BIOST 509: In-Class Exercise 6

Instructor: Amy Willis, Biostatistics, UW

Due date: 6:30pm on November 8, 2019 via Canvas

Instructions

Use the PSA dataset used in Homework 2 to answer the following questions. The dataset is available from Canvas in `Pages/Module 6 materials` or `Files/datasets`. **Answer the questions in full sentences.** This is an opportunity to practice your scientific and statistical communication!

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")

You may find reading through your Homework 2 responses helpful.

Optional but encouraged: Use R Markdown to create the pdf file with your results. See In-Class Exercise 5 for a brief introduction. The R Markdown file (".Rmd") used to produce the Homework 2 Solutions is available on Canvas under `Pages/Module 5 materials` to guide you.

Questions

1. Create an indicator variable for relapse within the first 24 months of hormone treatment. Let the indicator variable equal 1 if a relapse occurs within 24 months of treatment.
2. Perform a logistic regression analysis with relapse within 24 months as the response. Use tumor grade as the predictor, treating tumor grade as a continuous variable in your model.
3. Perform a logistic regression analysis with relapse within 24 months as the response. Use tumor grade as the predictor, treating tumor grade as a categorical variable in your model.
4. Is there any evidence of an association between relapse within 24 months and tumor grade? Provide evidence to support your conclusion.
5. (*Optional but encouraged, especially for students who have previous exposure to logistic regression*) Interpret the coefficient(s) on tumor grade in the models that you fit in Questions 2 and 3.