CSE 312

Foundations of Computing II

Lecture 10: Linearity of Expectation

PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Rachel Lin, Hunter Schafer

Slide Credit: Based on Stefano Tessaro's slides for 312 19au
incorporating ideas from Alex Tsun's and Anna Karlin's slides for 312 20su and 20au

Last Class:

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

Today:

- Linearity of Expectation
- Indicator Random Variables

Kandinsky

Expectation of Random Variable

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

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

or equivalently

$$\Pr(X = x)$$

$$= \Pr(\{\omega : X(\omega) = x\})$$

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

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

Linearity of Expectation (Idea)

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

- Every day you catch **X** fish, with **E[X] = 3.**
- Every day your friend catches **Y** fish, with **E[Y] = 7.**

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

E[Z] = E[X + Y] = E[X] + E[Y] = 3 + 7 = 10

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

E[5Z - 20] = 5E[Z] - 20 = 5 x 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 $X_1, \ldots, X_n$,

$$\mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n).$$

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)

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

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

$\sum_z z \cdot P(X+Y=z)$

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

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

$$= \mathbb{E}(X) + \mathbb{E}(Y)$$

Example – Coin Tosses

we flip n coins, each one heads with probability p

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

The brute force method

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

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

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

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

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

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

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

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.

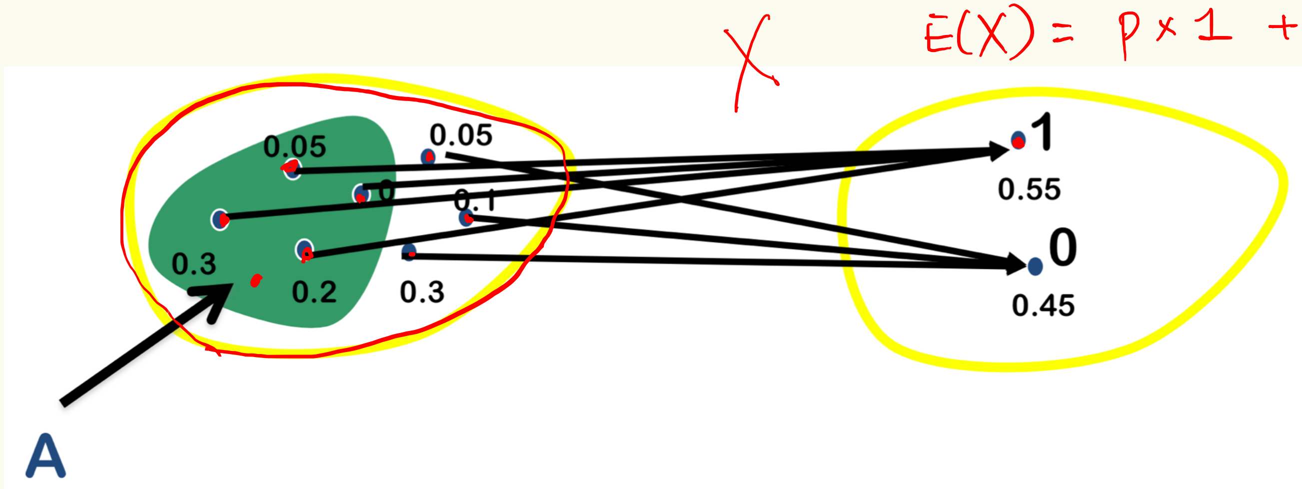
Indicator random variable

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

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

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

$E(X) = p \times 1 + (1-p) \times 0 = p$



Example – Coin Tosses

H H H T T T

$X_1=1$ $X_2=1$ $X_3=1$ $X_4=0$ $X_5=0$ $X_6=0$

we flip n coins, each one heads with probability p

3

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

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

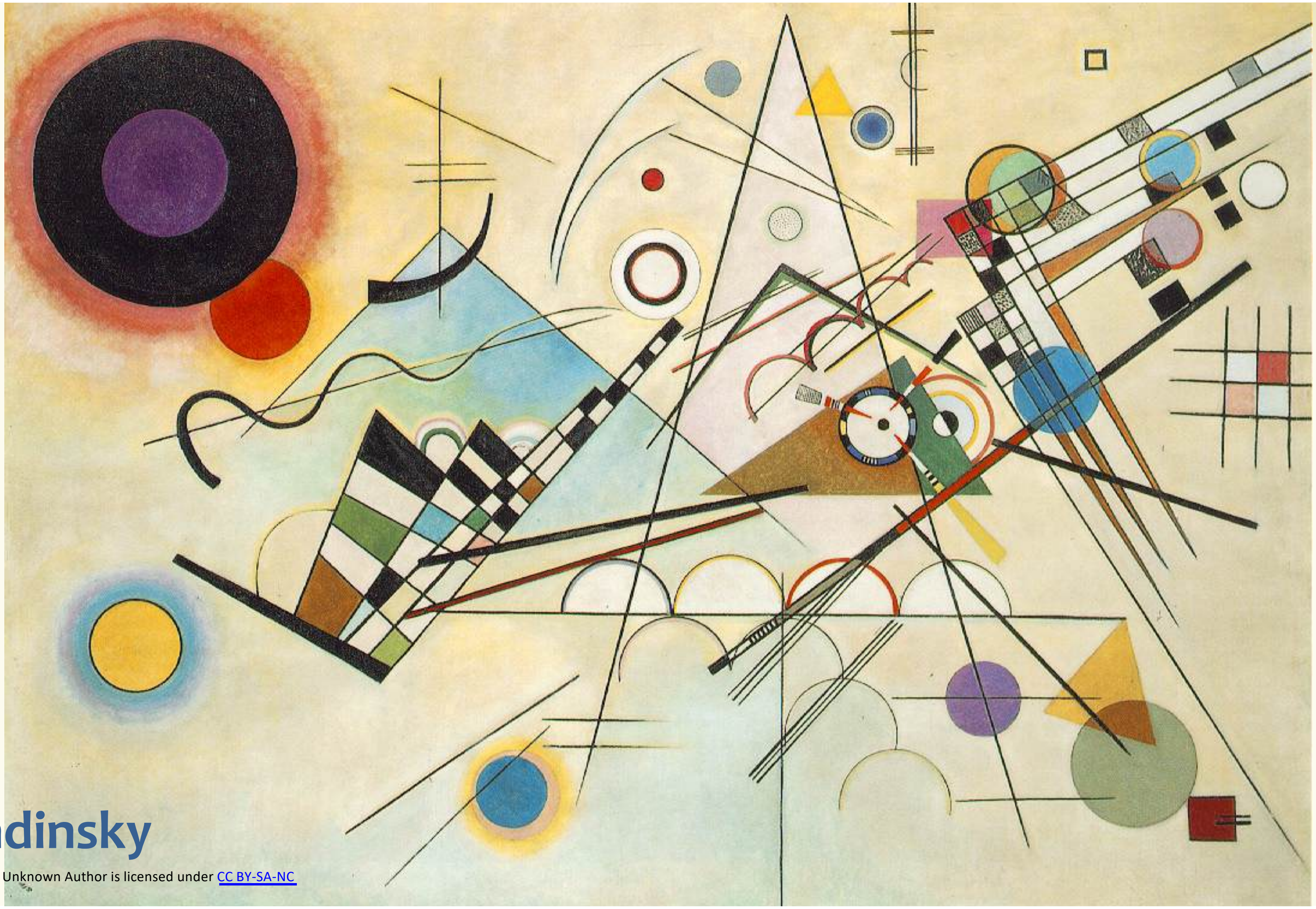
Fact. $Z = X_1 + \cdots + X_n$

Linearity of Expectation:

$$\mathbb{E}(Z) = \mathbb{E}(X_1 + \cdots + X_n) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = n \cdot p$$

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

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



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 $\mathbb{E}(X)$?

$X_i = \begin{cases} 1 & \text{if student } i \text{ gets HW back.} \\ 0 & \text{if " } i \text{ does not get HW back.} \end{cases}$

$\Pr(\omega)$	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

Poll: pollev.com/rachel312

Decompose: $\boxed{\text{What is } X_i\text{?}}$ $X = \sum_{i=1}^{n} X_i$

LOE: $E(X) = \sum_{i=1}^{n} E(X_i) = n \times \frac{1}{n} = 1$ $E(X_i) = P(X_i = 1) = \frac{1}{n}$

Conquer: What is $\mathbb{E}(X_i)$? A. $\frac{1}{n}$ B. $\frac{1}{n-1}$ C. ½

Pairs with same birthday

$(1,2)$ $(1,3)$ $(10,100)$

1 0

X

- In a class of m students, on average how many pairs of people have the same birthday?

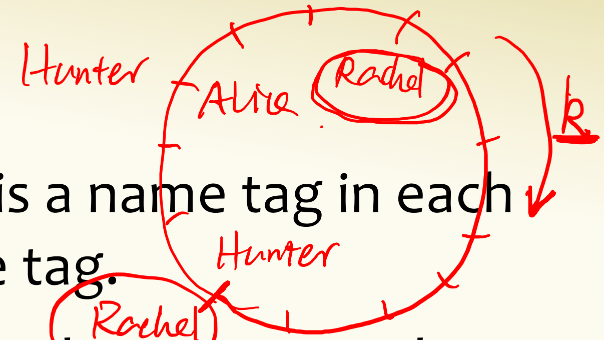
$\binom{m}{2}$

Decompose: $X_{ij} = \begin{cases} 1 & \text{if } i, j \text{ have same bday} \\ 0 & \text{otherwise.} \end{cases}$ $X = \sum_{1 \le i < j \le m} X_{ij}$

LOE: $E(X) = \sum_{ij} E(X_{ij})$.

Conquer: $E(X_{ij}) = P(X_{ij} = 1) = \frac{365}{365 \times 365} = \frac{1}{365}$

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 front of their own name tag.

What is E(X)?

Decompose: $X_i = \begin{cases} 1 & \text{if } i \text{ sits infront of name tag} \\ 0 & \text{if } i \text{ not sit in front of name tag} \end{cases}$ $\qquad X = \sum_{i=1}^{n} X_i$

LOE: $E(X) = \sum_{i=1}^{n} E(X_i) = n \cdot \frac{1}{n-1}$

Conquer: $E(X_i) = P(X_i = 1) = \frac{1}{n-1}$

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 <u>not</u> have $\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$

Linearity is special!

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

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

- $\mathbb{E}(XY) \neq \mathbb{E}(X)\mathbb{E}(Y)$
- $\mathbb{E}(X/Y) \neq \mathbb{E}(X)/\mathbb{E}(Y)$
- $\mathbb{E}(X^2) \neq \mathbb{E}(X)^2$

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

Expectation of $g(X)$

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

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

or equivalently

$$\mathrm{E}[g(X)] = \sum_{x \in X(\Omega)} g(x) \cdot \Pr(X = x)$$

Kandinsky

Take Home FUN Example – Coupon Collector Problem

Say each round we get a random coupon $X_i \in \{1, \dots, n\}$, how many rounds (in expectation) until we have one of each coupon?

Formally: Outcomes in Ω are sequences of integers in $\{1, \dots, n\}$ where each integer appears at least once (+ cannot be shortened).

Example, $n = 3$:

$$\Omega = \{(1,2,3), (1,1,2,3), (1,2,2,3), (1,2,3), (1,1,1,3,3,3,3,3,3,2), \dots\}$$

$$\mathbb{P}\big((1,2,3)\big) = \frac{1}{3} \cdot \frac{1}{3} \cdot \frac{1}{3} \qquad \mathbb{P}\big((1,1,2,2,2,3)\big) = \left(\frac{1}{3}\right)^6 \qquad \dots$$

Example – Coupon Collector Problem

Say each round we get a random coupon $X_i \in \{1, \dots, n\}$, how many rounds (in expectation) until we have one of each coupon?

$T_i =$ # of rounds until we have accumulated i distinct coupons

[Aka: length of the sampled ω]

Wanted: $\mathbb{E}(T_n)$

Hard to think about T_n directly,
Can we decompose T_n as a sum of simpler random variables?

$Z_i = T_i - T_{i-1}$

# of rounds needed to go from $i - 1$ to i coupons

Example – Coupon Collector Problem

T_i = # of rounds until we have accumulated i distinct coupons

Wanted: $\mathbb{E}(T_n)$

$Z_i = T_i - T_{i-1}$

$$T_n = T_1 + (T_2 - T_1) + (T_3 - T_2) + \cdots + (T_n - T_{n-1}) = T_1 + Z_2 + \cdots + Z_n$$

$$\Downarrow$$

$$\begin{aligned}\mathbb{E}(T_n) &= \mathbb{E}(T_1) + \mathbb{E}(Z_2) + \mathbb{E}(Z_3) + \cdots + \mathbb{E}(Z_n) \\ &= \underline{1 + \mathbb{E}(Z_2) + \mathbb{E}(Z_3) + \cdots + \mathbb{E}(Z_n)}\end{aligned}$$

Wanted: $\mathbb{E}(Z_i)$

Example – Coupon Collector Problem

$T_i =$ # of rounds until we have accumulated i distinct coupons

$Z_i = T_i - T_{i-1}$

Wanted: $\mathbb{E}(Z_i)$

If we have accumulated $i-1$ coupons, the number Z_i of attempts needed to get the i-th coupon is **geometric** with parameter $p = 1 - \frac{(i-1)}{n}$.

$$\mathbb{p}_{Z_i}(1) = p \qquad \mathbb{p}_{Z_i}(2) = (1-p)p \qquad \cdots \qquad \mathbb{p}_{Z_i}(i) = (1-p)^{i-1}p$$

$$\mathbb{E}[Z_i] = \frac{1}{p} = \frac{n}{n-i+1}$$

Expectation of geometric distribution shown in last lecture,
for the example #coin tosses to see first head

Example – Coupon Collector Problem

$T_i =$ # of rounds until we have accumulated i distinct coupons

$$Z_i = T_i - T_{i-1} \qquad \mathbb{E}(Z_i) = \frac{1}{p} = \frac{n}{n-i+1}$$

n-th **harmonic number**

$$H_n = \sum_{i=1}^{n} \frac{1}{i}$$

$$\ln(n) \leq H_n \leq \ln(n) + 1$$

$$\mathbb{E}(T_n) = 1 + \mathbb{E}(Z_2) + \mathbb{E}(Z_3) + \cdots + \mathbb{E}(Z_n)$$

$$= 1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}$$

$$= n \cdot \left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} + 1\right) = n \cdot H_n \approx n \cdot \ln(n)$$