DSC 40A

Theoretical Foundations of Data Science I

Lecture 20-21: Combinatorics

Agenda

- How do we count the number of outcomes, besides enumerating them all?
  - How many outcomes are possible if a die is rolled 100 times?
  - How many different ways are there to shuffle 52 cards?
  - How many ways are there to choose a jury of 12 people from a panel of 100?

- Many probability questions can be solved by counting, or combinatorics.

- We'll learn how to count sequences and sets.

Question

Answer at q.dsc40a.com

Remember, you can always ask questions at q.dsc40a.com!
If the direct link doesn't work, click the "Lecture Questions" link in the top right corner of dsc40a.com.

Combinatorics

Sequences vs. Sets

Sequences list, tuple	Sets collection of elements
Order matters	Order does not matter
Repetitions allowed (with replacement)	No repetitions allowed (without replacement)
Elements listed in order	Elements listed in no particular order within curly braces
Ex: $2, 4, 5 \neq 4, 2, 5$	Ex: $\{2, 4, 5\} = \{4, 2, 5\}$
Ex: $2, 2, 2 \neq 2, 2$	Ex: $\{2, 2, 2\} = \{2, 2\} = \{2\}$
Ex: $1, 3, 4 = 1, 3, 4$	Ex: $\{1, 3, 4\} = \{1, 3, 4\}$

Sequences

Sequences
Order matters
Repetitions allowed
Elements listed in order
Ex: 2, 4, 5 ≠ 4, 2, 5
Ex: 2, 2, 2 ≠ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

Example 1: sampling w/ replacement

draw a card, put it back, repeat four more times

$$(A\heartsuit, 2\clubsuit, 6\spadesuit, A\heartsuit, 3\diamondsuit)$$

$$\neq (2\clubsuit, 6\spadesuit, A\heartsuit, 3\diamondsuit, A\heartsuit)$$

Example 2:

flip a coin 100 times

$$(H, T, T, H, \ldots, H, T, T, T)$$

Sequences

Sequences
Order matters
Repetitions allowed
Elements listed in order
Ex: 2, 4, 5 $\neq$ 4, 2, 5
Ex: 2, 2, 2 $\neq$ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

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

5% A. 8^{10}

75% B. 10^8

20% C. 8!

5% D.

$$\underbrace{10 \cdot 10 \cdot 10 \cdot \ldots \cdot 10}_{8 \text{ times}} = 10^8$$

A _ _ _ _ _ _ _ _

10 options {0, 1, …, 9}

Sequences

Sequences
Order matters
Repetitions allowed
Elements listed in order
Ex: 2, 4, 5 $\neq$ 4, 2, 5
Ex: 2, 2, 2 $\neq$ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

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

A. 8^{10}

B. 10^{8}

C. 8!

D.

P is the population you can draw from and $n =$ |P| is the size of that population (number of elements). with replacement

How many sequences of length *k* are there?

$$\underbrace{|P| \cdot |P| \cdots |P|}_{k \text{ times}} = |P|^k = n^k$$

Sequences

Sequences
Order matters
Repetitions allowed
Elements listed in order
Ex: 2, 4, 5 ≠ 4, 2, 5
Ex: 2, 2, 2 ≠ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

<u>Exponential growth</u>

Flip a coin *n* times

n	# of Sequences of Length *n*
5	$2^5 = 32$
10	$2^{10} = 1024$
15	$2^{15} = 32{,}758$
20	$2^{15} \approx 1$ million
50	$2^{50} \approx$ # of grains of sand on Earth

Sequences

Sequences
Order matters
Repetitions allowed
Elements listed in order
Ex: 2, 4, 5 $\neq$ 4, 2, 5
Ex: 2, 2, 2 $\neq$ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

How many ways to select a president, vice president, and secretary from a group of 8 people? Nobody can serve in more than one role. $\Rightarrow$ sampling without replacement

$\underset{p}{\underline{5}}$ $\underset{vp}{\underline{2}}$ $\underset{s}{\underline{7}}$

president 8 options
vice president: 7 options
secretary : 6 options

$8 \cdot 7 \cdot 6$

$n = 8$
$k = 3$

$\underset{p}{\underline{5}}\ \underline{2}\ \underset{s}{\underline{7}} \neq \underset{p}{\underline{7}}\ \underline{2}\ \underset{s}{\underline{5}}$

Sequences

Sequences
Order matters
Repetitions allowed
Elements listed in order
Ex: 2, 4, 5 $\neq$ 4, 2, 5
Ex: 2, 2, 2 $\neq$ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

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

$n = 8$ (# elements to choose from)

$k = 3$ (# distinct elements to choose)

$P(8,3) = 8 \cdot 7 \cdot 6$

$P(n,k) = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k-1) = \frac{n!}{(n-k)!}$

Sequences where repetitions are not allowed are Permutations.

Sets

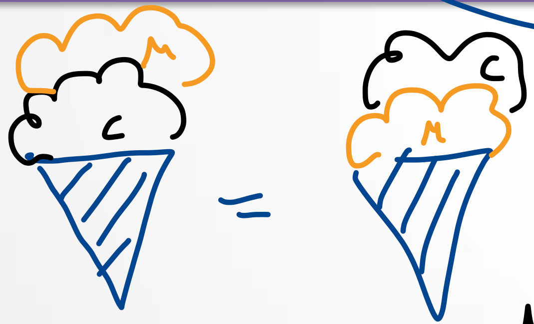
There are 24 ice cream flavors. How many ways can you pick 2 different flavors?

A. 24

$^2/_3$ B. 24*23

C. 24*24

$^1/_3$ D. 12*23



$\Rightarrow$ set

Use sequences

$CM \neq MC$

$\#\text{sequences} = 24 \cdot 23$

$$\#\text{sets} = \frac{\#\text{sequences}}{\#\text{orderings}} = \frac{24 \cdot 23}{2} = 12 \cdot 23 \Rightarrow \#\text{seqs} = \#\text{sets} \cdot \#\text{orderings}$$

Sets
Order does not matter
No repetitions allowed
Elements listed in no particular order within curly braces
Ex: $\{2, 4, 5\} = \{4, 2, 5\}$
Ex: $\{2, 2, 2\} = \{2, 2\} = \{2\}$
Ex: $\{1, 3, 4\} = \{1, 3, 4\}$

Sets

How many ways to select a committee of 3 from a group of 8?

$$\# \text{sets} = \frac{\# \text{sequences}}{\# \text{orderings}} = \frac{8 \cdot 7 \cdot \cancel{6}}{\cancel{3 \cdot 2} \cdot 1} = 8 \cdot 7 = 56$$

$n = 8$ # elements to choose from

$k = 3$ # elements to select

$$C(8,3) = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = \frac{8!/5!}{3!} = \frac{8!}{5!\,3!}$$

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

Sets
Order does not matter
No repetitions allowed
Elements listed in no particular order within curly braces
Ex: {2, 4, 5} = {4, 2, 5}
Ex: {2, 2, 2} = {2, 2} = {2}
Ex: {1, 3, 4} = {1, 3, 4}

C = combinations

Permutations vs. Combinations

Permutations	Combinations
Order matters	Order does not matter
No repetitions allowed (without replacement)	No repetitions allowed (without replacement)
Counts the number of **sequences of k distinct elements** chosen from n possible elements	Counts the number of **sets of size k** chosen from n possible elements
$P(n,k) = (n)(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$	"n choose k" $C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
How many ways to select a president, vice president, and secretary from a group of 8 people? P(8,3)	How many ways to select a committee of 3 from a group of 8? C(8,3)

Permutations vs. Combinations

Permutations
Order matters
No repetitions allowed (without replacement)
Counts the number of **sequences of k distinct elements** chosen from n possible elements
$P(n,k) = (n)(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$
How many ways to select a president, vice president, and secretary from a group of 8 people? P(8,3)

Example 1:

draw a card, **don't** put it back, repeat four more times

$(A\heartsuit, 2\clubsuit, 6\spadesuit, 7\heartsuit, 3\diamondsuit)$

Example 2:

rank 2 best cities to live in out of list of 10

SD, LA

Permutations vs. Combinations

Combinations
Order does not matter
No repetitions allowed (without replacement)
Counts the number of **sets of size k** chosen from n possible elements
$C(n,k) = \binom{n}{k} = \frac{n!}{k!(n-k)!}$
How many ways to select a committee of 3 from a group of 8? C(8,3)

Example 1:

draw a hand of 5 cards from a deck of 52

{A♥, 2♣, 6♠, 7♥, 3♦}

Example 2:

Select 5 student from the class

Owen Luis

Leah Nayan

Dillan

Sampling Without Replacement

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students?

Monday was sequences

Today using sets!

S = sets of length 5 chosen from 20 students

selected student = 17

$$\text{Prob}\left(\begin{array}{l}\#17 \text{ is included} \\ \text{in set of 5} \\ \text{selected students}\end{array}\right) = \frac{\#\text{ sets including } 17}{\#\text{ sets of 5 out of 20}}$$

Sampling Without Replacement

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?

\# sets of 5 out of 20

$$C(20,5) = \binom{20}{5}$$

Sampling Without Replacement

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 a particular person?

\# sets with 5 students including 17

Key! choosing 4 students out of 19

$$= C(19,4) = \binom{19}{4}$$

Sampling Without Replacement

Using the complement. If you draw a sample of size 5 at random without replacement from a population of size 20, how many different **sets** of individuals **do not** include a particular person?

sets of 5 students $\underbrace{\text{out of 20 not including 17}}_{\text{chosen from 19 students}}$

$= C(19, 5)$

from Theory Meets Data by Ani Adhikari, Chapter 4

Sampling Without Replacement

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students? $P(A) = 1 - P(\bar{A})$

$$P\left(\begin{matrix}\text{set of 5 students}\\ \text{includes 17}\end{matrix}\right) = \frac{\#\text{sets incl. 17}}{\#\text{sets}}$$

$$= \frac{C(19,4)}{C(20,5)}$$

$$= \frac{\frac{19!}{4!\,\cancel{15!}}}{\frac{20!}{5!\,\cancel{15!}}} = \frac{19!}{20!} \cdot \frac{5!}{4!} = \frac{1}{20} \cdot \frac{5}{1} = \frac{1}{4}$$

using complement

$$\frac{\#\text{sets} - \#\text{sets not incl. 17}}{\#\text{sets}}$$

$$= \frac{C(20,5) - C(19,5)}{C(20,5)}$$

from Theory Meets Data by Ani Adhikari, Chapter 4

Summary

- To calculate *P(E)* we need to find |*E*|
- We need to count sequences or sets
- Must decide if order matters
- When elements are distinct: permutations vs. combinations

$$P(n,k) = (n)(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$$

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

- **Next time:** more examples