Textbooks:

1) Papoulis
2) Bertsekas
3) Grimmett

Prof. Gowtham Kurri

Probability And Random Processes

Grading:

- Assignments 15%
- Quiz 1 & 2 15% + 15%
- In-Class Quiz 5%
- Midsem 20%
- Endsem 30%

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
2) Probability Space
3) Continuity of Probability
4) Conditional Probability, Independence
5) Bayes Theorem & Total Probability Theorem
6) Counting

→ 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 outcomes 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

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

- **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 favourable outcomes

- Probability Law : Function that defines probability.

- **Set Theory :-**

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

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

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

$\bigcup_{i=1}^{\infty} A_i = \{ x \in A_i \text{ for some } i \in \mathbb{N} \}$ → By defn of $\cup$

$\Rightarrow \left(\bigcup_{i=1}^{\infty} A_i\right)^c = \{x \notin A_i \ \forall\, i \in \mathbb{N}\}$ reverse of the prev. statement

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

$\bigcap_{i=1}^{\infty} A_i^c = \{x \in A_i \ \forall\, i \in \mathbb{N}\} \longrightarrow$ By defn. of $\cap$

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

- **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 possibilities)

  - Countably infinite sample space : Tossing a coin until we see tails.
    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.

- <u>Axioms of σ - field</u>:

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

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

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

↳ if we take $A_i = \phi \ \forall \ i \geq$ some k, then we can use this statement for finite unions as well

↳ Same can be said for $\cap$, using (ii) and De-Morgans)

<u>Note</u>: $A \Delta B = A \setminus B \cup B \setminus A$

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

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

$A \in \mathcal{F}, B \in \mathcal{F} \Rightarrow A \cap B^c \in \mathcal{F}$, |||^{ly} $B \cap A^c \in \mathcal{F}$

$\therefore A \Delta B \in \mathcal{F}$

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

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

<u>Example</u>: Complete $\{\Omega, \phi, \{1\}, \{2,3\})$ for $\Omega = \{1, 2, 3, 4\}$

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

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

$$\sigma(\Omega, \phi, A, B) = \sigma(\underbrace{A \cap B, A \cap B^c, A^c \cap B, A^c \cap B^c}_{\text{Disjoint sets that collectively exhaust } \Omega.})$$

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

- Probability Law (P) :- $P : \mathcal{F} \longmapsto \mathbb{R}$ 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) $P(\phi) = 0$.

Proof: Let $A_i = \phi \ \forall i$

$$P\left(\bigcup_i A_i\right) = \sum_i P(A_i)$$

$$= P(\phi) = \sum P(\phi) \Rightarrow P(\phi) = 0$$

Example: Construct a probability law for rolling a die.

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

Defn $\begin{cases} P(\{i\}) = p_i \quad \forall \; i \in \{1,2,3,4,5,6\} \text{ st } \sum_{i=1}^{6} p_i = 1 \text{ \& } p_i \geq 0 \\ P(A) = \sum_{i \in A} P(\{i\}), \quad A \subseteq \Omega \end{cases}$

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

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

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

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

$$\begin{aligned} P(B) &= P((A \cap B) \cup (B \setminus A)) \\ &= P(A \cap B) + P(B \setminus A) \end{aligned}$$

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

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

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

$$\Rightarrow P(B) = P(A) + P(B) - P(A \cap B)$$
$$= P(A) + P(B/A) + \cancel{P(A \cap B)} - \cancel{P(A \cap B)}$$
$$\Rightarrow P(B) = P(A) + P(B/A)$$
$$\Rightarrow \underline{P(B) \geq P(A)}$$

- Continuity Of Probability :-

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

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

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

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

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

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

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.

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

Proof of Claim 1:

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

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

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

Assume $x \in B_j \Rightarrow x \in A_j \setminus \bigcup_{k=1}^{j-1} A_k$

$$\Rightarrow x \notin \bigcup_{k=1}^{j-1} A_k$$

But we proved $x \in A_i \longrightarrow$ Contradiction

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

Illy can be stated for $i>j$.

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

$\Rightarrow \quad B_i \cap B_j = \phi \quad \forall \; i \neq j$

<u>Claim 2:</u> $\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} B_i \quad \forall \; n \in \mathbb{N}$ —— ①

$\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$ —— ②

<u>Proof of Claim 2:</u>

For ①, if $n = 1$, $B_1 = A_1 \longrightarrow$ 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 $n = k+1$,

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

$$\bigcup_{i=1}^{k+1} B_i = B_{k+1} \cup C_k = \left(A_{k+1} \setminus \bigcup_{i=1}^{k} A_i\right) \cup C_k$$

$$= (A_{k+1} \setminus C_k) \cup C_k$$

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

$\Rightarrow \boxed{\bigcup_{i=1}^{k+1} B_i = \bigcup_{i=1}^{k+1} A_i} \Rightarrow$ ① 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:

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

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

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

$$= \lim_{n\to\infty} \sum_{i=1}^{n} P(B_i)$$

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

$$= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} B_i\right) \quad \text{[Additivity Axiom]}$$

$$= \lim_{n\to\infty} P\left(\bigcup_{i=1}^{n} A_i\right) \quad \text{[Claim 2]}$$

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

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

Corollary 1:

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

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

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

$= P(A|B) = kP(A \cap B)$, if $A = B$, $P(B|B) = kP(B \cap B)$

$\Rightarrow 1 = kP(B)$

$\Rightarrow k = 1/P(B)$

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

→ 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,

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

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

$$\left(A_i \cap A_j = \phi \;\forall\, i \neq j\right), \left(\bigcup_{i=1}^{n} A_i = \Omega\right)$$

→ <u>Total Probability Theorem</u> :-

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

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

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

<u>Proof</u>:

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

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

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

→ **Bayes' Theorem :-**

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

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

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

→ **Multiplication Rule :**

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

→ **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 : ${}^nP_k = \frac{n!}{(n-k)!}$

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

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

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

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

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

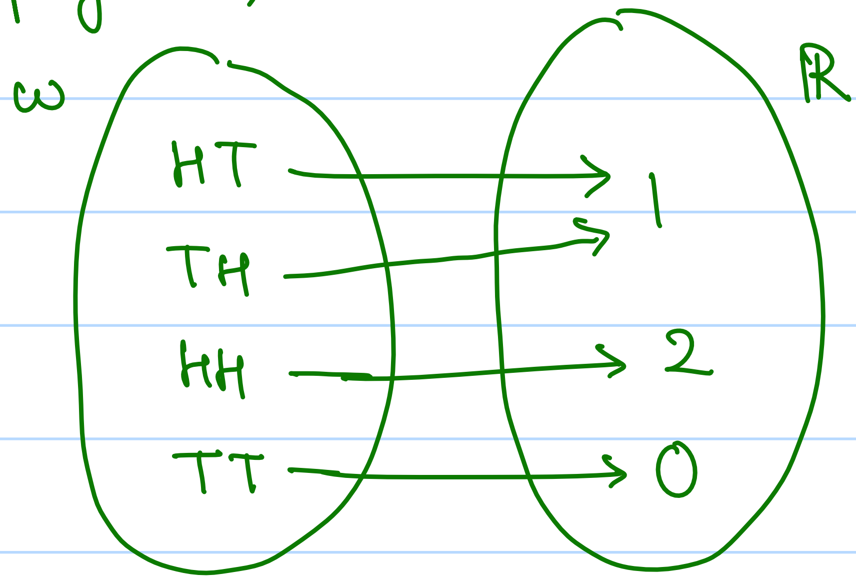
→ Random Variables :-

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

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

The mapping is,



$$\{x \le x_0\} = \begin{cases} \phi, & x_0 < 0 \\ \{TT\}, & x_0 \in [0,1) \\ \{HT, TH, TT\}, & x_0 \in [1,2) \\ \Omega, & x_0 \ge 2 \end{cases}$$

Theorem: $X : \Omega \mapsto \mathbb{R}$ 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) $X^{-1}(\{x_1\}) \in \mathcal{F} \quad \forall\, x_1 \in \mathbb{R}$

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) :

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

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

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

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

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

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

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,

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

Example: For,

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

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

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

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

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

- **Indicator R.V :-**

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

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

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

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

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

- $B_1 \cap B_2 = \phi$, then,

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

$B_1, B_2, B_3 \ldots$ 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) $\lim_{x \to \infty} F_X(x) = 1$, $\lim_{x \to -\infty} F_X(x) = 0$

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

$$\lim_{x \to -\infty} F_X(x) = \lim_{x \to -\infty} P_X(X \leq x) = P(\phi) = \underline{0}$$

P_X - Probability law of X

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

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

$$\lim_{\varepsilon \to 0^+} F_X(x+\varepsilon) = F_X(x) \quad \forall\, x \in \mathbb{R}$$

<u>Proof</u>:

$$\lim_{n\to\infty} F_X\left(x+\frac{1}{n}\right) = \lim_{n\to\infty} P\left(X \leq x+\frac{1}{n}\right)$$

Let $\left\{X \leq x+\frac{1}{n}\right\} = B_n \Rightarrow B_n$ is a decreasing sequence

WK by continuity of probability $P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n)$

$$\Rightarrow \lim_{n\to\infty} F_X\left(x+\frac{1}{n}\right) = P\left(\bigcap_{n=1}^{\infty}\left\{X \leq x+\frac{1}{n}\right\}\right)$$

$$= P(\{X \leq x\}) = F_X(x)$$

$$\therefore \lim_{n\to\infty} F_X\left(x+\frac{1}{n}\right) = F_X(x)$$

$$\Rightarrow \underline{RHL = F_X(x)}$$

4) $F_X(x) - \lim_{\varepsilon\to 0^+} F_X(x-\varepsilon) = P(\{X = x\})$

5) $P_X(x_1 \leq X \leq x_2) = P(\omega : x_1 \leq X(\omega) \leq x_2)$

$= P(X$

→ **Discrete Random Variable** :-

A random variable is discrete if $Range(X) \subseteq \mathbb{R}$ is finite / countably infinite.

∘ **Probability Mass Function** :

$P_X : X \longmapsto [0,1]$, given by

$$P_X(x) = P(X = x) = P(\{\omega : X(\omega) = x\})$$

• **Lemma:** For an rv st $x \in \{x_1, x_2, x_3 \dots\}$

i). $\sum_{i=1}^{\infty} P_X(x_i) = \sum_{i=1}^{\infty} P(X = x_i)$

$= P\left(\bigcup_{i=1}^{\infty} \{X = x_i\}\right)$ [Additivity]

$= P(\{X \in \{x_1, x_2, x_3 \dots\}) = P(\Omega)$

$\therefore \sum_{i=1}^{\infty} P_X(x_i) = \underline{\underline{1}}$

∘ **CDF of a Discrete R.V** :-

$$F_X(x) = P(\{X \leq x\}) = P\left(\bigcup_{i \in \mathbb{N} : x_i \leq x} \{X = x_i\}\right)$$

$$\Rightarrow \boxed{F_X(x) = \sum_{i \in \mathbb{N} : x_i \leq x} P_X(x_i)}$$

- **Functions on R.V :-**

Let $X : \omega \longmapsto \mathbb{R}$ and $Y : \mathbb{R} \longmapsto \mathbb{R}$, ie a function on X. Y is also a r.v.

Proof: $Y^{-1}(B) = \{ y : Y(y) \in B \} \quad \forall \; B \in \mathcal{B}(\mathbb{R})$

$= \{ X(x) : Y(X(x)) \in B \}$

$X^{-1}(Y^{-1}(B)) = \{ x : Y(X(x)) \in B \}$

$= \{ x : X(x) \in Y^{-1}(B) \}$

Since $X^{-1}(S) \in \mathcal{F} \quad \forall \; S \in \mathcal{B}(\mathbb{R})$

Better proof needed.

$Y^{-1}(B) \in X(\mathcal{F}) = \mathcal{F}' \quad \forall \; B \in \mathcal{B}(\mathbb{R})$

$\Rightarrow$ Y is a valid r.v

- PMF of y : $\boxed{P_Y(Y = y) = \sum_{Y(x) = y} P_X(X = x)}$

- **Expectation:** $\boxed{\mathbb{E}[X] = \sum_{x \in \mathbb{R}} x P_X(x)}$

- If $y = Y(x)$,

Law of the Unconscious Statistician : $\boxed{\mathbb{E}[Y] = \sum_{x \in \mathcal{X}} Y(x) P_X(x)}$

Proof: $\mathbb{E}[Y] = \sum_{y \in \mathcal{Y}} y P_Y(y) = \sum_{y \in \mathcal{Y}} y \sum_{Y(x) = y} P_X(x)$

$= \underbrace{\sum_{y \in \mathcal{Y}} \sum_{Y(x) = y}}_{\text{All possible } x} Y(x) \, P_X(x) = \sum_{x \in \mathcal{X}} Y(x) P_X(x)$

- **Variance:**

$$Var[X] = E\left[(X - E[X])^2\right]$$

$$\Rightarrow \quad Var[X] = E[X^2] - E[X]^2$$

- n^{th} Moment of an RV $= E[X^n]$. (defn)

- **Examples of RV:-**

1) **Bernoulli RV:** Binary o/p Rv. (ex: Coin Toss)

$$P(\{H\}) = p, \quad P(\{T\}) = 1-p$$

$$X(H) = 1, \quad X(T) = 0 \longrightarrow X \text{ is a Bernoulli R.v}$$

$$E[X] = p, \quad Var[X] = p - p^2$$

2) **Binomial RV:** (ex: Coin Tossed n times)

Any event ω : Sequence of H's and T's of length n.

$P(\{H\}) = p$, $P(\{T\}) = 1-p$. $X(\omega)$ = No. of heads in ω

$$\Rightarrow \quad P_X(k) = {}^nC_k \, p^k (1-p)^{n-k} \quad \forall \, k \in \mathbb{N}$$

$$E[X] = np \qquad Var[X] = np(1-p)$$

3) **Geometric RV:** (ex: Toss a coin till heads)

$X(\omega)$ = No. of coin tosses in ω, to get a head :)

$P_X(k) = p(1-p)^{k-1}$ $\quad E[X] = \frac{1}{p}$ $\quad Var[X] = \frac{1-p}{p^2}$

4) **Poisson RV:** $X \in \{0,1,2,\dots\}$

$$P_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0,1,2\dots \quad \lambda \in \mathbb{R}$$

$E[X] = Var[X] = \lambda$.

- Let $Y \sim$ Binomial (n,p), As $n \to \infty$ st $np = \lambda$, a constant, we have,

$$\lim_{n\to\infty} P_Y(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0,1,2,3,\dots$$

Proof:

$$\lim_{n\to\infty} P_Y(k) = \lim_{n\to\infty} {}^nC_k\, p^k(1-p)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n!}{(n-k)!\,k!} p^k(1-p)^{n-k}$$

$$= \lim_{n\to\infty} \frac{n!}{(n-k)!\,k!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n(n-1)(n-2)\cdots(n-k+1)}{n^k} \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \overbrace{\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\left(\frac{n-2}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)}^{1} \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} e^{\lim_{n\to\infty} \left(\frac{-\lambda}{n}\right)(n-k)}$$

$$= \frac{\lambda^k}{k!} e^{\lim_{n\to\infty} -\lambda + \frac{\lambda k}{n}}$$

$$\Rightarrow \lim_{n\to\infty} P_Y(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$(1+f(x))^{g(x)}$

$f(x) \to 0$

$g(x) \to \infty$

as $x \to \infty$,

then

$e^{g(x)\ln(1+f(x))}$

$= e^{\lim_{n\to\infty} g(x) f(x)}$

→ Multiple Random Variables :-

- (X, Y) are said to be jointly discrete if (X,Y) takes values in some countable subset of $\mathbb{R}^2$

Joint PMF $P_{X,Y}(x,y) = P(X=x \cap Y=y)$

$$P_X(x) = \sum_{y\in Y} P_{X,Y}(x,y)$$

Proof: $P_X(x) = P(X=x)$

$$= P\left(X=x \cap \bigcup_{y\in Y}(Y=y)\right)$$

$$= P\left(\bigcup_{y}(X=x \cap Y=y)\right)$$

$$= \sum_{y\in Y} P(X=x \cap Y=y) = \sum_{y\in Y} P_{XY}(x,y$$

- $P((XY) \in B) = \sum_{(x,y) \in B} P_{XY}(x,y) \quad \forall\ B \subseteq \mathbb{R}^2$

- <u>Functions on 2 Rv's</u> :-

  $\forall\ X : \Omega \longmapsto \mathbb{R}$, $Y : \Omega \longmapsto \mathbb{R}$ Rv's, $Z = g(X,Y)$,
  $g : \mathbb{R}^2 \longmapsto \mathbb{R}$ is also a Rv.

  $$P_Z(z) = \sum_{x,y\,:\,g(x,y) = z} P_{XY}(x,y)$$

- <u>Independence</u> :-

  2 Rv's X and Y are said to be independent if

  $$P_{XY}(x,y) = P_X(x)\,P_Y(y) \quad \forall\ x,y \in X,Y$$

- If $X, Y \in \{0,1\}$, then X, Y are independent.

- If X, Y are independent, then $E[XY] = E[X]\,E[Y]$

- Independent $\cong$ Uncorrelated

- n Rv's $X_1, X_2, X_3 \ldots X_n$ are said to be independent if,

  $$P_{X[1:n]}(x_{[1:n]}) = \prod_{i \in [1:n]} P_{X_i}(x_i)$$

- If X and Y are independent, $h(X)$ and $g(Y)$ are independent

• Let $Z = g(X,Y)$ be an rv, then,

$$E[Z] = \sum_{x,y} g(x,y)\, P_{XY}(x,y)$$

→ Properties of $E[X]$ and $Var[X]$:-

1) $Y = aX + b \Rightarrow E[Y] = aE[X] + b$

$$Var[Y] = a^2 Var[X]$$

2) $Y = X_1 + X_2 \Rightarrow E[Y] = E[X_1] + E[X_2] \longrightarrow$ Prove by LOTUS

$\Rightarrow E[\sum X_i] = \sum E[X_i]$

$\Rightarrow Var[Y] = Var[X_1] + Var[X_2] + 2\,Cov[X_1, X_2]$

Where,

$$Cov[X_1, X_2] = E[X_1 \cdot X_2] - E[X_1]E[X_2]$$

Definition: Correlation Coefficient $= \rho(X_1, X_2)$

$$= \frac{Cov[X_1, X_2]}{\sqrt{Var[X_1]}\sqrt{Var[X_2]}}$$

Theorem: $|\rho(X_1, X_2)| \leq 1$. Equality iff $X_1 = aX_2 + b$, $a, b \in \mathbb{R}$

ie X_1 and X_2 are linearly dependent.

Proof: By Cauchy-Schwartz Inequality,

$$E[XY] \leq \sqrt{E[X^2]E[Y^2]}$$

Proof of Cauchy-Schwartz,

$$E[(X-\alpha Y)^2] \geq 0 \quad (\text{Equality if } X = \alpha Y \text{ (lin. dep)})$$

$$\Rightarrow E[X^2] - 2\alpha E[XY] + \alpha^2 E[Y^2] \geq 0 \rightarrow \text{Is a Quadratic wrt } \alpha$$

$$\Rightarrow \sqrt{(2E[XY])^2 - 4(E[Y^2])(E[X^2])} \leq 0 \rightarrow \text{Discriminant} < 0$$

$$\Rightarrow 4E[XY]^2 - 4E[Y^2]E[X^2] \leq 0$$

$$\Rightarrow E[XY] \leq E[Y^2]E[X^2]$$

To Prove the theorem,

$$X \longrightarrow (X - E[X]) \quad Y \longrightarrow (Y - E[Y])$$

$$\Rightarrow E[(X-E[X])(Y-E[Y])]^2 \leq E[(X-E[X])^2]E[(Y-E[Y])^2]$$

$$\Rightarrow |Cov[X,Y]|^2 \leq \sqrt{Var[X]Var[Y]}$$

$$\Rightarrow \rho[X,Y]^2 \leq 1 \Rightarrow |\rho[X,Y]| \leq 1$$

3) $X, Y,$ et $Z = X + Y$

$$P_Z(z) = \sum_{\substack{x,y \\ \text{st } x+y \in z}} P_{XY}(x,y) = \sum_x P_{XY}(x, z-x)$$

$$= \sum_y P_{XY}(z-y, y)$$

$$\Rightarrow \underline{\underline{P_X * P_Y}} \quad \text{If } X, Y \text{ are independent}$$

Example: X, Y are Geometric RVs. Find $P_Z(z)$ where $Z = X + Y$ if X, Y are independent.

$$P_Z(z) = P_X(x) * P_Y(y)$$
$$= \sum_x P_X(x) P_Y(z-x)$$
$$= \sum_x p(1-p)^{x-1} q(1-q)^{z-x-1}$$
$$= pq \frac{(1-q)^{z-1}}{(1-p)} \sum_x \left(\frac{1-p}{1-q}\right)^x$$

If $p = q$,
$$P_Z(z) = (z-1) p^2 (1-p)^{z-2}$$

→ **Conditioning** :-

Conditioning an RV $X: \Omega \mapsto \mathbb{R}$ on an event $A \subseteq \Omega$

$$P_{X|A}(x) \triangleq P(X = x | A)$$
$$P_{X|A}(x) = \frac{P(X = x, A)}{P(A)}$$

- **Theorem:** If $A_1, A_2 \dots A_n$ forms a partition in Ω, then

$$P_X(x) = \sum_i P_{X|A_i}(x) P(A_i)$$

- Conditioning an RV X on another RV Y

$$P_{X|Y}(x) = P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

$$\Rightarrow P_{X|Y}(x) = \frac{P_{XY}(x, y)}{P_Y(y)}$$

- Conditional Expectation :-

$$E[X \mid Y = y] = \sum_{x \in X} x P_{X|Y}(x|y)$$

$$E[X \mid Y] = \sum_{y \in Y} \sum_{x \in X} x P_{X|Y}(x|y)$$

$$E[X \mid A] = \sum_{x \in X} x P_{X|A}(x) \quad \forall A \subseteq \Omega$$

$$E[g(X) \mid A] = \sum_{x \in X} g(x) P_{X|A}(x) \quad \forall A \subseteq \Omega$$

- Total Expectation Theorem :-

If $A_1, A_2 \ldots A_n$ form a partition of Ω, with $P(A_i) > 0 \ \forall i$, then

$$E[X] = \sum_{i=1}^{n} E[X \mid A_i] P(A_i)$$

Proof: $\sum_{i=1}^{n} E[X \mid A_i] P(A_i)$

$$= \sum_{i=1}^{n} \sum_{x \in X} x P_{X|A_i}(x) P(A_i)$$

$$= \sum_{x \in X} x \sum_{i=1}^{n} P_{X|A_i}(x)\, P(A_i)$$

$$= \sum_{x \in X} x \sum_{i=1}^{n} \frac{P(X = x \cap A_i)}{\cancel{P(A_i)}} \cancel{P(A_i)}$$

$$= \sum_{x \in X} x \sum_{i=1}^{n} P(X = x \cap A_i)$$

$$= \sum_{x \in X} x\, P(X = x)$$

$$= \sum_{x \in X} x\, P_X(x)$$

$$= \underline{\underline{E[X]}}$$

$|||^{ly}$ $\quad E[X] = \sum_{y} E[X \mid Y = y]\, P_Y(y)$

- <u>Conditional Expectance as a RV</u> :-

$$\phi(y) \triangleq E[X \mid Y = y] = \sum_{x \in X} x\, P_{X|Y}(x \mid y)$$

$\phi : \text{Range}(Y) \longmapsto \mathbb{R}$

$\phi(y)$ is a function of an RV. Therefore it is also an RV.

- <u>Theorem:</u> $\quad E[\phi(Y)] = E[X]$

$$E[\phi(Y)] = \sum_{y \in Y} \phi(y)\, P_Y(y)$$

$$= \sum_{y \in Y} E[X | Y = y] \, P_Y(y)$$

$$= E[X] \quad \text{(Total Expectation Theorem)}$$

- **Conditional Independence of RVs :-**

X and Y are conditionally independent given A with $P(A) > 0$ if

$$P_{X,Y|A}(x, y) = P_{X|A}(x) \, P_{Y|A}(y)$$

- **Conditional Variance :**

$$Var[X | Y = y] = E\left[(X - E[X|Y=y])^2 \mid Y = y \right]$$

Let $\psi(y) = Var[X | Y = y]$, $\psi(y)$ is an Rv.

$$Var[X | Y = y] = E[X^2 | Y = y] - (E[X | Y = y])^2 \longrightarrow$$

Same Proof as in Non - Condition - al Case

- **Law Of Total Variance :-**

$$Var[X] = E[\psi(Y)] \neq Var[\phi(Y)]$$

$$E[\psi(Y)] = \sum_{y \in Y} \psi(y) \, P_Y(y) \quad \text{(LOTUS)}$$

$$= \sum_{y \in Y} \left(E[X^2 | Y = y] - E[X | Y = y]^2 \right) P_Y(y)$$

$$= \sum_{y \in \mathcal{Y}} E[X^2 | Y = y] P_Y(y) - \sum_{y \in \mathcal{Y}} E[X | Y = y]^2 P_Y(y)$$

$$\sum_{y \in \mathcal{Y}} E[X^2 | Y = y] P_Y(y) = E[X^2] \quad (\text{LOTUS})$$

$$\sum_{y \in \mathcal{Y}} E[X | Y = y]^2 P_Y(y) =$$

→ **Memoryless Property of a Geometric RV :-**

$$P_X(k) = (1-p)^{k-1} p \quad , \quad k \in \mathbb{N}$$

$$P(X > n) = \sum_{k=n+1}^{\infty} (1-p)^{k-1} p$$

$$= p \frac{(1-p)^{n+1-1}}{1-(1-p)} = \underline{(1-p)^n}$$

Property: $P(X > m+n \mid X > m) = P(X > n)$

$$P(X > m+n \mid X > m) = \frac{P(X > m+n \cap X > m)}{P(X > m)}$$

$$= \frac{P(X > m+n)}{P(X > m)}$$

$$= \frac{(1-p)^{m+n}}{(1-p)^m}$$

$$= \underline{\underline{(1-p)^n}}$$

- CDF of an RV is given by,

$$F_X(x) = P(X < x) = \sum_{x_i \leq x} P_X(x_i)$$

$$\Rightarrow \quad P_X(x) = F_X(x) - F_X(x-1)$$

Example: $X = \max\{X_1, X_2, X_3\}$, X_1, X_2, X_3 are independent.

$$P_{X_i}(k) = \frac{1}{100} \quad \forall\, i \in \{1,2,3\}, \ \forall\, k \in [1:10]$$

$$P_X(k) = P(X = k) = P(\max\{X_1, X_2, X_3\} = k)$$

$$F_X(k) = P(X \leq k) = P(\max\{X_1, X_2, X_3\} \leq k)$$