Math 338 **Lab Assignment #3** Fall 2019

Let X be the random variable representing the outcome of rolling a fair six-sided die; that is, X = 1 if you roll a 1, X = 2 if you roll a 2, etc.

Recall that in Lab 2, you found the probability mass function of X to be:

X = x	P(X =x)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

Question #1 Compute the expected value (mean) of X.

The expected value would 1/6(1+2+3+4+5+6) = 3.5

Question #2 Compute the variance and standard deviation of X.

The variance would be 1/6((1−3.5)^2+(2−3.5)^2+(3−3.5)^2+(4−3.5)^2+(5−3.5)^2+(6−3.5)^2)) ~= 2.916666667

The standard deviation would be \sqrt{2.916666667} ~= 1.707825128

Now we are going to simulate the act of rolling a die. In RStudio, open a new Script (*File* ✉ *New File* ✉ *R Script* or *Ctrl + Shift + N*). The code below (in `this font`) will go in your script. It is best to type the lines (except for the introductory `>` or `+`) in the script window and then run them, rather than typing the lines directly into the Console. This will also allow you to save your script as a .R file at the end of the lab.

First, specify your sample space in a vector and create a corresponding vector of the associated probabilities:

```
> S <- c(1, 2, 3, 4, 5, 6)

> Prob <- c(1/6, 1/6, 1/6, 1/6, 1/6, 1/6)
```

Now, let's simulate rolling the die 10 times by randomly sampling a value of the sample space with probability specified in our probability vector:

```
> n_rolls <- 10

> set.seed(338)

> rolls <- sample(x = S, size = n_rolls, prob = Prob, replace = TRUE)

##note that by setting replace equal to TRUE, this replaces the value we
just sampled, thus making each experiment repetition INDEPENDENT.

> num_sixes <- sum(rolls == 6)

> mean(rolls)
```

Once all this code is written, highlight the entire set of code and click *Run* (or hit *Ctrl + Enter* to run it). You should see the code and the resulting output in the *Console* window.

Question #3 What is the relative frequency of the number 6 (the proportion of the time you got 6)? Is it close to the probability you computed earlier?

The relative frequency is 2 and that is not close to the number computed earlier.

Question #4 What is the mean value rolled? Is it close to the mean you computed earlier?

The mean value rolled from the Rscript program came out to be 3.7 which is very close to the calculated 3.5 in question 1.

Set the value of **`n_rolls`** to 100. Highlight that line and the lines after it, then click *Run*. This will now create a simulation of 100 die rolls from the same probability model (because you didn't change **`S`** or **`Prob`**) and obtain the number of 6's and the mean value rolled. Repeat for **`n_rolls`** = 1000, 10000, and 100000.

Question #5 Convert each value of **`num_sixes`** to a relative frequency and fill in table below:

Number of Rolls	Relative Frequency (Proportion of 6's)	Mean
10	2	3.7
100	18	3.65
1,000	169	3.558
10,000	1706	3.5094
100,000	16597	3.49271

Question #6 Describe what happens to the relative frequency of the occurrence of observing a 6 as the number of rolls increases from 10 to 100,000.

The relative frequency of the occurrence of observing a 6 increases exponentially as it approaches infinity.

Question #7 Describe what happens to the mean value of the rolls as the number of rolls increases from 10 to 100,000.

The mean value of the rolls as the number of rolls approaches infinity gets lower and lower.