Lecture 6

Random Variables and Distribution Functions

2. Random Variables (R.V)

A **Random Variable** is a function that assigns a real number to each outcome of a random experiment.

Types:

1. **Discrete Random Variable**
   - Takes values from a finite or countable set:
     $$X : x_1, x_2, \ldots, x_n$$
     with probabilities:
     $$p_i = P(X = x_i), \quad \sum_{i=1}^{n} p_i = 1$$
   - Probability distribution is defined by the **Probability Mass Function (PMF)**.
2. **Continuous Random Variable**
   - Takes values from a continuum (measured, not counted).
   - No jumps; continuous throughout.
   - Probability at a point is zero:
     $$P(X = x) = 0$$
   - The CDF is continuous everywhere.

3. Cumulative Distribution Function (CDF)

For any random variable X:
$$F_X(x) = P(X \leq x)$$
Relation: Distribution → Distribution Function → CDF → Probability Distribution.

Example (Discrete Case):

If:
$$X \in \{0, 1, 2\}$$
with:
$$P(X = 0) = \frac{1}{3}, \quad P(X = 1) = \frac{1}{3}, \quad P(X = 2) = \frac{1}{3}$$
The CDF will be a step function with jumps at $x = 0, 1, 2$.

4. Properties of CDF $F(x)$:

1. $\lim_{x \to -\infty} F(x) = 0, \quad \lim_{x \to \infty} F(x) = 1$
2. $F(x)$ is **non-decreasing**.
3. $F(x)$ is **right-continuous**.
4. For a discrete R.V.:
$$P(X = x) = F(x) - F(x^-)$$
(Point probability or PMF)

Summary Table:

Feature	**Discrete R.V.**	**Continuous R.V.**
Values	Finite / Countable	Infinite (measurable)
PMF	Yes	No
PDF	No	Yes
$P(X = x)$	> 0 possible	Always 0
CDF Type	Step function	Continuous function

Continuous Random Variables and Probability Density Function (PDF)

- For a continuous random variable, **every single point has probability zero**:
$$P(X = a) = 0$$
- Example: Probability of hitting a specific point on a board by a blindfolded person. The board contains *infinitely many points*, hence probability at any exact point is zero.
- **Continuous** $\Rightarrow$ Non-existence of probability for a single point.

Example (Real Line): On the number line, points are densely packed — probability at one point is zero.

Continuity as Approximation of Discrete Case

- Tossing 60 coins in 1 minute is discrete.
- Tossing 60 coins in 1 second $\Rightarrow$ can be approximated as continuous.

Probability Over Intervals

For a continuous random variable, probability is assigned to intervals, not points:
$$P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b)$$
The difference between open and closed intervals is irrelevant since $P(X = a) = 0$.

Probability Density Function (PDF)

We replace the **Probability Mass Function** (PMF) for discrete variables with the **Probability Density Function** for continuous variables.

- **Definition:** A function $f(x)$ is said to be a probability density function (PDF) of a continuous random variable X if:

$$f(x) \geq 0 \quad \text{for all } x \text{ (non-negative)}$$

  and

$$\int_{-\infty}^{\infty} f(x)\, dx = 1$$

- The probability distribution of X is given by $f(x)$ over $a \leq x \leq b$, provided:

$$\int_{a}^{b} f(x)\, dx = 1 \quad \text{(total probability)}$$

Example

$$f(x) = \begin{cases} 2x, & 0 < x < 1, \\ 0, & \text{otherwise.} \end{cases}$$

Here $f(x)$ is a valid PDF.

Probability between k_1 and k_2:

$$P(k_1 \leq X \leq k_2) = \int_{k_1}^{k_2} 2x\, dx$$

Example:

$$P\left(\frac{1}{4} < X < \frac{3}{4}\right) = \int_{1/4}^{3/4} 2x\, dx$$

Relationship Between PDF and CDF

- The probability for a continuous random variable X in an interval $[a, b]$ is given by the **area under the PDF curve** between a and b:

$$P(a \leq X \leq b) = \int_{a}^{b} f(x)\, dx$$

- The Cumulative Distribution Function (CDF) of X is defined as:

$$F_X(k) = P(X \leq k) = \int_{-\infty}^{k} f(x)\, dx$$

  where $f(x)$ is the **Probability Density Function (PDF)**.

From PDF to CDF

$$F_X(x) = \int_{-\infty}^{x} f(t)\, dt$$

The CDF is obtained by integrating the PDF from $-\infty$ to x.

From CDF to PDF

If $F(x)$ is differentiable, then:

$$f(x) = \frac{d}{dx} F_X(x)$$

This follows from the **Leibniz rule** for differentiation under the integral sign.