CSE 312

Foundations of Computing II

Lecture 8: More on random variables; expectation

Anonymous Questions: www.slido.com/1891306

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*: $X(\Omega)$ or Ω_X

Example: Returning Homeworks

s1 s2 s3

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

s1 s2 s3

$\Pr(\omega)$	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

$\Omega_X = \{0, 1, 3\}$

Example: Returning Homeworks

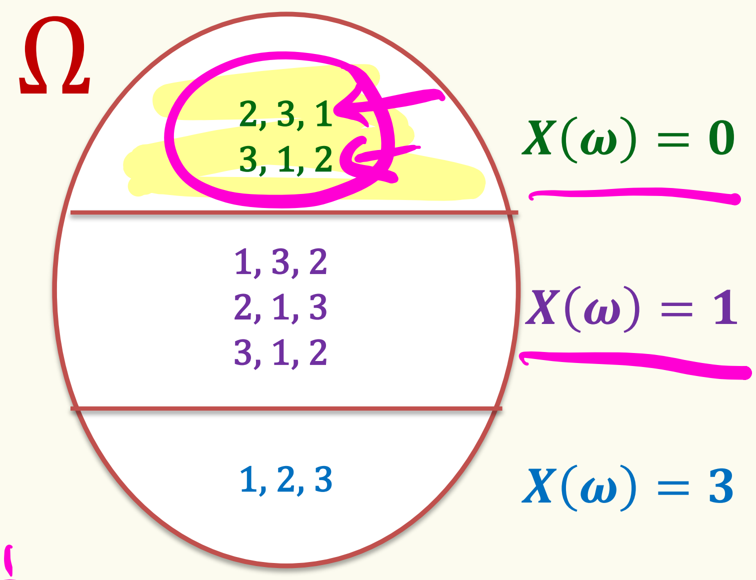
- Class with 3 students, randomly hand back homeworks. All permutations equally likely.
- Let X be the number of students who get their own HW

$\Pr(\omega)$	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

$X(\omega)=i$

$= \{\omega \mid X(\omega)=i\}$

$P(X(\omega)=0) = \frac{1}{3}$



Review Random Variables

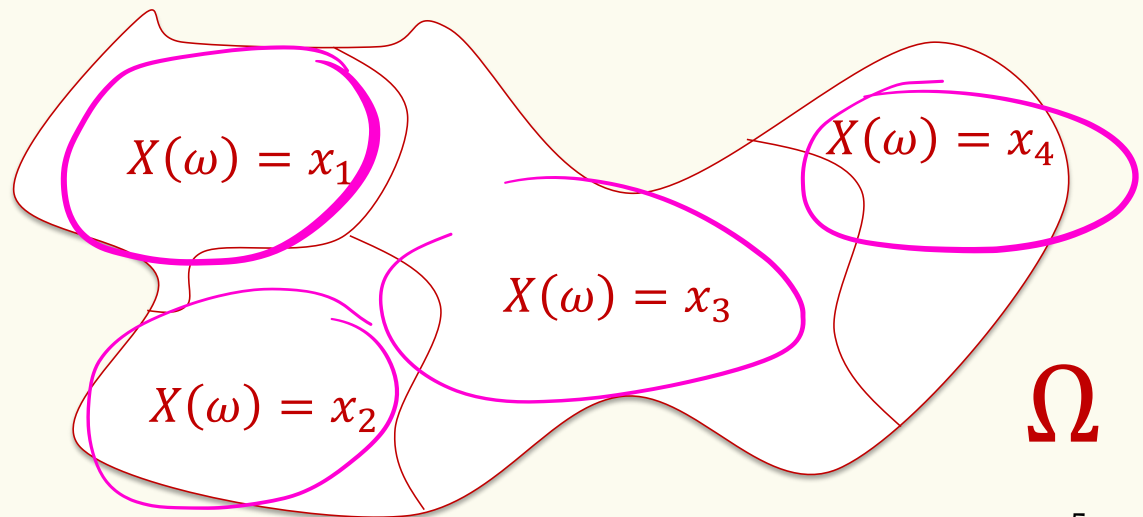
Definition. A **random variable (RV)** for a probability space (Ω, P) is a function $X: \Omega \to \mathbb{R}$.

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

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

Random variables **partition** the sample space.

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



$\Omega_X = \{x_1, x_2, x_3, x_4\}$

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 $X = x$

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

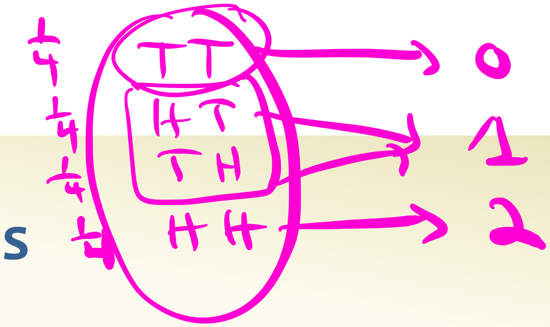
$$\boxed{\sum_{x \in \Omega_X} p_X(x) = 1}$$

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

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

r.v.

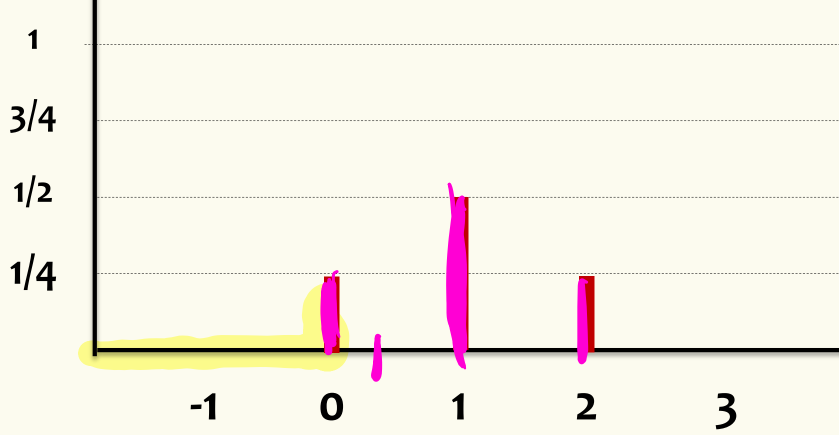
Example – Two fair independent coin flips



X = number of heads

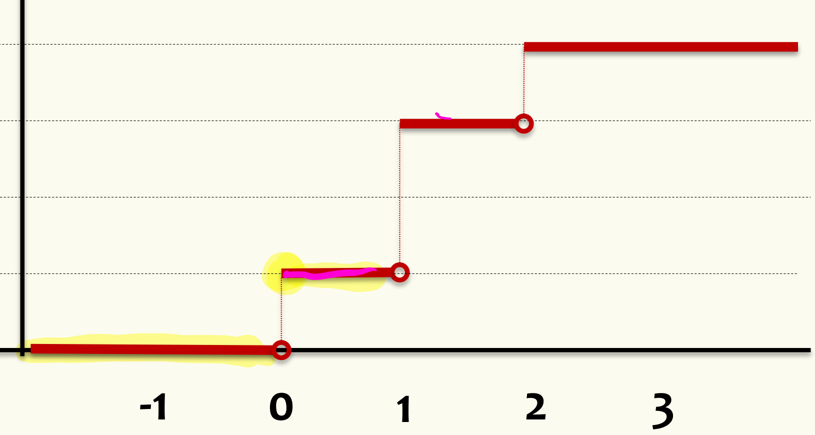
Probability Mass Function
PMF

p_X



Cumulative Distribution Function
CDF

F_X



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

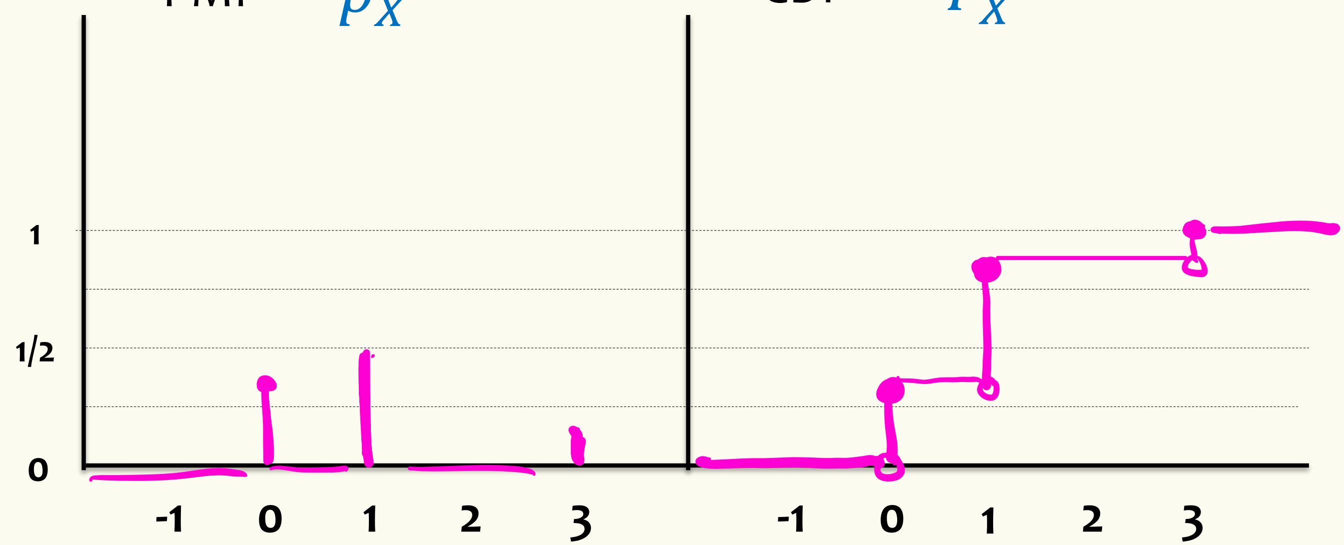
- Class with 3 students, randomly hand back homeworks.

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



$P(X = 0) = 1/3$

$P(X = 1) = 1/2$

$P(X = 3) = 1/6$

Example – Number of Heads

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

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

$|\Omega| = 2^n$

$X =$ # of heads

$\Omega_X = \{0, 1, 2, \dots, n\}$

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

$P(\underbrace{HH\ldots H}_{k}\underbrace{T\ldots T}_{n-k}) = p^k(1-p)^{n-k}$

$P(T\underbrace{H\ldots H}_{k}\underbrace{T\ldots T}_{n-k-1}) = p^k(1-p)^{n-k}$

Example – Number of Heads

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

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

$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)

$\Omega_X = \{0, 1, \ldots, 20\}$

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

$20 \cdot \frac{1}{4} =$

$X =$ number of heads

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

5

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

np.

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

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

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

Expectation

np

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

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

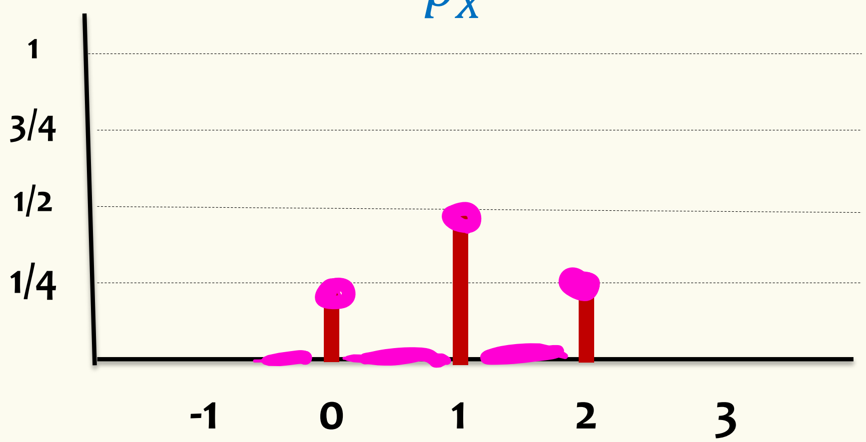
Example. Two fair coin flips

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

$X =$ number of heads

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

p_X



P(X=0) P(X=1) P(X=2)

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

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

Example: Returning Homeworks

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

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

- 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 $\mathbb{E}[X]$?

$\mathbf{Pr}(\omega)$	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

$$E(X) = 0 \cdot \tfrac{1}{6} + 0 \cdot \tfrac{1}{6} + 1 \cdot \tfrac{1}{6} + 1 \cdot \tfrac{1}{6} + 1 \cdot \tfrac{1}{6} + 3 \cdot \tfrac{1}{6}$$

$$= 0 \cdot P(X=0) + 1 \cdot P(X=1) + 3 \cdot P(X=3)$$

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

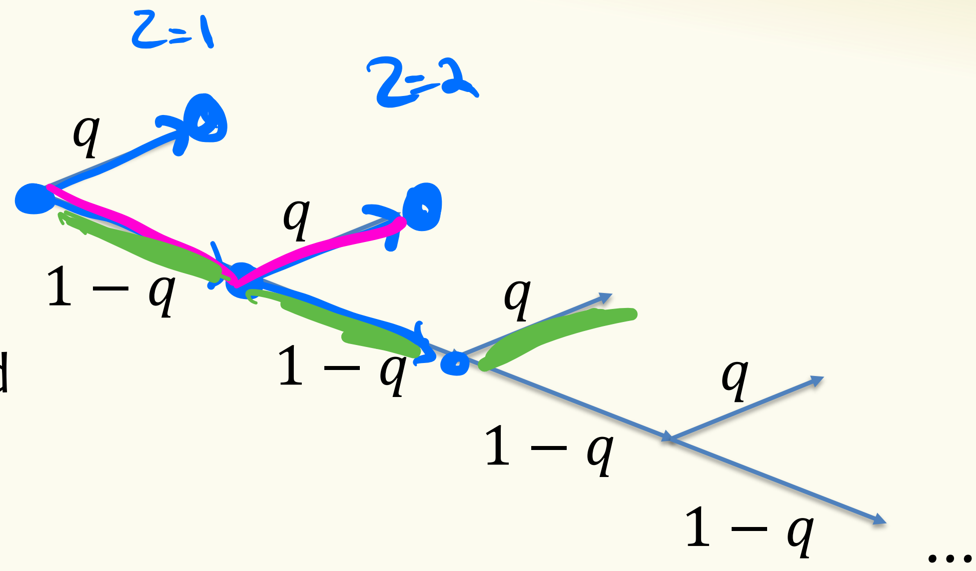
$\Pr(\omega)$	ω	$X(\omega)$
$1/6$	1, 2, 3	3
$1/6$	1, 3, 2	1
$1/6$	2, 1, 3	1
$1/6$	2, 3, 1	0
$1/6$	3, 1, 2	0
$1/6$	3, 2, 1	1

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

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

$E(Z) = \sum_{x \in \Omega_Z} x \, P(Z=x)$ $\Omega_X = \{1, 2, 3, \ldots\}$

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

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

$P(Z=1) = q$

$P(Z=2) = (1-q)q$

$P(Z=3) = (1-q)^2 q$

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

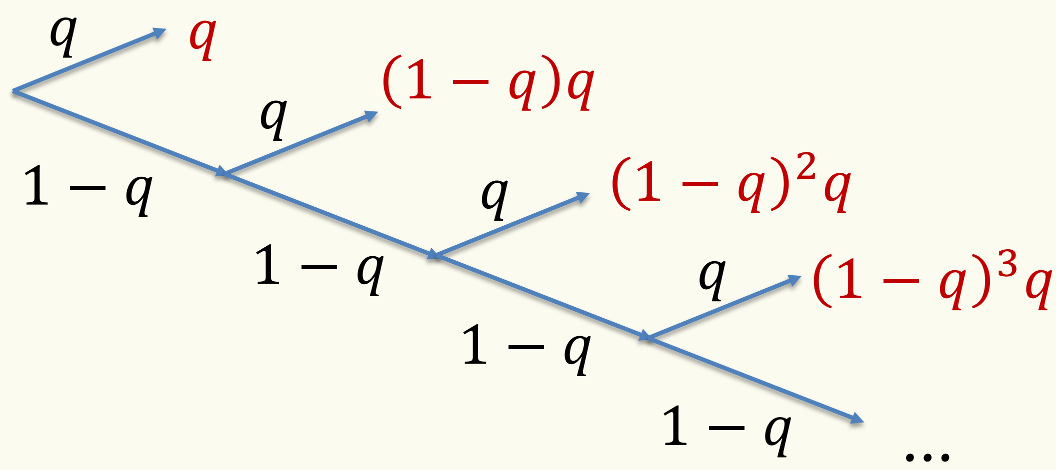
Example – Flipping a biased coin until you see heads

- 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…

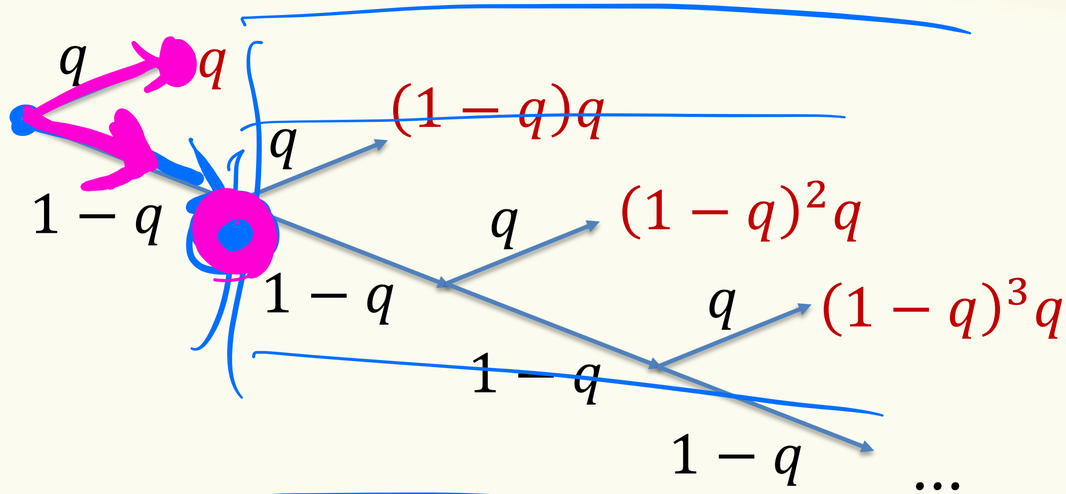
Example – Flipping a biased coin until you see heads

- 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

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

$$= q + (1-q) + (1-q)E(Z)$$

$$qE(Z) = 1 \implies E(Z) = \frac{1}{q}$$

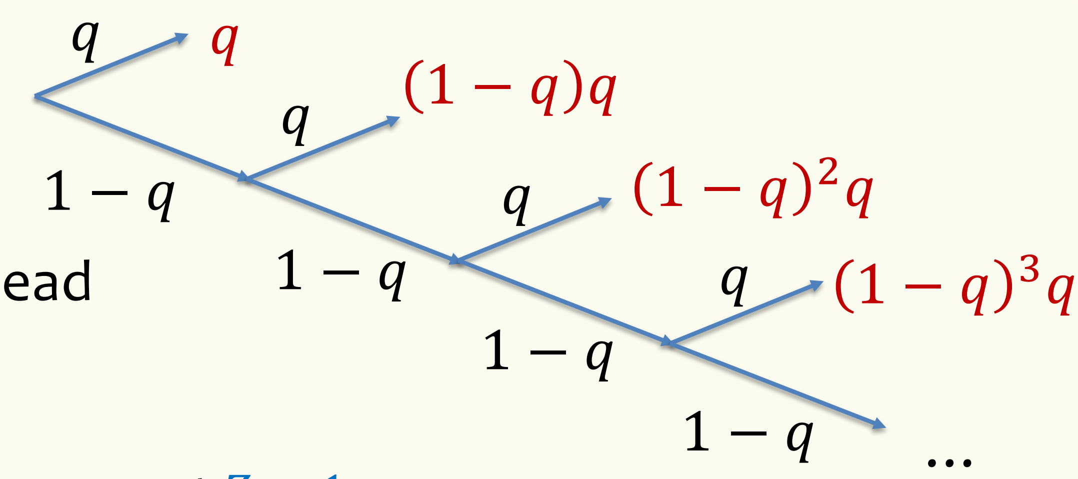
Example – Flipping a biased coin until you see heads

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

Example – Coin Tosses

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

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

$$P(Z=k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & k \in \{0,1,\dots,n\} \\ 0 & \text{otherwise.} \end{cases}$$

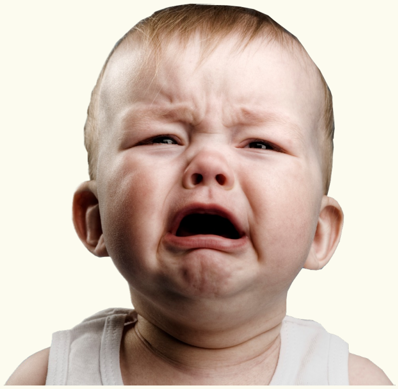
$$E(Z) = \sum_{k=0}^{n} k \, P(Z=k)$$

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

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

$\omega \to X(\omega)$

$\omega \to Y(\omega)$

$Z(\omega) = X(\omega) + Y(\omega)$

Linearity of Expectation

Theorem. For any two random variables X and Y

(no conditions whatsoever on the random variables)

(defined on common prob space)

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

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

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

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

Theorem. For any two random variables X and Y

(X, Y do not need to be independent)

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

Z

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

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

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

E(X)

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

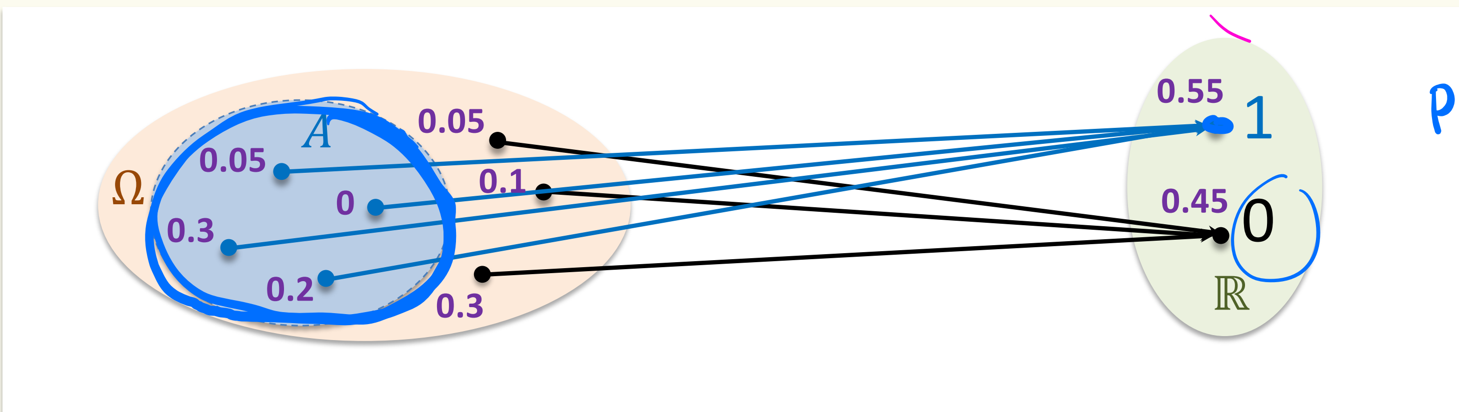
Often, X_i are **indicator** (0/1) random variables.

Indicator random variables – 0/1 valued

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

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

$$\boxed{\begin{aligned} P(X_A = 1) &= P(A) \\ P(X_A = 0) &= 1 - P(A) \end{aligned}}$$



$P(X_A=1)$

$$E(X_A) = \cancel{0 \cdot P(X_A=0)} + 1 \cdot P(X_A=1)$$

$$= P A$$