Linearity of Expectation

CSE 312 Su23
Lecture 8

LOTUS

Expectation

Expectation

The "expectation" (or "expected value") of a random variable X is:

$$\mathbb{E}[X] = \sum_{k} k \cdot \underbrace{\mathbb{P}(X = k)}$$

Intuition: The weighted average of values X could take on.

Weighted by the probability you actually see them.

How do we compute $\mathbb{E}[g(X)]$?

Law of the Unconscious Statistician (LOTUS)

Given a discrete RV $X: \Omega \to \mathbb{R}$, then the expected value of g(X) is

$$\mathbb{E}[g(X)] = \sum_{x \in \Omega_X} g(x) \cdot \mathbb{P}(X = x) = \sum_{x \in \Omega_X} g(x) \cdot p_X(x)$$

$$\mathbb{E}[X^2] = \sum_{x \in \Omega_X} x^2 \cdot \mathbb{P}(X = x)$$

Example

Suppose we roll a fair, 6-sided die. You win the cube of the number rolled, times 10, in dollars. How much do you expect to win?

Let X be the result of the die roll.

$Y = 10X^3$

LOTUS

$$\mathbb{E}[Y] = \mathbb{E}[10X^3] = \sum_{x \in \Omega_X} 10x^3 \cdot \mathbb{P}(X = x)$$

Example

Suppose we roll a fair, 6-sided die. You win the cube of the number rolled, times 10, in dollars. How much do you expect to win?

Let X be the result of the die roll. Then, our winnings are $10X^3$.

$\mathbb{E}[10X^3] = \sum_{x=1}^{6} 10x^3 \cdot \mathbb{P}(X = x)$

$= 10 \cdot \frac{1}{6} + 10 \cdot 2^3 \cdot \frac{1}{6} + 10 \cdot 3^3 \cdot \frac{1}{6} + 10 \cdot 4^3 \cdot \frac{1}{6} + 10 \cdot 5^3 \cdot \frac{1}{6} + 10 \cdot 6^3 \cdot \frac{1}{6}$

Independence of RVs

Independence

Independence

Two events A, B are independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

You'll sometimes see this called "statistical independence" to emphasize that we're talking about probabilities (not, say, physical interactions).

Equivalently,

If $\mathbb{P}(B) \neq 0,\ \mathbb{P}(A|B) = \mathbb{P}(A)$.

If $\mathbb{P}(A) \neq 0,\ \mathbb{P}(B|A) = \mathbb{P}(B)$

Independence for 3 or more events

For three or more events, we need two kinds of independence

Pairwise Independence

Events $A_1, A_2, \ldots, A_n$ are pairwise independent if
$\mathbb{P}(\boldsymbol{A_i} \cap \boldsymbol{A_j}) = \mathbb{P}(\boldsymbol{A_i}) \cdot \mathbb{P}(\boldsymbol{A_j})$ for all $\boldsymbol{i}, \boldsymbol{j},\ \boldsymbol{i} \neq \boldsymbol{j}$

Mutual Independence

Events $A_1, A_2, \ldots, A_n$ are mutually independent if
$\mathbb{P}(\boldsymbol{A_{i_1}} \cap \boldsymbol{A_{i_2}} \cap \cdots \cap \boldsymbol{A_{i_k}}) = \mathbb{P}(\boldsymbol{A_{i_1}}) \cdot \mathbb{P}(\boldsymbol{A_{i_2}}) \cdots \mathbb{P}(\boldsymbol{A_{i_k}})$
for every subset $\{\boldsymbol{i_1}, \boldsymbol{i_2}, \ldots, \boldsymbol{i_k}\}$ of $\{\boldsymbol{1}, \boldsymbol{2}, \ldots, \boldsymbol{n}\}$.

Independence of Random Variables

That's for events… what about random variables?

Independence (of random variables)

X and Y are independent if for all k, ℓ

$$\mathbb{P}(X = k, Y = \ell) = \mathbb{P}(X = k)\mathbb{P}(Y = \ell)$$

We'll often use commas instead of ∩ symbol.

Independence of Random Variables

The "for all values" is important.

Define $S =$"the sum of two dice" and $R =$"the value of the red die"

$$\mathbb{P}(S = 7, R = 5) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}(S = 7)\mathbb{P}(R = 5)$$

$(S = 7, R = 5)$: {(5,2)}

$(S = 7)$: {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)}

$(R = 5)$: {(5,1), (5,2), (5,3), (5,4), (5,5), (5,6)}

Independence of Random Variables

The "for all values" is important.

Define S ="the sum of two dice" and R ="the value of the red die"

But,

$\mathbb{P}(S = 2, R = 5) = 0 \neq \frac{1}{36} \cdot \frac{1}{6} = \mathbb{P}(S = 2)\mathbb{P}(R = 5)$ (for example)

$(S = 2)$: {(1,1)}

$(R = 5)$: {(5,1), (5,2), (5,3), (5,4), (5,5), (5,6)}

NOT independent.

Mutual Independence for RVs

A little simpler to write down than for events

Mutual Independence (of random variables)

$X_1, X_2, \ldots, X_n$ **are mutually independent if for all** $x_1, x_2, \ldots, x_n$

$$\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = \underbrace{\mathbb{P}(X_1 = x_1)\mathbb{P}(X_2 = x_2) \cdots \mathbb{P}(X_n = x_n)}$$

DON'T need to check all subsets for random variables…

But you do need to check all values (all possible x_i) still.

Why do we care?

Independent random variables have some special properties! We'll look at a lot of these in the next few days...

Linearity of Expectation

Linearity of Expectation

Linearity of Expectation

For any two random variables X and Y:

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

Note: X and Y do not have to be independent

Extending this to n random variables, $X_1, X_2, \ldots, X_n$

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

This can be proven by induction.

Linearity of Expectation - Proof

Linearity of Expectation

For any two random variables X and Y:

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

Note: X and Y do not have to be independent

Proof:

$\mathbb{E}[X+Y] = \Sigma_\omega P(\omega)(X(\omega)+Y(\omega))$

$= \Sigma_\omega P(\omega)X(\omega) + P(\omega)Y(\omega)$

$= \Sigma_\omega P(\omega)X(\omega) + \Sigma_\omega P(\omega)Y(\omega)$

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

defn $\mathbb{E}[X]$

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

Linearity of Expectation

Linearity of Expectation

For any two random variables X and Y:

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

Note: X and Y do not have to be independent

Constants are also fine:

$\mathbb{E}[c]$

For real numbers a, b, c

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

$$\mathbb{E}[X^2 + Y] = \mathbb{E}[X^2] + \mathbb{E}[Y]$$

Fishy Business

Say you and your friend go fishing everyday.

- You catch X fish, with $\mathbb{E}[X] = 3$
- Your friend catches Y fish, with $\mathbb{E}[Y] = 7$

- How many fish do both of you bring on an average day?

$$Z = X + Y$$

$$\mathbb{E}[Z] = \mathbb{E}[X+Y] \overset{\text{LOE}}{=} \mathbb{E}[X] + \mathbb{E}[Y]$$

$$= 3 + 7 = 10$$

Fishy Business

Say you and your friend go fishing everyday.

- You catch X fish, with $\mathbb{E}[X] = 3$
- Your friend catches Y fish, with $\mathbb{E}[Y] = 7$

- How many fish do both of you bring on an average day?

Let Z be the r.v. representing the total number of fish you both catch

$$\mathbb{E}[Z] = \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 3 + 7 = 10$$

Fishy Business

Say you and your friend go fishing everyday.

- You catch X fish, with $\mathbb{E}[X] = 3$
- Your friend catches Y fish, with $\mathbb{E}[Y] = 7$

- How many fish do both of you bring on an average day?

Let Z be the r.v. representing the total number of fish you both catch

z

$$\mathbb{E}[Z] = \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 3 + 7 = 10$$

- You can sell each for $10 per fish, but you need $15 (total) for expenses. What is your average profit?

$$A = 10Z - 15$$

$$\mathbb{E}[A] = \mathbb{E}[10Z - 15] = 10\,\mathbb{E}[Z] - 15$$

Fishy Business

Say you and your friend go fishing everyday.

- You catch X fish, with $\mathbb{E}[X] = 3$
- Your friend catches Y fish, with $\mathbb{E}[Y] = 7$

- How many fish do both of you bring on an average day?

Let Z be the r.v. representing the total number of fish you both catch

$$\mathbb{E}[Z] = \mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 3 + 7 = 10$$

- You can sell each for \$10 per fish, but you need \$15 (total) for expenses. What is your average profit?

$$\mathbb{E}[10Z - 15] = 10\mathbb{E}[Z] - 15 = 100 - 15 = 85$$

Coin Tosses

If we flip a coin twice, what is the expected number of heads that come up?

Coin Tosses

If we flip a coin twice, what is the expected number of heads that come up?

Let X be the r.v. representing the total number of heads

$$p_X(x) = \begin{cases} \frac{1}{4} & \text{if } x = 0 \quad \{TT\} \\ \frac{1}{2} & \text{if } x = 1 \quad \{HT, TH\} \\ \frac{1}{4} & \text{if } x = 2 \quad \{HH\} \end{cases}$$

Coin Tosses

If we flip a coin twice, what is the expected number of heads that come up?

Let X be the r.v. representing the total number of heads

$$p_X(x) = \begin{cases} \frac{1}{4} & \text{if } x = 0 \\ \frac{1}{2} & \text{if } x = 1 \\ \frac{1}{4} & \text{if } x = 2 \end{cases}$$

$$\mathbb{E}[X] = \Sigma_k \mathbb{P}(X = k) \cdot k = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 = 1$$

Repeated Coin Tosses

Now what if the probability of flipping a heads was p and that we wanted to find the total number of heads flipped when we flip the coin n times?

Let Y be the r.v. representing the total number of heads.

Make a prediction --- what should $\mathbb{E}[Y]$ be?

$\frac{1}{2} \cdot 2 = 1$

np

Repeated Coin Tosses

Now what if the probability of flipping a heads was p and that we wanted to find the total number of heads flipped when we flip the coin n times?

Let Y be the r.v. representing the total number of heads.

$$\mathbb{E}[Y] = \sum_{k=0}^{n} k \cdot \mathbb{P}(Y = k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

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

Ok, but what actually is it?
I don't have intuition for this formula.

Repeated Coin Tosses

Now what if the probability of flipping a heads was p and that we wanted to find the total number of heads flipped when we flip the coin n times?

$$\mathbb{E}[Y] = \sum_{k=0}^{n} k \cdot \mathbb{P}(Y = k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

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

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

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

$$= np(p + (1-p))^{n-1} = np$$

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

Binomial Theorem

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$

We did it! And all it took was a clever application of the binomial theorem, setup by a very non-obvious application of an obscure combinatorial identity. Ezpz.

Repeated Coin Tosses

Now what if the probability of flipping a heads was p and that we wanted to find the total number of heads flipped when we flip the coin n times?

$$\mathbb{E}[Y] = \sum_{k=0}^{n} k \cdot \mathbb{P}(Y = k) = \sum_{k=0}^{n} k \cdot \binom{n}{\cdot}$$

$$= \sum_{k=1}^{n} k \cdot \binom{n}{\cdot}$$

$$\binom{\cdot}{k-1} p^k (1-p)^{n-k}$$

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

$$= np(p + (1-p))^{n-1} = np$$

$$k\binom{n}{k} = n\binom{n-1}{k-1}$$

Binomial Theorem!

We did it! And all it took was a clever application of the binomial theorem, setup by a very non-obvious application of an obscure combinatorial identity. Ezpz.

No one wants to do proofs like this every time!

Indicator Random Variables

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

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

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

$\Omega_{X_A} = \{0, 1\}$

$$\mathbb{P}(X=x)\ p_X(x) = \underbrace{\begin{cases} \mathbb{P}(A) & \text{if } x = 1 \\ 1 - \mathbb{P}(A) & \text{if } x = 0 \\ 0 & \text{otherwise} \end{cases}}$$

$\mathbb{P}(X=2)$

Repeated Coin Tosses

The probability of flipping a heads is p and we want to find the total number of heads flipped when we flip the coin n times?

Let X be the total number of heads

What indicators can we define? What 'Booleans' have enough information to combine (add) and solve the problem?

H H T T H

1 1 0 0 1 = 3

X: total n flips

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip H} \\ 0 & \text{o.w.} \end{cases}$$

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

$$X = \sum_{i=1}^{n} X_i$$

Repeated Coin Tosses

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

The probability of flipping a heads is p and we want to find the total number of heads flipped when we flip the coin n times?

Let X be the total number of heads

Define X_i as follows:

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

$\mathbb{P}(X_i = 1) = p$

$\mathbb{P}(X_i = 0) = 1 - p$

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

$\Omega_{X_i} = \{0, 1\}$

def $\mathbb{E}[X]$

Repeated Coin Tosses

The probability of flipping a heads is p and we want to find the total number of heads flipped when we flip the coin n times?

Let X be the total number of heads

Define X_i as follows:

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

$$\mathbb{P}(X_i = 1) = p$$
$$\mathbb{P}(X_i = 0) = 1 - p$$

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

$$X = \sum_{i=1}^{n} X_i$$

Repeated Coin Tosses

The probability of flipping a heads is p and we want to find the total number of heads flipped when we flip the coin n times?

Let X be the total number of heads

Define X_i as follows:

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

$$\mathbb{P}(X_i = 1) = p$$
$$\mathbb{P}(X_i = 0) = 1 - p$$

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

By Linearity of Expectation,

$$\mathbb{E}[X] = \mathbb{E}[\textstyle\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \mathrm{n} \cdot \mathbb{E}[X_i] = np$$

Computing complicated expectations

We often use these three steps to solve complicated expectations

1. <u>Decompose</u>: Finding the right way to decompose the random variable into sum of simple random variables

$$X = X_1 + X_2 + \cdots + X_n$$

2. <u>LOE</u>: Apply Linearity of Expectation

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n]$$

3. <u>Conquer</u>: Compute the expectation of each X_i

Often, X_i are indicator random variables.

I, D

Pairs with the same birthday

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

Property.

$$X_i = \begin{cases} 1 & \text{if ith pair same bday} \\ 0 & \text{otherwise} \end{cases}$$

Decompose:

$$X = \sum_{i=1}^{\binom{m}{2}} X_i$$

LOE:

Conquer:

Pairs with the same birthday

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

Decompose:

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

LOE:

LOE

$$= \sum_{i=1}^{\binom{m}{2}} \underbrace{\mathbb{E}[X_i]}$$

Conquer:

$$= \sum_{i=1}^{\binom{m}{2}} \frac{1}{365} \quad \text{(next slide)} \quad = \binom{m}{2}\frac{1}{365}$$

$$\underbrace{\frac{1}{365} + \cdots + \frac{1}{365}}_{\binom{m}{2}}$$

Pairs with the same birthday

$X_i = \begin{cases} 1 & \text{if } i\text{th pair same} \\ 0 & \text{otherwise} \end{cases}$

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

$$\mathbb{E}[X_i] = \cancel{0 \cdot \mathbb{P}(X_i = 0)} + 1 \cdot \mathbb{P}(X_i = 1)$$

Decompose:

$$= \mathbb{P}(X_i = 1)$$

LOE:

$$= \mathbb{P}(\text{same bday})$$

Conquer:

$$= \frac{365 \cdot 1}{365 \cdot 365} = \frac{1}{365}$$

Pairs with the same birthday

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

Decompose: Let X be the number of pairs with the same birthday

Define X_{ij} as follows:

$$X_{ij} = \begin{cases} 1 & \text{if person i, j have the same bithday} \\ 0 & \text{otherwise} \end{cases}$$

$$X = \Sigma_{1 \le i < j \le m} X_{ij}$$

LOE:

$$\mathbb{E}[X] = \mathbb{E}\left[\Sigma_{1 \le i < j \le m} X_{ij}\right] = \Sigma_{1 \le i < j \le m} \mathbb{E}\left[X_{ij}\right]$$

Conquer:

$$\mathbb{E}\left[X_{ij}\right] = P\left(X_{ij} = 1\right) = \frac{365}{365 \cdot 365} = \frac{1}{365}$$

$$\mathbb{E}[X] = \binom{m}{2} \cdot \mathbb{E}\left[X_{ij}\right] = \binom{m}{2} \cdot \frac{1}{365}$$

Rotating the table $1 \le i \le n$

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)

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

$X_i = \begin{cases} 1 \\ \end{cases}$ ith person in front of nametag

Rotating the table

$$\frac{n}{n-1}$$

n people are sitting around a circular table. 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)

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

Decompose:

$$X_i = \begin{cases} 1 & \text{ith person in front} \\ 0 & \text{o.w.} \end{cases} \qquad X = \sum_{i=1}^{n} X_i$$

LOE:

Conquer:

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

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

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 $\mathbb{E}[X]$.

Decompose: Define X_i as follows:

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

Note: $X = \Sigma_{i=1}^{n} X_i$

LOE:

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

Conquer:

These X_i are not independent!
That's ok!!

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 $\mathbb{E}[X]$.

Decompose: Define X_i as follows:

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

LOE:

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

Conquer:

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