COMP 2804 — Assignment 4

Due: Wednesday December 7, before 4:30pm, in the course drop box in Herzberg 3115.

Assignment Policy: Late assignments will not be accepted. Students are encouraged to collaborate on assignments, but at the level of discussion only. When writing the solutions, they should do so in their own words. Past experience has shown conclusively that those who do not put adequate effort into the assignments do not learn the material and have a probability near 1 of doing poorly on the exams.

Important note: When writing your solutions, you must follow the guidelines below.

- The answers should be concise, clear and neat.
- When presenting proofs, every step should be justified.
- Assignments should be stapled or placed in an unsealed envelope.

Substantial departures from the above guidelines will not be graded.

Question 1: On the first page of your assignment, write your name and student number.

Question 2: As of this writing[1], Ma Long is the number 1 ranked ping pong player in the world. Simon Bose[2] also plays ping pong, but he is not at Ma's level yet. If you play a game of ping pong against Ma, then you win with probability p. If you play a game against Simon, you win with probability q. Here, p and q are real numbers such that $0 < p < q < 1$. (Of course, p is much smaller than q.) If you play several games against Ma and Simon, then the results are mutually independent.

You have the choice between the following two series of games:

1. *MSM*: First, play against Ma, then against Simon, then against Ma.
2. *SMS*: First, play against Simon, then against Ma, then against Simon.

For each $s \in \{MSM, SMS\}$, define the event

$$A_s = \text{"you play series } s \text{ and beat Ma at least once and beat Simon at least once"}$$

and the random variable

$$X_s = \text{the number of games you win when playing series } s.$$

- Determine $\Pr(A_{MSM})$ and $\Pr(A_{SMS})$. Which of these two probabilities is larger?

  Before you answer this question, spend a few seconds on guessing which one is larger.

[1] November 2016

[2] Jit's son

- Determine $\mathbb{E}(X_{MSM})$ and $\mathbb{E}(X_{SMS})$. Which of these two expected values is larger?

  Before you answer this question, spend a few seconds on guessing which one is larger.

Question 3: Consider the 8-element set $A = \{a, b, c, d, e, f, g, h\}$. We choose a 5-element subset B of A uniformly at random. Define the following random variables:

$$\begin{aligned} X &= |B \cap \{a, b, c, d\}|, \\ Y &= |B \cap \{e, f, g, h\}|. \end{aligned}$$

- Determine the expected value $\mathbb{E}(X)$ of the random variable X. Show your work.
- Are X and Y independent random variables? Justify your answer.

Question 4: You roll a fair die repeatedly and independently until the result is an even number. Define the random variables

$$X = \text{ the number of times you roll the die}$$

and

$$Y = \text{ the result of the last roll.}$$

For example, if the results of the rolls are $5, 1, 3, 3, 5, 2$, then $X = 6$ and $Y = 2$.

Prove that the random variables X and Y are independent.

Question 5: Elisa Kazan is having a party at her home. Elisa has a round table that has 52 seats numbered $0, 1, 2, \ldots, 51$ in clockwise order. Elisa invites 51 friends, so that the total number of people at the party is 52. Of these 52 people, 15 drink cider, whereas the other 37 drink beer.

In this exercise, you will prove the following claim: No matter how the 52 people sit at the table, there is always a consecutive group of 7 people such that at least 3 of them drink cider.

Note that this claim does not have anything to do with probability. In the rest of this exercise, you will use random variables to prove this claim.

From now on, we consider an arbitrary (which is not random) arrangement of the 52 people sitting at the table.

- Let k be a uniformly random element of the set $\{0, 1, 2, \ldots, 51\}$. Consider the consecutive group of 7 people that sit in seats $k, k+1, k+2, \ldots, k+6$; these seat numbers are to be read modulo 52. Define the random variable X to be the number of people in this group that drink cider. Prove that $\mathbb{E}(X) > 2$.

  Hint: Number the 15 cider drinkers arbitrarily as $P_1, P_2, \ldots, P_{15}$. For each i with $1 \leq i \leq 15$, define the indicator random variable

$$X_i = \begin{cases} 1 & \text{if } P_i \text{ sits in one of the seats } k, k+1, k+2, \ldots, k+6, \\ 0 & \text{otherwise.} \end{cases}$$

- For the given arrangement of the 52 people sitting at the table, prove that there is a consecutive group of 7 people such that at least 3 of them drink cider.

  Hint: Assume the claim is false. What is an upper bound on $\mathbb{E}(X)$?

Question 6: Let $n \geq 2$ be an integer. Consider the following random process that divides the integers $1, 2, \ldots, n$ into two sorted lists L_1 and L_2:

1. Initialize both L_1 and L_2 to be empty.
2. For each $i = 1, 2, \ldots, n$, flip a fair coin. If the coin comes up heads, then add i at the end of list L_1. Otherwise, add i at the end of the list L_2. (All coin flips made during this process are mutually independent.)

We now run algorithm $\textsc{Merge}(L_1, L_2)$ of Section 4.5 in the textbook. Define the random variable X to be the total number of comparisons made when running this algorithm: As in Section 4.5.2, X counts the number of times the line "**if** $x \leq y$" in algorithm $\textsc{Merge}(L_1, L_2)$ is executed. In this exercise, you will determine the expected value $\mathbb{E}(X)$ of the random variable X.

- Prove that $\mathbb{E}(X) = 1/2$ for the case when $n = 2$.
- Prove that $\mathbb{E}(X) = 5/4$ for the case when $n = 3$.
- Assume that $n \geq 2$. For each i and j with $1 \leq i < j \leq n$, define the indicator random variable

$$X_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are compared,} \\ 0 & \text{otherwise.} \end{cases}$$

  Prove that $\mathbb{E}(X_{ij}) = (1/2)^{j-i}$.

  Hint: Assume that i and j are compared. Can i and j be in the same list? What about the elements $i, i+1, \ldots, j-1$ and the element j?

- Determine $\mathbb{E}(X)$.

  Hint: $1 + x + x^2 + x^3 + \cdots + x^k = \frac{1-x^{k+1}}{1-x}$.

Question 7: By flipping a fair coin repeatedly and independently, we obtain a sequence of H's and T's. We stop flipping the coin as soon as the sequence contains either HH or TT. Define the random variable X to be the number of times that we flip the coin. For example, if the sequence of coin flips is $HTHTT$, then $X = 5$.

- Let $k \geq 2$ be an integer. Determine $\Pr(X = k)$.

- Determine the expected value $\mathbb{E}(X)$ of X using the formula

$$\mathbb{E}(X) = \sum_k k \cdot \Pr(X = k).$$

  Hint: In class, we have seen that for $0 < x < 1$, $\sum_{k=1}^{\infty} kx^k = \frac{x}{(1-x)^2}$.

- Determine $\Pr(X \geq 1)$.
- Let $k \geq 2$ be an integer. Determine $\Pr(X \geq k)$.
- According to Exercise 6.11 in the textbook, we have

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} \Pr(X \geq k).$$

  Use this formula to determine the expected value $\mathbb{E}(X)$ of X.