DSC 40A

Theoretical Foundations of Data Science I

Lecture 20-21: Combinatorics

Announcements

- Homework 5 due tonight
- Upcoming homework schedule:
  Homework 6 released Monday 11/18 and due 11/25

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

Today

- 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.

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 $\neq$ 4, 2, 5
Ex: 2, 2, 2 $\neq$ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

Example 1:
draw a card, put it back, repeat four more times

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

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

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 ≠ 4, 2, 5
Ex: 2, 2, 2 ≠ 2, 2
Ex: 1, 3, 4 = 1, 3, 4

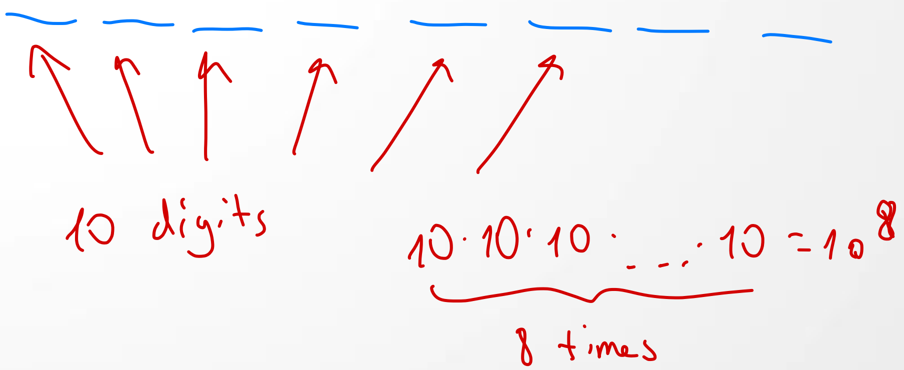
A UCSD PID starts with “A” then has 8 digits. How many UCSD PIDs are possible?

A. 8^{10}

B. 10^8

C. 8!

D.



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

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 |P| is the size of that population (number of elements). with replacement

How many sequences of length *k* are there?

$n = |P|$

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

Sequences

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

$n = 8$

$k = 3$

president : 8 options
vice-president : 7 options
secretary : 6 options
} without replacement

$8 \cdot 7 \cdot 6$

$\underset{p}{5}\ \underset{vp}{3}\ \underset{s}{2} \neq \underset{p}{2}\ \underset{vp}{3}\ \underset{s}{5}$

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

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 from)

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

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

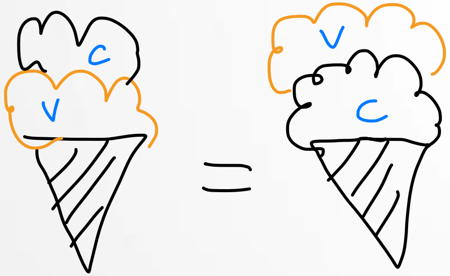
Sequences where repetitions are not allowed are Permutations.

Sets

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

A. 24
B. 24*23
C. 24*24
D. 12*23

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: $\{1, 3, 4\} = \{1, 3, 4\}$



$\Rightarrow$ set

use sequences

$CV \neq VC$

sequence: $24 \cdot 23$

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

$\Rightarrow \#\text{seq} = \#\text{sets} \cdot \#\text{orderings}$

Sets

$\{5, 3, 2\} = \{2, 3, 5\}$

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

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

$n = 8$ # 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
$P(n,k) = (n)(n-1)\ldots(n-k+1) = \frac{n!}{(n-k)!}$	"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
$P(n,k) = (n)(n-1)\dots(n-k+1) = \frac{n!}{(n-k)!}$
How many ways to select a president, vice president, and secretary from a group of 8 people? P(8,3)

without replacement

Example 1:
draw a card, **don't** put it back, repeat four more times

(A♥, 2♣, 6♠, 7♥, 3♦)

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

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?

Today: using sets

S = sets of students chosen from 20

selected student : #17

$$\text{Prob}\left(\begin{array}{c}\#17 \text{ is in our set of} \\ 5 \text{ selected students}\end{array}\right) = \frac{\#\text{ sets including } 17}{\#\text{ sets in } S \text{ chosen from } 20}$$

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?

# of sets of $k=5$ chosen from $n=20$

$C(20,5)$

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?

# of sets of length 5 chosen from 20 including person 17

key: the same as choosing 4 out of 19

$= C(19, 4)$

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 length 5 chosen from 20 not including 17

chosen from 19 students

$= C(19, 5)$

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?

$$P\left(\begin{array}{c}\text{set including 17} \\ \text{length 5 set}\end{array}\right) = \frac{\#\text{ sets with 17}}{\#\text{ sets}} \qquad \frac{\text{total \# sets} - \#\text{ sets without 17}}{\text{total \# sets}}$$

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

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

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

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

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

- **Next time:** more examples

DSC 40A

Theoretical Foundations of Data Science I

Today

- More examples of using combinatorics to solve probability questions.

Counting as a Tool for Probability

Example 7. What is the probability that a randomly generated bitstring of length 10 contains an equal number of zeros and ones?

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

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

Counting as a Tool for Probability

Example 8. What is the probability that a randomly generated bitstring of length 10 is the string 0011001101?

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

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

Counting as a Tool for Probability

Example 9. What is the probability that a fair coin flipped 10 times turns up an equal number of heads and tails?

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

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

Counting as a Tool for Probability

Example 10. What is the probability that a fair coin flipped 10 times turns up HHTTHHTTHT?

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

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

Counting as a Tool for Probability

Example 11. What is the probability that a biased coin with $Prob(H) = \frac{1}{3}$ flipped 10 times turns up an equal number of heads and tails?

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

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

Counting as a Tool for Probability

Example 12. What is the probability that a biased coin with $Prob(H) = \frac{1}{3}$ flipped 10 times turns up HHTTHHTTHT?

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

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

The Easy Way

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?

The Easy Way

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?

Another way to think of sampling without replacement:

1. randomly shuffle all 20 students
2. take the first 5

from Theory Meets Data by Ani Adhikari, Chapter 4

Practice Problems

Example 13. You were one of 238 individuals who reported for jury duty. If 54 of these people will be assigned to a courtroom, what is the probability that you get assigned to a courtroom?

Practice Problems

Example 13. You were one of 238 individuals who reported for jury duty. If 54 of these people will be assigned to a courtroom, what is the probability that you get assigned to a courtroom?

How many sets of 54 individuals include you?

A. C(238, 54)

B. C(237, 54)

C. C(238, 53)

D. C(237, 53)

Practice Problems

Example 14. You were one of 238 individuals who reported for jury duty. Suppose 28 of these individuals are doctors. If 54 of these people will be assigned to a courtroom, what is the probability that exactly 5 doctors get assigned to a courtroom?

Practice Problems

Example 15. What is the probability that your five-card poker hand is a straight?

Practice Problems

Example 16. Suppose you look at your first card as it is dealt, and you see that it is a Queen. What is the probability that your five-card hand is a straight?

Summary

- Counting is a useful tool for probability.
- Sometimes there's an easy way!
- **Next time:** Bayes Theorem