**STA 6166 – Fall 2013 – Exam 4 – PRINT Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Conduct all tests at  = 0.05 significance level**

Q.1. A multiple regression equation was fit for n = 45 observations using 5 independent variables X1, X2,…, X5 gave SS(Residual) = 780. What is the residual standard deviation (standard error of estimate)?

Q.2. A multiple regression equation was fit for n = 28 observations using 5 independent variables X1, X2,X3,X4 gave SS(Total) = 1500 and SS(Residual) = 600.

p.2.a. Calculate the value of the coefficient of determination.

p.2.b. Test the hypothesis that all the slopes are zero. H0: =0

Test Statistic \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Rejection Region: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q.3. Bob fits a regression model relating weight (Y) to weight (X1) for professional basketball players, with a dummy variable for males (X2 = 1 if Male, 0 if Female). Cathy fits a model **on the same dataset**, but she defines X2 = 1 if Female, 0 if Male.



Q.4. The ANOVA tables for fitting Y as a linear function of X are shown below. In the first table the “independent variables” include X1, the continuous variable, X2 and X3 as dummy variables to denote the three groups, and X12 and X13 representing the cross-products of X1 and the two dummy variables. The second table is the ANOVA table for fitting Y as a linear function of X1, X2, X3.





p.4.a. Complete the tables.

p.4.b. For the second model, test H0 = 0

Test Statistic \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Rejection Region \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

p.4.b. Is there significant evidence the slopes are not equal among the 3 groups?



Test Statistic \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Rejection Region \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Q.5. In the production of a certain chemical it is believed the yield, Y, can be increased by increasing the amount of a particular catalytic agent. Twenty trials were made with different amounts of the catalyst. Analysis of the yields, measured in grams, and amounts of the catalyst, X, in milligrams gave based on the following model:



‾x = 10 ‾y = 139.0 Σ(x - ‾x)2 = 500 Σ(y -‾y)2 = 895 Σ(x -‾x)(y -‾y) = 350

p.5.a. Compute the estimated slope.

p.5.b. Compute the estimated y-intercept.

p.5.c.  Compute the estimate of the residual standard deviation: Se

p.5.d. Compute a 95% Confidence Interval for 1:

Q.6. A regression model was fit, relating revenues (Y) to total cost of production and distribution (X) for a random sample of n=30 RKO films from the 1930s (the total cost ranged from 79 to 1530):



p.6.a. Obtain a 95% Confidence Interval for the **mean revenues for all movies** with total costs of x\*= 1000





p.6.b. Obtain a 95% Prediction Interval for **tomorrow’s new film** that had total costs of x\*= 1000

