Chapter 7

Time varying (or time-dependent) covariates

References:

Hosmer & Lemeshow	Chapter 7, Section 3
Collett	Chapter 8
Kleinbaum	Chapter 6
Allison	Chapter 5
Cox & Oakes	Chapter 8
Andersen & Gill	Page 168 (Advanced!)
Kalbfleisch & Prentice	Section 5.3

So far, we've been considering the following Cox PH model:

$$\begin{aligned} \lambda(t, \boldsymbol{Z}) &= \lambda_0(t) \ \exp(\boldsymbol{\beta}^T \boldsymbol{Z}) \\ &= \lambda_0(t) \ \exp\left(\sum \beta_j Z_j\right) \end{aligned}$$

where β_j is the parameter for the the j-th covariate (Z_j).

Important features of this model:

(1) the baseline hazard depends on t, but not on the covariates $Z_1, ..., Z_p$

(2) the hazard ratio $\exp(\boldsymbol{\beta}^T \boldsymbol{Z})$ depends on the covariates $Z_1, ..., Z_p$, but not on time t.

Now we want to relax the second assumption, and allow the hazard ratio to depend on time t.

7.1 A motivating example

Stanford Heart transplant example:

Variables:

- SURVIVAL - time from program enrollment until death or censoring
- DEAD - indicator of death (1) or censoring (0)
- TRANSPL - whether patient ever had transplant
  (1 if yes, 2 if no)
- SURGERY - previous heart surgery prior to program
- AGE - age at time of acceptance into program
- WAIT - time from acceptance into program until transplant surgery (=. for those without transplant)

Initially, a Cox PH model was fit for predicting survival time:

$$\lambda(t, \boldsymbol{Z}) \;=\; \lambda_0(t)\ \exp(\beta_1 * transpl + \beta_2 * surgery + \beta_3 * age)$$

However, this model could give misleading results, since patients who died more quickly had less time available to get transplants.

What does this model imply about the hazard ratio for death for those with vs without transplants?

A model with a time dependent indicator of whether a patient had a transplant at each point in time might be more appropriate:

$$\lambda(t, \boldsymbol{Z}) \;=\; \lambda_0(t)\;\exp(\beta_1 * trnstime + \beta_2 * surgery + \beta_3 * age)$$

where TRNSTIME = 1 if TRANSPL=1 and WAIT$< t$

or conversely TRNSTIME =0 if TRANSPL=2, WAIT=., or the transplant has not occurred yet, eg., WAIT$> t$.

SAS code for these two models

Time-independent covariate for TRANSPL:

```
proc phreg data=stanford;
  model survival*dead(0)=transpl surgery age;
run;
```

Time-dependent covariate for TRANSPL:

```
proc phreg data=stanford;
  model survival*dead(0)=trnstime surgery age;
  if wait>survival or wait=. then trnstime=0;
    else trnstime=1;
run;
```

[Note: there are several other options for coding the time-dependent covariate]

7.2 Extended Cox Model

If we add time-dependent covariates or interactions with time to the Cox proportional hazards model, then it is not a "proportional hazards" model any longer.

We refer to it as an "extended Cox model".

Comparison with a single binary predictor (like heart transplant):

- A standard Cox PH model would compare the survival distributions between those without a transplant (ever) to those with a transplant. A subject's transplant status at the end of the study would determine which category they were put into for the entire study follow-up.

  This wouldn't really be appropriate, because a patient would be put into the "risk set" of subjects with a transplant from the very start of follow-up, even though they might not get one until much later.

- An extended Cox model would compare the risk of an event between transplant and non-transplant at each event time, but would re-evaluate which risk group each person belonged in based on whether they'd had a transplant by that time.

Recidivism Example: (see Paul Allison's book)

Recidivism study:

432 male inmates were followed for one year after release from prison, to evaluate risk of re-arrest as function of financial aid (FIN), age at release (AGE), race (RACE), full-time work experience prior to first arrest (WEXP), marital status (MAR), parole status (PARO=1 if released with parole, 0 otherwise), and number of prior convictions (PRIO). Data were also collected on employment status over time during the year.

Time-independent model:

A time independent model might include the employment status of the individual at the beginning of the study (1 if employed, 0 if unemployed), or perhaps at any point during the year.

Time-dependent model:

However, employment status changes over time, and it may be the more recent employment status that would affect the hazard for re-arrest. For example, we might want to define a time-dependent covariate for each month of the study that indicates whether the individual was employed during the past month.

7.3 Extended Cox Model Framework

Framework:

For individual i, suppose we have their failure time, failure indicator, and a summary of their covariate values over time:

$$(X_i, \delta_i, \{Z_i(t), t \in [0, X_i]\}),$$

$\{Z_i(t), t \in [0, X_i]\}$ represents the **covariate path** for the i-th individual while they are in the study, and the covariates can take different values at different times.

Assumptions:

- conditional on an individual's covariate history, the hazard for failure at time t depends only on the value of the covariates **at that time**:

$$\lambda(t; \{Z_i(u), u \in [0, t]\}) = \lambda(t; Z_i(t))$$

- the Cox model for the hazard holds:

$$\lambda(t; Z_i(t)) = \lambda_0(t)\, e^{\beta Z_i(t)}$$

Survivor function:

As a result of the fact that the second term is also a function of time, we cannot factor it out of the integral when we calculate the survivor function:

$$S(t; Z) = \exp\{-\int_0^t \exp(\beta Z(u))\ \lambda_0(u)du\}$$

and depends on the values of the time dependent variables over the interval from 0 to t.

This is the classic formulation of the time-varying Cox regression survival model.

7.4 Applications and Examples

I. When **important covariates change** during a study

- **Framingham Heart study**

  5209 subjects followed since 1948 to examine relationship between risk factors and cardiovascular disease. A particular example:

  Outcome: time to congestive heart failure
  Predictors: age, systolic blood pressure, # cigarettes per day

- **Liver Cirrhosis study** (Andersen and Gill, p.528)

  Clinical trial comparing treatment to placebo for cirrhosis. The outcome of interest is time to death. Patients were seen at the clinic after 3, 6 and 12 months, then yearly.

  Fixed covariates: treatment, gender, age (at diagnosis)
  Time-varying covariates: alcohol consumption, nutritional status, bleeding, albumin, bilirubin, alkaline phosphatase and prothrombin.

- **Recidivism Study** (Allison)

II. For **cross-over studies**, or to indicate **change in treatment**

- **Stanford heart study** (Cox and Oakes p.129)
  Between 1967 and 1980, 249 patients entered a program at Stanford University where they were registered to receive a heart transplant. Of these, 184 received transplants, 57 died while waiting, and 8 dropped out of the program for other reasons. Does getting a heart transplant improve survival? Here is a sample of the data:

Waiting time	transplant?	survival post transplant	total survival	final status
49	2	.	.	1
5	2	.	.	1
0	1	15	15	1
35	1	3	38	1
17	2	.	.	1
11	1	46	57	1

Naive approach: Compare the total survival of transplanted and non-transplanted.

Problem: Length Bias!

III. For **testing the PH assumption**

For example, we can fit these two models:

(1) **Time independent covariate** Z_1

$$\lambda(t, \boldsymbol{Z}) \;=\; \lambda_0(t)\;\exp(\beta_1 * Z_1)$$

The hazard ratio for Z_1 is $\exp(\beta_1)$.

(2) **Time dependent covariate** Z_1

$$\lambda(t, \boldsymbol{Z}) \;=\; \lambda_0(t)\;\exp(\beta_1 * Z_1 + \beta_2 * Z_1 * t)$$

The hazard ratio for Z_1 is $\exp(\beta_1 + \beta_2 t)$.
(note: we may want to replace t by $(t - t_0)$, so that $\exp(\beta_1)$ represents HR at some convenient time, like the median survival time.)

A test of the parameter β_2 is a test of the PH assumption.

(How do we get a test? ...using the Wald test from the output of second model, or LR test formed by comparing the log-likelihoods of the two models)

7.5 Partial likelihood with time-varying covariates

Starting out just as before...
Suppose there are K distinct failure (or death) times, and let $(\tau_1, ...\tau_K)$ represent the K ordered, distinct death times. For now, assume there are no tied death times.

Risk Set: Let $\mathcal{R}(t) = \{i : x_i \geq t\}$ denote the set of individuals who are "at risk" for failure at time t.

Failure: Let i_j denote the label or identity of the individual who fails at time τ_j, including the value of their time-varying covariate during their time in the study

$$\{Z_{i_j}(t), t \in [0, \tau_j]\}$$

History: Let H_j denote the "history" of the entire data set, up to the j-th death or failure time, including the time of the failure, but not the identity of the one who fails, also including the values of all covariates for everyone up to and including time τ_j.

Partial Likelihood: We have seen previously that the partial likelihood can be written as

$$\begin{aligned} L(\beta) &= \prod_{j=1}^{d} P(i_j|H_j) \\ &= \prod_{j=1}^{d} \frac{\lambda(\tau_j; \boldsymbol{Z}_j(\tau_j))}{\sum_{\ell \in \mathcal{R}(\tau_j)} \lambda(\tau_j; \boldsymbol{Z}_\ell(\tau_j))} \end{aligned}$$

Under the Cox model assumption, this is:

$$L(\boldsymbol{\beta}) = \prod_{j=1}^{d} \frac{\exp(\boldsymbol{\beta} \boldsymbol{Z}_{jj})}{\sum_{\ell \in \mathcal{R}(\tau_j)} \exp(\boldsymbol{\beta} \boldsymbol{Z}_{\ell j})}$$

where $\boldsymbol{Z}_{\ell j}$ is a short-cut way to denote the value of the covariate vector for the ℓ-th person at the j-th death time, ie:

$$\mathbf{Z}_{\ell j} = \mathbf{Z}_\ell(\tau_j)$$

Inference (i.e. estimating the regression coefficients, constructing score tests, etc.) proceeds similarly to standard case. The main difference is that the values of Z will change at each risk set.

Old Example revisited: (see Chapter 4, slide 28)

Group 0: $4^+, 7, 8^+, 9, 10^+$
Group 1: $3, 5, 5^+, 6, 8^+$

Let Z_1 be group, and add another fixed covariate Z_2

ID	failure time	failure indicator	Z_1	Z_2	Contribution $e^{(\beta_1 Z_1 + \beta_2 Z_2)}$
1	3	1	1	1	$e^{\beta_1+\beta_2}$
2	4	0	0	1	e^{β_2}
3	5	1	1	1	$e^{\beta_1+\beta_2}$
4	5	0	1	0	e^{β_1}
5	6	1	1	1	$e^{\beta_1+\beta_2}$
6	7	1	0	0	1
7	8	0	0	1	e^{β_2}
8	8	0	1	0	e^{β_1}
9	9	1	0	1	e^{β_2}
10	10	0	0	0	1

Let's derive the partial likelihood contributions.

Ordered Failure Time (τ_j)	Individuals at risk	failure ID	Partial Likelihood contribution
3			
5			
6			
7			
9			

Ordered Failure Time (τ_j)	Individuals at risk	failure ID	Partial Likelihood contribution
3	1-10	1	$\frac{e^{\beta_1+\beta_2}}{2+2e^{\beta_1}+3e^{\beta_2}+3e^{\beta_1+\beta_2}}$
5	3-10	3	$\frac{e^{\beta_1+\beta_2}}{2+2e^{\beta_1}+2e^{\beta_2}+2e^{\beta_1+\beta_2}}$
6	5-10	5	$\frac{e^{\beta_1+\beta_2}}{2+e^{\beta_1}+2e^{\beta_2}+e^{\beta_1+\beta_2}}$
7	6-10	6	$\frac{1}{2+e^{\beta_1}+2e^{\beta_2}}$
9	9-10	9	$\frac{e^{\beta_2}}{1+e^{\beta_2}}$

Example continued:

Now suppose Z_2 (a completely different covariate) is a time varying covariate:

	failure	failure		$Z_2(t)$							
ID	time	indicator	Z_1	3	4	5	6	7	8	9	$\Rightarrow$ time
1	3	1	1	0							
2	4	0	0	1	1						
3	5	1	1	1	1	1					
4	5	0	1	0	0	0					
5	6	1	1	0	0	0	0				
6	7	1	0	0	0	0	1	1			
7	8	0	0	0	0	0	0	0	0		
8	8	0	1	0	0	0	0	1	1		
9	9	1	0	0	0	0	1	1	1	1	
10	10	0	0	0	1	1	1	1	1	1	

Example continued:

	failure	failure		$Z_2(t)$							Contribution $e^{\beta_1 Z_1 + \beta_2 Z_2(t)}$		
ID	time	indicator	Z_1	3	4	5	6	7	8	9	at $t=3$	at $t=5$	at $t=6$
1	3	1	1	0							e^{β_1}	-	-
2	4	0	0	1	1						e^{β_2}	-	-
3	5	1	1	1	1	1					$e^{\beta_1+\beta_2}$	$e^{\beta_1+\beta_2}$	-
4	5	0	1	0	0	0					e^{β_1}	e^{β_1}	-
5	6	1	1	0	0	0	0				e^{β_1}	e^{β_1}	e^{β_1}
6	7	1	0	0	0	0	1	1			1	1	e^{β_2}
7	8	0	0	0	0	0	0	0	0		1	1	1
8	8	0	1	0	0	0	0	1	1		e^{β_1}	e^{β_1}	e^{β_1}
9	9	1	0	0	0	0	1	1	1	1	1	1	e^{β_2}
10	10	0	0	0	1	1	1	1	1	1	1	e^{β_2}	e^{β_2}

Most of the contributions of the individual subjects remain the same from time $t = 3$ to time $t = 5$, but the contribution for subject 10 changes. Over time, other contributions change as well (for subjects #6 and #9 at time $t = 6$).

Let's see how calculating the partial likelihood changes!

Ordered Failure Time (τ_j)	Individuals at risk	failure ID	Partial Likelihood contribution
3	1-10	1	$\frac{e^{\beta_1}}{4+4e^{\beta_1}+e^{\beta_2}+e^{\beta_1+\beta_2}}$
5	3-10	3	$\frac{e^{\beta_1+\beta_2}}{3+3e^{\beta_1}+e^{\beta_2}+e^{\beta_1+\beta_2}}$
6	5-10	5	$\frac{e^{\beta_1}}{1+2e^{\beta_1}+3e^{\beta_2}}$
7	6-10	6	$\frac{e^{\beta_2}}{1+3e^{\beta_2}+e^{\beta_1+\beta_2}}$
9	9-10	9	$\frac{e^{\beta_2}}{2e^{\beta_2}}$

SAS solution to previous examples

```
Title 'PH regression:  small class example';
data ph;
    input time status group z3 z4 z5 z6 z7 z8 z9;
    cards;
3   1   1  0  .  .  .  .  .  .
4   0   0  1  1  .  .  .  .  .
5   1   1  1  1  1  .  .  .  .
5   0   1  0  0  0  .  .  .  .
6   1   1  0  0  0  0  .  .  .
7   1   0  0  0  0  1  1  .  .
8   0   0  0  0  0  0  0  0  .
8   0   1  0  0  0  0  1  1  .
9   1   0  0  0  0  1  1  1  1
10  0   0  0  1  1  1  1  1  1
run;
```

```
proc phreg ;
  model time*status(0)=group z3 ;
run;

proc phreg ;
  model time*status(0)=group z ;
  z=z3;
  if (time >= 4) then z=z4;
  if (time >= 5) then z=z5;
  if (time >= 6) then z=z6;
  if (time >= 7) then z=z7;
  if (time >= 8) then z=z8;
  if (time >= 9) then z=z9;
run;
```

SAS output from fitting both models

Model with z3:

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	16.953	13.699	3.254 with 2 DF (p=0.1965)
Score	.	.	3.669 with 2 DF (p=0.1597)
Wald	.	.	2.927 with 2 DF (p=0.2315)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
GROUP	1	1.610529	1.21521	1.75644	0.1851	5.005
Z3	1	1.360533	1.42009	0.91788	0.3380	3.898

Model with time-dependent Z:

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	16.953	14.226	2.727 with 2 DF (p=0.2558)
Score	.	.	2.725 with 2 DF (p=0.2560)
Wald	.	.	2.271 with 2 DF (p=0.3212)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
GROUP	1	1.826757	1.22863	2.21066	0.1371	6.214
Z	1	0.705963	1.20630	0.34249	0.5584	2.026

Time-varying covariates in R

The original data on page 4 may be stored as ('wide' format):

Table 1: A Toy Data Example

Subject ID	Group	Z1	Time1	Z2	Time2	Status
1	1	0	3			1
2	1	0	5			0
3	1	1	5			1
4	1	0	6			1
5	1	0	6	1	8	0
6	0	1	4			0
7	0	0	5	1	7	1
8	0	0	8			0
9	0	0	5	1	9	1
10	0	0	3	1	10	0

We first need to create a data set with start and stop values of time ('long' format):

```
id start stop status group   z
 1     0    3      1     1   0
 2     0    5      0     1   0
 3     0    5      1     1   1
 4     0    6      1     1   0
 5     0    6      0     1   0
 5     6    8      0     1   1
 6     0    4      0     0   1
 7     0    5      0     0   0
 7     5    7      1     0   1
 8     0    8      0     0   0
 9     0    5      0     0   0
 9     5    9      1     0   1
10     0    3      0     0   0
10     3   10      0     0   1
```

The R command to fit the Cox model would then be:
'coxph(Surv(time=start, time2=stop, status) ~ group + z, data)'.
This form of Surv() is also used to handle left truncated data, where 'time' is the truncation (entry) time Q, and 'time2' is the event time.

Results:

```
 Alive Dead Deleted
     9    5       0

      coef exp(coef) se(coef)     z     p
[1,] 1.827      6.21     1.23 1.487 0.137
[2,] 0.706      2.03     1.21 0.585 0.558

     exp(coef) exp(-coef) lower .95 upper .95
[1,]      6.21      0.161     0.559      69.0
[2,]      2.03      0.494     0.190      21.5

Likelihood ratio test= 2.73  on 2 df,   p=0.256
Efficient score test = 2.73  on 2 df,   p=0.256
```

Most other softwares handle time-dependent covariates similarly (Stata). SAS has multiple programming options (see Allison book).

7.6 The Stanford Heart Transplant data

```
data heart;
  infile 'heart.dat';
  input wait trans post surv status ;
run;

data heart;
  set heart;
  if trans=2 then surv=wait;
run;

*** naive analysis;
proc phreg;
  model surv*status(2)=tstat;
  tstat=2-trans;

*** analysis with time-dependent covariate;
proc phreg;
  model surv*status(2)=tstat;
  tstat = 0;
  if (trans=1 and surv >= wait) then tstat = 1;
run;
```

The second model took about twice as long to run as the first model, which is usually the case for models with time-dependent covariates.

RESULTS for Stanford Heart Transplant data:

Naive model with fixed transplant indicator:

Criterion	Covariates	Covariates	Model Chi-Square
-2 LOG L	718.896	674.699	44.198 with 1 DF (p=0.0001)
Score	.	.	68.194 with 1 DF (p=0.0001)
Wald	.	.	51.720 with 1 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
TSTAT	1	-1.999356	0.27801	51.72039	0.0001	0.135

Model with time-dependent transplant indicator:

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	1330.220	1312.710	17.510 with 1 DF (p=0.0001)
Score	.	.	17.740 with 1 DF (p=0.0001)
Wald	.	.	17.151 with 1 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
TSTAT	1	-0.965605	0.23316	17.15084	0.0001	0.381

7.7 Recidivism Example

Hazard for arrest within one year of release from prison.

Based on the model below, what are the important predictors of recidivism?

Model without employment status

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	1350.751	1317.496	33.266 with 7 DF (p=0.0001)
Score	.	.	33.529 with 7 DF (p=0.0001)
Wald	.	.	32.113 with 7 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
FIN	1	-0.379422	0.1914	3.931	0.0474	0.684
AGE	1	-0.057438	0.0220	6.817	0.0090	0.944
RACE	1	0.313900	0.3080	1.039	0.3081	1.369
WEXP	1	-0.149796	0.2122	0.498	0.4803	0.861
MAR	1	-0.433704	0.3819	1.290	0.2561	0.648
PARO	1	-0.084871	0.1958	0.188	0.6646	0.919
PRIO	1	0.091497	0.0287	10.200	0.0014	1.096

Recidivism Example: (cont'd)

Now, we use the indicators of employment status for each of the 52 weeks in the study, recorded as EMP1-EMP52.

We can fit the model in 2 different ways:

```
proc phreg data=recid;
  model week*arrest(0)=fin age race wexp mar parro prio employed
        / ties=efron;
  array emp(*) emp1-emp52;
  do i=1 to 52;
    if week=i then employed=emp(i);
  end;
run;

*** a shortcut;
proc phreg data=recid;
  model week*arrest(0)=fin age race wexp mar parro prio employed
        / ties=efron;
  array emp(*) emp1-emp52;
  employed=emp(week);
run;
```

Recidivism Example: Output

Model WITH employment as time-dependent covariate

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
FIN	1	-0.356722	0.1911	3.484	0.0620	0.700
AGE	1	-0.046342	0.0217	4.545	0.0330	0.955
RACE	1	0.338658	0.3096	1.197	0.2740	1.403
WEXP	1	-0.025553	0.2114	0.015	0.9038	0.975
MAR	1	-0.293747	0.3830	0.488	0.4431	0.745
PARO	1	-0.064206	0.1947	0.109	0.7416	0.938
PRIO	1	0.085139	0.0290	8.644	0.0033	1.089
EMPLOYED	1	-1.328321	0.2507	28.070	0.0001	0.265

Is current employment important?

Do the other covariates change much?

Can you think of any problem with using current employment as a predictor?

Another option for assessing impact of employment

Allison suggests using the employment status of the past week rather than the current week, as follows:

```
proc phreg data=recid;
  where week>1;
  model week*arrest(0)=fin age race wexp mar parro prio employed
        / ties=efron;
  array emp(*) emp1-emp52;
  employed=emp(week-1);
run;
```

The coefficient for EMPLOYED changes from -1.33 to -0.79, so the risk ratio is about 0.45 instead of 0.27. It is still highly significant with $\chi^2 = 13.1$.

Does this model improve the causal interpretation?

Other options for time-dependent covariates:

- multiple lags of employment status (week-1, week-2, etc.)
- cumulative employment experience (proportion of weeks worked)

7.8 Some cautionary notes

- Time-varying covariates must be carefully constructed to ensure interpretability
- There is no point adding a time-varying covariate whose value changes the same as study time ..... you will get the same answer as using a fixed covariate measured at study entry. For example, suppose we want to study the effect of age on time to death.

  We could

  1. use age at start of the study as a fixed covariate
  2. age as a time varying covariate

  However, the results will be the same! Why?

7.9 Using time-varying covariates to assess model fit

Suppose we have just fit the following model:

$$\lambda(t; \mathbf{Z}) = \lambda_0(t) \ \exp(\beta_1 Z_1 + \beta_2 Z_2 + \ldots \beta_p Z_p)$$

E.g., the nursing home data with gender, marital status and health.

Suppose we want to test the proportionality assumption on health (Z_p)

Create a new variable:

$$Z_{p+1}(t) = Z_p * \gamma(t)$$

where $\gamma(t)$ is a known function of time, such as

$$\begin{aligned} \gamma(t) &= t \\ &\text{or } \log(t) \\ &\text{or } e^{-\rho t} \\ &\text{or } I_{\{t>t^*\}} \end{aligned}$$

Then testing $H_0 : \beta_{p+1} = 0$ is a test for non-proportionality

Illustration: Colon Cancer data

```
*** model without time*covariate interaction;
proc phreg data=surv;
  model survtime*censs(1) = trtm stagen ;
```

Model without time*stage interaction

Event and Censored Values

Total	Event	Censored	Percent Censored
274	218	56	20.44

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	1959.927	1939.654	20.273 with 2 DF (p=0.0001)
Score	.	.	18.762 with 2 DF (p=0.0001)
Wald	.	.	18.017 with 2 DF (p=0.0001)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
TRTM	1	0.016675	0.13650	0.01492	0.9028	1.017
STAGEN	1	-0.701408	0.16539	17.98448	0.0001	0.496

```
*** model WITH time*covariate interaction;
proc phreg data=surv ;
  model survtime*censs(1) = trtm stagen tstage ;
  tstage=stagen*exp(-survtime/1000);
```

Model WITH time*stage interaction

Testing Global Null Hypothesis: BETA=0

Criterion	Without Covariates	With Covariates	Model Chi-Square
-2 LOG L	1959.927	1902.374	57.553 with 3 DF (p=0.0001)
Score	.	.	35.960 with 3 DF (p=0.0001)
Wald	.	.	19.319 with 3 DF (p=0.0002)

Analysis of Maximum Likelihood Estimates

Variable	DF	Parameter Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Square	Risk Ratio
TRTM	1	0.008309	0.13654	0.00370	0.9515	1.008
STAGEN	1	1.402244	0.45524	9.48774	0.0021	4.064
TSTAGE	1	-8.322371	2.04554	16.55310	0.0001	0.000

7.10 Piecewise Cox Model

The model we just saw is a non-proportional hazards model. In general, a non-proportional hazards model can be written

$$\lambda(t|\mathbf{Z}) = \lambda_0(t)exp\{\boldsymbol{\beta}(t)^T\mathbf{Z}\}$$

so that the regression effect of $\mathbf{Z}$ changes with time.

We can put different assumptions on $\boldsymbol{\beta}(t)$. We can model it as piecewise constant, linear or piecewise linear, or piecewise cubic (spline), etc.

When $\boldsymbol{\beta}(t)$ is piecewise constant, the non-PH model can be written as a Cox model with time-dependent covariates.

Suppose we are interested in comparing two treatments, and:

- HR=θ_1 during the interval $(0, t_1)$
- HR=θ_2 during the interval $[t_1, t_2)$
- HR=θ_3 during the interval $[t_2, \infty)$

Define the following covariates:

- X - treatment indicator
  ($X = 0 \rightarrow$ standard, $X = 1 \rightarrow$ new treatment)
- Z_2 - indicator of t during 2nd interval

$$Z_2(t) = \begin{cases} 1 & \text{if } t \in [t_1, t_2) \\ 0 & \text{otherwise} \end{cases}$$

- Z_3 - indicator of t during 3rd interval

$$Z_3(t) = \begin{cases} 1 & \text{if } t \in [t_2, \infty) \\ 0 & \text{otherwise} \end{cases}$$

The model for the hazard for individual i is:

$$\lambda_i(t) = \lambda_0(t) \exp\{\beta_1 x_i + \beta_2 x_i z_{2i}(t) + \beta_3 x_i z_{3i}(t)\}$$

What are the log hazard ratios for an individual on the new treatment relative to one on the standard treatment?

This type of model would probably be appropriate for the Rofecoxib study

Cardiovascular Events Associated with Rofecoxib (NEJM 2005)

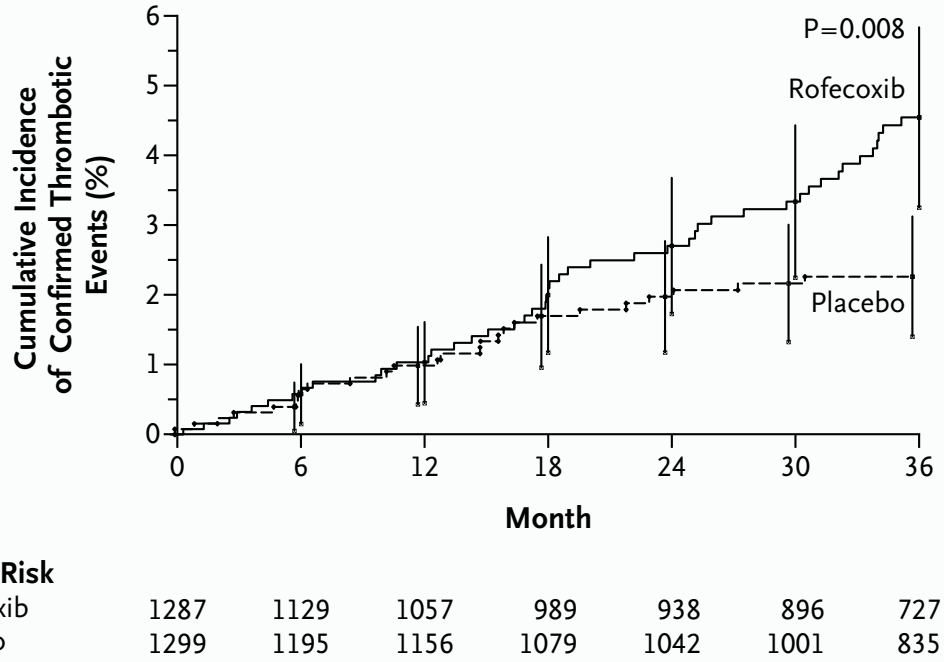

Figure 2. Kaplan–Meier Estimates of the Cumulative Incidence of Confirmed Serious Thrombotic Events.

Vertical lines indicate 95 percent confidence intervals.

How would you set up the time-dependent covariates for this example?

How many do you think you would need?

From the Statistical Analysis section of the article:

> A test of the proportional-hazards assumption was specified in the cardiovascular-analysis plan. This was accomplished by evaluating the interaction between the logarithm of time and the assigned treatment in the Cox proportional-hazards model

From the Results section "Incidence of Thrombotic Events" of the article:

> In a post hoc analysis, the difference between the two groups in the incidence of thrombotic events was evident in the second 18 months of the study, whereas the event rates were similar for the first 18 months (Fig. 2 and Table 3). The changing pattern of the treatment effect over time was confirmed by a failed test for proportionality of hazards (P=0.01).

See Collett, Chapter 10, for further details of piecewise models.

Contents