Extensions of the Proportional Hazards Model[1]

STA312 Spring 2019

[1] See last slide for copyright information.

Background Reading

- Section 8.2 in Chapter 8, and Chapter 9 in *Applied Survival Analysis Using R* by Dirk Moore
- *Modeling Survival Data: Extending the Cox Model* (2000) by Terry Thereau and Patricia Grambsch

Overview

1. Stratification
2. Time Dependent Coefficients
3. Frailty Models
4. Competing Risks

Stratification

- *Strata* are levels, or layers, like a cake.
- Think of a stratum as a sub-population.
- We often consider an independent random sample from each stratum.
- For example, companies in Canada, the U.S. and Mexico.
- For proportional hazards regression, it may not make sense to assume that the baseline hazard functions are the same in all the strata.
- Multi-center clinical trials, with different patient populations in each medical center.
- Assume a separate baseline hazard function in each stratum.

Partial Likelihood Function for a Stratified Model

There are k strata

$$\mathrm{PL}(\boldsymbol{\beta}) = \prod_{\ell=1}^{k} \left(\prod_{i=1}^{D} \frac{e^{\mathbf{x}_{i,\ell}^\top \boldsymbol{\beta}}}{\sum_{j \in R_{i,\ell}} e^{\mathbf{x}_{j,\ell}^\top \boldsymbol{\beta}}} \right)$$

- Separate baseline hazards are cancelling within the parentheses.
- Note that the parameter vector $\boldsymbol{\beta}$ is the same in all strata.
- This condition can be relaxed.
- And tested with a partial likelihood ratio test.
- But there is no direct test for differences between strata.

Sample Code for Stratification

```
coxph(Surv(time, delta) ~
     age + strata(ascites) + bili + protime + albumin)
```

Time Dependent Coefficients

- The regression coefficients β_j might depend on time: $\beta_j(t)$.

$$h(t) = h_0(t) \exp\{\mathbf{x}^\top \boldsymbol{\beta}(t)\}$$

- This is attractive, but maximum likelihood estimation of the function (actually, p functions) would require lots of failures at every possible time point.
- Solution: Estimate the function another way, and then put the estimate into the partial likelihood.

Schoenfeld Residuals

For each $0 < t_1, < \ldots < t_k < \ldots < t_D$

- $\mathbf{s}_k$ is the vector of p Schoenfeld residuals at time k.
- $s_{k,j}$ is the Schoenfeld residual for variable j at time k.
- $\mathbf{s}_k^*$ are the scaled Schoenfeld residuals.
- Grambsch and Thereau have shown $E(s_{k,j}^* + \widehat{\beta}_j) \approx \beta_j(t_k)$.
- They suggest using $s_{k,j}^* + \widehat{\beta}_j$ to estimate $\beta_j(t)$ at t_k.
- If a plot of the Schoenfeld residuals against time looks constant, no problem.
- If the plot shows a trend, it suggests β_j is a function of time.
- And the proportional hazards assumption is wrong.

Testing proportional hazards using Schoenfeld residuals

- Have a plot of the Schoenfeld residuals against time.
- Test whether the correlation equals zero.
- Transformations of the t axis (scaling) allow curves.
- For example, check correlation of the residuals against $\log(t)$.

Estimation for a fixed $\beta_j(t)$

Using partial likelihood

- $\beta_j(t)$ is assumed "known," but usually it's a guess based on residual plots.
- For simplicity, consider a single explanatory variable.
- Original model: $h(t_i) = h_0(t_i) \exp\{\beta x_i\}$.
- Create a time-varying covariate that just equals time, or a function of time $g(t)$ like $\log(t)$.
- Replace x_i by the "interaction term" $x_i \, g(t_i)$.
- Model for the hazard is now $h(t_i) = h_0(t_i) \exp\{\beta g(t_i) \, x_i\}$.
- The function $\beta(t) = \beta g(t_i)$.
- The β part is unknown, and is estimated as usual by maximum partial likelihood.
- So really you are assuming that the form of $\beta(t)$ is known, but only up to multiplication by a constant.

Sample Code for Time-Dependent Coefficients

```
coxph( Surv(time,status) ~ celltype + tt(karno),
          tt = function(x,t, ...){x*log(t)} )
```

```
loginter = function(x,t,...) {x*log(t)}

coxph(Surv(time,status) ~ celltype + tt(karno),
          tt = loginter)
```

Frailty Models

Within-cases, Random effects

- A single unit may contribute more than one event, like several seizures.
- Randomly assign one eye to experimental condition, one to control. Response variable is time to blindness.
- Some groups of patients are surely not independent, like several female relatives of a breast cancer patient.
- The reason for the term "frailty" is the idea that individuals (and units) have a characteristic that is their own relative chance of failure.
- Frail means weak – more likely to die.

The Frailty Model

Random effects

The hazard at time j for cluster i is $h_{i,j}(t_{i,j}) = h_0(t_{i,j})\,\omega_i \exp\{\mathbf{x}_{i,j}^\top \boldsymbol{\beta}\}$.

- $\omega_i > 0$ is a *random effect*.
- The clusters (individuals, families, whatever) are randomly sampled from some population, and the hazard is multiplied by the same quantity ω_i for every member of cluster i.
- If $\omega_i = 2$, it means every member of cluster i is quite frail. Their hazards are all multiplied by 2.
- Think of it as a "random shock."
- Shock is random because clusters are assumed to be randomly sampled from some population.
- So $\omega_i > 0$ comes from some (assumed) probability distribution.
- Gamma and log-normal are typical choices.
- For log-normal(0,σ^2), the parameter vector is $(\boldsymbol{\beta}, \sigma^2)$.

Log-Normal Random Effects

Instead of writing $h_{i,j}(t_{i,j}) = h_0(t_{i,j})\,\omega_i\,\exp\{\mathbf{x}_{i,j}^\top \boldsymbol{\beta}\}$

Another way to write the hazard is

$$h_{i,j}(t_{i,j}) = h_0(t_{i,j})\exp\{\sigma z_i + \mathbf{x}_{i,j}^\top \boldsymbol{\beta}\},$$

where z_i is standard normal.

- σ is like another regression coefficient.
- Interpretation: If the random effect is one standard deviation above the mean (so $z_i = 1$), then the hazard is multiplied by e^σ.

Sample Code for Frailty Models

`me` stands for mixed effects

```
install.packages("coxme",dependencies=TRUE) # Only need to do
library(coxme)

(Surv(age, brcancer) ~ mutant + (1|famID), data=ashkenazi)



coxph(Surv(y, uncens) ~ trt) # Just treatment

# Add random effect for medical center
coxme(Surv(y, uncens) ~ trt + (1|center))

# Random effect of treatment nested within medical center
coxme(Surv(y, uncens) ~ trt + (1 | center/trt))
```

Rich specification of mixed models as in `lmer`.

Copyright Information

This slide show was prepared by Jerry Brunner, Department of Statistics, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. The LaTeX source code is available from the course website:

http://www.utstat.toronto.edu/~brunner/oldclass/312s19