BIOST 509: Homework 2

Instructor: Amy Willis, Biostatistics, UW

Due date: 1:00pm on October 18, 2019 via Canvas

Instructions

All questions for this homework relate to the dataset "psa.txt", which we analysed in In-Class Exercise 1. The dataset is available from Canvas in `Pages/Week 1 materials` or `Files/datasets`.

As before, "psa.txt" contains data on 50 men having hormonally treated prostate cancer. The first line of the file contains the variable names separated by white space, and each successive line contains data pertinent to one of the 50 study subjects. Missing data are denoted with "NA". The variables are as follows:

- *ptid* = patient identifier
- *nadirpsa* = lowest PSA value attained post therapy (ng/ml)
- *pretxpsa* = PSA value prior to therapy (ng/ml)
- *ps* = performance status (0= worst, 100= best)
- *bss* = bone scan score (1= least disease, 3= most)
- *grade* = tumor grade (1= least aggressive, 3= most)
- *age* = patient's age (years)
- *obstime* = time observed in remission (months)
- *inrem* = indicator whether patient still in remission at last follow-up ("yes" or "no")

Read in the dataset and answer the following questions *using the `tidyverse` functions introduced in lecture* (`filter`, `group_by`, `summarize`, etc.). Submit your answers to the below questions in a R Script (.R), Word (.doc or .docx) or pdf file to Canvas. Don't forget to include both your code and answers in your response.

Questions

1. How many individuals have missing `pretxpsa` data?
2. Which individuals have missing `pretxpsa` data?
3. What is the mean `pretxpsa` value for individuals with non-missing data on this variable?
4. What was the minimum observed time in remission among observations in remission at last follow-up?
5. What is the mean nadir PSA level for individuals whose observed time in remission was less than 24 months? What is the mean nadir PSA level for individuals whose observed time in remission was *not* less than 24 months?
6. What is the mean tumour grade for individuals whose observed time in remission was less than 24 months? What is the mean tumour grade for individuals whose observed time in remission was *not* less than 24 months?
7. What is the mean age of individuals whose observed time in remission was less than 24 months? What is the mean age of individuals whose observed time in remission was *not* less than 24 months?
8. What is the mean bone scan score of individuals whose observed time in remission was less than 24 months? What is the mean bone scan score of in individuals whose observed time in remission was *not* less than 24 months?