Econ 184b, Econometrics, Assignment 1

The raw .rmd file for this assignment are available at: https://raw.githubusercontent.com/f1kidd/Econ184/main/Assignment/Assignment_1.Rmd

Note: For the math problems, you can either (1) type your solution and compile it with RMarkdown; (2) type your solution use another word-processing software and convert it into a PDF file; (3) write your solution on paper, scan it and make it into a legible PDF file. TAs can decide, at their discretion, that the submitted file is illegible and thus give zero credit to a question or the entire problem set. If you type your solutions (options 1 and 2 above), you will get **a bonus equal to 5% of your performance** on the problem set. No credit will be given if you only report the final answers without showing intermediate steps whenever appropriate.

For the programming problems, you need to submit both the compiled pdf/html file and the RMarkdown (.rmd) file on LATTE. You will get **an additional bonus equal to 5% of your performance** if your RMarkdown file is completely reproducible with minimal altercation (install packages, etc.). That is, our team (or anyone else) should be able to recompile your Rmarkdown file and reach *the same* result. You need to explicitly write out your answers - just showing programming outputs will receive zero credit.

Question 1:

The following table gives the joint probability distribution between employment status and college graduation among those either employed or looking for work (unemployed) in the working-age U.S. population for September 2017.

	Unemployed(Y=0)	Employed(Y=1)
Non-college grads(X=0)	0.026	0.576
College grads (X=1)	0.009	0.389

a. Compute E(Y)
b. The unemployment rate is the fraction of the labor force that is unemployed. Which one of the following E(X), E(Y), 1-E(X), 1-E(Y), E(X)*E(Y), E(XY), is the unemployment rate? Explain.
c. Calculate E(Y|X=1) and E(Y|X=0)
d. Calculate the unemployment rate for (i) college graduates, and (ii) non-college graduates.
e. A randomly selected number of this population reports being unemployed. What is the probability that this worker is not a college graduate?
f. Are educational achievement and employment status independent? Explain.

Question 2:

The random variable Y has a mean of 1 and a variance of 4. Let Z=(Y-1)/2. Show that E(Z)=0 and var(Z)=1.

Question 3:

Y follows normal distribution with mean=30, variance=100.

a. Compute the theoretical probability of $P(20 \leq Y \leq 50)$. (Hint: the "pnorm" function returns the distribution function of the normal distribution.)
b. Use simulations (10,000 repetitions) to compute $P(30 \leq Y \leq 45)$. (Hint: the rnorm function returns random numbers from the normal distribution. Your code could look like the following:)

```
# draw 10,000 number from the above distribution and store it.
a = 1
a
```

```
## [1] 1
```

```
# count the number of draws that falls in the given range.
```

Question 4:

From the class example, we know that the distribution with possible values of X 7, 77, and 777, and the following probability:

$P(X = 7) = 0.7$
$P(X = 77) = 0.2$
$P(X = 777) = 0.1$

has a mean $\mu = 98$, standard deviation $\sigma = 228.01$. Use 500 repetitions, each with a sample size of 100,

a. Verify the mean and the standard deviation of the above distribution
b. Plot the distribution of means for those 500 repetitions

Your code could look like the following:

```
N = 500
n = 100
# for(i in 1:N){
  # draw one such sample with n=100
  # store the mean of the sample
  # store the standard deviation of the sample
# }
# compute the mean of the sample means
# compute the mean of the sample standard deviations
```

```
# plot your data:
  # the hist() function is the easiest way to do so.
  # we will learn more sophisticated plotting skills later.
```