)f(t).
\end{aligned} \tag{4.14}$$

Em que a $f(t)$ é a função densidade de probabilidade da distribuição Weibull.

Resumo e discussão

Na Tabela 1, $f(t)$ representa a função densidade de probabilidade da distribuição Weibull, $S(t)$ a correspondente função de sobrevivência e $F(t)$ a função de distribuição acumulada.

Tabela 1 – Função de sobrevivência e função densidade para modelos de fração de cura sob diferentes mecanismos de ativação

Ativação	$S_{pop}(t)$	$f_{pop}(t)$
Primeira	$\frac{2(p^3 - 1)}{1 - p^3 S(t)} - \frac{3(p^2 - 1)}{1 - p^2 S(t)}$	$f(t) \left[\frac{2(p^3 - 1)p^3}{(1 - p^3 S(t))^2} - \frac{3(p^2 - 1)p^2}{(1 - p^2 S(t))^2} \right]$
Última	$\left[\frac{2(p^3 - 1)}{1 - p^3 F(t)} - \frac{3(p^2 - 1)}{1 - p^2 F(t)} \right]$	$\frac{2(p^3 - 1)p^3 f(t)}{(1 - p^3 F(t))^2} - \frac{3(p^2 - 1)p^2 f(t)}{(1 - p^2 F(t))^2}$
Aleatória	$1 + 2p^3 - 3p^2 + (3p^2 - 2p^3)\exp\left\{-\left(\frac{t}{\sigma}\right)^\alpha\right\}$	$(3p^2 - 2p^3)f(t)$

Note que, nos três mecanismos de ativação apresentados, a função de sobrevivência do modelo com fração de cura tende a um limite assintótico positivo quando $t \rightarrow \infty$. Esse comportamento reflete a presença de uma fração de indivíduos curados, ou seja, que nunca experimentarão o evento de interesse. A seguir, descreve-se o comportamento assintótico de $S_{pop}(t)$ para cada mecanismo:

- **Primeira ativação:**

A função de sobrevivência é dada por:

$$S_{pop}(t) = \frac{2(p^3 - 1)}{1 - p^3 S(t)} - \frac{3(p^2 - 1)}{1 - p^2 S(t)}.$$

Como $t \rightarrow \infty$, tem-se que $S(t) \rightarrow 0$, portanto:

$$\begin{aligned}\lim_{t \rightarrow \infty} S_{\text{pop}}(t) &= \frac{2(p^3 - 1)}{1} - \frac{3(p^2 - 1)}{1} = 2(p^3 - 1) - 3(p^2 - 1) \\ &= 2p^3 - 2 - 3p^2 + 3 = 2p^3 - 3p^2 + 1.\end{aligned}$$

• **Última ativação:**

A função de sobrevivência é:

$$S_{\text{pop}}(t) = 2 + 2p^3 - 3p^2 - \left[\frac{2(p^3 - 1)}{1 - p^3 F(t)} - \frac{3(p^2 - 1)}{1 - p^2 F(t)} \right]$$

Como $t \rightarrow \infty$, $F(t) \rightarrow 1$, então:

$$\begin{aligned}\lim_{t \rightarrow \infty} S_{\text{pop}}(t) &= 2 + 2p^3 - 3p^2 - \left[\frac{2(p^3 - 1)}{1 - p^3} - \frac{3(p^2 - 1)}{1 - p^2} \right] \\ &= 2 + 2p^3 - 3p^2 - [-2(p^3 - 1) + (-3)(p^2 - 1)] \\ &= 2 + 2p^3 - 3p^2 + 2(p^3 - 1) + 3(p^2 - 1) \\ &= (2p^3 + 2p^3) + (-3p^2 + 3p^2) + (2 + 3 - 2 - 3) = 2p^3 - 3p^2 + 1.\end{aligned}$$

• **Ativação aleatória:**

A função de sobrevivência é:

$$S_{\text{pop}}(t) = 1 + 2p^3 - 3p^2 + (3p^2 - 2p^3) \exp \left\{ - \left(\frac{t}{\sigma} \right)^\alpha \right\}.$$

Sabendo que $\exp \left\{ - \left(\frac{t}{\sigma} \right)^\alpha \right\} \rightarrow 0$ quando $t \rightarrow \infty$, tem-se:

$$\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 1 + 2p^3 - 3p^2 = 2p^3 - 3p^2 + 1.$$

Em todos os mecanismos de ativação considerados, observa-se que

$$\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = 2p^3 - 3p^2 + 1 = \mathbb{P}(M = 0),$$

em que $\mathbb{P}(M = 0)$ representa a probabilidade de não ativação de nenhum dos mecanismos latentes, ou seja, a proporção de indivíduos efetivamente curados (ou não suscetíveis ao evento de interesse).

Ainda, a incorporação de covariáveis no modelo com fração de cura é realizada por meio da parametrização do parâmetro p , associado à distribuição de Bilal. Nesse contexto, p é modelado como uma função das variáveis explicativas por meio de uma função logística. Tal escolha segue a abordagem consolidada na literatura de modelos com fração de cura, uma vez que a proporção de curados pode ser expressa diretamente em função de p . Assim, ao inserir a

estrutura de regressão em p , garante-se não apenas consistência com formulações clássicas, mas também uma interpretação análoga àquela em que a própria fração de cura é modelada como função das covariáveis. Dessa forma, a relação entre p_i e o vetor de covariáveis x_i é estabelecida por:

$$p_i = \frac{\exp(x_i^\top \beta)}{1 + \exp(x_i^\top \beta)} \quad \text{ou, equivalentemente,} \quad \log\left(\frac{p_i}{1 - p_i}\right) = x_i^\top \beta, \quad (4.15)$$

em que $\beta = (\beta_1, \dots, \beta_k)^\top$ denota o vetor de coeficientes de regressão e $x_i = (1, x_{i1}, \dots, x_{ik})^\top$ representa o vetor de covariáveis da i -ésima observação.

4.1 Estimação - Abordagem Clássica

O método de máxima verossimilhança (EMV) é uma das ferramentas mais fundamentais e amplamente utilizadas na inferência estatística moderna. Sua origem remonta ao início do século XX, sendo desenvolvido e sistematizado por Ronald Aylmer Fisher, um dos estatísticos mais influentes da história. Embora conceitos próximos à verossimilhança já tenham aparecido anteriormente, especialmente nos trabalhos de Carl Friedrich Gauss e Pierre-Simon Laplace no contexto da teoria dos erros e da probabilidade Bayesiana, foi Fisher quem estabeleceu a verossimilhança como base central para a inferência estatística frequentista (STIGLER, 1990; ALDRICH, 1997).

Em 1922, Fisher publicou o artigo seminal *On the mathematical foundations of theoretical statistics*, no qual introduziu formalmente a noção de função de verossimilhança como uma função dos parâmetros desconhecidos de um modelo estatístico, dada uma amostra observada (FISHER, 1922). A partir dessa função, ele propôs que os melhores estimadores seriam aqueles que maximizassem a verossimilhança, ou seja, que tornassem os dados observados mais prováveis sob o modelo considerado.

O EMV surgiu, assim, como uma alternativa robusta e sistemática aos métodos anteriores de estimação, como os momentos de Pearson (SALSBURG, 2024). Fisher mostrou que, sob certas condições regulares, os estimadores de máxima verossimilhança são consistentes, assintoticamente normais e eficientes (FISHER, 1922).

A partir da década de 1930, o método de máxima verossimilhança passou a ser explorado e expandido por outros estatísticos, como Jerzy Neyman, Egon Pearson, e mais tarde por Abraham Wald e Erich Lehmann, que estabeleceram os fundamentos da teoria da decisão e da inferência frequentista moderna (LEHMANN; CASELLA, 1998; NEYMAN; PEARSON, 1933).

Na segunda metade do século XX, com o avanço dos métodos computacionais, o uso do EMV tornou-se ainda mais difundido, sendo aplicado a modelos complexos e de alta dimensão,

como os modelos de regressão, séries temporais, modelos lineares generalizados e, mais recentemente, modelos de sobrevivência com fração de cura e estrutura latente. O método também foi fundamental para o desenvolvimento do algoritmo EM (*Expectation-Maximization*), introduzido por Dempster, Laird e Rubin em 1977, que oferece uma estratégia iterativa para maximizar a verossimilhança na presença de dados incompletos ou latentes (DEMPSTER; LAIRD; RUBIN, 1977).

Em síntese, a trajetória da estimação por máxima verossimilhança acompanha a própria consolidação da Estatística como disciplina teórica e aplicada. Concebida por Fisher em bases rigorosas, a EMV firmou-se como um dos pilares da inferência, aliando elegância formal e relevância prática. Sob hipóteses usuais de regularidade, seus estimadores são consistentes, assintoticamente normais e, no limite, eficientes (no sentido de Cramér–Rao), o que viabiliza rotinas inferenciais padronizadas – em especial testes de hipóteses e a construção de intervalos de confiança baseados na informação de Fisher.

4.1.1 Estimador de Máxima Verossimilhança

Em contextos de análise de sobrevivência, é comum a presença de dados censurados à direita, isto é, situações em que o tempo exato de ocorrência do evento de interesse não é observado para todos os indivíduos da amostra. Em vez disso, sabe-se apenas que o evento não ocorreu até um certo tempo t_i . Para lidar com essa característica dos dados, a função de verossimilhança precisa ser formulada de modo a incorporar tanto as observações completas quanto as censuradas. Seja δ_i uma variável indicadora tal que $\delta_i = 1$ se o tempo t_i for observado (isto é, o evento ocorreu) e $\delta_i = 0$ se t_i for censurado. A função de verossimilhança correspondente pode ser escrita como:

$$\mathcal{L}(\vartheta; \mathcal{D}) = \prod_{i=1}^n [f_{pop}(t_i; \vartheta)]^{\delta_i} [S_{pop}(t_i; \vartheta)]^{1-\delta_i}. \quad (4.16)$$

Em (4.16), $f_{pop}(t_i; \vartheta)$ representa a função densidade de probabilidade populacional e $S_{pop}(t_i; \vartheta)$ a função de sobrevivência populacional, ambas parametrizadas por um vetor ϑ . Assim, para indivíduos com eventos observados ($\delta_i = 1$), contribui-se com a densidade no instante observado t_i ; já para observações censuradas ($\delta_i = 0$), a contribuição é dada pela probabilidade de sobrevivência até aquele tempo. Essa formulação assegura que todas as informações disponíveis – tanto completas quanto parciais – sejam aproveitadas de maneira coerente no processo de estimação por máxima verossimilhança. O uso dessa estrutura é especialmente relevante em modelos com fração de cura, onde o comportamento assintótico da função de sobrevivência pode refletir a presença de indivíduos imunizados ou não suscetíveis ao evento.

Tomando o logaritmo de (4.16), tem-se a função log-verossimilhança:

$$\ell(\vartheta) = \sum_{i=1}^n \{ \delta_i \log(f_{pop}(t_i; \vartheta)) + (1 - \delta_i) \log(S(t_i; \vartheta)) \}. \quad (4.17)$$

Cujas propriedades analticas facilitam a inferncia. O estimador de mxima verossimilhana (EMV)  definido como:

$$\hat{\vartheta} = \arg \max_{\vartheta \in \Theta} \ell(\vartheta). \quad (4.18)$$

Sob condioes regulares, o EMV  consistente, assintoticamente Normal e eficiente (CASELLA; BERGER, 2024; BOLFARINE; SANDOVAL, 2001).

No Apndice A apresentam-se as construoes da verossimilhana e das funoes escore associadas aos trs mecanismos de ativao, considerando o caso sem covarivel.

4.1.2 Informaco de Fisher

A matriz de informaco de Fisher desempenha papel central na inferncia, pois quantifica a preciso dos estimadores de mxima verossimilhana (EMV) ao medir quanta informaco a amostra traz sobre os parmetros do modelo. Seja a funo score definida por:

$$U(\vartheta) = \frac{\partial \ell(\vartheta)}{\partial \vartheta},$$

em que $\ell(\vartheta)$  a log-verossimilhana. A informaco esperada de Fisher  dada por:

$$\mathcal{I}(\vartheta) = \mathbb{E}_{\vartheta} \left[U(\vartheta) U(\vartheta)^\top \right] = -\mathbb{E}_{\vartheta} \left[\frac{\partial^2 \ell(\vartheta)}{\partial \vartheta \partial \vartheta^\top} \right], \quad (4.19)$$

onde a segunda igualdade vale sob condioes regulares. A informaco observada  dada por :

$$\mathcal{I}_o(\vartheta) = -\frac{\partial^2 \ell(\vartheta)}{\partial \vartheta \partial \vartheta^\top},$$

que fornece um estimador prtico de $\mathcal{I}(\vartheta)$ quando avaliada em $\hat{\vartheta}$ (COX; HINKLEY, 1979).

4.1.3 Intervalos de Confiana via Informaco de Fisher

Assumindo condioes regulares, o estimador de mxima verossimilhana (EMV) apresenta distribuo assinttica Normal:

$$\sqrt{n}(\hat{\vartheta} - \vartheta) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathcal{I}^{-1}(\vartheta)), \quad (4.20)$$

em que $\mathcal{I}(\vartheta)$  a matriz de informaco de Fisher, definida como a esperanca da matriz hessiana negativa da log-verossimilhana (COX; HINKLEY, 1979; LEHMANN; CASELLA, 1998). Consequentemente, um intervalo de confiana assinttico para o parmetro ϑ_j  dado por:

$$\hat{\vartheta}_j \pm z_{\alpha/2} \sqrt{[\mathcal{I}^{-1}(\hat{\vartheta})]_{jj}/n}, \quad (4.21)$$

em que $z_{\alpha/2}$ representa o quantil superior da distribuição Normal padrão. Quando a matriz de informação esperada $\mathcal{I}(\vartheta)$ não está disponível de forma analítica, utiliza-se sua contraparte observada, dada por:

$$\mathcal{I}_{\text{obs}}(\hat{\vartheta}) = - \left. \frac{\partial^2 \ell(\vartheta)}{\partial \vartheta \partial \vartheta^\top} \right|_{\vartheta=\hat{\vartheta}}. \quad (4.22)$$

Essa substituição é justificada pelo fato de que, sob regularidade, a informação observada converge para a esperada quando $n \rightarrow \infty$ (BICKEL; DOKSUM, 2015; CASELLA; BERGER, 2024).

4.1.4 Considerações Computacionais

A obtenção do estimador de máxima verossimilhança (EMV) requer, em geral, a maximização da função de log-verossimilhança $\ell(\vartheta)$, o que não necessariamente pode ser feito de forma analítica. Na prática, essa maximização é realizada por meio de algoritmos iterativos de otimização, como o método de Newton-Raphson, BFGS (Broyden–Fletcher–Goldfarb–Shanno) ou outros métodos baseados em gradientes numéricos (NASH, 2018).

Ambientes computacionais como o R oferecem suporte robusto para essas tarefas, dispondo de funções como `optim()`, `nlm()` e `mle()`, que implementam diferentes algoritmos de maximização. No presente trabalho, a estimação clássica foi conduzida por meio da função `optim()`, dada sua flexibilidade, estabilidade numérica e adequação aos requisitos do modelo proposto.

4.1.5 Resíduos de Cox–Snell

A análise de resíduos constitui-se em uma das etapas mais relevantes no processo de avaliação de modelos estatísticos, uma vez que permite verificar a adequação das suposições teóricas assumidas e a qualidade do ajuste obtido a partir dos dados. No contexto da análise de sobrevivência, especialmente quando se considera a presença de fração de cura, os resíduos tornam-se ainda mais importantes, pois possibilitam investigar se a modelagem está captando de maneira apropriada as heterogeneidades da população em estudo (COLOSIMO; GIOLO, 2006).

Nos modelos clássicos de sobrevivência, diferentes resíduos podem ser empregados para avaliar o ajuste do modelo. Entre eles, destacam-se os resíduos de Cox–Snell (COX; SNELL, 1968), cuja finalidade principal é verificar se a distribuição teórica adotada descreve adequadamente os tempos observados. Eles oferecem uma medida global de ajuste, permitindo verificar se a distribuição assumida para o tempo de falha é consistente com os dados observados.

Para um indivíduo i , com tempo de falha (ou censura) T_i , o resíduo de Cox–Snell é definido como a função de risco acumulada estimada no tempo T_i :

$$r_i = \hat{H}(T_i) = -\log \hat{S}(T_i), \quad (4.23)$$

em que $\hat{H}(t)$ representa a função de risco acumulada estimada pelo modelo.

Conforme apresentado em (COLOSIMO; GIOLO, 2006), sob a hipótese de que o modelo está corretamente especificado, os resíduos r_i devem seguir aproximadamente uma distribuição exponencial com parâmetro 1, isto é,

$$r_i \sim \text{Exp}(1). \quad (4.24)$$

Entretanto, quando o objeto de interesse envolve modelos com fração de cura, torna-se necessário reconhecer que a população em análise é composta por duas subpopulações distintas: a parcela de indivíduos curados e a parcela suscetível ao evento. Nessa configuração, a função de sobrevivência global não coincide com a função de sobrevivência dos suscetíveis, o que inviabiliza a aplicação direta dos resíduos clássicos. Para contornar essa limitação, a literatura propôs adaptações específicas voltadas à avaliação de resíduos em modelos de mistura. Por exemplo, Wileyto *et al.* (2013) introduziram um resíduo derivado do tipo Schoenfeld para o caso de modelos paramétricos de mistura, enquanto Peng e Taylor (2017) propuseram metodologias fundamentadas em resíduos de Cox–Snell para analisar o ajuste em modelos de mistura padrão.

De modo complementar, novas contribuições têm buscado não apenas investigar as propriedades de resíduos já existentes, mas também propor alternativas de diagnóstico mais adequadas. Nesse sentido, Scolas *et al.* (2018) apresentaram métodos que utilizam resíduos de Cox–Snell adaptados para lidar tanto com censura à direita quanto com censura intervalar.

Nesta seção, tem-se por objetivo apresentar a análise dos resíduos de Cox–Snell no contexto do modelo com fração de cura e variáveis latentes de ativação, objeto central desta dissertação. Considerando, em particular, o modelo com ativação do tipo mínimo (primeira ativação), Fidelis (FIDELIS, 2019) realiza uma análise detalhada dos resíduos de Cox–Snell para diferentes modelos e, no tocante aos modelos com mecanismo de ativação, ele assume que o número de casos latentes segue uma distribuição de Poisson.

À luz do trabalho de Fidelis e dos demais estudos supracitados, discute-se, nesta seção, o comportamento dos resíduos de Cox–Snell para o modelo que adota a distribuição de Bilal na fração de cura, a distribuição Weibull para os tempos de sobrevivência e as três formas distintas de ativação: primeira, última e aleatória.

Conforme discutido no Capítulo 4, a função de sobrevivência populacional – isto é, a função de sobrevivência do modelo – apresenta um platô quando $t \rightarrow \infty$. Em outras palavras, a

curva de sobrevivência tende assintoticamente a uma reta horizontal, representando a proporção de indivíduos *não curados*.

Nesse contexto, a proporção de não curados é dada por:

$$\pi = 2p^3 - 3p^2 + 1 = (1 - p)^2(2p + 1).$$

Para a análise dos resíduos de Cox–Snell no âmbito dos modelos discutidos nesta dissertação, apresenta-se, nas páginas seguintes, um estudo detalhado sobre a distribuição desses resíduos. Para tal finalidade, enunciam-se e demonstram-se proposições que fundamentam os resultados obtidos.

Assim, apresenta-se a seguir a distribuição associada à função de sobrevivência. Para isso, define-se $U = S_{\text{pop}}(t)$.

Lema 4.1.1. Para $u \in (\pi, 1]$, com S_{pop} contínua e estritamente decrescente, vale a equivalência de eventos:

$$\{S_{\text{pop}}(T) \leq u\} \iff \{T \geq S_{\text{pop}}^{-1}(u)\}.$$

Demonstração. Como S_{pop} é estritamente decrescente e contínua, para todo $t \in [0, \infty)$ tem-se

$$t \geq a \iff S_{\text{pop}}(t) \leq S_{\text{pop}}(a),$$

e, em particular, para $a = S_{\text{pop}}^{-1}(u)$,

$$t \geq S_{\text{pop}}^{-1}(u) \iff S_{\text{pop}}(t) \leq u. \quad (*)$$

A equivalência (*) é ponto a ponto (isto é, para valores determinísticos de t). Aplicando-a ao valor $t = T(\omega)$ em cada ω do espaço amostral, obtém-se:

(\Rightarrow) Se $S_{\text{pop}}(T(\omega)) \leq u$, então, por (*) com $t = T(\omega)$, conclui-se $T(\omega) \geq S_{\text{pop}}^{-1}(u)$, isto é, $\omega \in \{T \geq S_{\text{pop}}^{-1}(u)\}$.

(\Leftarrow) Se $T(\omega) \geq S_{\text{pop}}^{-1}(u)$, então, novamente por (*), tem-se $S_{\text{pop}}(T(\omega)) \leq u$, isto é, $\omega \in \{S_{\text{pop}}(T) \leq u\}$.

Assim, os dois conjuntos coincidem:

$$\{S_{\text{pop}}(T) \leq u\} = \{T \geq S_{\text{pop}}^{-1}(u)\}.$$

□

Proposição 4.1.2. Seja a função de sobrevivência populacional $S_{\text{pop}}(t)$ contínua, estritamente decrescente, com $S_{\text{pop}}(0) = 1$ e $S_{\text{pop}}(t) \rightarrow \pi \in [0, 1)$ quando $t \rightarrow \infty$, então, para $U = S_{\text{pop}}(T)$, tem-se que U admite a seguinte função de distribuição acumulada:

$$F_U(u) = \begin{cases} 0, & u < \pi, \\ u, & \pi \leq u < 1, \\ 1, & u \geq 1. \end{cases}$$

Demonstrao. Note que S_{pop} é contnua, estritamente decrescente em $[0, \infty)$, com $S_{\text{pop}}(0) = 1$ e $S_{\text{pop}}(\infty) = \pi$. Logo, S_{pop} é bijetora de $[0, \infty)$ em $(\pi, 1]$ e admite inversa:

$$S_{\text{pop}}^{-1} : (\pi, 1] \rightarrow [0, \infty).$$

A funo de distribuico acumulada de U é:

$$F_U(u) = \mathbb{P}(U \leq u) = \mathbb{P}(S_{\text{pop}}(T) \leq u).$$

- Se $u < \pi$, como $U \geq \pi$ quase certamente, ento $F_U(u) = 0$.
- Se $u \geq 1$, como $U \leq 1$, temos $F_U(u) = 1$.
- Para $u \in (\pi, 1)$, pela monotonicidade de S_{pop} (Lema 4.1.1):

$$\{S_{\text{pop}}(T) \leq u\} \iff \{T \geq S_{\text{pop}}^{-1}(u)\}.$$

Portanto,

$$F_U(u) = \mathbb{P}(T \geq S_{\text{pop}}^{-1}(u)) = S_{\text{pop}}(S_{\text{pop}}^{-1}(u)) = u.$$

No ponto $u = \pi$, temos

$$F_U(\pi) = \mathbb{P}(S_{\text{pop}}(T) \leq \pi) = \mathbb{P}(T = \infty) = \pi,$$

correspondendo a uma massa pontual de tamanho π no valor $U = \pi$. □

Corolrio 4.1.3. A varivel $U = S_{\text{pop}}(T)$ pode ser descrita como:

$$U \stackrel{d}{=} \begin{cases} \pi, & \text{com probabilidade } \pi, \\ \text{Uniforme}(\pi, 1), & \text{com probabilidade } 1 - \pi. \end{cases}$$

Em particular, a parte contnua de U em $(\pi, 1)$ tem densidade constante igual a 1. Se definirmos a transformao:

$$V = \frac{U - \pi}{1 - \pi},$$

ento, condicionalmente a $T < \infty$, temos $V \sim \text{Uniforme}(0, 1)$.

A proposio 4.1.2 e o conseqente Corolrio 4.1.3 estabelecem a distribuico da varivel aleatria $U = S_{\text{pop}}(T)$, sob as hipteses de que S_{pop} é contnua e estritamente decrescente, com $S_{\text{pop}}(0) = 1$ e $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = \pi \in [0, 1)$. Esse resultado mostra que $U = S_{\text{pop}}(T)$, a qual em modelos sem frao de cura é uniformemente distribuda em $(0, 1)$, assume agora uma estrutura mista. De um lado, existe uma massa pontual em $u = \pi$, de tamanho π , que corresponde à proporo de indivduos curados, isto é, aqueles para os quais $T = \infty$. De outro lado, no intervalo $(\pi, 1)$, a varivel U é uniformemente distribuda, representando o comportamento dos indivduos suscetveis, isto é, aqueles que apresentam o evento em tempo finito.

Proposição 4.1.4. Seja a função de sobrevivência populacional S_{pop} contínua, estritamente decrescente, $S_{\text{pop}}(0) = 1$ e $\lim_{t \rightarrow \infty} S_{\text{pop}}(t) = \pi \in [0, 1)$. Defina $U = S_{\text{pop}}(T)$ e $H = -\log(S_{\text{pop}}(T)) = -\log(U)$. Então a função de distribuição de U é:

$$F_U(u) = \begin{cases} 0, & u < \pi, \\ u, & \pi \leq u < 1, \\ 1, & u \geq 1, \end{cases}$$

e a função de distribuição de H é dada por:

$$F_H(h) = \mathbb{P}(H \leq h) = \begin{cases} 0, & h < 0, \\ 1 - e^{-h}, & 0 \leq h < -\log(\pi), \\ 1, & h \geq -\log(\pi), \end{cases}$$

adotando-se a convenção $-\log(0) = +\infty$ (isto é, quando $\pi = 0$ o segundo caso vale para todo $h \geq 0$). Em particular, H possui densidade

$$f_H(h) = \frac{d}{dh} F_H(h) = \begin{cases} e^{-h}, & 0 < h < -\log(\pi), \\ 0, & \text{caso contrário,} \end{cases}$$

e massa pontual de tamanho π no ponto $h_0 = -\log(\pi)$ (se $\pi > 0$).

Demonstração. Da proposição anterior, $F_U(u) = 0$ para $u < \pi$, $F_U(u) = u$ para $\pi \leq u < 1$ e $F_U(u) = 1$ para $u \geq 1$, o que equivale a dizer que U tem massa π em $u = \pi$ e é uniforme em $(\pi, 1)$ com peso $1 - \pi$.

Para $h \in \mathbb{R}$,

$$F_H(h) = \mathbb{P}(H \leq h) = \mathbb{P}(-\log U \leq h) = \mathbb{P}(U \geq e^{-h}).$$

Se $h < 0$, então $e^{-h} > 1$ e o evento é vazio, logo $F_H(h) = 0$.

Se $0 \leq h < -\log \pi$, então $e^{-h} \in (\pi, 1]$ e, como F_U é contínua nesse intervalo,

$$F_H(h) = \mathbb{P}(U \geq e^{-h}) = 1 - F_U(e^{-h}) = 1 - e^{-h}.$$

Se $h \geq -\log \pi$, então $e^{-h} \leq \pi$ e, pela massa em $u = \pi$,

$$F_H(h) = \mathbb{P}(U \geq e^{-h}) = 1.$$

Derivando no intervalo contínuo $(0, -\log \pi)$ obtém-se $f_H(h) = e^{-h}$. No ponto $h_0 = -\log \pi$ há salto de tamanho $1 - (1 - \pi) = \pi$, isto é, $\mathbb{P}(H = h_0) = \pi$. Quando $\pi = 0$, desaparece a massa pontual e $F_H(h) = 1 - e^{-h}$ para $h \geq 0$, isto é, $H \sim \text{Exp}(1)$. \square

De modo análogo ao discutido por Peng e Taylor (PENG; TAYLOR, 2017) e Scolas et al. (SCOLAS *et al.*, 2018), ao se considerar a distribuição associada a $H(T)$ no contexto do modelo de mistura padrão, observa-se que, no caso do modelo com variáveis latentes de ativação, a distribuição correspondente a $H(T)$ assume caráter misto. Ademais, destaca-se que os resultados teóricos aqui obtidos são condizentes ao encontrado por Fidelis (FIDELIS, 2019) sob outras distribuições.

Ainda que a distribuição vinculada a $H(T)$ não seja exponencial padrão, como ocorre nos modelos clássicos, seu comportamento no intervalo $[0, -\log \pi)$ equivale ao de uma distribuição exponencial padrão. Dessa forma, a avaliação do ajuste do modelo pode ser realizada por meio dos resíduos de Cox–Snell, definidos como:

$$r_i = -\log \widehat{S}_{\text{pop}}, \quad i = 1, \dots, n.$$

Quando o ajuste do modelo é satisfatório, tais resíduos apresentam o mesmo padrão de comportamento de uma amostra truncada proveniente de uma distribuição exponencial padrão no intervalo $[0, -\log \pi)$.

Por fim, ressalta-se que os resultados obtidos são independentes do tipo de mecanismo de ativação considerado. As hipóteses assumidas na Proposição 4.1.4 dizem respeito exclusivamente ao comportamento da função de sobrevivência populacional $S_{\text{pop}}(t)$. Em razão disso, torna-se necessário demonstrar que, sob as ativações *primeira*, *última* e *aleatória*, a função de sobrevivência populacional apresentada na Tabela 1 é contínua e estritamente decrescente no intervalo $[0, \infty)$.

Proposição 4.1.5. Seja $S : [0, \infty) \rightarrow (0, 1]$ uma função de sobrevivência contínua, estritamente decrescente, com $S(0) = 1$ e $\lim_{t \rightarrow \infty} S(t) = 0$. Fixe $p \in (0, 1)$ e denote $F(t) = 1 - S(t)$.

Considere as três funções populacionais:

$$\text{Primeira: } S_{\min}(t) = \frac{2(p^3 - 1)}{1 - p^3 S(t)} - \frac{3(p^2 - 1)}{1 - p^2 S(t)},$$

$$\text{Última: } S_{\max}(t) = 2 + 2p^3 - 3p^2 - \left[\frac{2(p^3 - 1)}{1 - p^3 F(t)} - \frac{3(p^2 - 1)}{1 - p^2 F(t)} \right],$$

$$\text{Aleatória: } S_{\text{ale}}(t) = 1 + 2p^3 - 3p^2 + (3p^2 - 2p^3)S(t).$$

Então $S_{\min}, S_{\max}, S_{\text{ale}}$ são contínuas e estritamente decrescentes em $[0, \infty)$.

Demonstração. Continuidade. Como S e F são contínuas, basta checar que os denominadores das frações não se anulam. Para $k \in \{2, 3\}$, tem-se $0 < p^k S(t) < 1$ e $0 < p^k F(t) < 1$; logo $1 - p^k S(t) > 0$ e $1 - p^k F(t) > 0$. As expressões são composições e combinações algébricas de funções contínuas em conjuntos onde estão bem definidas; portanto, $S_{\min}, S_{\max}, S_{\text{ale}}$ são contínuas.

Monotonicidade. Como S é estritamente decrescente, $S'(t) < 0$ para quase todo t ; como $F = 1 - S$, tem-se $F'(t) = -S'(t) > 0$. Em cada caso, calcula-se a derivada da função de sobrevivência do modelo.

Aleatória. Escreva $S_{\text{ale}}(t) = c_0 + c_1 S(t)$, com $c_0 = 1 + 2p^3 - 3p^2$ e $c_1 = 3p^2 - 2p^3 = p^2(3 - 2p) > 0$. Assim

$$\frac{d}{dt} S_{\text{ale}}(t) = c_1 S'(t) < 0,$$

pois $c_1 > 0$ e $S'(t) < 0$. Logo S_{ale} é estritamente decrescente.

Primeira. Seja $g_1(u)$ tal que $S_{\text{min}}(t) = g_1(S(t))$, ou seja:

$$g_1(u) = \frac{2(p^3 - 1)}{1 - p^3 u} - \frac{3(p^2 - 1)}{1 - p^2 u},$$

derivando em relação a u :

$$\begin{aligned} g_1'(u) &= -2(1 - p^3) \frac{p^3}{(1 - p^3 u)^2} + 3(1 - p^2) \frac{p^2}{(1 - p^2 u)^2} \\ &= \frac{-2(1 - p^3)p^3(1 - p^2 u)^2 + 3(1 - p^2)p^2(1 - p^3 u)^2}{(1 - p^3 u)^2(1 - p^2 u)^2} \\ &= \frac{N(u)}{(1 - p^3 u)^2(1 - p^2 u)^2} \end{aligned}$$

em que:

$$N(u) = p^2 \left\{ 3(1 - p^2)(1 - p^3 u)^2 - 2p(1 - p^3)(1 - p^2 u)^2 \right\}.$$

A fim de analisar o sinal de $N(u)$, considere $h(u) = \frac{1 - p^2 u}{1 - p^3 u}$, $u \in [0, 1]$. Como $h'(u) = \frac{p^3 - p^2}{(1 - p^3 u)^2} < 0$, tem-se:

$$\frac{1 - p^2}{1 - p^3} \leq h(u) \leq 1 \quad \Rightarrow \quad \frac{(1 - p^2)^2}{(1 - p^3)^2} \leq \frac{(1 - p^2 u)^2}{(1 - p^3 u)^2} \leq 1.$$

Em particular,

$$(1 - p^2 u)^2 \leq (1 - p^3 u)^2.$$

Como o coeficiente que multiplica $(1 - p^2 u)^2$ em $N(u)$ é $-2p(1 - p^3) < 0$, substituir $(1 - p^2 u)^2$ por um *majorante* torna esse termo *mais negativo*. Portanto, usa-se precisamente a desigualdade $(1 - p^2 u)^2 \leq (1 - p^3 u)^2$, que fornecerá uma *cota por baixo* para $N(u)$:

$$\begin{aligned}
N(u) &= p^2 \left\{ 3(1-p^2)(1-p^3u)^2 - 2p(1-p^3)(1-p^2u)^2 \right\} \\
&\geq p^2 \left\{ 3(1-p^2)(1-p^3u)^2 - 2p(1-p^3)(1-p^3u)^2 \right\} \\
&\geq p^2(1-p^3u)^2 \left\{ 3(1-p^2) - 2p(1-p^3) \right\}.
\end{aligned}$$

Os fatores p^2 e $(1-p^3u)^2$ são positivos; ademais,

$$3(1-p^2) - 2p(1-p^3) = 3 - 3p^2 - 2p + 2p^4 > 0 \quad \text{para } p \in (0, 1).$$

Logo $N(u) > 0$ e, conseqüentemente, $g'_1(u) > 0$.

Pela regra da cadeia,

$$S'_{\min}(t) = g'_1(S(t))S'(t) < 0.$$

Última. Seja $g_2(u)$ uma função tal que $S_{\max}(t) = g_2(F(t))$, ou seja:

$$g_2(u) = 2 + 2p^3 - 3p^2 - \left[\frac{2(p^3 - 1)}{1 - p^3u} - \frac{3(p^2 - 1)}{1 - p^2u} \right].$$

Note que:

$$g_2(u) = 2 + 2p^3 - 3p^2 - g_1(u).$$

em que $g_1(u)$ é a função definida na demonstração da monotonicidade da primeira ativação. Derivando $g_2(u)$:

$$g'_2(u) = -g'_1(u) < 0$$

E, portanto, pela regra da cadeia:

$$S'_{\max}(t) = g'_2(F(t))F'(t) < 0.$$

Conclui-se que, nos três casos, S_{pop} é contínua e estritamente decrescente em $[0, \infty)$. \square

Por fim, a análise dos resíduos é realizada conforme descrito em (FIDELIS, 2019; FIDELIS; ORTEGA; CORDEIRO, 2024). Os resíduos são definidos por:

$$r_i = -\log \left\{ \widehat{S}_{\text{pop}}(t_i | x_i) \right\},$$

em que t_i representa o tempo observado e $\widehat{S}_{\text{pop}}(\cdot | x_i)$ corresponde à função de sobrevivência populacional ajustada. Essa função é obtida a partir da distribuição Weibull para os tempos

latentes e da distribuição discreta de Bilal para o número de causas M , com probabilidade p . Nos modelos que incluem covariáveis, essa probabilidade é modelada como:

$$p(x) = \text{logit}^{-1}(\beta_0 + \beta_1 x).$$

Com base nos pares (r_i, δ_i) , estima-se de forma não paramétrica a função de risco acumulada $\hat{H}(r)$ utilizando os estimadores de Kaplan–Meier ou de Nelson–Aalen. Posteriormente, constrói-se o gráfico de $\hat{H}(r)$ em função de r , juntamente com a reta identidade $y = x$. Sob especificação adequada do modelo, os pontos devem situar-se próximos a essa reta, enquanto discrepâncias sistemáticas sinalizam falhas no ajuste.

4.2 Estimação - Abordagem Bayesiana

Nesta etapa, conduz-se a inferência Bayesiana dos parâmetros do modelo BWeibcr. Serão explicitadas as questões técnicas necessárias à implementação: especificação das distribuições *a priori*, formulação da verossimilhança sob censura à direita, obtenção da distribuição *a posteriori* e estratégia computacional de amostragem (MCMC), além de procedimentos de diagnóstico e sumarização inferencial.

À semelhança da abordagem clássica, considera-se uma variável aleatória T (tempo de vida) sujeita à censura à direita. Para $i = 1, \dots, n$, sejam W_i o tempo real de falha e C_i o tempo de censura; observa-se

$$T_i = \min\{W_i, C_i\},$$

e define-se o indicador

$$\delta_i = \mathbb{I}\{W_i \leq C_i\},$$

com $\delta_i = 1$ quando ocorre falha observada e $\delta_i = 0$ quando a observação é censurada.

Admite-se, para cada unidade $i = 1, \dots, n$, um vetor de k covariáveis $x_i^\top = (1, x_{i1}, \dots, x_{ik})$, cujo primeiro componente é unitário (intercepto). O parâmetro p_i da distribuição Bilal (Seção 3.1) é modelado em função de x_i por meio da ligação logística:

$$\text{logit}(p_i) = x_i^\top \beta, \tag{4.25}$$

em que $\text{logit}(u) = \log\left(\frac{u}{1-u}\right)$ e $\beta^\top = (\beta_0, \beta_1, \dots, \beta_k)$ denota o vetor de parâmetros.

De forma equivalente,

$$p_i = \frac{\exp\{x_i^\top \beta\}}{1 + \exp\{x_i^\top \beta\}} = (1 + \exp\{-x_i^\top \beta\})^{-1}. \tag{4.26}$$

Essa parametrização permite incorporar a influência das covariáveis na estrutura do modelo por intermédio de p_i , capturando variações sistemáticas na fração de cura entre diferentes

perfis de indivduos. Quando o modelo   ajustado desconsiderando covari veis, a parametrizao de p_i em funo das covari veis n o   utilizada, de modo que p passa a ser tratado como um par metro  nico a ser estimado conjuntamente com os demais par metros do modelo.

A funo de verossimilhana, assim como definida na abordagem inferencial cl ssica, associada ao modelo   expressa por:

$$\mathcal{L}(\vartheta; \mathcal{D}) = \prod_{i=1}^n f_{\text{pop}}(t_i; \vartheta)^{\delta_i} S_{\text{pop}}(t_i; \vartheta)^{1-\delta_i}, \quad (4.27)$$

em que $\vartheta = (\beta, \alpha, \sigma)^\top$ representa o vetor de par metros do modelo; no caso particular em que n o se consideram covari veis, tem-se $\vartheta = (p, \alpha, \sigma)^\top$; e $\mathcal{D} = (t, \delta, x)$ denota o conjunto de dados observados, com $t = (t_1, \dots, t_n)^\top$ e $\delta = (\delta_1, \dots, \delta_n)^\top$.

O vetor $\beta = (\beta_0, \beta_1, \dots, \beta_k)^\top$ cont m os coeficientes de regress o associados  s covari veis, com suporte em \mathbb{R}^{k+1} ; o par metro $\alpha > 0$ corresponde ao par metro de forma da distribuio Weibull; $\sigma > 0$ representa o par metro de escala dessa mesma distribuio; e, no caso sem covari veis, o par metro $p \in (0, 1)$ est  associado   distribuio de Bilal, controlando a frao de cura.

As funo es $f_{\text{pop}}(\cdot; \vartheta)$ e $S_{\text{pop}}(\cdot; \vartheta)$ correspondem, respectivamente,   funo densidade de probabilidade populacional e   funo de sobreviv ncia populacional, cujas formulao es encontram-se detalhadas no Cap tulo 4 para o modelo BWeibcr sob os esquemas de primeira ativao,  ltima ativao e ativao aleat ria.

4.2.1 Distribuico es a priori e a posteriori

Para o desenvolvimento da infer ncia sob a abordagem Bayesiana, torna-se necess ria a especificao de distribuico es *a priori* para os par metros α , σ e β . No caso particular em que n o se consideram covari veis, o vetor de regress o β   substituído pelo par metro p , associado   distribuio de Bilal. Tais distribuico es desempenham o papel de incorporar informao es ou crenas pr vias acerca dos par metros, anteriores ao processo de observao emp rica.

Neste trabalho, assume-se que α , σ e β s o independentes *a priori*. Assim, a distribuico conjunta pode ser escrita como:

$$\pi(\vartheta) = \pi(\alpha)\pi(\sigma) \prod_{j=0}^k \pi(\beta_j), \quad (4.28)$$

em que $\vartheta = (\alpha, \sigma, \beta)$ corresponde ao vetor completo de par metros do modelo.

Dado que α e σ assumem suporte positivo, isto  , \mathbb{R}_+ , enquanto os coeficientes β_j pertencem a \mathbb{R} , para $j = 0, \dots, k$, especificam-se as seguintes distribuico es *a priori*:

$$\alpha \sim \text{Gama}(a, b), \quad \sigma \sim \text{Gama}(a, b), \quad \beta_j \sim N(0, \sigma^2), \quad j = 0, \dots, k,$$

em que $\text{Gama}(a, b)$ denota a distribuição Gama parametrizada por forma a e escala b , e $N(0, \sigma^2)$ representa a distribuição Normal de média nula e variância σ^2 .

Com o intuito de empregar distribuições *a priori* pouco informativas, optou-se por $a = 0,1$, $b = 0,01$ e $\sigma^2 = 10^3$.

Na ausência de covariáveis, ao parâmetro p é atribuída uma distribuição uniforme no intervalo $(0, 1)$, ou seja, $p \sim U(0, 1)$.

Em consonância com o exposto acima, a distribuição *a posteriori* é do tipo:

$$\pi(\vartheta | \mathcal{D}) \propto \pi(\mathcal{D} | \vartheta) \pi(\vartheta). \quad (4.29)$$

Na expressão (4.29), $\pi(\mathcal{D} | \vartheta)$ denota a mesma verossimilhança apresentada em (4.27). A distribuição *a posteriori* conjunta de ϑ resulta da combinação da distribuição *a priori*, dada em (4.28), com a função de verossimilhança $\mathcal{L}(\vartheta; \mathcal{D})$, apresentada em (4.27). No entanto, para os três esquemas de ativação considerados (primeira, última e aleatória), essa distribuição não admite expressão analítica fechada. Por essa razão, a inferência sobre os parâmetros é conduzida por meio de métodos de simulação de Monte Carlo via Cadeias de Markov (MCMC). No presente trabalho, esses procedimentos numéricos foram implementados no ambiente R (R Core Team, 2024) por meio do software JAGS (PLUMMER *et al.*, 2003), que permite a obtenção das distribuições *a posteriori* via amostragem eficiente e totalmente automatizada.

4.2.2 Critérios de comparação de modelos

A literatura dispõe de diversos procedimentos destinados tanto a avaliar o ajuste de um modelo a um conjunto de dados quanto a selecionar, entre alternativas candidatas, aquela com melhor desempenho. Neste trabalho, adota-se o critério LPML (*Logarithm of the Pseudo-Marginal Likelihood*), construído a partir das *conditional predictive ordinates* (CPO).

Para formalizar a CPO, seja \mathcal{D} o conjunto de dados completo, $\mathcal{D}^{(-i)}$ o conjunto obtido pela remoção da i -ésima observação e $\pi(\vartheta | \mathcal{D}^{(-i)})$ a densidade *a posteriori* de ϑ condicionada a $\mathcal{D}^{(-i)}$, para $i = 1, \dots, n$. A CPO associada à i -ésima observação é então definida por:

$$CPO_i = \int_{\Theta} f(t_i | \vartheta) \pi(\vartheta | \mathcal{D}^{(-i)}) d\vartheta = \left\{ \int_{\Theta} \frac{\pi(\vartheta | \mathcal{D})}{f(t_i | \vartheta)} d\vartheta \right\}^{-1}, \quad (4.30)$$

em que $f(t_i | \vartheta)$ denota a função densidade de probabilidade associada ao modelo considerado.

Na comparação entre modelos, valores mais altos de CPO_i sinalizam melhor aderência do modelo aos dados. Para o modelo aqui proposto, uma expressão analítica fechada de CPO_i não está disponível; procede-se, então, à estimação por amostragem de Monte Carlo a partir de uma cadeia MCMC da distribuição *a posteriori* $\pi(\vartheta | \mathcal{D})$. Seja $\vartheta^{(1)}, \dots, \vartheta^{(M)}$ uma amostra

de tamanho M após *burn-in*. Uma aproximação de Monte Carlo para CPO_i , conforme (CHEN; SHAO; IBRAHIM, 2012), é dada por:

$$\widehat{CPO}_i = \left\{ \frac{1}{M} \sum_{m=1}^M \frac{1}{f(t_i | \vartheta^{(m)})} \right\}^{-1}.$$

em que $f(t_i | \vartheta^{(m)})$ denota a função densidade de probabilidade associada ao modelo considerado avaliada sob cada $\vartheta^{(m)}$.

A estatística LPML (*Log Pseudo-Marginal Likelihood*) é dada por:

$$LPML = \sum_{i=1}^n \log(\widehat{CPO}_i),$$

em que \widehat{CPO}_i denota o *Conditional Predictive Ordinate* estimado para a i -ésima observação. Para fins comparativos, entre dois ou mais modelos, prefere-se aquele que apresentar o maior valor de LPML.

Em complemento, empregam-se três critérios adicionais de seleção: o EAIC (*Expected Akaike Information Criterion*, Brooks *et al.* (2002)), o EBIC (*Expected Bayesian/Schwarz Information Criterion*, (CARLIN; LOUIS *et al.*, 2000)) e o DIC (*Deviance Information Criterion*, (SPIEGELHALTER *et al.*, 2002)). Esses índices são construídos a partir da média *a posteriori* da *deviance*, $E\{D(\vartheta)\}$, que representa uma medida de ajuste do modelo.

Operacionalmente, $E\{D(\vartheta)\}$ pode ser estimada por uma amostra MCMC $\{\vartheta^{(m)}\}_{m=1}^M$ via:

$$\bar{D} = \frac{1}{M} \sum_{m=1}^M D(\vartheta^{(m)}),$$

em que $D(\vartheta) = -2 \sum_{i=1}^n \log(f(t_i | \vartheta))$ representa a *deviance*, e $f(\cdot)$ denota a função densidade de probabilidade associada ao modelo considerado.

Os critérios EAIC, EBIC e DIC podem ser calculados, respectivamente, da seguinte forma:

$$\widehat{EAIC} = \bar{D} + 2q, \quad \widehat{EBIC} = \bar{D} + q \log(n) \quad \text{e} \quad \widehat{DIC} = 2\bar{D} - \widehat{D},$$

em que q é o número de parâmetros do modelo e $\widehat{D} = D\left(\frac{1}{M} \sum_{m=1}^M \vartheta^{(m)}\right)$ corresponde à *deviance* avaliada no estimador de Monte Carlo da média *a posteriori*.

Na comparação entre modelos alternativos, o modelo preferido é aquele que apresenta os menores valores para os três critérios (EAIC, EBIC e DIC).

4.2.3 Análise Bayesiana de Influência por Deleção de Casos

O procedimento de deleção de casos (COOK; WEISBERG, 1982) constitui um recurso amplamente difundido para mensurar a sensibilidade do ajuste do modelo a observações individuais. Em paralelo, distintas estratégias de influência local vêm sendo propostas e aplicadas na literatura; ver, por exemplo, Suzuki, Cancho e Louzada (2016), Suzuki *et al.* (2017).

Neste trabalho, emprega-se a análise de influência por deleção de casos baseada em uma divergência genérica ψ . Seja P a distribuição *a posteriori* de ϑ obtida a partir dos dados completos e $P_{(-i)}$ a distribuição *a posteriori* correspondente quando a i -ésima observação é removida. Denotando por $D_\psi(P; P_{(-i)})$ a divergência ψ entre P e $P_{(-i)}$, a medida de divergência ψ é então expressa como:

$$D_\psi(P, P_{(-i)}) = \int_{\vartheta \in \Theta} \psi \left(\frac{\pi(\vartheta | \mathcal{D}^{(-i)})}{\pi(\vartheta | \mathcal{D})} \right) \pi(\vartheta | \mathcal{D}) d\vartheta,$$

em que $\psi(\cdot)$ é uma função convexa tal que $\psi(1) = 0$. Diversas escolhas para ψ são discutidas em Peng e Dey (1995). Neste trabalho, adota-se as seguintes: $\psi(z) = -\log(z)$, que define a divergência de Kullback–Leibler (K–L); $\psi(z) = (z - 1) \log(z)$, correspondente à distância J (ou versão simétrica da divergência K–L); e $\psi(z) = 0,5|z - 1|$, que define a distância variacional, também conhecida como norma L_1 .

Com base na expressão (4.30), a medida de divergência ψ pode ser reescrita da seguinte forma:

$$D_\psi(P, P_{(-i)}) = \mathbb{E}_{\vartheta | \mathcal{D}} \left[\psi \left(\frac{CPO_i}{f(t_i | \vartheta)} \right) \right]. \quad (4.31)$$

A partir da expressão (4.31), pode-se obter, por exemplo, a divergência de Kullback–Leibler (K–L):

$$\begin{aligned} D_{\text{K-L}}(P, P_{(-i)}) &= -\mathbb{E}_{\vartheta | \mathcal{D}} \{ \log(CPO_i) \} + \mathbb{E}_{\vartheta | \mathcal{D}} \{ \log[f(t_i | \vartheta)] \} \\ &= -\log(CPO_i) + \mathbb{E}_{\vartheta | \mathcal{D}} \{ \log[f(t_i | \vartheta)] \}. \end{aligned}$$

As estimativas de Monte Carlo para $D_\psi(P, P_{(-i)})$ são dadas por:

$$\widehat{D}_\psi(P, P_{(-i)}) = \frac{1}{Q} \sum_{q=1}^Q \psi \left(\frac{\widehat{CPO}_i}{f(t_i | \vartheta^{(q)})} \right).$$

Assim, a estimativa da divergência de Kullback–Leibler, $D_{\text{K-L}}(P, P_{(-i)})$, é obtida da seguinte forma:

$$\widehat{D}_{\text{K-L}}(P, P_{(-i)}) = -\log(\widehat{CPO}_i) + \frac{1}{Q} \sum_{q=1}^Q \log[f(t_i | \vartheta^{(q)})].$$

A divergência $D_\psi(P, P_{(-i)})$ pode ser interpretada como o efeito da exclusão da i -ésima observação dos dados completos sobre a distribuição *a posteriori* de ϑ . No entanto, como discutido por Peng e Dey (1995) e Weiss (1996), pode ser desafiador para o pesquisador estabelecer um ponto de corte para o valor dessa medida de divergência, a fim de determinar se um subconjunto de observações exerce ou não influência significativa.

Neste trabalho, segue-se a proposta de Peng e Dey (1995) e Weiss (1996), que estabelece uma analogia com uma moeda viesada de probabilidade de sucesso p . Nesse contexto, a divergência ψ entre duas moedas viesadas é expressa por:

$$D_\psi(g_0, g_1) = \int \psi \left(\frac{g_0(x)}{g_1(x)} \right) g_1(x) dx,$$

em que $g_0(x) = p^x(1-p)^{1-x}$ e $g_1(x) = 0,5$, para $x = 0$ ou 1 . Se $D_\psi(g_0, g_1) = d_\psi(p)$, pode-se verificar que d_ψ satisfaz a seguinte equação:

$$d_\psi(p) = \frac{\psi(2p) + \psi(2(1-p))}{2}.$$

Observa-se que d_ψ aumenta à medida que p se afasta de $0,50$. Além disso, $d_\psi(p)$ é simétrico em torno de $p = 0,50$, atingindo seu valor mínimo neste ponto, em que $d_\psi(0,5) = 0$ e $g_0 = g_1$.

Por exemplo, ao considerar $p > 0,80$ (ou $p < 0,20$) como indicativo de forte viés em uma moeda, obtém-se $d_{L_1}(0,80) = 0,30$. Dessa forma, estabelece-se que um caso i pode ser considerado influente quando $d_{L_1} > 0,30$.

De modo análogo, ao utilizar a divergência de Kullback–Leibler, uma observação é classificada como influente quando $d_{K-L} > 0,2231$. Já no caso da distância J , o ponto de corte corresponde a $d_J > 0,4159$.

Neste trabalho, adotam-se esses valores de referência como critérios para determinar se uma observação exerce influência significativa ou não.

RESULTADOS E APLICAÇÕES

5.1 Modelo Sem Covariável

Neste capítulo são apresentados os resultados referentes ao modelo sem covariável, descrito em maior detalhe na Seção 4. No mesmo contexto, apresentam-se os resultados de um estudo de simulação sob as abordagens clássica e Bayesiana, além da aplicação do modelo a dados reais e da análise diagnóstica correspondente.

5.1.1 Estudo de simulação

Um estudo de simulação constitui uma ferramenta fundamental para avaliar o desempenho de estimadores sob diferentes cenários, permitindo investigar suas propriedades em termos de viés, erro quadrático médio, cobertura de intervalos de confiança e estabilidade computacional. No contexto clássico, tais estudos possibilitam verificar a consistência das estimativas de máxima verossimilhança e a adequação das inferências assintóticas (DAVISON; HINKLEY, 1997). Já na abordagem Bayesiana, a simulação permite examinar a acurácia de métodos baseados em amostradores de cadeias de Markov e avaliar critérios de seleção de modelos, como o DIC e o LPML (GELMAN *et al.*, 1995). Assim, tanto na perspectiva clássica quanto na Bayesiana, a simulação atua como um instrumento central para validar metodologias estatísticas e oferecer suporte à aplicação prática em dados reais.

Para a composição dos elementos da amostra, procede-se inicialmente à geração do número de causas latentes M_i para cada unidade $i = 1, \dots, n$ – em que n representa o tamanho da amostra – assumindo $M_i \sim \text{Bilal}(p_i)$. Em seguida, para cada réplica $j = 1, \dots, 1000$, amostram-se os tempos latentes Z_{i1}, \dots, Z_{iM_i} , considerados independentes e identicamente distribuídos segundo uma distribuição Weibull(α, σ). O tempo de falha observado T_i é definido conforme o mecanismo de ativação: no modelo de *primeira ativação*, $Y_i = \min\{Z_{ik}\}$; no de *última ativação*, $Y_i = \max\{Z_{ik}\}$; e no de *ativação aleatória*, $Y_i = Z_{iR}$, em que R é selecionado de forma uniforme

em $\{1, \dots, M_i\}$. Em seguida, geram-se tempos de censura $C_i \sim \text{Unif}(0, \tau)$ e registram-se $T_i = \min(Y_i, C_i)$ e $\delta_i = \mathbb{I}(T_i \leq C_i)$. Os verdadeiros valores dos parâmetros são: $\alpha = 5$, $\sigma = 2$ e $p = 0,3$ com nível de censura de 85%.

(i) Abordagem Clássica

O estudo de simulação sem a inclusão de covariáveis foi realizado com 1.000 réplicas para os tamanhos amostrais $n \in \{50, 100, 200, 400\}$, fixando-se como valores verdadeiros $\alpha = 5$, $\sigma = 2$ e $p = 0,3$.

Obtiveram-se as estimativas pontuais dos parâmetros de interesse, bem como o viés médio, o erro quadrático médio (EQM) e a probabilidade de cobertura (PC) dos intervalos de confiança assintóticos, considerando um nível de confiança de 95%. Os intervalos foram construídos a partir da matriz de informação observada de Fisher. Todas as análises foram implementadas no software R por meio da função *optim*.

Tabela 2 – Estudo de simulação abordagem clássica sem covariáveis sob o esquema de primeira ativação

Tamanho Amostral (n)	Parâmetro	Estimativa	Viés	EQM	PC
50	p	0,3127	0,0127	0,0033	0,961
	α	5,0474	0,0474	1,5864	0,954
	σ	2,1013	0,1013	0,0343	0,954
100	p	0,3042	0,0042	0,0016	0,954
	α	5,1180	0,1180	1,0615	0,960
	σ	2,0634	0,0634	0,0170	0,937
200	p	0,2993	-0,0007	0,0008	0,956
	α	5,0609	0,0609	0,6339	0,958
	σ	2,0337	0,0337	0,0075	0,959
400	p	0,3010	0,0010	0,0004	0,951
	α	4,9825	-0,0175	0,2804	0,953
	σ	2,0242	0,0242	0,0036	0,961

Tabela 3 – Estudo de simulação abordagem clássica sem covariáveis sob o esquema de última ativação

Tamanho Amostral (n)	Parâmetro	Estimativa	Viés	EQM	PC
50	p	0,3157	0,0157	0,0034	0,952
	α	5,0742	0,0742	1,6845	0,954
	σ	2,0728	0,0728	0,0305	0,932
100	p	0,3064	0,0064	0,0015	0,966
	α	5,0747	0,0747	1,0152	0,956
	σ	2,0622	0,0622	0,0154	0,940
200	p	0,3016	0,0016	0,0007	0,968
	α	5,0011	0,0011	0,5482	0,958
	σ	2,0371	0,0371	0,0073	0,942
400	p	0,3011	0,0011	0,0004	0,966
	α	5,0032	0,0032	0,3488	0,931
	σ	2,0246	0,0246	0,0038	0,947

Tabela 4 – Estudo de simulação abordagem clássica sem covariáveis sob o esquema de ativação aleatória

Tamanho Amostral (n)	Parâmetro	Estimativa	Viés	EQM	PC
50	p	0,3895	0,0895	0,0440	0,990
	α	4,2228	-0,7772	1,3291	0,955
	σ	2,2313	0,2313	0,6797	0,911
100	p	0,3357	0,0357	0,0158	0,984
	α	4,4575	-0,5425	0,8122	0,968
	σ	2,1123	0,1123	0,2210	0,915
200	p	0,3136	0,0136	0,0030	0,970
	α	4,5663	-0,4337	0,5561	0,969
	σ	2,0625	0,0625	0,0561	0,956
400	p	0,3058	0,0058	0,0009	0,971
	α	4,7530	-0,2470	0,2689	0,982
	σ	2,0226	0,0226	0,0138	0,968

A Tabela 2 apresenta os resultados do estudo de simulação sob o esquema de primeira ativação, sem a inclusão de covariável. Observa-se que as estimativas de p , α e σ aproximam-se gradualmente dos valores verdadeiros à medida que o tamanho amostral aumenta. Os vieses tendem a se reduzir com o crescimento da amostra, acompanhados pela diminuição do erro quadrático médio (EQM), evidenciando a consistência assintótica dos estimadores. As taxas de cobertura (PC) mantêm-se em torno de 95%, o que demonstra a adequação dos intervalos de confiança baseados na informação observada de Fisher.

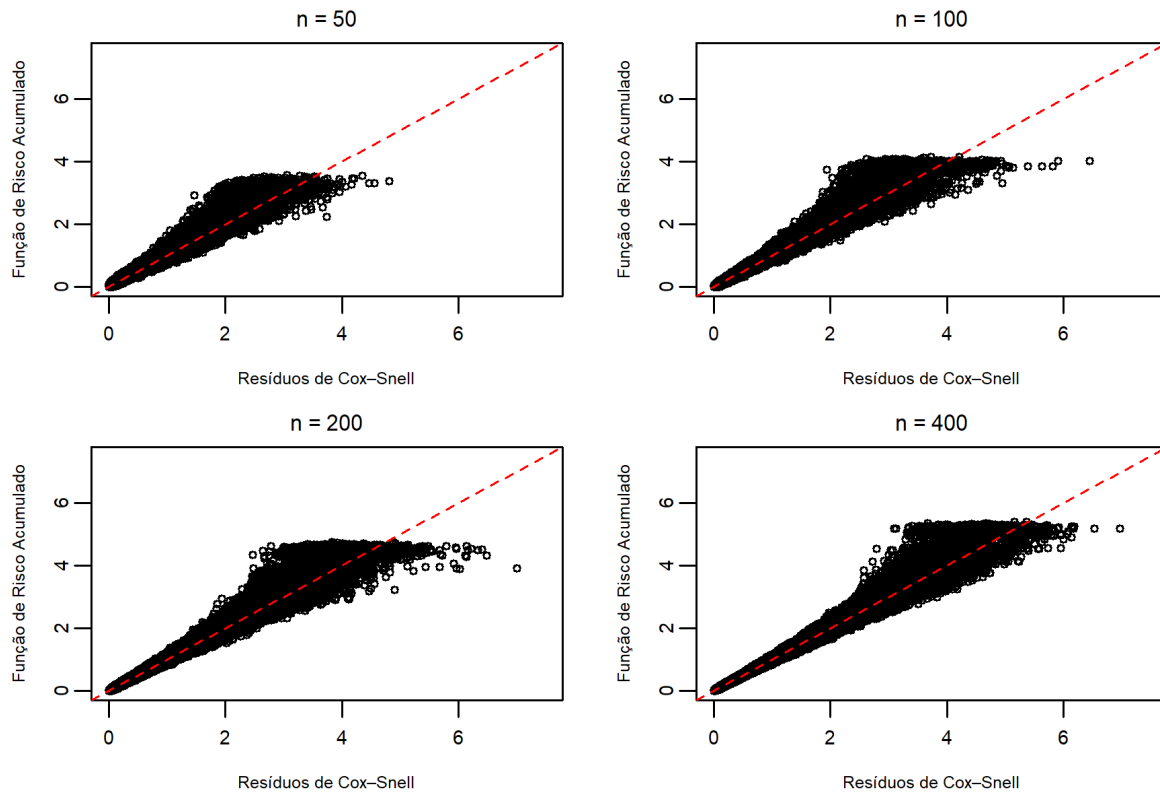
Na Tabela 3, referente ao esquema de última ativação, verifica-se padrão semelhante ao do caso anterior. As estimativas apresentam pequena discrepância para amostras reduzidas, mas convergem rapidamente para os valores verdadeiros à medida que n cresce. Tanto o viés quanto o EQM diminuem de forma consistente, enquanto as taxas de cobertura permanecem estáveis e próximas ao nível nominal, indicando bom desempenho inferencial do modelo.

Por fim, a Tabela 4, associada ao mecanismo de ativação aleatória, evidencia resultados alinhados aos esquemas anteriores, embora se perceba maior instabilidade em amostras pequenas, especialmente para os parâmetros α e σ (maior viés). Ainda assim, com o aumento do tamanho amostral, nota-se convergência das estimativas, redução do viés e manutenção de taxas de cobertura adequadas. Em termos gerais, os três mecanismos de ativação apresentam desempenho semelhante na ausência de covariável, distinguindo-se apenas em pequenas variações nas amostras menores.

Como análise de simulação complementar, avaliou-se o comportamento dos resíduos de Cox–Snell em dois cenários: (i) especificação correta – dados gerados pelo BWeibcr sob o mecanismo de última ativação, com parâmetros $\alpha = 5$, $\sigma = 2$ e $p = 0,95$; o ajuste foi realizado por máxima verossimilhança com o próprio modelo de última ativação, seguido do cálculo dos respectivos resíduos de Cox–Snell; (ii) especificação incorreta – dados também gerados sob última ativação, porém ajustados pela verossimilhança correspondente ao mecanismo de primeira

ativação. O estudo considerou 1.000 réplicas, tamanhos amostrais $n = 50, 100, 200$ e 400 e nível de censura de 50% .

Figura 8 – Resíduos de Cox-Snell do modelo bem especificado sem covariável



Fonte: Elaborada pelo autor.

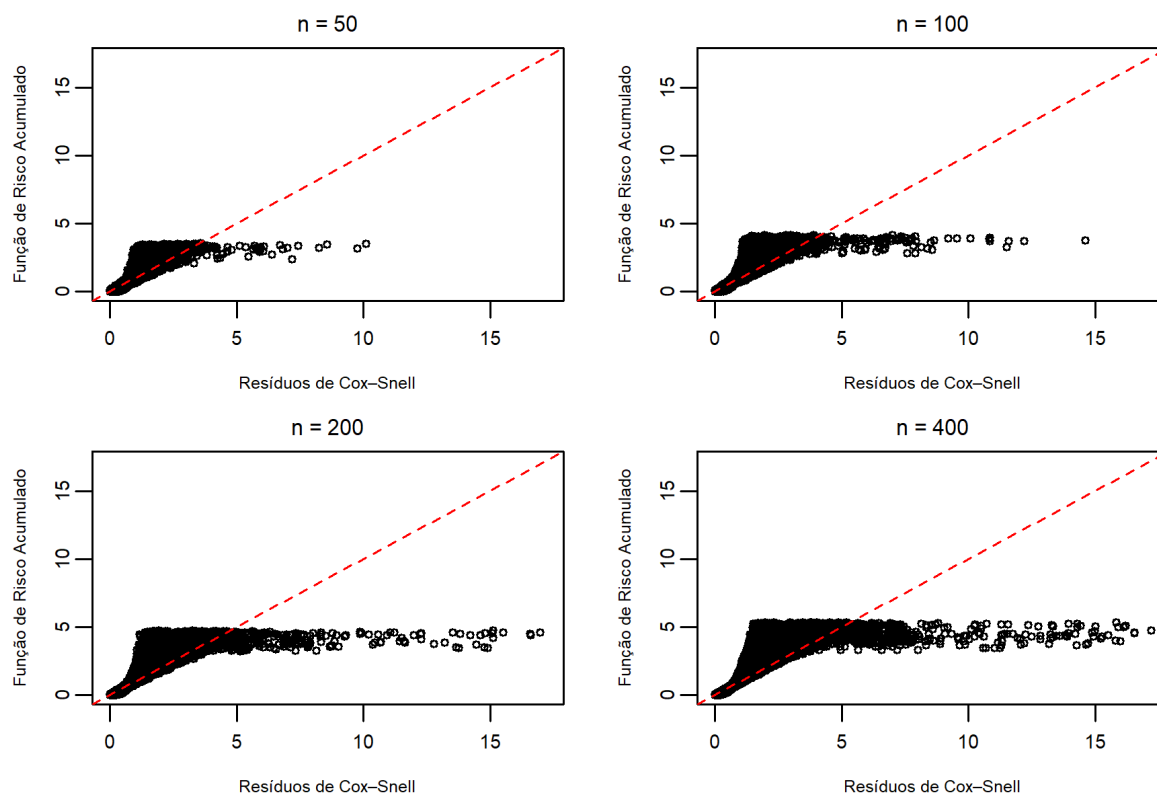
Na Figura 8, correspondente ao modelo corretamente especificado, a nuvem de pontos acompanha a reta identidade, sobretudo a partir de $n \geq 200$, com discreta sublinearidade nas caudas, compatível com o comportamento esperado dos resíduos de Cox-Snell. Já na Figura 9, referente ao ajuste com mecanismo de ativação incorreto, observa-se afastamento sistemático da referência, com concavidade marcada e caudas alongadas. O aumento do tamanho amostral não remove o desvio estrutural, evidenciando a não adequação do modelo aos dados.

(ii) Abordagem Bayesiana

Foi realizado um estudo de simulação com dois objetivos principais: (i) avaliar o comportamento das estimativas dos parâmetros nos modelos propostos; e (ii) examinar o desempenho da abordagem diagnóstica sugerida, considerando conjuntos de dados simulados contendo nenhum, um, dois ou três casos perturbados.

Para cada conjunto de dados, foram geradas duas cadeias MCMC de tamanho 25.000 para cada parâmetro, com descarte das primeiras 5.000 iterações a fim de eliminar a influência dos valores iniciais (*burn-in*). Com o propósito de reduzir a autocorrelação entre as amostras, adotou-se um espaçamento de 10 iterações, resultando em uma amostra final de tamanho 4.000, sobre a qual a inferência *a posteriori* foi conduzida. Para cada amostra, registraram-se a média a

Figura 9 – Resíduos de Cox-Snell do modelo mal especificado sem covariável



Fonte: Elaborada pelo autor.

posteriori dos parâmetros, bem como os valores para avaliação dos modelos. Os intervalos de credibilidade foram construídos a partir dos percentis 2,5% e 97,5% da distribuição *a posteriori*. As análises bayesianas foram implementadas no software R (R Core Team, 2024), utilizando o ambiente JAGS (*Just Another Gibbs Sampler*) (PLUMMER *et al.*, 2003).

A convergência das cadeias foi avaliada segundo os procedimentos recomendados por Cowles e Carlin (1996). Em todos os cenários, a convergência foi confirmada tanto pelo diagnóstico de Gelman–Rubin (GELMAN; RUBIN, 1992), cujos valores permaneceram muito próximos de 1 (isto é, ≤ 1.01), quanto pela inspeção visual dos gráficos de traços.

No primeiro estudo o modelo BWeibcr foi analisado sob três esquemas de ativação: primeira, última e aleatória. Os valores verdadeiros dos parâmetros foram fixados em $\alpha = 5$, $\sigma = 2$ e $p = 0,3$.

Foram conduzidos experimentos de simulação com 200 réplicas, a fim de avaliar o desempenho do modelo BWeibcr em cada um dos três esquemas de ativação, na ausência de covariáveis. Os tamanhos amostrais considerados foram $n = 50, 100$ e 200 .

A Tabela 5 apresenta os resultados do estudo de simulação Bayesiano sem covariáveis sob o esquema de primeira ativação. Para $n = 50$, observa-se superestimação de p (média em torno de 0,37) e subestimação de α (média próxima de 4,66), enquanto σ é ligeiramente

Tabela 5 – Estudo de simulação abordagem Bayesiana sem covariáveis sob o esquema de primeira ativação

Tamanho Amostral (n)	Parâmetro	Média	Viés	EQM	PC
50	p	0,3707	0,0707	0,0064	0,9344
	α	4,6581	-0,3419	1,1256	0,9713
	σ	2,3910	0,3910	0,2250	0,9754
100	p	0,3145	0,0145	0,0013	0,9654
	α	4,5765	-0,4235	1,0500	0,9352
	σ	2,1478	0,1478	0,0625	0,9698
200	p	0,3012	0,0012	0,0008	0,9580
	α	4,7509	-0,2491	0,6478	0,9180
	σ	2,0244	0,0244	0,0071	0,9700

Tabela 6 – Estudo de simulação abordagem Bayesiana sem covariáveis sob o esquema de última ativação

Tamanho Amostral (n)	Parâmetro	Média	Viés	EQM	PC
50	p	0,3693	0,0693	0,0063	0,9330
	α	4,4440	-0,5560	1,5962	0,9357
	σ	2,3398	0,3398	0,1823	0,9651
100	p	0,3159	0,0159	0,0014	0,9735
	α	4,7155	-0,2845	1,0345	0,9185
	σ	2,1368	0,1368	0,0393	0,9532
200	p	0,3050	0,0050	0,0017	0,9620
	α	4,8320	-0,1680	0,5895	0,9380
	σ	2,0480	0,0480	0,0137	0,9520

Tabela 7 – Estudo de simulação abordagem Bayesiana sem covariáveis sob o esquema de ativação aleatória

Tamanho Amostral (n)	Parâmetro	Média	Viés	EQM	PC
50	p	0,5234	0,2234	0,0557	0,7143
	α	4,5986	-0,4014	1,0841	0,9786
	σ	2,6331	0,6331	0,4802	0,9357
100	p	0,4438	0,1438	0,0276	0,8640
	α	5,0827	0,0827	1,2392	0,9880
	σ	2,4560	0,4560	0,3271	0,9080
200	p	0,3956	0,0956	0,0145	0,9115
	α	4,8380	-0,1620	1,1126	0,9714
	σ	2,3747	0,3747	0,2563	0,9193

superestimado (média $\approx 2,39$). À medida que o tamanho amostral aumenta, os vieses diminuem: p aproxima-se do valor verdadeiro, o viés de α reduz-se em magnitude (de cerca de $-0,42$ em $n = 100$ para $-0,24$ em $n = 200$) e σ passa a apresentar valores bastante próximos de 2. Os erros quadráticos médios (EQM) acompanham essa melhora, e as taxas de cobertura (PC) dos intervalos de credibilidade permanecem adequadas, próximas ao nível nominal de 95%.

A Tabela 6, referente ao mecanismo de última ativação, mostra comportamento semelhante. O parâmetro p apresenta leve superestimação, reduzindo conforme n aumenta. O parâmetro α é subestimado de forma mais acentuada em amostras pequenas (viés $\approx -0,56$ em

$n = 50$), mas essa discrepância diminui com o aumento amostral, atingindo viés em torno de $-0,17$ em $n = 200$. Já σ mostra viés positivo que decresce à medida que n cresce (de $\approx 0,34$ em $n = 50$ para $\approx 0,05$ em $n = 200$). Os EQM decrescem de maneira consistente e as taxas de cobertura mantêm-se estáveis e próximas ao nível nominal de 95%.

A Tabela 7 mostra que o esquema de ativação aleatória apresenta maior instabilidade em amostras pequenas, com superestimação acentuada de p e σ em $n = 50$ (médias de 0,52 e 2,63, respectivamente), além de taxa de cobertura reduzida para p (cerca de 0,71). À medida que o tamanho amostral aumenta, observa-se melhora nas estimativas, sobretudo em $n = 100$, quando α se aproxima bastante do valor verdadeiro e as taxas de cobertura retornam a níveis satisfatórios, próximos de 0,98. No entanto, mesmo em $n = 200$, persistem vieses moderados para p e σ , indicando que esse mecanismo, apesar da tendência de convergência, mantém maior sensibilidade e menor estabilidade relativa quando comparado aos demais esquemas de ativação. De modo geral, verifica-se um comportamento decrescente do viés e do erro quadrático médio à medida que se amplia o tamanho da amostra.

Em linhas gerais, verifica-se que todos os mecanismos de ativação convergem para os valores verdadeiros com o aumento do tamanho amostral, exibindo reduções consistentes no viés e no EQM.

No segundo estudo, foram considerados conjuntos de dados simulados contendo nenhum, um, dois ou três casos perturbados, com o objetivo de avaliar o desempenho das medidas de diagnóstico propostas.

Foi gerada uma amostra de tamanho $n = 400$ a partir do modelo BWeibcr sob o esquema de última ativação.

Para introduzir observações influentes no conjunto de dados, selecionaram-se um, dois ou três indivíduos, cujas variáveis resposta foram modificadas segundo a regra $t_i = t_i + 3S_y$, para $i = 100, 200$ e 300 , em que S_y corresponde ao desvio padrão dos t_i . Em seguida, procedeu-se ao ajuste do modelo BWeibcr sob o mesmo esquema de ativação.

Na Tabela 8 são apresentadas a estimativa média, o desvio padrão (DP) e a diferença relativa (DR, em porcentagem). O conjunto (a) corresponde ao cenário original, simulado sem perturbação, enquanto os conjuntos (b)–(h) representam cenários alternativos com casos perturbados, descritos na segunda coluna da tabela. A diferença relativa foi calculada a partir da comparação entre as estimativas obtidas nos conjuntos (b)–(h) e aquelas associadas ao conjunto (a), ela quantifica o impacto dessas perturbações sobre as estimativas pontuais, sendo calculada pela comparação entre os cenários perturbados e o cenário de referência, conforme

$$DR = \left| \frac{\hat{\theta}_{\text{com perturbação}} - \hat{\theta}_{\text{sem perturbação}}}{\hat{\theta}_{\text{sem perturbação}}} \right| \times 100.$$

Os resultados indicaram que as inferências *a posteriori* apresentaram certa sensibilidade

à inclusão das observações perturbadas.

Tabela 8 – Média, DP e DR das estimativas dos parâmetros obtidas a partir do ajuste do modelo BWeibcr sem covariável sob o esquema de última ativação

Nomes dos dados	Caso(s) perturbado(s)	Parâmetros		
		α	σ	p
		Média (DP) DR	Média (DP) DR	Média (DP) DR
a	nenhum	3,623 (0,202) -	1,070 (0,022) -	0,648 (0,019) -
b	100	2,645 (0,116) 26,994	1,036 (0,027) 3,178	0,663 (0,020) 2,315
c	200	2,515 (0,110) 30,582	1,031 (0,030) 3,645	0,665 (0,021) 2,623
d	300	2,309 (0,098) 36,268	1,015 (0,032) 5,140	0,668 (0,020) 3,086
e	{100, 200}	2,299 (0,103) 36,544	1,029 (0,033) 3,832	0,667 (0,020) 2,932
f	{100, 300}	2,170 (0,093) 40,105	1,018 (0,033) 4,860	0,669 (0,020) 3,241
g	{200, 300}	2,123 (0,093) 41,402	1,017 (0,035) 4,953	0,669 (0,020) 3,241
h	{100, 200, 300}	2,031 (0,088) 43,941	1,021 (0,037) 4,579	0,671 (0,019) 3,549

A Tabela 9 apresenta os valores dos critérios DIC, EAIC, EBIC e LPML para cada versão perturbada dos dados originais. Conforme esperado, o conjunto (a), correspondente ao cenário sem perturbações, apresentou o melhor ajuste ao modelo BWeibcr sob o esquema de última ativação.

Tabela 9 – Critérios Bayesianos de seleção de modelos no caso sem covariável

Nomes dos dados	Critérios Bayesianos			
	EAIC	EBIC	DIC	LPML
a	477,494	489,469	474,295	-237,168
b	558,550	570,525	555,500	-284,262
c	574,572	586,547	571,597	-293,204
d	598,857	610,831	595,903	-306,932
e	610,795	622,769	607,926	-309,144
f	627,874	639,848	624,817	-319,157
g	636,243	648,217	633,229	-323,050
h	658,936	670,910	655,880	-333,337

A amostra das distribuições *a posteriori* dos parâmetros do modelo BWeibcr, sob o esquema de última ativação, foi utilizada para o cálculo das medidas de divergência ψ definidas em (4.31).

Os resultados apresentados na Tabela 10 indicam que, antes da perturbação (conjunto de dados (a)), os casos selecionados não foram identificados como influentes por nenhuma das medidas de divergência ψ . Após a perturbação (conjuntos de dados (b)–(h)), entretanto, os valores das medidas aumentaram, caracterizando os casos perturbados como influentes.

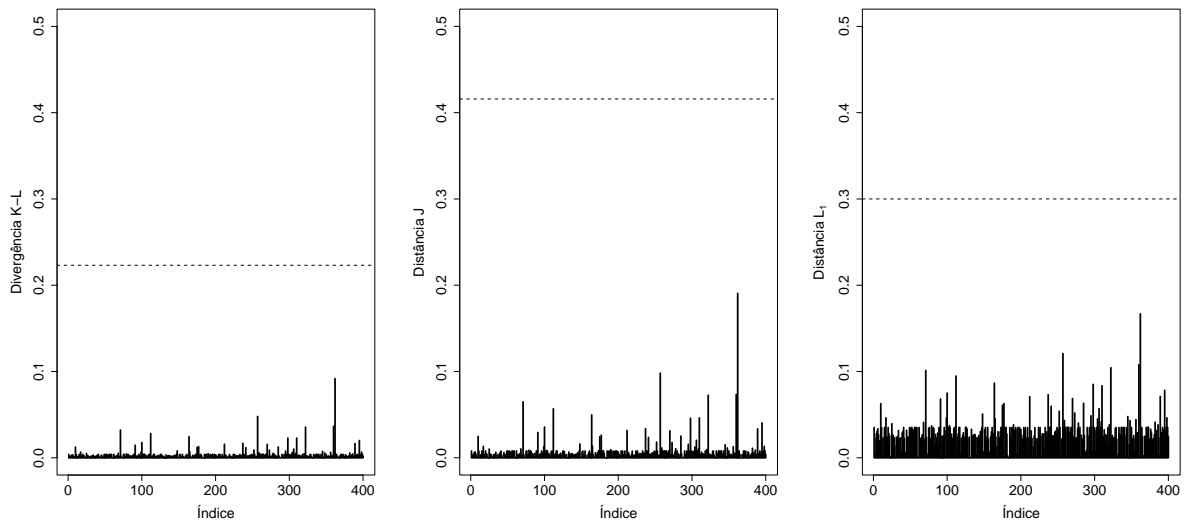
Tabela 10 – Medidas de divergência Bayesiana no caso sem covariável

Nome dos dados	Caso(s) perturbado(s)	Medidas de divergência		
		K-L	distância J	norma L_1
a	100	0,018	0,036	0,075
	200	0,003	0,007	0,033
	300	0,004	0,008	0,035
b	100	7,002	10,915	0,927
c	200	7,883	12,242	0,946
d	300	9,533	14,433	0,958
e	100	2,178	4,050	0,714
	200	3,574	6,278	0,814
f	100	1,425	2,884	0,621
	300	5,842	9,631	0,920
g	200	2,032	4,012	0,720
	300	4,939	8,731	0,885
h	100	0,835	1,727	0,493
	200	1,409	2,939	0,614
	300	3,773	7,484	0,852

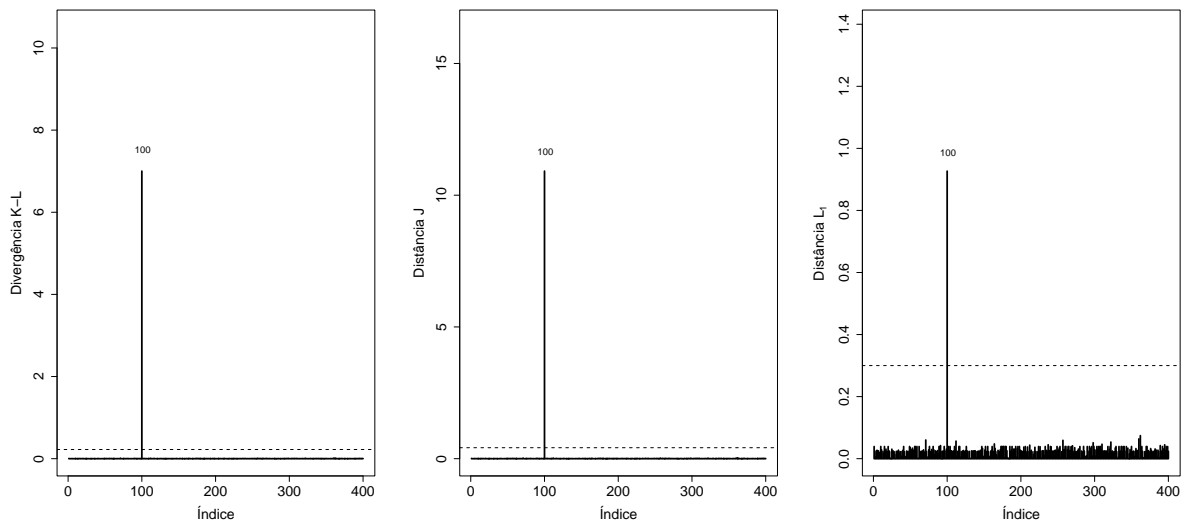
As Figuras 10, 11, 12 e 13 apresentam as medidas de divergência ψ para os conjuntos de dados (a), (b), (e) e (h), respectivamente. Na Figura 10, verifica-se que não foram identificados pontos influentes, conforme esperado, dado que nenhuma observação foi perturbada no conjunto (a). Por outro lado, nas Figuras 11, 12 e 13, as medidas de divergência identificaram corretamente as observações perturbadas como pontos influentes. Resultados semelhantes foram obtidos para os demais cenários apresentados no Apêndice B.

5.1.2 Aplicação a dados reais

O melanoma cutâneo representa uma das formas mais letais de câncer de pele, caracterizando-se pela proliferação maligna de melanócitos, células produtoras de melanina localizadas predominantemente na epiderme. Embora sua incidência seja consideravelmente menor quando comparada a outras neoplasias cutâneas, como o carcinoma basocelular – tipo mais comum de câncer de pele, originado nas células basais da camada mais profunda da epiderme, de crescimento lento, raramente metastático, mas capaz de provocar danos locais significativos se não tratado; associado principalmente à exposição solar crônica e mais prevalente em indivíduos de pele clara, especialmente em áreas expostas como rosto, pescoço e orelhas, sendo geralmente curável com diagnóstico precoce e excisão cirúrgica – e o carcinoma espinocelular – segundo tipo mais frequente de câncer de pele, com origem nas células escamosas da camada superficial

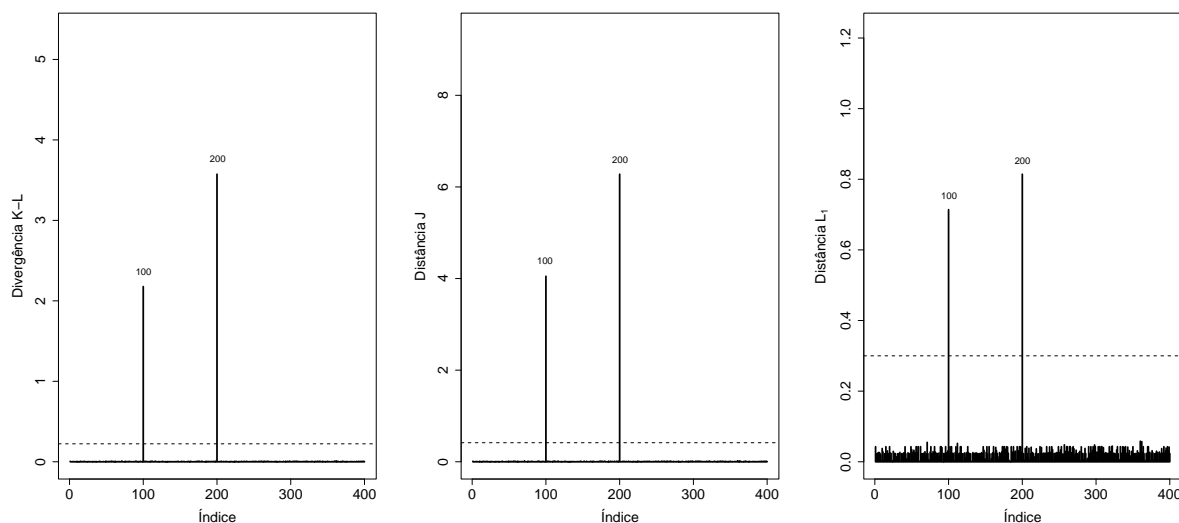
Figura 10 – Divergência ψ para o conjunto de dados (a) no caso sem covariável

Fonte: Elaborada pelo autor.

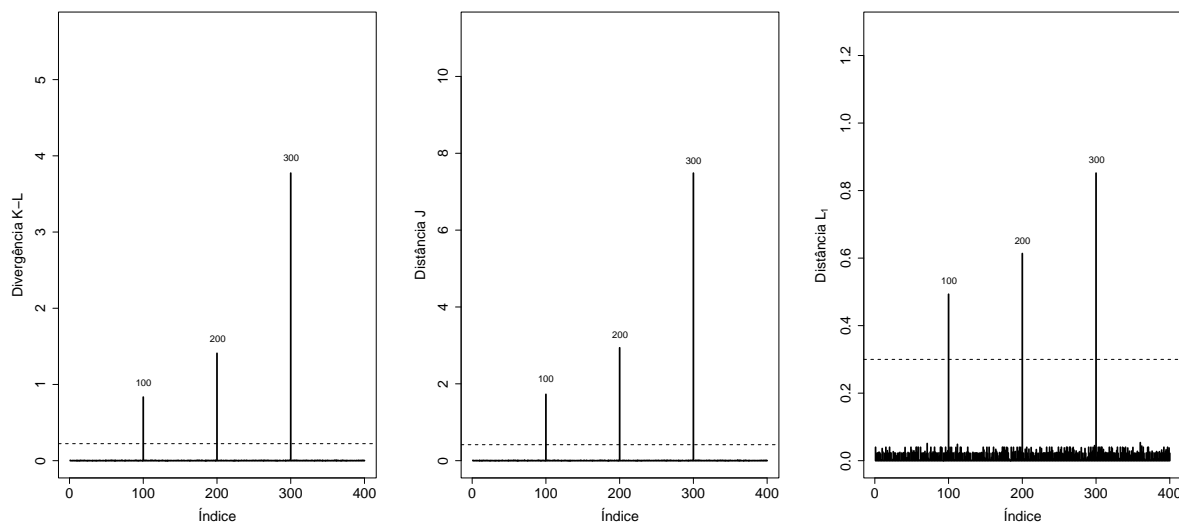
Figura 11 – Divergência ψ para o conjunto de dados (b) no caso sem covariável

Fonte: Elaborada pelo autor.

da epiderme, mais agressivo que o basocelular, com potencial de invasão profunda e metástases em estágios avançados; também relacionado à exposição solar acumulada, podendo surgir em cicatrizes, feridas crônicas ou áreas de inflamação persistente, e cujo tratamento precoce, geralmente cirúrgico, apresenta elevada taxa de cura – sua letalidade é desproporcionalmente elevada. Estima-se que, embora o melanoma represente apenas cerca de 4% dos cânceres de pele, seja responsável por mais de 75% dos óbitos decorrentes dessas patologias (ASHTON; WILKINSON, 2024). Tal discrepância evidencia a relevância clínica e estatística de investigações voltadas à identificação de seus fatores prognósticos e à compreensão da sua dinâmica de progressão.

Figura 12 – Divergência ψ para o conjunto de dados (e) no caso sem covariável

Fonte: Elaborada pelo autor.

Figura 13 – Divergência ψ para o conjunto de dados (h) no caso sem covariável

Fonte: Elaborada pelo autor.

Neste contexto, a base de dados utilizada nesta dissertação provém de um estudo clínico conduzido no Hospital Universitário de Odense, na Dinamarca, sob responsabilidade do cirurgião K.T. Drzewiecki e colaboradores. O estudo original visava investigar a relação entre características histopatológicas – características observadas em tecidos biológicos por meio de exame microscópico, após serem processados e corados em laboratório. Essas análises permitem identificar alterações estruturais nas células e tecidos, sendo fundamentais para o diagnóstico de doenças, como o câncer, e para avaliar a gravidade e o prognóstico das lesões – do tumor primário e a sobrevivência dos pacientes submetidos a tratamento cirúrgico para melanoma em

estágio clínico I (DRZEWIECKI; LADEFOGED; CHRISTENSEN, 1980). Os dados resultantes deste levantamento tornaram-se referência clássica na literatura biomédica e estatística, sendo amplamente utilizados como *benchmark* em obras metodológicas fundamentais da análise de sobrevivência, como o trabalho seminal de Andersen et al. (ANDERSEN *et al.*, 2012).

O conjunto, intitulado *Melanoma Survival Data*, encontra-se acessível no pacote `timereg` do software R, e compreende 205 observações referentes a pacientes acompanhados prospectivamente após a realização de cirurgia (DRZEWIECKI, 2004). As variáveis disponíveis abrangem aspectos clínicos, demográficos e de desfecho, possibilitando uma rica gama de análises estatísticas, especialmente no campo da modelagem do tempo até a ocorrência de eventos.

A seguir, apresenta-se uma descrição detalhada das variáveis incluídas na base de dados:

- **no**: identificador numérico atribuído a cada paciente. Esta variável não possui relevância clínica ou estatística direta, sendo empregada apenas para fins de indexação;
- **status**: variável categórica que descreve o estado vital do paciente ao final do período de observação. Os valores possíveis são: 1 – óbito decorrente de melanoma (evento de interesse), 2 – vivo ao fim do acompanhamento (censura à direita), e 3 – óbito por outras causas (censura por evento competitivo). Para a aplicação de modelos tradicionais de sobrevivência com falha única, é comum recodificar esta variável em uma forma binária, onde apenas o evento 1 é tratado como falha e os demais como censura;
- **days**: tempo de seguimento, expresso em dias, compreendido entre a data da cirurgia e o evento de interesse (ou censura). Esta variável constitui o eixo temporal fundamental para a análise de sobrevivência;
- **ulc**: variável binária que indica a presença (1) ou ausência (0) de ulceração na lesão tumoral primária;
- **thick**: espessura do tumor, medida em centésimos de milímetro;
- **sex**: variável binária indicativa do sexo biológico do paciente, sendo 0 para mulheres e 1 para homens.

A qualidade e completude dessa base tornam-na particularmente valiosa para a aplicação de modelos estatísticos voltados à análise do tempo até a ocorrência de eventos. As características de censura à direita e a existência de diferentes covariáveis explanatórias habilitam o uso de uma ampla gama de técnicas de modelagem estatística, tanto em sua forma clássica quanto em abordagens modernas com estruturas latentes ou não observáveis.

Nesta dissertação, adota-se uma abordagem moderna e flexível para a modelagem da sobrevivência, mediante a aplicação de um modelo com fração de cura e estrutura latente de ativação. A formulação matemática do modelo é apresentada no Capítulo 4, enquanto os

procedimentos de estimação clássica e Bayesiana são discutidos em detalhes nas Seções 4.1 e 4.2, respectivamente.

Para fins de análise, os dados originalmente expressos em dias foram convertidos para anos. O tempo de sobrevivência, medido em anos, foi selecionado como a variável resposta. Ao todo, tem-se 205 unidades amostrais com 72,2% de censura.

Na Tabela 11 apresentam-se medidas resumo da variável de interesse com dados censurados e não censurados.

Tabela 11 – Medidas resumo da variável de interesse

Estatística	Valor
Mínimo	0,0274
1° Quartil	4,1781
Mediana	5,4932
Média	5,8981
3° Quartil	8,3342
Máximo	15,2466
Desvio Padrão	3,0741

(i) Abordagem Clássica

Nesta parte, os modelos com diferentes abordagens de ativação serão ajustados por meio de inferência clássica, utilizando o método da máxima verossimilhança. Essa abordagem fundamenta-se nos teoremas clássicos da estatística, os quais asseguram, sob condições regulares, que os estimadores obtidos sejam consistentes, assintoticamente normais e eficientes. Com base nessas propriedades, serão calculados os erros padrão e, a partir dessas estimativas, os intervalos de confiança para os parâmetros. Além disso, métricas como o AIC, o BIC e o logaritmo da verossimilhança também foram computadas com o objetivo de comparar o desempenho dos modelos ajustados. Por fim, apresentam-se os resíduos de Cox-Snell detalhados anteriormente.

Tabela 12 – Estimativas de máxima verossimilhança do modelo sem covariáveis

Ativação	Parâmetro	Estimativa	Erro Padrão	IC (95%)
Primeira	α	1,6568	0,2150	(1,2353, 2,0783)
	σ	5,3327	0,8977	(3,5732, 7,0922)
	p	0,4079	0,0352	(0,3389, 0,4770)
Última	α	1,5399	0,2015	(1,1450, 1,9348)
	σ	4,4525	0,5557	(3,3633, 5,5417)
	p	0,4050	0,0330	(0,3402, 0,4697)
Aleatória	α	1,5998	0,2088	(1,1905, 2,0090)
	σ	4,8966	0,7138	(3,4976, 6,2955)
	p	0,4071	0,0344	(0,3397, 0,4745)

A Tabela 12 apresenta as estimativas de máxima verossimilhança do modelo BWeibcr sem covariáveis para os três mecanismos de ativação considerados. Observa-se que os valores

estimados de α e σ variam de forma moderada entre os mecanismos, indicando ajustes semelhantes quanto à forma e à escala da distribuição Weibull. As estimativas de p permanecem próximas de 0,40 em todos os casos, sugerindo uma fração de cura estável e consistente entre as diferentes estruturas de ativação.

Tabela 13 – Proporção de cura: $\pi = 2p^3 - 3p^2 + 1$

Ativação	π
Primeira	0,6365
Última	0,6408
Aleatória	0,6378

A Tabela 13 apresenta as estimativas das frações de indivíduos não suscetíveis. Observa-se que os valores obtidos são bastante semelhantes entre os diferentes mecanismos de ativação, com $\pi \approx 0,64$ tanto para Primeira, Última quanto para Aleatória, apresentando apenas pequena variação. Esse resultado evidencia que, na ausência de covariáveis, a proporção de indivíduos curados estimada pelo modelo mostra-se essencialmente robusta à escolha do mecanismo de ativação, o que se reflete na estabilidade observada em p .

Tabela 14 – Métricas dos modelos ajustados sem a presença de covariáveis

Ativação	AIC	BIC	LL
Primeira	458,5480	468,5171	-226,2740
Última	458,8786	468,8477	-226,4393
Aleatória	458,6796	468,6486	-226,3398

No que se refere às métricas de ajuste, a Tabela 14 apresenta os valores dos critérios AIC, BIC e da log-verossimilhança para os três mecanismos de ativação considerados. Verifica-se que os resultados são bastante próximos entre os modelos: a ativação Primeira apresenta os menores valores de AIC e BIC, bem como a maior log-verossimilhança (menos negativa). Entretanto, as diferenças observadas em AIC e BIC em relação às demais alternativas são inferiores a 0,35, magnitude que pode ser interpretada como evidência de suporte quase equivalente entre os modelos. Dessa forma, na ausência de covariáveis, não há indicação robusta para a escolha de um mecanismo específico apenas com base nesses critérios globais.

A seleção pode, portanto, apoiar-se em aspectos substantivos, estabilidade numérica ou diagnósticos residuais, uma vez que o desempenho dos modelos é praticamente indistinguível em termos de AIC e BIC. Ainda assim, do ponto de vista técnico, a ativação Primeira apresenta ligeira superioridade nas métricas avaliadas.

A Figura 14 apresenta as curvas de sobrevivência empíricas estimadas via Kaplan–Meier (em preto) juntamente com as curvas ajustadas pelo modelo com fração de cura (em vermelho), considerando os três mecanismos de ativação: Primeira, Última e Aleatória, todos na ausência de covariáveis. Visualmente, observa-se boa concordância entre as curvas ajustadas e as curvas empíricas ao longo de todo o horizonte temporal. De modo geral, os três mecanismos

capturam adequadamente o platô da curva de sobrevivência, associado à fração de indivíduos não suscetíveis, reforçando a robustez do modelo quanto à escolha da forma de ativação em cenários sem covariáveis. Esse resultado está em consonância com as estimativas próximas de π observadas na Tabela 13, evidenciando consistência entre a análise numérica e visual.

A Figura 15 apresenta os resíduos de Cox–Snell obtidos a partir dos modelos ajustados sem covariáveis, sob os três mecanismos de ativação: Primeira, Última e Aleatória. Cada painel mostra a função de risco acumulada estimada a partir dos resíduos sobreposta à reta identidade, que representa o comportamento esperado sob adequada especificação do modelo.

De modo geral, os três mecanismos produzem resíduos que se alinham de forma satisfatória à reta identidade, indicando ajuste consistente do modelo em todos os casos. Pequenas discrepâncias locais podem ser observadas, sobretudo em regiões intermediárias, mas não configuram desvios sistemáticos relevantes. Assim, os resultados sugerem que, na ausência de covariáveis, o modelo de fração de cura baseado na distribuição Weibull para os tempos latentes e na distribuição discreta de Bilal para M apresenta especificação adequada, independentemente do mecanismo de ativação adotado. Esse diagnóstico visual corrobora as métricas globais de ajuste e a estabilidade da estimativa da fração de não suscetíveis, reforçando a robustez do modelo.

(ii) Abordagem Bayesiana

Conforme discutido anteriormente, o modelo BWeibcr sob o esquema de primeira ativação apresentou o melhor ajuste aos dados de melanoma. Na sequência, o mesmo modelo é ajustado novamente ao conjunto de dados, agora sob a abordagem Bayesiana.

Foram simuladas duas cadeias de tamanho 25.000 para cada parâmetro, desconsiderando-se as primeiras 5.000 iterações a fim de eliminar a influência dos valores iniciais (*burn-in*). Para reduzir a autocorrelação entre as amostras, foi adotado um espaçamento de 10 iterações, resultando em uma amostra final de tamanho 4.000, utilizada para a realização da inferência *a posteriori*.

A convergência das cadeias foi avaliada de acordo com os procedimentos recomendados por Cowles e Carlin (1996). Em todos os cenários, a convergência foi confirmada pelo diagnóstico de Gelman–Rubin (GELMAN; RUBIN, 1992), com valores próximos de 1 (≤ 1.01), além da inspeção visual dos gráficos de traços.

O resumo *a posteriori* dos parâmetros, juntamente com os limites inferiores e superiores dos intervalos de credibilidade de 95%, encontra-se apresentado na Tabela 15.

Tabela 15 – Resumo *a posteriori* dos parâmetros do modelo BWeibcr sob o esquema de primeira ativação

Parâmetro	Média	DP	IC (95%)
α	1,552	0,188	(1,203, 1,933)
σ	5,437	0,282	(5,020, 5,958)
p	0,424	0,031	(0,365, 0,488)

As medidas de divergência ψ foram calculadas e encontram-se apresentadas na Tabela 29, localizada no Apêndice C, e ilustradas na Figura 16. Constata-se que, em todas as três medidas, nenhuma observação foi classificada como influente.

As Figuras 17 e 18 apresentam, respectivamente, os gráficos de traços e de autocorrelação dos parâmetros do modelo. Os gráficos de traços (Figura 17) e da autocorrelação (Figura 18) evidenciam um comportamento adequado das cadeias geradas pelo método MCMC. Nos gráficos de traços, observa-se rápida mistura das cadeias, sem tendência aparente ou padrões sistemáticos, o que indica boa exploração do espaço paramétrico. Já os gráficos de autocorrelação apresentam valores próximos de zero a partir de poucas defasagens, sugerindo baixa dependência entre as observações sucessivas. Em conjunto, esses resultados reforçam a convergência das cadeias e a confiabilidade das inferências *a posteriori* dos parâmetros do modelo.

Por fim, a Figura 19 apresenta o boxplot *a posteriori* da proporção de indivíduos curados.

5.2 Modelo Com Covariável

Do mesmo modo ao abordado na análise sem covariável, apresentada na Seção 5.1, examina-se, neste ponto, a inclusão de covariável no modelo. Essa inclusão se dá por meio da parametrização do parâmetro p da distribuição de Bilal, associado à fração de cura, de modo a vinculá-lo à covariável por um preditor linear (e.g., $p(x) = \text{logit}^{-1}(\beta_0 + \beta_1 x)$). Mantêm-se os tempos de falha descritos pela distribuição de Weibull e a variável latente M definida pela distribuição discreta de Bilal, enquanto a ligação com covariável possibilita quantificar sua influência sobre a dinâmica de suscetíveis e curados, permitindo ainda a comparação entre diferentes mecanismos de ativação (primeira, última e aleatória).

Com o objetivo de avaliar o desempenho do modelo de sobrevivência com fração de cura incorporando covariável, desenvolve-se inicialmente um estudo de simulação sob as abordagens clássica e Bayesiana. Na abordagem clássica, a inferência é realizada via estimadores de máxima verossimilhança e intervalos de confiança obtidos a partir da informação observada de Fisher, possibilitando testes formais sobre os coeficientes. Já na abordagem Bayesiana, a estimação é conduzida por meio de métodos de Monte Carlo via Cadeias de Markov (MCMC), permitindo explorar integralmente a incerteza dos parâmetros e avaliar a robustez dos mecanismos de ativação.

Na sequência, procede-se à aplicação do modelo a dados reais, considerando a presença de covariável, de modo a verificar sua capacidade de explicar diferenças nas taxas de cura observadas. Por fim, são apresentados resultados de diagnóstico e análise de influência, que possibilitam avaliar a adequação do ajuste e a sensibilidade das conclusões em relação às observações disponíveis.

5.2.1 Estudo de simulação

Para o cenário com covariável, a única diferença em relação ao processo de geração de dados descrito na Seção 5.1.1 é a inclusão de uma covariável e a parametrização correspondente de p_i ; assume-se, para $i = 1, \dots, n$, uma covariável $x_i \sim \text{Bernoulli}(0,5)$ e modela-se o parâmetro da distribuição Bilal por $\text{logit}(p_i) = \beta_0 + \beta_1 x_i$; utilizam-se como valores verdadeiros $\alpha = 5$, $\sigma = 2$, $\beta_0 = -0,5$ e $\beta_1 = 0,75$ com um nível de censura de 75%.

Avalia-se a capacidade de recuperar corretamente os parâmetros (α, σ) , os coeficientes de regressão (β_0, β_1) e, conseqüentemente, a fração de cura π . Os indicadores considerados incluem viés, EQM, e probabilidade de cobertura – tanto de intervalos de confiança quanto de intervalos de credibilidade.

(i) Abordagem Clássica

No estudo de simulação com covariável, foram realizadas 1.000 réplicas para os tamanhos amostrais $n \in \{50, 100, 200, 400\}$, adotando-se como valores verdadeiros dos parâmetros $\alpha = 5$, $\sigma = 2$, $\beta_0 = -0,5$ e $\beta_1 = 0,75$. Em cada réplica, obtiveram-se as estimativas pontuais, o viés médio, o erro quadrático médio (EQM) e as taxas de cobertura (PC) dos intervalos de confiança assintóticos, calculados com base na informação observada de Fisher.

Tabela 16 – Estudo de simulação abordagem clássica com covariáveis sob o esquema de primeira ativação

Tamanho Amostral (n)	Parâmetro	Estimativa	Viés	EQM	PC
50	β_0	-0,6677	-0,1677	1,6776	0,976
	β_1	0,9696	0,2196	1,8726	0,962
	α	5,6697	0,6697	2,2959	0,966
	σ	1,9973	-0,0027	0,0355	0,903
100	β_0	-0,5173	-0,0173	0,0920	0,974
	β_1	0,7884	0,0384	0,1658	0,948
	α	5,3967	0,3967	1,0493	0,956
	σ	1,9961	-0,0039	0,0153	0,922
200	β_0	-0,5093	-0,0093	0,0434	0,961
	β_1	0,7499	-0,0001	0,0774	0,957
	α	5,1518	0,1518	0,4302	0,936
	σ	1,9949	-0,0051	0,0071	0,947
400	β_0	-0,4963	0,0037	0,0209	0,949
	β_1	0,7560	0,0060	0,0366	0,940
	α	5,0713	0,0713	0,1894	0,949
	σ	2,0014	0,0014	0,0034	0,951

A Tabela 16 apresenta os resultados obtidos sob o esquema de primeira ativação. Nota-se que, para todos os tamanhos amostrais, as estimativas se aproximaram dos valores verdadeiros, com redução progressiva dos vieses e dos EQM à medida que n aumentou, refletindo a consistência assintótica do estimador de máxima verossimilhança. As taxas de cobertura mantiveram-se próximas ao nível nominal de 95%, com valores variando entre 0,90 e 0,98, indicando boa adequação dos intervalos de confiança mesmo em amostras menores.

Tabela 17 – Estudo de simulação abordagem clássica com covariáveis sob o esquema de última ativação

Tamanho Amostral (n)	Parâmetro	Estimativa	Viés	EQM	PC
50	β_0	-0,7015	-0,2015	1,6738	0,9800
	β_1	1,0323	0,2823	2,0423	0,9559
	α	5,6227	0,6227	2,9743	0,9659
	σ	1,9663	-0,0337	0,0188	0,9238
100	β_0	-0,5230	-0,0230	0,1151	0,9499
	β_1	0,7832	0,0332	0,2258	0,9399
	α	5,3383	0,3383	1,1200	0,9559
	σ	1,9938	-0,0062	0,0086	0,9379
200	β_0	-0,5065	-0,0065	0,0480	0,9439
	β_1	0,7580	0,0080	0,0985	0,9599
	α	5,1283	0,1283	0,4154	0,9519
	σ	1,9929	-0,0071	0,0042	0,9238
400	β_0	-0,4995	0,0005	0,0217	0,9570
	β_1	0,7532	0,0032	0,0452	0,9480
	α	5,0604	0,0604	0,1823	0,9620
	σ	1,9956	-0,0044	0,0019	0,9480

Tabela 18 – Estudo de simulação abordagem clássica com covariáveis sob o esquema de ativação aleatória

Tamanho Amostral (n)	Parâmetro	Estimativa	Viés	EQM	PC
50	β_0	-0,6060	-0,1060	1,1420	0,976
	β_1	1,0014	0,2514	2,5626	0,966
	α	5,5720	0,5720	2,0006	0,967
	σ	1,9850	-0,0150	0,0189	0,906
100	β_0	-0,5275	-0,0275	0,0960	0,961
	β_1	0,8095	0,0595	0,2486	0,948
	α	5,2651	0,2651	0,8447	0,951
	σ	1,9950	-0,0050	0,0082	0,937
200	β_0	-0,5181	-0,0181	0,0434	0,960
	β_1	0,7733	0,0233	0,0829	0,953
	α	5,1519	0,1519	0,3743	0,955
	σ	1,9956	-0,0044	0,0043	0,930
400	β_0	-0,5000	0,0001	0,0203	0,948
	β_1	0,7552	0,0052	0,0421	0,941
	α	5,0334	0,0334	0,1434	0,961
	σ	1,9974	-0,0026	0,0020	0,940

Na Tabela 17, referente ao esquema de última ativação, observa-se comportamento semelhante ao caso anterior. As estimativas tenderam aos valores verdadeiros e, com o crescimento amostral, tanto o viés quanto o EQM apresentaram diminuição sistemática. As taxas de cobertura permaneceram estáveis, majoritariamente acima de 0,92, sinalizando que a incerteza associada às estimativas foi adequadamente captada pelos intervalos de confiança. Em comparação com a primeira ativação, as diferenças foram marginais, reforçando a robustez do procedimento de estimação.

Por fim, a Tabela 18 reporta os resultados sob o esquema de ativação aleatória. Assim como nos casos anteriores, verifica-se a convergência das estimativas para os valores verdadeiros dos parâmetros, com declínio progressivo dos vieses e EQM conforme o aumento de n . As taxas de cobertura mantiveram-se próximas a 95%, com oscilações pequenas em torno desse valor. Em termos comparativos, o desempenho obtido foi bastante similar aos demais mecanismos de ativação, evidenciando que, sob a presença de covariável, não se observaram discrepâncias relevantes entre os esquemas avaliados.

Como análise de simulação complementar com covariável, avaliou-se o comportamento dos resíduos de Cox–Snell em dois cenários: (i) especificação correta – dados gerados pelo BWeibcr sob o mecanismo de última ativação, com covariável $x_i \sim \text{Bernoulli}(0,5)$ e parâmetros verdadeiros $\alpha = 5$, $\sigma = 2$, $\beta_0 = -0,5$ e $\beta_1 = 0,75$; o ajuste foi realizado por máxima verossimilhança no próprio modelo com covariável e última ativação, seguido do cálculo dos respectivos resíduos de Cox–Snell; (ii) especificação incorreta – dados também gerados sob última ativação, porém ajustados pela verossimilhança correspondente ao mecanismo de primeira ativação (mantendo a covariável no preditor). O estudo considerou 1.000 réplicas, tamanhos amostrais $n = 50, 100, 200$ e 400 e nível de censura de 72%.

Na Figura 20, referente ao modelo corretamente especificado com covariável, a nuvem de pontos permanece próxima da reta identidade, sobretudo a partir de $n \geq 200$, com discreta sublinearidade nas caudas – padrão coerente com o comportamento teórico dos resíduos de Cox–Snell. Em contraste, a Figura 21, correspondente ao ajuste com mecanismo de ativação incorreto, exhibe afastamento sistemático da referência, sobretudo para $n = 50$. Visualmente, a maior parte da nuvem de pontos nos quatro tamanhos amostrais concentra-se acima da reta identidade.

(ii) Abordagem Bayesiana

O estudo de simulação Bayesiano com covariável considerou uma covariável binária $X \sim \text{Bernoulli}(0,5)$ e 200 réplicas para $n \in \{50, 100, 200\}$, sob os valores verdadeiros $\alpha = 5$, $\sigma = 2$, $\beta_0 = -0,5$ e $\beta_1 = 0,75$. Os objetivos foram: (i) avaliar o comportamento das estimativas dos parâmetros; e (ii) examinar o desempenho da estratégia diagnóstica proposta em cenários com 0, 1, 2 ou 3 observações perturbadas. Para cada conjunto, executaram-se duas cadeias MCMC com 25.000 iterações; descartaram-se as 5.000 iniciais (*burn-in*) e aplicou-se *thinning* de 10, perfazendo amostra de tamanho 4.000 para *a posteriori*, usadas no cálculo das médias, intervalos de credibilidade de 95% e critérios de seleção de modelos.

A convergência foi monitorada conforme as recomendações de Cowles e Carlin (1996) e confirmada pelo diagnóstico de Gelman–Rubin ($\hat{R} \leq 1,01$), complementada pela inspeção visual dos traços (GELMAN; RUBIN, 1992).

A Tabela 19 apresenta os resultados do estudo de simulação Bayesiana com covariável sob o esquema de primeira ativação. Nota-se que, para $n = 50$, os vieses são relativamente

Tabela 19 – Estudo de simulação abordagem Bayesiana com covariável sob o esquema de primeira ativação

Tamanho Amostral (n)	Parâmetro	Média	Viés	EQM	PC
50	β_0	-0,4327	0,0673	0,2391	0,9487
	β_1	0,3757	-0,3743	18,1695	0,8974
	α	6,7059	1,7059	4,0852	0,9231
	σ	1,9975	-0,0025	0,0115	0,9744
100	β_0	-0,5080	-0,008	0,1092	0,9479
	β_1	0,9469	0,1969	0,2830	0,9446
	α	5,3955	0,3955	0,8108	0,9577
	σ	2,0021	0,0021	0,0073	0,9642
200	β_0	-0,5173	-0,0173	0,0444	0,9475
	β_1	0,7945	0,0445	0,0865	0,9475
	α	5,0507	0,0507	0,3637	0,9455
	σ	2,0098	0,0098	0,0041	0,9495

Tabela 20 – Estudo de simulação abordagem Bayesiana com covariável sob o esquema de última ativação

Tamanho Amostral (n)	Parâmetro	Média	Viés	EQM	PC
50	β_0	-0,2062	0,2938	7,3136	0,9415
	β_1	1,1520	0,4020	6,7561	0,9298
	α	4,8523	-0,1477	2,8630	0,9474
	σ	1,9051	-0,0949	0,0409	0,9532
100	β_0	-0,4850	0,0150	0,1609	0,9267
	β_1	0,8210	0,0710	0,2674	0,9311
	α	5,0663	0,0663	0,9782	0,9511
	σ	1,9985	-0,0015	0,0104	0,9511
200	β_0	-0,5065	-0,0065	0,0474	0,9427
	β_1	0,7671	0,0171	0,0979	0,9604
	α	4,9946	-0,0054	0,4068	0,9471
	σ	1,9987	-0,0013	0,0042	0,9383

elevados, em especial para β_1 e α , cujos valores médios distanciam-se bastante dos parâmetros verdadeiros. À medida que o tamanho amostral cresce, observa-se clara melhora, com redução progressiva do viés e do EQM, além dos percentuais de cobertura (PC) mais próximos do nível nominal de 95%. Essa tendência reforça a consistência assintótica do método Bayesiano empregado.

Na Tabela 20, correspondente ao mecanismo de última ativação, o comportamento geral é semelhante. Para amostras reduzidas ($n = 50$), destacam-se novamente vieses maiores, sobretudo em β_0 e β_1 , mas que diminuem conforme n aumenta. O parâmetro σ mantém boa estabilidade em todos os cenários, apresentando viés próximo de zero e elevadas taxas de cobertura. No geral, há convergência satisfatória para os valores verdadeiros quando a amostra é ampliada.

Já a Tabela 21, que contempla o esquema de ativação aleatória, apresenta resultados consistentes com os anteriores, embora com maior instabilidade em $n = 50$, sobretudo para α e β_1 , cujos vieses e EQM são mais elevados. Com o aumento da amostra, verifica-se maior proximidade

Tabela 21 – Estudo de simulação abordagem Bayesiana com covariável sob o esquema de ativação aleatória

Tamanho Amostral (n)	Parâmetro	Média	Viés	EQM	PC
50	β_0	-0,5360	-0,0360	0,3174	0,9048
	β_1	1,4003	0,6503	1,0237	0,8571
	α	7,1094	2,1094	6,7201	0,9048
	σ	1,8524	-0,1476	0,0320	0,9143
100	β_0	-0,5699	-0,0699	0,1302	0,9382
	β_1	0,9790	0,2290	0,3275	0,9618
	α	5,6020	0,6020	0,8859	0,9706
	σ	1,8575	-0,1425	0,0291	0,9662
200	β_0	-0,5720	-0,0720	0,0477	0,9437
	β_1	0,8012	0,0512	0,0914	0,9356
	α	5,3790	0,3790	0,3899	0,9678
	σ	1,8597	-0,1403	0,0236	0,9479

entre as estimativas médias e os valores verdadeiros, redução do EQM e taxas de cobertura dentro do esperado. Assim, apesar das diferenças iniciais, os três mecanismos de ativação apresentam desempenhos bastante semelhantes sob a abordagem Bayesiana, distinguindo-se apenas pela maior sensibilidade em amostras pequenas.

Assim como no modelo sem covariável (ver Seção 5.1.1), este segundo estudo com covariável tem por objetivo avaliar a robustez das medidas de diagnóstico sob perturbações controladas. Gerou-se uma amostra de tamanho 400 a partir do modelo BWeibcr sob o esquema de última ativação e, em seguida, o mesmo modelo foi ajustado aos dados perturbados.

Consideraram-se cenários com 0, 1, 2 e 3 observações influentes. Para introduzir as perturbações, adotou-se a regra

$$t_i \leftarrow t_i + 3S_y, \quad i \in \{100, 150, 325\},$$

em que S_y é o desvio-padrão dos tempos $\{t_i\}$; os índices foram incluídos conforme o número de casos perturbados do cenário. A Tabela 22 apresenta, para cada configuração, a estimativa média, o desvio-padrão (DP) e a diferença relativa (DR, em %), sendo o conjunto (a) a linha de base (sem perturbações) e (b)–(h) os cenários com perturbações; a DR foi calculada em relação a (a). Os resultados são análogos aos obtidos no caso sem covariável (cf. Seção 5.1.1): as inferências *a posteriori* indicaram alguma sensibilidade frente às perturbações introduzidas.

A Tabela 23 reporta os valores de DIC, EAIC, EBIC e LPML para cada versão perturbada dos dados. Como esperado, o conjunto (a) – sem perturbações – apresenta o melhor desempenho no ajuste do modelo BWeibcr sob o esquema de última ativação (menores DIC/EAIC/EBIC e maior LPML).

Assim como no caso sem covariável, empregou-se a amostra *a posteriori* do modelo BWeibcr sob última ativação para calcular as medidas de divergência ψ (Equação (4.31)). Como antes, a Tabela 24 indica que o cenário (a) não aponta observações influentes; após as

Tabela 22 – Média, DP e DR das estimativas dos parâmetros obtidas a partir do ajuste do modelo BWeiber com covariável sob o esquema de última ativação

Nomes dos dados	Caso(s) perturbado(s)	Parâmetros			
		α Média (DP) DR	σ Média(DP) DR	β_0 Média(DP) DR	β_1 Média(DP) DR
a	nenhum	5,416 (0,461)	2,073 (0,046)	-0,408 (0,180)	0,366 (0,222)
b	100	3,597 (0,276) 33,586	2,214 (0,078) 6,802	-0,315 (0,195) 22,794	0,363 (0,250) 0,820
c	150	3,295 (0,257) 39,162	2,256 (0,097) 8,828	-0,308 (0,201) 24,510	0,406 (0,250) 10,929
d	325	3,404 (0,274) 37,149	2,242 (0,089) 8,152	-0,310 (0,204) 24,020	0,389 (0,257) 6,284
e	{100,150}	2,955 (0,279) 45,439	2,405 (0,126) 16,015	-0,129 (0,323) 68,382	0,322 (0,330) 12,022
f	{100,325}	3,020 (0,257) 44,239	2,381 (0,122) 14,858	-0,202 (0,243) 50,490	0,370 (0,284) 1,093
g	{150,325}	2,851 (0,255) 47,360	2,442 (0,137) 17,800	-0,126 (0,300) 69,118	0,386 (0,327) 5,464
h	{100,150,325}	2,520 (0,382) 53,471	2,537 (0,256) 22,383	0,409 (0,811) 200,245	0,070 (0,605) 80,874

Tabela 23 – Critérios Bayesianos de seleção de modelos caso com covariável

Nomes dos dados	Critérios Bayesianos			
	EAIC	EBIC	DIC	LPML
a	349,181	365,146	344,985	-172,463
b	389,737	405,703	385,805	-196,927
c	398,189	414,155	394,390	-201,597
d	394,997	410,963	391,108	-201,833
e	409,759	425,725	405,897	-206,581
f	407,485	423,451	403,420	-204,488
g	412,018	427,984	408,130	-207,567
h	418,809	434,775	412,181	-209,580

perturbações (b)–(h), os valores de ψ aumentam e passam a sinalizar os casos alterados como influentes.

As Figuras 22, 23, 24 e 25 exibem as medidas de divergência ψ para os cenários (a), (b), (e) e (h), respectivamente. No cenário (a), sem perturbações, nenhuma observação foi sinalizada como influente, como esperado. Já nos cenários (b), (e) e (h), as medidas identificam corretamente as observações perturbadas. Resultados análogos são obtidos nos demais cenários; ver Apêndice B.

Tabela 24 – Medidas de divergência Bayesiana no caso com covariável

Nome dos dados	Caso(s) perturbado(s)	Medidas de divergência		
		K-L	distância J	norma L_1
a	100	0,014	0,029	0,068
	150	0,008	0,015	0,050
	325	0,007	0,014	0,047
b	100	4,658	8,211	0,889
c	150	5,022	8,647	0,891
d	325	6,897	11,933	0,954
e	100	1,469	3,102	0,642
	150	2,814	5,630	0,795
f	100	1,310	2,774	0,606
	325	2,061	4,377	0,727
g	150	2,441	4,940	0,774
	325	1,704	3,586	0,680
h	100	0,828	1,657	0,472
	150	1,128	2,508	0,580
	325	0,783	1,721	0,497

5.2.2 Aplicação a dados reais

Conforme apresentado na Seção 5.1.2, encontra-se ali a caracterização pormenorizada do conjunto de dados de melanoma; nesta etapa, essa base será empregada no processo de modelagem com a inclusão da covariável sexo.

(i) Abordagem Clássica

Com o objetivo de avaliar o desempenho do modelo de sobrevivência com fração de cura incorporando efeitos de covariável, procedeu-se ao ajuste sob a estatística clássica via máxima verossimilhança. Mantêm-se os tempos de falha modelados pela distribuição de Weibull e a variável latente M descrita pela distribuição discreta de Bilal, enquanto o parâmetro que governa M é ligado às covariáveis por um preditor linear (e.g., $p(x) = \text{logit}^{-1}(\beta_0 + \beta_1 x)$). Esse delineamento permite quantificar a influência das covariáveis sobre a dinâmica de suscetíveis/curados e comparar o comportamento do modelo entre diferentes mecanismos de ativação (primeira, última e aleatória). A inferência é conduzida por estimativas de máxima verossimilhança, com erros-padrão e intervalos de confiança obtidos a partir da informação observada de Fisher, possibilitando testes e conclusões formais sobre os coeficientes.

A Tabela 25 apresenta as estimativas de máxima verossimilhança para os parâmetros do modelo com tempos latentes Weibull (α, σ) e para os coeficientes (β_0, β_1) que ligam o parâmetro p da distribuição de M à covariável por meio de um preditor linear, $p(x) = \text{logit}^{-1}(\beta_0 + \beta_1 x)$. Define-se x como indicador de sexo feminino ($x = 0$ para homens e $x = 1$ para mulheres), de modo que β_1 quantifica a diferença entre mulheres e homens no parâmetro p . Intervalos de confiança que não contêm 0 indicam evidência de significância estatística ao nível de 5% de significância.

Tabela 25 – Estimativas de máxima verossimilhança do modelo

Ativação	Parâmetro	Estimativa	Erro Padrão	IC (95%)
Primeira	α	1,6662	0,2164	(1,2421, 2,0903)
	σ	5,3856	0,9250	(3,5726, 7,1986)
	β_0	-0,0413	0,2020	(-0,4372, 0,3546)
	β_1	-0,5491	0,2220	(-0,9843, -0,1139)
Última	α	0,8673	0,1264	(0,6195, 1,1150)
	σ	6,2002	1,2625	(3,7256, 8,6747)
	β_0	0,5221	0,4461	(-0,3523, 1,3965)
	β_1	1,2943	0,5732	(0,1708, 2,4178)
Aleatória	α	1,6014	0,2076	(1,1944, 2,0084)
	σ	4,8514	0,6894	(3,5002, 6,2025)
	β_0	-0,0579	0,2070	(-0,4636, 0,3478)
	β_1	-0,5437	0,2384	(-1,0110, -0,0764)

Os valores de forma e escala, α e σ respectivamente, são bem identificados em todos os esquemas de ativação. Observa-se $\alpha > 1$ nas ativações Primeira ($\hat{\alpha} \approx 1,67$) e Aleatória ($\hat{\alpha} \approx 1,60$), o que é compatível com risco latente crescente no tempo; já na ativação Última $\alpha < 1$ ($\hat{\alpha} \approx 0,87$), sugere risco latente decrescente. As estimativas de σ situam-se entre $\approx 4,85$ e $\approx 6,20$, com intervalos relativamente amplos, porém afastados de zero, indicando escalas plausíveis e consistentes entre os mecanismos.

O intercepto β_0 , em todos os esquemas, apresenta intervalos de confiança que incluem 0, indicando ausência de evidência de deslocamento sistemático no logito de p para o grupo de referência (homens, $x = 0$).

Efeito do sexo (β_1): como $x = 1$ para mulheres, o sinal de β_1 informa a direção da diferença mulher-homem em p .

- **Primeira ativação:** $\hat{\beta}_1 \approx -0,549$ com IC95% (-0,984; -0,114). Há evidência de efeito negativo para mulheres.
- **Última ativação:** $\hat{\beta}_1 \approx 1,294$ com IC95% (0,171; 2,418). Há evidência de efeito positivo para mulheres.
- **Ativação aleatória:** $\hat{\beta}_1 \approx -0,544$ com IC95% (-1,011; -0,076). Novamente observa-se efeito negativo.

Em síntese, o intercepto é indistinto de zero, enquanto o coeficiente β_1 é estatisticamente diferente de zero nos três cenários, mas com sinais que dependem do mecanismo de ativação: negativo nas ativações Primeira e Aleatória (menor p entre mulheres) e positivo na Última (maior p entre mulheres). Tal divergência indica sensibilidade das conclusões sobre o efeito do sexo à escolha do mecanismo de ativação, recomendando a avaliação conjunta com critérios de ajuste e diagnósticos de resíduos.

Tabela 26 – Proporção de cura: $\pi(x) = 2p(x)^3 - 3p(x)^2 + 1$

Ativação	Masculino	Feminino
Primeira	0,5155	0,7093
Última	0,3127	0,0532
Aleatória	0,5217	0,7128

A Tabela 26 reporta a proporção de cura $\pi(x) = 2p(x)^3 - 3p(x)^2 + 1$ por sexo e mecanismo de ativação, em que $p(x) = \text{logit}^{-1}(\beta_0 + \beta_1 x)$ e $x = 1$ para mulheres ($x = 0$ para homens). Como $\pi'(p) = 6p(p-1) \leq 0$ para $p \in [0, 1]$, maiores valores de p implicam menores proporções não suscetível. Observa-se que, nas ativações Primeira e Aleatória, as mulheres apresentam π substancialmente maior do que os homens (diferenças de $\approx 0,194$ e $\approx 0,191$, respectivamente), sugerindo p menor para $x = 1$ e, portanto, $\beta_1 < 0$ nesses cenários. Em contraste, na ativação Última obtém-se π muito menor para mulheres (0,0532) do que para homens (0,3127), diferença de $\approx 0,260$, compatível com p maior entre mulheres e $\beta_1 > 0$. Em síntese, a fração de curados é sensível ao mecanismo de ativação: Primeira e Aleatória indicam maior π para mulheres, enquanto Última aponta o oposto, reforçando que conclusões sobre o efeito do sexo dependem da suposição adotada para a ativação latente.

Tabela 27 – Métricas dos modelos ajustados

Ativação	AIC	BIC	LL
Primeira	454,3863	467,6783	-223,1931
Última	466,7647	480,0567	-229,3823
Aleatória	455,3279	468,6199	-223,6639

A Tabela 27 apresenta os valores das métricas AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) e logaritmo da verossimilhança (LL) para os três modelos ajustados, cada um com um mecanismo distinto de ativação latente: primeira, última e aleatória.

Observa-se que o modelo com mecanismo de ativação pela primeira causa apresentou o menor valor de AIC (454,3863) e BIC (467,6783), além de um logaritmo da verossimilhança ($-223,1931$) superior (menos negativo) em relação aos demais. Esses resultados indicam que, sob os critérios de ajuste e penalização por complexidade, o modelo com ativação pela primeira causa proporciona melhor equilíbrio entre qualidade de ajuste e parcimônia, sendo, portanto, o mais adequado entre os considerados.

O modelo com ativação aleatória apresentou desempenho intermediário, com métricas ligeiramente superiores às do modelo com ativação pela primeira causa, mas ainda inferiores às do modelo com ativação pela última causa. Já o modelo com ativação pela última causa foi o que apresentou os piores desempenhos em todas as métricas, com os maiores valores de AIC (466,7647) e BIC (480,0567), além do menor valor de log-verossimilhança ($-229,3823$).

Esses resultados quantitativos, aliados às análises inferenciais realizadas anteriormente, reforçam a importância da escolha apropriada do mecanismo de ativação na modelagem esta-

tística com fração de cura. O modelo com ativação pela primeira causa não apenas apresentou melhor ajuste segundo os critérios formais, como também resultou em estimativas estáveis e interpretações coerentes com a estrutura latente assumida, o que evidencia a robustez da abordagem adotada neste trabalho.

A Figura 26 apresenta o estimador de Kaplan-Meier (linhas tracejadas) juntamente com as curvas de sobrevivência ajustadas pelos modelos com fração de cura sob os três mecanismos de ativação considerados: primeira, última e aleatória. As curvas estão estratificadas por sexo (masculino e feminino), permitindo observar o comportamento diferencial da função de sobrevivência entre os grupos.

De maneira geral, observa-se que o modelo com ativação pela primeira causa apresenta curvas ajustadas que se alinham bem ao estimador de Kaplan-Meier em ambos os grupos, especialmente no início e no meio do tempo de seguimento. Essa boa aderência visual corrobora os resultados obtidos nas Tabelas 25 e 27, que indicaram esse modelo como o de melhor ajuste segundo os critérios de máxima verossimilhança, AIC e BIC.

O modelo com ativação pela última causa, por sua vez, apresenta curvas mais distantes do traçado empírico, sobretudo no grupo masculino, sugerindo menor aderência aos dados observados. Isso reforça os resultados da Tabela 27, na qual esse modelo apresentou o pior desempenho entre os três avaliados.

Por fim, o modelo com ativação aleatória demonstra desempenho intermediário, com curvas que se aproximam razoavelmente do estimador de Kaplan-Meier, especialmente para o grupo feminino. Esse comportamento é consistente com as métricas do modelo e com o fato de β_1 ter sido significativo, indicando sensibilidade à covariável sexo.

Na Figura 27 tem-se, para cada mecanismo de ativação, a avaliação gráfica da adequação global do modelo com covariável. Observa-se que, nas ativações Primeira e Aleatória, os pontos se alinham de forma consistente à reta identidade ao longo de quase todo o domínio, com discreta dispersão apenas na cauda. Esse comportamento é compatível com bom ajuste global e coaduna-se com os achados anteriores: $\alpha > 1$ (risco latente crescente), $\beta_1 < 0$ e maiores proporções de “não suscetíveis” entre mulheres, sem indícios de falta de ajuste sistemática.

Por outro lado, na ativação Última nota-se desvio perceptível em valores mais altos dos resíduos, com padrão que se afasta da reta identidade de modo persistente. Esse traço sugere possível má especificação sob esse mecanismo e está alinhado ao quadro já observado: $\alpha < 1$, inversão do sinal de β_1 (efeito positivo para mulheres) e forte redução da fração de “não suscetíveis” no sexo feminino. Em síntese, os diagnósticos visuais corroboram a adequação relativa dos ajustes sob Primeira e Aleatória e sinalizam fragilidade do esquema Última para os dados analisados.

(ii) Abordagem Bayesiana

À semelhança do cenário sem covariável, o modelo BWeibcr com mecanismo de primeira

ativação apresentou o melhor ajuste aos dados de melanoma. Em razão desse resultado, apresenta-se, em seguida, o ajuste do referido modelo no arcabouço da inferência Bayesiana.

Foram executadas duas cadeias independentes com 25.000 iterações; descartaram-se as 5.000 iniciais (burn-in) e aplicou-se espaçamento (thinning) de 10, resultando em 4.000 amostras para a inferência *a posteriori*. A convergência foi avaliada conforme Cowles e Carlin (1996) e confirmada pelo diagnóstico de Gelman–Rubin ($\hat{R} \leq 1,01$), além da inspeção visual dos traços (GELMAN; RUBIN, 1992). Os resumos *a posteriori*, com intervalos de credibilidade de 95%, encontram-se na Tabela 28.

Tabela 28 – Resumo *a posteriori* dos parâmetros do modelo BWeibcr sob o esquema de primeira ativação

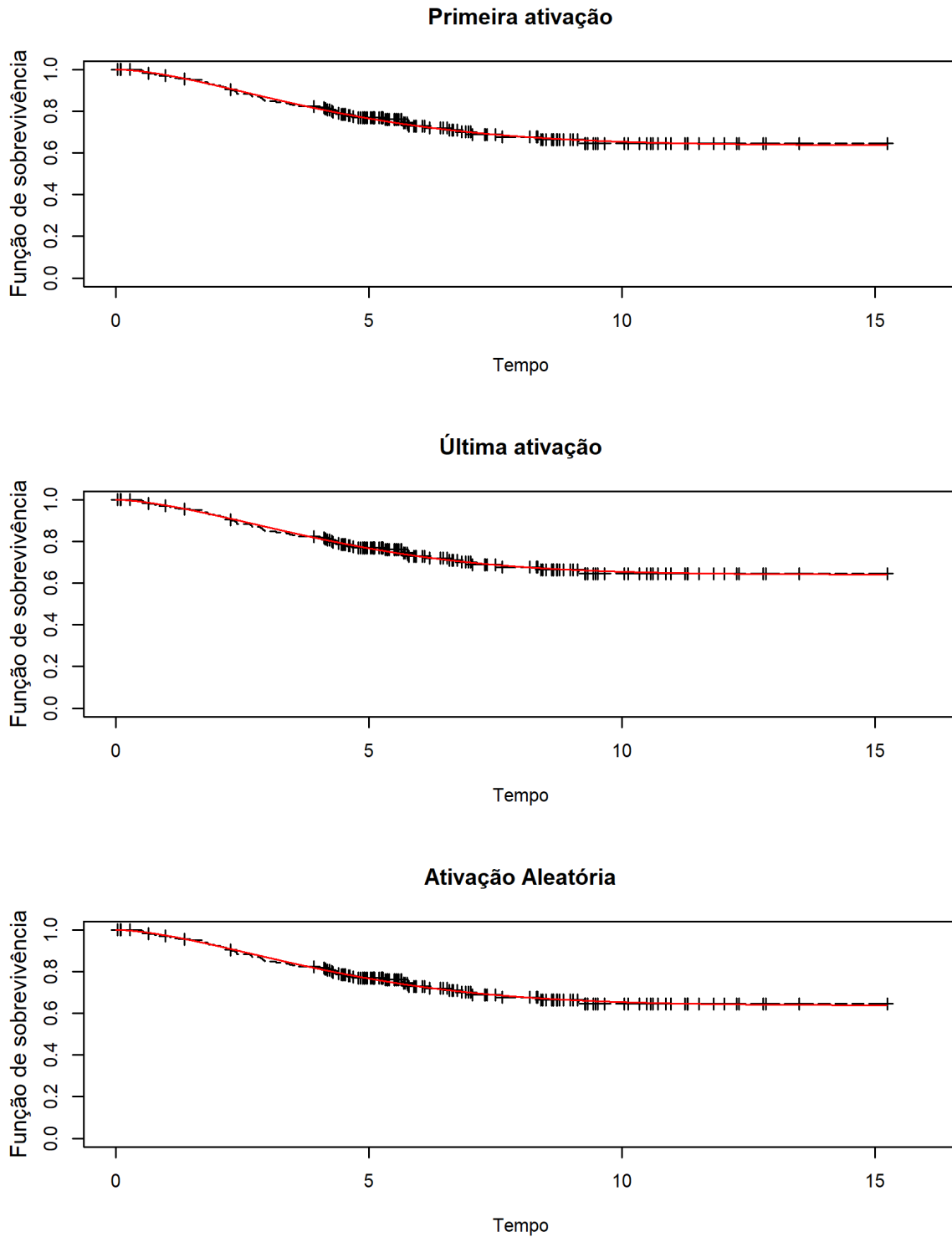
Parâmetro	Média	DP	IC (95%)
α	1,545	0,190	(1,193, 1,933)
σ	5,459	0,281	(5,019, 5,965)
β_0	0,034	0,199	(-0,343, 0,447)
β_1	-0,572	0,249	(-1,061, -0,080)

Computaram-se as medidas de divergência ψ ; seus valores constam da Tabela 30 (Apêndice C) e sua representação gráfica aparece na Figura 28. Nenhuma observação foi identificada como influente por nenhuma das três métricas.

As Figuras 29 e 30 apresentam, respectivamente, os traços e as funções de autocorrelação dos parâmetros. Os traços exibem mistura rápida, sem tendências ou padrões sistemáticos, indicando exploração adequada do espaço paramétrico. As autocorrelações decaem para valores próximos de zero após poucas defasagens, evidenciando baixa dependência serial. Em conjunto, os diagnósticos são compatíveis com a convergência das cadeias e sustentam a confiabilidade das inferências *a posteriori*.

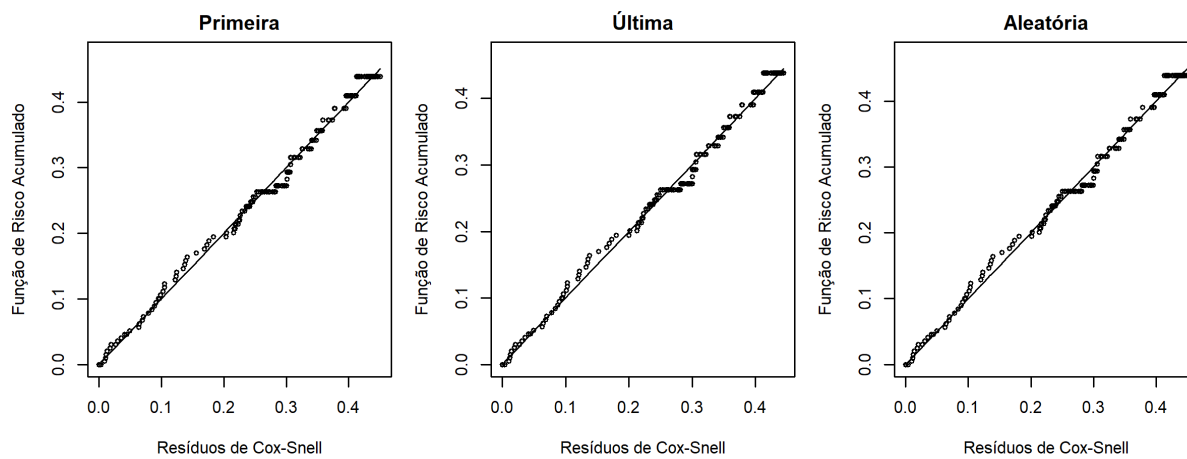
Por fim, a Figura 31 exibe o boxplot *a posteriori* da fração de curados (π), sintetizando sua tendência central e dispersão.

Figura 14 – Estimador de Kaplan–Meier e curvas de sobrevivência dos modelos sem covariável ajustados por inferência clássica sob diferentes esquemas de ativação

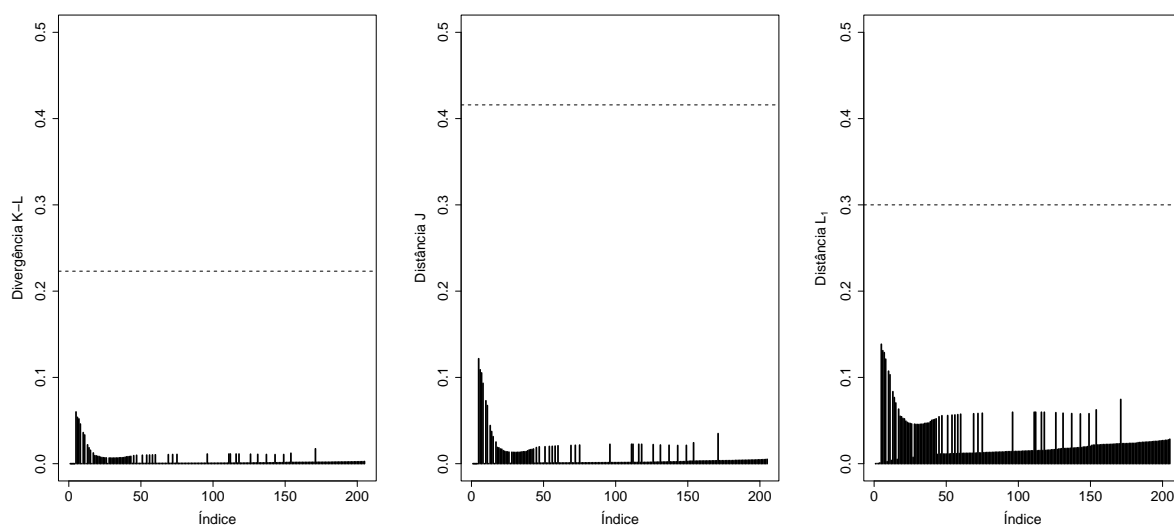


Fonte: Elaborada pelo autor.

Figura 15 – Resíduos do tipo Cox-Snell para os modelos ajustados sem covariável

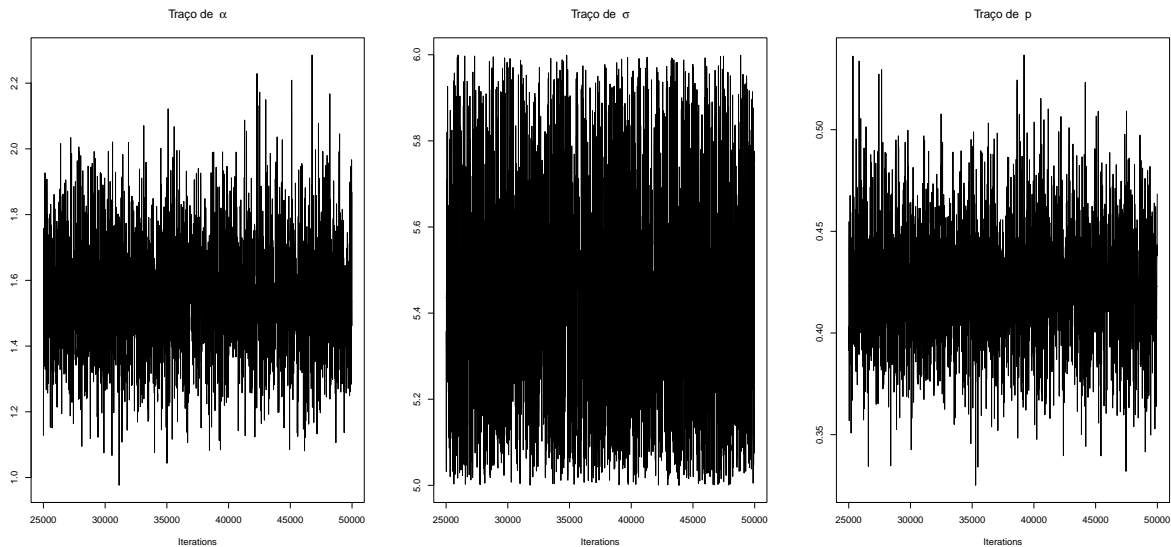


Fonte: Elaborada pelo autor.

Figura 16 – Divergência ψ para o conjunto de dados de melanoma

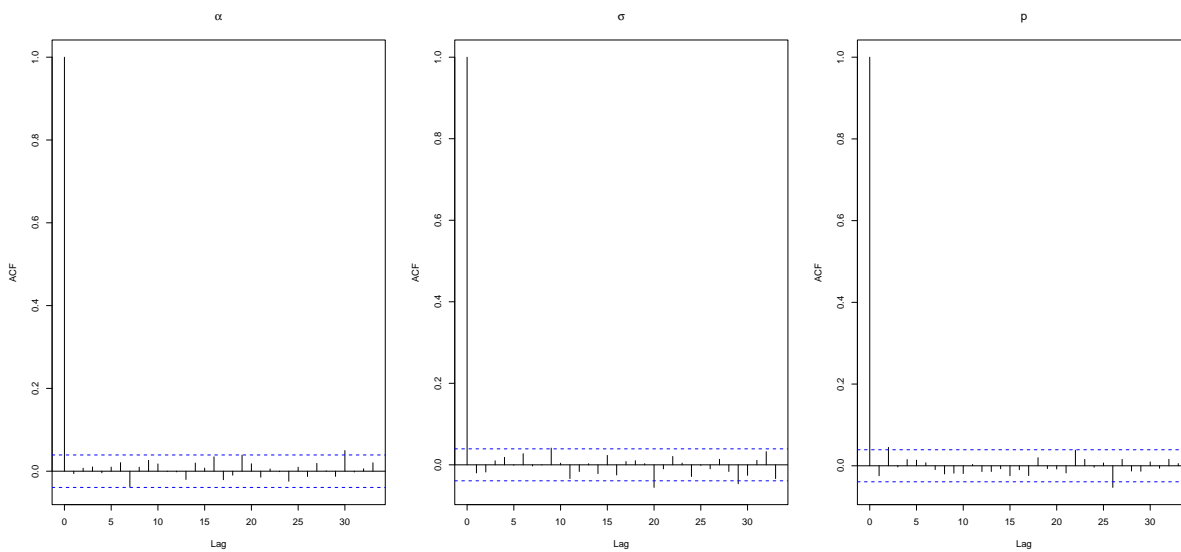
Fonte: Elaborada pelo autor.

Figura 17 – Gráfico de traços

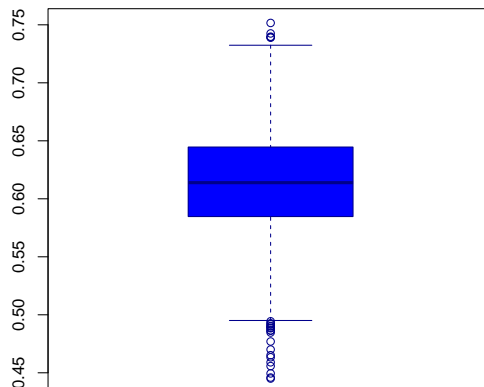


Fonte: Elaborada pelo autor.

Figura 18 – Gráfico ACF

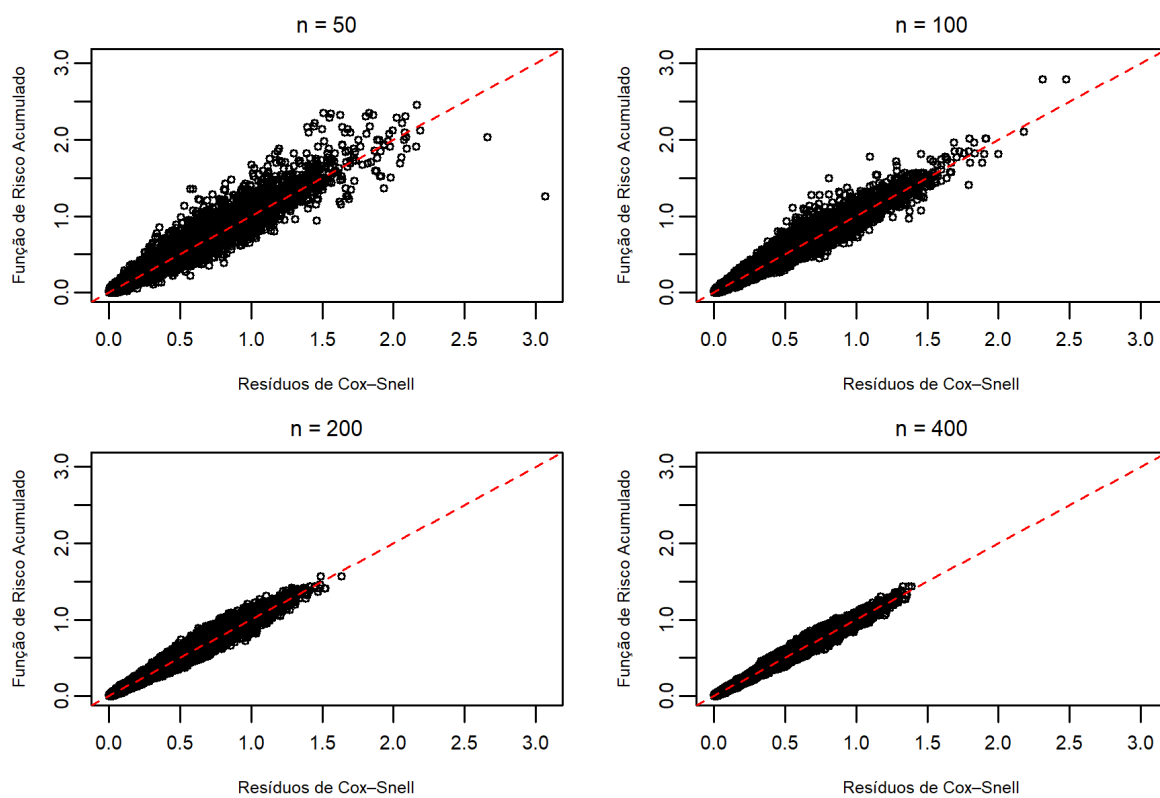


Fonte: Elaborada pelo autor.

Figura 19 – Box plot para a distribuição *a posteriori* da proporções de curados

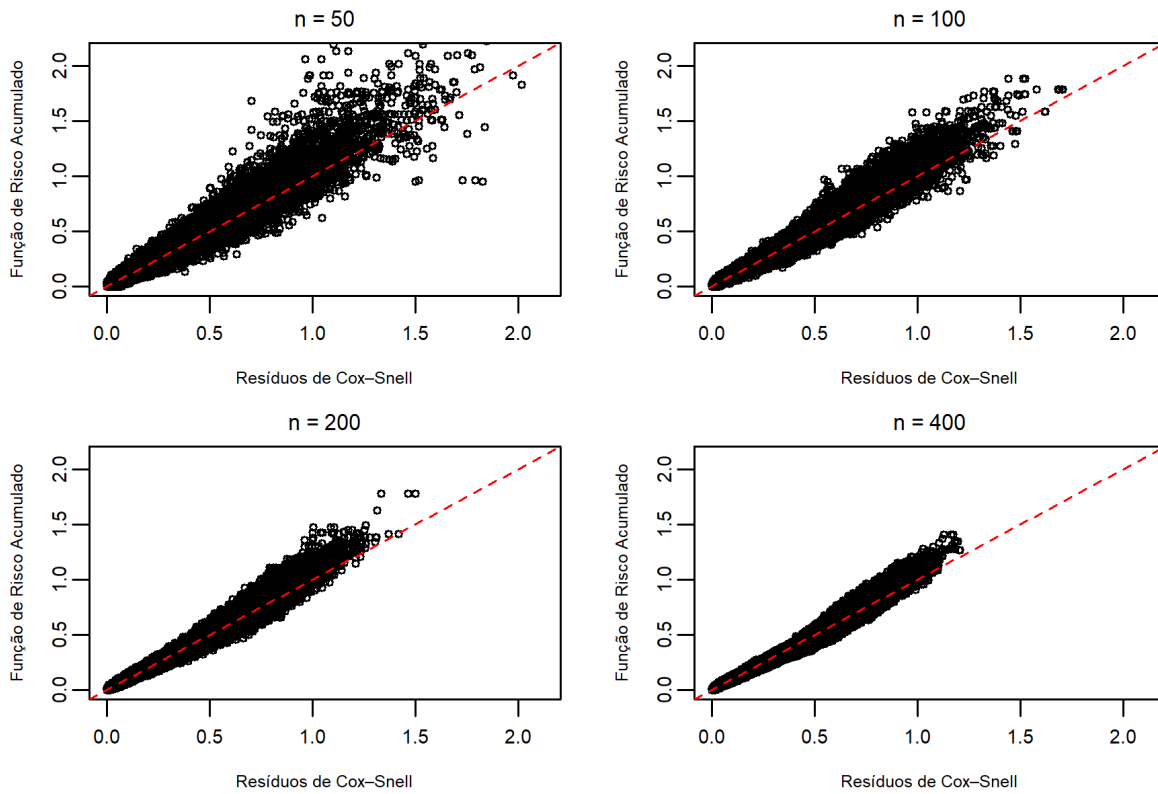
Fonte: Elaborada pelo autor.

Figura 20 – Resíduos de Cox–Snell do modelo bem especificado com covariável

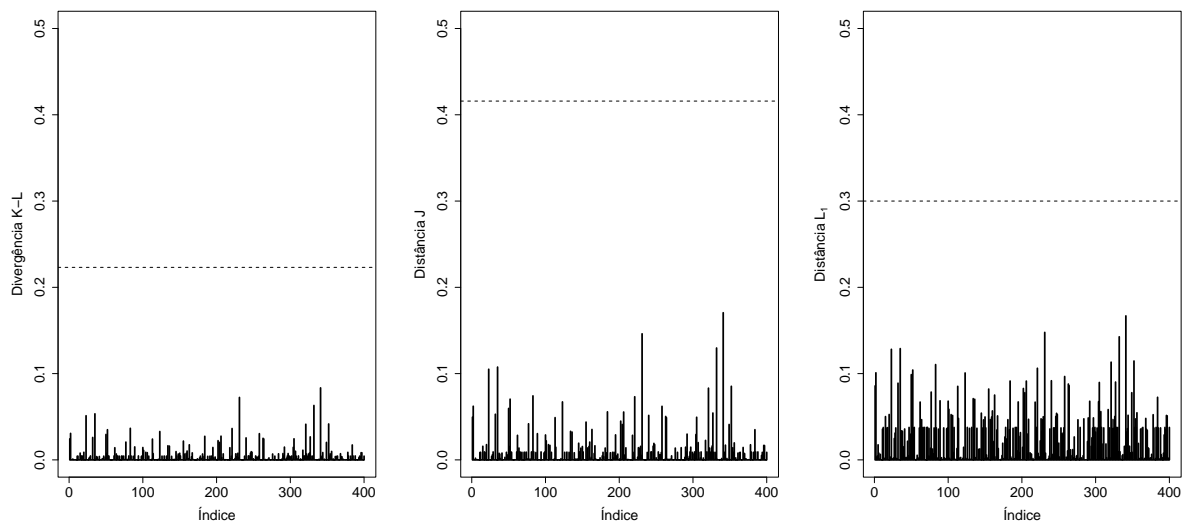


Fonte: Elaborada pelo autor.

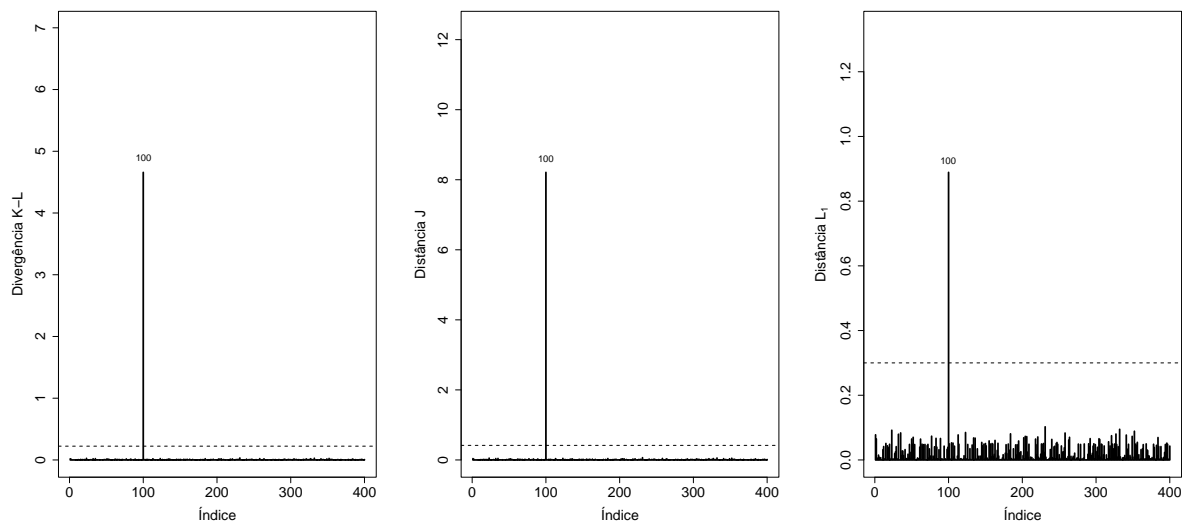
Figura 21 – Resíduos de Cox–Snell do modelo mal especificado com covariável



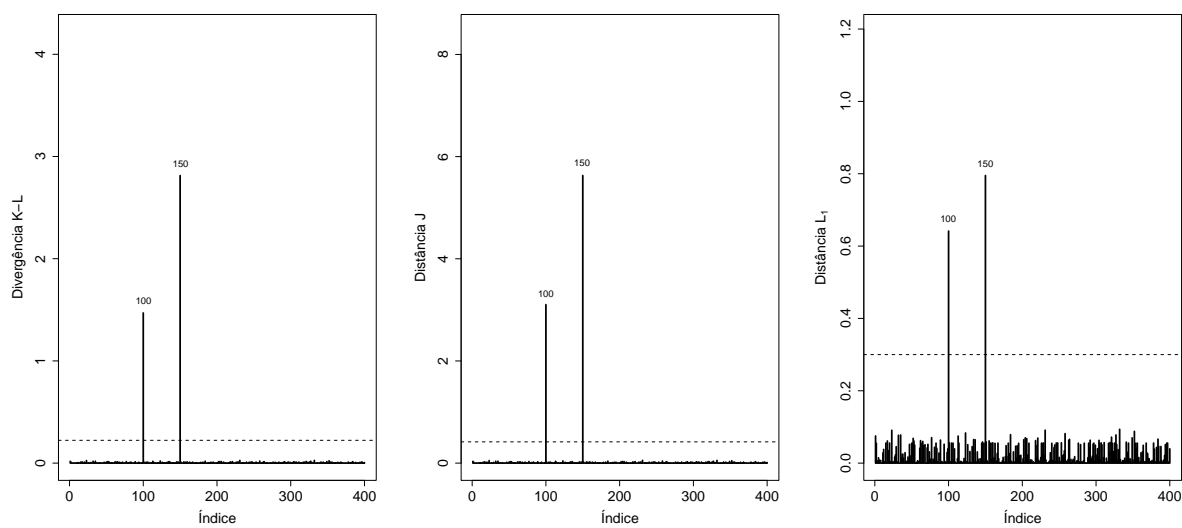
Fonte: Elaborada pelo autor.

Figura 22 – Divergência ψ para o conjunto de dados (a) no caso com covariável

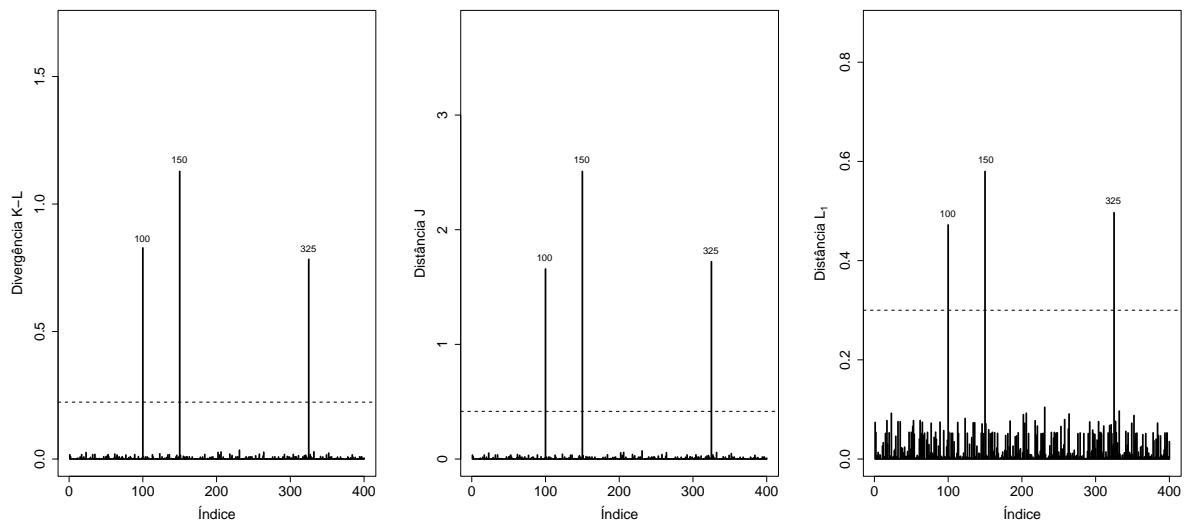
Fonte: Elaborada pelo autor.

Figura 23 – Divergência ψ para o conjunto de dados (b) no caso com covariável

Fonte: Elaborada pelo autor.

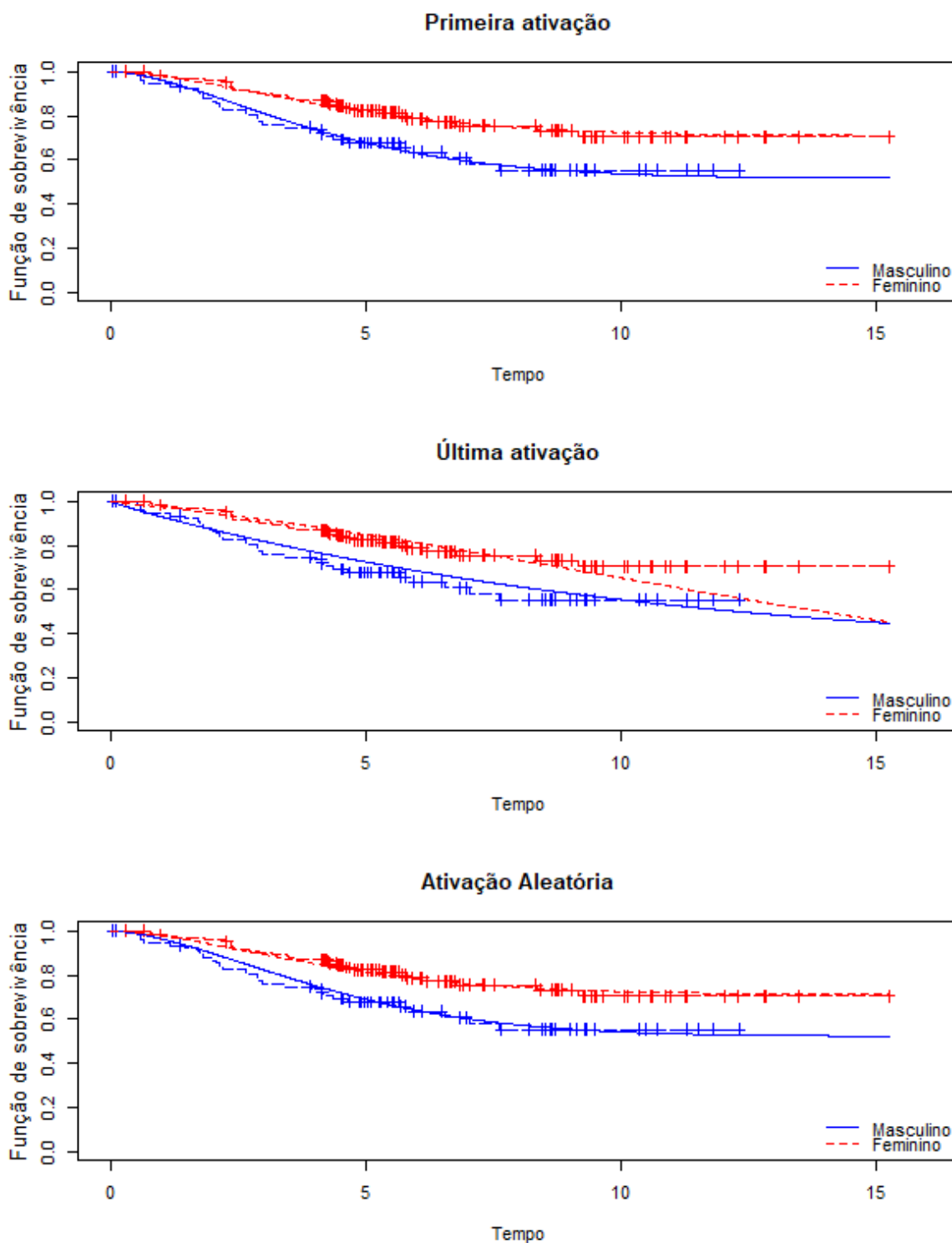
Figura 24 – Divergência ψ para o conjunto de dados (e) no caso com covariável

Fonte: Elaborada pelo autor.

Figura 25 – Divergência ψ para o conjunto de dados (h) no caso com covariável

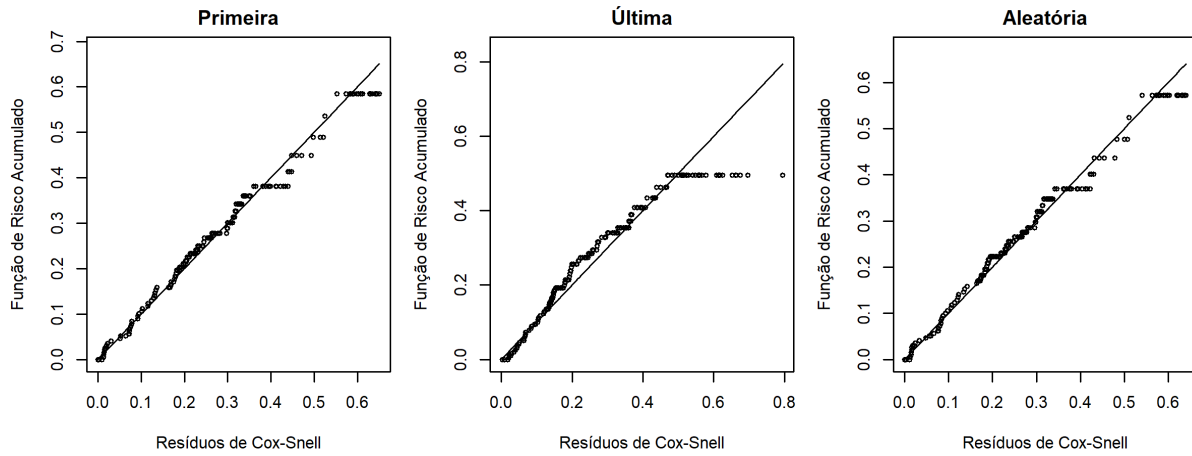
Fonte: Elaborada pelo autor.

Figura 26 – Estimador de Kaplan–Meier e curvas de sobrevivência dos modelos com covariável ajustados por inferência clássica sob diferentes esquemas de ativação

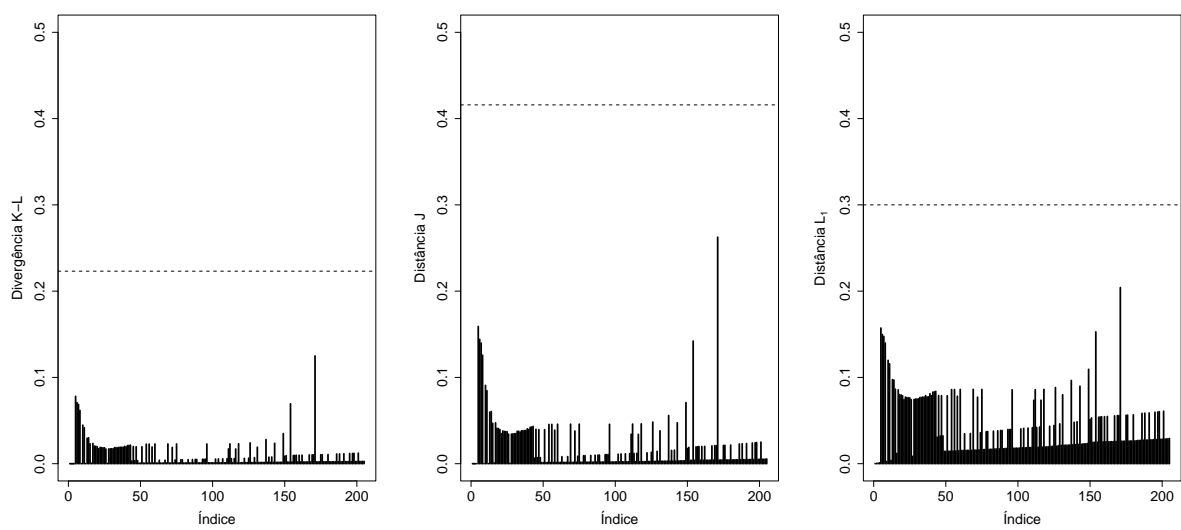


Fonte: Elaborada pelo autor.

Figura 27 – Resíduos do tipo Cox-Snell para os modelos ajustados com covariável

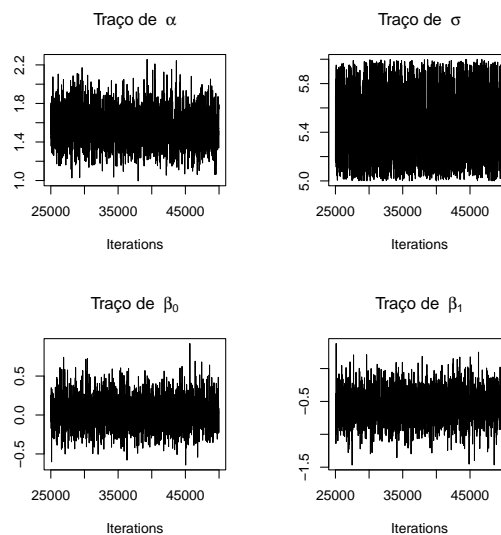


Fonte: Elaborada pelo autor.

Figura 28 – Divergência ψ para o conjunto de dados de melanoma com covariável

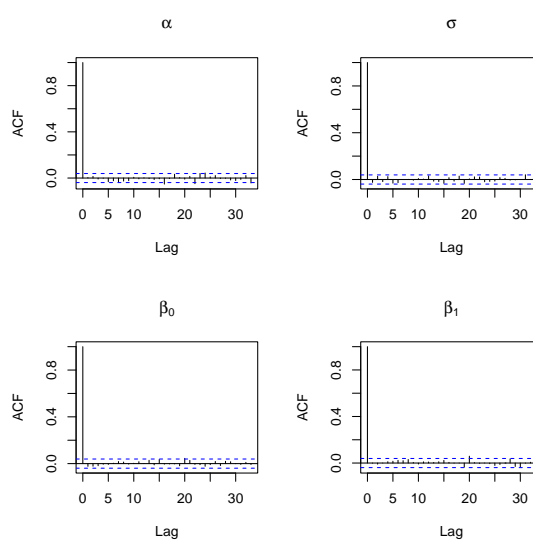
Fonte: Elaborada pelo autor.

Figura 29 – Gráfico de traços para o modelo com covariável

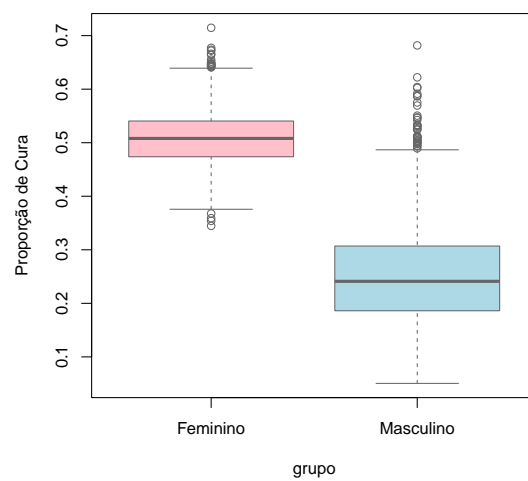


Fonte: Elaborada pelo autor.

Figura 30 – Gráfico ACF para o modelo com covariável



Fonte: Elaborada pelo autor.

Figura 31 – Box plot para a distribuição *a posteriori* da proporções de curados do modelo com covariável

Fonte: Elaborada pelo autor.

CONSIDERAÇÕES FINAIS

Este trabalho propôs e investigou um novo modelo de sobrevivência com fração de cura, denominado modelo Bilal Weibull com fração de cura (BWeibcr). A formulação combina a flexibilidade da distribuição Weibull para os tempos de falha com a distribuição discreta Bilal para o número de causas latentes M , permitindo captar heterogeneidades populacionais e diferentes mecanismos de ativação. A avaliação sob as perspectivas clássica e Bayesiana, por meio de estudos de simulação e de uma aplicação a dados reais, indicou que a abordagem é viável e informativa.

Nos experimentos de simulação, os estimadores de máxima verossimilhança apresentaram bom desempenho em termos de viés e precisão, sobretudo para tamanhos amostrais moderados e grandes. A análise por resíduos de Cox–Snell confirmou a adequação global dos ajustes, reforçando o papel dos diagnósticos no contexto de modelos com fração de cura. As comparações entre mecanismos de ativação mostraram desempenhos próximos em critérios usuais (AIC, BIC, DIC, EAIC, EBIC e LPML), com a ativação pela primeira ocorrência se destacando em alguns cenários.

No enfoque Bayesiano, a inferência via MCMC proporcionou uma quantificação abrangente da incerteza, oferecendo distribuições a posteriori informativas para os parâmetros. Critérios Bayesianos de comparação e medidas de influência, como a divergência ψ , mostraram-se úteis para avaliar a adequação e robustez, e as estimativas mantiveram-se estáveis em diferentes configurações de simulação.

A aplicação a dados de melanoma evidenciou a capacidade de ajuste e a interpretabilidade do modelo BWeibcr em contextos biomédicos, nos quais a fração de cura e a heterogeneidade latente desempenham papel central. A incorporação de covariáveis mostrou-se pertinente para identificar fatores prognósticos associados à proporção de curados, ampliando a utilidade prática do modelo.

Algumas limitações merecem destaque. A combinação de mecanismos de ativação

com fração de cura eleva o custo computacional, sobretudo em bases extensas ou com grande número de covariáveis. Além disso, embora a distribuição de Bilal tenha se mostrado adequada, alternativas discretas merecem investigação sistemática. Por fim, em amostras pequenas, observa-se perda de precisão inferencial, como esperado.

Como desdobramentos futuros, sugerem-se: (i) explorar outras distribuições para M e para os tempos de falha (p. ex., lognormal, log-logística, Gompertz), incluindo versões semiparamétricas; (ii) incorporar covariáveis dinâmicas e estruturas hierárquicas; (iii) desenvolver e comparar diagnósticos residuais específicos para distintos mecanismos de ativação; (iv) investigar estratégias computacionais mais eficientes (por exemplo: amostradores Hamiltonianos) e estudar o comportamento do modelo sob má especificação.

Em síntese, o modelo BWeibcr acrescenta uma alternativa flexível e interpretável ao repertório de modelos de sobrevivência com fração de cura, aliando rigor teórico, diagnóstico cuidadoso e evidência empírica. Os resultados apoiam seu uso em diversas aplicações, incentivando novas pesquisas sobre ativação latente, seleção de modelos e eficiência computacional.

REFERÊNCIAS

- AALEN, O. Nonparametric inference for a family of counting processes. **The Annals of Statistics**, JSTOR, p. 701–726, 1978. Citado na página 33.
- ABD-ELRAHMAN, A. M. Utilizing ordered statistics in lifetime distributions production: a new lifetime distribution and applications. **Journal of probability and statistical science**, v. 11, n. 2, p. 153–164, 2013. Citado nas páginas 23 e 45.
- ALDRICH, J. Ra fisher and the making of maximum likelihood 1912-1922. **Statistical science**, Institute of Mathematical Statistics, v. 12, n. 3, p. 162–176, 1997. Citado na página 67.
- ALTUN, E.; EL-MORSHEDY, M.; ELIWA, M. A study on discrete bilal distribution with properties and applications on integervalued autoregressive process. **REVSTAT-Statistical Journal**, v. 20, n. 4, p. 501–528, 2022. Citado nas páginas 23, 40, 46, 48, 50, 51, 61 e 64.
- ANDERSEN, P. K.; BORGAN, O.; GILL, R. D.; KEIDING, N. **Statistical models based on counting processes**. [S.l.]: Springer Science & Business Media, 2012. Citado na página 96.
- ASHTON, R.; WILKINSON, A. N. Melanoma crash course. **Canadian Family Physician**, The College of Family Physicians of Canada, v. 70, n. 4, p. 254–257, 2024. Citado na página 94.
- 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 36 e 37.
- BICKEL, P. J.; DOKSUM, K. A. **Mathematical statistics: basic ideas and selected topics, volumes I-II package**. [S.l.]: Chapman and Hall/CRC, 2015. Citado na página 70.
- BOAG, J. W. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, v. 11, n. 1, p. 15–53, 1949. Citado na página 37.
- BOLFARINE, H.; SANDOVAL, M. C. **Introdução à inferência estatística**. [S.l.]: SBM, 2001. v. 2. Citado na página 69.
- BROOKS, S.; SMITH, J.; VEHTARI, A.; PLUMMER, M.; STONE, M.; ROBERT, C. P.; TITTERINGTON, D.; NELDER, J.; ATKINSON, A.; DAWID, A. *et al.* Discussion on the paper by Spiegelhalter, Best, Carlin and van der Linde. **Journal of the Royal Statistical Society. Series B: Statistical Methodology**, Oxford University Press, v. 64, n. 4, p. 616–639, 2002. Citado na página 81.
- 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 39.
- CARLIN, B. P.; LOUIS, T. A. *et al.* **Bayes and empirical Bayes methods for data analysis**. [S.l.]: Chapman & Hall/CRC, 2000. Citado na página 81.

- CARREGARI, R. C. **Um novo modelo de sobrevivência Bell–Inversa Gaussiana com fração de cura**. Dissertação (Dissertação de Mestrado) — Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos, SP, 2021. Disponível em: <<https://repositorio.usp.br/item/003067159>>. Disponível em: <<https://repositorio.usp.br/item/003067159>>. Citado na página 39.
- CASELLA, G.; BERGER, R. **Statistical inference**. [S.l.]: Chapman and Hall/CRC, 2024. Citado nas páginas 69 e 70.
- CASTRO, M. d.; CANCHO, V. G.; RODRIGUES, J. A bayesian long-term survival model parametrized in the cured fraction. **Biometrical Journal: Journal of Mathematical Methods in Biosciences**, Wiley Online Library, v. 51, n. 3, p. 443–455, 2009. Citado nas páginas 23, 38 e 39.
- CASTRO, M. D.; CANCHO, V. G.; RODRIGUES, J. A hands-on approach for fitting long-term survival models under the gamlss framework. **Computer methods and programs in biomedicine**, Elsevier, v. 97, n. 2, p. 168–177, 2010. Citado na página 38.
- CHEN, M.-H.; IBRAHIM, J. G.; SINHA, D. A new bayesian model for survival data with a surviving fraction. **Journal of the American Statistical Association**, Taylor & Francis, v. 94, n. 447, p. 909–919, 1999. Citado na página 38.
- _____. Bayesian inference for multivariate survival data with a cure fraction. **Journal of Multivariate Analysis**, Elsevier, v. 80, n. 1, p. 101–126, 2002. Citado na página 38.
- CHEN, M.-H.; SHAO, Q.-M.; IBRAHIM, J. G. **Monte Carlo methods in Bayesian computation**. [S.l.]: Springer Science & Business Media, 2012. Citado na página 81.
- COLES, S.; BAWA, J.; TRENNER, L.; DORAZIO, P. **An introduction to statistical modeling of extreme values**. [S.l.]: Springer, 2001. v. 208. Citado na página 53.
- COLOSIMO, E. A.; GIOLO, S. R. **Análise de sobrevivência aplicada**. [S.l.]: Editora Blucher, 2006. Citado nas páginas 21, 22, 25, 26, 28, 31, 34, 70 e 71.
- COOK, R. D.; WEISBERG, S. **Residuals and Influence in Regression**. Boca Raton, FL: Chapman & Hall/CRC, 1982. Citado na página 82.
- COONER, F.; BANERJEE, S.; CARLIN, B. P.; SINHA, D. Flexible cure rate modeling under latent activation schemes. **Journal of the American Statistical Association**, Taylor & Francis, v. 102, n. 478, p. 560–572, 2007. Citado nas páginas 39, 62 e 63.
- COONER, F.; BANERJEE, S.; MCBEAN, A. M. Modelling geographically referenced survival data with a cure fraction. **Statistical methods in medical research**, Sage Publications Sage CA: Thousand Oaks, CA, v. 15, n. 4, p. 307–324, 2006. Citado nas páginas 23, 39 e 62.
- 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 89, 99, 103 e 111.
- 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 22.
- COX, D. R.; HINKLEY, D. V. **Theoretical statistics**. [S.l.]: CRC Press, 1979. Citado na página 69.

COX, D. R.; SNELL, E. J. A general definition of residuals. **Journal of the Royal Statistical Society: Series B (Methodological)**, Wiley Online Library, v. 30, n. 2, p. 248–265, 1968. Citado na página 70.

DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. [S.l.]: Cambridge university press, 1997. Citado na página 85.

DEMPSTER, A. P.; LAIRD, N. M.; RUBIN, D. B. Maximum likelihood from incomplete data via the em algorithm. **Journal of the royal statistical society: series B (methodological)**, Wiley Online Library, v. 39, n. 1, p. 1–22, 1977. Citado na página 68.

DRZEWIECKI, K. **The Melanoma Survival Data**. 2004. <<https://cran.r-project.org/package=timereg>>. Disponível no pacote timereg para R. Citado na página 96.

DRZEWIECKI, K.; LADEFOGED, C.; CHRISTENSEN, H. Biopsy and prognosis for cutaneous malignant melanomas in clinical stage i. **Scandinavian journal of plastic and reconstructive surgery**, Taylor & Francis, v. 14, n. 2, p. 141–144, 1980. Citado na página 96.

FAREWELL, V. T. The use of mixture models for the analysis of survival data with long-term survivors. **Biometrics**, JSTOR, p. 1041–1046, 1982. Citado na página 37.

_____. Mixture models in survival analysis: Are they worth the risk? **Canadian Journal of Statistics**, Wiley Online Library, v. 14, n. 3, p. 257–262, 1986. Citado na página 37.

FIDELIS, C.; ORTEGA, E.; CORDEIRO, G. **Residual Analysis for Poisson-Exponentiated Weibull Regression Models with Cure Fraction**. **Stats** 2024, 7, 492–508. 2024. Citado na página 77.

FIDELIS, C. R. **Resíduos em modelos de sobrevivência com fração de cura**. Dissertação (Dissertação de Mestrado) — Universidade Federal do Rio Grande do Norte, Centro de Ciências Exatas e da Terra, Natal, Brasil, 2019. Disponível em: <<https://repositorio.ufrn.br/items/0cd735c0-9d10-427c-b3c6-e19cfb2f4916>>. Citado nas páginas 71, 75 e 77.

FISHER, R. Statistical methods and scientific induction. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, Oxford University Press, v. 17, n. 1, p. 69–78, 1955. Citado na página 21.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character**, The Royal Society London, v. 222, n. 594-604, p. 309–368, 1922. Citado na página 67.

GALLARDO, D. I.; CASTRO, M. de; GÓMEZ, H. W. An alternative promotion time cure model with overdispersed number of competing causes: An application to melanoma data. **Mathematics**, MDPI, v. 9, n. 15, p. 1815, 2021. Citado na página 23.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian data analysis**. [S.l.]: Chapman and Hall/CRC, 1995. Citado na página 85.

GELMAN, A.; RUBIN, D. B. Inference from iterative simulation using multiple sequences. **Statistical science**, Institute of Mathematical Statistics, v. 7, n. 4, p. 457–472, 1992. Citado nas páginas 89, 99, 103 e 111.

- GHITANY, M.; MALLER, R. A.; ZHOU, S. Exponential mixture models with long-term survivors and covariates. **Journal of multivariate Analysis**, Elsevier, v. 49, n. 2, p. 218–241, 1994. Citado na página 37.
- IBRAHIM, J. G.; CHEN, M.-H.; SINHA, D.; IBRAHIM, J.; CHEN, M. **Bayesian survival analysis**. [S.l.]: Springer, 2001. v. 2. Citado nas páginas 23, 35, 36, 38 e 39.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous univariate distributions, volume 2**. [S.l.]: John wiley & sons, 1995. v. 2. Citado nas páginas 52 e 53.
- KAPLAN, E. L.; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American statistical association**, Taylor & Francis, v. 53, n. 282, p. 457–481, 1958. Citado na página 34.
- KAYE, D. H.; FREEDMAN, D. A. Reference guide on statistics. **Reference manual on scientific evidence**, p. 211–302, 2011. Citado na página 21.
- KIM, S.; CHEN, M.-H.; DEY, D. K. A new threshold regression model for survival data with a cure fraction. **Lifetime data analysis**, Springer, v. 17, n. 1, p. 101–122, 2011. Citado na página 39.
- KLEINBAUM, D. G.; KLEIN, M. **Survival analysis a self-learning text**. [S.l.]: Springer, 1996. Citado nas páginas 21, 25, 26 e 34.
- LAWLESS, J. F. **Statistical models and methods for lifetime data**. [S.l.]: John Wiley & Sons, 2011. Citado nas páginas 21, 22, 26 e 53.
- LEHMANN, E. L.; CASELLA, G. **Theory of point estimation**. [S.l.]: Springer, 1998. Citado nas páginas 67 e 69.
- LOUZADA, F.; CANCHO, V. G.; BARRIGA, G. D. The poisson-exponential regression model under different latent activation schemes. **Computational & Applied Mathematics**, SciELO Brasil, v. 31, p. 617–632, 2012. Citado na página 39.
- LOUZADA, F.; CANCHO, V. G.; YIQI, B. The log-weibull-negative-binomial regression model under latent failure causes and presence of randomized activation schemes. **Statistics**, Taylor & Francis, v. 49, n. 4, p. 930–949, 2015. Citado na página 39.
- LOUZADA-NETO, F.; MAZUCHELI, J.; ACHCAR, J. A. **Uma Introdução à análise de Sobrevivência e Confiabilidade**. [S.l.]: Sociedad Chilena de Estadística, 2001. Citado na página 34.
- LYU, M. R. *et al.* **Handbook of software reliability engineering**. [S.l.]: IEEE computer society press Los Alamitos, 1996. v. 222. Citado na página 53.
- MEEKER, W. Q.; ESCOBAR, L. A.; PASCUAL, F. G. **Statistical methods for reliability data**. [S.l.]: John Wiley & Sons, 2021. Citado na página 53.
- NASH, J. C. **Compact numerical methods for computers: linear algebra and function minimisation**. [S.l.]: Routledge, 2018. Citado na página 70.
- NELSON, W. Hazard plotting for incomplete failure data. **Journal of Quality Technology**, Taylor & Francis, v. 1, n. 1, p. 27–52, 1969. Citado na página 33.

- NEYMAN, J.; PEARSON, E. S. Ix. on the problem of the most efficient tests of statistical hypotheses. **Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character**, The Royal Society London, v. 231, n. 694-706, p. 289–337, 1933. Citado na página 67.
- ORTEGA, E. M.; CANCHO, V. G.; PAULA, G. A. Generalized log-gamma regression models with cure fraction. **Lifetime Data Analysis**, Springer, v. 15, n. 1, p. 79–106, 2009. Citado na página 38.
- PEARSON, K. The grammar of science. **Nature**, Nature Publishing Group UK London, v. 46, n. 1185, p. 247–247, 1892. Citado na página 21.
- PENG, F.; DEY, D. K. Bayesian analysis of outlier problems using divergence measures. **Canadian Journal of Statistics**, Wiley Online Library, v. 23, n. 2, p. 199–213, 1995. Citado nas páginas 82 e 83.
- PENG, Y.; TAYLOR, J. M. Residual-based model diagnosis methods for mixture cure models. **Biometrics**, Wiley Online Library, v. 73, n. 2, p. 495–505, 2017. Citado nas páginas 71 e 75.
- PERDONÁ, G. C.; LOUZADA-NETO, F. A general hazard model for lifetime data in the presence of cure rate. **Journal of Applied Statistics**, Taylor & Francis, v. 38, n. 7, p. 1395–1405, 2011. Citado na página 37.
- PLUMMER, M. *et al.* Jags: A program for analysis of bayesian graphical models using gibbs sampling. In: VIENNA, AUSTRIA. **Proceedings of the 3rd international workshop on distributed statistical computing**. [S.l.], 2003. v. 124, n. 125.10, p. 1–10. Citado nas páginas 80 e 89.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2024. Disponível em: <<https://www.R-project.org/>>. Citado nas páginas 80 e 89.
- 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 na página 41.
- SALSBURG, D. **Uma senhora toma chá...: como a estatística revolucionou a ciência no século XX**. [S.l.]: Editora Schwarcz-Companhia das Letras, 2024. Citado na página 67.
- SCOLAS, S.; LEGRAND, C.; OULHAJ, A.; GHOUGH, A. E. Diagnostic checks in mixture cure models with interval-censoring. **Statistical Methods in Medical Research**, SAGE Publications Sage UK: London, England, v. 27, n. 7, p. 2114–2131, 2018. Citado nas páginas 71 e 75.
- SOERJOMATARAM, I.; LOUWMAN, M. W.; RIBOT, J. G.; ROUKEMA, J. A.; COEBERGH, J. W. W. An overview of prognostic factors for long-term survivors of breast cancer. **Breast cancer research and treatment**, Springer, v. 107, p. 309–330, 2008. Citado na página 36.
- SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit. **Journal of the royal statistical society: Series b (statistical methodology)**, Wiley Online Library, v. 64, n. 4, p. 583–639, 2002. Citado na página 81.
- SPOSTO, R. Cure model analysis in cancer: an application to data from the children's cancer group. **Statistics in medicine**, Wiley Online Library, v. 21, n. 2, p. 293–312, 2002. Citado na página 37.

- STIGLER, S. M. **The history of statistics: The measurement of uncertainty before 1900**. [S.l.]: Harvard University Press, 1990. Citado na página 67.
- SUZUKI, A. K.; BARRIGA, G. D.; LOUZADA, F.; CANCHO, V. G. A general long-term aging model with different underlying activation mechanisms: Modeling, bayesian estimation, and case influence diagnostics. **Communications in Statistics-Theory and Methods**, Taylor & Francis, v. 46, n. 6, p. 3080–3098, 2017. Citado nas páginas 23, 39 e 82.
- SUZUKI, A. K.; CANCHO, V. G.; LOUZADA, F. The poisson–inverse-gaussian regression model with cure rate: a bayesian approach and its case influence diagnostics. **Statistical Papers**, Springer, v. 57, n. 1, p. 133–159, 2016. Citado nas páginas 23, 39 e 82.
- TOURNOUD, M.; ECOCHARD, R. Application of the promotion time cure model with time-changing exposure to the study of hiv/aids and other infectious diseases. **Statistics in medicine**, Wiley Online Library, v. 26, n. 5, p. 1008–1021, 2007. Citado na página 38.
- 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 23 e 38.
- WEIBULL, W. A statistical distribution function of wide applicability. **Journal of applied mechanics**, 1951. Citado nas páginas 22, 40, 52 e 61.
- WEISS, R. An approach to bayesian sensitivity analysis. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, Oxford University Press, v. 58, n. 4, p. 739–750, 1996. Citado na página 83.
- WILEYTO, E. P.; LI, Y.; CHEN, J.; HEITJAN, D. F. Assessing the fit of parametric cure models. **Biostatistics**, Oxford University Press, v. 14, n. 2, p. 340–350, 2013. Citado na página 71.
- YAKOVLEV, A. Y.; ASSELAIN, B.; BARDOU, V.; FOURQUET, A.; HOANG, T.; ROCHEFE-DIERE, A.; TSODIKOV, A. A simple stochastic model of tumor recurrence and its application to data on premenopausal breast cancer. **Biometrie et analyse de donnees spatio-temporelles**, Rennes France, v. 12, p. 66–82, 1993. Citado na página 38.
- 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 na página 38.

APÊNDICE

Este apêndice apresenta a formulação da verossimilhança e das funções escore associadas aos três mecanismos de ativação.

A.1 Verossimilhança e funções *score* sob cada mecanismo de ativação

Seja $S(t) = \exp\{-(t/\sigma)^\alpha\}$ a função de sobrevivência e $f(t) = \frac{\alpha}{\sigma} \left(\frac{t}{\sigma}\right)^{\alpha-1} \exp\{-(t/\sigma)^\alpha\}$ a densidade da distribuição Weibull. Considere observações (t_i, δ_i) , $i = 1, \dots, n$, com $\delta_i = 1$ para falha observada e $\delta_i = 0$ para censura à direita. A função de log-verossimilhança geral é dada por

$$\ell(\vartheta) = \sum_{i=1}^n \{ \delta_i \log f_{\text{pop}}(t_i) + (1 - \delta_i) \log S_{\text{pop}}(t_i) \},$$

em que $(S_{\text{pop}}(t), f_{\text{pop}}(t))$ variam conforme o mecanismo de ativação.

Para simplificar a notação, seja:

$$S_i = S(t_i), \quad F_i = 1 - S_i, \quad f_i = f(t_i), \quad \eta_i = \left(\frac{t_i}{\sigma}\right)^\alpha, \quad u_i = \log\left(\frac{t_i}{\sigma}\right).$$

Desse modo:

$$\frac{\partial \eta_i}{\partial \alpha} = \eta_i u_i, \quad \frac{\partial \eta_i}{\partial \sigma} = -\frac{\alpha}{\sigma} \eta_i, \quad \frac{\partial u_i}{\partial \sigma} = -\frac{1}{\sigma}.$$

Como $\log f_i = \log \alpha - \log \sigma + (\alpha - 1)u_i - \eta_i$, obtém-se:

$$\frac{\partial \log f_i}{\partial \alpha} = \frac{1}{\alpha} + u_i - \eta_i u_i \tag{A.1}$$

$$\frac{\partial \log f_i}{\partial \sigma} = \frac{\alpha}{\sigma} (\eta_i - 1). \tag{A.2}$$

Como $S_i = \exp(-\eta_i)$ e $F_i = 1 - S_i$,

$$\boxed{\frac{\partial S_i}{\partial \alpha} = -S_i \eta_i u_i}, \quad \boxed{\frac{\partial S_i}{\partial \sigma} = S_i \frac{\alpha}{\sigma} \eta_i}, \quad \boxed{\frac{\partial F_i}{\partial \cdot} = -\frac{\partial S_i}{\partial \cdot}}. \quad (\text{A.3})$$

A.1.1 Primeira ativação

Sejam $S_{\text{pop}}(t)$ e $f_{\text{pop}}(t)$ definidas conforme (4.2) e (4.4), respectivamente. Assim, obtém-se a seguinte expressão para a log-verossimilhança:

$$\ell(\vartheta) = \sum_{i=1}^n \left\{ \delta_i \left[\log f_i + \log(C_i - D_i) \right] + (1 - \delta_i) \log(A_i - B_i) \right\},$$

em que:

$$A_i = \frac{2(p^3 - 1)}{1 - p^3 S_i}, \quad B_i = \frac{3(p^2 - 1)}{1 - p^2 S_i}, \quad C_i = \frac{2(p^3 - 1)p^3}{(1 - p^3 S_i)^2}, \quad D_i = \frac{3(p^2 - 1)p^2}{(1 - p^2 S_i)^2}.$$

A.1.1.0.1 Funções escore.

$$\frac{\partial \ell(\vartheta)}{\partial p} = \sum_{i=1}^n \left\{ \delta_i \frac{\partial_p(C_i - D_i)}{C_i - D_i} + (1 - \delta_i) \frac{\partial_p(A_i - B_i)}{A_i - B_i} \right\},$$

em que:

$$\begin{aligned} \frac{\partial A_i}{\partial p} &= \frac{6p^2(1 - S_i)}{(1 - p^3 S_i)^2}, & \frac{\partial B_i}{\partial p} &= \frac{6p(1 - S_i)}{(1 - p^2 S_i)^2}, \\ \frac{\partial C_i}{\partial p} &= 6p^2 \left[\frac{2p^3 - 1}{(1 - p^3 S_i)^2} + \frac{2p^3(p^3 - 1)S_i}{(1 - p^3 S_i)^3} \right], & \frac{\partial D_i}{\partial p} &= 6p \left[\frac{2p^2 - 1}{(1 - p^2 S_i)^2} + \frac{2p^2(p^2 - 1)S_i}{(1 - p^2 S_i)^3} \right]. \end{aligned}$$

Para os parâmetros da Weibull:

$$\frac{\partial \ell(\vartheta)}{\partial \alpha} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{\partial \log f_i}{\partial \alpha} + \frac{\partial_\alpha(C_i - D_i)}{C_i - D_i} \right] + (1 - \delta_i) \frac{\partial_\alpha(A_i - B_i)}{A_i - B_i} \right\},$$

$$\frac{\partial \ell(\vartheta)}{\partial \sigma} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{\partial \log f_i}{\partial \sigma} + \frac{\partial_\sigma(C_i - D_i)}{C_i - D_i} \right] + (1 - \delta_i) \frac{\partial_\sigma(A_i - B_i)}{A_i - B_i} \right\},$$

com:

$$\frac{\partial A_i}{\partial \alpha} = \frac{2(p^3 - 1)p^3}{(1 - p^3 S_i)^2} \frac{\partial S_i}{\partial \alpha}, \quad \frac{\partial B_i}{\partial \alpha} = \frac{3(p^2 - 1)p^2}{(1 - p^2 S_i)^2} \frac{\partial S_i}{\partial \alpha},$$

$$\frac{\partial C_i}{\partial \alpha} = \frac{4(p^3 - 1)p^6}{(1 - p^3 S_i)^3} \frac{\partial S_i}{\partial \alpha}, \quad \frac{\partial D_i}{\partial \alpha} = \frac{6(p^2 - 1)p^4}{(1 - p^2 S_i)^3} \frac{\partial S_i}{\partial \alpha},$$

e análogas para σ substituindo $\partial S_i / \partial \alpha$ por $\partial S_i / \partial \sigma$.

A.1.2 Última ativação

Sejam $S_{\text{pop}}(t)$ e $f_{\text{pop}}(t)$ definidas conforme (4.9) e (4.11), respectivamente. Assim, obtém-se a seguinte expressão para a log-verossimilhança:

$$\ell(\vartheta) = \sum_{i=1}^n \left\{ \delta_i \left[\log f_i + \log(\bar{C}_i - \bar{D}_i) \right] + (1 - \delta_i) \log \left[(1 + p_0) - \bar{A}_i + \bar{B}_i \right] \right\},$$

com:

$$\bar{A}_i = \frac{2(p^3 - 1)}{1 - p^3 F_i}, \quad \bar{B}_i = \frac{3(p^2 - 1)}{1 - p^2 F_i}, \quad \bar{C}_i = \frac{2(p^3 - 1)p^3}{(1 - p^3 F_i)^2}, \quad \bar{D}_i = \frac{3(p^2 - 1)p^2}{(1 - p^2 F_i)^2},$$

$$\text{e } p_0 = 2p^3 - 3p^2 + 1.$$

A.1.2.0.1 Funções score.

$$\frac{\partial \ell(\vartheta)}{\partial p} = \sum_{i=1}^n \left\{ \delta_i \frac{\partial_p(\bar{C}_i - \bar{D}_i)}{\bar{C}_i - \bar{D}_i} + (1 - \delta_i) \frac{(6p^2 - 6p) - \partial_p \bar{A}_i + \partial_p \bar{B}_i}{(1 + p_0) - \bar{A}_i + \bar{B}_i} \right\},$$

onde:

$$\frac{\partial \bar{A}_i}{\partial p} = \frac{6p^2(1 - F_i)}{(1 - p^3 F_i)^2}, \quad \frac{\partial \bar{B}_i}{\partial p} = \frac{6p(1 - F_i)}{(1 - p^2 F_i)^2},$$

$$\frac{\partial \bar{C}_i}{\partial p} = 6p^2 \left[\frac{2p^3 - 1}{(1 - p^3 F_i)^2} + \frac{2p^3(p^3 - 1)F_i}{(1 - p^3 F_i)^3} \right], \quad \frac{\partial \bar{D}_i}{\partial p} = 6p \left[\frac{2p^2 - 1}{(1 - p^2 F_i)^2} + \frac{2p^2(p^2 - 1)F_i}{(1 - p^2 F_i)^3} \right].$$

Para α e σ :

$$\frac{\partial \ell(\vartheta)}{\partial \alpha} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{\partial \log f_i}{\partial \alpha} + \frac{\partial_\alpha(\bar{C}_i - \bar{D}_i)}{\bar{C}_i - \bar{D}_i} \right] + (1 - \delta_i) \frac{-\partial_\alpha \bar{A}_i + \partial_\alpha \bar{B}_i}{(1 + p_0) - \bar{A}_i + \bar{B}_i} \right\},$$

$$\frac{\partial \ell(\vartheta)}{\partial \sigma} = \sum_{i=1}^n \left\{ \delta_i \left[\frac{\partial \log f_i}{\partial \sigma} + \frac{\partial_\sigma(\bar{C}_i - \bar{D}_i)}{\bar{C}_i - \bar{D}_i} \right] + (1 - \delta_i) \frac{-\partial_\sigma \bar{A}_i + \partial_\sigma \bar{B}_i}{(1 + p_0) - \bar{A}_i + \bar{B}_i} \right\}.$$

A.1.3 Ativação aleatória

Considere as funções $S_{\text{pop}}(t)$ e $f_{\text{pop}}(t)$, dadas por (4.13) e (4.14), respectivamente. Assim, a log-verossimilhança assume a forma:

$$\ell(\vartheta) = \sum_{i=1}^n \left\{ \delta_i \left[\log(1 - p_0) + \log f_i \right] + (1 - \delta_i) \log \left[p_0 + (1 - p_0)S_i \right] \right\}.$$

A.1.3.0.1 Funções escore.

$$\frac{\partial \ell(\vartheta)}{\partial p} = \sum_{i=1}^n \left\{ \delta_i \left(-\frac{p'_0}{1-p_0} \right) + (1-\delta_i) \frac{p'_0(1-S_i)}{p_0 + (1-p_0)S_i} \right\}, \quad p'_0 = 6p^2 - 6p,$$

$$\frac{\partial \ell(\vartheta)}{\partial \alpha} = \sum_{i=1}^n \left\{ \delta_i \frac{\partial \log f_i}{\partial \alpha} + (1-\delta_i) \frac{(1-p_0)}{p_0 + (1-p_0)S_i} \frac{\partial S_i}{\partial \alpha} \right\},$$

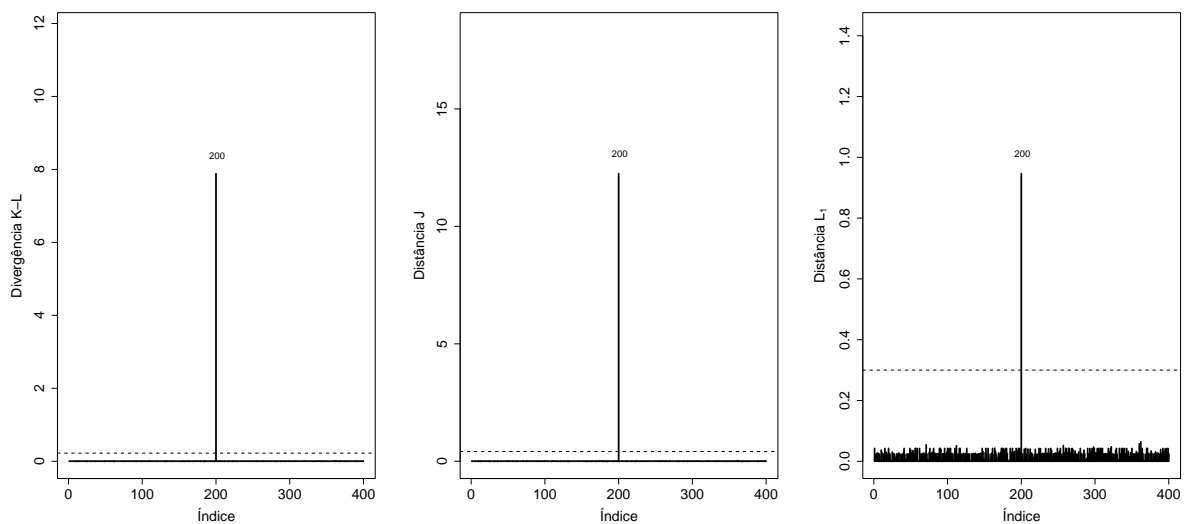
$$\frac{\partial \ell(\vartheta)}{\partial \sigma} = \sum_{i=1}^n \left\{ \delta_i \frac{\partial \log f_i}{\partial \sigma} + (1-\delta_i) \frac{(1-p_0)}{p_0 + (1-p_0)S_i} \frac{\partial S_i}{\partial \sigma} \right\}.$$

APÊNDICE

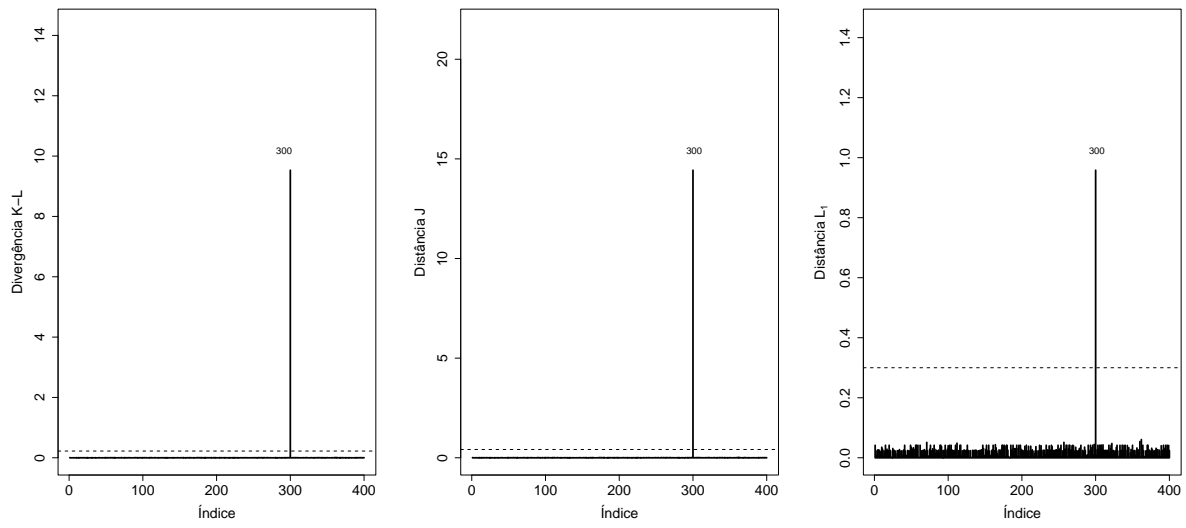
Este apêndice apresenta medidas de divergência ψ referentes aos estudos de simulação dos modelos sem e com covariável, descritos, respectivamente, nas Subseções 5.1.1 e 5.2.1.

As Figuras 32, 33, 34 e 35 exibem, respectivamente, as medidas de divergência ψ para os cenários (c), (d), (f) e (g) do modelo sem covariável. Em todas elas, as três medidas identificam corretamente as observações influentes.

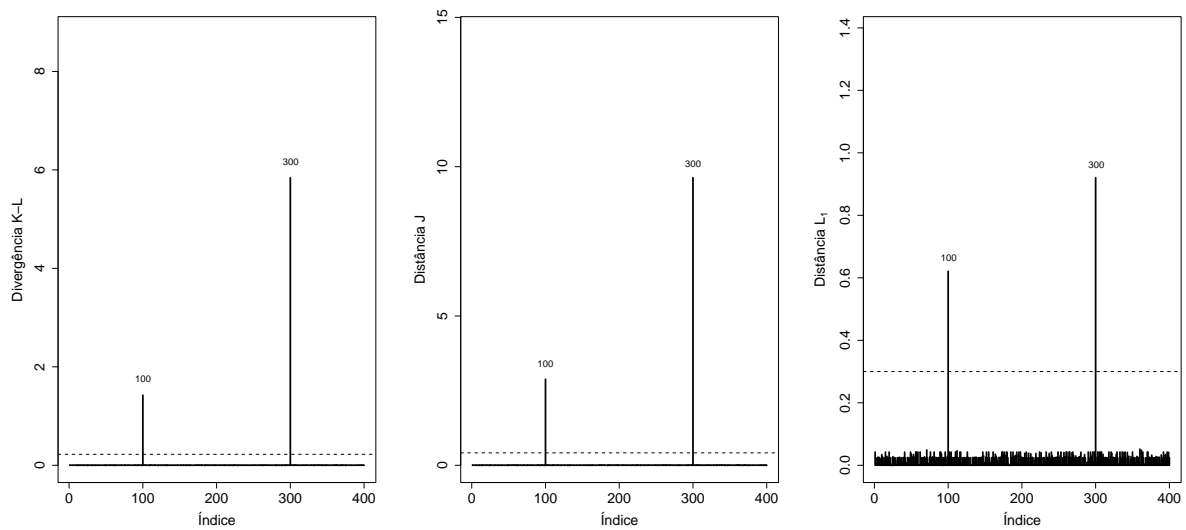
Figura 32 – Divergência ψ para o conjunto de dados (c) no caso sem covariável



Fonte: Elaborada pelo autor.

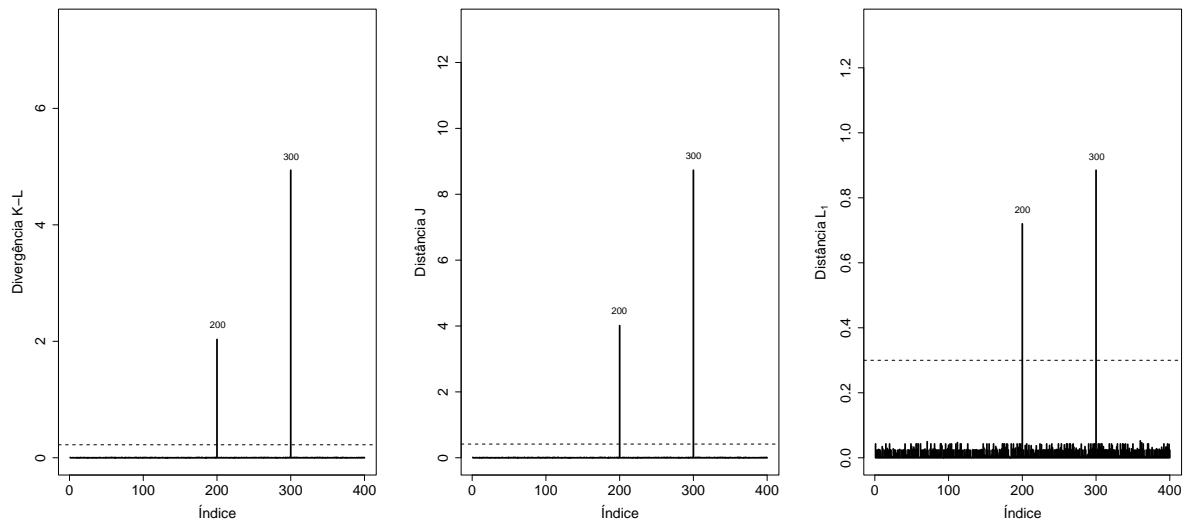
Figura 33 – Divergência ψ para o conjunto de dados (d) no caso sem covariável

Fonte: Elaborada pelo autor.

Figura 34 – Divergência ψ para o conjunto de dados (f) no caso sem covariável

Fonte: Elaborada pelo autor.

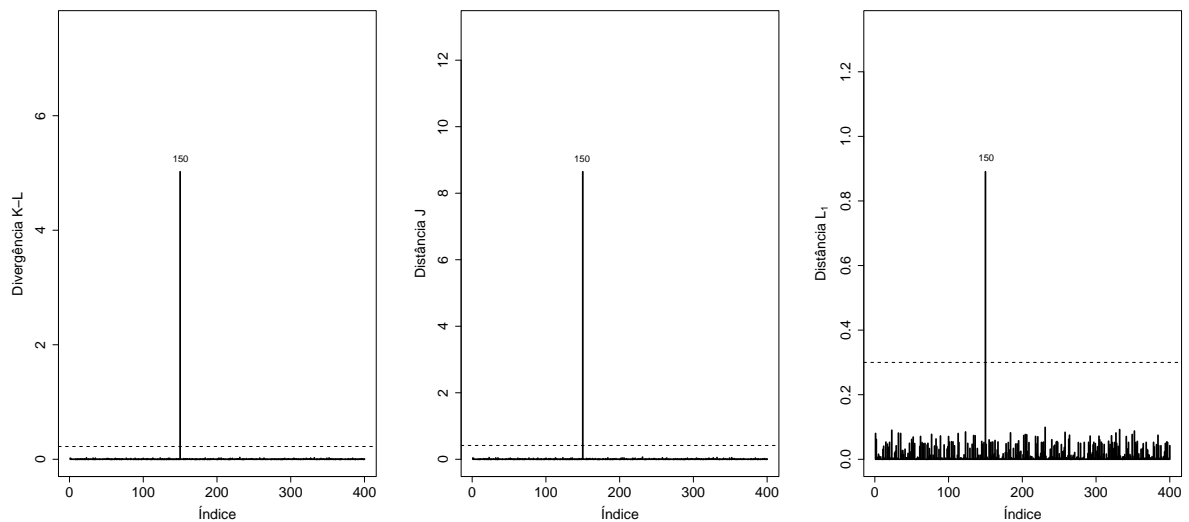
Figura 35 – Divergência ψ para o conjunto de dados (g) no caso sem covariável



Fonte: Elaborada pelo autor.

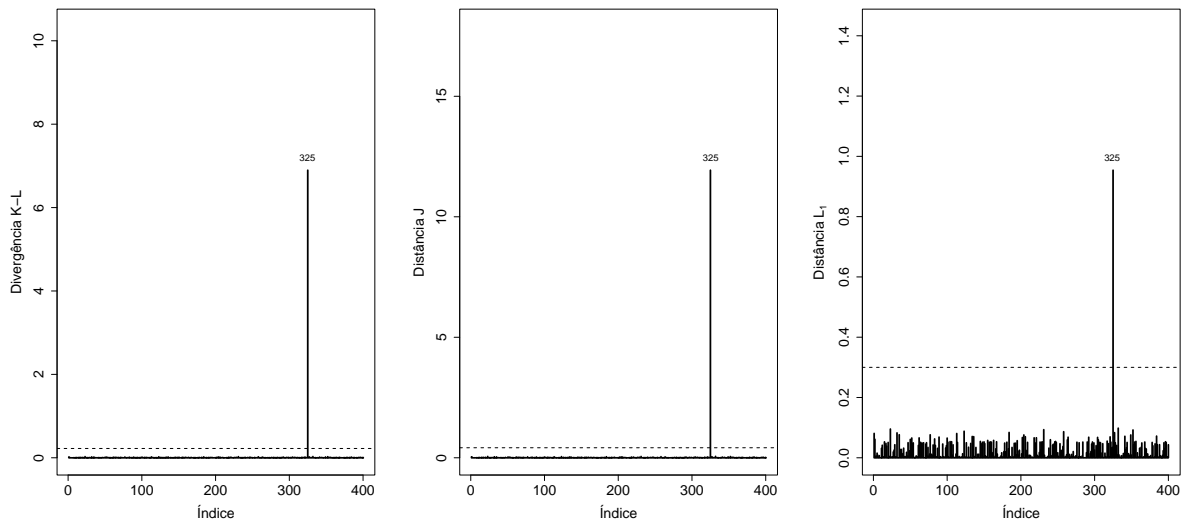
As Figuras 36, 37, 38 e 39 exibem, respectivamente, as medidas de divergência ψ para os cenários (c), (d), (f) e (g) no modelo com covariável. Em todas elas, as três medidas identificam corretamente as observações influentes.

Figura 36 – Divergência ψ para o conjunto de dados (c) no caso com covariável



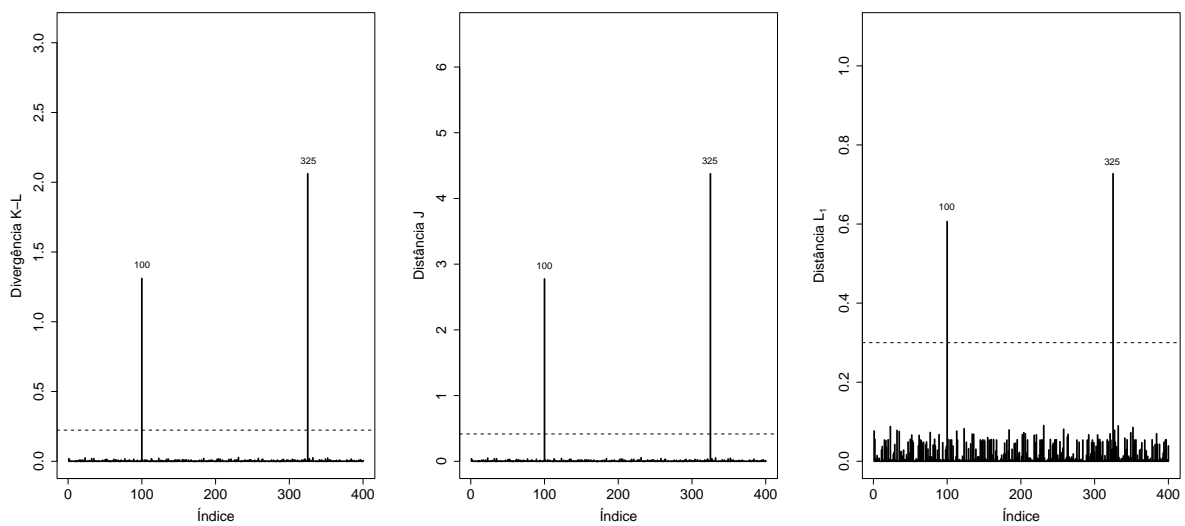
Fonte: Elaborada pelo autor.

Figura 37 – Divergência ψ para o conjunto de dados (d) no caso com covariável

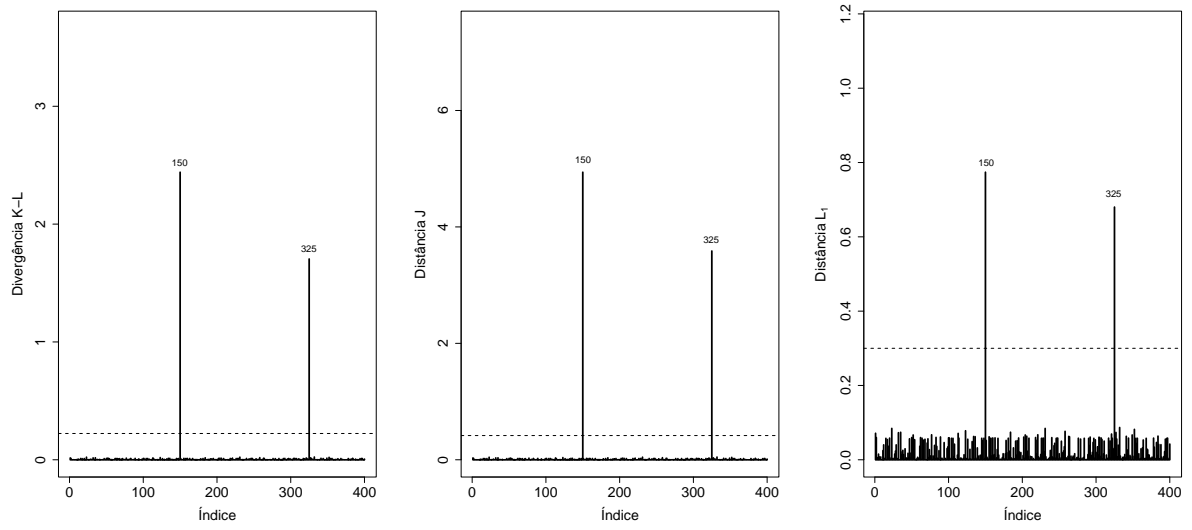


Fonte: Elaborada pelo autor.

Figura 38 – Divergência ψ para o conjunto de dados (f) no caso com covariável



Fonte: Elaborada pelo autor.

Figura 39 – Divergência ψ para o conjunto de dados (g) no caso com covariável

Fonte: Elaborada pelo autor.

APÊNDICE

C

APÊNDICE

Este apêndice reúne os resultados obtidos a partir da aplicação em dados reais, correspondentes às análises apresentadas nas Seções [5.1.2](#) e [5.2.2](#).

APÊNDICE

Código-fonte 1 – Código relacionado a distribuição Bilal discreta

```

1:
2: #' fdp Bilal
3: #'
4: dbilal <- function(x, p) {
5:   prob <- 2 * (p^3 - 1) * p^(3*x) - 3 * (p^2 - 1) * p^(2*x)
6:
7:   # Substituir NaN por zero
8:   if (is.vector(prob)) {
9:     prob[is.nan(prob)] <- 0 # Substituir NaN por zero em um
      vetor
10:  } else if (is.nan(prob)) {
11:    prob <- 0 # Substituir NaN por zero se for um número
12:  }
13:
14:   return(prob)
15: }
16:
17: #' fda Bilal
18: #'
19: pbilal <- function(x, p) return(
20:   1 - (3 - 2 * p^(x + 1)) * p ^ (2*x + 2)
21: )
22:
23: #' Simulacao Bilal
24: #'
25: rbilal <- function(n, p) {

```

```

26: simulated <- sample(0:1000, size = n, replace = TRUE, prob =
      dbilal(0:1000, p))
27: return(simulated)
28: }
29:
30: #' Esperanca Bilal
31: #'
32: exp_bilal <- function(p) {
33:   num <- p^2 * (p^2 + p + 3)
34:   den <- (p^2 + p + 1) * (1 - p^2)
35:   return(num/den)
36: }
37:
38: #' Variancia Bilal
39: #'
40: var_bilal <- function(p) {
41:   num <- p^2 * (3*p^4 + 4*p^3 - p^2 + 4*p + 3)
42:   den <- (p^2 + p + 1)^2 * (p^2 - 1)^2
43:   return(num/den)
44: }

```

Código-fonte 2 – Funções de log-verossimilhança sob os três mecanismos de ativação

```

1: #
2: # Feito por Renan Oliveira
3: # Primeira Ativação
4: #
5:
6: #' Função de verossimilhança para a primeira ativação
7: #' Distribuição Bilal
8: #'
9: lv_bilal_min <- function(vpar, x, status, y) {
10:   alpha = vpar[1]
11:   sigma = vpar[2]
12:   vbeta = vpar[3:4]
13:   p = exp(vbeta[1] + vbeta[2] * x) / (1 + exp(vbeta[1] + vbeta
      [2] * x))
14:
15:   suv = exp( - (y / sigma)^alpha)
16:
17:   # definindo a funcao de sobrevivencia
18:   spop = ((2*(p^3 - 1)) / (1 - p^3 * suv)) - ((3 * (p^2 - 1)) /

```

```

      (1 - p^2 * suv))
19:
20: # função densidade de probabilidade
21: parte1 = (2*alpha*(p^3)*suv*(y^(alpha-1))*(p^3-1)) / ((sigma^
      alpha) * ((1-(p^3)*suv)^2))
22: parte2 = (3*alpha*(p^2)*suv*(y^(alpha-1))*(p^2-1)) / ((sigma^
      alpha) * ((1-(p^2)*suv)^2))
23:
24: fpop = parte1 - parte2
25:
26: loglik = sum(status*log(fpop) + (1-status)*log(spop))
27: return(loglik)
28: }
29:
30: #
31: # Feito por Renan Oliveira
32: # Ultima Ativação
33: #
34:
35: #' Função de verossimilhança para a primeira ativação
36: #' Distrubuição Bilal
37: #'
38: lv_bilal_max <- function(vpar, x, status, y) {
39:   alpha = vpar[1]
40:   sigma = vpar[2]
41:   vbeta = vpar[3:4]
42:   p = exp(vbeta[1] + vbeta[2] * x) / (1 + exp(vbeta[1] + vbeta
      [2] * x))
43:
44:   suv = exp( - (y / sigma)^alpha)
45:   fac = 1 - exp( - (y / sigma)^alpha)
46:
47:   # definindo a funcao de sobrevivencia
48:   spop = 2*p^3 - 3*p^2 + 2 - (2*(p^3 - 1))/(1-(p^3)*fac) + (3*(
      p^2 - 1))/(1-(p^2)*fac)
49:
50: # função densidade de probabilidade
51: parte1 = (2*(p^3 - 1)*(p^3)*dweibull(y, alpha, sigma)) / ((1
      - p^3 * fac)^2)
52: parte2 = (3*(p^2 - 1)*(p^2)*dweibull(y, alpha, sigma)) / ((1
      - p^2 * fac)^2)
53:

```

```
54: fpop = parte1 - parte2
55:
56: loglik = sum(status*log(fpop) + (1-status)*log(spop))
57: return(loglik)
58: }
59:
60: #
61: # Feito por Renan Oliveira
62: # Ativação Aleatoria
63: #
64:
65: #' Função de verossimilhança para a primeira ativação
66: #' Distribuição Bilal
67: #'
68: lv_bilal_aleatoria <- function(vpar, x, status, y) {
69:   alpha = vpar[1]
70:   sigma = vpar[2]
71:   vbeta = vpar[3:4]
72:   p = exp(vbeta[1] + vbeta[2] * x) / (1 + exp(vbeta[1] + vbeta
       [2] * x))
73:
74:   suv = exp( - (y / sigma)^alpha) #1 - fac
75:   fac = 1 - suv#pweibull(y, alpha, sigma)
76:   fdp = (alpha/sigma) * ((y/sigma)^(alpha-1)) * suv #dweibull(
       y, alpha, sigma)
77:
78:   p0 = 2*(p^3) - 3*(p^2) + 1
79:
80:   # definindo a funcao de sobrevivencia
81:   spop = p0 + (1 - p0) * suv
82:
83:   # função densidade de probabilidade
84:   fpop = (1 - p0) * fdp
85:
86:   loglik = sum(status*log(fpop) + (1-status)*log(spop))
87:   return(loglik)
88: }
```
