Atividade 2

- Modelo de Cox estratificado

- <u>Definição</u>

- Suponha $X_1, \dots, X_p$ covariáveis que satisfazem a suposição de proporcionalidade dos riscos

$W_1, \dots, W_K$ covariáveis que não satisfazem a suposição de riscos proporcionais

Combinar $W_1, \dots, W_K$ em K^* estratos

- O modelo pode ser escrito como

$$\alpha_g(t|x) = \alpha_{0g}(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)$$

Em que:

$\alpha_g(t|x)$: É a função de taxa de falha associada a um indivíduo com covariável $\underline{X}$ pertencente ao extrato g, $g = 1, \dots, K^*$

$\alpha_{0g}(\cdot)$: Função de taxa basal associada ao extrato g

E ainda

$$S_g(t|x) = \exp\left\{-\int_0^t \alpha_g(u|\underline{x})\, du\right\} = \exp\left\{-A_{0g}(t) e^{x'\beta}\right\} = \left[S_{0g}(t)\right]^{e^{x'\beta}}$$

- Função de sobrevivência depende do estrato
- Uma função de taxa de falha para cada estrato

$$\hat{A}_{0g}(t) = \sum_{j: t_{(j)g} \le t} \frac{d_{jg}}{\sum_{l \in R_g(t_{(j)g})} e^{x'\beta}}, \quad g = 1, \dots, K^*$$

- <u>Justificativa de uso</u>

- Quando a suposição de riscos é violada para alguma covariável (um cruzamento entre as curvas das estimativas de Kaplan-Meier para tal covariável)
- Fornece um ajuste mais geral para as variáveis de confundimento
- Em comparação do modelo de de Cox estratificado com n modelos de Cox separados é que o mode estratificado tem menores erros-padrões das estimativas dos parâmetros

Verossimilhança Parcial

- Mesma ideia da construção da verossimilhança parcial do modelo de cox usual
- Constrói-se um conjunto de risco de pacientes em cada estrato

A verossimilhança parcial é dada por

$$L(\underset{\sim}{\beta}) = \prod_{g=1}^{K^*} \prod_{i=1}^{n_g} \frac{e^{X'_{g(i)}\underset{\sim}{\beta}}}{\sum_{j \in R_g(t_{(i)g})} e^{X'_j \underset{\sim}{\beta}}}$$

Em que n_g: nº de falhas no estrato g

$X_{g(i)}$: vetor de covariáveis associada à observação que falhou em $t_{(i)g}$

$R_g(t_{(i)g})$: indivíduos do estrato g em risco no instante $t_{(i)g}$

A log Verossimilhança

$$\ell(\underset{\sim}{\beta}) = \ln(L(\underset{\sim}{\beta})) = \sum_{g=1}^{K^*} \ln\left[\prod_{i=1}^{n_g} \frac{e^{X'_{g(i)}\beta}}{\sum_{j \in R_g(t_{(i)g})} e^{X'_j \beta}}\right] = \ell_1(\underset{\sim}{\beta}) + \ell_2(\underset{\sim}{\beta}) + \dots + \ell_{K^*}(\underset{\sim}{\beta})$$

$\hat{\underset{\sim}{\beta}}$ é obtido derivando $\ell(\underset{\sim}{\beta})$ em relação a $\underset{\sim}{\beta}$ e resolvendo a equação

$$\frac{\partial \ell(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = \underset{\sim}{0}$$