STAT 88: Lecture 1

Contents

Course resources:

- Course website: http://stat88.org

- Piazza: https://piazza.com/class/kdxsrxzejp77mg

- Gradescope: https://www.gradescope.com/courses/163710

Ice-breaker: Introduce yourself to your classmates (breakout rooms).

Discuss which is probability and which is statistics:
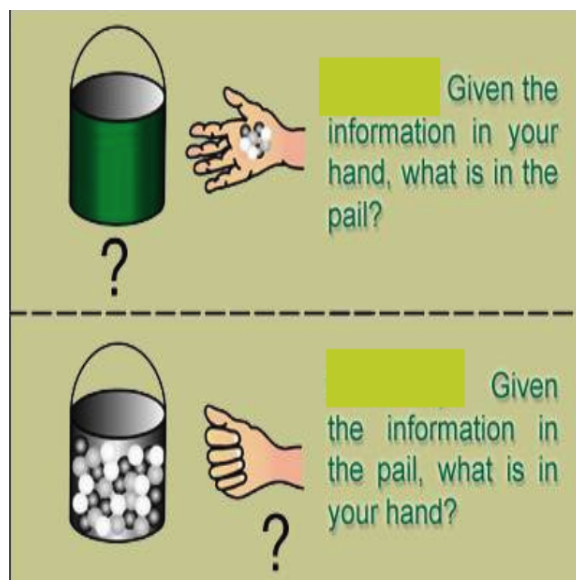
1.1. Probabilities as Proportions

Probability is a numerical measure of uncertainty.

Terminology

- Experiment: activity that has a set of possible outcomes and involves chance

  ex Rolling a six-sided die once

- Outcome space or Sample space: the set of all outcomes of an experiment

  ex $\{1,2,3,4,5,6\}$

- Event: a subset of outcome space

  ex The die shows a multiple of 3

Outcome space is denoted as Ω and $A \subseteq \Omega$ is an event.

Equally likely outcomes: for any $A \subseteq \Omega$, $P(A) = \frac{\#A}{\#\Omega}$.

<u>Example</u>: What is the chance that a die shows a multiple of 3?

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

- $A = \{3.6\}$

- $P(A) = \frac{2}{6} = \frac{1}{3}$

Example: Deck of cards:

4 suits : clubs, diamonds, hearts, spades
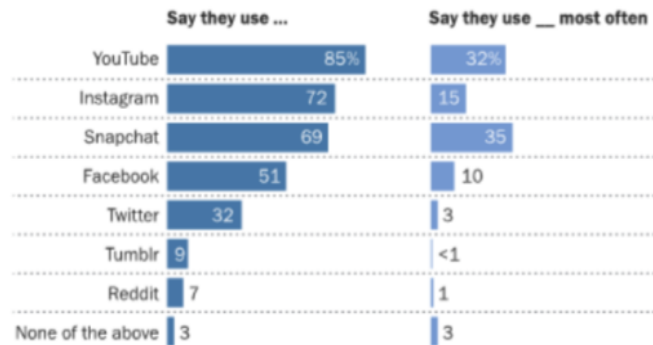
13 ranks : Ace, 2-10, J, Q, K

_____

52 cards

Suppose a deck of cards is shuffled and the top 2 cards are picked. What is the chance that you get at least one ace among the 2 cards? i.e.

$$A = \{(\text{ace}, \text{ace}), (\text{ace}, \text{nonace}), (\text{nonace}, \text{ace})\}.$$ What is $P(A)$?

Example: Teen use of online formats:

YouTube, Instagram and Snapchat are the most popular online platforms among teens

% of U.S. teens who ...

	Say they use ...	Say they use __ most often
YouTube	85%	32%
Instagram	72	15
Snapchat	69	35
Facebook	51	10
Twitter	32	3
Tumblr	9	<1
Reddit	7	1
None of the above	3	3

Note: Figures in first column add to more than 100% because multiple responses were allowed. Question about most-used site was asked only of respondents who use multiple sites; results have been recalculated to include those who use only one site. Respondents who did not give an answer are not shown.
Source: Survey conducted March 7-April 10, 2018.
"Teens, Social Media & Technology 2018"

PEW RESEARCH CENTER

Wish to understand student usage of online platform such as Facebook or Twitter.

Are these columns distribution?

Why the percents add up to more than 100 in the first column?

What is the chance that the chosen teen used FB most often? What is the chance that the chosen teen did not use FB most often?

What is the chance that the student used Facebook but used some other platform more often?

Conditioning We are given a new information: Someone picks a teen at random from the population. We know that the teen used Facebook.
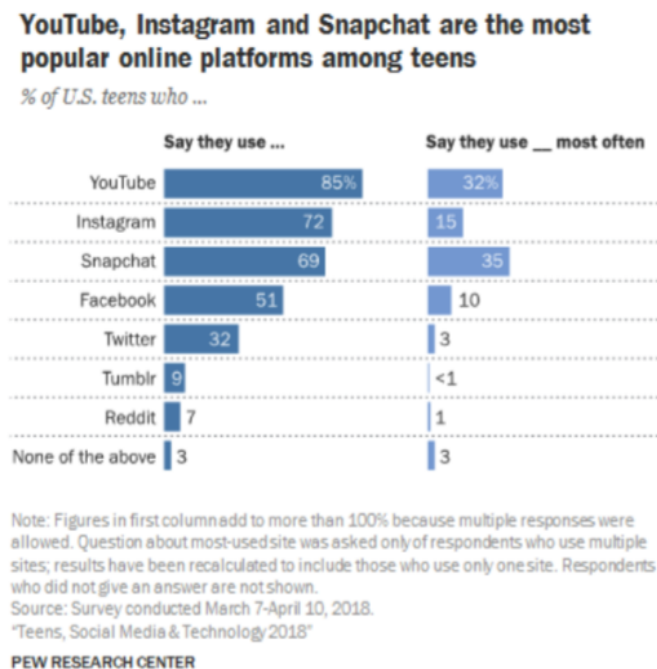
Given this information, what is the chance that the teen used Facebook most often?

To find a conditional probability:

- First restrict the set of all outcomes as well as the event to only the outcomes that satisfy the given condition

- Then calculate proportions accordingly

1.2. Exact Calculation or Bound

Sometimes we can only give a bound for probability of an event instead of the exact value.

YouTube, Instagram and Snapchat are the most popular online platforms among teens

% of U.S. teens who ...

	Say they use ...	Say they use __ most often
YouTube	85%	32%
Instagram	72	15
Snapchat	69	35
Facebook	51	10
Twitter	32	3
Tumblr	9	<1
Reddit	7	1
None of the above	3	3

Note: Figures in first column add to more than 100% because multiple responses were allowed. Question about most-used site was asked only of respondents who use multiple sites; results have been recalculated to include those who use only one site. Respondents who did not give an answer are not shown.
Source: Survey conducted March 7-April 10, 2018.
"Teens, Social Media & Technology 2018"

PEW RESEARCH CENTER

What is the chance that the selected teen used Facebook or Twitter most often?

What is the chance that the selected teen used Facebook or Twitter?

- Upper bound?

- Lower bound?

Extra problem: Two draws are made at random with replacement from the box

a) If you draw a 1, what is the chance that the second number is bigger than 2?

b) Find the chance the second number is bigger than twice the first.

1.3. Fundamental Rules

Formally, probability is a function on events, $P : A \subseteq \Omega \mapsto [0, 1]$, satisfying the following 3 axioms:

1. $P(A) \geq 0$ for all $A \subseteq \Omega$.

2. $P(\Omega) = 1$.

3. If A and B are mutually exclusive, i.e. $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$.