**Math 324 Extra-Credit Lab**

**Background**: recording values of two numerical variables for each individual, a question often arises if there is a linear dependence between the variables. A good starting point for such a study is an “XY scatterplot” – a chart where pairs of values for each individual are plotted as points on the xy-plane. Should the points fit a straight line, more or less, one can find the equation of the “best-fit regression line” and characterize the closeness of points to the line by the so-called “correlation coefficient r”. The equation can be used to compute the value of Y for a given X. The goal of this lab is to study that using the material of Ch.12 of our textbook and use Excel to present those concepts graphically.

1. Label A1”Error” and put 40 random numbers in A2:A41 using Normal Distribution with mu = 0 and sigma = 5. Label B1 “X-data” and put 40 random numbers in B2:B41 using Uniform Distribution on the interval [30, 50] (use a different seed number for that to avoid possible interdependence of the generated data). Label C1 “Y-data” and put 40 numbers in C2:C41: enter the formula of the model “C2=15+1.5\*B2+A2” in C2 and drag the formula down to C41.

2. Put the Descriptive Statistics Report on the X-data in the range from D1. Make a histogram of the X-data, put it in the range from F1. Do the same thing for the Y-data: Descriptive Statistics report in the range from D21 and the histogram in the range from F21.

3. Use the Regression tool in the Data Analysis toolkit on the XY-data in columns B and C. In the request form, check the Line Fits and Normal Probability Plot windows.

4. Make a verbal report that should include:

a. Your name, section, “XCR”, and both seed numbers used to generate random numbers;

b. Separate reports on the X and Y data: summary statistics, outliers and histograms;

c. Report on the XY association: direction, strength, coefficients of correlation and determination, formula of the regression line. Does the Normal Probability plot of residuals support the assumption of Normality?

5. Organize your printout into two parts: one for individual X, Y analyses, another – for the association analysis (graphs and reports). No more than three pages will be accepted. Sign.