TMATH 390 R Lab 5

You will submit your R labs online in Canvas as two separate documents. First you will submit the R script that you use to complete the lab. The R script should have a commented header that identifies you as the author, the class, date, and assignment. Then all of the CODE you use to complete the lab, with comments describing what the code is meant to accomplish. See the example R script on Canvas for guidelines. The R script will be graded on completeness and form (e.g., comments and header). The results of running the R script will be graded by what you produce in your lab document.

The second file will be a document that records the RESULTS of running the R code. Only the results that are indicated in the instructions should be included. This should be one document with all of the resulting calculations, graphics, or written interpretations. For your lab document make sure to label each problem, and to complete ALL parts of each problem. Your lab document should be completed in Word (or some other word processing program) then printed as a pdf for final submission.

Note: If your R script includes code to generate a graphic or calculation, and the resulting graphic or calculation is not included in your lab document, you will NOT be given credit for that problem!

Make sure to give your digital documents a reasonable title that includes either your name or initials, and the assignment label (e.g., RLab1_MCK.pdf; RLab1Script_MCK.R).

Objectives

1. Use R to calculate probabilities for the Poisson and binomial distributions.
2. Use R to explore how the shape of the Poisson and binomial distributions change with different parameter values.
3. Use R to draw random numbers from the Poisson and binomial distributions.

In this lab we will learn the R functionality for two of the thoeretical distributions we have learned in class: Poisson and binomial. We will also construct some boxplots and histograms.

As always when you start a new session, make sure to change your working directory to the class directory using setwd() or the windows.

C1. (4) **Submit your R script to Canvas**. You can upload them directly to your assignment as a *.R document. Make sure to include comments throughout that explain your code, and a commented header at the top that identifies you as the author, the assignment, and the date.

Poisson distribution

From Lecture 8 practice problems, recall we evaluated a Poisson distribution with a mean rate of 4.5 clams/12 hours.

In R, for a discrete distribution the “d” function (e.g., dpois) returns the probability (the height of the mass function) for a particular value of X (remember, for a continuous function this is not a probability– it is the height of the density function at that value of X). See the next page for an example.

For example:

```
# 0:15 returns a vector of integers: (0,1,2,...,15). First we create an
# object (x.vals) that contains this vector
x.vals=0:15
# The function dpois evaluates the Poisson mass function at
# each value of x and returns a vector of probabilities
pois1.vec=dpois(x = x.vals,lambda = 4.5)

# And here we plot those values, with x.vals on the x-axis and
# the Poisson probabilities on the y-axis. The argument type="b"
# specifies that we want both points and lines.
plot(0:15,pois1.vec,type="b",pch=16)
```

C2. (1) Copy and paste the plot generated using the code above.

C3. (1) Use R to find P(x=5) for Pois(4.5). Return your R command and the result.

You can evaluate dpois at multiple x simultaneously, and you can add them up using the sum function:

```
# evaluates the Poisson distribution at x = {0,1,2}, then sums
# the probabilities
dpois(0:2, 4.5)

## [1] 0.01110900 0.04999048 0.11247859

sum(dpois((0:2),4.5))

## [1] 0.1735781

ppois(2,4.5)

## [1] 0.1735781
```

C4. (2) Use this information to find P(x $\geq$ 3) (there are at least 3 clams) and explain how you got it.

C5. (2) Create and include a plot that explores how the Poisson distribution changes with different values of λ: {1,2,3,4,5,6,7,8}.

Below is some code to get you started. Make sure to increase the size of your plotting window so all of the graphs fit and are readable!

```
# use x.vals again, defined as 0:15
x.vals=0:15
# use par to set up a 3x3 graphing window
par(mfrow=c(3,3))
# now create a graph for lambda=1
plot(x.vals,dpois(x.vals,1),type="b")
# lambda=2
plot(x.vals,dpois(x.vals,2),type="b")
# now repeat for the additional values of lambda
```

C6. (1) Describe what happens to the shape of Poisson distribution with increasing λ (increasing expected value).

Binomial distribution

The relevant functions for the binomial distribution are dbinom pbinom. The have the syntax: dbinom(x,size,prob), pbinom(q,size,prob)

C7. (1) What do "size" and "prob" represent in these function calls?

C8. (1) Using R, determine the probability x $\leq$ 3, where X~binom(100,0.005). Hint: adapt the code from C4 above.

C9. (2) Create and include a plot that explores how the binomial distribution changes with different vaues of π and n. Plot all of the distributions on one sheet for the following combinations of n and π (with x = 0:20):

n	5	5	5	5	10	10	10	10	20	20	20	20
pi	0.01	0.1	0.5	0.9	0.01	0.1	0.5	0.9	0.01	0.1	0.5	0.9

The first few lines of code are given below. Complete for the remaining values of n and π.

```
# set up a plotting area with 3 rows and 4 columns.
par(mfrow=c(3,4))
plot(dbinom(0:20,5,0.01),type="b")  #n=5, pi=0.01
plot(dbinom(0:20,5,0.1),type="b")   #n=5, pi=0.1
plot(dbinom(0:20,5,0.5),type="b")   #n=5, pi=0.5
plot(dbinom(0:20,5,0.9),type="b")   #n=5, pi=0.9

# now complete the rows and columns with the remaining
# values of n and pi.
```

C10. (1) Describe what happens to the shape of the binomial distribution with increasing expected value ($n\pi$).

We can also produce random draws from our common distribution using the "r" version of each function (e.g., rnorm, rpois, rbinom), where "r" is for random. Look at the help files for each function and create histograms of 10000 random draws from the following distributions:

Here is some R code to get you started (repeated from Lab 1). Adapt this code for the remaining distributions.

```
# reset the plotting window to include only one plot
par(mfrow=c(1,1))
# now create the histogram and perform the
# random draws in one line
hist(rnorm(10000,0,1))
```

C11. (2) X~binom(50,0.5) (rbinom) (histogram)

C12. (2) X~Pois(3.2) (rpois) (histogram)