**MULTIVARIATE DATA ANALYSIS (*BIA 652)***

Spring 2017

Homework 1

**SIMPLE LINEAR REGRESSION**

1. As the manager of a bus company you are interested to know if there is a relationship between the annual maintenance cost of a bus and its age. Following data gives the age and maintenance cost of the buses in the compay.

Bus Maintenance Cost ($) Age (years)

Y X

1 859 8

2 682 5

3 471 3

4 708 9

5 1094 11

6 224 2

7 320 1

8 651 8

9 1049 12

1. Plot a scatter diagram
2. Determine the regression line
3. Determine Total Sum of Squares
4. Determine Sum of Squares about regression
5. Determine Sum of Squares due to regression
6. Show

Total Sum of Squares = Sum of Squares due to regression + Sum of Squares about regression

1. Find a 95% confidence interval for the slope
2. Find a 90% confidence interval for the intercept
3. Compute R2
4. Compute correlation between Maintenance Cost and Age
5. Compare R2 with correlation2
6. Forecast the annual maintenance cost for a bus that is 4.5years old and one that is 7 years old.

**CENTRAL LIMIT THEORM**

1. Consider a die with faces numbered 1,2,3,4,5, and 6. The probability that a face occurs from a random roll of the die is 1/6.
   1. What is the mean and standard deviation of a random roll of this die?
   2. Use random number generator to simulate 100 rolls of this die and calculate the following number:

(∑ Xi – n μ)/√n σ = (∑ Xi – 100 mean)/(√100 \* Standard deviation). Repeat b 1000 times and draw the histogram of these 1000 numbers.

* 1. Change the numbers on the face of die to 1, √2, 3, 16, 5, √6 and the probability of these numbers occurring to 5%, 10%, 20%, 30%, 25% , and 10% respectively. Repeat part a and b