Writing Project 2

Paper, due Thursday November 15 by 5 pm; submit blinded version by email to Prof. Levine

Brief Background:

Prostate cancer, along with skin cancer, is the most common cancer among men. It can often be treated if detected early enough. The Ohio State University Comprehensive Cancer Center lead a study to determine if baseline exam measurements can predict whether a tumor has penetrated the prostatic capsule. In a subset of the data appearing on our class Blackboard site, 153 of 380 subjects had a cancer that penetrated the capsule. The data set consists of the following variables.

Variable	Description	Coding
ID	Identification code	integer
capsule	Tumor penetration of prostatic capsule	1 = penetration
age	Age of subject	years
race	Race of the subject	1 = white, 2 = black
dpros	Results of digital	1 = no nodule
	rectal exam	2 = unilobar nodule (left)
		3 = unilobar nodule (right)
		4 = bilobar nodule
dcaps	Detection of capsular involvement	1 = no, 2 = yes
psa	Prostatic Specific Antigen value	*mg/ml*
gleason	Total Gleason score	scale of 0-10

PSA is a measure of a protein produced by prostate gland cells. Elevated levels may suggest prostate cancer and is thus used as a screening test. The Gleason score is a scale measuring the abnormality of cells. Larger values suggest higher risk of cancer.

Data Analysis Report:

Your report is restricted to 7 pages including all tables and figures, but excluding appendices and references. The report must be compiled in LaTeX using the `documentclass{article}` with 11 point font, one inch margins, and single spacing. Please see the course Blackboard site for the LaTeX preface. Please see the document "Generic Outline for Data Analysis Report" on the course Blackboard site for organization of the report. The audience of your paper are statisticians and statistically-savvy scientists. Statistical results should be appropriately interpreted and put within scientific context throughout. Do not present *R* output dumps; instead present results in well formatted tables and figures that allow you to effectively communicate your findings and aid the reader in interpreting your text. All figures and tables presented in the main document must be discussed in the main text.

Your report must include an appendix of code and supplemental material. The code should allow another student in the class to repeat the analysis you performed. The code must then include appropriate documentation and output for the user to understand your programming and perform quality control checks. The supplemental material may contain well labeled, concise, and neatly presented EDA and analysis results (i.e., tables and figures) that did not fit in the main document. Though the supplemental material should not be a critical part of the discussion and argument in the main text, it may support analysis decisions, findings, and conclusions presented in the main text.

The course Blackboard site presents the grading rubric for the writing project.