

Probability And Random Processes

→ Modules :

M1) Basics Of Probability

M2) Discrete Random Variables

M3) Continuous Random Variables

M4) Tail Bounds & Limit Theorems

M5) Random Processes

→ Module 1:

1) Approach to define probability

6) Counting

2) Probability Space

3) Continuity of Probability

4) Conditional Probability, Independence

5) Bayes Theorem & Total Probability Theorem

→ Classical Approach of Probability :-

- For an event E,

$$P(E) = \frac{\text{no. of outcomes favourable to } E}{\text{total no. of outcomes}}$$

- Issues :

- Fails when the outcomes are not equally likely.
- Fails when the no. of possible outcome is infinite.

→ Frequency Approach of Probability :-

- For an event E, where the experiment has been performed n times,

$$P(E) = \frac{n_E}{n}$$

n_E - no. of times E has occurred

$$\Rightarrow P(E) = \lim_{n \rightarrow \infty} \frac{n_E}{n}$$

- Issues :

- We cannot perform an experiment infinite times.
- The ratio may not converge.

→ Axiomatic Approach of Probability :-

- The approach depends on a set of axioms.
- Probability space is a set represented by 3 entities.

$$P.S = (\Omega, \mathcal{F}, P)$$



Sample Space Event Space Probability law

- Sample Space : Set of all outcomes of the random experiment.
- Event Space : Set of all favorable outcomes
- Probability Law : Function that defines probability.

• Set Theory :-

$$A \setminus B = \{x \in A \text{ st } x \notin B\}$$

$$A \cup B = \{x \in A \text{ or } x \in B\}$$

$$\text{Power} : \left(\bigcup_{i=1}^{\infty} A_i \right)^c = \bigcap_{i=1}^{\infty} A_i^c$$

$$\bigcup_{i=1}^{\infty} A_i = \{x \in A_i \text{ for some } i \in \mathbb{N}\} \rightarrow \text{By defn } \not\subseteq \cup$$

$$\Rightarrow \left(\bigcup_{i=1}^{\infty} A_i \right)^c = \{x \notin A_i \text{ } \forall i \in \mathbb{N}\} \quad \begin{matrix} \text{inverse of the} \\ \text{prev. statement} \end{matrix}$$

$$A_i^c = \{x \notin A_i\}$$

$$\bigcap_{i=1}^{\infty} A_i^c = \{x \in A_i \text{ } \forall i \in \mathbb{N}\} \rightarrow \text{By defn. of } n$$

$$\Rightarrow \bigcap_{i=1}^{\infty} A_i^c = \underline{\underline{\left(\bigcup_{i=1}^{\infty} A_i \right)^c}}$$

• Sample Space (Ω) :

- The elements of Ω , are all the possible outcomes of the random experiment.
- The elements of Ω must be mutually exclusive, (disjoint from each other) and collectively exhaustive (cover all possibility)
- Countably infinite sample space : Tossing a coin until we see tail.

Uncountably infinite sample space : Throwing a dart on a square, $\Omega = \{(x,y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$

• Event Space (\mathcal{F}) :

- An event is a subset of a sample space
- An event space is a set of subsets of Ω that form a

σ -field.

- Axioms of σ -field:

i) $\Omega \in \mathcal{F}$

ii) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (Closure under complements)

iii) $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$ (Closure under countable unions)

↳ if we take $A_i = \emptyset$ & $i \geq k$, then we can use

(this statement for finite union as well)

↳ Same can be said for \cap , using (ii) and De-Morgan's

Note: $A \Delta B = A \setminus B \cup B \setminus A$

iv) $A, B \in \mathcal{F} \Rightarrow A \Delta B \in \mathcal{F}$ → Not an axiom. Implied by (ii) - (iii)

$$A \Delta B = (A \cap B^c) \cup (B \cap A^c)$$

$$A \in \mathcal{F}, B \in \mathcal{F} \Rightarrow A \cap B^c \in \mathcal{F}, \text{ (by } B \cap A^c \in \mathcal{F}) \\ \therefore A \Delta B \in \mathcal{F}$$

• Smallest σ -field with E : $\{\Omega, \emptyset, E, E^c\}$

" " " " $A, B : \{\Omega, \emptyset, A, B, A^c, B^c, A \cup B, A \cup B^c, A \cap B, A \cap B^c, A^c \cup B, A^c \cap B, A^c \cap B^c, A^c \cup B^c, A \Delta B, (A \Delta B)^c\}$

Example: Complete $\{\Omega, \emptyset, \{1\}, \{2, 3\}\}$ for $\Omega = \{1, 2, 3, 4\}$

$$\mathcal{F} = \{\Omega, \emptyset, \{1\}, \{2, 3\}, \{2, 3, 4\}, \{1, 4\}, \{1, 2, 3\}, \{4\}\}$$

- Define the smallest σ -set of X, Y, Z as $\sigma(X, Y, Z)$, then,

$$\sigma(\Omega, \emptyset, A, B) = \sigma(A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c)$$

Disjoint sets that collectively exhaust Ω .

Any union of the 4 sets, have their complements as a union of the other sets.

- Probability Law (P) :- $P : \mathcal{F} \rightarrow \mathbb{R}$ is a probability law if it follows the below axioms.

i) $P(\Omega) = 1$ (Normalization)

ii) $P(E) \geq 0 \quad \forall E \in \mathcal{F}$ (Non-negativity)

iii) If A_1, A_2, \dots are mutually exclusive (disjoint), then,

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i)$$

Properties:

i) $P(\emptyset) = 0$.

Proof: Let $A_i = \emptyset \quad \forall i$

$$\begin{aligned} P\left(\bigcup_i A_i\right) &= \sum_i P(A_i) \\ &= P(\emptyset) = \sum P(\emptyset) \Rightarrow P(\emptyset) = 0 \end{aligned}$$

Example: Construct a probability law for rolling a die.

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

Defn

$$\left\{ \begin{array}{l} P(\{\text{i}\}) = p_i \quad \forall i \in \{1, 2, 3, 4, 5, 6\} \quad \text{st. } \sum_{i=1}^6 p_i = 1 \quad * p_i \geq 0 \\ P(A) = \sum_{i \in A} P(\{\text{i}\}), \quad A \subseteq \Omega \end{array} \right.$$

$$P(\Omega) = \sum_{i \in \Omega} P(\{\text{i}\}) = \sum_{i \in \Omega} p_i = 1 \quad (\text{Normalization})$$

$$p_i \geq 0 \Rightarrow \sum_{i \in A} p_i \geq 0 \quad \forall A \subseteq \Omega \Rightarrow P(A) \geq 0 \quad \forall A \subseteq \Omega$$

(Non-negativity)

By defn. additivity is satisfied.

ii) $P(A) \leq 1 \quad \forall A \subseteq \Omega$

iii) If $A \subseteq B$, then $P(A) \leq P(B)$

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

v) $P(A) + P(A^c) = 1$

Proof:

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

$P(A \cup B) = P(A \cup (B|A)) = P(A) + P(B|A)$

$$P(B) = P((A \cap B) \cup (B|A))$$

$$= P(A \cap B) + P(B|A)$$

$$\Rightarrow P(B|A) = P(B) - P(A \cap B)$$

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

iii) If $A \subseteq B$, $P(A \cup B) = P(B)$

$$\begin{aligned} \Rightarrow P(B) &= P(A) + P(B) - P(A \cap B) \\ &= P(A) + P(B/A) + P(\cancel{A \cap B}) - \cancel{P(A \cap B)} \end{aligned}$$

$$\Rightarrow P(B) = P(A) + P(B/A)$$

$$\Rightarrow \underline{\underline{P(B) \geq P(A)}}$$

• Continuity Of Probability :-

- A continuous function has a definite limit for all points in its domain.

- A function $f: \mathbb{R} \rightarrow \mathbb{R}$ if a sequence $x_n \rightarrow x^* \Rightarrow f(x_n) \rightarrow f(x^*)$ as $n \rightarrow \infty$ (Formal defn), ie,

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right)$$

- But, probability law is a function on sets. For such a function, continuity is defined as,

Let $A_1, A_2, A_3, \dots, n \in \mathbb{N}$ be a sequence of events, then

$$\lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right) = P\left(\lim_{n \rightarrow \infty} \bigcup_{i=1}^n A_i\right)$$

Proof:

Claim 1: Consider 3 sets A_1, A_2, A_3 . Then

$$A_1 \cup A_2 \cup A_3 = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2))$$

As a union of disjoint sets B_1, B_2, B_3 .

$$\Rightarrow B_i = A_i \setminus \bigcup_{k=1}^{i-1} A_k \longrightarrow \bigcap_{i=1}^{\infty} B_i = \emptyset$$

Proof of Claim 1:

To prove, $B_i \cap B_j = \emptyset \quad \forall i \neq j$

WLOG let $i < j$. Let $x \in B_i$

$$\Rightarrow x \in A_i \setminus \bigcup_{k=1}^{i-1} A_k \Rightarrow x \notin \bigcup_{k=1}^{i-1} A_k \quad \& \quad x \in A_i$$

$$\begin{aligned} \text{Assume } x \in B_j &\Rightarrow x \in A_j \setminus \bigcup_{k=1}^{j-1} A_k \\ &\Rightarrow x \notin \bigcup_{k=1}^{j-1} A_k \end{aligned}$$

But we proved $x \in A_i \longrightarrow \text{Contradiction}$

$\therefore x \in B_i \Rightarrow x \notin B_j, \text{ if } i < j$

Similarly can be stated for $i > j$.

$\therefore x \in B_i \Rightarrow x \notin B_j \text{ if } i > j$

$$\Rightarrow \underline{B_i \cap B_j = \emptyset \text{ } \forall i \neq j}$$

$$\underline{\text{Claim 2: } \bigcup_{i=1}^n A_i = \bigcup_{i=1}^n B_i \text{ } \forall n \in \mathbb{N}} \quad \textcircled{1}$$

$$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i \quad \longrightarrow \textcircled{2}$$

Proof of Claim 2:

For $\textcircled{1}$, if $n = 1$, $B_1 = A_1 \rightarrow$ trivially proved

Assume true for $n = k$, ie, $\bigcup_{i=1}^k A_i = \bigcup_{i=1}^k B_k := C_k$

To prove for $n = k+1$,

$$\bigcup_{i=1}^{k+1} A_i = A_{k+1} \cup C_k$$

$$\begin{aligned} \bigcup_{i=1}^{k+1} B_i &= B_{k+1} \cup C_k = (A_{k+1} \bigcup_{i=1}^k A_i) \cup C_k \\ &= (A_{k+1} \setminus C_k) \cup C_k \end{aligned}$$

$$\Rightarrow \boxed{\bigcup_{i=1}^{k+1} B_i = \bigcup_{i=1}^{k+1} A_i} \Rightarrow \textcircled{1} \text{ is valid.}$$

To extend the equality till ∞ , we can prove that any element in LHS will belong to RHS. (Induction is not valid for $n \rightarrow \infty$)

Final Proof: To prove:

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right)$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) \quad \text{As defined and proved above}$$

$$= \sum_{i=1}^{\infty} P(B_i) \quad [\text{Additivity Axiom}]$$

Applies only because summation is defined
of \mathbb{R} and limit is well
- defined for \mathbb{R} .

$$\begin{aligned} &= \lim_{n \rightarrow \infty} \sum_{i=1}^n P(B_i) \\ &= \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n B_i\right) \quad [\text{Additivity Axiom}] \\ &= \lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right) \quad [\text{Claim 2}] \end{aligned}$$

$$\Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \underbrace{\lim_{n \rightarrow \infty} P\left(\bigcup_{i=1}^n A_i\right)}$$

Note: A sequence $\{x_n\}$ is said to converge to x^* iff
 $\forall \varepsilon > 0$, $\exists n_0 \in \mathbb{N}$ et $|x_n - x^*| < \varepsilon \quad \forall n > n_0$.

Corollary 1:

i) $A_i \subseteq A_{i+1} \quad \forall i \Rightarrow P\left(\bigcup_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n)$

ii) $A_i \supseteq A_{i+1} \quad \forall i \Rightarrow P\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{n \rightarrow \infty} P(A_n)$

Corollary 2 : Union Bound for infinite events,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$$

→ For finite union, induction can be used for proof
for infinite, continuity.

→ Conditional Probability :-

• $P(A|B)$ is the probability of A given that B has occurred.

• $P(A|B) \propto P(A \cap B)$ If $P(A \cap B) = 0$, $P(A|B) = 0$

$$= P(A|B) = k P(A \cap B), \text{ if } A=B, P(B|B) = k P(B \cap B)$$
$$\Rightarrow l = k P(B)$$
$$\Rightarrow k = l/P(B)$$

$$\Rightarrow P(A|B) = P(A \cap B) / P(B)$$

→ Independence :

• 2 events A and B are independent if,

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

$$\Rightarrow P(A|B) = P(A)$$

- 3 events A, B, C are independent if,

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

$$P(B \cap C) = P(B)P(C)$$

$$P(A \cap C) = P(A)P(C)$$

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

- For n events,

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \quad \# I \subseteq \{1, 2, 3, \dots, n\}$$

- A collection of sets is a partition of Ω if they are mutually exclusive and exhaustive.

$$(A_i \cap A_j = \emptyset \quad \forall i \neq j), \quad (\bigcup_{i=1}^n A_i = \Omega)$$

→ Total Probability Theorem :-

Let $\{A_1, A_2, A_3, \dots, A_n\}$ be a partition of Ω , and B be any event. Then,

$$P(B) = \sum_i P(B \cap A_i)$$

$$\Rightarrow P(B) = \sum_i P(B|A_i)P(A_i), \quad \text{if } P(A_i) \neq 0 \quad \forall i$$

Proof:

$$B = \bigcup_{i=1}^n (B \cap A_i)$$

$$\Rightarrow P(B) = P\left(\bigcup_{i=1}^n (B \cap A_i)\right)$$

$$\Rightarrow P(B) = \sum_{i=1}^n P(B \cap A_i)$$

\rightarrow Bayes' Theorem :-

$\{A_1, A_2, A_3, \dots, A_n\}$ be a partition of Ω . B be any event.

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

$$\Rightarrow P(A_i | B) = \frac{P(B|A_i) P(A_i)}{\sum_i P(B|A_i) P(A_i)}$$

\rightarrow Multiplication Rule :

$$P(A_1 \cap A_2 \cap A_3 \dots \cap A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_2 \cap A_1) \dots P(A_n | A_1 \cap A_2 \dots \cap A_{n-1})$$

\rightarrow Conditional Independence :

2 events A and B are said to be independent given C if,

$$P(A \cap B | C) = P(A | C) \cdot P(B | C)$$

- If $P(C) = 1$, then A and B are truly independent, ie,

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

→ Counting Techniques :-

Suppose there are n objects

i) No. of k -length sequences : ${}^n P_k = \frac{n!}{(n-k)!}$

ii) No. of k -length sets : ${}^n C_k = \frac{n!}{k!(n-k)!}$

iii) No. of ways of choosing r sets of sizes $n_1, n_2, n_3, \dots, n_r$,

$$= {}^n C_{n_1} \cdot {}^{n-n_1} C_{n_2} \cdot {}^{n-n_1-n_2} C_{n_3} \cdots {}^{n-n_1-n_2-\cdots-n_{r-1}} C_{n_r}$$

$$= \frac{n!}{(n-n_1)! n_1!} \frac{(n-n_1)!}{(n-n_1-n_2)! n_2!} \cdots \cdots$$

$$= \frac{n!}{n_1! n_2! n_3! \cdots n_r!}$$

→ Random Variables :-

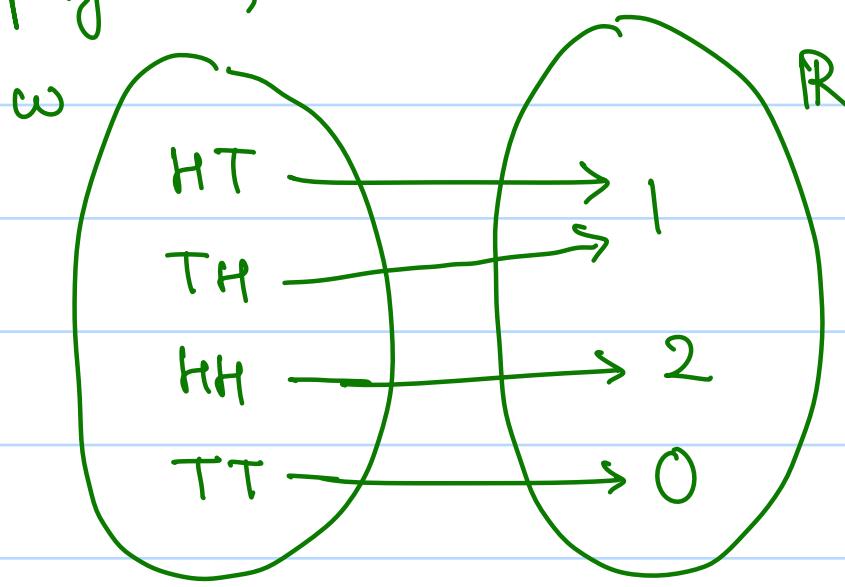
° A random variable is a function $X: \Omega \mapsto \mathbb{R}$ such that,

$$\{X \leq x\} \triangleq \{\omega \mid X(\omega) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R}$$

Example: $\Omega = \{\text{HT, TH, HH, TT}\}$

$X: \Omega \mapsto \mathbb{R}$ st $X(\omega) = \text{No. of heads}$

The mapping is,



$$\{x \leq x_0\} = \begin{cases} \emptyset, & x_0 < 0 \\ \{\text{TT}\}, & x_0 \in [0, 1) \\ \{\text{HT, TH, TT}\}, & x_0 \in [1, 2) \\ \Omega, & x_0 \geq 2 \end{cases}$$

Theorem: $X: \Omega \rightarrow \mathbb{R}$ on \mathcal{F} , then

i) $X^{-1}((-\infty, x]) \in \mathcal{F}$

ii) $X^{-1}([x_1, x_2]) \in \mathcal{F} \quad \forall x_1, x_2 \in \mathbb{R}$

iii) $X^{-1}(\{x\}) \in \mathcal{F} \quad \forall x \in \mathbb{R}$

Proof of (i):

wkt $X^{-1}(-\infty, x] \in \mathcal{F} \quad \forall x \in \mathbb{R}$

$$(-\infty, x) = \bigcup_{n=1}^{\infty} \left(-\infty, x - \frac{1}{n} \right]$$

$$X^{-1}\left(\left(-\infty, x - \frac{1}{n}\right]\right) \in \mathcal{F} \quad \forall n \Rightarrow X^{-1}\left(\bigcup_{n=1}^{\infty} \left(-\infty, x - \frac{1}{n} \right]\right) \in \mathcal{F}$$

Proof of (ii) :

$$\text{wkt } x^{-1}(-\infty, x_1] \in \mathcal{F}$$

$$x^{-1}(-\infty, x_2) \in \mathcal{F}$$

Since subtraction is closed in \mathcal{F} , $x^{-1}(-\infty, x_2) \in \mathcal{F}$

$$\Rightarrow x^{-1}([x_2, \infty)) \in \mathcal{F}$$

$$\Rightarrow x^{-1}(-\infty, x_1] \cap x^{-1}([x_2, \infty)) \in \mathcal{F}$$

$$\Rightarrow x^{-1}([x_2, x_1]) \in \mathcal{F}$$

iv) $x^{-1}((x_1, x_2)) \in \mathcal{F}$

$$(x_1, x_2) = (x_1, \infty) \cap (-\infty, x_2)$$

$$x^{-1}(-\infty, x_1] \in \mathcal{F} \Rightarrow x^{-1}(x_1, \infty) \in \mathcal{F}$$

$$x^{-1}([x_2, \infty)) \in \mathcal{F} \Rightarrow x^{-1}(-\infty, x_2) \in \mathcal{F}$$

$$\Rightarrow x^{-1}((x_1, \infty) \cap (-\infty, x_2)) \in \mathcal{F}$$

$$\Rightarrow x^{-1}((x_1, x_2)) \in \mathcal{F}$$

• Borel σ -Algebra :

Smallest σ -field on \mathbb{R} containing sets of the form $(-\infty, x]$ $\forall x \in \mathbb{R}$, ie,

$$B(\mathbb{R}) = \sigma((-\infty, x] \forall x \in \mathbb{R})$$

- Contains all possible subsets of \mathbb{R} .

• Cumulative Distribution Function (CDF) :-

- A CDF is a function $F_X : \mathbb{R} \mapsto [0,1]$ such that,

$$F_X(x) = P_X(\{X < x\})$$

Example: For,

$$\{X \leq x\} = \begin{cases} \emptyset, & x < 0 \\ \{\text{TT}\}, & x \in [0,1) \\ \{\text{TT, HT, TH}\}, & x \in [1,2) \\ \Omega, & x \geq 2 \end{cases}$$

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{4}, & x \in [0,1) \\ \frac{3}{4}, & x \in [1,2) \\ 1, & x \geq 2 \end{cases}$$

Example: Let $X(\omega) = c$, $c \in \mathbb{R}$. Then,

$$\{X \leq x\} = \begin{cases} \emptyset, & x < c \\ \Omega, & x \geq c \end{cases}$$

$$\Rightarrow F_X(x) = \begin{cases} 0, & x < c \\ 1, & x \geq c \end{cases}$$

• Indicator R.V :-

Consider an $A \in \mathcal{F}$. Then $I_A: \Omega \rightarrow \mathbb{R}$ such that,

$$I_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

$$\Rightarrow \{I_A < x\} = \begin{cases} \emptyset, & x < 0 \\ A^c, & x \in [0,1) \\ \Omega, & x \geq 1 \end{cases}$$

$\{I_A < x\} \in \mathcal{F} \quad \forall x \in \mathbb{R}$. So I_A is a valid r.v.

I_A is an indicator r.v for the event A .

• $B_1 \cap B_2 = \emptyset$, then,

$$I_{B_1 \cup B_2}(\omega) = I_{B_1}(\omega) + I_{B_2}(\omega) \quad \forall \omega \in \Omega$$

B_1, B_2, B_3, \dots form a partition of Ω , then,

$$I_{\bigcup_{i=1}^n B_i} = \sum_{i=1}^n I_{B_i} = 1$$

• Theorems of CDF :-

i) $\lim_{x \rightarrow \infty} F_X(x) = 1, \quad \lim_{x \rightarrow -\infty} F_X(x) = 0$

$$\lim_{x \rightarrow \infty} F_X(x) = \lim_{x \rightarrow \infty} P_X(X \leq x) = P_X(\mathbb{R}) = 1$$

$$\lim_{x \rightarrow -\infty} F_X(x) = \lim_{x \rightarrow -\infty} P_X(X \leq x) = P_X(\emptyset) = 0$$

2) $x < y \Rightarrow F_X(x) \leq F_X(y)$

3) $F_X(x)$ is always right continuous, ie,

$$\lim_{\epsilon \rightarrow 0^+} F_X(x + \epsilon) = F_X(x) \quad \forall x \in \mathbb{R}$$

