CS 109: Introduction to Probability for Computer Scientists

Will Monroe
Summer 2017

with materials by
Mehran Sahami and Chris Piech

June 26, 2017
Who am I?

Stanford undergrad
Who am I?

Graphics
Who am I?

Monroe et al. (2014)

Li et al. (2016)

Natural language processing

I'm going to the police station.

No, no, no, no, you're not going anywhere.

I need you to stay here.

I don’t know what you are talking about.

Where are you going?

I’ll come with you.

Why?
Who am I?

Call me Will! 😊

or Mr. Monroe 👍

Professor Monroe

Dr. Monroe

Master Monroe?
Who am I?

Reno, Nevada
Traditional probability setups
Why probability with computers?
Artificial Intelligence
Biocomputation
Theory
Computer Engineering
Systems
Information
HCI
Graphics
Randomized algorithms

Learning theory

Gene expression prediction

Monte Carlo rendering

Hypothesis testing

Language modeling

Caching, error correction

Load balancing, queueing

Monte Carlo rendering
Quiz question

https://b.socrative.com/login/student/

Room: CS109SUMMER17
Zika testing

0.08% of people have Zika
90% of people with Zika test positive
7% of people without Zika test positive

Someone tests positive. What's the probability they have Zika?

A) <0.5%  B) 1%  C) 7%  D) 41%  E) 90%
Zika testing

0.08% of people have Zika
90% of people with Zika test positive
7% of people without Zika test positive

Someone tests positive. What's the probability they have Zika?

Answer: B) 1%
Counting Combinatorial analysis
An experiment
General principle of counting

An experiment consisting of two or more separate parts has a number of outcomes equal to the product of the number of outcomes of each part.

\[ |A_1 \times A_2 \times \cdots \times A_n| = \prod_{i} |A_i| \]

shapes: 4

colors: 3

total: \(4 \cdot 3 = 12\)
General principle of counting

colors: 3

shapes: 4

total: $4 \cdot 3 = 12$
General principle of counting

\[ \text{total: } |A| \cdot |B| \]
Counting license plates

\[26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 17,576,000\]

\[10 \cdot 26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 175,760,000\]

(Approximate 2017 California population: 39,800,000)

Jerry “Woody”, Happoshu
General principle of counting

An experiment consisting of two or more separate parts has a number of outcomes equal to the product of the number of outcomes of each part.

\[ |A_1 \times A_2 \times \cdots \times A_n| = \prod_{i} |A_i| \]

- shapes: 4
  - total: 4 · 3 = 12
- colors: 3
  - total: 4 · 3 = 12
Sum rule of counting

If two sets don't overlap, the total number of elements in the two sets is the sum of the number of elements of each set.

\[ |A \cup B| = |A| + |B| \quad \text{if} \quad A \cap B = \emptyset \]

3 + 4 = 7
Principle of Inclusion/Exclusion

The **total number of elements** in two sets is the sum of the number of elements of each set, **minus** the number of elements in both sets.

$$|A \cup B| = |A| + |B| - |A \cap B|$$

3 + 4 - 1 = 6
Inclusion/Exclusion intuition
Bit strings

How many 4-bit strings are there that start or end with 1?

A

B

$|A| = 2^3$

$|B| = 2^3$

$|A \cap B| = 2^2$
Bit strings

How many 4-bit strings are there that start or end with 1?

\[ \begin{align*}
2^3 + 2^3 - 2^2 &= 12
\end{align*} \]
Principle of Inclusion/Exclusion

The total number of elements in two sets is the sum of the number of elements of each set, minus the number of elements in both sets.

\[ |A \cup B| = |A| + |B| - |A \cap B| \]

3 + 4 - 1 = 6
Floor and ceiling

\[
\begin{bmatrix}
x \\
\end{bmatrix}
\quad
\begin{bmatrix}
x \\
\end{bmatrix}
\]
Floor and ceiling

$\lfloor x \rfloor$  $\lceil x \rceil$
Floor and ceiling

\[ \lfloor 2.5 \rfloor = 2 \]

\[ \lfloor 1 \rfloor = 1 \]
Floor and ceiling

\[ \lfloor 2.5 \rfloor = 3 \]

\[ \lceil 1 \rceil = 1 \]
If $m$ objects are placed in $n$ buckets, then **at least one bucket** must contain at least $\lceil \frac{m}{n} \rceil$ objects.
Pigeonhole Principle intuition

More objects than buckets = Some objects have to share!

\[ n = 5 \text{ buckets} \]

\[ m = 7 \text{ objects} \]
Pigeonhole Principle intuition

More pigeons than holes = Some pigeons have to share!

\[ m = 7 \text{ pigeons} \]

\[ n = 5 \text{ holes} \]
Pigeonhole Principle intuition

More pigeons than holes = Some pigeons have to share!

\[ n = 5 \text{ holes} \]

\[ m = 7 \text{ pigeons} \]
String hashing

$n = 100$ buckets

$m = 950$ strings

Is there a bucket with at least 10 entries?

A) Yes
B) No
C) Don't know

https://b.socrative.com/login/student/
Room: CS109SUMMER17
String hashing

$m = 950$ strings

$n = 100$ buckets

Is there a bucket with at least 9 entries?

A) Yes
B) No
C) Don't know
String hashing

$n = 100$ buckets

$m = 950$ strings

Is there a bucket with at least 11 entries?

A) Yes
B) No
C) Don't know

https://b.socrative.com/login/student/
Room: CS109SUMMER17
String hashing

\[ n = 100 \text{ buckets} \]

\[ m = 950 \text{ strings} \]

Can there be a bucket that is empty?

A) Yes
B) No
C) Don't know
Break time!
Course Logistics
Course Website

https://cs109.stanford.edu/
Teaching assistants

Michela Meister

Yuling Liu

Brendan Corcoran
Prereqs

**CS 106B/X**
- Programming
- Recursion
- Hash tables
- Binary trees

**CS 103**
- Proofs (induction)
- Set theory

**MATH 51/CME 100**
- Multivariate differentiation
- Multivariate integration
- Basic linear algebra (what's a vector?)
Contacting us

Office hours

Forum

Email

ps109-sum1617-staff@lists.stanford.edu
Units

Undergrads and SCPD students: 5

Grad students: 3, 4, 5

(same amount of work! ~10 hours/week for assignments)
Recorded lectures

https://mvideox.stanford.edu/
Late days

[updated July 14]

free 24-hour “extensions”
Grade components

45%  6 problem sets

20%  Midterm exam
      Tuesday, July 25, 7:00-9:00pm
      [email staff ASAP with conflicts]

30%  Final exam
      Saturday, August 19, 12:15-3:15pm
      [no alternate time!]

5%   Participation: Piazza, feedback
Honor code

You may:

Talk with the course staff
Talk with classmates (cite collaboration)
Look up general material online

You may not:

Copy answers:
  from classmates
  from former students
  from the Internet
Look at previous quarters' solutions
A very brief history of AI

further reading:
Russell & Norvig (2009)
Early AI optimism

1952

1962

Evans (1962)
Early AI optimism

"...in a visible future—the range of problems [AIs] can handle will be coextensive with the range to which the human mind has been applied."

–Herbert Simon, 1957
(quoted in Russell & Norvig, 2003)
“The spirit is willing but the flesh is weak”

“Спирт готов но мясо слабо.”

“The vodka is good but the meat is rotten”
BRACE YOURSELVES

WINTER IS COMING
Expert systems (1980s)

Lauritzen & Spiegelhalter (1988)
The statistical revolution (1990s)

examples → probability
Neural nets return (2000s)
Speech recognition

who is the current president of France?
Image classification

Easiest classes
- red fox (100)
- hen-of-the-woods (100)
- ibex (100)
- goldfinch (100)
- flat-coated retriever (100)

Hardest classes
- muzzle (71)
- hatchet (68)
- water bottle (68)
- velvet (68)
- loupe (66)

- hook (66)
- spotlight (66)
- ladle (65)
- restaurant (64)
- letter opener (59)

right: Russakovsky et al. (2015)
Solving speech recognition?

Teacher: ...includes—including ace.
Student: 40.
Teacher: (writing)—including—ace.
Student: Isn't it 4 out of 40...?
Solving object recognition?

\[ x + 0.007 \times \text{sign}(\nabla_x J(\theta, x, y)) = x + \epsilon\text{sign}(\nabla_x J(\theta, x, y)) \]

\( x \)  

“panda”  

57.7\% confidence

\( x \) + 0.007 \times \text{sign}(\nabla_x J(\theta, x, y)) \)

“nematode”  

8.2\% confidence

\( x + \epsilon\text{sign}(\nabla_x J(\theta, x, y)) \)

“gibbon”  

99.3\% confidence

Goodfellow et al. (2015)
Solving question answering?

Search query: how long has theresa may been pm

Result: 1.72 trillion pm