

PhD Course in Basic Biostatistics Exercises, day 8

Exercise 8.1

The file *meldates.dta* contains the data on survival with malignant melanoma from the lecture day 8. Below we will go through the analysis of these data in Stata.

First we setup the data as survival data in Stata using time since operation (in years) as the time scale. This can be done in two ways in Stata. Either we can use the dates directly:

```
stset enddate, failure(status==1) enter(startdate) ///
origin(startdate) scale(365.25)
```

Or first we can calculate the time since operation

```
generate survtime=(enddate-startdate)/365.25
stset survtime, failure(status==1)
```

After this we can estimate the survival function and the cumulative mortality proportion by the Kaplan-Meier method:

1. Run the commands

sts graph, gwood

And

sts graph, gwood fail

Note: In Stata 11 the option *gwood* has changed to *ci*.

And recognize the plot on pages 10 and 11.

Look at the graph: What is the risk of dying within 7 years after the operation?

Use the command *sts list, at(0 7)* to get the precise value.

Look at the graph: The 25-percentile of the time to death?

Use the command *stci, p(25)* to get the precise value.

We will now compare the mortality among men and women.

2. Plot the survival function for men and women separately like on page 21.

Find the 25-percentile of the time to death for men and women. Compare survival for men and women using the log-rank test.

Check the validity of the assumption of proportional hazards by making the plots on pages 31 and 31:

3. Run the Stata commands
`stcoxkm, by(sex)`
 and
`stphplot, by(sex) nolnt noneg`
(Comment: `nolnt` uses time as x-scale and `noneg` plots the log (-log(survival))).
 Is it reasonable to assume that the hazard ratio is constant over time?

Finally we can estimate the log hazard ratio and the hazard ratio by the commands:
`stcox i.sex, nohr` and `stcox i.sex`

4. Find all estimates, confidence intervals, and tests shown on page 29.

Exercise 8.2

The file `azat.dta` contains the data from example 26.2 in *Kirkwood & Stern* (2003). It is data on 184 patients from a randomized trial of Azathioprine for primary cirrhosis. The primary outcome was death.

1. Plot the Kaplan-Meier survival curves for the patients with and without central cholestasis.
2. Is the survival equal among patients with and without central cholestasis?
3. Use a Cox regression to estimate the hazard ratio between the two types of patients. Compare your estimate with the one found on page 289 in the book.
4. Check the assumption behind the Cox regression you just fitted, and write a short summary of the influence of the prognosis of central cholestasis.

The purpose of the study was to estimate the possible effect of Azathioprine on mortality.

5. Analyze the data for the patients without central cholestasis in order to asses the possible effect of the treatment.
 What are your conclusions?

Exercise 8.3

During an experiment, twenty rats had toxin injected in two doses (low and high). After each injection the diameter of the lesion of the skin was measured (in mm). Ten rats were prior to the experiment treated with progesterone, while the other ten rats were considered controls. You will find the data in the `toxin.dta` Stata-file

1. Analyze the data in order to see how the preliminary treatment with progesterone influences the *difference* in effect of high and low dose toxin?
2. Write a short conclusion containing information on size of the possible effect modification of the difference, and a discussion on the validity of the assumptions behind the statistical

analysis.

3. Explain how you would analyze the data if you wanted to see if the response depends on the preliminary treatment with progesterone.