The Stratified Proportional Hazards Model

Statistics 255 - Survival Analysis

Presented February 23, 2016
The Stratified Proportional Hazards Model

Example of two ways to stratify

▶ Suppose a confounder $C$ has 3 levels on which we would like to stratify when comparing $\lambda(t|E)$ and $\lambda(t|\bar{E})$ where $E$ is an indicator of “exposure”

▶ Dummy variable stratification: Consider

$$x_E = \begin{cases} 1 & E \\ 0 & \bar{E} \end{cases} \quad c_2 = \begin{cases} 1 & \text{level 2} \\ 0 & \text{else} \end{cases} \quad c_3 = \begin{cases} 1 & \text{level 3} \\ 0 & \text{else} \end{cases}$$

▶ Now suppose we fit the following proportional hazards model:

$$\lambda(t) = \lambda_0(t)e^{\beta_E x_E + \beta_2 c_2 + \beta_3 c_3}$$
The Stratified Proportional Hazards Model

Example of two ways to stratify

- From this, we have

<table>
<thead>
<tr>
<th>Level 1 of C:</th>
<th>exposed : $\lambda(t) =$</th>
<th>unexposed : $\lambda(t) =$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2 of C:</td>
<td>exposed : $\lambda(t) =$</td>
<td>unexposed : $\lambda(t) =$</td>
</tr>
<tr>
<td>Level 3 of C:</td>
<td>exposed : $\lambda(t) =$</td>
<td>unexposed : $\lambda(t) =$</td>
</tr>
</tbody>
</table>

$\Rightarrow$ HR =

*Key point: Stratum-to-stratum differences in \( \log \lambda(t) \) for unexposed have arbitrarily different magnitudes, but they are parallel (assumption).
The Stratified Proportional Hazards Model

Example of two ways to stratify

- **“True” stratification**: Alternatively we could also consider the following proportional hazards model:

\[ \lambda(t) = \lambda_{0i}(t)e^{\beta E x_E}, \quad i = 1, 2, 3 \]

- In this case we have different baseline hazards for each stratum

- Note: This is the implicit model for the stratified logrank test
The Stratified Proportional Hazards Model

Example of two ways to stratify

- From the “true stratification" model, we have

  **Level 1 of C:**
  - exposed: \( \lambda(t) = \)
  - unexposed: \( \lambda(t) = \)
  \( \Rightarrow \text{HR=} \)

  **Level 2 of C:**
  - exposed: \( \lambda(t) = \)
  - unexposed: \( \lambda(t) = \)
  \( \Rightarrow \text{HR=} \)

  **Level 3 of C:**
  - exposed: \( \lambda(t) = \)
  - unexposed: \( \lambda(t) = \)
  \( \Rightarrow \text{HR=} \)

*Key point: Stratum-to-stratum differences in \( \log \lambda(t) \) for unexposed are completely arbitrary (ie. no assumptions).
The Stratified Proportional Hazards Model

Example of two ways to stratify

Graphically, we can think of the difference as follows:

- **Dummy Variable**
  \[ \lambda(t) = \lambda_0(t)e^{\beta X + \beta_2 c_2 + \beta_3 c_3} \]

- **“True” Stratification**
  \[ \lambda(t) = \lambda_0(t)e^{\beta X} \]

Ex: 6-MP study

Partial likelihood for the stratified model

Assessing the PH assumption

Summary
The Stratified Proportional Hazards Model

The stratified proportional hazards regression model

- The general stratified proportional hazards regression model is:

\[
\lambda_i(t \mid x_i, \text{stratum } g) = \lambda_g(t) \exp(\beta^T x_i)
\]

for strata \( g = 1, \ldots, G \).

- A \textit{semi-parametric} model with a non-parametric baseline hazard function \textit{within each stratum}

- Alternatively, it can be useful because the hazard can difficult to model as a function of complicated set of adjustment covariates
The Stratified Proportional Hazards Model

## Considerations

- Under dummy variable stratification, the proportional stratum-to-stratum hazards assumption may not be correct
  - The confounder (adjustment/stratification variable) may be inadequately controlled, and may still confound

- Proportionality assumption can be checked by examining stratum-specific log $\Lambda(t)$ plots (coming up)

- True stratification is a more thorough adjustment as long as observations within each level are homogenous
  - If the confounder can be measured continuously, control for confounding might be achieved with the continuous covariate adjustment
The Stratified Proportional Hazards Model

Considerations

- If the confounder is controlled using true stratification, there is no way to estimate a summary relative risk comparing two levels of the confounder.

- True stratification generally requires more data to obtain the same precision in coefficient estimates (bias-variance trade-off).
The Stratified Proportional Hazards Model

Example: 6-MP Data

- Example: Recall that in the 6-MP leukemia data (Section 1.2 in K & M):
  - 21 patient pairs
  - Interest on the effect of the treatment 6-MP on survival, and
  - Matching was on remission status (remstat) and hospital (pairid)

- Recall also, that matched pairs are like many strata

- Goal: Estimate the relative risk for 6-MP versus placebo, adjusting for hospital and remission status
The Stratified Proportional Hazards Model

Example: 6-MP Data

- We can conduct the stratified analysis accounting for matching by including a `strata()` term

```r
> fit <- coxph(Surv(time, irelapse) ~ sixmp + strata(pairid),
  data=sixmpLong)
> summary(fit)
Call:
coxph(formula = Surv(time, irelapse) ~ sixmp + strata(pairid),
  data = sixmpLong)
n= 42

coef  exp(coef) se(coef)      z  Pr(>|z|)     
sixmp -1.792    0.167    0.624    -2.87   0.0041 **

---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

exp(coef) exp(-coef) lower .95 upper .95
sixmp     0.167     6.0 0.0491    0.566

Rsquare= 0.247  (max possible= 0.5 )
Likelihood ratio test= 11.9  on 1 df,  p=0.000565
Wald test     = 8.26  on 1 df,  p=0.00406
Score (logrank) test = 10.7  on 1 df,  p=0.00106
```
The Stratified Proportional Hazards Model

Example: 6-MP Data

Conclusion: the *within-stratum* estimated relative risk for 6-MP versus placebo is 0.17, 95% CI [0.05,.57], indicating that the risk of relapse is about 6 times higher in the placebo than in the 6-MP group, after adjusting for any effects due to hospital and remission status.
The Stratified Proportional Hazards Model

Example: 6-MP Data

▶ For comparison, let's consider the unmatched (incorrect) analysis

```r
> fit2 <- coxph( Surv(time, irelapse) ~ sixmp, data=sixmpLong )
> summary( fit2 )
Call:
coxph(formula = Surv(time, irelapse) ~ sixmp, data = sixmpLong)

n= 42

coef exp(coef)  se(coef)     z  Pr(>|z|)  
sixmp  -1.572  0.208  0.412 -3.81  0.00014 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1  1

exp(coef) exp(-coef) lower .95 upper .95
sixmp  0.208   4.82    0.0925   0.466

Rsquare= 0.322  (max possible= 0.988 )
Likelihood ratio test= 16.4  on 1 df,  p=5.26e-05
Wald test     = 14.5  on 1 df,  p=0.000138
Score (logrank) test = 17.2  on 1 df,  p=3.28e-05
```
Example: 6-MP Data

▶ **(Incorrect)** Conclusion: the estimated relative risk for 6-MP versus placebo is 0.208, indicating that the risk of relapse is about 4.8 times higher in the placebo than in the 6-MP group.

▶ Incorrect because not at all clear what the reference population is
Partial Likelihood for the Stratified PH Model

Recall that the partial likelihood for the ordinary proportional hazards model is

\[ L_P = \prod_{\text{failure times } j} \left( \frac{\exp(\beta^T x_{(j)})}{\sum_{i \in R_{(j)}} \exp(\beta^T x_i)} \right) \]

\( L_P \) compares the covariate \( x_{(j)} \) of the failed subject to the covariates \( x_i \) in the risk set \( R_{(j)} \) at time \( t_{(j)} \).

The stratified partial likelihood combines partial likelihoods \( L_{Pg} \) across strata \( g = 1, \ldots, G \):

\[ L_P = \prod_{g=1}^{G} L_{Pg} \]
Partial Likelihood for the Stratified PH Model

- Alternatively, who is in the “effective” risk set at time $t_{(j)}$?
  - A: The subject failing at $t_{(j)}$ and other subjects not yet failed, but also in the same stratum (i.e., in stratum $g$)

- Suppose $t_{(j)}$ is a failure time for a subject in stratum $g$

- Define $R_{(j)}^g$ to be the set of subjects in stratum $g$ and at risk for failure at $t_{(j)}$

- Then the stratified partial likelihood (in the absence of ties) is

$$L_P = \prod_{\text{failure times } j} \left( \frac{\exp(\beta^T x_{(j)})}{\sum_{i \in R_{(j)}^g} \exp(\beta^T x_i)} \right)$$
Assessing the PH Assumption

Example: Bone marrow transplant data

- Ex: Bone marrow transplant data (Section 1.3 in K & M)
  - Question: For leukemia patients, does “high-risk” versus “low-risk” sub-classification have any prognosticative meaning or value, after adjusting for other baseline clinical information?
  - Data: time (in days) of disease-free survival
    - \textit{waittime}: the waiting time to transplant (days)
    - \textit{fab}: FAB grade of disease (4 or 5, coded 0 or 1)
    - \textit{i.mtx}: MTX, a graph-vs-host disease prophylactic treatment
    - \textit{agep}: age of the patient
    - \textit{aged}: age of the donor
    - \textit{g}: Indicator of whether patient has acute myeloid leukemia (AML)
  - Prior model building already performed...
### Assessing the PH Assumption

**Example: Bone marrow transplant data**

- **Ex: Bone marrow transplant data (Section 1.3 in K & M)**
  - **Question:** Adjusting for other covariates, does the effect of `imtx` follow a proportional hazards model?
  - **Method:** If the proportional hazards assumption holds, the log cumulative hazards for the two MTX groups, adjusting for other covariates $x_i$, should be roughly parallel.
Assessing the PH Assumption

Example: Bone marrow transplant data

- Consider strata defined by $\text{imt} \times = 0$ or 1 and model

$$\lambda_i(t \mid x_i, \text{imt} \times_i) = \lambda_{\text{imt} \times_i}(t) \exp(\beta^T x_i)$$

where $x_i$ contains the other covariates

- Recall from Lecture 4: If proportional hazards holds for $\text{imt} \times$, then

$$\log \Lambda_{\text{imt} \times = 1}(t) = \log(\phi) + \log \Lambda_{\text{imt} \times = 0}(t)$$

where $\phi = \exp(\beta^T x)$

- So, if the baseline log cumulative hazards are not parallel this is evidence of non-proportional hazards
Assessing the PH Assumption

Example: Bone marrow transplant data

- R will estimate the \textit{within-stratum} baseline hazard when a stratified analysis is fit

```r
> ##
> ######
> ##### Read in BMT data
> ######
> ##

> ##
> ######
> Fit the stratified cox model and plot
> ######
> ##
> fit <- coxph( Surv(tnodis,inodis) ~ agep*aged + g + strata(imtx),
+ data=bmt )
> plot( survfit(fit), fun="cloglog",
+ xlab="Time from Transplantation (logarithmic scale)",
+ ylab="Log-Cumulative Hazard Function" )
```
Assessing the PH Assumption

Example: Bone marrow transplant data
Assessing the PH Assumption

<table>
<thead>
<tr>
<th>Example: Bone marrow transplant data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conclusion: Hard to tell but not strong evidence of non-proportional hazards’</td>
</tr>
<tr>
<td>⇒ This is a good look at gross departures, but it is far from a formal test...More later!</td>
</tr>
</tbody>
</table>
The Stratified Proportional Hazards Model

Summary

- $\beta_k$ is the change in the log-hazard function, i.e. the log-relative risk, associated with a unit difference in $x_k$, for subjects in the same stratum and holding other covariates constant.

- Therefore, the contrast expressed by $\beta_k$ is adjusted for all other covariates in the model as well as any variables contained in the stratification variable.

- The key assumption is that the effect of covariate $x_{ik}$ is the same in each stratum.

  - If it is not, i.e. there is an interaction between the stratification variable and the predictor of interest, stratum-specific estimates should be presented.
### Summary

- With $p$ covariates in the model, plots based on a model stratified on values of *new* covariate $x_{i,p+1}$ can help assess the validity of the proportional hazards assumption for $x_{i,p+1}$.

- The stratified partial likelihood model works by only allowing subjects from the same stratum to be in the same risk set.