CE077 - Análise de Sobrevivência

Modelo de Regressão de Cox

Silva, J.L.P.

Abril, 2024

Objetivos do Módulo

Objetivos do Módulo

- Introduzir o modelo semiparamétrico de riscos proporcionais de Cox como alternativa aos modelos paramétricos.
- Discutir estimação do modelo com base na função de verossimilhança parcial, testes de hipóteses e diagnóstico/adequação.
- Discutir a interpretação dos coeficientes estimados e as suposições do modelo de riscos proporcionais.
- Apresentar extensões do modelo de Cox: covariáveis repetidas no tempo e violação da suposição básica de taxas proporcionais.
- Exemplificar o uso do modelo por meio de aplicações de dados reais.
- Por fim, discutir direções para outras situações não abordadas.

Introdução

Introdução

Discutimos a modelagem de dados de sobrevivência no contexto paramétrico por meio de distribuições como a exponencial, Weibull ou log-normal.

Um modelo bastante utilizado na prática por sua versatilidade é o modelo de taxas de falhas proporcionais de Cox, ou simplesmente modelo de Cox (Cox, 1972):

- Abriu uma nova fase na modelagem de dados de sobrevivência.
- Atualmente é o modelo mais popular na análise de dados de sobrevivência na área clínica/biológica.
- Permite incorporar facilmente covariáveis dependentes do tempo, que ocorrem com frequência em várias áreas de aplicação.

Modelo de Regressão de Cox

Modelo de Regressão de Cox

- O modelo de Cox assume a seguinte forma para a função de taxa de falha:

  $$\lambda(t) = \lambda_0(t) g(x'\beta),$$

  em que $g(\cdot)$ é uma função não-negativa tal que $g(0) = 1$.

- O componente não-paramétrico, $\lambda_0(t)$, não é especificado e é uma função não-negativa do tempo.

- $\lambda(t) = \lambda_0(t)$, função de base, pode ser obtida para $x = 0$.

- O componente paramétrico é frequentemente usado na seguinte forma multiplicativa:

  $$g(x'\beta) = \exp(x'\beta) = \exp(\beta_1 x_1 + \ldots + \beta_p x_p),$$

  em que β é o vetor de parâmetros associado às covariáveis.

Modelo de Regressão de Cox

- A suposição de taxas de falhas proporcionais significa que, para dois indivíduos diferentes i e j, temos que,

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0 \exp(x_i'\beta)}{\lambda_0 \exp(x_j'\beta)} = \exp\left\{x_i'\beta - x_j'\beta\right\}$$

não depende do tempo.

- Ou seja, se um indivíduo no início do estudo tem taxa de morte igual a duas vezes a taxa de um segundo indivíduo, então esta razão é a mesma para todo o período de acompanhamento.

Ajustando o Modelo de Cox

Modelo de Regressão de Cox

Modelo de Cox:

$$\lambda(t|x) = \lambda_0(t) \exp\left\{x'\beta\right\}.$$

- O modelo de regressão de Cox é caracterizado pelos coeficientes β's, que medem os efeitos das covariáveis sobre a função de taxa de falha.
- Queremos fazer inferência nos coeficientes β a partir das observações amostrais.
- A presença do componente não-paramétrico $\lambda_0(t)$ na função de verossimilhança, traz dificuldades ao processo inferencial.

Modelo de Regressão de Cox

Sabe-se que

$$L(\beta) = \prod_{i=1}^{n} \left[f(t_i|x_i)\right]^{\delta_i} \left[S(t_i|x_i)\right]^{1-\delta_i}$$
$$= \prod_{i=1}^{n} \left[\lambda(t_i|x_i)\right]^{\delta_i} \left[S(t_i|x_i)\right].$$

No modelo de Cox

$$S(t_i|x_i) = \exp\left\{-\int_0^{t_i} \lambda_0(u) \exp\left\{x_i'\beta\right\} du\right\} = \left[S_0(t_i)\right]^{\exp\left\{x_i'\beta\right\}}.$$

Assim, aplicando este resultado na função de verossimilhança, temos

$$L(\beta) = \prod_{i=1}^{n} \left[\lambda_0(t_i) \exp\left\{x_i'\beta\right\}\right]^{\delta_i} \left[S_0(t_i)\right]^{\exp\left\{x_i'\beta\right\}}.$$

Verossimilhança Parcial

A função de verossimilhança usual do modelo de Cox envolve o componente não paramétrico $\lambda_0(t)$.

Cox (1975) desenvolveu uma alternativa que condiciona no conhecimento da história passada de falhas e censuras, eliminando o componente não paramétrico da função de verossimilhança (parcial).

Considerando que, para um amostra de tamanho n, existam $k \leq n$ falhas distintas nos tempos $t_1 < t_2 < \ldots < t_k$, a construção da verossimilhança parcial é obtida do argumento condicional: a probabilidade condicional da i-ésima unidade vir a falhar no tempo t_i conhecendo quais observações estão sob risco em t_i.

Verossimilhança Parcial

No tempo de falha t_i, a contribuição para a verossimilhança parcial é

$$\begin{aligned}
L_i(\beta) =& P(\text{indivíduo } i \text{ falhar} | \text{uma falha em } R(t_i)) \\
=& \frac{P(\text{indivíduo } i \text{ falhar} | \text{em risco em } t_i)}{\sum_{j \in R(t_i)} P(\text{indivíduo } j \text{ falhar} | \text{em risco em } t_i)} \\
=& \frac{\lambda_i(t|x_i)}{\sum\limits_{j \in R(t_i)} \lambda_j(t|x_j)} \\
=& \frac{\lambda_0(t) \exp\{x_i'\beta\}}{\sum_{j \in R(t_i)} \lambda_0(t) \exp\left\{x_j'\beta\right\}} \\
=& \frac{\exp\{x_i'\beta\}}{\sum_{j \in R(t_i)} \exp\left\{x_j'\beta\right\}}.
\end{aligned}$$

Condicional à história de falhas e censuras até o tempo t_i, o componente não paramétrico $\lambda_0(t)$ desaparece.

Verossimilhança Parcial

A função de verossimilhança parcial é dada pelo produto de todos os termos da expressão anterior associados aos tempos distintos de falha.

Assim,

$$
\begin{aligned}
L(\beta) &= \prod_{i=1}^{k} L_i(\beta) \\
&= \prod_{i=1}^{k} \frac{\exp\{x_i'\beta\}}{\sum_{j \in R(t_i)} \exp\left\{x_j'\beta\right\}} \\
&= \prod_{i=1}^{n} \left(\frac{\exp\{x_i'\beta\}}{\sum_{j \in R(t_i)} \exp\left\{x_j'\beta\right\}} \right)^{\delta_i}.
\end{aligned}
$$

Verossimilhança Parcial

O EMVP (Estimador de Máxima Verossimilhança Parcial) é o valor de $\hat{\beta}$ que maximiza $L(\beta)$.

EMVP é obtido resolvendo-se o sistema de equações definido por $U(\beta) = 0$, em que $U(\beta)$ é o vetor escore de primeiras derivadas da função $l(\beta) = \log\ L(\beta)$.

Isto é,

$$U(\beta) = \sum_{i=1}^{n} \delta_i \left[x_i - \frac{\sum_{j \in R(t_i)} x_j \exp\left\{ x_j' \hat{\beta} \right\}}{\sum_{j \in R(t_i)} \exp\left\{ x_j' \hat{\beta} \right\}} \right] = 0.$$

Verossimilhança Parcial

- A função de verossimilhança parcial, $L(\beta)$ estabelecida anteriormente, não pressupõe a possibilidade de empates nos tempos observados de falha.

- Empates podem, contudo, ocorrer nos tempos de falhas devido a medições imprecisas.

- Aproximações para $L(\beta)$, quando ocorrem empates, foram propostas, dentre outros, por Breslow (1972), Peto (1972), Efron (1977).

- A função `coxph` do pacote `survival` do `R` assume a aproximação de Efron (1977) como padrão.

Verossimilhança Parcial

Sob certas condições de regularidade, os estimadores de máxima verossimilhança parcial são consistentes e assintoticamente normais.

Para fazer inferências no modelo de Cox é possível, então, usar as estatísticas de Wald, da Razão de Verossimilhança e Escore.

O teste de Wald é o mais usado para testar hipóteses relativas a um único parâmetro, isto é

$$H_0 : \beta_j = \beta_{0j},\ j = 1, \ldots, p.$$

Interpretação dos Coeficientes

Interpretação dos Parâmetros

- No modelo de Cox, o efeito das covariáveis é de acelerar ou desacelerar a função de taxa de falha.

- A propriedade de taxas de falhas proporcionais do modelo é utilizada para interpretar os coeficientes estimados.

- Tomando a razão das taxas de falha de dois indivíduos i e j que têm os mesmos valores para as covariáveis com exceção da l-ésima, tem-se

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \exp\left\{\beta_l(x_{il} - x_{jl})\right\},$$

que é interpretado como a razão de taxas de falhas.

Interpretação dos Parâmetros

- Por exemplo, suponha que x_l seja uma covariável dicotômica indicando pacientes hipertensos. A taxa de morte entre os hipertensos é $\exp(\beta_l)$ vezes a taxa daqueles com pressão normal, mantidas fixas as outras covariáveis.

- Uma interpretação similar é obtida para covariáveis contínuas. Se, por ex., o efeito de idade é significativo e $e^{\hat{\beta}} = 1,05$ para este termo, tem-se com o aumento de 1 ano na idade, que a taxa de morte fica aumentada em 5%.

- Estimativa para $\exp(\beta_l)$ é obtida utilizando a propriedade de invariância do estimador de máxima verossimilhança parcial. O intervalo de 95% de confiança é dado por: $\exp\left\{\hat{\beta} \pm 1,96 \times \widehat{EP}(\hat{\beta})\right\}$.

Funções Relacionadas a $\lambda_0(t)$

Funções Relacionadas a $\lambda_0(t)$

As funções relacionadas a $\lambda_0(t)$ referem-se basicamente a

$$\Lambda_0(t) = \int_0^t \lambda_0(u)du \quad \text{e} \quad S_0(t) = \exp\left\{-\Lambda_0(t)\right\}.$$

A função de sobrevivência

$$S(t_i|x_i) = [S_0(t_i)]^{\exp\left\{x_i'\beta\right\}}$$

também é importante para o cálculo de percentis associados a grupos de indivíduos.

Como $\lambda_0(t)$ não é especificado parametricamente, os estimadores para estas quantidades são de natureza não-paramétrica.

Funções Relacionadas a $\lambda_0(t)$

Um estimador simples proposto para a função de taxa de falha basal acumulada $\Lambda_0(t)$, proposto por Breslow (1972) é uma função escada, com saltos nos distintos tempos de falha, e é expresso por:

$$\widehat{\Lambda}_0(t) = \sum_{j:t_j<t} \frac{d_j}{\sum\limits_{l \in R_j} \exp\left\{x_l' \beta\right\}},$$

em que d_j é o número de falhas em t_j.

Consequentemente, as funções de sobrevivência $S_0(t|x)$ e $S(t|x)$ podem ser estimadas por, respectivamente,

$$\widehat{S}_0(t) = \exp\left\{-\widehat{\Lambda}_0(t)\right\} \quad \text{e} \quad \widehat{S}(t) = \left[\widehat{S}_0(t)\right]^{\exp\{x'\beta\}}.$$

Funções Relacionadas a $\lambda_0(t)$

Na ausência de covariáveis, temos

$$\widehat{\Lambda}_0(t) = \sum_{j:t_j<t} \left(\frac{d_j}{n_j} \right),$$

que é o estimador de Nelson-Aalen descrito anteriormente.

Por este fato, o estimador é também referenciado como estimador de Nelson-Aalen-Breslow.

Adequação do Modelo de Cox

Adequação do Modelo de Cox

- O modelo de Cox não se ajusta a qualquer situação e, como qualquer outro modelo estatístico, requer o uso de técnicas para avaliar a sua adequação.

- A violação da suposição básica de proporcionalidade das taxas de falhas pode acarretar em sérios vícios na estimação dos coeficientes do modelo (Struthers e Kalbfleisch, 1986).

- As técnicas de avaliação do modelo são baseadas em resíduos, como em outros modelos.

- Os resíduos de Schoenfeld (1982) são atualmente os mais utilizados para verificar a adequação do modelo de Cox, em especial, a suposição de proporcionalidade das taxas de falhas.

Resíduos de Schoenfeld (1982)

- Para o i-ésimo indivíduo, correspondente a um evento, com covariáveis $x_i = (x_{i1}, \dots, x_{ip})'$, o vetor de resíduos de Schoenfeld denotado por $r_i = (r_{i1}, \dots, r_{ip})'$ e definido para cada componente $r_{iq},\ q = 1, \dots, p$, por:

$$r_{iq} = x_{iq} - \frac{\sum_{j \in R(t_i)} x_{jq} \exp\left\{x_j' \hat{\beta}\right\}}{\sum_{j \in R(t_i)} \exp\left\{x_j' \hat{\beta}\right\}}.$$

- Os resíduos padronizados de Schoenfeld são dados por:

$$s_i^* = [\mathcal{I}(\hat{\beta})]^{-1} \times r_i,$$

em que $\mathcal{I}(\hat{\beta})$ é a matriz de informação observada.

Avaliação da Proporcionalidade das Taxas – Gráfico

- Grambsch e Therneau (1994) sugerem a utilização de s_i^* para avaliar a suposição de proporcionalidade dos riscos.

- Considere o modelo de Cox dinâmico:

$$\lambda(t) = \lambda_0(t) \exp\left\{x'\beta(t)\right\},$$

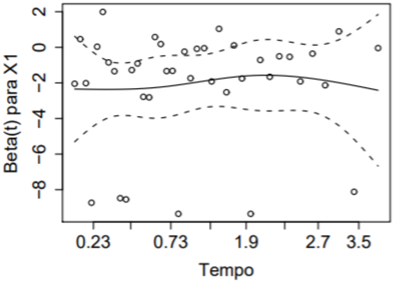
a restrição $\beta(t) = \beta$ corresponde à proporcionalidade das taxas.

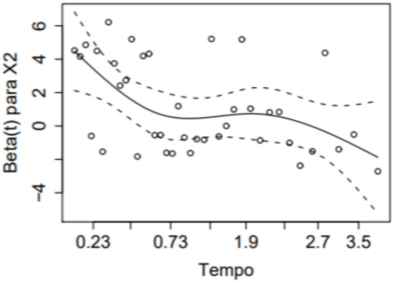
  - Se a suposição de proporcionalidade é válida, o gráfico de $\beta_q(t) \times t$ deve ser uma linha horizontal.
  - Sugestão: usar o gráfico de $(s_{iq}^* + \hat{\beta}) \times t$ ou alguma função do tempo, $g(t)$, $(q = 1, \ldots, p)$.

Avaliação da Proporcionalidade das Taxas – Gráfico

- Para auxiliar na detecção de uma possível falha da suposição de riscos proporcionais, uma curva suavizada (*splines*), com bandas de confiança, é adicionada a este gráfico.

- As figuras abaixo ilustram estes gráficos (primeira, adequada e a segunda, inadequada) para $g(t) = t$.





Avaliação da Proporcionalidade das Taxas - Testes

Testes de hipóteses associado aos resíduos de Schoenfeld:

- O coeficiente de correlação de Pearson (ρ) entre os resíduos padronizados de Schoenfeld e $g(t)$ para cada covariável é uma dessas medidas.

- Valores de ρ próximos de zero mostram evidências a favor da suposição de riscos proporcionais.

- Estão disponíveis testes globais e locais para avaliação da proporcionalidade dos riscos.

Avaliação da Proporcionalidade das Taxas - Testes

Um teste para a hipótese global de proporcionalidade de riscos no modelo de Cox, e assumindo que $g_q(t) = g(t)$, é obtido a partir da forma quadrática:

$$T = \frac{(g - \bar{g})' S^* \mathcal{I} S^{*'} (g - \bar{g})}{d \sum_k (g_k - g)^2} \sim \chi^2_p,$$

em que $\mathcal{I}$ é a matriz de informação observada, d é o número de falhas e $S^* = dR\mathcal{I}^{-1}$, sendo R a matriz $d \times q$ dos resíduos de Schoenfeld não padronizados.

Avaliação da Proporcionalidade das Taxas - Testes

Para testar a hipótese de taxas proporcionais para a q-ésima covariável $(q = 1, \ldots, p)$ utiliza-se a estatística de teste:

$$T_q = \frac{d\left(\sum_k (g_k - g)s^*_{qk}\right)^2}{\mathcal{I}_q^{-1}\sum_k (g_k - g)^2} \sim \chi^2_1,$$

em que $\mathcal{I}_q^{-1}$ é o q-ésimo elemento da diagonal do inverso da matriz de informação observada.

Algumas opções estão disponíveis $g(t)$ na função `cox.zph()`, sendo `km` (uma versão contínua à esquerda da curva de sobrevivência de Kaplan-Meier) a opção padrão.

Avaliação de Outros Aspectos do Modelo de Cox

Além da suposição de proporcionalidade, outros aspectos do modelo podem ser também avaliados.

Dentre eles, se destacam:

- Avaliação da melhor forma funcional de uma dada covariável.
- Presença de potenciais indivíduos atípicos (*outliers*).
- Influência que cada indivíduo exerce no modelo ajustado.

As técnicas de diagnóstico se baseiam, essencialmente, nos resíduos *martingal* e *deviance*, que foram definidos anteriormente para os modelos paramétricos.

Pontos Atípicos e Forma Funcional das Covariáveis

Os resíduos martingal, modificação dos resíduos de Cox-Snell, são frequentemente usados para tais finalidades.

Quando os dados apresentam censuras à direita e todas as covariáveis são fixadas no início do estudo, estes são definidos, para $i = 1, \ldots, n$, por:

$$\hat{m}_i = \delta_i - \widehat{\Lambda}_0(t_i) \exp\left\{\sum_{k=1}^{p} x_{ik}\hat{\beta}_k\right\} = \delta_i - \hat{e}_i, \quad i = 1, \ldots, n.$$

- Para, por exemplo, a covariável contínua x_q, o gráfico de $\hat{m}_i$ versus x_{iq} é utilizado para que se possa avaliar a forma funcional desta covariável.
- Na prática, as interpretações desses gráficos não são muito simples em razão da distribuição assimétrica desses resíduos.

Pontos Atípicos e Forma Funcional das Covariáveis

Outro resíduo geralmente usado para detecção de pontos atípicos (*outliers*) é o resíduo deviance, definido para o modelo de Cox por

$$\hat{d}_i = \text{sinal}(\hat{m}_i)\left[-2\left(\hat{m}_i + \delta_i \log(\delta_i - \hat{m}_i)\right)\right]^{1/2},$$

que não são tão assimétricos como os resíduos martingal.

O gráfico de $\hat{d}_i$ versus o preditor linear $\sum_{l=1}^{p} x_{il}\beta_l,\ i = 1, \ldots, n$ é utilizado, nesse caso, para avaliar a presença de dados atípicos.

Pontos Influentes

Uma medida global de efeito das observações pode ser obtida por:

$$\Delta\beta_i = (\hat{\beta} - \hat{\beta}_{(i)}) \left[\mathcal{I}(\hat{\beta})\right]^{-1} (\hat{\beta} - \hat{\beta}_{(i)}), \quad i = 1, \ldots, n,$$

com $\Delta\beta_i$ a mudança no vetor de coeficientes estimados obtida pela remoção, uma de cada vez, das observações.

O gráfico desses resíduos versus i pode ser útil na detecção de observações influentes.

Esta medida é usualmente chamada de D-Cook na literatura.

Pontos Influentes

Também é possível obter tais resíduos para cada covariável.

Para a q-ésima covariável ($q = 1, \ldots, p$) no modelo de Cox, tem-se, para $i = 1, \ldots, n$,

$$\Delta\beta_{i,q} = (\hat{\beta}_q - \hat{\beta}_{q(i)}) \left[\mathcal{I}(\hat{\beta})\right]^{-1}_{qq} (\hat{\beta}_q - \hat{\beta}_{q(i)}), \quad i = 1, \ldots, n,$$

A matriz D de dimensão $n \times p$ composta das p colunas dos resíduos é denominada de resíduos *dfbetas*.

Gráficos dos respectivos resíduos associados a cada covariável versus os valores da respectiva covariável são usados para a identificação de pontos influentes.

Exemplos de Aplicação

Exemplos de Aplicação

Aplicação 1: sobrevida de pacientes com leucemia aguda.

Aplicação 2: sobrevida de pacientes com câncer de encéfalo.

Aplicação 3: dados de aleitamento materno.

Aplicação 4: dados de câncer de laringe.

Extensões do Modelo de Cox

Extensões do Modelo de Cox

- Algumas situações práticas envolvem covariáveis que são monitoradas durante o estudo, e seus valores podem mudar ao longo desse período. Tais covariáveis são chamadas de dependentes do tempo e o modelo de Cox pode ser estendido para incorporá-las.

- Em outras situações, a suposição de taxas proporcionais é violada e o modelo de Cox não é adequado. Modelos alternativos existem para enfrentar esta situação. Um deles é uma extensão do próprio modelo de Cox chamado de modelo de riscos proporcionais estratificado.

Modelo de Cox com Covariáveis Dependentes do Tempo

Covariáveis Dependentes do Tempo

- Covariáveis que alteram seu valor ao longo do período podem ser incorporadas ao modelo de regressão de Cox:

$$\lambda(t|x(t)) = \lambda_0(t) \exp\left\{x'(t)\beta\right\}.$$

- Definido desta forma, este modelo não é mais de taxas de falhas proporcionais pois a razão das funções de taxa de falha no tempo t para dois indivíduos quaisquer i e j fica sendo

$$\frac{\lambda(t|x_i(t))}{\lambda(t|x_j(t))} = \exp\left\{x_i'(t)\beta - x_j'(t)\beta\right\}.$$

que é dependente do tempo.

Covariáveis Dependentes do Tempo

- A interpretação dos coeficientes β do modelo deve considerar o tempo t. Cada coeficiente $\beta_l, l = 1, \ldots, p$, pode ter a interpretação usual, mantendo as demais covariáveis fixas no mesmo tempo.

- O ajuste do modelo é obtido estendendo-se o logaritmo da função de verossimilhança parcial. Isto é feito usando-se:

$$U(\beta) = \sum_{i=1}^{n} \delta_i \left[x_i(t_i) - \frac{\sum_{j \in R(t_i)} x_j(t_i) \exp\left\{ x_j'(t_i)\hat{\beta} \right\}}{\sum_{j \in R(t_i)} \exp\left\{ x_j'(t_i)\hat{\beta} \right\}} \right] = 0.$$

- Os estimadores obtidos são consistentes e assintoticamente normais, sob certas condições de regularidade. As estatísticas de Wald e da razão de verossimilhança são usadas de forma usual.

Modelo de Cox Estratificado

Modelo de Cox Estratificado

- O modelo de Cox não pode ser usado se a suposição de RTP for violada. Nestes casos, uma solução é estratificar os dados de modo que a suposição seja válida em cada estrato.

- A análise estratificada consiste em dividir os dados de sobrevivência em m estratos, de acordo com uma indicação de violação da suposição.

- O modelo é então expresso como

$$\lambda(t|x_{ij}) = \lambda_{0_j}(t) \exp\left\{x_{ij}'\beta\right\},$$

para $j = 1, \ldots, m$ e $i = 1, \ldots, n_j$, sendo n_j o número de observações no j-ésimo estrato. As funções de base $\lambda_{0_1}, \ldots, \lambda_{0_m}$ são arbitrárias e completamente não relacionadas.

Inferência no Modelo de Cox Estratificado

- A estratificação não cria nenhuma complicação na estimação do vetor de parâmetros β.

- Uma função verossimilhança parcial é construída para cada estrato e a estimação é baseada na soma dos logaritmos das funções de verossimilhanças parciais:

$$\ell(\beta) = [\ell_1(\beta) + \ldots + \ell_m(\beta)] \,,$$

com $\ell_j(\beta) = \log(L_j(\beta))$ obtida usando somente os dados dos indivíduos no j-ésimo estrato, e $\ell(\beta)$ é maximizada com respeito a β.

- As propriedades assintóticas dos estimadores são obtidas a partir dos estimadores do modelo não estratificado (Colosimo, 1997).

Modelo de Cox Estratificado

- O modelo de Cox estratificado assume que as covariáveis atuam de modo similar na função de taxa de falha base de cada estrato. Ou seja, β é assumido ser comum para todos os estratos.

- Esta suposição pode ser testada usando, por exemplo, o teste da razão de verossimilhanças:

$$TRV = -2\left[\ell(\hat{\beta}) - \sum_{j=1}^{m} \ell_j(\hat{\beta}_j)\right],$$

sendo $\ell(\hat{\beta})$ o logaritmo da função de verossimilhança parcial sob o modelo que assume β's comuns e $\sum_{j=1}^{m} \ell_j(\hat{\beta}_j)$ o logaritmo dessa função sob o modelo que assume β's distintos em cada estrato.

Modelo de Cox Estratificado

- Sob a hipótese nula e para grandes amostras, a estatística segue distribuição qui-quadrado com $(m-1)p$ graus de liberdade em que m é o número de estratos e p a dimensão do vetor β.

- O modelo estratificado deve somente ser usado caso realmente necessário, ou seja, na presença de violação da suposição de taxas de falha proporcionais.

- O uso desnecessário da estratificação acarreta em uma perda de eficiência das estimativas obtidas.

Exemplos de Aplicação

Exemplos de Aplicação

Aplicação 5: sinusite em pacientes HIV.

Aplicação 6: hormônio de crescimento.

Outras Extensões

Outras Extensões

E se nada der certo?

- A suposição de taxas de falhas proporcionais não é satisfeita.
- Não podemos utilizar o modelo estratificado.
- Os modelos paramétricos não se ajustam aos dados.

Uma solução: O modelo aditivo (não-paramétrico) de Aalen, discutido no capítulo 7 do Livro ASA.

Como lidar com respostas altamente discretas?

- Poucos empates: aproximações para a verossimilhança parcial (Breslow e Efron).
- Muitos empates: aproximações ficam ruins. Temos que reconhecer a natureza discreta das respostas.

Uma solução: Modelos Discretos e Censura Intervalar (Capítulo 8).

Outras Extensões

Como lidar com respostas dependentes: mesmo indivíduo ou conglomerado?

- Modelos de efeitos aleatórios (Modelo de Fragilidade).
- Modelos do tipo GEE. Corrigindo a variância dos estimadores. Capítulo 9 do Livro ASA.

Como lidar com respostas dependentes: eventos recorrentes?

- Processo de Contagem (Processo de Poisson).
- O objetivo do estudo deve nortear a modelagem estatística.
  - Identificação de marcadores de risco: usar "gap times": modelos de fragilidade e GEE;
  - Caracterização de sistemas reparáveis: tempo global: modelagem de processos de contagem (lei de potência, mais utilizado).

Diferentes Desenhos e Objetivos

Desenho do Estudo

- Assume-se que a amostra foi coletada em um esquema de "amostragem aleatória simples";
- O estudo longitudinal pode ser observado uma unica vez: "Current status data".

Objetivo do Estudo

- Usualmente o objetivo é identificar marcadores ou comparar grupos;
- Predições.
  - Predição Clínica;
  - Caracterização de produtos industriais: Testes de Degradação.

Outras Extensões

Conceitos Importantes em Análise de Sobrevivência

- Indivíduo "sob risco": assume-se que ao ser censurado a observação estava sob risco do evento.
- Censura não informativa: T e C são independentes.

Referências

Referências

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

Colosimo, E. A., *Análise de Sobrevivência Aplicada*. Notas de Aula. Disponível em https://www.est.ufmg.br/~enricoc/metodos_estatisticos_analise_sobrevivencia.htm.