Please download the activity slide for today! ☺

Expectation

CSE 312 Summer 21
Lecture 9

Announcements

Office Hours:

Kushal will be doing Tuesday Office Hours at 7 pm.

Justin will be doing Wednesday Office Hours at 7 pm.

Links will be updated on the calendar and on the pinned Ed post.

Probability Mass Function

Try It Yourself

There are 20 balls, numbered 1,2,…,20 in an urn.

You'll draw out a size-three subset. (i.e. without replacement)

$\Omega = \{$size three subsets of $\{1, \dots, 20\}$ $\}$, $\mathbb{P}()$ is uniform measure.

Let X be the largest value among the three balls.

If outcome is $\{4,2,10\}$ then $X = 10$.

Write down the pmf of X.

Try It Yourself

There are 20 balls, numbered 1,2,…,20 in an urn.

You'll draw out a size-three subset. (i.e. without replacement)

Let X be the largest value among the three balls.

$$p_X(x) = \begin{cases} \binom{x-1}{2} / \binom{20}{3} & \text{if } x \in \mathbb{N},\ 3 \leq x \leq 20 \\ 0 & \text{otherwise} \end{cases}$$

Good check: if you sum up $p_X(x)$ do you get 1?

Good check: is $p_X(x) \geq 0$ for all x? Is it defined for all x?

Cumulative Distribution Function

Describing a Random Variable

The most common way to describe a random variable is the PMF.

But there's a second representation:

The cumulative distribution function (CDF) gives the probability $X \leq x$

More formally, $\mathbb{P}(\{\omega: X(\omega) \leq x\})$

Often written $F_X(x) = \mathbb{P}(X \leq x)$

$$F_X(x) = \sum_{i: i \leq x} p_X(i)$$

Try It Yourself

What is the CDF of X where X be the largest value among the three balls? (Drawing 3 of the 20 without replacement)

$$P(X \le 10) = \frac{\binom{10}{3}}{\binom{20}{3}}$$

$$p_X(x) = \begin{cases} \binom{x-1}{2} \Big/ \binom{20}{3} & x \in \mathbb{N},\ 3 \le x \le 20 \\ 0 & \text{otherwise} \end{cases}$$

$$P(X = 10) = \frac{\binom{10-1}{2}}{\binom{20}{3}}$$

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

Try It Yourself

$F_X(20) = \binom{20}{3} / \binom{20}{3} = 1$

What is the CDF of X where X be the largest value among the three balls? (Drawing 3 of the 20 without replacement)

$\lfloor x \rfloor$ - floor of x

$$F_X(x) = \begin{cases} 0 & \text{if } x < 3 \\ \binom{\lfloor x \rfloor}{3} / \binom{20}{3} & \text{if } 3 \leq x \leq 20 \\ 1 & \text{otherwise} \end{cases}$$

$P(X \leq 10.5) = P(X \leq 10)$

Try It Yourself

What is the CDF of X where X be the largest value among the three balls? (Drawing 3 of the 20 without replacement)

$$F_X(x) = \begin{cases} 0 & \text{if } x < 3 \\ \binom{\lfloor x \rfloor}{3} / \binom{20}{3} & \text{if } 3 \leq x \leq 20 \\ 1 & \text{otherwise} \end{cases}$$

Good checks: Is $F_X(-\infty) = 0$? Is $F_X(\infty) = 1$? If not, something is wrong.

Is $F_X(x)$ increasing? If not, something is wrong.

Is $F_X(x)$ defined for all real number inputs? If not, something is wrong.

Two descriptions

PROBABILITY MASS FUNCTION

Defined for all $\mathbb{R}$ inputs.

Usually has "0 otherwise" as an extra case.

$\sum_x p_X(x) = 1$

$0 \leq p_X(x) \leq 1$

$\sum_{z:z\leq x} p_X(z) = F_X(x)$

CUMULATIVE DISTRIBUTION FUNCTION

Defined for all $\mathbb{R}$ inputs.

Usually has "0 otherwise" and 1 otherwise" extra cases

Non-decreasing function

$0 \leq F_X(x) \leq 1$

$\lim_{x\to-\infty} F_X(x) = 0$

$\lim_{x\to\infty} F_X(x) = 1$

Expectation

Expectation

Expectation

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

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

Intuition:

The weighted average of values that X can take on, weighted by the probability you see them.

Coin Tosses

Flip a fair coin twice (independently)

Let X be the number of heads.

$\Omega = \{TT, TH, HT, HH\}$, $\mathbb{P}()$ is a uniform measure.

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

$$0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4}$$

$$= 1$$

$$\mathbb{E}[X] = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 = 0 + \frac{1}{2} + \frac{1}{2} = 1$$

Biased Die Rolls

We roll a biased die such that it shows a 6 with probability $\frac{1}{3}$, and values 1,2,…,5 each with probability $\frac{2}{15}$.

6 ?

3 ?

4 ?

Let X be the value of the die. What is $\mathbb{E}[X]$?

$$\frac{1}{3} \cdot 6 + \frac{2}{15} \cdot 5 + \frac{2}{15} \cdot 4 + \frac{2}{15} \cdot 3 + \frac{2}{15} \cdot 2 + \frac{2}{15} \cdot 1$$

$$= 2 + \frac{2}{15} \cdot (5 + 4 + 3 + 2 + 1) = 2 + \frac{30}{15} = 4$$

$\mathbb{E}[X]$ is not just the most likely outcome!

Activity!

Let X be the result of the roll of a fair die. What is $\mathbb{E}[X]$?

Let Y be the sum of two (independent) die rolls. What is $\mathbb{E}[Y]$?

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

Activity!

3.5

3.5

Let X be the result of the roll of a fair die. What is $\mathbb{E}[X]$?

3.5

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

$$= \frac{21}{6} = 3.5$$

$\mathbb{E}[X]$ is not necessarily a possible outcome from our samples.

This is fine as the expectation is an average and not exact value.

Activity!

Let Y be the sum of two (independent) die rolls. What is $\mathbb{E}[Y]$?

$$\frac{1}{36}\cdot 2+\frac{2}{36}\cdot 3+\frac{3}{36}\cdot 4+\frac{4}{36}\cdot 5+\frac{5}{36}\cdot 6+\frac{6}{36}\cdot 7+\frac{5}{36}\cdot 8+\frac{4}{36}\cdot 9+\frac{3}{36}\cdot 10+\frac{2}{36}\cdot 11+\frac{1}{36}\cdot 12$$

$$= 7$$

Do you see a relation between $\mathbb{E}[X]$ and $\mathbb{E}[Y]$?

$Y = 2X$

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

Important Note

$X \to$ random variable.

$\mathbb{E}[X]$ is not random. It is a number.

You do not need to run an experiment to know what it is.

In the die roll example, on experimentation the random variable X can take values $\{1, 2, 3, 4, 5, 6\}$, whereas we know before any roll of the die that the expected value of the die roll will be 3.5

Linearity of Expectation

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

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

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

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

This can be proven by induction.

Linearity of Expectation - Proof

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

$$\begin{aligned}\mathbb{E}[\boldsymbol{X} + \boldsymbol{Y}] &= \Sigma_\omega P(\omega)(X(\omega) + Y(\omega)) \\ &= \Sigma_\omega P(\omega)X(\omega) + \Sigma_\omega P(\omega)Y(\omega) \\ &= \mathbb{E}[\boldsymbol{X}] + \mathbb{E}[\boldsymbol{Y}]\end{aligned}$$

$$E[X] = \sum_{\omega \in \Omega} P(\omega) \cdot X(\omega)$$

$$E[X] = \sum_{k \in X(\Omega)} k \cdot P(X = k)$$

Linearity of Expectation

Linearity of Expectation

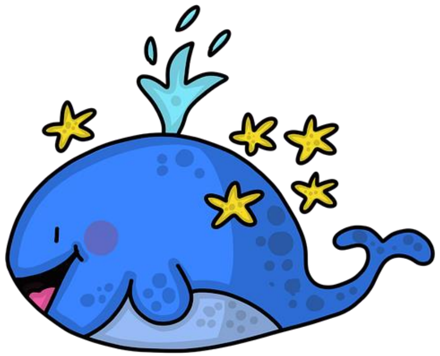
For any two random variables X and Y:

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

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

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

Fishy Business

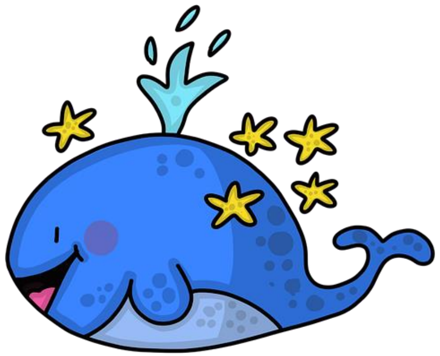


Say you and your friend go fishing everyday.

- 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?

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 $\mathbb{E}[Y] = 7$

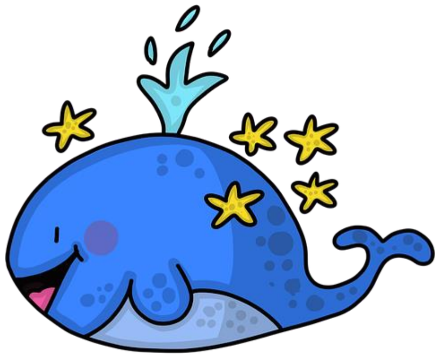
$Z = X + Y$

- 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?

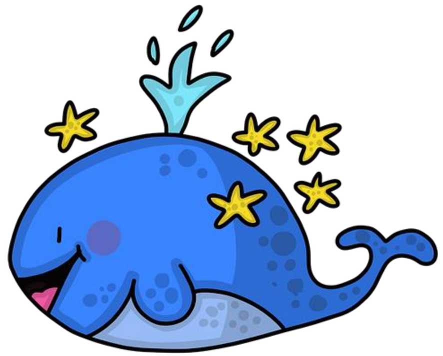
Let Z be the r.v. representing the total number of fish you both catch

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

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

$$E[10Z - 15] = 10E[Z] - 15 = 100 - 15 = 85$$

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, but you need $15 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

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

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

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

$$\mathbb{E}[\boldsymbol{X}] = \Sigma_\omega P(\omega)X(\omega) = \frac{1}{4} \cdot 0 + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 = \mathbf{1}$$

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?

If Y is the r.v. representing the total number of heads that come up.

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

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

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?

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

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

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k} \qquad \left[k\binom{n}{k} = n\binom{n-1}{k-1}\right]$$

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

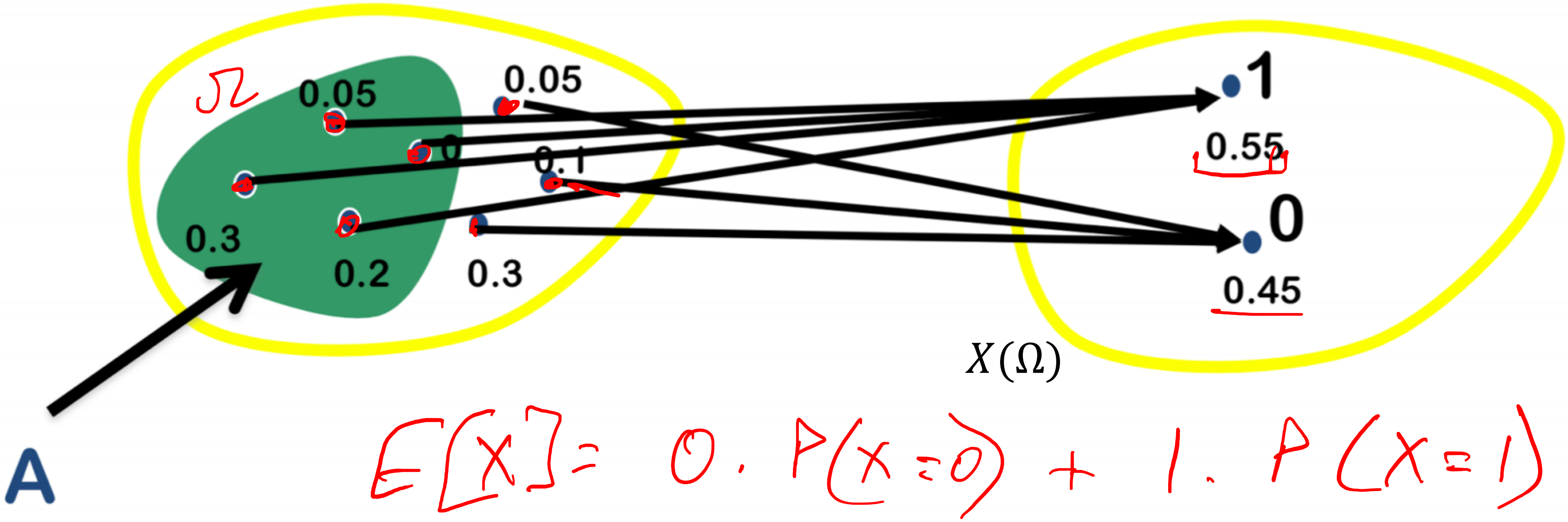
$$= np(p + (1-p))^{n-1} = \boldsymbol{np}$$

Indicator Random Variables

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

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

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



$E[X] = 0 \cdot P(X=0) + 1 \cdot P(X=1) = P(X=1)$

Repeated Coin Tosses (contd)

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

Repeated Coin Tosses (contd)

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

Let X be the total number of heads

Let us define X_i as follows:

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

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

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

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

Repeated Coin Tosses (contd)

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

Let X be the total number of heads

Let us define X_i as follows:

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

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

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

By Linearity of Expectation,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n] = \boldsymbol{np}$$

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

A	B	C	D	
D	C	B	A	← original
A	D	C	B	A, C
B	A	D	C	— X
C	B	A	D	← B, D

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: X_i as the ith person ending up in front of their own nametag.

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

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

LOE:

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

Conquer:

$$E[X_i] = \frac{1}{n-1}$$

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

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 $\mathbb{E}[X]$.**

<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?

Decompose: Let us define X as the number of pairs with the same birthday

Let us define X_k as follows:

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

LOE:

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

Conquer:

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

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

More Practice: Random Variables

More Practice

Suppose you flip a coin until you see a heads for the first time.

Let X be the number of trials (including the heads).

What is the pmf of X?

The cdf of X?

$\mathbb{E}[X]$?

More Practice

Suppose you flip a coin until you see a heads for the first time.

Let X be the number of trials (including the heads)

What is the pmf of X? $p_X(x) = \frac{1}{2^x}$ for $x \in \mathbb{Z}^+$, 0 otherwise

The cdf of X? $F_X(x) = 1 - \frac{1}{2^{\lfloor x \rfloor}}$ for $x \geq 0$, 0 for $x < 0$.

$\mathbb{E}[X]$? $\sum_{i=1}^{\infty} \frac{i}{2^i} = 2$

More Random Variable Practice

Roll a fair die n times. Let Z be the number of rolls that are $5s$ or $6s$.

What is the pmf?

Don’t try to write the CDF…it’s a mess…

Or try for a few minutes to realize it isn’t nice.

More Random Variable Practice

Roll a fair die n times. Let Z be the number of rolls that are $5s$ or $6s$.

What's the probability of getting exactly z 5's/6's?

We need to know which z of the n rolls are 5's/6's. And then multiply by the probability of getting exactly that outcome

$$p_Z(z) = \begin{cases} \binom{n}{z} \cdot \left(\frac{1}{3}\right)^z \left(\frac{2}{3}\right)^{n-z} & \text{if } z \in \mathbb{Z}, 0 \leq z \leq n \\ 0 & \text{otherwise} \end{cases}$$

More Independence

Independence of events

Recall the definition of independence of **events**:

Independence

Two events A, B are independent if

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

Independence for 3 or more events

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

Pairwise Independence

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

Mutual Independence

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

Pairwise Independence vs. Mutual Independence

Roll two fair dice (one red one blue) independently

$R =$"red die is 3"

$B =$"blue die is 5"

$S =$"sum is 7"

How should we describe these events?

Pairwise Independence

$\mathbb{P}(R \cap B)\ ?= \mathbb{P}(R)\mathbb{P}(B)$

$\frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6}$ Yes! (These are also independent by the problem statement)

$\mathbb{P}(R \cap S)\ ?= \mathbb{P}(R)\mathbb{P}(S)$

$\frac{1}{36}\ ?= \frac{1}{6} \cdot \frac{1}{6}$ Yes!

$\mathbb{P}(B \cap S)\ ?= \mathbb{P}(B)\mathbb{P}(S)$

$\frac{1}{36}\ ?= \frac{1}{6} \cdot \frac{1}{6}$ Yes!

R, B, S are pairwise independent

Since all three pairs are independent, we say the random variables are pairwise independent.

Mutual Independence

$\mathbb{P}(R \cap B \cap S) = 0$

if the red die is 3, and blue die is 5 then the sum is 8 (so it can't be 7)

$\mathbb{P}(R)\mathbb{P}(B)\mathbb{P}(S) = \left(\frac{1}{6}\right)^3 = \frac{1}{216} \neq 0$

R, B, S are not mutually independent.

Checking Mutual Independence

It's not enough to check just $\mathbb{P}(A \cap B \cap C)$ either.

Roll a fair 8-sided die.

Let A be $\{1,2,3,4\}$

B be $\{2,4,6,8\}$

C be $\{2,3,5,7\}$

$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{2\}) = \frac{1}{8}$

$\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

Checking Mutual Independence

It's not enough to check just $\mathbb{P}(A \cap B \cap C)$ either.

Roll a fair 8-sided die.

Let A be $\{1,2,3,4\}$

B be $\{2,4,6,8\}$

C be $\{2,3,5,7\}$

$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\{2\}) = \frac{1}{8}$

$\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

But A and B aren't independent (nor are B, C; though A and C are independent). Because there's a subset that's not independent, A, B, C are not mutually independent.

Checking Mutual Independence

To check mutual independence of events:

Check **every** subset.

To check pairwise independence of events:

Check **every** subset of size two.

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.

We say that the event “the sum is 7” is independent of “the red die is 5”

What about $S =$“the sum of two dice” and $R =$“the value of the red die”

Independence of Random Variables

The "for all values" is important.

We say that the event "the sum is 7" is independent of "the red die is 5"

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

NOT independent.

$\mathbb{P}(S = 2, R = 5) \neq \mathbb{P}(S = 2)\mathbb{P}(R = 5)$ (for example)

Independence of Random Variables

Flip a coin independently $2n$ times.

Let X be "the number of heads in the first n flips."

Let Y be "the number of heads in the last n flips."

X and Y are 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** $x_1, x_2, \ldots, x_n$

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

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

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