Chapter 7

Extensions of the Proportional Hazards Model

Assumptions of the Cox Model

- The Cox proportional hazards model is

$$h(t \mid \beta, x) = h_0(t) \cdot e^{\beta x}$$

- When we fit this model we are assuming
  - A common baseline hazard: $h_0(t)$
  - "Proportional hazards" – the effect of a covariate x does not depend on time
  - The value of x is constant over time
- We can assess whether the proportional hazards assumption is reasonable, but what should we do if it isn't appropriate to assume proportional hazards?

The Proportional Hazards Assumption

- One way we saw to assess the proportional hazards assumption is to include an interaction between time and a covariate in the model and check for statistical significance
  - $\log\big(h(t|\beta, x)\big) = \log(h_0(t)) + \beta x + \gamma x t$
  - Test the hypothesis H_0: $\gamma = 0$ against H_1: $\gamma \neq 0$.

- If the coefficient for the interaction term is statistically significant, then we can just keep that covariate in our model
  - This approach is sensitive to the function of time chosen
  - Results can be difficult to interpret in some cases

The Proportional Hazards Assumption

- Example: Interested in the effect of a treatment on risk of relapse in leukemia patients
- Model:

$$\log(h(t \mid \beta, x)) = \log(h_0(t)) + \beta_L \cdot x_L + \beta_T \cdot x_T + \beta_G \cdot x_G + \gamma \cdot x_G \cdot \log(t)$$

x_L + log WBC, x_T = 1 for new treatment, x_G = 1 for female

  - $\hat{\gamma}$ = -1.91, p = 0.038 so the proportional hazards assumption for gender may not be appropriate.
  - But what does it mean to say that the effect of gender is changing over time?

The Stratified Proportional Hazards Model

- Another way of adjusting for non-proportional hazards is to use a stratified Cox model.
  - Use the covariate that violates the proportional hazards assumption to form strata
  - Fit a Cox proportional hazards model within each strata
  - Proportional hazards assumption assumed to hold for each of the remaining covariates
  - The covariate you used to stratify your data is not included in the model

The Stratified Proportional Hazards Model

- Requirements:
  - Covariate needs to be categorical so that we can stratify
    - If covariate is continuous, can break up into categories first
  - Because we can't include the covariate in the model, shouldn't stratify using a covariate that you actually care about
- When you stratify using a covariate, you are essentially taking its effect and its interaction with time and absorbing them into the baseline hazard function
  - Each stratum will have it's own baseline hazard function

The Stratified Proportional Hazards Model

- The stratified Cox proportional hazards model
  - Assume there are *S* strata
  - The hazard function for the r^{th} stratum is

    $$\log(h_r(t \mid \beta, x)) = \log(h_{r0}(t)) + \beta x, \; r = 1, 2, \ldots, S$$

  - Note that β is assumed to be the same for all of the strata
    - Can include an interaction between x and the stratifying covariate to get around this assumption

The Stratified Proportional Hazards Model

- To calculate the partial likelihood for the stratified Cox proportional hazards model:
  - First find the partial likelihood for each stratum

$$L_{rp}(\beta) = \prod_{i \in Stratum\ r} \left(\frac{e^{\beta x_i}}{\sum_{\substack{j \in Stratum\ r \\ \cap\ j \in R(t_i)}} e^{\beta x_j}} \right)^{c_i}$$

Subjects who both belong to stratum *r* and are still at risk at time t_i

  - Multiple the partial likelihood for each stratum together

$$L_p(\beta) = \prod_{r=1}^{S} L_{rp}(\beta)$$

The Stratified Proportional Hazards Model

- Example: In the leukemia data set, the treatment effect is our primary concern.
  - The proportional hazards assumption is not justified for gender
    - We have reason to believe that the shape of the baseline hazard function differs for men and women
  - We want to control for the effect of gender when estimating the treatment effect, but we're not that interest in the effect of gender itself
- The model:
  - Hazard for men: $h_M(t \mid \beta , x) = h_{M0}(t) \cdot e^{\beta \cdot Trt}$
  - Hazard for women: $h_W(t \mid \beta , x) = h_{W0}(t) \cdot e^{\beta \cdot Trt}$

The Stratified Proportional Hazards Model

- Using this model, we can't calculate a hazard ratio for men vs women.
- We've also assumed the effect of treatment is the same for men and women
  - If we think the effect of the treatment might vary by gender, we can include an interaction term

The Stratified Proportional Hazards Model

- Advantages of stratifying:
  - Easy way to address non-proportionality in a covariate
  - Do not have to choose a function of time to use in an interaction between the covariate and time
  - May be easier to interpret than an interaction term with time
  - Can stratify over multiple variables

The Stratified Proportional Hazards Model

- Disadvantages of stratifying:
  - Can only stratify using a covariate that is of secondary importance
  - No way to test for the effect of the stratifying covariate
    - It is not legitimate to compare the log-likelihoods for models with and without a stratifying variable.
  - When estimating stratum-specific covariates, you break your sample into *S* smaller ones so will lose power

Time-Varying Covariates

- So far we have assumed that the values of the covariates for each subject do not change over the course of the study or observation period
- We may want to include a covariate whose can change. These covariates are called time-dependent or time-varying covariates
  - Example: Outcome is the risk of relapse for patients with multiple sclerosis.
  - Can give patients an MRI scan at the start of the study to assess their lesion count, but may also continuing scanning patients at regular intervals to keep an up-to-date lesion count

Time-Varying Covariates

- Two basic types of time-varying covariates:
  - Internal: changes in the covariate are subject specific
    - Examples: blood pressure, lesion count, white blood cell count
    - Usually have to monitor the subject in order to know that a change in the value of the covariate occurred.
  - External: changes in the covariate occur at the study or environmental level
    - Examples: season, year of recruitment, treatment crossover
    - Know when these changes occur because they usually have nothing to do with the subject

Time-Varying Covariates

- Notation:
  - Fixed covariate: x
  - Time-varying covariate: x(t) = value of covariate *x* at time *t*
- x(t) can be defined using any information about the individual up and including time *t*
  - For example, if you want to create a time-dependent covariate using employment history, your covariate values could be
    - Whether the person is currently employed
    - Whether person was employed in the previous month
    - Number of months worked in the past year
  - Can't use information from the 'future'

Time-Varying Covariates

- Cox model that includes time-varying covariates

$$\log\big(h(t|\beta, x)\big) = \log(h_0(t)) + \beta \cdot x(t)$$

  - Notice that while the value of x may change over time, the effect of x (i.e., β) is assumed to be constant.
  - However, the hazard ratio for comparing two individuals is now a function of time

Partial Likelihood with Time-Varying Covariates

- Estimation of the model parameters when we have time-varying covariates is still done by maximizing the partial likelihood

$$L_p(\beta) = \prod_{i=1}^{n} \left(\frac{e^{\beta \cdot x_i(t_i)}}{\sum_{j \in R(t_i)} e^{\beta \cdot x_j(t_i)}} \right)^{c_i}$$

- So to calculate the partial likelihood, we need to know the value of the time-varying covariate at each observed event time for all of the people who were still at risk at that time.

Example

- Example: Have a sample of patients who are eligible for heart transplant. Some will later be able to get transplants, others will not.
  - Want to know if people who get transplants will survive longer than those who do not.
  - Naïve (incorrect) approach:

$$\log(h(t \mid \beta,x)) = \log(h_0(t)) + \beta \cdot \text{Transplant}$$

- Transplant status must be treated as time-varying
  - At the start of the study, no one has had a transplant yet
  - If you treat transplant status as a fixed covariate, you are looking into the future to assign subjects to groups

Example

- Let x(t) be the time-varying covariate for transplant status
  - x(t) = 0 if patient has not received a transplant by time *t*
  - x(t) = 1 if patient has received a transplant either on or before time *t*
- For patients who died or left the study before getting a transplant, x(t) always equals zero.
- Value of x(t) only changes for those who got transplants
  - For example, for a patient who got a transplant on Day 5:

$$x(t) = \begin{cases} 0 & \text{if } t \leq 5 \\ 1 & \text{if } t > 5 \end{cases}$$

  - This patient is considered to be a non-transplant patient until Day 5

Discrete Time-Varying Covariates

- Transplant status is an example of a discrete time-varying covariate.
- Discrete time-varying covariates start at a particular value and stay at that value until some intermediate event occurs. The value of the covariate then changes.
  - Discrete time-varying covariates are often indicators for whether the intermediate event has occurred.
    - Start at zero and then change to one when the event occurs
  - Example: Studying time to death following hospital admission for heart failure.
    - x(t) = 0 while subject is in the hospital, x(t) = 1 once subject has been discharged from the hospital

Continuous Time-Varying Covariates

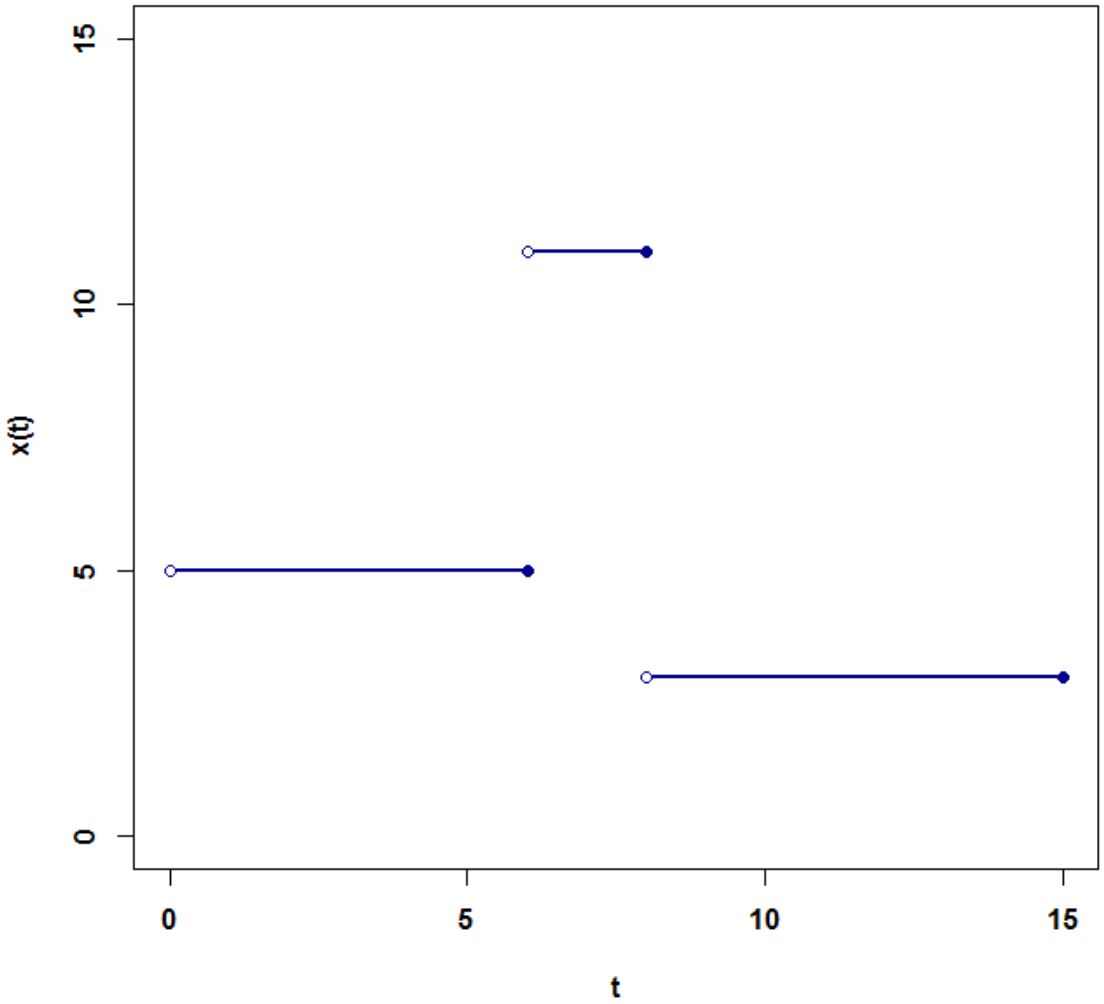
- Continuous time-varying covariates are used when the value of the covariate represents a series of measurements taken over time.
  - Example: blood pressure measurements taken daily
- Using continuous time-varying covariates is usually much more difficult than discrete time-varying covariates.
  - Generally, it is not possible to monitor a subject continuously for values of a covariate
  - Instead, we usually record values of the covariate at certain times

Continuous Time-Varying Covariates

- Continuous time-varying covariates are usually more difficult to use in SAS.
  - SAS has a ‘counting process’ syntax that can be used in PROC PHREG that is useful for continuous time-varying covariates
  - May require lots of additional programming
- To account for time-dependent covariates, multiple records are created for each subject
  - Records are created to match intervals where the value of the covariate is constant.
    - $(t_1, t_2]$
  - Need to have a variable indicating whether the observation was censored at the end of the interval (i.e., at time t_2)

Continuous Time-Varying Covariates

- Example:
  - Measurements of covariate *x* were taken at baseline (t = 0) and at two follow-up visits (one at t = 6, the other at t = 8).
  - Treat the value of *x* as constant between follow-up visits.



Continuous Time-Varying Covariates

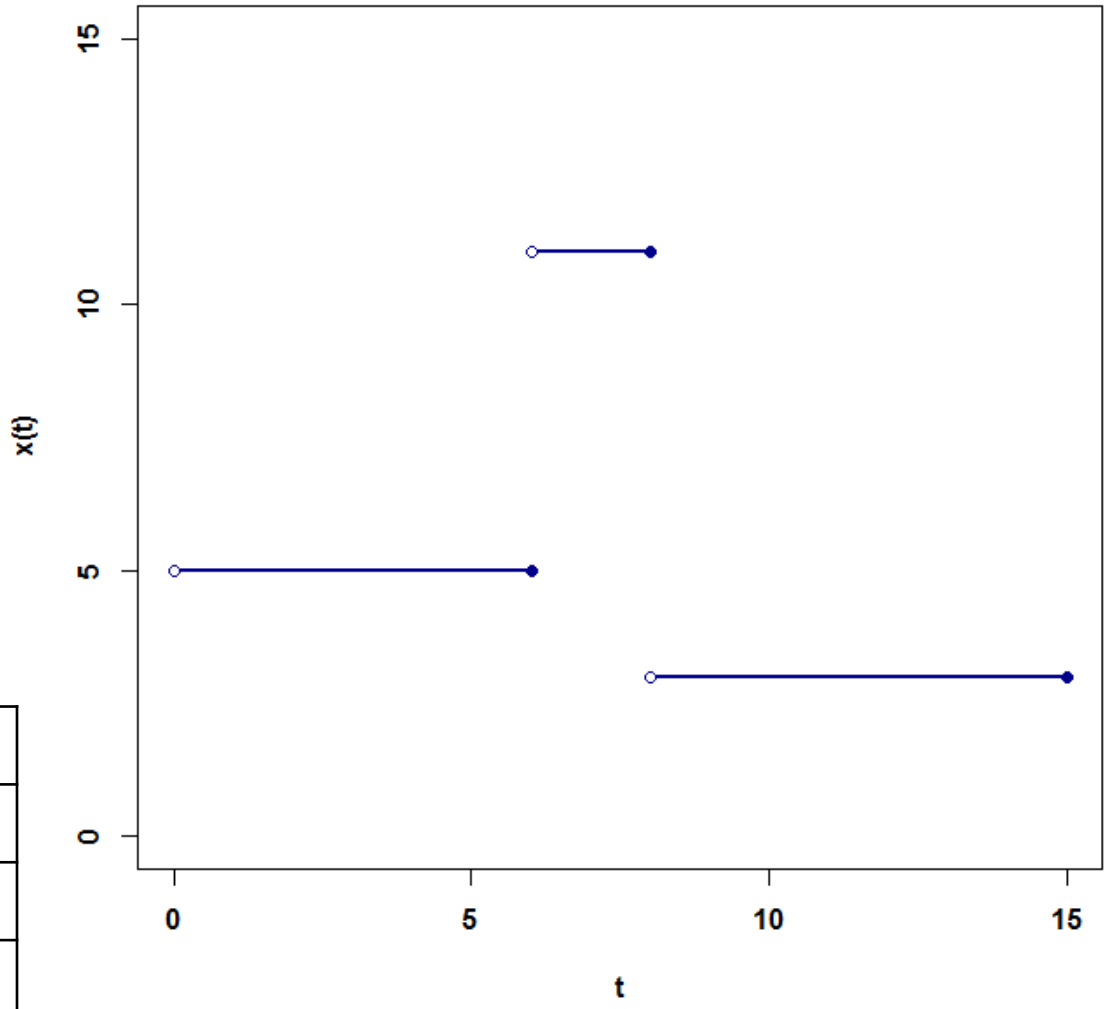
- Original observation:

ID	Months	Died	X0	Visit1
1	15	1	5	6

X1	Visit2	X2
11	8	3

- Counting process:

ID	Start	Stop	Died	X
1	0	6	0	5
1	6	8	0	11
1	8	15	1	3



Counting Process Syntax

- Notes on the counting process syntax in SAS:
  - Intervals have the form $(t_1, t_2]$ or $t_1 < t \leq t_2$.
  - SAS does not allow "zero-length" intervals
    - Example of a zero-length interval: (15, 15].
    - Most common zero-length interval is (0, 0].
    - SAS will ignore any zero-length intervals but will not give you an error message.
  - One way to deal with zero-length intervals is to add a very small number, ε, to t_2.
    - ε needs to be small enough that $t_2 + \varepsilon$ is still less than any times occurring after t_2.

Counting Process Syntax

- Example: Transplant data set
  - One subject died on Day 0 and two subjects had transplants on Day 0.

ID	Days	Died	Transplant	Wait	TransplantDays	Age	Surgery	TissueScore
3	15	1	Yes	0	15	30	No	2
15	0	1	No	.	.	53	Yes	.
45	44	1	Yes	0	44	36	No	1

  - Data rearranged to use counting process syntax:

ID	Start	Stop	TransplantStatus	Censor
3	0	0	0	0
3	0	15	1	1
15	0	0	0	1
45	0	0	0	0
45	0	44	1	1

Counting Process Syntax

- Fit model using a programming step:

Number of Observations Read	103
Number of Observations Used	103

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
103	75	28	27.18

Criterion	Without Covariates	With Covariates
-2 LOG L	596.651	596.475

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TransplantStatus	1	0.12567	0.30108	0.1742	0.6764	1.134

- Fit model using the counting process syntax:

Number of Observations Read	172
Number of Observations Used	169

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
169	74	95	56.21

Criterion	Without Covariates	With Covariates
-2 LOG L	587.382	587.206

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TransplantStatus	1	0.12567	0.30108	0.1742	0.6764	1.134

Counting Process Syntax

- Fit model using the counting process syntax with no zero-length intervals:

Number of Observations Read	172
Number of Observations Used	172

Summary of the Number of Event and Censored Values			
Total	Event	Censored	Percent Censored
172	75	97	56.40

Criterion	Without Covariates	With Covariates
-2 LOG L	596.651	596.475

Analysis of Maximum Likelihood Estimates						
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
TransplantStatus	1	0.12567	0.30108	0.1742	0.6764	1.134

ID	Start	Stop	TransplantStatus	Censor
3	0	1×10^{-5}	0	0
3	1×10^{-5}	15	1	1
15	0	1×10^{-5}	0	1
45	0	1×10^{-5}	0	0
45	1×10^{-5}	44	1	1

Censoring and Truncation

- So far we have assumed that our observations are only subject to right censoring.
- We've also implicitly assumed that all subjects were at risk of the event starting at time t=0 and continued to be at risk until either the event or censoring occurred.
- We might have a data set where these assumptions are not reasonable
  - Can the Cox proportional hazards model be extended to cover these cases?

Left Truncation

- Left truncation: survival time has to exceed some value in order for the subject to be included in the study
- Given that left truncation has taken place, the subjects who are in our study should be treated as not having been at risk between time t = 0 and the truncation point
  - By design, the event can't occur between the start time and the truncation time for subjects in the study.

Left Truncation

- The Cox proportional hazards model can be used with left truncated data.
  - Subjects do not contribute to the partial likelihood unless they are actually at risk (i.e., they have passed the truncation point).
- The counting process syntax in SAS can be used to specify when each subject is at risk.

Interval Censoring

- Interval Censoring: exact survival time is unknown, but we can identify an interval in which the time occurred.
  - Example: Outcome of interest is time to relapse for MS patients. Patients are seen every month and asked whether they had a relapse since the last visit.

- Can fit a Cox proportional hazards model with interval censoring by proposing binomial random variables that indicate whether the event took place in each interval

Interval Censoring

- To fit a Cox proportional hazards model with interval censoring
  - Need to identify a finite set of intervals that are common to all subjects
    - Example: Monthly follow-up visits.
  - Fit a logistic regression with the complementary log-log function as the link function
    - Like the standard Cox PH model, the regression model should not include an intercept term.
  - Model parameters are still interpreted as in a proportional hazards model.

Censoring and Truncation

- What about right truncation and left censoring?
  - Right truncation: survival time has to be less than some value in order for the subject to be included in the study.
  - Left censoring: exact survival time is unknown, but we know it is greater than some value.

- The Cox proportional hazards model cannot be easily extended to these situations.