19.07.2023
Wednesday

Lecture-1

Date: Page:

STATISTICAL METHODS

These methods are used for extracting the info. where uncertainity is involved

If o/p is fixed ⇒ NO uncertainity (ON/OFF Switch)

If o/p is not fixed ⇒ ~~Is~~ uncertainity (Tossing coin)

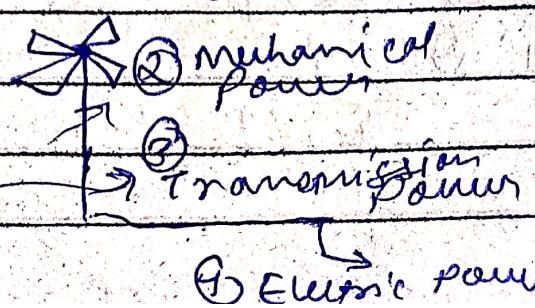
Statistical methods & ~~signs~~ science used for ~~not~~ decision making or extracting the knowledge where uncertainity is involved :-

Experiment

Deterministic (o/p known) — Switch on/off the bulb

Non-deterministic (Random) (o/p not known) — Tossing of a coin ↓ "uncertainity"

Natural Phenomenon ↓ All are random ↓ Current speed of wind = v_1 — What will be after 5 min? → Not known



Efficiency = η_i

η_1 η_2 η_3

Date: Page:

$$\boxed{P = \eta_1 \eta_2 \eta_3 \rho v^3}$$

ρ → air density

v → air speed

η_1, η_2, η_3 Remains constant unless everything is new

ρ & v changes uncertain

Most volatile = v

$\therefore P \propto v^3$ → change ઉમેરવાની ટેવ !

$P_1 = K_1 v_1$
$P_2 = K_2 v_1$

$P_1 \neq P_2$ even when $v_1 = v_1$ other factors can change.

① Study of methods to collect & organise numeric data related to phenomenon & experiment

② Analyze the numeric info.

③ Interpret the → mind

(i) Individual → coin, Tuesday, Tossing a coin

(ii) Variability : People or object included in the study

(iii) Data : → Face of the coin → Mind Speed & Power

Head / Tails

Electric

Date: _____ Page: _____

Wind → Wind Speed → Value of this

Turbine → Power Electric → Value of this

Variables

- **Qualitative**
  - There is no ordering in this.
  - Tossing a coin :- Head, Tails
  - Religion of a man
- **Quantitative**
  - Rolling a dice
  - **Discrete**
    - Outcome can be listed
  - **Continuous**
    - Values lie betn the interval (in Range)

Random Variable: Whose values are depending on the chance.

Defined on Random experiment.

→ Instances : Sample Space of exp.

Random exp.: Tossing pair of coin

Sample Space : {HH, HT, TH, TT}.

Date: Page:

Random variable = x = No. of head] defination of variable

Can only be defined on sample space of Random experiment.

Real valued function

$x = 0, 1, 2$ → Instances / Events of R.V.

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

$1/4$ $1/2$ $1/4$ → Probabilities

For each instance there is probability

R.V. is the real valued fn defined on sample space of Random exp.

x = head in a single — Not a R.V. X (Boolean true or false)

x = No. of Heads — R.V. (Counting is involved)

(1,1) (1,2) (1,3) (1,4) (1,5) (1,6)

(2,2)

(3,3)

(4,4)

(5,5)

(6,6)

	1	2	3	4	5	6
1	0	1	2	3	4	5
2	1	0	1	2	3	4
3	2	1	0	1	2	3
4	3	2	1	0	1	2
5	4	3	2	1	0	1
6	5	4	3	2	1	0

RV. = x = abs. diff of two numbers

$x = 0, 1, 2, 3, 4, 5$ (6 instances)

$\frac{1}{6}$ $\quad \frac{\cancel{10}\,5}{\cancel{36}\,18}$ $\quad \frac{8}{36}, \frac{6}{36}, \frac{4}{36}, \frac{2}{36}$

Probability f^n can be defined on x these are called instances. Prob. density f^n

If sum of the probabilities of all the instances should be equal to 1

each i/p $\Rightarrow$ unique o/p

element of SS are covered

For each i/p of SS there is unique o/p

~~A~~ ~~Each~~

Each element of SS should be name in a single instance

In this case $\sum P_i = 1$

$P_i > 0$ (atleast one element is considered)

20.07.2023
Thursday.

Lecture-2.

Date: Page:

Random variable is a mapping from a sample space Ω to the relevant space of numbers.

↓ collection of
All possible outcomes

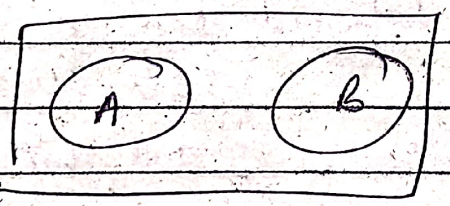
Event :- An evt. associated with a random exp. is a subset of SS
↳ Instance

E : Evt. $\bar{E}$ = evt.

A : Evt. B : Evt. $A \cap B$: Evt.
$A \cup B$: Evt.

Mutually Exclusive

A & B → no two common evts. in A & B.



↓
Random variable's instances always ME
(part of 1 instance)

Collutively Exhaustive ~~Who~~

When it covers all elements of SS & no other possibility.

Probability Distrib^n

list of all valid outcomes with associated probabilities.

Each observation belongs to one instance & all should be covered. $P_i > 0$

ME & CE

Date: Page:

Probability Density f^n A f^n that represents PD of discrete random variable. $\Sigma P_i = 1$, $\Sigma P_i > 0$

Ex:

Observations : 60

Instances :	5 ft.	5.5 ft.	6 ft.
Probability			

Can be represented by

Prob. Histogram

Mean & Standard Deviation of a Discrete RV

→ exact value

Avg. :- When no. of students are fixed.

Mean :- In form of group with probability we calculate mean.

→ expected value

(Group form).

~~Contin~~ interval based

$a < x < b$ $P(x) =$

No. of observations are very large.

Variance :- Avg. deviation of all instances w.r.t. expected value.

Variance of discrete R.V. $= V(x) = \Sigma [x_i - \mu]^2 P(x_i)$

$$\Rightarrow \boxed{\begin{aligned} V(x) &= \Sigma [x_i]^2 P(x_i) - \mu^2 \\ &= \text{Mean of squares} - \text{Sq. of mean} \end{aligned}}$$

Date: Page:

$$P[x_i]^2 + P[\mu]^2 - 2 x_i \mu$$
$$P[x_i]^2 + P[\mu]^2 - 2 \mu \mu$$
$$P[x_i]^2 - P[\mu]^2$$

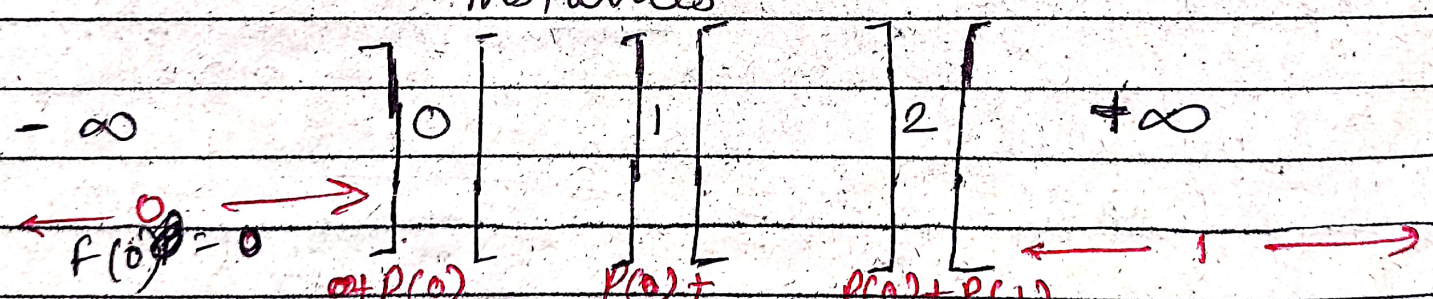
Standard Deviation $\sigma = \sqrt{\Sigma [x_i - \mu]^2 P(x_i)}$

Probability Distribution fn cumulative sum of probabilities from $-\infty$ to ∞
[all points]

valid instances = 0, 1, 2

∴ for all other points make $P = 0$

→ gives prob. of all valid & invalid instances



$$F(a) = P(x \le a) = \sum_{x_i \le a} P(x_i), \quad x_i \in X$$

union of all instances less than a

21.07.2023
Friday

Lecture 03

Date: Page:

Sum of 2 number is 5 → valid event
Abs diff. of 2 num. → Not valid
Abs. → is 5 → valid evt.

→ Union: Abs. diff 2 num. is 5 & sum of 2 number is 5
∴ Valid evt.

All instances are part of SS but vice versa is not true.

Instances are always evts but all evts are not instances

All the ~~cases~~ evts where we can generate real value
→ Random variable.
" boolean value
→ X RV X

For ~~Sam~~ same RV we can't find intersection of instances (ME)

For diff. RV we can find ∩ of instances
~~Cmay~~ :

No. of ways events can occur

Date: Page:

Probability for Equally likely Outcomes = $\frac{f}{N}$

Total no. of possible outcomes

① Random variable notat^n

X = denote total no. of tails in 4 tosses.

$\{X = 3\}$. Total no. of tails obtained in 4 tosses

②

No. of tails

$P(0) = \frac{1}{16}$ — — — —

$^4C_1 = 4$

$P(1) = \frac{4}{16} = \frac{1}{4}$

$P(2) = \frac{6}{16} = \frac{3}{8}$ — — — — $^4C_2 = \frac{4 \times 3}{2} = \frac{6}{16}$

$P(3) = \frac{3}{8}$ $^4C_3 =$

③ $X = N(H) - N(T)$

	1	1	-1
3 HHH	HHT	HTH	HTT
THH	THT	TTH	TTT
1	-1	-1	-3

① $\{X = 3, 1, -1, -3\}$

$P(3) =$

Date: Page:

(i) Probability density fn

$P(X = -3) = \frac{1}{8}$

$P(X = -1) = \frac{3}{8}$

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

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

(ii) Mean $= \frac{1}{8} \times (-1) + \frac{3}{8} \times (-3) + \frac{3}{8} \times 3 + \frac{1}{8} \times 1$

$\Sigma P_i(x_i)$

$= \cancel{\frac{1+9+9+1}{8} = \frac{20}{8} = \frac{5}{2}}$

0

(iii) Standard deviation variance $= \sigma^2 = \Sigma P(x_i) \times (x_i)^2 - \mu^2$

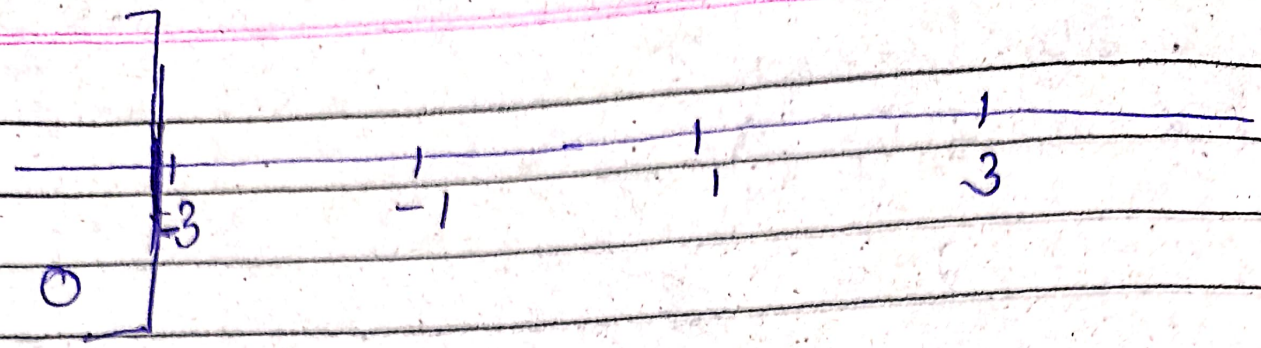
$= \cancel{\frac{9 \times 3}{8}}$

$= (-3)^2 \times \frac{1}{8} + (-1)^2 \times \frac{3}{8} + (1)\frac{3}{8} + \frac{1}{8} \times 3^2$

$= \frac{9+3+3+9}{8} = 3$

(iv) $P(\cancel{x < -3}) =$

~~Distribution fn~~



$F(-3) = P(x \leq -3) = \frac{1}{8}$

$F(-1) = P(x \leq -1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}$

$F(+1) = P(x \leq 1) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8}$

$F(+3) = P(x \leq 3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$

$$F(x) \begin{cases} 0 & x < -3 \\ 1/8 & -3 \leq x < -1 \\ 1/2 & -1 \leq x < 1 \\ 7/8 & 1 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Q)

$$P(x) = \begin{cases} k & \text{if } x = 0 \\ 2k & \text{if } x = 1 \\ 3k & \text{if } x = 2 \\ 0 & \text{otherwise} \end{cases}$$

① $\sum P_i = 1$

$k + 2k + 3k = 1 \Rightarrow 6k = 1 \Rightarrow k = \frac{1}{6}$

Date: Page:

(2) $P(x=0) = \frac{1}{6}$

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

$P(x=2) = \frac{3}{6}$

(i) $P(x<2) = \underset{x=0}{\frac{1}{6}} + \underset{x=1}{\frac{2}{6}} = \frac{3}{6} = \frac{1}{2}$

(ii) $P(x \leq 2) = 1$ (all the points

(iii) $P(0 < x \leq 2) = \frac{2}{6}$ $(x=1)$

(3) Distribution Function:-

$F(0) = P(x \leq 0) = \frac{1}{6}$

$F(1) = P(x \leq 1) = \frac{2}{6} = \frac{1}{3}$

$F(2) = P(x \leq 2) = \frac{3}{6} = \frac{1}{2}$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{6} & 0 \leq x < 1 \\ \frac{1}{3} & 1 \leq x < 2 \\ \frac{1}{2} & x \geq 2 \end{cases}$$

Date: Page:

Q)

$$P(x) = \begin{cases} 1/4 & x = -2 \\ 1/4 & x = 3 \\ 1/2 & x = 6 \\ 0 & \text{Otherwise} \end{cases}$$

(i) $P(1 < x \leq 4) = P(x=3) = \frac{1}{4}$

(ii) $P(x > 5) = P(6) = \frac{1}{2}$

(iii) $P(2x - 3 \geq 1) = P(3) + P(6) = \frac{1}{4} + \frac{2}{4}$

$2x - 3 \geq 1$ $\qquad = \frac{3}{4}$

$x \geq 2$

(iv) $P(x^2 - 2x \leq 3) = P(3) = \frac{1}{4}$

$x^2 - 2x - 3 \leq 0$

$(x-3)(x+1) \leq 0$

$x \in [-1, 3]$

Prob. Mass Distribution : Integration (Area)
Prob density Distribu^n : Isolated Pts.

Discrete RV.

2 Coin Toss : x (No of Heads) =

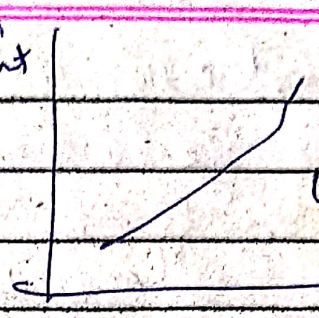
0	1	2
$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Date: Page:

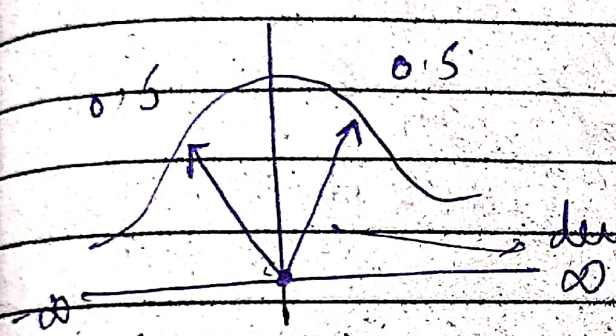
$r \in [-1, 1]$
correlation coefficient

x	y
2	1
5	3
7	9
⋮	⋮

+ve relation
$x \uparrow$ $y \uparrow$



Linear Regression



Continuous RV
$\int_{-\infty}^{\infty} f(x)\, dx = 1$

density Discrete RV

Symmetric
Normal

B.D. $^nC_r\, p^r q^{n-r}$
$p + q = 1$

Poisson D. $p \to 0$
$n \to \infty$

$y = mx + c$
Single variable

$y = m_1 x_1 + m_2 x_2 + \dots + m_n x_n + c$

Q) Three coin are tossed simultaneously
① find the P(getting atleast 2 Heads)
$\Rightarrow \frac{4}{8} = \frac{1}{2}$

{HHH THH
HHT THT
HTH TTH
HTT TTT}

② at most 2 heads
$\Rightarrow \frac{7}{8}$

Q) A bag contain 4 red balls, 6 yellow balls & 3

26.07.2023
Wednesday

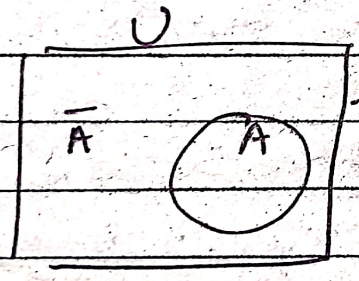
Date: Page:

balls are drawn randomly then what is probability that are of same color?

3

$$= \frac{{}^4C_3 + {}^6C_3}{{}^{10}C_3}$$

$$= 4 + \frac{\frac{6 \times 5 \times 4}{3 \times 2}}{\frac{10 \times 9 \times 8}{3 \times 2}} \qquad \frac{24}{10 \times 12} = \frac{1}{5}$$



→ can't define RV here

Qualitative outs

↓

Qualitative ← Toss winning captain will won the match

T F

n no. of matches how many times tossing winning won the match

n+1 instances $\{0, 1, 2, 3, \ldots, n\}$

Quantitative

Tossing a coin equally likely

P(batting if batting is preferred) $> \frac{1}{2}$

Date: Page:

Indian captain choosing batting after winning toss (good in bat)

→ collect historical data (observations)

Unconditional Prob. :- No. extra cond^n.
Calculated w.r.t. Universe

conditional Prob. :- $P(A|B) = \frac{P(A \cap B)}{P(B)}$ $\frac{P(A \cap B)}{P(B)}$

P(A) w.r.t. P(B)

Looking $A \cap B$ w.r.t. B

~~P(A)~~ $A \cap B$ w.r.t. U

① Prob.] uncond^n
② Prior Prob.]
③ Likelihood] cond^n
④ Posterior Prob.]

Some obs. are required (historical data)
Parameters
Observation needs to be done.

Toss winning captain will elect the bat first. → (I)
The stadium is Wankhedi stadium. → (II)
P(that (I) when the stadium (II))

Sample space: All people played in Wankhedi.
easy observation → Check how (I) satisfies in this & this observed.
tough to interrogate (We can interrogate)
P(that match is played on Wankhedi with that toss winning captain elect the bat first) → Collect all info and record

When given is less we can interrogate

Date: Page:

Q) It is given that in a population one in 50,000 has meningitis. 1 in 20 has stiff neck. How to compute no. of M. patients having stiff neck and no. of stiff neck patients having M.

$\rightarrow M$

$P(M) = \frac{1}{50000}$ $\quad P(N) = \frac{1}{20}$

$\cancel{P(M|N) = \frac{P(M \cap N)}{P(N)} = \frac{\frac{1}{50000} \times \frac{1}{20}}{\frac{1}{20}}}$ = Given

likelihood

$P(N|M) =$ M is small. If likelihood is $\frac{1}{2}$

$\rightarrow$ A given B.

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

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

Bayes formula

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

about one should be likelihood

$P(N|M) = \frac{1}{2}$ $\quad$ Neck given meningitis

$P(M|N) = \frac{P(N|M) \times P(M)}{P(N)}$ $\quad \frac{1}{2} \times \frac{1}{50000}$

$= \frac{1}{2} \times \frac{1}{50000}$

$= \frac{\frac{1}{100000}}{\frac{1}{20}} = \frac{1}{5000}$

Date: Page:

$P(A|B) \rightarrow$ Post erior

$P(B) = x$. $P(A) = ?$

highly populated.
(nearly equal to Uniform)

$A \rightarrow A_1$
$\rightarrow A_2$
$\vdots$
A_K

ME & CE

'R.V.

$P(A_i) = ?$

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

$P(A_i|B) > 0$ $\quad \sum P(A_i|B) = 1$

$$1 = \frac{P(B|A_1) * P(A_1)}{P(B)} + \frac{P(B|A_2) * P(A_2)}{P(B)} + \dots \frac{P(B|A_K) * P(A_K)}{P(B)}$$

$$\Rightarrow \boxed{\sum P(B|A_i) * P(A_i) = 1}$$

Total Probability

① ~~Every~~ all B should be covered by A.

#) Blood test

+ve Res $\frac{1}{2}$ $\quad$ -ve Res $\frac{1}{2}$

in general

on diseased person
$P(+ve\ res) \approx 1$

on not diseased person
$P(+ve\ res) \approx 0$

Date: Page:

Q) The Prob. of a person choosen at random having a disease is 0.05. The choosen person is administr-ated a diagonistic test which is not perfect. The prob. that the diagonostic test gives a positive result when the disease is present is 0.7 & a negative result when disease is not present is 0.95. Find out the P(person) having the disease when the diagnostic test gives a +ve result.

Sol:

$P(D) = 0.05 \Rightarrow P(\bar{D}) = P(ND) = 1 - 0.05 = 0.95$

~~P(+ve) = 0.7 P(~~

$P(+ve \mid D) = P(P \mid D) = 0.7$ 70% accuracy

$P(N \mid ND) = 0.95$ 95% chances (better)

$P(P \mid ND)$

$P(P)$ $P(N)$ → ~~dy~~ unknowns

$$P(D \mid P) = \frac{P(P \mid D) \times P(D)}{P(P)}$$

$P(P)$ → Total Probability → P → D, ND

$$P(P) = P(P/D) \times P(D) + P(P/ND) \times P(ND)$$

$$= \cancel{0.05 \times 0.0} \; 0.7 \times 0.05 + (0.05) \times 0.95$$

$$= 0.0825$$

$$P(D \mid P) = \frac{0.7 \times 0.05}{0.0825} = 0.4242$$

Date: Page:

Sol:

100%

30%	30%	40%
P_1	P_2	P_3
Defective rate 2%	3%	4%

~~4% → Defective rate~~

$P(P_2|D)$

$P = P(P_2 \cap D)$

$$P(P_2|D) = \frac{P(P_2 \cap D)}{P(D)} = \frac{0.02 \times 0.3}{0.02 \times 0.3 + 0.03 \times 0.3 + 0.04 \times 0.4}$$

Total probability

$$= \frac{3 \times 3}{\underset{6}{2 \times 3} + \underset{9}{3 \times 3} + \underset{16}{4 \times 4}} = \frac{9}{31}$$

$$x^*|c = \underset{(y_i)}{\operatorname{argmax}} \left[P(x_i/c) \right] \forall \text{ values of } x$$

(Always Posterior)

$$= \underset{(y_i)}{\operatorname{argmax}} \left[\frac{P(c/x_i) \times P(x_i)}{P(c)} \right] \forall \text{ values of } x$$

Date: Page:

BINOMIAL COEFFICIENT

Two events successes & failures.

Out of n trials (n+1) successes are possible.

Instance (0, 1, 2, 3, ..., n-1)

BD is used to describe the probability distribution of the no. of successes in n independent trial of some experiment.

$$\boxed{P(X=x) = {}^nC_x (p)^x (1-p)^{n-x}}$$

$$\boxed{\begin{array}{l} \text{mean} = np \\ \text{Variance} = npq \end{array}}$$

POISSON

$$\frac{e^{-\lambda}\lambda^x}{x!}$$

Q) In a particular minority 40% students smoke, 6 minority students are selected to know their views on smoking. Find Prob. that

(i) None of those selected smoke

(ii) All those selected smoke

(iii) Atleast half of the 6 smoke.

Sol: ① Identify the success

smoking

T

40% S — 60% NS

② $p = 0.4$ $q = 0.6$

③ $n = 6$

(i) ${}^nC_x p^x q^{n-x} = 1 \times (0.6)^6$

$= 0.046656$