Lesson 3: Defining Probability

TB sections 2.1

Nicky Wakim

2024-10-07

Learning Objectives

1. Define probability and explain the Law of Large Numbers within examples
2. Define relationships between events and their probability properties (including disjoint events, non-disjoint events, complements, and independent events)
3. Calculate an unknown probability in a word problem using the probability properties

Where are we?

Data
- Collecting data
- Categorical vs. Numeric
- Summary statistics
- Data visualization

Probability
- Probability rules
- Independence, conditional
- Random variables and probability distributions
- Linear combinations
- Binomial, Normal, and Poisson

Sampling Variability, and Statistical Inference
- Sampling distributions
- Central Limit Theorem
- Confidence Intervals
- Hypothesis tests

Inference for continuous data/outcomes
- One sample t-test
- 3+ independent samples
- Simple linear regression / correlation
- 2 sample tests: paired and independent
- Power and sample size
- Non-parametric tests

Inference for categorical data/outcomes
- One proportion test
- Fisher's exact test
- Non-parametric tests
- Chi-squared test
- 2 proportion test
- Power and sample size

R
- Basics
- Reproducibility
- Quarto
- Packages
- Data visualization
- Data wrangling
- R Projects
- …

Let's start with an example!

Example: Rolling fair 6-sided dice

Suppose you roll a fair 6-sided die.

1. What is the probability that you roll a 4?
2. What is the probability that you roll an even number?
3. What is the probability that you did not roll a 3?

① 6 possible sides

$P(\text{getting } 4) = \frac{1}{6}$

\# outcomes in event

total # possible outcomes

② 2, 4, 6

$P(\text{even}) = \frac{3}{6} = \frac{1}{2}$

③ 1, 2, 4, 5, 6

$P(\text{not } 3) = \frac{5}{6}$

Learning Objectives

1. Define probability and explain the Law of Large Numbers within examples
2. Define relationships between events and their probability properties (including disjoint events, non-disjoint events, complements, and independent events)
3. Calculate an unknown probability in a word problem using the probability properties

What is a probability?

Definition: Probability

How likely something will happen.

- On a more technical note, the probability of an outcome is the proportion of times the outcome would occur if the random phenomenon could be observed an infinite number of times.

- We can think of flipping a coin. There are two possible outcomes (heads or tails). The probability of getting heads is 0.5.

What is a probability? with the Law of Large Numbers

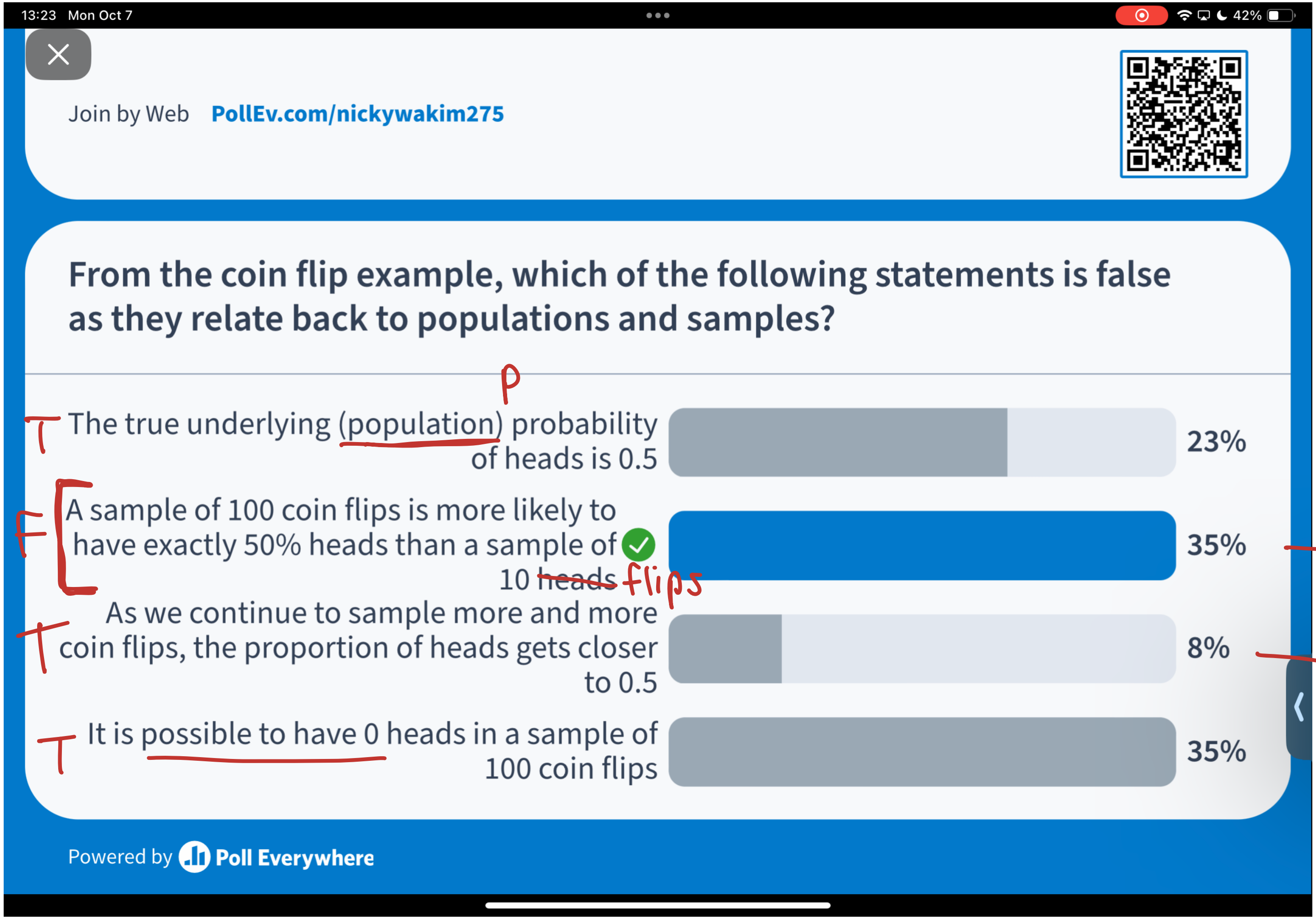
- We can think of flipping a coin. There are two possible outcomes (heads or tails). The probability of getting heads is 0.5.
  - If we flip the coin 10 times, it is not certain that we will get 5 heads. However, if we flip it infinite times, we will get heads 50% of the flips.

- Fun "Seeing Theory" demonstration!

Law of large numbers

As more observations are collected, the proportion of occurrences, $\hat{p}$, with a particular outcome converges to the true probability p of that outcome.

Poll Everywhere Question 1



the prob of exactly one event (50 Heads)

converging towards population prop.

45 - 55 heads

Some probability notation

- Probability typically defined as a proportion
  - Takes values between 0 and 1

$\hat{p}$ or p

$0 \leq p \leq 1$

- Probability can also be expressed as a "percent chance," taking values between $\boxed{\text{0\% and 100\%}}$

- If we want to discuss the probability of an event, say A, we would write $P(A)$
  - We can write: $A = \{\text{rolling a 1}\}$, with associated probability $P(A)$
  - OR we can write $P(\text{rolling a 1})$

Let A = rolling a 1

$P(A)$ $P(\text{rolling a 1})$

Learning Objectives

1. Define probability and explain the Law of Large Numbers within examples

2. Define relationships between events and their probability properties (including disjoint events, non-disjoint events, complements, and independent events)

3. Calculate an unknown probability in a word problem using the probability properties

Disjoint / mutually exclusive events

Disjoint / mutually exclusive events

Two events or outcomes are called **disjoint** or **mutually exclusive** if they cannot both happen at the same time.

ONE COIN FLIP

event 1: flipping a H

event 2: flipping a T

> disjoint

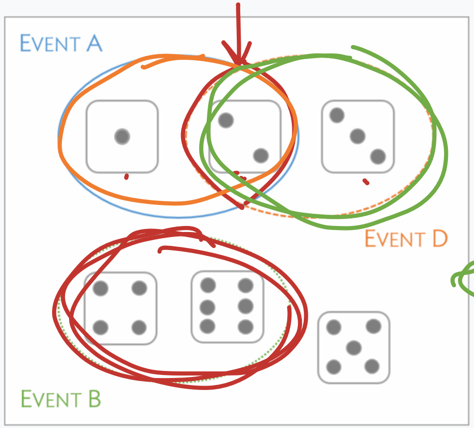
cannot have H & T in 1 coin flip

Poll Everywhere Question 2

13:35 Mon Oct 7 38%

Join by Web **PollEv.com/nickywakim275**

Let's say we roll a die with events A, B, and D: A is the event of rolling a 1 or 2, B is the event of rolling a 4 or 6, and D is the event of rolling a 2 or 3. Which events are disjoint?



A and B		4%
~~A and D~~		~~18%~~
B and D		4%
~~Both (A and B) and (A and D)~~		~~4%~~
Both (A and B) and (B and D) ✓		61%
~~Both (A and D) and (B and D)~~		~~0%~~
~~All events are disjoint~~		~~11%~~

Powered by **Poll Everywhere**

Probability for disjoint events

A & B disj: $P(A \text{ or } B) = P(A \cup B) = P(A) + P(B)$

Probability rule for disjoint events

If A_1 and A_2 represent two disjoint outcomes, then the probability that either one of them occurs is given by

$$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2)$$

If there are k disjoint outcomes $A_1, \ldots, A_k$, then the probability that either one of these outcomes will occur is

$$P(A_1) + P(A_2) + \cdots + P(A_k)$$

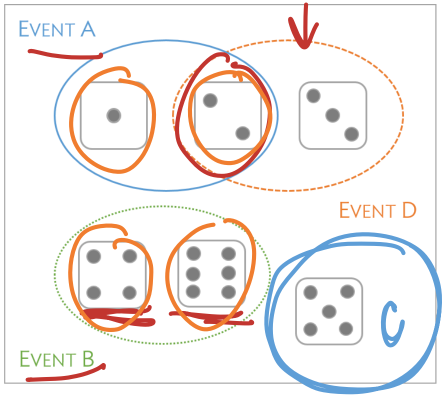
- From the poll everywhere question with the die, what is the probability of event A or B?

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

$P(A \text{ or } B) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$

A & B & C disjoint

$P(A \cup B \cup C) = P(A) + P(B) + P(C)$



Probabilities when events are not disjoint

- When events are not disjoint, we cannot use the previous addition rule for probabilities!!
- We must use a general rule that recognizes the potential overlap between events

General probability addition rule

If A and B are any two events, disjoint or not, then the probability that at least one of them will occur is

$P(A \cap B)$

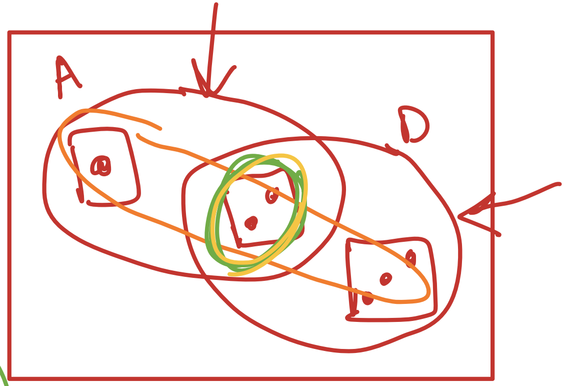
$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B),$$

where $P(A \text{ and } B)$ is the probability that both events occur.

$P(A \cap B)$

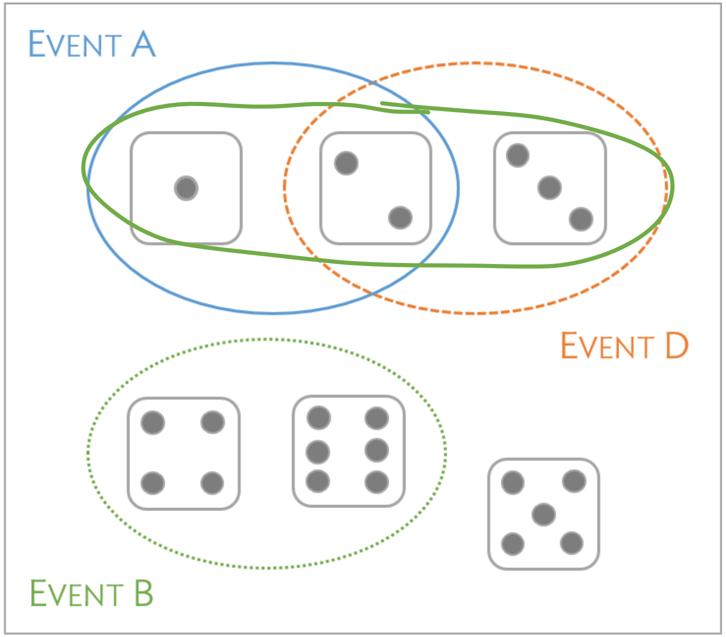
$P(A \text{ or } D) = P(A) + P(D) - P(A \cap D)$

$= \frac{2}{6} + \frac{2}{6} - \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$



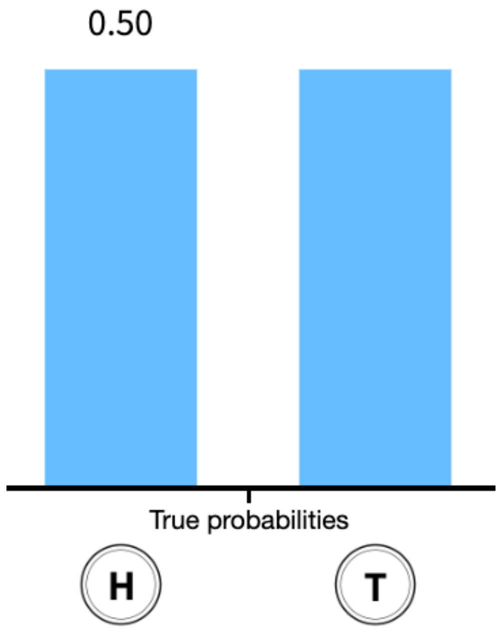
Think back to our die

- Event A and D are not disjoint, they share an outcome of rolling a 2
- How do we find the probability of event A or event D?



Probability distributions

- A **probability distribution** consists of all disjoint outcomes and their associated probabilities
- We've already seen one in our heads and tails example

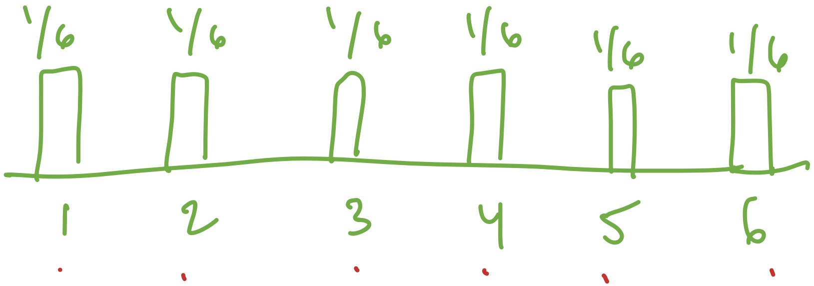


Rules for a probability distribution

A probability distribution is a list of all possible outcomes and their associated probabilities that satisfies three rules:

1. The outcomes listed must be disjoint
2. Each probability must be between 0 and 1 ✓
3. The probabilities must total to 1

die:

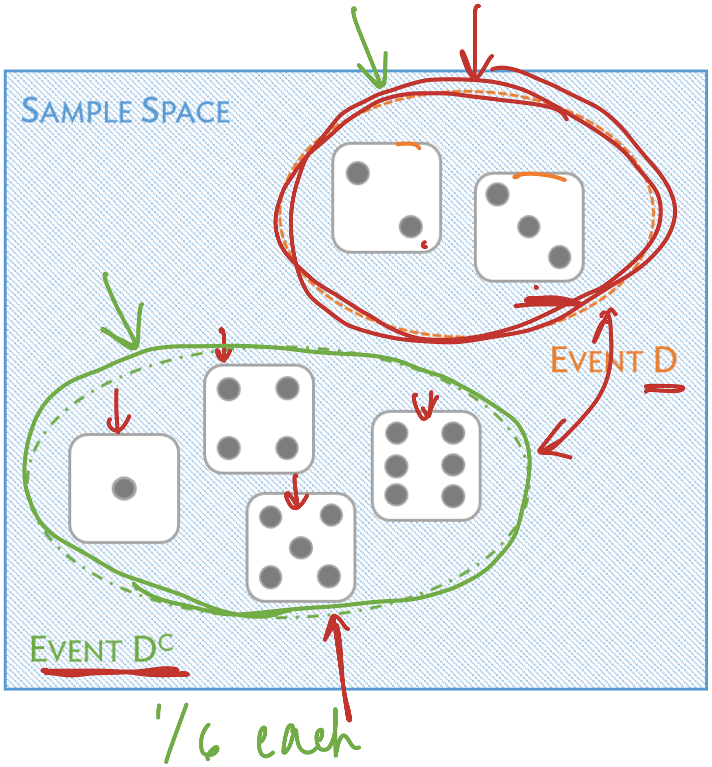


$6 \cdot \frac{1}{6} = 1$

Complement of an event

We need two math definitions for this:

- **Sample space:** denoted as S is the set of all possible outcomes
- **Complement:** complement of an event, say D, represents all the outcomes in the sample space that are not in D
  - Complement is denoted as D^c or D'



Complement

The complement of event A is denoted A^c, and A^c represents all outcomes not in A. A and A^c are mathematically related:

$$P(A) + P(A^c) = 1, \quad \text{i.e.} \quad P(A) = 1 - P(A^c).$$

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

$P(D) + P(D') = 1$

$\frac{2}{6} = \frac{4}{6} = 1$

Independence

- Two processes are **independent** if knowing the outcome of one provides no information about the outcome of the other
- For example, if we flip two different coins and one lands on heads, what does that tell us about the other coin?

Multiplication Rule for independent processes

If A and B represent events from two different and independent processes, then the probability that both A and B occur is given by:

$P(A \cap B)$

$$P(A \text{ and } B) = P(A)P(B).$$

Similarly, if there are k events A_1, ..., A_k from k independent processes, then the probability they all occur is

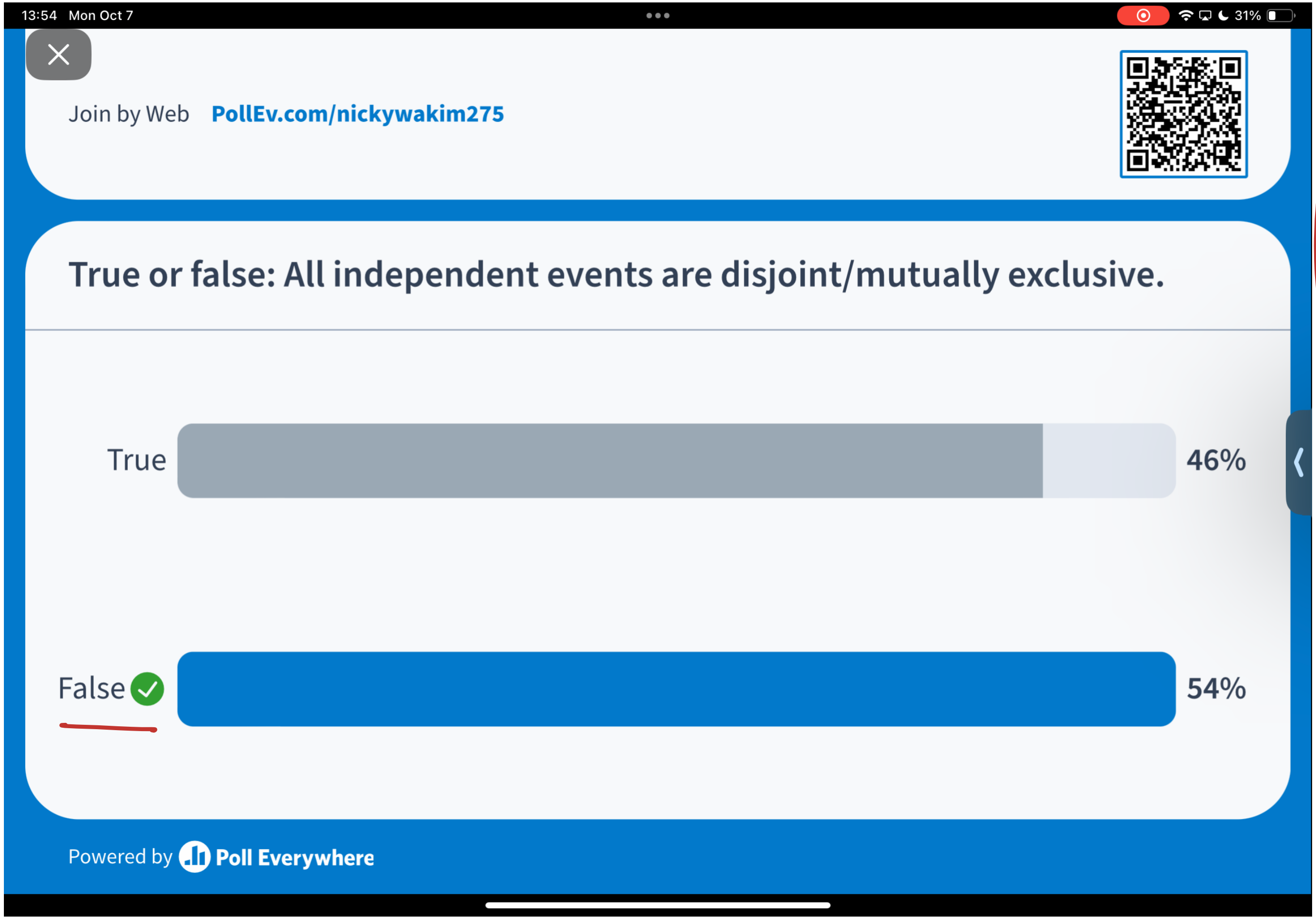
$$P(A_1)P(A_2)\cdots P(A_k).$$

2 coins:
prob of heads
in both

A = head in coin 1
B = head in coin 2

$$P(A \cap B) = P(A)P(B) = \frac{1}{2}\cdot\frac{1}{2} = 1/4$$

Poll Everywhere Question 3



die:

A: 1 or 2

B: 4 or 5

$$\left[P(A \cap B) = 0 \right]$$

indep:

$$\underline{P(A \cap B)} = \underline{P(A)}\,\underline{P(B)}$$

$$\frac{2}{6} \cdot \frac{2}{6}$$

If you do not know A & B are independent, then you likely need to use:

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

(and know 3 out of 4 of these probabilities)

~~if ind $= P(A)P(B)$~~

Learning Objectives

1. Define probability and explain the Law of Large Numbers within examples
2. Define relationships between events and their probability properties (including disjoint events, non-disjoint events, complements, and independent events)
3. Calculate an unknown probability in a word problem using the probability properties

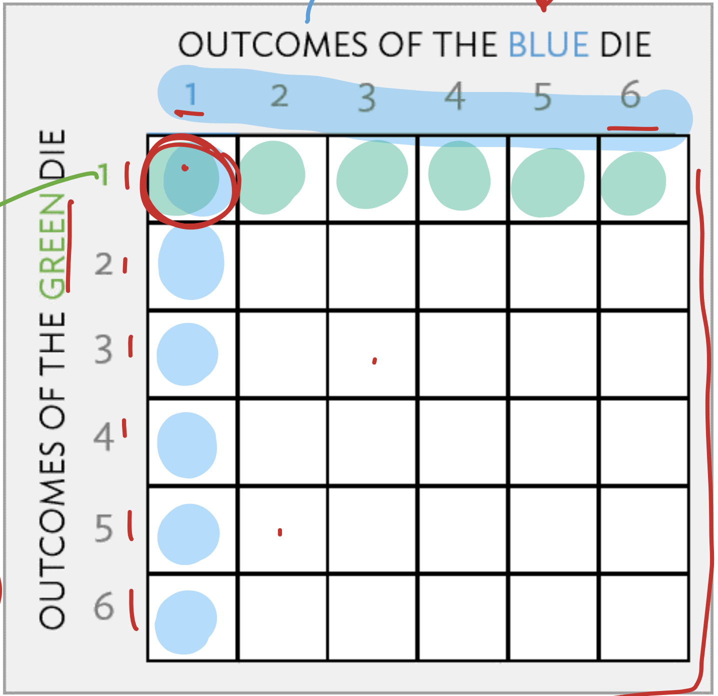
Example rolling two dice

What is the probability that both dice will be 1?

$P(\text{Blue is } 1 \;\&\; \text{green is } 1) = \frac{1}{36}$ $\cap$

$P(1) = \frac{1}{6}$

$P(1) = \frac{1}{6}$



$P(\text{Blue is } 1 \;\&\; \text{green is } 1)$

$= P(\text{Blue is } 1) \cdot P(\text{Green is } 1)$

$= \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

General steps for probability word problems

1. Define the events in the problem and make a Venn Diagram
2. Translate the words and numbers into probability statements → $P(A) = \frac{2}{3}$
3. Translate the question into a probability statement
4. Think about the various definitions and rules of probabilities. Is there a way to define our question's probability statement (in step 3) using the probability statements with assigned values (in step 2)?
5. Plug in the given numbers to calculate the answer!

Weekly medications

Example 3

If a subject has an

- 80% chance of taking their medication *this* week,
- 70% chance of taking their medication *next* week, and
- 10% chance of *not* taking their medication *either* week,

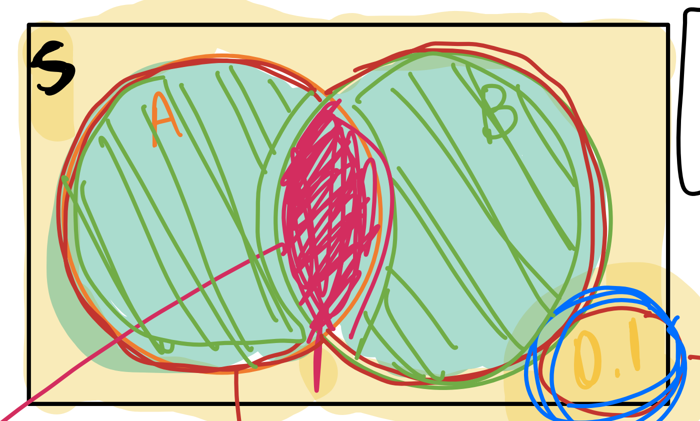
~~then find the probability of them taking their medication exactly one of the two weeks.~~

probability that we take it both weeks

Hint: Draw a Venn diagram labelling each of the parts to find the probability.

Let A = take med this week

B = take med next week



$P(A) = 0.8$

$P(B) = 0.7$

$\rightarrow (A \cup B)^c$

pink is intersection

$\underbrace{P(A \cup B)}_{F} = 1 - \underbrace{P((A \cup B)^c)}_{F^c}$ ✓

$$P(A \cup B) = 1 - 0.1 = 0.9$$

$P(A)$ $P(B)$ $P(A \cap B)$

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

$$0.9 = 0.8 + 0.7 - P(A \cap B)$$

$$+P(A \cap B) \quad -0.9 \quad + P(A \cap B)$$

$$-0.9$$

$$P(A \cap B) = 0.8 + 0.7 - 0.9 = 0.6$$

The probability that they take meds both weeks is 0.6

2 wk:

		This wk med	This wk not med
next	med		
next	not med		