Computer Practical Day IV
Exercises

To illustrate interpretation of Cox model results, we consider a cohort of 312 participants in a placebo-controlled clinical trial of D-penicillamine (DPCA) for primary biliary cirrhosis (PBC) (Dickson et al., 1989). PBC destroys bile ducts in the liver, causing bile to accumulate. Tissue damage is progressive and ultimately leads to liver failure. Time from diagnosis to end-stage liver disease ranges from a few months to 20 years. During the approximate ten year follow-up period, 125 study participants died. Predicting survival in PBC patients is important for clinical decision making. The investigators collected data on age as well as baseline laboratory values and clinical signs including serum bilirubin levels, enlargement of the liver (hepatomegaly), accumulation of water in the legs (edema), and visible veins in the chest and shoulders (spiders) all signs of liver damage. Data are in the SPSS file **pbc.sav**.

Fit a Cox model

1. Fit a Cox model for the effects of treatment with DPCA (rx) and bilirubin (bilirubin) on mortality risk in the PBC cohort. Report the hazard ratio for treatments and explain the meaning.

2. Patients in the PBC study underwent a liver biopsy to determine their level of tissue damage. The scores ranged from 1 to 4, with increasing values reflecting greater damage.

   Fit a Cox model with predictor *histology*. Make sure that the variable is categorical (option *categorical*). Note that SPSS takes by default as reference category the last group. Run the analysis by taking as reference category the first group. Click reference category "first"and then click "change".

3. Age at enrollment of participants in the PBC study was recorded in years. Fit a Cox model with age as predictor and estimate the hazard ratio and its confidence interval.

4. Estimate the hazard ratio and the confidence interval for an increase in age of five years.

5. We now examine the association between bilirubin levels and survival. Fit first a Cox model with bilirubin as predictor.

6. Fit a Cox model with bilirubin, hepatomegaly, edema and spiders as predictors. What is the adjusted hazard ratio for bilirubin?

Interaction

We now check for interaction between treatment with (rx) and the presence of liver enlargement or hepatomegaly (hepatom). This analysis examines the hypothesis that treatment is differentially effective according to this baseline covariate. Interaction is handled by including product terms in the model.

1. Define the variable rxhepa as the product of rx and hepatom (you can do this directly in SPSS by selecting the two variables with the command control and the click the option "$> a * b >$"; or you can first create the target variable: rxhepa; numeric expression: rx*hepatom). Fit a Cox model allowing for an interaction between treatment and hepatomegaly (and the main term included).

2. Does there appear to be an interaction between treatment and presence of liver enlargement?

3. Give an interpretation of the interaction hazard ratio $\exp(\beta_{rxhepa})$.

Predicted survival for specific covariate patterns

The estimated survival function $\{\hat{S}_0(t)\}^{\exp(\hat{\beta}_1 Z_1 + \ldots + \hat{\beta}_p Z_p)}$ is also useful for making predictions for specific covariate patterns.

1. Plot the predicted survival curve for a PBC patient with mean values for hepatomegaly and bilirubin as covariates

2. Can you estimate from this curve the mean survival?

Stratified Cox Model and test for the proportional hazard assumption

We want to model the effect of edema among patients with PBC.

1. Estimate Kaplan-Meier with variable **edema** as factor. Plot the survival and cumulative hazard functions. Test the equality of the survival distributions. By looking at the Kaplan-Meier what can you conclude about the proportional hazard assumption. Is that fulfilled?

2. Fit a Cox model by using **edema** and age as explanatory variables. Again plot the survival and cumulative hazard functions. Do you think this is a good model?

3. Now fit a stratified Cox model with **edema** as stratified variable and age as predictor (in SPSS: age as covariate, edema as strata). Plot the survival and cumulative hazard functions. Do you think that this is a good model?

4. As you have seen in question 1 the proportional hazard assumption is violated. Test the proportional hazard assumption for the variable edema by using a graphical method. Use the transformation $\log(-\log(S(t))$ (SPSS option in Cox regression by "plot type", click "Log minus log"). What does the plot obtained here suggest about the proportional hazard assumption?