Random Variable

A random variable is a variable whose value depends on the outcome of a random experiment. It acts like a bridge between the random world of chance (represented by the sample space) and the world of numbers. Here's a breakdown with examples:

Imagine a coin toss:

- The sample space (S) is {Heads, Tails}. These are the two possible outcomes of the experiment.

Random Variable Example 1: Number of Heads

- Let's define a random variable named "X" that represents the number of heads obtained when the coin is tossed.
- X can take on the values 0 (if tails) or 1 (if heads).
- This random variable assigns a numerical value to each outcome in the sample space.

There are two main types of random variables used in probability and statistics: discrete and continuous. Each type is suited for representing different kinds of random phenomena.

1. Discrete Random Variables:

Imagine a counter that can only click to specific whole numbers. A discrete random variable is like that counter – it can only take on a finite or countable number of distinct values. Here are some characteristics:

- **Values:** The variable can take on specific, separated values. For example, the number of times you win a coin toss in 3 attempts (0, 1, 2, or 3 wins).
- **Sample Space:** The set of all possible outcomes is finite or countable.
- **Examples:**

    - Number of successful attempts to fix a machine before it works (0, 1, 2, 3, etc.).
    - Number of people arriving at a bus stop in a minute (0, 1, 2, 3, etc.).
    - The outcome of rolling a die (1, 2, 3, 4, 5, or 6).

Probability Mass Function (PMF):

Now, imagine assigning a probability (like a percentage chance) to each color marble being drawn. The PMF helps you do this mathematically.

- **Definition:** The probability mass function (PMF), denoted by $f(x)$ or $P(X = x)$, is a function that assigns a probability value (between 0 and 1) to each possible value a discrete random variable (X) can take.
- **Properties:**
    - Non-negative probabilities: $f(x) \geq 0$ for all possible values of x (probability can't be negative).
    - Summation Rule: The sum of the probabilities for all possible values of x equals 1. $\Sigma f(x) = 1$ (ensures all outcomes are considered with a total probability of 1).

Example:

Consider rolling a fair die. Let X be the number appearing on the top face.

- X can take values 1, 2, 3, 4, 5, or 6.
- The PMF for X can be expressed as:
    - $f(1) = P(X = 1) = 1/6$ (probability of rolling a 1)
    - $f(2) = P(X = 2) = 1/6$ (probability of rolling a 2)

How PMF helps with Discrete Random Variables:

The PMF is like a roadmap for understanding the likelihood of each outcome in a discrete random experiment. It allows you to:

- **Calculate probability of specific outcomes:** Using the die example, the PMF helps you find the probability of rolling a 3 ($f(3) = 1/6$).

- **Compare likelihoods:** You can see that each number on the die has an equal probability (1/6) of appearing.
- **Visualize the distribution:** The PMF can be used to create a probability histogram, a bar graph showing the probability for each outcome.

The cumulative distribution function (CDF)

- **Definition:** The CDF, $F(x)$, represents the probability that the random variable X will take on a value less than or equal to a specific value x. In simpler terms, it tells you the chance of getting a value up to x, including x itself.
- **Notation:** $F(x) = P(X \leq x)$

Relationship with PMF and PDF:

- **Discrete Case:** For a discrete random variable, the CDF is related to the PMF ($f(x)$) by the following:
  - $F(x) = \Sigma f(t)$ for all values $t \leq x$ (summation of probabilities from the minimum value up to x).

2. Continuous Random Variable

Continuous random variables are like those endlessly flowing rivers – their values can take on any value within a specific, continuous range. Unlike a bucket with distinct water levels (discrete), a continuous random variable allows for infinite possibilities within that range. Here's a breakdown to understand them better:

Imagine measuring the following:

- The weight of a person (can take any value between a minimum and maximum weight).
- The time it takes to complete a marathon (can take any value within a certain range of seconds).
- The amount of rain on a given day (can take any value from 0 to a very large amount).

Probability Density Function (PDF):

Since a continuous random variable can take on any value, calculating the probability of getting an exact value (like a specific weight) is zero. The PDF comes to the rescue!

- **Imagine a wave:** The height of the wave at a point represents the relative likelihood of a value in that range. A higher wave indicates a more probable range of values.
- **Definition:** The probability density function (PDF), denoted by f(x), describes the relative likelihood of a continuous random variable (X) taking on a specific value within a certain range.
- **Properties:**
  - Non-negative: $f(x) \geq 0$ for all possible values of x (probability can't be negative).
  - Total area under the curve equals 1: This ensures that the probability of the variable falling somewhere within the entire range is 1 (like all the water eventually flows somewhere in the river).

The Cumulative Distribution Function (CDF):

In the realm of continuous random variables, the cumulative distribution function (CDF) acts like a bridge between the concept of probability and the continuous flow of values. While the probability density function (PDF) describes the relative likelihood of a specific value, the CDF tells you the probability of a continuous random variable taking on a value **less than or equal to** a specific value (x).

Relationship with PDF and PDF:

- **Continuous Case:** For a continuous random variable, the CDF is related to the PDF (f(x)) by the following:
  - $F(x) = \int f(t)\,dt$ from minus infinity to x (integration of the probability density function from minus infinity up to x).