

Math 222 - Project 7

Due Friday, March 30

1. The Great Britain Office of Population Census and Surveys once collected data on a random sample of 170 married couples in Britain, recording the age (in years) and heights (in inches) of the husbands and wives. The data is in the file `marriedHeights.csv`
 - (a) Plot the relationship between the heights of husbands and wives in the data set. How strong is the correlation?
 - (b) Find the least squares regression line. Add a plot of the least squares regression line to the plot in part (a), and also give the formula for the least squares regression line.
 - (c) What is the slope of the least squares regression line? Explain what it means in words for this example.
 - (d) How tall are the wives of men who are 6 feet tall, on average? Express your answer using a confidence interval.
 - (e) If Bob is 6 ft. tall, you cannot be 95% sure that his wife is in the confidence interval above. Make a prediction interval for the height of his wife, and explain how it differs from the confidence interval in part (e).
2. Many people believe that gender, weight, drinking habits, and many other factors are much more important in predicting blood alcohol content (BAC) than simply considering the number of drinks a person consumed. Here we examine data from sixteen student volunteers at Ohio State University who each drank a randomly assigned number of cans of beer. These students were evenly divided between men and women, and they differed in weight and drinking habits. Thirty minutes later, a police officer measured their blood alcohol content (BAC) in grams of alcohol per deciliter of blood. The data is in the file `bac.csv`
 - (a) Make a scatterplot for this data. Use it to describe the relationship between the number of cans of beer and BAC.
 - (b) Write the equation of the regression line. Interpret the slope and intercept in context.
 - (c) How much of the variability in BAC can be explained simply by the number of beers a student has drunk?
 - (d) Make a 95% confidence interval for the slope of the regression line. Explain clearly what this confidence interval tells us about the effect of each extra beer a person drinks.
 - (e) Suppose we visit a bar, ask people how many drinks they have had, and also take their BAC. Do you think the relationship between number of drinks and BAC would be as strong as the relationship found in the Ohio State study? Why or why not?