CSE 312

Foundations of Computing II

Lecture 8: More on random variables; expectation

Last Class:

- Random Variables
- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)

Today:

- Recap
- Expectation
- Linearity of Expectation
- Indicator Random Variables

Kandinsky

Review Random Variables

Definition. A **random variable (RV)** for a probability space (Ω, P) is a function $X: \Omega \to \mathbb{R}$.

The set of values that X can take on is its *range/support*: Ω_X

Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

$\mathbf{Pr}(\boldsymbol{\omega})$	$\boldsymbol{\omega}$	$\boldsymbol{X}(\boldsymbol{\omega})$
$1/6$	$1, 2, 3$	3
$1/6$	$1, 3, 2$	1
$1/6$	$2, 1, 3$	1
$1/6$	$2, 3, 1$	0
$1/6$	$3, 1, 2$	0
$1/6$	$3, 2, 1$	1

Review Random Variables

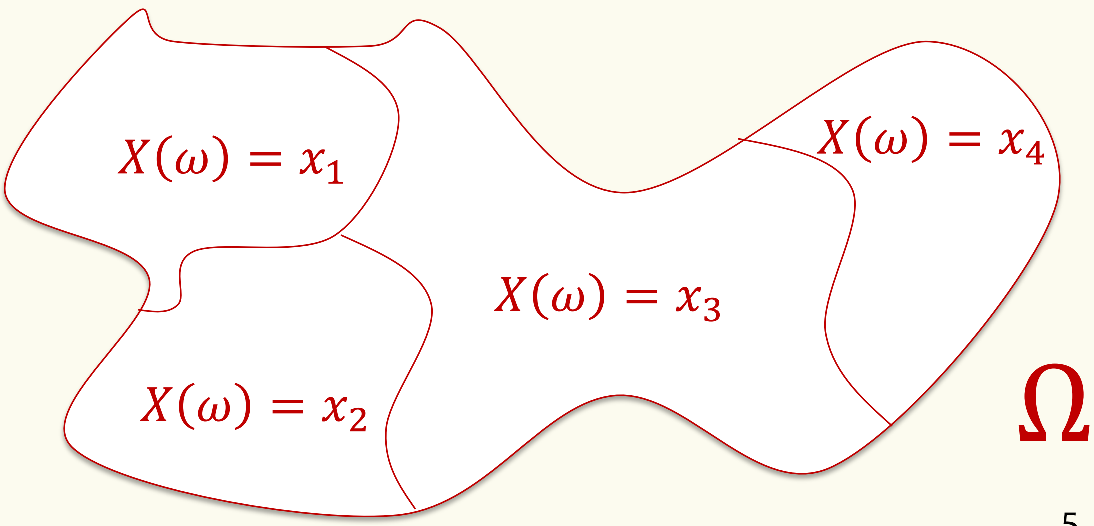
Definition. A **random variable (RV)** for a probability space (Ω, P) is a function $X: \Omega \to \mathbb{R}$.

The set of values that X can take on is its *range/support*: $\boldsymbol{X(\Omega)}$

$\{X = x_i\} = \{\omega \in \Omega \mid X(\omega) = x_i\}$

Random variables **partition** the sample space.

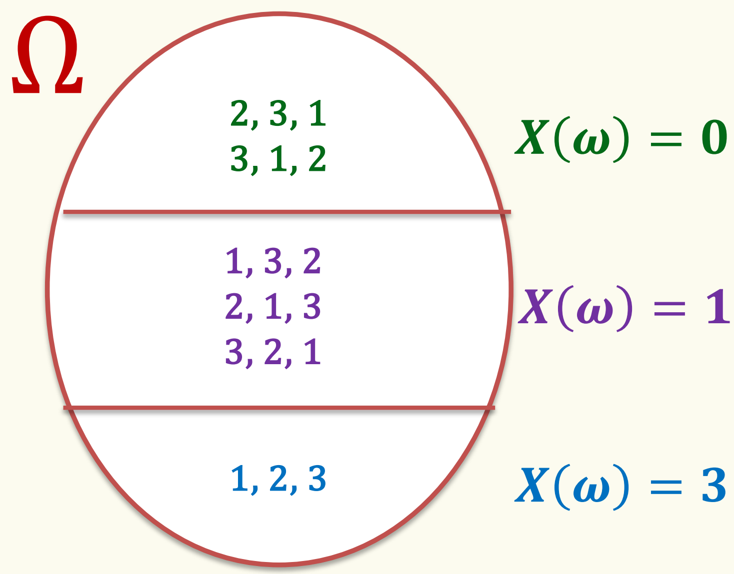
$\Sigma_{x \in X(\Omega)} P(X = x) = 1$



Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

$\Pr(\omega)$	ω	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1



Review PMF and CDF

Definitions:

For a RV $X: \Omega \to \mathbb{R}$, the probability mass function (pmf) of X specifies, for any real number x, the probability that $X = x$

$$p_X(x) = P(X = x) = P(\{\omega \in \Omega \mid X(\omega) = x\})$$

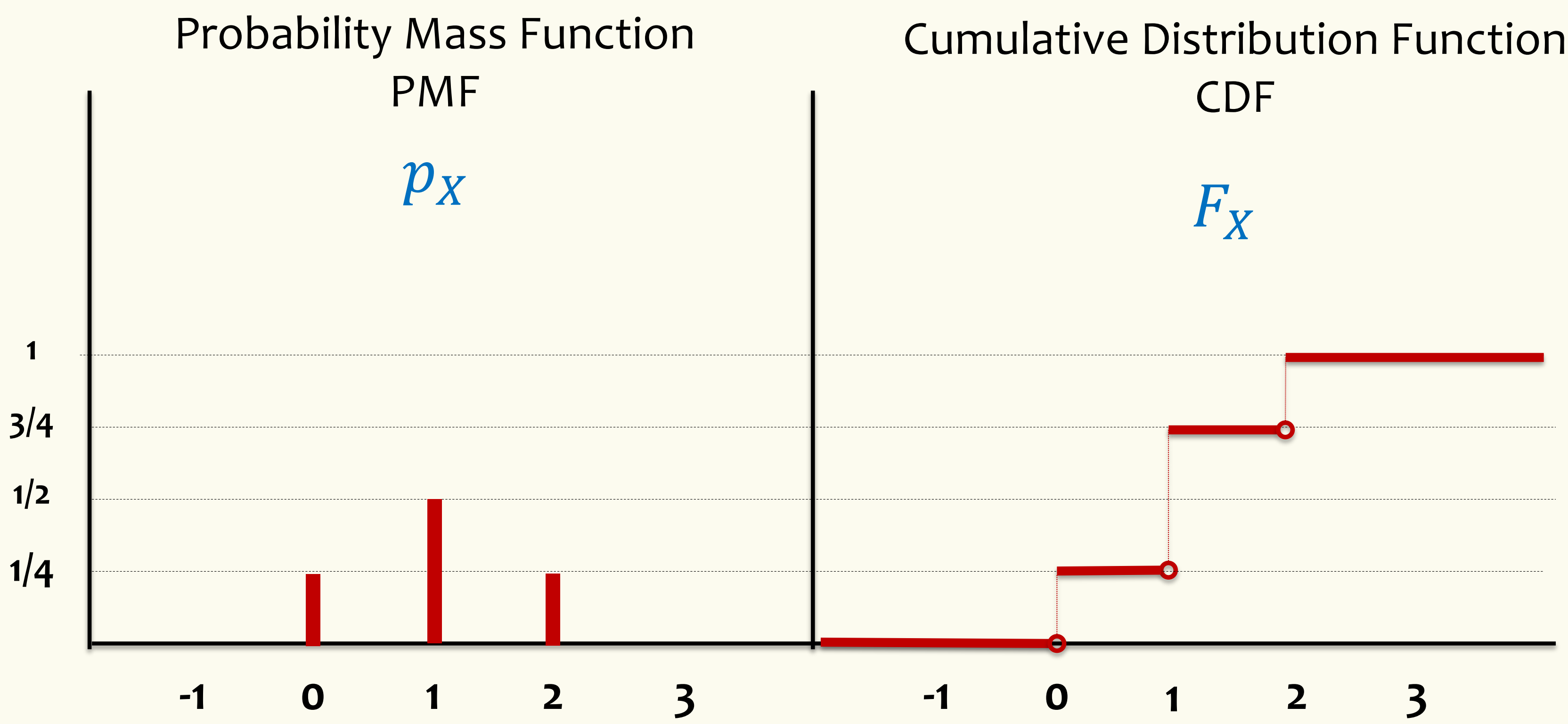
$$\sum_{x \in \Omega_X} p_X(x) = 1$$

For a RV $X: \Omega \to \mathbb{R}$, the cumulative distribution function (cdf) of X specifies, for any real number x, the probability that $X \leq x$

$$F_X(x) = P(X \leq x)$$

Example – Two fair independent coin flips

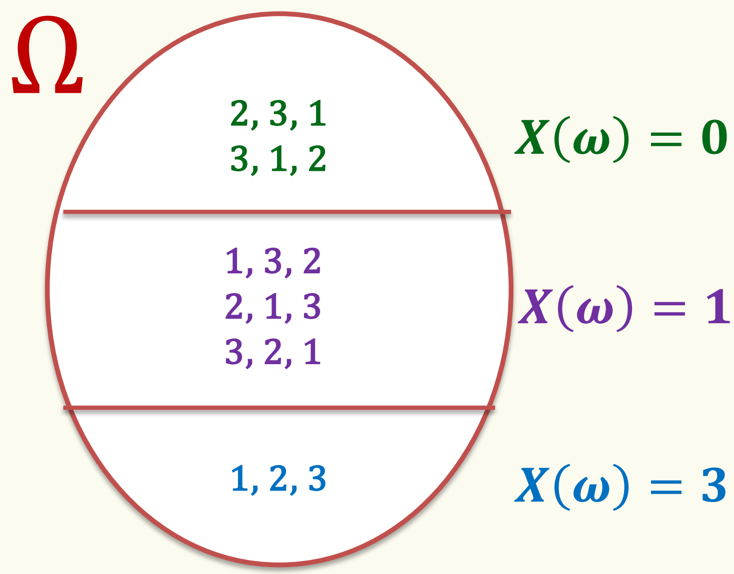
X = number of heads



Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

$\Pr(\omega)$	ω	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1



Example: Returning Homeworks

$\Pr(\omega)$	ω	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

- Class with 3 students, randomly hand back homeworks.

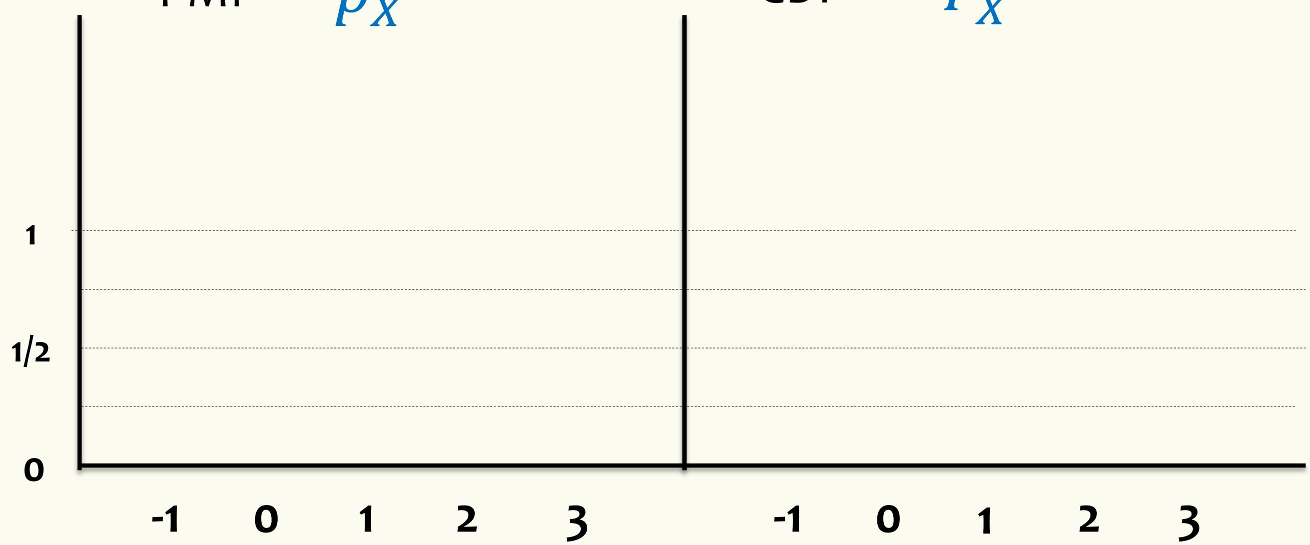
All permutations equally likely.

- **Let X be the number of students who get their own HW**

Probability Mass Fn
PMF p_X

Cumulative Distribution Fn
CDF F_X

$$\mathbf{p_X(0) = P(X = 0) = 1/3}$$
$$\mathbf{p_X(1) = P(X = 1) = 1/2}$$
$$\mathbf{p_X(2) = P(X = 3) = 1/6}$$



Example – Number of Heads

We flip n coins, independently, each heads with probability p

$\Omega = \{\mathrm{HH} \cdots \mathrm{HH}, \mathrm{HH} \cdots \mathrm{HT}, \mathrm{HH} \cdots \mathrm{TH}, \dots, \mathrm{TT} \cdots \mathrm{TT}\}$

$X =$ # of heads

$$p_X(k) = P(X = k) =$$

Example – Number of Heads

We flip n coins, independently, each heads with probability p

$\Omega = \{\mathrm{HH} \cdots \mathrm{HH}, \mathrm{HH} \cdots \mathrm{HT}, \mathrm{HH} \cdots \mathrm{TH}, \dots, \mathrm{TT} \cdots \mathrm{TT}\}$

$X =$ # of heads

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

# of sequences with k heads

Prob of sequence w/ k heads

Agenda

- Random Variables
- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)
- Expectation

Expectation (Idea)

Example. Toss a coin 20 times independently with probability ¼ of coming up heads on each toss.

$X =$ number of heads

How many heads do you ***expect*** to see?

What if you toss it independently *n* times and it comes up heads with probability *p* each time?

Review Expected Value of a Random Variable

Definition. Given a discrete RV $X: \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of X is

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

or equivalently

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

Intuition: "Weighted average" of the possible outcomes (weighted by probability)

Expectation

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

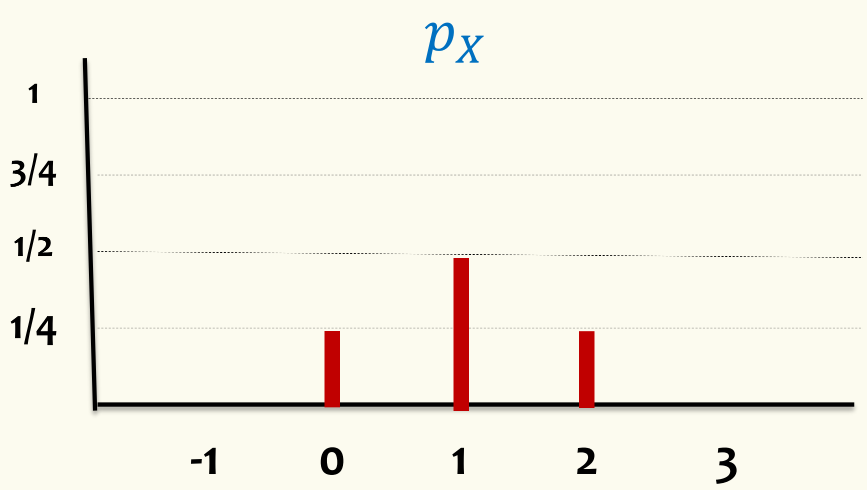
$$\mathbb{E}[X] = \sum_{x \in \Omega_X} x \cdot P(X = x)$$

Example. Two fair coin flips

$\Omega = \{\mathrm{TT}, \mathrm{HT}, \mathrm{TH}, \mathrm{HH}\}$

X = number of heads

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



$$\mathbb{E}[X] = X(TT)\,P(TT) + X(HT)\,P(HT) + X(TH)\,P(TH) + X(HH)\,P(HH)$$

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

$$\mathbb{E}[X] = 0 \cdot p_X(0) + 1 \cdot p_X(1) + 2 \cdot p_X(2)$$

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

Example: Returning Homeworks

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

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

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW
- What is $\mathbb{E}[X]$?

$\Pr(\omega)$	ω	$X(\omega)$
$1/6$	$2, 3, 1$	0
$1/6$	$3, 1, 2$	0
$1/6$	$1, 3, 2$	1
$1/6$	$3, 2, 1$	1
$1/6$	$2, 1, 3$	1
$1/6$	$1, 2, 3$	3

Example: Returning Homeworks

- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

$\Pr(\omega)$	ω	$X(\omega)$
$1/6$	$1, 2, 3$	3
$1/6$	$1, 3, 2$	1
$1/6$	$2, 1, 3$	1
$1/6$	$2, 3, 1$	0
$1/6$	$3, 1, 2$	0
$1/6$	$3, 2, 1$	1

$$\mathbb{E}[X] = 3 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 0 \cdot \frac{1}{6} + 1 \cdot \frac{1}{6}$$

$$= 6 \cdot \frac{1}{6} = 1$$

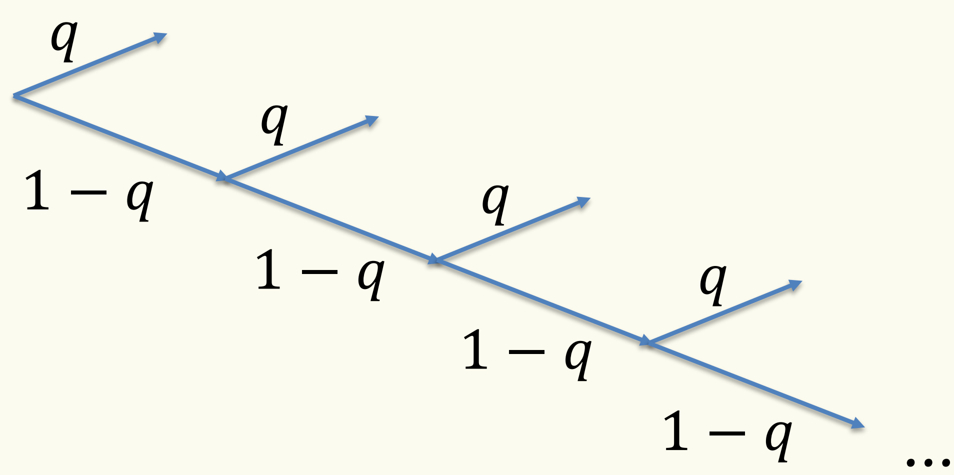
Example – Flipping a biased coin until you see heads

- Biased coin, each flip indep:

  $P(H) = q > 0$

  $P(T) = 1 - q$

- Z = # of coin flips until first head



$P(Z = 0) =$

$P(Z = 1) =$

$P(Z = 2) =$

$P(Z = i) =$

$\mathbb{E}[Z] =$

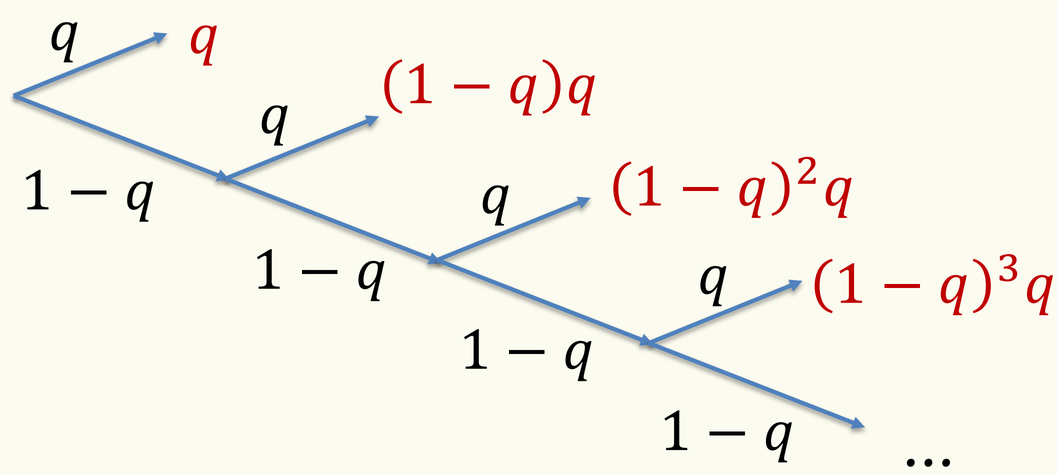
Example – Flipping a biased coin until you see heads

- Biased coin, each flip indep:

  $P(H) = q > 0$

  $P(T) = 1 - q$

- Z = # of coin flips until first head



$$P(Z = i) = q\,(1-q)^{i-1}$$

$$\mathbb{E}[Z] = \sum_{i=1}^{\infty} i \cdot P(Z = i) = \sum_{i=1}^{\infty} i \cdot q(1-q)^{i-1}$$

Converges, so $\mathbb{E}[Z]$ is finite

Can calculate this directly but…

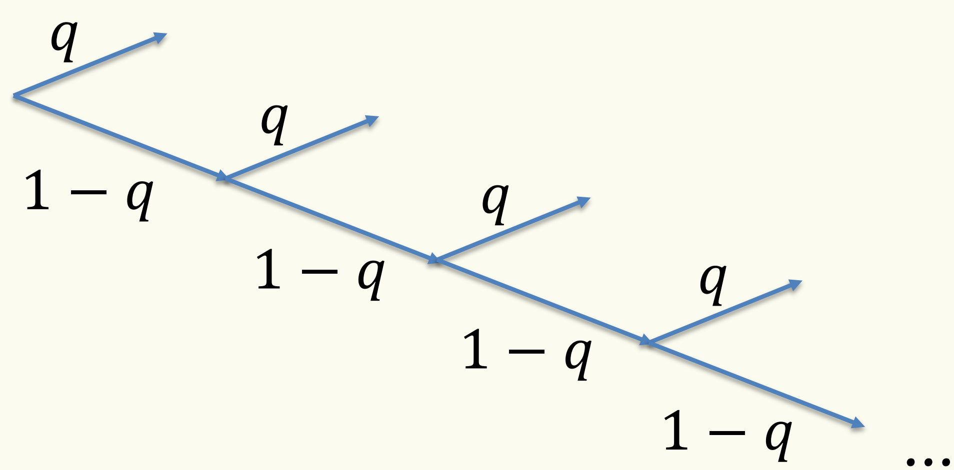
Example – Flipping a biased coin until you see heads

- Biased coin, each flip indep:

  $P(H) = q > 0$

  $P(T) = 1 - q$

- Z = # of coin flips until first head



Another view: If you get heads first try you get $Z = 1$;
If you get tails you have used one try and have the same experiment left

$\mathbb{E}[Z] =$

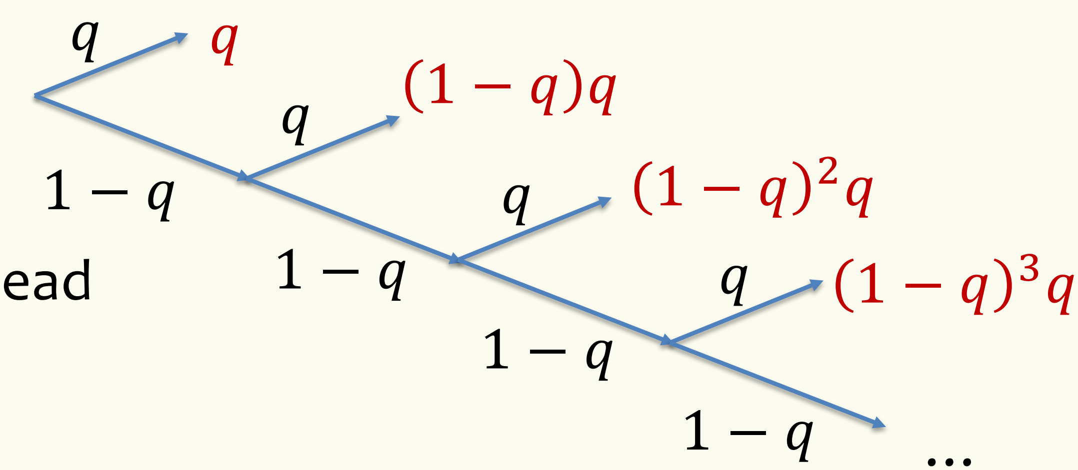
Example – Flipping a biased coin until you see heads

- Biased coin:

  $P(H) = q > 0$

  $P(T) = 1 - q$

- Z = # of coin flips until first head



Another view: If you get heads first try you get $Z = 1$;
If you get tails you have used one try and have the same experiment left

$$\mathbb{E}[Z] = q \cdot 1 + (1-q)(1 + \mathbb{E}(Z))$$

Solving gives $q \cdot \mathbb{E}[Z] = q + (1-q) = 1$ Implies $\mathbb{E}[Z] = 1/q$

Example – Coin Tosses

We flip n coins, each toss independent, probability p of coming up heads.

Z is the number of heads, what is $\mathbb{E}(Z)$?

Example – Coin Tosses

We flip n coins, each toss independent; heads with probability p,

Z is the number of heads, what is $\mathbb{E}[Z]$?

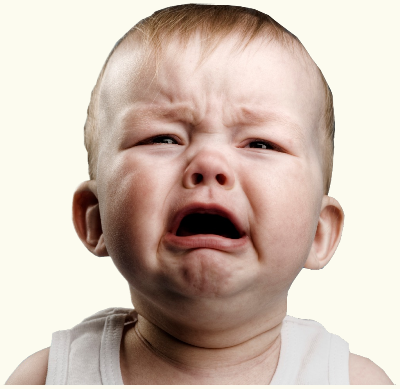
$$\mathbb{E}[Z] = \sum_{k=0}^{n} k \cdot P(Z = k) \quad = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=0}^{n} k \cdot \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \quad = \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!\,(n-1-k)!} p^k (1-p)^{(n-1)-k}$$

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



Can we solve it more elegantly, please?

Agenda

- Random Variables
- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)
- Expectation
- Linearity of Expectation

Linearity of Expectation

Theorem. For any two random variables X and Y
(no conditions whatsoever on the random variables)

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

Or, more generally: For any random variables $X_1, \ldots, X_n$,

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

Because: $\mathbb{E}[X_1 + \cdots + X_n] = \mathbb{E}[(X_1 + \cdots + X_{n-1}) + X_n]$

$$= \mathbb{E}[X_1 + \cdots + X_{n-1}] + \mathbb{E}[X_n] = \cdots$$

Linearity of Expectation – Proof

Theorem. For any two random variables X and Y

(X, Y do not need to be independent)

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

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

$$= \sum_{\omega} P(\omega)X(\omega) + \sum_{\omega} P(\omega)Y(\omega)$$

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

Using LOE to compute complicated expectations

Often boils down to the following three steps:

- Decompose: Finding the right way to decompose the random variable into sum of simple random variables

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

- LOE: Apply linearity of expectation.

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

- Conquer: Compute the expectation of each X_i

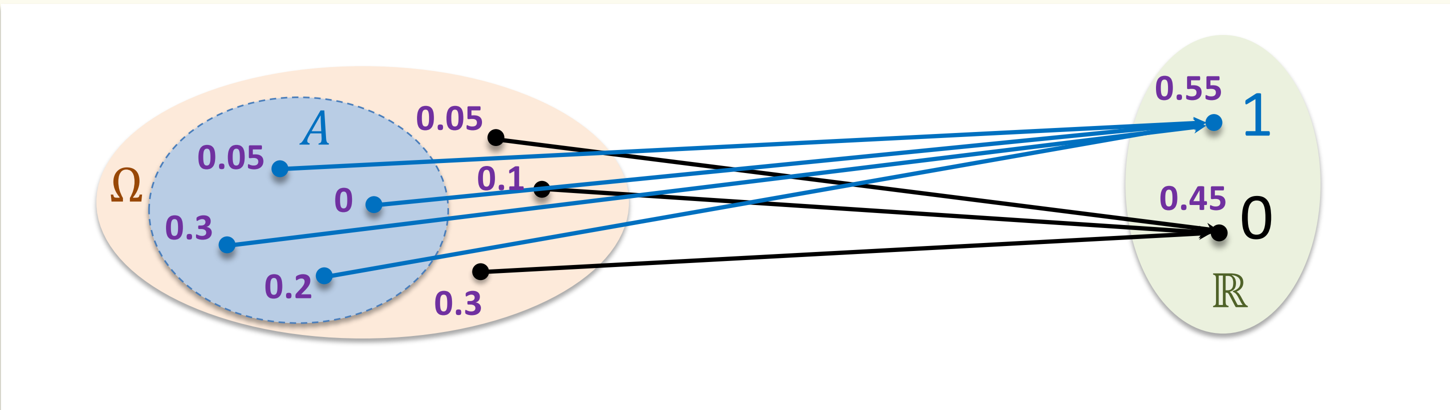
Often, X_i are **indicator** (0/1) random variables.

Indicator random variables – 0/1 valued

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

$$X_A = \begin{cases} 1 & \text{if event } A \text{ occurs} \\ 0 & \text{if event } A \text{ does not occur} \end{cases}$$

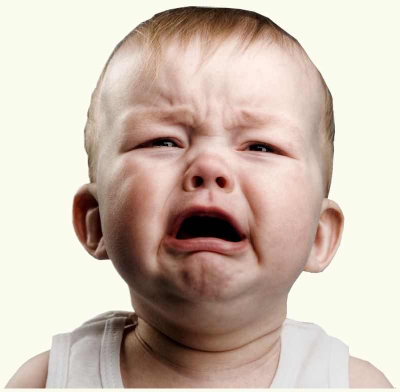
$$P(X_A = 1) = P(A)$$
$$P(X_A = 0) = 1 - P(A)$$



Example – Coin Tosses – The brute force method

We flip n coins, each one heads with probability p,

Z is the number of heads, what is $\mathbb{E}[Z]$?



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

$$= \sum_{k=0}^{n} k \cdot \frac{n!}{k!\,(n-k)!} p^k (1-p)^{n-k} \quad = \sum_{k=1}^{n} \frac{n!}{(k-1)!\,(n-k)!} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!\,(n-k)!} p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!\,(n-1-k)!} p^k (1-p)^{(n-1)-k}$$

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

Can we solve it more elegantly, please?

Computing complicated expectations

Often boils down to the following three steps:

- Decompose: Finding the right way to decompose the random variable into sum of simple random variables

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

- LOE: Apply linearity of expectation.

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

- Conquer: Compute the expectation of each X_i

Often, X_i are indicator (0/1) random variables.

Example – Coin Tosses

We flip n coins, each toss independent, comes up heads with probability p

Z is the number of heads, what is $\mathbb{E}[Z]$?

$$X_i = \begin{cases} 1, & i^{\text{th}} \text{ coin flip is heads} \\ 0, & i^{\text{th}} \text{ coin flip is tails.} \end{cases}$$

Fact. $Z = X_1 + \cdots + X_n$

Outcome	X_1	X_2	X_3	Z
TTT	0	0	0	0
TTH	0	0	1	1
THT	0	1	0	1
THH	0	1	1	2
HTT	1	0	0	1
HTH	**1**	**0**	**1**	**2**
HHT	1	1	0	2
HHH	1	1	1	3

Example – Coin Tosses

We flip n coins, each toss independent, comes up heads with probability p

Z is the number of heads, what is $\mathbb{E}[Z]$?

- $X_i = \begin{cases} 1, & i^{\text{th}} \text{ coin flip is heads} \\ 0, & i^{\text{th}} \text{ coin flip is tails.} \end{cases}$

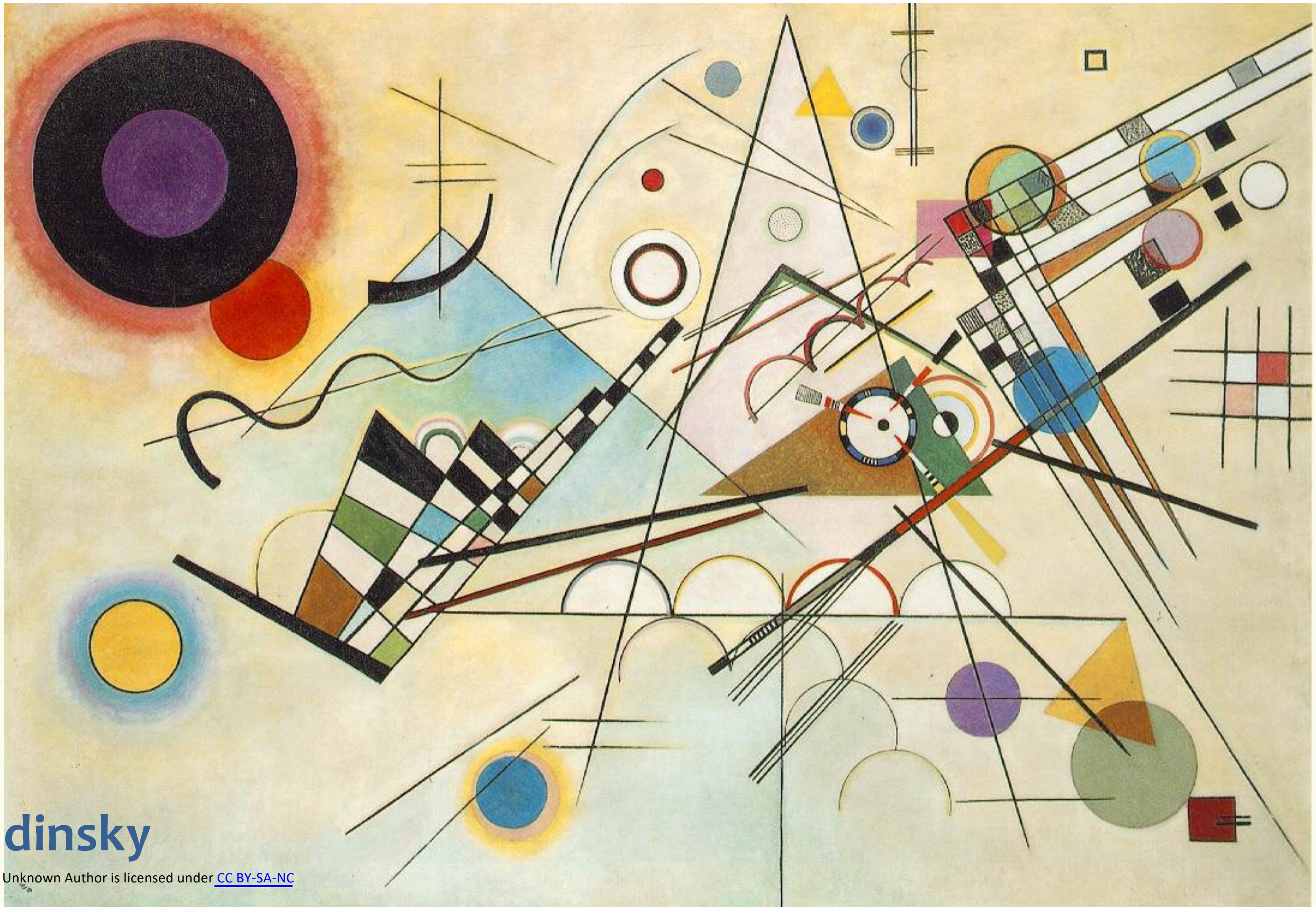
Fact. $Z = X_1 + \cdots + X_n$

Linearity of Expectation:

$$\mathbb{E}[Z] = \mathbb{E}[X_1 + \cdots + X_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] \quad = n \cdot p$$

$P(X_i = 1) = p$
$P(X_i = 0) = 1 - p$

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



Kandinsky

Example: Returning Homeworks

- Class with n students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

What is $\mathbb{E}[X]$? Use linearity of expectation!

$\Pr(\omega)$	ω	$X(\omega)$
1/6	1, 2, 3	3
1/6	1, 3, 2	1
1/6	2, 1, 3	1
1/6	2, 3, 1	0
1/6	3, 1, 2	0
1/6	3, 2, 1	1

Decompose: Find the right way to decompose the random variable into sum of simple random variables

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

LOE: Apply linearity of expectation.

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

Conquer: Compute the expectation of each X_i and sum!

Example: Returning Homeworks

- Class with n students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

What is $\mathbb{E}[X]$? Use linearity of expectation!

$\Pr(\omega)$	ω	$X(\omega)$
$1/6$	$1, 2, 3$	3
$1/6$	$1, 3, 2$	1
$1/6$	$2, 1, 3$	1
$1/6$	$2, 3, 1$	0
$1/6$	$3, 1, 2$	0
$1/6$	$3, 2, 1$	1

Decompose:

LOE:

Conquer:

Example: Returning Homeworks

- Class with n students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

What is $\mathbb{E}[X]$? Use linearity of expectation!

$\Pr(\omega)$	ω	$X(\omega)$
$1/6$	$1, 2, 3$	3
$1/6$	$1, 3, 2$	1
$1/6$	$2, 1, 3$	1
$1/6$	$2, 3, 1$	0
$1/6$	$3, 1, 2$	0
$1/6$	$3, 2, 1$	1

Decompose: What is X_i?

$X_i = 1$ iff i^{th} student gets own HW back

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

Conquer: $\mathbb{E}[X_i] = \frac{1}{n}$

Therefore, $\mathbb{E}[X] = n \cdot \frac{1}{n} = 1$

Pairs with the same birthday

- In a class of m students, on average how many pairs of people have the same birthday (assuming 365 equally likely birthdays)?

Pairs with the same birthday

- In a class of m students, on average how many pairs of people have the same birthday (assuming 365 equally likely birthdays)?

Decompose: Indicator events involve **pairs** of students (i, j) for $i \neq j$
$X_{ij} = 1$ iff students i and j have the same birthday

LOE: $\binom{m}{2}$ indicator variables X_{ij}

Conquer: $\mathbb{E}[X_{ij}] = \frac{1}{365}$ so total expectation is $\frac{\binom{m}{2}}{365} = \frac{m(m-1)}{730}$ pairs

Linearity of Expectation – Even stronger

Theorem. For any random variables $X_1, \dots, X_n$, and real numbers $a_1, \dots, a_n \in \mathbb{R}$,

$$\mathbb{E}[a_1X_1 + \cdots + a_nX_n] = a_1\mathbb{E}[X_1] + \cdots + a_n\mathbb{E}[X_n].$$

Linearity of Expectation – Even stronger

Theorem. For any random variables $X_1, \dots, X_n$, and real numbers $a_1, \dots, a_n \in \mathbb{R}$,

$$\mathbb{E}[a_1 X_1 + \cdots + a_n X_n] = a_1 \mathbb{E}[X_1] + \cdots + a_n \mathbb{E}[X_n].$$

Very important: In general, we do <u>not</u> have $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Linearity is special!

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

E.g., $X = \begin{cases} +1 \text{ with prob } 1/2 \\ -1 \text{ with prob } 1/2 \end{cases}$

Then: $\mathbb{E}[X^2] \neq \mathbb{E}[X]^2$

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

Expected Value of $g(X)$

Definition. Given a discrete RV $X : \Omega \to \mathbb{R}$, the **expectation** or **expected value** or **mean** of $g(X)$ is

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

or equivalently

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

Example: from concept check

- Toss a die; each side equally likely. X is the number showing
- $Y = X \bmod 4$
- What is $\mathbb{E}[X], \mathbb{E}[Y]$?

$\Pr(\omega)$	ω	X
$1/6$	1	1
$1/6$	2	2
$1/6$	3	3
$1/6$	4	4
$1/6$	5	5
$1/6$	6	6

Kandinsky