

Homework 4 (Due 2/17/2014, 10:10am)

1. The Forced Expiratory Volume (FEV, measured in liters) is a primary indicator of lung function and corresponds to the volume of air that can forcibly be blown out in the first second after full inspiration. The Large.FEV data file contains the FEV values of a large sample of children, along with some categorical descriptor of each individual.

- a) Make a histogram of FEV and describe the distribution.
- b) Make separate histograms of FEV for girls (sex = 0) and for boys (sex = 1), using the same horizontal scale for both. Describe the differences between the two groups.
- c) Now create a new variable based on age: preschool for ages 3–5, elementary for ages 6–10, middle for ages 11–13, and highschool for ages 14 and above. Then make separate histograms using the same horizontal scale for each of the four age groups. What do you notice?
- d) Lastly, create eight separate histograms with the same horizontal scale for the eight combinations of sex and age group. Does this last analysis change your previous conclusions? If so, explain how.

2. Continue your work from question 1 with the large data set Large.FEV. We now study adolescents (middle school and high school: ages 11 and over) separately from the younger children (preschool and elementary school: ages 10 and below).

- a) Among the younger children, boys and girls are similar enough to be studied together. Plot the relationship between FEV and height for these younger children and obtain the least-squares regression equation. What percent of the variation in FEV is explained by this model? Predict the FEV of child who is 5 feet tall (60 inches).
- b) Plot the relationship between FEV and height for adolescent girls and obtain the least-squares regression equation. What percent of the variation in FEV is explained by this model? Predict the FEV of an adolescent girl who is 68 inches tall.
- c) Plot the relationship between FEV and height for adolescent boys and obtain the least-squares regression equation. What percent of the variation in FEV is explained by this model? Predict the FEV of an adolescent boy who is 68 inches tall.
- d) Why we should analyze the boys and girls separately?

3. (Graduate Only) Follow the lecture notes and derive the variance of the mean response and the variance of the predicted value.

4. Vitamin D is needed for the body to use calcium. An experiment is designed the effects of calcium and vitamin D supplements on the bones of first-year college students. The outcome measure is the total body mineral content, a measure of bone health. Three doses of calcium will be used: 0, 200, and 400mg/day. The does of vitamin D will be 0,50,and 100 international units per day. The calcium and vitamin D will be given in a single tablet. All tablets including those with no calcium and no vitamin D will look identical. Subjects for the study will be 90 men and 90 women.

- a) What are the factors and the treatments for this experiment?
- b) Draw a picture explaining how you would randomize the 180 college students to the treatments.
- c) Use R to carry out randomization.

5. (bonus 8pts) "First Monday" is a peer-reviewed journal on the internet. They recently published two articles concerning Facebook and academic performance. Visit their website, firstmonday.org, look at the first three articles in Volumn14, number 5-4, May 2009. Identify the key controversial issues that involve the use of statistics addressed in these articles and write a report summarizing the facts as you see them.