Section 3 – Solutions

Review

1) Conditional Probability. $\mathbb{P}(\mathcal{B} \mid \mathcal{A}) =$ ____________

$\mathbb{P}(\mathcal{B} \mid \mathcal{A}) = \frac{\mathbb{P}(\mathcal{B} \cap \mathcal{A})}{\mathbb{P}(\mathcal{A})}$

2) Independent Events. Two events $\mathcal{A}, \mathcal{B}$ are **independent** if $\mathbb{P}(\mathcal{A} \cap \mathcal{B}) =$ ____________.

$\mathsf{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(\mathcal{A})\,\mathbb{P}(\mathcal{B})$

If $\mathbb{P}(\mathcal{A}) \neq 0$, this is equivalent to $\mathbb{P}(\mathcal{B} \mid \mathcal{A}) =$ ____________.

$\mathsf{P}(\mathcal{B} \mid \mathcal{A}) = \mathbb{P}(\mathcal{B})$

If $\mathbb{P}(\mathcal{B}) \neq 0$, this is equivalent to $\mathbb{P}(\mathcal{A} \mid \mathcal{B}) =$ ____________.

$\mathsf{P}(\mathcal{A} \mid \mathcal{B}) = \mathbb{P}(\mathcal{A})$

3) Partition. Nonempty events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ partition the sample space Ω iff

(1) ________________________, (2) ________________________

(1) $\bigcup_{i=1}^{n} \mathcal{E}_i = \Omega$, (2) $\forall i \forall i \neq j, \mathcal{E}_i \cap \mathcal{E}_j = \varnothing$

4) Bayes Rule. For any events $\mathcal{A}$ and $\mathcal{B}$, $\mathbb{P}(\mathcal{A}|\mathcal{B}) =$ _______________ .

$\mathbb{P}(\mathcal{A} \mid \mathcal{B}) = \frac{\mathbb{P}(\mathcal{B} \mid \mathcal{A})\mathbb{P}(\mathcal{A})}{\mathbb{P}(\mathcal{B})}$

5) Chain Rule: Suppose $\mathcal{A}_1, \ldots, \mathcal{A}_n$ are events. Then,

$$\mathbb{P}(\mathcal{A}_1 \cap \ldots \cap \mathcal{A}_n) = \underline{\hspace{3cm}}$$

$\mathbb{P}(\mathcal{A}_1 \cap \ldots \cap \mathcal{A}_n) = \prod_{i=1}^{n} \mathbb{P}(\mathcal{A}_i \mid \mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_{i-1})$

6) Law of Total Probability (LTP): Suppose $\mathcal{E}_1, \ldots, \mathcal{E}_n$ is a partition of Ω and let $\mathcal{B}$ be any event. Then $\mathbb{P}(\mathcal{B}) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \cap \mathcal{E}_i) =$ ____________

$\mathbb{P}(\mathcal{B}) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \cap \mathcal{E}_i) = \sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \mid \mathcal{E}_i)\,\mathbb{P}(\mathcal{E}_i)$

7) Bayes Theorem with LTP: $\mathcal{E}_1, \ldots, \mathcal{E}_n$ is a partition of Ω and let $\mathcal{B}$ be any event. Then

$$\mathbb{P}(\mathcal{E}_1 \mid \mathcal{B}) = \frac{\underline{\hspace{3cm}}}{\sum_{i=1}^{n} \underline{\hspace{3cm}}} .$$

$\mathbb{P}(\mathcal{E}_1 \mid \mathcal{B}) = \frac{\mathbb{P}(\mathcal{B} \mid \mathcal{E}_1)\mathbb{P}(\mathcal{E}_1)}{\sum_{i=1}^{n} \mathbb{P}(\mathcal{B} \mid \mathcal{E}_i)\mathbb{P}(\mathcal{E}_i)}$.

Task 1 – Naive Bayes

Most of Section 3 will be an introduction to an application of Bayes' Theorem called the Naive Bayes Classifier.

Task 2 – Flipping Coins

We consider two independent tosses of the same coin. The coin is "heads" one quarter of the time.

a) What is the probability that the second toss is "heads" given that the first toss is "tails"?

Consider the probability space with sample space $\Omega = \{HH, TT, HT, TH\}$. Because heads come $1/4$ of the time, and tails $3/4$, we have $\mathbb{P}(HH) = 1/4 \times 1/4 = 1/16$, $\mathbb{P}(HT) = \mathbb{P}(TH) = 3/4 \times 1/4 = 3/16$ and finally $\mathbb{P}(TT) = 9/16$.

Then, let $\mathcal{A}$ be the event that the first coin is tails, and let $\mathcal{B}$ be the event that the second coin is heads. Then,

$$\mathbb{P}(\mathcal{B}|\mathcal{A}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{B})}{\mathbb{P}(\mathcal{A})} .$$

Note that $\mathcal{A} = \{TT, TH\}$ and $\mathcal{B} = \{HH, TH\}$, and thus

$$\mathbb{P}(\mathcal{A}) = \mathbb{P}(TT) + \mathbb{P}(TH) = 9/16 + 3/16 = 12/16 = 3/4$$
$$\mathbb{P}(\mathcal{A} \cap \mathcal{B}) = \mathbb{P}(TH) = 3/16 .$$

Therefore, $\mathbb{P}(\mathcal{B}|\mathcal{A}) = (3/16)/(3/4) = 1/4$.

It is important to realize that this exactly what we would have expected – indeed, we model the coins to be independent.

b) What is the probability that the second toss is "heads" given that at least one of the tosses is "tails"?

Here, $\mathcal{B}$ is as as in **a)**. We define $\mathcal{C} = \{TH, TT, TT\}$, and we want $\mathbb{P}(\mathcal{B}|\mathcal{C})$. Note that

$$\mathbb{P}(\mathcal{C}) = 1 - \mathbb{P}(HH) = 15/16$$
$$\mathbb{P}(\mathcal{B} \cap \mathcal{C}) = \mathbb{P}(TH) = 3/16 .$$

Therefore,

$$\mathbb{P}(\mathcal{B} \mid \mathcal{C}) = \frac{\mathbb{P}(\mathcal{A} \cap \mathcal{C})}{\mathbb{P}(\mathcal{C})} = \frac{3/16}{15/16} = \frac{3}{15} = \frac{1}{5} .$$

c) In the probability space of this task, give an example of two events that are disjoint but not independent.

$\mathcal{E}_1 = \{TT\}$ and $\mathcal{E}_2 = \{HH\}$ are disjoint, but not independent. Indeed, $\mathbb{P}(\mathcal{E}_1 \cap \mathcal{E}_2) = \mathbb{P}(\varnothing) = 0$, but each event occurs with positive probability, and so $\mathbb{P}(\mathcal{E}_1) \cdot \mathbb{P}(\mathcal{E}_2) > 0$.

d) In the probability space of this task, give an example of two events that are independent but not disjoint.

$\mathcal{E}_1 = \{TH, HH\}$ and $\mathcal{E}_2 = \{TH, TT\}$ are not disjoint, but are independent.

Task 3 – Balls from an Urn – Take 2

Say an urn contains three red balls and four blue balls. Imagine we draw three balls without replacement. (You can assume every ball is uniformly selected among those remaining in the urn.)

a) What is the probability that all three balls are all of the same color?

The experiment is modeled with $\Omega = r, b^3$. Probabilities are assigned as we have seen in class, by assuming every draw is uniform among the remaining balls. Then, note that $\mathbb{P}(rrr) = 3/7 \cdot 2/6 \cdot 1/5 = 1/35$ and $\mathbb{P}(bbb) = 4/7 \cdot 3/6 \cdot 2/5 = 4/35$. Therefore, the probability that they all have the same color is $1/35 + 4/35 = 1/7$.

b) What is the probability that we get more than one red ball given the first ball is red?

Let $\mathcal{R}$ be the event that the first ball is red. It is not hard to see that $\mathbb{P}(\mathcal{R}) = \frac{3}{7}$. (This can be computed more explicitly.) We also consider the event $\mathcal{M}$ that we have more than one red ball. Let $\mathcal{M}$ be the event that more than one ball is red. We need to now compute the probability $\mathbb{P}(\mathcal{M} \cap \mathcal{R})$, but note that by the law of total probability

$$\mathbb{P}(\mathcal{M} \cap \mathcal{R}) = \mathbb{P}(\mathcal{R}) - \mathbb{P}(\mathcal{M}^c \cap \mathcal{R}) = 3/7 - \mathbb{P}(\mathcal{M}^c \cap \mathcal{R}) \ .$$

Note that $\mathcal{M}^c \cap \mathcal{R}$ is the event that the first ball is red, and both remaining balls are blue. In particular,

$$\mathbb{P}(\mathcal{M}^c \cap \mathcal{R}) = \frac{3}{7} \cdot \frac{4}{6} \cdot \frac{3}{5} = \frac{6}{35} \ .$$

Thus, $\mathbb{P}(\mathcal{M} \cap \mathcal{R}) = 3/7 - 6/35 = 9/35$, and

$$\mathbb{P}(\mathcal{M}|\mathcal{R}) = \frac{\mathbb{P}(\mathcal{M} \cap \mathcal{R})}{\mathbb{P}(\mathcal{R})} = \frac{9/35}{3/7} = \frac{3}{5} \ .$$

Task 4 – Game Show

Corrupted by their power, the judges running the popular game show America's Next Top Mathematician have been taking bribes from many of the contestants. During each of two episodes, a given contestant is either allowed to stay on the show or is kicked off. If the contestant has been bribing the judges, they will be allowed to stay with probability 1. If the contestant has not been bribing the judges, they will be allowed to stay with probability 1/3, independent of what happens in earlier episodes. Suppose that 1/4 of the contestants have been bribing the judges. The same contestants bribe the judges in both rounds.

a) If you pick a random contestant, what is the probability that they are allowed to stay during the first episode?

Let S_i be the event they she stayed during the i-th episode. By the Law of Total Probability conditioning on whether the contestant bribed the judges we get,

$$\mathbb{P}(S_1) = \mathbb{P}(\text{Bribe})\,\mathbb{P}(S_1 \mid \text{Bribe}) + \mathbb{P}(\text{No bribe})\,\mathbb{P}(S_1 \mid \text{No bribe}) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} = \boxed{\frac{1}{2}}$$

b) If you pick a random contestant, what is the probability they are allowed to stay during both episodes?

Let S_i be defined as before. Staying during both episodes is equivalent to the contestant staying in episodes 1 and 2, so the event $S_1 \cap S_2$. By the Law of Total Probability, we get:

$$\mathbb{P}(S_1 \cap S_2) = \mathbb{P}(\text{Bribe})\,\mathbb{P}(S_1 \cap S_2 \mid \text{Bribe}) + \mathbb{P}(\text{No bribe})\,\mathbb{P}(S_1 \cap S_2 \mid \text{No bribe}) \tag{1}$$

We know a contestant is guaranteed to stay on the show, given that they are bribing the judges, hence:

$$\mathbb{P}(S_1 \cap S_2 \mid \text{Bribe}) = 1$$

On the other hand, if they have not been bribing judges, then the probability they stay on the show is 1/3, independent of what happens on earlier episodes. By conditional independence, we have:

$$Pr(S_1 \cap S_2 \mid \text{No bribe}) = Pr(S_1 \mid \text{No bribe})Pr(S_2 \mid \text{No bribe}) = \frac{1}{3} \cdot \frac{1}{3}$$

Plugging our results above into equation (1) gives us:

$$\mathbb{P}(S_1 \cap S_2) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{3} \cdot \frac{1}{3} = \boxed{\frac{1}{3}}$$

c) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that they get kicked off during the second episode?

By the definition of conditional probability and the Law of Total Probability,

$$\mathbb{P}(\overline{S_2} \mid S_1) = \frac{\mathbb{P}(S_1 \cap \overline{S_2})}{\mathbb{P}(S_1)} = \frac{\mathbb{P}(S_1 \cap \overline{S_2} \mid \text{Bribe})\mathbb{P}(\text{Bribe}) + \mathbb{P}(S_1 \cap \overline{S_2} \mid \text{No bribe})\mathbb{P}(\text{No bribe})}{\mathbb{P}(S_1)}$$

We have already computed $P(S_1)$ in part (a). We compute the numerator term by term. Given that a contestant is bribing the judges, they are guaranteed to stay on the show. As such:

$$\mathbb{P}(S_1 \cap \overline{S_2} \mid \text{Bribe}) = \mathbb{P}(S_1 \mid \text{Bribe}) \cdot \mathbb{P}(\overline{S_2} \mid \text{Bribe}) = 1 \cdot 0 = 0$$

On the other hand, if they have not been bribing judges, the probability they leave the show is $2/3$ (by complementing). We can then write:

$$\mathbb{P}(S_1 \cap \overline{S_2} \mid \text{No bribe}) = \mathbb{P}(S_1 \mid \text{No bribe}) \cdot \mathbb{P}(\overline{S_2} \mid \text{No bribe}) = \frac{1}{3} \cdot \frac{2}{3}$$

We can now evaluate our initial expression:

$$\mathbb{P}(\overline{S_2} \mid S_1) = \frac{0 \cdot \frac{1}{4} + \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{3}{4}}{\frac{1}{2}} = \frac{1/6}{1/2} = \boxed{\frac{1}{3}}$$

d) If you pick a random contestant who was allowed to stay during the first episode, what is the probability that they were bribing the judges?

Let B be the event that they bribed the judges. By Bayes' Theorem,

$$\mathbb{P}(B \mid S_1) = \frac{\mathbb{P}(S_1 \mid B)\mathbb{P}(B)}{\mathbb{P}(S_1)} = \frac{1 \cdot \frac{1}{4}}{\frac{1}{2}} = \boxed{\frac{1}{2}}$$

Task 5 – Allergy Season

In a certain population, everyone is equally susceptible to colds. Each person, in particular, catches a cold with probability 0.2.

The number of colds suffered by each person during each winter season ranges from 0 to 4, with probability 0.2 for each value (see table below). A new cold prevention drug is introduced that, for people for whom the drug is effective, changes the probabilities as shown in the table. Unfortunately, the effects of the drug last only the duration of one winter season, and the drug is only effective in 20% of people, independently.

number of colds	no drug or ineffective	drug effective
0	0.2	0.4
1	0.2	0.3
2	0.2	0.2
3	0.2	0.1
4	0.2	0.0

a) Sneezy decides to take the drug. Given that he gets 1 cold that winter, what is the probability that the drug is effective for Sneezy?

Let E be the event that the drug is effective for Sneezy, and C_i be the event that he gets i colds the first winter. By Bayes' Theorem,

$$\mathbb{P}(E \mid C_1) = \frac{\mathbb{P}(C_1 \mid E)\mathbb{P}(E)}{\mathbb{P}(C_1 \mid E)\mathbb{P}(E) + \mathbb{P}(C_1 \mid \overline{E})\mathbb{P}(\overline{E})} = \frac{0.3 \times 0.2}{0.3 \times 0.2 + 0.2 \times 0.8} = \frac{3}{11}$$

b) The next year he takes the drug again. Given that he gets 2 colds in this winter, what is the updated probability that the drug is effective for Sneezy?

Here, we need to consider the idea that Sneezy's drug effectiveness is no longer 20% like the general population. We will need to reduce our sample to the event of one occurrence of cold in previous season, and effectiveness of the drug given the one occurrence.Let the reduced sample space for part (b) be C_1 from part (a), so that $\mathbb{P}_{C_1}(E) = \mathbb{P}_\Omega(E|C_1)$. Let D_i be the event that he gets i colds the second winter. By Bayes' Theorem,

$$\mathbb{P}(E \mid D_2) = \frac{\mathbb{P}(D_2 \mid E)\mathbb{P}(E)}{\mathbb{P}(D_2 \mid E)\mathbb{P}(E) + \mathbb{P}(D_2 \mid \overline{E})\mathbb{P}(\overline{E})} = \frac{0.2 \times \frac{3}{11}}{0.2 \times \frac{3}{11} + 0.2 \times \frac{8}{11}} = \frac{3}{11}$$

c) Why is the answer to (b) the same as the answer to (a)?

The probability of two colds whether or not the drug was effective is the same. Hence knowing that Sneezy got two colds does not change the probability of the drug's effectiveness.

Task 6 – Coins

There are three coins, C_1, C_2, and C_3. The probability of "heads" is 1 for C_1, 0 for C_2, and p for C_3. A coin is picked among these three uniformly at random, and then flipped a certain number of times.

a) What is the probability that the first n flips are tails?

We have

$$1/3 \cdot 0 + 1/3 \cdot 1 + 1/3 \cdot (1-p)^n = \frac{1}{3} + \frac{1}{3}(1-p)^n \ .$$

b) Given that the first n flips were tails, what is the probability that C_1 was flipped / C_2 was flipped / C_3 was flipped?

We use Bayes Rule, and obtain

$$\mathbb{P}\left(C_1 \mid n \text{ tails}\right) = \frac{1/3 \cdot 0}{1/3 + 1/3(1-p)^n} = 0$$

$$\mathbb{P}\left(C_2 \mid n \text{ tails}\right) = \frac{1/3 \cdot 1}{1/3 + 1/3(1-p)^n} = \frac{1}{1 + (1-p)^n}$$

$$\mathbb{P}\left(C_3 \mid n \text{ tails}\right) = \frac{1/3 \cdot (1-p)^n}{1/3 + 1/3(1-p)^n} = \frac{(1-p)^n}{1 + (1-p)^n}$$

Task 7 – Parallel Systems

A parallel system functions whenever at least one of its components works. Consider a parallel system of n components and suppose that each component works with probability p independently.

a) What is the probability the system is functioning?

Let C_i be the event component i is working, and F be the event that the system is functioning.
For the system to function, it is sufficient for any component to be working. This means that the only case in which the system does not function is when none of the components work. We can then use complementing to compute $\mathbb{P}(F)$, knowing that $\mathbb{P}(C_i) = p$. We get:

$$\mathbb{P}(F) = 1 - \mathbb{P}(F^C) = 1 - \mathbb{P}(\bigcap_{i=1}^{n} C_i^C) = 1 - \prod_{i=1}^{n} \mathbb{P}(C_i^C) = 1 - \prod_{i=1}^{n} (1 - \mathbb{P}(C_i)) = 1 - \prod_{i=1}^{n} (1-p) = \boxed{1 - (1-p)^n}$$

Note that $\mathbb{P}(\bigcap_{i=1}^n C_i^C) = \prod_{i=1}^n \mathbb{P}(C_i^C)$ due to independence of C_i (components working independently of each other). Note also that $\prod_{i=1}^n a = a^n$ for any constant a.

b) If the system is functioning, what is the probability that component 1 is working?

We know that for the system to function only one component needs to be working, so for all i, we have $\mathbb{P}(F \mid C_i) = 1$. Using Bayes Theorem, we get:

$$\mathbb{P}(C_1|F) = \frac{\mathbb{P}(F|C_1)\mathbb{P}(C_1)}{\mathbb{P}(F)} = \frac{1 \cdot p}{1-(1-p)^n} = \boxed{\frac{p}{1-(1-p)^n}}$$

c) If the system is functioning and component 2 is working, what is the probability that component 1 is working?

$$\mathbb{P}(C_1|C_2, F) = \mathbb{P}(C_1|C_2) = \mathbb{P}(C_1) = p$$

where the first equality holds because knowing C_2 and F is just as good as knowing C_2 (since if C_2 happens, F does too), and the second equality holds because the components working are independent of each other.

More formally, we can use the definition of conditional probability along with a careful application of the chain rule to get the same result. We start with the following expression:

$$\mathbb{P}(C_1 \mid C_2, F) = \frac{\mathbb{P}(C_1, C_2, F)}{\mathbb{P}(C_2, F)} = \frac{\mathbb{P}(F \mid C_1, C_2) \cdot P(C_1 \mid C_2)\mathbb{P}(C_2)}{\mathbb{P}(F \mid C_2) \cdot \mathbb{P}(C_2)}$$

We note that the system is guaranteed to work if any one component is working, so $\mathbb{P}(F \mid C_1, C_2) = \mathbb{P}(F|C_2) = 1$. We also note that components work independently of each other, hence $\mathbb{P}(C_1|C_2) = \mathbb{P}(C_1)$. With that in mind, we can rewrite our expression so that:

$$\mathbb{P}(C_1 \mid C_2, F) = \frac{1 \cdot \mathbb{P}(C_1) \cdot \mathbb{P}(C_2)}{1 \cdot \mathbb{P}(C_2)} = \mathbb{P}(C_1) = \boxed{p}$$

Task 8 – Marbles in Pockets

A girl has 5 blue and 3 white marbles in her left pocket, and 4 blue and 4 white marbles in her right pocket. If she transfers a randomly chosen marble from left pocket to right pocket without looking, and then draws a randomly chosen marble from her right pocket, what is the probability that it is blue?

Let $W_{\_}, B_{\_}$ denote the event that we choose a white marble or a blue marble respectively, with subscripts L, R indicating from which pocket we are picking – left and right, respectively.
We know that we will pick from the left pocket first, and right pocket second. We can then use the Law of Total Probability conditioning on the color of the transferred marble so that:

$$\mathbb{P}(B_R) = \mathbb{P}(W_L) \cdot \mathbb{P}(B_R|W_L) + \mathbb{P}(B_L) \cdot \mathbb{P}(B_R|B_L) = \frac{3}{8} \cdot \frac{4}{9} + \frac{5}{8} \cdot \frac{5}{9} = \boxed{\frac{37}{72}}$$

Task 9 – A game

Pemi and Shreya are playing the following game: A 6-sided die is thrown and each time it's thrown, regardless of the history, it is equally likely to show any of the six numbers.

- If it shows 5, Pemi wins.
- If it shows 1, 2, or 6, Shreya wins.
- Otherwise, they play a second round and so on.

What is the probability that Shreya wins on the 4th round?

Let S_i be the event that Shreya wins on the i-th round and let N_i be the event that nobody wins on the i-th round. Then we are interested in the event

$$N_1 \cap N_2 \cap N_3 \cap S_4.$$

Using the chain rule, we have

$$\begin{aligned}\mathbb{P}(N_1, N_2, N_3, S_4) &= \mathbb{P}(N_1) \cdot \mathbb{P}(N_2|N_1) \cdot \mathbb{P}(N_3|N_2, N_3) \cdot \mathbb{P}(S_4|N_1, N_2, N_3) \\ &= \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{2}.\end{aligned}$$

In the final step, we used the fact that if the game hasn't ended, then the probability that it continues for another round is the probability that the die comes up 3 or 4, which has probability $1/3$.

Task 10 – Another game

Leiyi and Luxi are playing a tournament in which they stop as soon as one of them wins n games. Luxi wins each game with probability p and Leiyi wins with probability $1-p$, independently of other games. What is the probability that Luxi wins and that when the match is over, Leiyi has won k games?

Since the match is over when someone wins the n^{th} game, and Luxi won the match, Luxi won the last game.

Before this, Luxi must've won $n-1$ games and Leiyi must've won k games. Therefore, the probability that we reach a point in time when Luxi has won $n-1$ games and Leiyi has won k games is: $p^{n-1} \cdot (1-p)^k \cdot \binom{n-1+k}{k}$. The binomial coefficient counts the number of ways of picking the k games that Leiyi has won out of $n-1+k$ games.

At that point in time, we want Luxi to win the next game so that she has won n games. This happens with probability p, independent of previous outcomes. Therefore, our final probability is:

$$p^{n-1} \cdot (1-p)^k \cdot \binom{n-1+k}{k} \cdot p = p^n \cdot (1-p)^k \cdot \binom{n-1+k}{k}$$

Task 11 – Random Variables

(The material for this problem will most likely be covered early next week.)

Assume that we roll a fair 3-sided die three times. Here, the sides have values $1, 2, 3$.

a) Describe the PMF of the random variable X giving the sum of the first two rolls.

We have $\mathbb{P}(X=2) = 1/9$, $\mathbb{P}(X=3) = 2/9$, $\mathbb{P}(X=4) = 3/9$, $\mathbb{P}(X=5) = 2/9$, and $\mathbb{P}(X=6) = 1/9$.

b) Give the expectation $\mathbb{E}[X]$.

We give a direct proof here, and note that

$$\mathbb{E}[X] = 1/9 \cdot (2+6) + 2/9 \cdot (3+5) + 3/9 \cdot 4 = (8+16+12)/9 = 4\ .$$

c) Compute $\mathbb{P}(X > 3)$.

$\mathbb{P}(X > 3) = 3/9 + 2/9 + 1/9 = 6/9 = 2/3.$

d) Let Y be the random variable describing the sum of the three rolls. Compute $\mathbb{P}(X = 5 \mid Y = 7)$.

First, $\mathbb{P}(X = 5 \mid Y = 7) = \mathbb{P}(X = 5, Y = 7) / \mathbb{P}(Y = 7)$. Then, $\mathbb{P}(X = 5, Y = 7) = 2/27$, whereas

$$\mathbb{P}(Y = 7) = 3/27 + 2/27 + 1/27 = 2/9\ .$$