

Lecture 16 - Conditional Probability, Sequences and Permutations



DSC 40A, Spring 2023

Announcements

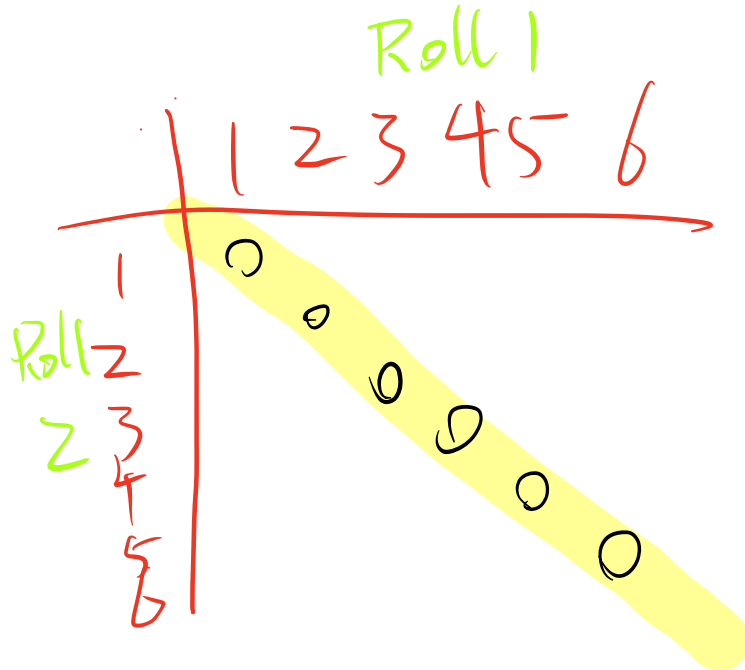
- ▶ HW4 due tonight
 - ▶ HW4 Question 6 is **mandatory**
 - ▶ Look at Campuswire pinned post for more hints.
- ▶ Homework 5 is released, due **Next Wednesday at 11:59pm.**
- ▶ **Important:** We've posted **many** probability resources on the [resources tab of the course website](#). These will no doubt come in handy.
 - ▶ No more DSC 40A-specific readings, though the Probability Roadmap was written specifically for students of this course.

Agenda

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

Example: rolling a die

- ▶ Suppose we roll the die twice. What is the probability that the two rolls have different faces?



Total possible outcome:

$$6 \times 6 = 36$$

Two roll same

6.

$$\text{Two roll different: } 1 - \frac{6}{36} = \frac{5}{6}$$

Conditional probability



Last time

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

① Not

- ▶ For any two events A and B → ② OR

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

If A and B are mutually exclusive, this simplifies to

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

→ This can be zero Empty

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

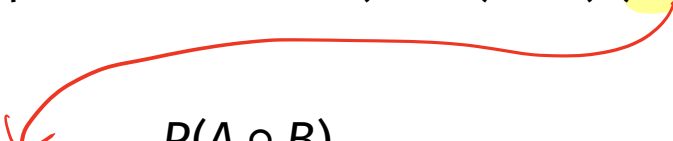
③ AND

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

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

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

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


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**.

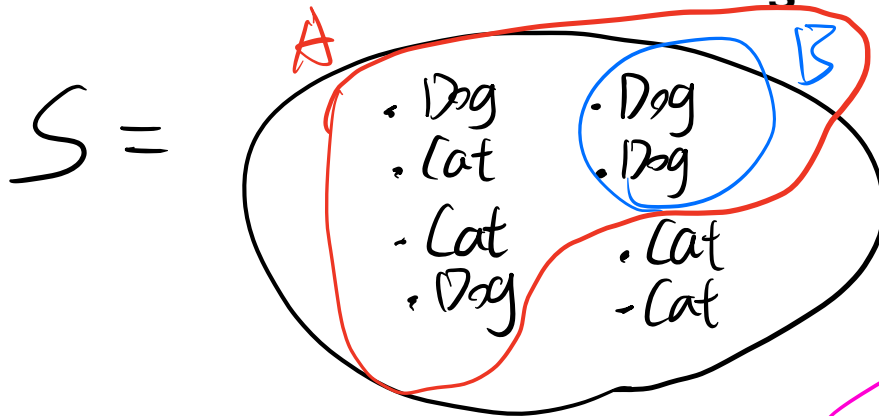
$S =$

• older
• younger

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

Example: pets

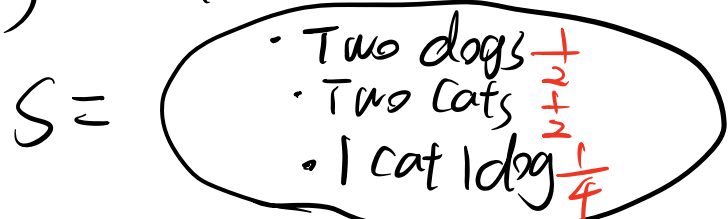
Let's now compute the probability that both pets are dogs given that **at least one of them is a dog**.



A include B

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(B)}{P(A)} = \frac{1/4}{3/4} = \frac{1}{3}$$

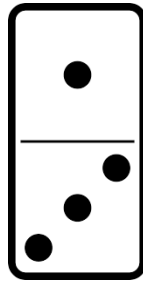
Peel off order info:



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	0	△	△	△△	△	△	
1	x	0	△	△△	△	△	
2	x	x	0	△	△△	△	
3	x	x	x	0	△△	△	
4	x	x	x	x	0	△△	
5	x	x	x	x	x	0	△
6	x	x	x	x	x	x	0



These two are the same

Total # of tiles:

$$7 \times 7 = 49$$

$$49 - 7 \text{ (diag)} = 42$$

$$42 / 2 = 21$$

$$21 + 7 = 28$$

Example: dominoes (source: 538)

diagonals

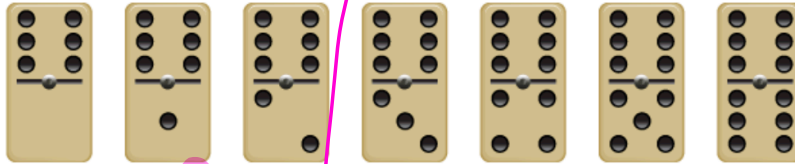


Question 1: What is the probability of drawing a “double” from a set of dominoes — that is, a tile with the same number on both sides?

$$P(\text{Double} | AD) = \frac{7}{28} = \frac{1}{4}$$

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	0	△	△	△	△	△	△
1	x	0	△	△	△	△	△
2	x	x	0	△	△	△	△
3	x	x	x	0	△	△	△
4	x	x	x	x	0	△	△
5	x	x	x	x	x	0	△

Column of 6: 7 entries

double: 1

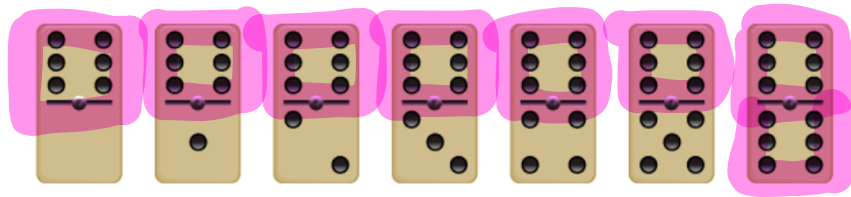
$$\text{Probability} = \frac{1}{7}$$

6 | + + + + +

Example: dominoes (source: 538)

→ We do not know which side is uncovered

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?



← 6 on top and on bottom are two diff. elements

7 dominoes with 8 samples

a double contains 2 samples

$$\text{Probability: } \frac{2}{8} = \frac{1}{4}$$

See 538's explanation here.

uncover 6 \Rightarrow 8 samples
 uncover 5 \Rightarrow 8 samples
 uncover 4 \Rightarrow 8 samples
 ...

$8 \times 7 = 56$ elements
 if we uncover
 one sample.

Simpson's Paradox

$$\begin{aligned}
 P(66 | \text{uncover } 6) &= \frac{P(66 \cap \text{uncover } 6)}{P(\text{uncover } 6)} \\
 &= \frac{P(66)}{P(\text{uncover } 6)} = \frac{1/28}{8/56} = \frac{2/56}{8/56} = \frac{2}{8} \\
 &= \frac{1}{4}
 \end{aligned}$$

\nearrow
uncover
6
 \uparrow
uncover
any #

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

*different
difficulty*

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.

Special case of permutations

- ▶ Suppose we have n people. The total number of ways I can rearrange these n people in a line is

- ▶ This is consistent with the formula

$$P(n, n) = \frac{n!}{(n - n)!} = \frac{n!}{0!} = \frac{n!}{1} = n!$$

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.