

# Lecture 16 - Conditional Probability, Sequences and Permutations



DSC 40A, Spring 2023

## Announcements

- ▶ Discussion is tonight at 7pm or 8pm in FAH 1101. Groupwork 6 is due **tonight at 11:59pm**.
- ▶ Homework 5 is released, due **Tuesday at 11:59pm**.
- ▶ **Important:** We've posted **many** probability resources on the [resources tab of the course website](#). These will no doubt come in handy.
  - ▶ No more DSC 40A-specific readings, though the Probability Roadmap was written specifically for students of this course.

today's slides: on Campuswire

# Agenda

- ▶ Conditional probability.
- ▶ Simpson's Paradox.
- ▶ Sequences and permutations.

## Conditional probability

## Last time

- ▶  $\bar{A}$  is the complement of event A.  $P(\bar{A}) = 1 - P(A)$ .

① complement rule (NOT)

- ▶ For any two events A and B

② addition rule (OR)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

If A and B are mutually exclusive, this simplifies to

no overlap

$$P(A \cup B) = P(A) + P(B).$$

- ▶ The probability that events A and B both happen is

③ multiplication rule (AND)

$$P(A \cap B) = P(A)P(B|A).$$

conditional

- ▶  $P(B|A)$  is the conditional probability of B occurring, given that A occurs. If  $P(B|A) = P(B)$ , then events A and B are independent.

$P(B)$

# Conditional probability

- ▶ The probability of an event may **change** if we have additional information about outcomes.
- ▶ Starting with the multiplication rule,  $P(A \cap B) = P(A)P(B|A)$ , we have that

$$P(B|A) = \frac{P(A \cap B)}{P(A)} \quad \leftarrow$$

assuming that  $P(A) > 0$ .

## Example: pets

Suppose a family has two pets. Assume that it is **equally likely that each pet is a dog or a cat**. Consider the following two probabilities:

1. The probability that both pets are dogs given that **the oldest is a dog**.
2. The probability that both pets are dogs given that **at least one of them is a dog**.

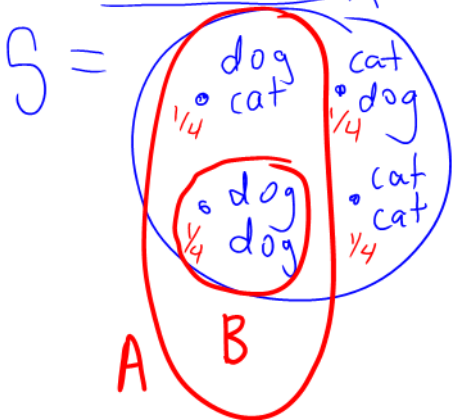
### Discussion Question

Are these two probabilities equal?

- a) Yes, they're equal
- b) No, they're not equal

## Example: pets

Let's compute the probability that both pets are dogs given that the oldest is a dog.

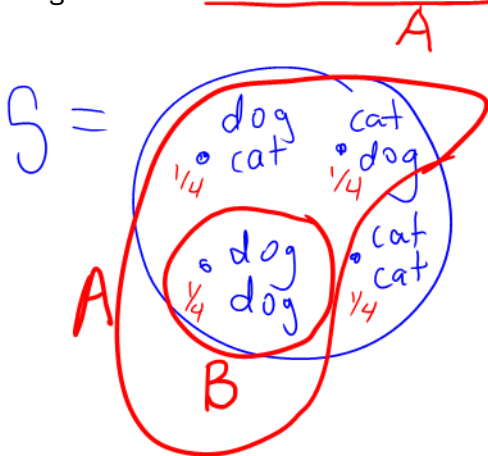


$$\begin{aligned} P(B|A) &= \frac{P(A \text{ and } B)}{P(A)} \\ &= \frac{1/4}{1/2} \\ &= 1/2 \end{aligned}$$



## Example: pets

Let's now compute the probability that both pets are dogs given that at least one of them is a dog.



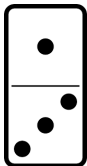
$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

$$= \frac{\frac{1}{4}}{\frac{3}{4}}$$

$$= \frac{1}{3}$$

## Example: dominoes (source: 538)

In a set of dominoes, each tile has two sides with a number of dots on each side: zero, one, two, three, four, five, or six. There are 28 total tiles, with each number of dots appearing alongside each other number (including itself) on a single tile.

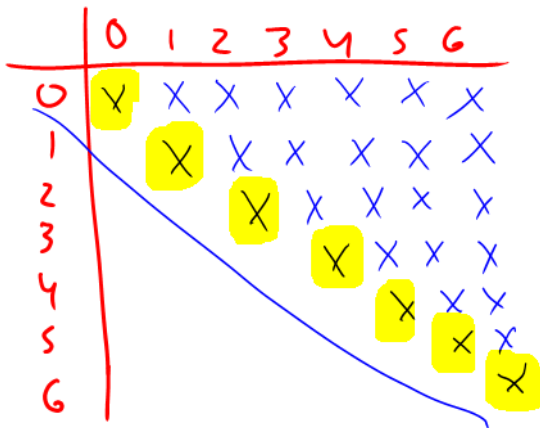


	0	1	2	3	4	5	6	
0	X	X	X	X	X	X	X	→ 7
1		X	X	X	X	X	X	→ 6
2			X	X	X	X	X	→ 5
3				X	X	X	X	→ 4
4					X	X	X	→ 3
5						X	X	→ 2
6							X	→ 1

The table is a lower triangular matrix of 'X' marks. A blue diagonal line runs from the top-left to the bottom-right. To the right of the matrix, blue arrows point from each row to a number (7, 6, 5, 4, 3, 2, 1). A large blue bracket on the far right groups these numbers from 1 to 7.

## Example: dominoes (source: 538)

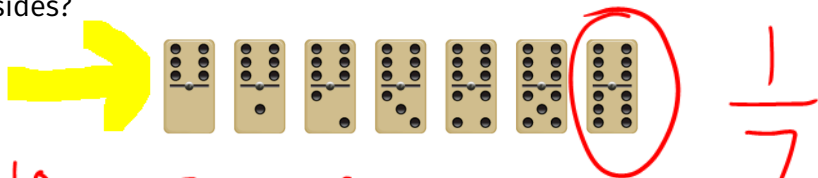
**Question 1:** What is the probability of drawing a “double” from a set of dominoes – that is, a tile with the same number on both sides?



$$\frac{7}{28} = \frac{1}{4}$$

## Example: dominoes (source: 538)

**Question 2:** Now your friend picks a random tile from the set and tells you that at least one of the sides is a 6. What is the probability that your friend's tile is a double, with 6 on both sides?



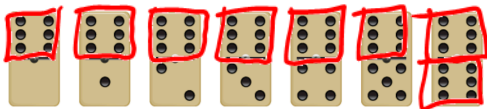
	0	1	2	3	4	5	6
0	X	X	X	X	X	X	X
1		X	X	X	X	X	X
2			X	X	X	X	X
3				X	X	X	X
4					X	X	X
5						X	X
6							X

direct interpretation

$$P(66 | \text{at least one } 6) = \frac{P(66 \text{ AND } \text{at least one } 6)}{P(\text{at least one } 6)}$$
$$= \frac{1/28}{7/28} = 1/7$$

## Example: dominoes (source: 538)

**Question 3:** Now you pick a random tile from the set and uncover only one side, revealing that it has six dots. What is the probability that this tile is a double, with six on both sides?



$\frac{1}{4}$   
 $\frac{1}{7}$

set of domino halves

$28 \cdot 2 = 56$  domino halves

8 of these domino halves have 

direct interpretation: of 8 halves, 2 of them also have six dots across from it

See 538's explanation here.

$$\frac{2}{8} = \frac{1}{4} \rightarrow$$

$$P(66 \mid \text{uncovered side had } 6)$$

$$= P(66 \text{ and uncovered side is } 6)$$

---

$$P(\text{uncovered side is } 6)$$

$$= \frac{1/28}{8/56}$$

← of 28 dominoes, only 1 is 

6
6

$$= 1/4$$

← of all 58 halves they could have uncovered, 8 have 

6
---

## Simpson's Paradox

## Simpson's Paradox (source: [nih.gov](http://nih.gov))

	Treatment A	Treatment B
Small kidney stones	81 successes / 87 (93%)	234 successes / 270 (87%)
Large kidney stones	192 successes / 263 (73%)	55 successes / 80 (69%)
Combined	273 successes / 350 (78%)	289 successes / 350 (83%)

### Discussion Question

Which treatment is better?

- a) Treatment A for all cases.
- b) Treatment B for all cases.
- c) Treatment A for small stones and B for large stones.
- d) Treatment A for large stones and B for small stones.



## Simpson's Paradox ([source: nih.gov](https://www.nih.gov))

	Treatment A	Treatment B
<b>Small kidney stones</b>	81 successes / 87 (93%)	234 successes / 270 (87%)
<b>Large kidney stones</b>	192 successes / 263 (73%)	55 successes / 80 (69%)
<b>Combined</b>	273 successes / 350 (78%)	289 successes / 350 (83%)

**Simpson's Paradox** occurs when an association between two variables exists when the data is divided into subgroups, but reverses or disappears when the groups are combined.

- ▶ See more in DSC 80.

## Sequences and permutations

# Motivation

- ▶ Many problems in probability involve counting.
  - ▶ Suppose I flip a fair coin 100 times. What's the probability I see 34 heads?
  - ▶ Suppose I draw 3 cards from a 52 card deck. What's the probability they all are all from the same suit?
- ▶ In order to solve such problems, we first need to learn how to count.
- ▶ The area of math that deals with counting is called **combinatorics**.

## Selecting elements (i.e. sampling)

- ▶ Many experiments involve choosing  $k$  elements randomly from a group of  $n$  possible elements. This group is called a **population**.
  - ▶ If drawing cards from a deck, the population is the deck of all cards.
  - ▶ If selecting people from DSC 40A, the population is everyone in DSC 40A.
- ▶ Two decisions:
  - ▶ Do we select elements with or without **replacement**?
  - ▶ Does the **order** in which things are selected matter?

## Sequences

- ▶ A **sequence** of length  $k$  is obtained by selecting  $k$  elements from a group of  $n$  possible elements **with replacement**, such that **order matters**.
- ▶ **Example:** Draw a card (from a standard 52-card deck), put it back in the deck, and repeat 4 times.
- ▶ **Example:** A UCSD PID starts with “A” then has 8 digits. How many UCSD PIDs are possible?

# Sequences

In general, the number of ways to select  $k$  elements from a group of  $n$  possible elements such that **repetition is allowed** and **order matters** is  $n^k$ .

(Note: We mentioned this fact in the lecture on clustering!)

# Permutations

- ▶ A **permutation** is obtained by selecting  $k$  elements from a group of  $n$  possible elements **without replacement**, such that **order matters**.
- ▶ **Example:** Draw 4 cards (without replacement) from a standard 52-card deck.
- ▶ **Example:** How many ways are there to select a president, vice president, and secretary from a group of 8 people?

# Permutations

- ▶ In general, the number of ways to select  $k$  elements from a group of  $n$  possible elements such that **repetition is not allowed** and **order matters** is

$$P(n, k) = (n)(n - 1)\dots(n - k + 1)$$

- ▶ To simplify: recall that the definition of  $n!$  is

$$n! = (n)(n - 1)\dots(2)(1)$$

- ▶ Given this, we can write

$$P(n, k) = \frac{n!}{(n - k)!}$$



## Discussion Question

UCSD has 7 colleges. How many ways can I rank my top 3 choices?

- a) 21
- b) 210
- c) 343
- d) 2187
- e) None of the above.



**Summary, next time**

## Summary

- ▶ The **conditional probability** of  $B$  given  $A$  is

$$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

- ▶ A **sequence** is obtained by selecting  $k$  elements from a group of  $n$  possible elements with replacement, such that order matters.
  - ▶ Number of sequences:  $n^k$ .
- ▶ A **permutation** is obtained by selecting  $k$  elements from a group of  $n$  possible elements without replacement, such that order matters.
  - ▶ Number of permutations:  $P(n, k) = \frac{n!}{(n-k)!}$ .
- ▶ **Next time: combinations**, where order doesn't matter.