R Tutorial for STAT 350 Lab 4

Author: Leonore Findsen; Jeremy Troisi

1. Generate sampling distributions of the mean

The assignment describes the Central Limit Theorem and how we will be illustrating its effect in this lab. Please check the details in the assignment document. The tutorial elaborates a little more on the procedure you will follow even though the software package doesn't calculate it using this method.

1) Generate 1000 random samples, each of size n from each of the distributions. In this illustration, we are putting these values into the columns A, B, … In the assignment, you will be asked to implement the steps using different given values of n.

2) Average each row and store the result in another dataset variable. For example, in the following table, we are averaging six observations (columns) from 5 different random samples (rows) and putting the average of each row into column G.

	A	B	C	D	E	F	G (average)
1	-0.30	-1.28	0.24	1.28	1.20	1.73	0.478
2	-2.18	-0.23	1.10	-1.09	-0.69	-1.69	-0.797
3	-1.85	-0.98	-0.77	-2.12	-0.57	-0.40	-1.115
4	0.13	-0.37	-0.33	-0.37	1.34	-0.09	0.052
5	-0.19	-0.51	1.97	0.87	2.38	-0.65	0.645

3) From column G, generate the histogram, the QQ plot, mean, and standard deviation.

code:

```
SRS <- 1000 #the number of repeats (not to be changed)

# n: the number of columns that are being averaged over
n <- 1

# I strongly suggest that you change the titles to include the distribution.
# The number of columns that are being averaged over will automatically
# be updated.
title <- paste("Normal distribution: averaged over ", n)

# calculates the average data
data.vec <- rnorm(SRS*n, mean = 0, sd = 1) #creates random normal data
data.mat <- matrix(data.vec, nrow = SRS) #separates the data into rows

#apply(your data table/matrix, row/column index, function):
# The apply() function can apply the function of your choice to each row or
#     each column of your data table/matrix.
# Use 1 for the second argument ("row/column index") to apply the function to
#     each row; and use 2 for each column.
avg <- apply(data.mat, 1, mean) #performs the averaging

# Please complete the code for histogram, QQ plot, mean, and SD.
```

R Tutorial for STAT 350 Lab 4

Author: Leonore Findsen; Jeremy Troisi

The data that you want to analyze is in the variable '`avg`.' You will need to include the rest of the code as appropriate to produce what is required in the assignment. Please see Lab 3 for details.

2. Generate random samples

In Lab 3, we have already discussed how to generate a normal distribution, exponential, beta, uniform, and t-distributions.

Use the following commands to generate the other two distributions in this lab:

```
data.vec <- rgamma(SRS*n, 3, rate = 2) #creates random data for gamma

data.vec <- rpois(SRS*n, 3) #creates random data for Poisson
```

Remember to use the appropriate data set when you are generating the histogram and the normal probability plot. **I would also suggest changing the title so that you know what is being plotted.**