Lecture 19 – Conditional Probability, Combinatorics

DSC 40A, Fall 2022 @ UC San Diego

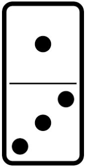
Mahdi Soleymani, with help from **many others**

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.



= 28

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?

$S = \{ \text{All possible Dominos} \}$

$|S| = 28$

$A = \{ 00, 11, 22, \ldots, 66 \}$

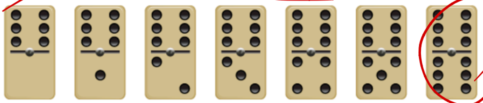
$|A| = 7$

$P(A) = \frac{7}{28} = \frac{1}{4}$

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?

7



1

$\frac{1}{7}$

$S = \{\text{All possible tiles}\}$ $\quad |S| = 28$

$B = \{06, 16, 26, \ldots, 66\}$

$B = \{\text{tiles with at least one } 6\}$ $\quad |B| = 7$

$A = \{66\}$

$A \cap B = A$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{28}}{\frac{7}{28}}$$

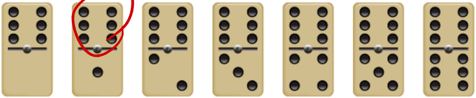
$$= \frac{1}{7}$$

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?

B

$T_1, \ldots, T_{28}$



$\frac{1}{4}$

$A = \{(T_7, S_1), (T_7, S_2)\}$

All half tiles

$S = \{ (T_1, S_1) \cdots \}$

$|S| = 2 \times 28 = 56$

$B = \{ (T_1, S_1), (T_2, S_1), (T_3, S_1) \cdots, (T_7, S_1), (T_7, S_2) \}$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{56}}{\frac{8}{56}} = \frac{1}{4}$$

$|A \cap B| \leftarrow 2$, $|S| \leftarrow 56$, $|B| \leftarrow 8$, $|S| \rightarrow 56$

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? $\cancel{\frac{34}{100}}$
  - 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

Coin 5 times TTHHT ≠ HHTTT

- 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.

$\underline{52}\ \underline{52}\ \underline{52}\ \underline{52} = 52^4$ # possibilites

- **Example:** A UCSD PID starts with "A" then has 8 digits. How many UCSD PIDs are possible?

A $\underline{10}\ \underline{10}\ \underline{10}\ \underline{10}\ \underline{10}\ \underline{10}\ \underline{10}\ \underline{10} = 10^8$

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.

n^k

→ order matters

→ with replacement

(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?

A B C D E F G H

$$\underset{P}{\underline{8}}\ \underset{VP}{\underline{7}}\ \underset{S}{\underline{6}} = \frac{8\cdot 7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}{5\cdot 4\cdot 3\cdot 2\cdot 1}$$

$$= \frac{8!}{5!}$$

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

$$P(n, k) = (n)(n-1)...(n-k+1)$$

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

$$n! = (n)(n-1)...(2)(1)$$

- 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) $\boxed{210}$

C) 343

D) 2187

E) None of the above

~~E E E~~

E
S
M

$P(n,k) = P(7,3)$

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

$P_{(7,3)} = \frac{7!}{(7-3)!} = \frac{7\cdot 6\cdot 5\ \cancel{4!}}{\cancel{4!}} = 210$

Special case of permutations

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

$\underline{n}\ \underline{n-1}\ \underline{n-2}\ \ldots\ \underline{1} = n!$

- This is consistent with the formula

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

$k = n$

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?

Strawberry & Vanila

Vanila & strawberry

$\underline{24}$ $\underline{23}$ $= 24 \times 23$

double count

$\# \text{ ways} = \dfrac{24 \times 23}{2} =$

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

3 flavors out of 24

C, S, V

C V S
V S C
⋮

6 $=3!$

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

- 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?

$\underline{8}\ \underline{7}\ \underline{6}$

$P(8,3)$

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

$\binom{n}{k} = \binom{8}{3}$

- 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?

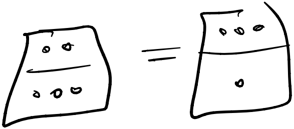
A) $\binom{7}{2}$

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

C) $P(7,2)$

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

$$\frac{7 \times 6}{2} = 21 \quad \times$$



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

$$\binom{7}{2} + \binom{7}{1} = 21 + 7 = 28$$

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?

$S = \{\text{Permutations of 5 students}\}$

$|S| = P(20, 5)$

$A = \{\text{Perms. of students including Billy}\}$

$$P(A) = \frac{|A|}{|S|} = \frac{}{P(20, 5)}$$

$\Rightarrow$ $\underline{B}$ $\underline{19}$ $\underline{18}$ $\underline{17}$ $\underline{16}$ $= P(19,4)$

$\underline{19}$ $\underline{B}$ $\underline{18}$ $\underline{17}$ $\underline{16}$

__ __ $\underline{B}$ __ __

__ __ __ $\underline{B}$ __

__ __ __ __ $\underline{B}$

$$|A| = 5 \times P(19,4)$$

$$\frac{|A|}{|S|} = \frac{5 \times P(19,4)}{P(20,5)} = \frac{5 \cdot \cancel{19 \cdot 18 \cdots}}{20 \cdot \cancel{19 \cdots}} = 1/4$$

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

$P(A) = 1 - P(\bar{A})$

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?

$S = \{ \text{All perm. of } 5 \} \Rightarrow |S| = P(20, 5)$

$A = \{ \ldots \}$

$$P(A) = 1 - P(\bar{A}) = 1 - \frac{\#\text{ Perms. of } 5 \text{ students out of } 19}{|S|}$$

$$= 1 - \frac{P(19,5)}{P(20,5)} = 1 - \frac{\cancel{19 \cdot 18 \cdot 17 \cdot 16} \cdot 15}{20 \cdot \cancel{19 \cdots 16}} = 1 - \frac{15}{20}$$

$$= 1 - 3/4 = 1/4$$

Selecting students (Method 3: using combinations)

$\{A, B, C, D, E\} = \{B, A, D, C, E\}$

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?

$S = \{\text{Combinations of 5 students out of 20}\}$

$|S| = \binom{20}{5}$ $\quad |W| = \binom{19}{4}$

$W = \{\text{All combinations Billy included}\}$

$$P(W) = \frac{|W|}{|S|} = \frac{\binom{19}{4}}{\binom{20}{5}} = \frac{\frac{19!}{4!\,15!}}{\frac{20!}{5!\,15!}} = \frac{\cancel{19!}}{\cancel{15!}\,\cancel{4!}} \cdot \frac{\cancel{15!}\,\cancel{5!}}{\cancel{20!}} = \frac{5}{20} = \frac{1}{4}$$

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?

Combinations of 5 out of 20

$$\binom{5}{20}$$

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?

Billy is already picked.

4 other choices left

combinations of 4 out of $20-1=19$

$\binom{19}{4}$

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?

$$P(W) = \frac{\binom{19}{4}}{\binom{20}{5}} = \frac{5}{20} = \frac{1}{4}$$

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?

1. Give all 20 students a number
2. shuffle

$$\underbrace{\square\ \square\ \square\ \cdots}_{\text{First 5 in the line}}\ \cdots\ \square\ \square\ \square$$

$$\frac{5}{20} = \frac{1}{4}$$

With vs. without replacement

5 times

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

47719448

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

$$P(\text{Billy}) = 1 - P(\text{No Billy})$$

$$= 1 - \left(\frac{19}{20}\right)\left(\frac{19}{20}\right)\left(\frac{19}{20}\right)\left(\frac{19}{20}\right)\left(\frac{19}{20}\right)$$

$$= 1 - \left(\frac{19}{20}\right)^5 \approx 0.226 < \frac{1}{4} = 0.25$$

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