Probability

Sean McGrath

June 17, 2025

Yale SCHOOL OF PUBLIC HEALTH
Big Data Summer Immersion

Outline: Part 1

- Basic Set Theory
- Probability Measures and Basic Properties
- Discrete Probability
- Counting Tools

Outline: Part 2

- Conditional Probability
- Independence
- Law of Total Probability

Part 1

What is Probability?

- We often use probability in day-to-day life
  - **Weather Forecasting**: "What is the chance of rain in New Haven tomorrow?"
  - **Health Decisions**: "What is the lifetime risk of developing breast cancer?"

- Notions of probability often affects how we make decisions
  - Should I bring an umbrella or not?
  - Should I get cancer screening?

- How do we **define** and **interpret** probability?

What is probability?

Nobody has responded yet.

Hang tight! Responses are coming in.

What is Probability

- **Informal Definition:**
  - Mathematical theory that describes and characterizes uncertainty

- **Interpretation:**
  - The interpretation of probability is not universally agreed on!
    - **Frequentist interpretation:** Frequency at which an event occurs
    - **Bayesian interpretation:** Belief in the chance of an event occurring

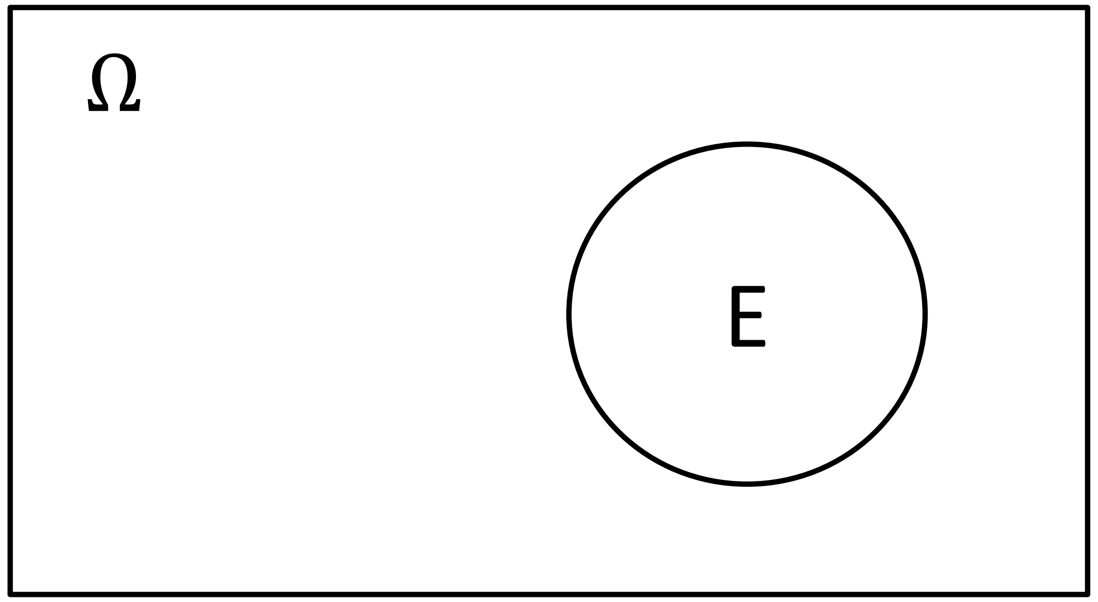
- **Towards a more formal definition:**
  - A more formal definition of probability requires an understanding of set theory, which we explore next

Sample Space and Events

- **Sample Space**: The set of all possible outcomes of a random experiment. We will denote it by Ω.

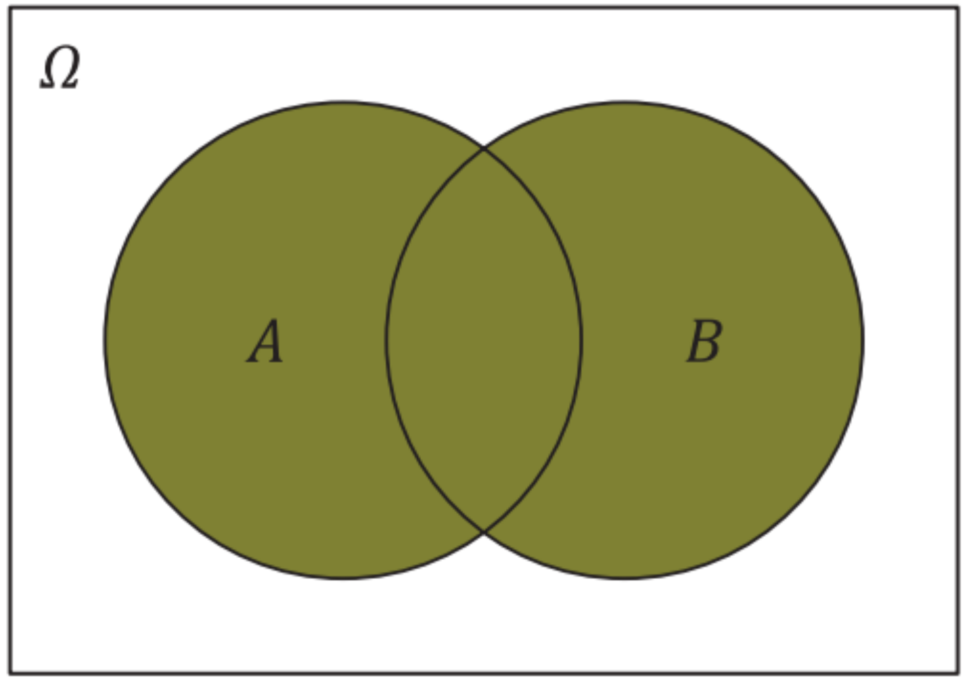
- **Event**: Any collection of possible outcomes of a random experiment, i.e., a subset of the sample space S. We will denote it by E.

- **Example**: Tossing a coin
  - Sample space: heads and tails, denoted by $\Omega=\{H,T\}$
  - Some events: $E = \{H\}$, $E = \{T\}$, $E = \{H,T\}$



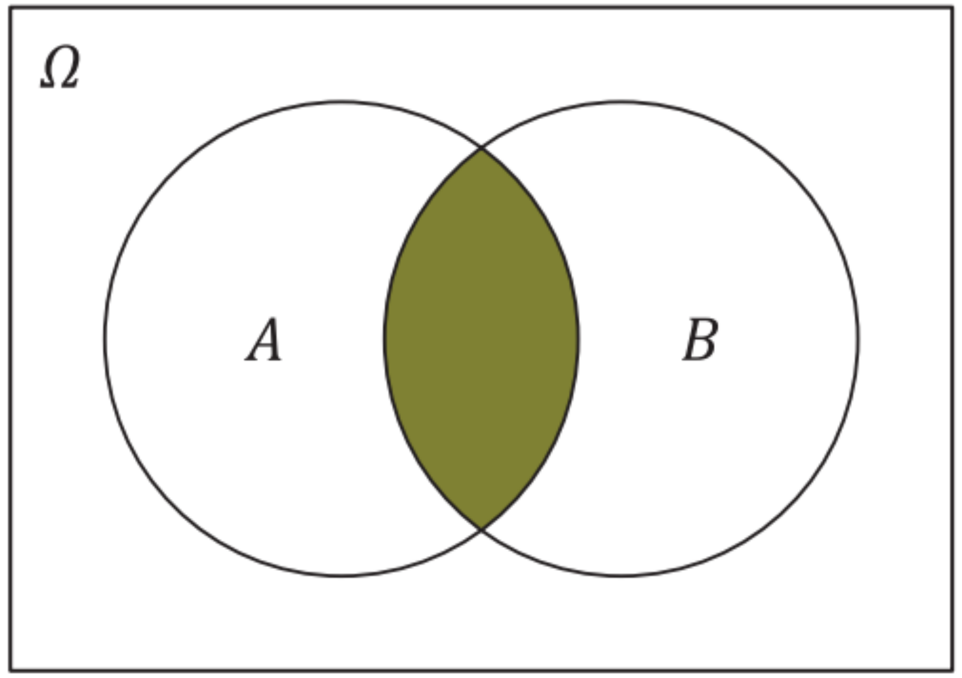
Set Operations

Union



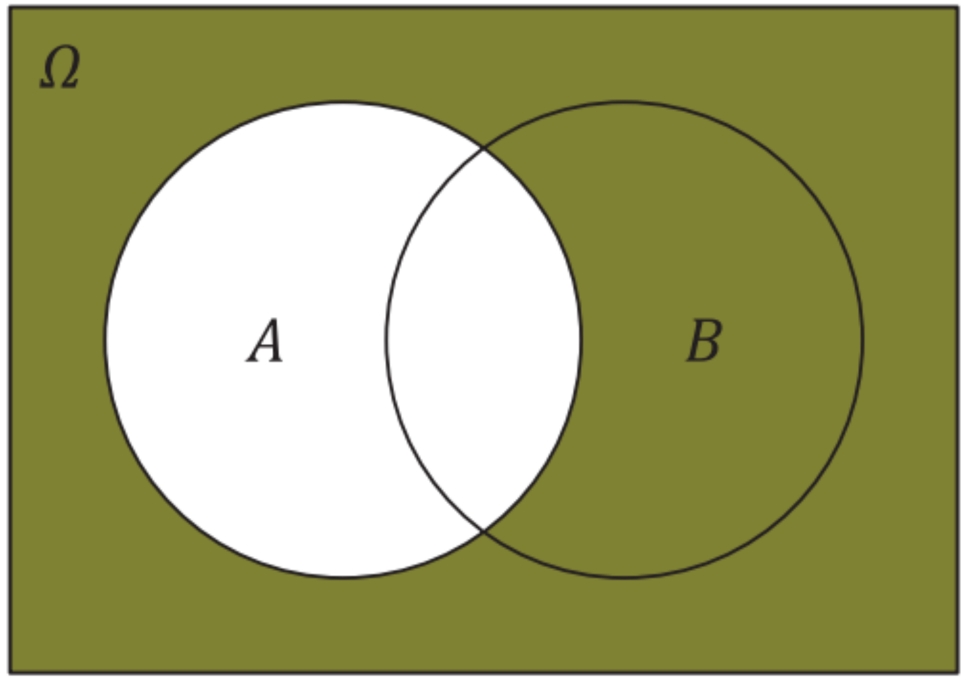
$A \cup B = \{x: x \in A \text{ or } x \in B\}$

Intersection



$A \cap B = \{x: x \in A \text{ and } x \in B\}$

Complement

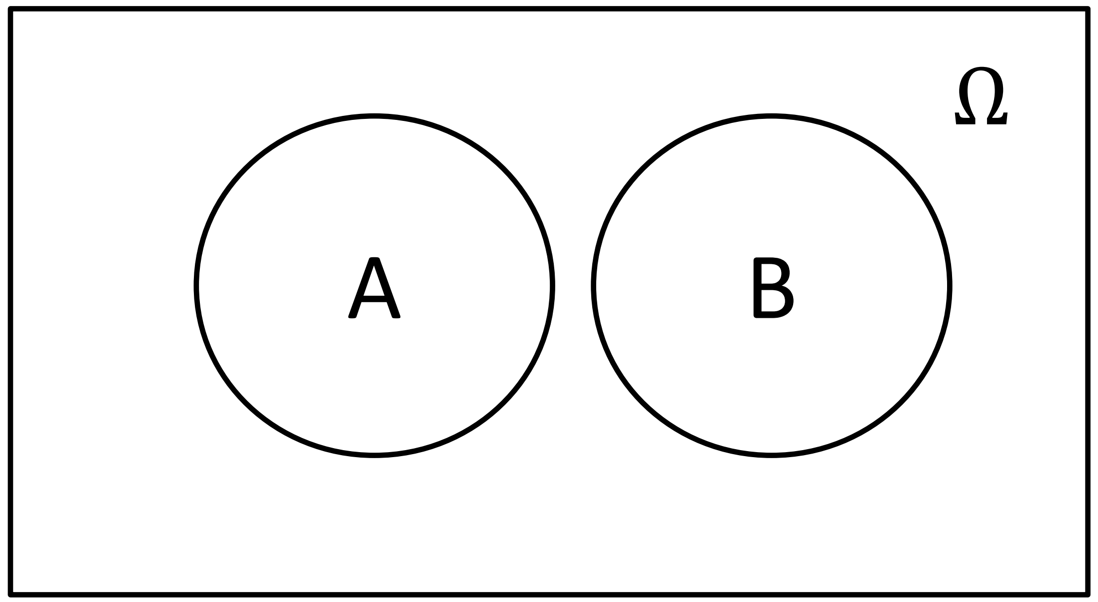


$A^C = \{x: x \notin A\}$

Cardinal, J. S., Journal of Computer Applications in Archaeology (2019).

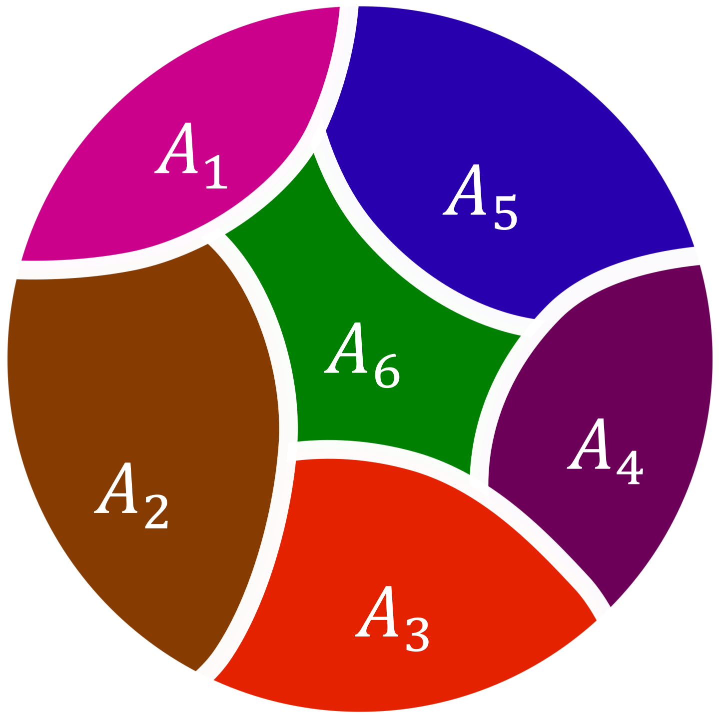
Disjoint Sets and Partitions

Disjoint Sets



$A \cap B = \emptyset$, i.e., A and B have no elements in common

Partition



If $A_1, A_2, \ldots$ are pairwise disjoint and $\bigcup_i A_i = \Omega$

A More Formal Definition of Probability

- A function P is called a **probability measure** (on a sample space Ω) if:
  1. $0 \leq P(A) \leq 1$ for all $A \in \Omega$
  2. $P(\emptyset) = 0$ and $P(\Omega) = 1$
  3. If $A_1, A_2, \ldots$ is a disjoint sequence of sets, then

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

- P(A) is referred to as the **probability** of event A

Some Properties of Probability

- **Complement Property:**
$$P(A^c) = 1 - P(A)$$

- **Subset Property:** If $A \subset B$, then $P(A) \leq P(B)$

- **Set Subtraction property:**
$$P(A \cap B^c) = P(A) - P(A \cap B)$$

- **Union Property:**
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- **Exercise:** Prove any/all of the above properties using the definition of a probability measure

Probability on a Finite Set

- Let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_N\}$ be a finite set. Suppose that $P(\omega_i) = p_i$. Then, we can find $P(A)$ by

$$P(A) = \sum_{i:\omega_i \in A} p_i$$

- This gives us a way to compute probabilities for any event!

- **Special Case:** If all elements in Ω have equal probability, then

$$P(A) = \frac{n(A)}{N}$$

where n(A) denotes the number of elements of A

Examples

1. Suppose that a fair coin is thrown twice. What is the probability of getting at least one heads?

Solution:

- Our sample space is $\Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}$ and event of interest is $A = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}\}$
- Each element in Ω is equally likely
- Since A has three elements and Ω has four elements, $\mathrm{P}(A) = 3/4$

2. A lottery is operated as follows. From the numbers 1,…, 44, a person may choose any six for their lottery ticket (numbers do not need to be distinct). What is the probability of randomly selecting the winning ticket?

Counting Tools: Multiplication Principle

- From our last example, enumerating all possibilities is too tedious/infeasible
  - We need methods for counting to solve many probability problems

- **Multiplication Principle:** Suppose an experiment consists of k separate tasks, where the 1st task can be done n_1 ways, the 2nd task can be done n_2 ways, and so on. The experiment can be done in $n_1 \times n_2 \times \cdots \times n_k$ ways
  - **Example from last slide:**
    - The first number can be chosen 44 ways
    - The second number can be chosen 44 ways, and so on
    - Thus, the experiment can be done in 44^6 ways

Counting Tools: Ordering and Replacement Considerations

- Two key considerations in counting problems
  - **Ordering**: Does the order of the objects matter?
  - **Replacement**: Are objects taken with or without replacement?

- Our lottery example was one that was **ordered, with replacement**
  - Let's consider other variations on this problem

- **Variation 1: Ordered, without replacement**
  - Consider that lottery tickets cannot include repeat numbers
  - How many possible tickets are there?

$$44 \times 43 \times 42 \times 41 \times 40 \times 39$$

Counting Tools: Ordering and Replacement Considerations

- **Variation 2: Unordered, without replacement**
  – Suppose that the order of the lottery ticket numbers does not matter and tickets cannot include repeat numbers
  – How many possible tickets are there?

$$\frac{44 \times 43 \times 42 \times 41 \times 40 \times 39}{6 \times 5 \times 4 \times 3 \times 2 \times 1}$$

  – Denoted by $\binom{44}{6}$, read as "44 choose 6"

- **Variation 3: Unordered, with replacement**
  – Suppose that the order of the lottery ticket numbers do not matter
  – How many possible tickets are there?
    - (Optional exercise)

What is the probability that two people in our class (29 people) share a birthday?

(0, 0.2]	0%
(0.2, 0.4]	0%
(0.4, 0.6]	0%
(0.6, 0.8]	0%
(0.8, 1]	0%

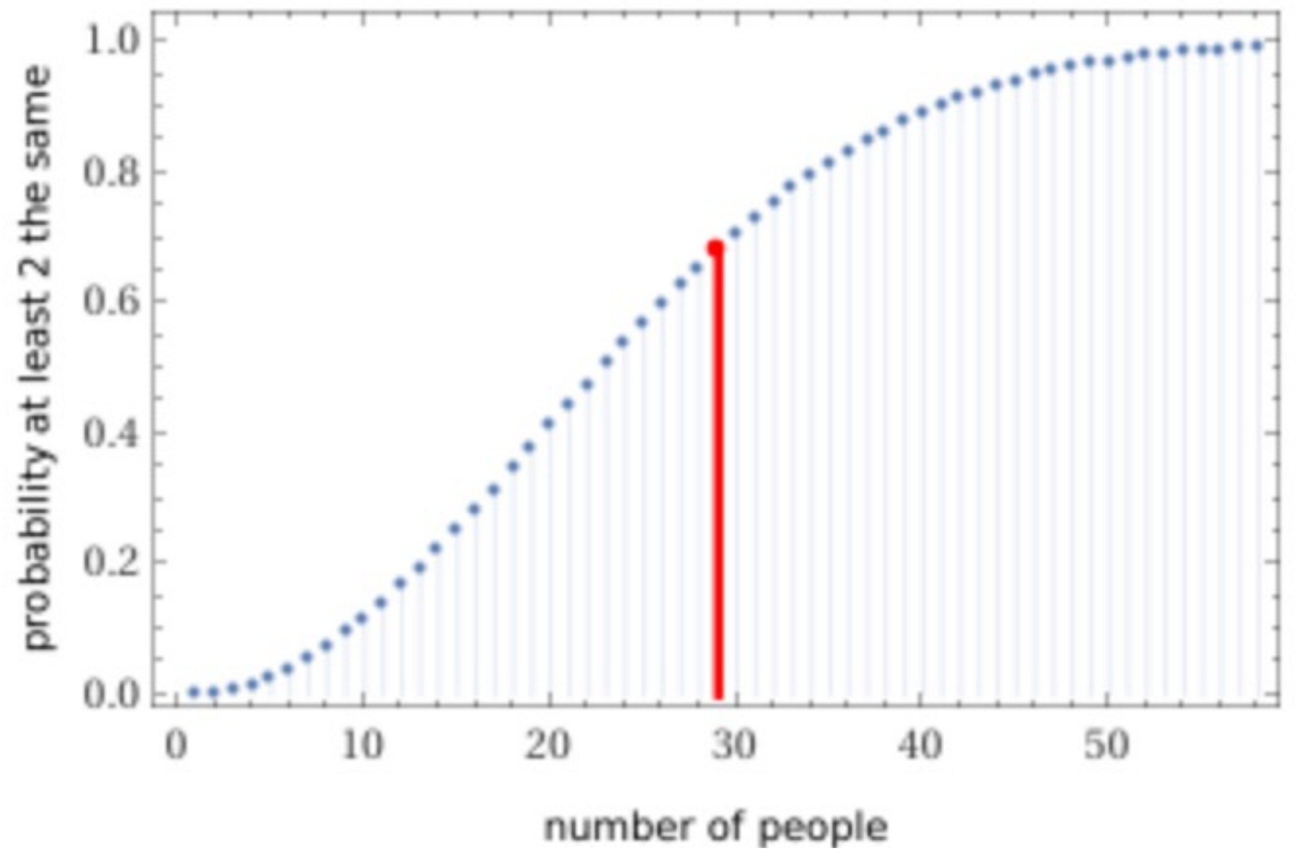
Solution

- P(at least one birthday shared) = 1 – P(no birthdays shared)
- How many possible birthday combinations are there?
  - 365^{29}
- How many possible ways in which nobody shares a birthday?
  - $365 \times 364 \times \cdots \times (365-29)$
- Therefore,

$$\text{P(no birthdays shared)} = \frac{365 \times 364 \times \cdots \times (365-28)}{365^{29}}$$

and so

$$\text{P(at least one birthday shared)} = 1 - \frac{365 \times 364 \times \cdots \times (365-28)}{365^{29}} \approx 0.681$$



http://wolframalpha.com

Part 2

What is the probability that a randomly selected individual in the USA is 6 feet tall?

Nobody has responded yet.

Hang tight! Responses are coming in.

Now, suppose that we are given the information that the randomly selected individual is an NBA (baseketball) player. What is the probability that he is at least 6 feet tall?

Nobody has responded yet.

Hang tight! Responses are coming in.

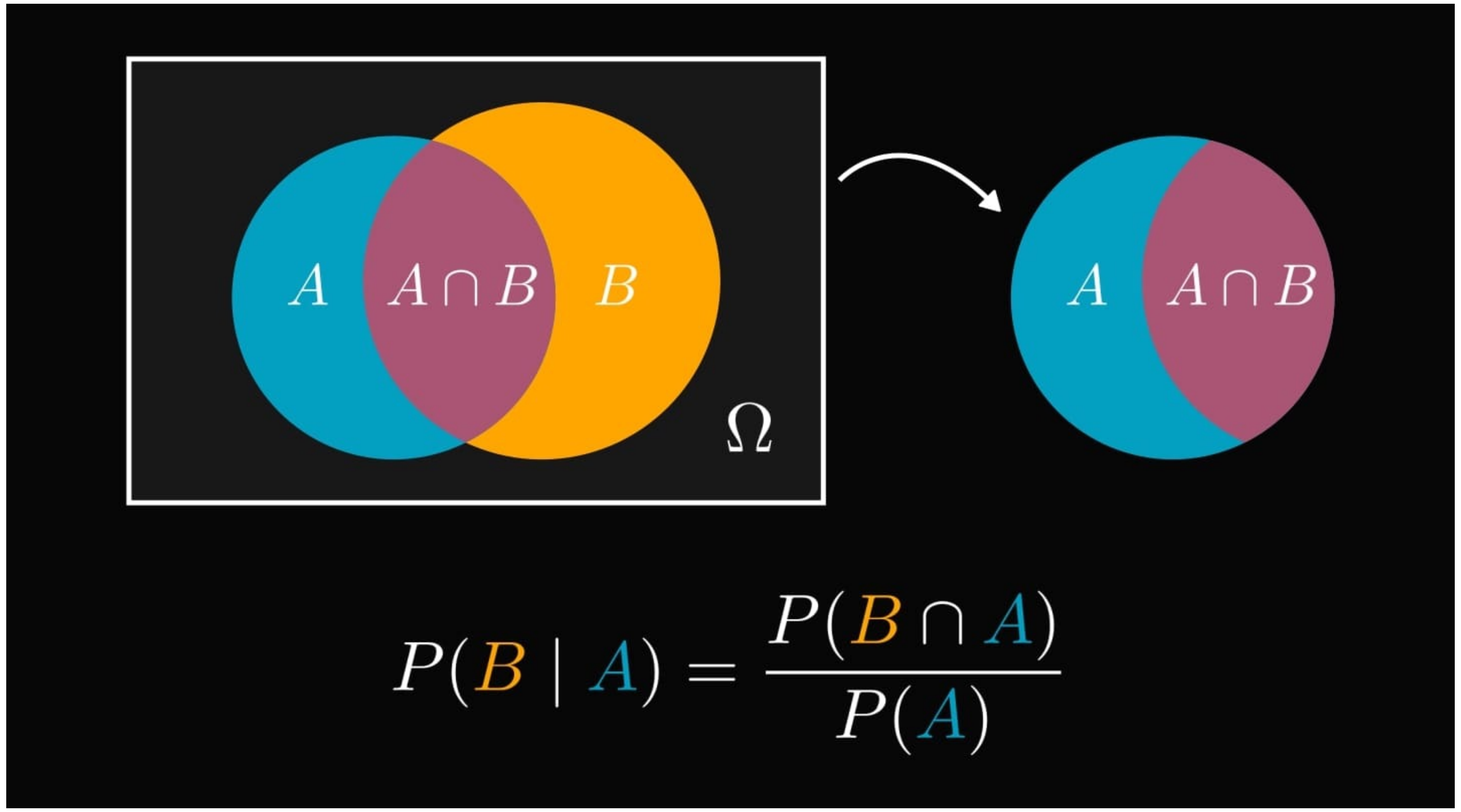
Conditional Probability

- **Conditional probability**: Probability of an event B given a second event A (i.e., knowing that event A happened), denoted by $P(B \mid A)$

- **Previous Example:**
  - B: selecting someone that is at least 6 feet tall
  - A: selecting an NBA player
  - In the first poll, we wanted to compute $P(B)$
  - In the second poll, we wanted to compute $P(B \mid A)$

Interpretation of Conditional Probability

- Interpretation of P(B|A):
  - We know that A has happened
  - That is, we are reducing the sample space from Ω to A

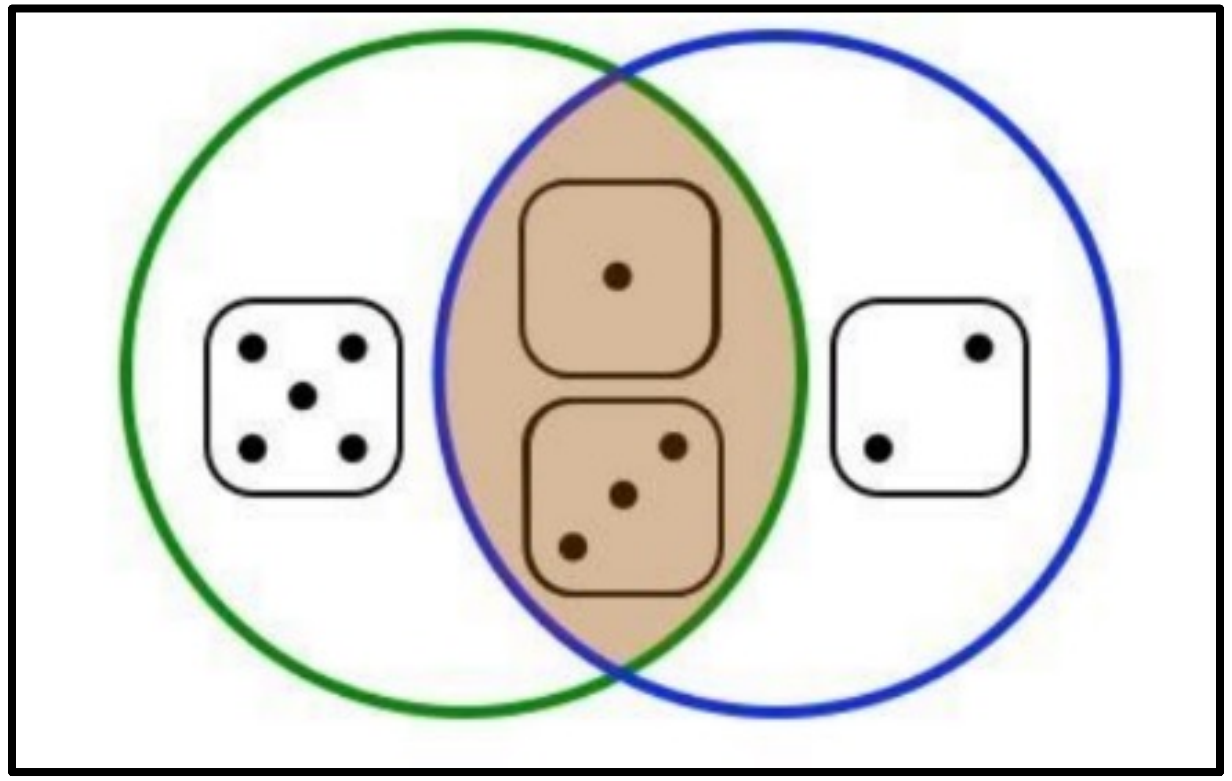


https://tivadardanka.com/blog/conditional-probability

Example

- **Question:** Consider the experiment of rolling a fair die. What is the probability of rolling a number less than 4 conditional on rolling an odd number?

- **Solution:**
  - Sample Space: $\Omega = \{1, 2, 3, 4, 5, 6\}$
  - A: Rolling an odd number, i.e. $\mathrm{A} = \{1, 3, 5\}$
  - B: Rolling a number less than 4, i.e. $\mathrm{B} = \{1, 2, 3\}$
  - $\mathrm{P(B|A)} = \frac{2}{3}$



Independence

- **Independence**: Events A and B are independent if and only if $P(A \cap B) = P(A)P(B)$

- **Interpretation:** Event B does not give any information about event A, and vice versa
  - **Reason:**

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

- **Application:** This property is crucial in statistical inference problems (later!)

Law of Total Probability

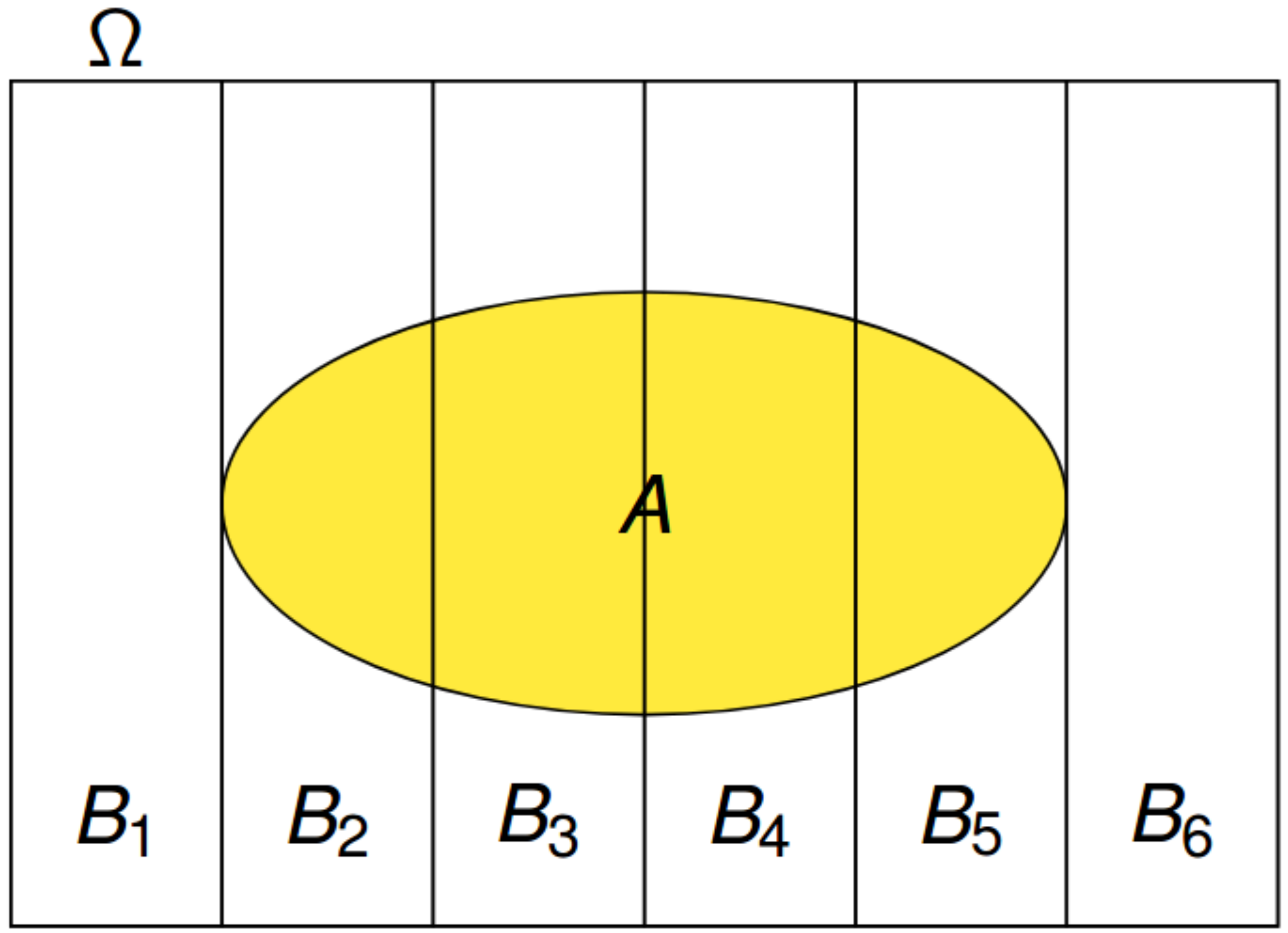
- **Motivation:** To find the probability of an event, it often helps to decompose the event into simpler events

- **Law of Total Probability:** Suppose that $B_1, B_2, \ldots$ form a partition of the sample space. Then,

$$P(A) = \sum_i P(A \cap B_i)$$

Further, by conditional probability,

$$P(A) = \sum_i P(A|B_i)P(B_i)$$



https://towardsdatascience.com/probability-theory-continued-infusing-law-of-total-probability-4abfca6e65bb

Example

- Suppose you have three bags that each contain 100 marbles.
  - Bag 1 has 75 **red** and 25 **blue** marbles;
  - Bag 2 has 60 **red** and 40 **blue** marbles;
  - Bag 3 has 45 **red** and 55 **blue** marbles.

- You choose one of the bags at random and pick a marble also at random from the chosen bag. What is the probability that the chosen marble is red?

Solution

- **Defining events:**
  - A: Chosen marble is red
  - B_i: Bag i is chosen.
- **Known Quantities:**
  - $P(A \mid B_1) = 0.75$
  - $P(A \mid B_2) = 0.60$
  - $P(A \mid B_3) = 0.45$
  - $P(B_i) = \frac{1}{3}$ for all i since we choose the bag at random.
- Using the law of total probability, we can write

$$P(A) = P(A \mid B_1)P(B_1) + P(A \mid B_2)P(B_2) + P(A \mid B_3)P(B_3) = 0.60.$$

Further Reading

- **Reference:** *Statistical Inference*, 2nd Edition, by G. Casella and R. L. Berger, Duxbury: Thomson Learning Inc (2002).

- **Topics:** Further details on
  - Axioms of probability
  - Properties of probability measures
  - Counting techniques
  - Probability in discrete and continuous settings
  - and other topics