

**MODELOS DE RISCOS
ADITIVO-MULTIPLICATIVOS EM
ANÁLISE DE SOBREVIVÊNCIA**

Tamy Harumy Moraes Tsujimoto

DISSERTAÇÃO APRESENTADA AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO TÍTULO DE
MESTRE EM CIÊNCIAS

Orientador: Antonio Carlos Pedroso de Lima

Durante o desenvolvimento deste trabalho a autora recebeu auxílio financeiro da FAPESP

São Paulo, abril de 2015

Modelos de riscos aditivo-multiplicativos em análise de sobrevivência

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 30/04/2015. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Antonio Carlos Pedroso de Lima - IME-USP
- Prof^a. Dr^a. Gisela Tunes da Silva - IME-USP
- Prof. Dr. Giovanni Loiola da Silva - IST-Portugal

*Aos meus pais Elaine e Luiz
e aos meus irmãos Thaís e Leonardo.*

Resumo

Este trabalho visa estudar o modelo de Cox-Aalen, que incorpora efeitos fixos e variantes no tempo por meio das estruturas aditiva e multiplicativa. Os principais modelos aditivos e multiplicativos que englobam efeitos fixos e variantes no tempo são revisados com enfoque de processos de contagem e um teste do tipo escore para auxiliar na escolha entre a estrutura aditiva e multiplicativa é proposto e avaliado por meio de estudos de simulação. Aplicação e comparação dos principais modelos discutidos em dados do Instituto do Câncer do Estado de São Paulo (ICESP) sugerem que o modelo estudado pode ser uma opção interessante na análise de dados de sobrevivência em que as ferramentas usuais não são adequadas.

Palavras-chave: efeito dependente do tempo, teste escore, Aranda-Ordaz, Cox-Aalen

Abstract

This dissertation aims to study the Cox-Aalen regression model, which incorporates fixed and time-varying effects through the additive and multiplicative structures. The primary additive and multiplicative risk models that deal with fixed and time-dependent effects are reviewed with the counting process framework and a score test to assist the decision between the additive and multiplicative structures is proposed and evaluated with a simulation study. The application of the models to the Instituto do Câncer do Estado de São Paulo (ICESP) dataset suggests that the model under study can be considered as an interesting option to analyse survival data when the usual techniques are not appropriate.

Keywords: time-dependent effect, score test, Aranda-Ordaz, Cox-Aalen.

Agradecimentos

Primeiramente, gostaria de agradecer meu orientador Antonio Carlos Pedroso de Lima, com quem tive a chance de trabalhar não só neste trabalho, mas também em outras oportunidades como o meu projeto de iniciação científica e o meu trabalho de conclusão de curso. Sou grata por todo crescimento acadêmico proporcionado por ele, desde disciplinas ministradas e apresentações em congressos até um estágio de pesquisa no exterior. Agradeço pelas críticas construtivas, pela motivação nos momentos frustrantes e por me lembrar sempre que valores como família e saúde devem sempre prevalecer.

Agradeço também o professor Giovani Loiola da Silva, com quem tive a oportunidade de trabalhar durante meu estágio de pesquisa na University of Western Ontario (Londres, ON, Canada). Sou grata pela paciência e atenção dadas ao meu trabalho e pela sua agradável companhia durante minha estada. Agradeço também a professora Charmaine Dean, que além de possibilitar minha estada na University of Western Ontario, me recebeu e me incluiu em seu grupo de estudos durante meu estágio de pesquisa.

Gostaria também de agradecer todos os professores que participaram da formação acadêmica, em especial os professores Julio da Motta Singer, Elisabeti Kira, Nelson Tanaka, Lucia Barroso e Luis Gustavo Esteves.

Sou imensamente grata aos meus pais Luiz e Elaine, pelos valores ensinados e por nunca duvidarem da minha capacidade. Agradeço aos meus irmãos Thaís e Leonardo, por sempre me lembrarem do valor de abraços e piadas entre irmãos. Sou grata também aos meus tios Claudio e Tiemi e aos meus avós Luiz e Lucia por todo apoio e acolhimento nos momentos em que mais precisei. Agradeço também meu tio Carlos, que me ajudou de todas as maneiras possíveis, em todas as fases da minha vida.

Agradeço ao meu namorado Victor, que me acompanha desde a graduação, pelo carinho, paciência e pela certeza de que poderei sempre contar com conselhos e um abraço em momentos difíceis. Agradeço também à minha amiga Tuany pelo companheirismo, pelas longas conversas e

pelas risadas em momentos inoportunos. Agradeço à minha amiga Damaris pelas ótimas histórias que colecionamos durante a graduação.

Agradeço também à Locus Danças pelas horas de descontração e relaxamento que foram cruciais para o bom desempenho desta dissertação. Em especial agradeço aos meus amigos Rafael Coelho, Erika Takano, Derek Williams e Priscilla Sayuri.

Por fim, agradeço à Fundação de Amparo à Pesquisa do Estado de São Paulo pelo apoio financeiro durante a elaboração deste trabalho.

TAMY HARUMY MORAES TSUJIMOTO

São Paulo, abril de 2015

Sumário

Introdução	1
1 Motivação	3
2 Modelos de riscos aditivos vs multiplicativos - efeitos fixos no tempo	9
2.1 Modelo de riscos multiplicativos	9
2.2 Modelo de riscos aditivos	11
2.3 Teste de modelos aditivos e multiplicativos para dados agrupados	12
2.3.1 Modelos de regressão para dados agrupados	13
2.3.2 Família de transformações assimétricas e modelo proposto	14
2.3.3 Teste escore	15
2.3.4 Estudos de simulação	19
3 Modelos de riscos aditivos vs multiplicativos - efeitos variantes no tempo	25
3.1 Modelo de Cox generalizado	25
3.2 Modelo de riscos aditivos de Aalen	26
3.2.1 Modelo semiparamétrico de riscos aditivos	30
3.2.2 Estimação da função de sobrevivência	32
4 Modelo de Cox-Aalen	33
4.1 Modelo	33
4.2 Estimação	34
4.3 Propriedades assintóticas e inferenciais	36
4.3.1 Para os riscos relativos β	37
4.3.2 Para os coeficientes acumulados $A(t)$	37
4.4 Técnicas de diagnóstico	39
4.5 Estimação da função de sobrevivência	40

5	Aplicação	43
5.1	Modelo de Cox	44
5.2	Modelo semiparamétrico de riscos aditivos	45
5.3	Modelo de Cox-Aalen	47
5.3.1	Modelo tradicional	49
5.3.2	Modelo proposto	50
5.4	Comparação dos modelos	51
5.5	Discussão	53
6	Pesquisas futuras	55
A	Conceitos básicos de processos de contagem	57
B	Gráficos dos ajustes dos modelos	63
	Referências Bibliográficas	73

Introdução

O modelo de [Cox \(1975\)](#) tem sido amplamente utilizado em estudos cujo principal interesse está na análise do tempo até a ocorrência de um determinado evento de interesse. Este modelo considera que a função taxa de falha em um dado instante t é dada por

$$h^*(t; \mathbf{x}) = h_0(t) \exp\{\mathbf{x}^\top \boldsymbol{\beta}\},$$

em que $\mathbf{x} = (x_1, \dots, x_p)^\top$ é o vetor p -dimensional de covariáveis, $\boldsymbol{\beta}$ é o respectivo vetor de parâmetros e $h_0(t)$ é a função taxa de falha basal, um componente não paramétrico que não precisa ser especificado previamente. A suposição básica desse modelo é a de que os riscos são proporcionais, isto é, a razão das taxas de falha de dois indivíduos que diferem pelas covariáveis deve se manter contante no tempo.

No entanto, casos de covariáveis com efeitos que dependem do tempo não são raros em situações práticas. Neste caso, o modelo de Cox passa a não ser a ferramenta mais adequada devido à violação da suposição de riscos proporcionais. É o caso do estudo conduzido pelo Instituto do Câncer do Estado de São Paulo (ICESP), em que 793 pacientes oncológicos em estágio terminal foram acompanhados de 2010 a 2012. Neste estudo, foram encontrados indícios de fuga da suposição de proporcionalidade para algumas covariáveis, possivelmente explicados pela presença de efeitos dependentes do tempo.

Uma possível abordagem para dados desta natureza é a generalização do modelo de Cox

$$h^*(t; \mathbf{x}) = h_0(t) \exp\{\mathbf{x}^\top \boldsymbol{\beta}(t)\},$$

que incorpora efeitos dependentes do tempo e vem sendo estudado por diversos autores, tais como [Brown \(1975\)](#), [Taulbee \(1979\)](#), [Murphy e Sen \(1991\)](#), [Zucker e Karr \(1990\)](#), [Grambsch e Therneau \(1994\)](#), entre outros.

Outra alternativa é o modelo de riscos aditivos de [Aalen \(1980\)](#)

$$h^*(t; \mathbf{x}(t)) = \mathbf{x}^\top(t)\boldsymbol{\alpha}(t),$$

que incorpora efeitos dependentes do tempo por meio da estrutura aditiva e apresenta técnicas mais simples de estimação.

No entanto, a escolha entre as estruturas aditiva e multiplicativa pode não ser trivial, uma vez que algumas covariáveis podem ser melhor explicadas pela estrutura aditiva, e outras, pela estrutura multiplicativa. Além disso, muitas vezes os estudos englobam tanto covariáveis com efeitos fixos quanto variantes no tempo. Dessa forma, o uso de modelos que combinem as estruturas aditiva e multiplicativa, englobando efeitos fixos e variantes no tempo pode ser interessante.

No modelo de Cox-Aalen, proposto por [Scheike e Zhang \(2002\)](#), a taxa de falha basal do modelo de Cox é substituído pelo modelo de riscos aditivos de Aalen

$$h^*(t; \mathbf{x}(t); \mathbf{z}(t)) = \left[\mathbf{x}^\top(t)\boldsymbol{\alpha}(t) \right] \exp\{\mathbf{z}^\top(t)\boldsymbol{\beta}\}.$$

Neste caso, tem-se um modelo com uma parcela aditiva, com efeitos dependentes do tempo, e uma outra parcela multiplicativa, com efeitos fixos no tempo, combinadas por meio do produto.

Neste trabalho, tem-se como objetivo aprofundar o conhecimento do modelo de Cox-Aalen quanto à estimação, propriedades assintóticas e técnicas de diagnóstico. Além disso, deseja-se propor uma metodologia para auxiliar na etapa de seleção de modelos da aplicação aos dados do estudo do ICESP, que é a principal motivação deste trabalho e está descrito no Capítulo 1.

Para tal, é realizado primeiramente um estudo dos conceitos básicos de processos de contagem no Apêndice A. Em seguida, são discutidos os modelos aditivos e multiplicativos com efeitos fixos no tempo e é proposto um teste do tipo score para auxiliar na escolha da melhor estrutura no Capítulo 2. No Capítulo 3, são discutidos os modelos aditivos e multiplicativos que incorporam efeitos dependentes do tempo, dando-se maior ênfase para o modelo aditivo de Aalen. No Capítulo 4, aprofunda-se o conhecimento acerca do modelo de Cox-Aalen e é realizada a aplicação e a comparação dos principais modelos discutidos aos dados do estudo do ICESP no Capítulo 5. Por último, pesquisas futuras são discutidas no Capítulo 6.

Capítulo 1

Motivação

A principal motivação desta dissertação é proveniente de um estudo conduzido pelo Instituto do Câncer do Estado de São Paulo (ICESP) envolvendo 793 pacientes oncológicos em estágio terminal, acompanhados no período de 2010 a 2012. Como todos os pacientes foram encaminhados à Unidade de Tratamento Intensivo (UTI) em algum momento do acompanhamento, o estudo visa avaliar o tempo (em dias) desde a internação na UTI até o óbito do paciente, de acordo com alguns fatores de interesse.

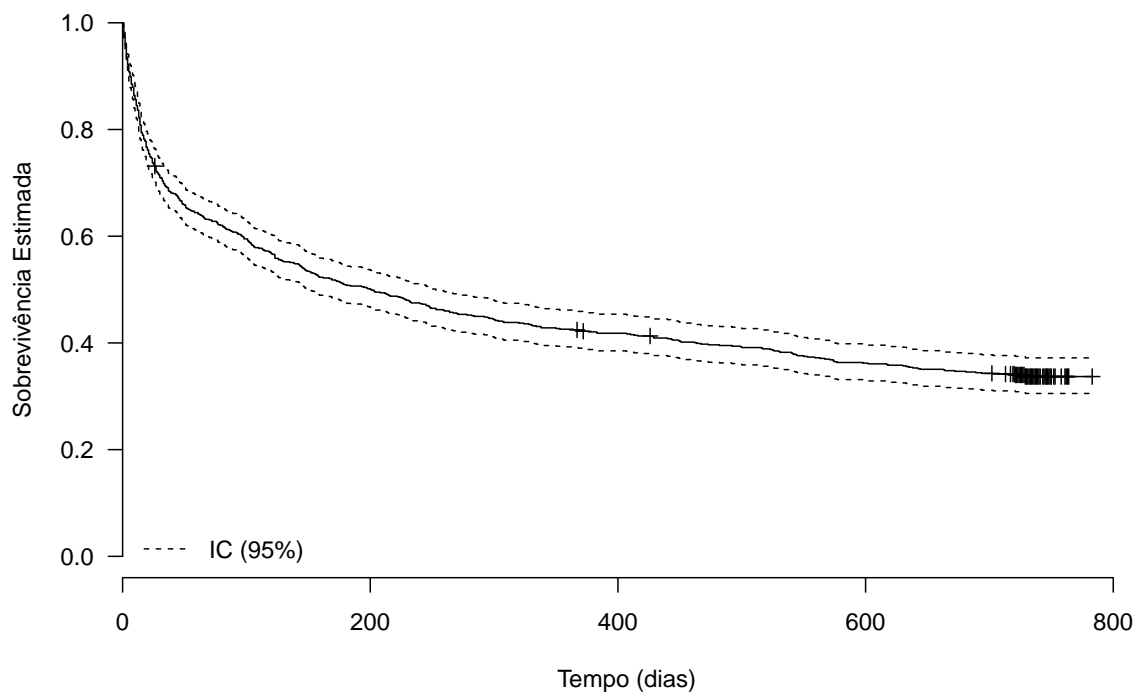


Figura 1.1: Curva de Kaplan Meier dos tempos de sobrevivência dos pacientes do ICESP.

Apesar do diagnóstico de câncer terminal dos pacientes, observa-se, pela curva de Kaplan Meier dos tempos de sobrevivência na Figura 1.1, que cerca de 34% dos pacientes não apresentam o evento ao final do estudo. O banco de dados é composto por 224 variáveis, das quais apenas um subconjunto com as principais variáveis de interesse é utilizado para fins de aplicação neste trabalho. São elas:

- Sexo (Feminino; Masculino)
- Idade (anos);
- Índice de Massa Corpórea (kg/m^2);
- Status do câncer (Ativo - recente; Ativo - recaída/progressão; Controlado/remissão);
- Tipo de admissão na UTI (Cirúrgico; Não cirúrgico);
- Extensão do câncer (Limitado; Localmente avançado; Leucemias e Metástase à distância);
- Tratamento prévio: Cirurgia (Sim; Não);
- Tratamento prévio: Radioterapia (Sim; Não);
- Tratamento prévio: Quimioterapia (Sim; Não);
- Presença de delirium (Não; Sim/sedado);
- Escala de capacidade funcional ECOG (0 a 4; quanto maior, pior o estado de saúde do paciente);
- Escore de qualidade de vida no instante de admissão na UTI (0 a 1; quanto maior, melhor a qualidade de vida do paciente);
- Escore SAPS III - índice de mortalidade para pacientes admitidos em UTIs (Adimensional; quanto maior, maior a gravidade do paciente admitido).

Primeiramente, foi realizada uma análise preliminar por meio de gráficos de Kaplan Meier, em conjunto com testes não paramétricos de Log-rank e Wilcoxon (Lawless, 2011), a fim de verificar a influência de cada uma das covariáveis no tempo de sobrevivência dos pacientes. Na Tabela 1.1 a seguir, são apresentados os níveis descritivos dos testes, que foram realizados de maneira univariada, bem como a ordem de pior para melhor prognóstico de cada uma das covariáveis consideradas.

Tabela 1.1: Níveis descritivos dos testes não paramétricos de Logrank e Wilcoxon e ordem de pior prognóstico para as variáveis do ICESP.

Variável	Log-rank	Wilcoxon	Ordem de pior prognóstico
Sexo	0,369	0,223	-
Idade	0,083	0,078	-
IMC	< 0,001	< 0,001	Menores valores de IMC 1 - Ativo - recaída/progressão
Status câncer	< 0,001	< 0,001	2 - Controlado/remissão/ 3 - Ativo - diagnóstico recente 1 - Não cirúrgico
Tipo de admissão	< 0,001	< 0,001	2 - Cirúrgico 1 - Leucemias/Metástase à distância
Extensão câncer	< 0,001	< 0,001	2 - Localmente avançado 3 - Limitado
Cirurgia	< 0,001	< 0,001	1 - Não 2 - Sim
Radioterapia	< 0,001	0,001	1 - Sim 2 - Não
Quimioterapia	< 0,001	< 0,001	1 - Sim 2 - Não
Presença delirium	< 0,001	< 0,001	1 - Sim/sedado 2 - Não
Escala de capacidade funcional ECOG	< 0,001	< 0,001	Maiores valores da escala
Escore de qualidade de vida	< 0,001	< 0,001	Menores valores do escore
Escore SAPS III	< 0,001	< 0,001	Maiores valores do escore

Além da análise preliminar descrita anteriormente, foram construídos gráficos de resíduos de Schoenfeld (1982) para cada uma das variáveis, a fim de verificar a adequabilidade da suposição de proporcionalidade dos riscos. A partir disso, observou-se que para algumas variáveis esta suposição não parece adequada, como no caso do Escore SAPS III, cujo gráfico dos resíduos é apresentado na Figura 1.2.

Sabe-se que a curva suavizada dos resíduos de Schoenfeld pode ser vista como uma estimativa do coeficiente ao longo do tempo (Grambsch e Therneau, 1994). Dessa forma, se a proporcionalidade dos riscos é válida, então espera-se que esta curva seja próxima de uma linha horizontal, como no gráfico de resíduos do Escore de qualidade de vida, apresentado na Figura 1.3. No entanto, os resíduos do Escore SAPS III apresentam uma certa inclinação negativa, indicando uma possível dependência temporal do efeito desta variável.

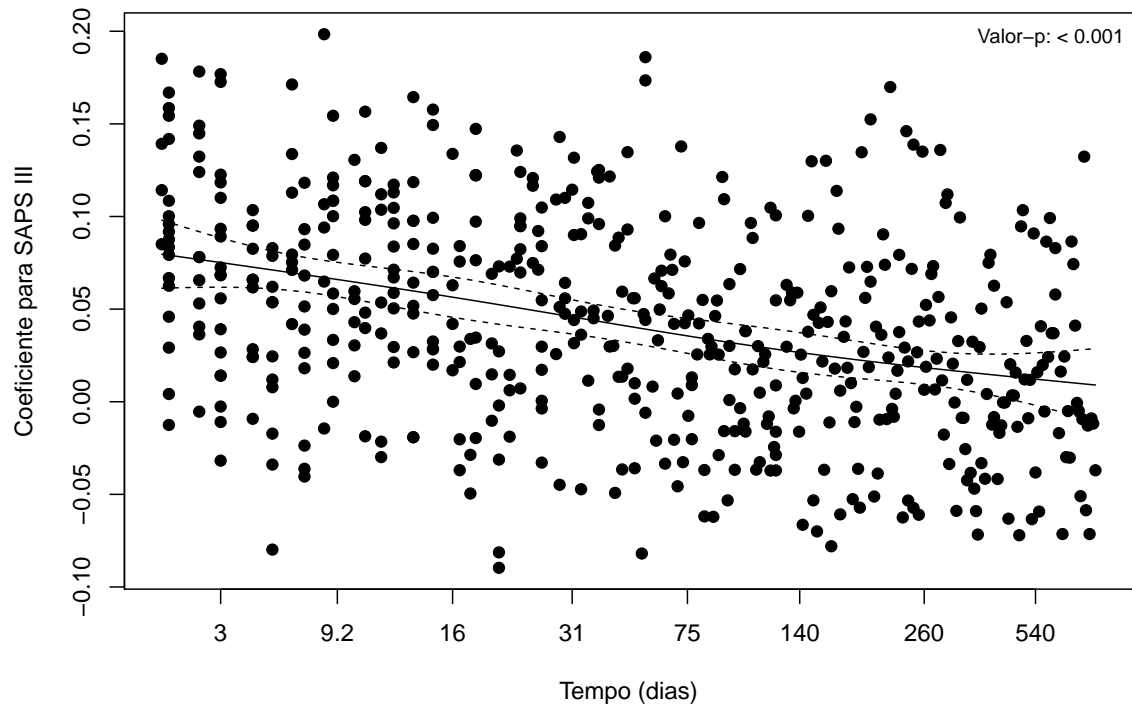


Figura 1.2: Resíduos de Schoenfeld e nível descritivo do teste de proporcionalidade dos riscos para o Escore SAPS III.

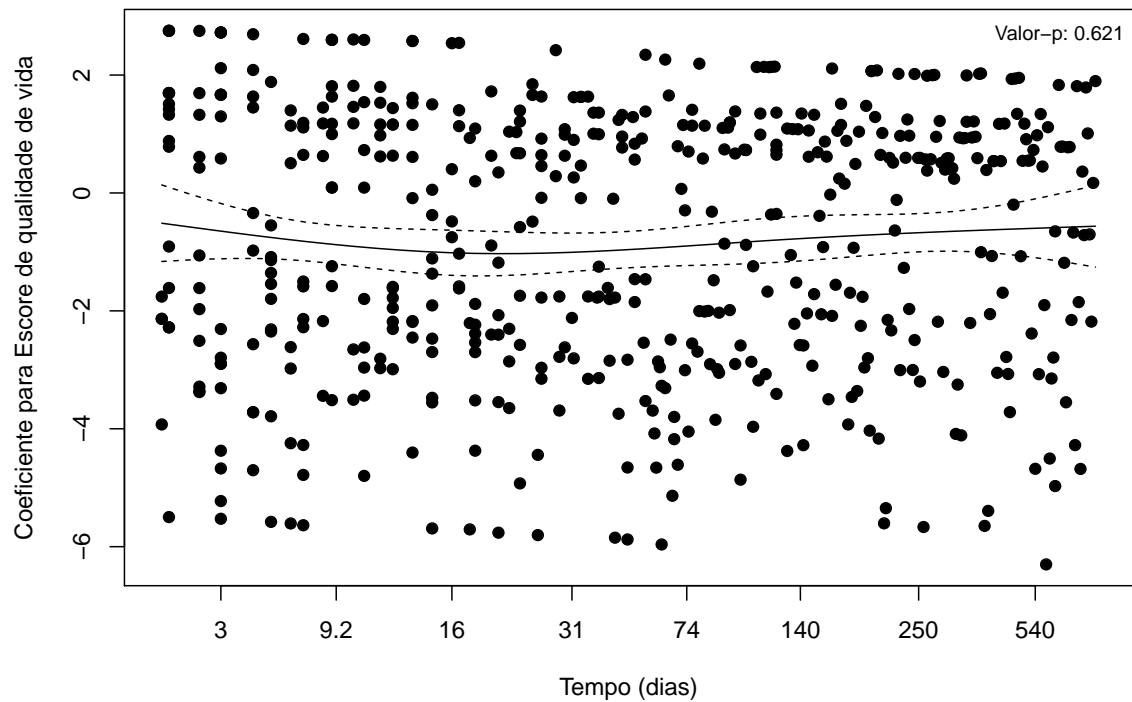


Figura 1.3: Resíduos de Schoenfeld e nível descritivo do teste de proporcionalidade dos riscos para o Escore de qualidade de vida.

Como o recurso gráfico pode envolver interpretações subjetivas, também foram realizados testes de hipótese para auxiliar na verificação da proporcionalidade dos riscos. Nos gráficos das Figuras 1.2 e 1.3 são apresentados os níveis de significância do teste, cuja estatística é baseada

na correlação linear entre os resíduos padronizados de Schoenfeld e o tempo (Colosimo e Giolo, 2006), em que valores próximos de zero levam à não rejeição da hipótese. A seguir é apresentada na Tabela 1.2 os níveis descritivos do teste global de proporcionalidade para todas as variáveis consideradas neste trabalho:

Tabela 1.2: Estatística e nível descritivo do teste global de proporcionalidade dos riscos.

Variável	Estatística	Valor-p
Sexo	1,660	0,197
Idade	4,290	0,038
IMC	0,170	0,680
Status câncer	1,259	0,533
Tipo de admissão	42,500	< 0,001
Extensão câncer	0,036	0,982
Cirurgia	9,090	0,003
Radioterapia	13,100	< 0,001
Quimioterapia	0,255	0,613
Presença delirium	15,900	< 0,001
Escala de capacidade funcional ECOG	0,453	0,501
Escala de qualidade de vida	0,244	0,621
Escala SAPS III	59,000	< 0,001

A falha na suposição de proporcionalidade dos riscos pode ser explicada pela possível dependência temporal dos efeitos das covariáveis. Na Tabela 1.2, verifica-se que o conjunto considerado neste trabalho é composto tanto por variáveis cujos efeitos são fixos no tempo, quanto por variáveis cujos efeitos são dependentes do tempo. Dessa forma, o estudo de modelos que acomodem dados desta natureza é o principal objetivo desta dissertação.

Neste trabalho, deseja-se aprofundar o conhecimento do modelo de risco aditivo-multiplicativo de Cox-Aalen, que admite tanto efeitos dependentes do tempo, por meio da estrutura aditiva, quanto fixos no tempo, por meio da estrutura multiplicativa. Além disso, são discutidos os principais modelos aditivos e multiplicativos com efeitos fixos e dependentes do tempo e é proposto um teste do tipo escore que irá auxiliar na etapa de seleção do modelo estudado.

Capítulo 2

Modelos de riscos aditivos vs multiplicativos - efeitos fixos no tempo

Neste Capítulo, os modelos de Cox (1975) e de riscos aditivos de Lin e Ying (1994) são brevemente discutidos sob a abordagem de processos de contagem, cujos conceitos básicos são apresentados no Apêndice A no final desta dissertação.

Os dois modelos apresentam diferentes formulações para a relação entre a taxa de falha e as covariáveis, e a escolha da melhor estrutura nem sempre é trivial. Dessa forma, desenvolve-se também um teste do tipo escore para dados agrupados a fim de auxiliar na escolha da estrutura mais adequada.

2.1 Modelo de riscos multiplicativos

O modelo de regressão de Cox (1975) tem sido amplamente utilizado na análise de dados de sobrevivência. Neste modelo, assume-se que a intensidade do processo de contagem $N(t)$ tem a seguinte forma:

$$\begin{aligned}h(t; \mathbf{x}) &= Y(t)h^*(t; \mathbf{x}) \\h^*(t; \mathbf{x}) &= h_0(t) \exp\{\mathbf{x}^\top \boldsymbol{\beta}\},\end{aligned}\tag{2.1}$$

em que $Y(t)$ é o indicador de risco, $h_0(t)$ é a função taxa de falha basal, $\mathbf{x} = (x_1, \dots, x_p)^\top$ é o vetor p -dimensional de covariáveis e $\boldsymbol{\beta}$, o seu respectivo vetor de parâmetros. A suposição básica desse modelo é a de que os riscos são proporcionais, isto é, a razão das taxas de falha de dois indivíduos que diferem pelas covariáveis deve se manter contante no tempo.

A estimação dos parâmetros neste modelo é realizada pelo método de máxima verossimilhança parcial. Para a construção da função verossimilhança parcial, consideram-se n observações $(N_i(t), Y_i(t), \mathbf{x}_i), i = 1, \dots, n$ com as formulações dadas no Apêndice e definem-se

$$N_{\bullet}(t) = \sum_{i=1}^n N_i(t) \quad M_{\bullet}(t) = \sum_{i=1}^n M_i(t).$$

Supondo que β é conhecido, pode-se construir um estimador do tipo Nelson-Aalen (desenvolvido no Apêndice A) para a função de taxa de falha acumulada basal $H_0(t)$, a partir da decomposição de Doob-Meyer:

$$\begin{aligned} dN_{\bullet}(t) &= \sum_{i=1}^n h_i(t)dt + dM_{\bullet}(t) \\ \Leftrightarrow dN_{\bullet}(t) &= dH_0(t) \sum_{i=1}^n Y_i(t) \exp(\mathbf{x}_i^{\top} \beta) + dM_{\bullet}(t) \\ \Leftrightarrow \frac{dN_{\bullet}(t)}{\sum_{i=1}^n Y_i(t) \exp(\mathbf{x}_i^{\top} \beta)} &= dH_0(t) + \frac{dM_{\bullet}(t)}{\sum_{i=1}^n Y_i(t) \exp(\mathbf{x}_i^{\top} \beta)}, \end{aligned}$$

para $\sum_{i=1}^n Y_i(t) \exp(\beta^{\top} \mathbf{x}_i) \neq 0$.

Definindo $J(t) = \mathbb{I}_{\{\sum_{i=1}^n Y_i(t) \exp(\beta^{\top} \mathbf{x}_i) > 0\}}$ e estabelecendo que $0/0 = 0$, tem-se que o estimador do tipo Nelson-Aalen para a taxa de falha acumulada basal será:

$$\hat{H}_0(t, \beta) = \int_0^t \frac{J(u)}{\sum_{i=1}^n Y_i(u) \exp(\mathbf{x}_i^{\top} \beta)} dN_{\bullet}(u). \quad (2.2)$$

Segundo [Andersen et al. \(1993\)](#), sabe-se que a função de verossimilhança para $t \in [0, \tau]$ será da forma

$$\begin{aligned} &\prod_{i=1}^n \prod_{t \leq \tau} \left[(h_i(t)dt)^{dN_i(t)} \exp \left(- \int_0^{\tau} h_i(t)dt \right) \right. \\ &\left. \prod_{i=1}^n \prod_{t \leq \tau} \left[dH_0(t) Y_i(t) \exp(\mathbf{x}_i^{\top} \beta) \right]^{dN_i(t)} \exp \left(- \int_0^{\tau} Y_i(t) \exp(\mathbf{x}_i^{\top} \beta) dH_0(t) \right) \right]. \end{aligned} \quad (2.3)$$

Substituindo o estimador de Nelson-Aalen obtido em (2.2) na função de verossimilhança (2.3), tem-se a função de verossimilhança parcial

$$\prod_{i=1}^n \prod_{t \leq \tau} \left(\frac{Y_i(t) \exp(\mathbf{x}_i^{\top} \beta)}{\sum_{i=1}^n Y_i(t) \exp(\mathbf{x}_i^{\top} \beta)} \right)^{dN_i(t)},$$

a partir da qual pode-se obter a seguinte função escore parcial para a estimação de β :

$$U(\beta) = \sum_{i=1}^n \int_0^{\tau} \left[\mathbf{x}_i - \frac{\sum_{i=1}^n Y_i(t) \exp(\beta^\top \mathbf{x}_i) \mathbf{x}_i}{\sum_{i=1}^n Y_i(t) \exp(\beta^\top \mathbf{x}_i)} \right] dN_i(t).$$

Pode-se verificar ainda que

$$n^{1/2}(\hat{\beta} - \beta) \xrightarrow{\mathcal{D}} N_p(\mathbf{0}, \Sigma^{-1}),$$

em que Σ pode ser estimada por $\hat{\Sigma} = n^{-1}I(\hat{\beta})$, com $I(\hat{\beta}) = -\partial U(\beta)/\partial \beta$.

2.2 Modelo de riscos aditivos

O modelo de riscos aditivos proposto por [Lin e Ying \(1994\)](#), que pode ser visto como um caso particular do modelo proposto por [Aalen \(1980\)](#), possui a seguinte formulação para o processo de intensidade:

$$\begin{aligned} h(t; \mathbf{x}) &= Y(t)h^*(t; \mathbf{x}) \\ h^*(t; \mathbf{x}) &= h_0(t) + \mathbf{x}^\top \boldsymbol{\alpha}, \end{aligned} \tag{2.4}$$

em que $Y(t)$ é o indicador de risco, $h_0(t)$ é a função taxa de falha basal, $\mathbf{x} = (x_1, \dots, x_p)^\top$ é o vetor p -dimensional de covariáveis e $\boldsymbol{\alpha}$, o seu vetor de parâmetros.

A estimação dos parâmetros é feita de maneira semelhante a do modelo de Cox, em que se constrói o estimador do tipo Nelson-Aalen para a taxa de falha basal e, posteriormente, substitui-se a estimativa obtida na função verossimilhança. Considerando-se que $\boldsymbol{\alpha}$ é conhecido, tem-se que:

$$\begin{aligned} dN_{\bullet}(t) &= \sum_{i=1}^n h_i(t)dt + dM_{\bullet}(t) \\ \Leftrightarrow dN_{\bullet}(t) &= dH_0(t) \sum_{i=1}^n Y_i(t) + \sum_{i=1}^n Y_i(t) \mathbf{x}_i^\top \boldsymbol{\alpha} dt + dM_{\bullet}(t) \\ \Leftrightarrow \frac{dN_{\bullet}(t)}{\sum_{i=1}^n Y_i(t)} &= dH_0(t) + \frac{\sum_{i=1}^n Y_i(t) \mathbf{x}_i^\top \boldsymbol{\alpha} dt}{\sum_{i=1}^n Y_i(t)} + \frac{dM_{\bullet}(t)}{\sum_{i=1}^n Y_i(t)}, \end{aligned}$$

para $\sum_{i=1}^n Y_i(t) \neq 0$.

Definindo $J(t) = \mathbb{I}_{\{\sum_{i=1}^n Y_i(t) > 0\}}$ e estabelecendo que $0/0 = 0$, tem-se que o estimador do tipo

Nelson-Aalen para a taxa de falha acumulada basal será:

$$\hat{H}_0(t, \alpha) = \int_0^t \frac{J(u) [dN_{\bullet}(u) - \sum_{i=1}^n Y_i(u) \mathbf{x}_i^{\top} \alpha du]}{\sum_{i=1}^n Y_i(u)}.$$

De maneira análoga ao modelo de Cox, pode-se construir a seguinte função escore para α :

$$U(\alpha) = \sum_{i=1}^n \int_0^{\tau} \left[\mathbf{x}_i - \frac{\sum_{i=1}^n Y_i(t) \mathbf{x}_i}{\sum_{i=1}^n Y_i(t)} \right] [dN_i(t) - Y_i(t) \mathbf{x}_i^{\top} \alpha dt],$$

da qual pode-se obter a seguinte expressão para o estimador de α :

$$\hat{\alpha} = \left\{ \sum_{i=1}^n \int_0^{\tau} Y_i(t) \left[\mathbf{x}_i - \frac{\sum_{i=1}^n Y_i(t) \mathbf{x}_i}{\sum_{i=1}^n Y_i(t)} \right]^{\otimes 2} dt \right\}^{-1} \left\{ \sum_{i=1}^n \int_0^{\tau} Y_i(t) \left[\mathbf{x}_i - \frac{\sum_{i=1}^n Y_i(t) \mathbf{x}_i}{\sum_{i=1}^n Y_i(t)} \right] dN_i(t) \right\},$$

em que $a^{\otimes 2} = aa^{\top}$.

A variância do estimador pode ser consistentemente estimada por

$$\hat{\Sigma} = A^{-1} B A^{-1},$$

em que

$$A = n^{-1} \sum_{i=1}^n \int_0^{\tau} Y_i(t) \left[\mathbf{x}_i - \frac{\sum_{i=1}^n Y_i(t) \mathbf{x}_i}{\sum_{i=1}^n Y_i(t)} \right]^{\otimes 2} dt$$

$$B = n^{-1} \sum_{i=1}^n \int_0^{\tau} \left[\mathbf{x}_i - \frac{\sum_{i=1}^n Y_i(t) \mathbf{x}_i}{\sum_{i=1}^n Y_i(t)} \right]^{\otimes 2} dN_i(t).$$

Pode-se mostrar que $n^{1/2}(\hat{\alpha} - \alpha)$ possui distribuição Normal p -variada com vetor de média zero e matriz de covariância Σ , estimada por $\hat{\Sigma}$.

2.3 Teste de modelos aditivos e multiplicativos para dados agrupados

Alguns estudos em análise de sobrevivência são conduzidos de modo que não é possível observar o instante exato, mas sim o intervalo de tempo em que ocorreu o determinado evento de interesse. Neste caso tem-se a censura do tipo intervalar, cujo caso particular, denominado dados agrupados, ocorre quando todos os indivíduos do estudo são avaliados nos mesmos instantes ao longo do acompanhamento.

A metodologia proposta nesta seção é baseada em modelos para dados agrupados em conjunto com a família de transformações assimétricas proposta por [Aranda-Ordaz \(1983\)](#), que pos-

sui como casos especiais os modelos aditivo e multiplicativo discutidos anteriormente. A partir disso, desenvolve-se um teste do tipo score a fim de auxiliar na verificação da melhor estrutura para modelar um dado conjunto de covariáveis. [Colosimo et al. \(2000\)](#) utilizam a mesma abordagem para a escolha entre os modelos de riscos proporcionais e de chances proporcionais.

2.3.1 Modelos de regressão para dados agrupados

Suponha n observações do tempo T até a ocorrência de um determinado evento de interesse e $\mathbf{x} = (x_1, \dots, x_p)^T$, o vetor de covariáveis associadas. Suponha ainda que os tempos de sobrevivência são agrupados em k intervalos $I_i = [t_{i-1}, t_i)$, $i = 1, \dots, k$, em que $0 = t_0 < t_1 < \dots < t_k = \infty$, e que toda censura ocorre apenas ao final dos intervalos. Defina ainda R_i como sendo o grupo de indivíduos em risco no instante t_{i-1} e δ_{li} , como indicador de falha do l -ésimo ($l = 1, \dots, n$) indivíduo no intervalo I_i .

Para a construção da função de verossimilhança para dados agrupados, faz-se necessária a definição da probabilidade $p_i(\mathbf{x}_l)$ do l -ésimo indivíduo falhar no intervalo I_i , condicional ao fato de estar em risco no instante t_{i-1} e às covariáveis associadas \mathbf{x}_l :

$$p_i(\mathbf{x}_l) = \mathbb{P}(T_l \leq t_i | T_l \geq t_{i-1}; \mathbf{x}_l) = 1 - \frac{S(t_i)}{S(t_{i-1})}, \quad (2.5)$$

em que $S(\cdot)$ é a função de sobrevivência. De acordo com [Lawless \(2011\)](#), tem-se que a função de verossimilhança para dados agrupados é dada por

$$\prod_{i=1}^k \prod_{l \in R_i} [p_i(\mathbf{x}_l)]^{\delta_{li}} [1 - p_i(\mathbf{x}_l)]^{1 - \delta_{li}}, \quad (2.6)$$

em que o indicador de falha δ_{li} pode ser visto como uma variável aleatória com distribuição de Bernoulli de parâmetro $p_i(\mathbf{x}_l)$.

Sabe-se ainda que $S(t) = \exp\left\{-\int_0^t h^*(u) du\right\}$ e, sendo assim, a probabilidade $p_i(\mathbf{x}_l)$ em (2.5) pode assumir diferentes formas dependendo da estrutura adotada na função taxa de falha. Para as estruturas definidas em (2.1) e (2.4), obtêm-se os seguintes modelos para dados agrupados:

- Modelo de riscos aditivos

$$\begin{aligned} p_i(\mathbf{x}_l) &= 1 - \left[\frac{S_0(t_i)}{S_0(t_{i-1})} \right] \exp\left\{-(t_i - t_{i-1})\boldsymbol{\alpha}^\top \mathbf{x}\right\} \Leftrightarrow \\ -\log[1 - p_i(\mathbf{x}_l)] &= -\log\left[\frac{S_0(t_i)}{S_0(t_{i-1})} \right] + (t_i - t_{i-1})\boldsymbol{\alpha}^\top \mathbf{x}; \end{aligned} \quad (2.7)$$

- Modelo de riscos multiplicativos

$$p_i(\mathbf{x}_l) = 1 - \left[\frac{S_0(t_i)}{S_0(t_{i-1})} \right]^{\exp(\boldsymbol{\beta}^\top \mathbf{x})} \Leftrightarrow$$

$$\log \{-\log[1 - p_i(\mathbf{x}_l)]\} = \log \left\{ -\log \left[\frac{S_0(t_i)}{S_0(t_{i-1})} \right] \right\} + \boldsymbol{\beta}^\top \mathbf{x}; \quad (2.8)$$

em que $S_0(\cdot)$ é a função de sobrevivência basal.

2.3.2 Família de transformações assimétricas e modelo proposto

Aranda-Ordaz (1983) propõe em seu artigo a seguinte família de transformações assimétricas, que depende de um parâmetro extra λ :

$$V_\lambda(p) = \begin{cases} \lambda^{-1} \{ [-\log(1-p)]^\lambda - 1 \}, & \lambda \neq 0 \\ \log [-\log(1-p)], & \lambda = 0 \end{cases}.$$

A partir desta família de transformações, pode-se propor o seguinte modelo:

$$V_\lambda [p_i(\mathbf{x}_l)] = \xi_i + \boldsymbol{\gamma}^\top \mathbf{x}_l = \eta_{li},$$

em que ξ_i $i = 1, \dots, k$ são os efeitos dos intervalos e $\boldsymbol{\gamma}$ é o vetor p -dimensional de parâmetros das covariáveis associadas. Note que este modelo apresenta a estrutura multiplicativa quando $\lambda = 0$, e a estrutura (essencialmente) aditiva quando $\lambda = 1$, supondo-se que os tamanhos dos intervalos são fixos, isto é, $t_i - t_{i-1} = \Delta$ $i = 1, \dots, k$. A inversa da função de ligação é facilmente obtida e é dada por:

$$p_i(\mathbf{x}_l) = V_\lambda^{-1}(\eta_{li}) = \begin{cases} 1 - \exp \{ -(1 + \lambda \eta_{li})^{1/\lambda} \}, & \lambda \neq 0, \quad \lambda \eta_{li} > -1 \\ 1 - \exp \{ -\exp(\eta_{li}) \}, & \lambda = 0 \\ 0, & \lambda \neq 0, \quad \lambda \eta_{li} \leq -1. \end{cases} \quad (2.9)$$

Dada esta formulação de modelo, é possível verificar qual a melhor estrutura (definida por λ) para explicar a relação entre função taxa de falha e as covariáveis. Em particular, neste trabalho deseja-se construir testes de hipóteses para discriminar entre as estruturas aditiva ($\lambda = 1$) e multiplicativa ($\lambda = 0$). Para tal, opta-se por um teste do tipo score, uma vez que é mais vantajoso computacionalmente devido a sua construção ser baseada em cálculos sob as hipóteses nulas.

2.3.3 Teste escore

Nesta etapa, são construídos testes do tipo escore para as hipóteses $\mathbb{H}_0 : \lambda = 0$ e $\mathbb{H}_0 : \lambda = 1$. Para tal, seja $\boldsymbol{\theta} = (\lambda, \boldsymbol{\xi}^\top, \boldsymbol{\gamma}^\top)^\top$ o vetor $(1 + k + p)$ -dimensional de parâmetros, em que $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)^\top$ e $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_p)^\top$. Sabe-se que a estatística do teste escore proposto por Rao (1948) é dada por:

$$S_R = \mathbf{U}^\top(\hat{\boldsymbol{\theta}}_0) \mathbf{I}^{-1}(\hat{\boldsymbol{\theta}}_0) \mathbf{U}(\hat{\boldsymbol{\theta}}_0),$$

em que $\hat{\boldsymbol{\theta}}_0$ é a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sob a hipótese nula, e $\mathbf{U}(\boldsymbol{\theta})$ e $\mathbf{I}(\boldsymbol{\theta})$ são a função escore e a informação de Fisher, dadas por:

$$\mathbf{U}(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \begin{bmatrix} \mathbf{U}_\lambda(\boldsymbol{\theta}) \\ \mathbf{U}_\xi(\boldsymbol{\theta}) \\ \mathbf{U}_\gamma(\boldsymbol{\theta}) \end{bmatrix},$$

$$\mathbf{I}(\boldsymbol{\theta}) = \mathbb{E} \left[-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \right] = \begin{bmatrix} \mathbf{I}_{\lambda\lambda}(\boldsymbol{\theta}) & \mathbf{I}_{\lambda\xi}(\boldsymbol{\theta}) & \mathbf{I}_{\lambda\gamma}(\boldsymbol{\theta}) \\ \mathbf{I}_{\xi\lambda}(\boldsymbol{\theta}) & \mathbf{I}_{\xi\xi}(\boldsymbol{\theta}) & \mathbf{I}_{\xi\gamma}(\boldsymbol{\theta}) \\ \mathbf{I}_{\gamma\lambda}(\boldsymbol{\theta}) & \mathbf{I}_{\gamma\xi}(\boldsymbol{\theta}) & \mathbf{I}_{\gamma\gamma}(\boldsymbol{\theta}) \end{bmatrix},$$

com $\ell(\boldsymbol{\theta})$ sendo a o logaritmo função de verossimilhança definida na expressão (2.6):

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^k \sum_{l \in R_i} \{ \delta_{li} \log p_i(\mathbf{x}_l) + (1 - \delta_{li}) \log [1 - p_i(\mathbf{x}_l)] \}.$$

A seguir, são calculadas as expressões gerais para os termos do vetor de escore e da matriz de informação de Fisher. Para tal, defina θ_q ($q = 1, \dots, p + k + 1$) como sendo o q -ésimo componente do vetor de parâmetros $\boldsymbol{\theta}$ e considere $p_{li} = p_i(\mathbf{x}_l)$ para maior simplicidade de notação.

- Função escore

$$\begin{aligned} \mathbf{U}_{\theta_q}(\boldsymbol{\theta}) &= \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_q} \\ &= \sum_{i=1}^k \sum_{l \in R_i} \left\{ \frac{\delta_{li}}{p_{li}} \frac{\partial p_{li}}{\partial \theta_q} - \frac{(1 - \delta_{li})}{1 - p_{li}} \frac{\partial p_{li}}{\partial \theta_q} \right\} \\ \mathbf{U}_{\theta_q}(\boldsymbol{\theta}) &= \sum_{i=1}^k \sum_{l \in R_i} \left\{ \frac{\delta_{li} - p_{li}}{p_{li}(1 - p_{li})} \frac{\partial p_{li}}{\partial \theta_q} \right\}. \end{aligned} \tag{2.10}$$

- Informação de Fisher

$$\begin{aligned}
I_{\theta_j \theta_q}(\boldsymbol{\theta}) &= \mathbb{E} \left[-\frac{\partial U_{\theta_q}(\boldsymbol{\theta})}{\partial \theta_j} \right] \\
&= \mathbb{E} \left[-\frac{\partial}{\partial \theta_j} \sum_{i=1}^k \sum_{l \in R_i} \left\{ \frac{\delta_{li} - p_{li}}{p_{li}(1-p_{li})} \frac{\partial p_{li}}{\partial \theta_q} \right\} \right] \\
&= \mathbb{E} \left[-\sum_{i=1}^k \sum_{l \in R_i} \left\{ \frac{\delta_{li} - p_{li}}{p_{li}(1-p_{li})} \frac{\partial^2 p_{li}}{\partial \theta_j \partial \theta_q} - \frac{p_{li}(1-p_{li}) + (\delta_{li} - p_{li})(1-2p_{li})}{p_{li}^2(1-p_{li})^2} \frac{\partial p_{li}}{\partial \theta_j} \frac{\partial p_{li}}{\partial \theta_q} \right\} \right] \\
&= \mathbb{E} \left[\sum_{i=1}^k \sum_{l \in R_i} \left\{ \frac{p_{li}(1-p_{li}) + (\delta_{li} - p_{li})(1-2p_{li})}{p_{li}^2(1-p_{li})^2} \frac{\partial p_{li}}{\partial \theta_j} \frac{\partial p_{li}}{\partial \theta_q} - \frac{\delta_{li} - p_{li}}{p_{li}(1-p_{li})} \frac{\partial^2 p_{li}}{\partial \theta_j \partial \theta_q} \right\} \right]
\end{aligned}$$

Como $\delta_{li} \sim \text{Bernoulli}(p_{li})$ tem-se, utilizando a propriedade de linearidade da esperança, que:

$$I_{\theta_j \theta_q}(\boldsymbol{\theta}) = \sum_{i=1}^k \sum_{l \in R_i} \frac{1}{p_{li}(1-p_{li})} \frac{\partial p_{li}}{\partial \theta_j} \frac{\partial p_{li}}{\partial \theta_q}. \quad (2.11)$$

Dessa forma, para obter os componentes do vetor escore e da matriz de informação de Fisher, basta conhecer as derivadas parciais de $p_{li} = 1 - \exp \{-(1 + \lambda \eta_{li})^{1/\lambda}\}$ em relação a cada um dos parâmetros de interesse em $\boldsymbol{\theta}$, no caso geral $\lambda \neq 0$ e $\lambda \eta_{li} > -1$:

$$\frac{\partial p_{li}}{\partial \lambda} = \exp \{-(1 + \lambda \eta_{li})^{1/\lambda}\} (1 + \lambda \eta_{li})^{1/\lambda} \left[\frac{\eta_{li}}{\lambda(1 + \lambda \eta_{li})} - \frac{\log(1 + \lambda \eta_{li})}{\lambda^2} \right] \quad (2.12)$$

$$\frac{\partial p_{li}}{\partial \xi_i} = \exp \{-(1 + \lambda \eta_{li})^{1/\lambda}\} (1 + \lambda \eta_{li})^{1/\lambda-1}, \quad i = 1, \dots, k \quad (2.13)$$

$$\frac{\partial p_{li}}{\partial \gamma_j} = \exp \{-(1 + \lambda \eta_{li})^{1/\lambda}\} x_{lj} (1 + \lambda \eta_{li})^{1/\lambda-1}, \quad j = 1, \dots, p \quad (2.14)$$

Para a obtenção das expressões no caso particular $\lambda = 0$, basta tomar os limites das expressões acima para $\lambda \rightarrow 0$.

Como o interesse nesta etapa está no teste do parâmetro λ , a estatística escore assume a seguinte forma simplificada:

$$S_R = U_\lambda^2(\hat{\boldsymbol{\theta}}_0) I^{\lambda\lambda}(\hat{\boldsymbol{\theta}}_0),$$

em que $I^{\lambda\lambda}(\boldsymbol{\theta})$ é o elemento na posição $\lambda\lambda$ da inversa da matriz de informação de Fisher.

Utilizando-se as expressões dadas em (2.10) e (2.12), obtém-se a seguinte expressão geral para a função escore de λ :

$$U_\lambda(\boldsymbol{\theta}) = \sum_{i=1}^k \sum_{l \in R_i} \left\{ \left(\frac{\delta_{li}}{p_{li}} - 1 \right) (1 + \lambda \eta_{li})^{1/\lambda} \left[\frac{\eta_{li}}{\lambda(1 + \lambda \eta_{li})} - \frac{\log(1 + \lambda \eta_{li})}{\lambda^2} \right] \right\},$$

que possui as seguintes expressões avaliadas nas hipóteses de interesse $\lambda = 0$ e $\lambda = 1$:

$$U_{\lambda}(\boldsymbol{\theta}_0) = U_{\lambda}(0, \boldsymbol{\xi}_0^{\top}, \gamma_0^{\top}) = -\frac{1}{2} \sum_{i=1}^k \sum_{l \in R_i} \left\{ \left[\frac{\delta_{li}}{1 - \exp\{-\exp(\eta_{li})\}} - 1 \right] \eta_{li}^2 \exp(\eta_{li}) \right\}$$

$$U_{\lambda}(\boldsymbol{\theta}_0) = U_{\lambda}(1, \boldsymbol{\xi}_0^{\top}, \gamma_0^{\top}) = \sum_{i=1}^k \sum_{l \in R_i} \left\{ (1 + \eta_{li}) \left[\frac{\eta_{li}}{1 + \eta_{li}} - \log(1 + \eta_{li}) \right] \left[\frac{\delta_{li}}{1 - \exp\{-(1 + \eta_{li})\}} \right] \right\},$$

em que $\boldsymbol{\xi}_0$ e γ_0 são os valores dos parâmetros $\boldsymbol{\xi}$ e γ sob a hipótese nula.

Para o cálculo do componente $I^{\lambda\lambda}(\boldsymbol{\theta}_0)$, particiona-se a matriz $I(\boldsymbol{\theta}_0)$ em

$$I(\boldsymbol{\theta}_0) = \begin{bmatrix} I_{\lambda\lambda}(\boldsymbol{\theta}) & I_{\lambda\phi}(\boldsymbol{\theta}) \\ I_{\phi\lambda}(\boldsymbol{\theta}) & I_{\phi\phi}(\boldsymbol{\theta}) \end{bmatrix},$$

em que

$$I_{\lambda\phi}(\boldsymbol{\theta}_0) = I_{\phi\lambda}^{\top}(\boldsymbol{\theta}_0) = \begin{bmatrix} I_{\lambda\xi}(\boldsymbol{\theta}) & I_{\lambda\gamma}(\boldsymbol{\theta}) \end{bmatrix}$$

$$I_{\phi\phi} = \begin{bmatrix} I_{\xi\xi}(\boldsymbol{\theta}) & I_{\xi\gamma}(\boldsymbol{\theta}) \\ I_{\gamma\xi}(\boldsymbol{\theta}) & I_{\gamma\gamma}(\boldsymbol{\theta}) \end{bmatrix}.$$

Pode-se então obter a seguinte matriz inversa particionada

$$I^{-1}(\boldsymbol{\theta}_0) = \begin{bmatrix} I^{\lambda\lambda}(\boldsymbol{\theta}) & I^{\lambda\phi}(\boldsymbol{\theta}) \\ I^{\phi\lambda}(\boldsymbol{\theta}) & I^{\phi\phi}(\boldsymbol{\theta}) \end{bmatrix},$$

em que o componente $I^{\lambda\lambda}(\boldsymbol{\theta})$ utilizado no cálculo da estatística escore é dado por (Rao, 2009):

$$I^{\lambda\lambda}(\boldsymbol{\theta}) = I_{\lambda\lambda}(\boldsymbol{\theta}) - I_{\lambda\phi}(\boldsymbol{\theta}_0) I_{\phi\phi}^{-1}(\boldsymbol{\theta}) I_{\phi\lambda}(\boldsymbol{\theta}_0).$$

As expressões gerais das componentes da matriz de informação de Fisher são obtidas a partir da expressão (2.11) em conjunto com as relações (2.12), (2.13) e (2.14) e são omitidas aqui por apresentarem formulações muito extensas. São apresentadas a seguir, apenas as expressões das componentes da matriz de informação de Fisher avaliadas sob as hipóteses de interesse. Para $\lambda = 0$, tomou-se o limite para $\lambda \rightarrow 0$ das expressões gerais obtidas.

- Sob $\mathbb{H}_0 : \lambda = 0$

$$I_{\lambda\lambda}(\boldsymbol{\theta}_0) = -\frac{1}{4} \sum_{i=1}^k \sum_{l \in R_i} \frac{\exp(2\eta_{li})\eta_{li}^4}{1 - \exp\{\exp(\eta_{li})\}}$$

$$I_{\lambda\xi_i}(\boldsymbol{\theta}_0) = \frac{1}{2} \sum_{l \in R_i} \frac{\exp(2\eta_{li})\eta_{li}^2}{1 - \exp\{\exp(\eta_{li})\}}, \quad i = 1, \dots, k$$

$$I_{\lambda\gamma_j}(\boldsymbol{\theta}_0) = \frac{1}{2} \sum_{i=1}^k \sum_{l \in R_i} \frac{\exp(2\eta_{li})x_{lj}\eta_{li}^2}{1 - \exp\{\exp(\eta_{li})\}}, \quad j = 1, \dots, p$$

$$I_{\xi_i\xi_s}(\boldsymbol{\theta}_0) = \begin{cases} -\sum_{l \in R_i} \frac{\exp(2\eta_{li})}{1 - \exp\{\exp(\eta_{li})\}}, & i = s = 1, \dots, k \\ 0, & i \neq s \end{cases}$$

$$I_{\xi_i\gamma_j}(\boldsymbol{\theta}_0) = -\sum_{l \in R_i} \frac{\exp(2\eta_{li})x_{lj}}{1 - \exp\{\exp(\eta_{li})\}}$$

$$I_{\gamma_j\gamma_v}(\boldsymbol{\theta}_0) = -\sum_{i=1}^k \sum_{l \in R_i} \frac{\exp(2\eta_{li})x_{lj}x_{lv}}{1 - \exp\{\exp(\eta_{li})\}}, \quad j, v = 1, \dots, p$$

- Sob $\mathbb{H}_0 : \lambda = 1$

$$I_{\lambda\lambda}(\boldsymbol{\theta}_0) = -\sum_{i=1}^k \sum_{l \in R_i} \frac{[(1 + \eta_{li}) \log(1 + \eta_{li}) - \eta_{li}]^2}{1 - \exp(1 + \eta_{li})}$$

$$I_{\lambda\xi_i}(\boldsymbol{\theta}_0) = \sum_{l \in R_i} \frac{(1 + \eta_{li}) \log(1 + \eta_{li}) - \eta_{li}}{1 - \exp(1 + \eta_{li})}, \quad i = 1, \dots, k$$

$$I_{\lambda\gamma_j}(\boldsymbol{\theta}_0) = \sum_{i=1}^k \sum_{l \in R_i} \frac{x_{lj} [(1 + \eta_{li}) \log(1 + \eta_{li}) - \eta_{li}]}{1 - \exp(1 + \eta_{li})}, \quad j = 1, \dots, p$$

$$I_{\xi_i\xi_s}(\boldsymbol{\theta}_0) = \begin{cases} -\sum_{l \in R_i} \frac{1}{1 - \exp(1 + \eta_{li})}, & i = s = 1, \dots, k \\ 0, & i \neq s \end{cases}$$

$$I_{\xi_i\gamma_j}(\boldsymbol{\theta}_0) = -\sum_{l \in R_i} \frac{x_{lj}}{1 - \exp(1 + \eta_{li})}$$

$$I_{\gamma_j\gamma_v}(\boldsymbol{\theta}_0) = -\sum_{i=1}^k \sum_{l \in R_i} \frac{x_{lj}x_{lv}}{1 - \exp(1 + \eta_{li})}, \quad j, v = 1, \dots, p$$

A partir das expressões apresentadas, pode-se construir a estatística de teste em ambas as hipóteses de interesse.

2.3.4 Estudos de simulação

Nesta etapa, são apresentados os estudos de simulação realizados a fim de avaliar empiricamente as propriedades dos testes apresentados. Para tal, foram considerados 5 intervalos de tamanho 0,5 e apenas uma covariável indicadora de grupo, gerada a partir de uma distribuição Bernoulli com parâmetro 0,5. Os dados foram gerados com $\xi = (-0,8; -0,6; -0,4; -0,2; 0,2)^\top$ e $\beta = 0,5$, de acordo com a seguinte regra:

1. Para cada intervalo em que a observação estiver em risco, gerar uma Bernoulli com parâmetro dado em (2.9) representando a falha;
2. Para cada intervalo em que a observação estiver em risco, gerar uma Bernoulli com parâmetro p_c representando a censura;
3. Se a falha (censura) ocorrer antes da censura (falha), então esta observação será considerada falha (censura). No caso em que a falha e a censura ocorrem no mesmo intervalo, a observação será considerada falha com probabilidade 0,7;
4. Repetir os passos 1-3 até que obtenha uma amostra de tamanho desejado.

Sabe-se que a estatística escore converge assintoticamente para uma distribuição Qui-quadrado sob a hipótese nula (Cox e Hinkley, 1979). Sendo assim, foram geradas 1000 réplicas de amostras de tamanhos 50, 100, 500 e 1000 com as estruturas aditiva e multiplicativa e foram calculados os valores das estatísticas sob a hipótese nula. Estes valores foram então comparados com os quantis teóricos da distribuição Qui-quadrado com um grau de liberdade e são apresentados nos seguintes gráficos do tipo Quantil-Quantil:

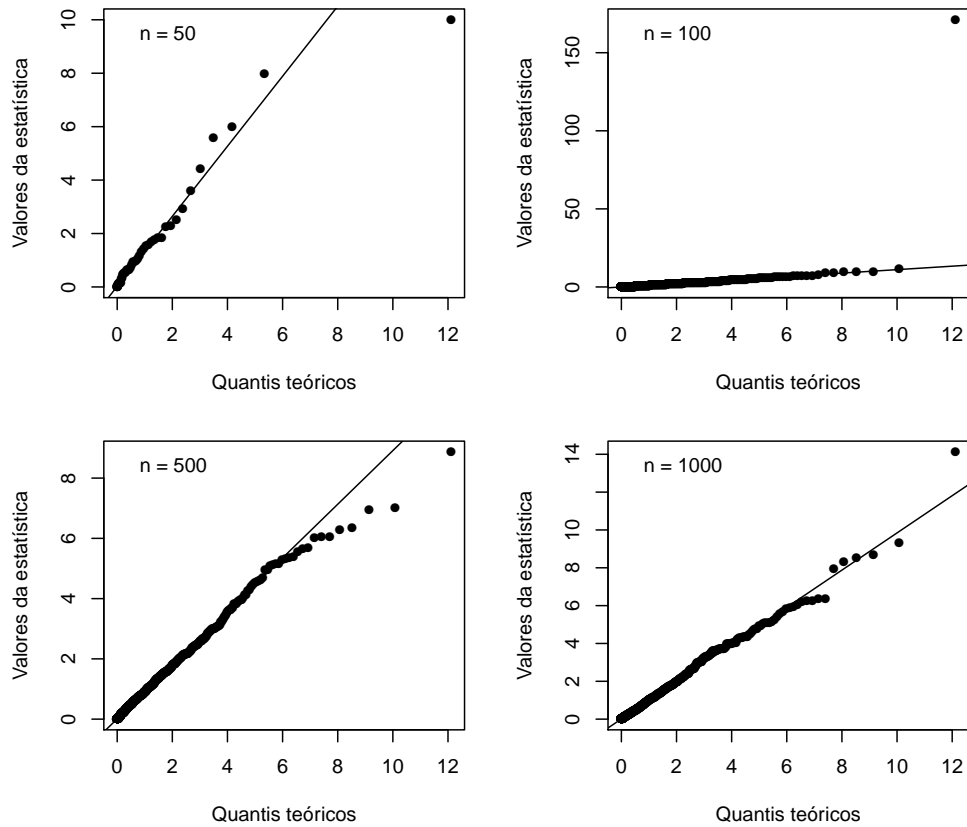


Figura 2.1: Gráfico do tipo QQ para os valores das estatísticas sob a hipótese nula para dados com estrutura multiplicativa.

A partir das Figuras 2.1 e 2.2, verifica-se que os valores das estatísticas score são próximas da distribuição Qui-quadrado com um grau de liberdade, tendo uma fuga acentuada apenas para os dados com estrutura aditiva de tamanho amostral igual a 50.

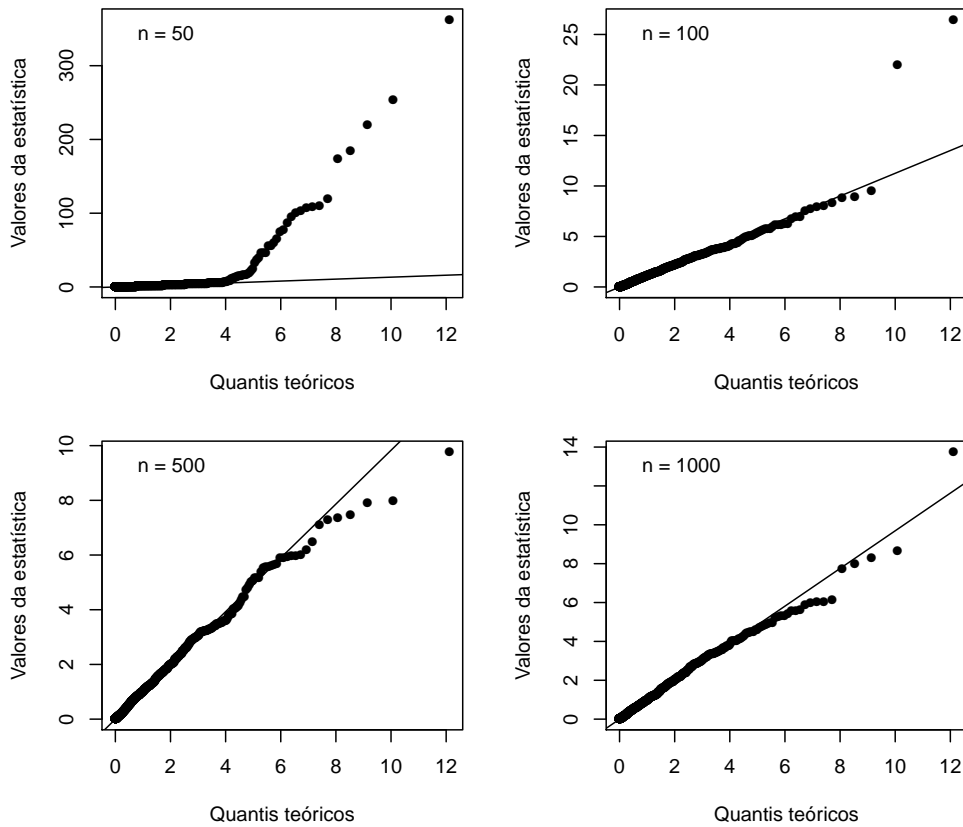


Figura 2.2: Gráfico do tipo QQ para os valores das estatísticas sob a hipótese nula para dados com estrutura aditiva.

Além da análise da distribuição da estatística score apresentada anteriormente, também foram avaliadas as taxas de rejeição (com nível de significância igual a 5%) para amostras de tamanhos 50, 100, 500 e 1000 e percentuais de censura iguais a 0, 25 e 50%, geradas com as estruturas aditiva e multiplicativa e submetidas aos dois testes de interesse. As taxas de rejeição em 1000 réplicas de cada uma das configurações são apresentadas na Tabela 2.1 a seguir:

Tabela 2.1: Taxas de rejeição dos testes escore para diferentes tamanhos amostrais (n) e percentuais de censura (p_c).

n	p_c	Teste $H_0 : \lambda = 0$ (Multiplicativo)		Teste $H_0 : \lambda = 1$ (Aditivo)	
		Dados Aditivos	Dados Multiplicativos	Dados Aditivos	Dados Multiplicativos
50	0	10,10%	5,60%	10,80%	17,30%
	0,25	8,20%	8,50%	30,40%	43,70%
	0,50	-	-	68,80%	79,00%
100	0	19,60%	5,10%	5,60%	7,80%
	0,25	14,10%	7,80%	12,00%	26,40%
	0,50	-	-	42,30%	64,60%
500	0	75,80%	6,60%	4,10%	13,00%
	0,25	60,60%	5,90%	6,10%	8,80%
	0,50	40,40%	5,00%	8,50%	14,00%
1000	0	96,70%	4,90%	4,60%	23,80%
	0,25	86,20%	5,70%	5,80%	12,40%
	0,50	73,40%	4,20%	5,70%	7,80%

Por meio da Tabela 2.1, verifica-se que o teste para riscos multiplicativos parece se comportar de maneira satisfatória, uma vez que as taxas de rejeição para dados com riscos multiplicativos são próximas de 5%, com valores mais discrepantes para tamanhos de amostra menores (50 e 100) e percentual de censura de 25%. Para dados com riscos aditivos, as taxas de rejeição se mostram crescentes à medida que se aumenta o tamanho amostral, sendo este crescimento menos acentuado, mas ainda sim satisfatório, para maiores percentuais de censura. Para as amostras de tamanho 50 e 100, não foi possível calcular as taxas de rejeição para o percentual de censura de 50%.

Quanto ao teste para riscos aditivos, pode-se observar que este não apresenta o mesmo comportamento que o teste para riscos multiplicativos, uma vez que as taxas de rejeição para dados com riscos aditivos se aproximam de 5% apenas para maiores tamanhos amostrais (500 e 1000) e os percentuais de censura parecem afetar de maneira mais expressiva nas amostras de menores tamanhos (50 e 100). Além disso, avaliando as taxas de rejeições para dados com riscos multiplicativos, verifica-se que as taxas não são tão altas quanto as encontradas no teste para riscos multiplicativos e também não parece haver uma tendência clara para diferentes tamanhos amostrais e percentuais de censura.

Por fim, foi realizada também uma análise da função poder para avaliar o comportamento dos testes para diferentes valores do parâmetro λ e de percentuais de censura (0, 15 e 30%). Para tal, foram gerados 1000 réplicas de amostras de tamanho 500 para cada valor de $\lambda \in [-0,5; 1,0]$

(com saltos de 0, 1) e foram calculadas as taxas de rejeição (com nível de significância de 5%) para ambos os testes desenvolvidos.

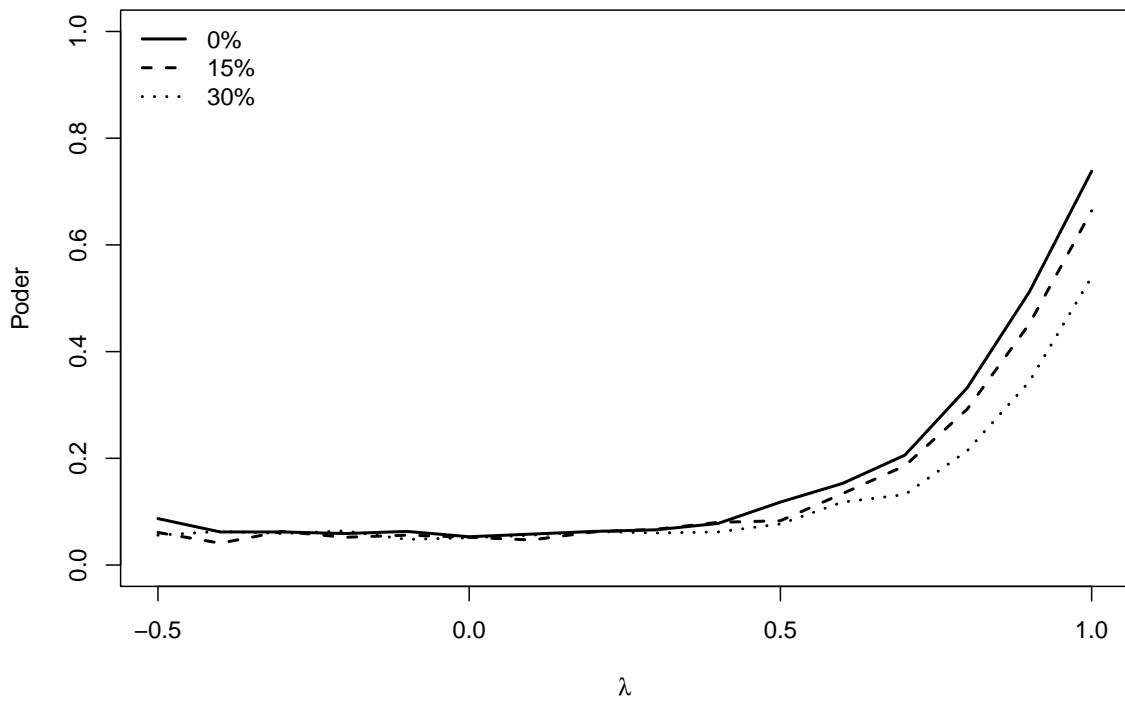


Figura 2.3: Gráfico de poder para o teste $H_0 : \lambda = 0$.

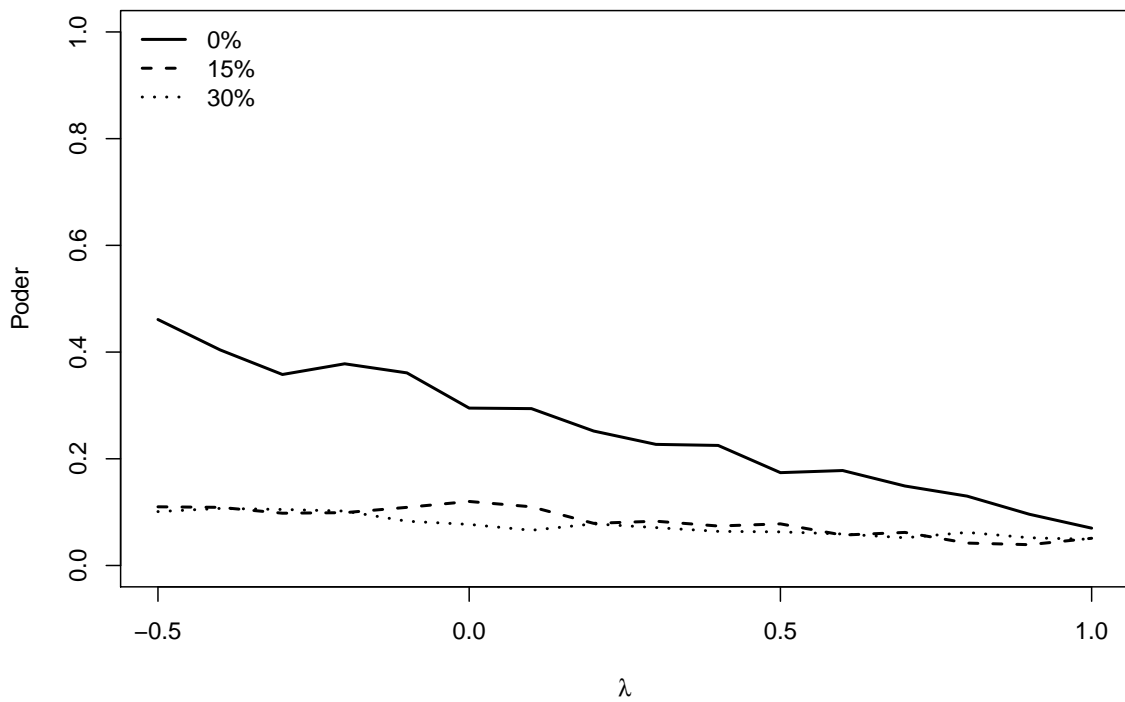


Figura 2.4: Gráfico de poder para o teste $H_0 : \lambda = 1$.

Por meio da Figura 2.3, verifica-se que as curvas de poder do teste para riscos multiplicativos ficam em torno de 5% para valores entre $-0,5$ e $0,5$ do parâmetro λ e que a partir do valor $0,5$, as curvas passam a crescer até próximo de 80% para $\lambda = 1$, sendo este crescimento menos acentuado à medida que se aumenta o percentual de censura da amostra.

Quanto à Figura 2.4, pode-se verificar novamente que o comportamento do teste para riscos aditivos difere do teste para riscos multiplicativos, uma vez que apenas a curva de amostras sem censura apresenta um aumento, com pico em torno de 45%, à medida que os valores do parâmetro de distanciam de 1. As demais curvas parecem apresentar apenas um ligeiro aumento, com valores entre 5 e 10%.

Capítulo 3

Modelos de riscos aditivos vs multiplicativos - efeitos variantes no tempo

Neste Capítulo, os principais modelos de riscos aditivos e multiplicativos que incorporam efeitos dependentes do tempo são discutidos com a abordagem de processos de contagem. Será dada uma maior ênfase para o modelo de riscos aditivos de [Aalen \(1980\)](#), uma vez que este modelo irá contribuir posteriormente na composição do modelo de Cox-Aalen, que é o principal interesse deste trabalho.

3.1 Modelo de Cox generalizado

O modelo de Cox discutido no Capítulo anterior assume que os riscos relativos são constantes ao longo do tempo. Entretanto, essa suposição não é adequada quando o efeito de pelo menos uma covariável varia no tempo, como por exemplo, tratamentos que perdem sua eficácia com o passar do tempo. Neste contexto, a generalização do modelo de Cox que admita efeitos dependentes do tempo pode ser interessante. Este modelo assume que o processo de intensidade tem a seguinte forma:

$$h(t; \mathbf{x}) = Y(t)h^*(t; \mathbf{x})$$
$$h^*(t; \mathbf{x}) = h_0(t) \exp\{\mathbf{x}(t)^\top \boldsymbol{\beta}(t)\},$$

em que $Y(t)$ é o indicador de risco no instante t , $\mathbf{x}(t) = (x_1(t), \dots, x_p(t))^\top$ é o vetor p -dimensional de covariáveis possivelmente dependentes do tempo e $\boldsymbol{\beta}(t) = (\beta_1(t), \dots, \beta_p(t))^\top$, seu respectivo vetor de parâmetros dependentes do tempo.

O modelo apresentado é estudado por diversos autores na literatura. [Brown \(1975\)](#), [Taulbee \(1979\)](#), [Stablein et al. \(1981\)](#) e [Moreau et al. \(1985\)](#) consideram formas simplificadas para $\boldsymbol{\beta}(t)$, de maneira que a dimensão do espaço paramétrico se mantenha finito. [Murphy e Sen \(1991\)](#) propõem um estimador baseado no método de *sieves*, particionando o tempo em intervalos e supondo que $\boldsymbol{\beta}(t)$ é constante dentro de cada uma das K_n partições, que devem ser especificadas previamente.

[Zucker e Karr \(1990\)](#) permitem que o espaço paramétrico do modelo seja de dimensão infinita, propondo uma técnica de estimação baseada na função de verossimilhança parcial penalizada. Esta verossimilhança é composta por duas parcelas, em que a primeira é referente à verossimilhança parcial do modelo de Cox usual, e a segunda, a uma função de penalidade para se obter estimativas suavizadas e com menores variâncias. Nesta metodologia porém, a escolha de parâmetros de suavização é necessária.

[Grambsch e Therneau \(1994\)](#) sugerem a estimação de $\boldsymbol{\beta}(t)$ por meio da suavização dos resíduos padronizados de Schoenfeld, que é somada posteriormente às estimativas dos coeficientes sob o modelo de Cox usual. [Winnett e Sasieni \(2003\)](#) utilizam a mesma ideia para estimar os parâmetros iterativamente, de maneira que a cada iteração se estime o coeficiente por meio da suavização dos resíduos de Schoenfeld e, a partir do modelo ajustado, se calcule os novos resíduos de Schoenfeld.

[Cai e Sun \(2003\)](#) propõem estimação dos parâmetros por meio da verossimilhança parcial local, em que se considera uma janela em torno de cada instante de tempo onde o coeficiente é aproximado por uma função linear. Abordagens Bayesianas do modelo são desenvolvidas em [Gamerman \(1991\)](#) e [Sargent \(1997\)](#).

3.2 Modelo de riscos aditivos de Aalen

O modelo proposto por [Aalen \(1980\)](#) assume que a intensidade do processo de contagem $N(t)$, $t \in [0, \tau]$ possui a seguinte forma:

$$\begin{aligned} h(t; \mathbf{x}(t)) &= Y(t)h^*(t; \mathbf{x}(t)) \\ h^*(t; \mathbf{x}(t)) &= \mathbf{x}(t)^\top \boldsymbol{\alpha}(t), \end{aligned} \tag{3.1}$$

em que $Y(t)$ é o indicador de risco no instante t , $\mathbf{x}(t) = (1, x_1(t), \dots, x_{p-1}(t))^\top$ é o vetor p -dimensional de covariáveis possivelmente dependentes do tempo e $\boldsymbol{\alpha}(t) = (\alpha_0(t), \alpha_1(t), \dots, \alpha_{p-1}(t))^\top$, seu respectivo vetor de parâmetros dependentes do tempo.

A estimação dos parâmetros $\boldsymbol{\alpha}(t)$ neste modelo é feita através do coeficiente acumulado

$$\begin{aligned} \mathbf{A}(t) &= (A_1(t), \dots, A_p(t))^\top \\ A_j(t) &= \int_0^t \alpha_j(u) du, \quad j = 1, \dots, p. \end{aligned}$$

Para tal, sejam $(N_i(t), Y_i(t), \mathbf{x}_i(t))$, $i = 1, \dots, n$ e defina os vetores n -dimensionais $\mathbf{N}(t)$, $\mathbf{h}(t)$, $\mathbf{M}(t)$ e a matriz $\mathbf{X}(t)$ de dimensão $n \times p$, dadas por:

$$\begin{aligned} \mathbf{N}(t) &= (N_1(t), \dots, N_n(t))^\top \\ \mathbf{M}(t) &= (M_1(t), \dots, M_n(t))^\top \\ \mathbf{h}(t) &= (h_1(t), \dots, h_n(t))^\top \\ \mathbf{X}(t) &= (Y_1(t)\mathbf{x}_1(t), \dots, Y_n(t)\mathbf{x}_n(t))^\top. \end{aligned}$$

Pela decomposição de Doob-Meyer, tem-se que

$$\begin{aligned} d\mathbf{N}(t) &= \mathbf{h}(t)dt + d\mathbf{M}(t) \\ &= \mathbf{X}(t)\boldsymbol{\alpha}(t)dt + d\mathbf{M}(t) \\ &= \mathbf{X}(t)d\mathbf{A}(t) + d\mathbf{M}(t). \end{aligned}$$

Como $d\mathbf{M}(t)$ são martingais de média zero, pode-se estimar $d\mathbf{A}(t)$ por técnicas de regressão linear multivariada. Para tal, define-se a inversa generalizada de $\mathbf{X}(t)$ como

$$\mathbf{X}^-(t) = \left[\mathbf{X}^\top(t)\mathbf{W}(t)\mathbf{X}(t) \right]^{-1} \mathbf{X}^\top(t)\mathbf{W}(t), \quad (3.2)$$

em que $\mathbf{W}(t) = \text{diag}\{W_1(t), \dots, W_n(t)\}$ é uma matriz diagonal $n \times n$ de pesos conhecidos. Estabelecendo que $\mathbf{X}^-(t)$ é zero quando não existir e definindo $J(t)$, que possui valor um se a inversa generalizada existir e zero caso contrário. A inversa generalizada deve ser tal que

$$\mathbf{X}^-(t)\mathbf{X}(t) = J(t)\mathbf{I}_p,$$

em que \mathbf{I}_p é uma matriz identidade de dimensão p .

Sendo assim, tem-se o seguinte estimador de mínimos quadrados generalizados (Huffer e McKeague, 1991):

$$d\widehat{\mathbf{A}}(t) = \mathbf{X}^{-}(t)d\mathbf{N}(t),$$

que pode ser escrito como

$$\begin{aligned}\widehat{\mathbf{A}}(t) &= \int_0^t \mathbf{X}^{-}(u)d\mathbf{N}(u) \\ &= \int_0^t \mathbf{X}^{-}(u) [\mathbf{X}(t)d\mathbf{A}(t) + d\mathbf{M}(t)] \\ &= \int_0^t J(u)d\mathbf{A}(u) + \int_0^t \mathbf{X}^{-}(u)d\mathbf{M}(u).\end{aligned}\tag{3.3}$$

Como $\int_0^t \mathbf{X}^{-}(u)d\mathbf{M}(u)$ também é um ruído de média zero, tem-se que o estimador dado em (3.3) é não viciado. Suas propriedades assintóticas são apresentadas no Teorema a seguir:

Teorema 1 *Se as seguintes condições de regularidade estiverem satisfeitas,*

1. $\sup_{t \in [0, \tau]} \mathbb{E} [Y_i(t)W_i^2(t)x_{ij}(t)x_{ik}(t)x_{il}(t)] < \infty$ para todo $j, k, l = 1, \dots, p$;
2. $\mathbf{r}_2(t) = \mathbb{E} [Y_i(t)W_i(t)\mathbf{x}_i^{\otimes 2}(t)]$ é não singular para todo $t \in [0, \tau]$.

tem-se, para $n \rightarrow \infty$, que

$$n^{1/2} [\widehat{\mathbf{A}}(t) - \mathbf{A}(t)] \xrightarrow{D} U$$

em $D[0, \tau]^p$, em que U é um processo Gaussiano martingal com função de covariância $\Phi(t)$, dada por

$$\begin{aligned}\Phi(t) &= \int_0^t \phi(u)du, \\ \phi(t) &= \mathbf{r}_2^{-1}(t)\mathbb{E} [Y_i(t)W_i^2(t)\mathbf{x}_i^{\otimes 2}(t)\mathbf{x}_i^{\top}(t)\boldsymbol{\alpha}(t)] \mathbf{r}_2^{-1}(t),\end{aligned}$$

que pode ser consistentemente estimada por

$$\widehat{\Phi}(t) = n \int_0^t \mathbf{X}^{-}(u)\text{diag} \{d\mathbf{N}(u)\} [\mathbf{X}^{-}(u)]^{\top}.$$

Sendo assim, pode-se construir o seguinte intervalo de confiança pontual com nível de confiança $(1 - \alpha)$ para $A_j(t)$:

$$\left[\widehat{A}_j(t) \pm n^{-1/2} c_{\alpha/2} \widehat{\Phi}_{jj}^{1/2}(t) \right],$$

em que $\widehat{\Phi}_{jj}^{1/2}(t)$ é o j -ésimo ($j = 1, \dots, p$) elemento de $\widehat{\Phi}(t)$ e $c_{\alpha/2}$ é o quantil de ordem $1 - \alpha/2$ da distribuição Normal padrão.

Além da estimação, deseja-se realizar também a inferência dos parâmetros por meio de testes de hipótese. Duas hipóteses de interesse neste modelo seriam

$$\mathbb{H}_{01} : \alpha_j(t) = 0$$

$$\mathbb{H}_{02} : \alpha_j(t) = \gamma,$$

para $t \in [0, \tau]$, que podem ser escritas em função do coeficiente acumulado como

$$\mathbb{H}_{01} : A_j(t) = 0$$

$$\mathbb{H}_{02} : A_j(t) = \gamma t.$$

A estatística de teste para a hipótese \mathbb{H}_{01} é desenvolvida em [Scheike \(2002\)](#), bem como as técnicas de reamostragem necessárias para a obtenção dos seus quantis. Esta estatística é baseada no desvio máximo da estimativa do coeficiente acumulado, ponderado pela variância de $n^{1/2} (\hat{\mathbf{A}}(t) - \mathbf{A}(t))$:

$$\sup_{t \in [0, \tau]} \left| \frac{n^{1/2} \hat{A}_j(t)}{\hat{\Psi}_{jj}^{1/2}(t)} \right|,$$

em que $\hat{\Psi}_{jj}(t)$ é o j -ésimo elemento da diagonal da matriz $\hat{\Psi}(t)$, dada por

$$\hat{\Psi}(t) = n^{-1} \sum_{i=1}^n \hat{\epsilon}_i^{\otimes 2}(t)$$

$$\hat{\epsilon}_i(t) = \int_0^t \left[n^{-1} \mathbf{X}^\top(u) \mathbf{X}(u) \right]^{-1} \mathbf{x}_i(u) d\hat{M}_i(u)$$

$$\hat{M}_i(t) = N_i(t) - \int_0^t Y_i(u) \mathbf{x}_i(u)^\top d\hat{\mathbf{A}}(t).$$

De maneira semelhante, as estatísticas para testar a hipótese \mathbb{H}_{02} são baseadas no desvio da estimativa do coeficiente acumulado com o seu valor sob a hipótese nula:

$$T_{2S} = n^{1/2} \sup_{t \in [0, \tau]} \left| \hat{A}_j(t) - \hat{A}_j(\tau) \frac{t}{\tau} \right|$$

$$T_{2I} = n \int_0^\tau \left(\hat{A}_j(t) - \hat{A}_j(\tau) \frac{t}{\tau} \right)^2 dt.$$

Para ambas as estatísticas, a obtenção dos quantis pode ser feita por técnicas de reamostragem. No entanto, [Martinussen e Scheike \(2006\)](#) apresentam uma abordagem utilizando a transformação de [Khmaladze \(1982\)](#) em ambas as estatísticas, de maneira que é possível obter seus

quantis em tabelas padrão. Os testes baseados nas estatísticas transformadas são denominados teste de Kolmogorov-Smirnov e de Cramér-von Mises e estão implementados no pacote estatístico `timereg`, do *software* R, utilizado na etapa de aplicação deste trabalho.

3.2.1 Modelo semiparamétrico de riscos aditivos

Uma vez testados os coeficientes quanto à dependência do tempo, pode-se chegar à conclusão de que algumas covariáveis podem ter seus efeitos considerados constantes no tempo. Sendo assim, [McKeague e Sasieni \(1994\)](#) propõem em seu artigo o chamado modelo semiparamétrico de riscos aditivos, que assume que o processo de intensidade possui a seguinte estrutura:

$$\begin{aligned} h(t; \mathbf{x}(t); \mathbf{z}(t)) &= Y(t)h^*(t; \mathbf{x}(t); \mathbf{z}(t)) \\ h^*(t; \mathbf{x}(t); \mathbf{z}(t)) &= \mathbf{x}(t)^\top \boldsymbol{\alpha}(t) + \mathbf{z}(t)^\top \boldsymbol{\gamma}, \end{aligned}$$

em que $Y(t)$, $\boldsymbol{\alpha}(t)$ e $\mathbf{x}(t)$ possuem as mesmas formulações apresentadas no modelo de Aalen, $\mathbf{z}(t) = (z_1(t), \dots, z_q(t))^\top$ é vetor q -dimensional de covariáveis possivelmente dependentes do tempo e $\boldsymbol{\gamma}$, seu respectivo vetor de parâmetros fixos no tempo.

Para a estimação dos parâmetros, além das notações definidas anteriormente, define-se a matriz $\mathbf{Z}(t)$ de dimensão $n \times q$, dada por:

$$\mathbf{Z}(t) = (Y_1(t)z_1(t), \dots, Y_n(t)z_n(t))^\top.$$

Utilizando novamente a decomposição de Doob-Meyer, tem-se que

$$\begin{aligned} d\mathbf{N}(t) &= \mathbf{h}(t)dt + d\mathbf{M}(t) \\ &= \mathbf{X}(t)d\mathbf{A}(t) + \mathbf{Z}(t)\boldsymbol{\gamma}dt + d\mathbf{M}(t). \end{aligned}$$

Como $d\mathbf{M}(t)$ são ruídos de média zero, os parâmetros $\mathbf{A}(t)$ e $\boldsymbol{\gamma}$ podem ser estimados por meio das seguintes equações de mínimos quadrados:

$$\mathbf{X}^\top(t)\mathbf{W}(t)[d\mathbf{N}(t) - \mathbf{h}(t)dt] = 0 \quad (3.4)$$

$$\int \mathbf{Z}^\top(t)\mathbf{W}(t)[d\mathbf{N}(t) - \mathbf{h}(t)dt] = 0, \quad (3.5)$$

que podem ser resolvidas sucessivamente da seguinte maneira:

1. Resolvendo a equação (3.4) para γ fixo, tem-se que

$$d\widehat{\mathbf{A}}(t) = \mathbf{X}^{-}(t) [d\mathbf{N}(t) - \mathbf{Z}(t)\gamma dt]; \quad (3.6)$$

2. Substituindo a solução (3.6) em (3.5) e resolvendo-a para γ , tem-se

$$\widehat{\gamma} = \left[\int_0^\tau \mathbf{Z}^\top(t) \mathbf{H}(t) \mathbf{Z}(t) dt \right]^{-1} \int_0^\tau \mathbf{Z}^\top(t) \mathbf{H}(t) d\mathbf{N}(t), \quad (3.7)$$

em que

$$\mathbf{H}(t) = \mathbf{W}(t) [\mathbf{I}_n - \mathbf{X}(t) \mathbf{X}^{-}(t)];$$

3. Por fim, substituindo a solução (3.7) no estimador obtido em (3.6) obtém-se:

$$\widehat{\mathbf{A}}(t) = \int_0^t \mathbf{X}^{-}(u) [d\mathbf{N}(u) - \mathbf{Z}(u) \widehat{\gamma} du]. \quad (3.8)$$

As propriedades assintóticas dos estimadores são apresentadas no seguinte Teorema:

Teorema 2 *Sejam $\tilde{\mathbf{x}}_i(t) = (\mathbf{x}_i(t), \mathbf{z}_i(t))$. Se as seguintes condições de regularidade estiverem satisfeitas,*

1. $\sup_{t \in [0, \tau]} \mathbb{E} [W_i(t) Y_i(t) \tilde{x}_{ij}(t) \tilde{x}_{ik}(t) \tilde{x}_{il}(t)] < \infty$ para todo $j, k, l = 1, \dots, p + q$;
2. $\mathbf{r}_2(t) = \mathbb{E} [Y_i(t) W_i(t) \tilde{\mathbf{x}}_i^{\otimes 2}(t)]$ é não singular para todo $t \in [0, \tau]$.

tem-se, para $n \rightarrow \infty$, que

$$n^{1/2} (\widehat{\gamma} - \gamma) \xrightarrow{D} N_q(\mathbf{0}, \Sigma),$$

$$n^{1/2} [\widehat{\mathbf{A}}(t) - \mathbf{A}(t)] \xrightarrow{D} U,$$

em $D[0, \tau]^p$, em que U é um processo Gaussiano de média zero com função de covariância $\Phi(t)$.

As variâncias assintóticas Σ e $\Phi(t)$ podem ser consistentemente estimadas por

$$\begin{aligned} \widehat{\Sigma} &= C_1^{-1} [M_1](\tau) C_1^{-1} \\ \widehat{\Phi}(t) &= [M_2](t) + P(t) C_1^{-1} [M_1](\tau) C_1^{-1} P^\top(t) \\ &\quad - [M_2, M_1](t) C_1^{-1} P^\top(t) - P(t) C_1^{-1} [M_1, M_2](t), \end{aligned}$$

em que

$$\begin{aligned}
C_1 &= n^{-1} \int_0^\tau \mathbf{Z}^\top(t) \mathbf{H}(t) \mathbf{Z}(t) dt \\
P(t) &= \int_0^t \mathbf{X}^-(u) \mathbf{Z}(u) du \\
[M_1](t) &= n^{-1} \int_0^t \mathbf{Z}^\top(u) \mathbf{H}(u) \text{diag} \{d\mathbf{N}(u)\} \mathbf{H}(u) \mathbf{Z}(u) \\
[M_2](t) &= n \int_0^t \mathbf{X}^-(u) \text{diag} \{d\mathbf{N}(u)\} [\mathbf{X}^-(u)]^\top \\
[M_1, M_2](t) &= \int_0^t \mathbf{Z}^\top(u) \mathbf{H}(u) \text{diag} \{d\mathbf{N}(u)\} [\mathbf{X}^-(u)]^\top.
\end{aligned}$$

3.2.2 Estimação da função de sobrevivência

Nesta última etapa, apresenta-se o processo de estimação da curva de sobrevivência para o modelo semiparamétrico de riscos aditivos. Vale ressaltar que o modelo aditivo de Aalen se enquadra como um caso especial quando $q = 0$. Para tal, assume-se um indivíduo com covariáveis \mathbf{x}_0 e z_0 fixos no tempo. Sabe-se que sua função de sobrevivência será da forma:

$$S_0(t) = \exp \left\{ -\mathbf{x}_0^\top \mathbf{A}(t) - z_0^\top t\gamma \right\},$$

que pode ser estimado por

$$\hat{S}_0(t) = \exp \left\{ -\mathbf{x}_0^\top \hat{\mathbf{A}}(t) - z_0^\top t\hat{\gamma} \right\}.$$

Também é possível obter uma banda de confiança aproximada com nível de confiança de $(1 - \alpha)$ utilizando técnicas de reamostragem descritas em [Martinussen e Scheike \(2006\)](#):

$$\left[\hat{S}_0(t) \pm C_\alpha n^{-1/2} \hat{Q}^{1/2}(t) \right],$$

em que

$$\begin{aligned}
\hat{Q}(t) &= \hat{S}_0^2(t) n^{-1} \sum_{i=1}^n \left\{ \mathbf{x}_0^\top \hat{\epsilon}_{3,i}(t) + t z_0^\top \hat{\epsilon}_{2,i} \right\}^2 \\
\hat{\epsilon}_{2,i} &= \int_0^\tau \left\{ \mathbf{z}_i^\top - \mathbf{Z}^\top(t) \mathbf{X}^\top(t) \left[\mathbf{X}^\top(t) \mathbf{X}(t) \right]^{-1} \mathbf{x}_i^\top \right\} d\hat{M}_i(t) \\
\hat{\epsilon}_{3,i}(t) &= \hat{\epsilon}_{4,i} - P(t) C_1^{-1} \hat{\epsilon}_{2,i} \\
\hat{\epsilon}_{4,i}(t) &= \int_0^t \left[n^{-1} \mathbf{X}^\top(u) \mathbf{X}(u) \right]^{-1} \mathbf{x}_i^\top d\hat{M}_i(u),
\end{aligned}$$

e C_α é o quantil de ordem $(1 - \alpha)$ da distribuição baseada em reamostragem de $n^{1/2}(\hat{S}_0(t) - S_0(t))$.

Capítulo 4

Modelo de Cox-Aalen

Nos Capítulos anteriores, foram apresentados os modelos de riscos aditivos e multiplicativos, discutindo-se também suas versões mais gerais, que acomodam efeitos dependentes do tempo. No entanto, a escolha da estrutura mais adequada nem sempre é trivial, uma vez algumas covariáveis podem ser melhor explicadas pela estrutura aditiva, e outras, pela estrutura multiplicativa.

Sendo assim, será apresentado neste Capítulo um estudo aprofundado do modelo de Cox-Aalen, que combina as estruturas aditiva e multiplicativa por meio do produto, além de permitir efeitos fixos e variantes no tempo.

4.1 Modelo

O chamado modelo de Cox suavizado proposto por [Dabrowska \(1997\)](#) possui a seguinte formulação para a intensidade:

$$\begin{aligned}h(t; \mathbf{x}(t); \mathbf{z}(t)) &= Y(t)h^*(t; \mathbf{x}(t); \mathbf{z}(t)) \\h^*(t; \mathbf{x}(t); \mathbf{z}(t)) &= h_0(t, \mathbf{x}(t)) \exp\{\mathbf{z}(t)^\top \boldsymbol{\beta}\},\end{aligned}$$

em que $Y(t)$ é o indicador de risco no instante t , $h_0(t, \mathbf{x}(t))$ é a função taxa de falha basal, caracterizada pelo vetor $(p - 1)$ -dimensional de covariáveis basais possivelmente dependentes do tempo $\mathbf{x}(t) = (x_1(t), \dots, x_{p-1}(t))^\top$, $\mathbf{z}(t) = (z_1(t), \dots, z_q(t))^\top$ é o vetor q -dimensional de covariáveis possivelmente dependentes do tempo e $\boldsymbol{\beta}$, o seu respectivo vetor de parâmetros.

A estimação deste modelo é realizada por técnicas baseadas em verossimilhança parcial suavizada, fazendo com que a escolha de um parâmetro de suavização seja necessária. Além disso, algumas adaptações são necessárias caso se deseje incluir covariáveis basais não contínuas.

Utilizando-se a expansão de Taylor de primeira ordem da função taxa de falha basal em torno do ponto $\mathbf{x}(t) = \mathbf{0}$, tem-se

$$h_0(t, \mathbf{x}(t)) \approx h_0(t, \mathbf{0}) + \mathbf{x}(t)^\top \mathbf{h}'_0(t, \mathbf{0}),$$

em que $h_0(t, \mathbf{0})$ representa a taxa de falha basal para $\mathbf{x}(t) = \mathbf{0}$ e $\mathbf{h}'_0(t, \mathbf{0})$, o vetor de dimensão $p - 1$ de funções variantes no tempo. Dessa forma, tem-se que a taxa de falha basal do modelo de Cox suavizado pode ser aproximado pelo modelo de riscos aditivos de Aalen (3.1). Sendo assim, Scheike e Zhang (2002) propõem um modelo, denominado Cox-Aalen, em que o termo basal do modelo de Cox suavizado é substituído pelo modelo de riscos aditivos de Aalen:

$$h(t; \mathbf{x}(t); \mathbf{z}(t)) = Y(t) \left[\mathbf{x}(t)^\top \boldsymbol{\alpha}(t) \right] \exp\{\mathbf{z}(t)^\top \boldsymbol{\beta}\}. \quad (4.1)$$

Este modelo, apesar de ser menos geral que o modelo proposto por Dabrowska, apresenta processos de estimação mais simples e comporta um número maior de covariáveis (inclusive discretas) no termo não paramétrico. A seguir são discutidos os processos de estimação dos parâmetros e suas propriedades inferenciais, bem como as técnicas de diagnósticos desenvolvidas para o modelo de Cox-Aalen.

4.2 Estimação

Sejam $(N_i(t), Y_i(t), \mathbf{x}_i(t), \mathbf{z}_i(t))$, $i = 1, \dots, n$ e considere os vetores n -dimensionais $\mathbf{N}(t)$, $\mathbf{M}(t)$ e $\mathbf{h}(t)$ com as mesmas formulações apresentadas na seção 3.2, em que $h_i(t)$ tem a estrutura dada em (4.1). Além disso, defina as matrizes $\mathbf{Y}(\boldsymbol{\beta}, t)$ e $\mathbf{Z}(t)$ de dimensões $n \times p$ e $n \times q$, tais que

$$\begin{aligned} \mathbf{Y}(\boldsymbol{\beta}, t) &= \left(Y_1(t) \exp\{\mathbf{z}_1(t)^\top \boldsymbol{\beta}\} \mathbf{x}_1(t), \dots, Y_n(t) \exp\{\mathbf{z}_n(t)^\top \boldsymbol{\beta}\} \mathbf{x}_n(t) \right)^\top \\ \mathbf{Z}(t) &= (\mathbf{z}_1(t), \dots, \mathbf{z}_n(t))^\top. \end{aligned}$$

A partir da estrutura apresentada, deseja-se estimar os parâmetros $\mathbf{A}(t) = \int_0^t \boldsymbol{\alpha}(u) du$ e $\boldsymbol{\beta}$. Nota-se, primeiramente, que para $\boldsymbol{\beta}$ conhecido pode-se obter o estimador de $\mathbf{A}(t)$ de acordo com o estimador de Huffer e McKeague (1991), construído a partir da decomposição de Doob-Meyer

$$\begin{aligned} d\mathbf{N}(t) &= \mathbf{h}(t)dt + d\mathbf{M}(t) \\ &= \mathbf{Y}(\boldsymbol{\beta}, t)d\mathbf{A}(t) + d\mathbf{M}(t) \end{aligned}$$

Como $dM(t)$ são martingais de média zero, pode-se estimar $dA(t)$ por técnicas de regressão linear multivariada. Para tal, define-se a inversa generalizada de $Y(\beta, t)$ como

$$Y^-(\beta, t) = \left[Y(\beta, t)^\top W(t) Y(\beta, t) \right]^{-1} Y(\beta, t)^\top W(t), \quad (4.2)$$

em que $W(t) = \text{diag} \{W_1(t), \dots, W_n(t)\}$ é uma matriz diagonal $n \times n$ de pesos conhecidos. Estabelecendo que $Y^-(\beta, t)$ é zero quando não existir e definindo $J(t)$, que possui o valor um se a inversa generalizada existir e zero caso contrário. A inversa generalizada deve ser tal que

$$Y^-(\beta, t) Y(\beta, t) = J(t) \mathbf{I}_p,$$

em que \mathbf{I}_p é uma matriz identidade de dimensão p .

Sendo assim, tem-se o seguinte estimador de mínimos quadrados ponderados de $A(t)$

$$\hat{A}(\beta, t) = \int_0^t Y^-(\beta, u) dN(u). \quad (4.3)$$

A estimação também pode ser dada por meio de equações de estimação. Para tal, constrói-se a função de verossimilhança de acordo com [Andersen et al. \(1993\)](#):

$$\begin{aligned} & \prod_{i=1}^n \prod_{t \leq \tau} (h_i(t) dt)^{dN_i(t)} \exp \left(- \int_0^\tau h_i(t) dt \right) \\ &= \prod_{i=1}^n \prod_{t \leq \tau} \left[\left(Y_i(t) \exp \left\{ z_i(t)^\top \beta \right\} x_i(t)^\top dA(t) \right)^{dN_i(t)} \right] \\ & \times \exp \left(- \sum_{i=1}^n \int_0^\tau Y_i(t) x_i(t)^\top \alpha(t) \exp \left\{ z_i(t)^\top \beta \right\} dt \right). \end{aligned}$$

Assim, a função de log-verossimilhança será dada por

$$\begin{aligned} & \sum_{i=1}^n \int_0^\tau dN_i(t) \log \left(Y_i(t) \exp \left\{ z_i(t)^\top \beta \right\} x_i(t)^\top dA(t) \right) \\ & - \sum_{i=1}^n \int_0^\tau Y_i(t) \exp \left\{ z_i(t)^\top \beta \right\} x_i(t)^\top dA(t), \end{aligned}$$

da qual pode-se obter as seguintes funções escore:

$$\begin{aligned} U(\beta) &= U(\beta, \tau) = \int_0^\tau Z^\top(t) [dN(t) - Y(\beta, t) dA(t)] \\ U(\alpha) &= Y(\beta, t)^\top \text{diag} \{1/h_i(t)\} [dN(t) - Y(\beta, t) dA(t)]. \end{aligned}$$

A partir das equações de estimação apresentadas, observa-se que o estimador de máxima verossimilhança para $\mathbf{A}(t)$ possui a mesma estrutura que o estimador de mínimos quadrados ponderados apresentado em (4.3) com pesos $W_i(t) = 1/h_i(t) = Y_i(t) \exp(-\mathbf{z}_i(t)^\top \boldsymbol{\beta}) / \mathbf{x}_i(t)^\top \boldsymbol{\alpha}(t)$. Sendo assim, utiliza-se o estimador (4.3) adotando-se pesos $w_i(t) = Y_i(t) \exp(-\mathbf{z}_i(t)^\top \boldsymbol{\beta}) / g_i(t)$, em que $g_i(t)$ é uma função conhecida que não depende do parâmetro $\boldsymbol{\beta}$. Dessa forma, no caso particular em que $g_i(t) = \mathbf{x}_i(t)^\top \boldsymbol{\alpha}(t)$, obtém-se o estimador de máxima verossimilhança.

Inserindo o estimador dado em (4.3) na equação de estimação de $\boldsymbol{\beta}$ tem-se

$$U(\boldsymbol{\beta}, \tau) = \int_0^\tau \mathbf{Z}^\top(t) [\mathbf{I}_n - \mathbf{Y}(\boldsymbol{\beta}, t)\mathbf{Y}^{-}(\boldsymbol{\beta}, t)] d\mathbf{N}(t).$$

O processo de estimação será composto por duas etapas. Na primeira, obtém-se os valores iniciais $\hat{\mathbf{A}}^{(0)}(t)$ e $\hat{\boldsymbol{\beta}}^{(0)}$ tomando-se pesos tais que $g_i(t) = 1$, estimando $\hat{\boldsymbol{\beta}}^{(0)}$ de maneira que $U(\hat{\boldsymbol{\beta}}^{(0)}, \tau) = \mathbf{0}$ e, posteriormente, substituindo o valor encontrado em $\hat{\mathbf{A}}(\hat{\boldsymbol{\beta}}^{(0)}, t)$. A partir dos valores iniciais obtidos na primeira etapa, repete-se os passos (1) – (3) apresentados a seguir até a convergência.

1. Estimar pesos $\hat{w}_i^{(j)}(t)$ de máxima verossimilhança tais que $g_i^{(j)}(t) = \mathbf{x}_i(t)^\top \hat{\boldsymbol{\alpha}}^{(j)}(t)$;
2. Estimar $\hat{\boldsymbol{\beta}}^{(j)}$ tal que $U(\hat{\boldsymbol{\beta}}^{(j)}, \tau) = \mathbf{0}$
3. Calcular $\hat{\mathbf{A}}(\hat{\boldsymbol{\beta}}^{(j)}, t) = \int_0^t \mathbf{Y}^{-}(\hat{\boldsymbol{\beta}}^{(j)}, u) d\mathbf{N}(u)$.

4.3 Propriedades assintóticas e inferenciais

Nesta etapa são apresentados os principais resultados assintóticos, construídos a partir de argumentos baseados em expansão da série de Talylor em torno no verdadeiro valor do parâmetro. Inicialmente são apresentados os resultados para os riscos relativos $\boldsymbol{\beta}$ e, posteriormente, para os coeficientes acumulados $\mathbf{A}(t)$. Para maiores detalhes, sugere-se a leitura de [Martinussen e Scheike \(2006\)](#).

4.3.1 Para os riscos relativos β

Utilizando a expansão de Taylor de primeira ordem da função escore $U(\beta, \tau)$ em torno do verdadeiro valor β_0 , tem-se

$$\begin{aligned} U(\hat{\beta}, \tau) &\approx U(\beta_0, \tau) + U'(\beta_0, \tau)(\hat{\beta} - \beta_0) \\ &= U(\beta_0, \tau) - I(\beta_0, \tau)(\hat{\beta} - \beta_0) \\ \Leftrightarrow \hat{\beta} - \beta_0 &\approx I^{-1}(\beta_0, \tau)U(\beta_0, \tau) \\ \Leftrightarrow \sqrt{n}(\hat{\beta} - \beta_0) &\approx \left[\frac{1}{n}I(\beta_0, \tau) \right]^{-1} \left[\frac{1}{\sqrt{n}}U(\beta_0, \tau) \right]. \end{aligned}$$

Para demonstrar o Teorema 3 a seguir, é necessário mostrar 1 - Normalidade assintótica de $n^{-1/2}U(\beta_0, \tau)$ baseada no Teorema Limite Central para martingais; 2 - Consistência de $n^{-1}I(\beta_0, \tau)$ e 3 - Consistência de $\hat{\beta}$. Estas verificações são feitas no artigo de [Scheike e Zhang \(2002\)](#).

Teorema 3 *Sob certas condições de regularidade, tem-se que $\sqrt{n}(\hat{\beta} - \beta_0)$ converge para uma distribuição normal multivariada com vetor de médias iguais a zero e matriz de variâncias que pode ser consistentemente estimada por*

$$\hat{\Sigma} = nI^{-1}(\hat{\beta}, \tau) \left[U(\hat{\beta}) \right] (\tau) I^{-1}(\hat{\beta}, \tau),$$

em que

$$\begin{aligned} [U(\beta)](\tau) &= \int_0^\tau \mathbf{Z}^\top(t) [\mathbf{I}_n - \mathbf{Y}(\beta, t)\mathbf{Y}^{-}(\beta, t)] \text{diag} \{d\mathbf{N}(t)\} [\mathbf{I}_n - \mathbf{Y}(\beta, t)\mathbf{Y}^{-}(\beta, t)]^\top \mathbf{Z}(t). \\ I(\beta, \tau) &= -\frac{\partial}{\partial \beta} U(\beta, \tau) \\ &= \int_0^\tau \mathbf{Z}^\top(t) \text{diag} \{ \mathbf{Y}(\beta, t)\mathbf{Y}^{-}(\beta, t)d\mathbf{N}(t) \} \mathbf{Z}(t) \\ &\quad - \int_0^\tau \mathbf{Z}^\top(t)\mathbf{Y}(\beta, t)\mathbf{Y}^{-}(\beta, t)\text{diag} \{ \mathbf{Y}(\beta, t)\mathbf{Y}^{-}(\beta, t)d\mathbf{N}(t) \} \mathbf{Z}(t). \end{aligned}$$

4.3.2 Para os coeficientes acumulados $A(t)$

Foi visto anteriormente que o estimador para $A(t)$ é dado por:

$$\begin{aligned} \hat{A}(\hat{\beta}, t) &= \int_0^t \mathbf{Y}^{-}(\hat{\beta}, u)d\mathbf{N}(u) \\ &= \int_0^t \mathbf{Y}^{-}(\hat{\beta}, u)d\mathbf{H}(u) + \int_0^t \mathbf{Y}^{-}(\hat{\beta}, u)d\mathbf{M}(u). \end{aligned}$$

Para β fixo, denota-se o compensador de $\hat{\mathbf{A}}(\beta, t)$ por $\mathbf{A}^*(\beta, t)$. Tem-se então que

$$\hat{\mathbf{A}}(\hat{\beta}, t) - \mathbf{A}(t) = \left[\hat{\mathbf{A}}(\hat{\beta}, t) - \hat{\mathbf{A}}(\beta_0, t) \right] + \left[\hat{\mathbf{A}}(\beta_0, t) - \mathbf{A}^*(\beta_0, t) \right] + [\mathbf{A}^*(\beta_0, t) - \mathbf{A}(t)].$$

A segunda parcela do lado direito da equação resulta em $\int_0^t \mathbf{Y}^-(\beta_0, u) d\mathbf{M}(u)$. Além disso, utilizando a expansão de Taylor na primeira parcela, tem-se que

$$\hat{\mathbf{A}}(\hat{\beta}, t) - \mathbf{A}(t) \approx (\hat{\beta} - \beta_0)^\top \int_0^t \frac{\partial}{\partial \beta} \mathbf{Y}^-(\beta_0, u) d\mathbf{N}(u) + \int_0^t \mathbf{Y}^-(\beta_0, u) d\mathbf{M}(u) + [\mathbf{A}^*(\beta_0, t) - \mathbf{A}(t)].$$

Assim, para demonstrar o Teorema 4 a seguir, basta mostrar que 1 - $H(\beta, t) = \int_0^t \partial/\partial \beta \mathbf{Y}^-(\beta_0, u) d\mathbf{N}(u)$ converge uniformemente em probabilidade; 2 - os termos martingais convergem conjuntamente em distribuição e 3 - $[\mathbf{A}^*(\beta_0, t) - \mathbf{A}(t)]$ converge uniformemente em probabilidade para zero. Estas verificações podem ser vistas em [Scheike e Zhang \(2002\)](#).

Teorema 4 *Sob certas condições de regularidade, tem-se que $\sqrt{n} \left(\hat{\mathbf{A}}(\hat{\beta}, t) - \mathbf{A}(t) \right)$ converge em distribuição para um processo Gaussiano com função de variância $\Phi(t)$, que pode ser consistentemente estimada por*

$$\hat{\Phi}(t) = n \left(\frac{1}{n} \mathbf{H}(\hat{\beta}, t)^\top \hat{\Sigma} \mathbf{H}(\hat{\beta}, t) + [\mathbf{M}_A](t) + \mathbf{C}(t) \right),$$

em que $\mathbf{C}(t)$ é dado por

$$\mathbf{C}(t) = \mathbf{H}(\hat{\beta}, t)^\top \mathbf{I}^{-1}(\hat{\beta}, \tau) [\mathbf{U}, \mathbf{M}_A](t) + [\mathbf{M}_A, \mathbf{U}](t) \mathbf{I}^{-1}(\hat{\beta}, \tau) \mathbf{H}(\hat{\beta}, t)$$

com $[\mathbf{U}, \mathbf{M}_A](t)$ e $[\mathbf{M}_A](t)$ estimados, respectivamente, por

$$\int_0^\tau \mathbf{Z}^\top(t) \left[\mathbf{I}_n - \mathbf{Y}(\beta, t) \mathbf{Y}^-(\beta, t) \right] \text{diag} \{d\mathbf{N}(t)\} \left(\mathbf{Y}^-(\beta, t) \right)^\top \\ \int_0^\tau \mathbf{Y}^-(\beta, t) \text{diag} \{d\mathbf{N}(t)\} \left(\mathbf{Y}^-(\beta, t) \right)^\top.$$

Observação: se forem utilizados os pesos de máxima verossimilhança, o termo $\mathbf{C}(t)$ será assintoticamente nulo.

A partir dos resultados apresentados, pode-se testar $H_0 : \beta = \beta_0$ a partir de estatísticas de Wald, Razão de verossimilhanças e Escore. Para os riscos acumulados, propõem-se as seguintes

estatísticas para testar a significância e a dependência do parâmetro no tempo, respectivamente:

$$\begin{aligned} & \sup_{t \in [0, \tau]} |\hat{A}_j(t)|, \\ & \sup_{t \in [0, \tau]} \left| \hat{A}_j(t) - \frac{\hat{A}_j(\tau)}{\tau} \right|, \end{aligned}$$

em que $\hat{A}_j(t)$ representa o j -ésimo componente de $\hat{\mathbf{A}}(t)$. Em ambos os casos, os níveis descritivos devem ser obtidos a partir de técnicas de simulação. Uma possível abordagem de simulação é discutida em [Lin et al. \(1993\)](#).

4.4 Técnicas de diagnóstico

Nesta seção discute-se as principais técnicas de diagnóstico para avaliar a adequabilidade do modelo de Cox-Aalen. Para tal, utilizam-se os seguintes resíduos martingais:

$$\begin{aligned} M_i(t) &= N_i(t) - H_i(t) \\ &= N_i(t) - \int_0^t Y_i(s) \exp\{\mathbf{z}_i(s)^\top \boldsymbol{\beta}\} \mathbf{x}_i(s)^\top d\mathbf{A}(s). \end{aligned}$$

Dessa forma, utilizando-se a relação (4.3), pode-se estimar o vetor de resíduos martingais por

$$\begin{aligned} \hat{\mathbf{M}}(t) &= \mathbf{N}(t) - \int_0^t \mathbf{Y}(\hat{\boldsymbol{\beta}}, s) \mathbf{Y}^{-}(\hat{\boldsymbol{\beta}}, s) d\mathbf{N}(s) \\ &= \int_0^t [\mathbf{I}_n - \mathbf{Y}(\hat{\boldsymbol{\beta}}, s) \mathbf{Y}^{-}(\hat{\boldsymbol{\beta}}, s)] d\mathbf{N}(s). \end{aligned}$$

Sob o modelo de Cox-Aalen, tem-se que os resíduos martingais serão assintoticamente equivalentes a um processo de média zero. Dessa forma, seu comportamento pode ser investigado de diversas maneiras. Uma possível maneira discutida em [Martinussen e Scheike \(2006\)](#) é a soma dos resíduos de acordo com os m níveis das covariáveis. Para tal, define-se a seguinte matriz $n \times m$

$$\mathbf{K}(t) = (\mathbf{K}_1(t), \dots, \mathbf{K}_n(t))^\top,$$

e define-se o processo residual acumulado como

$$n^{-1/2} \mathbf{M}_K(t) = \int_0^t \mathbf{K}^\top(s) d\hat{\mathbf{M}}(t),$$

que pode ser avaliado graficamente ao longo do tempo. Além disso, pode-se calcular a estatística

$$\sup_{t \in [0, \tau]} |M_{Kj}(t)|, \quad j = 1, \dots, m,$$

cuja distribuição deve ser aproximada por meio de técnicas de reamostragem.

Para verificar a suposição de proporcionalidade das covariáveis incluídas na parcela multiplicativa do modelo, considera-se o processo de escore acumulado $n^{-1/2}U(\hat{\beta}, t)$, bem como a estatística de teste

$$\sup_{t \in [0, \tau]} |U_j(\hat{\beta}, t)|, \quad j = 1, \dots, q.$$

4.5 Estimação da função de sobrevivência

Nesta etapa, considera-se um indivíduo com vetores de covariáveis \mathbf{x}_0 e \mathbf{z}_0 fixos no tempo, por simplicidade. Tem-se então que a função taxa de falha e a função de sobrevivência são dadas por

$$\begin{aligned} h^*(t) &= \mathbf{x}_0^\top \boldsymbol{\alpha}(t) \exp(\mathbf{z}_0^\top \boldsymbol{\beta}) \\ S_0(t) &= S_0(t; \mathbf{A}, \boldsymbol{\beta}) = \exp \left\{ -\mathbf{x}_0^\top \mathbf{A}(t) \exp(\mathbf{z}_0^\top \boldsymbol{\beta}) \right\}, \end{aligned}$$

em que a função de sobrevivência é estimada por

$$\hat{S}_0(t) = \hat{S}_0(t; \hat{\mathbf{A}}, \hat{\boldsymbol{\beta}}) = \exp \left\{ -\mathbf{x}_0^\top \hat{\mathbf{A}}(t) \exp(\mathbf{z}_0^\top \hat{\boldsymbol{\beta}}) \right\}.$$

Utilizando a expansão de Taylor de primeira ordem em torno do verdadeiro valor dos parâmetros, tem-se que

$$\begin{aligned} \hat{S}_0(t; \hat{\mathbf{A}}, \hat{\boldsymbol{\beta}}) - S_0(t; \mathbf{A}, \boldsymbol{\beta}) &\approx -S_0(t; \mathbf{A}, \boldsymbol{\beta}) \begin{bmatrix} \exp(\mathbf{z}_0^\top \boldsymbol{\beta}) \mathbf{x}_0^\top & \mathbf{x}_0^\top \mathbf{A}(t) \exp(\mathbf{z}_0^\top \boldsymbol{\beta}) \mathbf{z}_0^\top \end{bmatrix} \begin{bmatrix} \hat{\mathbf{A}}(t) - \mathbf{A}(t) \\ \hat{\boldsymbol{\beta}} - \boldsymbol{\beta} \end{bmatrix} \\ &\approx -S_0(t; \mathbf{A}, \boldsymbol{\beta}) \left[\exp(\mathbf{z}_0^\top \boldsymbol{\beta}) \mathbf{x}_0^\top (\hat{\mathbf{A}}(t) - \mathbf{A}(t)) + \mathbf{x}_0^\top \mathbf{A}(t) \exp(\mathbf{z}_0^\top \boldsymbol{\beta}) \mathbf{z}_0^\top (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \right]. \end{aligned}$$

Utilizando as propriedades assintóticas de $\hat{\mathbf{A}}$ e $\hat{\boldsymbol{\beta}}$ vistas anteriormente em conjunto com o método delta funcional, tem-se que $n^{1/2} \left(\hat{S}_0(t; \hat{\mathbf{A}}, \hat{\boldsymbol{\beta}}) - S_0(t; \mathbf{A}, \boldsymbol{\beta}) \right)$ converge para um processo Gaussiano de média zero e função de variância Q .

Pode-se demonstrar que, sob certas condições de regularidade, $n^{1/2} \left(\hat{S}_0(t; \hat{\mathbf{A}}, \hat{\boldsymbol{\beta}}) - S_0(t; \mathbf{A}, \boldsymbol{\beta}) \right)$

tem a mesma distribuição assintótica de

$$\Delta_S(t) = S_0(t)n^{1/2} \sum_{i=1}^n \hat{\epsilon}_{4i}(t)G_i$$

em que G_1, \dots, G_n são variáveis aleatórias independentes com distribuição Normal padrão e

$$\begin{aligned} \hat{\epsilon}_{4i}(t) &= \exp(\mathbf{z}_0^\top \hat{\boldsymbol{\beta}}) \mathbf{x}_0^\top \hat{\epsilon}_{2i}(t) + \mathbf{x}_0^\top \hat{\mathbf{A}}(t) \exp(\mathbf{z}_0^\top \hat{\boldsymbol{\beta}}) \mathbf{z}_0^\top \mathbf{I}^{-1}(\hat{\boldsymbol{\beta}}, \tau) \hat{\epsilon}_{1i}(t) \\ \hat{\epsilon}_{2i}(t) &= \hat{\epsilon}_{3i}(t) + H^\top(\hat{\boldsymbol{\beta}}, t) \mathbf{I}^{-1}(\hat{\boldsymbol{\beta}}, \tau) \hat{\epsilon}_{1i}(t) \\ \hat{\epsilon}_{3i}(t) &= \int_0^t \left[\mathbf{Y}^\top(\hat{\boldsymbol{\beta}}, s) \mathbf{W}(s) \mathbf{Y}(\hat{\boldsymbol{\beta}}, s) \right]^{-1} \mathbf{x}_i^\top(s) d\hat{M}_i(s) \\ \hat{\epsilon}_{1i}(t) &= \int_0^t \left\{ \mathbf{z}_i^\top(s) - \mathbf{Z}^\top(s) \mathbf{Y}^\top(\hat{\boldsymbol{\beta}}, s) \left[\mathbf{Y}^\top(\hat{\boldsymbol{\beta}}, s) \mathbf{W}(s) \mathbf{Y}(\hat{\boldsymbol{\beta}}, s) \right]^{-1} \right\} d\hat{M}_i(s). \end{aligned}$$

Além disso, tem-se que

$$\hat{Q}(t) = n \hat{S}_0^2(t) \sum_{i=1}^n \hat{\epsilon}_{4i}^2(t).$$

Sendo assim, pode-se construir a seguinte banda de confiança de nível de confiança de $(1 - \alpha)$

$$\left[\hat{S}_0(t) \pm C_\alpha n^{-1/2} \hat{Q}^{1/2}(t) \right],$$

em que C_α é o quantil de ordem $(1 - \alpha)$ de

$$\sup_{t \in [0, \tau]} \left| \frac{\Delta_S^k(t)}{\hat{Q}^{1/2}(t)} \right|,$$

com $\Delta_S^k(t)$ sendo a k -ésimo ($k = 1, \dots, K$) processo reamostrado.

Capítulo 5

Aplicação

Neste Capítulo é realizada a aplicação dos modelos discutidos aos dados do estudo do ICESP apresentado no Capítulo 1. Serão avaliados os seguintes modelos:

Modelo 1: Modelo de Cox

Modelo 2: Modelo semiparamétrico de riscos aditivos

Modelo 3: Modelo de Cox-Aalen - abordagem tradicional

Modelo 4: Modelo de Cox-Aalen - abordagem proposta

Conforme visto anteriormente, os modelos descritos possuem diferentes estruturas (aditiva e multiplicativa) e admitem efeitos fixos e, com exceção do Modelo 1, variantes no tempo. Sendo assim, deseja-se avaliar as principais diferenças quanto à interpretação dos parâmetros, bem como comparar o ajuste dos modelos.

Para auxiliar na decisão se o efeito de uma variável será considerado fixo ou variante no tempo (Modelos 2, 3 e 4) e se a relação será dada de forma multiplicativa ou aditiva (Modelos 3 e 4), serão utilizados o teste global de proporcionalidade de riscos, apresentado anteriormente na Tabela 1.2 do Capítulo 1, e o teste escore desenvolvido na seção 2.3.

O Modelo 1 é ajustado supondo-se que os efeitos de todas as covariáveis são fixos no tempo, ignorando as evidências de fuga da suposição de proporcionalidade encontradas anteriormente. No Modelo 2, assume-se que as covariáveis com falha na suposição de proporcionalidade possuem efeitos dependentes do tempo, e que as demais possuem efeitos fixos no tempo.

Os Modelos 3 e 4 são ambos com a formulação do modelo de Cox-Aalen. No entanto, o Modelo 3, que ilustra a abordagem adotada tradicionalmente, tem as covariáveis com fuga da suposição de proporcionalidade alocadas na parcela aditiva com efeitos dependentes do tempo, e as demais,

na parcela multiplicativa com efeitos fixos no tempo. Já no Modelo 4, propõe-se alocar as covariáveis com fuga da proporcionalidade na parcela aditiva com efeitos dependentes do tempo, assim como no Modelo 3, e avaliar a estrutura mais adequada para as demais covariáveis por meio do teste escore desenvolvido na seção 2.3.

A seguir são apresentados os ajustes dos modelos considerados. As variáveis Sexo e Idade não são consideradas por não se mostrarem influentes de acordo com os testes não paramétricos. Além disso, para que a comparação entre os modelos seja possível, não foi realizada a redução dos modelos. Para o ajuste dos modelos, foram utilizados os pacotes `survival` e `timereg` do *software R*.

Para os modelos que possuem efeitos dependentes do tempo, adota-se a estratégia de se avaliar a significância do efeito para, posteriormente, avaliar a hipótese de dependência do tempo dentre os efeitos que se mostraram significativos no primeiro teste.

5.1 Modelo de Cox

A interpretação do modelo de Cox é feita por meio de riscos relativos, que somada à flexibilidade do modelo por permitir que a taxa de falha basal não tenha uma estrutura pré definida, faz com que este modelo seja a principal ferramenta para a análise de dados de sobrevivência. A seguir, é apresentada na Tabela 5.1 o ajuste do modelo de Cox para os dados do ICESP.

Tabela 5.1: Ajuste do modelo de Cox para os dados do ICESP.

Variável	Coef.	E.P	R.R	Z (p)
IMC	-0,049	0,010	0,952	-5,147 (< 0,001)
Status do câncer (Ativo - recaída/progressão)	0,017	0,111	1,018	0,157 (0,875)
Status do câncer (Controlado/remissão)	-0,371	0,202	0,690	-1,833 (0,067)
Tipo de admissão (Não cirúrgico)	0,313	0,128	1,367	2,448 (0,014)
Extensão do câncer (Limitado)	-0,153	0,139	0,858	-1,100 (0,271)
Extensão do câncer (Localmente avançado)	0,086	0,121	1,090	0,713 (0,476)
Cirurgia (Sim)	-0,119	0,096	0,887	-1,244 (0,213)
Quimioterapia (Sim)	-0,092	0,109	0,912	-0,842 (0,400)
Radioterapia (Sim)	0,147	0,109	1,158	1,349 (0,177)
Presença de delirium (Sim/sedado)	0,225	0,103	1,252	2,184 (0,029)
Escala de capacidade funcional ECOG	0,124	0,049	1,132	2,513 (0,012)
Escala de qualidade de vida	-0,296	0,141	0,744	-2,099 (0,036)
Escore SAPS III	0,029	0,005	1,030	6,219 (< 0,001)

Coef.: Coeficiente; E.P: Erro Padrão; R.R: Risco relativo; Z: Estatística Z; p: Nível descritivo.

Por meio da Tabela 5.1, verifica-se que apenas as variáveis IMC, Tipo de admissão, Presença de delirium, Escala de capacidade funcional ECOG, Escala de qualidade de vida e Escore SAPS III se mostraram significativas, em que o aumento de uma unidade no índice de massa corpórea acarreta uma redução de 4,8% na taxa de falha; pacientes com admissão não cirúrgica apresentam taxa de falha 36,7% maior que pacientes com admissão cirúrgica; pacientes que apresentam delirium ou estão sedados possuem taxa de falha 25% maior que pacientes que não apresentam; o aumento de uma unidade na Escala de capacidade funcional ECOG acarreta um aumento de 13,2% na taxa de falha; o aumento de uma unidade na Escala de qualidade de vida acarreta uma redução de 25,6% na taxa de falha; o aumento de uma unidade no Escore SAPS III acarreta um aumento de 3% na taxa de falha; mantendo-se constantes as demais variáveis.

No entanto, é importante ressaltar que este modelo ignora a fuga da suposição de proporcionalidade de algumas covariáveis, apresentada na Tabela 1.2 do Capítulo 1.

5.2 Modelo semiparamétrico de riscos aditivos

Neste modelo, as covariáveis com fuga na suposição de proporcionalidade de acordo com o teste global apresentado na Tabela 1.2 são alocadas na parcela com efeitos dependentes do tempo (não paramétrica). São elas Tipo de admissão, Cirurgia, Radioterapia, Presença de delirium e Escore SAPS III. As demais covariáveis serão alocadas na parcela cujos efeitos são fixos no tempo.

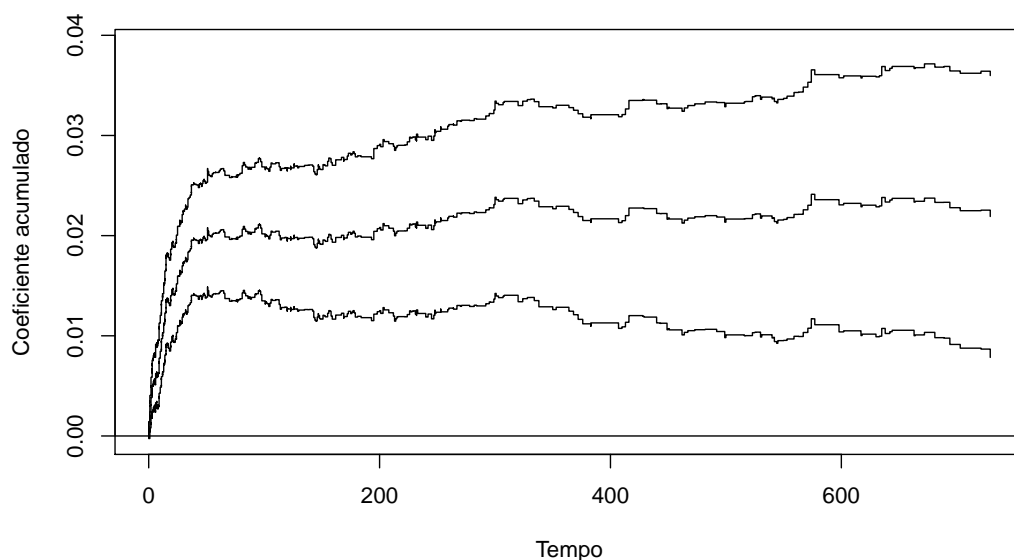


Figura 5.1: Estimativa do coeficiente acumulado de Escore SAPS III para o modelo semiparamétrico de riscos aditivos.

A interpretação do modelo é realizada em duas etapas. Na primeira, interpreta-se os resultados para os efeitos dependentes do tempo por meio dos coeficientes acumulados e, na segunda, avalia-se os efeitos fixos no tempo. Na Figura 5.1 é apresentada a estimativa do coeficiente acumulado do Escore SAPS III, juntamente com o intervalo de confiança pontual com nível de significância de 95%. Para verificar o comportamento do coeficiente ao longo do tempo, basta avaliar as inclinações da derivada da curva de coeficiente acumulado. Para o Escore SAPS III, por exemplo, observa-se que o coeficiente é positivo no início do estudo, mas decai para zero ao longo do tempo, ou seja, o aumento do Escore SAPS III acarreta um aumento na taxa de falha no início do estudo, mas este efeito vai se tornando nulo à medida que o tempo passa.

Os gráficos do intercepto e dos demais coeficientes se encontram no Apêndice B. O intercepto nesse caso representa a taxa de falha basal para pacientes com todas as covariáveis nulas. Vale ressaltar também que os intervalos de confiança apresentados são úteis apenas para conclusões pontuais. Na Tabela 5.2, são apresentadas as estatísticas e níveis descritivos dos testes de significância (Sup) e de dependência do tempo (K.S e C.v.M) dos parâmetros.

Tabela 5.2: Ajuste do modelo semiparamétrico de riscos aditivos para os dados do ICESP.

Variável	Sup (p)	K.S (p)	C.v.M (p)
Intercepto	5,030 (< 0,001)	0,582 (< 0,001)	70,600 (< 0,001)
Cirurgia (Sim)	1,670 (0,737)	0,127 (0,183)	4,080 (0,094)
Tipo de admissão (Não cirúrgico)	3,700 (0,004)	0,316 (0,001)	32,700 (0,001)
Radioterapia (Sim)	4,110 (< 0,001)	0,212 (0,112)	6,410 (0,252)
Presença de delirium (Sim/sedado)	4,120 (0,001)	0,338 (< 0,001)	20,600 (0,006)
Escore SAPS III	7,070 (< 0,001)	0,019 (< 0,001)	0,095 (< 0,001)

Sup: Estatística baseada no supremo; K.S: Estatística de Kolmogorov-Smirnov; C.v.M: Estatística de Cramér-von Mises; p: Nível descritivo.

Avaliando primeiramente o teste de significância dos parâmetros, verifica-se que apenas a variável Cirurgia não se mostra significativa. Dentre as demais variáveis, observa-se que apenas o efeito de Radioterapia pode ser considerado fixo no tempo.

Na Tabela 5.3 a seguir são apresentadas as estimativas dos coeficientes fixos no tempo:

Tabela 5.3: Ajuste do modelo semiparamétrico de riscos aditivos para os dados do ICESP.

Variável	Coef. ($\times 10^{-4}$)	Coef.	
IMC	-0,748	0,167	-4,49 (< 0,001)
Status do câncer (Ativo - recida/progressão)	-1,130	2,310	-0,490 (0,624)
Status do câncer (Controlado/remissão)	-8,080	3,800	-2,130 (0,033)
Extensão do câncer (Limitado)	-5,580	3,100	-1,800 (0,072)
Extensão do câncer (Localmente avançado)	-1,870	3,150	-0,592 (0,554)
Quimioterapia (Sim)	-0,961	2,620	-0,367 (0,713)
Escala de capacidade funcional ECOG	4,230	1,370	3,080 (0,002)
Escala de qualidade de vida	-5,410	3,070	-1,770 (0,077)

Coef.: Coeficiente; E.P: Erro Padrão; Z: Estatística Z; p: Nível descritivo.

Avaliando-se a Tabela acima, verifica-se que apenas os efeitos de IMC, de Status do Câncer (Controlado/remissão) e de Escala de capacidade funcional ECOG são significativos, em que o aumento de uma unidade do IMC acarreta uma redução estimada de 0,0000748 na taxa de falha; pacientes com Status do câncer classificado como "Controlado/remissão" possuem taxa de falha estimada 0,000808 menor que a taxa de falha estimada de pacientes classificados como "Ativo - recente"; o aumento de uma unidade na Escala de capacidade funcional ECOG acarreta um aumento estimado de 0,000423 na taxa de falha, mantendo-se constantes as demais características.

5.3 Modelo de Cox-Aalen

Nesta etapa, são construídos dois modelos com a formulação de Cox-Aalen. No primeiro (Modelo 3), considera-se que as covariáveis com fuga na suposição de proporcionalidade possuem efeitos dependentes do tempo por meio da estrutura aditiva, e que as demais covariáveis possuem efeitos fixos no tempo por meio da estrutura multiplicativa.

No segundo modelo (Modelo 4), também considera-se que as covariáveis com fuga na suposição de proporcionalidade dos riscos possuem efeitos dependentes do tempo por meio da estrutura aditiva, assim como no Modelo 3. No entanto, para as demais covariáveis, que não possuem indícios de que tenham efeitos dependentes do tempo, realiza-se o teste escore para riscos multiplicativos desenvolvido na Seção 2.3 e aquelas que tiverem a hipótese rejeitada são alocadas na parcela aditiva, com seu efeito dependente do tempo. Para a realização dos testes, foram consideradas 10 partições de tamanho 80. O teste escore para riscos aditivos não foi utilizado nesta etapa por não apresentar o comportamento tão satisfatório quanto o teste para riscos multiplicativos nos estudos de simulação realizados no Capítulo 2.

Na Tabela 5.4 a seguir, são apresentados os valores das estatísticas e os níveis descritivos do teste global de proporcionalidade e do teste escore de riscos proporcionais.

Tabela 5.4: Estatística (Nível descritivo) do teste global de proporcionalidade e do teste escore para riscos multiplicativos.

Variável	Riscos proporcionais	Riscos Multiplicativos
IMC	0,170 (0,680)	7,165 (0,007)
Status do câncer	1,259 (0,533)	0,011 (0,916)
Tipo de admissão	42,100 (< 0,001)	-
Extensão do câncer	0,036 (0,982)	0,294 (0,588)
Cirurgia	9,090 (0,003)	-
Radioterapia	13,100 (< 0,001)	-
Quimioterapia	0,255 (0,613)	3,430 (0,064)
Presença de delirium	15,900 (< 0,001)	-
Escala de capacidade funcional ECOG	0,453 (0,501)	0,078 (0,781)
Escala de qualidade de vida	0,244 (0,621)	2,540 (0,111)
Escore SAPS III	59,000 (< 0,001)	-

Por meio da Tabela 5.4, verifica-se que a hipótese de riscos multiplicativos é rejeitada apenas para a variável IMC. Dessa forma, os Modelos 3 e 4 diferem apenas com relação a essa variável, que é alocada na parcela multiplicativa no Modelo 3 e na parcela aditiva no Modelo 4.

A interpretação do modelo de Cox-Aalen é realizada em duas etapas. Na primeira, avalia-se a parcela referente ao modelo de Aalen por meio dos coeficientes acumulados de maneira análoga ao Modelo 2. Na segunda etapa, avalia-se a parcela referente ao modelo de Cox por meio dos riscos relativos de maneira análoga às interpretações do Modelo 1. A seguir são apresentados os ajustes do modelo tradicional e do modelo proposto.

5.3.1 Modelo tradicional

Na Tabela 5.5 a seguir é apresentado o ajuste da parcela aditiva do modelo de Cox-Aalen.

Tabela 5.5: Ajuste da parcela aditiva do modelo de Cox-Aalen para os dados do ICESP.

Variável	Sup (p)	K.S (p)
Intercepto	2,880 (0,066)	1,760 (0,008)
Cirurgia (Sim)	1,660 (0,742)	0,438 (0,188)
Tipo de admissão (Não cirúrgico)	2,200 (0,278)	0,880 (0,076)
Radioterapia (Sim)	2,660 (0,144)	0,683 (0,224)
Presença de delirium (Sim/sedado)	2,320 (0,268)	0,897 (0,034)
Escore SAPS III	3,210 (0,018)	0,059 (< 0,001)

Sup: Estatística baseada no supremo; K.S: Estatística de Kolmogorov-Smirnov; p: Nível descritivo.

Avaliando o teste de significância (Sup), verifica-se que apenas o Escore SAPS III se mostra significativo além do intercepto, que representa a taxa de falha basal de pacientes com todas as covariáveis nulas. Quanto à dependência do tempo (K.S), verifica-se que a hipótese de que o efeito de Escore SAPS III seja fixo no tempo é rejeitada. Avaliando o gráfico do coeficiente acumulado estimado de SAPS III na Figura B.12 do Apêndice, verifica-se que o comportamento é semelhante ao encontrado no modelo semiparamétrico aditivo. O gráfico dos demais coeficientes acumulados estimados também são apresentados no Apêndice B.

Na Tabela 5.6 a seguir são apresentados o ajuste da parcela multiplicativa.

Tabela 5.6: Ajuste da parcela multiplicativa do modelo de Cox-Aalen para os dados do ICESP.

Variável	Coef.	E.P	R.R	Z (p)
IMC	-0,048	0,011	0,953	-4,790 (< 0,001)
Status do câncer (Ativo - recaída/progressão)	0,000	0,110	1,000	-0,003 (0,998)
Status do câncer (Controlado/remissão)	-0,342	0,202	0,710	-1,680 (0,093)
Extensão do câncer (Limitado)	-0,215	0,144	0,807	-1,680 (0,092)
Extensão do câncer (Localmente avançado)	0,026	0,125	1,027	0,240 (0,810)
Quimioterapia (Sim)	-0,033	0,106	0,968	-0,321 (0,748)
Escala de capacidade funcional ECOG	0,137	0,053	1,147	2,690 (0,007)
Escala de qualidade de vida	-0,227	0,147	0,797	-1,640 (0,101)

Por meio da Tabela 5.6, verifica-se que apenas as covariáveis IMC e Escala de capacidade funcional ECOG se mostraram significativas, em que o aumento de uma unidade no índice de massa corpórea acarreta uma redução de 4,7% na taxa de falha e o aumento estimado de uma unidade na Escala de capacidade funcional ECOG acarreta um aumento estimado de 14,7% na

taxa de falha, mantendo-se constantes as demais características.

5.3.2 Modelo proposto

Na Tabela 5.7 a seguir é apresentado o ajuste da parcela aditiva do modelo de Cox-Aalen.

Tabela 5.7: Ajuste da parcela aditiva do modelo de Cox-Aalen para os dados do ICESP.

Variável	Sup (p)	K.S (p)
Intercepto	2,840 (0,066)	0,340 (0,302)
IMC	3,960 (0,002)	0,021 (0,014)
Cirurgia (Sim)	1,660 (0,744)	0,144 (0,176)
Tipo de admissão (Não cirúrgico)	2,640 (0,108)	0,287 (0,042)
Radioterapia (Sim)	3,460 (0,008)	0,229 (0,144)
Presença de delirium (Sim/sedado)	3,050 (0,056)	0,305 (0,002)
Escore SAPS III	4,800 (< 0,001)	0,019 (< 0,001)

Avaliando a Tabela 5.7 acima, verifica-se que neste modelo as covariáveis IMC, Radioterapia e Escore SAPS III são significativos. Destas covariáveis, apenas Radioterapia possui efeito que pode ser considerado fixo no tempo. Por meio da Figura B.14 do coeficiente acumulado estimado de IMC, verifica-se que o coeficiente é negativo no início do estudo, mas vai se aproximando de zero ao longo do tempo. O comportamento do coeficiente do Escore SAPS III apresentado na Figura B.19 é semelhante aos comportamentos encontrados nos modelos anteriores.

Tabela 5.8: Ajuste da parcela multiplicativa do modelo de Cox-Aalen para os dados do ICESP.

Variável	Coef.	E.P	R.R	Z (p)
Status do câncer (Ativo - recaída/progressão)	0,003	0,110	1,003	0,033 (0,974)
Status do câncer (Controlado/remissão)	-0,343	0,202	0,710	-1,720 (0,086)
Extensão do câncer (Limitado)	-0,229	0,145	0,795	-1,780 (0,074)
Extensão do câncer (Localmente avançado)	0,015	0,126	1,015	0,137 (0,891)
Quimioterapia (Sim)	-0,044	0,105	0,957	-0,430 (0,667)
Escala de capacidade funcional ECOG	0,132	0,054	1,141	2,620 (0,009)
Escala de qualidade de vida	-0,230	0,147	0,795	-1,670 (0,096)

Avaliando a Tabela 5.8 de ajuste da parcela multiplicativa do modelo, verifica-se apenas a Escala de capacidade funcional ECOG se mostra significativa, em que o aumento de uma unidade na Escala de capacidade funcional ECOG acarreta um aumento estimado de 14,1% na taxa de falha, mantendo-se constantes as demais características.

5.4 Comparação dos modelos

Nesta etapa, deseja-se comparar os modelos considerados quanto ao ajuste. Para tal, propõe-se comparar as curvas de sobrevivência estimada de cada um dos modelos com a curva de Kaplan-Meier geral.

Para construir as curvas de sobrevivência estimada, foi considerado um perfil médio (comum a todos os modelos), em que se considera a média para as covariáveis contínuas e as frequências relativas no lugar do valor um das variáveis indicadoras das covariáveis categóricas. É importante ressaltar que esta metodologia foi considerada apenas para que as curvas de sobrevivência estimada pudessem ser comparadas, não tendo interesse na interpretação das curvas estimadas deste perfil médio, que pode não fazer sentido.

Na Figura 5.2 a seguir, tem-se as estimativas das curvas de sobrevivência do perfil médio para todos os modelos considerados, juntamente com a curva de Kaplan Meier geral dos dados do ICESP.

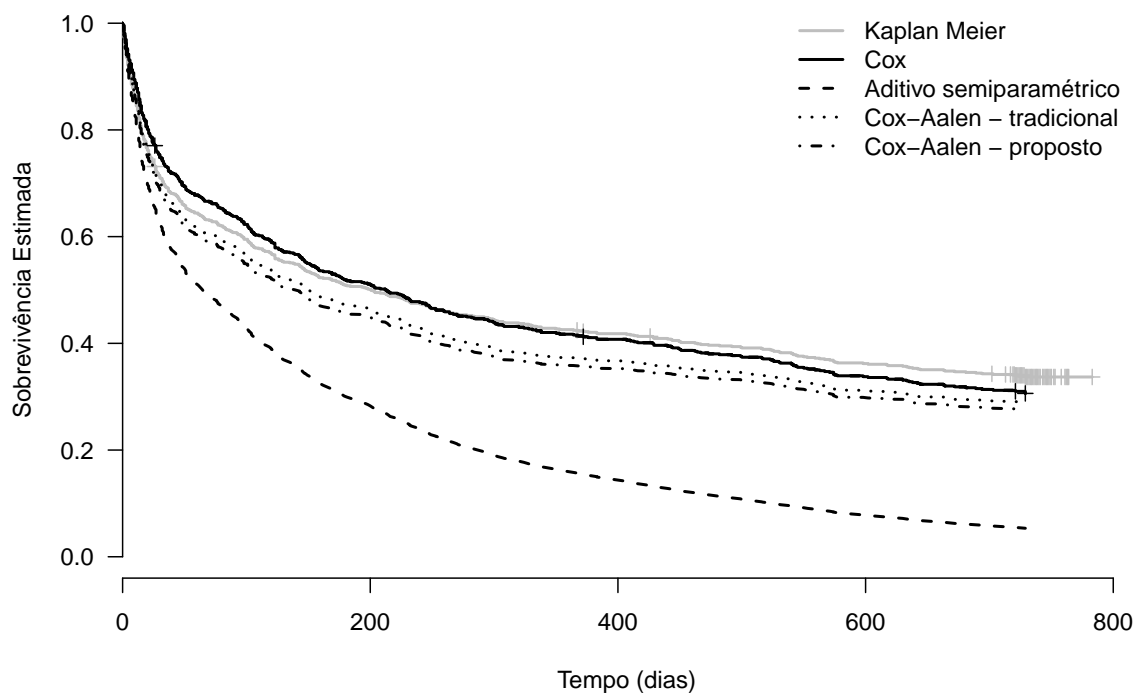


Figura 5.2: Curvas de Kaplan Meier e de sobrevivências estimadas pelos modelos considerados.

Avaliando a Figura 5.2, pode-se perceber que a curva de sobrevivência estimada mais próxima da curva geral de Kaplan Meier é a do modelo de Cox, ainda que este modelo não leve em consideração a fuga da suposição de proporcionalidade de algumas covariáveis. Em seguida, tem-se as curvas estimadas sob o modelo de Cox-Aalen, que estão bem próximas uma da outra, com a

curva do modelo tradicional ligeiramente mais próxima da curva de Kaplan Meier geral. A sobrevivência estimada que mais se distancia da curva de Kaplan Meier é a do modelo semiparamétrico aditivo, indicando que o uso exclusivo da estrutura aditiva pode não ser o mais adequado para estes dados.

Além do gráfico de curvas de sobrevivência estimada, também foram avaliadas as amplitudes dos intervalos/bandas de confiança para a função de sobrevivência estimada de cada um dos modelos considerados. Para as estimativas de Kaplan-Meier e do modelo de Cox, foram construídos três intervalos de confiança. O primeiro, denominado "Padrão", é baseado no fato de que para t fixo, $\hat{S}(t) - S(t)$ converge para uma distribuição Normal com média zero e variância $\Sigma(t) = \text{Var}(S(t))$, que pode ser estimada pela fórmula de Greenwood (1926). Assim, o intervalo de confiança padrão tem a seguinte forma

$$\left[\hat{S}(t) \pm z_{\alpha/2} \hat{\Sigma}(t)^{1/2} \right].$$

No entanto, o intervalo padrão pode conter limites inferiores a zero ou superiores a um. Sendo assim, outros dois intervalos, utilizando as transformações log e log-log, são propostos. A partir do método delta, tem-se que

$$\begin{aligned} \exp \left[\log \hat{S}(t) \pm z_{\alpha/2} \frac{\hat{\Sigma}(t)^{1/2}}{\hat{S}(t)} \right] &= \left[\hat{S}(t) \exp \left\{ \pm z_{\alpha/2} \frac{\hat{\Sigma}(t)^{1/2}}{\hat{S}(t)} \right\} \right] \\ \exp(-\exp) \left[\log(-\log \hat{S}(t)) \pm z_{\alpha/2} \frac{\hat{\Sigma}(t)^{1/2}}{\hat{S}(t) \log \hat{S}(t)} \right] &= \left[\hat{S}(t)^{\exp \left\{ \pm z_{\alpha/2} \frac{\hat{\Sigma}(t)^{1/2}}{\hat{S}(t) \log \hat{S}(t)} \right\}} \right]. \end{aligned}$$

Na Tabela 5.9 a seguir são apresentados os limites inferior e superior, bem como as amplitudes dos intervalos nos instantes t_p tal que $\hat{S}(t_p) = p$. Pode-se verificar que, conforme esperado, as amplitudes dos intervalos aumentam à medida que os instantes de tempo avaliados aumentam. No instante $t_{0.95}$ o modelo de Cox apresenta a menor amplitude, no instante $t_{0.75}$ os intervalos obtidos por meio do Kaplan-Meier e do modelo aditivo apresentam as menores amplitudes e, no instante $t_{0.50}$ o modelo aditivo apresentou a menor amplitude. Os modelos de Cox-Aalen apresentam maior amplitude em todos os instantes avaliados.

Tabela 5.9: Limites e amplitudes dos intervalos/bandas de confiança para a função de sobrevivência estimada de cada um dos modelos aplicados aos dados do ICESP.

		$t_{0.95}$			$t_{0.75}$			$t_{0.50}$		
Modelo		LI	LS	Δ	LI	LS	Δ	LI	LS	Δ
KM	Padrão	0,944	0,972	0,028	0,718	0,778	0,060	0,463	0,533	0,070
	Log	0,944	0,972	0,028	0,718	0,779	0,061	0,465	0,535	0,070
	Log-log	0,941	0,970	0,029	0,716	0,777	0,061	0,463	0,533	0,070
Cox	Padrão	0,939	0,965	0,026	0,717	0,779	0,062	0,462	0,539	0,077
	Log	0,939	0,965	0,026	0,718	0,780	0,062	0,463	0,540	0,077
	Log-log	0,937	0,964	0,027	0,716	0,778	0,062	0,461	0,538	0,077
Aditivo		0,936	0,965	0,029	0,720	0,781	0,061	0,468	0,535	0,067
Cox-Aalen	Tradicional	0,936	0,966	0,030	0,715	0,783	0,067	0,457	0,542	0,085
	Proposto	0,935	0,965	0,031	0,715	0,783	0,068	0,458	0,543	0,085

LI: Limite inferior; LS: Limite superior; Δ : Amplitude do intervalo.

5.5 Discussão

As diferentes estruturas (aditiva e multiplicativa) adotadas levam a diferentes formas de interpretação dos parâmetros. No modelo de Cox, esta interpretação é realizada por meio de riscos relativos, que parece ser mais informativa quando comparada à interpretação por meio de variação de taxa de falha dos efeitos fixos no tempo do modelo semiparamétrico aditivo.

No entanto, o modelo de Cox não leva em consideração as evidências de falha da suposição de proporcionalidade dos riscos encontradas com o teste global. Dessa forma, este modelo pode não ser o mais adequado, além de se perder informações interessantes quanto ao comportamento dos parâmetros dependentes do tempo, como as encontradas no modelo semiparamétrico de riscos aditivos.

Dessa forma, o modelo de Cox-Aalen parece ser bastante interessante no que diz respeito à interpretação dos parâmetros, uma vez que mantém a interpretação por riscos relativos para covariáveis com efeitos fixos no tempo e ao mesmo tempo, traz informações interessantes com relação ao comportamento dos efeitos de covariáveis em que a suposição de riscos proporcionais não é adequada.

Quanto à estimação da curva de sobrevivência, a estimativa do modelo de Cox se mostrou mais próxima da curva de Kaplan-Meier, com amplitudes de intervalos medianas. Já o modelo aditivo, apesar de apresentar menores amplitudes de intervalo, apresentou a curva estimada mais afastada da curva de Kaplan-Meier. Por fim, os modelos de Cox-Aalen apresentaram estimativas mais próximas da curva de Kaplan-Meier, mas com maiores amplitudes das bandas de confiança.

Com relação ao tempo de execução, tem-se que os modelos que incorporam efeitos dependentes do tempo apresentam tempo de processamento ligeiramente superior ao do modelo de Cox. O modelo de Cox-Aalen (proposto) apresenta ainda o tempo adicional para a realização do teste escore, que aparentemente não trouxe grandes mudanças no ajuste do modelo.

Sendo assim, de maneira geral o modelo tradicional de Cox ainda parece ser uma opção robusta na modelagem dos dados do ICESP, ainda que não se leve em consideração a fuga da suposição de proporcionalidade de algumas variáveis. O uso de modelos que incorporem efeitos dependentes do tempo neste caso foi uma possível solução que trouxe resultados interessantes com relação ao comportamento temporal dos efeitos destas variáveis. O modelo aditivo de Aalen, no entanto, se mostrou pior que o modelo de Cox-Aalen quanto à estimação da curva de sobrevivência, segundo os critérios adotados neste trabalho.

Capítulo 6

Pesquisas futuras

O Modelo de Cox-Aalen com abordagem proposta do capítulo anterior utiliza o teste de proporcionalidade para avaliar as evidências de efeitos dependentes do tempo e o teste escore (desenvolvido na Seção 2.3) para auxiliar na classificação das variáveis que não apresentam fuga da suposição de proporcionalidade dentre as estruturas aditiva e multiplicativa. No entanto, as variáveis em que a hipótese de riscos multiplicativos é rejeitada são alocadas na parcela aditiva, com seus parâmetros dependentes do tempo, ainda que não haja evidências para tal. Uma abordagem mais interessante seria o uso da extensão do modelo de Cox-Aalen, citado em [Scheike e Zhang \(2002\)](#), em que a função taxa de falha basal é substituída pelo modelo semiparamétrico de riscos aditivos discutido na Subseção 3.2.1:

$$h(t) = Y(t) \left[\mathbf{x}_1(t)^\top \boldsymbol{\alpha}(t) + \mathbf{x}_2(t)^\top \boldsymbol{\gamma} \right] \exp\{\mathbf{z}(t)^\top \boldsymbol{\beta}\}.$$

Dessa forma, poder-se-ia construir um modelo em que as variáveis com fuga na suposição de proporcionalidade fossem alocadas na parcela aditiva com seus efeitos dependentes do tempo, e as demais seriam avaliadas por meio do teste escore para a escolha da melhor estrutura. No caso de rejeição da hipótese de riscos multiplicativos, a variável seria então alocada na parcela aditiva com seu efeito fixo no tempo e, caso contrário, seria alocada na parcela multiplicativa. Para tal, seria necessário primeiramente o estudo mais aprofundado deste modelo tais como suas técnicas de estimação e propriedades assintóticas, além da construção de rotinas computacionais para o seu ajuste.

Foi observado também no Capítulo 2 que o teste escore para riscos aditivos ($\lambda = 1$) não se comporta tão bem quanto o teste para riscos multiplicativos ($\lambda = 0$). Dessa forma, o entendimento desta diferença comportamental dos testes por meio de um estudo aprofundado de conceitos

de testes de hipóteses para parâmetros de famílias de distribuições poderia ser interessante. Se houver entendimento e possível melhoria do teste escore para riscos aditivos, poder-se-ia propor construções mais refinadas do modelo de Cox-Aalen semiparamétrico utilizando-se os dois testes escore desenvolvidos.

Uma outra alternativa para esta metodologia seria por meio da abordagem bayesiana, podendo-se construir intervalos de credibilidade para o parâmetro λ para avaliar a estrutura mais adequada para os dados.

Por fim, no capítulo anterior, os principais modelos discutidos foram aplicados aos dados do ICESP e comparados por meio da estimativa da função de sobrevivência. Uma possível alternativa para a comparação destes modelos seria o uso de critérios baseados em penalizações da função log-verossimilhança de acordo com o número de parâmetros do modelo. Alguns critérios já foram propostos, como por exemplo o critério AIC de Akaike (1974) e o critério BIC de Schwarz (1978), e vêm sendo amplamente utilizados para auxiliar na seleção de modelos cujos parâmetros são fixos no tempo. No entanto, o uso desta metodologia para modelos com efeitos dependentes do tempo parece não estar totalmente desenvolvido e poderia ser de grande utilidade na seleção de modelos que incorporam efeitos dependentes do tempo.

Apêndice A

Conceitos básicos de processos de contagem

Considere T^* e C duas variáveis aleatórias independentes e não negativas, em que T^* representa o tempo de ocorrência de um evento de interesse e C , o seu respectivo tempo de censura. Em situações práticas, o que efetivamente se observa é a variável aleatória $T = \min(T^*, C)$, em conjunto com o indicador de falha $\delta = \mathbb{I}(T^* \leq C)$. Assume-se ainda que T^* possui função de distribuição F , com função densidade $f(t) = dF(t)/dt$ e função de sobrevivência $S(t) = 1 - F(t) = \mathbb{P}(T^* > t)$.

Sendo assim, define-se a função taxa de falha como

$$h^*(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t \leq T^* < t + \Delta t | T^* \geq t)}{\Delta t} = \frac{f(t)}{S(t)}.$$

A abordagem baseada em processos de contagem traz ferramentas úteis e poderosas para o estudo de modelos de análise de sobrevivência. A seguir são apresentados os conceitos básicos de processos de contagem, que será a abordagem adotada nos demais capítulos desta dissertação.

Suponha $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade em que Ω é o espaço amostral, \mathcal{F} é a σ -álgebra de subconjuntos de Ω e $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$, uma medida de probabilidade.

Definição 1 *Um processo estocástico é uma família de variáveis aleatórias $X = \{X(t) : t \in \mathcal{I}\}$, indexados em \mathcal{I} e definidos em \mathcal{F} .*

Definição 2 *Uma família de sub- σ -álgebras $\{\mathcal{F}_t : t \geq 0\}$ de \mathcal{F} é crescente se $s \leq t$ implica em $\mathcal{F}_s \subset \mathcal{F}_t$, $s, t \in \mathcal{I}$. Uma família crescente de sub- σ -álgebras é denominada filtragem.*

Definição 3 Um espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$ equipado com uma filtração $\{\mathcal{F}_t : t \geq 0\}$ é denominado base estocástica e é representado por $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, \mathbb{P})$.

Definição 4 Um processo estocástico $X = \{X(t) : t \geq 0\}$ é adaptado à filtração $\{\mathcal{F}_t : t \geq 0\}$ se para $\forall t \geq 0$, $X(t)$ é \mathcal{F}_t -mensurável.

Definição 5 Um processo de contagem é um processo estocástico $\{N(t) : t \geq 0\}$ adaptado à filtração $\{\mathcal{F}_t : t \geq 0\}$, tal que

1. $N(0) = 0$;
2. $N(t) < \infty$;
3. O processo é contínuo à direita;
4. O processo possui apenas descontinuidades de tamanho 1.

No contexto de Análise de Sobrevivência descrito anteriormente, os valores observados do par (T, δ) passam a ser representados pelo par $(N(t), Y(t))$, em que

$$N(t) = \mathbb{I}_{(T \leq t, \delta=1)} \quad \text{e} \quad Y(t) = \mathbb{I}_{(T \geq t)}.$$

Dessa forma, define-se o incremento $dN(t) = N((t + \Delta t)-) - N(t-)$, com $N(t-) = \lim_{s \uparrow t} N(s)$, como sendo o indicador de falha no intervalo $[t, t + \Delta t)$ e a variável $Y(t)$ é vista como o indicador de risco no instante t .

Seguindo [Fleming e Harrington \(2011\)](#), tem-se que

$$h^*(t) = \lim_{\Delta t \rightarrow 0} \frac{\mathbb{P}(t \leq T^* < t + \Delta t | T^* \geq t)}{\Delta t}$$

$$\Rightarrow h^*(t)\Delta t \approx \mathbb{P}(t \leq T^* < t + \Delta t | T^* \geq t).$$

Como T^* e C são independentes,

$$h^*(t)\Delta t \approx \mathbb{P}(t \leq T^* < t + \Delta t | T^* \geq t, C \geq t)$$

$$\Rightarrow h^*(t)\Delta t \approx \mathbb{P}(dN(t) = 1 | Y(t) = 1)$$

$$\Rightarrow h^*(t)\Delta t \approx \mathbb{E}(dN(t) | Y(t) = 1).$$

Assim, a função taxa de falha dá a ideia da taxa média de mudança de $N(t)$ no intervalo $[t, t + \Delta t)$, dado que o tempo de falha e de censura excedem t .

Além disso,

$$\begin{aligned}
\mathbb{E}(N(t)) &= \mathbb{P}(T \leq t, \delta = 1) = \mathbb{P}(T^* \leq t; T^* \leq C) \\
&= \int_0^t \mathbb{P}(C \geq u) f(u) du \\
&= \int_0^t \mathbb{P}(C \geq u) S(u) \frac{f(u)}{S(u)} du \\
&= \int_0^t \mathbb{P}(C \geq u) \mathbb{P}(T^* > u) h^*(u) du \\
&= \int_0^t \mathbb{P}(C \geq u; T^* > u) h^*(u) du \\
&= \int_0^t \mathbb{P}(T \geq u) h^*(u) du \\
&= \int_0^t \mathbb{E}(Y(u)) h^*(u) du \\
&= \mathbb{E} \left(\int_0^t Y(u) h^*(u) du \right) = \mathbb{E}(H(t)).
\end{aligned}$$

Sendo assim, o processo $M(t) = N(t) - H(t)$ possui média zero. Além disso, $h(t) = Y(u)h^*(u)$ é definido como processo de intensidade de $N(t)$.

Definição 6 Seja $X = \{X(t) : t \geq 0\}$ um processo de contagem contínuo à direita e $\{\mathcal{F}_t : t \geq 0\}$, uma filtragem. Então X será um martingal com respeito a $\{\mathcal{F}_t : t \geq 0\}$ se

1. X é adaptado à \mathcal{F}_t ;
2. $\mathbb{E}|X(t)| < \infty, \forall t < \infty$;
3. $\mathbb{E}[X(t+s)|\mathcal{F}_t] = X(t)$ ou, equivalentemente, $\mathbb{E}[dX(t)|\mathcal{F}_{t-}] = 0 \quad \forall s, t > 0$.

X será um submartingal (supermartingal) se

$$\mathbb{E}[X(t+s)|\mathcal{F}_t] \geq (\leq) X(t).$$

Definindo a filtragem

$$\mathcal{F}_t = \sigma \{N(u), \mathbb{I}_{(T \leq u, \delta=0)} : 0 \leq u \leq t\},$$

tem-se que

$$\begin{aligned}
\mathbb{E} [dM(t)|\mathcal{F}_{t-}] &= \mathbb{E} [dN(t)|\mathcal{F}_{t-}] - \mathbb{E} [dH(t)|\mathcal{F}_{t-}] \\
&= Y(t)h^*(t)dt - dH(t) \\
&= Y(t)h^*(t)dt - Y(t)h^*(t)dt \\
&= 0,
\end{aligned}$$

e, portanto, $M(t)$ é martingal com respeito a \mathcal{F}_t .

Definição 7 Seja $X = \{X(t) : t \geq 0\}$ um processo estocástico e $\{\mathcal{F}_t : t \geq 0\}$ uma filtragem. Se $X(t)$ é \mathcal{F}_t -mensurável, então $X(t)$ é denominado processo \mathcal{F}_t -previsível.

Teorema 5 (Decomposição de Doob-Meyer) Seja $(\Omega, \mathcal{F}, \{\mathcal{F}_t : t \geq 0\}, \mathbb{P})$ uma base estocástica e $X(t)$ um submartingal contínuo à direita. Então existe um processo martingal contínuo à direita M e um processo crescente, contínuo à direita $H(t)$, com $\mathbb{E}(H(t)) < \infty$, tal que

$$X(t) = M(t) + H(t) \quad \forall t \geq 0.$$

Corolário 1 Seja $X = \{X(t) : t \geq 0\}$ um processo de contagem \mathcal{F}_t -adaptado com $\mathbb{E}(N(t)) < \infty, \forall t \geq 0$. Então existe um único processo crescente e previsível $H(t)$ tal que $H(0) = 0$ q.c, $\mathbb{E}(H(t)) < \infty$ e

$$M(t) = X(t) - H(t)$$

é um martingal de média zero.

O processo $\{H(t) : t \geq 0\}$ é também denominado de compensador de $X(t)$.

A decomposição de um processo de contagem em um martingal mais um compensador é uma ferramenta muito útil na modelagem de processos de contagem, uma vez que o compensador pode ser visto como a parte sistemática do modelo e o martingal, como um ruído de média zero. Dessa forma, torna-se possível a derivação de propriedades assintóticas para os estimadores semi e não paramétricos.

Para ilustrar o uso da decomposição de Doob-Meyer na construção de estimadores, será construído a seguir o estimador de Nelson-Aalen para a função de risco acumulada. Para tal, definimos

$$N_{\bullet}(t) = \sum_{i=1}^n N_i(t) \quad Y_{\bullet}(t) = \sum_{i=1}^n Y_i(t) \quad M_{\bullet}(t) = \sum_{i=1}^n M_i(t).$$

O processo $N_{\bullet}(t)$ pode ser decomposto em

$$N_{\bullet}(t) = \int_0^t Y_{\bullet}(u)h^*(u)du + M_{\bullet}(t)$$

$$dN_{\bullet}(t) = Y_{\bullet}(t)h^*(t)dt + dM_{\bullet}(t)$$

$$\frac{dN_{\bullet}(t)}{Y_{\bullet}(t)} = h^*(t)dt + \frac{dM_{\bullet}(t)}{Y_{\bullet}(t)},$$

se $Y_{\bullet}(t) \neq 0$. Foi visto anteriormente que $dM_i(t)$ é um processo de média zero e, sendo assim, o processo $dM_{\bullet}(t)$ também pode ser visto como um ruído. Se definir $J(t) = \mathbb{I}_{\{Y_{\bullet}(t) > 0\}}$ e estabelecer que $0/0 = 0$, tem-se que

$$\int_0^t \frac{J(u)}{Y_{\bullet}(u)} dN_{\bullet}(u) = \int_0^t J(u)h^*(u)du + \int_0^t \frac{J(u)}{Y_{\bullet}(u)} dM_{\bullet}(u).$$

Como a parcela $\int_0^t \frac{J(u)}{Y_{\bullet}(u)} dM_{\bullet}(u)$ pode ser vista como um ruído, tem-se que o estimador de Nelson-Aalen para a função taxa de falha acumulada é dado por

$$\hat{H}^*(t) = \int_0^t \frac{J(u)}{Y_{\bullet}(u)} dN_{\bullet}(u).$$

Apêndice B

Gráficos dos ajustes dos modelos

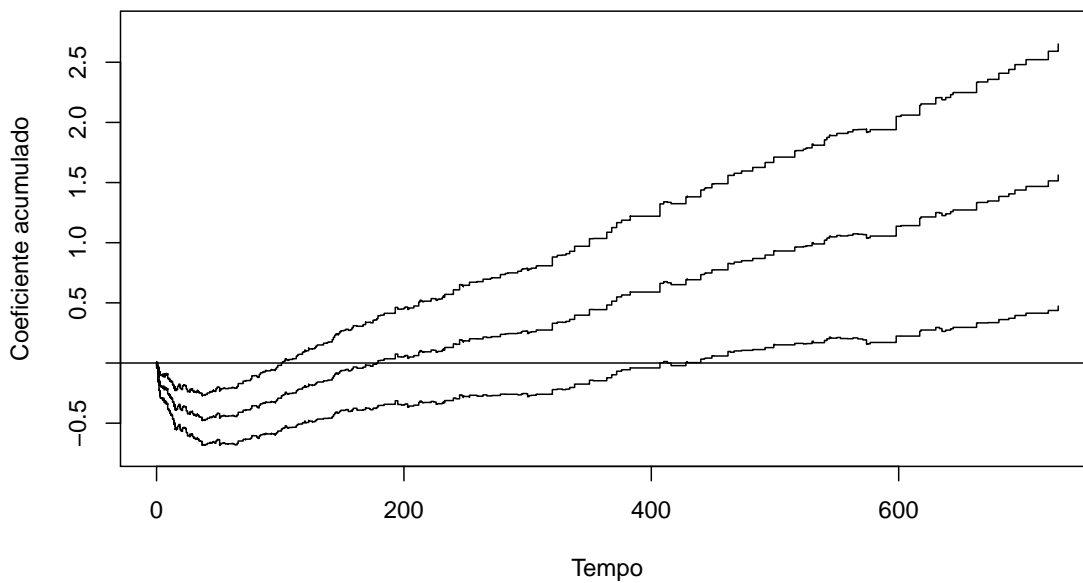


Figura B.1: Estimativa do intercepto para o modelo semiparamétrico de riscos aditivos.

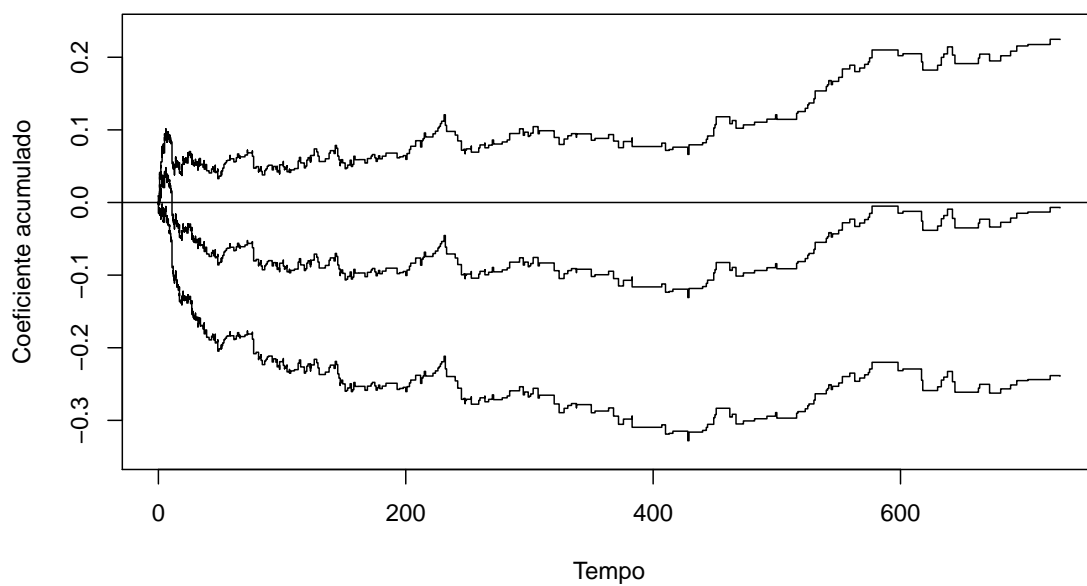


Figura B.2: Estimativa do coeficiente acumulado de Cirurgia (Sim) para o modelo semiparamétrico de riscos aditivos.

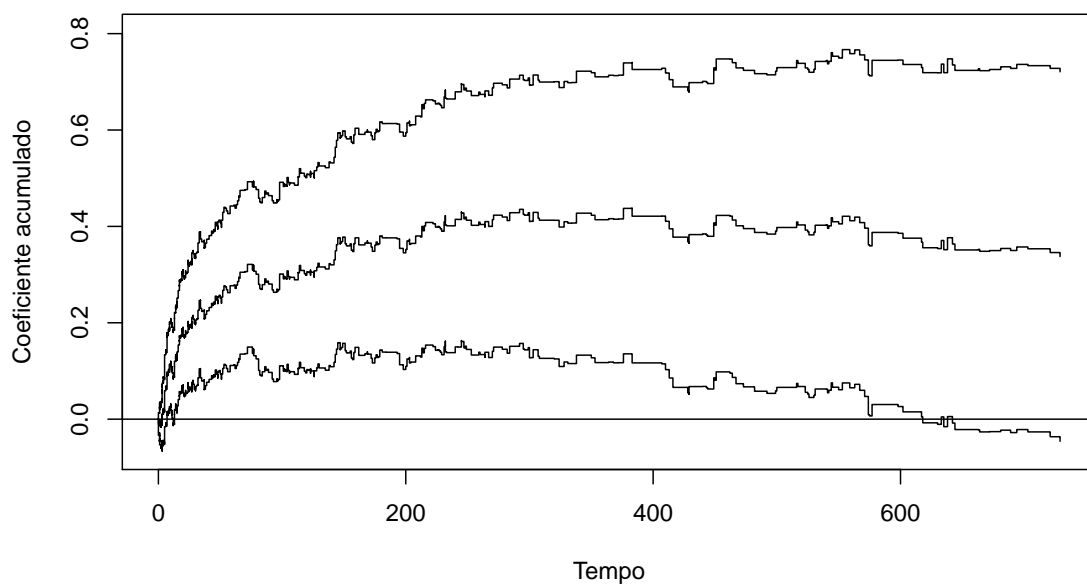


Figura B.3: Estimativa do coeficiente acumulado de Tipo de admissão (Não cirúrgico) para o modelo semiparamétrico de riscos aditivos.

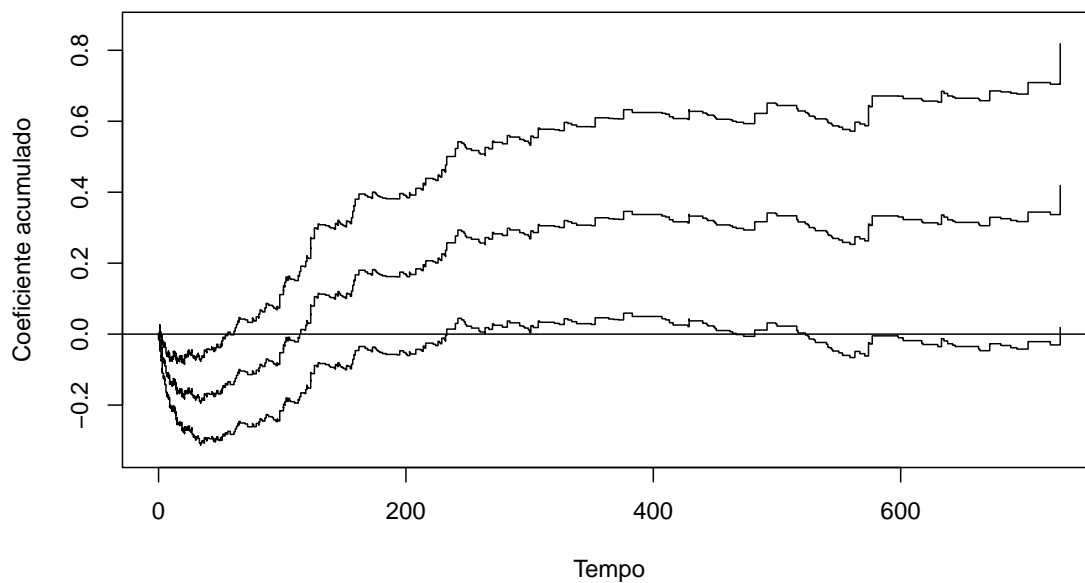


Figura B.4: Estimativa do coeficiente acumulado de Radioterapia (Sim) para o modelo semiparamétrico de riscos aditivos.

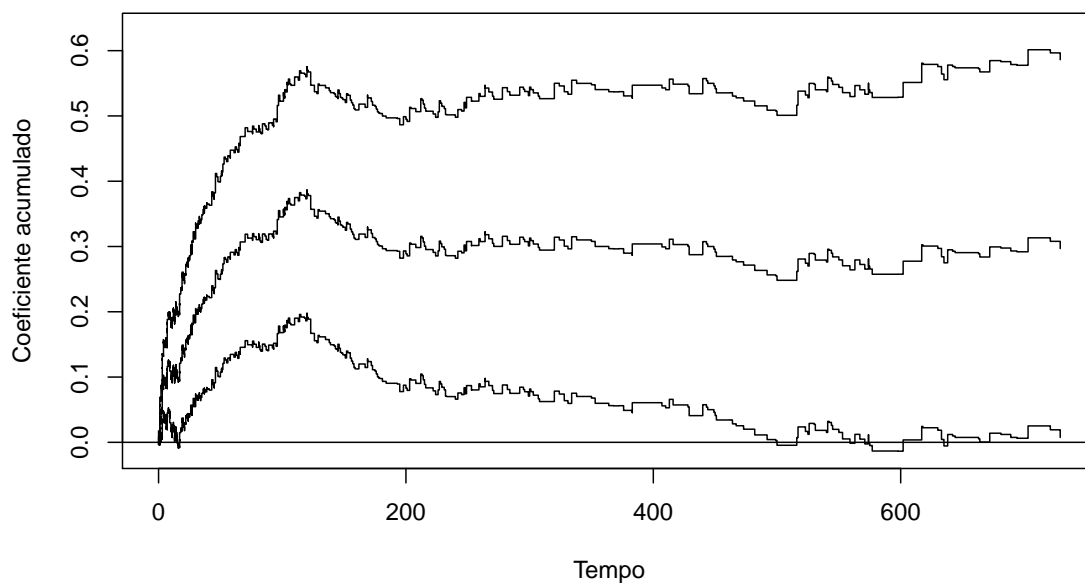


Figura B.5: Estimativa do coeficiente acumulado de Presença de delirium (Sim/sedado) para o modelo semiparamétrico de riscos aditivos.

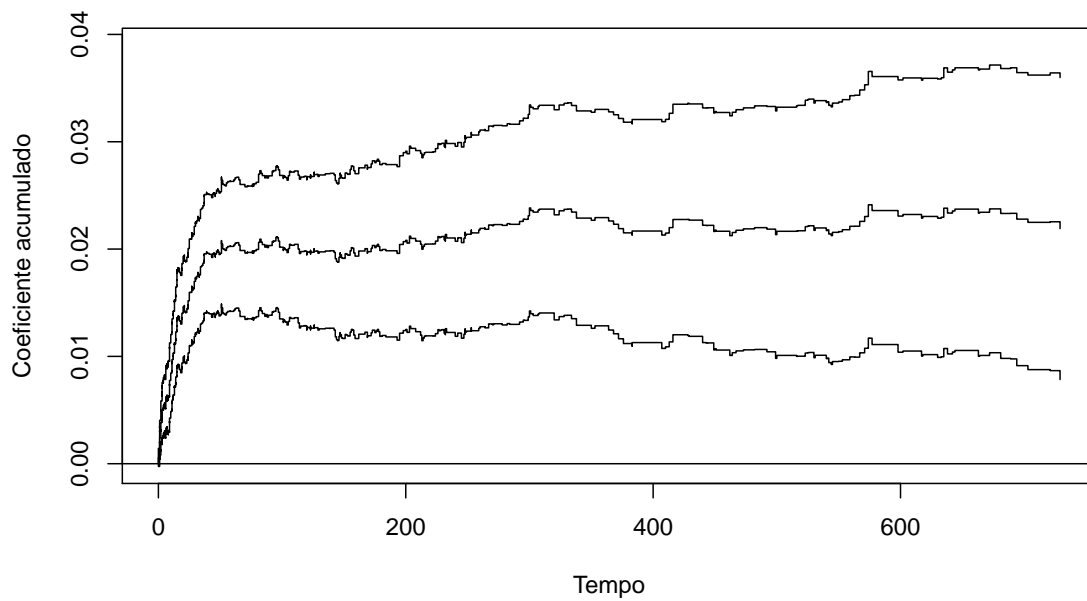


Figura B.6: Estimativa do coeficiente acumulado de Escore Saps III para o modelo semiparamétrico de riscos aditivos.

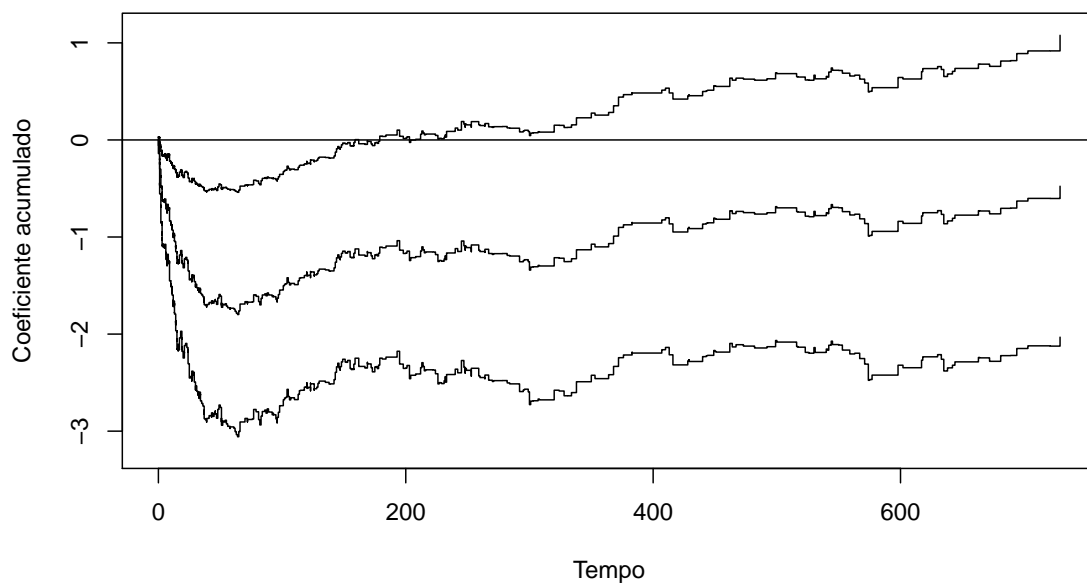


Figura B.7: Estimativa do intercepto para o modelo de Cox-Aalen - tradicional.

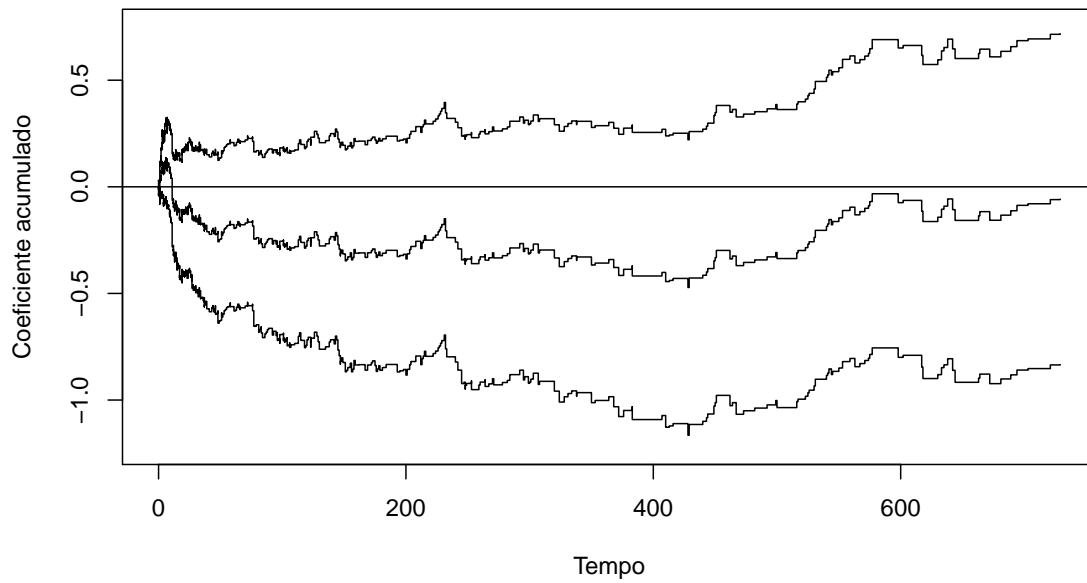


Figura B.8: Estimativa do coeficiente acumulado de Cirurgia (Sim) para o modelo de Cox-Aalen - tradicional.

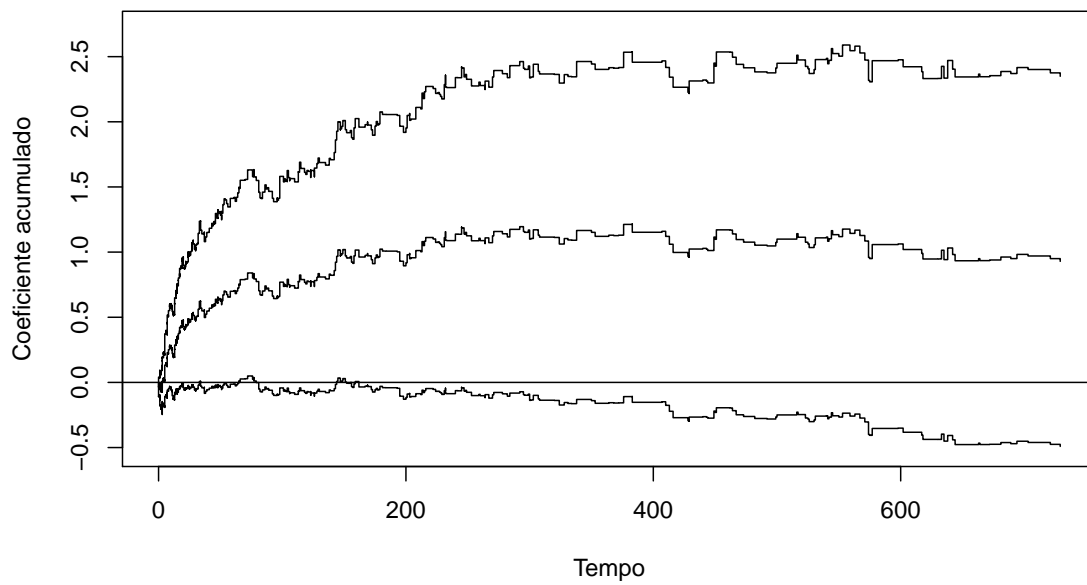


Figura B.9: Estimativa do coeficiente acumulado de Tipo de admissão (Não cirúrgico) para o modelo de Cox-Aalen - tradicional.

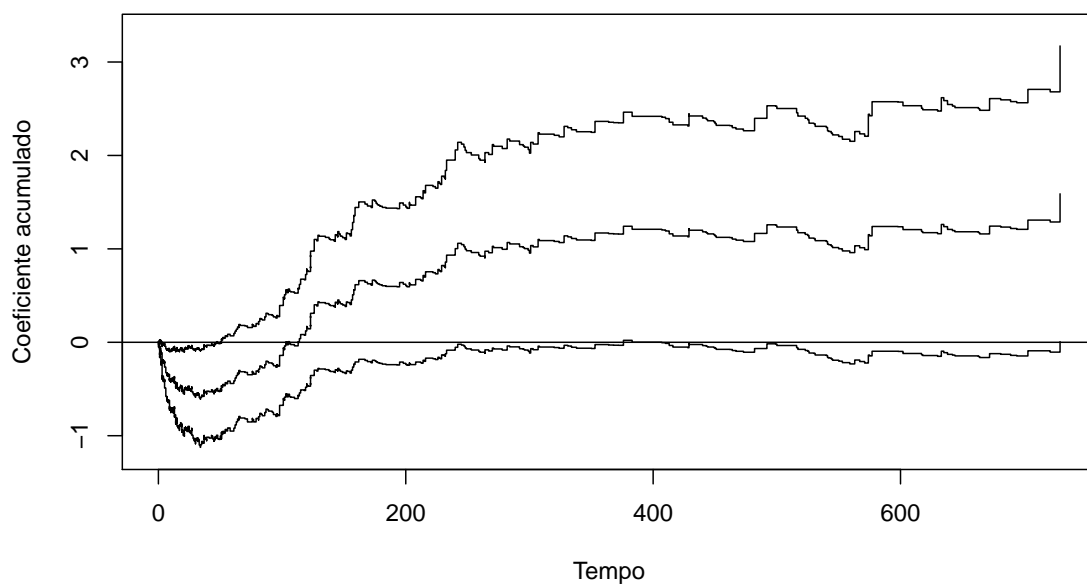


Figura B.10: Estimativa do coeficiente acumulado de Radioterapia (Sim) para o modelo de Cox-Aalen - tradicional.

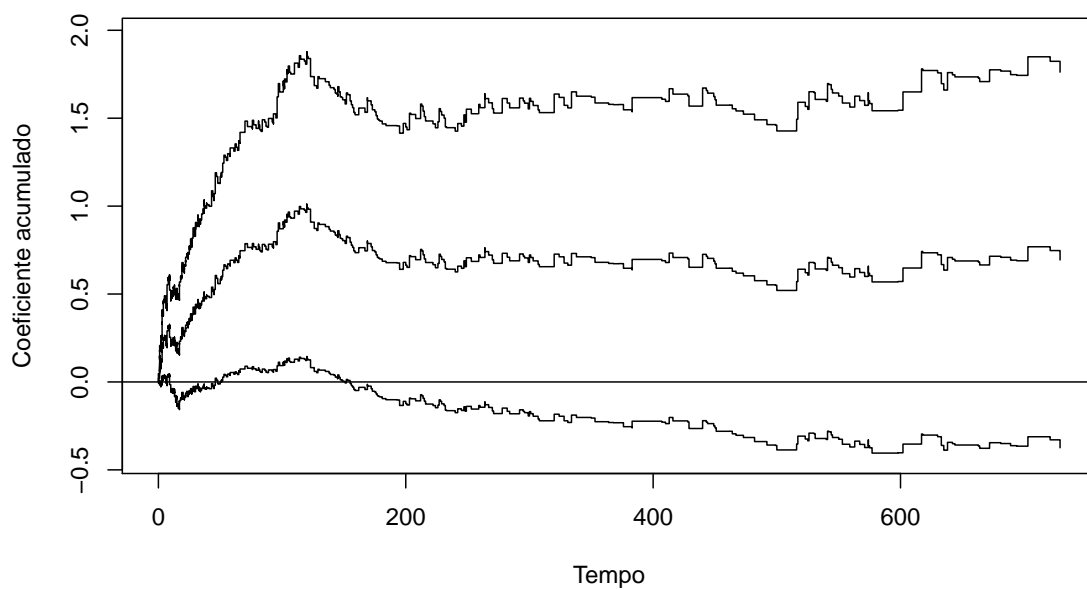


Figura B.11: Estimativa do coeficiente acumulado de Presença de delirium (Sim/sedado) para o modelo de Cox-Aalen - tradicional.

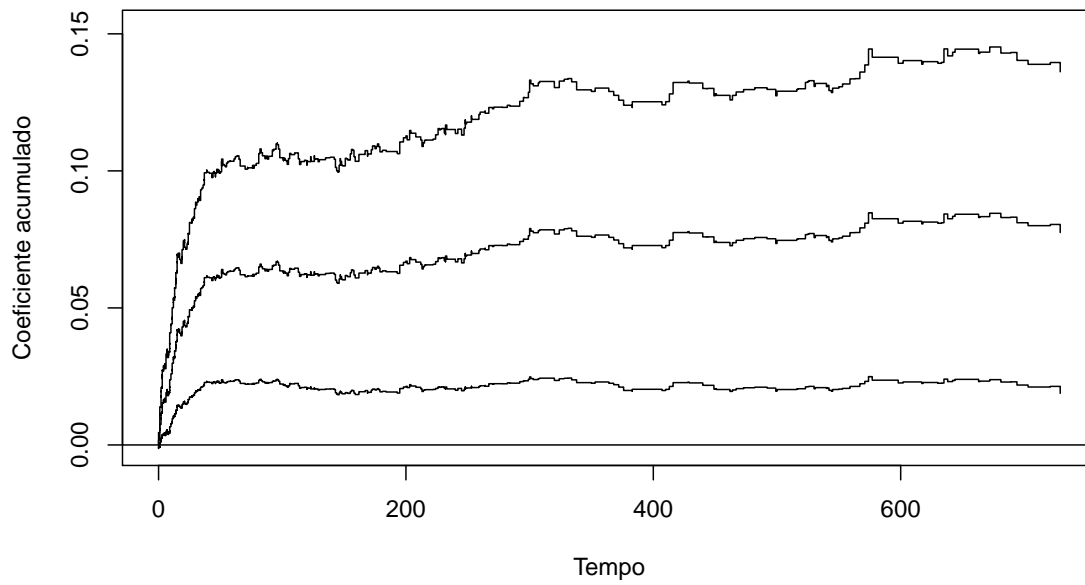


Figura B.12: Estimativa do coeficiente acumulado de Escore SAPS III para o modelo de Cox-Aalen - tradicional.

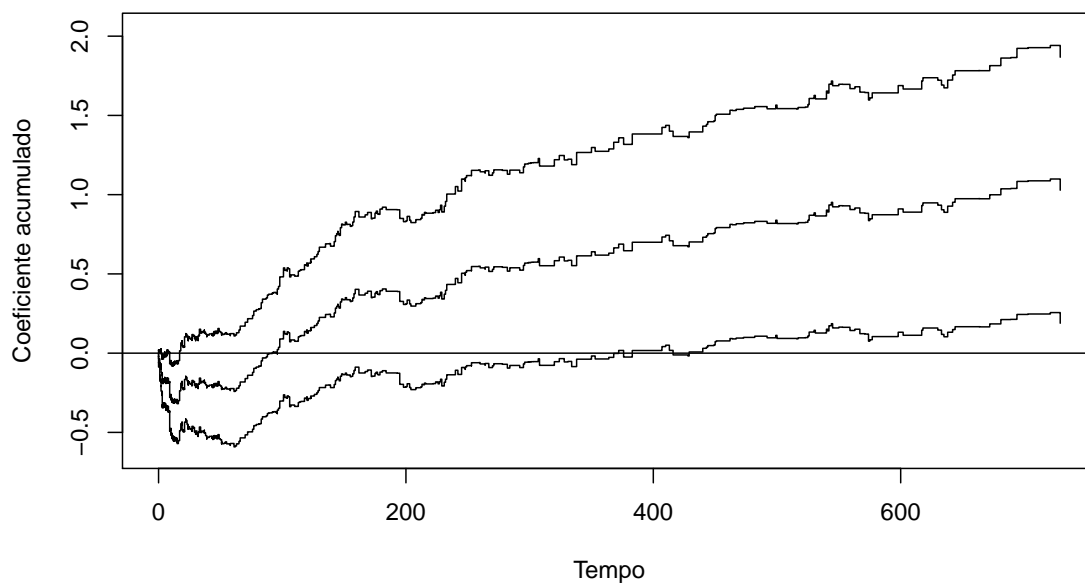


Figura B.13: Estimativa do intercepto para o modelo de Cox-Aalen - proposto.

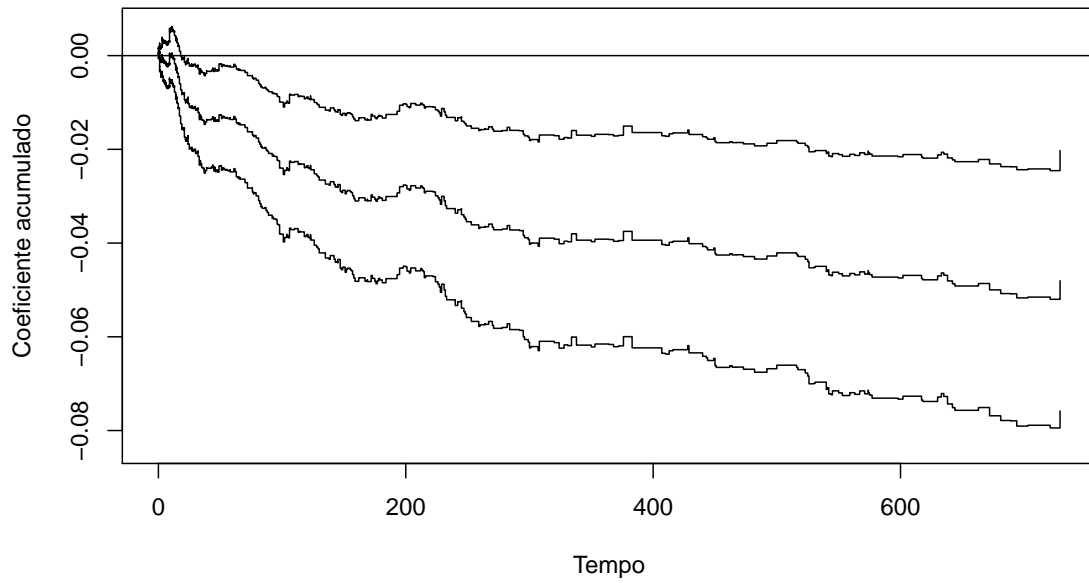


Figura B.14: Estimativa do coeficiente acumulado de IMC para o modelo de Cox-Aalen - proposto.

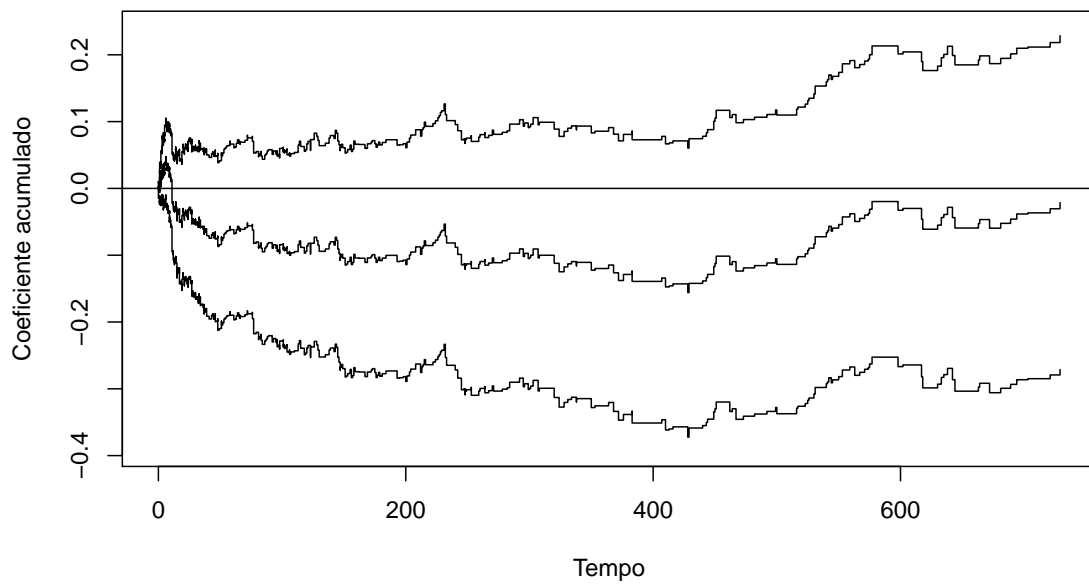


Figura B.15: Estimativa do coeficiente acumulado de Cirurgia (Sim) para o modelo de Cox-Aalen - proposto.

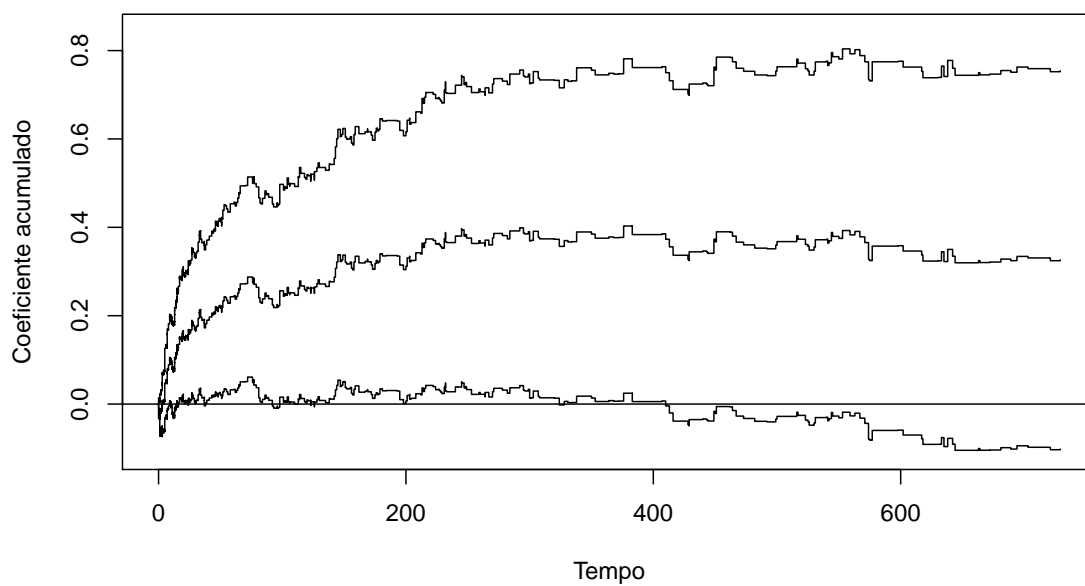


Figura B.16: Estimativa do coeficiente acumulado de Tipo de admissão (Não cirúrgico) para o modelo de Cox-Aalen - proposto.

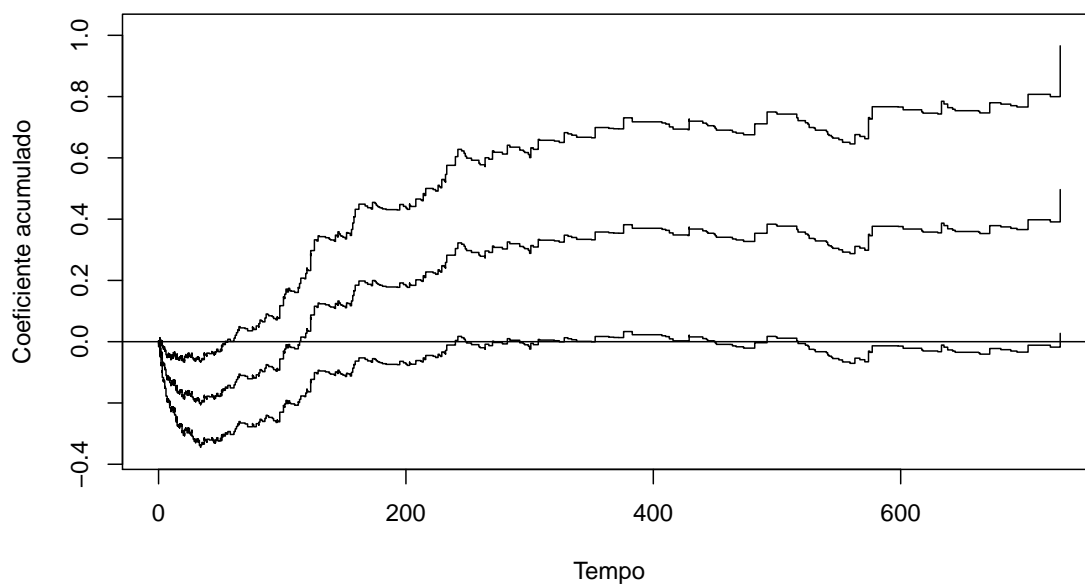


Figura B.17: Estimativa do coeficiente acumulado de Radioterapia (Sim) para o modelo de Cox-Aalen - proposto.

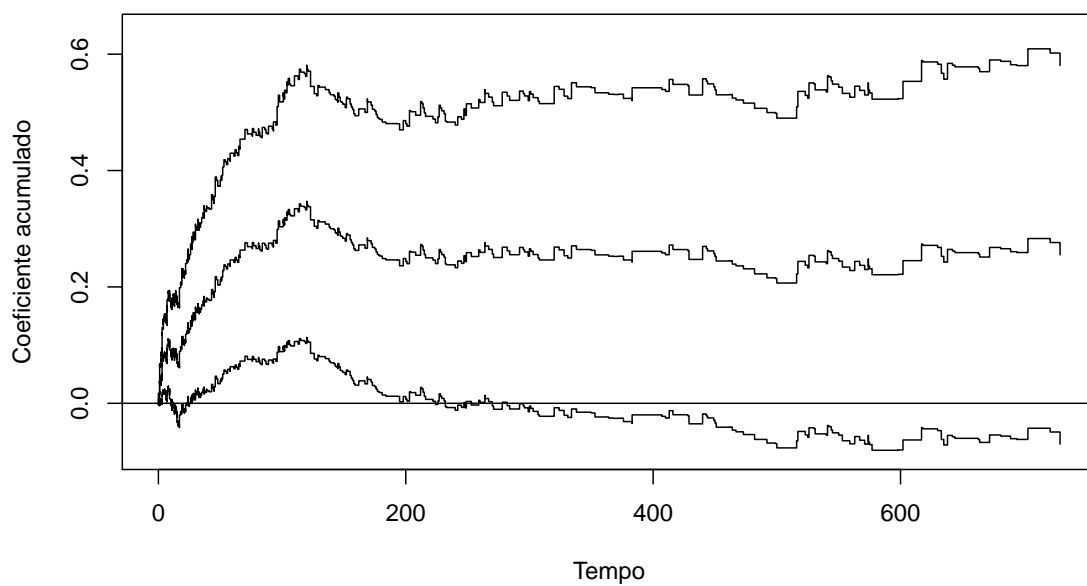


Figura B.18: Estimativa do coeficiente acumulado de Presença de delirium (Sim/sedado) para o modelo de Cox-Aalen - proposto.

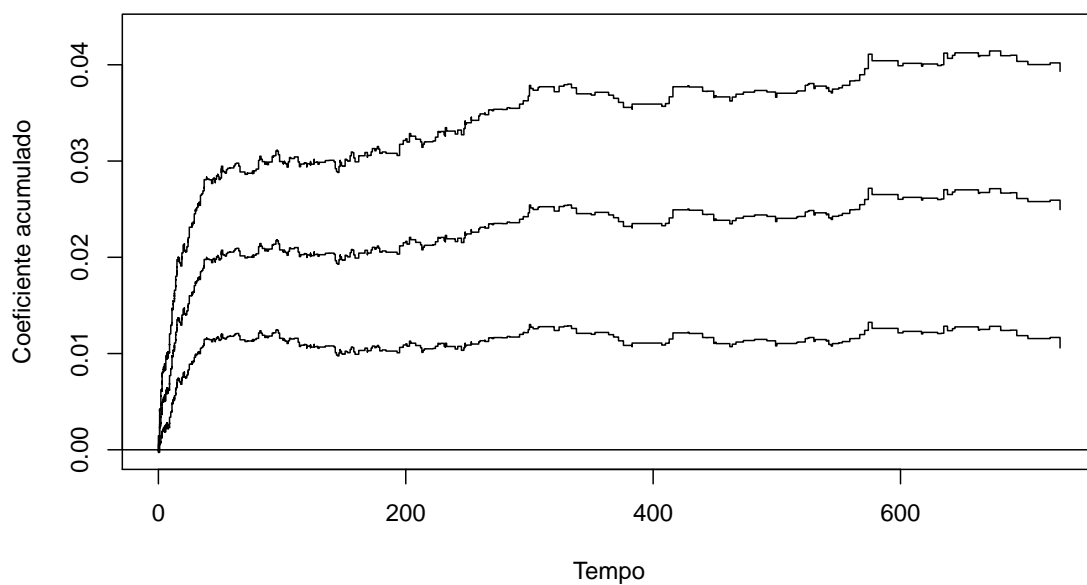


Figura B.19: Estimativa do coeficiente acumulado de Escore SAPS III para o modelo de Cox-Aalen - proposto.

Referências Bibliográficas

- Aalen (1980)** OO Aalen. A model for nonparametric regression analysis of counting processes. *Lecture notes in statistics*, 2:1–25. Citado na pág. 2, 11, 25, 26
- Akaike (1974)** Hirotugu Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723. Citado na pág. 56
- Andersen et al. (1993)** Per K Andersen, Ornnull Borgan, Richard D Gill e Niels Keiding. *Statistical models based on counting processes*. Springer. Citado na pág. 10, 35
- Aranda-Ordaz (1983)** Francisco J Aranda-Ordaz. An extension of the proportional-hazards model for grouped data. *Biometrics*, 39(1):109–117. Citado na pág. 12, 14
- Brown (1975)** Charles C Brown. On the use of indicator variables for studying the time-dependence of parameters in a response-time model. *Biometrics*, 31(4):863–872. Citado na pág. 1, 26
- Cai e Sun (2003)** Zongwu Cai e Yanqing Sun. Local linear estimation for time-dependent coefficients in cox's regression models. *Scandinavian Journal of Statistics*, 30(1):93–111. Citado na pág. 26
- Colosimo et al. (2000)** Enrico A Colosimo, Lician VAS Chalita e Clarice GB Demétrio. Tests of proportional hazards and proportional odds models for grouped survival data. *Biometrics*, 56(4):1233–1240. Citado na pág. 13
- Colosimo e Giolo (2006)** Enrico Antônio Colosimo e Suely Ruiz Giolo. Análise de sobrevivência aplicada. Em *ABE-Projeto Fisher*. Edgard Blücher. Citado na pág. 7
- Cox (1975)** David R Cox. Partial likelihood. *Biometrika*, 62(2):269–276. Citado na pág. 1, 9
- Cox e Hinkley (1979)** David R Cox e David V Hinkley. *Theoretical statistics*. CRC Press. Citado na pág. 19
- Dabrowska (1997)** Dorota M Dabrowska. Smoothed cox regression. *The Annals of Statistics*, 25(4):1510–1540. Citado na pág. 33
- Fleming e Harrington (2011)** Thomas R Fleming e David P Harrington. *Counting processes and survival analysis*, volume 169. John Wiley & Sons. Citado na pág. 58
- Gamerman (1991)** Dani Gamerman. Dynamic bayesian models for survival data. *Applied Statistics*, 40(1):63–79. Citado na pág. 26
- Grambsch e Therneau (1994)** P.M. Grambsch e T.M. Therneau. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3):515–526. Citado na pág. 1, 5, 26
- Greenwood (1926)** Major Greenwood. The natural duration of cancer. Em *Reports on Public Health and Medical Subjects*. London: HMSO. Citado na pág. 52

- Huffer e McKeague (1991)** Fred W Huffer e Ian W McKeague. Weighted least squares estimation for aalen's additive risk model. *Journal of the American Statistical Association*, 86(413):114–129. Citado na pág. 28, 34
- Khmaladze (1982)** Estate V Khmaladze. Martingale approach in the theory of goodness-of-fit tests. *Theory of Probability & Its Applications*, 26(2):240–257. Citado na pág. 29
- Lawless (2011)** Jerald F Lawless. *Statistical models and methods for lifetime data*, volume 362. John Wiley & Sons. Citado na pág. 4, 13
- Lin et al. (1993)** Danyu Y Lin, Lee-Jen Wei e Zhiliang Ying. Checking the cox model with cumulative sums of martingale-based residuals. *Biometrika*, 80(3):557–572. Citado na pág. 39
- Lin e Ying (1994)** DY Lin e Zhiliang Ying. Semiparametric analysis of the additive risk model. *Biometrika*, 81(1):61–71. Citado na pág. 9, 11
- Martinussen e Scheike (2006)** Torben Martinussen e Thomas H Scheike. *Dynamic regression models for survival data*. Springer. Citado na pág. 29, 32, 36, 39
- McKeague e Sasieni (1994)** Ian W McKeague e Peter D Sasieni. A partly parametric additive risk model. *Biometrika*, 81(3):501–514. Citado na pág. 30
- Moreau et al. (1985)** T Moreau, J O'quigley e M Mesbah. A global goodness-of-fit statistic for the proportional hazards model. *Applied Statistics*, 34(3):212–218. Citado na pág. 26
- Murphy e Sen (1991)** Susan A Murphy e Pranab K Sen. Time-dependent coefficients in a cox-type regression model. *Stochastic Processes and their Applications*, 39(1):153–180. Citado na pág. 1, 26
- Rao (1948)** Radhakrishna C Rao. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. Em *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 44, páginas 50–57. Cambridge Univ Press. Citado na pág. 15
- Rao (2009)** Radhakrishna C Rao. *Linear statistical inference and its applications*, volume 22. John Wiley & Sons. Citado na pág. 17
- Sargent (1997)** Daniel J Sargent. A flexible approach to time-varying coefficients in the cox regression setting. *Lifetime Data Analysis*, 3(1):13–25. Citado na pág. 26
- Scheike (2002)** Thomas H Scheike. The additive nonparametric and semiparametric aalen model as the rate function for a counting process. *Lifetime Data Analysis*, 8(3):247–262. Citado na pág. 29
- Scheike e Zhang (2002)** Thomas H Scheike e Mei-Jie Zhang. An additive–multiplicative cox–aalen regression model. *Scandinavian Journal of Statistics*, 29(1):75–88. Citado na pág. 2, 34, 37, 38, 55
- Schoenfeld (1982)** David Schoenfeld. Partial residuals for the proportional hazards regression model. *Biometrika*, 69(1):239–241. Citado na pág. 5
- Schwarz (1978)** Gideon Schwarz. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464. Citado na pág. 56
- Stablein et al. (1981)** Donald M Stablein, Walter H Carter e Joel W Novak. Analysis of survival data with nonproportional hazard functions. *Controlled Clinical Trials*, 2(2):149–159. Citado na pág. 26
- Taulbee (1979)** John D Taulbee. A general model for the hazard rate with covariables. *Biometrics*, 35(2):439–450. Citado na pág. 1, 26
- Winnett e Sasieni (2003)** Angela Winnett e Peter Sasieni. Iterated residuals and time-varying covariate effects in cox regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(2):473–488. Citado na pág. 26

Zucker e Karr (1990) David M Zucker e Alan F Karr. Nonparametric survival analysis with time-dependent covariate effects: a penalized partial likelihood approach. *The Annals of Statistics*, páginas 329–353. Citado na pág. [1](#), [26](#)