

|  |
| --- |
| Computational Assignment #2: Statistical Inference in Linear Regression  *MSDS 410* |

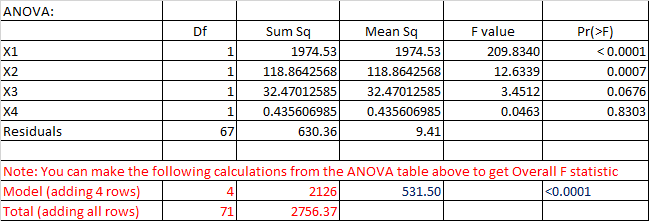
In this **assignment you will review model output from R and perform hypothesis tests and computations related to statistical inference for linear regression**. Show all work in your computations. A good practice is to **write down the generic formula for any computation and then fill in the values needed for the computation from the problem statement.** Throughout this assignment keep all decimals to **four places**, i.e. X.xxxx. You are expected to use correct **notation** and **terminology**, and to be clear, complete and concise with all **interpretations** of results.

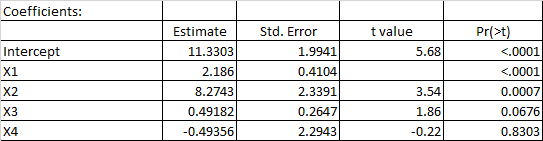
Any computations that involve “**the log function**”, denoted by log(x), ***are always meant to mean the natural log function (which will show as ln() on a calculator).*** The only time that you should ever use a log function other than the natural logarithm is if you are given a specific base.

**MECHANICS AND COMPUTATIONS (100 points)**

**Model 1:** Consider the following R output for a regression model which we will refer to as Model 1. (Note 1: In the ANOVA table, I have added 2 rows – (1) Model df and Model SS - which is the sum of the rows corresponding to all the 4 variables (2) Total DF and Total SS - which is the sum of all the rows;

Note 2: **The F test corresponding to the Model denotes the overall significance test.** In R output, you will see that at the bottom of the Coefficients table)



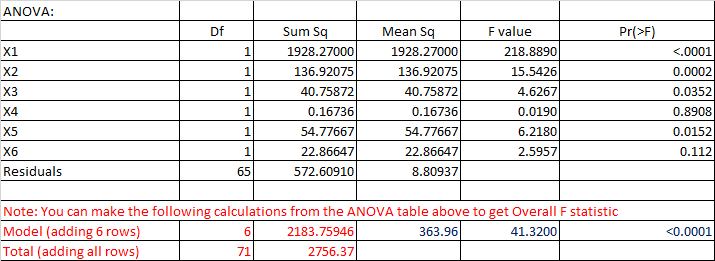


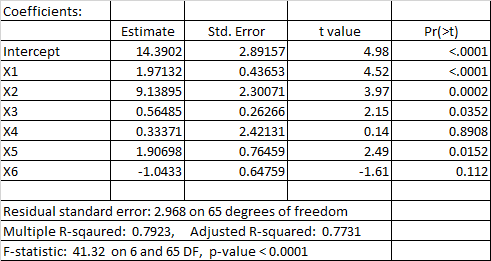




1. (10 points) How many observations are in the sample data?
2. (10 points) Write out the null and alternate hypotheses for the t-test for Beta1.
3. (10 points) Compute the t- statistic for Beta1. Conduct the hypothesis test and interpret the result.
4. (10 points) Compute the R-Squared value for Model 1, using information from the ANOVA table. Interpret this statistic.
5. (10 points) Compute the Adjusted R-Squared value for Model 1. Discuss why Adjusted R-squared and the R-squared values are different.
6. (10 points) Write out the null and alternate hypotheses for the Overall F-test.
7. (10 points) Compute the F-statistic for the Overall F-test. Conduct the hypothesis test and interpret the result.

**Model 2:** Now let’s consider the following R output for an alternate regression model which we will refer to as Model 2.







1. (10 points) Now let’s consider Model 1 and Model 2 as a pair of models. Does Model 1 nest Model 2 or does Model 2 nest Model 1? Explain.
2. (10 points) Write out the null and alternate hypotheses for a nested F-test using Model 1 and Model 2.
3. (10 points) Compute the F-statistic for a nested F-test using Model 1 and Model 2. Conduct the hypothesis test and interpret the results.

**Assignment Document:**

Results should be presented and discussed in the numerical order of the questions given. The report should not contain unnecessary results or information. Tables are highly effective for summarizing data across multiple models. The document MUST be submitted in pdf or html format. Please use the naming convention: CompAssign2\_YourLastName.pdf.