1. (20 pts.) Random variables, sample space and events
Consider the random experiment of flipping a coin 4 times.

1. (2 pts.) Define the appropriate random variables.
   Answer: The random variables are
   \[ X_i = \text{value of } i^{th} \text{ coin flip} \quad i \in \{1, 2, 3, 4\} \]

2. (2 pts.) List all the outcomes in the sample space \(\Omega\). How many are there?
   Answer: There are 16:
   \[
   \begin{align*}
   &HHHH \ HTHH \ THHH \ TTHH \\
   &HHHT \ HTHT \ THHT \ TTHT \\
   &HHTH \ HTTH \ THTH \ TTTH \\
   &HHTT \ HTTT \ THTT \ TTTT \\
   \end{align*}
   \]

3. (2 pts.) Let \(A\) be the event that the first flip is a Heads. List all the outcomes in \(A\). How many are there?
   Answer: This is the event \(X_1 = H\). There are 8 outcomes in this event:
   \[
   \begin{align*}
   &HHHH \ HTHH \\
   &HHHT \ HTHT \\
   &HHTH \ HTTH \\
   &HHTT \ HTTT \\
   \end{align*}
   \]

4. (2 pts.) Let \(B\) be the event that the third flip is a Heads. List all the outcomes in \(B\). How many are there?
   Answer: This is the event \(X_3 = H\). There are 8 outcomes in this event:
   \[
   \begin{align*}
   &HHHH \ HTHH \ THHH \ TTHH \\
   &HHHT \ HTHT \ THHT \ TTHT \\
   \end{align*}
   \]

5. (3 pts.) Let \(C\) be the event that the first flip and the third flip are both Heads. List all the outcomes in \(C\). How many are there?
   Answer: This is the event that \(X_1 = H\) and \(X_3 = H\). There are 4 outcomes in this event:
   \[
   \begin{align*}
   &HHHH \ HTHH \ THHH \ TTTH \\
   &HHHT \ HTHT \ THHT \ TTHT \\
   \end{align*}
   \]
6. **(3 pts.)** Let $D$ be the event that the first flip or the third flip is a Heads. List all the outcomes in $D$. How many are there?

**Answer:** This is the event that $X_1 = H$ or $X_3 = H$. There are 12 outcomes in this event:

- HHHH
- HTHH
- THHH
- TTHH
- HHTH
- HTTH
- THHT
- TTHT
- HHTT
- HTTT

7. **(3 pts.)** Are the events $A$ and $B$ disjoint? Express the event $C$ in terms of $A$ and $B$. Express the event $D$ in terms of $A$ and $B$.

**Answer:** No, $A$ and $B$ are not disjoint. The relationships are:

- $C = A \cap B$
- $D = A \cup B$

**Comment:** Even though the event $A$ is about the first flip and the event $B$ is about the third flip, the two events are not disjoint at all! This is a very common mistake and should be avoided at all cost!

8. **(3 pts.)** Suppose now the coin is flipped $n \geq 3$ times instead of 4 flips. Define the random variables, compute $|\Omega|, |A|, |B|, |C|, |D|$.

**Answer:**

The random variables are

- $X_i =$ value of $i^{th}$ coin flip $\quad i \in \{1, \ldots, n\}$

- $|\Omega| = 2^n$
- $|A| = 2^{n-1}$
- $|B| = 2^{n-1}$
- $|C| = 2^{n-2}$
- $|D| = 2^n - 2^{n-2}$

For $|D|$, note that a outcome is not in $D$ if the first and third flips are both Tails. The number of such outcomes is $2^{n-2}$. Subtracting this number from the total number of outcomes ($2^n$) yields the answer above.

2. **(20 pts.) More Probability Models**

Suppose you have two coins, one is biased with a probability of $p$ coming up Heads, and one is biased with a probability of $q$ coming up Heads. Answer the questions below, but you don’t need to provide justifications.
1. Suppose \( p = 1 \) and \( q = 0 \).

(a) (5 pts.) You pick one of the two coins randomly and flip it. You repeat
this process \( n \) times, each time randomly picking one of the two coins and
then flipping it. Let \( X_i \) be the result of the \( i^{th} \) flip for \( i = 1, \ldots, n \). Describe
the sample space and give a reasonable probability assignment to model the
situation.

Answer: If \( \omega \) is an outcome, i.e., a sequence of Heads and Tails of length
\( n \), then

\[ P(\omega) = \frac{1}{2^n}. \]

Explanation: Whenever we pick the coin biased with \( p = 1 \), we always
get Heads. Similarly, when we pick the coin biased with \( q = 0 \), we always
get Tails. For a given flip, we are equally likely to use each coin, so the flip
is equally likely to be Heads or Tails. Moreover, since we randomly pick
the coin for each flip, all sequences are equally likely. Thus, we model all
sequences as having the same probability:

\[ P(\omega) = \frac{1}{|\Omega|} = \frac{1}{2^n}. \]

Comment: For the curious, here is a generalization. Suppose the proba-
bility of picking the first coin is \( r \) and the probability of picking the second
coin is \( 1 - r \). Then, the probability of getting Heads on any given flip is \( r \),
so the probability of a sequence depends on the number of Heads and Tails.
In particular, if \( \omega \) is a sequence with \( k \) Heads and \( n - k \) Tails,

\[ P(\omega) = r^k(1 - r)^{n-k}. \]

(The problem didn’t ask you to solve this generalization.)

(b) (5 pts.) Now you pick one of the two coins randomly, but flip the same
coin \( n \) times. Again let \( X_i \) be the result of the result of the \( i^{th} \) flip. Identify
the sample space for this experiment together with a reasonable probability
assignment to model the situation. Is your answer the same as in the
previous part?

Answer: No, it is different. The sample space includes only two outcomes,
all Heads (\( HHH \ldots H \)) and all Tails (\( TTT \ldots T \)). The probabilities are

\[ P(HHH \ldots H) = P(TTT \ldots T) = \frac{1}{2}. \]

Equivalently, we can say that the sample space includes all possible se-
quences of \( n \) Heads and Tails, but the probability of a sequence is 0 if it is
not \( HHH \ldots H \) or \( TTT \ldots T \).
Explanation: In this case, only two sequences are possible:

\[ HHH \ldots H \quad \text{or} \quad TTT \ldots T, \]

i.e., we either get all Heads (if we choose the coin biased with probability \( p = 1 \)) or we get all Tails (if we choose the coin biased with probability \( q = 0 \)). Since we are equally likely to pick either coin, we model this as

\[ P(HHH \ldots H) = P(TTT \ldots T) = \frac{1}{|Ω|} = \frac{1}{2}. \]

Comment: This answer can also be generalized, if the probability of picking the first coin is \( r \). Then we get the following probability assignment:

\[ P(HHH \ldots H) = r \quad P(TTT \ldots T) = 1 - r. \]

2. Repeat the above two questions for arbitrary values of \( p \) and \( q \). Express your answers in terms of \( p \) and \( q \).

(a) (5 pts.) Answer: If \( \omega \) is a sequence with \( k \) Heads and \( n - k \) Tails, then

\[
P(\omega) = \left( \frac{1}{2}p + \frac{1}{2}q \right)^k \times \left( 1 - \left( \frac{1}{2}p + \frac{1}{2}q \right) \right)^{n-k}.
\]

Explanation: Assume each coin is chosen with probability \( \frac{1}{2} \) and consider a single flip. Since the probability of getting Heads with the first coin is \( p \) and the probability of getting heads with the second coin is \( q \), the probability of getting Heads on this flip is \( \frac{1}{2}p + \frac{1}{2}q \). Similarly, the probability of getting Tails is \( 1 - \left( \frac{1}{2}p + \frac{1}{2}q \right) \).

We can then construct the probability of a sequence by multiplying the probabilities for each flip. Thus, the probability of getting a sequence with \( k \) heads and \( n - k \) tails is the probability of getting heads, to the \( k \)th power, times the probability of getting tails, to the \( n - k \)th power (the probability for each toss gets multiplied due to independence).

To provide a more rigorous explanation of the probability of Heads in the first flip (this is not required for a full mark solution), denote \( R \) as the event that the coin with bias \( p \) is chosen and \( \bar{R} \) as the event that the coin with bias \( q \) is chosen. Since \( R \) and \( \bar{R} \) are disjoint and \( R \cup \bar{R} = Ω \) (i.e., \( R \) and \( \bar{R} \) form a partition of all possible samples), the probability of getting Heads is

\[
P(H) = P(H \cap R) + P(H \cap \bar{R})
= P(H|R)P(R) + P(H|\bar{R})P(\bar{R})
= P(H|R)\frac{1}{2} + P(H|\bar{R})\frac{1}{2}
= \frac{1}{2}p + \frac{1}{2}q.
\]
Since we get Tails whenever we don’t get Heads, the probability of Tails is

\[ 1 - \left( \frac{1}{2}p + \frac{1}{2}q \right) = \frac{1}{2}(1-p) + \frac{1}{2}(1-q) . \]

**Comment:** A generalization: If the probability of getting the coin with bias \( p \) is \( r \), the probability of getting Heads is \( r \cdot p + (1-r) \cdot q \), and thus the probability of a sequence \( \omega \) with \( k \) heads and \( n-k \) tails is

\[ P(\omega) = (r \cdot p + (1-r) \cdot q)^k \times (1 - (r \cdot p + (1-r) \cdot q))^{n-k} . \]

(b) **(5 pts.) Answer:** If \( \omega \) is a sequence with \( k \) Heads and \( n-k \) Tails, then

\[ P(\omega) = \frac{1}{2}p^k(1-p)^{n-k} + \frac{1}{2}q^k(1-q)^{n-k} . \]

**Explanation:** If we only used the coin with bias \( p \),

\[ P(\omega) = p^k(1-p)^{n-k} . \]

Likewise, if we only used the coin with bias \( q \), we would have

\[ P(\omega) = q^k(1-q)^{n-k} . \]

Since we pick the coin before flipping the sequence, assuming each coin is chosen with probability \( \frac{1}{2} \), we get

\[ P(\omega) = \frac{1}{2}p^k(1-p)^{n-k} + \frac{1}{2}q^k(1-q)^{n-k} . \]

More formally (this is not required for a full mark solution), let \( \omega \) be a sequence of length \( n \) and let \( h(\omega) \) be the number of Heads in the sequence (thus, \( n - h(\omega) \) is the number of Tails). As before, let \( R \) be the event that we use the coin with bias \( p \) and \( \bar{R} \) be the event that we use the coin with bias \( q \). Then

\[ P(\omega|R) = p^{h(\omega)}(1-p)^{n-h(\omega)} \]
\[ P(\omega|R^c) = q^{h(\omega)}(1-q)^{n-h(\omega)} \]

and we have

\[ P(\omega) = P(\omega \cap R) + P(\omega \cap R^c) \]
\[ = P(\omega|R)P(R) + P(\omega|R^c)P(R^c) \]
\[ = P(\omega|R)\frac{1}{2} + P(\omega|R^c)\frac{1}{2} \]
\[ = \frac{1}{2}p^{h(\omega)}(1-p)^{n-h(\omega)} + \frac{1}{2}q^{h(\omega)}(1-q)^{n-h(\omega)} . \]
Comment: Generalizing to the case where we pick the first coin with probability $r$, we get

$$P(\omega) = r \cdot p^h(\omega)(1 - p)^{n-h(\omega)} + (1 - r) \cdot q^h(\omega)(1 - q)^{n-h(\omega)}.$$ 

3. (20 pts.) DNA Sequencing

The DNA of a bacteria is a length 1 million circular sequence containing the nucleotides A,G,C, T. We sequence the DNA by sampling 100,000 reads of length 100 randomly from the DNA. We are interested in particular in a specific gene, which is a specific length 1000 subsequence of the DNA.

1. (6 pts.) What is the probability that the first two reads overlap?

**Answer:** The random variables are

$$X_i \in \{1, \ldots, 1,000,000\} = \text{starting location of } i\text{-th read } i = 1, \ldots, 100,000.$$ 

Each outcome of $(X_1, \ldots, X_{100,000})$ is equally probable (they all have probability $1,000,000^{-100,000}$). The event where the first two reads overlap is given by "$|X_1 - X_2| \leq 100 - 1$ or $(X_1 + 1,000,000) - X_2) \leq 100 - 1$ or $|X_1 - (X_2 + 1,000,000)| \leq 100 - 1$". Note that since each read has length 100, two reads overlap if their locations are within a distance 99. Since the sequence is circular, we also measure the distance between $X_1 + 1,000,000$ and $X_2$ (also $X_1$ and $X_2 + 1,000,000$).

An easier way to calculate the probability is to assume $X_1 = 1$ by symmetry. Then the event becomes "$X_2 \leq 100$ or $X_2 \geq 1,000,000 - 98$". The probability is

$$\frac{100 + (1,000,000 - (1,000,000 - 98) + 1)}{1,000,000} = \frac{199}{1,000,000}.$$ 

2. (7 pts.) What is the probability that a specific nucleotide of the DNA is covered by at least one read?

**Answer:** By symmetry, we assume the nucleotide of interest is at location 100 (hence it is covered by reads at locations 1, \ldots, 100). The event is "$X_i \leq 100$ for some $i$". We first compute the probability of its complement "$X_i > 100$ for all $i$". Now the probability that "$X_i > 100$" for a fixed $i$ is $(1,000,000-100)/1,000,000$. By independence, the probability of "$X_i > 100$ for all $i$" is

$$\frac{(1,000,000 - 100)^{100,000}}{1,000,000^{100,000}} = \left(\frac{9999}{10000}\right)^{100,000}.$$ 

The probability of "$X_i \leq 100$ for some $i$" is

$$1 - \left(\frac{9999}{10000}\right)^{100,000}.$$
Comment: Using the approximation \((1 - 1/x)^x \approx e^{-1}\) for large \(x\), we have

\[
1 - \left( \frac{9999}{10000} \right)^{100,000} = 1 - \left( \left( 1 - \frac{1}{10000} \right)^{10000} \right)^{10} \approx 1 - e^{-10} \approx 0.9999546.
\]

3. (7 pts.) What is the probability that there is at least one read that is contained in the gene of interest?

Answer: By symmetry, we assume the gene of interest is at location 1, . . . , 1000 (hence reads at locations 1, . . . , 901 are contained in it). The event is \(X_i \leq 901\) for some \(i\). By a similar calculation as the previous part, the probability is

\[
1 - \left( \frac{1,000,000 - 901}{1,000,000} \right)^{100,000}.
\]

Comment: Using the approximation \((1 - 1/x)^x \approx e^{-1}\) for large \(x\), we have

\[
1 - \left( \frac{1,000,000 - 901}{1,000,000} \right)^{100,000} \approx 1 - e^{-90.1},
\]

which is extremely close to 1.

4. (20 pts.) A new game

You have two quarters and a table with a row of squares marked like this:

<table>
<thead>
<tr>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
</table>

Before the game begins, you get to place each quarter on one square. You can put either both quarters on the same square, or you can put them on two different squares: your choice.

Then, you roll two fair dice, sum up the numbers showing on the dice to get a number from 2–12, and if there’s a quarter on the square labelled with that number, remove it from the table. (If there are two quarters on that square, remove only one of them.)

Now roll the two fair dice a second time, again getting a number from 2–12, and again removing a single quarter from the square with that number, if there’s a quarter there.

At this point, the game is over. If you removed both quarters, you win; if any quarter remains on the table, you lose.

1. (8 pts.) What’s the probability of winning, if you put two quarters on the square labelled 5?
Answer: The basic random variables are \( X_1, X_2 \), the numbers on the two dice in the first roll, and \( Y_1, Y_2 \) the numbers on the two dice on the second roll. The sample space is

\[
\Omega = \{(i, j), (k, \ell) : 1 \leq i, j, k, \ell \leq 6\},
\]

All outcomes are equally likely, so each outcome has probability \( \frac{1}{6^4} \).

If we place both quarters on 5, we win if \( X_1 + X_2 = 5 \) and \( Y_1 + Y_2 = 5 \).

\( X_1 + X_2 = 5 \) if we roll \( (1, 4), (2, 3), (3, 2), \) or \( (4, 1) \) in the first roll. So, there are \( 4 \times 4 = 16 \) outcomes in \( \Omega \) in which we can roll two dice so that they sum up to 5 in both of the two rolls. Since each of these outcomes have probability \( \frac{1}{6^4} \), the probability of winning if we place both coins on 5 is:

\[
\frac{16}{6^4} = \frac{1}{81}.
\]

2. (12 pts.) What’s your best strategy? In other words, what’s the best place to put your two quarters, if you want to maximize the probability of winning? State where you should put your two quarters. Then, calculate the probability that you win, if you put your two quarters there.

Answer: There are two ways to place the coins if we want to maximize the probability of winning. We can either place the coins on locations 6 and 7, or on locations 7 and 8 (either gives the same chance of winning, and both are optimal). We will calculate the probability that we win by placing the coins on 6 and 7.

The random variables and the sample space is as in part 1. If we place the coins on locations 6 and 7, we have two ways to win:

- We win if \( X_1 + X_2 = 6 \) and \( Y_1 + Y_2 = 7 \).
- We win if \( X_1 + X_2 = 7 \) and \( Y_1 + Y_2 = 6 \).

These two ways are disjoint, and by symmetry the probability of these two events are the same. Hence, probability of winning by placing the coins on 6 and 7 is just twice the probability of the event that \( X_1 + X_2 = 6 \) and \( Y_1 + Y_2 = 7 \).

We will calculate the probability of this later event. Let’s call it \( E \).

There are 6 ways to roll two dice so that they add up to 7, and there are 5 ways to roll two dice so that they add up to 6. Therefore the number of outcomes in \( E \) is \( 6 \times 5 = 30 \). Hence, the probability of \( E \) is \( \frac{30}{6^4} = \frac{5}{216} \).

Hence, the probability of winning is \( \frac{5}{108} \).

The analysis for placing the coins on locations 7 and 8 is similar, and the probability of winning in that case is the same.

Comment: Note that if we place both coins on 7, then we need both rolls of two dice to sum up to 7. The probability of winning in this case is \( \frac{5}{216} \times \frac{5}{216} \), which
is less than the probability of winning if we place the coins on locations 6 and 7 (or on 7 and 8).

What happened? It’s tempting to think that the best strategy is to place both coins on 7, since 7 is the most likely sum to get after rolling two dice. However, surprisingly, that isn’t the best possible strategy. In effect, placing the coins on locations 6 and 7 gives us almost twice as many ways to win, which more than compensates for the slight reduction in the probability of rolling a 6 (compared to the probability of rolling a 7).

5. (20 pts.) Monty Hall Again

1. (2 pts.) In the original Monty Hall problem, there were only two strategies - the sticking strategy where you would stick with your original pick after the host opens the first door, or the switching strategy, where you would change your original pick after Monty opens the first door. Now, there are two stages where a decision needs to be made, and at each stage there are two choices. Either you can stick with your current door, or you can switch to another door (at stage one where there are two possible doors to choose from, choose one of them randomly). So, in all, there are $2 \cdot 2 = 4$ different strategies.

2. (18 pts.) Following the suggestion from the lecture notes, we will define the random variables and the sample space in a similar way as given in the lecture notes. This time we will define $X$ as the prize door, $Y_1$ is the door that the contestant initially picks, $Z_1$ is the door that is opened by the contestant in the first round, $Y_2$ is the next door that the contestant picks (it could be the same as $Y_1$ if the contestant decides to stick with the original door), and $Z_2$ is the door that is opened in the second round. Each outcome in the sample space is a 5-tuple.

Knowing all these, we can decide if a particular strategy leads to winning or not. Once again we realize that there are certain tuples which are not possible in our probability space. Moreover, as the example below will show, the probability space itself depends on which of the four strategies the contestant decides to follow. Note that this is an important difference between the game with three doors and the one with four doors.

Let’s consider some events. First suppose $X = 1$ and $Y_1 = 2$. At this point, the only choice for $Z_1$ is either 3 or 4. Suppose that $Z_1 = 3$. Now, if the contestant has decided to use a stick-stick or stick-switch strategy, then we must have $Y_1 = Y_2 = 2$. However, if the contestant is using a switch-stick or switch-switch strategy, then $Y_2$ can be either 1 or 4.

Another point to realize is that we can heavily cut down on the number of possibilities we need to enumerate by considering the symmetry of the problem.
In particular, the probability that a certain strategy leads to winning will only depend on whether $Y_1 = X$ or $Y_1 \neq X$, and not on the actual values of $X$ and $Y_1$.

(a) **(4 pts.) Probability of winning by sticking, sticking.**

This is the easiest probability to calculate, since in this strategy the contestant never changes his/her choice of the door. Thus, this leads to a win if and only if $Y_1 = X$ i.e. the contestant chose the prize door to begin with. Since $Y_1 = X$ for 4 out of the 16 possible pairs of values these 2 random variables take on, and all pairs are equally likely, we get

$$P(\text{win by sticking, sticking}) = \frac{1}{4}$$

(b) **(4 pts.) Probability of winning by sticking, switching**

Consider what happens under this strategy, when $X = Y_1$. Since the contestant decided to stick in the first round, we must have $Y_2 = X$. However, now the contestant will switch in the second round and thus end at a door which is not equal to $i$. So this strategy will never lead to a win if $X = Y_1$. On the other hand, consider the case when $Y_1 \neq X$. Let us consider the case $X = 1$ and $Y_1 = 2$ for concreteness. In this case the host will open the door $Z_1 = 3$ or $Z_1 = 4$. Since the contestant decides to stick in the first round, we will have that $Y_2 = Y_1 = 2$. However, then the host will be forced to open the door only remaining door which does not contain a prize and is not chosen by the contestant i.e. if $Z_1 = 3$ then $Z_2 = 4$ and vice-versa. Now if the contestant switches in the second round, she will definitely win the prize, since the prize door is the only one remaining to switch to. Thus,

$$P(\text{win by sticking, switching}) = \Pr(X = Y_1) = \frac{3}{4}.$$ 

(c) **(4 pts.) Probability of winning by switching, sticking**

Under this strategy also, if we start out with a situation when $X = Y_1$, then the contestant can never win. This is because the contestant will first choose $Y_2 \neq Y_1$ (since he switches in the first round) and then sticks to it. However, if $X = Y_1$ then we will have $Y_2 \neq X$ and so she will stick to the wrong choice.

Consider now the case when $Y_1 \neq X$, again taking $X = 1$ and $Y_1 = 2$ for concreteness. $Z_1$ will again be either 3 or 4. This is now the moment of decision - the contestant will choose a value of $Y_2$ which he will later stick to. Since she is choosing randomly between the two unopened doors, one of which has the prize, she can win with probability $\frac{1}{2}$, if she started out in the case with $Y_1 \neq X$. Hence,
\[ P(\text{win by switching, sticking}) = \frac{1}{4} \cdot 0 + \frac{3}{4} \cdot \frac{1}{2} = \frac{3}{8} \]  

(1)

(d) (4 pts.) Probability of winning by switching, switching

As before, we break the analysis into two case. Consider the case when the contestant starts out with \( Y_1 = X \), say both equal to 1. There are three other doors which do not contain the prize. Now, \( Z_1 \) will be one of these doors (say \( Z_1 = 3 \)). Also, since the contestant switches, he will switch to one of these three doors other than \( Z_1 \) (say \( Y_2 = 4 \)), since \( Z_1 \) is already open. Now, the host will be forced to open the third non-prize door (i.e. \( Z_2 = 2 \)). The contestant will switch to only door left, which is 1, and hence will win in this case.

Now consider the case when \( Y_1 \neq X \), say \( X = 1 \) and \( Y_1 = 2 \). Then the host will open a door \( Z_1 \), which is neither of 1 or 2 (say \( Z_1 = 3 \)). Now, the contestant has a choice of two doors to switch to. One of these is the prize door 1 and the other is a non-prize door (4 in our example). If the contestant chooses the door 1 at this round, she will not win the game since she will switch again at the next round. On the other hand, if she switches to a non-prize door, then the host will be forced to open 2. Then, in the next round, the only choice for switching will be 1, which will result in a win. Hence, if the contestant starts with \( Y_1 \neq X \), then the probability of winning is \( \frac{1}{2} \) since she wins if and only if she switches to a non-prize door in the first round. Thus,

\[ P(\text{win by switching, switching}) = \frac{1}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{2} = \frac{5}{8} \]

(2 pts.) The best strategy is to stick in the first round and switch in the second round.