Assignment #4: Statistical Inference in Linear Regression (50 points)

This assignment will be made available in both pdf and Microsoft docx format. Answers should be typed into the docx file, saved, and converted into pdf format for submission. **Color your answers in green so that they can be easily distinguished from the questions themselves.**

Throughout this assignment keep all decimals to four places, i.e. X.xxxx.

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.

In this assignment we will review model output from SAS and perform the computations related to statistical inference for linear regression. By performing this computations we are ensuring that we understand how the numbers in this SAS output are computed.

Model 1: Let's consider the following SAS output for a regression model which we will refer to as Model 1.

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	2126.00904	531.50226		<.0001
Error	67	630.35953	9.40835		
Corrected Total	71	2756.36857			

Root MSE	3.06730	R-Square	
Dependent Mean	37.26901	Adj R-Sq	
Coeff Var	8.23017		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	11.33027	1.99409	5.68	<.0001
X1	1	2.18604	0.41043		<.0001
X2	1	8.27430	2.33906	3.54	0.0007
X3	1	0.49182	0.26473	1.86	0.0676
X4	1	-0.49356	2.29431	-0.22	0.8303

Number in Model	C(p)	R-Square	AIC	BIC	Variables in Model
4	5.0000	0.7713	166.2129	168.9481	X1 X2 X3 X4

(1) (5 points) How many observations are in the sample data?

72 observations

(2) (5 points) Write out the null and alternate hypotheses for the t-test for Beta1

H0: Bj = 0 Ha: Bj not equal to 0

The null hypothesis is that the co-efficient is not significantly different than zero.

The alternative hypothesis is that the co-efficient is significantly different than zero

(3) (5 points) Compute the t- statistic for Beta1.

5.3262

(4) (5 points) Compute the R-Squared value for Model 1.

0.7713

(5) (5 points) Compute the Adjusted R-Squared value for Model 1.

0.8173

(6) (5 points) Write out the null and alternate hypotheses for the Overall F-test.

H0: B1 = B2=… Bp=0

Ha: B1 not = 0 or b2 not equal to 0 or Bp not = to 0

The null hypothesis is that the co-efficient are not jointly significantly different from zero.
The alternative hypothesis is that the the co-efficient are jointly significantly different from zero.

(7) (5 points) Compute the F-statistic for the Overall F-test.

56.4926

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

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	6	2183.75946	363.95991	41.32	<.0001
Error	65	572.60911	8.80937		
Corrected Total	71	2756.36857			

Root MSE	2.96806	**R-Square**	0.7923
Dependent Mean	37.26901	**Adj R-Sq**	0.7731
Coeff Var	7.96388		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	14.39017	2.89157	4.98	<.0001
X1	1	1.97132	0.43653	4.52	<.0001
X2	1	9.13895	2.30071	3.97	0.0002
X3	1	0.56485	0.26266	2.15	0.0352
X4	1	0.33371	2.42131	0.14	0.8908
X5	1	1.90698	0.76459	2.49	0.0152
X6	1	-1.04330	0.64759	-1.61	0.1120

Number in Model	C(p)	R-Square	AIC	BIC	Variables in Model
6	7.0000	0.7923	163.2947	166.7792	X1 X2 X3 X4 X5 X6

(8) (5 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.

Model 1 is nested within model 2, because model two has all the terms of model 1 plus it has additional terms in it, model 1 is the reduced model and model 2 is the full model.

(9) (5 points) Write out the null and alternate hypotheses for a nested F-test using Model 1 and Model 2.

Hypotheses: H0: $\beta_{k+1} = \beta_{k+2} = \ldots = \beta_{k+p} = 0$

Ha: At least one $\beta \neq 0$ (beta not equal to zero)

(10) (5 points) Compute the F-statistic for a nested F-test using Model 1 and Model 2.

Test statistic: F = (SSER−SSEC)/ # of additional β's SSEC /[n−(k+p+1)]

=630.35953 -572.60911/2/67-65/572.690/65

= **3.2769**