# COMP 2804B — Assignment 4

**Due:** Sunday December 3rd, 11:59 pm.

**Assignment Policy:**

- Your assignment must be submitted as a single .pdf file through Brightspace. Typesetting (using Latex, Word, Google docs, etc) is recommended but not required. Marks will be deducted for illegible or messy solutions. This includes but is not limited to excessive scribbling, shadows, blurry photos, messy handwriting, etc.
- **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", or "my dog ate my laptop charger".**
- 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:** Lecture 16 - Infinite Sample Spaces. Lecture 18 - Expected Value.
You and your friend decide to play a dice game. You each know a little about probability - for instance, you know that if you roll 2 dice the most likely number to come up is 7. You suggest the following game. You keep rolling the dice until one of the following two events happens:

- If the roll is 3 or 11, your friend wins.
- If a 7 comes up *twice* before a 3 or 11 is rolled, you win.

Each roll is mutually independent.

a) What is the probability that you win and what is the probability that your friend wins?

**Solution:** [Solution 1]We split the game into two rounds. The first round lasts until a 3, 7, or 11 is rolled. The second round lasts until the game ends. Let $A$ be the event "your friend wins.". Let $A_1$ be the event "your friend wins at the end of round 1" and let $A_2$ be the event "your friend does not win at the end of round 1 but wins at the end of round 2. Then, by the Sum Rule, $\Pr(A) = \Pr(A_1) + \Pr(A_2)$.

Let $B_1$ be the event "round 1 ends with a 3 or 11. Let $B_2$ be the event "round 2 ends with a 3 or 11. The roll that ends round 1 ends with a 3 or 7 or 11, so

$$\Pr(B_1) = \Pr\left(\text{"roll 3 or 11} \mid \text{"roll 3, 11, or 7"}\right) = \frac{2/36 + 2/36}{2/36 + 2/36 + 6/36} = 2/5$$

(If you want to do this more formally, you would condition on the number $i$ of rounds by defining, for each $i \geq 1$ the events $C_i$ = "round 1 ends on the $i$th roll" and $B_{i,1} = C_i \cap B_1$ and use the Sum Rule to get $\Pr(B_1) = \sum_{i=1}^{\infty} \Pr(B_{i,1}) = \sum_{i=1}^{\infty} \Pr(C_i) \cdot \Pr(B_{i,1} \mid C_i) = 2/5$.)

The events $A_1$ and $B_1$ are the same, i.e., $A_1 = B_1$, so $\Pr(A_1) = \Pr(B_1) = 2/5$. Now we can write $A_2$ as $A_2 = \overline{A_1} \cap B_2$, so

$$\Pr(A_2) = \Pr(\overline{A_1}) \cdot \Pr(B_2 \mid \overline{A_1}) = \Pr(\overline{A_1}) \cdot \Pr(B_1) = 3/5 \cdot 2/5 = 6/25 \ .$$

The switch from $\Pr(B_2 \mid \overline{A_1})$ to $\Pr(B_1)$ comes from the fact that the second round, if it occurs, is exactly the same as the first round. You keep rolling dice until you get a 3, 7, or 11. So we finish with

$$\Pr(A) = \Pr(A_1) + \Pr(A_2) = 2/5 + 6/25 = 16/25 \ .$$

The probability that your friend wins is $\Pr(A)$, so the probability that you win is $\Pr(\overline{A}) = 1 - \Pr(A) = 9/25$.

**Note:** There is another way to see this. Imagine that you always play two rounds, no matter how the first round ends, but the rules are the same. So for you to win the first round has to end with a 7 and the second round has to end with a 7. In the notation above, the event $\overline{B_1} \cap \overline{B_2}$ is the event "you win." The probability of each of these events is 3/5 and they are independent, so $\Pr(\overline{B_1} \cap \overline{B_2}) = (3/5)^2 = 9/25$.

**Solution:** [Solution 2:] We will start by computing the probability on individual die rolls. Let $D_7$ be the event that you roll 7 on two dice. Let $D_3$ be the event that you roll a 3 or 11 (we will shorten $3or11$ to 3 for all subscripts so that our calculations are neater). Let $D_0$ be the event that you roll anything other than $3, 7$ or 11, and thus no one wins that round. A simple calculation tells us that $\Pr(D_7) = \frac{3}{18}, \Pr(D_3) = \frac{2}{18}$, and $\Pr(D_0) = \frac{13}{18}$.

The easiest way to answer this question is to find the probability that your friend wins and use the complement rule. However, we will find the probability that you win directly, since we will need some of these calculations later.

Let $W$ be the event that you win. There are two (disjoint) cases to consider. Let $W_{73}$ be the case where you roll one 7 before you roll a 3 or 11. Let $W_3$ be the event that no 7's are rolled. Then $W = W_{73} \cup W_3$, and $\Pr(W) = \Pr(W_{73}) + \Pr(W_3)$ (since they are disjoint). We will also define the event $W_7$ is that we roll a 7 before a 3 or 11.

Finding $\Pr(W_3)$ is straightforward. It is an infinite sample space where each outcome has the form $D_0^n D_3, n \geq 0$. Thus

$$\begin{aligned}
\Pr(W_3) &= \sum_{n=0}^{\infty} \Pr(D_0^n D_3) \\
&= \sum_{n=0}^{\infty} \Pr(D_0)^n \Pr(D_3) \\
&= \sum_{n=0}^{\infty} \left(\frac{13}{18}\right)^n \frac{2}{18} \\
&= \frac{2}{18} \sum_{n=0}^{\infty} \left(\frac{13}{18}\right)^n \\
&= \frac{2}{18} \left(\frac{1}{1 - \frac{13}{18}}\right) \\
&= \frac{2}{18} \left(\frac{18}{5}\right) \\
&= \frac{2}{5}.
\end{aligned}$$

The outcomes in $W_{73}$ have the form $D_0^n D_7 D_0^m D_3, n \geq 0, m \geq 0$. To find $\Pr(W_{73})$ we will divide each outcome into two distinct substrings. Like in class, we will play the game in two "rounds". The first round is until 7 comes up once, so all substrings of the form $D_0^n D_7, n \geq 0$. We will denote this $W_{7,1}$ (for rolling a 7 first in round one). The second round we roll until 3 or 11 comes up, so all substrings of the form $D_0^m D_3, m \geq 0$. We will denote this $W_{3,2}$ (for rolling a 3 or 11 first in round two). We can observe that, if we consider the second round independently, $\Pr(W_{3,2}) = \Pr(W_3)$ that we computed above. We are left to compute $W_{7,1}$.

$$\begin{aligned}
\Pr(W_{7,1}) &= \sum_{n=0}^{\infty} \Pr(D_0^n D_7) \\
&= \sum_{n=0}^{\infty} \Pr(D_0)^n \Pr(D_7) \\
&= \sum_{n=0}^{\infty} \left(\frac{13}{18}\right)^n \frac{3}{18}
\end{aligned}$$

$$= \frac{3}{18} \sum_{n=0}^{\infty} \left( \frac{13}{18} \right)^n$$
$$= \frac{3}{18} \left( \frac{1}{1 - \frac{13}{18}} \right)$$
$$= \frac{3}{18} \left( \frac{18}{5} \right)$$
$$= \frac{3}{5}.$$

Here we *should* use the law of total probability, but we will accept an informal analysis. Which is that, assuming we are in case $W_{73} = W_{7,1} \cap W_{3,2}$, then these two "rounds" are independent, since all die rolls are independent. Thus

$$\begin{aligned} \Pr(W_{73}) &= \Pr(W_{7,1} \cap W_{3,2}) \\ \Pr(W_{73}) &= \Pr(W_{7,1}) \cdot \Pr(W_{3,2}) \\ &= \frac{2}{5} \cdot \frac{3}{5} \\ &= \frac{6}{25}. \end{aligned}$$

Thus the probability that you win is

$$\begin{aligned} \Pr(W) &= \Pr(W_{73} \cup W_3) \\ &= \Pr(W_{73}) + \Pr(W_3) \\ &= \frac{6}{25} + \frac{10}{25} \\ &= \frac{16}{25}. \end{aligned}$$

Then the probability that your friend wins is $\frac{9}{25}$.

b) You decide to play for money. Every time a 3 or 11 comes up, you pay your friend \$3. Every time a 7 comes up, he pays you \$2. What is your expected winnings per die roll?

**Solution:** It looks like I reversed you and your friend. Regardless the calculations are the same, so whoever they assigned which value to does not matter. Let $X$ be the amount you win on a given roll. Then

$$\begin{aligned} E(X) &= \sum_k k \cdot \Pr(X = k) \\ &= 2 \cdot \Pr(D_7) - 3 \cdot \Pr(D_3) \\ &= 2 \cdot \frac{1}{6} - 3 \cdot \frac{1}{9} \\ &= 0. \end{aligned}$$

This game pays even money.

c) How many dice rolls would you expect there to be before a 3, 7, or 11 is rolled?
**Solution:** Let $D_{73}$ be the event that you roll a 3, 7, or 11. Then $\Pr(D_{73}) = \frac{5}{18}$. This is a geometric distribution, so if $X$ is the number of die rolls,

$$E(X) = \frac{1}{\Pr(D_{73})} = \frac{1}{\frac{5}{18}} = \frac{18}{5}$$

.

d) How many dice roll would you expect before someone wins the original game? That is, how many dice rolls would you expect before a 7 is rolled twice or a 3 or 11 is rolled once?
**Solution:** Let $X$ be the random variable that counts the number of dice rolls that take place during the game. Let $X_1$ be the random variable that counts the number of dice rolls that take place in the first round. Let $X_2$ be the random variable that is either

- 0 if the first round ends by rolling a 3 or 11; or
- the number of dice rolls that take place in the second round.

Then $X = X_1 + X_2$. From the previous question, already know that $\mathrm{E}(X_1) = 18/5$. For $X_2$, we get

$$\begin{aligned}
\mathrm{E}(X_2) &= \sum_{i=0}^{\infty} i \cdot \Pr(X_2 = i) \\
&= \sum_{i=1}^{\infty} i \cdot \Pr(X_2 = i) && \text{(since first term is } 0 \cdot \Pr(X_2 = 0) = 0) \\
&= \sum_{i=1}^{\infty} i \cdot \Pr(X_2 = i \cap X_2 \neq 0) \\
&= \sum_{i=1}^{\infty} i \cdot \Pr(X_2 = i \cap \overline{B_1}) && \text{(since } B_1 \text{ implies } X_2 = 0) \\
&= \sum_{i=1}^{\infty} i \cdot \Pr(\overline{B_1}) \cdot \Pr(X_2 = i \mid \overline{B_1}) && \text{(definition of conditional probability)} \\
&= \frac{3}{5} \sum_{i=1}^{\infty} i \cdot \Pr(X_2 = i \mid \overline{B_1}) && \text{(from the answer to (a))} \\
&= \frac{3}{5} \sum_{i=1}^{\infty} i \cdot \Pr(X_1 = i) && \text{(explained below)} \\
&= \frac{3}{5} \mathrm{E}(X_1) && \text{(definition of expected value)} \\
&= \frac{3}{5} \cdot 185 = \frac{54}{25} \mathrm{E}(X_1) && \text{(from answer to (c))}
\end{aligned}$$

To see why the second-last equality holds, notice that round 2, if it takes place is played in exactly the same way as round 1. The event "round 2 takes place" is exactly the event $\overline{B_1}$. Thus $\Pr(X_2 = i \mid \overline{B_1}) = \Pr(X_1 = i)$.

Therefore, $\mathrm{E}(X) = \mathrm{E}(X_1) + \mathrm{E}(X_2) = 18/5 + 54/25 = 144/25$.

Again, another way to see this is to imagine that the always play both rounds, but we only count the length, $X_2$ of the second round if the first round ends in a 7, which happens with probability 3/5. Therefore $\mathrm{E}(X) = (18/5) + (3/5) \cdot (18/5) = 144/25$.

The next question uses a standard deck of cards. A standard deck is a deck of 52 cards. If $R = \{A, 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K\}$ is the set of possible ranks and $S = \{\heartsuit, \diamondsuit, \spadesuit, \clubsuit\}$ are the set of suits, then a deck of 52 cards $D = S \times R = \{(\heartsuit, A), (\diamondsuit, A), (\spadesuit, A), (\clubsuit, A), (\heartsuit, 2), (\diamondsuit, 2), (\spadesuit, 2), (\clubsuit, 2), ..., (\spadesuit, K), (\clubsuit, K)\}$. The ranks $J$, $Q$, and $K$ are known as "face cards".

**Question 3:** Lecture 17 - Random Variables, Lecture 18 - Expected Value.

Let $D$ be a standard deck of cards. For a card $c \in D$, let the value of $X(c)$ be equal to its rank if the rank is an integer, 11 if the rank is $A$ and 10 if the rank is in $\{J, Q, K\}$. For example, if $X(c)$ random variable mapping the card to its value, then $X(\heartsuit, A) = 11$, $X(\spadesuit, A) = 11$, $X(\spadesuit, 7) = 7$, $X(\spadesuit, 8) = 8$, $X(\spadesuit, K) = 10$, etc. Assume you are playing Blackjack. You are dealt 2 cards, and the dealer deals themselves 1 card up and 1 card down so you cannot see it (the card that is face down is known as the *hole card*). One strategy in Blackjack is to always play as if the dealer's hole card is worth 10. Let

$$\begin{aligned} V &= \text{The sum of the values of the two cards in your hand.} \\ F &= \text{The value of the dealer's card + 10} \\ Z &= \text{F - V} \end{aligned}$$

a) What is $E(Z)$?
**Solution:** Using linearity of expectation we realize $E(Z) = E(F) - E(V)$. Let $X_D$ be the value of the card in the dealer's hand. Then $E(F) = E(X_D+10) = E(X_D)+E(10)$. To determine $E(X_D)$, we note that the dealer has one card chosen uniformly at random from the 52 in the deck. Thus

$$\begin{aligned} E(X_D) &= \sum_{k=2}^{11} k \cdot \Pr(X_D = k) \\ &= \sum_{k=2}^{9} k \cdot \frac{4}{52} + 11 \cdot \frac{4}{52} + 10 \cdot \frac{12}{52} \\ &= \frac{85}{13} && \text{according to Wolfram Alpha.} \end{aligned}$$

Therefore

$$E(F) = \frac{85}{13} + \frac{130}{13} = \frac{215}{13}.$$

Let $X_1$ and $X_2$ be the first and second cards in your hand, and note that $E(X_1) = E(X_2) = E(X_D)$, since all cards were drawn uniformly at random. Using linearity of expectation:

$$\begin{aligned} E(V) &= E(X_1) + E(X_2) \\ &= \frac{85}{13} + \frac{85}{13} \\ &= \frac{170}{13} \approx 13.07. \end{aligned}$$

Thus $E(Z) = E(F) - E(V) = \frac{215}{13} - \frac{170}{13} = \frac{45}{13}$.

b) Are $F$ and $Z$ independent random variables?
**Solution:** We will say no and provide a counter-example. Consider the events $F = 12$ and $Z = 17$. If $Z = 17$ then it must be that $F = 21$ and $V = 4$. Which means that $F = 12 \cap Z = 17 = \emptyset$. Since $F = 12$ can happen if the dealer has a 1, $\Pr(F = 12) > 0$. $Z = 17$ is also non-empty, thus $\Pr(F = 12 \cap Z = 17) \neq \Pr(F = 12) \cdot \Pr(Z = 17)$ and $F$ and $Z$ are not independent.

**Question 4:** Lectures 19, 20.

You are in a class of 200 people. Let $X$ be the number of different birthdays among these 200 people (assuming no one was born on February 29, i.e., on a leap year). Determine the expected value $E(X)$ of $X$.

Hint: Use indicator random variables. **Solution:** There are 365 days in a year. We will define the indicator random variables for each day as follows:

$$X_i = \begin{cases} 1 \text{ if at least one person was born on day } i \\ 0 \text{ otherwise} \end{cases}$$

Since $X_i$ is an indicator random variable, $E(X_i) = Pr(X_i = 1)$, which is the probability that, out of 200 students, at least one has their birthday on day $i$. We will use the complement rule, thus $Pr(X_i = 1) = 1 - Pr(X_i = 0)$, and $X_i = 0$ is the event that no one is born on day $i$. Let $a_{ij}$ be the event that person $j$ was born on day $i$. $Pr(a_{ij}) = \frac{1}{365}$, and thus $Pr(\overline{a}_{ij}) = \frac{364}{365}$. Since each birthday is determined independently, the probability of being born on day $i$ over 200 students is

$$\begin{aligned} Pr(X_i = 0) &= Pr(\overline{a}_{i1} \wedge \overline{a}_{i2} \wedge \ldots \wedge \overline{a}_{i200}) \\ &= Pr(\overline{a}_{i1}) \cdot Pr(\overline{a}_{i2}) \cdot \ldots \cdot Pr(\overline{a}_{i200}) \\ &= \left(\frac{364}{365}\right)^{200}. \end{aligned}$$

Thus $Pr(X_i = 1) = 1 - Pr(X_i = 0) = 1 - \left(\frac{364}{365}\right)^{200}$. Linearity of expectation tells us:

$$\begin{aligned} E(X) &= E(X_1 + X_2 + ... + X_{365}) \\ &= E(X_1) + E(X_2) + ... + E(X_{365}) \\ &= \sum_{j=1}^{365} Pr(X_i = 1) \\ &= \sum_{j=1}^{365} \left(1 - \left(\frac{364}{365}\right)^{200}\right) \\ &= 365 \cdot \left(1 - \left(\frac{364}{365}\right)^{200}\right) \approx 154.14. \end{aligned}$$

The rest of the questions are meant to be a combination of techniques that you have seen, and thus they don't necessarily correspond cleanly to one or two Lectures. Lectures are mentioned where appropriate.

**Question 5:** Consider a fair 6-sided die.

a) Roll the die twice. What is the expected value of the highest number?
**Solution:** Let $X$ be the highest roll out of two dice. As an example, consider if the highest die is 4. To count $X = 4$, we look at two cases: either one die is 4, or both are. There is 1 way to have both dice be 4, so $\Pr(4, 4) = \frac{1}{6^2}$. To find the probability of the other case, we first choose one of the two dice to be 4. Then the other die can have any number $1 - 3$. Then $\Pr(X = 4) = \frac{1+2\cdot 3}{6^2}$. If we generalize this we get $\Pr(X = i) = \frac{1+2\cdot(i-1)}{6^2}$. Thus

$$\begin{aligned} E(X) &= \sum_{i=1}^{6} i \cdot \Pr(X = i) \\ &= \sum_{i=1}^{6} i \cdot \left(\frac{1 + 2 \cdot (i-1)}{6^2}\right) \\ &= \frac{161}{36} \approx 4.48. \end{aligned}$$

b) Roll the die once. If the number is $> 3$ keep that number. Otherwise roll the die again and keep the highest of the two rolls. What is the expected value?
**Solution:** We will sum over the range. Note that if the highest number is 1, 2 or 3, it is the same calculation as above. Consider that the highest number is $i > 3$. Then we either rolled $i$ on the first roll (event $F_i$), or we rolled $1 - 3$ on the first roll, and $i$ on the second roll (event $S_i$). Thus $\Pr(X = i) = \Pr(F_i) + \Pr(S_i) = \frac{1}{6} + \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{4}$.

$$\begin{aligned} E(X) &= \sum_{k=1}^{6} k \cdot \Pr(X = k) \\ &= 1 \cdot \frac{1}{36} + 2 \cdot \frac{1+2}{36} + 3 \cdot \frac{1+2\cdot 2}{36} + 4 \cdot \left(\frac{1}{4}\right) + 5 \cdot \left(\frac{1}{4}\right) + 6 \cdot \left(\frac{1}{4}\right) \\ &= \frac{157}{36} \approx 4.37. \end{aligned}$$

c) Roll the die once. If the number is $> 4$ keep that number. Otherwise roll the die again and keep the highest of the two rolls. What is the expected value? What strategy gives us the highest value on average?
**Solution:** Consider that the highest number is $i > 4$. Then we either rolled $i$ on the

first roll (event $F_i$), or we rolled $1-4$ on the first roll, and $i$ on the second roll (event $S_i$). Thus $\Pr(X=i) = \Pr(F_i) + \Pr(S_i) = \frac{1}{6} + \frac{2}{3} \cdot \frac{1}{6} = \frac{5}{18}$.

$$\begin{aligned}
E(X) &= \sum_{k=1}^{6} k \cdot \Pr(X=k) \\
&= 1 \cdot \frac{1}{36} + 2 \cdot \frac{1+2}{36} + 3 \cdot \frac{1+2 \cdot 2}{36} + 4 \cdot \frac{1+2 \cdot 3}{36} + 5 \cdot \left(\frac{5}{18}\right) + 6 \cdot \left(\frac{5}{18}\right) \\
&= \frac{40}{9} \approx 4.44.
\end{aligned}$$

Thus the first strategy gives the highest average value.

**Question 6:** In this question we will consider bitstrings. A bit is called *lonely* if it is a 1 and every adjacent bit is a 0. A bit is *not lonely* if it is a 1 and it is adjacent to at least one other 1.

a) Consider a random bitstring of length 10. What is the expected number of lonely bits?
**Solution:** We will use indicator random variables.

$$X_i = \begin{cases} 1 \text{ if bit } i \text{ is lonely} \\ 0 \text{ otherwise.} \end{cases}$$

Since these are indicator random variables, $E(X_i) = \Pr(X_i = 1)$. The first and last bit are lonely if they are a 1 and the bit beside them is a 0. So

$$\Pr(X_1 = 1) = \Pr(X_{10} = 1) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}.$$

Bits 2 through 9 are lonely if they are a 1 and both adjacent bits are 0. Thus $\Pr(X_i = 1) = \frac{1}{8}, 2 \leq i \leq 9$. Thus the expected number of lonely bits are

$$\begin{aligned}
E(X) &= E(X_1) + E(X_2) + ...E(X_{10}) \\
&= 2 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} \\
&= \frac{3}{2}.
\end{aligned}$$

b) We choose a bitstring uniformly at random from all bitstrings of length 5 with exactly three 1's. What is the expected number of lonely bits?
**Solution:** Since this bitstring is short and specific, we can enumerate the bitstrings with lonely bits. There are $\binom{5}{3} = 10$ bitstrings of length 5 with 3 ones. The ones with lonely bits are listed below:

| | |
|---|---|
| 10110 | 1 lonely bit |

| | |
|---|---|
| 10011 | 1 lonely bit |
| 11001 | 1 lonely bit |
| 01101 | 1 lonely bit |
| 01011 | 1 lonely bit |
| 11010 | 1 lonely bit |
| 10101 | 3 lonely bits |

Thus (ignoring $k = 0$):

$$\begin{aligned} E(X) &= \sum_k k \cdot \Pr(X = k) \\ &= 1 \cdot \frac{6}{10} + 3 \cdot \frac{1}{10} \\ &= \frac{9}{10}. \end{aligned}$$

c) What is the expected number of bits that are not lonely?
**Solution:** The easiest thing to do is use linearity of expectation. Let $Y$ be the number of not lonely bits. Let $Z = X + Y$. Since we know that each 1 is either lonely or not lonely, we know $Z = 3$. Thus

$$\begin{aligned} E(Z) &= E(X + Y) \\ E(Z) &= E(X) + E(Y) \\ 3 &= \frac{7}{10} + E(Y) \\ E(Y) &= \frac{23}{10}. \end{aligned}$$

d) We choose a bitstring uniformly at random from all bitstrings of length 10 with exactly four 1's. What is the expected number of lonely and not lonely bits?
**Solution:** We iterate over the domain, and break it into cases. Thus we use $E(X) = \sum_{\forall k} k \cdot \Pr(X = k)$. The sample space is $\binom{10}{4} = 210$.

- $k = 1$: That means there are three bits together and one bit separated. The patterns would either be $0^*10^+110^*$ or $0^*1110^+110^*$, where $0^*$ represents zero or more 0's, and $0^+$ represents one or more 0's. To count the number of strings that fall into this category, we can map this to a linear equation $x_1 + x_2 + x_3 = 6, x_1 \geq 0, x_2 \geq 1, x_3 \geq 0$. Note that $x_1, x_2$, and $x_3$ represent the number of 0's to the left of all 1's, between the two separate substrings of 1's (where there must be at least one 0), and to the right of all 1's. We know that the number of solutions to that equation is the same as the number of solutions to $x_1 + x_2' + x_3 = 5, x_1 \geq 0, x_2' \geq 0, x_3 \geq 0$ (subtract one from $x_2$ and the right hand side), which is $\binom{7}{2}$. So there are $2 \cdot \binom{7}{2} = 42$ such strings with $k = 1$ lonely bits.

- $k = 2$: We have patterns such as $0^*10^+10^+110^*$. There are 3 possible ways to arrange these sets of 1's (the two 1's come first, second, or third). And the number of ways to arrange the 0's between them is $x_1 + x_2 + x_3 + x_4 = 6, x_1 \geq 0, x_2 \geq 1, x_3 \geq 1, x_4 \geq 0$. The number of solutions to this equation is equal to the number of solutions to $x_1 + x_2 + x_3 + x_4 = 4, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0$, which is $\binom{7}{3}$. Thus there are $3 \cdot \binom{7}{3} = 105$ strings with $k = 2$ lonely bits.
- $k = 3$: There is no way to have $k = 3$ lonely bits.
- $k = 4$: The pattern is $0^*10^+10^+10^+10^*$ and it is the only such pattern. The number of ways to arrange the 0's are $x_1 + x_2 + x_3 + x_4 + x_5 = 3, x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0$, which is $\binom{7}{4} = 35$. Thus there are 35 strings with 4 lonely bits.

If we iterate over the range we get

$$\begin{aligned} E(X) &= \sum_k k \cdot \Pr(X = k) \\ &= \frac{1 \cdot 42 + 2 \cdot 105 + 3 \cdot 0 + 4 \cdot 35}{210} \\ &= \frac{28}{15} \approx 1.87. \end{aligned}$$

All that for two marks.

**Question 7:** We have a fair, 6-sided die. We roll this die until the sum of all rolls is $\geq 2$. Let $X$ be the number of rolls, and let $Y$ be the sum of all the rolls.

a) What is $E(X)$? Use the formula $E(X) = \sum_{\forall k} k \cdot \Pr(X = k)$.

**Solution:** The value of $X$ is always a positive integer. We can check that $\Pr(X = 1) = 5/6$ since $X = 1$ if and only if we roll a $2, 3, 4, 5$, or 6 on our first roll. On the other hand, if $X \neq 1$ then $X = 2$ since our first two rolls will always sum to at leat 2, so $\Pr(X = 2) = 1 - \Pr(X = 1) = 1/6$. Therefore

$$\mathrm{E}(X) = 1 \cdot \Pr(X = 1) + 2 \cdot \Pr(X = 2) = 5/6 + 2/6 = 7/6 \ .$$

b) What is $E(Y)$? Use the formula $E(Y) = \sum_{\forall k} k \cdot \Pr(Y = k)$.

**Solution:** We know that $Y$ is never less than 2 or greater than 7, so we need to evaluate $\Pr\{Y = k\}$ for each $k \in \{2, \ldots, 7\}$. For $k \in \{2, 3, 4, 5, 6\}$, there are exactly two ways to get $k$. We could roll $k$ on the first roll or we could roll a 1 followed by $k-1$. So, by the Sum Rule

$$\Pr(X = k) = 1/6 + 1/36 = 7/36$$

for $k \in \{2, 3, 4, 5, 6\}$. On the other hand, there is only way $Y = 7$. We first have to roll a 1 then roll a 6, so $\Pr(Y = 7) = 1/36$. So

$$\mathrm{E}(Y) = \sum_{k=2}^{7} k \Pr(X = k) = \sum_{k=2}^{6} 7k/36 + 7/36 = 147/36 = 49/12$$

c) Let $D$ be the value of a single die roll. We have seen in class that $E(D) = 3.5$. What is $E(X) \cdot E(D)$?

**Solution:** $\mathrm{E}(X) \cdot \mathrm{E}(D) = 7/6 \cdot 7/2 = 49/12$.

d) This is an example of Wald's Identity. Wald's Identity tells us that if $X$ is the number of die rolls, and the value of $X$ depends on a *stopping condition* (which it does in this case), then the expected sum $E(Y) = E(X) \cdot E(D)$. Find $E(X)$ if $X$ is the number of rolls until the sum is $\geq 3$. Then find the corresponding value $E(Y)$ using Wald's Identity.

**Solution:** We follow the same strategy as in Part (a):

(a) We get $X = 1$ exactly when our first roll is $3, 4, 5$ or $6$ so $\Pr(X = 1) = 4/6 = 24/36$.

(b) To get $X = 2$ we could start by rolling 1 followed by a $2, 3, 4, 5$ or $6$ or we could start by rolling 2 followed by anything. So $\Pr(X = 2) = (1/6) \cdot (5/6) + 1/6 = 11/36$.

(c) We get $X = 3$ when $X \neq 1$ and $X \neq 2$, so $\Pr(X = 1) = \Pr(X \neq 1 \cap X \neq 2) = 1 - \Pr(X = 1 \cup X = 2) = 1 - (\Pr(X = 1) + \Pr(X = 2)) = 1 - (11/36 + 24/36) = 1 - 35/36 = 1/36$.

If you found that confusing, then notice that $X = 3$ if and only if the first two rolls are both 1, which happens with probability $1/36$.

So

$$\mathrm{E}(X) = \sum_{k=1}^{3} k \Pr(X = k) = 1 \cdot 24/36 + 2 \cdot 11/36 + 3 \cdot 1/36 = 49/36\ .$$

So, by Wald's Identity, $\mathrm{E}(Y) = \mathrm{E}(X) \cdot \mathrm{E}(D) = (49/36) \cdot (7/2) = 343/72$.

**Question 8:** You should review Lecture 13 - Conditional Probability and Lecture 18 - Expected Value.

If $X$ is a random variable that can take any value $n$ where $n$ is an integer and $n \geq 1$, and if $A$ is an event, then the *conditional expected value* $E(X|A)$ is defined as

$$E(X|A) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k|A).$$

You roll a fair six-side die repeatedly until you see the number 6. Define the random variable $X$ to be the number of die rolls (including the last roll where you see 6). We have seen in class that $E(X) = 6$. Let $A$ be the event

$A =$ "You do **not** roll 6 on the first two rolls".

Determine the conditional expected value $E(X|A)$.

**Solution:** Our intuition tells us that the first answer should be $2 + \mathrm{E}(X) = 2 + 6 = 8$. Let's see if this agrees with what we get from the definition. For each integer $i \geq 1$, let $D_i$ be the result of the $i$th dice roll. For each $k \in \{1, 2\}$, $\Pr(X = k \mid A) = 0$. For $k \geq 3$,

$$\begin{aligned}
\Pr(X = i \mid A) &= \Pr(D_1 \neq 6 \cap D_2 \neq 6 \cap D_3 \neq 6 \cap \cdots \cap D_{k-1} \neq 6 \cap D_k = 6 \mid A) \\
&= \frac{\Pr(D_1 \neq 6 \cap D_2 \neq 6 \cap D_3 \neq 6 \cap \cdots \cap D_{k-1} \neq 6 \cap D_k = 6 \cap A)}{\Pr(A)} \\
&= \frac{\Pr(D_1 \neq 6 \cap D_2 \neq 6 \cap D_3 \neq 6 \cap \cdots \cap D_{k-1} \neq 6 \cap D_k = 6)}{\Pr(A)} \\
&= \frac{(5/6)^{k-1}\dot{(}1/6)}{\Pr((5/6)^2)} \\
&= (5/6)^{k-3} \cdot (1/6)
\end{aligned}$$

So,

$$\begin{aligned}
\mathrm{E}(X \mid A) &= \sum_{k=1}^{\infty} k \cdot \Pr(X = k \mid A) \\
&= \sum_{k=3}^{\infty} k \cdot \Pr(X = k \mid A) \\
&= \sum_{k=3}^{\infty} k \cdot (5/6)^{k-3} \cdot 1/6 \\
&= \sum_{k=1}^{\infty} (k+2) \cdot (5/6)^{k-1} \cdot 1/6
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{6}\left(\sum_{k=1}^{\infty} k \cdot (5/6)^{k-1} + 2\sum_{k=1}^{\infty} (5/6)^{k-1}\right) \\
&= \frac{1}{6}\left(\sum_{k=1}^{\infty} k \cdot (5/6)^{k-1} + 2\sum_{k=0}^{\infty} (5/6)^{k}\right) \\
&= \frac{1}{6}\left(\sum_{k=1}^{\infty} k \cdot (5/6)^{k-1} + 12\right) \\
&= \frac{1}{6}\left(36 + 12\right) = 48/6 = 8 \ .
\end{aligned}
$$