CSE 312

Foundations of Computing II

Lecture 8: Linearity of Expectation

Review Random Variables

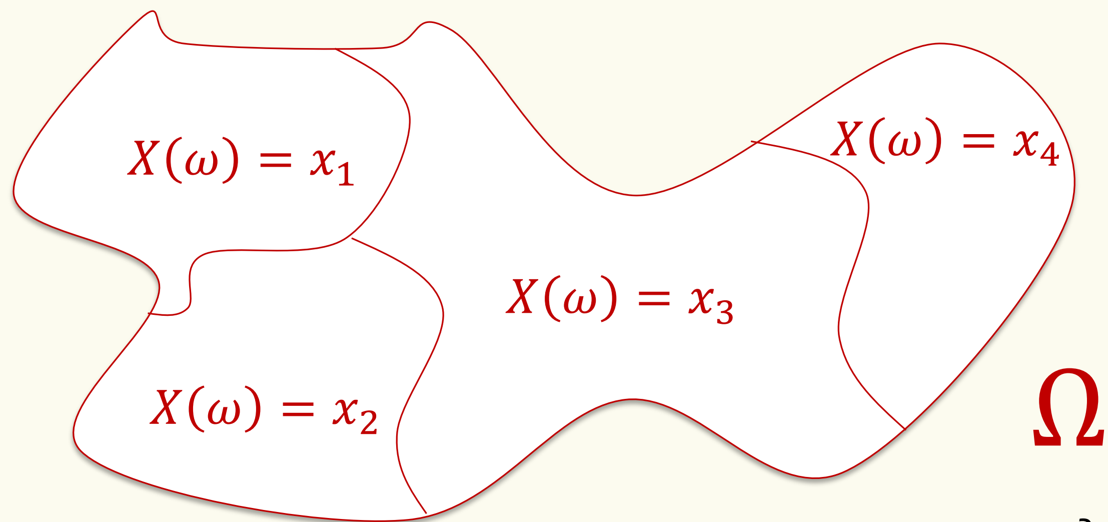
Definition. A **random variable (RV)** for a probability space (Ω, P) is a function $X: \Omega \to \mathbb{R}$.

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

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

Random variables **partition** the sample space.

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



Agenda

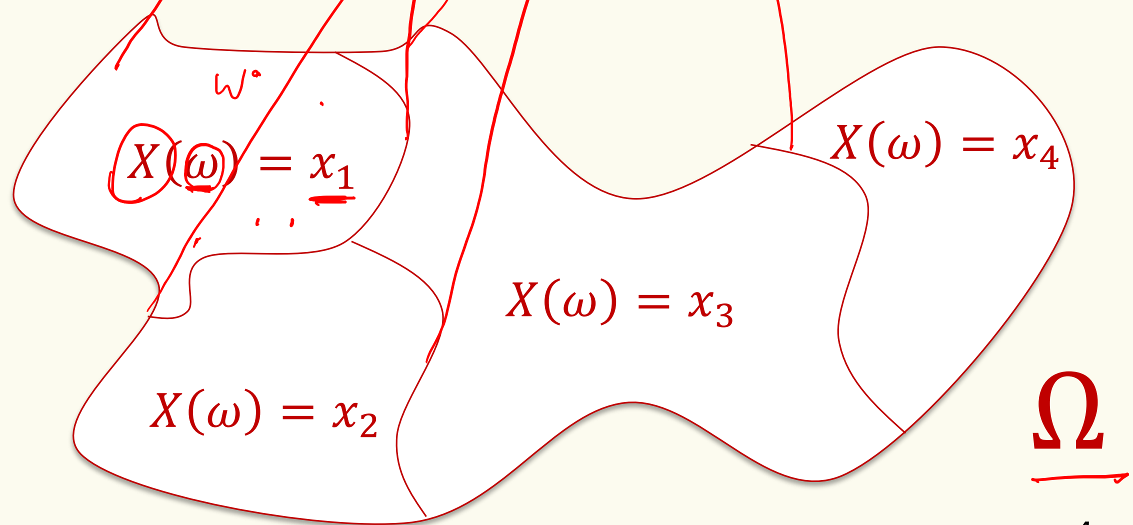
- Random Variables
- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)
- Expectation
- Properties of Expectation

Probability Mass Function (PMF)

Definition. For a RV $X: \Omega \to \mathbb{R}$, the function $p_X: X(\Omega) \to \mathbb{R}$ defined by $p_X(x) = P(X = x)$ is called the **probability mass function (PMF)** of X

Random variables **partition** the sample space.

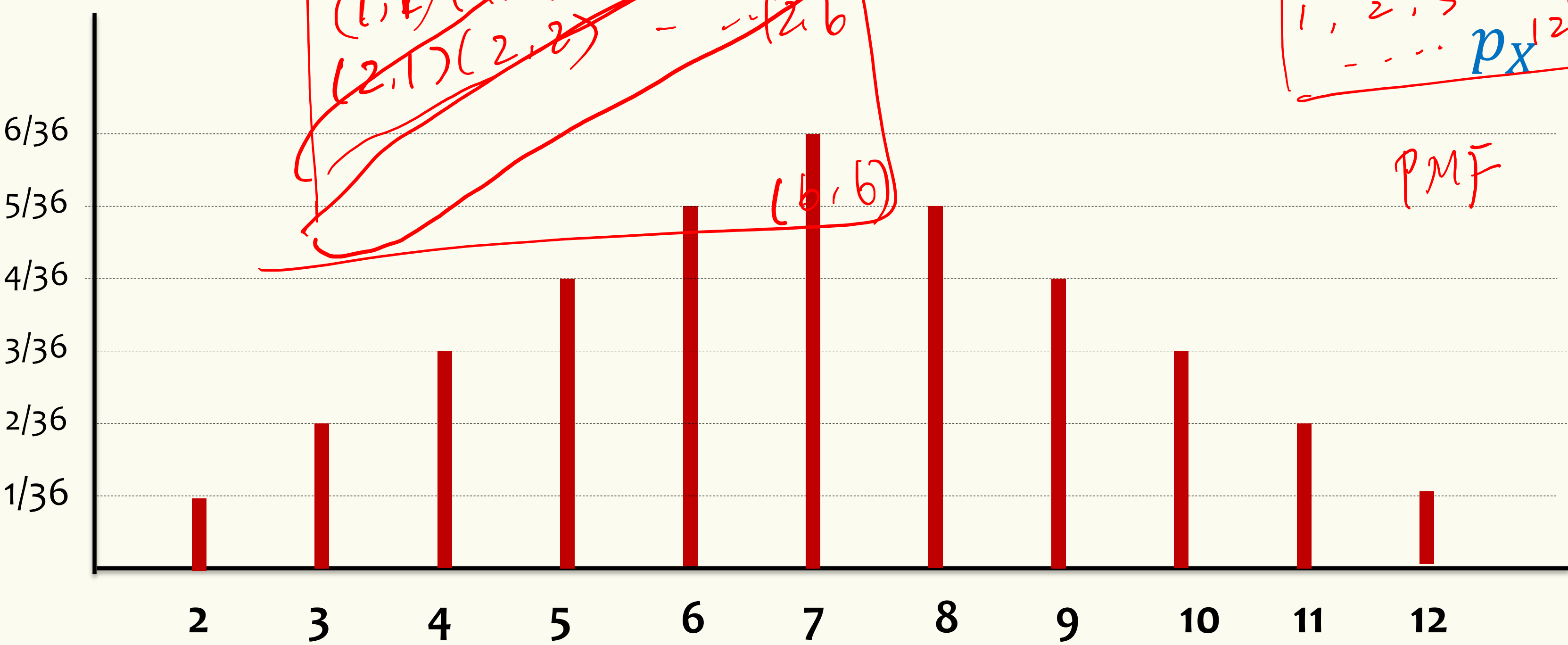
$$\left(\sum_{x \in \Omega_X} \underbrace{P(X = x)}_{p_X(x)}\right) = 1$$



Example – Two Fair Dice

X = sum of two dice throws

p_X



Example – Number of Heads

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

$\Omega = \{\text{HH} \cdots \text{HH}, \text{HH} \cdots \text{HT}, \text{HH} \cdots \text{TH}, \ldots, \text{TT} \cdots \text{TT}\}$

$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

Events concerning RVs

We already defined $P(X = x) = P(\{X = x\})$ where

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

Sometimes we want to understand other events involving RV X

– e.g. $\{X \leq x\} = \{\omega \in \Omega \mid X(\omega) \leq x\}$ which makes sense for any $x \in \mathbb{R}$

Cumulative Distribution Function (CDF)

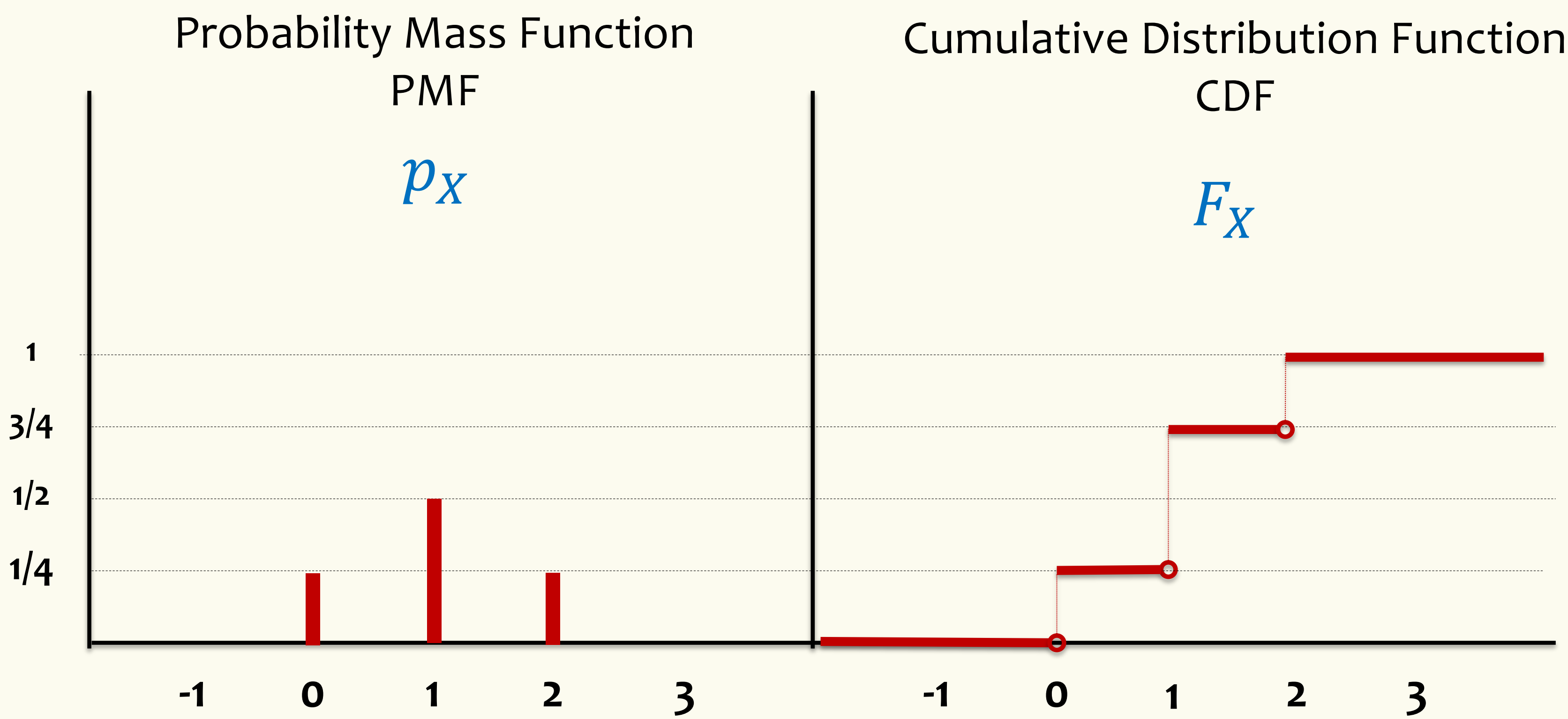
Definition. For a RV $X: \Omega \to \mathbb{R}$, the cumulative distribution function of X is the function $F_X: \mathbb{R} \to [0,1]$ that specifies for any real number x, the probability that $X \le x$.

$p_X(x) = P(X = x)$

That is, F_X is defined by $F_X(x) = P(X \le x)$

Example – Two fair coin flips

X = number of heads



Agenda

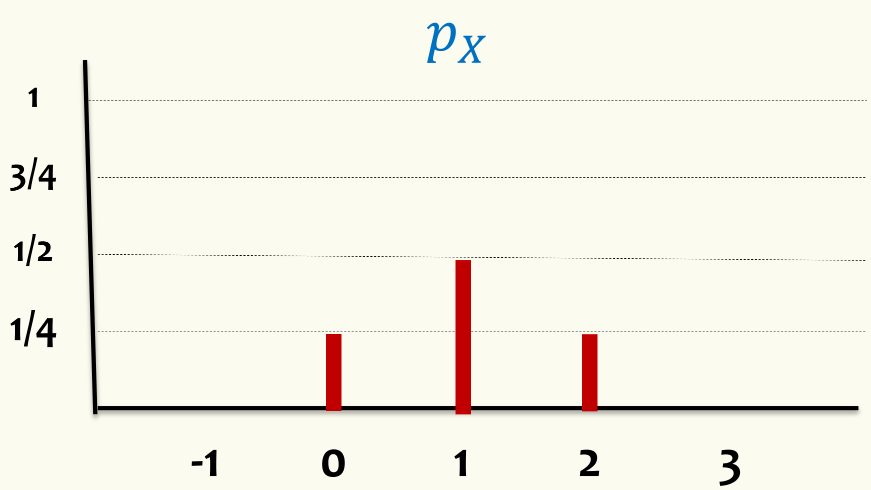
- Random Variables
- Probability Mass Function (PMF)
- Cumulative Distribution Function (CDF)
- Expectation
- Properties of Expectation

Expectation (Idea)

Example. Two fair coin flips

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

$X =$ number of heads



- What is the number of heads do we ***expect*** to see in two fair coin flips?

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

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

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

Expected Value

Definition. The expected value of a (discrete) RV X is

$$\mathbb{E}[X] = \sum_x x \cdot p_X(x) = \sum_x x \cdot P(X = x)$$

Example. Value X of rolling one fair die

$$p_X(1) = p_X(2) = \cdots = p_X(6) = \frac{1}{6}$$

$$\mathbb{E}[X] = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3.5$$

For the equally-likely outcomes case, this is just the average of the possible outcomes!

Expectation

Example. Two fair coin flips

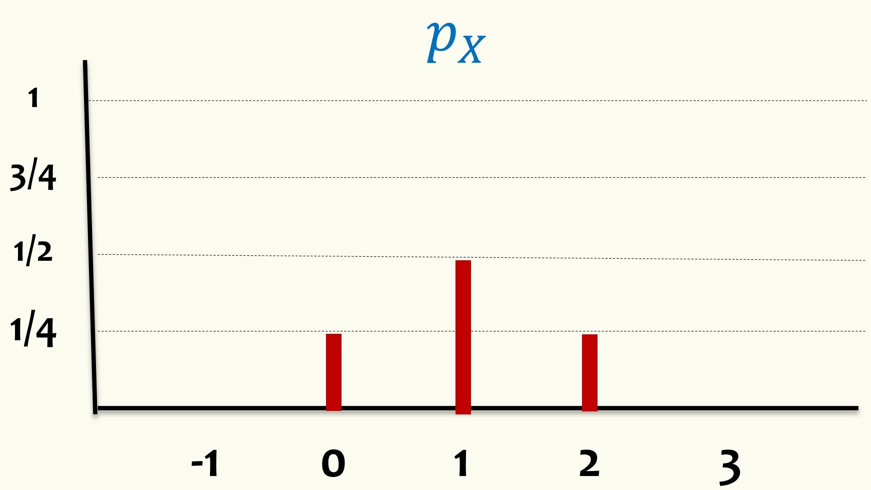
$\Omega = \{TT, HT, TH, HH\}$

X = number of heads

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

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

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



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

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

0

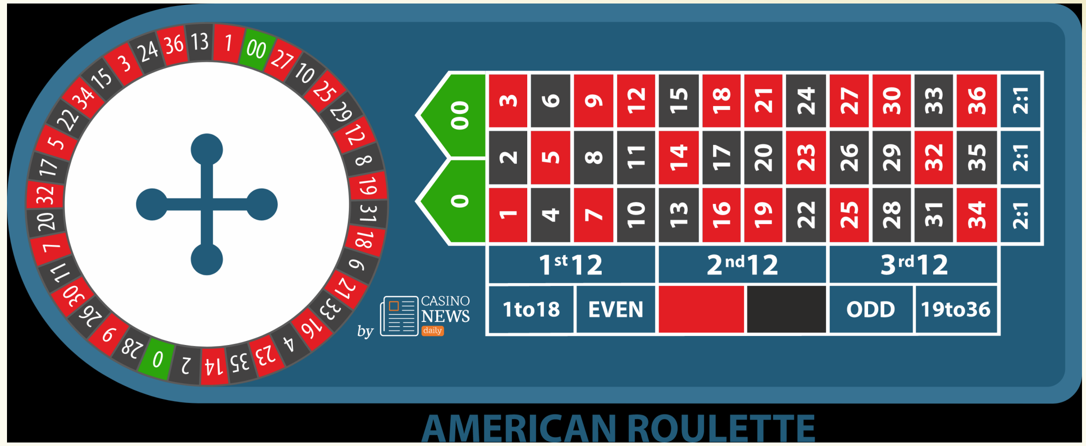
Roulette (USA)

Ω:

Numbers 1-36

- 18 Red
- 18 Black

Green 0 and 00



Note 0 and 00 are not EVEN

RVs for gains from some bets:

RV RED: If Red number turns up $+1$, if Black number, 0, or 00 turns up -1

$$\mathbb{E}[\text{RED}] = (+1) \cdot \frac{18}{38} + (-1) \cdot \frac{20}{38} = -\frac{2}{38} \approx -5.26\%$$

RV 1st12: If number 1-12 turns up $+2$, if number 13-36, 0, or 00 turns up -1

$$\mathbb{E}[1^{st}12] = (+2) \cdot \frac{12}{38} + (-1) \cdot \frac{26}{38} = -\frac{2}{38} \approx -5.26\%$$

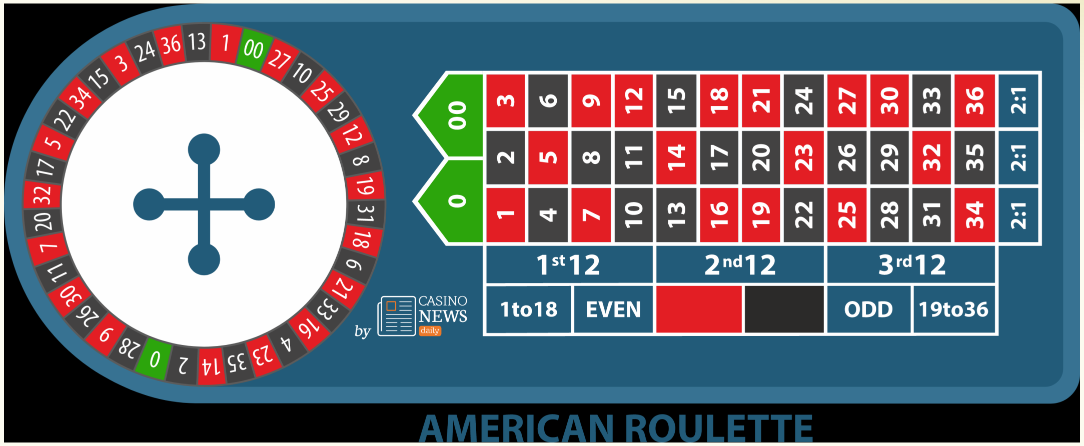
Roulette (USA)

Ω:

Numbers 1-36

- 18 Red
- 18 Black

Green 0 and 00



Note 0 and 00 are not EVEN

An even worse bet:

RV BASKET: If 0, 00, 1, 2, or 3 turns up +6 otherwise -1

$$\mathbb{E}[\text{BASKET}] = (+6)\cdot\frac{5}{38} + (-1)\cdot\frac{33}{38} = -\frac{3}{38} \approx -7.89\%$$

Frequenst Interpretation

"If X is how much you win playing the game in one round. How much would you expect to win, <u>on average</u>, per game, when repeatedly playing?"

Answer: $\mathbb{E}[X]$

$x_1 \leftarrow X \quad x_2 \leftarrow X \quad x_3 \leftarrow X \quad \ldots$

$$\mathbb{E}[X] = \lim_{t \to \infty} \left(\sum_{i=1}^{t} x_i \right) / t$$

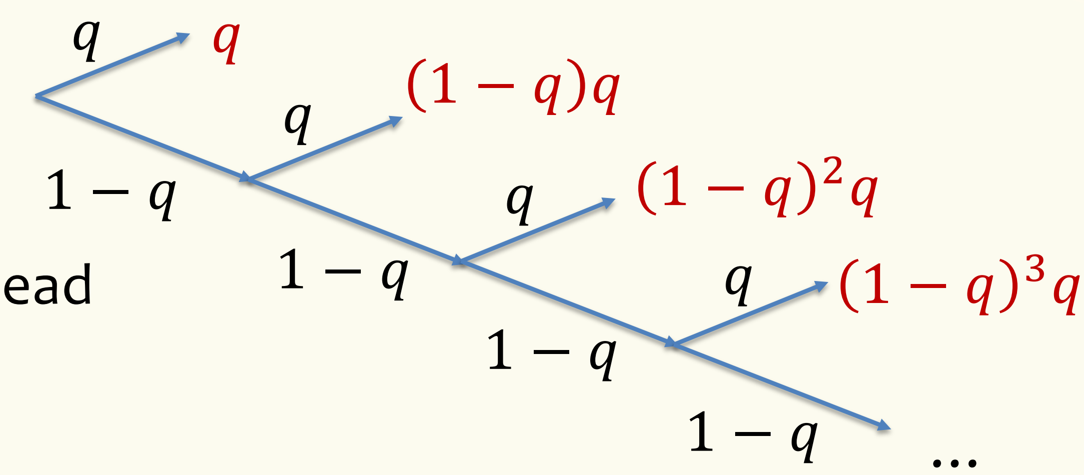
Example – Flipping a biased coin until you see heads

- Biased coin:

  $P(H) = q > 0$

  $P(T) = 1 - q$

- Z = # of coin flips until first head



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

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

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

Can calculate this directly but…

Agenda

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

Linearity of Expectation

Theorem. For any two random variables X and Y

(X, Y do not need to be independent)

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

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

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

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)

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

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

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

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

Example – Coin Tosses

We flip n coins, each one heads with probability p

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

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 $\mathbb{E}[Z]$?

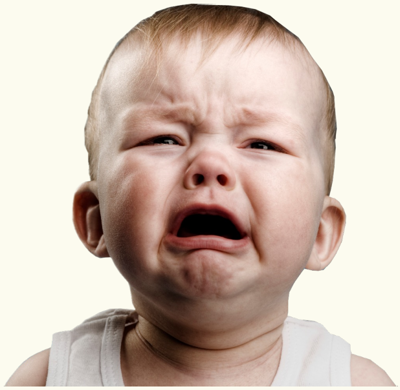
$$\mathbb{E}[Z] = \sum_{k=0}^{n} k \cdot P(Z = k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

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

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

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

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



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

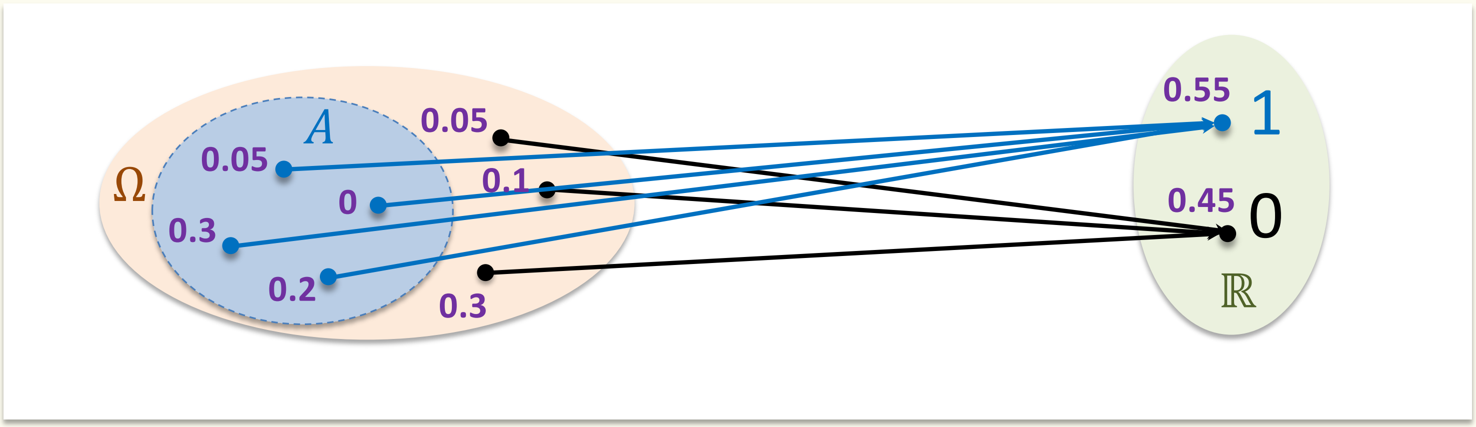
Often, X_i are indicator (0/1) random variables.

Indicator random variables

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

We flip n coins, each one heads with probability p

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

T H H ... H

0 1 1 ... 1

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

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

Linearity of Expectation:

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

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

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

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: What is $\mathbb{E}[X_i]$? A. $\frac{1}{n}$ B. $\frac{1}{n-1}$ C. $\frac{1}{2}$

Poll: **pollev.com/rachel312**

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 $a_1, \dots, a_n \in \mathbb{R}$,

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

Very important: In general, we do not 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

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