Homework Assignment #8

CS 3753

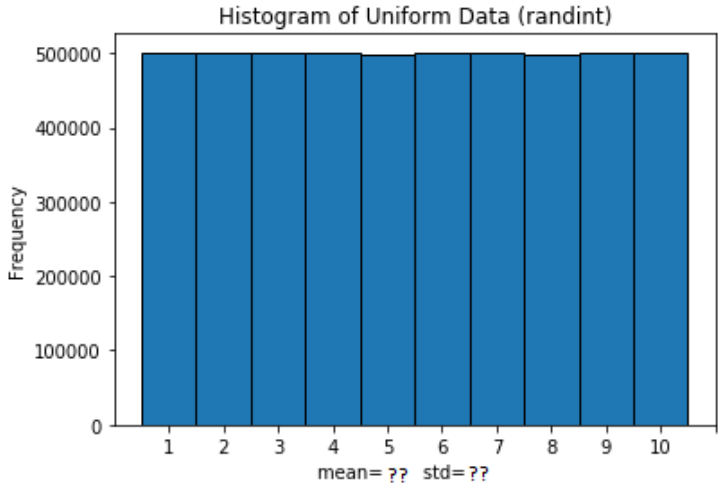
Individual work

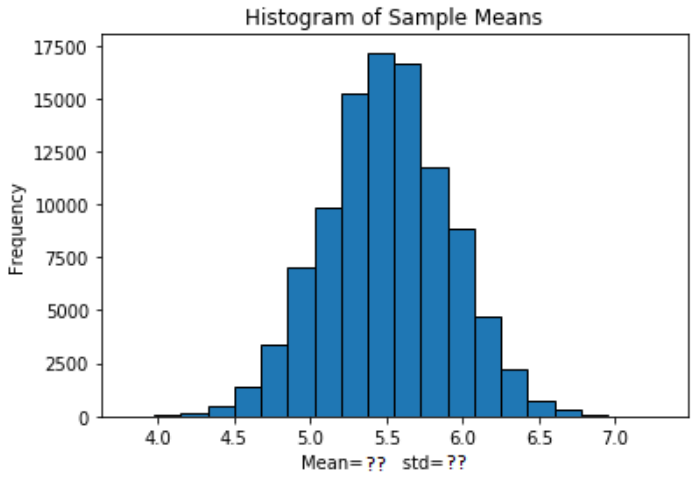
Submit one Jupyter notebook named as yourLastName_HW08 with appropriate code cell to solve the following questions. Make sure that each question is included in your notebook as a markdown cell above your answer. Include any specific direction/instruction to run your script in comments. The plotting questions to be included in your Jupyter notebook.

Note: if you want to include your script as separate .py files, you may submit all source files with a Jupyter notebook in a zipped file (compressed). In the notebook, inside the code cell of a question, write %run Qx. Where x is the number of the question.

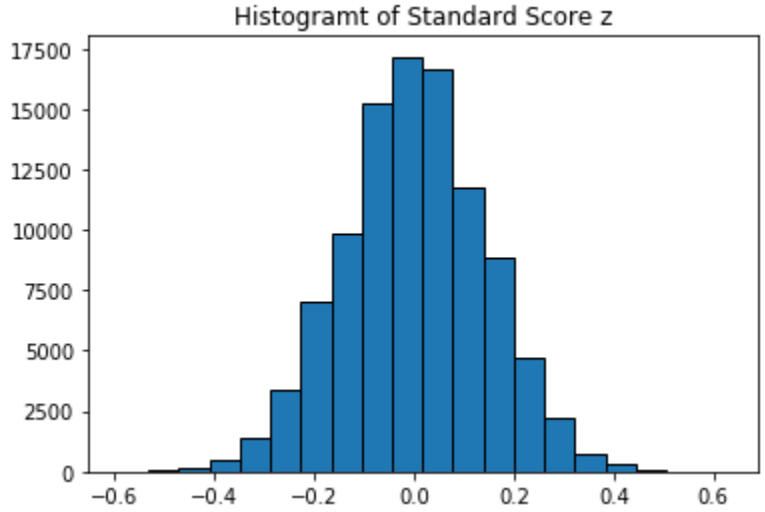
Due date: 05/03/2019 at 11:59PM

Q1) (30 points) In this question, we will verify whether the Central Limit Theorem is correct or not for our data set. Generate 5 million random numbers in [1, 10] using "np.randint". Using np.choice() method, draw 100,000 samples, each of size 50, without replacement. Plot the histogram for your original population and calculate μ and σ. For each sample, calculate $\bar{x}$. You will have 100,000 values. Plot the histogram for these values and calculate the mean and standard deviation for them. Is the mean equal to μ? Is the standard deviation equal to σ/sqrt(50)? In your plots, replace the '??' with your calculations.

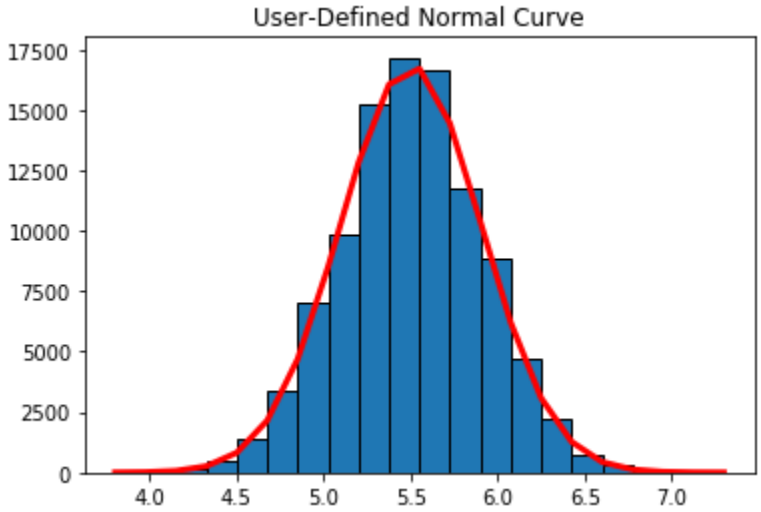




Q2) (15 points) Convert the sample means to standard score (z) and plot it.



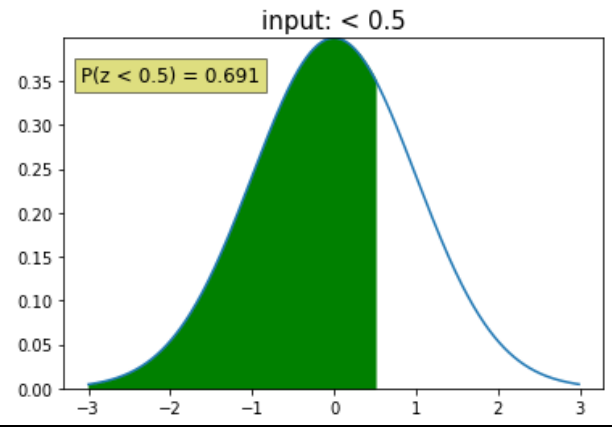
Q3) (15 points) Write your own function to plot the normal distribution curve for sample means from Q1. Don't use any built-in function that can be used to generate/plot the normal curve in red. You may use f(x) below to generate the curve.



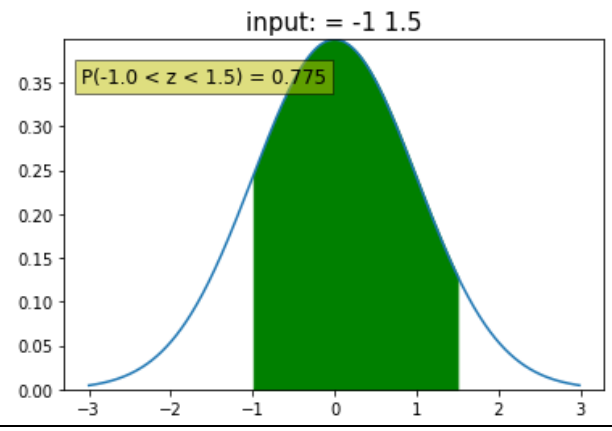
$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}}$$

Q4) (40 points) Create a standard normal distribution (mean=0, Std Deviation=1) curve in python and then shade area to the left, right or the middle of z-score(s) based on the input from the user. The user will provide the input as a sequence of 2 or 3 space-separated string. The first character is either (<, >, or =). The second and third inputs are actually numerical. The only case the program accepts 3 inputs is when the first character is '='. The first character defines which area to shade under the curve. If < is used, then your program will shade the area left to the provided number after <. If > is used, your program will shade the area right to the provided number. If = is used, your program expects two numbers to shade the area in-between. The numbers are given in increasing value. Thus, first number is used as the lower limit of the area to shade and the second number is used as an upper limit. Calculate the probability (P) for the shaded area and add it as a text into the plot. To shade the area, you may use plt.fill_between() method. This method fills the area between two given horizontal curves. To draw the curve for the standard normal distribution, use the user-defined function you wrote in Q3. Where x is a np.arange(-3,3,0.01). Below find some examples.

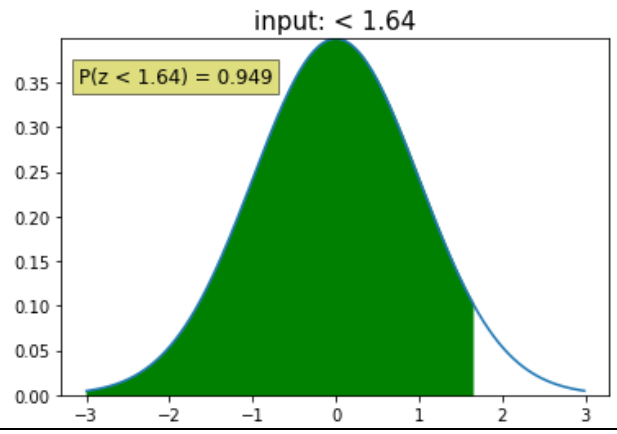
```
Enter your input: < 0.5
```



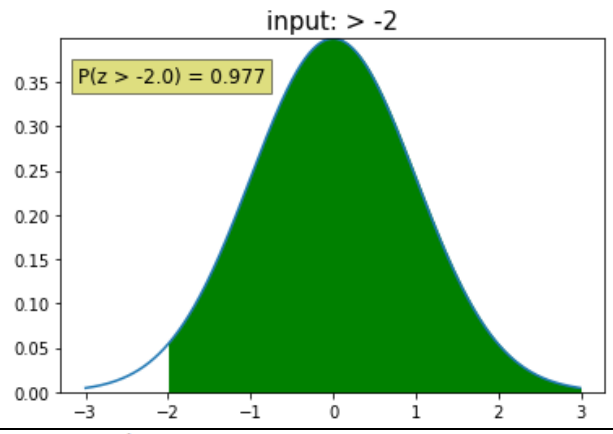

```
Enter your input: = -1 1.5
```



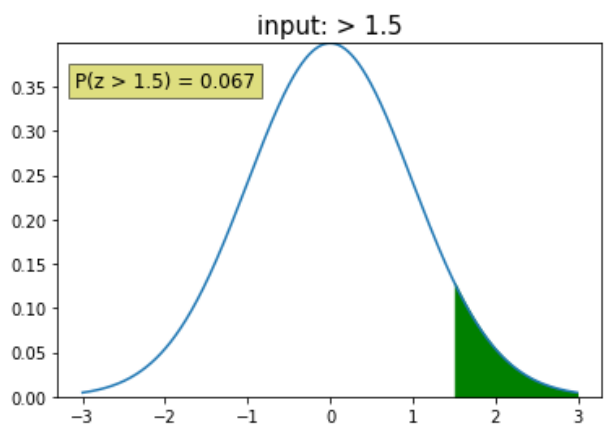

```
Enter your input: < 1.64
```




```
Enter your input: > -2
```




```
Enter your input: > 1.5
```




```
Enter your input: < -1
```