Lecture 8
Stratified Cox Model

So far, we've been considering the following Cox model (with possibly time-dependent covariates):

$$\lambda(t|\mathbf{Z}(t)) = \lambda_0(t)\ \exp\{\boldsymbol{\beta}'\mathbf{Z}(t)\}$$

Here the baseline hazard $\lambda_0(t)$ is common to all the individuals in a study.

But there are cases where this appears to be too strong an assumption. For example, in colon cancer, stage II and stage III diseases have very different prognosis. In a therapeutic study, we may expect similar amount of improvement by using a new treatment, but the baseline hazard for stage II and III cancer should be allowed to be different.

Suppose that there is a factor with K levels. In a **stratified Cox model**, the hazard for an individual from stratum k is

$$\lambda_k(t|\mathbf{Z}(t)) = \lambda_{0k}(t)\ \exp\{\boldsymbol{\beta}'\mathbf{Z}(t)\},$$

where $\lambda_{0k}(t)$ is the baseline hazard for stratum k, $k = 1, ..., K$.

What is the interpretation of $\boldsymbol{\beta}$?

In the colon cancer example, suppose $Z = 1$ for new treatment, 0 for old treatment, and stratum $k = 1$ for stage II, $k = 2$ for stage III. The hazard ratio of new vs. old treatment is still e^{β} within each stratum. But the baseline hazards (i.e. for old treatment group) for the 2 strata are different, and so are the hazards for the new treatment group between the 2 strata (why?).

Inference under the stratified Cox model is still carried out via the partial likelihood. But now, the contribution to the likelihood at the time when individual i from stratum k fails, is computed only within stratum k.

That is, it is the conditional probability of choosing individual i to fail, given the risk set and the stratum k, and that one failure is to occur:

$$\frac{e^{\boldsymbol{\beta}'\mathbf{Z}_{ki}(X_{ki})}}{\sum_{j \in R_k(X_{ki})} e^{\boldsymbol{\beta}'\mathbf{Z}_{kj}(X_{ki})}}$$

The partial likelihood is the product over all failures from all strata:

$$L(\boldsymbol{\beta}) = \prod_{k=1}^{K} \prod_{i=1}^{n_k} \left\{ \frac{e^{\boldsymbol{\beta}'\mathbf{Z}_{ki}(X_{ki})}}{\sum_{j \in R_k(X_{ki})} e^{\boldsymbol{\beta}'\mathbf{Z}_{kj}(X_{ki})}} \right\}^{\delta_{ki}}$$

where n_k is the number of subjects in stratum k.

Another way to see it is

$$L(\boldsymbol{\beta}) = \prod_{k=1}^{K} L_k(\boldsymbol{\beta})$$

where $L_k(\boldsymbol{\beta})$ is the partial likelihood from stratum k.

In coxph() there is a control parameter 'strata='.

This also solves the problem of <u>stratified log-rank test</u> for P-sample ($P \geq 2$) comparison that we did not quite talk about before. (How?)

Stratification is a way to deal with non-PH

Consider 3 models, where 'RENAL' indicates whether there is normal renal function:

$$\lambda_k(t|\mathbf{Z}) = \lambda_{0k}(t)\ \exp\{\beta_1 \cdot \text{TREAT},\} \tag{1}$$

where $k = 0, 1$, stratified by RENAL.

$$\lambda(t|\mathbf{Z}) = \lambda_0(t)\ \exp\{\beta_1 \cdot \text{TREAT},\} \tag{2}$$

$$\lambda(t|\mathbf{Z}) = \lambda_0(t)\ \exp\{\beta_1 \cdot \text{TREAT} + \beta_2 \cdot \text{RENAL}.\} \tag{3}$$

How do these models compare?

When do we want to use stratified models?