CSE 312

Foundations of Computing II

Lecture 9: Linearity of Expectation

Last Class:

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

Today:

- An Expectation Example
- Linearity of Expectation
- Indicator Random Variables

Kandinsky

Review Random Variables

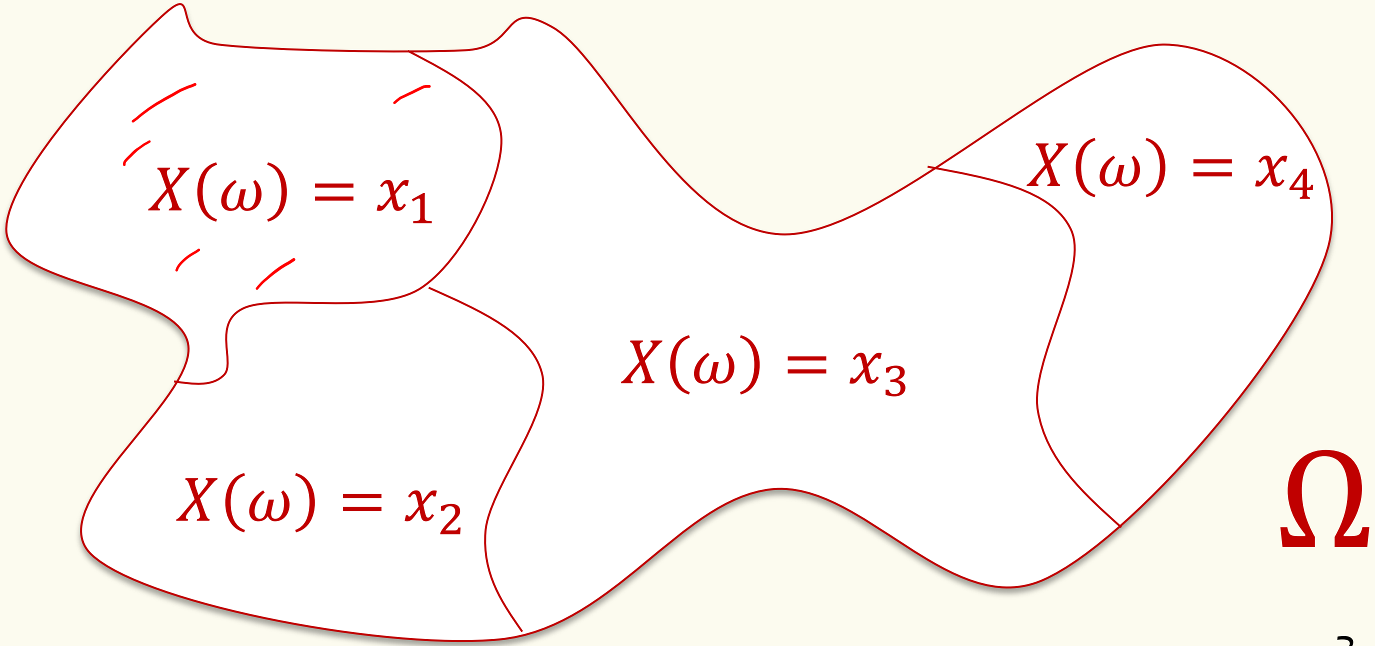
Definition. A **random variable (RV)** for a probability space (Ω, P) is a function $X: \Omega \to \mathbb{R}$.

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

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

Random variables **partition** the sample space.

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



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 $X \le x$

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

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

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

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

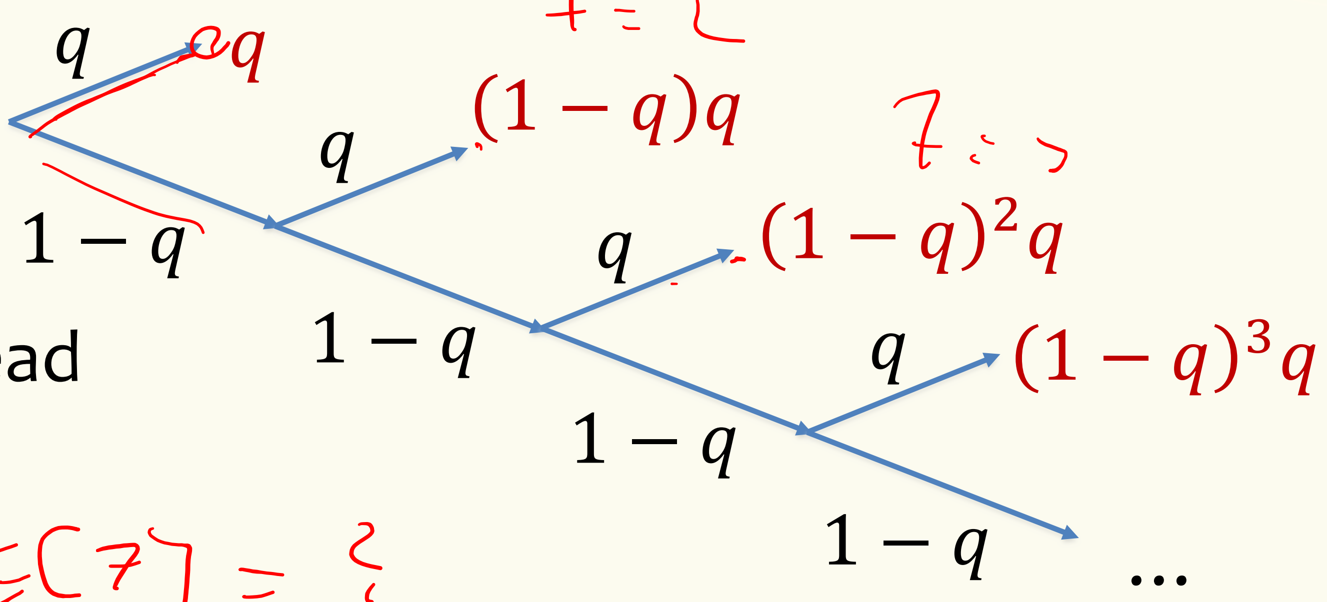
Example – Flipping a biased coin until you see heads

- 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…

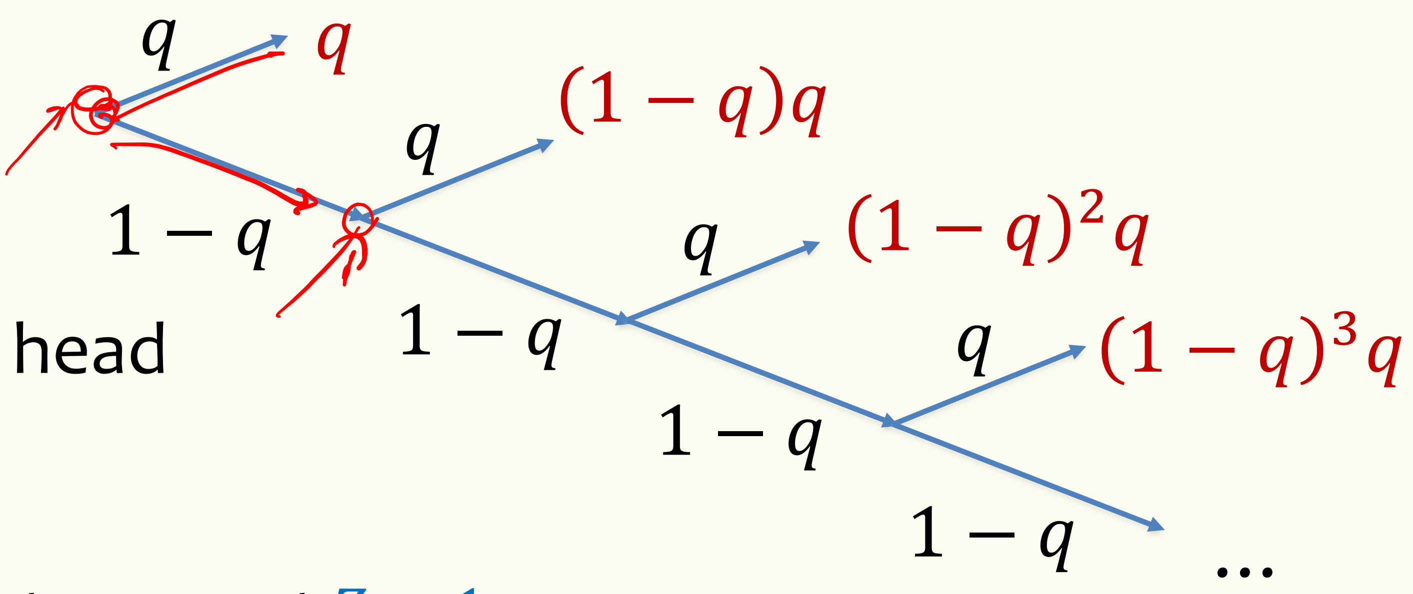
Example – Flipping a biased coin until you see heads

- 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

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

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

Implies $\mathbb{E}[Z] = 1/q$

Linearity of Expectation (Idea)

Let's say you and your friend sell fish for a living.

- Every day you catch X fish, with $\mathbb{E}[X] = 3$.
- Every day your friend catches Y fish, with $\mathbb{E}[Y] = 7$.

How many fish do the two of you bring in ($Z = X + Y$) on an average day?

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

You can sell each fish for \$5 at a store, but together you need to pay \$20 in rent. How much profit do you expect to make?

$$\mathbb{E}[5Z - 20] = 5\mathbb{E}[Z] - 20 = 5 \times 10 - 20 = 30$$

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 $\mathbb{E}[Z]$?

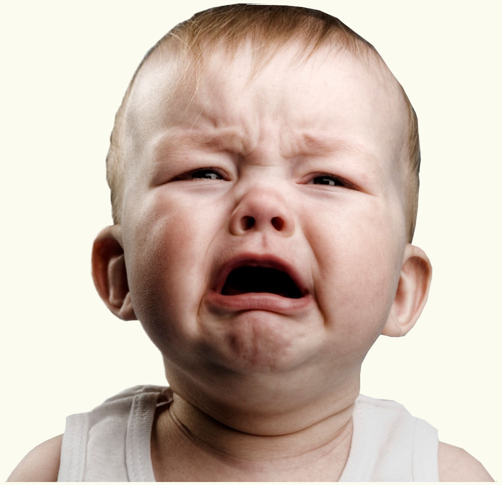
$$\mathbb{E}[Z] = \sum_{k=0}^{n} k \cdot P(Z = k) = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

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

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

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

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



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 $\mathbb{E}[Z]$?

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

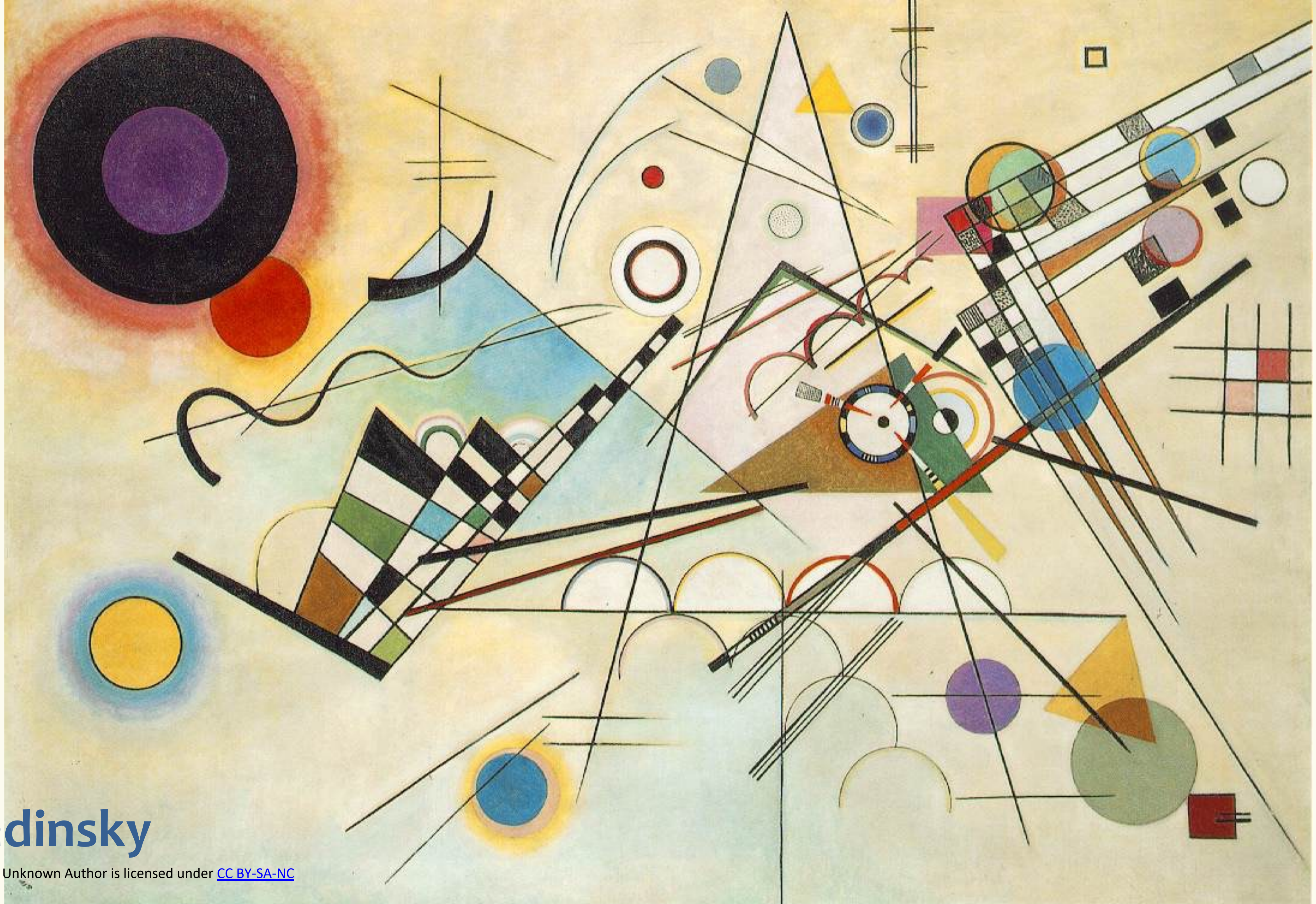
Fact. $Z = X_1 + \cdots + X_n$

Linearity of Expectation:

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

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

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



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: What is X_i?

$X_i = 1$ iff i^{th} student gets own HW back (0 o/.w)

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

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

$= P(X_i = 1)$

$= \frac{(n-1)!}{n!}$

$= \frac{1}{n}$

Poll: **pollev.com/stefanotessaro617**

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}

$P(X_{ij}=1) = \frac{365}{365^2} = \frac{1}{365}$

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 $X_1, \ldots, X_n$, and real numbers $a_1, \ldots, a_n \in \mathbb{R}$,

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

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

Linearity is special!

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

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

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

Kandinsky