PSTAT 175: Lab A

Please complete this work using `R`. Type up a lab report as described in the guidelines and submit it as a .pdf file on Gauchospace.

1. For this question, we will use the data in the file vets.txt which measured the time-to-event for subjects at a VA medical clinic.

   (a) Read in the first column of data as a vector and call it `vet.time`

   (b) The second column is 1 when an event occurred for that patient and 0 when that patient was censored. Read in the second column as a vector `vet.cns`

   (c) Load the library `survival`, and then create a survival vector for this data using the `Surv` function. Call this vector `vet.surv`. Print out the result.

   (d) Calculate the mean of `vet.time`. Why is this a biased estimate of the average time until an event?

   (e) Calculate `sum(vet.cns)` and `sum(vet.time*vet.cns)`. Give an interpretation for these statistics. In particular, what happens when you add up or multiply by a vector of 1s and 0s.

   (f) We can run the function `mean(vet.surv)` Why is this different from the answer above? What does this function calculate?

2. The `retire.txt` data set contains data about the number of months (`retire$time`) that subjects lived at a retirement home. The vector `retire$death` is coded 1 if the subject died and 0 if they left the home for other reasons.

   Use the `R` code

   ```
   retire <- read.table("retire.txt", header=TRUE,skip=2)
   ```

   to read in the data.

   (a) Construct a survival object `ret.surv` from this data using the `Surv` function.

   (b) The function `survfit` calulates a Kaplan–Meier estimate of the survival function. The syntax in this case is

   ```
   survfit(ret.surv ~ 1)
   ```

   (c) Use the `plot` function on the results from `survfit` to generate a picture of the estimate of the survival function. Include a clear, properly labeled plot in your report.

   (d) Use the `summary` function on the results from `survfit` to generate a summary of the survival function. What does this summary tell us about the probability of surviving past 50 months in this sample?