

MATH 739/839

Homework 1

Total Points: 25 Undergraduates/30 Graduates & Honor Students

Homework Instructions: The following problems are intended to give you an opportunity to apply the concepts discussed in lecture on simple linear regression. Please be sure to follow all instructions, answer each question completely and include the relative statistical software output. You may discuss questions with other students but must write up your own solutions. Copying will not be accepted.

Submission Instructions: You will be submitting your homework in myCourses. Save the file for the homework and upload to myCourses with the title *MATH739_839_HMWK1_Name*. Be sure to input your name into the title of the file.

The gasoline milage of 32 automobiles can be found in **MilesPerGallonMPV.JMP**.

- 1) (15 pts) The dataset **MilesPerGallonMPV.JMP** contains the gasoline mileage for 32 automobiles. Using the statistical software package of preference, answer the following questions.
 - a) (3 pts.) Fit a simple linear regression model (Fit Y by X platform) relating the response MPG (miles per gallon) to the regressor Displacement (engine displacement in cubic inches). From the scatterplot, does the straight-line model appear a good approximation to the relationship? Briefly explain. Identify the slope and the intercept estimates of the estimated regression line. Interpret the values if appropriate.
 - b) (2 pts.) From the ANOVA table does the regression appear significant? Briefly discuss why. Please include a copy of the ANOVA table with your answer.
 - c) (2 pts.) From the Parameter Estimates table are both the intercept and slope significant? Explain. Please include a copy of the Parameter Estimates table with your answer.
 - d) (2 pts.) From the Parameter Estimates table find 95% CI's for the slope and intercept. Based on the two CI's do both the slope and intercept appear to be significant (nonzero)? Explain. Please include a copy of the Parameter Estimates table with your answer.
 - e) (1 pt.) What percent of the total variation in MPG is explained by Displacement? Identify and interpret the percent. Please include the Summary of Fit table with your answer.
 - f) (2 pt.) What is the predicted MPG at Displacement = 275 in^3 . Find a 95% confidence interval for the true mean of MPG at Displacement = 275 in^3 . There are several ways of solving the problem. Please provide software output.
 - g) (2 pt.) Find a 95% prediction interval for possible future MPG values at Displacement = 275 in^3 . There are several ways of solving the problem. Please provide software output.
 - h) (1 pt.) Compare the two intervals found in part f and in part g and explain the difference; i.e., which one is wider and why is it wider?
- 2) (10 pts) The purity of oxygen produced by a fractional distillation process is thought to be related to the percentage of hydrocarbons in the main condenser of the processing unit. The twenty samples can be found in the data set: **OxyDistillationMPV.JMP**. Using the statistical software package of preference, answer the following questions.
 - a) (2 pts.) Using Fit Y by X, fit a SLR model with Purity% as the response and Hydrocarbon% as the regressor. From the scatterplot, does the straight-line model appear a good approximation to the relationship? Briefly explain. Identify the slope and the intercept estimates of the estimated regression line. Interpret the values if appropriate.

MATH 739/839

Homework 1

Total Points: 25 Undergraduates/30 Graduates & Honor Students

- b) (2 pts.) Based on the ANOVA table and the Parameter Estimates table does the regression appear to be significant? Briefly discuss why. Please include relevant output with your answer.
 - c) (2 pts.) Using the Parameter Estimates table find a 95% CI for the true slope value. Does the CI indicate a significant relationship between Purity% and Hydrocarbon%? Explain. Please include relevant output.
 - d) (2 pts.) Calculate R^2 by hand and confirm the value to the one computed by the software package. Do not forget to include the relevant output.
 - e) (2 pts.) What is the predicted Purity% at Hydrocarbon% = 1.0? Find a 95% CI for the true mean Purity% at Hydrocarbon% = 1.0. There are several ways of solving the problem. Please provide software output.
- 3) (5 pts) (Required for 839 students & Honor students. Extra credit for 739) Suppose that we are fitting a straight line and wish to make the standard error of the slope as small as possible. Suppose that the “region of interest” for x is $-1 \leq x \leq 1$. Where should the observations x_1, x_2, \dots, x_n be taken? Discuss the practical aspects of this data collection plan. Software analysis is not needed for this problem.