

## Probability and Statistics

### Random Experiments and Sample Spaces -

↓  
Experiments involving  
Randomness  
Coin toss, dice roll, etc.

→  $\Omega$  → Set of all possible outcomes  
of a random experiment. It could be  
finite or infinite.

$$\Omega_c = \{\text{Head, Tail}\}, \Omega_d = \{1, 2, 3, 4, 5, 6\}$$

$$2\text{-coins} \rightarrow \Omega_{2c} = \Omega_c \times \Omega_c = \{H, T\} \times \{H, T\}$$

### Outcomes and Events

↓  
(omega)  $\omega \in \Omega$  is called a sample point  
or a possible outcome

A subset  $A \subseteq \Omega$  is called an event.

Events in the coin toss experiment  $C_1 = \{T\} \quad [ \subseteq \Omega_c ]$

Events in the die roll  $D_1 = \{6\}, D_2 = \{1, 3, 5\} \quad [ \subseteq \Omega_d ]$

Events are any subset of  $\Omega$ , even null sets, but  $P(\emptyset) = 0$

Probability of an event  $A = P(A)$ .

It may or may not be possible to measure / assign  $P$  for every  
subset  $A$ .

~~$P_c(T) = 0.5$~~

Probability measure  $P$  is a set function. It acts on sets and measures  
the probability of such sets

Set Theory (01):-

$A^c$  = complement of A

$\emptyset$  = denotes empty set [belongs to every set]

$A \cup B$  = A union B

$A \cap B$  = A intersection B

~~A \ B~~

$A \setminus B$  = A minus B =  $A \cap B^c$

Symmetric

$A \Delta B$  =  $(A \setminus B) \cup (B \setminus A)$

M.E = Mutually Exclusive

~~H.W~~ → Identity (complement),

$|A|$  = no. of elements in A = cardinality of A.

Inclusion - Exclusion Principle =  $|A \cup B| = |A| + |B| - |A \cap B|$

Countable Sets & Uncountable Sets

Monotone Seq →  $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$  [Increasing seq]  
 $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$  [Decreasing]

$$I_n = [0, 1 - \frac{1}{n}]$$

$$D_n = [0, \frac{1}{n}]$$

Cartesian product :-  $A \times B = \{(a, b) : a \in A, b \in B\}$

Powerset  $[\mathcal{P}(A)]$  :- Set of all possible subsets of A.

$$\text{Ans} \quad |\mathcal{P}(A)| = 2^{|A|} \quad [\text{Only for discrete elements}]$$

$$\mathcal{P}([0, 1]) = \{(a, b) : a \leq b, a, b \in [0, 1]\}$$

↳ range, not just 2 elements

Functions:-

Fns are maps from elements in the domain D to the range R

$$f: D \rightarrow R$$

$$f: \mathbb{R} \rightarrow \mathbb{R} \Rightarrow f(x) = x$$

Set fns are those fns ~~who~~ who act on sets; Dis a collect

of sets.

IP is a set for.

Axioms of probability:-

$$1) \text{ } P(\emptyset) = 0, P(\Omega) = 1$$

$$2) \text{ For a set } A \subseteq \Omega, P(A) \in [0, 1]$$

3) For a disjoint collection of events  $A_1, A_2, \dots, A_j$  where  $A_i \subseteq \Omega$ , then:

$$P\left(\bigcup_{i=1}^j A_i\right) = \sum_{i=1}^j P(A_i) \quad [j \text{ can be } \infty] \rightarrow \Omega \text{ must be countable}$$

$$\text{where } A_i \cap A_j = \emptyset, i \neq j$$

In general, the domain of IP is  $\mathcal{P}(\Omega)$ .

~~P. probability of impossible event = 0.~~

But

Counter example for  $\mathcal{P}(\Omega)$  being satisfactory

1) Pick a number randomly (uniformly) of  $t$  from the real line.

2)  $\Omega = \mathbb{R}$ , hence  $P(\mathbb{R}) = 1$ ,

3) Domain =  $\mathcal{P}(\mathbb{R})$

4)  $P: \mathcal{P}(\mathbb{R}) \rightarrow [0, 1]$

5) IP has the property that sets of equal length have equal probability

6) We know that  $\mathbb{R} = \bigcup_{n=-\infty}^{\infty} [n, n+1]$  where  $[n, n+1] \in \mathcal{P}(\mathbb{R})$

[Countable union since  $n$  is an integer]

7)  $P(\mathbb{R}) = 1 = \sum P([n, n+1]) = 0 / \infty$  depending

8) This is a contradiction! [Happens when picking segments of unit length]

Powerset is thus a bad choice

Not all set for can be calibrated to measure all possible subset of the sample space.

Our choice of domain must have certain features.

Let domain =  $\Omega$

then  $\emptyset, \Omega \in \mathcal{F}$ ; if  $B \in \mathcal{F} \rightarrow B^c \in \mathcal{F}$  etc.

$\mathcal{F}$

Bags that meet these criteria are called sigma-algebrae

Sigma algebra / Event space  $[\mathcal{F}]$  associated with a set  $\Omega$  is a collection of subsets of  $\Omega$  that satisfy

$\rightarrow \emptyset \in \mathcal{F}$  and  $\Omega \in \mathcal{F}$

$\rightarrow A \in \mathcal{F} \rightarrow A^c \in \mathcal{F}$

$\rightarrow A_1, A_2, \dots, A_n \in \mathcal{F} \rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

$\rightarrow$  The  $\sigma$ -algebra is closed under formation of complements and countable unions and countable intersections (by De Morgan's)

$\rightarrow$  When  $\Omega$  is countable and finite, then  $\mathcal{P}(\Omega)$  is a  $\sigma$ -algebra

If  $\Omega$  is countable & finite, we will consider  $\mathcal{P}(\Omega)$  as the domain.

Probability Space =  $\{\Omega, \mathcal{F}, P\}$

Probability measure:- A probability measure  $P$  on the measurable space  $(\Omega, \mathcal{F})$  is a set function  $P: \mathcal{F} \rightarrow [0, 1]$  s.t.

1)  $P(\emptyset) = 0, P(\Omega) = 1$

2) For disjoint collection of events  $A_1, A_2, \dots$  from  $\mathcal{F}$ , we have

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

P Space for  $U[0, 1]$

$$\Omega = [0, 1]$$

If we generate a  $\sigma$ -algebra, but we get a borel-sigma algebra  $\mathcal{B}[0, 1]$

## Borel $\sigma$ -algebra $[\mathcal{B}[0,1]]$

Defined when  $\Omega = [0,1]$ , as the  $\sigma$ -algebra generated by closed sets of the form  $[a,b]$  where  $a \leq b$  &  $a,b \in [0,1]$

$$(a,b] = \bigcup_{n=1}^{\infty} \left[ a + \frac{1}{n}, b - \frac{1}{n} \right] \quad \left\{ \begin{array}{c} a+\frac{1}{n} \\ b-\frac{1}{n} \end{array} \right\}$$

$$(a,b] = \bigcup_{n=1}^{\infty} (a, b + \frac{1}{n})$$

$[\mathcal{B}[0,1]]$

Borel  $\sigma$ -algebra  $\uparrow$ : is the  $\sigma$ -algebra generated by sets of the form  $[a,b]$  or  $(a,b)$  or  $(a,b]$  or even  $[a,b)$  where  $a \leq b$  &  $a,b \in [0,1]$

If  $\Omega = \mathbb{R}$ , the  $\mathcal{B}(\mathbb{R})$  is the  $\sigma$ -algebra generated by open sets of the form  $(a,b)$  where  $a < b$  and  $a,b \in \mathbb{R}$ .

How to define  $\mathcal{B}(\mathbb{R}^2)$

(Consequences of the Probability Axioms:-

i)  $P(A^c) = 1 - P(A)$

$$P(A \cup A^c) = P(A) + P(A^c) = 1$$

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

~~P(AUB)~~  $\rightarrow$  Prove it

iii) If  $A \subseteq B$ , prove that  $P(A) \leq P(B)$  [ $A \subseteq B$  can be interpreted as  $A \rightarrow B$ ]

If  $A \subseteq B \rightarrow B = A + \lambda$ .

iv)  $P\left(\bigcup_{i=1}^{\infty} B_i\right) \leq \sum_{i=1}^{\infty} P(B_i)$  [Boole's ineq.]

Diff b/w Impossible event vs Zero probability event

$\emptyset$

finite sized set

$$\text{In } U[0,1], P(\omega = 0.5) = 0$$

Every experiment outcome  $\omega$  of this experiment is a zero probability event.  
This implies that zero outcome events could occur.

$\emptyset$  on the other hand, can never occur, and are hence impossible events.

$$P(\omega \in [0, 0.25] \cap [0.75, 1]) = 0$$

Limits and Continuity:-

Limits:-

Let  $a_1, a_2, \dots, a_n$  be a sequence with limit L.

Then  $\forall \epsilon, \exists N_\epsilon$  s.t.  $\forall n > N_\epsilon, |a_n - L| \leq \epsilon$

For a fn.  $f(x)$ ;  $\lim_{x \rightarrow c} f(x) = f(c)$

The limit exists only if  $\lim_{x \rightarrow c^+} f(x) = \lim_{x \rightarrow c^-} f(x)$ .

All points  $x$  close to  $c$  are  $\epsilon$  close to  $f(c)$ .

Continuity:-

$LHL = RHL = f(x)$ . If the fn. is cont at all pts. it is said to be continuous. As  $x \rightarrow c$ ,  $f(x) \rightarrow f(c)$

But what about IP?

For a continuous set fn. S as  $A_n \rightarrow A$ , we have  $S(A_n) \rightarrow S(A)$

Sequence of sets :-

Given  $(\Omega, \mathcal{F})$ , if  $A_1 \subset A_2 \subset \dots$  is an increasing sequence of events defined on  $\mathcal{F}$  and  $\bigcup_{n=1}^{\infty} A_n = A \in \mathcal{F}$ , then we say that the given seq of sets  $A_n$  are increasing to  $A$  ( $A_n \uparrow A$ )

Similarly when  $A_1 \supset A_2 \supset \dots$  is a decreasing seq of events and  $\bigcap_{n=1}^{\infty} A_n = A$ , then we have  $A_n \downarrow A$

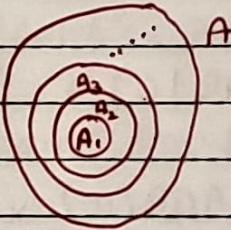
Alt notation:- increasing seq of sets  $A_n \Rightarrow \lim_{n \rightarrow \infty} A_n$  for  $\bigcup_{n=1}^{\infty} A_n$

decreasing seq of sets  $A_n \Rightarrow \lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$

Lemma:- For a sequence of events of the type  $A_n \uparrow A$  or  $A_n \downarrow A$ , we have  $\lim_{n \rightarrow \infty} P(A_n) = P(A)$

Proof:-

$$A = \bigcup_{i=1}^{\infty} A_i$$



i) Consider an increasing seq.

define  $F_n = A_n - A_{n-1}$

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} F_n = \sum_{n=1}^{\infty} F_n = \lim_{n \rightarrow \infty} \sum_{i=1}^{\infty} P(F_i)$$

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

Hence Proved.

ii) Decreasing seq. [H.W.]

Conditional Probability :-

Ex:- Given the die outcome is odd, what is  $P(1)$ ? [ $\frac{1}{3}$ ]  
Given  $\bar{\omega} \in [0, 0.5]$ , what is  $P(\bar{\omega} \in [0, 0.25])$

Defn:- The conditional probability of event A is defined as  
 $P(B|A) = \frac{P(A \cap B)}{P(A)}$  when  $P(A) > 0$

Theorem:-  $P(A|B) \cdot P(B) = P(B|A)P(A)$

$$\text{or } \frac{P(A \cap B)}{P(B)} \times P(B) = \frac{P(B \cap A)}{P(A)} \times P(A)$$

$\Phi$  Bayes Rule :-  $P(B|A) = \frac{P(A|B) \times P(B)}{P(A)}$

i) What is  $P(A|B \cap C)$

ii)  $P(A \cap B) = P(A|B) \times P(B)$

iii)  $P(A \cap B|C) = P(A|B \cap C) \times P(B|C)$

$$\frac{P(A \cap B \cap C)}{P(C)} \times \frac{P(C)}{P(B \cap C)} \times \frac{P(B \cap C)}{P(C)}$$

$$P(A|B \cap C) \times P(B|C)$$

iv)  $P(A \cap B \cap C) = P(A) \times P(B|A) \times P(C|A \cap B)$

$$\frac{P(A \cap B \cap C)}{P(A \cap B)} \times \frac{P(A \cap B)}{P(A)} \times \frac{P(A)}{P(A)}$$

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

Q) Draw 4 cards without replacement. What is the  $P(\text{9c, 8d, Ks, Kc})$

$$P(9c) \times P(8d | 9c) \times P(Ks | 9c \cap 8d) \times P(Kc | 9c \cap 8d \cap Ks)$$

$$= \frac{1}{52} \times \frac{1}{51} \times \frac{1}{50} \times \frac{1}{49} = \frac{48!}{52!}$$

When every outcome is equally likely in a finite sample space  $\Omega$   
 $\Rightarrow P(B|A) = \frac{|A \cap B|}{|A|}$

Law of Total Probability:-

$$A = (A \cap B) \cup (A \cap B^c), \therefore P(A) = P(A \cap B) + P(A \cap B^c)$$

$$\text{Same as } P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

Let  $B_1, B_2, \dots, B_n$  be the partition of the sample space  $\Omega$   
 Then for any event A we have

$$P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A|B_i)P(B_i)$$

Q) 3 bags with M marbles total. Bag i has  $R_i$  red and  $B_i$  blue  
 for  $i \in [1, 3]$ . Find  $P(\text{marble} = \text{red from rand bag})$

$$\frac{1}{3} \times \sum_{i=R_i+B_i}^3 R_i = \frac{1}{3} \times \sum_{i=1}^3 P(\text{Red} | B_i) \times P(B_i)$$

Independence:-

An event A is independent (w.r.t on event B) if  
 $P(A|B) = P(A)$

Ex:- Simultaneously Tossing a coin and rolling a die

$$\Omega_1 = \{\text{H, T}\} \quad \Omega_2 = \{1, 2, 3, 4, 5, 6\}$$

$$\Omega = \Omega_1 \times \Omega_2$$

$F = \text{Powerset } (\Omega)$

$$P(\{\text{H, 6}\}) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$$

$$P(\{\text{T, odd}\}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$$

$$P(A \cap B) = P(A) P(B) \Leftrightarrow \text{Independence}$$

If A and B are independent

$$P(A^c) = 1 - P(A) ; P(B^c) = 1 - P(B)$$

$$P(A^c) \times P(B^c) = 1 - P(A) - P(B) + P(A \cap B) = 1 - P(A \cup B) = (P(A \cup B))^c = P(A^c \cap B^c)$$

Where to now?

$A^c \text{ and } B^c$  are independent

$$\text{Prove } P\left(\bigcup_{i=1}^n A_i\right) = 1 - \prod_{i=1}^n (1 - P(A_i))$$

## Mutual and Pairwise independence

A collection of events  $\{A_i, i \in I\}$  are said to be **mutually independent** if the  $P(\bigcap_{j \in J} A_j) = \prod_{j \in J} P(A_j)$  for any subset  $J \subseteq I$

A collection of events  $\{A_i, i \in I\}$  are said to be **pairwise independent** if any pair of events from the collection are independent  $\Downarrow$  all pairs.

**Ex:-** Pick a random number from  $\{1, \dots, 10\}$

$$P(A) = \{w \mid w < 7; w \in \{1, \dots, 10\}\} = \frac{3}{5}$$

$$P(B) = \{w \mid w < 8; w \in \{1, \dots, 10\}\} = \frac{7}{10}$$

$$P(C) = \{w \mid w \% 2 = 0; w \in \{1, \dots, 10\}\} = \frac{1}{2}$$

$$P(A \cap B) = P(A) \rightarrow \text{not independent}$$

$$P(A \cap C) = P(A) \times P(C) \rightarrow \text{independent}$$

$$P(B \cap C) = \frac{3}{10} \neq P(B) \times P(C) \rightarrow \text{not independent}$$

Correlation b/w events :-

Two events A & B are +vely correlated iff  $P(A|B) > P(A)$

Two events A & B are -vely corrrelated iff  $P(A|B) < P(A)$

A & B have same correlation as  $A^c$  &  $B^c$  (Prove)

A & B have opp correlation as  $A^c$  &  $B^c$  (Prove)

Mutual Exclusivity :-

Two events are mutually exclusive if  $P(A \cap B) = 0$

The occurrence of one implies the other cannot occur.

Two events with non zero probabilities cannot be both mutually exclusive and independent.  
 $P(A \cap B) = 0 = P(A)P(B) \rightarrow \text{contradiction.}$

If A & B are ME

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{P(A)}{P(B^c)}$$

When  $A \subset B$ , A & B are not ME nor I

Zero Probability Events are always independent

Let  $P(E) = 0$

Then for any set F, we want to show that  $P(E \cap F) = 0$

But  $E \cap F$  has two choices

- (a)  $E \cap F = \emptyset$
- (b)  $E \cap F \subset E$

In either case  $P(E \cap F) = 0$  or  $P(E \cap F) \leq P(E) = 0$   
 $P(E \cap F) = 0 \quad \leftarrow$

Conditional Independence:-

Two events A & B are conditionally independent of C with  
 $P(C) > 0$  if  $P((A \cap B)|C) = P(A|C) \cdot P(B|C)$

It thus follows that  $P(A|B \cap C) = P(A|C)$

Ex:- Given a fair coin and a fake coin (2 heads). Pick a random coin and toss twice.

A:- First toss  $\rightarrow H = 3/4$

B:- Second toss  $\rightarrow H = \text{still } 3/4$

~~Event C~~ :- First coin is chosen =  $\frac{1}{2}$

$$\begin{aligned} P(A \cap B) &= P(A \cap B | C) + P(A \cap B | C^c) \\ &= \underbrace{\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}}_{\text{Fair coin}} + \underbrace{\frac{1}{2}}_{\text{Unfair coin}} \\ &= \frac{1}{8} \neq \frac{3}{4} \times \frac{3}{4} \rightarrow \text{Not independent} \\ &\quad (C \text{ coin is chosen and it is tossed twice; not one coin being chosen and tossing twice.}) \end{aligned}$$

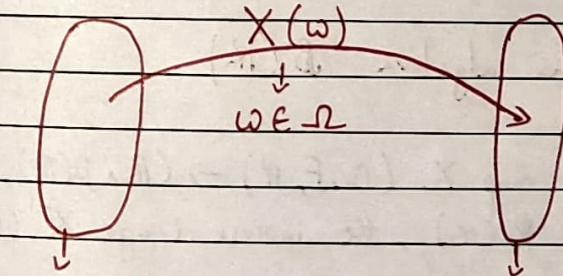
A & B are conditionally independent on C, not  $C^c$

### Random Variables

Given a random experiment with associated  $(\Omega, \mathcal{F}, P)$ , it is sometimes difficult to deal directly with  $\omega \in \Omega$ .  $\omega$  is not a number, it is a sequence (vector) of nos.

Random variables are devices that help us make a mapping from  $(\Omega, \mathcal{F}, P)$  to  $(\Omega', \mathcal{F}', P_X)$ , as this would help analyze functions of sample pts rather than any sample pt.

Assume you want to count no. of 6s in 10 rolls



$\Omega$  = set of all  $\omega$  vectors

$\Omega' = \{1, \dots, 10\} : \mathcal{F}'$ ; the Powerset  $(\Omega')$

$P \oplus P_X$

A random variable  $X$  is a fn  $X: \Omega \rightarrow \Omega'$  that transforms the probability space  $(\Omega, \mathcal{F}, P)$  to  $(\Omega', \mathcal{F}', P_X)$  and is  $(\mathcal{F}, \mathcal{F}')$ -measurable.

The map  $X: \Omega \rightarrow \Omega'$  implies  $X(\omega) \in \Omega' \quad \forall \omega \in \Omega$

For the event  $B \in \mathcal{F}'$ , the preimage  $X^{-1}(B)$  is defined as  $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$

for every  $B \in \mathcal{F}'$ , we have  $X^{-1}(B) \in \mathcal{F}$

$X^{-1}$  is a one to many mapping from  $\mathcal{F}' \rightarrow \mathcal{F}$

$P_X(B) =$  Probability that  $X$  results in  $\{B\}$  while being acted upon by  $\omega \in \Omega$

= Probability of  $P(\omega)$  where  $\omega \in \Omega ; X(\omega) \in B$

$\mathcal{F}$  could have elements that have no ~~B~~ images in  $\mathcal{F}'$ . But  $\forall f \in \mathcal{F}'$ ,  $f$  has a pre image in  $\mathcal{F}$

$P_X$  is called induced probability measure

Conventions:-

i)  $\Omega' = \mathbb{R}$

ii)  $\mathcal{F}' = \text{Borel sigma algebra } \mathcal{B}(\mathbb{R})$

A rand var  $X$  is a map  $X: (\Omega, \mathcal{F}, P) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$  such that for each  $B \in \mathcal{B}(\mathbb{R})$ , the inverse image  $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$  satisfies

$$X^{-1}(B) \in \mathcal{F}$$

$$P_X(B) = P(\omega \in \Omega : X(\omega) \in B) = P(X^{-1}(B))$$

If  $\Omega'$  is countable, then the random variable is called a discrete random variable. Only here is  $\mathcal{F}'$  used as the Powerset.

Capital letters X, Y, Z etc. for rand var, and x, y, z etc. are the realizations of said rand var.

### Discrete Random Variables :-

Ex:- Roll two dice and calculate the sum

$$\begin{aligned}\Omega &= \text{set of all tuples } (x, y) \text{ where } x, y \in [1, 6] \\ \Omega' &= \text{integers } [2, 12]\end{aligned}$$

$\mathcal{F}$  &  $\mathcal{F}'$  are Powersets

$$\{x = 3\} \in \mathcal{F}' \rightarrow P_x(3) = P(\{1, 2\}, \{2, 1\})$$

In general for  $x \in \Omega'$   $P_x(x) = P(\{(a, b) | a, b \in \Omega, a+b=x\})$

$$P_x(x) = P_x(\{x\})$$

$\hookrightarrow$  fn [not a probability measure]  $\rightarrow$  makes  $x$  a set and passes  
as arg to  $P_x$ .  
 $\hookrightarrow$  Probability mass fn.

$$P_x(x) = \begin{cases} x-1/36 & \forall x \in \text{int}[2, 7] \\ (13-x)/36 & \forall x \in \text{int}[8, 12] \end{cases}$$

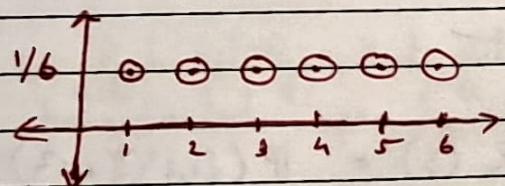
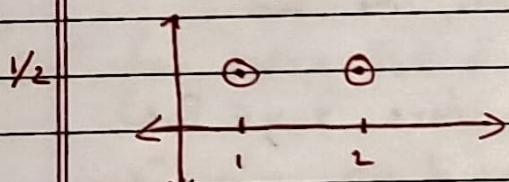
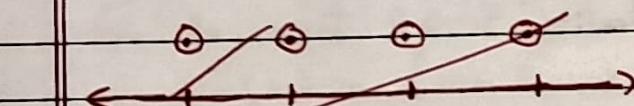
Multiple  $X_i$ 's can have the same Probability Mass fn.

For example, for the exp of rolling a die four times, then divide and conquer using  $P_x(x)$  mentioned above.

Only if  $X_i$ 's are independent

The fn  $p(x) := P_x(\{x\})$  for  $x \in \Omega'$  is called as a probability Mass Function (PMF) of a random variable  $X$ .

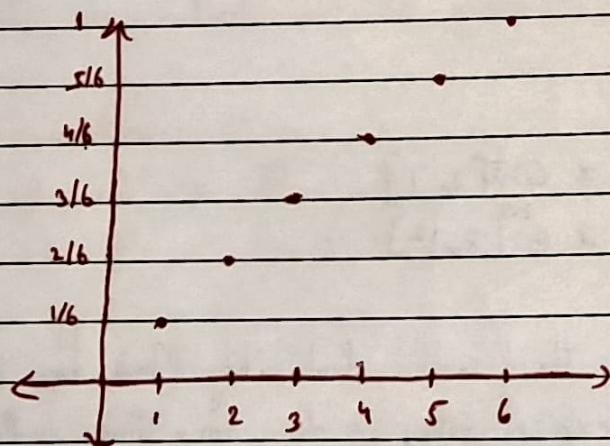
Q) Plot PMFs for  $X$  corresponding to coin tosses or dice rolls.



The Cumulative Distribution Function (CDF) is defined as  

$$F_X(x_i) = \sum_{x \leq x_i} p_x(x) = P\{\{ \omega \in \Omega : X(\omega) \leq x_i \}\}$$

$F_X$  (Rolling dice)



$\Omega'$

The X-axis is only  $X$  not  $\Omega$ . Hence the mapping  $X$  is identical to  $\Omega$ .

## Expectation and Moments:-

The mean or expectation of a random variable  $X$  is denoted by  $E(X)$  and is given by  $E(X) = \sum_{x \in \Omega} x p_x(x)$

The expected value need not be part of  $\Omega$

Depends on how  $X$  is defined.

First moment of a rand var.

The  $n^{\text{th}}$  moment of a random variable  $X$  is denoted by  $E(X^n)$  and is given by  $E(X^n) = \sum_{x \in \Omega} x^n p_x(x)$

For a fn  $g(\cdot)$  of a rand var  $X$ ,  $E(g(X)) = \sum_{x \in \Omega} g(x) p_x(x)$

fns of random variables are random variables with their own probability mass fns.

## Consistency of PMF:-

PMF:  $p_x(x) = P(\{\omega \in \Omega : X(\omega) = x\})$  for  $x \in \Omega'$

To prove  $\sum_{x \in \Omega'} p_x(x) = 1$

$$\sum_{x \in \Omega'} p_x(x) = \sum_{x \in \Omega'} P(\{\omega \in \Omega : X(\omega) = x\})$$

$$= P\left(\bigcup_{x \in \Omega'} \{\omega \in \Omega : X(\omega) = x\}\right)$$

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

Expectance is Linear :-

$$\text{Let } Y = ax + b$$

$$\text{Then } E(Y) = E(ax + b)$$

$$= \sum_{x \in \Omega} (ax + b) p_x(x)$$

$$= a \left( \sum_{x \in \Omega} x p_x(x) \right) + b$$

$$= a E(X) + b$$