Lab 4 (100 pts. + 25 pts. BONUS) – Central Limit Theorem
Objectives: A better understanding of the Central Limit Theorem

This is a group lab so only one report should be submitted per group. There should be 3 – 4 people in each group. It is acceptable that each person does one or two distributions and then discuss the results with the rest of their group to write a combined summary statement. More than one software package may be used in this lab.

To help you understand the Central Limit Theory, you are going to be simulating the distribution of the mean ($\bar{X}$) for four different distributions: normal, uniform, gamma and Poisson. The distribution of $\bar{X}$ is called a sampling distribution. For each of the population distributions, you will be averaging different numbers of randomly generated distributions from the same parent population. This value which is called n will be given to you and may or may not be different for the different population distributions. You will learn in Chapter 7 that the sampling mean and standard devotion are:

$$\mu_{\bar{X}} = \mu_X \qquad \text{Equations 1}$$
$$\sigma_{\bar{X}} = \frac{\sigma_X}{\sqrt{n}}$$

where $\mu_{\bar{X}}$ is the mean of the sampling distribution, μ_X (or μ) is the mean of the population, $\sigma_{\bar{X}}$ is the standard deviation of the sampling distribution, σ_X (or σ) is the standard deviation of the population and *n* is the number of samples averaged. When *n* is large, the distribution of $\bar{X}$ is approximately normal, that is

$$\bar{X} \dot{\sim} \mathcal{N}\left(\mu, {\sigma^2}/{n}\right) \qquad Equation\ 2$$

In this lab, you will create 1000 random samples of size *n* for each of the distributions below. The number of samples that you will be averaging, *n* depends on the given distribution. For each of the 1000 random samples, compute the sample mean, $\bar{x}$. So you will have 1000 $\bar{x}$'s for each given sample size *n*.

For each of the distributions::

1. (5 pts.) Code
   You only need to provide one code listing for each distribution.

2. (10 pts) Histogram/normal quantile plots
   For each of the values of *n*, submit a histogram (with the two colored lines) and a normal quantile plot. For each of the graph pairs, indicate whether the situation is normal or not.

3. (5 pts.) Summary table
   This table contains the experimental mean and standard deviation calculated from the data (output is required) and the theoretical mean and standard deviation calculated from Equations 1 (with work for one of the values for each distribution). The format for this table for Part A is below.

   For standard normal part (A);

n	experimental mean of your 1000 $\bar{x}$	theoretical mean (Equations 1)	experimental standard deviation of your 1000 $\bar{x}$	theoretical standard deviation (Equations 1)
1				
2				
6				
10				

4. (5 pts.) Concluding remarks.
   In this part, you are to write a conclusion (complete sentences in English) summarizing the results of Parts 2 and 3. There should be one sentence summarizing what happens to the shape as n increases and what value of n is considered 'large' (Part 2). The second sentence should contain whether Equations 1 are valid for all values of n (Part 3).

The distributions with the values of n (the number of samples to average) are below: I have included the population means and standard deviations for the distribution that we have not yet covered in class.

A. (25 points) Standard Normal Distribution. n = 1, 2, 6 and 10.

B. (25 points) Uniform distribution over the interval (0,3). n = 1, 2, 9 and 16.

C. (25 points) Gamma distribution with parameters α = 2 and β = 1. n = 1, 5, 10, 20, 30, 40, until the shape becomes normal. This distribution has population mean and standard deviation of $\mu = 2$, $\sigma = \sqrt{2}$.

D. (25 points) Poisson distribution with parameter λ = 2. n = 1, 5, 10, 20, 30, 40, until the shape becomes normal.

E. (BONUS: 25 points) Exponential with parameter λ = 2. n = 1, 5, 10, 20, 30, 40, 50, until the shape becomes normal.