ASSIGNMENT #5 – Statistical Sampling and Random Value Generation

DUE: At the beginning of your recitation section, October 31-November 2, 2006.

NOTE: This assignment is to be performed in groups of 2 or 3, all from your recitation section. The same group will work together on Assignment #4. Before leaving recitation today, sign up as a group. Turn in a single write-up, with all names listed on the first page.

In a manufacturing environment, statistical sampling is often used to evaluate component or product characteristics, whether for incoming inspection or to monitor specific manufacturing processes. The results of this sampling process are used to estimate the probability that a component or product, chosen at random from the entire "pool," will meet or fail to meet specifications. The validity of the sampling process depends on a number of factors, including sample size.

For this assignment, you will use MATLAB or EXCEL to simulate the sampling process. The principal difference is that you will know the actual characteristics of the "pool" which generated the samples. Thus, instead of trying to estimate those characteristics, you will be observing the effectiveness of the sampling process.

Using either MATLAB or EXCEL, generate 1000 random numbers (the "samples") for each of the following sets of characteristics. You may need to use the "help" function in EXCEL, along with some basic principles of probability, to figure out how to generate the random samples, particularly for the normal distribution.

1. Uniform distribution between 560 and 640.

2. Gaussian (normal) distribution with a mean of 600 and a standard deviation of 13.

3. Gaussian (normal) distribution with a mean of 600 and a standard deviation of 7.

For each of these three sets, compute the actual minimum, maximum, mean, and standard deviation for the sample set after the first 10, 50, 100, and 500 samples, as well as for the entire set of 1000 samples. Generate a table similar to the following for each of the three sets. (Obviously, your values will differ.)

# values	Min	Avg	max	stdev
10	51.775	55.554	59.604	2.973
50	50.137	54.645	59.718	3.108
100	50.055	54.838	59.946	2.977
500	50.055	54.864	59.958	2.935
1000	50.006	54.880	59.958	2.900

Additionally, plot histograms for the sample data after 10, 100, and 1000 samples for each case. Use a range of 520 to 680 for the x axis, with 40 bins (bin size of 4). Clearly label each table and plot.

On a cover page, describe the method you used to generate each of the sample sets (e.g., the MATLAB or EXCEL formulas) and briefly summarize your observations regarding the use of sampling as a method to determine the characteristics of a larger pool of devices. Attach all of your tables and histogram plots.