I SECICAT

I SEMINÁRIO DE CIÊNCIAS CONTÁBEIS E ATUARIAIS DA UFPB

05 E 06 DE NOVEMBRO DE 2015

TEMA: A INFORMAÇÃO CONTÁBIL E ATUARIAL PARA TOMADA DE DECISÃO

ANÁLISE DE SOBREVIVÊNCIA

UFPB / CCSA / DFC / I SECICAT
LUIZ CARLOS SANTOS JÚNIOR

ROTEIRO

1 - Conceitos básicos e exemplos

2 - Técnicas não-paramétricas

3 - Modelos probabilísticos ou paramétricos

1 CONCEITOS BÁSICOS E EXEMPLOS

INTRODUÇÃO

A variável resposta é, geralmente, **tempo de falha**.

A principal característica de dados com sobrevivência é a presença de **censura**.

Dados truncados.

CARACTERIZANDO DADOS DE SOBREVIVÊNCIA

TIPOS DE CENSURA

Censuras do Tipo I (período pré-estabelecido de tempo)

Censuras do Tipo II (número pré-estabelecido de indivíduos).

Censura aleatória ocorrem quando o “paciente” é “retirado” no decorrer do estudo sem ter ocorrido a falha.

REPRESENTAÇÃO DOS DADOS DE SOBREVIVÊNCIA

$$\delta_i = \begin{cases} 1 \; se \; t_i \text{ é } um \; tempo \; de \; falha \\ 0 \; se \; t_i \text{ é } um \; tempo \; censurado \end{cases}$$

ESPECIFICANDO O TEMPO DE SOBREVIVÊNCIA

Função de Sobrevivência

$$S(t) = P(T \geq t) = 1 - P(T \leq t) = 1 - F(t)$$

ESPECIFICANDO O TEMPO DE SOBREVIVÊNCIA

Função Taxa de Falha

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t}$$

ESPECIFICANDO O TEMPO DE SOBREVIVÊNCIA

Função de Taxa de Falha Acumulada

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

ESPECIFICANDO O TEMPO DE SOBREVIVÊNCIA

Tempo Médio e Vida Média Residual

$$t_m = \int_0^\infty S(t)dt$$

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

Observe que $vmr(0) = t_m$

ESPECIFICANDO O TEMPO DE SOBREVIVÊNCIA

Relações entre as Funções

$$\lambda(t) = \frac{f(t)}{S(t)}$$

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

$$S(t) = e^{-\Lambda(t)} = e^{-\int_0^t \lambda(u)du}$$

2 – TÉCNICAS NÃO-PARAMÉTRICAS

INTRODUÇÃO

Os **objetivos** de uma análise estatística envolvendo dados de sobrevivência estão geralmente relacionados, em medicina, à **identificação de fatores de prognóstico para uma doença** ou à **comparação de tratamentos** em um estudo clínico enquanto controlado por outros fatores.

O ESTIMADOR DE KAPLAN MEIER (OU LIMITE PRODUTO)

O estimador de KM se reduz a estimar q_j

$$q_j = \frac{nº\ de\ falhas\ em\ t_{j-1}}{nº\ de\ observações\ sob\ risco\ em\ t_{j-1}} = \frac{d_j}{n_j}$$

$$\hat{S}(t) = \prod_{j:t_j<t} \left(\frac{n_j - d_j}{n_j}\right) = \prod_{j:t_j<t} \left(1 - \frac{d_j}{n_j}\right) = \prod_{j:t_j<t} \left(1 - q_j\right)$$

$$\hat{S}(t) = \prod_{j:t_j<t} p_j$$

O ESTIMADOR DE KAPLAN MEIER (OU LIMITE PRODUTO)

Variância de $\widehat{S}$(t)

$$\hat{V}(\hat{S}(t)) = \left[\hat{S}(t)\right]^2 \sum_{j:t_j} \frac{d_j}{n_j(n_j - d_j)}$$

Intervalo de confiança

$$\hat{S}(t) = \pm\, z_{\frac{\alpha}{2}} \sqrt{V(\hat{S}(t))}$$

COMPARAÇÃO DE CURVAS DE SOBREVIVÊNCIA

O objetivo principal do estudo normalmente é comparar os grupos. Para isso, é possível a utilização de diversos testes, tais quais o de Logrank (MANTEL, 1966), o de Peto, o de Wilcoxon etc.

$$T = \frac{\left[\sum_{j=1}^{k}\left(d_{2j} - w_{2j}\right)\right]^2}{\sum_{j=1}^{k}\left(V_j\right)}$$

Sob a hipótese nula H_0: $S_1(t) = S_2(t)$ para todo t no período de acompanhamento, tem distribuição qui-quadrado com 1 grau de liberdade para grandes amostras.

3 - MODELOS PROBABILÍSTICOS OU PARAMÉTRICOS

DISTRIBUIÇÃO EXPONENCIAL

$$f(t) = \frac{1}{\alpha} exp\left\{-\frac{t}{\alpha}\right\}, t \geq 0;$$

$$S(t) = exp\left\{-\frac{t}{\alpha}\right\};\ h(t) = \frac{1}{\alpha}\ , t \geq 0$$

“LINEARIZAÇÃO NA EXPONENCIAL”

$$-ln(S(t)) = \frac{1}{\alpha}t$$

DISTRIBUIÇÃO WEIBULL

$$f(t) = \frac{\gamma}{\alpha^{\gamma}} t^{\gamma-1} exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\}, t \geq 0;\ \alpha, \gamma > 0$$

$$S(t) = exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\};\ \ h(t) = \frac{\gamma}{\alpha^{\gamma}}\ t^{\gamma-1}, t \geq 0$$

“LINEARIZAÇÃO NA WEIBULL”

$$-ln\big(S(t)\big) = \left(\frac{t}{\alpha}\right)^{\gamma}$$

OUTROS MODELOS PROBABILÍSTICOS

Existem outras distribuições de probabilidade apropriadas para modelar o tempo de falha. Dentre elas, podem ser citadas as distribuições log-gama, Rayleigh, normal inversa e **Gompertz**.

Diversos textos apresentam a popular **função de taxa de falha do tipo banheira**, que descreve o comportamento do **tempo de vida de seres humanos**. Para esta função, distinguem-se três regiões distintas: período de falhas prematuras (**mortalidade infantil**), período de vida útil (**fase adulta**) e período de desgaste (**velhice**).

•

ESCOLHA DO MODELO PROBABILÍSTICO

Métodos Gráficos

a) Comparação da função de sobrevivência do modelo proposto com o estimador de Kaplan-Meier.

b) Linearização dos parâmetros.

ESCOLHA DO MODELO PROBABILÍSTICO

Testes de Hipóteses

Teste de Razão de Verossimilhança

H_0: o modelo de interesse é adequado.

$$TRV \; = \; -2log\left[\frac{L\left(\hat{\theta}_M\right)}{L\left(\hat{\theta}_G\right)}\right] = 2\left[logL\left(\hat{\theta}_G\right) - logL\left(\hat{\theta}_M\right)\right]$$

Que, sob H_0, tem aproximadamente uma distribuição qui-quadrado com número de graus de liberdade igual a diferença do número de parâmetros ($\left(\hat{\theta}_G \; e \; \hat{\theta}_M\right)$ dos modelos sendo comparados.

Referência Básica

- COLOSIMO, E. A.; GIOLO, S. R. Analise de Sobreviv ´ encia Aplicada ˆ . Sao Paulo: Edi- ˜ tora Edgard Blucher Ltda, ABE-Projeto Fisher, 2006, 367p