MAE0514 PROVA 2 . PARTE 1 , 19/07/2021

Rubens Santos Andrade Filho, 10370336 ,

1. a) Uma variável binária adequada é

$$X_i = \begin{cases} 1, & \text{se o paciente } i \text{ tem tumor grande} \\ 0, & \text{'' '' '' moderado} \end{cases}$$

e a forma do modelo semiparamétrico de Cox fica

$$\alpha_i(t_i \mid X_i = x_i) = \alpha_0(t_i) e^{x_i \beta}$$

em que $\alpha_0(t_i) = \alpha(t_i \mid X_i = 0)$ é a função de risco basal.

1\. b) Sejam $t_{(1)} < t_{(2)} < \ldots < t_{(D)}$ tempos de falha ordenados, assumindo que não há empates. E seja $R(t_{(i)})$ o conjunto de indivíduos em risco em $t_{(i)}$.

Para esse problema, a forma da verossimilhança parcial de Cox é

$$L(\beta) \propto \prod_{i=1}^{d} \frac{e^{x_{(i)}\beta}}{\sum_{j \in R(t_{(i)})} e^{x_{(j)}\beta}}$$

em que $x_{(i)}$ refere-se à variável do item a) do paciente com tempo de falha $t_{(i)}$ $\blacksquare$

c) O escore é $U(\beta) = \dfrac{\partial l(\beta)}{\partial \beta}$, com $l(\beta) = \log L(\beta)$.

Bom,

$$l(\beta) = \log L(\beta) = \sum_{i=1}^{d} \log \left[\frac{e^{x_{(i)}\beta}}{\sum_{j \in R(t_{(i)})} e^{x_{(j)}\beta}} \right] =$$

$$= \sum_{i=1}^{d} x_{(i)}\beta - \sum_{i=1}^{d} \log \left[\sum_{j \in R(t_{(i)})} e^{x_{(j)}\beta} \right]$$

Calculemos $U(\beta)$:

$$U(\beta) = \frac{\partial l(\beta)}{\partial(\beta)} = \sum_{i=1}^{d} x_{(i)} - \sum_{i=1}^{d} \frac{u_i'}{u_i}, \quad \text{com } u_i = \sum_{j \in R(t_{(i)})} e^{x_{(j)}\beta}$$

$$\text{e } u_i' = \sum_{j \in R(t_{(i)})} x_{(j)} e^{x_{(j)}\beta}$$

$$U(\beta) = \sum_{i=1}^{d} x_{(i)} - \sum_{i=1}^{d} \left[\frac{\sum_{j \in R(t_{(i)})} x_{(j)} e^{x_{(j)}\beta}}{\sum_{j \in R(t_{(i)})} e^{x_{(j)}\beta}} \right]$$

$\blacksquare$

1.d) A matriz de informação observada é

$$I(\beta) = -\frac{\partial^2 \ell(\beta)}{\partial\beta\,\partial\beta^t} = -\frac{\partial^2 \ell(\beta)}{\partial\beta^2} \quad \text{no nosso caso.}$$

$$= -\frac{\partial U(\beta)}{\partial\beta}$$

Para simplificar a notação façamos $x_i = x_{(i)}$.

$$I(\beta) = +\frac{\partial}{\partial\beta}\left[\sum_{i=1}^{d}\left(\frac{\sum_{j\in R(t(i))} x_j e^{x_j\beta}}{\sum_{j\in R(t(i))} e^{x_j\beta}}\right)\right]$$

$$= \sum_{i=1}^{d}\left[\frac{\left(\sum_{j\in R(t(i))} x_j^2 e^{x_j\beta}\right)\cancel{\left(\sum_{j\in R(t(i))} e^{x_j\beta}\right)} - \left(\sum_{j\in R(t(i))} x_j e^{x_j\beta}\right)^2}{\left(\sum_{j\in R(t(i))} e^{x_j\beta}\right)^2}\right]$$

$$= \sum_{i=1}^{d}\left[\frac{\sum_{j\in R(t(i))} x_j^2 e^{x_j\beta}}{\sum_{j\in R(t(i))} e^{x_j\beta}} - \frac{\left(\sum_{j\in R(t(i))} x_j e^{x_j\beta}\right)^2}{\left(\sum_{j\in R(t(i))} e^{x_j\beta}\right)^2}\right] \quad \blacksquare$$

Scanned with CamScanner

e) I) $$U(0) = \sum_{i=1}^{d} x_i - \sum_{i=1}^{d} \left[\frac{\sum_{j \in R(t(i))} x_j e^{x_j 0}}{\sum_{j \in R(t(i))} e^{x_j 0}} \right]$$

$$= \sum_{i=1}^{d} x_i - \sum^{d} \left[\frac{\sum_{j \in R(t(i))} x_j}{n_i} \right]$$

em que $n_i = \sum_{j \in R(t(i))} 1 = |R(t(i))|$ é o número de idivíduos em risco em $t(i)$.

Como $x_j = 1$ se o paciente j tem tumor grande, então $n_{i1} := \sum_{j \in R(t(i))} x_j$ é o número de indivíduos em risco em $t(i)$ do grupo com tumor grande (em $t(i)$). E seja $n_{i0} = n_i - n_{i1}$ o número de indivíduos em risco do grupo com tumor moderado.

LOGO,

$$U(0) = \sum_{i=1}^{d} x_i - \sum^{d} \frac{n_{i1}}{n_i}$$

$$= \sum_{i=1}^{d} \left[x_i - \frac{n_{i1}}{n_i} \right] \blacksquare$$

e) II) $$I(0) = \sum_{i=1}^{d}\left[\frac{\sum\limits_{j\in R(t(i))} x_j^2 e^{x_j 0}}{\sum\limits_{j\in R(t(i))} e^{x_j 0}} - \frac{\left(\sum\limits_{j\in R(t(i))} x_j e^{x_j 0}\right)^2}{\left(\sum\limits_{j\in R(t(i))} e^{x_j 0}\right)^2}\right]$$

$$= \sum_{i=1}^{d}\left[\frac{n_{i1}}{n_i} - \frac{n_{i1}^2}{n_i^2}\right]$$

$$= \sum_{i=1}^{d}\left[\frac{n_{i1}}{n_i}\left(1 - \frac{n_{i1}}{n_i}\right)\right]$$

observando que

$$\sum_{j\in R(t(i))} x_j^2 = \sum_{j\in R(t(i))} x_j = n_{i1}$$

III) Obtemos a estatística do teste escore:

$$S = \frac{U^2(0)}{I(0)} = \frac{\left(\sum\limits_{i=1}^{d}\left[x_i - \frac{n_{i1}}{n_i}\right]\right)^2}{\sum\limits_{i=1}^{d}\left[\frac{n_{i1}}{n_i}\left(1 - \frac{n_{i1}}{n_i}\right)\right]}$$

que tem distribuição $\chi^2_{(1)}$ sob $H_0: \beta = 0$

2.a)

$$T_i|U \sim \text{Weibull}(\lambda = Ue^{x^T\beta}, \gamma)$$

$$\alpha(t|U,x) = Ue^{x^T\beta}\gamma t^{\gamma-1} \quad , \quad U>0$$

A função de sobrevivência de $T|U$ é

$$S(t|U) = e^{-\lambda t^{\gamma}} = \exp\{-Ue^{x^T\beta}t^{\gamma}\}$$

A função de sobrevivência marginal de T é:

$$S(t) = P(T>t) = \int_t^{\infty} f_T(s)\,ds = \int_t^{\infty}\int_0^{\infty} f_{T|U}(s|u) f_U(u)\,du\,ds =$$

$$= \int_0^{\infty}\left(\int_t^{\infty} f_{T|U}(s|u)\,ds\right) f_U(u)\,du = \int_0^{\infty} S(t|u) f_U(u)\,du =$$

$$= \int_0^{\infty} \exp\{-U\underbrace{e^{x^T\beta}t^{\gamma}}_{s}\} f_U(u)\,du = \int_0^{\infty} e^{-sU} f_U(u)\,du =$$

$$= E\left[e^{-sU}\right] = \mathcal{L}(s) = e^{-s^{\theta}} =$$

$$= e^{-e^{\theta x^T\beta}t^{\theta\gamma}}$$

Com isso, $T \sim \text{Weibull}(\lambda = e^{\theta x^T\beta}, \gamma = \theta\gamma)$

2.b) Agora, $\alpha(t|U,x) = \alpha_0(t) U e^{x'\beta}$, em que $\alpha_0(t)$ é a função de risco basal.

Então, $S(t|U) = \exp\left\{-\int_0^t \alpha(s|U)\,ds\right\}$

$$= \exp\left\{-\int_0^t \alpha_0(s) U e^{x'\beta} ds\right\}$$

$$= \exp\left\{-U e^{x'\beta} \int_0^t \alpha_0(s)\,ds\right\}$$

Análogo ao item a), a f. de sobrevivência ~~de~~ marginal de T é $S(t) = P(T>t) = \int_0^\infty S(t|u) f_U(u)\,du =$

$$= \int_0^\infty \exp\Big\{-U \underbrace{e^{x'\beta}\int_0^t \alpha_0(s)\,ds}_{s}\Big\} f_U(u)\,du =$$

$$= E[e^{-sU}] = \mathcal{L}(s) = e^{-s^\theta} =$$

$$= \exp\left\{-\left(e^{x'\beta}\int_0^t \alpha_0(s)\,ds\right)^\theta\right\}$$

$$S(t) = \exp\left\{-e^{\theta x'\beta}\left(\int_0^t \alpha_0(s)\,ds\right)^\theta\right\}$$ //

A função ~~de~~ ~~sobrevivência~~ ~~risco é~~ densidade de probabilidades é

$$f(t) = -\frac{dS(t)}{dt} = \cancel{-}\exp\left\{-e^{\theta x'\beta}\left(\int_0^t \alpha_0(s)\,ds\right)^\theta\right\}.$$
$$\left(\cancel{-}e^{\theta x'\beta}\theta\left(\int_0^t \alpha_0(s)\,ds\right)^{\theta-1}\cdot \alpha_0(t)\right)$$

$$f(t) = \underbrace{\exp\left\{-e^{\theta x'\beta}\left(\int_0^t \alpha_0(s)\,ds\right)^\theta\right\}}_{S(t)} e^{\theta x'\beta}\theta\left(\int_0^t \alpha_0(s)\,ds\right)^{\theta-1}\alpha_0(t)$$

Como $f(t) = S(t)\alpha(t)$, a função de riscos é

$$\alpha(t) = \theta\,\alpha_0(t)\, e^{\theta x'\beta}\left[\int_0^t \alpha_0(s)\,ds\right]^{\theta-1}$$ ∎

Scanned with CamScanner

2. c) Pro caso de $\theta = 1$, $\alpha(t|U) = \alpha(t)$.

Então ao ignorar a existência da variável de fragilidade U no modelo é o caso de considerar $\theta = 1$.

2,d) (T_1, T_2)

$(T_1 \perp\!\!\!\perp T_2) \mid U$

$\alpha_j(t|U) = \alpha_0(t)\, U \lambda_j$, $j=1,2$

$U>0$, $\mathcal{L}(s) = E\left[e^{-sU}\right] = e^{-s^\theta}$

$$F_{T_1|U, T_2|U}(t_1,t_2) \overset{|U}{=} F_{T_1|U}(t_1)\, F_{T_2|U}(t_2) = \left(1 - S_{T_1|U}(t_1)\right)\left(1 - S_{T_2|U}(t_2)\right)$$

$$\Rightarrow S_{T_1|U, T_2|U}(t_1,t_2|U) = S_{T_1|U}(t_1) + S_{T_2|U}(t_2) + S_{T_1|U}(t_1) \cdot S_{T_2|U}(t_2)$$

com $S_{T_j|U}(t_j) = \exp\Big\{-U\lambda_j \underbrace{\int_0^\infty \alpha_0(t)\, ds}_{A_0(t)}\Big\}$, $j=1,2$

Análogamente ao item b),

$$S_{T_1,T_2}(t_1,t_2) = \int_0^\infty S_{T_1|U, T_2|U}(t_1,t_2|U)\, f(u)\, du =$$

$$= \int_0^\infty \exp\{-U\lambda_1 A_0(t)\} f(u)\, du + \int_0^\infty \exp\{-U\lambda_2 A_0(t)\} f(u)\, du$$

$$+ \int_0^\infty \exp\{-U A_0(t)(\lambda_1+\lambda_2)\}\, f(u)\, du$$

$$= e^{-(\lambda_1 A_0(t))^\theta} + e^{-(\lambda_2 A_0(t))^\theta} + e^{-[(\lambda_1+\lambda_2)A_0(t)]^\theta}$$