Avaliação de funções Kernel no modelo de Cox com efeitos dinâmicos

Milena de Souza Reis

Dissertação Apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do grau
de
Mestre em Ciências

Área de Concentração: Estatística

Orientador: Prof. Dr. Antonio Carlos Pedroso de Lima

São Paulo, abril de 2008

Avaliação de funções Kernel no modelo de Cox com efeitos dinâmicos

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Milena de Souza Reis
e aprovada pela Comissão Julgadora.

Banca Examinadora:

- Prof. Dr. Antonio Carlos Pedroso de Lima (orientador) - IME-USP.
- Profa. Dra. Clélia Maria de Castro Toloi - IME-USP.
- Prof. Dr. Edwin Moises Marcos Ortega - ESALQ-USP.

Agradecimentos

Agradeço a todos que contribuíram para a elaboração deste trabalho, em especial:

Ao Prof. Antonio Carlos, pela orientação, paciência e disposição em me ajudar durante o desenvolvimento do trabalho.

Aos meus pais Rute e Edmar, que sempre me incentivaram e possibilitaram a realização de meus objetivos.

Ao Alessandro Rocha, pelos momentos felizes e pelo apoio durante esse período de dedicação aos estudos.

À Darcymara Morais, que me apoiou e me permitiu ausentar da empresa para me dedicar ao mestrado.

Aos professores do Departamento de Estatística do IME-USP que muito contribuíram para minha formação.

Resumo

O modelo de regressão de Cox é extensamente utilizado em estudos nos quais o objetivo é analisar a relação entre as covariáveis e o tempo até a ocorrência de um evento de interesse. O modelo de riscos proporcionais assume que os coeficientes da regressão são constantes. Entretanto, verificamos em diferentes conjuntos de dados que essa suposição não é satisfeita, isto é, as covariáveis do modelo apresentam efeitos no tempo.

Neste trabalho apresentamos uma metodologia para a análise do modelo de Cox com coeficientes dependentes do tempo. Os coeficientes são estimados através da suavização da função de verossimilhança parcial ponderada por uma função *Kernel*. Os métodos apresentados são ilustrados através de um exemplo de dados reais e de uma simulação.

Abstract

The Cox regression model is widely used in analyses where the purpose is to evaluate the relation between covariates and time until a studied event. The proportional hazards model assumes that the regression coefficients are time invariant. However, we can verify in different datas that this assumption is not satisfied. That is, the covariates present temporal effects.

In this work we present a methodology to analyse the Cox model with time-varying coefficients. The coefficients are estimated through a weighted partial likelihood function. All the methods showed are ilustrated with a real example and a simulation.

Sumário

Capítulo 1

Introdução

A análise de sobrevivência tem como objetivo estudar o tempo até a ocorrência de um evento de interesse. O termo sobrevivência é utilizado predominantemente na área médica, na qual são estudados os tempos até a morte do paciente, até sua cura ou até a recidiva de uma determinada doença.

As técnicas de análise de sobrevivência também são empregadas em outras áreas como, por exemplo, ciências sociais, onde o objetivo é analisar a duração de eventos de desemprego, casamentos e migrações. A engenharia tem auxiliado no desenvolvimento da técnica através da chamada Análise de Confiabilidade. Nessa área, o principal interesse é estudar o tempo de falha de máquinas e componentes eletrônicos, de forma a obter informações sobre a durabilidade dos produtos.

Outros exemplos da aplicação da técnica podem ser encontrados em instituições financeiras, nas quais o modelo de sobrevivência é utilizado para avaliar, por exemplo, o tempo até o cancelamento de cartões de crédito. Considerando que o mercado está cada vez mais competitivo, tornou-se fundamental para as empresas identificar clientes propensos ao cancelamento de serviços (cartões, planos, assinaturas etc) para que sejam realizadas campanhas de retenção e fidelização.

Existem determinados aspectos nos conjuntos de dados que favorecem a utilização de modelos de sobrevivência, como dados censurados e não-normalidade. A presença de dados censurados à direita indica que toda a informação referente à variável resposta se resume ao conhecimento de que o tempo de falha é superior àquele observado.

O procedimento mais utilizado para modelar a relação entre as covariáveis e o tempo até a ocorrência do evento estudado (variável resposta) é o modelo de riscos proporcionais de Cox, no qual a razão das taxas de falha de dois indivíduos diferentes é constante no tempo.

O modelo de riscos proporcionais de Cox assume que os coeficientes da regressão são constantes. Entretanto, muitas vezes verificamos que os parâmetros variam e o efeito das covariáveis no tempo de falha pode ser de grande interesse. Por exemplo, em um estudo médico que compara um novo tratamento com o usual, pode ser verificado que a nova droga apresenta bons resultados no início do tratamento, mas perde sua eficácia ao longo do tempo.

O objetivo do modelo de Cox com coeficientes dependentes do tempo é analisar situações como a descrita acima, em que os coeficientes da regressão não são constantes. Esse modelo tem auxiliado principalmente em estudos médicos, nos quais o principal objetivo é comparar o efeito de diferentes drogas e suas eficácias durante os tratamentos.

Dada a ampla utilização e necessidade de estudos do efeito temporal das covariáveis, é apresentada nesta dissertação uma proposta para estimação dos coeficientes dependentes do tempo para o modelo de Cox. O procedimento utilizado para estimação é baseado na suavização da função de verossimilhança parcial, ponderada por uma função *Kernel* (Tian et al, 2005).

Em seguida, são expostas idéias para estimação de bandas de confiança e também para estimação da função de sobrevivência, que pode ser aplicada para um indivíduo com um determinado conjunto de valores para as covariáveis presentes no modelo.

Todas as técnicas discutidas são ilustradas com o conjunto de dados de cirrose biliar primária da Mayo Clinic. Além da aplicação aos dados de cirrose, é apresentada uma simulação com o objetivo de verificar qual função *Kernel* produz as melhores estimativas dos coeficientes, de acordo com o comportamento da razão das taxas de falha (crescente ou decrescente).

O trabalho é desenvolvido na seguinte seqüência: no Capítulo 2 é descrito o conjunto de dados utilizado. No Capítulo 3 são apresentados os modelos de Cox de riscos proporcionais e com coeficientes dependentes do tempo.

Em seguida, no Capítulo 4, é exposto um resumo da função *Kernel* e são desenvolvidas as estimativas para os coeficientes e para as bandas de confiança. Os resultados da simulação e do ajuste do modelo de Cox com coeficientes dependentes do tempo são apresentados no Capítulo 5. E, por fim, as conclusões e trabalhos futuros são descritos no Capítulo 6.

Capítulo 2

Motivação

2.1 Dados de cirrose biliar - Mayo Clinic

O conjunto de dados analisado consiste de informações de cirrose biliar primária (PBC) coletadas na Mayo Clinic no período de Janeiro de 1974 a Maio de 1984 (Fleming e Harrington, 1991). A doença estudada é uma doença crônica fatal do fígado de causa desconhecida. Devido ao fato da PBC ser uma doença rara (prevalência estimada de 50 casos para milhão de pessoas), o conjunto de dados estudado é de extremo valor para a área médica. O objetivo do estudo é avaliar o efeito da droga *D-penicillamine (DPCA)*.

O tempo de sobrevivência de 312 pacientes como também dados de outras covariáveis de interesse foram acompanhados pelo período de 10 anos. Dentre as informações que foram coletadas dos pacientes podemos citar: idade, sexo, bilirrubina (mg/dl), presença de edema, triglicérides (mg/dl), albumina (mg/dl) etc.

No período de estudo, dos 312 pacientes analisados, 19 foram submetidos ao transplante de fígado e 125 morreram em decorrência da doença. Ao término das análises, foi verificado que a droga *D-penicillamine (DPCA)* não apresentava resultados eficazes no tratamento da cirrose biliar primária. Isto é, não aumentava o tempo de sobrevivência dos pacientes

tratados com a droga quando comparado com o grupo de placebo. Nos anos 80, o índice de transplantes de fígado realizados com sucesso cresceu significativamente, e portanto, o transplante tornou-se uma alternativa para a PBC.

Fleming e Harrington (1991) ajustaram o modelo de Cox de riscos proporcionais para este conjunto de dados e identificaram as seguintes variáveis como importantes preditoras para o modelo final: idade, log(albumina), log(bilirrubina), edema e log(protrombina).

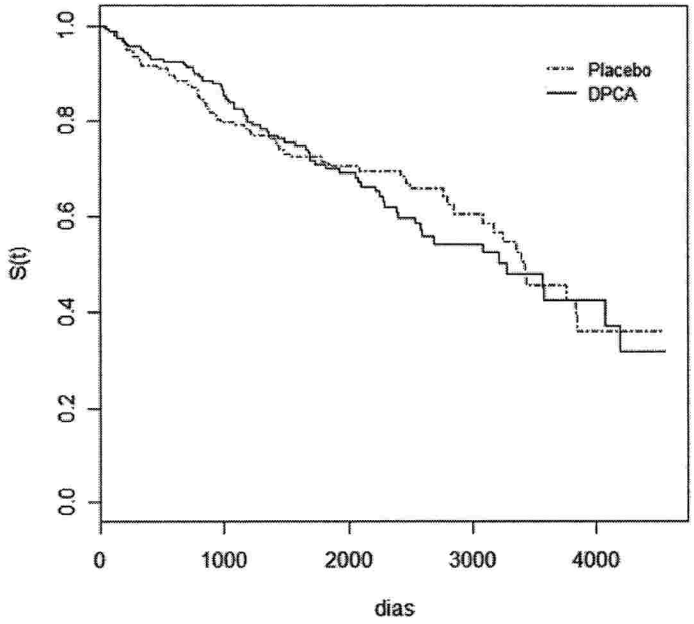
Com o auxílio dos gráficos dos resíduos de Schoenfeld, Fleming e Harrington (1991) também indicaram que as variáveis edema e log(protrombina) não satisfaziam a suposição de riscos proporcionais. A seguir serão apresentados os resultados para o ajuste baseado no modelo de Cox.

2.2 Análise dos dados

A primeira análise foi realizada com o objetivo de verificar se a droga DPCA aumentava o tempo de sobrevivência dos pacientes. Para isto, foi utilizado o modelo de Cox de riscos proporcionais com a covariável X, para a qual $X = 1$ corresponde ao tratamento e $X = 0$ ao placebo.

As funções de sobrevivência estimadas são apresentadas no gráfico abaixo:

Gráfico 2.1: Funções de sobrevivência



Podemos observar que, ao longo do tempo, a separação das curvas de sobrevivência dos dois grupos é bem pequena, o que mostra que a droga *D-penicillamine* não apresenta resultados satisfatórios no tratamento da cirrose biliar primária.

Os resultados a seguir apresentam o ajuste do modelo de Cox de riscos proporcionais com as covariáveis consideradas significativas por Fleming e Harrington (1991):

Tabela 2.1: Estimativa dos coeficientes

Variável	coef	exp(coef)	se(coef)	z	p
idade	0.0332	1.0338	0.00865	3.84	1.2e-04
log(bilirrubina)	0.8801	2.4110	0.09874	8.91	0.0e+00
log(albumina)	-3.0599	0.0469	0.72404	-4.23	2.4e-05
edema	0.7859	2.1943	0.29897	2.63	8.6e-03
log(protrombina)	3.0140	20.3690	1.02395	2.94	3.2e-03

Variável	exp(coef)	exp(-coef)	lower .95	upper .95
idade	1.0338	0.9673	1.0164	1.051
log(bilirrubina)	2.4110	0.4148	1.9868	2.926
log(albumina)	0.0469	21.3260	0.0113	0.194
edema	2.1943	0.4557	1.2213	3.943
log(protrombina)	20.3690	0.0491	2.7377	151.549

Rsquare = 0.472
Likelihood ratio test = 199 on 5 df, p = 0
Wald test = 198 on 5 df, p = 0
Score (logrank) test = 270 on 5 df, p = 0

De acordo com os autores, o modelo obtido acima é biologicamente interpretável. O coeficiente negativo para a covariável log(albumina) é consistente com o fato de que com o avanço da PBC, a capacidade do fígado produzir albumina diminui.

Além disso, podemos verificar que a covariável que mais influencia no risco de morte é a log(protrombina).

Para avaliar a suposição de riscos proporcionais do modelo de Cox são propostas duas abordagens: coeficiente de correlação de Pearson e método gráfico.

O coeficiente de correlação de Pearson (ρ) entre os resíduos padronizados de Schoenfeld e uma função do tempo $g(t)$, dentre elas t ou $log(t)$, para cada uma das covariáveis pode ser utilizado como uma medida para verificar se a suposição de riscos proporcionais do modelo está satisfeita.

Tabela 2.2: Verificação da suposição de riscos proporcionais

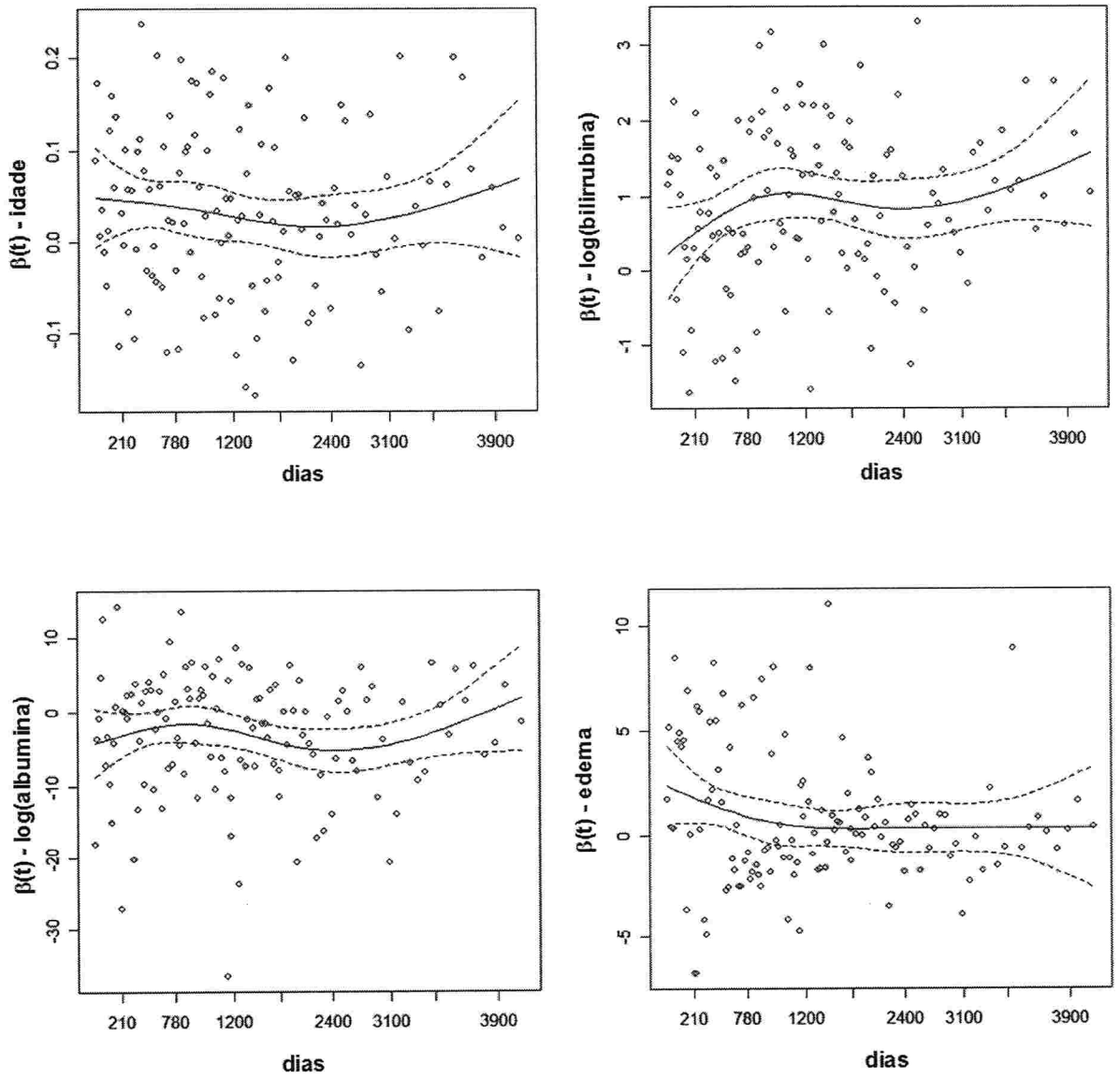
Variável	ρ	χ^2	p-valor
idade	-0.0364	0.1448	0.7036
log(bilirrubina)	0.1413	2.3266	0.1272
log(albumina)	-0.0196	0.0502	0.8227
edema	-0.1481	2.6540	0.1033
log(protrombina)	-0.1792	3.2523	0.0713
Global	NA	8.9292	0.1119

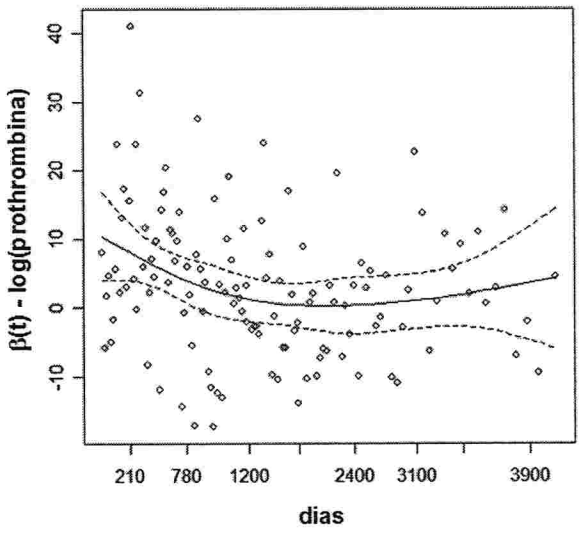
Considerando o nível de significância de 10%, verificamos que há evidências de que as variáveis edema e log(protrombina) não satisfazem a suposição de riscos proporcionais.

Desta forma, é possível o ajuste do modelo de Cox com coeficientes dependentes do tempo aos dados de cirrose biliar primária.

A seguir são apresentados os gráficos de resíduos de Schoenfeld. Para que a suposição de riscos proporcionais seja válida, para cada uma das covariáveis, o gráfico de $\beta(t)$ versus t deve ser uma linha horizontal. A técnica gráfica envolve conclusões subjetivas pois depende da interpretação dos gráficos, portanto, os resultados devem ser analisados com cautela.

Gráfico 2.2: Estimativas de $\beta(t)$ em função de t





Para o desenvolvimento do modelo consideramos satisfeita a suposição de riscos proporcionais. Entretanto, podemos observar que há evidências de que as covariáveis edema e log(protrombina) não satisfazem essa suposição.

Portanto, considerando a não adequabilidade da suposição de riscos proporcionais e a possibilidade da obtenção de melhores estimativas para o modelo, será apresentada neste trabalho uma técnica de ajuste do modelo de Cox com coeficientes dependentes do tempo. Desta forma, será possível avaliar o efeito temporal de cada uma das covariáveis.

Capítulo 3

Modelo de Cox

Neste capítulo é descrita a técnica de análise de sobrevivência, que é uma das técnicas estatísticas que mais se desenvolveram nos últimos anos, com ampla aplicação principalmente na área médica. Conforme mencionado anteriormente, em análise de sobrevivência a variável resposta é o tempo até a ocorrência de um evento de interesse (tempo de falha).

Uma característica da técnica é a presença de dados censurados. Nessas ocasiões, por alguma razão não associada à análise (mudança de cidade, morte por outro motivo não relacionado ao estudo etc), o acompanhamento do paciente é interrompido. Isso significa que toda informação disponível do indivíduo se concentra no fato de que o tempo de falha é superior ao tempo observado.

Um importante objetivo na área médica é avaliar a relação de diferentes covariáveis coletadas do paciente com o tempo de sobrevivência. Para isso, a forma mais eficiente para a análise dos dados é o emprego de um modelo de regressão apropriado para acomodar dados censurados.

3.1 Taxas de falha proporcionais

O modelo de regressão de Cox com riscos proporcionais (Cox, 1972) é extensamente utilizado para modelar a relação de covariáveis com o tempo de falha. A suposição básica para o uso do modelo é que as taxas de falha sejam proporcionais, ou seja, a razão das taxas de falha de dois indivíduos diferentes é constante no tempo.

A análise de resíduos, como por exemplo, de Schoenfeld é usualmente utilizada para verificar a suposição de riscos proporcionais. Se estiver satisfeita, o gráfico de $\boldsymbol{\beta}(t)$ versus $g(t)$ deve ser uma linha horizontal. Maiores detalhes podem ser encontrados em Schoenfeld (1982).

Para desenvolver o modelo de Cox, seja $\boldsymbol{X}_i$ o vetor de covariáveis do indivíduo i, para $i = 1, ..., n$. A função de taxa de falha para o indivíduo i é dada por:

$$\alpha_i(t) = \alpha_0(t)e^{\boldsymbol{X}_i'\boldsymbol{\beta}}$$

em que, α_0 é uma função não-negativa chamada de taxa de falha padrão e $\boldsymbol{\beta}$ é um vetor coluna de dimensão p, cujos elementos são os coeficientes da regressão.

Devido ao fato de que a razão das taxas de falha de dois indivíduos com vetores fixos de covariáveis $\boldsymbol{X}_i$ e $\boldsymbol{X}_j$ é constante no tempo,

$$\frac{\alpha_i(t)}{\alpha_j(t)} = \frac{\alpha_0(t)e^{\boldsymbol{X}_i'\boldsymbol{\beta}}}{\alpha_0(t)e^{\boldsymbol{X}_j'\boldsymbol{\beta}}} = \frac{e^{\boldsymbol{X}_i'\boldsymbol{\beta}}}{e^{\boldsymbol{X}_j'\boldsymbol{\beta}}} = e^{\boldsymbol{\beta}(\boldsymbol{X}_i' - \boldsymbol{X}_j')}$$

o modelo é denominado modelo de riscos proporcionais de Cox.

Para a estimação dos coeficientes $\boldsymbol{\beta}$'s que medem os efeitos das covariáveis, utiliza-se em geral o método de máxima verossimilhança parcial, baseado na chamada função de verossimilhança parcial. Para descrever essa verossimilhança, considere uma amostra de n indivíduos com vetor de covariáveis $\boldsymbol{X}_i$, $i = 1, ..., n$. Admita que $t_1, ..., t_n$ sejam os instantes

de falha ou censura observados e $\delta_1, ..., \delta_n$ os correspondentes indicadores de falha, isto é $\delta_i = 1$, se falha ou $\delta_i = 0$, se censura.

A verossimilhança parcial (Colosimo, E. e Giolo, S., 2006) é dada por:

$$\boldsymbol{L}(\boldsymbol{\beta}) = \prod_{i=1}^{n} \left(\frac{e^{\boldsymbol{X}_i'\boldsymbol{\beta}}}{\sum_{j \in R(t_i)} e^{\boldsymbol{X}_j'\boldsymbol{\beta}}} \right)^{\delta_i}$$

em que $R(t_i)$ é o conjunto dos índices de observações em risco em t_i^- (o instante imediatamente anterior a t). Entendemos por observações em risco aquelas não censuradas e que não apresentaram o evento de interesse até t^-.

Os valores de $\boldsymbol{\beta}$ que maximizam a função de verossimilhança parcial $\boldsymbol{L}(\boldsymbol{\beta})$ são obtidos a partir do sistema de equações definido por $\boldsymbol{U}(\boldsymbol{\beta}) = 0$, em que $\boldsymbol{U}(\boldsymbol{\beta})$ é a função score apresentada a seguir

$$\log \boldsymbol{L}(\boldsymbol{\beta}) = \sum_{i=1}^{n} \delta_i \left[\boldsymbol{X}_i\boldsymbol{\beta} - \log \left(\sum_{j \in R(t_i)} e^{\boldsymbol{X}_j\boldsymbol{\beta}} \right) \right]$$

$$\boldsymbol{U}(\boldsymbol{\beta}) = \frac{\partial \log \boldsymbol{L}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \delta_i \left[\boldsymbol{X}_i - \frac{\sum_{j \in R(t_i)} \boldsymbol{X}_j e^{\boldsymbol{X}_j'\hat{\boldsymbol{\beta}}}}{\sum_{j \in R(t_i)} e^{\boldsymbol{X}_j'\hat{\boldsymbol{\beta}}}} \right] = 0.$$

3.2 Coeficientes dependentes do tempo

O modelo de Cox de riscos proporcionais assume que os coeficientes da regressão são constantes no tempo. Entretanto, muitas vezes verificamos que esses parâmetros variam, e o efeito temporal das covariáveis nos tempos de falha pode ser de interesse.

Considere T^0 o tempo de falha, C a variável de censura e $\boldsymbol{X}$ um vetor de p covariáveis. Seja $\{(T_i^0, C_i, \boldsymbol{X}_i), i = 1, ..., n\}$ n réplicas independentes de $(T^0, C, \boldsymbol{X}(\cdot))$. Para o i-ésimo indivíduo, o que de fato observamos é $\{(T_i, \delta_i, \boldsymbol{X}_i)\}$, em que $T_i = min(T_i^0, C_i)$ e $\delta_i = 1$ se $T_i = T_i^0$ e 0 caso contrário. A função de taxa de falha do modelo de Cox com coeficientes dependentes do tempo pode ser definida como:

$$\alpha_i(t) = \alpha_0(t) e^{\boldsymbol{X}_i' \boldsymbol{\beta}(t)}$$

com $\boldsymbol{\beta}(t)$ uma função de t e $\alpha_0(t)$ a função de taxa de falha padrão.

Estamos interessados em estimar a função de coeficientes $\{\boldsymbol{\beta}(t), t > 0\}$ baseado em $\{(T_i, \delta_i, \boldsymbol{X}_i), i = 1, ..., n\}$.

A seguir serão apresentadas, resumidamente, diferentes abordagens da estimação dos coeficientes dependentes do tempo do modelo de Cox. Zucker e Karr (1990) propõem a utilização de uma função de verossimilhança parcial penalizada. A função é composta por dois termos: o primeiro termo corresponde à função de verossimilhança do modelo de Cox de riscos proporcionais. O segundo termo refere-se à uma função de penalidade, que possui o objetivo de reduzir a variância das estimativas obtidas.

Em Verweij e van Houwelingen (1995) é discutida uma proposta não-paramétrica para a estimação dos coeficientes do modelo. O vetor de coeficientes $\boldsymbol{\beta}(t)$ é substituído por uma função definida no domínio de tempo discreto. Para cada instante de tempo especificado, um vetor de parâmetros é utilizado para medir o efeito das covariáveis. Além disso, é

utilizada uma função de penalidade que, como em Zucker e Karr (1990) é subtraída do logarítmo da função de verossimilhança. Verweij e van Houwelingen (1995) também discutem sobre a escolha do parâmetro de suavização da função de penalidade, que deve ser escolhido de acordo com o critério de AIC (*Akaikes´s Information Criterion*).

Cai e Sun (2003) desenvolveram uma técnica de verossimilhança local parcial para estimar os coeficientes dependentes do tempo do modelo de Cox. A idéia central é a expansão da técnica de ajuste utilizada em gráficos de dispersão suavizados, combinada com a função de verossimilhança parcial. Em uma janela em torno de cada instante de tempo, os coeficientes $\boldsymbol{\beta}(t)$ devem ser aproximados por uma função linear utilizando a expansão de Taylor de primeira ordem. Desta forma é possível obter uma estimativa para a verossimilhança parcial. O valor estimado da função linear no tempo t será a estimativa dos coeficientes suavizados em t.

Outras abordagens para estimação dos coeficientes do modelo podem ser encontradas em Winnett e Sasieni (2003), que apresentam estimativas não-paramétricas para os coeficientes baseadas nos resíduos de Schoenfeld. A idéia central é ajustar o modelo de Cox de riscos proporcionais e calcular as estimativas de $\boldsymbol{\beta}(t)$ através da proposta de Grambsch e Therneau (1994) de que $\beta_j(t) \approx \hat{\beta}_j + E(s_j^*)$, em que $\hat{\beta}_j$ são as estimativas do modelo de Cox de riscos proporcionais e s^* são os resíduos de Schoenfeld.

O modelo de Cox com coeficientes dependentes do tempo também foi estudado por outros autores. Dentre as principais referências podemos citar: Hastie e Tibshirani (1993), Martinussen et al (2002) e Murphy e Sen (1991).

Na tentativa de contribuir para o tema, este trabalho se propõe a apresentar uma estrutura para o modelo de Cox com coeficientes dependentes do tempo, utilizando como base o artigo de Tian et al (2005).

Para o processo de estimação dos coeficientes, será apresentada uma técnica de verossimilhança parcial ponderada por uma função *Kernel*. A função *Kernel* é utilizada para

suavizar a contribuição de pontos remotos presentes nos dados e conseqüentemente reduzir a variância das estimativas. Outras referências da técnica que será discutida podem ser encontradas em Valsecchi, Silvestri e Sasieni (1996) e em Cai e Sun (2003).

Resumidamente, a técnica se propõe a obter estimativas dos parâmetros através da maximização do logarítmo da função de verossimilhança parcial a cada instante de tempo t fixado. Como é de interesse fazer inferências sobre uma função no tempo, também são apresentadas bandas de confiança para examinar o efeito temporal das covariáveis. Para tal análise, é utilizada a técnica de *strong approximation* apresentada por Bickel e Rosenblatt (1973) e Yandell (1983).

Além disso, é proposto um método de estimação da função de sobrevivência para indivíduos com um determinado conjunto de valores para as covariáveis presentes no modelo. Todas as propostas apresentadas serão ilustradas com o conjunto de dados de cirrose biliar primária da Mayo Clinic.

Na seqüência, uma simulação é apresentada para se estudar qual função *Kernel* apresenta as melhores estimativas para os coeficientes da regressão, de acordo com o comportamento da razão das taxas de falha.

Capítulo 4

Estimação dos Parâmetros

A abordagem mais popular para estudar modelos de sobrevivência envolve o uso de ferramentas de processos estocásticos, mais especificamente, processos de contagem. Neste capítulo discutimos a estimação de parâmetros de interesse, sob a ótica de tais processos. Iniciamos introduzindo a linguagem e a notação nos moldes de Andersen et al (1993).

4.1 Processos de contagem

Considere a situação em que n indivíduos são observados ao longo do tempo até a ocorrência de algum evento de interesse. Sejam $T_1^0, T_2^0, ..., T_n^0$ os tempos de falha desses indivíduos e $C_1, C_2, ..., C_n$ os respectivos tempos de censura, variáveis aleatórias não-negativas. Na prática, o que efetivamente se observa pode ser representado por

$$T_i = min(T_i^0, C_i) \quad \text{e}$$

$$\delta_i = I\{T_i = T_i^0\}$$

para $i = 1, ..., n$ em que δ_i é o indicador de falha definido anteriormente.

A função de taxa de falha é definida por

$$\alpha(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}.$$

Essa função pode ser interpretada como a taxa instantânea de falha no instante t.

Também podemos definir a função de taxa de falha acumulada por

$$A(t) = \int_0^t \alpha(s) ds.$$

A função de sobrevivência é definida como a probabilidade de um indivíduo não falhar até o instante t e pode ser expressa por

$$S(t) = P(T > t) = \exp\left\{ - \int_0^t \alpha(s) ds \right\}.$$

Logo, temos que $S(t) = e^{-A(t)}$.

As análises a seguir são ilustradas sob a abordagem de processos de contagem. O desenvolvimento de modelos estatísticos baseados em processos de contagem para analisar dados de sobrevivência foi originalmente introduzido por Aalen (1978) e depois sistematizado por Andersen et al (1993).

Esta abordagem proporciona ferramentas bastante poderosas que são capazes de generalizar diversas situações importantes na análise de dados de sobrevivência.

Um processo de contagem pode ser entendido como um conjunto de variáveis aleatórias indexadas no tempo $\{N(t) : t \geq 0\}$, tal que

(i) $N(t) = 0$, $t = 0$;

(ii) $N(t) < \infty$;

(iii) o processo é contínuo à direita;

(iv) $N(t)$ tem apenas descontinuidades de tamanho 1.

O exemplo em que n indivíduos são acompanhados ao longo do tempo até a ocorrência de um evento de interesse pode ser convenientemente representado por um processo de contagem multivariado $\boldsymbol{N}(t) = \{N_1(t), N_2(t), ..., N_n(t)\}$, sendo

$$N_i(t) = I\{T_i \leq t, \delta_i = 1\}, \quad i = 1, ..., n.$$

Seja $dN(t) = N(t) - N(t - dt)$ o incremento do processo $N(t)$ num intervalo de comprimento dt. Observe que $N(t)$ e $N(t - dt)$ só serão diferentes quando ocorrer uma falha no intervalo $(t - dt, t]$ e então

$$dN(t) = \begin{cases} 1, & \text{se falha em } (t - dt, t] \\ 0, & \text{caso contrário} \end{cases}$$

Por último, podemos definir a variável $Y_i(t) = I\{T_i \geq t\}$ que indica se o *i-ésimo* indivíduo está ou não em risco no instante t^-. Logo, $Y(t) = \sum_{i=1}^{n} Y_i(t)$ representa o total de indivíduos em risco no instante t^-.

Na teoria de processos de contagem, dizemos que o comportamento ao longo do tempo de um processo $N(t)$ é controlado por seu processo de intensidade $\lambda(t)$. Dada a história anterior, o processo $\lambda(t)dt$ é definido pela probabilidade condicional de que $N(t)$ salte em um pequeno intervalo de comprimento dt.

De acordo com Andersen et al (1993), o processo de intensidade de $N(t)$ é dado por $\lambda(t) = \alpha(t)Y(t)$, o que corresponde ao *modelo de intensidade multiplicativo.*

Da mesma forma que na função de taxa de falha acumulada, temos que o processo de intensidade acumulado é definido por

$$\Lambda(t) = \int_0^t \lambda(s)ds.$$

Com as definições apresentadas anteriormente, temos os conceitos necessários para a continuação do trabalho.

Em estudos envolvendo dados de sobrevivência não é possível realizar uma análise descritiva usual, pois a presença de censuras invalida esse tipo de tratamento aos dados. Desta forma, as medidas de interesse para o estudo que devem ser estimadas são as funções de taxa de falha e de sobrevivência.

Um dos estimadores que pode ser empregado nas análises é o estimador de Nelson-Aalen. Este estimador baseia-se no fato de que a função de sobrevivência é função da taxa de falha acumulada.

A seguir o estimador de Nelson-Aalen é definido e estudado, e é apresentada uma forma para obtenção de uma estimativa de $\alpha(t)$. Andersen et al (1993) apresenta um estudo detalhado sobre esse estimador, descrevendo suas propriedades.

Considere um intervalo contínuo $\mathcal{T}$ da forma $[0, \tau]$ ou $[0, \tau)$ para um determinado tempo τ, $0 < \tau < \infty$. Seja o processo de contagem multivariado $\boldsymbol{N}(t) = ((N_1(t), \ldots, N_n(t)), t \in \mathcal{T})$, satisfazendo o modelo multiplicativo, isto é, o processo de intensidade é dado por:

$$\lambda_h(t) = \alpha_h(t)Y_h(t), \qquad h = 1, ..., n$$

com $\alpha_h(\cdot)$ e $Y_h(\cdot)$ definidos anteriormente.

Para derivar heuristicamente um estimador para $A_h(t) = \int_0^t \alpha_h(s)ds$, utilizamos o processo

$$M_h(t) = N_h(t) - \int_0^t \alpha_h(s)Y_h(s)ds,$$

de tal forma que

$$dN_h(t) = \alpha_h(t)Y_h(t)dt + dM_h(t)$$

em que $dM_h(t)$ pode ser considerado como um ruído aleatório. Em Andersen et al (1993) podemos encontrar maiores detalhes de $M(t)$, que é definido como um processo martingal de média zero.

Desta forma, um estimador para $A_h(t)$ é:

$$\hat{A}_h(t) = \int_0^t Y_h^{-1}(s)dN_h(s)$$

A seguir é apresentado um método não-paramétrico muito utilizado para estimação destas quantidades através da suavização do estimador de Nelson-Aalen com o auxílio da função *Kernel*.

4.2 Função *Kernel*

A função *Kernel* apresenta extensa utilização na suavização de funções de verossimilhança para a obtenção de estimativas mais suaves e que não apresentem a interferência de pontos aberrantes eventualmente presentes nos dados. Um estimador de $\alpha(t)$ pode ser obtido através da suavização dos incrementos do estimador de $\hat{A}(\cdot)$ e pode ser definido

como:

$$\hat{\alpha}(t) = \frac{1}{b}\int_{\mathcal{T}} K\left(\frac{t-s}{b}\right) d\hat{A}(s)$$

A função *Kernel* $K(\cdot)$ é uma função limitada, com suporte em [-1, 1] e integral 1. A janela b é um parâmetro positivo que deve ser escolhido de acordo com a aplicação. Um *Kernel* freqüentemente utilizado é o *Kernel* Epanechnikov, que será descrito adiante. Maiores detalhes podem ser encontrados em Staniswalis (1989) e Andersen et al (1993).

Se denotarmos $Z_1 < Z_2 < \cdots$ os sucessivos instantes de salto do processo de contagem $\boldsymbol{N}(\cdot)$, então $\hat{\alpha}(t)$ é expressa como:

$$\hat{\alpha}(t) = \frac{1}{b}\sum_j K\left(\frac{t-Z_j}{b}\right)\frac{1}{Y(Z_j)}$$

onde $t - b \leq Z_j \leq t + b$

Podemos notar que $\hat{\alpha}(t)$ é uma média ponderada dos incrementos $Y(Z_j)^{-1}$ do estimador de Nelson-Aalen no intervalo $[t-b, t+b]$.

Conforme comentado anteriormente, $K(\cdot)$ é qualquer função que satisfaça as propriedades descritas acima. Algumas funções são mais populares, entre elas:

- *Kernel* Uniforme, caracterizada por

$$K_U(x) = \frac{1}{2}, \qquad -1 \leq x \leq 1$$

- *Kernel* Epanechnikov, dada por

$$K_E(x) = \frac{3}{4}(1-x^2), \qquad -1 \leq x \leq 1$$

- *Kernel* Bi-ponderada, expressa por

$$K_B(x) = \frac{15}{16}(1 - x^2)^2, \qquad -1 \leq x \leq 1$$

- *Kernel* Triangular, que é escrita como

$$K_T(x) = 1 - |x|, \qquad -1 \leq x \leq 1$$

Maiores detalhes referentes à função *Kernel* também podem ser encontrados em Wand e Jones (1995) e Silverman (1986).

Um ponto importante é que as estimativas dependem da escolha da janela b, isto é, quanto maior a janela, mais suavizadas serão as estimativas, o que implica em valores cada vez mais viciados.

Um método frequentemente citado na literatura para a escolha da janela de suavização é o método da validação cruzada (Efron e Tibshirani, 1993), que será descrito na próxima seção. Existem outras abordagens para a definição do parâmetro de suavização, como por exemplo, Estatística C_p e Bootstrap. O objetivo desses métodos é produzir medidas de avaliação da qualidade de ajuste dos parâmetros de um modelo.

4.3 Estimação dos parâmetros via *Kernel*

Nesta seção é apresentada a estimação dos coeficientes dependentes do tempo do modelo de Cox, através da utilização da função de verossimilhança parcial ponderada.

Considere T^0 o tempo de falha, C a correspondente variável de censura e $\boldsymbol{X}(t)$ um vetor de p covariáveis dependentes ou não do tempo. Condicionalmente a $\boldsymbol{X}(\cdot)$, assumimos que T^0 e C são independentes. Sejam $\{(T_i^0, \boldsymbol{X}_i(\cdot), C_i), i = 1, ..., n\}$ n cópias independentes de $\{(T^0, \boldsymbol{X}(\cdot), C)\}$. Para o *i-ésimo* indivíduo, podemos observar $\{(T_i, \boldsymbol{X}_i, \delta_i)\}$, onde $T_i =$

$min(T_i^0, C_i)$ e $\delta_i = 1$ *se* $T_i = T_i^0$ e 0 caso contrário. A função de taxa de falha do modelo de Cox com coeficientes dependentes do tempo é definida como

$$\alpha_i(t) = \alpha_0(t)e^{\boldsymbol{X}_i'(t)\boldsymbol{\beta}(t)}, \ t \geq 0$$

Estamos interessados em estimar $\{\boldsymbol{\beta}(t), t > 0\}$ baseado em $\{(T_i, \boldsymbol{X}_i(\cdot), \delta_i), i = 1, ..., n\}$.

A idéia proposta por Cai e Sun (2003) para estimar os coeficientes dependentes do tempo do modelo de Cox é uma extensão da técnica de ajuste linear utilizada nos gráficos de dispersão combinada com a função de verossimilhança parcial.

Em uma janela em torno de cada instante de tempo t, cada um dos coeficientes de $\boldsymbol{\beta}(t)$ deve ser aproximado por uma função linear através da expansão da série de Taylor de primeira ordem. As estimativas para a função linear devem ser obtidas com a utilização da função de verossimilhança parcial e dos tempos de falha observados na janela definida.

Os valores estimados para a função linear no instante t são as estimativas dos coeficientes suavizados de $\boldsymbol{\beta}(t)$ para o instante t.

De acordo com os autores, este método também pode ser utilizado como uma ferramenta de diagnóstico para a identificação de covariáveis que não satisfazem à suposição de riscos proporcionais.

A seguir é apresentado o método de estimação proposto para os coeficientes dependentes do tempo do modelo de Cox.

Temos que $\boldsymbol{\beta}(t) = (\beta_1(t), \beta_2(t), \ldots, \beta_p(t))$.

Desta forma, para um ponto s na vizinhança de t, através da expansão de 1^a ordem em série de Taylor temos que

$$\beta_j(s) \approx \beta_j(t) + a_j(t)(s - t)$$

em que $a_j(t) = \partial\beta_j(t)/\partial t$.

Seja $\boldsymbol{\theta}$ definido como $\boldsymbol{\theta}(t) = (\beta_1(t), \ldots, \beta_p(t), a_1(t), \ldots, a_p(t))$. Além disso, seja $K(\cdot)$ a função *Kernel* e $h = h_n > 0$ o parâmetro de suavização. Portanto, a função de verossimilhança parcial é dada por

$$\boldsymbol{L}(\boldsymbol{\theta}, t) = \prod_{i=1}^{n} \left\{ \frac{e^{\boldsymbol{\theta}' \boldsymbol{X}_i(s)}}{\sum_{j=1}^{n} Y_j(s) e^{\boldsymbol{\theta}' \boldsymbol{X}_j(s)}} \right\}^{dNi(s)}.$$

Considerando o log da função de verossimilhança, temos

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}, t) = \sum_{i=1}^{n} \left\{ \boldsymbol{\theta}' \boldsymbol{X}_i(s) - \log \left(\sum_{j=1}^{n} Y_j(s) e^{\boldsymbol{\theta}' \boldsymbol{X}_j(s)} \right) \right\} dN_i(s).$$

A função de verossimilhança parcial ponderada pela *Kernel* proposta por Cai e Sun (2003) é dada por

$$\boldsymbol{\mathcal{L}}(\boldsymbol{\theta}, t) = \frac{1}{nh_n} \sum_{i=1}^{n} \int_0^{\tau} K\left(\frac{s-t}{h_n} \right) \times \left\{ \boldsymbol{\theta}' \boldsymbol{X}_i(s) - \log \left(\sum_{j=1}^{n} Y_j(s) e^{\boldsymbol{\theta}' \boldsymbol{X}_j(s)} \right) \right\} dN_i(s).$$

onde a função *Kernel* $K(\cdot)$, como descrito anteriormente, é uma função densidade de probabilidade simétrica com suporte em [-1, 1] e média 0.

Para $t \in [h_n, \tau - h_n]$, seja $\hat{\boldsymbol{\theta}}$ o maximador da função acima com relação a $\boldsymbol{\theta}$. Portanto, o estimador de máxima verossimilhança local parcial de $\boldsymbol{\beta}(t)$, denotado por $\hat{\boldsymbol{\beta}}(t)$, é o vetor formado pelos primeiros p componentes de $\hat{\boldsymbol{\theta}}$. E os últimos p componentes de $\hat{\boldsymbol{\theta}}$ são as estimativas das derivadas de $\boldsymbol{\beta}(t)$.

Dado que $t < h_n$ e $t > \tau - h_n$, seja $\hat{\beta}(t) = \hat{\beta}(h_n)$ e $\hat{\beta}(\tau - h_n)$ respectivamente. Desta

forma, temos $\hat{\beta}(t)$ bem definido no intervalo $[0, \tau]$, o que possibilita o estudo das propriedades deste estimador.

Cai and Sun (2003) provaram a consistência de $\hat{\beta}(t)$ para cada t fixado. Em Tian et al (2005) e Gill (1984) também podemos encontrar uma demostração mais detalhada de que $\hat{\beta}(t)$ apresenta consistência uniforme.

Um dos problemas mais complexos na metodologia considerada é a escolha da janela h_n. Neste trabalho, é apresentado o método de *validação cruzada em k-partes* proposto por Efron e Tibshirani (1993). A seguir, o método é apresentado com maiores detalhes. Este método é uma ferramenta muito utilizada para estimar o erro da previsão em modelos de regressão. O erro é dado por um valor que mede a qualidade de ajuste da previsão da variável resposta de acordo com o modelo que foi ajustado.

Nos modelos de regressão, o erro da previsão refere-se à esperança da diferença ao quadrado da futura resposta e do valor obtido pelo modelo:

$$PE = E(y - \hat{y})^2$$

Usualmente, para o cálculo desses valores, utilizamos a mesma amostra de dados tanto para o ajuste do modelo como para sua avaliação, o que pode resultar em estimativas do erro predito viciadas. Com o objetivo de se obter estimativas mais realistas para o erro, o ideal seria que os dados fossem divididos em dois grupos: um grupo para teste e um para desenvolvimento (treinamento).

Entretanto, na maioria das vezes, temos conjuntos de dados pequenos e dados adicionais dos experimentos estudados não são fáceis de serem disponibilizados por razões de custo e viabilidade, por exemplo. Desta forma, o método de validação cruzada tem como proposta usar parte da amostra para o ajuste do modelo e uma outra parte para sua avaliação e teste. Com grandes bases de dados, uma prática comum é a divisão dos dados em duas partes de mesmo tamanho.

Já com bases menores, o método de validação cruzada permite que as informações disponíveis sejam melhores utilizadas. A seguir é apresentada uma introdução deste procedimento. Outros detalhes do método podem ser encontrados em Cai et al (2000).

Tabela 4.1: Etapas do Método de Validação Cruzada

Etapa 1: dividir a base de dados em k partes iguais
Etapa 2: fixar a parte k e ajustar o modelo nas outras $(k-1)$ partes
Etapa 3: calcular o erro do modelo obtido na k-ésima parte
Etapa 4: repetir o processo acima para todas as outras $(k-1)$ partes
Etapa 5: combinar as k estimativas do erro predito

Suponha que a base de dados seja dividida em k partes. Seja $k(i)$ a parte da amostra contendo a informação i. Denote por $\hat{y}_i^{-k(i)}$ o valor ajustado para a observação i a partir da $k(i)$-*ésima* parte excluída dos dados. Então, o erro predito é dado por

$$CV = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i^{-k(i)}\right)^2$$

Freqüentemente, k é escolhido como $k = n$, resultando na *validação cruzada para cada elemento*. Para cada observação i, o modelo é reajustado desconsiderando a observação e o valor predito para a i-*ésima* observação é calculado e denotado por $\hat{y}_i^{-i}$.

Esse ajuste é repetido para cada observação e então é calculada a média da soma de quadrados da validação cruzada:

$$CV = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i^{-i}\right)^2$$

A partir da equação acima, os parâmetros serão definidos de acordo com a minização

de CV.

No artigo de Tian et al (2005) são propostos dois critérios diferentes para avaliação do erro da previsão. No primeiro, o erro é obtido através do logaritmo da função de verossimilhança parcial e para o segundo critério, o erro é baseado no resíduo martingal. Em seguida, a janela ótima h_n é escolhida de acordo com minimização dos erros.

4.4 Estimação dos intervalos

Após a estimação dos parâmetros apresentada na seção anterior, podemos determinar intervalos e bandas de confiança para o parâmetro $\boldsymbol{\beta}(t)$. A abordagem detalhada do desenvolvimento e obtenção das estimativas pode ser encontrada em Tian et al (2005).

Estamos interessados em estimar intervalos de confiança para o contraste $\boldsymbol{a}'\boldsymbol{\beta}(t)$ em um determinado tempo t fixado, onde $\boldsymbol{a}$ é um vetor de constantes conhecidas de dimensão p.

Através da utilização da expansão em série de Taylor da função score $U\{\hat{\boldsymbol{\beta}}(t), t\}$ em torno de $\boldsymbol{\beta}(t)$, temos

$$(nh_n)^{1/2}\{\hat{\boldsymbol{\beta}}(t) - \boldsymbol{\beta}(t)\} \approx I^{-1}\{\hat{\boldsymbol{\beta}}(t), t\}U\{\boldsymbol{\beta}(t), t\}$$

Utilizando a proposta da função score apresentada em Cai e Sun (2003), temos que

$$U\{\boldsymbol{\beta}(t), t\} \approx (nh_n)^{-1/2}\sum_{i=1}^{n}\int_0^{\tau}[\boldsymbol{X}_i(s) - E\{\boldsymbol{\beta}(t), s\}] \times K\left(\frac{s-t}{h_n}\right)dM_i(s)$$

em que

$$M_i(s) = N_i(s) - \int_0^s Y_i(t)\alpha_i(t)dt.$$

Para qualquer $t \in [h_n, \tau - h_n]$, a distribuição de $(nh_n)^{1/2}\{\hat{\boldsymbol{\beta}}(t) - \boldsymbol{\beta}(t)\}$ é aproximadamente normal com média 0 e matriz de covariância $\boldsymbol{I}^{-1}\{\hat{\boldsymbol{\beta}}(t), t\}\int_{-1}^{1}K^2(s)ds$.

Desta forma, para qualquer vetor $\boldsymbol{a}$ dado, de dimensão p, o intervalo de confiança para o contraste $\boldsymbol{a}'\boldsymbol{\beta}(t)$ pode ser obtido através da aproximação em grandes amostras da distribuição de $\hat{\boldsymbol{\beta}}(t)$.

Devido ao fato de que estamos interessados no comportameno temporal de $\{\boldsymbol{a}'\boldsymbol{\beta}(t)\}$, a construção e análise de bandas de confiança em determinado intervalo fixado, como por exemplo, $[b_1, b_2] \subseteq [h_n, \tau - h_n]$ são mais informativas que os intervalos.

Uma abordagem muito utilizada para a obtenção das bandas de confiança é através da aproximação em grandes amostras da distribuição da função de S dada por

$$S = \sup_{t \in [b_1, b_2]} \hat{w}(t) |\boldsymbol{a}' \{\hat{\boldsymbol{\beta}}(t) - \boldsymbol{\beta}(t)\}|$$

em que $\hat{w}(t)$ é uma função peso.

Seja c_α o percentil $100(1-\alpha)$ para a distribuição acima, sendo $0 < \alpha < 1$. Portanto, a banda de confiança para $\{\boldsymbol{a}'\boldsymbol{\beta}(t), t \in [b_1, b_2]\}$ é dada por

$$\{\boldsymbol{a}'\hat{\boldsymbol{\beta}}(t) \pm c_\alpha \hat{w}(t)^{-1}, b_1 \leq t \leq b_2\}.$$

Para se obter uma aproximação da distribuição de S deve ser utilizada a técnica de *strong approximation* apresentada em detalhes por Bickel e Rosenblatt (1973) e Yandell (1983).

Além da estimação dos parâmetros do modelo e das respectivas bandas de confiança, podemos também estudar a função de sobrevivência de um determinado indivíduo com um conjunto de covariáveis.

A função de sobrevivência é uma das principais funções probabilísticas utilizadas para

descrever estudos de sobrevivência. A função é definida como a probabilidade de uma observação não falhar até um certo tempo t, isto é, a probabilidade de uma observação sobreviver ao tempo t, que pode ser escrita como $S(t) = P(T \geq t)$.

A estimativa da função de sobrevivência pode ser obtida a partir do estimador de Breslow, considerando a função de taxa de falha acumulada, dada por:

$$\hat{S}(t) = \exp[-\hat{\Lambda}\{\hat{\boldsymbol{\beta}}(\cdot), t\}]$$

em que

$$\hat{\Lambda}\{\boldsymbol{\beta}(\cdot), t\} = \sum_{i=1}^{n} \int_0^t \frac{dN_i(s)}{\sum_{j=1}^{n} Y_j(s) e^{\boldsymbol{\beta}(s)' X_j(s)}}$$

Tian et al (2005) também abordam a construção de intervalos e bandas de confiança. No Capítulo 5 serão apresentados os resultados do ajuste da função de sobrevivência do modelo de Cox com coeficientes dependentes do tempo.

Capítulo 5

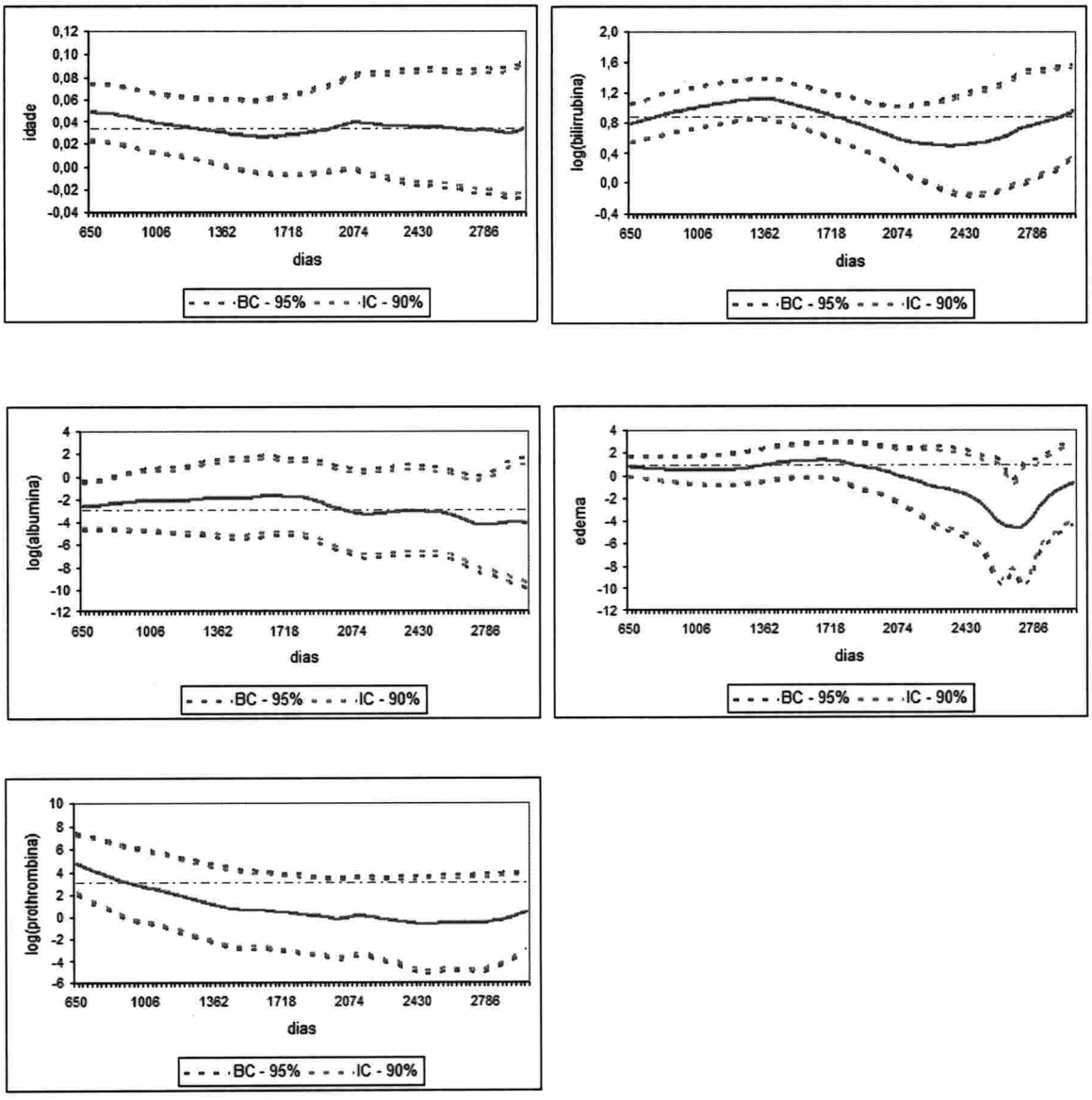
Resultados e Simulação

5.1 Modelo de Cox com coeficientes dependentes do tempo

A seguir serão apresentados os gráficos com as estimativas no tempo dos coeficientes do modelo de Cox apresentado no Capítulo 4 e também as estimativas dos intervalos e bandas de confiança utilizando a função *Kernel* Epanechnikov. Para tal análise, foram utilizados os dados de cirrose biliar primária descritos no Capítulo 2.

Os ajustes dos coeficientes foram obtidos com o auxílio do *software R*. A programação encontra-se no Apêndice A - *Programação em R*.

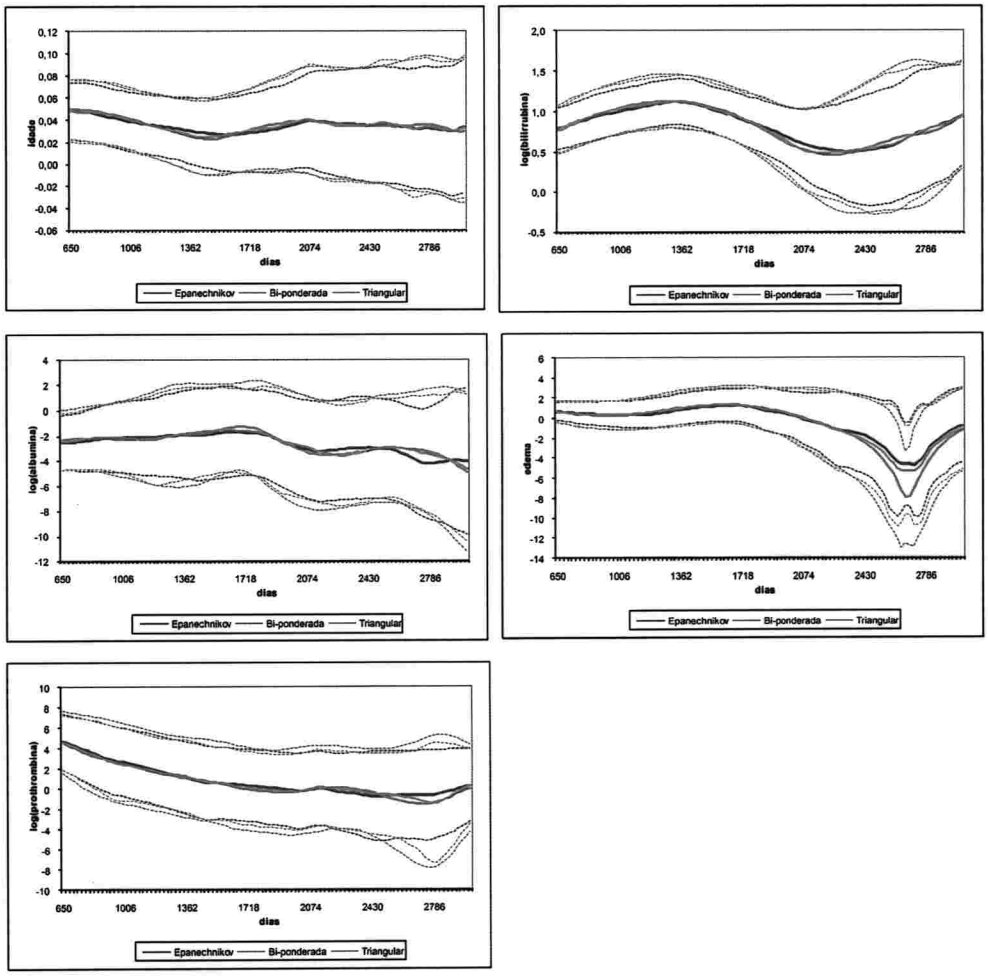
Gráfico 5.1: Estimativas de $\beta(t)$



Como visto anteriormente, os resultados dos gráficos sugerem forte efeito da covariável log(protrombina) na função de taxa de falha dos pacientes, mas que diminui gradualmente com o tempo.

Com o objetivo de comparar as estimativas resultantes de diferentes funções *Kernel*, serão apresentados a seguir os gráficos obtidos a partir do ajuste do modelo de Cox com coeficientes dependentes do tempo para as funções *Kernel* Bi-ponderada, Epanechnikov e Triangular e as respectivas bandas de confiança.

Gráfico 5.2: Estimativas de $\beta(t)$ para diferentes funções *Kernel*



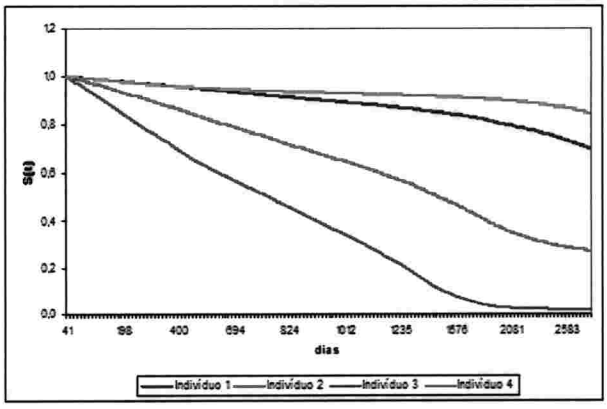
É possível verificar que as estimativas obtidas pelas diferentes funções Kernel apresentam valores muito semelhantes.

A seguir serão apresentados exemplos da estimativa da função de sobrevivência para diferentes indivíduos, considerando o modelo de Cox com coeficientes dependentes do tempo.

Tabela 5.1: Exemplos de valores das covariáveis presentes no modelo

Indivíduo	idade	edema	log(albumina)	log(bilirrubina)	log(prothrombina)
1	55	0.0	1.4085	0.000	2.2721
2	67	0.5	1.1817	0.7884	2.4069
3	57	1.0	1.0919	2.1400	2.5095
4	32	0.0	1.2641	-0.6931	2.3025

Gráfico 5.3: Funções de sobrevivência estimadas



5.2 Descrição da Simulação

O objetivo da simulação é avaliar a função *Kernel* que apresenta as melhores estimativas para o modelo de Cox com coeficientes dependentes do tempo, considerando razões de taxas de falha crescentes e decrescentes.

Na simulação também serão avaliados diferentes tamanhos de amostra e percentuais de censura.

Para obtermos razões de taxas de falha crescentes e decrescentes, consideramos inicialmente a variável aleatória T com distribuição de $Weibull(\rho, \lambda)$. As funções densidade de probabilidade e de sobrevivência são dadas por:

$$f(t) = \rho\lambda^{\rho}t^{\rho-1}e^{-(\lambda t)^{\rho}}, \quad t \geq 0, \rho \ e \ \lambda > 0$$

$$S(t) = e^{-(t\lambda)^{\rho}}$$

Seja X a covariável assumindo os valores 0 ou 1. Podemos definir a função de taxa de falha como $\alpha(t) = \lambda(X)^{\rho(X)}\rho(X)t^{\rho(X)-1}$, com os parâmetros de escala e forma definidos por $\lambda(X) = \zeta e^{\beta X}$ e $\rho(X) = e^{\gamma X}$.

Desta forma, para $X = 0$ temos que $\lambda(0) = \zeta$ e $\rho(0) = 1$ e para $X = 1$, $\lambda(1) = \zeta e^{\beta}$ e $\rho(1) = e^{\gamma}$.

Logo, as funções de taxa de falha para os valores de X são dadas por $\alpha(t/X = 0) = \zeta$ e $\alpha(t/X = 1) = e^{\gamma}(\zeta e^{\beta})^{e^{\gamma}}t^{e^{\gamma}-1}$ e, portanto, a razão das taxas de falha (RR) é definida por:

$$RR = \frac{\alpha(t/X = 1)}{\alpha(t/X = 0)} = \frac{e^{\gamma}(\zeta e^{\beta})^{e^{\gamma}}t^{e^{\gamma}-1}}{\zeta} = e^{\gamma}\zeta^{e^{\gamma}-1}e^{\beta e^{\gamma}}t^{e^{\gamma}-1}$$

$$RR = \kappa(\zeta, \gamma, \beta)t^{e^{\gamma}-1}$$

Desta forma, temos que:

- Para $\gamma > 0 \quad \rightarrow \quad e^{\gamma} - 1 > 0 \quad \rightarrow$ Razão crescente e
- Para $\gamma < 0 \quad \rightarrow \quad e^{\gamma} - 1 < 0 \quad \rightarrow$ Razão decrescente

Para gerar conjuntos de dados com razão de taxas de falha crescente ou decrescente, de acordo com definições acima, devemos seguir as etapas abaixo:

1. Fixar os parâmetros ζ, γ e β, de acordo com o comportamento que se deseja obter (RR crescente ou decrescente).

2. Gerar n (tamanho da amostra) valores para a covariável X, aleatoriamente escolhidos dentre $\{0, 1\}$.

3. Para valores da covariável $X = 0$, os tempos de falha serão gerados com distribuição $Exp(\zeta)$, pois sabemos que $\alpha(t) = \zeta$ e $S(t) = e^{-(\zeta t)}$. Em seguida, geramos n valores de uma $Uniforme[0, 1]$: $u_1, u_2, ..., u_n$ e utilizando o Método da Distribuição Inversa, temos que $u_j = S(t_j) = e^{-(\zeta t_j)}$ e $t_j = S^{-1}(u_j)$.

Portanto, $ln(u_j) = -\zeta t_j \quad \rightarrow \quad t_j = -\frac{1}{\zeta} ln(u_j)$.

4. Para valores da covariável $X = 1$, os tempos de falha serão gerados com distribuição $Weibull(\rho, \lambda)$, pois $S(t) = e^{-(\lambda t)^{\rho}}$. Com o auxílio da distribuição $U[0, 1]$ e do Método da Distribuição Inversa, obtemos:

$u_j = S(t_j) = e^{-(\lambda t)^{\rho}}$

$ln(u_j) = -(\lambda t_j)^{\rho}$

$ln(-ln(u_j)) = \rho ln(\lambda) + \rho ln(t_j)$

Portanto, $\rightarrow t_j = e^{\frac{1}{\rho}ln(-ln(u_j))-ln(\lambda)} = e^{\frac{1}{e^{\gamma}}ln(-ln(u_j))-ln(\zeta e^{\beta})}$

Observações:

- valor apropriado para ρ : entre 0,5 e 3
- $E(T) = \frac{1}{\lambda}\Gamma[1 + (1/\rho)]$
- $Var(T) = \frac{1}{\lambda^2}[\Gamma[1 + (2/\rho)] - \Gamma[1 + (1/\rho)]^2]$

Para gerar valores da variável de censura, de acordo com definições anteriores, devemos desenvolver as seguintes etapas:

1. Para a covariável $X = 0$, definimos:

$T \sim Exp(\zeta)$

$C \sim U[0, \tau]$

Seja p_c o percentual de censura dos dados que será utilizado na simulação (10%, 20%, ...).

$$p_c = P(T > C) = EP(T > C/T) = \int_0^{+\infty} P(C < t/T = t).\zeta e^{-\zeta t} dt$$

$$= \int_0^{+\infty} \frac{t}{\tau}\zeta e^{-\zeta t} dt = \frac{1}{\tau}E(T) = \frac{1}{\tau\zeta}$$

Portanto,

$$\tau = \frac{1}{\zeta p_c}$$

2. Para a covariável $X = 1$ temos:

$T \sim Weibull(\rho, \lambda)$

$C \sim U[0, \tau]$

$$p_c = P(T > C) = EP(T > C/T) = \int_0^{+\infty} P(C < t/T = t).f_T dt$$

$$= \int_0^{+\infty} \frac{t}{\tau} f_T dt = \frac{1}{\tau} E(T) = \frac{1}{\tau\lambda} \Gamma(1 + (1/\rho))$$

Portanto,

$$\tau = \frac{\Gamma(1 + (1/\rho))}{\lambda p_c}$$

Para avaliar a qualidade de ajuste das diferentes funções *Kernel*, foram definidas as medidas de erro de EQM e Disc (medida de discrepância com o objetivo de apresentar o vício).

n = tamanho da amostra, $n = 1, ..., n_1$, $\quad n_1 = 100, 200, 500$

k = número de repetições, $k = 100$

j = instantes de tempo para estimação dos coeficientes, $j = 1, ..., j_1$

$$E\hat{Q}M = \frac{1}{100} \sum_{k=1}^{100} \sum_{j=1}^{j_1} (\hat{\beta}_k(T_j) - \bar{\beta}(T_j))^2$$

$$\hat{Disc} = \frac{1}{100} \sum_{k=1}^{100} sup_{1 \leq j \leq 50} \left| \hat{\beta}_k(T_j) - \bar{\beta}(T_j) \right|$$

onde,

$$\bar{\beta}(T_j) = \frac{1}{100}\sum_{k=1}^{100}\hat{\beta}_k(T_j)$$

A seguir serão apresentados os resultados das simulações, comparando os ajustes dos modelos com diferentes funções *Kernel*, tamanhos de amostra e percentuais de censura.

5.3 Resultados

Objetivo: Avaliar a função *Kernel* que apresenta o melhor ajuste na estimação dos parâmetros do modelo de Cox com coeficientes dependentes do tempo. Para a simulação foram consideradas funções de razão de taxa de falha com comportamento crescente e decrescente.

Os seguintes parâmetros foram considerados na simulação: tamanho da amostra ($n = 100, 200, 500$), percentual de censura ($p_c = 0\%, 10\%, 30\%, 60\%$) e funções *Kernel* (Epanechnikov, Bi-ponderada e Triangular).

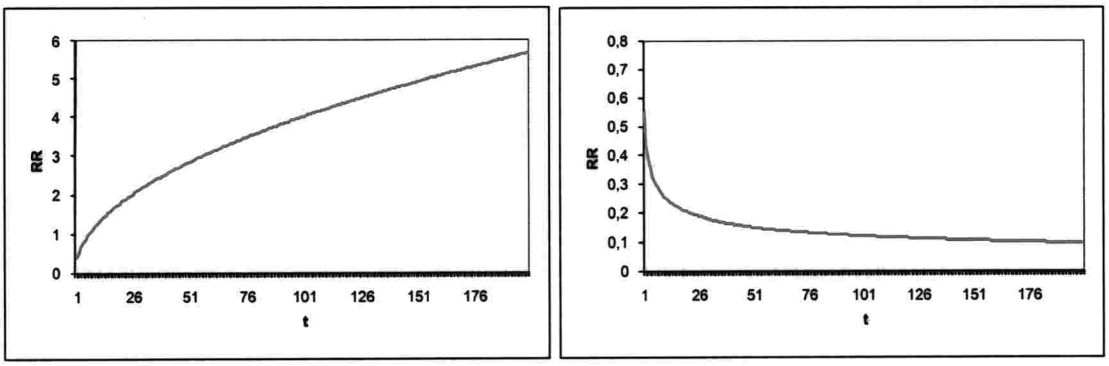
Para a geração dos dados foram fixados os seguintes valores, de acordo com o comportamento das razões das taxas de falha:

Tabela 5.2: Valores fixados para os parâmetros

Parâmetro	RR Crescente	RR Decrescente
ζ	0.05	0.05
β	0.20	0.80
γ	0.41	-0.40

Abaixo são apresentados os gráficos das razões de taxa de falha obtidos a partir da simulação.

Gráfico 5.4: Razões de taxa de falha obtidas



A seguir são apresentadas as medidas de EQM e Disc para as diferentes funções Kernel.

Tabela 5.3: Resultados para razão de taxa de falha crescente

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
0%	Epanechnikov	6.77	0.53	2.96	0.34	1.17	0.21
	Bi-ponderada	15.86	0.83	6.29	0.46	1.78	0.23
	Triangular	8.68	0.60	3.90	0.37	1.41	0.22

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
10%	Epanechnikov	19.20	0.90	5.63	0.45	1.61	0.24
	Bi-ponderada	19.38	0.89	6.86	0.51	2.50	0.31
	Triangular	10.95	0.70	5.44	0.44	1.61	0.26

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
30%	Epanechnikov	71.55	1.45	8.59	0.60	2.65	0.32
	Bi-ponderada	34.04	1.41	19.77	0.71	4.83	0.44
	Triangular	78.90	2.03	8.38	0.64	2.88	0.34

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
60%	Epanechnikov	-	-	442.29	8.05	51.02	1.92
	Bi-ponderada	-	-	421.76	7.82	39.19	1.63
	Triangular	-	-	364.41	6.86	21.52	1.10

Tabela 5.4: Resultados para razão de taxa de falha decrescente

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
0%	Epanechnikov	17.96	0.75	5.53	0.40	2.04	0.23
	Bi-ponderada	29.66	0.95	10.01	0.54	4.21	0.32
	Triangular	24.69	0.82	6.86	0.45	2.51	0.25

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
10%	Epanechnikov	33.96	1.10	8.30	0.50	2.75	0.28
	Bi-ponderada	33.18	1.01	12.61	0.59	4.52	0.34
	Triangular	26.72	1.01	9.32	0.47	2.99	0.29

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
30%	Epanechnikov	81.58	1.27	9.33	0.62	3.27	0.43
	Bi-ponderada	44.20	1.32	59.20	1.03	6.43	0.44
	Triangular	75.51	1.45	10.45	0.58	3.84	0.36

		$n = 100$		$n = 200$		$n = 500$	
p_c	Kernel	EQM	Disc	EQM	Disc	EQM	Disc
60%	Epanechnikov	-	-	389.83	6.13	23.55	1.42
	Bi-ponderada	-	-	352.78	6.90	40.76	1.99
	Triangular	-	-	384.47	6.42	35.11	1.30

Para a razão de taxas de falha crescente é possível verificar que as funções *Kernel* Epanechnikov e Triangular apresentaram bons resultados para dados com até 30% de censura. As simulações considerando percentual de censura de 60% apresentaram resultados pouco satisfatórios. Para $n = 100$ e $p_c = 60\%$ não foi possível a obtenção das estimativas devido ao tamanho do conjunto de dados.

Conforme esperado, quanto maior a amostra, observa-se a diminuição dos valores de EQM obtidos. Para $n = 500$ as estimativas obtidas a partir da Epanechnikov e Triangular também apresentaram as menores medidas de erro.

Considerando a razão de taxas de falha decrescente, podemos observar que a *Kernel* Epanechnikov apresentou os melhores ajustes para a maioria das simulações. Assim como para o comportamento crescente, os resultados obtidos com dados com percentual de censura de 60% não apresentaram fortes conclusões.

A literatura faz referências de que a *Kernel* Epanechnikov é a mais utilizada, mas devido à semelhança das estimativas obtidas, a escolha da função deve ser feita considerando também outros aspectos como, por exemplo, esforço computacional.

As simulações apresentadas sugerem o uso da *Kernel* Epanechnikov para dados com razão de taxas de falha decrescente e *Kernel* Epanechnikov ou Triangular para dados com razão crescente.

Capítulo 6

Considerações Finais

Neste trabalho foi apresentada uma técnica para a obtenção das estimativas do modelo de Cox com coeficientes dependentes do tempo a partir da função de verossimilhança parcial ponderada por uma função *Kernel*.

O objetivo principal foi apresentar uma nova abordagem para estudos que tenham interesse em avaliar o efeito temporal das covariáveis no tempo de sobrevivência analisado. Além disso, foi proposta uma comparação das estimativas dos coeficientes obtidas com o ajuste de diferentes funções *Kernel*.

Na aplicação a dados reais, foi utilizado o conjunto de dados de cirrose biliar primária da Mayo Clinic. Os resultados sugerem que existe um forte efeito da covariável log(protrombina) para $t < 1200$ dias que diminui gradualmente com tempo.

O modelo de Cox com coeficientes dependentes do tempo tem auxiliado principalmente a área médica, na qual o principal objetivo é comparar o efeito de diferentes drogas e suas eficácias durante os tratamentos.

Para as estimativas dos coeficientes no tempo obtidas a partir de funções *Kernel* distintas, verificamos que os valores obtidos foram semelhantes para as funções Epanechnikov, Bi-ponderada e Triangular.

Corforme mencionado anteriormente é aconselhável a utilização de métodos que auxiliem na escolha do valor das janelas de tempo, para que os resultados não sejam viciados.

Na simulação foi possível comparar as estimativas resultantes da aplicação de diferentes funções *Kernel* considerando razões de taxa de falha crescentes e decrescentes.

De acordo com os resultados, a função *Kernel* Epanechnikov apresentou as menores medidas de erro, tanto para a razão de taxa de falha crescente como decrescente. Como já era esperado, quanto maior o tamanho da amostra, menor o erro. Já para o percentual de censura, quanto menor o seu valor, melhores os ajustes. De forma geral, foi verificado que as funções *Kernel* apresentaram resultados muito próximos para as funções de taxa de falha crescente e decrescente.

Como sugestão para trabalhos futuros, o estudo de taxas de convergência, métodos de diagnóstico e também a comparação do modelo proposto com outros métodos para estimação dos coeficientes dependentes do tempo do modelo de Cox.

Apêndice A

Programação em R

Neste apêndice serão apresentados os códigos dos programas em R que foram utilizados para as análises descritas no trabalho, dentre eles: estimação dos parâmetros, estimação da função de sobrevivência, gráficos das estimativas dos coeficientes, gráficos de resíduos e resultados das simulações.

A programação envolvendo as estimativas para o modelo de Cox com coeficientes dependentes do tempo foi cedida por Lu Tian, autor do artigo *On the Cox Model with Time-Varying Regression Coeficients*, publicado em Março de 2005 no *Journal of the American Statistical Association.*

A.1 Modelo de Cox

- Coeficientes do modelo de Cox de riscos proporcionais

```
cirrose <- read.table("C:\ Documents and Settings \ Windows \ Meus Documentos \ cir-
rose.txt", h = T)
attach(cirrose)
require(survival)
summary(cirrose)
ajuste <- coxph(Surv(Number Days, Status) ~ Age + log(Bilirubin) + log(Albumin) +
log(Prothrombin) + Edema, data = cirrose))
summary(ajuste)
residuos <- resid(ajuste, type = "scaledsch")
cox.zph(ajuste, transform = "identity")
testerp <- cox.zph(ajuste, transform = "identity")
```

- Coeficientes dependentes do tempo

```
# Input
# (y, delta): tempo de sobrevivência observado e variável de censura
# x: matriz de covariáveis n x p
# d: janela
# grid: pontos onde β(t) será estimado
# B: número de reamostras para obtenção das bandas e intervalos de confiança
# control: critério de convergência


# Output
# vb: estimativas de β(t)
# vbv: matriz de covariância estimada para cada β(t)
```

```
# cut.point1: ponto de corte para construção das bandas de confiança (95%)
# cut.point2: ponto de corte para construção dos intervalos de confiança (90%)

# Função para obter as estimativas de β(t)
coxph.time <- function(y, delta, x, d, grid, B = 500, control = 1e-9)
{
x <- as.matrix(x)
n <- length(y)
p <- length(x[1, ])
grid.lenght <- length(grid)

index <- order(y)
y <- y[index]
delta <- delta[index]
x <- x[index, ]

vb <- vx <- matrix(0, ncol = grid.length, nrow = p)
ptb <- array(0, c(n, p, grid.length))
varm <- array(0, c(p, p, grid.length))

beta.est0 <- coxph(Surv(y, delta) ~ x)$coef

for(j in 1 : grid.length)
{
grid.point <- grid[j]
weight <- 0.75 * (1 - ((y - grid.point)/d)^2) * (abs(y - grid.point) < d)

index1 <- min((1 : n)[weight > 0])
index 2 <- max((1 : n)[weight > 0])
xt <- x[n : index1, ]
```

```
xtf <- x[index2 : index1, ]
nt1 <- n - index1 + 1
nt2 <- index2 - index1 + 1

deltanew <- (delta * weight)[index2 : index1]
beta.est <- beta.est0

error <- 1
while (error > control)
   { effect <- as.vector(xt % * % beta.est)
   effect <- effect - mean (effect)
   s0 <- exp(effect)
   s1 <- t(xt * s0)

   cums0 <- cumsum(s0)[(nt1 - nt2 + 1) : nt1]
   cums1 <- apply(s1, 1, cumsum)[(nt1 - nt2 + 1) : nt1, ]
   expectation <- cums1 / cums0
  difference <- (xtf - expectation) * deltanew
   score <- apply(difference, 2, sum)

   temp <- (cumsum(deltanew[nt2 : 1] / cums0[nt2 : 1]))[nt2 : 1]
   temp <- c(rep(temp[1], nt1 - nt2), temp)

   I1 <- t(xt * (temp * s0)) %*% xt
   I2 <- t(expectation * deltanew) %*% expectation
   v <- solve(I2 - I1)
   beta.est <- as.vector(beta.est - v %*% score)
   error <- max(abs(score))
   }
ptb[(index2 : index1), , j] <- difference
```

```
varm [ , , j] <- v
vb[ , j] <- beta.est
}

rdiff <- matrix(0, nrow = p, ncol = grid.length)
mdata <- array(0, c(p, grid.length, B))

for (s in 1 : B)
{ g <- rnorm(n)
for (i in 1 : grid.length)
   { rdiff[ , i] <- varm[ , , i] %*% apply(ptb[ , , i] * g, 2, sum) }
mdata[ , , s] <- rdiff
}

for(j in 1 : grid.length)
   { vx[ , j] <- sqrt(diag((n - 1) * var(t(varm[ , , j] %*% t (ptb[ , , j])))))) }

m <- matrix(0, ncol = B, nrow = p)
for (s in 1 : B)
{ m[ , s] <- apply(abs(mdata[ , , s] / vx), 1, max) }

cut.point1 <- cut.point2 <- rep(0, p)
for(i in 1 : p)
   { cut.point1[i] <- quantile(m[i, ], 0.95)
   cut.point2 <- quantile(m[i, ], 0.90) }

return(list(vb = vb, vbv = vx, cut.point1 = cut.point1, cut.point2 = cut.point2))
}
```

- Função de Sobrevivência

```
# Input
# (y, delta): tempo de sobrevivência observado e variável de censura
# x: matriz de covariáveis n x p
# d: janela
# x0: vetor de covariáveis, condicional a função de sobrevivência que será estimada
# maxtime: valor máximo do intervalo de tempo onde a função de sobrevivência será
estimada
# control: critério de convergência

# Output
# time: intervalo de tempo onde a função foi estimada
# surv: estimativa da função de sobrevivência

# Função para obter a estimativa de S(t)
coxph.prd <- function(y, delta, x, x0, d, maxtime, control = 1e-8)
{
x <- as.matrix(x)
x <- t(t(x) - x0)
n <- length(y)
p <- length(x[1, ])

index <- order(y)
y <- y[index]
delta <- delta[index]
x <- x[index, ]

grid <- y[delta * ( y <= maxtime) == 1]
grid.length <- length(grid)
```

```
beta.est0 <- coxph(Surv(y, delta) ~ x)$coef
prd <- rep(0, grid.length)

for (j in 1 : grid.length)
{
grid.point <- grid[j]
weight <- 0.75 * (1 - ((y - grid.point) / d) ^2) * (abs(y - grid.point) < d)

index1 <- min((1 : n)[weight > 0])
index2 <- max((1 : n)[weight > 0])

xt <- x[n : index1, ]
xtf <- x[index2 : index1, ]
xt <- t(t(xt) - apply(xt, 2, mean))
xtf <- t(t(xtf) - apply(xtf, 2, mean))

nt1 <- n - index1 + 1
nt2 <- index2 - index1 + 1

deltanew <- (delta * weight)[index2 : index1]

beta.est <- beta.est0
error <- 1
while(error > control)
   { s0 <- as.vector(exp(xt %*% beta.est))
   s1 <- t(xt * s0)

   cums0 <- cumsum(s0)[(nt1 - nt2 + 1) : nt1]
   cums1 <- apply(s1, 1, cumsum)[(nt1 - nt2 + 1) : nt1, ]
   expectation <- cums1 / cums0
```

```
    difference <- (xtf - expectation) * deltanew
    score <- apply(difference, 2, sum)

    temp <- (cumsum(deltanew[nt2 : 1] / cums0[nt2 : 1]))[nt2 : 1]
    temp <- c(rep(temp[1], nt1 - nt2), temp)

    I1 <- t(xt * (temp * s0)) %*% xt
    I2 <- t(expectation * deltanew) %*% expectation
    v <- solve(I2 - I1)
    beta.est <- as.vector(beta.est - v %*% score)
    error <- max(abs(score))
    print(c(grid.point, error))
    }
    prd[j] <- 1 / sum(exp(x[y >= grid.point, ] %*% beta.est))
}
prd <- cumsum(prd)
list(time = grid, surv = exp(-prd))
}
```

- Exemplo

```
library( survival)
data (pbc)
y = pbc$time
delta = pbc$status
x = cbind(pbc$age, pbc$albumin)
d = 650
grid = seq(d, 3000, length = 100)
fit = coxph.time(y, delta, x, d = 650, grid = seq(d, 3000, length = 100), B = 250)
```

```
plot(grid, fit$vb[1, ], type = "l")


x0 = apply(x, 2, mean)
fit = coxph.prd(y, delta, x, x0, d, maxtime = 3000)
plot(fit$time, fit$surv, type = "l")
```

A.2 Simulação

- Razão das taxas de falha

```
lin <- 200
zeta <- 0.05
gama <- 0.41
beta <- 0.2
del <- 1
var <- as.integer(runif(lin, 0, 2))
u <- runif(lin, 0, 1)
t <- ifelse(var < 0.5, -(1/zeta)*logb(u,exp(1)), exp(((1/exp(gama))*(logb(-logb(u,exp(1)),
exp(1))))-logb(zeta*exp(beta),exp(1))))
dados <- cbind(var, t, del)
fit <- coxph(Surv(t, del) ~ var, data = dados)
summary(fit)
residuals(fit, type = "scaledsch")
cox.zph(fit, transform = "identity")
plot(cox.zph(fit))
lines(c(0,1), c(0,0), type = "l")
```

- Dados para simulação

```
require(survival)
lin <- 500
rep <- 100
k <- 1
final <- array(0, c(lin, 3, rep))
result <- array(0, c(50, 1, rep))
```

```
for (k in 1:rep)
zeta <- 0.05
gama <- 0.41
beta <- 0.2
pc <- 0.0
lambda <- zeta*exp(beta)
var <- as.integer(runif(lin, 0, 2))
u <- runif(lin, 0, 1)
tempo <- ifelse(var < 0.5, -(1/zeta)*logb(u,exp(1)),
exp(((1/exp(gama))*(logb(-logb(u,exp(1)),
exp(1))))-logb(zeta*exp(beta),exp(1))))
dados <- cbind(var, u, tempo)
tal.exp <- 1/(zeta*pc)
u.exp <- runif(lin, 0, tal.exp)
tal.wei <- (factorial(1 + 1/exp(gama)))/(lambda*pc)
u.wei <- runif(lin, 0, tal.wei)
dados <- cbind(var, u, tempo, u.exp, u.wei)
u.exp.c <- ifelse(tempo > u.exp, 0, 1)
u.wei.c <- ifelse(tempo > u.wei, 0, 1)
del <- ifelse(var < 0.5, u.exp.c, u.wei.c)
dados <- cbind(var, u, tempo, u.exp, u.wei, u.exp.c, u.wei.c, del)
dados <- cbind(var, tempo, del)
final[,,k] <- dados
y <- final[, 2, k]
delta <- final[, 3, k]
x <- cbind(final[, 1,k])
coxph.time <- function(y, delta, x, d, grid, B, control = 1e-9)
fit = coxph.time(y, delta, x, d, grid, B = 20)
vb <- as.matrix(vb)
result[, , k] <- vb
media <- result
```

```
beta.mean <- as.matrix(apply(media, 1, mean))
calc <- (media - beta.mean)^2
calc1 <- apply(calc, 2, sum)
calc2 <- sum(calc1)
EQM <- calc2/rep
EQM
aux <- (media - beta.mean)/beta.mean
aux1 <- apply(aux, 2, sum)
aux2 <- sum(aux1)
VR <- aux2/rep
VR
plot(beta.mean)
```

Referências Bibliográficas

Aalen, O. O. (1978), *Nonparametric Inference for a Family of Counting Processes*, The Annals of Statistics, 6, 701-726.

Andersen P., Borgan, O., Gill, R., and Keiding, N. (1993), *Statistical Models Based on Counting Processes*, New York: Springer-Verlag.

Bickel, P., and Rosenblatt, R. (1973), *On Some Global Measures of the Deviations of Density Function Estimates*, The Annals of Statistics, 1, 1071-1095.

Cai, Z., Fan, J., and Li, R. (2000), *Efficient Estimation and Inference for Varying-Coefficient Models*, Journal of the Statistical Association, 95, 888-902.

Cai, Z., and Sun, Y. (2003), *Local Linear Estimation for Time-Dependent Coefficients in Cox´s Regression Models*, Scandinavian Journal of Statistics, 30, 93-111.

Colosimo, E. A., e Giolo, S. R. (2006), *Análise de Sobrevivência Aplicada*, São Paulo: Edgard Blücher.

Cox, D. R. (1972), *Regression Models and Life Tables*, Journal of the Royal Statistical Society, Ser. B, 34, 187-220.

Efron, B., and Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, New York: Chapman & Hall.

Fleming, T. R., and Harrington, D. P. (1991), *Counting Processes and Survival Analysis*, New York: Wiley.

Gill, R. (1984), *Understanding Cox´s Regression Model: A Martingale Approach*, Journal of the American Statistical Association, 386, 441-447.

Grambsch, P. M., and Therneau, T. M. (2000), *Modeling Survival Data: Extending the Cox Model*, New York: Springer-Verlag.

Hastie, T., and Tibshirani, R. (1993), *Varying-Coefficient Models*, Journal of the Royal Statistical Society, Ser. B, 55, 757-796.

Martinussen, T., Scheike, T. H., and Skovgaard, I. M. (2002), *Efficient Estimation of Fixed and Time-Varying Covariate Effects in Multiplicative Intesity Models*, Scandinavian Journal of Statistics, 29, 57-74.

Murphy, S., and Sen, P. (1991), *Time-Dependent Coefficients in a Cox Type Regression Model*, Stochastic Process and their Applications, 39, 153-180.

Schoenfeld, D. (1982), *Partial Residuals for the Proportional Hazards Regression Model*, Biometrika, 69, 239-241.

Silverman, B. W. (1986), *Density Estimation*, New York: Chapman & Hall.

Staniswalis, J. (1989), *The Kernel Estimate of a Regression Function in Likelihood-Based Models*, Journal of the American Statistical Association, 84, 276-283.

Tian, L., Zucker, D., and Wei, L. J. (2005), *On the Cox Model with Time-Varying Regression Coefficients*, Journal of the American Statistical Association, 469, 172-183.

Valsecchi, M., Silvestri, D., and Sasieni, P. (1996), *Evaluation of Long-Term Survival: Use of Diagnostics and Robust Estimators with Cox´s Proportional Hazards Model*, Statistics in Medicine, 15, 2763-2780

Verweij, P., and van Houwelingen, H. (1995), *Time-Dependent Effects of Fixed Covariates in Cox Regression*, Biometrics, 51, 1550-1556.

Wand, M. P., and Jones, M. C. (1995), *Kernel Smoothing*, New York: Chapman & Hall.

Winnett, A., and Sasieni, P. (2003), *Iterated Residuals and Time-Varying Covariate Effect in Cox Regression*, Journal of the Royal Statistical Society, Ser. B, 65, 473-488.

Yandell, B. (1983), *Nonparametric Inference for Rates with Censored Survival Data*, The Annals of Statistics, 11, 1119-1135.

Zucker, D., and Karr, A. (1990), *Nonparametric Survival Analysis with Time-Dependent Covariate Effects: A Penalized Partial Likelihood Approach*, The Annals of Statistics, 18, 329-353.