

CS 260: Foundations of Data Science

Prof. Thao Nguyen

Fall 2026



HAVERFORD
COLLEGE

Admin

- Lab 4 due tonight



Admin

- **Mid-semester feedback form** (link on Piazza)
- **Practice midterm solutions** posted
- **No lab** tomorrow

Outline for today

- Review
- Introduction to probability

Outline for today

- Review
- Introduction to probability

Runtime of matrix operations

- **Multiplication**

$$\begin{matrix} A \\ \uparrow \\ n \times p \end{matrix} \begin{matrix} B \\ \uparrow \\ p \times m \end{matrix} = \begin{bmatrix} \text{ith row} \end{bmatrix} \begin{bmatrix} \text{jth column} \end{bmatrix} = \begin{bmatrix} \text{i,j entry} \end{bmatrix} \quad n \times m$$

entries: $O(nm)$
one entry: $O(p)$ } Runtime = $O(nmp)$

- **Inverse**

Matrix must be square, $(n \times n) \rightarrow O(n^3)$

Analytic solution to linear regression

$$\textcircled{11} \quad \vec{w} = \underbrace{\underbrace{(X^T X)^{-1}}_{(d)} \underbrace{X^T}_{(c)}}_{(a)} \underbrace{\vec{y}}_{(b)}$$

(p+1) x (p+1) + (p+1) x n

(c) $X^T \Rightarrow O(np)$

(a) multiply (p+1) x n + n x (p+1) $\Rightarrow O(p^2 n)$

(b) invert (p+1) x (p+1) matrix $\Rightarrow O(p^3)$
inverse is cubic

(c) multiply (p+1) x n + (n x 1) $\Rightarrow O(pn)$

(d) multiply (p+1) x (p+1) + (p+1) x 1 $\Rightarrow O(p^2)$

$\Rightarrow O(p^2 n + p^3 + pn + p^2) \Rightarrow \boxed{O(p^2 n + p^3)} \star$

⑧

$\Rightarrow$
 $\Rightarrow$
 $\Rightarrow$

Outline for today

- Review
- Introduction to probability

Intro to Probability

- The **probability** of an **event** e has a number of epistemological interpretations
- Assuming we have **data**, we can count the number of times e occurs in the dataset to estimate the probability of e , $P(e)$

$$P(e) = \frac{\text{count}(e)}{\text{count}(\text{all events})}$$

- If we put all events a bag, shake it up, and choose one at random (called **sampling**), how likely are we to get e ?

Intro to Probability



- Suppose we flip a fair coin
- What is the probability of heads, $P(e = H)$?
- We have “all” of two possibilities, $e \in \{H, T\}$
- $$P(e = H) = \frac{\text{count}(H)}{\text{count}(H) + \text{count}(T)}$$

Intro to Probability



- Suppose we have a fair 6-sided die
- What's the probability of getting "1"?

$$\frac{\textit{count}(s)}{\textit{count}(1) + \textit{count}(2) + \textit{count}(3) + \cdots + \textit{count}(6)} = \frac{1}{1 + 1 + 1 + 1 + 1 + 1} = \frac{1}{6}$$

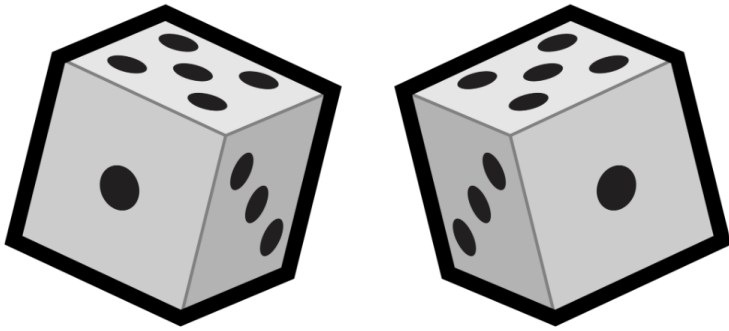
Intro to Probability



- What about a die with only three numbers {1,2,3}, each of which appears twice?
- What's the probability of getting "1"?

$$P(e = 1) = \frac{\textit{count}(1)}{\textit{count}(1) + \textit{count}(2) + \textit{count}(3)} = \frac{2}{2 + 2 + 2} = \frac{1}{3}$$

Intro to Probability



- Suppose we have two fair 6-sided dice
- What's the probability of the sum of the rolls being “7”?

$$\frac{\text{sum}(7)}{\text{all_combinations}} = \frac{6}{6 \times 6} = \frac{1}{6}$$

Intro to Probability



- The set of all probabilities for an event e is called a **probability distribution**
- Each coin toss is an independent event (Bernoulli trial)

Intro to Probability



- Which is greater, $P(HHHHHH)$ or $P(HHTHHH)$?
- Since the events are independent, they are equal

Intro to Probability

Probability Axioms

1. Probabilities of events must be no less than 0. $P(e) \geq 0$ for all e .

2. The sum of all probabilities in a distribution must sum to 1.

That is, $P(e_1) + P(e_2) + \cdots + P(e_n) = 1$

Or, more succinctly, $\sum_{e \in E} P(e) = 1$

Intro to Probability

Joint Probability

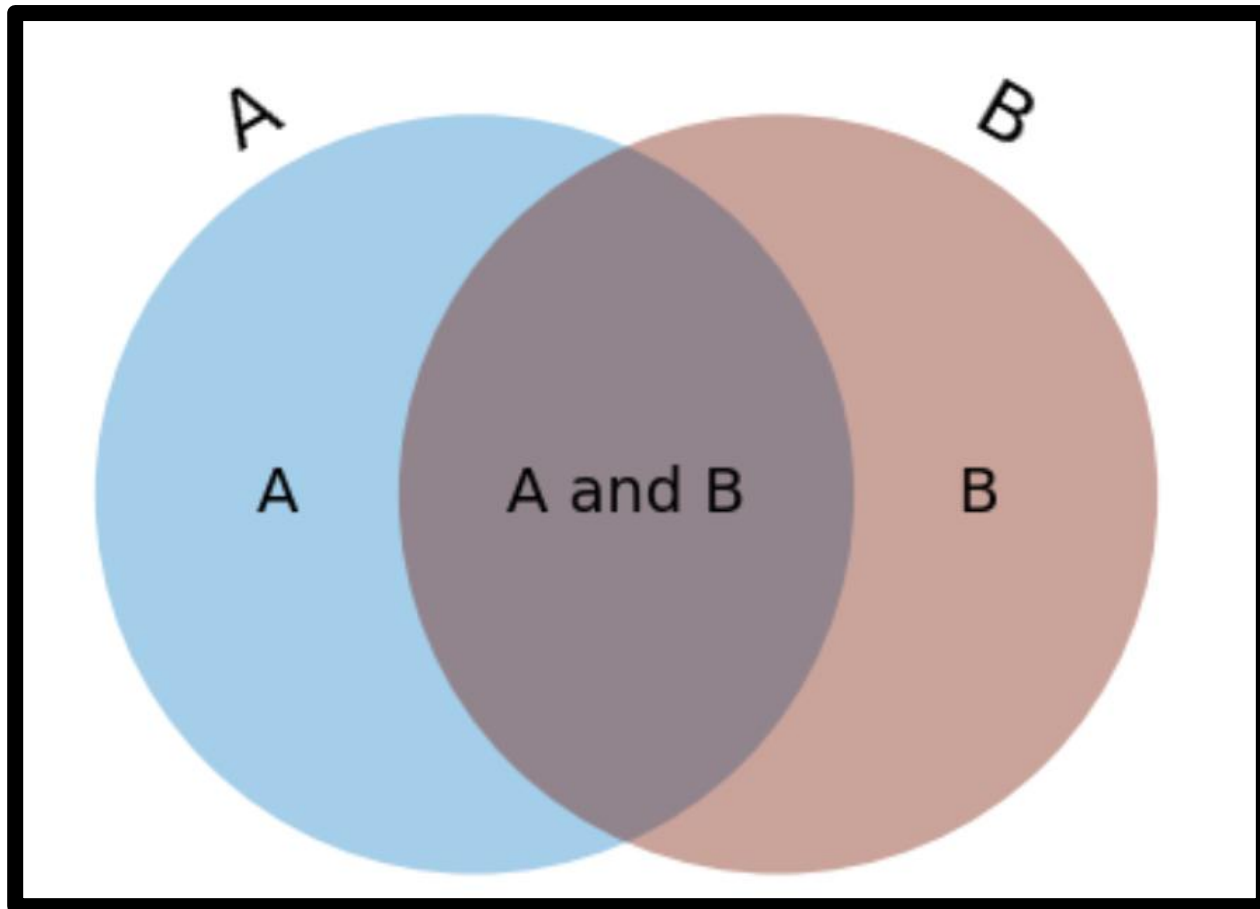
The probability that two independent events e_1 and e_2 both occur is given by their product.

$$P(e_1 \wedge e_2) = P(e_1 \cap e_2) = P(e_1)P(e_2) \text{ when } e_1 \cap e_2 = \emptyset$$

- Intuitively, think of every probability as a *scaling* factor.
- You can think of a probability as the fraction of the probability space occupied by an event e_1
 - $P(e_1 \wedge e_2)$ is the fraction of e_1 's probability space wherein e_2 also occurs.
 - So, if $P(e_1) = \frac{1}{2}$ and $P(e_2) = \frac{1}{3}$, then $P(e_2, e_1)$ is a third of a half of the probability space or $\frac{1}{3} \times \frac{1}{2}$

Intro to Probability

Joint Probability



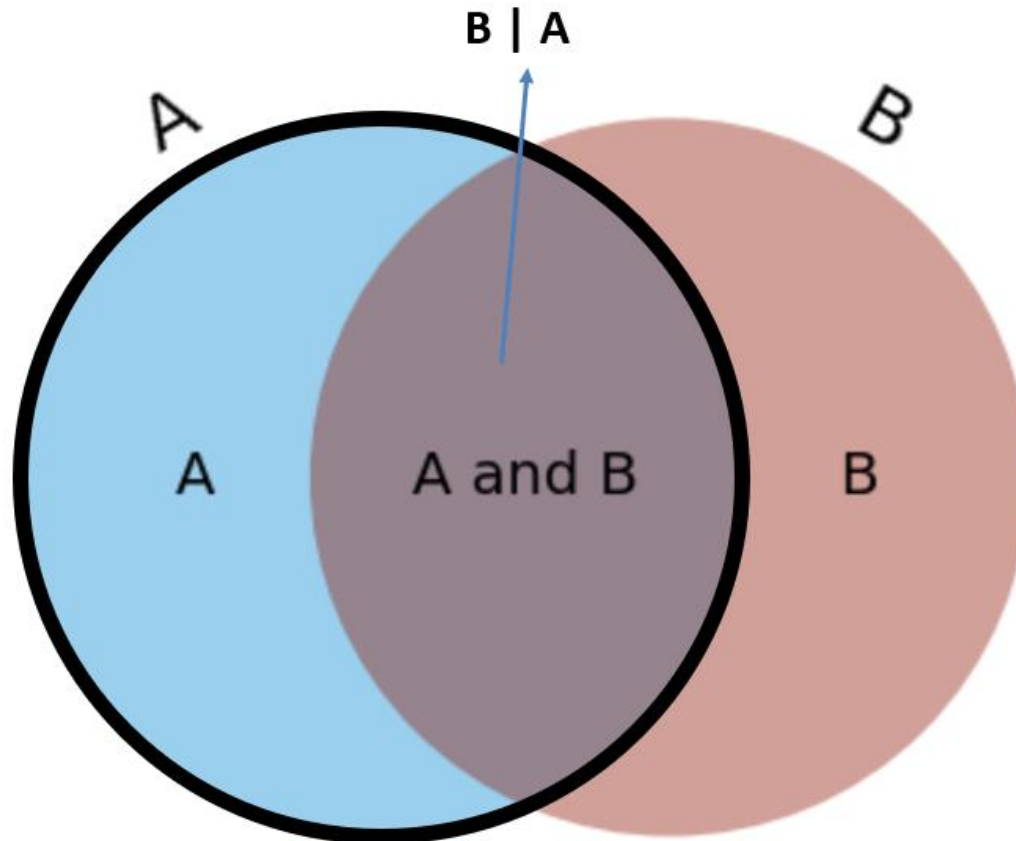
Intro to Probability

Conditional Probability

- A conditional probability is the probability that one event occurs given that we take another for granted.
- The probability of e_2 given e_1 is $P(e_2 \mid e_1)$
- This is the probability that e_2 will occur given that we take for granted that e_1 occurs.

Intro to Probability

Conditional Probability



Intro to Probability

Marginal Probability Distributions

Given a discrete joint probability distribution function $P(X, Y)$, how would we find $P(X)$?

- “Marginalize out” the Y (sum over all $y \in Y$)
- Discrete Case: $p(x) = \sum_{y \in Y} P(x, y)$
- Continuous Case: $p(x) = \int p(x, y) dy$

Intro to Probability

Marginal Probability Distributions

Example: $P(u) = P(u, \text{rain}) + P(u, \overline{\text{rain}})$

