STAT 308 – Homework 4

For the problems in which calculations are needed, please include your `R` code with your answers, otherwise you will not be given full credit. Please upload your assignment by Thursday, September 29, 11:59 pm in a pdf file to Sakai.

- 1. Suppose we perform a simple linear regression where

$$n = 50, \bar{x} = -0.208, \bar{y} = 1.516, s_x = 2.354, s_y = 3.185$$

$$\hat{\beta}_0 = 1.745, \hat{\beta}_1 = 1.102, s_{Y|X} = 1.868, s_{\hat{\beta}_0} = 0.265, s_{\hat{\beta}_1} = 0.113$$

– a. Calculate the sample correlation coefficient, r, and the r^2.

– b. Calculate the estimate of the regression variance $s^2_{y|x}$. (Hint: Intuitively, this is the variance of Y that is not explained through the linear model with X.)

- 2. Use the following incomplete ANOVA table to answer the following questions.

	df	Sum Sq	Mean Sq	F value	Pr(>F)
Model	1	1.47		0.18	
Error					
Total	34				

– a. What is the mean squares for the model (MSM)?

– b. What is the mean squared error (MSE)?

– c. What are the error degrees of freedom?

– d. What is the sum of squared errors (SSE)?

– e. What is the p-value used to test for a significant linear relationship between X and Y?

- 3. Reconsider the dataset `AdRevenue.csv` as well as our simple linear regression model of ad revenue (in millions of dollars) based on circulation (in millions).

– a. Obtain an ANOVA table for this model.

– b. Using the ANOVA table, perform a hypothesis test for a linear relationship between ad revenue and circulation. Be sure to properly state your hypotheses, your test statistic, p-value, and a decision and conclusion at $\alpha = 0.05$.

– c. What is the distribution the test statistic follows under H_0? In other words, what is the distribution we use to calculate the p-value?

– d. Using the ANOVA table, calculate the value of r^2. Interpret this value in the context of the problem.