

# Statistical Methods II AEMA-610

## Assignment 1

Date 2014/09/19  
Date due 2014/09/26

We carry out a study to look at the effect of several variables (X1, X2, X3 and X4) on Y. We obtain the following data:

Observation	Y	X1	X2	X3	X4
1	51.4	0.2	17.8	24.6	18.9
2	72.0	1.9	29.4	20.7	8.0
3	53.2	0.2	17.0	18.5	22.6
4	83.2	10.7	30.2	10.6	7.1
5	57.4	6.8	15.3	8.9	27.3
6	66.5	10.6	17.6	11.1	20.8
7	98.3	9.6	35.6	10.6	5.6
8	74.8	6.3	28.2	8.8	13.1
9	92.2	10.8	34.7	11.9	5.9
10	97.9	9.6	35.8	10.8	5.5
11	88.1	10.5	29.6	11.7	7.8
12	94.8	20.5	26.3	6.7	10.0
13	62.8	0.4	22.3	26.5	14.3
14	81.6	2.3	37.9	20.0	0.5

NOTE. Show sufficient calculations. Use a 1% probability level for all statistical tests.

Q1. We consider that X1, X2, X3 and X4 may affect Y in a linear manner. Specify a suitable multiple regression model (only linear terms, no interactions).

Define the terms and parameters in your model. 4 points

Q2. From the equation developed in Q1 write out the equation for each observation. 2 points

Q3. Write out these same equations, but in matrix format. Clearly label the dimensions of all matrices. 2 points

Q4. Write out the Normal Equations, in matrix format, and also as a set of simultaneous equations. 4 points.

Q5. Obtain, using SAS PROC/IML, an estimate of  $\sigma^2$ , estimates (and standard errors) for b. 6 points

Q6. Specify in words and in matrix form (formal statistical notation) your Null and Alternative Hypotheses for testing of each of X1, X2 X3 and X4. 4 points

Q7. Construct a suitable Analysis of Variance table. Test, using F-tests, whether or not any of the regression covariables can be considered to be statistically significant or not. Clearly answer, in words, what hypotheses you accept or reject Clearly indicate the tabulated F values you use for each test. 8 points

Use PROC IML to compute the necessary values for Q4-7. Clearly label and cross-reference your written answers to your SAS Program Editor, log and output results so that I can follow your computations.