MAE5783 - Análise de Sobrevivência
1º semestre/2016

Professora:

Gisela Tunes da Silva
tunes@ime.usp.br gitunes@gmail.com
Sala 298 - Bloco A
Telefone: 3091-6188

Horário:

Terça-feira, 8:00 - 10:00
Quinta-feira, 10:00 - 12:00

Programa

1. Introdução e conceitos básicos:
   - Função de taxa de falha, sobrevivência, equivalências;
   - Censuras e truncagem; tipos de censura.
2. Conceitos básicos de processos estocásticos de contagem sob o enfoque de análise de sobrevivência:
   - Filtragem, propriedade martingal, etc;
   - Resultados utilizados no estudo de propriedades de estimadores e estatísticas de teste.
3. Modelos paramétricos e estimação de máxima verossimilhança para amostras censuradas:
   - Desenvolvimento de propriedades assintóticas para o caso de uma amostra;
   - Estimação paramétrica da função de sobrevivência e outras quantidades de interesse.
4. Estimação não-paramétrica da função de sobrevivência e da função de risco acumulada:
   - Estimador de Kaplan-Meier e suas propriedades assintóticas.
5. Testes não-paramétricos para uma ou mais amostras na presença de observações censuradas:
   - O teste de logrank (ponderado).
6. Modelos paramétricos de Regressão:
   - Modelos de tempos de vida acelerados;
   - Modelos de locação-escala;
   - Modelos paramétricos de riscos proporcionais.
7. Modelo semiparamétrico de Cox:
   - Estimação e testes envolvendo covariáveis;
   - Teoria assintótica;
   - Resíduos;
   - Modelo de Cox estendido.
8. Alguns modelos multivariados:
   - Modelo de Fragilidade;
   - Riscos competitivos.

Bibliografia

1. Aalen, O. O., Borgan, O. e Gjessing, H. (2008). **Survival and event history analysis: a process point of view**. New York: Springer. Processos Pontuais de Contagem
2. Andersen, P. K., Borgan, O., Gill, R. DS. e Keiding, N. (1993). **Statistical Models Based on Counting Processes**. New York: Springer.
3. Colosimo, E. A. e Giolo, S. R. (2006). **Análise de Sobrevivência Aplicada**. São Paulo: Edgard Blucher.
4. Cox, D. R. and Oakes, D. (1984). **Analysis of Survival Data**. New York: Chapman and Hall.
5. Fleming, T. R. e Harrington, D. P. (1991). **Counting Processes and Survival Analysis**. New York: John Wiley and Sons. Processos Pontuais de Contagem
6. Hosmer, D. W., Lemeshow, S. e May, S. (2008). **Applied Survival Analysis: regression modeling of time-to-event data**. New York: John Wiley and Sons.
7. Kalbfleisch, J. D. e Prentice, R. L. (1980). **The Statistical Analysis of Time Failure Data**. New York: John Wiley and Sons.
8. Klein, J.P. e Moeschberger, M. L. (2003). **Survival analysis: techniques for censored and truncated data**. New York: Springer, 2ª edição.
9. Kleinbaum, D. G. (1996). **Survival Analysis: a Self-Learning Text**. New York: Springer.
10. Lee, E. T. e Wang, J. W. (2003). **Statistical Methods for Survival Data Analysis**. New York: John Wiley and Sons.

Sistema de apoio

1. Acessar MAE5783

senha: MAE5783-Ideafix-2016

Será utilizado o Sistema PACA: paca.ime.usp.br.

AVISOS IMPORTANTES:

- Os alunos matriculados devem se cadastrar no sistema e se inscrever na disciplina;
- Os alunos devem manter o e-mail cadastrado no sistema atualizado;
- Informações e avisos importantes, listas de exercícios e material de apoio serão disponibilizados no sistema apenas.

Avaliação

A avaliação será baseada em:

- Listas de Exercícios;
- Seminário (obrigatório para alunos de doutorado);
- Duas provas.

Média Final:

- Média das listas de exercícios e seminário: ponderação 40%.
- Médias de provas: ponderação 60%.

Análise de Sobrevivência

Aula 01.

23/02/2016

Sugestão de data:

8 ou 10/03 - Definir data da P1

Prova 1 -

28/04 - Quinta-feira

26/04 - ?

19/04 - ?

Resumo Extendido Com Bibliografia duas semanas antes do seminário.

14/06 - Não haveria aula (terça)

16/06 - Prova 2 (Quinta)

21/06 (terça) - Entregar o 2º parte da P2.

23/06 - Quinta

24/06 - Sexta

MAE 5783 - Análise de Sobrevivência

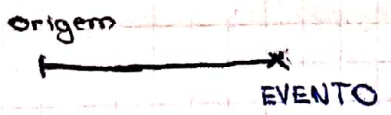
Lista 1 (pelo fim de semana)

" Área que se preocupa com fenômenos que podem ser caracterizados por tempo até ocorrência de um determinado evento"

* A diferença nesta vem do fato de observações incompletas "sobre o tempo de vida de algumas observações"

Antes da censura, interesse no

TEMPO ATÉ A OCORRÊNCIA DE UM EVENTO



o EVENTO deve estar muito bem definido.

o ORIGEM deve ser definida de forma precisa.

Antes de qqer coisa, devemos nos perguntar: Qual a ORIGEM e qual o EVENTO?

Exemplo: CEA. Dados de tratamento para mulheres viciadas em TABACO.

Queram estudar Tempo de Permanência no Tratamento.

Algumas mulheres

- terminaram o tratamento por terem alta;
- ainda em trat. após o término
- Abandono

O objetivo do estudo está muito relacionado com a definição do evento.

O evento 1 está bem definido.

Exemplo: Ocorrência de Anemia em doadores de sangue

Dependendo do prob., de 2 a 3 meses para definir a origem e o evento.
Por vezes, podemos estar ~~preocupado~~ imerso em riscos competitivos.

TERMINOLOGIA

. EVENTO DE INTERESSE: FALHA

Ex: Tempo de PEGA
Transp. medula óssea.

TEMPO DE FALHA: tempo até a ocorrência de um evento.

Observe que é v.a. não negativa: $T \geq 0$.

CARACTERIZAÇÃO DO TEMPO DE SOBREVIVÊNCIA

- Caso Contínuo

Seja T uma v.a. não-negativa, contínua, com função densidade de probabilidade

$$f(t), \quad t \geq 0.$$

A Função de Sobrevivência é definida por

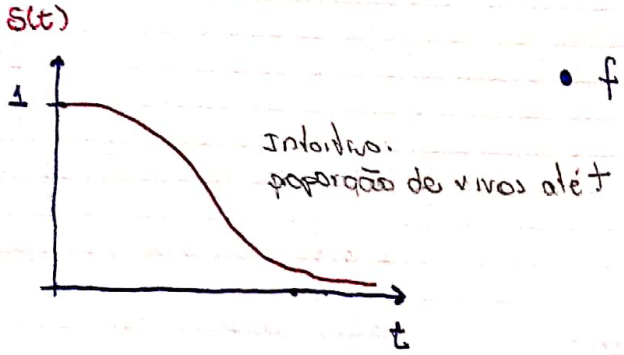
$$S(t) = \mathbb{P}(T > t) = \int_t^{\infty} f(u)\,du$$

Às vezes se define $\geq$ (Segundo Prof., maior parte def. $>$)

Note que

- $S(\cdot)$ é contínua e estritamente decrescente
- $S(0) = 1$
- $S(t) \to 0$ quando $t \to +\infty$
- $S(t) = 1 - F(t)$, com que $F(t)$ é função de distribuição acumulada.

$$F(t) = \mathbb{P}(T \leq t) = \int_0^t f(u)\,du$$



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

Resultado:

$$E(T) = \int_0^\infty S(t)\,dt$$

Barry James, pg 112.

Duas formas:

1. Escrever $S(t)$ como $\int_t^\infty f(w)\,dw$ e derivar resultado
2. Partir da def. de $E(T)$ e, via integração por partes, obter relação usando resultado do Barry James (pg. 112)

Função de Taxa de Falha ou Função de Risco

A função de taxa de falha é definida por

$$\alpha(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t+\Delta t \,/\, T \geq t)}{\Delta t}$$

observe que

- $\alpha(t) \geq 0 \quad \forall t$
- $\alpha(t)$ não é um probabilidade, é uma taxa.

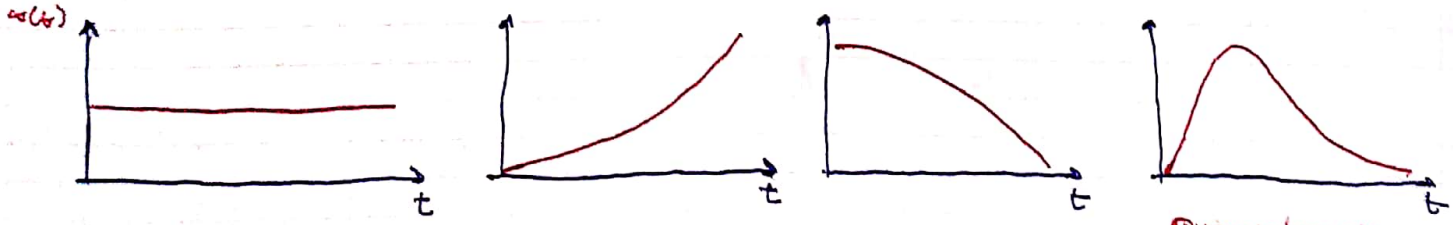
Interpretação intuitiva para $\alpha(t)$.

Para Δt pequeno

$$\alpha(t)\Delta t \approx P(\text{morrer em } (t, t+\Delta t) \,/\, \text{Vivo em } t^-)$$

: força de mortalidade instantânea

$\alpha(t) \rightarrow$ "força" de mortalidade



Quimioterapia (Pico de chance ...)

Uma idéia da forma de $\alpha(t)$ pode auxiliar na ~~construção~~ escolha do modelo paramétrico adequado.

Função de Risco Acumulada ou Função de Taxa de Falha Acumulada

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

(Função de Risco integrada)
outro termo

- Relações Importantes.

- $\alpha(t) = \dfrac{f(t)}{S(t)}$

- $S(t) = \exp\left\{-\int_0^t \alpha(s)\,ds\right\} = e^{-A(t)}$

- $f(t) = \alpha(t)\exp\left\{-\int_0^t \alpha(s)\,ds\right\} = \alpha(t)\,e^{-A(t)}$.

- A função de risco definem a dist. da v.a., de modo que posso modelar ou a f. de sobr., ou a acumulada, ou então a função de risco"

Análise de Sobrevivência

Aula 02

25/02/2016

Resumo Aula 01 • Apresentação da área e exemplos

• Caracterização do tempo de sobrevivência (T v.a. ñ negativa contínua)

• $S(t) = \mathbb{P}(T > t) = 1 - F(t)$

• $\alpha(t) = \lim_{\Delta t \to 0} \dfrac{\mathbb{P}(t \leq T \leq t + \Delta t \,/\, T \geq t)}{\Delta t}$

• $\alpha(t) = \dfrac{f(t)}{S(t)}$

— · — " —

Caso Discreto

Seja T uma v.a. não negativa assumindo os valores t_j, $j \geq 1$, com

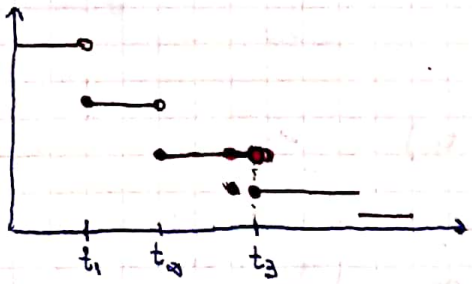
$$\mathbb{P}(T = t_j) = p_j, \quad j = 1, 2, 3, \ldots$$

Suponha $0 \leq t_1 < t_2 < t_3 < \ldots$, ou seja, considere os tempos ordenados.

A função de Sobrevivência é definida por

$$S(t) = \mathbb{P}(T > t) = \sum_{j : t_j > t} p_j$$

Observe que



• $S(t)$ é descontínua com saltos em t_j

• $S(t) \to 0$ com $t \to +\infty$

• $S(t)$ é não crescente

• $S(\cdot)$ é constante entre instantes de falhas sucessivos

Resultado. Se T é v.a. discreta assumindo valores inteiros não negativos, então

$$E(T) = \sum_{j=1}^{\infty} \mathbb{P}(T \geq j) = \sum_{j=0}^{\infty} \mathbb{P}(T > j)$$

Função De Taxa de Falha (Caso Discreto)

$$\alpha_j = \alpha(t_j) = \mathbb{P}(T = t_j \,/\, T \geq t_j)$$

$$= \frac{\mathbb{P}(T = t_j \cap T \geq t_j)}{\mathbb{P}(T \geq t_j)} = \mathbb{P}(T = t_j) \Big/ \mathbb{P}(T > t_{j-1}) = p_j \Big/ S(t_{j-1})$$

$\therefore$

$$\alpha_j = \frac{p_j}{S(t_{j-1})}, \quad j = 1, 2, 3, \ldots$$

Como $p_j = P(T = t_j) = S(t_{j-1}) - S(t_j)$, então

$$\alpha_j = \frac{S(t_{j-1}) - S(t_j)}{S(t_{j-1})} = 1 - \frac{S(t_j)}{S(t_{j-1})}. \qquad (*)$$

Observe ainda que, para t qualquer,

$$t_j \leq t < t_{j+1},$$

$$S(t) = S(t_j) = \frac{S(t_j)}{S(t_{j-1})} \frac{S(t_{j-1})}{S(t_{j-2})} \cdots \frac{S(t_1)}{S(t_0)},$$

com $S(t_0) = 1$. Daí, por $(*)$, segue que

$$S(t) = (1-\alpha_j)(1-\alpha_{j-1}) \cdots (1-\alpha_1)$$

$$= \prod_{j: t_j \leq t} (1-\alpha_j)$$

• Relação entre $S(\cdot)$ e $\alpha(\cdot)$ no caso discreto

A função de risco acumulada é definida por

$$A(t) = \sum_{t_j \leq t} \alpha(t_j)$$

Obs: Alguns autores preferem definir

$$A(t) = -\sum_{t_j \leq t} \ln(1-\alpha(t_j))$$

para que a relação

$$S(t) = e^{-A(t)}$$ seja válida.

— " — " —

Exemplos de Modelos Paramétricos

1. Exponencial $\quad f(t) = \lambda e^{-\lambda t}, \quad S(t) = e^{-\lambda t}, \quad \alpha(t) = \frac{f(t)}{S(t)} = \lambda$: constante

$$A(t) = \lambda t$$

2. Weibull $(p; \theta)$

- $f(t) = p\theta^p t^{p-1} e^{-(\theta t)^p}$, $p, \theta > 0$, $t \geq 0$.
- $S(t) = e^{-(\theta t)^p}$
- $\alpha(t) = p\theta^p t^{p-1}$
  - $p < 1$: decrescente
  - $p = 1$: constante
  - $p > 1$: crescente
- $A(t) = (\theta t)^p$

Observo que

$$\ln(A(t)) = p\ln\theta + p\ln t = p\ln\theta + pg, \quad g = \ln(t).$$

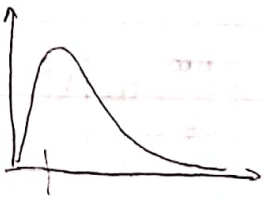
relação linear (útil p/ ajuste das dist.)

3. Distribuição Log-Logística

$$f(t) = \frac{p}{\lambda^p} t^{p-1} \left[1 + (t/\lambda)^p\right]^{-2}$$

$$S(t) = \frac{1}{(1 + (t/\lambda)^p)}$$

$$\alpha(t) = \frac{p\, t^{p-1}}{\lambda^p (1 + (t/\lambda)^p)}$$



Período de taxa de falha baixa

- Em geral, se não Weibull e Cox, próxima opção é Log-Logística.

Outros Conceitos.

- Vida Mediana e Tempo de Vida Média Residual

No caso contínuo, o quantil de ordem p é o valor t_p tal que

$$S(t_p) = 1 - p \quad e \quad F(t_p) = p.$$

O tempo mediano é $t_{0.5}$.

No caso discreto, o quantil p é dado pelo menor t tal que

$$S(t_p) \leq 1-p \qquad \text{(Klein, ...)}$$

$$t_p = \inf\{t : S(t) \leq (1-p)\}$$

Tempo de Vida médio residual

$$vmr(t) = E(T-t \mid T>t)$$ - quanto tempo médio de vida me resta, dado que sobrevive t.

Resultado.

$$vmr(t) = \frac{\int_t^\infty (u-t) f(u)\, du}{S(t)},$$

$$vmr(t) = \frac{\int_t^\infty S(u)\, du}{S(t)}$$

Exemplo: Exponencial

$$f(t) = \lambda e^{-\lambda t} \rightarrow E(T) = 1/\lambda$$

Observe que

$$f(u \mid T>t) = \frac{f(u)}{S(t)} = \frac{\lambda e^{-\lambda u}}{e^{-\lambda t}} = \lambda e^{-\lambda(u-t)}, \quad u>t$$

$$vmr(t) = \int_t^\infty (u-t) \frac{\lambda e^{-\lambda u}}{e^{\lambda t}}\, du = \frac{1}{\lambda}$$ falta de memória

Caso Discreto

$$vmr(t) = E[T-t \mid T>t]$$

$$vmr(t) = \sum_{t_j>t} \frac{(t_j-t)\, P(T=t_j)}{P(T>t)} = \frac{1}{S(t)} \sum_{t_j>t} (t_j-t)[S(t_{j-1}) - S(t_j)]$$

$$= \frac{1}{S(t)}\Big[(t_k-t) S(t_{k-1}) + (-1)(t_k-t) S(t_k) + (t_{k+1}-t) S(t_k) + \cancel{(-1)}$$
$$+ (-1)(t_{k+1}-t) S(t_{k+1}) + (t_{k+2}-t) S(t_{k+1})$$
$$+ (-1)(t_{k+2}-t) S(t_{k+2}) + (\quad \ldots \Big]$$

Continuação Aula 02 25/02/2016

com $t_{k-1} \leq t \leq t_k$,

$$vmr(t) = \frac{1}{S(t)} \left[(t_k - t)\, S(t_{k-1}) + \sum_{j \geq k} (t_{j+1} - t_j)\, S(t_j) \right]$$

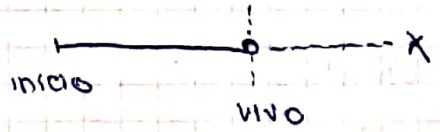
Observações Incompletas e Construção da Verossimilhança

- Tipos de Censura
  - à direita → {Tipo I, Tipo II, Aleatória}
  - à esquerda
  - intervalar

Ainda há a truncagem (Difícil ver se há). Quando algumas observações são desconsideradas por não poderem participar do estudo. (População com + 50 anos. Truncagem: muitos morrem antes dos 50. Eles deveriam ser inclusos?). Nesse caso, é preciso uma correção na f. veross.

Censura à Direita

Dizemos que uma observação é censurada à direita se sabe-se que o evento de interesse ocorreu após algum instante.

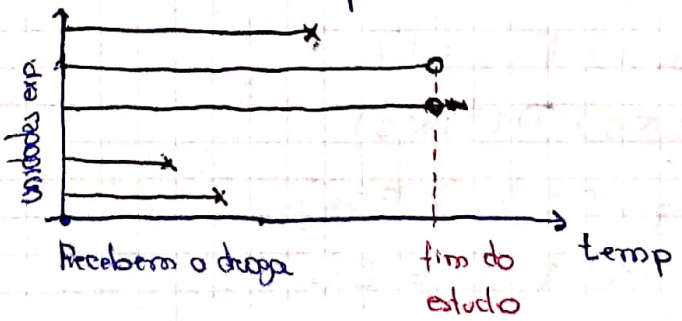


Censura Tipo I

O evento é observado somente se ocorre antes do final do estudo.

Exemplo: Experimento realizado com ratos no NCTR.

- Os ratos receberam determinadas doses de uma droga cancerígena.
- Os ratos são acompanhados até o óbito ou fim do estudo.



Notação: T é o tempo de falha; C é o tempo de censura

Assuma que T seja contínua com função dens. de prob. $f(t)$.

- Suponha n observações em estudo
- Suponha também $T_1, T_2, \dots, T_n$ v.a. ind. e ident. distribuídas.

Considere que os dados estão sujeitos do tipo I.

As quantidades observáveis são

$$Z_i = \min(T_i, c)$$

$$\delta_i = I(T_i \leq c) = \begin{cases} 1, & \text{se falha} \\ 0, & \text{se censura} \end{cases}$$

Temos portanto uma amostra aleatória $(Z_1, \delta_1), (Z_2, \delta_2), \dots, (Z_n, \delta_n)$.

Construção da Verossimilhança

A verossimilhança é construída com base na distribuição conjunta de (Z_i, δ_i).

Note que

$$f(z_i, \delta_i = k) = f(z_i / \delta_i = k) f(\delta_i = k), \quad k = 0, 1$$
$$= f(z_i / \delta_i = k) \mathbb{P}(\delta_i = k), \quad k = 0, 1$$

Para $k = 0$,

$$f(z_i, \delta_i = 0) = f(z_i / \delta_i = 0) \mathbb{P}(\delta_i = 0)$$
$$= f(z_i / T_i > c) \mathbb{P}(T_i > c)$$
$$= \mathbb{P}(\underbrace{Z_i = c}_{1} / T_i > c) \mathbb{P}(T_i > c)$$
$$= \mathbb{P}(T_i > c)$$
$$= S(c)$$

Para $k = 1$,

$$f(z_i, \delta_i = 1) = f(z_i / \delta_i = 1) \mathbb{P}(\delta_i = 1)$$
$$= f(z_i / T_i \leq c) \mathbb{P}(T_i \leq c)$$
$$= f_{T / T \leq c}(z_i / T_i \leq c) \mathbb{P}(T_i \leq c)$$
$$= \frac{f(z_i)}{\cancel{\mathbb{P}(T_i \leq c)}} \cancel{\mathbb{P}(T_i \leq c)} = f(z_i)$$

Podemos escrever

$$f(z_i, \delta_i = k_i) = \begin{cases} f(z_i), & \text{se } k_i = 1 \\ S(c), & \text{se } k_i = 0 \end{cases} \quad \text{ou}$$

$$f(z_i, \delta_i) = (f(z_i))^{\delta_i} (S(c))^{1-\delta_i}$$

A função de verossimilhança é dada por

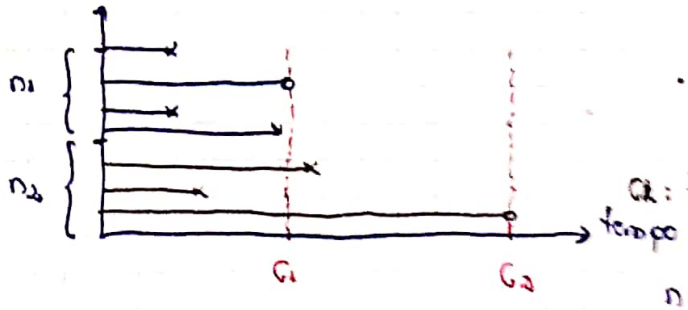
$$L(\theta) = \prod_{i=1}^{n} (f(z_i))^{\delta_i} (S(c))^{1-\delta_i},$$

$$L(\theta) = \prod_{i=1}^{n} (f(z_i))^{\delta_i} (S(z_i))^{1-\delta_i}$$

- Caso + simples: Censura à direita do tipo I.

→ Variações da censura do tipo I

- Censura Tipo I Progressiva
  - n unidades em estudo acompanhadas a partir de um mesmo instante
  - O estudo tem mais de um instante de censura



- Quando cara manter indivíduos

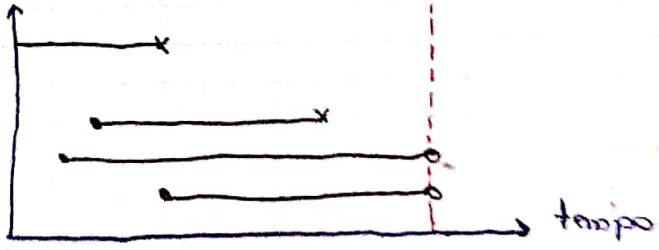
c_i: tempo de censura dos n_i indivíduos

$n = n_1 + n_2$

$$L(\theta) = \prod_{i=1}^{n_1} (f(z_i))^{\delta_i} (S(c_1))^{1-\delta_i}$$

$$\times \prod_{i=n_1+1}^{n_1+n_2} (f(z_i))^{\delta_i} (S(c_2))^{1-\delta_i}$$

- Censura do Tipo I Generalizada



O tempo de censura pode variar de unidade para unidade amostral, mas o tempo de censura não é aleatório. Basta transladar, original, determinado pelo fim de vida

Observamos

$$z_i = \min(T_i, c_i)$$

$$\delta_i = I(T_i \leq c_i)$$

c_i: constantes conhecidas

$$\rightarrow L(\theta) = \prod_{i=1}^{n} (f(z_i))^{\delta_i} (S(c_i))^{1-\delta_i}$$

Exemplo:

Análise de Sobrevivência

$S(t) = \mathbb{P}(T > t)$

$\alpha(t)$, $\mu = E(T)$, $vmr(t)$

Censuras
- à direita
  - Tipo I
  - Tipo II
  - Censura Aleatória
- à esquerda
- intervalar

Aula 03

01/03/2016

Censura Tipo I

$$L(\theta) = \prod_{i=1}^{n} f(z_i)^{\delta_i} S(c_i)^{1-\delta_i}$$

$$\begin{cases} Z_i = \min(T_i, C_i) \\ \delta_i = J(T_i \leq C_i) \end{cases}$$

$$L(\theta) = \prod_{i=1}^{n} f(z)^{\delta_i} S(z_i)^{1-\delta_i}$$

Exemplo. Sejam $T_1, \dots, T_n$ tempos de falhas independentes com

$$T_j \sim \exp(\theta)$$

Assuma que observamos

$$\begin{cases} Z_i = \min(T_i, C_i) \\ \delta_i = J(T_i \leq C_i) \end{cases}, \quad i = 1, \dots, n$$

Queremos encontrar o E.M.V. de θ.

$$L(\theta) = \prod_{i=1}^{n} f(z_i)^{\delta_i} S(z_i)^{1-\delta_i}$$

$$= \prod_{i=1}^{n} (\theta e^{-\theta z_i})^{\delta_i} e^{-\theta z_i (1-\delta_i)}$$

$$= \theta^{\sum_{i=1}^{n} \delta_i} e^{-\theta \sum_{i=1}^{n} z_i}$$

$$\ell(\theta) = \left(\sum_{i=1}^{n} \delta_i\right) \ln \theta - \theta \sum_{i=1}^{n} z_i$$

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{\sum \delta_i}{\theta} - \sum z_i = 0 \Leftrightarrow \quad \hat{\theta} = \sum_{i=1}^{n} \delta_i \Big/ \sum_{i=1}^{n} z_i$$

$$\hat{\theta} = \frac{r}{\sum_{i=1}^{r} z_i}$$, r: nº de falhas (correção em rel. ao usual $\hat{\theta} = 1/\bar{T}$)

(Nem sempre se consegue a I. Fisher)
(Acaba-se usando Inf. Observada)

Temos ainda que

$$\mu = E(T) = 1/\theta \Rightarrow \hat{\mu} = 1/\hat{\theta} = \frac{\sum_{i=1}^{r} z_i}{r}$$

Para t_0 fixado,

$$S(t_0) = e^{-\theta t_0}$$

Assim,

$$\widehat{S(t_0)} = e^{-\hat{\theta} t_0}.$$

Estimador para a mediana

$$S(m_d) = 1/2$$

$$e^{-\theta m_d} = 1/2 \Rightarrow \theta m_d = \log 2 \Rightarrow m_d = \log 2/\theta \Rightarrow$$

$$\hat{m}_d = \log 2 / \hat{\theta}$$

Observação:

$$L(\theta) = \prod_{i=1}^{n} f(z_i)^{\delta_i} (S(z_i))^{1-\delta_i}$$

$$= \prod_{i=1}^{n} (\alpha(z_i) S(z_i))^{\delta_i} (S(z_i))^{1-\delta_i}$$, uma vez que $f(t) = \alpha(t) S(t)$

$$L(\theta) = \prod_{i=1}^{n} \alpha(z_i)^{\delta_i} S(z_i)$$

Censura do Tipo II

O experimento é conduzido até que um número pré-determinado de falhas ocorra.

Exemplo: 20 máquinas são colocadas em teste em um mesmo instante
- O experimento termina quando a 15ª máquina falhar.

20 máquinas → 15 falhas
→ 15 censuras

Tempo de censura é igual ao tempo de falha da 15ª máquina a falhar.

Genericamente,

- n unidades experimentais;
- r falhas
- $(n-r)$ censuras

Sejam $T_1, T_2, \dots, T_n$ os tempos de falhas associados a cada máquina. Assuma $T_1, \dots, T_n$ v.a. i.i.d. com

$$T_j \sim \begin{matrix} f(\cdot) \\ S(\cdot) \end{matrix} \quad , j = 1, 2, \dots, n$$

Obtenção da Função de Verossimilhança

Sejam $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(r)}$ as estatísticas de ordem

No experimento, observamos de fato as r primeiras estatísticas de ordem.

- Suponha inicialmente $r = 1$:

$$T_{(1)} = \min(T_1, T_2, \dots, T_n)$$

Vamos obter a densidade de $T_{(1)}$:

$$f_{T_{(1)}}(t) = f_{(1)}(t) = \lim_{\Delta \to 0} \frac{F_{T_{(1)}}(t+\Delta) - F_{T_{(1)}}(t)}{\Delta}$$

$$= \lim_{\Delta \to 0} \frac{P(t \leq T_{(1)} \leq t+\Delta)}{\Delta}$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbb{P}\left(\begin{matrix} 1 \text{ dos } T_i\text{'s} \in (t; t+\Delta] \text{ e} \\ \text{o restante do } T_i\text{'s} > t \end{matrix} \right)$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta} \begin{bmatrix} \mathbb{P}(T_1 \in (t, t+\Delta]) \mathbb{P}(T_2 > t) \dots \mathbb{P}(T_n > t+\Delta) + \\ \mathbb{P}(T_2 \in (t, t+\Delta]) \mathbb{P}(T_1 > t) \dots \mathbb{P}(T_n > t+\Delta) + \\ \mathbb{P}(T_3 \in (t, t+\Delta]) \mathbb{P}(T_1 > t) \mathbb{P}(T_2 > t) \dots \mathbb{P}(T_n > t+\Delta) \\ \vdots \\ \mathbb{P}(T_n \in (t, t+\Delta]) \mathbb{P}(T_1 > t) \dots \mathbb{P}(T_{n-1} > t+\Delta) \end{bmatrix}$$

$$= n \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbb{P}(T_1 \in (t, t+\Delta]) (S(t+\Delta))^{n-1}$$

$$f_{T_{(1)}}(t) = n f(t) (S(t))^{n-1}. \quad //$$

. Suponha $r=2$

Para $t_1 < t_2$,

$$f_{T_{(1)},T_{(2)}}(t_1,t_2) = \lim_{\substack{\Delta_1 \to 0 \\ \Delta_2 \to 0}} \frac{1}{\Delta_1 \Delta_2} \mathbb{P}\left(\begin{array}{l} 1 \text{ dos } T_i\text{'s} \in (t_1; t_1+\Delta_1] \\ 1 \text{ dos } T_i\text{'s} \in (t_2; t_2+\Delta_2] \\ \text{o restante dos } T_i\text{'s} > t_2+\Delta_2 \end{array} \right)$$

$$= \lim_{\substack{\Delta_1 \to 0 \\ \Delta_2 \to 0}} \frac{1}{\Delta_1 \Delta_2} \left[\begin{array}{l} n(n-1)\, \mathbb{P}(T_1 \in (t_1, t_1+\Delta_1]) \times \mathbb{P}(T_2 \in (t_2, t_2+\Delta_2]) \\ \times \mathbb{P}(T_3 > t_2+\Delta_2) \times \dots \times \mathbb{P}(T_n > t_2+\Delta_2). \end{array} \right]$$

$$f_{T_{(1)},T_{(2)}}(t_1,t_2) = n(n-1) \lim_{\substack{\Delta_1 \to 0 \\ \Delta_2 \to 0}} \left[\frac{\mathbb{P}(T_1 \in (t_1, t_1+\Delta_1])}{\Delta_1} \times \frac{\mathbb{P}(T_2 \in (t_2, t_2+\Delta_2])}{\Delta_2} \left(S(t_2+\Delta_2)\right)^{n-2} \right]$$

$$= n(n-1)\, f(t_1)\, f(t_2)\, \left(S(t_2)\right)^{n-2}.$$

. Para o caso geral, temos

$$f_{T_{(1)},T_{(2)},\dots,T_{(r)}}(t_1,t_2,\dots,t_r) = \frac{n!}{(n-r)!} \left(\prod_{j=1}^{r} f(t_j) \right) \left(S(t_r) \right)^{n-r}$$

A função de verossimilhança é dada por

$$L(\theta) = \frac{n!}{(n-r)!} \left(\prod_{j=1}^{r} f(t_j) \right) \left(S(t_r) \right)^{n-r} \propto \underbrace{\left(\prod_{i=1}^{r} f(t_j) \right)}_{\text{observados}} \underbrace{\left(S(t_r) \right)^{n-r}}_{\text{censurados}}$$

Censura do Tipo II Progressiva

Esquema: n unidades amostrais em teste.

- Após r_1 falhas, tem-se $(n-r_1)$ unidades em observação.
- Dessas $(n-r_1)$, retiro n_1 do estudo. (Censuro n_1)
- Restam $(n-r_1-n_1)$ unidades em observação
- Acompanho as unidades até r_2 falhas acontecerem $\therefore$ n. falhas r_1+r_2

Aula 08.

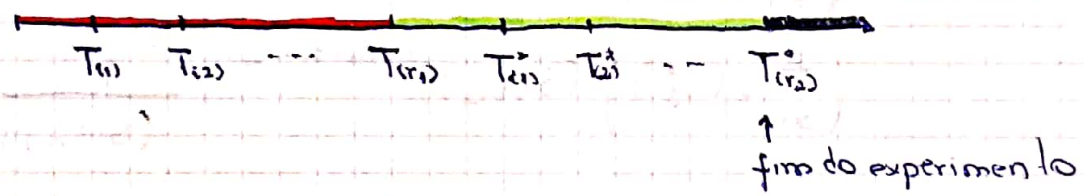
<u>Obtenção da função de verossimilhança</u>

Temos $T_1, \dots, T_n$ tempos de falhas independentes com

$$T_j \sim f(\cdot)$$
$$S(\cdot)$$

Sejam: $T_{(1)}, T_{(2)}, \dots, T_{(r_1)}$ - tempos de falha na 1ª fase do experimento

- $T^*_{(1)}, T^*_{(2)}, \dots, T^*_{(r_2)}$: tempos de falha ordenados na 2ª fase do experimento



A função de verossimilhança é baseada na distribuição conjunta

$$h(t_1, t_2, \dots, t_{r_1}, t_1^*, t_2^*, \dots, t_{r_2}^*) = h_1(t_1, \dots, t_{r_1})\, h_2(t_1^*, \dots, t_{r_2}^* \mid t_1, \dots, t_{r_1})$$

Pelo visto anteriormente,

$$h_1(t_1, \dots, t_{r_1}) = \frac{n!}{(n-r_1)!}\left[\prod_{j=1}^{r_1} f(t_j)\right](S(t_{r_1}))^{n-r_1}$$

Suponho inicialmente $r_2 = 1$. Note que $t_1^* > t_{r_1}$.

$$h_2(t_1^* \mid t_1, t_2, \dots, t_{r_1}) = \lim_{\Delta \to 0} P(T^*_{(1)} \in (t_1^*, t_1^* + \Delta] \mid T^*_{(1)} > t_{r_1}) / \Delta$$

$$= \lim_{\Delta \to 0} \frac{1}{\Delta}\left[\begin{array}{l} P(T_1^* \in (t_1^*, t_1^* + \Delta] \mid T_1^* > t_{r_1}) \\ \times P(T_2^* > t_1^* + \Delta \mid T_2^* > t_{r_1}) \\ \vdots \\ \times P(T^*_{n-r_1-n_1} > t_1^* + \Delta \mid T^*_{n-r_1-n_1} > t_{r_1}) \end{array}\right] \times (n - r_1 - n_1)$$

$$= (n - r_1 - n_1) \lim_{\Delta \to 0} \frac{P(T_1^* \in (t_1^*, t_1^* + \Delta])}{\Delta S(t_{r_1})} \cdot \left(\frac{S(t_1^* + \Delta)}{S(t_{r_1})}\right)^{n - r_1 - n_1 - 1}$$

$\therefore$

$$h_2(t_i^* / t_1, \dots, t_{r_1}) = \frac{f(t_i^*)}{S(t_{r_1})} \cdot \left(\frac{S(t_i^*)}{S(t_{r_1})}\right)^{n-r_1-n_1-1}$$

Para um r_2 genérico,

$$h_2(t_1^*, \dots, t_{r_2}^* / t_1, \dots, t_{r_2}) = \frac{(n-n_1-r_1)!}{(n-n_1-r_1-r_2)!} \left(\prod_{j=1}^{r_2} f(t_j^*)\right) (S(t_{r_2}))^{n-n_1-r_1-r_2}$$

$$\times \frac{1}{S(t_{r_1})^{n-n_1-r_1}}$$

Portanto,

$$h(t_1, \dots, t_{r_1}, t_1^*, \dots, t_{r_2}^*) \propto \left(\prod_{j=1}^{r_1} f(t_j)\right) (S(t_{r_1}))^{n-r_1}$$

$$\left(\prod_{j=1}^{r_2} f(t_j^*)\right) (S(t_{r_2}))^{n-n_1-r_1-r_2} \cdot \frac{1}{(S(t_{r_1}))^{n-n_1-r_1}}$$

Assim,

$$L(\theta) \propto \left(\prod_{j=1}^{r_1} f(t_j)\right) (S(t_{r_1}))^{n_1} \left(\prod_{j=1}^{r_2} f(t_j^*)\right) (S(t_{r_2}))^{n-n_1-r_1-r_2}$$

. Censura Aleatória (Caso + comum em caso em que se consideram pacientes, 90% dos casos com poc acontecem isso)

As unidades amostrais estão sujeitas a outros eventos aleatórios que não são de ~~interesse~~ interesse e que podem inibir a observação do evento de interesse.

Exemplo: Sobrevida de pacientes com doença cardíaca grave.

Início: data de entrada no Protocolo

Evento: óbito

→ 350 pacientes selecionados

→ Perda de acompanhamento devido a

- mudança de cidade
- mudança de plano de saúde
- etc. { abandono do Tratamento

, Cuidado: Às vezes, a causa pode estar rel. com a falta. O paciente já se sente melhor ou acha que ñ está surtindo efeito

Temos.

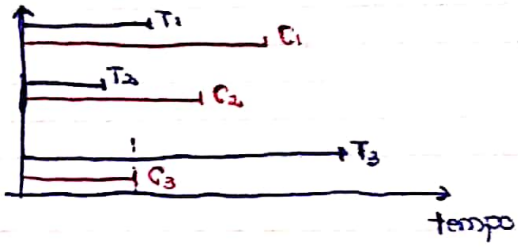
T_i : tempo de óbito

C_i : tempo de censura

Assumindo que

C_i é variável aleatória

observamos



$$Z_i = \min(T_i, C_i)$$

$$\delta_i = I(T_i \leq C_i)$$

Assumimos que

$T_i \sim f(.)$; $C_i \sim g(.)$

$\delta(.)$ $\qquad$ $G(c) = \mathbb{P}(C \geq c)$

Análise de Sobrevivência

Censura

- à direita
  - Tipo I $\rightarrow \prod_{i=1}^{n} f(z_i)^{\delta_i} S(z_i)^{1-\delta_i}$
  - Tipo II $\rightarrow \propto \left(\prod_{j=1}^{r} f(t_j)\right)\left(S(t_r)\right)^{n-r}$

$$\propto \prod_{i=1}^{n} f(z_i)^{\delta_i} S(z_i)^{1-\delta_i}$$

Aula 04 03/03/16

<u>Censura Aleatória</u>

Temos

T_i : tempo de falha

C_i : tempo de censura

Assuma que

$T_i \sim f(\cdot)$; $C_i \sim g(\cdot)$ e

$S(\cdot)$ $G(\cdot) =$ func sob. da cens.

Observamos

$$Z_i = \min(T_i, C_i)$$

$$\delta_i = I(T_i \leq C_i)$$

↳ Construção da Função de Verossimilhança

Suposições adicionais

(1) Independência : as variáveis T_i e C_i são independentes, para todo i

(2) Não-informatividade : A distribuição de C_i não envolve parâmetros da densidade de T_i

(Klein, Kolb fleisch, Enrico)

Precisamos obter a distribuição conjunta de (Z_i, δ_i)

- Para $\delta = 1$

$$P(z_i, \delta_i = 1) \overset{\lim(?)}{=} P(Z_i \in (z_i, z_i + \Delta]; \delta_i = 1)$$

$$= P(Z_i \in (z_i, z_i + \Delta], T_i \leq C_i)$$

$$= P(T_i \in (z_i, z_i + \Delta], C_i > z_i + \Delta)$$

$$\overset{ind}{=} P(T_i \in (z_i, z_i + \Delta]) P(C_i > z_i + \Delta) = f(z_i) \cdot G(z_i)$$

Para $\delta_i = 0$.

$$P(z_i, \delta_i = 0) = P(Z_i \in (z_i, z_i + \Delta], C_i < T_i)$$
$$= P(C_i \in (z_i, z_i + \Delta], T_i > z_i + \Delta)$$
$$\overset{ind}{=} P(C_i \in (z_i, z_i + \Delta]) \times P(T_i > z_i + \Delta)$$
$$= g(z_i)\, S(z_i)$$

A função de Verossimilhança é dada por

$$L(\theta) = \prod_{i=1}^{n} \left(f(z_i) G(z_i) \right)^{\delta_i} \left(g(z_i) S(z_i) \right)^{1-\delta_i}$$

$$= \left(\prod_{i=1}^{n} f(z_i)^{\delta_i} S(z_i)^{1-\delta_i} \right) \left(\prod_{i=1}^{n} \underbrace{g(z_i)^{1-\delta_i} G(z_i)^{\delta_i}} \right)$$

A dist. da Cens. não envolve parâmetro da dist. de T

$$\rightarrow \quad L(\theta) \propto \prod_{i=1}^{n} f(z_i)^{\delta_i} S(z_i)^{1-\delta_i}.$$

Suposi. Censura Ñ Informativa

Já há uma versão do E.i KM para dados com Cens Inf.

- Tempo de sob. ajustado pela qualidade de vida
Com base num quest., encontra-se um escore $0 \longleftrightarrow 1$ (péssima, qual. perf.)
Após isso, pondera-se o tempo por tal escore.

Exemplo. Log-normal.

$T \sim$ log-normal se $X = \ln T \sim N(\mu, \sigma^2)$

Note que

$$E(T) = E(e^X) = e^{\mu + \sigma^2/2}$$
$$Var(T) = (e^{\sigma^2} - 1)\, e^{2\mu + \sigma^2}$$
$$f_T(t) = \frac{1}{\sqrt{2\pi\sigma^2}\, t} \exp\left\{ -\frac{1}{2} \left(\frac{\ln t - \mu}{\sigma} \right)^2 \right\}, \quad t > 0.$$

$$S_T(t) = \int_t^{\infty} f(u)\, du = P(T > t) = P(e^X > t) = P(X > \ln t)$$
$$= 1 - \Phi\left(\frac{\ln t - \mu}{\sigma} \right)$$

$\Phi(\cdot)$: F.D. acum. da normal padrão

Função de Verossimilhança

$$L(\mu,\sigma^2) \propto \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}\, z_i} \exp\left\{ -\frac{1}{2} \left(\frac{\ln z_i - \mu}{\sigma} \right)^2 \right\} \right)^{\delta_i}$$

$$\times \left[1 - \Phi\left(\frac{\ln z_i - \mu}{\sigma} \right) \right]^{1-\delta_i}$$

$$\propto \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^{\sum \delta_i} \left[\prod_{i=1}^{n} \left(\frac{1}{z_i} \right)^{\delta_i} \right] \exp\left\{ -\frac{1}{2\sigma^2} \sum \delta_i (\ln z_i - \mu)^2 \right\} \times$$

$$\times \prod_{i=1}^{n} \left(1 - \Phi\left(\frac{\ln z_i - \mu}{\sigma} \right) \right)^{1-\delta_i}$$

$$\ell(\mu,\sigma^2) = -\frac{\sum \delta_i}{2} \log(2\pi\sigma^2) - \sum_{i=1}^{n} \delta_i \log z_i$$

$$- \frac{1}{2\sigma^2} \sum_{i=1}^{n} \delta_i (\ln z_i - \mu)^2$$

$$+ \sum_{i=1}^{n} (1-\delta_i) \log\left[1 - \Phi\left(\frac{\ln z_i - \mu}{\sigma} \right) \right]$$

. Caso não houvesse censura, estimação seria trivial. Com censura, se torna infactível.

→ Métodos numéricos p/ estimação de μ e σ^2.

Observação. Se T e C não são independentes, a ver. construída não é correta.

Podemos escrever

efeito aleatório

$$L(\theta) \propto \prod_{i=1}^{n} \left[\left(-\frac{\partial S(t, z_i)}{\partial t} \right)_i \Big|_{t=z_i} \right]^{\delta_i} \left[\left(-\frac{\partial S(z, c)}{\partial c} \right) \Big|_{c=z_i} \right]^{1-\delta_i}$$

. Assume ind. entre $\{(T_i, C_i)\}_{i=1}^{n}$ ind. $T_i \perp\!\!\!\perp C_i$.

em que

$S(t,c)$ denota a função de sob. cons. de T e C.

. Censura à Esquerda.

Um tempo de sobrevivência T é considerado censurado à esquerda se é menor do que um tempo de censura C_L, porém não se observa o instante exato.

Exemplo:

(1) Estudo conduzido em uma escola p/ avaliar a idade do 1º uso de maconha.

(2) Estudo com bebês (CEA)

- Selecionaram mães de bebês (prematu. e no tempo)
  - O bebê ja senta
  - " " " come sozinho
  - - - - - -

Deseja-se verificar a idade (em semanas) que bebês conseguem fazer certa atividade.

Alguns bebês já conseguiam a atividade quando o estudo foi iniciado.

- Eventos repetidos recorrentes.

Sejam

T_i: tempo de sobrevida

C_L: tempo de censura

Em geral, observamos

$$Z = \max(T_i, C_L)$$

$$\varepsilon = I(T \geq C_L) = \begin{cases} 1, & \text{se } T \geq C_L \quad \text{(falha observada)} \\ 0, & \text{se } T < C_L \quad \text{(censura)} \end{cases}$$

Construção da Função de Verossimilhança

Suponha

- $T_i \sim f(\cdot), S(\cdot)$
- $C_{L_i} \sim g, G$
- $T_i \perp C_{L_i}$ (ind.)
- Censura ñ informativa

Precisamos obter a distribuição conjunta de (Z_i, ε_i).

Para $\varepsilon_i = 1$,

$$\begin{aligned} P(z_i; \varepsilon_i = 1) &= P(Z_i \in (z_i, z_i + \Delta], T_i \geq C_{L_i}) \\ &= P(T_i \in (z_i, z_i + \Delta], C_{L_i} < z_i) \\ &= P(T_i \in (z_i, z_i + \Delta]) \, P(C_{L_i} < z_i) \\ &= f(z_i)(1 - G(z_i)) \end{aligned}$$

- Para $\varepsilon_i = 0$, $P(z_i, \varepsilon_i = 0) = P(Z_i \in (z_i, z_i + \Delta], \varepsilon_i = 0)$

$= P(Z_i \in (z_i, z_i+\Delta], T_i < C_{L_i}) = P(C_{L_i} \in (z_i, z_i+\Delta], T_i < z_i)$

$= P(C_{L_i} \in (z_i, z_i+\Delta]) P(T_i < z_i)$

$= g(z_i)(1 - S(z_i))$

A função de verossimilhança é dada por

$$L(\theta) = \prod_{i=1}^{n} \left(f(z_i)(1-G(z_i)) \right)^{\varepsilon_i} \times \left(g(z_i)(1-S(z_i)) \right)^{1-\varepsilon_i}$$

$$L(\theta) \propto \prod_{i=1}^{n} f(z_i)^{\varepsilon_i} (1-S(z_i))^{1-\varepsilon_i}.$$

Exemplo: $T_1, \dots, T_n$ v.a. i.i.d. com dist. exponencial(θ).
Assuma dados sujeitos a cens. à esquerda.

$$L(\theta) \propto \prod_{i=1}^{n} (f(z_i))^{\varepsilon_i} (1-S(z_i))^{1-\varepsilon_i}$$

$$\propto \prod_{i=1}^{n} (\theta e^{-\theta z_i})^{\varepsilon_i} (1-e^{-\theta z_i})^{1-\varepsilon_i}$$

$$\propto \theta^{\sum \varepsilon_i} e^{-\theta \sum \varepsilon_i z_i} \prod_{i=1}^{n} (1-e^{-\theta z_i})^{1-\varepsilon_i}$$

$$\ell(\theta) = (\sum \varepsilon_i) \log\theta - \theta \sum \varepsilon_i z_i + \sum_{i=1}^{n} (1-\varepsilon_i) \log(1-e^{-\theta z_i})$$

$$\frac{d}{d\theta} \ell(\theta) = \frac{\sum \varepsilon_i}{\theta} - \sum \varepsilon_i z_i + \sum_{i=1}^{n} \frac{(1-\varepsilon_i)}{1-e^{-\theta z_i}} e^{-\theta z_i} z_i$$

$\searrow$ Métodos Numéricos.

Observação: Podemos ter dados "duplamente" censurados. Podemos representar

$$Z = \max(\min(T, C_R), C_L)$$

$$\delta = \begin{cases} 1, & \text{se } Z = T \\ 0, & \text{se } Z = C_R \\ -1, & \text{se } Z = C_L \end{cases}$$

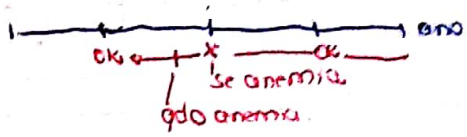
Censura Intervalar

Ocorre quando sabemos que o evento ocorreu em determinado intervalo, mas não se sabe o momento exato.

Exemplo: Dos bebês,

- Dados de anemia (Cjto dados real - Tempo até ocorr. de anemia.

Pessoas doando sangue



Sejam $(a_i, b_i] = I_i$ o intervalo de censura, $i = 1, 2, \dots, n$.

a_i: último instante sem evento

b_i: "primeiro instante com evento". (o pac. já apres. o evento)

Sabemos que $T \in (a_i, b_i]$.

A informação que temos é que

$$\mathbb{P}(T \in (a_i, b_i]) = S(a_i) - S(b_i)$$
$$= F(b_i) - F(a_i)$$

Observe que:

$b_i = +\infty \rightarrow I_i = (a_i, +\infty) \Rightarrow$ censura à direita (acont. depois de a_i)

$a_i = 0 \rightarrow I_i = (0; b_i] \rightarrow$ censura à esquerda (acont. antes de b_i)

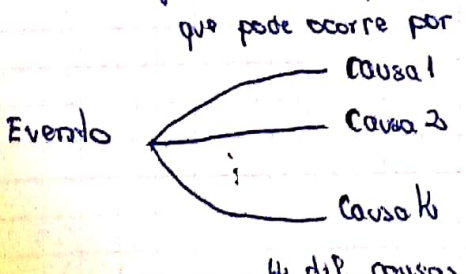
No caso de haver apenas observações com censura intervalar, a função de verossimilhança será

$$L(\theta) = \prod_{i=1}^{n} (S(a_i) - S(b_i))$$

(Est. de Turnble)

Modelos de regressão: bs bastantes + bom
- Em
- Mét. Imputação

Obs: Riscos Competitivos (Tópico planej. pl fim do curso)



- Riscos Comp.: o tempo ocorre por um causa, e NÃO OCORRERÁ POR OUTRA (morre só uma vez, por uma só causa)

≠

- Censura: o tempo é censurado, MAS O EVENTO IRÁ OCORRER

tal causa ñ é censura

Função de Incidência Acumulada

Análise de Sobrevivência

Prova 1 - 28/04 - Quinta-feira

Aula 05

Até aqui, os tipos de censura e as respectivas construções das funções de verossimilhança.

- Censura à direita
  - Tipo I
  - Tipo II
  - Aleatória
- Censura à esquerda
- Censura Intervalar

Discutimos tbm sobre riscos competitivos.

Truncagem

Ocorre quando parte da população de interesse não pode ser observada.

É possível observar apenas unidades amostrais que experimentam o evento dentro de uma "janela observável" (Y_L, Y_R).

Unidades experimentais com evento fora da janela NÃO SÃO OBSERVÁVEIS e não há informação nenhuma sobre elas.

* Mais com a amostragem do que qqer outra coisa
* Int. relacionado com o evento de interesse

- Se $Y_R = +\infty \rightarrow$ janela $= (Y_L; +\infty)$ - Truncagem à esquerda
- Se $Y_L = 0 \rightarrow$ janela $= (0; Y_R)$ - Truncagem à direita

Exemplos

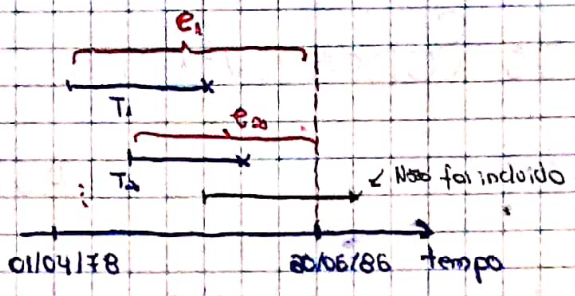
(1) Estudos de HIV com pacientes infectados por transfusão de sangue

Dados disponíveis: 1º de abril de 1978 à 30 de junho de 1986

T_i: tempo decorrido entre transfusão e início dos sintomas

Esquema amostral.

Foram incluídos pacientes que até a data do fim do estudo apresentaram sintomas da doença.



Pacientes infectados mas que em 30/06/86 não apresentaram os sintomas não foram incluídos.

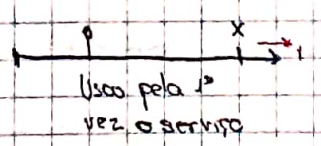
→ Truncagem à Direita

(2) CHANNINGHOUSE: ASILO na California

Estudo de Jan/64 à Jul/75.

→ objetivos do estudo: Idade de óbito de indivíduos que adquiriram o programa de saúde oferecido pelo centro.

Forma de seleção da amostra: foram selecionados apenas usuários que residiam no asilo; não foram incluídos pessoas (em geral mais novas) que adquiriram o programa, porém vieram a óbito antes de terem a oportunidade de utilizar os serviços.



Truncagem à Esquerda. (Aqui pode ter censura tbm)

Construção da Verossimilhança (Truncagem)

(A) Truncagem à direita

Defino

T_i: tempo de falha

e_i: tempo de truncagem

Observamos, portanto

$$P(T_i \in (t_i, t_i + \Delta] \,/\, T_i < e_i)$$

Mas

$$P(T_i \in (t_i, t_i+\Delta] \mid T_i < e_i) = \frac{P(T_i \in (t_i, t_i+\Delta])}{P(T_i < e_i)}$$

$$= \frac{f(t_i)}{1-S(e_i)}$$

A função de verossimilhança para dados truncados à direita (sem censura) fica dada por

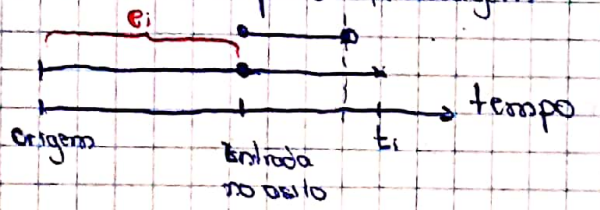
$$L(\Theta) = \prod_{i=1}^{n} \frac{f(t_i)}{1-S(e_i)}$$

(B) Truncagem à Esquerda

Defina

T_i: tempo de falha

e_i: tempo de truncagem



Observamos T_i por que T_i é maior do que e_i.

Temos que considerar a variável truncada, com função de sobrevivência dada por

$$S_{TRUNC}(t) = P(T_i > t \mid T_i > e_i) = P(T_i > t) / S(e_i)$$

$$\rightarrow f_{TRUNC}(t) = f(t) / S(e_i)$$

Para dados com truncagem à esquerda e sujeitos a censura à direita, a função de verossimilhança é dada por

$$L(\Theta) = \prod_{i=1}^{n} f_{TR}(z_i)^{\delta_i} \left(S_{TR}(z_i)\right)^{1-\delta_i}$$

$$= \prod_{i=1}^{n} \left(\frac{f(z_i)}{S(e_i)}\right)^{\delta_i} \left(\frac{S(z_i)}{S(e_i)}\right)^{1-\delta_i}$$

- Forma Geral da Função de Verossimilhança

Sejam

D: conjunto de índices de indivíduos cujo evento é observado

R: conjunto de índices de indivíduos ~~cuja~~ censurados à direita.

L: " " " " " " à esquerda.

I: " " " " " " com censura intervalar.

A função de verossimilhança pode ser escrita na forma.

$$L(\underline{\theta}) \propto \prod_{i \in D} f(t_i) \prod_{i \in R} S(c_i) \prod_{i \in L} (1 - S(e_i)) \prod_{i \in I} (S(a_i) - S(b_i)) \leftarrow$$

Processos de Contagem (Fleming ...) (cap) Flem, Andersen, S Kalb.)

Abordagem muito usada para obtenção de propriedades assintóticas de estimadores.

Noções Básicas

Considere inicialmente o caso sem censura.

Sejam

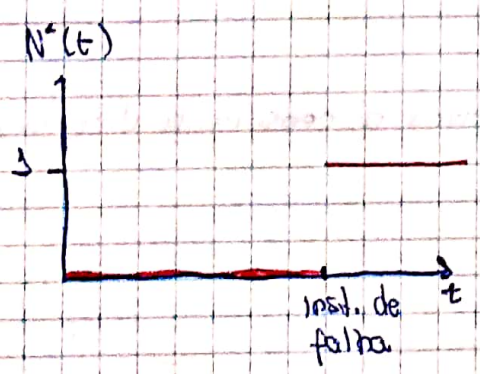
T: tempo de falha

$\alpha(t)$: função de risco associada a T

Defina

$$N^*(t) = I(T \leq t) = \begin{cases} 1, & \text{se } T \leq t \\ 0, & \text{caso contrário} \end{cases}$$

Gráfico



Considere a possibilidade de censura à direita. Sejam

T: tempo de falha

C: tempo de censura

Observamos

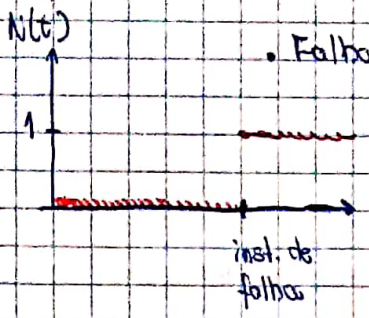
$$Z = \min(T, C)$$

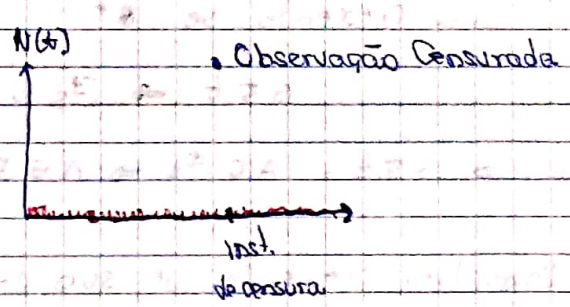
$$\delta = I(T \leq C)$$

Definimos

$$N(t) = I(Z \leq t, \delta = 1) = \begin{cases} 1, & \text{se } Z \leq t, \delta = 1 \\ 0, & \text{c.c.} \end{cases}$$

Graficamente



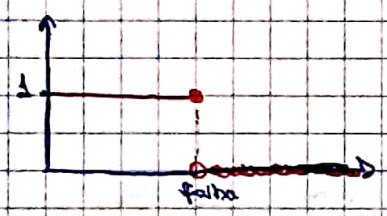


Defina ainda

$$Y(t) = I(Z \geq t) = \begin{cases} 1, & \text{se } Z \geq t \\ 0, & \text{se } Z < t \end{cases}$$

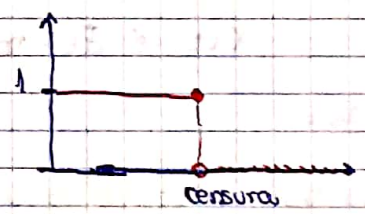
Graficamente

• Falha observada ($\delta = 1$)



Indicador de indivíduo em risco

• Obs. Censurada



Com $N(t)$ e $Y(t)$, temos toda a informação. É uma outra forma de descrever o problema!

Conceitos Básicos

Seja $(\Omega, \mathcal{F}, \mathbb{P})$ um espaço de probabilidade.

Ω: espaço amostral

$\mathcal{F}$: σ-álgebra

$\mathbb{P}$: medida de probabilidade

Definição: Um processo estocástico é uma família de variáveis aleatórias

$$X = \{X(t) : t \in \mathcal{T}\}$$

definidas no espaço de probabilidade $(\Omega, \mathcal{F}, \mathbb{P})$.

Observe que se $\mathcal{T} = \{0, 1, 2, \ldots\}$, temos um processo em tempo discreto e se $\mathcal{T} = [0, \infty)$, X é um processo em tempo contínuo.

Definição: Uma família de sub σ-álgebras $\{\mathcal{F}_t : t \geq 0\}$ de uma σ-álgebra $\mathcal{F}$ é dita crescente se

$$s \leq t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t .$$

(ou seja, p/ $s \leq t$, $A \in \mathcal{F}_s \Rightarrow A \in \mathcal{F}_t$).

Uma família crescente de sub σ-álgebras é denominada FILTRAGEM.

Definição: A filtragem $\mathcal{N}_t = \sigma\{X(s) : s \leq t\}$ é chamada de Filtragem Natural ou história de X.

Definição: Um processo estocástico $\{X(t) : t \in \tau\}$ é ADAPTADO à filtragem $\mathcal{F}_t$ (dizemos $\mathcal{F}_t$-adaptado) se $\forall t \geq 0$,

$$X(t) \text{ é } \mathcal{F}_t\text{-mensurável.}$$

Informalmente: Ao ler

$X(t)$ conhecido quando $\mathcal{F}_t$ é dado, ou seja, $X(t)$ é especificado por $\mathcal{F}_t$.

Observação: É claro que $\{X(t) : t \in \tau\}$ é adaptado a sua história natural ou filtragem natural.

Análise de Sobrevivência

Aula 06

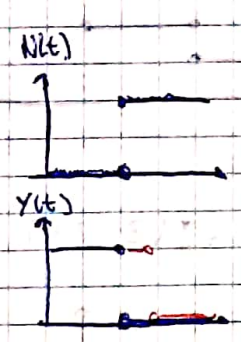
10/03/16

Processo de Contagem

T: tempo de falha

$Z = \min(T, C)$ $\rightarrow$ $N(t) = I(Z \leq t, \delta = 1)$, $t \geq 0$

$\delta = I(T \leq C)$ $\quad Y(t) = I(Z \geq t)$



$(\Omega, \mathcal{F}, P)$: espaço de probabilidade

Filtragem: família crescente de sub σ-álgebras

Filtragem natural ou "história de X"

$$\mathcal{H}_t = \sigma\{X(s) : s \leq t\}$$

Definição: Processo de contagem.

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

(i) $N(0) = 0$

(ii) $N(t) < \infty \quad \forall t$ (q.c.)

(iii) As trajetórias de $N(t)$ são contínuas à direita

(iv) $N(t)$ é constante com "saltos" ou descontinuidades de tamanho $+1$.

Exemplos

(1) T_i: tempo de falha ; C_i: tempo de censura, $i = 1, 2, \ldots, n$.

$\delta_i = I(T_i \leq C_i)$

Definimos

$$N_i(t) = I(Z_i \leq t, \delta_i = 1)$$

$N_i(t)$ são proc. de contagem, pois

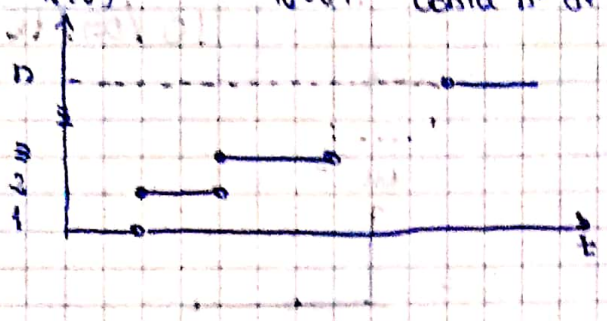
(i) ✓ $N_i(0) = 0$ $\quad$ (iii) ✓

(ii) ✓ $N_i(t) \leq 1 < \infty$ $\quad$ (iv) ✓

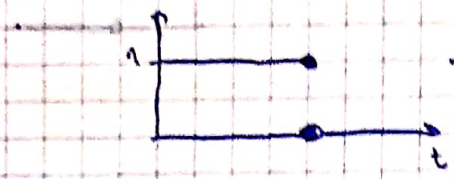
(2) Definimos

$$\bar{N}(t) = \sum_{i=1}^{n} N_i(t)$$

$\bar{N}(t)$ "conta nº de falhas até t"



(3) $Y_1(t) = J(Z; Z \geq t)$ não é um proc. de contagem



Traj. não é contínua à direita

$Y_1(0) = 1 \neq 0$.

Processo de Intensidade

→ Ideia Intuitiva

O processo de interesse está associado com a probabilidade condicional de que um evento ocorra em $[t, t+dt)$ dado o passado anteriormente a esse intervalo.

Podemos escrever

$$dN(t) = N(t) - N(t^-) \quad \text{(definição)}$$

Observe que

$$dN(t) = \begin{cases} 0 \\ 1, \end{cases} \text{ ou}$$

Portanto, podemos escrever o processo de intensidade como

$$\lambda(t)\,dt = \mathbb{P}(dN(t) = 1 \mid \text{Passado})$$

$$[= E(dN(t) \mid \text{Passado})]$$

$$= P(t \leq Z < t+dt, \delta = 1 \mid \text{Passado})$$

$$= P(t \leq Z \leq t+dt, \delta = 1 \mid \text{inf. até } t^-)$$

Notação: $\lambda(t)$
$\Lambda(t)$
$d\Lambda(t)$
Iguais são para definições

Informação até t^-: conhecer o processo até t^- (inst. imed. ant. a t)

$\Rightarrow$ conheço $Y(t^-)$ $\Rightarrow$ conheço $Y(t)$

contínua à esquerda (está viva ou não em f. do risco (ñ falhou, ñ censurado) até t)

Observe que

$$Y(t) = \begin{cases} 0, & \text{se falhou ou censura antes de } t \\ 1, & \text{se em risco em } t. \end{cases}$$

- Se falha ou censura já aconteceu até $t^- \Rightarrow Y(t) = 0$.

$$\begin{aligned} \lambda(t)\,dt &= \mathbb{P}(dN(t) = 1 \mid \text{passado}) \\ &= \mathbb{P}(dN(t) = 1 \mid Y(t) = 0) \\ &= 0. \end{aligned}$$

- Se observação não falhou e não foi censurado até $t^- \Rightarrow Y(t) = 1$

$$\begin{aligned} \lambda(t)\,dt &= \mathbb{P}(dN(t) = 1 \mid Y(t) = 1) \\ &= \mathbb{P}(t \leq Z \leq t + dt, \delta = 1 \mid Y(t) = 1) \\ &= \mathbb{P}(t \leq T \leq t + dt \mid T \geq t, C \geq t) \\ &\overset{\text{Censura indep.}}{=} \mathbb{P}(t \leq T \leq t + dt \mid T \geq t) \\ &\approx \alpha(t)\,dt, \quad \text{se } Y(t) = 1 \end{aligned}$$

Podemos escrever

$$\lambda(t)\,dt = Y(t)\,\alpha(t)\,dt$$

ou

$$\lambda(t) = Y(t)\,\alpha(t).$$

- Função de risco. f. determinística
- $\Lambda(t)$: processo.

Definição: O processo de intensidade associado ao processo $N(t)$ é definido como

$$d\Lambda(t) = E(dN(t) \mid \mathcal{F}_{t^-})$$

em que $dN(t) = N(t) - N(t^-)$, e

$$\mathcal{F}_t = \sigma\{N(u), I(Z \leq u, \delta = 0) : 0 \leq u \leq t\} \text{ ou}$$

$$\mathcal{F}_t = \sigma\{N(u), Y(u^+) : 0 \leq u \leq t\}$$

$\mathcal{F}_t$: toda história do processo até t.

Definição: O Processo de Intensidade Acumulado é dado por

$$\Lambda(t) = \int_0^t d\Lambda(t)$$

Podemos escrever

$$\Lambda(t) = \int_0^t \lambda(u)\,du = \int_0^t Y(u)\alpha(u)\,du$$

Observação:

$\Lambda(t)$ fornece aproximadamente o número esperado de saltos em $[0,t)$

De fato:

$$\begin{aligned}
E(N(t)) &= P(N(t)=1) \\
&= P(Z \leq t, \delta = 1) \\
&= P(T \leq t, C \geq t) \\
&\overset{\text{Exercício}}{=} \int_0^t G(u^-) f(u)\,du \quad , \text{com } G(u) = P(C \geq u) \\
&= \int_0^t G(u^-) \left(\frac{f(u)}{S(u)}\right) S(u)\,du \\
&= \int_0^t P(C \geq u) P(T \geq u) \left(\frac{f(u)}{S(u)}\right) du
\end{aligned}$$

$$\begin{aligned}
E(N(t)) &= \int_0^t P(C \geq u) P(T \geq u) \alpha(u)\,du \\
&= \int_0^t P(C \geq u, T \geq u) \alpha(u)\,du \\
&= \int_0^t P(Z \geq u) \alpha(u)\,du \\
&= \int_0^t E(I(Z \geq u)) \alpha(u)\,du \\
&= E\left(\int_0^t I(Z \geq u) \alpha(u)\,du\right)
\end{aligned}$$

$$E(N(t)) = E\left(\int_0^t Y(u)\alpha(u)\,du\right)$$

$$= E(\Lambda(t))$$

$$\Rightarrow E(N(t)) = E(\Lambda(t)).$$

Definição Martingal

Seja $X = \{X(t) : t \geq 0\}$ um proc. estocástico contínuo à direita e $\{\mathcal{F}_t : t \geq 0\}$ uma filtragem definidos em um mesmo espaço de probabilidade.

Se

1. X é adaptado à $\mathcal{F}_t$;
2. $E|X(t)| < \infty \ \forall t < \infty$
3. $E(X(t+s) \mid \mathcal{F}_s) = X_s$.

então X é um martingal.

Obs: Se (1) e (2) na def. anterior valem, então X é um martingal se e só se

$$E(dX(t) \mid \mathcal{F}_{t^-}) = 0 \quad \text{(informal)}$$

Prova:

- Suponha $E(dX(t) \mid \mathcal{F}_{t^-}) = 0$. Então

$$E(X(t+s) \mid \mathcal{F}_s) - X_s = E(X(t+s) - X(s) \mid \mathcal{F}_s)$$

$$= E\left(\int_s^{t+s} dX(u) \,\middle|\, \mathcal{F}_s\right)$$

$$= \int_s^{t+s} E(dX(u) \mid \mathcal{F}_s)$$

$$= \int_s^{t+s} E\left(\underbrace{E(dX(u) \mid \mathcal{F}_{u^-})}_{0} \,\middle|\, \mathcal{F}_s\right)$$

$$= 0$$

- Suponha agora $E(X(t+s) \mid \mathcal{F}_s) = X(s) \Rightarrow$

$$E(X(t) \mid \mathcal{F}_{t^-}) = X(t^-) \Rightarrow$$

$$E(X(t) - X(t^-) \mid \mathcal{F}_{t^-}) = 0 \Rightarrow$$

$$E(dX(t) \mid \mathcal{F}_{t^-}) = 0.$$

Observação

Se $N(t)$ é um processo de contagem, então

$$\int_0^t f(w)\, dN(w) = \int_s^t f\, dN$$

é a representação da soma dos valores de f nos instantes em $(s,t]$ em que $N(\cdot)$ salta.

Exemplo: Passeio Aleatório.

Seja $Y_1, Y_2, \ldots$ v.a. ind. com mesma dist. c $E(Y_1)=0$.

Suponha tbm que Y_i assuma dois valores apenas: -1 e $+1$.

Seja $X(n) = \sum_{i=0}^{n} Y_i$, com $Y_0 = 0$.

Seja

$$\mathcal{F}_n = \sigma\{X(0), X(1), \ldots, X(n)\}$$
$$= \sigma\{Y_0, Y_1, \ldots, Y_n\}$$

Assim:

(1) X_n é adaptado ao $\mathcal{F}_n$.

(2) $E(|X(n)|) = E(|\sum_{i=0}^{n} Y_i|) < \infty$

(3) $E(X(n+k) / \mathcal{F}_n) = E(X(n+k) / Y_0, Y_1, \ldots, Y_n)$

$$= E(\sum_{i=1}^{n} Y_i + \sum_{i=n+1}^{n+k} Y_i / Y_0, Y_1, \ldots, Y_n)$$

$$= \sum_{i=1}^{n} Y_i + E(\sum_{i=n+1}^{n+k} Y_i / Y_0, \ldots, Y_n)$$

ind.

$$= \sum_{i=1}^{n} Y_i = X(n)$$

Definição: Processo Previsível

Um processo estocástico X é Previsível com relação a filtragem $\mathcal{F}_t$ se, para cada t, o valor de X é especificado ou é uma função de $\mathcal{F}_{t^-}$.

Mais formalmente, $X(t)$ é previsível se $X(t)$ é $\mathcal{F}_{t^-}$ - mensurável.

Resultados:

(1) Todo processo contínuo à esquerda é previsível.

(2) Se X é previsível, então

$$E(X(t) / \mathcal{F}_{t^-}) = X(t).$$

Exemplo: T: tempo de falha

C: tempo de censura

$Z = \min(T, C)$

$Y(t) = I(Z \geq t)$

→ $Y(t)$ é um processo previsível

Vimos ainda que

$$d\Lambda(t) = \underbrace{Y(t)}_{\text{processo previsível}} \underbrace{\alpha(t)dt}_{\text{função determinística}}$$

→ $d\Lambda(t)$ é um Processo Previsível

Logo,

$$E(d\Lambda(t) \mid \mathcal{F}_{t-}) = d\Lambda(t)$$

Vimos que $N(t) = I(Z \leq t, \delta = 1)$ é um proc. de contagem.

Defina agora

$$M(t) = N(t) - \Lambda(t), \quad \Lambda(t) = \int_0^t \lambda(u)du = \int_0^t Y(u)\alpha(u)\,du$$

Resultado: $M(t)$ é um martingal.

De fato,

(1) $M(t) = N(t) - \Lambda(t)$

$N(t)$ é adaptado à $\mathcal{F}_t = \sigma\{N(u), I(Z \leq u, \delta = 0); 0 \leq u \leq t\}$

$\Lambda(t)$ é adaptado à ↗

→ $M(t)$ é adaptado à $\mathcal{F}_t$.

(2) $E(|M(t)|) = E(|N(t) - \Lambda(t)|)$

Como $|a+b| \leq |a| + |b|$, então

$$E(|M(t)|) \leq E(|N(t)|) + E(|-\Lambda(t)|)$$

$$\leq 1 + E(|\Lambda(t)|)$$

$$E(|M(t)|) \leq 1 + E\left(\int_0^t \lambda(s)\,ds\right)$$

$$\leq 1 + E\left(\int_0^t Y(s)\alpha(s)\,ds\right)$$

$$= 1 + E\left(\int_0^t J(Z \geq s)\alpha(s)\,ds\right)$$

$$= 1 + E\left(\int_0^t J(T \geq s) J(C \geq s)\alpha(s)\,ds\right)$$

$$\leq 1 + E\left(\int_0^t J(T \geq s)\alpha(s)\,ds\right)$$

$$= 1 + \left(\int_0^t E(I(T \geq s))\alpha(s)\,ds\right.$$

$$= 1 + \int_0^t P(T \geq s)\alpha(s)\,ds$$

$$= 1 + \int_0^t S(s)\alpha(s)\,ds$$ Teo. Le Boletto

$$= 1 + \int_0^t f(s)\,ds \leq 1 + 1 = 2 \quad \therefore E(|M(t)|) < \infty$$

(3) $E(dM(t)/\mathcal{F}_{t^-}) = E(dN(t) - d\Lambda(t)/\mathcal{F}_{t^-})$

$$= E(dN(t)/\mathcal{F}_{t^-}) - E(d\Lambda(t)/\mathcal{F}_{t^-})$$

$$\overset{\text{def}}{=} d\Lambda(t) - E(d\Lambda(t)/\mathcal{F}_{t^-})$$

$$= d\Lambda(t) - d\Lambda(t) = 0$$

$$\therefore E(dM(t)/\mathcal{F}_{t^-}) = 0 \quad \blacksquare$$

0	1	2	3	4	...	k
p_0	p_1	p_2	p_3	p_4	...	p_k

$$P(A_0, A_0^c, \ldots, A_0^c) = P(A_0)^k (1 - P(A_0))^k$$

A. S.

Aula 07

15/03/16

Proc. de Contagem

$$N(t) = I(Z \leq t, \delta = 1)$$

Proc. de Intensidade

$$\Lambda(t) = \int_0^t \lambda(u)\,du$$

$$= \int_0^t Y(u)\alpha(u)\,du$$

Martingal

1. $\mathcal{F}_t$ - mens.
2. $E|X(t)| < \infty$
3. $E(X(t+s)/\mathcal{F}_s) = X(s) \Leftrightarrow E(dX(t)/\mathcal{F}_{t^-}) = 0$

Proc. Previsível

$$Y(t) = I(Z \geq t)$$

Definimos

$$M(t) = N(t) - \Lambda(t)$$

→ $M(t)$ é um martingal de média zero.

$$\Lambda(t) = \int_0^t Y(u)\alpha(u)\,du$$

Observação:

$\lambda(t) = Y(t)\alpha(t)$ se supomos censura independente.

Resultado: Fleming, pg 26.

Seja T um tempo de falha e C o tempo de censura. Defina

$$Z = \min(T, C)$$

$$\delta = I(T \leq C)$$

$$N(t) = I(Z \leq t, \delta = 1)$$

O processo dado por

$$M(t) = N(t) - \int_0^t I(Z \geq u)\alpha(u)\,du$$

é um martingal com respeito a filtragem

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

se e somente se

$$\alpha(t) = \left(- \partial \mathbb{P}(T \geq u, C \geq t) / \partial u \Big|_{u=t} \right) \Big/ \mathbb{P}(T \geq t, C \geq t)$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \mathbb{P}(t \leq T \leq t + \Delta t \mid T \geq t, C \geq t) \quad \forall t \text{ com } \mathbb{P}(Z \geq t) > 0.$$

Decomposição de Doob-Meier

Definição: Submartingal.

Seja o processo $X = \{X(t) : t \geq 0\}$ contínuo à direita e $\mathcal{F}_t$ uma filtragem definidos no mesmo espaço de probabilidade se

(1) X é adaptado à $\mathcal{F}_t$;

(2) $E(|X(t)|) < \infty \ \forall t$; e

(3) $E(X(t+s) \mid \mathcal{F}_s) \geq X(s)$

então X é um submartingal.

Obs: Um submartingal é um processo que tende a aumentar quando o tempo passa.

Teorema. Decomposição de Doob-Meyer.

Seja X um sub-martingal contínuo à direita e não negativo adaptado à filtragem $\mathcal{F}_t$. Então, existe um martingal M contínuo à direita e um processo previsível A tal que

$$E(A(t)) < \infty \text{ e}$$

$$X(t) = M(t) + A(t), \ \forall t \geq 0$$

se $A(0) = 0$ e $X = M' + A'$ é uma outra decomposição com $A'(0) = 0$, então $\forall t \geq 0$

$$\mathbb{P}(M'(t) \neq M(t)) = 0 \rightarrow \mathbb{P}(A'(t) \neq A(t))$$

Corolário. Seja $\{N(t) : t \geq 0\}$ um processo de contagem adaptado à filtragem $\{\mathcal{F}_t : t \geq 0\}$ com $E(N(t)) < \infty \ \forall t$.

Então existe um único processo $\Lambda(t)$ não decrescente e previsível tal que

- $\Lambda(0) = 0$;
- $E(\Lambda(t)) < \infty$, $\forall t$
- $M(t) = N(t) - \Lambda(t)$ é um martingal

$\Lambda(t)$ é conhecido como compensador de $N(t)$.

Exemplo.

$$N(t) = I(Z \leq t, \delta = 1)$$

Vimos que

$$M(t) = N(t) - \int_0^t Y(u)\alpha(u)\,du$$

$$= N(t) - \Lambda(t)$$

é um martingal.

Pela Decomposição de Doob-Meier,

$$\Lambda(t) = \int_0^t Y(u)\alpha(u)\,du$$

é o único compensador de $N(t)$. Observe que

$$\underbrace{N(t)}_{\text{observável}} = M(t) + \Lambda(t)$$

$$= \underbrace{\Lambda(t)}_{\text{parte sistemática}} + \underbrace{M(t)}_{\text{erro aleatório}}$$

Definição: Processo de Variação Previsível

Para um martingal contínuo à direita e adaptado à $\{\mathcal{F}_t : t \geq 0\}$ define-se o proc. de variação previsível $\langle M \rangle(t)$ por

$$\langle M \rangle(t) = \lim_{n \to \infty} \sum_{k=1}^{n} Var(\Delta M_k \mid \mathcal{F}_{(k-1)t/n})$$

em que o intervalo $[0,t]$ é particionado em n intervalos de comprimento t/n e

$$\Delta M_k = M(kt/n) - M((k-1)t/n)$$

Informalmente, temos

$$d\langle M\rangle(t) = Var(dM(t) \mid \mathcal{F}_{t^-})$$

ou

$$\langle M\rangle(t) = \int_0^t Var(dM(u) \mid \mathcal{F}_{u^-})$$

<u>Observação</u>:

$$Var(dM(t) \mid \mathcal{F}_{t^-}) = E((dM(t))^2 \mid \mathcal{F}_{t^-}) - (E(dM(t) \mid \mathcal{F}_{t^-}))^2$$

$E(dM(t) \mid \mathcal{F}_{t^-}) = 0$ pois $M(t)$ é um martingal

$$Var(dM(t) \mid \mathcal{F}_{t^-}) = E((dM(t))^2 \mid \mathcal{F}_{t^-}). \quad (*)$$

Considere o processo

$$M^2(t) - \langle M\rangle(t) = U(t)$$

Observe que

$$\begin{aligned} d(M^2(t)) &= M^2(t) - M^2(t^-) \\ &= (M(t))^2 - (M(t^-))^2 \\ &= (M(t) - M(t^-) + M(t^-))^2 - (M(t^-))^2 \\ &= (dM(t) + M(t^-))^2 - (M(t^-))^2 \\ &= (dM(t))^2 + 2\,dM(t)\,M(t^-) + (M(t^-))^2 - (M(t^-))^2 \\ &= (dM(t))^2 + 2\,dM(t)\,M(t^-) \end{aligned}$$

Vamos calcular

$$E(dM^2(t) \mid \mathcal{F}_{t^-}) = E((dM(t))^2 \mid \mathcal{F}_{t^-}) + 2E(dM(t)M(t^-) \mid \mathcal{F}_{t^-})$$

Observe que

$$E(dM(t)\,M(t^-) \mid \mathcal{F}_{t^-}) = M(t^-)\,\underbrace{E(dM(t) \mid \mathcal{F}_{t^-})}_{0,\ \text{uma vez } M(t) \text{ é mart.}}$$

$$= M(t^-) \cdot 0 = 0$$

$\therefore$

$$E(d(M^2(t)) \mid \mathcal{F}_{t^-}) = E((dM(t))^2 \mid \mathcal{F}_{t^-}) \overset{(*)}{=} Var(dM(t) \mid \mathcal{F}_{t^-}) \overset{def}{=} d\langle M\rangle(t)$$

15/08/2016

Aula 07

Podemos escrever

$$E(d(M^2(t))/\mathcal{F}_{t^-}) - d\langle M\rangle(t) = 0. \Rightarrow$$

$$E(d(M^2(t)) - d\langle M\rangle(t)/\mathcal{F}_{t^-}) = 0 \Rightarrow$$

$$E(d(M^2(t) - \langle M\rangle(t))/\mathcal{F}_{t^-}) = 0 \Rightarrow$$

$$E(dU(t)/\mathcal{F}_{t^-}) = 0.$$

Resultado: Seja M um martingal contínuo à direita com relação a filtragem $\{\mathcal{F}_t : t \geq 0\}$ e assuma $E(M^2(t)) < \infty$ $\forall t \geq 0$.

Então existe um único processo previsível $\langle M\rangle(t)$ denominado Processo de Variação Previsível de M tal que.

- $\langle M\rangle(0) = 0$
- $E(\langle M\rangle(t)) < \infty$ $\forall t$ e
- $\{M^2(t) - \langle M\rangle(t) : t \geq 0\}$ é um martingal contínuo à direita.

Obs: Assuma $M(0) = 0$. Então

$$E(M(t)) = E(E(M(t)/\mathcal{F}_0)) \overset{\text{pois } M(t) \text{ é um martingal}}{=} E(M(0)) = 0$$

$\Rightarrow Var(M(t)) = E(M^2(t))$.

Pelo resultado, $\langle M\rangle(0) = 0$ e temos ainda que $M^2(0) = 0$.

Então,

$$E(M^2(t) - \langle M\rangle(t)) = E(E(M^2(t) - \langle M\rangle(t)/\mathcal{F}_0))$$
$$= E(M^2(0) - \langle M\rangle(0)) = 0$$

$$\Rightarrow E(M^2(t)) = E(\langle M\rangle(t))$$
$$\Downarrow$$
$$Var(M(t)) = E(\langle M\rangle(t)) \quad \text{(Kalb fleisch ...)}$$

Exemplo:

T: tempo de falha

C: tempo de censura

$Z = \min(T, C)$

$N(t) = I(Z \leq t, \delta = 1)$

$Y(t) = I(Z \geq t)$

Vimos que

$$M(t) = N(t) - \Lambda(t)$$

é um martingal de média zero com

$$\Lambda(t) = \int_0^t Y(u)\alpha(u)\,du$$

Vimos que

$$E(dN(t) \mid \mathcal{F}_{t^-}) = Y(t)\alpha(t)\,dt = d\Lambda(t)$$

Assim,

$$\begin{aligned} d\langle M \rangle(t) &= Var(dM(t) \mid \mathcal{F}_{t^-}) \\ &= Var(dN(t) - d\Lambda(t) \mid \mathcal{F}_{t^-}) \\ &= Var(dN(t) \mid \mathcal{F}_{t^-}) \end{aligned}$$

const, uma vez que é previsível

Relembre que $dN(t)$ é v.a. binária. Então

$$\begin{aligned} d\langle M \rangle(t) &= E((dN(t))^2 \mid \mathcal{F}_{t^-}) - (E(dN(t) \mid \mathcal{F}_{t^-}))^2 \\ &= E(dN(t) \mid \mathcal{F}_{t^-}) - [E(dN(t) \mid \mathcal{F}_{t^-})]^2 \\ &= d\Lambda(t) - (d\Lambda(t))^2 \end{aligned}$$

$d\langle M \rangle(t) \approx d\Lambda(t)$ ou

$$\langle M \rangle(t) = \Lambda(t)$$

Definição: Sejam M_1 e M_2 dois martingais contínuos à direita e assuma $E(M_i(t))<\infty$, $\forall t\geq 0$, $i=1,2$.

O processo de Covariância Previsível é dado por

$$\langle M_1, M_2\rangle(t) = \lim_{n\to\infty} \sum_{k=1}^{n} Cov(\Delta M_{1k}, \Delta M_{2k} \mid \mathcal{F}_{(k-1)t/n})$$

em que o intervalo $[0,t]$ é particionado em n subintervalos de comprimento t/n e

$$\Delta M_{ik} = M_i(tk/n) - M_i(t(k-1)/n), \; i=1,2$$

Podemos escrever

$$d\langle M_1, M_2\rangle(t) = Cov(dM_1(t), dM_2(t) \mid \mathcal{F}_{t^-})$$

Resultado: Sejam M_1 e M_2 dois martingais contínuos à direita com respeito a uma filtragem $\mathcal{F}_t$ e assuma $E(M_i(t))^2<\infty$ $\forall t\geq 0$, $i=1,2$.

Então existe um processo previsível contínuo à direita $\langle M_1, M_2\rangle$ chamado de processo de covariância com $\langle M_1, M_2\rangle(0)=0$,

$$E(\langle M_1, M_2\rangle(t))<\infty \text{ t.q.}$$

1. $\langle M_1, M_2\rangle$ é a dif. de dois processos crescentes previsíveis contínuos à direita
2. $M_1 M_2 - \langle M_1, M_2\rangle$ é um martingal

Definição: Processo de Variação Quadrática

Para um martingal contínuo à direita e adaptado a $\{\mathcal{F}_t; t\geq 0\}$, define-se o Processo de Variação Quadrática $[M](t)$ por

$$[M](t) = \lim_{n\to\infty} \sum_{k=1}^{n} (\Delta M_k)^2$$

com

$$\Delta M_k = M(tk/n) - M(t(k-1)/n)$$

sendo que o intervalo $(0,t]$ é particionado em n subintervalos de comprimento t/n

Podemos escrever

$$[M](t) = \sum_{s\leq t} (dM(s))^2$$ soma é nos inst. de descontinuidade de M

Resultados:

- $M^2(t) - [M](t)$ é um martingal de média zero
- $Var(M(t)) = E([M](t))$

Exemplo:

$$N(t) = I(Z \leq t, \delta = 1)$$

$$Y(t) = I(Z \geq t)$$

$M(t)$ é um martingal

$$M(t) = N(t) - \Lambda(t)$$

Vimos que

$$\langle M \rangle (t) = \Lambda(t)$$

Vamos obter $[M](t)$.

$$[M](t) = \sum_{s \leq t} (dM(s))^2$$

$$= \sum_{s \leq t} (dN(s))^2 = \sum_{s \leq t} dN(s) = N(t).$$

$$\therefore [M](t) = N(t).$$

Análise de Sobrevivencia

Aula 08

17/03/16

$Z = \min(T, C)$

$\delta = I(T \leq C)$

$N(t) = I(Z \leq t, \delta = 1)$

$Y(t) = I(Z \geq t)$

$M(t) = N(t) - \Lambda(t)$

$\Lambda(t) = \int_0^t I(Z \geq u)\alpha(u)\,du$

$M(t)$ é um martingal

$\langle M \rangle (t) = \int_0^t Var(dM(u) / \mathcal{F}_{u^-})$

$E(\langle M \rangle (t)) = Var(M(t))$

$[M](t) = \sum_{s \leq t} (dM(s))^2$ (Proc. de var. opcional)

pontos de descontinuidade de M

$Var(M(t)) = E([M](t))$

Integrais Estocásticas

Uma possível definição de integral estocastica é

$$L(t) = \int_0^t H(s)\,dM(s)$$

$$= \lim_{n \to \infty} \sum_{k=1}^{n} H_k \Delta M_k$$

em que o intervalo $(0, t]$ é particionado em subintervalos de tamanho t/n,

$$H_k = H\left(\frac{(k-1)t}{n}\right), \quad \Delta M_k = M\left(\frac{tk}{n}\right) - M\left(\frac{t(k-1)}{n}\right).$$

$H(t)$ é um processo <u>Previsível</u> e $M(t)$ um martingal com relação à história $\mathcal{F}_t$.

Considere que

$$M(t) = N(t) - \Lambda(t), \text{ com}$$

$N(t)$: processo previsível

$\Lambda(t)$: compensador

Neste caso,

$$L(t) = \int_0^t H(s)\, dM(s) =$$

$$= \int_0^t H(s)\, d(N(s) - \Lambda(s))$$

$$\therefore L(t) = \int_0^t H(s)\, dN(s) - \int_0^t H(s)\, d\Lambda(s)$$

$$= \int_0^t H(s)\, dN(s) - \underbrace{\int_0^t H(s)\lambda(s)\, ds}_{\text{int. de Riemann}}$$

$$= \sum_{\substack{t_j:\ \text{saltos de } N(t) \\ t_j \leq t}} H(t_j) - \int H(s)\lambda(s)\, ds$$

Resultado:

$L(t)$ em (*) é um martingal (sob cond. gerais).

De fato:

$$E(dL(t) / \mathcal{F}_{t^-}) = E(H(t)\, dM(t) / \mathcal{F}_{t^-}) = H(t)\, E(dM(t) / \mathcal{F}_{t^-}) = 0$$

Além disso,

$$E(L(0)) = E\left(\int_0^0 H(s)\, dM(s)\right) = 0 \quad \text{e}$$

$$E(L(t)) = E\left(E(L_t / \mathcal{F}_0)\right) = E(L(0)) = 0.$$

$L(t)$ é um martingal de média zero.

Processo de Variação Previsível de $L(t)$

$$d\langle L \rangle(t) \overset{\text{def}}{=} Var(dL(t) / \mathcal{F}_{t^-})$$

$$= Var(H(t)\, dM(t) / \mathcal{F}_{t^-})$$

$$= H^2(t)\, Var(dM(t) \mid \mathcal{F}_{t^-})$$

$$= H^2(t)\, d\langle M \rangle(t).$$

ou

$$\langle L \rangle(t) = \int_0^t H^2(s)\, d\langle M \rangle(s)$$

Processo de Variação Quadrática

$$[L](t) = \sum_{s \le t} (dL(s))^2 = \sum_{s \le t} (H(s)\, dM(s))^2$$

$$= \sum_{s \le t} H^2(s)\, (dM(s))^2.$$

$$= \sum_{s \le t} H^2(s)\, dN(s)$$

$$\Rightarrow [L](t) = \int_0^t H^2(s)\, dN(s)$$

Considere que

$$M(t) = N(t) - \Lambda(t), \text{ com}$$

$N(t)$: processo previsível

$\Lambda(t)$: compensador

Neste caso,

$$L(t) = \int_0^t H(s)\, dM(s) \overset{(*)}{=} \int_0^t H(s)\, d(N(s) - \Lambda(s))$$

Processo de Wiener

O proc. de Wiener padrão é um proc. estocástico satisfazendo

- $W(0) = 0$ e $E(W(t)) = 0 \ \forall t$.
- $W(t)$ tem incrementos independentes, ou seja, $W(t) - W(u)$ é ind. de $W(w)$, $0 \le w \le t$.
- $W(t) - W(s)$ tem dist. normal com

$$E(W(t) - W(s)) = 0 \quad e$$

$$Var(W(t) - W(s)) = t - s.$$

- W tem trajetórias contínuas.

Seja agora $V(t)$ uma função estritamente crescente, contínua, com $V(0) = 0$.

Considere o processo

$$U(t) = W(V(t))$$

Então $U(t)$

- tem trajetórias contínuas
- tem incrementos independentes
- $W(V(t)) - W(V(s))$ tem dist. normal com média zero e variância $V(t)-V(s)$.

$U(t)$ é dito um Martingal Gaussiano com processo de variação previsível

$$\langle U \rangle (t) = V(t)$$

Processos de Contagem Multivariados

Definição: O processo $\{N_1, N_2, \dots, N_n\}$ é chamado de processo de contagem multivariado se cada N_i é um processo de contagem e dois processos não "saltam" no mesmo instante.

Teorema: Seja $\{N_1, N_2, \dots, N_n\}$ um processo de contagem multivariado e Λ_i seus respectivos compensadores, $i = 1, 2, \dots, n$. Defina

$$U_l(t) = \sum_{i=1}^{n} \int_0^t H_{i,l} \, dM_i(u), \quad l = 1, 2, \dots, K$$

$H_{i,l}$ são processos previsíveis limitados. Então:

(i) $U_l(t)$ é um martingal

(ii) $E(U_l(t)) = 0 \quad \forall \, 0 \leq t < \infty$;

(iii) $Cov(U_l(s), U_{l'}(t)) = \sum_{i=1}^{n} \int_0^{min(t,s)=t\wedge s} E\left[H_{i,l}(u) H_{i,l'}(u) \, d\Lambda_i(u) \right]$

De (iii), decorre que

$$Var(U_l(t)) = \sum_{i=1}^{n} \int_0^t E(H_{i,l}^2(u) \, d\Lambda_i(u))$$

$$Cov(U_l(s), U_l(t)) = \sum_{i=1}^{n} \int_0^{t\wedge s} E(H_{i,l}^2(u) \, d\Lambda_i(u))$$

Aula 08

Teorema de Rebolledo (versão para processos de contagem)

Sejam

- $\underline{N(t)} = (N_1(t), N_2(t), \dots, N_n(t))$ processo de contagem multivariado
- $\Lambda(t) = (\Lambda_1(t), \Lambda_2(t), \dots \Lambda_n(t))$ os respectivos compensadores

Para processos previsíveis,

$$H_{ji}(t) \quad , j = 1, 2, \dots, k$$
$$i = 1, 2, \dots, n$$

defina os martingais

$$U_j^{(n)}(t) = \sum_{i=1}^{n} \int_0^t H_{ji}(u)\,(dN_i(u) - d\Lambda_i(u)), \quad j = 1, 2, \dots, k$$

$$U_{j\varepsilon}^{(n)}(t) = \sum_{i=1}^{n} \int_0^t H_{ji}(s)\, I(|H_{ji}(s)| > \varepsilon)\,(dN_i(s) - d\Lambda_i(s)) \quad \text{e}$$

$$\underset{\sim}{U}^{(n)}(t) = (U_1^{(n)}(t), U_2^{(n)}(t), \dots, U_k^{(n)}(t)) \quad \text{vetor } k \times 1$$

Considere as condições

(a) $\langle \underline{U}^{(n)} \rangle (t) \xrightarrow[n \to \infty]{P} V(t)$

$\langle U_l^{(n)}, U_{l'}^{(n)} \rangle \xrightarrow[n \to \infty]{P} v_{ll'}(t)$

$V(t)$: matriz determinística e positiva definida

(b) $\langle U_{j\varepsilon}^{(n)} \rangle (t) \xrightarrow[n \to \infty]{P} 0 \quad , j = 1, 2, \dots, k$

Então, se (a) e (b) são válidas,

$$\underset{\sim}{U}^{(n)}(t) \xrightarrow{D} \underset{\sim}{U}^{(\infty)} \text{, martingal gaussiano com}$$

$\langle U^{(\infty)} \rangle (t) = V(t)$.

Em particular, para t_0 fixado:

$$\underset{\sim}{U}^{(n)}(t_0) \xrightarrow{D} N(0, V(t_0))$$

Obs:

$U_{j\delta}^{(n)}(t)$: contém os saltos de $U_j^{(n)}(t)$ maiores que δ (em valor absoluto)

Estimação Não Paramétrica de Função de Risco Acumulada e Função de Sobrevivência

- Estimador de Nelson-Aalen para $\Lambda(t)$

Sejam

T: tempo de falha com função densidade de prob. $f(\cdot)$, função de sobrevivência $S(\cdot)$ e função de risco $\alpha(t)$

C: tempo de censura

Assuma $T \perp C$ e defina

$$\delta = I(T \leq C), \quad Z = \min(T, C)$$

Assuma que temos disponível uma amostra aleatória (Z_i, δ_i), $i = 1, 2, \ldots, n$, de (Z, δ).

Sejam

$t_1 < t_2 < \ldots < t_k$: tempos de falha, ou seja, instantes em que pelo menos uma falha aconteceu. Note que $k \leq n$ (pode haver empates).

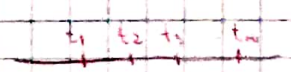
Para $j = 1, 2, \ldots, k$, defina

d_j: nº de falhas em t_j

r_j: nº de indivíduos em risco em t_j.

Indivíduos em risco em t_j: aqueles que não falharam e não foram censurados até t_j.

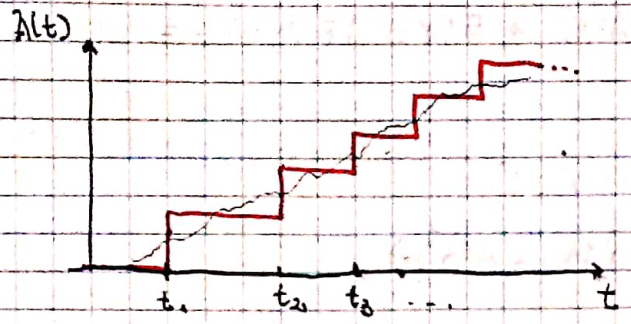
Intuitivamente:



$\hat{\alpha}(t_j) = d_j / r_j$ O estimador de Nelson-Aalen é definido por

$$\hat{A}(t) = \sum_{j: t_j \le t} d_j / r_j$$

1 ✓ (Maior perg)
2 ✓
3 ✓ (X ST fe Var (Sob) Lig)
✗ Dim + Tolerância
✗
✗



$N_i(t) = I(z_i \le t, \delta_i = 1)$

$Y_i(t) = I(z_i \ge t)$, $i = 1, 2, \dots, n$

Expressão do estimador de Nelson-Aalen utilizando notação de processos de contagem

Defina

$\bar{N}(t) = \sum_{i=1}^{n} N_i(t)$: nº de falhas até t.

$\bar{Y}(t) = \sum_{i=1}^{n} Y_i(t)$: nº de indivíduos em risco em t.

Então:

$$\hat{A}(t) = \sum_{j: t_j \le t} \frac{r_j}{d_j} = \int_0^t d\bar{N}(u) / \bar{Y}(u)$$

Propriedades do Estimador de Nelson Aalen

Seja

$$A^*(t) = \int_0^t I(\bar{Y}(u) > 0)\, \alpha(u)\, du \qquad \text{(Hem., 92)}$$

e

$$\pi_i(t) = P(z_i \ge t) =: \pi(t)$$

Resultados:

Supondo $A(t) < \infty \ \forall t \ge 0$, temos

① $E(\hat{A}(t) - A^*(t)) = 0$

② $E(\hat{A}(t) - A(t)) = -\int_0^t (1-\pi(s))^n \, dA(s) \ge -(1-\pi(t))^n A(t)$

③ $\sigma_n^2 \overset{def}{=} E\left(\left[\sqrt{n}\,(\hat{A}(t) - A^*(t))\right]^2\right) = E\left(n \int_0^t \frac{I(\bar{Y}(s) > 0)}{\bar{Y}(s)} \alpha(s)\, ds\right)$

Observação:

(1) $A^*(t) = \int_0^t I(Y(u)>0)\,\alpha(u)\,du$

$$\leq \int_0^t \alpha(u)\,du = A(t) \quad , \quad A^*(t) \leq A(t) \Rightarrow E(A^*(t)) \leq A(t) = E(A(t))$$

(2) $E(\hat{A}(t) - A(t)) = E(\hat{A}(t) - A^*(t) + A^*(t) - A(t))$

$$= \cancelto{0}{E(\hat{A}(t) - A^*(t))} + E(A^*(t) - A(t))$$

$$= E(A^*(t) - A(t)) \leq 0.$$

Análise de Sobrevivência

Aula 09

29/03/15

Na aula passada, encerramos vendo que

(2) $E[\hat{A}(t) - A(t)] = E[A^*(t) - A(t)] \leq 0$

"O estimador de Nelson-Aalen tem vício negativo, que tende à zero qdo $n \nearrow \infty$.

$E[\hat{A}(t) - A(t)] = -\int_0^t (1-\pi(s))^n dA(s) \geq -(1-\pi_0(t))^n A(t)$ " é ≤ 0 e tbm maior ou igual à uma qtde negativa

(3) $\sigma_*^2 \overset{def}{=} E\left(\left[\sqrt{n}\,(\hat{A}(t) - A^*(t))\right]^2\right) =$

$= E\left(n \int_0^t \frac{I(\bar{Y}(s)>0)}{\bar{Y}(s)} \alpha(s)\, ds\right)$

Provas:

1. $\hat{A}(t) - A^*(t) = \int_0^t \frac{d\bar{N}(u)}{\bar{Y}(u)} - \int_0^t I(\bar{Y}(u)>0)\alpha(u)\,du$

$= \int_0^t \left(\frac{d\bar{N}(u)}{\bar{Y}(u)} - I(\bar{Y}(u)>0)\alpha(u)\,du \right)$

$= \int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)} \left(d\bar{N}(u) - \underbrace{\bar{Y}(u)\alpha(u)\,du}_{d\bar{\Lambda}(u)} \right)$

$= \int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\, d\bar{M}(u)$

$\bar{M}(u) = \bar{N}(u) - \bar{\Lambda}(u)$

"Prov. que $\bar{M}(t)$ é um mart."

Defino ainda

$H(u) = \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}$ é um processo previsível

$\Rightarrow \hat{A}(t) - A^*(t) = \int_0^t H(u)\, d\bar{M}(u)$, como $H(u)$ é um proc. prev. e $\bar{M}(t)$ é um martingal, então $\hat{A}(t) - A^*(t)$ é um martingal de média zero

$$\Rightarrow E[\hat{A}(t) - A^*(t)] = 0 \quad \blacksquare$$

2. $E[\hat{A}(t) - A(t)] = E[\hat{A}(t) - A^*(t) + A^*(t) - A(t)]$

$$= \cancelto{0}{E[\hat{A}(t) - A^*(t)]} + E[A^*(t) - A(t)]$$

$$= \int_0^t [E(I(\bar{Y}(u) > 0) - 1]\, \alpha(u)\, du$$

$$= \int_0^t [\mathbb{P}(\bar{Y}(u) > 0) - 1]\, \alpha(u)\, du$$

$$= -\int_0^t [1 - \mathbb{P}(\bar{Y}(u) > 0)]\, \alpha(u)\, du$$

$$= -\int_0^t \mathbb{P}(\bar{Y}(u) = 0)\, \alpha(u)\, du$$

$\bar{Y}(u) = \sum_{i=1}^{n} Y_i(u)$; $P(\bar{Y}(u) = 0) = P(Y_1 = 0, Y_2 = 0, \dots, Y_n = 0) = \prod_{i=1}^{n} \mathbb{P}(Y_i = 0)$

$$= \prod_{i=1}^{n} (1 - \mathbb{P}(Y_i > 0)) = \prod_{i=1}^{n} (1 - \pi(t)) = [1 - \pi(t)]^n$$

$$\pi_i(t) = \mathbb{P}(Z_i > t) = \mathbb{P}(Y_i(t) > 0)$$

$$\Rightarrow E[\hat{A}(t) - A(t)] = -\int_0^t (1 - \pi(u))^n \alpha(u)\, du = -\int_0^t (1 - \pi(u))^n\, dA(u).$$

Agora, vamos provar a desigualdade:

Note que:

$$1 - \pi(u) = \mathbb{P}(Z_i < u) \leq P(Z_i \leq t) = 1 - \pi(t), \quad \forall u \leq t.$$

$$1 - \pi(u) \leq 1 - \pi(t) \Rightarrow -(1 - \pi(u)) \geq -(1 - \pi(t)), \quad \forall u \leq t$$

Assim:

$$-\int_0^t (1 - \pi(u))^n \alpha(u)\, du \geq -\int_0^t (1 - \pi(t))^n \alpha(u)\, du =$$

$$= -(1 - \pi(t))^n \int_0^t \alpha(u)\, du$$

$$= -(1 - \pi(t))^n A(t)$$

$$E[\hat{A}(t) - A(t)] = -\int_0^t (1 - \pi(u))^n\, dA(u) \geq -(1 - \pi(t))^n A(t) \quad \blacksquare$$

3. $\sigma_*^2(t) \overset{def}{=} E[(\sqrt{n}\,(\hat{A}(t) - A^*(t))^2] = Var(\sqrt{n}\,(\hat{A}(t) - A^*(t)))$

Seja $L(t) = \sqrt{n}\,(\hat{A}(t) - A^*(t))$. Então

$$L(t) = \sqrt{n}\int_0^t \frac{d\bar{N}(u)}{\bar{Y}(u)} - \sqrt{n}\int_0^t I(\bar{Y}(u)>0)\alpha(u)\,du$$

(ideia é escrever em termos de proc. de contagem, chegar à algo como martingais)

$$= \sqrt{n}\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\left(d\bar{N}(u) - \bar{Y}(u)\alpha(u)\,du\right)$$

$$= \sqrt{n}\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\,d\bar{M}(u)$$

então: $L(t) = \int_0^t H(u)\,d\bar{M}(u)$, com $H(u) = \sqrt{n}\,\frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}$ "proc. previsível"

$\Rightarrow L(t)$ é um martingal de média zero.

Podemos escrever

$$dL(t) = H(t)\,d\bar{M}(t),$$

observe que queremos calcular $E[(L(t))^2]$. Como $L(t)$ é um martingal de média zero,

$$E((L(t))^2) = Var(L(t)) \overset{*}{=} E(\langle L\rangle(t))$$ * por resultados anteriores

Temos ainda o resultado:

$$\langle L\rangle(t) = \int_0^t H^2(u)\,d\langle\bar{M}\rangle(u) = \int_0^t \left(\sqrt{n}\,\frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\right)^2 d\langle\bar{M}\rangle(u)$$

Pode-se demostrar que.

$$d\langle\bar{M}\rangle(u) = d\bar{\Lambda}(u) = \bar{Y}(u)\alpha(u)\,du.$$

$$\langle L\rangle(t) = \int_0^t n\,\frac{I(\bar{Y}(u)>0)}{(\bar{Y}(u))^2}\,\bar{Y}(u)\alpha(u)\,du$$

$$= n\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\alpha(u)\,du \Rightarrow E[\langle L\rangle(t)]$$

$$= E\left(n\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\alpha(u)\,du\right)$$

Resultado. (Fleming pag 93-94)

$$\sigma_+^2(t) \approx Var(\sqrt{n}\,\hat{A}(t))$$

Defina agora.

$$\hat{\sigma}^2(t) = \int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)/n}\, d\hat{A}(u)$$

$$= \int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)/n} \frac{d\bar{N}(u)}{\bar{Y}(u)} \Rightarrow \hat{\sigma}^2(t) = \int_0^t \frac{d\bar{N}(u)}{(\bar{Y}(u))^2/n}$$

Resultado: $E(\hat{\sigma}^2(t) - \sigma_0^2(t)) = 0$

Prova:

$$E(\hat{\sigma}^2(t) - \sigma_0^2(t)] = E\left[\int_0^t \left(\frac{d\bar{N}(u)}{(\bar{Y}(u))^2/n} - n\frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\, dA(u)\right)\right]$$

$$= E\left[\int_0^t \frac{n I(\bar{Y}(u)>0)}{(\bar{Y}(u))^2}\left(d\bar{N}(u) - \bar{Y}(u)\, dA(u)\right)\right]$$

$$= E\left[\int_0^t \underbrace{\frac{n I(\bar{Y}(u)>0)}{(\bar{Y}(u))^2}}_{\text{proc. prev.}} \cdot \underbrace{d\bar{M}(u)}_{\text{martingal}}\right] = 0$$

Distribuição Assintótica (Aalen, 89)

Assuma que exista uma função $\pi(t)$ t.q.

$$\frac{\bar{Y}(t)}{n} \xrightarrow{P} \pi(t), \ \forall t \in [0, \tau], \ \tau = \sup\{t : \pi(t) > 0\}.$$

Então para $\pi(t) > 0$, temos:

(•) $$\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\, dA(u) \xrightarrow{P} \int_0^t \frac{dA(u)}{\pi(u)} = \sigma^2(t)$$

(•) $$\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)/n} \cdot I\left[\left|\sqrt{n}\, \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\right| > \varepsilon\right] d\,A(u) \xrightarrow{P} 0.$$

Teorema: O processo $\sqrt{n}\,(\hat{A}(t) - A(t))$, $0 \le t \le \tau$, converge em distribuição para um martingal Gaussiano de média zero e função de covariância

$$\sigma^2(t) = \int_0^t dA(u) \Big/ \pi(u)$$

que pode ser consistentemente estimado por

$$\hat{\sigma}^2(t) = \int_0^t \frac{d\bar{N}(u)}{\bar{Y}^2(u)/n}$$

Em particular, para t_0 fixado,

Aula 09

29/03

$$\sqrt{n}\,(\hat{A}(t_0) - A(t_0)) \xrightarrow{D} N(0, \sigma^2(t_0)).$$

Estimador de Kaplan-Meier para $S(t)$

Seja $(T_1, \delta_1), (T_2, \delta_2), \dots, (T_n, \delta_n)$ uma a.a. de (T, δ) em que

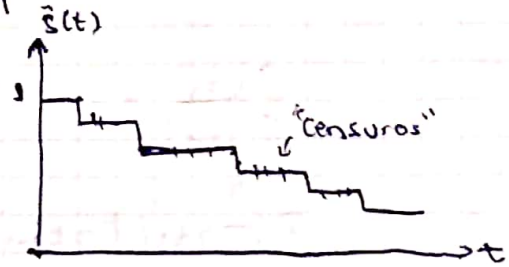
$$T = \min(T^*, C)$$

$$\delta = I(T \leq C).$$

Denote por $t_1 < t_2 < \dots < t_n$ os instantes distintos de falha.

O Estimador de Kaplan-Meier de $S(t)$ é dado por

$$\hat{S}(t) = \prod_{j:\, t_j \leq t} \frac{r_j - d_j}{r_j},$$



em que

d_j: Nº de falhas em t_j

r_j: Nº de indivíduos em risco em t_j.

Podemos escrever:

$$\hat{S}(t) = \prod_{j:\, t_j \leq t} \left(1 - \frac{dN(t_j)}{\bar{Y}(t_j)}\right),$$

"assim se escreve o est. de KM usando a notação de proc. de contagem"

em que

$$\bar{N}(t) = \sum_{i=1}^{n} N_i(t), \quad N_i(t) = I(T_i \leq t, \delta_i = 1)$$

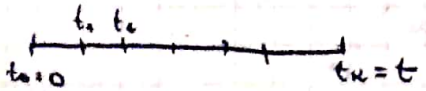
$$\bar{Y}(t) = \sum_{i=1}^{n} Y_i(t), \quad Y_i(t) = I(T_i \geq t)$$

- Representação usando integral produto.

Vimos que $S(t) = e^{-A(t)}$. Observe que podemos escrever:

$$S(t) = P(T > t_n)\, P(T > t_2 \mid T > t_1) \cdot P(T > t_3 \mid T > t_2) \dots P(T > t_k \mid T > t_{k-1})$$

em que particionamos $(0, t)$ em k subintervalos. Assim



$$S(t) = \prod_{j=1}^{k} S(t_j \mid t_{j-1}) \quad (*).$$

Relembre agora que

$$\alpha(t)\,dt \approx \mathbb{P}(t \leq T \leq t+dt \,/\, T \geq t) = \frac{\mathbb{P}(t \leq T \leq t+dt)}{\mathbb{P}(T \geq t)} \approx \frac{f(t)}{S(t^-)}$$

Podemos ainda escrever.

$$\underbrace{\alpha(t)\,dt}_{dA(t)} \approx (-dS(t)) \cdot \frac{1}{S(t^-)} \sim dA(t) \approx -\frac{1}{S(t^-)}\,dS(t) \quad \text{ou}$$

$$dS(t) \approx -S(t^-)\,dA(t)$$

Essa igualdade justifica a aproximação

$$S(t_j) - S(t_{j-1}) \approx -S(t_{j-1})\,(A(t_j) - A(t_{j-1}))$$

$$\frac{S(t_j)}{S(t_{j-1})} - 1 \approx -(A(t_j) - A(t_{j-1}))$$

$$\frac{S(t_j)}{S(t_{j-1})} = \mathbb{P}(T > t_j \,/\, T > t_{j-1}) = S(t_j \,/\, t_{j-1})$$

$$S(t_j \,/\, t_{j-1}) \approx 1 - (A(t_j) - A(t_{j-1})) \quad (**)$$

$$\approx 1 - (dA(t_j))$$

De $(*)$,

$$S(t) = \prod_{j=1}^{k} S(t_j \,/\, t_{j-1}) \approx \prod_{j=1}^{k} (1 - dA(t_j)) \quad \left(S(t) = \prod_{u \leq t} (1 - dA(u))\right)$$

Definição: Integral Produto:

Para uma função $B(t)$ (contínua à direita e c/ lim. à esq.), a Integral produto é definida por

$$\prod_{u \leq t} [1 + dB(u)] = \lim_{m \to 0} \prod_{j=1}^{k} (1 + B(t_j) - B(t_{j-1})), \text{ em que}$$

$$m = \max_j (t_j - t_{j-1}) \text{ : comprimento do maior intervalo}$$

Observe que, no caso contínuo,

$$e^{-\alpha(u)du} = 1 - \alpha(u)du \quad (\text{ex. de Taylor})$$

Então, informalmente, temos

$$S(t) = \prod_{u \leq t} (1 - dA(u)) = \prod_{u \leq t} (1 - \alpha(u)\,du)$$

$$= \lim_{m\to 0} \prod_{j=1}^{k} (1-\alpha(t_j)\,dt_j) = \lim_{m\to 0} \prod_{j=1}^{k} e^{-\alpha(t_j)dt_j}$$

$$= \lim_{m\to 0} e^{-\sum_{j=1}^{k} \alpha(t_j)dt_j} = \exp\left\{-\int_0^t \alpha(u)\,du\right\} = e^{-A(t)}$$

Em resumo:

- Caso absolutamente contínuo.

$$S(t) = \prod_{u\le t} (1-dA(u))$$

$$= \exp\left\{-\int_0^t \alpha(u)\,du\right\} = \exp\left\{-\int_0^t dA(u)\right\}$$

- Caso Discreto:

$$S(t) = \prod_{u\le t} (1-dA(u)) = \prod_{u\le t} (1-\alpha_u), \text{ onde } \alpha_u = P(T=u \,/\, T \geq u).$$

Análise de Sobrevivência

Aula 10

31/03/2016

Relação entre o estimador de KM e o estimador de Nelson Aalen.

Vimos que

$$\hat{A}(t) = \int_0^t \frac{d\bar{N}(u)}{\bar{Y}(u)} \Rightarrow d\hat{A}(t_j) = \frac{dN(t_j)}{\bar{Y}(t_j)}$$

Partindo do estimador de Nelson Aalen, podemos pensar em um estimador de $S(t)$.

$$\tilde{S}(t) = \prod_{u \leq t} (1 - d\hat{A}(u)) = \prod_{t_j \leq t} (1 - d\hat{A}(t_j)) =$$

$$= \prod_{t_j \leq t} \left(1 - \frac{d\bar{N}(t_j)}{\bar{Y}(t_j)}\right) = \hat{S}(t)$$

Estudo das propriedades de $S(t)$.

Resultado: Se $S(t) > 0$, então

$$\frac{\hat{S}(t)}{S(t)} = 1 - \int_0^t \frac{\hat{S}(s^-)}{S(s)} \left(\frac{d\bar{N}(s)}{\bar{Y}(s)} - dA(s)\right)$$

Prova: Fleming

Corolário. Para qualquer $S(t) > 0$,

$$\hat{S}(t) - S(t) = -S(t) \int_0^t \frac{\hat{S}(u^-)}{S(u)} \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} d\bar{M}(u) + B(t),$$

$$B(t) = I(T_0 < t) \cdot \frac{\hat{S}(T_0)(S(T_0) - S(t))}{S(T_0)}, \quad T_0 = \inf\{s : \bar{Y}(s) = 0\}$$

Prova:

$$\frac{\hat{S}(t)}{S(t)} = 1 - \int_0^t \frac{\hat{S}(s^-)}{S(s)} \left(\frac{d\bar{N}(s)}{\bar{Y}(s)} - dA(s)\right) \Leftrightarrow$$

$$\hat{S}(t) - S(t) = -S(t) \int_0^t \frac{\hat{S}(s^-)}{S(s)} \left(\frac{d\bar{N}(s)}{\bar{Y}(s)} - dA(s)\right)$$

$$= -S(t)\int_0^t \frac{\hat{S}(s^-)}{S(s)}\left(\frac{d\bar{N}(s)}{\bar{Y}(s)} I_{(\bar{Y}(s)>0)} - I_{(\bar{Y}(s)>0)} dA(s) - I(\bar{Y}(s)=0)\, dA(s)\right)$$

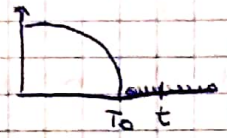
$$= -S(t)\int_0^t \frac{\hat{S}(s^-)}{S(s)} \frac{I(\bar{Y}(s)>0)}{\bar{Y}(s)}\left(d\bar{N}(s) - \bar{Y}(s)\, dA(s)\right) - \frac{\hat{S}(s^-)}{S(s)} I(\bar{Y}(s)=0)\, dA(s)$$

$\therefore$

$$\hat{S}(t) - S(t) = -S(t)\int_0^t \frac{\hat{S}(s^-)}{S(s)} \frac{I(\bar{Y}(s)>0)}{\bar{Y}(s)}\, d\bar{M}(s) + B(t),$$

$$B(t) = S(t)\int_0^t \frac{\hat{S}(u)}{S(u)} I(\bar{Y}(u)=0)\, dA(u) = S(t)\int_{T_0}^t \frac{\hat{S}(u^-)}{S(u)}\, dA(u),$$

$T_0 = \inf\{u : \bar{Y}(u) = 0\}$, se $T_0 < t$



$$B(t) = I(T_0 < t)\, \hat{S}(T_0)\frac{[S(T_0) - S(t)]}{S(T_0)}$$ (pg 98, Fleming) ■

Resultado: Seja

$$\pi(t) = \mathbb{P}(Z_i > t),\quad Z_i = \min(T_i, C_i).$$

Então

$$0 \leq E(\hat{S}(t) - S(t)) \leq (1 - S(t))\,(1 - \pi(t))^n \qquad \text{(superestima)}$$

Prova. Pelo Corolário,

$$E(\hat{S}(t) - S(t)) = \underbrace{E\left[-S(t)\int_0^t \frac{\hat{S}(u^-)}{S(u)} \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\, d\bar{M}(u)\right]}_{(I)} + \underbrace{E(B(t))}_{(II)}$$

(I) Seja

$$L(t) = \int_0^t S(t)\frac{\hat{S}(u^-)}{S(u)} \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\, \overset{\text{Martingal}}{d\bar{M}(u)}$$

$L(t)$ é um martingal de média zero $\Rightarrow$ (I) $= 0$.

(II) $$E(B(t)) = E\left[I(T_0 < t)\,\hat{S}(T_0)\frac{(S(T_0) - S(t))}{S(T_0)}\right]$$

$$\leq E\left[I(T_0 < t)\,\hat{S}(T_0)\left(1 - \frac{S(t)}{S(T_0)}\right)\right]$$

$$\leq E\left[I(T_0 < t)\, 1\left(1 - \frac{S(t)}{S(T_0)}\right)\right]$$

Observe ainda que

$$S(T_0) \leq 1 \Rightarrow \frac{1}{S(T_0)} \geq 1 \Rightarrow \frac{S(t)}{S(T_0)} \geq S(t) \Rightarrow -S(t) \geq -\frac{S(t)}{S(T_0)}$$

$$\Rightarrow 1 - S(t) \geq 1 - \frac{S(t)}{S(T_0)}$$

Então

$$E(B(t)) \leq E\left(J(T_0 < t)\left(1 - \frac{S(T_0)}{S(t)}\right)\right) \leq E(J(T_0 < t)(1 - S(t)))$$

$$\leq (1 - S(t))\, E(J(T_0 < t))$$

$$\Rightarrow E(J(T_0 < t)) = P(T_0 < t)$$

Relembre que $T_0 = \inf\{s : \bar{Y}(s) = 0\}$

$$P(T_0 < t) = \mathbb{P}(\bar{Y}(t) = 0) = \mathbb{P}(Y_1 = 0, Y_2 = 0, \dots, Y_n = 0)$$

$$= \mathbb{P}(Z_1 < t)\, \mathbb{P}(Z_2 < t) \dots \mathbb{P}(Z_n < t)$$

$$= (1 - \pi_1(t))(1 - \pi_2(t)) \dots (1 - \pi_n(t)) = (1 - \pi(t))^n$$

$$E(B(t)) \leq (1 - S(t))\, E(J(T_0 < t)) \leq (1 - S(t))(1 - \pi(t))^n$$

Observe que

$$B(t) = \begin{cases} 0, & \text{se } T_0 > t \\ > 0, & \text{se } T_0 < t. \end{cases} \Rightarrow E(B(t)) \geq 0 \qquad \blacksquare$$

Observações.

- $E(\hat{S}(t) - S(t)) \geq 0 \Rightarrow E(\hat{S}(t)) \geq S(t)$
- Quando $n \to +\infty$, se $\pi(t) > 0$,

$$E(B(t)) \to 0$$

Portanto,

$$E(\hat{S}(t)) \xrightarrow{n \to \infty} 0.$$

Teorema: Seja $t \in [0, \tau]$ e assuma que

$$-\int_0^t \frac{J(\bar{Y}(u) > 0)}{\bar{Y}(u)}\, \alpha(u) \xrightarrow{P} 0 \text{ qdo } n \to +\infty$$

$$\cdot \int_0^t (1 - I(\bar{Y}(u) > 0))\, \alpha(u)\, du \xrightarrow{P} 0 \quad \text{com } n \to +\infty$$

Então

$$\sup_{s \in [0,\tau]} |\hat{S}(s) - S(s)| \xrightarrow{P} 0. \qquad \text{Consistência Uniforme.}$$

Teorema. Seja $t \in [0, \tau]$ e assuma que

. Para todo $s \in [0, \tau]$,

$$n \int_0^s \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} \alpha(u)\, du \xrightarrow{P} \sigma^2(s) \quad \text{qdo } n \to +\infty$$

$$\cdot\, n \int_0^t \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} \alpha(u)\, I\left(\left| \sqrt{n}\, \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} \right| > \varepsilon \right) du \xrightarrow[n \to +\infty]{P} 0$$

$$\cdot \sqrt{n} \int_0^t (1 - I(\bar{Y}(u) > 0))\, \alpha(u)\, du \xrightarrow{P} 0 \quad \text{com } n \to +\infty$$

Então

$\sqrt{n}\,(\hat{S}(\cdot) - S(\cdot))$ converge para um proc. gaussiano com função de covariância

$$Cov(W(s), W(t)) = S(t) S(s) \int_0^{t \wedge s} \frac{dA(u)}{\pi(u)}$$

$t \wedge s = \min(s, t)$ e $\pi(u) = P(Y \geq u)$.

Mais precisamente,

$$\sqrt{n}\,(\hat{S}(\cdot) - S(\cdot)) \xrightarrow{D} -S(\cdot)\, U(\cdot)$$

em que

$U(\cdot)$ é um martingal gaussiano com $U(0) = 0$ e

$$Cov(U(s), U(t)) = \int_0^{t \wedge s} \frac{dA(u)}{\pi(u)}$$

Observações

. Para t_0 fixado

$$\sqrt{n}\,(\hat{S}(t_0) - S(t_0)) \xrightarrow{D} N(0, \sigma^2_{S(t_0)}),$$

$$\sigma^2_{S(t_0)} = S^2(t_0) \int_0^{t_0} \frac{dA(u)}{\pi(u)}$$

Seja $\sigma^2(t) = \int_0^t \frac{dA(u)}{\pi(u)}$

Podemos estimar $\sigma^2(t)$ por

Aula 10

$$\hat{\sigma}^2(t) = \int_0^t \frac{d\bar{N}(u)}{\bar{Y}(u)} \frac{1}{\bar{Y}(u)/n}$$

Assim,

$$\widehat{Var}(\hat{S}(t_0)) = \frac{1}{n} \hat{S}(t_0) \int_0^{t_0} \frac{d\bar{N}(u)}{\bar{Y}^2(u)/n}$$

- outro estimador (mais usado na literatura) para $Var(\hat{S}(t))$

Vimos que

$$Var(\sqrt{n}\, \hat{S}(t)) = S^2(t) \int_0^t \frac{dA(u)}{\pi(u)}$$

Observe que

$$dA(t) = \alpha(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt} S(t) \Big/ S(t)$$

$$dA(t) = -\frac{dS(t)}{S(t)}$$

Assim

$$Var(\sqrt{n}\, \hat{S}(t)) = S^2(t) \int_0^t -\frac{dS(t)}{S(t)} \frac{1}{\pi(u)}$$

e

$$\widehat{Var}(\sqrt{n}\, \hat{S}(t)) = \widehat{S^2}(t) \int_0^t -\frac{d\hat{S}(t)}{\hat{S}(t)} \frac{1}{\bar{Y}(u)/n}$$

Estimador de Greenwood

Intervalos de Confiança Ponto a Ponto

- Sem Transformação

Seja

$$\hat{\sigma}^2(t_0) = \int_0^{t_0} -\frac{d\hat{S}(u)}{\hat{S}(u)\bar{Y}(u)}$$

Vimos que

$$\widehat{Var}(\hat{S}(t_0)) = \hat{S}^2(t_0)\, \hat{\sigma}^2(t_0)$$

Assim, para t_0 fixado, $(1-\alpha)\%$ confiança, um I.C para $S(t_0)$ é dado por

$$\hat{S}(t_0) \pm z_{1-\alpha/2}\, \hat{\sigma}(t_0)\, \hat{S}(t_0)$$

Vantagem: Simplicidade

Desvantagem: LI pode ser negativo
LS " " maior que 1.

- Transformação LOG-LOG (pg 75 Fábio)

A. S.

Aula 11

05/09/16

Estimador Kaplan-Meier

$$\hat{S}(t) = \prod_{j:t_j \le t} \frac{r_j - d_j}{r_j}$$

$$= \prod_{j:t_j \le t} \left(1 - \frac{d\bar{N}(t_j)}{\bar{Y}(t_j)} \right)$$

Intervalos de Confiança para S(t)

→ Sem Transformação

$$\hat{S}(t_0) \pm z_{1-\alpha/2}\, \hat{\sigma}(t_0)\, \hat{S}(t_0),$$

$$\hat{\sigma}^2(t_0) = \int_0^{t_0} \frac{-d\hat{S}(u)}{S(u)\bar{Y}(u)}$$

Transformação Log-Log

Para t_0 fixado, defina

$$\theta_0 = \log(-\log S(t_0)) = g(S(t_0))$$

e

$$\hat{\theta}_0 = \log(-\log \hat{S}(t_0)) = g(\hat{S}(t_0))$$

$\hat{S}(\cdot)$: estimador de KM de $S(\cdot)$

Aplicamos o método delta a

$$g(\hat{S}(t_0)) = \hat{\theta}_0.$$

$$g(\hat{S}(t_0)) \approx g(S(t_0)) + g'(S(t_0))(\hat{S}(t_0) - S(t_0))$$

$$\rightarrow g(\hat{S}(t_0)) - g(S(t_0)) \approx g'(S(t_0))(\hat{S}(t_0) - S(t_0))$$

$$Var(\hat{\theta}_0) \approx (g'(S(t_0)))^2\, Var(\hat{S}(t_0))$$

observe que, neste caso,

$$g(x) = \log(-\log x) \rightarrow g' = -1/x\log x$$

$$\hat{Var}(\hat{\theta}_0) \approx \left(-\frac{1}{\hat{S}(t_0)\log\hat{S}(t_0)}\right)^2 \hat{\sigma}^2(t_0)\,\hat{S}^2(t_0)$$

$$\approx \frac{\sigma^2(t_0)}{[\log \hat{S}(t_0)]^2}$$

Assim, temos o seguinte IC. para $\theta_0 = g(S(t_0))$

$$\log(-\log \hat{S}(t_0)) \pm z_{1-\alpha/2} \frac{\hat{\sigma}(t_0)}{-\log \hat{S}(t_0)}$$

Um intervalo para $S(t_0)$ fica dado por

$$y = g(x) = \log(-\log x)$$

$$\rightarrow x = \exp\{-e^y\}$$

$$\left[\exp\left\{-\exp\left(\hat{\theta}_0 - \left(\frac{\hat{\sigma}(t_0)}{-\log\hat{S}(t_0)}\right)z_{1-\alpha/2}\right)\right\};\ \exp\left\{-\exp\left(\hat{\theta}_0 + \left(\frac{\hat{\sigma}(t_0)}{-\log\hat{S}(t_0)}\right)z_{1-\alpha/2}\right)\right\}\right]$$

ou

$$[\hat{S}(t_0)^{\theta^{-k}};\ \hat{S}(t_0)^{\theta^{k}}],\quad k = z_{1-\alpha/2}\sqrt{Var(\hat{\theta}_0)},\ \text{ou}$$

$$[\hat{S}(t_0)^{1/\theta},\ \hat{S}(t_0)^{\theta}],\quad \theta = \exp\{z_{1-\alpha/2}\sqrt{Var(\hat{\theta}_0)}\}$$

→ Borgan e Liestol

Propuseram um intervalo com transformação ARCO-SENO

$$IC(S(t);\gamma) = (LI, LS),$$

$$LI = \text{sen}^2\left[\max\left(0;\ \text{arcsen}(\hat{S}(t)^{1/2}) - (1/2)z_{1-\gamma/2}\,\hat{\sigma}_s(t)\left(\frac{\hat{S}(t)}{1-\hat{S}(t)}\right)^{1/2}\right)\right]$$

$$LS = \text{sen}^2\left[\min\left(\pi/2;\ \text{arcsen}(\hat{S}(t)^{1/2}) + 1/2\, z_{1-\gamma/2}\,\hat{\sigma}_s(t)\left(\frac{\hat{S}(t)}{1-\hat{S}(t)}\right)^{1/2}\right)\right]$$

$$\hat{\sigma}_s^2(t) = \sum_{t_j \le t} \frac{d_j}{n_j(n_j - d_j)}$$

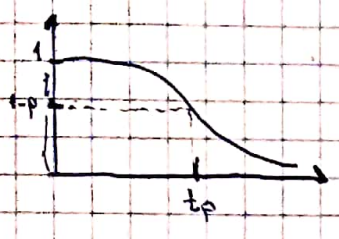
Calcular quantis!

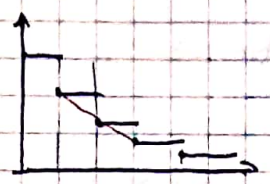
Prob. média: $\lim_{n \to \infty}$

Estimação de Percentis

Podemos definir o percentil p como sendo o valor t_p tal que

$$t_p = \inf\{t : S(t) \leq 1-p\}$$





Um estimador para t_p é dado por

$$\hat{t}_p = \inf\{t : \hat{S}(t) \leq 1-p\}$$ (Klein Moeschberger Lemeshow)

Colosimo e Gido. Devemos estimar t_p por meio de uma interpolação linear.

Instantes de falha	$\hat{S}(t)$
⋮	⋮
t_L	$\hat{S}(t_L)$
t_U	$\hat{S}(t_U)$
⋮	⋮

Encontramos t_L e t_U tais que

$\hat{S}(t_L) > (1-p)$ e

$\hat{S}(t_U) \leq (1-p)$

Resolvemos a equação

$$\frac{t_U - t_L}{\hat{S}(t_U) - \hat{S}(t_L)} = \frac{\hat{t}_p - t_L}{(1-p) - \hat{S}(t_L)}$$

Pode-se demonstrar que um estimador $\hat{t}_p$ é dado por

$$\widehat{Var}(\hat{t}_p) = \widehat{Var}(\hat{S}(\hat{t}_p)) \,/\, (\hat{f}(\hat{t}_p))^2$$

$\widehat{Var}(\hat{S}(\hat{t}_p))$: pode ser obtido pela fórmula de Greenwood.

$\hat{f}(\hat{t}_p)$: estimador da função densidade de prob.

Sugestão:

$$\hat{f}(\hat{t}_p) = (\hat{S}(a) - \hat{S}(b)) / (b-a)$$

$a = \max\{t : \hat{S}(t) \geq (1-p) + 0{,}05\}$

$b = \min\{t : \hat{S}(t) \leq (1-p) - 0{,}05\}$ (Hosmer & Lemeshow)

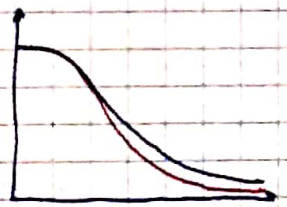
arbitrário

Teste de Log-Rank

Ideia: "comparar curvas de sobrev. dist. do tempo de falha". Procedimento não paramétrico, sem fazer suposições sobre

Comparação de Duas Populações

Queremos comparar duas funções de sobrevivência.



$S_1(t)$, $S_2(t)$

População 1: amostra aleatória (Z_{1j}, δ_{1j}), $j = 1, 2, \ldots, n_1$;

População 2: amostra aleatória (Z_{2j}, δ_{2j}), $j = 1, 2, \ldots, n_2$,

com

$Z_{ij} = \min(T_{ij}, C_{ij})$, C_{ij}: tempo de censura

$\delta_{ij} = I(T_{ij} \leq C_{ij})$ T_{ij}: " - falha $i = 1, 2$, $j = 1, 2, \ldots, n_i$

Assuma

- independência entre os tempos de falha
- censura indep.
- $T_{1j} \sim S_1(\cdot)$ e $T_{2j} \sim S_2(\cdot)$

Queremos testar as hipóteses

limitação p/ tornar possível a comparação

$H_0: S_1(t) = S_2(t) \quad \forall t \in [0, \tau]$

$H_A: S_1(t) \neq S_2(t)$ p/ algum $t \in [0, \tau]$

τ: usualmente, é considerado o maior instante de tempo no qual os dois grupos têm pelo menos uma observação em risco.

Aula

05/04

O teste de log-rank é o mais poderoso para hipóteses do tipo

$$H_0: S_1(t) = S_2(t)$$
$$H_A: S_1(t) = [S_2(t)]^c \qquad t \in [0, \tau]$$

observe que, se $S_1(t) = (S_2(t))^c$, temos

$$\alpha_1(t) = -\frac{d}{dt} S_1(t) \cdot \frac{1}{S_1(t)}$$

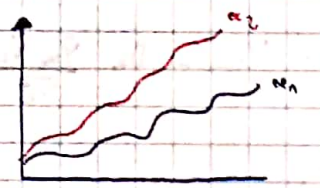
$$= -\frac{d}{dt} (S_2(t)^c) \frac{1}{(S_2(t))^c}$$

$$= -c(S_2(t))^{c-1} \frac{dS_2(t)}{dt} \frac{1}{(S_2(t))^c}$$

$$\alpha_1(t) = c \cdot \underbrace{-\frac{d}{dt} S_2(t) \frac{1}{S_2(t)}}_{\alpha_2(t)}$$

$\alpha_1(t) = c\,\alpha_2(t)$ ou

$\frac{\alpha_1(t)}{\alpha_2(t)} = c$, que não depende de t



Pop 1 e Pop 2 tem "riscos proporcionais".

Construção do Teste

Seja $n = n_1 + n_2$ e considere $T_1^0 < T_2^0 < \dots < T_L^0$: instantes em que alguma falha ocorreu na amostra combinada.

Para cada instante T_k^0, $k = 1, 2, \dots, L$, defina:

D_{ik}: nº de falhas observadas em T_k^0 na amostra i

Y_{ik}: nº de indivíduos em risco na amostra i no instante T_k^0, $i = 1, 2$

Para cada instante de falha T_k^0, podemos construir uma tabela

Eventos	Grupo 1	Grupo 2	Total
Falhas	D_{1k}	D_{2k}	D_k
Sobrevi.	$Y_{1k} - D_{1k}$	$Y_{2k} - D_{2k}$	$Y_k - D_k$
Em risco	Y_{1k}	Y_{2k}	Y_k

Note que

- Dado $\bar{Y}_{1k}$, temos $D_{1k} / \bar{Y}_{1k} \sim$ binomial $(\bar{Y}_{1k}, p_{1k})$

- Dado $\bar{Y}_{2k}$,

$$D_{2k} / \bar{Y}_{2k} \sim \text{binomial}(\bar{Y}_{2k}, p_{2k})$$

p_{ik}: prob. de óbito no grupo i e em T_k^0.

Se H_0 é verdadeira,

$$p_{1k} = p_{2k}.$$

Qual a distribuição de

$$D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k \overset{H_0}{\sim} \text{Hipergeométrica}$$

Da dist. hipergeométrica, temos que

- $E(D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k) = E_{1k} = e_{1k} = \frac{\bar{Y}_{1k}}{\bar{Y}_k} D_k = \frac{D_k}{\bar{Y}_k} \cdot \bar{Y}_{1k}$

- $Var(D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k) = V_{1k} = D_k \frac{\bar{Y}_{1k} \bar{Y}_{2k}}{\bar{Y}_k^2} \frac{(\bar{Y}_k - D_k)}{(\bar{Y}_k - 1)}$

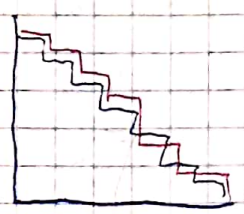
O teste de log-rank é baseado na ideia de comparar o observado com o esperado se H_0 é verdadeira.

Assumindo independência entre os tempos de falha, temos a estatística do teste é

$$LR = \left[\sum_{j=1}^{L} (D_{1j} - E_{1j})\right]^2 \Big/ \left[\sum_{j=1}^{L} V_{1j}\right]$$

ideal quando riscos proporcionais

$$LR \xrightarrow[n \to \infty]{D} \chi^2_{(1)}$$



cruzam: grave indica risco não prop.

Observação:

Note que a estatística é baseada no quadrado da soma das diferenças. Para T_k^0, observo ainda:

$$D_{1k} - E_{1k} = D_{1k} - D_k \frac{\bar{Y}_{1k}}{\bar{Y}_k} = \frac{\bar{Y}_{1k} \bar{Y}_{2k}}{\bar{Y}_k} \left(\left(\frac{D_{1k}}{\bar{Y}_{1k}}\right) - \left(\frac{D_{2k}}{\bar{Y}_{2k}}\right) \right)$$

Exemplo.

Grupo 1: 3 5 7 9$^+$ 18

Grupo 2. 8$^+$ 12 19 20 20$^+$

Tempos de falha. 3 5 7 12 18 19 20

$T_1^0 = 3$

	G1	G2	Total
F	1	0	1
S	4	5	9
Em Risco	5	5	10

$T_2^0 = 5$

	G1	G2	Total
F	1	0	1
S	3	5	8
Em risco	4	5	9

⋮

Em suma,

T_k^0	Y_{1k}	D_{1k}	$\bar{Y}_{1k}$	D_{1k}	E_{1k}	$V_{1,k}$
3	10	1	5	1	0,5	0,25
5	9	1	4	1	0,444	0,247
7	8	1	3	1	0,375	0,239
12	5	1	1	0	0,2	0,160
18	4	1	1	1	0,25	0,183
19	3	1	0	0	0	0
20	2	1	0	0	0	0

$$LR = \left(\sum_{j=1}^{L} (D_{ik} - E_{ik})\right)^2 / \sum V_{ik}$$

$$= \left(\sum_{j=1}^{L} D_{ik} - \sum_{j=1}^{L} E_{ik}\right)^2 / \sum V_{ik}$$

$$= (\text{soma } ① - \text{soma } ②)^2 / \text{soma } ③$$

$$= (4 - 1{,}769)^2 / 1{,}0787 = 4{,}612$$

$\chi^2_{(1)} = 3{,}84 \quad (\alpha = 5\%)$

$\Rightarrow$ então há evidencias para rejeição de H_0.

Análise de Sobrevivência

Aula 12

07/04/16

Teste de Log Rank

$$\begin{cases} H_0 \quad S_1(t) = S_2(t) \quad , \quad 0 \leq t \leq \tau \\ H_A \quad \text{existe alguma dif.} \end{cases}$$

$$LR = \left[\sum_{j=1}^{L} (D_{1j} - E_{1j}) \right]^2 \Big/ \left[\sum_{j=1}^{L} V_{1j} \right] \xrightarrow{D} \chi_{(1)}$$

$T_1^\circ, T_2^\circ, \ldots, T_L^\circ$ instantes distintos de falha

Generalizações

- Inclusão de ponderações

$$WLR = \left[\sum_{j=1}^{L} w_j (D_{1j} - E_{1j}) \right]^2 \Big/ \left[\sum_{j=1}^{L} w_j^2 V_{1j} \right] \xrightarrow{D} \chi^2_{(1)} \quad , \; w_j \text{ pesos}$$

Testes Mais Utilizados

Peso	Nome
$W_j = 1$	Log-Rank
$W_j = \bar{Y}_j$	Gehan ou Wilcoxon [dá mais peso pl o começo com mais intensidade]
$W_j = \sqrt{\bar{Y}_j}$	Tarone-Ware
$W_j = (\hat{S}(t_{j-1}))^p (1-\hat{S}(t_{j-1}))^q$	Família Fleming Harrington [Não precisamos $p+q=1$. $\rightarrow 0 \leq p, q \leq 1$.]

$\uparrow$ peso no começo $\qquad \uparrow$ peso na cauda

$$W_j = \frac{\bar{Y}_j}{\bar{Y}_{j+1}} \tilde{S}(t_{j-1}) \qquad \text{Peto-Prentice} \quad , \; \tilde{S}(t) = \prod_{t_k \leq t} \left(1 - \frac{d_k}{\bar{Y}_{k+1}} \right)$$

→ $\hat{S}(t)$ e $\tilde{S}(t)$ são obtidas da amostra combinada.

Formulação com Processos de Contagem

Seja...

$$N_{ij}(t) = I(Z_{ij} \leq t, \delta_{ij}), \quad i = 1,2, \quad j = 1,2,\dots,n_i$$

$$Y_{ij}(t) = I(Z_{ij} > t)$$

$$\bar{Y}_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t), \quad i = 1,2$$

Relembre que

$$LR = \left(\sum_{k=1}^{L} (D_{1k} - E_{1k})\right)^2 \Big/ \left[\sum_{k=1}^{L} V_k\right] = \frac{(Num)^2}{Den}$$

$$V_{1k} = D_k \frac{\bar{Y}_{1k}\bar{Y}_{2k}}{\bar{Y}_k^2} \cdot \frac{(\bar{Y}_k - D_k)}{(\bar{Y}_k - 1)}$$

$$E_{1k} = D_k \frac{\bar{Y}_{1k}}{\bar{Y}_{1k} + \bar{Y}_{2k}}$$

$T_1^0 < T_2^0 < \dots < T_L^0$: instantes ordenados de falha

Seja $Num = \sum_{k=1}^{L} (D_{1k} - E_{1k}) = \sum_{k=1}^{L} D_{1k} - \sum_{k=1}^{L} E_{1k}$

Observa-se que

$$\sum_{k=1}^{L} D_{1k} = \sum_{j=1}^{n_1} \int_0^{\tau} dN_{1j}(s)$$

$$\sum_{k=1}^{L} E_{1k} = \sum_{k=1}^{L} D_k \frac{\bar{Y}_{1k}}{\bar{Y}_{1k} + \bar{Y}_{2k}} = \sum_{k=1}^{L} (D_{1k} + D_{2k}) \frac{\bar{Y}_{1k}}{\bar{Y}_{1k} + \bar{Y}_{2k}} =$$

$$= \sum_{k=1}^{L} D_{1k} \frac{\bar{Y}_{1k}}{\bar{Y}_{1k} + \bar{Y}_{2k}} + \sum_{k=1}^{L} D_{2k} \frac{\bar{Y}_{1k}}{\bar{Y}_{1k} + \bar{Y}_{2k}}$$

$$\sum_{k=1}^{L} E_{1k} = \sum_{j=1}^{n_1} \int_0^{\tau} \frac{\bar{Y}_1(u)}{\bar{Y}_1(u) + \bar{Y}_2(u)} dN_{1j}(u) + \sum_{j=1}^{n_2} \int_0^{\tau} \frac{\bar{Y}_1(u)}{\bar{Y}_1(u) + \bar{Y}_2(u)} dN_{2j}(u)$$

$$\bar{Y}(u) = \bar{Y}_1(u) + \bar{Y}_2(u)$$

Portanto,

$$Num = \sum_{j=1}^{n_1} \int_0^{\tau} dN_{1j}(u) - \sum_{j=1}^{n_1} \int_0^{\tau} \frac{\bar{Y}_1(u)}{\bar{Y}(u)} dN_{1j}(u) - \sum_{j=1}^{n_2} \int_0^{\tau} \frac{\bar{Y}_1(u)}{\bar{Y}(u)} dN_{2j}(u)$$

$$Num = \sum_{j=1}^{n_1} \int_0^{\tau} \frac{\bar{Y}_1(u) + \bar{Y}_2(u) - \bar{Y}_1(u)}{\bar{Y}(u)} \, dN_{1j}(u) - \sum_{j=1}^{n_2} \int_0^{\tau} \frac{\bar{Y}_1(u)}{\bar{Y}(u)} \, dN_{2j}(u)$$

Pela Decomposição de Doob-Meier

$$M_{1j}(t) = N_{1j}(t) - \int_0^{t} Y_{1j}(u)\,\alpha_1(u)\,du$$

$$M_{2j}(t) = N_{2j}(t) - \int_0^{t} Y_{2j}(u)\,\alpha_2(u)\,du$$

$$Num = \sum_{j=1}^{n_1} \int_0^{\tau} \overset{W(u)}{\cdots} \frac{\bar{Y}_2(u)}{\bar{Y}(u)} \, dM_{1j}(u) - \sum_{j=1}^{n_2} \int_0^{\tau} \overset{W(u)}{-} \frac{\bar{Y}_1(u)}{\bar{Y}(u)} \, dM_{2j}(u)$$

$$- \int_0^{\tau} \overset{W(u)}{+} \frac{\bar{Y}_2(u)}{\bar{Y}(u)} \overset{\bar{Y}_1(u)}{\left(\sum_{j=1}^{n_1} Y_{1j}(u) \right)} \alpha_1(u)\,du + \int_0^{\tau} \overset{W(u)}{-} \frac{\bar{Y}_1(u)}{\bar{Y}(u)} \underbrace{\left(\sum_{j=1}^{n_2} Y_{2j}(u) \right)}_{\bar{Y}_2(u)} \alpha_2(u)\,du \quad \text{sob } H_0$$

Sob H_0,

$$\alpha_1(t) = \alpha_2(t) \quad \forall t \in [0, \tau]$$

e os últimos termos são iguais

$$Num = \sum_{j=1}^{n_1} \int_0^{\tau} \left(\frac{\bar{Y}_2(u)}{\bar{Y}(u)} \right)^{\overset{\text{proc. previsível}}{W(u)}} dM_{1j}(u) - \sum_{j=1}^{n_2} \int_0^{\tau} \left(\frac{\bar{Y}_1(u)}{\bar{Y}(u)} \right)^{\overset{\text{proc. prev.}}{W(u)}} dM_{2j}(u)$$

pode por expo pero desde que seja previsível

→ Num é um martingal de média zero:

$$\mathbb{E}(Num) = 0 \quad \text{sob } H_0$$

Poderse demonstrar que o processo de variação previsível de NUM é dado por

$$\langle Num \rangle(t) = \sum_{j=1}^{n_1} \int_0^{t} W^2(u) \frac{\bar{Y}_2(u)}{\bar{Y}^2(u)} \bar{Y}_1(u)\,\alpha_1(u)\,du$$

$$+ \sum_{j=1}^{n_2} \int_0^{t} W^2(u) \frac{\bar{Y}_1(u)}{\bar{Y}^2(u)} \bar{Y}_{2j}(u)\,\alpha_2(u)\,du$$

Sob H_0, $\alpha_1(t) = \alpha_2(t) = \alpha(t)$,

$$\langle Num \rangle(t) = \int_0^{\tau} W^2(u) \frac{\bar{Y}_1(u)\,\bar{Y}_2(u)}{\bar{Y}^2(u)} \alpha(u)\,du$$

Por resultados anteriores

$$Var(Num) = E(\langle (Num) \rangle)$$

Um estimador para $Var(Num)$ é

$$\widehat{Var}(Num) = \int_0^{\tau} W^2(u) \frac{\bar{Y}_1(u)\bar{Y}_2(u)}{\bar{Y}(u)} \, d\hat{A}(u)$$

estimador de Nelson-Aalen de $A(t)$ da amostra combinada.

$$= \int_0^{\tau} W^2(u) \frac{\bar{Y}_1(u)\bar{Y}_2(u)}{\bar{Y}(u)} \left(\frac{d\bar{N}(u)}{\bar{Y}(u)} \right)$$

$$= \int_0^{\tau} W^2(u) \frac{\bar{Y}_1(u)\bar{Y}_2(u)}{\bar{Y}^2(u)} \, d\bar{N}(u)$$

Vimos que

$$V_{1k} = D_k \frac{\bar{Y}_{1k}\bar{Y}_{2k}}{\bar{Y}_k^2} \cdot \frac{(\bar{Y}_k - D_k)}{(\bar{Y}_k - 1)}$$

Se não há empates,

$$V_{1k} = \frac{\bar{Y}_{1k}\bar{Y}_{2k}}{\bar{Y}_k^2}.$$

Com ponderação $V_{1k} = W_k^2 \frac{\bar{Y}_{1k}\bar{Y}_{2k}}{\bar{Y}_k^2}$.

Ainda, no denominador da est. de teste WLR

$$\sum_{k=1}^{L} W_k^2 V_{1k} = \sum_{k=1}^{L} W_k^2 \frac{\bar{Y}_{1k}\bar{Y}_{2k}}{\bar{Y}_k^2}$$

Resultado: Se existe $y_1(t)$ e $y_2(t)$ tais que

$$\frac{\bar{Y}_1(t)}{n} \xrightarrow{P} y_1(t) \quad e \quad \frac{\bar{Y}_2(t)}{n} \xrightarrow{P} y_2(t)$$

$y_1(t)$ e $y_2(t)$ funções determinísticas, então

$$WLR \xrightarrow{D} \chi^2_{(1)} \quad \text{(Aalen, livro)}$$

Consistência

Definição: Seja Q_n, $n = 1, 2, \ldots$ uma seq. de est. de teste p/ testar a hip. H_0.
Sejam R_n, $n = 1, 2, \ldots$ as regiões de rejeição associadas a Q_n com um nível designif. α. A seq. Q_n é consistente para uma hipótese alternativa H_A se

$$\lim_{n \to \infty} P(Q_n \in R_n) = 1 \text{ se } H_A \text{ é verdadeira} \quad \text{(Fleming, pg 266)}$$

Analise de Sobrevivência

Aula 13

12/04/2016

$H_0: S_1(t) = S_2(t)$

H_A: Existe alguma diferença

Est. Log-Rank: $$WLR = \left[\sum_{j=1}^{k} W_j (D_{1j} - Z_{1j})^2\right] \Big/ \left(\sum_{j=1}^{k} W_j^2 V_{1j}\right) \xrightarrow{D} \chi^2_{(1)}$$

Consistência

Q_n, $n = 1, 2, 3, \ldots$ sequência de est. de teste.

R_n: regiões de rejeição de H_0

G_n é consistente se

$$\lim_{n \to +\infty} P(Q_n \in R_n) = 1 \text{ se } H_A \text{ é verdadeira}$$

Consistência do Teste de LR

$$LR = \frac{(NUM)^2}{Var(NUM)} = \frac{1/n\,(NUM)^2}{1/n\,Var(NUM)} = \frac{(NUM/\sqrt{n})^2}{Var(NUM/\sqrt{n})}$$

Podemos escrever

$$\frac{NUM}{\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{j=1}^{n_1} \int_0^{\tau} \frac{\bar{Y}_2(t)}{\bar{Y}(t)}\, dM_{1j}(t) - \frac{1}{\sqrt{n}} \sum_{j=1}^{n_2} \frac{\bar{Y}_1(t)}{\bar{Y}(t)}\, dM_{2j}(t) =$$

$$\frac{1}{\sqrt{n}} NUM = \ldots = \ldots = \underbrace{\frac{1}{\sqrt{n}} \int_0^{\tau} \frac{\bar{Y}_1(t)\,\bar{Y}_2(t)}{\bar{Y}(t)} (\alpha_1(t) - \alpha_2(t))\, dt}_{(*)}$$

Se H_A é verdadeira, $(*)$ não é zero. Temos ainda que

$$(*) = \frac{n_1}{n} \cdot \frac{n_2}{n} \sqrt{n} \int_0^{\tau} \frac{\bar{Y}_1(t)/n_1 \;\; \bar{Y}_2(t)/n_2}{\bar{Y}(t)/n} (\alpha_1(t) - \alpha_2(t))\, dt$$

Vimos que $(*) \to +\infty$ se

Pag 266, Fleming

$\alpha_1(t) \geq \alpha_2(t)$ com desigualdade estrita para algum t (ou $\alpha_1(t) \leq \alpha_2(t)$ com desigualdade estrita)

Generalização Para K Grupos

Suponha agora que queremos comparar G grupos. Para cada grupo, seja $S_g(t)$ a função de sobrevivência do grupo g, $g = 1, \dots, G$.

Queremos testar as hipóteses

$$H_0: S_1(t) = S_2(t) = \dots = S_G(t), \ \forall t \leq \tau$$

H_1: Existe alguma diferença

Sejam $T^*_{(1)}, T^*_{(2)}, \dots, T^*_{(L)}$ instantes de falha ordenados da amostra combinada.

Para cada instante de falha podemos obter

$T^*_{(k)}$

	1	2	3	...	G	Total
Falhas	D_{1k}	D_{2k}	D_{3k}		D_{Gk}	D_k
Sobrev.	$\bar{Y}_{1k} - D_{1k}$	$\bar{Y}_{2k} - D_{2k}$	$\bar{Y}_{3k} - D_{3k}$	...	$\bar{Y}_{Gk} - D_{Gk}$	$\bar{Y}_k - D_k$
Em risco	$\bar{Y}_{1k}$	$\bar{Y}_{2k}$	$\bar{Y}_{3k}$	...	$\bar{Y}_{Gk}$	$\bar{Y}_k$

Fixados os marginais, pode-se trabalhar com a distribuição de $(D_{1k}, \dots, D_{(G-1)k})$ Hipergeométrica Multivariada.

O Valor esperado de D_{gk} é:

$$e_{gk} = D_k \frac{\bar{Y}_{gk}}{\bar{Y}_k}, \quad g = 1, 2, \dots, G; \ k = 1, 2, \dots, L$$

Seja

$d_k = (D_{1k}, D_{2k}, \dots, D_{(G-1)k})$: vetor $(G-1) \times 1$

$e_k = (e_{1k}, e_{2k}, \dots, e_{(G-1)k})$: vetor $(G-1) \times 1$

A variância de D_{gk} é:

$$(V_k)_{gg} = \frac{\bar{Y}_{gk}(\bar{Y}_k - \bar{Y}_{gk}) D_k (\bar{Y}_k - D_k)}{\bar{Y}_k^2 (\bar{Y}_k - 1)} \quad , g = 1, \dots, G$$

e a covariância de D_{gk} e D_{hk} é

$$(V_k)_{gh} = -\frac{\bar{Y}_{gk} \bar{Y}_{hk} D_k (\bar{Y}_k - D_k)}{\bar{Y}_k^2 (\bar{Y}_k - 1)}$$

$$V_k = \begin{pmatrix} (V_k)_{11} & (V_k)_{12} & \cdots & (V_k)_{1(G-1)} \\ \vdots & (V_k)_{22} & \cdots & (V_k)_{2(G-1)} \\ \vdots & & \ddots & \\ (V_k)_{(G-1)1} & & & (V_k)_{(G-1)(G-1)} \end{pmatrix}$$

Defina agora

$$V = \sum_{k=1}^{L} V_k .$$

Seja $\underset{\sim}{w}_k = (w_{1(k)}, w_{2(k)}, \dots, w_{(G-1)k})$ vetor de pesos.

$W_k = \text{diag}(\underset{\sim}{W}_k)$ = matriz diagonal com pesos na diagonal

A estatística de teste é dada por

$$WLR = \left[\sum_{k=1}^{L} W_k (\underset{\sim}{d}_k - \underset{\sim}{e}_k) \right]^T \times \left[\sum_{k=1}^{L} W_k V_k W_k \right]^{-1} \times \left[\sum_{k=1}^{L} W_k (\underset{\sim}{d}_k - \underset{\sim}{e}_k) \right]$$

Sob H_0, $WLR \xrightarrow{D} \chi^2_{(G-1)}$.

Teste de Log Rank Estratificado

Suponha que você tenha uma outra covariável observada nos dados. Considere que essa variável tem M níveis ou categorias.

Suponha que o interesse está em testar a igualdade das funções de sobrevivência dos dois grupos (sexo, por exemplo) dentro de cada nível da nova variável (trat., por exemplo).

Assim queremos testar

$$H_0: S_1^{(1)}(t) = S_2^{(1)}(t)$$
$$S_1^{(2)}(t) = S_2^{(2)}(t) \qquad \forall t \leq \tau$$
$$S_1^{(3)}(t) = S_2^{(3)}(t)$$
$$\vdots$$
$$S_1^{(M)}(t) = S_2^{(M)}(t)$$

subíndice: categoria do tratamento

superíndice: sexo, ...

ou $S_1^{(s)}(t) = S_2^{(s)}(t)$, $s = 1, 2, \ldots, M$ $\forall t \leq \tau$.

H_a: Existe alguma diferença

Relembre que

$$LR = \frac{(N \cup M)^2}{\sum V_{1k}} = \frac{\left(\sum w_k (D_{1k} - E_{1k})\right)^2}{\left(\sum w_k^2 V_{1k}\right)}$$

Podemos usar

$$U / \tilde{\sigma}_{LR} \xrightarrow{D} N(0,1)$$

No teste estratificado, calculamos

- U para cada estrato $\rightarrow U_s$
- $\tilde{\sigma}^2_{LR}$ para cada estrato $\rightarrow \tilde{\sigma}^2_s$

A estatística é

$$T = \sum_{s=1}^{M} U_s \Big/ \left(\sum_{s=1}^{M} \tilde{\sigma}_s^2 \right)^{1/2} \xrightarrow{D} N(0,1) \quad \text{sob } H_0.$$

* Teste de Renyi

Klein, pg. 233

Aula 13

É uma alternativa quando há cruzamento das funções de risco. Suponha que queremos testar

$$H_0: S_1(t) = S_2(t) \quad , t \in [0, \tau].$$

Ideia: $T_1^\circ < T_2^\circ < \dots < T_L^\circ$: instantes de falha ordenados da amostra combinada.

Calculamos para cada T_k°, $k = 1, 2, \dots, L$.

$$Z(T_k^\circ) = \sum_{j: T_j^\circ \leq T_k^\circ} \left(D_{1j} - D_j \frac{Y_{1j}}{Y_j} \right), \quad k = 1, 2, \dots, L,$$

e obtemos

$$Z(T_1^\circ), Z(T_2^\circ), Z(T_3^\circ), \dots, Z(T_L^\circ)$$

Seja $\sigma^2(\tau) = \sum_{k=1}^{L} V_{1k}$: denominador da estatística de LR.

A estatística de teste é dada por

$$Q = [\sup\{|Z(t)|, t \leq \tau\}] / \sigma(\tau).$$

Sob H_0, pode-se demonstrar que Q tem a mesma distribuição de

$$\sup\{|B(x)|, 0 \leq x \leq 1\}$$

$B(\cdot)$: mov. browniano padrão

Rejeita-se H_0 p/ valores grandes de Q.

- Exemplo Computacional
  - Usa library ~~surv Misc~~ survMisc. ! require (Km.ci)) bandas de confiança
  - KMSurv e GGKM. – KM mais bonitinho

surdiff → log-rank

Pwr GSD, com outros pesos

Teste Renyi – no survMisc.

Análise de Sobrevivência

Aula 14

14/04/16

Modelos de Regressão Paramétricos

Assuma que temos uma amostra aleatória de $(Z_i, \delta_i, \underset{\sim}{X}_i)$, com

$$Z_i = \min(T_i, C_i)$$

$$\delta_i = I(T_i < C_i)$$

$\underset{\sim}{X}_i$ = vetor de p variáveis,

"assuma exist. do intercepto"

$$\underset{\sim}{X}_i = (1, x_{1i}, x_{2i}, \dots, x_{(p-1)i})^T$$

Assumimos que

$$T_i / \underset{\sim}{X}_i \sim f(.)$$

$$S(.)$$

$$\alpha(.)$$

"Assumir que $T_i / \underset{\sim}{X}_i$ tem uma dada distribuição com forma conhecida que dependem apenas de $\underset{\sim}{X}_i$"

$f(.)$, $S(.)$ e $\alpha(.)$ dependem do vetor de parâmetros desconhecidos.

* Modelo de Regressão Exponencial

Relembre que

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

Podemos ter o modelo

$$f_i(t) = \lambda_i e^{-\lambda_i t}, \quad \lambda_i = e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}}, \text{ i.e.}$$

$$f_i(t) = (e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}}) e^{-(e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}})t}, \quad \underset{\sim}{\beta} : \text{vetor de parâmetros desconhecidos}$$

"Historinha: $y_i = X_i'\beta + e_i$, $e_i \sim N(0, \sigma^2)$

$\Rightarrow Y_i \sim N(X_i'\beta, \sigma^2)$

$Y_i \sim f.\exp$ MLG

$EY_i = \mu_i = g^{-1}(X_i'\beta)$ $\begin{cases} Y_i \sim N(.) \\ E(Y_i) = \mu_i = X_i'\beta \\ Var(Y_i) = \sigma^2 \end{cases}$ "

Observe que, neste caso, a função de risco é dada por

$$\alpha(t / X_i) = e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}}$$

Note ainda que, para indivíduos diferentes i e j, temos

$$\frac{\alpha(t/\underset{\sim}{X}_i)}{\alpha(t/\underset{\sim}{X}_j)} = e^{(\underset{\sim}{X}_i - \underset{\sim}{X}_j)\underset{\sim}{\beta}}$$

"Interpretação dos parâmetros por aqui. Um acréscimo de uma unidade na covariável x_j implica no aumento de e^{β_j} no risco"

que não depende de t.

⇒ Dizemos que é um modelo de riscos proporcionais

Outra representação do modelo Exponencial

Seja ξ_i uma v.a. com distribuição exponencial (1).

$$f_\xi(\xi_i) = e^{-\xi_i}.$$

Podemos escrever

$$T_i = \frac{\xi_i}{\lambda_i} = e^{-\underset{\sim}{X}_i'\underset{\sim}{\beta}}\,\xi_i$$

Observe que se $T = \xi/\lambda$, i.e., $\xi = \lambda T$,

$$f(t) = f_\xi(\xi(t)) \left| \frac{\partial \xi(t)}{\partial t} \right|$$

$$= e^{-t\lambda} \cdot \lambda = \lambda e^{-\lambda t}.$$

⇒ $T_i \sim$ exponencial (λ_i)

$$\lambda_i = e^{\underset{\sim}{X}_i'\underset{\sim}{\beta}}.$$

Defina agora

$$Y_i = \ln T_i$$

$$= \ln\left(\frac{\xi_i}{\lambda_i}\right)$$

$$= -\log \lambda_i + \ln \xi_i$$

$$= -\underset{\sim}{X}_i'\underset{\sim}{\beta} + \ln \xi_i = -\underset{\sim}{X}_i'\underset{\sim}{\beta} + w_i, \quad w_i = \ln \xi_i$$

Como $\varepsilon_i \sim \exp(1)$,

$$S_{\varepsilon_i}(\varepsilon) = P(\varepsilon_i > \varepsilon) = e^{-\varepsilon}.$$

Temos portanto que

$$S_W(w) = P(W_i > w) = P(\log \varepsilon_i > w_i) = P(\varepsilon_i > e^{w_i}).$$

$$S_W(w) = \exp\{-e^{w_i}\}$$

$$f_W(w) = e^{w_i} \exp\{-e^{w_i}\}, \text{ i.e.}$$

$W_i \sim$ Valor extremo padrão.

Assim as seguintes formulações são equivalentes.

(1) • $T_i \mid \underset{\sim}{X}_i \sim$ exponencial $(e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}})$

(2) • $Y_i = \log T_i = \underset{\sim}{X}_i' \underset{\sim}{\gamma} + W_i$, $W_i \sim$ valor extremo padrão

Observa que $\underset{\sim}{\gamma} = -\underset{\sim}{\beta}$ nesse caso

"R - SurvReg solta $\underset{\sim}{\gamma}$, isto é, considera a formulação (2). Nesse caso, a Dist. da pos. não é direta. Há que ter cuidados"

A função de sobrevivência de Y_i ficou dada por

$$\begin{aligned} S(y \mid x_i) &= P(Y > y \mid X_i) \\ &= P(\underset{\sim}{X}_i' \underset{\sim}{\gamma} + W_i > y) \\ &= P(W_i > y - \underset{\sim}{X}_i' \underset{\sim}{\gamma}) \\ &= \exp\left\{-e^{y - \underset{\sim}{X}_i' \underset{\sim}{\gamma}}\right\} \end{aligned}$$

$$f(y \mid \underset{\sim}{X}_i) = e^{y - \underset{\sim}{X}_i' \underset{\sim}{\gamma}} \exp\left\{-e^{y - \underset{\sim}{X}_i' \underset{\sim}{\gamma}}\right\}$$

Definição: Um modelo de sobrevivência é dito um modelo de vida acelerado se

$$S(t \mid \underset{\sim}{X}) = S_0(\psi_x t) \quad \forall t, \psi_x > 0$$

em que

$S_0(\ )$ é uma função de sobrevivencia basal, que não depende de X e

ψ_x é conhecido como "fator de aceleração"

No caso do modelo exponencial, observe que

$$
\begin{aligned}
S(t/\underline{X}_i) &= P(T_i > t / \underline{X}_i) \\
&= P(Y_i > \ln t / \underline{X}_i) \\
&= P(\underline{X}_i \gamma + w_i > \ln t / \underline{X}_i) \\
&= P(w_i > \ln t - \underline{X}_i \gamma / \underline{X}_i) \\
&= P(e^{w_i} > t e^{-\underline{X}_i \gamma} / \underline{X}_i) \\
&= P(\varepsilon_i > t e^{-\underline{X}_i \gamma})
\end{aligned}
$$

$$S(t/\underline{X}_i) = e^{-t e^{-\underline{X}_i \gamma}}$$

Defina

$$S_0(t) = e^{-t}$$

Observe que

$$S(t/\underline{X}_i) = e^{-t e^{-\underline{X}_i \gamma}} = S_0(t e^{-\underline{X}_i \gamma})$$

Neste caso,

$$\psi_x = e^{-\underline{X}_i \gamma}$$

$$S(t/\underline{X}_i) = S_0(\psi_x t)$$

$\Rightarrow$ O modelo exponencial é um modelo de vida acelerado.

Exemplo:

Considere uma situação simples, com uma covariável apenas, indicadora de tratamento

$$X = \begin{cases} 0, & \text{se grupo A} \\ 1, & \text{" " B} \end{cases}$$

Temos que:

• Se $X_i = 0 \Rightarrow$

$$Y_i = \ln T_i = \gamma_0 + \gamma_1 \cdot 0 + W_i$$
$$= \gamma_0 + W_i$$

• Se $X_i = 1 \Rightarrow$

$$Y_i = \ln T_i = \gamma_0 + \gamma_1 + W_i$$

$$S_A(t) = P(T > t \mid X = 0) = P(e^{\gamma_0 + W_i} > t)$$
$$= P(\xi_i > t) = S_0(t) \quad , \quad \xi_i = e^{\gamma_0 + W_i}$$

• $S_B(t) = P(T > t \mid X = 1)$

$$= P(e^{\gamma_0 + \gamma_1 + W_i} > t \mid X = 1)$$
$$= P(e^{\gamma_0 + W_i} > e^{-\gamma_1} t \mid X = 1)$$
$$= P(\xi_1 > e^{-\gamma_1} t)$$

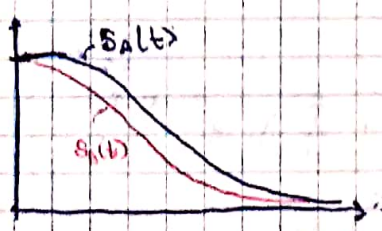
$\xi_1 = e^{\gamma_0 + W_i}$.

$$S_B(t) = S_0(e^{-\gamma_1} t) =$$

$$S_B(t) = S_A(\underbrace{e^{-\gamma_1}}_{\psi} t)$$

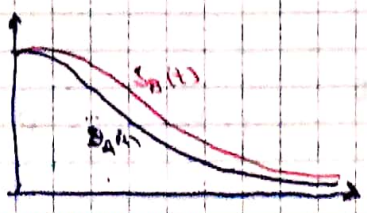
$$S_B(t) = S_A(\psi t) \quad , \quad \psi = e^{-\gamma_1}.$$

~~O grupo~~ • Se $\psi > 1$, $e^{-\gamma_1} > 1$,



O grupo se comporta como o "futuro" do grupo A.

• Se $\psi < 1$



O grupo se comporta como o "~~futuro~~ passado" do grupo A.

Estimação no Modelo de Regressão Exponencial

Normalmente, a estimação é feita considerando-se o modelo

$$Y_i^* = \ln(T_i) = \underset{\sim}{X}_i' \underset{\sim}{\gamma} + W_i \quad , \quad W_i \sim \text{valor extremo padrão}$$

Observamos

$$Y_i = \ln(\min(T_i, C_i))$$

$$\delta_i = I(T_i \leq C_i)$$

$\underset{\sim}{X}_i$ = vetor de covariáveis (incluindo o intercepto)

Para o caso de dados com censura à direita, a função de verossimilhança fica dada por

$$L(\underset{\sim}{\gamma}) = \prod_{i=1}^{n} (f(y_i / \underset{\sim}{X}_i))^{\delta_i} (S(y_i / \underset{\sim}{X}_i))^{1-\delta_i}$$

$$\overset{(*)}{=} \prod_{i=1}^{n} \left[\exp\left\{(y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}) + e^{-y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}}\right\}\right]^{\delta_i} \cdot \left[\exp\left\{-e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}}\right\}\right]^{1-\delta_i}$$

$$L(\underset{\sim}{\gamma}) = \prod_{i=1}^{n} (e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}})^{\delta_i} \exp\{-e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}}\}$$

$$= e^{\sum_{i=1}^{n} \delta_i (y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma})} \exp\left\{-\sum_{i=1}^{n} e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}}\right\}$$

A função de log-verossimilhança é dada por

$$\ell(\underset{\sim}{\gamma}) = \sum_{i=1}^{n} \delta_i (y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}) - \sum_{i=1}^{n} (e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}})$$

Observe que

$$U(\underset{\sim}{\gamma}) = \frac{\partial \ell(\underset{\sim}{\gamma})}{\partial \underset{\sim}{\gamma}} = -\sum_{i=1}^{n} \delta_i X_i + \sum_{i=1}^{n} \underset{\sim}{x}_i e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}}$$

Precisamos métodos numéricos para obtenção de $\hat{\underset{\sim}{\gamma}}$.

Temos ainda que

$$\frac{\partial^2 \ell(\underset{\sim}{\gamma})}{\partial \gamma_j \gamma_k} = -\sum_{i=1}^{n} x_{ij} x_{ik} e^{y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma}} \quad , \quad j, k = 1, \ldots, p$$

$$-E\left(\frac{\partial^2 \ell(\gamma)}{\partial \gamma_j \partial \gamma_i}\right) = +\sum_{i=1}^{n} x_{ij} x_{ik} \, E(e^{y_i}) \, e^{-x_i \gamma}$$

$$E(e^{y_i}) = E(E(e^{y_i} / \delta_i))$$

$$= P(\delta_i = 0) \, E(e^{y_i} / \delta_i = 0) + P(\delta_i = 1) \, E(e^{y_i} / \delta_i = 1)$$

→ $E(e^{y_i})$ envolve a distribuição da censura.

Terminologia

$-\frac{\partial}{\partial \gamma \partial \gamma'} \ell(\gamma)$: informação OBSERVADA

"Não se pode calcular o esp, o que torna impossível o método escore de Fisher - esp envolve a censura. Logo, o ideal é método de Newton-Raphson"

Analise de Sobrevivência

Aula 15

19/05/2016

Modelo de Regressão Weibull

Relembre que, para a distribuição Weibull, temos

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

$$\alpha(t) = \theta \rho t^{\rho-1}$$

Podemos considerar o modelo

$$S(t/X_i) = e^{-\theta_i t^{\rho}}, \quad \theta_i = e^{X_i'\beta} \quad \text{ou}$$

$$\alpha(t/X_i) = \theta_i \rho t^{\rho-1} = (e^{X_i'\beta})\rho t^{\rho-1}$$

Neste caso, temos

$$f(t/X_i) = (e^{X_i'\beta})\rho t^{\rho-1} e^{-(e^{X_i'\beta})t^{\rho}}$$

Observe que, para duas observações distintas i e j temos

$$\frac{\alpha(t/X_i)}{\alpha(t/X_j)} = \frac{e^{X_i'\beta}\rho t^{\rho-1}}{e^{X_j'\beta}\rho t^{\rho-1}} = \exp\{(X_i - X_j)'\beta\},$$

que não depende de t.

$\Rightarrow$ O modelo Weibull é um modelo de riscos proporcionais.

Outra Representação do Modelo Weibull

Considere o modelo

$$Y = \ln T = X'\beta + \sigma W,$$

W no valor extremo padrão, sem

$$f_W(w) = e^{w} \exp\{-e^{w}\}; \quad S(w) = \exp\{-e^{w}\}$$

Nesse caso, temos que

$$S(y/\underline{x}) = P(Y > y / \underline{x})$$
$$= P(\underline{x}'\underline{\beta} + \sigma W > y / \underline{x})$$
$$= P\left(W > \frac{y - \underline{x}'\underline{\beta}}{\sigma} / \underline{x}\right)$$
$$= \exp\left\{-e^{\frac{y - \underline{x}'\underline{\beta}}{\sigma}}\right\}.$$

Vamos encontrar

$$S_T(t/\underline{x}) = P(T > t / \underline{x})$$
$$= P(\ln T > \ln t / \underline{x})$$
$$= P(Y > \ln t / \underline{x})$$
$$= \exp\left\{-e^{\left[\frac{\ln t - \underline{x}'\underline{\beta}}{\sigma}\right]}\right\} = \exp\left\{-t^{1/\sigma} e^{-\underline{x}'\underline{\beta}/\sigma}\right\}$$

Observamos que

$$T \sim \text{Weibull} \quad e \quad \rho = 1/\sigma \quad e \quad \beta = \frac{1}{\sigma}\underline{\beta}$$

Função de Verossimilhança

Consideramos o modelo

$$Y_i^* = \ln T_i = \underline{x}_i'\underline{\beta} + \sigma W_i \quad , \quad W_i \sim \text{valor extrema padrão}$$

Observamos

$$Y_i = \ln(\min(T_i, C_i))$$
$$\delta_i = I(T_i \leq C_i)$$
$$X_i = \text{vetor de covariáveis}$$

A função de verossimilhança fica dada por

$$L(\underline{\beta}, \sigma) = \prod_{i=1}^{n} (f(y_i / \underline{x}_i))^{\delta_i} (S(y_i / \underline{x}_i))^{1-\delta_i}$$
$$= \exp\left\{\sum_{i=1}^{n} \frac{\delta_i(y_i - \underline{x}_i'\underline{\beta})}{\sigma}\right\} \exp\left\{-\sum_{i=1}^{n} e^{(y_i - \underline{x}_i'\underline{\beta})/\sigma}\right\}$$

A função de log-verossimilhança é

$$\ell(\underset{\sim}{\gamma}, \sigma) = \sum_{i=1}^{n} \frac{\delta_i (y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma})}{\sigma} - \sum_{i=1}^{n} e^{(y_i - \underset{\sim}{x}_i' \underset{\sim}{\gamma})/\sigma}$$

→ utilizamos algoritmo numérico para obtenção de $\hat{\underset{\sim}{\gamma}}$ e $\hat{\sigma}$.

* Modelo de Regressão Log-Logístico

Considere o modelo

$$Y = \ln T = \underset{\sim}{x}' \underset{\sim}{\gamma} + \sigma W,$$

$W \sim$ distribuição logística padrão, ou seja,

$$f_W(w) = \frac{e^{w}}{(1+e^{w})^2}, \quad S(w) = \frac{1}{1+e^{w}}$$

Neste caso, temos que

$$\begin{aligned} S_T(t/\underset{\sim}{x}) &= P(T > t / \underset{\sim}{x}) = P(Y > \ln t / \underset{\sim}{x}) \\ &= P(\underset{\sim}{x}' \underset{\sim}{\gamma} + \sigma W > \ln t / \underset{\sim}{x}) \\ &= P\left(W > \frac{\ln t - \underset{\sim}{x}' \underset{\sim}{\gamma}}{\sigma} \Big/ \underset{\sim}{x}\right) \\ &= \frac{1}{1 + \exp\left\{\frac{\ln t - \underset{\sim}{x}' \underset{\sim}{\gamma}}{\sigma}\right\}} \\ &= \frac{1}{1 + t^{1/\sigma} e^{-\underset{\sim}{x}' \underset{\sim}{\gamma}/\sigma}} \end{aligned}$$

De modo análogo, pod. obt. a log-verossi: os parâmetros são est. por algum método numérico

* Modelo de Regressão Log-Normal

Se $T \sim$ log-normal (μ, σ^2), então

$$f(t) = \frac{1}{\sigma t \sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(\ln t - \mu)^2\right\}$$

Podemos escrever

$$Y = \log T \sim N(\mu, \sigma^2)$$

$$Y = \log T = \mu + \sigma \varepsilon \quad , \quad \varepsilon \sim N(0,1)$$

A incorporação de covariáveis pode ser feita da seguinte forma.

$$Y_i = \log T_i = \underset{\sim}{X}_i' \underset{\sim}{\gamma} + \sigma \varepsilon_i \quad , \quad \varepsilon_i \sim N(0,1)$$

※ Modelos de Locação-Escala

Observe que todos os modelos paramétricos podem ser escritos na forma

$$Y_i = \log T_i = \underset{\sim}{X}_i' \underset{\sim}{\gamma} + \sigma \varepsilon_i \quad , \quad \varepsilon_i \sim \text{distribuição na forma padrão} \qquad (1)$$

Dizemos que um modelo é de <u>locação-escala</u> quando podemos representá-lo na forma (1).

* Todos os gd distribuídos são de locação-escala.

※ Interpretação dos Parâmetros

No caso da dist. Weibull,

$$\frac{\alpha(t/\underset{\sim}{X}_i)}{\alpha(t/\underset{\sim}{X}_j)} = \frac{\rho t^{\rho-1} e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}}}{\rho t^{\rho-1} e^{\underset{\sim}{X}_j' \underset{\sim}{\beta}}} = \exp\left\{ (\underset{\sim}{X}_i - \underset{\sim}{X}_j)' \underset{\sim}{\beta} \right\}$$

$$\underset{\sim}{X}_i = \begin{pmatrix} 1 \\ X_{1i} \\ X_{2i} \end{pmatrix} \begin{matrix} \\ \to \text{cont.} \\ \to \text{bicolas.} \end{matrix} \qquad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

Assuma que

$$\underset{\sim}{X}_j = \begin{pmatrix} 1 \\ X_1^{(0)} \\ X_{2j} \end{pmatrix} \; ; \; \underset{\sim}{X}_i = \begin{pmatrix} 1 \\ X_1^{(1)} \\ X_{2j} \end{pmatrix}$$

$$\frac{\alpha(t/\underset{\sim}{X}_i)}{\alpha(t/\underset{\sim}{X}_j)} = e^{\beta_1} > 1$$

$$(e^{\beta} - 1) \% \%$$

- Log-logístico, razão de chances

A.S. - Aula 15

19/04/16

Folha 02

- Normal: razão de medianas.

* Resultados Assintóticos para modelos paramétricos. (Kalbfleisch, pg 27)

Assuma que temos uma amostra aleatória $(Z_i, \delta_i, \underset{\sim}{X}_i)$, $i = 1, 2, \ldots, n$ com

$$Z_i = \min(T_i, C_i)$$

$$\delta_i = I(T_i \leq C_i)$$

$\underset{\sim}{X}_i$: vamos a ~~pressupor~~ vetor de p covariáveis

Definimos

$$N_i(t) = I(Z_i \leq t, \delta_i = 1)$$

$$Y_i(t) = I(Z_i \geq t)$$

Considere o processo de intensidade

$$\Lambda_i(t/\underset{\sim}{X}_i) = \int_0^t Y_i(u)\, \alpha(u/X_i)\, du$$

$$= \int_0^t \lambda_i(u)\, du \quad , \quad \lambda_i(u) = Y_i(u)\, \alpha_i(t/\underset{\sim}{X}_i)$$

$\alpha_i(t/\underset{\sim}{X}_i)$: função conhecida que depende de parâmetros desconhecidos.

A função de verossimilhança é dada por

$$L(\underset{\sim}{\theta}) = \prod_{i=1}^{n} (f(z_i/\underset{\sim}{X}_i))^{\delta_i} (S(z_i/X_i))^{1-\delta_i}$$

$$= \prod_{i=1}^{n} (\alpha(z_i/\underset{\sim}{X}_i)\, S(z_i/\underset{\sim}{X}_i))^{\delta_i}\, S(z_i/\underset{\sim}{X}_i)^{1-\delta_i}$$

$$= \prod_{i=1}^{n} (\alpha(z_i/\underset{\sim}{X}_i))^{\delta_i}\, S(z_i/\underset{\sim}{X}_i)$$

A função de log-verossimilhança é dada por

$$\ell(\underset{\sim}{\theta}) = \sum_{i=1}^{n} \delta_i \ln(\alpha(z_i/\underset{\sim}{X}_i)) + \sum_{i=1}^{n} \ln(S(z_i/\underset{\sim}{X}_i))$$

Note que podemos escrever

$$S(z/\underset{\sim}{X}_i) = \exp\left\{-\int_0^z \alpha(u/X_i)\,du\right\}$$

$$= \exp\left\{-\int_0^\tau Y_i(u)\alpha(u/\underset{\sim}{X}_i)\,du\right\}$$

τ: inst. de temp. suf. grande.

Portanto,

$$\ell(\underset{\sim}{\theta}) = \sum_{i=1}^{n}\int_0^\tau \ln(\alpha(u/\underset{\sim}{X}_i))\,dN_i(u)$$

$$+\sum_{i=1}^{n}(-1)\int_0^\tau Y_i(u)\alpha(u/\underset{\sim}{X}_i)\,du$$

A função escore $U(\underset{\sim}{\theta},\tau) = \frac{\partial}{\partial\underset{\sim}{\theta}}\ell(\underset{\sim}{\theta})$ é dada por

$$U(\underset{\sim}{\theta},\tau) = \frac{\partial}{\partial\underset{\sim}{\theta}}\ell(\underset{\sim}{\theta})$$

$$= \sum_{i=1}^{n}\int_0^\tau \frac{\partial}{\partial\underset{\sim}{\theta}}\ln(\alpha(u/\underset{\sim}{X}_i))\,dN_i(u)$$

$$-\sum_{i=1}^{n}\int_0^\tau Y_i(u)\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\alpha(u/X_i)\right)du$$

$$= \sum_{i=1}^{n}\int_0^\tau\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\ln(\alpha(u/\underset{\sim}{X}_i))\right)dN_i(u)$$

$$-\sum_{i=1}^{n}\int_0^\tau Y_i(u)\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\alpha(u/\underset{\sim}{X}_i)\frac{1}{\alpha(u/\underset{\sim}{X}_i)}\right)\alpha(u/X_i)\,du$$

$$U(\underset{\sim}{\theta},\tau) = \sum_{i=1}^{n}\int_0^\tau\left(\frac{\partial}{\partial\underset{\sim}{\theta}}\ln(\alpha(u/X_i))\right)dN_i(u)$$

$$-\sum_{i=1}^{n}\int_0^\tau Y_i(u)\left(\frac{d}{d\underset{\sim}{\theta}}\ln(\alpha(u/X_i))\right)\alpha(u/X_i)\,du$$

$$U(\theta,\tau) = \sum_{i=1}^{n} \int_0^{\tau} \left(\frac{\partial}{\partial\theta} \ln \alpha(u/\underset{\sim}{X}_i)\right) \underbrace{\left(dN_i(u) - Y_i(u)\,\alpha(u/X_i)\,du\right)}_{dM_i(u)}$$

Relembre que

$$M_i(t) = N_i(t) - \Lambda_i(t/\underset{\sim}{X}_i) \text{ é um martingal}$$

$$\Lambda_i(t)(\underset{\sim}{X}_i) = \int_0^t Y_i(u)\,\alpha(u/\underset{\sim}{X}_i)\,du$$

$$U(\theta,\tau) = \sum_{i=1}^{n} \int_0^{\tau} \left[\frac{\partial}{\partial\theta} \ln \alpha(u/\underset{\sim}{X}_i)\right] dM_i(u)$$

Relembre também que

$$L(t) = \int_0^t \underbrace{H(u)}_{\text{proc. previsível}}\, dM(u) \quad \text{é um martingal}$$

Como $\frac{\partial}{\partial\theta} \ln\alpha(u/\underset{\sim}{X}_i)$ é processo previsível

$U(\theta,\tau)$ é um martingal.

Processo de Variação Previsível de $U(\theta)$ $(\langle M\rangle)$

Relembre que $dM(t) = \text{var}(dM(t)/\mathcal{F}_{t^-})$

- $\text{Var}(M(t)) = E(\langle M\rangle(t))$
- $\langle M\rangle(t) = \Lambda(t)$
- $\langle L\rangle(t) = \int_0^t H^2(u)\, d\langle M\rangle(u)$

- $L(t) = \int_0^t H(u)\, dM(u)$

Pode-se demonstrar que

$$\langle U(\theta)\rangle(t) = \sum_{i=1}^{n} \int_0^t \left(\frac{\partial}{\partial\theta} \ln\alpha(u/\underset{\sim}{X}_i)\right)^{\otimes 2} Y_i(u)\,\alpha(u/\underset{\sim}{X}_i)\,du,$$

em que

$$b^{\otimes 2} = b\, b'$$

Seja

$$I(\theta) = -\frac{\partial^2 L(\theta)}{\partial\theta\, \partial\theta^T}$$, matriz de inf. observada.

Podemos verificar que

$$I(\theta) = \sum_{i=1}^{n} \int_0^{\tau} \left(\frac{\partial}{\partial\theta} \ln(\alpha(u|x_i)) \right)^{\otimes 2} dN_i(t)$$

$$\underbrace{- \sum_{i=1}^{n} \int_0^{\tau} \left(\frac{\partial^2}{\partial\theta\, \partial\theta^T} \ln(\alpha(u|x_i)) \right) \alpha(u|x_i)\, dM_i(u)}_{\text{Martingal de medio cero}} \Rightarrow$$

$$E(I(\theta)) = \sum_{i=1}^{n} \int_0^{\tau} \left(\frac{\partial}{\partial\theta} \ln \alpha(u|x_i) \right)^{\otimes 2} E(dN_i(u))$$

$$= \sum_{i=1}^{n} \int_0^{\tau} \left(\frac{\partial \ln \alpha(u|x_i)}{\partial\theta} \right)^{\otimes 2} E\left[\underbrace{E\left(dN_i(u) \mid \mathcal{F}_{u^-} \right)}_{\text{Def. proc. intensidade}} \right]$$

$$E(I(\theta)) = \sum_{i=1}^{n} \int_0^{\tau} \left(\frac{\partial}{\partial\theta} \ln \alpha(u|x_i) \right)^{\otimes 2} E\left(Y_i(u) \right) \alpha(u|x_i)\, du \qquad (2)$$

Comparando (1) e (2):

$$E(I(\theta)) = E\left(\langle U(\theta) \rangle (t) \right) = Var(U(\theta, \tau))$$

Resultado: Kalbfleisch, pg 180

Assuma válidas as condições

(i) $\frac{1}{n} \langle U(\theta_0) \rangle (\tau) \xrightarrow{P} I(\theta_0)$,

com

θ_0: verdadeiro valor do vetor de parâmetros

(ii) Para qualquer $\varepsilon > 0$,

A. S. - Aula 15

19/04/16

Folha 03

$$\frac{1}{n}\sum_{i=1}^{n}\int_{0}^{\tau}\left(\frac{\partial \ln \alpha(t/x_i)}{\partial \theta_j}\right)^2 I\left(\left|\sqrt{n}\,\frac{\partial \ln \alpha(t/x_i)}{\partial \theta_0}\right| > \varepsilon\right)$$

$$\times\, Y_i(t)\,\alpha(t/x_i)\,dt \xrightarrow{P} 0$$

Então vale o resultado

$$(1/\sqrt{n})\,(U(\theta_0)(\tau) \xrightarrow{D} N(0, \Sigma(\theta_0))$$

Análise de Sobrevivência

Aula 16

26/04/2016

Prova 10-14; Sala B138

$Z_i = \min(T_i, C_i)$; $\delta_i = I(T_i \leq C_i)$

$\underline{X}_i$: vetor de p covariáveis

Definimos

$$N_i(t) = I(Z_i \leq t, \delta_i = 1)$$

$$Y_i(t) = I(Z_i \geq t)$$

Processo de intensidade

$$\Lambda_i(t \mid \underline{X}_i) = \int_0^t Y_i(u)\alpha(u/\underline{X}_i)\,du$$

função de risco

$$L(\underline{\theta}) = \prod_{i=1}^{n} (\alpha(z_i/\underline{X}_i))^{\delta_i} S(z_i/\underline{X}_i)$$

$$\ell(\underline{\theta}) = \sum_{i=1}^{n} \delta_i \ln \alpha(z_i/\underline{X}_i) + \sum_{i=1}^{n} \ln S(z_i/\underline{X}_i)$$

$$U(\theta) = -\frac{\partial \ell(\theta)}{\partial \theta} = \sum_{i=1}^{n} \int_0^{\tau} \frac{\partial}{\partial \underline{\theta}} \ln \alpha(u/\underline{X}_i)$$

$$\langle U(\theta)\rangle(t) = \sum_{i=1}^{n} \int_0^t \left(\frac{\partial}{\partial \theta} \ln \alpha(u/\underline{X}_i)\right)^{\otimes 2} Y_i(u)\alpha(u/\underline{X}_i)\,du$$

$$\underline{b}^{\otimes 2} = \underline{b}\,\underline{b}^T$$

$$E(I(\underline{\theta})) = E(\langle U(\underline{\theta})\rangle(\tau)) = Var(U(\theta)(\tau))$$

$I(\underline{\theta})$: matriz de informação observada, $I(\underline{\theta}) = -\frac{\partial}{\partial \underline{\theta}} U(\underline{\theta})$

Resultado. Kalbfleisch, pg 180. Ver Andersen, pgs 420-424 (Aula Passada - correção do resultado).

Assuma válidas as condições

(1) $\frac{1}{n}\langle U(\underline{\theta}_0)\rangle(\tau) \xrightarrow{P} \Sigma(\underline{\theta}_0)$; $\underline{\theta}_0$: verdadeiro valor do vetor de parâmetros

$\Sigma(\underline{\theta}_0)$: matriz positiva definida

(ii) Para qualquer $\delta > 0$,

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^t \left(\frac{\partial}{\partial\theta_j}\ln\alpha(t/\underline{X}_i)\right)^2 \times I\left(\left|(\sqrt{n})^{-1}\frac{\partial}{\partial\theta_j}\ln\alpha(t/\underline{X}_i)\right| > \delta\right) \times$$

$$\times Y_i(u)\alpha(u/\underline{X}_i)\,du \xrightarrow{P} 0.$$

Então vale o resultado

$$(1/\sqrt{n})\, U(\underline{\theta}_0)(\tau) \xrightarrow{D} N(0, \Sigma(\underline{\theta}_0))$$

Considere agora a expansão de Taylor $U(\hat{\underline{\theta}})$ ao redor de $\underline{\theta}_0$:

o $\hat{\underline{\theta}}$ é o EMV, que zera $U(\cdot)$.

$$U(\hat{\underline{\theta}}) \approx U(\underline{\theta}_0) + U'(\underline{\theta}_0)(\hat{\underline{\theta}} - \underline{\theta}_0)$$

$$\Downarrow$$

$$U'(\underline{\theta}_0)(\hat{\underline{\theta}} - \underline{\theta}_0) \approx -U(\underline{\theta}_0)$$

$$\parallel$$

$$-I(\underline{\theta}_0)(\hat{\underline{\theta}} - \underline{\theta}_0) \approx -U(\underline{\theta}_0)$$

$$(\hat{\underline{\theta}} - \underline{\theta}_0) \approx (I(\underline{\theta}_0))^{-1}\, U(\underline{\theta}_0)$$

Podemos escrever

$$(\hat{\underline{\theta}} - \underline{\theta}_0) \approx \left(\frac{I(\underline{\theta}_0)}{n}\right)^{-1} \frac{U(\underline{\theta}_0)}{n}$$

$$\sqrt{n}\,(\hat{\underline{\theta}} - \underline{\theta}_0) \approx \underbrace{\left(\frac{I(\underline{\theta}_0)}{n}\right)^{-1}}_{\downarrow P \;\; (\Sigma(\underline{\theta}_0))^{-1}} \underbrace{\left(\frac{U(\underline{\theta}_0)}{\sqrt{n}}\right)}_{\downarrow D \;\; N(0, \Sigma(\underline{\theta}_0))}$$

Por Slutzky,

$$\sqrt{n}\,(\hat{\underline{\theta}} - \underline{\theta}_0) \xrightarrow{D} N(0, (\Sigma(\underline{\theta}_0))^{-1})$$ Teo. VI.1.2 pg 124, Andersen.

Testes de hip: Wald; R. de Verossimilhança; Escore.

Métodos de Diagnósticos Para Modelos Paramétricos

(Klein, Enrico)

→ Método Gráfico Descritivo

Vimos que

- Exponencial

$$\alpha(t) = \lambda \quad , \quad A(t) = \lambda t$$

- Weibull

$$\alpha(t) = \rho\theta t^{\rho-1} \quad , \quad A(t) = \theta t^{\rho} \quad , \quad \ln A(t) = \ln\theta + \rho \ln t$$

- Log-logística

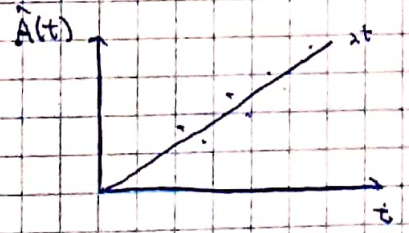
$$\alpha(t) = \frac{\gamma t^{\gamma-1}\lambda}{1+\lambda t^{\gamma}} \quad , \quad A(t) = \ln(1+\lambda t^{\gamma})$$

"Idéia: estimo a f. de risco por Nelson-Aalen e comparo com o modelo teórico. Como $\alpha(t)$ e $A(t)$ são, por vezes, complexas, considero-se em primeiro momento uma transformação, a ~~fim de encontrar uma relação~~ h"

Idéia do Método. Estimar $A(t)$ usando o estimador de Nelson-Aalen (separadamente para cada nível de uma covariável categorizada).

→ Adequabilidade da distribuição Exponencial

Gráfico de $\hat{A}(t)$ versus t.

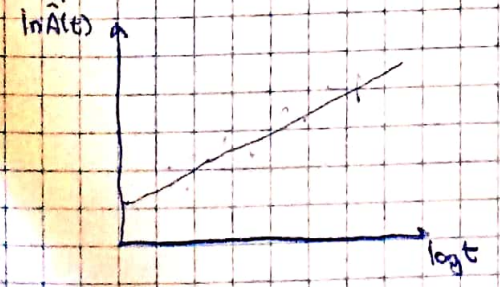


Comportamento esperado se o mod. exp. for adequado

Reta passando pela origem

→ Adequabilidade da distribuição Weibull

Gráfico: $\ln \hat{A}(t)$ versus $\log t$



Comportamento esperado: reta com inclinação ρ e intercepto $\log\theta$.

→ Adequabilidade da distribuição Log-logística.

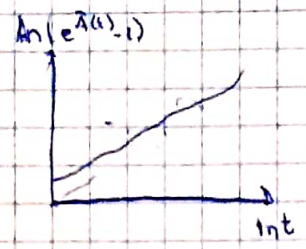
Gráfico: $\ln(e^{\hat{A}(t)} - 1)$ versus $\log t$

observe que, para a dist. Log-logística,

$$e^{A(t)} = 1 + \lambda t^{\gamma} \Leftrightarrow$$

$$e^{A(t)} - 1 = \lambda t^{\gamma} \Leftrightarrow$$

$$\ln(e^{A(t)} - 1) = \ln\lambda + \gamma \ln t.$$



Comportamento esperado: reta.

* Resíduos Padronizados

São utilizados para avaliar a qualidade geral do ajuste.

São calculados com base na representação log-linear do modelo.

$$Y_i = \ln T_i = X_i \underset{\sim}{\gamma} + \sigma W_i, \quad i = 1, 2, \ldots, n.$$

Os resíduos padronizados são dados por

$$s_i = \frac{\ln Z_i - X_i \hat{\gamma}}{\hat{\sigma}},$$

$Z_i = \min(T_i, C_i)$.

Qual é o comportamento esperado dos resíduos padronizados?

→ Modelo Exponencial $s_i, i = 1, \ldots, n$ devem constituir uma amostra Censurada de variáveis com dist. valor extremo padrão.

→ Modelo Weibull: idem?

→ Modelo log-logístico:

s_i devem constituir uma amostra Censurada da distribuição logística padrão.

→ Modelo log-normal:

s_i devem constituir uma amostra censurada da dist. normal padrão.

Definimos

$$\hat{e}_i^{+} = e^{s_i}$$

Análise de Sobrevivência

Aula 16

Folha 02

Distribuição de $\hat{e}_i^*$:

→ Exponencial → exp(1)

→ Weibull → exp(1)

→ log-logístico ~ log-logística padrão

→ log normal ~ log-normal padrão

Procedimentos Sugeridos

1. Obtemos o estimador de KM dos resíduos $\hat{e}_i^*$. Comparamos graficamente o estimador de KM obtido e a curva teórica.

- gráfico de KM juntamente com a curva teórica.
- gráfico de 'quantis'

$$\hat{S}(e_i^*) \text{ versus } S_T(\hat{e}_i^*)$$

* Resíduos de Cox-Snell

São também utilizados para a avaliação da adequação geral do modelo.

No caso de modelos paramétricos, a análise com base nos resíduos padronizados é equivalente a análise dos resíduos de Cox-Snell.

São definidos por

$$\hat{r}_i = \hat{A}_M(z_i \mid X_i)$$

$\hat{A}(\cdot)$ é o estimador de $A(\cdot)$ com base no modelo.

Exemplos:

. Exponencial

$$A(t) = \lambda_i t = t e^{X_i \beta t}$$

Assim,

$$\hat{r}_i = z_i e^{X_i \hat{\beta}} \quad (\beta \text{ com intercepto})$$

- Weibull

$$A(t/x) = e^{x'\beta} t^{\rho} \Rightarrow$$

$$\hat{r}_i = e^{x_i'\hat{\beta}} z_i^{\hat{\rho}} \quad (\beta \text{ com intercepto})$$

- Log. logística

$$\alpha(t/x) = \frac{\rho t^{\rho-1} e^{x'\beta}}{1 + e^{x'\beta} t^{\rho}}$$

$$\Rightarrow \hat{r}_i = -\log\left(\frac{1}{1 + (e^{x_i'\hat{\beta}}) z_i^{\hat{\rho}}}\right)$$

$$= -\log\left(\frac{1}{1 + [\exp(-x_i'\hat{\gamma})]^{1/\hat{\sigma}} z_i^{1/\hat{\sigma}}}\right)$$

- Log-normal

$$\hat{r}_i = -\log\left(1 - \Phi\left(\frac{\ln z_i - x_i'\hat{\gamma}}{\hat{\sigma}}\right)\right) \quad \text{(mod. locação escala)}$$

$\Phi(.)$: função de dist. acumulada da $N(0,1)$

Como verificar a adequabilidade do modelo?

Seja T uma v.a. representando o tempo de falha e $A(.)$ a função de taxa de falha acumulada da dist. de T. Defina

$$V = A(T)$$

Temos que

$$S_V(v) = P(V > v) = P(A(T) > v)$$

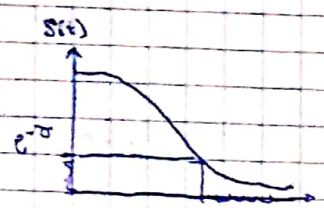
$$= P(-A(T) < -v)$$

$$= P(e^{-A(T)} \leq e^{-v})$$

$$= P(S(T) \leq e^{-v})$$

$$= P(T > S^{-1}(e^{-v}))$$

$$= S(S^{-1}(e^{-v})) = e^{-v}$$



$\therefore$ $V \sim$ exponencial (1)

Procedimento

- obtemos os ~~resultados~~ resíduos de Cox-Snell

Analisamos $\hat{r}_1, \hat{r}_2, \ldots, \hat{r}_n$ como se fossem uma amostra aleatória censurada:

- obtemos o estimador $\hat{A}_r(t)$ de função de risco acumulada dos resíduos (Nelson-Aalen)
- Comparamos $\hat{A}_r(t)$ com $A(t) = t$.

Problema dos resíduos de Cox-Snell.

- Não fornecem informação sobre a causa do problema quando o modelo não está ~~adequado~~ bem.
- O resultado

$$A(t) \sim \exp(1)$$

é válido para os verdadeiros parâmetros. Quando usamos $\hat{\beta}$ no lugar de β, a aproximação pela exponencial pode não ser boa.

Análise de Sobrevivência

Aula 17

03/05/2015

Métodos de Diagnóstico

→ Método Descritivo

→ Resíduos padronizados

$$s_i = \frac{\ln Z_i - \hat{\alpha}' \hat{\gamma}}{\hat{\sigma}}$$

$$\hat{\varepsilon}_i^* = e^{s_i}$$

Tanto os padronizados quanto o de Cox-Snell apenas indicam uma ideia da qualidade (em caso contrário, não indica nada sobre a adequada) e, ainda assim, com estimativas, e pode influenciar os resultados. O de Cox-Snell é amplamente usado, mas há casos limitantes.

→ Resíduo de Cox-Snell

$$r_i = \hat{A}_M(z_i | x_i)$$

T vem da cond. S ~, $V = A(T) \sim \exp(1)$. → Ideia do ajuste

Vamos continuar: Res. Martingal e Deviance.

Resíduo Martingal

No caso paramétrico, o resíduo martingal é definido por

$$M_i = \delta_i - \hat{r}_i,$$

em que $\hat{r}_i$ é o resíduo de Cox-Snell.

"Se trata de um resíduo assimétrico, com comport. estranho, valendo $1-\hat{r}_i$ quando falha e $-\hat{r}_i$ qdo censura"

O arg. e: M_i fornece informação sobre o número de eventos observados nos dados, mas não previstos no modelo.

→ Só pra ver análise funcional

Principal uso: verificação da melhor funcional da variável.

Resíduos Deviance

Os resíduos Deviance são dados por

$$D_i = \text{Sinal}(M_i)\left[-2(M_i + \delta_i \log(\delta_i - M_i))\right]^{1/2}.$$

"Nos livros, diz que foram propostos para obter mais simetria, não usa antes dos martingais. Ainda assim, comportamento estranho. O ideal é usá-lo para identificar pontos atípicos"

Foram propostos como uma tentativa de tornar os resíduos martingais mais simétricos em torno de zero.

Principal uso: detecção de pontos atípicos.

Gráficos Sugeridos

$\begin{pmatrix} \text{Resíduo} \\ \text{Deviance} \end{pmatrix} \times \begin{matrix} \text{tempo} \\ \text{\# observação} \end{matrix}$

→ Slides: câncer de laringe.

, pacote survminer → comando ggsurvplot constrói gráfico km pelo ggplot-2

* Aj. exponencial

survreg(..., ..., dis="exp")

eba - solta coef. com outra parametrização

No gráfico Resíduos Dev. x tempo, em geral, os resíduos foram entre $-3-\varepsilon$ e $3+\varepsilon$, então ok. Caso tenha algo acima de 3 ou ab. de -3, outlier, p.ex.

Modelo ~~de Cox~~ Semiparamétrico de Cox

"Na área médica, é + usado. Talvez 99% das vezes. Até os médicos o preferem, ainda que os par. possam ser melhores."

Seja (Z_i, δ_i, X_i), $i=1, \dots, n$ uma amostra aleatória com

$$Z_i = \min(T_i, C_i) \quad \text{e} \quad \delta_i = I(T_i \leq C_i)$$

Aqui,

X_i: vetor de covariáveis (sem intercepto)

No modelo semiparamétrico de Cox, supomos que a função de risco é dada por

$$\alpha(t/X_i) = \underbrace{\alpha_0(t)}_{\text{parte não paramétrica}} \underbrace{e^{X_i'\beta}}_{\text{parte paramétrica}},$$

depende de um vetor de parâmetros não conhecidos

em que

$\alpha_0(t)$ é uma função arbitrária do tempo, não especificada e β é um vetor de parâmetros desconhecidos.

Observe que se

- $\alpha_0(t)$ é especificada ⇒ modelo paramétrico.
- $\alpha_0(t)$ é arbitrária ⇒ modelo semiparamétrico.

Note ainda que

$$\frac{\alpha(t/X_i)}{\alpha(t/X_j)} = \frac{\alpha_0(t)e^{X_i'\beta}}{\alpha_0(t)e^{X_j'\beta}} = e^{(X_i'-X_j')\beta}$$ ⇒ É um modelo de riscos proporcionais.[1]

1. Por isso, tbm é conh. como modelo de riscos proporcionais.

De maneira mais geral, outras funções podem ser usadas na função de taxa de falha

$$\alpha(t/X_i) = \alpha_0(t)\, g(X)$$

- $g(X) = e^{X'\beta}$ (m. de Cox é o + usado)

- $g(\underline{X}) = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p = \underline{X}'\underline{\beta}$ (Há que por restrições p/ que $\alpha(t/\underline{X}) > 0$)

Interpretação dos parâmetros.

$$\frac{\alpha(t/\underline{X}_i)}{\alpha(t/\underline{X}_j)} = e^{(\underline{X}_i - \underline{X}_j)'\underline{\beta}}$$

⇒ Cov. Binária: $\frac{\alpha(t/\underline{X}_i)}{\alpha(t/\underline{X}_j)} = e^{\beta_p}$.

"O risco de óbito aumenta em tantos porcentos quanto aumenta-se uma un. na cov."

⇒ Cov. contínua: $\frac{\alpha(t/X = x_0 + 1)}{\alpha(t/X = x_0)} = e^{\beta}$

Análise de Sobrevivência

Aula 18

05/05/2016

Modelo Semiparamétrico de Cox e Extensões

$Z_i = \min(T_i, C_i)$

$\delta_i = I(T_i \leq C_i)$

$\underset{\sim}{X}_i$ = vetor de covariáveis (sem intercepto)

$$\alpha(t/\underset{\sim}{X}) = \underbrace{\alpha_0(t)}_{\text{não espec.}} e^{\underbrace{\underset{\sim}{X}_i'\underset{\sim}{\beta}}_{\text{parte paramétrica}}}$$

$$\frac{\alpha(t/\underset{\sim}{X}_i)}{\alpha(t/\underset{\sim}{X}_j)} = e^{(\underset{\sim}{X}_i - \underset{\sim}{X}_j)'\underset{\sim}{\beta}}$$

Não se inclui o intercepto pois, nesse caso, $\alpha(t/\underset{\sim}{X}) = \underbrace{\alpha_0(t)\, e^{\beta_0}}_{\alpha^*(t) \to \text{fun. de falha basal end. parâm. } \beta_0} e^{\beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p}$

Vamos agora à estimação dos parâmetros. Problema na est. da verossim.: a parcial e a perfilada.

Estimação Dos Parâmetros No Modelo de Cox

1ª Abordagem Fleming, pg 137. Verossimilhança Parcial de Cox

Em primeiro momento, uma ideia geral:

Suponha que, em um caso geral, o vetor $\underset{\sim}{Y}$ de variáveis resposta possa ser particionado em uma sequência de pares

$$(V_1, W_1, V_2, W_2, V_3, W_3, \dots, V_n, W_n)$$

A função de verossimilhança para um vetor de parâmetros Θ pode ser escrita na forma

$$L(\underset{\sim}{\Theta}) = f_{V_1, W_1, \dots, V_n, W_n}(v_1, w_1, v_2, w_2, \dots, v_n, w_n);$$

ou, de modo equivalente,

$$L(\underset{\sim}{\Theta}) = f_{V_1}(v_1)\, f_{W_1/V_1}(w_1/v_1)\, f_{V_2/V_1, W_1}(v_2/v_1, w_1) \times f_{W_2/V_1, W_1, V_2}(w_2/v_1, w_1, v_2) \times$$

$$\times \dots \times f_{W_n/V_1, W_1, V_2, W_2, \dots, V_{n-1}, W_{n-1}}(w_n/v_1, w_1, v_2, w_2, \dots, v_{n-1}, w_{n-1}, v_n)\, f_{W_n/V_1, W_1, \dots, V_{n-1}, W_{n-1}, V_n}(w_n/v_1, w_1, \dots, v_{n-1}, w_{n-1}, v_n)$$

Temos

$$L(\underline{\theta}) = \left[\prod_{i=1}^{n} f_{V_i / V_1, W_1, \dots, V_{i-1}, W_{i-1}} (v_i / v_1, w_1, \dots, v_{i-1}, w_{i-1}) \right]$$

$$\times \left[\prod_{i=1}^{n} f_{W_i / V_1, W_1, \dots, V_{i-1}, W_{i-1}, V_i} (w_i / v_1, w_1, \dots, v_{i-1}, w_{i-1}, v_i) \right]$$

Seja

$$P_i = (V_1, W_1, \dots, V_{i-1}, W_{i-1})$$

$$Q_i = (V_1, W_1, \dots, V_{i-1}, W_{i-1}, V_i)$$

Daí.

$$L(\underline{\theta}) = \left[\prod_{i=1}^{n} f_{V_i / P_i} (v_i / p_i) \right] \underbrace{\left[\prod_{i=1}^{n} f_{W_i / Q_i} (w_i / q_i) \right]}_{(*)}$$

Cox denominou (*) de função de verossimilhança parcial.

Verossimilhança Parcial Para o Modelo de Cox.

Suponha que observamos

$$(Z_i, \delta_i, \underline{X}_i) \quad , i = 1, \dots, n$$

Sejam

$$T_1^0 < T_2^0 < \dots < T_L^0$$

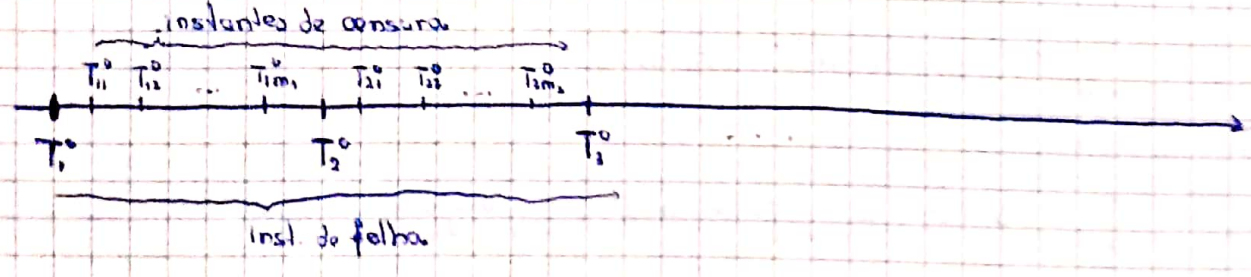
os instantes de falha.

Assuma que não há empates. Denote por (j) a observação que falhou em T_j, $j = 1, \dots, L$.

Sejam

$$T_{j1}^0, T_{j2}^0, \dots, T_{jm_j}^0 \text{ os instantes de censura em } [T_j^0, T_{j+1}^0).$$

Seja (jk) a observação censurada em T_{jk}^0, $k = 1, 2, \dots, m_j$.



Para construir a função de verossimilhança parcial, fazemos

$$V_j = \{ T_j^0, T_{(j-1)1}^0, T_{(j-1)2}^0, \dots, T_{(j-1)m_{j-1}}^0, (j-1, 1), (j-1, 2), \dots, (j-1, m_{j-1}) \}$$

$$W_j = \{ (j) \}$$

A função de verossimilhança parcial dada por (*), neste caso, é

$$L(\underset{\sim}{\beta}) = \prod_{j=1}^{L} f_{W_j|\underset{\sim}{Q}_j}(w_j|\underset{\sim}{q}_j)$$

$$= \prod_{j=1}^{L} P(W_j=(j)/\underset{\sim}{Q}_j)$$

Observe que

$$P(W_j=(j)) = P\left(\begin{array}{c|c} \text{indivíduo } (j) & \text{Toda a hist. até} \\ \text{falhar em } T_j^0=t_j & T_j^{0-} \text{ e uma falha} \\ & \text{ocorreu em } T_j^0 \end{array}\right)$$

tempo de falha do p.es. indivíduo

$$= \frac{P(\text{indivíduo } (j) \text{ falhar / história})}{P(\text{1 falha em } T_j^0=t_j \text{ / história})} \quad \rightarrow T_{(j)} \geq t_j$$

$$= \frac{P(T_{(j)} \in [t_j, t_j+\Delta)/\mathcal{F}_{t^-})\left[\prod_{l \in R_j-(j)} P(T_l \notin [t_j, t_j+\Delta)/\mathcal{F}_{t^-}\right]}{\sum_{l \in R_j} P(T_l \in [t_j, t_j+\Delta)/\mathcal{F}_{t_j^-})\left[\prod_{l \in R_j-l} P(T_l \notin [t_j, t_j+\Delta)/\mathcal{F}_{t_j^-})\right]}$$

$$\approx \frac{\alpha(t_j/\underset{\sim}{X}_{(j)})\,\Delta \cdot \boxed{\left[\prod_{l \in R_j-(j)} (1-\alpha(t_j/\underset{\sim}{X}_l)\Delta\right]} \approx 1}{\Delta \sum_{l \in R_j} \alpha(t_j/\underset{\sim}{X}_l)\boxed{\left[\prod_{l \in R_j-l} (1-\alpha(t_j/\underset{\sim}{X}_l)\Delta\right]} \approx 1}$$

$$= \frac{\alpha(t_j/\underset{\sim}{X}_{(j)})}{\sum_{l \in R_j} \alpha_0(t_j)\, e^{\underset{\sim}{X}_l'\underset{\sim}{\beta}}}, \text{ i.e.}$$

$$P(W_j=(j)) = \frac{e^{\underset{\sim}{X}_{(j)}'\underset{\sim}{\beta}}}{\sum_{l \in R_j} e^{\underset{\sim}{X}_{(l)}'\underset{\sim}{\beta}}}$$

A função de verossimilhança parcial de Cox é dada por

$$L(\underset{\sim}{\beta}) = \prod_{j=1}^{L} \frac{e^{\underset{\sim}{X}_{(j)}'\underset{\sim}{\beta}}}{\sum_{l \in R(t_j)} e^{\underset{\sim}{X}_l'\underset{\sim}{\beta}}}.$$

Análise de Sobrevivência

Aula 19 — 30 de maio de 2016

Comentário sobre a prova:

Ex. 4 $$\int_0^t \frac{d\hat{S}(u)}{\bar{Y}(u)\hat{S}(u)} = \sum_{t_k} \frac{\hat{S}(t_k) - \hat{S}(t_{k-1})}{\bar{Y}(u)\hat{S}(t_k)} = \sum_{k \le st} \frac{d_k}{(r_k - d_k) r_k}$$

21/06 - 2 sem. (William, Marcos)

23/06 - 3 sem. (Rodrigo, Simone / Francisco, Hugo / Jeniffer)

24/06 - 3 seminários (Fabio; Alino / Auro; Pedro)

09/06 - Envio do resumo do seminário

Modelo de Cox (Continuação)

O mod. semiparamétrico de Cox é dado por

$$\alpha(t \mid \underset{\sim}{X}) = \underbrace{\alpha_0(t)}_{\text{N par.}} \underbrace{e^{\underset{\sim}{X}'\underset{\sim}{\beta}}}_{\text{par.}}$$

Dada a dificuldade em função da parte não paramétrica, usamos

Abordagem 1: "Verossim. parcial de Cox"

A função de Verossimilhança parcial de Cox:

$$L(\underset{\sim}{\beta}) = \prod_{j=1}^{L} \frac{e^{\underset{\sim}{X}_{(j)}'\underset{\sim}{\beta}}}{\sum_{i \in R(t_j)} e^{\underset{\sim}{X}_i'\underset{\sim}{\beta}}}$$

$\underset{\sim}{X}_{(j)}$: valor de covariáveis da observação que falhou em $t_{(j)}$.

Abordagem 2: "Verossim. perfilada" (Aula de hoje)

2ª Abordagem

Trabalha-se com a função de verossimilhança perfilada.

Temos que

$$\alpha(t \mid \underset{\sim}{X}) = \alpha_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}}$$

$$S(t \mid \underset{\sim}{X}) = \exp\left\{ -A_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}} \right\}$$

$$A_0(t) = \int_0^t \alpha_0(u)\,du$$

$$L(\underset{\sim}{\beta}, \alpha_0(t)) = \prod_{i=1}^{n} \left[(\alpha_0(t))^{\delta_i} \left(e^{x_i'\beta}\right)^{\delta_i} \exp\left\{ -A_0(t)\, e^{x_i'\beta} \right\} \right]$$

T_i^* : tempo de falha

$T_i = \min(T_i^*, C_i)$

$\delta_i = I(T_i^* \leq C_i)$

Problema. $\alpha_0(t)$ e $A_0(t)$ são f. desconhecidas.

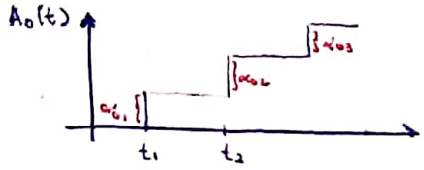
Alternativa:

Aproximar $A_0(t)$ e $\alpha_0(t)$ por funções descontínuas.

Fazemos

$$\alpha_{0j} = \alpha_0(t_j), \quad j = 1, 2, \ldots, L \quad \text{(instantes distintos de falha)}$$

$$A_{0j} = \sum_{t_{(k)} \leq t_j} \alpha_{0k}$$



"Assim, transf. o prob de estimar uma funç. no de estimar tamanho de saltos (agora finito)."

Podemos reescrever a função de verossimilhança em termos de α_{0j}:

$$L(\underset{\sim}{\beta}, \underset{\sim}{\alpha}_0) = \prod_{i=1}^{n} (\alpha_{0i_i})^{\delta_i} \left(e^{x_i'\beta}\right)^{\delta_i} \exp\left\{ -A_{0i_i}\, e^{x_i'\beta} \right\},$$

"Depois, eu retirei os ( ) da notação - estudeiro"

$\alpha_{0i_i} = \alpha(t_{(i)})$; $A_{0i_i} = A_0(t_{(i)}) = \sum_{t_{(k)} \leq t_{(i)}} \alpha_{0k}$ e $\underset{\sim}{\alpha}_0 = (\alpha_{01}, \alpha_{02}, \ldots, \alpha_{0L})^T$.

Ideia geral da Função de Verossimilhança Perfilada.

Assuma que o vetor de parâmetros a ser estimado em $(\underset{\sim}{\theta}, \underset{\sim}{\eta})$, em que $\underset{\sim}{\theta}$ é o vetor de interesse e $\underset{\sim}{\eta}$ é o vetor de parâmetros de incômodo.

Denotando por $L(\underset{\sim}{\theta}, \underset{\sim}{\eta})$ a função de verossimilhança completa, a função de verossimilhança PERFILADA para $\underset{\sim}{\theta}$ é dada por

$$PL(\underset{\sim}{\theta}) = \sup_{\underset{\sim}{\eta}} L(\underset{\sim}{\theta}, \underset{\sim}{\eta})$$

- No modelo de Cox, a função de log-verossimilhança fica dada por

$$\ell(\underset{\sim}{\beta}, \underset{\sim}{\alpha_0}) = \sum_{i=1}^{n} \delta_i \log(\alpha_{0i}) + \sum_{i=1}^{n} \delta_i \underset{\sim}{x}_i' \underset{\sim}{\beta} - \sum_{i=1}^{n} A_{0i}\, e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}}$$

$$= \sum_{i=1}^{n} \delta_i \log(\alpha_{0i}) + \sum_{i=1}^{n} \delta_i \underset{\sim}{x}_i' \underset{\sim}{\beta} - \sum_{i=1}^{n} \Big(\sum_{t_{(j)} \leq t_{(i)}} \alpha_{0j} \Big) e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}}$$

α_{0k}

$t_{(1)} \quad t_{(k)} \quad t_{(L)}$

Assuma que não há empates. Vamos derivar $\ell(\underset{\sim}{\beta}, \underset{\sim}{\alpha_0})$ em relação à α_{0k}:

$$\frac{\partial}{\partial \alpha_{0k}} \ell(\underset{\sim}{\beta}, \underset{\sim}{\alpha_0}) = \frac{1}{\alpha_{0k}} - \sum_{i \in R(t_{(k)})} e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}} \quad , \quad k = 1, 2, \ldots, L$$

Igualando-se a zero, obtemos

$$\frac{1}{\hat{\alpha}_{0k}} = \sum_{i \in R(t_{(k)})} e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}} \Rightarrow \hat{\alpha}_{0k} = \frac{1}{\sum_{i \in R(t_{(k)})} e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}}}$$

Temos também que

$$\hat{A}_{0k} = \sum_{t_{(j)} \leq t_{(k)}} \left(\frac{1}{\sum_{i \in R(t_{(j)})} e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}}} \right)$$

O próximo passo é substituir α_{0j} por $\hat{\alpha}_{0j}$ na função de verossimilhança, obtendo uma verossimilhança perfilada:

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{n} (\hat{\alpha}_{0i})^{\delta_i} \left(e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}}\right)^{\delta_i} \exp\left\{ -\hat{A}_{0i}\, e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}} \right\}$$

$$= \left[\prod_{j=1}^{L} \hat{\alpha}_{0j}\, e^{\underset{\sim}{x}_{(j)}' \underset{\sim}{\beta}} \right] \times \prod_{i=1}^{n} \exp\left\{ -\hat{A}_{0i}\, e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}} \right\}$$

$$= \left[\prod_{j=1}^{L} e^{\underset{\sim}{x}_{(j)}' \underset{\sim}{\beta}} \frac{1}{\sum_{k \in R(t_{(j)})} e^{\underset{\sim}{x}_k' \underset{\sim}{\beta}}} \right] \times \exp\left\{ -\sum_{r=1}^{n} \hat{A}_{0r}\, e^{\underset{\sim}{x}_r' \underset{\sim}{\beta}} \right\}$$

Observe que

$$\sum_{i=1}^{n} \hat{A}_{0i}\, e^{\underline{X}_i' \underline{\beta}} = \sum_{i=1}^{n} (\hat{\alpha}_{01} + \hat{\alpha}_{02} + \dots + \hat{\alpha}_{0i})\, e^{\underline{X}_i' \underline{\beta}}$$

$$= (\hat{\alpha}_{01} + \hat{\alpha}_{02} + \hat{\alpha}_{03} + \hat{\alpha}_{04} \qquad)\, e^{\underline{X}_1' \underline{\beta}} +$$

$$(\hat{\alpha}_{01} + \hat{\alpha}_{02} \qquad)\, e^{\underline{X}_2' \underline{\beta}} +$$

$$(\hat{\alpha}_{01} + \hat{\alpha}_{02} + \hat{\alpha}_{03} + \hat{\alpha}_{04} + \hat{\alpha}_{05} \qquad)\, e^{\underline{X}_3' \underline{\beta}} + \dots \quad , \text{ i.e.,}$$

$$\sum_{i=1}^{n} \hat{A}_{0i}\, e^{\underline{X}_i' \underline{\beta}} = \sum_{j \in R(t_{(1)})} \hat{\alpha}_{01}\, e^{\underline{X}_j' \underline{\beta}} + \sum_{j \in R(t_{(2)})} \hat{\alpha}_{02}\, e^{\underline{X}_j' \underline{\beta}} + \dots + \sum_{j \in R(t_{(L)})} \hat{\alpha}_{0L}\, e^{\underline{X}_j' \underline{\beta}}$$

$$= \hat{\alpha}_{01} \sum_{j \in R(t_{(1)})} e^{\underline{X}_j' \underline{\beta}} + \hat{\alpha}_{02} \sum_{j \in R(t_{(2)})} e^{\underline{X}_j' \underline{\beta}} + \dots + \hat{\alpha}_{0L} \sum_{j \in R(t_{(L)})} e^{\underline{X}_j' \underline{\beta}}$$

$$= \left(\frac{1}{\sum_{j \in R(t_{(1)})} \cancel{e^{\underline{X}_j' \underline{\beta}}}} \right) \sum_{j \in R(t_{(1)})} \cancel{e^{\underline{X}_j' \underline{\beta}}} + \dots + \left(\frac{1}{\sum_{j \in R(t_{(L)})} \cancel{e^{\underline{X}_j' \underline{\beta}}}} \right) \sum_{j \in R(t_{(L)})} \cancel{e^{\underline{X}_j' \underline{\beta}}} = L$$

Portanto,

$$L(\underline{\beta}) = \left(\prod_{i=1}^{L} \frac{e^{\underline{X}_{(i)}' \underline{\beta}}}{\sum_{k \in R(t_{(i)})} e^{\underline{X}_k' \underline{\beta}}} \right) \cdot e^{-L} \propto \prod_{i=1}^{L} \frac{e^{\underline{X}_{(i)}' \underline{\beta}}}{\sum_{k \in R(t_{(i)})} e^{\underline{X}_k' \underline{\beta}}} \quad , \text{ a f.v. parcial de Cox.}$$

O logaritmo da função de verossimilhança é dada por

$$\ell(\underline{\beta}) = \log L(\underline{\beta}) = \sum_{i=1}^{L} \log \left(\frac{e^{\underline{X}_{(i)}' \underline{\beta}}}{\sum_{k \in R(t_{(i)})} e^{\underline{X}_k' \underline{\beta}}} \right)$$

$$= \sum_{i=1}^{L} \left[\underline{X}_{(i)}' \underline{\beta} - \log \left(\sum_{k \in R(t_{(i)})} e^{\underline{X}_k' \underline{\beta}} \right) \right]$$

O estimador de máxima verossimilhança parcial é obtido da maneira usual.

$$U(\underline{\beta}) = \frac{\partial \ell(\underline{\beta})}{\partial \underline{\beta}} = \sum_{i=1}^{L} \underline{X}_{(i)} - \sum_{i=1}^{L} \left(\frac{\sum_{j \in R(t_{(i)})} \underline{X}_j\, e^{\underline{X}_j' \underline{\beta}}}{\sum_{j \in R(t_{(i)})} e^{\underline{X}_j' \underline{\beta}}} \right)$$

Podemos ainda escrever

$$U(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \delta_i \left(\underset{\sim}{X}_i - \left(\sum_{j \in R(t_i)} \underset{\sim}{X}_j' e^{\underset{\sim}{X}_i' \underset{\sim}{\beta}} \right) \Big/ \left(\sum_{j \in R(t_i)} e^{\underset{\sim}{X}_j' \underset{\sim}{\beta}} \right) \right).$$

O método de Newton-Raphson é muito usado para a obtenção de $\hat{\beta}$.

Análise de Sobrevivência

Aula 20 12/05/2016

Função de Verossimilhança Parcial de Cox.

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{k} \frac{e^{\underset{\sim}{X}'_{(i)}\underset{\sim}{\beta}}}{\sum_{l \in R(t_{(i)})} e^{\underset{\sim}{X}'_{l}\underset{\sim}{\beta}}},$$

em que $t_{(1)} < t_{(2)} < \dots < t_{(k)}$ são os tempos de falha.

Aproximações da Verossimilhança Parcial na Presença de Empates

1ª Aproximação: Breslow (1974)

Sejam $t^*_{(1)} \leq t^*_{(2)} \leq \dots \leq t^*_{D^*}$ os instantes ordenados de falha (note que podemos ter empates)

Fazemos

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{D^*} \left(\frac{e^{\underset{\sim}{X}'_{(i)}\underset{\sim}{\beta}}}{\sum_{j \in R(t_{(i)})} e^{\underset{\sim}{X}'_{j}\underset{\sim}{\beta}}} \right) \qquad (*)$$

Sejam $t_{(1)} < t_{(2)} < \dots < t_D$ os instantes de falha distintos.

Sejam $d_1, d_2, \dots, d_D$ o número de falhas em cada instante.

A função de verossimilhança $(*)$ pode ser reescrita como

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{D} \frac{e^{\underset{\sim}{X}'_{(i1)}\underset{\sim}{\beta}} e^{\underset{\sim}{X}'_{(i2)}\underset{\sim}{\beta}} \dots e^{\underset{\sim}{X}'_{(id_i)}\underset{\sim}{\beta}}}{\left[\sum_{j \in R(t_{(i)})} e^{\underset{\sim}{X}'_{j}\underset{\sim}{\beta}}\right]^{d_i}}$$

$$= \prod_{i=1}^{D} \frac{e^{\underset{\sim}{s}'_i\underset{\sim}{\beta}}}{\left[\sum_{j \in R(t_{(i)})} e^{\underset{\sim}{X}'_{j}\underset{\sim}{\beta}}\right]^{d_i}}, \quad \text{e} \quad \underset{\sim}{s}_i = \sum_{k \in D_i} \underset{\sim}{X}_k,$$

D_i: conjunto de indivíduos que falharam em $t_{(i)}$.

Essa aproximação funciona bem quando há poucos empates.

2ª Aproximação: EFRON (1977)

EFRON propôs uma aproximação mais próxima a uma verossimilhança "exata".

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{D} \frac{e^{s_i' \underset{\sim}{\beta}}}{\prod_{j=1}^{d_i} \left[\sum_{k \in R(t_{(i)})} e^{x_k' \underset{\sim}{\beta}} - \frac{(j-1)}{d_i} \sum_{k \in D_i} e^{x_k' \underset{\sim}{\beta}} \right]}.$$

Exemplo: Seja $t^{\circ}_{(4)}$ um instante de falha.

Assuma.

- 5 indivíduos em risco em $t^{\circ}_{(4)}$
- 2 falhas em $t^{\circ}_{(4)}$ → indivíduos 1 e 2 falharam.

Denominador da aprox. de EFRON.

$$Den = \left(\sum_{k \in R(t^{\circ}_{(4)})} e^{x_k' \underset{\sim}{\beta}} - 0 \right) \left(\sum_{k \in R(t^{\circ}_{(4)})} e^{x_k' \underset{\sim}{\beta}} - \frac{1}{2} \left(e^{x_1' \underset{\sim}{\beta}} + e^{x_2' \underset{\sim}{\beta}} \right) \right)$$

3ª alternativa: Kalbfleisch, pg 105. Constrói-se a verossim. para cada uma das $D_i!$ possibilidade e extrai-se uma média aritmética.

4ª alternativa: Cox. É a verossimilhança parcial obtida com base em um modelo discreto.

Klein, pg. 259.

Algumas correções são feitas e obtém-se

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{D} \frac{e^{s_i' \underset{\sim}{\beta}}}{\sum_{q \in Q_i} e^{s_q^{*} \underset{\sim}{\beta}}}, \quad \text{em que}$$

Q_i: conjunto de todos os subconjuntos de $R(t_{(i)})$ com d_i elementos

$s_q^{*} := \sum_{j \in q} x_j$

Exemplos:

i. Se não há empates

Suponha que $R(t_{(i)}) = \{obs. 2, 3, 7, 11\}$, $d_i = 1$. Daí,

$$Q_i = \{\{2\}, \{3\}, \{7\}, \{11\}\} \text{ e}$$

$$\sum_{q \in Q_i} e^{s_q' \beta} = \sum_{j \in R(t_{(i)})} e^{x_j' \beta}$$

ii. Suponha

$R(t_{(i)}) = \{2, 3, 7, 11\}$ e $d_i = 2$. Agora,

$$Q_i = \{\{2,3\}, \{2,7\}, \{2,11\}, \{3,7\}, \{3,11\}, \{7,11\}\} \text{ e}$$

$$\sum_{q \in Q_i} e^{s_q' \beta} = e^{(x_2 + x_3)'\beta} + e^{(x_2 + x_7)'\beta} + e^{(x_2 + x_{11})'\beta} + e^{(x_3 + x_7)'\beta} + e^{(x_3 + x_{11})'\beta} + e^{(x_7 + x_{11})'\beta}$$

Estimação da Função de Risco Acumulada e Função de Sobrevivência

Na 2ª "derivação" da função de verossimilhança parcial (como uma ver. perfilada) chegamos a

$$\hat{\alpha}_{0k} = \frac{1}{\sum_{j \in R(t_{(k)})} e^{x_j' \beta}}.$$

Com isso, obtemos

$$\hat{A}_{0k} = \sum_{s: t_{(s)} \leq t_{(k)}} \left(\frac{1}{\sum_{l \in R(t_{(s)})} e^{x_l' \hat{\beta}}} \right)$$

O Estimador de Breslow para $A_0(t)$ é dado por

$$\hat{A}_0(t) = \sum_{j:\, t_{(j)} \leq t} \left(\frac{d_j}{\sum_{l \in R(t_{(j)})} e^{\underset{\sim}{X}_l' \underset{\sim}{\beta}}} \right)$$

em que

$t_{(1)} < t_{(2)} < \ldots < t_{(D)}$ são os instantes distintos de falha ordenados,

d_j: nº de falhas em $t_{(j)}$

$R(t_{(j)})$: nº de indivíduos em risco em $t_{(j)}$.

Relembre agora que

$$\alpha(t / \underset{\sim}{X}) = \alpha_0(t)\, e^{\underset{\sim}{X}' \underset{\sim}{\beta}}$$

$$\Rightarrow S(t / \underset{\sim}{X}) = \exp\left\{ -A_0(t)\, e^{\underset{\sim}{X}' \underset{\sim}{\beta}} \right\}.$$

$$S(t / \underset{\sim}{X}) = \left(e^{-A_0(t)} \right)^{e^{\underset{\sim}{X}' \underset{\sim}{\beta}}} = \left(S_0(t) \right)^{e^{\underset{\sim}{X}' \underset{\sim}{\beta}}}.$$

Assim,

$$\underline{\underline{\hat{S}(t / X) = \left(\hat{S}_0(t) \right)^{e^{\underset{\sim}{X}' \hat{\underset{\sim}{\beta}}}}, \quad \hat{S}_0(t) = e^{-\hat{A}_0(t)}}}.$$

Observação: Outros estimadores para $A_0(t)$ e $S_0(t)$ foram propostos.

- Kolbfleisch e Prentice (1973)

$$\hat{A}_0(t) = \sum_{t_{(i)} \leq t} \left[1 - \left(1 - \frac{e^{\underset{\sim}{X}_{(i)}' \hat{\underset{\sim}{\beta}}}}{\sum_{l \in R(t_{(i)})} e^{\underset{\sim}{X}_{(l)}' \hat{\underset{\sim}{\beta}}}} \right)^{e^{-\underset{\sim}{X}_i' \hat{\underset{\sim}{\beta}}}} \right], \quad \text{se não há empates.}$$

- Tbm possível para quando há empates - ver livro.

Extensão do Modelo de Cox: Covariáveis Dependentes do Tempo

Suponha que observamos os dados

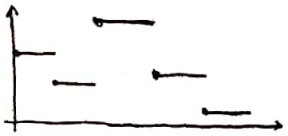
$$\left(Z_i, \delta_i, \{ \underset{\sim}{X}_i(t) : 0 \leq t \leq Z_i \} \right)$$

$Z_i = \min(T_i, C_i)$

$\delta_i = I(T_i \leq C_i)$

$$\underset{\sim}{X}_i(t) = \begin{pmatrix} X_{1i}(t) \\ X_{2i}(t) \\ \vdots \\ X_{pi}(t) \end{pmatrix}$$

<u>Suposição</u>: Conhecemos $X_i(t)$ para todo t t.q. $0 \leq t \leq Z_i$ (período que o indivíduo está em observação)



Exemplos:

1. Estudo de leucemia

- **evento**: óbito ; início: transplante de medula;
- covariáveis fixas: sexo, tipo de leucemia, idade de início da doença
- covariáveis dependentes do tempo: DECH crônica , DECH aguda

$$X_c(t) = \begin{cases} 0, & \text{ñ desenvolveu DECH crônica} \\ 1, & \text{se já desenvolveu.} \end{cases}$$

Aula 21 — Análise de Sobrevivência — 17 de maio de 2016

Extensão do Modelo de Cox: Covariáveis Dependentes do Tempo

$(Z_i, \delta_i, \{X_i(t): 0 \leq t \leq Z_i\})$, $Z_i = \min(T_i, C_i)$ e $\delta_i = I(T_i \leq C_i)$

$$\underset{\sim}{X}_i(t) = \begin{pmatrix} X_1(t) \\ X_2(t) \\ \vdots \\ X_p(t) \end{pmatrix}$$, vetor de covariáveis dependentes do tempo.

$$\alpha(t/X(t)) = \alpha_0(t)\, e^{X(t)\beta}$$

Exemplo: Exposição ao fumo x câncer de pulmão

Evento. ocorrência do câncer de pulmão.

T: medido em anos

$X(t)$: # de maços (total) consumidos até t.

Suponha que para um determinado indivíduo, tem-se.

- Iniciou o vícios aos 20 anos
- fuma 350 maços por anos

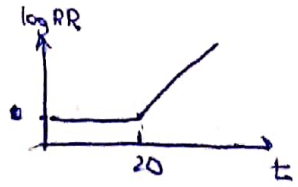
Então

$$X_1(t) = \begin{cases} 0, & \text{se } t < 20 \\ 350(t-20), & \text{se } t \geq 20 \end{cases}$$

Vamos calcular o RR desse indivíduo e um que nunca fumou:

$$RR = \frac{\alpha(t/X_1(t))}{\alpha(t/X(t)=0)} = \frac{\alpha_0(t)\, e^{X'(t)\beta}}{\alpha_0(t)} = e^{X'(t)\beta} = \begin{cases} 1, & \text{se } t < 20 \\ e^{350(t-20)\beta}, & t \geq 20 \end{cases}$$

$$\log RR = \begin{cases} 0, & t < 20 \\ 350(t-20)\beta, & t \geq 0 \end{cases}$$



O RR não é constante no tempo.

Tipos de Variáveis Dependentes no Tempo (Kalbfleisch)

A. Variáveis Extremas:

- Poluição do ar
- Temperatura

B. Fixas no início do estudo e com caminhos conhecidos - Idade.

→ Variáveis Internas

Exemplos: marcadores biológicos

$X(t)$: associada com gravidade ou estadiamento do câncer

$$X(t) = \begin{cases} 1, \text{ muito ruim} \\ 2, \text{ ruim} \\ 3, \text{ pouco deb.} \\ 4, \text{ sem evidências clínicas da doença.} \end{cases}$$

X_T: tratamento.

Função de Verossimilhança Parcial

Modelo de Cox: $\alpha(t/X(t)) = \alpha_0(t)\, e^{X'(t)\beta}$

No modelo: sem covariáveis dependentes do tempo, vimos que

$$l(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \delta_i \left(\underset{\sim}{X}_i' \underset{\sim}{\beta} - \log \sum_{k \in R(t_{(i)})} e^{\underset{\sim}{X}_k' \underset{\sim}{\beta}} \right)$$

Definimos

$$N_i(t) = I(Z_i \leq t, \delta_i = 1)$$

$$Y_i(t) = I(Z_i \geq t)$$

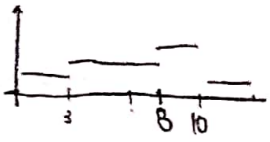
A função de log-verossimilhança pode ser reescrita como

Aula 21 Análise de Sobrevivência, 18 de maio

$$\ell(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \int_0^{\tau} \left(\underset{\sim}{X}_i' \underset{\sim}{\beta} - \log \sum_{j=1}^{n} Y_j(u) e^{\underset{\sim}{X}_j' \underset{\sim}{\beta}} \right) dN_i(u)$$

$$= \int_0^{\tau} \sum_{i=1}^{n} \underset{\sim}{X}_i' \underset{\sim}{\beta} \, dN_i(u) - \int_0^{\tau} \sum_{i=1}^{n} \left[\log\left(\sum_{j=1}^{n} Y_j(u) e^{\underset{\sim}{X}_j' \underset{\sim}{\beta}} \right) \right] dN_i(u)$$

$\underset{\sim}{X}_i'(u)$ $\qquad$ $\underset{\sim}{X}_j'(u)$



No R,

ind	Inic	Fim	δ
1	0	3	0
1	3	8	0
1	8	10	1

Fazendo todo o desenv., chega-se à mesma forma com $\underset{\sim}{X}_j'(u)$ em função de u.

$$\ell(\underset{\sim}{\beta}) = \sum_{i=0}^{n} \delta_i \left(\underset{\sim}{X}_i'(z_i) \underset{\sim}{\beta} \right) - \sum_{i=1}^{n} \delta_i \log \left(\sum_{j \in R(t_{(i)})} e^{\underset{\sim}{X}_j'(z_i) \underset{\sim}{\beta}} \right)$$

ou

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{L} \frac{e^{\underset{\sim}{X}_{(i)}'(t_{(i)}) \underset{\sim}{\beta}}}{\sum_{k \in R(t_{(i)})} e^{\underset{\sim}{X}_k'(t_{(i)}) \underset{\sim}{\beta}}}, \quad \hat{\underset{\sim}{\beta}} \text{ é o valor que maximiza } L(\underset{\sim}{\beta}).$$

Propriedades Assintóticas de $\hat{\beta}$

Defina

$$S^{(0)}(\underset{\sim}{\beta}, t) = \sum_{j=1}^{n} Y_j(t) e^{\underset{\sim}{X}_j'(t) \underset{\sim}{\beta}}.$$

Assim, podemos escrever

$$\ell(\underset{\sim}{\beta}) = \int_0^{\tau} \sum_{i=1}^{n} \underset{\sim}{X}_i'(t) \underset{\sim}{\beta} \, dN_i(t) - \int_0^{\tau} \sum_{i=1}^{n} \left(\log\left(S^{(0)}(\beta, t) \right) \right) dN_i(t)$$

Defina também

$$S^{(1)}(\underset{\sim}{\beta}, t) = \sum_{j=1}^{n} Y_j(t) \underset{\sim}{X}_j(t) e^{\underset{\sim}{X}_j' \underset{\sim}{\beta}} \quad \text{: vetor de } p \text{ linhas}$$

$$S^{(2)}(\underset{\sim}{\beta}, t) = \sum_{j=1}^{n} Y_j(t) \underset{\sim}{X}_j(t)^{\otimes 2} e^{\underset{\sim}{X}_j' \underset{\sim}{\beta}}, \quad V^{\otimes 2} = \underset{\sim}{V} \underset{\sim}{V}^T \text{ matriz } p \times p.$$

Derivando-se $\ell(\underset{\sim}{\beta})$ com relação a $\underset{\sim}{\beta}$, obtemos $U(\underset{\sim}{\beta})$.

$$U(\beta) = \int_0^{\tau} \sum_{i=1}^{n} \left(\underset{\sim}{X}_i(t) - \frac{1}{S^{(0)}(\underset{\sim}{\beta}, t)} \cdot \frac{\partial S^{(0)}(\underset{\sim}{\beta}, t)}{\partial \underset{\sim}{\beta}} \right) dN_i(t)$$

$$= \int_0^{\tau} \sum_{i=1}^{n} \left(\underset{\sim}{X}_i(t) - \frac{S^{(1)}(\underset{\sim}{\beta}, t)}{S^{(0)}(\underset{\sim}{\beta}, t)} \right) dN_i(t)$$

Seja $M_i(t) = N_i(t) - \Lambda_i(t)$ um martingal associado à $N_i(t)$. Podemos escrever

$$U(\underset{\sim}{\beta}) = \int_0^\tau \sum_{i=1}^n \left(\underset{\sim}{X}_i(t) - \frac{S^{(1)}(\underset{\sim}{\beta}, t)}{S^{(0)}(\underset{\sim}{\beta}, t)} \right) (dM_i(t) + d\Lambda_i(t))$$

$$= \sum_{i=1}^n \int_0^t \left(\underset{\sim}{X}_i(t) - \frac{S^{(1)}(\underset{\sim}{\beta}, t)}{S^{(0)}(\underset{\sim}{\beta}, t)} \right) dM_i(t) + \underbrace{\int_0^\tau \sum_{i=1}^n \left(\underset{\sim}{X}_i(t) - \frac{S^{(1)}(\underset{\sim}{\beta}, t)}{S^{(0)}(\underset{\sim}{\beta}, t)} \right) d\Lambda_i(t)}_{\text{II}}$$

Obs: $\mathcal{F}_t = \sigma\{N(u), I(Z \leq u, \delta = 0), \underset{\sim}{X}(u) : 0 \leq u \leq t\}$

$$\text{II} = \sum_{i=1}^n \int_0^\tau \left(\underset{\sim}{X}_i(t) - \frac{S^{(1)}(\underset{\sim}{\beta}, t)}{S^{(0)}(\underset{\sim}{\beta}, t)} \right) \alpha_0(t) \underbrace{e^{\underset{\sim}{X}_i'(t)\underset{\sim}{\beta}} dt}_{d\Lambda_i(t)}$$

$$= \int_0^\tau \left(\overbrace{\sum_{i=1}^n X_i(t) e^{\underset{\sim}{X}_i'(t)\underset{\sim}{\beta}} Y_i(t)}^{S^{(1)}(\beta, t)} \right) \alpha_0(t)\, dt - \int_0^\tau \frac{S^{(1)}(\beta, t)}{S^{(0)}(\beta, t)} \overbrace{\left(\sum_{i=1}^n e^{\underset{\sim}{X}_i'(t)\underset{\sim}{\beta}} Y_i(t) \right)}^{S^{(0)}(\beta, t)} \alpha_0(t)\, dt$$

$$= 0.$$

Portanto,

$$U(\underset{\sim}{\beta}) = \sum_{i=1}^n \int_0^\tau \underbrace{\left(X_i(t) - \frac{S^{(1)}(\underset{\sim}{\beta}, t)}{S^{(0)}(\underset{\sim}{\beta}, t)} \right)}_{\text{proc. previsíveis}} dM_i(t)$$

$\Rightarrow U(\underset{\sim}{\beta})$ é martingal.

- Matriz de informação observada

$$I(\underset{\sim}{\beta}) = -\frac{\partial U(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = \sum_{i=1}^n \int_0^\tau \left(-\frac{S^{(1)}(\beta, t)}{(S^{(0)}(\beta, t))^2} + \frac{S^{(2)}(\beta, t)}{S^{(0)}(\beta, t)} \right) dN_i(t)$$

Aula 21 Análise de Sobrevivência, 17 de maio

$$I(\beta) = \int_0^\tau V(\beta,t) \sum_{i=1}^n dN_i(t).$$

com

$$V(\beta,t) = \frac{S^{(2)}(\beta,t)}{S^{(0)}(\beta,t)} - \left(\frac{S^{(1)}(\beta,t)}{S^{(0)}(\beta,t)}\right)^{\otimes 2}$$

Resultados: (Proc. de Variação Previsível)

- β_0 : verdadeiro valor de β
- $\langle U(\beta_0)\rangle(t) = \int_0^\tau V(\beta_0,t) S^{(0)}(\beta,t)\alpha_0(t)\,dt$
- $I(\beta_0) = \langle U(\beta_0)\rangle(\tau) + \int_0^\tau V(\beta_0,t)\,dM_i(t)$
- $E(I(\beta_0)) = E[\langle U(\beta_0)\rangle(\tau)]$

Relembrando:

- convergência $n^{-1/2}U(\beta_0)$,
- consistência de $\hat\beta$
- consistência da matriz observada,
- Expansão de Taylor de $U(\beta_0)$

Em Kalbfleisch, pg 179, as condições nec. para o Teo. Rebolledo são discutidas com mais detalhes.

(C1) Existem funções $s^{(s)}(\beta,t)$, $s=0,1,2$, que satisfazem

a) $\sup_{t\in[0,\tau),\ \beta\in B} \| n^{-1} S^{(s)}(\beta,t) - s^{(s)}(\beta,t)\| \xrightarrow[n\to\infty]{P} 0$

B: vizinhança de β_0 (verdadeiro valor)

$\|A\|$, maior valor absoluto de um escalar, vetor ou matriz.

b) $s^{(0)}(\beta,t)$ é limitado fora de zero para $t\in[0,\tau]$

c) Para $s=0,1,2$, $s^{(s)}(\beta,t)$ é função contínua de β e

$$s^{(1)}(\underset{\sim}{\beta},t) = \frac{\partial s^{(0)}(\underset{\sim}{\beta},t)}{\partial \underset{\sim}{\beta}}$$

$$s^{(2)}(\underset{\sim}{\beta},t) = \frac{\partial^2 s^{(0)}(\underset{\sim}{\beta},t)}{\partial \underset{\sim}{\beta}\,\partial \underset{\sim}{\beta}^T}$$

d) $I(\beta,t) = \int_0^\tau v(\beta,t)\, s^{(0)}(\underset{\sim}{\beta},t)\, \alpha_0(t)\, dt$ é positiva definida, com

$$v(\underset{\sim}{\beta},t) = \frac{s^{(2)}(\underset{\sim}{\beta},t)}{s^{(0)}(\underset{\sim}{\beta},t)} - e(\underset{\sim}{\beta},t)\, e^T(\underset{\sim}{\beta},t).$$

$$e(\underset{\sim}{\beta},t) = \frac{s^{(1)}(\underset{\sim}{\beta},t)}{s^{(0)}(\underset{\sim}{\beta},t)}$$

Resultado: Sob as condições (C1),

$$\hat{\beta} \xrightarrow{P} \underset{\sim}{\beta}.$$

Condição:

C2) $\exists\, \delta > 0$ tal que, com $n \to +\infty$,

$$n^{-1/2} \sup_{i,t} \left(|\underset{\sim}{X}_i(t)|\, Y_i(t)\, I\left(\underset{\sim}{X}_i' \underset{\sim}{\beta} > -\delta \right) \Delta_i(t) \right) \xrightarrow{P} 0.$$

Resultado:

Sob as cond. (C1) e (C2),

- $n^{-1/2}\, U(\beta_0, \tau) \xrightarrow{D} N(0, I)$
- $\dfrac{\mathcal{I}(\hat{\beta})}{n} \xrightarrow{P} I.$

Vamos agora fazer a expansão de Taylor de $U(\hat{\beta})$ em torno de β_0:

$$\cancel{U(\hat{\beta})}^{\,0} \approx U(\underset{\sim}{\beta}_0) + U'(\beta_0)(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}_0)$$

$$U'(\beta_0)(\hat{\beta} - \beta_0) \approx -U(\beta_0)$$

$$\sqrt{n}(\hat{\beta} - \hat{\beta}_0) \approx -(U'(\beta_0))^{-1} \frac{U(\beta_0)}{\sqrt{n}} \cdot n$$

$$\approx (-U'(\beta_0)/n)^{-1} \underbrace{(n^{-1/2})\, U(\beta_0)}_{\xrightarrow{D} N}$$

$$-\frac{U'(\beta_0)}{n} = J(\beta_0) \xrightarrow{P} \Sigma$$

Pelo teorema de Slutzk,

$$\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{D} N(0, \Sigma^{-1}).$$

Σ^{-1} pode ser consistentemente estimada por $(I(\hat{\beta})/n)^{-1}$.

Modelo de Cox Estratificado

"Útil qdo não há prop. dos riscos"

Exemplo: Transplante de medula óssea

Início: transplante

evento: óbito

Covariáveis: sexo, idade, escore de Karnofsky, tratamento, tipo de Tratamento (autólogo x alogênico)

Notação. Suponha que observamos $(Z_i, \delta_i, \underline{X}_i, X_{Ei})$, $i = 1, 2, \dots, n$.

$Z_i = \min(T_i, C_i)$

$\delta_i = I(T_i \leq C_i)$

$\underline{X}_i$ é um vetor de p variáveis

X_{Ei}: Variável que indica estrato (grupo) ao qual a observação pertence (s categorias)

Suponha que a população esteja dividida em s estratos

→ Suposições do modelo de Cox estratificado.

→ Riscos proporcionais dentro dos estratos

→ funções de risco basal diferentes para cada estrato.

Assim, o modelo de Cox estratificado é dado por

$$\alpha_j(t \mid \underline{X}) = \alpha_{0j}(t)\, e^{\underline{X}'\underline{\beta}}, \quad j = 1, \dots, s,$$

em que

$\alpha_{0j}(t)$: função de risco basal para observações do estrato j

$\underline{\beta}$: vetor de parâmetros desconhecidos

Observação: Note que $\underline{\beta}$ é comum a todos os estratos.

Exemplo: Considere $p=2$ e uma cov. binária

$$X = \begin{cases} 1, & \text{se trat. novo} \\ 0, & \text{se trat. padrão} \end{cases}$$

Considere o modelo

$$\alpha_j(t/X) = \alpha_{0j}(t)\, e^{X\beta}, \quad j=1,2.$$

Estrato 1

$$RR_1 = \frac{\alpha_1(t/X=1)}{\alpha_1(t/X=0)} = \frac{\alpha_{01}(t)\, e^{1\beta}}{\alpha_{01}(t)\, e^{0}} = e^{\beta}$$

Estrato 2

$$RR_2 = \frac{\alpha_2(t/X=1)}{\alpha_2(t/X=0)} = \frac{\alpha_{02}(t)\, e^{\beta}}{\alpha_{02}(t)\, e^{0}} = e^{\beta}$$

Risco relativo de pacientes de estratos diferentes:

$$RR_{12} = \frac{\alpha_1(t/X=1)}{\alpha_2(t/X=0)} = \frac{\alpha_{01}(t)\, e^{\beta}}{\alpha_{02}(t)} = e^{\beta}\, \frac{\alpha_{01}(t)}{\alpha_{02}(t)}.$$

"Uma alternativa: modelo para métricas (ideal) ao encontrar semiparamétrico e alternar tempo para comparar grupos"

que não é constante em t.

→ Os riscos <u>não</u> são proporcionais

Função de Verossimilhança Parcial

Ideia é semelhante ao caso usual.

Diferença: constrói-se o conjunto de risco de pacientes em cada estrato.

A função de verossimilhança parcial fica dada por

$$L(\underline{\beta}) = \prod_{g=1}^{s} L_g(\underline{\beta}) = \prod_{g=1}^{s} \left(\prod_{i=1}^{d_g} \frac{e^{X_{(i)}\beta}}{\sum_{l \in R_i^{(g)}} e^{X_l \beta}} \right)$$

$R_k^{(g)}$: grupo de indivíduos em risco em $t_{(k)}^g$, $g = 1, \dots, s$; $k = 1, \dots, L(g)$

$L(g)$: nº de instantes de falha distintos no estrato g.

$X_{(k)g}$: vetor de covariáveis de observação que falhou em $t_{(k)}^g$.

A função de log-verossimilhança parcial é

$$l(\underset{\sim}{\beta}) = \frac{\partial L(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = \sum_{g=1}^{s} l_g(\underset{\sim}{\beta})$$

O E.M.V.P. é ~~dado~~ obtido maximizando-se $l(\underset{\sim}{\beta})$.

Podemos pensar em um modelo

$$\alpha_g(t \mid \underset{\sim}{X}) = \alpha_{0g}(t)\, e^{\underset{\sim}{X}' \underset{\sim}{\beta}_g} \quad , \quad g = 1, \dots, s$$

$$L(\theta) = \prod_{k=1}^{L(1)} \frac{e^{X'_{(k)1}\beta_1}}{\sum_{l \in R_k^{(1)}} e^{X'_{(l)1}\beta_1}} \cdot \prod_{k=1}^{L(2)} \frac{e^{X'_{(k)2}\beta_2}}{\sum_{l \in R_k^{(2)}} e^{X'_{(l)2}\beta_2}} \cdot \dots \cdot \prod_{k=1}^{L(s)} \frac{e^{X'_{(k)s}\beta_s}}{\sum_{l \in R_k^{(s)}} e^{X'_{(l)s}\beta_s}}$$

$$L(\hat{\theta}) = \prod_{g=0}^{s} L_g(\hat{\underset{\sim}{\beta}}_g)$$

$$l(\hat{\theta}) = \sum_{g=1}^{s} l_g(\hat{\underset{\sim}{\beta}}_g)$$

Suponha que queremos testar

H_0: $\beta_1 = \beta_2 = \dots = \beta_s$

H_1: existe alguma diferença

Procedimento:

(1) ajustar o modelo (I) e obter $\hat{\beta}$ e $l_I(\hat{\underset{\sim}{\beta}})$

(2) Ajustar o modelo (II) (ou seja, um modelo de Cox para cada estrato) e obter

$$\ell_g(\hat{\beta}_g) \quad e \quad \ell_{II}(\hat{\theta}) = \ell_1(\hat{\beta}_1) + \hat{\ell}_2(\hat{\beta}_2) + \dots + \hat{\ell}_s(\hat{\beta}_s)$$

Calculamos a estatística do teste da RV:

$$\Lambda = -2\left(\ell_I(\hat{\beta}) - \ell_{II}(\hat{\theta})\right) \xrightarrow[\text{sob } H_0]{D} \chi^2_{(s-1)p}$$

Teste de Wald.

$$\alpha_g(t \mid \underset{\sim}{X}) = \alpha_{0g}(t)\, e^{\underset{\sim}{X}' \underset{\sim}{\beta}_g}, \quad g = 1, \dots, s$$

$H_0: \beta_{gk} = \beta_{\ell k}, \quad \ell \neq g.$

$H_A: \beta_{gk} \neq \beta_{\ell k}, \quad \ell, g = 1, \dots, s, \quad \ell \neq g.$

$$\underset{\sim}{\beta}_g = \begin{pmatrix} \beta_{g1} \\ \beta_{g2} \\ \vdots \\ \beta_{gp} \end{pmatrix}, \quad g = 1, \dots, s$$

$$\hat{\underset{\sim}{\beta}}_g \overset{aprox}{\sim} N(\underset{\sim}{\beta}_g, I^{-1}(\underset{\sim}{\beta}_g))$$

$$\hat{\beta}_{gk} \overset{aprox}{\sim} N(\beta_{gk}, V_{gk}) \quad e \quad \hat{\beta}_{\ell k} \overset{aprox.}{\sim} N(\beta_{\ell k}, V_{\ell k})$$

$$Z_k = \frac{\hat{\beta}_{gk} - \hat{\beta}_{\ell k}}{\sqrt{\hat{V}_{gk} + \hat{V}_{\ell k}}} \underset{\text{sob } H_0}{\overset{aprox}{\sim}} N(0,1), \quad \text{ou} \quad Z_k^2 \sim \chi^2_{(1)} \text{ sob } H_0.$$

Podemos ter um modelo da forma

$$\alpha_g(t \mid \underset{\sim}{X}) = \alpha_{0g}(t) \exp\left(\underset{\sim}{X}_f' \underset{\sim}{\beta}_f + \underset{\sim}{X}_I' \underset{\sim}{\beta}_g\right), \quad g = 1, 2, \dots, s$$

$\underset{\sim}{X}_f$: vetor de covariáveis com mesmo efeito em todos os estratos

$\underset{\sim}{X}_I$: vetor de covariáveis com efeitos diferentes em cada estrato.

$$\Theta' = (\underset{\sim}{\beta}_f', \underset{\sim}{\beta}_1, \underset{\sim}{\beta}_2, \dots, \underset{\sim}{\beta}_s).$$

Exercício: Como construir $\underset{\sim}{W}_i$ e $\underset{\sim}{W}$ matriz de planej. de forma que $\underset{\sim}{W}_i' \Theta = \underset{\sim}{X}_{fi}' \underset{\sim}{\beta}_f + \underset{\sim}{X}_{Ii}' \underset{\sim}{\beta}_g$, g fixado

Estimação da Função de Sobrevivência

Temos que

$$S_g(t/\underset{\sim}{x}) = \exp\left\{ -\int_0^t \alpha_{0g}(u)\, e^{\underset{\sim}{x}'\underset{\sim}{\beta}}\, du \right\}$$

$$S_g(t/\underset{\sim}{x}) = \exp\left\{ -A_{0g}(t)\, e^{\underset{\sim}{x}'\underset{\sim}{\beta}} \right\} = \left(e^{-\Lambda_{0g}(t)} \right)^{e^{\underset{\sim}{x}'\underset{\sim}{\beta}}}$$

$$= \left(S_{0g}(t) \right)^{e^{\underset{\sim}{x}'\underset{\sim}{\beta}}}, \quad g = 1, \dots, s.$$

Assim,

$$\hat{S}_g(t/\underset{\sim}{x}) = \left(\hat{S}_{0g}(t) \right)^{e^{\underset{\sim}{x}'\hat{\underset{\sim}{\beta}}}}$$

com

$$\hat{S}_{0g}(t) = \exp\left\{ -\sum_{k:\, t^{(g)}_{(k)} \le t} \left(\frac{d^{g}_{(k)}}{\sum_{j \in R^{(g)}_k} e^{\underset{\sim}{x}_j'\underset{\sim}{\beta}}} \right) \right\}$$

$d^{(g)}_k$: nº de falhas no estrato g em $t^{g}_{(k)}$

$R^{(g)}_k$: indivíduos em risco em $t^{g}_{(k)}$

$t^{g}_{(1)} < t^{g}_{(2)} < \dots < t^{g}_{(k_g)}$: instante de falha distintos do estrato g.

Adequação do Modelo de Cox

→ Qualidade geral do ajuste

→ Verificação da suposição de proporcionalidade dos riscos

→ Detecção de pontos atípicos

Métodos Gráficos Para Avaliação da Suposição de Proporcionalidade dos Riscos

(A) Método Gráfico Descritivo

Ideia: estudar as curvas de sobrevivência em cada comb. de covariáveis

Ex: X_1 = sexo (F/M)

X_4 = Trat. (Novo / Padrão)

Temos 4 combinações.

$$\begin{pmatrix} FN & MN \\ FP & MP \end{pmatrix}$$

Assuma válida a suposição de proporcionalidade dos riscos. Então vale

$$S(t/\underset{\sim}{X}) = \exp\left\{ -\int_0^t \alpha_0(t)\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}}\, du \right\} = e^{-A_0(t)\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}}}$$

$$\rightarrow \log S(t/\underset{\sim}{X}) = -A_0(t)\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}} \rightarrow \log(-\log S(t/\underset{\sim}{X})) = \underset{\sim}{X}'\underset{\sim}{\beta} + \log(A_0(t))$$

Esse resultado justifica o procedimento:

(i) Obter as estimativas de KM para cada ~~um dos riscos~~ grupo das combinações das covariáveis (ou estimativa de $A(t)$ via Nelson-Aalen)

(ii) Calculamos $\log(-\log \hat{S}_g(t))$, $g = 1, \dots, s$

(iii) Construímos, em um msm gráfico, as curvas

$\log(-\log \hat{S}_g(t))$ versus t.

Se a suposição de proporcionalidade dos riscos é válida, as curvas devem ser aproximadamente "paralelas".

Se $\underset{\sim}{\theta} = \begin{pmatrix} \underset{\sim}{\beta_f} \\ \underset{\sim}{\beta_1} \\ \vdots \\ \underset{\sim}{\beta_s} \end{pmatrix}$ Como construir $\underset{\sim}{W_i}$ e $\underset{\sim}{W}$ matriz de planejamento, de forma que obtenhamos:

$$\underset{\sim}{W_i'}\underset{\sim}{\theta} = \underset{\sim}{X_{fi}'}\underset{\sim}{\beta_s} + \underset{\sim}{X_{Ii}'}\underset{\sim}{\beta_g} \quad , g \text{ fixada.}$$

É se trata de manipular a matriz de planejamento!

Estimação de Função de Sobrevivência

Temos que $S_g(t|\underset{\sim}{X}) = \exp\left\{-\int_0^t \alpha_{0g}(u).e^{\underset{\sim}{X}'\underset{\sim}{\beta}}du\right\}$

$$S_g(t|\underset{\sim}{X}) = \exp\left\{-A_{0g}(t)e^{\underset{\sim}{X}'\underset{\sim}{\beta}}\right\} = \left(e^{-A_{0g}(t)}\right)^{e^{\underset{\sim}{X}'\underset{\sim}{\beta}}} = \left(S_{0g}(t)\right)^{e^{\underset{\sim}{X}'\underset{\sim}{\beta}}} \quad , g=1,2,\dots,s.$$

Assim, $\hat{S}_g(t|\underset{\sim}{X}) = \left(\hat{S}_{0g}(t)\right)^{e^{\underset{\sim}{X}'\hat{\underset{\sim}{\beta}}}}$, com: $\hat{S}_{0g}(t) = \exp\left\{-\sum_{k:t^g_{(k)}\le t}\left(\frac{d^g_{(k)}}{\sum_{j\in R^{(g)}_k} e^{\underset{\sim}{X}_j'\underset{\sim}{\beta}}}\right)\right\}$

$d_k^{(g)}$: nº de falhas no estrato g em $t^g_{(k)}$.

$R_k^{(g)}$: indivíduos em risco em $t^g_{(k)}$.

$t^g_{(1)} \le t^g_{(2)} \le \dots \le t^g_{(k_g)}$: instantes de falha distintos do estrato g.

Adequação do modelo de Cox:
- → Qualidade geral do ajuste.
- → Verificação da suposição de proporcionalidade dos riscos.
- → Detecção de pontos atípicos.

Métodos gráficos para Avaliação da suposição de proporcionalidade dos Riscos:

(A) **Método gráfico Descritivo:** Ideia: Estudar as curvas de sobrevivência em % combinação de covariáveis.

Exemplo:
X_1 = sexo (F/M)
X_2 = Trat (Novo/padrão), temos 4 combinações: $\begin{pmatrix} FN & MN \\ FP & MP \end{pmatrix}$

Assuma válida a suposição de proporcionalidade dos riscos. Então vale:

$$S(t|\underset{\sim}{X}) = \exp\left\{-\int_0^t \alpha_0(t)e^{\underset{\sim}{X}'\underset{\sim}{\beta}}du\right\} = e^{-A_0(t)e^{\underset{\sim}{X}'\underset{\sim}{\beta}}}$$

$$\Rightarrow \log S(t|\underset{\sim}{X}) = -A_0(t)\,e^{\underset{\sim}{X}'\underset{\sim}{\beta}}$$

$$\Rightarrow \log(-\log S(t|\underset{\sim}{X})) = \underset{\sim}{X}'\underset{\sim}{\beta} + \log(A_0(t))$$

Esse resultado justifica o procedimento:

(i) Obter as estimativas de KM para 1 grupo das combinações das covariáveis (ou estimativas de $A(t)$ via Nelson-Aalen).

(ii) Calculamos $\log(-\log \hat{S}_g(t))$, $g=1,2,\dots,s$.

(iii) Construímos, em um mesmo gráfico as curvas: $\log(-\log \hat{S}_g(t))$ versus t.
Se a suposição de proporcionalidade dos riscos é válida, as curvas devem ser aproximadamente "paralelas".

Aula 23 24/05/2016

(B) Valores Esperados por valores observados (Contexto semelhante ao caso anterior).

Procedimento:

(i) Obter as curvas de KM para cada nível da combinação das covariáveis (sem suposição de proporcionalidade dos riscos) → valores "observados".

(ii) Estimar $S(t|x)$ pelo modelo de Cox (com a suposição de prop. dos riscos). Devemos usar os mesmos grupos do item (i) como covariáveis.

Se a suposição de proporcionalidade dos riscos for razoável, então devemos ter:

$$\hat{S}_g^{KM}(t) \approx \hat{S}(t|X=g), \quad g=1,2,\dots,s$$

(iii) Colocamos em um mesmo gráfico $\hat{S}_g^{KM}(t)$ e $\hat{S}(t|X=g)$ e comparamos as curvas. Se temos uma grande diferença entre elas (diferenças g (as septas) então não são de riscos proporcionais (os hubem cruzamento de as curvas, tempo).

(C) Usando estratificação. Seja X o vetor de covariáveis. Escrevemos: $\underset{\sim}{X} = (X_1, X_2^T)^T$. Se X_1 é variável contínua, devemos categorizá-la. Assumir que X_1 tenha "s" categorias.

Procedimento:

(i) Ajustar um modelo de Cox estratificado por X_1, com covariáveis $\underset{\sim}{X_2}$. Seja $\hat{A}_{0g}(t)$ a função de risco basal acumulada estimada para o estrato g; $g=1,2,\dots,s$.

$$\hat{A}_{0g}(t) = \sum_{j:\, t_{(j)} \le t} \frac{d_j^{(g)}}{\sum_{l \in R_g(t_{(j)})} e^{\underset{\sim}{X_2'}\underset{\sim}{\beta}}}, \quad g=1,2,\dots,s$$

(ii) Assuma válida a suposição de proporcionalidade dos riscos para X_1:

$$\alpha(t|\underset{\sim}{X}) = \alpha_0(t) \cdot e^{\underset{\sim}{X'}\underset{\sim}{\beta}} = \alpha_0(t) e^{\underset{\sim}{X_1'}\underset{\sim}{\beta_1}} e^{\underset{\sim}{X_2'}\underset{\sim}{\beta_2}} \dots, \quad g=1,2,\dots,s$$

$$\Rightarrow A(t|\underset{\sim}{X}) = \underbrace{A_0(t) e^{\underset{\sim}{X_1'}\underset{\sim}{\beta_1}}}_{\uparrow} \cdot e^{\underset{\sim}{X_2'}\underset{\sim}{\beta_2}}, \quad \underset{\sim}{\beta_1} = \begin{pmatrix} \beta_{1(2)} \\ \beta_{1(3)} \\ \vdots \\ \beta_{1(s)} \end{pmatrix}$$

idéia: Comparar as "g" curvas estimadas $\hat{A}_{0g}(t)$ com uma dele tipo por exemplo para $g=1$.

Se ajustamos um modelo de Cox estratificado por X_1, devemos ter:

$$A(t|X_1=1, X_2) = A_{01}(t) e^{\underset{\sim}{X_2'}\underset{\sim}{\beta}}$$
$$A(t|X_1=2, X_2) = A_{02}(t) e^{\underset{\sim}{X_2'}\underset{\sim}{\beta}}$$
$$\vdots$$
$$A(t|X_1=s, X_2) = A_{0s}(t) e^{\underset{\sim}{X_2'}\underset{\sim}{\beta}}, \quad \text{com} \quad A_{0g}(t) = A_0(t) e^{\beta_{1(g)}}, \quad g=1,2,\dots,s$$

$\beta_{1(1)} = 0$.

Observe que se a suposição de proporcionalidade dos riscos é válida para cada X_1, então: → $\log \hat{A}_{0g}(t) \approx \beta_{1(g)} + \log A_0(t)$. Podemos portanto construir o gráfico:

$\log \hat{A}_{01}(t)$
$\log \hat{A}_{02}(t)$ versus t. Assim se a suposição de prop. de riscos é válida, as curvas devem ser paralelas.
$\vdots$
$\log \hat{A}_{0s}(t)$

Alternativamente, podemos construir gráficos:

$\log(\hat{A}_{02}(t)) - \log(\hat{A}_{01}(t))$
$\log(\hat{A}_{03}(t)) - \log(\hat{A}_{01}(t))$ versus t.
$\vdots$
$\log(\hat{A}_{0s}(t)) - \log(\hat{A}_{01}(t))$

→ Se a suposição de riscos proporcionais é válida então as curvas devem ser paralelas ao eixo x.

* Resíduos No modelo de Cox:

① **Resíduo de Cox-Snell** - Semelhante ao caso paramétrico.

Relembre o resultado: T v.a. contínua com função de distribuição acumulada $F(t)$, $t \geq 0$.

Defina: $V = -\log S(T) = A(T)$, $S(t) = 1 - F(t)$, então: $V \sim \text{expon}(1)$.

No Modelo de Cox,

$$\alpha(t|X_i) = \alpha_0(t)\, e^{X_i'\beta}$$

$$\Rightarrow A(t|X_i) = A_0(t)\, e^{X_i'\beta}$$

Pelo resultado, devemos ter:

$$A_0(T_i)\, e^{X_i'\beta} \sim \text{expon}(1)$$

O resíduo de Cox-Snell é definido por:

$$\hat{e}_i = \hat{A}_0(z_i)\, e^{X_i'\hat{\beta}}, \quad z_i = \min(T_i, C_i)$$

$\hat{\beta}$ = estimador de máx. verossimil. parcial de β.

$\hat{A}_0(t)$ = estimador de Breslow de $A_0(t)$.

Propriedades dos resíduos de Cox-Snell:

- $e_i \geq 0$
- Se o modelo for adequado: $\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n$ deve se comportar aproximadamente como uma amostra censurada da. Variável aleatória com distribuição exp(1).

(Se não houver censuras → o ajuste, se deve ajustar os resíduos a um exponencial e ver se realmente está próximo a 1), para saber se o ajuste está bom ou não.

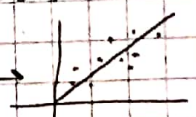
Procedimento seguido:

(1) Calcular os resíduos $\hat{e}_i$, obtemos: $(\hat{e}_1, \delta_1), (\hat{e}_2, \delta_2), \ldots, (\hat{e}_n, \delta_n)$

(2) Obter o estimador Nelson-Aalen dos resíduos $(\hat{e}_i, \delta_i)$, $i = 1, 2, \ldots, n$.

(3) Construir o gráfico de:

$$\hat{A}_e(e_i) \text{ vs. } e_i \longrightarrow$$



Se o modelo for adequado, os pontos devem se aproximar de uma reta passando pela origem e inclinação 1.

Observação: ⊙ Para um modelo de Cox estratificado, os resíduos de Cox-Snell são dados por:

$$\hat{e}_j = \hat{A}_{0g}(z_j)\, e^{X_j'\hat{\beta}}$$

, em que: $\hat{A}_{0g}(\cdot)$ é o estimador da função de risco basal do estrato g (grup. ao qual a observação j pertence).

⊙ Para o modelo de Cox com covariáveis dep. do tempo, tem-se que: $\hat{e}_j = \hat{A}_0(z_j)\, e^{X_j'(z_j)\hat{\beta}}$

(Crítica de resíduo Cox-Snell): Las mismas que para modelos paramétricos. (Ver Problemas de resíduos Cox-Snell) (el verdadero residuo es mayor que $\hat{e}_j$ por tanto se descarta observ. censurada).

Principalmente: Verificar a qualidade geral do ajuste. ¡OJO!

② **Resíduo Martingal**. Relembre que definimos: $N_i(t) = I(z_i \leq t, \delta_i = 1)$

$Y_i(t) = I(z_i \geq t)$, $i = 1, 2, \ldots, n$. Vimos ainda que:

$$M_i(t) = N_i(t) - \int_0^t \alpha_0(u)\, e^{X_i'(u)\beta}\, Y_i(u)\, du$$

é um martingal de média zero. O processo associado ao resíduo martingal é dado por:

$$\hat{M}_i(t) = N_i(t) - \int_0^t Y_i(u)\, e^{X_i'\hat{\beta}}\, d\hat{A}_0(u)$$

, em que: $\hat{A}_0(\cdot)$ é o estimador de Breslow dado por:

$$\hat{A}_0(t) = \sum_{j:\, t_{(j)} \leq t} \frac{d_j}{\sum_{i \in R(t_{(j)})} e^{X_i'\hat{\beta}}}$$

: sem covariáveis dependente do tempo.

O Resíduo Martingal é definido como o processo $\hat{M}_i(t)$ no instante final do estudo, ou seja,

$$\hat{M}_i(\tau) = \hat{M}_i(\infty) = N_i(\tau) - \int_0^\tau Y_i(u)\, e^{X_i'\hat{\beta}}\, d\hat{A}_0(u)$$

. No caso do modelo sem covariáveis dependentes do tempo, Chegamos a:

$$\hat{M}_i = \delta_i - e^{X_i'\hat{\beta}} A_0(z_i), \quad i = 1, 2, \ldots, n, \quad z_i = \min(T_i, C_i)$$

Propriedades do resíduo Martingal

① $E(M_i)=0$ (quando M_i é avaliado no verdadeiro valor de β e $\lambda_0(t)$).

② $\sum_{i=1}^{n} \hat{M}_i = 0$

③ $Cov(M_i, M_j)=0$, $i \neq j$.

④ $Cov(\hat{M}_i, \hat{M}_j) < 0$, porém $cov(\hat{M}_i, \hat{M}_j) \xrightarrow[n \to +\infty]{} 0$

⑤ $\hat{M}_i \in (-\infty; 1)$, $i=1,2,\dots,n$

⑥ $\hat{M}_i$ tem distribuição assimétrica.

Principal uso: Encontrar a forma funcional de uma covariável. Suponha que $\underset{\sim}{X}$ pode ser particionada em (X_1, X^*), em que X_1: escalar, X^*: vetor de (p-1) linhas.

Ideia: Assuma que queremos obter a forma funcional de X_1, ou seja, queremos encontrar $f(X_1)$ que melhor explique o efeito de X_1 no risco de ocorrência do evento. Se o modelo de Cox é verdadeiro, temos que:

$$\alpha(t|X_1, X^*) = \lambda_0(t)\, e^{f(X_1)}\, e^{X^{*\prime}\beta}$$ em que:

$f(.)$ é uma função de X_1, para explicar o seu efeito no risco do evento.

Suponha que um modelo SEM a covariável X_1 seja ajustado. Neste caso, o resíduo martingal será dado por:

$$\hat{M}_1^* = \delta_1 - \hat{A}_0(z_i)\, e^{X_i^{*\prime}\hat{\beta}^*}$$ → "Resíduo martingal do modelo errado"

$\alpha(t|X^*) = \lambda_0(t)\, e^{X^{*\prime}\beta^*}$ ← modelo errado

então por Therneau e Grambsch, demonstram que: $E(\hat{M}_i^*/X_{1i}) \approx [f(X_{1i}) - \log(g^*)]\,\omega$,

em que: ω = prob de eventos / n.

g^* = alguma média de $\exp(f(X_{1i}))$

Podemos escrever: $E(\hat{M}_i^*(X_{1i}) \approx K_1 f(X_{1i}) - K_2$, Assim o gráfico de $\hat{M}_i^*$ vs. X_{1i} deve fornecer informação sobre $f(.)$.

Procedimento:
- Ajustar um modelo de Cox sem a covariável em estudo
- Obter o resíduo martingal desse modelo.
- Construir o gráfico de $\hat{M}_i^*$ vs. X_{1i}. → (Caso categórica)

Em geral, coloca-se uma curva suavizada no gráfico para avaliar a tendência. (Se Linear) mais dos tem efeito então os pontos devem ser paralelos ao eixo x!

31/05/16: **Resíduo Deviance:** (Livro de Klein)

- São utilizados para detecção de pontos atípicos. Correspondem a uma transformação do resíduo martingal. São definidos por:

$$\hat{D}_i = \text{Sinal}(\hat{M}_i)\left[-2(\hat{M}_i + \delta_i \log(\delta_i - \hat{M}_i))\right]^{1/2}$$

Procedimento Sugerido

Construção do gráfico de: $\hat{D}_i$ vs. $\underset{\sim}{X}_i'\hat{\beta}$ (preditor linear)

Pontos atípicos → devem apresentar valor absoluto desse resíduo alto.

④ Resíduo Escore

Relembre que o logaritmo da função de verossimilhança parcial é dado por:

$$\ell(\underset{\sim}{\beta}) = \int_0^{\tau} \sum_{i=1}^{n} \left[\underset{\sim}{X}_i'(u)\underset{\sim}{\beta} - \log\left(\sum_{j=1}^{n} Y_j(u)\, e^{\underset{\sim}{X}_j'(u)\underset{\sim}{\beta}} \right) \right] dN_i(u)$$

Aula 29

Análise de Sobrevivência

31/05/2016

Resíduos Cox-Snell

Resíduo Martingal

$$\hat{M}_i = \delta_i - \int_0^{+\infty} Y_i(u) e^{X_i'(u)\hat{\beta}} \, d\hat{A}_0(u)$$

No caso sem covariáveis dependentes do tempo

$$\hat{M}_i = \delta_i - \hat{e}_i,$$

$\hat{e}_i$: resíduo de Cox-Snell

$$\hat{e}_i = \hat{A}_0(T_i) \, e^{X_i'\hat{\beta}}$$

Resíduo Deviance

São utilizados para detecção de pontos atípicos. Correspondem a uma transformação do resíduo martingal.

São definidos por

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

Procedimento Sugerido:

Construção do gráfico de

$\hat{D}_i$ versus $\underset{\sim}{X}_i' \underset{\sim}{\beta}$ (preditor linear)

Pontos atípicos: devem apresentar valor absoluto desse resíduo alto.

1

4. Resíduo Escore

Relembre que o log da função de verossimilhança parcial é dado por

$$l(\underset{\sim}{\beta}) = \int_0^{\tau} \sum_{i=1}^{n} \left[\underset{\sim}{X}_i'(u)\underset{\sim}{\beta} - \log\left(\sum_{j=1}^{n} Y_j(u) e^{\underset{\sim}{X}_j'(u)\underset{\sim}{\beta}} \right) \right] dN_i(u)$$

e o vetor escore é dado por

$$\underset{(p\times 1)}{U(\underset{\sim}{\beta})} = \int_0^{\tau} \sum_{i=1}^{n} \left[\underset{\sim}{X}_i(u) - \bar{\underset{\sim}{X}}(\underset{\sim}{\beta}, u) \right] dN_i(u).$$

com

$$\bar{\underset{\sim}{X}}(\underset{\sim}{\beta}, u) = \frac{\sum_{j=1} Y_j(u) \underset{\sim}{X}_j e^{\underset{\sim}{X}_j'\underset{\sim}{\beta}}}{\left(\sum_{j=1}^{n} Y_j(u) e^{\underset{\sim}{X}_j'\underset{\sim}{\beta}} \right)}$$

Vimos que

$$\underset{\sim}{U}(\underset{\sim}{\beta}_0) = \int_0^{\tau} \sum_{i=1}^{n} \left(\underset{\sim}{X}_i(u) - \bar{\underset{\sim}{X}}(\underset{\sim}{\beta}_0, u) \right) dM_i(u)$$

O Processo Escore para o i-ésimo indivíduo é definido por

$$\underset{\sim}{U}_i(\underset{\sim}{\beta}, t) = \int_0^{t} \left(\underset{\sim}{X}_i(u) - \bar{\underset{\sim}{X}}(\underset{\sim}{\beta}, u) \right) dM_i(u),$$

$$\underset{\sim}{U}_i(\underset{\sim}{\beta}, t) = \begin{pmatrix} U_{i1}(t) \\ U_{i2}(t) \\ \vdots \\ U_{ip}(t) \end{pmatrix}.$$

O Resíduo Escore do indivíduo i para a k-ésima covariável é definido por

$$S_{ik}(t) = \int_0^{t} \left(X_{ik}(u) - \bar{X}_k(u) \right) d\hat{M}_i(u)$$

$$\bar{X}_k(u) = \frac{\sum_{j=1}^{n} Y_j(u)\, X_{jk}(u)\, e^{X_j'\hat{\beta}}}{\sum_{j=1}^{n} Y_j(u)\, e^{X_j'\hat{\beta}}}$$

Propriedade: $\sum_{i=1}^{n} S_{ik}(\tau) = 0.$

O Processo Escore para a k-ésima covariável é definido por

$$U_k(t) = \sum_{i=1}^{n} S_{ik}(t)$$

Observação: No caso <u>sem</u> covariáveis dependentes do tempo,

$$U_k(t) = \sum_{j:\, \text{falhas} \le t} \left(X_{jk} - \bar{X}_k(Z_j) \right)$$

"Klein → Mart., Pr Escore
Enrico → Sheff ..."

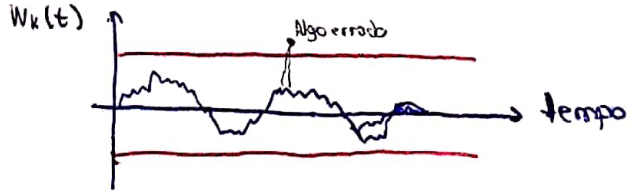
Dois usos principais são sugeridos.

<u>Uso 1</u>: Klein

É motivado pelo resultado:

$$W_k(t) = U_k(t) \,/\, \text{erro padrão}(U_k(t)) \quad \text{: ponte Browniana}$$
$$= U_k(t) \div \text{erro padrão}(\hat{\beta}_k)$$

Conhece-se a dist. de sup (Valor absoluto ($W_k(t)$)). Espera-se



Procedimento: Gráfico de $W_k(t)$ versus t. Se o modelo é adequado, $W_k(t)$ deve se comportar como uma "ponte Browniana". Colocamos no gráfico, em geral, limites

para o sup (valor abs de Brownian bridge). Se o gráfico apresentar picos fora dos limites, o modelo não é adequado.

<u>Uso 2</u>: Detecção de pontos influentes.

Considere a situação sem covariáveis dependentes do tempo. Neste caso,

$$S_{ik} = S_{ik}(\infty) = \delta_i \left(x_{ik} - \bar{x}_k(z_i) \right) - \sum_{t_{(j)} \leq z_i} \left\{ \left[x_{ik} - \bar{x}_k(t_{(j)}) \right] e^{x_i' \hat{\beta}} \times \left(\hat{A}_0(t_{(j)}) - \hat{A}_0(t_{(j-1)}) \right) \right\}$$

com

$$\bar{x}_k(t) = \left(\sum_{h \in R(t)} x_{hk} \, e^{x_h' \beta} \right) \Big/ \left(\sum_{h \in R(t)} e^{x_h' \beta} \right)$$

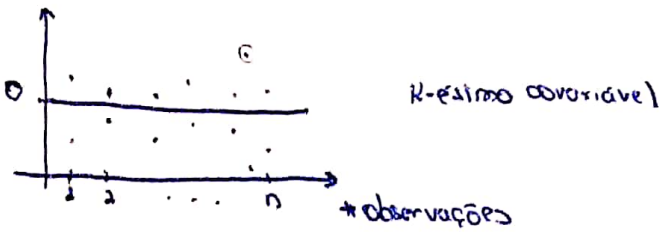
Motivação: Seja $\hat{\beta}_{(i)}$ a estimativa de m.v. parcial de β excluindo-se a i-ésima observação. Observe que se

$(\hat{\beta}_{(i)} - \hat{\beta})$ está próximo de zero,

a i-ésima observação não tem muita influência na estimativa de β. Podemos obter

$$\begin{pmatrix} \hat{\beta} - \hat{\beta}_{(1)} \\ \hat{\beta} - \hat{\beta}_{(2)} \\ \vdots \\ \hat{\beta} - \hat{\beta}_{(n)} \end{pmatrix}$$

O gráfico de $(\hat{\beta}_k - \hat{\beta}_{(i)k})$, $k = 1, \ldots, p$) fornece informação sobre os pontos mais influentes.



Resultado: Pode-se demonstrar que

$$(\underset{\sim}{\beta} - \hat{\underset{\sim}{\beta}}_{(i)}) \approx I^{-1}(\hat{\underset{\sim}{\beta}}) \begin{pmatrix} S_{i1} \\ S_{i2} \\ \vdots \\ S_{ip} \end{pmatrix},$$

em que

$I(\underset{\sim}{\beta})$: informação observada.

Gráficos Sugeridos:

$$\left(I^{-1}(\hat{\beta}) \begin{pmatrix} S_{i1} \\ S_{i2} \\ \vdots \\ S_{ip} \end{pmatrix} \right)$$

versus # observação

(p gráficos)

↓

Resíduo escore padronizado

Resíduos de Cox-Snell; Martingal; Deviance e Escore

Resíduo Schoenfeld

Os resíduos Scho são definidos apenas para as observações que falharam. São definidos por

$$r_{ik} = \delta_i \left(X_{ik} - \bar{X}_k(Z_i) \right),$$

$$\bar{X}_k(Z_i) = \left(\sum_{j \in R(Z_i)} X_{jk} e^{X_j' \beta} \right) \Big/ \sum_{j \in R(Z_i)} e^{X_j' \beta}$$

Observe que temos, nesse caso, um resíduo para cada observação que falhou por covariável. Temos ainda que

$$\sum_i \underset{\sim}{r}_i = 0, \quad \underset{\sim}{r}_i = (r_{i1}, r_{i2}, \ldots, r_{ip})^T$$

Os resíduos Schoenfeld podem ser padronizados. Os resíduos padronizados são dados por

$$\underset{\sim}{r}_i^* = [\hat{Var}(\underset{\sim}{r}_i)]^{-1} \underset{\sim}{r}_i \quad (\text{"scaled residuals"}),$$

$$\hat{Var}(\underset{\sim}{r}_i) \approx (m \, Var(\hat{\beta}))^{-1},$$

(Ref. Colosimos, mas cuidado ~~porque~~ há um "erriinho" lá.)

ou

$$\approx I(\hat{\underset{\sim}{\beta}}) / m$$

m: nº total de falhas. (Grombsch e Therneau (1994))

Uma aproximação para $\underset{\sim}{r}_i^*$ é

$$r_i^* = m\,\widehat{Var}(\hat{\underline{\beta}})\,r_i = m\,I^{-1}(\hat{\underline{\beta}})\,r_i$$

Principal uso: avaliação da suposição de proporcionalidade dos riscos

Motivação: Seja $\underline{X} = (X_1, \underline{X}^*)$ o vetor de covariáveis.

Suponha que o seguinte modelo seja adequado

$$\alpha(t|\underline{X}) = \alpha_0(t)\exp\left\{\underline{X}^*\underline{\beta}^* + X_1(\beta_1 + \gamma g(t))\right\} \qquad (1)$$

em que

$g(t)$: função conhecida do tempo

Observe que o modelo (1) não é um modelo de riscos proporcionais

Observe ainda que

$\gamma = 0 \Rightarrow$ o modelo (1) é de riscos proporcionais

Suponha que agora que ajustamos um modelo de riscos proporcionais de Cox para dados em que a verdadeira estrutura seja dada por (1).

Ajustamos o modelo

$$\alpha(t|\underline{X}) = \alpha_0(t)\,e^{X_1\beta_1 + \underline{X}_2^{*\prime}\beta_2}$$

Seja

r_i^*: resíduo padronizado de Schoenfeld

Calculando com base no modelo de Riscos Proporcionais.

* Resultado: $E(r_{i1}^* + \hat{\beta}_1) = (\beta_1 + \gamma g(t)) = \beta_1(t)$ ←

ou

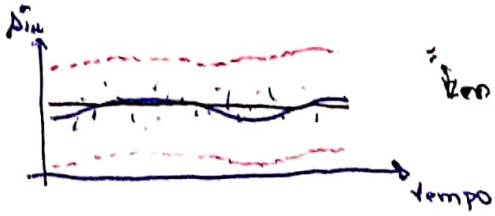
$E(\beta_{i1}^*) = \beta_1(t)$, $\beta_{i1}^* = r_{i1}^* + \hat{\beta}_1$, "rescaled residuals"

De forma mais geral,

$$E(r^*_{ik} + \hat{\beta}_k) = \hat{\beta}_k(t) \qquad , \; s^*_{ik} = r^*_{ik} + \hat{\beta}_k .$$

Procedimento Usual

Gráfico de s^*_{ik} versus tempos t:



"Em geral, ver se consigo traçar uma reta entre as bandas"

Se a tendência dos resíduos se aproxima de uma reta paralela ao eixo do tempo, então pode ser razoável.

$$\beta_k(t) = \beta_k \quad (\text{constante em } t)$$

e o modelo de riscos prop. é adequado.

Testes baseados nos Resíduos de Schoenfeld para verificar a sup. de prop. dos riscos

1) Teste global, para todas as variáveis.

Hipóteses do teste:

H_0: $\underline{\beta}$ não depende do tempo, ou seja, $\underline{\beta}(t) = \underline{\beta}$;

H_A: $\beta_q(t) = \beta_q + \theta_q g(t)$,

para pelo menos um dos q's, $q = 1, \dots, p$.

Observe que H_0 é equivalente a

$$H_0^T: \underset{\sim}{\theta} = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_p \end{pmatrix} = \underset{\sim}{0}.$$

Estatística do teste

$$U = \frac{(\underset{\sim}{g} - \bar{g})^T_{1\times m} \, R^*_{m\times p} \, I_{p\times p} \, R^{*T}_{p\times m} \, (\underset{\sim}{g} - \bar{g})_{m\times 1}}{m \sum_{k=1}^{m} (g_k - \bar{g})^2}, \quad \underset{\sim}{g} = \begin{pmatrix} g(t_1) \\ g(t_2) \\ \vdots \\ g(t_m) \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

, vetor de m linhas

m: nº de falhas

$$\bar{g} = \left(\sum_{k=1}^{m} g(t_k)\right) \cdot \frac{1}{m} = \frac{g_1 + g_2 + \dots + g_m}{m}, \text{ escalar}$$

R^* : matriz formada pelos resíduos $(m \times p)$ de Schoenfeld padronizados.

I : matriz de informação observada.

Sob a hipótese nula, pode-se demonstrar que

$$U \sim \chi^2_{(p)}$$

Rejeita-se H_0 para valores grandes de U.

ii) Teste para uma covariável apenas.

$$H_0: \beta_q(t) = \beta_q \text{ ou } \beta_q = 0$$

$$H_A: \beta_q(t) = \beta_q + \theta_q \, g(t)$$

A estatística de teste é dada por

$$U_q = \frac{m \left[\sum_{k=1}^{m} (g_k - \bar{g}) r_{kq}^{*} \right]^2}{I_q^{(-1)} \sum_{k=1}^{m} (g_k - \bar{g})^2},$$

$I_q^{(-1)}$: q-ésimo elemento da diagonal de I^{-1}.

Sob H_0,

$$U_q \sim \chi^2_{(1)}$$

Rejeita-se H_0 p/ valores grandes de U_q.

"Quando não se verifica a prop,"

formas são

- Modelo Aditivo de Aalen
- Modelos paramétricos
- outros

Não há um + adequado

Resíduos martingal: transf. das variáveis

Deviance: pontos atípicos

Escore: pontos influentes (dfbeta)

Schoenfeld: p/ proporcionalidade coxph (- - - , transform="tg")

forma da g(t)

Aula 26 — Análise de Sobrevivência — 07/06/2016

Modelo de Fragilidade

"Incorporar uma possível dependência entre indivíduos (mesma família, por exemplo)"

→ Contexto Univariado

Seja T uma v.a. não negativa. Suponha que condicional a v.a. U ($U \geq 0$), T tenha distribuição com função de risco dada por

$$\alpha(t|U) = \alpha_0(t)\, U\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}}$$

v.a. de fragilidade. Aumenta/diminui o risco.

$\underset{\sim}{X}$: vetor de p covariáveis

$\underset{\sim}{\beta}$: vetor de parâmetros desconhecidos

$\alpha_0(t)$: função de risco basal.

Observe que podemos definir

$$U = e^{W}$$

e

$$\alpha(t|W) = \alpha_0(t)\, e^{W}\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}}.$$

$W \sim \text{Normal}(\mu, \sigma^2)$, $e^W \sim \text{Lognormal}(\mu, \sigma^2)$

A função de sobrevivência condicional é dada por

$$S(t|U, \underset{\sim}{X}) = \exp\left\{-\int_0^t \alpha_0(s)\, U\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}}\, ds\right\}$$

$$= \exp\left\{-A_0(t)\, U\, e^{\underset{\sim}{X}'\underset{\sim}{\beta}}\right\}$$

"Costuma-se dizer que U é v.a. latente, uma vez que não a observamos. Solução: integrar em U"

A função de sobrevivência marginal de T fica dada por

$$S(t) = P(T > t) = \int_t^{+\infty} f_T(s)\, ds = \int_t^{+\infty} \int_0^{+\infty} \left(f_{T|U}(s|u) f(u) \right) du\, ds$$

$$= \int_0^{+\infty} \left(\int_t^{+\infty} f_{T|U}(s|u) f(u)\, ds \right) du$$

$$= \int_0^{+\infty} S_{T|U}(s|u) f(u)\, du$$

No caso de modelos de riscos proporcionais, temos

$$S(t) = P(T > t) = \int_0^{+\infty} e^{-(A_0(t) e^{X\beta} u)} f(u)\, du \qquad (*)$$

Definição: Seja V uma v.a. não negativa. A transformada de Laplace de V é definida por

$$\mathcal{L}_V(s) = E(e^{-sV}) = \int_0^{+\infty} e^{-sv} f(v)\, dv$$

Observe que, por $(*)$,

$$S(t) = \int_0^{+\infty} e^{-(A_0(t) e^{X\beta} u)} f(u)\, du$$

$$= \mathcal{L}_U(A_0(t) e^{X\beta}) = \mathcal{L}_U(A(t|\underset{\sim}{X})) \quad , \quad A(t|\underset{\sim}{X}) = A_0(t) e^{X\underset{\sim}{\beta}}$$

Exemplo:

Seja $U \sim \text{Gama}(\alpha, \beta)$, então

$$\mathcal{L}_U(s) = E(e^{-sU}) = \int_0^{+\infty} e^{-su} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta u} u^{\alpha-1}\, du$$

$$= \frac{\beta^\alpha}{(\beta+s)^\alpha} \int_0^{+\infty} \frac{(\beta+s)^\alpha}{\Gamma(\alpha)} e^{-(\beta+s)u} u^{\alpha-1}\, du = \frac{\beta^\alpha}{(\beta+s)^\alpha} = \frac{1}{(1+s/\beta)^\alpha}$$

Assuma agora que

$$\alpha(t/\underset{\sim}{X}, U) = \alpha_0(t)\, U e^{\underset{\sim}{X}'\underset{\sim}{\beta}}, \quad U \sim \text{Gama}(1/\theta, 1/\theta)$$

Nesse caso,

$$EU = 1/\theta \,/\, 1/\theta = 1 \quad \text{e} \quad Var(U) = \theta.$$

$$S(t/\underset{\sim}{X}) = \mathcal{L}_U(A(t/\underset{\sim}{X})) = \frac{1}{(1 + A(t/\underset{\sim}{X})\theta)^{1/\theta}} = \frac{1}{(1 + \theta A_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}})^{1/\theta}}.$$

A função densidade de probabilidade é dada por

$$f(t/\underset{\sim}{X}) = \frac{\alpha_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}}}{(1 + \theta A_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}})^{1/\theta + 1}}$$

$$\alpha(t/\underset{\sim}{X}) = \frac{f(t/\underset{\sim}{X})}{S(t/\underset{\sim}{X})} = \alpha_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}}$$

No caso geral, pode-se demonstrar ainda que

$$f(t) = -(\alpha_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}})\, \mathcal{L}'(A_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}})$$

$$\alpha(t) = -(\alpha_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}}) \frac{\mathcal{L}'(A_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}})}{\mathcal{L}(A_0(t) e^{\underset{\sim}{X}'\underset{\sim}{\beta}})}$$

Exemplos:

(1) Faça $U \sim$ Normalinversa.

A densidade é dada por

$$f(u) = \frac{\sqrt{\lambda}}{\sqrt{2\pi u^3}} \exp\left(-\frac{\lambda}{2\mu^2 u}(u-\mu)^2\right), \quad \lambda > 0,\ \mu > 0,\ u > 0$$

Pode-se demonstrar que

$$\mathcal{L}(s) = \exp\left\{ \frac{\lambda}{\mu}\left(1 - \sqrt{\frac{1+2\mu^2 s}{\lambda}}\right)\right\}$$

$E(U) = \mathcal{L}'(0) = \mu = 1.$

$Var(U) = \mathcal{L}''(0) - (\mathcal{L}'(0))^2 = \mu^3/\lambda.$

$$S(t/\underset{\sim}{X}) = \exp\left\{ \lambda\left(1 - \sqrt{\frac{1+2A_0(t)e^{\underset{\sim}{X}'\underset{\sim}{\beta}}}{\lambda}}\right)\right\}$$

(2) Faça $U = e^W$, $W \sim N(0,\sigma^2)$

Neste caso, não conseguimos expressão analítica para $S(t/\underset{\sim}{X})$.

$$\alpha(t/\underset{\sim}{X}, W) = \alpha_0(t)\,(\overset{U}{e^W})\,e^{\underset{\sim}{X}'\underset{\sim}{\beta}} \quad , \quad E(e^W) = e^{\mu+\sigma^2/2} = e^{\sigma^2/2} \neq \overset{\text{Na Soma}}{1}$$

Construção da Função de Verossimilhança para Modelos de Fragilidade Gama

Considere o contexto usual de análise de sobrevivência.

$$T_i = \min(T_i^*, C_i) \; ; \; \delta_i = I(T_i^* \leq C_i)$$

$\underset{\sim}{X}_i$: vetor de covariáveis fixas.

Assuma um modelo de fragilidade dado por

$$\alpha(t/X_i, U_i) = \alpha_0(t)\,e^{\underset{\sim}{X}_i'\underset{\sim}{\beta}}\,U_i, \quad i = 1, \ldots, n$$

com

$U_i \sim \text{Gama}(1/\theta, 1/\theta)$, independentes, não observáveis.

A função de verossimilhança fica dada por

$$L(\beta, \underline{\gamma}, \theta) = \prod_{i=1}^{n} (f(t/\underline{x}_i))^{\delta_i} (S(t/\underline{x}_i))^{1-\delta_i} = \prod_{i=1}^{n} (\alpha(t/\underline{x}_i))^{\delta_i} S(t/\underline{x}_i),$$

$\underline{\gamma}$ os parâmetros em $\alpha_0(t)$, $f(t/\underline{x})$, $\alpha(t/\underline{x})$, $S(t/\underline{x})$ distribuições marginais de T integrada com relação a U.

$$L(\beta, \underline{\gamma}, \theta) = \prod_{i=1}^{n} \left(\frac{\alpha_0(t_i) e^{\underline{x}_i'\beta}}{1 + \theta A_0(t) e^{\underline{x}_i'\beta}} \right)^{\delta_i} \frac{1}{(1 + \theta A_0(t_i) e^{\underline{x}_i'\beta})^{1/\theta}}$$

Antes, utilizava-se o algoritmo EM. No entanto, havia prob de conver., além da mesma ser lenta. Therneau propôs a verso penalizada e mostrou que ores da mesma equivale ao EM. Pela penaliz. conv. rápida, tudo ok!

Modelos de Fragilidade Compartilhada

Exemplo: Ninhadas de rato (1977)

- 15 ninhadas de ratos
- Em cada ninhada, 3 fêmeas

Variável resposta: tempo até o aparecimento do tumor.

No contexto geral, suponha um estudo com n indivíduos. Essas observações podem ser agrupadas em G grupos com n_i observações cada.

Observe que

$$n_1 + n_2 + \dots + n_G = n.$$

Seja T_{ij} o tempo de falha associado à j-ésima observação do grupo i.

No modelo de fragilidade, temos

$$\alpha(t \mid X_{ij}, U_i) = \alpha_0(t)\, U_i\, e^{X_{ij}'\beta} \quad , \; i = 1, \dots, G \; ; \; j = 1, \dots, n_i$$

Suposição: Dado U_i, T_{ij}^* são independentes.

No modelo de fragilidade gama,

$$U_i, \; i = 1, \dots, G \quad \text{i.i.d. Gamas}^{(\nu_0, \nu_0)} \text{ com } E(U_i) = 1 \text{ e } Var(U_i) = \theta$$

* Coeficiente τ de Kendall (Medida de Concordância)

Sejam

(T_{11}, T_{12}) Pares de v. aleatórios

(T_{21}, T_{22})

(T_{31}, T_{32}) $\qquad (T_{i1}, T_{i2})$ indep. de (T_{j1}, T_{j2}), $i \neq j$

$\vdots$

(T_{n1}, T_{n2})

n réplicas de T_1 T_2

O coeficiente τ de Kendall é definido por

$$\tau = E\left[\text{sinal}\left((T_{11} - T_{21})(T_{12} - T_{22}) \right) \right]$$

$$\text{sinal}(x) = \begin{cases} -1 & \text{se } x < 0 \\ 0 & \text{se } x = 0 \\ 1 & \text{se } x > 0 \end{cases}$$

Análise de Sobrevivência

Aula 27 14/06/2016

Recordar aula passada (Foto)

Coeficiente τ de Kendall (medida de concordância)

Sejam

$$\left.\begin{array}{c}(T_{11}, T_{12}) \\ (T_{21}, T_{22}) \\ \vdots \\ (T_{n1}, T_{n2})\end{array}\right\} \text{ pares de v.a. com } (T_{i1}, T_{i2}) \text{ ind. de } (T_{j1}, T_{j2}),\ i \neq j.$$

O coeficiente τ de Kendall é definido por

$$\tau = E\Big(\text{sinal}\,(T_{11} - T_{21})(T_{12} - T_{22})\Big)$$

$$\text{sinal}(x) = \begin{cases} -1, & \text{se } x < 0 \\ 0, & \text{se } x = 0 \\ 1, & \text{se } x > 0 \end{cases}$$

Para variáveis aleatórias contínuas

$$\begin{aligned}\tau &= E\big(\text{sinal}\,(T_{11} - T_{21})(T_{12} - T_{22})\big) \\ &= P\big((T_{11} - T_{21})(T_{12} - T_{22}) > 0\big) - P\big((T_{11} - T_{21})(T_{12} - T_{22}) < 0\big) \\ &= P\big((T_{11} - T_{21})(T_{12} - T_{22}) > 0\big) - \big[1 - P\big((T_{11} - T_{21})(T_{12} - T_{22}) > 0\big)\big] \\ &= 2p - 1,\end{aligned}$$

em que $p = P\big((T_{11} - T_{21})(T_{12} - T_{22}) > 0\big)$

Observação. Assuma T_1 e T_2 independentes, ou seja, T_{i1} é independente de T_{i2}.

Temos

$$P((T_{21}-T_{21})(T_{12}-T_{22})>0) = P((T_{11}-T_{21})>0, (T_{12}-T_{22})>0)$$
$$+ P((T_{11}-T_{21})<0, (T_{12}-T_{22})<0)$$

Pelo indep.

$$P((T_{11}-T_{21})>0, (T_{12}-T_{22})>0) = P(T_{11}-T_{21}>0)P(T_{12}-T_{22}>0) = 1/4,$$

$\Rightarrow \tau = 0.$

Logo como indep. $\Rightarrow \tau = 0$, $\tau \neq 0 \Rightarrow$ não indep.

Resultado

Para os modelos de fragilidade compartilhada com

$$\alpha_{ij}(t \mid X_{ij}, U_i) = \alpha_0(t) U_i e^{X_{ij}\beta}.$$

Vale o resultado

$$\tau = 4\left(\int_0^{+\infty} s\, \mathcal{L}''(s)\, \mathcal{L}(s)\, ds\right) - 1$$

em que $\mathcal{L}(\cdot)$ é a transformada de Laplace.

Exemplo: Considere o modelo de fragilidade gama com $U_i \sim \text{Gama}(1/\theta; 1/\theta)$.

$$\mathcal{L}_U(s) = \frac{1}{(1+\theta s)^{1/\theta}} \quad \text{(resultados anteriores)}$$

$$\mathcal{L}_U'(s) = -\frac{1}{(1+\theta s)^{1/\theta+1}}, \quad \mathcal{L}_U''(s) = \frac{1+\theta}{2(2+\theta)}$$

$$\int_0^{\infty} s\, \mathcal{L}''(s)\, \mathcal{L}'(s)\, ds \overset{\text{por partes}}{=} \frac{1+\theta}{2(2+\theta)}; \quad \tau = 4 \cdot \frac{\cancel{2}(1+\theta)}{\cancel{2}(2+\theta)} - 1 = \frac{\theta}{\theta+2}$$

$\tau = \dfrac{\theta}{\theta+2}$, θ variância do ef. aleatório. Induz uma dep. positiva.

Análise de Sobrevivência 14/06/16

Modelo de Fragilidade Gama

T_{ij}^* : tempo de falha da j-ésima observação do grupo i

$T_{ij} = \min(T_{ij}^*, C_{ij})$; $\delta_{ij} = I(T_{ij}^* \leq C_{ij})$ $\quad \underset{\sim}{X}_{ij}$: vetor de covariáveis

$j = 1, \ldots, n_i$; $i = 1, \ldots, G$.

Modelo: $\alpha_{ij}(t / X_{ij}, U_i) = \alpha_0(t)\, e^{X_{ij}\beta}$, $U_i \sim \text{Gama}(1/\theta, 1/\theta)$

Suposição: Dado U_i, T_{ij}^* são independentes.

Construção da Verossimilhança

$$L(\underset{\sim}{\beta}) = \prod_{i=1}^{G} L_i(t_{i1}, t_{i2}, \ldots, t_{in_i}) = \prod_{i=1}^{G} L_i$$

$$f_i(t_{i1}, t_{i2}, \ldots, t_{in_i}) = \int_0^{+\infty} f(t_{i1}, \ldots, t_{in_i} / U_i)\, f_U(u_i)\, du_i$$

$$L_{FULL}(\underset{\sim}{\beta}, \theta, A_0) = \prod_{i=1}^{G} \left[\prod_{j=1}^{n_i} (\alpha_{ij}(t_{ij} / X_{ij}, U_i))^{\delta_{ij}}\, S_{ij}(t_{ij} / U_i, X_{ij}) \right] \cdot g(U_i)$$

No modelo de fragilidade gama

$$L_{FULL}(\beta, \theta, A_0) = \prod_{i=1}^{G} \left\{ \left[\prod_{j=1}^{n_i} (\alpha_0(t_{ij})\, U_i\, e^{X_{ij}'\beta})^{\delta_{ij}}\, e^{-A_0(t_{ij}) U_i e^{X_{ij}'\beta}} \right] \frac{(1/\theta)^{1/\theta}}{\Gamma(1/\theta)}\, U_i^{1/\theta - 1}\, e^{-U_i/\theta} \right\}$$

Func de Veross. Parcial Penalizada

Relembre que, no modelo de Cox, o log da veross. parcial é

$$\ell(\underset{\sim}{\beta}) = \sum_{i=1}^{D} \left(\underset{\sim}{x}_{(i)}' \underset{\sim}{\beta} - \log \sum_{j \in R_i} e^{\underset{\sim}{x}_j' \underset{\sim}{\beta}} \right)$$

$$= \sum_{i=1}^{n} \delta_i \left(\underset{\sim}{x}_i' \underset{\sim}{\beta} - \log \left(\sum_{j \in R} e^{\underset{\sim}{x}_j' \underset{\sim}{\beta}} \right) \right)$$

Therneau e Grambsch propuseram a f.v. parcial penalizada:

$$PPL(\underset{\sim}{\beta}, \underset{\sim}{w}, \theta) = \ell(\underset{\sim}{\beta}, \underset{\sim}{w}) - g(\underset{\sim}{w}, \theta),$$

com $\underset{\sim}{w} = (w_1, \ldots, w_G)$ e $w_i = \log U_i$, ou $U_i = e^{w_i}$.

$$\ell(\underset{\sim}{\beta}, \underset{\sim}{w}) = \sum_{i=1}^{G} \sum_{j=1}^{n_i} \delta_{ij} \left(\underset{\sim}{x}_{ij}' \underset{\sim}{\beta} + w_i - \log \left(\sum_{l \in R(t_{ij})} e^{\underset{\sim}{x}_{il}' \underset{\sim}{\beta} + w_l} \right) \right),$$

$g(\underset{\sim}{w}, \theta)$: função de penalidade

Modelo de Fragilidade Gama

$U_1, \ldots, U_G$ i.i.d. gama$(1/\theta; 1/\theta)$, $W_i = \log U_i$.

$$f(w) = \frac{(1/\theta)^{1/\theta}}{\Gamma(1/\theta)} e^{w/\theta} \exp\left(-\frac{e^{w}}{\theta} \right)$$

$$f(w_1, \ldots, w_G) = \frac{(1/\theta)^{G/\theta}}{(\Gamma(1/\theta))^{G}} e^{\sum w_i/\theta} \exp\left(-\frac{1}{\theta} \sum_{i=1}^{G} e^{w_i} \right)$$

$$\log f(w_1, \dots, w_G) = -\frac{G}{\theta}\log\theta - G\log\Gamma(1/\theta) + \frac{1}{\theta}\sum_{i=1}^{G}(w_i - e^{w_i})$$

A proposta de Therneau é usar

$$g(\underline{w}, \theta) = \frac{1}{\theta}\sum_{i=1}^{G}(w_i - e^{w_i})$$

Processo de Estimação

É dividido em 3 etapas

(i) Maximização em relação a $\underline{\beta}$ e $\underline{w}$

(ii) Estimação de θ

(i) Maximização em relação à $\underline{\beta}$ e $\underline{w}$.

$$U(\underline{\beta}_q) = \frac{\partial PPL}{\partial \beta_q} = \sum_i \sum_j \delta_{ij}\left(X_{ij(q)} - \frac{\sum_{k,l \in R(t_{ij})} X_{kl(q)}\, e^{\underline{X}'_{kl}\underline{\beta} + \overset{\text{offset}}{w_k}}}{\sum_{k,l \in R(t_{ij})} e^{\underline{X}'_{kl}\underline{\beta} + w_k}} \right)$$

Para w_i, temos

$$U(w_i) = \frac{\partial PPL}{\partial w_i} = d_i - \sum_{k=1}^{G}\sum_{l=1}^{n_k} \delta_{kl}\left(\frac{\sum_{m \in R(t_{k,l})} e^{\underline{X}'_{im}\underline{\beta} + w_i}}{\sum_{g,m \in R(t_{k,l})} e^{\underline{X}'_{gm}\underline{\beta} + w_g}} \right) - \frac{1}{\theta}(1 - e^{w_i})$$

Assim, para um dado valor de θ, obtemos

$\hat{\underline{\beta}}$ e $\hat{w}_i$ fazendo-se $U(\underline{\beta}) = 0$ e $U(\underline{w}) = 0$.

(ii) Estimação de θ

Inicialmente obtemos $\hat{A}_0(t)$ com base no estimador de Breslow (com offset w_i)

Frag. Gama: obtemos a f. veros. marginal (integrando-se com relação a U_i)

Substituimos

$$\beta \text{ por } \hat{\beta} \quad e \quad A_0(t) \text{ por } \hat{A}_0(t) \quad e \quad \theta \text{ por } \hat{\theta}.$$

e maximizamos com relação à θ.

~~[illegible]~~

Algoritmo:

1. Especificamos val. iniciais para θ e $\underset{\sim}{w}$.
2. Obtemos $\hat{\beta}$ consid. os valores de $\underset{\sim}{w}$ e θ atualizados
3. Atualizamos os val. de $\underset{\sim}{w}$ utilizando as equações e os com obtidos, com o θ atualizado e $\hat{\beta}$ obtido em 2.
4. obtemos (2) e (3) até convergência

   (2) e (3): loop interno
5. Atualizamos θ de acordo com (11)
6. Rep. (2) a (5) até a converg.

Modelo de Frag. Log. Normal.

$$f(\underset{\sim}{w}) = \frac{1}{(2\pi\sigma^2)^{G/2}} e^{-\sum_{i=1}^{G} w_i^2 / 2\sigma^2}$$

$$\log f(\underset{\sim}{w}) = -\frac{G}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{G} w_i^2$$

Ripatti e Palmgren (2000).

$$PPL(\underset{\sim}{\beta}, \underset{\sim}{w}, \sigma^2) = l(\underset{\sim}{\beta}, \underset{\sim}{w}) - g(\underset{\sim}{w}, \sigma^2)$$

$$g(w, \sigma^2) = -\sum_{i=1}^{G} w_i^2 / 2\sigma^2$$

Processo de maximização: igual ao caso anterior, modificando a est. de σ^2. Chegam na expressão

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{G} \hat{w}_i^2 + t(H_{22}^{-1})}{G}, \quad t(H_{22}^{-1}) \cdot \text{traço de } H_{22}^{-1}, \quad H_{22}^{-1} : \text{inversa de } H_{22}$$

$$H_{22} = -\partial^2 PPL(\underset{\sim}{\beta}, \underset{\sim}{w}, \sigma^2)$$

ANÁLISE DE SOBREVIVÊNCIA - LISTA 1

Maicon Aparecido Pinheiro

1. Seja T uma variável aleatória contínua não negativa com função densidade de probabilidade $f_T(t)$, função de Sobrevivência $S_T(t)$ e função de taxa de falha (ou função de risco) $\alpha_T(t)$.

a) Mostre que

$$\alpha_T(t) = \frac{f_T(t)}{S_T(t)}.$$

Resposta: Uma vez que T é contínua, $\mathbb{P}(T \geq t) = \mathbb{P}(T > t) = S_T(t)$. Dito isso, basta ver que

$$\begin{aligned}
\alpha_T(t) &= \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(t \leq T \leq t + \Delta t \,|\, T \geq t\right)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(t \leq T \leq t + \Delta t, T \geq t\right)}{\mathbb{P}(T \geq t)\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(t \leq T \leq t + \Delta t\right)}{\mathbb{P}(T > t)\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\mathbb{P}\left(T \leq t + \Delta t\right) - \mathbb{P}\left(T \leq t\right)}{S_T(t)\Delta t} \\
&= \frac{1}{S_T(t)} \lim_{\Delta t \to 0} \frac{F_T(t + \Delta t) - F_T(t)}{\Delta t} \\
&= \frac{1}{S_T(t)} \frac{\mathrm{d}}{\mathrm{d}t} F_T(t) \\
&= \frac{f_T(t)}{S_T(t)}
\end{aligned}$$

Q1 1.0
Q2 1.0
Q3 0.5
Q4 1.5
Q5 2.0
Q6 2,4
Q7 1.5
9,9

emque ~~onde~~ $F_T(t)$ é a função distribuição acumulada de T.

b) Mostre que

$$S_T(t) = \exp - \int_0^t \alpha_T(s) ds.$$

Resposta: Do item anterior, sabemos que

$$\alpha_T(t) = \frac{1}{S_T(t)} \frac{\mathrm{d}}{\mathrm{d}t} F_T(t),$$

emque ~~onde~~ $F_T(t)$ é a função distribuição acumulada de T. Como $F_T(t) = 1 - S_T(t)$, segue que

$$\begin{aligned}
\alpha_T(t) &= \frac{1}{S_T(t)} \frac{\mathrm{d}}{\mathrm{d}t} \left(1 - S_T(t)\right) = \frac{1}{S_T(t)} \frac{\mathrm{d}}{\mathrm{d}t} \left(-S_T(t)\right) \\
&= -\frac{\mathrm{d}}{\mathrm{d}t} S_T(t) \Big/ S_T(t) \\
&= -\frac{\mathrm{d}}{\mathrm{d}t} \ln S_T(t).
\end{aligned}$$

Integrando ambos os lados da equação acima, temos

$$\begin{aligned}
\int_0^t \alpha_T(s)\mathrm{d}s = -\int_0^t \frac{\mathrm{d}}{\mathrm{d}t}\ln S_T(t)\mathrm{d}s \quad &\Leftrightarrow \\
\int_0^t \alpha_T(s)\mathrm{d}s = -\left(\ln S_T(t) - \ln S_T(0)\right) \quad &\Leftrightarrow \\
\int_0^t \alpha_T(s)\mathrm{d}s = -\ln S_T(t) \quad &\Leftrightarrow \\
\ln S_T(t) = -\int_0^t \alpha_T(s)\mathrm{d}s \quad &\Leftrightarrow \\
S_T(t) = \exp\left\{-\int_0^t \alpha_T(s)\mathrm{d}s\right\}.
\end{aligned}$$

2. Mostre que, para T uma variável aleatória contínua e não negativa, vale o resultado

$$\mathrm{E}(T) = \int_0^\infty S(t)\mathrm{d}t.$$

Mostre também que, para T variável aleatória discreta assumindo valores inteiros não negati vale o resultado

$$\mathrm{E}(T) = \sum_{j=0}^{\infty} P(T > j) = \sum_{j=1}^{\infty} P(T \geq j).$$

Resposta: Em primeiro momento, vamos averiguar que

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

Para tal, note que para cada $b > 0$,

$$\int_0^b tf(t)\mathrm{d}t = tF(t)\Big|_0^b - \int_0^b F(t)\mathrm{d}t = bF(b) - \int_0^b F(t)\mathrm{d}t = \int_0^b [F(b) - F(t)]\,\mathrm{d}t,$$

onde a primeira passagem decorre de uma integração por partes com $u = t$ e $\mathrm{d}v = f(t)\mathrm{d}t$, o implica em $\mathrm{d}u = \mathrm{d}t$ e $v = F(t)$. Como $F(b) \leq 1$ e $1 - F(t) \geq 0$, temos

$$\int_0^b tf(t)\mathrm{d}t = \int_0^b [F(b) - F(t)]\,\mathrm{d}t \leq \int_0^b [1 - F(t)]\,\mathrm{d}t,\ \forall b > 0,$$

de modo que

$$\int_0^\infty tf(t)\mathrm{d}t = \lim_{b\to\infty}\int_0^b tf(t)\mathrm{d}t \leq \int_0^\infty [1 - F(t)]\,\mathrm{d}t = \int_0^\infty S(t)\mathrm{d}t.$$

Por outro lado, seja $\lambda > 0$. Se $b > \lambda$, então

$$\begin{aligned}
\int_0^b [F(b) - F(t)]\,\mathrm{d}t &\geq \int_0^\lambda [F(b) - F(t)]\,\mathrm{d}t \\
&= \int_0^\lambda [F(b) - 1]\,\mathrm{d}t + \int_0^\lambda [1 - F(t)]\,\mathrm{d}t \\
&= \lambda[F(b) - 1] + \int_0^\lambda S(t)\mathrm{d}t,
\end{aligned}$$

e, portanto,

$$\int_0^\infty tf(t)\mathrm{d}t = \lim_{b\to\infty}\int_0^b tf(t)\mathrm{d}t = \lim_{b\to\infty}\int_0^b [F(b)-F(t)]\,\mathrm{d}t$$
$$\geq \int_0^\lambda S(t)\mathrm{d}t + \lim_{b\to\infty}\lambda\,[F(b)-1] = \int_0^\lambda S(t)\mathrm{d}t.$$

Já que vale para todo $\lambda > 0$, temos

$$\int_0^\infty tf(t)\mathrm{d}t \geq \lim_{\lambda\to\infty}\int_0^\lambda S(t)\mathrm{d}t = \int_0^\infty S(t)\mathrm{d}t.$$

Das duas desigualdades obtidas, segue que

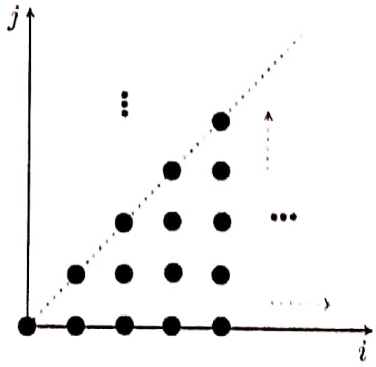
$$\mathrm{E}(T) = \int_0^\infty S(t)\mathrm{d}t.$$

Agora, se T é uma variável aleatória não negativa discreta, basta ver que

$$\sum_{i=0}^\infty P(T>i) = \sum_{i=1}^\infty P(T\geq i) = \sum_{i=1}^\infty\sum_{j=i}^\infty P(T=j).$$

Uma vez que a probabilidade é uma função não negativa, pelo Teorema de Tonneli podemos alterar a ordem das somas respeitando os limites de variação, de modo que

$$\sum_{i=0}^\infty P(T>i) = \sum_{i=1}^\infty\sum_{j=i}^\infty P(T=j) = \sum_{j=1}^\infty\sum_{i=1}^j P(T=j) = \sum_{j=1}^\infty jP(T=j) = E\,[T]\,.$$



3. A função do tempo médio de vida residual para uma variável aleatória contínua T é definida como $m(t) = E(T-t\,|\,T>t)$. Denotando por $S(t)$ a função de sobrevivência de T, mostre que

$$m(t) = \frac{\int_0^t S(x)dx}{S(t)}$$

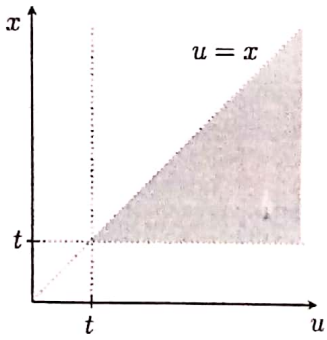
Resposta: Em primeiro momento, note que a condicional de T dado a ocorrência do evento $\{T>t\}$ é dada por

$$f_{T|T>t}(u) = \frac{f_T(u)}{S_T(t)}, \quad u>t.$$

Daí, da definição de esperança condicional, segue que

$$\begin{aligned}
E(T-t\,|\,T>t) &= \int_t^{\infty}(u-t)\frac{f_T(u)}{S_T(t)}\mathrm{d}u \\
&= \frac{1}{S_T(t)}\int_t^{\infty}(u-t)f_T(u)\mathrm{d}u \\
&= \frac{1}{S_T(t)}\int_t^{\infty}\left(\int_t^{u}\mathrm{d}x\right)f_T(u)\mathrm{d}u \\
&= \frac{1}{S_T(t)}\int_t^{\infty}\left(\int_x^{\infty}f_T(u)\mathrm{d}u\right)\mathrm{d}x \\
&= \frac{\int_t^{\infty}S_T(x)\mathrm{d}x}{S_T(t)},
\end{aligned}$$

onde na segunda passagem usamos o fato de que $u-t=\int_t^u \mathrm{d}x$, e na terceira o Teorema de Fubi para trocar a ordem de integração, observando os novos limites da integral.



4. Suponha que a variável aleatória Y tenha distribuição normal com média μ e variância σ assuma que $T=e^Y$.

a) Obtenha uma expressão para a função de sobrevivência de T, ou seja, $P(T>t)$.

Resposta: Em primeiro momento, note que T toma valores maiores ou iguais a zero, u vez que $T=e^Y$. Observado isso, basta ver que

$$S_T(t)=P(e^Y>t)=P(Y>\ln t)=P\left(\frac{Y-\mu}{\sigma}>\frac{\ln t-\mu}{\sigma}\right)=1-\Phi\left(\frac{\ln t-\mu}{\sigma}\right),\quad t>$$

onde $\Phi(\cdot)$ é a função distribuição acumulada da distribuição normal padrão.

b) Obtenha uma expressão para a função de taxa de falha associada a T, ou seja, $\alpha(t)$.

Resposta: Não há uma forma explícita para a taxa de falha associada a T, mas poder representá-la a partir da relação vista no item *a)* do Exercício 1 dessa lista, de modo que

$$\alpha_T(t)=\frac{f_T(t)}{S(t)}=\frac{1}{t\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left(\frac{\ln t-\mu}{\sigma}\right)^2\right\}\Bigg/\left[1-\Phi\left(\frac{\ln t-\mu}{\sigma}\right)\right],\quad t>0,$$

uma vez que

$$f_T(t)=\frac{\mathrm{d}}{\mathrm{d}t}F_T(t)=\frac{\mathrm{d}}{\mathrm{d}t}\Phi\left(\frac{\ln t-\mu}{\sigma}\right)=\phi\left(\frac{\ln t-\mu}{\sigma}\right)\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\ln t-\mu}{\sigma}\right)=\phi\left(\frac{\ln t-\mu}{\sigma}\right)\frac{1}{t\sigma},\ t>$$

onde $\phi(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Phi(x)$ é a densidade da distribuição normal padrão no ponto x, isto é,

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\}, \quad x \in \mathbb{R}.$$

c) Calcule $E(T^r)$, a média e a variância de T.

Resposta: Como $T = e^Y$, segue que

$$E(T^r) = E\left((e^Y)^r\right) = E(e^{rY}) = M_Y(r) = \exp\left\{\mu r + \frac{1}{2}\sigma^2 r^2\right\}, \; r \in \mathbb{R},$$

onde $M_Y(\cdot)$ é a função geradora de momentos da variável aleatória Y, que é bem conhecida. Logo,

$$E(T) = \exp\left\{\mu + \frac{\sigma^2}{2}\right\} \quad \text{e} \quad E(T^2) = \exp\left\{2(\mu + \sigma^2)\right\},$$

e

$$Var(T) = E(T^2) - (E(T))^2 = e^{2(\mu+\sigma^2)} - e^{2\mu+\sigma^2} = e^{2\mu+\sigma^2}\left(e^{\sigma^2} - 1\right).$$

5. Considere o artigo

Birnbaum, Z. W. e Saunders, S. C. (1969). A new family of life distributions. *Journal of Applied Probability* 6: 319-327

e o material adicional disponibilizado.

a) Descreva como os autores derivaram a distribuição de Birnbaum-Saunders (seção 2 do artigo original).

Resposta: Antes da formulação matemática necessária e de descrever como os autores derivaram a distribuição de Birnbaum-Saunders, vamos apresentar, baseado no artigo, o contexto.

Contexto: Suponha que um determinado material é submetido a um conjunto finito de cargas e de modo cíclico, isto é, as cargas são pré-fixados e se repetem de ciclo para ciclo. Tais cargas desgastam o material causando fissuras. Com o decorrer desse mecanismo, a fissura principal ultrapassa um certo nível de resistência crítico ω. Nesse instante, ocorre a falha do material.[1] O tempo até a falha do material (em sua versão contínua) é uma variável aleatória com distribuição de Birbaum Saunders.

Vamos à formulação matemática e à derivação da distribuição. Para tal, denotemos por $l_1, l_2 \ldots$ as cargas às quais o material é submetido no decorrer do processo. Como já mencionado, supõe-se que a sequência de cargas ocorre de maneira cíclica, isto é, para algum $m > 1$ (número de cargas aplicadas em cada ciclo) e para todo $i = 1, \ldots, m$,

$$l_{jm+i} = l_{km+i}, \quad \forall j \neq k.$$

No $j+1$-ésimo ciclo ocorre a exposição do material às cargas $(l_{jm+1}, l_{jm+2}, \ldots, l_{jm+m})$. Agora, vamos supor que

[1] Tal processo é conhecido na literatura como processo de fadiga.

1. O incremento X_i na fissura decorrente da $i-$ésima oscilação da carga é uma vari aleatória com distribuição que depende só da fissura atual causada pela tensão no pectivo ciclo. Nessas condições, a extensão da fissura durante o $(j+1)$-ésimo c é

$$Y_{j+1} = X_{jm+1} + \ldots + X_{jm+m}, \quad \text{para } j = 0, 1, \ldots,$$

onde X_{jm+i} é o incremento da fissura decorrente da carga l_i aplicada na i-ésima oscila no $j+1$-ésimo ciclo. Segue dessa suposição que as extensões da fissura por ciclos independentes.

2. O incremento total Y_j associado ao $j-$ésimo ciclo é uma variável aleatória com médi e variância σ^2 para $j = 1, 2, \ldots$. Nossa notação será

$$W_n = \sum_{j=1}^{n} Y_j,$$

que tem função distribuição

$$H_n(w) = P[W_n \leq w], \quad n = 1, 2, \ldots.$$

Segue de considerações elementares de probabilidade que a distribuição de C, o núm de ciclos necessários até a falha, no caso em que a falha é definida como o ultrapasse comprimento crítico ω pela primeira vez, é dada por

$$P[C \leq n] = P[W_n > w] = 1 - H_n(\omega).$$

Agora,

$$P[C \leq n] = P\left[\sum_{i=1}^{n} \frac{Y_j - n\mu}{\sigma\sqrt{n}} > \frac{\omega - n\mu}{\sigma\sqrt{n}}\right]$$

e uma vez que as Y_j's são variáveis aleatórias i.i.d., usando o Teorema Limite Cen para n grande, a equação acima pode ser aproximada pela distribuição normal.

3. A distribuição de C é

$$P[C \leq n] = 1 - \Phi\left[\frac{\omega - n\mu}{\sigma\sqrt{n}}\right] = \Phi\left[\frac{\mu\sqrt{n}}{\sigma} - \frac{\omega}{\sigma\sqrt{n}}\right]$$

A partir daí, Birnbaum e Saunders substituem n não equação acima por uma variável ı não negativa t. Dessa forma, se obtém T, uma extensão contínua da variável discreta C, · representa tempo total até a falha do material.

A função de distribuição da variável aleatória T fica dada por

$$F(t; \beta, \alpha) = P(T \leq t) = \Phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)\right], t > 0,$$

com $\alpha = (\sigma/\sqrt{\omega\mu}) > 0$ e $\beta = (\omega/\mu) > 0$.

b) Com base em (a), obtenha a função densidade de probabilidade de uma variável com dis buição Birnbaum-Saunders.

Resposta: A partir da $F(t;\beta,\alpha)$ obtida no item a) podemos obter a função densidade de probabilidade de uma variável com distribuição Birnbaum-Saunders por meio da relação

$$
\begin{aligned}
f(t;\beta,\alpha) = \frac{\mathrm{d}}{\mathrm{d}t}F(t;\beta,\alpha) &= \phi\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}}-\sqrt{\frac{\beta}{t}}\right)\right] \times \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{\alpha}\left(\sqrt{\frac{t}{\beta}}-\sqrt{\frac{\beta}{t}}\right)\right] \\
&= \frac{1}{2\alpha\sqrt{2\pi}}\left[\frac{1}{\sqrt{\beta t}}+\frac{\sqrt{\beta}}{t^{3/2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\left(\sqrt{\frac{t}{\beta}}-\sqrt{\frac{\beta}{t}}\right)^2\right\} \\
&= \frac{(t+\beta)}{2t^{3/2}\alpha\sqrt{2\pi\beta}}\exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t}{\beta}+\frac{\beta}{t}-2\right)\right\}, \quad t>0.
\end{aligned}
$$

c) Obtenha gráficos da função densidade de probabilidade, função de taxa de falha e função de sobrevivência para algumas combinações de valores dos parâmetros da distribuição (considere quatro ou cinco combinações diferentes e coloque as curvas de todas as combinações no mesmo gráfico). Você pode utilizar o software de sua preferência, porém adicione o código utilizado para a geração dos gráficos.

Resposta: Usarei o software R.

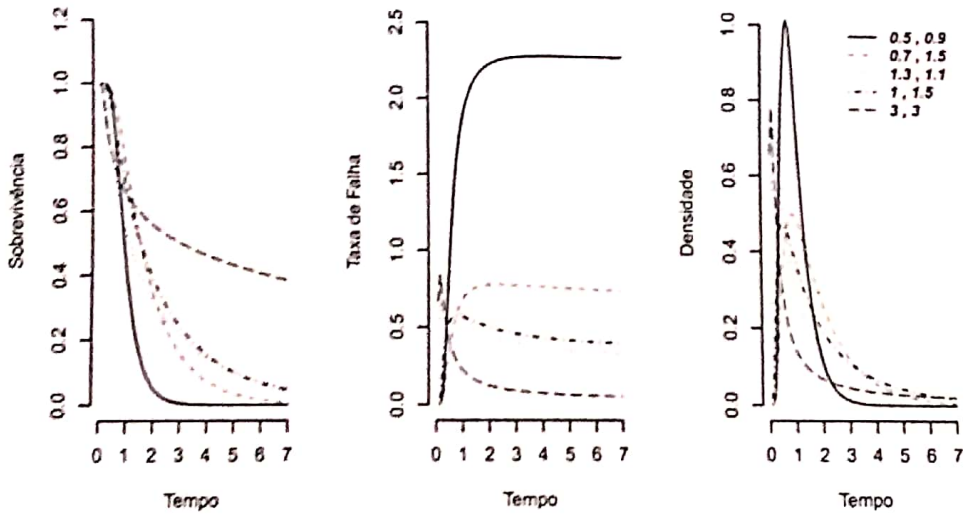


Figura 1: *Funções de sobrevivência, de taxa de falha e de densidade da Birbaun-Saunders para cinco combinações de* (α,β).

O código é dado abaixo.

```
dens.bs<-function(x,a,b){((x+b)/(2*(x^(1.5))*a*sqrt(2*pi*b)))*
                    exp(-(1/(2*a^2))*(x/b+b/x-2))}
surv.bs<-function(x,a,b){1-pnorm((1/a)*(sqrt(x/b)-sqrt(b/x)))}
rate.bs<-function(x,a,b){(dens.bs(x,a,b))/(surv.bs(x,a,b))}


fun.bs<-list(surv.bs,rate.bs,dens.bs)

alpha<-c(0.5,0.7,1.3,1,3)
```

```
beta<-c(0.9,1.5,1.1,1.5,3)
yc<-c(1.2,2.5,1)
yname<-c('Sobrevivência','Taxa de Falha','Densidade')
color<-c('black','tomato','forestgreen','navyblue','mediumorchid4')

par(mfrow=c(1,3))

for (j in 1:3) {
        f<-function(x){fun.bs[[j]](x,alpha[1],beta[1])}
        curve(f(x),0,7,ylim=c(0,yc[j]),ylab = yname[j],xlab='Tempo',
              bty='n',col=color[1],lty=1)
        for (i in 2:5) {
                f<-function(x){fun.bs[[j]](x,alpha[i],beta[i])}
                curve(f(x),0,7,ylim=c(0,yc[j]),,ylab = yname[j],
                      xlab='Tempo',col=color[i],bty='n',lty=i,add=TRUE)
        }
}

legenda<-c()
for (i in 1:5)
{
        a<-as.character(alpha[i])
        b<-as.character(beta[i])
        legenda[i]<-paste(a,',',b)
}

legend('topright',inset=0.02,legenda,cex=.85,bty='n',lty=c(1,2,3,4,5),
       text.font=4,col=color)
```

6. Suponha que o tempo de vida T tenha distribuição exponencial com parâmetro λ e que o ten de censura tenha também distribuição exponencial com parâmetro θ. Sejam $Z = \min(T; C$ $\delta = I(T \leq C)$. Assuma ainda que T e C sejam independentes.

a) Obtenha $P(\delta = 1)$.

Resposta: Basta notar que os eventos $\{\delta = 1\}$ e $\{T \leq C\}$ são equivalentes, de modo qu

$$
\begin{aligned}
P(\delta = 1) = P(T \leq C) &= \int_0^\infty \int_0^c f_{T,C}(t,c)\mathrm{d}t\mathrm{d}c \\
&= \int_0^\infty \int_0^c \lambda e^{-\lambda t}\theta e^{-\theta c}\mathrm{d}t\mathrm{d}c \\
&= \int_0^\infty \theta e^{-\theta c}\left[1 - e^{-\lambda t}\right]\Big|_0^c \mathrm{d}c \\
&= \int_0^\infty \theta e^{-\theta c}\mathrm{d}c - \int_0^\infty \theta e^{-(\theta+\lambda)c}\mathrm{d}c \\
&= 1 - \frac{\theta}{\theta+\lambda}\int_0^\infty (\theta+\lambda)e^{-(\theta+\lambda)c}\mathrm{d}c \\
&= 1 - \frac{\theta}{\theta+\lambda} \\
&= \frac{\lambda}{\theta+\lambda}.
\end{aligned}
$$

Daí, decorre que δ tem distribuição Bernoulli com parâmetro $\lambda/(\lambda+\theta)$.

b) Encontre a distribuição de Z.

Resposta: Uma vez que T e C são não negativas, Z também o é. Dito isso, note que

$$\begin{aligned} P(Z \geq z) &= P(T \geq z, C \geq z) \\ &= P(T \geq z)P(C \geq z) \\ &= e^{-\lambda z}e^{-\theta z} \\ &= e^{-(\lambda+\theta)z}, \quad z \geq 0. \end{aligned}$$

Daí,

$$F_Z(s) = P(Z \leq z) = 1 - e^{-(\lambda+\theta)z}, \quad z \geq 0,$$

isto é, Z tem distribuição exponencial com parâmetro $\lambda+\theta$.

c) Mostre que Z e δ são independentes.

Resposta: Pelo itens *a)* e *b)*, sabemos que Z tem distribuição exponencial com parâmetro $\lambda+\theta$ e que δ tem distribuição de Bernoulli com probabilidade de sucesso $\lambda/(\lambda+\theta)$, de modo que quando $\Delta \to 0$,

$$\begin{aligned} P\left(Z \in (z, z+\Delta)\right) \times P(\delta = i) &= (\lambda+\theta)e^{-(\lambda+\theta)z} \times \left(\frac{\lambda}{\lambda+\theta}\right)^i \left(\frac{\theta}{\lambda+\theta}\right)^{1-i} \\ &= \lambda^i\theta^{1-i}e^{-(\lambda+\theta)z}, \quad z > 0, i \in \{0, 1\}. \end{aligned}$$

Por outro lado, quando $\Delta \to 0$, para $\delta = 1$,

$$\begin{aligned} P\left(Z \in (z, z+\Delta), \delta = 1\right) &= P\left(Z \in (z, z+\Delta), T \leq C\right) \\ &= P\left(T \in (z, z+\Delta), C > z+\Delta\right) \\ &= P\left(T \in (z, z+\Delta)\right) P\left(C > z+\Delta\right) \\ &= \lambda e^{-\lambda z}e^{-\theta z} \\ &= \lambda^1\theta^{1-1}e^{-(\lambda+\theta)z}, \end{aligned}$$

e para $\delta = 0$,

$$\begin{aligned} P\left(Z \in (z, z+\Delta), \delta = 1\right) &= P\left(Z \in (z, z+\Delta), C \leq T\right) \\ &= P\left(C \in (z, z+\Delta), T > z+\Delta\right) \\ &= P\left(C \in (z, z+\Delta)\right) P\left(T > z+\Delta\right) \\ &= \theta e^{-\theta z}e^{-\lambda z} \\ &= \lambda^0\theta^{1-0}e^{-(\lambda+\theta)z}, \end{aligned}$$

Como a probabilidade conjunta pode ser escrita como o produto das probabilidades marginais, segue que Z e δ são variáveis aleatórias independentes.

d) Seja $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ uma amostra aleatória de (Z, δ). Obtenha o estimador de máxima verossimilhança de λ, seu valor esperado e sua variância.

Resposta: Do item *c*), sabemos que para $i = 1, \ldots, n$,

$$f_{Z_i,\delta_i}(z_i, \delta_i) = \lambda^{\delta_i}\theta^{1-\delta_i}e^{-(\lambda+\theta)z_i}.$$

Da independência entre os pares (Z_i, δ_i), segue que

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{n} f_{Z_i,\delta_i}(z_i, \delta_i) = \prod_{i=1}^{n} \lambda^{\delta_i}\theta^{1-\delta_i}e^{-(\lambda+\theta)z_i} \\ &= \prod_{i=1}^{n} \lambda^{\delta_i}e^{-\lambda z_i}\theta^{1-\delta_i}e^{-\theta z_i} = \lambda^{\sum_{i=1}^{n}\delta_i}e^{-\lambda\sum_{i=1}^{n} z_i}\theta^{n-\sum_{i=1}^{n}\delta_i}e^{-\theta\sum_{i=1}^{n} z_i}, \end{aligned}$$

de modo que,

$$l(\lambda) = \left(\sum_{i=1}^{n}\delta_i\right)\ln\lambda - \lambda\sum_{i=1}^{n} z_i + \left(n - \sum_{i=1}^{n}\delta_i\right)\ln\theta - \theta\sum_{i=1}^{n} z_i.$$

Supondo que a suposição de Não Informatividade seja válida, isto é, a distribuição de C_i envolve parâmetros da densidade de T_i, segue que

isso já está dito

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\lambda}l(\lambda) = \frac{\sum_{i=1}^{n}\delta_i}{\lambda} - \sum_{i=1}^{n} z_i = 0 \Leftrightarrow \\ \sum_{i=1}^{n}\delta_i = \lambda\sum_{i=1}^{n} z_i, \end{aligned}$$

donde decorre que o estimador de máxima verossimilhança para λ $(\hat{\lambda})$ é dado por

$$\hat{\lambda} = \frac{\sum_{i=1}^{n}\delta_i}{\sum_{i=1}^{n} Z_i}.$$

Para encontrar $E(\hat{\lambda})$ e $Var(\hat{\lambda})$ vamos, em um primeiro momento, por uma questão de notac tomar

$$r = \sum_{i=1}^{n}\delta_i \text{ e } Y = \left(\sum_{i=1}^{n} Z_i\right)^{-1}.$$

Agora, $\hat{\lambda} = rY$. Como $(Z_1, \delta_1), \ldots, (Z_n, \delta_n)$ consiste de uma amostra aleatória de $(Z$ segue que $\{\delta_i\}_{i=1}^{n}$ é uma sequência de variáveis de Bernoulli independentes com parâm comum $\lambda/(\lambda+\theta)$ e $\{Z_i\}_{i=1}^{n}$ uma sequência de variáveis exponenciais independentes co mesmo parâmetro $\lambda + \theta$. Daí, por resultados já bem conhecidos, segue que

$$r \sim \text{Binomial}(n, \lambda/(\lambda+\theta)),$$

e que

$$\sum_{i=1}^{n} Z_i \sim \text{Gamma}(n, \lambda+\theta),$$

o que implica que

$$Y = \left(\sum_{i=1}^{n} Z_i\right)^{-1} \sim \text{Gamma Inversa}(n, \lambda+\theta).$$

Supondo $n > 2$, alguns resultados decorrentes são

$$E(r) = n\left(\frac{\lambda}{\lambda+\theta}\right), \quad Var(r) = n\left(\frac{\lambda}{\lambda+\theta}\right)\left(\frac{\theta}{\lambda+\theta}\right);$$
$$E(Y) = \frac{\lambda+\theta}{n-1}, \quad Var(Y) = \frac{(\lambda+\theta)^2}{(n-1)^2(n-2)}.$$

Para encontrar $E(\hat{\lambda})$, vamos usar as propriedades da esperança condicional, de modo que

$$E(\hat{\lambda}) = E(rY) = E(E(rY \mid r)) = E\left(\frac{\lambda+\theta}{n-1}r\right) = \frac{\lambda+\theta}{n-1}E(r) = \left(\frac{\lambda+\theta}{n-1}\right)\left(\frac{n\lambda}{\lambda+\theta}\right) = \frac{n\lambda}{n-1}.$$

aqui vai usar indep!

Para encontrar $Var(\hat{\lambda})$, usaremos o resultado

$$Var(rY) = E\left(Var(rY \mid r)\right) + Var\left(E(rY \mid r)\right),$$

de modo que

$$\begin{aligned}
Var(\hat{\lambda}) &= E\left(Var(rY \mid r)\right) + Var\left(E(rY \mid r)\right) \\
&= E\left(\frac{(\lambda+\theta)^2}{(n-1)^2(n-2)}r^2\right) + Var\left(\frac{\lambda+\theta}{n-1}r\right) \\
&= \frac{(\lambda+\theta)^2}{(n-1)^2(n-2)}E(r^2) + \left(\frac{\lambda+\theta}{n-1}\right)^2 Var(r) \\
&= \frac{(\lambda+\theta)^2}{(n-1)^2(n-2)}\left(Var(r) + (E(r))^2\right) + \left(\frac{\lambda+\theta}{n-1}\right)^2 Var(r) \\
&= \frac{(\lambda+\theta)^2}{(n-1)^2(n-2)}\left(\frac{n\lambda}{\lambda+\theta}\right)^2 + \left[\frac{(\lambda+\theta)^2}{(n-1)^2(n-2)} + \left(\frac{\lambda+\theta}{n-1}\right)^2\right]\frac{n\lambda\theta}{(\lambda+\theta)^2} \\
&= \frac{n\lambda^2}{(n-1)^2(n-2)} + \frac{n\lambda\theta}{(n-1)^2}\left[\frac{1}{n-2} + 1\right] \\
&= \frac{n^2\lambda^2}{(n-1)^2(n-2)} + \frac{n\lambda\theta}{(n-1)^2}\left[\frac{n-1}{n-2}\right] \\
&= \frac{n^2\lambda^2}{(n-1)^2(n-2)} + \frac{n\lambda\theta}{(n-1)(n-2)} \\
&= \frac{n\lambda(\lambda n + n\theta - \theta)}{(n-1)^2(n-2)}.
\end{aligned}$$

acho que falta

7. Dada uma covariável X, suponha que o logaritmo do tempo de sobrevivência, denotado por Y, segue um modelo linear com erro seguindo a distribuição logística, ou seja,

$$Y = \log T = \mu + \beta X + \sigma W,$$

em que W é uma variável aleatória com densidade dada por

$$f(w) = \frac{e^w}{(1+e^w)^2}, \quad -\infty < w < \infty.$$

a) Para uma observação com covarável X, encontre a função de sobrevivência condicional de T dado X, denotada por $S(t \mid X)$.

Resposta: Dada uma observação com covarável X, temos que

$$\begin{aligned} S(t \mid X) = \mathbb{P}(T > t \mid X) &= \mathbb{P}(\ln T > \ln t \mid X) \\ &= \mathbb{P}(\mu + \beta X + \sigma W > \ln t \mid X) \\ &= \mathbb{P}(\sigma W > \ln t - \mu - \beta X) \\ &= \mathbb{P}\left(W > (\ln t - \mu - \beta X)/\sigma\right) \\ &= \int_{(\ln t - \mu - \beta X)/\sigma}^{+\infty} \frac{e^w}{(1+e^w)^2} \mathrm{d}w. \end{aligned}$$

Para resolver a integral acima, vamos considerar a transformação $u = 1 + e^w$, donde dec $e^w = u - 1$ e que $\mathrm{d}u = e^w \mathrm{d}w$, isto é, $\mathrm{d}w = \mathrm{d}u/(u-1)$. Ademais, quando $w = (\ln t - \mu - \beta X)/\sigma$, $u = 1 + e^{(\ln t - \mu - \beta X)/\sigma}$, e quando $w \to +\infty$, $u \to +\infty$. Desse modo, pelo método substituição, a integral acima pode ser reescrita, de modo que

$$\begin{aligned} S(t \mid X) &= \int_{1+e^{(\ln t - \mu - \beta X)/\sigma}}^{+\infty} \frac{u-1}{u^2} \frac{\mathrm{d}u}{u-1} \\ &= \int_{1+e^{(\ln t - \mu - \beta X)/\sigma}}^{+\infty} \frac{1}{u^2} \mathrm{d}u \\ &= \left[-\frac{1}{u}\right]\Bigg|_{1+e^{(\ln t - \mu - \beta X)/\sigma}}^{+\infty} \\ &= \frac{1}{1+e^{(\ln t - \mu - \beta X)/\sigma}}, \quad t > 0. \end{aligned}$$

b) A chance de que um indivíduo apresente o evento de interesse antes de t é dada por

$$(1 - S(t \mid X))/S(t \mid X).$$

Obtenha uma expressão para a chance para este modelo.

Resposta: Usando do resultado obtido no item *a)* e a definição dada no enunciado, a cha de que um indivíduo apresente o evento de interesse antes de t é dada por

$$\begin{aligned} (1 - S(t \mid X))/S(t \mid X) &= \left[1 - \frac{1}{1+e^{(\ln t - \mu - \beta X)/\sigma}}\right] \Bigg/ \left[\frac{1}{1+e^{(\ln t - \mu - \beta X)/\sigma}}\right] \\ &= \left[\frac{1+e^{(\ln t - \mu - \beta X)/\sigma} - 1}{1+e^{(\ln t - \mu - \beta X)/\sigma}}\right] \Bigg/ \left[\frac{1}{1+e^{(\ln t - \mu - \beta X)/\sigma}}\right] \\ &= e^{(\ln t - \mu - \beta X)/\sigma}, \quad t > 0. \end{aligned}$$

c) Para dois indivíduos com valores diferentes da covariável, mostre que a razão de suas chai não depende de t.

Resposta: Denotemos por X_1 e X_2 as covariáveis associadas aos indivíduos 1 e 2 em c sideração, de modo respectivo. Pelo item *b)*, sabemos que a chance de que o indivídu $i \in \{1, 2\}$, apresente o evento de interesse antes de t, nas condições dadas, é de

$$(1 - S(t \mid X_i))/S(t \mid X_i) = e^{(\ln t - \mu - \beta X_i)/\sigma}, \quad t > 0.$$

Logo, a razão das chances desses indivíduos ou é dado pela divisão entre a chance par indivíduo 1 e a chance para o indivíduo 2, que vale

$$\left[e^{(\ln t - \mu - \beta X_1)/\sigma}\right] \Big/ \left[e^{(\ln t - \mu - \beta X_2)/\sigma}\right] = e^{(\ln t - \mu - \beta X_1 + \ln t + \mu + \beta X_2)/\sigma} = e^{-\beta(X_1 - X_2)/\sigma}, \quad t > 0$$

e não depende de t; ou então é dado pela divisão entre a chance para o indivíduo 2 e a chance para o indivíduo 1, que nada mais é do que o inverso da razão dada acima, isto é,

$$e^{-\beta(X_2-X_1)/\sigma}, \quad t > 0,$$

que também não depende de t.

Q1 2,0
Q2 1,0
Q3 0,5
Q4 2,0
Q5 1,0
Q6 1,0
Q7 1,2

8,8

MAE5783 - Análise de Sobrevivência

2ª Lista de Exercícios — Folha 01

Aluno: Maicon Aparecido Pinheiro Professora: Gisela Tunes da Silva

01. Seja T uma variável aleatória com distribuição exponencial com taxa de falha (função de risco) dada por $\alpha(t) = 3$, $t > 0$. Defina

$$\tilde{T} = \min\{T, 4\} \quad e \quad D = I(T \leq 4)$$

e considere o processo de contagem

$$N(t) = I(\tilde{T} \leq t, D = 1).$$

a. Esboce o gráfico de $N(t)$ em função de t para $D=1$ e para $D=0$.

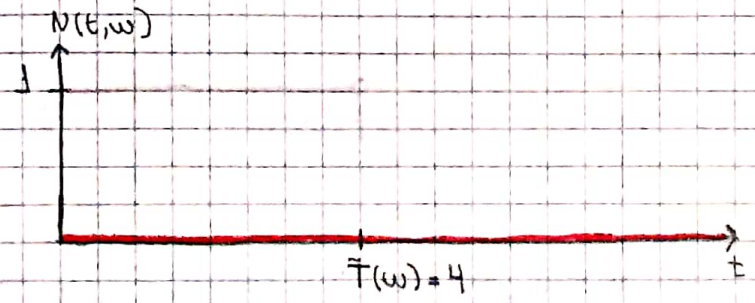
Para tornar a redação mais clara, denotemos por $(\Omega, \mathcal{F}, P)$ o espaço de probabilidade onde está definido o processo de contagem $(N(t) : t \geq 0)$ - o mesmo onde estão definidas as variáveis aleatórias T, $\tilde{T}$ e D.

Há uma certa dificuldade para esboçar o gráfico de $N(t)$ pois, na verdade, para cada $\omega \in \Omega$, temos um gráfico (ou trajetória), que denotaremos por $N(t, \omega)$.

No entanto, dado o evento $\{D=1\}$ ($\{D=0\}$), tais trajetórias apresentam um padrão em comum, que será exibido por representações gráficas logo abaixo.

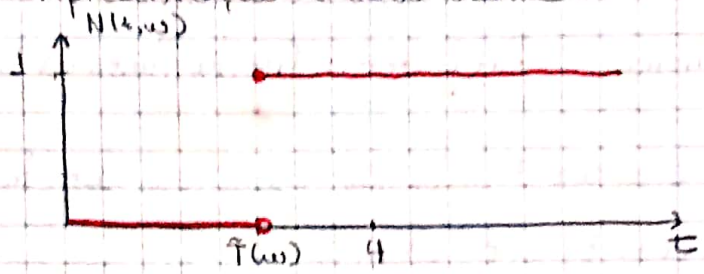
• $D=0$

Para todo $\omega \in \Omega$ em que $D(\omega) = 0$, temos $\tilde{T}(\omega) = 4$ e $N(t, \omega) = 0$ $\forall t \geq 0$, uma vez que $T(\omega) > 4$ e $D(\omega) = 1$ não é observada. Assim, para todo $\omega \in \{D=0\}$, a trajetória de $N(t)$ é dada em vermelho no gráfico abaixo.



• $D=1$

Quando $\omega \in \{D=1\}$, a trajetória $N(t,\omega)$ apresenta um salto que ocorre justamente em $\tilde{T}(\omega) = T(\omega)$ [se $\omega \in \{D=1\}$, $T(\omega) \leq 4$, i.e., $\tilde{T}(\omega) = T(\omega)$], que é no máximo igual à 4. Uma representação é dada abaixo.



, com $\tilde{T}(\omega) = T(\omega) \leq 4$.

b) Obtenha uma expressão para o processo de intensidade $\lambda(t)$ de $N(t)$ e esboce um gráfico de $\lambda(t)$ para $D=1$ e para $D=0$.

Do que já foi visto na sexta aula do curso, sabemos que

$$\lambda(t) = Y(t)\alpha(t), \text{ com } Y(t) = I(\tilde{T} \geq t), \ t \geq 0$$

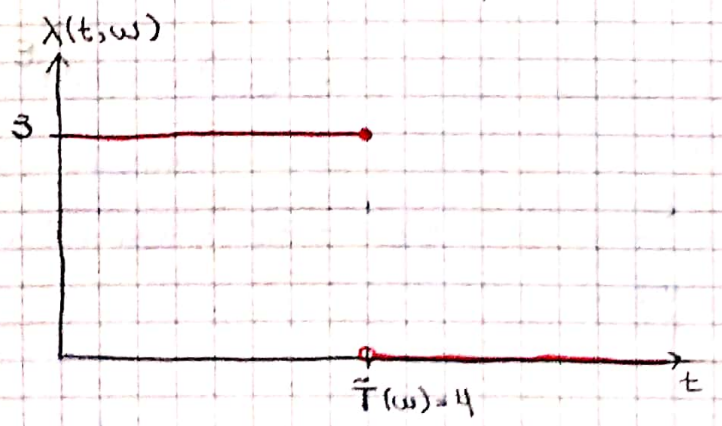
Nesse caso, em $\alpha(t) = 3$, segue que

$$\lambda(t) = 3Y(t), \quad t \geq 0.$$

Notando que $(\lambda(t) : t \geq 0)$ também está definido em $(\Omega, \mathcal{F}, \mathbb{P})$ (item anterior), vamos apresentar os gráficos que representam $\lambda(t,\omega)$ (trajetória do processo $\lambda(t)$) considerando a partição de $\{D=1\}$, $\{D=0\}$ de Ω.

• $D=0$

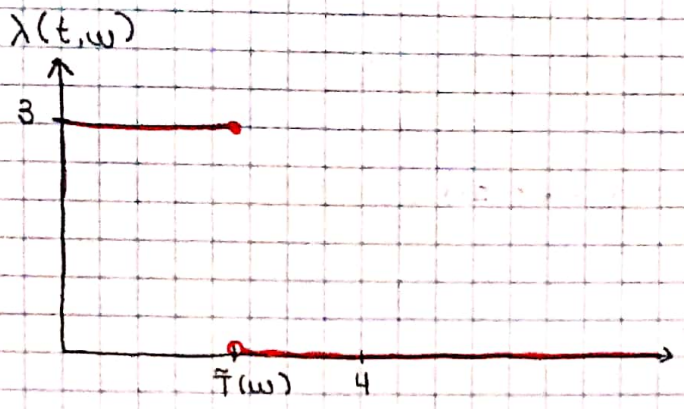
Para todo $\omega \in \Omega$ t.q. $D(\omega) = 0$, temos $\tilde{T}(\omega) = 4$. Nesse caso, $\lambda(t,\omega)$ apresenta um único comportamento, em que o salto (para baixo) ocorre em $\tilde{T}(\omega) = 4$



• $D=1$

Quando $\omega \in \{D=1\}$, $\tilde{T}(\omega) = T(\omega)$ é menor igual à 4. O salto de $\lambda(t,\omega)$

nesse instante.



$\tilde{T}(\omega) = T(\omega) \leq 4$.

c) Esboce um gráfico de $\Lambda(t) = \int_0^t \lambda(u)\,du$ em t para $D=1$ e para $D=0$.

- $D=0$

Quando $\omega \in \{D=0\}$,

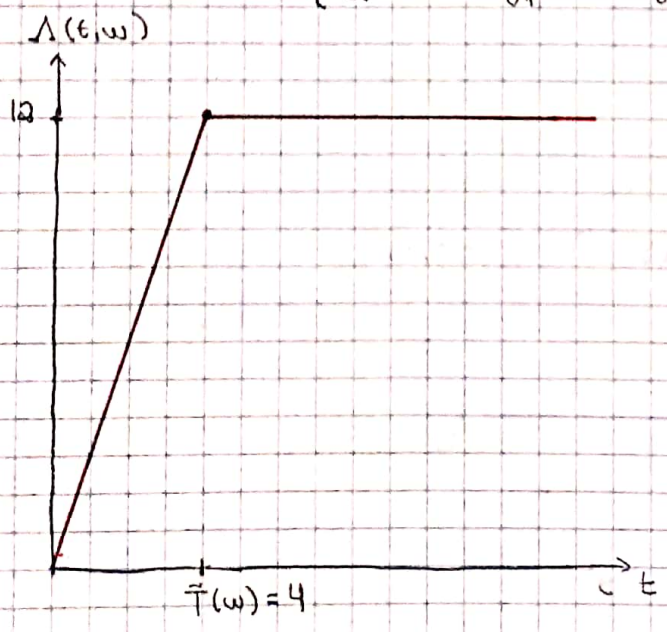
$$\lambda(t,\omega) = \begin{cases} 3, & t \leq 4 \\ 0, & \text{c.c.} \end{cases}$$

Daí

$$\Lambda(t,\omega) = \int_0^t \lambda(u)\,du$$

$$= \begin{cases} \int_0^t 3\,du = 3t, & \text{se } t \leq 4 \\ \int_0^4 3\,du + \int_4^t 0\,du = \int_0^4 3\,du = 12, & t > 4. \end{cases}$$

Assim,



- $D=1$

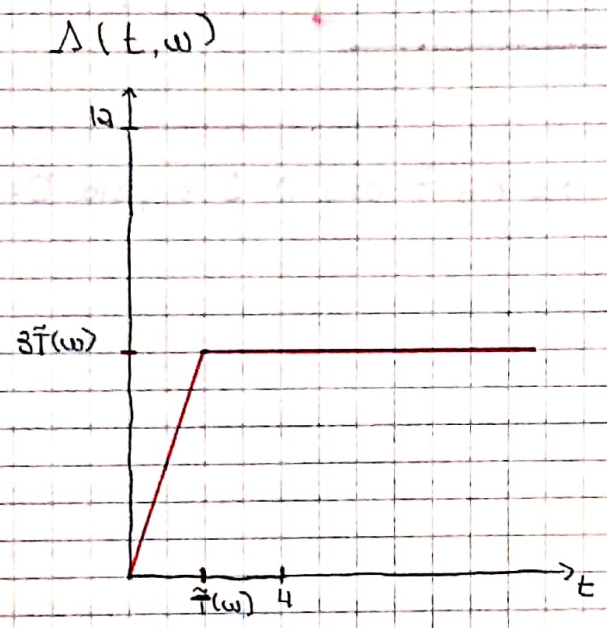
Quando $\omega \in \{D=1\}$, o salto em $\lambda(t,\omega)$ ocorre em $\tilde{T}(\omega) = T(\omega) \leq 4$.

Assim,

$$\Lambda(t,\omega) = \int_0^t \lambda(u,\omega)\,du$$

$$= \begin{cases} \int_0^t 3\,du = 3t, & \text{se } t \leq \tilde{T}(\omega) \\ \int_0^{\tilde{T}(\omega)} 3\,du + \int_0^t 0\,du = 3\tilde{T}(\omega), & \text{se } t > \tilde{T}(\omega) \end{cases}$$

Logo,



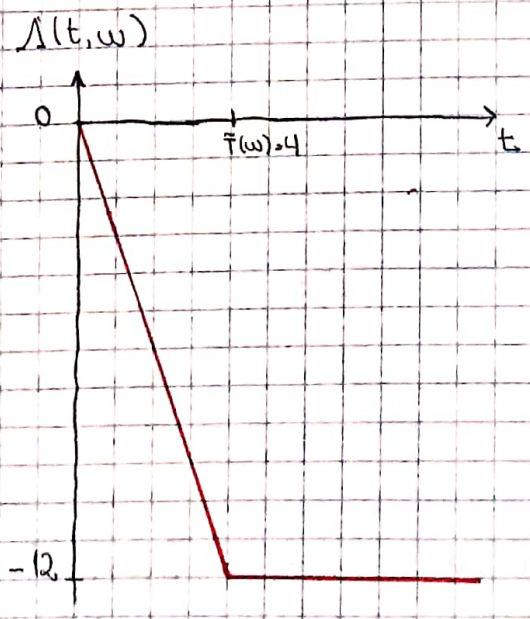
$\tilde{T}(\omega) = T(\omega) \leq 4$.

d) Seja $M(t) = N(t) - \Lambda(t)$. Esboce o gráfico de $M(t)$ em função de t para $D=1$ e para $D=0$.

. $D=0$

Do itens anteriores, segue que

$$M(t,\omega) = N(t,\omega) - \Lambda(t,\omega)$$
$$= -\Lambda(t,\omega) \quad , t \geq 0$$



MAE 5783 - Análise de Sobrevivência **Folha 02**

Aluno: Maicon Aparecido Pinheiro **Professora:** Gisela Tunes da Silva

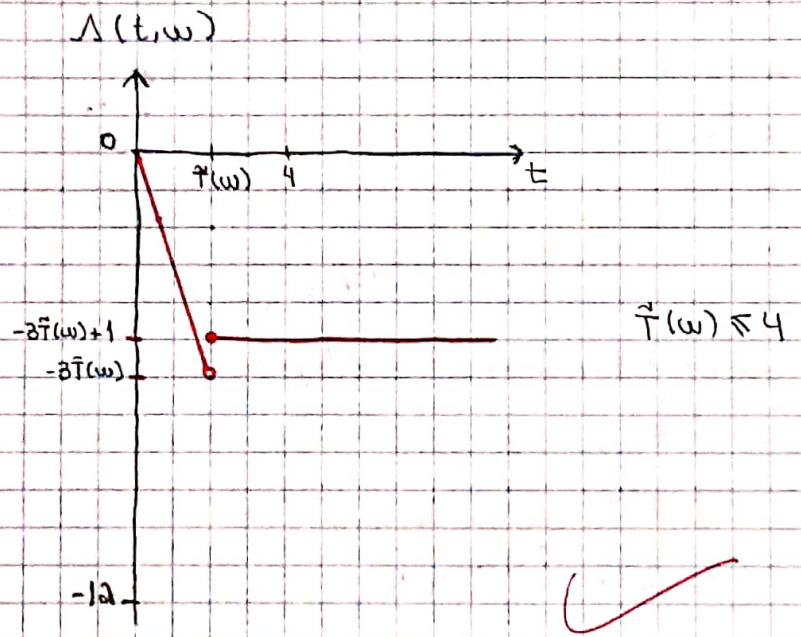
01 (...)

. $D = 1$.

Nesse caso,

$$M(t,\omega) = N(t,\omega) - \Lambda(t,\omega)$$

$$= \begin{cases} -\Lambda(t,\omega) = -3t, & t < \tilde{T}(\omega) \\ 1 - \Lambda(t,\omega) = 1 - 3\tilde{T}(\omega), & \text{se } t \geq \tilde{T}(\omega) \end{cases}$$



02. Um processo de Poisson com taxa λ, denotado por $N^*(t)$, é definido por

(i) $N^*(0) = 0$;

(ii) $N^*(t) - N^*(s)$ tem distribuição de Poisson com parâmetro $\lambda(t-s)$ para qualquer $0 \leq s \leq t$;

(iii) $N^*(t_2) - N^*(t_1)$, $N^*(t_3) - N^*(t_2)$, ..., $N^*(t_n) - N^*(t_{n-1})$ são independentes para quaisquer $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$.

a) Denotando por $\mathcal{N}_t$ a filtragem natural de $N^*(t)$, mostre que $\{N^*(t) - \lambda t : t \geq 0\}$ é um martingal.

Basta verificar que:

1. $\{N^*(t) - \lambda t : t \geq 0\}$ é adaptado à $\mathcal{N}_t$.

De fato, $\forall B \in \mathcal{B}(\mathbb{R})$,

$$\{N^*(t) - \lambda t\}^{-1}(B) = \{\omega \in \Omega : (N^*(t) - \lambda t)(\omega) \in B\}$$
$$= \{\omega \in \Omega : N^*(t)(\omega) \in B \oplus \lambda t\} \in \mathcal{N}_t,$$

pois, do enunciado, $N^*(t) \in \mathcal{N}_t$, com $B \oplus \lambda t = \{x + \lambda t : x \in B\} \in \mathcal{B}(\mathbb{R})$.

2. $E|N^*(t) - \lambda t| < \infty \quad \forall t < \infty$

Não foi mencionado, mas $\lambda > 0$ ($N^*(t) \sim \text{Poisson}(\lambda)$). Daí, basta ver que

$$E|N^*(t) - \underbrace{\lambda t}_{\geq 0}| \leq E|N^*(t)| \overset{(1)}{=} EN^*(t) \overset{(2)}{=} \lambda t < +\infty, \quad \forall t < +\infty,$$

onde as passagens (1) e (2) decorrem de $N^*(t) \sim \text{Poisson}(\lambda t)$ ser não negativa e ter esperança $E(N^*(t)) = \lambda t$.

3. $E(N^*(t+s) - \lambda(t+s) / \mathcal{F}_s) = N^*(s) - \lambda s$.

Sabemos que $N^*(s)$ e $N^*(t+s) - N^*(s)$ são independentes. Ademais

$$N^*(t+s) = (N^*(t+s) - N^*(s)) + N^*(s) \quad \text{para } s, t \in \mathbb{R}_+.$$

Agora, basta ver que

$$\begin{aligned}
E(N^*(t+s) - \lambda(t+s) / \mathcal{F}_s) &\overset{a}{=} E(N^*(t+s) - N^*(s) + N^*(s) / \mathcal{F}_s) - \lambda(t+s) \\
&\overset{b}{=} E(N^*(t+s) - N^*(s) / \mathcal{F}_s) + E(N^*(s) / \mathcal{F}_s) - \lambda(t+s) \\
&\overset{c}{=} E[N^*(t+s) - N^*(s)] + N^*(s) - \lambda(t+s) \\
&= EN^*(t+s) - EN^*(s) + N^*(s) - \lambda(t+s) \\
&\overset{d}{=} \lambda(t+s) - \lambda s + N^*(s) - \lambda(t+s) \\
&= N^*(s) - \lambda s
\end{aligned}$$

a. Reescrita de $N^*(t+s)$ e extração da constante $\lambda(t+s)$ da esperança condicional por ser independente de $\mathcal{F}_s$.

b. Linearidade da esperança condicional

c. Independência de $N^*(t+s) - N^*(s)$ e $N^*(s)$ implica $E[N^*(t+s) - N^*(s) | \mathcal{F}_s] = E[N^*(t+s) - N^*(s)]$

d. $N^*(t) \sim \text{Poisson}(\lambda t) \Rightarrow E(N^*(t)) = \lambda t \quad \forall t \in \mathbb{R}$.

Logo como $\{N^*(t) - \lambda t\}$ satisfaz as 3 propriedades em questão, segue que $\{N^*(t) - \lambda t\}$ é um martingal com respeito à filtragem $\mathcal{N}_t$.

b) Mostre que qualquer martingal tem valor esperado constante.

Seja X um martingal adaptado à filtragem $\{\mathcal{F}_t : t \geq 0\}$. Uma vez que X é martingal

$$E(X(t+s)/\mathcal{F}_s) = X_s \quad , \quad s,t \in \mathbb{R}_+ \qquad (1)$$

Daí, como para todo $t \geq 0$ e $s \geq 0$,

$$E(X(t+s)) = E(E(X(t+s)/\mathcal{F}_s)), \qquad (2)$$

segue substituindo (1) em (2) que

$$E(X(t+s)) = E(X_s) \quad \forall\, t \geq 0, s \geq 0. \qquad (3)$$

Tomando $s=0$ em (3), segue o resultado, isto é,

$$E(X(t)) = E(X(0)) \quad \forall\, t \geq 0.$$

c) Use a e b para obter $E[N^*(t)]$.

Sabemos por a) que $\{N^*(t) - \lambda t : t \geq 0\}$ é um martingal com respeito a $\mathcal{N}_t$.

Assim, por b), segue que

$$(1) \qquad E(N^*(t) - \lambda t) = E(N^*(0) - \lambda \cdot 0) = E[N^*(0)] \overset{(i)}{=} E[0] = 0, \ \forall t \geq 0.$$

Logo, pela linearidade da esperança, segue de (1) que

$$0 = E(N^*(t) - \lambda t) = E(N^*(t)) - \lambda t \quad , \text{ isto é,}$$

$$E[N^*(t)] = \lambda t \quad , \ \forall t \geq 0,$$

como devia ser, já que $N^*(t) \sim \text{Poisson}(\lambda t)$.

03. Sejam T_i, $i = 1, \ldots, n$, tempos de falha sujeitos à censura à direita. Denote por C_i os tempos de censura e defina $Z_i = \min(T_i, C_i)$ e $\delta_i = I(T_i \leq C_i)$. Considere ainda que se tem informação sobre uma covariável numérica x_i e que a função de taxa de falha para a i-ésima observação é dada por

$$\alpha_i(t) = \alpha_0(t)\, e^{x_i \beta},$$

em que β é um parâmetro desconhecido. Para cada observação, defina os processos

$$N_i(t) = I(Z_i \leq t, \delta_i = 1)$$

$$Y_i(t) = I(Z_i \geq t)$$

Escreva a seguinte expressão utilizando a notação de processos de contagem

$$\sum_{i=1}^{n} \delta_i \beta x_i - \sum_{i=1}^{n} \delta_i \log\left[\sum_{j \in R_i} e^{x_j \beta}\right]$$

em que R_i denota o conjunto de indivíduos em risco no instante Z_i.

Ao ver que

$$N_i(t) = \begin{cases} 1, & \text{se ocorreu a falha do } i\text{-ésimo indivíduo antes de } t \\ 0, & \text{se não ocorreu.} \end{cases}$$

Como

$$\delta_i = \begin{cases} 1, & \text{se o } i\text{-ésimo indivíduo falhou} \\ 0, & \text{" " " " " não falhou} \end{cases}$$

podemos tomar

$$\delta_i = \lim_{t \to \infty} N_i(t), \text{ ou,}$$

de outra forma,

$$\delta_i = \int_0^{\tau} dN_i(u) \quad \text{com } \tau > \max\{T_i, C_i : i = 1, \dots, n\}$$

Note que τ existe quase certamente, pois T_i e C_i são finitos quase certamente, $i = 1, \dots, n$.

Ademais,

$$Y_i(t) = \begin{cases} 1, & \text{se o } i\text{-ésimo indivíduo não falhou nem foi censurado até } t \text{ (estáem risco em } t\text{)} \\ 0, & \text{c.c.} \end{cases}$$

com $i = 1, 2, \dots, n$

Outra forma de ver se o i-ésimo indivíduo está em risco em t e averiguar se

$$\int_0^t dY_i(u) = 0 \quad \text{ou, de modo equivalente,}$$

$$1 - \int_0^t dY_i(u) = 1.$$

isso só ajuda aproximamos se Y_i for constante em $[0, t]$ ou $Y_i(u)$ for constante se $\Delta_i(u) = 0$ $\forall u \in [0, t]$ pois $\int_0^t dY_i(u) \neq \int_0^t Y_i(u)\, du$ e $\int_0^t dY_i(u) = 0$

Assim,

$$\sum_{i=1}^n \delta_i \beta x_i - \sum_{i=1}^n \delta_i \log\left[\sum_{j \in R_i} e^{x_j \beta}\right]$$

$$= \sum_{i=1}^n \left[\int_0^{\tau} dN_i(u)\right] \beta x_i - \sum_{i=1}^n \left[\int_0^{\tau} dN_i(u)\right] \log\left[\sum_{j=1}^n \left(1 - \int_0^{Z_i} dY_j(u)\right) e^{x_j \beta}\right]$$

$$= \beta \sum_{i=1}^n \int_0^{\tau} x_i \, dN_i(u) - \sum_{i=1}^n \left[\int_0^{\tau} dN_i(u)\right] \log\left[\sum_{j=1}^n \left(1 - \int_0^{Z_i} dY_j(u)\right) e^{x_j \beta}\right]$$

Obs: Não consegui verificar se há e quais seriam as condições de regularidade para a troca da ordem dos operadores, a fim de simplificar ainda mais a expressão.

04. Assuma que (T, C) tenha distribuição conjunta dada por $P(T > t, C > s) = H(t, s)$, com

$$H(t, s) = \exp\{-\lambda t - \mu s - \theta t s\}$$

em que $0 \leq \theta \leq \lambda \mu$.

MAE 5783 - Análise de Sobrevivência **Folha 03**

Aluno: Marcon Aparecido Pinheiro Professora: Gisela Tunes da Silva

Q4 (...)

a) Encontre a função de sobrevivência marginal $H_1(t) = \mathbb{P}(T > t)$ e a correspondente função de taxa de falha $\alpha_1(t)$. Encontre também a função de sobrevivência marginal $H_2(t) = \mathbb{P}(C > s)$ e a correspondente taxa de falha $\alpha_2(s)$.

Em primeiro momento, reparamos que T e C são não-negativas, uma vez que

$$\mathbb{P}(T \geq 0, C \geq 0) = e^{-\lambda \cdot 0 - \mu \cdot 0 - \theta \cdot 0 \cdot 0} = e^{-0} = 1.$$

Dito isso, basta ver que

$$H_1(t) = \mathbb{P}(T > t) = \mathbb{P}(T > t, C \geq 0) = e^{-\lambda t - \mu \cdot 0 - \theta t \cdot 0} = e^{-\lambda t}, \quad t \geq 0$$

donde decorre que

$$\alpha_1(t) = -\frac{d}{dt} \ln H_1(t) = -\frac{d}{dt}(-\lambda t) = \lambda.$$

De modo análogo,

$$H_2(s) = \mathbb{P}(C > s) = \mathbb{P}(T \geq 0, C > s) = e^{-\lambda \cdot 0 - \mu s - \theta \cdot 0 \cdot s} = e^{-\mu s}, \quad s \geq 0$$

e

$$\alpha_2(s) = -\frac{d}{ds} \ln H_2(s) = -\frac{d}{ds}(-\mu s) = \mu.$$

Em suma,

$$H_1(t) = e^{-\lambda t} \text{ e } \alpha_1(t) = \lambda, \quad \lambda \geq 0, \ t \geq 0$$

$$H_2(s) = e^{-\mu t} \text{ e } \alpha_2(s) = \mu, \quad \mu \geq 0, \ t \geq 0.$$

b) Calcule

$$\alpha_1^* = \frac{-\frac{\partial \mathbb{P}(T \geq u, C \geq t)}{\partial u}\Big|_{u=t}}{\mathbb{P}(T \geq t, C \geq t)}$$

e

$$\alpha_2^* = \frac{\frac{\partial \mathbb{P}(T \geq t, C \geq u)}{\partial u}\Big|_{u=t}}{\mathbb{P}(T \geq t, C \geq t)}$$

Em primeiro momento, notemos que

$$\mathbb{P}(T \geq t, C > s) = \lim_{(m,n) \to (0,0)} \mathbb{P}(T \geq t + m, C > s + n)$$

$$= \lim_{(m,n) \to (0,0)} e^{-\lambda t} e^{-\lambda m} e^{-\mu s} e^{-\mu n} e^{-\theta t} e^{-\theta m} e^{-\theta s} e^{-\theta n}$$

$$= e^{-\lambda t} e^{-\mu s} e^{-\theta s} \lim_{(m,n) \to (0,0)} e^{-\lambda m} e^{-\mu n} e^{-\theta m} e^{-\theta n}$$

$$= \mathbb{P}(T > t, C > s), \text{ isto é,}$$

o vetor (T, C) é absolutamente contínuo com respeito a medida de Lebesgue em $\mathbb{R}^2$

Daí,

$$\alpha_1^* = \frac{-\frac{\partial \mathbb{P}(T \geq u, C \geq t)}{\partial u}\Big|_{u=t}}{\mathbb{P}(T \geq t, C \geq t)} = \frac{-\frac{\partial}{\partial u}\left[e^{-\lambda u - \mu t - \theta u t}\right]\Big|_{u=t}}{e^{-\lambda t - \mu t - \theta t^2}}$$

$$= \frac{-e^{-\lambda u - \mu t - \theta u t}(-\lambda - \theta t)\big|_{u=t}}{e^{-\lambda t - \mu t - \theta t^2}}$$

$$= \frac{e^{-\lambda t - \mu t - \theta t^2}(\lambda + \theta t)}{e^{-\lambda t - \mu t - \theta t^2}} = \lambda + \theta t, \quad t \geq 0.$$

e

$$\alpha_2^* = \frac{\frac{\partial \mathbb{P}(T \geq t, C \geq u)}{\partial u}\Big|_{u=t}}{\mathbb{P}(T \geq t, C \geq t)} = \frac{-\frac{\partial}{\partial u}\left[e^{-\lambda t - \mu u - \theta t u}\right]\Big|_{u=t}}{e^{-\lambda t - \mu t - \theta t^2}}$$

$$= \frac{-e^{-\lambda t - \mu u - \theta t u}(-\mu - \theta t)\big|_{u=t}}{e^{-\lambda t - \mu t - \theta t^2}}$$

$$= \frac{e^{-\lambda t - \mu t - \theta t^2}(\mu + \theta t)}{e^{-\lambda t - \mu t - \theta t^2}} = \mu + \theta t, \quad t \geq 0.$$

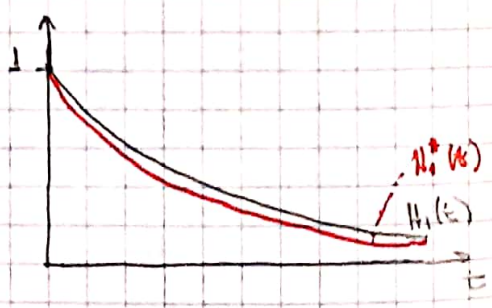
c) Assuma $\lambda = 1$, $\mu = 2$ e $\theta = 2$. Compare $H_1(t)$ com $H_1^*(t) = \exp\left\{-\int_0^t \alpha_1^*(u)\,du\right\}$ e compare $H_2(t)$ com $H_2^*(t) = \exp\left\{-\int_0^t \alpha_2^*(u)\,du\right\}$.

Por a) e b),

$$H_1(t) = e^{-\lambda t} = e^{-t}$$
$$\alpha_1^*(t) = \lambda + \theta t = 1 + 2t, \quad ; \quad t \geq 0$$

Daí,

$$H_1^*(t) = \exp\left\{-\int_0^t 1 + 2u\,du\right\}$$
$$= \exp\left\{-\left[u + u^2\right]\Big|_0^t\right\}$$
$$= \exp\{-t - t^2\}$$
$$= e^{-t}e^{-t^2} = e^{-t^2} \times H_1(t), \quad t \geq 0$$



Como $e^{-t^2} \leq 1$ $\forall t \geq 0$, temos $H_1^*(t) \leq H_1(t)$ $\forall t \geq 0$. De fato, $H_1^*(t) = e^{-t^2} H_1(t)$.

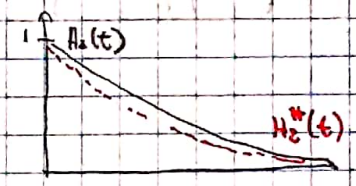
De modo análogo, por a) e b),

$$H_2(t) = e^{-\mu t} = e^{-2t}$$
$$\alpha_2(t) = \mu + \theta t = 2 + 2t \quad ; \quad t \geq 0$$

Assim,

$$H_2^*(t) = \exp\left\{-\int_0^t 2 + 2u\,dt\right\} = \exp\left\{-\left[2u + u^2\right]\Big|_0^t\right\} = \exp\{-2t - t^2\} = e^{-t^2}e^{-2t}$$

Como $e^{-t^2} \leq 1$, $H_2^*(t) \leq H_2(t)$, $\forall t \geq 0$. De modo mais específico, $H_2^*(t) = e^{-t^2} H_2(t)$.



d) Assuma T e G independentes com distribuições dadas pelas marginais obtidas em a). Calcule α_1^*, $H_1^*(t)$, α_2^* e $H_2^*(t)$ neste caso.

Nessas condições,

$$\mathbb{P}(T \geq t, G \geq s) = \mathbb{P}(T \geq t)\,\mathbb{P}(G \geq s) = e^{-\lambda t} e^{-\mu s}, \quad t \geq 0, s \geq 0.$$

Assim, $\forall t \geq 0$,

$$\alpha_1^* = \frac{-\frac{\partial}{\partial u} e^{-\lambda u - \mu t}\Big|_{u=t}}{e^{-\lambda t} e^{-\mu t}} = \frac{e^{-\mu t} \lambda e^{-\lambda u}\Big|_{u=t}}{e^{-\lambda t} e^{-\mu t}} = \frac{\lambda e^{-\mu t} e^{-\lambda t}}{e^{-\lambda t} e^{-\mu t}} = \lambda = \alpha_1(t) \quad e$$

$$H_1^*(t) = \exp\left\{-\int_0^t \lambda\, du\right\} = \exp\left\{-\lambda u\Big|_0^t\right\} = e^{-\lambda t} = H_1(t);$$

$$\alpha_2^*(t) = \frac{-\frac{\partial}{\partial u} e^{-\lambda t} e^{-\mu u}\Big|_{u=t}}{e^{-\lambda t} e^{-\mu t}} = \frac{e^{-\lambda t} \mu e^{-\mu u}\Big|_{u=t}}{e^{-\lambda t} e^{-\mu t}} = \frac{\mu e^{-\lambda t} e^{-\mu t}}{e^{-\lambda t} e^{-\mu t}} = \mu = \alpha_2(t),$$

$$H_2^*(t) = \exp\left\{-\int_0^t \mu\, du\right\} = \exp\left\{-\mu u\Big|_0^t\right\} = e^{-\mu t} = H_2(t)$$

05. Seja $X = \{X(t) : t \geq 0\}$ um processo estocástico contínuo à direita adaptado à filtragem $\{\mathcal{F}_t : t \geq 0\}$. Dizemos que $X(t)$ é um submartingal se

(i) $X(t)$ é adaptado à $\{\mathcal{F}_t : t \geq 0\}$

(ii) $E|X(t)| < +\infty \ \forall t < +\infty$;

(iii) $E(X(t+s) \mid \mathcal{F}_s) \geq X_s$.

a. Suponha que $X(t)$ é um processo crescente, adaptado a $\mathcal{F}_t$, com $E|X(t)| < +\infty$ para todo t. Mostre que $X(t)$ é um martingal.

Do enunciado, temos as condições (i) e (ii) satisfeitas. Basta verificar que

$$E[X(t+s) \mid \mathcal{F}_s] \geq X_s.$$

De fato, se $X(t)$ é crescente, então $X(t) \geq X(s)$ $\forall s \leq t$, $s, t \in \mathbb{R}_+$, i.e.,

$$X(t) - X(s) \geq 0 \quad \forall s < t, \ s, t \in \mathbb{R}_+,$$

donde segue que $E(X(t+s)-X(s)/\mathcal{F}_s) \geq 0$ e, portanto,

$$E[X(t+s)/\mathcal{F}_s] = E(X(t+s)-X(s)+X(s)/\mathcal{F}_s)$$
$$= E(X(t+s)-X(s)/\mathcal{F}_s) + E(X(s)/\mathcal{F}_s)$$
$$= \underbrace{E(X(t+s)-X(s)/\mathcal{F}_s)}_{\geq 0} + X(s)$$
$$\geq X(s) \qquad \forall s<t,\ s,t \in \mathbb{R}_+$$

Como vale (i), (ii) e (iii), temos que $X(t)$ é um submartingal.

b) Seja $M(t)$ um martingal com $EM^2(t) < \infty\ \forall t < \infty$. Prove que $M^2(t)$ é submartingal.

Temos que verificar

(i) $M^2(t)$ é adaptado à $\{\mathcal{F}_t : t \geq 0\}$, a filtragem subjacente.

Como $M(t) \in \mathcal{F}_t\ \forall t \geq 0$ e $M^2(t) = (M(t))^2$, segue que $\forall B \in \mathcal{B}(\mathbb{R})$,

$$(M^2(t))^{-1}(B) = \{\omega \in \Omega : M^2(\omega) \in B\}$$
$$= \{\omega \in \Omega : M(\omega) \in g^{-1}(B)\} \in \mathcal{F}_t,\ \forall t \geq 0$$

onde $g^{-1}(B)$ é a imagem inversa da função $g(x)=x^2$ no conjunto B, que pertence à $\mathcal{B}(\mathbb{R})$, uma vez que a imagem inversa de um conjunto formado a partir de uniões limitadas de int. semi-abertos por uma função contínua tbm será um conjunto com tais propriedades.

(ii) $E|M^2(t)| = EM^2(t) \overset{\text{Enunciado}}{<} +\infty\ \forall t < +\infty$, $(|x^2| = x^2)$.

(iii) $E(M^2(t+s)/\mathcal{F}_s) \geq M^2(s)$

Para tal, basta usar a desigualdade de Jensen para esperanças condicionais tomando como função convexa $\varphi(x)=x^2$, isto é,

$$E(\varphi(M(t+s)/\mathcal{F}_s) \geq \varphi(E(M(t+s)/\mathcal{F}_s) \quad \Leftrightarrow \quad \varphi(x)=x^2$$
$$E(M^2(t+s)/\mathcal{F}_s) \geq (E(M(t+s)/\mathcal{F}_s))^2 \quad \Leftrightarrow \quad M(t) \text{ é martingal}$$
$$E(M^2(t+s)/\mathcal{F}_s) \geq (M(s))^2 = M^2(s).$$

Logo, como valem (i), (ii) e (iii), $M^2(t)$ é submartingal.

06. Os dados deste exercício referem-se a um estudo com pacientes tratados na Mayo Clinic com câncer de ovário. Os dados são referentes a tempos (em dias) até o óbito do paciente e os pacientes foram divididos em dois grupos: que apresentavam tumor grande e os que

MAE 5783 - Análise de Sobrevivência **Folha 04**

Aluno: Maicon Aparecido Pinheiro **Professora:** Gisela Tunes da Silva

06 (...) apresentavam tumor pequeno. Os tempos censurados à direita são denotados por um sinal "+".

Tamanho do Tumor	Tempo (dias)
Grande	28 89 175 195 309 377+ 393+ 421+ 447+ 462 709+ 744+ 770+ 1106+ 1206
Pequeno	34 88 137 199 280 291 299+ 300+ 309 351 358 369 369 370 375 382 392 429+ 451 1119+

Com base nos dados apresentados abaixo, obtenha a função de sobrevivência estimada (Kaplan-Meier) e a função de taxa de falha acumulada estimada (pelo estimador de Nelson-Aalen) para cada grupo, sem auxílio computacional. Represente graficamente as curvas obtidas.

- Grupo dos pacientes que apresentavam tumor grande.

Nesse, 15 observações e 7 instantes de falha, que são $t_1 = 28$, $t_2 = 89$, $t_3 = 175$, $t_4 = 195$, $t_5 = 309$, $t_6 = 462$ e $t_7 = 1206$, e 8 censuras. Vamos aos cálculos:

t_j	d_j	r_j	$r_j - d_j$	$(r_j - d_j)/r_j$	$\hat{S}(t_j)$	d_j/r_j	$\hat{\Lambda}(t_j)$
28	1	15	14	14/15	14/15	1/15	0,0667
89	1	14	13	13/14	13/15	1/14	0,1381
175	1	13	12	12/13	12/15	1/13	0,2150
195	1	12	11	11/12	11/15	1/12	0,2984
309	1	11	10	10/11	10/15	1/11	0,3893
462	1	6	5	5/6	5/9	1/6	0,5559
1206	1	1	0	0	0	1	1,5559

Lembrete:

d_j: # falhas em t_j

r_j: # ind. em risco em t_j

- Estimado K-M

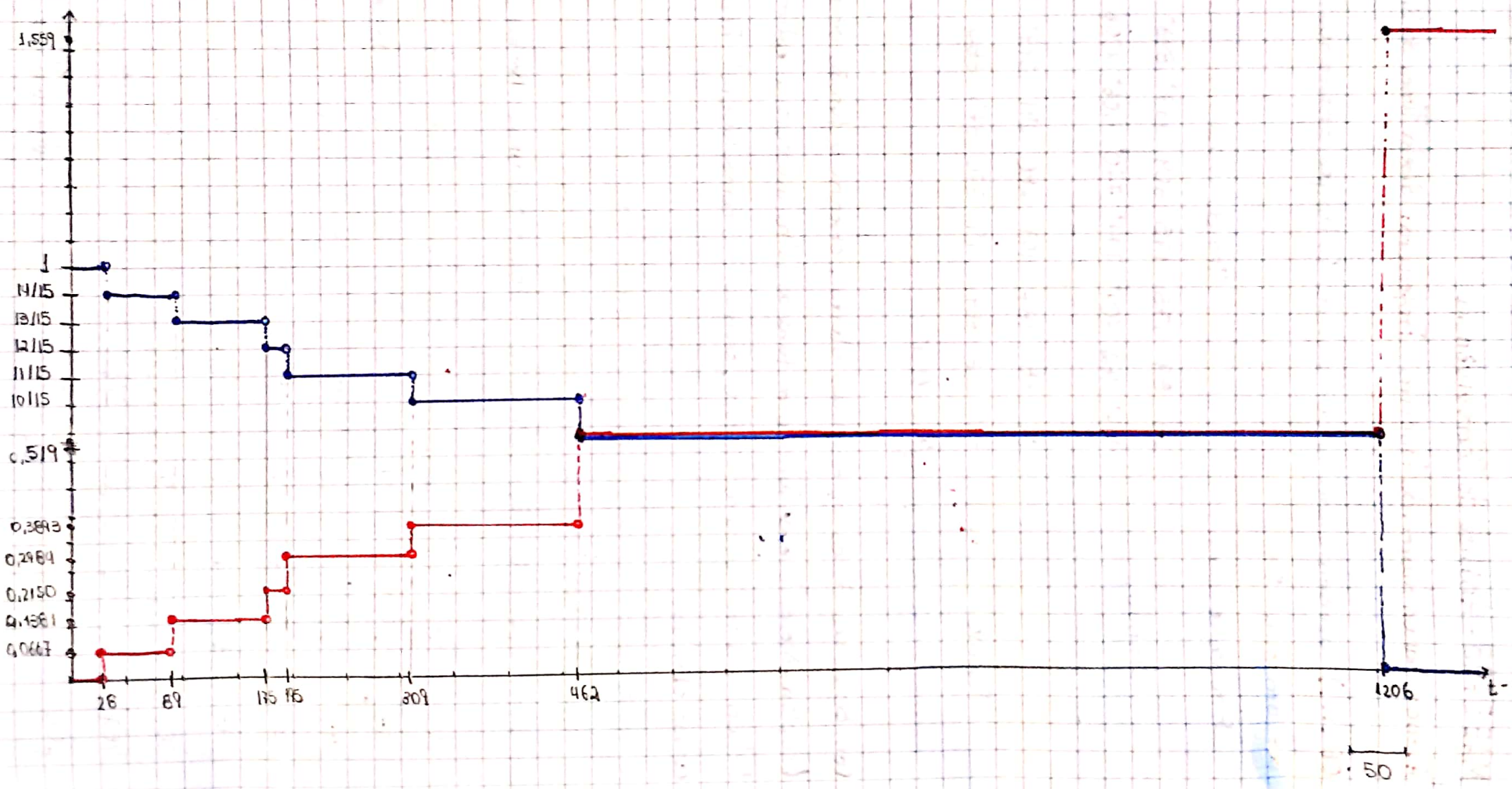
$$\hat{S}(t) = \prod_{j: t_j \le t} \frac{r_j - d_j}{r_j}$$

- Estimado Nelson-Aalen

$$\hat{\Lambda}(t) = \sum_{j: t_j \le t} \frac{d_j}{r_j}$$

Vamos à representação gráfica de $\hat{S}(t)$ e $\hat{\Lambda}(t)$.

$\hat{S}(t), t \geq 0$ e $\hat{A}(t), t \geq 0$.



• Grupo dos pacientes que apresentavam tumor pequeno.

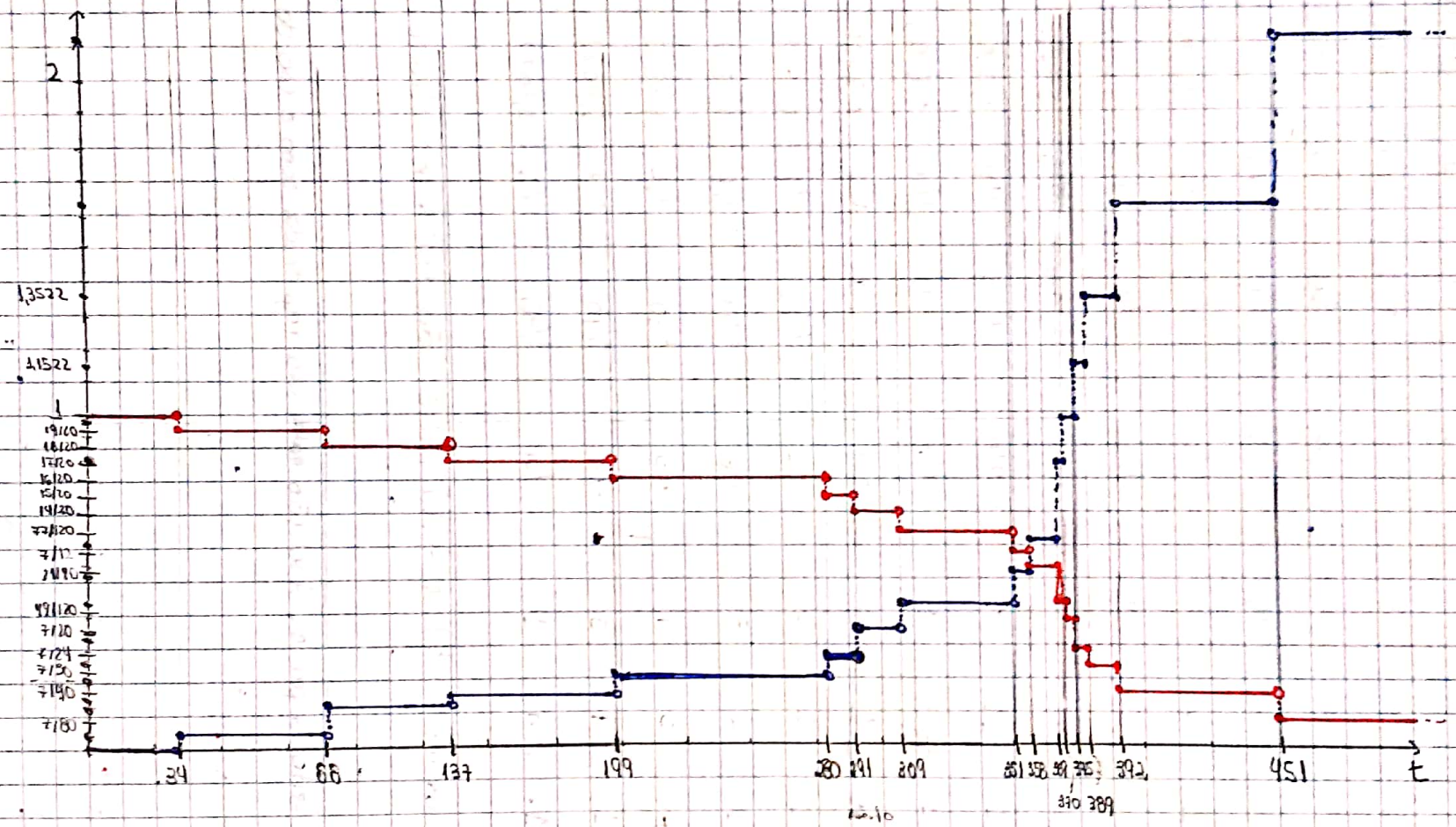
Nesse, 20 observações e 15 instantes de falha e 4 censuras.

t_j	d_j	r_j	$r_j - d_j$	$(r_j - d_j)/r_j$	$\hat{S}(t_j)$	d_j/r_j	$\hat{A}(t_j)$
34	1	20	19	19/20	19/20	1/20	0,0500
88	1	19	18	18/19	18/20	1/19	0,1026
137	1	18	17	17/18	17/20	1/18	0,1582
199	1	17	16	16/17	16/20	1/17	0,2170
280	1	16	15	15/16	15/20	1/16	0,2795
291	1	15	14	14/15	14/20	1/15	0,3462
309	1	12	11	11/12	77/120	1/12	0,4295
351	1	11	10	10/11	7/12	1/11	0,5202
358	1	10	9	9/10	21/40	1/10	0,6202
369	2	9	7	7/9	49/120	2/9	0,8426
370	1	7	6	6/7	7/20	1/7	0,9855
375	1	6	5	5/6	7/24	1/6	1,1522
382	1	5	4	4/5	7/30	1/5	1,3522
392	1	4	3	3/4	7/40	1/4	1,6022
451	1	2	1	1/2	7/80	1/2	2,1022

Os gráficos de $\hat{S}(t)$ e $\hat{A}(t)$ $t \geq 0$ são dados no outro lado dessa página.

07. Será feito à parte. A solução ~~será~~ será registrada com o uso do computador.

$\hat{S}(t), t \geq 0.$ $\hat{A}(t), t \geq 0.$



25

ANÁLISE DE SOBREVIVÊNCIA - LISTA 2

Maicon Aparecido Pinheiro - EXERCÍCIO 7

7. Um estudo com pacientes diagnosticados com câncer de língua foi conduzido para verificar o efeito prognóstico de alterações no número de cromossomos (aneuploidia) no DNA das células do tumor. Tecidos do tumor dos pacientes foram analisados segundo uma determinada técnica para determinar se o tumor continha células com aneuploidia (anormal) ou diploidia (normal). Os dados estão disponíveis no arquivo `Lista2-cancerLingua.txt` e a descrição dos dados no arquivo `Lista2-cancerLingua.des`

a) Obtenha o estimador de Kaplan-Meier dos dados dos pacientes considerados (todos os pacientes) - você pode utilizar um software e apresentar o gráfico do estimador. Construa também curvas de Kaplan-Meier para pacientes segundo perfil do DNA do tumor (com aneuploidia/normal).

Resposta: Para a confecção dessa figura, utilizei o pacote `OIsurv`, no software *R*.

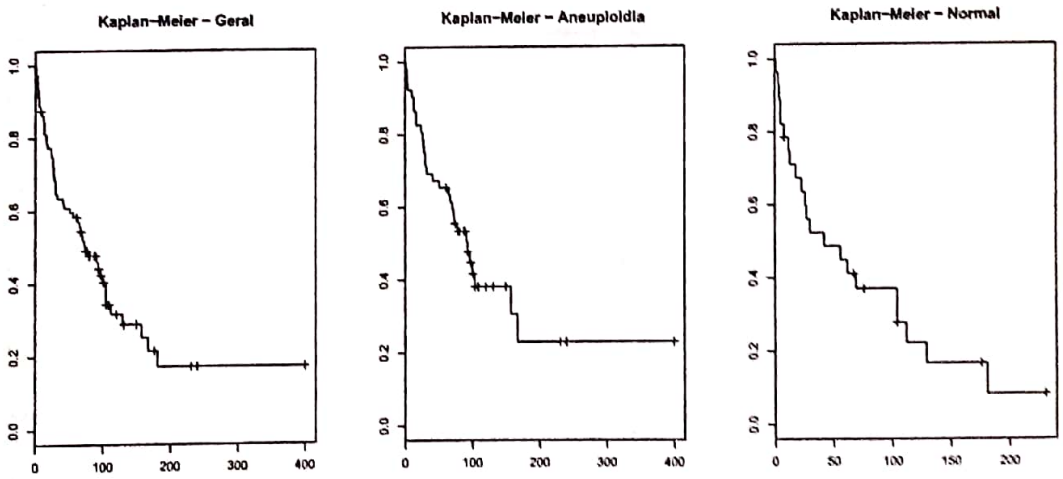


Figura 1: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier, para todos os pacientes, para os com perfil do DNA do tumor com aneuploidia e normal, da esquerda para a direita.*

O código é dado abaixo.

```
install.packages("OIsurv")
library(OIsurv)

setwd('C:\\Users\\Maicon Pinheiro\\Desktop\\2016\\4. Análise de Sobrevivência
      \\Lista 2\\Gabarito Questão 7')
tongue<-read.table('Ex7.txt')
names(dados)<-c('type','time','delta')

my.surv.object <- Surv(time, delta)
my.surv.object.1 <- Surv(time[type==1], delta[type==1])
```

```
my.surv.object.2 <- Surv(time[type==2], delta[type==2])
detach(tongue)

survfit(my.surv.object ~ 1)
survfit(my.surv.object.1 ~ 1)
survfit(my.surv.object.2 ~ 1)

par(mfrow=c(1,3))
plot(survfit(my.surv.object ~ 1),conf.int=FALSE,
     main="Kaplan-Meier - Geral")
plot(survfit(my.surv.object.1 ~ 1),conf.int=FALSE,
     main="Kaplan-Meier - Aneuploidia ")
plot(survfit(my.surv.object.2 ~ 1),conf.int=FALSE,
     , main="Kaplan-Meier - Normal")
```

c) Obtenha intervalos de confiança com transformação logarítmica para $S(t)$, com coeficiente de confiançaa igual a 95%, e apresente o gráfico (juntamente com a curva de Kaplan-Meier).

Resposta: Para a confecção dessa figura, utilizei o pacote `OIsurv`, no software *R*.

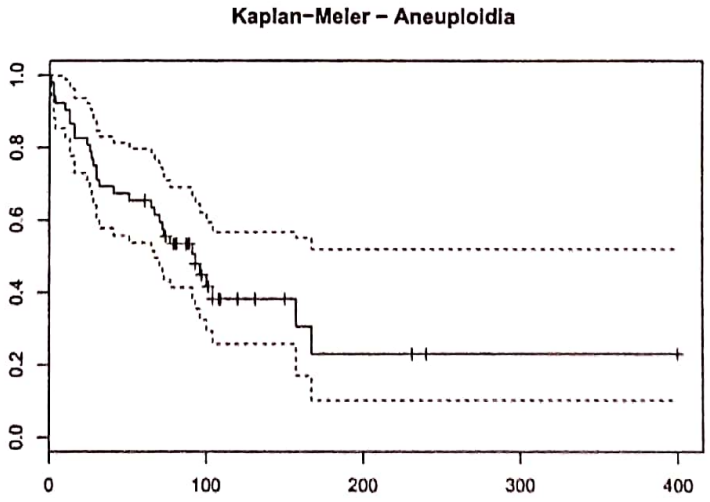


Figura 2: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier, para todos os pacientes, para os com perfil do DNA do tumor com aneuploidia e normal, da esquerda para a direita, com intervalos de confiança com transformação logarítmica para $S(t)$, com coeficiente de confiança igual a 95%.*

O código adicional para a geração dessa figura é dado abaixo.

```
par(mfrow=c(1,1))
plot(survfit(my.surv.object.1 ~ 1,conf.int = 0.95,conf.type = "log"),
     main="Kaplan-Meier - Aneuploidia ")
```

d) Consulte a seção 4.4 do livro de Klein e Moeschberger, façaa uma breve descrição da obtenção de alguma banda de confiança para o estimador de Kaplan-Meier e obtenha a banda de confiança para o estimador de KM. Compare com os intervalos pontuais obtidos em (c).

?

Dica: Bandas de confiança estão implementadas no SAS e no R (biblioteca km.ci).

Resposta:

Resposta: Para a confecção dessa figura, utilizei o pacote `OIsurv`, no software *R*.

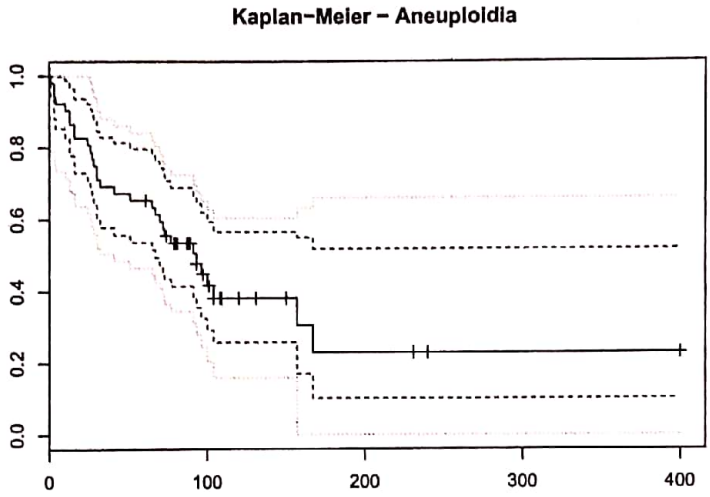


Figura 3: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier, para todos os pacientes, para os com perfil do DNA do tumor com aneuploidia e normal, da esquerda para a direita, com intervalos de confiança com transformação logarítmica e bandas de confiança (linha mais clara) para para $S(t)$, com coeficiente de confiança igual a 95%.*

O código adicional para a geração dessa figura é dado abaixo.

```
my.cb.1 <- confBands(my.surv.object.1, confLevel=0.95, type="hall")

plot(survfit(my.surv.object.1 ~ 1,conf.int = 0.95,conf.type = "log"),
     main="Kaplan-Meier - Aneuploidia ")
lines(my.cb.1$time, my.cb.1$lower, lty=3, type="s")
lines(my.cb.1$time, my.cb.1$upper, lty=3, type="s")
```

Faltou o item (b) ?

Q1 1,5
Q2 2,5
Q3 0,5
Q4 1,5
Q5 1,2
7,2

ANÁLISE DE SOBREVIVÊNCIA - LISTA 3

Maicon Aparecido Pinheiro

1. Um produtor de requeijão deseja comparar dois tipos de embalagens (A e B) para seu produto. Ele deseja saber se existe diferença na durabilidade de seu produto com relação às embalagens. O produto dele é vendido a temperatura ambiente e sem conservantes. O evento de interesse é o aparecimento de algum tipo de fungo no produto e os dados estão apresentados na tabela abaixo.

Embalagem	Tempo (horas)
A	31 33 36 40 40 42 43 44 44 46 46 47 48 48
	49 50 50 60 60 60 60 60^+ 60^+ 60^+ 60^+ 60^+
B	47 48 48 48 49 49 49 49 50 50 50 50 53 53
	53 53 54 54 54 55 55^+ 55^+ 55^+ 55^+ 55^+ 55^+

a) Obtenha as curvas de Kaplan-Meier do tempo até o aparecimento de fungos para as duas embalagens.

Resposta: As curvas e o respectivos código são dados abaixo.

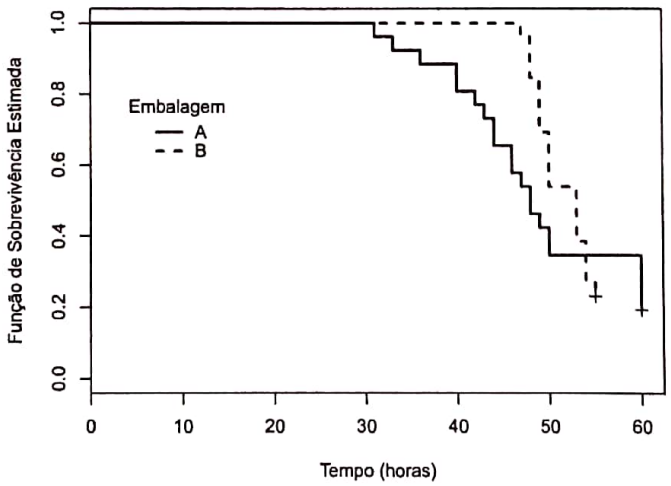


Figura 1: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier para o tempo até o aparecimento de fungos, por tipo de embalagem.*

Por uma análise prévia das curvas, há dúvidas sobre qual das embalagens é a mais adequada. Para tempos menores, parece que B apresenta o melhor desempenho. No entanto, para tempos maiores, o cenário se inverte.

Para construir a Figura 1 usamos o pacote `Survival` do sofware R através do seguinte código:

```
rm(list=ls())
require(survival)
setwd('C:/Users/Maicon Pinheiro/Desktop/2016/2. Análise de Sobrevivência/
      Lista 3/Gabarito Lista 3')

dados = read.table('Dados.txt', h=T)
attach(dados)

dados2 = Surv(time=Tempo, event=Falha, type='right')

# Curvas de Kaplan-Meier
km = survfit(dados2~Grupo)
plot(km, xlab='Tempo (horas)', ylab='Função de Sobrevivência Estimada',
     col=c('springgreen3','steelblue'), lty=1:2, lwd=2)
legend(x=5, y=0.8, paste(c('A', 'B')), title='Embalagem',
       col=c("springgreen3", "steelblue"), bty='n', lty=1:2, lwd=2, y.intersp=1)
```

b) Teste a igualdade das curvas de sobrevivência associadas às duas embalagens utilizando o teste de log-rank com um nível de significância de 10%.

Resposta: Denotemos por $S_A(t)$ e $S_B(t)$ as funções de sobrevivência do tempo até o aparecimento de fungos com o uso das embalagens A e B, respectivamente. Queremos testar as hipóteses

$$H_0 : S_A(t) = S_B(t), \quad \forall t \in [0, 60]$$
$$H_1 : S_A(t) \neq S_B(t), \quad \text{para algum } t \in [0, 60].$$

Para tal, usaremos o teste log-rank. Em aula, foi visto que o teste log-rank é o mais poderoso quando os dois grupos possuem riscos proporcionais, o que não será averiguado nessa solução. Apenas iremos realizar o teste, conforme solicitado no item. O fiz adicionando ao código dado no item *a*) o seguinte comando:

```
#Teste LogRank
survdiff(dados2~Grupo)
```

Obtive o seguinte resultado:

```
Call:
survdiff(formula = dados2 ~ Grupo)

         N Observed Expected (O-E)^2/E (O-E)^2/V
Grupo=A 26       21       19     0.201     0.476
Grupo=B 26       20       22     0.175     0.476

 Chisq= 0.5  on 1 degrees of freedom, p= 0.49
```

Logo, ao nível de significância de $\alpha = 0,1$, não há evidências para rejeitarmos H_0 ($p = 0,49 > 0,1$), isto é, não há diferença significativa entre as duas embalagens quando ao tempo de aparecimento dos fungos.

c) Teste a igualdade das curvas de sobrevivência associadas às duas embalagens utilizando o teste da família de Fleming-Harrington com $p = 1$ e $q = 0$ e também com $p = 0,5$ e $q = 0$, com um nível de significância de 10%.

Resposta: Para $p = 1$ e $q = 0$, o código para o teste é dado por

```
## p=1, q=0
survdiff(dados2~Grupo, rho=1)
```

e temos como resultado

```
Call:
survdiff(formula = dados2 ~ Grupo, rho = 1)

         N Observed Expected (O-E)^2/E (O-E)^2/V
Grupo=A 26     15.2     11.4     1.286      3.53
Grupo=B 26     11.1     14.9     0.984      3.53

 Chisq= 3.5  on 1 degrees of freedom, p=0.0602.
```

Nesse caso, ao nível de significância de $\alpha = 0,1$, rejeitamos H_0 ($p = 0,0602 < 0,1$), isto é, há evidência significativa de diferença entre as embalagens quanto a durabilidade do requeijão.

Para $p = 0.5$ e $q = 0$, o código para o teste é dado por

```
## p=1, q=0
survdiff(dados2~Grupo, rho=0.5)
```

e temos como resultado

```
Call:
survdiff(formula = dados2 ~ Grupo, rho = 0.5)

         N Observed Expected (O-E)^2/E (O-E)^2/V
Grupo=A 26     17.6     14.4     0.692        1.8
Grupo=B 26     14.7     17.9     0.557        1.8

 Chisq= 1.8  on 1 degrees of freedom, p=0.18.
```

Nesse caso, ao nível de significância de $\alpha = 0,1$, não rejeitamos H_0 ($p = 0,18 > 0,1$), isto é, não há evidência significativa de diferenças entre as embalagens quanto a durabilidade do requeijão.

A diferença entre os resultados decorre dos pesos dados às caudas das curvas. No primeiro teste ($p = 1$ e $q = 0$) damos um peso maior ao início da curva do que no primeiro teste ($p = 0,5$ e $q = 0$), onde de fato, pela Figura 6?, a embalagem B apresenta melhor comportamento do a embalagem A.

2. Os dados do arquivo `viarapida.xls` referem-se a um estudo feito pelo Instituto do Coração, do Hospital das Clínicas - FMUSP, para comparar dois tipos de protocolo de recuperação: via-rápida

e convencional. O protocolo via-rápida tem como filosofia uma maior integração entre as várias equipes que atendem os pacientes com a finalidade de tentar diminuir o tempo de internação, melhorando a recuperação e diminuindo custos. Considere, para as questões que seguem, apenas as seguintes variáveis:

- Delta - indicador de censura (0 - censura, 1 - saído centro cirúrgico);
- Idade - idade do paciente na internação;
- Protocolo - sendo 0 convencional e 1 via rápida;
- Raça - sendo 1 Branco, 2 negro e 3 amarelo;
- Sexo - sendo 0 feminino e 1 masculino;
- Tipac - sendo 0 congênito e 1 coronariano; e
- T-CC - tempo entre a entrada e a saída do centro cirúrgico (horas).

(*a*) Utilizando o programa de sua preferência, construa curvas de Kaplan-Meier para o T-CC com a finalidade de comparar, descritivamente, o sexo, a raça, o tipo de paciente (Tipac) e o protocolo com relação ao tempo de permanência no centro cirúrgico.

Resposta: As curvas de Kaplan-Meier para o tempo de permanência no centro cirúrgico para os fatores gênero, raça, tipo de paciente e protocolo são dadas na Figura ?? logo abaixo.

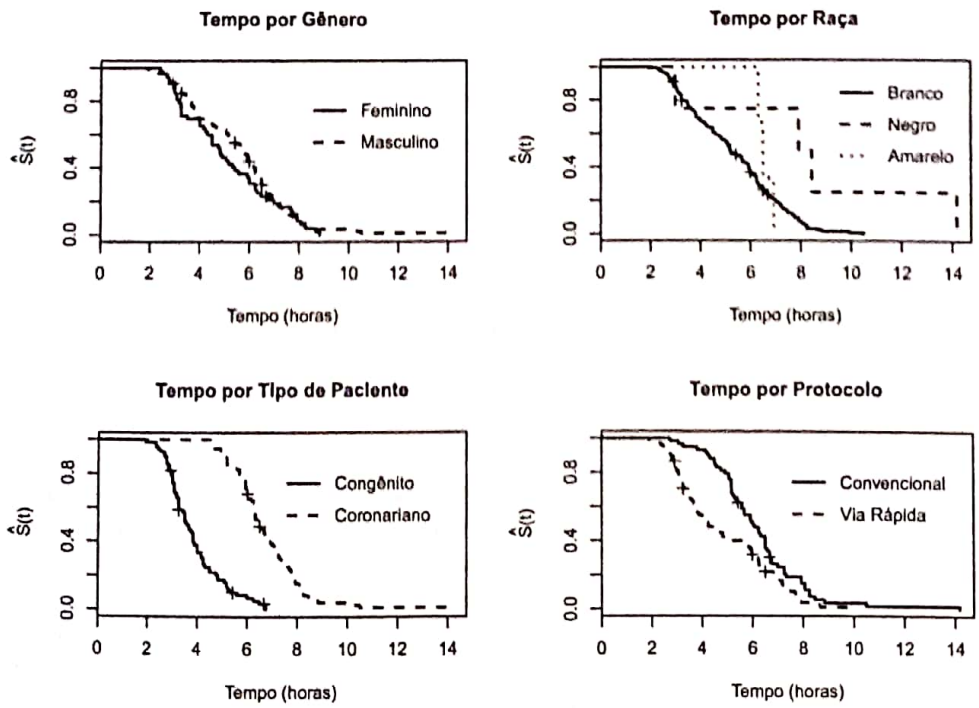


Figura 2: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier para o tempo de permanência no centro cirúrgico para os fatores gênero, raça, tipo de paciente e protocolo.*

Dada a proximidades das curvas estimadas, parece que pacientes dos sexos feminino e masculino permanecem o mesmo tempo no centro cirúrgico. Para a variável raça, a comparação

é complicada em função das poucas observações nas categorias negro e amarelo. Observamos as maiores diferenças para os fatores tipo de paciente e protocolo, de modo que pacientes coronarianos e pacientes convencionais tendem a permanecer mais tempo no centro cirúrgico.

Para construir a Figura 02, utilizei o seguinte código pelo software R:

```
rm(list=ls())
require(survival)
setwd('C:/Users/Maicon Pinheiro/Desktop/2016/2. Análise de Sobrevivência/
      Lista 3/Gabarito Lista 3')

dados = read.table('viarapida.txt', dec=",",h=T)
attach(dados)

dados2 = Surv(time=tcc, event=delta, type='right')

par(mfrow=c(2,2), mar=c(5, 4.5, 4, 2))
#Curvas de K-M
km1 = survfit(dados2~sexo)
plot(km1, xlab='Tempo (horas)', ylab=expression(paste(hat(S), '(t)')),
     col=c('firebrick2','dodgerblue4'), lty=1:2, lwd=2)
legend(x=8, y=0.9, paste(c('Feminino', 'Masculino')),
       col=c('firebrick2','dodgerblue4'), bty='n', lty=1:2, lwd=2,
       y.intersp=1.5)
title('Tempo por Gênero',cex.main=1.1)

km2 = survfit(dados2~raça)
plot(km2, xlab='Tempo (horas)', ylab=expression(paste(hat(S), '(t)')),
     col=c('firebrick2','dodgerblue4','gray20'), lty=1:3, lwd=2)
legend(x=9, y=1.0, paste(c('Branco', 'Negro', 'Amarelo')),
       col=c('firebrick2','dodgerblue4','gray20'), bty='n', lty=1:3, lwd=2,
       y.intersp=1.5)
title('Tempo por Raça',cex.main=1.1)

km3 = survfit(dados2~tipac)
plot(km3, xlab='Tempo (horas)', ylab=expression(paste(hat(S), '(t)')),
     col=c('gray27','dodgerblue4'), lty=1:2, lwd=2)
legend(x=7, y=0.9, paste(c('Congênito', 'Coronariano')),
       col=c('gray27','dodgerblue4'), bty='n', lty=1:2, lwd=2, y.intersp=1.5)
title('Tempo por Tipo de Paciente',cex.main=1.1)

km4 = survfit(dados2~protocolo)
plot(km4, xlab='Tempo (horas)', ylab=expression(paste(hat(S), '(t)')),
     col=c('firebrick2','gray23'), lty=1:2, lwd=2)
legend(x=7, y=0.9, paste(c('Convencional', 'Via Rápida')),
       col=c('firebrick2','gray23'), bty='n', lty=1:2, lwd=2, y.intersp=1.5)
title('Tempo por Protocolo',cex.main=1.1)
```

(*b*) Construa testes estatísticos não paramétricos para averiguar a significância estatística dos fatores do item anterior. Interprete os resultados.

Resposta: Denotemos por $S_{k,i}(t)$ a função de sobrevivência do tempo de permanência no

centro cirúrgico para a $i-$ésima categoria do fator k, $k \in$ {gênero, raça, tipo de paciente, protocolo}. Denotando por $|k|$ e $\tau(k)$ o número de categorias e o maior tempo observado para o fator k, respectivamente, queremos testar

$$H_0 : S_{k,1}(t) = \ldots = S_{k,|k|}(t), \quad \forall t \in [0, \tau(k)]$$
$$H_1 : S_{k,i}(t) \neq S_{k,j}(t), \quad \text{para algum } i, j \text{ em } \{1, \ldots, |k|\} \text{ e } t \in [0, \tau(k)]..$$

Para fazê-lo, usaremos o teste log-rank. Pelo software R, basta adicionar ao código do item anterior e executar os seguintes comandos:

```
survdiff(dados2~sexo)
survdiff(dados2~raça)
survdiff(dados2~tipac)
survdiff(dados2~protocolo)
```

Usando três casas decimais, temos, em suma, os seguintes resultados:

Tabela 1: *Resultados do Teste Log-Rank.*

Fatores	χ^2	p-valor
Gênero	1	0,312
Raça	6,3	0,044
Tipo de Paciente	108	< 0,001
Protocolo	9,8	0,002

Assim, assumindo um nível de significância de 5%, vê-se que não há diferenças significativas entre as duas categorias de gênero quanto ao tempo de permanência no centro cirúrgico e que esse tempo de permanência difere segundo os demais fatores, sendo a diferença ainda mais evidentes para os fatores tipo de paciente e protocolo.

(c) Considerando apenas o protocolo e métodos gráficos baseados na função de taxa de falha acumulada não paramétrica, procure averiguar qual dos modelos, exponencial, Weibull ou log-normal é mais adequado aos dados.

Resposta: Para avaliar a adequabilidade de cada um dos modelos paramétricos citados no enunciado, vamos considerar a linearização da função de risco acumulada estimada por Nelson Aalen, que denotaremos por $\hat{A}(t)$, pra cada uma das duas categorias da variável protocolo:

Modelo Exponencial: Se $T \sim$Exp(λ), então a função de risco acumulada é dada por $A(t) = \lambda t$, $t \geq 0$. Logo, caso o modelo exponencia fosse adequado, é esperado um comportamento linear no gráfico de $\hat{A}(t)$ contra t. Para estes dados, o gráfico de $\hat{A}(t)$ contra t fica dada por:

Modelo Weibull: Se $T \sim$Weibull(ρ, θ) (parametrização utilizada em sala de aula), então $\ln(A(t)) = \ln\theta + \rho \ln t$, $t \geq 0$. Logo, caso o modelo Weibull fosse adequado, é esperado um comportamento linear no gráfico de $\ln\left(\hat{A}(t)\right)$ contra $\ln t$.Para estes dados, o gráfico de $\hat{A}(t)$ contra t fica dada por:

Modelo Log-Normal: Se $T \sim$Log-Normal(μ, σ^2), temos que

$$S(t) = 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right), \quad t > 0.$$

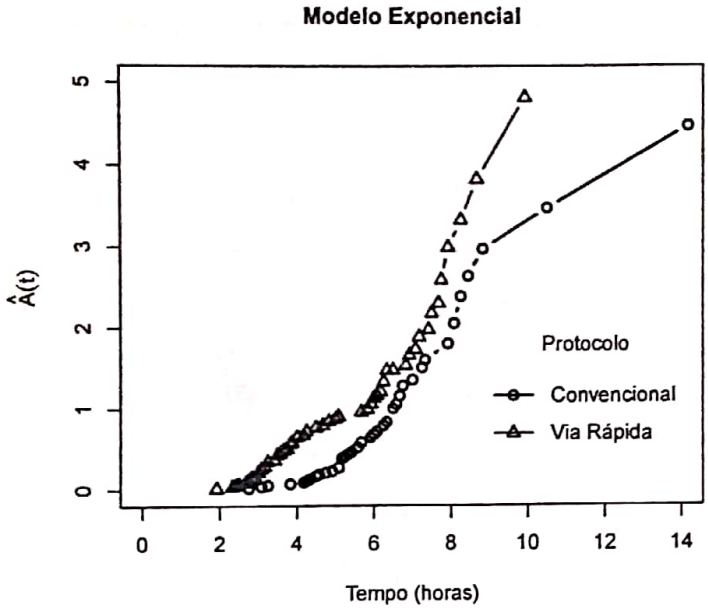


Figura 3: *Função de risco acumulada estimada por Nelson-Aalen versus tempo de permanência no centro cirúrgico por protocolo.*

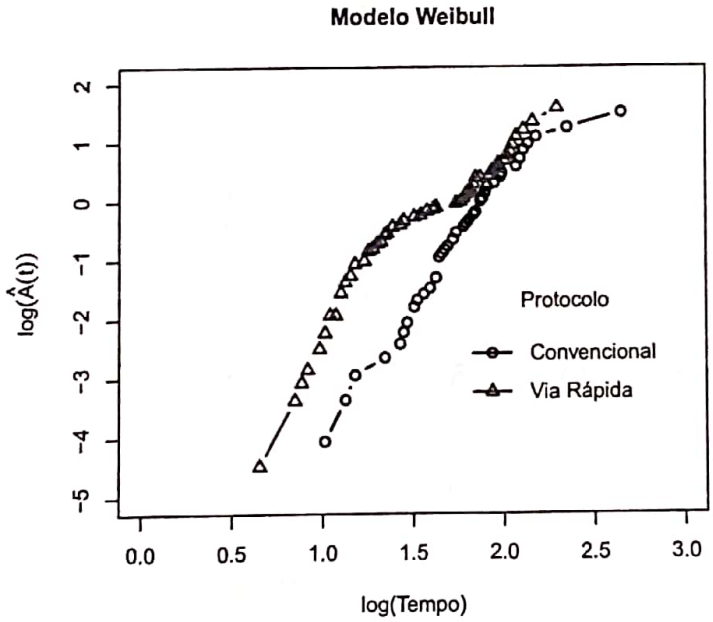


Figura 4: *Logaritmo da função de risco acumulada estimada por Nelson-Aalen versus tempo de permanência no centro cirúrgico por protocolo.*

Por outro lado, sabemos que $e^{-A(t)} = S(t)$. Das duas últimas equações, segue que

$$e^{-A(t)} = 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \Leftrightarrow 1 - e^{-A(t)} = \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \Leftrightarrow$$

$$\Phi^{-1}\left(1 - e^{-A(t)}\right) = \frac{\ln t}{\sigma} - \frac{\mu}{\sigma} \Leftrightarrow \Phi^{-1}\left(1 - e^{-A(t)}\right) = a + b \ln t,$$

em que $a = -\mu/\sigma$ e $b = 1/\sigma$. Logo, se o modelo log-normal for adequado, é esperado um comportamento linear no gráfico de $\Phi^{-1}\left(1 - e^{-\hat{A}(t)}\right)$ versus $\ln t$. Para estes dados, o gráfico de $\hat{A}(t)$ contra t fica dada por:

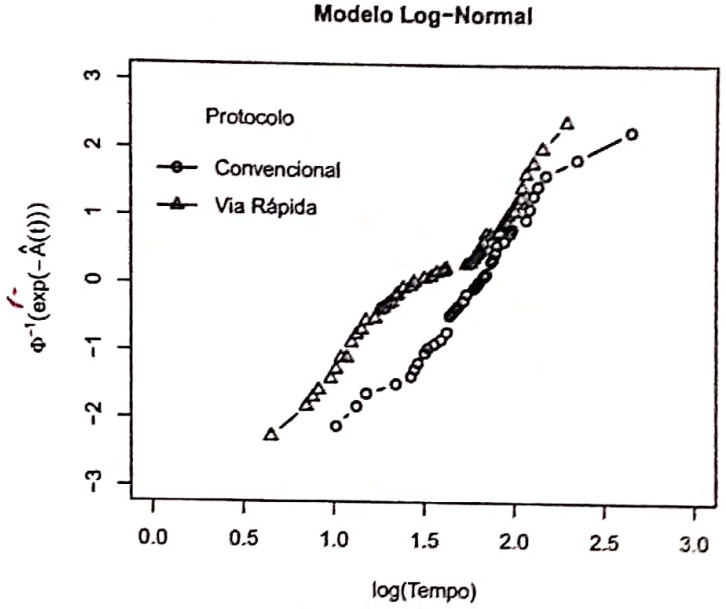


Figura 5: *Linearização da função de risco acumulada estimada por Nelson-Aalen considerando a log-normal como modelo teórico.*

No modelo exponencial, o esperado seriam duas retas passando pela origem, o que não ocorre para nenhuma das categorias (Figura ??). Assim, o modelo exponencial não é adequado para estes dados. Para os modelos Weibull e log-normal, os devidos gráficos apresentam a forma linear para cada uma das categorias, de modo que ambos devem ser considerados.

Para contruir as Figuras 03, 04 e 05, usamos o seguinte código:

```
#Estimativa Nelson-Aalen - protocolo
A1 = cumsum(km4$n.event[1:37]/km4$n.risk[1:37])
t1 = km4$time[1:37]

A2 = cumsum(km4$n.event[38:85]/km4$n.risk[38:85])
t2 = km4$time[38:85]

par(mfrow=c(1,1))
par(mar=c(5, 5, 4, 2))

## Modelo exponencial
plot(A1~t1, t='b', col='gray23', pch=1, lwd=2, xlab='Tempo (horas)',
     ylab=expression(hat(A)(t)), xlim=c(0,14), ylim=c(0,5))
lines(A2~t2, t='b', col='lawngreen', pch=2, lwd=2)
legend(x=9, y=2, paste(c('Convencional', 'Via Rápida')), title='Protocolo',
       pch=1:2, col=c('gray23','lawngreen'), bty='n', lty=1, lwd=2,
```

```
        y.intersp=1.5)
title('Modelo Exponencial',cex.main=1.1)

## Modelo Weibull
plot(log(A1)~log(t1), t='b', col='gray23', pch=1, lwd=2, xlab='log(Tempo)',
     ylab=expression(log(hat(A)(t))), xlim=c(0,3), ylim=c(-5,2))
lines(log(A2)~log(t2), t='b', col='lawngreen', pch=2, lwd=2)
legend(x=1.8, y=-1.4, paste(c('Convencional', 'Via Rápida')), title='Protocolo',
        pch=1:2, col=c('gray23','lawngreen'), bty='n', lty=1, lwd=2,
        y.intersp=1.5)
title('Modelo Weibull', cex.main=1.1)

## Modelo Log-Normal
plot(qnorm(1-exp(-A1))~log(t1), t='b', col='gray23', pch=1, lwd=2,
     xlab='log(Tempo)', ylab=expression(paste(Phi^-1, (exp(-hat(A)(t))))),
     xlim=c(0,3), ylim=c(-3,3))
lines(qnorm(1-exp(-A2))~log(t2), t='b', col='lawngreen', pch=2, lwd=2)
legend(x=0, y=2.7, paste(c('Convencional', 'Via Rápida')), title='Protocolo',
        pch=1:2, col=c('gray23','lawngreen'), bty='n', lty=1, lwd=2,
        y.intersp=1.5)
title('Modelo Log-Normal',cex.main=1.1)
```

(*d*) Compare o ajuste dos modelos exponencial, Weibull e log-normal utilizando os resíduos de Cox-Snell. Qual o modelo que você recomenda?

Resposta: Como visto em aula, os resíduos de Cox-Snell são definidos por

$$\hat{r}_i = \hat{A}_M(z_i \mid \mathbf{X}_i),$$

em que $\hat{A}_M(\cdot)$ é o estimador de $A(t \mid \mathbf{X}_i)$ dado pelo modelo M e $Z_i = \min(T_i, C_i)$, com T_i e C_i os tempos de falha e de censura, respectivamente, do $i-$ésimo indivíduo, $i = 1, \ldots, 146$, $(n = 146)$. Nesse exercício, consideraremos todas as variáveis enunciadas no começo do exercício, de modo que que $\mathbf{X}_i$ é o vetor das covariáveis idade, raça, sexo, protoloco e tipac, observadas para o $i-$ésimo indivíduo.

Como visto em aula,

Modelo Exponencial: a função de risco acumulada é $A(t \mid \mathbf{X}_i) = e^{\mathbf{X}_i'\beta}t$, de modo que $\hat{r}_i = z_i e^{\mathbf{X}_i'\hat{\beta}}$, $i = 1, \ldots, 146$.

Modelo Weibull: a função de risco acumulada é $A(t|\mathbf{X}_i) = e^{\mathbf{X}_i'\beta}t^{\rho}$, de modo que $\hat{r}_i = z_i^{\hat{\rho}} e^{\mathbf{X}_i'\hat{\beta}}$, $i = 1, \ldots, 146$.

Modelo log-normal: a função de risco acumulada é

$$A(t \mid \mathbf{X}_i) = -\ln\left(1 - \Phi\left(\frac{\ln t - \mathbf{X}_i'\gamma}{\sigma}\right)\right),$$

de modo que

$$\hat{r}_i = -\ln\left(1 - \Phi\left(\frac{\ln t - \mathbf{X}_i'\hat{\gamma}}{\hat{\sigma}}\right)\right), \quad i = 1, \ldots, 146$$

.

Nas figuras a seguir, apresentamos as curvas de sobrevivência estimadas pelo estimador de Kaplan-Meier para os resíduos de Cox-Snell em cada um dos modelos de regressão considerados. Ademais, em cada uma das figuras, segue também a curva da função de sobrevivência teórica da exponencial padrão, a qual seria a distribuição exata de $\hat{r}(t)$ quando avaliados nos verdadeiros parâmetros. Quanto mais próximas forem as duas curvas, temos mais indícios de que o modelo escolhido é adequado.

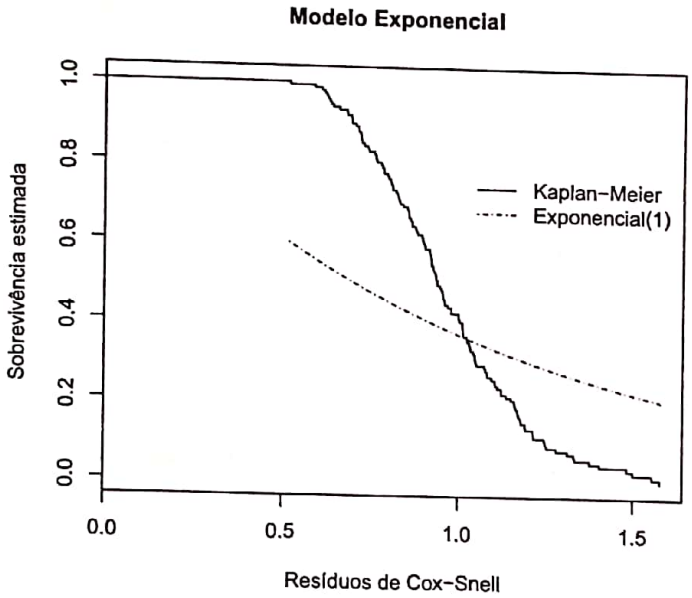


Figura 6: *Sobrevivência estimada dos resíduos de Cox-Snell e da exponencial padrão para o modelo de regressão exponencial.*

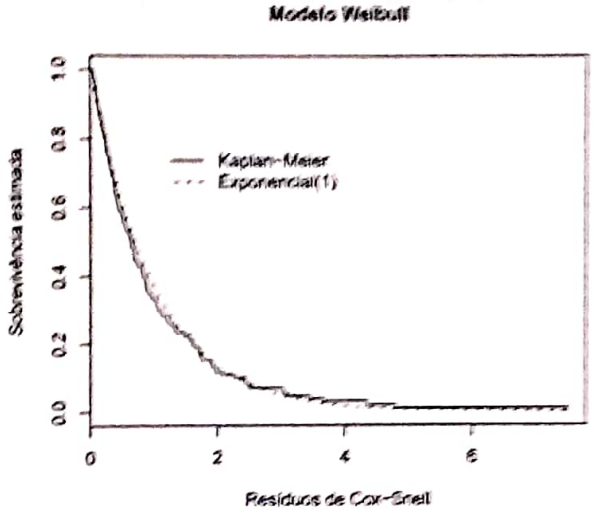


Figura 7: *Sobrevivência estimada dos resíduos de Cox-Snell e da exponencial padrão para o modelo de regressão Weibull.*

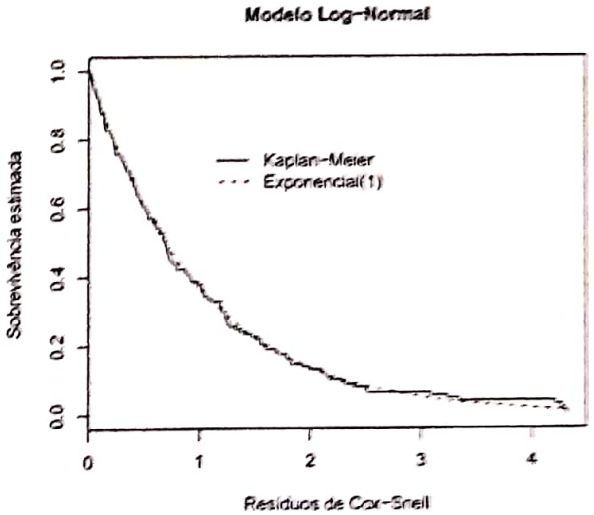


Figura 8: *Sobrevivência estimada dos resíduos de Cox-Snell e da exponencial padrão para o modelo de regressão log-normal.*

Pela análise gráfica, nota-se que o modelo log-normal é o que melhor se ajusta aos dados, embora o ajuste do modelo Weibull também seja razoável.

Para construir as Figuras 06, 07 e 08, considerou-se o seguinte código:

```
data = dados[-69,]
tcc = data$tcc
delta = data$delta

raca.br = ifelse(data$raça==1, 1, 0)
raca.ne = ifelse(data$raça==2, 1, 0)

data2 = Surv(time=data$tcc, event=data$delta)
data2

## Modelo Exponencial
aj1<-survreg(data2~data$protocolo + data$sexo + data$idade + data$tipac
             + raca.br + raca.ne, dist='exponential')

s1 = 1/aj1$scale
xb1 = aj1$linear.predictors

e1 = tcc*exp(-xb1) #Resíduos de Cox-Snell para o modelo Exponencial

ekm1 = survfit(Surv(e1,delta)~1)
t = ekm1$time
st = ekm1$surv #Função de Sobrevivência estimada por KM
sexp = exp(-t) #Função de sobrevivência no modelo Exponencial

plot(ekm1,conf.int=F,mark.time=F, xlab="Resíduos de Cox-Snell",
     ylab="Sobrevivência estimada")
lines(t,sexp,lty=4,col='steelblue4')
legend(1.0,0.8,lty=c(1,4),c("Kaplan-Meier","Exponencial(1)"),bty="n",
       col=c(1,'steelblue4'))
title('Modelo Exponencial',cex.main=1.1)


## Modelo Weibull
aj2 = survreg(data2~data$protocolo + data$sexo + data$idade
              + data$tipac + raca.br + raca.ne, dist='weibull')

xb2 = aj2$linear.predictors
s2 = aj2$scale
e2 = (tcc*exp(-xb2))^(1/s2) #Resíduos de Cox-Snell para o modelo Weibull

ekm2 = survfit(Surv(e2,delta)~1)
t = ekm2$time
st = ekm2$surv
sexp = exp(-t)

plot(ekm2,conf.int=F,mark.time=F, xlab="Resíduos de Cox-Snell",
     ylab="Sobrevivência estimada")
```

cuidado com raça no modelo!

```
lines(t,sexp,lty=4,col='steelblue4')
legend(1.0,0.8,lty=c(1,4),c("Kaplan-Meier","Exponencial(1)"),bty="n",
       col=c(1,'steelblue4'))
title('Modelo Weibull',cex.main=1.1)


## Modelo Log-Normal
aj3 = survreg(data2~data$protocolo+data$sexo+data$idade+data$tipac
              +raca.br+raca.ne, dist='lognormal')

xb3 = aj3$linear.predictors
s3 = aj3$scale
e3 = -log(1-pnorm((log(data$tcc)-xb3)/s3)) #Resíduos Cox-Snell Log-Normal


ekm3 = survfit(Surv(e3,data$delta)~1)
t = ekm3$time
st = ekm3$surv #Kaplan Meier dos residuos Cox-snell
sexp = exp(-t) #Sobrevivência da exponencial(1)

plot(ekm3,conf.int=F,mark.time=F, xlab="Resíduos de Cox-Snell",
     ylab="Sobrevivência estimada")
lines(sexp~t,lty=4,col='steelblue4')
legend(1.0,0.8,lty=c(1,4),c("Kaplan-Meier","Exponencial(1)"), bty="n",
       col=c(1,'steelblue4'))
title('Modelo Log-Normal',cex.main=1.1)
```

(*e*) Com base no modelo obtido no item anterior, faça o teste paramétrico para comparar os protocolos com relação ao tempo de permanência no centro cirúrgico.

Resposta: No item anterior, vimos que o modelo log-normal é, dentre os considerados, o mais adequado. Portanto, para testar o efeito do fator protocolo na variável TCC, vamos considerar tal modelo e realizar o teste da razão de verossimilhanças (TRV) utilizando a aproximação da distribuição da estatítica do teste pela distribuição qui-quadrado com um grau de liberdade. O p-valor obtido foi menor que $0,001$ (Estatística RV=14,665), ou seja, a um nível de 5% de significância, rejeitamos a hipótese nula de que o protocolo não afeta o tempo de permanência no centro cirúrgico. Para a obtenção desses resultados, consideramos o seguinte código:

```
aj4 = survreg(data2~data$sexo+data$idade+data$tipac+raca.br+raca.ne,
              dist='lognormal')

LL1 = summary(aj3)$loglik[2] #Modelo Completo
LL2 = summary(aj4)$loglik[2] #Modelo Sem o fator protocolo

RV = 2*(LL1-LL2)      #menos duas vezes log da RV
pv = 1-pchisq(RV, 1) #P-valor TRV  - Qui Qudrado 1 gl
```

3. Os dados no arquivo `HOD NHL.csv` são referentes a 43 pacientes com linfoma de Hodgking ou linfoma não Hodgking, submetidos a transplante de medula óssea. Alguns pacientes receberam

transplante de doador aparentado compatível (transplante alogênico) e outros receberam transplante autólogo (ou seja, a própria medula óssea do paciente é coletada e posteriormente infundida). Nos dados também há informação sobre o escore de Karnofsky, que é uma medida de perfomance que classifica os pacientes segundo o bem-estar dos pacientes. O objetivo principal do estudo é a comparação dos tipos de transplante, considerando-se o tempo (dias) livre de doença (i.e., tempo antes de ocorrência da recorrência ou óbito).

(*a*) Construa curvas de Kaplan-Meier para o tempo de sobrevivência dos pacientes com linfoma para cada um dos grupos de tratamento. Categorize a variável escore de Karnofsky, criando uma variável com duas categorias: escore < 80 ou escore ≥ 80. Construa gráficos de Kaplan-Meier para essa nova variável categorizada. Comente.

Resposta: Por tratamento: Pela categorização da variável escore de Karnofsky:

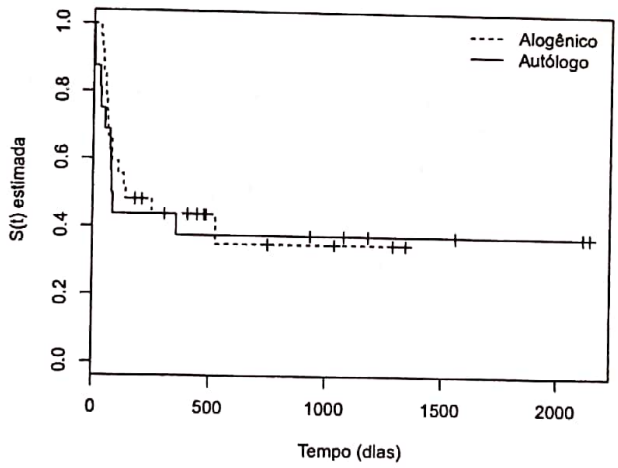


Figura 9: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier para o tempo de sobrevivência dos pacientes com linfoma para cada um dos grupos de tratamento.*

Pela análise gráfica, fica evidente que o tempo é maior para pacientes com escore de Karnofsky acima de 80 e parece haver um maior tempo de sobrevida para pacientes que realizaram o transplante autólogo.

O código utilizado é dado a seguir:

```
rm(list=ls())
require(survival)
setwd('C:/Users/Maicon Pinheiro/Desktop/2016/2. Análise de Sobrevivência/
      Lista 3/Gabarito Lista 3')

dados = read.table('HOD_NHL.txt',h=T)
dados\$transplante<-ifelse(dados\$transplante==2,0,1) # 1 alogênico e 0 autólogo
attach(dados)

as.factor(transplante)
```

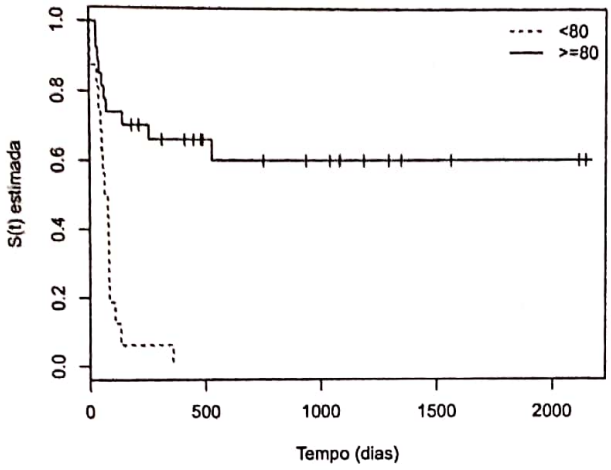


Figura 10: *Funções de sobrevivência estimadas pelo estimador de Kaplan-Meier para o tempo de sobrevivência dos pacientes com linfoma pela categorização da variável escore de Karnofsky.*

```
as.factor(d)
table(transplante,d)


require(survival)
KM <- survfit(Surv(tempo,d) ~ transplante)
plot(KM,lty=c(2,1),xlab="Tempo (dias)",ylab="S(t) estimada")
legend('topright',lty=c(2,1),bty='n',c("Alogênico","Autólogo"))

k<-ifelse(dados\$Karnofsky<80,0,1) # 1 << 80 e 0 >= 80
as.factor(k)

KM2 <- survfit(Surv(tempo,d) ~ k)
plot(KM2,lty=c(2,1),xlab="Tempo (dias)",ylab="S(t) estimada")
legend('topright',lty=c(2,1),bty='n',c("<80",">=80"))
```

(*b*) Ajuste um modelo de regressão Weibull apenas com a variável tratamento. Apresente as estimativas de máxima verossimilhança dos parâmetros considerando a representação do modelo Weibull como um modelo de locação-escala e também como um modelo de riscos proporcionais.

Resposta:

?

```
#Ajuste modelo
mwei <- survreg(Surv(tempo,d) ~ transplante,dist='weibull')
#TRV para \gamma_1=1
mwei2 <- survreg(Surv(tempo,d) ~ 1,dist='weibull')
```

0

4. Deseja-se comparar duas populações de tempos de vida. Uma amostra de tamanho n ($r \leq n$ falhas) foi obtida da população 1, que tem distribuição exponencial com média α. Uma amostra

de tamanho m ($s \leq m$ falhas) foi obtida da população 2, que tem distribuição exponencial com média $\alpha + \Delta$).

(a) Estabeleça as hipóteses que deseja testar.

Resposta: Denotemos por $S_i(\cdot)$ a função de sobrevivência associada à população i, $i = 1, 2$. Queremos testar as hipóteses:

$$\begin{aligned} H_0 &: S_1(t) = S_2(t), \quad \forall t \in [0, \tau] \\ H_1 &: S_1(t) \neq S_2(t), \quad \text{para algum } t \in [0, \tau], \end{aligned}$$

em que τ é o maior valor entre as $m + n$ observações. Nesse caso, onde as funções de sobrevivência $S_1(\cdot)$ e $S_2(\cdot)$ ficam determinadas pelos parâmetros α e $\alpha + \Delta$, respectivamente, uma forma mais econômica de apresentar as hipóteses dadas acima é:

$$\begin{aligned} H_0 &: \Delta = 0, \\ H_1 &: \Delta \neq 0. \end{aligned}$$

Era para considerar a nova direto aqui o enunciado tá ficou bom.

(b) Escreva a função de verossimilhança para $\boldsymbol{\theta} = (\alpha, \Delta)^T$. Calcule o vetor escore e a matriz de informação observada.

Resposta: Sob a suposição de independência entre os $m + n$ indivíduos oriundos da união das duas amostras, rotulemos a amostra combinada de modo que as primeiras n observações sejam da população 1 e as demais m da população 2, as primeiras r e s observações em cada bloco representam tempos de falha, seguidas por sua vez pelas respectivas u e v censuras, $r + u = n$ e $s + v = m$. Ademais, façamos a rotulação de modo que dentre as u censuras apareçam as u_1 que são à direita seguidas da u_2 à esquerda seguidas das u_3 intervalares. Do mesmo modo, para as outras v censuras. Nessas condições, sendo

D : conjunto de índices de indivíduos cujo evento foi observado;
R : conjunto de índices de indivíduos censurados à direita;
L : conjunto de índices de indivíduos censurados à esquerda;
I : conjunto de índices de indivíduos com censura intervalar,

a função de verossimilhança pode ser escrita na forma

$$\begin{aligned} L(\boldsymbol{\theta}) \propto \prod_{j \in D} f(t_j) \prod_{j \in R} S(c_j) \prod_{j \in L} (1 - S(e_j)) \prod_{j \in I} (S(a_j) - S(b_j)) = \left[\prod_{j=1}^{r} \lambda e^{-\lambda t_j} \prod_{j=n+1}^{n+s} (\lambda + \Delta) e^{-(\lambda+\Delta) t_j} \right] \times \\ \times \left[\prod_{j=r+1}^{r+u_1} e^{-\lambda c_j} \prod_{j=n+s+1}^{n+s+v_1} e^{-(\lambda+\Delta) c_j} \right] \times \left[\prod_{j=r+u_1+1}^{r+u_1+u_2} (1 - e^{-\lambda e_j}) \prod_{j=n+s+v_1+1}^{n+s+v_1+v_2} (1 - e^{-(\lambda+\Delta) e_j}) \right] \times \\ \times \left[\prod_{j=r+u_1+u_2+1}^{n} (e^{-\lambda a_j} - e^{-\lambda b_j}) \prod_{j=n+s+v_1+v_2+1}^{n+m} (e^{-(\lambda+\Delta) a_j} - e^{-(\lambda+\Delta) b_j}) \right]. \end{aligned}$$

Assim, segue que

$$l(\boldsymbol{\theta}) = a + \sum_{j=1}^{r} \ln\lambda - \lambda \sum_{j=1}^{r} t_j + \sum_{j=n+1}^{n+s} \ln(\lambda+\Delta) - (\lambda+\Delta) \sum_{j=n+1}^{n+s} t_j - \lambda \sum_{j=r+1}^{r+u_1} c_j -$$
$$- (\lambda+\Delta) \sum_{j=n+s+1}^{n+s+v_1} c_j + \sum_{j=r+u_1+1}^{r+u_1+u_2} \ln(1-e^{-\lambda e_j}) + \sum_{j=n+s+v_1+1}^{n+s+v_1+v_2} \ln(1-e^{-(\lambda+\Delta)e_j}) +$$
$$+ \sum_{j=r+u_1+u_2+1}^{n} \ln(e^{-\lambda a_j} - e^{-\lambda b_j}) + \sum_{j=n+s+v_1+v_2+1}^{n+m} \ln(e^{-(\lambda+\Delta)a_j} - e^{-(\lambda+\Delta)b_j}),$$

isto é,

$$l(\boldsymbol{\theta}) = a + r\ln\lambda - \lambda \sum_{j=1}^{r} t_j + s\ln(\lambda+\Delta) - (\lambda+\Delta) \sum_{j=n+1}^{n+s} t_j - \lambda \sum_{j=r+1}^{r+u_1} c_j -$$
$$- (\lambda+\Delta) \sum_{j=n+s+1}^{n+s+v_1} c_j + \sum_{j=r+u_1+1}^{r+u_1+u_2} \ln(1-e^{-\lambda e_j}) + \sum_{j=n+s+v_1+1}^{n+s+v_1+v_2} \ln(1-e^{-(\lambda+\Delta)e_j}) +$$
$$+ \sum_{j=r+u_1+u_2+1}^{n} \ln(e^{-\lambda a_j} - e^{-\lambda b_j}) + \sum_{j=n+s+v_1+v_2+1}^{n+m} \ln(e^{-(\lambda+\Delta)a_j} - e^{-(\lambda+\Delta)b_j}),$$

em que a é uma constante. Nessas condições, temos

$$\frac{\partial}{\partial\lambda} l(\boldsymbol{\theta}) = \frac{r}{\lambda} - \sum_{j=1}^{r} t_j + \frac{s}{\lambda+\Delta} - \sum_{j=n+1}^{n+s} t_j - \sum_{j=r+1}^{r+u_1} c_j + \sum_{j=r+u_1+1}^{r+u_1+u_2} \frac{e^{-\lambda e_j}}{1-e^{-\lambda e_j}} +$$
$$+ \sum_{j=n+s+v_1+1}^{n+s+v_1+v_2} \frac{e^{-(\lambda+\Delta)e_j}}{1-e^{-(\lambda+\Delta)e_j}} - \sum_{j=r+u_1+u_2+1}^{n} \frac{e^{-\lambda a_j} - e^{-\lambda b_j}}{e^{-\lambda a_j} - e^{-\lambda b_j}}$$
$$- \sum_{j=n+s+v_1+v_2+1}^{n+m} \frac{e^{-(\lambda+\Delta)a_j} - e^{-(\lambda+\Delta)b_j}}{e^{-(\lambda+\Delta)a_j} - e^{-(\lambda+\Delta)b_j}}$$
$$= \frac{r}{\lambda} - \sum_{j=1}^{r} t_j + \frac{s}{\lambda+\Delta} - \sum_{j=n+1}^{n+s} t_j - \sum_{j=r+1}^{r+u_1} c_j +$$
$$+ \sum_{j=r+u_1+1}^{r+u_1+u_2} \frac{e^{-\lambda e_j}}{1-e^{-\lambda e_j}} + \sum_{j=n+s+v_1+1}^{n+s+v_1+v_2} \frac{e^{-(\lambda+\Delta)e_j}}{1-e^{-(\lambda+\Delta)e_j}} - u_3 - v_3,$$

e

$$\frac{\partial}{\partial\Delta} l(\boldsymbol{\theta}) = \frac{s}{\lambda+\Delta} - \sum_{j=n+1}^{n+s} t_j + \sum_{j=n+s+v_1+1}^{n+s+v_1+v_2} \frac{e^{-(\lambda+\Delta)e_j}}{1-e^{-(\lambda+\Delta)e_j}} - v_3,$$

o que determina o vetor escore $\mathbf{U}(\boldsymbol{\theta})$, uma vez que

$$\mathbf{U}(\boldsymbol{\theta}) = \left(\frac{\partial}{\partial\lambda} l(\boldsymbol{\theta}), \frac{\partial}{\partial\Delta} l(\boldsymbol{\theta}) \right).$$

Ademais, temos que

$$\frac{\partial^2}{\partial\Delta^2} l(\boldsymbol{\theta}) = \frac{\partial^2}{\partial\Delta\partial\lambda} l(\boldsymbol{\theta}) = \frac{\partial^2}{\partial\lambda\partial\Delta} l(\boldsymbol{\theta}) = -\frac{s}{(\lambda+\Delta)^2} + \sum_{j=n+s+v_1+1}^{n+s+v_1+v_2} \frac{2e^{-2(\lambda+\Delta)e_j} - e^{-(\lambda+\Delta)e_j}}{(1-e^{-(\lambda+\Delta)e_j})^2}$$

e

$$\frac{\partial^2}{\partial\lambda^2}l(\boldsymbol{\theta}) = \frac{\partial^2}{\partial\Delta^2}l(\boldsymbol{\theta}) - \frac{r}{\lambda^2} + \sum_{j=r+u_1+1}^{r+u_1+u_2} \frac{2e^{-2\lambda e_j} - e^{-\lambda e_j}}{(1-e^{-\lambda e_j})^2},$$

o que determina a matriz de informação, uma vez que a mesma é dada por $-\frac{\partial^2}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'}l(\boldsymbol{\theta})$.

(c) Obtenha uma expressão para o teste da razão de verossimilhanças para as hipóteses apresentadas em (a).

Resposta: Por definição, a razão de verossimilhanças é dada por

$$RV = \frac{\sup_{\Delta=0} L(\boldsymbol{\theta})}{\sup_{\Delta\neq 0} L(\boldsymbol{\theta})} = \sup_{\Delta\neq 0}\left[\prod_{j=n+1}^{n+s} \frac{\lambda e^{\Delta t_j}}{(\lambda+\Delta)} \times \prod_{j=n+s+1}^{n+s+v_1} e^{\Delta c_j} \times \right.$$
$$\left. \times \prod_{j=n+s+v_1+1}^{n+s+v_1+v_2} \frac{(1-e^{-\lambda e_j})}{(1-e^{-(\lambda+\Delta)e_j})} \times \prod_{j=n+s+v_1+v_2+1}^{n+m} \frac{e^{-\lambda a_j} - e^{-\lambda b_j}}{e^{-(\lambda+\Delta)a_j} - e^{-(\lambda+\Delta)b_j}}\right].$$

No caso em que $r = n$ e $s = m$, temos um caso bem conhecido: teste da igualdade dos parâmetros $\lambda_1 = \lambda$ e $\lambda_2 = \lambda + \Delta$ de duas exponenciais, a partir das respectivas amostras de tamanhos $n_1 = n$ e $n_2 = m$. Nesse caso, o sup no denominador ocorre nos estimadores de máxima verossimilhança de λ_1 e λ_2, que são fáceis de serem encontrados, de modo que a razão de verossimilhança fica dada por

$$\frac{(n+m)^{n+m}}{n^n m^m} \frac{(\sum_{j=1}^n t_j)^n (\sum_{j=n+1}^{n+m} t_j)^m}{(\sum_{j=1}^n t_j + \sum_{j=n+1}^{n+m} t_j)^{n+m}}.$$

5. Mostre que, se X é uma variável dicotômica, então a função escore obtida com base na verossimilhança parcial avaliada em $\beta = 0$ (ou seja, $U(0)$), usada para testar a hipótese nula $H_0 : \beta = 0$, se reduz ao numerador da estatística do teste de log-rank.

Resposta: Como visto em aula, a função de verossimilhança parcial para o modelo de Cox é dada por

$$L(\beta) = \prod_{j=1}^{l} \frac{e^{\mathbf{X}'_{(j)}\beta}}{\sum_{k\in\mathbb{R}(t_{(j)})} e^{\mathbf{X}'_{(k)}\beta}},$$

donde decorre que

$$l(\beta) = \sum_{j=1}^{l} \ln\left(\frac{e^{\mathbf{X}'_{(j)}\beta}}{\sum_{k\in\mathbb{R}(t_{(j)})} e^{\mathbf{X}'_{(k)}\beta}}\right) = \sum_{j=1}^{l}\left[\mathbf{X}'_{(j)}\beta - \ln \sum_{k\in\mathbb{R}(t_{(j)})} e^{\mathbf{X}'_{(k)}\beta}\right].$$

Assim, temos

$$U(\beta) = \frac{\partial}{\partial\beta}l(\beta) = \sum_{j=1}^{l}\mathbf{X}'_{(j)} - \sum_{j=1}^{l}\left(\frac{\sum_{k\in\mathbb{R}(t_{(j)})}\mathbf{X}'_{(k)}e^{\mathbf{X}'_{(k)}\beta}}{\sum_{k\in\mathbb{R}(t_{(j)})} e^{\mathbf{X}'_{(k)}\beta}}\right),$$

de modo que

$$
\begin{aligned}
U(0) &= \sum_{j=1}^{l} \mathbf{X}'_{(j)} - \sum_{j=1}^{l} \left(\frac{\sum_{k \in \mathbb{R}(t_{(j)})} \mathbf{X}'_{(k)}}{\sum_{k \in \mathbb{R}(t_{(j)})} 1} \right) \\
&= \sum_{j=1}^{l} \mathbf{X}'_{(j)} - \sum_{j=1}^{l} \frac{\sum_{k \in \mathbb{R}(t_{(j)})} \mathbf{X}'_{(k)}}{|\mathbb{R}(t_{(j)})|} \\
&= \sum_{j=1}^{l} \left[\mathbf{X}'_{(j)} - \frac{\sum_{k \in \mathbb{R}(t_{(j)})} \mathbf{X}'_{(k)}}{|\mathbb{R}(t_{(j)})|} \right],
\end{aligned}
$$

que se reduz a uma soma de diferenças entre um valor observado e um valor esperado. A primeira parcela de cada diferença dentro da soma vale 1 ou 0 (variável dicotômica) e isso representa o observado e a segunda toma um valor entre 0 e 1 indicando a proporção de indivíduos com a covariável 1 ou 0, respectivamente, dentre os que ainda estão em risco e essa faz o papel de valor esperado.

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Marcon Aparecido Pinheiro Professora: Gisela Tunes da Silva
Fábio Hideto Oki

1. Leia o artigo "Understanding Cox' regression model: a martingal approach", publicado no Joural of the American Statistical Association, no ano de 1984, volume 79, páginas 441-447. Prepare um resumo de aproximadamente duas ou três páginas do artigo, enfatizando e descrevendo com mais detalhes os seguintes pontos:

- A demonstração de que o processo escore, devidamente padronizado e calculado em termos do verdadeiro valor do vetor de parâmetros, é um martingal de média zero
- O desenvolvimento utilizado para mostrar que o escore (calculado a partir do processo escore anterior) converge em distribuição para uma distribuição normal.
- A técnica utilizada para encontrar a distribuição assintótica do estimador de máxima verossimilhança parcial com base nos resultados para o escore.

Em suma, o artigo discute de modo claro e com forte apelo intuitivo toda a formulação do modelo de Cox, tanto na especificação mais elementar quanto a feita via processos de contagem e processos previsíveis. Além disso, apresenta de modo elucidativo toda a teoria sobre martingais necessária ao estudo da convergência do estimador do parâmetro β no modelo de Cox. Conforme solicitado, iremos resumir o artigo dando enfoque aos três pontos citados no enunciado. Para tornar a tarefa mais simples e organizada, seguiremos com a estrutura do artigo, isto é, apresentaremos o conteúdo do resumo na mesma ordem vista no artigo. Em função do pouco espaço, faremos nas devidas formulações e passaremos a maior parte do tempo discorrendo sobre a quinta seção do artigo. Vamos à isso

O modelo de Regressão de Cox (1972) é de grande sucesso entre os estatísticos aplicados. Sua elegância combina as vantagens das abordagens paramétrica e não paramétrica à referência estatística, e se adapta muito bem à dados de sobrevida. Ao incluir variações no tempo, covariáveis aleatórias, ele se torna uma ferramenta altamente flexível para a construção de modelos. Vamos à primeira especificação do modelo de regressão de Cox.

Sejam T_i, $i=1, \dots, n$ v.a's independentes, contínuas e não negativas representando o tempo de vida de n indivíduos. Cada uma das T_i's pode ser observada apenas no intervalo $[0, c_i]$, onde $(c_i)_{i=1}^{n}$ são os respectivos tempos de censura. O modelo de regressão de Cox consiste basicamente em supor que a taxa de falha p/ o i-ésimo indivíduo tem uma forma especial dada por

$$\lambda_i(t) = \lambda_0(t) \exp(\beta_0' z_i(t)), \quad i=1, \dots, n, \qquad (1)$$

onde β_0' é a transposta de um vetor coluna β_0 de p coeficientes desconhecidos, z_i é um vetor coluna de p possíveis covariáveis variáveis no tempo e λ_0 é uma taxa de falha basal fixada e desconhecida para o indivíduo com $z \equiv 0$ (a parte não paramétrica do modelo). As observações para o i-ésimo indivíduo consistem de $\min(T_i, c_i)$, $\delta_i = I(T_i \leq c_i)$ e $z_i(t)$, $t \in [0, \min(T_i, c_i)]$. O interesse é estimar ou testar hipóteses acerca de β_0. Para tal, Cox propôs que a inferência estatística acerca de β_0 poderia ser derivada da seguinte função de verossimilhança parcial:

$$(2) \qquad \mathcal{L}(\beta) = \prod_{i: T_i \leq c_i} \left(\frac{e^{\beta' z_i(T_i)}}{\sum_{j \in R(T_i)} e^{\beta' z_j(T_i)}} \right), \quad R(t) = \{i : T_i \geq t \text{ e } c_i \geq t\}$$

conj. de ind. em risco em t

Dado $R(T_i)$ e T_i, a probabilidade de que a falha ocorra para o indivíduo i.

Assim como no artigo, ressaltamos que a expressão acima pode ser obtida ainda em generalizações do modelo permitir que os dados sejam observados de forma mais ou menos pois de censura. No entanto, em casos mais complexos, não é claro que as propriedades permaneçam as mesmas.

Agora vamos à formulação do mesmo modelo por meio de processos de contagem e processos de contagem.

Alunos: Fábio H. Oki
Marcon A. Pinheiro

Nesse caso, define-se o processo de contagem

$$N_i(t) = \mathbb{1}\{T_i \leq t, T_i \leq C_i\}. \qquad (3)$$

e modela-se seu processo de intensidade Λ_i da seguinte forma.

$$\Lambda_i(t)dt = Y_i(t)\,\lambda_0(t)\,e^{\beta_0' Z_i(t)}\,dt, \qquad (4)$$

em que $Y_i(t) = \mathbb{1}\{T_i \geq t, C_i \geq t\}$. O processo de contagem multivariado $\tilde{N} = \{N_i(t); 0 \leq t < +\infty, i = 1, \ldots, n\}$ pode ser visto como a contagem de ocorrências de falhas dos n indivíduos em estudo conforme o tempo decorre. Uma extensão óbvia é tomar $\tilde{N}$ cujo processo de intensidade $\tilde{\Lambda} = \{\Lambda_i(t): 0 \leq t < +\infty, i = 1, \ldots, n\}$ satisfaz

$$\Lambda_i(t)dt = Y_i(t)\,\lambda_0(t)\,e^{\beta_0' \underbrace{Z_i(t)}_{\text{covariável aleatória}}}\,dt \qquad (5)$$

Aqui, não se exige mais que N_i realize apenas um salto nem que Y_i seja da forma dada acima. Tudo que se requer é que N_i, Y_i e Z_i sejam processos que possam ser observados e que Y_i e Z_i sejam processos previsíveis. Nessa nova notação, $L(\beta_0) = \lim_{u \to \infty} L(\beta, u)$, com

$$L(\beta, u) = \prod_{\substack{t \geq 0 \\ t \leq u}} \prod_{i=1}^{n} \left(\frac{Y_i(t)\,e^{\beta' Z_i(t)}}{\sum_{j=1}^{n} Y_j(t)\,e^{\beta' Z_j(t)}} \right)^{dN_i(t)}$$

(produto finito sobre todos os i e t's em que N_i salta, caso contrário, $dN_i(t) = 0$)

(6)

O interesse continua na estimação de β. Agora, é preciso do uso da teoria de martingais. Na verdade, alguns resultados dessa teoria: a integral estocástica de um proc. prev. com respeito à um martingal é um martingal; e teoremas centrais do limite, o que fornece condições sob as quais um martingal M será aproximadamente distribuído como uma normal e terá incrementos independentes, de modo que o processo se assemelha ao movimento Browniano reescalado. No contexto do modelo de reg. de Cox, verifica-se que

$$\Lambda_i(t)dt = E[dN_i(t) \mid \mathcal{F}_{t-}],$$ $\mathcal{F}_{t-}$ σ-álgebra indicando a hist. do proc. antes de t. (7)

e que $dM_i(t) = dN_i(t) - \Lambda_i(t)dt$ tem $E[dM_i(t) \mid \mathcal{F}_{t-}] = 0$, donde segue que

$$M_i(t) = N_i(t) - \int_0^t \Lambda_i(s)\,ds \qquad (8)$$

é um martingal. Também é preciso definir o processo de variação previsível $\langle M \rangle$ de M que é t.q.

$$d\langle M \rangle(t) = E[(dM(t))^2 \mid \mathcal{F}_{t-}] = Var[dM(t) \mid \mathcal{F}_{t-}], \qquad (9)$$

previsível e não decrescente. Pode ser visto como soma de variâncias condicionais.

Pode ser verificado que

$$\langle M_i \rangle(t) = \int_0^t \Lambda_i(s)\,ds \quad \text{e que} \quad \langle M_i, M_j \rangle = E[dM_i(t)\,dM_j(t) \mid \mathcal{F}_{t-}] = 0, i \neq j, \forall t \qquad (10)$$

Ademais, se H é processo previsível, M é um martingal,

(11) $$M'(t) = \int_0^t H(s)\,dM(s) \text{ é um martingal}$$

Fábio H. OK!

Por fim, acerca da teoria de martingais, menciona-se teoremas centrais do limite, que são de extrema importância para o estudo do comportamento assintótico de $\hat{\beta}$. Aqui, vamos seguir com mais detalhes. Um movimento Browniano reescalado é um processo $W = \{W(t) : t \geq 0\}$ com as seguintes propriedades. As realizações $W(\cdot)$ são funções contínuas e são nulas em $t=0$. Para quaisquer $t_1, \ldots, t_n$ $W(t_1), \ldots, W(t_n)$ visto como vetor aleatório tem dist. normal multivariada de médias nulas e tem-se que $W(t) - W(s) \perp W(s)$ $\forall s < t$. Pela indep. dos incrementos, a variância condicional de $dW(t)$ dado o passado de W em $[0,t)$ não depende do passado. Ademais, a esperança condicional do mesmo é zero. Desse modo, W é um martingal contínuo com processo de variação previsível $\langle W \rangle$ igual a uma função determinística A.

Tais propriedades caracterizam a distribuição de W (Gaussiana). Assim, não seria surpresa se uma seq. de martingais $M^{(n)}$, $n \geq 1$ t.q. a) os saltos de $M^{(n)}$ ficam menores quando $n \to +\infty$ ($M^{(n)}$ se torna aprox. contínuo); b) $\langle M^{(n)} \rangle \xrightarrow{P} A(t)$, qdo $n \to +\infty$, A uma função fixa, então $M^{(n)} \xrightarrow{D} W$, qdo $n \to +\infty$.

Dito isso, pode-se seguir às propriedades assintóticas de $\hat{\beta}$. Nesse caso, a ideia é usar o Teorema Central do Limite para martingais. De modo mais específico, considera-se o vetor escore $D \log L(\beta_0)$ como soma (ou integral) sobre os instantes t.

Nesse caso, segue que

Basta aplicar log em $L(\beta_0, u)$, derivá-la e multiplicá-la por $n^{-1/2}$.

$$\frac{1}{n^{1/2}} D \log L(\beta_0, u) = \frac{1}{n^{1/2}} \sum_{i=1}^{n} \sum_{t \leq u} \left(Z_i(t) - \frac{\sum_{j=1}^{n} Y_j(t) Z_j(t) e^{\beta_0' Z_j(t)}}{\sum_{j=1}^{n} Y_j(t) e^{\beta_0' Z_j(t)}} \right) dN_i(t)$$

$$= \sum_{i=1}^{n} \int_{t=0}^{u} n^{-1/2} \left(Z_i(t) - E_0(t) \right) dN_i(t), \quad \text{com } E_0(t) = \frac{1/n \sum_{j=1}^{n} Y_j(t) Z_j(t) e^{\beta_0' Z_j(t)}}{1/n \sum_{j=1}^{n} Y_j(t) e^{\beta_0' Z_j(t)}}$$

apenas outra forma de reescrever, já que $\int_{t=0}^{u} f\,dN_i(t) = \sum_{t \in O_i} f(t_k)$, O_i = conj. dos inst. de salto de $N_i(\cdot)$

(12) $$= \sum_{i=1}^{n} \int_{t=0}^{u} n^{-1/2} \left(Z_i(t) - E_0(t) \right) dM_i(t)$$, uma vez que

$$dM_i(t) = dN_i(t) - \Lambda_i(t)\,dt \quad \text{e} \quad \sum_{i=1}^{n} (Z_i(t) - E_0(t)) \Lambda_i(t) = \sum_{i=1}^{n} Z_i(t) Y_i(t) \lambda_0(t) e^{\beta_0' Z_i(t)}$$

$$- E_0(t) \cdot \sum_{i=1}^{n} Y_i(t) \lambda_0(t) e^{\beta_0' Z_i(t)} = 0$$

Agora, $n^{-1/2}(Z_i(t) - E_0(t))$ é um vetor de processos previsíveis (depende apenas de constantes e dos proc. previsíveis Y_j e Z_j, $j \geq 1$), de modo que $n^{-1/2} D \log L(\beta_0, t)$, visto como

MAE5783 - Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Maicon A. Ribeiro
Fábio H. Oki

Cont. 1.

um processo em t, é a soma de n martingais, isto é, também é um martingal. Agora, tudo se resume a verificar que

$$M^{(n)}(t) = \frac{1}{n^{1/2}} D \log L(\beta_0, t) \xrightarrow{\mathcal{D}} W, \qquad (\star)$$

quando $n \nearrow +\infty$. Na verdade, para a demonstração de $(\star)$, se considera a versão multivariada da abordagem discutida antes, uma vez que $M^{(n)}(t)$ é, para cada $t \geq 0$, um vetor. Ressaltamos que a versão multivariada em pauta existe e que não seria necessária se β e $Z_i(t)$ fossem escalares. Assim como no artigo, vamos assumir que esse seja o caso, para simplicidade. Até agora, não estabelecemos muito bem o significado dos saltos de $M^{(n)}$ convergirem a algo muito pequeno quando $n \nearrow +\infty$: para entendê-lo, o autor considera o caso especial em que $|Z_i(t)| \leq C < +\infty$ $\forall i$ e t e p/ alguma cte C (essa condição não é necessária p/ o resultado final). Neste caso é fácil ver que o integrando $Z_i(t) - E_0(t)$ em (12) é também limitado por C. Cada M_i tem saltos de tamanho ± 1, coincidentes com os de N_i. Uma vez que não há múltiplos saltos no mesmo instante, os saltos de $M^{(n)}$ são dominados por $n^{-1/2} C$, que converge à zero quando $n \nearrow +\infty$. Esta é a condição (1).

Para a condição (2), devemos avaliar o processo $\langle M^{(n)} \rangle$. Usando os resultados da seção anterior (também já visto em aula), segue que

$$\langle M^{(n)} \rangle = \int_0^t \frac{1}{n} \sum_{i=1}^{n} (Z_i(s) - E_0(s))^2 \Lambda_1(s)\, ds$$

$$= \int_0^t \left(\frac{1}{n} \sum_{i=1}^{n} Z_i(s)^2 Y_i(s) e^{\beta_0' Z_i(s)} - \frac{\left(\frac{1}{n} \sum_{i=1}^{n} Z_i(s) Y_i(s) e^{\beta_0' Z_i(s)} \right)^2}{\frac{1}{n} \sum_{i=1}^{n} Y_i(s) e^{\beta_0' Z_i(s)}} \right) \lambda_0(s)\, ds$$

Desse modo, $\langle M^{(n)} \rangle(t)$ pode ser expresso em termos de uma média simples de $Y_i(s) Z_i(s)^r e^{\beta Z(s)}$, $r = 0, 1 \text{ e } 2$, o que converge em probabilidade à uma constante, como menciona o autor, o que satisfaz (2) e verifica, por fim, (*).

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Marcos Aparecido Pinheiro
Fábio Hideto Oki

Professora: Gisela Tunes da Silva

2. Suponha que um pesquisador esteja interessado em comparar dois tratamentos (tratamento 0 e tratamento 1) em relação a sobrevivência de pacientes desde o início da terapia. Neste estudo, as observações foram pareadas e foram incluídos Q pares. Dentro de cada par, os tratamentos 0 e 1 foram atribuídos de maneira aleatória. Assuma um modelo de Cox estratificado para esse problema, ou seja,

$$\alpha_q(t \mid Z_{qi}) = \alpha_{0q}(t) \exp\{\beta Z_{qi}\}, \quad i=1,2, \quad q=1,\dots,Q$$

em que i indica o paciente do par q e Z_{qi} é a variável indicadora do tratamento. Sem perda de generalidade, pode-se escrever $(Z_{q1}, Z_{q2}) = (0,1)$ para todo $q=1,\dots,Q$.

d. Estabeleça a hipótese nula de não haver efeito do tratamento na sobrevivência dos pacientes.

Cada um dos Q estratos contém 2 indivíduos: o que recebeu o tratamento e o que não recebeu. Supondo, sem perda de generalidade,

$$(Z_{q1}, Z_{q2}) = (0,1), \quad q=1,\dots,Q,$$

temos

$$\alpha_q(t \mid Z_{q1}) = \alpha_{0q}(t)\, e^{\beta Z_{q1}} = \alpha_{0q}(t) \quad , q=1,\dots,Q$$
$$\alpha_q(t \mid Z_{q2}) = \alpha_{0q}(t)\, e^{\beta Z_{q2}} = \alpha_{0q}\, e^{\beta}$$

Nessas condições, supondo β comum para todos os estratos, a hipótese nula

de não haver efeitos do tratamento na sobrevivência do paciente é dada por

$$H_0 : \beta = 0. \qquad (1)$$

Para ser mais geral, poderíamos pensar em uma interação entre as covariáveis e estratos. Nesse caso, considera-se o parâmetro $\Theta = (\beta_1, \ldots, \beta_Q)$, em que β_q é o parâmetro associado ao modelo para o estrato q, $q = 1, \ldots, Q$, i.e.,

$$\alpha_q(t/Z_i) = \alpha_{0q}(t)\, e^{\beta_q Z_{qi}} \quad , i = 1, 2, \quad q = 1, \ldots, Q,$$

e a hipótese nula de não haver efeitos do tratamento na sobrevivência do paciente é dada por

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_Q = 0. \qquad (2)$$

b) Assumindo censura independente e não informativa, encontre a estatística do teste de escore para testar a hipótese nula estabelecida no item a), incluindo a possibilidade de empates entre os tempos de falha de um por (utilize a informação observada para o cálculo da estatística de teste). Simplifique ao máximo a expressão obtida.

Nesse caso, a estatística de teste é dada por

$$S = U(0)^T [I(0)]^{-1} U(0), \quad (\text{Basta testar } (1)) \qquad (3)$$

em que U é o vetor escore associado à verossimilhança construída incluindo a possibilidade de empates e I a matriz de informação observada, como indicado no enunciado.

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Marcon A. Pinheiro
Fábio H. Oki

Continuação 2 b).

A fim de caracterizar S, vamos primeiramente obter a função de verossimilhança. Para tal, sejam $I_1 \subseteq \{1, \dots, Q\}$ o conjunto dos índices de estratos em que não houve empates e façamos $I_2 = \{1, \dots, Q\} \setminus I_1$ (conjunto dos índices de estratos em que ocorrem empates). Ademais seja $Q_1 = |I_1|$ e $Q_2 = |I_2|$, $Q_1 + Q_2 = Q$.

Com essa formulação, podemos escrever a função de verossimilhança da seguinte forma:

$$L(\beta) = \left(\prod_{q \in I_1} L_q(\beta) \right) \left(\prod_{q \in I_2} L_q(\beta) \right) \qquad (4)$$

Ou ainda, poderíamos expressá-la da seguinte modo

$$L(\beta) = \left(\prod_{q=1}^{Q_1} L_q(\beta) \right) \left(\prod_{q=Q_1+1}^{Q_1+Q_2} L_q(\beta) \right), \qquad (5)$$

o que segue ao reordenar os fatores (ou estratos), de modo que os primeiros sejam aqueles em que não houve empates. De (5), segue que

$$\ell(\beta) = \sum_{q=1}^{Q_1} \ell_q(\beta) + \sum_{q=Q_1+1}^{Q} \ell_q(\beta) \qquad (6)$$

No primeiro somatório não há empates, de modo que

$$\ell_q(\beta) = \sum_{k=1}^{L(q)} \left[\beta Z_{qj} - \ln \left(\sum_{l \in R_j^{(q)}} e^{\beta Z_{ql}} \right) \right], \quad q \in I_1 \text{ ou } q = 1, \dots, Q_1 \text{ em (6)}$$

$$= \sum_{k=1}^{2} \delta_{qk} \left[\beta Z_{qk} - \ln \left(\sum_{l \in R_k^{(q)}} e^{\beta Z_{ql}} \right) \right], \qquad (7)$$

com δ_{qi} o indicador de falha, $L(q)$, o número de instantes de falhas distintos no estrato q, $R_j^{(q)}$ o grupo de indivíduos em risco em t_j^q.

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Marcos A. Pinheiro
Fábio H. Oki

Continuação 2b)

No segundo somatório em (6), ocorrem as falhas e em tempos iguais, demodo que

$$\ell_q(\beta) \overset{*}{=} \frac{e^{\beta s_1}}{\sum_{j\in R(t_1)} e^{\beta Z_{qj}}} = \frac{e^{\beta(\overset{0}{Z_{q1}}+\overset{1}{Z_{q2}})}}{\sum_{k=1}^{2} e^{\beta Z_{qk}}} = \frac{e^{\beta}}{e^{\beta\cdot 0}+e^{\beta\cdot 1}} = \frac{e^{\beta}}{1+e^{\beta}}, \quad q\in I_2 \text{ ou } q=Q_1+1,\dots,Q \text{ em}$$

* Usando a aproximação de Breslow, já que a mesma funciona bem quando há poucos empates. (8)

Agora, vamos à obtenção de $U(\beta)$ e $I(\beta)$. Por definição,

$$U(\beta) = \frac{d}{d\beta}\ell(\beta) = \frac{d}{d\beta}\left(\sum_{q=1}^{Q_1}\ell_q(\beta) + \sum_{q=Q_1+1}^{Q}\ell_q(\beta)\right) \qquad (9)$$

$$= \sum_{q=1}^{Q_1}\frac{d}{d\beta}\ell_q(\beta) + \sum_{q=Q_1+1}^{Q}\frac{d}{d\beta}\ell_q(\beta)$$

e

$$I(\beta) = -\frac{d}{d\beta^2}\ell(\beta) = \sum_{q=1}^{Q_1}\frac{d}{d\beta^2}\ell_q(\beta) + \sum_{q=Q_1+1}^{Q}\frac{d}{d\beta^2}\ell_q(\beta) \qquad (10)$$

Substituindo em (9) e (10) as expressões obtidas em (7) e (8), temos

$$U(\beta) = \sum_{q=1}^{Q_1}\sum_{k=1}^{2}\delta_{qk}\frac{d}{d\beta}\left[\beta Z_{qk} - \ln\left(\sum_{\ell\in R_k^{(q)}} e^{\beta Z_{q\ell}}\right)\right] + \sum_{q=Q_1+1}^{Q}\frac{d}{d\beta}\left(\frac{e^{\beta}}{1+e^{\beta}}\right)$$

$$= \sum_{q=1}^{Q_1}\sum_{k=1}^{2}\delta_{qk}Z_{qk} - \sum_{q=1}^{Q_1}\sum_{k=1}^{2}\delta_{qk}\frac{\sum_{\ell\in R_k^{(q)}} Z_{q\ell}e^{\beta Z_{q\ell}}}{\sum_{\ell\in R_k^{(q)}} e^{\beta Z_{q\ell}}} + (Q-Q_1)\left[\frac{e^{\beta}+e^{2\beta}-e^{2\beta}}{1+2e^{\beta}+e^{2\beta}}\right] \qquad (11)$$

e, naturalmente,

$$I(\beta) = -\frac{d}{d\beta^2}\ell(\beta) = -\frac{d}{d\beta}\left(\frac{d}{d\beta}\ell(\beta)\right) = -\frac{d}{d\beta}U(\beta) = -\frac{d}{d\beta}\left(\sum_{q=1}^{Q_1}\sum_{k=1}^{2}\delta_{qk}\frac{\sum_{\ell\in R_k^{(q)}} Z_{q\ell}e^{\beta Z_{q\ell}}}{\sum_{\ell\in R_k^{(q)}} e^{\beta Z_{q\ell}}}\right) + (Q-Q_1)\frac{(e^{\beta}-e^{2\beta})}{(1+e^{\beta})^2} \qquad (12)$$

e

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Maicon A. Pinheiro

Fábio H. Oki

Continuação 2b.

Para tornar as expressões (11) e (12) mais agradáveis, vamos estudá-las um pouco mais. Façamos assim:

$I_1^1 \subseteq \{1, \dots, Q_1\}$: cjto dos índices dos estrados em ambos os obs. foram censura

$I_1^2 \subseteq \{1, \dots, Q_1\} / I_1^1$: " em que houve uma falha depois de uma censura

$I_1^3 \subseteq \{1, \dots, Q_1\} / (I_1^1 \cup I_1^2)$: " em que houve uma falha antes de uma censura

$I_1^4 = \{1, \dots, Q_1\} / (I_1^1 \cup I_1^2 \cup I_1^3)$: " em que ambos falharam

e definamos $Q_1^j = |I_1^j|$, $j = 1, \dots, 4$. Assim, (11) fica dada por

$$U(\beta) = \left(\sum_{q=Q_1^1+1}^{Q_1^1+Q_1^2} Z_{q^*} \underbrace{- \cancelto{0}{\sum_{q=Q_1^1+1}^{Q_1^1+Q_1^2}} \frac{Z_{q^*} e^{\beta Z_{q^*}}}{e^{\beta Z_{q^*}}}}_{\text{em que houve uma falha depois de uma censura } Z_{q^*} \text{ é a cov. para a obs. que falhou}} \right) + \left(\sum_{q=Q_1^1+Q_1^2+1}^{Q_1^1+Q_1^2+Q_1^3} Z_{q^*} - \sum_{q=Q_1^1+Q_1^2+1}^{Q_1^1+Q_1^2+Q_1^3} \frac{e^{\beta}}{1+e^{\beta}} \right)$$

$$+ \underbrace{\sum_{q=Q_1^1+Q_1^2+Q_1^3+1}^{Q_1} 1}_{Q_1^4} - \sum_{q=Q_1^1+Q_1^2+Q_1^3+1}^{Q_1} \underbrace{\left(\frac{\sum_{l \in I_1^{4'}} Z_{ql} e^{\beta Z_{ql}}}{\sum_{l \in I_1^{4'}} e^{\beta Z_{ql}}} + \frac{\sum_{l \in I_1^{4''}} Z_{ql} e^{\beta Z_{ql}}}{\sum_{l \in I_1^{4''}} e^{\beta Z_{ql}}} \right)}_{-V} + (Q - Q_1) \frac{e^{\beta}}{(1+e^{\beta})^2}$$

$$= \sum_{q=Q_1^1+Q_1^2+1}^{Q_1^1+Q_1^2+Q_1^3} Z_{q^*} - Q_1^3 \frac{e^{\beta}}{1+e^{\beta}} + Q_1^4 - V + (Q - Q_1) \frac{e^{\beta}}{(1+e^{\beta})^2}. \qquad (13)$$

Para simplicar V, vamos considerar $I_1^{4'} \subset I_1^4$ o cjto dos índices em I_1^4 em que a falha ocorreu primeiro para que estava sem o tratamento e façamos $I_1^{4''} = I_1^4 / I_1^{4'}$, com cardinalidades $Q_1^{4'} := |I_1^{4'}|$ e $Q_1^{4''} = |I_1^{4''}|$. Nesse caso,

$$V = \sum_{q=Q_1^1+Q_1^2+Q_1^3+1}^{Q_1^1+Q_1^2+Q_1^3+Q_1^{4'}} \left(\frac{e^{\beta}}{1+e^{\beta}} + \frac{e^{\beta}}{e^{\beta}} \right) + \sum_{q=Q_1^1+Q_1^2+Q_1^3+Q_1^{4'}+1}^{Q} \left(\frac{e^{\beta}}{1+e^{\beta}} \right) = Q_1^{4'} + Q_1^{4''} \frac{e^{\beta}}{1+e^{\beta}} \qquad (14)$$

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Marcon A. Pinheiro

Fábio H. Oki

Continuação 2 b.

Logo, substituindo (14) em (13), temos que

$$U(\beta) = \sum_{q=Q_1^1+Q_1^2+1}^{Q_1^1+Q_1^2+Q_1^3} Z_{q^*} - (Q_1^3+Q_1^4)\frac{e^{\beta}}{1+e^{\beta}} + (Q_1^4 - Q_1^{4'}) + (Q-Q_1)\frac{e^{\beta}}{(1+e^{\beta})^2} \qquad (15)$$

E assim, temos que

$$I(\beta) = (Q_1^3+Q_1^4)\frac{e^{\beta}}{(1+e^{\beta})^2} - (Q-Q_1)\frac{(e^{\beta}-e^{3\beta})}{(1+e^{\beta})^3} \qquad (15)$$

donde segue que

$$U(0) = \sum_{q=Q_1^1+Q_1^2+1}^{\overbrace{Q_1^1+Q_1^2+Q_1^3}^{Q_1^{3'}}} Z_{q^*} - \frac{Q_1^3}{2} - \frac{Q_1^4}{2} + \frac{Q}{4} - \frac{Q_1}{4} \quad e \qquad (16)$$

$$I(0) = \frac{Q_1^3}{4} + \frac{Q_1^4}{4} \qquad (17)$$

$Q_1^{3'}$ denota o número de estratos em que o indivíduo sob trat. falhou antes do que não que recebeu, que por sua vez, foi censurado. Assim a estatística do teste fica dada por

$$S = \left(\frac{2Q_1^{3'}}{2} - \frac{Q_1^3}{2} - \frac{Q_1^4}{2} + \frac{Q/2}{2} - \frac{Q_1/2}{2}\right)\frac{4}{Q_1^3+Q_1^4}\left(\frac{2R(I_1^3) - Q_1^3 - Q_1^4 + Q/2 - Q_1/2}{2}\right)$$

$$= \frac{\left(2Q_1^{3'} - \overbrace{Q_1^3}^{Q_1^{3'}+Q_1^{3''}} - Q_1^4 - Q/2 - Q_1/2\right)^2}{Q_1^3 + Q_1^4}, \quad \text{lembrete.}$$

$$= \frac{\left((Q_1^{3'} - Q_1^{3''}) - Q_1^4 - Q_2/2\right)^2}{Q_1 + Q_1^4}, \quad Q_2 = Q - Q_1$$

\# estrat. em que houve empates

$Q_1^{3''}$: # estratos em que o ind sem trat. falhou antes do sob trat., que foi censurado.

$Q_1^{3'}$ # estratos em que o ind. sob trat. falhou antes do que estava sem tratamento, que foi censur

Q_1 : # estratos em que não houve empates

Q_1^3 : # estratos em que houve uma falha antes de uma censura

Q_1^4 : # estratos em que ambos falharam

Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Maicon A. Pinheiro

Fábio H. Oki

Continuação 2b.

Isto é,

$$S = \frac{\left(Q_1^4 + Q_2/2 + Q_1^{3''} - Q_1^{3'} \right)^2}{Q_1 + Q_1^4} \qquad (18)$$

c) Discuta a informação que não é utilizada na estatística do teste encontrada.

Pela obtenção de S dada em 2.b), vemos que não se considera a informação dos estratos em que ambos os indivíduos foram censurados, o que já era esperado, uma vez que a função de verossimilhança parcial de Cox é uma soma sobre os tempos de falha, antes que qualquer outra soma que envolva as censuras. No entato, é bem razoável, já que não há como dizer se o tratamento é ou não adequado nesse caso. Não se observa discrepâncias entre o que recebeu o trat. e o que não.

Também são desconsiderados os estratos em que se observa uma falha depois de uma censura, o que faz sentido: não há como dizer que o que foi censurado sobreviveria menos que aquele que teve a falha.

Outro fato interessante é que S é composta basicamente por contagens de estratos em que se observam determinadas regras.

MAE 5783 - Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Maicon Aparecido Dinheiro
Fábio Hideto Oki

Professora: Gisela Tunes da Silva

3. Leia o artigo "Proportional hazards tests and diagnostics based on weighted residuals" publicado na Biometrika em 1994, volume 81, número 3, por Terry Therneau e Patricia Grambsch. Explique detalhadamente os resultados apresentados na seção 2 do artigo.

Antes da tarefa em si, vamos recapitular alguns conceitos e introduzi-los conforme a notação dada na seção 1 do artigo. Isso é necessário para o bom entendimento do que será visto na seção 2.

Como usual, associamos de modo respectivo n processos de contagem independentes

$$\{N_i(t) : t \geq 0,\ i = 1, \dots, n\}$$

aos n indivíduos em estudo. O processo de intensidade para $N_i(t)$ é caracterizado por

$$Y_i(t) \exp\{\underset{\sim}{\beta}' \underset{\sim}{Z}_i(t)\} d\Lambda_0(t), \qquad (1)$$

onde

$$Y_i(t) = \begin{cases} 1, & \text{se o indivíduo } i \text{ está em risco em } t; \\ 0, & \text{se " " " foi censurado ou falhou antes de } t, \end{cases}$$

$\underset{\sim}{\beta} = (\beta_1, \dots, \beta_p)' \in \mathbb{R}^p$ um vetor de p parâmetros e $\underset{\sim}{Z}_i$ o vetor de p-covariáveis (em aula, usamos a notação X_i) e $d\Lambda_0(t)$ é uma função taxa de falha não especificada. Apenas recapitulando o que também foi visto em aula, podemos estimar $\underset{\sim}{\beta}$ maximizando o logaritmo neperiano da função de verossimilhança parcial de Cox

$$\sum_{i=1}^{n} \int_0^{+\infty} \left[Y_i(t)\, \underset{\sim}{\beta}' \underset{\sim}{Z}_i(t) - \ln\left\{ \sum_{j=1}^{n} Y_j(t)\, e^{\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)} \right\} \right] dN_i(t). \qquad (2)$$

↳ 1a

MAE 5883 - Análise de Sobrevivência

4ª Lista de Exercícios

Alunos: Marcos Aparecido Pinheiro
Fábio Hideto Oki

Professora: Gisela Tunes da Silva

Continuação Exercício 3:

$$F_t = \sigma\{Z_i(u), N_i(u), Y_i(u+) : 0 \leq u \leq t, i = 1, \dots, n\},$$

e $g_j(t)$, $j = 1, \dots, p$, um vetor de processos previsíveis com respeito à F_t. Suponha que se deseja testar a hipótese de riscos proporcionais como em (1) contra a alternativa de parâmetros variando no tempo com processo de intensidade caracterizado por

$$\begin{aligned} h_i(t)\,dt &= Y_i(t) \exp\{\beta'(t) Z_i(t)\}\, d\Lambda_0(t) \\ &= Y_i(t) \exp[\{\beta + G(t)\theta\}' Z_i(t)]\, d\Lambda_0(t) \end{aligned} \qquad (9)$$

Aqui, há uma generalização da abordagem de Schoenfeld (1982), que considerava a estrutura de não proporcionalidade para apenas uma covariável.

Primeiramente, assuma (9) e β conhecido. Então

$$\begin{aligned} r_k(\beta) &\overset{def}{=} Z_{(k)} - M(\beta, t_k) \\ &= Z_{(k)} - M(\beta, t_k) + [\overbrace{M(\beta(t_k), t_k) - M(\beta(t_k), t_k)}^{0}] \\ &= [Z_{(k)} - M(\beta(t_k), t_k)] + [M(\beta(t_k), t_k) - M(\beta, t_k)], \end{aligned} \qquad (10)$$

que é a soma dos resíduos de Schoenfeld do verdadeiro modelo (v.a. com média 0) e a diferença entre as médias ponderadas das covariáveis sob os modelos verdadeiro e nulo, respectivamente. Ao expandir $M\{\beta(t_k), t_k\}$ em um termo pela expansão de Taylor em torno de $\beta(t_k) = \beta$ na segunda parcela da soma acima obtemos

$$r_k(\beta) = [Z_{(k)} - M(\beta(t_k), t_k)] + [M(\beta, t_k) + M'(\beta, t_k)[\beta(t_k) - \beta] - M(\beta, t_k)] \qquad (11)$$

donde segue que

$$E(r_k(\beta) \mid F_{t_k}) \approx V(\beta, t_k) G_k \theta, \qquad (12)$$

em que $G_k = G(t_k)$ é uma matriz diagonal com $(G)_{jj} = g_j(t)$.

Seja $r_k^* = r_k^*(\beta) = V^{-1}(\beta, t_k) r_k(\beta)$ os resíduos de Schoenfeld escalados. Então,

$$\begin{aligned} E(r_k^* \mid F_{t_k}) &= V^{-1}(\beta, t_k)\, E(r_k(\beta) \mid F_{t_k}) \\ &\approx V^{-1}(\beta, t_k)\, V(\beta, t_k)\, G_k \theta \\ &= G_k \theta, \end{aligned} \tag{13}$$

$$\mathrm{var}(r_k^*) = V^{-1}(\beta, t_k)\, V\{\beta(t_k), t_k\}\, V^{-1}(\beta, t_k) \approx V^{-1}(\beta, t_k) \tag{14}$$

Aqui não o faremos, mas pode-se demonstrar que os r_ks são não correlacionados. Estas duas últimas equações sugerem um modelo padrão linear para r_k^*. Com $V_k = V(\beta, t_k)$, o mínimo quadrado generalizado fornece

$$\hat{\theta} = \left(\sum G_k V_k G_k\right)^{-1} \sum G_k r_k . \tag{15}$$

Isto leva a um teste χ^2_p com p graus de liberdade:

$$\left(\sum G_k r_k\right)' \left(\sum G_k V_k G_k\right)^{-1} \left(\sum G_k r_k\right) \tag{16}$$

para testar $H_0: \theta = 0$.

Agora, assuma β desconhecido e seja $\hat{\beta}$ o estima de máxima verossimilhança parcial sob H_0. Seja $\hat{V}_k = V(\hat{\beta}, t_k)$ e $\hat{r}_k = r_k(\hat{\beta})$. Como acima, temos $E(\hat{V}_k^{-1} \hat{r}_k) \approx G_k \theta$. No entanto, uma vez que $\sum \hat{r}_k = 0$, os resíduos são correlacionados com $\mathrm{cov}(\hat{r}_l, \hat{r}_m)$ consistentemente estimados por $\delta_{lm} V_l - V_l \left(\sum V_k\right)^{-1} V_m$ assintoticamente sob H_0. Desse modo, $\mathrm{var}(\hat{V}_k^{-1} \hat{r}_k) \approx V_k^{-1} - \left(\sum_l \hat{V}_l\right)^{-1}$. Min. quadrados generalizados leva à

$$\hat{\theta} = D^{-1} \sum G_k \hat{r}_k , \tag{17}$$

em que $D = \sum G_k \hat{V}_k G_k - \left(\sum G_k \hat{V}_k\right)\left(\sum \hat{V}_k\right)^{-1}\left(\sum G_k \hat{V}_k\right)'$ sob H_0 a variância assintótica de $n^{-1/2} \sum G_k \hat{r}_k$ pode ser consistentemente estimada por $n^{-1} D$, levando a

$$T(G) = \left(\sum G_k \hat{r}_k\right)' D^{-1} \left(\sum G_k \hat{r}_k\right) \underset{\text{usando-se teor. martingales}}{\sim} \chi^2_p \text{ qdo } n \to \infty \tag{18}$$

Por fim, em função do $g(t)$ considerado, derivam-se 5 testes.

1.º $g(t)$ uma função específica do tempo, então $T(G)$ é um teste de escore para a adição da variável dependente no tempo $g(t) Z$ no modelo;

Análise de Sobrevivência — Alunos: Maicon A. Pinheiro

4ª Lista de Exercícios — Fábio H. Oki

Continuação 3.

2. Se g é uma função simples em intervalos não sobrepostos com os intervalos e constantes escolhidos de antemão, $T(G)$ é o teste de escore proposto por O'Quigley e Pessione (1989).

3. Se $g(t) = \bar{N}(t^-)$, então $T(G)$ é a covariância entre os resíduos escalados de Schoenfeld e o rank dos tempos de eventos. Nesse caso, o teste resultante é similar ao de Harrel (1986).

4. Lin (1991) sugere comparar $\hat{\beta}$ e $\hat{\beta}_g$ da eq. de estimação ponderada $\sum G_i r_k = 0$ com $g(t)$ uma das funções de peso usualmente escolhidas para testes de log rank. Lin mostrou que $\hat{\beta} - \hat{\beta}_g \xrightarrow{D} N(0, \Sigma)$, quando $n \to +\infty$, Σ obtida pela teoria de martingais.

5. $g_j(t_1) = 0$ e $g_j(t_{k+1}) = a_j^2 \hat{r}_{jk}$, $j = 1, \dots, p$. Esta leva à estatística de teste de Nagelkerke, Oosting e Hart (1984).

O ponto principal é que cada um dos testes acima podem ser diretamente visualizados como um simples teste de tendência aplicado ao gráfico de $g(t)$ versus $\hat{r}^*$, o que é bem mais simples em termos práticos.

Definição. (Processo de Contagem) Um processo de contagem é um processo estocástico $\{N(t) : t \geq 0\}$ adaptado a $\mathcal{F}_t$ tal que

i. $N(0) = 0$

ii. $N(t) < +\infty \ \forall t$ (q.c.)

iii. As trajetórias de $N(t)$ são contínuas à direita.

iv. $N(t)$ é constante com "saltos" ou descontinuidades de tamanho $+1$.

Ex: $N_i(t) = I(Z_i \leq t, \delta_i = 1)$, $N(t) = \sum_{i=1}^{n} N_i(t)$

Definição: O Processo de Intensidade associado ao processo $N(t)$ é definido como

$$d\Lambda(t) = E[dN(t) / \mathcal{F}_{t^-}]$$

em que $dN(t) = N(t) - N(t^-)$ e $\mathcal{F}_t = \sigma\{N(u), I(Z \leq u, \delta = 0) : 0 \leq u \leq t\}$ ou $\mathcal{F}_t = \sigma\{N(u), Y(u^+) : 0 \leq u \leq t\}$.

Definição. O Processo de Intensidade Acumulado é dado por

$$\Lambda(t) = \int_0^t d\Lambda(t)$$

Podemos escrever

$$\Lambda(t) = \int_0^t \lambda(u)\,du = \int_0^t Y(u)\alpha(u)\,du.$$

Obs: $\Lambda(t)$ fornece, aproximadamente, o número esperado de saltos em $(0, t)$.

Definição: (Martingal) Seja $X = \{X(t) : t \geq 0\}$ um processo estocástico contínuo à direita e $\{\mathcal{F}_t : t \geq 0\}$ uma filtragem definidas em um mesmo espaço de probabilidade.

Se

1. X é adaptado à $\mathcal{F}_t$;
2. $E|X(t)| < +\infty \ \forall t < +\infty$; , então X é um martingal.
3. $E(X(t+s) / \mathcal{F}_s) = X_s$

Obs: Se (1) e (2) valem, então X é um martingal se e só se

$$E(dX(t) / \mathcal{F}_{t^-}) = 0$$

Definição. (Processo Previsível) Um processo estocástico X é Previsível com relação à filtragem $\mathcal{F}_t$ se, para cada t, o valor de X é especificado ou é uma função de $\mathcal{F}_{t^-}$.

Mais formalmente, $X(t)$ é previsível se $X(t)$ é $\mathcal{F}_{t^-}$-mensurável.

Resultados: 1. Todo processo ~~previsível~~ contínuo à esquerda é previsível

2. Se X é previsível, então $E(X(t) \mid \mathcal{F}_{t^-}) = X(t)$

Ex: $Y(t) = I(Z \geq t)$, $Z = \min(T, C)$, é um proc. prev.

Vimos ainda que

$$d\Lambda(t) = \underbrace{Y(t)}_{\text{proc. prev.}} \underbrace{\alpha(t)}_{\text{f. det.}} dt$$

$\Rightarrow d\Lambda(t)$ é um Proc. Previsível. Logo,

$$E[d\Lambda(t) \mid \mathcal{F}_{t^-}] = d\Lambda(t)$$

Resultado: $M(t) := N(t) - \Lambda(t)$ é um martingal.

. Fleming, pg 26.

Definição. Submartingal

Seja o processo $X = \{X(t) : t \geq 0\}$ cont. à direita e $\mathcal{F}_t$ uma filtragem definidos no mesmo espaço de prob. Se

1. X é adaptado a $\mathcal{F}_t$;
2. $E(|X(t)|) < +\infty \ \forall t < +\infty$; e
3. $E(X(t+s) / \mathcal{F}_s) \geq X(s)$

então X é submartingal.

Teorema. (Decomposição de Doob-Meyer) Seja X um submartingal contínuo à direita e não-negativo adaptado à filtragem $\mathcal{F}_t$. Então, existe um martingal M cont. à direita e um proc. previsível A tal que

$$E(A(t)) < +\infty \text{ e}$$

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

Se $A(0) = 0$ e $X = M' + A'$ é uma outra decomposição com $A'(0) = 0$, então $\forall t \geq 0$,

$$\mathbb{P}(M'(t) \neq M(t)) = \mathbb{P}(A'(t) \neq A(t)) = 0.$$

Corolário. Seja $\{N(t), t \geq 0\}$ um proc de contagem adaptado à filtragem $\{\mathcal{F}_t : t \geq 0\}$ com $E(N(t)) < \infty$ $\forall t$. Então existe um único processo $\Lambda(t)$ não decrescente e previsível t.q.

. $\Lambda(0) = 0$; . $E(\Lambda(t)) < +\infty \ \forall t \geq 0$ e $M(t) = N(t) - \Lambda(t)$ é um martingal.

$\Lambda(t)$ é conhecido como compensador de $N(t)$.

Definição: Processo de Variação Previsível

Para um martingal contínuo à direita e adaptado à $\{\mathcal{F}_t, t \geq 0\}$, define-se o proc. de variação previsível $\langle M \rangle (t)$ por

$$\langle M \rangle (t) = \lim_{n \to \infty} \sum_{k=1}^{n} Var(\Delta M_k / \mathcal{F}_{(k-1)t/n})$$

em que o intervalo $(0, t]$ é particionado em n intervalos de comprimento t/n e

$$\Delta M_k = M(kt/n) - M((k-1)t/n)$$

Resultado: $M^2(t) - [M](t)$ é mart. media zero

$Var(M(t)) = E([M](t))$

Ex: $M(t) = N(t) - \Lambda(t)$; $\langle M\rangle(t) = \Lambda(t)$

$[M](t) = \sum_{s\le t} (dM(s))^2 = \sum_{s\le t} (dN(s))^2 = \sum_{s\le t} dN(s) = N(t)$

Informalmente, $d\langle M\rangle(t) = Var(dM(t)/\mathcal{F}_{t^-})$ ou $\langle M\rangle(t) = \int_0^t Var(dM(u)/\mathcal{F}_{u^-})$

Obs: $Var(dM(t)/\mathcal{F}_{t^-}) = E((dM(t))^2/\mathcal{F}_{t^-}) - (\underbrace{E(dM(t)/\mathcal{F}_{t^-})}_{0})^2$

$= E(dM(t)^2/\mathcal{F}_{t^-})$

Resultado: Seja M um martingal cont. à direita com relação à filtragem $\{\mathcal{F}_t : t\ge 0\}$ e assuma $E(M^2(t)) < +\infty \ \forall t\ge 0$. Então existe um único processo previsível $\langle M\rangle(t)$ denominado Processo de Variação Previsível de M tal que

- $\langle M\rangle(0) = 0$; • $E(\langle M\rangle(t)) < +\infty \ \forall t$ e • $\{M^2(t) - \langle M\rangle(t) : t\ge 0\}$ é um martingal cont. à direita.

Obs: Assuma $M(0) = 0$. Então $E[M^2(t) - \langle M\rangle(t)] = 0 \Rightarrow E(M^2(t)) = Var(M(t)) = E(\langle M\rangle(t))$

Ex: $M(t) = N(t) - \Lambda(t)$, $\Lambda(t) = \int_0^t Y(u)\lambda(u)du$.

$d(\langle M\rangle(t)) = Var(dM(t)/\mathcal{F}_{t^-}) = Var(dN(t) - \overset{\text{pode uma vez que é previsível}}{d\Lambda(t)}/\mathcal{F}_{t^-}) = Var(dN(t)/\mathcal{F}_{t^-})$

$dN(t)$ é binária, de modo que

$$d(\langle M\rangle(t)) = E(dN(t)/\mathcal{F}_{t^-}) - [E(dN(t)/\mathcal{F}_{t^-})]^2 = d\Lambda(t) - (d\Lambda(t))^2$$

$$d(\langle M\rangle(t)) \approx d\Lambda(t) \quad \text{ou} \quad \langle M\rangle(t) = \Lambda(t)$$

Definição: Sejam M_1 e M_2 dois mart. cont. à direita e assuma $E(M_i(t)) < +\infty$, $\forall t\ge 0$, $i = 1,2$. O processo de Covariância Previsível é dado por

$$\langle M_1, M_2\rangle(t) = \lim_{n\to\infty} \sum_{k=1}^{n} Cov(\Delta M_{1k}, \Delta M_{2k} / \mathcal{F}_{(k-1)t/n})$$

em que $[0,t]$ é part. em n subint. de comp. t/n e $\Delta_{ik} = M_i(tk/n) - M_i(t(k-1)/n)$, $i = 1,2$.

Podemos escrever

$$d\langle M_1, M_2\rangle(t) = Cov(dM_1(t), dM_2(t)/\mathcal{F}_{t^-})$$

Resultado: M_1 e M_2 dois mart. cont. à direita c.r. $\{\mathcal{F}_t : t\ge 0\}$ e assuma $E(M_i^2(t)) < +\infty \ \forall t\ge 0$, $i = 1,2$. Então Existe um proc. previsível cont. à direita $\langle M_1, M_2\rangle$ chamado de proc. de covariância com $\langle M_1, M_2\rangle(0) = 0$,

$$E(\langle M_1, M_2\rangle(t)) < +\infty \text{ t.q.}$$

1. $\langle M_1, M_2\rangle$ é a dif. de dois proc. crescentes previsíveis contínuos à direita
2. $M_1 M_2 - \langle M_1, M_2\rangle$ é um martingal

Def. (Proc. Variação Quadrat.) M um mart. contínuo à direita e adapt. à $\{\mathcal{F}_t, t\ge 0\}$, define-se o PVQ $[M](t)$

$$[M](t) = \lim_{n\to\infty} \sum_{k=1}^{n} (\Delta M_k)^2 \quad \text{com } \Delta M_k = M(tk/n) - M(t(k-1)/n) \text{ em que } [0,t]$$

Podemos escrever

$$[M](t) = \sum_{s\le t} (dM(s))^2$$ soma é nos int. de descont. de M

Integrais Estocásticas. Uma possível def. é

$$L(t) = \int_0^t H(s)\,dM(s) = \lim_{n\to\infty} \sum_{k=1}^{n} H_k \Delta M_k$$

em que $(0,t]$ é part. em subintervalos de tamanho t/n.

$$H_k = H\left(\frac{(k-1)t}{n}\right), \quad \Delta M_k = M\left(\frac{tk}{n}\right) - M\left(\frac{t(k-1)}{n}\right)$$

$H(t)$ é um proc. previsível e $M(t)$ é um mart. relação a história $\mathcal{F}_t$.

Resultado: $L(t)$ é um mart. (sob cond. gerais).

$E(dL(t)/\mathcal{F}_{t^-}) = E(H(t)\,dM(t)/\mathcal{F}_{t^-}) = H(t)\,E(dM(t)/\mathcal{F}_{t^-}) = 0$. Além disso, tem média zero.

Proc. de Variação Previsível de $L(t)$: $d\langle L\rangle(t) \overset{def}{=} Var(dL(t)/\mathcal{F}_{t^-})$

$= Var(H(t)\,dM(t)/\mathcal{F}_{t^-})$

$= H^2(t)\,Var(dM(t)/\mathcal{F}_{t^-})$

$= H^2(t)\,d\langle M\rangle(t)$, ou $\langle L\rangle(t) = \int_0^t H^2(s)\,d\langle M\rangle(s)$

Proc. de Var. Opcional:

$$[L](t) = \sum_{s\le t} (dL(s))^2 = \sum_{s\le t} (H(s)\,dM(s))^2 = \sum_{s\le t} H^2(s)(dM(s))^2$$

$$= \sum_{s\le t} H^2(s)\,dN(s) = \int_0^t H^2(s)\,dN(s)$$

Processo de Wiener (Padrão): Um proc. est. satisf. • $W(0)=0$ e $E(W(t))=0\ \forall t$; • $W(t)$ tem incrementos indep., ou seja, $W(t)-W(u)$ é ind. de $W(u)$, $0\le u<t$; • $W(t)-W(s) \sim N(0, t-s)$; • W tem traj. contínuas.

Seja agora $V(t)$ uma f. est. cresc., contínua com $V(0)=0$. Considere o proc.

$$U(t) = W(V(t)).$$

Então $U(t)$

• tem traj. contínuas; • tem incrementos ind.; • $W(V(t)) - W(V(s)) \sim N(0, V(t)-V(s))$.

$U(t)$ é dito um martingal Gaussiano com proc. de var. previsível $\langle U\rangle(t) = V(t)$.

Proc. de Contagem Multivariados

Def. O proc. $\{N_1, \dots, N_n\}$ é chamado de proc. de cont. mult. se cada N_i é um proc. de cont. e dois proc. não "saltam" no mesmo inst.

Teo: Seja $\{N_1, \dots, N_n\}$ um proc. de ctg. mult. e Λ_i seus resp. compens., $i=1,\dots,n$. Defina

$$U_\ell(t) = \sum_{i=1}^{n} \int_0^t H_{i,\ell}\,dM_i(u), \quad \ell = 1, 2, \dots, k$$

$H_{i,\ell}$ são proc. previsíveis limitados. Então i. $U_\ell(t)$ é martingal; (ii) $E(U_\ell(t)) = 0\ \forall\, 0\le t<\infty$;

iii. $Cov(U_\ell(s), U_\ell(t)) = \sum_{i=1}^{n} \int_0^{min(t,s)} E[H_{i,\ell}(u) H_{i,\ell'}(u)\,d\Lambda(u)]$

De (iii) decorre que $Var(U_\ell(t)) = \sum_{i=1}^{n} \int_0^t E(H_{i,\ell}^2(u)\,d\Lambda_i(u))$

$$Cov(U_\ell(s), U_\ell(t)) = \sum_{i=1}^{n} \int_0^{min} E(H_{i,\ell}^2(u)\,d\Lambda_i(u))$$

Teo. de Rebolledo. (Versão p/ proc. de contagem)

Sejam $\underline{N(t)} = (N_1(t), N_2(t), \dots, N_n(t))$ um proc. cont. mult. e $\Lambda(t) = (\Lambda_1(t), \dots, \Lambda_n(t))$ as resp. compensadores. Para proc. previsíveis

$$H_{ji}(t),\ j = 1, 2, \dots, k;\ i = 1, 2, \dots, n$$

defina os martingais

$$U_j^{(n)}(t) = \sum_{i=1}^{n} \int_0^t H_{ji}(u)\,(dN_i(u) - d\Lambda_i(u)),\ j = 1, 2, \dots, k$$

$$U_{j0}^{(n)}(t) = \sum_{i=1}^{n} \int \dots$$

$$\underset{\sim}{U}^{(n)}(t) = (U_1^{(n)}(t), U_2^{(n)}(t) \dots$$

Considere as cond.

(a) $\langle \underset{\sim}{U}^{(n)} \rangle (t) \xrightarrow[n \to +\infty]{P} V(t)$

(b) $\langle U_{j0}^{(n)} \rangle (t) \xrightarrow{P} 0 \dots$

Se a e b vál., $\underset{\sim}{U}^{(n)}(t) \to \underset{\sim}{U}^{(\infty)}$ mart. gaussiano com $\langle U^{(\infty)} \rangle (t) = V(t)$. Em part., para t_0 fixado

$$\underset{\sim}{U}^{(n)}(t_0) \xrightarrow{D} N(0, V(t_0))$$

Estimador de Nelson-Aalen (Não paramétrico da Função de Risco Acumulada)

Sejam T: o tempo de falha com fdens. prob. $f(\cdot)$, fi. sobr. $S(\cdot)$ e C tempo de censura, $T \perp C$.
Defina

$$\delta = I(T \leq C) \quad e \quad Z = \min(T, C)$$

Assuma que temos uma amo. (Z_i, δ_i), $i = 1, 2, \dots, n$ de (Z, δ). Sejam $t_1 < t_2 < \dots < t_k$ tempos de falha, ou seja, inst. em q. pelo menos uma falha ocorr. Note que $k \leq n$ (pode haver empate)

Para $j = 1, 2, \dots, k$, defina

d_j: nº de falhas em t_j

r_j: # indivíduos em risco em t_j^- (que não falharam ou fi censurados até t_j^-)

Intuitivamente, $\hat{\alpha}(t_j) = \dfrac{d_j}{r_j}$. O estimador de Nelson-Aalen é dado por

$$\hat{A}(t) = \sum_{j:\, t_j \leq t} d_j / r_j = \int_0^t d\bar{N}(u) / \bar{Y}(u) \qquad \bar{N}(u) = \sum_{i=1}^{n} N_i(t)$$

= nº de falhas até t

$\bar{Y}(t) = \sum_{i=1}^{n} Y_i(t)$ = ind. risco em t.

Propriedades do Estimador de Nelson-Aalen

Sejam $A^*(t) = \int_0^t I(\bar{Y}(u) > 0)\,\alpha(u)\,du$ e $\pi_i(t) = P(Z_i \geq t) =: \pi(t)$ (Flem. 92)

Supondo $A(t) < \infty\ \forall t \geq 0$, temos

① $E(\hat{A}(t) - A^*(t)) = 0$, ② $E(\hat{A}(t) - A(t)) = -\int_0^t (1 - \pi(s))^n\, dA(s) \geq -(1 - \pi(t))^n A(t)$

③ $\sigma_t^2 \overset{def}{=} E\left(\left[\sqrt{n}\,(\hat{A}(t) - A^*(t))\right]^2\right) = E\left(n \int_0^t \frac{I(\bar{Y}(s) > 0)}{\bar{Y}(s)}\,\alpha(s)\,ds\right)$

Obs: (1) e (2) ver notas
$A^*(t) \leq A(t)$ $E(\hat{A}(t) - A^*(t)) \leq 0$

. O estimador de N-A tem viés neg., que tende à zero qdo n→∞ e tem uma qtde neg. $-(1-\alpha_1(t))^n A(t)$

Demonstremos das propriedades do est. N-A.

Na aula 14, falamos sobre o modelo de Regressão Exponencial. Apresentamos suas duas formulações ($\lambda_i = e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}}$. 1. $T_i / \underset{\sim}{x}_i \sim exp(e^{\underset{\sim}{x}_i' \underset{\sim}{\beta}})$ 2. $Y_i = \ln T_i = \underset{\sim}{x}_i' \underset{\sim}{\gamma} + W_i$, $\underset{\sim}{\gamma} = -\underset{\sim}{\beta}$, $W_i \sim VEP$). Vimos que é um modelo de vida acelerado e derivamos a f.v. do mesmo. ✓

15. Modelos Weibull, Log-logístico, Log-Normal, Locação-Escala, Resultados Assint. para mod. paramétricos, ✓

16. Método de Diagnóstico p/ modelos paramétricos: Método gráfico descritivo; Resíduos Padronizados, Resíduos de Cox-Snell ✓

17. Resíduos Martingal, Resíduos Deviance

→ Modelo Cox!

18. Estimação dos Parâmetros no Mod. de Cox

1. Veross. Parcial de Cox

19. 2. Veross. Perfilada

20. Aprox. na presença de Empates

Breslow (boa qdo poucos empates)

EFRON (correção: ver. exata)

Kalbf (média arit. pl cada das DL poss.)

4ª alt. (Com base mod. discreto)

Estimação F Risco Acumulada e F. sobrevivência

Extensões do Modelo Cox: Cov. dep. no tempo

21. Cov. Dep do Tempo; Prop Assintóticas $\hat{\beta}$

22. Cox Estratificado; Adequação Cox; Mel. Graf Prop Riscos

23. Resíduos no Mod. Cox

Cox-Snell - Martingal

24. Deviance

23. Schoenfeld: Testes baseados Res. Sch. plausível

sup prop riscos

~~26. Resíduos~~

Modelo Apareceu Primeiro, $S(t) = e^{-A(t)}$, $A(t) = \int_0^t \alpha(s)\,ds$, $f(t) = \alpha(t)S(t)$.

Se T cont. val $t_1, t_2, \ldots$, D é $t_1, t_2, \ldots$, $S(t) = \sum_{j: t_j > t} p_j$, $p_j = P(T > t_j)$; $\alpha_j = \alpha(t_j) = P(T = t_j \mid T > t_j) = p_j / S(t_{j-1})$. $(S(t_{j-1}) - S(t_j)) / S(t_{j-1}) = 1 - S(t_j)/S(t_{j-1})$. P/ $t_j \le t \le t_{j+1}$, $S(t) = S(t_j) = \prod_{j: t_j \le t} (1 - \alpha_j)$, $A(t) = \sum_{j: t_j \le t} \alpha(t_j)$ ou $-\sum_{j: t_j \le t} \ln(1 - \alpha(t_j))$

vida média: $t_{0.5}$: $S(t_{0.5}) = 0.5$ e $F(t_{0.5}) = 0.5$; t_p: $S(t_p) = 1 - p$ e $F(t_p) = p$ \

$t_p = \inf\{t : S(t) \le 1 - p\}$. Idade v. m.r.: vmr$(t) = E[T - t \mid T > t] = \int_t^\infty (u - t) f(u)\,du / S(t) = \int_t^\infty S(u)\,du / S(t)$ \ vmrt$(t) = \sum_{j: t_j > t} (t_j - t) P(T = t_j \mid T > t) / P(T > t) = [(t_k - t) S(t_{k-1}) + \sum_{j=k+1}^{n+1} (t_{j-1} - t_j) S(t_j)] / S(t)$. **Censura** é direita $[T, p]$, Π, Aleatória J, à esquerda e intervalar (ev. ocorre após, antes, dentro int. entre (a, b)). J (ev. obs. ocorre antes do estudo) **Exemplo**; $L(\theta) = \prod (f(y_i))^{\delta_i} (S(c_i))^{1-\delta_i}$

I Progress.: o estado de um int. cens. $L(\theta) = \prod (f(y_i))^{\delta_i} (S(c_i))^{1-\delta_i} \prod_{j=1}^{n_{int}} (f(y_j))^{\delta_j} (S(c_j))^{1-\delta_j}$

I Geral: obs. de cens. pelo vários de vários plos. $L(\theta) = \prod (f(y_i))^{\delta_i} (S(c_i))^{1-\delta_i} \prod_{j=1}^{n_{int}} (f(y_j))^{\delta_j} S(c_j)$

II: o exp. é cont. até se perder, falhas corrig.: n vi, R, r falhas, $(n-r)$ cens. $T_1 \le \ldots \le T_r$

Obs: $T_{r+1} \in T_{r+2} \in \ldots \le T_{r+r}$; $r = 1$, $f_{T_{r+1}}(t) = f_{T_1}(t) = \lim_{\Delta \to 0} [F_{T_1}(t+\Delta) - F_{T_1}(t)] / \Delta = \lim_{\Delta \to 0} P(t < T_1 \le t + \Delta) / \Delta = \lim_{\Delta \to 0} P(1 \text{ obs } T_i : \in (t, t+\Delta] \text{ e } r \text{ obs } T_i > t + \Delta) = n f(t) S(t)^{n-1}$

$r = 2$, $t_1 < t_2$, $f_{T_{r+1}, T_{r+2}}(t_1, t_2) = \lim_{\Delta \to 0} P(1 \text{ obs } T_i : \in (t_1, t_1 + \Delta], 1 \text{ obs } T_i : \in (t_2, t_2 + \Delta]$ e r obs $T_i > t_2 + \Delta_2) = n(n-1) f(t_1) f(t_2) (S(t_2))^{n-2}$. Emg. $f_{T_{r+1} \ldots T_{r+r}}(t_1, \ldots, t_r) = (n!/(n-r)!) (\prod_{i=1}^{r} f(t_i)) (S(t_r))^{n-r} \to L(\theta) \propto (\prod_{i=1}^{r} f(t_i)) (S(t_r))^{n-r}$

$= \prod$ Prog.: n ind. em teste. Após r_1 falhas, $(n - r_1)$ cenados, n_1 do estado, tempo r_2 falhas ocorrem

Sejam $T_{r_1+1}, \ldots, T_{r_1+r_2}$ tempos falh. m^{1^a} etapa e $T_{r_1+1}, \ldots, T_{r_1+r_2}$ ordenados e A_j devem ser e base no dist. cons. $h(t_{r_1+1}, \ldots, t_{r_2}, t_1, \ldots, t_{r_1}) = h_1(t_1, \ldots, t_{r_1}) h_2(t_{r_1+1}, \ldots, t_{r_2} \mid t_1, \ldots, t_{r_1})$

Suponha $r_2 = 1$. Note $f_i = h(t_i \mid t_1, \ldots, t_{r_1}) = \lim_{\Delta \to 0} P(T_i \in (t_i, t_i + \Delta] \mid T_i > t_{r_1}) / \Delta = (n - r_1 - n_1) \times \lim_{\Delta \to 0} [P(T_i \in (t_i, t_i + \Delta] / T_i > t_{r_1}) \times P(T_{2 > t_i + \Delta} / T_2 > t_{r_1}) \times P(T_{n-r_1-n_1} > t_{r_1} + \Delta / T_{n-r_1-n_1} > t_{r_1})] / \Delta = (n - r_1 - n_1) \frac{f(t_i)}{S(t_{r_1})} (S(t_i) / S(t_{r_1}))^{n-r_1-1}$

P/ r_2 genérico, $h_2(t_i, \ldots, t_{r_2} \mid t_1, \ldots, t_{r_1}) = [(n - n_1 - r_2)! / (n - n_1 - r_1 - r_2)!] (\prod_{i=1}^{r_2} f(t_i)) \times (S(t_{r_2}))^{n-n_1-r_1-r_2} \times (1 / S(t_{r_1})^{n-n_1-r_1})$. Assim $L(\theta) \propto \prod_{i=1}^{r_1} f(t_i) S(t_{r_1})^{n-r_1} \times \prod_{i=1}^{r_2} f(t_i) (S(t_{r_2}))^{n-n_1-r_1-r_2}$. **Cens. Al.** As vezes estão seg. o det. eventos p. ñ são de interesse eq. pdem inib. a obs. do ev. de int. T_i: tempo de obito; C_i: tem censura é v.a.

Obs $Z = \min(T_i, C_i)$, $\delta = I(T_i \le C_i)$; Const. f.v. Supos. ind. $T_i \perp C_i$ $\forall i, j$. $\bar{N}$ inf.: Ad. dist de C_i ñ cons. param. de T_i. Precis obter a qta de (Z_i, δ_i). Para $\delta = 1$, $P(Z_i \in (z_i, z_i + \Delta), \delta_i = 1) = P(Z_i \in (z_i, z_i + \Delta), T_i \le C_i) = f(z_i) G(z_i)$ $T \perp C$ C cont. $P(Z_i, \delta_i = 0) = g(z_i) S(z_i) \Rightarrow L(\theta) = \prod_{i=1}^{n} (f(z_i))^{\delta_i} S(z_i)^{1-\delta_i} \prod_{i=1}^{n} g(z_i)^{1-\delta_i} G(z_i)^{\delta_i}$ ou J.

Cens. Int. T cens. à esq. se ocorrer do um t. de censura C_i, mas não se obser. o t. exato $Z = \max(T_i, C_i)$, $\delta = I(T \ge C_i) = \begin{cases} 1, & T \ge C_i \text{ falha} \\ 0, & T < C_i \text{ cens.} \end{cases}$ Prec. achar qta (Z_i, δ_i). P/ $\delta_i = 1$, $P(Z_i, \delta_i = 1) = f(z_i)(1 - G(z_i))$; $P(Z_i, \delta_i = 0) = g(z_i)[1 - S(z_i)] \Rightarrow L(\theta) = \prod (f(z_i))(1 - G(z_i))^{\delta_i} g(z_i)(1 - S(z_i))^{1-\delta_i} \propto \prod f(z_i)^{\delta_i} (1 - S(z_i))^{1-\delta_i}$ \ Pod. ter dado dup. cens.: $Z = \max(\min(T, C_1), C_2)$, $\delta = 1$, $Z = T$; 0 se $Z = C_2$, -1 se $Z = C_1$ \ **Cens. Int.** ocorre qdo sob. q evento ocorre determ. intervalo, mas ñ se sob. mom. ex. Sejam $(a_i, b_i] = I_i$ inst. de cens. $i = 1 - n$. a_i: última inst. conhecido e b_i prim. inst. com ev. Sob $T \in (a_i, b_i]$, $L(\theta) = \prod (S(a_i) - S(b_i))$

Interseções: Ocorre qdo a pop. de int. ñ pod. ser observada. É possível obs. apenas v.a. dentro de um seg. obs. $(Y_L, Y_R]$. V.a. c/ evento fora do seg. ñ é OBS e não há inf. nenhuma sobre elas. $(Y_L, \infty) \Rightarrow$ trun. à esq.; $(0, Y_R)$ à direita \ Vcross. **Trunc. à Dire** T_i: temp. falhas, C_i: t. truncagem. Obs $P(T_i \in (t, t+\Delta] \mid T_i \le C_i) = f(t_i) / [1 - S(c_i)] \Rightarrow L(\theta) = \prod f(t_i) / (1 - S(c_i))$; **Trunc. à Esq** Obs. T_i pq $T_i > C_i$. Tomo que cons. y_i trun. com f.s. dado por $S_{trunc}(t) = P(T_i > t \mid T_i > C_i) = P(T_i > t) / S(C_i) \Rightarrow f_{trunc}(t) = f(t) / S(C_i)$. P/ dados com cens. à esq. y_i à cens. à direita $L(\theta) = \prod f_{tra}(y_i)^{\delta_i} (S_{tra}(y_i))^{1-\delta_i}$ \ **Fi Geral** D: cto índices de ind. cujo evento é obs.; R: com censura à direita; L: à esq.; I: intervalo $\Rightarrow L(\theta) = \prod_{i \in D} f(t_i) \prod_{i \in R} S(c_i) \prod_{i \in L} (1 - S(c_i)) \prod_{i \in I} (S(a_i) - S(b_i))$ \ **Proc. de Contagem**

Cons. $N(t) = \sum (Z_i \le t, \delta_i = 1) = Y(t) = \sum (Z_i > t)$, tempo do todo o inf. **Def**: A filtra $\mathcal{F}_t = \sigma\{X(s) : s \le t\}$ é dita natural de X. **Def.** $\{X(t), t \in T\}$ é adap. à $\mathcal{F}_t$ se $\forall t \ge 0$, $X(t)$ é $\mathcal{F}_t$-mens. **Def** (Proc. Cons.) $\{N(t) : t \ge 0\}$ adap. à $\mathcal{F}_t$ t.q. $N(0) = 0$, $N(t) < \infty$ $\forall t$ q. c. as traj. de $N(t)$ são cont., $N(t)$ é ete com saltos ou dex. de tamanho $+1$. **Def.** O proc. de intensidade assoc. à $N(t)$ é def. $d\Lambda(t) = E(dN(t) / \mathcal{F}_{t^-})$ emque $dN(t) = N(t) - N(t^-)$ e $\mathcal{F}_{t^-} = \sigma(N(u), I(E \le u, \delta = 0), 0 \le u \le t)$ ou $\mathcal{F}_{t^-} = \{N(u), Y(u), 0 \le u \le t\}$. **Def** DP. O acumulado é dado $\Lambda(t) = \int_0^t d\Lambda(t) = \int_0^t \lambda(u)\,du = \int_0^t Y(u) \alpha(u)\,du$ (Form. aprox. ent. saltos em $[0, t]$) **Def.** (Mart.) $\{X(t), t \ge 0\}$ c.d. e $\mathcal{F}_{t^-}$, $t \ge 0$ def. cond. esp. Se X é adap. à $\mathcal{F}_t$; $E|X(t)| < \infty$ $\forall t < \infty$ e $E(X(t+s) / \mathcal{F}_s) = X_s$. Se (1) e (2) bv. X é mart. $\Leftrightarrow E(dX(t) / \mathcal{F}_{t^-}) = 0$ **Def** (Proc. Previs.) X é prev. c.r. à filt. $\mathcal{F}_t$, se p/ cada $t > 0$ val. de X é esp. ou ref. de $\mathcal{F}_{t^-}$ ($X_t \in \mathcal{F}_{t^-}$). **Res.** 1. Todo proc. Cont. é prev. 2. Se X prev. $E(X(t) / \mathcal{F}_{t^-}) = X(t)$. **Ex.** $Y(t) = I(\mathcal{F}_{t^-})$, $d\Lambda(t) = Y(t)\alpha(t)\,dt \to E(d\Lambda(t) / \mathcal{F}_{t^-}) = d\Lambda(t)$. $M(t) = N(t) - \Lambda(t)$ é mart.

Def. (Sahm). (3) $E(X(t+s) / \mathcal{F}_s) \ge X_s$. **Teo.** (D. Doob-Meyer) Seja X um sub c.d. e ñ negativo $\in \mathcal{F}_t$. Então $\exists$ um mart. M c.d. com proc. prev. A t.q. $E(A(t)) < \infty$ e $X(t) = M(t) + A(t)$, $\forall t \ge 0$. Se $A(0) = 0$ e $X = M' + A'$ é outra dec. com $A'(0) = 0$, então $\forall t \ge 0$ $P(M'(t) \ne M(t)) = P(A'(t) \ne A(t)) = 0$. **Cons.** Seja $\{N(t), t \ge 0\}$ um p.c. $\in \{\mathcal{F}_t, t \ge 0\}$ com $E(N(t)) < \infty$ $\forall t$. Então $\exists$ um único proc. $\Lambda(t)$ ñ decr. e cont. (prev. t.q. $\Lambda(0) = 0$, $E(\Lambda(t)) < \infty$ $\forall t$ e $M(t) = N(t) - \Lambda(t)$ é mart. $\Lambda(t)$ é compensador de $N(t)$. **Def.** Proc. de Var. Previsível: P/ um mart. c.d. $\in \{\mathcal{F}_t, t \ge 0\}$ define-se o proc. de var. prev. $\langle M \rangle(t) = \lim_{n \to \infty} \sum_{k=1}^{n} Var(\Delta M_k / \mathcal{F}_{t_{k-1}})$ com q $(0, t)$ é part. em n int. de comp. t/n e $\Delta M_k = M(kt/n) - M((k-1)t/n)$

Obs: Informalmente $d\langle M \rangle(t) = Var(dM(t) / \mathcal{F}_{t^-})$ ou $\langle M \rangle(t) = \int_0^t Var(dM(t) / \mathcal{F}_{t^-})$. $Var(dM(t) / \mathcal{F}_{t^-}) = E(dM(t)^2 / \mathcal{F}_{t^-}) - (E(dM(t) / \mathcal{F}_{t^-}))^2 = E(dM(t)^2 / \mathcal{F}_{t^-})$. **Res.** Seja M um mart. c.d. $\in \{\mathcal{F}_t, t \ge 0\}$ e ass. $E(M^2(t)) < \infty$ $\forall t \ge 0$. Então $\exists$ um único p. prev. $\langle M \rangle(t)$ cen. proc. var. Prev. t.q. $\langle M \rangle(0) = 0$; $E(\langle M \rangle(t)) < \infty$ $\forall t$ e $\{M^2(t) - \langle M \rangle(t), t \ge 0\}$ é mart. c.d. **Obs**: Assumo $M(0) = 0 \Rightarrow E(M^2(t)) = E(\langle M \rangle(t)) = Var(M(t)) = E(\langle M \rangle(t))$. **Ex.** $d\langle M \rangle(t) = Var(dM(t) / \mathcal{F}_{t^-}) = Var(dN(t) - d\Lambda(t) / \mathcal{F}_{t^-}) = Var(dN(t) / \mathcal{F}_{t^-}) = E(dN(t) / \mathcal{F}_{t^-}) - [E(dN(t) / \mathcal{F}_{t^-})]^2 = d\Lambda(t) - (d\Lambda(t))^2$

$d\langle M \rangle(t) \approx d\Lambda(t)$ ou $\langle M \rangle(t) = \Lambda(t)$. **Def.** M_1 e M_2 c.d. e ass. $E(M_i) < \infty$ $\forall t \ge 0$. O proc. de Cov. Prev. é dado $\langle M_1, M_2 \rangle(t) = \lim_{n \to \infty} \sum_{k=1}^{n} Cov(\Delta M_{1k}, \Delta M_{2k} / \mathcal{F}_{(k-1)t/n})$ emque ... e $\Delta_{1k} = M_1(t^k/n) - M_1(t^{(k-1)}/n)$. Podemos ex. $d\langle M_1, M_2 \rangle(t) = Cov(dM_1(t), dM_2(t) / \mathcal{F}_{t^-})$

Res. M_i $(i = 1, 2)$ dois mart. c.d. $\in \{\mathcal{F}_t\}$ e ass. $E(M_i^2(t)) < \infty$ $\forall t \ge 0 \Rightarrow \exists$ pro. prev. c.d. $\langle M_1, M_2 \rangle$ ch. pro. de cov. prev. com $\langle M_1, M_2 \rangle(0) = 0$, $E(\langle M_1, M_2 \rangle(t)) < \infty$ t.q. 1. $\langle M_1, M_2 \rangle$ é dif. de 2 pro. crescentes prev. 2. $M_1 M_2 - \langle M_1, M_2 \rangle$ é mart. **Def.** Proc. de Var. Quadr. P/ um m. c.d. $\in \{\mathcal{F}_t\}$ def. PVQ $[M](t) = \lim_{n \to \infty} \sum_{k=1}^{n} (\Delta M_k)^2$ com $\Delta M_k = M(kt/n) - M((k-1)/n)$ emq. ... Podemos ex. $[M](t) = \sum_{s \le t} (dM(s))^2$ com $\Delta M_k = M(kt/n) - M((k-1)/n)$

Res. $M^2(t) - [M](t)$ é mart. de média zero. 1. $Var(M(t)) = E[M](t)$. **Ex.** $[M](t) = \langle M \rangle(t)$. $[M](t) = \sum_{s \le t} (dM(s))^2 = \sum_{s \le t} (dN(s))^2 = \sum_{s \le t} dN(s) = N(t)$. **Int. Estocástica**: Uma poss. def. é $L(t) = \int_0^t H(s)\,dM(s) = \lim_{n \to \infty} \sum_{k=1}^{n} H_k \Delta M_k$ emque ... com $H_k = H(\frac{k-1}{n})$ e $\Delta M_k = \ldots$, $H(t)$ é um p.a. prev. e $M(t)$ um m. $\in \mathcal{F}_t$. **Res.** $L(t)$ é mart. (inh. cond. geral) $E(dL(t) / \mathcal{F}_{t^-}) = H(t) E(dM(t) / \mathcal{F}_{t^-}) = 0$. Além disso, tem média zero.

Proc. Var. Prev. $L(t)$ $d\langle L \rangle(t) \overset{t \to 0}{=} Var(dL(t) / \mathcal{F}_{t^-}) = H^2(t) Var(dM(t) / \mathcal{F}_{t^-}) = H^2(t)\,d\langle M \rangle(t)$ ou $\langle L \rangle(t) = \int_0^t H^2(s)\,d\langle M \rangle(s)$. P.V.Q. $L(t)$ $[L](t) = \sum_{s \le t} (dL(s))^2 = \sum_{s \le t} (H(s)\,dM(s))^2 = \sum H^2(s)\,dN(s) = \int_0^t H^2(s)\,dN(s)$. **Wiener**: Um p.c. com $W(0) = 0$ é $W(t) = 0$ $\forall t$; $W(t)$ tem inc. ind., o.s., $W(t) - W(s) \perp W(u)$, $0 \le u \le t$ e $W(t) - W(s) \sim N(0, t-s)$ e W tem traj. cont. \ Seja agora $V(t)$ um f. est. proc. com $V(0) = 0$. Consid. o proc. $U(t) = W(V(t)) \Rightarrow U(t)$ tem traj. cont., inc. ind.; $W(V(t)) - W(V(s)) \sim N(0, V(t) - V(s))$. $U(t)$ é dito M.G. com PVP $\langle U \rangle(t) = V(t)$.

P.C.M $\{N_1 \ldots N_k\}$ se cada N_i é p.0 e 2ñ sall. nos mesmos inst. **Teo.** Seja $\{N_1 \ldots N_k\}$ um p.c.m. e Λ_i seus resp. comp. **Def.** $U_i(t) = \sum_{j=1}^{n} \int_0^t H_{ij}\,dM_j(u)$, $i = 1 \ldots k$, H_{ij} p. prev. limitados $\Rightarrow$ i. U_i é mart. 1. $E(U_i(t)) = 0$ $\forall t$, ii) $Cov(U_i(t), U_j(t)) = \sum_{j=1}^{n} \int_0^{t \wedge s} E[H_{ij}(u) H_{j'}(u)\,d\langle \Lambda_j \rangle(u)]$. **T. Deteriorada**. Sejam $\underline{N}(t) = (N_1(t) \ldots N_k(t))$ p.c.m. e $\Lambda(t) = (\Lambda_1(t) \ldots \Lambda_k(t))$ r. comp. e pp. $H_{ij}(t)$, $i, j = 1 \ldots k$, $i = 1 \ldots n$, defina os mart. $U_{ij}^{(n)}(t) = \sum_{j=1}^{n} \int H_{ij}(s) I(|H_{ij}(s)| \ge \epsilon)(dN_i(s) - d\Lambda_i(s))$ e $U_{ij}^{(n)} = \sum_{j=1}^{n} \int_0^t H_{ij}(u)(dN_i(u) - d\Lambda_i(u))$ e $U^{(n)}(t) = (U_1^{(n)}(t) \ldots U_k^{(n)}(t))$ v. ka. L. Cons. asscad.

a) $\langle U_{ij}^{(n)} \rangle(t) \xrightarrow{P} V_{ij}(t)$, $\langle U_{ij}^{(n)}, U_{ij'}^{(n)} \rangle \xrightarrow{P} V_{ij'}(t)$; $V(t)$ m p.d; b) $\langle U_{ij}^{(n)} \rangle(t) \ge 0$, $i = 1 - k \Rightarrow$ Se a e b) val. $U^{(n)}(t) \xrightarrow{D} U^{(\infty)}$ m.G. e $\langle U^{(\infty)} \rangle(t) = V(t)$. Em part. para t_0 fixo, $U^{(n)}(t_0) \xrightarrow{D} N(0, V(t_0))$. **Est N-A.** Seja T_0 tempo de falh. com dens. $f(t)$ e sob $S(t) = G$ f. cens. $T \perp C$. Def. $\delta = I(T \le C)$ e $Z = \min(T, C)$. Assumo q. temos $(Z_i, \delta_i)_{i=1}^{n}$ o.a. Sejam $t_1, t_2 \ldots t_k$ t. de falhas. Para $j = 1 \ldots k$, ch. $d_j = \#$ falhas em t_j e $r_j = \#$ ind. risco em t_j. Int. cut. $\hat{\Lambda}(t_j) = d_j / r_j \Rightarrow \hat{\Lambda}(t) = \sum_{j: t_j \le t} d_j / r_j = \int_0^t d\hat{N}(u) / \hat{Y}(u)$; $\hat{N}(t) = \sum N_i(t)$ # falhas até t. $\hat{Y}(t) = \sum Y_i(t)$ # risco em t.

Prop. Est N-A. Seja $\hat{A}(t) = \int_0^t I(\hat{Y}(u) > 0) \alpha(u)\,du$ e $\hat{\pi}(t) = \hat{Y}(t)/n$. $\hat{A}(t) \le \infty$ $\forall t \ge 0$, temos ① $E(\hat{A}(t) - A(t)) = 0$. ② $E(\hat{A}(t) - A(t))^2 = \int_0^t E(\hat{A}(t) - A(t))^2\,dA(s) \ge -(1 - \pi(t))^n A(t)$. ③ $\sqrt{n}(\hat{A}(t) - A(t)) \xrightarrow{D} N(0, \int_0^t \frac{1}{\pi(s)}\,dA(s))$

Obs: (1) $A^*(t) \ge A(t)$ (2) $E(\hat{A}(t) - A(t)) \ge 0$. Vício EMA tende a neg. $\to 0$ e tem $\sigma^2 = (1 - \pi(t))^n$

Res: $\sigma_n^2(t) \approx Var(\sqrt{n}\hat{A}(t))$. **Dis. Ass. Aalen Ass.** $\exists$ func. $\pi(t)$ t.q. $\hat{\pi}(t) \to \pi(t)$ $\forall t \in [0, \tau)$. $\tau = \sup\{t : \pi(t) > 0\}$. Então para $\pi(t) > 0$ temos

$\int_{\Gamma(u)} dA(u) \Rightarrow \int \cdots \xrightarrow{P} 0$ **Teo** O proc. $F_n(\hat{S}(t)-S(t))$, $0 \le t \le T$, conv. em dist. à um m.G. de média zero e f.c.v. $\sigma^2(t) = \int_0^t \frac{dA(u)}{\pi(u)}$ que pode ser estimado por $\hat{\sigma}^2(t) = \int_0^t dN(u)/\bar{Y}^2(u)/n$. Em part., p/ lo fixado $\sqrt{n}\,(\hat{A}(t_0) - A(t_0)) \xrightarrow{D} N(0, \sigma^2(t_0))$ **Est. K-M.** Seja $(T_i, \delta_i)_{i=1}^n$ uma a.a. de (T, δ), $T = \min(T^*, C)$, $\delta = I(T^* \le C)$. Denote por $t_1 < t_2 < \dots < t_k$ os tempos dist. de falha. O EKM de $S(t)$: $\hat{S}(t) = \prod_{s_i \le t} \frac{r_i - d_i}{r_i} = \prod_{s_i \le t} \left(1 - \frac{dN(t_i)}{\bar{Y}(t_i)}\right)$

Rep. Os int. prod. $S(t) = e^{-A(t)}$. Obs que $S(t) = P(T > t) = P(T > t_1 | T > 0)\, P(T > t_2 | T > t_1) \cdots P(T > t_k | T > t_{k-1}) = \prod_{i=1}^{k} S(t_i | t_{i-1}) = \prod_{i=1}^{k} (1 - dA(t_i)) = \prod_{u \le t} (1 - dA(u))$ Em resumo

Caso Abs. cont.: $S(t) = \prod_{u \le t} (1 - dA(u)) = \exp\{-\int_0^t \alpha(u)du\} = \exp\{-\int_0^t dA(u)\}$

Caso Discreto: $S(t) = \prod_{u \le t} (1 - dA(u)) = \prod_{u \le t} (1 - \alpha(u))$. **Rel. N-A e KM** Vimos que $\hat{A}(t) = \int_0^t dN(u)/\bar{Y}(u) \to dA(t) = dN(t_i)/\bar{Y}(t_i)$. Daí, $\hat{S}(t) = \prod_{u \le t} (1 - d\hat{A}(u)) = \prod_{t_i \le t} (1 - d\hat{A}(t_i)) = \prod_{t_i \le t} (1 - dN(t_i)/\bar{Y}(t_i))$

Propried. $S(t)$. Se $S(t) > 0$: $\hat{S}(t)/S(t) = 1 - \int_0^t [\hat{S}(s^-)/S(s)]\,(dN(s)/\bar{Y}(s) - dA(s))$

Corol. P/ qq ω $S(t) > 0$, $\hat{S}(t) - S(t) = -S(t) \int_0^t [\hat{S}(u^-)/S(u)]\,[I(\bar{Y}(u) > 0)/\bar{Y}(u)]\, dM(u)$

$B(t) = \{(T \wedge t)\, I(T_0 \le t)\, [S(T_0) - \hat{S}(T_0)]/S(T_0)$. $T_0 = \inf\{s : \bar{Y}(s) = 0\}$. $+B(t)$

Resul. Seja $Z(t) = P(Z > t)$, $Z = \min(T, C) \Rightarrow 0 \le t \le (\hat{S}(t) - S(t)) \le (1 - S(t)) \times (1 - \pi(t))^n$ (superior) **Teo.** Seja $t \in [0, T]$ e ass. que $\int_0^t \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} \alpha(u) \xrightarrow{P} 0$. Então $\sup_{u \in [0,t]} |\hat{S}(u) - S(u)| \xrightarrow{P} 0$ consist. uniforme.

Teo. Seja $t \in [0, T]$ e ass. que $\forall s \in [0, T]$, $n \int_0^t \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} \alpha(u) du \to \sigma^2(s)$ e $n \int_0^t \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)} \alpha(u)\, I\left(\left|\sqrt{n}\, \frac{I(\bar{Y}(u) > 0)}{\bar{Y}(u)}\right| > \varepsilon\right) du \xrightarrow{P} 0$, $\sqrt{n} \int_0^t \frac{(1 - I(\bar{Y}(u) > 0))}{\alpha(u)} du \xrightarrow{P} 0$

$\Rightarrow \sqrt{n}\,(\hat{S}(\cdot) - S(\cdot)) \xrightarrow{D}$ proc. G. com f.c.v. $Cov(W(s), W(t)) = S(t)S(s) \int_0^{t \wedge s} dA(u)/\pi(u)$, $\pi(u) = P(Z > u)$. Mais precisa,

$\sqrt{n}\,(\hat{S}(\cdot) - S(\cdot)) \xrightarrow{D} -S(\cdot)\, U(\cdot)$ em $\mathcal{U}$ e m.G. com $U(0) = 0$ e

$Cov(U(s), U(t)) = \int_0^{t \wedge s} dA(u)/\pi(u)$. **Obs.** Para t_0 fixado,

$\sqrt{n}\,(\hat{S}(t_0) - S(t_0)) \xrightarrow{D} N(0, \sigma^2_{S(t_0)})$, $\sigma^2_{S(t_0)} = S^2(t_0) \int_0^{t_0} \frac{dA(u)}{\pi(u)}$.

Seja $\sigma^2(t) = \int_0^t dA(u)/\pi(u)$. Pode est. $\sigma^2(t)$ por $\hat{\sigma}^2(t) = \int_0^t \frac{dN(u)}{\bar{Y}(u)} \frac{1}{\bar{Y}(u)}/n$. Assim $\hat{V}(\hat{S}(t_0)) = \frac{1}{n} \hat{S}^2(t_0) \int_0^{t_0} dN(u)/(\bar{Y}^2(u)/n)$.

Outro est. consistente é lit. Vimos que $Var(\sqrt{n}\,\hat{S}(t)) \to S^2(t) \int_0^t dA(u)/\pi(u)$. Obs q. $dA(t) = d(t)/S(t) = -dS(t)/S(t)$. Assim $Var(\sqrt{n}\,\hat{S}(t)) \to S^2(t) \int_0^t -dS(t)/S(t)\, 1/\pi(u) = Var(\sqrt{n}\,\hat{S}(t)) \to \hat{S}^2(t) \int_0^t -\frac{dS(t)}{S(t)} \frac{1}{\bar{Y}(u)/n}$ (E. Greenwood)

J.C. Pontos p/ int. confiança. Seja $\hat{\sigma}^2(t_0) = \int_0^{t_0} -d\hat{S}(u)/\hat{S}(u)\bar{Y}(u)$. Vimos que $\hat{Var}(\hat{S}(t_0)) = \hat{S}^2(t_0)\, \hat{\sigma}^2(t_0)$. Assim p/ lo fixado, $(1-\alpha)\%$ de conf. IC p/ $S(t_0)$ é dado por $\hat{S}(t_0) \pm z_{1-\alpha/2}\, \hat{\sigma}(t_0)\, \hat{S}(t_0)$ – LI e LS podem extrapolar

Transf. log-log. Para t_0 fixa, def. $\theta_0 = \log(-\log S(t_0)) = g(S(t_0))$ e $\hat{\theta}_0 = \log(-\log \hat{S}(t_0)) = g(\hat{S}(t_0))$. Aplic. o m. delta a $g(\hat{S}(t_0)) = \hat{\theta}_0$: $g(\hat{S}(t_0)) \approx g(S(t_0)) + g'(S(t_0))(\hat{S}(t_0) - S(t_0)) \Rightarrow g(\hat{S}(t_0)) - g(S(t_0)) \approx g'(S(t_0))(\hat{S}(t_0) - S(t_0))$; $Var(\hat{\theta}_0) \approx g'(S(t_0))^2\, Var(\hat{S}(t_0))$.

Nesse caso $g(x) = \log(-\log x) \Rightarrow g'(x) = -1/x \log x \Rightarrow$

$\hat{Var}(\hat{\theta}_0) \approx (-1/[S(t_0)\log S(t_0)])^2\, \hat{\sigma}^2(t_0)\, \hat{S}^2(t_0) \approx \hat{\sigma}^2(t_0)/(\log \hat{S}(t_0))^2$

Assim IC p/ $\theta = g(S(t_0))$

$\log(-\log \hat{S}(t_0)) \pm z_{1-\alpha/2}\, \hat{\sigma}(t_0)/[-\log \hat{S}(t_0)]$, (Omitindo $S(t_0)$ na escala)

$[\exp\{\exp(\hat{\theta}_0 \pm (\frac{z_{1-\alpha/2}\, \hat{\sigma}(t_0)}{-\log \hat{S}(t_0)})\}]$ ou $[\hat{S}(t_0)^{e^{u_L}}, \hat{S}(t_0)^{e^{u_U}}]$,

$u_L = z_{1-\alpha/2} \sqrt{\hat{Var}(\hat{\theta}_0)}$ ou $[\hat{S}(t_0)^{1/c}, \hat{S}(t_0)^{c}]$, $c = \exp\{z_{1-\alpha/2} \sqrt{\hat{Var}(\hat{\theta}_0)}\}$

Bergan e Linstot propõem int. confiança arco-seno. **Estimação Percentis.** Def. definir o perc. p como $t_p = \inf\{t : S(t) \le 1-p\}$. Uma est. p/ t_p é $\hat{t}_p = \inf\{t : \hat{S}(t) \le 1-p\}$ (xi conv.). P/ calcular isso, devemos usar i.?. Enc. t_L e t_U t.q. $\hat{S}(t_L) > 1-p$ e $\hat{S}(t_U) \le 1-p$. Res.: $\frac{t_U - t_L}{\hat{S}(t_L) - \hat{S}(t_U)} \frac{t_p - t_L}{(1-p)\hat{S}(t_L)}$

Test Log Rank. Seja $n = n_1 + n_2$, e cons. $T_1^* < T_2^* < \dots < T_k^*$ (ordenação) algumas falhas ocorrem em um comb. P/ cada inst. T_k^* est. I defina D_{ik}: # de falhas obs. em T_k^* na am. i, $\bar{Y}_{ik}$: # seg. no inst. T_k^* na am. i. Para cada inst. de falha T_k^*, podemos construir a tab.

Grupos	G1	G2	Total
Falhas	D_{1k}	D_{2k}	D_k
Sobrev.	$\bar{Y}_{1k} - D_{1k}$	$\bar{Y}_{2k} - D_{2k}$	$\bar{Y}_k - D_k$
Em risco	$\bar{Y}_{1k}$	$\bar{Y}_{2k}$	$\bar{Y}_k$

Note que: Dado $\bar{Y}_{1k}$, $D_{1k} / \bar{Y}_{1k} \sim Bin(\bar{Y}_{1k}, p_{1k})$; p_{1k}: prob. falha no grupo 1 em T_k^*.

Se H_0 vdd., $p_{1k} = p_{2k}$. Nesta cond., $D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k \sim$ hipergeom.; Daí, temos $E(D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k) = E_{1k} = \bar{Y}_{1k} \frac{D_k}{\bar{Y}_k} = \frac{D_k}{\bar{Y}_k} \bar{Y}_{1k}$.

$Var(D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k) = V_{1k} = D_k \frac{\bar{Y}_{1k}}{\bar{Y}_k} \bar{Y}_{2k} \frac{(\bar{Y}_k - D_k)}{(\bar{Y}_k - 1)}$. Obs. O T.L.R.

É baseado na (dens. de comp.) o obs. com o sup. se ele é vdd. Assim, indo ext. 1 (a), temos est. de teste dado

$LR = [\sum_{i=1}^{k} (D_{1i} - E_{1i})]^2 / [\sum_{i=1}^{k} V_{1i}] \xrightarrow[H_0]{D} \chi^2_{(1)}$

Ex: G_1: 3 5 7+ 9+ 18 $\quad G_2$: 8+ 12 19 20 20+

T_i^*	S_1	S_2	T
3	1	0	1
5	4	2	6
R	3	3	10

T_k^*	Y_{1k}	D_{1k}	$\bar{Y}_k$	E_{1k}	V_{1k}
3	10	1	1	0.5	0.25
5	9	1	1	0.444	0.247
7	8	1	1	0.333	0.222
12	5	1	1	0.2	0.16
18	4	1	1	0.25	0.183
19	3	1	1	0	0
20	2	1	1	0	0

Generaliz. Indução de perda de sg. (S.)

$WLR = [\sum_{i=1}^{k} w_i (D_{1i} - E_{1i})]^2 / [\sum_{i=1}^{k} w_i^2 V_{1i}] \xrightarrow{D} \chi^2_{(1)}$ w_i: pesos.

$w_i = \bar{Y}_i$ – Gehan ou Wilcoxon, $w_i = \sqrt{\bar{Y}_i}$ – Tarone-Ware, $w_i = [\hat{S}(t_{i-1})]^p (1 - \hat{S}(t_{i-1}))^q$ – Fam. Hom. Harrington e Fleming, $w_i = \frac{\bar{Y}_i}{\bar{Y}_{i+1}} \hat{S}(t_{i-1})$ – Peto e Prentice, $\hat{S}(t) = \prod_{t_i \le t} (1 - d_i/\bar{Y}_{i+1})$

Cons. Seja G_n: $n \ge 1$ uma seq. de est. de test. p/ test. H_0. Seja $R_n = R G_n$. Assim a G_n com n sig. α. A seq. G_n é cons. p/ H_0 se $\lim_{n \to \infty} P(G_n \le c_\alpha | H_0) = \alpha$ e H_A é vdd.

Gen p/ k-Grupos. Tr. estes os mais pode ser feito com a distrib. $(D_{1k}, \dots, D_{(k-1)k}) \sim$ hip. multiv.

Test G-Restral. Cor. obs. ressaltados com mod. Suponha int. teste iguald. dos f_i p/ cada um dos níveis (1, ..., k). Queremos testar $H_0: S_1^{(1)} = S_2^{(1)}$ (1 e 2), ..., H_0 vs. H_A não H_0.

T. Reny. Ok, q. dá teste cruz. dos f. risco. Suponha que q. t e t $S_1(t) \le S_2(t)$ $t \in [0, T]$. Considere $T_1^* < T_2^* < \dots < T_k^*$. Para cada T_i^* calc. $Z(T_i^*)$: $\sum_{i=1}^{k} (D_{1i} - D_{1i}\bar{Y}_{1i}/\bar{Y}_i)^2 / \dots$ L. reb. $(Z(T_1^*), \dots, Z(T_k^*))$

A est. de teste é dada por $G = [\sup|Z(t)|, t \le T] / \sigma(t) \xrightarrow{D} \sup|B(t)|$, $t \in [0,1]$

M.R. Par. Assumimos (Z_i, δ_i, X_i) com $Z_i = \min(T_i, C_i)$, $X_i = (1, x_{i1}, \dots, x_{ip-1})^T$. Ass. T_i / X_i ~ f.

M.R. Exp. Podemos ter $f_i(t) = \lambda_i e^{-\lambda_i t}$, $\lambda_i = e^{X_i^T \beta}$, i.e., $f_i(t) = (e^{X_i^T \beta}) e^{-t e^{X_i^T \beta}}$. Nesse $\alpha(t / X_i) = e^{X_i^T \beta}$ (risco prop.). Cut. o prop. $\varepsilon_i \sim Exp(1)$, $T_i = \varepsilon_i / \lambda_i = e^{-X_i^T \beta}$. Def. $Y_i = \ln T_i = -X_i^T \beta + \ln \varepsilon_i = -X_i^T \beta + W_i$, W_i: v.a. perdão. Assegurado, temos que: $T_i / X_i \sim Exp(e^{X_i^T \beta})$, $Y_i = \log T_i = X_i^T \beta + W_i$, $W_i \sim Exp(1, -\beta)$

Def. Um mod. de sob. é dito de vida acel. Se $S(t / X_i) = S_0(\psi_i t)$ $\forall t$, $\psi_i > 0$ com $S_0(\cdot)$ s. f.s. base sol. q. indep. de X_i e ψ_i é fator aceleração.

Est. Mod. Reg. Exp. $Y_i = \ln(T_i) = X_i^T \beta + W_i$. Ass. $\ln(\min(T_i, C_i))$. P/ cens. à dir., $L(\beta) = \prod (f_i(y_i / X_i))^{\delta_i} S(y_i / X_i)^{1-\delta_i} = e^{\sum \delta_i (y_i - X_i^T \beta)} \exp\{-\sum e^{y_i - X_i^T \beta}\} \Rightarrow l(\beta) = \dots$ $\partial l(\beta)/\partial \beta = -\sum \delta_i X_i + \sum e^{y_i - X_i^T \beta}$ – mal num. $(\partial l(\beta)/\partial \beta)^T (\inf(\beta))$

M.R. Weibull. Pode cons. $S_T(t / X) = \exp\{-t^{1/\sigma} e^{-X^T \beta/\sigma}\}$, T. Weibull $p = 1/\sigma$ e $\beta_i = \alpha/\sigma \beta_i$. F.V. $Y_i = \ln T_i = X_i^T \beta + \sigma W_i$. $L(\beta, \sigma) = \exp\{\sum \delta_i (y_i - X_i^T \beta)/\sigma\} \exp\{-\sum e^{(y_i - X_i^T \beta)/\sigma}\}$

Mod. Log-log. – cons. mod. $Y = \ln T = X^T \beta + \sigma W$, $W \sim S(t) = 1/[1 + t^{1/\sigma} e^{-X^T \beta/\sigma}] \Rightarrow$ log-var. $\dots$

Log-Norm. $Y = \log T \sim N(\mu, \sigma)$. $Y = \log T = \mu + \sigma B$, $B \sim N(0,1)$. $Y_i = \log T_i = X_i^T \beta + \sigma \varepsilon_i$, $\varepsilon_i \sim N(0,1)$

Log-Exp. podem ser estim. na forma $Y_i = \log T_i = X_i^T \beta + \sigma \varepsilon_i$, ε_i dist. p/ ε_i

$Var(XY) = E(Var(XY)/X)) + Var(E(XY/X))$

Inf. obs. $J(\hat{\theta}) = -\frac{\partial}{\partial \theta \partial \theta^T} l(\theta)$

$S_W(w) = \exp(-e^{w})$, $f_W(w) = e^{w} \exp(-e^{w})$

Marcos Aparecido Pinheiro | Modelos de Regressão Paramétricos

1. Exponencial
2. Weibull
3. Log. logístico
4. Log. normal

Função da verossim. a partir da relação $Y_i = \ln T_i = x_i'\gamma + \sigma W_i$, $W_i \sim$ YEP

Diagnóstico: Método Descritivo $\sum_{\alpha}$
- Resíduos Padronizados: $\hat{\varepsilon}_i^* = \hat{\varepsilon}^{\delta_i}$ ~ FP
- Resíduos de Cox Snell: $\hat{r}_i = \hat{\Lambda}_m(t/x) \sim Exp(1)$
- Res. Martingal (funcional do cov.)
- Res. Deviance (pontos atípicos)

Interpretação dos parâmetros

1. e 2. $(e^{\beta_i} - 1)\%$ aumento no risco 3. Razão de Chances 4. Razão Medianas

Prop. Assint. dos Estimadores: $Z_i = \min(T_i, C_i)$, $\delta_i = I(T_i \le C_i)$, x_i: v. paravar.

Def.: $N_i(t) = I(Z_i \le t, \delta_i = 1)$, $Y_i(t) = I(Z_i \ge t)$ e considere o processo de intensidade $\Lambda_i(t/x_i) = \int_0^t Y_i(u)\alpha(u/x_i)\,du$. Daí, $L(\theta) = \prod_i [\alpha(z_i/x_i)]^{\delta_i} \times E(z_i/x_i) \Rightarrow U(\theta) = -\partial/\partial\theta\, \ell(\theta) = \sum_i \int_0^\infty \partial/\partial\theta \ln\alpha(u/x_i)$ e $\tau U(\theta)\tau(t) = \sum_i \int_0^t (\partial/\partial\theta \ln\alpha(u/x_i))^{\otimeq 2} Y_i(u)\alpha(u/x_i)\,du$. Segue tbm que $E(I(\theta)) = E(\tau U(\theta)\tau(t)) = Var(U(\theta)(t))$, $I(\theta) = -\partial/\partial\theta\, U(\theta)$

Result. Assumindo vál. das cond. $(1/\sqrt{n})U(\theta_0)(t) \xrightarrow{D} N(0, \Sigma(\theta_0))$ por Slutz-Ky, $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{D} N(0, (\Sigma(\theta_0))^{-1})$.

D. Met. Gráf. Descritivo. No exp., $A(t) = \lambda t$. Estimo $\hat{A}(t)$ e plot. vs t. Se adequado, uma reta pela origem. No Weibull, $\alpha(t) = \rho\theta t^{\rho-1}$, $A(t) = \theta t^\rho$ $\Rightarrow \ln A(t) = \ln\theta + \rho \ln t$, tot $\hat{A}(t)$, faço $\ln(\hat{A}(t))$ e plot. contra t. Se ade. reta com incl. ρ e interc. $\log\theta$. No log-logist. $\alpha(t) = \theta t^{\rho-1}\lambda$, $A(t) = \ln(1+\lambda t^\rho)$ $\Rightarrow \ln(\hat{A}(t) - 1)$ vs t. Na log-normal $S(t) = 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)^{1+\lambda t^\rho}$ $\Rightarrow$ $e^{-A(t)} = 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right) \Rightarrow \Phi^{-1}(1 - e^{-A(t)}) = \frac{\ln t}{\sigma} - \frac{\mu}{\sigma}$.

Res. Padronizados: $s_i = \frac{\ln z_i - x_i'\hat{\gamma}}{\hat{\sigma}}$, calculados a partir de $Y_i = x_i'\gamma + \sigma W_i$, $i = 1, \dots, n$

Então 1º, 2º: s_i, $i = 1, \dots, n$, a.a. censurada extremo padrão 3: logístico padrão 4: normal padrão. Ademais, $\hat{\varepsilon}_i^* = \hat{\varepsilon}^{\delta_i}$. Em 1º, 2º $\hat{\varepsilon}_i \sim exp(1)$, 3 log-logística 4: log-normal. Proc: Obter KM plot $\hat{\varepsilon}_i^*$ e comp. curva teórica: $\hat{S}(\hat{\varepsilon}_i^*)$ vs $S_T(\hat{\varepsilon}_i^*)$

Cox-Snell: $\hat{r}_i = \hat{A}_m(z_i/x_i)$. $\hat{A}_m(\cdot)$ é est. de $A(\cdot)$ com base no modelo 1. $A(t) = \lambda t \Rightarrow t e^{x'\beta} \Rightarrow \hat{r}_i = z_i e^{x_i'\hat{\beta}}$ (p/ com intercepto) 2. $A(t/x) = e^{x'\beta} t^\rho$ $\Rightarrow \hat{r}_i = e^{x_i'\hat{\beta}} t^{\hat{\beta}}$ (β_0 com interc.) 3. $\alpha(t/x) = \frac{\rho t^{\rho-1} e^{x'\beta}}{1 + e^{x'\beta} t^\rho} \Rightarrow \hat{r}_i = -\log\left(\frac{1}{1 + (e^{x_i'\hat{\beta}} z_i)^{\hat{\rho}}}\right)$ $= -\log\left(\frac{1}{1 + (\exp(-x_i'\hat{\gamma}) z_i)^{1/\hat{\sigma}}}\right)$ 4. $\hat{r}_i = -\log(1 - \Phi((\log z_i - x_i'\hat{\gamma})/\hat{\sigma}))$

$Y = A(T) \sim Exp(1) \Rightarrow$ Proc: $\hat{r}_1, \dots, \hat{r}_n$ a.a. censurada: $\hat{A}_r(t)$ vs $A(t) = t$.

Prob: Não forn. inf. sobre a causa mal ajuste e $A(T) \sim exp(1)$ qdo β e α ddeveiro.

Martingal: $M_i = \delta_i - \hat{r}_i$. Deviance: $D_i = \text{Sinal}(M_i)[-2(M_i + \delta_i \log(\delta_i - M_i))]^{1/2}$

Determinações atípicas: traçar vs tempo, por observação

Modelo de Fragilidade: Seja T uma v.a. não negativa. Sup. que cond. à U, $(U > 0)$, T tenha f. risco $\alpha(t/U) = \alpha_0(t) U e^{x'\beta}$, x_i: v. paravariáveis, β: vet. par. desc. e $\alpha_0(t)$ f. risco basal. Observe que podemos definir $V = e^W$ e $\alpha(t/W) = \alpha_0(t) e^W e^{x'\beta}$. A f. de sob. cond. à da da por $S(t/U, x) = \exp\left\{-\int_0^t \alpha_0(s) U e^{x'\beta} ds\right\} = \exp\{-A_0(t) U e^{x'\beta}\}$. $S(t) = \mathbb{P}(T > t) = \int_0^\infty S_{T|U}(s/u) f(u)\,du$. No caso de modelos de riscos proporcionais, $S(t) = \int_0^\infty e^{-(A_0(t) e^{x'\beta} u)} f(u)\,du = \mathcal{L}_U(\cdot)$

Def. Var. neg. Atrib. Laplace é def. $\mathcal{L}_V(s) = E(e^{-sV}) = \int_0^\infty e^{-sv} f(v)\,dv$

Por ex., $S(t) = \mathcal{L}_U(A_0(t) e^{x'\beta}) = \mathcal{L}_U(A(t/x))$, $A(t/x) = A_0(t) e^{x'\beta}$.

Ex.: $U \sim Gama(\alpha, \beta)$, $\mathcal{L}_U(s) = \beta^\alpha / (\beta + s)^\alpha = 1/(1 + s/\beta)^\alpha$

Assumo agora $\alpha(t/x, U) = \alpha_0(t) U e^{x'\beta}$, $U \sim Gama(1/\theta, 1/\theta)$. Nesse caso $EU = \frac{1/\theta}{1/\theta} = 1$ e $Var U = \theta$ e $S(t/x) = \mathcal{L}_U(A(t/x)) = \frac{1}{[1 + \theta A_0(t) e^{x'\beta}]^{1/\theta}}$. A f.d. é $f(t/x) = \frac{(\alpha_0(t) e^{x'\beta})}{[1 + \theta A_0(t) e^{x'\beta}]^{1/\theta + 1}}$

$\alpha(t/x) = f(t/x)/S(t/x) = \alpha_0(t) e^{x'\beta}$.

No caso geral, pode-se dem. que $f(t) = (-\alpha_0(t) e^{x'\beta}) \mathcal{L}'(A_0(t) e^{x'\beta})$

$\alpha(t) = -(\alpha_0(t) e^{x'\beta}) \mathcal{L}'(A_0(t) e^{x'\beta}) / \mathcal{L}(A_0(t) e^{x'\beta})$. Ex. $U \sim$ Inversa Gaus.

$f(u) = \sqrt{\frac{\lambda}{2\pi u^3}} \exp\left\{-\frac{\lambda}{2\mu^2 u}(u - \mu)^2\right\}$, $\lambda > 0$, $\mu > 0$, $u > 0$, $\mathcal{L}(s) = \exp\left\{\frac{\lambda}{\mu}\left[1 - \sqrt{1 + \frac{2\mu^2 s}{\lambda}}\right]\right\}$

$EU = \mathcal{L}'(0) = \mu = 1$, $Var U = \mathcal{L}''(0) - (\mathcal{L}'(0))^2 = \mu^3/\lambda$. $S(t/x) = \mathcal{L}(A_0(t) e^{x'\beta})$

Ex. $U = e^W$, $W \sim N(0, \sigma^2)$. Nesse caso, a h. expressão ant. (Integral) $S(t/x)$. $\alpha(t/x, w) = \alpha_0(t) e^w e^{x'\beta}$. $E(U) = E(e^W) = e^{\mu + \sigma^2/2} = e^{\sigma^2/2} \neq 1$ pois $U \sim$ Gama.

Contra-Verossimilhança: $T_i = \min(T_i^\circ, C_i)$, $\delta_i = I(T_i^\circ \le C_i)$, x_i: cov. Assuma $\alpha(t/x_i, U_i) = \alpha_0(t) e^{x_i'\beta} U_i$, $i = 1, \dots, n$ com $U_i \sim Gama(1/\theta, 1/\theta)$

$L(\beta, \gamma, \theta) = \prod_i (f(t_i/x_i))^{\delta_i} (S(t_i/x_i))^{1-\delta_i} = \prod_i (\alpha(t_i/x_i))^{\delta_i} S(t_i/x_i)$

γ o par. em $\alpha_0(t)$, $f(t/x)$, $\alpha(t/x)$, $S(t/x)$ dist. marg. de T (int. geral).

$L(\beta, \gamma, \theta) = \prod_i \left(\frac{\alpha_0(t_i) e^{x_i'\beta}}{1 + \theta A_0(t) e^{x_i'\beta}}\right)^{\delta_i} \frac{1}{(1 + \theta A_0(t) e^{x_i'\beta})^{1/\theta}}$

Mod. Fragilidade Compartilhada. As n obs. podem ser agrupadas em G grupos com n_i obs. cada, $n_1 + \dots + n_G = n$. Seja T_{ij} o tempo de falha assoc. à j-ésima obs. do grupo i. No model. frag. temos $\alpha(t/x_{ij}, U_i) = \alpha_0(t) U_i e^{x_{ij}'\beta}$, $i = 1, \dots, G$, $j = 1, \dots, n$

Supos.: Dado U_i, T_{ij} são indep. No mod. frag. Gama, $U_i \sim Gama(1/\theta, 1/\theta)$

$EU_i = 1$, $Var U_i = \theta$, $i = 1, \dots, G$.

Coef. τ Kendall: $\tau = E[\text{sinal}((T_{11} - T_{21})(T_{12} - T_{22}))]$ $\begin{matrix}(T_{11}, T_{12})\\(T_{21}, T_{22})\end{matrix}\}$ replicas

Para v.a. cont., $\tau = 2p - 1$, $p = P((T_{11} - T_{21})(T_{12} - T_{22}) > 0)$

Se T_1 e T_2 são ind., $P((T_{11} - T_{21})(T_{12} - T_{22}) > 0) = \underbrace{P(A > 0, B > 0)}_{} + P(A < 0, B < 0)$ $= \underbrace{P(A > 0)P(B > 0)}_{1/4} + \underbrace{P(A < 0)P(B < 0)}_{1/4}$

$\Rightarrow \tau = 0$.

Result. P/ mod. frag. compart. com $\alpha_{ij}(t/x_{ij}, U_i) = \alpha_0(t) U_i e^{x_{ij}'\beta}$, vale o result. $\tau = 4\left(\int_0^\infty s \mathcal{L}''(s) \mathcal{L}(s)\,ds\right) - 1$.

Ex.: $U_i \sim Gama(1/\theta, 1/\theta) \Rightarrow \tau = 4\frac{(1+\theta)}{2(1+2\theta)} - 1 = \theta/(\theta + 2)$ $\Rightarrow$ dep. pos. vs

M.F.G: $T_{ij}^\circ = \min(T_{ij}^\circ, C_{ij})$, $\delta_{ij} = I(T_{ij}^\circ \le C_{ij})$, x_{ij} cov., $j = 1, \dots, n_i$, $i = 1, \dots, G$

modelo: $\alpha_{ij}(t/x_{ij}, U_i) = \alpha_0(t) e^{x_{ij}'\beta} U_i$, $U_i \sim Gama(1/\theta, 1/\theta)$. Sup. Dado U_i, T_{ij}° ind.

$L(\beta) = \prod_i^G L_i(t_{i1}, \dots, t_{in_i}) = \prod_i^G L_i$; $f_i(t_{i1}, \dots, t_{in_i}) = \int_0^\infty f(t_{i1}, \dots, t_{in_i}/u) f_U(u)\,du$

$L_{FULL}(\beta, \theta, A_0) = \prod_i^G \left[\prod_{j=1}^{n_i} (\alpha_{ij}(t_{ij}/x_{ij}, U_i))^{\delta_{ij}} S(t_{ij}/U_i, x_{ij})\right] f_U(U_i)$

No MFG $= \prod_i^G \left[\prod_{j=1}^{n_i} (\alpha_0(t_{ij}) U_i e^{x_{ij}'\beta})^{\delta_{ij}} e^{-A_0(t_{ij}) U_i e^{x_{ij}'\beta}}\right] \frac{(\theta U_i)^{1/\theta} U_i^{\theta-1} e^{-U_i/\theta}}{\Gamma(\theta)}$

F.V. Parcial Penalizada. Therneau e Grambsch prop. f.v. parcial pen.

$PPL(\beta, w, \theta) = \ell(\beta, w) - g(w, \theta)$, $w = (w_1, \dots, w_G)$, $w_i = \log U_i$, $U_i = e^{w_i}$

$\ell(\beta, w) = \sum_i^G \sum_j^{n_i} \delta_{ij}\left(x_{ij}'\beta + w_i - \log\left(\sum_{k \in R(t_{ij})} e^{x_{kj}'\beta + w_k}\right)\right)$, $g(w, \theta)$: função de penalidade. Prop. Thern. $g(w, \theta) = \frac{1}{\theta}\sum_i^G (w_i - e^{w_i})$ offset

Proc. Est. 2 etapas: i. Max. β e w ii. Est. maxim. θ

Algorit. Esp. val. in. $\theta \in w$
1. Obtemos $\hat{\beta}$ com $\ell(\beta, w)$ total
2. Atualiz. w util. expressões em Cotutelis e Ballerini 2
4. (2) e (3) até conv.
3. Atualiz. θ por $\ell_p(\theta)$ e (5) até conv.

$U(\beta, w) = \sum_i \sum_j \delta_{ij}\left(x_{ij} - \frac{\sum_{k \in R(t_{ij})} x_{kj} e^{x_{kj}'\beta + w_k}}{\sum_{k \in R(t_{ij})} e^{x_{kj}'\beta + w_k}}\right)$

$U(w, \beta) = d_i - \sum_i^G \sum_j^{n_i} \delta_{ij}\left(\frac{\sum_{m \in R(t_{ij})} e^{x_{im}'\beta + w_i}}{\sum_{\beta, m \in R(t_{ij})} e^{x_{jm}'\beta + w_j}}\right)$

Mod. Semiparamétrico de Cox. $Z_i = \min(T_i, C_i)$, $\delta_i = I(T_i \le C_i)$, x_i: cov. (sem intercepto). Modelo: $\alpha(t/x) = \underbrace{\alpha_0(t)}_{\text{Não especificada}} \underbrace{e^{x'\beta}}_{\text{paramétrica}}$

tbm conh. mod. riscos proporcionais

Verossim. Parcial de Cox. Suponha que o vetor Y de variáveis possa ser particionado em uma seq. de pares $(V_1, W_1, V_2, W_2, \dots, V_n, W_n)$. Nesse caso, $L(\theta) = f_{V_1, W_1, \dots, V_n, W_n}(v_1, w_1, \dots, v_n, w_n)$ cov. bmente

$L(\theta) = f_{V_1}(v_1) f_{W_1|V_1}(w_1/v_1) f_{V_2|V_1, W_1}(v_2/v_1, w_1) \times f$

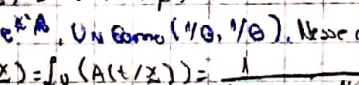
$\times \dots \times f_{V_n|V_1, W_1, V_2, W_2, \dots, V_{n-1}, W_{n-1}}(v_n/v_1, w_1, \dots, v_{n-1}, w_{n-1}) f_{W_n|\dots}(w_n/\dots)$

Temos $L(\theta) = \left[\prod f_{V_i|V_1, W_1, \dots, V_{i-1}, W_{i-1}}(v_i/v_1, w_1, \dots, v_{i-1}, w_{i-1})\right]$

$\times \left[\prod f_{W_i|V_1, W_1, \dots, V_{i-1}, W_{i-1}, V_i}(w_i/\dots)\right]$. Fazendo $P_i = (V_1, W_1, \dots, V_{i-1}, W_{i-1}, V_i)$ e $Q_i = (V_1, W_1, \dots, V_{i-1}, W_{i-1}, V_i, W_i)$

$L(\theta) = \left[\prod f_{V_i|P_i}(v_i/P_i)\right]\left[\prod f_{W_i|Q_i}(w_i/Q_i)\right]$ F.V. Parcial

Verossim. Parc. P/ o Mod. Cox. Obs. (Z_i, δ_i, x_i), $i = 1, \dots, n$. Sejam T_i°, C_i° os inst. de falha. Assuma q. há empates. Seja (j) a obs. que falha em $T_{(j)}^\circ$, $j = 1, \dots, L$. Sejam $T_{j1}^\circ, \dots, T_{jm_j}^\circ$ os inst. de cens. em $[T_{(j)}^\circ, T_{(j+1)}^\circ)$.



Fazemos $V_j = \{T_{(j)}^\circ, T_{(j)+1}^\circ, \dots, T_{(j+1)-1}^\circ\}$ $\{(j-1, j), (j-1, 2), \dots, (j-1, m_{j-1})\}$ e $W_j = \{(j)\}$

A f.v. parcial ficaria dada por $L(\beta) = \prod_{j=1}^L f_{W_j|V_j}(w_j/v_j) = \prod_{j=1}^L P(W_j = (j)/V_j) = \prod_{j=1}^L \frac{e^{x_{(j)}'\beta}}{\sum_{l \in R(t_j)} e^{x_l'\beta}} = \prod_{j=1}^L \frac{e^{x_{(j)}'\beta}}{\sum_{k \in R(t_j)} e^{x_k'\beta}}$, em riscos ent.

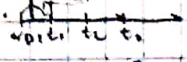
Em suma, $L(\beta) = \prod_{j=1}^L \frac{e^{x_{(j)}'\beta}}{\sum_{l \in R(t_j)} e^{x_l'\beta}}$, $x_{(j)}$: vet. cov. dos obs. que falham em $t_{(j)}$

2. Veros. Perfilada. Temos $\alpha(t/\underline{x}) = \alpha_0(t)e^{\underline{x}'\beta}$ e $S(t/\underline{x}) = e^{-A_0(t)e^{\underline{x}'\beta}}$

$A_0(t) = \int_0^t \alpha_0(u)du \Rightarrow L(\beta, \alpha_0(t_i)) = \prod_{i=1}^{n} \left[(\alpha_0(t_i))^{\delta_i}(e^{\underline{x}_i'\beta})^{\delta_i} e^{-A_0(t_i)e^{\underline{x}_i'\beta}}\right]$

Façamos T_i^*: tempo de falha, $T_i = \min(T_i^*, C_i)$, $\delta_i = I[T_i^* \le C_i]$

Problema: $\alpha_0(t)$ e $A_0(t)$ são f. desconh. Altern: Aprox. por f. de escada (nível).



Fazemos $\alpha_{0j} = \alpha_0(t_j)$, $j=1 \dots L$ (t. dist. de falha) e $A_{0j} = \sum_{t_{(m)} \le t_j} \alpha_{0m}$. Agora temos $L(\beta, \underline{\alpha}_0) = \prod_{i=1}^{n} (\alpha_{0i})^{\delta_i}(e^{\underline{x}_i'\beta})^{\delta_i} \exp(-A_{0i}e^{\underline{x}_i'\beta})$.

$\alpha_{0i} = \alpha(t_{(i)})$; $A_{0i} = A_0(t_{(i)}) = \sum_{\alpha_{0m} \le t_{(i)}} \alpha_{0m}$ e $\underline{\alpha}_0 = (\alpha_{01}, \dots, \alpha_{0L})^T$.

Ideia da F.V. Perf.: Assuma $(\underline{\theta}, \underline{\eta})$ vet. de par., em que $\underline{\theta}$ é o vet. interes. e $\underline{\eta}$ o de incomod. Denot. por $L(\underline{\theta}, \underline{\eta})$ a f.v. completa, a f.v. Perfilada pl $\underline{\theta}$ é dada por $PL(\underline{\theta}) = \sup_{\underline{\eta}} L(\underline{\theta}, \underline{\eta})$.

Normal Cox. $\ell(\beta, \underline{\alpha}_0) = \sum_{i=1}^{n} \delta_i \log \alpha_{0i} + \sum_{i=1}^{n} \delta_i \underline{x}_i'\beta - \sum_{i=1}^{n} A_{0i}e^{\underline{x}_i'\beta}$

$= \sum_{i=1}^{L} \delta_i \log(\alpha_{0i}) + \sum_{i=1}^{n} \delta_i \underline{x}_i'\beta - \sum_{i=1}^{n} \left(\sum_{\alpha_{0j}(t_{(j)}) \le t_{(i)}} \alpha_{0j}\right) e^{\underline{x}_i'\beta}$. Assuma q̃ há empates.

$\frac{\partial}{\partial \alpha_{0k}} \ell(\beta, \underline{\alpha}_{0k}) = \frac{1}{\alpha_{0k}} - \sum_{i \in R(t_{(k)})} e^{\underline{x}_i'\beta}$, $k = 1, \dots, L$. Igual a zero $\Rightarrow$ $\frac{1}{\hat{\alpha}_{0k}} = \sum_{i \in R(t_{(k)})} e^{\underline{x}_i'\beta} \Rightarrow \hat{\alpha}_{0k} = \frac{1}{\sum_{i \in R(t_{(k)})} e^{\underline{x}_i'\beta}}$

Temos então $\hat{A}_{0k} = \sum_{t_{(j)} \le t_{(k)}} \hat{\alpha}_{0k}$.

O próximo passo é sub. α_{0j} por $\hat{\alpha}_{0j}$ na f.v., obt. a perfilada:

$L(\beta) = \left(\prod_{i=1}^{L} \frac{e^{\underline{x}_{(i)}'\beta}}{\sum_{k \in R(t_{(i)})} e^{\underline{x}_k'\beta}}\right) \times e^{-L} \times \prod_{i=1}^{L} \frac{e^{\underline{x}_{(i)}'\beta}}{\sum_{k \in R(t_{(i)})} e^{\underline{x}_k'\beta}}$, a fr. Breslow e Cox.

$\ell(\beta) = \sum_{i=1}^{L} \left[\underline{x}_{(i)}'\beta - \log\left(\sum_{k \in R(t_{(i)})} e^{\underline{x}_k'\beta}\right)\right]$

$\Rightarrow U(\beta) = \sum_{i=1}^{L} \underline{x}_{(i)} - \sum_{i=1}^{L} \left(\sum_{j \in R(t_{(i)})} \underline{x}_j e^{\underline{x}_j'\beta} \Big/ \sum_{j \in R(t_{(i)})} e^{\underline{x}_j'\beta}\right)$

$= \sum_{i=1}^{n} \delta_i \left(\underline{x}_i - \left(\sum_{j \in R(t_i)} \underline{x}_j e^{\underline{x}_j'\beta}\right) \Big/ \left(\sum_{j \in R(t_i)} e^{\underline{x}_j'\beta}\right)\right)$ Newt. Raph

F.V.P. Cox: $L(\beta) = \prod_{j=1}^{L} \left[e^{\underline{x}_{(j)}'\beta} \Big/ \sum_{k \in R(t_{(j)})} e^{\underline{x}_k'\beta}\right]$, $t_{(1)} < \dots < t_{(L)}$ inst. falha.

Em pres. de Empates:

1º Breslow (1974): $t_{(1)} \le t_{(2)} \le \dots \le t_{(D)}$ inst. de falha (pos. ter empates)

$L(\beta) = \prod_{i=1}^{D} \left[e^{\underline{x}_{(i)}'\beta} \Big/ \sum_{j \in R(t_{(i)})} e^{\underline{x}_j'\beta}\right)$. Se denot. $t_{(1)} < t_{(2)} < \dots < t_{(D)}$ os inst. de falha dist. e $d_1, \dots, d_D$ o nº de falhas em cada instante, a

$L(\beta) = \prod_{i=1}^{D} \frac{e^{\underline{x}_{(i)}'\beta} \cdots e^{\underline{x}_{(i+d_i-1)}'\beta}}{\left[\sum_{j \in R(t_{(i)})} e^{\underline{x}_j'\beta}\right]^{d_i}} = \prod_{i=1}^{D} \frac{e^{\underline{s}_i'\beta}}{\left[\sum_{j \in R(t_{(i)})} e^{\underline{x}_j'\beta}\right]^{d_i}}$,

$\underline{s}_i = \sum_{k \in D_i} \underline{x}_k$, D_i = cjto indiv. que falhar. em $t_{(i)}$.

2. EFRON: (+ pq V. exata) $L(\beta) = \prod_{i=1}^{D} \left\{e^{\underline{s}_i'\beta} \Big/ \prod_{j=1}^{d_i} \left[\sum_{k \in R(t_{(i)})} e^{\underline{x}_k'\beta} - \frac{(j-1)}{d_i} \sum_{k \in D_i} e^{\underline{x}_k'\beta}\right]\right\}$

4. Cox (base mod. discreto): $L(\beta) = \prod_{i=1}^{D} e^{\underline{s}_i'\beta} \Big/ \sum_{q \in Q_i} e^{\underline{s}_q'\beta}$, Q_i = cjto de todos subcjtos de $R(t_{(i)})$ com d_i elementos e $\underline{s}_q = \sum_{j \in q} \underline{x}_j$.

Estimação F. Risco Acumulada e F.S.: $\hat{A}_0(t) = \sum_{j: t_{(j)} \le t} \left(d_j \Big/ \sum_{k \in R(t_{(j)})} e^{\underline{x}_k'\beta}\right)$

Com $t_{(1)} < \dots < t_{(D)}$ inst. dist. de falha, d_j: nº de falhas em $t_{(j)}$, $R(t_{(j)})$: vet. ind. em risco em $t_{(j)}$.

$S(t/\underline{x}) = \exp\{-A_0(t)e^{\underline{x}'\beta}\} \Rightarrow S(t/\underline{x}) = (e^{-A_0(t)})^{e^{\underline{x}'\beta}} = (S_0(t))^{e^{\underline{x}'\beta}}$

$\Rightarrow \hat{S}(t/\underline{x}) = (\hat{S}_0(t))^{e^{\underline{x}'\beta}}$, $\hat{S}_0(t) = e^{-\hat{A}_0(t)}$. Tem há outros est. p $A_0(t)$

Extensão M. Cox: Cov. Dep. do Tempo. Observemos $(z_i, \delta_i, \{\underline{x}_i(t): 0 \le t \le z_i\})$

$\underline{x}_i(t)' = (x_{1i}(t), x_{2i}(t), \dots, x_{pi}(t))$. Supr. cont. $\underline{x}_i(t)$ $\forall t$ t.q. $0 \le t \le z_i$.

Ñ vale mais riscos proporcionais (a RR não dep. do tempo). Externas, Internas, Fixamo tempo

F.V.P.: $\alpha(t/\underline{x}(t)) = \alpha_0(t)e^{\underline{x}'(t)\beta}$ $\ell(\beta) = \sum_{i=1}^{n} \int_0^{\tau} (\underline{x}_i(u)\beta - \log \sum_{j=1}^{n} Y_j(u)e^{\underline{x}_j'(u)\beta}) dN_i(u)$

$\ell(\beta) = \sum_{i=1}^{n} \delta_i (\underline{x}_i(z_i)\beta) - \sum_{i=1}^{n} \delta_i \log\left(\sum_{j \in R(t_{(i)})} \exp(\underline{x}_j(z_i)\beta)\right)$ ou ainda

$L(\beta) = \prod_{i=1}^{L} e^{\underline{x}_{(i)}(t_{(i)})\beta} \Big/ \sum_{k \in R(t_{(i)})} e^{\underline{x}_k(t_{(i)})\beta}$. $\hat{\beta}$ é o máximizador.

Prop. assint. $\hat{\beta}$. $U(\beta)$ é mart. ogal. $U(\beta) = \sum_{i=1}^{n} \int_0^{\tau} \left(\underline{x}_i(t) - \frac{S^{(1)}(\beta,t)}{S^{(0)}(\beta,t)}\right) dM_i(t)$

$S^{(0)}(\beta,t) = \sum_{j=1}^{n} Y_j(t)e^{\underline{x}_j'\beta}$, $S^{(1)}(\beta,t) = \sum_{j=1}^{n} Y_j(t)\underline{x}_j(t)e^{\underline{x}_j'\beta}$, $S^{(2)} = \sum_{j=1}^{n} Y_j(t)\underline{x}_j(t)\underline{x}_j'e^{\underline{x}_j'\beta}$,

$V^{(2)} = \underline{x}\underline{x}^T$; $I(\beta) = \int_0^{\tau} V(\beta,t) \sum dN_i(t)$, $V(\beta,t) = S^{(2)}/S^{(0)} - (S^{(1)}/S^{(0)})^2$.

Sob cond. $\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{D} N(0, I^{-1})$, I^{-1} pode ser aproxim. estimado por $(I(\hat{\beta})/n)^{-1}$.

Cox Estratificado: Obs. $(z_i, \delta_i, \underline{x}_i, x_{ki})$, $\underline{x}_i$: p cov. x_k estrat. é dada por $\alpha_j(t/\underline{x}) = \alpha_{0j}(t)e^{\underline{x}'\beta}$, $j = 1, \dots, S$, $\alpha_{0j}(t)$: f.r. basal p/ cada estrato. β: vetor comum par.

F.V. Parcial: $L(\beta) = \prod_{g=1}^{S} L_g(\beta) = \prod_{g=1}^{S} \left(\prod_{k=1}^{L(g)} \left(e^{\underline{x}_{(k)}'\beta} \Big/ \sum_{j \in R_k(g)} e^{\underline{x}_j'\beta}\right)\right)$, em que

$R_k^{(g)}$: ind. em risco em $t_{(k)}^g$, $g = 1, \dots, S$, $k = 1, \dots, L(g)$. $L(g)$: nº inst. falha dist. estrato g.

$\underline{x}_{(k)g}$: vet. cov. obs. falha em $t_{(k)}^g$. $\ell(\beta) = \sum_{g=1}^{S} \ell_g(\beta)$

Possa pensar no modelo $\alpha_g(t/\underline{x}) = \alpha_{0g}(t)\exp(\underline{x}'\beta_g)$ e $\ell(\underline{\beta}) = \sum_{g=1}^{S} \ell(\hat{\beta}_g)$.

$\underline{\beta} = (\beta_1, \dots, \beta_S)$. Suponha que se teste $H_0: \beta_1 = \beta_2 = \dots = \beta_S$ vs H_A: $\exists$ alg dif.

Proc: ajusta (I) e obt. $\hat{\beta}$ e $\ell_I(\hat{\beta})$. Ajusta (II) e obt. $\ell_g(\hat{\beta}_g)$ e $\ell_{II}(\hat{\underline{\beta}}) = \sum_{g=1}^{S} \ell_g(\hat{\beta}_g)$.

Calcula $\Lambda = -2(\ell_I(\hat{\beta}) - \ell_{II}(\hat{\underline{\beta}})) \xrightarrow[\text{sob } H_0]{D} \chi^2_{(S-1)p}$. Teste Wald: $H_0: \beta_{gk} = \beta_{hk}$, $1 \le g$ vs $H_A: \beta_{gk} \ne \beta_{hk}$. $\hat{\beta}_g \overset{as}{\sim} N(\beta_g, I^{-1}(\beta_g))$, $\hat{\beta}_{gk} \sim N(\beta_{gk}, V_{gk})$

$Z_k = (\hat{\beta}_{gk} - \hat{\beta}_{hk}) / \sqrt{V_{gk} + V_{hk}} \overset{as}{\sim} N(0,1)$. Podemos ter ainda $\alpha(t/\underline{x})$:

$\alpha_{0g}(t)\exp(\underline{x}_c'\beta_c + \underline{x}_i'\beta_g)$, $\underline{\beta}^T = (\beta_c', \beta_1, \beta_2, \dots, \beta_S)$, β_c: parte fixa.

$S_g(t/\underline{x}) = \exp(-A_{0g}(t)e^{\underline{x}'\beta}) = (S_{0g}(t))^{e^{\underline{x}_c'\beta_c + \underline{x}_i'\beta_g}} \Rightarrow \hat{S}_g(t) = (\hat{S}_{0g}(t))^{e^{\underline{x}_c'\beta_c}}$.

$S_{0g}(t) = \exp\left\{-\sum_{k: t_{(k)}^g \le t} \left(\frac{d_{(k)}^g}{\sum_{j \in R_k(g)} e^{\underline{x}_j'\beta}}\right)\right\}$, $d_{(k)}^g$: nº falhas estrato g em $t_{(k)}^g$, $t_{(1)}^g < \dots < t_{(L)}^g$

Adequação Modelo Cox. (Aval. geral ajuste; Verif. sup. prop. dos riscos; Detec. pto atip.)

Método Graf. Sup. Prop. Riscos. (A) Descritivo: $S(t/\underline{x}) = e^{-A_0(t)e^{\underline{x}'\beta}} \Rightarrow$

$\log(-\log S(t/\underline{x})) = \underline{x}'\beta + \log(A_0(t))$. Proc. Obter KM p/ cada grupo das combinações das cov. (cov. cat.) do $A(t)$ e $N(t)$. Calculamos $\log(-\log \hat{S}_g(t))$, $g = 1, \dots$. Construímos em um mesmo gráfico versus tempo. Se sup. for válida, as curvas devem ser ≈ paralelas

B. (Esperados vs Observados) $\hat{S}_g^{KM}(t)$ e $\hat{S}(t/\underline{x}=g)^{\exp(\underline{x}'\hat{\beta})}$ no mesmo gráfico, devem estar

C. (Estratificação) $\underline{x} = (\underline{x}_1, \underline{x}_2)'$. Categorizo $\underline{x}_1$ e ajusta modelo Cox estrati. por $\underline{x}_1$ com cov. $\underline{x}_2$. Seja $\hat{\alpha}_g(t)$ a f.r. basal est. m. para o estrato g. Se vale a sup. prop. p/ $\underline{x}_2$, $A(t/\underline{x}) = A_0(t)e^{\underline{x}_1'\beta_1}e^{\underline{x}_2'\beta_2}$, $\beta = (\beta_1, \beta_2)'$ (no estratif. não vale se for igual ao ñ estratif.) $\log \hat{A}_{0g}(t) \approx \beta_{0g} + \log A_0(t)$. $\log(\hat{A}_{0g}(t))$ vs t p/ cada

Resid. no Mod. Cox. Cox-Snell: $\hat{\varepsilon}_i = \hat{A}_0(z_i)e^{\underline{x}_i'\hat{\beta}}$, $A_0(T_i)e^{\underline{x}_i'\beta} \sim Exp(1)$

pois $A(T) = A(T/\underline{x}) = A_0(t)e^{\underline{x}_i'\beta} = V \sim Exp(1)$. Semelhante para martingal.

Proc. $(\hat{\varepsilon}_i, \delta_i) - (\hat{\varepsilon}_0, S_0)$ $\hat{A}_0(\varepsilon_i)$ vs ε_i. Obs: No estr. $\hat{\varepsilon}_j = \hat{A}_{0g}(z_j)e^{\underline{x}_j'\hat{\beta}}$, $j \in g$.

Dep. do tempo, $\hat{\varepsilon}_j = \hat{A}_0(z_j)e^{\hat{\alpha}_j(z_j)\hat{\beta}}$. Resíduo Martingal: $\hat{M}_i = \delta_i - \int_0^{\infty} Y_i(u)e^{\underline{x}_i'(u)\hat{\beta}} d\hat{A}_0(u)$. No caso sem cov. dep. tempo, $\hat{M}_i = \delta_i - \hat{\varepsilon}_i$,

$\hat{\varepsilon}_i = \hat{A}_0(z_i)e^{\underline{x}_i'\hat{\beta}}$ (res. Cox-Snell). Proc: Ajusta o MCox sem cov. em estudo e obter resíduo martingal deste modelo. Construir o graf. $\hat{M}_i$ vs X_{11} (forma funcional)

Deviance: Detec. Pts Atip. $\hat{D}_i = \text{sinal}(\hat{M}_i)[-2(\hat{M}_i + \delta_i \log(\delta_i - \hat{M}_i))]^{1/2}$.

Proced: Construção $\hat{D}_i$ versus $\underline{x}_i'\hat{\beta}$ (preditor linear). Escore: Relembre que $\ell(\beta) = \int_0^{\tau} \sum_{i=1}^{n} \left[\underline{x}_i'(u)\beta - \log\left(\sum_{j=1}^{n} Y_j(u)e^{\underline{x}_j'(u)\beta}\right)\right] dN_i(u)$, de modo

$U(\beta) = \int_0^{\tau} \sum_{i=1}^{n} [\underline{x}_i(u) - \bar{\underline{x}}(\beta, u)] dN_i(u)$, $\bar{\underline{x}}(\beta, u) = \frac{\sum_{j=1}^{n} Y_j(u)\underline{x}_j e^{\underline{x}_j'\beta}}{\sum_{j=1}^{n} Y_j(u)e^{\underline{x}_j'\beta}}$.

Vimos que $U(\beta_0) = \int_0^{\tau} \sum (\underline{x}_i(u) - \bar{\underline{x}}(\beta_0, u)) dM_i(u)$. O P. Escore p/ o resíduo ind. é def. por $U_i(\beta, \tau) = \int_0^{\tau} (\underline{x}_i(u) - \bar{\underline{x}}(\beta, u)) dN_i(u)$.

O Resíduo Escore dor. ind. p/ a k-ésima cov. é def. por

$S_{ik}(t) = \int_0^{t} (X_{ik}(u) - \bar{X}_k(u)) d\hat{M}_i(u)$, $\bar{X}_k(u) = \frac{\sum_{j=1}^{n} Y_j(u)X_{jk}(u)e^{\underline{x}_j'\hat{\beta}}}{\sum_{j=1}^{n} Y_j(u)e^{\underline{x}_j'\hat{\beta}}}$.

Prop. $\sum_{i=1}^{n} S_{ik}(\tau) = 0$. Proced: $W_k(t)$ vs t, $W_k(t) = U_k(t)$ / erro padrão (β_k).

$U_k(t) = \sum_{i=1}^{n} S_{ik}(t) = \sum_{j: \text{falha} \le t} (X_{jk} - \bar{X}_k(z_j))$. Se a mod. adequado, deve se comportar ponte browniana. Proc: $I^{-1}(\hat{\beta}) \begin{pmatrix} S_{i1} \\ \vdots \\ S_{ip} \end{pmatrix}$ vs observação $(\hat{\beta} - \hat{\beta}_{(i)}) \approx I^{-1}(\hat{\beta})(S_{i1}, \dots, S_{ip})^T$ → inf. perciv. no fim p/ S_{ip} padronizado

Resíduo Schoenfield.

Def. apenas p/ obs que falharam: $r_{ik} = \delta_i(x_{ik} - \bar{x}_k(z_i))$, $\bar{x}_k(z_i) = \left(\sum_{j \in R(z_i)} x_{jk} e^{\underline{x}_j'\hat{\beta}}\right) \Big/ \sum_{j \in R(z_i)} e^{\underline{x}_j'\hat{\beta}}$. Um ex. p/ ilustrar que

falhas. $\sum \underline{r}_i = \underline{0}$, $\underline{r}_i = (r_{i1}, \dots, r_{ip})^T$. Podem ser padronizados:

$\underline{r}_i^* = [\widehat{Var}(\underline{r}_i)]^{-1}\underline{r}_i$, $Cov(\underline{r}_i) \approx I(\hat{\beta})/m$, m: total de falhas. Avaliação

Proporcional. dos. Motivação: $\underline{x}_i = (t_i, \underline{x}^T)$, $\alpha(t/\underline{x}) = \alpha_0(t)\exp[\underline{x}'\beta + \underline{x}_1(\beta_1 + \gamma_1 g(t))]$

é adequado, $g(t)$ função conhecida do tempo. Este é RRop. só se $\gamma_1 = 0$. Suponha agora que ajusta Cox para este modelo: $\alpha(t/\underline{x}) = \alpha_0(t)e^{\underline{x}_1'\beta_1 + \underline{x}'\beta}$. Seja $\hat{r}_i^*$: resíd. pad. Sch.

Regr. $E(r_{ik}^* + \hat{\beta}_k) = (\beta_k + \gamma_k g(t))$ ou $E(s_{ik}^*) = \beta_k(t)$, $s_{ik}^* = r_{ik}^* + \hat{\beta}_k$

o resíd. reescalados. $E(r_{ik}^* + \hat{\beta}_k) = \beta_k(t)$. Proc. s_{ik}^* vs t. Teste baseado

Res. Schoenf. H_0: β não dep. tempo vs H_A: $\beta_g(t) = \beta_g + \theta_g g(t)$, p/ pelo menos um g.

$H_0 = \underline{0} = (0, \dots, 0)^T$ vs $\underline{0}$. Est. teste: $U = (\underline{g} - \bar{\underline{g}})^T_{1 \times m} R^*_{m \times p} I_{p \times p} R^{*T}_{p \times m} (\underline{g} - \bar{\underline{g}})_{m \times 1}$

$\underline{g} = (g(t_1), \dots, g(t_m))$, $\bar{\underline{g}} = \sum_{m}^{m} \underline{g}(t_i)$, R^*: matriz dos resíd. pad. de Schoenf. $m \sum_{i=1}^{m} (g_i - \bar{g})^2$

$\alpha(t) = \lim_{\Delta t \to 0} \frac{P(t \le T \le t+\Delta t / T > t)}{\Delta t} = -\frac{d}{dt} \ln S(t)$; $S(t) = e^{-A(t)}$

$A(t) = \int_0^t \alpha(s)\,ds$; $f(t) = \alpha(t) S(t) = \alpha(t) e^{-A(t)}$

No c. discreto, Tassume val. $t_j, j \ge 1$, $0 \le t_1 \le t_2 \le t_3 \dots$ $S(t) = \sum_{j: t_j > t} p_j$

$p_j = P(T = t_j)$

$\alpha_j = \alpha(t_j) = P(T = t_j / T \ge t_j) = p_j / S(t_{j-1}) = (S(t_{j-1}) - S(t_j)) / S(t_{j-1})$

$= 1 - S(t_j)/S(t_{j-1})$. Para t qqer $t_j \le t < t_{j+1}$,

$S(t) = S(t_j) = \frac{S(t_j)}{S(t_{j-1})} \frac{S(t_{j-1})}{S(t_{j-2})} \dots \frac{S(t_1)}{S(t_0)} = \prod_{j: t_j \le t} (1 - \alpha_j)$

$A(t) = \sum_{t_j \le t} \alpha(t_j)$ aux. acumulado $-\sum_{j \le t} \ln(1 - \alpha(t_j))$

Vida média: $t_{0.5}$: $S(t_{0.5}) = 0.5$ e $F(t_{0.5}) = 0.5$.

t_p : $S(t_p) = 1-p$ e $F(t_p) = p$ / $t_p = \inf\{t : S(t) \le 1-p\}$

Tempo de vida m. residual: $vmr(t) = E[T - t / T > t] = \frac{\int_t^{+\infty} (u-t) f(u)\,du}{S(t)}$

$= \frac{\int_t^{+\infty} S(u)\,du}{S(t)}$ / $vmr(t) = \sum_{t_j > t} \frac{(t_j - t) P(T = t_j)}{P(T > t)} = \frac{1}{S(t)} \left[(t_k - t) S(t_k) + \sum_{j > k} (t_{j+1} - t_j) S(t_j) \right]$

Cens
- à direita → {Tipo I, Tipo II, Aleatória}
- à esquerda
- intervalar

Tem há truncagem, qdo obs. são desc. por não fazer parte do estudo

Sabe-se que o ev. do inst. ocorreu após algum instante
" " antes
" " entre uma janela observável

VI: ev. obs. se t ≤ c ocorre antes do f. do estudo

$L(\theta) = \prod (f(z_i))^{\delta_i} (S(c_i))^{1-\delta_i}$

I: Proj. o estudo tem r de um inst de censura

$L(\theta) = \prod^{r_1} (f(z_i))^{\delta_i} (S(c_i))^{1-\delta_i} \prod_{i=r_1+1}^{r_2} (f(z_i))^{\delta_i} (S(c_i))^{1-\delta_i}$

I. Generalizada: O t de cens. pode ser de v.a. p/ v.a. mas fixo a tal.

$L(\theta) = \prod (f(z_i))^{\delta_i} (S(c_i))^{1-\delta_i} \overset{obs}{=} \prod (\alpha(z_i))^{\delta_i} S(c_i)$

Tipo II: O exp. é conduzido até que um nº pré-det. de falhas ocorra.

Gen. n v.i., r falhas, (n-r) cens. $T_1, \dots, T_n$ t. d. falha

$T_{(1)} \le T_{(2)} \le \dots \le T_{(r)}$ est. ordem. Obs. r pr. est. ordem

$r = 1 \to T_{(1)} = \min(T_1, \dots, T_n)$. $f_{T_{(1)}}(t) = f_{T_{(1)}}(t) = \lim_{\Delta \to 0} \frac{F_{T_{(1)}}(t+\Delta) - F(t)}{\Delta}$ ← fim

$= \lim_{\Delta \to 0} P(t \le T \le t+\Delta)/\Delta = \lim_{\Delta \to 0} \frac{P(1 \text{ dos } T_i\text{'s} \in (t, t+\Delta] \text{ e os } T_i > t)}{\Delta}$

$= n \lim_{\Delta \to 0} P(T_1 \in (t, t+\Delta]) (S(t+\Delta))^{n-1} \Rightarrow f_{T_{(1)}}(t) = n f(t) S(t)^{n-1}$

$r = 2$ $t_1 \le t_2$

$f_{T_{(1)}T_{(2)}}(t_1, t_2) = \lim_{\substack{\Delta_1 \to 0 \\ \Delta_2 \to 0}} \frac{1}{\Delta_1 \Delta_2} P\left(\begin{array}{l} 1 \text{ dos } T_i\text{'s} \in (t_1, t_1+\Delta_1], 1 \text{ dos } T_i\text{'s} \in (t_2, t_2+\Delta_2] \\ \text{os } T_i\text{'s} > t_2 + \Delta_2 \end{array} \right)$

$= n(n-1) P(T_1 \in (t_1, t_1+\Delta_1]) P(T_2 \in (t_2, t_2+\Delta_2]) P(T_3 > t_2 + \Delta_2) \dots P(T_n > t_2 + \Delta_2)$

$= G(n-1) f(t_1) f(t_2) (S(t_2))^{n-2}$

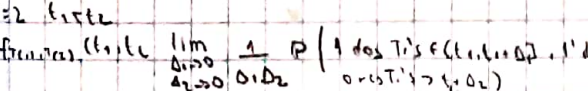
$f_{T_{(1)} \dots T_{(r)}}(t_1, \dots, t_r) = \frac{n!}{(n-r)!} \left(\prod_{i=1}^{r} f(t_i)\right) (S(t_r))^{n-r}$

$L(\theta) = \frac{n!}{(n-r)!} \left(\prod f(t_i)\right) (S(t_r))^{n-r} \propto \left(\prod_{i=1}^{r} f(t_i)\right) (S(t_r))^{n-r}$

Tipo III Progressiva: n v.i. contínuo. Após r_1 falhas, $(n-r_1)$ v.i. c/ mais obs. Dessas $(n-r_1)$, retira n_1 do estudo (censura n_1) → Resta o exp. $(n-r_1-n_1)$ v.i. acomp. até r_2 falhas acontecer com n falhas, $r_1 < r_2$

Sejam $T_{(1)}, \dots, T_{(r_1)}$ tempos de falhas na 1ª fase do experimento

$T_{(1)}^*, \dots, T_{(r_2)}^*$ " " " " " " " "



A f. de verossim. é bas. na dist. cond. $h(t_1, t_2, \dots, t_{r_1}, t_1^*, \dots, t_{r_2}^*)$

$= h(t_1, \dots, t_{r_1}) h_2(t_1^*, \dots, t_{r_2}^* / t_1, \dots, t_{r_1})$

Mas, ord.

$h(t_1, \dots, t_{r_1}) = \frac{n!}{(n-r_1)!} \left[\prod_1^{r_1} f(t_i)\right] (S(t_{r_1}))^{n-r_1}$

Suponha inic. $r_2 = 1$. Note $t_i^* > t_{r_1}$.

$h_2(t_i^* / t_1, \dots, t_{r_1}) = \lim_{\Delta \to 0} P[T_{(i)}^* \in (t_i^*, t_i^* + \Delta] / T_{(i)} > t_{r_1}] / \Delta$

$= (n - r_1 - n_1) \times \lim_{\Delta \to 0} \frac{1}{\Delta} \left[\begin{array}{l} P(T_i^* \in (t_i^*, t_i^* + \Delta] / T_i^* > t_{r_1}) \\ \times P(T_2^* > t_1 + \Delta / T_2 > t_{r_1}) \\ \times P(T_{n-r_1-n_1}^* > t_1 + \Delta / T_{n-r_1-n_1}^* > t_{r_1}) \end{array} \right]$

$= (n - r_1 - n_1) \lim_{\Delta \to 0} \frac{P(T_i^* \in (t_i^*, t_i^* + \Delta])}{\Delta S(t_{r_1})} \left(\frac{S(t_i^* + \Delta)}{S(t_{r_1})}\right)^{n-r_1-n_1-1}$

$= (n - r_1 - n_1) \frac{f(t_i^*)}{S(t_{r_1})} \left(\frac{S(t_i^*)}{S(t_{r_1})}\right)^{n-r_1-n_1-1}$

P/ r_2 genérico

$h_2(t_1^*, \dots, t_{r_2}^* / t_1, \dots, t_{r_2}) = \frac{(n - n_1 - r_1)!}{(n - n_1 - r_1 - r_2)!} \left(\prod_{i=1}^{r_2} f(t_i^*)\right) (S(t_{r_2}))^{n-n_1-r_1-r_2} \times \frac{1}{S(t_{r_1})^{n-n_1-r_1}}$

Assim

$L(\theta) \propto \left(\prod_{i=1}^{r_1} f(t_i)\right) (S(t_{r_1}))^{n_1} \left(\prod_{i=1}^{r_2} f(t_i^*)\right) (S(t_{r_2}))^{n-n_1-r_1-r_2}$

Censura Aleatória: As v.a. estão sujeitas a outros eventos q. não são de interesse e que podem inibir a obs. do ev. de interesse

T_i tempo óbito. C_i tempo censura é v.a.

Obs. $Z_i = \min(T_i, C_i)$, $\delta = I(T_i \le C_i)$

Construção da f.v. Suposições: 1- Ind. $T_i \perp C_i$ $\forall i$. 2. Não inf. A dist de C_i não depende param. da dens. de T_i. Precisa obter a dist cond. de (Z_i, δ_i)

Para $\delta = 1$, $P(z_i, \delta_i = 1) = P(Z_i \in (z_i, z_i + \Delta], \delta_i = 1)$ Int. f, S

$= P(Z_i \in (z_i, z_i + \Delta], T_i \le C_i)$ Cong. G

$= P(T_i \in (z_i, z_i + \Delta], C_i > z_i + \Delta) = f(z_i) G(z_i)$

P/ $\delta = 0$, $P(z_i, \delta_i = 0) = g(z_i) S(z_i)$

Daí,

$L(\theta) = \prod^n (f(z_i) G(z_i))^{\delta_i} (g(z_i) S(z_i))^{1-\delta_i} = \left(\prod f(z_i)^{\delta_i} G(z_i)^{1-\delta_i}\right) \left(\prod g(z_i)^{1-\delta_i} G(z_i)^{\delta_i}\right)$ ← não inf. sobre p. de T.

$\Rightarrow L(\theta) \propto \prod_{i=1}^n f(z_i)^{\delta_i} S(z_i)^{1-\delta_i}$

Censura à Esquerda: Um t do sob. T é considerado censurado à esq. se é menor do que um tempo de censura C_L, porém não se observa o inst. exato. T: tempo, C_L t de censura

$Z = \max(T, C_L)$, $\delta = I(T > C_L) = \begin{cases} 1, & T > C_L \text{ (not obs.)} \\ 0, & T \le C_L \text{ (censura)} \end{cases}$

Construção da f. Versos. $T_i \sim f(\cdot), S(\cdot)$; $C_i \sim g, G$, $T_i \perp C_i$, são inf.

Prec. cong. (Z_i, δ_i)

P/ $\delta_i = 1$, $P(z_i, \delta_i = 1) = P(Z_i \in (z_i, z_i + \Delta], T_i > C_i) = P(T_i \in (z_i, z_i + \Delta], C_i < T_i) \overset{ind}{=} f(z_i)(1 - G(z_i))$

P/ $\delta = 0$, $P(z_i, \delta_i = 0) = P(Z_i \in (z_i, z_i + \Delta], \delta_i = 0) = P(C_i \in (z_i, z_i + \Delta], T_i \le C_i)$

$= P(C_i \in (z_i, z_i + \Delta]) P(T_i \le C_i) = g(z_i)(1 - S(z_i))$

$L(\theta) = \prod^n (f(z_i)(1 - G(z_i)))^{\delta_i} (g(z_i)(1 - S(z_i)))^{1-\delta_i}$

$\propto \prod f(z_i)^{\delta_i} (1 - S(z_i))^{1-\delta_i}$

Obs. Podemos ter dados dup cens.: $Z = \max(\min(T, C_R), C_L)$

$\delta = 1$, se $Z = T$; 0, se $Z = C_R$; -1 se $Z = C_L$

Censura Intervalar: ocorre qdo sabemos que evento ocorreu em um intervalo, mas não se sabe o mom. exato

Sejam $(a_i, b_i] = I_i$, p. int. de cens., $i = 1, 2, \dots, n$. a_i: último inst. sem evento, b_i: prim. inst. com evento. Sob $T \in (a_i, b_i]$, I. int. que ocorre

$P(T \in (a_i, b_i]) = S(a_i) - S(b_i) = F(b_i) - F(a_i)$

$L(\theta) = \prod (S(a_i) - S(b_i))$

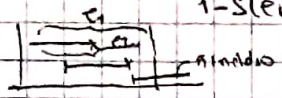
Truncagem: Ocorre qdo a pop. de interesse não pode ser observada. É possível obs. apenas v.a. que exp. o evento dentro de uma janela observável (Y_L, Y_R). V.a. somente fora dessa janela são obs. e não há inf. nenhuma sobre elas

$(Y_L, +\infty)$ → truncagem esq. $(0, Y_R)$ à direita

Hipótese: Trunc à direita: T_i: t. de falha; e_i: tempo de trunc.

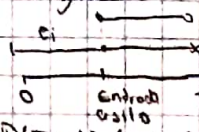
Obs: portanto $P(T_i \in (t_i, t_i+\Delta] \mid T_i \le e_i) = \frac{f(t_i)}{1-S(e_i)}$

$L(\theta) = \prod_i f(t_i)/(1-S(e_i))$



Trun. à Esq. T_i: t. falha, e_i: tempo truncagem

Obs: T_i pq $T_i > e_i$. Temos que cons. a v. trunc. com f.s. dada por



$S_{TRUNC}(t) = P(T_i > t \mid T_i > e_i) = P(T_i > t)/S(e_i)$

$f_{TRUNC}(t) = f(t)/S(e_i)$

Dl dados contrunc à esq suj. a cens. à direita,

$L(\theta) = \prod f_{TR}(z_i)^{\delta_i} (S_{TR}(z_i))^{1-\delta_i} = \prod \left(\frac{f(z_i)}{S(e_i)}\right)^{\delta_i} \left(\frac{S(z_i)}{S(e_i)}\right)^{1-\delta_i}$

Forma geral: D: cjto de ind. de ind. cujo ev. é observado

R: " " " " cens. à direita

L: " " " à esquerda

I: " " int.

$L(\theta) = \prod_{i \in D} f(t_i) \prod_{i \in R} S(t_i) \prod_{i \in L} (1-S(e_i)) \prod_{i \in I} (S(a_i) - S(b_i))$

Proc. de cont. Com $N(t) = \mathbb{1}(Z \le t, \delta_i = 1)$ e $Y(t) = \mathbb{1}(Z_i \ge t)$ + temos toda a inf.

Def: A fil $\mathcal{H}_t = \sigma\{X(s) : s \le t\}$ é cha natural ou hist de X.

Def: $\{X(t) : t \in T\}$ é Adapt à $\mathcal{F}_t$ se $\forall t \ge 0$ $X(t)$ é $\mathcal{F}_t$-mens.

Def. P. Cond. $\{N(t), t \ge 0\}$ adap à $\mathcal{F}_t$ t.q. $N(0) = 0$, $N(t) < +\infty$ $\forall t$ q.c., As trajs. de $N(t)$ são cont. à dir., $N(t)$ é de consoltos ou dest de tam = 1

Def. O Proc de Int associ. ad à $N(t)$ é def $d\Lambda(t) = E(dN(t) \mid \mathcal{F}_{t^-})$ em que $dN(t) = N(t) - N(t^-)$ e $\mathcal{F}_t = \sigma\{N(u), I(z \le u, \delta = 0), 0 \le u \le t\}$ ou $\mathcal{F}_t = \{\mathcal{H}(u), Y(u)\}$ ou

Def. O proc. int acumulado é dado $\Lambda(t) = \int_0^t d\Lambda(t) = \int_0^t \lambda(u)\,du =$
$= \int_0^t Y(u)\alpha(u)\,du$ (Form aprox. ao esp de saltos em $[0,t]$)

Def (Martg.) $\{X(t) : t \ge 0\}$ é e. cont. à dir. t. $\mathcal{F}_t$ $t \ge 0$ def mon. esp. prob se X é adap à $\mathcal{F}_t$; $E|X(t)| < +\infty$ $\forall t < +\infty$ e $E(X(t+s) \mid \mathcal{F}_t) = X_s$, X martg. Se (1) e (2) são vál, X é mart $\Leftrightarrow E(dX(t) \mid \mathcal{F}_{t^-}) = 0$.

Def. (Proc Previsível) X é prev. com rel. à filtra $\mathcal{F}_t$ se, p cada t, o valor de X é especificado ou é uma f. de $\mathcal{F}_{t^-}$ ($X(t) \in \mathcal{F}_{t^-}$)

Result. Todo proc. cont. à esq é previsível 2. Se X é prev. $E(X(t) \mid \mathcal{F}_{t^-}) = X(t)$

$Y(t) = J(Z \ge t)$ é prev. $d\Lambda(t) = \underset{\text{prev}}{Y(t)}\,\underset{\text{dete}}{\alpha(t)}\,dt$ é previsível $\Rightarrow E(dM(t) \mid \mathcal{F}_{t^-}) = d\Lambda(t)$

$M(t) = N(t) - \Lambda(t)$ é martingal

Def: (Submartingal) $\{X(t) : t \ge 0\}$ c. à direita e $\mathcal{F}_t$ uma filtr. def em um EP. X é adap à $\mathcal{F}_t$, 2. $E|X(t)| < +\infty$ $\forall t < +\infty$, 3. $E(X(t+s) \mid \mathcal{F}_s) \ge X_s \Leftrightarrow$ X é submart.

Teo (Dec. de Doob-Meyer) Seja X um submart. c.d. c não negativo, adaptado à filtração $\mathcal{F}_t$, então $\exists$ um mart M c.d. e um proc. prev. A t.q. $E(A(t)) < +\infty$ e $X(t) = M(t) + A(t)$ $\forall t \ge 0$. Se $A(0) = 0$ e $X = M' + A'$ é outra decomp com $A'(0) = 0$, então $\forall t \ge 0$ $P(M'(t) \ne M(t)) = P(A'(t) \ne A(t)) = 0$

Corolário: Seja $\{N(t) : t \ge 0\}$ um proc. cont. adap à f.f. $\{\mathcal{F}_t, t \ge 0\}$ com $E(N(t)) < +\infty$ $\forall t$. Então $\exists$ um único proc. $\Lambda(t)$ $\uparrow$ decrescente e prev. t.q. $\Lambda(0) = 0$, $E(\Lambda(t)) < +\infty$ $\forall t$ e $M(t) = N(t) - \Lambda(t)$ é mart. $\Lambda(t)$ é com como compensador de $N(t)$.

Def: Proc de Var. Previsível. Dl um mart. c.d. adap à $\{\mathcal{F}_t : t \ge 0\}$ def-se o proc. de var. prev. $\langle M \rangle(t) := \lim_{n \to \infty} \sum_{k=1}^{n} Var(\Delta M_k \mid \mathcal{F}_{t_{k-1}})$ em que $(0,t)$ é part. em n int. de comp t/n e $\Delta M_k = M(kt/n) - M((k-1)t/n)$

Obs: Inf. $d\langle M \rangle(t) = Var(dM(t) \mid \mathcal{F}_{t^-})$ ou $\langle M \rangle(t) = \int_0^t Var(dM(u) \mid \mathcal{F}_u)$

$Var(dM(t) \mid \mathcal{F}_{t^-}) = E((dM(t))^2 \mid \mathcal{F}_{t^-}) - (E(dM(t) \mid \mathcal{F}_{t^-}))^2 = E((dM(t))^2 \mid \mathcal{F}_{t^-})$

Res. Seja M um mart. c.d c.r. à $\{\mathcal{F}_t, t \ge 0\}$ e assume $E(M^2(t)) < +\infty$ $\forall t \ge 0$ então $\exists$ um único proc. prev. $\langle M \rangle(t)$ den. proc. de Var. Prev. s.t. $\langle M \rangle(0) = 0$ e $E(\langle M \rangle(t)) < +\infty$ $\forall t$ e $\{M^2(t) - \langle M \rangle(t) : t \ge 0\}$ é mart. c.d.

Obs: Assume $M(0) = 0 \Rightarrow E(M^2(t) - \langle M \rangle(t)) = 0 \Rightarrow E(M^2(t)) = E(\langle M \rangle(t))$

Ex. $d\langle M \rangle(t) = Var(dM(t) \mid \mathcal{F}_t) = Var(dN(t) - \underset{\text{dete}}{d\Lambda(t)} \mid \mathcal{F}_t) = Var(dN(t) \mid \mathcal{F}_t)$

$dN(t)$ é biná. $\Rightarrow d\langle M \rangle(t) = E(dN(t) \mid \mathcal{F}_t) - (E(dN(t) \mid \mathcal{F}_t))^2 = d\Lambda(t) - (d\Lambda(t))^2$

$d\langle M \rangle(t) \approx d\Lambda(t)$ ou $\langle M \rangle(t) = \Lambda(t)$

Def. M_1 e M_2 mart. c.d. e ou $E(M_i(t)) < +\infty$ $\forall t \ge 0$, $i = 1,2$. O proc. de Covar. Esperado Previsível é dado por $\langle M_1, M_2 \rangle(t) := \lim_{n \to \infty} \sum_{k=1}^{n} Cov(\Delta M_{1k}, \Delta M_{2k} \mid \mathcal{F}_{t_{k-1}})$ em q. ..., e $\Delta_{ik} = M_i(tk/n) - M_i(t(k-1)/n)$, $i = 1,2$. Pode-se escrever $d\langle M_1, M_2 \rangle(t) = Cov(dM_1(t), dM_2(t) \mid \mathcal{F}_{t^-})$

Result. M_i, $i = 1,2$ dois mart c.d. c.r. à $\{\mathcal{F}_t, t \ge 0\}$ ass. $E(M_i^2(t)) < +\infty$ $\forall t \ge 0$, $i = 1,2$. Então $\exists$ um proc. prev. cont. à dir. $\langle M_1, M_2 \rangle$ chamado proc. de var. com $\langle M_1, M_2 \rangle(0) = 0$, $E(\langle M_1, M_2 \rangle(t)) < +\infty$ t.q. $\Delta \langle M_1, M_2 \rangle$ é def. de 2 proc. crescentes, possível e. $M_1 M_2 - \langle M_1, M_2 \rangle$ é um mart. igual.

Def (Proc. Var. Quadrático) Pl um mart. c.d. e adap à $\{\mathcal{F}_t, t \ge 0\}$, def. PVQ $[M](t)$ $[M](t) = \lim_{n \to \infty} \sum_{k=1}^{n} (\Delta M_k)^2$ com $\Delta M_k = M(tk/n) - M((k-1)/n)$ em que

Podemos escrever $[M](t) = \sum_{s \le t} (dM(s))^2$ ← soma de pos(ns) do proc. de M_i

Res. $M^2(t) - [M](t)$ é mart. nulo em 0; $Var(M(t)) = E[[M](t)]$

Ex: $M(t) = N(t) - \Lambda(t) \Rightarrow \langle M \rangle(t) = \Lambda(t)$. $[M](t) = \sum_{s \le t} (dM(s))^2 =$
$= \sum_{s \le t} (dN(s))^2 = \sum_{s \le t} dN(s) = N(t)$.

Int. Estocástica Uma proc def. é $L(t) = \int_0^t H(s)\,dM(s) = \lim_{n \to \infty} \sum_{k=1}^{n} H_k \Delta M_k$ em que $(0,t]$ ... com $H_k = H((k-1)t/n)$ e $\Delta M_k = M(tk/n) - M((k-1)t/n)$, $H(t)$ é um proc. prev. e $M(t)$ um mart. c.r. à $\mathcal{F}_t$.

Result. $L(t)$ é um mart. (sob cond gerais): $E(dL(t) \mid \mathcal{F}_t) = E(H(t)\,dM(t) \mid \mathcal{F}_t)$
$= H(t)\,E(dM(t) \mid \mathcal{F}_t) = H(t) \cdot 0 = 0$. Além disso, tem média zero.

Proc. de Var. Prev. de $L(t)$: $d\langle L \rangle(t) \overset{def}{=} Var(dL(t) \mid \mathcal{F}_{t^-}) = Var(H(t)\,M(t) \mid \mathcal{F}_{t^-})$
$= H^2(t)\,Var(dM(t) \mid \mathcal{F}_{t^-}) = H^2(t)\,d\langle M \rangle(t)$ ou $\langle L \rangle(t) = \int_0^t H^2(s)\,d\langle M \rangle(s)$

Proc. de Var. Quad. de $L(t)$: $[L](t) = \sum_{s \le t} (dL(s))^2 = \sum_{s \le t} (H(s)\,dM(s))^2 = \sum_{s \le t} H^2(s)(dM(s))^2$
$= \sum_{s \le t} H^2(s)\,dN(s) = \int_0^t H^2(s)\,dN(s)$

Proc. de Wiener: (Padrão) Um proc est. sat. $W(0) = 0$ e $E W(t) = 0$ $\forall t$, $W(t)$ tem inc. ind., ass. $W(t) - W(s)$ é ind. de $W(u)$, $0 \le u \le t$, e $W(t) - W(s) \sim N(0, t-s)$ e W tem trajs. cont. c.q.s. / Seja a sol. $V(t)$ uma função est. crescente com $V(0) = 0$. Considere o proc. $U(t) = W(V(t)) \Rightarrow$ $U(t)$ tem trajs. cont. e inc. ind., $W(U(t)) - W(U(s)) \sim W(0, V(t) - V(s))$, $U(t)$ é dito Martingal Gaussiano com proc. de var. Previsível $\langle U \rangle(t) = V(t)$.

Proc. Conta Mult.v. O proc $\{N_i\}$ N_i é chamado de p de cont mult. se cada N_i é p. de con e dois p c não saltam mesm int.

Teo Seja $(N_1, \dots, N_k)$ um p. c.m e Λ_i seus resp compens., $i = 1, \dots, k$. Defina $U_i(t) = \sum_{j=1}^{n} \int_0^t H_{j,i}\,dM_j(u)$, $i = 1, \dots, k$, $H_{j,i}$ são p. prev. e limitado $\Rightarrow$ i. $U_i(t)$ é mart. ii) $E[U_i(t)] = 0$ $\forall 0 \le t < +\infty$, iii) $Cov(U_i(s), U_i(t)) = \sum_{j=1}^{n} \int_0^{t \wedge s} E[H_{j,i}(u) H_{j,i}(u)\,d\Lambda(u)]$

Teo. Bebo-Ledu. Seja $N_1(t) = (N_1(t), \dots, N_n(t))$ p.c.m e $\Lambda(t) = (\Lambda_1(t), \dots, \Lambda_n(t))$ os resp. comp. P. proc prev. $H_{ji}(t)$, $j = 1, 2, \dots, k$, $i = 1, \dots, n$, def. no es. mart. seguir

$U_j^{(n)}(t) = \sum_{i=1}^{n} \int_0^t H_{ji}(u)(dN_i(u) - d\Lambda_i(u))$, $j = 1, 2, \dots, k$

$U_{jt}^{(n)}(t) = \sum_{i=1}^{n} \int H_{ji}(s)\,I(Y_i(s) > 0)(dN_i(s) - d\Lambda_i(s))$ e

$U^{(n)}(t) = (U_1^{(n)}(t), U_2^{(n)}(t), \dots, U_k^{(n)}(t))$ vetor l.x.

Cons. do cont. a) $\langle U_j^{(n)} \rangle(t) \xrightarrow[n \to \infty]{P} V(t)$, $\langle U_j^{(n)}, U_l^{(n)} \rangle \xrightarrow[n \to \infty]{P} V_{jl}(t)$

$V(t)$: matriz determinística, c.p. positiva definida

b) $\langle U_{jt}^{(n)} \rangle(t) \xrightarrow[n \to \infty]{P} 0$, $j = 1, 2, \dots, k$. Então se (a) e (b) são vál.

$U^{(n)}(t) \xrightarrow{D} U^{(\infty)}$ proc. gaussiano com

$\langle U^{(\infty)} \rangle(t) = V(t)$. Em part, para t fixo $U^{(n)}(t) \xrightarrow{D} N(0, V(t))$

Ex). N-A. Seja T o tempo de falha, condicionado $f(\cdot)$ e sob $S(\cdot)$ e C t com $T \perp C$

Def. $\delta = I(T \le C)$ e $Z = \min(T, C)$. Assuma que temos $(z_1, \delta_1), \dots, (z_n, \delta_n)$ a.a. Sejam $t_1 < t_2 < \dots < t_k$ tempos de falha. Para $j = 1, \dots, k$ def.

d_j = # falhas em t_j e r_j = # de ind. em risco em t_j $\Rightarrow$ h.a.v.

$Q(t_j) = d_j / r_j \Rightarrow \hat{A}(t) = \sum_{j: t_j \le t} d_j / r_j = \int_0^t dN(u)/Y(u)$

$N(t) = \sum N_i(t)$ = # de falhas até t $Y(t) = \sum Y_i(t)$ = # ind. em risco

Prop. Est. N-A. Seja $A(t) = \int_0^t I(Y(u) > 0)\lambda(u)\,du$ e $M(t) = N(t) - A(t)$. Supondo $A(t) < +\infty$ $\forall t \ge 0$, temos

(1) $E(\hat{A}(t) - A(t)) = 0$, (2) $E(\hat{A}(t) - A(t))^2 = \int_0^t (1-n(u))^2 \lambda(u)\,du + \int_0^t (n(u))^2 \lambda(u)$

2) $\sigma_n^2 = \frac{t}{t} \cdot \mathrm{Var}(\sqrt{n}(\hat{A}(t), \hat{N}(t))^T = E\left(n \int_0^t \frac{I(\bar{Y}(s)>0)}{\bar{Y}(s)} \alpha(s)\, ds\right)$

Obs. (1) $\hat{N}(t) \le \hat{A}(t)$ (2) $E(\hat{A}(t) - \hat{N}(t)) \ge 0$

O est. N-A tem v. neg. que → o qdo $n \to \infty$ e t bom? p tbe neg. - (t-s)(s)

Aos. $\sigma_n^2(t) \approx \mathrm{Var}(\sqrt{n}\,\hat{A}(t))$

Dist. Assint. Aalen. Ass. qo ∃ função $\pi(t)$ t.q. $\bar{Y}(t)/n \xrightarrow{P} \pi(t)\ \forall t \in [0,\tau]$, $\tau = \sup\{t : \pi(t) > 0\}$

Então para q $\pi(t) > 0$, temos

$\int_{\bar{Y}(u)} dA(u) \to \int \cdot \xrightarrow{P} 0$

Teor. O proc. $\sqrt{n}(\hat{A}(t) - A(t))$, $0 \le t \le \tau$, conv. em dist. p/ um m. gauss. de média zero e f. cov.

$$\sigma^2(t) = \int_0^t dA(u)/\pi(u).$$

que pode ser consist. estimado por

$$\hat{\sigma}^2(t) = \int_0^t d\hat{N}(u)/\bar{Y}^2(u)/n$$

Em p/ pto fixado, $\sqrt{n}(\hat{A}(t_0) - A(t_0)) \xrightarrow{D} N(0, \sigma^2(t_0))$

Est. KM.

Seja $(T_1, \delta_1) \ldots (T_n, \delta_n)$ m.a.a. de (T, δ), $T = \min(T^*, C)$, $\delta = I(T^* \le C)$.

Denote por $\{t_1 < t_2 < \ldots t_{k_0}\}$ dist. de falha, $0 \le k_0$

a q.c.m. de $S(t)$ é dado por

$$\hat{S}(t) = \prod_{j: t_j \le t} \frac{r_j - d_j}{r_j} = \prod_{j: t_j \le t} \left(1 - \frac{d\hat{N}(t_j)}{\bar{Y}(t_j)}\right)$$

Ela não é do prod. $S(t) = e^{-A(t)}$. Obs que

$S(t) = P(T > t_1) P(T > t_2 | T > t_1) P(T > t_3 | T > t_2) \ldots P(T > t_k | T > t_{k-1})$

$= \prod_{j=1}^{k} S(t_j / t_{j-1}) = \prod_{j=1}^{k} (1 - dA(t_j))$ $(S(t) = \prod_{0 \le t} (1 + dA(t)))$

Temos

Caso Abs. cont.: $S(t) = \prod_{0 \le t} (1 - dA(u)) = \exp\left\{-\int_0^t \alpha(u)\, du\right\}$

$= \exp\left\{-\int_0^t dA(u)\right\}$

Caso Discreto:

$$S(t) = \prod_{u \le t} (1 - dA(u)) = \prod_{u \le t} (1 - \alpha(u)) \text{ onde } \alpha(u) = P(T = u | T \ge u)$$

Rel. N-A e KM. Vimos que $\hat{A}(t) = \int_0^t \frac{d\hat{N}(u)}{\bar{Y}(u)} \Rightarrow d\hat{A}(t_j) = \frac{d\hat{N}(t_j)}{\bar{Y}(t_j)}$

Dem.

$\hat{S}(t) = \prod_{u \le t} (1 - d\hat{A}(u)) = \prod_{t_j \le t} (1 - d\hat{A}(t_j)) = \prod_{t_j \le t} \left(1 - \frac{d\hat{N}(t_j)}{\bar{Y}(t_j)}\right)$ $\hat{S}(t)$

Prop. $\hat{S}(t)$. Se $S(t) > 0$, $\frac{\hat{S}(t)}{S(t)} = 1 - \int_0^t \frac{\hat{S}(s)}{S(s)} \left(\frac{d\hat{N}(s)}{\bar{Y}(s)} - dA(s)\right)$

Corol. P/ qq t $S(t) > 0$, $\hat{S}(t) - S(t) = -S(t) \int_0^t \frac{S(u^-)}{S(u)} \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)} d\hat{N}(u) + B(t)$

$B(t) = I(T_0 \le t) \frac{\hat{S}(T_0)(S(T_0) - S(t))}{S(T_0)}$ $T_0 = \inf\{s : \bar{Y}(s) = 0\}$

Result. Seja $\pi(t) = P(T \ge t)$, $T = \min(T^*, C) \to$

$0 \le E(\hat{S}(t) - S(t)) \le (1 + S(t))(1 - \pi(t))^n$ (superestima)

Teo. Seja $t \in [0, \tau]$ e assuma que $\int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)} \alpha(u) \xrightarrow{P} 0$ p/ $n \to \infty$

Então $\sup_{s \in [0,t]} |\hat{S}(s) - S(s)| \xrightarrow{P} 0$ consis. uniforme

Teo. Seja $t \in [0, \tau]$ e assuma que. Para todo $s \in [0, t]$,

$n \int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)} \alpha(u)\, du \xrightarrow{P} \sigma^2(s)$ p/ $n \to \infty$

$n \int_0^t \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)} \alpha(u)\, I\left(\left|\sqrt{n} \frac{I(\bar{Y}(u)>0)}{\bar{Y}(u)}\right| > \epsilon\right) du \xrightarrow[n \to \infty]{P} 0$

$\sqrt{n} \int_0^t (1 - I(\bar{Y}(u)>0)) \alpha(u)\, du \xrightarrow{P} 0$ com $n \to +\infty \Rightarrow$

$\sqrt{n}(\hat{S}(\cdot) - S(\cdot)) \to$ proc. gauss. com f. cover. $\mathrm{Cov}(W(s), W(t))$

$= S(t) S(s) \int_0^{s \wedge t} dA(u)/\pi(u)$, $\pi(u) = P(T \ge u)$

na prática, $\sqrt{n}(\hat{S}(\cdot) - S(\cdot)) \xrightarrow{D} -S(\cdot) U(\cdot)$ em que

$U(\cdot)$ é m. gauss. com $U(0) = 0$ e $\mathrm{Cov}(U(s), U(t)) = \int_0^{s \wedge t} \frac{dA(u)}{\pi(u)}$

Obs. Para t fixado

$\sqrt{n}(\hat{S}(t_0) - S(t_0)) \xrightarrow{D} N(0, \hat{\sigma}_S(t_0))$,

$\hat{\sigma}^2(t) = \hat{S}^2(t_0) \int_0^{t_0} \frac{dA(u)}{\pi(u)}$. Seja $\sigma^2(t) = \int_0^t \frac{dA(u)}{\pi(u)}$. Podemos

est. mar $\sigma^2(t)$ por $\hat{\sigma}^2(t) = \int_0^t \frac{d\hat{N}(u)}{\bar{Y}(u)} \frac{1}{\bar{Y}(u)/n}$. Assim

$$\hat{V}(\hat{S}(t_0)) = \frac{1}{n} \hat{S}(t_0) \int_0^{t_0} \frac{d\hat{N}(u)}{\bar{Y}^2(u)/n}$$

Outro est. t usado na lit. Vimos que $\mathrm{Var}(\sqrt{n}\,\hat{S}(t)) = S^2(t) \int_0^t \frac{dA(u)}{\pi(u)}$

Obs que $dA(t) = \alpha(t) = \frac{f(t)}{S(t)} = -dS(t)/S(t)$

Assim

$\mathrm{Var}(\sqrt{n}\,\hat{S}(t)) = S^2(t) \int_0^t -dS(t)/S(t) \cdot 1/\pi(u)$ c

$\widehat{\mathrm{Var}}(\sqrt{n}\,\hat{S}(t)) = \hat{S}^2(t) \int_0^t -\frac{d\hat{S}(t)}{\hat{S}(t)} \frac{1}{\bar{Y}(u)/n}$ (t. Greenwood)

Int. Conf. Ponto a Ponto sem transf.

Seja $\hat{\sigma}^2(t_0) = \int_0^{t_0} -\frac{d\hat{S}(u)}{\hat{S}(u)\bar{Y}(u)}$. Vimos q $\widehat{\mathrm{Var}}(\hat{S}(t_0)) = \hat{S}^2(t_0)\hat{\sigma}^2(t_0)$

Assim, p/ t fixado, $(1-\alpha)\%$ dos IC p/ $S(t_0)$ é dado por

$\hat{S}(t_0) \pm z_{1-\alpha/2}\, \hat{\sigma}(t_0) \hat{S}(t_0)$ - LI e LS podem extrapolar

Transf. log-log. Para t fixado, defina $\theta_0 = \log(-\log S(t_0)) = g(S(t_0))$ e $\hat{\theta}_0 = \log(-\log \hat{S}(t_0)) = g(\hat{S}(t_0))$. Aplicamos o método do Delta a

$g(\hat{S}(t_0)) = \hat{\theta}_0$. $g(\hat{S}(t_0)) \approx g(S(t_0)) + g'(S(t_0))(\hat{S}(t_0) - S(t_0))$

$\Rightarrow g(\hat{S}(t_0)) - g(S(t_0)) \approx g'(S(t_0))(\hat{S}(t_0) - S(t_0))$

$\mathrm{Var}(\hat{\theta}_0) \approx g'(S(t_0))^2 \mathrm{Var}(\hat{S}(t_0))$

Nosso caso, $g(x) = \log(-\log x) \Rightarrow g'(x) = -1/x\log x$

$\widehat{\mathrm{Var}}(\hat{\theta}_0) \approx (-1/[\hat{S}(t_0)\log \hat{S}(t_0)])^2 \hat{\sigma}^2(t_0)\hat{S}^2(t_0) \approx \hat{\sigma}^2(t_0)/(\log \hat{S}(t_0))^2$

Assim IC para $\theta_0 = g(S(t_0))$

$\log(-\log \hat{S}(t_0)) \pm z_{1-\alpha/2}\, \hat{\sigma}(t_0) / -\log \hat{S}(t_0)$

Um int. para $S(t_0)$ fica dado por

$\left[\exp\left\{-\exp\left(\theta_0 - \left(\frac{\hat{\sigma}(t_0)}{-\log \hat{S}(t_0)}\right) z_{1-\alpha/2}\right)\right\}, \exp\left\{-\exp\left(\theta_0 + \left(\frac{\hat{\sigma}(t_0)}{-\log \hat{S}(t_0)}\right) z_{1-\alpha/2}\right)\right\}\right]$

ou $[\hat{S}(t_0)^{e^{-K}}, \hat{S}(t_0)^{e^K}]$, $K = z_{1-\alpha/2}\sqrt{\mathrm{Var}(\hat{\theta}_0)}$ ou

$[\hat{S}(t_0)^{\theta}, \hat{S}(t_0)^{\theta}]$, $\theta = \exp\{z_{1-\alpha/2}\sqrt{\mathrm{Var}(\hat{\theta}_0)}\}$

Berger e Liesta prop. int. contrast/ crossing

Estimação Percentis. Pode def. o percentil p como $t_p = \inf\{t : S(t) \le 1-p\}$

Um est. p/ t_p é $\hat{t}_p = \inf\{t : \hat{S}(t) \le 1-p\}$ (Klein, Hos lemes) holds

p/ calcular, devemos est. (ppor int.) linear. Encont. t_L e t_U t.q.

$\hat{S}(t_L) > (1-p)$ e $\hat{S}(t_U) \le 1-p$. Resol. $\frac{t_U - t_L}{\hat{S}(t_U) - \hat{S}(t_L)} = \frac{\hat{t}_p - t_L}{(1-p) - \hat{S}(t_L)}$

Teste do log-Rank. Comparar curvas de sobrev.

Pop. 1: (Z_{1j}, δ_{1j}), $j = 1 \ldots n_1$; Pop 2: (Z_{2j}, δ_{2j}), $j = 1 \ldots n_2$

Ass. ind. ent tempos de falha cens. ind. $T_{1j} \sim S_1(\cdot)$ $T_{2j} \sim S_2(\cdot)$

Quer testar as hipót.

$H_0: S_1(t) = S_2(t)\ \forall t \in [0, \tau]$ t vasta maior

$H_A: S_1(t) \ne S_2(t)$ p/ algum $t \in [0, \tau]$ tempo 2 grupos pode ser cruzado

O teste log-R é o mais pod. p/ hip. do tipo $H_0: S_1^\gamma = S_2(t)$, em que

$H_A: S_2(t) = [S_1(t)]^\gamma$, $t \in [0, \tau]$ (o que é tempo prop. cons.)

Construção do Teste. Seja $n = n_1 + n_2$ e com $T_1^* < T_2^* < \ldots < T_k^*$ instants

que alg. falha ocorra na amostra combinada. Para cada inst T_k^*, $k = 1, \ldots, k$

defina D_{1k}: nº de falhas obs. em T_k^* na amostra 1

Y_{1k}: " " indivíduos em risco " " " 1 em inst T_k^*

Para cada inst de falha T_k^*, podemos cont. uma tab.

Grupos	G1	G2	Tot
Falhas	D_{1k}	D_{2k}	D_k
Sobrev	$Y_{1k} - D_{1k}$	$Y_{2k} - D_{2k}$	$Y_k - D_k$
Em risco	Y_{1k}	Y_{2k}	Y_k

Note que

Dado Y_{1k}, D_{1k} / Y_{1k}, Y_{2k}, D_k

Dado Y_k, D_k / Y_{1k}, Y_{2k}, D_{1k}

p_{1k}: prob. obs. falha p/ 1 em T_k^*

Se H_0 vdd, $p_{1k} = p_{2k}$. Agora, condiz. cond.

$D_{1k} / Y_{1k}, Y_k, D_k \overset{H_0}{\sim}$ Hipergeométrica

Sob hip. temos que

$E(D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k) = E_{1k} = e_{1k} = \frac{\bar{Y}_{1k}}{\bar{Y}_k} D_k = \frac{D_k}{\bar{Y}_k} \bar{Y}_{1k}$.

$Var(D_{1k} / \bar{Y}_{1k}, \bar{Y}_{2k}, D_k) = V_{1k} = D_k \frac{\bar{Y}_{1k}}{\bar{Y}_k^2} \bar{Y}_{2k} \frac{(\bar{Y}_k - D_k)}{(\bar{Y}_k - 1)}$

O LR e bas. na ideia de comp. o obs com esp se H_0 é vdd

Assumindo ind. ent. t. falha, temos o qui-dois, pé

$$LR = \left[\sum_{j=1}^{L} (D_{1j} - E_{1j})\right]^2 / \left[\sum_{j=1}^{L} V_{1j}\right] \sim LR \xrightarrow[n \to \infty]{D} \chi^2_{(1)}$$

Ex. G1: 3 5 7 9* 18

G2: 8* 12 19 20 20*

$T_i^* = 3$

	G1	G2	Total
F	1	0	1
S	4	5	9
Em risco	3	5	10

T_k^*	Y_k	D_k	$\bar{Y}_{1k}$	D_{1k}	E_{1k}	V_{1k}
3	10	1	5	1	0.5	0.25
5	9	1	4	1	0.444	0.247
7	8	1	3	1	0.375	0.239
12	5	1	1	0	0.2	0.160
18	4	1	1	1	0.25	0.183
19	3	1	0	0	0	0
20	2	1	0	0	0	0

Generalizações. Inclusão de ponderações. $WLR = \left[\sum_{j=1}^{L} w_j (D_{1j} - E_{1j})\right]^2 / \left[\sum_{j=1}^{L} w_j^2 V_{1j}\right] \xrightarrow{D} \chi^2_{(1)}$, w_j pesos não aleat.

$w_j = 1 \rightarrow LR$

$w_j = \bar{Y}_j \rightarrow$ Gehan ou Wilcoxon

$w_j = \sqrt{\bar{Y}_j}$ - Tarone-Ware

$w_j = (\hat{S}(t_{j-1}))^p (1 - \hat{S}(t_{j-1}))^q \rightarrow$ Fam. Fleming-Harrington, $\hat{S}$ é $\hat{S}$ ob. a mod. comb.

$p, q \geq 1$, $0 \leq p, q \leq 1$.

$w_j = \frac{\bar{Y}_j}{\bar{Y}_{j+1}} \hat{S}(t_{j-1}) \rightarrow$ Peto-Prentice, $\hat{S}(t) = \prod_{t_k \leq t} (1 - d_k / \bar{Y}_{k+1})$

Const. Seja Q_n, $n = 1, 2, \ldots$ uma seq. de est. de test p/ testar a hip. H_0. Seja R_n os reg. rej. assoc. Q_n com nível de sig. α. A seq. Q_n é consist. para uma hip. alternat. H_A se ~~[illegible]~~ $\lim_{n \to \infty} P(Q_n \in R_n) = 1$ se H_A é vdd

Gen. P K-Grupos. Ci. dado o mesmo, pode se trab. com o dol de $(Q_{1n}, \ldots, Q_{n-1,K}) \sim N$ p/ g. multivar.

Teste Log-Rank. Estrat. ficção: Considér. obs. nos dados com M categorias. Suponha que mt. teste iguald. das f. de sob. dos 2 grupos (ind. p.e) dentro cada nível (strat. p.e.). Queremos testar $H_0: S_1^{(s)}(t) = S_2^{(s)}(t)$, $s = 1, 2, \ldots, M$ vs. H_A: $\exists$ desigualdade $\forall t \in \tau$.

Ti. Ranq. alt. qdo há cruz. das f. de risco. Suponha que seja testar $H_0: S_1(t) = S_2(t)$, $t \in [0, \tau]$. Idéia: $T_1^* < T_2^* < \ldots < T_L^*$ inst. de fal. ord. comb. Calc. para cada T_k^*, $k = 1, \ldots, L$, $Z(T_k^*) = \sum_{j: T_j^* \leq T_k^*} \left(D_{1j} - D_{1j} \frac{Y_{2j}}{Y_j}\right)$, $k = 1, 2, \ldots, L$ e obt. $Z(T_1^*), \ldots, Z(T_L^*)$

A est. de t. é dada por $Q = [\sup\{|Z(t)|, t \leq \tau\}] / \sigma(\tau)$. Sob H_0, pode-se dem. que $Q \xrightarrow{D} \sup(|B(x)|, 0 \leq x \leq 1)$. Rej. H_0 vários p. Q

M. Reg. Param. As. temos um com. (Z_i, δ_i, X_i) com $Z_i = \min(T_i, C_i)$, $\delta_i = I(T_i \leq C_i)$, $X_i = (1, x_{i1}, \ldots, x_{i(p-1)})^T$ (covariáveis). Assumo $T_i / X_i \sim f$

M. Reg. Exponencial. Podemos ter $f_i(t) = \lambda_i e^{-\lambda_i t}$, $\lambda_i = e^{X_i' \beta}$, i.e., $f_i(t) = (e^{X_i' \beta}) e^{-(e^{X_i' \beta} t)}$. Note $\lambda(t / X_i) = e^{X_i' \beta}$, relevante $\lambda(t / X_i) / \lambda(t / X_j)$ não dep. de t (modelo de riscos proporc.). Outra prop. $C_i \sim Exp(1)$, $T_i = C_i / \lambda_i = e^{-X_i' \beta} C_i$. Def. $Y_i = \ln T_i = -X_i' \beta + \ln \xi_i = -X_i' \beta + W_i$, W_i não p.

As seg. formul. são eq. $T_i / X_i \sim Exp(e^{X_i' \beta})$. $Y_i = \log T_i = X_i' \gamma + W_i$, $W_i \sim i.i.d.$ p. $(\gamma = -\beta)$

Def. Um mod. de sob é dito modelo de vida acel. se $S(t / X) = S_0(\psi_x t)$, $\forall t$, $\psi_x > 0$ e onde $S_0(\cdot)$ é uma f. s. basal, que não dep. do X e ψ_x é conhec. como fator de acel.

Est. Mod. Reg. Exp. $Y_i = \ln(T_i) = X_i' \gamma + W_i$, $Y_i = \ln(\min(T_i, C_i))$, $\delta_i = I(T_i \leq C_i)$, X_i. Para o c. à direita, $L(\gamma) = \prod (f(y / X))^{\delta} (S(y / X))^{1-\delta}$

$= e^{\sum \delta_i (y_i - X_i' \gamma)} \exp\{-\sum e^{y_i - X_i' \gamma}\} \Rightarrow U(\gamma) = \partial \ell(\gamma) / \partial \gamma = -\sum \delta_i X_i + \sum X_i e^{y_i - X_i' \gamma} \rightarrow$ não num. $-\partial / \partial \gamma \partial \gamma' \ell(\gamma)$ inf. observ.

M. R. Weibull. Podemos cons. $S(t / X_i) = e^{-\theta_i t^{\rho}}$, $\theta_i = e^{X_i' \beta}$ ou $\lambda(t / X_i) = \theta_i \rho t^{\rho - 1} = (e^{X_i' \beta}) \rho t^{\rho - 1}$ (riscos prop.). Cons. o mod. $Y = \ln T = X' \gamma + \sigma W$, $W \sim$ v.a.e.p. $S_T(t / X) = \exp\{-t^{1/\sigma} e^{-X_i' \beta / \sigma}\}$, $T \sim$ Weibull, $\rho = 1/\sigma$, $\beta = 1/\sigma \gamma$. E.V. $Y_i = \ln T_i = X_i' \gamma + \sigma W_i$.

$L(\gamma, \sigma) = \exp\{\sum \delta_i (y_i - X_i' \gamma) / \sigma\} \exp\{-\sum e^{(y_i - X_i' \gamma) / \sigma}\}$

M. R. Log-Log. Cons. mod. $Y = \ln T = X' \gamma + \sigma W \Rightarrow S_T(t / X) = 1 / (1 + t^{1/\sigma} e^{-X' \gamma / \sigma}) \rightarrow$ log-ver. $\Rightarrow$ par. pder mt. num.

M. R. Log-Norm. $Y = \log T \sim N(\mu, \sigma^2) \Rightarrow Y = \log t = \mu + \sigma \varepsilon$, $\varepsilon \sim N(0, 1)$, $\Rightarrow$ i.e. das v. pode ser fei. $Y_i = \log T_i = X_i' \gamma + \sigma \varepsilon_i$, $\varepsilon_i \sim N(0, 1)$

M. Loglog-Locale. Obs. que todos os mod. p. podem ser esc. na forma $Y_i = \log T_i = -X_i' \gamma + \sigma \varepsilon_i$, ε_i v.a. dist. sp. | pode ser (mod. loc. escala)