What is probability?

Rain with 30% chance today

Toss a coin — 50% H

experiment: roll a 6-sided die and record the number

outcome: 1, 2, 3, 4, 5, 6

event: includes:

- single outcome e.g. $\{1\}$
- more outcomes, e.g. $\{\text{multiples of } 3\} = \{3, 6\}$
- nothing, e.g. $\{\text{getting a } 7\} = \emptyset$
- everything, e.g. $\{1, 2, 3, 4, 5, 6\}$.

probability (of a outcome / a event)

- prob. of getting 6 $= \frac{1}{6} = P(\{6\}) = P(6)$
- prob. of getting 3 or 6 $= \frac{2}{6}$ $= P(\{3,6\})$ ← $\#\{3,6\}$ ← $\#\{1,2,\dots,6\}$

Assumption: equally likely outcomes. (only works for finitely many outcomes)

Calculation: prob. $= \dfrac{\#\text{ of target outcomes}}{\text{total }\#\text{ of outcomes}}$

More examples:

- roll a die twice (2, 5)
- draw a card from a standard deck of 52 cards ♡5

Seriously

outcome (ω – omega)

outcome space: (Ω – Omega)

the set containing all possible outcomes as its element.

event (A, B, C, etc.): a subset of Ω,

i.e. contains part of the outcomes.

could be $\emptyset$, or Ω itself.

probability (of the event) $P(A)$, $P(\cdots)$

3 Axioms of probability. → $\boxed{\text{simplest}}$ set of rules made by scientists so that the theory works.

1. non-negativity: for any event A, $P(A) \geq 0$
2. one: $P(\Omega) = 1$
3. addition: for any two events A, B, suppose $A \cap B = \emptyset$ then $P(A \cup B) = P(A) + P(B)$

3* generalized addition: for any events $A_1, A_2, \dots, A_n$ "mutually exclusive", then

$$P(\underbrace{\boxed{A_1 \cup A_2 \cup \cdots}}_{A} \cup \overset{B}{\boxed{A_n}}) = P(A_1) + P(A_2) + \cdots + P(A_n) = \sum_{i=1}^{n} P(A_i)$$

$$= P(\underbrace{A_1 \cup A_2 \cup \cdots}_{A} \cup \underbrace{A_{n-1}}_{B}) + P(A_n)$$

$$= P(A_1 \cup A_2 \cup \cdots \cup A_{n-2}) + P(A_{n-1}) + P(A_n)$$

Consequences from the 3 Axioms:

1. Complement.

Note that for any event A, $\begin{cases} A \cap A^c = \emptyset \\ A \cup A^c = \Omega \end{cases}$

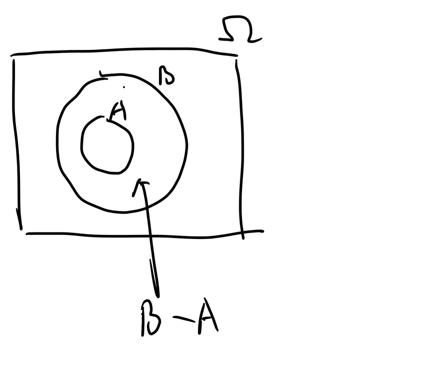
$$1 \overset{(A2)}{=} P(\Omega) = P(A \cup A^c) \overset{(A3)}{=} P(A) + P(A^c)$$

$$\Rightarrow \begin{cases} P(A) = 1 - P(A^c) \overset{(A1)}{\leq} 1 \\ P(A^c) = 1 - P(A) \overset{(A1)}{\leq} 1 \end{cases} \Rightarrow P(A) \leq 1$$

2. Difference

For events $A \subseteq B$,

note that $\begin{cases} A \cap (B-A) = \emptyset \\ A \cup (B-A) = B. \end{cases}$



$$P(B) = P(A \cup (B-A)) \overset{(A3)}{=} P(A) + P(B-A)$$

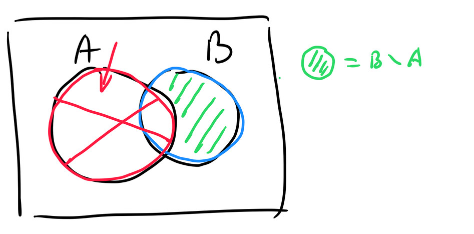
$$\Rightarrow P(B-A) = P(B) - P(A) \overset{(A1)}{\leq} P(B), \text{ for } A \subseteq B.$$

$$\Rightarrow \begin{cases} P(B-A) \leq P(B) \\ P(A) \leq P(B) \end{cases} \text{ for } A \subseteq B.$$

3. Addition inequality.

For any two events A, B

$P(A) + P(B) \geq P(A \cup B)$.



proof: $P(A \cup B) = P(A) + P(\underbrace{B \setminus A}_{\subseteq B})$

$\leq P(A) + P(B)$

$B \setminus A$

$B - A$ ← This also means $A \subseteq B$.

For any events $A_1, A_2, \cdots A_n$

$$P(A_1 \cup A_2 \cup A_3 \cdots \cup A_n) \leq \sum_{i=1}^{n} P(A_i)$$

use vs. use most often

left: add up to > 100%

$\boxed{\text{right: add up to 100\%}}$

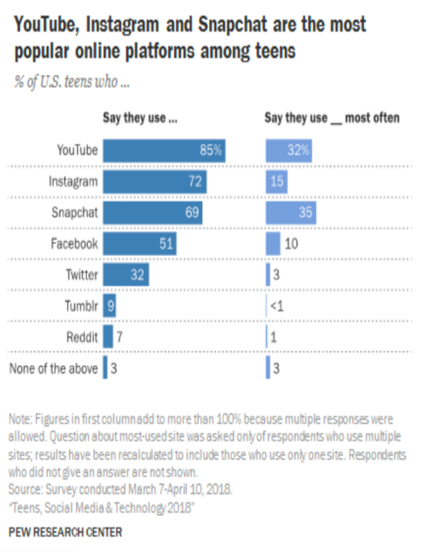
"distribution"

can only belong to one category

experiment: select a U.S. teen "uniformly" at random

equally likely to pick each one.

outcome: what platforms he uses and what he uses most often.



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

Direct information

1. prob. that he uses FB

$P(FB) = 51\%$

2. prob. that he uses FB most often

$P(FB \text{ most often}) = 10\%$

Indirect information

3. addition for combination of two distinct groups

prob. that he uses FB most often or he uses Twitter most often.

Addition Axiom $= 10\% + 3\% = 13\%$

$\boxed{\text{Think: prob. that he use FB or Twi. or both? } 51\% + 32\% \text{ ?}}$

4. Complement:

prob. that he do not use FB

$P(FB^c) = 1 - P(FB) = 1 - 51\% = 49\%$

5. Subtraction

prob. that he uses FB but does not use FB most often.

$P(\{FB\} - \{FB \text{ most often}\}) = P(FB) - P(\cdots) = 51\% - 10\% = 41\%$

hidden: if he uses FB most often, he uses FB.

6. Conditioning.

what if we know he uses FB?

prob. that he uses FB most often provided he uses FB.

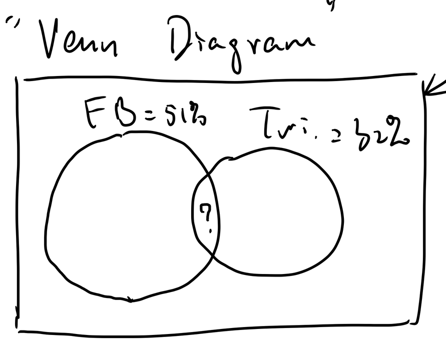
$= \dfrac{10\%}{51\%}$ ← new total number

Treat this as if we are considering a smaller group of U.S. Teens

total % $100\% \rightarrow 51\%$.

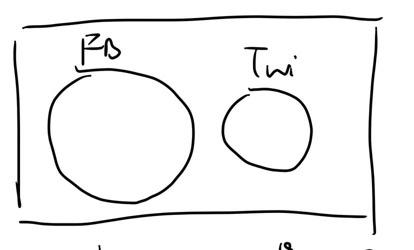
Uncertain information.

"Venn Diagram" world of outcomes $= 100\%$



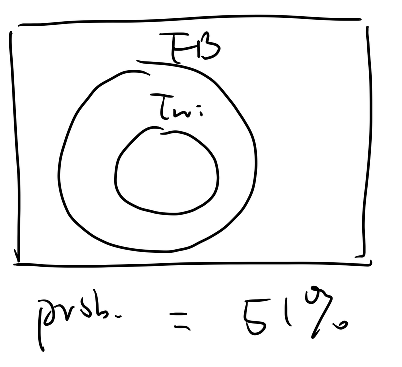
Think about extreme cases.

Case 1:



prob. $= 51\% + 32\% = 83\%$.

Case 2:

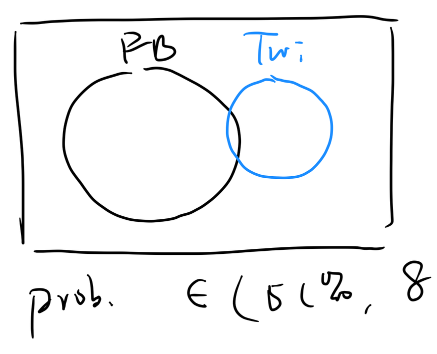


prob. $= 51\%$

A = FB

B = Twi.

General case

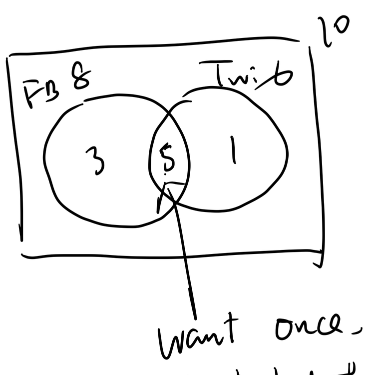


$$\underset{51\%}{P(A)} \leq P(A \cup B) \leq \underset{83\%}{P(A) + P(B)}$$

prob. $\in (51\%, 83\%)$

Chapter 2.

Inclusion-exclusion. (originated from counting)



want once, calculated twice ⇒ remove once.

$\#\{FB \text{ or } Twi\} = 3 + 5 + 1 = 9$

$\#\{FB \text{ or } Twi\} = \#\{FB\} + \#\{Twi\} - \#\{FB \,\&\, Twi\}$

$= 8 + 6 - 5$

$= 9$.

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

know 3 ⇒ know 4th.

Roll a fair 6-sided die twice.

What is the prob. that we get at least one 6?

outcome: (a, b), $a, b = 1, 2, \cdots, 6$

events (we care): $A = \{\text{first roll is } 6\} = \{(6, b) : b = 1, \cdots 6\}$

$B = \{\text{second roll is } 6\} = \{(a, 6) : a = 1, \cdots 6\}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{6}{36} + \frac{6}{36} - \frac{1}{36} = \frac{11}{36}.$$