This homework is due Thursday, November 6th at the start of class. Late homework will not be accepted except for medical emergencies (with proof). Note that if a question says to explain your answer, you will get no credit without some explanation.

Reading: Chapter 7.

1. How well does the number of beers a student drinks predict his or her blood alcohol content? Sixteen student volunteers at Ohio State University drank a randomly assigned number of cans of beer. Thirty minutes later, a police officer measured their blood alcohol content (BAC). The data are in bac.JMP. The students were equally divided between men and women and differed in weight and usual drinking habits. Because of this variation, many students don’t believe that number of drinks predicts blood alcohol content well.

(a) Make a scatterplot of the data. Does the graph suggest that a simple linear regression model might be appropriate? Find the equation of the least squares regression line for predicting blood alcohol from number of beers and add this line to your plot. Briefly summarize what your data analysis shows.

(b) Give a 95% confidence interval for the slope of the regression line assuming that the simple linear regression model holds.

(c) Is there significant evidence that drinking more beers increases blood alcohol on the average in the population of all students? State hypotheses, give a p-value and state your conclusion.

(d) Calculate a 95% confidence interval for the mean blood alcohol content of a student who drinks six beers. Calculate a 95% prediction interval for the mean blood alcohol content of a student who drinks six beers.

(e) Do you trust the above confidence interval for the mean response and prediction interval for a student who drinks six beers, even though there are no students in the sample who drank six beers? Explain.

(f) Predict the blood alcohol content of a student who drinks 20 beers based on the least squares line. Do you trust this prediction? Explain.

(g) The police are trying to establish guidelines for how much beer one can drink and still drive safely. Which do you think are more relevant - (i) confidence intervals for the mean blood alcohol of a student who drinks \( x \) number of beers or (ii) prediction intervals for the mean blood alcohol content of a student who drinks \( x \) number of beers? Explain.

2. One of the applications of regression analysis is to the situation where two variables \( x \) and \( y \), are related, but one of them – say, \( y \) - is difficult to measure. We can finesse such a problem by simply measuring the \( x \)-variable and estimating \( y \) via the regression function. It may be extremely difficult, for example, to measure the volume of an irregularly shaped object but very simple to weigh it. The data in childvolume.jmp shows the weight in kilograms and the volume in cubic decimeters of 18 children between the ages of 5 and 8.

(a) Find the coefficients of the least squares line.

(b) Find a 95% confidence interval for the mean volume of children with weight 14 pounds, \( \mu \{Y|14.0\} \).

(c) Find a 95% prediction interval for the volume of a child weighing 14.0 pounds.
(d) A doctor needs to estimate the volume of a sick child weighing 14.0 pounds in order to determine the appropriate treatment for the child. Which would be more useful to the doctor – the confidence interval from (b) or the prediction interval from (c)? Explain.

3. *The Statistical Sleuth*, 7.22. The meat processing study is case study 7.1.2. The formula for $SE(\hat{\beta}_0)$ is in Section 7.3.5 rather than Section 7.3.4.

4. *The Statistical Sleuth*, 7.29. Write a summary report in the form of Case Study 7.1.1, containing a summary of statistical findings and scope of inferene section.