CS109: Probability for Computer Scientists

Lisa Yan
September 23, 2019
Yes, my undergrad was here...

...But now I’m here!!!

PhD: Tools to understand student learning

My interests over time

 Networks, Data Science

 Create technology

 Help people

 Create technology to help people
Why I like probability

• I like data

• I want to help people

• Probability helps me help people with data

• Also Pokemon

\[ a = \frac{(3 \times HP_{\text{max}} - 2 \times HP_{\text{current}}) \times \text{rate} \times \text{bonus}_{\text{ball}}}{3 \times HP_{\text{max}}} \times \text{bonus}_{\text{status}} \]
Teaching team
What about you?
Today’s plan

Course Mechanics

Why you should take CS109

Counting!
Course mechanics (light version)

• For more info, read the Administrivia handout

• Course website:
  http://cs109.stanford.edu/
Prerequisites

**CS106B/X**
- Programming
- Recursion
- Hash tables
- Binary trees

**MATH 51/CME 100**
- Multivariate differentiation
- Multivariate integration
- Basic facility with linear algebra (vectors)

**CS103**
(co-requisite OK)
- Proofs (induction)
- Set theory
- Math maturity

**Important!**
Staff contact

- Piazza

- Email cs109@cs.stanford.edu

- Working office hours

- Contact Lisa for course level issues, extensions, etc.
How many units should I take?

Hours per week = Units × 3
Average about 10 hours / week for assignments

Start Here

Are you an undergrad? No Yes

Do you want to take CS109 for fewer units? Yes No

Yes

3 Units -or- 4 Units

5 Units
Where you learn

• Lectures (not videotaped)
• Lecture notes (on website)
• Textbook readings (optional)
• Discussion Section
• Problem Sets
Class breakdown

45%  6 Problem Sets

20%  Midterm
Tuesday, October 29th, 7:00–9:00pm

30%  Final
Wednesday, December 11th, 3:30–6:30pm

5%  Participation
   • Weekly concept checks (due Mondays 1pm)
   • Section participation
Problem Sets

Late Days:

2

[class days]
(for Problem Sets only)

Review session this Friday
(time/location TBA)
Stanford Honor Code

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

NOT permitted:
• Copy answers:
  from classmates
  from former students
  from previous quarters
• Copy answers from the internet
  Besides, these are usually incorrect
Questions on logistics?
Today’s plan

Course Mechanics

Why you should take CS109

Counting!
Traditional View of Probability
CS view of probability

http://www.site.com

But wait...
There’s MORE!!
Machine Learning

= Machine (compute power)
+ Probability
+ Data
Machine Learning Algorithm

Data $\xrightarrow{\text{Build a } \textit{probabilistic model}}$ Do one thing
Classification
Where is this useful?

A machine learning algorithm performs **better than** the best dermatologists.

Developed in 2017 at Stanford.

The last remaining board game
Image tagging
Self-driving cars
Augmented Reality Machine Translation

Automatic machine translation on Google Translate
Voice assistants

What can I help you with?

Alexa  Siri  Google Now  Cortana
Probability is *more* than just machine learning.
Probability and medicine
Probability and art
Probability and climate
Probabilistic analysis of algorithms
Probability at your fingertips
Probability and philosophy
Probability for good

How do we identify systemic biases in our data and incorporate human judgment into our probabilistic models?

Algorithms of Oppression, Safiya Umoja Noble. 2018
We’ll get there!
Probability is not always intuitive.
Zika test

A patient takes a Zika test that returns positive. What is the probability that they have the Zika virus?

- 0.8% of people have the virus
- Test has 90% positive rate for people with the virus
- Test has 7% positive rate for people without the virus

Correct answer: 9%
Probability = Important + Needs Studying
Today’s plan

Course Mechanics

Why you should take CS109

Counting!
01: Counting
What is Counting?

An experiment in probability:

Counting: How many possible outcomes can occur from performing this experiment?
What is Counting?

Roll 6 \{1, 2, 3, 4, 5, 6\}

Roll even only 3 \{2, 4, 6\}

Roll 36 \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (4, 6), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5), (5, 6), (6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\}
Sum Rule of Counting

If the outcome of an experiment can be either from Set $A$, where $|A| = m$, 
or Set $B$, where $|B| = n$, 
where $A \cap B = \emptyset$, 
Then the number of outcomes of the experiment is $|A| + |B| = m + n$. 
Video streaming application

Your application has distributed servers in 2 locations.

If a server request is sent to the application, how large is the set of servers it can get routed to?

Goal
Outcome server is in either San Jose or Boston

Define
\( A \) : San Jose
\( B \) : Boston
Note: \( A \cap B = \emptyset \)

Solve
\( |A| + |B| = m + n \)

\( |A| + |B| = 150 \) servers
Product Rule of Counting

If an experiment has two parts, where

The first part’s outcomes are from Set $A$, where $|A| = m$, and
The second part’s outcomes are from Set $B$, where $|B| = n$,

Then the number of outcomes of the experiment is

$|A||B| = mn$. 

Two-step experiment

$\rightarrow A \rightarrow B$
Dice

How many possible outcomes are there from rolling two 6-sided dice?

Goal
Outcome roll contains an outcome from both die 1 and die 2

Define
$A$: Die 1 outcomes
$B$: Die 2 outcomes

Solve
$|A| \times |B| = 36$

36 outcomes
**TOP DEFINITION**

**kick it up a notch**

To make things more intense, exciting, or interesting.

(introduced by chef Emeril Lagasse in reference to spicing up his recipes.)
Inclusion-Exclusion Principle

If the outcome of an experiment can be either from Set $A$ or set $B$, where $A$ and $B$ may overlap,

Then the total number of outcomes of the experiment is $|A \cup B| = |A| + |B| - |A \cap B|$. 

One experiment

A

B only

Sum Rule of Counting:
A special case
Transmitting bytes over a network

An 8-bit string is sent over a network.
• The receiver only accepts strings that either start with 01 or end with 10.

How many 8-bit strings will the receiver accept?

Define

A : 8-bit strings starting with 01
B : 8-bit strings ending with 10

1. What is |A| ?
   A. $2^8$
   B. $2^6$
   C. $2^4$
   D. 0

2. What is |A ∩ B| ?
   A. $2^8$
   B. $2^6$
   C. $2^4$
   D. 0

Inclusion-Exclusion Principle

$|A \cup B| = |A| + |B| - |A \cap B|$
Transmitting bytes over a network

An 8-bit string is sent over a network.
• The receiver only accepts strings that either start with 01 or end with 10.

How many 8-bit strings will the receiver accept?

Define

A : 8-bit strings starting with 01
B : 8-bit strings ending with 10

1. What is |A|?
   A. $2^8$
   B. $2^6$
   C. $2^4$
   D. 0

2. What is |A ∩ B|?
   A. $2^8$
   B. $2^6$
   C. $2^4$
   D. 0

Inclusion-Exclusion Principle

$|A ∪ B| = |A| + |B| - |A ∩ B|$
Transmitting bytes over a network

An 8-bit string is sent over a network.

- The receiver only accepts strings that either start with 01 or end with 10.

How many 8-bit strings will the receiver accept?

Define

- \( A \): 8-bit strings starting with 01
- \( B \): 8-bit strings ending with 10

1. What is \( |A| \)?

   B. \( 2^6 \)

2. What is \( |A \cap B| \)?

   C. \( 2^4 \)

Solve

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

\[
= 2^6 + 2^6 - 2^4 = 112 \text{ outcomes}
\]
General Principle of Counting

If an experiment has $r$ steps, such that

Step $i$ has $n_i$ outcomes for all $i = 1, \ldots, r$,

Then the number of outcomes of the experiment is

$$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i.$$
License plates

How many CA license plates are possible if...

6-part experiment:

A-Z → A-Z → A-Z → digit → digit → digit

$$26 \times 26 \times 26 \times 10 \times 10 \times 10$$

$$= 17,576,000$$

2-part experiment:

digit → 6-place license plate experiment

$$10 \times 17,576,000$$

$$= 175,760,000$$

General Principle of Counting

$$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i$$
Floors and ceilings

**Floor function**

\[ [x] \]

The largest integer \( \leq x \)

Check it out:

\[ [1/2] = 0 \quad [2.9] = 2 \quad [8.0] = 8 \quad [-1/2] = -1 \]

\[ [1/2] = 1 \quad [2.9] = 3 \quad [8.0] = 8 \quad [-1/2] = 0 \]

**Ceiling function**

\[ [x] \]

The smallest integer \( \geq x \)
Pigeonhole Principle

For positive integers $m$ and $n$,

if $m$ objects are placed in $n$ buckets,

then at least one bucket must contain

at least $\lceil m/n \rceil$ objects.

Example:

$m$ objects = 10 pigeons

$n$ buckets = 9 pigeonholes

At least one pigeonhole must

contain $\lceil m/n \rceil = 2$ pigeons.

Bounds: an important part of CS109
Balls and urns

\[ n \text{ balls} \]

\[ r \text{ urns (buckets)} \]

\[ \geq 1 \text{ bucket must contain at least } \left\lfloor \frac{m}{n} \right\rfloor \text{ objects} \]
Balls and urns    Hash Tables and strings

Consider a hash table with 100 buckets.
950 strings are hashed and added to the table.

1. Is it guaranteed that at least one bucket contains \textit{at least} 10 entries?
2. Is it guaranteed that at least one bucket contains \textit{at least} 11 entries?
3. Is it possible to have a bucket with \textit{no entries}?
Balls and urns  Hash Tables and strings

Consider a hash table with 100 buckets.  
950 strings are hashed and added to the table.

\[ n = 100 \]
\[ m = 950 \]

1. Is it guaranteed that at least one bucket contains **at least** 10 entries?  
Yes

2. Is it guaranteed that at least one bucket contains **at least** 11 entries?  
No

3. Is it possible to have a bucket with **no entries**?  
Sure
Takeaways from this lecture

Inclusion-Exclusion Principle (generalized Sum Rule)

If the outcome of an experiment can be either from Set $A$ or set $B$, where $A$ and $B$ may overlap, then the total number of outcomes of the experiment is

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

General Principle of Counting (generalized Product Rule)

If an experiment has $r$ steps, such that step $i$ has $n_i$ outcomes for all $i = 1, \ldots, r$, then the total number of outcomes of the experiment is

$$n_1 \times n_2 \times \cdots \times n_r = \prod_{i=1}^{r} n_i.$$
Unique 6-digit passcodes

How many unique 6-digit passcodes are possible?

Steps:
1. First digit in passcode 10 outcomes
2. Second digit in passcode 10 outcomes
   ...
6. Sixth digit in passcode 10 outcomes

Total $= n_1 \times n_2 \times \cdots \times n_6$
$= 10 \times 10 \times 10 \times 10 \times 10 \times 10$
$= 10^6$ passcodes
Unique 6-digit passcodes with six smudges

How many unique 6-digit passcodes are possible if a phone password uses each of six distinct numbers?
Sort $n$ indistinct objects
Sort $n$ distinct objects

Ayesha  Tim  Irina  Joey  Waddie
Sort $n$ distinct objects

Steps:
1. Choose 1\textsuperscript{st} can \hspace{1cm} 5 options
2. Choose 2\textsuperscript{nd} can \hspace{1cm} 4 options
   \hspace{1cm} ...
5. Choose 5\textsuperscript{th} can \hspace{1cm} 1 option

Total \hspace{1cm} = 5 \times 4 \times 3 \times 2 \times 1 \hspace{1cm} = 120
Permutations

A permutation is an ordered arrangement of distinct objects.

The number of unique orderings (permutations) of \( n \) distinct objects is

\[
 n! = n \times (n - 1) \times (n - 2) \times \cdots \times 2 \times 1.
\]
Unique 6-digit passcodes with six smudges

How many unique 6-digit passcodes are possible if a phone password uses each of six distinct numbers?

Total = 6!

= 720 passcodes
Unique 6-digit passcodes with five smudges

How many unique 6-digit passcodes are possible if a phone password uses each of five distinct numbers?

Steps:
1. Choose digit to repeat (5 outcomes)
2. Create passcode (permute 4 distinct, 2 indistinct)

Total = \( 5 \times \frac{6!}{2!} \) = 1,800 passcodes