

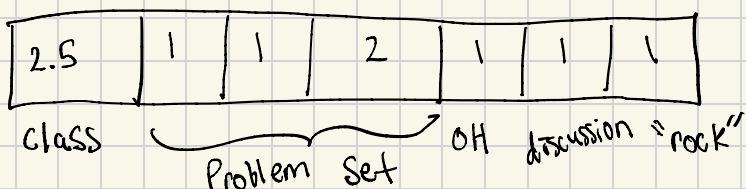
This class:

- probability + linear algebra
- data science at scale
- beautiful math

} challenging course!

Prereqs: linear algebra, probability, algorithms

Time Expectation: 12.5 hr/week



- Earlier = better
- More drawing = better
- More talking = better

- discord for communication
- reading for each lecture
- these slides are online

Plan:

- Math Review
- Streaming and Sketching
- Linear Algebra and Spectral Methods

Course Goal: Use randomness to solve problems more efficiently

Probability

Random variable X

e.g., $X = \text{outcome of fair dice}$

For $x \in \{1, 2, 3, 4, 5, 6\}$,

$$\Pr(X=x) = 1/6$$

What is X on average?

Expectation of X :

$$\mathbb{E}[X] = \sum_x x \Pr(X=x)$$

$$= \sum_{x=1}^6 x \cdot 1/6$$

$$= 1/6(1+2+\dots+6)$$

$$= 21/6$$

How close is X to its expectation?

Variance of X :

$$\text{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

$$= \sum_x (x - \mathbb{E}[X])^2 \Pr(X=x)$$

=

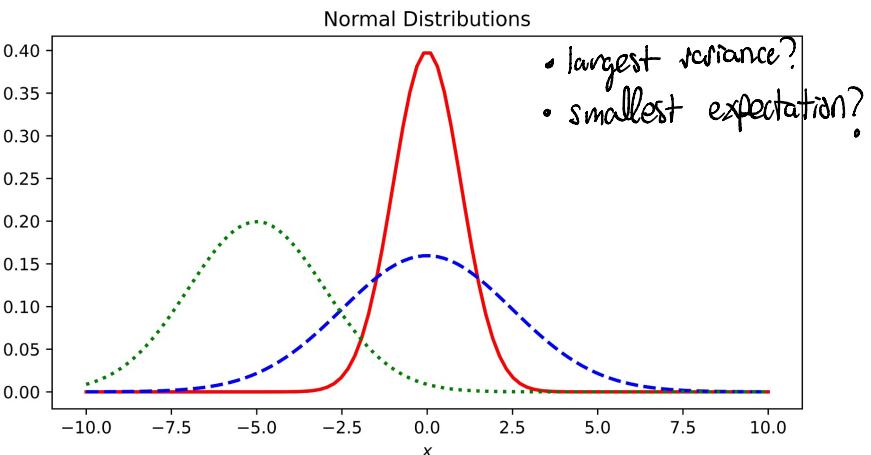
Multiplying by Scalars

$$\mathbb{E}[\alpha X] = \sum_x \alpha x \Pr(X=x)$$

$$= \alpha \sum_x x \Pr(X=x)$$

$$= \alpha \mathbb{E}[X]$$

$$\text{Var}(\alpha X) =$$



Independent Random Variables

A, B are events

$A =$ dice shows 1 or 2

$B =$ dice roll is odd

$$\Pr(A \cap B) = \Pr(A) \Pr(B | A)$$

$$= \Pr(B) \Pr(A | B)$$

“and” “given”

$$\Pr(B | A) = \Pr(B) \text{ iff } A, B \text{ independent}$$

“if and only if”

$$\Leftrightarrow \Pr(A \cap B) = \Pr(A) \Pr(B) \text{ iff } A, B \text{ independent}$$

Random Variables X, Y are indep

iff $\Pr(X=x \cap Y=y) = \Pr(X=x) \Pr(Y=y)$
for all possible outcomes x, y

Independent?

- Heads on 1st flip, Heads on 2nd
- Ace on 1st draw, Ace on 2nd
- Rain, carrying an umbrella

Example A, B indep?

Linearity of Expectation

Thm: $E[X+Y] = E[X] + E[Y]$

$$\begin{aligned} \text{Proof: } E[X+Y] &= \sum_x \sum_y (x+y) \Pr(X=x \cap Y=y) \\ &= \sum_x \sum_y x \Pr(X=x \cap Y=y) + \sum_x \sum_y y \Pr(X=x \cap Y=y) \\ &= \sum_x x \sum_y \Pr(X=x \cap Y=y) + \sum_y y \sum_x \Pr(Y=y \cap X=x) \\ &= \sum_x x \Pr(X=x) + \sum_y y \Pr(Y=y) \\ &= E[X] + E[Y] \end{aligned}$$

Expected number of shared birthdays?

Useful Facts

If X, Y indep, $E[XY] = E[X]E[Y]$

$$\text{Var}(X) = E[X^2] - E[X]^2$$

If X, Y indep, $\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$