**Statistics 251: Lab 5 Handout**

First Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Last Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Student Number: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Please write down your answers neatly and do show your work. Give your R code as well as the plots.**

1. Observe the shape of the **probability density function** of X. **(5 min)**
2. Draw a random sample of size 500 and construct a histogram. [Hint: replace the *frequency* on the vertical axis with the *density* using prob=TRUE in the hist command]. Comment about the shape of the histogram (Skewed or approximately symmetric? If the histogram is skewed, in which direction is it?).
3. Does the shape of the histogram represent the original distribution from which the sample was drawn? [Hint: use curve function to add the curve that represents the original distribution to the histogram, adjust from and to values from previous histogram horizontal axis].
4. Observe the **distribution of *sample means*** of “small” sample sizes. **(10 min)**
   1. Draw 1000 samples of size 5, find the *sample mean* of each sample and draw a histogram of the *sample means*. [Hint: use a matrix to store each sample and use the apply function to find the mean of each sample].
   2. Draw a density plot of the sample means obtained above. [Hint: use plot(density(\_)) function]. Comment about the shape of the distribution of *sample means* (Skewed or approximately symmetric? If the histogram is skewed, in which direction is it?).
5. Observe the **distribution of *sample means*** of “large” sample sizes. **(10 min)**
   1. Draw 1000 samples of size 50, find the *sample mean* of each sample and draw a histogram of the *sample means*.
   2. Draw a density plot of the *sample means* obtained above. Comment about the shape of the distribution of *sample means* (Skewed or approximately symmetric? If the histogram is skewed, in which direction is it?).
6. What is the **effect of sample size** on the distribution of *sample means*? **(5 min)**
   1. Compare the *center* and *spread* of the two distributions of the *sample means*. Are they approximately the same or different? If different, how does each differ?
      1. Center
      2. Spread
   2. What can you conclude about the effect of sample size on the distribution of sample means?