CS 322 Assignment 2 Probability, Coding and Simulation

There are three math problems (1, 2, and 3) and one coding problem (4). For the first three problems, you must solve them by hand, and your handwriting should be clear and legible. Take a picture of each solution and convert it to either a PDF or JPEG file. If you prefer, you may type the formulas and equations in Word or LaTeX instead. For the coding problem, you must submit an actual executable file as usual.

Academic honesty: You must complete the work on your own. You may discuss questions with classmates or visit office hours. However, if you and a classmate submit identical solutions, it will be treated as cheating and will receive zero credit—unless you can prove the similarity is an extraordinary coincidence. For more complex assignments or projects, collaboration may be allowed and even encouraged.

Folder submission: Place all files in one folder. Name the folder exactly as your name appears in NEIUPORT. For example, if your name is Jessica Fatima, submit the folder as Jessica_Fatima.

Due date: Thursday (10/2).

Grace period: You may submit up to three days late without a penalty of 20%.

Late policy: Submissions more than one week past the due date will not be accepted or graded.

Problem 1 (20 points)
Let X be a discrete random variable with the following PMF:

$$P_X(x) = \begin{cases} 0.1 & \text{for } x = 0.2 \\ 0.2 & \text{for } x = 0.4 \\ 0.2 & \text{for } x = 0.5 \\ 0.3 & \text{for } x = 0.8 \\ 0.2 & \text{for } x = 1 \\ 0 & \text{otherwise} \end{cases}$$

a. Find R_X, the range of the random variable X.

b. Find $P(X < 0.5)$.

c. Find $P(0.25 < X < 0.75)$.

d. Find $P(X = 0.2 \mid X < 0.6)$.

Problem 2 (25 points)

I roll two dice and observe two numbers X and Y.

a. Find R_X, R_Y and the PMFs of X and Y.

b. Find P(X = 2, Y = 6).

c. Find P(X > 3 | Y = 2).

d. Let Z = X + Y. Find the range and PMF of Z.

e. Find P(X = 4 | Z = 8).

Problem 3 (20 points)

We roll a fair die until the first six appears.
Let X be the number of rolls **before** the first six (so $X \geq 0$).

Define the event E: *all of the rolls before the six are even numbers (2 or 4)*.

Question: Find $P(X = x \mid E)$.

Hint: This one is a bit harder, but here is an important clue: use the Bayesian Proposition

$$P(X = x \mid E) = \frac{P(E \mid X = x)\, P(X = x)}{\sum_{k=0}^{\infty} P(E \mid X = k)\, P(X = k)}$$

Problem 4 (35 points)

We have spent quite some time working probability out by hand. Let us try something different: simulation.

A random walk is like taking steps on a number line. You start at 0. At each step, you flip a coin:

- heads → move right (+1)
- tails → move left (–1)

After many steps (say x), your final position could be far to the left, far to the right, or near the middle.

By definition, a random variable is a real-valued function on the sample space. The final position after x steps is therefore a random variable. How do you figure out its probability mass function? That can be a challenging question. Later on, we will learn some theory to work it out. For now, let's use simulation to explore it.

Goal: Use simulation to see that the distribution of the final position after X steps in a simple symmetric random walk is approximately symmetric about 0 when visualized with a histogram. Later on, you will see that it will become a normal distribution if the number of steps is infinite.

Hint: To simulate a single step forward or backward when flipping a coin, use random.choice([-1, 1]) from Python's random library. Think of it as a drunk person taking a walk of, say, 1000 steps and ending up somewhere. Repeat this walk 10,000 times. Each outcome seems random, but the collection of outcomes reveals a pattern when plotted as a histogram. The histogram is an intuitive way to approximate the probability mass function.

Implementation note: You will need at least two loops—one loop to simulate a single walk which is inside a function called 'one_walk', and another loop to repeat the random walk many times which is inside a function called 'many_walks'.