

CSE 312 – Section 4 Solutions

Spring 2026

Review of Main Concepts

- **Random Variable (rv):** A numeric function $X : \Omega \rightarrow \mathbb{R}$ of the outcome.
- **Range/Support:** The support/range of a random variable X , denoted Ω_X , is the set of all possible values that X can take on.
- **Discrete Random Variable (drv):** A random variable taking on a countable (either finite or countably infinite) number of possible values.
- **Probability Mass Function (pmf) for a discrete random variable X :** a function $p_X : \Omega_X \rightarrow [0, 1]$ with $p_X(x) = \mathbb{P}(X = x)$ that maps possible values of a discrete random variable to the probability of that value happening, such that $\sum_x p_X(x) = 1$.
- **Cumulative Distribution Function (CDF) for a random variable X :** a function $F_X : \mathbb{R} \rightarrow \mathbb{R}$ with $F_X(x) = \mathbb{P}(X \leq x)$
- **Expectation (expected value or mean):** The expectation of a discrete random variable is defined to be $\mathbb{E}[X] = \sum_x x p_X(x) = \sum_x x \mathbb{P}(X = x)$.
- **Law of the unconscious statistician (LOTUS):** The expectation of a function of a discrete random variable $g(X)$ is $\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$.
- **Linearity of Expectation (LoE):** Let X and Y be random variables, and $a, b, c \in \mathbb{R}$. Then, $\mathbb{E}[aX + bY + c] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$. Also, for any random variables X_1, \dots, X_n ,

$$\mathbb{E}[X_1 + X_2 + \dots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n].$$

- **Variance:** Let X be a random variable and $\mu = \mathbb{E}[X]$. The variance of X is defined to be $\text{Var}(X) = \mathbb{E}[(X - \mu)^2]$. Notice that since this is an expectation of a non-negative random variable $((X - \mu)^2)$, variance is always non-negative. With some algebra, we can simplify this to $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
- **Standard Deviation:** Let X be a random variable. We define the standard deviation of X to be the square root of the variance, and denote it $\sigma = \sqrt{\text{Var}(X)}$.
- **Property of Variance:** Let $a, b \in \mathbb{R}$ and let X be a random variable. Then, $\text{Var}(aX + b) = a^2 \text{Var}(X)$.
- **Independence:** Random variables X and Y are independent iff

$$\forall x \forall y, \quad \mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x) \mathbb{P}(Y = y)$$

In this case, we have $\mathbb{E}[XY] = \mathbb{E}[X] \mathbb{E}[Y]$ (the converse is not necessarily true).

- **i.i.d. (independent and identically distributed):** Random variables X_1, \dots, X_n are i.i.d. (or iid) iff they are independent and have the same probability mass function.
- **Variance of Independent Variables:** If X is independent of Y , $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. This depends on independence, whereas linearity of expectation always holds. Note that this combined with the above shows that $\forall a, b, c \in \mathbb{R}$ and if X is independent of Y , $\text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y)$.

Announcements & Plan for Section

Announcements

- PSet 2 grades were released and can be viewed on Gradescope. Regrade requests will close on 4/25. We highly recommend taking a look through feedback on the psets, the common errors on Ed and the solutions that were posted (link on Ed).
- PSet 3 was due earlier this week.
- PSet 4 is released — start early.
- This week's focus: Linearity of Expectation, LOTUS, and Variance, leading into discrete random variables.

Plan for Section

- **Content Review (Problem 1)**
Quick check of core definitions: PMF vs CDF, expectation, and variance.
- **Problem 3 (3-sided die)**
Build intuition for PMFs and practice computing expectation both directly and using linearity.
- **Problem 8 (Hungry washing machine)**
Key application of indicator variables and linearity of expectation — avoiding messy counting.
- **Problem 11 (Hat check)**
Classic example showing linearity of expectation works even without independence.
- **Problem 5 (if time)**
Reinforces thinking about ranges/support of random variables.
- **Problem 6 (if time)**
Additional practice connecting PMFs, expectations, and variance.

1 Content Review

- a) True or false: the range of a random variable X is the set of probabilities corresponding to the possible values X can take on.

False. The range (or *support*) of a random variable is the set of all possible *values* it can take on, not the probabilities of those values.

b) What is the relationship between standard deviation and variance of a random variable X ?

- $\sigma = (\text{Var}(X))^2$
- $\sigma = \text{Var}(X^2)$
- $\text{Var}(X) = \sigma^2$

$\text{Var}(X) = \sigma^2$ (or equivalently, $\sigma = \sqrt{\text{Var}(X)}$).

c) Let X be the random variable representing the the sum of a 3-dice roll of 6-sided dice. Which function would you use to determine the probability that $X = 7$?

- CDF (cumulative distribution function)
- PMF (probability mass function)

PMF. We use the PMF when we want to find the probability of a specific, exact value of a discrete random variable occurring.

d) Let X be the random variable representing the outcome of taking the sum of a 3-dice roll of 6-sided dice. Which function would you use to determine the probability that $X \leq 7$?

- CDF (cumulative distribution function)
- PMF (probability mass function)

CDF. The CDF directly gives us the cumulative probability $\mathbb{P}(X \leq x)$.

e) A random variable X has the PMF

$$p_X(x) = \begin{cases} 1/4 & x = -1 \\ 1/4 & x = 0 \\ 1/2 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

What is $\mathbb{E}[X]$?

- 1/4
- 3/4
- 1
- 2

3/4.

$$\mathbb{E}[X] = \sum_{x \in \Omega_X} xp_X(x) = (-1) \left(\frac{1}{4}\right) + (0) \left(\frac{1}{4}\right) + (2) \left(\frac{1}{2}\right) = -\frac{1}{4} + 0 + 1 = \frac{3}{4}$$

f) A random variable X has the PMF

$$p_X(x) = \begin{cases} 1/4 & x = -1 \\ 1/4 & x = 0 \\ 1/2 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

What is $\text{Var}(X)$?

- $3/4$
- 1
- $((1/4) + 2) - ((3/4)^2) = 27/16$
- $((1/4) + 2) + ((3/4)^2) = 45/16$

27/16.

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\ &= \left(\sum_{x \in \Omega_X} x^2 p_X(x) \right) - \left(\frac{3}{4} \right)^2 \\ &= \left((-1)^2 \cdot \frac{1}{4} + (0)^2 \cdot \frac{1}{4} + (2)^2 \cdot \frac{1}{2} \right) - \frac{9}{16} \\ &= \left(\frac{1}{4} + 0 + 2 \right) - \frac{9}{16} \\ &= \frac{9}{4} - \frac{9}{16} = \frac{27}{16} \end{aligned}$$

g) A random variable X has the CDF

$$F_X(x) = \begin{cases} 0 & x < 1 \\ \frac{1}{5} & 1 \leq x < 2 \\ \frac{1}{2} & 2 \leq x < 3 \\ \frac{3}{5} & 3 \leq x < 4 \\ 1 & x \geq 4 \end{cases}$$

What is $p_X(2)$?

- $3/10$
- $1/2$
- $7/10$
- $1/5$

3/10.

$$\begin{aligned} p_X(2) &= \mathbb{P}_X(X \leq 2) - \mathbb{P}_X(X \leq 1) \\ &= F_X(2) - F_X(1) \\ &= \frac{1}{2} - \frac{1}{5} = \frac{3}{10} \end{aligned}$$

2 Identify that Range!

Identify the support/range Ω_X of the random variable X , if X is...

- a) The sum of two rolls of a six-sided die.

X takes on every integer value between the minimum sum ($1+1=2$) and the maximum sum ($6+6=12$).
 $\Omega_X = \{2, 3, \dots, 12\}$

- b) The number of lottery tickets I buy until I win it (assuming I buy at least one lottery ticket).

X takes on all positive integer values. (It is possible, though increasingly unlikely, that I never win the lottery, so there is no upper bound).
 $\Omega_X = \{1, 2, 3, \dots\}$

- c) The number of heads in n flips of a coin with $0 < \mathbb{P}(\text{head}) < 1$.

X takes on every integer value between the minimum number of heads (0), and the maximum possible (n).
 $\Omega_X = \{0, 1, \dots, n\}$

- d) The number of heads in n flips of a coin with $\mathbb{P}(\text{head}) = 1$.

Since $\mathbb{P}(\text{head}) = 1$, we are mathematically guaranteed to get exactly n heads in n flips.
 $\Omega_X = \{n\}$

- e) The number of whole minutes I wait at the bus stop for the next bus.

The number of whole minutes is discrete and will take on values between the minimum waiting time (0, the bus is already here), and the maximum theoretically unbounded waiting time.
 $\Omega_X = \{0, 1, 2, \dots\}$

3 3-sided Die

Let the random variable X be the sum of two independent rolls of a fair 3-sided die. (If you are having trouble imagining what that looks like, you can use a 6-sided die and change the numbers on 3 of its faces.)

a) What is the probability mass function of X ?

First let us define the range of X . A three-sided die can take on values $\{1, 2, 3\}$. Since X is the sum of two rolls, the minimum is 2 and the maximum is 6. Thus, $\Omega_X = \{2, 3, 4, 5, 6\}$.

We can then define the PMF of X . To that end, we define two random variables R_1, R_2 with R_1 being the roll of the first die, and R_2 being the roll of the second die. Then, $X = R_1 + R_2$. Note that $\Omega_{R_1} = \Omega_{R_2} = \{1, 2, 3\}$ and $\mathbb{P}(R_i = x) = 1/3$ for all valid x .

We can compute the probability of each sum by adding up the probabilities of the specific roll combinations (R_1, R_2) that yield that sum:

- $X = 2$: $(1,1) \rightarrow \frac{1}{3} \cdot \frac{1}{3} = \frac{1}{9}$
- $X = 3$: $(1,2), (2,1) \rightarrow \frac{2}{9}$
- $X = 4$: $(1,3), (2,2), (3,1) \rightarrow \frac{3}{9}$
- $X = 5$: $(2,3), (3,2) \rightarrow \frac{2}{9}$
- $X = 6$: $(3,3) \rightarrow \frac{1}{9}$

Summarizing this in a piece-wise PMF:

$$p_X(k) = \begin{cases} 1/9 & k = 2 \\ 2/9 & k = 3 \\ 3/9 & k = 4 \\ 2/9 & k = 5 \\ 1/9 & k = 6 \\ 0 & \text{otherwise} \end{cases}$$

b) Find $\mathbb{E}[X]$ directly from the definition of expectation.

$$\begin{aligned} \mathbb{E}[X] &= \sum_{k=2}^6 k \cdot p_X(k) = 2 \left(\frac{1}{9}\right) + 3 \left(\frac{2}{9}\right) + 4 \left(\frac{3}{9}\right) + 5 \left(\frac{2}{9}\right) + 6 \left(\frac{1}{9}\right) \\ &= \frac{2 + 6 + 12 + 10 + 6}{9} = \frac{36}{9} = \boxed{4} \end{aligned}$$

c) Find $\mathbb{E}[X]$ again, but this time using Linearity of Expectation.

Let R_1 be the roll of the first die, and R_2 the roll of the second. Then, $X = R_1 + R_2$.
By Linearity of Expectation, we get:

$$\mathbb{E}[X] = \mathbb{E}[R_1 + R_2] = \mathbb{E}[R_1] + \mathbb{E}[R_2]$$

We first compute the expected value of a single roll:

$$\mathbb{E}[R_1] = \sum_{i=1}^3 i \cdot \mathbb{P}(R_1 = i) = 1 \left(\frac{1}{3}\right) + 2 \left(\frac{1}{3}\right) + 3 \left(\frac{1}{3}\right) = \frac{6}{3} = 2$$

Since the dice are identical, $\mathbb{E}[R_2] = 2$ as well.

Plugging this into our expression for X :

$$\mathbb{E}[X] = 2 + 2 = \boxed{4}$$

d) What is $\text{Var}(X)$?

We know from the definition of variance that:

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

We computed $\mathbb{E}[X] = 4$ in the previous parts. Now we compute $\mathbb{E}[X^2]$ using the PMF:

$$\begin{aligned} \mathbb{E}[X^2] &= \sum_{x=2}^6 x^2 p_X(x) = \frac{2^2(1) + 3^2(2) + 4^2(3) + 5^2(2) + 6^2(1)}{9} = \frac{4 + 18 + 48 + 50 + 36}{9} \\ &= \frac{156}{9} = \frac{52}{3} \end{aligned}$$

Plugging this into our variance equation gives us:

$$\text{Var}(X) = \frac{52}{3} - (4)^2 = \frac{52}{3} - \frac{48}{3} = \boxed{\frac{4}{3}}$$

4 Kit Kats Again

Suppose we have N candies in a jar, K of which are Kit-Kats. Suppose we draw (without replacement) until we have (exactly) k Kit-Kats, where $k \leq K \leq N$. Let X be the number of draws until the k^{th} Kit-Kat. What is Ω_X , the range of X ? What is $p_X(n) = \mathbb{P}(X = n)$? (X is called a “negative hypergeometric” random variable).

Since it takes a minimum of k draws to get k Kit-Kats, the lowest possible value is k . The absolute maximum number of draws occurs if we draw every single non-Kit-Kat candy in the

jar ($N - K$ of them) before we finally get our k -th Kit-Kat. Therefore, the maximum number of draws is $(N - K) + k$.

$$\Omega_X = \{k, k + 1, \dots, N - K + k\}$$

There are two completely valid ways to find the PMF.

Method 1: Pure Combinatorics

Imagine drawing all N candies one after the other without replacement. There are $\binom{N}{K}$ equally likely ways to position the K Kit-Kats among the N total draws. For the k -th Kit-Kat to occur exactly on the n -th draw, two independent conditions must be met:

- a) We must position exactly $k - 1$ Kit-Kats somewhere among the first $n - 1$ draws: $\binom{n-1}{k-1}$ ways.
- b) There is a Kit-Kat at the n -th draw.
- c) We must position the remaining $K - k$ Kit-Kats somewhere among the final $N - n$ draws: $\binom{N-n}{K-k}$ ways.

By the product rule, we multiply these independent choices and divide by the total sample space:

$$p_X(n) = \mathbb{P}(X = n) = \frac{\binom{n-1}{k-1} \binom{N-n}{K-k}}{\binom{N}{K}}$$

Method 2: Conditional Probability

Let B be the event: "We draw exactly k Kit-Kats in our first n draws." Then

$$\mathbb{P}(B) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}$$

Let A be the event: "The n -th draw is a Kit-Kat." We want to find $\mathbb{P}(A \cap B)$, which means we drew exactly k Kit-Kats in our first n draws AND the very last of those n draws was a Kit-Kat (making it the k -th one).

Conditioned on event B , we know we have a set of n drawn candies where exactly k of them are Kit-Kats. Because the order in which those n candies were drawn is uniformly random, the probability that the *last* candy drawn was one of the Kit-Kats is simply $\frac{k}{n}$. Therefore, $\mathbb{P}(A | B) = \frac{k}{n}$.

Applying the Chain Rule ($\mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A | B)$) gives:

$$p_X(n) = \mathbb{P}(X = n) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}} \cdot \frac{k}{n}$$

(Note: While these expressions look entirely different, with a bit of algebra, you can verify that they are equal!)

5 Hungry Washing Machine

You have 10 pairs of socks (so 20 socks in total), with each pair being a different color. You put them in the washing machine, but the washing machine eats 4 of the socks chosen at random.

Every subset of 4 socks is equally probable to be the subset that gets eaten. Let X be the number of complete pairs of socks that you have left.

- a) What is the range of X , Ω_X (the set of possible values it can take on)? What is the probability mass function of X ?

The washing machine eats exactly 4 socks. The outcomes can be categorized by how those 4 eaten socks were distributed among the pairs:

- The 4 socks belong to 4 completely different pairs. This ruins 4 pairs, leaving $10 - 4 = 6$ complete pairs.
- The 4 socks belong to 1 complete pair and 2 other single socks. This ruins 3 pairs, leaving $10 - 3 = 7$ complete pairs.
- The 4 socks belong to exactly 2 complete pairs. This ruins 2 pairs, leaving $10 - 2 = 8$ complete pairs.

Thus, the range is

$$\Omega_X = \{6, 7, 8\}$$

Because every subset of 4 socks is equally likely, our sample space size is $|\Omega| = \binom{20}{4} = 4845$. We can compute $\mathbb{P}(X = k) = \frac{|X=k|}{|\Omega|}$ for each k :

For $k = 6$: We must pick 4 out of the 10 pairs of socks from which the machine will eat a single sock ($\binom{10}{4}$ ways). For each of these 4 pairs, the machine has 2 specific socks to pick from ($\binom{2}{1}^4 = 2^4$ ways). $|X = 6| = \binom{10}{4}2^4 = 210 \cdot 16 = 3360$.

For $k = 7$: We first pick 1 out of the 10 pairs to be eaten in its entirety ($\binom{10}{1}$ ways). We then pick 2 out of the remaining 9 pairs from which it will eat a single sock ($\binom{9}{2}$ ways). For each of these 2 pairs, there are 2 specific socks to pick from ($\binom{2}{1}^2 = 2^2$ ways). $|X = 7| = \binom{10}{1}\binom{9}{2}2^2 = 10 \cdot 36 \cdot 4 = 1440$.

For $k = 8$: We simply pick 2 out of the 10 pairs of socks to be eaten entirely ($\binom{10}{2}$ ways). $|X = 8| = \binom{10}{2} = 45$.

(Self-check: $3360 + 1440 + 45 = 4845$, which perfectly matches $|\Omega|$.)

Summarizing as a PMF:

$$p_X(k) = \begin{cases} \frac{3360}{4845} & k = 6 \\ \frac{1440}{4845} & k = 7 \\ \frac{45}{4845} & k = 8 \\ 0 & \text{otherwise} \end{cases}$$

- b) Find $\mathbb{E}[X]$ from the definition of expectation.

$$\begin{aligned}\mathbb{E}[X] &= \sum_{k \in \Omega_X} k \cdot p_X(k) = 6 \left(\frac{3360}{4845} \right) + 7 \left(\frac{1440}{4845} \right) + 8 \left(\frac{45}{4845} \right) = \frac{20160 + 10080 + 360}{4845} \\ &= \frac{30600}{4845} = \boxed{\frac{120}{19}}\end{aligned}$$

- c) Find $\mathbb{E}[X]$ using Linearity of Expectation. (see [this video for a walkthrough of this problem!](#))

Step 1: Define Indicator Random Variables

For $i \in \{1, \dots, 10\}$, let X_i be an indicator random variable such that $X_i = 1$ if pair i survived completely intact, and 0 otherwise. We can express our total intact pairs as $X = \sum_{i=1}^{10} X_i$.

Step 2: Find Expectation of the Indicator

The expected value of an indicator variable is simply the probability the event occurs: $\mathbb{E}[X_i] = \mathbb{P}(X_i = 1)$. For pair i to survive intact, the washing machine must have chosen its 4 eaten socks exclusively from the *other* 18 available socks.

$$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{\binom{18}{4}}{\binom{20}{4}} = \frac{3060}{4845} = \frac{12}{19}$$

Step 3: Apply Linearity of Expectation

By Linearity of Expectation, the expectation of the sum is the sum of the expectations:

$$\mathbb{E}[X] = \mathbb{E} \left[\sum_{i=1}^{10} X_i \right] = \sum_{i=1}^{10} \mathbb{E}[X_i] = 10 \cdot \left(\frac{12}{19} \right) = \boxed{\frac{120}{19}}$$

- d) Which way was easier? Doing both (a) and (b), or just (c)?

Part (c) was vastly easier and less prone to algebraic error. In this specific problem, computing the PMF (parts a and b) was manageable because there were only 3 possible values in the range. However, if you had 50 pairs of socks and the machine ate 20 of them, the PMF would have dozens of complex cases, making part (b) a nightmare. Linearity of Expectation (c) scales effortlessly regardless of the numbers involved!

6 Hat Check

At a reception, n people give their hats to a hat-check person. When they leave, the hat-check person gives each of them a hat chosen at random from the hats that remain. What is the expected number of people who get their own hats back? (Notice that the hats returned to two people are not independent events: if a certain hat is returned to one person, it cannot also be returned to the other person.)

Note: See [this video for a walkthrough of this problem!](#)

Let X be the total number of people who get their own hats back.

Step 1: Define Indicator Random Variables

For $i \in \{1, \dots, n\}$, let X_i be an indicator variable where $X_i = 1$ if person i gets their own hat back, and 0 otherwise. Thus, $X = \sum_{i=1}^n X_i$.

Step 2: Find Expectation of the Indicator

The expected value of an indicator is $\mathbb{E}[X_i] = \mathbb{P}(X_i = 1)$. The sample space $|\Omega|$ is all $n!$ possible ways to distribute the n hats among the n people. The event we care about is person i receiving hat i . If we lock person i with hat i , there are $(n - 1)$ remaining people and $(n - 1)$ remaining hats, which can be distributed in $(n - 1)!$ ways.

$$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \frac{(n - 1)!}{n!} = \frac{1}{n}$$

Step 3: Apply Linearity of Expectation

Notice that even though the random variables X_i are heavily dependent on each other, *Linearity of Expectation holds regardless of independence!*

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n \frac{1}{n} = n \cdot \frac{1}{n} = \boxed{1}$$

On average, exactly 1 person will get their own hat back, whether there are 10 people at the party or 10,000 people!

7 Frogger

A frog starts on a 1-dimensional number line at 0. At each second, independently, the frog takes a unit step right with probability p_1 , to the left with probability p_2 , and doesn't move with probability p_3 , where $p_1 + p_2 + p_3 = 1$. After 2 seconds, let X be the location of the frog.

a) Find $p_X(k)$, the probability mass function for X .

Let an outcome be defined by the sequence of the two steps. For example, LN means a Left step followed by No step. Because the frog takes at most 2 steps, the range of X is $\{-2, -1, 0, 1, 2\}$. We calculate the probability of each final position by adding the probabilities of the step sequences that land there:

- $X = -2$: Must step Left then Left (LL). $\mathbb{P}(LL) = p_2 \cdot p_2 = p_2^2$
- $X = -1$: Can be (LN) or (NL). $\mathbb{P}(LN \cup NL) = p_2p_3 + p_3p_2 = 2p_2p_3$
- $X = 0$: Can be (NN), (LR), or (RL). $\mathbb{P}(NN \cup LR \cup RL) = p_3^2 + p_2p_1 + p_1p_2 = p_3^2 + 2p_1p_2$
- $X = 1$: Can be (RN) or (NR). $\mathbb{P}(RN \cup NR) = p_1p_3 + p_3p_1 = 2p_1p_3$
- $X = 2$: Must step Right then Right (RR). $\mathbb{P}(RR) = p_1 \cdot p_1 = p_1^2$

Summarizing as a PMF:

$$p_X(k) = \begin{cases} p_2^2 & k = -2 \\ 2p_2p_3 & k = -1 \\ p_3^2 + 2p_1p_2 & k = 0 \\ 2p_1p_3 & k = 1 \\ p_1^2 & k = 2 \\ 0 & \text{otherwise} \end{cases}$$

b) Compute $\mathbb{E}[X]$ from the definition.

$$\begin{aligned} \mathbb{E}[X] &= \sum_{k=-2}^2 k \cdot p_X(k) \\ &= (-2)(p_2^2) + (-1)(2p_2p_3) + (0)(p_3^2 + 2p_1p_2) + (1)(2p_1p_3) + (2)(p_1^2) \\ &= 2p_1^2 + 2p_1p_3 - 2p_2^2 - 2p_2p_3 \\ &= 2p_1(p_1 + p_3) - 2p_2(p_2 + p_3) \end{aligned}$$

Since $p_1 + p_2 + p_3 = 1$, we know $(p_1 + p_3) = 1 - p_2$ and $(p_2 + p_3) = 1 - p_1$. Substituting these in:

$$\begin{aligned} &= 2p_1(1 - p_2) - 2p_2(1 - p_1) \\ &= 2p_1 - 2p_1p_2 - 2p_2 + 2p_1p_2 \\ &= 2(p_1 - p_2) \end{aligned}$$

c) Compute $\mathbb{E}[X]$ again, but using Linearity of Expectation.

Let Y_1 be the amount you moved on the first step (taking values $-1, 0, 1$), and Y_2 the amount you moved on the second step. Then $X = Y_1 + Y_2$.

The expected value of a single step is:

$$\mathbb{E}[Y_1] = (-1)(p_2) + (0)(p_3) + (1)(p_1) = p_1 - p_2$$

Because the steps are identical, $\mathbb{E}[Y_2] = p_1 - p_2$.

By Linearity of Expectation:

$$\mathbb{E}[X] = \mathbb{E}[Y_1 + Y_2] = \mathbb{E}[Y_1] + \mathbb{E}[Y_2] = (p_1 - p_2) + (p_1 - p_2) = 2(p_1 - p_2)$$

8 Balls in Bins

Let X be the number of bins that remain empty when m balls are distributed into n bins randomly and independently. For each ball, each bin has an equal probability of being chosen. (Notice that two bins being empty are not independent events: if one bin is empty, that decreases the probability

that the second bin will also be empty. This is particularly obvious when $n = 2$ and $m > 0$.) Find $\mathbb{E}[X]$.

Step 1: Define Indicator Random Variables

For $i \in \{1, \dots, n\}$, let X_i be an indicator variable where $X_i = 1$ if bin i is empty, and 0 otherwise. The total number of empty bins is $X = \sum_{i=1}^n X_i$.

Step 2: Find Expectation of the Indicator

$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1)$. For bin i to be empty, *every single one* of the m balls must go into a bin other than bin i . The probability that a specific ball goes into a different bin is $(n - 1)/n$. Since the balls are distributed independently, we multiply this probability m times:

$$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1) = \left(\frac{n-1}{n}\right)^m$$

Step 3: Apply Linearity of Expectation

Despite the indicator variables being dependent on one another, LoE still holds:

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}[X_i] = n \cdot \left(\frac{n-1}{n}\right)^m$$

9 Fair Game?

You flip a fair coin independently and count the number of flips until the first tail, including that tail flip in the count. If the count is n , you receive 2^n dollars. What is the expected amount you will receive? How much would you be willing to pay at the start to play this game?

The expected reward is ∞ .

Let N be the number of flips until the first tail. The probability of getting exactly $n - 1$ heads followed by 1 tail is: $p_N(n) = \left(\frac{1}{2}\right)^{n-1} \cdot \frac{1}{2} = \frac{1}{2^n}$ for $n \in \{1, 2, \dots\}$.

We want to find the expected value of our reward, which is a function of N , specifically $g(N) = 2^N$. Using the Law of the Unconscious Statistician (LOTUS):

$$\mathbb{E}[2^N] = \sum_{n=1}^{\infty} 2^n \cdot p_N(n) = \sum_{n=1}^{\infty} 2^n \cdot \frac{1}{2^n} = \sum_{n=1}^{\infty} 1 = \infty$$

In theory, a fully rational actor aiming to maximize expected value should be willing to pay any finite amount of money to play this game (this is the famous St. Petersburg Paradox). However, in reality, you would likely be nervous to pay a lot. For instance, if you pay \$1000 to play, you will lose money unless the first 9 flips are all heads (which has a probability of less than 0.2%). With high probability you will lose your buy-in, and with an incredibly low probability you will win a massive amount of money.

10 Symmetric Difference

Suppose A and B are random, independent (possibly empty) subsets of $\{1, 2, \dots, n\}$, where each subset is equally likely to be chosen as A or B . Consider $A\Delta B = (A \cap B^C) \cup (B \cap A^C) = (A \cup B) \cap (A^C \cup B^C)$, i.e., the set containing elements that are in exactly one of A and B . Let X be the random variable that is the size of $A\Delta B$. What is $\mathbb{E}[X]$?

Step 1: Define Indicator Random Variables

For $i \in \{1, 2, \dots, n\}$, let X_i be the indicator variable where $X_i = 1$ if element $i \in A\Delta B$, and 0 otherwise. The total size of the set is $X = \sum_{i=1}^n X_i$.

Step 2: Find Expectation of the Indicator

$\mathbb{E}[X_i] = \mathbb{P}(X_i = 1)$. Because subsets are chosen uniformly at random, any element i has a $1/2$ probability of being in A , and independently, a $1/2$ probability of being in B . Element i is in $A\Delta B$ if it is in A but not B , OR in B but not A .

$$\begin{aligned}\mathbb{P}(X_i = 1) &= \mathbb{P}(i \in A)\mathbb{P}(i \notin B) + \mathbb{P}(i \notin A)\mathbb{P}(i \in B) \\ &= \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}\end{aligned}$$

Step 3: Apply Linearity of Expectation

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n \mathbb{E}[X_i] = \sum_{i=1}^n \frac{1}{2} = \frac{n}{2}$$

11 Practice

- Let X be a random variable with $p_X(k) = ck$ for $k \in \{1, \dots, 5\} = \Omega_X$, and 0 otherwise. Find the value of c that makes X follow a valid probability distribution and compute its mean and variance ($\mathbb{E}[X]$ and $\text{Var}(X)$).
- Let X be *any* random variable with mean $\mathbb{E}[X] = \mu$ and variance $\text{Var}(X) = \sigma^2$. Find the mean and variance of $Z = \frac{X - \mu}{\sigma}$. (When you're done, you'll see why we call this a "standardized" version of X !)
- Let X, Y be independent random variables. Find the mean and variance of $X - 3Y - 5$ in terms of $\mathbb{E}[X], \mathbb{E}[Y], \text{Var}(X)$, and $\text{Var}(Y)$.
- Let X_1, \dots, X_n be independent and identically distributed (iid) random variables each with mean μ and variance σ^2 . The sample mean is $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Find the mean and variance of \bar{X} . If you use the independence assumption anywhere, **explicitly label** at which step(s) it is necessary for your equalities to be true.

- For X to follow a valid probability distribution, its PMF must sum to 1: $\sum_{k \in \Omega_X} p_X(k) =$

1.

$$\sum_{k=1}^5 ck = c \sum_{k=1}^5 k = c \cdot (1 + 2 + 3 + 4 + 5) = 15c = 1$$

Therefore, $c = 1/15$.

We can now use the definition of expectation:

$$\mathbb{E}[X] = \sum_{k=1}^5 k \left(\frac{k}{15} \right) = \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2}{15} = \frac{1 + 4 + 9 + 16 + 25}{15} = \frac{55}{15} = \frac{11}{3} \approx 3.667$$

And compute $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = \sum_{k=1}^5 k^2 \left(\frac{k}{15} \right) = \frac{1^3 + 2^3 + 3^3 + 4^3 + 5^3}{15} = \frac{1 + 8 + 27 + 64 + 125}{15} = \frac{225}{15} = 15$$

And finally the variance of X :

$$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 15 - \left(\frac{11}{3} \right)^2 = \frac{135}{9} - \frac{121}{9} = \frac{14}{9} \approx 1.556$$

b) By Linearity of Expectation, we know that $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$.

$$\mathbb{E}[Z] = \mathbb{E} \left[\frac{X - \mu}{\sigma} \right] = \mathbb{E} \left[\frac{1}{\sigma} X - \frac{\mu}{\sigma} \right] = \frac{1}{\sigma} \mathbb{E}[X] - \frac{\mu}{\sigma}$$

Since we are given $\mathbb{E}[X] = \mu$:

$$\mathbb{E}[Z] = \frac{\mu}{\sigma} - \frac{\mu}{\sigma} = 0$$

For the variance, we know the property $\text{Var}(aX + b) = a^2 \text{Var}(X)$.

$$\text{Var}(Z) = \text{Var} \left(\frac{X - \mu}{\sigma} \right) = \text{Var} \left(\frac{1}{\sigma} X - \frac{\mu}{\sigma} \right) = \left(\frac{1}{\sigma} \right)^2 \text{Var}(X)$$

Since we are given $\text{Var}(X) = \sigma^2$:

$$\text{Var}(Z) = \frac{1}{\sigma^2} (\sigma^2) = 1$$

(This is why it's called standardized: it shifts the distribution to be centered exactly at 0, with a clean variance of exactly 1!)

c) Using Linearity of Expectation:

$$\mathbb{E}[X - 3Y - 5] = \mathbb{E}[X] - 3\mathbb{E}[Y] - 5$$

For variance, we know that the variance of a sum of *independent* random variables is the sum of their variances, and that $\text{Var}(aX + b) = a^2 \text{Var}(X)$.

$$\text{Var}(X - 3Y - 5) = \text{Var}(X) + \text{Var}(-3Y) = \text{Var}(X) + 9\text{Var}(Y)$$

d) Using Linearity of Expectation (which holds regardless of independence):

$$\mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} \sum_{i=1}^n \mathbb{E}[X_i] = \frac{1}{n}(n\mu) = \mu$$

Using the properties of variance, and specifically noting that **because the X_i 's are independent**, the variance of their sum is the sum of their variances:

$$\text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) \stackrel{\text{indep.}}{=} \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$$

12 Coin Flipping

Suppose we have a coin with probability p of heads. Suppose we flip this coin until we flip a head for the first time. Let X be the number of times we flip the coin *up to and including* the first head. What is $\mathbb{P}(X = k)$, for $k = 1, 2, \dots$? Verify that $\sum_{k=1}^{\infty} \mathbb{P}(X = k) = 1$, as it should. (You may use the fact that $\sum_{j=0}^{\infty} a^j = \frac{1}{1-a}$ for $|a| < 1$). (This is called a *geometric* random variable.)

If the k^{th} flip is our first head, the first $k-1$ flips must be tails (each with probability $(1-p)$), and the k^{th} flip must be a head (which occurs with probability p). Therefore:

$$\mathbb{P}(X = k) = (1-p)^{k-1}p$$

To verify the probabilities sum to 1, we pull the constant p out of the summation:

$$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = \sum_{k=1}^{\infty} (1-p)^{k-1}p = p \cdot \sum_{k=1}^{\infty} (1-p)^{k-1}$$

We can re-index the summation by setting $j = k-1$ (so our lower bound $k=1$ turns into $j=0$):

$$p \cdot \sum_{j=0}^{\infty} (1-p)^j = p \left(\frac{1}{1-(1-p)} \right) = p \left(\frac{1}{p} \right) = 1$$

13 More Coin Flipping ...

Suppose we have a coin with probability p of heads. Suppose we flip this coin n times independently. Let X be the number of heads that we observe. What is $\mathbb{P}(X = k)$, for $k = 0, \dots, n$? Verify that $\sum_{k=0}^n \mathbb{P}(X = k) = 1$, as it should. (This is called a "binomial" random variable.)

The answer is:

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

To see why, observe that for a specific, given sequence of flips containing exactly k heads, the probability of that sequence occurring is $p^k (1-p)^{n-k}$. Since there are exactly $\binom{n}{k}$ such possible valid sequences, the total probability of observing exactly k heads is $\binom{n}{k} p^k (1-p)^{n-k}$.

To verify the sum:

$$\sum_{k=0}^n \mathbb{P}(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k}$$

By the Binomial Theorem, this expression factors perfectly to:

$$(p + (1-p))^n = (1)^n = 1$$