Linearity and Variance

CSE 312 Summer 21
Lecture 10

Announcements

Sections will be merged.

Link updated on Ed. Use Section AB link.

Problem Set 2 grades should be up later today.

Real World 1 is due a week from today! 😥

Concept Check will be released later at night.

Linearity of Expectation

Linearity of Expectation

Linearity of Expectation

For any two random variables X and Y:

$$\mathbb{E}[\boldsymbol{X} + \boldsymbol{Y}] = \mathbb{E}[\boldsymbol{X}] + \mathbb{E}[\boldsymbol{Y}]$$

More generally, for random variables X and Y and scalars a, b and c :

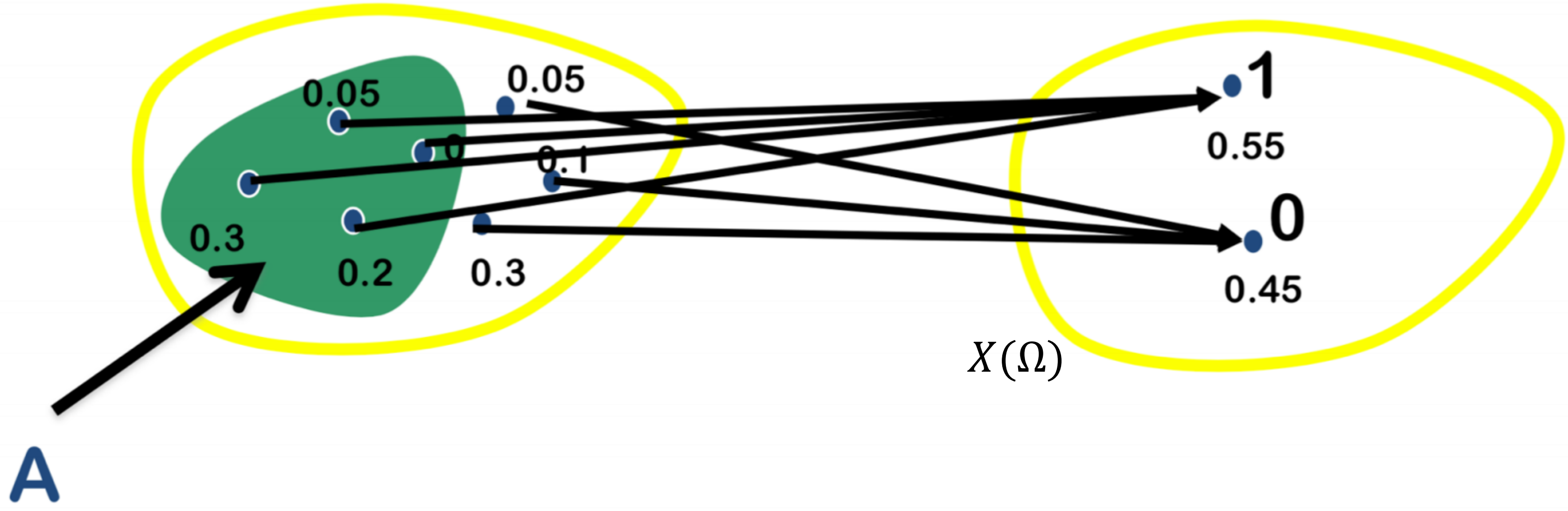
$$\mathbb{E}[a\boldsymbol{X} + b\boldsymbol{Y} + c] = a\mathbb{E}[\boldsymbol{X}] + b\mathbb{E}[\boldsymbol{Y}] + c$$

Indicator Random Variables

For any event A, we can define the indicator random variable X for A

$$X = \begin{cases} 1 & \text{if event A occurs} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{P}(X = 1) = \mathbb{P}(A)$$
$$\mathbb{P}(X = 0) = 1 - \mathbb{P}(A)$$



Computing complicated expectations

We often use these three steps to solve complicated expectations

1. Decompose: Finding the right way to decompose the random variable into sum of simple random variables
$$X = X_1 + X_2 + \cdots + X_n$$
2. LOE: Apply Linearity of Expectation
$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]$$
3. Conquer: Compute the expectation of each X_i

Often X_i are indicator random variables

Rotating the table

n people are sitting around a circular table. There is a name tag in each place. Nobody is sitting in front of their own name tag.

Rotate the table by a random number k of positions between 1 and n-1 (equally likely)

X is the number of people that end up in front of their own name tag. **Find $\mathbb{E}[X]$.**

<u>Decompose:</u>

<u>LOE:</u>

<u>Conquer:</u>

Rotating the table

n people are sitting around a circular table. There is a name tag in each place. Nobody is sitting in front of their own name tag.

Rotate the table by a random number k of positions between 1 and n-1 (equally likely)

X is the number of people that end up in front of their own name tag. **Find $\mathbb{E}[X]$.**

<u>Decompose:</u> Let us define X_i as follows:

$$X_i = \begin{cases} 1 & \text{if person } i \text{ sits infront of their own name tag} \\ 0 & \text{otherwise} \end{cases} \qquad X = \Sigma_{i=1}^{n} X_i$$

<u>LOE:</u>

$$\mathbb{E}[X] = \Sigma_{i=1}^{n} \mathbb{E}[X_i]$$

<u>Conquer:</u>

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

Pairs with the same birthday

In a class of m students, on average how many pairs of people have the same birthday?

Decompose:

LOE:

Conquer:

Fill out the poll everywhere so Kushal knows how long to explain
Go to pollev.com/cse312su21

Pairs with the same birthday

In a class of m students, on average how many pairs of people have the same birthday?

<u>Decompose:</u> Let us define X as the number of pairs with the same birthday

Let us define X_k as follows:

$$X_k = \begin{cases} 1 & \text{if the } k\text{th pair have the same birthday} \\ 0 & \text{otherwise} \end{cases} \qquad X = \Sigma_k^{\binom{m}{2}} X_k$$

<u>LOE:</u>

$$\mathbb{E}[X] = \Sigma_k^{\binom{m}{2}} \mathbb{E}[X_k]$$

<u>Conquer:</u>

$$\mathbb{E}[X_k] = P(X_k = 1) = \frac{365}{365 \cdot 365} = \frac{1}{365}$$

$$\mathbb{E}[X] = \binom{m}{2} \cdot \mathbb{E}[X_k] = \binom{m}{2} \cdot \frac{1}{365}$$

Shawarma Orders

If n Avengers order unique shawarmas, and if I randomly deliver the orders. How many Avengers get their own order?

Decompose:

LOE:

Conquer:

Fill out the poll everywhere so Kushal knows how long to explain
Go to pollev.com/cse312su21

Shawarma Orders

If n Avengers order unique shawarmas, and if I randomly deliver the orders. How many Avengers get their own order?

<u>Decompose:</u> Let us define X as the number of Avengers who get their own order

Let us define X_i as follows:

$$X_i = \begin{cases} 1 & \text{if the } i\text{th Avenger gets their own order} \\ 0 & \text{otherwise} \end{cases} \qquad X = \Sigma_{i=1}^{n} X_i$$

<u>LOE:</u>

$$\mathbb{E}[X] = \Sigma_{i=1}^{n} \mathbb{E}[X_i]$$

<u>Conquer:</u>

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

$$\mathbb{E}[X] = n \cdot \mathbb{E}[X_i] = 1$$

Linearity is special!

Linearity is special!

In general, $\mathbb{E}[g(X)] \neq g(\mathbb{E}[X])$

Example, for $X = \begin{cases} 1 & with\ prob\ 1/2 \\ -1 & \text{otherwise} \end{cases}$

- $\mathbb{E}[XY] \neq \mathbb{E}[X] \cdot \mathbb{E}[Y]$
- $\mathbb{E}\left[\frac{X}{Y}\right] \neq \frac{\mathbb{E}[X]}{\mathbb{E}[Y]}$
- $\mathbb{E}[X^2] \neq \mathbb{E}[X]^2$

So how do we calculate $\mathbb{E}[g(X)]$?

Expectation of $g(X)$

Expectation of $g(X)$

Given a discrete random variable X, the expectation or expected value of $g(X)$ is:

$$\mathbb{E}[g(X)] = \Sigma_{\omega \in \Omega}\, g\big(X(\omega)\big) \cdot \mathbb{P}(\omega)$$

or

$$\mathbb{E}[g(X)] = \Sigma_{x \in \mathrm{X}(\Omega)}\, g(x) \cdot \mathbb{P}(X = x)$$

Variance

Where are we?

A random variable is a way to summarize what outcome you saw.

The Expectation of a random variable is its average value.

A way to summarize a random variable

Expectation is **linear**

$\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

$X + Y$ is a random variable – it's a function that outputs a number given an outcome (or, here, a combination of outcomes).

Variance

Another one number summary of a random variable.

But wait, we already have expectation, what's this for?

A gamble to think about

Would you play either of these games?

I have a (biased) coin which flips heads with probability 1/3.

<u>Game 1:</u> I will flip the biased coin

- If a heads comes up, you win $2
- If a tails comes up, you pay me $1

<u>Game 2:</u> I will flip the biased coin

- If a heads comes up, you win $10
- If a tails comes up, you pay me $5

A gamble to think about

Would you play either of these games?

I have a (biased) coin which flips heads with probability 1/3.

Game 1: I will flip the biased coin

- If a heads comes up, you win \$2
- If a tails comes up, you pay me \$1

Game 2: I will flip the biased coin

- If a heads comes up, you win \$10
- If a tails comes up, you pay me \$5

You are probably thinking about the average profit you get from these, or if you can get a higher payout than me

A gamble to think about

Would you play either of these games?

<u>Game 1:</u> Win \$2, lose \$1

<u>Game 2:</u> Win \$10, lose \$5

Use random variables X_1 and X_2 to represent your profit from playing game 1 and game 2 respectively and then find the expected profit.

$$\mathbb{E}[X_1] = \mathbb{E}[X_2] = 0$$

Both games have an average profit of \$0. But, somehow one of the two is more enticing to some of you. Even with a \$0 average profit you believe that one of the games is more volatile.

What is the difference?

Expectation tells you what the average will be…

But it doesn’t tell you how “extreme” your results could be.

Nor how likely those extreme results are.

Game 2 has more extreme results when compared to Game 1.

In expectation they “cancel out” but if you can only play once…

…it would be nice to measure that.

Designing a Measure – Try 1

Well let's measure how far all the events are away from the center, and how likely they are

$$\sum_{\omega} \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])$$

What happens with Game 1?

$$\frac{1}{3} \cdot (2 - 0) + \frac{2}{3} \cdot (-1 - 0)$$
$$\frac{2}{3} - \frac{2}{3} = 0$$

What happens with Game 2?

$$\frac{1}{3} \cdot (10 - 0) + \frac{2}{3} \cdot (-5 - 0)$$
$$\frac{10}{3} - \frac{10}{3} = 0$$

Designing a Measure – Try 2

How do we prevent cancelling? Squaring makes everything positive.

$\sum_{\omega} \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])^2$

What happens with Game 1?

$$\frac{1}{3} \cdot (2 - 0)^2 + \frac{2}{3} \cdot (-1 - 0)^2$$
$$\frac{4}{3} + \frac{2}{3} = 2$$

What happens with Game 2?

$$\frac{1}{3} \cdot (10 - 0)^2 + \frac{2}{3} \cdot (-5 - 0)^2$$
$$\frac{100}{3} + \frac{50}{3} = 50$$

We say that X_2 has "higher variance" than X_1.

Why Squaring?

Why not absolute value? Or Fourth power?

Squaring is nicer algebraically.

Our goal with variance was to talk about the spread of results. Squaring makes extreme results even more extreme.

Fourth power over-emphasizes the extreme results (for our purposes).

Variance

Variance

The variance of a random variable X is

$$\mathrm{Var}(X) = \sum_{\omega} \mathbb{P}(\omega) \cdot (X(\omega) - \mathbb{E}[X])^2 = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

The first forms are the definition.

We can simplify the formula with an algebra trick.

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

Variance and Standard Deviation

Standard Deviation

The standard deviation of a random variable X is

$$\sigma(X) = \sqrt{Var(X)}$$

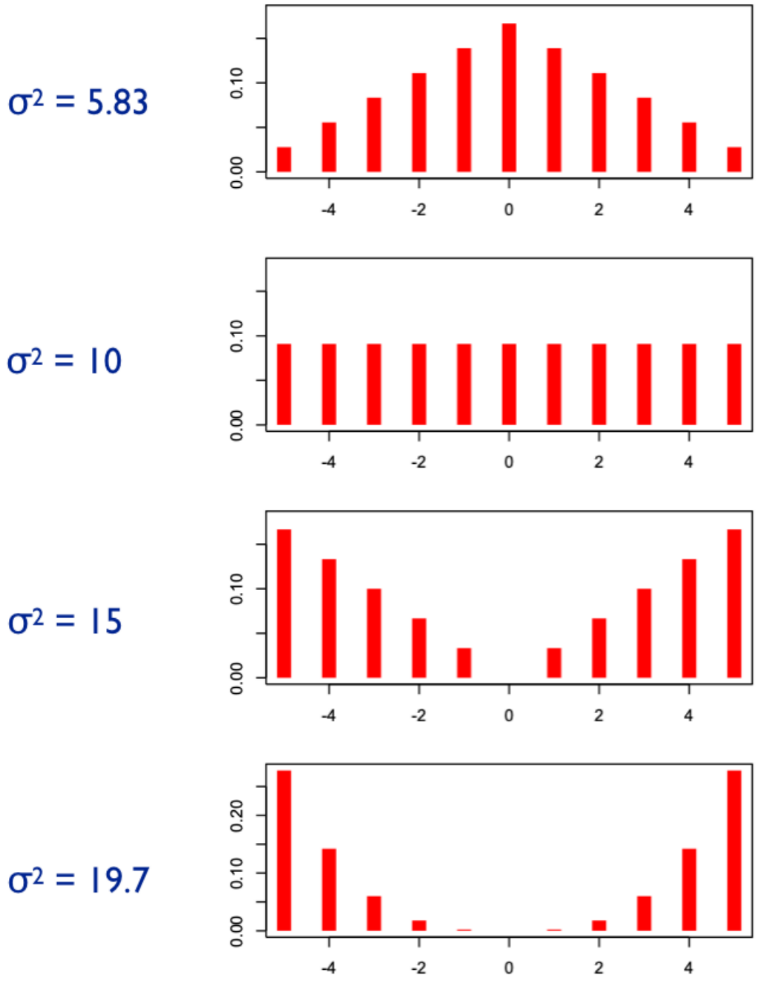
Variance (or standard deviation) is a quantity that measures, in expectation, how “far” the random variable is from its expectation.

Variance in Pictures

Captures how much “spread” there is in a pmf

All the PMFs graphed on the right have the same expectation of 0.

Expectation does not give the complete picture.



Variance of a die

Let X be the result of rolling a fair die.

$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}[(X - 3.5)^2]$

$= \frac{1}{6}(1 - 3.5)^2 + \frac{1}{6}(2 - 3.5)^2 + \frac{1}{6}(3 - 3.5)^2 + \frac{1}{6}(4 - 3.5)^2 + \frac{1}{6}(5 - 3.5)^2 + \frac{1}{6}(6 - 3.5)^2$

$= \frac{35}{12} \approx 2.92.$

Or $\mathbb{E}\left[X^2\right] - (E[X])^2 = \sum_{k=1}^{6} \frac{1}{6} \cdot k^2 - 3.5^2 = \frac{91}{6} - 3.5^2 \approx 2.92$

Variance of n Coin Flips

Flip a coin n times, where it comes up heads with probability p each time (independently). Let X be the total number of heads.

We saw last time $\mathbb{E}[X] = np$.

$$X_i = \begin{cases} 1 & \text{if flip } i \text{ is heads} \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} p = np.$$

Variance of n Coin Flips

Flip a coin n times, where it comes up heads with probability p each time (independently). Let X be the total number of heads.

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

$$\mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \sum_{\omega} \mathbb{P}(\omega)(X(\omega) - np)^2$$

$$= \sum_{k=0}^{n} \binom{n}{k} \cdot p^k (1-p)^{n-k} \cdot (k - np)^2$$

Algebra time?

Variance

If X and Y are independent, then
$$\mathbf{Var}(X+Y) = \mathbf{Var}(X) + \mathbf{Var}(Y)$$

We'll talk about what it means for random variables to be independent soon…

For now, in this problem X_i is independent of X_j for $i \neq j$ where

$$X_i = \begin{cases} 1 & \text{if flip } i \text{ was heads} \\ 0 & \text{otherwise} \end{cases}$$

Variance of n Coin Flips

$\mathrm{Var}(X) = \mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$

What's the $\mathrm{Var}(X_i)$?

$\mathbb{E}\left[(X_i - \mathbb{E}[X_i])^2\right]$

$= \mathbb{E}[(X_i - p)^2]$

$= p(1-p)^2 + (1-p)(0-p)^2$

$= p(1-p)[(1-p)+p] = p(1-p).$

OR

$Var(X_i) = \mathbb{E}\left[X_i^2\right] - \mathbb{E}[X_i]^2 = \mathbb{E}\left[X_i^2\right] - p^2 = p - p^2 = p(1-p)$.

Plugging In

$\text{Var}(X) = \text{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \text{Var}(X_i)$

What's the $\text{Var}(X_i)$?

$p(1-p)$.

$\text{Var}(X) = \sum_{i=1}^{n} p(1-p) = np(1-p)$.

Variance simplification algebra trick

$\mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2 - 2X\mathbb{E}[X] + (\mathbb{E}[X])^2\right]$ expanding the square

$= \mathbb{E}\left[X^2\right] - \mathbb{E}\left[2X\mathbb{E}[X]\right] + \mathbb{E}[(\mathbb{E}[X])^2]$ linearity of expectation.

$= \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]\mathbb{E}[X] + \mathbb{E}[(\mathbb{E}[X])^2]$ linearity of expectation.

$= \mathbb{E}\left[X^2\right] - 2\mathbb{E}[X]\mathbb{E}[X] + (\mathbb{E}[X])^2$ expectation of a constant is the constant

$= \mathbb{E}\left[X^2\right] - 2(\mathbb{E}[X])^2 + (\mathbb{E}[X])^2$

$= \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$

So, $\mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - (\mathbb{E}[X])^2$.

Facts about Variance

Facts about Variance

$\mathrm{Var}(X + c) = \mathrm{Var}(X)$

Proof:

$$\mathrm{Var}(X + c) = \mathbb{E}\left[(X + c)^2\right] - \mathbb{E}[X + c]^2$$

$$= \mathbb{E}\left[X^2\right] + \mathbb{E}[2Xc] + \mathbb{E}\left[c^2\right] - (\mathbb{E}[X] + c)^2$$

$$= \mathbb{E}\left[X^2\right] + 2c\mathbb{E}[X] + c^2 - \mathbb{E}[X]^2 - 2c\mathbb{E}[X] - c^2$$

$$= \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$$

$$= \mathrm{Var}(X)$$

Facts about Variance

$\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$

Proof:

$$Var(aX) = \mathbb{E}\left[(aX)^2\right] - (\mathbb{E}[aX])^2$$

$$= a^2\mathbb{E}\left[X^2\right] - (a\mathbb{E}[X])^2$$

$$= a^2\mathbb{E}\left[X^2\right] - a^2\mathbb{E}[X]^2$$

$$= a^2\left(\mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2\right)$$

$$= a^2Var(X)$$

More Practice: Linearity

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_R, to the left with probability p_L, and doesn't move with probability p_S, where $p_L + p_R + p_S = 1$. After 2 seconds, let X be the location of the frog. Find $\mathbb{E}[X]$.

Frogger – Brute Force

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_R, to the left with probability p_L, and doesn't move with probability p_S, where $p_L + p_R + p_S = 1$. After 2 seconds, let X be the location of the frog. **Find $\mathbb{E}[X]$.**

$$p_X(x) = \begin{cases} p_L^2 & x = -2 \\ 2p_L p_S & x = -1 \\ 2p_L p_R + p_s^2 & x = 0 \\ 2p_R p_S & x = 1 \\ p_R^2 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E}[\boldsymbol{X}] = \Sigma_\omega P(\omega)X(\omega) = (-2)p_L^2 + (-1)2p_L p_S + 0 \cdot \left(2p_L p_R + p_s^2\right) + (1)2p_R p_S + (2)p_R^2 = \mathbf{2}(\boldsymbol{p_R} - \boldsymbol{p_L})$$

Frogger – LOE

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_R, to the left with probability p_L, and doesn't move with probability p_S, where $p_L + p_R + p_S = 1$. After 2 seconds, let X be the location of the frog. **Find $\mathbb{E}[X]$.**

Let us define X_i as follows:

$$X_i = \begin{cases} -1 & \text{if the frog moved left on the } i\text{th step} \\ 0 & \text{otherwise} \\ 1 & \text{if the frog moved right on the } i\text{th step} \end{cases}$$

$$\mathbb{E}[X_i] = -1 \cdot p_L + 1 \cdot p_R + 0 \cdot p_S = (p_R - p_L)$$

By Linearity of Expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{2} X_i\right] = \sum_{i=1}^{2} \mathbb{E}[X_i] = \mathbf{2(p_R - p_L)}$$