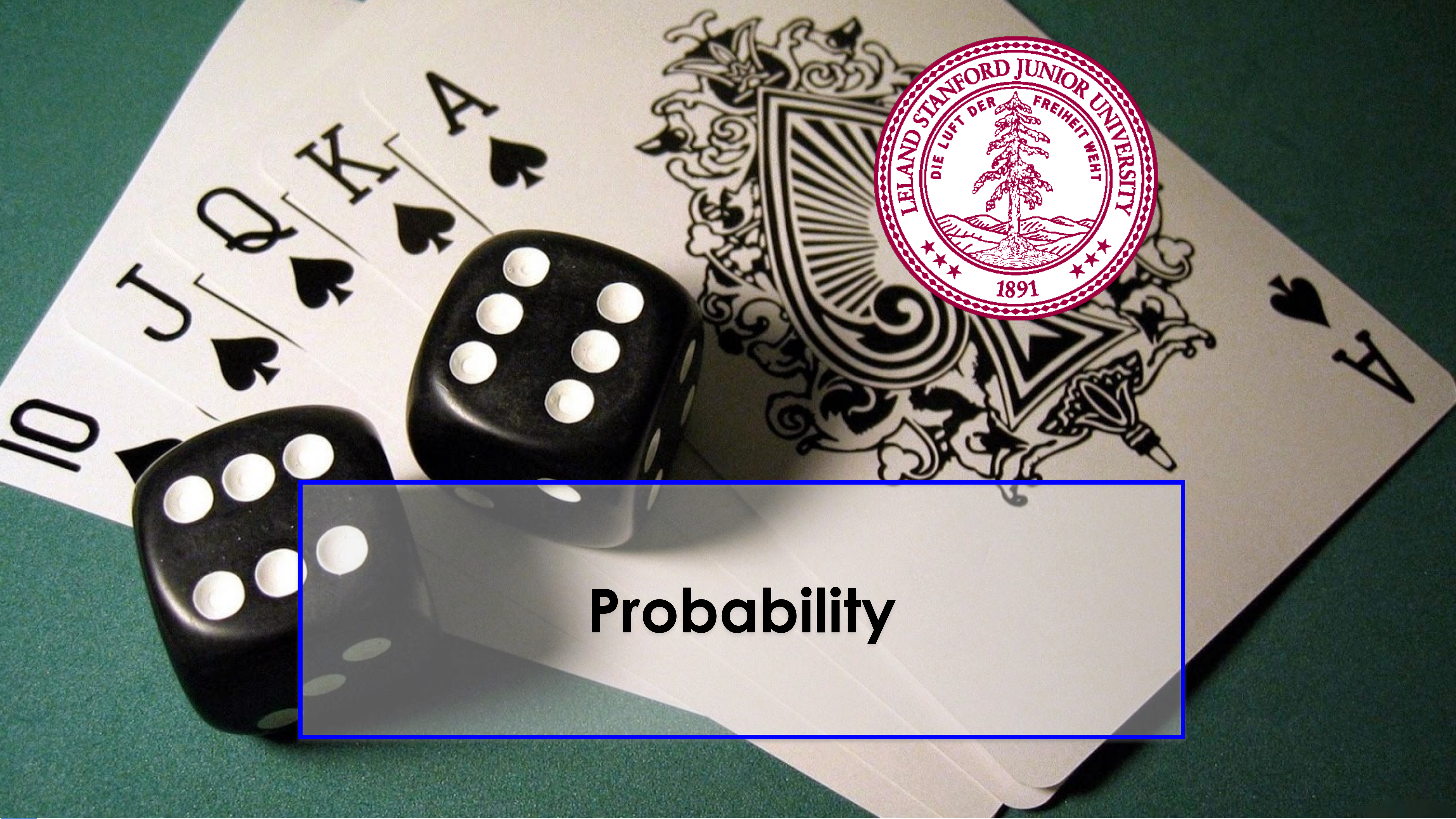
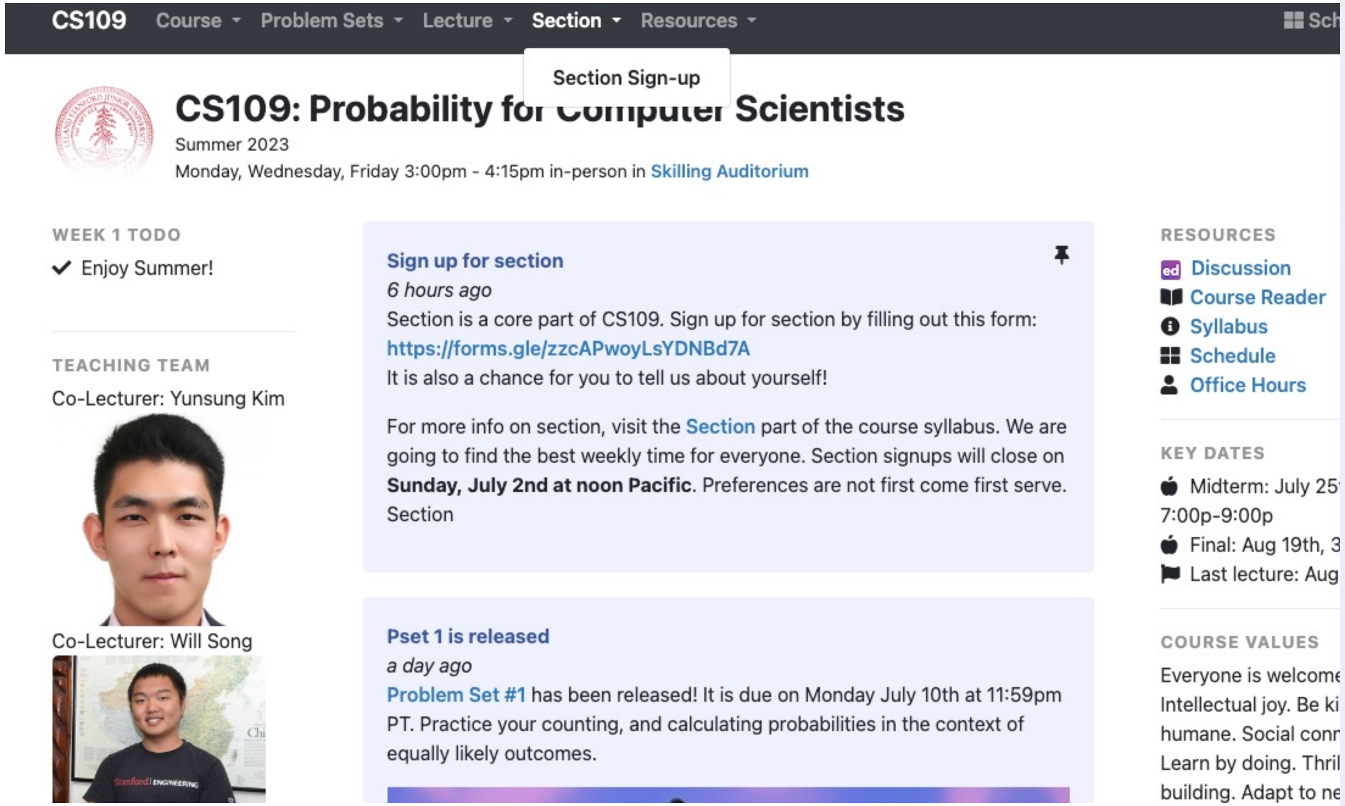


Probability

Announcements

- Sign up for sections by Sunday.
- Python Review Session next Thursday.
- Pset 1 due in 2 Mondays from now (July 7th)
  - Get started early! Pset 2 is planned to be out next Friday
- We are going to make history today.

Sign up for Sections by **Sunday noon**



CS109 Course Problem Sets Lecture Section Resources Sch

Section Sign-up

CS109: Probability for Computer Scientists

Summer 2023

Monday, Wednesday, Friday 3:00pm - 4:15pm in-person in Skilling Auditorium

WEEK 1 TODO

✓ Enjoy Summer!

TEACHING TEAM

Co-Lecturer: Yunsung Kim

Co-Lecturer: Will Song

Sign up for section

6 hours ago

Section is a core part of CS109. Sign up for section by filling out this form: https://forms.gle/zzcAPwoyLsYDNBd7A

It is also a chance for you to tell us about yourself!

For more info on section, visit the Section part of the course syllabus. We are going to find the best weekly time for everyone. Section signups will close on **Sunday, July 2nd at noon Pacific**. Preferences are not first come first serve. Section

Pset 1 is released

a day ago

Problem Set #1 has been released! It is due on Monday July 10th at 11:59pm PT. Practice your counting, and calculating probabilities in the context of equally likely outcomes.

RESOURCES

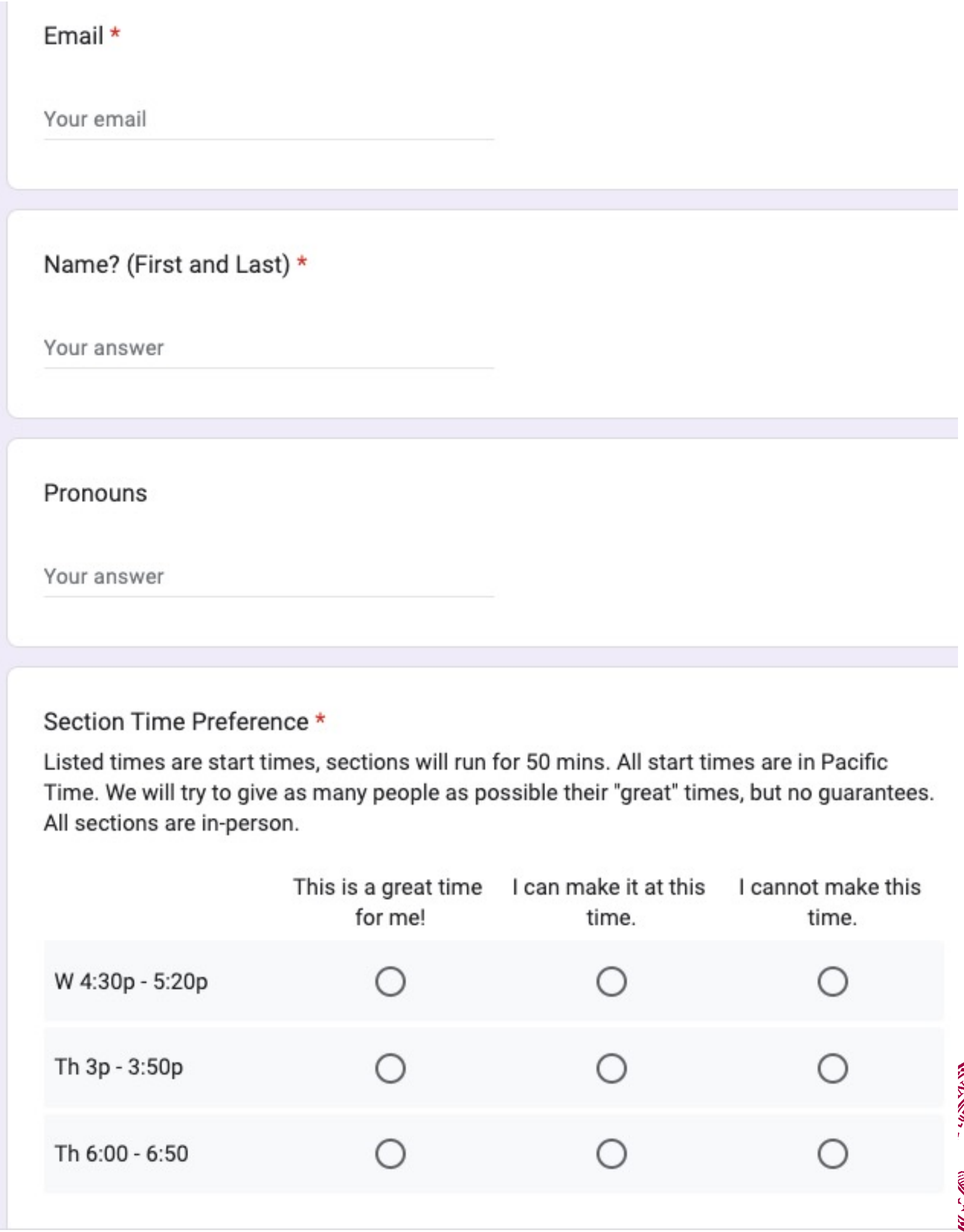
- Discussion
- Course Reader
- Syllabus
- Schedule
- Office Hours

KEY DATES

- Midterm: July 25
  7:00p-9:00p
- Final: Aug 19th, 3
- Last lecture: Aug

COURSE VALUES

Everyone is welcome. Intellectual joy. Be ki humane. Social conr Learn by doing. Thril building. Adapt to ne



Email *

Your email

Name? (First and Last) *

Your answer

Pronouns

Your answer

Section Time Preference *

Listed times are start times, sections will run for 50 mins. All start times are in Pacific Time. We will try to give as many people as possible their "great" times, but no guarantees. All sections are in-person.

	This is a great time for me!	I can make it at this time.	I cannot make this time.
W 4:30p - 5:20p	○	○	○
Th 3p - 3:50p	○	○	○
Th 6:00 - 6:50	○	○	○

PSet FAQ

Q: Can I check my answers (many) times?

A: Yes. No penalty. We will know if you try all solutions.

Q: Will you just grade us on if the answers are correct?

A: No! You must explain your answer. Graded on style and correctness.

Q: How much explanation should we provide?

A: As detailed as the course reader worked examples

Q: Must I do all my work on the app?

A: No

PSet Example Answer

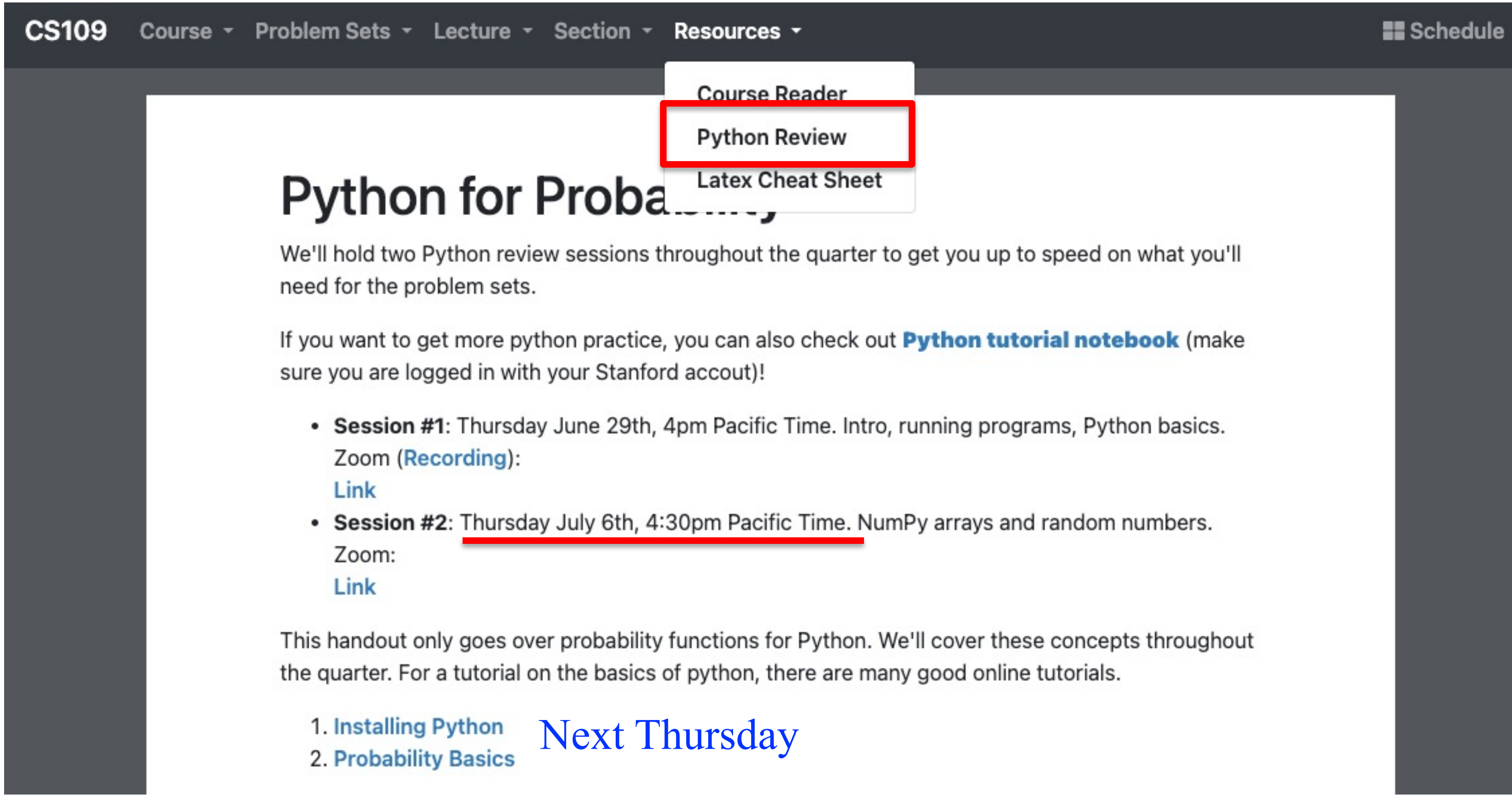
Enigma Machine

Two wires: How many ways are there to place exactly two wires? Recall that wires are not considered distinct. Each letter can have at most one wire connected to it, thus you couldn't have a wire connect 'K' to 'L' and another one connect 'L' to 'X'

There are $\binom{26}{2}$ ways to place the first wire and $\binom{24}{2}$ ways to place the second wire. However, since the wires are indistinct, we have double counted every possibility. Because every possibility is counted twice we should divide by 2:

$$\text{Total} = \frac{\binom{26}{2} \cdot \binom{24}{2}}{2} = 44,850$$

Python (NumPy) Review Session



CS109 Course Problem Sets Lecture Section Resources Schedule

Course Reader
Python Review
Latex Cheat Sheet

Python for Proba...

We'll hold two Python review sessions throughout the quarter to get you up to speed on what you'll need for the problem sets.

If you want to get more python practice, you can also check out **Python tutorial notebook** (make sure you are logged in with your Stanford accout)!

- **Session #1**: Thursday June 29th, 4pm Pacific Time. Intro, running programs, Python basics. Zoom (Recording):
  Link
- **Session #2**: Thursday July 6th, 4:30pm Pacific Time. NumPy arrays and random numbers. Zoom:
  Link

This handout only goes over probability functions for Python. We'll cover these concepts throughout the quarter. For a tutorial on the basics of python, there are many good online tutorials.

1. Installing Python
2. Probability Basics

Next Thursday

Online and will be recorded!

Review

Counting Rules

Counting operations on *n* objects

- **Sort, order matters {perms}**
  - Distinct: $n!$
  - Some Distinct: $\frac{n!}{n_1!n_2!\dots}$
- **Choose *k* {combinations}**
  - Distinct: $\binom{n}{k} = \frac{n!}{k!(n-k)!}$
- **Put in *r* buckets**
  - Distinct: r^n
  - None Distinct: $\frac{(n+r-1)!}{n!(r-1)!}$

Counting Rules

Counting operations on *n* objects

- **Sort, order matters {perms}**
  - Distinct: $n!$
  - Some Distinct: $\dfrac{n!}{n_1!n_2!\dots}$
- **Choose *k* {combinations}**
  - Distinct: $\dbinom{n}{k} = \dfrac{n!}{k!(n-k)!}$
- **Put in *r* buckets**
  - Distinct: r^n
  - None Distinct: $\dfrac{(n+r-1)!}{n!(r-1)!}$

Poker Hands

- How many unique hands of 5 cards are there in a 52 card deck?

Choosing 5 distinct cards from a deck of 52 cards → $\binom{52}{5}$

End Review

Sample Space

- **<u>Sample space</u>**, S, is set of all possible outcomes of an experiment
  - Coin flip: S = {Head, Tails}
  - Flipping two coins: S = {{H, H}, {H, T}, {T, H}, {T, T}}
  - Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}
  - # emails in a day: $S = \{x \mid x \in \mathbf{Z}, x \geq 0\}$ {non-neg. ints}
  - YouTube hrs. in day: $S = \{x \mid x \in \mathbf{R}, 0 \leq x \leq 24\}$

Event Space

- **Event**, E, is some subset of S $\{E \subseteq S\}$
  - Coin flip is heads: $E = \{\text{Head}\}$
  - $\geq$ 1 head on 2 coin flips: $E = \{\{H, H\}, \{H, T\}, \{T, H\}\}$
  - Roll of die is 3 or less: $E = \{1, 2, 3\}$
  - # emails in a day $\leq$ 20: $E = \{x \mid x \in \mathbf{Z}, 0 \leq x \leq 20\}$
  - Wasted day {$\geq$ 5 YT hrs.}: $E = \{x \mid x \in \mathbf{R}, 5 \leq x \leq 24\}$

Note: When Ross uses: $\subset$, he really means: $\subseteq$

Event Space

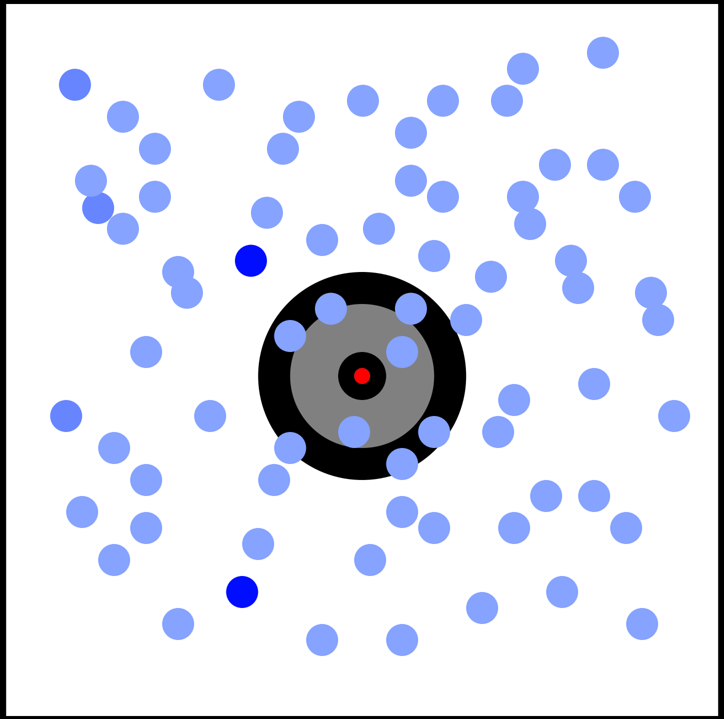
Sample Space, S

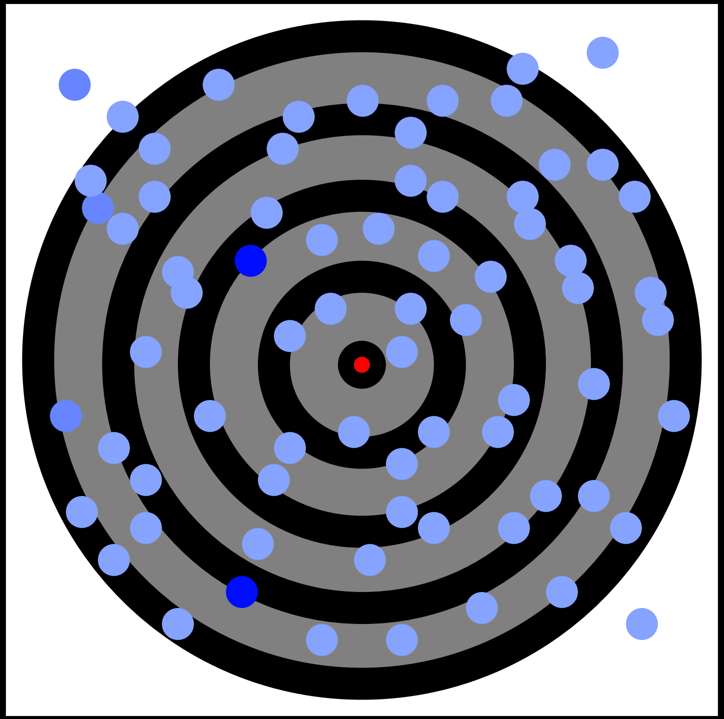
- Coin flip
  $S = \{\text{Heads}, \text{Tails}\}$
- Flipping two coins
  $S = \{(\text{H,H}), (\text{H,T}), (\text{T,H}), (\text{T,T})\}$
- Roll of 6-sided die
  $S = \{1, 2, 3, 4, 5, 6\}$
- # emails in a day
  $S = \{x \mid x \in \mathbb{Z}, \ x \geq 0\}$
- TikTok hours in a day
  $S = \{x \mid x \in \mathbb{R}, 0 \leq x \leq 24\}$

Event, E

- Flip lands heads
  $E = \{\text{Heads}\}$
- ≥ 1 head on 2 coin flips
  $E = \{(\text{H,H}), (\text{H,T}), (\text{T,H})\}$
- Roll is 3 or less:
  $E = \{1, 2, 3\}$
- Low email day (≤ 20 emails)
  $E = \{x \mid x \in \mathbb{Z}, \ 0 \leq x \leq 20\}$
- Wasted day (≥ 5 TT hours):
  $E = \{x \mid x \in \mathbb{R}, 5 \leq x \leq 24\}$

What is a probability?






90
MLB 040

[suspense]

Number between 0 and 1

A number to which we ascribe meaning

$$P(E)$$

* Our belief that an event E occurs

A number to which we ascribe meaning

$$\Pr(E)$$

* Our belief that an event E occurs

Axioms of Probability

Recall: S = all possible outcomes. E = the event.

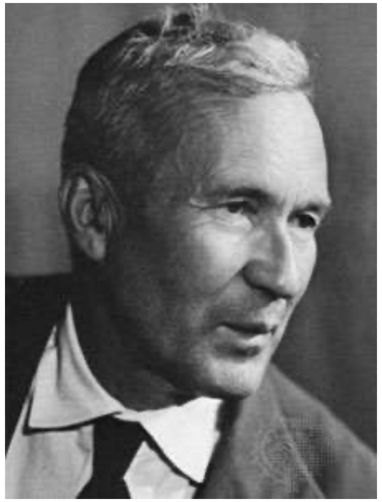
- Axiom 1: $0 \leq P(E) \leq 1$
- Axiom 2: $P(S) = 1$
- Axiom 3: If events E and F are mutually exclusive:

$$P(E \cup F) = P(E) + P(F)$$

Axioms of Probability

Recall: S = all possible outcomes. E = the event.

Kolmogorov, 1933



- Axiom 1: $0 \leq P(E) \leq 1$
- Axiom 2: $P(S) = 1$
- Axiom 3: If events E and F are mutually exclusive:

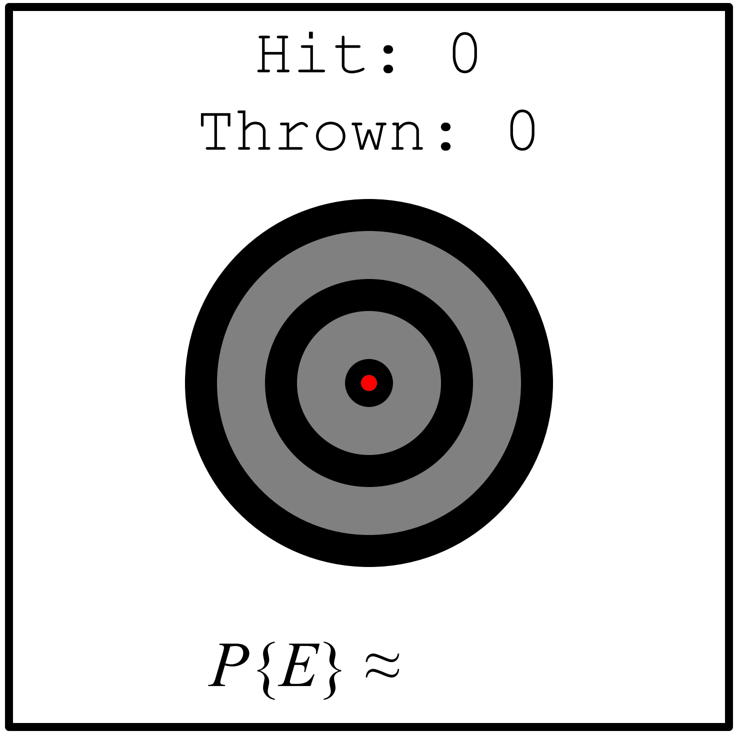
$$P(E \cup F) = P(E) + P(F)$$

Real-life meaning of Probability

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

Real-life meaning of Probability

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$



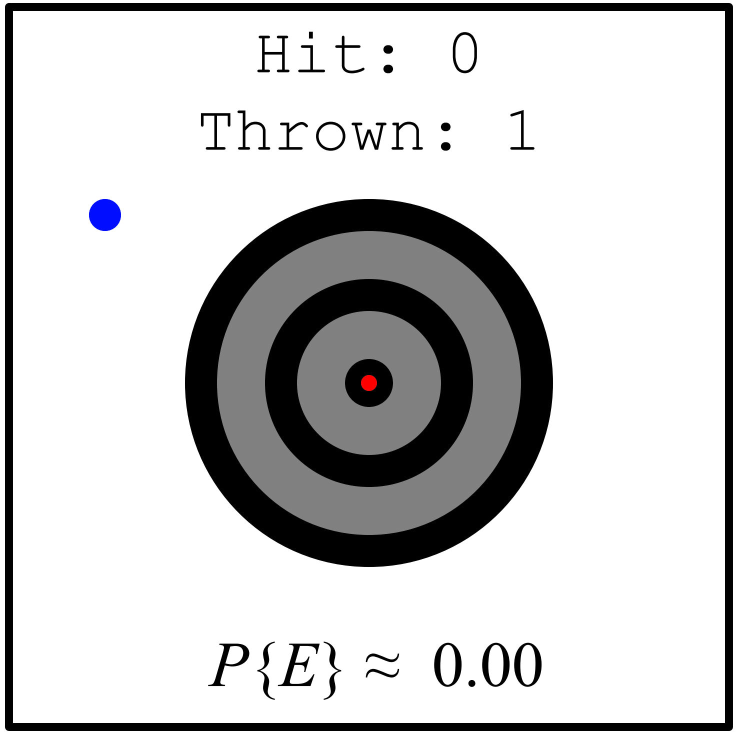
n is the number of trails

The "event" *E* is that you hit the target

Real-life meaning of Probability

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

n is the number of trails

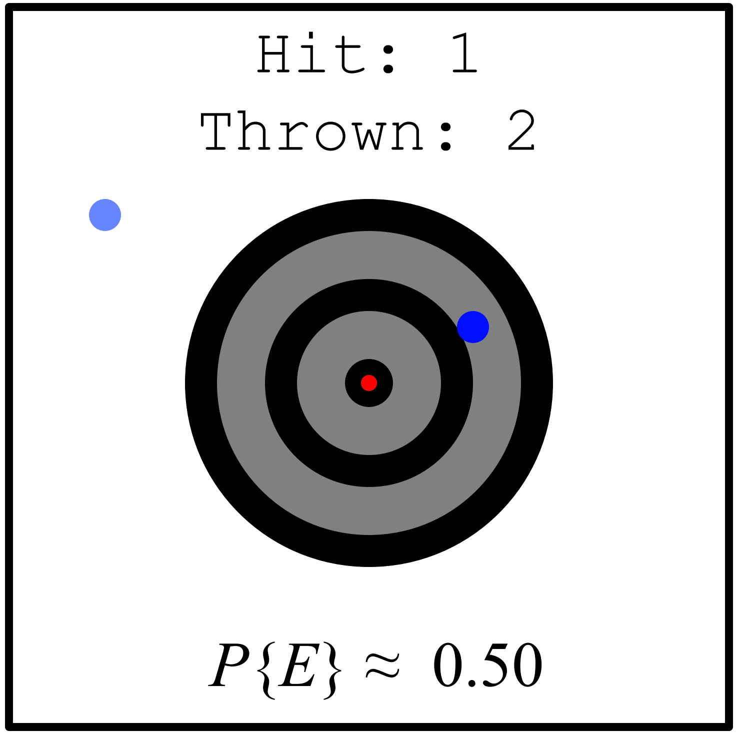


The "event" *E* is that you hit the target

Real-life meaning of Probability

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

n is the number of trails

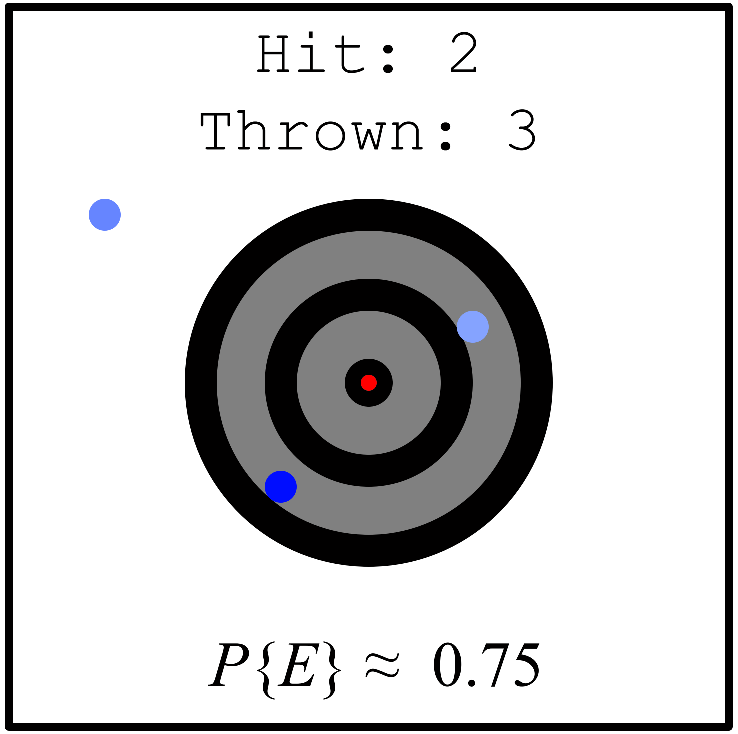


The "event" E is that you hit the target

Real-life meaning of Probability

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

n is the number of trails

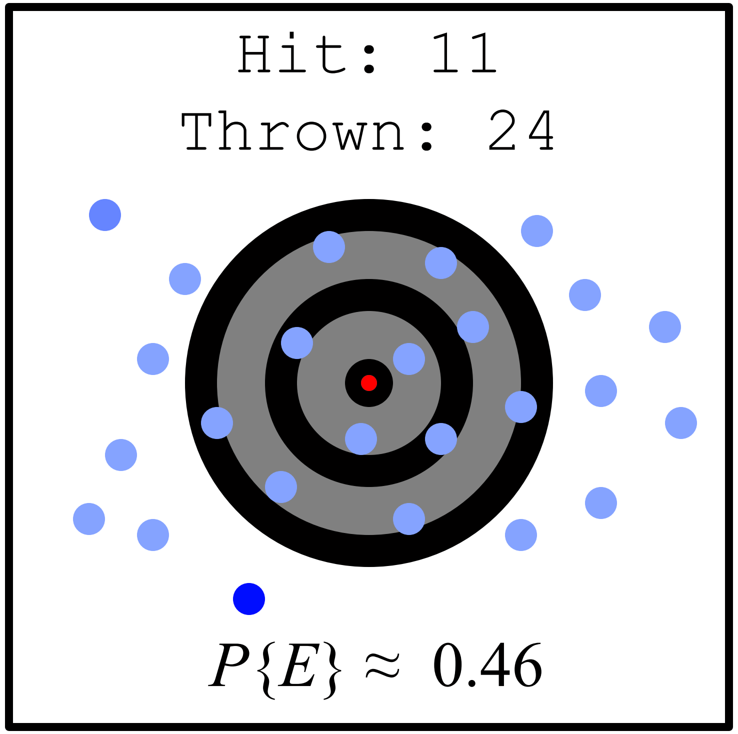


The “event” *E* is that you hit the target

Real-life meaning of Probability

$$P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$$

n is the number of trails



The "event" E is that you hit the target

Special Case of Probability

Equally Likely Outcomes

Equally Likely Outcomes

Some sample spaces have equally likely outcomes.

- Coin flip: S = {Head, Tails}
- Flipping two coins: S = {{H, H}, {H, T}, {T, H}, {T, T}}
- Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}

If we have equally likely outcomes, then P{Each outcome} $= \frac{1}{|S|}$

Therefore $$P(E) = \frac{\text{\# outcomes in E}}{\text{\# outcomes in S}} = \frac{|E|}{|S|}$$ {by Axiom 3}

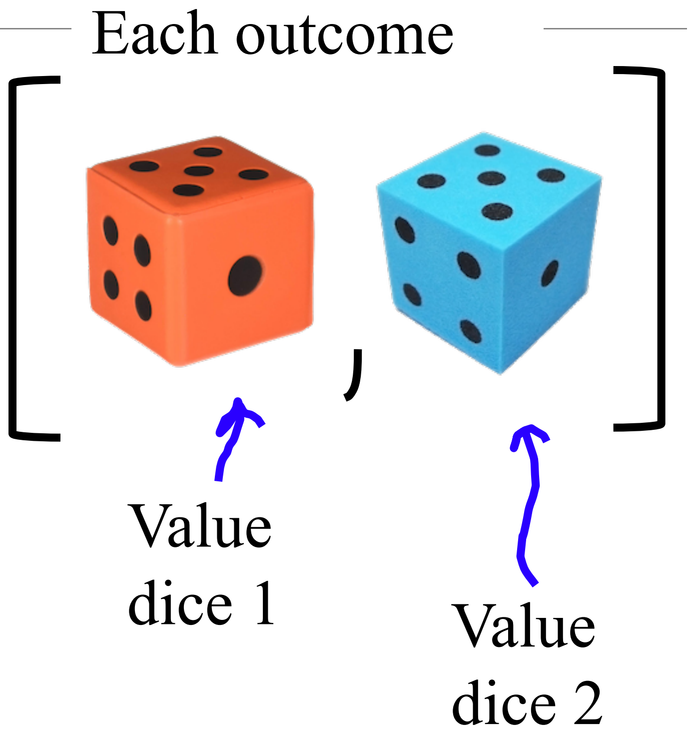
Not Everything is Equally Likely

- Play lottery.
  - What is P{Win}?

- S = {Lose, Win}
- E = {Win}
- P{Win} = |E|/|S| = 1/2 = 50%

Sum of Two Die = 7?

Roll two 6-sided dice. What is P[sum = 7]?



S = { [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] }

1. Choose a sample space S (hopefully one that's equally likely)
2. Define the event set E that is of interest

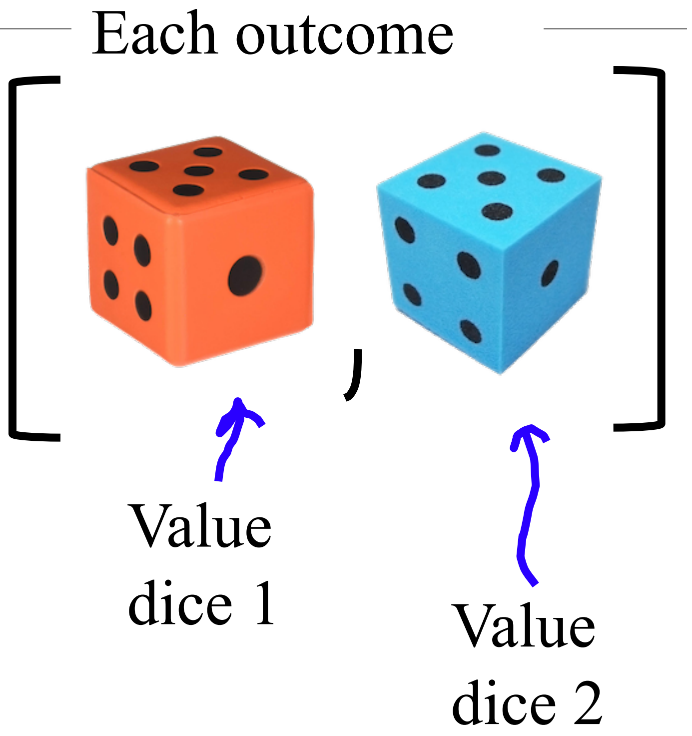
Sum of Two Die = 7?

Roll two 6-sidex dice. What is probability the sum = 7?

Let E be the event that the sum is 7

S = {

[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] }



E = *in blue*

$$P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.16\bar{6}$$

Is it correct?

$$P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.16\bar{6}$$

Sum of Two Die = 7?

```python
import random
from tqdm import tqdm

N_TRIALS = 10000000  # getting close to infinity
TARGET_SUM = 7       # do the two dice sum to 6?

def main():
    n_events = 0
    for i in tqdm(range(N_TRIALS)):
        dice_total = run_experiment()
        if dice_total == TARGET_SUM:
            n_events += 1
    pr_e = n_events / N_TRIALS
    print(f'after {N_TRIALS} trials')
    print('P(E) ≈ ', pr_e)

def run_experiment():
    d_1 = roll_dice()
    d_2 = roll_dice()
    return d_1 + d_2

def roll_dice():
    # give me a random dice roll
    # alternatively random.randint(1, 7)
    return random.choice([1,2,3,4,5,6])

if __name__ == '__main__':
    # this starts the program in main
    main()
```

$$P(E) = \frac{|E|}{|S|} = \frac{6}{36} = 0.16\bar{6}$$

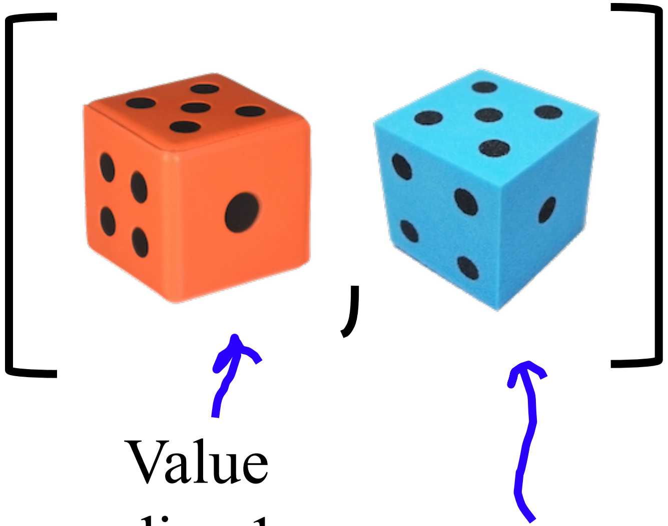
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL

```
piech@Chriss-MBP-2 3 % python dice_soln.py
after 10000000 trials
P(E) =  0.1666913
```

Sum of Two Die = 2?

Roll two 6-sidex dice. What is probability the sum = 7?
Let E be the event that the sum is 7

Each outcome



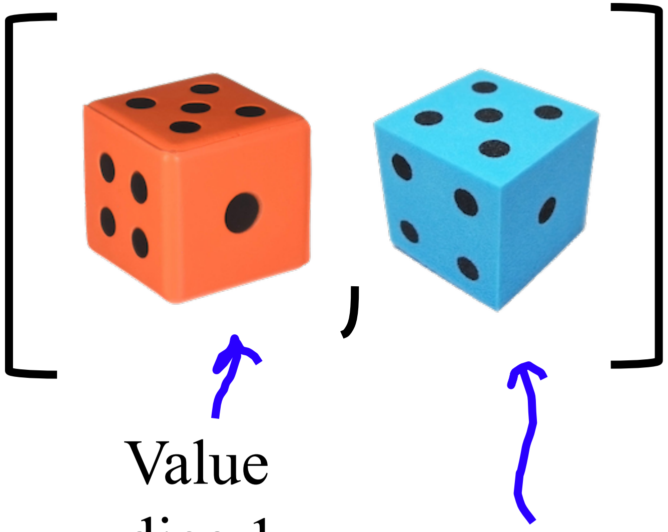
S = { [1,1] [1,2] [1,3] [1,4] [1,5] [1,6]
[2,1] [2,2] [2,3] [2,4] [2,5] [2,6]
[3,1] [3,2] [3,3] [3,4] [3,5] [3,6]
[4,1] [4,2] [4,3] [4,4] [4,5] [4,6]
[5,1] [5,2] [5,3] [5,4] [5,5] [5,6]
[6,1] [6,2] [6,3] [6,4] [6,5] [6,6] }

E =

Sum of Two Die = 2?

Roll two 6-sidex dice. What is probability the sum = 2?
Let E be the event that the sum is 2

Each outcome



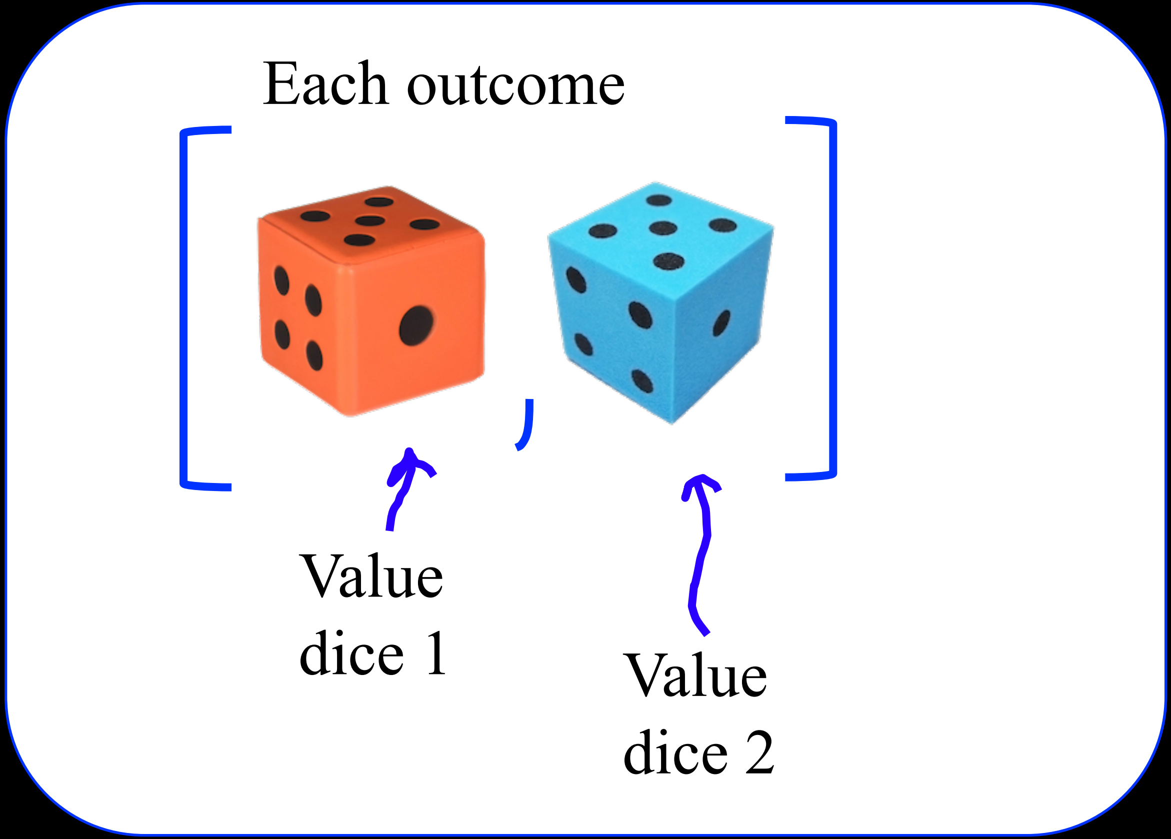
S = {

[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] }

E = *in red*

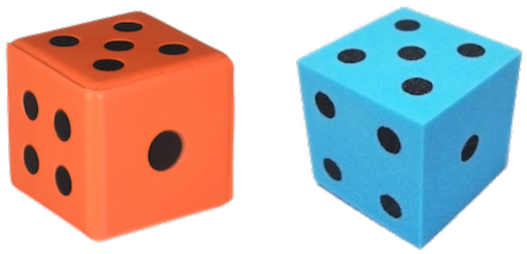
$$P(E) = \frac{|E|}{|S|} = \frac{1}{36} = 0.02\overline{7}$$

Other ways to make a Sample Space?



Sum of Two Die: Three options for the sample space

Value dice 1



Value dice 2

Think of the die as **distinct**

[5 , 5]

Value dice 1

Value dice 2

Think of the die as **indistinct**

{ 5 , 5 }

Value of a dice

Value of a dice

Just look at the sum

10

Sum of Two Die = 7? Bug: Die are Indistinct

Each outcome

Just look at the sum

Roll two 6-sidex dice. What is probability the sum = 7?
Let E be the event that the sum is 7

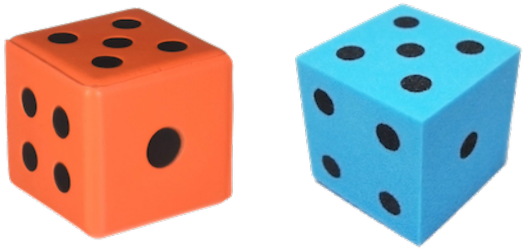
S = { ~~1~~ 2 3 4 5 6
7 8 9 10 11 12 }

E = *in red*

$$P(E) = \frac{|E|}{|S|} = \frac{1}{11} = 0.0\overline{909}$$

Sum of Two Die: Three options for the sample space

Value dice 1



Value dice 2

Think of the die as **distinct**

[5 , 5]

Value dice 1

Value dice 2

Think of the die as **indistinct**

{ 5 , 5 }

Value of a dice

Value of a dice

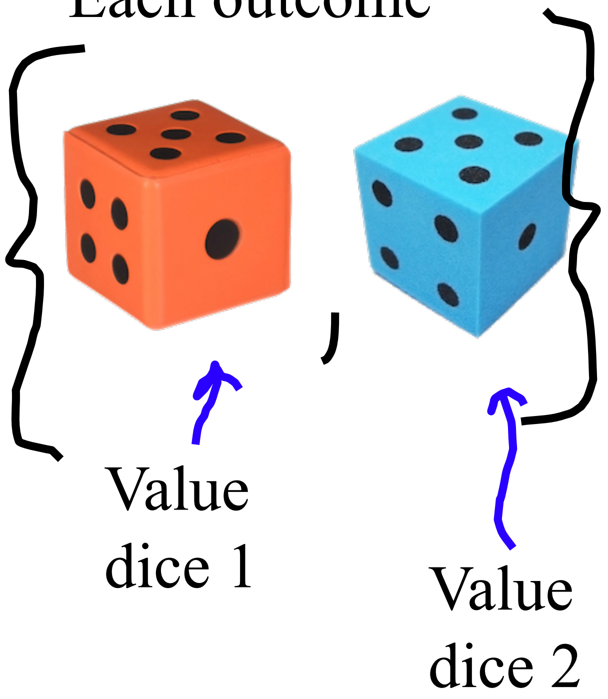
Just look at the sum

10

Sum of Two Die = 7? Bug: Die are Indistinct

Roll two 6-sidex dice. What is P(sum = 7)?

S = { {1,1} {1,2} {1,3} {1,4} {1,5} **{1,6}**
{2,2} {2,3} {2,4} **{2,5}** {2,6}
{3,3} **{3,4}** {3,5} {3,6}
{4,4} {4,5} {4,6}
{5,5} {5,6}
{6,6} }

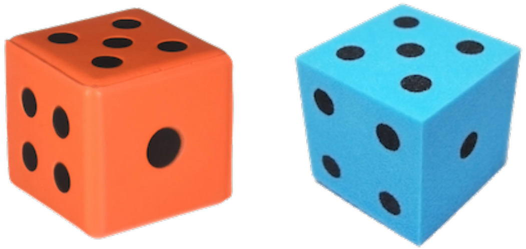


E = *in blue*

$$P(E) = \frac{|E|}{|S|} = \frac{3}{20} = 0.15?$$

Sum of Two Die: Three options for the sample space

Value dice 1



Value dice 2

Think of the die as **distinct**

[5 , 5]

Value dice 1

Value dice 2

✓

Think of the die as **indistinct**

{ 5 , 5 }

Value of a dice

Value of a dice

✗

Just look at the sum

10

✗

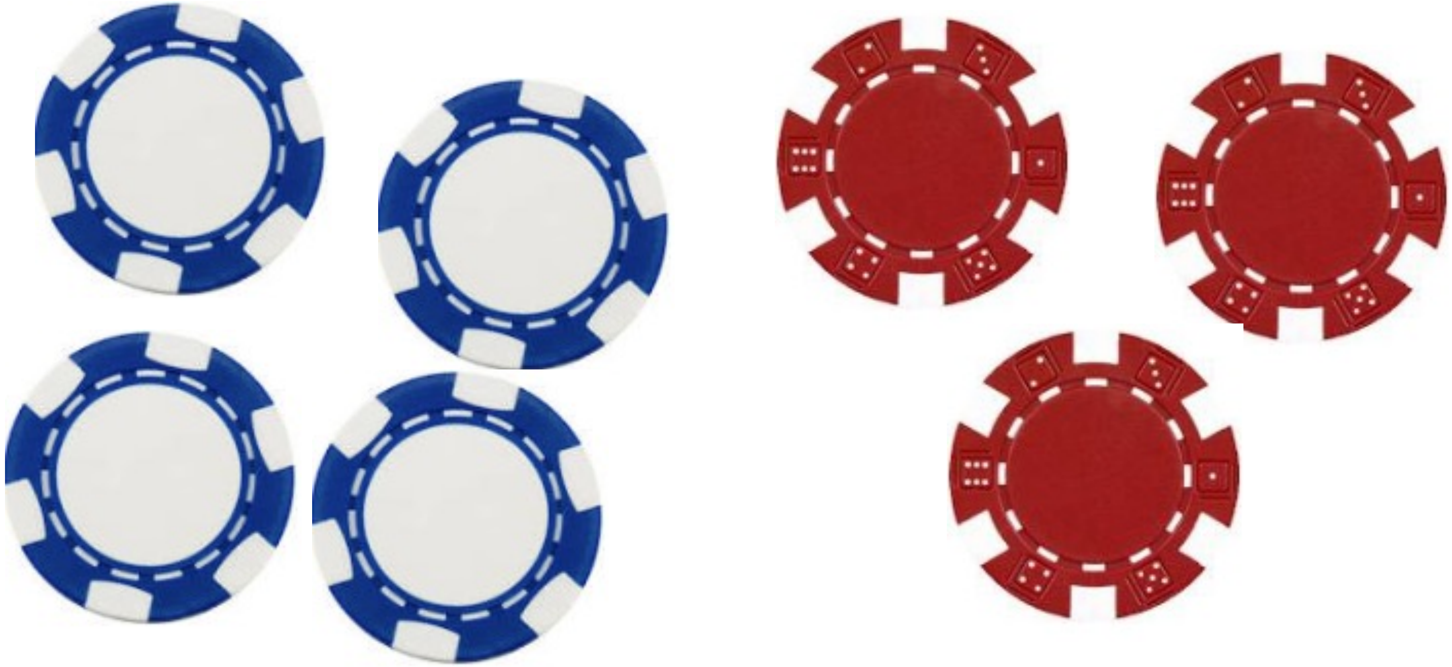
Sum of Two Die: Three options for the sample space

To get equally likely outcomes, it often helps to think of <u>items as distinct</u>, rather than indistinct.

Casino Chips

- 4 blue chips ($10) and 3 red chips ($50). 3 chips are drawn.
  - What is P(3 chips are worth $110)? =P(1 blue chip and 2 red chips)

Equally likely sample space? Thought experiment



4 blue

3 red

The Choice of Sample Space is Yours!

	Distinct	Indistinct
Unordered	$\{B_1, R_2, R_3\}$ $\{B_1, B_2, B_3\}$	{2 red, 1 blue} {3 blues}
Ordered	$[B_1, R_2, R_3]$ $[B_1, B_2, B_3]$	[blue, red, red] [blue, blue, blue]

Which choice will lead to equally likely outcomes?

pigs and cows

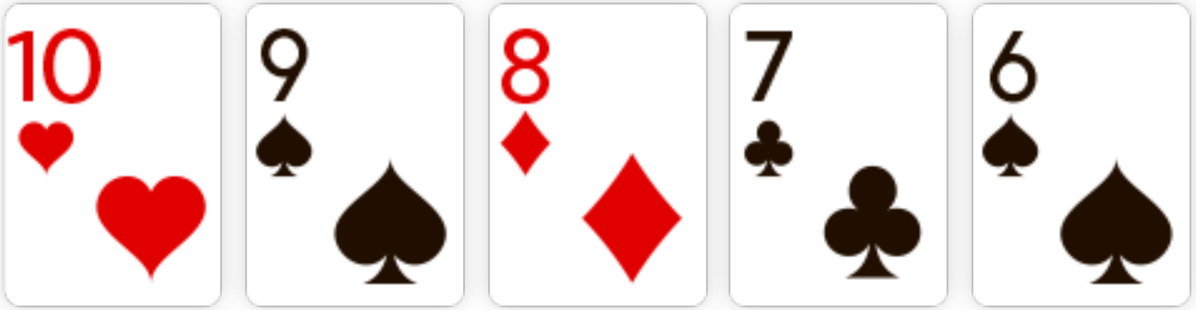
- 4 blues and 3 reds in a Bag. 3 drawn.
  - What is P(1 blue and 2 red drawn)?
- Ordered and Distinct:
  - Pick 3 ordered items: |S| = 7 * 6 * 5 = 210
  - Pick blue as either 1st, 2nd, or 3rd item:
    |E| = {4 * 3 * 2} + {3 * 4 * 2} + {3 * 2 * 4} = 72
  - P(1 blue, 2 red) = 72/210 = 12/35
- Unordered:
  - |S| = $\binom{7}{3}$ = 35
  - |E| = $\binom{4}{1}\binom{3}{2}$ = 12
  - P(1 blue, 2 red) = 12/35

Make indistinct items **distinct** to get equally likely sample space outcomes

*You will need to use this "trick" with high probability

Straight Poker Hand

- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - What is P(straight)?



Straight Poker Hand

- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - What is P(straight)?

What is an example of one outcome?

Is each outcome equally likely?

$$|S| = \binom{52}{5}$$

$$|E| = 10 \cdot \binom{4}{1}^5$$

$$P(\text{straight}) = \frac{|E|}{|S|} = \frac{10 \cdot \binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$$

Straight Poker Hand

- Consider 5 card poker hands.
  - "straight" is 5 consecutive rank cards of any suit
  - "straight flush" is 5 consecutive rank cards of same suit
  - What is P(straight, but not straight flush)?

$$|S| = \binom{52}{5}$$

$$|E| = 10\binom{4}{1}^5 - 10\binom{4}{1}$$

$$P(\text{straight}) = \frac{|E|}{|S|} = \frac{10\binom{4}{1}^5 - 10\binom{4}{1}}{\binom{52}{5}} \approx 0.00392$$

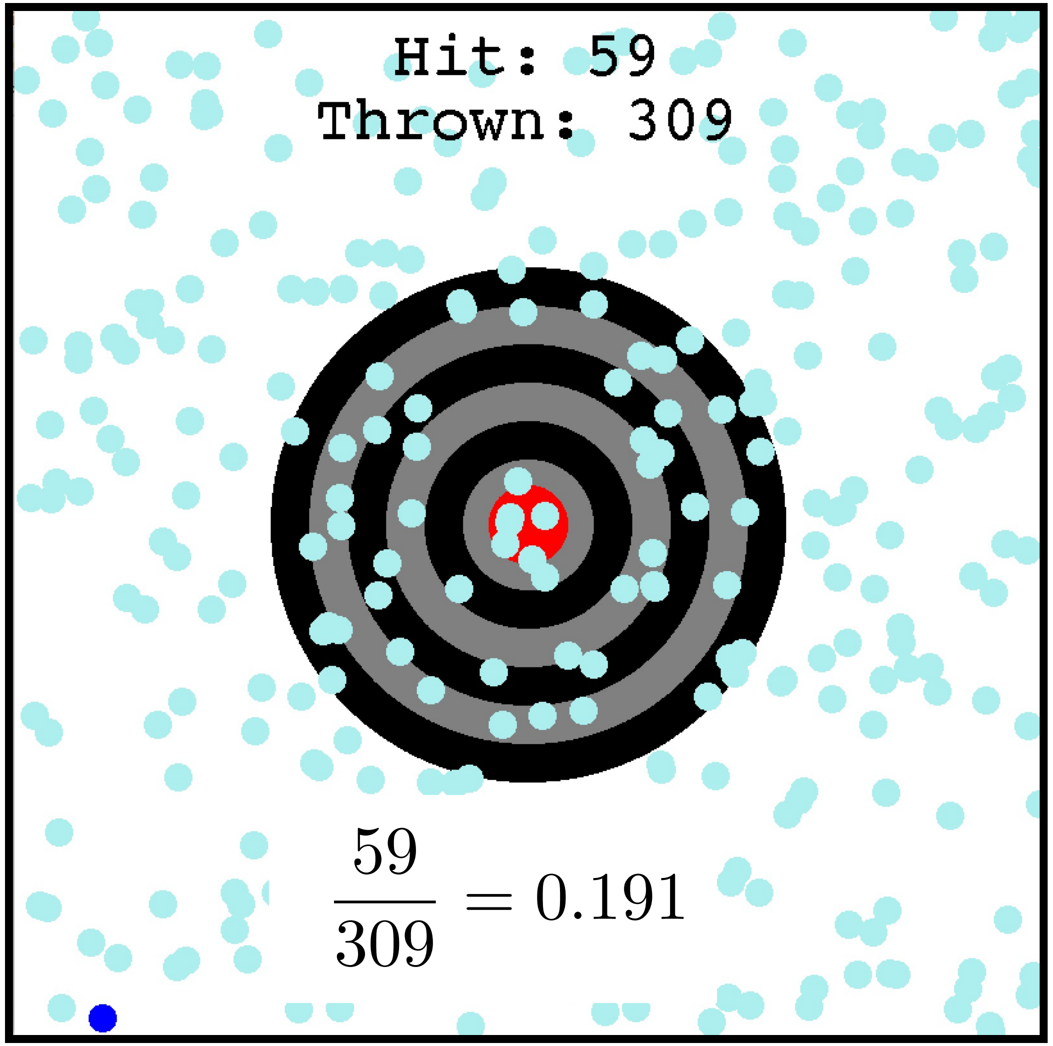
When approaching an "**equally likely probability**" problem, start by defining **sample spaces** and **event spaces**.

Chip Defect Detection

- n chips manufactured, 1 of which is defective.
- k chips randomly selected from n for testing.
  - What is P{defective chip is in k selected chips}?
- $|S| = \binom{n}{k}$
- $|E| = \binom{1}{1}\binom{n-1}{k-1}$
- P(defective chip is in k selected chips)

$$= \frac{\binom{1}{1}\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\frac{(n-1)!}{(k-1)!(n-k)!}}{\frac{n!}{k!(n-k)!}} = \frac{k}{n}$$

Target Revisited



$$\frac{59}{309} = 0.191$$

Screen size = 800x800
Radius of target = 200

The dart is equally likely to land anywhere on the screen.

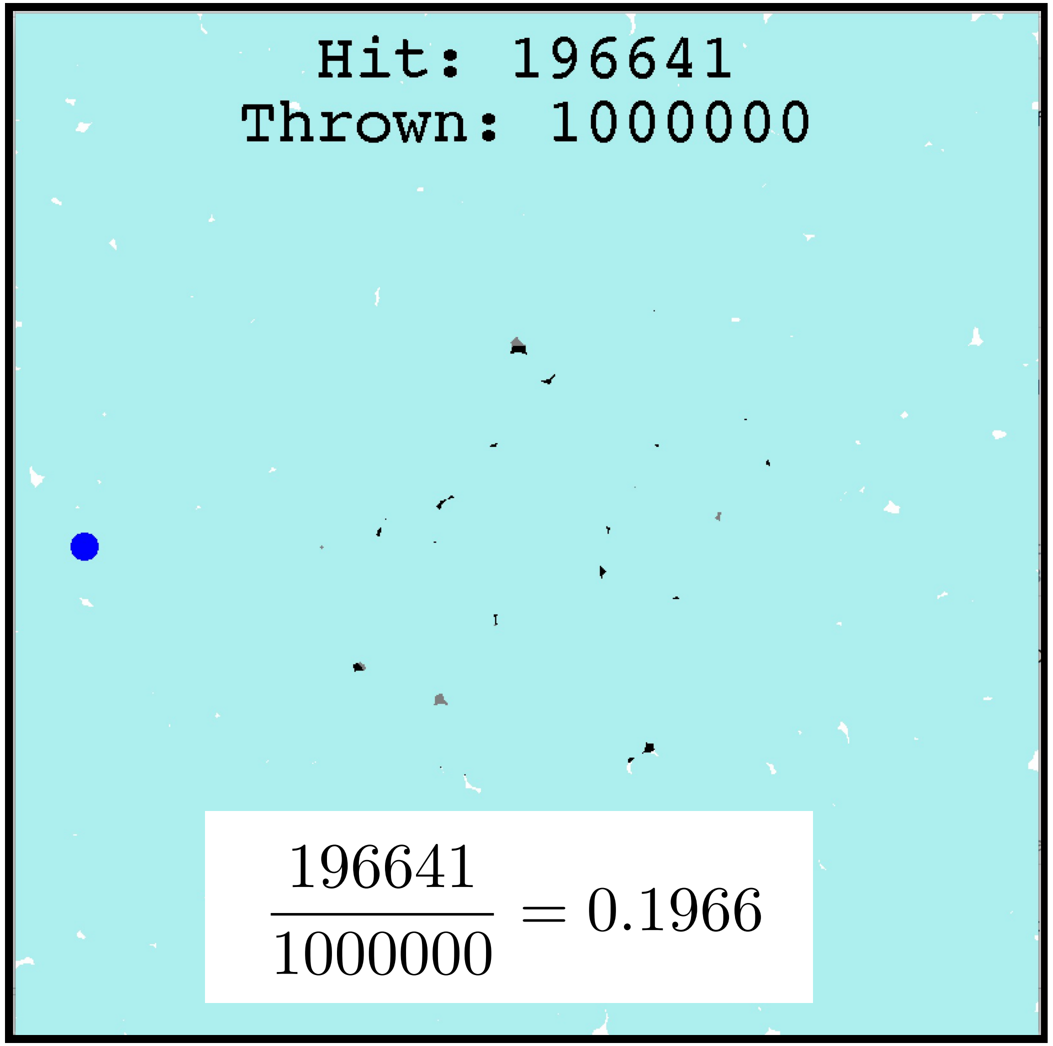
What is the probability of hitting the target?

$$|S| = 800^2$$

$$|E| = \pi 200^2$$

$$p(E) = \frac{\pi \cdot 200^2}{800^2} \approx 0.1963$$

Target Revisited



$$\frac{196641}{1000000} = 0.1966$$

Screen size = 800x800
Radius of target = 200

The dart is equally likely to land anywhere on the screen.

What is the probability of hitting the target?

$$|S| = 800^2$$

$$|E| = \pi 200^2$$

$$p(E) = \frac{\pi \cdot 200^2}{800^2} \approx 0.1963$$

Let it find you.

SERENDIPITY

the effect by which one accidentally stumbles upon something truely wonderful, especially while looking for something entirely unrelated.

WHEN YOU MEET YOUR BEST FRIEND

Somewhere you didn't expect to.

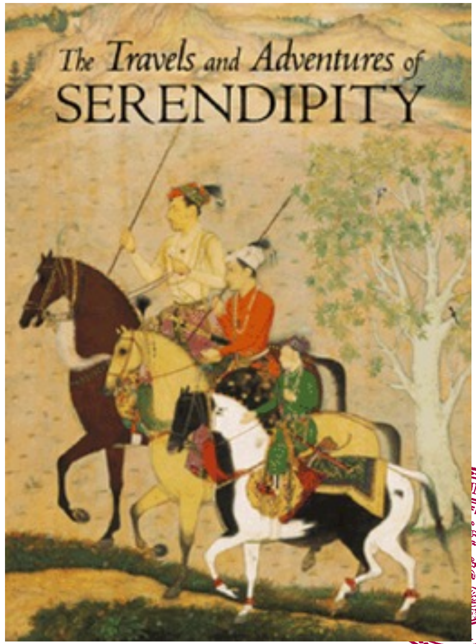
Serendipity

- Say the population of Stanford is 17,000 people
  - You are friends with 80 people?
  - Walk into a room, see 62 random people.
  - What is the probability that you see someone you know?
  - Assume you are equally likely to see each person at Stanford

$$|S| = \binom{17{,}000}{62}$$

$$|E^C| = \binom{17{,}000 - 80}{62}$$

$$P(\text{see someone you know})$$
$$= P(\text{see 1 or more friends})$$
$$= 1 - P(\text{don't see anyone you know})$$



$$P(E) = 1 - P(E^C) = 1 - \frac{|E^C|}{|S|} \approx 0.1914$$

Many times it is easier to calculate $P(E^C)$.

$$P(E^C) = 1 - P(E)$$

(We'll prove this in just a bit)

Back to 3 Axioms

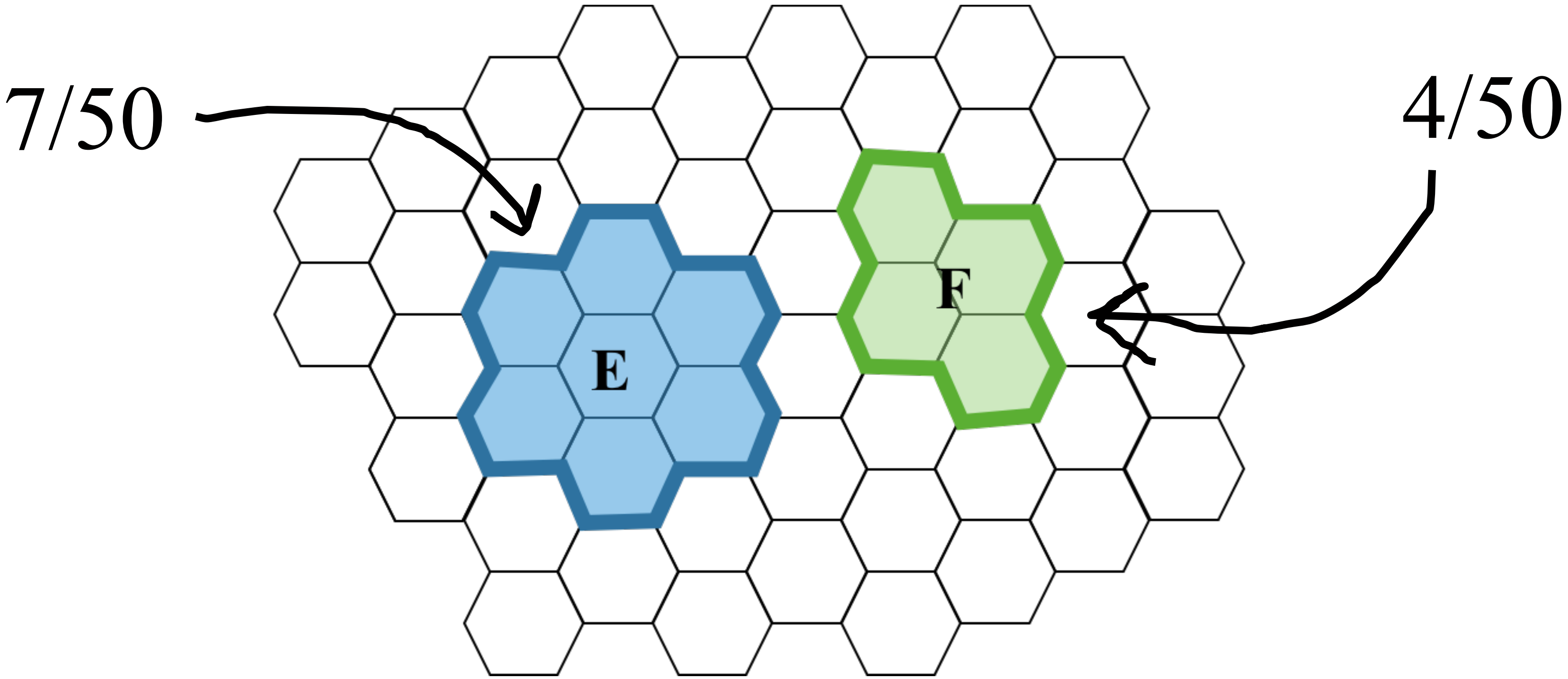
Axioms of Probability

Recall: S = all possible outcomes. E = the event.

- Axiom 1: $0 \leq P(E) \leq 1$
- Axiom 2: $P(S) = 1$
- Axiom 3: If events E and F are mutually exclusive:

$$P(E \cup F) = P(E) + P(F)$$

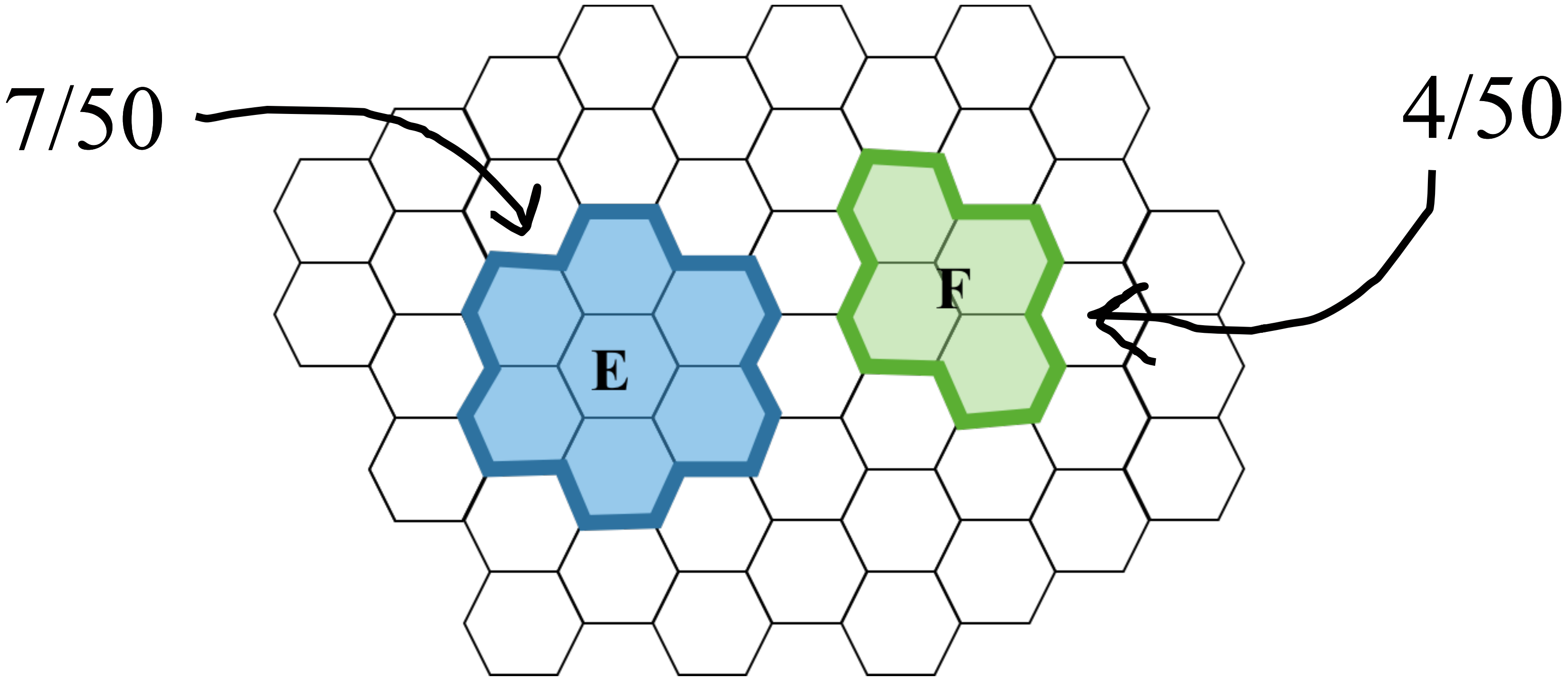
Mutually Exclusive Events



If events are mutually exclusive, probability of OR is simple:

$$P(E \cup F) = P(E) + P(F)$$

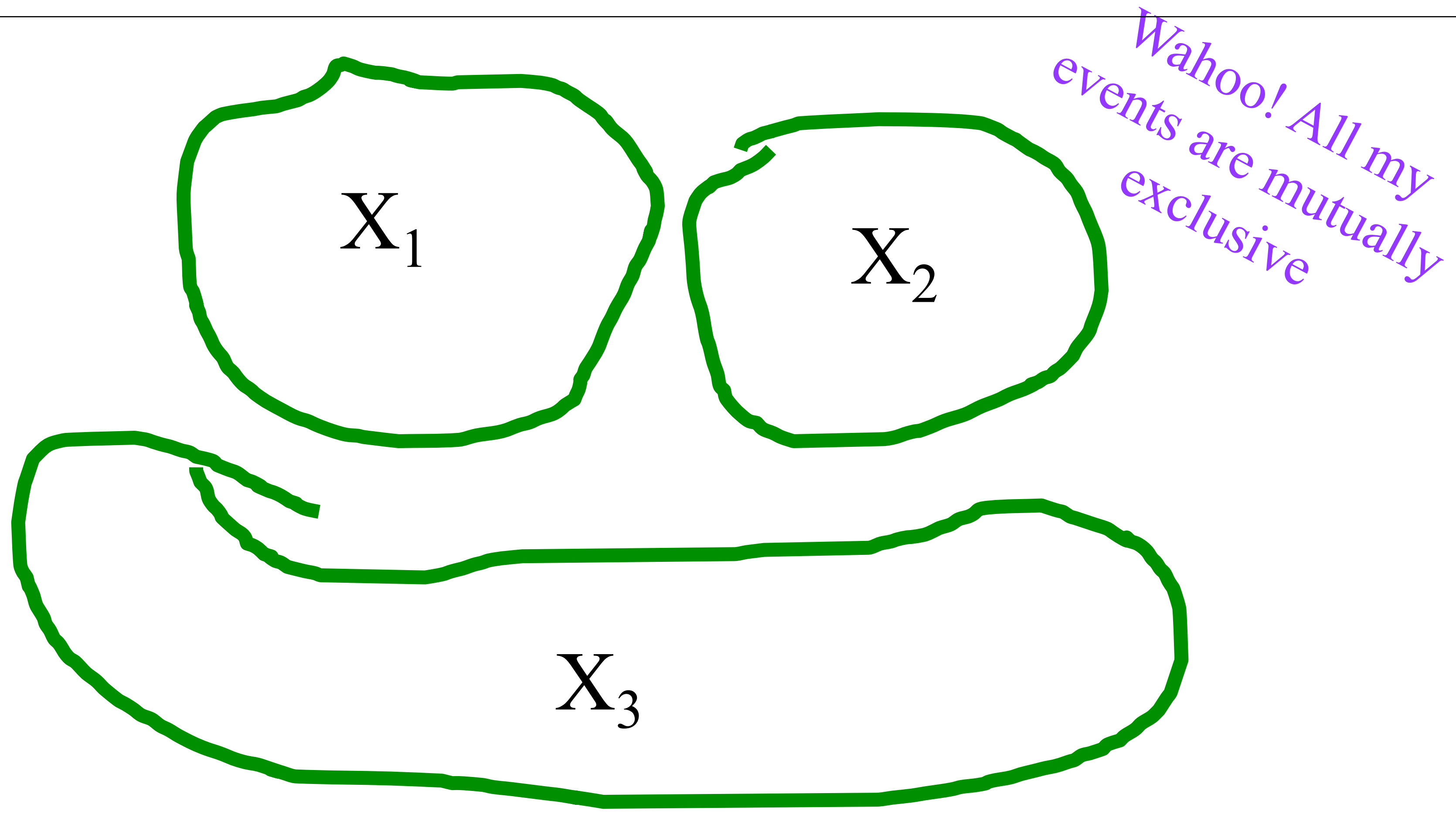
Mutually Exclusive Events



If events are mutually exclusive, probability of OR is simple:

$$P(E \cup F) = \frac{7}{50} + \frac{4}{5} = \frac{11}{50}$$

Probability of "or"



$$P(X_1 \cup X_2 \cup \cdots \cup X_n) = \sum_{i=1}^{n} P(X_i)$$

If events are *mutually exclusive* probability of OR is easy!

$$P(E^c) = 1 - P(E)?$$

$$P(E \cup E^c) = P(E) + P(E^c)$$

Axiom 3. Since E and E^c are mutually exclusive

$$P(S) = P(E) + P(E^c)$$

Since everything must either be in E or E^c

$$1 = P(E) + P(E^c)$$

Axiom 2

$$P(E^c) = 1 - P(E)$$

Rearrange

Trailing the dovetail shuffle to it’s lair – Persi Diaconosis

Trailing the Dovetail Shuffle to Its Lair

The Annals of Applied Probability
1992, Vol. 2, No. 2, 294–313

TRAILING THE DOVETAIL SHUFFLE TO ITS LAIR

BY DAVE BAYER[1] AND PERSI DIACONIS[2]

Columbia University and Harvard University

We analyze the most commonly used method for shuffling cards. The main result is a simple expression for the chance of any arrangement after any number of shuffles. This is used to give sharp bounds on the approach to randomness: $\frac{3}{2}\log_2 n + \theta$ shuffles are necessary and sufficient to mix up n cards.

Key ingredients are the analysis of a card trick and the determination of the idempotents of a natural commutative subalgebra in the symmetric group algebra.

1. Introduction. The dovetail, or riffle shuffle is the most commonly used method of shuffling cards. Roughly, a deck of cards is cut about in half and then the two halves are riffled together. Figure 1 gives an example of a riffle shuffle for a deck of 13 cards.

A mathematically precise model of shuffling was introduced by Gilbert and Shannon [see Gilbert (1955)] and independently by Reeds (1981). A deck of n cards is cut into two portions according to a binomial distribution; thus, the chance that k cards are cut off is $\binom{n}{k}/2^n$ for $0 \le k \le n$. The two packets are then riffled together in such a way that cards drop from the left or right heaps with probability proportional to the number of cards in each heap. Thus, if there are A and B cards remaining in the left and right heaps, then the chance that the next card will drop from the left heap is $A/(A+B)$. Such shuffles are easily described backwards: Each card has an equal and independent chance of being pulled back into the left or right heap. An inverse riffle shuffle is illustrated in Figure 2.

Experiments reported in Diaconis (1988) show that the Gilbert–Shannon–Reeds (GSR) model is a good description of the way real people shuffle real cards. It is natural to ask how many times a deck must be shuffled to mix it up. In Section 3 we prove:

THEOREM 1. *If n cards are shuffled m times, then the chance that the deck is in arrangement π is $\binom{2^m+n-r}{n}/2^{mn}$, where r is the number of rising sequences in π.*

Rising sequences are defined and illustrated in Section 2 through the analysis of a card trick. Section 3 develops several equivalent interpretations of

Received January 1990; revised May 1991.

[1]Partially supported by the Alfred P. Sloan Foundation, by ONR contract N00014-87-K0214 and by NSF Grant DMS-90-06116.

[2]Partially supported by NSF Grant DMS-89-05874.

AMS 1980 *subject classifications.* 20B30, 60B15, 60C05, 60F99.

Key words and phrases. Card shuffling, symmetric group algebra, total variation distance.

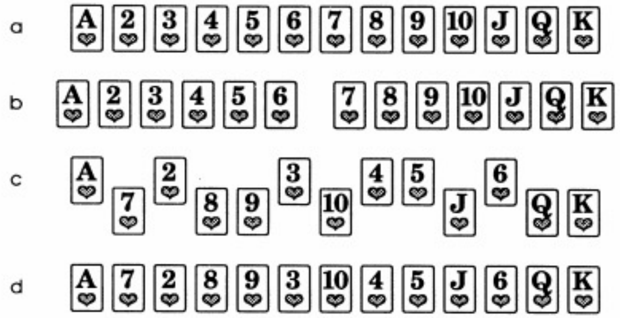


FIG. 1. *A riffle shuffle. (a) We begin with an ordered deck. (b) The deck is divided into two packets of similar size. (c) The two packets are riffled together. (d) The two packets can still be identified in the shuffled deck as two distinct "rising sequences" of face values.*

the GSR distribution for riffle shuffles, including a geometric description as the motion of n points dropped at random into the unit interval under the baker's transformation $x \to 2x \pmod 1$. This leads to a proof of Theorem 1.

Section 3 also relates shuffling to some developments in algebra. A permutation π has a descent at i if $\pi(i) > \pi(i+1)$. A permutation π has r rising sequences if and only if π^{-1} has $r-1$ descents. Let

$$A_k = \sum_{\pi \text{ has } k \text{ descents}} \pi$$

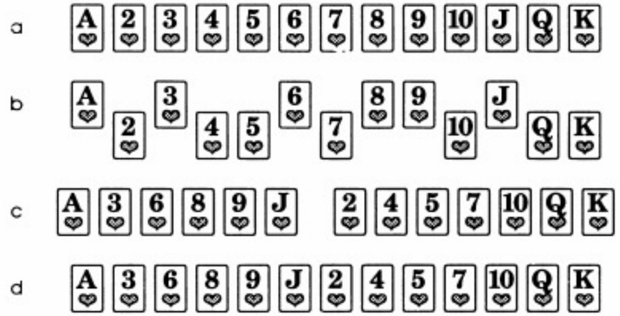


FIG. 2. *An inverse riffle shuffle. (a) We begin with a sorted deck. (b) Each card is moved one way or the other uniformly at random, to "pull apart" a riffle shuffle and retrieve two packets. (c) The two packets are placed in sequence. (d) The two packets can still be identified in the shuffled deck; they are separated by a "descent" in the face values. This shuffle is inverse to the shuffle diagrammed in Figure 1.*

Probability of "or"

- What is the probability that in the *n* shuffles seen since the start of time, yours is unique?
  - $|S| = \{52!\}^n$
  - $|E| = \{52! - 1\}^n$
  - P{no deck matching yours} = $\{52!-1\}^n/\{52!\}^n$

- For $n = 10^{20}$,
  - P{deck matching yours} < 0.000000001

* Assume 7 billion people have been shuffling cards once a second since cards were invented