ISSN 0871-3413 • ©*ArquiMed*, 2007

Análise de Sobrevivência

Métodos Não Paramétricos

Joana Bastos*, Cristina Rocha†
**Serviço de Higiene e Epidemiologia, Faculdade de Medicina da Universidade do Porto; †DEIO e CEAUL, Faculdade de Ciências da Universidade de Lisboa*

INTRODUÇÃO

Como em qualquer análise estatística, a análise descritiva dos dados deve ser o primeiro passo no estudo de tempos de vida. Dada a existência de censura, os dados de sobrevivência são resumidos de forma conveniente através de estimativas da função de sobrevivência e da função de risco. Neste artigo, apresentaremos o estimador de Kaplan-Meier para a função de sobrevivência e o estimador de Nelson-Aalen para a função de risco cumulativa. Trata-se de métodos não paramétricos, porque a estimação é feita sem que se faça nenhuma suposição sobre a distribuição de probabilidade do tempo de sobrevivência. Iremos, também, descrever alguns métodos não paramétricos para a comparação de curvas de sobrevivência, nomeadamente o teste log-rank e o teste de Gehan.

O ESTIMADOR DE KAPLAN-MEIER DA FUNÇÃO DE SOBREVIVÊNCIA

Quando os dados não são censurados, a função de sobrevivência, num dado instante t, poderá ser estimada a partir dos tempos de vida observados, como sendo a proporção de indivíduos que sobreviveram para além do instante t. Esta função designa-se por função de sobrevivência empírica e define-se do seguinte modo:

$$\hat{S}(t) = \frac{\text{numero de observaçoes} > t}{n},$$

, em que n representa a dimensão da amostra.

Este método de estimar a função de sobrevivência não se deve aplicar quando existe censura, uma vez que despreza informação relativa a qualquer indivíduo cujo tempo de sobrevivência seja superior a t, mas que tenha sido censurado antes desse instante.

Em 1958, Kaplan e Meier propuseram um estimador não-paramétrico para a função de sobrevivência, quando estamos na presença de uma amostra censurada. Este estimador é designado por estimador de Kaplan-Meier (K-M) ou estimador produto-limite.

Sejam $t_{(1)} < ... < t_{(r)}$ os instantes de morte distintos numa amostra de dimensão n ($r \le n$), o número de mortes ocorridas em $t_{(i)}$ e $n_{(i)}$ o número de indivíduos em risco em $t_{(i)}$. O estimador de Kaplan-Meier da função de sobrevivência, define-se da seguinte forma:

$$\hat{S}(t) = \prod_{i:t_{(i)} \le t} \left(\frac{n_i - d_i}{n_i}\right) = \prod_{i:t_{(i)} \le t} \left(1 - \frac{d_i}{n_i}\right)$$

Observe-se que:

- $\hat{S}(t) = 1$ para $0 \le t < t_{(1)}$;
- $\hat{S}(t) = 0$ para $t \ge t_{(r)}$, se $t_{(r)}$ é a maior observação registada.
- Se a maior observação registada for um tempo censurado t* , então $\hat{S}(t)$ nunca toma o valor zero e não está definido para t > t*.

A variância estimada de $\hat{S}(t)$ é dada por $Var(\hat{S}(t)) \approx \left[\hat{S}(t)\right]^2 \sum_{j=1}^{k} \frac{d_j}{n_j(n_j - d_j)}$ para $t_{(k)} \le t < t_{(k+1)}$. Este resultado é conhecido por "fórmula de Greenwood".

INTERVALOS DE CONFIANÇA PARA A FUNÇÃO DE SOBREVIVÊNCIA NUM PONTO

Uma vez estimado o desvio padrão de $\hat{S}(t)$ pode-se construir um intervalo de confiança para o valor da função de sobrevivência, num dado instante t. Este intervalo é obtido assumindo que o estimador da função no instante t tem distribuição normal com valor médio S(t) e variância estimada $Var(\hat{S}(t))$. Então um intervalo de 100 (1-α) % de confiança para é dado por $\left[\hat{S}(t) - z_{\frac{\alpha}{2}}\sqrt{Var(\hat{S}(t))}, \hat{S}(t) + z_{\frac{\alpha}{2}}\sqrt{Var(\hat{S}(t))}\right]$, em que $z_{\frac{\alpha}{2}}$ representa o quantil de probabilidade $1 - \frac{\alpha}{2}$ da distribuição normal centrada e reduzida, ou seja, da distribuição N(0,1). No caso do estimador de Kaplan-Meier, em que o desvio padrão é $SD\{\hat{S}(t)\} \approx \hat{S}(t)\sqrt{\sum_{j=1}^{k} \frac{d_j}{n_j(n_j - d_j)}}$, um intervalo de 100 (1-α) % de confiança para é dado por

$$\left[\hat{S}_{KM}(t) - z_{\frac{\alpha}{2}}\hat{S}_{KM}(t)\sqrt{\sum_{j=1}^{k} \frac{d_j}{n_j(n_j - d_j)}}, \hat{S}_{KM}(t) + z_{\frac{\alpha}{2}}\hat{S}_{KM}(t)\sqrt{\sum_{j=1}^{k} \frac{d_j}{n_j(n_j - d_j)}}\right]$$

.

Os intervalos de confiança obtidos por este processo são simétricos, sendo possível que os limites do intervalo de confiança sejam maiores que um ou menores que zero, nos instantes em $\hat{S}(t)$ que está próximo desses valores. Uma solução pragmática para este problema, é a substituição de qualquer limite maior que um por 1.0 e qualquer limite menor que zero por 0.0.

Uma alternativa é transformar S (t) num valor pertencente ao intervalo (-∞,+∞), e obter um intervalo de confiança para o valor transformado. Existem várias transformações possíveis, sendo duas delas a transformação logística, $\log\left[\frac{S(t)}{1-S(t)}\right]$, e a transformação complementar log-log, $\log\left[-\log(S(t))\right]$.

Quando se aplica a transformação complementar log-log, obtém-se um intervalo de confiança assimétrico S(t) para a partir de um intervalo simétrico para $\log\left[-\log(S(t))\right]$. Este método, além de assegurar que os limites do intervalo de confiança são positivos e menores ou iguais a um, é bastante preciso, porque $\log\left[-\log(\hat{S}(t))\right]$ tem uma distribuição mais próxima da normal que $\hat{S}(t)$. Assim, o intervalo de 100(1-α) % de confiança para $\log\left[-\log(S(t))\right]$ é dado por $\left[W - z_{\frac{\alpha}{2}}\sqrt{Var(W)},\ W + z_{\frac{\alpha}{2}}\sqrt{Var(W)}\right]$, em que $W = \log\left[-\log(\hat{S}(t))\right]$, $z_{\frac{\alpha}{2}}$ representa o quantil de probabilidade $1-\frac{\alpha}{2}$ da distribuição N(0,1) e $Var(W) \approx \frac{1}{\left[\log(\hat{S}(t))\right]^2}\sum_{j=1}^{k}\frac{d_j}{n_j(n_j-d_j)}$.

O ESTIMADOR DE NELSON-AALEN PARA A FUNÇÃO DE RISCO CUMULATIVA

O estimador proposto por Kaplan e Meier (1958) é o estimador mais utilizado para a estimação não paramétrica da função de sobrevivência. Frequentemente, interessa-nos também estimar a função de risco cumulativa. Um estimador natural de H(t) será $\hat{H}(t) = -\log\hat{S}(t)$, onde $\hat{S}(t)$ é o estimador de Kaplan-Meier. No entanto, um estimador alternativo, sugerido por Nelson (1972) e estudado por Aalen (1978), é uma outra opção que se torna cada vez mais usual. Este estimador é designado por estimador de Nelson-Aalen.

Sejam $t_{(1)} < ... < t_{(r)}$ os instantes de morte distintos numa amostra de dimensão n ($r \leq n$), $d_{(i)}$ o número de mortes ocorridas em $t_{(i)}$ e $n_{(i)}$ o número de indivíduos em risco em $t_{(i)}$. O estimador de Nelson-Aalen define-se por $\hat{H}_{NA}(t) = \sum_{i:\, t_{(i)} \leq t}\frac{d_i}{n_i}$.

O estimador de Nelson-Aalen estima directamente a função de risco cumulativa, embora como é óbvio se possa obter também uma estimativa da função de sobrevivência.

Assim sendo, o estimador de Nelson-Aalen para a função de sobrevivência, também conhecido por estimador de Breslow, é dado por $\hat{S}_{NA}(t) = \exp\left(-\sum_{i:\, t_{(i)} \leq t}\frac{d_i}{n_i}\right)$.

Embora o estimador de Nelson-Aalen apresente um melhor comportamento, para pequenas amostras, do que o estimador de Kaplan-Meier, em muitas circunstâncias as estimativas serão muito semelhantes, principalmente quando ainda existem muitos indivíduos em risco. Como o estimador de Kaplan-Meier é uma generalização da função de sobrevivência empírica, a sua utilização será preferível.

O estimador de Kaplan-Meier com estratificação

Em análise de sobrevivência interessa analisar os factores endógenos ou exógenos aos indivíduos que contribuem para a ocorrência do acontecimento de interesse, isto é, características como o sexo, a idade, a utilização de determinado fármaco, entre outras, podem ter um papel primordial no tempo de sobrevivência, e irão originar curvas de sobrevivência distintas. A estratégia utilizada, com base no estimador de Kaplan-Meier, para comparar as diferentes curvas correspondentes aos vários grupos, é a estratificação. Esta estratificação consiste na divisão do conjunto total de observações em grupos distintos, de acordo com as covariáveis de interesse, e na estimação das funções de sobrevivência, separadamente para cada um dos grupos.

Testes para comparação de curvas de sobrevivência

A representação gráfica da estimativa de Kaplan-Meier com estratificação, para a função de sobrevivência, permite ter uma ideia do comportamento das curvas de sobrevivência, nos respectivos grupos. No entanto, para avaliar se existe uma diferença significativa entre as várias curvas deve-se recorrer aos testes de hipóteses. Existem vários testes não paramétricos adequados a esta comparação, sendo o teste log-rank e o teste de Gehan, para dois grupos, os mais utilizados em análise de sobrevivência.

As hipóteses que pretendemos testar são, H_0: $S_1(t) = S_2(t)$ vs H_1: $S_1(t) \neq S_2(t)$.

O teste log-rank

O teste log-rank compara a distribuição da ocorrência dos acontecimentos observados em cada grupo, com a distribuição que seria esperada, se a incidência fosse igual em todos os grupos. Se a distribuição observada for equivalente à distribuição esperada, dizemos que a função de sobrevivência dos indivíduos pertencentes ao grupo, coincide com a função de sobrevivência dos indivíduos em geral (a covariável não exerce influência

sobre a sobrevivência).

A estatística de teste é dada por $W_L = \frac{U_L^2}{V_L}$ em que $U_L = \sum_{j=1}^{r}(d_{1j} - e_{1j})$ e $e_{1j} = \frac{n_{1j}\, d_j}{n_j}$ representa o número esperado de indivíduos que morrem no instante $t_{(j)}$ no grupo 1; $V_L = Var(U_L) = \sum_{j=1}^{r} v_{1j}$ e $v_{1j} = \frac{n_{1j}\, n_{2j} d_j (n_j - d_j)}{n^2{}_j (n_j - 1)}$.

Sob a validade de H_0, W_L tem distribuição assintótica de Qui-quadrado com 1 grau de liberdade.

Rejeita-se H_0 para valores grandes da estatística de teste, isto é, se $W_L > \chi_{1-\alpha}$, em que $\chi_{1-\alpha}$ representa o quantil de probabilidade da distribuição Qui-quadrado com 1 grau de liberdade.

O teste de Gehan

O teste de Gehan, ou teste de Wilcoxon generalizado, baseia-se numa estatística semelhante à utilizada no teste log-rank e que é a seguinte:

$$U_G = \sum_{j=1}^{r} n_j (d_{1j} - e_{1j}).$$

Verificamos, então, que a diferença $(d_{1j} - e_{1j})$ é ponderada por n_j. Assim sendo, vai ser atribuído menor peso às diferenças $(d_{1j} - e_{1j})$ correspondentes aos instantes onde o número total de indivíduos em risco é pequeno, isto é, aos maiores tempos de sobrevivência. Por isso, este teste é menos sensível do que o teste log-rank a diferenças entre o número observado e o número esperado de mortes que se verifiquem na cauda direita da distribuição.

A variância da estatística é dada por $V_G = \sum_{j=1}^{r} n_j^2 v_{1j}$, então a estatística de teste é $W_G = \frac{U_G^2}{V_G}$.

Sob a validade de H_0, W_G tem distribuição assintótica de Qui-quadrado com 1 grau de liberdade.

Comparação entre o teste log-rank e o teste de Gehan

Notemos que o teste log-rank e o teste de Gehan pertencem ambos a uma classe de testes sugerida por Tarone e Ware (1977), em que a estatística de teste é dada por $\frac{\left\{\sum_{j=1}^{r} w_j (d_{1j} - e_{1j})\right\}^2}{\sum_{j=1}^{r} w_j{}^2 v_{1j}}$, com W_J constantes, que quando:

- $w_j = 1$, j = 1, ..., r é o teste log-rank;
- $w_j = n_j$, j = 1, ..., r é o teste de Gehan.

O teste log-rank é o mais potente na detecção de afastamentos da hipótese de igualdade das distribuições que sejam do tipo *riscos proporcionais*. Quando as funções de risco se cruzam, o teste log-rank pode não conseguir detectar diferenças significativas entre as curvas de sobrevivência, pelo que se deve utilizar o teste de Gehan.

No sentido de facilitar a opção por um dos dois testes, podemo-nos apoiar no facto de que se as funções de risco são proporcionais, então as respectivas funções de sobrevivência para os dois grupos não se cruzam. De facto, seja $h_1(t)$ a função de risco de um indivíduo do grupo 1 no instante t e $h_2(t)$ a função de risco de um indivíduo do grupo 2 no mesmo instante. Se as funções de risco forem proporcionais, então $h_1(t) = \varphi h_2(t)$, em que φ é uma constante que não depende de t. Ora, $h_1(t) = \varphi h_2(t) \Rightarrow \exp\left\{-\int_0^t h_1(u)du\right\} = \exp\left\{-\int_0^t \varphi h_2(u)du\right\}$, como $S(t) = \exp(-\int_0^t h(u)\, du)$ então, $S_1(t) = [S_2(t)]^{\varphi}$.

Uma vez que a função de sobrevivência toma valores entre zero e um, este resultado mostra que $S_1(t)$ é maior ou menor que $S_2(t)$ consoante φ seja maior ou menor que a unidade, para qualquer instante t. Isto significa que, se as duas funções de risco são proporcionais, então as verdadeiras funções de sobrevivência não se cruzam. Esta é uma condição necessária, mas não suficiente, para a proporcionalidade das funções de risco. No entanto, uma avaliação informal da hipótese de riscos proporcionais pode ser feita através da representação gráfica das estimativas das funções de sobrevivência. Se as estimativas das funções de sobrevivência não se cruzam, a hipótese de riscos proporcionais pode ser justificada e o teste log-rank é o teste apropriado. Como é óbvio, as estimativas das funções de sobrevivência podem-se cruzar, embora as verdadeiras funções de risco sejam proporcionais.

Outro método gráfico, que permite avaliar com mais rigor a hipótese de riscos proporcionais, é a representação gráfica das funções, $\log\left[-\log(\hat{S}(t))\right]$, para cada grupo. De facto, se $h_1(t)$ e $h_2(t)$ são funções de risco proporcionais então, $S_1(t) = [S_2(t)]^{\varphi} \Rightarrow -\log S_1(t) = \varphi[-\log S_2(t)] \Rightarrow$.

$$\Rightarrow \log[-\log S_1(t)] = \log\varphi + \log[-\log S_2(t)].$$

Portanto, o logaritmo das funções de risco cumulativas de dois indivíduos, pertencentes a grupos diferentes, apresentam uma distância constante e igual a $\log\varphi$. Então, se $\hat{S}_1(t)$ e $\hat{S}_2(t)$ são estimativas de $S_1(t)$ e de $S_2(t)$ não baseadas na hipótese de riscos proporcionais (obtidas utilizando o estimador de Kaplan-Meier), o gráfico de $\log\left[-\log\hat{S}_1(t)\right]$ *versus* t tenderá a ser paralelo ao gráfico de $\log\left[-\log\hat{S}_2(t)\right]$ *versus* t, quando $h_1(t)$ e $h_2(t)$ e são proporcionais.

REFERÊNCIAS

1 - Aalen O, Nonparametric inference for a family of counting processes. Annals of Statistics 1978;6:701-26.

2 - Carvalho M, Andreozzi V, Codeço C, Barbosa MT, Shimukura S. Análise de sobrevida. Teoria e Aplicações em Saúde. Rio de Janeiro: Editora Fiocruz; 2005.

3 - Collett D. Modelling Survival Data in Medical Research. 2nd ed. London: Chapman & Hall; 2003.
4 - Kalbfleisch JD, Prentice RL. The Statistical Analysis of Failure Time Data. New York: Wiley; 1980.
5 - Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 1958;53:457-81.
6 - Marubini E, Valsecchi MG. Analysing Survival Data from Clinical Trials and Observational Studies. Chichester: Wiley, 1995.
7 - Nelson W, Theory and applications of hazard plotting for censored failure data. Technometrics 1972;14:946-65.
8 - Tarone R, Ware J. On distribution-free tests for equality of survival distributions. Biometrika 1977;64:156-60.

Correspondência:
Dr.ª Joana Bastos
Serviço de Higiene e Epidemiologia
Faculdade de Medicina da Universidade do Porto
Alameda Prof. Hernâni Monteiro
4200-319 Porto

e-mail: joanabas@med.up.pt