Lecture 19 – Conditional Probability, Combinatorics

DSC 40A, Fall 2022 @ UC San Diego

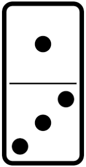
Mahdi Soleymani, with help from **many others**

Agenda

- Finish conditional probability examples.
- Sequences, permutations, and combinations.
- Practice problems.

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.

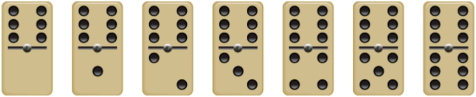


Example: dominoes (source: 538)

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?

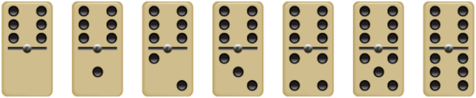
Example: dominoes (source: 538)

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?



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?



Simpson's Paradox (source: 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.

To answer, go to `menti.com` **and enter 4771 9448.**

Simpson's Paradox (source: 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%)

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, permutations, and combinations

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 first 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:** 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

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

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

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

- 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

To answer, go to `menti.com` and enter 4771 9448.

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

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

Combinations

- A **combination** is a set of k items selected from a group of n possible elements **without replacement**, such that **order does not matter**.

- **Example:** There are 24 ice cream flavors. How many ways can you pick two flavors?

From permutations to combinations

- There is a close connection between:
  - the number of permutations of k elements selected from a group of n, and
  - the number of combinations of k elements selected from a group of n

$$\#\text{ combinations} = \frac{\#\text{ permutations}}{\#\text{ orderings of } k \text{ items}}$$

- Since # permutations = $\frac{n!}{(n-k)!}$ and # orderings of k items = $k!$, we have

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

Combinations

In general, the number of ways to select k elements from a group of n elements such that **repetition is not allowed** and **order does not matter** is

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

The symbol $\binom{n}{k}$ is pronounced "n choose k", and is also known as the **binomial coefficient**.

Example: committees

- How many ways are there to select a president, vice president, and secretary from a group of 8 people?

- How many ways are there to select a committee of 3 people from a group of 8 people?

- If you're ever confused about the difference between permutations and combinations, **come back to this example**.

Discussion Question

A domino consists of two faces, each with anywhere between 0 and 6 dots. A set of dominoes consists of every possible combination of dots on each face.

How many dominoes are in the set of dominoes?

A) $\binom{7}{2}$

B) $\binom{7}{1} + \binom{7}{2}$

C) $P(7, 2)$

D) $\frac{P(7,2)}{P(7,1)} 7!$

To answer, go to `menti.com` and enter 4771 9448.

More examples

Counting and probability

- If S is a sample space consisting of equally-likely outcomes, and A is an event, then $P(A) = \frac{|A|}{|S|}$.
- In many examples, this will boil down to using permutations and/or combinations to count $|A|$ and $|S|$.
- **Tip:** Before starting a probability problem, always think about what the sample space S is!

Selecting students — overview

We're going to start by answering the same question using several different techniques.

Question 1: There are 20 students in a class. Billy is one of them. Suppose we select 5 students in the class uniformly at random **without replacement**. What is the probability that Billy is among the 5 selected students?

Selecting students (Method 1: using permutations)

Question 1: There are 20 students in a class. Billy is one of them. Suppose we select 5 students in the class uniformly at random **without replacement**. What is the probability that Billy is among the 5 selected students?

Selecting students (Method 2: using permutations and the complement)

Question 1: There are 20 students in a class. Billy is one of them. Suppose we select 5 students in the class uniformly at random **without replacement**. What is the probability that Billy is among the 5 selected students?

Selecting students (Method 3: using combinations)

Question 1: There are 20 students in a class. Billy is one of them. Suppose we select 5 students in the class uniformly at random **without replacement**. What is the probability that Billy is among the 5 selected students?

Selecting students (Method 3: using combinations)

Question 1, Part 1 (Denominator): If you draw a sample of size 5 at random without replacement from a population of size 20, how many different **sets** of individuals could you draw?

Selecting students (Method 3: using combinations)

Question 1, Part 2 (Numerator): If you draw a sample of size 5 at random without replacement from a population of size 20, how many different **sets** of individuals include Billy?

Selecting students (Method 3: using combinations)

Question 1: There are 20 students in a class. Billy is one of them. Suppose we select 5 students in the class uniformly at random **without replacement**. What is the probability that Billy is among the 5 selected students?

Selecting students (Method 4: "the easy way")

Question 1: There are 20 students in a class. Billy is one of them. Suppose we select 5 students in the class uniformly at random **without replacement**. What is the probability that Billy is among the 5 selected students?

With vs. without replacement

Discussion Question

We've determined that a probability that a random sample of 5 students from a class of 20 **without replacement** contains Billy (one student in particular) is $\frac{1}{4}$.

Suppose we instead sampled **with replacement**. Would the resulting probability be equal to, greater than, or less than $\frac{1}{4}$?

A) Equal to
B) Greater than
C) Less than

To answer, go to `menti.com` **and enter 3779 0977.**

Summary

Summary

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

- A **combination** is obtained by selecting k elements from a group of n possible elements without replacement, such that order does not matter.
  - Number of combinations: $\binom{n}{k} = \frac{n!}{(n-k)!k!}$.