TMATH 390 R Lab 4: Normal and Exponential Distributions

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 normal and exponential distributions.
2. Use R to draw random numbers from the normal and exponential distributions.

C1. (4) **Submit your R scripts to Canvas**. You can upload them directly to your assignment as *.R documents.

Normal distribution

X~N(μ,σ): $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

Z~N(0,1): $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

```r
# The function dnorm returns the height of the normal density
# function at a given value of x, for a normal distribution with
# a given mean and standard deviation (sd)
dnorm(x = 1,mean = 0, sd = 1)
```

C2. (1) For what special normal distribution does the above call return the value of the normal density function?

C3. (1) Report the result of the dnorm command as executed above, and compare it to the value obtained from a hand-calculation of the normal density function with the same mean and standard deviation.

We can also use R to determine the cumulative area under the curve for the normal distribution using the pnorm (“p” for probability) function. This returns the value of the cumulative distribution function (cdf) for the normal distribution. This is especially helpful because we cannot evaluate the integral analytically, we have to solve it numerically. R does this for us.

C4. (1) Create a graph of the cumulative distribution function for the standard normal distribution. Use the code below to help you.

```
# first create a sequence of x values for the graph
# starting at -4, to 4, by 0.1
x.vals=seq(-4,4,0.1)

# now create a vector of the corresping values for the cdf
y.vals=pnorm(x.vals,mean=0,sd=1)

# and make the graph
plot(x.vals,y.vals,xlab="Z",ylab="F(x)",main="cdf for the standard normal")
```

C5. (1) Describe the shape of the graph you created in C4.

C6. (2) For each R command below, report the result (include the command in your report). Compare the result to the probability obtained using the standard normal table.

```
round(pnorm(2.37, mean=0, sd=1), 5)

pnorm(-0.5, mean=0, sd=1)

pnorm(2.37, mean=0, sd=1) - pnorm(-0.5, mean=0, sd=1)

pnorm(18, mean=22, sd=4.5)

pnorm(23, mean=22, sd=4.5)
```

We are often also interested in determining a normal distribution quantile associated with a given probability. We use the r function qnorm ("q" for quantile) to return the normal quantile associated with a given cumulative probability.

C7. (1) For the R command below, report the result (include the command in your report). Compare the result to the quantile obtained using the standard normal table.

```
qnorm(0.25, mean=0, sd=1)
```

C8. (2) Now use R to determine the value that represents the 97th percentile **standard normal distribution**, and the 97th percentile of a normal distribution with a **mean of 5.8 and a standard deviation of 2.5 inches** (distribution of January precipitation from lecture 7 practice problems). Return you R code and the solution.

We can also produce random draws from our common distributions using the "r" version of each function (e.g., rnorm), where "r" is for random. Look at the help file for this function and create a histogram of 10000 random draws from the standard normal distribution:

C9. (1) X~N(0,1). **Copy and paste the histogram into your report.**

Here is some R code to get you started.

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

C10. (1) Now produce a histogram of random draws from the distribution of January precipitation. X~N(5.8,2.5)

Exponential distribution

The exponential distribution is a "waiting time" distribution–it is often used to model processes such as time to failure (the time a component is operating until it fails). The rate parameter λ is the mean number of failures per set time interval. The density function for the exponential distribution is:

$$f(x) = \lambda e^{-\lambda x}, x \geq 0$$

We can use R to evaluate this density function, to calculate probabilities, quantiles, and random numbers associated with this density function. The function dexp returns the density function, pexp cumulative probabilities, qexp quantiles, and rexp random draws.

C11. (1) Use the function pexp to sketch the exponential distribution where λ = 1/8500.

```
# create a sequence of x values starting from 0, ending at
# 10000, by 0.1
x.vals=seq(0,100000,0.1)

# now create a vector of y-values which is the exponential
# density function evaluated at those value of x
y.vals=dexp(x.vals,rate=1/8500)

# now graph:
plot(x.vals,y.vals,xlab="X",ylab="f(x)",main="X~exp(1/8500)")
```

C12. (1) Using R, what is the probability x<1000 where x~exp(1/8500)? Return the solution and your R code in your lab document. Hint: use pexp.

C13. (1) Using R, what is P(1000<x<10000) for this exponential distribution? Return the solution and your R code in your lab document.

C14. (1) Using R, produce a histogram of 10000 random draws of an exponential distribution with a mean of 1/8500 events per hour. Hint: use rexp.