Math 150 - Methods in Biostatistics - Homework 10

Solutions

Due: Wednesday, April 10, 2019, in class

```
knitr::opts_chunk$set(message=FALSE, warning=FALSE, fig.height=3.5, fig.width=8,
                      fig.align = "center")
library(tidyverse)
library(broom)
library(tidylog)
library(survival)
library(survminer)
```

1. Chp 9, A45 (pg 311)

Note1 the R manual corresponding to your text has code which allows you to plot the hazard function. You may use that code or you can use the code (see R code in the class notes, set `fun="cumhaz"`) to plot the cumulative hazard function using `ggsurvplot`.

Note2 Remember that when you are loading in the data, the missing values are coded as `*`.

Use the software instructions provided to plot the estimated hazard rates for the college graduation data.

2. Chp 9, A46 (pg 311)

Although the estimated hazard curve may not exhibit a distinguishable pattern, discuss some important features of the curve.

3. Chp 9, A47 (pg 311)

Indicate periods of time during their college career when students are at their lowest and highest risk of graduating college. Does your answer match your common understanding of when students typically graduate from college?

4. Chp 9, E9

Sketch hazard functions that would correspond to the following time-to-event random variables (You may want to do a little background research.)

(a) Lifetime of an individual measured from birth (don't assume anything about the health or demographics of this person).

(b) Time until death after surgery to remove a cancerous tumor.

Be sure to label the time axis, and mark time points appropriately. Briefly explain your reasons for any changes in the shape of the hazard function over time.

5. Chp 9, E10

The graphs displayed in Figure 9.19 are population cumulative hazard functions for three distributions of the time-to-event random variable, T. For each one, sketch a possible corresponding hazard function h(t). Be sure to label the same time points on your sketches as are provided on the graphs of H(t).