Seminar in Statistics: Survival Analysis
Chapter 6

Extension of the Cox Proportional Hazards Model for Time-Dependent Variables

Tulasi Agnihotram, Gaby Binder, Fabian Frei
4. April 2011

1 Review

Cox PH-Model: $h(t, \mathbf{X}) = h_0(t) \cdot \exp[\sum_{i=1}^{p} \beta_i X_i]$
Hazard ratio: $\hat{HR}(\mathrm{t}) = \frac{\hat{h}(t, X^*(t))}{\hat{h}(t, X(t))} = \exp[\sum_{i=1}^{p} \hat{\beta}_i [X_i^* - X_i]]$
PH assumption: The hazard ratio is independent of time: $\frac{h(t,X^*)}{h(t,X)} = \theta$
Methods for checking the PH assumption:

- Graphical
- Time-dependent covariates
- Goodness-of-fit test

What can be done if the PH assumption is not met:

- Use a stratified Cox procedure
- Use an extended Cox model

2 Time-dependent Variables

Definition: Any variable whose values differ over time.
Example:

- Race: time-independent
  Race $\times$ t: time-dependent

There are 3 types of variables:

- Defined variables
- Internal variables
- Ancillary variables

3 The Extended Cox Model for Time-dependent Variables

$$h(t, \mathbf{X}(t)) = h_0(t) \cdot \exp[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{i=1}^{p_2} \delta_i X_i(t)]$$

$\mathbf{X}$(t) = $(X_1, \cdots, X_{p_1}, X_1(t), \cdots, X_{p_2}$(t)) denotes all predictors
X_i denotes the i^{th} time-independent variable
X_i(t) denotes the i^{th} time-dependent variable

- The ML procedure is used to estimate the regression coefficients
- The model assumes that the hazard at time t depends on the value of X_i(t) at the SAME time t
- We can modify the model to allow lag-time

4 The Hazard Ratio for the Extended Cox Model

Extended hazard ratio:

$$\hat{HR}(t) = \frac{\hat{h}(t, X^*(t))}{\hat{h}(t, X(t))} = \exp[\sum_{i=1}^{p_1} \hat{\beta}_i [X_i^* - X_i] + \sum_{i=1}^{p_2} \hat{\delta}_i [X_i^*(t) - X_i(t)]]$$

The HR depends on time. That's why the PH assumption is not satisfied.

5 Assessing Time-independent Variables that do not satisfy the PH Assumption

General formula for assessing the PH assumption:

$$h(t, \mathbf{X}(t)) = h_0(t) \cdot \exp[\sum_{i=1}^{p_1} \beta_i X_i + \sum_{i=1}^{p_2} \delta_i X_i g_i(t)]$$

Here the g_i(t) is a function of time. It is important which form we choose for g_i(t) in the model.
We do a test for assessing the PH assumption using the Likelihood ratio test

$$\text{LR} = -2\log(L_{\text{PH model}}) - (-2\log(L_{\text{ext. Cox model}})) \sim \chi_p^2 \text{ under } H_0$$

Nullhypothesis H_0: $\delta_1 = \cdots = \delta_p = 0$
If the test is significant then the extended Cox model is preferred.
We can choose g_i(t) as a heaviside function g_i(t) = $\begin{cases} 1 & \text{if } t \geq t_0 \\ 0 & \text{if } t < t_0 \end{cases}$

In a model, we can use one or more heaviside functions. If we use heaviside functions, the HR yields constant values for different time intervals.

6 An Application of the Extended Cox Model: Treatment of Heroin Addiction

We compare two methadone maintenance clinics. Clinic 2 has always higher retention probabilities than clinic 1. The difference is very significant after one year of treatment. Because the two curves in the -ln(-ln(S)) plot are not parallel, the variable clinic doesn't satisfy the PH assumption. Two extended Cox models were considered:

- Heaviside functions to obtain two distinct hazard ratios. One for less than one year and the other for greater than one year
- A time-dependent variable that allows for the two survival curves to diverge over time

7 The Extended Cox Likelihood

Example:

	TIME	STATUS	SMOKE
Barry	2	1	1
Gary	3	1	0
Harry	5	0	0
Larry	8	1	1

Cox PH model: $h(t) = h_0(t)e^{\beta_1 SMOKE}$
Cox Likelihood:

$$\begin{aligned} L = & [\frac{h_0(2)e^{\beta_1}}{h_0(2)e^{\beta_1} + h_0(2)e^0 + h_0(2)e^0 + h_0(2)e^{\beta_1}}] \\ & \times [\frac{h_0(3)e^0}{h_0(3)e^0 + h_0(3)e^0 + h_0(3)e^{\beta_1}}] \\ & \times [\frac{h_0(8)e^{\beta_1}}{h_0(8)e^{\beta_1}}] \end{aligned}$$

Cox extended model: $h(t) = h_0(t)e^{\beta_1 SMOKE + \beta_2 SMOKE \times TIME}$
Extended Cox Likelihood:

$$\begin{aligned} L = & [\frac{h_0(2)e^{\beta_1 + 2\beta_2}}{h_0(2)e^{\beta_1 + 2\beta_2} + h_0(2)e^0 + h_0(2)e^0 + h_0(2)e^{\beta_1 + 2\beta_2}}] \\ & \times [\frac{h_0(3)e^0}{h_0(3)e^0 + h_0(3)e^0 + h_0(3)e^{\beta_1 + 3\beta_2}}] \\ & \times [\frac{h_0(8)e^{\beta_1 + 8\beta_2}}{h_0(8)e^{\beta_1 + 8\beta_2}}] \end{aligned}$$

References

1. Kleinbaum, D.G. and Klein, M. (2005). *Survival Analysis. A self-learning text.* Springer.

2. ftp://stat.ethz.ch/WBL/Source-WBL-2/R/TK.R.functions.R

3. http://www.stat.ucla.edu/ david/teac/surv/time-dep-coxph.R