Self Assignment: Basic Probability Concepts I

2024

Introduction

Probability is a measure of how likely an event is to occur. In computer science, probability plays a crucial role in areas such as machine learning, cryptography, and algorithms. This note introduces the fundamental concepts of probability, using simple examples to explain each concept clearly.

1 Sample Space and Events

The **sample space** (denoted as S) is the set of all possible outcomes of an experiment. An **event** is a subset of the sample space. For example, if you flip a coin, the sample space is:

$$S = \{\text{Heads}, \text{Tails}\}$$

An event could be flipping a head, so the event is:

$$E = \{\text{Heads}\}$$

Example: Consider rolling a six-sided die. The sample space is:

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

An event could be rolling an even number, so the event is:

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

2 Probability of an Event

The probability of an event E is a number between 0 and 1 that describes how likely it is for the event to occur. The probability of an event is given by the formula:

$$P(E) = \frac{\text{Number of favorable outcomes}}{\text{Total number of outcomes}}$$

For a fair six-sided die, the probability of rolling a 4 is:

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

Properties of Probability:

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$ (the probability of the entire sample space is 1)
3. $P(\emptyset) = 0$ (the probability of the empty set is 0)

3 Union and Intersection of Events

For two events A and B in the sample space S, the **union** of A and B (denoted $A \cup B$) represents the event that either A or B occurs. The **intersection** of A and B (denoted $A \cap B$) represents the event that both A and B occur.

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

Example: Let A be the event of rolling an even number, and B be the event of rolling a number greater than 4 when rolling a six-sided die.

$$A = \{2, 4, 6\}, \quad B = \{5, 6\}$$

The union of A and B is:

$$A \cup B = \{2, 4, 5, 6\}$$

The intersection of A and B is:

$$A \cap B = \{6\}$$

4 Conditional Probability

The **conditional probability** of event A given that event B has occurred is denoted by $P(A|B)$ and is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) \neq 0$$

Example: Consider a deck of 52 playing cards. Let A be the event of drawing a red card, and B be the event of drawing a heart. Since all hearts are red, we have:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$$

This means that if you know a heart was drawn, then the probability of drawing a red card is 1.

5 Independence of Events

Two events A and B are **independent** if the occurrence of one does not affect the probability of the other. Mathematically, A and B are independent if:

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

Example: If you roll a die and flip a coin, the outcome of the die roll does not affect the outcome of the coin flip. Hence, rolling a 4 and flipping a heads are independent events.

6 Bayes' Theorem

Bayes' Theorem is a fundamental result in probability theory that allows us to reverse conditional probabilities. It is given by:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \quad P(B) \neq 0$$

Bayes' Theorem is especially useful in machine learning and decision-making processes.

Example: Suppose 1% of people have a rare disease, and a test for the disease is 90% accurate (i.e., 90% of people with the disease test positive, and 90% of people without the disease test negative). What is the probability that a person who tests positive actually has the disease?

Let D be the event that the person has the disease, and T be the event that the person tests positive. We are interested in $P(D|T)$. Using Bayes' Theorem:

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)}$$

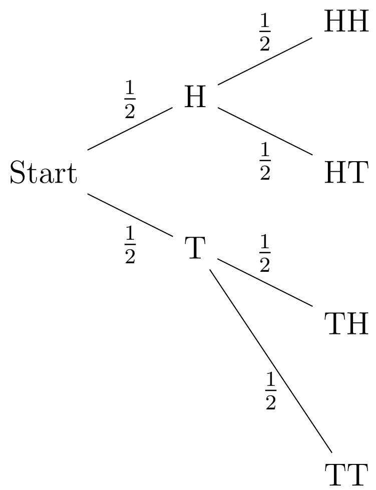
Where:

$$P(T) = P(T|D)P(D) + P(T|\neg D)P(\neg D)$$

Substitute the given values to calculate the result.

7 Illustration: Probability Tree Diagram

A probability tree diagram can help visualize conditional probabilities. Here's a simple example of a tree for flipping a coin twice:



Each branch represents an outcome, and the probabilities of the outcomes can be multiplied to find the total probability for each path in the tree.

8 Random Variables

A **random variable** is a function that assigns numerical values to the outcomes of a random experiment. Random variables help us quantify uncertain events and analyze them mathematically. They can be classified into two types: **discrete** and **continuous** random variables.

8.1 Discrete Random Variables

A **discrete random variable** takes on a countable number of distinct values. These variables usually result from experiments with a finite or countable set of outcomes.

Example 1: Rolling a Die

Consider the random experiment of rolling a fair six-sided die. The sample space is:

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

Let X be the random variable that represents the outcome of the die roll. Here, X is a discrete random variable because it can take one of the six distinct values: $1, 2, 3, 4, 5$, or 6.

We can represent the probability distribution of X using a **probability mass function** (PMF), which specifies the probability for each value of the random variable:

$$P(X = x) = \frac{1}{6} \quad \text{for } x = 1, 2, 3, 4, 5, 6$$

Example 2: Number of Heads in a Coin Toss

Consider tossing a fair coin three times. The sample space for this experiment is:

$$S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$$

Let Y be the random variable representing the number of heads obtained. The possible values of Y are $0, 1, 2, 3$, since you can get between zero and three heads in three coin tosses.

The probability distribution of Y is given by:

$$P(Y = 0) = \frac{1}{8}, \quad P(Y = 1) = \frac{3}{8}, \quad P(Y = 2) = \frac{3}{8}, \quad P(Y = 3) = \frac{1}{8}$$

This table represents a **discrete probability distribution** because Y takes on a finite number of possible outcomes.

8.2 Continuous Random Variables

A **continuous random variable** takes on an uncountable number of values, usually within a certain range. These variables result from measurements where the outcome can be any value within an interval.

Example: Measuring Height

Let Z be the random variable representing the height of a person selected at random from a population. In this case, Z is a continuous random variable because height can take any value within a certain range (e.g., 150 cm to 200 cm). Unlike discrete variables, we cannot list all the possible values of Z because there are infinitely many potential values within this interval.

For continuous random variables, we use a **probability density function** (PDF) instead of a PMF. The PDF describes the likelihood of the random variable taking a specific value within an interval. For a continuous random variable Z, the probability that it takes a value within an interval $[a, b]$ is given by:

$$P(a \leq Z \leq b) = \int_a^b$$