PROBABILITY

what we will study

→ Basic definitions of prob. (axioms)

→ sample space, events

→ Baye's theorem

→ Random variables

→ prob. distributions

what we'll not study

combinatorics type of prob. is in discrete maths

↓

Permutation & comb.

PROBABILITY ?

what does it mean to say –

Probability of appearing head on toss of fair coin is 1/2.

does this mean, if you toss 2 times, you'll get 1 time head. No
4 times — 2 times head. No
100 times — 50 times head. No

However, if you toss the same coin 4000 times, the outcomes will be close to half heads and half tails.

English statistician Karl Pearson who tossed a coin 24,000 times with a result of 12,012 heads, one of the authors tossed a coin 2,000 times. The results were 996 heads.

→ Probability tells how likely an event is to occur. enough no. of times ≈ equally likely

$$\lim_{n\to\infty} \frac{n(H)}{n} = \frac{1}{2}$$

n(H) → no. of times head; n → total tosses

50% chances ⇒ head

50% chances ⇒ tails

ques. Toss a coin 100 times, is it possible to get 100 heads : $\underbrace{H, H, H, \dots H}_{100\ \text{times}}$

↳ yes it is possible but the prob. is very low, $\left(\frac{1}{2}\right)^{100}$

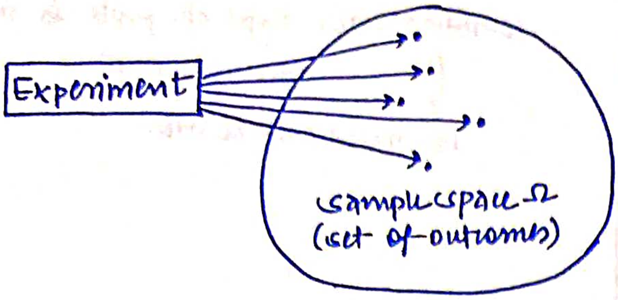
ques. Is it possible :

100 times tossed → 40 times head
→ 60 times tail

↳ yes it is also possible

Sample Space

- Every probabilistic model involves an underlying process, called the experiment, that will produce exactly one out of several possible outcomes.
- The set of all possible outcomes is called the sample space of the experiment.

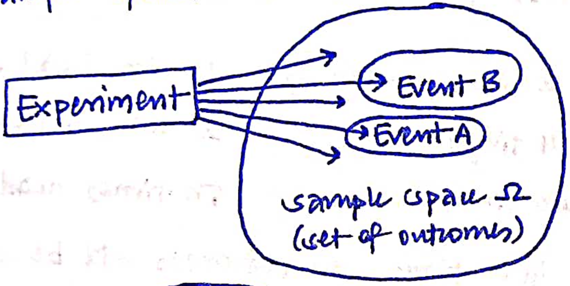


ex Tossing a coin → {H, T}

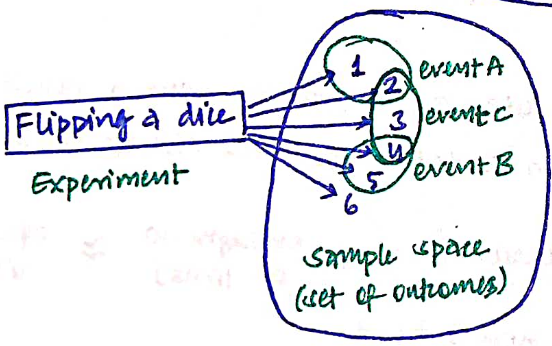
$\underbrace{\text{Tossing a coin}}_{\text{experiment}} \rightarrow \{H, T\}$ ↖ All possible outcomes.

Event

A subset of the sample space, that is, a collection of possible outcomes, is called an event.



ex



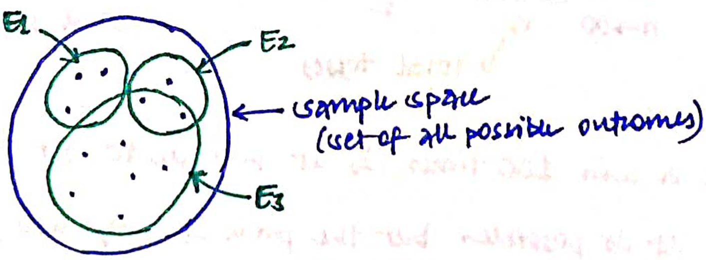
event A : {1, 2}
event B : {4, 5}
event C : {2, 3, 4}

o→ $E_1 = \{1, 2\}$; $P(E_1) = 2/6$

$E_2 = \{3, 4, 5\}$; $P(E_2) = 3/6$

$E_3 = \{5\}$; $P(E_3) = 1/6$

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

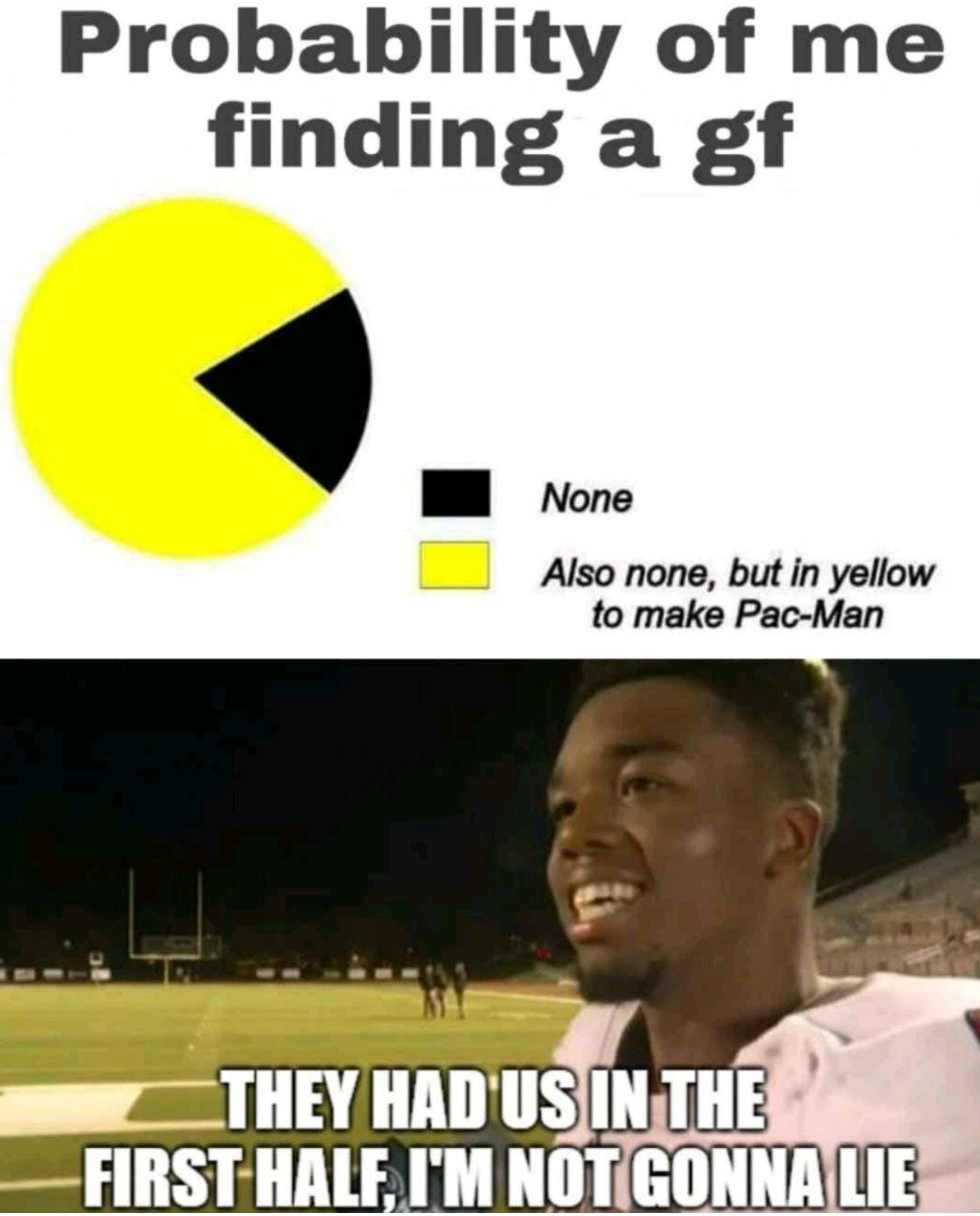


event is subset of sample space
E Ω

∴ total no. of Events possible

= no. of subsets of Ω = 2^{Ω}

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

(ex) Experiment : Flipping a coin

Sample space, $\Omega = \{H, T\}$

All possible events $= 2^{\Omega}$

$$\{\{H\}, \{T\}, \{H,T\}, \{\}\}$$

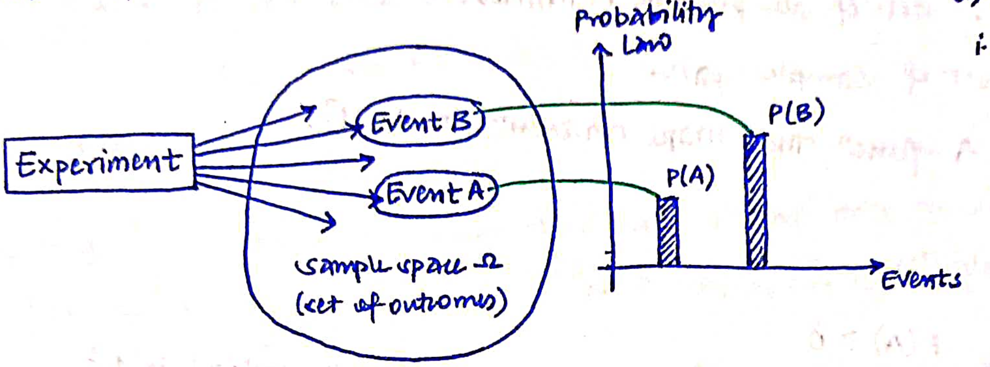
$$P(H) = 1/2 \quad P(T) = 1/2 \quad 1 \quad 0$$

Probability

Func^n that maps an event to $[0,1]$. → i.e. $P : 2^{\Omega} \to [0,1]$

→ prob. is defined over event.

i.e. $P(\{\})$ ← event



• The main ingredients of a Probabilistic model.

Properties of Sample Space

- must have "list" (set) of possible outcomes.
- list must be :
  - Mutually exclusive
  - Collectively Exhaustive

→ Sample space for tossing a coin : $\{H, T\}$

Sample space for tossing 2 coins (or tossing coin 2 times) : $\{HH, HT, TH, TT\}$

Sample space for rolling a dice : Ω (or S) $= \{1, 2, 3, 4, 5, 6\}$

→ if the outcome of an experiment consists in the determination of the sex of a newborn child, then $S = \{g, b\}$ where the outcome g means the child is a girl. b " boy.

→ if the outcome of an experiment is the order of finishing in a race among the 7 horses having post positions 1, 2, 3, 4, 5, 6 and 7, then

$$S = \{\text{all } 7! \text{ permutations of } (1,2,3,4,5,6,7)\}$$

The outcome (2, 3, 1, 6, 5, 4, 7) means, for instance, that the number 2 horse comes in first, then the no 3 horse, then the no 1 horse, and so on.

→ sample space depends on what we are interested in.

(ex) sample space of tossing coin 2 times : $\{HH, HT, TH, TT\}$

but if we are only interested in what comes last ; then $S = \{H, T\}$

(ex) if we are interested in order of all horses : $7!$ permutations

knowing who has came first ⇒ $S = \{1,2,3,4,5,6,7\}$

knowing 1st two positions ⇒ $S = \{12, 13, 14, 15, 16, 17, 23, 24, 25, 26, 27, \ldots\}$

7×6 possible outcomes.

$\boxed{\text{NOTE}}$ ↦

→ Sample Space : set of all possible outcomes.

→ Event : subset of sample space.

→ Probability : A func^n that maps an event to $[0,1]$.

Axioms of Probability

1. Nonnegativity : $P(A) \geq 0$
2. Normalization : $P(\Omega) = 1$ {Probability of all the outcomes combined is 1}
3. Additivity : If $A \cap B = \phi$, then $P(A \cup B) = P(A) + P(B)$

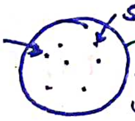
→ S (or Ω) $= \{\underbrace{E_1, E_2, E_3}_{\text{mutually exclusive events}}\}$; $P(\Omega) = P(E_1) + P(E_2) + P(E_3) = 1$

→ $P(\{H\} \text{ or } \{T\}) = P(\{H\}) + P(\{T\})$ or ↓ $\cup$ ← union

$P(H \cup T) = P(H) + P(T)$

H&T are disjoint

→ $S = \{s_1, s_2, s_3, \ldots, s_n\}$



disjoint (mutually exclusive)

set of all possible outcomes (exhaustive)

$P(\{s_1, s_2, \ldots, s_n\}) = P(s_1) + P(s_2) + \ldots + P(s_n)$

4. Countable Additivity Axiom : (extended 3rd Axiom) if no. of events are infinite

If $A_1, A_2, \ldots$ are disjoint events, then

$$P(A_1 \cup A_2 \cup \ldots) = P(A_1) + P(A_2) + \ldots$$

→ let, all outcomes be equally likely, then

$$P(A) = \frac{\text{no. of elements of } A}{\text{total no. of sample points}}$$

$\hookrightarrow$ $S = \{s_1, s_2, s_3, \dots, s_n\}$ $\qquad$ $P(\{s_1\}) = P(s_1)$ ← abuse of notation

$P(s_1) = P(s_2) = \dots = P(s_n)$ (equally likely outcomes)

$\hookrightarrow$ $P(\{s_1\} \cup \{s_2\} \cup \{s_3\} \cup \dots \{s_n\}) = 1$ (using 2nd Axiom)

$\Rightarrow P(s_1) + P(s_2) + \dots + P(s_n) = 1$

$\Rightarrow P(s_1) = \frac{1}{n}$

Now, if $A = \{s_1, s_2, \dots, s_k\}$
↑ an event

$\Rightarrow P(A) = P(s_1) + P(s_2) + \dots + P(s_k)$

$\qquad\qquad = 1/n + 1/n + \dots + 1/n$

$P(A) = \frac{k}{n}$ $\begin{array}{l} \rightarrow \text{no. of elements in } A \\ \rightarrow \text{total elements} \end{array}$

→ if everything is not equally likely then this doesn't hold true

ex $\{s_1, s_2, s_3, s_4, s_5, s_6\}$

$\quad 1/6 \quad 1/6 \quad 2/6 \quad 1/12 \quad 1/6 \quad 1/3$

Now, ~~$P(s_1, s_2, s_3) = \frac{3}{6} = \frac{1}{2}$~~ ← no. of elements

$P(s_1, s_2, s_3) = \frac{1}{6} + \frac{1}{6} + \frac{2}{6} = \frac{4}{6} = \frac{2}{3}$

ques. If 2 dice are rolled, what is the probability that the sum of the upturned faces will equal 7 ?

$\hookrightarrow$ sum = 7

Possible outcomes : (1,1) (1,2) (1,3) (1,4) (1,5) (1,6)
(2,1) (2,2) . . . .
36 outcomes

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

$\Rightarrow P(A) = \frac{6 \leftarrow n(A)}{36 \leftarrow n} = \frac{1}{6}$

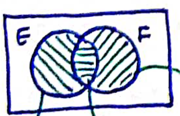
$P(A) = \frac{n(A)}{n}$ ← when every outcome is equally likely

The Inclusion-Exclusion Principle

For any 2 events, E and F :

$P(E \cup F) = P(E) + P(F) - P(E \cap F)$

proof



$c = F \cap \bar{E}$

$a = E \cap \bar{F}$ $\qquad$ $b = E \cap F$

now, $P(E \cup F) = P(a \cup b \cup c) = LHS = P(a) + P(b) + P(c)$

$P(E) = P(a \cup b) \qquad P(F) = P(b \cup c) \qquad P(E \cap F) = P(b)$

RHS : $P(a \cup b) + P(b \cup c) - P(b)$

$= P(a) + \cancel{P(b)} + P(b) + P(c) - \cancel{P(b)} = LHS$ ✓

Hence Proved.

→ For 3 events E, F & G, that formula mathematically expands to,

$P(E \cup F \cup G) = P(E) + P(F) + P(G)$
$\qquad - P(E \cap F) - P(E \cap G) - P(G \cap F)$
$\qquad + P(E \cap F \cap G)$

De Morgan's Law for Probability

For any 2 events E and F :

$$P(\overline{E \cap F}) = P(\bar{E} \cup \bar{F})$$

OR

$$P(\overline{E \cup F}) = P(\bar{E} \cap \bar{F})$$

→ $P(E \cap F) = 1 - P(\overline{E \cap F})$ $\quad (P(A) + P(\bar{A}) = 1)$

$= 1 - P(\bar{E} \cup \bar{F})$ $\quad$ (Demorgan's law)

→ $P(E \cup F) = 1 - P(\overline{E \cup F})$ $\quad (P(A) + P(\bar{A}) = 1)$

$= 1 - P(\bar{E} \cap \bar{F})$ $\quad$ (Demorgans law)

GATE 1997
ques. The prob. that it will rain today is 0.5, And the prob that it will rain tomorrow is 0.6. The prob that it will rain either today or tomorrow is 0.7. What is the prob that it will rain Today and tomorrow?

↳ $P(\text{today} \cup \text{tomorrow}) = P(\text{today}) + P(\text{tomorrow}) - P(\text{today} \cap \text{tomorrow})$

$0.7 = 0.5 + 0.6 - P(\text{today} \cap \text{tomorrow})$

$P(\text{today} \cap \text{tomorrow}) = 0.6 + 0.5 - 0.7 = 0.4$ ✓

GATE 08
ques. A sample space has 2 events, A and B such that probabilities

$P(A \cap B) = \frac{1}{2}$ $\quad P(\bar{A}) = \frac{1}{3}$ $\quad P(\bar{B}) = \frac{1}{3}$. $\quad P(A \cup B) = ?$

↳ $P(A) = 1 - P(\bar{A}) = 1 - \frac{1}{3} = \frac{2}{3}$ $\quad P(B) = 1 - P(\bar{B}) = 1 - \frac{1}{3} = \frac{2}{3}$

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

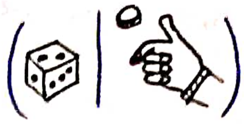
$= \frac{2}{3} + \frac{2}{3} - \frac{1}{2} = \frac{5}{6}$ ✓

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Conditional Probability

Probability of an event based on some partial information.

"change in Belief"

P 

if the tossing of coin has happened, then what is the prob. of occuring 2 on the dice

ex INDIA v/s Australia

(assume fairplay conditions & equally good teams)

Before start of play : What is the chance of India winning? = 0.5

India scores 395 batting first : what is the chance of India winning? > 0.5

- what exactly happenned here?

A : event that India will win $P(A) = 0.5$

B : India scores 395 runs.

P(A) changes once we know that event B has occurred $P(A|B) > 0.5$

$$P(A|B) \neq P(A)$$

NOTE → $P(A|B) > P(A)$, $<$, $=$ } there is no relation b/w $P(A)$ & $P(A|B)$

anything can happen

Example of conditioning

Dice tossing : We consider the following situation

- we throw 2 dice.
- We look for P (sum of 2 faces is 9)

(3,6), (6,3)
(4,5), (5,4)

- Without prior info : $P(\text{sum of 2 faces is } 9) = \frac{4}{36} = \frac{1}{9}$

- with additional info :

If first face is = 4. Then

- Only 6 possible results : (4,1), (4,2), (4,3), (4,5), (4,6)
- Among them, only (4,5) give sum = 9
- Prob. of having sum = 9 becomes $\frac{1}{6}$

Infected by covid : Suppose 1% of population has covid
↳ $\frac{1}{100}$

$\therefore$ $P(\text{someone is having covid}) = \frac{1}{100}$ ← without prior info

$P(\text{person is having covid} \mid \text{Person is having symptoms of covid}) \approx 1$ ← with prior info ques. A die has been rolled where additional info is given that the experiment has produced an even face.

what is the prob. ~~of~~ that 2 occurred ?

$P(\text{occuring } 2) \rightarrow \frac{1}{6}$

$P(\text{occuring } 2 \mid \text{even face}) = \frac{1}{3}$

$S = \{2, 4, 6\}$

		S_1	S_2	S_3	S_4	S_5 (occuring 5)	S_6 (occuring 6)
~~occuring 2~~ die rolled	:	1/6	1/6	1/6	1/6	1/6	1/6
even face has occured	:	0	1/3	0	1/3	0	1/3

} b/oz the prob. masses has been redistributed

$$\frac{\frac{1}{6}+\frac{1}{6}+\frac{1}{6}}{1/2} = \frac{1/2}{1/2}$$

$$2\times\frac{1}{6} + 2\times\frac{1}{6} + 2\times\frac{1}{6} = 1$$

old sample space

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

$P(2) = \frac{1}{6}$

New sample space

$S' = \{2, 4, 6\}$

$P(2) = \frac{1}{3}$

$P(2 \mid \text{even face}) = \frac{1}{3}$

ex) ~~&~~ We have the sample space as:

	S_1	S_2	S_3	S_4	S_5	S_6
Initial prob. distribution	1/4	1/8	1/4	1/8	1/8	1/8
Now, additional info is that, S_2, S_4 & S_6 has not occurred	?	0	?	0	?	0
	$\frac{1}{4}\times\frac{8}{5}$		$\frac{1}{4}\times\frac{8}{5}$		$\frac{1}{8}\times\frac{8}{5}$	
	$\frac{2}{5}$		$\frac{2}{5}$		$\frac{1}{5}$	

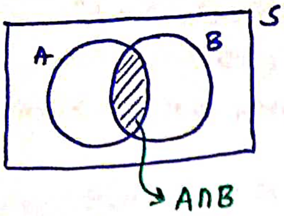
← We have to redistribute the prob distribution

$$\frac{\frac{1}{4}+\frac{1}{4}+\frac{1}{8}}{5/8} = \frac{5/8}{5/8} = 1$$

Now, therefore $P(S_1 \mid \{S_1, S_3, S_5\}) = \frac{2}{5}$

$P(S_2 \mid \{S_1, S_3, S_5\}) = 0$

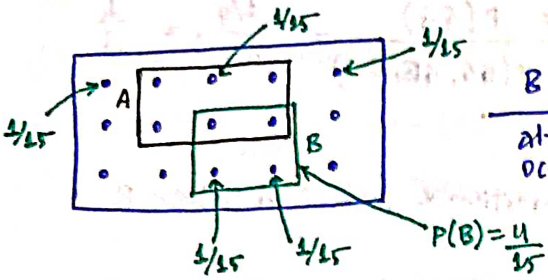
Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



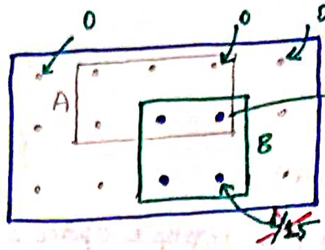
$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

prob. that A has occurred given that B has already occurred.

lets understand this intutively,



B has already occurred →



or, intutively prob. of happening of one outcome out of 4 equally possible outcomes $= \frac{1}{4}$

$\frac{1}{15} \times \frac{15}{4} = \frac{1}{4} \Rightarrow$ correct ✓

every outcome in the sample space is equally possible

we always work here

outcomes when event B doesnot happen is not possible

we dont need this new sample space. This is just for intution.

$$\frac{\frac{1}{15}+\frac{1}{15}+\frac{1}{15}+\frac{1}{15}}{\underbrace{4/15}_{P(B)}} = \frac{4/15}{4/15} = 1$$

Formula :

$$P(A|B) = \frac{P(A\cap B)}{P(B)} = \frac{2/15}{4/15} = \frac{1}{2}$$

Using intution:

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

- Using intution is only for understanding we wont use it to solve questions as it is a bit confusing sometimes. → we the prob mass distribution is not equally among outcomes.

  Always use formula.

ques. A die has been rolled where additional info is given that the experiment has produced no. less than 4.
What is the prob that 2 occurred ?

↳ $\{1,2,3\} \leftarrow$ given

Intutively, $P(2|\{1,2,3\}) = \frac{1}{3}$

Formula : $P(2|\{1,2,3\}) = \dfrac{P(\{2\}\cap\{1,2,3\})}{P(\{1,2,3\})} = \dfrac{P(\{2\})}{P(\{1,2,3\})} = \dfrac{1/6}{1/2} = \dfrac{1}{3}$

ques. consider a family with 2 children. We are interested in the children's genders. Our sample space is $S = \{(G,G),(G,B),(B,G),(B,B)\}$. Assume all 4 possible outcomes are equally likely.

(a) What is the prob. that both children are girls given that the 1st is a girl ?

Intutively: $\{(G,G),(G,B)\}$ → $\frac{1}{2}$ ✓

Formulae : $P(\{GG\}|\{GG,GB\}) = \dfrac{P(GG)}{P(GG,GB)} = \dfrac{1/4}{1/4+1/4} = \dfrac{1/4}{1/2} = \dfrac{1}{2}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

(b) We ask the father : "Do you have atleast one daughter ?". He responds "yes!".
Given this extra info, what is the prob that both children are girls ?
In other words, what is the prob that both children are girls given that we know atleast one of them is girl ?

↳ $\{GG, GB, BG\}$

Intutively :→ $\frac{1}{3}$

Using formula : $P(GG \mid \{GG, GB, BG\})$

$$= \frac{P(GG)}{P(GG, GB, BG)} = \frac{1/4}{3/4} = \frac{1}{3}$$

Formal def of conditional prob :

If A and B are 2 events in a sample space S, then the conditional prob of A given B is defined as

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \text{, when } P(B) > 0 \rightarrow \frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(B)}{n(S)}} = \frac{n(A\cap B)}{n(B)}$$

$$P(A|B) = \frac{\text{no. of elements of } A\cap B}{\text{no. of elements of } B}$$ (when all outcomes are equally likely)

↠ conditional prob can also be viewed as a prob law on a new universe B, bcoz all of the conditional prob is concentrated on B.

Qun. A fair dice is tossed twice

E = { The sum of the face values of these 2 tosses is 5. } ,

(a) Calculate P(E).

(b) given that the event E happens, what is the prob that the face value of the first toss is less than that of the 2nd toss ?

↳ The sample space S, $S = \{(i,j) : 1 \le i \le 6,\ 1 \le j \le 6\}$

in total = 6×6 = 36 elements.

(a) The event E is $E = \{(1,4), (2,3), (3,2), (4,1)\}$

and $P(E) = \frac{4}{36} = \frac{1}{9}$

(b) The conditional prob.

intutively, $S' = \{(1,4), (2,3), (3,2), (4,1)\}$

$\therefore$ Prob $= \frac{2}{4} = \frac{1}{2}$

using formulae,

$$P(\{\ \} \mid \{(1,4), (2,3), (3,2), (4,1)\})$$

$$= \frac{\{(1,4), (2,3)\}}{4/36}$$

$$= \frac{2/36}{4/36} = \frac{1}{2}$$ ✓

1,2	2,3	3,4	4,5	
1,3	2,4	3,5	4,6	
1,4	2,5	3,6		
1,5	2,6			
1,6				5,6

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques. Suppose we flip 3 fair coins. Let A stand for the event that 1^{st} 2 coins are both heads. Let B stands for the event that 3^{rd} coin is different from the 2^{nd} coin.

(a) Find P(A), P(B).

$$P(A) = P(HHT) + P(HHH) = \frac{2}{8} = \frac{1}{4}$$

$$P(B) = P(HHT) + P(HTH) + P(THT) + P(TTH) = \frac{4}{8} = \frac{1}{2}$$

(b) Find the conditional probability P(A|B).

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A \text{ and } B)}{P(B)} = \frac{1/8}{1/2} = \frac{1}{4}$$

$$P(A \text{ and } B) = P(HHT) = \frac{1}{8}$$

Multiplication Rule (or Product Rule)

(Directly follows from conditional probability definition)

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

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

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

$\rightarrow$ $P(A \cap B) = P(A,B) = P(AB)$ (notation)

$P(A,B) = P(A) \cdot P(A|B)$

$P(B,A) = P(B) \cdot P(B|A)$

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

(MSQ) ques. Which of the following is correct. $P(A,B,C) = ?$

(a) $P(C) \cdot P(B|C) \cdot P(A|B,C)$ ✓ $\rightarrow P(A,B,C)$

(b) $P(A) \cdot P(B|A) \cdot P(C|A,B)$ ✓ $\rightarrow P(C,B,A)$

(c) $P(B) \cdot P(A|B) \cdot P(C|A,B)$ ✓ $\rightarrow P(C,A,B)$

(d) $P(B) \cdot P(C|B) \cdot P(A|B,C)$ ✓ $\rightarrow P(A,C,B)$

All are correct, as $P(A,B,C) = P(C,B,A) = P(C,A,B),\ P(A,C,B)$ ✓

$\rightarrow$ $P(A,B,C,D) = P(D) \cdot P(C|D) \cdot P(B|C,D) \cdot P(A|B,C,D)$

TREE DIAGRAM

(for conditional probability)] → useful in sequential models

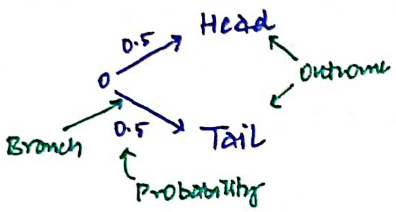
Many experiments have an inherently sequential character, such as for ex :→

(i) Tossing a coin 3 times.

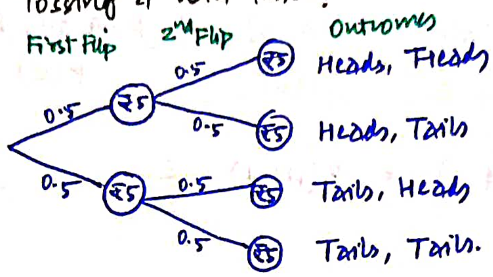
(ii) Observing the value of a stock on 5 successive days.

(iii) Receiving 8 successive digits at a communication receiver

•→ Tree diagram for coin toss



•→ Tossing a coin twice!



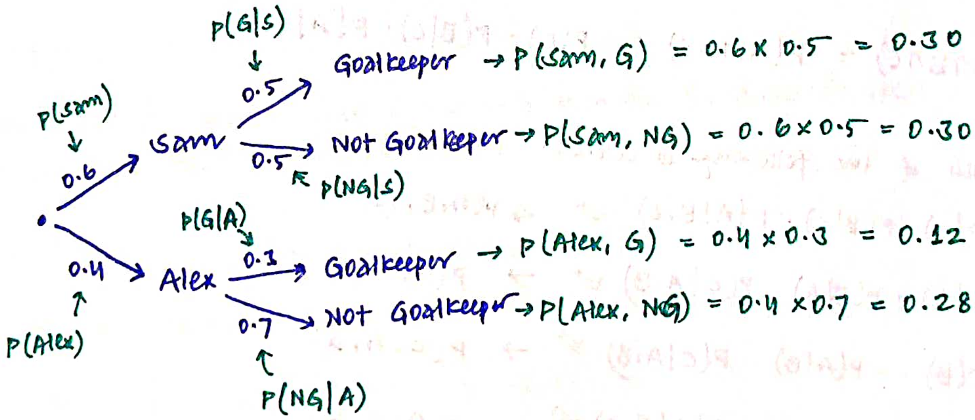
ques. You are off to soccer, & love being the Goalkeeper, but that depends who is the coach today :

- with coach Sam the prob of being Golkeeper is 0.5
- — " — Alex ———— " ———— 0.3

Probability of Sam being coach = 0.6.

So, whats the prob you'll be a goalkeeper today?

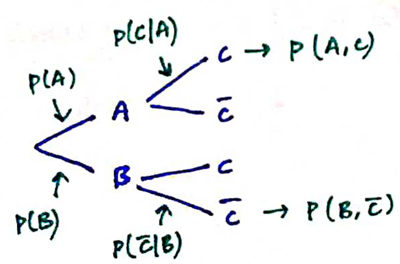
↳



$\therefore P(G) = P(Sam, G) + P(Alex, G)$

$= 0.30 + 0.12 = \boxed{0.42}$

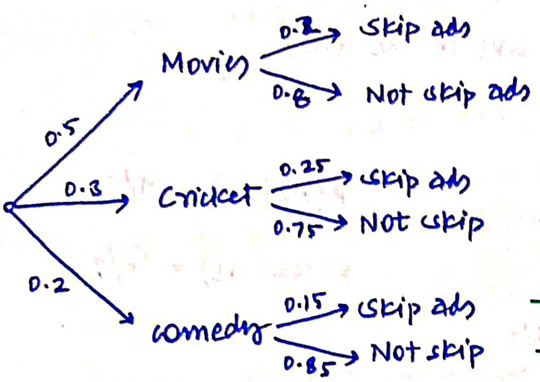
•→



$P(A,C) = P(A) \cdot P(C|A)$ ← so this tree diagram makes sense mathematically too

ques. On a sunday evening, 50% watch a movie, 30% watch a cricket match and 20% watch comedy show. The % of people skipping ads are 20%, 25% & 15% respectively.

P(watch cricket and see ads) = ?



→ P(movie, skip Ads) = 0.5 × 0.2 = 0.10

→ P(movie, not skip Ads) = 0.5 × 0.8 = 0.40

→ P(cricket, skip Ads) = 0.3 × 0.25 = 0.075

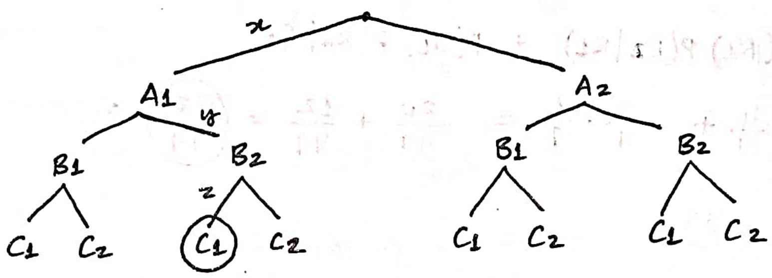
→ P(cricket, not skip Ads) = 0.3 × 0.75 = 0.225 ✓

→ P(comedy, skip Ads) = 0.2 × 0.15 = 0.03

→ P(comedy, not skip Ads) = 0.2 × 0.85 = 0.17

P(watch cricket and see ads) = 0.225 ✓

Conceptual ques.



(i) The prob 'x' represents

(a) $P(A_1)$ ✓

(b) $P(A_1 | B_2)$

(c) $P(B_2 | A_1)$

(d) $P(C_1 | B_2 \cap A_1)$

(ii) The prob 'y' represents

(a) $P(B_2)$

(b) $P(A_1 | B_2)$

(c) $P(B_2 | A_1)$ ✓

(d) $P(C_1 | B_2 \cap A_1)$

(iii) The prob 'z' represents

(a) $P(C_1)$

(b) $P(B_2 | C_1)$

(c) $P(C_1 | B_2)$

(d) $P(C_1 | B_2 \cap A_1)$ ✓

(iv) The circled node represents the event

(a) C_1

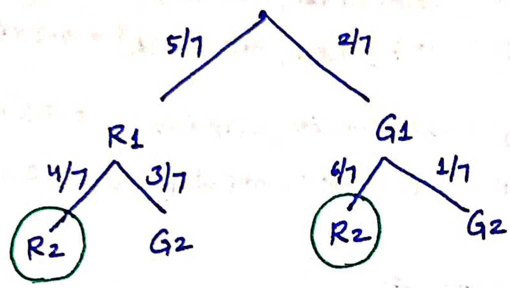
(b) $B_2 \cap C_1$

(c) $A_1 \cap B_2 \cap C_1$ ✓

(d) $C_1 | B_2 \cap A_1$.

ques. An urn contains 5 red balls and 2 green balls. A ball is drawn. If it's green, a red ball is added to the urn & if its red a green ball is added to the urn. (The original ball is not returned to the urn). Then a second ball is drawn. What is the prob that the 2nd ball is red?

↳ The sequence of the actions are:
first draw ball 1 (and add appropriate ball to the urn) and then draw ball 2.



$P(R_2|R_1) = 4/7$

$P(R_2|G_1) = 6/7$

$$\therefore \quad P(R_2) = P(R_2, R_1) + P(R_2, G_1)$$

$$= P(R_1)\, P(R_2|R_1) + P(G_1)\, P(R_2|G_1)$$

$$= \frac{5}{7} \cdot \frac{4}{7} + \frac{2}{7} \cdot \frac{6}{7} = \frac{20}{49} + \frac{12}{49} = \frac{32}{49} \checkmark$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Conditional Probabilities satisfy the 3 Axioms :

(i) Non Negative : $P(A) \geq 0$ $\quad P(A|B) \geq 0 \rightarrow \frac{P(A \cap B)}{P(B)} \geq 0.$

conditional prob is always non-ve.

(ii) Normalization : earlier we used to say that prob of sample space = 1.

i.e. $P(\Omega) = 1$.

prob of sample space given some event B.

Now, it will be like $P(\Omega|B) = 1$ $\rightarrow P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$

(iii) Additivity : if A_1 and A_2 are disjoint then,

earlier : $P(A_1 \cup A_2) = P(A_1) + P(A_2)$

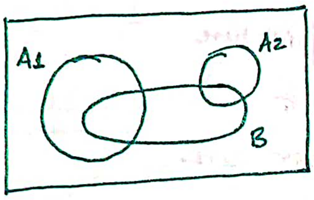
now : $P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B)$

proof : $P(A_1 \cup A_2 | B) = \frac{P((A_1 \cup A_2) \cap B)}{P(B)}$ distributive

$= \frac{P((A_1 \cap B) \cup (A_2 \cap B))}{P(B)}$ disjoint sets

$= \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)}$

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



Axioms of Probability

1. Non negativity :
$P(A) \geq 0$

2. Normalization :
$P(\Omega) = 1$

3. Additivity :
$P(A_1 \cup A_2) = P(A_1) + P(A_2)$

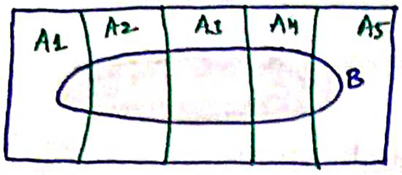
Axioms of conditional probability

1. Non negativity :
$P(A|B) \geq 0$

2. Normalization :
$P(\Omega|B) = 1$

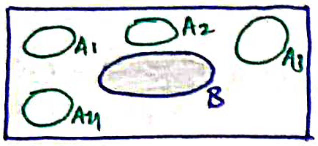
3. Additivity : A_1, A_2 are disjoint sets.
$P(A_1 \cup A_2 | B) = P(A_1|B) + P(A_2|B)$

Divide & conquer



$P(B) = P((B\cap A_1) \cup (B\cap A_2) \cup (B\cap A_3) \cup (B\cap A_4) \cup (B\cap A_5))$

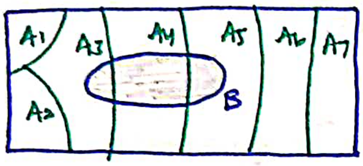
o→



⇝ $P(B) = P((B\cap A_1) \cup (B\cap A_2) \cup (B\cap A_3) \cup (B\cap A_4))$ ✗

we cannot write like this in this case

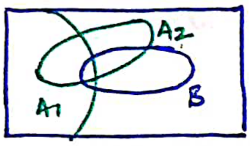
$\boxed{\text{NOTE}}$ → To write this, we should have all the A_i's as mutually exhaustive



⇝ $A_1 \cup A_2 \cup \dots A_n = \Omega$ & all are disjoint

then $B = (B\cap A_1) \cup (B\cap A_2) \cup \dots\dots (B\cap A_n)$ ✓

o→



⇝ $(B\cap A_1)$ & $(B\cap A_2)$ are not disjoint here

bcz A_1 & A_2 are not disjoint.

so can't use divide & conquor here

$\boxed{\text{NOTE}}$ → if A_1, A_2, A_3 are making partition then,

for any event B :

{ 1. A_1, A_2, A_3 are mutually exhaustive
2. A_1, A_2, A_3 are mutually disjoint. }

$$P(B) = P((B\cap A_1) \cup (B\cap A_2) \cup (B\cap A_3))$$

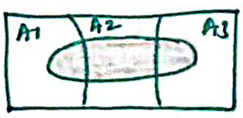
$$= P(B\cap A_1) + P(B\cap A_2) + P(B\cap A_3)$$

Marginalization (also called Total Probability)

If A_1, A_2 and A_3 partition the sample space then for any event B :→

$$P(B) = P(B\cap A_1) + P(B\cap A_2) + P(B\cap A_3)$$

Very easy to prove using venn diagram :.



⇝ $P(B) = P((B\cap A_1) \cup (B\cap A_2) \cup (B\cap A_3))$

Rules in Probability

- If outcomes are equally likely,

$$P(A) = \frac{\text{no. of elements in } A}{\text{total no. of sample points}}$$

- Inclusion Exclusion principle :→

for any 2 events A, B :

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

- De-Morgan's Law :→

for any 2 events A, B :

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

- $P(A) + P(\bar{A}) = 1$

- Total Probability :→

If A_1, A_2, A_3 partition the sample space then for any even B :

$$P(B) = P(B\cap A_1) + P(B\cap A_2) + P(B\cap A_3)$$

- Bayes' Theorem :→

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

$$= \frac{P(A\cap B)}{P(B\cap A_1) + P(B\cap A_2) + P(B\cap A_3)}$$

Rules in conditional Probability

$$P(A|B) = \frac{\text{no. of element of } A\cap B}{\text{no. of elements of } B}$$

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

$$P(\overline{(A\cup B)}|E) = P(\bar{A}\cap\bar{B}|E)$$

$$P(A|B) + P(\bar{A}|B) = 1$$

$$P(B|A) = P(B\cap A_1|A) + P(B\cap A_2|A) + P(B\cap A_3|A)$$

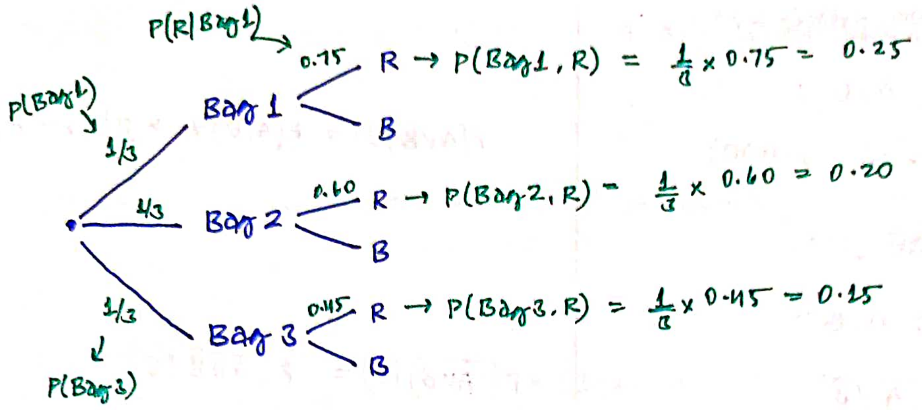
$$P(A|B,E) = \frac{P(A\cap B|E)}{P(B|E)}$$

$$= \frac{P(A\cap B|E)}{P(B\cap A_1|E) + P(B\cap A_2|E) + P(B\cap A_3|E)}$$

Qus. I have 3 bags that each contain 100 marbles :

- Bag 1 has 75 red and 25 blue marbles.
- Bag 2 60 40
- Bag 3 45 55

I chose one of the bag at random & then pick a marble from the chosen bag, also at random. What is the prob. that the chosen marble is red ?

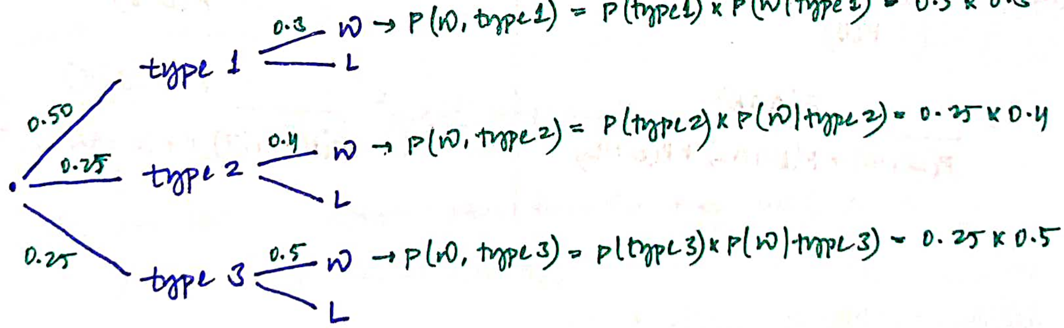


$$P(R) = 0.25 + 0.20 + 0.15 = 0.60$$

$P(R|Bag1) = 0.25$

Qus. ↳ in baye's thm we'll learn to answer $P(Bag1|R)$ → The picked marble is red, what is the prob it is from Bag 1

Qus. You enter a chess tournament where your probability of winning a game is 0.3 against half the players (call them type 1), 0.4 against a quarter of the players (call them type 2), and 0.5 against the remaining quarter of the players (call them type 3). You play a game against a random chosen opponent. What is the prob of winning?



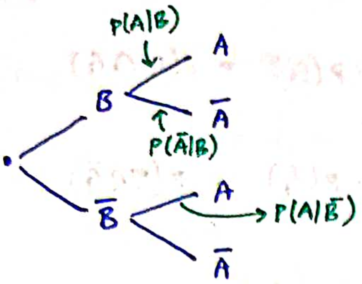
$$P(W) = 0.5 \times 0.3 + 0.25 \times 0.4 + 0.25 \times 0.5$$

$$= 0.375$$

True/ False

(i) $P(A|B) + P(\bar{A}|B) = 1$ True

(ii) $P(A|B) + P(A|\bar{B}) = 1$ False



$\Big\}$ $P(A|B) + P(\bar{A}|B) = 1$

$P(A|B)$ & $P(A|\bar{B})$ are of complete different branches of different parents.
They are not related, hence $P(A|B) + P(A|\bar{B}) = 1$ is false

(iii) $P(A|\bar{B}) + P(\bar{A}|B) = 1$ False
for the similar reason

ques. Which of the following is true?

(a) $P(\bar{B}|A) = 1 - P(B|A)$

$P(B) + P(\bar{B}) = 1$ (in prob)

II^{ly}, $P(B|A) + P(\bar{B}|A) = 1$ (in conditional prob)

$\Rightarrow$ TRUE

(b) $P(\bar{B}|A) = P(B) - P(B|A)$

$P(\bar{B}|A) + P(B|A) = 1$

$\Rightarrow P(B) = 1$ This need not be True

$\Rightarrow$ FALSE

(c) $P(E,F) + P(\bar{E} \cup \bar{F}) = 1$

$P(E \cap F) + P(\bar{E} \cup \bar{F})$

Now, $P(E \cap F) = 1 - P(\overline{E \cap F})$ demorgan's law
$= 1 - P(\bar{E} \cup \bar{F})$

$\Rightarrow$ TRUE

(d) $P(E \cup F) = 1 - P(\bar{E}|\bar{F})\, P(\bar{F})$

$1 - P(\overline{E \cup F})$

$\Rightarrow 1 - P(\bar{E}|\bar{F}) \cdot P(\bar{F})$ TRUE

bcoz, $P(\overline{E \cup F}) = P(\bar{E} \cap \bar{F})$ factorization
$= P(\bar{F}) \cdot P(\bar{E}|\bar{F})$

(e) $P(E,F|E) = P(E,F|F)$

$P(E \cap F|E)$

$= \dfrac{P(E \cap F \cap E)}{P(E)}$

$= \dfrac{P(E \cap F)}{P(E)}$

$P(E \cap F|F)$

$= \dfrac{P(E \cap F \cap F)}{P(F)}$

$= \dfrac{P(E \cap F)}{F}$

LHS $\neq$ RHS

FALSE

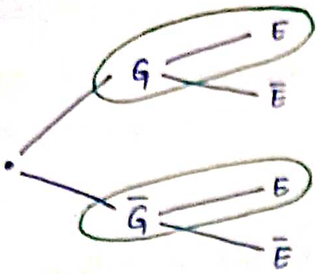
(f) $P(E|F)\,P(F) + P(\bar{E}|F)\,P(F) = P(F)$

$P(E,F) + P(\bar{E},F)$

$= P(F)$

TRUE

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



→ Total probability for E to happen,

$$P(E) = P(G \cap E) + P(\bar{G} \cap E)$$

another intuition to this is,

$$P(E) = P(E|G) \cdot P(G) = P(E \cap G)$$
$$+$$
$$P(E|\bar{G}) \cdot P(\bar{G}) = P(E \cap \bar{G})$$

ques (T/F)

$$P(E|F) = P(E|F,G)\, P(G|F) + P(E|F,\bar{G})\, P(\bar{G}|F)$$

↳ in RHS, there is an extra variable 'G', so we can use concept of total prob for it.

⇒ in normal prob: $P(E) = P(E \cap G) + P(E \cap \bar{G})$

in conditional prob: $P(E|F) = P(E \cap G|F) + P(E \cap \bar{G}|F)$

↳ in normal prob: $P(E) = P(E \cap G) + P(E \cap \bar{G})$ — factorization

$= P(G) \cdot P(E|G) + P(\bar{G}) \cdot P(E|\bar{G})$ — using 'given F' both the sides

||ly in conditional prob: $P(E|F) = P(G|F) \cdot P(E|G,F) + P(\bar{G}|F) \cdot P(E|\bar{G},F)$

Hence proved

TRUE

OR

Method 2:

$$P(E|F) = P(E|F,G)\, P(G|F) + P(E|F,\bar{G})\, P(\bar{G}|F)$$

LHS: $\dfrac{P(E \cap F)}{P(F)}$

RHS: $\dfrac{P(E,F,G)}{\cancel{P(F,G)}} \cdot \dfrac{\cancel{P(F,G)}}{P(F)} + \dfrac{P(E,F,\bar{G})}{\cancel{P(F,\bar{G})}} \cdot \dfrac{\cancel{P(\bar{G},F)}}{P(F)}$

$$= \frac{P(E,F,G) + P(E,F,\bar{G})}{P(F)} = \frac{P(E,F)}{P(F)}$$

LHS = RHS → TRUE

Method 3:

$$P(E|F) = \underbrace{P(E|F,G)}\, P(G|F) + \underbrace{P(E|F,\bar{G})}\, P(\bar{G}|F)$$

$$\frac{P(E,G|F)}{P(G|F)} \cdot P(G|F) + \frac{P(E,\bar{G}|F)}{P(\bar{G}|F)} \cdot P(\bar{G}|F)$$

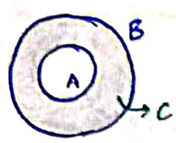
$$= P(E,G|F) + P(E,\bar{G}|F)$$

$$= P(E|F)$$

$$P(E|G) = \frac{P(E,G)}{P(G)}$$

ques. True/False

If A and B are events and $A \subseteq B$ then
$P(A) \leq P(B)$.

↳ 

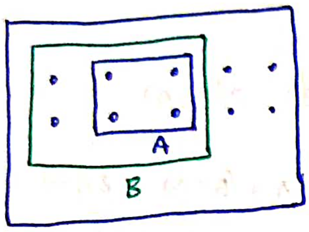
$B = A \cup C$

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

$P(B) = P(A) + \underbrace{P(C)}_{\geq 0}$

we know that $P(C) \geq 0$, therefore

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

ex 

$P(B) = \frac{6}{10}$ $\qquad A \subseteq B$

$P(A) = \frac{4}{10}$ $\qquad P(A) \leq P(B)$

ques. True/False

$P(E, F, G) \leq \min\{P(E), P(F), P(G)\}$

↳ $E \cap F \cap G \subseteq E \quad \Rightarrow \quad P(E \cap F \cap G) \leq P(E)$

$E \cap F \cap G \subseteq F \qquad P(E, F, G) \leq P(F)$

$E \cap F \cap G \subseteq G \qquad P(E, F, G) \leq P(G)$

ques. True/False

If the occurance of event F makes event E more likely, then the occurence of E necessarily makes F also more likely

↳ i.e. $P(E|F) \geq P(E) \quad \Rightarrow \quad P(F|E) \geq P(F)$

given : $P(E|F) \geq P(E)$

$\Rightarrow \frac{P(E,F)}{P(F)} \geq P(E)$

$\Rightarrow \frac{P(E,F)}{P(E)} \geq P(F)$

$\Rightarrow P(F|E) \geq P(F)$ ← Hence Proved

TRUE

ques. Which of the following is equal to $P(B \cap C | A)$?

(a) $P(B | C \cap A)$ (c) $P(B | C \cap A)\, P(C | A)$

(b) $\dfrac{P(B|C)}{P(A)}$ (d) $P(B|C)\, P(C|A)$

↳
$$P(B \cap C | A) = P_A(B \cap C)$$
$$= P_A(B|C)\, P_A(C)$$
$$= P(B | C \cap A)\, P(C|A)$$

Thus, option (C) is correct

Girls Method ↦ We can extend the normal formulae

$$P(B, C) = P(C) \cdot P(B|C)$$

Now, putting conditional prob formulae

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

∴ option (c) is correct

ques. Which of the following is true?

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

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

→ option (A) is correct

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

ques. if we wish to express $P(B|A)$ in terms of only B and $\bar{A}$, show that

$$P(B|A) = \frac{P(B) - P(B|\bar{A})\, P(\bar{A})}{1 - P(\bar{A})}.$$

↳
$$P(B|A) = \frac{P(A, B)}{P(A)}$$
$$= \frac{P(A, B) + P(\bar{A}, B) - P(\bar{A}, B)}{P(A)}$$
$$= \frac{P(B) - P(B|\bar{A})\, P(\bar{A})}{1 - P(\bar{A})} \qquad \{P(A, B) + P(\bar{A}, B) = P(B)\}$$

ques. (T/F) $P(A, B | C, D, E) = \dfrac{P(A, B, C, E | D)}{P(C, E | D)}$ TRUE

in RHS it is given 'D'. → so, lets skip 'D' first from LHS & then evaluate.

↳
$$P(A, B | C, E) = \frac{P(A, B, C, E)}{P(C, E)} \qquad \{\text{writing formula without 'D'}\}$$

Adding 'given D' to both the sides now :→

$$\rightarrow P(A, B | C, E, D) = \frac{P(A, B, C, E | D)}{P(C, E | D)}$$

→ ||'ly we can say that :

$$P(A, B | C, D, E) = \frac{P(A, B, E | C, D)}{P(E | C, D)} \checkmark$$

∵ $P(A, B | E) = \dfrac{P(A, B, E)}{P(E)}$

$$P(A, B | C, D, E) = \frac{P(A, B, C, D, E)}{P(C, D, E)}$$

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

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

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

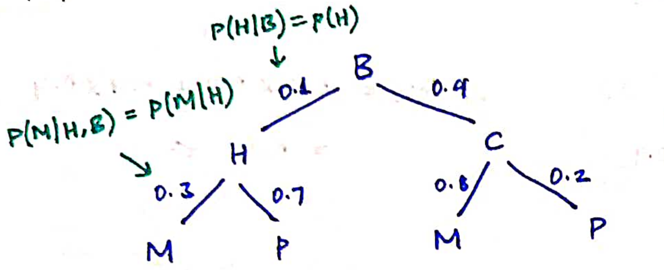
Live Session 1 : GATE PYQs Tree Method

ques. A person starts his journey from Bengaluru with 0.1 probability he goes to Hyderabad & with 0.9 prob he goes to Chennai.
If he is in Hyderabad then 0.3 probability he goes to Mumbai and with 0.7 prob he goes to Pune.
If he is in Chennai then with 0.8 prob he goes to Mumbai and with 0.2 prob he goes to Pune.
Prob that he will goes to Pune?

Prob that he will go to Mumbai?

If person is in Mumbai what is the prob that he came from Hyderabad? bcoz Bang. is just a starting point

↳



$P(M|H,B) = P(M|H)$ & $P(H|B) = P(H)$

bcoz according to the ques, once you reach to Hyd. you dont care from where you have come

& $P(M \cap H) = P(M|H)\,P(H)$
$= 0.3 \times 0.1 = 0.03$

(i) $P(\text{Pune}) = 0.1 \times 0.7 + 0.9 \times 0.2 = 0.25$

(ii) $P(\text{Mumbai}) = 0.1 \times 0.3 + 0.9 \times 0.8 = 0.75$

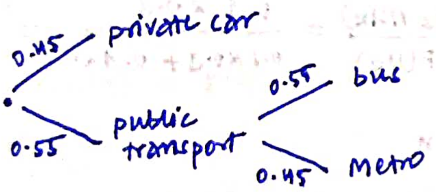
(iii) $P(H|M) = ?$

↳ $P(H|M) = \dfrac{P(H \cap M)}{P(M)} = \dfrac{0.03}{0.75} = 0.04$

ISRO 2016
ques. A person on a trip has a choice between a private car and public transport. The prob of using a private car is 0.45. While using public transport, the further choice available are bus & metro. Out of which the probability of commuting by a bus is 0.55. In such a situation, the prob of using a car, bus and metro respectively would be ?

(a) 0.45, 0.30 & 0.25 (b) 0.45, 0.25 and 0.30

(c) 0.45, 0.55 & 0 (d) 0.45, 0.35 and 0.20



$P(\text{private car}) = 0.45$

$P(\text{Bus}) = 0.55 \times 0.55 = 0.3025$

P(Bus, public)

$P(\text{Metro}) = 0.45 \times 0.55 = 0.25$

P(Metro, public)

GATE 08

ques. Aishwarya studies either CSE or Maths everyday. If she studies CSE on a day, then the prob that she'll study Maths the next day is 0.6.

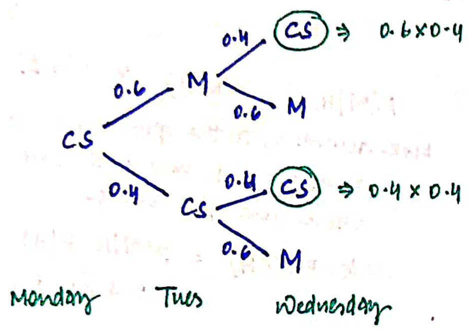
If she studies maths on a day, then the prob that she studies CSE the next day is 0.4.

Given that she studies CSE on Monday, what is the prob that she studies CSE on Wednesday?

(A) 0.24 (B) 0.36 (C) 0.4 (D) 0.6

S

$\Rightarrow$



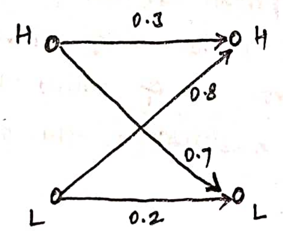
$\therefore$ Ans $= 0.6 \times 0.4 + 0.4 \times 0.4$

$= 0.4$

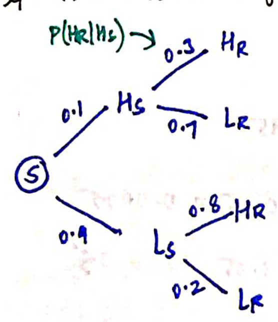
GATE 21

ques. A sender (S) transmits a signal, which can be one of the 2 kinds: H & L with probabilities 0.1 and 0.9 respectively, to a Receiver (R).

In the graph below, the weight of the edge (u,v) is the prob of receiving v when u is transmitted, where $u, v \in \{H, L\}$. For ex, the prob that the received signal is L, given the transmitted signal was H, is 0.7.



if the received signal is H, the prob that the transmitted signal was H is ______.



we want to find: $P(H_S|H_R)$

$$P(H_S|H_R) = \frac{P(H_S \cap H_R)}{P(H_R)} = \frac{0.1 \times 0.3}{0.1 \times 0.3 + 0.9 \times 0.8} = \frac{3}{75}$$

$$= 0.04$$

GATE IT 2006

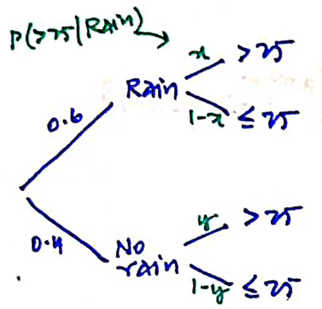
ques. In a certain town, the prob that it'll rain in the afternoon is known to be 0.6. Moreover, meterological data indicates that if the temp at noon is less than or equal to 25°C, the prob that it'll rain in the afternoon is 0.4. The temp at noon is equally likely to be above 25°C or at/below 25°C. What is the prob that it'll rain in the afternoon on a day when the temp at noon is above 25°C?

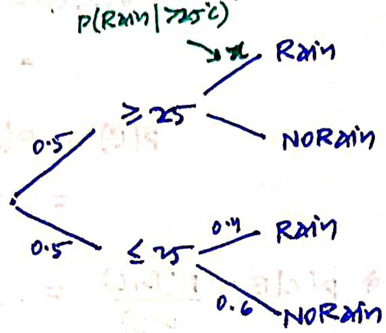
(A) 0.4 (B) 0.6 (C) 0.8 (D) 0.9

$P(\text{rain}) = 0.6$ $\qquad P(>25°C) = P(\le 25°C) = 0.5$

$P(\text{rain} \mid \le 25°C) = 0.4$ $\qquad P(\text{rain} \mid > 25°C) = ?$

⇒ we can make 2 tree diagrams from the given info:





$P(>25°C) = 0.5 = P(\le 25°C)$ $\qquad P(\text{rain}) = 0.6$

$\Rightarrow 0.6x + 0.4y = 0.5$

$0.6(1-x) + 0.4(1-y) = 0.5$ } → using total probability ← { ⇒ $0.5 \times x + 0.5 \times 0.4 = 0.6$

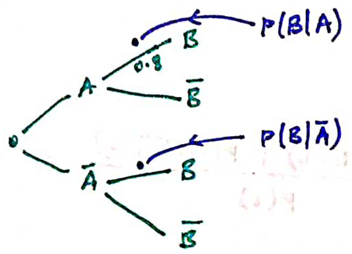
⇒ $x = 1.2 - 0.4 = 0.8$

But using these eq^ns we can't find x & y so we gotta use the other tree diagram.

⇒ $\boxed{P(\text{Rain} \mid > 25°C) = 0.8}$ ✓

→ $P(\underset{B}{\text{Rain}} \mid \underset{A}{> 25°C}) = 0.8$

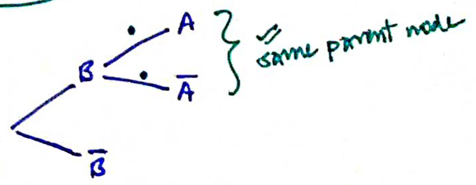
then $P(\underset{B}{\text{Rain}} \mid \underset{\bar{A}}{\le 25°C}) = \cancel{0.2}$ You don't know ⤳ bcoz $P(B|A) + P(B|\bar{A}) \neq 1$



ques. T/F

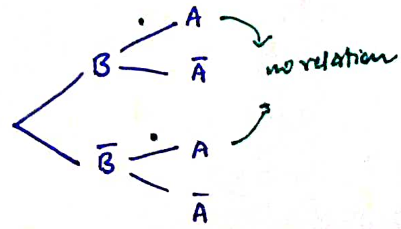
(i) $P(A|B) + P(\bar{A}|B) = 1$

TRUE



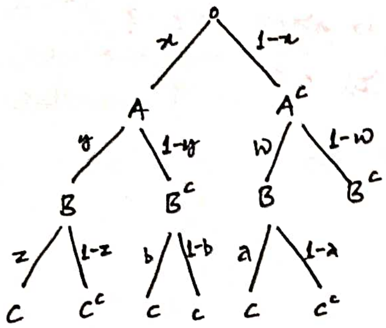
(ii) $P(A|B) + P(A|\bar{B}) = 1$

FALSE



Live Session 2 : GATE PYQs and Baye's Theorem

Someone asked this ques.



$P(C|B) = ?$

$P(C|B,A) = z$

$P(A,B,C) = xyz$

Now, $P(C|B) = \dfrac{P(B,C)}{P(B)}$

$P(B,C) = P(A,B,C) + P(A^c,B,C)$ {using marginalization}

$= xyz + (1-x)wa$

$P(B) = P(A,B) + P(A^c,B)$ {using marginalization}

$= xy + (1-x)w$

$\Rightarrow P(C|B) = \dfrac{P(B,C)}{P(B)} = \dfrac{xyz + (1-x)wa}{xy + (1-x)w}$

$P(C|A) = ?$

$\boxed{\text{Method 1}} \rightarrow$ $P(C|A) = P(C,B|A) + P(C,B^c|A)$ ✓

$= \dfrac{P(A,B,C)}{P(A)} + \dfrac{P(A,B^c,C)}{P(A)}$

$= \dfrac{xyz}{x} + \dfrac{x(1-y)b}{x} = yz + (1-y)b$

$P(C|A) = \cancel{P(C|A,B) + P(C|A,B^c)}$

we need to do marginalization wrt C.

$\boxed{\text{Method 2}} \rightarrow$ $P(C|A) = \dfrac{P(C,A)}{P(A)} = \dfrac{P(C,A,B) + P(C,A,B^c)}{x} = \dfrac{xyz + x(1-y)b}{x}$

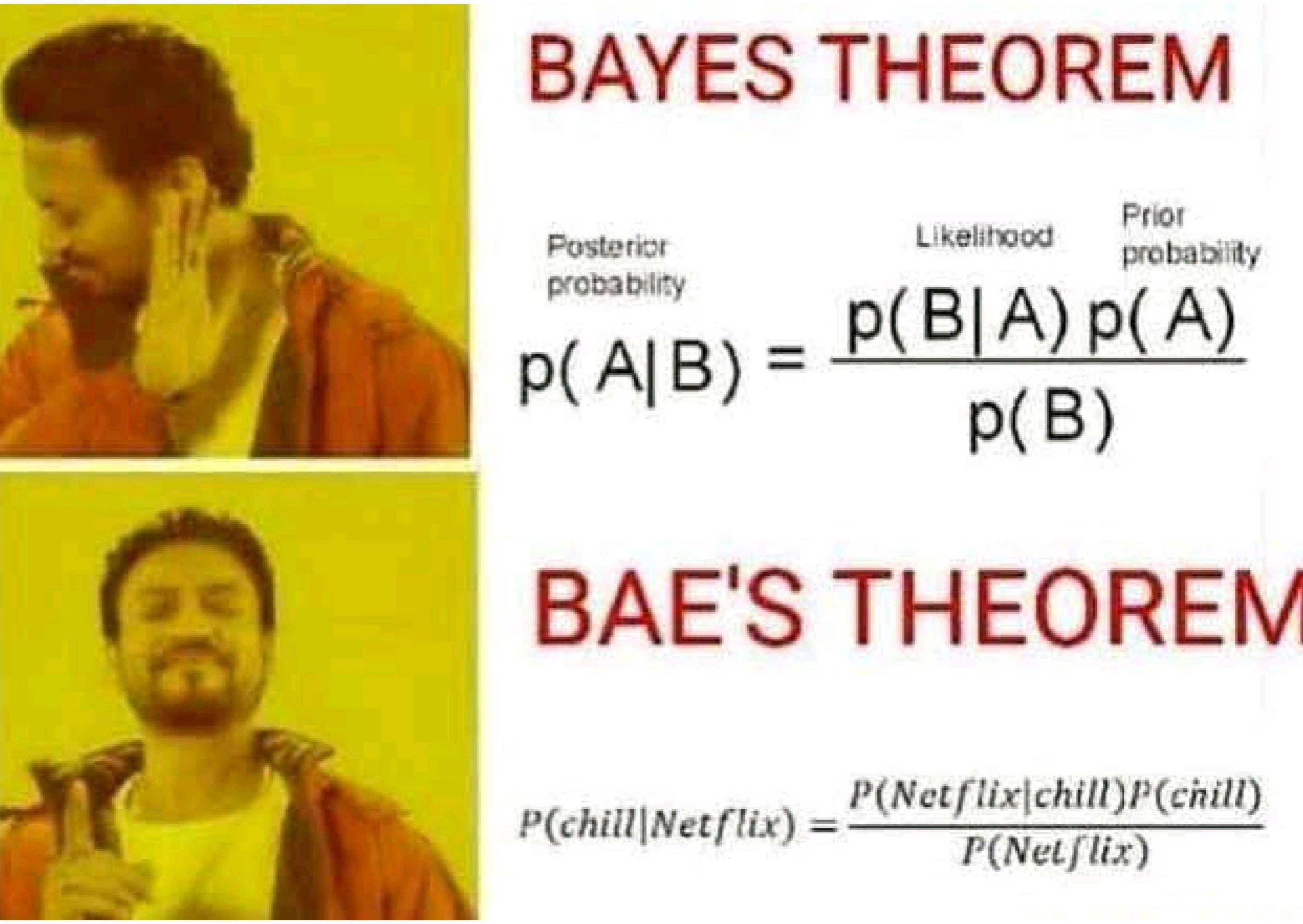
$= yz + (1-y)b$ ✓

$P(A|C) = ?$

$P(A|C) = \dfrac{P(A \cap C)}{P(C)} = \dfrac{P(A,B,C) + P(A,B^c,C)}{P(C)} = \dfrac{xyz + x(1-y)b}{P(C)}$

$P(C) = P(C,A) + P(C,A^c)$

$= P(A,B,C) + P(A,B^c,C) + P(A^c,B,C) + P(A^c,B^c,C)$

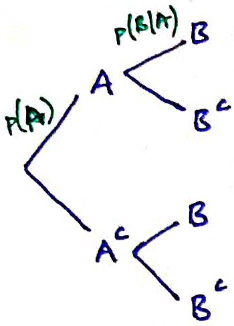


Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

BAYE'S THEOREM

→ it is for the backward reasoning, not for the forward reasoning which have done until now.

ex. $P(A|B)$, $P(A^c|B)$ etc $P(B|A)$, $P(B|A^c)$ etc



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

$$= \frac{P(A,B)}{P(A,B) + P(A^c,B)}$$

ex

Red	Blue

Bag 1

Red	Blue

Bag 2 → marbles

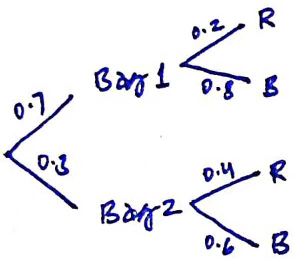
Reverse type of question

$P(\text{Bag}1) = 0.7$ $P(\text{Bag}2) = 0.3$

$P(\text{Bag}1|R) = ?$

you have chosen the Red marble what is the prob it has come from Bag 1

$P(R|\text{Bag}1) = 0.2$ $P(R|\text{Bag}2) = 0.4$



$$P(\text{Bag}1|R) = \frac{P(R \wedge \text{Bag}1)}{P(R)}$$

$$= \frac{0.2 \times 0.7}{0.7 \times 0.2 + 0.3 \times 0.4} = \frac{7}{13} = 0.54$$

→ Formal def of Baye's Theorem:

Let $A_1, A_2, \ldots A_n$ be disjoint events that form a partition of the sample space, and assume that $P(A_i) > 0$, for all i. Then, for any event B such that $P(B) > 0$, we have

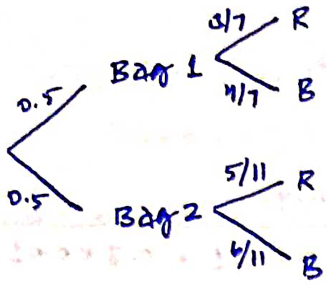
$$P(A_i|B) = \frac{P(A_i)\,P(B|A_i)}{P(B)} \quad \left\{ \text{or } \frac{P(A_i, B)}{P(B)} \right\}$$

$$= \frac{P(A_i)\,P(B|A_i)}{P(A_1)\,P(B|A_1) + P(A_2)\,P(B|A_2) + \cdots + P(A_n)\,P(B|A_n)}$$

Ai's are making a partition so we have marginalized (Total prob)

Questions

(i)

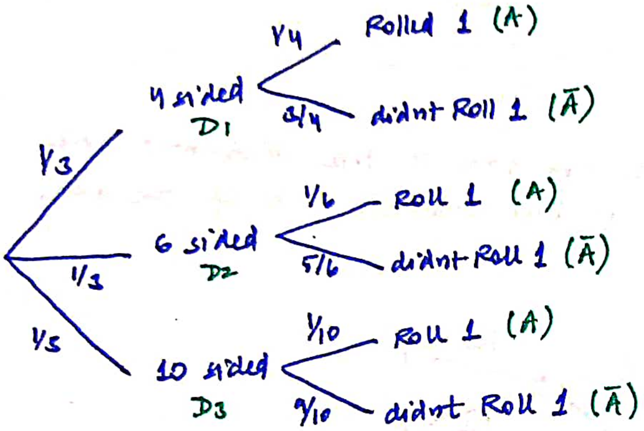


$P(Bag2 \mid R) = ?$

$$P(Bag2 \mid R) = \frac{P(R, Bag2)}{P(R)}$$

$$= \frac{0.5 \times \frac{5}{11}}{0.5 \times \frac{3}{7} + 0.5 \times \frac{5}{11}} = 0.37$$

(ii)



Let A be event that Rolled 1 & dies be D1, D2 & D3.

$\therefore$ $P(D1 \mid A) = ?$

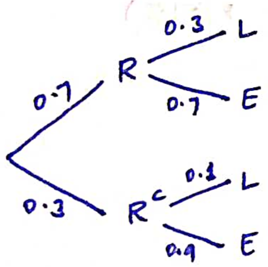
$$P(D1 \mid A) = \frac{P(A, D1)}{P(A)}$$

$$= \frac{\frac{1}{3} \times \frac{1}{4}}{\frac{1}{3} \times \frac{1}{4} + \frac{1}{3} \times \frac{1}{6} + \frac{1}{3} \times \frac{1}{10}} = \frac{15}{31}$$

(iii) L ≡ Late, R ≡ Rain, E ≡ Early, $P(L \mid R) = 0.3$

$P(L \mid R^c) = 0.1$

$P(R) = 0.7$



(a) $P(E) = P(R, E) + P(R^c, E)$ {Marginalisation}

$= P(R)\,P(E \mid R) + P(R^c)\,P(E \mid R^c)$

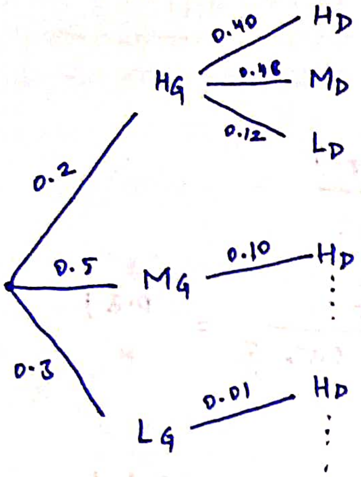
$= 0.7 \times 0.7 + 0.3 \times 0.9 = 0.76$

(b) $P(R \mid E) = \dfrac{P(R, E)}{P(E)}$

$$= \frac{P(R, E)}{P(R, E) + P(R^c, E)} = \frac{0.7 \times 0.7}{0.7 \times 0.7 + 0.3 \times 0.9} = \frac{49}{76} = 0.6447$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

(iv)



$P(HG|HD) = ?$

$$P(HG|HD) = \frac{P(HG, HD)}{P(HD)}$$

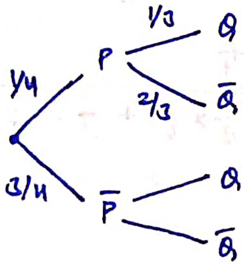
$$= \frac{0.2 \times 0.4}{0.2 \times 0.4 + 0.5 \times 0.1 + 0.3 \times 0.01}$$

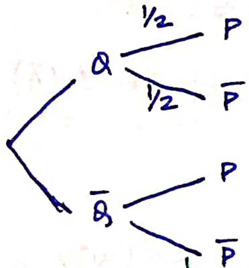
$$= \frac{8}{13.3} \approx 0.60$$

(v) $P(P) = \frac{1}{4}$ $\quad P(B|P) = \frac{1}{3}$

$P(P|B) = \frac{1}{2}$ $\quad P(\bar{P}|\bar{B}) = ??$

we can have 2 tree diagrams here as our understanding of the qus





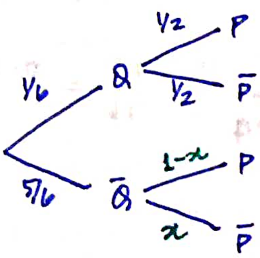
this is what we need to find $P(\bar{P}|\bar{B})$

we need to find out $P(B)$ first;

$$\Rightarrow P(P, B) = \underset{1/4}{P(P)} \cdot \underset{1/3}{P(B|P)} = P(B) \cdot \underset{1/2}{P(P|B)}$$

$$\Rightarrow P(B) = \frac{\frac{1}{4} \times \frac{1}{3}}{1/2} = \frac{1}{6}$$

Now, the 2nd tree will become



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

$P(\bar{P}) = P(B, \bar{P}) + P(\bar{B}, \bar{P})$

$$\frac{3}{4} = \frac{1}{6} \times \frac{1}{2} + \frac{5}{6} \times x \Rightarrow \boxed{x = \frac{4}{5} = 0.8}$$

$$\therefore \boxed{P(\bar{P}|\bar{B}) = 0.8}$$

(vi) $P(\text{odd}) = 0.9 \times P(\text{even})$ $P(\text{even} \mid >3) = 0.75$ $P(>3) = ?$

$\hookrightarrow \{2, 4, 6\}$

$\hookrightarrow$ let, ~~P(even)~~ $P(2) = P(4) = P(6) = x$

$\therefore \; P(\text{even}) = 3x$

$\Rightarrow \; P(\text{odd}) = 0.9 \times 3x = \cancel{0.27}\, 2.7x$

$P(\text{odd}) + P(\text{even}) = 1$

or $2.7x + 3x = 1$

$x = \dfrac{1}{5.7} = P(2) = P(4) = P(6)$

Now, $P(\text{even} \mid >3) = 0.75$

$$\Rightarrow \frac{P(\text{even}, >3)}{P(>3)} = \frac{2x}{P(>3)} = 0.75$$

$$\Rightarrow P(>3) = \frac{0.75}{2x} = \frac{0.75}{\frac{2}{5.7}} = 0.468$$

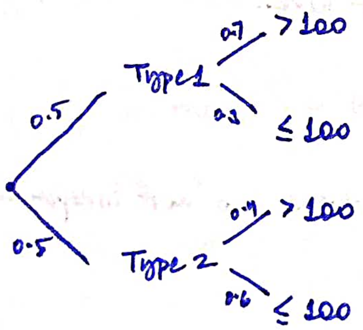
OR

$P(\text{odd}) + P(\text{even}) = 0.9x + x = 1 \Rightarrow x = \dfrac{1}{1.9}$ $P(\text{even}) = \dfrac{1}{1.9}$

$$P(\text{even} \mid >3) = \frac{P(\text{even}, >3)}{P(>3)} = \frac{P(>3 \mid \text{even})\, P(\text{even})}{P(>3)}$$

$$\Rightarrow 0.75 = \frac{\frac{2}{3} \times \frac{1}{1.9}}{P(>3)} \Rightarrow P(>3) = 0.468$$

(vii)



$P(>100) = P(\text{Type 1}, >100) + P(\text{Type 2}, >100)$

$= 0.5 \times 0.7 + 0.5 \times 0.4$

$= 0.55$

Independence of Events

till now we have been seeing, $P(A|B) = \frac{P(A,B)}{P(B)}$ — given that B has happened, the prob of A.

the occurance of B updates the prob of occurance of A. our beliefs are changing

• Do we always update our beliefs ?

A : I had a sandwich for breakfast.

B : It will rain today

If A occurs, will you update your belief about B ?

↳ NO, occurance of A provides no info about B.

$\Rightarrow P(B|A) = P(B)$ There is no impact of occurance of A over B.

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

→ In general,

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

For independent events,

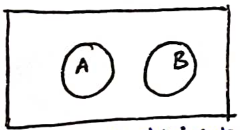
$\boxed{P(A \cap B) = P(A) \cdot P(B)}$

comparing both $\Rightarrow \boxed{P(B) = P(B|A)}$

↳ $P(A) > 0$ bcoz A is in denominator

qun. Are these events Independent ? from the diag. they may look independent but they are not

NO

(i) 

A, B are disjoint
↓
They are Mutually Exclusive

if A has occurred, the prob of occurance of B is 0, B can never occur
↳ so, A clearly conveys info about B

so, they are extremely dependent.

$P(B|A) = 0$ $\quad$ $||^{ly}$, $P(A|B) = 0$.

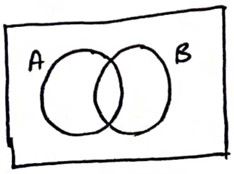
$\Rightarrow P(B|A) = P(B)$ is not present here.

using formulae, $P(A \cap B) = P(A) \cdot P(B)$

LHS : $P(A \cap B) = P(\underbrace{\{\ \}}_{\text{empty set}}) = 0$

RHS: $P(A) \cdot P(B) > 0$ — $P(A) > 0$, $P(B) > 0$

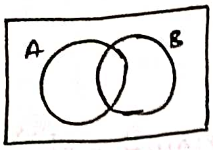
} LHS ≠ RHS , so not independent

(ii) 

for indep: $P(A \cap B) = P(A) \cdot P(B)$

→ we cant say anything here.

(iii) $P(A) = 0.30$ $P(B) = 0.40$ $P(A\cap B) = 0.12$

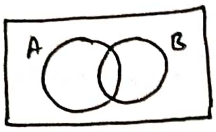


$P(A\cap B) = 0.12$

$P(A)\cdot P(B) = 0.3 \times 0.4 = 0.12$

$\Rightarrow P(A\cap B) = P(A)\cdot P(B) \rightarrow$ Independent ✓

(iv) $P(A) = 0.30$ $P(B) = 0.40$ $P(A\cap B) = 0.20$



$P(A)\cdot P(B) = 0.12 \neq P(A\cap B)$

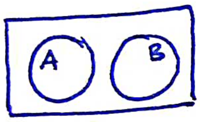
$\Rightarrow$ Not Independent

o→ Use $P(A\cap B) = P(A)\cdot P(B)$ to check indep. events.
Venn diag doesnt reveal anything

Mutually Exclusive Events and Independent Events.

Mutually Exclusive events:

$P(A\cap B) = 0$



Independent Events:

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

o→ Two ME events are indep. iff $P(A) = 0$ or $P(B) = 0$

Inclusion-Exclusion principle: $P(A\cup B) = P(A) + P(B) - P(A\cap B)$

for ME events:

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

for Indep. Events:

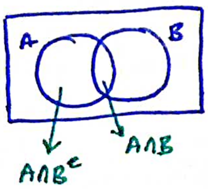
$P(A\cup B) = P(A) + P(B) - P(A)\cdot P(B)$

$\boxed{\text{NOTE}}$ → If A and B are indep. then A and B^c are also indep.

Intuitive Arg :→

If A doesnot change your belief about the likelihood that B occurred then A should not change your belief about the likelihood that B doesnot occur.

Formal Proof :→ we want to prove if $P(A\cap B) = P(A)(B)$, then $P(A\cap B^c) = P(A)\, P(B^c)$



Proof: $A = (A\cap B) \cup (A\cap B^c)$

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

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

$P(A\cap B^c) = P(A)\,[1 - P(B)]$

$P(A\cap B^c) = P(A)\cdot P(B^c)$ ✓ Hence proved

NOTE →

A, B are indep. $\Leftrightarrow$ A^c, B are indep

$\Updownarrow$ $\qquad$ $\Leftrightarrow$

A, B^c are indep $\qquad$ A^c, B^c are indep.

Independence of 3 events

indep of 2 events : $P(A\cap B) = P(A)\cdot P(B)$

events A, B and C are said to be independent if all of the following eq's hold :

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

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

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

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

Indep. of 4 events A, B, C, D

any pair of 2 : $P(A\cap B) = P(A)\cdot P(B)$

any pair of 3 : $P(A\cap B\cap C) = P(A)\cdot P(B)\cdot P(C)$

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

ques. True/False

(i) If collection of events are independent then they are pairwise independent.

TRUE

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

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

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

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

} if this is given → then A, B; A, C; B, C are gonna be independent

collection of events are independent

(ii) If collection of events are pairwise independent then they are independent

FALSE

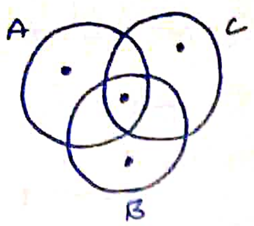
given: $P(A\cap B) = P(A)\cdot P(B)$

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

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

but $P(A\cap B\cap C) = P(A)\cdot P(B)\cdot P(C)$ is not given, so not independent only pairwise independent

(ex)



→ A, B, C pairwise indep? (yes)

→ A, B, C independent? (NO)

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

$P(B \cap C) = P(B)\, P(C)$ ✓
$\frac{1}{4} = \frac{2}{4} \cdot \frac{2}{4}$

$P(A \cap C) = P(A)\, P(C)$ ✓
$\frac{1}{4} = \frac{2}{4} \cdot \frac{2}{4}$

But $P(A \cap B \cap C) = P(A)\, P(B)\, P(C)$ ✗ → doesnot hold.
$\frac{1}{4} = \frac{2}{4} \cdot \frac{2}{4} \cdot \frac{2}{4}$

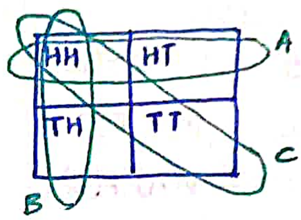
Independence and Pair wise Independence

• Two independent fair coin tosses

Event A : First toss is H

Event B : ~~First~~ second toss is H.

Event C : 1st and 2nd toss gives same result.



- $P(A) = P(B) = 1/2$
- $P(C) = \frac{2}{4} = 1/2$

→ $P(A \cap B) = P(A) \cdot P(B)$
$1/4 \quad 2/4 \quad 2/4$

$P(A \cap C) = P(A) \cdot P(C)$
$1/4 \quad 1/2 \quad 1/2$

$P(B \cap C) = P(B) \cdot P(C)$
$1/4 \quad 1/2 \quad 1/2$

→ $P(A \cap B \cap C) \neq P(A) \cdot P(B) \cdot P(C)$
$1/4 \quad 1/2 \quad 1/2 \quad 1/2$

→ A, B, C are pairwise indep.
But not indep.

The point is that just knowing about A or just knowing about B tells us nothing about C, but knowing what happened with both A & B gives us info about C. (infact, in this case gives us perfect info about C).

→ Pairwise Indep doesnot imply Independence.

Pairwise

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

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

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

$\not\Rightarrow$

3-wise

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

K-wise

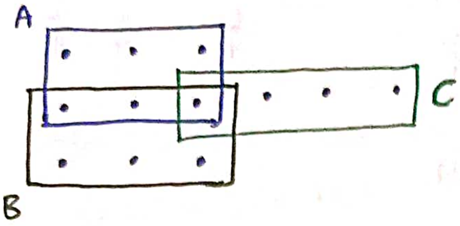
qun. T/F

If we have $P(A\cap B\cap C) = P(A).P(B).P(C)$ then we can also say pairwise indep.

i.e

$$P(A\cap B\cap C) = P(A)P(B)P(C) \Rightarrow \begin{array}{ll} P(A\cap B) = P(A)P(B) & \text{and} \\ P(A\cap C) = P(A)P(C) & \text{and} \\ P(B\cap C) = P(B)P(C) & \end{array}$$

↳ eg



$$P(A\cap B\cap C) = P(A)P(B)P(C)$$
$$\tfrac{1}{12} \qquad \tfrac{6}{12}\times\tfrac{6}{12}\times\tfrac{4}{12}$$

$$\underset{\frac{3}{12}}{P(A\cap B)} \neq \underset{\frac{6}{12}}{P(A)}\ \underset{\frac{6}{12}}{P(B)}$$ ||'y for $P(A\cap C)$ & $P(B\cap C)$

$\boxed{\text{NOTE}}$ → Neither pairwise independence imply independence nor implied by independence

2-wise

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

$P(B\cap C) = P(B)P(C)$ $\quad \not\Rightarrow \quad \not\Leftarrow$

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

3-wise

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

o→ K-wise $\not\Leftrightarrow$ K'-wise

eg Consider 2 events A & B such that $P(A\cap B) \neq P(A)P(B)$ i.e they are dependent events.

Now consider a 3rd event $C = \phi$.

$P(A\cap B\cap C) = 0$ since $C = \phi$ and $P(A)\cdot P(B)\cdot P(C) = 0$ since $P(C) = 0$.

So, Although every set of 3 events in this ~~situation~~ collection (there is only one set of 3 events) has the independence property, this collection is not pairwise independent

qun. Are A and A^c independent if $0 < P(A) < 1$?

Answer 1: Intutively, no. If you know A happens, then you know A^c does not happen

Answer 2: Formally, $P(A\cap A^c) = P(\phi) = 0$.

if $0 < P(A) < 1$, then $P(A)P(A^c) \neq 0$

o→ A & A^c are extremely dependent

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

CONDITIONAL INDEPENDENCE

Two events A and B are independent given C :

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

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

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

Now,
$P(A \cap B \mid C) = P(A \mid C)\, P(B \mid A, C)$

comparing them

OR

$$P(B \mid C) = P(B \mid A, C)$$

→ A and B are independent given C.

Earlier Normally, it was like $P(B) = P(B \mid A)$.

Earlier, Independence	Conditional Independence
$P(A \cap B) = P(A)\, P(B)$ — ①	$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C)$ — ①
$P(A \cap B) = P(A) \cdot P(B \mid A)$ — ②	$P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid A, C)$ — ②
$= P(B) \cdot P(A \mid B)$ — ③	$= P(B \mid C) \cdot P(A \mid B, C)$ — ③
1 & 2 $\Rightarrow$ $P(B) = P(B \mid A)$	1 & 2 $\Rightarrow$ $P(B \mid C) = P(B \mid A, C)$
1 & 3 $\Rightarrow$ $P(A) = P(A \mid B)$	1 & 3 $\Rightarrow$ $P(A \mid C) = P(A \mid B, C)$

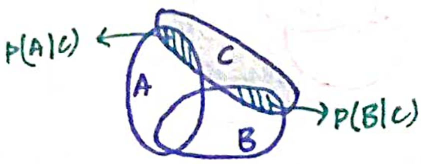
NOTE →

if A and B are indep then
- A and B' are indep
- A' and B are indep
- A' and B' are indep

if A and B are indep given C then
- A and B' are indep given C
- A' and B are indep given C
- A' and B' are indep given C

Qus. T/F . $P(A \cap B) = P(A) \cdot P(B) \Rightarrow P(A \cap B \mid C) = P(A \mid C) \cdot P(B \mid C)$? FALSE

ex Assume A and B are indep i.e. $P(A \cap B) = P(A) \cdot P(B)$



here, $P(A \cap B \mid C) = 0$
& $\underset{>0}{P(A \mid C)} \cdot \underset{>0}{P(B \mid C)} > 0$

$\Big\} \Rightarrow P(A \cap B \mid C) \neq P(A \mid C)\, P(B \mid C)$

so, A & B are not independent given C.

earlier A & B were indep but after C they got dependent.

(ex) Independence doesn't imply conditional Independence

Suppose that my friends Alice & Bob are the only 2 people who ever call me. Each day, they decide independently whether to call me : letting A be the event that Alice calls and B be the event that Bob calls

A & B are unconditionally independent

But suppose that I hear the phone ringing now. conditional on this observation, A & B are no longer independent : if the phone call isn't from Alice, it must be from Bob. In other words, letting R be the event that the phone is Ringing,

we have $P(B|R) < 1$, $P(B|A^c, R) = 1$

so, B and A^c are not conditionally indep given R, and likewise for A & B.

→ suppose there are only 2 people Alice & Bob who calls me. And both of them calls me independently

A : Alice's call

B : Bob's call

R : Phone Ring

$P(A \cap B) = P(A) P(B)$ (given that both calls independently)

Suppose that phone rings then –

$P(A|R)$ be the prob that Alice has called given the phone rings

we can say that

$P(A | B^c, R) = 1 \neq P(A|R)$

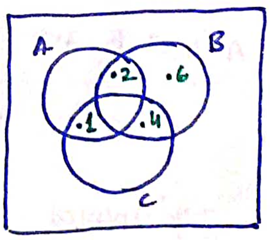
so, $P(A|R) \neq P(A|B^c, R)$ ⇒ $\boxed{\text{not conditional Indep}}$

→ earlier A & B we indep, but after C they just got dependent.

$P(A \cap B) = P(A) P(B)$ $\not\Rightarrow$ $\not\Leftarrow$ $P(A \cap B | C) = P(A|C) P(B|C)$

(ex) consider rolling a die & let,

$A = \{1, 2\}$ $B = \{2, 4, 6\}$ $C = \{1, 4\}$



Then, we have $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{2}$

$P(A \cap B) = \frac{1}{6} = P(A) \cdot P(B)$

Thus, A & B are indep., but we have

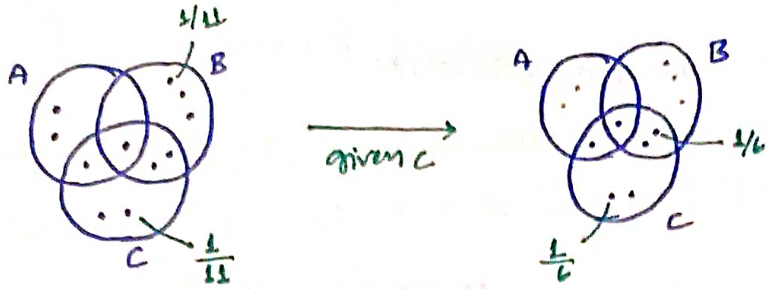
$P(A|C) = \frac{1}{2}$, $P(B|C) = \frac{1}{2}$; $P(A \cap B | C) = P(\{2\} | C) = 0$

Thus, $P(A \cap B | C) \neq P(A|C) P(B|C)$

which means A & B are not conditionally indep given C.

(ex) Conditional Independence doesn't imply Independence

i.e. $P(A \cap B | C) = P(A|C)\, P(B|C) \not\Rightarrow P(A \cap B) = P(A)\,P(B)$



$$P(A \cap B | C) = P(A|C)\; P(B|C)$$
$$\tfrac{1}{6} = \tfrac{1}{3} \times \tfrac{1}{2}$$

even though They are conditionally independent

$$P(A \cap B) \neq P(A)\; P(B)$$
$$\tfrac{1}{11} \neq \tfrac{4}{11} \times \tfrac{6}{11}$$

they are not indep.

qun. T/F . Given the following factorization →

$$P(A,B,C) = P(C)\, P(B|C)\, P(A|C)$$

can we say A & B are independent given C ? TRUE

↳ $P(A,B,C) = P(C)\; P(B|C)\; \underline{P(A|B,C)}$ (general formula)

but here we have, $P(A|C)$

comparing both, $P(A|C) = P(A|B,C)$ ⇒ A & B are conditional independent.

⇒ $A \perp B | C$

qun. Suppose A and B are conditionally independent given C then which of the following is/are true ?

a) $P(A \cap B) = P(A)\; P(B)$ ✗

b) $P(A,B,C) = P(A|B,C)\; P(B|C)\; P(C)$ ✓

c) $P(A,B,C) = P(A|C)\; P(B|C)\; P(C)$ ✓

d) $P(A,B|C) = P(A|C)\; P(B|C)$ ✓

↳ $P(A,B,C) = P(C) \cdot P(B|C) \cdot P(A|B,C)$ ⟶ if A & B are indep given C, then

$$P(A|B,C) = P(A|C)$$

⇒ $P(A,B,C) = P(C)\; P(B|C)\; P(A|C)$

$P(\underbrace{A,B}, C) = P(C) \cdot P(A,B|C)$

can consider whole thing as one event

Qun. Mark True to all the expressions that are equal to P(A|B), given no independence assumption.

i) $P(A, C|B) + P(A, C^c|B)$ TRUE

$\hookrightarrow$ $P(A|B) = P(A, C|B) + P(A, C^c|B)$ Using marginalization

(ii) $\dfrac{P(A, C|B)}{P(C|B)}$ FALSE

$\hookrightarrow$ $\dfrac{P(A, C|B)}{P(C|B)} = \dfrac{\frac{P(A, B, C)}{P(B)}}{\frac{P(C, B)}{P(B)}} = \dfrac{P(A, B, C)}{P(B, C)} = P(A|B, C) \neq P(A|B)$

(iii) $P(A|B, C) + P(A|B, C^c) \neq P(A|B)$ FALSE

$\hookrightarrow$ this is not equal to P(A|B), bcoz marginalization doesnt work on the given that side.
if it was $P(A, C|B) + P(A, C^c|B)$ then it were true.

(iv) $\dfrac{P(A, B, C) + P(A, B, C^c)}{P(B, C) + P(B, C^c)}$ TRUE

$\hookrightarrow$ $P(A|B) = \dfrac{P(A, B)}{P(B)} = \dfrac{P(A, B, C) + P(A, B, C^c)}{P(B, C) + P(B, C^c)}$ Using marginalization

(v) $\dfrac{P(B|A) * P(A|C)}{P(B, C) + P(B, C^c)}$ FALSE

$\hookrightarrow$ $\dfrac{P(B|A)\, P(A|C)}{P(B)} \neq P(A|B)$

(vi) $\dfrac{P(A|C, B)\, P(C|A, B)}{P(C|B)}$ FALSE

$\hookrightarrow$ $\dfrac{\frac{P(A, B, C)}{P(B, C)} \cdot \frac{P(A, B, C)}{P(A, B)}}{\frac{P(B, C)}{P(B)}} \neq P(A|B)$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques. Mark all expressions that are equal to $P(A,B,C)$, given that $A \perp B$ (A & B are indep).

$\hookrightarrow P(A,B,C) = P(A) \cdot P(B|A) \cdot P(C|A,B)$ general formula

$A \perp B$ means

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

or $P(A|B) = P(A)$

or $P(B|A) = P(B)$

(i) $P(A)\,P(B)\,P(C|A,B)$ TRUE

$\hookrightarrow P(A,B,C) = P(A)\,P(B|A)\,P(C|A,B)$

$= P(A)\,P(B)\,P(C|A,B)$

(ii) $P(C)\,P(A|C)\,P(B|C)$ not equal

$\hookrightarrow P(A,B,C) = P(C) \cdot P(A|C) \cdot P(B|A,C)$

now $P(B|A,C) \neq P(B|C)$

so given expression is FALSE

(iii) $P(A)\,P(B|A)\,P(C|A,B)$ TRUE

$\hookrightarrow P(A,B,C) = P(A)\,P(B|A)\,P(C|A,B)$

(iv) $P(A|C) \cdot P(C|B) \cdot P(B)$ not equal

$\hookrightarrow P(A,B,C) = P(B)\,P(C|B)\,P(A|B,C)$

now, $P(A|B,C) \neq P(A|C)$

so given expression is FALSE

(v) $P(A)\,P(C|A)\,P(B|C)$ not equal

$\hookrightarrow P(A,B,C) = P(A)\,P(C|A)\,P(B|A,C)$

$P(B|C) \neq P(B|A,C)$

so given exp is FALSE

(vi) $P(A,C)\,P(B|A,C)$ TRUE

$\hookrightarrow P(B,\underbrace{A,C}_{\text{take this as single event}}) = P(A,C)\,P(B|A,C)$

ques. if $A \perp B | C$. then $P(A,B|C) = ?$

(i) $P(A|C)\,P(B|C)$

TRUE $P(A,B|C) = P(A|C)P(B|C)$

(ii) $\dfrac{P(C,A|B)\,P(B)}{P(C)}$ TRUE

$\hookrightarrow P(A,B|C) = \dfrac{P(A,B,C)}{P(C)} = \dfrac{P(B)\,P(C,A|B)}{P(C)}$

(iii) $\dfrac{P(C)\,P(B|C)\,P(A|C)}{P(C|A,B)}$ FALSE

$\hookrightarrow \dfrac{\cancel{P(C)} \cdot \dfrac{P(B,C)}{\cancel{P(C)}} \cdot \dfrac{P(A,C)}{P(C)}}{\dfrac{P(C,A,B)}{P(A,B)}} \neq \dfrac{P(A,B,C)}{P(C)}$

(iv) $P(A|B)\,P(B|C)$

FALSE

$\hookrightarrow$ it should have been $P(A|C)\,P(B|C)$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

RANDOM VARIABLES INTRO

Even though this chapter may appear to cover a lot of new ground, but this is not really the case.

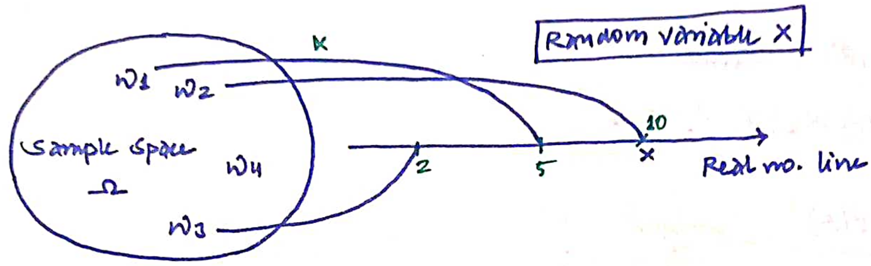
In this chapter, we'll simply take concepts from chapter 1 (probabilities, conditioning, independencies etc) & apply them on Random variables, rather than events.

The only genuine new concepts relate to means & variances.

Random Variable

↳ Function that maps an outcome to the Real no.

w_1, w_2, w_3, w_4 are outcomes.



$x(w_1) = 5$

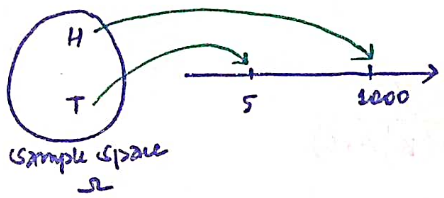
$x(w_2) = 10$

$x(w_3) = 2$

ex tossing a coin

possible outcomes : $\{H, T\}$

let X be the Random variable



$X(H) = 1000$

$X(T) = 5$

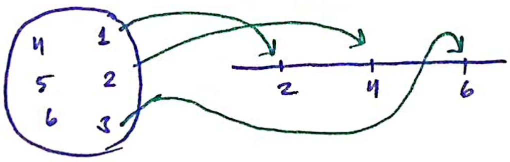
$X = 1000 \leftarrow \{H\}$

ex Rolling a dice

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

y : RV

y : twice of the no. shown on face



$y(1) = 2$ $\quad y = 6 \leftarrow \{3\}$

$y(2) = 4$

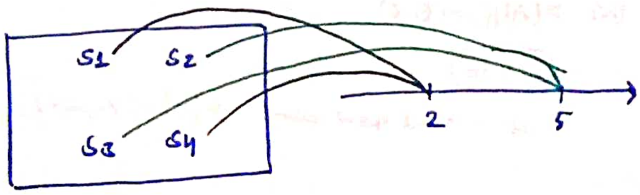
Why are we calling it "Random variable" when it is a func^n?

↳ 'Random' bcoz we are assigning some func^n to Random experiment (outcomes of Random exp is sample space Ω)

'Variable' bcoz we want to treat this as variable.

we dont want to write $y(1)=2$, but $y = 2$
$x(H) = 1000$, but $X = 1000$ } mathematical convenience

ex



$X=5$ is a set. $\{s_2, s_3\}$

$X=2$ is a set $\{s_1, s_4\}$

$X=5 \rightarrow \{s_2, s_3\}$ → associated set of outcomes related to $X=5$

$\boxed{X = 5 \equiv \{w \mid X(w) = 5\}}$

$X \leq 5 \equiv \{s_1, s_2, s_3, s_4\}$

$X < 5 \equiv \{s_1, s_4\}$

$X < 10 \equiv \{s_1, s_2, s_3, s_4\}$

$X > 0 = \{s_1, s_2, s_3, s_4\}$

ex) tossing a coin

	$X(\cdot)$
H	5
T	10

} func^n mapping Head to 5, & Tail to 10.

now, $P(X=5)$

$= P(\{\omega \mid X(\omega)=5\})$

$= P(\{H\})$

$= 1/2$

ex) Rolling a dice

→

	$P(\cdot)$	$X(\cdot)$
1	1/6	1
2	1/6	1
3	1/6	1
4	1/6	2
5	1/6	2
6	1/6	2

- $P(\text{outcome is } 3) = 1/6$
- $P(X=1)$
  $= P(\{\omega \mid X(\omega)=1\})$
  $= P(\{1,2,3\})$
  $= \frac{1}{6}+\frac{1}{6}+\frac{1}{6} = \frac{3}{6} = \frac{1}{2}$
- $P(X=2) = 3/6 = 1/2$

→

	$P(\cdot)$	$X(\cdot)$
1	1/12	1
2	1/6	2
3	1/12	1
4	4/12	1
5	4/24	2
6	4/24	1

- $P(X=1) = P(1)+P(3)+P(4)+P(6)$
  $= \frac{1}{12}+\frac{1}{12}+\frac{4}{12}+\frac{4}{24}$
  $= \frac{8}{12} = 0.67$

ex) Possible outcomes in sample space $= \{1, 3, 7, 12\}$

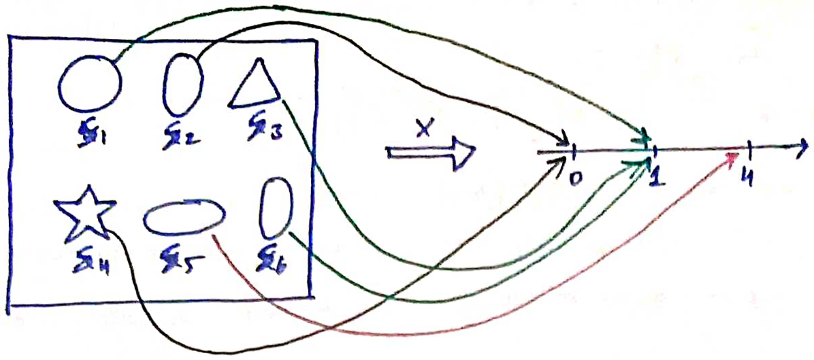
	$P(\cdot)$	$X(\cdot)$
1	1/2	1
3	1/4	1
7	1/8	1
12	1/8	1

$P(X=1) \longrightarrow P(\{\omega \mid X(\omega)=1\})$
$= 1$

$P(X=2) = 0$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ex



Assume $\{s_1, s_2, \dots s_6\}$ is equally likely i.e. $P(s_1) = P(s_2) = \dots = P(s_6) = 1/6$

- $P(X < 1) = P(\{s_2, s_4\}) = P(s_2) + P(s_4) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$
- $P(X < 2) = 5/6$
- $P(1 < X < 2) = 0$
- $P(X \le 4) = 1$

  ↳ $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$ ✓

$\boxed{\text{NOTE}}$ → Any condition on RV is an event.

ex Rel $X < 2 \equiv \{w \mid X(w) < 2\}$

↖ set of some outcomes ≡ EVENT

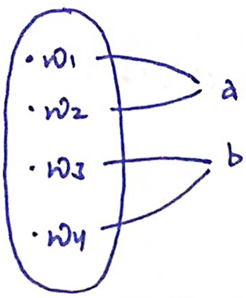
$X = 3 \equiv \{w \mid X(w) = 3\}$

$X = 2$ is an event ✓
$X < 2$ is an event ✓
$X = 3$ is an event ✓

qus. T/F.

Suppose we have 2 events $X = a$, $X = b$. We can say that both of the events are mutually exclusive. → there is no common outcome.

↳ Let.



$X = a$
↳ $\{w_1, w_2\}$

$X = b$
↳ $\{w_3, w_4\}$

RV is a func^n. A func^n cannot give multiple outputs to an input & every input must have an output

$f(x) = x^2$ $\qquad f: R \to R^+$

$f(2) = 4$
$f(3) = 9$
$f(-2) = 4$

$2 \to 4$, $-2 \to 4$ func^n ✓

$2 \to 4$, $2 \to 5$ ← not even a func^n

qus. T/F

X maps any outcome to either a, b or c

Thus $X=a$, $X=b$ and $X=c$, these 3 events are mutually exclusive & exhaustive.

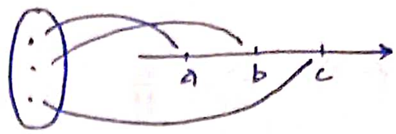
covers whole samplespace

TRUE

According to the func^n def it is True

Any outcome is either mapped to $X=a$ or $X=b$ or $X=c$.

so all the outcomes are getting covered.



- $X=a$ is an event ✓
- $X=a$, $X=b$ are mutually exclusive ✓
- $X=a, X=b, X=c$ are mutually exhaustive ✓ ↳ covers everything in the sample space.

(suppose only 3 possibilities)

$X=a, X=b, X=c$ are partitioning the sample space

ex $P(y=2) \equiv P(y=2, \underbrace{X=a}) + P(y=2, X=b) + P(y=2, X=c)$ Marginalization

this is nothing just events as seen before

{ earlier we use to marginalize with events.

ex $P(y=2) \equiv P(y=2, A_1) + P(y=2, A_2) + P(y=3, A_3)$ }

events

qus. T/F . $P(X=5)$ is same as $\sum_k P(X=5, y=k)$

↳ TRUE

ex Coin Toss.

let X be a RV which takes value 1 on Head and 0 on tail.

what is the prob of X being 1 ?

		X
H	1/2	1
T	1/2	0

$P(X=1) = 1/2$

Coin Toss Twice

X = no. of H after both flips

X_1 = no. of H on 1st flip

X_2 = no. of H on 2nd flip

P(·)		X	X_1	X_2	y
1/4	HH	2	1	1	2
1/4	HT	1	1	0	1
1/4	TH	1	0	1	1
1/4	TT	0	0	0	0

} outcomes when $y<2$

$P(X=2) = ?$ $P(X_1+X_2 < 2) = ?$

$P(X=2) = 3/4$

$P(\underbrace{X_1+X_2} < 2) = 1/4 + 1/4 + 1/4 = 3/4$

it is also a RV but we call it as func^n of RV → $y = X_1 + X_2$

i.e $y(\omega) = X_1(\omega) + X_2(\omega)$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

eg Two Dice Tossed

X = "sum of the scores on the 2 dice"

Possible values of X ?

↳ $\boxed{2 \text{ to } 12}$

now $P(X = 2) = \frac{1}{36}$ — (1, 1) $P(X=2) = P(\{(1,1)\}) = \frac{1}{36}$

$P(X = 3) = \frac{2}{36} = \frac{1}{18}$

↳ (1,2), (2,1)

ques. $P(X = x) = \begin{cases} 1/9, & \text{if } x \text{ is an integer in the range } [-4, 4]. \\ 0, & \text{otherwise} \end{cases}$

let $y = |X|$. Find out $P(y = 2)$. let $Z = X^2$. Find $P(Z = 16)$

PMF of X

X	$P(X=k)$	y	Z
−4	1/9	4	16
−3	1/9	3	9
−2	1/9	2	4
−1	1/9	1	1
0	1/9	0	0
1	1/9	1	1
2	1/9	2	4
3	1/9	3	9
4	1/9	4	16

- $P(X < 0) = 4/9$
- $P(y = 2) = 2/9$ $P(y=2) = P(X = 2 \text{ or } X = -2)$

$P(y = 2) = P(X=2) + P(X=-2)$

$= 1/9 + 1/9$

$= 2/9$

$X = 2 \rightarrow y = 2$
$X = -2 \rightarrow y = 2$

- $P(Z = 16) = P(X = 4 \text{ or } X = -4)$

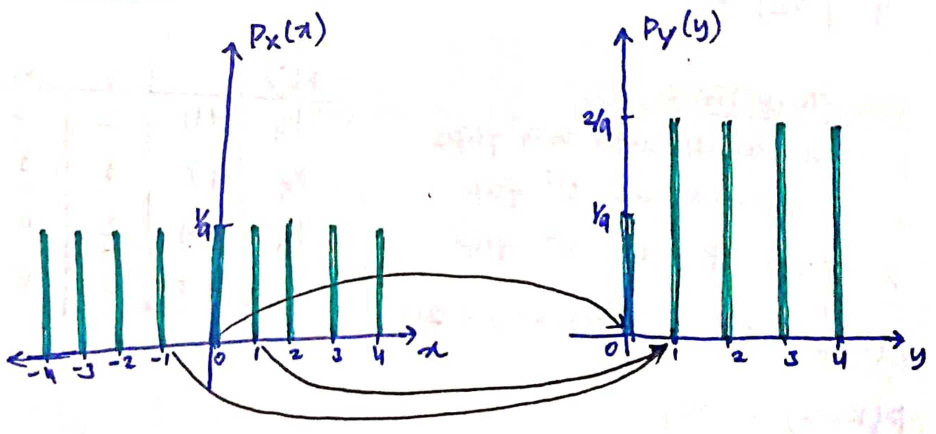
$= 1/9 + 1/9 = 2/9$

$X = 4 \rightarrow Z = 16$
$X = -4 \rightarrow Z = 16$

→ PMF for Z (prob. Mass function)

	Z	$P(\cdot)$
$X=0 \rightarrow$	0	1/9
$X=1, X=-1 \rightarrow$	1	2/9
$X=2, X=-2 \rightarrow$	4	2/9
$X=3, X=-3 \rightarrow$	9	2/9
$X=4, X=-4 \rightarrow$	16	2/9

→ in graphical representation



→ The PMFs of X and $y = |X|$

ques. let X and Z are 2 RV.

$y = X \oplus Z$. Find $P(y=y)$

	$X=0$	$X=1$
$P(X=x)$	p	$1-p$

	$Z=0$	$Z=1$
$P(Z=z)$	q	$1-q$

$X: \; P(X=0)=p \qquad P(X=1)=1-p \qquad y = X \oplus Z \qquad P(y=0)=?$

$Z: \; P(Z=0)=q \qquad P(Z=1)=1-q \qquad\qquad\qquad P(y=1)=?$

Method 1 →

X	Z	$y = X \oplus Z$	
0	0	0	pq
0	1	1	$p(1-q)$
1	0	1	$(1-p)q$
1	1	0	$(1-p)(1-q)$

$P(y=0) = pq + (1-p)(1-q)$

$P(y=1) = p(1-q) + (1-p)q$

Method 2 →

$P(y=0) = P\big((X=0, Z=0) \text{ or } (X=1, Z=1)\big)$ bcoz both events are disjoint

$= P(X=0, Z=0) + P(X=1, Z=1)$

bcoz X and Z are indep RV.

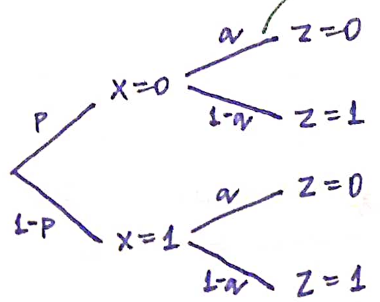
$= P(X=0) \cdot P(Z=0) + P(X=1) \cdot P(Z=1)$

$= pq + (1-p)(1-q)$

$||^{ly} \; P(y=1) = p(1-q) + (1-p)q$

Method 3 → <u>TREE Method</u>

$P(Z=0 \mid X=0) = P(Z=0)$: bcoz X and Z are independent



: $\boxed{y=0} = pq$

: $\boxed{y=1} = p(1-q)$

: $\boxed{y=1} = (1-p)q$

: $\boxed{y=0} = (1-p)(1-q)$

$\therefore P(y=0) = pq + (1-p)(1-q)$

$P(y=1) = p(1-q) + (1-p)q$

ques. variation :

$P(y=0 \mid X=0) = ?$

Method 1 →

X	Z	y
0	0	0
0	1	1
1	0	1
1	1	0

$P(y=0 \mid X=0)$

$= P(X \oplus Z = 0 \mid X=0)$

$= P(Z=0 \mid X=0) \rightarrow \dfrac{P(X=0, Z=0)}{P(X=0)} = \dfrac{pq}{p} = q$

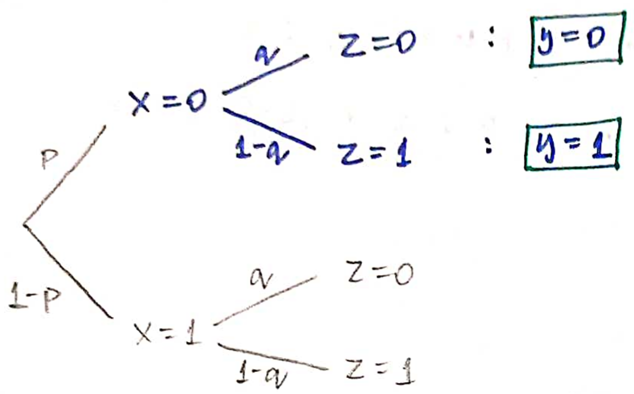
$= q$

$\boxed{\text{Method 2}}$:

	$x=0, z=0$	$x=0, z=1$	$x=1, z=0$	$x=1, z=1$
Earlier : prob.	pa	$p(1-a)$	$(1-p)a$	$(1-p)(1-a)$
updated when prob given $x=0$:	$\frac{pa}{p} = a$	$\frac{p(1-a)}{p} = (1-a)$	0	0

prob has been redistributed : $a + (1-a) = 1$ ✓

$\boxed{\text{Method 3}}$: TREE Method



$P(y=0 \mid x=0) = a$

$P(y=1 \mid x=0) = 1-a$

whatever written with pencil is unnecessary and has been removed.

$\boxed{\text{Method 4}}$: Using Plain Formula

$$P(y=0 \mid x=0) = \frac{P(y=0, x=0)}{P(x=0)} \leftarrow p$$

x & z are indep.
y & x are not indep.
so, we can't write
$P(y=0, x=0) \cancel{= P(y=0)} \cdot P(x=0)$

now, $P(y=0, x=0)$ ↘ bcoz $y = x \oplus z$

$= P(x \oplus z = 0, x=0)$

$= P(z=0, x=0) = \underset{a}{P(z=0)} \cdot \underset{p}{P(x=0)} = ap$

$$\therefore P(y=0 \mid x=0) = \frac{P(y=0, x=0)}{P(x=0)} = \frac{pa}{p} = a$$

ques. Consider 2 RV X and Y with joint prob distribution given in the table below:

	y=0	y=1	y=2
x=0	1/6	1/4	1/8
x=1	1/8	1/6	1/6

(i) $P(X=0, Y=1) = 1/4$

(ii) $P(X=0, Y \le 1) = 1/6 + 1/4 = 5/12$

PMF of X

(iii) $P(X=0) = 1/6 + 1/4 + 1/8 = 13/24$

(iv) $P(X=1) = 1/8 + 1/6 + 1/6 = 11/24$

$P(X=0) = \underset{1/6}{P(X=0, Y=0)} + \underset{1/4}{P(X=0, Y=1)} + \underset{1/8}{P(X=0, Y=2)}$

PMF of Y

(v) $P(Y=0) = 1/6 + 1/8 = 7/24$

(vi) $P(Y=1) = 1/4 + 1/8 = 3/8$

(vii) $P(Y=2) = 1/8 + 1/6 = 7/24$

(viii) $P(Y=1 | X=0) =$

Using the intution (Reduced domain)

	y=0	y=1	y=2
x=0	1/6 ↓ 4/13	1/4 ↓ 6/13	1/8 ↓ 3/13
~~x=1~~	~~1/8~~	~~1/6~~	~~1/6~~

← these are previous prob, they will be redistributed now

it is removed bcoz conditional prob is given

$P(Y=1 | X=0) = 6/13$

$\underbrace{\frac{1/4}{1/6 + 1/4 + 1/8}}_{\text{redistribution of prob.}} = \frac{1/4}{13/24}$

Using Formula directly:

$$P(Y=1|X=0) = \frac{P(Y=1, X=0)}{P(X=0)} = \frac{1/4}{13/24} = \frac{6}{13}$$

(ix) Are X and Y independent?

we need to check for all the values there is no other way

- $\underset{1/6}{P(X=0, Y=0)} = \underset{\frac{13}{24} \times \frac{7}{24}}{P(X=0) \cdot P(Y=0)}$ ✗ (1/6 ≠) no need to further check for other values. one counter example is sufficient } → X & Y are dependent

- $P(X=0, Y=1) = P(X=0) \cdot P(Y=1)$

- $P(X=0, Y=2) = P(X=0) \cdot P(Y=2)$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques. I toss a coin twice and define X to be the no of heads I observe. Then, I toss the coin 2 more times and define Y to be the no. of heads that I observe this time → last 2 tosses

Find $P((X<2)$ and $(Y>1))$.

	X	Y
HH	2	2
HT	1	1
TH	1	1
TT	0	0

← for simplicity

Actually I would be like this

	X	Y
HH HH	2	2
HH HT	2	1
HH TH	2	1
HH TT	2	0
⋮	⋮	⋮

$$P(X<2,\ Y>1) = P(\underset{\text{HT, TH, HH}}{X<2}) \cdot P(\underset{\text{HH}}{Y>1})$$ } → bcoz X & Y are independent bcoz what happen in 1^{st} 2 tosses doesnt effect what happen in 2^{nd} 2 tosses. Both are independent tosses.

$$= \frac{3}{4} \cdot \frac{1}{4}$$

$$= \frac{3}{16}$$

ques. Given the joint prob. distribution.

(a) Find distribution of X and Y.

Y \ X	1	2	3	4
1	1/16	0	1/8	1/16
2	1/32	1/32	1/4	0
3	0	1/8	1/16	1/16
4	1/16	1/32	1/16	1/32

distribution of Y (PMF of Y)

$P(Y=1) = P(Y=1, X=1) + P(Y=1, X=2) + P(Y=1, X=3) + P(Y=1, X=4)$

$= 1/16 + 0 + 1/8 + 1/16 = 1/4$

$P(Y=2) = 1/32 + 1/32 + 1/4 = 5/16$

$P(Y=3) = 1/8 + 1/16 + 1/16 = 1/4$

$P(Y=4) = 1/16 + 1/32 + 1/16 + 1/32 = 3/16$

||'ly we can find for X:

PMF of X

$P(X=1) = 1/16 + 1/32 + 0 + 1/16 = 5/32$

$P(X=2) = 0 + 1/32 + 1/8 + 1/32 = 3/16$

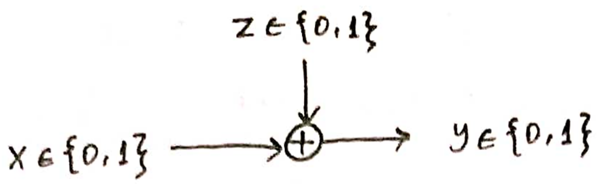
$P(X=3) = 1/8 + 1/4 + 1/16 + 1/16 = 1/2$

$P(X=4) = 1/16 + 0 + 1/16 + 1/32 = 5/32$

(b) Find $P(Y=1 \mid X=1)$

$$= \frac{P(Y=1, X=1)}{P(X=1)} = \frac{1/16}{5/32} = \frac{2}{5}$$

ques. consider the following binary channel



$P(x=0) = p$

$P(z=0) = q$

$y = (x+z) \bmod 2 = x \oplus z$, and x and z are independent. Find $P(y=0)$.

↳ $P(x=0) = p$

$P(z=0) = q$

$y = x \oplus z$; $P(y=0) = P(x \oplus z = 0)$

$= P(x=0, z=0) + P(x=1, z=1)$ } x & z are independent

$= pq + (1-p)(1-q)$

or using the table :

x	z	y	
0	0	0	pq
0	1	1	
1	0	1	
1	1	0	(1-p)(1-q)

$P(y=0) = pq + (1-p)(1-q)$

$P(y=0 \mid x=0) = P(x \oplus z = 0 \mid x=0)$ putting x=0 in $x \oplus z = 0$

$= P(0 \oplus z = 0) = P(z=0) = q$

$P(y=1 \mid x=0) = P(x \oplus z = 1 \mid x=0) = P(0 \oplus z = 1) = P(z=1) = 1-q$

$P(y=0 \mid x=1) = P(x \oplus z = 0 \mid x=1) = P(1 \oplus z = 0) = P(z=1) = 1-q$

$P(y=1 \mid x=1) = P(x \oplus z = 1 \mid x=1) = P(1 \oplus z = 1) = P(z=0) = q$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

GATE 17

Ques. Suppose X_i for $i=1,2,3$ are independent & identically distributed RV whose PMF are $P_r[X_i=0] = P_r[X_i=1] = \frac{1}{2}$ for $i=1,2,3$.

Define another RV, $Y = X_1X_2 \oplus X_3$, where $\oplus$ denotes XOR.

Then, $P_r[Y=0 \mid X_3=0] = ?$

↳ $Y = X_1X_2 \oplus X_3$ $\qquad P(X_i=1) = P(X_i=0) = 1/2$

$\boxed{\text{Method 1}}$ →

$P(Y=0 \mid X_3=0)$

$= P(X_1X_2 \oplus X_3 = 0 \mid X_3=0)$ $\xrightarrow{\text{or}}$ $P(X_1X_2 \oplus 0 = 0) = P(X_1X_2=0)$

$= P(X_1X_2 = 0 \mid X_3=0) = P(X_1X_2=0)$

Now, $P(X_1X_2=0) = P(X_1=0, X_2=0) + P(X_1=0, X_2=1) + P(X_1=1, X_2=0)$

$= 1/2 \cdot 1/2 + 1/2 \cdot 1/2 + 1/2 \cdot 1/2$ ← bcoz X_1, X_2 are indep.

$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

$\boxed{\text{Method 2}}$ → Reduced Domain

X_1	X_2	X_3	Y	initial prob.	prob redistributed
0	0	0	0	→ 1/8	→ 1/4
0	0	1	1		
0	1	0	0	→ 1/8	→ 1/4
0	1	1	1		
1	0	0	0	→ 1/8	→ 1/4
1	0	1	1		
1	1	0	1	→ 1/8	→ 1/4
1	1	1	0		

$\therefore P(Y=0 \mid X_3=0)$

$= \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4}$

$\boxed{\text{Method 3}}$ →

- $\xrightarrow{1/2}$ $X_3 = 0$
  - $\xrightarrow{1/2}$ $X_2 = 0$
    - $\xrightarrow{1/2}$ $X_1 = 0 \to Y = 0$
    - $\xrightarrow{1/2}$ $X_1 = 1 \to Y = 0$
  - $\xrightarrow{1/2}$ $X_2 = 1$
    - $\xrightarrow{1/2}$ $X_1 = 0 \to Y = 0$
    - $\xrightarrow{1/2}$ $X_1 = 1 \to Y = 1$
- $\xrightarrow{1/2}$ $X_3 = 1$
  - $X_2 = 0$
  - $X_2 = 1$

$P(Y=0) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4}$

$= \frac{3}{4}$

whatever written with pencil is unnecessary and has been removed.

$\boxed{\text{Method 4}} \rightarrow$

$$P(y=0 \mid x_3=0) = \frac{P(y=0,\ x_3=0)}{P(x_3=0)}$$

$$= \frac{P(x_1x_2 \oplus x_3 = 0,\ x_3=0)}{P(x_3=0)} =$$

$$= \frac{P(x_1x_2=0,\ x_3=0)}{P(x_3=0)}$$

$$= \frac{P(x_1x_2=0)\cdot \cancel{P(x_3=0)}}{\cancel{P(x_3=0)}}$$

$$= P(x_1x_2=0) = \frac{3}{4}$$

$\boxed{\text{Method 5}} \rightarrow$

x_1	x_2	x_3	y	
0	0	0	0	$\rightarrow 1/8$
0	0	1	1	$\rightarrow 1/8$
0	1	0	0	$\rightarrow 1/8$
0	1	1	1	$\rightarrow 1/8$
1	0	0	0	$\rightarrow 1/8$
1	0	1	1	$\rightarrow 1/8$
1	1	0	1	$\rightarrow 1/8$
1	1	1	0	$\rightarrow 1/8$

$$P(y=0 \mid x_3=0)$$

$$= \frac{P(y=0,\ x_3=0)}{P(x_3=0)}$$

$$= \frac{1/8 + 1/8 + 1/8}{1/2} = \frac{3/8}{1/2} = \frac{3}{4}$$

TYPES OF RANDOM VARIABLES

- **Discrete RV** → takes discrete values.

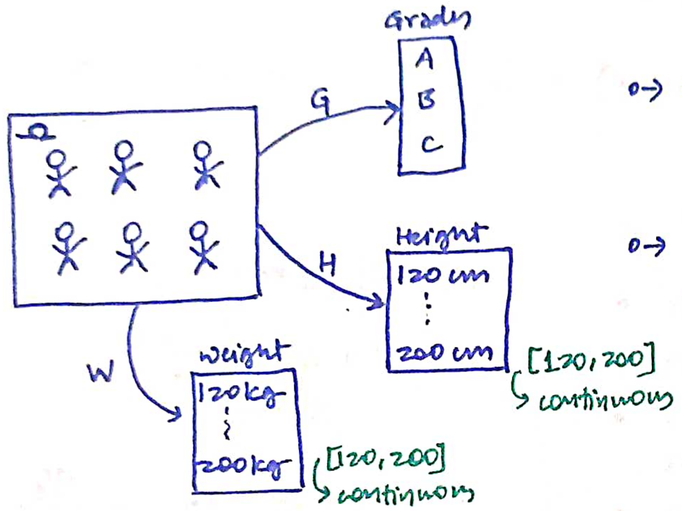
  A RV is called discrete if takes either finite or countably infinite no. of values.

  ex) The no. of sixes in the two rolls → 1, 2, 0

- **Continuous RV** → takes continuous values.

  A RV that can take continuous and uncountably infinite no. of values.

  ex) consider an experiment of choosing a point from the interval $[-1, 1]$.



→ Grades, Nationality are discrete values

  G → DRV

→ Height, weight are continuous values.

  H, W → CRV

→ When you think of a RV, immediately ask yourself
- what are the possible values?
- what are their probabilities?

(Intuitive) question

Write down the type of RV in the following table.

RV description	Range	DRV or CRV?
(i) X, the no. of heads in n flips of a fair coin.	$\{0, 1, \ldots, n\}$	DRV
(ii) N, the no. of people born this year.	$\{0, 1, 2, \ldots\}$	DRV
(iii) F, the no. of flips of a fair coin upto and including my first head.	$\{1, 2, 3, \ldots\}$ H, TH, TTH, TTT…H	DRV
(iv) B, the amount of time I wait for the next bus in seconds.	$[0, \infty)$ there could be partial seconds waited.	CRV
(v) C, the temp. in celcius of liquid water.	$(0, 100)$ the temp can be any real no. in this range	CRV

DISCRETE RANDOM VARIABLE

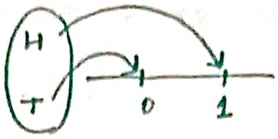
PROBABILITY MASS FUNCTION (PMF)

listing down prob. of each value for a DRV is called PMF.

(ex) Experiment : Flipping the coin once.

X = no. of heads in a flip.

		$P(\cdot)$
T	$X=0$	1/2
H	$X=1$	1/2



ques1. Suppose that a coin is tossed twice. let X represent the no. of heads that can come up. Write down PMF of X.

↳ X : 0, 1, 2

	X	$P(\cdot)$
TT →	0	1/4
HT, TH →	1	2/4 = 1/2
HH →	2	1/4

} PMF of X

	X
HH	2
HT	1
TH	1
TT	0

ques2. Check whether the funcⁿ given by

$P(X=x) = \frac{x+2}{25}$ for $x = 1, 2, 3, 4, 5$ is valid PMF of RV X.

↳ every prob. should be ≥ 0 and ≤ 1

↳ sum of all prob. should be 1.

↳ $P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5)$

$= \frac{3}{25} + \frac{4}{25} + \frac{5}{25} + \frac{6}{25} + \frac{7}{25}$

$= \frac{25}{25} = 1$ ✓

ques3. Suppose that a pair of fair dice are to be tossed, & let the RV X denote the sum of the points. Obtain the Prob distribution of X.

PMF for X

X	2	3	4	5	6	7	8	9	10	11	12
$P(\cdot)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

NOTE → Calculation of the PMF of a RV X.

For each possible value x of X :

1. collect all the possible outcome that give rise to the event $\{X = x\}$.
2. Add their probabilities to obtain $P(X=x)$.

ques 4. The RV X has prob. funcⁿ

$$P(X=x) = \begin{cases} kx & , x = 1,2 \\ k(x-2) & , x = 3,4,5 \end{cases}$$

where k is constant.

Find the value of k.

↳ $P(X=1) + P(X=2) + P(X=3) + P(X=4) + P(X=5) = 1$

$k + 2k + k + 2k + 3k = 1$

$9k = 1$

$\Rightarrow \boxed{k = 1/9}$

ques 5. Suppose that B, & C are DRV each taking values 1, 2 & 3, with joint PMF $P(B,C)$ given by the following table.

Note that one of the entries in the table is missing

B \ C	1	2	3
1	1/20	2/20	0
2	2/20	5/20	1/20
3	3/20	2/20	k ⟶ let it be 'k'

Determine the value of the missing entry $P(3,3)$

↳ $P(C=1) + P(C=2) + P(C=3) = 1$

$\left(\frac{1}{20} + \frac{2}{20} + \frac{3}{20}\right) + \left(\frac{2}{20} + \frac{5}{20} + \frac{2}{20}\right) + \left(0 + \frac{1}{20} + k\right) = 1$

$\frac{6}{20} + \frac{9}{20} + \frac{1}{20} + k = 1$

$k = 1 - \frac{16}{20} = \frac{4}{20}$ $\boxed{k = \frac{4}{20}}$

ques 6 Let X be a DRV taking the values $\{1, 2, \ldots, n\}$ all with equal prob.

Let Y be another DRV taking values in $\{1, 2, \ldots, n\}$.

Assume that X, Y are independent. Find $P(X=Y)$.

$X = 1, 2, \ldots, n$ each with prob. $1/n$

$Y = 1, 2, \ldots, n$ nothing is said about prob of Y

$P(X=Y) = P(X=Y=1) + P(X=Y=2) + P(X=Y=3) + \ldots + P(X=Y=n)$

$= P(X=1, Y=1) + P(X=2, Y=2) + \ldots + P(X=n, Y=n)$

$= P(X=1) \cdot P(Y=1) + P(X=2) \cdot P(Y=2) + \ldots + P(X=n) \cdot P(Y=n)$

$= \frac{1}{n} P(Y=1) + \frac{1}{n} P(Y=2) + \ldots + \frac{1}{n} P(Y=n)$

$= \frac{1}{n}\left[P(Y=1) + P(Y=2) + \ldots P(Y=n)\right] = \frac{1}{n} \times 1 = \frac{1}{n}$ ✓

(or) simply,

$$P(x=y) = \sum_{k=1}^{n} P(x=k, y=k) = \sum_{k=1}^{n} P(x=k)\, P(y=k)$$

$$= \frac{1}{n} \sum_{k=1}^{n} P(y=k) = \frac{1}{n}$$

(or)

$$P(X=Y) = \sum_{k=1}^{n} P(X=Y \mid Y=k)\, P(Y=k)$$

$$= \sum_{k=1}^{n} \frac{P(X=Y, Y=k)}{P(Y=k)}\, P(Y=k)$$

$$= \sum_{k=1}^{n} P(X=Y, Y=k) = \sum_{k=1}^{n} P(X=Y)\, P(Y=k) = \frac{1}{n} \sum_{k=1}^{n} P(Y=k) = \frac{1}{n}$$

ques 7 Let X and Y be DRV with Range(X) = {0, 1, 2} and Range(Y) = {1, 2} with joint distribution given by the chart below.

	X=0	X=1	X=2
Y=1	0.1	0.2	0.1
Y=2	0.3	0.2	0.1

What is the value of $P(X=1 \mid Y=2)$?

$$P(X=1 \mid Y=2) = \frac{P(X=1, Y=2)}{P(Y=2)} = \frac{0.2}{0.3+0.2+0.1} = \frac{0.2}{0.6} = 0.333.$$

ques 8 A DRV, Y, has probability mass func^n

$$P(Y=y) = c(y-3)^2, \quad y = -2, -1, 0, 1, 2.$$

Find the value of the constant c.

↳ writing the PMF as table we have

y	-2	-1	0	1	2
P(Y=y)	25c	16c	9c	4c	c

The sum of PMF is 1.

$$\Rightarrow \sum_{y=-2}^{2} P(Y=y) = 55c = 1$$

$$\Rightarrow \boxed{c = \frac{1}{55}}$$

PMF of a RV which is func^n of RV

ques 9 Suppose we roll a dice and let X be the value that is showing. Find PMF of X^2?

↳ PMF of X

X	$P(\cdot)$
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6

PMF of X^2

possible values of X^2

	X^2	$P(\cdot)$
$X=1 \rightarrow$	1	1/6
$X=2 \rightarrow$	4	1/6
$X=3 \rightarrow$	9	1/6
$X=4 \rightarrow$	16	1/6
$X=5 \rightarrow$	25	1/6
$X=6 \rightarrow$	36	1/6

$\rightarrow P(X^2=1) = P(X=1) = 1/6$

$\rightarrow P(X^2=4) = P(X=2) = \frac{1}{6}$

ques. 10 Suppose X has PMF $P(X=k) = \frac{k}{10}$ for $k = 1, 2, 3, 4$. Find the PMF of $X^2 - 4X + 3$.

PMF of X

X	$P(\cdot)$
1	1/10
2	2/10
3	3/10
4	4/10

$y = x^2 - 4x + 3$

$x=1 \Rightarrow y=0$

$x=2 \Rightarrow y=-1$

$x=3 \Rightarrow y=0$

$x=4 \Rightarrow y=3$

PMF of y

	y	$P(\cdot)$
$x=1$, $x=3 \rightarrow$	0	4/10
$x=2 \rightarrow$	-1	2/10
$x=4 \rightarrow$	3	4/10

$P(y=0) = P(x=1) + P(x=3) = \frac{1}{10} + \frac{3}{10} = \frac{4}{10}$

ques 11 PMF of X is given below:

$P(X=-2) = \frac{1}{8}$, $P(X=-1) = \frac{1}{4}$, $P(X=1) = \frac{3}{8}$, $P(X=2) = \frac{1}{4}$. Let $Y = X^2$. Find out PMF of Y.

PMF of X

X	$P(\cdot)$
-2	1/8
-1	1/4
1	3/8
2	1/4

$y = x^2$

$x=-2$, $x=2 \Rightarrow y=4$

$x=-1$, $x=1 \Rightarrow y=1$

PMF of y

	y	$P(\cdot)$
$x=1$, $x=-1$	1	$\frac{1}{4} + \frac{3}{8} = \frac{5}{8}$
$x=-2$, $x=2$	4	$\frac{1}{8} + \frac{1}{4} = \frac{3}{8}$

ques 12 $\quad P(X=x) = \dfrac{x^2}{a}\ ,\ x \in \{-3,-2,-1,1,2,3\}$

(a) Find a.

The sum of PMF is 1

$$\Rightarrow \quad P(X=-3) + P(X=-2) + P(X=-1) + P(X=1) + P(X=2) + P(X=3) = 1$$

$$\frac{9}{a} + \frac{4}{a} + \frac{1}{a} + \frac{1}{a} + \frac{4}{a} + \frac{9}{a} = 1$$

$$\frac{28}{a} = 1 \quad \Rightarrow \quad a = 28$$

(b) Find PMF of $Z = X^2$

$Z = X^2$	$P(\cdot)$	
9	$\frac{18}{a}$	$[P(X=3) + P(X=-3)]$
4	$\frac{8}{a}$	$[P(X=2) + P(X=-2)]$
1	$\frac{2}{a}$	$[P(X=1) + P(X=-1)]$

$$P(Z = X^2) = \begin{cases} \frac{18}{28}, & z = 9. \\ \frac{8}{28}, & z = 4. \\ \frac{2}{28}, & z = 1. \end{cases}$$

EXPECTATION of DRV

- A single no to summarize PMF.
- weighted average (in proportion to probabilities) of the possible values of x.

→ The PMF of a RV X provides us with several nos, the probablities of all the possible values of x. It would be desirable to summarize this info in a single representative no. This is accomplished by the expectation of x, which is a weighted (in proportion to probabilities) avg of the possible values of x.

Expectation of DRV : $E(x) = \sum k \cdot P(X=k)$

qus. If we toss a coin and X is 1 if we have heads and 0 if we have tails, what is the expectation of X ?

↳ $x = 1$ H

$x = 0$ T

$E[x] = ?$

x	P(·)
1	1/2
0	1/2

$k \cdot P(X=k)$

$E[x] = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$

$E[x] = 1 \cdot P(X=1) + 0 \cdot P(X=0)$

qus. A fair blue die has faces numbered 1, 1, 3, 3, 5 and 5. The RV B represents the score when the die is rolled. Find E(B).

$E[x] = \sum k \cdot P(X=k)$

↳

B	P(·)
1	2/6
3	2/6
5	2/6

$E[B] = 1 \cdot \frac{2}{6} + 3 \cdot \frac{2}{6} + 5 \cdot \frac{2}{6}$

$= \frac{2}{6}(1+3+5) = \frac{2}{6} \cdot 9 = 3$

{↳ $P(B = E[B]) = 2/6$ } but it is not useful. we may or may not get the prob of expected value

qus. In a game a fair die is rolled, and the amount of money won and lost is determined by the no. rolled. The amount are given in the table below:

Number	1	2	3	4	5	6
winnings	-8	3	1	-8	1	5

Find the expected amount of money won or lost
↳ E[x]

↳ X = amount of money won or lost

	x	P(·)
1	-8	1/6
2	3	1/6
3	1	1/6
4	-8	1/6
5	1	1/6
6	5	1/6

$E[x] = \frac{1}{6}(-8) + \frac{1}{6}(3) + \frac{1}{6}(1) + \frac{1}{6}(-8) + \frac{1}{6}(1) + \frac{1}{6}(5)$

$= \frac{1}{6}(-8+3+1-8+1+5)$

$= -1$

ques. let S and Q are 2 RV. Their joint prob distribution is given below :→

	S=0	S=1	S=2	
Q=0	0.1	0.1	0	0.2
Q=1	0.1	0.2	0	0.3
Q=2	0	0	0.5	0.5

→ $P(Q=0) = 0.2$

Find the $E[Q]$?

PMF for Q :

Q	P(·)
0	0.2
1	0.3
2	0.5

$$E(Q) = 0 \times 0.2 + 1 \times 0.3 + 2 \times 0.5$$
$$= 0.3 + 1$$
$$= 1.3$$

ques A biased die with 6 faces is rolled. The DRV X represents the score on the uppermost face. The prob. distribution of X is shown in the table below:

x	1	2	3	4	5	6
P(X=x)	a	a	a	b	b	0.3

Given that $E(x) = 4.2$ Find the value of a and b.

↳ Sum of all probabilities = 1

$\Rightarrow \quad 3a + 2b + 0.3 = 1$

$3a + 2b = 0.7$ — ①

$E(X) = 4.2$

$\Rightarrow 1 \cdot a + 2 \cdot a + 3 \cdot a + 4 \cdot b + 5 \cdot b + 6 \cdot 0.3 = 4.2$

$6a + 9b = 4.2 - 1.8$

$6a + 9b = 2.4$ — ②

$$\begin{aligned} 6a + 9b &= 2.4 \\ -6a \pm 4b &= -1.4 \\ \hline 5b &= 1 \end{aligned}$$

$\Rightarrow b = 1/5 = 0.20 \quad \Rightarrow a = \frac{0.7 - 0.4}{3} = 0.1$

$\boxed{a = 0.1, b = 0.2}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Revisiting prob definition

prob of head = 1/2 , then what does this mean ?

Do we mean that if we toss 2 times then we'll definitely be getting one head ?

2 times - 1 Head NO
4 times - 2 Heads NO
8 times - 4 Heads NO

→ it means, if you toss $\boxed{\text{enough}}$ no. of times then you'll be getting half of times head.

2000 times toss ⇝ 996 heads. {A mathematician did it}

SAMPLE probabilities v/s Real probabilities

let's suppose we have a biased coin,

$P(H) = 1/3$, $P(T) = 2/3$ } underlying prob. distribution — Real Probabilities (fix, Absolute, no one can change it)

Now, we start tossing the coin and we get : (This process is called sampling the data)

bcz we collected some samples.

H, H, T, T, T, H, T, T, T, H, H, T ⇝ $P(H) = \frac{5}{12}$ $P(T) = \frac{7}{12}$

H, H, T, T, T, H, T, T, T, H, H, T, T, H, T ⇝ $P(H) = \frac{6}{15}$ $P(T) = \frac{9}{15}$

} sample distribution

→ The sample distribution will keep on changing but if we collect enough samples, we'll keep getting closer to the Real underlying distribution and this is exactly what we desire, to approach the real probabilities.

underlying distribution : $P(H) = 1/3$
(no body knows) $P(T) = 2/3$

sample distribution
(calculated by sampling the data)
i.e. by doing the experiment and then calculating

tossed the coin 9000 times. → 2990 Heads, → 6010 Tails

$P(H) = \frac{2990}{9000}$ ≈ Real prob.

→ Why most companies are hungry for more and more data ?

↳ you have data ⟶ calculate (sample) prob.

Sample prob $\xrightarrow{\text{with more \& more data}}$ Real prob.

- in Real world, no one knows underlying distribution.
- we can just collect more & more data
- We calculate probabilities using data.
- if you have more data, you get better probabilities.

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

EXPECTATION v/s AVERAGE

What does expectation Really mean? How Expectation is different from Average?

Suppose you are playing a soccer game & the #goals you scored in successive matches are:

→ 1, 3, 6, 1, 3, 6, 5, 1, 5, 1, 3, 6

$$\text{Avg of all goals} = \frac{1+3+6+1+3+6+5+1+5+1+3+6}{12}$$

rewriting

$$= 1\left(\frac{4}{12}\right) + 3\left(\frac{3}{12}\right) + 6\left(\frac{3}{12}\right) + 5\left(\frac{2}{12}\right)$$

this is nothing but prob of scoring 5 goals in a match.

is it real prob? NO

it is sample prob.

if we take closer look, these are nothing but probabilities not real probabilities, but sample probabilities.

so, the Avg is nothing but, expectation, calculated with sample probabilities.

→ 1, 3, 6, 1, 3, 6, 5, 1, 5, 1, 3, 6, 5, 6, 1

$$\text{Avg of all goals} = \frac{1+3+6+1+3+6+5+1+5+1+3+6+5+6+1}{15}$$

rewriting

$$= 1\left(\frac{5}{15}\right) + 3\left(\frac{3}{15}\right) + 6\left(\frac{4}{15}\right) + 5\left(\frac{3}{15}\right)$$

Again this is avg, but if you look at it in rewritten form, it is expectation, calculated with sample probabilities.

sample probabilities.

→ for 12 samples : $1\left(\frac{4}{12}\right) + 3\left(\frac{3}{12}\right) + 6\left(\frac{3}{12}\right) + 5\left(\frac{2}{12}\right) \rightarrow$ <u>Avg</u>.

15 samples : $1\left(\frac{5}{15}\right) + 3\left(\frac{3}{15}\right) + 6\left(\frac{4}{15}\right) + 5\left(\frac{3}{15}\right) \rightarrow$ <u>Avg</u>

if you collect more and more samples (the sample probabilities will tend to real probabilities)

Real probabilities.

$1 \cdot P(1) + 3 \cdot P(3) + 6 \cdot P(6) + 5 \cdot P(5) \Rightarrow$ <u>Expectation</u>

Sample probabilities $\xrightarrow{\text{more \& more data}}$ Real probabilities

Avg $\xrightarrow{\text{enough data}}$ Expectation

→ Expectation is just Average on enough data.

but for our purpose, we just say that Expectation is just Avg.

↑ to avoid getting into details as we assume we have enough data.

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ex ~~If~~ we roll a die, X = Number appears on top

$E[X] = ?$

$X =$	1	2	3	4	5	6
$P(\cdot) =$	1/6	1/6	1/6	1/6	1/6	1/6

$$E[X] = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6}$$

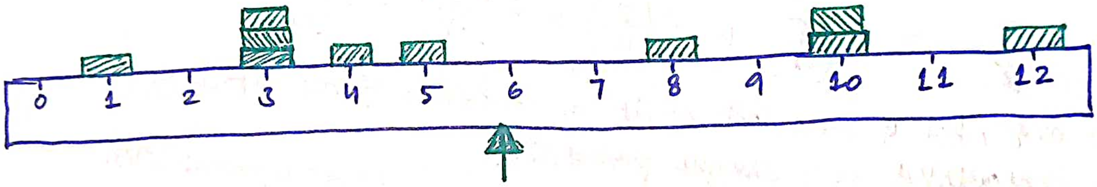
$$= \frac{1+2+3+4+5+6}{6} = 3.5$$

This 3.5 intutively means that

- if you roll a die, on an avg you'll get 3.5.
  i.e if you roll a die and take avg of all the outputs it will come out to be 3.5.
- if you roll a die enough no. of times, the avg of all the outputs will be 3.5.

∘ you wont even get 3.5 ever, you'll just get integers {1, 2, – 6} but if you take avg of all, then you'll get 3.5.

→ EXPECTATION as CENTRE OF MASS (Physics)



$P(1) = \frac{1}{10}$ $P(3) = \frac{3}{10}$ $P(4) = \frac{1}{10}$ $P(5) = \frac{1}{10}$ $P(8) = \frac{1}{10}$ $P(10) = \frac{2}{10}$ $P(12) = \frac{1}{10}$

$$E[X] = 1\times\frac{1}{10} + 3\times\frac{3}{10} + 4\times\frac{1}{10} + 5\times\frac{1}{10} + 8\times\frac{1}{10} + 10\times\frac{2}{10} + 12\times\frac{1}{10}$$

$$= \frac{1+9+4+5+8+20+12}{10} = \frac{59}{10} = 5.9$$

ques. T/F if $a \le X \le b$ then $a \le E[X] \le b$.

TRUE

expectation is nothing but avg.
Avg will always be somewhere b/w the Range of nos & never outside that Range.

ex

→ intuitively, expectation will be closer to 1, bcoz 1 is very heavy.

X	$P(\cdot)$
1	0.9
2	0.09
12	0.01

$$\left.\begin{aligned} E[X] &= 1\times 0.9 + 2\times 0.09 + 12\times 0.01 \\ &= 1.20 \end{aligned}\right\}\text{calculated.}$$

very close to 1.
bcoz $P(1)$ is very high.

MIT Book Question

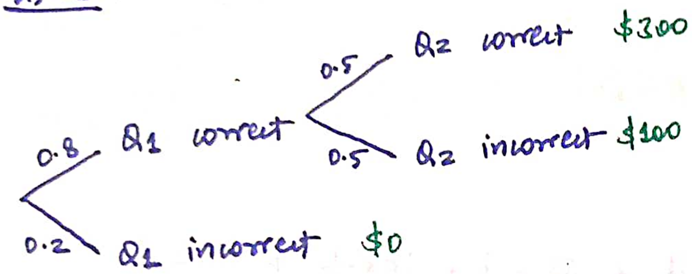
Consider a quiz where a person is given 2 questions & must decide which question to answer 1st. Ques 1 will be answered correctly with a prob. of 0.8 and the person will then receive as price \$100, while ques. 2 will be answered correctly with prob 0.5, & the person will then receive as price \$200.

If the 1st ques attempted is answered incorrectly, the quiz terminates, i.e, the person is not allowed to attempt the 2nd ques. If the 1st ques is answered correctly, the person is allowed to attempt the 2nd ques.

Which ques to be answered 1st to maximize the expected value of the total prize received?

Q1 :	0.8	100 \$
Q2 :	0.5	200 \$

Case 1 :



• $\boxed{\text{Q1 answered 1st}}$

X : Amount won.

PMF of X :

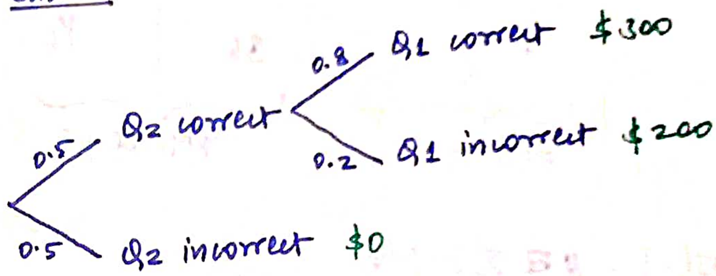
X	P(·)
0	0.2
100	0.8×0.5 = 0.4
300	0.8×0.5 = 0.4

$\therefore E[X] = 0\times0.2 + 100\times0.4 + 300\times0.4$

$= \$160.$

∴ if we start with ques 1, expected money won = \$160

Case 2 :



• $\boxed{\text{Q2 answered 1st}}$

PMF of X :

X.	P(·)
0	0.5
200	0.5×0.2 = 0.1
300	0.5×0.8 = 0.4

$E[X] = 0\times0.5 + 200\times0.1 + 300\times0.4$

$= \$140.$

if we start with ques 2, expected money won = \$140.

∴ It is better to start with Q1. ← Answer

*A very similar question is asked by GATE in 2021. take a look at it also.

Expectation of RV which is a funcⁿ of RV

ques. Suppose we roll a die & let X be the value that is showing. We want to find the expectation $E[X^2]$?

X	$P(X=k)$	X^2	$P(X^2=k)$
1	1/6	1	1/6
2	1/6	4	1/6
3	1/6	9	1/6
4	1/6	16	1/6
5	1/6	25	1/6
6	1/6	36	1/6

PMF of X — PMF of X^2

$$E[y] = \sum y \cdot P(y)$$

$$E[X^2] = \sum k P(X^2=k) = 1\times\frac{1}{6} + 4\times\frac{1}{6} + 9\times\frac{1}{6} + 25\times\frac{1}{6} + 16\times\frac{1}{6} + 36\times\frac{1}{6}$$

$$E[X^2] = \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$$

$\boxed{\text{Method 2}}$ →

$$E[X^2] = \sum_k X^2 \cdot P(X=k)$$

$$\Rightarrow E[X^2] = \sum_k X^2 \cdot P(X=k)$$

$$= 1^2\cdot P(X=1) + 2^2\cdot P(X=2) + 3^2\cdot P(X=3)$$
$$4^2\cdot P(X=4) + 5^2\cdot P(X=5) + 6^2\cdot P(X=6)$$

$$= 1^2\cdot\frac{1}{6} + 2^2\cdot\frac{1}{6} + 3^2\cdot\frac{1}{6} + 4^2\cdot\frac{1}{6} + 5^2\cdot\frac{1}{6} + 6^2\cdot\frac{1}{6}$$

$$= \frac{1+4+9+16+25+36}{6} = \frac{91}{6}$$

$$E[y] = \sum_k y \cdot P(X=k)$$

$$E[\sin x] = \sum_k \sin x \cdot P(X=k)$$

$$E[x^2+3x+2] = \sum_k (x^2+3x+2)\cdot P(X=k)$$

$$E[g(x)] = \sum_k g(x) \cdot P(X=k)$$

this stays the same no need to modify for every funcⁿ.

i.e we dont need to find the PMF g(x) to plug in the values in calculating $E[g(x)]$.

ques.

X	$P(\cdot)$
-2	$1/5$
-1	$1/5$
0	$1/5$
1	$1/5$
2	$1/5$

$E[X^2] = ?$

$\boxed{\text{Method 1}} \to$ possible values of $X^2 = 0, 1, 4$

finding PMF of X^2

X^2	$P(\cdot)$
0	$1/5$
$-1, 1 \to 1$	$2/5$
$-2, 2 \to 4$	$2/5$

$$E[X^2] = 0 \cdot \frac{1}{5} + 1 \cdot \frac{2}{5} + 4 \cdot \frac{2}{5}$$
$$= 2$$

$\boxed{\text{Method 2}} \to$

$$E[X^2] = \sum_k x^2 P(X=k)$$
$$= (-2)^2 \cdot P(X=-2) + (-1)^2 P(X=-1) + 0^2 \cdot P(X=0) + (1)^2 \cdot P(X=1) + 2^2 \cdot P(X=2)$$
$$= 4 \cdot \frac{1}{5} + 1 \cdot \frac{1}{5} + 0 \cdot \frac{1}{5} + 1 \cdot \frac{1}{5} + 4 \cdot \frac{1}{5} = 2$$

$\boxed{\text{NOTE}} \to$ Expected value Rule for Func^n's of RV

Let X be a RV with PMF $P_X(x)$, and $g(X)$ be a real valued func^n of X.

Then, the expected value of the RV $g(X)$ is given by

$$E[g(X)] = \sum_x g(x) P_X(x)$$

ques. PMF of X is given below:

$P(X=-2) = \frac{1}{8}$ $\quad P(X=-1) = \frac{1}{4}$ $\quad P(X=1) = \frac{3}{8}$ $\quad P(X=2) = \frac{1}{4}$.

Let $Y = X^2$. Find out expectation of Y.

$$E[X^2] = \sum_k x^2 P(X=k)$$
$$= (-2)^2 \cdot \frac{1}{8} + (-1)^2 \cdot \frac{1}{4} + (1)^2 \cdot \frac{3}{8} + (2)^2 \cdot \frac{1}{4}$$
$$= 4 \cdot \frac{1}{8} + \frac{1}{4} + \frac{3}{8} + 4 \cdot \frac{1}{4} = \frac{17}{8}$$

ques. let X be a DRV with $P_X(k) = \frac{1}{5}$ for $k = -1, 0, 1, 2, 3$. let $y = 2|X|$.

Find the range and PMF of y.

↳ $y = 2|X|$ ⇒ ~~Range of y~~ :

$\cancel{\{0, 2, 4, 6\}}$

$X = -1 \Rightarrow y = 2|-1| = 2$

$X = 0 \Rightarrow y = 2|0| = 0$

$X = 1 \Rightarrow y = 2|1| = 2$

$X = 2 \Rightarrow y = 2|2| = 4$

$X = 3 \Rightarrow y = 2|3| = 6$

PMF of y

y	$P(\cdot)$
0	$1/5$
2	$1/5 + 1/5 = 2/5$
4	$1/5$
6	$1/5$

$X = -1$

$X = 2$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

EXPECTATION Tricky Questions

LAW OF TOTAL EXPECTATION :

$$E[X] = E_1 P(A) + E_2 P(A^c)$$

we can break the expectation into 2 parts just like law of total probability

- $A \to E_1$ if A happens expectation is E_1
- $A^c \to E_2$ if A doesn't happens expectation is E_2

ques 1

- $\frac{3}{6} = \frac{1}{2}$: >3 (4, 5, 6) : +100 rupees
- $\frac{3}{6} = \frac{1}{2}$: ≤ 3 (1, 2, 3) : −50 rupees

$$\Rightarrow E = \frac{1}{2}(100) + \frac{1}{2}(-50)$$

$$= \frac{50}{2} = 25 \text{ rupees}$$

ques 2

prof

- $2/5$: Alice : 10 errors
- $1/5$: Claire : 50 errors
- $2/5$: Betsy : 20 errors

$$\Rightarrow E[X] = \frac{2}{5}(10) + \frac{1}{5}(50) + \frac{2}{5}(20)$$

$$= 22 \text{ errors}$$

ques 3 Expected height of randomly chosen male = $5'11'' = 5 + \frac{11}{12}$ feet

———— " ———— female = $5'5'' = 5 + \frac{5}{12}$ feet

$P(\text{male}) = \frac{49.8}{100} = 0.498$ $\quad P(\text{female}) = 0.502$

- 0.498 : Male : $5 + \frac{11}{12}$ feet
- 0.502 : Female : $5 + \frac{5}{12}$ feet

$$\Rightarrow E = 0.498 \times \left(5 + \frac{11}{12}\right) + 0.502\left(5 + \frac{5}{12}\right)$$

$$= 5.665 \text{ ft} \approx 5'8''$$

ques 4

1	2
3	4

- only pick ①, ② → −10 rupees
- pick ③, ④ atleast once → +10 rupees

1st ball, 2nd ball, Amt won/lost

- 1st ball 1 ($1/4$):
  - $1/3$: 2 → −10
  - $1/3$: 3 → 10
  - $1/3$: 4 → 10
- 1st ball 2 ($1/4$):
  - $1/3$: 1 → −10
  - $1/3$: 3 → 10
  - $1/3$: 4 → 10
- 1st ball 3 ($1/4$) → 10
- 1st ball 4 ($1/4$) → 10

no need to divide further as we are already winning ₹10 no matter what

$E[X] =$

$$\left[\cancel{\frac{1}{4}\cdot\frac{1}{3}(-10)} + \cancel{\frac{1}{4}\cdot\frac{1}{3}(10)} + \frac{1}{4}\cdot\frac{1}{3}\cdot 10\right]$$

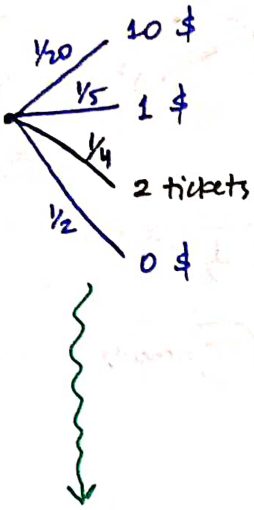
$$+ \left[\cancel{\frac{1}{4}\cdot\frac{1}{3}(-10)} + \cancel{\frac{1}{4}\cdot\frac{1}{3}(10)} + \frac{1}{4}\cdot\frac{1}{3}\cdot 10\right]$$

$$+ \left[\frac{1}{4}\cdot 10\right] + \left[\frac{1}{4}\cdot 10\right]$$

$$E[X] = \frac{10}{12} + \frac{10}{12} + \frac{10}{4} + \frac{10}{4} = \frac{20}{3} = 6.67 \text{ rupees}$$

Recursive way to Find Expectation

ques 5. $E \rightarrow$ expected no. of prize from one ticket

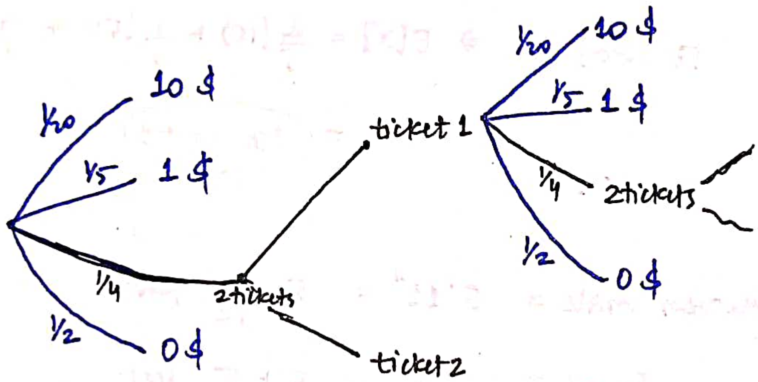


$$E = \frac{1}{20}\times 10 + \frac{1}{5}\times 1 + \frac{1}{4}\times 2E + \frac{1}{2}\times 0$$

$$\Rightarrow E = \frac{1}{2} + \frac{1}{5} + \frac{E}{2} + 0$$

$$\Rightarrow \frac{E}{2} = \frac{7}{10}$$

$$\Rightarrow \boxed{E = 1.4\ \$}$$



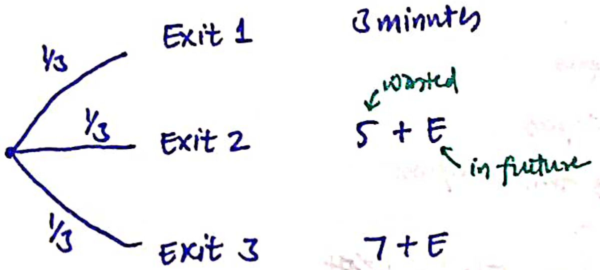
E: expected money with the one ticket

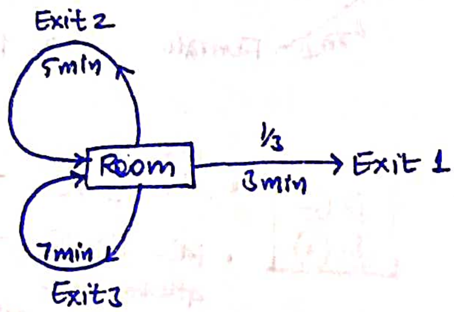
if we are winning 2 tickets, itself then we are winning 2E $.

$$E = \frac{1}{20}(10) + \frac{1}{5}(1) + \frac{1}{4}\underbrace{\boxed{E_1}}_{2E} + \frac{1}{2}\times 0 \qquad E_1 = 2E$$

$$\boxed{E = 1.4\ \$}$$

ques 6. E : expected time to exit from room $\left\{\text{time taken for mouse to leave maze}\right\}$





$$\Rightarrow E = \frac{1}{3}\times 3 + \frac{1}{3}(5+E) + \frac{1}{3}(7+E)$$

$$\Rightarrow 3E = 15 + 2E$$

$$\Rightarrow \boxed{E = 15 \text{ minutes}}$$

ques 7 Consider the following recursive func^n.

```
int near()
{   int b = randomInteger(1,4);
    if (b==1)      return 2;
else if (b==2)     return 4;
else if (b==3)     return (6 + near());
else    return (8 + near());
}
```

Let, y = value returned by near().

What is $E[y]$?

↳

prob	b	y
1/4	1	2
1/4	2	4
1/4	3	6+E
1/4	4	8+E

$$E = \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 4 + \frac{1}{4}\cdot(6+E) + \frac{1}{4}\cdot(8+E)$$

$$E = \frac{2+4+(6+E)+(8+E)}{4}$$

$$4E = 2E + 20$$

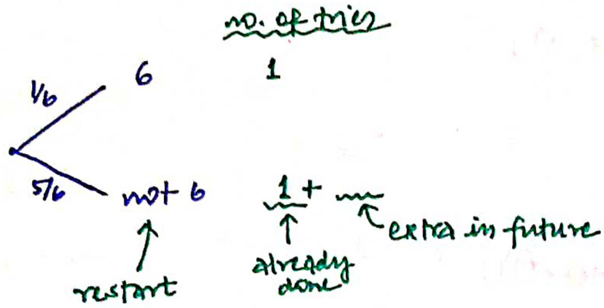
$$\boxed{E = 10}$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques 8. X : Number of tries to get 6

$E[X] = ?$

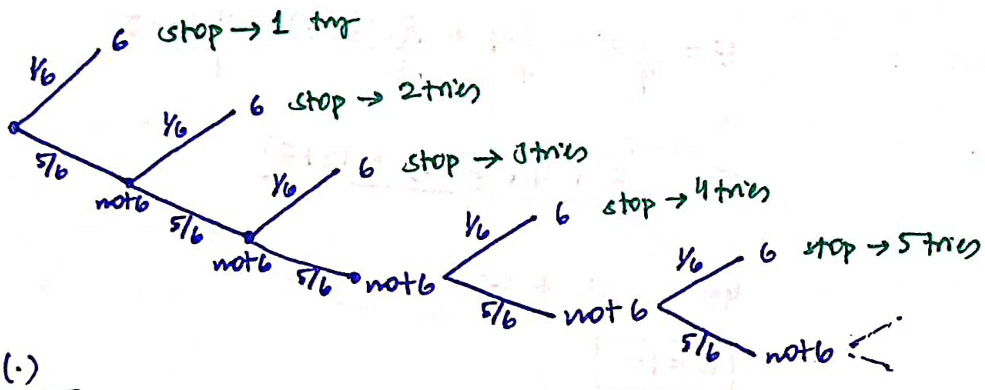
Method 1 :



$$\Rightarrow E = \frac{1}{6}\cdot 1 + \frac{5}{6}(1+E)$$

$$E = \frac{1}{6} + \frac{5}{6} + \frac{5}{6}E \Rightarrow \boxed{E=6}$$

Method 2 :



X	P(·)
1	$1/6$
2	$\frac{5}{6}\cdot\frac{1}{6}$
3	$\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}$
4	$\left(\frac{5}{6}\right)^3\cdot\frac{1}{6}$
⋮	⋮

$$\Rightarrow E = 1\cdot\frac{1}{6} + 2\cdot\frac{5}{6}\cdot\frac{1}{6} + 3\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6} + 4\cdot\left(\frac{5}{6}\right)^3\cdot\frac{1}{6} + \cdots$$

$$E = \frac{1}{6}\underbrace{\left[1 + 2\cdot\frac{5}{6} + 3\cdot\left(\frac{5}{6}\right)^2 + 4\cdot\left(\frac{5}{6}\right)^3 + \cdots\right]}_{AGP}$$

$$x = 5/6$$

$$E = \frac{1}{6}\left[1 + 2x + 3x^2 + 4x^3 + \cdots\right]$$

$$S = 1 + 2x + 3x^2 + 4x^3 + \cdots$$

$$xS = \quad x + 2x^2 + 3x^3 + \cdots$$

$$S - xS = 1 + x + x^2 + x^3 + \cdots$$

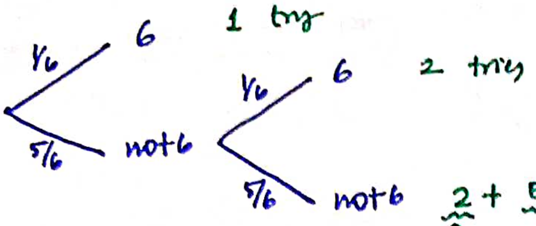
$$S(1-x) = \frac{1}{1-x} \Rightarrow S = \frac{1}{(1-x)^2}$$

$$\Rightarrow E = \frac{1}{6}\cdot S$$

$$E = \frac{1}{6}\left(\frac{1}{(1-x)^2}\right)$$

$$E = \frac{1}{6}\,\frac{1}{(1/6)^2} \Rightarrow E = \frac{1}{6}\times 6^2 \Rightarrow \boxed{E=6}$$

Method 3:



$\underbrace{\underset{\text{wasted}}{2} + \underset{\text{future}}{E}}$

- '2' tosses you have already done
- you will do 'E' more tosses to get '6'.

$E = \frac{1}{6}\cdot 1 + \frac{5}{6}\cdot\frac{1}{6}\cdot 2 + \frac{5}{6}\cdot\frac{5}{6}\cdot(2+E)$

$\Rightarrow E = \frac{1}{6} + \frac{10}{6^2} + \frac{25}{6^2}(2+E)$

$\Rightarrow 36E = 6 + 10 + 50 + 25E$

$11E = 66$

$\boxed{E = 6}$

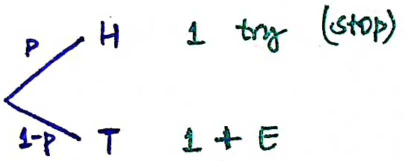
ques 9 GATE IT 2006

$$\left\{\begin{array}{ll} H & N=1 \\ TH & N=2 \\ TTH & N=3 \\ TTTH & N=4 \end{array}\right\}$$

$P(H) = p$

N : no. of tosses till the 1st head appears

Method 1:



$E = p\cdot 1 + (1-p)(1+E)$

$\cancel{E} = \cancel{p} + 1 - \cancel{p} + \cancel{E} - Ep$

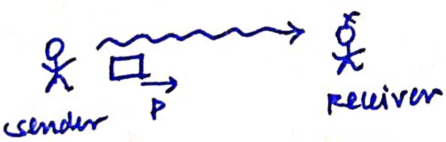
$Ep = 1 \Rightarrow \boxed{E = 1/p}$

NOTE → Template of such questions

Head or getting 6 or anything $\Rightarrow$ success

- $P(\text{success}) = p$
- How many expected trials s.t. you get succeeded? $= \frac{1}{p}$ ✓

eg GATE PYQ in CN



sender is sending data pkt to the receiver with probability = p

P(receiving packet successfully) = p

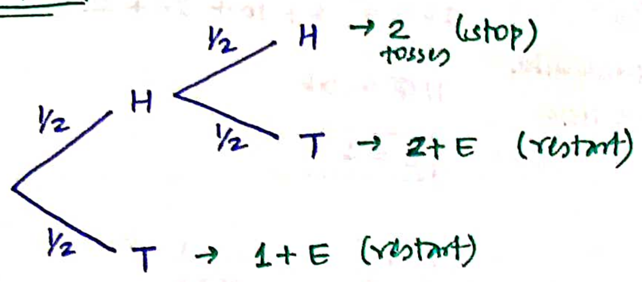
Expected no. of tries (packets) to send just 1 packet successfully $= \frac{1}{p}$ ✓

ques 10. TIFR 2015 A-6

$P(H) = P(T) = 1/2$

Two heads in a Row :

$\boxed{\text{Method 1}}$:



$$E = \frac{1}{2}\cdot\frac{1}{2}\cdot 2 + \frac{1}{2}\cdot\frac{1}{2}\cdot(2+E) + \frac{1}{2}\cdot(1+E)$$

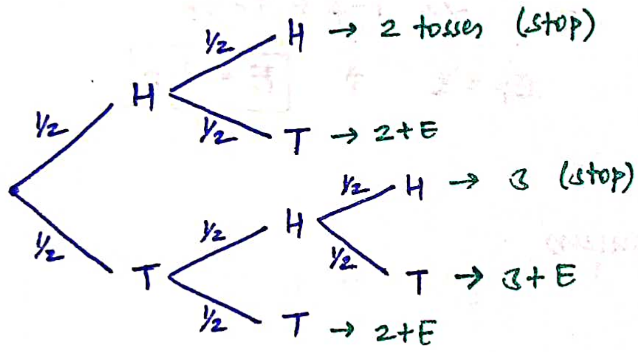
$$E = \frac{1}{2} + \frac{1}{2} + \frac{E}{4} + \frac{1}{2} + \frac{E}{2}$$

$$E = 1 + \frac{1}{2} + \frac{3E}{4}$$

$$\frac{E}{4} = \frac{3}{2} \Rightarrow \boxed{E = 6}$$

$\boxed{\text{Method 2}}$: infinite series

$\boxed{\text{Method 3}}$: (partially growing tree)



$$E = \frac{1}{2}\cdot\frac{1}{2}\cdot 2 + \frac{1}{2}\cdot\frac{1}{2}(2+E) + \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot 3 + \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot(3+E) + \frac{1}{2}\cdot\frac{1}{2}\cdot(2+E)$$

$$E = \frac{1}{2} + \frac{1}{2}(2+E) + \frac{3}{8} + \frac{3}{8} + \frac{E}{8}$$

$$E = \frac{1}{2} + 1 + \frac{E}{2} + \frac{3}{4} + \frac{E}{8}$$

$$E - \frac{5E}{8} = \frac{9}{4} \Rightarrow \frac{3E}{8} = \frac{9}{4} \Rightarrow \boxed{E = 6}$$

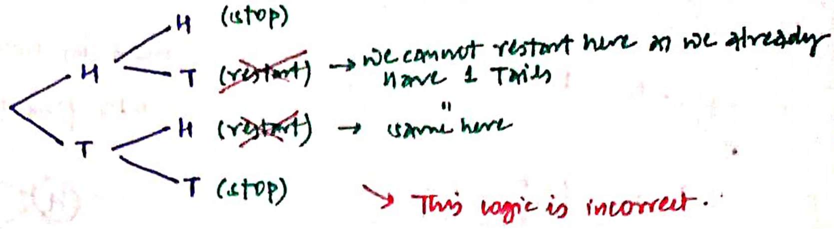
ques 11. GATE 2015

Same ques as 10. $\boxed{\text{Ans : } E[X] = 6 \text{ tosses}}$

ques12. GATE IT 2005

either 2 heads or 2 tails

HH or TT



$\frac{1}{2}$ H → $1 + E_1$

$\frac{1}{2}$ T → $1 + E_2$

$\Rightarrow E = \frac{1}{2}(1+E_1) + \frac{1}{2}(1+E_2)$

E_1 : tosses we need for HH or TT if I already have Head in my hand.

E_2 : tosses we need for HH or TT if I already have Tail in my hand.

now, we'll calculate these E_1 and E_2,

E_1 :

(H): $\frac{1}{2}$ H → 1 ; $\frac{1}{2}$ T → $1+E_2$

$\Rightarrow E_1 = \frac{1}{2}\cdot 1 + \frac{1}{2}\cdot(1+E_2)$

$E_1 = 1 + \frac{E_2}{2}$ — (1)

E_2 :

(T): $\frac{1}{2}$ H → $1+E_1$; $\frac{1}{2}$ T → 1

$\Rightarrow E_2 = \frac{1}{2}\cdot(1+E_1) + \frac{1}{2}\cdot 1$

$E_2 = 1 + \frac{E_1}{2}$ — (2)

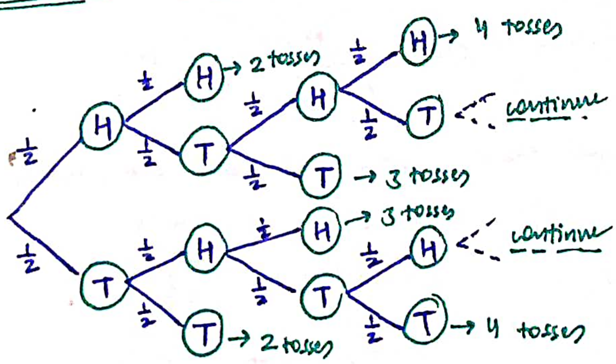
using (1), (2) we get

$E_1 = 2$, $E_2 = 2$

now, put the values in : $E = \frac{1}{2}(1+E_1) + \frac{1}{2}(1+E_2)$

$= \frac{1}{2}\cdot 3 + \frac{1}{2}\cdot 3 = 3 \Rightarrow \boxed{E = 3}$

$\boxed{\text{Method 2}}$: expand till infinity



no. of tosses	Prob.
2	$2\left(\frac{1}{2}\right)^2$
3	$2\cdot\left(\frac{1}{2}\right)^3$
4	$2\cdot\left(\frac{1}{2}\right)^4$
⋮	⋮

$$E = 2\Big[\underbrace{2\times\left(\tfrac{1}{2}\right)^2 + 3\times\left(\tfrac{1}{2}\right)^3 + 4\times\left(\tfrac{1}{2}\right)^4 + \cdots}_{S}\Big]$$

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

$\boxed{E = 3}$

$S = 2\left(\frac{1}{2}\right)^2 + 3\left(\frac{1}{2}\right)^3 + \cdots\cdots$

$\frac{1}{2}S = \quad 2\left(\frac{1}{2}\right)^3 + \cdots$

$\frac{S}{2} = 2\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 + \cdots$

$\Rightarrow \frac{S}{2} = \frac{1}{2} + \frac{1/8}{1-1/2} \Rightarrow \frac{S}{2} = \frac{3}{4} \Rightarrow S = \frac{3}{2}$

ques 13. No. of tosses until you get HT

$\frac{1}{2}$ H $\quad 1+E_1$

$\frac{1}{2}$ T $\quad 1+E$ (restart)

E_1: we have head in hand, and looking for HT.

let's find E_1,

(H) H $\rightarrow 1+E_1$

T $\rightarrow 1$

$$E_1 = \frac{1}{2}(1+E_1) + \frac{1}{2}\cdot 1$$

$$E_1 = 1 + \frac{E_1}{2}$$

$$E_1 = 2$$

$$\therefore E = \frac{1}{2}(1+E_1) + \frac{1}{2}(1+E)$$

$$E = \frac{1}{2}(1+2) + \frac{1}{2} + \frac{E}{2}$$

$$E = \frac{3}{2} + \frac{1}{2} + \frac{E}{2}$$

$$\boxed{E = 4}$$ ✓

o→ For HT or TH, we need how many tosses?

(start) H $\quad 1+E_1$

T $\quad 1+E_2$

E_1 : (H) H $\rightarrow 1+E_1$

T $\rightarrow 1$

$$E_1 = \frac{1}{2}(1+E_1) + \frac{1}{2}\cdot 1$$

$$E_1 = 2$$

E_2 : (T) H $\rightarrow 1$

T $\rightarrow 1+E_2$

$$E_2 = \frac{1}{2}\cdot 1 + \frac{1}{2}(1+E_2)$$

$$E_2 = 2$$

$$E = \frac{1}{2}(1+E_1) + \frac{1}{2}(1+E_2)$$

$$= \frac{1}{2}\cdot 3 + \frac{1}{2}\cdot 3$$

$$\boxed{E = 3}$$ ✓

NOTE :→ no. of tosses for :→

- TT → 6
  HH → 6
- HH or TT → 3
- HT → 4
- TH → 4
- HT or TH → 3

CUMULATIVE DISTRIBUTION FUNCTION (CDF)

The CDF of a RV 'X' may be defined as the probability that the RV 'X' takes a value "less than or equal to x".

$$\boxed{\text{CDF}: F_X(x) = P(X \le x)}$$

- CDF can be defined for discrete and continuous RV.

denoted as : $F_X(x) / F(x)$

$$\boxed{F(X) = P(X \le x)}$$

(ex) $X = 1, 2, 3, 4$

$$P(X \le 3) = P(X=1) + P(X=2) + P(X=3)$$

ques. Find the cummulative distribution func^n of the total of heads obtained in 4 tosses of a balanced coin.

↳ X = no. of heads in 4 tosses

$$P(X=0) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \left(\frac{1}{2}\right)^4$$

(TTTT)

$$P(X=1) = {}^4C_1 \cdot \left(\frac{1}{2}\right)^4$$

HTTT
THTT
TTHT
TTTH

$$P(X=2) = {}^4C_2 \left(\frac{1}{2}\right)^4$$

_ _ _ _

choose 2 positions to put 'H'

$$P(X=3) = {}^4C_3 \left(\frac{1}{2}\right)^4$$

$$P(X=4) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \left(\frac{1}{2}\right)^4$$

HHHH

⇒

X	P(·)	F
0	$(1/2)^4$	$F(0) = P(X \le 0) = P(0) = \frac{1}{16}$
1	$4\cdot(1/2)^4$	$F(1) = P(X \le 1) = P(0) + P(1) = 1/16 + 4/16 = 5/16$
2	$6\cdot(1/2)^4$	$F(2) = P(X \le 2) = P(0) + P(1) + P(2) = 11/16$
3	$4\cdot(1/2)^4$	$F(3) = P(X \le 3) = P(0) + P(1) + P(2) + P(3) = 15/16$
4	$(1/2)^4$	$F(4) = P(X \le 4) = P(0) + P(1) + P(2) + P(3) + P(4) = 1$

Hence, the cumulative distribution func^n is given by

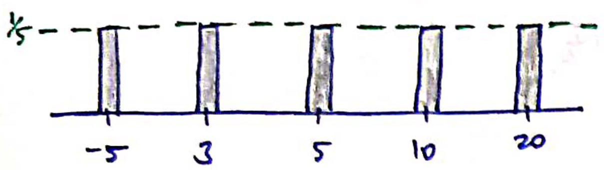
$$F(x) = \begin{cases} 0 & \text{for } x < 0, \\ 1/16 & \text{for } 0 \le x < 1, \\ 5/16 & \text{for } 1 \le x < 2, \\ 11/16 & \text{for } 2 \le x < 3, \\ 15/16 & \text{for } 3 \le x < 4, \\ 1 & \text{for } x \ge 4 \end{cases}$$

Observe that this distribution is defined not only for the values taken on the given RV, but for all real numbers.

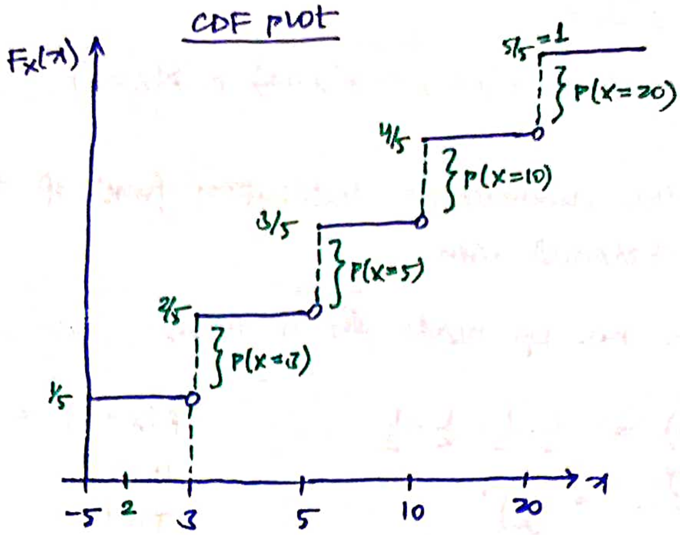
For instance, we can write $F(1.7) = \frac{5}{16}$ and $F(100) = 1$, $F(0.5) = 1/16$ although the prob of getting atmost 100 heads, atmost 1.7 heads or atmost 0.5 heads in 4 tosses of a balanced coin may not be of any real significance.

- $F(2.5) = P(X \le 2.5) = P(X \le 2) = F(2)$
- $F(10) = P(X \le 10) = P(X \le 4) = F(4) = 1$

PMF plot



CDF plot



$\boxed{\text{NOTE}}$ → as x varies from $-\infty$ to ∞ the graph of CDF i.e $F_X(x)$ resembles a staircase with upward steps having height $P(X=x_i)$ at each $x = x_i$. But note one thing that the graph of $F_X(x)$ remains constant b/w 2 steps or event.

Q→ Can we find PMF from CDF (in DRV) ?

↳ yes ✓

(ex) If X has the distribution funcn $F(1) = 0.25$, $F(2) = 0.61$, $F(3) = 0.83$, $F(4) = 1$ for $x = 1, 2, 3, 4$, find the prob distribution of X.

↳ we have;

$$P(1) = F(1) = 0.25,$$

$$P(X=2) = F_X(X=2) - F_X(X=1) = 0.61 - 0.25 = 0.36,$$

$$P(X=3) = F_X(X=3) - F_X(X=2) = 0.83 - 0.61 = 0.22$$

$$P(4) = F(4) - F(3) = 1 - 0.83 = 0.17$$

⇒

X	P(X=x)
1	0.25
2	0.36
3	0.22
4	0.17

GATE CSE 2012

ques. Consider a RV X that takes values +1 and −1 with prob. 0.5 each. The values of the CDF F(x) at x = −1 and +1 are

(A) 0 and 0.5 (B) 0 and 1 (C) 0.5 and 1 (D) 0.25 and 0.75

↳ ~~P(X=0).~~ $P(X=-1) = 0.5$ $P(X=1) = 0.5$

option (C) correct

$F_X(x=-1) = P(X=-1) = 0.5$

$F_X(x=1) = P(X=1) + P(X=-1) = 0.5 + 0.5 = 1$

VARIENCE

- It measures how far a set of no. is spread out from their average
- Mean is one way to summarize data and varience is one other way. Together they serve as good summary of data.
- Mean/Expectation : one no. that summarize PMF. It gives us an idea about the avg of the data in the distribution.
- Varience : gives us data about how spread away the data is from the mean. It is average squared distance from mean. It is always non-negative.

eg

Dataset-1 (PMF-1)

−5 0 5

$P(x=-5)=\frac{1}{3}$ $P(x=0)=\frac{1}{3}$ $P(x=5)=\frac{1}{3}$

- Avg of these points $= \frac{-5+0+5}{3} = 0$

↑ Mean/Expectation

Dataset-2 (PMF-2)

−1 0 1

$P(x=-1)=\frac{1}{3}$ $P(x=0)=\frac{1}{3}$ $P(x=1)=\frac{1}{3}$

- Avg of these points $= \frac{-1+0+1}{3} = 0$

- Varience is average squared distance from mean.

- Varience $= \frac{(Avg-(-5))^2 + (Avg-0)^2 + (Avg-5)^2}{3}$

$= \frac{(0+5)^2 + 0^2 + (0-5)^2}{3}$

$= \frac{50}{3} = 16.67$

- Varience $= \frac{(Avg-(-1))^2 + (Avg-0)^2 + (Avg-1)^2}{3}$

$= \frac{(0+1)^2 + 0^2 + (0-1)^2}{3}$

$= \frac{2}{3} = 0.67$

Varience will be low bcoz points are closer to mean.

Expectation/Mean of both the datasets is 0.
but it visible how different both datasets are,
so using just expectation we cannot tell much about a dataset
Hence varience is also important

GAMBLING EXAMPLE

Game I

$-10^6\$$ (lose) $+10^6\$$ (win)

$P(\text{lose } 10^6\$) = 1/2$ $P(\text{win } 10^6\$) = 1/2$

Expected money won/loss = 0

$$\text{Varience} = \frac{(10^6)^2 + (10^6)^2}{2} = 10^{12}$$

here varience tells that the risk of loss is very high

Game II

-10 (lose) $+10$ (win)

$P(\text{lose } 10\$) = 1/2$ $P(\text{win } 10\$) = 1/2$

Expected money won/loss = 0

$$\text{Varience} = \frac{(10)^2 + (10)^2}{2} = 100$$

the risk of loss is low.

Sampelling from a probability distribution.

$P(H) = 1/3$, $P(T) = 2/3$ } underlying Prob. distribution → nobody knows this → so we'll do the experiment & collect data.

and lets suppose we'll create the perfect sample:

HTTHTTHTT $\rightsquigarrow$ $P(H) = \frac{3}{9} = \frac{1}{3}$, $P(T) = \frac{2}{3}$ sample prob. = real prob.

o→ $P(x=i) = \frac{1}{6}$ on rolling a dice
i = 1, 2, 3, 4, 5, 6

$$E[x] = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + 3\cdot\frac{1}{6} + 4\cdot\frac{1}{6} + 5\cdot\frac{1}{6} + 6\cdot\frac{1}{6} = 3.5$$

now, Expectation using Sample data :→
Forget about prob. distribution

Sample data : 1, 2, 3, 4, 5, 6 → perfect sample

$$\text{Avg} = \frac{1+2+3+4+5+6}{6} = 3.5$$

$$\text{Varience} = \frac{(1-3.5)^2 + (2-3.5)^2 + (3-3.5)^2 + (4-3.5)^2 + (5-3.5)^2 + (6-3.5)^2}{6} = 2.916$$

$\boxed{\text{NOTE}}$ →

	Calculating using data	Calculate using prob distribution
Expectation	$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$	$\sum_k k\, P(x=k)$
Varience	squared distance from mean $\frac{(x_i - \bar{x})^2}{n}$	$E((x - E[x])^2)$ $\{E((x-\bar{x})^2)$ $\bar{x} = E[x]\}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

$\boxed{\text{NOTE}}$ → The most imp quality associated with a RV x, other than the mean, is its varience, which is denoted by $Var(x)$ or $v(x)$ and its defined as the expected value of the RV $(x - E[x])^2$. i.e,

$$v(x) = E[(x - E[x])^2].$$

Now,

$$E[(x-\bar{x})^2] = E[x^2 - 2x\bar{x} + \bar{x}^2]$$

{ bcoz $\bar{x}$ → expectation (mean (average)); $\bar{x} = E[x]$ }

$$= E[x^2 - 2x\,E[x] + (E(x))^2]$$

$$= E[x^2] - E[2x\,E(x)] + E[(E(x))^2]$$

$$= E[x^2] - 2E(x)\,E[x] + [E(x)]^2 E[1]$$

{ bcoz $E[ax] = aE[x]$ }
$E(x)$ is a RV or a constant

$$= E[x^2] - 2[E[x]]^2 + [E(x)]^2$$

$$\Rightarrow \boxed{V(x) = E[x^2] - [E[x]]^2}$$

Simpler formula for Varience :

$$v(x) = E[(x - E[x])^2]$$

← we'll not use this again

Both formulas are same

$$v(x) = E[x^2] - (E[x])^2$$

← we'll always use this

ques. Consider the RV X, which has the PMF

$$P_X(x) = \begin{cases} 1/9, & \text{if } x \text{ is an integer in the range } [-4,4], \\ 0, & \text{otherwise.} \end{cases}$$

Find out the mean and varience.

↳ mean = expectation, $E[x] = 0$

$$v(x) = E[(x - E[x])^2]$$

$$= E[(x-0)^2] = E[x^2]$$

$$\Rightarrow v(x) = E[x^2] = \frac{60}{9} \leftarrow \text{Varience}$$

$$E[\_] = \sum_k \_\cdot P(x=k)$$

$$E[x^2] = \sum_k x^2 \cdot P(x=k)$$

$$= 2\left((1)^2\tfrac{1}{9} + (2)^2\cdot\tfrac{1}{9} + (3)^2\cdot\tfrac{1}{9} + (4^2)\cdot\tfrac{1}{9}\right) + 0^2\cdot\tfrac{1}{9}$$

$$= \frac{60}{9}$$

$\boxed{\text{Method 2}}$ → using the perfect sample :

forget prob dis, focus on data : $-4, -3, -2, -1, 0, 1, 2, 3, 4$ ← perfect sample

mean (expectation) $= 0$, varience $= \dfrac{2(16+9+4+1)}{9} = \dfrac{60}{9}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

$\boxed{\text{Method 3}}$:→

The mean E[x] is equal to 0. This can be seen from the symmetry of the PMF of X around 0, & can also be verified from the definition.

$$E[X] = \sum_x x P_X(x) = \frac{1}{9} \sum_{x=-4}^{4} x = 0$$

Now,

let $Z = (X - E[X])^2 = X^2$.

$$P_Z(z) = \begin{cases} 2/9 & \text{if } z = 1, 4, 9, 16 \\ 1/9 & \text{if } z = 0 \\ 0 & \text{otherwise} \end{cases}$$

The varience of X is then obtained by

$$V(X) = E[Z] = \sum_z z P_Z(z) = 0 \cdot \frac{1}{9} + 1 \cdot \frac{2}{9} + 4 \cdot \frac{2}{9} + 9 \cdot \frac{2}{9} + 16 \cdot \frac{2}{9} = \frac{60}{9}$$

ques For each RV; x, y, z and w. compute the mean and varience.

(i)

value x	1	2	3	4	5
pmf $p(x)$	1/5	1/5	1/5	1/5	1/5

$$E[X] = \text{mean} = \frac{1+2+3+4+5}{5} = 3$$

$$E[X^2] = 1^2 \cdot \frac{1}{5} + 2^2 \cdot \frac{1}{5} + 3^2 \cdot \frac{1}{5} + 4^2 \cdot \frac{1}{5} + 5^2 \cdot \frac{1}{5} = \frac{55}{5} = 11$$

$$V(X) = E[X^2] - (E[X])^2$$
$$= 11 - (3)^2$$
$$= 2$$

(ii)

value y	1	2	3	4	5
pmf $p(y)$	1/10	2/10	4/10	2/10	1/10

$$E[Y] = 1 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{4}{10} + 4 \cdot \frac{2}{10} + 5 \cdot \frac{1}{10} = 3$$

$$E[Y^2] = 1^2 \cdot \frac{1}{10} + 2^2 \cdot \frac{2}{10} + 3^2 \cdot \frac{4}{10} + 4^2 \cdot \frac{2}{10} + 5^2 \cdot \frac{1}{10} = 10$$

$$V(Y) = E[Y^2] - [E(Y)]^2$$
$$= 10 - (3)^2 = 1$$

(iii)

value z	1	2	3	4	5
pmf $p(z)$	5/10	0	0	0	5/10

$$E[z] = 1\cdot\frac{5}{10} + \underbrace{2\cdot 0 + 3\cdot 0 + 4\cdot 0}_{=0} + 5\cdot\frac{5}{10} = \frac{30}{10} = 3.0$$

$$E[z^2] = 1^2\cdot\frac{5}{10} + 5^2\cdot\frac{5}{10} = \cancel{126}\ \frac{130}{10} = 13$$

$$V(x) = E[z^2] - (E[z])^2$$
$$= 13 - (3)^2 = 4$$

(iv)

value w	1	2	3	4	5
pmf $p(w)$	0	0	1	0	0

$$E[w] = 0\cdot 1 + 0\cdot 2 + 1\cdot 3 + 0\cdot 4 + 0\cdot 5 = 3$$

$$E[w^2] = 1\cdot 3^2 = 9$$

$$V(w) = E[w^2] - (E[w])^2$$
$$= 9 - (3)^2 = 0$$

ques. A random integer N b/w 1 and 5 (inclusive) is chosen. Suppose that N has PMF P_N given by $P_N(n) = k\cdot n^2$ for $n = 1, 2, 3, 4, 5$

for some constant k.

(a) compute $E[N]$. Your final answer should not include the letter k.

$$\sum_{n=1}^{5} P_N(n) = 1 \Rightarrow k + 4k + 9k + 16k + 25k = 1$$
$$k = 1/55$$

$$\therefore \quad P_N(n) = \frac{n^2}{55}$$

now, $$E[N] = 1\cdot\frac{1^2}{55} + 2\cdot\frac{2^2}{55} + 3\cdot\frac{3^2}{55} + 4\cdot\frac{4^2}{55} + 5\cdot\frac{5^2}{55} = \frac{225}{55} = \frac{45}{11}$$

(b) compute $V(N)$.

now, $$E[N^2] = 1^2\cdot\frac{1^2}{55} + 2^2\cdot\frac{2^2}{55} + 3^2\cdot\frac{3^2}{55} + 4^2\cdot\frac{4^2}{55} + 5^2\cdot\frac{5^2}{55} = \frac{89}{5}$$

$$V(N) = E[N^2] - (E(N))^2$$
$$= \frac{89}{5} - \left(\frac{45}{11}\right)^2 = \frac{644}{605}$$

qus. Consider a sequence of independent tosses of a coin that is biased so that it comes up heads with prob. 3/4 and tails with prob 1/4.
let x_i be 1 if the ~~i^{th}~~ ith toss comes up heads and 0 otherwise.

(a) compute $E[X_1]$ and $Var[X_1]$

$P(H) = 3/4 \quad P(T) = 1/4$

$P(X_i = 1) = 3/4$

$X_1 = 1$ if 1^{st} toss is head
$X_2 = 1$ if 2^{nd} toss is head
$X_3 = 1$ if 3^{rd} toss is head

$P(X_1 = 1) = P(1^{st}$ toss is head$) = 3/4$

PMF of X_1 :

$P(X_1 = 1) = 3/4$

$P(X_1 = 0) = 1/4$

$$E[X_1] = 1 \cdot \frac{3}{4} + 0 \cdot \frac{1}{4} = \frac{3}{4}$$

now, $E[X_1^2] = 1^2 \cdot \frac{3}{4} + 0^2 \cdot \frac{1}{4} = \frac{3}{4}$

$$Var[X_1] = E[X_1^2] - (E[X_1])^2 = \frac{3}{4} - \left(\frac{3}{4}\right)^2 = \frac{3}{16}$$

(b) compute $Var[X_1 + 2X_2 + 3X_3 + 4X_4]$

it must be true only when X_1, X_2, X_3, X_4 are independent

$$= V[X_1] + V[2X_2] + V[3X_3] + V[4X_4]$$

$$= V[X_1] + 2^2 V[X_2] + 3^2 V[X_3] + 4^2 V[X_4]$$

$$= \frac{3}{16} + 4 \cdot \frac{3}{16} + 9 \cdot \frac{3}{16} + 16 \cdot \frac{3}{16}$$

$$= \frac{3}{16}(1 + 4 + 9 + 16) = \frac{90}{16} = \frac{45}{8}$$

qus. Suppose that $P(X=0) = 1 - P(X=1)$. If $E[X] = 5\,Var(X)$, Find $P(X=1)$.
Assume non-zero probabilities :-

$P(X=0) = 1-p$

$P(X=1) = p$

$E[X] = 1 \cdot p + 0 \cdot (1-p) = p$

$E[X^2] = 1^2 \cdot p + 0^2 \cdot (1-p) = p$

$V(X) = E[X^2] - (E[X])^2$
$= p - p^2 = p(1-p)$

Now, $E[X] = 5\,V(X)$

$\Rightarrow p = 5p(1-p)$

$p - 5p(1-p) = 0$

$p(1 - 5(1-p)) = 0$

$p = 0$ or $5(1-p) = 1$

$p = 4/5$

$\Rightarrow \boxed{P(X=1) = \frac{4}{5}}$

Ques. let D_1 and D_2 be the outcomes (in $\{1,2,3,4,5,6\}$) of 2 independent fair die rolls.
let y_i be the RV which is equal to 1 if $D_1 = i$ and 0 otherwise
compute the $Var\left[\sum_{i=1}^{6} y_i\right]$.

↳ $D_1 \Rightarrow$ outcome of 1^{st} die roll

$D_2 \Rightarrow$ outcome of 2^{nd} die roll

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

now, y_i be the RV which is equal to 1 if $D_1 = i$ or 0 otherwise.

$y_1 = 1$ if $D_1 = 1$

$y_2 = 1$ if $D_1 = 2$

y_1	y_2	y_3	y_4	y_5	y_6
↓	↓	↓	↓	↓	↓
1	1	1	1	1	1
1^{st} roll get 1	if 2^{nd} roll get 2	if 3^{rd} roll is 3		if 5^{th} roll is 5	

y_1	y_2	y_3	y_4	y_5	y_6
1 / 0	1 / 0	1 / 0	1 / 0	1 / 0	1 / 0
$D_1 = 1$	$D_1 = 2$	$D_1 = 3$	$D_1 = 4$	$D_1 = 5$	$D_1 = 6$

$\Rightarrow$ value of y_i is totally based on first roll.
i.e if ~~$D_1=2$~~ $D_1 = 5$, $y_5 = 1$ and everything else is 0, no 2 y's will be 1 together.

$\Rightarrow$ At any time exactly one of the $y = 1$

y_1	y_2	y_3	y_4	y_5	y_6		
1	0	0	0	0	0	$\Leftarrow D_1 = 1$	1^{st} roll is 1
0	1	0	0	0	0	$\Leftarrow D_1 = 2$	$\to 1^{st}$ roll is 2
0	0	1	0	0	0	$\Leftarrow$ ~~D_1~~ $D_1 = 3$	$\to 1^{st}$ roll is 3

NOW, $V\left[\sum_{i=1}^{6} y_i\right]$

$= V[y_1 + y_2 + y_3 + y_4 + y_5 + y_6]$ → This sum is constant.

$= V[1] = 0$ ← Answer

$$\left\{ \begin{aligned} V[1] &= E[1^2] - (E[1])^2 \\ &= E[1] - 1^2 \\ &= 1 - 1 = 0 \end{aligned} \right\}$$

o→ $\boxed{V[\text{any constant}] = 0}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques GATE CSE 2011

If the difference b/w the expectation of the square of a RV $(E[X^2])$ and the square of the expectation of the RV $(E[X])^2$ is denoted by R, then

(A) $R = 0$ (B) $R > 0$ (C) $R \geq 0$ (D) $R < 0$

↳ $\underbrace{R}_{\text{Varience}} = E[X^2] - (E[X])^2 = E\Big[\underbrace{\overbrace{(X - E(X))^2}^{y}}_{\geq 0}\Big]$

$\therefore \; y \geq 0 \Rightarrow E[y] \geq E[0]$

$\Rightarrow E[y] \geq 0$

$\Rightarrow \boxed{R \geq 0}$

Properties of Expectation and Varience

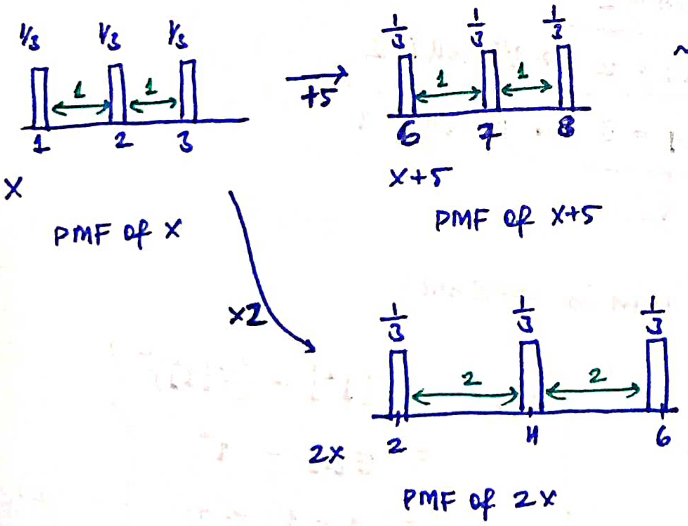
- $y = aX + b$

$E[y] = a\,E[X] + b$

$V[y] = V[aX] = a^2 V[X]$

- $E[X_1 + X_2] = E[X_1] + E[X_2]$ (Always True)

$V[X_1 + X_2] = V[X_1] + V[X_2]$ (if X_1 and X_2 are Independent)



⇝ in this expectation will change
$E[X+5] = E[X] + 5$

but varience will be same, as spread is still the same
$V[X+5] = V[X]$

⇝ in this expectation will change
$E[2X] = 2E[X]$

Varience will also change as spread has changed
$V[2X] = 2^2 V[X]$

qun. True/False

(i) $V[X-Y] = V[X] - V[Y]$; X, Y are indep.

FALSE

$V[X+Y] = V[X] + V[Y]$ (formula)

$V[X-Y] = V[X] + V[-Y]$

$= V[X] + (-1)^2 V[Y]$

$V[X-Y] = V[X] + V[Y]$

(ii) $V[aX+bY+c] = a^2V[X] + b^2V[Y] + c$; X, Y are indep

FALSE

$V[aX+bY+c] = a^2V[X] + b^2V[Y]$ as $\underbrace{V[c]=0}_{\text{constant}}$

(iii) $V[aX+bY+c] = a^2V[X] + b^2V[Y] + V[c]$; X, Y are indep

TRUE

(iv) $V[aX+bY+c] = a^2V[X] + b^2V[Y]$ for any RVs X, Y

FALSE True only if X, Y are independent

(v) $E[aX+bY+c] = aE[X] + bE[Y] + c$ for any RVs X, Y

TRUE

STANDARD DEVIATION

$$S.D = \sqrt{\text{Variance}}$$

RV X has a unit (marks, rupees etc)

but the unit of variance is unit².

hence we need SD.

$$\text{Variance} = \frac{(\ )^2 + (\ )^2 + (\ )^2}{3} \quad \text{unit}^2$$

ex) P (winning \$10) = 1/2 ← we did this example before

P (losing \$10) = 1/2

E[money win or lose] = 0

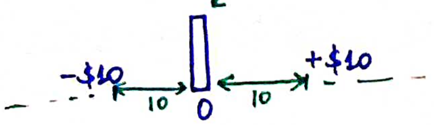
$\{\frac{1}{2}\cdot 10 + \frac{1}{2}(-10) = 0\}$

Variance( ) = 100

$\boxed{SD = 10}$

it means from the expected point we'll be deviating 10 here and there.

Avg winning/losing amount



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Discrete Random Variables

Famous examples of DRVs

- Bernoulli ⇝ $X=1 \quad p$, $X=0 \quad 1-p$
- Binomial
- Poisson
- Uniform

with probability

$X=1$ for some outcomes w.p. p

$X=0$, for some outcomes w.p. $1-p$

DRV

Any RV

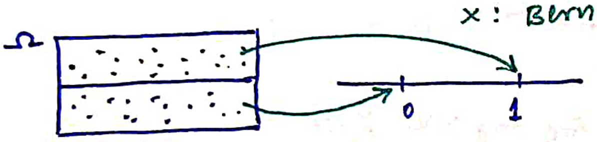
$$Y = \begin{cases} 0 & \text{w.p } p_1 \\ 1 & \text{w.p } p_2 \\ 2 & \text{w.p } p_3 \end{cases} \qquad p_1+p_2+p_3 = 1$$

Bernoulli RV

- For its simplicity, the Bernoulli RV is very important.
- It can take 2 values, 1 and 0. It takes on a 1 if an experiment with prob. p resulted in success and a 0 otherwise.

i.e $\quad X = \begin{cases} 1 & \text{w.p } p \\ 0 & \text{w.p } 1-p \end{cases}$

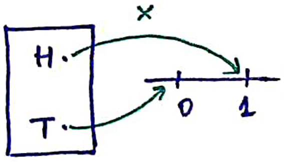
X : Bernoulli RV



ex) Toss a coin:

$X(H) = 1$

$X(T) = 0$



$\boxed{P(X=1) = P(H)}$

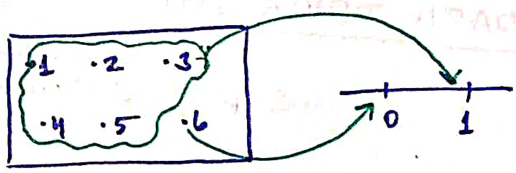
getting head is a success

$p(\text{success}) = p = p(H)$

ex) Rolling a die:

X: Bernoulli RV

$1, 2, 3, 4, 5, 6$ → 1 ; 6 → 0



NOTE → if we have some experiment where we can classify every outcome as success or failure then we say we can have Bernoulli RV.

Ques. Consider a RV X, which takes values 1 or 0.

$$X = \begin{cases} 1 & \text{if a head} \\ 0 & \text{if a tail} \end{cases} \qquad P(X=x) = \begin{cases} p & \text{if } x=1 \\ 1-p & \text{if } x=0 \end{cases}$$

Find the expectation and variance of X?

↳ $E[X] = 1\cdot p + 0(1-p) = p$

$E[X^2] = 1^2\cdot p + 0^2(1-p) = p$

$$V(X) = E[X^2] - (E[X])^2 = p - p^2 = p(1-p).$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

NOTE → <u>Remember</u> **BERNOULLI RV**

1) PMF : $X = \begin{cases} 1 & \text{w.p } p \\ 0 & \text{w.p } 1-p \end{cases}$

2) Expectation : $E[X] = p$

3) Varience : $V[X] = p(1-p)$

BINOMIAL RV

Repeated Independent trials of Bernoulli

Trials : n

ex) <u>tossing n times</u>

no. of heads = k

no. of tails = $n-k$

<u>performing experiment n times</u>

no. of success = k

no. of failures = $n-k$

ques. we do toss 5 times.

X : no. of heads occurred.

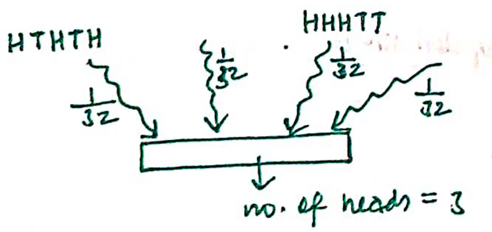
$P(X=3) = ?$

↳ $P(HTHTH) = \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{32}$

$\cancel{P(X=3) = \frac{1}{32}}$ Wrong as HTHTH is not the only sequence in which H is coming 3 times

HHHHH
TTTTT
HTHTH
⋮

we need to consider all the combinations in which 'H' is coming 3 times. ——————



$\therefore P(X=3) = {}^5C_3 \cdot \frac{1}{32}$

$= 10 \cdot \frac{1}{32} = \frac{5}{16}$

ques. A fair coin is tossed n times.

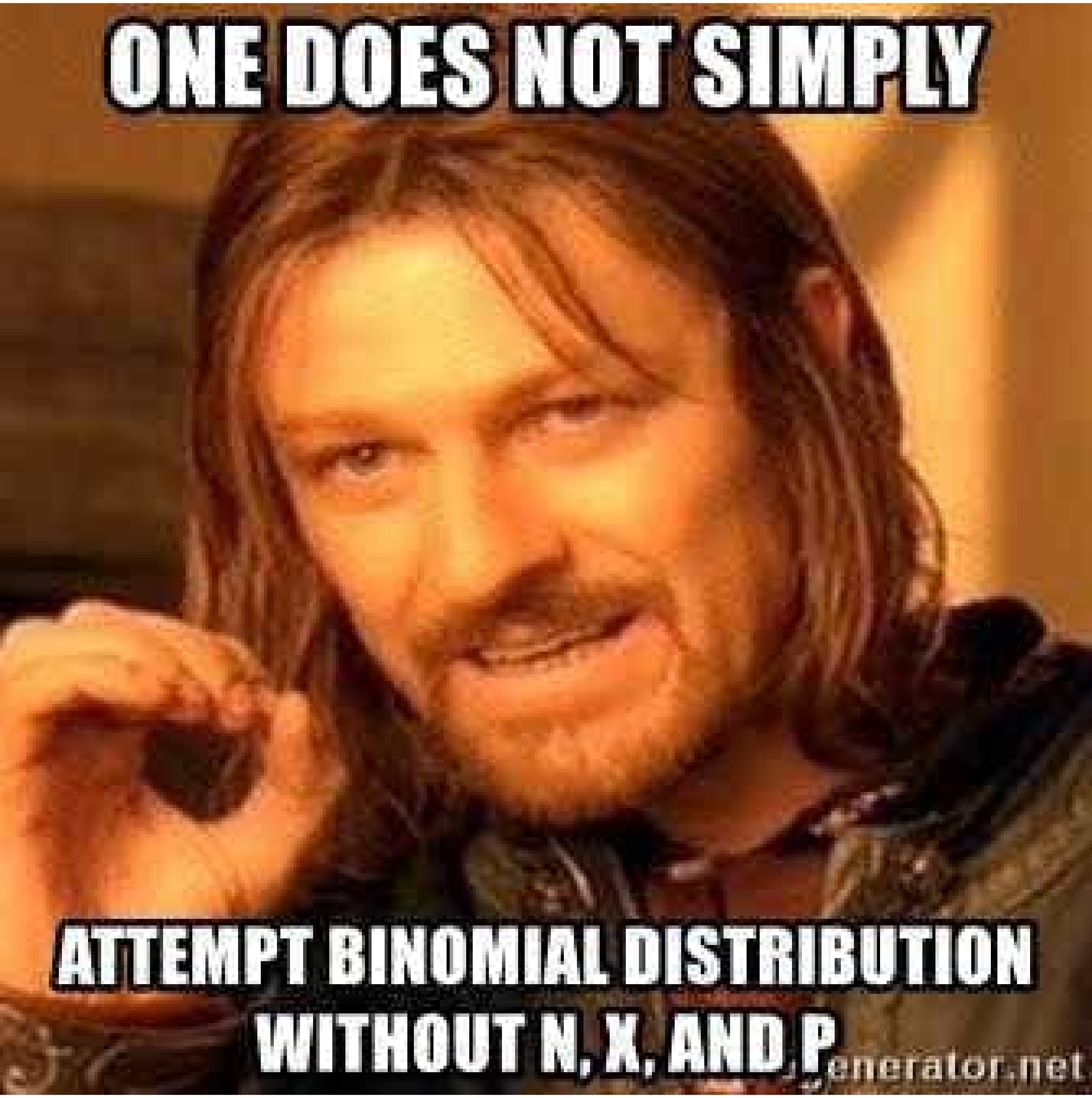
$P(H) = p$

X = no. of H

what will be the $p(X=k)$?

↳ $P(X=k) = {}^nC_k\, p^k (1-p)^{n-k}$, $k = 0, 1, 2, \dots, n$

$\boxed{\text{Binomial RV}}$



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Bernoulli	Binomial
$X = 0, 1$	$X = 0, 1, \ldots, n$
PMF : $P(X=1) = p$	PMF : $P(X=k) = {}^nC_k\, p^k (1-p)^{n-k}$

•→ Repeated independent trials of Bernoulli = Binomial

qun. A fair coin is tossed n times.

$P(H) = p$

$X = \#$heads

What will be the $P(X=k)$?

Find out the Expectation & Varience of Binomial RV.

↳ $P(X=k) = {}^nC_k \cdot p^k (1-p)^{n-k}$

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

X	$P(X=k)$
0	$(1-p)^n$
1	$n \cdot p \cdot (1-p)^{n-1}$
2	${}^nC_2 \cdot p^2 \cdot (1-p)^{n-2}$
⋮	⋮

$\boxed{\text{Method 1}}$ →

$$E[X] = \sum_{k=0}^{n} k \cdot P(X=k)$$

after solving this

$$E[X] = \sum_{k=0}^{n} k \cdot {}^nC_k\, p^k (1-p)^{n-k} = np$$

(skipping derivation)

$\boxed{\text{Method 2}}$ → n times tossing a coin

X : No of heads → $0, 1, 2, \ldots, n$

$E[X]$ → expected no. of heads

we're doing n trials

Let's define a seperate RV for each trial.

no. of Heads in 1st trial ↓ X_1 → 1, 0 ; no. of Heads in 2nd trial ↓ X_2 → 1, 0 ; X_3 → 1, 0 ; ....., X_n → 1, 0

no. of heads ↓

$X = X_1 + X_2 + X_3 + \ldots + X_n$

↑ Binomial

$X_1 = 1$
$X_2 = 0$
$X_3 = 1$
$X_4 = 1$
} no. of heads = 3.

We have broken down our Binomial RV into n indep. Bernoulli RVs.

$$\boxed{X = X_1 + X_2 + \ldots + X_n}$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

now, $X = X_1 + X_2 + \dots + X_n$

$X_1 = \begin{cases} 0, & \text{no head in } 1^{st} \text{ trial} \\ 1, & \text{head in } 1^{st} \text{ trial} \end{cases}$

Expectation

$E[X] = E[X_1 + X_2 + \dots + X_n]$

$= \underbrace{E[X_1]}_{p} + \underbrace{E[X_2]}_{p} + \dots + \underbrace{E[X_n]}_{p}$

$\boxed{E[X] = np}$

Varience

$V[X] = V[X_1 + X_2 + \dots + X_n]$ ↘ X_i's are independent

$= \underbrace{V[X_1]}_{p(1-p)} + \underbrace{V[X_2]}_{p(1-p)} + \dots + \underbrace{V[X_n]}_{1-p}$

$= np(1-p) = npq \qquad (q = 1-p)$

$\boxed{V[X] = npq}$

we donot need the independent assumption for the expected value, since it is a linear func^n of RVs, but we need it for varience

NOTE → for Binomial RV

X : no. of success in n trials

Range of X : $0, 1, 2, \dots, n$

PMF of X : $P(X=k) = {}^nC_k \, p^k (1-p)^{n-k}$

$E[X] = n\,p$ (n → no. of trials, p → prob. of success)

$V[X] = npq \qquad (q = 1-p)$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques. 30 people are invited to a party. Each person accepts the invitation, independently of all others, with probability $1/3$.

let X be the no. of accepted invitation.

compute the following:

(a) $E[X]$ (b) $V[X]$ (c) $E[X^2]$ (d) $E[X^2-4X+5]$

↳ $\overset{\text{binomial}}{X} = X_1 + X_2 + \dots + X_n$ for each person $= X_i$

X_i: 1 (Accept), 0 (Reject)

$E[X] = \underset{n}{30} \times \underset{p}{\frac{1}{3}} = 10$

$V[X] = npq$

$= np(1-p) = 30 \times \frac{1}{3} \times \frac{2}{3} = \frac{20}{3}$

$E[X^2]$; $V[X] = E[X^2] - (E[X])^2$

$\frac{20}{3} = E[X^2] - 100$

$E[X^2] = \frac{20}{3} + 100 = \frac{320}{3}$

$E[X^2 - 4X + 5] = E[X^2] - 4E[X] + 5$

$= \frac{320}{3} - 4 \times 10 + 5 = \frac{215}{3}$

ques. Suppose that a fair die is rolled 72000 times. Each roll turns up a uniformly Random member of the set $\{1,2,3,4,5,6\}$ and the rolls are independent of each other. For each $j \in \{1,2,3,4,5,6\}$ let X_j be the no. of times that the die comes up j.

X_1: no. of 1's, X_2: no. of 2's, X_3: no. of 3's, X_4: no. of 4's, X_5: no. of 5's, X_6: no. of 6's

(a) compute $E[X_3]$ and $V[X_3]$.

you are rolling a dice 72000 times and X_3: no. of $\underset{\text{success}}{\text{3's}}$

in general template
X: No of success

$E[X_3] = n\underset{\text{prob. of success}}{p} = 72000 \times \frac{1}{6}$

$= \boxed{12000}$

$V[X_3] = npq = 72000 \times \frac{1}{6} \times \frac{5}{6}$ $\left(q = 1-p = \frac{5}{6}\right)$

$= \boxed{10000}$

(b) compute $V[X_1 + X_2]$.

X_1 : no of 1's.

X_2 : no of 2's.

$y = X_1 + X_2$ = no. of 1's and 2's.

success: either 1 comes or 2 comes in any trial

$P(\text{success}) = 2/6 = 1/3$

$E[y] = n \cdot p = 72000 \times \frac{1}{3} = 24000$

$V[y] = n \cdot p \cdot q = 72000 \times \frac{1}{3} \times \frac{2}{3} = \boxed{16000}$

(c) $E[X_1 + X_2 + X_3] = ?$

success: either 1, 2 or 3 can come

$P(\text{success}) = 1/2 \quad (= 3/6)$

$E[X_1 + X_2 + X_3] = np = 72000 \times \frac{1}{2} = 36000$

OR

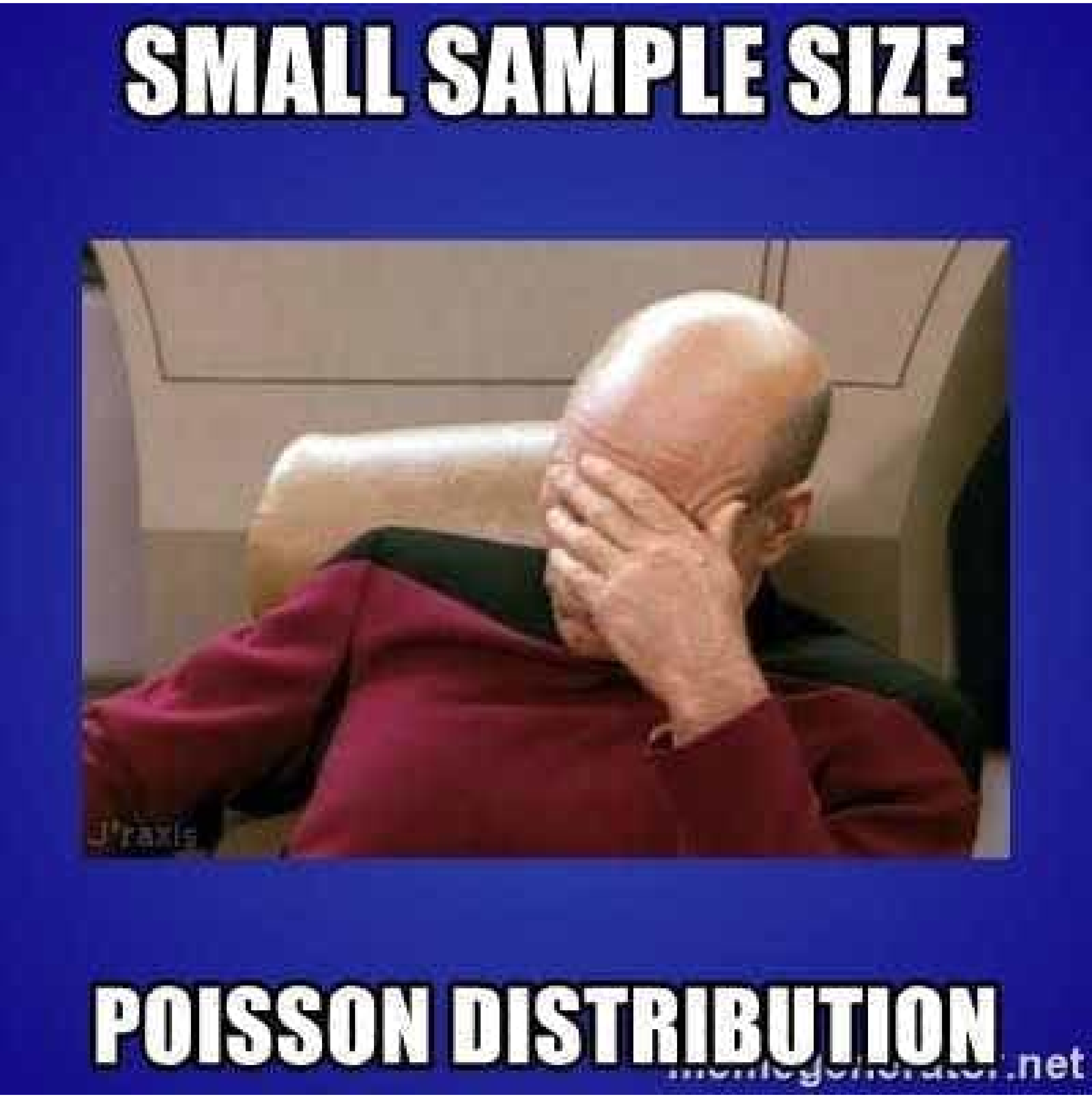
{we dont need any assumption here about dep. or indep.}

$E[X_1 + X_2 + X_3] = E[X_1 + X_2] + E[X_2]$

$= 24000 + 12000 = 36000$

we already did   we already did

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Poisson Distribution

A R.V X has the Poisson distribution with parameter λ, where $\lambda > 0$, if the PMF of X is

$$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots$$

e = Euler's constant ≈ 2.71828

X can be seen as representing the no of events occuring in a fixed interval where events occur randomly throughout the interval.

Bernoulli : $X = 0, 1$

Binomial : $X = 0, 1, 2, \ldots, n$

Poisson : $X = 0, 1, 2, \ldots \infty$

Binomial and Poisson are same, its just that we put $n = \infty$ in terms of Poisson.

Binomial with $n = \infty$ is Poisson.

$$\underbrace{{}^nC_k\, p^k (1-p)^{n-k}}_{n=\infty} \quad \} \; n\uparrow, p\downarrow \quad = \frac{e^{-\lambda}\lambda^k}{k!}$$

and $p\downarrow$, p is very small

Mean and Varience of Poisson RV

$$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!} \qquad k = 0, 1, 2, \ldots$$

$$E[X] = \sum_{k=0}^{\infty} k \cdot P(X=k) = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda}\lambda^k}{k!} = \lambda$$

$$\boxed{E[X] = V[X] = \lambda}$$

$$V[X] = E[X^2] - (E[X])^2 = \lambda$$

ques. On an avg there are 2 accidents per day, then what is prob for 4 accidents per day?

$E[X] = \lambda = 2 \leftarrow$ 2 accidents per day

$$P(X=4) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-2}2^4}{4!} \approx 0.09$$

NOTE → More precisely, the Poisson PMF with parameter λ is a good approximation for a binomial PMF with parameters n and p, provided $\lambda = np$, n is very large, and p is very small. i.e.,

$$\frac{e^{-\lambda}\lambda^k}{k!} \approx {}^nC_k\, p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.$$

In this case, using the Poisson PMF may result in simpler models & calculations.

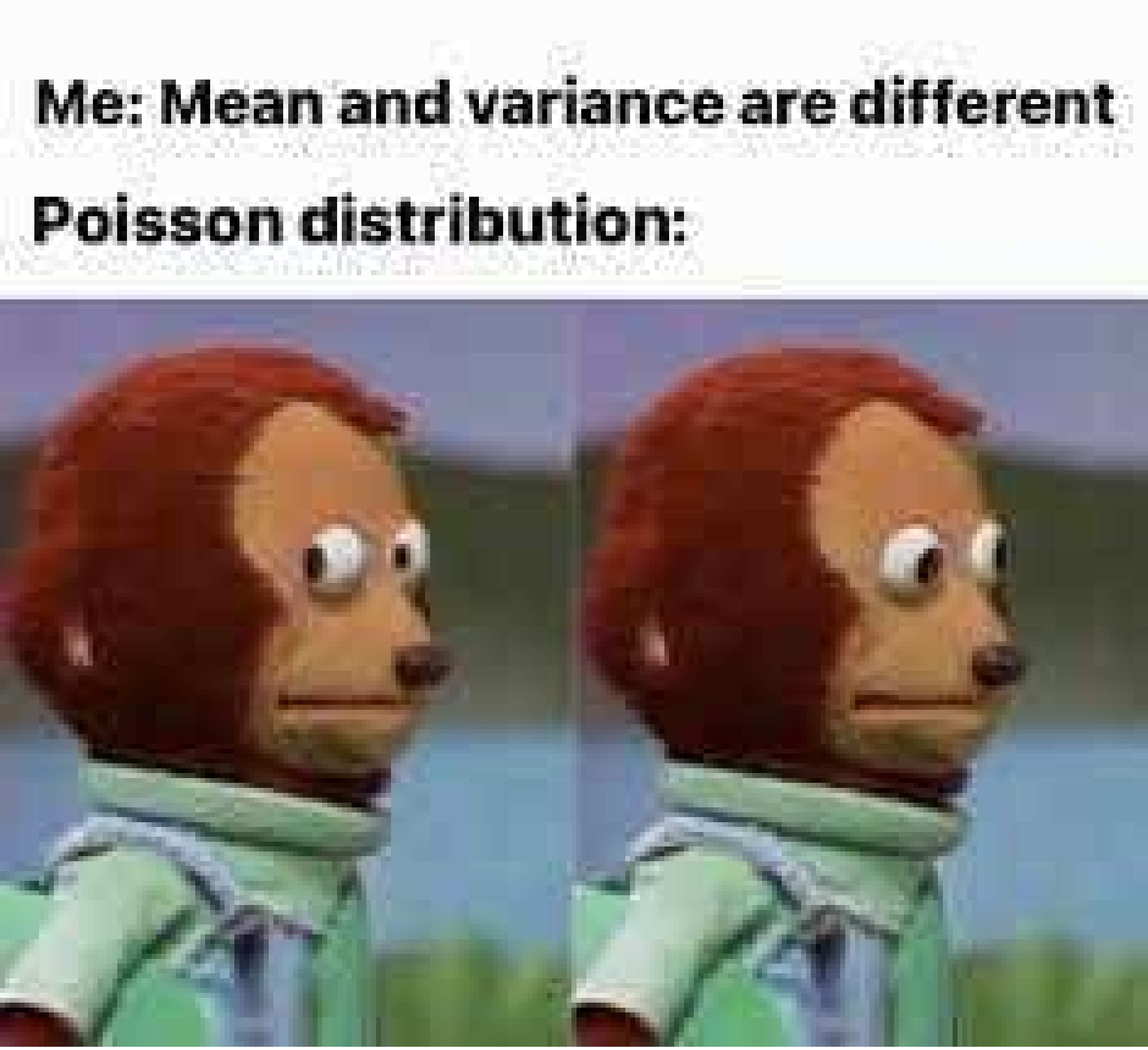
For ex, let $n = 100$ and $p = 0.01$. The prob of $k = 5$ successes in $n = 100$ trials is calculated using the binomial PMF as $\;{}^{100}C_5 (0.01)^5 (1-0.01)^{95} = \underline{0.00290}$.

using the Poisson PMF with $\lambda = np = 100 \cdot 0.01 = 1$, this prob is approximated by

$$e^{-1} \cdot \frac{1}{5!} = \underline{0.00306}.$$

POISSON DISTRIBUTION BE LIKE

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

How do I identify a Poisson RV ?

Some examples of RV that generally obey the Poisson prob. law :

1) The no. of misprints on a page (or a group of pages) of a book.
2) The no. of people in a community who survive to age 100.
3) The no of wrong telephone nos that are dialed in a day
4) The no of packages of dog biscuits sold in a particular store each day
5) The no. of customers entering a post office on a given day
6) The no of vacancies occuring during a year in a federal judicial system.
7) The no of α-particles discharged in a fixed period of time from some radioactive material.

ques. Suppose that the no. of typographical errors on a single page of this book has a poisson distribution with parameter $\lambda = \frac{1}{2}$. Calculate the prob. that there is atleast one error on this page

↳ on a single page avg errors $(\lambda) = \frac{1}{2}$

$X \rightarrow$ no. of errors on this page

now, $P(X \geq 1) = 1 - P(X < 1)$

$= 1 - P(X=0)$

$= 1 - e^{-1/2} \approx \boxed{0.393}$

$P(X=0) = \frac{e^{-\lambda}\lambda^k}{k!} = \frac{e^{-1/2}(1/2)^0}{0!}$

$= e^{-1/2}$

ques. On an avg there are 2 accidents per day, what is the prob for 4 accidents in 2 days?

↳ $\lambda = 2 \rightarrow$ per day

now $\lambda = 4 \rightarrow$ per 2 days

we have to modify the λ as per question

$P(X=4) = \frac{e^{-\lambda}\lambda^k}{k!}$

$= \frac{e^{-4}(4)^4}{4!} \approx 0.195$

ques. Patients arrive at a hospital accident and emergency department at random at a rate of 6 per hour.

(a) Find the prob that, during any 90 min period, the no. of patients arriving at the hospital accident & emergency department is

(i) exactly 7 $P(X=7)$

(ii) atleast 3 $P(X \geq 3)$

↳ given $\lambda = 6$

60 min — 6 patients

90 — $\frac{6}{60} \times 90 = 9$ patients

$\boxed{\lambda = 9}$ ↑ updated λ

$$P(X=7) = \frac{e^{-\lambda}\lambda^k}{k!} = \boxed{\frac{e^{-9}\,9^7}{7!}}$$

$\frac{e^{-\lambda}\lambda^k}{k!}$

$$P(X \geq 3) = 1 - P(X \leq 2)$$

$$= 1 - [P(X=0) + P(X=1) + P(X=2)]$$

$$\downarrow \quad e^{-\lambda} \qquad \downarrow \quad e^{-\lambda}\lambda \qquad \downarrow \quad \frac{e^{-\lambda}\lambda^2}{2!}$$

(b) A patient arrives at 11.30 am. Find the prob that the next patient arrives before 11.45 am.

↳ exactly one patient in 15 min.

60 — 6

15 — $\frac{6}{60} \times 15 = 3/2$

$\boxed{\lambda = 3/2}$ ↑ updated λ

$$P(X=1) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$$= e^{-\lambda}\lambda$$

$$\boxed{= \frac{3}{2}e^{-3/2}}$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques GATE CSE 2013

Suppose p is the no. of cars per minute passing through a certain road junction b/w 5 PM and 6 PM, and p has poisson distribution with mean 3. what is the prob of observing fewer than 3 cars during any given minute in this interval?

(A) $\frac{8}{(2e^3)}$ (B) $\frac{9}{(2e^3)}$ (C) $\frac{17}{(2e^3)}$ (D) $\frac{26}{(2e^3)}$

↳ p: no. of cars per minute

$\lambda = 3$ cars per minute

$$P(p<3) = P(p=0) + P(p=1) + P(p=2)$$

$$= e^{-\lambda} + e^{-\lambda}\lambda + \frac{e^{-\lambda}.\lambda^2}{2!}$$

$$= e^{-3}\left(1 + 3 + \frac{3^2}{2!}\right) = \frac{17}{2e^3}$$

ques. GATE CSE 2017

If a RV X has a poisson distribution with mean 5, then the expectation $E[(x+2)^2] = ?$

↳ $E[(x+2)^2] = E[x^2 + 4x + 4]$

$= E[x^2] + 4E[x] + 4$

$= 30 + 4(5) + 4$

$= 54$

in poisson, $E[x] = V[x] = \lambda = 5$

$V[x] = E[x^2] - (E[x])^2$

$5 = E[x^2] - 25$

$E[x^2] = 30$

ques. GATE CSE 2007

In a multiuser operating system on an avg, 20 requests are made to use a particular resource per hour. The arrival of requests follows a poisson distribution; The prob. that either 1, 3 or 5 requests are made in 45 min is given by:

(A) $6.9 \times 10^6 \times e^{-20}$ (B) $1.02 \times 10^6 \times e^{-20}$ (C) $6.9 \times 10^3 \times e^{-20}$ (D) $1.02 \times 10^3 \times e^{-20}$

↳ 60 min → 20 requests

45 min → $\frac{20}{60} \times 45 = 15$ requests

$\lambda = 15$ modified λ

$P(x=k) = \frac{e^{-\lambda}\lambda^k}{k!}$

now, $P(x=1) + P(x=3) + P(x=5)$

$$= e^{-\lambda}.\lambda + \frac{e^{-\lambda}.\lambda^3}{3!} + \frac{e^{-\lambda}.\lambda^5}{5!}$$

$$= e^{-15}\left(15 + \frac{(15)^3}{3!} + \frac{(15)^5}{5!}\right)$$

$$= 6905.625 \times e^{-15}$$

$$= 6.9 \times 10^3 \times e^{-20} \times e^5 = 6.9 \times 151.63 \times 10^3 \times e^{-20}$$

$$= \boxed{1.02 \times 10^6 \times e^{-20}}$$

DISCRETE UNIFORM RV

A RV X has a discrete uniform distribution if each of the 'n' values in its range, say $x_1, x_2, \ldots, x_n$, has equal probability. Then,

$$p(x_i) = \frac{1}{n}$$

where $p(x)$ represents the PMF.

ex $\{A, B, C, D\}$

$p(A) = p(B) = p(C) = p(D) = \frac{1}{4}$

ex $\{1, 2, 3, 4\}$

$p(1) = p(2) = p(3) = p(4) = \frac{1}{4}$

ex $\{2, 50, 103, 110\}$

$p(2) = p(50) = p(103) = p(110) = \frac{1}{4}$

ex $\{1, 2, 3, 4, \ldots, n\}$

$p(1) = p(2) = \ldots = p(n) = \frac{1}{n}$

ex $\{2, 4, 6, 8, \ldots, 2n\}$

$p(2) = p(4) = \ldots = p(2n) = \frac{1}{n}$

ex Roll a dice and let X be the upward face showing. Find mean & varience

$X = 1, 2, 3, 4, 5, 6$

PMF: $p(X = i) = \frac{1}{6}$

$$E[X] = \sum_x x \cdot p(x) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3.5$$

$$\cancel{E[X]} \quad E[X^2] = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} = \frac{81}{6}$$

$$V[X] = E[X^2] - (E[X])^2$$

$$= \frac{81}{6} - 3.5$$

ques. If X is uniformly distributed on the set $\{1, 2, 3, \ldots, N\}$, then

(i) $P(x) = ?$

$$P(x) = \frac{1}{N} \qquad x \in \{1, 2, 3, \ldots, N\}$$

(ii) $$E[X] = \frac{1+2+3+\cdots+N}{N} = \frac{\frac{N(N+1)}{2}}{N} = \frac{N+1}{2}$$

(iii) $$V[X] = E[X^2] - (E[X])^2$$

$$E[X^2] = \frac{1^2+2^2+\cdots+N^2}{N} = \frac{\frac{N(N+1)(2N+1)}{6}}{N} = \frac{(N+1)(2N+1)}{6}$$

$$V[X] = \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4}$$

$$= \frac{(N+1)}{2} \cdot \frac{(N-1)}{6} = \frac{N^2-1}{12}$$

ques. If X is a uniformly distributed RV on the set $\{a, a+1, \ldots, b-1, b\}$ then Find out

(i) $$P(X=x) = \frac{1}{(b-a+1)} \rightarrow \text{total no. of elements in set}$$

(ii) $$E[X] = \frac{\frac{(b-a+1)}{2}(a+b)}{(b-a+1)} = \frac{(a+b)}{2}$$

(iii) $V[X] = ?$

$$E[X^2] = \frac{a^2+(a+1)^2+\cdots+(b-1)^2+b^2}{(b-a+1)} = \frac{\frac{b(b+1)(2b+1)}{6} - \frac{(a-1)a(2a-1)}{6}}{(b-a+1)}$$

$$= \frac{2b^3+3b^2+b-2a^3+3a^2-a}{6(b-a+1)}$$

now, $$V[X] = E[X^2] - (E[X])^2 = \frac{(b-a+1)^2-1}{12}$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

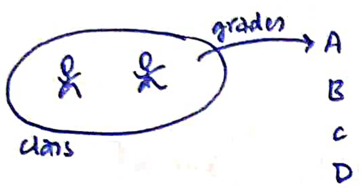
Discrete RVs	Continuous RVs
1. Bernoulli	1. Uniform
2. Binomial	2. Exponential
3. Poisson	3. Normal
4. Uniform	

CONTINUOUS RV

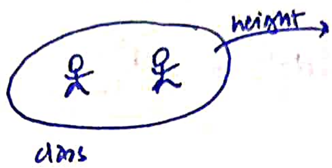
Discrete RV v/s **Continuous RV**

X takes countably finite values — X will take uncountable values.

(ex)



(ex)



120 cm
⋮
200 cm } continuous values

- In this we have PMF — we have PDF here

Probability Density Function (PDF) (denoted by $f(x)$)

A RV X has a PDF $f(x)$ if; $\boxed{P(a \le X \le b) = \int_a^b f(x)\,dx}$ $\forall a, b$.

→ if $a=b \Rightarrow \int_a^a f(x)\,dx = 0$

→ For valid pdf -

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

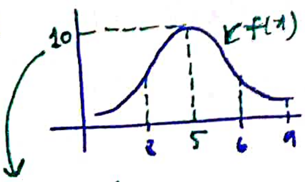
→ $P(X \le b) = \int_{-\infty}^{b} f(x)\,dx$

What is $f(x)$, intuitively? → prob. density function

↳ it tells likelihood.

$f(x)$ cannot tell you the prob directly but comparably it can tell which range is more likely

(ex)



$P(2 \le X \le 6)$ is more than $P(6 \le X \le 9)$

prob cannot be more than 1.
so $f(x)$ doesnt tells us prob.

$\boxed{f(5) = 10}$ is not prob.

→ Point probability is zero.

mathematically,

$P(X=a) \Rightarrow P(a \le X \le a) = 0$ Prob. of any point is zero.

$\int_a^a f(x) = 0$

$\Rightarrow P(X=a) = 0$

$\rightarrow P(a \le X \le b) = P(a < X \le b) = P(a \le X < b) = P(a < X < b)$

intuitively,

in DRV,

$P(X=6) = \frac{30}{100}$ ← no. of times $X=6$, ← total

in CRV, 0 ———— 10

$P(X=3.0012786390) = \frac{3}{\infty} = 0$ ← finite

ques. let X be a continuous RV with the following PDF.

$$f(x) = \begin{cases} ce^{-x} & , x \ge 0 \\ 0 & , \text{otherwise} \end{cases}$$

where c is a +ve constant

(a) Find c.

(b) Find $P(1<X<3)$.

↳ $\int_{-\infty}^{\infty} f(x) = 1 = \int_{-\infty}^{0} f(x) + \int_{0}^{\infty} f(x) = 1$ ↗0

$\Rightarrow \int_0^\infty ce^{-x} = 1 \Rightarrow -c\left[e^{-x}\right]_0^\infty = -c\left[\underbrace{\frac{1}{e^\infty}}_{0} - \underbrace{\frac{1}{e^0}}_{1}\right] = 1$

$\Rightarrow \boxed{c = 1}$

now, $P(1<X<3) = \int_1^3 f(x)\,dx$

$= \int_1^3 e^{-x}\,dx = \left[-e^{-x}\right]_1^3 = \boxed{e^{-1} - e^{-3}}$

ques. GATE CSE 2016

A PDF on the interval $[a,1]$ is given by $1/x^2$. and outside this interval the value of the funcⁿ is 0. The value of a is ___.

↳ $$f(x) = \begin{cases} 1/x^2, & a \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

$\int_a^1 f(x)\,dx = 1 \Rightarrow \int_a^1 \frac{1}{x^2}\,dx = 1 \Rightarrow \left[-\frac{1}{x}\right]_a^1 = 1 \Rightarrow -\frac{1}{1} + \frac{1}{a} = 1$

$\boxed{a = \frac{1}{2}}$

Discrete RV

- X can take only discrete values in a set

  eg : $X = \{0, 1, 2, \ldots, n\}$

- PMF $\quad P(X=x)$

- CDF

  $F(k) = P(X \le k) = \sum_{x \le k} P(X=x)$

- Mean

  $E[X] = \sum x P(x)$

- Varience

  $V(X) = E[X^2] - (E[X])^2$

Continuous RV

- X can take all possible values in an interval of real line

  eg $\quad X \in [0, 1]$

- PDF $\quad f(x)$

- CDF

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

- Mean

  $E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$

- Varience

  $V(X) = E[X^2] - (E[X])^2$

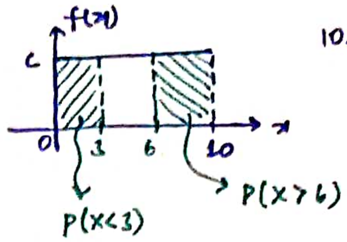
  $E[X^2] \rightarrow \int_{-\infty}^{\infty} x^2 f(x)\, dx$

UNIFORM DISTRIBUTION

ques. If X is uniformly distributed over $(0, 10)$, calculate the prob that

(a) $X < 3$ (b) $X > 6$ (c) $3 < X < 3$

$$\Rightarrow \quad f(x) = \begin{cases} 1/10 , & x \in [0, 10] \\ 0 , & \text{otherwise} \end{cases}$$



$10 \times c = 1$

$c = \frac{1}{10}$

$P(X<3) = \frac{1}{10} \times 3 = \frac{3}{10}$

$P(X>6) = \frac{1}{10} \times (10-6) = \frac{4}{10}$

$P(3<X<8) = \frac{1}{10} \times (8-3) = \frac{5}{10}$

OR

$P(X<3) = \int_0^3 \frac{1}{10}\, dx = \frac{3}{10}$

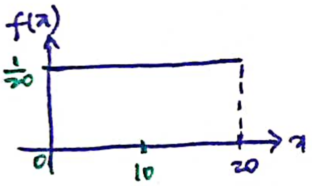
$P(X>6) = \int_6^{10} \frac{1}{10}\, dx = \frac{4}{10}$

$P(3<X<8) = \int_3^8 \frac{1}{10}\, dx = \frac{1}{2}$

ques. For the uniform PDF described ... with $a=0$ and $b=20$, $f(x) = \frac{1}{20} = 0.05$ for $x \in [0, 20]$.

$P(1 \le X \le 15) = ?$

$E[X] = ?$

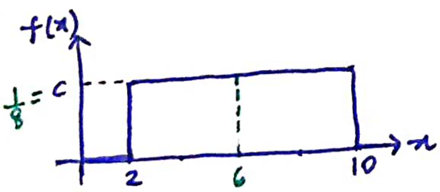


$P(1 \le X \le 15) = \frac{1}{20} \times (15-1) = \frac{14}{20}$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx = \frac{1}{20} \int_0^{20} x\, dx = \frac{1}{20} \left[\frac{x^2}{2} \right]_0^{20} = \cancel{\frac{1}{20}} \times \frac{\cancel{20} \times 20}{2}$$

$E[X] = 10$

ques. Bus is uniformly late b/w 2 and 10 minutes. How long can you expect to wait? with what SD? If its >7 mins late, you'll be late for work. What's the prob of you being late?



$c \times (10-2) = 1$

$c = \frac{1}{8}$

- $E[X] = \frac{10+2}{2} = 6$ mins
- $V[X] = \frac{(b-a)^2}{12} = \frac{8^2}{12}$

$\sigma = \sqrt{\frac{8^2}{12}} = \frac{8}{\sqrt{12}} = 2.31$ mins

- $P(7<X<10) = \frac{1}{8} \times (10-7) = \frac{3}{8}$

GEOMETRIC DISTRIBUTION

[Position are fixed] → no combination of p & q

$$\left.\begin{aligned} x &= 1 \rightarrow p \\ &= 2 \rightarrow qp \\ &= 3 \rightarrow qqp \\ &\;\;\vdots \end{aligned}\right\}$$

$$\boxed{P(x) = p^1 q^{x-1}}$$

Mean:

$E[x] = \sum x\, P(x)$

$= 1 \times p + 2 \times p \times q + 3 \times p q^2 + \cdots$

$= p\,[1 + 2q + 3q^2 + \cdots]$

$= p\,[1-q]^{-2}$

$= p \cdot p^{-2} = \boxed{1/p}$

$p + q = 1$

$p = 1 - q$

$$\boxed{E[x] = \frac{1}{p}}$$

Variance: $\boxed{V[x] = \frac{q}{p^2}}$

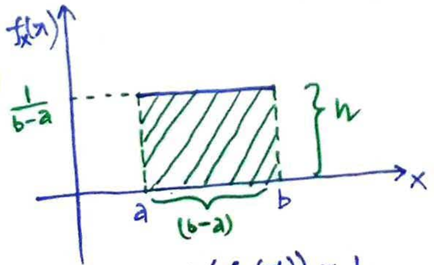
UNIFORM DISTRIBUTION

When a RV 'x' is equally likely to be observed in a finite range & is likely to have zero value outside this finite range, then the RV is said to have a uniform distribution.

→ $X \sim \text{uniform}(a, b)$ means that $x \in [a, b]$, $a \rightarrow \min^m$ value, $b \rightarrow \max^m$ value

→ X is a continuous RV which is in between a & b.

PDF



$A(f_x(x)) = 1$

$(b-a) \cdot h = 1 \Rightarrow h = \frac{1}{(b-a)}$

→ X = continuous RV

→ $\int_{-\infty}^{\infty} f_x(x) = 1$

pdf

$$f_x(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & \text{otherwise} \end{cases}$$

CDF

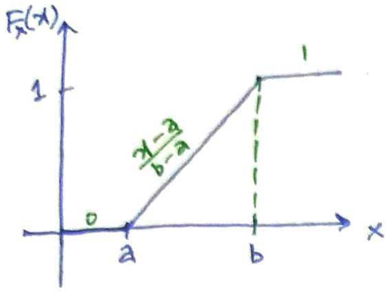
$F(X<a) = 0$

$F(X>b) = 1$

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(x)\,dx$$

$$= \int_{-\infty}^{a} 0\,dx + \int_{a}^{x} \frac{1}{b-a}\,dx \quad ; \quad a<x<b$$

$$= \frac{1}{b-a}\left[x\right]_a^x = \frac{x-a}{b-a} \quad ; \quad a<x<b$$



$$F_X(x) = \begin{cases} 0 & \text{if } x<a \\ \frac{x-a}{b-a}, & x \in [a,b] \\ 1, & x \ge b \end{cases}$$

Mean:

$$\mu = E[X]$$

$$= \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-\infty}^{a} x \cdot 0\,dx + \int_{a}^{b} x \cdot \frac{1}{(b-a)}\,dx + \int_{b}^{\infty} x \cdot 0\,dx$$

$$= \left[\frac{x^2}{2(b-a)}\right]_a^b = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

$$\boxed{\mu = E[X] = \frac{a+b}{2}}$$

Variance:

$$V[X] = E[X^2] - [E(X)]^2$$

$$= \int_a^b x^2 \frac{1}{b-a}\,dx + \left(\frac{a+b}{2}\right)^2 =$$

$$= \left[\frac{x^3}{3}\right]_a^b \times \frac{1}{(b-a)} + \frac{(a+b)^2}{4} = \frac{b^3-a^3}{(b-a)\,3} + \frac{(a+b)^2}{4}$$

$$= \frac{1}{12}(b-a)^2$$

$$\boxed{V[X] = \frac{1}{12}(b-a)^2}$$

Ques. A R.V is uniformly distributed over the interval 2 to 10. Its variance will be -

(A) $\frac{16}{3}$ (B) 6 (C) $\frac{256}{9}$ (D) 36

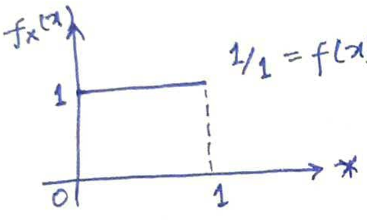
Sol^n

$$V(x) = \frac{(b-a)^2}{12} = \frac{(10-2)^2}{12} = \frac{64}{12} = \boxed{\frac{16}{3}}$$

Ques. X is uniformly distributed random variable that takes values between 0 & 1. The value of $E[x^3]$ will be,

(A) 0 (B) 1/8 (C) 1/4 (D) 1/2

Sol^n



$1/1 = f(x)$, $0 < x < 1$

Now, CRV ✓

$$E[x^3] = \int_{x=-\infty}^{\infty} x^3 f(x)\,dx$$

$$= \int_{x=0}^{1} x^3 \cdot 1\,dx = \left[\frac{x^4}{4}\right]_0^1 = \boxed{\frac{1}{4}}$$

Some real life examples of Poisson's Distribution

1) No. of calls received at a call centre per hour

↳ Requirement & Assumptions :

a) Calls happen at a constant rate (λ)

b) occurance of one call doesnot affect another.

c) Two or more events/calls cannot accur simultaneously.

2) No. of ~~pt~~ patents arriving at the emergency b/w 10 PM - 11 PM.

3) No. of customers at the counter per hr.

4) No. of insurance claims in a year.

5) No. of goals in a sports event in b/w 2 teams.

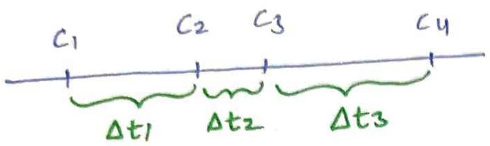
6) No. of visitors on your website per minute.

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

EXPONENTIAL DISTRIBUTION

→ Continuous Distribution → continuous R.V

→ lets take an example of a poisson process for example the calls received at a call centre.
↳ generates poisson distributed RV



we don't know the distribution b/w Δt_i's (uneven distribution) so we need exponential distribution.

If c_1, c_2, c_3, c_4 ... etc follows the poisson distribution/ process then the time interval in between the two poisson events follows exponential distribution, in the above figure Δt_1, Δt_2, Δt_3 etc follow exponential distribution.

P (10 calls in 1 hr) → can be answer using poisson's distribution

but, P (next event/call occurs in 3 min) → can't be answered with poisson's distribution But Exponential distribution can.

↓ Assume c_2 has happened & ~~next~~ next event is c_3

similar to

$P(\Delta t_2 < 3 \text{ min})$

★ Poisson distribution measures the no. of events per unit time.
↳ parameter is 'λ' (rate) ⇒ $X \sim \text{Poisson}(\lambda)$

Exponential distribution is the distribution of time intervals b/w events in a poisson process.
↳ parameter is 'λ' (rate) ⇒ $X \sim \text{Expo}(\lambda)$

So, ~~if~~ the Δt_i's (time intervals b/w events), if we think of them as RV ~~[illegible]~~ 'x' it follows Exponential distribution. $X \sim \text{Expo}(\lambda)$.

★ **PDF**

$$\boxed{f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}}$$

CDF

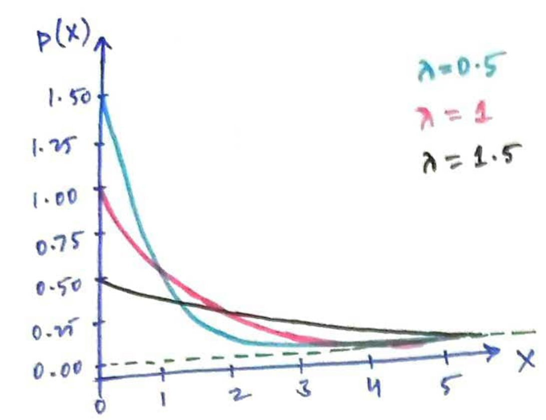
$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx = \cancel{\int_{-\infty}^{0} 0\,dx}^{\,0} + \int_{0}^{x} \lambda e^{-\lambda x}\,dx$$

$$= \lambda \int_0^x e^{-\lambda x}\,dx = \lambda \left(\frac{e^{-\lambda x}}{-\lambda}\right)_0^x = 1 - e^{-\lambda x}, x \geq 0$$

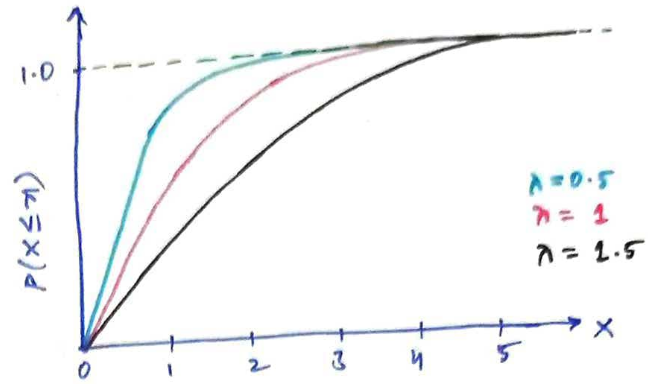
$$\boxed{F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}}$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

PDF for Exponential distribution



CDF for Exponential distribution



- **MEAN** : $\mu = E(X)$

$$= \int_0^\infty \underset{\mathrm{I}}{x}\,(\underset{\mathrm{II}}{\lambda e^{-\lambda x}})\,dx \quad \text{using Integration by parts}$$

$$= \left[-x e^{-\lambda x}\right]_0^\infty + \int_0^\infty e^{-\lambda x}\,dx = 0 + \left[\frac{e^{-\lambda x}}{-\lambda}\right]_0^\infty = \frac{1}{\lambda}$$

$$\boxed{\mu = E[x] = \frac{1}{\lambda}}$$

- **VARIANE** : $V(x) = E(x^2) - [E(x)]^2$

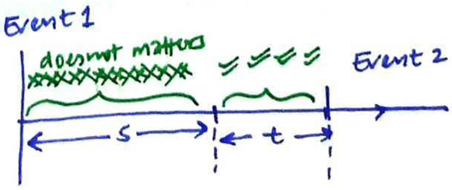
$$= E(x^2) - \mu^2 = \int_0^\infty x^2 \lambda e^{-\lambda x} - \frac{1}{\lambda^2} \quad \text{on doing integration by parts.}$$

$$= \frac{1}{\lambda^2}$$

$$\boxed{V(x) = \frac{1}{\lambda^2}}$$

★⇒ Memoryless Property of Exponential Distribution

$$\boxed{P(X > s+t \,/\, X > s) = P(X > t)} \quad \forall s,t \geq 0$$



Prob that time interval b/w 2 events is 's+t' given that time 's' has already passed

$$= \frac{P(X > s+t \wedge X > s)}{P(X > s)} = \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{1 - P(X \leq s+t)}{1 - P(X \leq s)} \Rightarrow \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t}$$

$$= P(X > t)$$

Ques. Assume that the duration in minutes of a telephone conversation follows the exponential distribution.

$f(x) = \frac{1}{5} e^{-x/5}, \ x \geq 0$. The prob that the conversation will exceed 5 minutes is :

(A) $1/e$ (B) $1 - 1/e$ (C) $1/e^2$ (D) $1 - 1/e^2$

Sol^n prob that the conv. will exceed 5 min

$$\hookrightarrow P(X > 5) = \int_{x=5}^{\infty} f(x)\, dx$$

$$= \int_{5}^{\infty} \frac{1}{5} e^{-x/5} dx = \frac{1}{5}\left[\frac{e^{-x/5}}{-1/5}\right]_5^{\infty}$$

$$= -1\left[e^{-x/5}\right]_5^{\infty} = -1\left[e^{-\infty} - e^{-5/5}\right] = e^{-1} = \boxed{1/e}$$

Prob. Dist

$\swarrow$ Discrete RV

1) Bernoulli D.
2) Binomial D.
3) Hyper Geometric D.
4) Poisson D.
5) Geometric D.

$\searrow$ Continuous RV

1) Uniform D.
2) Exponential D.
3) Normal D.

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

PROBABILITY DISTRIBUTION

NORMAL / GAUSSIAN DISTRIBUTION

→ used for continuous RV. ⇒ continuous distribution.

→ most widely used & popular distribution.

→ ex : height, length of leaves, weights of people in a set of population. ↗ All are natural phenomenon

If not normally distributed they are approximately Normal or Gaussian distributed

→ It is found in so many natural phenomenons that a lot of mathematicians / staticians when prove something like a theoritical model, & have to make an assumption on the distribution of a R.V 'x', if they donot know exactly whether it is a Binomial or poisson or any other. The most typical assumption they make that 'x' follows Normal distribution. → so very useful ✓

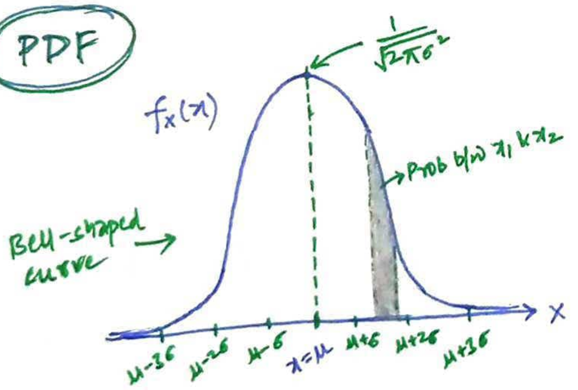
→ A R.V 'x' is normally distributed, represented as.

$$X \sim N(\underbrace{\mu}_{\text{mean}}, \underbrace{\sigma^2}_{\text{variance}})$$

μ & σ^2 should be finite.

→ σ = std. dev = $+\sqrt{\text{variance}}$

PDF



$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}}, \quad \sigma \neq 0$$

① X → continuous R.V

② μ = mean

σ^2 = Variance ⇒ $\sqrt{\sigma^2} = \sqrt{V}$ = SD

σ = Standard Deviation

③ Symmetric curve (about, $x = \mu$)

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

$$\Rightarrow \int_{\mu}^{\infty} f_X(x)\,dx = 0.5 \quad ; \quad \int_{-\infty}^{\mu} f_X(x)\,dx = 0.5$$

Now, $P[x_1 < x < x_2] = \int_{x_1}^{x_2} f_X(x)\,dx$

$$= \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \, dx$$

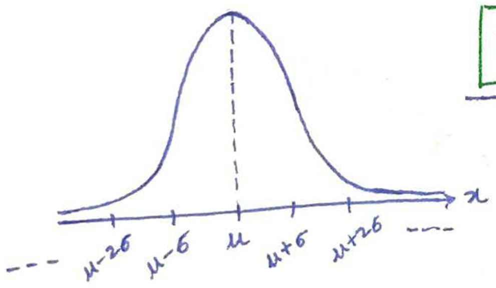
it is difficult to solve so we use a trick ~~of~~ "standard Normal distribution"

TRICK N·D (difficult) $\longrightarrow$ Z·D (Easy)

★ Standard Normal Distribution : If $X \sim N(\mu, \sigma^2)$ then we can convert it to another normal variable 'z' with $\mu = 0$ & $\sigma^2 = 1$ by the following transformation :

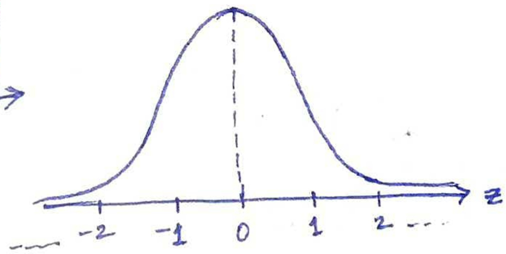
$$\boxed{z = \frac{x-\mu}{\sigma} \sim N(0,1)}$$

N·D (difficult) $\xrightarrow{z = \frac{x-\mu}{\sigma}}$ Z·D (easy)



[Normal Dist.]

$$\boxed{z = \frac{x-\mu}{\sigma}} \longrightarrow$$



[Z Dist.] standard Normal Dist.

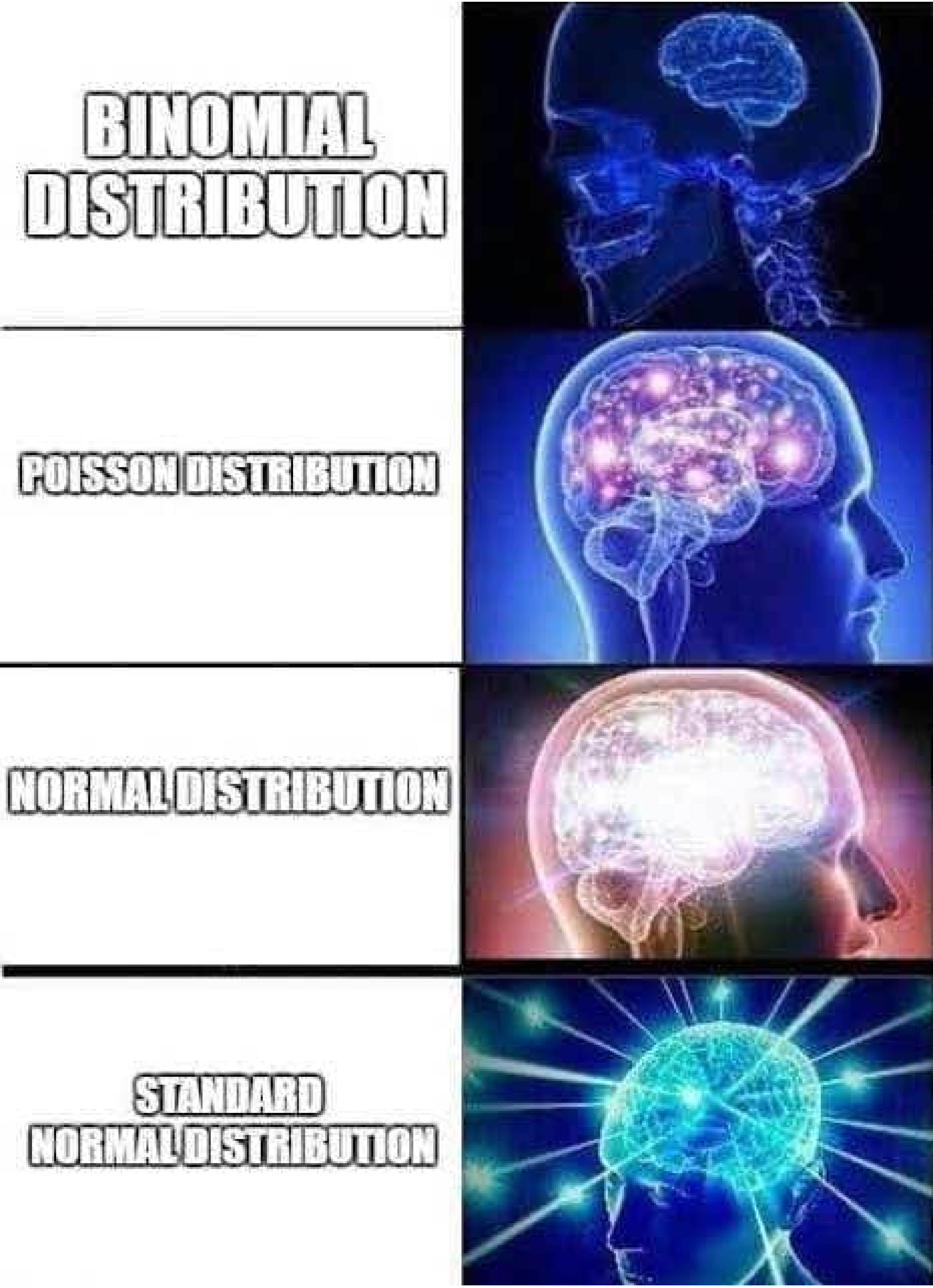
$z = \frac{x-\mu}{\sigma}$ $\Rightarrow \sigma z = x - \mu$

$\therefore x = \mu + \sigma z \Rightarrow dx = 0 + \sigma dz$

Now, $P[x_1 < x < x_2] = P\left[\frac{x_1-\mu}{\sigma} < \frac{x-\mu}{\sigma} < \frac{x_2-\mu}{\sigma}\right] = P[z_1 < z < z_2]$

$$= \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\cancel{\sigma}} e^{-\frac{z^2}{2}} \cdot \cancel{\sigma}\, dz$$

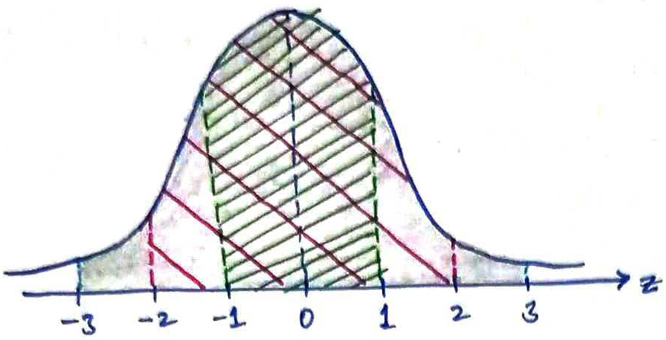
$$= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-z^2/2}\, dz$$ ✓



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Z distribution :

↳ 99% chances are that our prob will be in the region, $-3 < z < 3$.



$P(-1 < z < 1) = 0.6827$ $P(-3 < z < 3) = 0.9973$

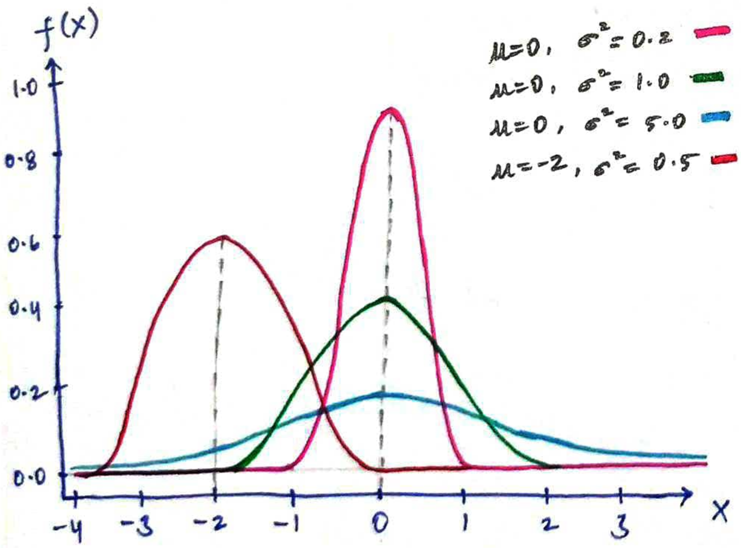
$P(-2 < z < 2) = 0.9545$

→ $P(z > 0) = 0.5$
→ $P(z < 0) = 0.5$ } b/w z symmetric graph & overall area = 1.

→ $P(0 < z < 1) = ??$

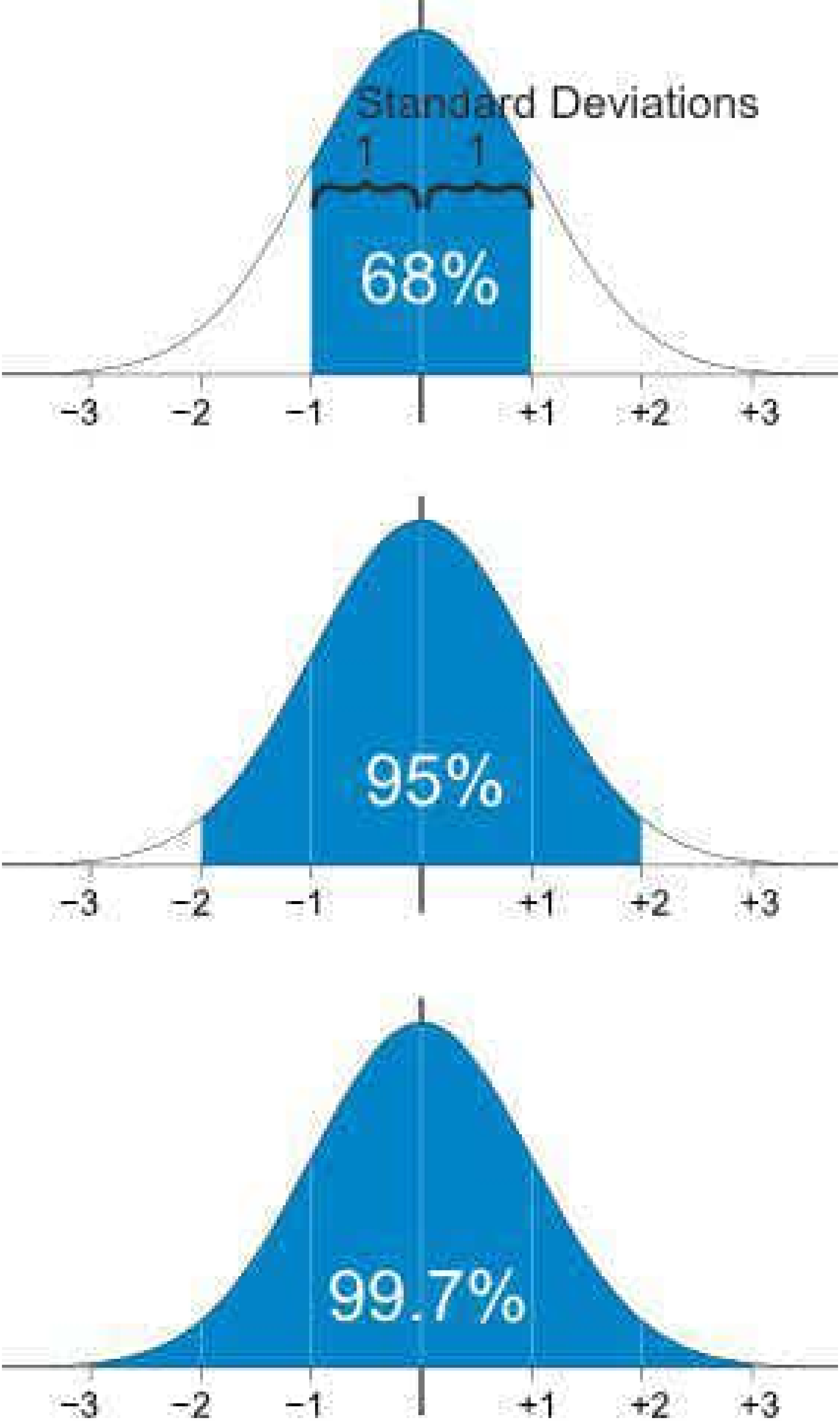
↳ plot is symmetric ⇒ $P(0 < z < 1) = \frac{P(-1<z<1)}{2} = \frac{0.6827}{2} = 0.34135$ ✓

★ PDF for Normal distribution — variations



As variance increases, the peakedness of the curves decreases in the PDF
→ variance ↑ peak ↓

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques. A nationalized bank has found that the daily balance available in its savings accounts follows a normal distribution with a mean of Rs 500. & a standard deviation of Rs 50. The percentage of savings account holders, who maintain an average daily balance more than Rs 500 is ___.

Solⁿ ① Normal distribution ✓

$\mu = 500$

$\sigma = 50$

$$P(X > 500) = P(X > \mu) = \int_{\mu}^{\infty} f(x)\,dx = 0.5$$

$$\% = 0.5 \times 100 = \boxed{50\%}$$

Method 2 : $P[X > 500]$

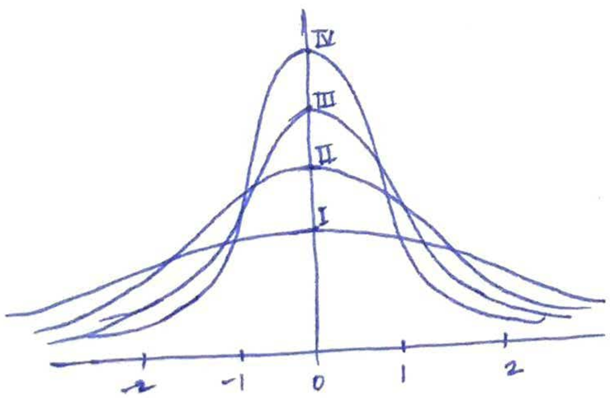
convert to z distribution

$$= P\left[\frac{x-\mu}{\sigma} > \frac{500-\mu}{\sigma}\right] = P\left[z > \frac{500-500}{50}\right] = P[z > 0]$$

$$= 0.5$$

$$\% = 0.5 \times 100 = \boxed{50\%}$$

ques. Which one has the lowest variance



(A) I (B) II

(C) III (D) IV.

Solⁿ Variance highest → Peak lowest

Variance lowest → Peak highest

↳ IV ✓ → lowest variance.

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Ques. The annual precipitation data of a city is normally distributed with mean & standard deviation as 1000 mm & 200 mm, respectively. The probability that the annual precipitation will be more than 1200 mm is.

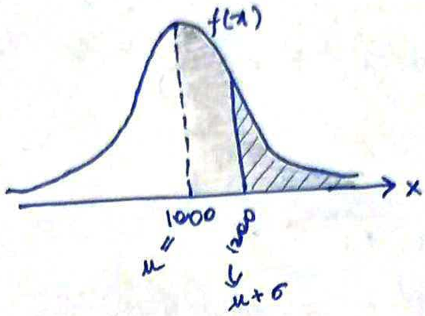
(A) < 50 % (B) 50 % (C) 75 % (D) 100 %

Soln ① Normal dist ✓

$\mu = 1000$

$\sigma = 200$

$p(x > 1200) = ??$



$P[X > \mu] = 50\%$

↳ $P[X > 1200]$ would obviously $< 50\%$.

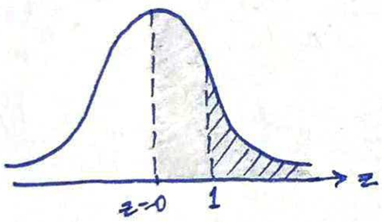
So (A) is correct

Method 2 z-distribution

$$P[X > 1200] = P\left[\frac{X-\mu}{\sigma} > \frac{1200-\mu}{\sigma}\right]$$

$$\boxed{Z = \frac{X-\mu}{\sigma}}$$

$$= P\left[Z > \frac{1200-1000}{200}\right] = P[Z > 1]$$



$$P[Z > 1] = P[Z > 0] - P[0 < Z < 1]$$

$$= 0.5 - \frac{0.6827}{2} \approx \boxed{0.16}$$

Ques. $f_X(x) = \frac{1}{\sqrt{8\pi}} e^{-\left\{\frac{(x-1)^2}{8}\right\}}$, $-\infty < x < \infty$ then $\int_1^{\infty} f_X(x)\,dx =$

(A) 0 (B) 1/2 (C) $1 - \frac{1}{e}$ (D) 1

Soln $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ ⇒ $2\pi\sigma^2 = 8\pi$ ⇒ $\sigma^2 = 4$

$\sigma = 2$ ($\sigma \neq -2$ bcoz S.D can't be -ve)

⇒ $\mu = 1$

Now, $\int_1^{\infty} f_X(x)\,dx = \int_\mu^{\infty} f_X(x)\,dx = \boxed{0.5}$ ✓

ques. Consider the following 2 Normal distributions.

$$f_1(x) = e^{-\pi x^2} \qquad f_2(x) = \frac{1}{2\pi} e^{\left\{-\frac{1}{4\pi}(x^2+2x+1)\right\}}$$

If μ & σ denote the mean & standard deviation, respectively, then

(A) $\mu_1 < \mu_2$ & $\sigma_1^2 < \sigma_2^2$

(B) $\mu_1 < \mu_2$ & $\sigma_1^2 > \sigma_2^2$

(C) $\mu_1 > \mu_2$ & $\sigma_1^2 < \sigma_2^2$

(D) $\mu_1 > \mu_2$ & $\sigma_1^2 > \sigma_2^2$

Sol^n ND : $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

(i) $f_1(x) = e^{-\pi x^2}$

$\Rightarrow \cancel{-}\pi x^2 = \cancel{-}\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_1^2}$

$\Rightarrow \pi(x-0)^2 = \frac{1}{2\sigma_1^2}(x-\mu_1)^2$

$\Rightarrow \pi = \frac{1}{2\sigma_1^2}$, $\mu_1 = 0$ ✓

$\hookrightarrow \sigma_1^2 = \frac{1}{2\pi}$ ✓

(ii) $f_2(x) = \frac{1}{2\pi} e^{\left\{-\frac{1}{4\pi}(x^2+2x+1)\right\}}$

$\Rightarrow \cancel{-}\frac{1}{2\,\cancel{4}\pi}(x^2+2x+1) = \cancel{-}\frac{(x-\mu_2)^2}{\cancel{2}\sigma_2^2}$

$\Rightarrow \frac{[x-(-1)]^2}{2\pi} = \frac{(x-\mu_2)^2}{\sigma_2^2}$

$\Rightarrow \mu_2 = -1$ ✓ , $\sigma_2^2 = 2\pi$ ✓

$\Rightarrow \boxed{\mu_1 > \mu_2 ,\ \sigma_1^2 < \sigma_2^2}$

ques. $I = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x^2/8}\, dx$

Sol^n Method 1 : using Euler's Gamma function : $\left\{ \int_0^\infty e^{-t} t^{n-1} dt = \Gamma n \right\}$

$\Rightarrow$ let, $\frac{x^2}{8} = t \Rightarrow x^2 = 8t \Rightarrow 2x\,dx = 8\,dt$

$\hookrightarrow x = \sqrt{8t}$ $\qquad dx = \frac{4\,dt}{x} \Rightarrow dx = \frac{4\,dt}{\sqrt{8t}}$

x	0	∞
t	0	∞

After changing variable & substituting :→

$\Rightarrow I = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-t} \frac{4\,dt}{\sqrt{8t}} \Rightarrow I = \frac{4}{\sqrt{8}\sqrt{2\pi}} \int_0^\infty e^{-t} t^{-1/2}\, dt$

$\Rightarrow I = \frac{1}{\sqrt{\pi}} \int_0^\infty e^{-t} t^{-1/2}\, dt \Rightarrow I = \frac{1}{\sqrt{\pi}} \times \Gamma\frac{1}{2} \Rightarrow I = \frac{1}{\cancel{\sqrt{\pi}}} \times \cancel{\sqrt{\pi}} \Rightarrow \boxed{I = 1}$

$-\frac{1}{2} = n-1$

$\Rightarrow n = 1 - \frac{1}{2} = \frac{1}{2}$

Method 2 : $I = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-x^2/8}\, dx \Rightarrow I = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/8}\, dx$

In Normal dis : $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

comparing powers of exponent :→

$$\Rightarrow \frac{-x^2}{8} = -\frac{(x-\mu)^2}{2\sigma^2}$$

$$\Rightarrow \frac{(x-0)^2}{4} = \frac{(x-\mu)^2}{\sigma^2} \longrightarrow \mu = 0, \quad \sigma^2 = 4 \Rightarrow \sigma = 2$$

Now,

$$I = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/8}\, dx$$

'σ' term is missing here, so we'll add it.

$$I = \sqrt{2^2} \int_0^\infty \frac{1}{\sqrt{2\pi}\sqrt{2^2}} e^{-x^2/8}\, dx$$

$$I = 2 \int_0^\infty \frac{1}{\sqrt{2\pi(2)^2}} e^{-\frac{(x-0)^2}{2(2)^2}}\, dx$$

$$I = 2 \int_\mu^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = 2 \times \frac{1}{2} = \boxed{1} \checkmark$$

ques. An exponential random variable is given such that its CDF is given by →

$$F_Z(z) = \begin{cases} 1 - e^{-z} & ; z \geq 0 \\ 0 & ; \text{otherwise} \end{cases}$$

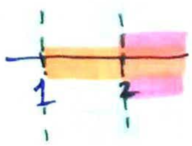
Find : probability ⇒ $P[z>2 / z>1]$.

pdf : $f(z) = \frac{d}{dz}[cdf]$

$$\Rightarrow f(z) = \frac{d}{dz}[1 - e^{-z}] = e^{-z}$$

Solⁿ $P[A/B] = \frac{P[A \cap B]}{P[B]}$

$$\Rightarrow P[z>2 / z>1] = \frac{P[z>1,\ z>2]}{P[z>1]}$$



$$= \frac{P[z>2]}{P[z>1]} = \frac{\int_{z=2}^\infty f(z)\, dz}{\int_{z=1}^\infty f(z)\, dz} = \frac{\int_2^\infty e^{-z}\, dz}{\int_1^\infty e^{-z}\, dz} = e^{-1} = \boxed{0.37}$$

(Or) Method 2 making use of the cdf directly

$$\frac{P[z>2]}{P[z>1]} = \frac{1 - P[z \leq 2]}{1 - P[z \leq 1]} = \frac{1 - F_Z(z=2)}{1 - F_Z(z=1)} = \frac{1 - [1 - e^{-2}]}{1 - [1 - e^{-1}]} = \frac{e^{-2}}{e^{-1}} = e^{-1} = \boxed{0.37}$$

ques. If calls arrive at a telephone exchange such that the time of arrival of any call is independent of the time ~~interval~~ of arrival of earlier or future calls, the probability distribution function of the total no. of calls in a fixed time interval will be.

(A) Poisson (B) Gaussian (C) Exponential (D) Gamma

Solⁿ Total no. of calls are specific ⇒ discrete RV

⇒ Poisson ✓ bcoz Gaussian, Exponential → continuous RV

Gamma is not present

ques. The second moment of a Poisson-distributed RV is 2. The mean of the RV is ____ .

Solⁿ $E[X^n]$, $n=1$ (first moment)

$n=2$ (second moment)

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

Now, In Poisson's : $E[X] = V[X] = \lambda$ (let)

$\hookrightarrow V[X] = E[X^2] - [E[X]]^2$

$\Rightarrow V[X] = 2 - (E[X])^2$

$\Rightarrow \lambda = 2 - \lambda^2 \Rightarrow \lambda^2 + \lambda - 2 = 0$

$\Rightarrow \lambda = 1, \cancel{-2}$ ↳ variance can't be -ve

$\Rightarrow$ variance = $\boxed{\text{Mean} = 1}$

ques. The prob of a resistor being defective is 0.02. There are 50 such resistors in a circuit. The prob of 2 or more defective resistors in the circuit is.

Solⁿ $P(D) = 0.02$

$n = 50$

$P(X \geq 2) = ?$

$P(X \geq 2) = 1 - P(X < 2)$

$= 1 - P(X=0) + P(X=1)$

Method 1 [Binomial]

$P(X) = {}^nC_x \, p^x q^{n-x}$

$n = 50$

$p = 0.02 \quad q = 1 - 0.02 = 0.98$

$x = 0, 1$

$\Rightarrow 1 - P(X=0) + P(X=1)$

$\Rightarrow 1 - {}^{50}C_0 p^0 q^{50} + {}^{50}C_1 p^1 q^{49}$

$\Rightarrow 1 - (0.98)^{50} + 50 \times 0.02 \times (0.98)^{49}$

$\Rightarrow P[X \geq 2] = 1 - 0.73577$

$= \boxed{0.264}$

Method 2 [Poisson]

↳ $\lambda = np$

$\lambda = 50 \times 0.02 = 1$

$P(X) = \dfrac{e^{-\lambda}\lambda^x}{x!} = \dfrac{e^{-1} \cdot 1^x}{x!}$

$P(X \geq 2) = 1 - P[X < 2]$

$= 1 - (P(X=0) + P(X=1))$

$= 1 - e^{-1}\left[\dfrac{1^0}{0!} + \dfrac{1^1}{1!}\right]$

$= 1 - e^{-1}(2) = \boxed{0.26}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

ques. Two random variables X & Y are distributed according to

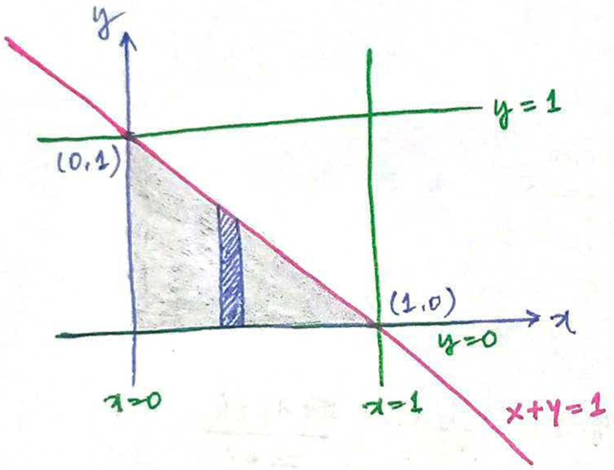
$$f_{X,Y}(x,y) = \begin{cases} (x+y), & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

The prob $P(X+Y \le 1)$ is ____.

Sol^n

$$= \int_{x=}\int_{y=} f(x,y)\,dx\,dy$$

→ here we are having difficulty in finding $f(x,y)$ & the limits of integration so we have to use the concept of calculus.



using MARIO strip concept

$x = 0$ to 1

$y = 0$ to $1-x$

Now,

$$\int_{x=0}^{1}\int_{y=0}^{1-x} (x+y)\,dx\,dy = \boxed{0.33}$$

ques. If X & Y are RV. such that $E[2X+Y]=0$ & $E[X+2Y]=33$, then $E[X]+E[Y] = ??$

Sol^n

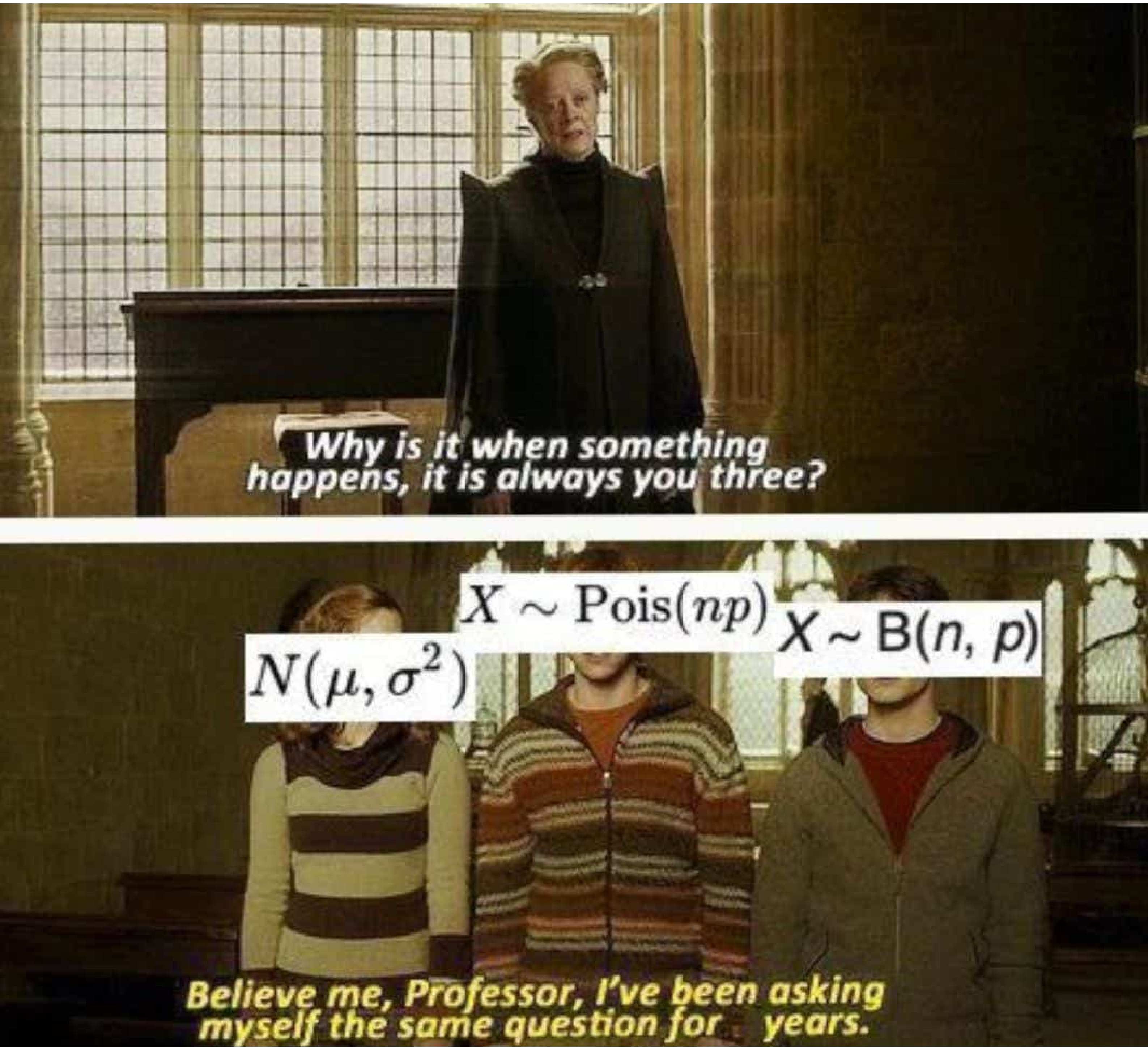
$E[2X+Y]=0 \rightarrow E[2X]+E[Y]=0 \rightarrow 2E[X]+E[Y]=0$

$E[X+2Y]=33 \rightarrow E[X]+E[2Y]=33 \rightarrow E[X]+2E[Y]=33$

$3E[X]+3E[Y]=33$

$\Rightarrow \boxed{E[X]+E[Y]=11}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

STATISTICS & PYQs

1) MEAN = Average

↳ [Un Grouped Data]

(ex) : {1, 8, 3, 2, 6}

$$\text{Avg} = \frac{\sum x_i}{n} = \frac{1+8+3+2+6}{5} = \frac{20}{5} = \boxed{4}$$

↳ [Grouped Data]

(ex)

x_i	f_i
8	1
6	1
5	3
7	2
4	2
9	1

f_i → frequency → repetition

$\xrightarrow[\text{as ungrouped data}]{\text{can be written}}$ {8, 6, 5, 5, 5, 7, 7, 4, 4, 9}

$$\text{Avg} = \frac{\sum x_i}{n} = \frac{\sum x_i f_i}{\sum f_i}$$

$$= \frac{8+6+15+14+8+9}{10} = \frac{60}{10} = \boxed{6}$$

Ques.

Flow rate (litre/sec)	7.5 to 7.7	7.7 to 7.9	7.9 to 8.1	8.1 to 8.3	8.3 to 8.5	8.5 to 8.7
Frequency	1	5	35	17	12	10

Mean flow rate of the liquid is

(A) 8.00 l/s (B) 8.06 l/s (C) 8.16 l/s (D) 8.26 l/s

Sol^n take avg of flow rate as $\left(\frac{\text{upper value} + \text{lower value}}{2}\right)$ ex: $x_1 = \frac{7.5+7.7}{2} = 7.6$

Now,
$$\text{Avg} = \frac{\sum x_i f_i}{\sum f_i} = \frac{(1\times7.6)+(5\times7.8)+(35\times8)+(17\times8.2)+(12\times8.4)+(10\times8.6)}{80}$$

$$= \frac{652.8}{80}$$

$$= \boxed{8.16\ l/s}$$

Modified table will be using $x_i = \frac{(x_i)_{upper} + (x_i)_{lower}}{2}$

x_i	7.6	7.8	8	8.2	8.4	8.6
f_i	1	5	35	17	12	10

2) MEDIAN

- Step 1 : Sort all the values (either Ascending / Descending).
- Step 2 : Select the middle value.

ex) {2, 1, 3}

↳ 1, 2, 3

Median = 2

- Step 3 : If we have even no. of values, we need to take avg of middle 2 elements.

ex) {1, 4, 2, 3}

↳ 1, $\boxed{2, 3}$, 4

Median = $\frac{2+3}{2}$ = 2.5

Median → n = odd → $\left(\frac{N+1}{2}\right)^{th}$ term

Median → n = even → $\frac{(N/2)^{th} + (N/2+1)^{th}}{2}$

ex) {1, 2, 7, 6, 8}

↳ 1, 2, 6, 7, 8 ⇒ Median = $\left(\frac{5+1}{2}\right)^{th}$ = 3^{rd} term ↳ $\boxed{6}$

Median

Mean = $\frac{1+2+6+7+8}{5}$ = $\boxed{4.8}$

ques. The spot speeds (expressed in km/h) observed at a road section are 66, 62, 45, 79, 32, 51, 56, 60, 53 & 49. The median speed is ??

soln In Ascending order : 32, 45, 49, 51, 53, 56, 60, 62, 66, 79

n = 10 ⇒ Median → $\frac{(10/2)^{th} + (10/2+1)^{th}}{2}$ = $\frac{5^{th}+6^{th}}{2}$ = $\frac{53+56}{2}$

= $\boxed{54.5}$

ques. The marks obtained by a set of students are:

38, 84, 45, 70, 75, 60, 48

The mean & median marks, respectively are

(A) 45 & 75 (B) 55 & 48 (C) 60 & 60 (D) 60 & 70

sol^n Mean = $\frac{38+84+45+70+75+60+48}{7} = \frac{420}{7} = \underline{60}$

In Ascending order: 38, 45, 48, 60, 70, 75, 84

Median = $\left(\frac{N+1}{2}\right)^{th}$ term = $\left(\frac{7+1}{2}\right)^{th}$ = 4^{th} term → $\underline{60}$

3) MODE

The value that is observed most frequently i.e. has highest freq^n in x_i

ex)

x_i	f_i
1	1
2	2
3	3
7	4
10	2

→ 7 is mode of the data set

ex)

Height	frequency
150	10
155	20
160	70
170	15
180	4

→ 160 is mode of the data set.

ex) {1, 2, 2, 3, 3, 3} → $\boxed{\text{Mode} = 3}$ → UNIMODEL → only one value of mode

ex) {1, 2, 2, 3, 3} → $\boxed{\text{Mode} = 2 \,\&\, 3}$ → BIMODEL → 2 values of mode

ex) {1, 2, 2, 3, 3, 4, 4} → $\boxed{\text{Mode} = 2, 3 \,\&\, 4}$ → TRIMODEL

ex) {1, 2, 2, 3, 3, 4, 4, 5, 5} → $\boxed{\text{Mode} = 2, 3, 4 \,\&\, 5}$ → MULTIMODEL

ex) {1, 2, 3, 4, 5} → All have same frequency

↳ $\boxed{\text{No Mode}}$

ques. Marks obtained by 100 students in an examination are given in the table

S NO.	Marks obtained	No. of students
1	25	20
2	30	20
3	35	40
4	40	20

What would be the mean, median and mode of the marks obtained by the students?

(A) Mean : 33 , Median : 35 , Mode : 40
(B) Mean : 35 , Median : 32.5 , Mode : 40
(C) Mean : 33 , Median : 35 , Mode : 35
(d) Mean : 35 , Median : 32.5 , Mode : 35

Sol^n No of students → f_i
Marks obtained → x_i

⇒ Mode = 35 as its frequency is max^m. {35 came 40 times which is highest}

⇒ Mean : $\dfrac{\sum x_i f_i}{\sum f_i} = \dfrac{25(20) + 30(20) + 35(40) + 40(20)}{100} = \dfrac{3300}{100} = 33$

⇒ Median :

1 2 3 ----- 20 → 25 marks
21 ---------- 40 → 30 marks
41 ---------- 80 } 35 marks
81 ---------- 100 → 40 marks

$n = 100$

$$\Rightarrow \frac{\left(\frac{100}{2}^{th}\right) + \left(\frac{100}{2}+1\right)^{th}}{2} = \frac{50^{th} + 51^{th}}{2} = \frac{35+35}{2}$$

⇒ Median = 35

(M2) for Median using Cummulative frequency Method

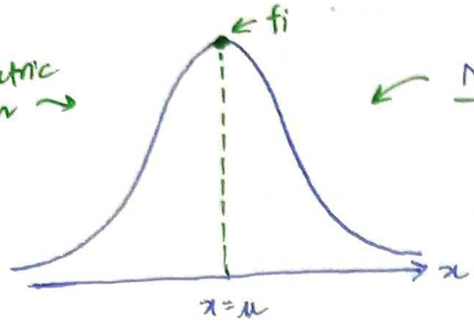
S.NO	x_i	f_i	cf
1	25	20	20
2	30	20	40
3	35	40	80 ← 50^th & 51^th term will be here
4	40	20	100

$$\text{Median} \rightarrow \frac{50^{th} + 51^{th}}{2} = \frac{35+35}{2} = 35$$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

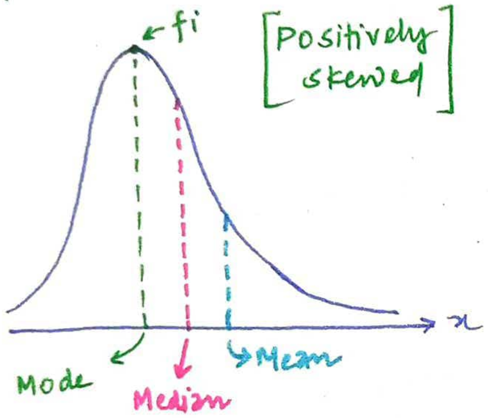
Normal distribution

Symmetric graph



Mode = Median = Mean

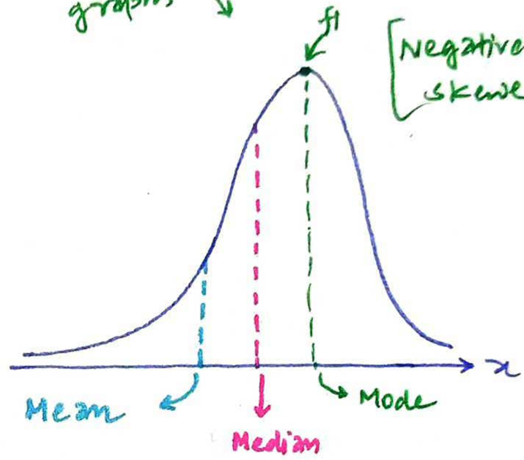
Assymetric graphs



[Positively skewed]

Mode ≤ Median ≤ Mean

Assymetric graphs



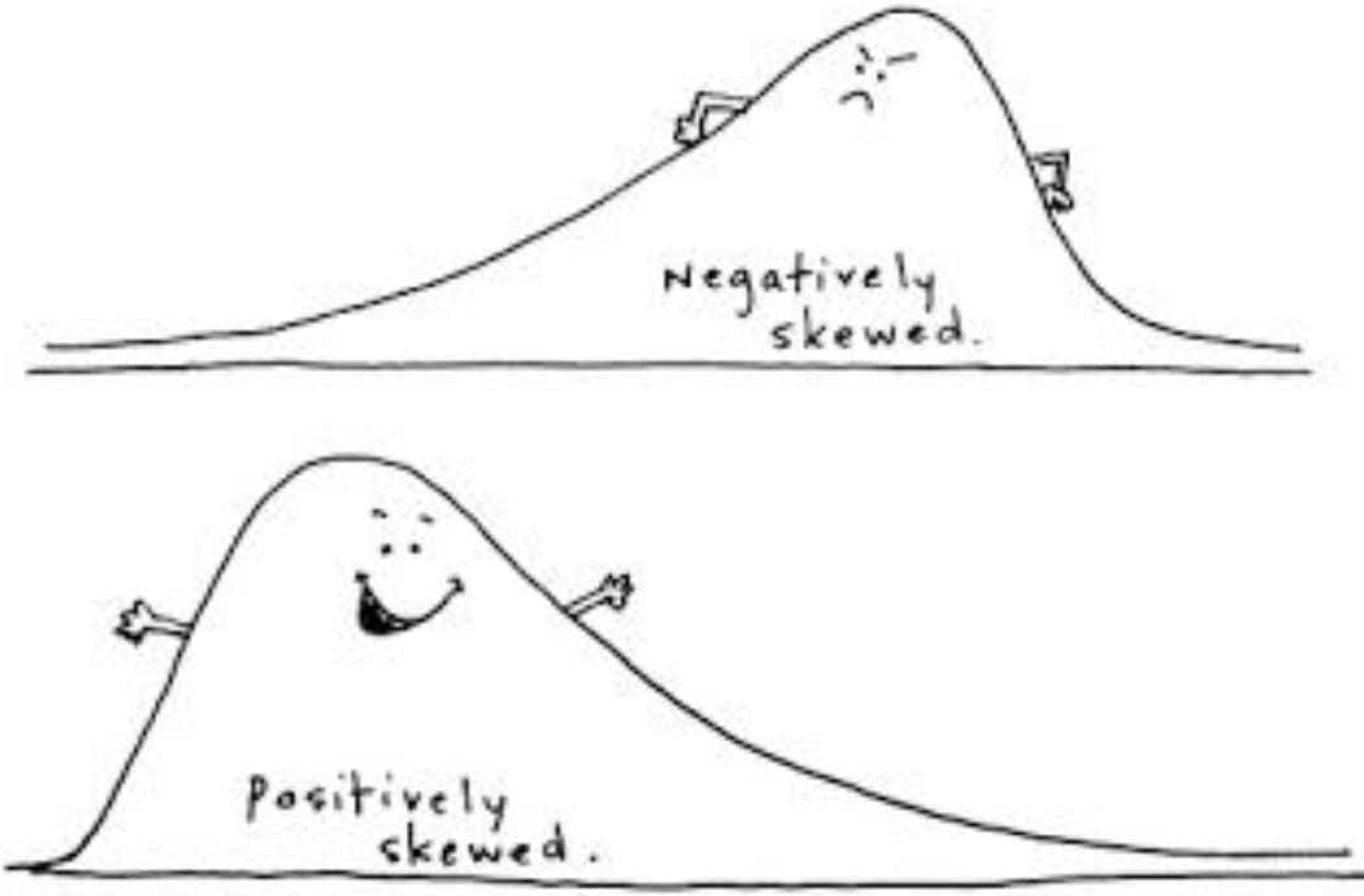
[Negatively skewed]

Mode ≥ Median ≥ Mean

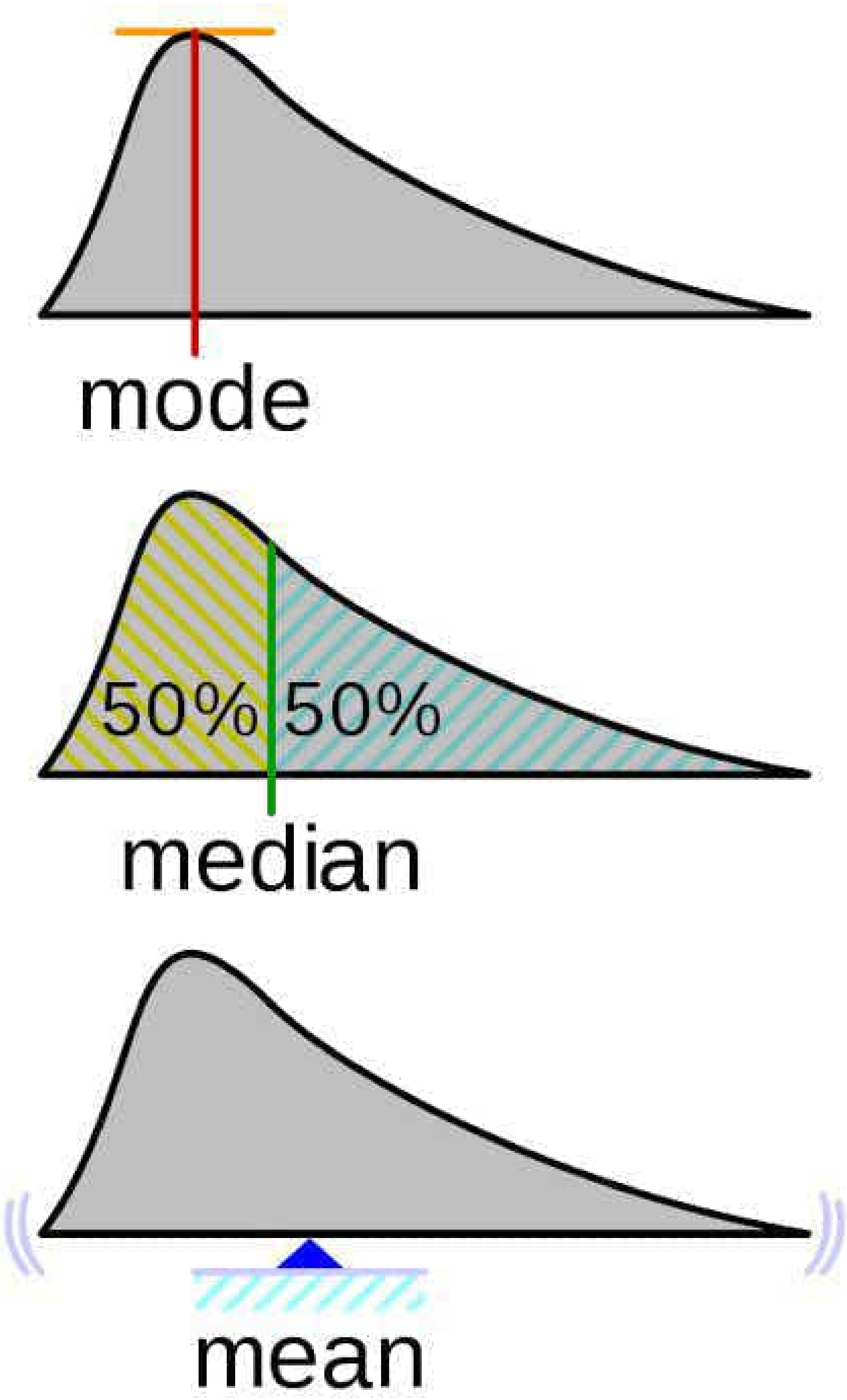
D ean = D iean

$\boxed{1 \text{ Mode} + 2 \text{ Mean} = 3 \text{ Median}}$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.



Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.


mode
50%
50%
median
mean

Variance & std Deviation of a data set

$x = \{2, 3, 4, 7, 9\}$

$$\boxed{\text{Variance}(\sigma^2) = \frac{\Sigma(x-\mu)^2}{n}}$$

$$\text{std Dev.} = \sigma = \sqrt{\text{Variance}} = \sqrt{\frac{\Sigma(x-\mu)^2}{n}}$$

Method 1:

x	μ	$x-\mu$	$(x-\mu)^2$
2	5	-3	9
3	5	-2	4
4	5	-1	1
7	5	2	4
9	5	4	16

↳ $\mu = \frac{2+3+4+7+9}{5} = 5$

↳ $\Sigma(x-\mu)^2 = 9+4+1+4+16 = 34$

$\Rightarrow$ Variance, $\sigma^2 = \frac{\Sigma(x-\mu)^2}{n} = \frac{34}{5} = 6.8$

$\Rightarrow$ std dev. $\sigma = \sqrt{6.8}$

Method 2:

$$\text{Variance} = E[x^2] - (E[x])^2 \rightarrow \text{Mean}, \mu = \frac{\Sigma x_i}{n} = \frac{\Sigma x_i f_i}{\Sigma f_i}$$

$$\boxed{\text{Variance} = \frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2}$$

$$\boxed{= \frac{\Sigma x_i^2 f_i}{\Sigma f_i} - \left(\frac{\Sigma x_i f_i}{\Sigma f_i}\right)^2}$$

x	x^2
2	4
3	9
4	16
7	49
9	81

$\Sigma x_i = 25$ $\quad \Sigma x_i^2 = 159$

$n = 5$

$\Rightarrow V = \frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2$

$V = \frac{159}{5} - \left(\frac{25}{5}\right)^2$

$\sigma^2 \rightarrow \boxed{V = 6.8}$

std dev, $\sigma = \sqrt{V} = \sqrt{6.8}$

$= \boxed{2.607}$

ques which data set has more VARIATION among these two,

(1) [1, 2, 3] (2) [101, 102, 103]

sol^n ⇒ Coefficient of Variation = $\frac{\sigma}{\mu}$

data set 1

x_i	x_i^2
1	1
2	4
3	9

$\Sigma x_i = 6$ $\Sigma x_i^2 = 14$ $n = 3$

Variance, $\sigma_1^2 = \frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2$

$= \frac{14}{3} - \left(\frac{6}{3}\right)^2 = \frac{14}{3} - 4 = \frac{2}{3}$

$\sigma_1 = \sqrt{\frac{2}{3}}$

$\mu_1 = \frac{\Sigma x_i}{n} = \frac{6}{3} = 2$

$\mu_1 = 2$

data set 2

x_i	μ	$(x_i - \mu)^2$
101	102	1
102	102	0
103	102	1

$n = 3$

Variance, $\sigma_2^2 = \frac{\Sigma(x_i - \mu)^2}{n} = \frac{2}{3}$

$\sigma_2 = \sqrt{\frac{2}{3}}$

$\mu_2 = \frac{\Sigma x_i}{n} = \frac{306}{3} = 102$

$\mu_2 = 102$

Now, coefficient of variations,

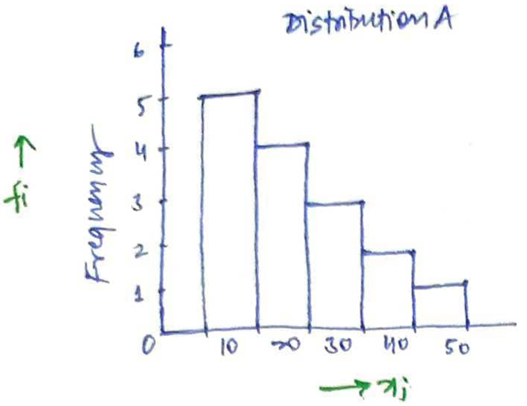
$\frac{\sigma_1}{\mu_1} = \frac{\sqrt{2/3}}{2}$ $\frac{\sigma_2}{\mu_2} = \frac{\sqrt{2/3}}{102}$

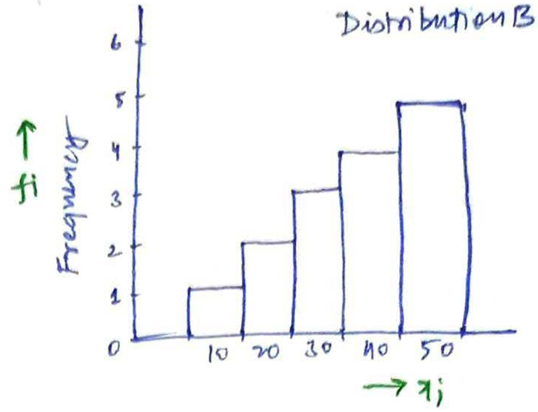
$\left(\frac{\sigma_1}{\mu_1}\right) > \left(\frac{\sigma_2}{\mu_2}\right)$

coefficient of variation of data set 1 > data set 2

⇒ Data set 1 has more variation

ques. For the distribution given below:





Which of the following is correct for the above distributions?

(a) std dev of A is significantly lower than std dev of B.

(b) std dev of A is slightly lower than std dev of B.

(c) std dev of A is same as std dev of B.

(d) std dev of A is significantly higher than std dev of B.

Sol^n

Distribution A

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
10	5	50	500
20	4	80	1600
30	3	90	2700
40	2	80	3200
50	1	50	2500
$\Sigma \Rightarrow$	15	350	10500

Distribution B

x_i	f_i	$x_i f_i$	$x_i^2 f_i$
10	1	10	100
20	2	40	800
30	3	90	2700
40	4	160	6400
50	5	250	12500
$\Sigma \Rightarrow$	15	550	22500

$$\sigma_1 = \sqrt{\frac{\Sigma x_i^2 f_i}{n} - \left(\frac{\Sigma x_i f_i}{n}\right)^2} \quad , \quad \sigma_2 = \sqrt{\frac{\Sigma x_i^2 f_i}{n} - \left(\frac{\Sigma x_i f_i}{n}\right)^2}$$

$n = \Sigma f_i = 15$ $\quad$ $n = \Sigma f_i = 15$

$$\sigma_1 = \sqrt{\frac{10500}{15} - \left(\frac{350}{15}\right)^2} = \underline{12.472} \quad , \quad \sigma_2 = \sqrt{\frac{22500}{15} - \left(\frac{550}{15}\right)^2} = \underline{12.472}$$

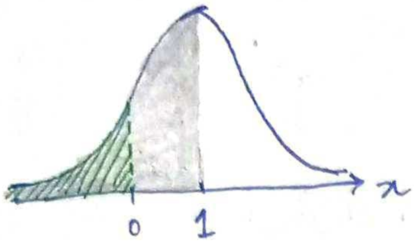
$$\Rightarrow \boxed{\sigma_1 = \sigma_2}$$

option (c) ✓

ques. let X be a normal RV with mean 1 & variance 4. The prob $P[X<0]$ is

(A) 0.5 (B) greater than zero & less than 0.5 (C) greater than 0.5 & less than 1.0 (D) 1.0

Sol^n



$\mu = 1 \Rightarrow P[X<1] = 0.5$

bcoz prob is always +ve

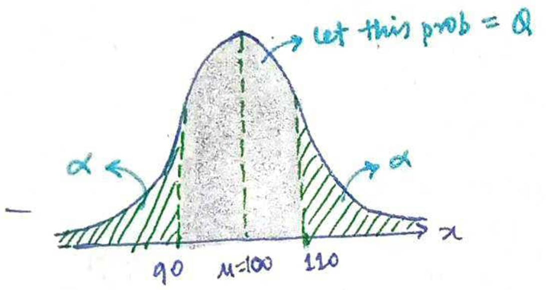
we want $P[X<0] \rightarrow 0 < P[X<0] < P[X<1]$

$$\boxed{0 < P[X<0] < 0.5}$$

ques. For a R.V X $(-\infty < x < \infty)$, following normal distribution, the mean is $\mu = 100$. If the prob is $P = \alpha$ for $x \geq 110$, then the prob of X lying between 90 & 110, i.e., $P(90 \leq x \leq 110)$ will be equal to.

(A) $1-2\alpha$ (B) $1-\alpha$ (C) $1-\frac{\alpha}{2}$ (D) 2α

Sol^n



$1 = \alpha + Q + \alpha$

$1 = 2\alpha + Q$

$$\boxed{Q = 1-2\alpha}$$

ques. let U & V be 2 independent zero mean Gaussian RV of variances $\frac{1}{4}$ & $\frac{1}{9}$ respectively. The prob $P(3V \geq 2U)$ is

(A) 4/9 (B) 1/2 (C) 2/3 (D) 5/9

Sol^n $P[3V \geq 2U] = P[3V - 2U \geq 0]$

let a normal RV, $Y = 3V - 2U$

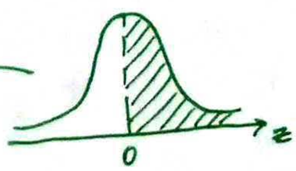
$\Rightarrow P[Y \geq 0]$

convert from Nor. dis. to Z-dis.

$= P\left[\frac{Y-\mu}{\sigma} \geq \frac{0-\mu}{\sigma}\right]$ this is mean of 'Y' not 'V' or 'U'.

$= P\left[Z \geq \frac{-\mu}{\sigma}\right]$

$= P\left[Z \geq \frac{0}{\sigma}\right] = P[Z \geq 0] = \boxed{1/2}$



$Y = 3V - 2U$

$\mu = E[Y] = E[3V - 2U]$

$= E[3V] - E[2U]$

$= 3\cancel{E[V]}^{0} - 2\cancel{E[U]}^{0}$

given V, U are zero mean RV

$\mu = E[Y] = 0$

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.

Bhanwar se lado tund lehron se uljho, kaha tak chaloge kinare-kinare.