Cox Proportional Hazards Model

Data (input): $(\widetilde{T}_i, \Delta_i, X_{1,i}, \ldots, X_{p,i})$ $(i = 1, \ldots, n)$

- T = survival time
- C = censoring time
- $\widetilde{T} = \min(T, C)$
- $\Delta = I(T \leq C)$
- $I(\cdot)$ = indicator function

Model:

$$\lambda(t|\mathbf{X}_i) = \lambda_0(t) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_i}$$

$$\mathbf{X}_i = \begin{bmatrix} X_{1,i} \\ \vdots \\ X_{p,i} \end{bmatrix}, \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}$$

$\lambda_0(\cdot)$ = arbitrary function

Maximum (Partial) Likelihood Estimation:

$$Likelihood: \; L(\boldsymbol{\beta}) = \prod_{i=1}^{n} \left\{ \frac{e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_i}}{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j}} \right\}^{\Delta_i}$$

$$Log-likelihood: \; l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta}) = \sum_{i=1}^{n} \Delta_i \left[\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_i - \log \left\{ \sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j} \right\} \right]$$

$$Score\ function: \; \mathrm{U}(\boldsymbol{\beta}) = \frac{\partial l(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} \Delta_i \left[\mathbf{X}_i - \frac{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j} \mathbf{X}_j}{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j}} \right]$$

$$Information\ matrix: \; \mathrm{V}(\boldsymbol{\beta}) = -\frac{\partial l^2(\boldsymbol{\beta})}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}^{\mathrm{T}}}$$

$$= \sum_{i=1}^{n} \Delta_i \left[\frac{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j} \mathbf{X}_j^{\otimes 2}}{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j}} - \left\{ \frac{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j} \mathbf{X}_j}{\sum_{j=1}^{n} I(\widetilde{T}_j \geq \widetilde{T}_i) e^{\boldsymbol{\beta}^{\mathrm{T}} \mathbf{X}_j}} \right\}^{\otimes 2} \right]$$

$\mathbf{a} = \mathbf{a}\mathbf{a}^{\mathrm{T}}$

$$\textit{Maximum likelihood estimator (MLE)}: \ \mathbf{U}(\widehat{\beta}) = 0$$

Solve the equation by the Newton-Raphson algorithm; see logistic regression

Convergence criterion:

$$\left| \frac{l(\beta^{new}) - l(\beta^{old})}{l(\beta^{old})} \right| < 10^{-8}$$

- halving

- other criteria: change in β estimates

$$\mathrm{Cov}(\widehat{\beta}) = \mathbf{V}^{-1}(\widehat{\beta})$$

$$\mathrm{Var}(\widehat{\beta}_j) = j\text{th diagonal element of } \mathrm{Cov}(\widehat{\beta}), \quad j = 1, \ldots, p$$

$$\mathrm{SE}(\widehat{\beta}_j) = \mathrm{Var}^{1/2}(\widehat{\beta}_j)$$

Output: same as logistic regression