

Clinical Trials & Survival Analysis

Micro Session 2, Wednesday 25 April, 10-12, Computer room 72

This session is an opportunity to try some of the practical exercises on Survival Data contained on the task sheets and more substantial exercises. Some of the data sets are on the course web page at <http://nickfieller.staff.shef.ac.uk/tampere/index.html> and you can download them (recommended) or open them directly in R or SPSS.

Survival Analysis in SPSS

The facilities for survival analysis in SPSS are found through the menus in Analyze>Survival. We shall be concerned with the second and third items in the list: Kaplan-Meier and Cox Regression. The first of these is useful for drawing Kaplan-Meier estimates of the survival curve and performing log rank tests. The second is for non-parametric proportional hazards regressions models.

1. Download the file `tumour.sav` found under Tumour remission times (SPSS) from the course web page and read it into SPSS. These are the data given in the example 2.4.3 on P31 of the course notes.
2. Open the dialogue box Analyze>Survival>Kaplan-Meier, enter the variable time into the box Time, the variable censor into the box Status and then click on Define Event and enter 1 in the sub-dialogue box against Single Value to indicate that censor=1 means the event has occurred.
3. Next click on Options and tick the box Survival under Plots followed by Continue to close the sub-dialogue and then OK to run the analysis. This should produce the KaplanMeier calculations and a plot of the Kaplan-Meier estimate.
4. Next download the file `braintu.sav` found under Brain tumour survival times (SPSS) and load into SPSS. Follow the instructions in 2 and 3 above and additionally enter group into the box Factor. Also under the Compare Factor sub-dialogue box. Click continue and OK to run the analysis to produce a Kaplan-Meier plot with two estimates displayed and also the value of the log rank statistic.
5. Other survival data sets are `kidneytumour.sav`, `lcancer.sav`, `leukrem.sav`, `livcancer.sav` and `lymphoma.sav` but do not spend too much time on this if you are happy with the SPSS facilities and instead move on to the section on R.

Survival Analysis in R

Start **R** and change the working directory to the folder where you will store the Rdata files. If the installation of **R** has accepted the option to associate Rdata file with running **R** then they can be opened by double clicking (in Window Explorer). Otherwise (and on a Mac), navigate to the directory containing the file and use the `load()` function: `load("tumour.Rdata")`. The facilities for survival analysis in **R** are in a special library called `survival` so the first step is always to open this library.

1. Type `library(survival)`
2. Type `library(help=survival)`. This will give information about the package and in particular a list of all the commands available in the package. We will be concerned mostly with the following:

<code>Surv</code>	Package a survival variable
<code>coxph</code>	Proportional Hazards Regression
<code>plot.survfit</code>	Plot method for <code>survfit</code> .
<code>strata</code>	Identify strata variables.
<code>survdiff</code>	Test Survival Curve Differences
<code>survfit</code>	Compute a survival Curve for Censored Data
<code>survreg</code>	Regression for a parametric survival model

3. Download `tumour.Rdata` from the course web page under Tumour remission times and load into **R** with `load("tumour.Rdata")`
4. Type `tumour` and it should list the data on page 31.
5. Type `attach(tumour)` to make sure that **R** will recognise the names of individual variables.
6. Next, create a 'survival object' with the command `Surv()` (note the first capital letter) and call the object `tumour.sv` (or a name of your choice):

```
tumour.sv<- Surv(time,censor,type="right")
```

7. The calculate the Kaplan-Meier estimate with `survfit()` (note the small letters), then summarize it and plot it with:

```
tumourSurv<-survfit(tumour.sv ~ 1)
summary(tumourSurv)
plot(tumourSurv)
```

This should give the Kaplan-Meier calculations and a plot of the estimate survival curve. (as in section 2.4.4.1).

8. Next, reproduce example 3.2.3.1 (p63 in course notes): download `braintu.Rdata` and load into **R**, attach the data set, create a Survival object called `braintu.sv` and then produce Kaplan-Meier estimates with

```
brainSurv<-survfit(braintu.sv~group)
```

and then summarize and plot it.

```
summary(brainSurv)
```

```
plot(brainSurv)
```

9. Type `survdiff(brain.sv~group)` to obtain the log rank test [and compare with the SPSS version if you have tried that].
10. Other things you can try are reproducing other examples from the course notes and the some of the exercises (in **R** or SPSS or both). The simplest questions are 2, 3, 4 and 5.