

Stochastic Models (Online)

Homework 3

Set: November 3rd, 2016.

Due: November 20th, 2016

1. The data on the survival times of 65 multiple myeloma patients were recorded. The data are available in the **emplik** R package and are called **myeloma**. A Cox regression model has been proposed to compare the survival of different patients, and particular interest is in the effect of the platelets variable and its effect on survival.

The data can be prepared for analysis using the following code:

```
library(emplik)
library(survival)

data(myeloma)
colnames(myeloma) <- c("time","vstatus","logBUN","HGB","platelet","age",
,"logWBC","FRAC","logPBM","protein","SCALC")
myeloma <- as.data.frame(myeloma)
```

Using the `coxph()` function, develop an appropriate model for predicting the survival of a myeloma patient.

An example of fitting such a model is given below:

```
fit <- coxph(Surv(time,vstatus)~platelet+age+platelets:age, data=myeloma)
```

where the model allows for platelet, age and an interaction between platelets and age in the model.

Carefully explain why you selected a particular model. Summarize the results of the chosen model and the effect of each covariate in this model.

2. The following are the survival times in weeks of a group of 14 hepatitis patients on a steroid treatment: 4 13 3 6 4+ 3 1 1 10+ 6 6 20+ 3 15

It is assumed that the survival function can be estimated from the data using a Kaplan-Meier estimate.

- (a) Construct a Kaplan-Meier estimate of the survival function and plot the estimate of the survival function
- (b) An alternative estimate of the survival function is called the Nelson-Aalen estimate.

The Kaplan-Meier estimate takes the form:

$$\hat{S}_{KM}(t_j) = \prod_{k=1}^j (1 - \lambda_k) = \prod_{k=1}^j \left(1 - \frac{d_k}{n_k}\right).$$

The Nelson-Aalen estimate takes the form:

$$\hat{S}_{NA}(t_j) = \exp \left(- \sum_{k=1}^j \lambda_k \right) = \prod_{k=1}^j \exp(-\lambda_k) = \prod_{k=1}^j \exp \left(- \frac{d_k}{n_k} \right).$$

Contrast the estimated survival functions found using the Kaplan-Meier and Nelson-Aalen approach.