Stat 452
Spring 2021

Homework 2, due Wed, Feb 10 by midnight

Please do this assignment in R Markdown and submit the Rmd file in Moodle.

1. Problem 40, Chap. 22, p. 703 (Fish Diet). In addition, do part (e):

(e) Construct and plot Bonferroni simultaneous confidence intervals for the relative risk of prostate cancer for each of the top 3 diet groups vs. the control group. See Chapter 22 notes. In the command `pairwiseCI` in package pairwiseCI, you can ask for pairwise CI's for each group vs. control group (rather than all pairwise CI's) by adding the option `control=` with the name of the control group in quotes (see `?pairwiseCI`). What do these CI's tell you?

2. The data set Body_fat_complete.csv (which we looked at last semester) is based on a study of 250 men to see if percent body fat, which is difficult to measure accurately, could be estimated from easy-to-measure variables, like waist size and height. Percent body fat (Pct.BF) was determined by an involved, but accurate, method and 13 other variables were measured. Assuming these men are representative of men of similar ages, look at predicting Pct.BF from Waist size (in inches) only.

(a) Graphically assess whether a linear regression model for predicting Pct.BF from waist size seems appropriate and whether the assumptions for inference are satisfied. That means a scatterplot, residual plot, etc.

(b) Report the estimated slope and a 95% confidence interval (using inline code) and interpret in the context of the problem.

(c) Report and interpret an estimate and 95% confidence interval for the mean body fat percentage for men with 40-inch waists.

(d) Report and interpret an estimate and 95% prediction interval for the body fat percentage for a man with a 40-inch waist.

3. (Grad problem) The data set Respiratory.csv contains the respiratory rates of 618 children between the ages of 0.1 and 3 years. There are two goals: 1) estimate (including a confidence interval) how respiratory rate changes with age, 2) provide a useful plot or chart that a physician could use to assess a normal range of respiratory rates for children of any age between 0 and 3. Consider transformations if necessary. Justify the model you choose.