Module 3 Live Coding Assignment: Fundamentals of Statistics

Throughout this assignment, please remember to import all the necessary libraries at the beginning of every exercise.

Question 1

Simulate the coin toss experiment using Python for the following cases:

1. Suppose you have a fair coin and you toss it once, write a code that simulates this experiment. Print the result to screen. **Hint:** You can code 1 for head and 0 for tail.
2. Set a random seed equal to 42. Write a code that simulates the tossing of a fair coin 17 times. Print the result to screen.
3. Set a random seed equal to 42. Repeat the above experiment 70 times and display how many times you get head in each experiment. Save the result is a vector "x" and print it to screen. Plot your results.

Question 2

In the experiment in Question 1.3, we tossed a fair coin 17 times. Then, we repeated the experiment 100 times and measured how many successes/heads we observed. Write a code to count how many times you see 0 heads, 1 head, ..., 17 heads with our fair coin toss and save the result to "head_count". Print "head_count" to screen. Use the observed successes to estimate the probability of getting "x3" successes in n=17 coin tosses and save your result to "probs70". Finally, plot your result.

Question 3

Repeat the experiment in Question 1.3 by tossing the coin 7000 times. Use the observed successes to estimate the probability of getting heads in n=17 coin tosses. Use the variables "x5", "head_count2" and "probs7000" instead of "x3", "head_count" and "probs70", respectively.

Question 4

Simulate the roll the dice experiment using Python for the following cases:

1. Suppose you have a fair dice and you roll it once. Save your result to "roll_once" and print it to screen. **Hint:** Use the randint function in the random library.

2. Suppose you have a fair dice and you roll it 20 times. Save your result to "roll20" and print it to screen. Finally, plot the distribution of 20 dice rolls using an histogram.

Question 5

This question is about the T-test. Set the seed equal to 14. Generate two random sets of data called "data1" and "data2" with 100 data points from N(3, 50) and N(3, 51), respectively. Compute the calculated t-statistic and two-tailed p-value and assign them to "t_stat" and "p_val". respectively. Print to screen your results.

Question 6

This question is about Q-Q-plots. Diplay the Q-Q plots for the following cases:

1. A sample of 4000 normal random data points.

2. A sample of 4000 exponential random data points.