ST745, Spring 2016

Homework 5 Due: Tuesday, 03/29/2016

1. The following table gives a small data set of survival times and a covariate z:

patient id	survival time (in years)	z
1	6^+	2
2	5	3
3	7^+	2
4	9^+	1
5	8	4
6	3	2

where + means a right censored observation. Assuming a proportional hazards model

$$\lambda(t|z) = \lambda_0(t)\mathrm{e}^{z\beta},$$

do the following:

   (a) Write down the partial likelihood of β.

   (b) Plot the log partial likelihood of β in [-2, 2], and convince yourself that this function is concave (Hint: using the griding method to make the plot).

   (c) Find $\hat{\beta}$ that maximize this log partial likelihood function, calculate the second derivative of the log partial likelihood function at $\hat{\beta}$.

   (d) Use `Phreg` in SAS to fit the above proportional hazards model to the data. How do your results compare to those from the SAS output?

2. We showed in class that the score test for comparing two treatments and two sample log rank test are equivalent when there are no ties in the censored survival times. This equivalence is also true for the situation where there are more than two treatments. In this problem, you are asked to show part of this when there are three treatments. Namely, suppose we have the following proportional hazards model

$$\lambda(t|\cdot) = \lambda_0(t)\mathrm{e}^{Z_1\beta_1+Z_2\beta_2},$$

where Z_1 and Z_2 are two dummy variables created for 3 treatments:

$$Z_1 = \begin{cases} 1 & \text{if treatment} = 1 \\ 0 & \text{otherwise} \end{cases}$$

$$Z_2 = \begin{cases} 1 & \text{if treatment} = 2 \\ 0 & \text{otherwise} \end{cases}$$

Given data $(x_i, \delta_i, z_{1i}, z_{2i})$ for $i = 1, 2, ..., n$ where there are **NO** ties in the censored survival times, show that the score vector $(\frac{\partial \ell}{\partial \beta_1}, \frac{\partial \ell}{\partial \beta_2})^T$ for testing

$$H_0 : \beta_1 = \beta_2 = 0$$

is identical to the vector of the 3-sample log rank test (with $w(x) = 1$) given on page 32 of lecture note 4.

3. Do problem 8.4 on page 288 of the textbook. The data set can be found in the textbook (page 240, Problem 7.7).