

# Probability Distribution

## Algebraic Variables :

In **algebra**, a variable (like **x** or **y**) is a symbol that represents an unknown value. For example:

- In the equation  $x + 2 = 5$ , the variable **x** is unknown. We solve to find that **x = 3**.

Think of algebraic variables as placeholders for numbers.

---

## Random Variables in Statistics and Probability:

A Random Variable is a **set of possible values** from a **random experiment**.

In **statistics and probability**, a **random variable** represents the possible outcomes of a random event. Instead of being a single unknown value, it can take **different values based on chance**.

### Example of Random Variables:

#### 1. Rolling a Die:

- Experiment: Roll a die.
- Random Variable (X): The number that shows up.
- Possible Values of X: {1, 2, 3, 4, 5, 6}.

Here, **X** is the random variable because its value depends on the random roll of the die.

#### 2. Flipping a Coin:

- Experiment: Flip a coin.
- Random Variable (Y): The result of the flip.
- Possible Values of Y: {Heads, Tails}.

Y is a random variable because the result is random and can either be "Heads" or "Tails."

## Types of Random Variables?

Random variables can be broadly classified into **two types** based on the nature of their possible outcomes:

---

### 1. Discrete Random Variables

A **discrete random variable** can take on a **countable** number of distinct values (whole numbers or specific points).

#### Key Features:

- Outcomes are **separate** and **distinct**.
- Often arises in situations where you count outcomes.

#### Examples:

- **Rolling a Die:**  
The random variable XXX (outcome) has possible values: {1, 2, 3, 4, 5, 6}.
  - **Flipping a Coin:**  
The random variable YYY (result) has possible values: {Heads, Tails}.
  - **Number of Emails Received in a Day:**  
The random variable ZZZ can take values like: {0, 1, 2, 3, ...}.
- 

### 2. Continuous Random Variables

A **continuous random variable** can take on an **infinite number of values** within a given range. These values are typically measurements.

## Key Features:

- Outcomes are **not countable** (can be any value within a range).
- Often arises in situations where you measure something.

## Examples:

- **Height of a Person:**  
The random variable HHH can take values like: 160.5 cm, 170.2 cm, etc.
- **Time to Complete a Task:**  
The random variable TTT could be 5.3 seconds, 10.7 seconds, etc.
- **Temperature in a City:**  
The random variable WWW might range continuously between 20°C and 35°C.

## 1. What are Probability Distributions?

A probability distribution is a **list of all of the possible outcomes** of a **random variable** along with their corresponding probability values.

A **probability distribution** is a way of showing all the possible outcomes of a random event and how likely each one is to happen.

Think of it like a table or a graph that tells you:

- **What can happen** (possible outcomes).
- **How often it might happen** (probabilities).

---

### Example: Rolling a Die

When you roll a die, the possible outcomes are:

1, 2, 3, 4, 5, 6.

Each outcome has an equal chance of happening (1 out of 6). So, the **probability distribution** is:

- Probability of rolling 1 =  $1/6 = 0.167$
  - Probability of rolling 2 =  $1/6 = 0.167$
  - Probability of rolling 3 =  $1/6 = 0.167$
  - ... (and so on for 4, 5, 6).
- 

### **Example: Flipping a Coin**

When flipping a coin, there are two possible outcomes: **Heads** and **Tails**. The **probability distribution** is:

- Probability of Heads =  $1/2 = 0.5$
  - Probability of Tails =  $1/2 = 0.5$ .
- 

### **Why Are Probability Distributions Useful?**

They help us:

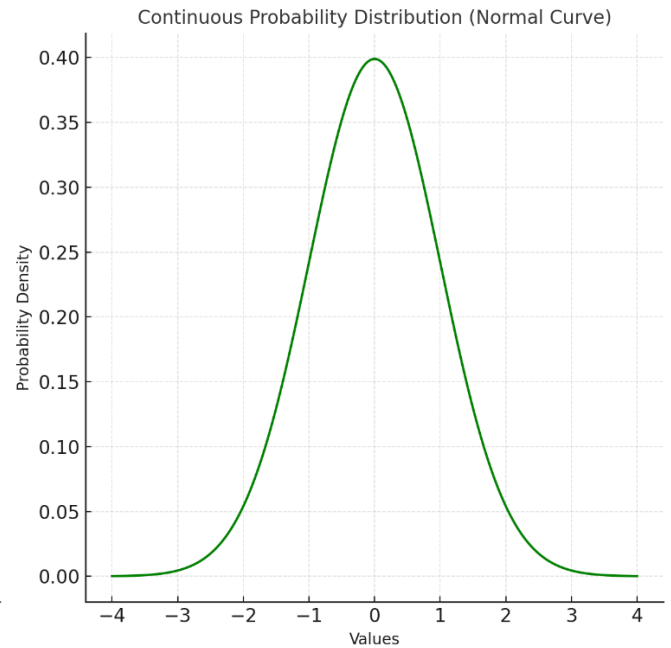
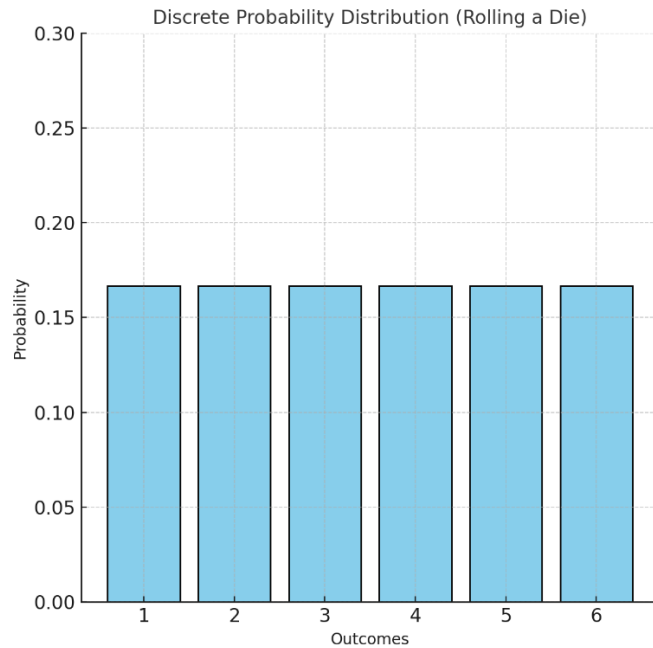
- Understand all possible outcomes of a random experiment.
- Calculate the chances of specific results happening.

For example, if you're playing a game or predicting the weather, knowing the probability distribution can help you make better decisions.

---

### **Visualizing Probability Distributions**

- **Discrete Distribution:** Bar graphs (e.g., for dice rolls or coin flips).
- **Continuous Distribution:** Curves (e.g., height or temperature measurements).



### Problem with Distribution?

In many scenarios, the number of outcomes can be much larger and hence a table would be tedious to write down. Worse still, the number of possible outcomes could be infinite,

in which case, good luck writing a table for that.

**Example** - Height of people, Rolling 10 dice together

**Solution** - Function?

What if we use a mathematical function to model the relationship between outcome and probability?

Note - A lot of time Probability Distribution and Probability Distribution Functions are used interchangeably.

# Probability distribution function

## Problem with Probability Distribution

Probability distributions are easy to understand when there are only a few possible outcomes. But in some cases:

1. **Too many outcomes:** For example, rolling 10 dice at the same time creates thousands of possible results.
2. **Infinite outcomes:** For example, the height of people can be any value, like 5.1 feet, 5.15 feet, 5.155 feet, and so on.

In these cases, writing a table to list all outcomes and their probabilities is **too hard or impossible**.

---

### Example:

1. **Height of People:** Everyone's height is slightly different, and there are infinite possibilities. Writing them all down isn't practical.
  2. **Rolling 10 Dice Together:** There are so many possible results that creating a table would take forever.
- 

### Solution: Use a Function

Instead of writing a table, we can use a **mathematical function** to represent the relationship between the outcomes and their probabilities.

- For example, in a **normal distribution** (like heights or test scores), we use this function:  
$$f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$
  
This function gives us the probability of any outcome directly without needing a table.
- 

### Important Note:

Sometimes, people use **Probability Distribution** and **Probability Distribution Function (PDF)** to mean the same thing.

- **Probability Distribution:** Describes all outcomes and their probabilities.
- **Probability Distribution Function (PDF):** A specific formula or function that calculates probabilities.

## Types of Probability Distribution function:

**Probability Density Function (PDF):**

**Probability Mass Function (PMF):**

### **Probability Mass Function (PMF):**

PMF stands for Probability Mass Function. It is a mathematical function that describes the probability distribution of a discrete random variable.

The PMF of a discrete random variable assigns a probability to each possible value of the random variable. The probabilities assigned by the PMF must satisfy two

conditions:

- a. The probability assigned to each value must be non-negative (i.e., greater than or equal to zero).
- b. The sum of the probabilities assigned to all possible values must equal 1.

### **Example: Rolling a Die**

Let's say we roll a fair 6-sided die. The random variable  $XXX$  represents the number on the die.

- The possible values of  $XXX$  are:  $\{1, 2, 3, 4, 5, 6\}$ .
- The PMF assigns a probability to each value.

Since the die is fair, the PMF is:


### **Verifying the Conditions:**

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

This means:

- $P(X = 1) = \frac{1}{6}$
- $P(X = 2) = \frac{1}{6}$
- ...
- $P(X = 6) = \frac{1}{6}$

### Verifying the Conditions:

1. All probabilities ( $\frac{1}{6}$ ) are greater than or equal to 0. 
2. The sum of probabilities is:

$$\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1.$$



---

### Why is PMF Important?

PMF helps us understand the chances of different outcomes for **discrete variables**, like dice rolls, coin flips, or number of customers arriving in an hour. It gives a clear picture of the probability distribution!



## What is the Cumulative Distribution Function (CDF)?

The Cumulative Distribution Function (CDF) is a function that gives the probability that a random variable  $X$  will take a value less than or equal to  $x$ .

In simple words, the CDF tells us:

- How much of the total probability is "accumulated" up to a specific value  $x$ .

The formula is:

$$F(x) = P(X \leq x)$$

This means  $F(x)$  gives the probability of  $X$  being less than or equal to  $x$ .

### Example: Rolling a Die

Let  $X$  be the outcome of rolling a fair 6-sided die, and we calculate the CDF for each value:

1.  $F(1) = P(X \leq 1) = P(X = 1) = \frac{1}{6}$
2.  $F(2) = P(X \leq 2) = P(X = 1) + P(X = 2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6}$
3.  $F(3) = P(X \leq 3) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{3}{6}$
4. Similarly,  $F(4) = \frac{4}{6}$ ,  $F(5) = \frac{5}{6}$ ,  $F(6) = 1$ .

**Key Idea:** The CDF always increases as  $x$  increases and reaches 1 at the maximum possible value of  $X$ .

## Probability Density Function (PDF)?

The Probability Density Function (PDF) is a mathematical function used to describe the probability distribution of a continuous random variable. Unlike discrete variables, we don't assign exact probabilities to specific values but describe the "density" of probability over a range of values.

---

### 1. Why Probability Density and not just Probability?

For continuous random variables:

- There are infinite possible values within any range (e.g., heights, temperatures).
- The probability of any exact value (like exactly 5 feet) is 0 because the range is infinitely small.

Instead, we use density to describe how the probabilities are distributed over a range of values.

- Example: The probability of height being between 5.1 and 5.2 feet can be calculated using the PDF.

---

## 2. What Does the Area of the PDF Graph Represent?

The area under the PDF curve represents probability for a range of values.

- The total area under the curve equals 1 (because the total probability is always 1).
- The area between two points (e.g., from  $a$  to  $b$ ) gives the probability of the random variable falling in that range.

---

## 3. How to Calculate Probability? (Hinglish Explanation)

### 1. Step 1: Identify the Range

Decide kis range ke liye probability nikalni hai (e.g.,  $P(a \leq X \leq b)$  ).

### 2. Step 2: Use the PDF Function

Aapko jo mathematical function diya gaya hai (PDF), usko use karke graph ke neeche ka area calculate karo.

### 3. Step 3: Integrate the PDF

Continuous variables ke liye integration ka use karte hain to find the area under the curve.

$$P(a \leq X \leq b) = \int_a^b f(x) dx$$

Yahaan  $f(x)$  PDF ko represent karta hai.

## 4. How is graph calculated?

### Density Estimation?

Density estimation is a way to figure out the shape of the underlying distribution of data points. It helps us understand where the data is concentrated and how it is spread out.

In simple words:

- It estimates the Probability Density Function (PDF) from a set of data.
  - It's like guessing the pattern of data by looking at where most of the points are and how they're distributed.
- 

### Why Use Density Estimation?

Density estimation is useful for many tasks, such as:

1. Data Analysis: Understanding patterns in data.
  2. Data Visualization: Creating smooth curves or visual representations of data distributions.
  3. Machine Learning: Estimating probabilities or modeling the likelihood of events.
- 

### Types of Density Estimation

#### 1. Parametric Methods:

- Assume that the data follows a specific type of distribution, like a normal (bell-shaped) distribution.

- Example: If we think data fits a normal distribution, we estimate parameters like mean ( $\mu$ ) and standard deviation ( $\sigma$ ).

## 2. Non-Parametric Methods:

➔ But sometimes the distribution is not clear or it's not one of the famous distributions.

- No assumption about the type of distribution.
- The method estimates the distribution directly from the data.
- Examples:
  - Histogram Estimation: Divides data into bins and counts how many points fall into each bin.
  - Kernel Density Estimation (KDE): Creates a smooth curve by placing small bumps (kernels) at each data point.
  - Gaussian Mixture Models (GMMs): Combines multiple normal distributions to model complex data.

---

### Example: Density Estimation in Everyday Terms

Imagine you're running a coffee shop and want to know the busiest times of the day based on when customers arrive.

- Histogram Estimation: You divide the day into 1-hour bins and count the number of customers in each hour.

- Kernel Density Estimation: Instead of fixed bins, you create a smooth curve that shows peaks during busy times like morning or evening.

**Density estimation** helps us **visualize and analyze data** without making strict assumptions.

- **Parametric methods:** Good for simple, well-known distributions.
- **Non-parametric methods:** More flexible and work for complex or unknown distributions.

