

COMP 2804 — Solutions Assignment 4

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

Solution:

- Name: Lionel Messi
- Student number: 10

Question 2: As of this writing¹, Ma Long is the number 1 ranked ping pong player in the world. Simon Bose² 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 = “you play series s and beat Ma at least once and beat Simon at least once”

and the random variable

X_s = 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.
- 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.

Solution: I am going to write W_M if you win a game against Ma and L_M if you lose against Ma. Similarly, I will write W_S and L_S .

The event A_{MSM} is equivalent to the following results:

$W_M W_S W_M$ or $W_M W_S L_M$ or $L_M W_S W_M$.

¹November 2016

²Jit's son

Since these three possible results are pairwise disjoint, and using that the results are independent of each other, we get

$$\begin{aligned}
\Pr(A_{MSM}) &= \Pr(W_M W_S W_M \text{ or } W_M W_S L_M \text{ or } L_M W_S W_M) \\
&= \Pr(W_M W_S W_M) + \Pr(W_M W_S L_M) + \Pr(L_M W_S W_M) \\
&= pqp + pq(1-p) + (1-p)qp \\
&= pq(2-p).
\end{aligned}$$

By a symmetric argument, we get

$$\Pr(A_{SMS}) = pq(2-q).$$

Since $p < q$, we have $2-q < 2-p$, implying that

$$\Pr(A_{SMS}) = pq(2-q) < pq(2-p) = \Pr(A_{MSM}).$$

At first sight, this is surprising, because it shows that the series MSM is a better choice for you than SMS . Think about it this way: Your goal is to beat both Ma and Simon at least once. It is easy to beat Simon and it is very hard to beat Ma. In the series MSM , you have two chances of beating Ma, whereas in the series SMS , you have only one chance to beat him. Having two chances to win at least once is better than having only one chance.

Next we determine $\mathbb{E}(X_{MSM})$: The possible values for this random variable are 0, 1, 2, 3.

1. $X_{MSM} = 0$: This is equivalent to $L_M L_S L_M$, which happens with probability

$$(1-p)(1-q)(1-p) = (1-p)^2(1-q).$$

2. $X_{MSM} = 1$: This is equivalent to

$$W_M L_S L_M \text{ or } L_M W_S L_M \text{ or } L_M L_S W_M,$$

which happens with probability

$$p(1-q)(1-p) + (1-p)q(1-p) + (1-p)(1-q)p = 2p(1-p)(1-q) + (1-p)^2q.$$

3. $X_{MSM} = 2$: This is equivalent to

$$W_M W_S L_M \text{ or } W_M L_S W_M \text{ or } L_M W_S W_M,$$

which happens with probability

$$pq(1-p) + p(1-q)p + (1-p)qp = 2p(1-p)q + p^2(1-q).$$

4. $X_{MSM} = 3$: This is equivalent to $W_M W_S W_M$, which happens with probability

$$ppq = p^2q.$$

This gives us

$$\begin{aligned}
\mathbb{E}(X_{MSM}) &= 0 \cdot \Pr(X_{MSM} = 0) + 1 \cdot \Pr(X_{MSM} = 1) + 2 \cdot \Pr(X_{MSM} = 2) + 3 \cdot \Pr(X_{MSM} = 3) \\
&= 1 \cdot (2p(1-p)(1-q) + (1-p)^2q) + 2 \cdot (2p(1-p)q + p^2(1-q)) + 3 \cdot (p^2q) \\
&= \text{(do a boring calculation)} \\
&= 2p + q.
\end{aligned}$$

By a symmetric argument, we get

$$\mathbb{E}(X_{SMS}) = 2q + p.$$

Here is an easier way to determine $\mathbb{E}(X_{MSM})$: Define indicator random variables X_1, X_2, X_3 , where

$$X_i = \begin{cases} 1 & \text{if you win the } i\text{-th game,} \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\begin{aligned}
\mathbb{E}(X_1) &= \Pr(X_1 = 1) = \Pr(W_M) = p, \\
\mathbb{E}(X_2) &= \Pr(X_2 = 1) = \Pr(W_S) = q,
\end{aligned}$$

and

$$\mathbb{E}(X_3) = \Pr(X_3 = 1) = \Pr(W_M) = p.$$

Since $X_{MSM} = X_1 + X_2 + X_3$, we get

$$\begin{aligned}
\mathbb{E}(X_{MSM}) &= \mathbb{E}(X_1 + X_2 + X_3) \\
&= \mathbb{E}(X_1) + \mathbb{E}(X_2) + \mathbb{E}(X_3) \\
&= p + q + p \\
&= 2p + q.
\end{aligned}$$

Which of the two expected values is larger: Since $p < q$, we have

$$\mathbb{E}(X_{MSM}) = 2p + q = p + p + q < p + q + q = 2q + p = \mathbb{E}(X_{SMS}).$$

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.

Solution: The possible values for the random variable X are 1, 2, 3, 4.

1. $X = 1$: This happens if exactly one element of B belongs to $\{a, b, c, d\}$ and the other four elements of B belong to $\{e, f, g, h\}$. The probability of this happening is equal to

$$\frac{\binom{4}{1} \binom{4}{4}}{\binom{8}{5}}.$$

2. $X = 2$: This happens if exactly two elements of B belong to $\{a, b, c, d\}$ and the other three elements of B belong to $\{e, f, g, h\}$. The probability of this happening is equal to

$$\frac{\binom{4}{2} \binom{4}{3}}{\binom{8}{5}}.$$

3. $X = 3$: This happens if exactly three elements of B belong to $\{a, b, c, d\}$ and the other two elements of B belong to $\{e, f, g, h\}$. The probability of this happening is equal to

$$\frac{\binom{4}{3} \binom{4}{2}}{\binom{8}{5}}.$$

4. $X = 4$: This happens if exactly four elements of B belong to $\{a, b, c, d\}$ and the other element of B belongs to $\{e, f, g, h\}$. The probability of this happening is equal to

$$\frac{\binom{4}{4} \binom{4}{1}}{\binom{8}{5}}.$$

This gives us

$$\begin{aligned} \mathbb{E}(X) &= 1 \cdot \Pr(X = 1) + 2 \cdot \Pr(X = 2) + 3 \cdot \Pr(X = 3) + 4 \cdot \Pr(X = 4) \\ &= \text{(plug in the values and do a boring calculation)} \\ &= 5/2. \end{aligned}$$

Here is an easier way to determine $\mathbb{E}(X)$: Define indicator random variables X_a, X_b, X_c, X_d , where

$$X_a = \begin{cases} 1 & \text{if } a \in B, \\ 0 & \text{otherwise.} \end{cases}$$

The variables X_b, X_c , and X_d are defined similarly. Then

$$\mathbb{E}(X_a) = \Pr(a \in B) = \frac{\binom{7}{4}}{\binom{8}{5}} = 5/8.$$

By the same arguments,

$$\mathbb{E}(X_b) = \mathbb{E}(X_c) = \mathbb{E}(X_d) = 5/8.$$

Since $X = X_a + X_b + X_c + X_d$, we get

$$\begin{aligned}\mathbb{E}(X) &= \mathbb{E}(X_a + X_b + X_c + X_d) \\ &= \mathbb{E}(X_a) + \mathbb{E}(X_b) + \mathbb{E}(X_c) + \mathbb{E}(X_d) \\ &= 5/8 + 5/8 + 5/8 + 5/8 \\ &= 5/2.\end{aligned}$$

Are the random variables X and Y independent? Note that $X + Y = 5$, no matter which subset B is chosen. In other words, if we know the value of X , then we also know the value of Y . For example, X and Y cannot both be equal to 1:

$$\Pr(X = 1 \text{ and } Y = 1) = 0,$$

whereas

$$\Pr(X = 1) \cdot \Pr(Y = 1) = \frac{4}{\binom{8}{5}} \cdot \frac{4}{\binom{8}{5}} \neq 0.$$

It follows that the random variables X and Y are not independent.

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

$X =$ the number of times you roll the die

and

$Y =$ 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.

Solution: The possible values for X are $1, 2, 3, 4, \dots$ and the possible values for Y are $2, 4, 6$.

Let $k \geq 1$ be an integer. To prove that X and Y are independent, we have to show that

$$\Pr(X = k \wedge Y = 2) = \Pr(X = k) \cdot \Pr(Y = 2),$$

$$\Pr(X = k \wedge Y = 4) = \Pr(X = k) \cdot \Pr(Y = 4),$$

and

$$\Pr(X = k \wedge Y = 6) = \Pr(X = k) \cdot \Pr(Y = 6).$$

I will show this for $Y = 4$; the other two equalities can be handled in exactly the same way.

1. What is $\Pr(Y = 4)$? The last roll results in one of 2, 4, and 6. Therefore,

$$\Pr(Y = 4) = 1/3.$$

2. What is $\Pr(X = k)$? This happens if the rolls give the sequence

$$\underbrace{\text{odd, odd}, \dots, \text{odd}}_{k-1 \text{ times}}, \text{even}.$$

Since the result is odd with probability $1/2$, and even with probability $1/2$, and since the rolls are independent, we get

$$\Pr(X = k) = (1/2)^k.$$

3. What is $\Pr(X = k \wedge Y = 4)$? This happens if the rolls give the sequence

$$\underbrace{\text{odd, odd}, \dots, \text{odd}}_{k-1 \text{ times}}, 4.$$

Since the result is odd with probability $1/2$, and 4 with probability $1/6$, and since the rolls are independent, we get

$$\Pr(X = k \wedge Y = 4) = (1/2)^{k-1} \cdot 1/6 = (1/2)^k \cdot 1/3.$$

4. We conclude that

$$\Pr(X = k \wedge Y = 4) = \Pr(X = k) \cdot \Pr(Y = 4).$$

Question 5: Elisa Kazan is having a party at her home. Elisa has a round table that has 52 seats numbered $0, 1, 2, \dots, 51$ 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 $\{0, 1, 2, \dots, 51\}$. Consider the consecutive group of 7 people that sit in seats $k, k+1, k+2, \dots, k+6$; 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, \dots, P_{15} . For each i with $1 \leq i \leq 15$, define the indicator random variable

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

- 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 $\mathbb{E}(X)$?

Solution: First note that

$$X = \sum_{i=1}^{15} X_i.$$

Let us consider P_i and let m be P_i 's seat number. Then $X_i = 1$ if and only if $k \in \{m - 6, m - 5, \dots, m\}$ (seat numbers are to be read modulo 52). Thus, there are 7 choices for k for which $X_i = 1$. It follows that

$$\mathbb{E}(X_i) = \Pr(X_i = 1) = 7/52.$$

This gives

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}\left(\sum_{i=1}^{15} X_i\right) \\ &= \sum_{i=1}^{15} \mathbb{E}(X_i) \\ &= \sum_{i=1}^{15} 7/52 \\ &= 15 \cdot 7/52 \\ &= 2 + 1/52 \\ &> 2. \end{aligned}$$

Now for the second part: Assume the claim is false. Then, no matter which value of k we choose, the value of the random variable X is at most 2. But then, $\mathbb{E}(X) \leq 2$, which contradicts the first part of the question.

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

1. Initialize both L_1 and L_2 to be empty.
2. For each $i = 1, 2, \dots, n$, 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 $\text{MERGE}(L_1, L_2)$ 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 $\text{MERGE}(L_1, L_2)$ 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 $i, i+1, \dots, j-1$ and the element j ?

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

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

Solution: We start with the case when $n = 2$. We go through all 4 possibilities for the lists L_1 and L_2 :

1. $L_1 = (1, 2)$ and L_2 is empty:
 - (a) In this case, $X = 0$.
 - (b) This case happens with probability $1/4$.
2. L_1 is empty and $L_2 = (1, 2)$:
 - (a) In this case, $X = 0$.
 - (b) This case happens with probability $1/4$.
3. $L_1 = (1)$ $L_2 = (2)$:
 - (a) In this case, $X = 1$.
 - (b) This case happens with probability $1/4$.
4. $L_1 = (2)$ $L_2 = (1)$:
 - (a) In this case, $X = 1$.
 - (b) This case happens with probability $1/4$.

From this, we get

$$\begin{aligned} \mathbb{E}(X) &= 0 \cdot \Pr(X = 0) + 1 \cdot \Pr(X = 1) \\ &= 0 + 1(1/4 + 1/4) \\ &= 1/2. \end{aligned}$$

Next we do the case when $n = 3$. We go through all 8 possibilities for the lists L_1 and L_2 :

1. $L_1 = (1, 2, 3)$ and L_2 is empty:
 - (a) In this case, $X = 0$.
 - (b) This case happens with probability $1/8$.
2. $L_1 = (1, 2)$ and $L_2 = (3)$:
 - (a) In this case, $X = 2$.
 - (b) This case happens with probability $1/8$.
3. $L_1 = (1, 3)$ and $L_2 = (2)$:
 - (a) In this case, $X = 2$.
 - (b) This case happens with probability $1/8$.
4. $L_1 = (2, 3)$ and $L_2 = (1)$:
 - (a) In this case, $X = 1$.
 - (b) This case happens with probability $1/8$.
5. $L_1 = (1)$ and $L_2 = (2, 3)$:
 - (a) In this case, $X = 1$.
 - (b) This case happens with probability $1/8$.
6. $L_1 = (2)$ and $L_2 = (1, 3)$:
 - (a) In this case, $X = 2$.
 - (b) This case happens with probability $1/8$.
7. $L_1 = (3)$ and $L_2 = (1, 2)$:
 - (a) In this case, $X = 2$.
 - (b) This case happens with probability $1/8$.
8. L_1 is empty and $L_2 = (1, 2, 3)$:
 - (a) In this case, $X = 0$.
 - (b) This case happens with probability $1/8$.

From this, we get

$$\begin{aligned}
 \mathbb{E}(X) &= 0 \cdot \Pr(X = 0) + 1 \cdot \Pr(X = 1) + 2 \cdot \Pr(X = 2) \\
 &= 0 + 1(1/8 + 1/8) + 2(1/8 + 1/8 + 1/8 + 1/8) \\
 &= 5/4.
 \end{aligned}$$

For the next part of the question, the main observation is the following:

Claim 1 $X_{ij} = 1$ if and only if

1. $i, i+1, \dots, j-1$ are all in L_1 and j is in L_2 , or
2. $i, i+1, \dots, j-1$ are all in L_2 and j is in L_1 .

Proof. Assume that $X_{ij} = 1$. Then i and j are compared and, therefore, they must be in different lists. We may assume that i is in L_1 and j is in L_2 . We proceed by contradiction. We assume that k is such that $i+1 \leq k \leq j-1$ and k is in L_2 . At the moment when k is removed from L_2 ,

- i has already been removed from L_1 and
- j is still in L_2 .

Therefore, i and j are never compared. This is a contradiction.

To prove the converse, assume that $i, i+1, \dots, j-1$ are all in L_1 and j is in L_2 . At the moment when i is removed from L_1 ,

- i has just been compared to the first element in L_2 ,
- the result of this comparison was that i is the smaller of the two,
- j is still in L_2 .

Therefore, the first element in L_2 can only be j . Thus, i and j are compared and $X_{ij} = 1$. ■

All of $i, i+1, \dots, j-1$ are in L_1 and j is in L_2 is the same as all coin flips for $i, i+1, \dots, j-1$ come up heads, and the coin flip for j comes up tails; this happens with probability $(1/2)^{j-i+1}$. Therefore,

$$\mathbb{E}(X_{ij}) = \Pr(X_{ij} = 1) = 2 \cdot (1/2)^{j-i+1} = (1/2)^{j-i}.$$

Since

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij},$$

we get

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}\left(\sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}\right) \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathbb{E}(X_{ij}) \\ &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n (1/2)^{j-i} \end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{n-1} \left(\frac{1}{2} + \left(\frac{1}{2} \right)^2 + \cdots + \left(\frac{1}{2} \right)^{n-i} \right) \\
&= \sum_{i=1}^{n-1} \frac{1}{2} \cdot \left(1 + \frac{1}{2} + \cdots + \left(\frac{1}{2} \right)^{n-i-1} \right) \\
&= \sum_{i=1}^{n-1} \left(1 - \left(\frac{1}{2} \right)^{n-i} \right) \\
&= (n-1) - \sum_{i=1}^{n-1} \left(\frac{1}{2} \right)^{n-i} \\
&= (n-1) - \left(\frac{1}{2} + \left(\frac{1}{2} \right)^2 + \cdots + \left(\frac{1}{2} \right)^{n-1} \right) \\
&= (n-1) - \frac{1}{2} \cdot \left(1 + \frac{1}{2} + \cdots + \left(\frac{1}{2} \right)^{n-2} \right) \\
&= (n-1) - \left(1 - \left(\frac{1}{2} \right)^{n-1} \right) \\
&= n - 2 + \left(\frac{1}{2} \right)^{n-1}.
\end{aligned}$$

You may check that for $n = 2$ and $n = 3$, this is the same answer as we got above.

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 $\Pr(X = k)$.

Solution: We have $X = k$ if and only if the coin flips result in

$$\underbrace{\cdots HTHTHT}_{k-2 \text{ flips}} HH$$

or

$$\underbrace{\cdots THTHTH}_{k-2 \text{ flips}} TT.$$

Since the coin flips are independent, we get

$$\Pr(X = k) = 2 \cdot (1/2)^k.$$

- 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 $0 < x < 1$, $\sum_{k=1}^{\infty} kx^k = \frac{x}{(1-x)^2}$.

Solution: The possible values for X are $2, 3, 4, \dots$. This gives

$$\begin{aligned}
 \mathbb{E}(X) &= \sum_{k=2}^{\infty} k \cdot \Pr(X = k) \\
 &= \sum_{k=2}^{\infty} 2k \cdot (1/2)^k \\
 &= 2 \sum_{k=2}^{\infty} k \cdot (1/2)^k \\
 &= 2 \left(\sum_{k=1}^{\infty} k \cdot (1/2)^k - 1/2 \right) \\
 &= 2 \left(\frac{1/2}{(1 - 1/2)^2} - 1/2 \right) \\
 &= 3.
 \end{aligned}$$

- Determine $\Pr(X \geq 1)$.

Solution: We flip the coin at least twice. Thus, no matter what the coin flips are, we have $X \geq 2$. It follows that

$$\Pr(X \geq 1) = 1.$$

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

Solution: We have $X \geq k$ if and only if the coin flips result in

$$\underbrace{HTHTHT \dots}_{k-1 \text{ flips}}$$

or

$$\underbrace{THTHTH \dots}_{k-1 \text{ flips}}.$$

Since the coin flips are independent, we get

$$\Pr(X \geq k) = 2 \cdot (1/2)^{k-1} = (1/2)^{k-2}.$$

- 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 .

Solution:

$$\begin{aligned}\mathbb{E}(X) &= \Pr(X \geq 1) + \sum_{k=2}^{\infty} \Pr(X \geq k) \\ &= 1 + \sum_{k=2}^{\infty} (1/2)^{k-2} \\ &= 1 + (1 + 1/2 + 1/4 + 1/8 + 1/16 + \cdots) \\ &= 1 + 2 \\ &= 3.\end{aligned}$$

This is the same answer as we got before!