Stat 333: Applied Regression Analysis
Spring 2015
Due Date: Friday, February 6 in class

Relevant text chapters: Ch. 1

Instructions: You may (and are encouraged to) discuss homework problems with other students, but the solutions that you should provide should be your own and not directly copied from another student. Show your work wherever possible, and write out the formulas you used to arrive at your solutions. If you have any questions or difficulties, you are more than welcome to come see me in my office. Note that datasets for textbook problems can be found on the CD that came with the textbook or can be downloaded from

https://netfiles.umn.edu/users/nacht001/www/nachtsheim/index.html

1. An alternate expression for the slope coefficient of the simple linear regression model is

$$\hat{\beta}_1 = r\frac{s_Y}{s_X}$$

where r is the Pearson correlation coefficient given by

$$r = \frac{SS_{xy}}{\sqrt{SS_{xx}SS_{yy}}}$$

and s_Y and s_X are the sample standard deviations of *Y* and *X*, respectively. Show that this alternate formulation is equivalent to what we learned in class, that is, prove that

$$r\frac{s_Y}{s_X} = \frac{SS_{XY}}{SS_{XX}}$$

2. Prove the following properties of the residuals e_i and the fitted values $\hat{Y}_i$:

a) $\sum_i e_i = 0$

b) $\sum_i Y_i = \sum_i \hat{Y}_i$

c) $\sum_i X_i e_i = 0$

d) $\sum_i \hat{Y}_i e_i = 0$

3. In class, we used the woodpecker nest cavity data to demonstrate that using a centered model $Y_i = \beta_0{}^* + \beta_1(X_i - \bar{X}) + \varepsilon_i$ resulted in the estimator of $\beta_0{}^*$ being $\bar{Y}$. Using this same data, fit the model $Y_i - \bar{Y} = \beta_1(X_i - \bar{X}) + \varepsilon_i$, where both the predictor and the response variables are centered. Print your output. What is your intercept of the Y-axis for this model?

4. Ch.1, Problem 1.19 (p.35)
Note: Plotting the estimated regression function can be obtained in R using the command abline(lm.fit), assuming that lm.fit is the name that you gave to the output from the function lm. For this problem make sure to print out your R output from summary(lm.fit)

5. Ch. 1, Problem 1.23 (p. 36). Also for this problem, print out the ANOVA table.

6. Ch. 1, problem 1.41 (p.38), part a). If we assume for this model that ε_i are independent and follow a normal distribution with mean 0 and variance σ^2, show that the estimator of β_1 is unbiased, that is $E(\hat{\beta}_1) = \beta_1$