1. Probability Distributions

Definition

A **probability distribution** is a mathematical function that provides the probabilities of occurrence of different possible outcomes in an experiment. It is classified into **discrete** and **continuous** distributions.

1.1 Discrete Probability Distributions

A **discrete random variable** takes on a countable number of distinct values. Examples include the number of heads in coin flips or the number of defective items in a batch.

Common discrete distributions include:

- **Bernoulli Distribution**: Represents a single trial with two outcomes (Success/Failure) with a probability **p** of success.
- **Binomial Distribution**: Models the number of successes in a fixed number of independent Bernoulli trials.
- **Poisson Distribution**: Estimates the number of events occurring in a fixed interval of time or space with a known average rate and independently of the time since the last event.

1.2 Continuous Probability Distributions

A **continuous random variable** can take any value within a given range. Examples include the height of individuals or the time required to complete a task.

Common continuous distributions include:

- **Normal Distribution**: Symmetric, bell-shaped distribution described by its mean and standard deviation. The empirical rule states approximately 68% of data lies within 1 standard deviation, 95% within 2, and 99.7% within 3.
- **Exponential Distribution**: Often used to model the time between events in a Poisson process.
- **Uniform Distribution**: All outcomes are equally likely within a specific range.

1.3 Key Concepts

- **Probability Mass Function (PMF)**: For discrete variables, gives the probability that a random variable is exactly equal to some value.

- **Probability Density Function (PDF)**: For continuous variables, represents the relative likelihood of a random variable to take a particular value.
- **Cumulative Distribution Function (CDF)**: Describes the probability that a random variable will be less than or equal to a particular value.
- **Expected Value (Mean)**: The average or central value of the random variable.
- **Variance and Standard Deviation**: Measure the spread or dispersion of the distribution.

---

2. Conditional Probability

Definition

The **conditional probability** of an event **A**, given that **B** has occurred, is the probability of **A** occurring under the condition that **B** is true. It is denoted by **P(A|B)**.

Formula

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

Where:

- **P(A|B)** is the conditional probability of **A** given **B**.
- **P(A ∩ B)** is the joint probability of **A** and **B** occurring together.
- **P(B)** is the probability of **B** occurring.

Key Points

- If **A** and **B** are **independent**, then **P(A|B) = P(A)**.
- If **P(B) = 0**, the conditional probability **P(A|B)** is **undefined**.
- **Mutually exclusive events** (cannot occur together) have a **P(A ∩ B) = 0**.

---

3. Bayes' Theorem

Definition

Bayes' Theorem provides a way to update the probability of a hypothesis based on new evidence. It relates the conditional and marginal probabilities of random events.

Formula

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

Where:

- **P(A|B)**: The probability of event **A** given **B** (posterior probability).
- **P(B|A)**: The probability of event **B** given **A** (likelihood).
- **P(A)**: The probability of event **A** (prior probability).
- **P(B)**: The probability of event **B** (marginal probability).

Applications

- **Medical Testing**: Calculating the probability of having a disease given a positive test result.
- **Spam Filtering**: Determining whether an email is spam based on certain keywords.
- **Risk Assessment**: Evaluating the probability of risks given historical data.
