03: Intro to Probability

Jerry Cain
April 1, 2022

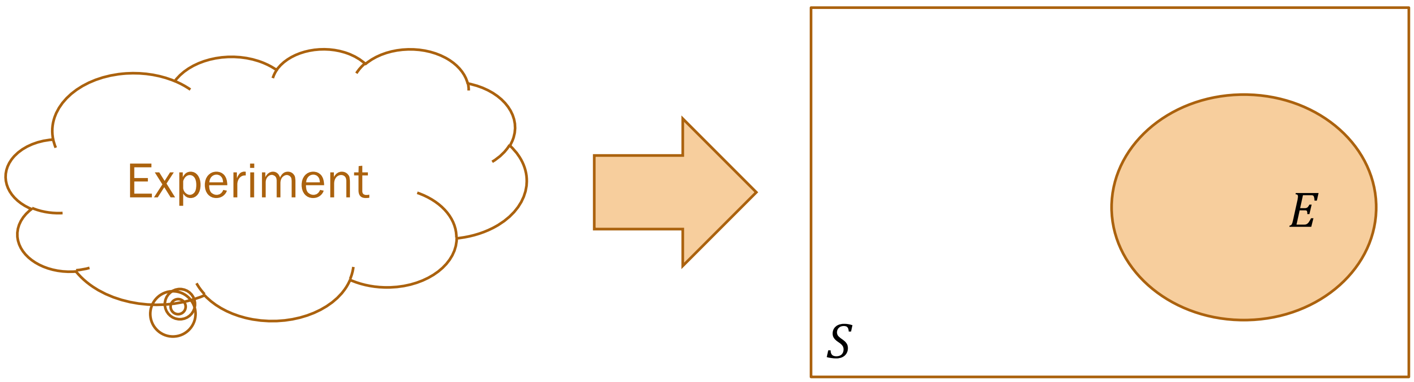
Quick slide reference

Today's discussion thread: https://edstem.org/us/courses/21301/discussion/1324046

Defining Probability

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
  $E = \{\text{Heads}\}$
- ≥ 1 head in two 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?

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

*our belief that an event E occurs.

What is a probability?

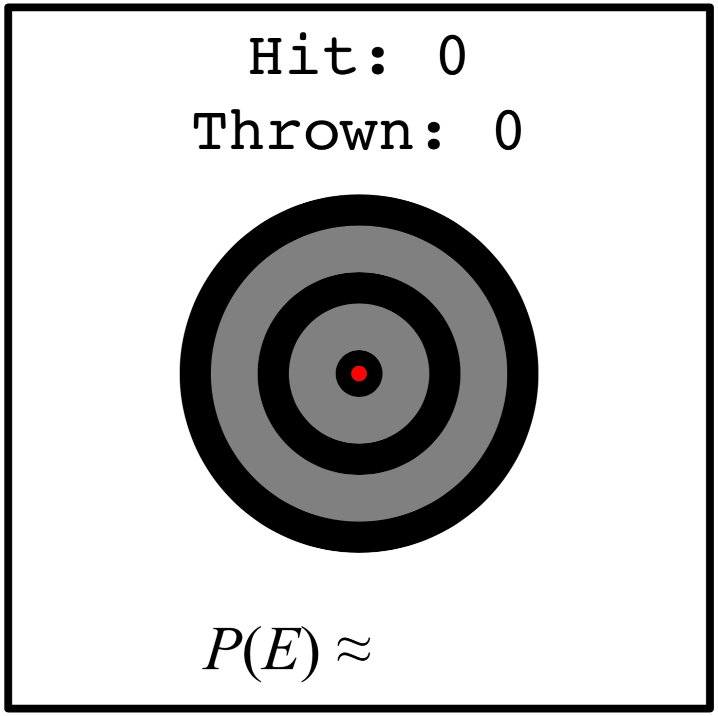
frequentist

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

n = # of total trials

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

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



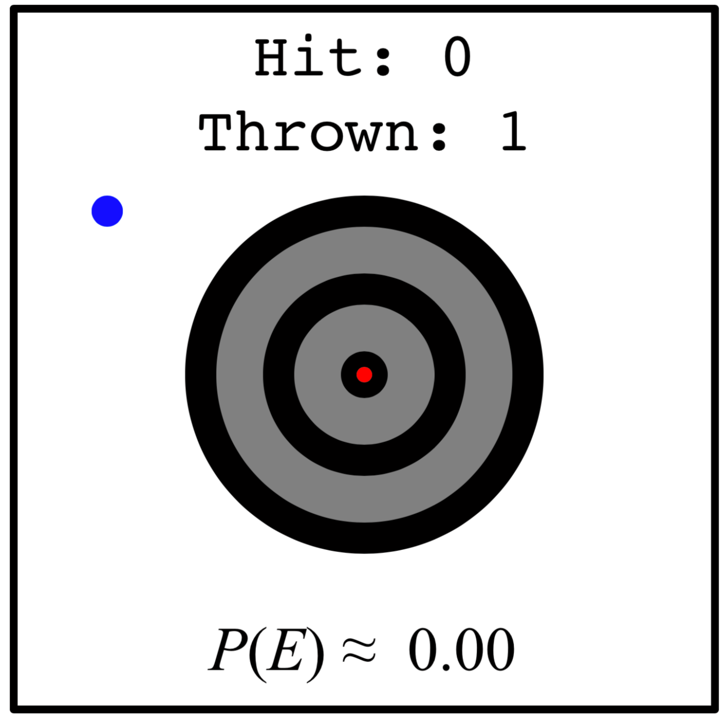
What is a probability?

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

n = # of total trials

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

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



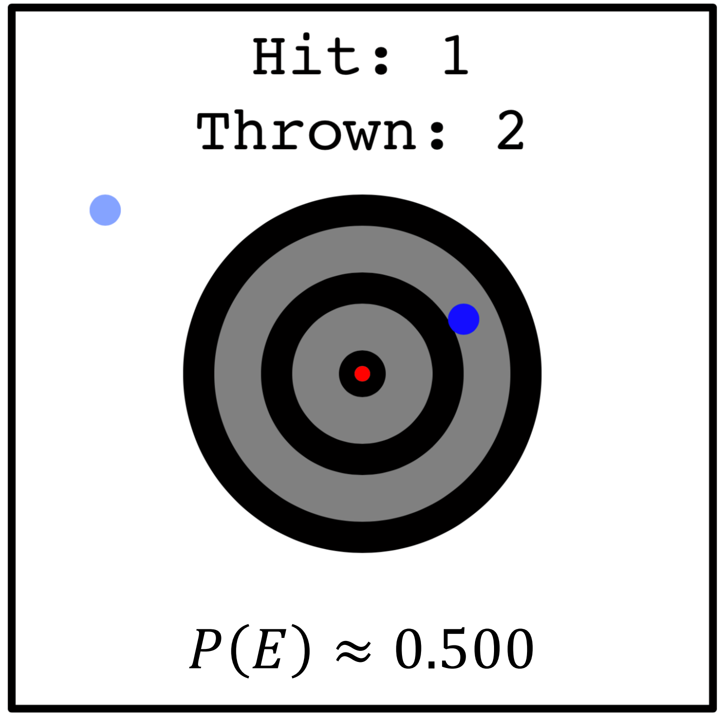
What is a probability?

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

n = # of total trials

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

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



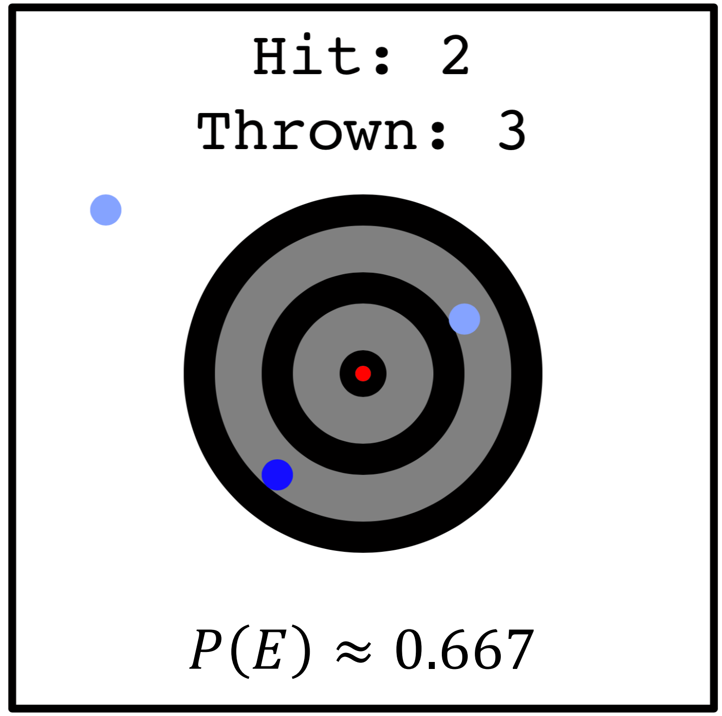
What is a probability?

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

n = # of total trials

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

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



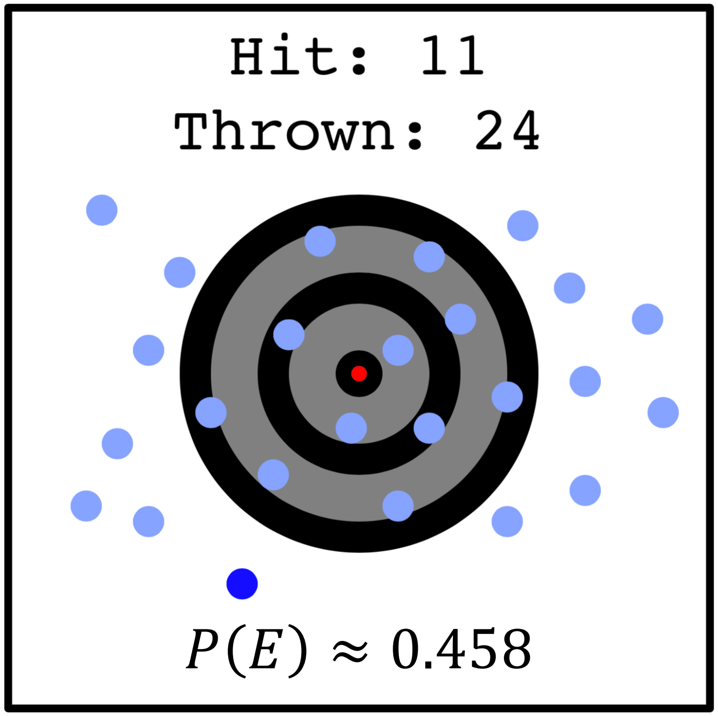
What is a probability?

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

n = # of total trials

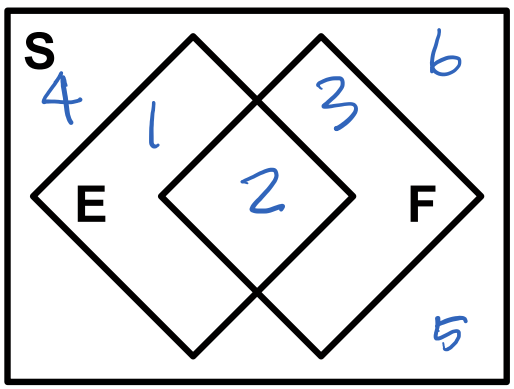
$n(E)$ = # trials where E occurs

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



Axioms of Probability

Quick review of sets



E and F are events in S.

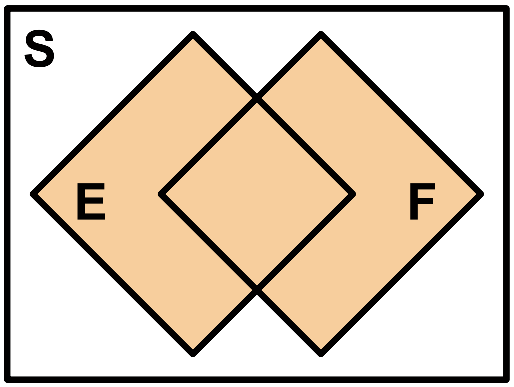
Experiment:

Die roll

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

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

Quick review of sets



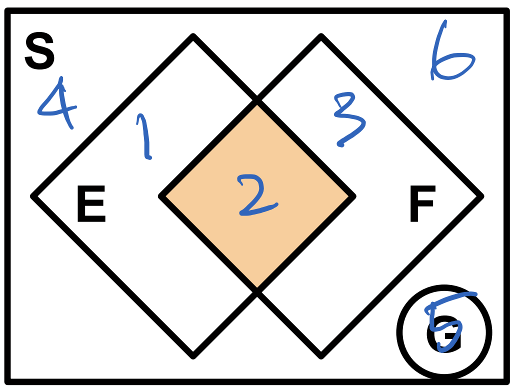
E and F are events in S.
Experiment:
Die roll
$S = \{1, 2, 3, 4, 5, 6\}$
Let $E = \{1, 2\}$, and $F = \{2,3\}$

def Union of events, $E \cup F$

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 $F = \{2,3\}$

def Intersection of events, $E \cap F$

The event containing all outcomes in E ***and*** F.

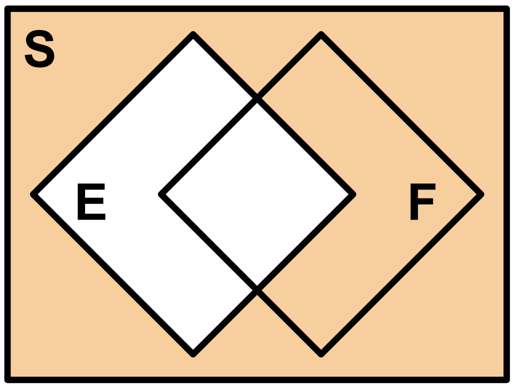
$G = \{5\}$

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

def Mutually exclusive events F and G means that $F \cap G = \emptyset$

Quick review of sets

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



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: $P(S) = 1$

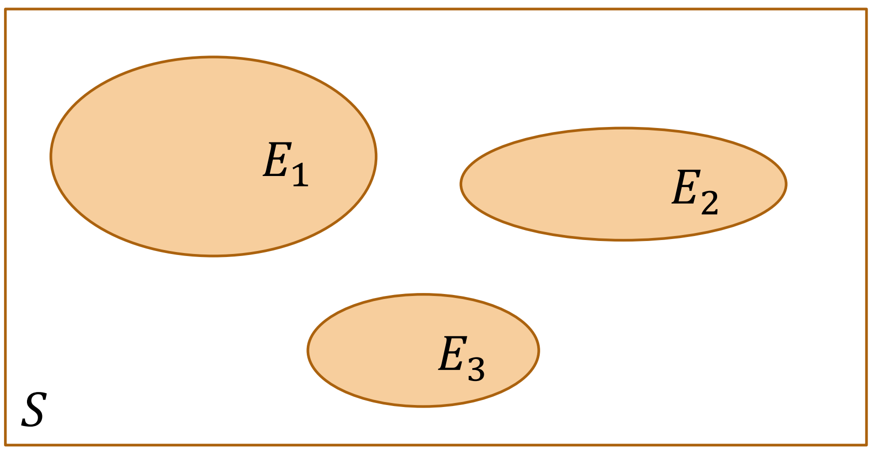
Axiom 3: If E and F are mutually exclusive ($E \cap F = \emptyset$), then $P(E \cup F) = P(E) + P(F)$

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 $E_1, E_2, \ldots$:

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



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

Equally Likely Outcomes

Equally Likely Outcomes

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

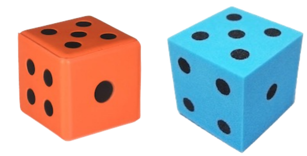
- Coin flip: S = {Head, Tails}
- Flipping two coins: S = {(H, H), (H, T), (T, H), (T, T)}
- Roll of 6-sided die: S = {1, 2, 3, 4, 5, 6}

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)

Roll two dice

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



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

$P(E) = P(E_{(1,6)}) + P(E_{(2,5)}) + \dots$

$S = \{$ (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)$\}$

$\frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} \cdots \frac{1}{6}$

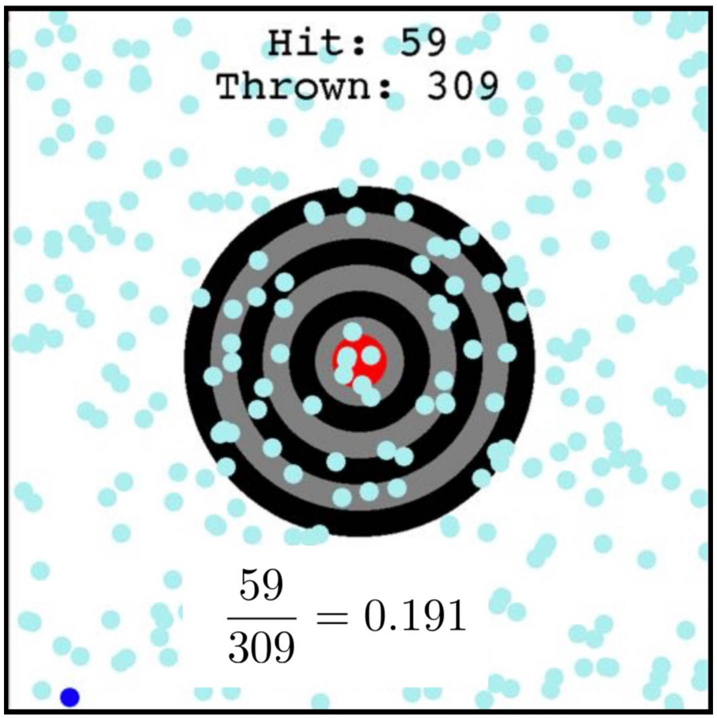
$E = \{(1,6), (2,5), (3,4), \dots, (6,1)\}$

Target revisited

Target revisited

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

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



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

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?

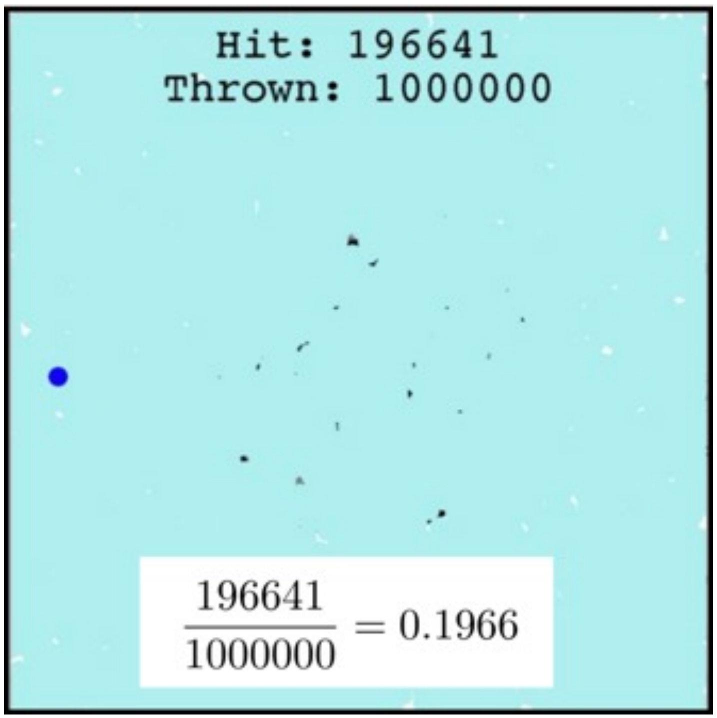
$|S| = 800^2$ $\quad |E| \approx \pi \cdot 200^2$

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

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 $P(\text{win})$?

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

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

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

calottery

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

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.

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. 0

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

(No)

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

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

E:
- css: $4 \cdot 3 \cdot 2 = 24$
- scs: $3 \cdot 4 \cdot 2 = 24$
- ssc: $3 \cdot 2 \cdot 4 = 24$

→ 72

$P(E) = \frac{72}{210} \overset{/6}{=} \frac{12}{35}$

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
- E = 1 distinct cat, 2 distinct sharks

$$\binom{4}{0}\binom{3}{3} + \binom{4}{1}\binom{3}{2} + \binom{4}{2}\binom{3}{1} + \binom{4}{3}\binom{3}{0}$$

$$|S| = \binom{7}{3} = 35$$

$$|E| = \binom{4}{1}\binom{3}{2} = 4 \cdot 3 = 12$$

$$P(E) = \frac{|E|}{|S|} = \frac{12}{35}$$

Excercises

Summary so far

Counting tasks on n objects

- **Sort objects (permutations)**
  - Distinct (distinguishable): $n!$
  - Some distinct: $\frac{n!}{n_1!\,n_2!\cdots n_r!}$
- **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: $\frac{(n+r-1)!}{n!\,(r-1)!}$

Equally likely outcomes:

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

Combinatorics

Probability

Counting? Probability? Distinctness?

We choose **3 books** from a set of
4 distinct (distinguishable) and **2 indistinct** (indistinguishable) books.
Each set of 3 books is equally likely.

Let event E = our choice excludes one or both indistinct books.

1. How many distinct outcomes are in E?

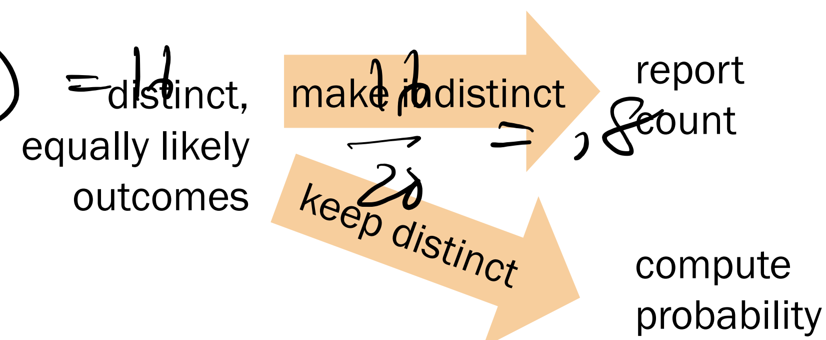
   $\binom{4}{2} = 6$, $\binom{4}{3} = 4$ $\Rightarrow$ 10

2. What is $P(E)$?

   $|E_{dist}| = 2 \cdot \binom{4}{2} + \binom{4}{3} = 16$

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

   $\frac{16}{20} = .8$



CS109 Mixer

Then check out the two questions on the next slide (Slide 33). Take two minutes to speak and brainstorm.

Poker Straights and Computer Chips

1. Consider equally likely 5-card poker hands.
   - Define "poker straight" as 5 consecutive rank cards of any suit

   What is P(Poker straight)?

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

2. 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?)

1. Any Poker Straight

Consider equally likely 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
⋮
10 J Q K A

Define

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

Compute $\quad P(\text{Poker straight}) = \frac{|E|}{|S|} = \frac{10(4)^5}{\binom{52}{5}} = 0.00394$

2. 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)

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

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

Compute

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

2. 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

$\frac{k}{n}$

Define

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

Mid Slide Deck Stretch

Announcements

possibly 6pm

Problem Set #1

Out: this past Wednesday
Due: Friday 4/8, 3:15pm
Covers: through today

Python tutorial

When: Monday, 4/4 at 7:00pm
Where: online via Zoom
Recorded? yes
Notes: to be posted online

Section sign-ups

Form released: already out
Form due: Saturday 4/2, noon
Results: latest Monday

Getting help

Office hours: they've started!
https://cs109.stanford.edu/office-hours

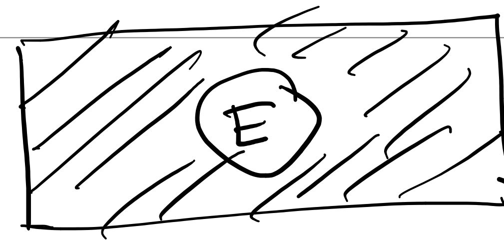
Interesting probability news

"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."

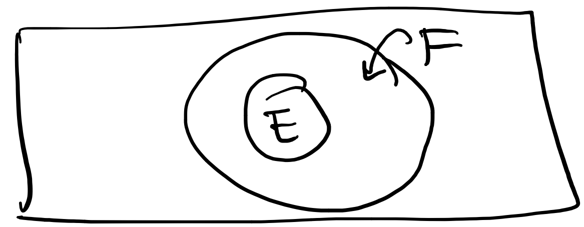
https://actu.epfl.ch/news/decoding-beethoven-s-music-style-using-data-scienc/

Corollaries of Probability

3 Corollaries of Axioms of Probability



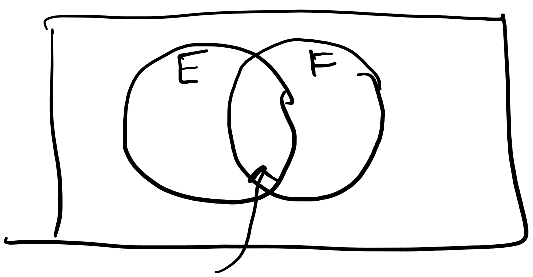
Corollary 1: $P(E^C) = 1 - P(E)$

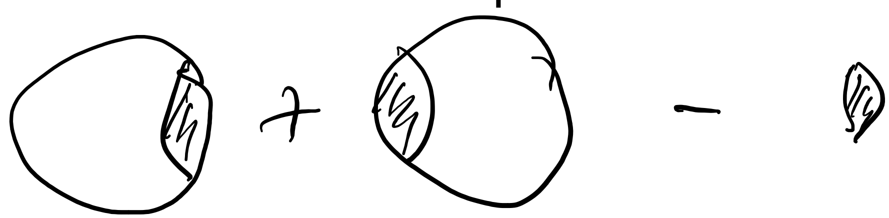


Corollary 2: If $E \subseteq F$, then $P(E) \leq P(F)$

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

(Inclusion-Exclusion Principle for Probability)





Selecting Programmers

- P(student programs in Python) = 0.28
- P(student programs in C++) = 0.07
- P(student programs in Python and C++) = 0.05.

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

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

$$P((Y \cup D)^c) = 1 - P(Y \cup D)$$

$$= 1 - (P(Y) + P(D) - P(Y \cap D))$$

$$= 1 - (.28 + .07 - 0.5)$$

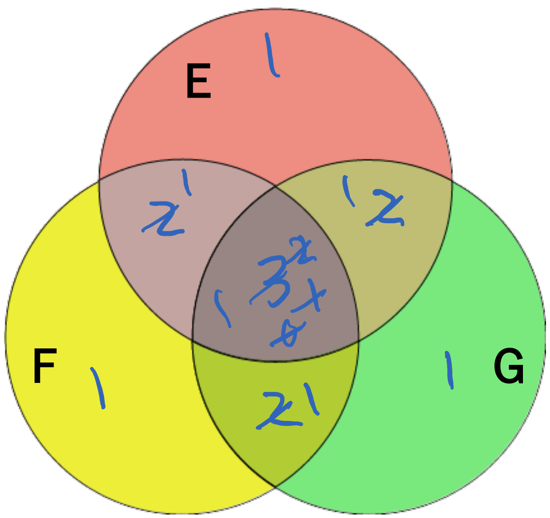
$$.3$$

$$= .7$$

Inclusion-Exclusion Principle (Corollary 3)

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

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



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

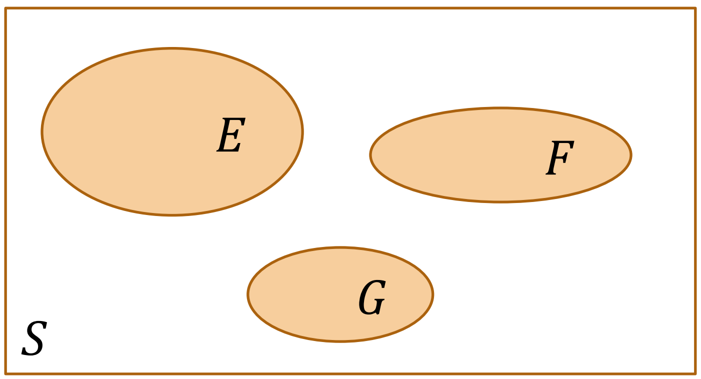
r = 1: $P(E) + P(F) + P(G)$

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

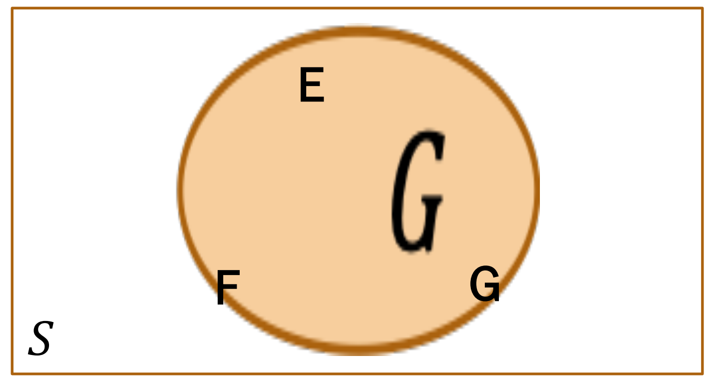
r = 3: $+ P(E \cap F \cap G)$

Takeaway: Union of events

Axiom 3,
Mutually exclusive events

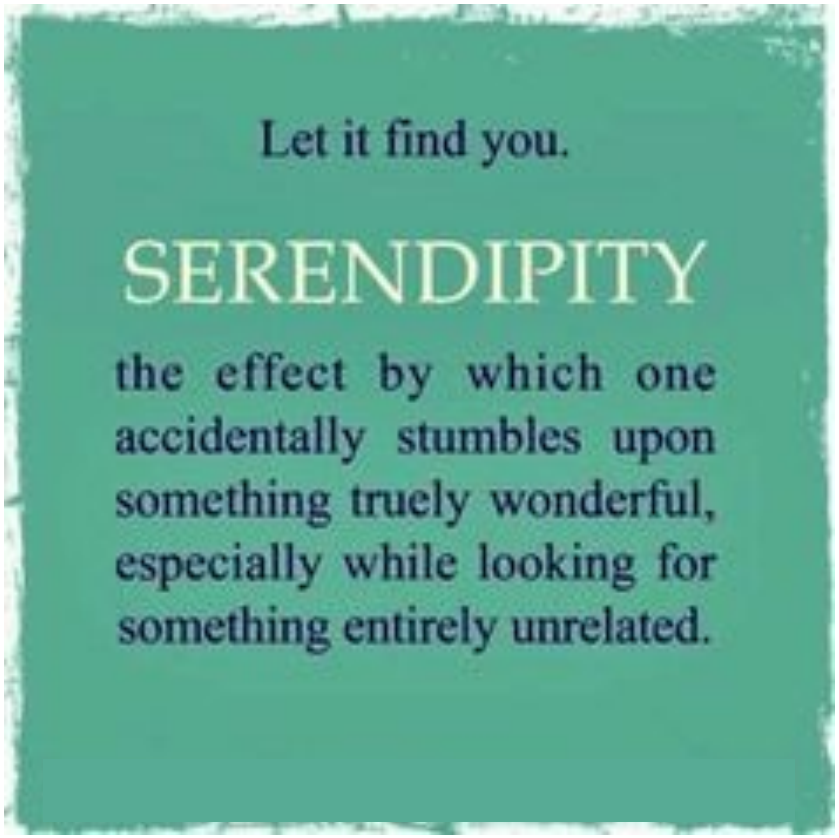


Corollary 3,
Inclusion-Exclusion Principle



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 $r = 100$ people.
- Walk into a room, see $k = 223$ random people.
- Assume each group of k Stanford people is equally likely to be in the room.

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

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

CS109 Mixer

Slide 48 is a question to think over by yourself (~30 seconds).

Serendipity

- The population of Stanford is $n = 17{,}000$ people.
- You are friends with $r = 100$ people.
- Walk into a room, see $k = 223$ random people.
- Assume each group of k Stanford people is equally likely to be in the room.

What is the probability that you see at least one friend in the room?

Define

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

What strategy would you use?

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

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 $k = 223$ random people.
- Assume each group of k Stanford people is equally likely to be in the room.

What is the probability that you see at least one friend in the room?

$|S| = \binom{17000}{223}$

Define

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

$|E^C| = \binom{100}{0}\binom{16900}{223} =$

$1 - \frac{\binom{16900}{223}}{\binom{17000}{223}}$

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, at least one pair of them share the same birthday?

For you to think about (and discuss in your first 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)?

51

$|E_{AS}| = 51!$

$|E_{2C}| = 51!$

Have a good weekend!