# References :-


1. Athanasis Papoulis, "Probability and Statistics", Prentice - Hall International Inc.

2. ~~Erwin~~ ~~John E. Freund~~'s, "Mathematical statistics with Applications", Pearson Education, ~~ISBN~~ : 81-7758-755-2
Irwin Miller, Marylees Miller, ISBN, 2006.

3. ~~Pro~~ Michael Mitzenmacher, Eli Upfal, "Probability and computing : Randomized Algorithms and Probabilistic Analysis", Cambridge University Press, 2005.

4. Henry Stark, John W. Woods, "Probability and Random Processes, with applications to signal processing", Pearson Education International, 2002.

5. Oliver C. Ibe, "Fundamentals of Applied Probability and Random Processes", (2nd Edition), Academic Press, 2014.

6. A.M. Mathai, "Basic Probability And Statistics - I* Probability and Random Variables", Module 6.

## Probability :- For computing probability (In classical) we need some idea of permutation and combination in order to find out the total number of elements in sample space as well as total number of elements in event space.

~~Permutations:~~

**Permutation :-** To Permute ⇒ To rearrange

Number of Permutation ⇒ Number of such ~~arrange~~ rearrangements.

**E.g. :-** Problem of filling some positions with some objects.

Let $r$ – seats and $n$ – individual seated on these $r$ seats.
(Position) (object)

Let $r = 2$ $\quad$ $n = 5$

$n_1 \quad n_2 \quad n_3 \quad n_4 \quad n_5$

$r_1 \quad r_2$

– $r_1$ can be assigned any one of the $n_1, n_2, n_3, n_4, n_5$

~~when~~ first seat ~~is already~~ is filled (i.e. $r_1$)

There are five choices to fill $r_1$

– When first position, $r_1$, is filled by any one of the five objects, then second position will be filled by only four object in four different ways.

⇒ Thus total number of choices for filling these two seats $5 \times 4 = 20$.

- If first seat is given to $n_1$ then the sequence possible for two seats:

$n_1 n_2, \; n_1 n_3, \; n_1 n_4, \; n_1 n_5$

- $n_2$ is given first ~~first~~ seat, then sequence.

$n_2 n_1, \; n_2 n_3, \; n_2 n_4, \; n_2 n_5$

- $n_3$ is given first seat then

$n_3 n_1, \; n_3 n_2, \; n_3 n_4, \; n_3 n_5$

- $n_4$ is given first seat then

$n_4 n_1, \; n_4 n_2, \; n_4 n_3, \; n_4 n_5$

- $n_5$ is given first seat then

$n_5 n_1, \; n_5 n_2, \; n_5 n_3, \; n_5 n_4$

"This is stated as the "number of permutations of five, taken two at a time".

$\Leftrightarrow$ "Total number of ordered set of two from a set of fiver or the total number of sequences of two items taken from a set of five items.

In General :- If there are $n$-individuals and $r$ seats to be filled then the total number of choices for filling up these $r$ seats with $n$-individuals is $n(n-1)(n-2)\cdots(n-(r-1))$

Notation :- $P(n, r) = {}_nP_r$ = Total number of ~~permutation~~ permutations of $n$, taken $r$ at a time

Definition (Permutations) :— The total number of permutations of n objects, taken r at time or total number of ordered sets of 'r' items from the set of 'n' items is given by:

$$\boxed{P(n,r) = n(n-1)(n-2)(n-3),\cdots(n-r+1)}$$

e.g. :→ Total number of permutations of 5 items, taken 3 at a time is:

$$5\times4\times3 = 60$$

– Total number of permutations of 5, taken 5 at a time or all is $5\times4\times3\times2\times1 = 120$

Note :– $P(1,1) = 1$, $P(2,1) = 2\times1 = 2$, $P(n,1) = n(n-1)(n-2)\cdots(n-\cancel{1}+1) = n$

$P(10,2) = 10\times9 = 90$, $P(4,4) = 4\times3\times2\times1 = 24 = 4!$

Notation :– $n!$ ≡ factorial n or n factorial

Definition :— $n! = 1\times2\times3\times4\times\cdots\times n$,

$0! = 1$ (Convention)

Note :– ① $P(n,r) = n(n-1)(n-2)\cdots(n-r+1)$

$$= \frac{n(n-1)(n-2)\cdots(n-(r-1))(n-r)(n-r-1)\cdots2\cdot1}{(n-r)(n-r-1)\cdots2\cdot1}$$

multiplying and dividing by $(n-r)(n-r-1)\cdots2\cdot1$

$$\boxed{P(n,r) = \frac{n!}{(n-r)!}}$$ → Factorial Representation of $P(n,r)$

② $P(n,n) = \frac{n!}{(n-n)!} = \frac{n!}{0!}$

But from original definition

$P(n,n) = n(n-1)(n-2)\cdots(n-n+1)$
$= n(n-1)(n-2)\cdots2\cdot1$

$P(n,n) = n!$

Thus we need convention $0! = 1$, when we want to use factorial representation of Permutation

[ Mathematical conventions are ~~Mathematical~~ convenient assumptions, which will not contradict or interfere with any of the mathematical derivations or computations]

Ex :- How many different 3-letter words can be made by using all the alphabets in the word (1) "can"; (2) How many different 4-letter words can be made by using all the alphabet of word "good" (3) How many 11-letter words can be made ~~by~~ using all alphabets in the word "Mississippi".

Soln :- (1) The different words are the following
can, cna, anc, acn, nac, nca

There $6 = 3!$ ~~letter~~ such words

$$P(n,r) = P(3,3) = 3!$$

(2) "good" ⇒ In this letter 'o' repeated two times.
↳ If these o's are different say $o_1, o_2$, Then total number of words possible

$$P(4,4) = 4! = 4 \times 3 \times 2 \times 1 = 24$$

But $o_1 o_2$ or $o_2 o_1$ gives the same sequence oo, $o_1$ and $o_2$ are permuted $2! = 2$ ways
↳ All these permutation generate the same words

⇒ Thus total number of distinct permutation/word is

$$\frac{4!}{2!} = \frac{4 \times \cancel{3} \times \cancel{2} \times 1}{2 \times 1} = 12$$

(3) In (3) letter "s" is repeated 4 times
"i" is repeated 4 times
"p" is repeated 2 times.

Hence total number of distinct word possible are

$$\frac{11!}{4!\,4!\,2!} = \frac{11\times10\times9\times8\times7\times6\times5\times\cancel{4!}}{\cancel{4!}\,4!\,2!}$$

$$= \frac{11\times10\times9\times\cancel{8}\times7\times\cancel{6}\times5}{\cancel{4\times3\times2}\times1\times\cancel{2}\times1}$$

$$= 11\times10\times9\times7\times5$$

$$= 990\times7\times5$$

$$= 6930\times5$$

$$\boxed{\frac{11!}{4!\,4!\,2!} = 34650}$$

Ex :- How many number plates can be made containing only 3 digits if ① repetition of numbers is allowed, ② no repetition is allowed.

Solution :- ① A number plate of three digits ⇒ filling three position with 10 numbers 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.

⇒ The first position is filled in 10 different ways.

⇒ The second position is also filled in 10 ———

( Since repetition of numbers are allowed)

⇒ Similarly the third position is also filled in 10 different ways.

Thus total number of plates possible is: $10\times10\times10 = 10^3 = 1000$ when number is allowed.

② When repetition is not allowed then

- $I^{st}$ position is filled in 10 different ways.
- $II^{nd}$ position is filled in 9 different ways
- $III^{rd}$ position is filled in 8 different ways

∴ total number of plates possible is:

$10\times9\times8 = 720$,

## Combinations:

- In permutation we are interested in an ordered set or ordered subsets from given set ~~of~~ of objects.
- In Combination we are not interested in the order but only in the subset only. e.g.

- 3 letters a, b, c and if interested ordered sub set of two letters
  ab, ac, ba, bc, ca, cb or there are $3 \times 2 = 6$ ordered sets.
- Subset of two letters then the subsets are:
  $\{a, b\}, \{a, c\}, \{b, c\}$.
  because whether the sequences ab or ba are same subset of the letter 'a' and 'b'.

**How many subsets of r elements possible from a set of n distinct elements?**

If a subset have r element then we can ordered them in $r!$ ways to get all ordered pair.

Total ordered pair ~~taking~~ of r element out of n element is $P(n, r)$.

∴ ~~Therefore~~

~~Since~~ ∵ $r!$ ordered pair = 1 Subset ~~of Set~~
≡ 1 unordered sequence

∴ $P(n,r)$ ordered pair $= \frac{P(n,r)}{r!}$ unordered sequence

∴ Therefore total number of ~~permutation~~ combination of n taken r element at time is $\frac{P(n,r)}{r!}$

$$= \frac{n!}{r!\,(n-r)!}$$

This is denoted as ${}^nC_r$, ${}_nC_r$, $\binom{n}{r}$, $C(n,r)$

Therefore

$$ {}^nC_r = {}_nC_r = \frac{n!}{r!(n-r)!} $$

Definition:— The number of combination of $n$, taken $r$ at a time or the number of possible subset of '$r$' distinct elements from a set of $n$, ~~elements~~ distinct elements, is given by:

$$ {}^nC_r = {}_nC_r = \frac{P(n,r)}{r!} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!} $$

Clearly:— ${}^nC_n = 1$, ${}^nC_1 = {}^nC_{n-1} = n$

— ${}^nC_2 = {}^nC_{n-2} = \frac{n(n-1)}{2!}$

— ${}^nC_r = {}^nC_{n-r}$ ∋ $r = 0, 1, 2, -- n.$

— ${}^nC_r = {}^{n-1}C_r + {}^{n-1}C_{r-1}$

Note (1):— ${}^nP_r$ and ${}^nC_r$ (permutations and combinations) hold both $n$ and $r$ must be non-negative i.e. positive.

(2) On implementation ~~in~~ through computer ~~program~~ program we use basic definition not factorial since factorial computation go out of range ~~of~~ of computer representation of a number and we get wrong result.

Ex :- A box contain 7 identical marbles, except for color, of which 4 red and 3 are green. Two marbles are selected at random (a) one by one with replacement (b) one by one without replacement (c) Two marbles together

(i) Compute the number of sample points in these cases; (ii) Compute the probabilities of the getting sequence (RG) ( R ≡ Red, G ≡ Green) in (a) and (b)

(iii) Compute the probabilities of getting exactly one red and one green marbles in (a) (b) (c)

Solution :- (a) (i) It is same as filling two positions with seven (7) objects.

- First position filled by 7 ways.
- Since marble is put back ⇒ The second position is also filled by 7 ways.

Therefore, there are $7 \times 7 = 49$ total possible ways we can select the two marbles with replacement policy.

∴ Total number of sample point in sample space = 49.

(b) (i) - First position is filled by 7 ways.

- Since marble is not put back, therefore there are six ways to fill the second place.

Thus total possible ways to ~~fill the~~ select the two marbles are $7 \times 6 = 42$.

Therefore total number of sample point in sample space is 42.

(c) (i) Here we looking subsets of 2 items from a set of a 7 (seven) items. Hence sample space consists of such subsets of 2 items and total number of sample point is

$$^7C_2 = \frac{7!}{2!5!} = \frac{7 \times 6 \times \cancel{5!}}{2! \times \cancel{5!}}$$

$$^7C_2 = \frac{7 \times \cancel{6}^3}{\cancel{2}} = 21$$

$$\therefore \boxed{^7C_2 = 21}$$

a-(ii) :- Compute the probability of RG sequence.

All elementary event in (a) have probability $\frac{1}{49}$ ~~(since expri~~ (since experiment is random and symmetric)

Since Here restriction is that first place is R and Second place is G. Therefore,

- First place is filled by four ways.
- Second place is filled by three ways.

$\therefore$ Total number of RG sequence is $4 \times 3 = 12$

Thus total number of sample points favorable to ~~event~~ the event = 12.

b-ii) This is also the case in b-(ii). Therefore probability of an event in cases

a-(ii) $= \frac{12}{49}$

b-(ii) $\frac{12}{42}$

a-(iii) :- Sequence RG or GR
event contain RG or GR sequence but not RR, GG.

RG ≡ first place can be fill in 4 ways and second place by 3 ways.

$\equiv 4 \times 3 = 12$

GR ≡ First place can be fill in three ways and second place will be fill by four ways

$\equiv 3 \times 4$

GR $= 12$ ways.

∴ Hence total number of sample point favorable to the event $= 12 + 12 = 24$

∴ Probability of an event $= \frac{24}{49}$

b-(iii) This is same as a-(iii)

∴ Probability of an event $= \frac{24}{42} = \frac{4}{7}$

(c)

c-(ii) :- Total number of sample points favorable to the event of getting exactly one red and one green marbal is following: one red can come from 4 red in $^4C_1 = 4$ ways, One green come from 3 green in $^3C_1 = 3$ ways.

Thus total number of events favorable to the even $= 4 \times 3 = 12$

∴ ~~Total~~ Probability of event getting exactly one red and one green $= \frac{12}{21} = \frac{4}{7}$

Sampling without replacement (b) and taking subset of two produced the same result

~~$^4C_2 = \frac{4!}{2!\,2!} = \frac{4 \times 3 \times 2!}{2!} = 6$~~ ~~$^3C_2 = \frac{3!}{2!\,1!} = 3$~~

# Conditional Probability :—

Example :— Consider we are interested in sum of the numbers that appears when two dice are tossed. Let the event that the sum of two tosses is 7, and observe that the first ~~toss~~ is 4.

⟹ Find the probabity of the event that the sum of the two ~~dice~~ is 7 provided that first ~~dice~~ is four.

— In absence of any information about the outcome of the ~~occurence~~ of the first ~~dtoss~~ and second toss, there are 36 Sample point in the sample space.

— And in this case no restriction of the outcome of the first toss ~~and~~ there are following sample point ~~wh~~ in which sum is 7

$\{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

→ If we put the restriction that on first toss outcome is 4, then we get following six sample points favorable to this event

$\{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$

⟹ After putting restriction on first outcome the sample space reduced.

— Let we denote this event by $B$

$\therefore B = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$

— Let $A$ denote the et ~~sum~~ event that sum of two dice is 7, i.e.,

$A = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$

Conditional Event :- ~~Happ~~ Occurrence of event A provided that event B.

$$A|B \equiv \text{Event A provided that event B}$$

Conditional Probability :-

Probability of $A|B$ is called conditional probability and read as " conditional probability of event A ~~given~~ given event B".

- The conditional probability of event A given event B is denoted by $P(A|B)$, is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \; ; \quad P(B) \neq 0$$

Now $A \cap B = \{(4,3)\}$

$B = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$

$\therefore \; P(B) = P(\{4,1\}) + P(\{4,2\}) + P(\{4,3\}) + P(\{4,4\}) + P(\{4,5\}) + P(\{4,6\})$

$$P(B) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36}$$

$$P(B) = \frac{6}{36}$$

$$\therefore \boxed{P(B) = \frac{1}{6}}$$

$$P(A \cap B) = P(\{4,3\}) = \frac{1}{36}$$

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

$$\boxed{P(A|B) = \frac{1/36}{1/6} = \frac{1}{6}}$$

Note :- ① For putting condition on event (say A)

$\Downarrow$

We reducing the sample space ~~on conditi~~ of conditioned event (A)

$\Downarrow$

~~Then possible all out come is the c~~

P

Then sample space of condition event (A) is the total possible out come of conditioning event (B).

$$\therefore \quad P(A|B) = \frac{\text{no. of favorable out come of A w.r.t. B}}{\text{no. of total out come of event B}}$$

$$= \frac{\#(A \cap B)}{\#(B)}$$

In previous example $= \frac{1}{6}$.

② In classical probability (frequentist) : Actually ~~ever~~ ~~P(A)~~ Probability of any event is conditioned on total possible outcome. ($\Omega$)

$$P(A) = P(A|\Omega) = \frac{P(A \cap \Omega)}{P(\Omega)}$$

$$= \frac{P(A)}{P(\Omega)}$$

$$= \frac{\dfrac{\#\{\text{favorable out come to event A} \cancel{\text{w.r.t. } \Omega}\}}{\cancel{\#(\Omega)}}}{\dfrac{\#\{\text{favorable out come to event } \Omega\}}{\cancel{\#\Omega}}}$$

$$= \frac{\#(A)}{\#(\Omega)}$$

$$= \frac{|A|}{|\Omega|}$$

$$\boxed{P(A) = P(A|\Omega) = \frac{N_A}{N_\Omega}}$$

Ex :- A bag contains 8 red balls, 4 green, and 8 yellow balls. A ball is drawn at random from the bag and it is found ~~that~~ not to be one of the red balls. What is the probability that it is a green ball.

Solution :-

$G$: Event that a selected ball is green

$\bar{R}$: Event that it is not a red ball.

Total number of sample point in sample space ($\Omega$)

$= 8+4+8$

$= 20$

$P(G) = \frac{n_G}{n_\Omega} = \frac{\#(G)}{\#(\Omega)} = \frac{4}{20}$ [Since there are four green balls out of twenty balls]

$P(G) = \frac{1}{5}$ —— (1)

$\bar{R}$ : Event that ~~is~~ selected ball is not ~~gree~~ red ball

: Selected ball is either green or yellow

: Total number of sample point in $\bar{R}$ is sum of green balls and yellow ball

: $4 + 8 = 12$

$\therefore P(\bar{R}) = \frac{\#(\bar{R})}{\#(\Omega)} = \frac{\cancel{12}\,3}{\cancel{20}\,5} = \frac{3}{5}$ —— (2)

$\therefore P(G|\bar{R})$ = Probability of getting green ball provided that it is not red ball

= Probability of getting green ball provided that it comes from the ~~sam~~ subset of sample containing green and yellow ball

$= \frac{\#(G)}{\#(G \cup Y)} = \frac{4}{12} = \frac{1}{3}$ —— (3)

Now using definition of conditional probability

$$P(G|\bar{R}) = \frac{P(G \cap \bar{R})}{P(\bar{R})}$$

$$= \frac{P(G)}{P(\bar{R})} \quad [\because G \cap \bar{R} = G] \quad \because G \subseteq \bar{R}$$

$$= \frac{\frac{4}{\cancel{20}}}{\frac{12}{\cancel{20}}}$$

$$= \frac{\cancel{4}}{\cancel{12}} = \frac{1}{3}$$

$$\therefore \boxed{P(G|\bar{R}) = \frac{1}{3}} \quad — (4)$$

$$P(G \cap \bar{R}) = P(G) = \frac{1}{5} \quad — (5)$$

$\therefore$ ~~P(G given R̄ has occurred) =~~

$$P(G \text{ given that } \bar{R} \text{ has occurred}) = \frac{1}{3} \quad — (6)$$

Now from (2), (5) and (6) we have

$$P(G \cap \bar{R}) = \frac{1}{5} = P(G \text{ given that } \bar{R} \text{ has occurred}) \times P(\bar{R})$$

$$= \frac{1}{3} \times \frac{3}{5} = \frac{1}{5}$$

**Ex:** A fair coin ~~was~~ tossed two times. Given that the first toss resulted in heads, what is the probability that both tosses resulted in head?

**Solution** Since coin is fair $\Rightarrow$ Experiment is symmetric, the four sample point in sample space $\Omega = \{HH, HT, TH, TT\}$. Since symmetry in experiment, therefore we can assign equal probability to each outcome i.e $\frac{1}{4}$.
sample point

Let B : Event that the first toss resulted in head

$B = \{HT, HH\}$

A : Event that both toss resulted are head.

$A = \{HH\}$

Now $P(B) = \frac{2}{4} = \frac{1}{2}$ —— (1)

P( A given that B occurred) = ?

Since event B occurred ⇒ The first toss is head, thus there are two possibility on second toss probability

⇒ Thus chance of getting head at next chance is $\frac{1}{2}$.

∴ P( A given that B occurred ) $= \frac{1}{2}$ —— (2)

What is $A \cap B$ here ?

$A \cap B = \{HH\}$

∴ $P(A \cap B) = \frac{1}{4}$ —— (3)

$P(A \cap B) =$ P( A given that B occurred) × $P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

Example - A box contains seven identical ~~marbels~~ marbles, except for color, of which 4 are red and 3 are ~~green~~ green. Marbles are taken at random, one by one, without ~~replaceme~~ replacement. What is probability ~~that~~ (a): The second marbles taken ~~are~~ is green, given that the first marble removed is a red marble? (b) the second marble is green.

Solution:- A: Event that the first marble removed is ~~Green~~. red.

B: Event that the second marble is Green.

(a) It is already known that first marble removed is red, thus there are only 6 marbles left in the box out of which 3 are green and hence required probability $\frac{3}{6} = \frac{1}{2}$, i.e.

$$P(B \text{ given that } A \text{ has occurred}) = \frac{3}{6} = \frac{1}{2} \quad \text{--- (4)}$$

What is $A \cap B$?

$A \cap B$ ≡ This ~~is~~ an event that the first marble is red and second marble is green

≡ Getting the sequence RG.

≡ filling two space by 7 objects in which 4 are ~~in one typ.~~ typ1 (red), and 3 are type(2) green

such that first place is occupied by type-1

= For first place we have 4 ways to fill. X second place is filled by 3 ways

$= 4 \times 3 = 12$,

= Thus we have 12 sample point in $A \cap B$.

∴ ~~$P(A \cap B) = 12$~~

And total number of ordered sequences are $= 7 \times 6$
$= 42$

Thus number of sample point in sample space = 42.

$\therefore P(A \cap B) = \frac{12}{42} = \frac{\cancel{4}}{}$

$P(A \cap B) = \frac{2}{7}$ —— (2)

$P(A)$ ? i.e. Probability that the first marble removed is red.

~~P(A)~~ = Event A is consists of the ordered sequence RG, and RR.

There are $4 \times 3 = 12$ RG sequence.

$\cancel{4} \times 3 = 12$ RR sequence

$\therefore$ Thus total number of sample point in A is $= 12 + 12 = 24$.

$\therefore P(A) = \frac{24}{42} = \frac{4}{7}$

Now

$P(A \cap B) = \frac{2}{7} = P(B \text{ given } A \text{ has occurred}) \times P(A)$

$= \frac{1}{\cancel{2}} \times \frac{\cancel{4}^{2}}{7}$

$= \frac{2}{7}$

$\because P(A \cap B) = P(B \text{ given } A \text{ has occurred}) \times P(A)$

Note :- ① From above discussion we ~~see~~ examine that the probability of the type B given A or the probability of an ~~of~~ event given that some other event has occurred.
— In some cases that information will change the prob. probability, ~~of a~~ i.e. the conditional i.e, the probability of a conditional statement and that of an unconditional statement may different.

$B/A$ = B given that A has occurred.

Notation :— $P(B/A)$ = Probability of B given A

= The conditional probability of B given that A has already ~~occur~~ occurred, where A and B are two events in the same sample space.

$P(B|A)$ can be defined in terms of the probability of simultaneous occurrence and ~~marginal~~ the marginal probability or the probability of conditioned event.

Definition :— The conditional probability of B given A is the probability of the simultaneous occurrence of B and A divided by the probability of A, when $P(A) \neq 0$. That is:

$$\left.\begin{aligned} P(B|A) &= \frac{P(B\cap A)}{P(A)}, \quad P(A)\neq 0 \\ \Rightarrow P(A\cap B) &= P(B|A)\cdot P(A) \end{aligned}\right\} \text{—} ①$$

$$\left.\begin{aligned} P(A|B) &= \frac{P(A\cap B)}{P(B)}, \quad P(B)\neq 0 \\ P(A\cap B) &= P(A|B)\cdot P(B) \end{aligned}\right\} \text{—} ②$$

Now from ① and ② we have

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

Note ① The conditional probabity satisfy the all three axioms of Kolmogorov Axioms of a probability measure on $(\Omega, \mathcal{A})$, so that $P(B|A)$ is a probability space in its own right (and hence triplet $(\Omega, \mathcal{A}, P(A|B))$ is a probability in its own right)

Note ② The conditional probability has the following intuitive property.

(i) $A \cap B = \phi \Rightarrow P(A|B) = 0$ and $P(B|A) = 0$

(ii) $A \subset B \Rightarrow A \cap B = A$, $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)}$

$P(A|B) = \frac{P(A)}{P(B)} \Rightarrow P(A|B) \geq P(A)$ and also $P(A) \leq P(B)$

$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$

(iii) $B \subset A$

$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

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

$\therefore P(B|A) \geq P(B)$ and $P(B) \leq P(A)$

## Product Rule for Conditional Probability :-

Now from definition of conditional Probability

$$P(A_2 \cap A_1) = P(A_2|A_1) \cdot P(A_1)$$

Similarly

$$P(A_3 \cap A_2 \cap A_1) = P(A_3|A_2 \cap A_1) P(A_2 \cap A_1) =$$

$$P(A_3 \cap A_2 \cap A_1) = P(A_3|A_2 \cap A_1) P(A_2|A_1) \cdot P(A_1)$$

$P(A_2 \cap A_1) \neq 0$, $P(A_1) \neq 0$

Generalizing this product rule :

$$P(A_n \cap A_{n-1} \cap A_{n-2} \cap \cdots A_2 \cap A_1) = P(A_n|A_{n-1} \cap A_{n-2} \cap \cdots \cap A_1) \cdot P(A_{n-1}|A_{n-2} \cap A_{n-3} \cap \cdots \cap A_1) \cdots P(A_2|A_1) \cdot P(A_1)$$

$P(A_{n-1} \cap A_{n-2} \cap \cdots \cap A_1) \neq 0, \cdots, P(A_1) \neq 0.$

Ex 8— A box contains 4 red and 3 green identical marbles. Marbles are taken at random one by one. (a) without replacement (b) with replacement. What is probability of getting

(i) The sequence RRG
(ii) The sequence RGR
(iii) Exactly two red and one green marble.

Solution — (i) (a) Marbles be selected at random without replacement

A: Event that first marble is red.

B: Event that the second marble is red.

C: Event that the third marble is green

$\therefore$ Sequence RRG $\equiv A \cap B \cap C$

Now using product rule of conditional probability.

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

$P(A) = \frac{4}{7}$ ; since there are 7 marbles out of which 4 are red and the marbles are picked at random.

If one red marble is removed, then probability of getting another red marble is $\frac{3}{6}$, because there are only 6 marbles left out of which 3 are red

$\therefore \quad P(B|A) = \frac{3}{6} = \frac{1}{2}$

If the two red marbles are removed then there are 5 marbles are left out in which 3 are green.

$\therefore \quad P(C|A \cap B) = \frac{3}{5}$

$\therefore P(A \cap B \cap C) \equiv P(C|A \cap B) \cdot P(B|A) \cdot P(A) = \frac{3}{5} \times \frac{3}{6} \times \frac{4}{7}$
$\equiv P(A) \cdot P(B|A) \cdot P(C|A \cap B) = \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5}$
$= \frac{6}{35}$

$\therefore \boxed{P\{(RRG)\} = \frac{6}{35}}$

(ii)(a) R G R

A: Event that first marble is red.

B: Even that second marble is green.

C: ~~Third~~ Event that the third marble is red.

$RGR \equiv A \cap B \cap C$

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

$P(A) = \frac{4}{7}$, $P(B|A) = \frac{3}{6}$, $P(C|A \cap B) = \frac{3}{5}$

$\therefore P(A \cap B \cap C) = \frac{3}{5} \times \frac{\cancel{3}}{\cancel{6}_2} \times \frac{\cancel{4}^2}{7}$

$\boxed{P(A \cap B \cap C) = \frac{6}{35}} \Rightarrow \boxed{P(A \cap B \cap C) = \frac{6}{35}}$

(iii) (a) Two red from 4 red marbles can be drawn in $^4C_2$ ways $= \frac{4!}{2!\,2!} = 6$

One Green from 3 green marbles can be drawn from $^3C_1$ ways $= \frac{3!}{1!\,2!} = 3$.

$\therefore$ Total possible way drawing 2 red and 1 green marbles $= 6 \times 3 = 18$

$\therefore$ Total number of events in this event is $= 18$

$\therefore$ Probability of this event $= \frac{18}{35}$

(i) (b) With replacement

If marbles are returned ~~each~~ each time then probability remains the same.

Then probability getting red in any trial

$$P(A) = P(B) = \frac{4}{7}$$

And probability getting in any trial $P(C) = \frac{3}{7}$

~~The probability of ge~~

⇒ The occurrence or non-occurrence of an event in the first trial does not affect the probability of occurrence of an event in the second trial, as so on.

$$\because P\{(RRG)\} = P(A \cap B \cap C) = P(C|B \cap A) \cdot P(B|A) \cdot P(A)$$

$$= \frac{3}{7} \times \frac{4}{7} \times \frac{4}{7}$$

$$= \frac{3}{7} \times \left(\frac{4}{7}\right)^2.$$

(ii) (b) using similar argument

$$P\{(RGR)\} = \cancel{\left(\frac{4}{7}\right) \times \frac{3}{7}} \; \frac{4}{7} \times \frac{3}{7} \times \frac{4}{7}$$

$$= \frac{3}{7} \times \left(\frac{4}{7}\right)^2$$

(iii) (c) using similar argument as in (i)(b)

RRG, RG,R, GRR

$$\left(\frac{4}{7}\right)^2 \times \frac{3}{7} \quad \left(\frac{4}{7}\right)^2 \times \frac{3}{7} \quad \left(\frac{4}{7}\right)^2 \times \left(\frac{3}{7}\right)$$

$$\therefore P\{(RRG) \cup (RGR) \cup (GRR)\} = \cancel{3 \times 4}$$

$$= 3 \times \left(\frac{4}{7}\right)^2 \times \frac{3}{7}$$

Ex :- ~~let in~~ 1 Suppose in previous example three marbles are taken togather at random. What is probability of getting exactly 2 red and one green marbles?

Solution :- This is a matter of selecting sub set of three elements. The total number of samble points possible

$$^7C_3 = \frac{7!}{3!(7-3)!} = \frac{7!}{3!\,4!}$$

$$^7C_3 = \frac{7\times6\times5\times\cancel{4!}}{3!\,\cancel{4!}}$$

$$^7C_3 = \frac{7\times\cancel{6}\times5}{\cancel{3\times2}} = 35$$

The total number of sample points favorable to the event is [ RRG, RGR, GRR ]

$$^4C_2 \times {}^3C_1 = \frac{4\times3}{2\times1} \times \frac{3}{3}$$

$$= 6\times3$$

$$= 18$$

- Because red ~~fruits~~ marbles can come from the set of red marbles and there are 4 red and a subset of 2 is taken, which can be done in $^4C_2$ ways. - Similarly one green marble can be selected $^3C_1$ ways.

- For each selection of red, the green is selected in ~~$^3C_1$~~ $^3C_1$ ways. and vice-versa.

- Thus total number of sample points favorable to the event is the product of two combinations.

Hence probability (iii) (a)

$$\frac{{}^4C_2 \times {}^3C_1}{{}^7C_3} = \frac{18}{35}$$

With replacement :- Probability of getting red at any trial $\frac{4}{7}$, probability of getting green marbal is $\frac{3}{7}$ at any ~~trail~~.

Total number of ways selecting three marbles is ${}^7C_3$.

Hence answer to (iii)(b) ~~be~~ is :

$$\cancel{\binom{7}{3}} \quad \cancel{{}^7C_3 \times \left(\frac{4}{7}\right)^2 \times \left(\frac{3}{7}\right)}$$

$$= \cancel{35 \left(\frac{4}{7}\right)^2 \times \left(\frac{3}{7}\right)}$$

$$= \frac{\left(\frac{4}{7}\right)^2 \times \left(\frac{3}{7}\right)}{35} = \boxed{\frac{\left(\frac{4}{7}\right)^2 \times \left(\frac{3}{7}\right)}{{}^7C_3}}$$

## Statistical Independence or Product Probability Property (PPP) :-

Let $(\Omega, \mathcal{A}, P)$ be probability space. Let A and B are events. If

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

then events A and B are said to be independent events or said to satisfy the product probability property.

equivalently The event A is said to be independent of the event B if and only if

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

$$\Rightarrow \frac{P(A\cap B)}{P(A)} = P(B)$$

equivalently $P(B|A) = P(B)$ $\Rightarrow P(B|A) = P(B)$ $\begin{bmatrix} P(A) \neq 0 \\ P(B) \neq 0 \end{bmatrix}$

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

Thus independence is symmetric property and holds mutually.

Note :- ① $\Omega$ and $\phi$ are independent of any event A.

$$P(A\cap\Omega) = P(A) = P(A)\cdot 1 = P(A)\cdot P(\Omega)$$

$$P(A\cap\phi) = P(\phi) = 0 = P(A)\cdot 0 = P(A)\, P(\phi)$$

② We define three events A, B, and C to be an independent ~~event~~ collection if they satisfy four conditions :-

$$\boxed{P(A\cap B) = P(A)\cdot P(B),\ P(A\cap C) = P(A)\cdot P(C),\ P(B\cap C) = P(B)\cdot P(C) \text{ and } P(A\cap B\cap C) = P(A)\cdot P(B)\cdot P(C)}$$

then the events A, B, and C are called mutually independent

- If first three condition satisfied then events ~~is~~ are called pairwise independent events.

In general A set of events $A_1, A_2, \ldots, A_k$ are said to be mutually independent events if for all subsets of the set $\{A_1, A_2, A_3, \ldots, A_k\}$ the product probability ~~two holds~~ property holds or the probability of intersection is the product of the probability of individual events, that is:

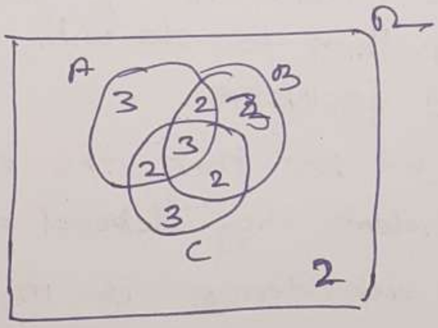
$$\boxed{P(A_{i_1} \cap A_{i_2} \cap \cdots A_{i_r}) = P(A_{i_1}) \cdot P(A_{i_2}) \cap \cdots P(A_{i_r})}$$

for all different subscripts $(i_1, i_2, \ldots, i_r)$ ; $r = 2, 3, \ldots k$

For every intersection of two, three, ..., k distinct events the probability of the intersection is the product of the individual probabilities.

Note: ① Pair wise independence need not imply the mutual independence.

e.g.




- Sample space Ω with symmetry in the outcomes and with 20 sample points, three events A, B, C, is given. The numbers in the various regions indicate the number of points falling in each region. Each of the sample points has a probability of $\frac{1}{20}$ each.

Total of 10 sample points fall in the each A, B, and C. , in $A \cap B$, $A \cap C$ and $B \cap C$ have ~~three~~ five sample point, three sample point fall in $A \cap B \cap C$, and two sample point in the complementary region of $A \cup B \cup C$. Therefore,

$$P(A) = \frac{10}{20} = \frac{1}{2} = P(B) = P(C);$$

$$P(A \cap B) = \frac{5}{20} = \frac{1}{4} = P(A) \cdot P(B)$$

$$P(A \cap C) = \frac{5}{20} = \frac{1}{4} = P(A) \cdot P(C)$$

$$P(B \cap C) = \frac{5}{20} = \frac{1}{4} = P(B) \cdot P(C)$$

$$P(A \cap B \cap C) = \frac{3}{20} \neq P(A) \cdot P(B) \cdot P(C) = \frac{1}{8}$$

$$\therefore \boxed{P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C)}$$

Thus A, B, C are pairwise independent, But not mutually independent

Note (2) :- Independent events should not confuse with mutually exclusive events.

- The phrase "independent" creates wrong impression in the minds as if events ~~are independent~~ have nothing to do with each other or they are mutually exclusive.
- When we say that the events A and B are ~~indep~~ independent they depend on each other a lot, ~~the~~ dependency is in the form of a product probability ~~of intersection~~ or the probability of intersection is the product of the individual probabilities, that is:

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

— Thus observe that A and B depend on each other through this product probability property (PPP). Hence Prof A.M. MATHAI suggested to replace ~~"Indep~~ "Independence of events" with event satisfying product probability property

Note (3):— "Independence ~~event~~ of events has nothing to do with mutually exclusiveness of events."

— Two events can be mutually exclusive and not ~~indep~~ independent

— Two event can be mutually exclusive and independent

— Two event ~~is~~ not mutually exclusive and independent

— Two event not mutually exclusive and not independent

## Origin of word Independent Event:—

This has originated with the conditional statements.

From defn of conditional Probability of A given ~~as~~ B as:

$$P(A|B) = \frac{P(A\cap B)}{P(B)}, \text{ for } P(B) \neq 0$$

Now if PPP hold then $P(A\cap B) = P(A)\cdot P(B)$. Then in this case

$$\boxed{P(A|B) = \frac{P(A)\cdot P(B)}{P(B)} = P(A) \text{ when } P(B) \neq 0}$$

↳ This means that conditional probability of A given B is the same as the marginal or ~~uncondition~~ unconditional probability of A when PPP holds.

↳ This means the probability of an event A is not affected by the occurrence or non-occurrence of B, in this sense word "independent" came in.

Ex :- A red die and ~~the~~ blue die are rolled together. What is the probability that we obtain 4 on the red die and 2 on blue die.

Solution :- R : " 4 on the red die."

B : " 2 on blue die.".

Thus we need $P(R \cap B)$ ?

Since outcome of one die does not affect the outcome of the other die., the events R and B are independent.

i.e. $P(R|B) = P(R)$

and $P(B|R) = P(B)$

$\therefore$ prob of the event we have

$$P(R \cap B) = P(R) \cdot P(B) \quad \text{---} (1)$$

Since symmetry of the experiment, therefore

$P(R) = \frac{1}{6}$, $P(B) = \frac{1}{6}$

$$\therefore \boxed{P(R \cap B) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}}$$

Ex :- Two coin are tossed. Let A denote the event " at most one head on the two tosses." and let B denote the event " one head and one tail in both tosses." Are A and B independent events ?

Solution :- The sample space of the experiment is $\Omega = \{HH, HT, TH, TT\}$. Now, the two events are defined as follows: $A = \{HT, TH, TT\}$ and $B = \{HT, TH\}$, also $A \cap B = \{HT, TH\}$

Thus,

$$P(A) = \frac{3}{4}$$

$$P(B) = \frac{2}{4} = \frac{1}{2}$$

$$P(A \cap B) = \frac{2}{4} = \frac{1}{2}$$

$$P(A) \cdot P(B) = \frac{3}{4} \times \frac{1}{2} = \frac{3}{8}$$

$$P(A) \cdot P(B) = \frac{3}{8}$$

Since $\boxed{P(A \cap B) \neq P(A) \cdot P(B)}$, we conclude that the events A and B are not independent.

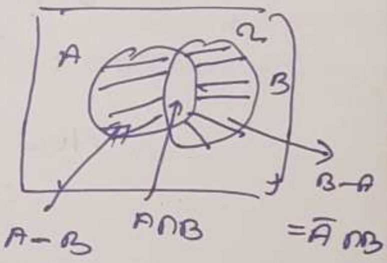
~~Moreover A⊆B~~ Moreover $B \subseteq A$, which confirms that they cannot be independent.

**Result :-** If A and B are independent events, then so are events A and $\bar{B}$, events $\bar{A}$ and B, and events $\bar{A}$ and $\bar{B}$

**Proof :-** Event A can be written as

$$A = (A \cap B) \cup (A - B)$$
$$= (A \cap B) \cup (A \cap \bar{B})$$

since $(A \cap B) \cap (A \cap \bar{B}) = \phi$




$\therefore$ Event $A \cap B$ and $A \cap \bar{B}$ are mutually exclusive. (from venn diagram) $= A \cap \bar{B}$

$\therefore P(A) = P(A \cap B) + P(A \cap \bar{B})$ [From the ~~law~~ of Axiom (ii)]

$= P(A) \cdot P(B) + P(A \cap \bar{B})$ [ Since A and B are independent $P(A \cap B) = P(A) \cdot P(B)$]

Thus we have

$$P(A \cap \bar{B}) = P(A) - P(A) \cdot P(B)$$
$$= P(A)[1 - P(B)]$$

① — $\boxed{P(A \cap \bar{B}) = P(A)\, P(\bar{B})}$ Thus from PPP ~~law~~ ~~Gre~~ rule for an event A and $\bar{B}$ ~~are~~ is satisfied

$\therefore$ $\boxed{A \text{ and } \bar{B} \text{ are independent}}$

- To show $\bar{A}$ and B are independent.

Now from Venn diagram we have:

$$B = (A \cap B) \cup (B - A) = (A \cap B) \cup (\bar{A} \cap B)$$

Since $(A \cap B) \cap (\bar{A} \cap B) = \phi$, therefore ~~A∩B~~ events $A \cap B$ and $\bar{A} \cap B$ are mutually exclusive.

$$\therefore \quad P(B) = P(A \cap B) + P(\bar{A} \cap B) \quad [\text{From Axiom (iii)}$$

$$= P(A) \cdot P(B) + P(\bar{A} \cap B) \quad [\text{since } A \text{ and } B \text{ are independent } \therefore P(A \cap B) = P(A) \cdot P(B)$$

$$P(B) = P(A) P(B) + P(\bar{A} \cap B)$$

$$\therefore \quad P(\bar{A} \cap B) = P(B) - P(A) \cdot P(B)$$

$$= P(B)[1 - P(A)]$$

$$\boxed{P(\bar{A} \cap B) = P(\bar{A}) \cdot P(B)} \quad \text{--- (2)}$$

Thus from PPP rule for an event $\bar{A}$ and B are independent.

- To show events $\bar{A}$ and $\bar{B}$ are independent.

Start with $\quad \bar{A} = (\bar{A} \cap B) \cup (\bar{A} \cap \bar{B})$

Now ~~P(Ā)~~ = since $\bar{A} \cap B$ and $\bar{A} \cap \bar{B}$ are independent

$$\therefore \quad P(\bar{A}) = P(\bar{A} \cap B) + P(\bar{A} \cap \bar{B})$$

$$= P(\bar{A}) \cdot P(B) + P(\bar{A} \cap \bar{B}) \quad [\text{since } \bar{A} \text{ and } B \text{ are independent}]$$

$$P(\bar{A} \cap \bar{B}) = P(\bar{A}) - P(\bar{A}) \cdot P(B)$$

$$= P(\bar{A})[1 - P(B)]$$

$$\boxed{P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B})} \quad \text{--- (3)}$$

Thus from PPP rule we have event $\bar{A}$ and $\bar{B}$ are independent.

Ex :— A and B are two ~~mutually~~ independent events defined in the same sample space. They have the following probabilities: $P(A) = x$ and $P(B) = y$. Find the probabilities of the following events in terms of $x$ and $y$.

a) Neither event A nor event B occurs.

b) Event A occurs but event B does not occur.

c) Either event A occurs or event B does not occur.

Solution :— Since event A and B are independent $\Rightarrow$

Event $\bar{A}$ and B independent, Event A and $\bar{B}$ are independent, and event $\bar{A}$ and $\bar{B}$ are also independent.

(a) Event A does not occur $\Rightarrow$ $\bar{A}$ occur

Event B does not occur $\Rightarrow$ $\bar{B}$ occur

Probability that ~~event~~ neither event A occur nor event B occurs, is the probability that event A does not occur and event B does not occur, which is given by

$$P_{\bar{A}\bar{B}} = P(\bar{A} \cap \bar{B}) = P(\bar{A}) \cdot P(\bar{B})$$

[Since A and B are independent $\Rightarrow$ $\bar{A}$ and $\bar{B}$ are independent]

$$P_{\bar{A}\bar{B}} = P(\bar{A}) \cdot P(\bar{B})$$

$$= (1-P(A))(1-P(B))$$

$$\boxed{P_{\bar{A}\bar{B}} = (1-x)(1-y)}$$

(b) Event A occurs but event B does not occur is the probability that event A occurs and event B does not occur, which is given by:

$$P_{A\bar{B}} = P(A \cap \bar{B}) = P(A)P(\bar{B})$$

[Since independence of A and B imply independence of A and $\bar{B}$]

$$P_{A\bar{B}} = P(A) \cdot [1 - P(B)]$$

$$\boxed{P_{A\bar{B}} = x(1-y)}$$

(e) The probability that either event A occurs or event B does not occur is given by

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

$$= P(A) + P(\bar{B}) - P(A) \cdot P(\bar{B})$$

[Since if A and B are independent event then A and $\bar{B}$ are also independent]

$$\therefore \cancel{P(A \cup \bar{B}) = P(A) + P(\bar{B}) \cdot (1 - P(A))}$$

$$P(A \cup \bar{B}) = x + (1-y) - x \cdot (1-y)$$

$$= \cancel{x} + 1 - y - \cancel{x} + xy$$

$$P(A \cup \bar{B}) = 1 - y + xy$$

$$\boxed{P(A \cup \bar{B}) = 1 - y(1-x)}$$

Ex :- Jim and Bill are marksmen. Jim can hit a target with a probability of 0.8 while Bill can hit target with probability of 0.7. If both fire at a target at the same time, what is the probability that the target is hit at least once ?

Solution :-

J : Event that Jim hit the target

$\bar{J}$ : Event that Jim does not hit the target

B : Event that Bill hits a target

$\bar{B}$ : Event that Bill does not hit a target

- Since outcome of Bill's ~~hit~~ shot does not affect the outcome of Jim shots vice-versa. $\Rightarrow$ The events J and B are independent

Since J and B are independen $\Rightarrow$ J and $\bar{B}$, $\bar{J}$ and B, and $\bar{J}$ and $\bar{B}$ are independent

Thus the probability that the target is hit atleast once is the probability of the union of its being hit once and at being hit twice

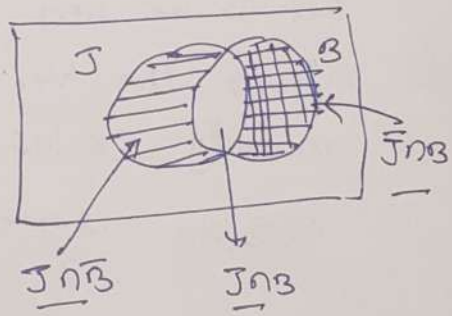
Event
$$E = \underbrace{\{J \cap B\}}_{\text{Both hit}} \cup \underbrace{\{J \cap \bar{B}\}}_{\text{J hit B not hit}} \cup \underbrace{\{\bar{J} \cap B\}}_{\text{J not B hit hit}}$$

$$p = P(E) = P(J \cap B) + P(J \cap \bar{B}) + P(\bar{J} \cap B)$$
$$= P(J)\cdot P(B) + P(J)\cdot P(\bar{B}) + P(\bar{J})\cdot P(B)$$
$$= 0.8 \times 0.7 + 0.8 \times 0.3 + 0.2 \times 0.7$$

$$\therefore p = 0.56 + 0.24 + 0.14$$
$$\boxed{p = 0.94}$$

~~Note that this probability is complement of the probability that neither of them hits the target, i.e.~~ ~~Probability that neither of them hits the target~~ ~~$P(\bar{J} \cap \bar{B})$~~ ? ⇒ complement of the above

~~$\therefore p = 1 - P(\bar{J} \cap \bar{B})$~~
~~$= 1 - P(\bar{J}) \cdot P(\bar{B})$~~
~~$= 1 - (0.2)(0.3)$~~
~~$= 1 - 0.06$~~
~~$\boxed{p = 0.94}$~~




~~Event $E = J \cup B = (J \cap \bar{B}) \cup ($~~

Event $E = J \cup B = (J \cap B) \cup (J \cap \bar{B}) \cup (\bar{J} \cap B)$;
$J \cup B$, $J \cap \bar{B}$, $\bar{J} \cap B$ are mutually exclusive

Event $\bar{E}$ = complement of E = neither of them hit the target
$= \overline{(J \cup B)} = \bar{J} \cap \bar{B}$

$\therefore P(\bar{E}) = P(\bar{J} \cap \bar{B}) = P(\bar{J}) \cdot P(\bar{B})$ [since $\bar{J}$ and $\bar{B}$ are independent given that J and B are independent
$= 0.2 \times 0.3 = 0.06$

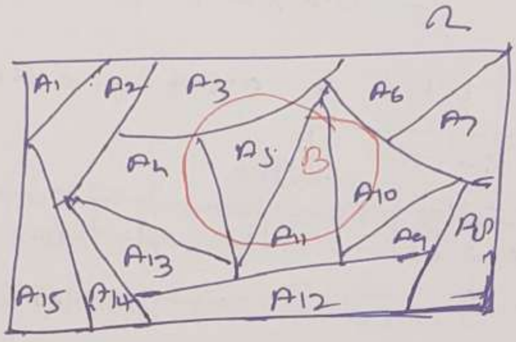
Now $P(\bar{E}) = 1 - P(E)$
$= 1 - 0.94$
$= 0.06$

## Total Probability :

Two important results of conditional probability namely: ① Total Probability
② Baye's Theorem.

Both deals with the partitioning of the sample space.

- Let sample space, $\Omega$ be partitioned into mutually ~~each~~ exclusive and totally exhaustive events $A_1, A_2, A_3, \ldots, A_k$. ~~and let B be any other event~~




Let B be any ~~sample~~ event in the sample space & clearly from the Venn diagram event B is partitioned into mutually exclusive event/pieces: $B\cap A_1, B\cap A_2, B\cap A_3, \ldots, B\cap A_k$ some of them may be empty. Thus

$$\Omega = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_k ;$$

$$A_i \cap A_j = \phi, \text{ for all } i \neq j = 1, 2, 3, \ldots, k$$

$$B = (B\cap A_1) \cup (B\cap A_2) \cup (B\cap A_3) \cup \cdots \cup (B\cap A_k),$$

$$(B\cap A_i) \cap (B\cap A_j) = B\cap A_i \cap A_j = B\cap\phi ,$$

since for $i \neq j = 1, 2, \ldots k$, $A_i \cap A_j = \phi$

$$= \phi$$

$\therefore$ $B\cap A_i$ and $B\cap A_j$ ~~is mutually excl~~ is exclussive event and collective exhaustive w.r.t B.

$\therefore P(B) =$

Hence by third Axiom of probability

$$P(B) = P(B\cap A_1) + P(B\cap A_2) + P(B\cap A_3) + \cdots + P(B\cap A_k) \quad \text{--- (1)}$$

Now using conditional probability we have

$$P(B) = P(B|A_1)\cdot P(A_1) + P(B|A_2)\cdot P(A_2) + P(B|A_3)\cdot P(A_3) + \cdots + P(B|A_k)\cdot P(A_k) \quad \text{--- (2)}$$

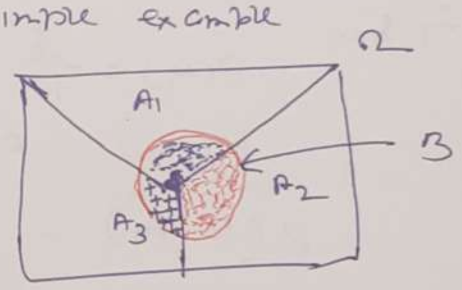
$$P(B) = \sum_{i=1}^{k} P(B|A_i)\,P(A_i)$$

$P(A_i) \neq 0$, for $i = 1, 2, 3, \ldots, k$.

The equation (2) is known as law of Total Probability of event B.

— Clearly Total probability law connect the conditional probability with marginal probability, where marginal probabilities are non-zeros.

Ex[ Start Explanation with simple example



Example 2.7:— Dr. Joy is not a very good ~~med~~ medical practitioner. If a patient goes to him the chance that he will diagnose the patient's symptoms properly 30%. Even if diagnosis is correct his treatment is such that the chance of the patient dying is 60% and if the diagnosis is wrong the chance of patient dying is 95%. What is the probability that a patient going to Dr. Joy dies during the treatment?

Solution:—

$A_1$: Event of a correct diagnosis.

$A_2$: Event of wrong diagnosis.

Then $A_1 \cap A_2 = \phi$, and $\Omega = A_1 \cup A_2$

B: Event of a patient of Dr. Joy dying

Then

$$P(A_1) = \frac{3}{10} = 0.3.$$

$$P(A_2) = 0.7$$

$$P(B|A_1) = 0.6$$

$$P(B|A_2) = 0.95$$

$P(B) = ?$

By using Total Probability law we have.

$$P(B) = P(B|A_1)P(A_1) + P(B|A_2)P(A_2)$$

$$= (0.6)(0.3) + (0.95)(0.7)$$

$$= 0.18 + 0.665$$

$$P(B) = 0.845$$

or the chance of dying the patient is 84.5%

Ex: 2.8 Mr. Narayanan is civil engineer with Kerala Government. He asked to design an over bridge (sky way). The chance that his design is going to be faulty is 60% and the chance that his design will be correct is 40%. The chance of the over bridge collapsing if the design is faulty is 90% otherwise, due to other causes, the chance of the over bridge collapsing is 20%. What is the chance that an over bridge built by Mr. Narayanan will collapse?

Solution:

$A_1$: Event that the design is falty

$A_2$: Event that the design is not falty.

Then $A_1 \cap A_2 = \phi$ and $A_1 \cup A_2 = \Omega$ (Sure event)

$B$: Event that overbridge is collapsing.

Then we are given following:

$P(A_1) = 0.6$, $P(A_2) = 0.4$, $P(B|A_1) = 0.9$,

$P(B|A_2) = 0.2$

Now compute $P(B)$? Now from the law of total probability

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2)$$

$$= 0.9 \times 0.6 + 0.2 \times 0.4$$

$$\therefore P(B) = 0.54 + 0.08$$

$$\boxed{P(B) = 0.62}$$

There is a 62% chance of the over bridge designed by Mr. Narayanan collapsing.

Ex:- A student buys 1000 chips ~~&~~ from supplier A, 2000 chips from supplier B, and 3000 chips from suplier C. He tested chips and found that the probability that the chip is defective depends on the supplier from where ~~he~~ it was bought. Specifically, given that a chip come from suplier A, the probability that it is defective is ~~0.10~~ 0.05, given that a chip came from supplier B, the probability that it is defective 0.10, and given that a chip come from supplier C, the probability that it is defective is 0.10.

If the chips from the three suppliers ~~mix~~ are mixed togather and one of them is selected at random, what is probability that it is defective?

Solun:- P(A) : Probability that randomaly selected chip come from supplier A.

P(B) : ——————— ———————
——————— supplier B

P(C) : ———————
——————— supplier C.

D: Selected chip is defective. selected randomaly

P(D|A) : chip defective come from supplier A

P(D|B) : ——————— ——————— B

P(D|C) : ——————— C

$P(D|A) = 0.05$

$P(D|B) = 0.10$

$P(D|C) = 0.10$

$$P(A) = \frac{1000}{1000+2000+3000} = \frac{1}{6}$$

$$P(B) = \frac{2000}{1000+2000+3000} = \frac{1}{3}$$

$$P(C) = \frac{3000}{1000+2000+3000} = \frac{1}{2}$$

$$\therefore P(D) = P(D|A) \cdot P(A) + P(D|B) \cdot P(B) + P(D|C) \cdot P(C)$$

$$= (0.05)\left(\frac{1}{6}\right) + (0.10)\left(\frac{1}{3}\right) + (0.10)\left(\frac{1}{2}\right)$$

$$\boxed{P(D) = 0.09167}$$

**Ex:** In previous example, given that a randomly selected chip is defective, what is the probability that ~~it~~ it came from supplier A.

**Soln:** Bayes formula

Cond. notation is used as in previous:

$\therefore$ Conditional event. $A|D$

$$\therefore P(A|B) = \frac{P(D \cap A)}{P(D)} = \frac{P(D|A) \cdot P(A)}{P(D)} = \frac{P(D|A) \cdot P(A)}{\cancel{P(A|D)P(D)+} \; P(D|A) \cdot P(A) + P(D|B) \cdot P(B) + P(D|C) \cdot P(C)}$$

$$= \frac{0.05 \times \frac{1}{6}}{(0.05)\left(\frac{1}{5}\right) + (0.10)\left(\frac{1}{3}\right) + (0.10)\left(\frac{1}{2}\right)}$$

$$\boxed{P(A|D) = 0.0909}$$

# Baye's Rule:-

- Let sample space $\Omega$ is partitioned into mutually exclusive and ~~collecti~~ totally exhaustive events $A_1, A_2, \ldots, A_k$ and let $B$ be any event in the sample space.

- Then using Total probability law we get

$$P(B) = \sum_{i=1}^{k} P(B|A_i) \cdot P(A_i) \; ; \quad P(A_i) \neq 0, \; i=1,2,\ldots k$$

- Now look the intersection of $B$ with $A_j$'s. For example $B \cap A_1$, now from definition of conditional probability

$$P(B \cap A_i) = P(A_i|B)\, P(B) \; ; \quad P(B) \neq 0$$

Therefore,

$$P(A_1|B) = \frac{P(B \cap A_1)}{P(B)} \, , \quad P(B) \neq 0$$

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B)}$$

$$\therefore P(A_1|B) = \frac{P(B|A_1)\, P(A_1)}{P(B|A_1) P(A_1) + P(B|A_2) \cdot P(A_2) + \cdots + P(B|A_k)\, P(A_k)}$$

$$\text{so} \quad \boxed{P(A_1|B) = \frac{P(B|A_1)\, P(A_1)}{\sum_{i=1}^{k} P(B|A_i) \cdot P(A_i)}} \; ; \quad P(A_i) \neq 0, \; i=1,2,3,\ldots,k$$

This equation is known as Baye's rule ~~or~~ or Baye's law or Bayes theorem. It is named after ~~th~~ christian priest Rev Bayes, who ~~dt~~ discovered this rule.

Note -

(1) $P(A_j|B)$ : posterior probabilities
$\equiv$ probability that event $A_j$ computed after observing the event $B$.

$P(A_j)$ : called prior probability of $A_j$

$\equiv$ Probability of $A_j$ computed before ~~obs~~ observing the event $B$.

* So Bayes rule connect the posterior probability with the prior probability.

(2) Baye's rules also provide inverse reasoning or establishes a connection between inverse probability of the type $P(A_1|B)$ and $P(B|A_1)$, where ~~the~~ one can interprected as the probability from cause to effect and other from effect to ~~cause~~ cause.

(3) This is used for interpreting following type of phenomenon:

- If patient dies and if ~~reati~~ relatives of the patient feel that the medical doctor attending to the patient was incompetent or the hospital was negligent then they would ~~have~~ like to have an estimate of the chance that the patient died due to negligence, or incompetence of doctor etc.

- What is ~~the~~ probability that the diagnosis was wrong given that the patient died?

- In the case of bridge collapsing the concerned general public may ~~be~~ want to know the chance that the engineer's design was in fact faulty in the light of bridge collapsing.

Ex Based on 2.7 In Example 2.7 what is the probability that Dr. Joy's diagnosis was wrong in the light of a patient of Dr. Joy dying?

Solun:- Here we are asked to compute the probability $P(A_2|B)$.

But
$$P(A_2|B) = \frac{P(B|A_2) \cdot P(A_2)}{\sum_{i=1}^{2} P(B|A_i)\, P(A_i)}$$

$$= \frac{P(B|A_2) \cdot P(A_2)}{P(B|A_1)\, P(A_1) + P(B|A_2) \cdot P(A_2)}$$

$$P(A_2|B) = \frac{(0.95) \times (0.7)}{0.845}$$

$$= \frac{0.665}{0.845}$$

$$P(A_2|B) = \frac{133}{169} \simeq 0.787$$

Ex Based on 2.8:- What is probability that the design of Mr. Narayanan was faulty, in the light of an over bridge designed by him collapsing.

Solun:- Here we asked to compute the probability $P(A_1|B)$. From Baye's rule we have:

$$P(A_1|B) = \frac{P(B|A_1) \cdot P(A_1)}{P(B)}$$

$$= \frac{P(B|A_1) \cdot P(A_1)}{P(B|A_1)\, P(A_1) + P(B|A_2) \cdot P(A_2)}$$

$$= \frac{(0.9)(0.6)}{(0.62)} = \frac{54}{62} \simeq 0.87$$

There is approximately 97% chance that Mr. Narayanan's design was faulty.

[ But there is no way of taking any action against Mr. Narayanan due to job security in the ~~b.p. of~~ Government setup even ~~though~~ if several pepole died due to collapse of the over bridge.]

Entropy :— Another concept associated with partitioning of a sample space or a system

$$(A_1, p_1), (A_2, p_2), \ldots, (A_k, p_k)$$

where $A_1, A_2, \ldots A_k$ are mutually exclusive and totally exhastive events (A partitioning of sample space) and $p_1, p_2, \ldots, p_k$ the associated probabilities, that is,

$P(A_i) = p_i$, $i = 1, 2, 3, \ldots, k$ such that $p_j \geq 0$, $p_1 + p_2 + p_3 + \cdots + p_k = 1$, is the concept of entropy or information or uncertainty. This can be explained with a simple for $k = 2$.

– Suppose Mr. ~~Ajit~~ Nimbus is contesting an election to be the chairman of the local township. Suppose that the only two possibilities are that either he wins or he does not win. Thus we have two events $A_1, A_2$ such that $A_1 \cap A_2 = \phi$, $A_1 \cup A_2 = \Omega$, where $\Omega$ is sure event.

* Three local news papers predicting his chaces of ~~winin~~ wining.
    - First news paper gave 50 – 50 chance of his wining.

- The second gave 80-20 chance and third gave
- 60-40 chance. That is, if $A_1$ is the event of ~~wining~~ winning and $p = P(A_1)$ the true probability of ~~wining~~ winning, then the three estimates for ~~this~~ $p$ are 0.5, 0.8, 0.6 respectively. We have three schemes here:

Scheme 1: $\begin{pmatrix} A_1 & A_2 \\ 0.5 & 0.5 \end{pmatrix}$

Scheme 2: $\begin{pmatrix} A_1 & A_2 \\ 0.8 & 0.2 \end{pmatrix}$

Scheme 3: $\begin{pmatrix} A_1 & A_2 \\ 0.6 & 0.4 \end{pmatrix}$

- In Scheme 1 there is quite a lot of uncertainity about the win, because it is 50-50 situation with maximum uncertainity, 50% chance of winning.
- In Scheme 2 the uncertainity much less because it is a 80-20 situation.
- Whatever be the "uncertainity" one can say this much that in ~~section 3~~ scheme 3 the uncertainity in inbetween the situations in ~~section~~ scheme 1 and scheme 2.

* Lack of "uncertainity" is the information content in ~~the~~ a scheme.

Is any mathematical measure for this "information" content in a scheme?

Answer! Yes, it has a lot application in practical situations such as sending a wireless message from one point and it is captured at another point.

→ One would like to make sure that the message is fully captured in every aspect or ~~atleast~~ at least the information content is maximum.

e.g. of photo is transmitted we would like to capture it in all its full details.

C.E. Shannon in 1948 came up with a measure of "uncertainity" or "information" in scheme. He developed it for communication networks. The measure is:

$$\boxed{S = -C \sum_{i=1}^{k} p_i \ln p_i}$$

where $C$ is constant and ln is natural logarithm

- He developed it by putting forward desirable properties as axioms or postulates or assumptions and then deriving the expression mathematically

  ↳ A whole discipline is developed and it is known as "Information Theory", with wide range of applications in almost all fields.

- The measure is simply called "entropy" in order to avoid possible misinterpretation if the term "information" or "uncertainity" is used.

* A mathematical Theory of communication:

By C.E. Shannon, The Bell System Technical Journal Vol. 27, pp. 379-423, 623-656, July, october 1948

# Some More Problem on Baye's Rule :

**Ex:-** In chip buy problem, given that randomally selected chip is defective, what is the probability that it came from supplier A?

**Solution:-** Using same notation as in previous example, the probability that the randomly selected chip came from supplier A, given that it is defective, is given by:

$$P(A|D) = \frac{P(D \cap A)}{P(D)} = \frac{P(D|A)\cdot P(A)}{P(D|A)P(A) + P(D|B)\cdot P(B) + P(D|C)\cdot P(C)}$$

$$P(A|D) = \frac{(0.05)\cdot(\frac{1}{6})}{(0.05)(\frac{1}{6}) + (0.10)(\frac{1}{3}) + (0.10)(\frac{1}{2})}$$

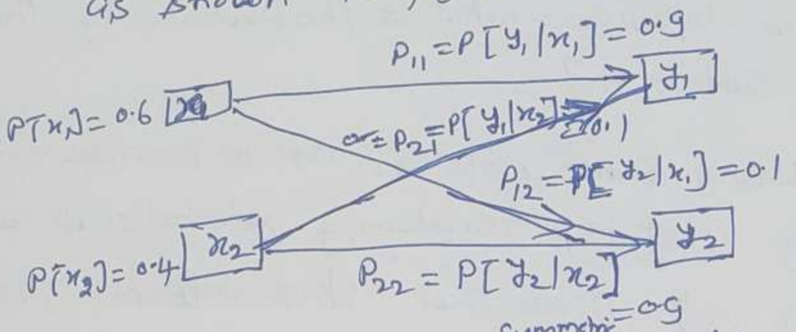
$$\boxed{P(A|D) = 0.0909}$$

**Ex —(The Binary Symmetric channel) :—** A discrete channel is characterised by an input alphabet $X = \{x_1, x_2, x_3, \ldots, x_n\}$ an output alphabet $Y = \{y_1, y_2, \ldots, y_n\}$, and a set of conditional probabilities (called transition ~~probability~~ probabilities), $P_{ij}$, which are defined as follows:

$$P_{ij} = P[y_j | x_i] = P[\text{receiving symbol } y_j \text{, given that symbol } x_i \text{ was transmitted}]$$

$$i = 1, 2 \ldots n \; ; \; j = 1, 2, 3, \ldots, m$$

The binary channel is a special case of the discrete channel where $n = m = 2$. It can be represented as shown in figure below:




(Binary symmetric Channel)

$$P = \begin{array}{c c} & \begin{array}{cc} y_1 & y_2 \end{array} \\ \begin{array}{c} x_1 \\ x_2 \end{array} & \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \end{array}$$

In binary channel, an error occurs if $y_2$ is received when $x_1$ is transmitted or $y_1$ is received when $x_2$ is transmitted. Thus, the probability of error, $P_e$, is given by:

$$P_e = P[x_1 \cap y_2] + P[x_2 \cap y_1]$$
$$= P[y_2|x_1] \cdot P[x_1] + P[y_1|x_2] \cdot P[x_2]$$
$$P_e = P[x_1]\, P_{12} + P[x_2] \cdot P_{21}$$

If $P_{12} = P_{21}$, we say that the channel is a binary symmetrical channel (BSC). Also, if in BSC $P[x_1] = p$, then $P[x_2] = 1 - p = q$

Consider the binary symmetric channel given in above figure: with $P[x_1] = 0.6$, $P[x_2] = 0.4$ Compute the following:

(a) The probability that $x_1$ was transmitted, given that $y_2$ was received.

(b) The probability that $x_2$ was transmitted, given that $y_1$ was received.

(c) The probability that $x_1$ was transmitted, given that $y_1$ was received.

(d) The probability that $x_2$ was transmitted, given that $y_2$ was received.

(e) The unconditional probability of error.

Solution :- $P[y_1]$ : Probability that $y_1$ was received

$P[y_2]$ : Probability that $y_2$ was received.

Then

(a) The probability that $x_1$ was transmitted, given that $y_2$ was received is given by :

$$P[x_1|y_2] = \frac{P[x_1 \cap y_2]}{P[y_2]} = \frac{P[y_2|x_1] \cdot P[x_1]}{P[y_2|x_1] P[x_1] + P[y_2|x_2] P[x_2]}$$

$$P[x_1|y_2] = \frac{(0.1)(0.6)}{(0.1)(0.6) + (0.9)(0.4)}$$

$$\boxed{P[x_1|y_2] = 0.143}$$

(b) The probability that $x_2$ was transmitted, given that $y_1$ was received is given by

$$P[x_2|y_1] = \frac{P[x_2 \cap y_1]}{P[y_1]} = \frac{P[y_1|x_2] P[x_2]}{P[y_1|x_2] P[x_2] + P[y_1|x_1] P[x_1]}$$

$$= \frac{(0.1)(0.4)}{(0.1)(0.4) + (0.9)(0.6)} = 0.069$$

(c) The probability that $x_1$ was transmitted, given that $y_1$ was received is given by:

$$P[x_1|y_1] = \frac{P[x_1 \cap y_1]}{P[y_1]}$$

$$= \frac{P[y_1|x_1]\,P[x_1]}{P[y_1|x_1]\,P[x_1] + P[y_1|x_2] \cdot P[x_2]}$$

$$= \frac{(0.9)(0.6)}{(0.9)(0.6) + (0.1)(0.4)}$$

$$= 0.931$$

$$= 1 - P[x_2|y_1]$$

(d) The probability that $x_2$ was transmitted, given that $y_2$ was received is given by:

$$P[x_2|y_2] = \frac{P[x_2 \cap y_2]}{P[y_2]}$$

$$P[x_2|y_2] = \frac{P[y_2|x_2]\,P[x_2]}{P[y_2|x_1]\,P[x_1] + P[y_2|x_2] \cdot P[x_2]}$$

$$= \frac{(0.9)(0.4)}{(0.1)(0.4) + (0.9)(0.4)}$$

$$= 0.857 = 1 - P[x_1|y_2]$$

(e) The unconditional probability of error is given by

$$P_e = P[x_1]\,P[y_2|x_1] + P[x_2]\,P[y_1|x_2]$$

$$= P[x_1]\,P_{12} + P[x_2]\,P_{21}$$

$$= (0.6)(0.1) + (0.4)(0.1)$$

$$\boxed{P_e = 0.1}$$

Ex :-

The quaterback for a certain football team has a good game with probability 0.6 and a bad game with ~~proba~~ probability 0.4. When he has a good game, he throws an interception with a probability of 0.2; and when he has a bad game, he throws an interception with probability of 0.5. Given that he threw an interception in a particular game, what is the probability he had a good game?

Solution :- Let G: denote the event that the ~~quarte~~ quarterback has a good game and B: the event that he has bad game. Similarly, let I denote the event that he throws an interception. Then we have that :

$$P[G] = 0.6 \; ; \; P[I/G] = 0.2$$

$$P[B] = 0.4 \; ; \; P[I/B] = 0.5$$

$$P[G|I] = \frac{P[G \cap I]}{P[I]}$$

According Baye's formula,

$$P[G|I] = \frac{P[I/G] \cdot P[G]}{P[I/G] \cdot P[G] + P[I|B] \cdot P[B]}$$

$$= \frac{(0.2)(0.6)}{(0.2)(0.6) + (0.5)(0.4)}$$

$$= \frac{0.12}{0.32}$$

$$\boxed{P[G|I] = 0.375}$$

Ex:- Two events A and B are such that $P[A \cap B] = 0.15$, $P[A \cup B] = 0.65$, and $P[A|B] = 0.5$. Find $P[B|A]$.

Solution:- $P[A \cup B] = P[A] + P[B] - P[A \cap B]$

$$\Rightarrow \quad 0.65 = P[A] + P[B] - 0.15$$

$$\therefore \quad P[A] + P[B] = 0.80$$

also $P(A \cap B) = P[A|B] \cdot P[B]$

$$\Rightarrow \quad P[B] = \frac{P[A \cap B]}{P[A|B]}$$

$$P[B] = \frac{0.15}{0.5} = 0.30$$

Thus $P[A] = 0.80 - P[B]$ ~~$= 1 - 0.30 =$~~

$$= 0.80 - 0.30$$

$$\Rightarrow \boxed{P[A] = 0.50}$$

Since $P[A \cap B] = P[B|A] \cdot P[A]$

we have that

$$P[B|A] = \frac{0.15}{0.50} = \underline{0.30}$$

Ex :-

A student went to the post office to send a priority mail to his parents. He gave the postal lady a bill he believed was \$20. However, the postal lady gave him change based on her belief that she received a \$10 bill from the student. The student started to dispute the change. Both the student and the postal lady are honest but may make mistakes. If the postal lady's drawer contains thirty \$20 bills and twenty \$10 bills, and the postal lady correctly identifies bills 90% of the time, what is the probability that the student's claim valid?

Solution:- Event

A : Student gave \$10 bill

B : Event that the student gave \$20. bill.

V : Event that the student claim is valid.

L : Event that the postal lady said that the student gave her a \$10 bill

Since there in draweer 30, \$20 bills and 20 \$10 bills

The probability that the student gave ~~\$10 bills~~ to postal lady \$20 bill is $= \frac{30}{30+20} = \frac{30}{50} = \frac{3}{5} = 0.6$

and the probability that it was \$10 bill was:

$$= \frac{20}{30+20} = \frac{20}{50} = \frac{2}{5} = 0.4$$

$$\therefore P[L] = P[L|A] \cdot P[A] + P[L/B] P[B]$$

$$= (0.90)(0.40) + (0.1)(0.6)$$

$$\boxed{P[L] = 0.42}$$

Therefore, the probability that the student's claim is valid is the probability that he gave a \$20 bill, given that the postal lady said that the student gave her \$10 bill. Using Baye's formula we ~~obta~~ ~~ob~~ obtain:

$$P[V|L] = \frac{P[V \cap L]}{P[L]} = \frac{P[L|V]\,P[V]}{P[L]}$$

$$= \frac{(0.10)(0.60)}{0.42}$$

$$P[V|L] = \frac{1}{7} = 0.1429$$

Ex:- An aircraft maintenance company bought an equipment for detecting structural defects in air-crafts. Tests indicate that 95% of the time the equipment detects defects when they actually exist and 1% of the time it gives a false alarm [that is, it indicates the presence of a structural defect when in fact there is none]. If 2% of the aircrafts actually have structural defects, what is the probability that an aircraft actually has a structural defect given that the equipment indicates that it has structural defect

Solution:- D: Event that an aircraft has a ~~stru~~ structural defect

B: Event that ~~there~~ the test indicates that there ~~is~~ is structural defect

We find $P[D|B] = ?$

We use Baye's formula:

$$P[D|B] = \frac{P[D\cap B]}{P[B]} = \frac{P[B|D]\cdot P[D]}{P[B|D]\cdot P[D] + P[B|\bar{D}]\,P[\bar{D}]}$$

$$= \frac{(0.95)(0.02)}{(0.95)(0.02) + (0.01)(0.98)}$$

$$= 0.660$$

Thus, only 66% of the aircrafts that the equipment diagnoses as having structural defects actually have structural defects.

# Tree Diagram :-

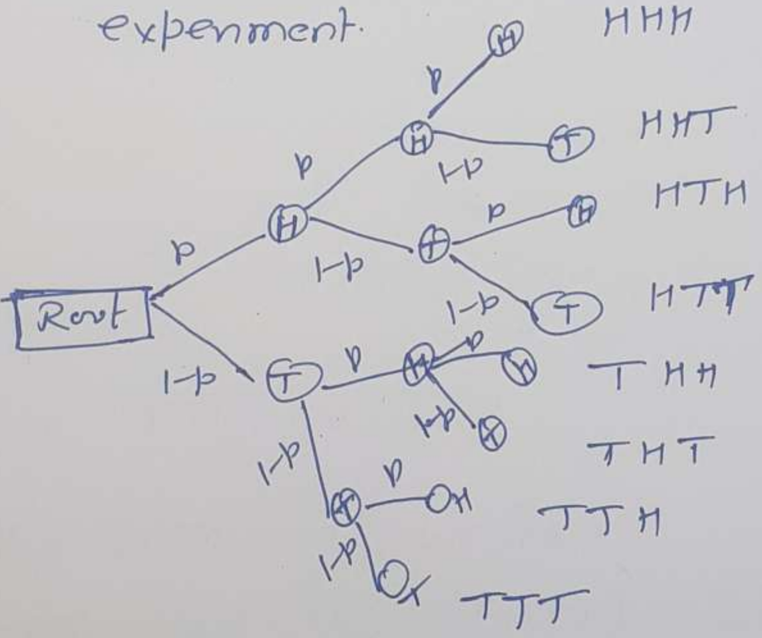
- Conditional probability are used to model experiments that take place in stages
- output of such experiments are conveniently ~~rep~~ represented by a tree diagram.
  - ↳ A tree is a connected graph that contains no circuit ~~(loop)~~ (or loop)
- Every two nodes in the tree have a unique path connecting them.
- Line segments called branches or edges ~~inter~~ interconnect the nodes.
- Each branch split into other branch or terminate

## Modelling Random Experiment with tree

- Node of the tree represent an events of ~~experiment~~ experiment
- Number of branches that emanate from a node represents the number of events that ~~occur~~ can occur, given that the event represented by ~~node~~ that node occurs.
- The node that has no preceedecessor is called root, and successor node is called child, and any node that has no successor node is called ~~child~~ leaf of the tree.
- Event of interest are usually defined at the leaves by tracing outcomes of the experiment from root to each leaf.

- The conditional probability appear on the branches leading from the node representing an event to the nodes representing the next events of the experiment
- A path through the tree corresponds to a possible outcome of the experiment
- Thus, the product of all the branch probabilities from the root of the tree to any node is equal to the probability of event represented by that node.

Ex:- Consider an experiment that consists of three tosses of a coin. Let $p$ denote the probability of heads in a toss, then $1-p$ is the probability of tails in a toss. Tree diagram of the experiment.




Let A : "First toss came up heads" = {HHH, HHT, HTH, HTT}

B : " Second toss came up tails."

= {HTH, HTT, TTH, TTT}

$P[A] = P[HHH] + P[HHT] + P[HTH] + P[HTT]$

$$P[A] = p^3 + p^2(1-p) + p^2(1-p) + p(1-p)^2$$
$$= p^3 + 2p^2(1-p) + p(1-p)^2$$
$$= p^3 + 2p^2 - 2p^3 + p(1+p^2 - 2p)$$
$$= \cancel{p^3} + \cancel{2p^2} - \cancel{2p^3} + p + \cancel{p^3} - \cancel{2p^2}$$
$$P[A] = p$$

$P[B] = P[HTH] + P[HTT] + P[TTH] + P[TTT]$

$$= \cancel{(1-p)^2 p}$$
$$= p^2(1-p) + p(1-p)^2 + p.(1-p)^2 + (1-p)^3$$
$$= p^2(1-p) + 2p(1-p)^2 + (1-p)^3$$
$$\cancel{= p(1-p)(p + 2 - 2p) + (1-p)^3}$$
$$\cancel{= p(1-p)(}$$
$$= p^2(1-p) + 2p(1-p)^2 + (1-p)^3$$
$$= p^2(1-p) + (1-p)^2(2p + 1 - p)$$
$$= p^2(1-p) + (1-p)^2(p+1)$$
$$= (1-p)[p^2 + (1-p)(1+p)]$$
$$= (1-p)[\cancel{p^2} + 1 - \cancel{p^2}]$$
$$P[B] = (1-p)$$

$A \cap B = \{HTH, HTT\}$ we have

$$P(A \cap B) = P[HTH] + P[HTT] = \cancel{(1-p)^2 p}$$
$$= p^2(1-p) + p(1-p)^2$$
$$= p(1-p)[\cancel{p} + 1 - \cancel{p}]$$
$$P(A \cap B) = p(1-p)$$

$\cancel{P(A \cup B)} =$

Now $A \cup B = \{HHH, HHT, HTH, HTT, TTH, TTT\}$

$$P[A \cup B] = P[HHH] + P[HHT] + P[HTH] + P[HTT] + P[TTH] + P[TTT]$$

$$P[A \cup B] = p^3 + 2p^2(1-p) + 2p(1-p)^2 + (1-p)^3$$
$$= p^2\{p + 2(1-p)\} + (1-p)^2\{2p + 1 - p\}$$
$$= p^2\{2 - p\} + (1-p)^2\{1 + p\}$$
$$= p^2\{1 + (1-p)\} + (1-p)^2(1+p)$$
$$= p^2 + p^2(1-p) + (1-p)^2(1+p)$$
$$= p^2 + (1-p)[p^2 + (1-p)(1+p)]$$
$$= p^2 + (1-p)[\cancel{p^2} + 1 - \cancel{p^2}]$$
$$= p^2 + 1 - p$$
$$= 1 - p + p^2$$

$$\boxed{P(A \cup B) = 1 - p(1-p)}$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= \cancel{p + (1-p) - 1 + p(1-p)}$$
$$= \cancel{p + (1-p)(1-p)}$$
$$= \cancel{p +} 1 -$$
$$= \cancel{p} + 1 - \cancel{p} - p(1-p)$$

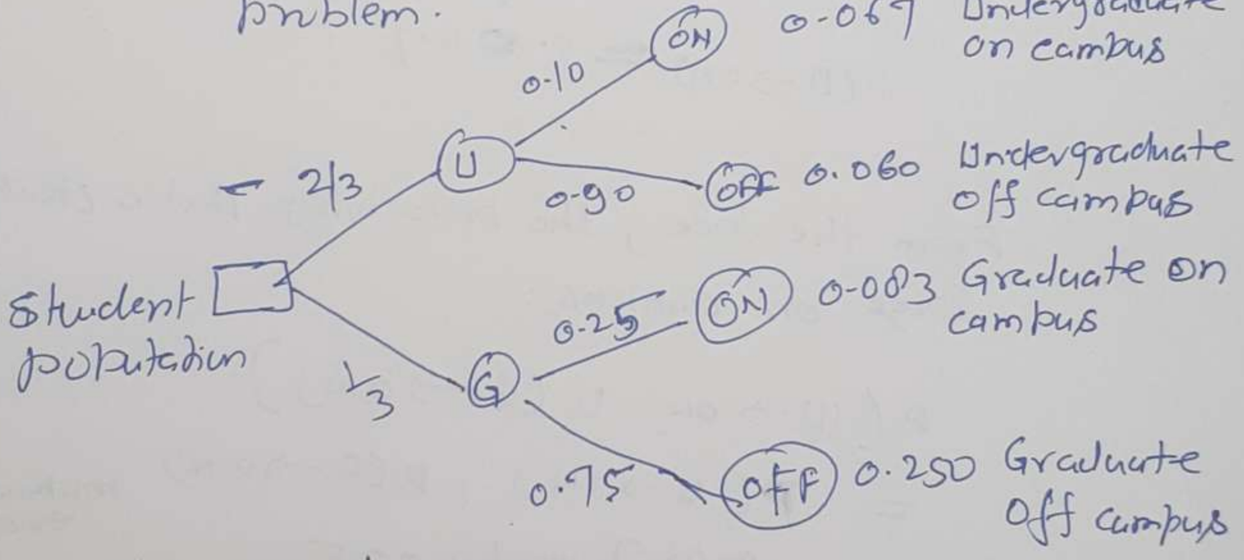
$$\boxed{P(A \cup B) = 1 - p(1-p)}$$

**Example:—** A university has twice as many undergraduate students as graduate students. 25% of the graduate students live on campus and 10% of the undergraduate students live on campus.

(a) If a student is chosen at random from the student population, what is the probability that the student is an undergraduate student living on campus?

(b) If a student living on campus is chosen at random, what is the probability that the student is a graduate student?

**Solution:—** Use tree diagram to solve the problem.




Let no. of graduate student $= x$

$\therefore$ no of undergraduate student $= 2x$

$\therefore$ Total student population $= x + 2x$
$= 3x$

$\therefore$ Proportion of UG student in population $= \dfrac{2x}{3x}$
$= \dfrac{2}{3}$

probation of graduate student ~~p~~ in population

$= \frac{x}{3x}$

$= \frac{1}{3}$

In tree  G : graduate student

U : undergraduate student

ON : living in campus

OFF : living off campus

(a) P(U → ON) : Probability that undergraduate student live in campus

$$P(U \to ON) = P(U) \times P(ON)$$

$$= \frac{2}{3} \times 0.10$$

$$= \frac{0.20}{3}$$

$$\boxed{P(U \to ON) \approx 0.067}$$

(b) From the tree, the probability that a student lives on campus :

$$P((U \to ON) \cup (G \to ON))$$

$$= P(U \to ON) + P(G \to ON) \quad \text{mutually exclusive}$$

$$= 0.067 + \frac{1}{3} \times 0.25$$

$$\approx 0.067 + 0.083$$

$$= 0.15$$

Thus, the probability that a randomally selected student living on campus is a graduate student

$$\approx \frac{P(G \to ON)}{P(U \to ON) + P(G \to ON)} = \frac{0.083}{0.15} = 0.55$$

We can use Baye's theorem to solve the problem as follows:

$$P(G/ON) = \frac{P(ON/G)\,P(G)}{P(ON/U)\,P(U) + P(ON/G)\,P(G)}$$

$$= \frac{(0.25)(\frac{1}{3})}{(0.25)(\frac{1}{3}) + (0.10)(\frac{2}{3})}$$

$$P(G/ON) = \frac{5}{9}$$

$$\boxed{P(G/ON) = 0.55}$$

Ex:- A multiple-choice-exam consists of 4 choices per question. On 75% of the questions, Pat thinks she knows the answer; and on the other 25% of the questions, she just guesses random. Unfortunately even when she thinks she knows the answer, Pat is right only 80% of the time.

(a). What is the probability that her answer to an arbitrary question is correct?

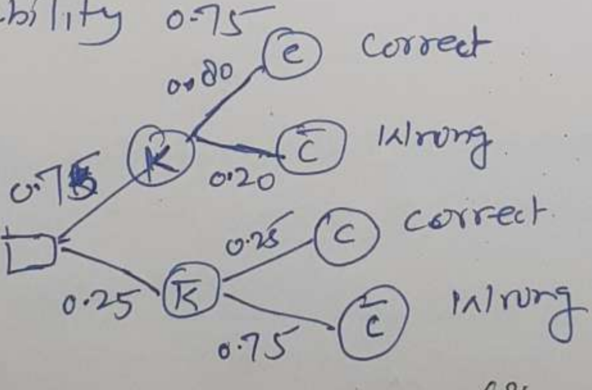
(b) Given that her answer to a question is correct, what is the probability that it was a lucky guess? (This means that it is among the questions whose answers she guessed at random)

Solution :- We can use tree diagram as follows:

K : Event that "Pat thinks she knows" the answer

$\bar{K}$ : Event that " Pat does not know". the answer

- Under event K she is correct (C) with probability 0.80 and not correct ($\bar{C}$) with probability (0.20).
- under the event $\bar{K}$, she is correct with probability 0.25 because she is equally likely to chose any of the 4 answer, therefore, she is not correct with probability 0.75



(a) The probability that Pat's answer to an arbitrary question is correct is given by

$$P[\text{Correct Answer}] = P[K \to C] + P[\overline{K} \to C]$$

$$= (0.75) \times (0.8) + (0.25) \times (0.25)$$

$$\boxed{P[\text{Correct Answer}] = 0.6625}$$

This can also be obtained by the direct method as follows:

$$P(\text{Correct Answer}) = P[\text{Correct Answer} \mid K] \cdot P[K] + P[\text{Correct Answer} \mid \overline{K}] \cdot P(\overline{K})$$

$$= (0.80) \cdot (0.75) + (0.25)(0.25)$$

$$\boxed{P[\text{Correct Answer}] = 0.6625}$$

(b) Given that she gets a question correct, the probability that it was lucky guess is given by:

~~P[ Lucky Guess | Correct Answer] = P[K̄ → C]~~

$$P[\text{Lucky Guess} \mid \text{Correct Answer}] \cdot P[\text{Correct Answer}] = P[\overline{K} \to C]$$

$$P[\text{Lucky Guess} \mid \text{Correct Answer}] = \frac{P[\overline{K} \to C]}{P[\text{Correct Answer}]}$$

$$= \frac{(0.25)(0.25)}{0.6625}$$

$$= \frac{0.0625}{0.6625}$$

$$\boxed{P[\text{Lucky Guess} \mid \text{Correct Answer}] = 0.0943}$$

# Random Variable

[ How set operation is converted to point operation ]

Since probability $f^n$ P in probability space $(\Omega, \mathcal{A}, P)$ is a set function ($P: \mathcal{A} \longrightarrow [0,1]$), ~~it is very difficult~~ (set function $\equiv$ domain is set) and range is $[0,1]$. Since domain of P is set function it is ~~very difficult~~ not possible ~~convenient~~ to perform arithmetic and algebraic operations and very well established real analysis and calculus to study the properties of P, Moreover, in practice one frequently observes some function of elementary events

- In order to define point function equivalent to P, we must represent the domain of P i.e, event space $\mathcal{A}$, with the help of real number
- Basic building blocks of an event space are the an outcome of an experiment, thus we first maps the sample space set outcomes to ~~R~~ real number $\mathbb{R}$.

Ex: Consider an experiment tossing a coin twice:

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

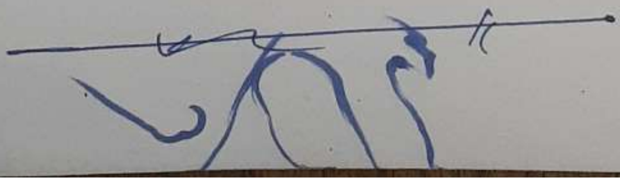
where H and T denotes head and tail, respectively. ~~Now~~

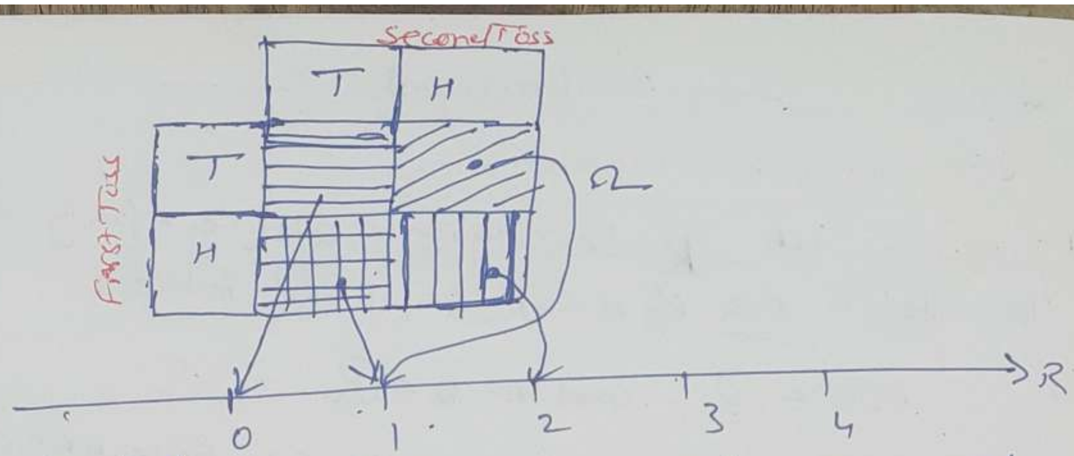
- Now we try to map $\Omega$ to set of real number R

Now consider the $f^n$ $X: \Omega \longrightarrow R$ as follows:

$$X(\omega) = \text{number of } H^{\text{'s}} \text{ in } \omega.$$

Then $X(HH) = 2$, $X(HT) = X(TH) = 1$, $X(TT) = 0$






— How to induce event space into a real numbers and corresponding set operation to point operation with the help of X?

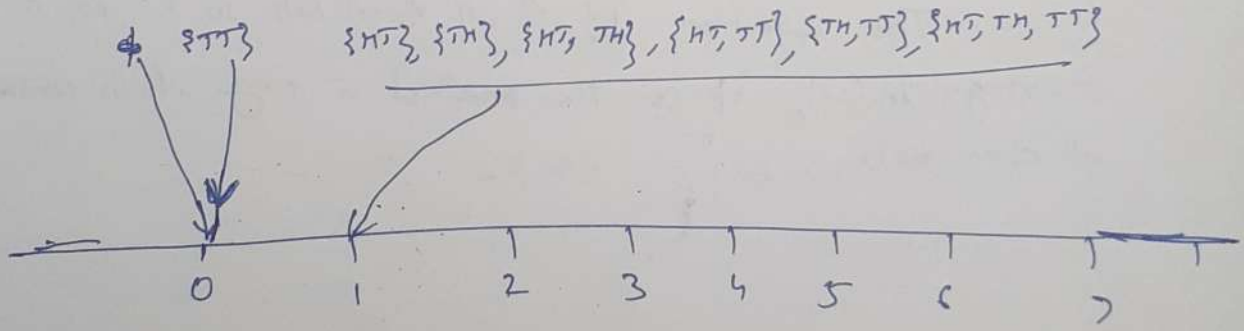
Now event space corresponding above experiment

$$A = \begin{cases} \phi, \{HH\}\ \{HT\},\ \{TH\}\ \{TT\}, \\ \{HH,HT\}, \{HH,TH\}, \{HH,TT\}, \{HT,TH\}, \{HH,TT\}, \{TH,TT\}, \\ \{HH,HT,TH\}, \{HH,HT,TT\}, \{HT,TH,TT\}, \{TH,TT,HH\}, \\ \{HH,HT,TH,TT\} \end{cases}$$

Random Variable

consider the events in which number of head less than equal to 1. These events in event space are $\phi$, {HT}{TH}, {TT}, ~~{HT,TH}~~, ~~{TH,TT}~~, ~~{HH,HT}~~ ~~{HH,TH}~~ ~~{HH,TT}~~ ~~{HT,TH}~~ {HT,TH}, {HT,TT}, {TH,TT}, {HT,TH,TT}

— In many cases, one is not interested in the occurrence of the particular outcome in the sample space. Rather, one would like to know the number of heads in two tosses or in more general in n tosses.




∴ {~~...~~} Thus event is mapped in ~~an~~ ~~a~~ interval [0, 1]

more to (3) similarly we can map event = no. of head less than equal 2 [0, 2]

# Random variable :

## Some mathematical definition :-

① :- **Ring** :- Let $\mathcal{R}$ be collection of sets (i.e. A family of sets) is called a ring if $A \in \mathcal{R}$ and $B \in \mathcal{R}$ implies:

$$A \cup B \in \mathcal{R}, \text{ and } A - B \in \mathcal{R} \quad \text{--- ①}$$

Since $A \cap B = A - (A - B)$, we also have $A \cap B \in \mathcal{R}$ if $\mathcal{R}$ is ring.

② **σ- Ring** :- A ring $\mathcal{R}$ is called a σ-ring if

$$\bigcup_{n=1}^{\infty} A_n \in \mathcal{R}, \text{ whenever } A_n \in \mathcal{R} \ (n=1,2,3,\cdots)$$

Note :- Since $\bigcap_{n=1}^{\infty} A_n = A_1 - \bigcup_{n=1}^{\infty} (A_1 - A_n)$

we also have $\bigcap_{n=1}^{\infty} A_n \in \mathcal{R}$

if $\mathcal{R}$ is a σ-ring.

③ **Borel set** :- We say that $E$ is a Borel set of $E$ can be obtained by a countable number of operations, starting from open sets, each operation is consisting of taking unions, intersections, or complements.

Note :- The collection $\mathcal{B}$ of all Borel sets in $R^p$ is a σ-ring; infact, it is the smallest σ-ring which contains all open sets.

③

Thus we see that X maps the event ~~in the~~ on the interval of $\mathbb{R}$.

* Previously discussed event; "the number of head is less than equal to 1" can be equivalently say that "number of ~~eve~~ head is less than 2".

↪ In this case X maps the ~~event~~ event on the interval $[0, 2)$

* "The number of head less than equal to two" can be equivalently say that "The number of head less than three".

↪ In this case X maps the event on the interval $[0, 3)$.

– In general :- We can associate an event in $\Omega$ over sample space $\Omega$ with an interval in $\mathbb{R}$:

(i) Singleton subset of $\mathbb{R}$. $\{X = a\}$

(ii) $(a, b] = \{x \mid a < x \leq b\}$

$[a, b) = \{x \mid a \leq x < b\}$

$[a, b] = \{x \mid a \leq x \leq b\}$

$(a, b) = \{x \mid a < x < b\}$

(iii) $(-\infty, a] = \{x \mid -\infty < x \leq a\}$

$(-\infty, a) = \{x \mid -\infty < x < a\}$

$(a, \infty) = \{x \mid a < x < +\infty\}$

$[a, \infty) = \{x \mid a \leq x < \infty\}$

– From above discussion we see that following :

* X maps sample space $\Omega$ to real number $\mathbb{R}$

* with the help of X we map event space to ~~int~~ collection of intervals of $\mathbb{R}$ [say $(\mathcal{B}(\mathbb{R}))$] $\mathcal{A}$, ⓚ

* Now objective is to ~~introduce~~ induce some measure (say $P_{\mathcal{B}}$) on $\mathcal{B}$ with the help of probability (measure) function $P$ of probability space $(\Omega, \mathcal{A}, P)$, which satisfy the axioms of probability i.e. Thus we have another space $\underline{(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})}$ which is induced by $\underline{(\Omega, \mathcal{A}, P)}$.

$$
\begin{tikzcd}
 & {[0,1]} & \\
\mathcal{A} \; (\equiv \bigcup_i A_i) \arrow[ur, "P"] & \Omega \arrow[l] \arrow[r, "X"] & \mathbb{R} \; (\equiv \bigcup_i \mathcal{B}_i) \arrow[ul, "P_{\mathcal{B}}(\mathcal{B}_i)"'] \arrow[ll, bend left, "X^{-1}(\mathcal{B}_i)"]
\end{tikzcd}
$$

$$P_{\mathcal{B}}(\mathcal{B}_i) \equiv P\left(X^{-1}(\mathcal{B}_i)\right)$$

Here $A_j = X^{-1}(\mathcal{B}_i)$ for some $j$

* $X \to$ Partition $\Omega$ $\quad$ $X^{-1}(\mathcal{B}_i)$ are disjoint?

Note :- Given Probability measure ~~$P_B$~~ $P_{\mathcal{B}}$ on $\mathcal{B}$, we have Borel Probability space, $(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})$, whose events are Borel set, and which therefore include the open — and — closed interval in $\mathbb{R}$.

→ A Bit Mathematical Approach :—
(Measure Theoretic Approach)

Inverse Image or Pre-Image of a fn

Let $\Omega$ be sample space & $\mathbb{R}$ is real number, consider a function 'f' from $\Omega$ to $\mathbb{R}$,

$$f : \Omega \longrightarrow \mathbb{R},$$

Given a function f; the inverse image, or Pre-image, of a set of points ~~$M \in$~~ $M \subset \mathbb{R}$ is defined as:

$$f^{-1}(M) = \{ \omega \mid f(\omega) \in M \} \subset \Omega$$

By definition, a function f is said to be $\mathcal{A}$-measurable or Just measurable, if and only if,

$$f^{-1}(M) \in \mathcal{A}; \quad \forall M \in \mathcal{B}$$

- Given a probability space $(\Omega, \mathcal{A}, P)$ and an $\mathcal{A}$-measurable function f, we can legitimately apply the probability measure P to the pre-image of any Borel set $M \in \mathcal{B}$ as $f^{-1}(M)$ must be an event in $\mathcal{A}$. In this we we say that the function f is an $\mathcal{A}$-measurable random variable.

Definition of Random Variable :— Given sample space $\Omega$, with an associated $\sigma$-algebra, $\mathcal{A}$, of events in $\Omega$, a Random variable $X : \Omega \longrightarrow \mathbb{R}$, is an $\mathcal{A}$-measurable real-valued function on $\Omega$

* A random variable, X, on probability space $(\Omega, \mathcal{A}, P)$ induce a probability ~~space~~ measure on $(\mathbb{R}, \mathcal{B})$ via the definition

$$P_{\mathcal{B}}(M) \triangleq P\left(X^{-1}(M)\right), \quad \forall M \in \mathcal{B}.$$

$$\boxed{P_{\mathcal{B}} \equiv P X^{-1}}$$

It does not means ~~compo~~ composition of $f^n$.

$\Rightarrow$ Thus corresponding to the probability space $(\Omega, \mathcal{A}, P)$ ~~we have~~ and random variable $X$, we have the induced probability space $(\mathbb{R}, \mathcal{B}, P_{\mathcal{B}})$

Refered as Base space

underlying probability space

Distribution function :— Let the Borel set $M_x$ denote the closed half-space:

$$M_x = (-\infty, x] = \{\xi \mid -\infty < \xi \le x\}, \quad x \in \mathbb{R}$$

If $X$ is a random variable for the probability space $(\Omega, \mathcal{A}, P)$, its (Cumulative) Distribution function $F_X$ is defined by

$$F_X(x) \triangleq P_{\mathcal{B}}(M_x) = P(X^{-1}(M_x))$$

$$= P\left(\{\omega \mid X(\omega) \in M_x\}\right); \quad \forall x \in \mathbb{R}.$$

Note:- $\mathcal{A}$-measurability of random variable $X$ and existence of probability measure $P$ on $\mathcal{A}$-events are both crucial to this definition.

Rohtagi

Random Variable : Def^n(1) Let $(\Omega, \mathcal{A})$ be a sample space. A finite, single-valued function that maps $\Omega$ into $\mathbb{R}$ is called random variable (RV) if the inverse images under $X$ of all Borel sets in $\mathbb{R}$ are events, that is, if:

$$X^{-1}(B) = \{\omega \mid X(\omega) \in B\} \in \mathcal{A}, \text{ for all } B \in \mathcal{B}$$

~~Def^n 2/ Res~~
(~~Res~~

Def^n 2/Result :- $X$ is an RV if and only if for each $x \in \mathbb{R}$

$$\{\omega \mid X(\omega) \le x\} = \{X \le x\} \in \mathcal{A}$$

$= [X \le x]$ → some reference Both

Ex Note :- If $X$ is an RV, the sets $\{X = x\}$, $\{a < X \le b\}$, $\{X < x\}$, $\{a \le X < b\}$, $\{a < X < b\}$, $\{a \le X \le b\}$ are all events.

Ex 1 :- For any set $A \subseteq \Omega$, define:

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

$I_A$ : Indicator function of set $A$. $I_A$ is RV if and only if ~~A~~ $A \in \mathcal{A}$.

Ex 2 :- Let $\Omega = \{H, T\}$ and $\mathcal{A}$ be the class of all subset of $\Omega$. Define $X$ by $X(H) = 1$, $X(T) = 0$. Then

$$X^{-1}(-\infty, x] = \begin{cases} \phi, & \text{if } x < 0 \\ \{T\}, & \text{if } 0 \le x < 1 \\ \{H, T\}, & \text{if } x \ge 1 \end{cases}$$

we see that $X$ is R-V.

Ex:- Let $\Omega = \{HH, TT, HT, TH\}$ and $\mathcal{A}$ be the class of all subsets of ~~$\mathcal{A}$~~ $\Omega$. Define $X$ by

$$X(\omega) = \text{number of } H\text{'s in } \omega.$$

Then $X(HH) = 2$, $X(HT) = X(TH) = 1$, and $X(TT) = 0$

$$X^{-1}(-\infty, x] = \begin{cases} \phi & x < 0 \\ \{TT\} & 0 \le x < 1 \\ \{HT, TH\} & 1 \le x < 2 \\ \Omega, & x \ge 2 \end{cases}$$

Note:- $(\Omega, \mathcal{A})$ be discrete sample space, i.e., let $\Omega$ be a countable set of points and $\mathcal{A}$ be the class of all subset of $\Omega$. Then every numerical-valued function defined on $(\Omega, \mathcal{A})$ is an R.V.

Ex:- Let $\Omega = [0, 1]$ and $\mathcal{A} = \mathcal{B} \cap [0, 1]$ be the $\sigma$-field of Borel sets on $[0, 1]$. Define $X$ on $\Omega$ by

$$X(\omega) = \omega, \quad \omega \in [0, 1]$$

Clearly $X$ is RV. Any Borel subset of $\Omega$ is an event

Note:- Let $(\Omega, \mathcal{A})$ be a sample space, and $a, b$ be constants. Then ~~$aX+b$~~ $aX+b$ is also an RV on $(\Omega, \mathcal{A})$. Moreover, $X^2$ is an RV and so also is $\frac{1}{X}$ provided that $[X = 0] = \phi$.

## Probability Distribution of a Random Variable

**Result:-** The RV $X$ defined on the probability space $(\Omega, \mathcal{A}, P)$ induces a probability space $(\mathbb{R}, \mathcal{B}, Q)$ by means of the correspondence

$$Q(B) = P[X^{-1}(B)]$$
$$= P\{\omega \mid X(\omega) \in B\} \quad \text{for all } B \in \mathcal{B}$$

~~We write~~ Notation: $Q \equiv PX^{-1}$ and call $Q$ or $PX^{-1}$ the (probability) distribution of $X$.

**Result:-** Let $X$ be an R.V. defined on $(\Omega, \mathcal{A}, P)$. Define a point function $F_X()$ on $\mathbb{R}$, using above result

$$F_X(x) = Q(-\infty, x] = P\{\omega : X(\omega) \le x\} \quad \text{for all } x \in \mathbb{R}$$
$$= P[X \le x]$$

Then $F$ is called ~~distribution~~ the distribution function of RV $X$.

* That is, $F_X(x)$ denotes the probability that the random variable $X$ takes on a value that is less than or equal to $x$. Some properties of $F_X(x)$ are:

(i) $F_X(x)$ is a non-decreasing function, i.e. $x_1 < x_2$, then $F_X(x_1) \le F_X(x_2)$. Thus $F_X(x)$ can increase or stay level, but it ~~can not~~ cannot go down, as $x$ increases.

(ii) $0 \le F_X(x) \le 1$

(iii) $F_X(+\infty) = 1$

(iv) $F_X(-\infty) = 0$

(v) $P[a < X \le b] = F_X(b) - F_X(a)$

(vi) $P[X > a] = 1 - P[X \le a] = 1 - F_X(a)$

~~Note~~ :- $F_X(x)$ is also called cumulative distribution function (CDF). of r.v. X.
~~Examp~~

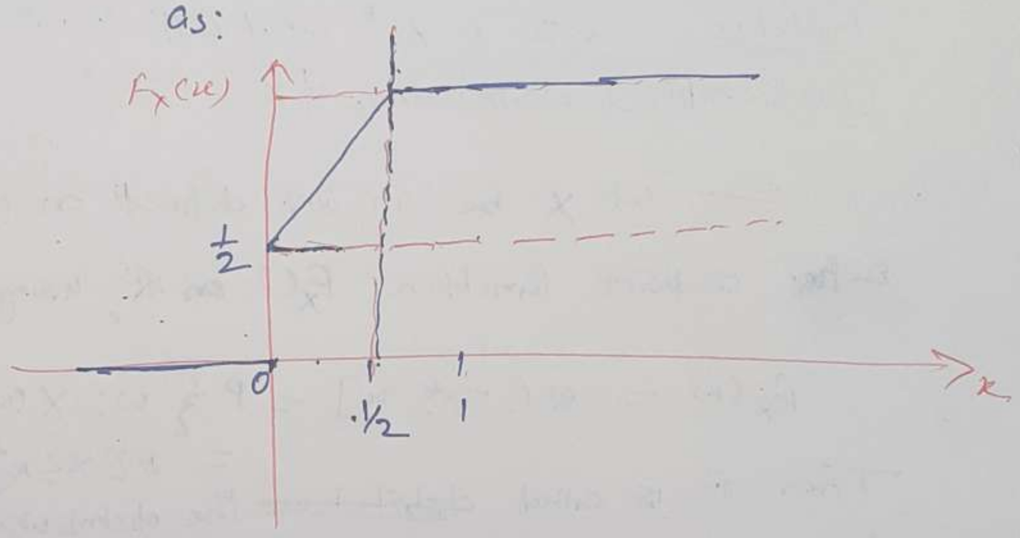
Example :- The CDF of random variable X is given by

$$F_X(x) = \begin{cases} 0 & ; \quad x<0 \\ x+\frac{1}{2} & ; \quad 0 \le x < \frac{1}{2} \\ 1 & ; \quad x \ge \frac{1}{2} \end{cases}$$

(a) :- Draw the graph of CDF

(b) Compute $P[X > \frac{1}{4}]$

Solution :- (a) The graph of the CDF is shown as:




(b) $P(X > \frac{1}{4}) = 1 - P(X \le \frac{1}{4})$

$= 1 - F_X(\frac{1}{4})$

$= 1 - (\frac{1}{4} + \frac{1}{2})$

$P[X > \frac{1}{4}] = 1 - \frac{3}{4}$

$$\boxed{P[X > \frac{1}{4}) = \frac{1}{4}}$$

# Discrete Random Variable :

Defn:- A discrete random variable is a random variable (r.v). that can take ~~on~~ at most a countable number of possible values. The ~~no~~ number can be finite or infinite; that is, the random variable can have a countably finite number of values or a countably infinite number of values.

[In simpler way: A discrete R.V. is function, $X : \Omega \to \mathbb{R}$, that takes a discrete set of values]

~~PMF~~ (P

## Probability mass function (PMF) / ~~Probability Function~~ / Probability Density function:

~~For~~ For a discrete random variable $X$, the probability mass function (PMF), $p_X(x)$, is defined as follows:

$$\boxed{p_X(x) = P[X = x], \quad \text{where } \sum_{x=-\infty}^{+\infty} p_X(x) = 1}$$

Note:-

The PMF is ~~non-zero~~ nonzero for at most a countable number of values of ~~x~~ $x$. In particular, if we assume that $X$ can only assume one of the values $x_1, x_2, \cdots, x_n$, then

$$p_X(x_i) \geq 0 \qquad i = 1, 2, \cdots, n$$
$$p_X(x_i) = 0, \qquad \text{otherwise}$$

CDF:- The CDF of $X$ can be represented ~~as~~ in terms of $p_X(x)$ as follows:

$$\boxed{F_X(x) = \cancel{\sum p(x_k)} \sum_{k \leq x} p(k)}$$

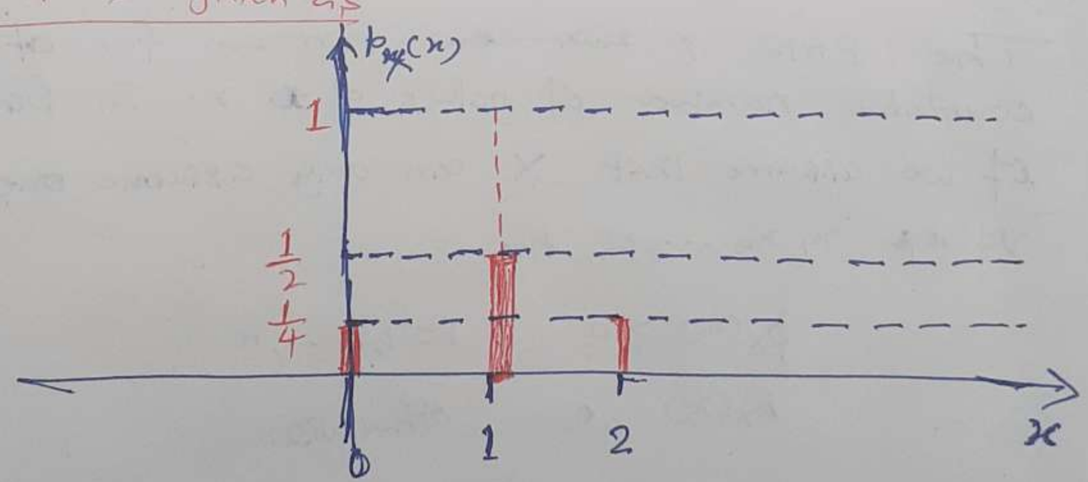
- The CDF of a discrete random variable is a series of step functions. That is, if $X$ takes on the values at $x_1, x_2, \ldots x_n, \ldots$ where $x_1 < x_2 < x_3 \ldots < x_n \ldots$, then the values of $F_X(x)$ is constant in the interval $x_{i-1}$ and $x_i$ and then takes jump of size $p_X(x_i)$ at $x_i = 1, 2, 3, \ldots$

Thus in this case, $F_X(x)$ represents the sum of probability masses we have encountered as we move from $-\infty$ to $x$.

Example:- Assume that $X$ has the PMF given by

$$p_X(x) = \begin{cases} \frac{1}{4}, & x=0 \\ \frac{1}{2}, & x=1 \\ \frac{1}{4}, & x=2 \\ 0, & \text{otherwise} \end{cases}$$

PMF is given as




CDF is given as follows:

$$F_X(x) = \begin{cases} 0; & x < 0 \\ \frac{1}{4}, & 0 \le x < 1 \\ \frac{3}{4}, & 1 \le x < 2 \\ 1, & x \ge 2 \end{cases}$$

## Discrete Random Variable :-

A R.V. that can take on at most a countable number of possible values. The number can be finite or infinite; that is, the random variable can have a countably finite number of values, or countably infinite number of values

## Probability mass function (pmf):

For discrete random variable $X$, the probability mass function (pmf), $p_X(x)$, is defined as follows:

$$p_X(x) = P[X = x]$$

where
$$\boxed{\sum_{x=-\infty}^{+\infty} p_X(x) = 1}$$

- The pmf ~~&~~ is ~~non zero~~ nonzero for at most a countable number of values of $x$. In particular we assume that one of the values $x_1, x_2, x_3, \dots, x_n$, then

$p_X(x_i) > 0 \quad i = 1, 2, \dots, n.$

$p_X(x_i) = 0 \quad$ otherwise.

- CDF of $X$ can be expressed ~~as~~ in terms of $p_X(x)$ as follows:

$$\boxed{F_X(x) = \sum_{k \le x} p_X(k)}$$

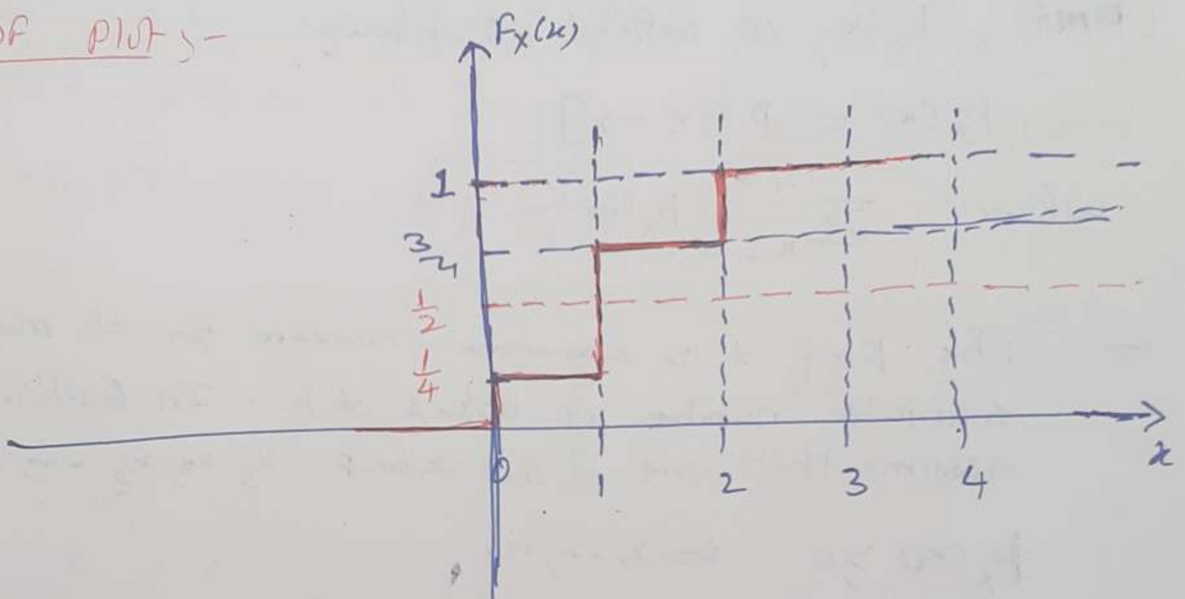
- CDF of discrete random variable is series of step functions. That is, if $X$ ~~take~~ takes values at $x_1, x_2, x_3, \dots x$ where $x_1 < x_2 < x_3 < \dots$, then $F_X(X)$ is constant in interval between $x_{i-1}$ and $x_i$ and then takes ~~value~~ a jump of size $p_X(x_i)$ at $x_i$, $i = 1, 2, 3 \dots$

$\Rightarrow$ $F_X(x) \equiv$ sum of all probability masses we have encountered as we move from $-\infty$ to $x$.

Ex:- Assume that $X$ has the PMF given by

$$P_X(x) = \begin{cases} \frac{1}{4} ; & x=0 \\ \frac{1}{2} ; & x=1 \\ \frac{1}{4} ; & x=2 \\ 0 ; & \text{otherwise} \end{cases}$$

CDF plot:-




Example:- Let the random variable $X$ denote the number of heads in the three tosses of a fair coin

(a) What is the PMF of $X$?

(b) Sketch the CDF of $X$.

Solution:- (a) Sample space of the experiment is:

$$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTT\}$$

The different events defined by the random variable $X$ are as follows:

$[X=0] = \{TTT\}$

$[X=1] = \{HTT, THT, TTH\}$

$[X=2] = \{HHT, HTH, THH\}$

$[X=3] = \{HHH\}$

$X : \Omega \to \mathbb{R}$

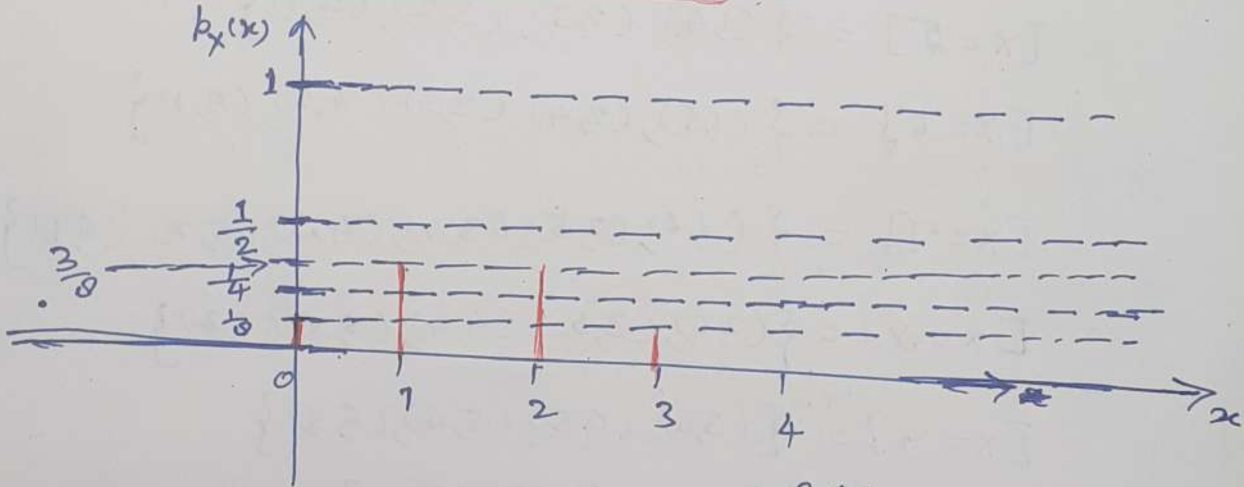
$X(\omega)$ = no of heads in $\omega$.

$\omega \in \Omega$

Since the eight sample points in $\Omega$ are symmetric i.e. equally likely, the PMF of X is as follows

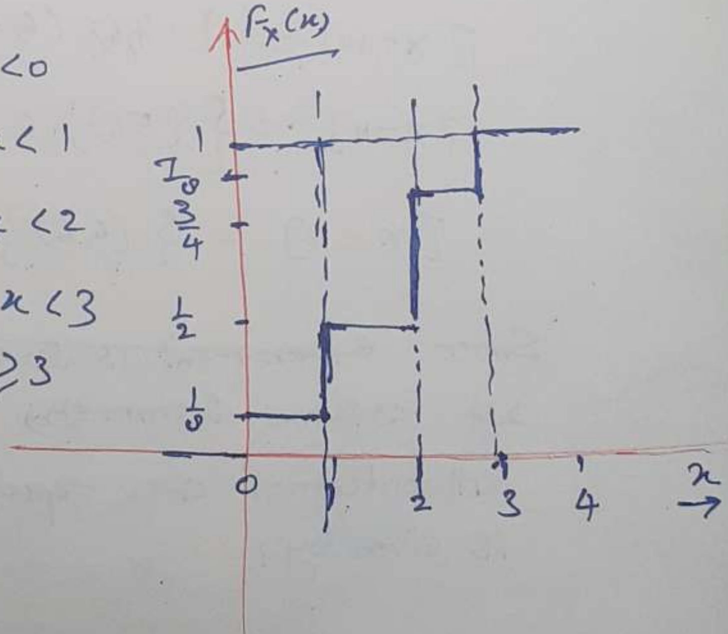
$$p_X(x) = \begin{cases} \frac{1}{8}; & x=0 \\ \frac{3}{8}; & x=1 \\ \frac{3}{8}, & x=2 \\ \frac{1}{8}, & x=3 \\ 0, & \text{otherwise} \end{cases}$$

Graphically $p_X(x)$ is represented by :-



CDF

$$F_X(x) = \begin{cases} 0, & x<0 \\ \frac{1}{8}, & 0 \le x < 1 \\ \frac{1}{2}, & 1 \le x < 2 \\ \frac{7}{8}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$



Ex :- Let the random variable $X$ denote the sum obtained in rolling a pair of fair dice. Determine the PMF of $X$.

Solution :- Let ~~the~~ the pair $[a, b]$ denote the outcomes of the roll, where 'a' is the outcome of one die and 'b' is the outcome of the other. Thus, the sum of the outcomes is $X = a+b$. The different events defined by random variable $X$ are as follows:

$$[X=2] = \{(1,1)\}$$
$$[X=3] = \{(1,2), (2,1)\}$$
$$[X=4] = \{(1,3), (2,2), (3,1)\}$$
$$[X=5] = \{(1,4), (2,3), (3,2), (4,1)\}$$
$$[X=6] = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$$
$$[X=7] = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}$$
$$[X=8] = \{(2,6), (3,5), (4,4), (5,3), (6,2)\}$$
$$[X=9] = \{(3,6), (4,5), (5,4), (6,3)\}$$
$$[X=10] = \{(4,6), (6,4), (5,5)\}$$
$$[X=11] = \{(5,6), (6,5)\}$$
$$[X=12] = \{(6,6)\}$$

Since ~~Experiment is symm~~

We assume symmetry in experiment $\Rightarrow$ ~~events~~ all outcomes are equally likely. Thus PMF of $X$ is given by:

$$p_X(x) = \begin{cases} \frac{1}{36}; & x=2 \\ \frac{2}{36}; & x=3 \\ \frac{3}{36}, & x=4 \\ \frac{4}{36}, & x=5 \\ \frac{5}{36}, & x=6 \\ \frac{6}{36}, & x=7 \\ \frac{5}{36}, & x=8 \\ \frac{4}{36}, & x=9 \\ \frac{3}{36}, & x=10 \\ \frac{2}{36}, & x=11 \\ \frac{1}{36}, & x=12 \\ 0, & \text{otherwise} \end{cases}$$

Ex :- The PMF of the number of components K of a system that fail is defined by

$$p_K(k) = \begin{cases} \binom{4}{k} (0.2)^k (0.8)^{4-k}, & k=0,1,2,3,4. \\ 0, & \text{otherwise} \end{cases}$$

(a) What is the CDF of K?

(b) What is the probability that less than 2 components of the system fail?

Solution :- (a) The CDF of K is given by

$$F_K(k) = P[K \le k]$$

$$= \sum_{m \le k} p_k(m) = \sum_{m=0}^{k} p_k(m)$$

$$= \sum_{m=0}^{k} \frac{4!}{(4-m)!\, m!} (0.2)^m (0.8)^{4-m}$$

$$F_K(k) = \begin{cases} 0 & k<0 \\ (0.8)^4, & 0 \le k < 1 \\ (0.8)^4 + 4(0.2)(0.8)^3, & 1 \le k < 2 \\ (0.8)^4 + 4(0.2)(0.8)^3 + 6(0.2)^2(0.8)^2; & 2 \le k < 3 \\ (0.8)^4 + 4(0.2)(0.8)^3 + 6(0.2)^2(0.8)^2 + 4(0.2)^3(0.8), & 3 \le k < 4 \\ (0.8)^4 + 4(0.2)(0.8)^3 + 6\cdot(0.2)^2(0.8)^2 + 4(0.2)^3(0.8) + (0.2)^4; & k \ge 4 \end{cases}$$

$$F_K(k) = \begin{cases} 0; & k<0 \\ 0.4096; & 0 \le k < 1 \\ 0.8192; & 1 \le k < 2 \\ 0.9728; & 2 \le k < 3 \\ 0.9984; & 3 \le k < 4 \\ 1.0; & k \ge 4 \end{cases}$$

(b) The probability that the less than 2 components of the system fail is the probability that ~~their~~ either no component fails or one component fail, which is given by:

$$P[K<2] = P[\{K=0\} \cup \{K=1\}]$$

$$= P[\{K=0\}] + P[\{K=1\}]$$

Events are mutually exclusive.

~~Events are mutuall.~~

$$= F_K(1)$$

$$P[K<2] = 0.8192$$

**Example :—** The PMF of the number N of customers that arrives at a local library within one hour interval is defined by

$$p_N(n) = \begin{cases} \frac{5^n}{n!} e^{-5}, & n = 0, 1, \cdots \\ 0 ; & \text{otherwise} \end{cases}$$

what is probability that atmost two customers arrive at the library with in one hour ?

**Solution :—** The probability that at most two customers arrive at the library with in one hour is the probability that 0 or 1 or 2 customers arrive at the library within one hour, which is:

$$P[N \leq 2] = P[\{N=0\} \cup \{N=1\} \cup \{N=2\}]$$
$$= P[N=0] + P[N=1] + P[N=2]$$
$$= p_N(0) + p_N(1) + p_N(2)$$

$$P[N \leq 2] = e^{-5}\left\{1 + 5 + \frac{25}{2}\right\}$$

$$\boxed{\begin{aligned} P[N \leq 2] &= 18.5\, e^{-5} \\ &= 0.1246 \end{aligned}}$$

where the second equality on the first line is due the fact that the three events are mutually exclusive

## obtaining the PMF from the CDF:-

we know that CDF ~~is relat~~ $F_x(x)$ is related to PMF, $p_x(x)$ as follows

$$\boxed{F_x(x) = \sum_{k \le x} p_x(k)}$$

- CDF of random variable have staircase plot, with jumps at those ~~points~~ values of the random variable where the PMF have non zero value.

  ↳ The size of jump at a value of r.v. ≡

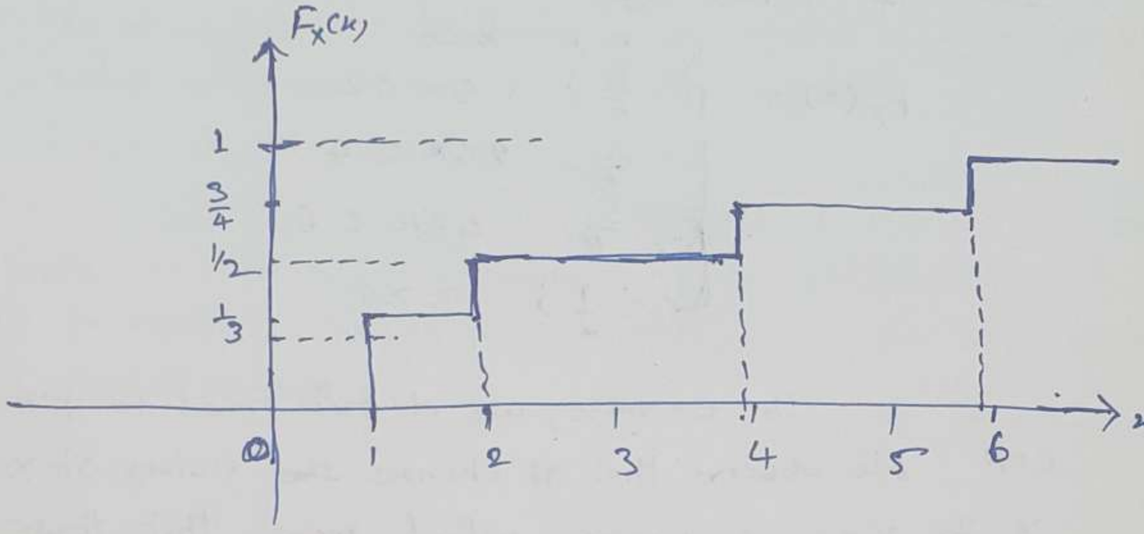
  The value of the PMF at the value.

- Thus PMF ~~takes~~ of random variable ~~takes value~~ by noting that the r.v. only takes on values that have ~~non-zero~~ nonzero probability at those points where ~~the~~ jumps occurs.

- The probability that the r.v. takes on any other value than where the jump occur is zero.

→ The probability that the random variable takes a value where a jump occurs is equal to the size of the jump.

Ex :- The plot of CDF of a discrete random variable X is shown as below:




Find the PMF X ?

Answer :- The r.v. takes on values with non-zero ~~probab~~ probability at $X=1$, $X=2$, $X=4$, and $X=6$. The size of jump at $X=1$ is $\frac{1}{3}$, the size of the jump at $X=2$ is $\frac{1}{2}-\frac{1}{3}=\frac{1}{6}$, the size of jump at $X=4$ is $\frac{3}{4}-\frac{1}{2}=\frac{1}{4}$ and the size of jump at $X=6$ is $1-\frac{3}{4}=\frac{1}{4}$. Thus the PMF of X is given by :-

$$p_X(x) = \begin{cases} \frac{1}{3}; & x=1 \\ \frac{1}{6}; & x=2 \\ \frac{1}{4}; & x=4 \\ \frac{1}{4}; & x=6 \\ 0; & \text{otherwise.} \end{cases}$$

## Example :-

Find the PMF of the discrete random variable X whose CDF is given by

$$F_X(x) = \begin{cases} 0 ; & x < 0 \\ \frac{1}{6} ; & 0 \le x < 2 \\ \frac{1}{2} ; & 2 \le x < 4 \\ \frac{5}{8} ; & 4 \le x < 6 \\ 1 ; & x \ge 6 \end{cases}$$

**Solution** :- In example, we do not need to plot the CDF. We observe that it changes the values at $X=0$, $X=2$, $X=4$ and $X=6$, which means that these are the values of the r.v. that have non-zero probability. The next task after isolating these values with non-zero probabilities is to determine their probabilities.

↪ The first value is $p_X(0)$, which is $\frac{1}{6}$.

At $X=2$ the size of jump is $\frac{1}{2} - \frac{1}{6} = \frac{1}{3} = p_X(2)$

Similarly at $X=4$ the size of jump is $\frac{5}{8} - \frac{1}{2} = \frac{1}{8} = p_X(4)$

Finally, at $X=6$ the size of jump is $1 - \frac{5}{8} = \frac{3}{8} = p_X(6)$

Therefore, PMF of X is given by

$$p_X(x) = \begin{cases} \frac{1}{6} ; & x = 0 \\ \frac{1}{3} ; & x = 2 \\ \frac{1}{8} ; & x = 4 \\ \frac{3}{8} ; & x = 6 \\ 0 , & \text{otherwise.} \end{cases}$$

## Continuous Random Variable :—

↳ continuous random variables can assume an uncountable set of possible values.

[ In otherwords, a continuous r.v. takes a range of values, which may be finite or infinite in extent. e.g. $[0,1]$, $[0,\infty)$, $(-\infty,\infty)$, $[a,b]$.

Defn :— A r.v. X to be a continuous r.v. if there ~~ex~~ exists a non-negative function $f_X(x)$, defined for all real $x \in (-\infty, \infty)$, having the property that for any set A of real numbers:

$$P[X \in A] = \int_A f_X(x)\,dx$$

The function $f_X(x)$ is called the probability density function (PDF) of the random variable X and is defined by

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

The properties of $f_X(x)$ are as follows:

1. $f_X(x) \geq 0$; that is, it is non-negative function.
2. Since X must assumes some value, $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$
3. $P[a \leq X \leq b] = \int_a^b f_X(x)\,dx$
4. $P[X < a) = P[X \leq a] = F_X(a) = \int_{-\infty}^{a} f_X(x)\,dx$

Note :— From property (3)

$$P[X=a] = P[a \leq X \leq a] = \int_a^a f_X(x)\,dx = 0$$

↳ Thus, probability that a continuous random variable will assume any fixed value is zero.

Note 2:- Probability density function $f_X(x)$ of a ~~cont~~ continuous random variable is the analogue of the probability mass function $p_X(x)$ of a discrete random variable. Here are two important differences:

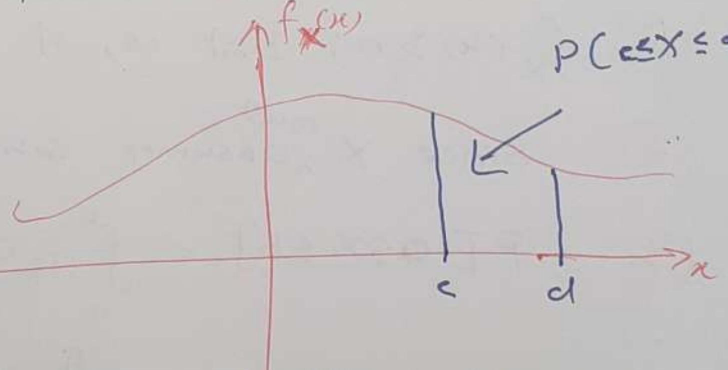
1. Unlike $p_X(x)$, the pdf $f_X(x)$ is not probability. We must have to integrate it to get probability

2. Since $f_X(x)$ is not a probability, there is no restriction that $f_X(x)$ be less than or equal to 1.

Note 3:- In property (2) we integrate over $(-\infty, \infty)$ since we do not know the range of values taken by X. Formally, this makes sense because we just define $f_X(x)$ to be zero outside of the range of X.

In practice, we would integrate between bounds given by the range of X.

## Graphical View of Probability




$P([c \le X \le d])$

$= P[c \le X \le d) =$ Area under the graph between c and d.

$$= \int_a^b f(x)\, dx$$

~~Example~~

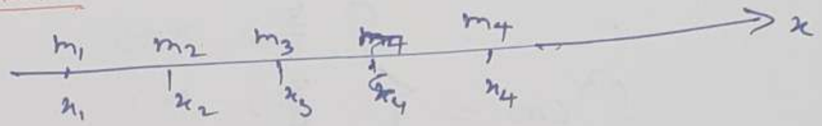
Example

## The term "Probability mass" and "probability density" :-

Why use term "mass" and "density" to describe the pmf and pdf? What is the difference between the two?

↳ These are analogous to the mass and density as in ~~Physical~~ Physics and calculas.

### Mass as a Sum :-



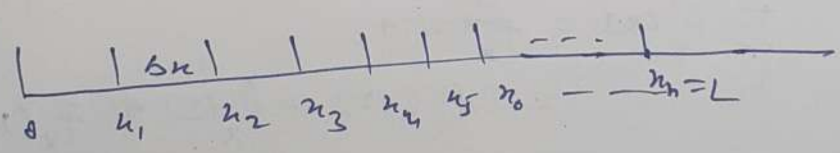

The total mass is $m_1 + m_2 + m_3 + m_4$.

We can define a "mass function" ~~$p(x) = p(x_1) + p(x_2) + p(x_3) + p(x_4)$~~ $p(x)$ with $p(x_j) = m_j$ for $j = 1, 2, 3, 4$ and $p(x) = 0$ otherwise. In this case total mass is $p(x_1) + p(x_2) + p(x_3) +$ ~~$p(x_4)$~~ $p(x_4)$.

↳ The "probability mass function" behaves in exactly same way, except it has the dimension probability instead of mass.

### mass a integral of density :-

rod have length L meters with density $f(x)$ Kg/meter.




mass of $i$th piece ≈ $f(x_i)\,\Delta x$

If the density varies ~~with the~~ continuously then

total mass $= \int_0^L f(x)\,dx$

⇒ probability density function also behaves the same way, except it has units of probability/(unit $x$) instead of kg/m.

Ex :- Check whether ~~show that~~ the followin ~~func~~ can be probability mass function for discrete random variable.

$$f_X^1(x) = \begin{cases} \frac{2}{3}, & x = -2 \\ \frac{1}{3}, & x = 5 \\ 0, & \text{elsewhere} \end{cases}$$

~~Solution :-~~

$$f_X^2(x) = \begin{cases} \frac{3}{4}, & x = -3, \\ \frac{2}{4}, & x = 0, \\ -\frac{1}{4}, & x = 2; \\ 0, & \text{elswhere.} \end{cases}$$

$$f_X^3(x) = \begin{cases} \frac{3}{5}, & x = 0, \\ \frac{3}{5}, & x = 1 \\ 0, & \text{elsewhere} \end{cases}$$

Solution :- consider $f_X^1(x)$ : for

A function $f_X(x)$ is valid pmf if

$$f_X(x_i) \geq 0$$

$$\sum_{x_i} f_X(x_i) = 1$$

Now from $f_X^1(x_i) \geq 0$, for $x_1 = -2$, $x_2 = 5$ and all other $x$'s.

$$\therefore \sum_{x_i} f_X^1(x_i) = \cancel{f_X^1(x_1)} +$$

$$= f_X(-2) + f_X(5) + f_X(x \neq -2 \text{ and } x \neq -5)$$

$$= \frac{2}{3} + \frac{1}{3} + 0$$

$$= \frac{3}{3}$$

$$= 1$$

$$\therefore \sum_{x_i} f_X^1(x_i) = 1$$

$\therefore$ $f_X^1(x)$ is valid pmf

Now ~~form~~ for $f_X^2(x)$:

clearly from defn of $f_X^2(x)$

$$\cancel{f_X^2(x_i) \geq 0 \quad \forall x_i}$$

and $$\sum_{x_i} f_X(x_i) = f_X(-3) + f_X(0) + f_X(2) +$$

$$= \frac{3}{4} + \frac{2}{4} + \left(-\frac{1}{4}\right) \neq 0$$

$$= \frac{5}{4} - \frac{1}{4}$$

$$\therefore \sum_{x_i} f_X(x_i) = \frac{4}{4} = 1$$

$$\therefore \boxed{\sum_{x_i} f_X(x_i) = 1}$$

But $f_X^2(2) = -\frac{1}{4} < 0$

$\therefore$ $f_X^2(x)$ is not valid pmf for any random variable

Now for $f_X^3(x)$ :-

$$f_X^3(0) = \frac{3}{5} > 0$$

$$f_X^3(1) = \frac{3}{5} > 0$$

$$f_X^3(x) = 0 \quad \text{for all other } x$$

Therefore $f_X^3(x) \geq 0 \quad \forall \text{all } x$

now $$\sum_{x_i} f_X^3(x_i) = f_X^3(0) + f_X^3(1) + 0$$

$$= \frac{3}{5} + \frac{3}{5} + 0$$

$$= \frac{6}{5} > 1$$

Thus $f_X^3(x)$ is not valid pmf for any r.v. $X$

Example :— check whether the following can be density functions of some random variables:

$$f_X^1(x) = \begin{cases} \frac{1}{b-a} ; & a \le x \le b, \quad b>a \\ 0 , & \text{elsewhere} \end{cases}$$

$$f_X^2(x) = \begin{cases} cx^4; & 0<x<1 \\ 0 ; & \text{otherwise} \end{cases}$$

$$f_X^3(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} ; & 0 \le x < \infty \\ 0 ; & \text{elsewhere} \end{cases}$$

$$f_X^4(x) = \begin{cases} x; & 0 \le x < 1 \\ 2-x; & 1 \le x \le 2 \\ 0; & \text{elsewhere} \end{cases}$$

Solution :— $f_X^1(x) \ge 0$, since it is either 0 or $\frac{1}{b-a}$ where $b-a>0$

∴ ~~Here~~ Hence first condⁿ is satisfied

$$\int_{-\infty}^{+\infty} f_X^1(x)\,dx = \int_{-\infty}^{a} f_X^1(x)\,dx + \int_{a}^{b} f_X^1(x)\,dx + \int_{b}^{+\infty} f_X^1(x)\,dx$$

$$\therefore \int_{-\infty}^{+\infty} f_X^1(x)\,dx = 0 + \int_{a}^{b} \frac{1}{b-a}\,dx + 0$$

$$= \int_{a}^{b} \frac{1}{b-a}\,dx$$

$$= \frac{1}{\cancel{(b-a)}} \cancel{(b-a)}$$

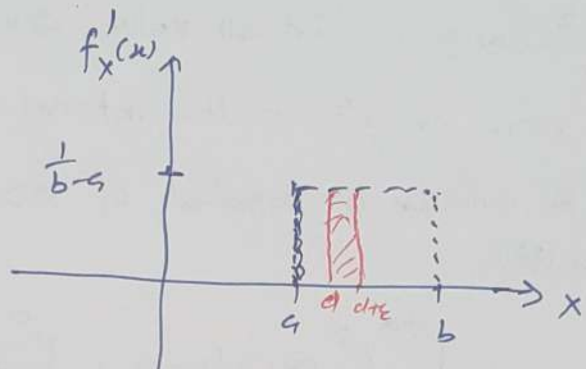
$$\therefore \boxed{\int_{-\infty}^{+\infty} f_X^1(x)\,dx = 1}$$

Thus second condition is also satisfied.

$\therefore$ $f'_X(x)$ is valid pdf.

Note:-

~~Consider $f^2_X(x)$~~ :




This density looks like rectangle and hence it is called a rectangular density.

Probability ⇒ Area under the curve

~~Thus x fa~~ Thus, probability that $x$ falls in the interval $d$ to $d+\varepsilon$ or $d \le x \le d+\varepsilon$ is given by the integral:

$$\int_d^{d+\varepsilon} \frac{1}{b-a}\, dx = \frac{\varepsilon}{b-a}$$

↳ If we take interval length $\varepsilon$ anywhere in the interval $a \le x \le b$ then the are will be the same as $\frac{\varepsilon}{b-a}$

↳ Thus, we can say that total are 1 is uniformally distributed ~~over the~~ over the pinterval $[a, b]$

↳ In this sense, this density $f'_X(x)$ is also called ~~a~~ "uniform density"

↳ unknown quantity $a$ and $b$ could be any constants free ~~for~~ of $x$. As long as $b>a$, $f_1(x)$ is density

$f_X^2(x)$ :—

$$f_X^2(x) = \begin{cases} cx^4 ; & 0 < x < 1 \\ 0, & \text{elsewhere} \end{cases}$$

∴ $f_X^2(x) \geq 0$ for all values of $x$, if $c > 0$, since either it is zero or $x^4$ in the interval $[0, 1]$ which is positive.

Thus condition (i) satisfied if $c > 0$. Now, let us check condition (ii)

$$\int_{-\infty}^{+\infty} f_X^2(x)\,dx = \int_{-\infty}^{0} f_X^2(x)\,dx + \int_{0}^{1} f_X^2(x)\,dx + \int_{1}^{+\infty} f_X^2(x)\,dx$$

$$= 0 + \int_0^1 cx^4\,dx + 0$$

$$\therefore \int_{-\infty}^{+\infty} f_X^2(x)\,dx = \int_0^1 cx^4\,dx = \left[c\frac{x^5}{5}\right]_0^1$$

$$\therefore \int_{-\infty}^{+\infty} f_X^2(x)\,dx = \frac{c}{5}$$

∴ Hence cond$^n$ (ii) is satisfied if $c = 5$

∴ For $c = 5$, $f_X^2(x)$ is a density function.

For $f_X^3(x)$:

$$f_X^3(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & 0 \leq x \infty \\ 0, & \text{elsewhere} \end{cases}$$

$f_X^3(x) \geq 0$ $\forall x$, when $\theta > 0$, since exponential function is never negative. Hence, $f_X^3(x)$ takes zero or a positive value only.

Now check for cond$^n$ (ii)

$$\int_{-\infty}^{+\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx = \int_{-\infty}^{0} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx + \int_{0}^{\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx$$

$$\therefore \int_{-\infty}^{+\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx = 0 + \int_{0}^{+\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx$$

$$\therefore \int_{-\infty}^{+\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx = \frac{1}{\theta}\int_{0}^{+\infty} e^{-\frac{x}{\theta}} dx = \left[-e^{-\frac{x}{\theta}}\right]_0^{\infty}$$

$$= -[0-(1)]$$

$$\therefore \boxed{\int_{-\infty}^{+\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} dx = 1}$$

Hence it is density for $\theta > 0$.

$\therefore f_X^3(x)$ is density as long as $\theta > 0$.

- Since density is associated with exponential function it is called exponential density.

~~$\frac{1}{\theta} < 0$~~

- If $\theta < 0$, then $\frac{1}{\theta} < 0$ even though exponential function remains positive $\Rightarrow$ The cond$^n$ (1) is violated.
- If $\theta < 0$, the exponent $-\frac{x}{\theta} > 0$, thereby integral from 0 ~~to~~ $\infty$ will be $\infty$. Thus condition (ii) will also ~~not~~ violated.

$\Rightarrow$ for $\theta < 0$ the $f_X^3(x)$ cannot be density

- When integration from 0 to $\infty$ the ~~ep~~ exponential function with positive exponent can not create density

Consider $f_X^4(x)$:

$\therefore$ $f_X^4(x)$ is zero or $x$ in $[0,1)$

and $2-x$ in $[1,2]$

Hence $f_X^4(x) \geq 0$ for all $x$ and condition ① ~~satisfy~~ satisfied.

$\therefore$ Now integral $\int_{-\infty}^{+\infty} f_X^4(x)\,dx$

$$= \int_{-\infty}^{0} f_X^4(x)\,dx + \int_0^1 f_X^4(x)\,dx + \int_1^2 f_X^4(x)\,dx + \int_2^{\infty} f_X^4(x)\,dx$$

$$= 0 + \int_0^1 x\,dx + \int_1^2 (2-x)\,dx + 0$$

$$= \left[\frac{x^2}{2}\right]_0^1 + \left[2x - \frac{x^2}{2}\right]_1^2$$

$$= \left[\frac{1}{2} - 0\right] + \left[\left(2\times 2 - \frac{2^2}{2}\right) - \left(2\times 1 - \frac{1^2}{2}\right)\right]$$

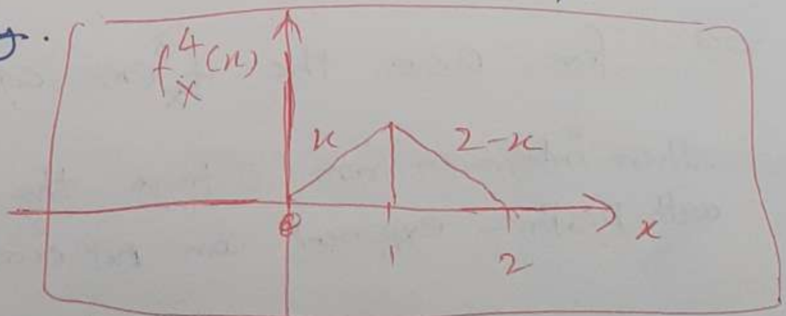
$$= \frac{1}{2} + \left[(4-2) - \left(2 - \frac{1}{2}\right)\right]$$

$$= \frac{1}{2} + \left[2 - \frac{3}{2}\right]$$

$$\therefore \int_{-\infty}^{+\infty} f_X^4(x)\,dx = \frac{1}{2} + \frac{1}{2}$$

$$= 1$$

$\therefore$ cond$^n$ (ii) satisfied. therefore $f_X^4(x)$ is valide density.




→ This is triangular density

## Definition (Parameter):–

Arbitrary constants sitting in a density or probability function are called parameters.

() uniform density a, and b are parameter

L exponential density $\theta$ is parameter.

## Defn (Normalizing constant):– If a constant sitting in a function is such that for specific value of this constant the function becomes a density or probability function then that constant is called the normalizing constant

e.g.

$$f_X^2(x) = \begin{cases} c\,x^4, & 0<x<1 \\ 0; & \text{otherwise} \end{cases}$$

In this c is a constant, but for $c=5$, $f_X^2(x)$ becomes density. This 'c' is called normalizing constant here.

## Definition (Degenerate Random Variable):– If whole probability mass is concentrated at one point then the r.v. is called a degenerate random variable or a mathematical variable.

e.

$$f_X(x) = \begin{cases} 1, & \text{if } x=b \\ 0, & \text{otherwise.} \end{cases}$$

Example :- Evaluate the distribution function for the following density

$$f_X^1(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & 0 \le x < +\infty \\ 0, & \text{otherwise} \end{cases}$$

$$f_X^2(x) = \begin{cases} x; & 0 \le x < 1 \\ 2-x; & 1 \le x < 2 \\ 0, & \text{otherwise} \end{cases}$$

Solution :- (i) The distribution function, by definition in the continuous case is:

$$F_X(t) = \int_{-\infty}^{t} f_X(x)\, dx$$

$$\Rightarrow \int_{-\infty}^{t} f_X^1(x)\, dx = \int_{-\infty}^{0} f_X^1(x)\, dx + \int_{0}^{t} f_X^1(x)\, dx$$

$$= 0 + \int_0^t \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dt$$

$$= \left[-e^{-\frac{x}{\theta}}\right]_0^t$$

$$\int_{-\infty}^{t} f_X^1(x)\, dx = 1 - e^{-\frac{t}{\theta}}, \quad 0 \le t < +\infty$$

$$= 0, \quad -\infty < t < 0$$

$$\therefore \boxed{F_X^1(t) = \begin{cases} 1 - e^{-t/\theta}, & 0 \le t < +\infty \\ 0, & -\infty < t < 0 \end{cases}}$$

(ii) $\int f_X^2(x)$ has to integrate in different pieces.

$$F_X^2(t) = 0, \quad \text{for } -\infty < t < 0$$

When $t \in (0, 1)$

then $\int_0^t f_X^2(x)\,dx = \left[\frac{x^2}{2}\right]_0^t = \frac{t^2}{2}$

$\therefore \quad F_X^2(t) = \frac{t^2}{2}, \quad 0<t<1$

When '$t$' is in interval $[1, 2]$ the integral up to 1, available from $\frac{t^2}{2}$ at $t=1$ which is $\frac{1}{2}$, plus the integral of the function $(2-x)$ from 1 to $t$ is to ~~be~~ be computed. That is

$$\frac{1}{2} + \int_1^t (2-x)\,dx = \frac{1}{2} + \left[2x - \frac{x^2}{2}\right]_1^t$$

$$= -1 + 2t - \frac{t^2}{2}$$

When $t$ is above 2 the total integral is one. Hence we have

$$F_X^2(t) = \begin{cases} 0, & -\infty<t<0 \\ \frac{t^2}{2}, & 0\leq t<1 \\ -1+2t-\frac{t^2}{2}, & 1\leq t<2 \\ 1, & t\geq 2 \end{cases}$$

## Mixed Random Variable :

→ R.V. where part of the probability mass is distributed on some individually distinct points (discrete case). but the remaining probability is distributed over a continuum (continuous case) points. Such random variables are called mixed cases.

**Ex:-** Compute the distribution function for the following probability function for a mixed case:

$$f_X(x) = \begin{cases} \frac{1}{2}, & x = -2 \\ x; & 0 \le x < 1 \\ 0, & \text{otherwise} \end{cases}$$