

TRIAL PROJECT IN STK4080/9080 AUTUMN 2014

From 1962 to 1969, 488 patients with liver cirrhosis at several hospitals in Copenhagen were included in a randomized clinical trial. The purpose of the study was to investigate whether patients treated with the hormone prednisone had a better survival than patients who got an inactive placebo treatment. 251 of the patients received prednisone while 237 received placebo.

In this project you shall study the effect of treatment with prednisone. You shall also investigate the effect on survival of the covariates sex, age, ascites (excess fluid in the abdomen), and prothrombin index (a measurement based on a blood test of some coagulation factors produced by the liver, in percent of the normal value).

You may read the data into R by the command:

```
cirrhosis=  
read.table("http://www.uio.no/studier/emner/matnat/math/STK4080/h14/cirrhosis.txt",header=T)
```

The data are organized with one line for each of the 488 patients, and with the following variables in the eight columns:

- **pas**: patient number (not to be used in the analysis)
- **status**: indicator for death/censoring (1=dead; 0=censored)
- **time**: time in days from start of treatment to death/censoring
- **treat**: treatment (0=prednisone; 1=placebo)
- **sex**: gender (0=female; 1=male)
- **asc**: ascites at start of treatment (0=none; 1=slight; 2=marked)
- **age**: age in years at start of treatment
- **prot**: prothrombin index

As indicated above, the purpose of the project is to determine the importance of the covariates treatment, sex, age, ascites, and prothrombin index on the mortality of the cirrhosis patients.

The analysis consists of three main parts:

a) Simple univariate analyses:

First you perform a simple univariate analysis for one covariate at a time by means of Nelson-Aalen and/or Kaplan-Meier plots and log rank tests, and discuss what you learn from these analyses. (When interpreting the Kaplan-Meier plots, it may be useful to give estimates of, e.g., 1 and 5 years survival and median survival time.) For the numeric covariates age and prothrombin index you should yourself chose a reasonable grouping.

Hint: In order to group age, e.g., into the groups: 49 years or less, 50-59 years, 60-69 years, and 70 years or more, you may create a new categorical covariate **agegroup** by the command:

```
cirrhosis$agegroup=cut(cirrhosis$age,breaks=c(0,49,59,69,100),labels=1:4)
```

b) Univariate regression:

Then you perform a Cox regression analysis for each of the covariates (one at a time), discuss the interpretation of the estimated regression coefficients, and discuss how the results of the Cox analyses compare with the analyses in a). In connection with the Cox regressions, you should in particular consider how the numeric covariates age and prothrombin index should be coded (as given on the data file, or suitably transformed, or grouped).

c) Multivariate regression:

Finally you perform a multivariate Cox regression analysis where the importance of all the covariates is studied simultaneously. In this analysis you should among other things:

- (i) decide which covariates are of importance for the mortality,
- (ii) decide whether the effect of treatment (or other covariates) depends on the values of the other covariates (interaction),
- (iii) check model assumptions.

Make sure that you make a careful interpretation of you “final model” and how the estimates of this model compares with the ones obtained in b). As an aid in the interpretation you may also include estimated survival curves for some selected combinations of the covariates.

The results of your analyses should be written up in a report. In this report you should describe the problem to be investigated and how the statistical analyses have been performed. As indicated above, you should also describe and discuss the results you arrive at. Note that it is preferable to present the results in nice tables, rather than just copying the output from R. Further you should not include R code in the main body of the report. But the R code you use to arrive at your results may be given in an appendix. (Only include code for results that are presented in your report, not all trials and errors.)

If you want to get comments on your work, you should email a pdf-file of your project report to borgan@math.uio.no no later than Friday October 24th.