Exercise 4, Prostate Cancer

Story

The data come from a study conducted by Stamey et al. (1989) where associations between the level of a prostate-specific antigen and a number of potentially explanatory covariates were studied. The aim of the study was to see whether any of the covariates could predict the level of prostate-specific antigen.

Data

Outcome

lpsa: level of prostate-specific antigen

Covariates

lcavol: log cancer volume

lweight: log prostate weight

age: age

lbph: log of the amount of benign prostatic hyperplasia

svi: seminal vesicle invasion (0 = no, 1 = yes)

lcp: log of capsular penetration

gleason: Gleason score (6,7,8,9)

pgg45: percent of Gleason scores

Exercise

- Read in the data prostate.txt and make descriptive plots of plot the lpsa, lcavol and svi.
- Fit an ANCOVA with lpsa as outcome and lcavol as continuous and svi as categorical covariates, with different slopes for svi=1 and svi=0
- Do a model check.
- Reduce the initial model until there only are significant covariates left. Estimate the parameters with 95% confidence intervals.
- Plot the estimated regression lines from the final model.
- If time look at the rest of the covariates and fit a general linear model, which covariates are statistically significant for the level of prostate-specific antigen?