Homework 2. Finish By 5/28.

Reading: Notes: Chapter 1, Chapter 2. **Videos:** 1.5, 1.6, 2.1, 2.2

Quick Calculations: Every homework will have a problem that focuses on quick calculations to help you get familiar with the mechanics of the concepts introduced that week. This will also help prepare you for exams, which will also include a similar problem.

Problem 2.1 (**Video 1.5**, **Lecture Problem**) Consider an experiment with sample space $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$. The outcomes have probabilities

$$\mathbb{P}[\{1\}] = \frac{1}{4} \quad \mathbb{P}[\{2\}] = \frac{1}{4} \quad \mathbb{P}[\{3\}] = \frac{1}{8} \quad \mathbb{P}[\{4\}] = \frac{1}{8}$$
$$\mathbb{P}[\{5\}] = \frac{1}{16} \quad \mathbb{P}[\{6\}] = \frac{1}{16} \quad \mathbb{P}[\{7\}] = \frac{1}{16} \quad \mathbb{P}[\{8\}] = \frac{1}{16}\ .$$

We also define the events

$$A = \{1, 3, 4\} \qquad B = \{2, 3, 4\} \qquad C = \{3, 4, 5, 6, 7, 8\}$$
$$D = \{2, 3, 5, 6\} \qquad E = \{2, 4, 6, 7\} \qquad F = \{5, 6, 7, 8\}\ .$$

For each of the following questions, give a "Yes" or "No" answer as well as your reasoning and calculations.

(a) Are the events A, B, and C independent? If not, are they at least pairwise independent?

(b) Are the events A and D independent?

(c) Are the events A and F independent?

(d) Are the events B and Ω independent?

(e) Are the events D, E, and F independent? If not, are they are least pairwise independent?

(f) Are the events A and D conditionally independent given C?

Solution:

(a) No, A, B, and C are not independent. However, they are pairwise independent. Here are the calculations:

$$\mathbb{P}[A] = \mathbb{P}\big[\{1, 3, 4\}\big] = \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}$$
$$\mathbb{P}[B] = \mathbb{P}\big[\{2, 3, 4\}\big] = \frac{1}{4} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}$$
$$\mathbb{P}[C] = \mathbb{P}\big[\{3, 4, 5, 6, 7, 8\}\big] = \frac{1}{8} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{2}$$
$$\mathbb{P}[A \cap B] = \mathbb{P}\big[\{3, 4\}\big] = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$$

$$\mathbb{P}[A \cap C] = \mathbb{P}\big[\{3,4\}\big] = \frac{1}{4}$$

$$\mathbb{P}[B \cap C] = \mathbb{P}\big[\{3,4\}\big] = \frac{1}{4}$$

$$\mathbb{P}[A \cap B \cap C] = \mathbb{P}\big[\{3,4\}\big] = \frac{1}{4}$$

$$\mathbb{P}[A \cap B \cap C] \neq \mathbb{P}[A]\mathbb{P}[B]\mathbb{P}[C] \implies \text{not independent}$$

$$\left.\begin{aligned} \mathbb{P}[A \cap B] &= \mathbb{P}[A]\mathbb{P}[B] \\ \mathbb{P}[A \cap C] &= \mathbb{P}[A]\mathbb{P}[C] \\ \mathbb{P}[B \cap C] &= \mathbb{P}[B]\mathbb{P}[C] \end{aligned}\right\} \implies \text{pairwise independent}$$

(b) No, A and D are not independent. Here are the calculations:

$$\mathbb{P}[A] = \frac{1}{2} \text{ from part (a)}$$

$$\mathbb{P}[D] = \mathbb{P}\big[\{2,3,5,6\}\big] = \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} = \frac{1}{2}$$

$$\mathbb{P}[A \cap D] = \mathbb{P}[3] = \frac{1}{8}$$

$$\mathbb{P}[A \cap D] \neq \mathbb{P}[A]\mathbb{P}[D] \implies \text{not independent}$$

(c) No, A and F are not independent. (Notice that they are mutually exclusive, which often means that they are dependent, except in the special case where either event has probability 0.) Here are the calculations:

$$\mathbb{P}[A] = \frac{1}{2} \text{ from part (a)}$$

$$\mathbb{P}[F] = \mathbb{P}\big[\{5,6,7,8\}\big] = \frac{1}{16} + \frac{1}{16} + \frac{1}{16} + \frac{1}{16} = \frac{1}{4}$$

$$\mathbb{P}[A \cap F] = \mathbb{P}[\emptyset] = 0$$

$$\mathbb{P}[A \cap F] \neq \mathbb{P}[A]\mathbb{P}[F] \implies \text{not independent}$$

(d) Yes, the events B and Ω (the sample space) are independent. In fact, the sample space Ω is independent of any event. Here are the calculations, just to check:

$$\mathbb{P}[\Omega] = 1$$

$$\mathbb{P}[B \cap \Omega] = \mathbb{P}[B]$$

$$\mathbb{P}[B \cap \Omega] = \mathbb{P}[B]\mathbb{P}[\Omega] \implies \text{independent}$$

(e) No, D, E, and F are not independent. They are also not pairwise independent. The issue is that D and E are not independent. Here are the calculations:

$$\mathbb{P}[D] = \frac{1}{2} \text{ from part (b)}$$

$$\mathbb{P}[E] = \mathbb{P}\big[\{2,4,6,7\}\big] = \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{16} = \frac{1}{2}$$
$$\mathbb{P}[D \cap E] = \mathbb{P}\big[\{2,6\}\big] = \frac{1}{4} + \frac{1}{16} = \frac{5}{16}$$
$$\mathbb{P}[D \cap E] \neq \mathbb{P}[D]\mathbb{P}[E] \implies \text{not independent}$$

(f) Yes, A and D are conditionally independent given C. (Recall that, from part (b), A and D are not independent. In general, conditional independence does not imply independence. Similarly, independence does not imply conditional independence.) Here are the calculations:

$$\mathbb{P}[C] = \frac{1}{2} \text{ from part (a)}$$
$$\mathbb{P}[A|C] = \frac{\mathbb{P}[A \cap C]}{\mathbb{P}[C]} = \frac{\mathbb{P}\big[\{3,4\}\big]}{\frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$
$$\mathbb{P}[D|C] = \frac{\mathbb{P}[D \cap C]}{\mathbb{P}[C]} = \frac{\mathbb{P}\big[\{3,5,6\}\big]}{\frac{1}{2}} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$
$$\mathbb{P}[A \cap D|C] = \frac{\mathbb{P}[A \cap D \cap C]}{\mathbb{P}[C]} = \frac{\mathbb{P}[3]}{\frac{1}{2}} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}$$
$$\mathbb{P}[A \cap D|C] = \mathbb{P}[A|C]\mathbb{P}[D|C] \implies \text{conditionally independent}$$

Problem 2.2 (**Video 1.5, 1.6**, **Lecture Problem**) Consider the following scenario. You play a simple game with probability of winning 1/4. You play this game repeatedly until your third loss, and then stop playing. Assume all games are independent.

(a) What is the probability of the following specific sequence of game outcomes: Win, Lose, Win, Lose, Lose?

Solution:

Let W_i be the event that you win the i^{th} game. By independence, we have that

$$\begin{aligned}\mathbb{P}[W_1 \cap W_2^{\mathsf{c}} \cap W_3 \cap W_4^{\mathsf{c}} \cap W_5^{\mathsf{c}}] &= \mathbb{P}[W_1]\,\mathbb{P}[W_2^{\mathsf{c}}]\,\mathbb{P}[W_3]\,\mathbb{P}[W_4^{\mathsf{c}}]\,\mathbb{P}[W_5^{\mathsf{c}}] \\ &= \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} = \frac{27}{1024}\end{aligned}$$

(b) How many different sequences of games are there that end after exactly 5 games? (Hint: you must lose the last game to stop. There aren't that many, so you can enumerate them.)

Solution:

First, notice that the fifth game must be a loss, since you stop at 5 games. Thus, we have 2 losses to place in the first 4 slots. This is done without replacement and order does not matter (since one loss is the same as any other). Thus, there are $\binom{4}{2} = 6$ different sequences of games that end at exactly 5 games.

(c) What is the probability of playing exactly 5 games?

Solution:

Each of the 6 possible sequences of exactly 5 games has probability $\frac{27}{1024}$. Thus, the probability of playing exactly 5 games is $\frac{6 \cdot 27}{1024} = \frac{81}{512}$.

(d) Given that you play exactly 5 games, what is the probability that your first game ended in a loss?

Solution:

First, note that each of the 6 possible sequences of exactly 5 games with 3 losses has the same probability. Thus, we can simply count how many ways there are to place 3 losses in 5 games, with one loss reserved for the fifth slot and another reserved for the first slot. There are $\binom{3}{1} = 3$ such sequences, and the probability of seeing one of these is just $\frac{3}{6} = \frac{1}{2}$.

(e) Now, let's generalize this a bit. Say the probability of winning an individual game is p and that you play until your m^{th} loss. What is the probability of playing exactly k games?

Solution:

A specific sequence that ends on the k^{th} game has probability $p^{k-m}(1-p)^m$ by independence. Since one loss is reserved for the last slot, there are $k-1$ free slots and $m-1$ losses to place. Thus, there are $\binom{k-1}{m-1}$ possible sequences of exactly k games. Finally, the probability of playing exactly k games is $\binom{k-1}{m-1}p^{k-m}(1-p)^m$. Note: If $k < m$ then the probability of playing exactly k games is zero since you can never get more losses than the number of games you play.

Problem 2.3 (**Video 1.6**) You would like to evaluate the probability of success for testing a batch of n widgets. To start out, let's assume that if there is a problem with the batch, exactly 1 out of the n widgets are defective. You are willing to test only k of the widgets (due to budget or times constraints).

(a) How many ways are there of testing k out of n widgets?

(b) How many ways are there of testing k widgets with the defective widget included?

(c) Use your answers from parts (a) and (b) to determine the probability of catching a defective batch.

(d) Evaluate your answer from part (c) for $n = 20$ and $k = 5$.

(e) Now, say that a defective batch contains exactly 2 defective widgets. How many ways are there of testing k widgets with *at least one* defective widget included? (You may assume that $k > 2$.)

(f) Use your answer from part (e) to determine the probability of catching a defective batch.

(g) Evaluate your answer from part (f) for $n = 20$ and $k = 5$.

Solution:

(a) There are $\binom{n}{k}$ ways of testing k out of n processors since order does not matter.

(b) There are $\binom{1}{1} \cdot \binom{n-1}{k-1} = \binom{n-1}{k-1}$ ways of testing k processors with the defective processor included. Intuitively, one of the choices is used on the defective processor and the remaining $k-1$ are spent on the $n-1$ non-defective processors.

(c) Let C be the event that we catch a defective batch of processors. The probability is simply the ratio of the number of ways to test with the defective processor included over the total number of ways to test,

$$\mathbb{P}[C] = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\frac{(n-1)!}{(n-1-(k-1))!(k-1)!}}{\frac{n!}{(n-k)!(k)!}} = \frac{\frac{(n-1)!}{(n-k)!(k-1)!}}{\frac{n!}{(n-k)!(k)!}} = \frac{\frac{k!}{(k-1)!}}{\frac{n!}{(n-1)!}} = \frac{k}{n} \,.$$

(d) Plugging in $n = 20$ and $k = 5$ into our formula from part (c), we get

$$\mathbb{P}[C] = \frac{5}{20} = \frac{1}{4} \,.$$

(e) The first step is to realize that we need to count the events that we test exactly 1 or 2 defective widgets separately and then add them up. The number of ways of testing k widgets with exactly $i < k$ defective widgets included is $\binom{2}{i} \cdot \binom{n-2}{k-i}$. Therefore, the total number of tests that include at least one defective widget is

$$\binom{2}{1} \cdot \binom{n-2}{k-1} + \binom{2}{2} \cdot \binom{n-2}{k-2} = 2 \cdot \binom{n-2}{k-1} + 1 \cdot \binom{n-2}{k-2} \,.$$

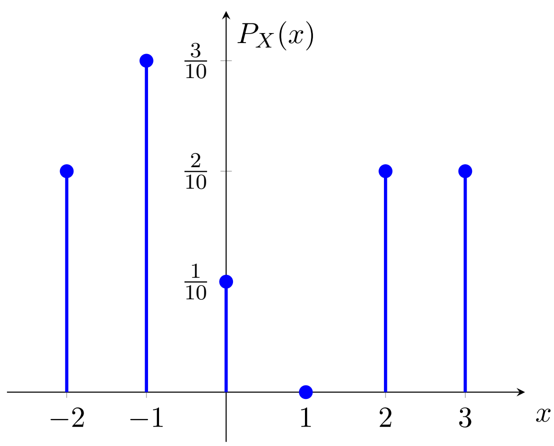
(f) Let C be the event that we catch a defective batch of widgets. The probability is simply the ratio of the number of ways to test with the defective widget included over the total number of ways to test,

$$\mathbb{P}[C] = \frac{2 \cdot \binom{n-2}{k-1} + 1 \cdot \binom{n-2}{k-2}}{\binom{n}{k}} \,.$$

(g) Plugging in $n = 20$ and $k = 5$ into our formula from part (c), we get

$$\mathbb{P}[C] = \frac{2 \cdot \binom{18}{4} + \binom{18}{3}}{\binom{20}{5}} = \frac{2 \cdot 3060 + 816}{15504} = \frac{6936}{15504} = \frac{17}{38} \approx 0.447 \; .$$

Problem 2.4 (**Video 2.1, 2.2**, **Lecture Problem**)



Consider the PMF above and let $A = \{-2, -1, 3\}$.

(a) Calculate the probability that X falls into A, $\mathbb{P}[X \in A]$.

Solution:

$$\mathbb{P}[X \in A] = \sum_{x \in A} P_X(x) = P_X(-2) + P_X(-1) + P_X(3) = \frac{2}{10} + \frac{3}{10} + \frac{2}{10} = \frac{7}{10}$$

(b) Calculate the probability that X^2 exceeds 1, $\mathbb{P}[X^2 > 1]$.

Solution:

First, notice that we can translate the condition $X^2 > 1$ into membership in a set $C = \{x \in R_X : x^2 > 1\} = \{-2, 2, 3\}$. Therefore, we have that

$$\begin{aligned}\mathbb{P}[X^2 > 1] &= \mathbb{P}[X \in C] \\ &= \sum_{x \in C} P_X(x) = P_X(-2) + P_X(2) + P_X(3) = \frac{2}{10} + \frac{2}{10} + \frac{2}{10} = \frac{3}{5}\end{aligned}$$

(c) Given that $\{X \in A\}$ occurs, what is the conditional probability that X^2 exceeds 1, $\mathbb{P}[X^2 > 1 | X \in A]$?

Solution:

By the definition of conditional probability,

$$\begin{aligned}\mathbb{P}[X^2 > 1 | X \in A] &= \mathbb{P}[X \in C | X \in A] \\ &\frac{\mathbb{P}[\{X \in C\} \cap \{X \in A\}]}{\mathbb{P}[X \in A]} \\ &\frac{\mathbb{P}[X \in (C \cap A)]}{\mathbb{P}[X \in A]} \,.\end{aligned}$$

Now, note that $C \cap A = \{-2, 3\}$ and

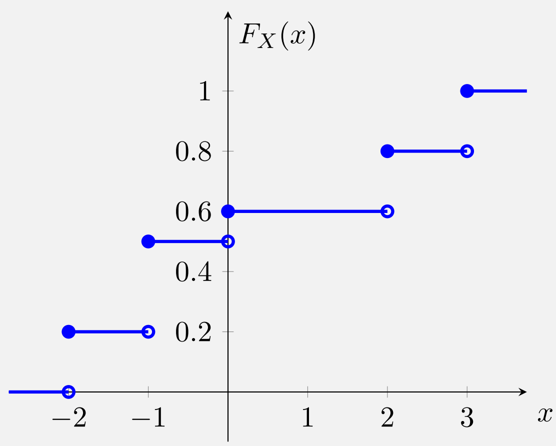
$$\mathbb{P}[X \in (C \cap A)] = \sum_{x \in C \cap A} P_X(x) = P_X(-2) + P_X(3) = \frac{2}{10} + \frac{2}{10} = \frac{2}{5} \,.$$

Combining this with our answer from part (a), we get

$$\mathbb{P}[X^2 > 1 | X \in A] = \frac{2/5}{7/10} = \frac{4}{7} \,.$$

(d) Determine the CDF $F_X(x)$.

Solution:



Problem 2.5 (**Video 1.5, 1.6, 2.1, 2.2**, **Quick Calculations**) Calculate each of the requested quantities.

(a) Let A and B be independent events with $\mathbb{P}[A] = 1/5$ and $\mathbb{P}[B] = 1/4$. Calculate $\mathbb{P}[A \cap B]$ and $\mathbb{P}[A \cup B]$.

(b) Let A_1, A_2, A_3 be events that are conditionally independent given B. *Additionally, assume that A_1, A_2, A_3 are conditionally independent given B^{c}.* Assume that $\mathbb{P}[A_i | B] = 1/4$ and $\mathbb{P}[A_i | B^{\mathsf{c}}] = 1/2$ for $i = 1, 2, 3$ and $\mathbb{P}[B] = 1/3$. Calculate $\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3 | B]$ and $\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3]$.

(c) Consider a packet of jellybeans that contains 9 jellybeans, of which 4 are lemon and the remaining 5 are raspberry. You reach in and pull out 3 jellybeans. What is the probability that they are all lemon? What is the probability that they are all raspberry?

(d) Let X be a random variable with PMF $P_X(x) = \begin{cases} 1/6 & x = -1, +1 \\ 2/3 & x = 0 \end{cases}$. Calculate $\mathbb{P}[X \neq 0]$ and $\mathbb{P}[X > 0 | X \neq 0]$.

(e) If the random variable Y has CDF $F_Y(y) = \begin{cases} 0 & y < 1 \\ 1/4 & 1 \leq y < 5 \\ 1 & 5 \leq y \end{cases}$, what is the PMF of Y?

Solution:

(a) Since A and B are independent, we have that $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B] = \frac{1}{5} \cdot \frac{1}{4} = \frac{1}{20}$. By the inclusion-exclusion property, we have that

$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cap B] = \frac{1}{5} + \frac{1}{4} - \frac{1}{20} = \frac{8}{20} = \frac{2}{5} \ .$$

(b) Since A_1, A_2, A_3 are conditionally independent given B, we know that $A_1, A_2^{\mathsf{c}}, A_3$ are also conditionally independent given B. Therefore, we have that

$$\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3 | B] = \mathbb{P}[A_1|B]\,\mathbb{P}[A_2^{\mathsf{c}}|B]\,\mathbb{P}[A_3|B] = \frac{1}{4} \cdot \left(1 - \frac{1}{4}\right) \cdot \frac{1}{4} = \frac{3}{64} \ .$$

To calculate $\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3]$, we will use the Law of Total Probability,

$$\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3] = \mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3|B]\,\mathbb{P}[B] + \mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3|B^{\mathsf{c}}]\,\mathbb{P}[B^{\mathsf{c}}] \ .$$

We can use the fact that A_1, A_2, A_3 are conditionally independent given B^{c} to get

$$\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3|B^{\mathsf{c}}] = \mathbb{P}[A_1|B^{\mathsf{c}}]\,\mathbb{P}[A_2^{\mathsf{c}}|B^{\mathsf{c}}]\,\mathbb{P}[A_3|B^{\mathsf{c}}] = \frac{1}{2} \cdot \left(1 - \frac{1}{2}\right) \cdot \frac{1}{2} = \frac{1}{8} \ .$$

Putting this together, we have

$$\mathbb{P}[A_1 \cap A_2^{\mathsf{c}} \cap A_3] = \frac{3}{64} \cdot \frac{1}{3} + \frac{1}{8} \cdot \frac{2}{3} = \frac{19}{192} \approx 0.099 \ .$$

(c) First, note that we this is an order-independent, without-replacement sampling problem. Therefore, if there are n items out of which we select k, we can make $\binom{n}{k}$ possible selections. The total number of ways to draw 3 out of 9 jellybeans is $\binom{9}{3} = 84$ and the number of ways to choose 3 out of 4 yellow jellybeans is $\binom{4}{3} = 4$. Thus, the probability of drawing all yellow jellybeans is $\frac{4}{84} = \frac{1}{21}$. Similarly, the number of ways to choose 3 out of 5 raspberry jellybeans is $\binom{5}{3} = 10$ and the probability of drawing all raspberry jellybeans is $\frac{10}{84} = \frac{5}{42}$.

(d) First, we have that $\mathbb{P}[X \neq 0] = P_X(-1) + P_X(+1) = 1/6 + 1/6 = 1/3$. Second, using the definition of conditional probability,

$$\mathbb{P}[X > 0 | X \neq 0] = \frac{\mathbb{P}[X > 0]}{\mathbb{P}[X \neq 0]} = \frac{P_X(+1)}{1/3} = \frac{1/6}{1/3} = \frac{1}{2} \ .$$

(e) The CDF jumps up at two points $y = 1$ and $y = 5$, and we use the heights of these jumps to determine the PMF values: $P_Y(y) = \begin{cases} 1/4 & y = 1 \\ 3/4 & y = 5 \end{cases}$