10/1/23

Probability and Stochastic Process
(Statistics)

event
chance
occurrence

Unit 1 : Probability, Expectation, Kurtosis
(Marginal Probability)

Unit 2 : Probability distribution function → Imp.

Unit 3 : Moment Generating function and least square method

Unit 4
Unit 5 } : Statistical method

Book : Gupta & Kapoor

Unit 1

Random Experiment

Prob (outcome) = Prob (H)
or
Prob (T)

Classical Definition of Probability = $\frac{\text{favourable}}{\text{total}}$

Ω = Sample Space
= Collection of ~~total~~ all possible outcomes of underlying random experiment

Eg: random Experiment = tossing a die

Ω = collection of all possible outcomes of underlying random experiment

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

events A, B, C : Subset of Ω

$A \subseteq \Omega$

Eg : $A = \{1\}$, $B = \{2\}$

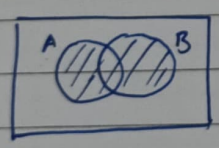
$P(A)$ $P(B)$ $P(A \cup B)$

Set theory

Set is a well-defined collection of objects

Operations on sets

① $A \cup B = \{x ; x \in A \text{ or } x \in B\}$



② $A \cap B = \{x ; x \in A \text{ and } x \in B\}$

If $A \cap B = \phi$, then A and B are known as disjoint or mutually exclusive events

null set

③ $\bar{A}$ or $A^c = \{x ; x \in S - A\}$

$\bar{A}$ or $A^c = \{x \,;\, x \in \Omega - A\}$, $\Omega - A = \Omega \cap A^c$

sample space

(4) $A - B = \{x \,;\, x \in A \;\&\; x \notin B\}$

(5) De-morgan's law

$$\overline{A \cup B} = \bar{A} \cap \bar{B}$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B}$$

Statements

(1) at least one of A or B — $A \cup B$

(2) Both the events A & B occur — $A \cap B$

(3) Neither A nor B — $\overline{A \cup B}$

(4) A occurs but not B — $A - B$

(5) Exactly one of the event A or B occur — $(A-B) \cup (B-A)$

events

HW

set

finite — countable — uncountable

11/1/23

Set

finite | countable | uncountable

$A = \{1, 2, 3\}$ $\quad A = \{x_1, x_2, x_3 \ldots\}$

$\updownarrow \quad \updownarrow$

$1 \quad 2$

Countable set : It has one-to-one correspondence with natural nos.

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

Q- Throw a coin until you get head

$\Omega = \{H, TH, TTH, TTTH, \ldots\}$

$\quad | \quad | \quad |$

$1 \quad 2 \quad 3$

↳ countable set

Q- Picking a no. from interval 0-1

$\Omega = [0, 1]$

↳ uncountable set (cannot map with natural nos.)

Q- consider a random experiment

then Ω : Universal set (A set containing all possible outcomes)

S : A class containing all subsets of Ω

members of Ω are called outcomes
members of S are called events

S : A class containing all subsets of Ω
$P \Rightarrow$ contains points from S

Events $\in S$ OR Events $\subseteq \Omega$

① # <u>classical approach</u>

$P(A)$ = is determined by a prior knowledge i.e. without doing any experiment

$$= \frac{\#(A)}{\#(\Omega)} = \frac{\text{Cardinality of set } A}{\text{Cardinality of set } \Omega}$$

$\#(\Omega) \rightarrow$ Not $S \Rightarrow$ Imp

$\#(A)$: cardinality of A
$\#(\Omega)$: Cardinality of Ω

Q- Experiment : Throwing two fair dice. Find the probability of event A.

(i) A = both dice show same no.
(ii) A = one die show even no.

(iii) A = the total of nos. is 8
(iv) A = first die shows 6

i) $1/6$

ii) $\frac{3}{6} \cdot \frac{1}{6} = \frac{1}{12}$

iii) (3,5) (5,3) (2,6) (6,2) (4,4)
$\frac{5}{36}$

iv) $1/6$

② Relative frequency (Statistical Approach)

↳ No Prior knowledge

↳ Experiment repeated several times

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

n : no. of times that we are performing the experiment

n_A : no. of times you are getting favourable result for occurence of A

~~③ Axiomatic Approach~~

Imp ③ Axiomatic Approach

Let Ω be the universal set and S be the class of subsets of Ω then we define P as probability measure ~~such that~~ such that

$$P : S \to [0, 1]$$

Such that

① $P(A) \geq 0$

② $P(\Omega) = 1$

③ Let $\{A_i, i = 1, 2, \ldots\}$ be disjoint sets then

$$Pr(A_1 \cup A_2 \ldots \cup A_n \ldots) = Pr(A_1) + Pr(A_2) + \ldots$$

example 1 : let us toss a coin

$$\Omega = \{H, T\}$$
$$S = \{\{H\}, \{T\}, \{H, T\}, \{\phi\}\}$$

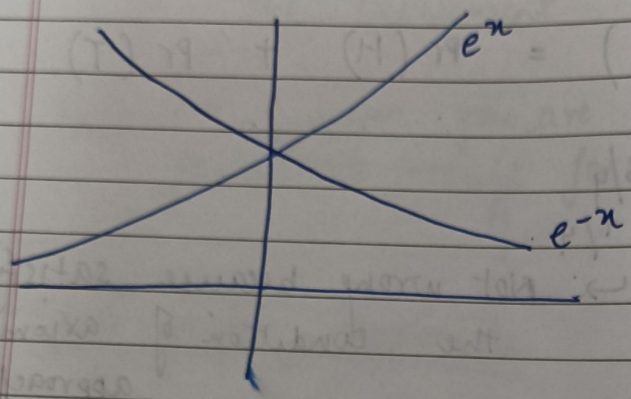
$$P(H) = 1/2$$
$$\& \quad P(T) = 1/2$$

$$Pr(\{H\} \cup \{T\}) = Pr(H) + Pr(T)$$

eg : if $P(H) = 3/4$
& $P(T) = 1/4$
↳ Not wrong because satisfying the condition of axiomatic approach.

Q- Let $\Omega = (0, \infty)$ and define

$$P(I) = \int_I e^{-x}\,dx$$, check if P is a probability measure according to its Axiomatic definition



Ans) i) $P(I) \geq 0$
as $e^{-x} \geq 0$
$\forall\, x \in (0, \infty)$

ii) $P(\Omega) = 1$

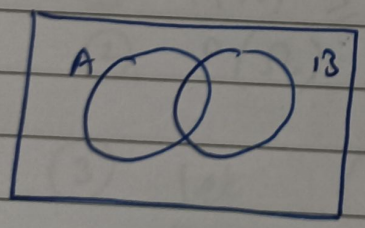
$$\int_0^\infty e^{-x}\,dx = 1$$

iii) $P(I_1 \cup I_2) = P(I_1) + P(I_2)$

$$\int_{I_1 \cup I_2} f(x)\,dx = \int_{I_1} f(x)\,dx + \int_{I_2} f(x)\,dx$$

<u>Additive Rule</u>

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



$A \cup (B - A)$

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

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

Write $A \cup B$ as union of disjoint sets

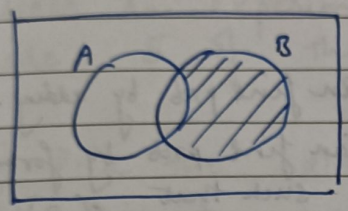
$$\cancel{A \cup (B - (A \cap B)) = A \cup B}$$

$$\cancel{P(A_1 \cup A_2) = P(A_1) + P(A_2)}$$

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

$$\cancel{[\because P(A - B) = P(A \cap B^c)]}$$

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



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

1

<u>Exhaustive Events</u> : $A_1, A_2, \ldots A_m$ are said to be exhaustive events if

$$A_1 \cup A_2 \ldots \cup A_m = \Omega$$

eg: For throwing a coin

$$A_1 = \{H\} \text{ and } A_2 = \{T\}$$

eg: For throwing a dice

$$A_1 = \{1, 3, 5\}, \quad A_2 = \{2, 4, 6\}$$

② # Mutually exclusive Events (ME)

Two events A_1, A_2 are said to be Mutually exclusive if

$$A_1 \cap A_2 = \phi$$

③ Equally likely : A_1 & A_2

Two events are said to be equally likely if $P(A_1) = P(A_2)$

Finite countable $\Omega = \{1, 2, \ldots, n\}$ ⇒ can find prob by adding simply

Infinite countable $\Omega = \{1, 2, \ldots\}$ ⇒ Can find prob by formula such that $P(\Omega) = 1$

Uncountable $\Omega = [0, 1]$

$$P([1,2]) = P(1) + P(2)$$

NOTE $\Omega = (a, b)$ & $A \subseteq \Omega$

then

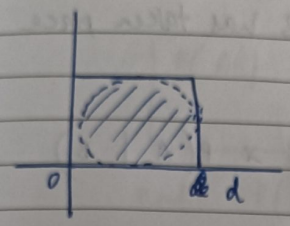
$$P(A) = \frac{\text{measure of } A}{\text{measure of } \Omega}$$

eg: If a random ~~the~~ ~~coin~~ point is chosen from an interval $[a, b]$ what is the prob that the point lies in the interval $[c, d]$ where $a \leq c \leq d \leq b$

Ans) $\dfrac{\text{length of } [c,d]}{\text{length of } [a,b]} = \dfrac{d-c}{b-a}$

Ans) $P = \frac{d-c}{b-a}$

eg:



$P = \frac{\text{Area of circle}}{\text{Area of square}} = \frac{\pi\left(\frac{a}{2}\right)^2}{a^2} = \frac{\pi}{4}$

$P(\text{shaded region}) = \frac{\pi}{4}$

Q. Consider an experiment of picking two balls from 10 balls where 5 of them are white and 5 are Red

i) Pick two balls ~~and~~ one by one with Replacement
ii) Pick two balls one by one without Replacement

A: first ball is white $\quad P(A) = 5/10$
B: second ball is white $\quad P(B) = 5/10$

i) $P(A \cap B) = \frac{5}{10} \times \frac{5}{10} = P(A) \times P(B)$

ii) $P(A \cap B) = \frac{5}{10} \times \frac{4}{9}$

$= \cancel{P(A) \times P(B|A)}$

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

↳ condition

$P(B|A)$ v/s $P(B)$

$P(B)$ when some more extra info is given (extra condition)

[Refined probability of B when A has taken place]

Conditional prob

$P(B \mid A) = P(B)$ when A has taken place

$P(A \mid B) = P(A)$ when B has taken place

$$P(A) \times P(B \mid A) = P(B) \times P(A \mid B)$$

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

$A = \{$ both coins show same face $\}$

$B = \{$ at least one coin shows H $\}$

$$P(A \mid B) = \cancel{\frac{P(A \cap B)}{P(B)}} = \frac{1}{3}$$

$A = \{HH, TT\}$

$B = \{HT, TH, HH\}$

$P(A \mid B)$ B has taken place already

$\therefore$ TT case is not there

Total outcomes $= \{HT, TH, HH\}$

same face prob $= 1/3$ (HH)

$$P(A \mid B) = \frac{1}{3}$$

$P(A \mid B) = \frac{1}{3}$ v/s $P(A) = \frac{1}{2}$

Refined probability (Given that B will happen)

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

$$\boxed{P(A \mid B) = \frac{P(A \cap B)}{P(B)}} \quad \text{where } P(B) \neq 0$$

and $$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \quad \text{where } P(A) \neq 0$$

Definition

Let $P(H) > 0$, then we define $P(A \mid H)$ as ~~Prob~~ Prob of A condition on H as

$$P(A \mid H) = \frac{P(A \cap H)}{P(H)}$$

Q- The probability of X, Y and Z becoming managers are $\frac{4}{9}$, $\frac{2}{9}$ and $\frac{1}{3}$ respectively.

The probability that the Bonus is introduced if X, Y & Z become managers are $\frac{3}{10}$, $\frac{1}{2}$ and $\frac{4}{5}$ respectively.

i) What is the prob that Bonus scheme is introduced ?

ii) If Bonus scheme is introduced, what is the probability that X has been appointed

↳ Conditional Probability

A_1 : X becomes manager
A_2 : Event when Y becomes manager
A_3 : Z becomes manager
B : Bonus is introduced.

have to write in the form of events

$P(A_1) = 4/9$
$P(A_2) = 2/9$
$P(A_3) = 1/3$

$\frac{\cancel{4}^{2}}{\cancel{9}_{3}} \times \frac{\cancel{3}}{\cancel{10}_{5}} + \frac{2}{9} \times \frac{1}{2} + \frac{1}{3} \times \frac{4}{5}$

$\frac{3}{10} + \frac{1}{2} + \frac{4}{5}$

$\frac{2}{15} + \frac{1}{9} + \frac{4}{15}$

$\frac{2}{5} + \frac{1}{9}$

$\frac{\cancel{26}^{2}}{\cancel{15}_{5}} + \frac{1}{9}$

$\frac{18 + 5}{\cancel{45}\ 9} \times \frac{\cancel{5}}{8} = \frac{23}{72}$

$\frac{3+5+8}{10} = \frac{\cancel{16}\ 8}{\cancel{10}\ 5}$

~~P(B)~~

Máam:

i) $P(B \mid A_1) = \frac{3}{10}$

$P(B \mid A_2) = \frac{1}{2}$

$P(B \mid A_3) = \frac{4}{5}$

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

↳ Mutually exclusive events

$$P(B) = P(B \mid A_1) \cdot P(A_1) + P(B \mid A_2) \cdot P(A_2) + P(B \mid A_3) \cdot P(A_3)$$

$$= \frac{\cancel{4}^{2}}{\cancel{9}_{3}} \times \frac{\cancel{3}}{\cancel{10}_{5}} + \frac{2}{9} \times \frac{1}{2} + \frac{1}{3} \times \frac{4}{5}$$

$$= \frac{2}{15} + \frac{1}{9} + \frac{4}{15} = \frac{6}{15} + \frac{1}{9} = \frac{2}{5} + \frac{1}{9}$$

$$= \frac{18+5}{45} = \frac{23}{45}$$

ii)

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

$$= \frac{P(B \mid A_1) \cdot P(A_1)}{P(B)} = \frac{\frac{\cancel{4}^{2}}{\cancel{9}} \times \frac{3}{\cancel{10}\,\cancel{2}}}{\frac{23}{\cancel{45}}}$$

$$= \frac{6}{23}$$

Independent Events

Two events A & B are said to be Independent if ~~P(A~~

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

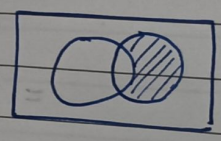
Result : If A & B are Independent then

i) A^c & B
ii) B^c & A
iii) A^c & B^c

are also independent

~~$P(A^c \cap B) = P(A) \cup P(B^c)$~~

i) $A^c \cap B = B - (A \cap B)$



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

Q- Consider families with 2 children and assume all the possible distribution of gender.
$\Omega = \{BB, BG, GB, GG\}$
Consider 2 events E and S.

E is an event that random chosen family has atmost 1 girl
F is an event that children of both gender.
Verify if E and F are independent or not.

$P(E) = 3/4$

$P(F) = 2/4 = 1/2$

$P(E \cap F) = 2/4 = 1/2$

$P(E \cap F) \neq P(E).P(F)$ $\therefore$ Not Independent

NOTE

A collection $\{A_1, A_2, \ldots\}$ is said to be mutually independent if any subcollection is Independent, i.e.

$$P(A_1 \cap A_2 \ldots \cap A_i) = P(A_1).P(A_2)\ldots P(A_i) \quad \forall i$$

If collection is pair-wise independent, it may or may not be mutually independent.

But if collection is mutually independent, then it is also ~~mutually indep~~ pair wise independent.

Mutually exclusive $\Rightarrow$ Pair wise independent
$\not\Leftarrow$

All combinations of i considered in mutually independent.

$$P(A_{i1} \cap A_{i2} \ldots \cap A_{in}) = P(A_{i1}).P(A_{i2})\ldots P(A_{in})$$

Take 4 marbles named as ~~A~~

$A_1\ A_2\ A_3$

A_1

A_2

A_3

Put the 4 marbles in the box and draw 1 at random.

Let E_i : denote an event that symbol A_i occurs on the marble

$i = 1, 2, 3$

Check whether $\{E_1, E_2, E_3\}$ is mutually independent or pairwise independent

$$P(E_1) = 2/4 = 1/2$$
$$P(E_2) = 2/4 = 1/2$$
$$P(E_3) = 2/4 = 1/2$$
$$P(E_1 \cap E_2 \cap E_3) = \frac{1}{4}$$

$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ $\therefore$ Pair wise independent

But $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \neq \frac{1}{4}$ $\therefore$ Not mutually independent

$$P(E_1) \cdot P(E_2) = P(E_1 \cap E_2)$$

$$P(E_1) \cdot P(E_2) \cdot P(E_3) \neq P(E_1 \cap E_2 \cap E_3)$$

Random variable

eg: Consider a Random experiment of throwing three coins

Ω		no. of heads	
$w_1 =$	HHH	3	$X(w_1) = 3$
$w_2 =$	HHT	2	$X(w_2) = 2$
$w_3 =$	HTT	1	$X(w_3) = 1$
$w_4 =$	HTH	2	
$w_5 =$	TTT	0	
$w_6 =$	THH	2	
$w_7 =$	THT	1	
$w_8 =$	TTH	1	

X : no. of heads
$0, 1, 2, 3$

$\{w : X(w) = 1\}$ For what w's, $X(w) = 1$

$= \{w_3, w_7, w_8\}$ collection of w's for which X maps to 1

$= \{HTT, THT, TTH\}$ $\Rightarrow$ Event

Prob of random variable $\equiv$ Prob of taking some event

~~Can~~
can define random variable for random experiment and can define more than 1 events for random variable.

→ Random variable is always a function that goes from Ω to R

$$X : \Omega \longrightarrow R$$
$$\Omega = \{w_1, w_2, \ldots w_8\}$$
$$X(w_1) = 3$$

→ Random variable is a function that takes input from Ω and maps on real line.

Q- Consider an experiment of rolling two dice defined for different random variables
(A, B, C, D)

Random variable

$$X : \Omega \rightarrow R$$

tossing two dice

X : Sum

$= 2, 3, \ldots\ldots 12$

$Pr\{w \in \Omega : X(w) = 2\} = Pr\{(1,1)\}$

$Pr\{w \in \Omega : X(w) = 3\} = Pr\{(1,2),(2,1)\}$

$Pr(w : 1 \leq X(w) \leq 3\}$

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

Random variable X — Discrete, Continuous, Mixed

Discrete Random variable

If it assumes only countable number of values

Continuous Random Variable

If it can assume any value within given range.

Q- (1) no. of ~~customers~~ arrivals — Discrete

2) Arrival time

3) no. of errors

4) Stock price

5) No. of deaths due to

No. of arrivals ⇒ Random variable ∵ there is a chance/prob that this many arrivals are there
↳ Discrete

Arrival time ⇒ Continuous
No. of errors ⇒ Discrete
Stock price ⇒ Continuous
No. of deaths due to ⇒ Discrete
Height / weight ⇒ Continuous

Probability Mass function (PMF)

Definition : If X is a discrete random variable, assumes countable numbers of values i.e. $x_1, x_2, x_3, \ldots$ then the function $P(x)$ ~~such~~ ~~that~~ defined as

$$P_X(x) = \begin{cases} P_i = P(X = x_i) & i = 1, 2 \\ 0 & \text{otherwise} \end{cases}$$

when $X = x_1$ prob $= P_1$
when $X = x_2$ prob $= P_2$

is called Probability Mass function if

(i) $\sum_{i=1}^{\infty} P_i = 1$ → Infinite series

if the series converges to 1, then the collection of P_i's is called PMF

ii) $P_i \geq 0$

For continuous random variable ⇒ Probability Density function

X : no. of heads (Discrete Random variable)

eg: Ω $\quad X = 0 = x_1,\ 1 = x_2,\ 2 = x_3,\ 3 = x_4$

HHH
HHT
HTH
HTT
THH
THT
TTH
~~THH~~ TTT

$X = 0 = x_1,\ 1 = x_2,\ 2 = x_3,\ 3 = x_4$

$P_1 = Pr(X=0) = \frac{1}{8}$

$P_2 = Pr(X=1) = \frac{3}{8}$

$P_3 = Pr(X=2) = \frac{3}{8}$

$P_4 = Pr(X=3) = \frac{1}{8}$

$P_X(x) = \left\{ P_1 = \frac{1}{8},\ P_2 = \frac{3}{8},\ P_3 = \frac{3}{8},\ P_4 = \frac{1}{8} \right.$

PMF $\because P_1 + P_2 + P_3 + P_4 = 1$

and $P_1 \geq 0,\ P_2 \geq 0,\ P_3 \geq 0,\ P_4 \geq 0$

Q1) A random variable has the following PMF

$X =$	0	1	2	3	4	5	6	7
$P(X) =$	0	k	$2k$	$3k$				

$X =$	0	1	2	3	4	5	6	7
$P(X) =$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

i) Find k

ii) $Pr(X < 6)$, $Pr(X \geq 6)$, $Pr(0 < X < 5)$

iii) If $P(X \leq a) > \frac{1}{2}$, find minimum value of a

i) $0 + 9k + 10k^2 = 1$

$10k^2 + 9k - 1 = 0$

$10k^2 + 10k - k - 1 = 0$

$10k(k+1) - 1(k+1) = 0$ $\quad k = -1, \ 1/10$

$k = \frac{1}{10}$

ii) $Pr(X < 6) = 0 + k + 2k + 2k + 3k + k^2$

$= k^2 + 8k$

$= \frac{1}{100} + \frac{8}{10} = 0.01 + 0.8 = 0.81$

$Pr(X \geq 6) = 2k^2 + 7k^2 + k = 9k^2 + k$

$= 0.09 + 0.1 = 0.19$

$Pr(0 < X < 5) = k + 2k + 2k + 3k$

$= 8k = 0.8$

iii) $P(X \leq a) > \frac{1}{2}$

$a = 4$

Q2) If $P(x) = \begin{cases} \frac{x}{15}, & x = 1, 2, 3, 4, 5 \\ 0 & \text{otherwise} \end{cases}$

Find i) $Pr(X = 1 \text{ or } 2)$

ii) $Pr\left(\frac{1}{2} < X < \frac{1}{5} \,\middle|\, X > 1\right)$

Eg 3 X = Sum of two dice

i) Construct its PMF

ii) Draw the charts of PMF

(x-axis : sum
y-axis : bar graph)

Probability Density function

Let X be a continuous Random Variable then a non-negative function $f_X(x)$ is called its Probability density function if

$$Pr(x_1 \le X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$$

such that

i) $$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

ii) $$f_X(x) \geq 0$$

$Pr(X = x_1) = 0$ ⇒ Mass at a given point is zero in continuous case

↳ ∵ then $\int_{x_1}^{x_1}$

Only for continuous RV

$$Pr(X < x_1) = Pr(X \le x_1)$$

as $Pr(X = x_1) = 0$

NOTE:

$$\boxed{\begin{array}{l} Pr(X < x_1) = Pr(X \le x_1) \text{ as } Pr(X = x_1) = 0 \text{ (for continuous RV)} \\ Pr(X < x_1) \ne Pr(X \le x_1) \text{ for discrete RV} \end{array}}$$

eg: let $f_X(x) = 6x(1-x) \quad 0 \le x \le 1$

i) Check if $f_X(x)$ is a PDF

ii) Determine a such that $Pr(X < a) = Pr(X > a)$

i) $\int_0^1 6x(1-x)\,dx = 6\int_0^1 (x - x^2)\,dx = 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1$

$= 6\left(\frac{1}{2} - \frac{1}{3}\right) = 6\left(\frac{3-2}{6}\right)$

$= 1$ Hence PDF

ii) $\int_0^a 6x(1-x)\,dx = 6\int x - x^2\,dx$

$= 6\left[\frac{x^2}{2} - \frac{x^3}{3}\right]$

$Pr(X < a) = Pr(X > a)$

$\int_0^a \cancel{6}x(1-x)\,dx = \int_a^1 \cancel{6}x(1-x)\,dx$

$\left(\frac{a^2}{2} - \frac{a^3}{3}\right) - 0 = \left(\frac{1}{2} - \frac{1}{3}\right) - \left(\frac{a^2}{2} - \frac{a^3}{3}\right)$

$2\left(\frac{a^2}{2} - \frac{a^3}{3}\right) = \frac{1}{6} \Rightarrow a^2\left(\frac{1}{2} - \frac{a}{3}\right) = \frac{1}{12}$

$2 \times \frac{(3a^2 - 2a^3)}{\cancel{6}} = \frac{1}{\cancel{6}}$

$6a^2 - 4a^3 = 1$

$4a^3 - 6a^2 + 1 = 0$

Eg: Let X be a continuous Random variable with Pdf

$$f_X(x) = \begin{cases} ax & 0 \le x \le 1 \\ a & 1 \le x \le 2 \\ -ax + 3a & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

i) find a

ii) compute $P(X \le 1.5)$ $\int_0^{1.5}$

20/1/23

Random variable X

Discrete — PMF — Σ — function which generates prob for discrete R.V.

continuous — PDF — $\int$ — function which generates prob for continuous R.V.

$\Rightarrow$ $X : 0, 1, 2$

$P(X = x) : \frac{1}{3}, \frac{1}{3}, \frac{1}{3}$

$Pr(X = 1 \text{ or } 2) = \frac{1}{3} + \frac{1}{3}$

$\Rightarrow$ $X \in [0, 1]$

$f_X(x) = e^{-x}$

$$Pr(0.5 \le X \le 0.7) = \int_{0.5}^{0.7} e^{-x}\, dx$$

PTO
→

Distribution function
(Cumulative Distribution function (CDF))

$$F_X(x) = Pr(w : X(w) \le x)$$

$$Pr(X \le 2)$$

If X is discrete

$$\boxed{\sum_{i=0}^{x} Pr(X = i)}$$

If X is continuous

$$\boxed{\int_{-\infty}^{x} f_X(x)\, dx}$$

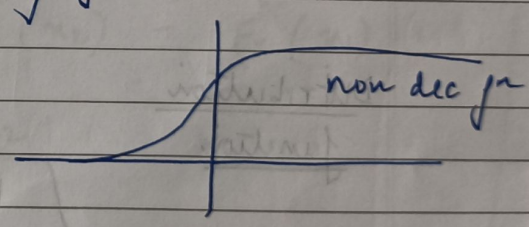
Note :

i) $F_X(-\infty) = 0$, $F_X(\infty) = 1$

ii) F is a non-decreasing function

For $x_1 \le x_2$

$F_X(x_1) \le F_X(x_2)$



iii) F is right continuous

$$\lim_{\varepsilon \to 0} F_X(x + \varepsilon) = F_X(x)$$

Q- Consider an experiment of throwing 3 coins. Consider a random variable X = no. of consecutive heads. Construct its Prob mass function and draw its distribution function.

$X = 0, 1, 2$

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

$P(X=0) = \{TTT, HTT, THT, TTH, HTH\} = 5/8$

$P(X=1) = \{HHT, THH\} = 2/8$

$P(X=2) = \{HHH\} = 1/8$

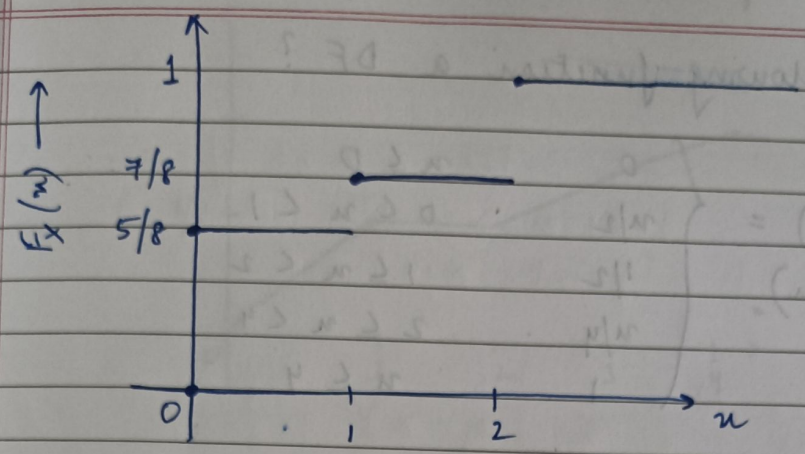
→ For discrete R.V., distribution function is a step function.

<u>PMF</u>

X	$Pr(X = x)$
0	5/8
1	2/8
2	1/8

<u>Distribution function</u>

$$F_X(x) = \begin{cases} Pr(X \le 0) & x = 0 \\ = Pr(X=0) = 5/8 & \\ Pr(X \le 1) & x = 1 \\ = Pr(X=0) + Pr(X=1) = 7/8 & \\ Pr(X \le 2) & x = 2 \\ = Pr(X=0) + Pr(X=1) + Pr(X=2) & \\ = 1 & \end{cases}$$



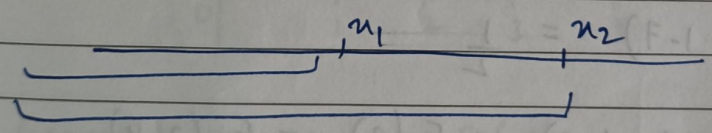
Q- Is the following function a distribution function?

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

NOTE:

① $Pr(x_1 \le X \le x_2) = Pr(X \le x_2) - Pr(X \le x_1)$

$= F_X(x_2) - F_X(x_1)$



② $Pr(X \le x) = 1 - Pr(X > x)$

OR

$Pr(X > x) = 1 - Pr(X \le x)$

$Pr(X > x) = 1 - F_X(x)$

PDF → integrate to find Prob
CDF → do not integrate, use formula to find prob (previous page NOTE)

Q- Is the following function a DF ?

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ x/2 & 0 \le x < 1 \\ 1/2 & 1 \le x < 2 \\ x/4 & 2 \le x < 4 \\ 1 & x \le 4 \end{cases}$$

$Pr(X \le x)$

Find out :

i) $Pr(X \le 0.7)$
ii) $Pr(X \le 1.7)$
iii) $Pr(2.4 \le X \le 3)$
iv) $Pr(2 \le X \le 4)$
v) $Pr(X \le 6)$

$$\boxed{\underset{\text{CDF}}{F_X(x)} = \int_{-\infty}^{x} \underset{\text{PDF}}{f_X(x)}\, dx}$$

Ans)

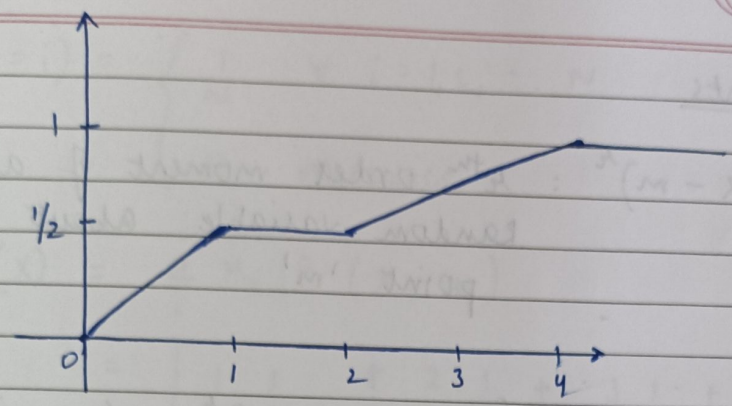
i) $Pr(X \le 0.7) = 0.7/2$

ii) $Pr(X \le 1.7) = \frac{1}{2}$

iii) $Pr(2.4 \le X \le 3) = F_X(3) - F_X(2.4)$
$= \frac{3}{4} - \frac{2.4}{4}$

iv) $Pr(2 \le X \le 4) = F_X(4) - F_X(2)$
$= 1 - \frac{2}{4}$

v) $Pr(X \le 6) = 1$

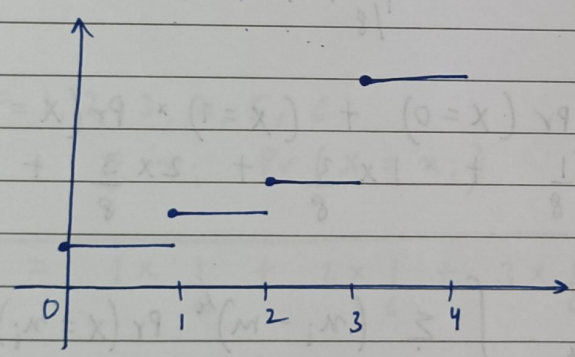


Expected values / variance / moments

X : marks of CSE 3 of Prob MSE

= 0, 1, ... 30

Freq 0 2 ..... 5 Total = 100 students



Expected values → E(X) ; central value

Variance : Error around mean value

23/1/23

Random variable

→ PMF / PDF

→ CDF $Pr(X \leq x)$

Moments

$E(X-m)^r$: r^{th} order moment of a random variable about a point 'm'

↓ expected value

$$E(X-m)^r = \begin{cases} \sum_i (x_i - m)^r \, Pr(X = x_i) & \text{if } X \text{ is a discrete R.V.} \\ \int (x-m)^r f_X(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

X	$Pr(X = x_i)$
0	1/8
1	3/8
2	3/8
3	1/8

$$(X=0) \times Pr(X=0) + (X=1) \times Pr(X=1)$$
$$= 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8}$$

$$\boxed{E(X-m)^r = \begin{cases} \sum_i (x_i - m)^r \, Pr(X = x_i) & \text{if } X \text{ is discrete} \\ \int (x-m)^r f_X(x)\, dx & \text{if } X \text{ is continuous} \end{cases}}$$

For $r=1$ and $m=0$

$$E(X) = \begin{cases} \sum_i x_i \, Pr(X = x_i) & X \text{ is discrete} \\ \int_{-\infty}^{\infty} x \, f_X(x) & X \text{ is continuous} \end{cases}$$

Mean value or expected value of X

Q- $Pr(X=i) = \begin{cases} \frac{1}{N} & \forall\ i = 1, 2, \ldots N \\ 0 & \text{otherwise} \end{cases}$. Find $E(X)$

Ans)

$$E(X) = \sum_i x_i \, Pr(X = x_i)$$

$$= \left[1 \cdot \frac{1}{N} + 2 \cdot \frac{1}{N} + 3 \frac{1}{N} + \ldots N \cdot \frac{1}{N} \right]$$

$$= \frac{1}{N} \left[1 + 2 + 3 + \ldots + N \right]$$

$$= \frac{1}{\cancel{N}} \cdot \frac{\cancel{N}(N+1)}{2} = \frac{N+1}{2}$$

Q-

a) Find the expectation of a number on a die when thrown

Ans) $X = 1, 2, 3, 4, 5, 6$

$$E(X) = \sum_i x_i \, Pr(X = x_i)$$

$$= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}$$

$$= \frac{1}{6} \left[1 + 2 + 3 + 4 + 5 + 6 \right] = \frac{21}{6}$$

$$= \frac{7}{2}$$

b) Two unbiased dice are thrown. Find the expected value of the sum of no. of points on them

PTO $\rightarrow$

$X = 2, 3, 4, 5, \ldots\ldots, 11, 12$

$Pr(X=2) = Pr(\{1,1\}) = \frac{1}{36}$

$Pr(X=3) = Pr\{(1,2),(2,1)\} = \frac{2}{36}$

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

$Pr(X=5) = Pr\{(2,3),(3,2),(1,4),(4,1)\} = \frac{4}{36}$

$Pr(X=6) = Pr\{(1,5),(5,1),(4,2),(2,4),(3,3)\} = \frac{5}{36}$

$Pr(X=7) = Pr\{(1,6),(6,1),(2,5),(5,2),(3,4),(4,3)\}$
$= \frac{6}{36}$

$Pr(X=8) =$
$Pr(X=9) =$
$Pr(X=10) =$
$Pr(X=11) =$
$Pr(X=12) =$

Q- $f_X(x) = \begin{cases} \frac{2}{x^3} ; & x \geq 1 \\ 0 & x < 0 \end{cases}$ Find $E(X)$

Ans)

$$E(X) = \int_1^\infty x \cdot \frac{2}{x^3}\, dx = 2\int_1^\infty \frac{1}{x^2}\, dx$$

$$= -2\left[\frac{1}{x}\right]_1^\infty$$

$$= 2$$

Q- Let $X(w) = I_A(w)$ → Indicator function Find $E(I_A(w))$

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

Ans) $E(I_A(w)) = 1 \times Pr(w \in A) + 0 \times \cancel{Pr(w \notin A)}$ $Pr(w \in \bar{A})$

$= Pr(w \in A)$

NOTE :

$E(g(X))$; where $g(X)$ is a function of X such that $g(X)$ is also a Random variable

$$E(g(X)) = \begin{cases} \sum_i g(x_i) \, Pr(X = x_i) & \text{when } X \text{ is a discrete RV} \\ \int g(x) \, f_X(x) \, dx & \text{when } X \text{ is continuous} \end{cases}$$

Q-

X	-3	6	9
$P(X)$	$1/6$	$1/2$	$1/3$

Find $E(X)$, $E(X^2)$ and $E(2X+1)^2$

Ans)

$E(X) = -3 \times \frac{1}{6} + 6 \cdot \frac{1}{2} + 9 \cdot \frac{1}{3} = -0.5 + 3 + 3 = 5.5$

$E(X^2) = (-3)^2 \times \frac{1}{6} + (6)^2 \times \frac{1}{2} + (9)^2 \cdot \frac{1}{3} = 1.5 + 18 + 27$
$= 46.5$

$E(2X+1)^2 = (-5)^2 \times \frac{1}{6} + (13)^2 \times \frac{1}{2} + (19)^2 \times \frac{1}{3}$

NOTE:

$$\boxed{\begin{array}{l} \text{i) } E(aX) = a\,E(X) \text{ where } a \text{ is a constant} \\ \text{ii) } E(X+b) = E(X) + b \text{ where } b \text{ is a constant} \end{array}}$$

Variance

$$\text{Variance} = E(X-\mu)^2 = \begin{cases} \sum_i (x_i - \mu)^2 \, Pr(X = x_i) & \text{if } X \text{ is discrete} \\ \int (x-\mu)^2 f_X(x)\, dx & \text{if } X \text{ is continuous} \end{cases}$$

2nd order moment about mean value

$\mu = E(X)$

$$\text{Standard deviation} = \sqrt{\text{variance}}$$

Variance = Average error around the mean value

Standard deviation = $\sqrt{\text{variance}}$

24/1/23

Q1) Suppose the life of a certain kind of radio has following PDF

$$f_X(x) = \begin{cases} \frac{100}{x^2} & \text{if } x \geq 100 \\ 0 & x < 100 \text{ otherwise} \end{cases}$$

i) Find CDF of x

ii) What is the prob that none of three such

tubes will have to be replaced during 150 hrs?

Ans) i) $F_X(x) = \begin{cases} 0 & , x < 100 \\ 1 - \frac{100}{x} & , 100 \leq x \end{cases}$

$$\int_{100}^{x} \frac{100}{x^2}\, dx = -100 \left[\frac{1}{x}\right]_{100}^{x} = -100\left[\frac{1}{x} - \frac{1}{100}\right]$$

$$= 1 - \frac{100}{x}$$

ii) Ans) $\left(\frac{1}{3}\right)^3$

$Pr(X \leq 150) = 1 - \frac{100}{150} = 1 - \frac{2}{3} = \frac{1}{3}$ → Pr(tubes will get replaced before 150 hrs)

$P(3 \text{ tubes}) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{8}{27}$

Q2) $F_X(x) = \begin{cases} 1 - (1+x)e^{-x} & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}$

i) Find the PDF for X
ii) $Pr(x \leq 1.5)$
iii) $Pr(X = 0)$
iv) Find mean and variance of X

Ans)
i) $\frac{d}{dx}\left(1 - (1+x)e^{-x}\right) = -\left[e^{-x} - e^{-x}(1+x)\right]$

$$= -\left[e^{-x} - e^{-x} - xe^{-x}\right]$$

$$= xe^{-x}$$

$$f_X(x) = \begin{cases} x e^{-x} & ; x \geq 0 \\ 0 & ; x < 0 \end{cases}$$

ii) $Pr(x \leq 1.5) = 1 - (1 + 1.5)\, e^{-1.5}$
$= 1 - (2.5)\, e^{-1.5}$

iii) $Pr(x = 0) = 0 . e^{-0} = 0$

Q3)

$$F_X(x) = \begin{cases} 0 & ; x \leq 1 \\ k(x-1)^4 & ; 1 < x \leq 3 \\ 1 & ; x > 3 \end{cases}$$

i) Find k

ii) Find mean & variance of X.

Ans 2) iv) $Var(X) = E(X - \mu)^2 \quad ; \mu = E(X)$

$= E(X^2 + \mu^2 - 2\mu X)$

($\because E(X+Y) = E(X) + E(Y)$
$E(a) = a$
↳ constant)

$= E(X^2) + E(\mu^2) - E(2\mu X)$

$= E(X^2) + \mu^2 - 2\mu E(X)$

$= E(X^2) + \mu^2 - 2\mu . \mu$

$= E(X^2) - \mu^2$

NOTE $\boxed{Var(X) = E(X^2) - (E(X))^2}$

So, $E(X) = \int_0^\infty x \cdot xe^{-x}\, dx$ — ①

$$E(X^2) = \int_0^\infty x^2 \cdot xe^{-x}\, dx$$

$$E(X^2) = \int_0^\infty x^3 e^{-x}\, dx \quad — ②$$

$$Var(X) = eq② - (eq①)^2$$

Ans 3)

$$f_X(x) = \begin{cases} 0 & ; \quad x \le 1 \\ 4k(x-1)^3 & ; \quad 1 < x \le 3 \\ 0 & ; \quad x > 3 \end{cases}$$

i) $\int_1^3 4k(x-1)^3 = 1$

$$\Rightarrow \cancel{4}k \left[\frac{(x-1)^4}{\cancel{4}}\right]_1^3 = 1 \quad \Rightarrow k(2)^4 = 1$$

$$\Rightarrow k = 1/16$$

Ma'am :

i) Directly put $x = 3$ in $k(x-1)^4 = 1$

So ans $\Rightarrow k(3-1)^4 = 1$

$k = 1/16$

ii) $E(X) = \int_1^3 x \cdot 4k(x-1)^3\, dx$

$$= \int_1^3 x \frac{4}{16}(x-1)^3\, dx = \int_1^3 \frac{x(x-1)^3}{4}\, dx$$

$$= \frac{1}{4}\int_1^3 (x^3 - 1 - 3x^2 + 3x)\, x\, dx$$

$$= \frac{1}{4} \int_1^3 (x^4 - x - 3x^3 + 3x^2)\, dx$$

$$= \frac{1}{4} \left[\frac{x^5}{5} - \frac{x^2}{2} - \frac{3x^4}{4} + \frac{\cancel{3}x^3}{\cancel{3}} \right]_1^3$$

$$= \frac{1}{4} \left[\frac{243}{5} - \frac{9}{2} - \frac{243}{4} + 27 - \frac{1}{5} + \frac{1}{2} + \frac{3}{4} - 1 \right]$$

$$= \frac{1}{4} \left[\cancel{243} \; \frac{242}{5} - \frac{240}{4} - \frac{8}{2} + 26 \right]$$

$$= \frac{1}{4} \left[\frac{242}{5} - 38 \right]$$

$$= \frac{1}{4} \left[\frac{242 - 190}{5} \right] = \frac{1}{\cancel{4}\,2} \left(\frac{\overset{26}{\cancel{52}}}{5} \right) = 2.6$$

Q4) The length of time that a certain person speaks on the telephone follows a given pdf

$$f_X(x) = \begin{cases} A e^{-x/5} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

i) Find A

ii) $E(X)$, $var(X)$

iii) What is prob that a person speaks more than 10 mins

iv) Less than 5 mins

v) between 5 mins and 10 mins

PTO
→

Ans) i) $\int_0^\infty A e^{-x/5} dx = 1$

$\Rightarrow \frac{A}{-1/5} \left[e^{-x/5}\right]_0^\infty = 1$

$\Rightarrow \frac{A}{-1/5} \left[\frac{1}{e^{\infty/5}} - 1\right] = 1$

$A = \frac{1}{5}$

ii) $E(X) = \int_0^\infty \frac{1}{5} e^{-x/5} \cdot x \, dx$

25/1/23

~~$\mu_r' = E(X-A)^r$~~

$\mu_r' = E(X-A)^r =$ r^{th} order moment about A when $A = \mu$

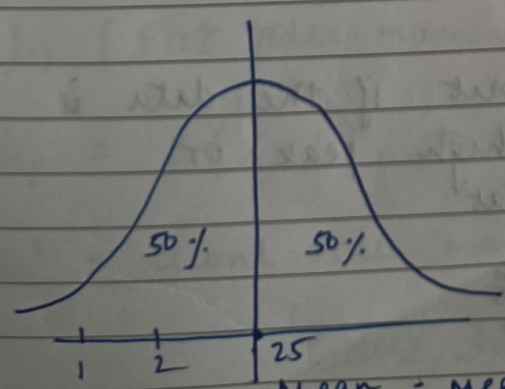
$\mu_r = E(X-\mu)^r =$ r^{th} central moment

Second order moment

$\mu_2 = \mu_2' - (\mu_1')^2$

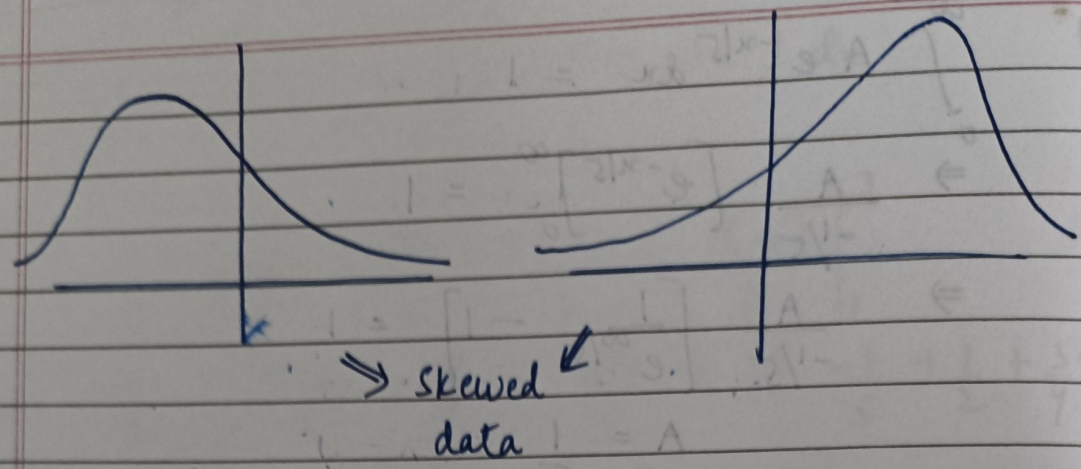
$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$

$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$



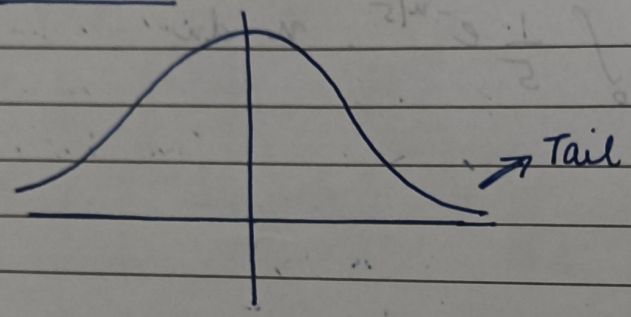
→ symmetric

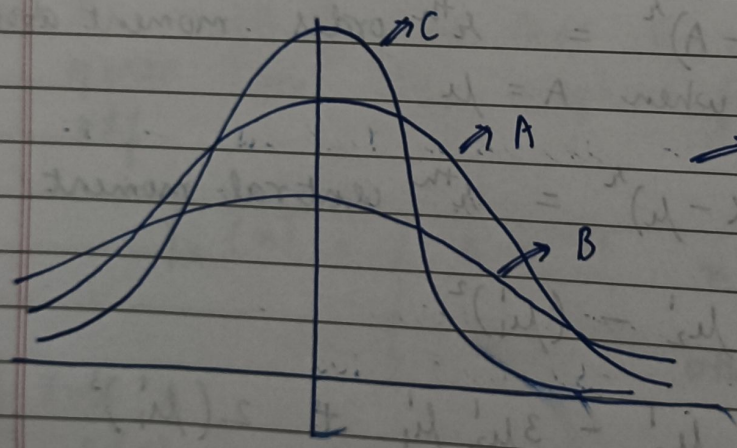
Mean = Median
= Mode



→ skewed data

Kurtosis





→ All three symmetric

By Kurtosis, can find out how the distribution is

Kurtosis : It tells about if the data is having high peak or more flat

PTO →

C = Leptokurtic curve
A = Normal curve
B = Platy Kurtic curve

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{E(X-\mu)^4}{[E(X-\mu)^2]^2} \quad ; \quad \gamma_2 = \beta_2 - 3$$

If $\gamma_2 = 0 \Rightarrow A \quad (\beta_2 = 3)$ Normal curve
$\gamma_2 > 0 \Rightarrow C \quad (\beta_2 > 3)$ leptokurtic curve
$\gamma_2 < 0 \Rightarrow B \quad (\beta_2 < 3)$ Platy Kurtic curve

$$\boxed{\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{E(X-\mu)^4}{[E(X-\mu)^2]^2}}$$

↳ can apply Kurtosis on any graph [Not just symmetric]

Q- The first four moments about a point 4 are -1.5, 17, -30 and 108

i) Find the first four moments about mean
ii) Find kurtosis

μ_1 (First order moment about mean)

$\mu_1 = E(X-\mu) = E(\mu-\mu) = 0$ [By default]

μ_2' = second order moment about any point A

↳ Here, taking moment about point 4

$$\mu_2 = \mu_2' - (\mu_1')^2$$
$$= 17 - (-1.5)^2$$
$$= 14.75$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$
$$= -30 - 3(17)(-1.5) + 2(-1.5)^3$$

Joint Distribution

$Pr(X \in A) =$ ~~~

$Pr(Y \in B) =$ ~~~

height

$Pr(h_1 \leq X \leq h_2)$

$Pr(w_1 \leq Y \leq w_2)$

$Pr(A \cap B) = Pr(X \in A \text{ and } Y \in B)$

↳ Joint Probability

Let us consider n random variables

$X_1, X_2, \ldots\ldots, X_N$

$Pr(a_1 \leq X_1 \leq a_2,\ b_1 \leq X_2 \leq b_2, \ldots$
$\ldots\ z_1 \leq X_N \leq z_2)$

= Joint probability

PTO
→

→ Let (X, Y) be collection of two discrete Random variables then their joint PMF is defined as

$$Pr(X = x_i, Y = y_i) = P_{ij} \quad \text{such that } P_{ij} \geq 0 \ \forall i, j$$

$$\sum_i \sum_j P_{ij} = 1$$

If (X, Y) be a collection of its Random variable then their joint pdf is

$$f_{XY}(x, y) \text{ such that } f_{XY}(x, y) \geq 0$$

and $$\int_x \int_y f_{XY}(x, y)\, dx\, dy = 1$$

Q-

Y \ X	1	2	3
4	1/8	1/24	1/12
6	1/4	1/4	0
8	1/8	1/24	1/12

→ Joint prob such that $X = 3, Y = 4$

↳ Joint Prob Mass function given (All Pr ≥ 0 and sum of Pr $= 1$)

i) $Pr(X \leq 2, Y \leq 6)$

$$= \frac{1}{8} + \frac{1}{24} + \frac{1}{4} + \frac{1}{4} = \frac{3+1+6+6}{24} = \frac{16}{24} = \frac{2}{3}$$

ii) $Pr(X = 1, Y \leq 8)$

$$= \frac{1}{8} + \frac{1}{4} + \frac{1}{8} = \frac{1+2+1}{8} = \frac{1}{2}$$

Either both X & Y are discrete or both X & Y are continuous

(X, Y) [X and Y both are discrete]

When X and Y both are discrete, its Joint Prob function is given in tabular form

$Y \backslash X$	x_1	x_2	.....	x_n
y_1	p_{11}	p_{12}		
y_2				
⋮				
y_m				p_{mn}

$$\sum_i \sum_j p_{ij} = 1$$

Joint PMF $\boxed{\sum_i \sum_j p_{ij} = 1}$

(X, Y) [X and Y both are continuous]

$$\boxed{\int_x \int_y f_{XY}(x, y) = 1}$$

Marginal PDF / PMF

Joint prob function (table) given, then

$\boxed{\text{PMF for } X = \text{Marginal PMF of } X}$

derived from Joint PMF

$X =$	x_1	x_2	.....	x_n
$Pr(X = x_i)$				

$$Pr(X = x_1) = (x_1 \cap y_1) \cup (x_1 \cap y_2) \cup \dots (x_1 \cap y_m)$$

X	x_1	x_2	$\dots$	x_n
y_1	P_{11}	P_{12}		
y_2				
$\vdots$				
y_m				P_{mn}

Prob that $X = x_1$

Each Row $\Rightarrow$ PMF for Y
Each Column $\Rightarrow$ PMF for X

$$Pr(X = x_1) = (x_1 \cap y_1) \cup (x_1 \cap y_2) \cup \dots (x_1 \cap y_m)$$

$$\boxed{Pr(X = x_j) = \sum_{i=1}^{m} P_{ij} \quad \forall j}$$

$X =$	x_1	x_2	$\dots$	x_n
$Pr(X = x_j)$	$\sum_{i=1}^{m} P_{i1}$	$\sum_{i=1}^{m} P_{i2}$	$\dots$	$\sum_{i=1}^{m} P_{in}$

$\hookrightarrow$ Marginal PMF of X

Y	y_1	y_2	$\dots$	y_m
$Pr(Y = y_i)$	$\sum_{j=1}^{n} P_{ij}$			

$\hookrightarrow$ Marginal PMF of Y

$Y \backslash X$	x_1	x_2	x_n	
y_1	P_{11}	P_{12}	P_{1n}	$\rightarrow \sum_{j=1}^{n} P_{ij}$
y_2	P_{21}			
$\vdots$				
y_m	P_{m1}		P_{mn}	$\sum_{j=1}^{m} P_{mj}$
Marginal PMF for X $\rightarrow$	$\sum_{i=1}^{m} P_{i1}$	$\sum_{i} P_{i2}$	$\sum_{i} P_{in}$	1

For Continuous :

IMP questions

Let $f_{XY}(x, y)$ be a joint PDF for (X, Y)

where $-\infty < x < \infty$ and $-\infty < y < \infty$

then the

$$\text{Marginal PDF for } X = \int_{y=-\infty}^{\infty} f_{XY}(x, y)\, dy$$

$$= f_X(x) \qquad -\infty < x < \infty$$

IMP (function of x)

(y eliminated because limits lag gaye)

$$\text{Marginal PDF for } Y = \int_{x=-\infty}^{\infty} f_{XY}(x, y)\, dx$$

$$= f_Y(y) \qquad -\infty < y < \infty$$

Q-

$y \backslash x$	0	1	$P(y)$
0	0.28	0.37	0.65
1	0.22	0.13	0.35
$P(x)$	0.5	0.5	1

Ans)

$X =$	0,	1
$P(x)$	0.5	0.5

↳ Table for Marginal PMF of x

$Y =$	0,	1
$P(Y)$	0.65	0.35

↳ Table for Marginal ~~PDF~~ PMF of Y

Q- $f_{XY}(x, y) = x + y$; $0 \le x \le 1$, $0 \le y \le 1$

① If $f_{XY}(x,y)$ is a joint PDF

i) $f_{XY}(x,y) \ge 0 \quad \forall \; 0 \le x, y \le 1$

ii) $\int_0^1 \int_0^1 (x+y)\, dx\, dy$

$= \int_0^1 \left[\frac{x^2}{2} + xy\right]_0^1 = \int_0^1 \frac{1}{2} + y \, dy$

$= \left[\frac{1}{2}y + \frac{y^2}{2}\right]_0^1$

$= 1$

Marginal PDF for $X = \int_0^1 (x+y)\, dy$

$$= \left[xy + \frac{y^2}{2} \right]_0^1$$

$$= x + \frac{1}{2}$$

Ans) $f_X(x) = \begin{cases} x + \frac{1}{2} \; ; & 0 \le x \le 1 \end{cases}$

(limits of x are also IMP)

Q-

X \ Y	1	2	3	4	5	6
0	0	0	$\frac{1}{32}$	$\frac{2}{32}$	$\frac{2}{32}$	$\frac{3}{32}$
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
2	$\frac{1}{32}$	$\frac{1}{32}$	$\frac{1}{64}$	$\frac{1}{64}$	0	$\frac{2}{64}$

i) Find the Marginal ~~PDF~~ PMF for X & Y and hence E(X), E(Y), var(X) and var(Y)

ii) $Pr(X \le 1, Y = 2)$

iii) $Pr(X \le 1 \mid Y = 2)$

iv) $Pr(X \le 1)$

v) $Pr(X < 3, Y \le 4)$

$<$ v/s $\leq$ $\Rightarrow$ Matters in Discrete
$\Rightarrow$ Does not matter in Continuous

X	0	1	2
$P(X)$	$\frac{8}{32} = 1/4$	$\frac{10}{16} = 5/8$	$\frac{8}{64} = 1/8$

$\Rightarrow$ Marginal PMF for X.

Y	1	2	3	4	5	6
$P(Y)$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{11}{64}$	$\frac{13}{64}$	$\frac{6}{32}$	$\frac{16}{64}$

$\Rightarrow$ Marginal PMF for Y

$$E(X) = 0 \times \frac{8}{32} + 1 \times \frac{10}{16} + 2 \times \frac{8}{64}$$

$$E(Y) = 1 \times \frac{3}{32} + 2 \times \frac{3}{32} + 3 \times \frac{11}{64} + 4 \times \frac{13}{64} + 5 \times \frac{6}{32} + 6 \times \frac{16}{64}$$

$$E(X^2) = 0^2 \times \frac{8}{32} + 1^2 \times \frac{10}{16} + 2^2 \times \frac{8}{64}$$

$$Var(X) = E(X^2) - (E(X))^2$$

ii) $Pr(X \leq 1, Y = 2) = 0 + \frac{1}{16}$

Ma'am:

$$pr(X \leq 1 \mid Y = 2)$$
$$= \frac{Pr(X \leq 1 \cap Y = 2)}{Pr(Y = 2)}$$
$$= \frac{Pr(X = 0 \cap Y = 2)}{Pr(Y = 2)} + \frac{Pr(X = 1 \cap Y = 2)}{Pr(Y = 2)}$$
$$= 0 + \frac{1/16}{0 + \frac{1}{16} + \frac{1}{32}}$$

$Pr(X \leq 1) = 1/4 + 5/8$

$Pr(X < 3, Y \leq 4) = 3/32 + 3/32 + 11/64 + 13/64$

Joint Distribution function $F_{XY}(x,y)$

or

Joint CDF

$$F_{XY}(x,y) = Pr(X \le x, Y \le y)$$

$$= \int_{-\infty}^{y} \int_{-\infty}^{x} f_{XY}(x,y)\, dx\, dy$$

or $$f_{XY}(x,y) = \frac{\partial^2}{\partial x\, \partial y} F_{XY}(x,y)$$

(partial derivative once w.r.t. x & then w.r.t. y)

Q- Let $$f_{XY}(x,y) = \frac{9(1+x+y)}{2(1+x)^4(1+y)^4}$$

$$-\infty < x, y < \infty$$

i) Find $f_X(x)$ and $f_Y(y)$

ii) $E(X)$ and $E(Y)$

Q- $f_{XY}(x,y) = 2$; $0 < x < 1$, $0 < y < x$

i) Check if $f_{XY}(x,y)$ is a joint PDF

ii) Find the marginal PDF for X and Y

iii) $E(X)$, $E(Y)$, $Var(X)$, $Var(Y)$

iv) $Pr(X \le \frac{1}{2})$, $Pr(X \le \frac{1}{2}, Y \le \frac{1}{2})$

i) $\int_0^1 \int_0^x 2 \, dy \, dx$

$\int_0^1 [2y]_0^x \, dx = \int_0^1 2x \, dx$

$= \cancel{2} \left[\frac{x^2}{\cancel{2}}\right]_0^1 = 1$

Yes, it is joint PDF

ii)

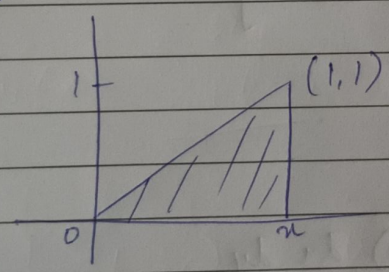
Marginal PDF for X $= \cancel{\int_0^x 2 \, dy = [2y]_0^x}$

$= \cancel{2x}$

Marginal PDF for Y $= \cancel{\int_0^1 2 \, dx = [2x]_0^1 = 2}$

Ma'am

ii)



$f_X(x) = \int_{y=0}^{x} f_{XY}(x,y) \, dy \qquad 0 < x < 1$

$f_Y(y) = \int_{x=y}^{1} f_{XY}(x,y) \, dx \qquad 0 < y < 1$

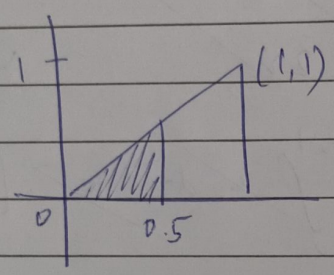
$\left.\begin{array}{l} 0 < x < 1 \\ 0 < y < x \end{array}\right\}$ x & y are dependent variables
∴ draw graph & then find limits for prob
↳ Imp

$$f_X(x) = \int_{y=0}^{x} f_{XY}(x,y)\,dy$$

$$= \int_0^x 2\,dy = 2x$$

iv) $$\Pr\left(X \le \frac{1}{2}\right) = \int_0^{1/2} 2x\,dx = \left[x^2\right]_0^{1/2}$$

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



x travels horizontally
y travels vertically

$$\int_{y=0}^{0.5} \int_{x=y}^{0.5} f_{XY}(x,y)\,dx\,dy$$

OR

$$\int_{x=0}^{0.5} \int_{y=0}^{x} f_{XY}(x,y)\,dx\,dy$$

Outer limits are always constants

iii) $$f_X(x) = \begin{cases} 2x & 0 < x < 1 \end{cases}$$

$$E(X) = \int_{x=0}^{1} x \times (\text{PDF of } X)\,dx$$

$$= \int_0^1 x \times 2x \, dx$$

$$= \int_0^1 2x^2 \, dx$$

$$= \left[\frac{2}{3} x^3\right]_0^1 = \frac{2}{3}$$

Unit 2

Special distributions

↳ All distributions are very imp.

1) Bernoulli distribution

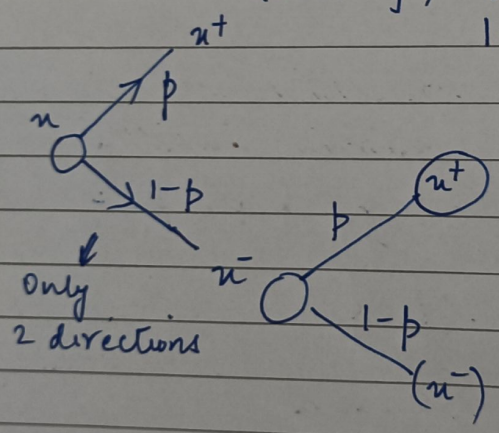
A random variable X is said to follow Bernoulli if it assumes following PMF

$$Pr(X = x) = \begin{cases} p & \text{if } x = 1 \\ 1-p = q & \text{if } x = 0 \end{cases}$$

↳ Random variable which takes only binary values (0 or 1)

Generally, $p \Rightarrow$ success

$1-p \Rightarrow$ failure



→ For every random variable, can find mean, variance and kurtosis

Mean $= p$

Variance $= E(X^2) - [E(X)]^2$

Kurtosis =

Imp

2) Binomial Distribution

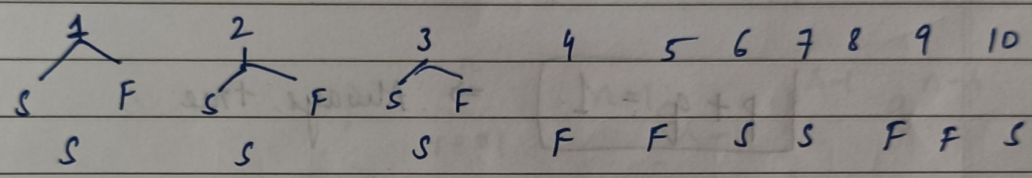
↳ Generalisation of Bernoulli distribution

↳ Binomial ≡ Repetition of Bernoulli R.V. finite no. of times

Bernoulli ⟶ Binomial

Bernoulli approaches Binomial when u repeat Bernoulli finite no. of times and each time u hv independent output.

n times



$Pr(S = 5)$

↳ These probabilities are found out by Binomial distribution

$$Pr(S=6) = \underbrace{p \times p \times p \times p \times p \times p}_{6} \times (q)^4$$

$$P(S \cap S \cap \ldots) = \underbrace{P(S)}_{\text{Prob of success}} \times P(S)$$

∵ trials done independently

$$= p \times p \times p \times q \times q \times p \times p \times q \times q \times p$$

$$= p^6 q^4$$

$$Pr(S=6) = {}^{10}C_6 \, p^6 \, q^{(10-6)}$$

to consider all arrangements

→ A Random variable X is said to follow Binomial distribution with n trials, p = Prob of success and q = Prob of failure $(p+q=1)$

$$Pr(X=r) = \begin{cases} {}^nC_r \, p^r \, q^{n-r} & ; r = 0, 1, 2, \ldots n \\ 0 & \text{otherwise} \end{cases}$$

$\boxed{p+q=1}$ ⇒ Always true

↳ prove it is a PMF

i) $\sum_{r=0}^{n} {}^nC_r \, p^r \, q^{n-r}$

$= (p+q)^n = 1$

ii) Find mean, variance

$\cancel{\sum_{r=0}^{n} {}^nC_r \, p^r \, q^{n-r}} = \cancel{(p+q)^n} = np.$

Mean $E(X) = \sum_{r=0}^{n} r \, {}^nC_r \, p^r \, q^{n-r}$

$$E(X) = \sum_{r=0}^{n} r \cdot {}^{n}C_{r}\, p^{r} q^{n-r}$$

$$= \sum_{r=0}^{n} r \cdot \frac{n!}{r!\,(n-r)!}\, p^{r} q^{n-r}$$

$$= \sum_{r=0}^{n} \frac{n!}{(r-1)!\,(n-r)!}\, p^{r} q^{n-r}$$

$$= \sum_{r=1}^{n} n \cdot \frac{(n-1)!}{(r-1)!\,(n-r)!}\, p^{r} q^{n-r}$$

$$= np \sum_{r=1}^{n} {}^{n-1}C_{r-1}\, p^{r-1} q^{n-r}$$

$$= np\,(p+q)^{n}$$

$$\boxed{\text{Mean} = np}$$

→ In Bernoulli, expected value = p

In Binomial, repeated n times

∴ expected value = np

$$E(X^2) = \sum_{r=0}^{n} r^2\, {}^{n}C_{r}\, p^{r} q^{n-r}$$

$$= \sum r^2 \frac{n!}{r!\,(n-r)!}\, p^{r} q^{n-r}$$

$$= \sum r \cdot \frac{n!}{(r-1)!\,(n-r)!}\, p^{r} q^{n-r}$$

$$= np \sum_{r=1}^{n} r \frac{(n-1)!}{(r-1)!\,(n-r)!}\, p^{r-1} q^{n-r}$$

$$= np^2 \sum_{r=2}^{n} \frac{(n-1)!}{(r-2)!\,(n-r)!}\, p^{r-2} q^{n-r}$$

$$= np^2$$

$$\text{Variance} = E(X) - E(X^2)$$

$$= np - np^2$$

$$= np(1-p) = npq$$

$$E(x^2) = \sum_{r=0}^{n} r^2 \, {}^nC_r \, p^r q^{n-r}$$

$$= \sum_{r=0} (r^2 - r + r) \, {}^nC_r \, p^r q^{n-r}$$

$$[r(r-1) + r] \, {}^nC_r \, p^r q^{n-r}$$

$$\sum \frac{\cancel{r(r-1)}(r+1)}{\cancel{r(r-1)}} \frac{n!}{(r-2)!\,(n-r)!} p^r q^{n-r}$$

$$= np^2$$

<u>Ma'am:</u>

$$E(x^2) = \sum_{r=0}^{n} r^2 \, {}^nC_r \, p^r q^{n-r}$$

$$= \sum_{r=0}^{n} \frac{n!}{(r-2)!\,(n-r)!}$$

$$= n(n-1)\,p^2 \sum_{r=2}^{n} {}^{n-2}C_{r-2} \, p^{r-2} q^{n-r}$$

Derivation ques can come in exam

$$= n(n-1)\,p^2 \, (p+q)^n$$

$$= n(n-1)\,p^2$$

$$\text{Variance} = E(x^2) - [E(x)]^2$$

$$= n(n-1)\,p^2 - (np)^2$$

$$= npq$$

Learn formulas for all distributions and formulas for mean, variance & kurtosis

Eg: A and B play a game in which chance of winning is in the ratio 3 : 2. Find A's chance of winning at least three games out of 5

$$p = 3/5 \quad q = 2/5$$

$$\cancel{1 - {}^5C_2 \cdot p^2 q^3 - {}^5C_0 p^0 q^5 - {}^5C_1 p^1 q^4 = 1 - \frac{5 \cdot 4 \cdot 3}{2 \cdot 1} \cdot \left(\frac{3}{5}\right)^2 \cdot \left(\frac{2}{5}\right)^3}$$

$$\cancel{= 1 - 10 \left(\frac{3}{5}\right)^2 \left(\frac{2}{5}\right)^3}$$

Ma'am:

Let X be a R.V. denote the no. of winning

$$Pr(X = r) = \begin{cases} {}^5C_r \, p^r q^{5-r} & r = 0, 1, 2, 3, 4, 5 \\ & p = \frac{3}{5}, \; q = \frac{2}{5} \end{cases}$$

Imp to write

$$Pr(X \geq 3) = {}^5C_3 \, p^3 q^2 + {}^5C_4 \, p^4 q^1 + {}^5C_5 \, p^5 q^0$$

$$= 10 \left(\frac{3}{5}\right)^3 \left(\frac{2}{5}\right)^2 + 5 \left(\frac{3}{5}\right)^4 \left(\frac{2}{5}\right)^1 + \left(\frac{3}{5}\right)^5$$

→ # Kurtosis of Binomial Distribution : ⇒ HW.

Q- The mean and variance of binomial distribution are 4 and $\frac{4}{3}$ respectively. Find $Pr(X \geq 1)$

$$np = 4$$

$$npq = \frac{4}{3} \quad \Rightarrow \quad q = 1/3 \quad \Rightarrow p = 2/3$$

$$n \cdot \frac{\cancel{2}}{3} = \cancel{4}\,2 \Rightarrow n = 6$$

$$Pr(X \geq 1) = 1 - Pr(X = 0)$$

$$= 1 - {}^6C_0 \, p^0 q^6 = 1 - \left(\frac{1}{3}\right)^6$$

HW

Q- If the prob of hitting target is $\frac{1}{5}$ and if 10 shots are fired, what is the conditional prob ~~that~~ of target being hit at least twice assuming that at least one hit is already scored.

Ans) 0.6999

Poisson distribution

A R.V. X is said to follow a poisson distribution with mean rate λ if it has following PMF

$$Pr(X = m) = \begin{cases} \frac{e^{-\lambda} \lambda^m}{m!} & m = 0, 1, 2, \ldots \infty \\ 0 & \text{otherwise} \end{cases}$$

In Binomial, till n
But in Poisson, going till ∞

Q- Mean, variance & kurtosis ?

4/2/23

Poisson Distribution

↳ (counting)

Def^n: (Imp) A R.V. X is said to follow Poisson distribution if its pmf is given as

$$Pr(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}, \quad x = 0, 1, 2, \ldots$$

$\lambda > 0$

↳ Parameter
(value if found can find all probabilities)

Eg: In Binomial → n, p
In Poisson → λ

$$\boxed{Pr(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}}, \quad x = 0, 1, 2 \ldots$$

$\lambda > 0$

λ: Average rate of arrival

$$\lambda = \frac{10}{\text{Program}}$$

$$X \sim P(\lambda)$$

Poisson variate X: denotes the occurence of an event or counting

λ: Average rate of arrival per unit time

i) $\frac{e^{-\lambda} \lambda^x}{x!} \geq 0$

ii) $$\sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!}$$

$$= e^{-\lambda} \cdot e^{\lambda}$$

$$= 1$$

Moments of the Poisson Distribution

$$\text{Mean} = E(X) = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{(x-1)!}$$

$$= \lambda \sum_{x=1}^{\infty} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} \times e^{\lambda}$$

$$\boxed{\text{Mean} = \lambda}$$

Mean = Average rate of arrival $= \lambda$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} (x^2 + x - x) \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} \{x(x-1) + x\} \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{x(x-1)\, e^{-\lambda}\lambda^{x}}{x!} + \sum_{x=0}^{\infty} \frac{x\, e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^{x}}{(x-2)!} + \lambda$$

$$= e^{-\lambda}\lambda^2 \sum_{x=0}^{\infty} \frac{\lambda^{x-2}}{(x-2)!} + \lambda$$

$$= \lambda^2 + \lambda$$

$$var(x) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

$$\boxed{Mean(x) = \lambda = var(x)}$$

Mean = variance

↳ In Poisson

v/s Binomial

Mean = np

variance = npq

Mean > variance

$$Kurtosis = \frac{\mu_4}{\mu_2^2}$$

Q. If X is a poisson variate such that

$$Pr(X=2) = 9\,Pr(X=4) + 90\,Pr(X=6)$$

find out the mean and standard deviation of X

Find λ : $Pr(X=2) = 9\,Pr(X=4) + 90\,Pr(X=6)$

$$\frac{e^{-\lambda}\lambda^2}{2!} = 9\frac{e^{-\lambda}\lambda^4}{4!} + 90\frac{e^{-\lambda}\lambda^6}{6!}$$

$$\lambda^4 + 3\lambda^2 - 4$$

$$\lambda^2 = \frac{-3 \pm 5}{2} = 2/2$$

$$\lambda = \pm 1 \Rightarrow \lambda = 1 \quad (\because \lambda > 0)$$

Average rate (counting) cannot be -ve

Q. A manufacturer who produces medicine bottles finds that 0.1% of the bottles are defective. The bottle are packed in boxes containing 500 bottles. A drug manufacturer buys 100 boxes from the producers of bottles. Using poisson distribution, how many boxes will contain

i) no defective ii) at least two defectives

$$\frac{0.1}{100} \times 500 = 0.5$$

↳ Usually, hv to figure out which type of distribution

λ can take any non-negative value (can be fraction)
↳ Theoretically

Practically, λ shd be whole no.

$$\lambda = \frac{0.1}{100} \times 500 = 0.5$$

Ans) X: Let X be a R.V. denoting no. of defective and follows Poisson distribution

Imp to write

with

$$\lambda = 0.1\% \times 500$$

i) $Pr(X=0) \times 100$

Note: $\lambda = np$

Ans) $p = 0.1\%$ $\quad n = 500$

$$\lambda = np = 500 \times 0.1\% = 0.5$$

i) $100 \times Pr(X=0) = 100 \times \frac{e^{-\lambda}\lambda^0}{0!} = 100 e^{-\lambda}$

$$= 100 e^{-0.5}$$

$$\approx 61$$

ii) $Pr(X \geq 2) = 1 - \{Pr(X=0) + Pr(X=1)\}$

$Pr(X=2) + Pr(X=3) + \ldots$

↳ Ans) $100 \times Pr(X \geq 2)$

Q2) In a book of 520 pages, 390 typo graphical errors occur.

Negative Binomial Distribution

Defn: A random variable X is said to follow a negative Binomial distribution if its pmf is given by

$$Pr(X=x) = \begin{cases} {}^{x+r-1}C_{r-1}\, p^r q^x & ; \quad x = 0,1,2,\ldots \\ 0 & \text{otherwise} \end{cases}$$

q = prob of failure

↳ Prob that there are x failure preceding the rth success in $(x+r)$ trials.

$p^{r-1} q^x \leftarrow p$ ← r^{th} success

(r^{th} success se pehle x failures and $(r-1)$ success if $(x+r)$ trials conducted)

↳ r^{th} position fixed for success ∴ taking $(r-1)$

(positions can change for $(r-1)$ success & x failure before r^{th} success → ∴ ${}^{x}C_{r}$)

$$x+r-1C_{r-1} = x+r-1C_{x}$$

$$= \frac{(x+r-1)(x+r-2)\cdots(r+1)\,r\,\cancel{(r-1)!}}{x!\,\cancel{(r-1)!}}$$

$$= \frac{(x+r-1)(x+r-2)\cdots(r+1)(r)}{x!}$$

$$= (-1)^x \frac{(-r)(-r-1)\cdots(-r-x+2)(-r-x+1)}{x!}$$

$$\boxed{x+r-1C_{r-1} = (-1)^x\, -rC_x}$$

v/s ${}^{r}C_{x}$ ↳ Binomial

$$^{n+r-1}C_{r-1} = (-1)^n \, {}^{-r}C_n$$

$$Pr(X=x) = \begin{cases} {}^{-r}C_x \, p^r (-q)^x & x = 0, 1, 2 \ldots \\ 0 & \text{otherwise} \end{cases}$$

Let $p = \frac{1}{Q}$; $q = \frac{P}{Q}$ so that $Q - P = \frac{1}{P} - \frac{q}{P}$

$$= \frac{1-q}{P} = 1$$

$$Pr(X=x) = \left\{ {}^{-r}C_x \, Q^{-r} \left(\frac{-P}{Q}\right)^x \right. \quad x = 0, 1, 2, \ldots$$

↳ can use this form

↳ better ∵ similar to mean & variance as in Binomial.

$$\text{mean} = rP = \frac{rq}{p}$$

$$\text{variance} = rPQ = \frac{rq}{p^2}$$

eg1 If the prob is 0.4 that a child exposed to a certain disease will catch it. What is the prob that the tenth child exposed to the disease will be third to catch it?

If position is fixed (nth) ↓ ve Binomial

$$^{n+r-1}C_{r-1} \, p^{r-1} q^n \, p^r$$

$p = 0.4$

$q = 0.6$

$$^{x+r-1}C_{r-1}\, p^{r-1} q^{x} p^{r}$$

$p = 0.4$

$q = 0.6$

$r = ?$ $\quad r = 3$ ($\because$ third position fixed)

$x = ?$

$x = 7$ ($\because x + r = 10$)

$$= {}^{9}C_{2}\, p^{2} q^{7}\, p^{1}$$

Eg2 Let X be the nos. of births in a family until the second daughter is born. If the prob of having a male child is 1/3, find the prob that the sixth child in the family is the second daughter?

$p = 1/3 \qquad q = 2/3$

4 male child

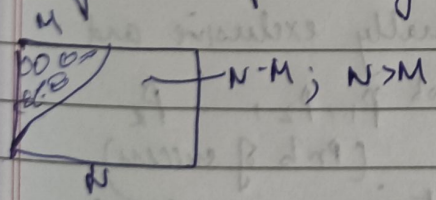
2 female child

$$= {}^{5}C_{4} \left(\frac{1}{3}\right)^{4} \left(\frac{2}{3}\right) \left(\frac{2}{3}\right)^{2}$$

Q- In a company 5% defective components are produced. What is the prob that at least 5 components are to be examined in order to get 3rd defective?

Hyper-Geometric Distribution

Defn: A discrete R.V. X is said to follow the hyper geometric distribution if its prob is given as



$$Pr(X=k) = \begin{cases} \dfrac{{}^{M}C_{k}\ {}^{N-M}C_{n-k}}{{}^{N}C_{n}} & ; k = 0, 1, 2, \ldots \min\{n, M\} \\ 0 & \text{otherwise} \end{cases}$$

$$Pr(X=k) = \begin{cases} {}^{M}C_{k} \; {}^{N-M}C_{n-k} & ; \; k = 0, 1, 2, \ldots \min\{n, M\} \\ 0 & \text{otherwise} \end{cases}$$

$$k = 0, 1, 2, \ldots \min\{n, M\}$$

$$\text{Mean} = E(X) = \sum_{k=0}^{n} k \frac{{}^{M}C_{k} \; {}^{N-M}C_{n-k}}{{}^{N}C_{n}}$$

$$\text{Mean} = \frac{nM}{N}$$

$$E(X^2) = \sum_{k=0}^{n} k^2 \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} = \frac{nM}{N}$$

Multinomial Distribution

↳ Generalization of Binomial distribution

p q

Suppose $E_1, E_2, \ldots E_k$ are k mutually exclusive and exhaustive events of a trial with prob $p_1, p_2, \ldots p_k$ (Prob of success)

Then the prob that E_1 occurs x_1 times, E_2 occurs x_2 times, ... E_k occurs x_k times.

$$\boxed{P(x_1, x_2, \ldots x_k) = \frac{n!}{x_1!\, x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}} \quad \sum x_i = n$$

$$\sum_{i=1}^{k} p_i = 1$$

Q- Consider an experiment of getting 4 types of outputs 1, 2, 3, 4 with equal prob.

Find the prob of getting output 1 : 5 times

output 2 :

Continuous Distribution

(1) Uniform distribution / Rectangle distribution

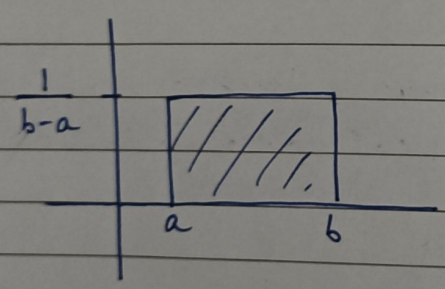
A R.V. X is said to follow a uniform distribution if it has a following pdf

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Eg $x = 1, 2, \ldots n$ $\Rightarrow$ In Discrete

$P(X = n) = \dfrac{1}{n}$

similarly in continuous, $f_X(x) = \dfrac{1}{b-a}$



$\rightarrow$ Mean $(\mu) = E(X)$ $\rightarrow$ centre of distribution (centre of rect $\Rightarrow \dfrac{b+a}{2}$)

$$= \int_a^b x \left(\frac{1}{b-a}\right) dx$$

$$= \left| \frac{x^2}{2} \right|_a^b \times \frac{1}{b-a}$$

$$= \frac{b^2-a^2}{2} \times \frac{1}{b-a}$$

$$= \frac{a+b}{2}$$

$$\boxed{\text{Mean} = \frac{a+b}{2}}$$

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

$$= \int_a^b x^2 \frac{1}{b-a}\,dx - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{b^3-a^3}{3} \cdot \frac{1}{b-a} - \left(\frac{a+b}{2}\right)^2 = \left(\frac{b^2+a^2+ab}{3}\right) - \left(\frac{a+b}{2}\right)^2$$

$$\boxed{\text{Var}(X) = \frac{(b-a)^2}{12}}$$

$X \sim U(a,b)$

↳ X is following Uniform distribution with parameters a and b

eg: $X \sim U(2,5)$

↳ learn Mean & variance

Q1) Let $X \sim U(a,b)$ with mean 1 and variance $\frac{4}{3}$. Find $Pr(X > 0)$

$$\frac{a+b}{2} = 1 \Rightarrow a+b = 2$$

$$\frac{(b-a)^2}{12} = \frac{4}{3} \Rightarrow (b-a)^2 = 16$$

$$\Rightarrow (2-a-a)^2 = 16 \Rightarrow (2-2a)^2 = 16$$

$$2a-2 = 4 \quad , \quad 2a-2 = -4$$

$$2a = 6 \Rightarrow a = 3 \qquad 2a = -2 \Rightarrow a = -1$$

$a = 3 \Rightarrow \cancel{a = 3/2}$ $\quad a = -1 \Rightarrow \cancel{a = -1/2}$

$b = -1 \quad \cancel{b = 1/2}$ ✗ $\quad b = 3 \Rightarrow \cancel{b = 5/2}$ ✓

$$\cancel{Pr(X > 0) = \int_0^{5/2} \frac{1}{b-a}\, dx} = \quad Pr(X > 0) = \int_0^3 \frac{1}{3-(-1)}\, dx$$

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

Q2) A subway train runs every half hour between midnight to 6'o clock in the morning. What is the prob a man entering the station during this period will have to wait for atleast 20. minutes?

X : waiting time (range of waiting time is 0 to 30)

$$X \sim U(0, 30) \; ; \; f_X(x) = \frac{1}{30}$$

$$Pr(X \geq 20) = \int_{20}^{30} \frac{1}{30}\, dx$$

$$= \frac{1}{30} x \Big|_{20}^{30} = \frac{1}{30}(30 - 20) = \frac{10}{30}$$

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

(2) Normal distribution / Gaussian distribution

$$X \sim N(\mu, \sigma^2) \; ; \; -\infty < \mu < \infty \; ; \; \sigma^2 \geq 0$$

↳ X is following Normal distribution with parameters μ and σ^2

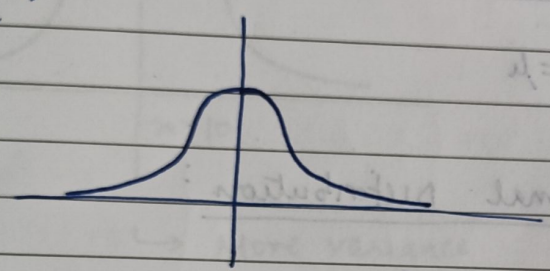
$X \sim N(\mu, \sigma^2)$; $-\infty < \mu < \infty$, $\sigma^2 \geq 0$

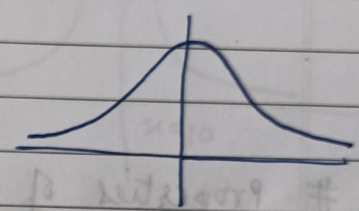
if it has the following pdf

$$f_X(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} & ; -\infty < x < \infty \\ 0 & \text{otherwise} \end{cases}$$

↳ learn formula

e^{-x^2}



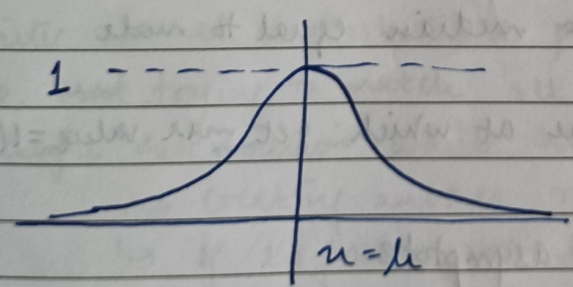


↳ Follows N.D if ① Non-negative

② $\int_{-\infty}^{\infty} = 1$

Mean & variance shd be μ and σ resp.

$e^{-x^2} = e^{-\left(\frac{x-0}{1}\right)^2} \equiv e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ where $\mu = 0$, $\sigma = 1$



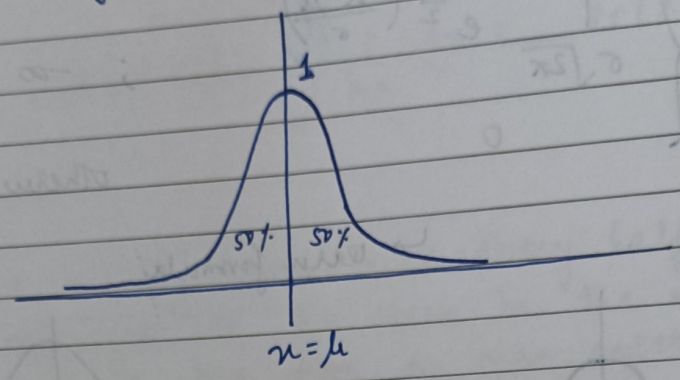
Max value = 1

$$\boxed{\begin{aligned} E(X) &= \mu \\ Var(X) &= \sigma^2 \end{aligned}}$$

↳ Parameters are itself μ and σ^2 resp.

Mean Variance

~~Pxo~~
Curve for Normal Distribution



Properties of Normal Distribution:

① The curve is bell-shaped and symmetric about $x=\mu$
(area equally distributed about $x=\mu$)

② Mean = Median = Mode

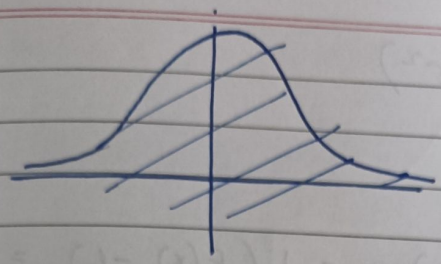
↳ Only distribution where mean is equal to median equal to mode

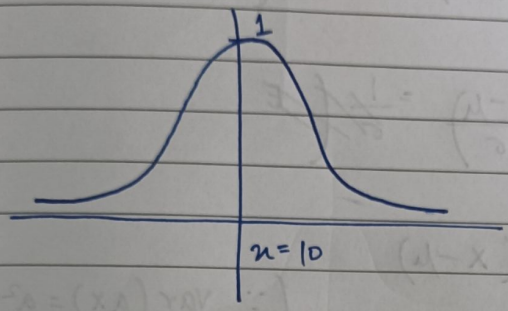
(Mode → value at which get max value = 1 (here))

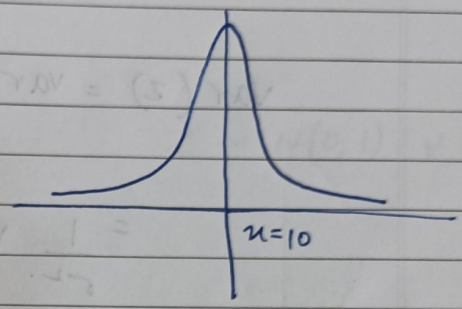
③ x-axis is the asymptote

④ Kurtosis = 3

↳ eg: X : height, weight follows Normal distribution
(bell-shaped graph)



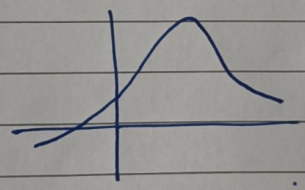




↳ More variance (More dispersed)

↳ Less variance

Mean coinciding but variance different



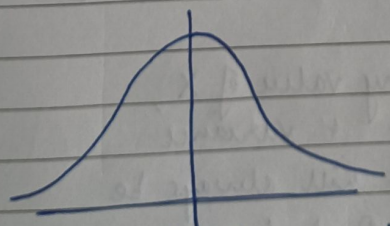
→ Variance same, mean different

Infinite Normal distributions possible if μ and σ^2 changes

So, ~~not~~ trying to match all distribution with a single distribution which has $\mu = 0$, $\sigma^2 = 1$.

↳ Creating another random variable Z.

↳ If ~~X~~ follows Normal dist. then Z also follows Normal dist.



$\mu = 0$

$\sigma^2 = 1$

↳ Matching with this graph

$$X \sim N(\mu, \sigma^2)$$

$$Z = \frac{X-\mu}{\sigma}$$

$$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\left(E(X) - \mu\right) = 0$$

$$var(Z) = var\left(\frac{X-\mu}{\sigma}\right) = \cancel{\frac{1}{\sigma}\left(E\right.}$$

$$= \frac{1}{\sigma^2} var(X-\mu)$$

$$= \frac{1}{\sigma^2} var(X)$$

$$= \frac{1}{\sigma^2} \times \sigma^2$$

$$= 1$$

$$\left(\because var(aX) = a^2 var(X) \quad var(X+b) = var(X) \quad \because var(b) = 0 \right)$$

const.

$$\boxed{\begin{aligned} E(Z) &= 0 \\ var(Z) &= 1 \end{aligned}}$$

X : Normal variate ; $X \sim N(\mu, \sigma^2)$

$Z = \frac{X-\mu}{\sigma^2}$; Standard Normal variate

$Z \sim N(0,1)$

↳ For any value of X, mean + variance of Z will always be 0 + 1 resp.

$Pr(x_1 \le X \le x_2)$ where $X \sim N(\mu, \sigma^2)$

$$\begin{cases} X_1 \sim N(\mu_1, \sigma_1^2) \\ X_2 \sim N(\mu_2, \sigma_2^2) \\ X_3 \sim N(\mu_3, \sigma_3^2) \\ \vdots \\ X_N \end{cases} \longrightarrow Z = \frac{X-\mu}{\sigma}$$

$$Z = \frac{X_i - \mu}{\sigma} \qquad Z \sim N(0,1) \ \forall i$$

↳ standard table for Z

$$Pr\left(\frac{x_1-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{x_2-\mu}{\sigma}\right)$$

$Pr(x_1 \le X \le x_2)$ where $X \sim N(\mu, \sigma^2)$

$$= Pr\left(\frac{x_1-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{x_2-\mu}{\sigma}\right)$$

$$= Pr\left(\frac{x_1-\mu}{\sigma} \le Z \le \frac{x_2-\mu}{\sigma}\right)$$

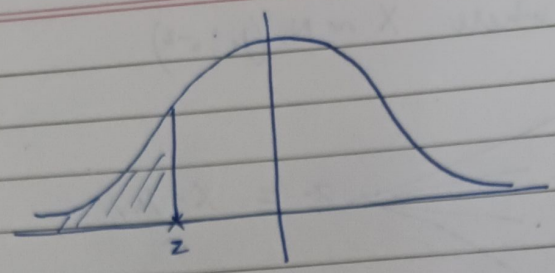
↳ Use Z-table

Q- Let $X \sim N(12, 16)$, $\mu = 12$, $\sigma^2 = 16$

a) $Pr(X \ge 20)$ b) $Pr(X \le 20)$ c) $Pr(0 \le X \le 12)$

$$Pr\left(\frac{X-12}{4} \ge \frac{20-12}{4}\right) = Pr(Z \ge 2)$$

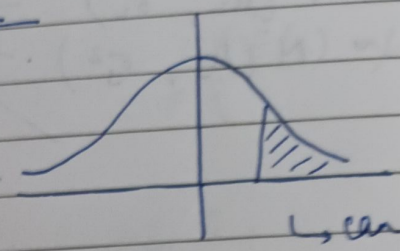
have to divide by σ (Not σ^2)



$Pr(Z = z)$

⇓

our z table follows this area

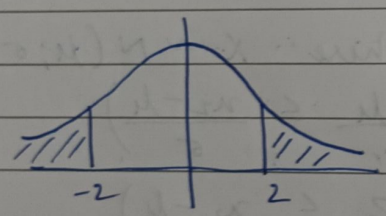


↳ can also take this

$Pr(X \geq 20)$

$= Pr\left(\frac{X-12}{4} \geq \frac{20-12}{4}\right) = Pr(Z \geq 2) = Pr(Z \leq -2)$

$= 1 - Pr(Z < 2)$



$Pr(Z = z)$

⇒ Both shaded regions will give same answer

postfix

Z	.00	.01	.02
-3.4			
-3.3			

prefix

$Z \leq -3.400$

$Pr(Z \leq -3.42)$

$Pr(Z \leq -3.4$
0.01

$= Pr(Z \leq -3.41)$

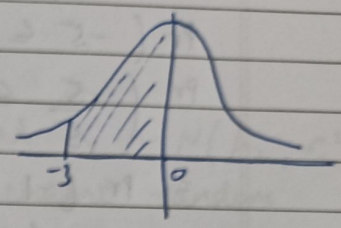
$Pr(Z \geq 2) = 1 - Pr(Z < 2)$
$= 1 - 0.9772$

$(Pr(Z \leq 2) = Pr(Z <$

$Pr(Z \geq 2) = Pr(Z \leq -2) = 0.0228$

$Pr(0 \leq X \leq 12)$

$= Pr\left(\frac{0-12}{4} \leq \frac{X-12}{4} \leq \frac{12-12}{4}\right)$



$= Pr(-3 \leq Z \leq 0)$

$= Pr(Z \leq 0) - Pr(Z \leq -3)$

$= 0.5 - 0.0013$

$= 0.4997$

$$\begin{array}{r} 0.5000 \\ 0.0013 \\ \hline 4997 \end{array}$$

Q- $X \sim N(30, 25 = \sigma^2)$ Find

a) $Pr(26 \leq X \leq 40)$

b) $Pr(|X-30| < 5)$

c) $Pr(|X-30| \geq 5)$

a) $Pr(26 \leq X \leq 40) = Pr\left(\frac{26-30}{5} \leq \frac{X-30}{5} \leq \frac{40-30}{5}\right)$

$= Pr\left(\frac{-4}{5} \leq Z \leq \frac{10}{5}\right)$

$= Pr(-0.8 \leq Z \leq 2)$

$= Pr(Z \leq 2) - Pr(Z \leq -0.8)$

$= 0.9772 - 0.2119$

$= 0.7653$

$$\begin{array}{r} 0.9772 \\ -0.2119 \\ \hline 7653 \end{array}$$

b) $Pr(|X-30| < 5)$

$= Pr\left(\frac{X-30}{5} < \frac{5}{5}\right)$

$= Pr(Z < 1) = 0.8413$

c) $Pr(|X-30| \leq 5) = Pr\left(\frac{X-30}{5} \leq \frac{5}{5}\right) = Pr(Z \leq 1)$

$= 0.8413$

b) $Pr(-5 < X-30 < 5)$

$Pr(-5 \leq X-30 \leq 5)$

$Pr\left(\frac{-5}{5} \leq \frac{X-30}{5} \leq \frac{5}{5}\right)$

$= Pr(-1 \leq Z \leq 1)$

$= Pr(Z \leq 1) - Pr(Z \leq -1)$

$= 0.8413 - 0.1587$

$= 0.6826$

c) ~~0.6826~~

Q. The mean yield for one-acre plot is 662 kilos with standard deviation of 32 kilos. Assume Normal Distribution. How many one-acre plot (in a batch of 1000) will have yield

(1) more than 700 kilos

(2) below 650 kilos

(3) between 550 to 600 kilos

Let X be a R.V. denoting yield

X: yield

$X \sim N(662, 32)$

$Pr(X \geq 700)$

$= Pr\left(\frac{X-662}{32} \geq \frac{700-662}{32}\right)$

$= Pr\left(Z \geq \frac{38}{32}\right) = Pr\left(Z \geq \frac{19}{16}\right)$

$= Pr(Z \geq 1.1875)$

$= 1 - Pr(Z \leq 1.1875)$
$= 1 - Pr(Z \leq 1.19)$
$= 1 - 0.8830$

Result: let $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ where X_1 and X_2 are independent random variables, then

$$\boxed{aX_1 + bX_2 \sim N(a\mu_1 + b\mu_2, a^2\sigma_1^2 + b^2\sigma_2^2)}$$

A and B are Independent Events if $Pr(A \cap B) = Pr(A) \times Pr(B)$

If X_1 & X_2 are two independent Random variables if

$$f_{X_1X_2}(x_1 x_2) = f_{X_1}(x_1) \times f_{X_2}(x_2)$$

(Joint Distribution = Product of their marginal pdfs)

eg: $X_1 \sim N(2, 9)$ $\quad X_2 \sim N(4, 16)$

$Pr(2X_1 - 3X_2 \geq 20) = ?$

$2X_1 - 3X_2 \sim N(\mu, \sigma^2)$

$\mu = 2 \times 2 - 3 \times 4$

$\sigma^2 = 4 \times 9 + \underset{(-3)^2}{9} \times 16 \quad = 2^2 \times 9 + (-3)^2 \times 16$

Result: let $X_1, X_2, X_3, \ldots X_n$ are independent Normal variates with $X_i \sim N(\mu_i, \sigma_i^2)$

then

$$a_1X_1 + a_2X_2 + \ldots + a_nX_n \sim N\left(\sum_i a_i\mu_i, \sum a_i^2\sigma_i^2\right)$$

Gamma distribution

A Random variable X is said to follow Gamma distribution with parameter $\lambda > 0$ if

$$f_X(x) = \begin{cases} \dfrac{e^{-x}\, x^{\lambda-1}}{\Gamma(\lambda)} & ; \lambda > 0 \; ; \; x > 0 \end{cases}$$

$$\boxed{\Gamma(\lambda) = \int_0^\infty e^{-x}\, x^{\lambda-1}\, dx}$$

Gamma function $\boxed{\Gamma(z) = \int_0^\infty e^{-x}\, x^{z-1}\, dx}$

$$\int_0^\infty f_X(x)\, dx = \int_0^\infty \frac{e^{-x}\, x^{\lambda-1}}{\Gamma(\lambda)}\, dx$$

$$= \frac{1}{\Gamma(\lambda)} \int_0^\infty e^{-x}\, x^{\lambda-1}\, dx$$

$$= \frac{1}{\Gamma(\lambda)} \times \Gamma(\lambda) = 1$$

Gamma function properties:

$$\Gamma(z) = (z-1)\,\Gamma(z-1)$$
$$= (z-1)(z-2)\,\Gamma(z-2)$$

$$\Gamma(n) = n!$$

$$\boxed{\text{Mean} = \lambda = \text{Variance}}$$

The daily consumption of milk in a city, in excess of 20,000 litres is approx distributed as a Gamma variate with parameter $a = \frac{1}{10,000}$ and $\lambda = 2$. The city has a daily stock of 30,000 litres. What is the prob that stock is insufficient on a particular day?

X: daily consumption

$Y = X - 20,000 \sim$ Gamma distribution with $a = \frac{1}{10,000}$ and $\lambda = 2$

$Pr(X \geq 30,000)$ or $Pr(Y > 10,000)$

$$= \frac{1}{\Gamma(2)} \int_{10000}^{\infty} \frac{1}{(10,000)^2} \times e^{-\frac{1}{10000} y} y^{2-1} \, dy$$

$$= \frac{1}{(10,000)^2} \int_{10,000}^{\infty}$$

Beta Distribution of first kind

A R.V. X is said to have a beta distribution of first kind with parameter $(\mu + \nu) > 0$ if its pdf is given by

$$f_X(x) = \begin{cases} \frac{1}{B(\mu, \nu)} x^{\mu-1} (1-x)^{\nu-1}, & \mu > 0, \nu > 0 \\ & 0 < x < 1 \end{cases}$$

$X \sim \beta_1(\mu, \nu)$

$B(\mu, \nu)$ is a Beta function $= \int_0^1 x^{\mu-1} (1-x)^{\nu-1} \, dx$

$$\int f_X(x)\, dx = 1$$

↳ reduces to a uniform $[0,1]$
where $\mu = \nu = 1$

$$B(\mu, \nu) = \frac{\Gamma(\mu)\,\Gamma(\nu)}{\Gamma(\mu+\nu)}$$

for μ, ν are integers
$\Gamma(\mu) = (\mu - 1)!$

The moments of Beta distribution

$$E(X^r) = \int_0^1 x^r f_X(x)\, dx$$

r^{th} order moment about the origin

$r = 1 \Rightarrow$ Mean

$r = 2 \Rightarrow E(X^2) \Rightarrow$ var

$$= \frac{1}{B(\mu,\nu)} \int_0^1 x^r x^{\mu-1} (1-x)^{\nu-1}\, dx$$

$$= \frac{1}{B(\mu,\nu)} \int_0^1 x^{r+\mu-1} (1-x)^{\nu-1}\, dx$$

$$= \frac{1}{B(\mu,\nu)} * B$$

$$E(X^2) = \frac{\mu(1+\mu)}{(\mu+\nu)(\mu+\nu+1)}$$

$$\text{var}(X) = E(X^2) - (E(X))^2 = \frac{\mu(1+\mu)}{(\mu+\nu)(\mu+\nu+1)} - \left(\frac{\mu}{\mu+\nu}\right)^2$$

$$\text{var}(X) = \frac{\mu\nu}{(\mu+\nu)^2(\mu+\nu+1)}$$

Beta distribution of second kind

~~A R.V.~~

If $X \sim \beta_2(\mu, \nu)$ then $X+1 \sim \beta_1(\mu, \nu)$

Exponential distribution

↳ learn formules, pdf

Cauchy distribution

The mean of Cauchy differentiation does not exist

$E(X) \rightarrow$ Improper Integration

Cauchy distribution ⇒ Distribution whose mean does not exist

chi-square distribution χ^2

A R.V. X is said to follow a chi-square distribution with n degree of freedom if its pdf is given as

total no. of sample − no. of constraints

degree of freedom = total no. of data − no. of conditions put on the data

Typically for $n > 30$, chi distribution follows Normal distr. (in case of large data size)

Large data size ⇒ Normal distr. used

If $X \sim N(\mu, \sigma^2)$

$$z = \left(\frac{X-\mu}{\sigma}\right)^2$$

then $z^2 \sim \chi^2$ (Degree of freedom = 1)

$X_1 \sim N(\mu, \sigma_1^2)$

$X_2 \sim N(\mu, \sigma_2^2)$

$z_1^2 + z_2^2 \sim \chi^2$ (degree of freedom = 2)

$$z_1 = \frac{X_1 - \mu_1}{\sigma_1} \quad \& \quad z_2 = \frac{X_2 - \mu}{\sigma_2}$$

where X_1 and X_2 are independent