HW2 Solutions

1. (12 pts.) Conditional probability

(a) (6 pts.) I have a bag containing either a $1 or $5 bill (with probability 1/2 for each of these two possibilities). I then add a $1 bill to the bag, so it now contains two bills. The bag is shaken, and you randomly draw a bill from the bag (without looking). Suppose it turns out to be a $1 bill. If a second student draws the remaining bill from the bag, what is the chance that it too is a $1 bill? Show your calculations.

Answer 1: The sample space is
\[ \Omega = \{ S1, D1, D5 \} , \]
with the following interpretation. The outcome \( S1 \) represents the case where the bag initially contains two $1 bills (same denomination) and you draw a $1 bill out of the bag. The outcome \( D1 \) represents the case where the bag initially contains a $1 and a $5 bill (different denomination) and you draw a $1 bill out of the bag. Finally, \( D5 \) represents the case where the bag initially contains $1 and $5, and you draw a $5 bill. By the assumption that the bag originally contains either $1 or $5 with equal probability,
\[ P(S1) = \frac{1}{2} . \]
Since the probabilities add up to 1 and \( D1, D5 \) have the same probability,
\[ P(D1) = \frac{1}{4} \quad P(D5) = \frac{1}{4} . \]
Let \( A \) be the event that you draw out a $1 bill, and \( B \) be the event that the second student draws a $1 bill. The problem is to compute the conditional probability \( P(B|A) \). To do this, we will need to know \( P(A \cap B) \) and \( P(A) \).
The event \( A \) occurs for outcomes \( S1 \) and \( D1 \). Therefore,
\[ P(A) = P(S1) + P(D1) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4} . \]
The event \( A \cap B \) occurs only for outcome \( S1 \), so
\[ P(A \cap B) = P(S1) = \frac{1}{2} . \]
Now we can calculate
\[ P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{1/2}{3/4} = \frac{2}{3} . \]

Answer 2: We can calculate it yet another way. Let the events \( A, B \) be defined as above. Both students would draw $1 bills iff the first bill added was a $1 bill. Thus \( P(A \cap B) = 1/2 \). Event \( A \) occurs if either (a) the first bill added was $1 (with probability 1/2), and you draw a $1 (with probability 1), or (b) the first bill was $5 (with probability 1/2) and you draw the $1 (with probability 1/2). So
\[ P(A) = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = 3/4 . \]
(b) (6 pts.) Your buddy found a website online where he could buy trick coins that are heads or tails on both sides. He puts three coins into a bag: one coin that is heads on both sides, one coin that is tails on both sides, and one that is heads on one side and tails on the other side. You shake the bag, draw out a coin at random, put it on the table without looking at it, then look at the side that is showing. Suppose you notice that the side that is showing is heads. What is the probability that the other side is heads? Show your work.

Answer: An outcome consists of us selecting a coin, and then choosing one of its sides. Since there are 3 coins each with 2 sides, our sample space \( \Omega \) has 6 elements

\[
\Omega = \{ (HH, H_1), (HH, H_2), (HT, H), (HT, T), (TT, T_1), (TT, T_2) \},
\]

where \((HH, H_1)\) refers to the outcome of drawing the coin with two heads and of looking at the first side. Similarly, \((HH, H_2)\) refers to the outcome of choosing the coin with two heads and of looking at the second side, \((HT, T)\) refers to the outcome of choosing the coin with both heads and tails and of looking at the tails side, etc. Let \( A \) be the event that we choose the \(HH\) coin, and let \( B \) be the event that we see a heads when we put the coin down on the table. We wish to compute \( P(A|B) \). Since \( A = \{ (HH, H_1), (HH, H_2) \} \), \( B = \{ (HH, H_1), (HH, H_2), (HT, H) \} \), and \( A \cap B = A \), we see

\[
P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{|A|/|\Omega|}{|B|/|\Omega|} = \frac{2/6}{3/6} = \frac{2}{3}.
\]

Comment: This might be a bit surprising. For instance, it’s tempting to think: if I see a heads, then this means I drew either the \(HH\) or the \(HT\) coin, and both should be equally likely, so it’s 50-50 whether the other side is heads or tails. However, the calculation above shows that this argument is incorrect. The reason is that if we draw the \(HH\) coin, there are two ways to end up with it showing heads (according to which of the two sides is on top), whereas if we draw the \(HT\) coin, there is only one way to end up with it showing heads, so given that I see heads, it’s twice as likely that I drew the \(HH\) coin rather than the \(HT\) coin.

2. (20 pts.) Random Variables

(a) (4 pts.) Define the basic random variables and give the sample space and assign probabilities to the outcomes.

Answer: We have 4 random variables \( X_i \) which is 1 if the \(i\)th coin toss is heads and 0 otherwise. Our sample space is composed of all possible outcomes of the coin flips.

\[
\Omega = \{ TTTT, HTTT, THTT, \ldots, HHHT, HHHH \}
\]

As seen in lecture, the probability of getting \( HTTT \) and the probability of getting \( TTHT \) are exactly the same: \( p(1-p)^3 = (1-p)^2 \cdot p(1-p) \) and in general, the probability of getting a sequence with \( r \) heads out of the 4 tosses is \( p^r(1-p)^{4-r} \).

(b) (4 pts.) Let \( X \) be the total number of Heads in the four flips. List the outcomes in the five events \( X = i, i = 0, 1, 2, 3, 4 \). Is \( X \) a random variable?

Answer: Indeed, \( X \) is a random variable. It takes every sample point in our sample space, and assigns it a real value. Our sample space and outcomes are shown in figure 1.

(c) (4 pts.) Are the events \( X = 1 \) and \( X = 2 \) disjoint? Are they independent? What about the events \( X = 1 \) and \( X \leq 2 \)?
Figure 1: The sample space for $X$.

**Answer:** The events $X = 1$ and $X = 2$ are indeed disjoint (it is impossible that there will be a total of both 1 heads and 2 heads in one outcome) and therefore $P(X = 1 \cap X = 2) = P[\emptyset] = 0$. These events however are not independent since $P(X = 1) \cdot P(X = 2) \neq 0 = P(X = 1 \cap X = 2)$. The events $X = 1$ and $X \leq 2$ are not disjoint. The event $X = 1$ is a subset of the event $X \leq 2$, therefore their intersection is not empty and is the event $X = 1$. They are not independent since $P(X = 1) \cdot P(X = 2) \neq P(X = 1 \cap X \leq 2)$.

(d) **(4 pts.)** Let $Y$ be the first flip when a Heads appears and $Y = 0$ if there is no Heads in the four flips. List the outcomes in the five events $Y = i, i = 0, 1, 2, 3, 4$.

**Answer:** Recall that a random variable on a sample space is a function that assigns to each sample point in the sample space a real number. The function $Y$ is not defined for the event in which no heads come out at all, which is a valid point in our sample space. Our sample space and outcomes are shown in figure 2.

Figure 2: The sample space for $Y$.

(e) **(4 pts.)** Are the events $X = 3$ and $Y = 4$ disjoint? Are they independent? What about the events $X = 2$ and $Y = 2$?
Answer: The events \( X = 3 \) and \( Y = 4 \) are indeed disjoint, since if heads only appears in the last flip, the total number of heads in all four flips is exactly one. Since both events \( Y = 4 \) and \( X = 3 \) have some positive probability (i.e. not 0) and \( P(Y = 4)P(Y = 3) \neq 0 = P(X = 3 \cap Y = 4) \), and thus these events are not independent.

The events \( X = 2 \) and \( Y = 2 \) are not disjoint. This is because their are outcomes that are both in \( X = 2 \) and in \( Y = 2 \), that is the set \((X = 2) \cap (Y = 2)\) is not empty. For example, the outcome \( THHT \) respects both properties of the having it land on heads the first time on the second flip, and the total number heads being 2. These events are not independent. We have that \( P(X = 2) = (\frac{1}{2})p^2(1 - p)^2 \) and \( P(Y = 2) = (1 - p)p \), and therefore \( P(X = 2) \cdot P(Y = 2) = (\frac{1}{2})p^3(1 - p)^3 \). On the other hand, the intersection of these events, \((X = 2) \cap (Y = 2)\) is \{\(THTH, THHT\)\}, each with probability \( p^2(1 - p)^2 \), and therefore \( P(X = 2 \cap Y = 2) = 2 \cdot p^2(1 - p)^2 \). In general, \((\frac{1}{2})p^3(1 - p)^3 \neq p^2(1 - p)^2\).

3. (20 pts.) Balls and Bins

Consider an experiment involving \( m \) balls and \( n \) bins, where each ball is independently placed in one of the bins randomly. We can define two sets of random variables for the balls and bins problem:

1. \( X_i \) = index of the bin where the \( i \)th ball lands, \( i = 1, \ldots, m \)
2. \( Y_i \) = number of balls in bin \( i \), \( i = 1, \ldots, n \)

(a) (4 pts.) Do the \( X_i \)'s contain the same information about the system as the \( Y_i \)'s? In other words, can one compute the \( Y_i \)'s given the \( X_i \)'s and vice versa? If not, which set of random variables contain more information, and give an example of an event that can be expressed in terms of one set of random variables but not the other.

Answer: No they do not contain the same information. The \( X_i \)'s contain more information than the \( Y_i \)'s. By knowing where each ball is located we can can calculate the number of balls in each bin (i.e. we can get \( Y_i \) from knowing all the \( X_i \)'s).

One example of an event that can be expressed in terms of the first set of random variables but not he second is, the event that the first ball lands in the second bin. This can be expressed as \( X_1 = 2 \) but it cannot be expressed in terms of the random variables \( Y_i \).

(b) (16 pts.) \( X_i \)'s are mutually independent by the nature of the experiment. Here we check if \( Y_i \)'s are also mutually independent.

(i) (5 pts.) Compute \( P(Y_i = 0), i = 1, \ldots, n \)

Answer: We can rewrite \( P(Y_i = 0) \) as \( P(X_1 \neq i, X_2 \neq i, \ldots, X_m \neq i) \). Now using the independence of the \( X_i \)'s we have

\[
P(Y_i = 0) = P(X_1 \neq i)P(X_2 \neq i) \cdots P(X_m \neq i)
\]

Assuming each bin is equally likely \( P(X_i \neq i) = 1 - 1/n \) so we conclude

\[
P(Y_i = 0) = (1 - 1/n)^m
\]

(ii) (6 pts.) Compute \( P(Y_i = 0, Y_j = 0) \) for \( i \neq j \) (You may want to separate out the two cases when \( n = 2 \) and \( n > 2 \))

Answer:

\( n = 2 \):
Since we only have two bins we only need to calculate $P(Y_1 = 0, Y_2 = 0)$. From the problem statement we know that there is at least 1 ball and it can only go into one of the two bins so we conclude that the event that neither bin has a ball is impossible. Specifically,

$$P(Y_1 = 0, Y_2 = 0) = 0$$

$n > 2$:

We can break the probability down using conditional probabilities as follows:

$$P(Y_i = 0, Y_j = 0) = P(Y_i = 0|Y_j = 0)P(Y_j = 0)$$

From the previous part we know $P(Y_j = 0) = (1 - 1/n)^m$. To calculate the conditional probability we exploit the fact that knowing bucket $j$ has no balls implies the balls must lie in the other $n - 1$ buckets. So

$$P(Y_i = 0|Y_j = 0) = \left(1 - \frac{1}{n-1}\right)^m$$

Multiplying the probabilities gives us

$$P(Y_i = 0, Y_j = 0) = \left(1 - \frac{1}{n-1}\right)^m(1 - 1/n)^m = ((n - 2)/n)^m$$

You can also directly compute this probability by arguing that all the balls must go into the $n - 2$ bins other than $i$ and $j$, giving us $((n - 2)/n)^m$.

(iii) (2 pts.) Are $Y_i$ and $Y_j$ independent?

**Answer:** No the random variables are not independent since

$$P(Y_i = 0|Y_j = 0) = \left(1 - \frac{1}{n-1}\right)^m \neq \left(1 - \frac{1}{n}\right)^m = P(Y_i = 0) \quad i \neq j$$

(iv) (3 pts.) What happens when $n$ is very large? Can you give some intuition for your answer?

**Answer:** When $n$ is very large $P(Y_i = 0|Y_j = 0) \approx P(Y_i = 0)$. Intuitively, the more bins you have, the less information knowing one of the bins is empty gives you about the others.

4. (13 pts.) Communication Channel

(a) (6 pts.) Find the probabilities of the output symbols.

**Answer:** Let $X$ be a random variable denoting the value of the input and $Y$ be a random variable denoting the output. Using the law of total probability and conditional probability we get

$$P(Y = 0) = P(Y = 0|X = 0)P(X = 0) + P(Y = 0|X = 2)P(X = 2)$$

$$= \frac{1}{2}(1 - \epsilon) + \frac{1}{4}\epsilon$$

$$= \frac{1}{2} - \frac{1}{4}\epsilon,$$

$$P(Y = 1) = \frac{1}{2}\epsilon + \frac{1}{4}(1 - \epsilon)$$

$$= \frac{1}{4} + \frac{1}{4}\epsilon,$$

$$P(Y = 2) = \frac{1}{4}\epsilon + \frac{1}{4}(1 - \epsilon)$$

$$= \frac{1}{4}.$$  

As expected, the sum of the probabilities of the events $B_i$ is 1.
(b) (7 pts.) Given that 1 is received, find the probabilities that the input was 0, 1, or 2

**Answer:** The conditional probabilities are

\[
P(X = 0|Y = 1) = \frac{P(X = 0, Y = 1)}{P(Y = 1)} = \frac{\frac{1}{2}e}{\frac{1}{4} + \frac{1}{4}e} = \frac{2e}{1+e},
\]

\[
P(X = 1|Y = 1) = \frac{\frac{1}{4}(1-e)}{\frac{1}{4} + \frac{1}{4}e} = \frac{1-e}{1+e},
\]

\[
P(X = 2|Y = 1) = 0.
\]

Again, note that the sum of the conditional probabilities is 1.

5. (20 pts.) DNA Sequencing In high throughput sequencing technologies, a DNA of length \(G\) symbols \(s_1, s_2, \ldots, s_G\) (each symbol one of the 4 possible nucleotides \(A,G,C,T\)) is sequenced by randomly sampling short subsequences called reads from it. For simplicity, we will assume that the genome is circular. See figure 1. Each read \(R_i\), of length \(L\) symbols, is uniformly sampled from the genome, and the locations of different reads are mutually independent. We sample \(N\) such reads. You can assume that \(G\) is significantly larger than \(L\).

(a) (12 pts.) Compute the probability that position \(i\) on the genome is not covered by any read.

**Answer:** Let \(E_i^{(R_j)}\) be the event that position \(i\) on the genome is not covered by read \(R_j\). We want to calculate \(P\left(\bigcap_{j=1}^{N} E_i^{(R_j)}\right)\). But since the location of each read is independent we can just multiply the probabilities of each event. Specifically,

\[
P\left(\bigcap_{j=1}^{N} E_i^{(R_j)}\right) = \prod_{j=1}^{N} P\left(E_i^{(R_j)}\right) = \prod_{j=1}^{N} (1 - L/G) = (1 - L/G)^N
\]

(b) (4 pts.) Give a non-trivial upper bound on the probability that at least one position on the genome is not covered by any reads.

**Answer:** Let \(A_i\) be the event that position \(i\) is not covered (i.e. the event for which we calculated the probability in part A), then

\[
P(\text{at least one pos. not covered}) = P(\bigcup_{i=1}^{N} A_i) \leq \sum_{i=1}^{N} P(A_i) = G(1 - L/G)^N
\]

Where we have used the union bound to get the upper bound.
(c) **(4 pts.)** Suppose \( G = 3 \times 10^9 \) and \( L = 100 \). How many reads do we need to sample to guarantee that the probability in part (b) is at most 1%?

**Answer:** Plugging in our values and solving for \( N \) we have

\[
3 \times 10^9 \left(1 - \frac{100}{3 \times 10^9}\right)^N \leq 0.01
\]

\[
N \log \left(1 - \frac{100}{3 \times 10^9}\right) \leq \log \left(\frac{0.01}{3 \times 10^9}\right)
\]

\[
N \geq -\frac{26.42705}{-3.333 \times 10^{-8}}
\]

\[
N \geq 7.93 \times 10^8
\]

Where \( \log(\cdot) \) is understood to be the natural logarithm (not that it particularly matters).

6. **(12 pts.) Learning in Monty Hall**

Consider the 4-door Monty Hall problem in HW1. Let us consider a model where the car is randomly behind one of the 4 doors but the contestant’s first choice is door 1 and the contestant follows the Stick-Stick strategy.

(a) **(2 pts.)** Compute and plot the probability distribution of the location of the car.

**Answer:** Let \( X \in \{1, 2, 3, 4\} \) be the location of the car, and \( Y_1, Y_2 \) be the assistant’s first and second door. By the assumption that the car is randomly behind one of the 4 doors, \( P(X = i) = 1/4 \) for \( i = 1, \ldots, 4 \).

(b) **(5 pts.)** Compute and plot the probability distribution of the location of the car, given that the assistant’s first pick is door 2.

**Answer:** We would compute \( P(X = i|Y_1 = 2) \). Apparently \( P(X = 2|Y_1 = 2) = 0 \).

\[
P(X = 1|Y_1 = 2) = \frac{P(X = 1, Y_1 = 2)}{P(Y_1 = 2)}
\]

\[
= \frac{P(X = 1)P(Y_1 = 2|X = 1)}{P(Y_1 = 2)}
\]

\[
= \frac{1/4}{1/3}
\]

\[
= 1/4.
\]

7
\[ P(X = 3 | Y_1 = 2) = \frac{P(X = 3)P(Y_1 = 2 | X = 3)}{P(Y_1 = 2)} = \frac{(1/4)(1/2)}{(1/3)} = 3/8. \]

Similarly \( P(X = 4 | Y_1 = 2) = 3/8. \)

(c) (5 pts.) Compute and plot the probability distribution of the location of the car, given that the assistant’s first pick is door 2 and her second pick is door 3.

**Answer:** We would compute \( P(X = i | Y_1 = 2, Y_2 = 3) \). Apparently \( P(X = 2 | Y_1 = 2, Y_2 = 3) = P(X = 3 | Y_1 = 2, Y_2 = 3) = 0. \)

\[ P(X = 1 | Y_1 = 2, Y_2 = 3) = \frac{P(X = 1)P(Y_1 = 2, Y_2 = 3 | X = 1)}{P(Y_1 = 2, Y_2 = 3)} = \frac{(1/4)(1/6)}{(1/6)} = 1/4. \]

\[ P(X = 4 | Y_1 = 2, Y_2 = 3) = 1 - P(X = 1 | Y_1 = 2, Y_2 = 3) = 3/4. \]
Yes, the contestant is learning more about the prize door based on the assistant’s actions. This can be seen in the fact that the probabilities are initially uniform but later start to concentrate on certain doors.