5.4 Proportional hazards models:

1. Introduction

Let $t_1, t_2, \cdots, t_n$ be the failure times associated with censor indicator $w_1, w_2, \cdots, w_n$ and the covariate vectors $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})$. Further, let $t_{(1)} \leq t_{(2)} \leq \cdots \leq t_{(m)}$ be the ordered uncensored failure times corresponding to $w_{(j)} = 1, j = 1, \cdots, m$, and $x_{(1)}, x_{(2)}, \cdots, x_{(m)}$ are the associated covariate vectors. Note (j) represents the label for the individual who dies at $t_{(j)}$.

Example (continue):

Suppose 3 covariates, # of cigarettes, gender and age, are of interest.

We have the following tables:

Labels:
$$(1) = 1, (2) = 4, (3) = 6, (4) = 8.$$

Failure times:

t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8
6	7	9.5	7	10	7	6	11

$t_{(1)} = t_1$	$t_{(2)} = t_4$	$t_{(3)} = t_6$	$t_{(4)} = t_8$
6	7	7	11

Covariates:

x_1	x_2	x_3	x_4
$(20, 1, 45)$	$(30, 1, 20)$	$(50, 0, 38)$	$(40, 1, 26)$
x_5	x_6	x_7	x_8
$(3, 0, 42)$	$(40, 0, 17)$	$(60, 1, 25)$	$(10, 0, 29)$

$x_{(1)} = x_1$	$x_{(2)} = x_4$	$x_{(3)} = x_6$	$x_{(4)} = x_8$
$(20, 1, 45)$	$(40, 1, 26)$	$(40, 0, 17)$	$(10, 0, 29)$

The proportional hazards model specifying the hazard at time *t* for an individual whose covariate vector is x is given by
$$\lambda(t) = \lambda_0(t) \cdot exp(x\beta),$$
where $\lambda_0(t)$ is referred to as the baseline hazard function. The model implies that the ratio of hazards for two individuals depends on the difference between their linear predictor at any time. For example, for individuals with covariate vectors x_1

and x_2, respectively, the ratio of hazards for the two individuals is

$$\frac{\lambda_0(t) \cdot exp(x_1\beta)}{\lambda_0(t) \cdot exp(x_2\beta)} = exp[(x_1 - x_2)\beta]$$

which only depends on the difference between their linear predictor.

The exact likelihood function,

$$L[\beta, \lambda_0(t)] = \prod_{I=1}^{n}[\lambda_0(t_i)]^{w_i}S(t_i)$$

$$= \prod_{I=1}^{n}[\lambda_0(t_i) \cdot exp(x_i\beta)]^{w_i} exp\left[-\int_0^{t_i}\lambda_0(x) \cdot exp(x_i\beta)dx\right]$$

depends on both the nonparametric function $\lambda_0(t)$ and the parameter β. Thus, it might be difficult to estimate $\lambda_0(t)$ and β simultaneously. To resolve the problem, one solution is to find a "modified function" involving only β. Then, we can estimate β or make statistical inference about β based on the modified likelihood function. Thus, the effect of the covariate vector x can be assessed.

2. Partial likelihood function

Denote $R(t)$ to be the set of individuals who are alive and uncensored at a time just prior to t. $R(t)$ is called the rik set.

Example (continue):

$R(6)$	$R(7)$	$R(9.5)$	$R(10)$	$R(11)$
$\{1, 2, 3, 4, 5, 6, 7, 8\}$	$\{2, 3, 4, 5, 6, 8\}$	$\{3, 5, 8\}$	$\{5, 8\}$	$\{8\}$

The partial likelihood function is

$$L_p(\beta) = \prod_{j=1}^{m}\frac{exp(x_{(j)}\beta)}{\sum_{l \in R(t_{(j)})}exp(x_l\beta)} = \prod_{i=1}^{n}\left[\frac{exp(x_i\beta)}{\sum_{l \in R(t_i)}exp(x_l\beta)}\right]^{w_i}.$$

The estimate $\widehat{\beta}$ can be obtained by numerically solving

$$\frac{\delta L_p(\beta)}{\delta\beta} = 0.$$

The variance-covariance matrix of $\widehat{\beta}$ can be estimated by the inverse of the observed information matrix evaluated at $\widehat{\beta}$, i.e.,

$$I^{-1}(\widehat{\beta}) = \left[\frac{-\delta^2 log[L_p(\beta)]}{\delta\beta^t\delta\beta}\right]^{-1}_{\beta=\widehat{\beta}}.$$

Thus, to test $H_0: \beta_k = 0$, **the Wald statistic**

$$\frac{\widehat{\beta}_k}{s.e.(\widehat{\beta}_k)} \sim N(0, 1)$$

under $H_0: \beta_k = 0$ **can be used, where** $s.e.(\widehat{\beta}_k)$ **is the standard error of the partial likelihood estimate** $\widehat{\beta}_k$.

Note:

Intuitively, given $R(t_{(j)})$,

$$\frac{P(patient\ (j)\ die\ at\ t_{(j)})}{\sum_{l \in R(t_{(j)})} P(patient\ l\ die\ at\ t_{(j)})}$$

$$= \frac{\lambda_{(j)}(t_{(j)})}{\sum_{l \in R(t_{(j)})} \lambda_l(t_{(j)})} = \frac{\lambda_0(t_{(j)}) exp(x_{(j)}\beta)}{\sum_{l \in R(t_{(j)})} \lambda_0(t_{(j)}) exp(x_l\beta)}$$

$$= \frac{exp(x_{(j)}\beta)}{\sum_{l \in R(t_{(j)})} exp(x_l\beta)}.$$

Example (continue):

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}.$$

The partial likelihood function is

$$L_p(\beta) = \prod_{j=1}^{4} \frac{exp(x_{(j)}\beta)}{\sum_{l \in R(t_{(j)})} exp(x_l\beta)} = \prod_{i=1}^{8} \left[\frac{exp(x_i\beta)}{\sum_{l \in R(t_i)} exp(x_l\beta)} \right]^{w_i}.$$

As $j = 1$,

$$\frac{exp(x_{(1)}\beta)}{\sum_{l \in R(t_{(1)})} exp(x_l\beta)} = \frac{exp(x_1\beta)}{\sum_{l \in R(6)} exp(x_l\beta)} = \frac{exp(x_1\beta)}{\sum_{l \in \{1,2,\cdots,8\}} exp(x_l\beta)} = \frac{exp(x_1\beta)}{\sum_{l=1}^{8} exp(x_l\beta)}.$$

As $j = 2$,

$$\frac{exp(x_{(2)}\beta)}{\sum_{l \in R(t_{(2)})} exp(x_l\beta)} = \frac{exp(x_4\beta)}{\sum_{l \in R(7)} exp(x_l\beta)} = \frac{exp(x_4\beta)}{\sum_{l \in \{2,3,4,5,6,8\}} exp(x_l\beta)}$$

$$= \frac{exp(x_4\beta)}{exp(x_2\beta) + exp(x_3\beta) + exp(x_4\beta) + exp(x_5\beta) + exp(x_6\beta) + exp(x_8\beta)}.$$

As $j = 3$,

$$\frac{exp(x_{(3)}\beta)}{\sum_{l \in R(t_{(3)})} exp(x_l\beta)} = \frac{exp(x_6\beta)}{\sum_{l \in R(7)} exp(x_l\beta)} = \frac{exp(x_6\beta)}{\sum_{l \in \{2,3,4,5,6,8\}} exp(x_l\beta)}$$

$$= \frac{exp(x_6\beta)}{exp(x_2\beta) + exp(x_3\beta) + exp(x_4\beta) + exp(x_5\beta) + exp(x_6\beta) + exp(x_8\beta)}.$$

As $j = 4$,

$$\frac{exp(x_{(4)}\beta)}{\sum_{l \in R(t_{(4)})} exp(x_l\beta)} = \frac{exp(x_8\beta)}{\sum_{l \in R(11)} exp(x_l\beta)} = \frac{exp(x_8\beta)}{\sum_{l \in \{,8\}} exp(x_l\beta)} = \frac{exp(x_8\beta)}{exp(x_8\beta)} = 1.$$

Thus,

$$L_p(\beta) = \frac{exp[(x_1 + x_4 + x_6)\beta]}{\left[\sum_{l=1}^{8} exp(x_l\beta)\right]\left[\sum_{l \in \{2,3,4,5,6,8\}} exp(x_l\beta)\right]^2}.$$