ANÁLISE DE CURVAS DE
SOBREVIVÊNCIA — O MODELO DE REGRESSÃO
DE COX

MARIA CECILIA MENDES BARRETO

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

ORIENTADOR:

Prof. Dr. Adolpho Walter Pimazoni Canton

- SÃO PAULO, MARÇO DE 1982 -

A
meus país
e
Nicolau Casagrande

"Conto os que morrem de Boubá,
de Tifo e de Verminose,
Conto os que morrem de Crupe,
de Chagas e Xistosomose.
Mas todos estes defuntos
Morrem de fato e de fome,
quer a chamemos de febre
ou de qualquer outro nome."

Ferreira Gullar

A G R A D E C I M E N T O S

Todo e qualquer trabalho não é o resultado do esforço e interesse de apenas uma pessoa. Por isso quero lembrar algumas que contribuíram de forma e intensidade, talvez, diferentes.

- Professor Doutor Adolpho Walter Pimazoni Canton, pela orientação e sugestões no decorrer deste processo de aprendizagem;
- Professor Doutor David Martins Dorigo, pelas sugestões e intervenções nos momentos oportunos;
- Doutor Eudes de Freitas Aquino, por ter me familiarizado com alguns dos problemas de alotransplante renal;
- Edina, Lúcia, Mônica, Mara, Márcia, Beatriz e Flávio que em épocas diferentes constituíram e constituem minha família em São Carlos;
- Antonio Aparecido Rosalem, pelo empenho sincero mas infrutífero na elaboração de programas de computador;
- Os colegas do Departamento de Computação e Estatística da Universidade Federal de São Carlos, pelas sugestões, paciência, ânimo e bom humor;
- Maria da Graça do Prado Juliano e Maria José Gualtieri da Costa, pelo rápido, eficiente e preciso trabalho de datilografia.

A todos, meus agradecimentos.

MARIA CECILIA MENDES BARRETO

Í N D I C E

ANÁLISE DE CURVAS DE SOBREVIVÊNCIA — O MODELO DE REGRESSÃO DE COX

PREFÁCIO	vii
 CAPÍTULO I — INTRODUÇÃO	 1
1.1. Considerações a Respeito da Análise de Curvas de Sobrevivência	1
1.2. Alguns Métodos de Análise de Curvas de Sobrevi vência	3
1.3. Formulação Geral dos Modelos de Regressão	5
1.4. Um Exemplo para Introduzir a Metodologia	9
1.5. Fazendo Inferências sobre o Parâmetro de Regressão ...	14
 CAPÍTULO II — O MODELO DE REGRESSÃO DE COX — ESTIMAÇÃO	 19
2.1. O Método do Produto-Limite	20
2.2. A Verossimilhança Parcial	22
2.3. Estimação Geral do Modelo Proposto por Cox	26
2.4. Algumas Propriedades dos Estimadores Baseados na Verossimilhança parcial	28
2.5. Estimação da Função de Sobrevivência	30
 CAPÍTULO III — EXTENSÃO DA ESTIMAÇÃO PARA O CASO DE EMPATES	 34
3.1. A Aproximação de Cox	35
3.2. A Abordagem de Kalbfleisch e Prentice	37
3.3. A Aproximação de Peto-Breslow	40
3.4. A Aproximação para Dados Agrupados	41
3.5. Desempate Aleatorizado	42

CAPÍTULO IV - ANÁLISE DO MODELO LINEAR GERAL DA RAZÃO DE RISCO PROPORCIONAL	44
4.1. Seleção de Covariáveis Através de Análise Gráfica	45
4.2. Estatísticas Empregadas na Seleção de Covariáveis	46
4.3. O Procedimento Stepwise	47
 CAPÍTULO V - APLICAÇÕES NUMÉRICAS	 51
5.1. Aspectos Gerais do Alotransplante Renal	52
5.2. Análise dos Dados de Doador Cadáver	54
5.3. Análise dos Dados de Doador Vivo	56
5.4. Alguns Aspectos Computacionais	62
 CAPÍTULO VI - CONCLUSÕES	 64
 APÊNDICE	 66
 REFERÊNCIAS BIBLIOGRÁFICAS	 81

P R E F Á C I O

O estudo do tempo de sobrevivência de indivíduos é de interesse em várias áreas de investigação científica. Por exemplo, no estudo de tratamentos em pacientes leucêmicos, o objetivo é comparar as curvas de sobrevivência entre diversos tratamentos.

Considerando um conjunto de indivíduos a partir de um instante bem definido, cada um deles é observado até que ocorra seu tempo de falha (morte) ou tempo de censura (perda). Para estes últimos sabemos apenas que seu tempo de falha é maior que o tempo de censura.

Muitas vezes o registro de informações adicionais na forma de covariáveis é de grande utilidade na análise de curvas

de sobrevivência. Em 1972, Cox introduziu uma metodologia específica que com a utilização de covariáveis na curva de sobrevivência nos permite estimar e testar hipóteses acerca dos subgrupos envolvidos. Essa metodologia é baseada nos estimadores de máxima verossimilhança da função de sobrevivência desenvolvidos por Kaplan e Meier (1958) que os denominaram estimadores do produto limite.

Neste trabalho nos propomos a estudar essa metodologia quando as variáveis não são dependentes do tempo de sobrevivência e aplicá-la em problemas experimentais reais.

No primeiro capítulo introduzimos alguns métodos de análise de curvas de sobrevivência e a formulação geral dos modelos de regressão. Usando parte dos resultados de uma pesquisa sobre alotransplante renal, introduzimos o modelo de regressão de razão de risco proporcional, ao final desse capítulo.

A teoria de estimação está formulada no segundo capítulo, e sua extensão para o caso de empates é apresentada no terceiro capítulo.

A análise geral do modelo linear geral da razão de risco proporcional é descrita no quarto capítulo.

No quinto capítulo temos a aplicação da metodologia de Cox em dois grupos de pacientes que sofreram alotransplante renal (Aquino, 1979). No primeiro grupo temos como doador cadáver, e no segundo, doador vivo. Usam-se como variáveis que possam influenciar a sobrevivência a idade ao fazer o transplante, o sexo e o tipo de esquema terapêutico imunossupressor.

Considerações finais sobre a metodologia de Cox são apresentadas no sexto capítulo.

C A P Í T U L O I

INTRODUÇÃO

1.1. CONSIDERAÇÕES A RESPEITO DA ANÁLISE DE CURVAS DE SOBREVIVÊNCIAS

O estudo em tábuas de sobrevivência é uma das técnicas estatísticas mais antigas e de grande utilidade na área de saúde e de seguros.

Existem duas formas de tábuas de sobrevivência que são largamente empregadas: as tábuas de vida de geração e as tábuas de vida corrente.

As tábuas de vida de geração, ou de coorte, ou ainda de seguimento, registram a experiência de mortalidade de um grupo de indivíduos durante um certo período de tempo, desde seu

nascimento até a morte do último indivíduo do grupo. Uma tábua de vida corrente ou convencional registra, por outro lado, a experiência de mortalidade de toda a população em um curto período de tempo, geralmente durante um ano.

Tanto em uma como em outra, a experiência de mortalidade pode ser decomposta em várias causas de morte. Assim podemos estudar a probabilidade de morte associada a causas específicas. Este estudo é conhecido como estudos de seguimento sob consideração de riscos competitivos. Chiang (1968) apresenta uma construção detalhada de tábuas de sobrevivência incluindo o caso de riscos competitivos.

Entretanto a análise de estudo de segmento tem se desenvolvido devido a seu papel relevante na tomada de decisões, principalmente na área médica. Littel (1952), Cutler e Ederer (1958) e Elveback (1958), entre outros estudaram esse problema propondo métodos que desencadearam toda uma abordagem específica.

Em estudos de segmento devemos saber exatamente o número de indivíduos expostos ao risco de morte. Esse problema aparece porque podemos ter indivíduos que podem ser considerados como observações perdidas. Em um primeiro caso temos os indivíduos que são retirados do segmento por razões como intervenção de outra doença, morte acidental, ou término do período de observação. Em outro caso classificamos os indivíduos que são eliminados do segmento por causa de migrações, ou falta de cooperação, ou ainda morrem por causas que não são de interesse do estudo, de modo que a condição clínica posterior ao último registro permanece desconhecida.

Diversos autores, entre eles Chiang (1968), consideram este último caso também como um risco competitivo. Outros, porém, não fazem distinção entre os dois casos (Cutler e Ederer, 1958).

Neste trabalho utilizaremos este último conceito, tratando todos esses indivíduos como simplesmente *censurados*, de tal sorte que a única coisa que sabemos é que o tempo de falha é maior que o tempo registrado.

Cabe notar que o simples abandono dessas observações acarretam uma sub-estimativa da curva de sobrevivência.

1.2. ALGUNS MÉTODOS DE ANÁLISE DE CURVAS DE SOBREVIVÊNCIA

Os mais antigos métodos de estimação da curva de sobrevivência são geralmente baseados em alguma suposição paramétrica a respeito da forma da distribuição. O método atuarial, entretanto, além de não fazer nenhuma suposição a esse respeito, também considera o emprego de observações censuradas (Elveback, 1958).

Em 1958, Kaplan e Meier formularam estimadores de máxima verossimilhança da função de sobrevivência considerando o caso de observações censuradas, sem fazer qualquer sorte de suposição a respeito da forma da função. Esses estimadores não paramétricos são conhecidos como estimadores do produto-limite. Em 2.1. veremos algumas diferenças e semelhanças entre o método atuarial e o do produto limite.

Com base no teste dos postos de Wilcoxon, Gehan (1965) fez uma generalização para comparar duas curvas de sobrevivên

cia, incluindo o caso de observações censuradas. Breslow (1970) generalizou o teste de Kruskal-Wallis para comparar k -curvas de sobrevivências, também sujeitas a observações censuradas. Entretanto Peto e Peto (1972) apresentaram uma generalização do teste de Wilcoxon que tem propriedades de invariância dos postos que a generalização de Gehan não possui, com relação aos vários tipos de censuras.

Com o objetivo de estimar e ajustar modelos de curvas de sobrevivência, Gehan e Siddiqui (1973) propuseram métodos de estimação por procedimentos de mínimos quadrados para as distribuições Exponencial, Linear, Gompertz ou Weibull. A escolha do modelo apropriado é feita através da investigação da verossimilhança para as quatro distribuições.

Usando modelos de regressão linear e mínimos quadrados ponderados, Kock, Johnson e Tolley (1972), entre outros, propuseram uma metodologia geral para testar hipóteses de contingência, baseada na metodologia geral de análise de dados categóricos através de modelos lineares, inicialmente proposta por Guizzle, Starmer e Koch (1969).

Quando a cada indivíduo está associado um vetor de informações as metodologias citadas anteriormente podem ser modificadas.

Feigl e Zelen (1965) apresentaram um método de estimação de distribuição de sobrevivência quando os tempos são supostos exponencialmente distribuídos com parâmetros iguais a uma função linear das covariáveis. Os coeficientes são estimados pelo método de máxima verossimilhança. Zippin e Armitage (1966) estenderam os resultados quando ocorre censura. Considerando a

razão de risco como o produto de duas funções: uma, de covariáveis e coeficiente de regressão desconhecida e a outra, uma função de tempo arbitrária e desconhecida, Cox (1972) sugeriu uma análise de tempos de falha, com a flexibilidade de introduzir observações censuradas. Baseados na verossimilhança parcial, Cox (1972, 1975) obteve os estimadores dos coeficientes de regressão pelo método de máxima verossimilhança. Inferências são feitas para comparar grupos e/ou selecionar variáveis que influenciem na distribuição de sobrevivência, a partir das propriedades assintóticas dos estimadores de máxima verossimilhança.

Muitas vezes em segmentos médicos é importante a informação do estágio da doença na análise de curvas de sobrevivência. Lagakos (1976) apresentou um modelo para tal situação na presença de uma só covariável. Beck (1979) estendeu esse modelo para o caso de várias covariáveis e na presença de riscos competitivos.

1.3. FORMULAÇÃO GERAL DOS MODELOS DE REGRESSÃO

Antes de introduzir alguns modelos de regressão vamos estabelecer a terminologia que empregaremos em todo este trabalho.

Considerando uma amostra de indivíduos de uma população podemos observá-los até que ocorra seu tempo de falha (morte). Sendo T uma variável aleatória que representa o tempo de falha de um indivíduo, a função sobrevivência é definida por:

$$F(t) = P[T \geq t] \quad (1.1)$$

que é uma função não crescente no tempo e $F(0) = 1$ e $F(\infty) = 0$.

A função densidade de probabilidade é obtida através da derivada de $F(t)$

$$f(t) = - \frac{\partial F(t)}{\partial t} \quad (1.2)$$

A razão de risco ou razão de falha específica é definida por:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P[t \leq T < t + \Delta t \mid t \leq T]}{\Delta t} = \frac{f(t)}{F(t)} \quad (1.3)$$

Dada uma das funções acima podemos obter as outras.

De uma maneira geral, vamos sempre considerar que temos n_0 indivíduos dos quais em n temos observado o tempo de falha e $n_0 - n$, o tempo de censura. Estudaremos o caso específico quando o tempo de falha é uma variável aleatória contínua.

Em cada um dos n_0 indivíduos temos associado um vetor $p \times 1$ de covariáveis, possivelmente dependentes do tempo. Para o j -ésimo, o vetor é

$$\underline{z}'_{(j)} = (z_{1j}, z_{2j}, \dots, z_{pj}) \quad (1.4)$$

Com o objetivo de estudar a relação entre a distribuição do tempo de sobrevivência e as covariáveis, de uma maneira geral podemos expressar a razão de falha por

$$\lambda(t, \underline{z}) = \lambda_0(t) g(\underline{\beta}, \underline{z}) \quad (1.5)$$

onde

$\lambda_0(t)$: função não negativa do tempo t

$g(\underline{\beta}, \underline{z})$: função das covariáveis $\underline{z}$, com parâmetros $\underline{\beta}$.

Por exemplo, motivados pelo problema de que a distri
buição de sobrevivência de pacientes leucêmicos é influenciada
pela contagem de glóbulos brancos no sangue na época do diagnós
tico, Feigl e Zelen (1965) apresentaram um método de estimação.
Assumindo que o tempo de sobrevivência de cada paciente segue
uma distribuição exponencial, cujo parâmetro pode ser expresso
por uma função linear das covariáveis, o modelo (1.5) é repre
sentado por

$$\lambda(t, \underline{z}) = (\beta_1 + \beta_2 z)^{-1} \quad (1.6)$$

onde em (1.5)

$$\lambda_0(t) = 1$$

e

$$g(\underline{\beta}, \underline{z}) = (\beta_1 + \beta_2 z)^{-1}$$

sendo z a contagem de glóbulos brancos na época do diagnóstico.

A metodologia por eles desenvolvida considera apenas
estudos de segmento onde não haja censura. Através do logarít
mo da função de sobrevivência, L , as estimativas de máxima ve
rossimilhança são obtidas resolvendo-se as equações

$$\frac{\partial L}{\partial \beta_\xi} = 0 \quad , \quad \xi = 1, 2 \quad (1.7)$$

por processos iterativos.

Zipin e Armitage (1966) generalizaram essa metodolo
gia permitindo a inclusão de observações censuradas.

Cox (1972) propôs uma metodologia de estimação do mo
delo (1.5) na forma

$$\lambda(t, \underline{z}) = \lambda_0(t) \exp(\underline{\beta}' \underline{z}) \quad (1.8)$$

onde

$\lambda_0(t)$: é uma função desconhecida

e

$$g(\underline{\beta}, \underline{z}) = \exp(\underline{\beta}' \underline{z})$$

Nesse trabalho as inferências sobre o parâmetro $\underline{\beta}$ são baseadas no conceito de verossimilhança condicional. Mais tarde Kalbfleisch e Prentice (1973) estenderam o conceito para verossimilhança marginal e Cox (1975), para verossimilhança parcial.

Prentice (1973) propôs uma metodologia de estimação de um caso particular do modelo de regressão de Cox, baseada na função de verossimilhança (completa). O modelo (1.5) é da forma

$$\lambda(t, \underline{z}) = \alpha \exp(\underline{\beta}' \underline{x}) \quad (1.9)$$

onde

$$\lambda_0(t) = \alpha$$

e

$$g(\underline{\beta}, \underline{z}) = \exp(\underline{\beta}' \underline{x})$$

e $\underline{x}$ é o valor $\underline{z}$ de covariáveis corrigido pelo vetor média amostral. Dessa maneira α pode ser interpretada como a razão de risco sob um conjunto médio de valores de covariáveis.

Linearizando o modelo (1.9) através da transformação logarítmica, a estimação pelo método de máxima verossimilhança é obtida.

Além do modelo exponencial proposto em (1.9), essa metodologia permite considerar o modelo com distribuição Weibull, isto é,

$$\lambda_0(t) = \lambda \alpha \quad t^{\alpha-1} \quad (1.10)$$

Nas duas próximas seções apresentamos informalmente o modelo de Cox usando parte dos resultados de um estudo sobre alotransplante renal humano (Aquino, 1979).

1.4. UM EXEMPLO PARA INTRODUIR A METODOLOGIA

Vinte e dois pacientes portadores de insuficiência renal crônica terminal que haviam realizado tratamento dialítico e posterior alotransplante renal com doador cadáver (Aquino, 1979) foram submetidos a dois esquemas terapêuticos imunossupressores.

Na tabela 1.1 temos registrado o tempo de sobrevivência ao transplante de cada indivíduo e o respectivo esquema terapêutico, que por simplicidade indicaremos por esquema A e esquema B.

Temos $n_0 = 22$ indivíduos, dos quais $n = 18$ temos registrado o tempo de falha e o restante, $n_0 - n = 4$, o tempo de censura.

Se associarmos a cada indivíduo o vetor de informação (t, z) , onde t é o tempo de falha ou censura e z é uma variável dicotômica que assume o valor zero se o indivíduo recebeu o esquema A e 1, caso tenha recebido o esquema B, o modelo

$$\lambda(t, z) = \lambda_0(t) \exp(\beta' z) \quad (1.8)$$

assume a forma

$$\lambda(t, z) = \begin{cases} \lambda_0(t) & \text{se o indivíduo recebeu o esquema A} \\ \lambda_0(t) e^{\beta} & \text{se o indivíduo recebeu o esquema B} \end{cases} \quad (1.11)$$

TABELA 1.1. Tempo de sobrevivência (dias) ao transplante renal de doador cadáver de 22 indivíduos submetidos a dois esquemas terapêuticos (Aquino, 1979)

nº de ordem do paciente	tempo de sobrevivência	esquema terapêutico
1	2435	A
2	1473	A
3	822*	B
4	552*	B
5	494*	B
6	381	A
7	288*	B
8	265	A
9	213	A
10	200	B
11	199	A
12	109	B
13	105	A
14	97	B
15	66	B
16	53	B
17	46	B
18	40	A
19	28	A
20	24	B
21	22	A
22	14	A

(*) - censura

Vamos fixar os tempos onde ocorreram falhas somente, e em cada um desses pontos formar subconjuntos dos indivíduos que possuem tempo de falha igual ou superior àquele indexado.

Na tabela 1.2. temos os subconjuntos com os respectivos números de ordem do paciente.

Parece razoável considerar a probabilidade de morte de um indivíduo no instante $t_{(i)}$ como sendo

$$\frac{\exp(\beta z_{(i)})}{\sum_{l \in R(t_{(i)})} \exp(\beta z_{(l)})}, \quad i = 1, \dots, 18 \quad (1.12)$$

onde em $R(t_{(i)})$ estão os indivíduos que possuem tempo de falha igual ou superior a $t_{(i)}$.

Podemos, então, dizer que a função de verossimilhança é da forma

$$\prod_{i=1}^{18} \frac{\exp(\beta z_{(i)})}{\sum_{l \in R(t_{(i)})} \exp(\beta z_{(l)})} \quad (1.13)$$

Convém notar que para construir a função de verossimilhança só precisamos ordenar as observações segundo o tempo de sobrevivência, já que z não depende do tempo, pois é a variável indicadora do esquema terapêutico empregado.

TABELA 1.2. Subconjuntos dos indivíduos que possuem tempo de sobrevivência igual ou maior àquele indexado

tempo de falha: $t_{(i)}$	subconjunto	nº de ordem dos indivíduos que pertencem ao subconjunto
14	$R(t_{(1)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
22	$R(t_{(2)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21
24	$R(t_{(3)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20
28	$R(t_{(4)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19
40	$R(t_{(5)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18
46	$R(t_{(6)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17
53	$R(t_{(7)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16
66	$R(t_{(8)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15
97	$R(t_{(9)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14
105	$R(t_{(10)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13
109	$R(t_{(11)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
199	$R(t_{(12)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11
200	$R(t_{(13)})$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
213	$R(t_{(14)})$	1, 2, 3, 4, 5, 6, 7, 8, 9

continuação da Tabela 1.2

tempo de falha: $t_{(i)}$	subconjunto	nº de ordem dos indivíduos que pertencem ao subconjunto
265	$R(t_{(15)})$	1, 2, 3, 4, 5, 6, 7, 8
381	$R(t_{(16)})$	1, 2, 3, 4, 5, 6
1473	$R(t_{(17)})$	1, 2
2435	$R(t_{(18)})$	1

1.5. FAZENDO INFERÊNCIAS SOBRE O PARÂMETRO DE REGRESSÃO

Para encontrar a estimativa de máxima verossimilhança de β vamos usar as técnicas usuais. O logaritmo de (1.13) é

$$L(\beta) = \sum_{i=1}^{18} z_{(i)} - \sum_{i=1}^{18} \ln \sum_{l \in R(t_{(i)})} \exp(\beta z_{(l)}) \quad (1.14)$$

Vamos calcular a primeira e segunda derivada de (1.14) que são, respectivamente,

$$U(\beta) = \frac{\partial L(\beta)}{\partial \beta} = \sum_{i=1}^{18} [z_{(i)} - A_i(\beta)] \quad (1.15)$$

e

$$\phi(\beta) = - \frac{\partial^2 L(\beta)}{\partial \beta^2} = \sum_{i=1}^{18} C_i(\beta) \quad (1.16)$$

onde

$$A_i(\beta) = \frac{\sum_{l \in R(t_{(i)})} z_l \exp(\beta z_{(l)})}{\sum_{l \in R(t_{(i)})} \exp(\beta z_{(l)})} \quad (1.17)$$

e

$$C_i(\beta) = \frac{\sum_{l \in R(t_{(i)})} z_l^2 \exp(\beta z_{(l)})}{\sum_{l \in R(t_{(i)})} \exp(\beta z_{(l)})} - A_i^2(\beta) \quad (1.18)$$

para $i = 1, \dots, 18$

Repare que (1.17) é a média dos z no conjunto $R(t_{(i)})$ usando ponderação exponencial e (1.18) é a variância amostral dos z no mesmo conjunto, também usando ponderação exponencial.

As estimativas de máxima verossimilhança podem ser calculadas usando processos iterativos em (1.15) e (1.16).

Para testar se existe diferença na sobrevivência dos indivíduos que receberam esquemas terapêuticos distintos, supomos que $U(0)$ tem distribuição normal com média zero e variância $\phi(0)$. Assim

$$U(0) / \sqrt{\phi(0)} \quad (1.19)$$

terá distribuição assintótica normal padrão sob a hipótese nula, $\beta = 0$.

Lembrando que a variável z é uma variável indicadora do esquema terapêutico B , as formas (1.15) e (1.16) tornam-se, no ponto $\beta = 0$,

$$U(0) = 7 - \sum_{i=1}^{18} A_i \quad (1.20)$$

e

$$\phi(0) = \sum_{i=1}^{18} A_i (1 - A_{(i)}) \quad (1.21)$$

onde A_i é a proporção dos indivíduos que têm tempo de falha igual ou superior a t_i e que receberam o esquema terapêutico B .

Usando os cálculos parciais da tabela 1.3 em (1.20) e (1.21) encontramos:

$$U(0) = - 1.2001$$

e

$$\phi(0) = 3.9774$$

O valor da estatística (1.19) é, portanto,

$$\frac{U(0)}{\sqrt{\phi(0)}} = - 0.6018$$

e o nível descritivo do teste é 54.85%.

Logo não existe diferença entre os dois esquemas terapêuticos em relação a curva de sobrevivência.

Apesar de β ter sido considerado estatisticamente igual a zero, calculamos o valor de β que maximiza (2.4) por processos iterativos e obtivemos

TABELA 1.3. Cálculos parciais para se testar a hipótese de que os dois esquemas terapêuticos produzem a mesma curva de sobrevivência

tempos de falha		$R(t, i)$		A_i		$A_i(i-A_i)$	
esquema A	esquema B	# que receberam esquema A	# que receberam esquema B	total			
14		11	11	22	0.5000	0.2500	
22		10	11	21	0.5238	0.2494	
	24	9	11	20	0.5500	0.2475	
28		9	10	19	0.5263	0.2493	1
40		8	10	18	0.5556	0.2469	16
	46	7	10	17	0.5882	0.2422	1
	53	7	9	16	0.5625	0.2461	
	66	7	8	15	0.5333	0.2489	
	97	7	7	14	0.5000	0.2500	
105		7	6	13	0.4615	0.2485	
	109	6	6	12	0.5000	0.2500	
199		6	5	11	0.4545	0.2517	
	200	5	5	10	0.5000	0.2500	

continuação da tabela 1.3

tempos de falha		$R(t, i)$			
esquema A	esquema B	# que receberam esquema A	# que receberam esquema B	total	$A_i \quad A_i(i-A_i)$
213		5	4	9	0.4444 0.2469
265		4	4	8	0.5000 0.2500
381		3	3	6	0.5000 0.2500
1473		2	0	2	0.0000 0.0000
2435		1	0	1	0.0000 0.0000

$$\hat{\beta} = - 0.3032$$

Como a variância assintótica de $\hat{\beta}$ é o inverso do negativo da informação de Fisher, encontramos:

$$\Sigma(\hat{\beta}) = [\phi(\hat{\beta})]^{-1} = 0.2555$$

e portanto o erro padrão é

$$ep(\hat{\beta}) = 0.5055$$

No próximo capítulo desenvolvemos a metodologia aqui empregada, sem muito rigor teórico.

C A P Í T U L O   I I

O MODELO DE REGRESSÃO DE COX - ESTIMAÇÃO

Neste capítulo apresentamos a estimação dos parâmetros de regressão no modelo proposto por Cox (1972). O método de estimação é uma generalização das estimativas do produto limite. A teoria desses estimadores (Kaplan e Meier, 1958) é mostrada na primeira seção. Logo a seguir temos o conceito de verossimilhança parcial (Cox, 1975) que permite fazer a formulação geral da verossimilhança apresentada na terceira seção.

Em segmentos médicos, um grupo de tratamento de pacientes e os valores de variáveis de prognóstico são tomadas em um determinado instante de tempo (primeiros sintomas, admissão no hospital, diagnóstico, tratamento, etc...) a partir do qual

a sobrevivência é medida em todos os pacientes. Os efeitos desses valores fixados na sobrevivência subsequente são examinados, mas nenhuma atenção é dada a mudanças no status do paciente que podem ocorrer durante o período de observação. Em alguns estudos, principalmente os de doenças crônicas, tais mudanças são passíveis de acontecer e são de fato o objetivo principal. É conveniente, portanto, saber a eficiência dos estimadores baseados na verossimilhança parcial quando as covariáveis são dependentes do tempo. Algumas considerações são feitas a esse respeito na seção quatro.

A visualização da função de sobrevivência é de interesse na análise crítica do modelo. Na última seção deste segundo capítulo apresentamos a estimação da função de sobrevivência que foi proposta por Cox (1972), mas difere da elaborada por Kaplan e Meier (1958) e por Kalbfleisch e Prentice (1973), pois nestes últimos a razão de risco é considerada contínua nos instantes onde não ocorrem falhas.

2.1. O MÉTODO DO PRODUTO LIMITE

Sendo $t_{(1)} < t_{(2)} < \dots < t_{(n)}$ os tempos de falha observados e ordenados de todos os indivíduos na amostra, a estimativa do produto limite da razão de risco e da função de sobrevivência são dadas, respectivamente, por (Kaplan e Meier, 1958).

$$\hat{\lambda}(t) = \sum_{r: t_r \leq t} \frac{1}{n-r} \quad (2.1)$$

e

$$\hat{F}(t) = \prod_{r: t_r \leq t} \frac{n-r}{n-r+1} \quad (2.2)$$

Considere que a amostra é formada por n_0 indivíduos nos quais n falhas são observadas e o restante, $n_0 - n$, censuras. Com relação aos indivíduos censurados, sabe-se apenas que o tempo de falha é maior que o tempo de censura. Suponha que os tempos de falha são

$$t_{(1)} < t_{(2)} < \dots < t_{(k)}$$

e $m_{(i)}$ é o número de falhas que ocorreram em $t_{(i)}$, $i = 1, \dots, k$,
 $\sum_i m_{(i)} = n$.

O conjunto de indivíduos sob risco no instante imediatamente anterior a t , $t-0$, será denotado por $R(t)$; nele estão todos os $r_{(i)}$ indivíduos cujo tempo de falha ou censura é no mínimo $t_{(i)}$.

A estimativa do produto limite da razão de risco é obtida por (Cox, 1972; Kaplan e Meier, 1958).

$$\hat{\lambda}(t) = \sum_{i=1}^k \frac{m_{(i)}}{r_{(i)}} \delta(t - t_{(i)}) \quad (2.3)$$

onde

$$\delta(t - t_{(i)}) = \begin{cases} 1 & \text{se } t \geq t_{(i)} \\ 0 & \text{se } t < t_{(i)} \end{cases}$$

e

$\frac{m_{(i)}}{r_{(i)}}$ é a frequência condicional esperada.

Assim podemos estimar a função de sobrevivência por

$$\hat{F}(t) = \prod_{t_{(i)} < t} \left(1 - \frac{m(i)}{r(i)}\right) \quad (2.4)$$

A relação dada em (2.2) é semelhante a estimativa atuarial, ou seja, quando não ocorrem censuras as duas estimativas coincidem. O mesmo não acontece em (2.4), pois na estimativa atuarial o número de indivíduos sob risco é igual ao número de indivíduos vivos no início do intervalo menos metade dos indivíduos censurados no intervalo entre duas falhas (Gehan, 1969).

As estimativas dadas nas relações (2.1) à (2.4) são estimadores de máxima verossimilhança para a família de todas as distribuições possíveis (Kaplan e Meier, 1958).

2.2. A VEROSSIMILHANÇA PARCIAL

Sendo $\underline{y}$ um vetor de observações da variável aleatória $\underline{Y}$ com função densidade $f_{\underline{Y}}(\underline{y}; \theta)$, considere as variáveis $(\underline{V}, \underline{W})$ que são transformações da variável $\underline{Y}$ que não dependem de parâmetros desconhecidos. Sejam $f_{\underline{V}}(\underline{v}; \theta)$ a verossimilhança marginal de $\underline{V}$ e $f_{\underline{W}|\underline{V}}(\underline{w}|\underline{v}; \theta)$, a verossimilhança condicional de $\underline{W}$ dado $\underline{V} = \underline{v}$.

Suponha que $\underline{Y}$ seja transformada na sequência

$$(X_1, S_1, X_2, S_2, \dots, X_m, S_m) \quad (2.5)$$

A verossimilhança completa dessa sequência é (Cox, 1975),

$$\prod_{j=1}^m f_{X_j | X^{(j-1)}, S^{(j-1)}}(x_j | x^{(j-1)}, s^{(j-1)}; \underline{\theta}) \times$$

$$\times \prod_{j=1}^m f_{S_j | X^{(j)}, S^{(j-1)}}(s_j | x^{(j)}, s^{(j-1)}; \underline{\theta}) \quad (2.6)$$

onde

$$x^{(j)} = (x_1, x_2, \dots, x_j)$$

e

$$s^{(j)} = (s_1, s_2, \dots, s_j)$$

O segundo produto em (2.6) é chamado verossimilhança parcial de S baseado na sequência $\{X_j, S_j\}$.

A verossimilhança condicional de W dado $V = v$ corresponde ao caso especial onde $S_1 = W$, $X_1 = V$ e a verossimilhança marginal de V , ao caso especial onde X_1 é uma constante conhecida, $S_1 = V$ e $X_2 = W$.

Tanto a verossimilhança condicional como a marginal são comuns em experimentos deduzidos, mas o mesmo não é verdadeiro em geral para a verossimilhança parcial. Isto por causa da maneira de condicionar os eventos (Cox, 1975).

A verossimilhança parcial é útil especialmente quando é mais simples que a verossimilhança completa, no sentido de envolver apenas parâmetros de interesse e não "nuisance parameters".

Exceto em casos especiais, a verossimilhança parcial de S baseada em $\{X_j, S_j\}$ não é a mesma que a verossimilhança marginal de S ou a verossimilhança condicional de S dado $X = x$. Para tanto note que a verossimilhança marginal de S é

(Cox, 1975)

$$\prod_{j=1}^m f_{S_j | S^{(j-1)}}(s_j | s^{(j-1)}; \theta) \quad (2.7)$$

que para ser igual a verossimilhança parcial, as sequências $\{X_j\}$ e $\{S_j\}$ devem ser independentes.

Da mesma maneira, a verossimilhança condicional de S dado $X = x$ é (Cox, 1975)

$$\prod_{j=1}^m f_{S_j | S^{(j-1)}, X}(s_j | s^{(j-1)}, x; \theta) \quad (2.8)$$

que para ser idêntica a verossimilhança parcial temos que ter para todo j , dado $S^{(j-1)} = s^{(j-1)}$, $X^{(j)} = x^{(j)}$, que S_j é independente de $(x_{j+1}, \dots)$.

Como estabelecido nas seções anteriores, temos n_0 indivíduos dos quais em n temos observado o tempo de falha e em $n_0 - n$, o tempo de censura. A cada um dos n_0 indivíduos temos associado o vetor $p \times 1$ de covariáveis possivelmente dependentes do tempo. Para o j -ésimo indivíduo, o vetor é

$$\tilde{z}'_{(j)} = (z_{1j}, z_{2j}, \dots, z_{pj}) \quad (1.4)$$

Considere os tempos de falhas distintos $t_{(1)} < t_{(2)} < \dots < t_{(k)}$, e $m_{(i)}$ o número de falhas que ocorreram em $t_{(i)}$, $\sum_i m_{(i)} = n$ e quando não ocorrem empates $k = n$ e $m_{(i)} = 1$.

No conjunto $R(t)$ de indivíduos sob risco em $t = 0$ estão todos os $r_{(i)}$ indivíduos cujo tempo de falha ou censura é no mínimo $t_{(i)}$.

Se considerarmos a sequência

$$(X_1, S_1, X_2, S_2, \dots, X_k, S_k) \quad (2.9)$$

como X_j para censuras e S_j para falhas, a metodologia apresentada em Cox (1972) pode ser aplicada, usando o conceito de verossimilhança parcial, se os mecanismos de censuras e falhas são independentes (Cox, 1975).

Dessa maneira, ao analisarmos o modelo

$$\lambda(t, \underline{z}) = \lambda_0(t) \exp(\underline{\beta}' \underline{z}) \quad (1.8)$$

cujas função densidade correspondente é

$$f(t) = \lambda_0(t) e^{\underline{\beta}' \underline{z}} \exp\left\{- \int_0^t \lambda_0(u) e^{\underline{\beta}' \underline{z}} du\right\} \quad (2.10)$$

utilizamos a verossimilhança parcial, já que $\lambda_0(t)$ é irrelevante ao estimarmos $\underline{\beta}$.

Para uma particular falha no tempo $t_{(i)}$ condicionada ao conjunto sob risco $R(t_{(i)})$, a probabilidade de que a falha está no i -ésimo indivíduo dos $r_{(i)}$ que estavam sob risco em $t_{(i)}$ é (Cox, 1972; Breslow, 1974)

$$\frac{\exp(\underline{\beta}' \underline{z}_{(i)})}{\sum_{l \in R(t_{(i)})} \exp(\underline{\beta}' \underline{z}_{(l)})} \quad (2.11)$$

supondo-se que não ocorrem empates.

Então a verossimilhança parcial é dada por (Cox, 1972; Kalbfleisch e Prentice, 1973; Cox, 1975; Kay, 1977)

$$\prod_{i=1}^k \frac{\exp(\underline{\beta}' \underline{z}_{(i)})}{\sum_{l \in R(t_{(i)})} \exp(\underline{\beta}' \underline{z}_{(l)})} \quad (2.12)$$

Na próxima seção apresentamos os estimadores de máxima verossimilhança dos parâmetros da regressão e as matrizes de covariância assintótica.

2.3. ESTIMAÇÃO GERAL DO MODELO PROPOSTO POR COX

As estimativas de máxima verossimilhança obtidas pela verossimilhança parcial, sob condições bem gerais, possuem as mesmas propriedades assintóticas dos estimadores obtidos através da verossimilhança tradicional (Cox, 1975).

O logarítmo da verossimilhança parcial (2.12) é

$$L(\underline{\beta}) = \sum_{i=1}^k \underline{\beta}' \underline{z}_{(i)} - \sum_{i=1}^k \ln \left[\sum_{l \in R(t_{(i)})} \exp\{\underline{\beta}' \underline{z}_{(l)}\} \right] \quad (2.13)$$

Cálculos diretos nos dão, para $\xi, \eta = 1, \dots, p$,

$$U_{\xi}(\underline{\beta}) = \frac{\partial L(\underline{\beta})}{\partial \beta_{\xi}} = \sum_{i=1}^k \{z_{\xi i} - A_{(\xi i)}(\underline{\beta})\} \quad (2.14)$$

e

$$\phi_{\xi\eta}(\underline{\beta}) = - \frac{\partial^2 L(\underline{\beta})}{\partial \beta_{\xi} \partial \beta_{\eta}} = \sum_{i=1}^k c_{(\xi\eta i)}(\underline{\beta}) \quad (2.15)$$

onde

$$A_{(\xi i)}(\underline{\beta}) = \frac{\sum_{l \in R(t_{(i)})} z_{\xi l} \exp(\underline{\beta}' \underline{z}_{(l)})}{\sum_{l \in R(t_{(i)})} \exp(\underline{\beta}' \underline{z}_{(l)})} \quad (2.16)$$

e

$$C_{(\xi\eta i)}(\underline{\beta}) = \left[\frac{\sum_{l \in R(t_{(i)})} z_{\xi l} z_{\eta l} \exp(\underline{\beta}' z_{(l)})}{\sum_{l \in R(t_{(i)})} \exp(\underline{\beta}' z_{(l)})} \right] - A_{(\xi i)}(\underline{\beta}) A_{(\eta i)}(\underline{\beta}) \quad (2.17)$$

A quantidade (2.16) é a média amostral dos $z_{(\xi)}$ no conjunto sob risco $R(t_{(i)})$ ponderada exponencialmente e (2.17) é a covariância entre z_{ξ} e z_{η} no conjunto $R(t_{(i)})$, também ponderada exponencialmente.

Chamado de $\phi(\underline{\beta})$ a matriz que tem por elementos os negativos das segundas derivadas parciais (2.15), ao calcularmos o valor esperado de seus elementos é necessário saber o tempo de falha dos indivíduos censurados. Entretanto essas informações são impossíveis de serem obtidas, por isso usamos argumentos assintóticos nos quais as variâncias e covariâncias populacionais são substituídas pelas amostrais (Cox, 1972).

As estimativas de $\underline{\beta}$ em (2.14) e (2.15) podem ser obtidas por processos iterativos (Newton-Raphson). Por exemplo, um bom valor inicial para o vetor de estimativas do vetor de parâmetros de regressão é $\underline{\beta}_0$ igual a zero (Breslow, 1974).

Substituindo o vetor estimado, $\hat{\underline{\beta}}$, em (2.15) e calculando seu inverso, temos a matriz de covariância assintótica das estimativas de máxima verossimilhança, $\Sigma(\hat{\underline{\beta}})$ (Cox, 1975; Kalbfleisch e Mc Intosh, 1977; Efron, 1977; Lubin, 1980).

Testes de significância sobre subconjuntos do conjunto de parâmetros podem ser feitos de várias maneiras. Por exemplo, comparando os logaritmos das verossimilhanças máximas alcançadas. Ou então, um teste relativamente simples para $\underline{\beta} = 0$

é considerando $U(\underline{\theta})$ como assintoticamente normal com média zero e matriz de covariância $\Sigma(\underline{\theta})$. Assim a estatística

$$U'(\underline{\theta}) \Sigma(\underline{\theta}) U(\underline{\theta}) \quad (2.18)$$

tem distribuição assintótica qui-quadrado com p graus de liberdade, onde

$$U(\underline{\theta}) = [U_{\xi}(\underline{\theta})] \quad (2.19)$$

e

$$\Sigma(\underline{\theta}) = \phi^{-1}(\underline{\theta}) = [\phi_{\xi\eta}(\underline{\theta})]^{-1} \quad (2.20)$$

Alguns testes de interesse são apresentados no capítulo IV e vemos que eles são baseados nesses dois conceitos.

O modelo de Cox (1.8) também é conhecido como modelo de regressão com razão de risco proporcional, quando as variáveis não são dependentes do tempo (Breslow, 1974; Kay, 1977; Prentice e Breslow, 1978; Prentice, Kalbfleisch e outros, 1978; Farewell, 1979; Lininger e outros, 1979; Schoenfeld, 1980; entre outros) pois envolve a suposição de razão de risco proporcional que tem sido usada na teoria de riscos competitivos (Chiang, 1968; David e Moeschberger, 1978).

2.4. ALGUMAS PROPRIEDADES DOS ESTIMADORES BASEADOS NA VEROSSIMILHANÇA PARCIAL

Desde que a estimação de $\underline{\beta}$ em (1.8) é válida para qualquer forma da razão de risco $\lambda_0(t)$, diversos trabalhos têm sido feitos calculando-se a eficiência desse estimador com relação a outros modelos.

Kalbfleisch (1974), por exemplo, verificou, no caso de uma covariável, que na vizinhança de $\beta = 0$, a estimação de β em (1.8) é assintoticamente tão eficiente quanto a estimação de β em (1.9), no modelo exponencial, e os resultados para pequenas amostras mostram uma rápida aproximação à eficiência completa com o aumento do tamanho das amostras.

Quando $\beta \neq 0$, usando formas aproximadas para a eficiência relativa da estimação de β em (1.8) e de β em (1.9), e simulando para alguns tamanhos de amostras, o mesmo autor verificou que para a maioria dos casos de importância prática em aplicações médicas, apenas pequena vantagem é atribuída aos estimadores baseados em (1.9) e (1.10), sugerindo que essa qualidade é atribuída às suposições mais severas desses modelos.

Os resultados acima são válidos apenas para o caso de modelo de regressão com somente uma covariável que não dependa do tempo, que possui observações não censuradas e dados não empata-

dos. O método de verossimilhança parcial (Cox, 1975) permite que a matriz de informação limitante para a técnica de Cox seja derivada da mesma maneira que para procedimentos de verossimilhança completa. Construindo um modelo compatível com (1.8) mas usando argumentos heurísticos, baseados na suposição de que o tempo de censura e funções de covariáveis são gerados por processos estocásticos fracamente ergóticos, Oakes (1977), derivou formas explícitas para a matriz de informação limitante. Com esses dois resultados, o autor mostrou que o método de Cox é assintoticamente mais eficiente apenas em circunstâncias especiais. Entretanto o montante da informação perdida em qualquer conjunto de dados é pequena.

Nesse mesmo trabalho, Oakes (1977) estendeu os resultados obtidos por Kalbfleisch (1974), já mencionados. No caso de β igual a zero, podendo existir indivíduos censurados, e no caso de $\beta \neq 0$, além de poder existir observações censuradas, as covariáveis podem ser dependentes do tempo.

Na mesma época Efron (1977) mostrou heurísticamente que a matriz de informação de Fisher real para $\hat{\beta}$ a partir da expressão da densidade (2.10) é assintoticamente igual a matriz de informação baseada na verossimilhança parcial. Isto equivale a mostrar que a estimativa de máxima verossimilhança baseada em (2.12) deve ser assintoticamente equivalente àquela baseada na verossimilhança completa. Esses resultados são exibidos para o caso de covariáveis dependentes do tempo, e parecem sofrer pequena alteração quando existe censura.

2.5. ESTIMAÇÃO DA FUNÇÃO DE SOBREVIVÊNCIA

Muitas vezes é necessário conhecer a forma da função de sobrevivência. O procedimento de estimação sugerido por Cox (1972) é baseado na generalização da estimativa da razão de risco (2.3) proposta por Kaplan e Meier (1958).

Assumindo-se que $\lambda_0(t)$ é idêntica a zero, exceto nos pontos onde ocorreram falhas, calcula-se uma estimativa de máxima verossimilhança em cada um desses pontos.

A contribuição de $\lambda_0(t)$ em $t_{(i)}$ é da forma

$$\frac{\pi_{(i)} \exp(-\beta' \tilde{z}_{(i)})}{1 - \pi_{(i)} + \pi_{(i)} \exp(-\beta' \tilde{z}_{(i)})} \delta(t - t_{(i)}) \quad (2.21)$$

onde

$$\delta(t - t_{(i)}) = \begin{cases} 1 & \text{se } t \geq t_{(i)} \\ 0 & \text{se } t < t_{(i)} \end{cases}$$

e

$\tilde{z}_{(i)}$: é constante arbitrária que na maioria dos casos pode ser a média dos $\tilde{z}$ em $R(t_{(i)})$

A estimativa de máxima verossimilhança de $\pi_{(i)}$ satisfaz

$$\hat{\pi}_{(i)} = \frac{m_{(i)}}{r_{(i)}} - \frac{\hat{\pi}_{(i)}(1 - \hat{\pi}_{(i)})}{r_{(i)}} \times \sum_{j \in R(t_{(i)})} \frac{\exp(\hat{\beta}'(\tilde{z}_{(j)} - \tilde{z}_{(i)})) - 1}{1 - \hat{\pi}_{(i)} + \hat{\pi}_{(i)} \exp(\hat{\beta}'(\tilde{z}_{(j)} - \tilde{z}_{(i)}))} \quad (2.22)$$

que pode ser resolvida por processos iterativos. Quando não existem covariáveis (2.22) reduz-se a estimativa da razão de risco do método do produto-limite (2.1).

Obtendo-se os valores de $\hat{\pi}_{(i)}$, para todo i , $i = 1, \dots, k$, a estimativa da função de sobrevivência é computada por

$$\hat{F}_{\tilde{z}}(t) = \prod_{t_{(i)} < t} \left\{ 1 - \frac{\hat{\pi}_{(i)} \exp(-\hat{\beta}'\tilde{z}_{(i)})}{1 - \hat{\pi}_{(i)} + \hat{\pi}_{(i)} \exp(-\hat{\beta}'\tilde{z}_{(i)})} \right\} \quad (2.23)$$

que também se reduz a estimativa do produto-limite (2.2) se não existem variáveis.

A versão de Kalbfleisch e Prentice (1973) para a estimativa da função de sobrevivência não é baseada no apelo à logística de (2.21), mas sim, é uma aproximação da razão de risco à função escada

$$\lambda_0(t) = \lambda_i, \quad t \in [b_{i-1}, b_i) = B_i, \quad i = 1, \dots, r \quad (2.24)$$

onde $b_0 = 0 < b_1 < b_2 < \dots < b_r = \infty$ é uma subdivisão conveniente da escala de tempo. Dada a estimativa $\hat{\beta}$ de máxima verossimilhança de β , as estimativas de máxima verossimilhança $\hat{\lambda}_i$ para λ_i (Kalbfleisch e Prentice, 1973) são dadas por:

para $i = 1, \dots, r-1$

$$\frac{d_i}{\sum_{l=1}^{q_i} (t_{(i)l} - b_{i-1}) \exp(\hat{\beta}' z_{(i)l}) + (b_i - b_{i-1}) \sum_{k=i+1}^r \sum_{l=1}^{q_k} \exp(\hat{\beta}' z_{(k)l})} \quad (2.25)$$

e para $i = r$

$$\frac{d_r}{\sum_{l=1}^{q_r} (t_{(r)l} - b_{r-1}) \exp(\hat{\beta}' z_{(r)l})} \quad (2.26)$$

onde $t_{(i)l}$, $l = 1, \dots, q_i$ são as observações (censuradas e não censuradas) em B_i com correspondentes covariáveis $z_{(i)l}$ das quais d_i são falhas.

A correspondente estimativa da função de sobrevivência é

$$\ln \hat{F}_{\hat{\beta}}(t) = -t \hat{\lambda}_{k(t)} - \sum_{i=1}^{k(t)} \{ \hat{\lambda}_i \exp(\hat{\beta}' \tilde{z}_i) - \hat{\lambda}_{k(t)} \} b_i \quad (2.27)$$

onde $k(t) = j$ se $t \in B_j$ e $\tilde{z}$ é um valor arbitrário conveniente da covariável.

Este procedimento de estimação da função de sobrevivência é similar àquele sugerido por Oakes (1972) e Breslow (1972). Em vez de selecionar os intervalos $B_1, B_2, \dots, B_r$ independentemente dos dados, eles sugeriram definir os pontos extremos dos intervalos as falhas observadas. Entretanto as curvas obtidas são bem próximas nas duas versões.

No próximo capítulo apresentamos a versão da metodologia de Cox quando o registro do tempo de falha de vários indivíduos coincidem.

C A P Í T U L O   I I I

EXTENSÃO DA ESTIMAÇÃO PARA O CASO DE EMPATES

Nas diversas aplicações práticas, tempos de sobrevi
vência podem ser entendidos como variáveis aleatórias contínuas
e dessa maneira o modelo (1.8) é suficientemente geral. Entretanto,
no registro de tempo de sobrevivência, erros de medida sem
pre são introduzidos e alguns empates podem resultar.

Em seu artigo original, Cox(1972) propôs uma aproximação
à verossimilhança parcial (2.12) inspirado no modelo logístico.
Outras aproximações surgiram (Peto, 1972; Kalbfleisch e
Prentice, 1973; Breslow, 1974, 1975) onde se tenta tornar factível,
computacionalmente, os cálculos para a estimação do parâmetro
de regressão, principalmente no caso onde existem muitos em

pates.

Neste capítulo apresentamos a versão original de Cox (1972) na primeira seção. A abordagem de Kalbfleisch e Prentice (1973) é mostrada na seção seguinte. A aproximação "grosseira" mas factível de Peto (19872) está descrita na terceira seção. Apresentamos também a versão do modelo de Cox (1.8) para o caso de dados agrupados na quarta seção e o controvertido desempate aleatorizado na quinta seção.

3.1. A APROXIMAÇÃO DE COX

Quando em nossos dados aparecem indivíduos distintos com o mesmo tempo de falha, a razão de risco

$$\lambda(t, \underline{z}) = \lambda_0(t) \exp(\underline{\beta}' \underline{z}) \quad (1.8)$$

deve ser modificado ou sua interpretação generalizada.

A aproximação de Cox para a verossimilhança parcial (2.12) é baseada no modelo logístico que considera a razão de risco em (1.8) generalizada para o caso discreto da seguinte forma

$$\frac{\lambda(t, \underline{z}) dt}{1 - \lambda(t, \underline{z}) dt} = \exp(\underline{\beta}' \underline{z}) \frac{\lambda_0(t) dt}{1 - \lambda_0(t) dt} \quad (3.1)$$

Em um particular tempo de falha $t_{(i)}$, a contribuição das $m_{(i)}$ falhas para a verossimilhança é (Cox, 1972).

$$\frac{\exp(\underline{\beta}' \underline{s}_{(i)})}{\sum_{l \in R(t_{(i)}, m_{(i)})} \exp(\underline{\beta}' \underline{s}_{(l)})} \quad (3.2)$$

onde

$\underline{s}_{(i)}$ é a soma dos vetores das covariáveis, $\underline{z}$, dos indivíduos que falharam em $t_{(i)}$

$R(t_{(i)}, m_{(i)})$ é a classe de todos os subconjuntos distintos de $R(t_{(i)})$ de tamanho $m_{(i)}$.

O logaritmo da verossimilhança parcial (2.13) neste caso torna-se

$$L_1(\underline{\beta}) = \sum_{i=1}^k \underline{\beta}' \underline{s}_{(i)} - \sum_{i=1}^k \ln \left[\sum_{\underline{l} \in R(t_{(i)}, m_{(i)})} \exp(\underline{\beta}' \underline{s}_{(l)}) \right] \quad (3.3)$$

As derivadas parciais de primeira e segunda ordem de (3.3) são respectivamente

$$U_{\xi}^{(1)}(\underline{\beta}) = \frac{\partial L_1(\underline{\beta})}{\partial \beta_{\xi}} = \sum_{i=1}^k \left[s_{\xi i} - A_{(\xi i)}^{(1)}(\underline{\beta}) \right] \quad (3.4)$$

e

$$\phi_{\xi \eta}^{(1)}(\underline{\beta}) = - \frac{\partial^2 L_1(\underline{\beta})}{\partial \beta_{\xi} \partial \beta_{\eta}} = \sum_{i=1}^k C_{(\xi \eta i)}^{(1)}(\underline{\beta}) \quad (3.5)$$

para $\xi, \eta = 1, \dots, p$

onde

$$A_{(\xi i)}^{(1)}(\underline{\beta}) = \frac{\sum_{\underline{l} \in R(t_{(i)}, m_{(i)})} s_{\xi l} \exp(\underline{\beta}' \underline{s}_{(l)})}{\sum_{\underline{l} \in R(t_{(i)}, m_{(i)})} \exp(\underline{\beta}' \underline{s}_{(l)})} \quad (3.6)$$

e

$$C_{(\xi\eta i)}^{(1)}(\beta) = \frac{\sum_{l \in R(t(i), m(i))} s_{\xi l} s_{\eta l} \exp(\beta' s_{(l)})}{\sum_{l \in R(t(i), m(i))} \exp(\beta' s_{(l)})} A_{(\xi i)}^{(1)}(\beta) A_{(\eta i)}^{(1)}(\beta) \quad (3.7)$$

para $\xi, \eta = 1, \dots, p; i = 1, \dots, k$

Cabe aqui notar que (3.3) é uma aproximação à verossimilhança real que tem uma forma escabrosa (Peto, 1972). Em cada um dos k-tempos de falhas é necessário calcular

$$\frac{r(i)!}{m(i)! (r(i) - m(i))!} \quad (3.8)$$

somas que correspondem a variação de l em (3.3) no conjunto $R(t(i), m(i))$. Se tivermos, por exemplo, 28 indivíduos no conjunto sob risco, e ocorrerem 4 mortes, teremos 20.475 termos.

Já que (3.3) é uma aproximação, Peto (1972) sugeriu uma outra, que apresentaremos na terceira seção. Antes disso veremos a generalização de (2.13) para o caso de empates proposta por Kalbfleisch e Prentice (1973), uma vez que o apelo logístico existente em (3.1), traz alguns problemas na interpretação da razão de risco no caso de empates. (Kalbfleisch e Prentice, 1972).

3.2. A ABORDAGEM DE KALBFLEISCH E PRENTICE

Existem n_0 indivíduos dos quais de n temos registrado o tempo de falha e de $n_0 - n$ temos o tempo de sobrevivência cen

surado. Em cada um dos k tempos de falha existem $m_{(i)}$ empates,
 $\sum_{i=1}^k m_{(i)} = n$.

Se ordenarmos os tempos de falhas, o vetor de postos não é totalmente possível de se obter; dos $m_{(i)}$ indivíduos que falharam em $t_{(i)}$, sabemos que seus postos são menores do que aqueles que falharam em $t_{(j)}$ $j > i$, mas não conhecemos o arranjo de seus postos. Visto sob este prisma, a verossimilhança marginal de β proposta por Kalbfleisch e Prentice (1973), é proporcional à probabilidade de que o vetor de postos seja um daqueles possíveis, dada a amostra. Essa probabilidade é a soma das probabilidades dos $m_1, \dots, m_k$ vetores de postos possíveis; seria então necessário calcular uma soma k -dimensional de termos do tipo:

$$\frac{\exp \left(\beta' \sum_{i=1}^n z_{(i)} \right)}{\prod_{i=1}^n \sum_{l \in R(t_{(i)})} \exp(\beta' z_{(i)})} \quad (3.9)$$

Se notarmos que os postos dos $m_{(i)}$ indivíduos que falharam em $t_{(i)}$ não são afetados pelos postos atribuídos aos $m_{(j)}$ que falharam em $t_{(j)}$, a soma k -dimensional reduz-se a um produto de k -somas unidimensionais. Sejam Q_i o conjunto das permutações dos símbolos $i_1, \dots, i_{m_i}$, que correspondem aos $m_{(i)}$ indivíduos que falharam em $t_{(i)}$ e seja $(p_1, \dots, p_{m_i})$ um elemento de Q_i considere também $R(t_{(i)})$, o conjunto de indivíduos sob risco em $t_{(i)}$ e $R(t_{(i)}, p_r)$ o conjunto diferença $R(t_{(i)}) - (p_1, \dots, p_{r-1})$. A verossimilhança marginal para β é proporcional a

$$L_2(\beta) \propto \prod_{i=1}^k \exp(\beta' s_{(i)}) \times$$

$$\times (p_1, \dots, p_{m_i}) \in Q_i \prod_{r=1}^{m_i} \left\{ \sum_{l \in R(t_{(i)}, p_r)} \exp(\beta' z_{(l)}) \right\}^{-1}$$

(3.10)

onde $s_{(i)}$ é a soma do vetor z de covariáveis dos indivíduos que falharam em $t_{(i)}$.

Os estimadores de máxima verossimilhança são obtidos encontrando-se o máximo de (3.10) de maneira usual.

Cabe aqui notar que em cada um dos k -tempos de falha é necessário calcular $m_{(i)}$ somas que correspondem ao número de elementos de Q_i . Se tivermos, por exemplo 28 indivíduos sob risco em $t_{(i)}$ e houver 4 mortes, teremos 24 termos que serão somados. Quanto maior o número de empates, maior trabalho teremos em calcular as equações de verossimilhança que originarem de (3.10).

A abordagem de Kalbfleisch e Prentice (1973) na estimação do vetor de parâmetros de regressão β não tem as propriedades assintóticas desejáveis já discutidas na seção 2.4 no caso de variáveis dependentes do tempo. Uma vez que um número maior de suposições são feitas na construção da verossimilhança marginal (3.9), parece essa ser o principal motivo das restrições na sua aplicabilidade (Cox, 1975).

3.3. A APROXIMAÇÃO DE PETO-BRESLOW

Em críticas feitas ao trabalho de Cox (1972), tanto Peto (1972), quanto Breslow (1972) sugeriram outra aproximação à generalização de (1.8) para o caso de empates.

A função de verossimilhança (2.13) sugerida por Cox (1972) não é exatamente a função de verossimilhança correta se o tempo for contínuo e os postos empatados representarem um leve agrupamento. A aproximação de Peto-Breslow também não é a função correta, mas é considerada tão boa aproximação quanto a de Cox, só que é a mais simples de ser manipulada até hoje sugerida.

O logarítmo da função de verossimilhança que deve ser maximizado para se estimar o vetor de parâmetros de regressão, β é (Breslow, 1975)

$$L_3(\beta) = \sum_{i=1}^k \left[\beta' s_{(i)} - m_{(i)} \ln \sum_{l \in R(t_{(i)})} \exp(\beta' z_{(l)}) \right] \quad (3.11)$$

onde $s_{(i)}$ é o vetor soma dos vetores de covariáveis dos $m_{(i)}$ indivíduos que falharam em $t_{(i)}$, $i = 1, \dots, k$.

Calculando as primeiras derivadas parciais de (3.11) obtemos:

$$U_{\xi}^{(3)}(\beta) = \frac{\partial L_3(\beta)}{\partial \beta_{\xi}} = \sum_{i=1}^k \left[s_{\xi i} - m_{(i)} \frac{\sum_{l \in R(t_{(i)})} z_{\xi l} \exp(\beta' z_{(l)})}{\sum_{l \in R(t_{(i)})} \exp(\beta' z_{(l)})} \right] \quad (3.12)$$

e o negativo das derivadas parciais a segunda de L_3 são:

$$\phi_{\xi \eta}^{(3)}(\beta) = - \frac{\partial^2 L_3(\beta)}{\partial \beta_{\xi} \partial \beta_{\eta}} = \sum_{i=1}^k m_{(i)} C_{(\xi \eta i)}^{(3)}(\beta) \quad (3.13)$$

onde $c_{(\xi \eta i)}(\beta)$ é como o definido em (2.17).

Segundo Peto (1972) a variação que existe entre as aproximações é menor que a variação aleatória que está sendo analisada.

3.4. A APROXIMAÇÃO PARA DADOS AGRUPADOS

Em experimentação animal, dados de tempos de falha são agrupados de uma maneira bem grosseira. Modelos de regressão de dados categóricos tais como o modelo logístico linear (3.1) podem ser usados para tais análises. Apesar do parâmetro de regressão no modelo logístico linear ter uma interpretação de risco relativo, o significado dos coeficientes de regressão, depende da escolha de intervalo de agrupamento.

Além do mais a aproximação de Cox para dados empata dos tem se mostrado não factível a não ser que os agrupamentos sejam em intervalos pequenos. A verossimilhança marginal de Kalbfleisch e Prentice (1973) simplifica bastante os cálculos, mas mesmo assim a dificuldade não é reduzida consideravelmente.

A abordagem Peto-Breslow considera uma função de verossimilhança (3.11), computacionalmente mais simples, mas que produz estimadores inconsistentes da função de risco relativo, na presença de tempos de falhas agrupados, com aumento do vício assintótico a medida que o agrupamento se torna grosseiro.

A versão de Prentice e Gloeckler (1978) para o modelo de razão de risco proporcional para dados agrupados é baseada na abordagem de Kalbfleisch e Prentice (1973) para dados agrupados.

Os tempos de falha são agrupados em intervalos de classe $B_i = [a_{i-1}, a_i)$, $i = 1, \dots, r$, com $Q_0 = 0$, $Q_r = \infty$. Generalizando um pouco os resultados de Kalbfleisch e Prentice, as covariáveis podem ser dependentes do tempo, mas são fixadas dentro de um intervalo em particular. A probabilidade de observar um tempo de falha $t_{(i)}$ em um particular indivíduo com vetor de regressão $\underline{z}$ é

$$\left[1 - \alpha_i^{exp(\underline{\beta}' \underline{z})} \right] \prod_{j=1}^{i-1} \alpha_j^{exp(\underline{\beta}' \underline{z})} \quad (3.14)$$

onde

$$\alpha_j = exp \left\{ \int_{a_{i-j}}^{a_j} \lambda_0(u) du \right\}$$

é a probabilidade de sobrevivência condicional em B_j com $\underline{z} = 0$.

Se considerarmos δ uma variável indicadora de falha ($\delta = 0$ se censura, $\delta = 1$ se falha), e o mecanismo de falha e censura independentes, a contribuição à verossimilhança é da forma

$$\left[1 - \alpha_i^{exp(\underline{\beta}' \underline{z})} \right]^\delta \prod_{j=1}^{i-1} \alpha_j^{exp(\underline{\beta}' \underline{z})} \quad (3.15)$$

A função de verossimilhança é o produto de (3.15) sob todos os indivíduos da amostra.

Reparametrizando convenientemente, Prentice e Gloeckler (1978) encontraram as equações de verossimilhança que por processos iterativos fornecem os estimadores de máxima verossimilhança dos parâmetros de regressão.

3.5. DESEMPATE ALEATORIZADO

Alguns métodos estatísticos mais sofisticados muitas vezes exigem recursos computacionais que nem sempre dispomos. O uso de técnicas mais simples é uma alternativa que deve ser usada sabendo-se das limitações de suas conclusões.

Entre as sugestões apresentadas para se obter uma forma de verossimilhança do modelo (1.8) quando ocorrem empates está o desempate aleatorizado (Breslow, 1975; Kay, 1977; Prentice e Gloeckler, 1978; Lubin, 1980).

Como o próprio nome diz os empates são desfeitos aleatoriamente. Assim sendo os postos dos tempos de sobrevivência empatados são colocados segundo um mecanismo aleatório.

Se os recursos computacionais são poucos, o uso de desempate aleatorizado é uma alternativa que pode ser utilizada se decidimos analisar dados de sobrevivência através da técnica de Cox. Este procedimento é indicado também se o número de empates é muito pequeno em relação ao número de observações, caso contrário a aproximação para dados empatados de Prentice e Gloeckler, é recomendada (Lubin, 1980).

No Apêndice podemos comparar as estatísticas obtidas quando usamos a aproximação de Peto-Breslow e quando usamos o desempate aleatorizado para o mesmo conjunto de observações. As conclusões são as mesmas e existem diferenças mínimas entre os dois tipos de abordagem.

No próximo capítulo apresentamos algumas estatísticas usadas em testes de interesse.

C A P Í T U L O   I V

ANÁLISE DO MODELO LINEAR GERAL DA RAZÃO DE RISCO PROPORCIONAL

Na análise de dados de sobrevivência utilizando o mo
delo de regressão de razão de risco proporcional, um dos objeti
vos primeiros é saber que covariáveis alteram a sobrevivência
dos indivíduos sob estudo.

Algumas técnicas gráficas são apresentadas na primei
ra seção. Convém lembrar que existem outras, por exemplo, usan
do a definição geral de resíduos proposta por Cox e Snell(1968),
Kay (1977) sugeriu que fosse feita uma análise de resíduos.

Na segunda seção temos as estatísticas usadas para
testar os parâmetros da regressão, baseadas nas sugestões de
Cox (1972, 1975). Existem outros testes além dos apresenta
dos. Por exemplo, Tarone (1975), derivou um teste de tendên

cia quando existe uma s covariável, motivado pela variação do nível de dose aplicada nos indivíduos sob estudo. Schoenfeld (1980) propôs testes de aderência qui-quadrado para o modelo de regressão de razão de risco proporcional.

O procedimento stepwire está explicado na terceira se
ção.

4.1. SELEÇÃO DE COVARIÁVEIS ATRAVÉS DE ANÁLISE GRÁFICA

Podemos visualizar a influência de covariáveis na so
brevivência dos indivíduos sob estudo ajustando modelos adequada
dos.

Assim, se tivermos, por exemplo, uma covariável dico
tômica que assume os valores 0 ou 1, o modelo (1.8) será

$$\lambda(t, \underline{z}) = \begin{cases} \lambda_0(t) & , \text{ se } z = 0 \\ \lambda_0(t) e^{\beta} & , \text{ se } z = 1 \end{cases} \quad (1.11)$$

Estimando o valor de β no modelo acima com a metodologia
descrita nos capítulos II e III, e também a função de sobrevi
vivência, ou pelo método de Cox (72), ou pelo método de Kalble
fleisch e Prentice (73), que estão descritos na seção 2.5, podemos
fazer o gráfico desta última. Teremos assim as funções de
sobrevivência:

$$\begin{cases} \hat{F}_0(t) & \text{se } z = 0 \\ \hat{F}_1(t) & \text{se } z = 1 \end{cases} \quad (4.1)$$

Se temos diversas variáveis dicotômicas podemos combiná-las convenientemente, formando grupos distintos. Para covariáveis discretas que assumem mais de dois valores, esses procedimentos são estendidos para dar informações a respeito da maneira pela qual elas atuam. Técnicas similares podem ser aplicadas para covariáveis contínuas se agrupamentos em intervalos de classes apropriados forem feitos.

Por exemplo, na aplicação do modelo de Cox na análise estatística efetuada por Kay (1977), a covariável idade foi dicotomizada em valores maiores que a média amostral e em valores não superiores.

4.2. ESTATÍSTICAS EMPREGADAS NA SELEÇÃO DE COVARIÁVEIS

Extendendo as propriedades assintóticas dos estimadores de máxima verossimilhança para a verossimilhança parcial, Cox (1972, 1975) sugeriu que testes de significância sobre parâmetros de regressão do modelo de razão de risco proporcional fossem baseados em dois princípios.

Um deles é a comparação dos logaritmos das verossimilhanças máximas alcançadas. Assim se a adição de u covariáveis ao modelo consegue aumentar a verossimilhança de L para L_u , então

$$-2 \left[\ln L - \ln L_u \right] \quad (4.2)$$

é uma variável aleatória que tem distribuição assintótica qui-quadrado com u graus de liberdade sob a hipótese nula de que os u coeficientes correspondentes a essas covariáveis sejam ze

ro.

O outro é baseado no vetor $U(\underline{\beta})$ de primeiras derivadas do logarítmo da verossimilhança (2.14) e na matriz $\phi(\underline{\beta})$ de menos as derivadas parciais a segunda (2.15) da mesma função. Como já citamos anteriormente na seção 2.3, um teste relativamente simples para $\underline{\beta} = 0$ é calculando o referido vetor e a referida matriz no ponto $\underline{\hat{\beta}} = \underline{0}$. Assim a estatística (Cox, 1972).

$$U(\underline{0}) \Sigma(\underline{0}) U'(\underline{0}) \quad (2.18)$$

tem distribuição assintótica qui-quadrado com p graus de liberdade. A extensão para o caso de empates discutido no capítulo III é imediata.

Para facilitar a explicação do procedimento da próxima seção, vamos considerar algumas estatísticas semelhantes a (2.18). Antes porém, convém lembrar que temos p covariáveis.

A estatística (Cox, 1975)

$$Q_{\xi} = \frac{U_{\xi}^2(0)}{\phi_{\xi}(0)} \quad (4.3)$$

tem distribuição assintótica qui-quadrado com um grau de liberdade; onde $U_{\xi}(0)$ denota a derivada do logarítmo da função de verossimilhança calculada no ponto $\beta_{\xi} = 0$ e $\phi_{\xi}(0)$ é o negativo da segunda derivada do logarítmo da função de verossimilhança, também calculada no ponto $\beta_{\xi} = 0$. A verossimilhança que consideramos em (4.3) só possui um parâmetro de regressão que corresponde a ξ -ésima covariável.

Vamos supor agora, sem perda de generalidade, que no modelo (1.8) existam θ covariáveis. O vetor $\underline{\beta}$ de parâmetros

está particionado em

$$\begin{pmatrix} \beta_{\theta-1} \\ \beta_{\theta} \end{pmatrix} \quad (4.4)$$

onde $\beta_{\theta-1}$ é o vetor dos primeiros $\theta-1$ parâmetros de regressão.

A estatística (Cox, 1975)

$$Q_{\theta} = \frac{U_{\theta}^2(\hat{\beta}_{\theta-1}, 0)}{\phi_{\theta}(\hat{\beta}_{\theta-1}, 0)} \quad (4.5)$$

tem distribuição assintótica qui-quadrado com um grau de liberdade; onde $U_{\theta}(\hat{\beta}_{\theta-1}, 0)$ é a derivada em relação a β_{θ} do logaritmo da função de verossimilhança do modelo que contém θ covariáveis calculada no ponto $\beta_{\theta-1} = \hat{\beta}_{\theta-1}$ e $\beta_{\theta} = 0$ e $\phi_{\theta}(\hat{\beta}_{\theta-1}, 0)$ é a derivada a segunda em relação a β_{θ} , da mesma função no mesmo ponto.

Se considerarmos as propriedades assintóticas dos estimadores de máxima verossimilhança completa, que parecem se estender para os de verossimilhança parcial, podemos ajustar um modelo com ϵ parâmetros e considerar a estatística (Cox, 1975).

$$\frac{(\hat{\beta}_{\psi} - \beta_{\psi})^2}{\Sigma(\hat{\beta}_{\psi})} \quad (4.6)$$

como tendo distribuição de qui-quadrado com um grau de liberdade, sob a hipótese nula de que β_{ψ} é igual um valor fixado, onde $\Sigma(\hat{\beta}_{\psi})$ é o elemento da diagonal de

$$\Sigma(\hat{\beta}) = \phi^{-1}(\hat{\beta}) = \left[\phi_{\xi\eta}(\hat{\beta}) \right]^{-1} \quad (4.7)$$

Convém notar que (2.18) não é só assintoticamente eficiente, mas também de maior poder local entre todos os procedimentos de testes de postos invariantes. Segundo Peto (1972), isto é verdade para qualquer tamanho de amostra, e a estatística $U(0)$ é, portanto, a melhor estatística de teste de postos invariantes para o problema.

O teste logrank (Peto e Peto, 1972) e o teste de Cox (1972) são casos especiais da generalização do teste de Wilcoxon feita por Prentice (1978) para testes sobre os coeficientes de regressão com observações censuradas.

Na próxima seção apresentamos o procedimento stepwise para seleção de covariáveis.

4.3. O PROCEDIMENTO STEPWISE

O modelo de regressão de razão de risco proporcional pode ser construído usando o procedimento stepwise que tanto pode utilizar as estatísticas baseadas no qui-quadrado das verossimilhanças (4.2) como no qui-quadrado das derivadas (4.3 e 4.5).

Vamos primeiramente apresentar o procedimento incluindo variáveis e depois retirando.

O procedimento incluindo variáveis começa de um modelo sem nenhuma variável ou com algumas que devem obrigatoriamente estar no modelo.

No primeiro caso, decide-se incluir primeiramente a

variável que contiver o maior valor da estatística (4.3) entre as p -possíveis. Ela permanecerá no modelo se for significativa a um nível especificado. Caso contrário, ela sairá e o modelo não conterá nenhuma variável.

Se já houver alguma variável no modelo, o objetivo será saturá-lo. Assim para adicionar outra variável, calculamos a estatística (4.5) para todas as variáveis que ainda não entraram. Aquela que tiver o maior valor e for significativa a um nível pré-fixado, entrará no modelo. Caso contrário o procedimento termina.

Dessa maneira, começamos com um modelo reduzido e vamos saturando-o de variáveis ou até que atinja um número pré-fixado de covariáveis, ou até que a última variável a entrar no modelo seja não significativa a um nível especificado.

O stepwise "backward" também pode ser executado, se for conveniente, de maneira semelhante descrita acima, mas partindo-se de um modelo saturado até chegar a um modelo reduzido. Em cada passo calcula-se a estatística (4.6) para todas as variáveis no modelo e verifica-se qual delas possui valor menor da estatística. Se ela não for significativa a um determinado nível sairá do modelo.

Estes procedimentos estão comentados em Bartolucci e Fraser (1977), enquanto que o procedimento stepwise utilizando as estatísticas (4.2) está ilustrado em Kay (1977).

No próximo capítulo apresentamos duas aplicações desses modelos em dois grupos de pacientes que sofreram alotransplante renal.

C A P Í T U L O   V

APLICAÇÕES NUMÉRICAS

Neste capítulo a metodologia de Cox é empregada nos dados do levantamento de alotransplante renal humano, realizado por Aquino (1979).

A população estudada consiste de 104 pacientes de am bos os sexos, portadores de insuficiência renal crônica termi nal, com depuração de creatina endôgena menor que 5 ml /min / 1,73 m³, que haviam realizado tratamento dialítico e posterior alotransplante renal com doador vivo ou cadáver, submetidos a diferentes esquemas terapêuticos imunossupressores, na Unidade de Transplante Renal do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, no pe ríodo de 14 de fevereiro de 1968 a 19 de julho de 1978, com pe

riodo de observação estendido até 31 de dezembro de 1978.

Todos os pacientes foram encaminhados à Unidade de Transplante Renal HC da FMRP/USP com diagnóstico anátomo-patológico de insuficiência renal crônica em fase final, com indicação de tratamento dialítico. Após período de observação variável porém suficiente para estabilização das condições clínicas gerais, indicava-se o alotransplante renal, ao mesmo tempo que se procedia junto aos familiares, averiguações para estudo subsequente dos eventuais doadores. Na inexistência destes ou impossibilidade por razões individuais ou médicas o paciente era programado para receber enxerto de cadáver.

Como o tempo de espera para encontrar um doador compatível é geralmente maior no caso de paciente com doador cadáver do que doador vivo, não há interesse de comparar a sobrevivência desses dois grupos.

Assim sendo, na primeira seção apresentamos aspectos gerais desses dois grupos. Na segunda seção temos os resultados da aplicação da metodologia de Cox no grupo que possui doador cadáver, e na terceira, doador vivo. Na quarta seção temos alguns aspectos computacionais.

5.1. ASPECTOS GERAIS DO ALOTRANSPLANTE RENAL

Os informes sobre os 104 pacientes estudados, desde a fase pré-transplante até a perda do enxerto, óbito ou término do período de observação (31/12/78), foram colhidos de várias fontes, tais como do arquivo central do HC-FMRP/USP, e registros da Unidade de Transplante Renal do mesmo hospital onde ha

viam informações sobre os resultados ambulatoriais mais recentes, cronologicamente subsequentes aos existentes no arquivo central.

O número de pacientes foi reduzido porque nem todos compareceram aos retornos ambulatoriais pré-fixados, alguns já haviam sido submetidos previamente ao alotransplante renal, ou outros receberam droga com atividade imunossupressora conhecida, diferente daquelas que compõe os protocolos terapêuticos em estudo, em qualquer fase evolutiva do transplante. Foram também excluídos os pacientes que tiveram sobrevida inferior a 8 dias ao alotransplante, impossibilitando avaliação de complicações e/ou ocorrências provocadas pelas drogas utilizadas.

Alguns desses motivos não deveriam levar a exclusão do estudo de segmento, tais como o de sobrevida inferior a 8 dias e não comparecimento aos retornos pré-fixados. Porém para a presente análise não foi possível recuperar essas informações.

Assim sendo 82 pacientes que haviam realizado alotransplante renal, sendo 22 com doador cadáver e 60 com doador vivo, receberam dois esquemas terapêuticos imunossupressores diferentes.

Esse tratamento pós operatório constitui-se de drogas que são basicamente as mesmas, mas cuja dosagem é diferente, às quais denominamos A e B.

As informações colhidas de cada um dos 82 indivíduos foram:

sexo

idade: idade à época do alotransplante renal.

tratamento: esquema terapêutico imunossupressor realizado (o es

quema A e o esquema B).

Para o grupo doador cadáver, a variável idade foi transformada na variável

$$idade_c = idade - 30$$

onde 30 é a mediana das idades dos indivíduos cujo doador foi cadáver.

A mesma transformação resultou $idad_v$, onde a mediana das idades é 26 para o grupo doador vivo.

Na próxima seção temos a análise do grupo doador cadáver.

5.2. ANÁLISE DOS DADOS DE DOADOR CADÁVER

Na tabela 5.1 temos os dados referentes aos pacientes do grupo doador cadáver com a codificação das covariáveis de acordo com o estabelecido na seção anterior.

O procedimento de análise do modelo utilizado foi o stepwise incluindo variáveis partindo do modelo sem nenhuma co variável.

A estatística (4.3) é então calculada para todas as covariáveis. Aquela que tiver o maior valor entrará no modelo, se o nível de significância dela for menor que 10%.

Na tabela 5.2 temos o resultado da análise. Com os resultados da tabela 5.2 chegamos a conclusão que nem idade, nem sexo e nem esquema terapêutico alteram a sobrevivência dos pacientes que sofreram alotransplante renal, no modelo de risco

TABELA 5.1: Informações pertinentes ao grupo doador cadáver de acordo com Aquino (1979)

Tempo de Sobrevivência (dias)		sexo	Idade (anos)
Esquema A	Esquema B		
14		F	22
22		M	40
	24	F	26
28		M	53
40		M	27
	46	M	23
	53	M	30
	66	M	19
	97	M	44
105		M	36
	109	M	64
199		M	20
	200	M	31
213		M	33
265		F	38
	288	M	20
381		M	15
	494*	M	54
	552*	M	24
	822*	F	38
1473		F	11
2435		M	30

(*) - Censura

proporcional.

TABELA 5.2 : Resultados passo a passo do procedimento stepwise incluindo variáveis para o grupo doador cadáver.

---

22 observações

18 observações não censuradas

$-2 \times (\text{logaritmo da verossimilhança para o modelo sem nenhuma covariável}) = 84,86$

Passo 0:

Covariável	Estatística $Q_{\xi}(4,3)$	Nível Descritivo
tratamento	0,36	0,5471
sexo	0,03	0,8575
idade	0,10	0,7502

Nenhuma variável obteve o nível de significância mínimo de 10% para entrar no modelo.

Na próxima seção temos a análise no grupo doador vivo.

5.3. ANÁLISE DOS DADOS DE DOADOR VIVO

Na tabela 5.3 temos os dados referentes aos pacientes do grupo doador vivo com a codificação das variáveis de acordo com o estabelecido na primeira seção deste capítulo.

TABELA 5.3 : Informações pertinentes ao grupo doador vivo de
acordo com Aquino (1979)

Tempo de Sobrevivência (dias)		sexo	Idade (anos)
Esquema A	Esquema B		
	13	M	54
	40	M	24
40		M	26
42		M	51
	76	M	34
89		M	12
	102	M	45
	121	M	29
	165*	M	17
	178	M	42
214		F	12
215		M	43
228		F	34
	233	M	43
	267*	M	18
	297	M	33
308		M	18
	319*	F	30
	320*	M	40
	340*	M	12
	347*	F	28
	353	M	42
	354*	M	39

continuação da Tabela 5.3.

Tempo de Sobrevivência (dias)		sexo	Idade (anos)
Esquema A	Esquema B		
	388*	F	15
396		F	22
	416*	M	24
	431*	M	22
	452*	M	26
	459	F	21
	473*	M	21
511		M	26
	522*	M	30
542		F	14
	585*	M	25
	620*	M	22
	641*	F	22
	676*	M	18
	690*	F	27
	741*	M	44
	753	M	39
	767*	M	24
	781*	M	14
	830*	F	22
	865*	M	36
	872*	M	36
956		M	25
970*		M	27

continuação da Tabela 5.3.

Tempo de Sobrevivência (dias)		sexo	Idade (anos)
Esquema A	Esquema B		
1227		M	44
1272*		F	20
1332		F	30
1343*		F	27
1393		M	39
1426*		M	30
1434*		F	21
1517*		F	20
1583		F	25
1588*		F	37
1609		F	32
1744*		M	12
1973		M	22

(*) - Censura

O procedimento de análise do modelo utilizado foi o stepwise retirando variáveis partindo de um modelo completo.

Em cada passo a estatística (4.6) é calculada para todas as variáveis no modelo. Aquela que tiver o menor valor sairá do modelo, se o nível de significância dela for maior que 5%, caso contrário permanece.

O processo é feito a cada novo modelo até que mais nenhuma covariável seja retirada. Aí então tenta-se colocar algu

ma variável usando a estatística (4.3) ou (4.5), conforme o caso, se o nível de significância dela for menor que 10%.

Na tabela 5.4 temos o resultado da análise.

TABELA 5.4 : Resultados passo a passo do procedimento stepwise retirando variáveis para o grupo doador vivo.

---

60 observações

27 observações

- $2 \times (\text{logaritmo da verossimilhança para o modelo contendo nenhuma covariável}) = 182,60$

Passo 0:

Variáveis no modelo: tratamento, sexo, idadev

- $2 \times (\text{logaritmo da verossimilhança do modelo}) = 174,15$

Qui-quadrado do modelo = 8.45 com 3 graus de liberdade

Nível descritivo = 3.75%

Variável	$\hat{\beta}$	$ep(\hat{\beta})$	Q_{β}	Nível descritivo
tratamento	-0,2829	0,4559	0,39	54,49%
sexo	-0,5220	0,4854	1,16	28,22%
idadev	0,0518	0,0221	5,51	1,89%

A variável tratamento não atingiu o nível 5% de significância para ficar no modelo, por isso será removida.

continuação da Tabela 5.4.

Passo 1:

Variáveis no modelo: sexo, idade

$$-2 \times (\text{logaritmo da verossimilhança do modelo}) = 174,53$$

Qui-quadrado do modelo = 8,07 com 2 graus de liberdade

$$\text{Nível Descritivo} = 1,77\%$$

<i>Variável</i>	$\hat{\beta}$	$ep(\hat{\beta})$	Q_{β}	<i>Nível Descritivo</i>
<i>sexo</i>	-0,4503	0,4750	0,90	0,3431
<i>idade</i>	0,0518	0,0222	5,44	0,0197

A variável sexo não atingiu o nível 5% de significância para ficar no modelo, por isso será removida.

Passo 2:

Variáveis no modelo: idade

$$-2 \times (\text{logaritmo da verossimilhança do modelo}) = 175,47$$

Qui-quadrado do modelo = 7,14 com 1 grau de liberdade

$$\text{Nível Descritivo} = 0,81\%$$

<i>Variável</i>	$\hat{\beta}$	$ep(\hat{\beta})$	Q_{β}	<i>Nível Descritivo</i>
<i>idade</i>	0,0582	0,0220	7,02	0,81%

A variável sexo atingiu o nível de 5% de significância para ficar no modelo.

continuação da Tabela 5.4.

Estatística Q_0 para variáveis, que não estão no modelo, serem incluídas:

<i>Covariável</i>	<i>Q_0</i>	<i>Nível descritivo</i>
<i>T</i>	<i>0.12</i>	<i>0.7295</i>
<i>sexo</i>	<i>0.82</i>	<i>0.3655</i>

Nenhuma variável encontrou o nível 10% de significância para entrar.

Com os resultados da tabela 5.4 chegamos a conclusão que somente a idade altera a sobrevivência dos pacientes que sofreram alotransplante renal, no modelo de razão de risco proporcional.

Na próxima seção apresentamos alguns aspectos computacionais de interesse.

5.4. ALGUNS ASPECTOS COMPUTACIONAIS

Ao se ajustar um modelo de forma (1.8) é relevante a utilização de programas especiais de computação.

Examinando o exemplo do capítulo I parece razoável calcular a estatística (2.18) sem auxílio de programas especiais. Mas estamos no caso onde não ocorre empates e a única covariável é binária 0-1. Estimando-se o parâmetro de regressão, no entanto, a equação de verossimilhança precisa ser resolvida utilizando-se processos iterativos.

Nas aplicações numéricas dos capítulos I, V, e do apêndice foram utilizados dois procedimentos que fazem parte do Statistical Analysis System (SAS) 79:

- o primeiro deles (COXREGR) ajusta todas as regressões possíveis com as p -covariáveis, num total de $2^p - 1$ modelos do tipo (1.8).

- o segundo procedimento (PHGLM) escolhe o melhor modelo pelo procedimento stepwise, incluindo ou eliminando as covariáveis.

Nos dois procedimentos as equações de verossimilhança são resolvidas pelo processo iterativo de Newton-Raphson, e no caso de empates utiliza-se a aproximação de Peto-Breslow.

C A P Í T U L O   V I

CONCLUSÕES

A metodologia apresentada neste trabalho é baseada, principalmente, nos desenvolvidos por Cox (1972, 1975), Peto (1972), Breslow (1974, 1975). Mais recentemente Prentice, Kalbfleisch e outros (1978) aplicaram-na no caso de riscos competitivos. Lubin (1980) usando a técnica de estimação de mínimos quadrados ponderados, adaptando-a do trabalho de Grizzle, Starmer e Koch (1969), apresentou estatísticas e regras de decisão de interesse para a análise no modelo de Cox.

Muitas vezes pode ser de interesse do pesquisador encontrar uma distribuição de probabilidade para a função desconhecida $\lambda_0(t)$, após a escolha das covariáveis que influenciam ou são de interesse no modelo de regressão. Assim seria possível

vel ajustar a Exponencial ou Weibull, por exemplo, à função $\lambda_0(t)$.

Entendemos, também, que uma análise de resíduos em modelos de regressão deve ser também utilizada para uma melhor escolha e julgamento de um modelo.

Uma das possibilidades para dar continuidade a este trabalho é o desenvolvimento de programas especiais em computação para o ajuste de $\lambda_0(t)$ e para a análise de resíduos para o modelo de regressão. Os artigos de Cox e Snell (1968), Kalbfleisch e Prentice (1973) e Kay (1977) apresentam aspectos de interesse nesse sentido.

A P Ê N D I C E

Os dados do exemplo aqui descrito foram extraídos do manual do SAS (Reinhardt, 1980) - procedimento COXREG.

Esse procedimento foi executado 3 vezes, ajustando todos os modelos de regressão possíveis, na forma de

$$\lambda(t, \underline{z}) = \lambda_0(t) \exp(\underline{\beta}' \underline{z}) \quad (1.8)$$

Como existem 3 covariáveis (Idade, RX, PS), foram obtidos $2^3 - 1 = 7$ modelos distintos.

A primeira execução foi feita com os dados originais. Na segunda execução, o tempo de falha, cuja unidade é dias, foi transformado em meses, surgindo vários empates. Estes últimos desempatados aleatoriamente e o procedimento executado novamente.

A comparação da primeira com a segunda execução reflete a invariância do procedimento quanto a transformações monotônicas crescentes. A diferença nos valores das estatísticas é atribuída exclusivamente a aproximação para dados empatados. A comparação da segunda com a terceira execução mostra a diferença entre as estimativas usando a aproximação de Peto-Breslow para dados empatados e o desempate aleatorizado. A comparação da primeira execução com a terceira fornece as diferenças introduzidas pelo desempate aleatorizado.

Nas tabelas são usados símbolos cujo significado vamos agora introduzir:

Covariável: variável ou variáveis independentes que são usadas no ajuste do modelo.

Original: resultados das estatísticas pertinentes a primeira execução do procedimento COXREG com os dados originais.

Transformado: resultados das estatísticas pertinentes a segunda execução do procedimento COXREG com o tempo de falha transformado em meses.

Desempate Aleatorizado: resultados das estatísticas pertinentes a terceira execução do procedimento COXREG com o tempo de falha transformado em meses e desempatado aleatoriamente.

$\hat{\beta}$: estimativa do vetor de parâmetros da regressão conforme a metodologia descrita nos capítulos II e III. A ordem das estimativas dos parâmetros segue a mesma ordem das covariáveis.

$U(0)$: valor da estatística formada pelo vetor de derivadas do logarítmo da verossimilhança parcial no ponto $\beta = 0$ (2.14).

$\phi(0)$: valor da estatística formada por menos a matriz de derivadas parciais à segunda do logarítmo da verossimilhança parcial no ponto $\beta = 0$ (2.15).

$\phi^{-1}(0)$: valor da estatística formada por menos a inversa da matriz de informação no ponto $\beta = 0$.

$\phi(\hat{\beta})$: valor da estatística formada por menos a matriz de informação no ponto $\beta = \hat{\beta}$.

$\Sigma(\hat{\beta})$: valor da estatística que corresponde a matriz de covariância assintótica das estimativas dos parâmetros de regressão.

$\loglik(0)$: valor do logarítmo da verossimilhança quando $\beta = 0$.

$\loglik(\hat{\beta})$: valor do logarítmo da verossimilhança calculado para $\beta = \hat{\beta}$.

χ_D^2 : valor da estatística usada para testar se o vetor de parâmetros é o vetor zero, baseado nas derivadas da verossimilhança parcial (2.18) que tem distribuição assintótica qui-quadrado com o número de graus de liberdade igual ao número de parâmetros de regressão no modelo. O outro valor que aparece é o nível descritivo do teste.

χ_L^2 : valor da estatística usada para testar se o vetor de parâmetros é o vetor zero, baseada na verossimilhança do modelo com todas as covariáveis e na verossimilhança do modelo sem nenhuma variável (4.3), que tem distribuição assintótica qui-quadrado com o número de graus de liberdade igual ao número de parâmetros de regressão no modelo. O outro valor que aparece é o nível descritivo do teste.

$\chi_{L,Y=0}^2$: valor da estatística usada para testar se o parâmetro correspondente a covariável Y é igual a zero baseada na verossimilhança do modelo completo e na verossimilhança do modelo sem a covariável Y (4.3), que tem distribuição assintótica qui-quadrado com o número de graus de liberdade igual ao número de covariáveis no modelo completo menos um. O outro valor que aparece é o nível descritivo do

teste.

$\chi^2_{L, Y=0/X}$: valor da estatística usada para testar se o parâmetro correspondente a covariável Y é igual a zero, quando X está no modelo. Essa estatística é baseada na verossimilhança do modelo contendo X e Y e na verossimilhança do modelo que só contém X (4.3), que tem distribuição assintótica qui-quadrado com número de graus de liberdade igual a diferença do número de parâmetros entre os dois modelos. O outro valor que aparece é o nível descritivo do teste.

iterações: número de iterações necessárias para estimar $\hat{\beta}$.

TABELA A-1: COVARIÁVEL IDADE

	original	transformado	desempate aleatorizado
$\hat{\beta}$	$1,61620 \times 10^{-1}$	$1,62471 \times 10^{-1}$	$1,62091 \times 10^{-1}$
$U(0)$	$1,00239 \times 10^2$	$0,998459 \times 10^2$	$1,00183 \times 10^2$
$\phi(0)$	$8,19602 \times 10^2$	$8,13719 \times 10^2$	$8,16506 \times 10^2$
$\phi^{-1}(0)$	$1,22010 \times 10^{-3}$	$1,22893 \times 10^{-3}$	$1,22473 \times 10^{-3}$
$\phi(\hat{\beta})$	$4,04191 \times 10^2$	$4,00804 \times 10^2$	$4,03175 \times 10^2$
$\Sigma(\hat{\beta})$	$2,47408 \times 10^{-3}$	$2,49499 \times 10^{-3}$	$2,48031 \times 10^{-3}$
$\loglik(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\loglik(\hat{\beta})$	$-2,78381 \times 10$	$-2,79560 \times 10$	$-2,78765 \times 10$
χ^2_D	$12,26 - 0,0005$	$12,25 - 0,0005$	$12,29 - 0,0005$
χ^2_L	$14,29 - 0,0002$	$14,31 - 0,0002$	$14,33 - 0,0002$
iterações	6	6	6

TABELA A-2: COVARIÁVEL RX

	original	transformado	desempate aleatorizado
$\hat{\beta}$	$-5,96380 \times 10^{-1}$	$-6,15012 \times 10^{-1}$	$-6,05045 \times 10^{-1}$
$U(0)$	-1,76647	-1,82118	-1,79261
$\Phi(0)$	2,93620	2,92908	2,93397
$\Phi^{-1}(0)$	$3,40577 \times 10^{-1}$	$3,41404 \times 10^{-1}$	$3,40835 \times 10^{-1}$
$\Phi(\hat{\beta})$	2,90228	2,90247	2,90345
$\Sigma(\hat{\beta})$	$3,44557 \times 10^{-1}$	$3,44534 \times 10^{-1}$	$3,44417 \times 10^{-1}$
$\loglik(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\loglik(\hat{\beta})$	$-3,44592 \times 10$	$-3,45519 \times 10$	$-3,45007 \times 10$
χ^2_D	1,06 — 0,3026	1,13 — 0,2873	1,10 — 0,2953
χ^2_L	1,05 — 0,3952	1,12 — 0,2903	1,08 — 0,2981
Iterações	4	4	4

TABELA A-3 : COVARIÁVEL PS

	original	transformado	desempate aleatorizado
$\hat{\beta}$	$3,98367 \times 10^{-1}$	$3,97102 \times 10^{-1}$	$4,07964 \times 10^{-1}$
$U(0)$	1,18217	1,17825	1,21159
$\phi(0)$	2,98847	2,90811	2,98934
$\phi^{-1}(0)$	$3,34619 \times 10^{-1}$	$3,34647 \times 10^{-1}$	$3,34522 \times 10^{-1}$
$\phi(\hat{\beta})$	2,90844	2,90811	2,91015
$\Sigma(\hat{\beta})$	$3,43826 \times 10^{-1}$	$3,43866 \times 10^{-1}$	$3,43625 \times 10^{-1}$
$\loglik(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\loglik(\hat{\beta})$	$-3,47505 \times 10$	$-3,48782 \times 10$	$-3,47961 \times 10$
χ^2_D	0,47 — 0,4941	0,46 — 0,4955	0,49 — 0,4835
χ^2_L	0,47 — 0,4935	0,47 — 0,4949	0,49 — 0,4830
iterações	4	4	4

TABELA A-4 : COVARIÁVEIS IDADE E PS

	original	transformado	desempate aleatorizado
$\hat{\beta}'$	$1,61501 \times 10^{-1}$	$1,62384 \times 10^{-1}$	$1,61918 \times 10^{-1}$
	$1,86619 \times 10^{-2}$	$1,33572 \times 10^{-2}$	$2,62010 \times 10^{-2}$
$U'(0)$	$1,00239 \times 10^2$	$0,998459 \times 10^2$	$1,00183 \times 10^2$
	$1,18217$	$1,17825$	$1,21159$
$\hat{\sigma}(0)$	$8,19602 \times 10^2$	$8,13719 \times 10^2$	$8,16506 \times 10^2$
	$1,01357 \times 10$	$1,01684 \times 10$	$1,00580 \times 10$
	$2,98847$	$2,98823$	$2,98934$
$\hat{\sigma}^{-1}(0)$	$1,27352 \times 10^{-3}$	$1,28350 \times 10^{-3}$	$1,27769 \times 10^{-3}$
	$-4,31924 \times 10^{-3}$	$-4,36751 \times 10^{-3}$	$-4,29894 \times 10^{-3}$
	$3,49268 \times 10^{-1}$	$3,49508 \times 10^{-1}$	$3,48987 \times 10^{-1}$
$\hat{\sigma}(\hat{\beta})$	$4,03563 \times 10^2$	$4,00358 \times 10^2$	$4,02311 \times 10^2$
	$2,55054$	$2,59969$	$2,63237$
	$2,80239$	$2,80197$	$2,80215$
$\hat{\Sigma}(\hat{\beta})$	$2,49226 \times 10^{-3}$	$2,51290 \times 10^{-3}$	$2,50101 \times 10^{-3}$
	$-2,26828 \times 10^{-3}$	$-2,33148 \times 10^{-3}$	$-2,34948 \times 10^{-3}$
	$3,58902 \times 10^{-1}$	$3,59054 \times 10^{-1}$	$3,59077 \times 10^{-1}$
$\log \text{lik}(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\log \text{lik}(\hat{\beta})$	$-2,78377 \times 10$	$-2,79557 \times 10$	$-2,78755 \times 10$

continuação da Tabela A-4.

	original	transformado	desempate aleatorizado
χ^2_D	12,26 — 0,0022	12,25 — 0,0022	12,29 — 0,0021
$\chi^2_{L, PS=0}$	$0,8 \times 10^{-5}$ — maior que 0,99	$0,6 \times 10^{-5}$ — maior que 0,99	$2,0 \times 10^{-5}$ — maior que 0,99
$\chi^2_{L, ID=0}$	13,83 — menor que 0,001	13,85 — menor que 0,001	13,84 — menor que 0,001
χ^2_L	14,29 — 0,0008	14,31 — 0,0008	14,33 — 0,0008
iterações	6	6	6

TABELA A-5 : COVARIÁVEIS: IDADE E RX

	original	transformado	desempate aleatorizado
$\hat{\beta}'$	$1,47327 \times 10^{-1}$	$-8,03973 \times 10^{-1}$	$1,47651 \times 10^{-1}$
			$-8,21725 \times 10^{-1}$
			$1,47542 \times 10^{-1}$
			$-8,09851 \times 10^{-1}$
$U'(0)$	$1,00239 \times 10^2$	$-1,76647$	$1,00183 \times 10^2$
			$-1,82118$
			$-1,79261$
$\hat{\sigma}(0)$	$8,19602 \times 10^2$	$8,13719 \times 10^2$	$8,16506 \times 10^2$
	$1,82057 \times 10$	$1,81663 \times 10$	$1,81481 \times 10$
			$2,93397$
$\hat{\sigma}^{-1}(0)$	$1,41499 \times 10^{-3}$	$1,42643 \times 10^{-3}$	$1,41995 \times 10^{-3}$
	$-8,77356 \times 10^{-3}$	$-8,84682 \times 10^{-3}$	$-8,78307 \times 10^{-3}$
			$3,95162 \times 10^{-1}$
$\hat{\sigma}(\beta)$	$4,82084 \times 10^2$	$4,79973 \times 10^2$	$4,81379 \times 10^2$
	$-5,66806$	$-5,69779$	$-5,72063$
			$2,57021$
$\hat{\Sigma}(\beta)$	$2,12955 \times 10^{-3}$	$2,13977 \times 10^{-3}$	$2,13380 \times 10^{-3}$
	$4,69692 \times 10^{-3}$	$4,74442 \times 10^{-3}$	$4,74931 \times 10^{-3}$
			$3,99644 \times 10^{-1}$
$\log \hat{\sigma}^2_k(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\log \hat{\sigma}^2_k(\beta)$	$-2,70419 \times 10$	$-2,71260 \times 10$	$-2,70696 \times 10$

continuação da Tabela A-5

	original	transformado	desempate aleatorizado
χ^2_D	18,56 — 0,0001	18,75 — 0,0001	18,68 — 0,0001
$\chi^2_{L, RX=0}$	1,59 — entre 0,30 e 0,50	1,66 — entre 0,30 e 0,50	1,6138 — entre 0,30 e 0,50
$\chi^2_{L, ID=0}$	15,33 — menor que 0,001	15,50 — menor que 0,001	15,45 — menor que 0,001
χ^2_L	15,89 — 0,0004	15,97 — 0,0003	15,95 — 0,0003
iterações	6	6	6

TABELA A-6 : COVARIÁVEIS RX e PS

	original	transformado	desempate aleatorizado
$\hat{\beta}$	$-5,78227 \times 10^{-1}$ $3,69797 \times 10^{-1}$	$-5,97103 \times 10^{-1}$ $3,67625 \times 10^{-1}$	$-5,84931 \times 10^{-1}$ $3,76525 \times 10^{-1}$
$U'(0)$	$-1,76647$ $1,18217$	$-1,82118$ $1,17825$	$-1,79261$ $1,21159$
$\sigma(0)$	$2,93620$ $-1,45431 \times 10^{-1}$	$2,92908$ $-1,45939 \times 10^{-1}$	$2,93397$ $-1,59272 \times 10^{-1}$
$\sigma^{-1}(0)$	$3,41400 \times 10^{-1}$ $1,66139 \times 10^{-2}$	$3,42237 \times 10^{-1}$ $1,67142 \times 10^{-2}$	$3,41823 \times 10^{-1}$ $1,82124 \times 10^{-2}$
$\hat{\sigma}(\beta)$	$2,90135$ $-1,39796 \times 10^{-1}$	$2,90040$ $-1,38744 \times 10^{-1}$	$2,90199$ $-1,51950 \times 10^{-1}$
$\Sigma(\beta)$	$3,45467 \times 10^{-1}$ $1,65968 \times 10^{-2}$	$3,45568 \times 10^{-1}$ $1,64780 \times 10^{-2}$	$3,45536 \times 10^{-1}$ $1,80384 \times 10^{-2}$
$\loglik(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\loglik(\hat{\beta})$	$-3,42577 \times 10$	$-3,43528 \times 10$	$-3,42919 \times 10$

continuação da Tabela A-6.

	original	transformado	desempate aleatorizado
χ^2_D	1,46 — 0,4808	1,53 — 0,4655	1,51 — 0,4696
$\chi^2_{L,PS=0}$	0,99 — entre 0,30 e 0,50	1,05 — entre 0,30 e 0,50	1,01 — entre 0,30 e 0,50
$\chi^2_{L,RX=0}$	0,40 — entre 0,30 e 0,50	0,40 — entre 0,30 e 0,50	0,42 — entre 0,30 e 0,50
χ^2_L	1,45 — 0,4833	1,52 — 0,4685	1,50 — 0,4723
iterrações	4	4	4

TABELA A-7: COVARIÁVEIS IDADE, RX e PS

	original	transformado	desempate aleatorizado
$\hat{\beta}$	$1,46989 \times 10^{-1}$ $-8,14585 \times 10^{-1}$ $1,03180 \times 10^{-1}$	$1,47323 \times 10^{-1}$ $-8,31907 \times 10^{-1}$ $0,98975 \times 10^{-1}$	$1,47178 \times 10^{-1}$ $-8,20648 \times 10^{-1}$ $1,08785 \times 10^{-1}$
$U'(0)$	$1,00239 \times 10^2$ $-1,76647$ $1,18217$	$0,998459 \times 10^2$ $-1,82118$ $1,17825$	$1,00183 \times 10^2$ $-1,79261$ $1,21159$
$\hat{\sigma}(0)$	$8,19602 \times 10^2$ $1,82057 \times 10$ $1,01357 \times 10$ $2,93620$ $-1,45431 \times 10^{-1}$ $2,98847$	$8,13719 \times 10^2$ $1,81663 \times 10$ $1,01684 \times 10$ $2,92908$ $-1,45939 \times 10^{-1}$ $2,98823$	$8,16506 \times 10^2$ $1,81481 \times 10$ $1,00580 \times 10$ $2,93397$ $-1,59272 \times 10^{-1}$ $2,98934$
$\hat{\sigma}^{-1}(0)$	$1,50183 \times 10^{-3}$ $-9,58738 \times 10^{-3}$ $-5,56013 \times 10^{-3}$ $4,02604 \times 10^{-1}$ $5,21087 \times 10^{-2}$ $3,56013 \times 10^{-1}$	$1,51535 \times 10^{-3}$ $-9,67875 \times 10^{-3}$ $-5,62913 \times 10^{-3}$ $4,04057 \times 10^{-1}$ $5,26682 \times 10^{-2}$ $3,56374 \times 10^{-1}$	$1,50753 \times 10^{-3}$ $-9,62800 \times 10^{-3}$ $-5,58525 \times 10^{-3}$ $4,03313 \times 10^{-1}$ $5,38832 \times 10^{-2}$ $3,56185 \times 10^{-1}$
$\hat{\sigma}(\beta)$	$4,78122 \times 10^2$ $-5,31248$ $8,69858 \times 10^{-1}$ $2,57027$ $2,50733 \times 10^{-1}$ $2,74818$	$4,76135 \times 10^2$ $-5,35489$ $8,73456 \times 10^{-1}$ $2,56944$ $2,50210 \times 10^{-1}$ $2,74843$	$4,77224 \times 10^2$ $-5,94371$ $9,00769 \times 10^{-1}$ $2,57040$ $2,41353 \times 10^{-1}$ $2,74616$
$\hat{\Sigma}(\beta)$	$2,14388 \times 10^{-3}$ $4,53700 \times 10^{-3}$ $-1,09472 \times 10^{-3}$ $4,02160 \times 10^{-1}$ $-3,81110 \times 10^{-2}$ $3,67693 \times 10^{-1}$	$2,15396 \times 10^{-3}$ $4,59641 \times 10^{-3}$ $-1,10298 \times 10^{-3}$ $4,02479 \times 10^{-1}$ $-3,81014 \times 10^{-2}$ $3,67664 \times 10^{-1}$	$2,14870 \times 10^{-3}$ $4,57008 \times 10^{-3}$ $-1,10845 \times 10^{-3}$ $4,02003 \times 10^{-1}$ $-3,68299 \times 10^{-2}$ $3,67744 \times 10^{-1}$
$\log \hat{L}(0)$	$-3,49849 \times 10$	$-3,51111 \times 10$	$-3,50421 \times 10$
$\log \hat{L}(k(\beta))$	$-2,70274 \times 10$	$-2,71126 \times 10$	$-2,70534 \times 10$
χ_D^2	$18,70$ ——— $0,0003$	$18,91$ ——— $0,0003$	$18,82$ ——— $0,0003$
$\chi_{L,ID=0}^2$	$14,46$ ——— menor que $0,001$	$14,48$ ——— menor que $0,001$	$14,48$ ——— menor que $0,001$

continuação da Tabela A-7.

$\chi^2_{L, RX=0}$	1,62 — entre 0,20 e 0,30	1,68 — entre 0,20 e 0,30	1,64 — entre 0,20 e 0,30
$\chi^2_{L, PS=0}$	0,03 — entre 0,80 e 0,90	0,03 — entre 0,80 e 0,90	0,02 — entre 0,80 e 0,90
$\chi^2_{L, ID=0/RX}$	15,45 — menor que 0,001	15,53 — menor que 0,001	15,45 — menor que 0,001
$\chi^2_{L, ID=0/PS}$	14,86 — menor que 0,001	14,88 — menor que 0,001	14,89 — menor que 0,001
$\chi^2_{L, RX=0/PS}$	1,62 — entre 0,10 e 0,20	1,69 — entre 0,20 e 0,30	1,65 — entre 0,20 e 0,30
χ^2_L	15,92 — 0,0012	16,00 — 0,0011	15,98 — 0,0011

iterações

6

6

REFERÊNCIAS BIBLIOGRÁFICAS

1. AQUINO, E.F. (1979) - *Alotransplante Renal Humano: estudo comparativo da sobrevida e das intercorrências com dois esquemas terapêuticos imunodepressores* - Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da USP para obtenção do Título de Mestre em Clínica Médica.
2. BARTOLUCCI, A.A. e FRASER, M.D. (1977) - *Comparative Step-up and Composite Tests for Selecting Prognostic Indicators Associated with Survival* - *Biometric Journal* 19, 437-448.
3. BECK, G.J. (1979) - *Stochastic Survival Models with Competing Risks and Covariates* - *Biometrics* 35, 427-38.
4. BRESLOW, N. (1970) - *A Generalized Kruskal-Wallis Test for Comparing Samples Subject to Unequal Patterns of Censoring* - *Biometrika* 57, 579-94.
5. BRESLOW, N. (1972) - *Contribuição à Discussão de Cox (1972)*.
6. BRESLOW, N. (1974) - *Covariance Analysis of Censored Survival Data* - *Biometrics* 30, 89-99.
7. BRESLOW, N. (1975) - *Analysis of Survival Data under a Proportional Hazards Model* - *International Statistical Review* 43, 45-57.

8. CHIANG, C.L. (1968) - *Introduction to Stochastic Processes in Biostatistics* - Wiley - New York.
9. COX, D.R. (1972) - *Regression Models and Life-Tables* - Journal of the Royal Statistical Society, B 34, 187-220.
10. COX, D.R. (1975) - *Partial Likelihood* - *Biometrika* 62, 269-76.
11. COX, D.R. e SNELL, E.J. (1968) - *A General Definition of Residuals* - *Journal of the Royal Statistical Society* B 30, 248-75.
12. CUTLER, S. e EDERER, F. (1958) - *Maximum Utilization of the Life Table Method in Analysing Survival* - *Journal of Chronical Disease* 8, 699-712.
13. DAVID, H.A. e MOESCHBERGER M.L. (1978) - *The Theory of Competing Risks* - Griffin's Statistical Monographs & Courses n° 39 - Charles Griffin & Company Ltd - London.
14. ELVEBACK, L. (1958) - *Estimation of Survivorship in Chronic Disease: The Actuarial Method* - *Journal of American Statistical Association* 53, 420 - 40.
15. EFRON, B. (1977) - *The Efficiency of the Cox's Likelihood Function for Censored Data* - *Journal of American Statistical Association* 72, 557-65.

16. FAREWELL, V.T. (1979) - *An Application of Cox's Proportional Hazard Model to Multiple Infection Data* - Journal of Applied Probability 28, 136-43.
17. FEIGL, P. e ZELLEN, M. (1965) - *Estimation of Exponential Survival Probabilities with Concomitant Information* - Biometrics 21, 826-38.
18. GEHAN, E.A. (1965) - *A Generalized Wilcoxon Test for Comparing Arbitrarily Single-Censored Samples* - Biometrika 52, 203-24.
19. GEHAN, E.A. (1969) - *Estimating Survival Functions from the Life-Table* - Journal of Chronical Disease 21, 629-44.
20. GEHAN, E.A. e SIDDIQUI, M.M. (1973) - *Simple Regression Methods for Survival Time Studies* - Journal of the American Statistical Association 68, 848-56.
21. GRIZZLE, J.E., STARMER, C.F. e KOCH, G.G. (1969) - *Analysis of Categorical Data by Linear Models* - Biometrics 25, 489-504.
22. KALBFLEISCH, J.D. (1974) - *Some Efficiency Calculations for Survival Distributions* - Biometrika 61, 31-38.
23. KALBFLEISCH, J.D. e MCINTOSH, A.A. (1977) - *Efficiency in Survival Distributions with Time-Dependent Covariates* - Biometrika 64, 47-50.

24. KALBFLEISCH, J.D. e PRENTICE, R.L. (1972) - *Contribuição à discussão de Cox* (1972).
25. KALBFLEISCH, J.D. e PRENTICE, R.L. (1973) - *Marginal Likelihoods Based on Cox's Regression and Life Model* - *Biometrika* 60, 267-78.
26. KAPLAN, E.L. e MEIER, P. (1958) - *Nonparametric Estimation from Incomplete Observations* - *Journal of American Statistical Association* 53, 457-81.
27. KAY, R. (1977) - *Proportional Hazard Regression Models and the Analysis of Censored Survival Data* - *Journal of Applied Probability* 26, 227-37.
28. KOCH, G.G., JOHNSON, W.D. e TOLLEY, H.D. (1972) - *A Linear Models Approach to the Analysis of Survival and Extent of Disease in Multidimensional Contingency Tables* - *Journal of the American Statistical Association* 67, 783-96.
29. LAGAKOS, S.W. (1976) - *A Stochastic Model for Censored-Survival Data in the Presence of an Auxiliary Variable* - *Biometrics* 32, 551-9.
30. LININGER, L.; GAIL M.H.; GREEN, S.B. e BYAR, D.P. (1979) - *Comparison of Four Test for Equality of Survival Curves in the Presence of Stratification and Censoring* - *Biometrika* 66, 419-28.

31. LITTELL, A.S. (1952) - *Estimation of the T-year Survival Rate from Follow-up Studies over a Limited Period of Time* - Human Biology 24, 87-116.
32. LUBIN, J.H. (1980) - *Analysis under Cox's Failure Time Model Using Weighted Least Squares* - Biometrics 36, 307-12.
33. OAKES, D. (1972) - *Contribuição à discussão de Cox* (1972).
34. OAKES, D. (1977) - *The Asymptotic Information in Censored Survival Data* - Biometrika 64, 441-8
35. PETO, R. (1972) - *Contribuição à discussão de Cox* (1972).
36. PETO, R. e PETO, J. (1972) - *Asymptotically Efficient Rank Invariant Test Procedures* - Journal of the Royal Statistical Society A 135, 185-206.
37. PRENTICE, R.L. (1973) - *Exponential Survivals with Censoring and Explanatory Variables* - Biometrika 60, 279-88.
38. PRENTICE, R.L. (1978) - *Linear Rank Tests with Right Censored Data* - Biometrika 65, 167-79.
39. PRENTICE, R.L. e BRESLOW, N.E. (1978) - *Retrospective Studies and Failure Time Models* - Biometrika 65, 153-8.
40. PRENTICE, R.L. e GLOECKLER, L.A. (1978) - *Regression Analysis of Grouped Survival Data with Applications to Breast*

Cancer Data - Biometrics 34, 57-67.

41. PRENTICE, R.L.; KALBFLEISCH, J.D.; PETERSON Jr., A.V.,
FAREWELL, V.T. e BRESLOW, N.E. (1978) - *The Analysis of
Failure Times in the Presence of Competing Risks* - Biome
trics 34, 541-54.
42. REINHARAT, P.S., ed. (1980) - *SAS Supplemental Library
User's Guide* - SAS Institut, Cary, N.C.
43. SCHOENFELD, D. (1980) - *Chi-Squared Goodness-of-fit Tests
for the Proportional Hazard Regression Model* - Biometri
cs 67, 143-53.
44. TARONE, R.E. (1975) - *Tests Trend in Life Table Analysis* -
Biometrika 62, 679-682.
45. ZIPPIN, C. e ARMITAGE, P. (1966) - *Use of Concomitant Varia*
bles and Incomplete Survival Information in the Estima
tion of an Exponencial Survival Parameter - Biometrics
22, 665-72.

ANÁLISE DE CURVAS DE
SOBREVIVÊNCIA — O MODELO DE REGRESSÃO DE COX

Página	Linha	Substituir	Por
22	↑ 5	$\underline{V} = \underline{\theta}$	$\underline{V} = \underline{v}$
23	↑11	deduzidos	derivados
38	↑10	$\exp(\underline{\beta}' \underline{z}_{(i)})$	$\exp(\underline{\beta}' \underline{z}_{(l)})$
38	↑ 3	Q_i considere	Q_i . Considere
42	↓ 9	a_{i-j}	a_{j-1}
42	↑ 7	$\prod_{j=1}^{i-1} \alpha_i \exp(\underline{\beta}' \underline{z})$	$\prod_{j=1}^{i-1} \alpha_j \exp(\underline{\beta}' \underline{z})$
47	↓ 8	$U(\underline{0}) \Sigma(\underline{0}) U'(\underline{0})$	$U'(\underline{0}) \Sigma(\underline{0}) U(\underline{0})$
48	↑11	é a	é menos a
48	↑ 4	$\frac{(\hat{\beta}_{\psi} - \beta_{\psi})^2}{\Sigma(\hat{\beta}_{\psi})}$	$Q_{\beta} = \frac{(\hat{\beta}_{\psi} - \beta_{\psi})^2}{\Sigma(\hat{\beta}_{\psi})}$
60	↓ 7	observações	observações não censu radas
61	↑ 2	sexo	idade

Em todo este trabalho toda vez que aparecer segmento,
substituir por seguimento.