03: Intro to Probability

Lisa Yan

April 10, 2020

Quick slide reference

Today's discussion thread: https://us.edstem.org/courses/109/discussion/24492

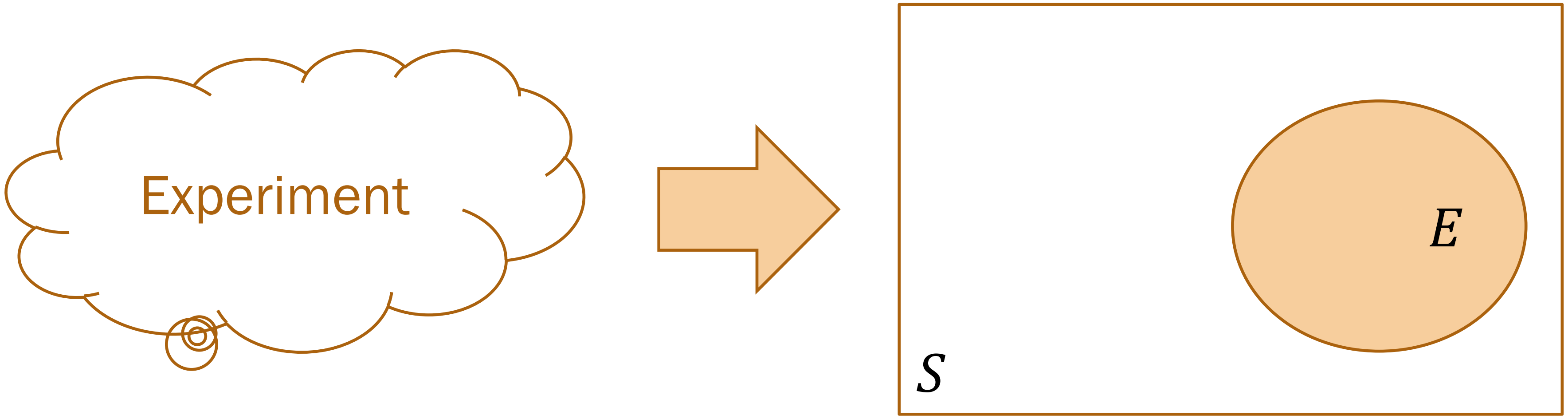
Defining Probability

Gradescope quiz, blank slide deck, etc.

http://cs109.stanford.edu/

Key definitions

An experiment in probability:



Sample Space, S: The set of all possible outcomes of an experiment

Event, E: Some subset of S ($E \subseteq S$).

Key definitions

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
  $E = \{\text{Heads}\}$
- $\geq$ 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 ($\leq$ 20 emails)
  $E = \{x \mid x \in \mathbb{Z}, \ 0 \leq x \leq 20\}$
- Wasted day ($\geq$ 5 TT hours):
  $E = \{x \mid x \in \mathbb{R}, 5 \leq x \leq 24\}$

What is a probability?

A number between 0 and 1
to which we ascribe meaning.*

*our belief that an event E occurs.

What is a probability?

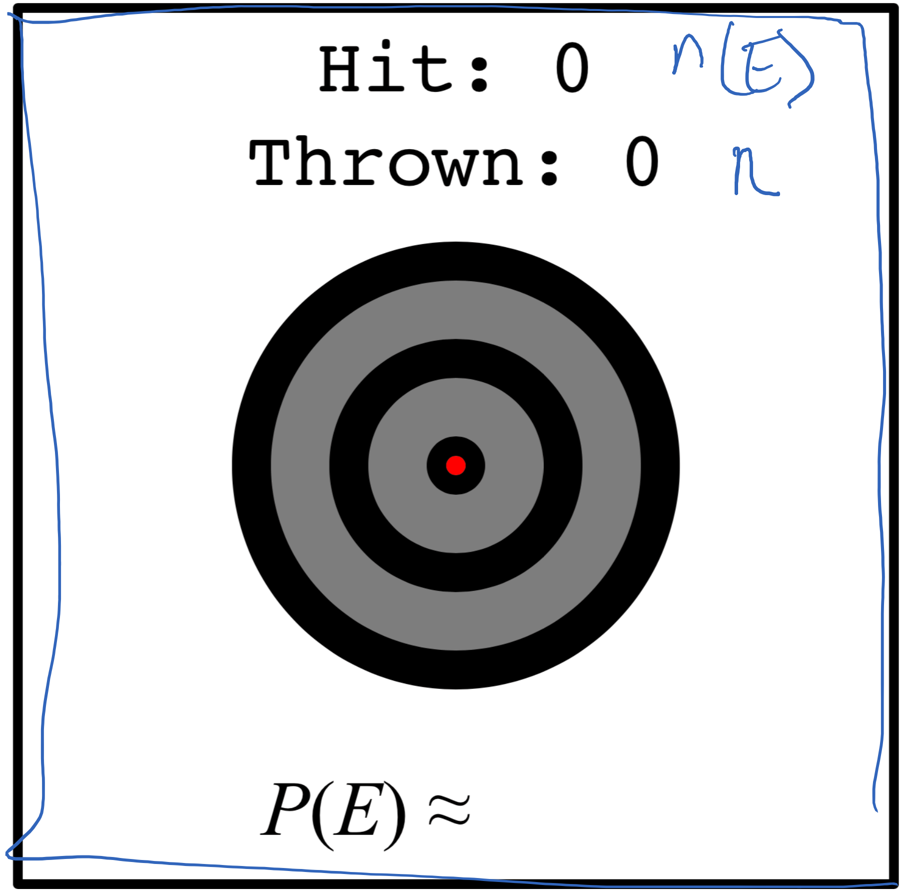
frequentist

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

n = # of total trials

$n(E)$ = # trials where E occurs

Let E = the set of outcomes where you hit the target.



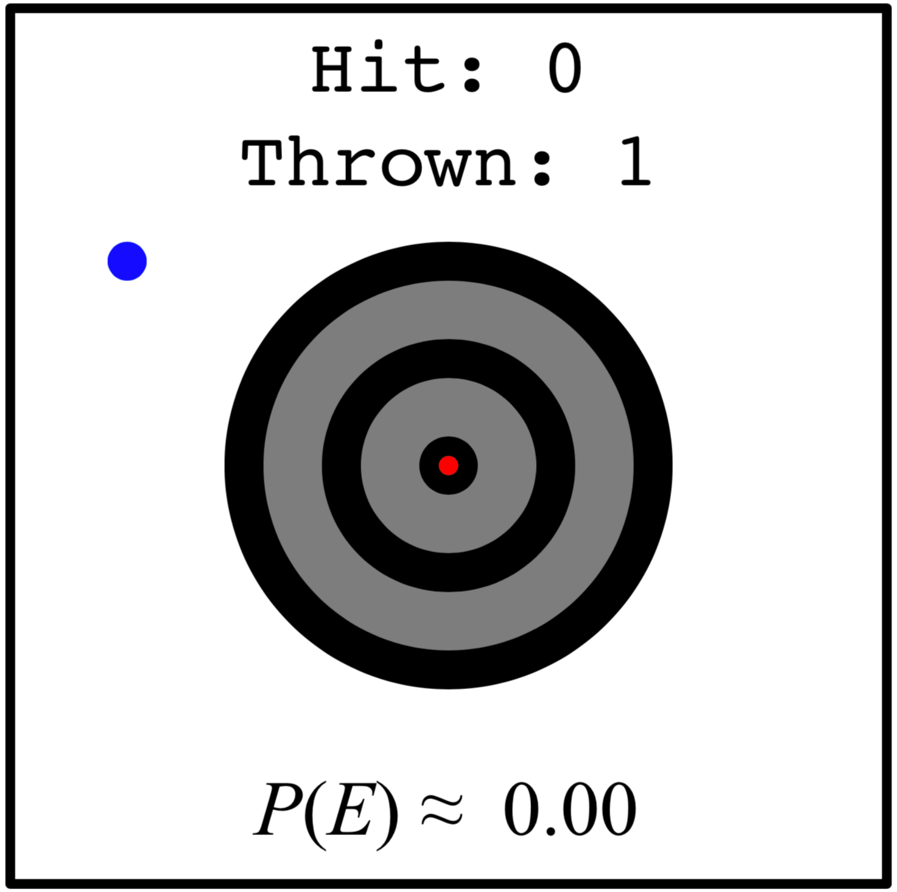
What is a probability?

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

n = # of total trials

$n(E)$ = # trials where E occurs

Let E = the set of outcomes where you hit the target.



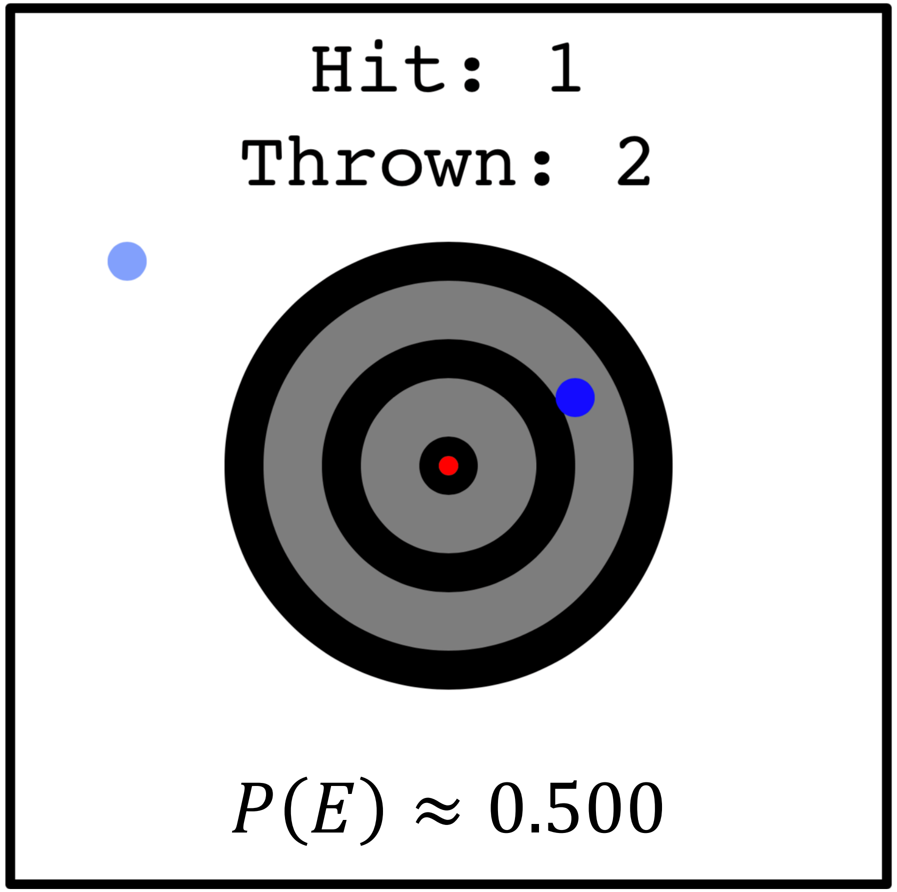
What is a probability?

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

n = # of total trials

$n(E)$ = # trials where E occurs

Let E = the set of outcomes where you hit the target.



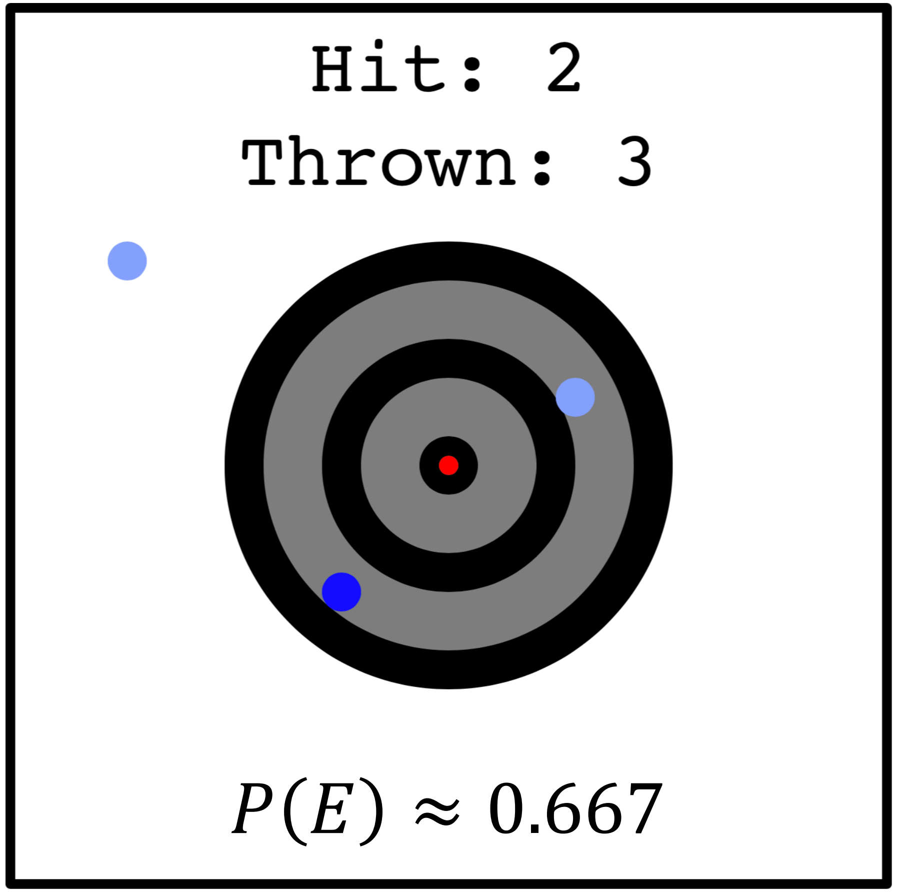
What is a probability?

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

n = # of total trials

$n(E)$ = # trials where E occurs

Let E = the set of outcomes where you hit the target.



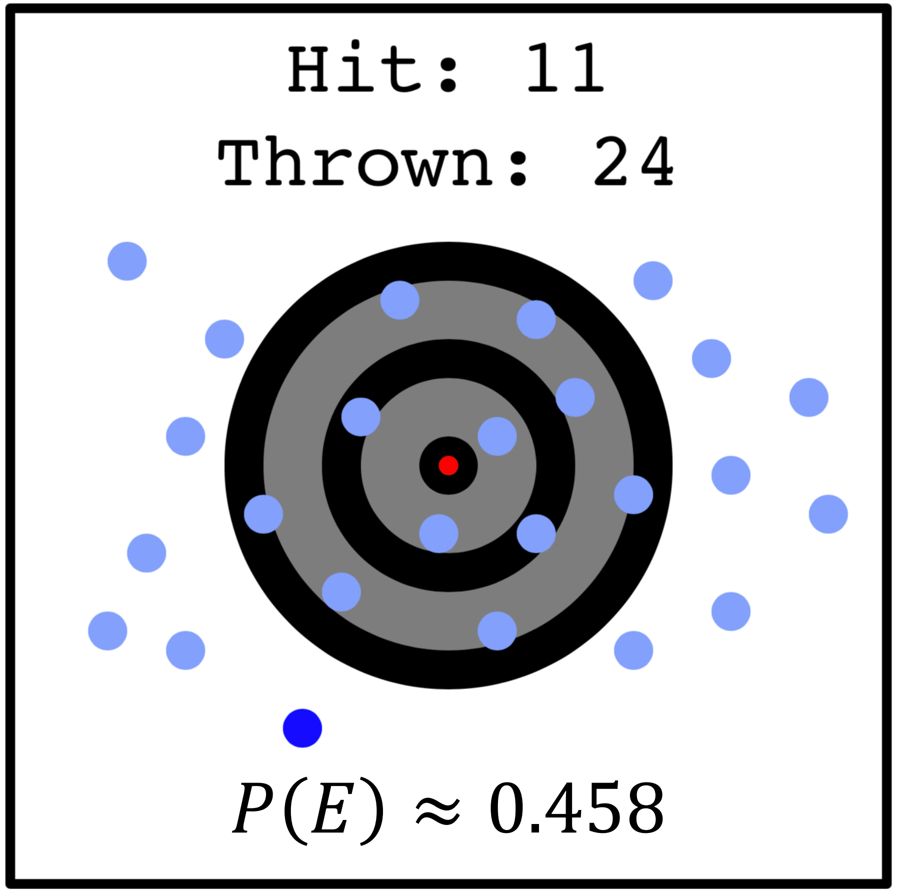
What is a probability?

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

n = # of total trials

$n(E)$ = # trials where E occurs

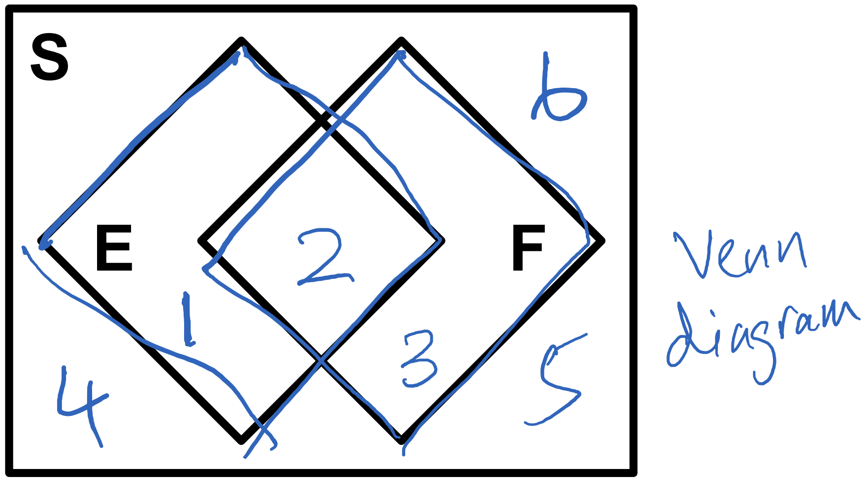
Let E = the set of outcomes where you hit the target.



Not just yet...

Axioms of Probability

Quick review of sets



E and F are events in S.

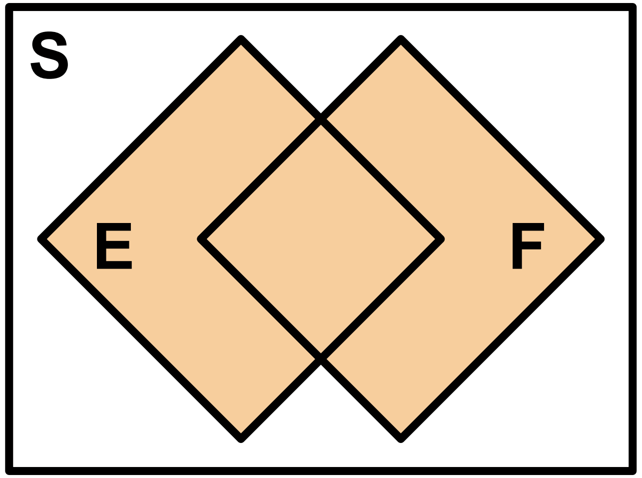
Experiment:

Die roll

$S = \{1, 2, 3, 4, 5, 6\}$

Let $E = \{1, 2\}$, and $F = \{2,3\}$

Quick review of sets



E and F are events in S.

Experiment:

Die roll

$S = \{1, 2, 3, 4, 5, 6\}$

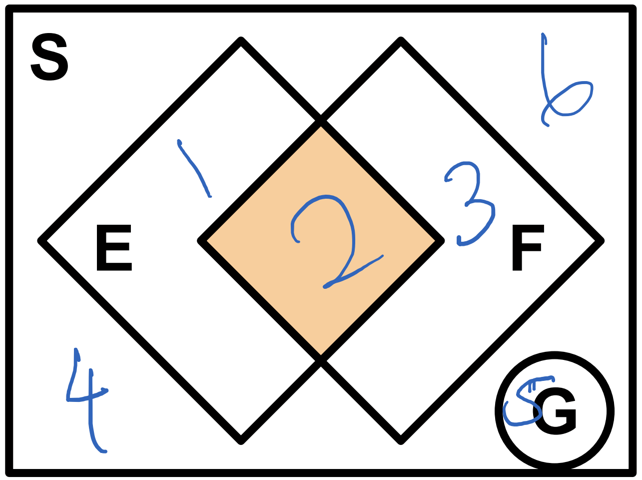
Let $E = \{1, 2\}$, and $F = \{2,3\}$

def Union of events, $E \cup F$ ↑ \cup

The event containing all outcomes in E or F.

$$E \cup F = \{1,2,3\}$$

Quick review of sets



E and F are events in S.

Experiment:

Die roll

$S = \{1, 2, 3, 4, 5, 6\}$

Let $E = \{1, 2\}$, and $F = \{2, 3\}$

<u>def</u> Intersection of events, $E \cap F$

The event containing all outcomes in E <u>*and*</u> F.

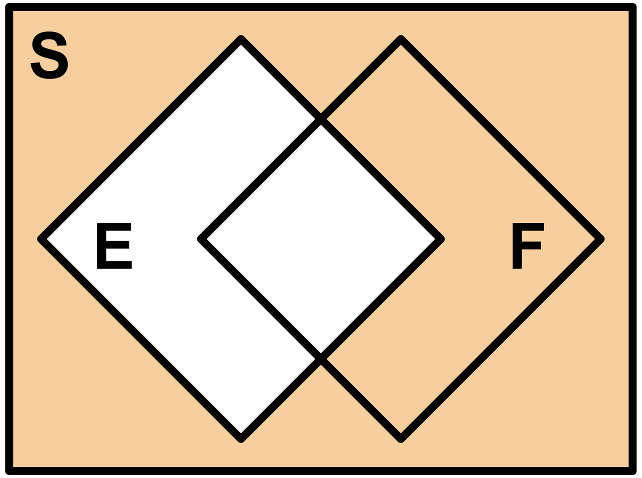
$$E \cap F = EF = \{2\}$$

\cap

<u>def</u> Mutually exclusive events F and G means that $F \cap G = \emptyset$

$G = \{5\}$

Quick review of sets



E and F are events in S.

Experiment:

Die roll

$S = \{1, 2, 3, 4, 5, 6\}$

Let $E = \{1, 2\}$, and $F = \{2,3\}$

<u>def</u> Complement of event E, E^C

The event containing all outcomes in that are ***not*** in E.

$$E^C = \{3, 4, 5, 6\}$$

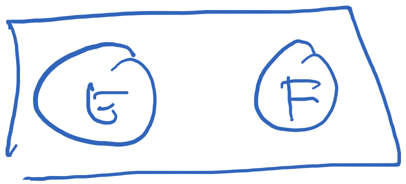
3 Axioms of Probability

Definition of probability: $P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$

Axiom 1: $0 \leq P(E) \leq 1$

Axiom 2: $P(S) = 1$

Axiom 3: If E and F are mutually exclusive ($E \cap F = \emptyset$), then $P(E \cup F) = P(E) + P(F)$

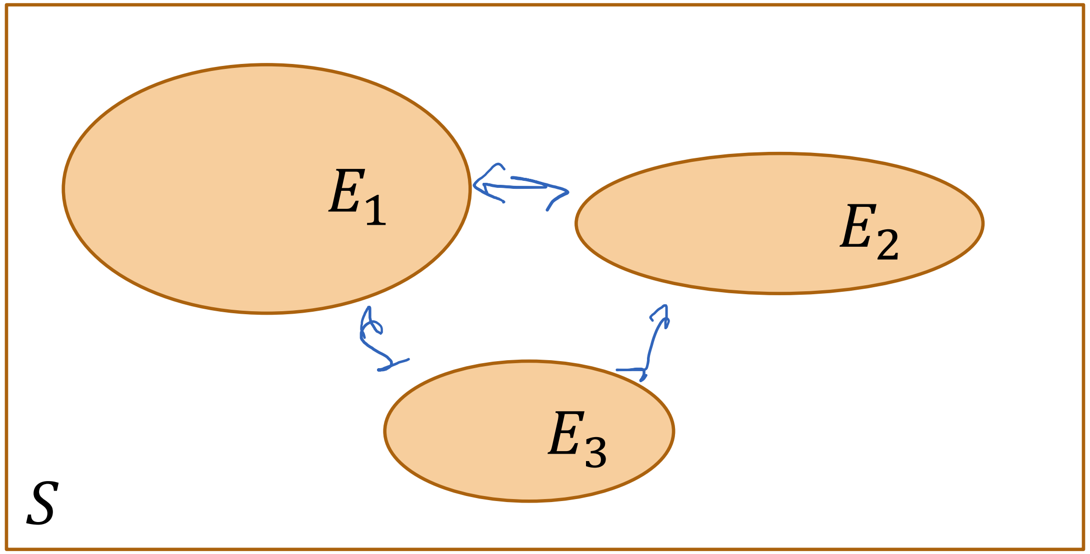


Axiom 3 is the (analytically) useful Axiom

Axiom 3: If E and F are mutually exclusive ($E \cap F = \emptyset$), then $P(E \cup F) = P(E) + P(F)$

More generally, for any sequence of mutually exclusive events $E_1, E_2, \ldots$:

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i)$$



$P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3)$

(like the Sum Rule of Counting, but for probabilities)

Equally Likely Outcomes

Equally Likely Outcomes

Some sample spaces have **equally likely outcomes**.

- fair Coin flip: S = {Head, Tails} $P(\text{Heads}) = 1/2$
- Flipping two fair coins: S = {(H, H), (H, T), (T, H), (T, T)} $P(\{(H,H)\}) = 1/4$
- Roll of fair 6-sided die: S = {1, 2, 3, 4, 5, 6} $P(\{3\}) = 1/6$

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

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

E: 3 or lower

$E = \{1, 2, 3\}$

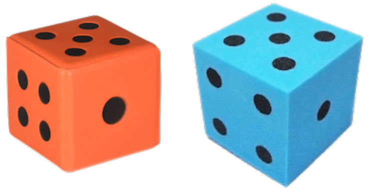
$E_1 = \{1\}, E_2 = \{2\}, E_3 = \{3\}, P(E_i) = \frac{1}{|S|}$

$$P(E) = P(E_1 \cup E_2 \cup E_3) = P(E_1) + P(E_2) + P(E_3) = 3 \cdot \frac{1}{|S|} = |E| \cdot \frac{1}{|S|} = \frac{|E|}{|S|}$$

Roll two dice

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

Roll two 6-sided fair 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)}

|S| = 36

$E =$ {(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)}

Sum is 7

|E| = 6

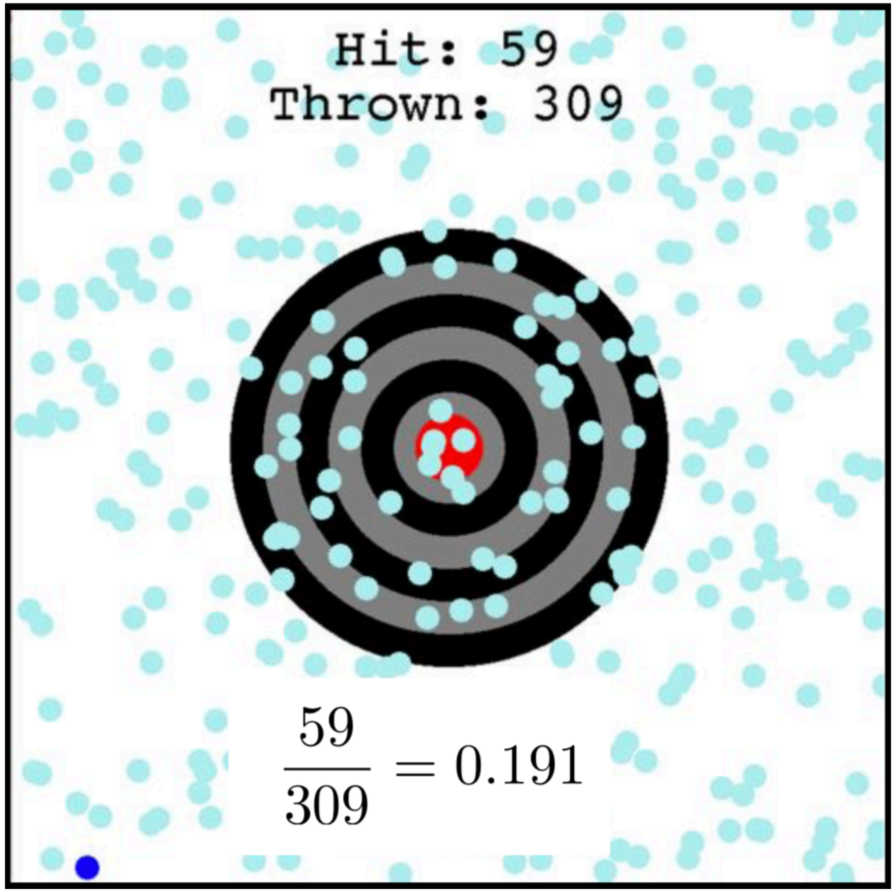
P(E) = 6/36 = 1/6

Target revisited

Target revisited

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

Let E = the set of outcomes where you hit the target.



Screen size = 800 ×800

Radius of target: 200

The dart is equally likely to land anywhere on the screen. What is $P(E)$, the probability of hitting the target?

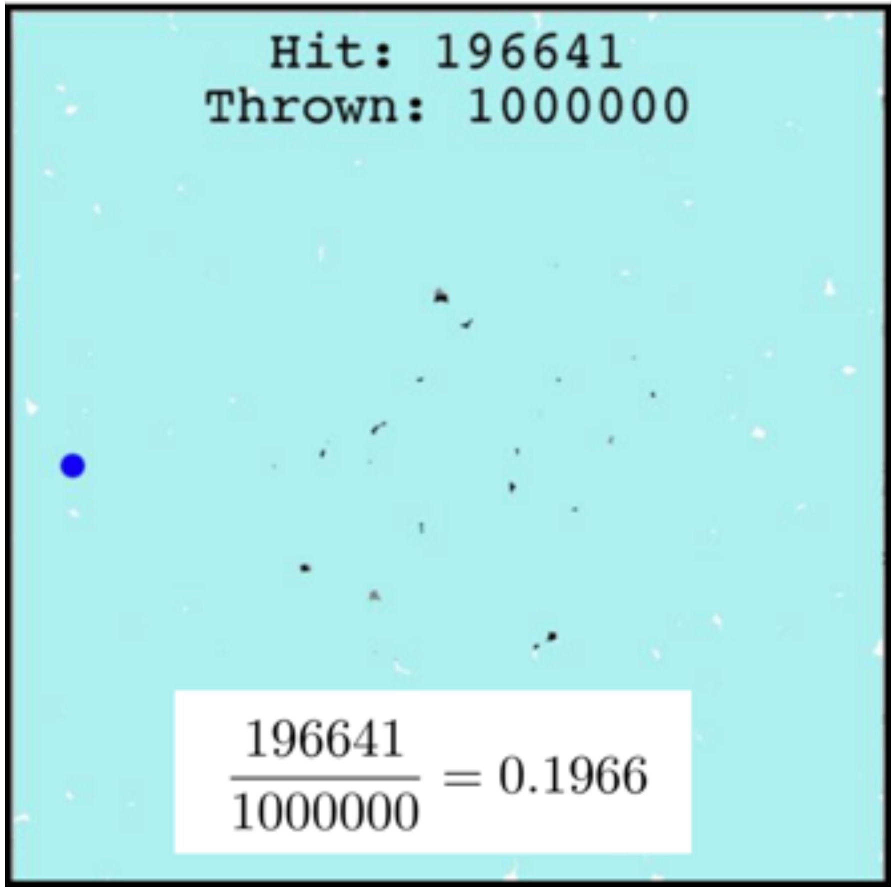
$|S| = 800^2$ $\quad\quad$ $|E| \approx \pi \cdot 200^2$

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

Target revisited

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

Let E = the set of outcomes where you hit the target.



Screen size = 800 ×800

Radius of target: 200

The dart is equally likely to land anywhere on the screen. What is $P(E)$, the probability of hitting the target?

$$|S| = 800^2 \qquad |E| \approx \pi \cdot 200^2$$

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

Not equally likely outcomes

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

Play the lottery.

What is $P(\text{win})$?

$S = \{\text{Lose}, \text{Win}\}$

$E = \{\text{Win}\}$

$$P(E) = \frac{|E|}{|S|} = \frac{1}{2} = 50\%?$$

41,416,355 tickets sold

1 winning

The hard part: defining outcomes consistently across sample space and events

$P(\text{win}) = 1/41416355$

Cats and sharks

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

4 cats and 3 sharks in a bag. 3 drawn.

What is P(1 cat and 2 sharks drawn)?

Make indistinct items distinct to get equally likely outcomes.

Note: Do indistinct objects give you an equally likely sample space?

CCC
CCS
CSS
SSS

(No)

A. $\frac{3}{7}$

B. $\frac{1}{4} \cdot \frac{2}{3}$

C. $\frac{4}{7} + 2 \cdot \frac{3}{6}$

D. $\frac{12}{35}$

E. Zero/other

Cats and sharks (ordered solution)

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

4 cats and 3 sharks in a bag. 3 drawn.

What is P(1 cat and 2 sharks drawn)?

Make indistinct items distinct to get equally likely outcomes.

Define

- S = Pick 3 distinct items
- E = 1 distinct cat, 2 distinct sharks

$7 \cdot 6 \cdot 5$ $\quad |S| = 210$

pick C first: $\underset{c}{4} \cdot \underset{s}{3} \cdot \underset{s}{2}$

pick C second $\underset{s}{3} \cdot \underset{c}{4} \cdot \underset{s}{2}$

pick C third $\underset{s}{3} \cdot \underset{s}{2} \cdot \underset{c}{4}$

$+$

$|E| = 72$

$P(E) = \frac{12}{35} = \frac{72}{210}$

Cats and sharks (unordered solution)

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

4 cats and 3 sharks in a bag. 3 drawn.

What is P(1 cat and 2 sharks drawn)?

Make indistinct items distinct to get equally likely outcomes.

Define

- S = Pick 3 distinct items

$|S| = \binom{7}{3} = \frac{7!}{3!4!} = 35$

- E = 1 distinct cat, 2 distinct sharks

$|E| = \binom{4}{1}\binom{3}{2} = 4 \cdot 3 = 12$ $P(E) = \boxed{\frac{12}{35}}$

Corollaries of Probability

Axioms of Probability

Review

Definition of probability: $P(E) = \lim_{n \to \infty} \frac{n(E)}{n}$

Axiom 1: $0 \leq P(E) \leq 1$

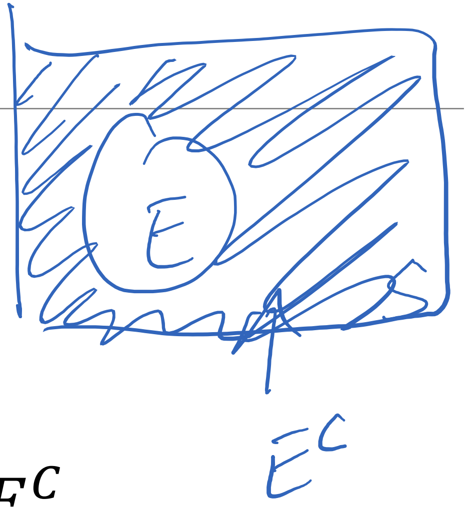
Axiom 2: $P(S) = 1$

Axiom 3: If E and F are mutually exclusive ($E \cap F = \emptyset$), then $P(E \cup F) = P(E) + P(F)$

3 Corollaries of Axioms of Probability

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

Proof of Corollary 1



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

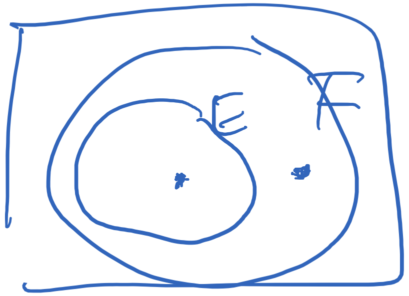
Proof:

E, E^C are mutually exclusive	Definition of E^C
$P(E \cup E^C) = P(E) + P(E^C)$	Axiom 3
$S = E \cup E^C$	Everything must either be in E or E^C, by definition
$1 = P(S) = P(E) + P(E^C)$	Axiom 2
$P(E^C) = 1 - P(E)$	Rearrange

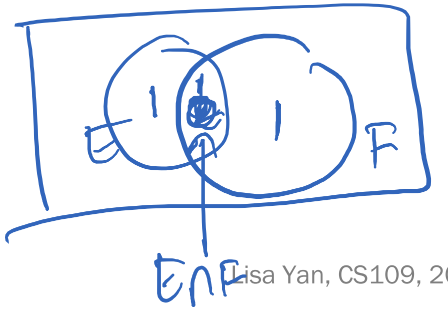
3 Corollaries of Axioms of Probability

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

Corollary 2: If $E \subseteq F$, then $P(E) \leq P(F)$



Corollary 3: $P(E \cup F) = P(E) + P(F) - P(EF)$
(Inclusion-Exclusion Principle for Probability)



Selecting Programmers

- P(student programs in Java) = 0.28 $= P(E)$
- P(student programs in Python) = 0.07 $= P(F)$
- P(student programs in Java and Python) = 0.05. $= P(E \cap F) = P(EF)$

What is P(student does not program in (Java or Python))?

1. Define events & state goal
2. Identify known probabilities
3. Solve

E: Java

F: Python

$P((E \cup F)^C)$ ←

(E: Java, F: Python)

Corollary 1: $P((E \cup F)^C) = 1 - P(E \cup F)$

Corollary 3: $P(E \cup F) = P(E) + P(F) - P(EF)$

$= 0.28 + 0.07 - 0.05$

$= 0.3$

$P((E \cup F)^C) = \boxed{0.7}$

Inclusion-Exclusion Principle (Corollary 3)

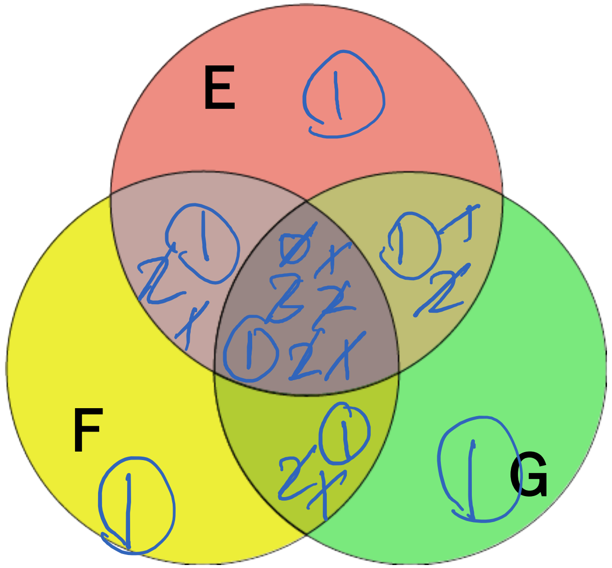
Corollary 3: $P(E \cup F) = P(E) + P(F) - P(EF)$

(Inclusion-Exclusion Principle for Probability)

General form:

$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_{r=1}^{n} (-1)^{(r+1)} \sum_{i_1 < \cdots < i_r} P\left(\bigcap_{j=1}^{r} E_{i_j}\right)$$

# of sets in intersection



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

r = 1: $P(E) + P(F) + P(G)$

r = 2: $- P(E \cap F) - P(E \cap G) - P(F \cap G)$

r = 3: $+ P(E \cap F \cap G)$

(live)
03: Intro to Probability

Lisa Yan

April 10, 2020

Reminders: Lecture with zoom

- Turn on your camera if you are able, mute your mic in the big room
- Virtual backgrounds are encouraged (classroom-appropriate)

Breakout Rooms for meeting your classmates

- ~~Just like sitting next to someone new~~ Our best approximation to sitting next to someone new

We will use Ed instead of Zoom chat

- Lots of activity and questions, thank you all!

Today's discussion thread: https://us.edstem.org/courses/109/discussion/24492

Holy crap, are all of the pre-lecture videos going to be this long??

(1) This course is packed to the brim with content, and the early half is definitely definition-heavy.
(2) Our videos will get closer and closer to the cumulative estimated 30 minutes as we get better at recording ☺

dang these breakout rooms are awkward

We know this cannot compare to an in-person discussion! Hopefully these will become smoother once we all adjust to the online format.

123
SESAME STREET

The Count

Chance The Rapper

Summary so far

Review

$P(E) = \frac{|E|}{|S|}$ Equally likely outcomes

Counting tasks on n objects

- Ordered — **Sort objects (permutations)**
  - Distinct (distinguishable): $n!$
  - Some distinct: $\frac{n!}{n_1!\, n_2! \cdots n_r!}$
- Unordered — **Choose k objects (combinations)**
  - Distinct
    - 1 group: $\binom{n}{k}$
    - r groups: $\binom{n}{n_1, n_2, \cdots, n_r}$
- **Put objects in r buckets**
  - Distinct: r^n
  - Indistinct: $\frac{(n+r-1)!}{n!\,(r-1)!}$

Indistinguishable? Distinguishable? Probability?

We choose 3 books from a set of 4 distinct (distinguishable) and 2 indistinct (indistinguishable) books.

6 books

Let event E = our choice does not include both indistinct books.

1. What is $|E|$?

   more likely! → one indistinct: $1 \cdot \binom{4}{2} = \frac{\binom{2}{1}\binom{4}{2}}{2} = 6$

   ~~neither~~ no indistinct: $\binom{4}{3} = 4$

   $|E| = 10$

2. What is $P(E)$?

   $S_{distinct}$: $|S_{dist}| = \binom{6}{3} = 20$

   $E_{distinct}$: $|E_{dist}| = \binom{2}{1}\binom{4}{2} + \binom{4}{3} = 12 + 4 = 16$

   $P(E) = \frac{|E_{dist}|}{|S_{dist}|} = \frac{16}{20} = 0.8$

distinguishable, equally likely outcomes

make indistinct → report count (unique outcomes)

keep distinct → compute probability (ELO)

Think, then Breakout Rooms

Then check out the question on the next slide (Slide 44). Post any clarifications here!

https://us.edstem.org/courses/109/discussion/24492

Think by yourself: 2 min

Breakout rooms: 5 min. Introduce yourself!

Poker Straights and Computer Chips

1. Consider 5-card poker hands.
   - "straight" is 5 consecutive rank cards of any suit

   What is P(Poker straight)?

- What is an example of an outcome?
- Is each outcome equally likely?
- Should objects be ordered or unordered?

2. Consider the "official" definition of a Poker Straight:
   - "straight" is 5 consecutive rank cards of any suit
   - straight flush" is 5 consecutive rank cards of same suit

   What is P(Poker straight, but not straight flush)?

3. Computer chips: n chips are manufactured, 1 of which is defective. k chips are randomly selected from n for testing.

   What is P(defective chip is in k selected chips?)

Any Poker Straight

1. Consider 5-card poker hands.
   - "straight" is 5 consecutive rank cards of any suit

   What is P(Poker straight)?

A 2 3 4 5
2 3 4 5 6
3 4 5 6 7
⋮
9 10 J Q K
10 J Q K A

Define
- S (unordered) $|S| = \binom{52}{5}$
- E (unordered, consistent with S) $|E| = 10 \cdot \binom{4}{1}^5$

Compute $P(\text{Poker straight}) =$ $P(E) = \dfrac{10 \cdot \binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$

"Official" Poker Straight

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(Poker straight, but not straight flush)?

Define

- S (unordered) $|S| = \binom{52}{5}$
- E (unordered, consistent with S) $|E| = 10 \cdot \binom{4}{1}^5 - 10 \cdot \binom{4}{1}$

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

Compute P(Official Poker straight) =

Chip defect detection

n chips are manufactured, 1 of which is defective.
k chips are randomly selected from n for testing.

What is P(defective chip is in k selected chips?)

Define

- S (unordered)
- E (unordered, consistent with S)

$$|S| = \binom{n}{k}$$

$$|E| = \binom{1}{1} \cdot \binom{n-1}{k-1}$$

defective non-defective

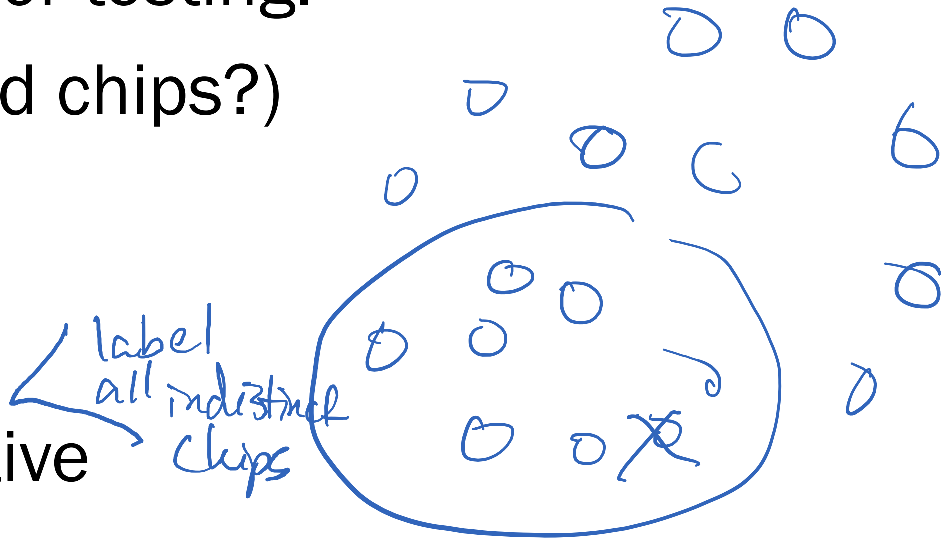
Compute

$$P(E) = \frac{|E|}{|S|} = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\frac{(n-1)!}{(k-1)!\,(n-k)!}}{\frac{n!}{k!\,(n-k)!}} = \frac{k}{n}$$

Chip defect detection, solution #2

n chips are manufactured, 1 of which is defective.
k chips are randomly selected from n for testing.

What is P(defective chip is in k selected chips?)



Redefine experiment

1. Choose k indistinct chips (1 way)
2. Throw a dart and make one defective

Define

- S (unordered)
- E (unordered, consistent with S)

$$\frac{k}{n}$$

Interlude for jokes/announcements

Announcements

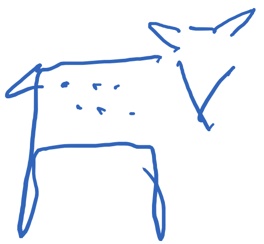
Section sign-ups

Preference form: ~~later today~~ out yesterday

Due: Saturday 4/11

Results: latest Monday

Please fill this out even if you don't plan to attend section this quarter.

Joke 

no eye deer

Handout: Calculation Reference

http://web.stanford.edu/class/cs109/handouts/H02_calculation_ref.pdf

Geometric series:

$$\sum_{i=0}^{n} x^i = \frac{1-x^{n+1}}{1-x}$$

$$\sum_{i=m}^{n} x^i = \frac{x^{n+1}-x^m}{x-1}$$

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \text{ if } |x| < 1$$

Integration by parts (everyone's favorite!):

Choose a suitable u and dv to decompose the integral of interest:

$$\int u \cdot dv = u \cdot v - \int v \cdot du$$

Interesting probability news

"The study finds that very few chords govern most of the music, a phenomenon that is also known in linguistics, where very few words dominate language corpora.... It characterizes Beethoven's specific composition style for the String Quartets, through a distribution of all the chords he used, how often they occur, and how they commonly transition from one to the other."

Find something cool, submit for extra credit on Problem Set #1 ☺

https://actu.epfl.ch/news/decoding-beethoven-s-music-style-using-data-scienc/

3 Corollaries of Axioms of Probability

Review

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

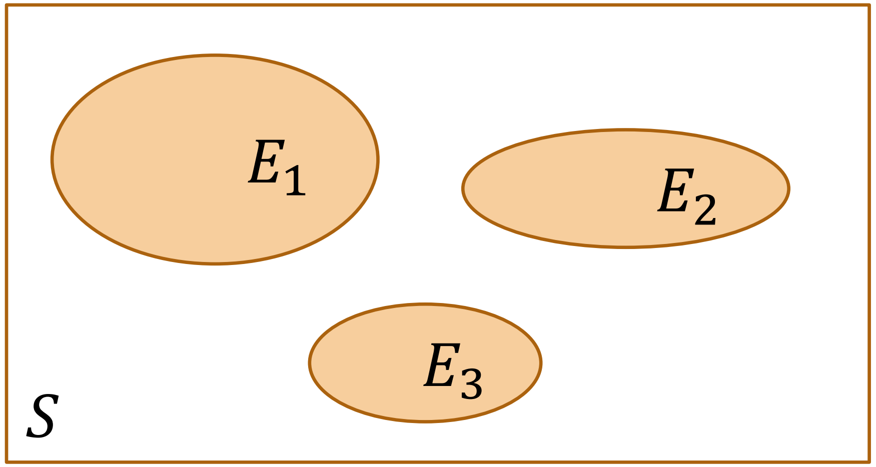
Corollary 2: If $E \subseteq F$, then $P(E) \leq P(F)$

Corollary 3: $P(E \cup F) = P(E) + P(F) - P(EF)$
(Inclusion-Exclusion Principle for Probability)

Takeaway: Mutually exclusive events

Review

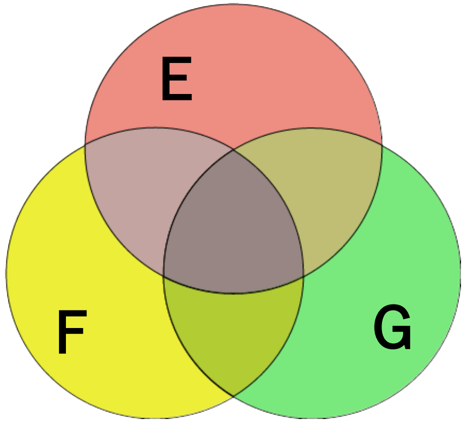
Axiom 3,
Mutually exclusive events



$$P\left(\bigcup_{i=1}^{\infty} E_i\right) =$$

$P(E_1) + P(E_2) + P(E_3) + \ldots$

Inclusion-Exclusion Principle



$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = P(E_1) + P(E_2) + P(E_3) + \ldots$$

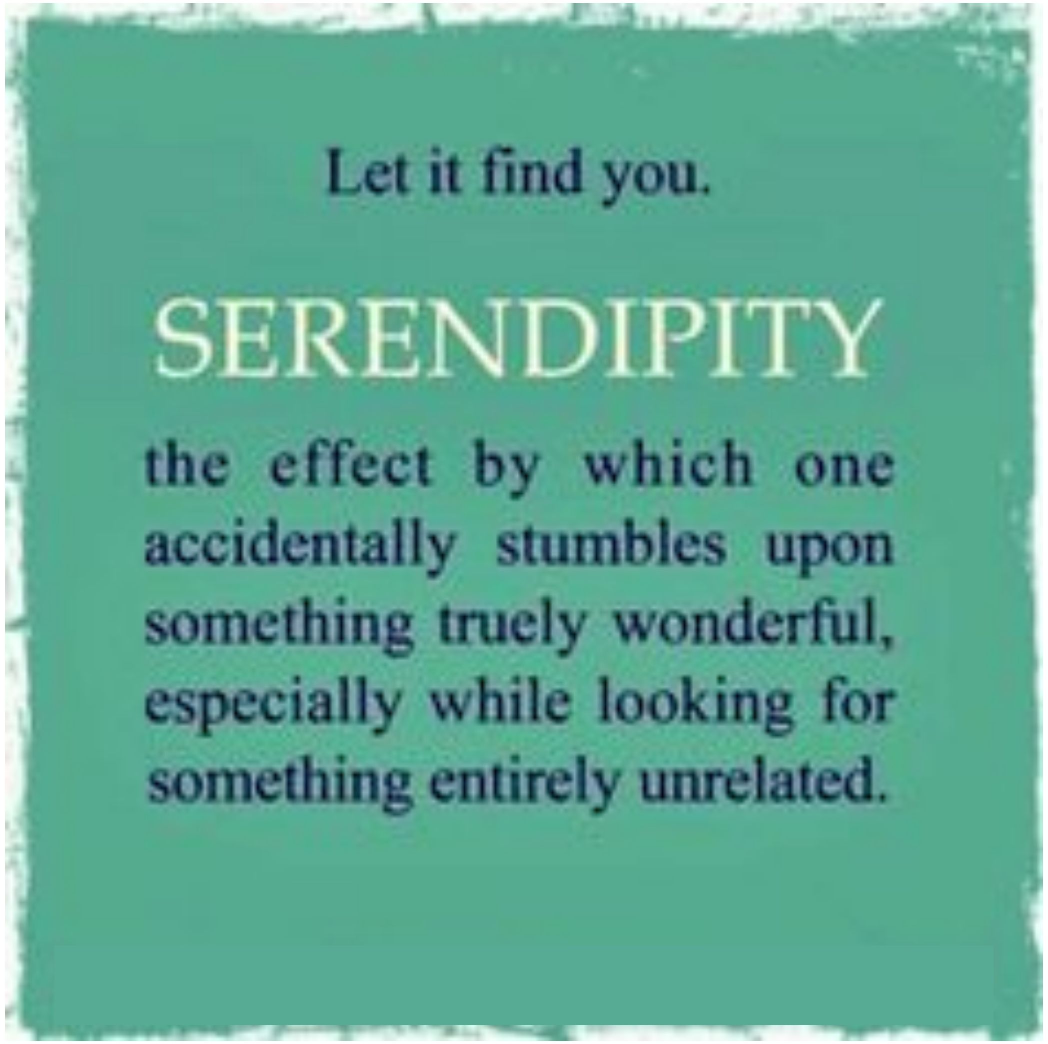
$- P(E_1 \cap E_2) - P(E_1 \cap E_3) - \ldots - P(E_2 \cap E_3) - \ldots$

$+ P(E_1 \cap E_2 \cap E_3) + P(E_1 \cap E_2 \cap E_4) + \ldots$

$-$

The challenge of probability is in defining events.
Some event probabilities are easier to compute than others.

Serendipity



Serendipity

- The population of Stanford is $n = 17{,}000$ people.
- You are friends with $r =$ people.
- Walk into a room, see $k = 360$ random people.
- Assume you are equally likely to see each person at Stanford.

What is the probability that you see someone you know in the room?

http://web.stanford.edu/class/cs109/demos/serendipity.html

Breakout Rooms

Check out the question on the next slide (Slide 57). Post any clarifications here!

https://us.edstem.org/courses/109/discussion/24492

Breakout rooms: 5 min. Introduce yourself if you haven't yet!

Serendipity

- The population of Stanford is $n = 17{,}000$ people.
- You are friends with $r = 100$ people.
- Walk into a room, see $k = 360$ random people.
- Assume you are equally likely to see each person at Stanford.

What is the probability that you see someone you know in the room?

≥ 1 friend?

Define

- S (unordered)
- E: ≥ 1 friend in the room

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

What strategy should you use?

A. $P(\text{exactly } 1) + P(\text{exactly } 2)$ $P(\text{exactly } 3) + \cdots$

$\binom{100}{1}\binom{\ }{99}$

B. $1 - P(\text{see no friends})$

Serendipity

- The population of Stanford is $n = 17{,}000$ people.
- You are friends with $r = 100$ people.
- Walk into a room, see $k = 360$ random people.
- Assume you are equally likely to see each person at Stanford.

What is the probability that you see someone you know in the room?

Define

- S (unordered)
- E: ≥ 1 friend in the room

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

$$P(E^c) = \frac{\binom{16900}{360}}{\binom{17000}{360}}$$

$$P(E) = 1 - \frac{\binom{16900}{360}}{\binom{17000}{360}}$$

It is often much easier to compute $P(E^c)$.

The Birthday ~~Paradox~~ Problem

What is the probability that in a set of *n* people, <u>at least one</u> pair of them will share the same birthday?

For you to think about (and discuss in section!)

Card Flipping

In a 52 card deck, cards are flipped one at a time.
After the first ace (of any suit) appears, consider the next card.

Is P(next card = Ace Spades) < P(next card = 2 Clubs)?

Once you think you have an answer, you can vote on pollEverywhere:

http://www.pollev.com/cs109

https://us.edstem.org/courses/109/discussion/24492

Check out Lectures Notes!

(by yourself)

In a 52 card deck, cards are flipped one at a time. After the first ace (of any suit) appears, consider the next card.

P(next card = Ace Spaces)
< P (next card = 2 Clubs)

P(next card = Ace Spaces)
> P (next card = 2 Clubs)

P(next card = Ace Spaces)
= P (next card = 2 Clubs)