COMP 2804 — Assignment 4

Due: Thursday November 29, before 11:55pm, through cuLearn.

Assignment Policy:

- Your assignment must be submitted as one single PDF file through cuLearn.
- **Late assignments will not be accepted. I will not reply to emails of the type "my internet connection broke down at 11:53pm" or "my scanner stopped working at 11:54pm".**
- You are encouraged to collaborate on assignments, but at the level of discussion only. When writing your solutions, you must do so in your 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.
- When writing your solutions, you must follow the guidelines below.
  - You must justify your answers.
  - The answers should be concise, clear and neat.
  - When presenting proofs, every step should be justified.

Question 1: Write your name and student number.

Question 2: Alexa[1], Tri[2], and Zoltan[3] play the OddPlayer game: In one round, each player flips a fair coin.

1. Assume that not all flips are equal. Then the coin flips of exactly two players are equal. The player whose coin flip is different is called the *odd player*. In this case, the odd player wins the game. For example, if Alexa flips tails, Tri flips heads, and Zoltan flips tails, then Tri is the odd player and wins the game.
2. If all three coin flips are equal, then the game is repeated.

Below, this game is presented in pseudocode:

[1] your friendly TA
[2] another friendly TA
[3] yet another friendly TA

Algorithm ODDPLAYER**:**

// all coin flips made are mutually independent
each player flips a fair coin;
if not all coin flips are equal
then the game terminates and the odd player wins
else ODDPLAYER
endif

- What is the sample space?
- Define the event

$$A = \text{“Alexa wins the game”}.$$

  Express this event as a subset of the sample space.
- Use your expression from the previous part to determine $\Pr(A)$.
- Use symmetry to determine $\Pr(A)$. Explain your answer in plain English and a few sentences.

  Hint: What is the probability that Tri wins the game? What is the probability that Zoltan wins the game?

Question 3: Consider the set $S = \{2, 3, 5, 30\}$. We choose a uniformly random element x from this set. Define the random variables

$$X = \begin{cases} 1 & \text{if } x \text{ is divisible by 2,} \\ 0 & \text{otherwise,} \end{cases}$$
$$Y = \begin{cases} 1 & \text{if } x \text{ is divisible by 3,} \\ 0 & \text{otherwise,} \end{cases}$$
$$Z = \begin{cases} 1 & \text{if } x \text{ is divisible by 5,} \\ 0 & \text{otherwise.} \end{cases}$$

- Is the sequence X, Y, Z of random variables pairwise independent? As always, justify your answer.
- Is the sequence X, Y, Z of random variables mutually independent? As always, justify your answer.

Question 4: Let a and b be real numbers. You flip a fair and independent coin three times. For $i = 1, 2, 3$, let

$$f_i = \begin{cases} a & \text{if the } i\text{-th coin flip results in heads,} \\ b & \text{if the } i\text{-th coin flip results in tails.} \end{cases}$$

Define the random variables

$$\begin{aligned} X &= f_1 \cdot f_2, \\ Y &= f_2 \cdot f_3. \end{aligned}$$

For each of the following questions, justify your answer.

- Assume that $a = b$. Are the random variables X and Y independent?
- Assume that $a = 0$ and $b \neq a$. Are the random variables X and Y independent?
- Assume that $a \neq 0$ and $b = -a$. Are the random variables X and Y independent?
- Assume that $a \neq 0$, $b \neq 0$, $a \neq b$, and $b \neq -a$. Are the random variables X and Y independent?

Question 5: You are given three fair dice. One die is red, one die is blue, and one die is green. You roll each die once, independently of the other two dice. Define the random variables

$$\begin{aligned} X_r &= \text{the value of the red die}, \\ X_b &= \text{the value of the blue die}, \\ X_g &= \text{the value of the green die}, \\ Y &= \max(X_r, X_b, X_g), \end{aligned}$$

and, for each integer k with $1 \leq k \leq 6$, the events

$$\begin{aligned} A_k &= \text{“}Y = k\text{”}, \\ B_k &= \text{“}Y \leq k\text{”}. \end{aligned}$$

- Determine the expected values $\mathbb{E}(X_r)$, $\mathbb{E}(X_b)$, and $\mathbb{E}(X_g)$.
- Let k be an integer with $1 \leq k \leq 6$. Determine $\Pr(B_k)$.
- Determine $\Pr(A_1)$.
- Let k be an integer with $2 \leq k \leq 6$. Express the event B_k in terms of the events A_k and B_{k-1}.
- Let k be an integer with $2 \leq k \leq 6$. Determine $\Pr(A_k)$.
- Determine the expected value $\mathbb{E}(Y)$ of the random variable Y.
- Is the following true or false?

$$\mathbb{E}(\max(X_r, X_b, X_g)) = \max(\mathbb{E}(X_r), \mathbb{E}(X_b), \mathbb{E}(X_g)).$$

Question 6: Let $n \geq 1$ be an integer and let $A[1 \dots n]$ be an array that stores a permutation of the set $\{1, 2, \dots, n\}$. If the array A is sorted, then $A[k] = k$ for $k = 1, 2, \dots, n$ and, thus,

$$\sum_{k=1}^{n} |A[k] - k| = 0. \tag{1}$$

If the array A is not sorted and $A[k] = i$, where $i \neq k$, then $|A[k]-k|$ is equal to the "distance" that the value i must move in order to make the array sorted. Thus, the summation in (1) is a measure for the "sortedness" of the array A: If the summation is small, then A is "close" to being sorted. On the other hand, if the summation is large, then A is "far away" from being sorted. In this exercise, you will determine the expected value of the summation in (1).

Assume that the array stores a uniformly random permutation of the set $\{1, 2, \dots, n\}$. For each $k = 1, 2, \dots, n$, define the random variable

$$X_k = |A[k] - k|,$$

and let

$$X = \sum_{k=1}^{n} X_k.$$

- Assume that $n = 1$. Determine the expected value $\mathbb{E}(X)$.
- Assume that $n \geq 2$. Is the sequence $X_1, X_2, \dots, X_n$ of random variables pairwise independent?
- Assume that $n \geq 1$. Let k be an integer with $1 \leq k \leq n$. Prove that

  $$\mathbb{E}(X_k) = \frac{n+1}{2} + \frac{k^2 - k - kn}{n}.$$

  Hint: Assume $A[k] = i$. If $1 \leq i \leq k$, then $|A[k] - k| = k - i$. If $k + 1 \leq i \leq n$, then $|A[k] - k| = i - k$. For any integer $m \geq 1$,

  $$1 + 2 + 3 + \cdots + m = \frac{m(m+1)}{2}.$$

- Assume that $n \geq 1$. Prove that

  $$\mathbb{E}(X) = \frac{n^2 - 1}{3}.$$

  Hint:

  $$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}.$$

Question 7: Let $b \geq 1$, $c \geq 1$, and $w \geq 1$ be integers, and let $n = b + c + w$. You are given b beer bottles $B_1, B_2, \ldots, B_b$, c cider bottles $C_1, C_2, \ldots, C_c$, and w wine bottles $W_1, W_2, \ldots, W_w$. Let $m \geq 1$ be an integer with $m \leq b$ and $m \leq n - b$.

All n bottles are in a box. From this box, you choose a uniformly random subset consisting of m bottles. Define the random variables

$$\begin{aligned} X &= \text{the number of beer bottles in the chosen subset,} \\ Y &= \text{the number of cider bottles in the chosen subset,} \\ Z &= \text{the number of wine bottles in the chosen subset.} \end{aligned}$$

- Determine the expected value $\mathbb{E}(X + Y + Z)$.
- Let k be an integer with $0 \leq k \leq m$. Prove that

$$\Pr(X = k) = \frac{\binom{b}{k}\binom{n-b}{m-k}}{\binom{n}{m}}.$$

- For each $i = 1, 2, \ldots, b$ and $j = 1, 2, \ldots, c$, define the indicator random variables

$$X_i = \begin{cases} 1 & \text{if } B_i \text{ is in the chosen subset,} \\ 0 & \text{otherwise.} \end{cases}$$

  and

$$Y_j = \begin{cases} 1 & \text{if } C_j \text{ is in the chosen subset,} \\ 0 & \text{otherwise.} \end{cases}$$

  Prove that

$$\mathbb{E}(X_i) = \mathbb{E}(Y_j) = \frac{m}{n}.$$

- Prove that

$$\sum_{k=0}^{m} \frac{k\binom{b}{k}\binom{n-b}{m-k}}{\binom{n}{m}} = \frac{bm}{n}.$$

- Let i and j be integers with $1 \leq i \leq b$ and $1 \leq j \leq c$. Are the random variables X_i and Y_j independent?
- Let i and j be integers with $1 \leq i \leq b$ and $1 \leq j \leq c$. Determine $\mathbb{E}(X_i \cdot Y_j)$.
- Let i and j be integers with $1 \leq i \leq b$ and $1 \leq j \leq c$. Is the following true or false?

$$\mathbb{E}(X_i \cdot Y_j) = \mathbb{E}(X_i) \cdot \mathbb{E}(Y_j).$$

Question 8: As you all know, Elisa Kazan is the President of the Carleton Computer Science Society. Elisa's neighborhood pub serves three types of drinks: cider, wine, and beer. Elisa likes cider[4] and wine[5], but does not like beer[6].

After a week of hard work, Elisa goes to this pub and repeatedly orders a random drink (the results of the orders are mutually independent). If she gets a glass of cider or a glass of wine, then she drinks it and places another order. As soon as she gets a pint of beer, she drinks it and takes a taxi home.

When Elisa orders one drink, she gets a glass of cider with probability 2/5, a glass of wine with probability 2/5, and a pint of beer with probability 1/5.

Define the random variables

$$\begin{aligned} X &= \text{the number of drinks that Elisa orders,} \\ Y &= \text{the number of different types that Elisa drinks.} \end{aligned}$$

If we denote cider by C, wine by W, and beer by B, then a possible sequence of drinks is $CCWCB$; for this case $X = 5$ and $Y = 3$. For the sequence $WWWB$, we have $X = 4$ and $Y = 2$.

- Determine the expected value $\mathbb{E}(X)$.
- Describe the sample space in terms of strings consisting of characters C, W, and B.
- Describe the event "$Y = 1$" in terms of a subset of the sample space.
- Use the result of the previous part to determine $\Pr(Y = 1)$.
- Describe the event "$Y = 2$" in terms of a subset of the sample space.
- Use the result of the previous part to determine $\Pr(Y = 2)$.
- Determine $\Pr(Y = 3)$.
- Use the results of the previous five parts to determine the expected value $\mathbb{E}(Y)$.
- Define the random variable

  $$Y_c = \begin{cases} 1 & \text{if Elisa drinks at least one glass of cider,} \\ 0 & \text{otherwise.} \end{cases}$$

  Determine the expected value $\mathbb{E}(Y_c)$.
- Define the random variable

  $$Y_w = \begin{cases} 1 & \text{if Elisa drinks at least one glass of wine,} \\ 0 & \text{otherwise.} \end{cases}$$

  Determine the expected value $\mathbb{E}(Y_w)$.

---

[4] I know this is true

[5] I am not sure if this is true

[6] I know this is true

- Express Y in terms of Y_c and Y_w.
- Use the results of the previous three parts to determine the expected value $\mathbb{E}(Y)$.

Question 9: Let $n \geq 2$ be an integer and consider n people $P_1, P_2, \ldots, P_n$. Each of these people has a uniformly random birthday, and all birthdays are mutually independent. (We ignore leap years.) Define the random variable

$$X = \text{ the number of indices } i \text{ such that } P_i \text{ and } P_{i+1} \text{ have the same birthday.}$$

Determine the expected value $\mathbb{E}(X)$.
Hint: Use indicator random variables.