PODS Lab 3: Probability Theory

Hamza Alshamy

Center for Data Science, NYU
ha2486@nyu.edu

02/07/25

1. **Basics of Probability**
2. **Conditional Probability**
3. **Bayes' Theorem**
4. **Random Variables**

Motivation: Why Do We Need Probability?

- **Goal:** Reason quantitatively about uncertain phenomena
- **Questions:**

Motivation: Why Do We Need Probability?

- **Goal:** Reason quantitatively about uncertain phenomena
- **Questions:**
  - Will it rain tomorrow?
  - Will the value of a stock drop tomorrow?
  - Who will win the 2026 FIFA world cup?

Main Concepts in Probability

Example: Selecting a Random Locker

Scenario: A gym has 6 lockers labeled 1 through 6. You randomly choose one to store your backpack.

- **Sample Space (Ω)**: The set of all possible outcomes.
  - The sample space is:

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

- **Event**: A subset of the sample space.
  - The event "Choosing an even-numbered locker" is:

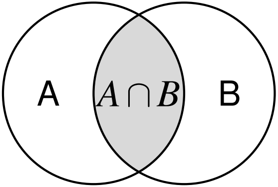
$$B = \{2, 4, 6\}$$

- **Probability (Frequentist)**: The long-run relative frequency of an event occurring.
  - Since 3 out of the 6 lockers are even-numbered, the probability is:

$$P(B) = \frac{\text{\# of outcomes in } B}{\text{\# of all possible outcomes}} = \frac{3}{6} = 0.5 \quad \text{(or 50\%)}$$

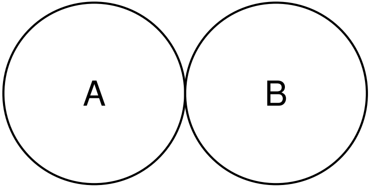
Basic Probability Rules: Union ($\cup$ "Or")

General Union Rule:



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

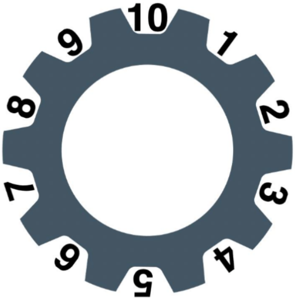
Mutually Exclusive Case:



$$P(A \cup B) = P(A) + P(B) \quad \text{(if } A \cap B = 0)$$

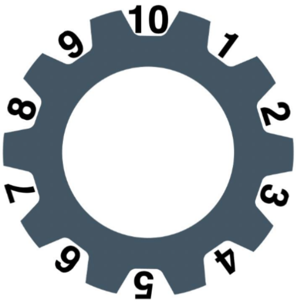
Mutually Exclusive Events

- **Goal:** Spin the gears
- **Sample Space:** $\Omega = \{1, 2, \ldots, 10\}$
- **Event:** The number at which the gear stops
- $P(8 \text{ and } 3) = P(8 \cap 3) = ?$



Mutually Exclusive Events

- **Goal:** Spin the gears
- **Sample Space:** $\Omega = \{1, 2, \ldots, 10\}$
- **Event:** The number at which the gear stops
- $P(8 \text{ and } 3) = P(8 \cap 3) = 0$
  - **Mutually exclusive:** Cannot happen at the same time!



Conditional Probability: Updating Beliefs

Key Idea: Conditional probability helps us revise our beliefs when new information is available.

Conditional Probability Formula

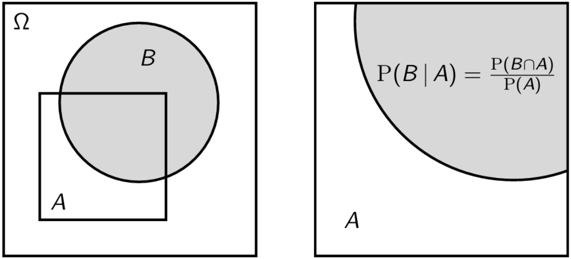
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

In words: What is the probability B Given that A has happened? how does this affect our belief in B?

Conditional Probability: Where does the formula come from?

Remember!

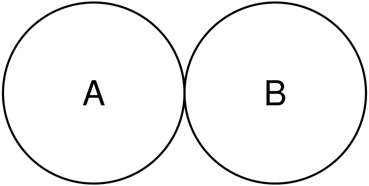
$$P(B) = \frac{\text{\# of outcomes in B}}{\text{\# of all outcomes}} \Rightarrow P(B \mid A) = \frac{P(B \cap A)}{P(A)}$$



- We're **restricting** the sample space to only cases where event A has occurred.
- In a sense, we are **zooming in** on the sample space!

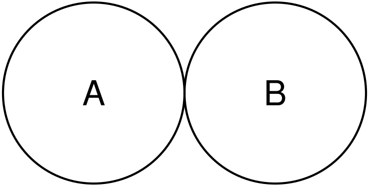
Relationship Between Independence and Mutual Exclusivity

- **Question:** If two events A and B are mutually exclusive, does that mean they are independent?



Relationship Between Independence and Mutual Exclusivity

- **Question:** If two events A and B are mutually exclusive, does that mean they are independent?



- **No!** If two events A and B are mutually exclusive, they are **NOT** independent.
- **Why?**
  - In this case, if we know event A happened, we know for a fact that B cannot happen. We are **gaining** information by knowing that event A happened.

Independence: Gaining No Extra Information

Definition: Two events A and B are **independent** if knowing that A occurred **does not** provide any information about whether B occurred.

1. Without Independence:

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

Rewriting:

$$P(A \cap B) = P(A) \cdot P(B \mid A)$$

2. With Independence:

$$P(B \mid A) = P(B)$$

Rewriting:

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

Summary of Probability

(1) Addition Rule

General: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Mutually Exclusive: $P(A \cup B) = P(A) + P(B)$

(2) Multiplication Rule

General: $P(A \cap B) = P(A) \cdot P(B \mid A)$
Independent: $P(A \cap B) = P(A) \cdot P(B)$

(3) Conditional Probability

General: $P(B \mid A) = \frac{P(A \cap B)}{P(A)}$
Independent: $P(B \mid A) = P(B)$

From Conditional Probability to Bayes' Rule

- **Bayes' Rule:** Allows us to "reverse" conditional probabilities.

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

Rewriting:

$$P(A \cap B) = P(B \mid A) \times P(A)$$

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

Rewriting:

$$P(A \cap B) = P(A \mid B) \times P(B)$$

Deriving Bayes' Rule

$$P(B \mid A) \times P(A) = P(A \mid B) \times P(B)$$

$$\implies \quad P(A \mid B) = \frac{P(B \mid A) \times P(A)}{P(B)}$$

Some Terminology

Bayes' Rule with Interpretation

$$\underbrace{P(A \mid B)}_{\text{Posterior}} = \frac{\overbrace{P(B \mid A)}^{\text{Likelihood}} \times \overbrace{P(A)}^{\text{Prior (of } A)}}{\underbrace{P(B)}_{\text{Prior (of } B)}}$$

- **Prior:** Our belief about event A before seeing the data (e.g., how likely we thought A was initially).
- **Likelihood:** The probability of observing the data B assuming A is true.
- **Posterior:** Our updated belief about A after seeing the data B.

Example: Updating Beliefs About a Coin

Scenario:

- **Prior Belief:** The coin is fair.
  - $P(\text{Heads}) = 0.5$
  - $P(\text{Tails}) = 0.5$
- **New Evidence:** We flip the coin 60 times and observe heads 70% of the time (42 heads).
- **Updated Belief (Posterior):**
  - Based on the data, we update our belief: The coin may be biased toward heads.

Key Idea: Prior $\rightarrow$ New Data $\rightarrow$ Posterior

What is a Random Variable (RV)?

- **Random Variable (RV):** A *function* that *maps* all elements of a sample space to the real numbers
  - **Mathematically:** $\Omega \to \mathbb{R}$
  - **RV Notation:** $\mathfrak{X}, X, \tilde{x}$
- **Intuition:** Sometimes, the sample space is not represented in numbers
  - The role of RVs is to map all those elements in the sample space Ω (that are, at times, not numbers) into real numbers $\mathbb{R}$
- **Relevance:**
  1. We will use RVs to describe probability distributions (next week!)
  2. We use RVs to deal with probabilities

Random Variables as Functions?

Example: Flipping a Coin ($\Omega \to \mathbb{R}$)

- **Scenario:** You flip a coin and want to reason about the outcome quantitatively.
- **Sample Space:**

Random Variables as Functions?

Example: Flipping a Coin ($\Omega \to \mathbb{R}$)

- **Scenario:** You flip a coin and want to reason about the outcome quantitatively.
- **Sample Space:** $\Omega = \{\text{Heads, Tails}\}$
- Define a Random Variable $\mathfrak{X}$ that "transforms" the elements in the sample space to real numbers:

Random Variables as Functions?

Example: Flipping a Coin ($\Omega \to \mathbb{R}$)

- **Scenario:** You flip a coin and want to reason about the outcome quantitatively.
- **Sample Space:** $\Omega = \{\text{Heads, Tails}\}$
- Define a Random Variable $\mathfrak{X}$ that "transforms" the elements in the sample space to real numbers:

$$\mathfrak{X}(\omega) := \begin{cases} 0, & \text{if } \omega = \text{Heads} \\ 1, & \text{if } \omega = \text{Tails} \end{cases}$$

- **In words:** The RV $\mathfrak{X}$ takes on the value 0 if the outcome is Heads and 1 if Tails
  - There you go! We have successfully mapped outcomes from the sample space Ω to real numbers $\mathbb{R}$

Random Variables and Probability

Example: Flipping a Coin

- **Question:** What is the probability of observing the outcome Heads (0)?

$$P_{\mathfrak{X}}(0) = P(\mathfrak{X} = 0) = \frac{\text{\# of outcomes in our event}}{\text{\# of all outcomes}} = \frac{1}{2} = 0.5$$

- **In words:** "What is the probability of the random variable $\mathfrak{X}$ taking the value 0?"
  - **NOT:** "What is the probability of getting heads?"
- **Key Idea:** The way we deal with probability it through RVs.
  - We ask "what is the probability of the random variable taking the value(s)?"

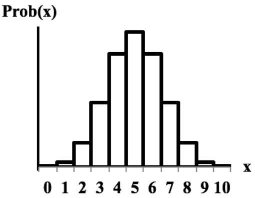
Two Types of Random Variables (RVs)

(1) Discrete

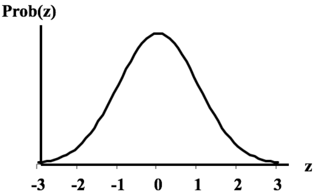
- Countable finite numbers of values
- **Example:** Goals scored in a soccer match

(2) Continuous

- Can take an infinite number of values, intervals
- **Example:** Marathon time



Binomial Distribution
Discrete Data & Discrete Probability Curve



Standard Normal Distribution
Continuous Data and Continuous Probability Curve