Unit → 2

Probability & Statistics

Random Experiment: → More than one outcome.
→ No prediction.

For ex → Tossing a coin is a Random Experiment.

Sample Space: → Set of All possible outcomes.

Event: A Subset of sample space is called event.

Types of Event:-

1. Impossible:- $A = \phi$
2. Sure:- $A =$ Sample Space
3. Simple Event:- Event E is simple ~~of~~ if it has one sample point.
4. Compound Event: Atleast two points.

Ex → $S = \{HT, TH, HH, TT\}$.

$A =$ Atleast one head.

$A = \{HT, TH, HH\}$.

Algebra of Events :-

1> Complementary Event :- $A \to$ Event

$$A' = \{x : x \notin A ; x \in S\}.$$

2> Event A or B :-

$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$

3> Event A and B :-

$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$

4> Event A but not B.

$$A - B = \{x \in A \ \& \ x \notin B\}.$$

5> Mutually Exclusive Events :-

$$A \cap B = \phi$$

6> Exhaustive Events :-

$$A \cup B = S$$

Conditional Probability :→ Let S be the sample space of any Random experiment and E & F are two associated events with S. Then the conditional probability of event E if F has occured is denoted by P(E/F) and given by

$$P(E/F) = \frac{P(E \cap F)}{P(F)} \quad \Big| \quad P(F/E) = \frac{P(F \cap E)}{P(E)}$$

Example:- A coins is tossed. Write the probability if atleast two heads appear last coin has tail.

S

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

E : At least two heads appears.
F : Tail at last coint

$E = \{HHH, HHT, HTH, THH\}$,

$F = \{HHT, HTT, TTT, THT\}$.

$E \cap F = \{HHT\}$

$$P(E/F) = \frac{P(E \cap F)}{P(F)} = \frac{\frac{1}{8}}{\frac{4}{8}} = \frac{1}{4} \text{ Ans.}$$

Properties of Conditional Probability :-

1) If E & F are two event & S is sample space then –

$$\boxed{P(S/F) = P(E/E) = 1}$$

2) A, B and F are three events & S is sample space.

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

$$P(F) \neq 0$$

Proof

$$P((A \cup B)/F) = \frac{P((A \cup B) \cap F)}{P(F)}$$

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

$$= \frac{P(A \cap F) + P(B \cap F) - P(A \cap F \cap B \cap F)}{P(F)}$$

$$= \frac{P(A \cap F)}{P(F)} + \frac{P(B \cap F)}{P(F)} - \frac{P((A \cap B) \cap F)}{P(F)}$$

$$\boxed{P((A \cup B)/F) = P(A/F) + P(B/F) - P(A \cap B/F)}$$

(iii) $P(E'/F) = 1 - P(E/F)$.

Proof:- We know that

$$P(S/F) = 1$$

$$S = E \cup E'$$

$$\Rightarrow P((E \cup E')/F) = 1$$

$$1 = P(E/F) + P(E'/F) - \cancel{P(S\cap B)}\, P(E \cap E'/F)$$

But E & E' are mutually exclusive events

$$\therefore \quad P(E \cap E') = \phi$$

So $\quad P(E/F) + P(E'/F) = 1$

$$P(E'/F) = 1 - P(E/F).$$

Probability of Independent Events :→

Let A and B are two events such that

$P(A \cap B) = P(A) \cdot P(B)$ then A and B are called independent events.

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

Note:→ If A & B are independent then

A' and B, A and B', A' and B' are also independent.

Ques:→ Two cards are drawn at Random without replacement from a pack of 52 cards. Find the probability that both the cards are black.

let A: First card is black.
B: Second card is black.

$\therefore P(A) = \frac{26}{52} = \frac{1}{2}$

Since card is not replaced so Total cards = 51, black = 25

$\therefore P(B) = \frac{25}{51}$

Since A and B are independent events

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

$= \cancel{26}\ \frac{1}{2} \times \frac{25}{51} = \frac{25}{102}$

Date: / /
Page :

Ques:→

Ques 1 → A box oranges is inspected by the examine. three randomly selected drawn without replacement. If all three oranges are good. Box is aproved for sale. Find Probability of 15 Out of which 12 are good and 3 are bad then box is aproved for sale.

Ans: ~~let~~

let Event

A: First orange is good.

B: Second Orange is good.

C: Third orange is good.

$$P(A) = \frac{12}{15} = \frac{4}{5}$$

$$P(B) = \frac{11}{14}$$

$$P(C) = \frac{10}{13}$$

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

$$\frac{4}{\cancel{5}} \times \frac{11}{\cancel{14}_7} \times \frac{\cancel{10}^{2}}{13} = \frac{44}{91}.$$

Partition of Sample Space A set of events $E_1, E_2, E_3 ---- E_n$ is said to be partition of sample space S if

1) $E_i \cap E_j = \phi$ (for all $i \& j = 1, 2 --- n$).

2) $E_1 \cup E_2 \cup E_3 --- E_n = S$

3) $P(E_i) > 0$ for $i = 1, 2, 3 --- n$.

Theorem of Total Probability

Let $E_1, E_2, ---- E_n$ be a partition of sample space S and for all $i = 1, 2, 3 --- n$ $P(E_i) > 0$. Let A be any event associated with S then.

$$P(A) = P(E_1) \cdot P(A/E_1) + P(E_2) \cdot P(A/E_2) + ---$$
$$--- P(E_n) \cdot P(A/E_n)$$

$$= \sum_{j=1}^{n} P(E_j) \cdot P(A/E_j)$$

Proof:-> Since $E_1, E_2 --- E_n$ is a partition

$\Rightarrow E_i \cap E_j = \phi$ and $E_1 \cup E_2 \cup --- E_n = S$.

We know that

$$A = A \cap S$$
$$= A \cap (E_1 \cup E_2 -- E_n)$$

Date: / /

Page :

$$A = (A \cap E_1) \cup (A \cap E_2) \cup \cdots \cup (A \cap E_n)$$

Since eact $E_i \cap E_j = \phi \Rightarrow E_i \cap A \neq E_j \cap A$.

$$\Rightarrow (E_i \cap A) \cap (E_j \cap A) = \phi$$

~~$P(A) = P(E_1 \cap A) + P(E_2 \cap A$~~

$$P(A) = P(A \cap E_1) + P(A \cap E_2) + P(A \cap E_2) + \cdots$$
$$\cdots P(A \cap E_n).$$

We know, $P(A/E_1) = \dfrac{P(A \cap E_1)}{P(E_1)}$.

$$\Rightarrow P(A \cap E_1) = P(A/E_1) \cdot P(E_1).$$

$$\boxed{P(A) = P(E_1) \cdot P(A/E_1) + P(E_2) \cdot P(A/E_2) \cdots \cdots P(E_n) \cdot P(A/E_n).}$$

Ques A person has taken under construction job. The probabilities are 0.65 that there will be strike, 0.80 that the construction job will be completed on time if there is no strike. 0.32 that the construction job will be completed on time if there is strike. Determine the probability that the construction job will be completed on time.

Date: / /

Page :

$P(S) = 0.65$

$P(S') = \cancel{0}\ 1 - 0.65 = 0.35$

$P(\cancel{A}\ C \cap S') = 0.80$

$P(C \cap S) = 0.32$

$P(C/S) = 0.32$

$P(S/S') = 0.80$

$P(C) = P(S) \cdot P(C/S) + P(S') \cdot P(C/S')$

$= 0.65 \times 0.32 + 0.35 \times 0.80$

$= 0.208 + 0.28$

$= 0.488$

Baye's Theorem :→ Let $E_1, E_2, \ldots E_n$ be the partition of sample space S i.e. $E_i \cap E_j = \phi$ for $(i, j = 1, 2 \ldots n)$ and $S = E_1 \cup E_2 \cup E_3 \ldots \cup E_n$, $P(E_i) > 0$ and A is any event of non-zero probability, then

$$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{P(E_1) \cdot P(A/E_1) + P(E_2) \cdot P(A/E_2) + \ldots + P(E_n) \cdot P(A/E_n)} \quad \left(\begin{array}{c}\text{where} \\ i = 1, 2 \ldots n\end{array}\right)$$

Proof :→ By condition probability

$$P(E_i/A) = \cancel{P(E_i A)} \frac{P(E_i \cap A)}{P(A)} = \frac{P(A \cap E_i)}{P(A)} \quad \text{①}$$

Now $P(A|E_i) = \dfrac{P(A \cap E_i)}{P(E_i)}$

$$P(A \cap E_i) = P(E_i)\, P(A/E_i)$$

$$\cancel{P(A \cap E_i)}\ P(E_i/A) = \frac{P(E_i) \cdot P(A|E_i)}{P(A)} \quad \text{②}$$

from total probability theorem

$$P(A) = P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A/E_2) \cdots P(E_n) \cdot P(A/E_n).$$

∴ eq ② becomes.

$$P(E_i/A) = \frac{P(E_i) \cdot P(A|E_i)}{P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A|E_2) + \cdots P(E_n) \cdot P(A/E_n)}$$

Date: / /

Page :

Ques 1:- Bag B_1 contains 3 Red and 4 Black balls. Another bag B_2 contains 5 Red and 6 Black balls. A ball is drawn at random from one of the bag and it is found to be red. Find the probability that It has draw from bag B2.

$$P(B_2/R) = ?$$

R : Ball is Red.

~~Bl Ball is~~

B_1 : Draw from bag 1.

B_2 : Draw from bag 2.

$$P(B_1) = \frac{1}{2}, \quad P(B_2) = \frac{1}{2}.$$

$$P(R|B_1) = \frac{3}{7}, \quad P(R|B_2) = \frac{5}{11}$$

By Baye's Theorem :-

$$P(B_2/R) = \frac{P(B_2) \cdot P(R/B_2)}{P(B_1) P(R/B_1) + P(B_2) \cdot P(R/B_2)}$$

$$= \frac{\frac{1}{2} \cdot \frac{5}{11}}{\frac{1}{2} \cdot \frac{3}{7} + \frac{1}{2} \times \frac{5}{11}}$$

$$= \frac{\frac{5}{22}}{\frac{3}{14} + \frac{5}{22}} = \frac{\frac{5}{22}}{\frac{33+35}{154}}$$

2 | 14, 22
7 | 7, 11
11 |

77
×2
154

$\Rightarrow \dfrac{\cancel{22}^{5}}{\cancel{154}_{7}}\cdot\dfrac{\cancel{68}}{68} = \dfrac{35}{68}$

Random Variable :→

It is a real valued function whose domain is Sample space i.e. if X is random Variable and S is sample space then

$X : S \rightarrow R\ ,\ X(x_i) = w_i$ where $x_i \in S$ & $w_i \in R$

Ex:→ Toss 3-coins,
Sample space: {HHH, HHT, HTH, THH, TTH, THT, HTT, TTT}

Some Questions:-

Ques:→ What is the probability that a leap year selected at random contains 53 Sundays.

A non-leap year has 365 days.

$\Rightarrow$ 7)365(52
35
15
14
1

$\Rightarrow$ 52 weeks & 1 day extra.

$\Rightarrow$ 366 days has 52 weeks & 2 days extra sunday may be come 2 time in diff pairs like-
S-M, M-T - - - S-S $P(\text{sunday}) = \frac{2}{7}$ &

Date: / /
Page :

Ques:→ If the letters of the word "Recommend" are arranged at random, what is the chance that there will be 4 letters b/w R and D.

Total letters 9.

Total outcomes − $^9C_2 = \frac{9 \times \cancel{8}^4}{\cancel{2} \times 1} = 36$

1) R b b b b D.
R b b b b D
R b b b b D
R b b b b D

Arrangemets are 2×4
↓
as R & D can change their position.

$P = \frac{2 \times 4}{36} = \frac{2}{9}$ Ans

Ques:→ Given a set of 100 +ve numbers from 1 to 100. What is the probability that a no. selected randomly is divisible by 3 or 5.

E_1 : No. divisble by 3.
3, 6, 9, 15, 18 --- 99.

$99 = 3 + (n-1)3$

$96 = (n-1)3$

$32 = n-1$

$\boxed{n = 33}$

$P(E_1) = \frac{33}{100}$.

E_2: No. divisble by 5.

$5, 10, 15 - - - 100.$

$$100 = 5 + (n-1)5$$

$$\frac{\cancel{95}^{19}}{\cancel{5}} = (n-1)$$

$$\boxed{n = 20}$$

$$P(E_2) = \frac{20}{100}$$

$$E_1 \cap E_2 = \{15, 30, 45 - - - 90\}$$

$$90 = 15 + (n-1)15$$

$$\frac{90-15}{15} = n-1$$

$$\frac{\cancel{75}^{5}}{\cancel{15}} = n-1$$

$$n = 6$$

$$P(E_1 \cap E_2) = \frac{6}{100}$$

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

$$= \frac{33}{100} + \frac{20}{100} - \frac{6}{100}$$

$$= \frac{53}{100} - \frac{6}{100}$$

$$= \frac{47}{100}$$

Ques:→ A computer manufacturer known from experience that order will be ready for shipment with probability = 0.80. and the order will be ready for shipment on time and delivered on time = 0.72. find the probability that order is delivered on time, given that it was ready for shipment.

R: Ready for shipment
D: Delivered on time

$$P(D/R) = \frac{P(D \cap R)}{P(R)}$$

$$= \frac{0.72}{0.80} = 0.9$$

Given $P(R) = 0.80$

$P(R \cap D) = \cancel{0.80}\ 0.72$

Ques:→ A Box has 1000 LED's of which 50 are defected if 3 LED's are selected at random, one after the other without replacement. What is the probability that all are defected.

$E_1 \rightarrow$ 1 LED's is defected
$E_2 \rightarrow$ 2nd " " "
$E_3 \rightarrow$ 3rd " " "

$$P(E_1) = \frac{50}{1000}$$

$$P(E_2) = \frac{49}{999}$$

$$P(E_2) = \frac{48}{998}$$

$$P(E_1 \cap E_2 \cap E_3) = \frac{\cancel{50}}{\cancel{1000}} \times \frac{49}{999} \times \frac{48}{998}$$

$$=$$

$$P(E_1 \cap E_2 \cap E_3) = \frac{1}{10} \times \frac{49}{999} \times \frac{24}{998}$$

Ques :→ A motor company has 3 plants, Plant I, Plant II, Plant III. Plant - I produces 35% of O/P. Plant II produces 20% of O/P & plant - III produces 45%. 1% of plant I O/P is defective. - - - - - -

	Car produces	Defective Output Cars
plant I	35%	1%
plant II	20%	1.8%
plant III	45%	2%

Annual O/P = 1000000 Cars.

A car is selected at random from annual output & it is found defective. What is the probability that it has come from plant - I.

E_1 : Plant I production.
E_2 : plant II production
E_3 : plant III production.
D : Defective cars.

$$P(E_1/D) = \frac{P(E_1) \cdot P(D/E_1)}{P(E_1) \cdot P(D/E_1) + P(E_2) \cdot P(D/E_2) + P(E_3) \cdot P(D/E_3)}$$

$P(E_1) = \frac{35}{100}$	$P(D/E_1) = \frac{1}{100}$
$P(E_2) = \frac{20}{100}$	$P(D/E_2) = \frac{1.8}{100}$
$P(E_3) = \frac{45}{100}$	$P(D/E_2) = \frac{2}{100}$

$$P(E_1/D) = \frac{\frac{35}{100} \times \frac{1}{100}}{\frac{35}{100} \times \frac{1}{100} + \frac{20}{100} \times \frac{1.8}{100} + \frac{45}{100} \times \frac{2}{100}}$$

$$= \frac{\frac{35}{\cancel{10000}}}{\frac{35 + 36 + 90}{\cancel{10000}}}$$

$$= \frac{35}{161} \text{ Ans}$$

Ques:→ Resident of a community are examined for cancer. The examine results are classified as +ve if maligrancy is there and -ve if there are no indication of maligrancy. If a person has cancer. The probability of Suspected malignancy is 0.98 and the probability of reporting cancer where none existed is 0.15. If 5% of the community has cancer, Then what is the probability of a person not having cancer, when exam is positive.

E_1: Person has cancer.
E_2: Person do not has cancer.
P: Result is +ve.

$P(E_1) = 5\% = \frac{5}{100} = 0.05$

$P(E_2) = \frac{95}{100} = 0.95$

~~P(P/E₁) = 0.98~~. $P(P/E_1) = 0.98$

$P(E_2) \cdot P(P/E_2) = 0.15$

$$P(E_1/P) = \frac{P(E_1)\, P(P/E_1)}{P(E_1) \cdot P(P/E_1) + P(E_2) \cdot P(P/E_2)}$$

$$= \frac{0.05 \times 0.98}{0.05 \times 0.98 + 0.95 \times 0.15}$$

$$= \frac{0.1425}{0.1915} = 0.744.$$

Scanned by CamScanner

Statistics:-

The science which deals with the Collection analysis and interpretation of the numerical data.

Measure of frequency- Count, frequency, percent
Measure of Central tendency- Mean, Median, Mode
Measure of dispersion or Variation - Variance, S.D, M.D about Mean
Measure of Position location - Quartile, Percentile, decile
Measure of shape of distribution - Skewness, Kurtosis
Measure of depence - Correlation Coefficient.

<u>Limitation of Statistics</u>:-

1) Not used for qualatative data.
2) Does not study individuals
3) Not exact
4) Can be measured

<u>Uses</u>:- Bussiness, Economics, Astronomy, medical Science, Psychology, Education.

For Ungrouped Data

Mean = $\frac{\Sigma x_i}{n} = \frac{x_1 + x_2 - - x_n}{n}.$

Median :- first of all arrange in ascending / deseding order

when $n \rightarrow$ even

$$\frac{\left(\frac{n}{2}\right)^{th} + \left(\frac{n}{2}+1\right)^{th}}{2}$$

$n \rightarrow$ odd $\left(\frac{n+1}{2}\right)^{th}$ term

Mode :- Highest frequency data.

Range → Maximum value - Minimum Value.

Mean Deviation :→ $\frac{\text{Sum of absolute values of deviation from a}}{\text{no. of observation}}$

$$M.D(\bar{x}) = \frac{\Sigma(x_i - \bar{x})}{n}$$

$$M.D(M) = \frac{\Sigma(x_i - M)}{n}$$

Date: / /

For Grouped Data

Mean $\Rightarrow$ $\dfrac{\Sigma f_i x_i}{\Sigma f_i}$

Mean $= a + \dfrac{\Sigma f_i d_i}{N} \times h$

Median $\Rightarrow$ find C.f.
~~thera~~ check $\rightarrow n$

Median $= l + \dfrac{\frac{N}{2} - C}{f} \times h$

$N \rightarrow$ total frequency

$C \rightarrow$ upper c.f.

$h \rightarrow$ class size.

$f \rightarrow$ frequency of (n th term in case Median).

$l \rightarrow$ lower limit

Variance :- $\sigma^2 = \dfrac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$

Standard Median :-

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2}$$

Coefficient of Variation :-

$$C.V. = \frac{\sigma}{\bar{x}} \times 100.$$

Note :- Series having greater C.V. is more variable

Quartile →

$$Q_i = \left(\frac{iN}{4}\right)^{th} \text{ item.}$$

$$Q_1 = \left(\frac{N}{4}\right)^{th} \text{ item}$$

$$Q_2 = \left(\frac{\cancel{2}N}{\cancel{4}2}\right)^{th} \text{ item}$$

$$Q_3 = \left(\frac{3N}{4}\right)^{th} \text{ item}$$

Percentile →

$$P_i = \left(\frac{iN}{100}\right)^{th} \text{ item.}$$

$$P_{25} = Q_1 \;,\; P_{50} = Q_2 \;,\; P_{75} = Q_3$$

Decile → $\left(\frac{iN}{10}\right)^{th}$

Grouped data :→

$$Q_i = l.b + \left[\frac{\frac{iN}{4} - c.f_b}{f_{Q_i}}\right) \times h$$

$l.b$ = lower boundry of quartile class.

N = total frequency.

$c.f_b$ = c.f. before the quartile class.

f_{Q_i} = frequency of the quartile class.

h = size of the class interval

i = i^{th} quartile.

Quartile class → $\left(\frac{iN}{4}\right)^{th}$ class interval

$$P_i = l.b + \frac{\left(\frac{iN}{100} - c.f_b\right) \times h}{f_{Q_i}}$$

Ques:- find Quartile and Percentile :-

x :	0	1	2	3	4	5	6	7	8
$f(x)$:	1	9	26	59	72	52	29	7	1
c.f :	1	10	36	95	167	219	248	255	256.

Quartile:-

$$Q_1 = \frac{1\times N}{4} = \frac{1\times 256}{4} = 64^{th} \text{ item} = 3$$

$$Q_2 = \frac{2\times N}{4} = \frac{256}{2} = 128^{th} \text{ item} = 4$$

$$Q_3 = \frac{3\times N}{4} = \frac{3\times 256}{4} = 3\times 64 = 192^{th} \text{ item} = 5$$

Percentile :-

$$P_{25} = \frac{25\times N}{100} = \frac{25\times 256}{100} = 64^{th} \text{ item} = 3 = Q_1$$

$$P_{50} = \frac{50\times N}{100} = \frac{256}{2} = 128^{th} \text{ item} = 4 = Q_2$$

$$P_{75} = \frac{75\times N}{100} = \frac{3\times 256}{4} = 192^{th} \text{ item} = 5 = Q_3$$

Hence,

$P_{25} = Q_1$

$P_{50} = Q_2$

$P_{75} = Q_3$.

Date: / /

Page :

Ques1:- Find Quartile & Percentile :→ Find Median.

Marks - Group	frequency	C.f	
0-10	4	4	
10-20	8	12	
20-30	11	23	
30-40	15	38	
40-50	12	50	
50-60	6	56	
60-70	3	59	

$Q_1 = \frac{iN}{4} = \frac{N}{4} = \frac{59}{4} = 15^{th}$ item

C.f → 23

$l = 20$

$f_{qi} = 11$

$C.f_b = 12$

$h = 10$

$$Q_1 = l + \frac{\left(\frac{iN}{4} - Cf_b\right) \times h}{f_{qi}}$$

$$= 20 + \frac{\left[\frac{59}{4} - 12\right] \times 10}{11}$$

$$= 20 + \left[\frac{59-48}{4 \times 11}\right] \times 10$$

$$= 20 + \frac{\cancel{11}}{\cancel{11}} \times \frac{10}{4}$$

$$= 20 + 2.5$$

$$= 22.5$$

$= Q_1$

Q_3

Date: / /

Page :

Q_2 or ~~Q_2~~ P_{50} is the median

for $P_{50} = \cancel{\frac{50N}{100}} = \cancel{\frac{2N}{4}} = \frac{N}{2} = \frac{59}{2}$

$= 30^{th}$ item approx.

$$P_{50} = l + \frac{\left(\frac{iN}{100} - c.f.\right)}{f} \times h$$

Here $l = 30$

$f = 15$

$Cf = 283$

$h = 10$

$$P_{50} = 30 + \frac{\left[\frac{59}{2} - 283\right]}{15} \times 10$$

$$= 30 + \frac{[59 - \cancel{48}]}{15 \times 2} \times 10$$

$$= 30 + \frac{13}{\cancel{15}_{3}} \times \cancel{10}^{2}$$

$$= 30 + \cancel{\frac{26}{3}} \quad \frac{\cancel{8.66}}{2} \quad 4.33$$

$= \cancel{38}$ $34.33 \Rightarrow$ Median.

$Q_2 = 34.33$.

Scanned by CamScanner

Mode In discrete Grouped data → Heighest frequency data.
In continous Grouped data.

$$l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h = l + \frac{f_1 - f_0}{(f_1 - f_0) - (f_2 - f_1)} \times h$$

approx.

1. first, Identify the heighest frequency
2. Class corresponding to " " = modal class.
3. l → lowest limit of M.C.
4. h → interval
f_1 → freqn of Modal class.
f_0 → " before " "
f_2 → " after " ".

Ques:→

Marks	Number of Students
0 - 10	4
10 - 20	10
20 - 30	25
30 - 40	15
40 - 50	10
50 - 60	6.

H.f. = 25
Modal class = 20-30
$l = 20$, $h = 10$, $f_1 = 25$, $f_0 = 10$, $f_2 = 15$

$$\text{Mode} = l + \frac{(f_1 - f_0)}{2f_1 - f_0 - f_2} \times h$$

$$= 20 + \frac{(25-10)}{2\times 25 - 10 - 15} \times 10$$

$$= 20 + \frac{15}{50-25} \times 10 = 20 + \frac{\cancel{150}^{6}}{\cancel{25}} = 26$$

Date: / /
Page:

Inter Percentile Range = Diff. b/w P_{25} & P_{75}

Moment :-

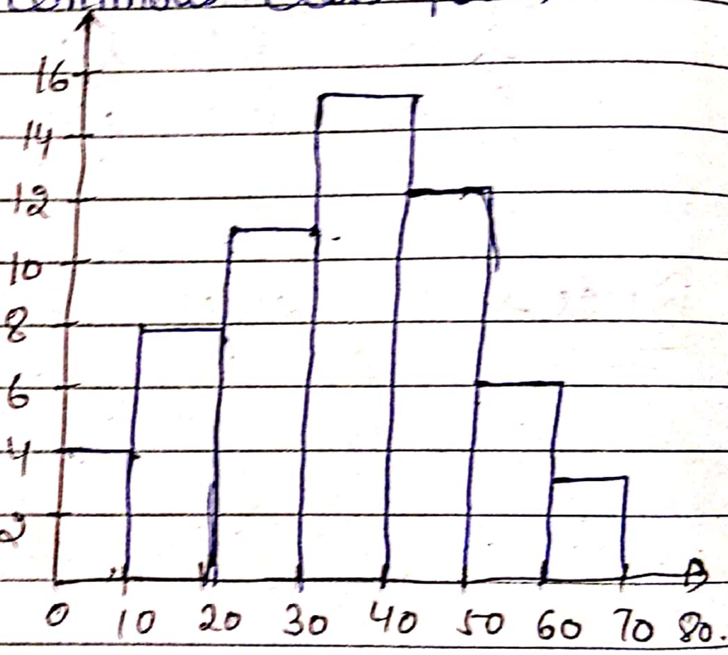
r^{th} moment about mean =

$$\mu_r = \frac{1}{N} \sum_{i=1}^{N} f_i (x_i - \bar{x})^r$$

Histogram :-

If data is given in continous class form like-

x	f
0 - 10	4
10 - 20	8
20 - 30	11
30 - 40	15
40 - 50	12
50 - 60	6
60 - 70	3

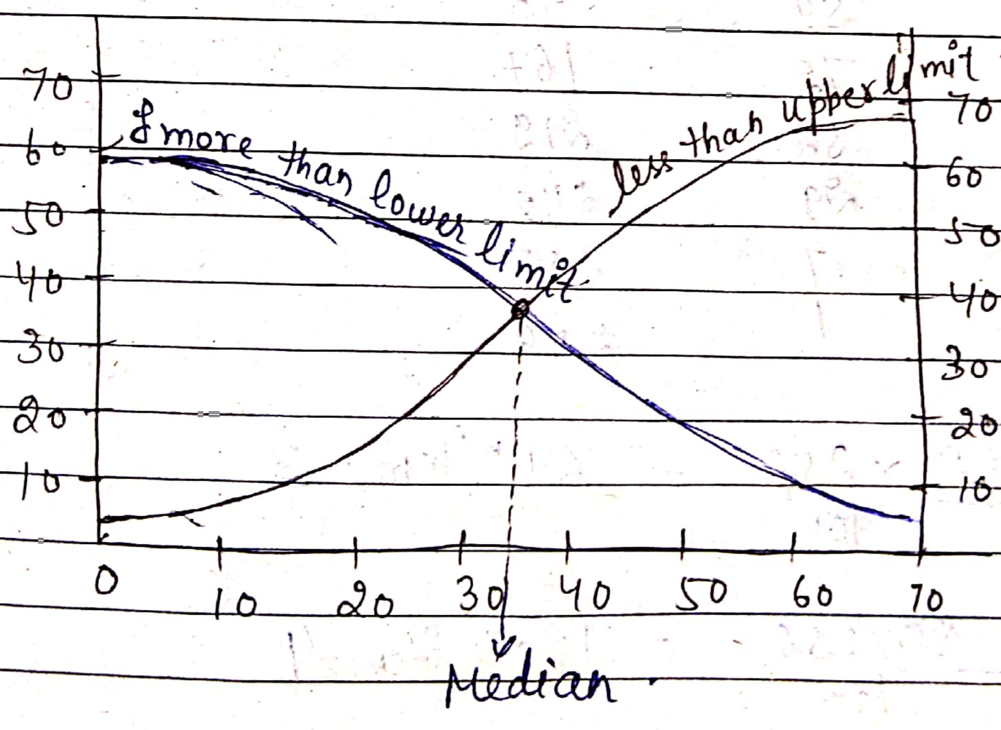


If we join mid points by using straight line it will form frequency polygon.
If we join by using free hand it will form frequency curve.

Ogive :→

If the plot of f less than upper limit vs f more than lower limit then it is called Ogive.

x	f	f less than upper limit	f more than ~~upper~~ lower limit
0-10	4	4	59
10-20	8	12	55 = (59-4)
20-30	11	23	47 = (55-8)
30-40	15	38	36 = (47-11)
40-50	12	50	21 = (36-15)
50-60	6	56	9 = (21-12)
60-70	3	59	6 = (9-3)



Box Plot:→ It is a rectangular box which gives graphical image of the concentration of the data. It shows how far are the extreme values from most of the data.

Box plots needs min value of data, max values, Ist Quartile (or 25^{th} percentile), 3^{rd} Quartile (or 75^{th} percentile).

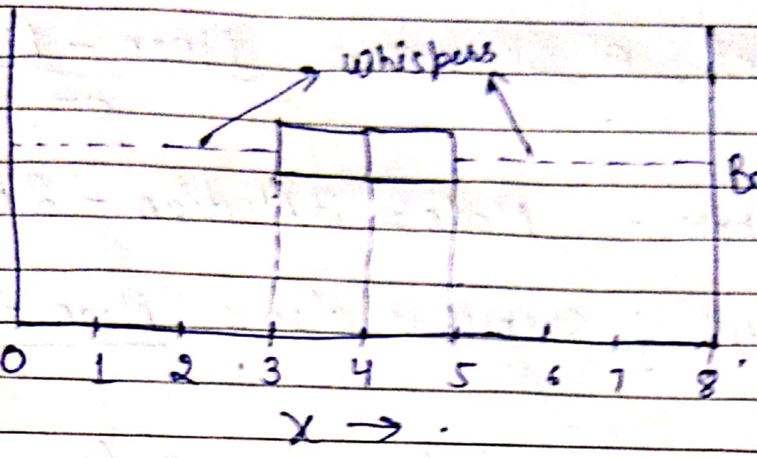
Percentiles are variable to summarize the tails (i.e. outer range of the distribution).

x	f	cf
0	1	1
1	9	10
2	26	36
3	59	95
4	72	167
5	52	219
6	29	248
7	7	255
8	1	256

$P_{25} = \frac{25 \times 256}{100} = 64^{th}$ item $= 3$

$P_{50} = \frac{50 \times 256}{100} = 128^{th}$ item $= 4$

$P_{75} = \frac{75 \times 256}{100} = 192^{th}$ item $= 5$.



Box-plot

Spread of percentile $P_{25} = 3-0 = 3$.
Spread of percentile $P_{50} = 4-3 = 1$
Spread of percentile $P_{75} = 5-4 = 1$

Skewness Means lack of Symmetry

A distribution is said to be skewed if

(i) Mean $\neq$ Median $\neq$ ~~Mode fall at~~
i.e. Mean, Median, Mode fall at different points.

(ii) Quartiles are not equidistant from median.
like in above example.
$P_{25} - P_{50} = 3 - 1 = 2$, $P_{50} - P_{75} = 1 - 1 = 0$
$\Rightarrow$ It is skewed.

(iii) Curve is stretched to one side.

Measure of Skewness.

1) S_k = Mean − Median
2) S_k = Mean − Mode
3) S_k = $(Q_3 - \text{Median}) - (\text{Median} - Q_1)$

Date: / /
Page:

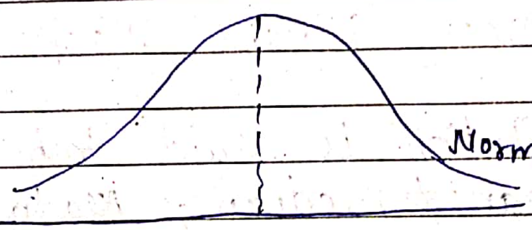
Coefficient of Skewness ⇒ $\dfrac{\text{Mean} - \text{Mode}}{\sigma}$

We know - Mode = 3 Median - 2 Mean

Coefficient of Skeness is also ⇒ $\dfrac{\text{Mean} - 3\,\text{Median} + 2\,\text{Mean}}{\sigma}$

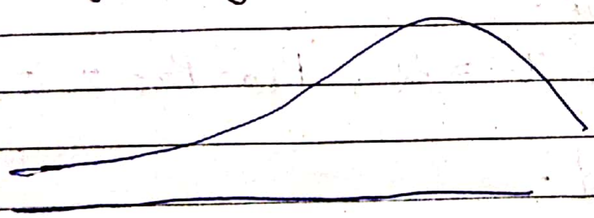
$$= \frac{3\,(\text{Mean} - \text{Median})}{\sigma}$$

Normal Curve :→



$M = M_d = M_o$

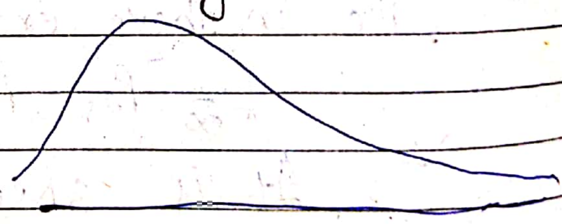
Negatively Skewed



$M < M_o$

$M < M_d$

Positively Skewed :



$M > M_o$

$M > M_d$

Kurtosis :→ Gives the convexity of the flatness of peakness of the Curve.

We find.

$$\beta_1 = \frac{\mu_3^2}{\mu_2^2}$$

$\mu_3 \rightarrow 3^{rd}$ Moment
$\mu_2 \rightarrow 2^{nd}$ Moment

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

$$\gamma_2 = \beta_2 - 3$$

① PLATYKURTIC Curve (more flat)

$\beta_2 < 3$

② Normal Curve $\beta_2 = 3$, $\gamma_2 = 0$.

③ LEPTOKURTIC Curve (More peaked), $\beta_2 > 3$