

Introduction to Probability

What is Probability?

- ▶ Numerical description of how likely an event is to occur.
- ▶ Formula: $P(A) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$
- ▶ $0 \leq P(A) \leq 1$

Key Concepts

Sample Space: Set of all possible outcomes.

Complement of an Event (A' or A^c): $P(A) + P(A') = 1$
($A' = \text{not } A$)

Odds in Favor: $\frac{\text{Number of favorable cases}}{\text{Number of unfavorable cases}}$

Odds Against: $\frac{\text{Number of unfavorable cases}}{\text{Number of favorable cases}}$

Equally Likely Events: $P(A) = P(B)$

Mutually Exclusive and Disjoint Events

- ▶ Mutually Exclusive: Two events are called mutually exclusive if the occurrence of one event means the other event cannot occur at the same time.

i.e. $P(A \cap B) = 0$

- ▶ Exhaustive Set of Events: $A_1 \cup A_2 = \text{Sample Space}$
- ▶ Independent Events: Events not depending on other events.

Addition Principle of Probability

1. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Non-Mutually Exclusive)
2. $P(A - B) = P(A) - P(A \cap B)$
3. For A, B, C three events:
 - ▶ $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$
4. $P(\text{Exactly two of the events}) = P(A \cap B) + P(B \cap C) + P(C \cap A) - 3P(A \cap B \cap C)$
5. $P(\text{Exactly one of the events}) = P(A) + P(B) + P(C) - 2P(A \cap B) - 2P(A \cap C) - 2P(B \cap C) + 3P(A \cap B \cap C)$

Multiplication Principle

- ▶ $P(A|B) = \frac{P(A \cap B)}{P(B)}$ (A happening given B already happened)
- 1. $P(A \cap B) = P(A) \cdot P(B|A)$
- 2. $P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$

Total Probability Theorem and Bayes' Theorem

These theorems handle partitioned events and updating beliefs with new evidence.

- **Total Probability Theorem:** $P(A) = \sum_{i=1}^n P(B_i) \cdot P(A|B_i)$, where B_i are exhaustive and mutually exclusive partitions. Example: $P(\text{Positive Test}) = P(\text{Positive|Disease}) * P(\text{Disease}) + P(\text{Positive|No Disease}) * P(\text{No Disease})$.
- **Bayes' Theorem:** $P(B|A) = \frac{P(A|B) \cdot P(B)}{\sum_{i=1}^n P(B_i) \cdot P(A|B_i)}$. Reverses conditionals to find posteriors from priors.
In data science, it's foundational for Bayesian inference, e.g., updating disease probability given a test result.

Bayes' is powerful for spam detection or recommendation systems, incorporating prior knowledge.

Random Variables (RV)

- ▶ Random Variable: A real-valued function which assigns a real number to each sample point in the sample space.
- ▶ Types:
 - ▶ Categorical R.V. (e.g., Gender of a person)
 - ▶ Numerical R.V.
 - ▶ Discrete R.V. (e.g., Number of people in family)
 - ▶ Continuous R.V. (e.g., Salary, Loan amount)

Random Variables (RV)

- **Definition:** A function assigning a real number to each sample point. Example: X = number of heads in two coin flips (values 0,1,2).
- **Types:**
 - **Categorical RV:** Non-numerical, e.g., gender (male/female).
 - **Numerical RV:**
 - **Discrete:** Countable values, e.g., family size (0,1,2,...).
 - **Continuous:** Infinite values in a range, e.g., height (any real number between 150-200 cm).

In data science, RVs represent features in datasets, enabling statistical modeling.

Probability Distribution Function

- ▶ Gives the probability of all the possible outcomes of any random variable.
- ▶ Two functions used to describe the probability distribution:
 - ▶ Probability Density Function (PDF) - Continuous R.V.
 - ▶ Probability Mass Function (PMF) - Discrete R.V.