Measures of central tendency.

- Central tendency is a tendency of a given set of observations to converge around a single central value and the central value represents characteristics of the entire mass of observations.
- A measure of the central single value is known as major of central tendency or measure of central location or measure of average.

Importance of measure of central tendency

1. It gives represtative value or summary value of the entire mass of observations under the study.
2. It condenses the given data. It facilitates the comparision between two or more data. It is useful for further statical anyalsis.

Types

It is broadly classified into two types :-

1. Mathematical averages (Means)
2. Positional averages.

Mathematical averages
(a) Arithmetic mean ($\bar{X}$)
(b) Geometric mean (Gm)
(c) Harmonic mean (Hm)

Positional averages
(a) Median (Me)
(b) Mode (Mo)

Note :- There are 5 measures of central tendency

(a) Arithemetic mean.

A.M is obtained by sum of observations divided by total number of observations is

$$\overline{X} = \frac{x_1 + x_2 + x_3 + \dots X_n}{n}$$

i.e $\overline{X} = \frac{\Sigma X}{n}$

$\Sigma X = n.\overline{X}$

Calculation of arithmetic mean

(A) Individual series

(i) $\overline{X} = \frac{\Sigma X}{n}$ (Direct method)

(ii) $\overline{X} = A + \frac{\Sigma d}{n}$ (Shortcut method)

where, d = X − A : Deviations taken from A

A : Assumed mean

(iii) $\overline{X} = A + \frac{\Sigma d}{n} \times h$ [Step deviation method]

where $d' = \frac{X-A}{h}$

h : Highest common factor of X

or : Common difference of X

(B) Discret Series

(i) $\overline{X} = \frac{\Sigma fX}{N}$ (Direct method)

(iii) $\overline{X} = A + \frac{\Sigma fd}{N}$ (Short cut method)

where d = X − A : Deviation taken from A

A : Assumed mean

(C) Continous Series.

(iii) $\bar{x} = A + \frac{\Sigma fd'}{N} \times h$ [Step-deviation method]

where $d' = \frac{X-A}{n}$

h = class width / size.

X : Mid value

(i) $\overline{X} = \frac{\Sigma fX}{N}$ (Direct method)

(ii) $\bar{x} = A + \frac{\Sigma fd}{N}$ (Short cut method)

where, d = X − A : Deviations taken from A.

A = Assumed mean

Question no: 3 .

Let :- X = No of goal score per match

f = No of matches.

X	Tally Bars	f	fx
0	II	2	0
1	III	3	3
2	卌	5	10
3	卌 I	7	21
4	卌	5	20
5	II	2	10
6	I	1	6

$\frac{\Sigma fx}{N}$

$= \frac{70}{25}$

$= 2.8$

(9)

Wave group	Total hours	Average no. of hrs. worked per workers	Mid value	No of workers (f)	$d' = \frac{X-150}{20}$ ($\frac{x-A}{n}$)	fd'
80-100	168	12	90	$\frac{168}{12} = 14$	−3	−42
100-120	170	10	110	17	−2	−32
120-140	225	9	130	25	−1	−25
140-160	272	8.5	150	32	0	0
160-180	126	7	170	18	1	18
180-200	91	6.5	190	14	2	28

Here; A = 150, h = 20 N = $\Sigma fd'$ =

$\bar{X} = A + \frac{\Sigma fd'}{N} \times h$

=

Exceeding 60 but no exceeding 70 i.e 61-70 71-80

Date

Page No.

(10) (x) Bonos (Rs)

(x) Bonos (Rs)	Salary group	Tally Bars	No of employee (f)	$d' = \frac{x-30}{5}$	fd'
10	61-70	|||	3	-4	-12
15	71-80	||||	4	-3	-12
20	81-90	~~||||~~	5	-2	-10
25	91-100	~~||||~~ ||	7	-1	-7
30	101-110	|||	3	0	0
35	111-120	~~||||~~	5	1	5
40	121-130	||	2	2	4
45	131-140	||	2	3	6
50	141-150	|	1	4	4

Here, $A = 30$, $h = 5$, $N = 32$, $\Sigma fd' = -22$

Average bonus per employee is,

$$\bar{X} = A + \frac{\Sigma fd'}{N} \times h$$

$$= 30 + \frac{-22}{32} \times 5$$

$$= 26.56.$$

(11)

Miles travelled (X)	No of villagers (f)	$d' = \frac{x-10}{2}$	fd'
			-
2	38	-4	-152
4	104	-3	-312
6	x	-2	$-2x$
8	78	-1	-78
10	48	0	0
12	42	1	42
14	y	2	24
16	24	3	72
18	16	4	64
20	2	5	10

Let x & y be the missing frequency of value of 6 & 14 respectively

Date

Page No.

Given,

$\bar{x} = 78$, $N = 520$

then, $\Sigma f = 352 + x + y$

$\therefore \Sigma f = N$

or, $852 + x + y = 520$

$\therefore x + y = 168$ —— (i)

Also,

$\Sigma fd' =$

$A = 10$, $h = 2$, $N = 520$.

$$\bar{X} = A + \frac{\Sigma fd'}{N} \times h$$

or, $\cancel{\bar{x}}\ 7.8 = 10 + \frac{(-354 - 2x + 2y)}{520} \times 2$

or, $109 = x - y$ —— (ii)

From (i) & (ii), we get.

$x + y = 168$

$x - y = 109$

$2x = 277$

$x = 138.5$

From eq^n (i) we get

$x + y = 168$

$138.5 + y = 168$

$y = 29.5$

(12)

Date
Page No.

Properties of Arithemetic mean

- If all the observations assumed by variable are constant then their A·M is also the same constant.
  Example :- $x = 3, 3, 3, 3, 3$
  $$\bar{x} = \frac{3+3+3+3+3}{5} = 3$$
- A·M is affected due to change of origin as well as scales.
- If two variables x and y are lineraly related by $y = a + bx$ then we can write $\bar{y} = a + b\bar{x}$ where a and b are constant
  Example :- Find A·M of y if A·M of x is 10 for $-3x + 5y = 20$.
  For $-3x + 5y = 20$
  $-3\bar{x} + 5\bar{y} = 20$
  $-3 \times 10 + 5\bar{y} = 20$
  $-30 + 5\bar{y} = 20$
  $y = \frac{50}{5}$ 10
  $\bar{y} = 10$ #
- Sum of deviations of all observations is 0 if the deviations are taken from their respective A.M. i.e $\Sigma(X-a) = 0$ if $a = \bar{X}$.
- Sum of squares of deviations of all observations is minimum if the deviations are taken from their respective mean i.e
  i.e $\Sigma(x-a)^2$ is minimum if $a = \bar{X}$.
- The A·M of all observations in A.P is given by $\frac{a+l}{2}$. Ex. A·M of first 'n' natural no. is given by $\frac{n+1}{2}$.

Date
Page No.

- If n_1 and n_2 be the number of observations of two groups, $\bar{x}_1$ and $\bar{x}_2$ are their respective A.M's then combined mean of the two groups taken together is given by $\bar{X} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

Example :-

Note 1
(i) Salary of each employees is increased by 10% of their average salary.
(ii) Salary of each employees is increased by 10% of their salary.
Here :- Total number of employees = 500
Average salary = Rs 32,000.

(i) Change of origin occurs in the first condition and the origin factor is 3200. i.e (10% of 32,000 = 3200)
(ii) $x + 10\%$ of $x = x + 0.1x = 1.1x$
Change of scale occurs in the second condition and the scale factor is 1.1.

Note 2 If each observations is increased or decreased by a constant value then their A·M is also increased or decreased by the same constant value.

Note 3 If each observation is mutiplied or divided by a constant value than their A·M is also multiplied or divided by the same constant value.

Note 4 If $\Sigma(x-a) + \Sigma(X-b) = 0$ then $\bar{x} = \frac{a+b}{2}$
$\Sigma(x-a) + \Sigma(X-b) = 0$

Date
Page No.

Eg. Find A·M of X if $\Sigma(X-10)=50$
and $\Sigma(X-12)=-50$

Ans, $\bar{X} = \frac{10+12}{2} = 11$

Note 5 If $\Sigma(x-a)^2 = \Sigma(x-b)^2$, then; $\bar{X} = \frac{a+b}{2}$

Example :- Find A·M of X if $\Sigma(x-20)^2=100$
and $\Sigma(X+30)^2=100$

Ans :- $\bar{X} = \frac{20-30}{2} = -5$ #
($20+(-30)$)

Combined mean
- → Pooled mean
- → Overall mean
- → Grouped mean

Date
Page No.

Pg 158

14 Worth of share

Worth of share	Price of share (X)	No. of share (F)
1200	10	120
1200	12	100
1200	15	80
1200	20	60
1200	24	50
$\Sigma fX = 6000$		$N = 410$

Mean $= \frac{\Sigma fX}{N} = \frac{6000}{410} = 14.63$.

15. Given
Mean weight of $(\bar{X}) = 60$ kg
$n = 150$

For weight of which $\bar{X}_1 = 70$ kg
For wt. of girls $\bar{X}_2 = 55$ kg

Let n_1 and n_2 be the nomber of boys and girls respectively.
Such that $n_1 + n_2 = 150$ —— (i)

$\therefore \bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2}$

or, $60 = \frac{70n_1 + 55n_2}{150}$

or, $9,000 = 70n_1 + 55(150-n_1)$

or, $9,000 = 70n_1 + 8250 - 55n_1$

or, $9000 - 8250 = 70n_1 - 55n_1$

or, $750 = 15n_1$

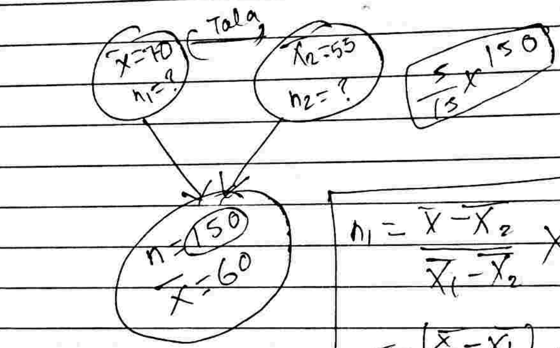
or, $n_1 = 50$

Again,

Putting value in eq (i)

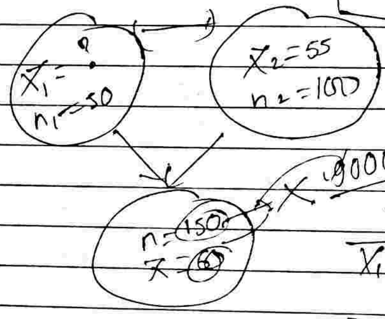
$n_2 = 150 - 50$

$= 100.$



$$n_1 = \frac{\bar{X}-\bar{X}_2}{\bar{X}_1-\bar{X}_2} \times n$$

$$n_2 = \frac{(\bar{X}-\bar{X}_1)}{(\bar{X}_1-\bar{X}_2)} \times n$$



$\times \frac{9000 - 5500}{50}$

$$\bar{X}_1 = \frac{n\bar{X} - n_2\bar{X}_2}{n - n_2}$$

$$\bar{X}_2 = \frac{n\bar{X} - n_1\bar{X}_1}{n - n_1}$$

(19) (i) Here, $n = 500$; $\bar{X} = Rs\ 200$

Wrong observations: (180, 20)

Correct observations: (80, 220)

$\Sigma X = n \cdot \bar{X}$

$= 500 \times 200$

$\therefore \Sigma X = 100,000$

$\Sigma X_C = 100000 - 180 - 20 + 80 + 220$

$= 100,100$

$$\therefore \bar{X}_C = \frac{\Sigma X_C}{X_C} = \frac{100100}{500} = 200.2$$

(iv) Weighted average

If a variable X assumes value X_1, $X_2, X_3, \ldots X_n$ and their respective weights $w_1, w_2, w_3, \ldots, w_n$ then weighted A.M is given by,

$$\bar{X}_w = \frac{w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n}{w_1 + w_2 + w_3 + \cdots + w_n}$$

i.e $\bar{X}_w = \frac{\Sigma wx}{\Sigma w}$

Q no 1

$$\frac{\Sigma fX}{N} = \frac{30+70+10+75+500+8+42+250+40+36}{10}$$

$= \frac{1061}{10}$

$= 106.1$ #

Q no 2

X	f	fX
50	1	50
60	3	180
70	5	350
80	7	560
90	6	540
100	2	200
110	1	110
	N = 25	fX = 1990

$$\therefore \bar{X} = \frac{\Sigma fX}{N} = \frac{1990}{25} = 79.6$$ #

$d' = \frac{X - A}{h}$ h = common factor, A → Assumed value.

Date
Page No.

Q no 4 we have,
Let missing frequency be n

X	F	Fn
10	3	30
12	7	84
14	n	$14n$
16	20	320
18	8	144
20	5	100
	$N = 43 + n$	$678 + 14n$

we have, $\bar{X} = 15.38$ so,

$$\bar{X} = \frac{\Sigma FX}{N}$$

$$15.38 = \frac{678 + 14n}{43 + n}$$

$$661.34 + 15.38n = 678 + 14n$$

$$\cancel{-}16.66 = \cancel{-}1.38n$$

$$12.07 = n$$

Q no 5

Vehicles	Mid point (x)	$d' = \frac{x-25}{10}$	F	Fd'
0-10	5	-2	3	-6
10-20	15	-1	14	-14
20-30	25	0	53	0
30-40	35	1	20	20
40-50	45	2	10	20
			$\Sigma f = N = 100$	$\Sigma fd = 20$

$$\text{Mean } (\bar{X}) = A + \frac{\Sigma fd'}{N} \times h$$

$$= 25 + \frac{20}{100} \times 10$$

$$= 27 \text{ vechicles}$$

Date
Page No.

6

Temp °C	Mid point (x)	$d' = \frac{x+5}{10}$	F	Fd'
-40-(-30)	-35	-3	10	-30
-30-(-20)	-25	-2	28	-56
-20 to -10	-15	-1	30	-30
-10 to 0	-5	0	42	0
0 to 10	5	1	65	65
10 to 20	15	2	180	360
20 to 30	25	3	10	30
			$\Sigma f = 365$	$\Sigma fd' = 339$

$$\text{Mean } (\bar{X}) = A + \frac{\Sigma fd'}{\Sigma f} \times h$$

$$= (-5) + \frac{339}{365} \times 10$$

$$= 4.29$$

12 (a)

Sales (x)	Mid value (x)	$d' = \frac{x-25}{10}$	F	Fd'
0-10	5	-2	5	-10
10-20	15	-1	25	-25
20-30	25	0	n	0
30-40	35	1	18	18
40-50	45	2	7	14
			$\Sigma f = 55 + n$	$\Sigma fd' = -3$

$$\text{Mean } (\bar{x}) = A + \frac{\Sigma fd'}{\Sigma f} \times h$$

$$24.625 = 25 + \frac{(-3)}{55+n} \times 10$$

$$24.625 = \frac{(1375 + 25n) + 80}{55 + n}$$

$$1354.375 + 24.625n = 1375 + 25n - 30$$

$$= 355 + 5n$$

~~999.375~~

1000.625 =

Date
Page No.

or, $9.375 = 0.375\,n$
or, $n = 25$ //

12(b)

CI	Mid Value(X)	$d' = \frac{x-35}{10}$	f	fd'
0-10	5	-3	5	-15
10-20	10	-2.5	10	-50
20-30	15	-2	n	-2n
30-40	35	0	30	0
40-50	45	1	y	y
50-60	55	2	10	20
			55+n+y	-45+2n+y

since total is 100

$55+n+y = 100$

$n = 45-y$

$\bar{x} = A + \frac{\Sigma fd}{f} \times h$

$33 = 15 + \frac{\cancel{100}\,90 + 3y}{55+n+y} \times 10$

$33 = \frac{15(55+n+y) + (\cancel{100}\,90 + 3y)\,10}{55+n+y}$

$33(55+n+y) = 15(55+n+y) + 900 + 30y$

$1815 + 33n + 33y = 825 + 15n + 15y + 900 + 30y$

$33n + 90 = \cancel{10} + 12y + 15n$

$33(45-y) = \cancel{10} + 12y + 15(45-y)$

$90 + 1485 - 33y = \cancel{10} + 12y + 675 - 15y$

$800 = 30y$

$\frac{80\cancel{0}}{3\cancel{0}} = y$

Date
Page No.

$33 = 15 + \frac{(-45+2n+y)}{10\cancel{0}} \times 1\cancel{0}$

$33 = 35 + \frac{(-45+2n+y)}{10}$

$33 = 350 + \frac{(-45+2n+y)}{10}$

$330 = 350 - 45 + 2n + y$

$25 = 2n + y$

$25 = 2(45-y) + y$

$25 = 90 - 2y + y$

12a Given Mean $(\bar{x}) = 24.625$

(a) Let, missing frequency be n

n	F	m	fm
0-10	5	5	25
10-20	25	15	375
-20	n	25	25n
	18	35	630
	7	45	315

Date
Page No.

12 (b) Here, We have given,
Mean of distribution ($\bar{x}$) is 33
Now, Let the missing frequency of CI 20-30 & 40-50
n & y resp. so,

C.I	F	Mid value	$d' = \frac{M-25}{10}$	Fd'
0-10	5	5	-2	-10
10-20	10	15	-1	-10
20-30	n	25	0	0
30-40	30	35	1	30
40-50	y	45	2	$2y$
50-60	10	35	3	30
	N=100			$\Sigma fd = 40+2y$

Now,
$\bar{X} = 33$
$A = 25$
$N = 100$ $\Sigma fd' = 40+2y$ $h = 10$
So, $\bar{x} = A + \frac{\Sigma fd}{N} \times h$
$33 = 25 + \frac{40+2y}{100} \times 10$
or, $3 = 25 + \frac{20+y}{5}$
or, $165 = 125 + 20 + y$
$\therefore y = 20$

And, Sum of frequency (N) = 100
$n + y + 55 = 100$
$n + 20 + 55 = 100$
$n = 25$ #.

Date
Page No.

(13) Soln:-
We have been given,
Mean weight of student ($\bar{x}$) = 119
Let weight of sixth student be a then,
We have,
$\bar{x} = \frac{\Sigma x}{N}$

or, $119 = \frac{115 + 109 + 129 + 117 + 114 + a}{6}$

or, $714 = 584 + a$
$\therefore a = 130$ lbs

(16) $\bar{X} = 45$ $N = 300$
$\bar{X}_1 = 70$ $\bar{X}_3 = 20$ $\bar{X}_2 = ?$
$n_1 = 100$ $n_3 = 100$ $n_2 = 100$

Now,
$\bar{x} = \frac{\bar{x}_1 n_1 + n\bar{x}_2 + n_3\bar{x}_3}{n + n_2 + n_3}$
or, $45 = \frac{70 \times 100 + \bar{X}_2 \times 100 + 20 \times 100}{300}$
or, $45 \times 300 = 7000 + 100\,\bar{X}_2 + 2000$
or, $100\,\bar{X}_2 = 4500$
or, $\bar{X}_2 = 45$ #.

$1^2+2^2+3^2+\cdots+n^2 = \frac{n(n+1)(2n+1)}{6}$

$1^3+2^3+3^3+\cdots+n^3 = \left[\frac{n(n+1)}{2}\right]^2$

$1+2+3+\cdots+n = \frac{n(n+1)}{2}$

Date
Page No.

Example :- Find weighted A.M of first 'n' natural numbers where the weights are their respective value.

Soln :-

X : 1, 2, 3, ---- n

W : 1, 2, 3, ------ n

$$\overline{X_W} = \frac{\sum WX}{\sum W}$$

$$= \frac{1^2+2^2+3^2+\cdots+n^2}{1+2+3+\cdots+n}$$

$$= \frac{\cancel{n}(n+1)(2n+1)}{3\cancel{6}\,\cancel{n}(n+1)}\cdot\cancel{2}$$

$$= \frac{2n+1}{3}$$

Note 1 :- If all the observations have not the same importance then we have to consider weighted average to get actual average.

Note 2 :- If all the observations have the same importance then simple average is sufficient to get actual average.

Note 3 :- Frequency is also treated as weight.

For a given set of observation median is the middle most value which divide total no. of observations into two equal parts. Since the given set of observations are arranged in either ascending or deconding order.

Date
Page No.

Example :-

(1) Median of 2, 8, 12, 15, 16

$M_e = 12$

Ex-2 :- Median of 8, 12, 14, 16, 17, 19 is given by

$$M_e = \frac{14+16}{2} = 15$$

Calculation of median.

Rank of median $= \frac{n+1}{2}$

(A) Individual series

~~Individ~~

Case I - if n is odd

M_e = A middle most value

OR

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

Case II - if n is even :-

M_e = a simple average of two middlemost values

$= \frac{1}{2}$ AM of $\left(\frac{n}{2}\right)^{th}$ and $\left(\frac{n}{2}+1\right)^{th}$ value

(B) Discrete series

Construct C.F distribution.

$M_e = \frac{N+1}{2}^{th}$ value

Example :- Find median of the following distribution

X:	10	20	30	40	50
F:	8	12	22	16	5

Soln :-

X	f	c.f
10	8	8
20	12	20
30	22	42
40	16	58
50	5	63

$N = 63$

Now, $Me = \left(\frac{N+1}{2}\right)^{th}$ term

$= \left(\frac{63+1}{2}\right)^{th}$ term

$= (32)^{th}$ ~~term~~ value

C.F just greater than 32 is 42 and it's respective value is 30.

∴ Median = 30.

ⓒ **Continuous series:-**

Construct c.f distribution; and median class.

$$\text{Median class} = \frac{N}{2}^{th} \text{ class}$$

$$\text{Median } (Me) = l + \frac{\left(\frac{N}{2} - c.f\right)}{f} \times h$$

(Interpolation formula)

Assumption:-

- It is applicable if classes of a frequency distribution are continuous with exclusive type, whatever maybe their class size.

Class interval

A. Exclusive type C.I or overlapping class

Example:- 10 - 20 → excludes here

Includes here ← 20 - 30

30 - 40

B. Inclusive type C.I or non-overlapping class

Ex 10 - 19

Included ← 20 - 29

30 - 39

L.C.L = 20

U.C.L = 29

Lower class boundary = 19.5

Upper C.B = 29.5

C. Open type C.I

below 20

20 - 30

30 - 40

above - 40 ↳ unspecified.

$$LCB = LCL \text{ (Lower class limit)} - \frac{d}{2}$$

$$UCB = UCL + \frac{d}{2}$$

29.5 - 4'

where,

d = Diff betn L.CL of a class and UCL of it's previous class.

Date
Page No.

(17) Men's salary — Women's salary — Total

$n_1\% = ?$ — $n_2\% = ?$ — $n_1 + n_2$

$\bar{X}_1 = Rs\ 620$ — $\bar{X}_2 = Rs\ 520$ — $\bar{X} = Rs\ 600$

We have,
$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2}$$
$$600 \cancel{30} = \frac{n_1 \times 620 + n_2 \times 520}{\cancel{600}\ n_1+n_2}$$
$$600n_1 + 600n_2 = 620n_1 + 520n_2$$
$$20n_1 = 80n_2$$
$$\frac{n_1}{n_2} = \frac{\cancel{80}\ 4}{\cancel{20}}$$
$$\frac{n_1}{n_2} = 4$$

Total ration = 4 + 1 = 5

$$n_1 = \frac{4}{5} \times 100\% = 80\%$$
$$n_2 = \frac{1}{5} \times 100\% = 20\%$$

Hence, the men & women are 80% and 20% resp.

(18) $\bar{X} = 100$

less than 60% — more than 60% — Total

$n_1 = 70$ — $n_2 = 30$ — $n_1 + n_2 = 100$

$\bar{X} = 50$

Marks	F	m	fm
0 - 20	16	10	160
20 - 40	24	30	720
40 - 60	30	50	~~800~~ 1500

$\Sigma fm = \cancel{9600}\ 2380$

So,
less than 60%.
$n_1 = 70$
$$\bar{X}_1 = \frac{\cancel{9600}}{\cancel{700}} = \frac{2380}{70}$$
$\cancel{=13.71}$. 34

$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2}$$
$$50 \cancel{13.71} = \frac{70 \times \cancel{13.71}\ 34 + n_2\bar{X}_2}{100}$$
$$5000 - 2380 = 30 \times \bar{X}_2$$
$$87.33 = \bar{X}_2 \ \#$$

(19)(i)

$N = 500$

$\bar{X} = 200$

$\Sigma X = \bar{X} \cdot N$

$= 500 \times 200$

$= 1,00,000$

$1,00,000 - 18 - 20 + 80 +$

220

$\cancel{-20}$

$= 100262$

$$\Sigma X_c = \frac{100262}{500}$$
$= 200.5 \#$

19(ii) Here, $n = 100$; $\bar{X} = \cancel{100}\ 65$

Wrong observation ($C\cancel{18}$ 40, 50).

$\Sigma X = n \cdot \bar{X}$

$= 100 \times 65$

$\Sigma X = 6500$

$\Sigma X_c = 6500 - 40 - 50$

$= 6410$

$$\therefore \bar{X}_c = \frac{\Sigma X_c}{X_c} = \frac{6410}{100} = 64.1 \ \#$$

Date
Page No.

Date
Page No.

(iii) $n = 80$ $\overline{x} = 40$

Wrong observations (54, 84)

$\Sigma X = n \times \overline{X}$

$= 80 \times 40$

$= 3,200$

$\Sigma X_c = 3200 - 54 - 84 = 3062$

$\Sigma \overline{X_c} = \frac{3200}{80} \frac{\Sigma X_c}{X_c}$

$= 38.275$ #

(iv) $X = 60, 75, 63, 59, 55$

$W = 1, 2, 1, 3, 3.$

$\overline{X}_w = \frac{\Sigma WX}{\Sigma W}$

$= \frac{60 \times 1 + 75 \times 2 + 63 \times 1 + 59 \times 3 + 55 \times 3}{1 + 2 + 1 + 3 + 3}$

$= \frac{615}{10}$

$= 61.5$ #

(20) $X : 36.00, 40.00, 44.00, 48.00$

$W : 14, 11, 9, 6$

$\overline{X}_w = \frac{\Sigma WX}{\Sigma W}$

$= \frac{36.00 \times 14 + 40.00 \times 11 + 44. \times 9 + 48 \times 6}{14 + 11 + 9 + 6}$

$= \frac{504 + 440 + 396 + 288}{40}$

Date
Page No.

$= \frac{1628}{40}$

$= 40.70$ #

(24)

Income 100 Rs	No of person	C.F
50-80	30	30
80-100	127	157
100-110	140	297
110-120	240	537
120-130	176	713
130-150	135	848
150-180	20	868
180-200	3	871
	N=871	

$N = 871$

Median class $= \frac{N}{2}$th class

$= \frac{871}{2}$th class

$= 435.5$

The C.F just greater than 435.5 is 537 and it's respe[ctive] value(class) is 110-120.

Here,

$l = 110$ $c.f = 297$ $\frac{N}{2} = 435.5$

$h = 10$ $f = 240$

median $= l + \frac{\left(\frac{N}{2} - c.f\right)}{f} \times h$

$= 110 + \left(\frac{435.5 - 297}{240}\right) \times 10$

$= 115.77$ #

23.

x	F	c.F
1	8	8
2	10	18
3	11	29
4	16	45
5	20	65
6	25	90
7	15	105
8	9	114
9	6	120

$N=120$

Median class = $\frac{N}{2}$th class = $\frac{120}{2} = 60$ just greater is 65 so class is 5.

Now,

$$Me = \left(\frac{N+1}{2}\right)$$

$$= \frac{120+1}{2}$$

$$= \frac{121}{2}$$

$= 60.5$ is 65 so it's 5.

$Me = 5.$

(22) ~~Simple average of two middle most value~~

(i) A middle most value since $n = $ odd (7) $Me = 85$.

(ii) Since even $n = 6$

Averge of two middle most value $\frac{80+90}{2}$

$Me = 85$

Properties of median

(i) If all the observations assumed by a variable are constant then their median is also the same constant. Eg:- $X = 12, 12, 12, 12, 12$ $Me = 12$

(ii) It is dependent on change of origin as well as scale.

(iii) If x and y are linearly related by $y = a + bx$ then Me of $y = a + b$ (Me of x)

i.e $y_{me} = a + b\, x_{me}$

* Find median of Y if median of X is 10 for

$2x - 5y = 20$

$\therefore 2x - 5y = -20$

or, $2x_{me} - 5y_{me} = -20$

or, $2 \times 10 - 5 y_{me} = -20$

or, $20 + 20 = 5y_{me}$

or, $40 = 5y_{me}$

$\therefore 8 = y_{me}$.

(iv) Sum of absolute deviations of all observations is minimum if the deviations are taken from their respective median i.e $\Sigma|X-a|$ is min if $a = Me$.

Partation values/Fractiles:-

For a given set of observations partation values are a number of values of the given observations which divide total number of observations into a number of equal parts. Since the given set of observations are arranged in ascending order only.

$750 \times \frac{50}{2}$ $300 \times \frac{50}{2}$

Date
Page No.

It is also termed as fractiles.
Some important partition values are :-

(A) Quartiles
(B) Deciles
(C) Percentiles

(A) Quartiles :-

Quartiles are three values which divide total number of observations into 4 equal parts. The three quartiles are :-

(i) First quartile or lower quartile (Q_1)
(ii) Second quartile or middle quartile (Q_2)
(iii) 3rd quartile or upper quartile (Q_3)

$\frac{n}{4}$ | $\frac{n}{4}$ | $\frac{n}{4}$ | $\frac{n}{4}$ $= n$
25% Q_1 25% Q_2 25% Q_3 25%

Notes :-

1. Extreme values of central 50% observations are Q_1 and Q_3.
2. Largest value of smallest 25% observations equals to Q_1.
3. Smallest value of ~~largest~~ top 25% observations is equals to Q_3.

(B) Deciles

Deciles are 9 values which divides total number of observations into 10 equal parts and the 9 deciles are :- $D_1, D_2, D_3 ---- D_9$
Such that :- $D_1 < D_2 < D_3 ---- < D_9$

$\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$ $\frac{n}{10}$
10% D_1 10% D_2 D_3 D_4 D_5 D_7 D_8 D_9
$D_5 \to Q_2$, Q_3

Date
Page No.

Notes

1. Largest value of smallest 30% observations is D_3.
2. Quartiles lie between D_3 and D_8 are Q_2 and Q_3

(C) Percentiles

Percentiles are 99 values which divide total number of observations into 100 equal parts and the 99 percentiles are $P_1, P_2, P_3 ---- P_{99}$ such that $P_1 < P_2 < P_3 --- < P_{99}$.

1% P_1 1% P_2 ---- P_{44} P_{50} P_{51} --- P_{98} P_{99}

Notes

1. Largest ~~term~~ value of smallest n% observations is P_n.
2. Smallest value of top n% observations is P_{100-n}

P_{100-n}

$50-\frac{n}{2}$ n% $50+\frac{n}{2}$

$P_{50-\frac{n}{2}}$ $P_{50+\frac{n}{2}}$

3. Limits of central n% observations are $P_{50-\frac{n}{2}}$ and $P_{50+\frac{n}{2}}$.

Example (i). If 20% students failed an exam then pass mark in term in percentile is given by

20% | P_{20}

Pass mark = P_{20}.

(ii) Lowest income of richest 10% people is P_{90}.
(iii) Limits of central 60% observations are P_{20} & P_{80}.

20 | 60% | 20
P_{20} P_{80}

Relationship

(A) $M_e = Q_2 = D_5 = P_{50}$

(B) $Q_1 = P_{25}$, $Q_3 = P_{75}$

(C) $D_1 = P_{10}$, $D_2 = P_{20}$, $D_3 = P_{30}$ ——, $D_9 = P_{90}$.

(21) In workshop A

$$\bar{X}_A = \frac{2.50\times2 + 3.50\times14 + 4.00\times20 + 3.00\times7 + 3.00\times6 + 2\times1}{2+14+20+7+6+1}$$

$$= \frac{5 + 49 + 80 + 21 + 18 + 2}{}$$

$$= \frac{175}{50}$$

$$= 3.5$$

$$\bar{X}_B = \frac{3.00\times18 + 3.00\times50 + 4.25\times8 + 3.50\times12 + 3.50\times10 + 5\times2}{18+50+8+12+10+2}$$

$$= \frac{54 + 150 + 34 + 42 + 35 + 10}{100}$$

$$= 3.25$$

So,

difference $= 3.5 - 3.25$

$= 0.25$.

The average rate of wages per worker is higher in A by 0.25 ie 25 paisa.

(25)

x	F	c.F
0-20	42	42
21-30	38	80
31-40	120	200
41-50	84	284
51-60	48	332
61-70	36	368
71-80	31	399
		N = 399

0.5-20.5

Median class $= \left(\frac{N}{2}\right)$th item

$= \left(\frac{399}{2}\right)$

$= 199.5$th term

The c.F just greater than 200 is 284 so the class is ~~41-50~~ 40.5-50.5.

Here,

$l = 40.5$, $c.F = 200$, $\frac{N}{2} = 199.5$.

$h = 10$, $F = 84$

Median $= l + \frac{\left(\frac{N}{2} - c.F\right)}{F} \times h$.

$= 41 + \left(\frac{199.5 - 200}{84}\right) \times 10$

$= 41 - 0.05$.

$= 40.95$.

0.5 - 20.5

20.5 - 30.5

30.5 - 40.5

40.5 - 50.5

50.5 - 60.5.

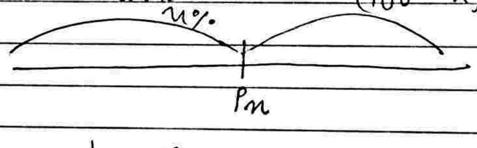
60.5 - 70.5

70.5 - 80.5

(ii) Let x be the percentage of failure

Given; x be the percentage of failure,

Pass mark = 35 $(100-x)\%$.



Pass mark = P_x

$\therefore P_x = 35$, which lies in the class 31-40

$\therefore P_x$ class = 31 - 40 (inclusive type)

= 30.5 - 40.5 (exclusive type)

$l = 30.5$, $h = 10$, $F = 120$, $c.f = 80$, $N = 399$

$$\therefore P_x = l + \frac{\left(\frac{xN}{100} - c.f\right)}{F} \times h$$

$$\text{or, } 35 = 30.5 + \frac{\left(\frac{399x}{100} - 80\right)}{120} \times 10$$

$$\text{or, } 4.5 = \frac{(3.99x - 80)}{12}$$

or, $33.58\% = x$.

Calculation

(A) Discrete series

(I) Quartiles

$$Q_i = \frac{i(N+1)}{4}^{th} \text{ value}; \; i = 1, 2, 3 \text{ ---}$$

(II) Deciles

$$D_i = \frac{i(N+1)}{10}^{th} \text{ value}; \; i = 1, 2, 3, \text{----} 9$$

(III) Percentiles

$$P_i = \frac{i(N+1)}{100}^{th} \text{ value}; \; i = 1, 2, 3 \text{--}$$

(B) Continuous series:-

(I) Quartiles

$$\left.\begin{aligned} Q_i \text{ class} &= \frac{iN}{4}^{th} \text{ class} \\ Q_i &= l + \frac{\left(\frac{iN}{4} - c.F\right)}{F} \times h \end{aligned}\right\} i = 1, 2, 3$$

(II) Deciles:

$$\left.\begin{aligned} D_i \text{ class} &= \frac{iN}{10}^{th} \text{ class} \\ D_i &= l + \frac{\left(\frac{iN}{10} - c.F\right)}{F} \times h \end{aligned}\right\} i = 1, 2, 3 \text{ ---} 9$$

(III) Percentiles

$$\left.\begin{aligned} P_i \text{ class} &= \frac{iN}{100}^{th \text{ term}} \text{ class} \\ P_i &= l + \frac{\left(\frac{iN}{100} - c.F\right)}{F} \times h \end{aligned}\right\} i = 1, 2, 3 \text{ --- } 99$$

31) (b) ~~25~~ P_{25} — 50% — P_{75} ~~25~~

Date
Page No.

(29) Here,

$Q_1 = \frac{n+1}{4}^{th}$ value $= \frac{12+1}{4}^{th} = 3.25^{th}$ value

$= 3^{th}$ value $+ 0.25\,(4^{th} - 3^{th})$ value

$= 40 + 0.25\,(41-40)$

$Q_1 = 40.25.$

$Q_2 = \frac{2(n+1)^{th}}{4}$ value

$= \frac{2\times 13}{4}$ value

$= 6.5^{th}$ value

$= 6^{th} + 0.5(7^{th} - 6^{th})$ values

$= 42 + 0.5(42-42)$

$= 42$

Simple frequency of given more than distribution is,

(32)

x	~~F~~	~~c.f~~	f	c.f
0-10	~~50~~	~~50~~	50-46=4	4
10-20	~~46~~	~~96~~	6	10
20-30	~~40~~		20	30
30-40	~~20~~		10	40
40-50	~~10~~		7	47
more than 50	~~3~~		3	50
				N = 50

Since 60% of the students pass this test then

Pass mark $= P_{40} = ?$

P_{40} — P_{100}

Calculation of P_{40}

P_{40} class $= \frac{40N}{100}^{th}$ class

$= \frac{40\times 50}{100}^{th}$ class.

Date
Page No.

$= 20^{th}$ class

C.F just greater than 20 is 30 and its corresponding class is 20 - 30

$\therefore P_{40}$ class $= 20 - 30$

$l = 20,\ h = 10,\ f = 20,\ c.f = 10,\ \frac{40N}{100} = 20$

$P_{40} = l + \frac{\left(\frac{40N}{100} - c.f\right)}{f} \times h$

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

$= 20 + \frac{10}{20} \times 10$

$P_{40} = 25$

$\therefore$ Pass mark $= 25.$

(34) (b) Total wages of workers, getting less than Rs1200 is given by :-

$\frac{50+80}{2}$

$\Sigma f x_1 = (30\times 65) + (127\times 90) + (140\times 105) + (240\times 115)$

$=$ Rs 55680 (in 00' Rs).

Also,

Total wedges of worker getting more than Rs 12,000 is given by :-

$\Sigma f X_2 = (176\times 125) + (135\times 140) + (20\times 165) + (3\times 190)$

$= 44770$ (in 00' Rs).

Total sum collected $= (5\%$ of $55680) + (16\%$ of $44770)$

$= 2784 + 4477.6$

$= 7261.$ (in Rs 00)

$= 7261 \times 00$

$= 726100$ #

* Graphical calculation of median and partion value

Example :-

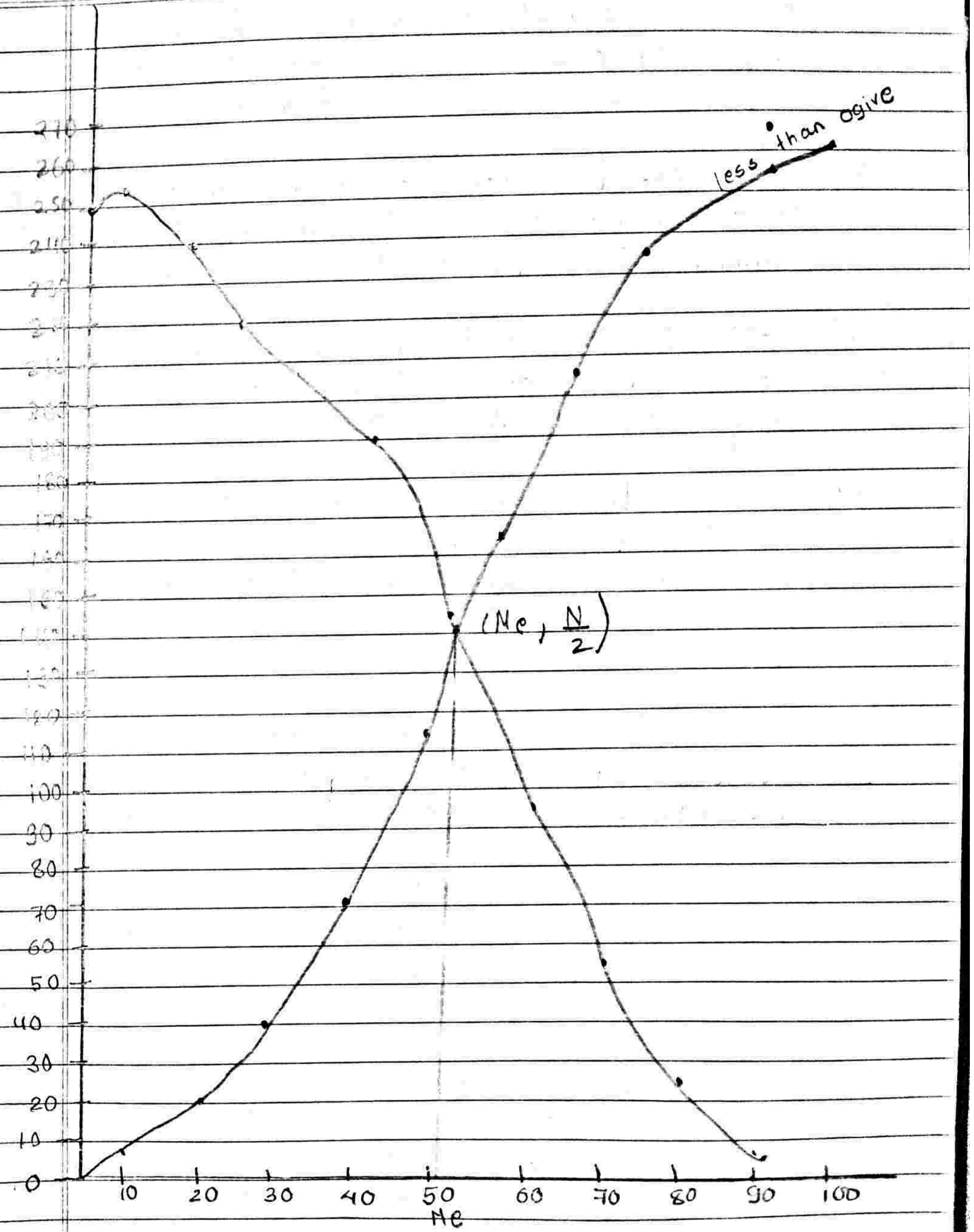
Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-
f :-	8	12	20	30	45	50	42	29	19	5

Less than type C.F distribution

less than	C.F.
10	8
20	20
30	40
40	70
50	115
60	165
70	207
80	236
90	255
100	260

More than type C.F distribution

More than	C.F
10 / 0	252 / 250
20	240
30	220
40	190
50	145
60	95
70	53
80	24
90	5



(i) Median and partation value can be obtained graphically from ogive (specially less than type) ogive for groped f.d only.

(ii) ~~Less than~~ Less than ogive moves from lower left corner to upper right on graph making approx. S shape or J-shape curve on graph.

(iii) Median is the foot of perpendicular (abscissa). Abscissa of a point on less than ogive whose respective ordinate is $\frac{N}{2}$.

(iv) Qi is the foot of perpendicular of a point on less the ogive whose respective ordinate is $\frac{iN}{4}$ where $i = 1, 2, 3$.

26

~~Monthly income~~ ~~No of families~~ ~~It should be converted into~~

Value	No of families (f)	C.F.
less than 100	40	40
100 - 200	89	129
200 - 300	148	277
300 - 400	64	341
400 to above	39	380

N = 380

Here $\frac{N}{2} = \frac{380}{2} = 190$. C.F just greater than 190 is 277

so it lies in 200 - 300 class

Where $l = 200$, $h = 100$, $f = 148$, $cf = 129$

Now,

$$\text{Median} = l + \left(\frac{\frac{N}{2} - c.f}{f}\right) \times h$$

$$= 200 + \left(\frac{\cancel{380}\,190 - 129}{148}\right) \times 100$$

$$= 200 + \frac{61}{148} \times 100$$

$$= 241.2 \ \#$$

27(a) Men who earn less than Rs 2.00 = 100

Rs 2 to Rs 2.24 = 450

Rs 2.25 to Rs 2.49 = 35% of 2000 = 700

Rs 2.50 to Rs 2.74 = 370

Rs 2.75 to Rs 2.99 = 12% of 2000 = 240

more than Rs 3 = 140.

Calculation.

Earning (Rs)	Workers (f)	C.F
Less than 200	100	100
Rs 2 to Rs 2.24	450	550
Rs 2.25 to Rs 2.49	700	1250
Rs 2.50 to Rs 2.74	370	1620
Rs 2.75 to Rs 2.99	240	1860
more than Rs. 3.	140	2000

N = 2000

Here $\frac{N}{2} = \frac{2000}{2} = 1000$. C.f just greater than 1000 is 1250. So class is Rs 2.25 - 2.49

Date
Page No.

Here,

$l = 2.25 \quad F = 700 \quad \frac{N}{2} = 1000$

$h = 0.24 \quad c.f = 550.$

$$Me = l + \frac{\left(\frac{N}{2} - c.f\right)}{F} \times h$$

$$= 2.25 + \left(\frac{1000 - 550}{700}\right) \times 0.24$$

$$= Rs\ 2.40 \ \#$$

(b) Simple frequency of given above distribution is,

Wages (Rs)	No. of labors (f)	c.f
0 - 10	150 (650-500)	150
10 - 20	75	225
20 - 30	50	275
30 - 40	75	350
40 - 50	25	375
50 - 60	25.	400
~~70 and above~~ 60 - 70	150	550
70 and above	100	650

$N = 650$

So, $\frac{N}{2} = \frac{650}{2} = 325$ so class is 30-40

$\frac{N}{2} = 325 \quad F = 75 \quad c.f = 275 \quad h = 10$

$l = 30$

Now,

$$Md = l + \frac{\left(\frac{N}{2} - c.f\right)}{F} \times h$$

$$= 30 + \frac{(325 - 275)}{75} \times 10$$

$$= 36.67 \ \#$$

Date
Page No.

(28)

Expenditure (Rs)	No of families (f)	c.f
39.5 - 59.5	50	50
59.5 - 79.5.	x	$50 + x$
79.5 - 99.5	500	$50 + x + 500$
99.5 - 119.5	y	$550 + x + y$
119.5 - 129.5	50	$600 + x + y$
~~119.5 - 129.5.~~		

$N = 600 + x + y. = 1000$

$$\frac{N}{2} = \frac{\cancel{600 + x + y}\ 1000}{2} = \cancel{300 + x + y} = 500$$

We have,

c.f just greater than 500 is $550 + x$ so the median lies betn 79.5 - 99.5.

So, $l = 79.5 \quad h = 20 \quad F = 500 \quad c.f = 50 + x + 500$

$\frac{N}{2} = 500$

We have,

$$Me = l + \frac{\left(\frac{N}{2} - c.f\right)}{F} \times h$$

$$87 = 79.5 + \frac{(500 - (50 + x))}{\cancel{500}\ 25} \times 20$$

$$87 = \frac{1987.5 + 450 - x}{25.}$$

$2175 = 1987.5 + 450 - x$

$\cancel{-} 262.5 = \cancel{-}x$

$263 = x.$

We have,

$600 + x + y = 1000$

$600 + 263 + y = 1000$

$y = 137 \ \#.$

Date
Page No.

30)

Marks (n)	No of students (f)	C.F
20-25	10	10
25-30	20	30
30-35	20	50
35-40	15	65
40-45	15	80
45-50	20	100

N = 100

For P_{45}.

$$\frac{kN}{100} = \frac{45 \times \cancel{100}}{\cancel{100}} = 45.$$

just greater in c.f is 50 so lies betn 30-35.
where, $l = 30$ F = 20 c.F = 30 h = 5

$$P_{45} = l + \frac{\frac{45N}{100} - c.F}{F} \times h$$

$$= 30 + \frac{45 - 30}{20} \times 5.$$

$$= 30 + \frac{15}{\cancel{20}_4} \times \cancel{5}^1$$

$$= 30 + \frac{15}{4}$$

$$= 33.75 \#$$

For P_{57}

$$\frac{kN}{100} = \frac{57 \times \cancel{100}}{\cancel{100}} = 57$$

C.F just greater 65 so class 35-40, $l = 35$, F = 15 c.F = 50

$$P_{57} = l + \frac{\frac{57N}{100} - c.F}{F} \times h$$

$$= 35 + \frac{57 - 50}{15} \times 5.$$

$$= 37.33 \#.$$

Date
Page No.

31)

Wages (Rs)	No of persons (F)	C.F
30-40	1	1
40-50	3	4
50-60	11	15
60-70	21	36
70-80	43	79
80-90	32	111
90-100	9	120

N = 120.

For lower quartiles i = 1. Q_1 $\frac{iN}{4}$

$$\frac{iN}{4} = \frac{1 \times \cancel{120}^{30}}{\cancel{4}} = 30.$$

C.f just greater than 30 is 36 so it lies betn 60-70.
$l = 60$ h = 10 F = 21 c.f = 15

$$Q_1 = l + \left(\frac{\frac{iN}{4} - c.F}{F}\right) \times h.$$

$$= 60 + \left(\frac{30 - 15}{21}\right) \times 10$$

$$= 60 + 7.14$$

$$= 67.14.$$

For Q_3 (Upper quartile) Q_3 $\frac{iN}{4}$ = 3

$$Q_3 \text{ class} = \frac{iN}{4} = \frac{3 \times \cancel{120}^{30}}{\cancel{4}} = 90.$$

C.F just greater the 90 is 111 so 80-90
$l = 80$ h = 10 F = 32 c.f = 79.

$$Q = l + \frac{\left(\frac{iN}{4} - c.F\right)}{F} \times h$$

$$= 80 + \frac{(90 - 79)}{32} \times 10$$

$$= 83.44 \#$$

For 7th decile.

D_i class; $\frac{iN}{10}$ th class $= \frac{7\times 12\cancel{0}}{1\cancel{0}} = 84$.

C.f just greater than 84 is 111 cl. 80 - 90.

$l = 80 \quad h = 10 \quad f = 32 \quad c.f = 79$

$$D_7 = l + \frac{\left(\frac{iN}{10} - c.f\right)}{f} \times h.$$

$$= 80 + \frac{(84 - 79)}{32} \times 10$$

$$= 81.56.$$

For percentile,

$$P_{60} = \frac{kN}{100} = \frac{60\times 12\cancel{0}}{1\cancel{00}} = 72$$

C.f just greater is 79 so class is 70-80

$l = 70 \quad h = 10 \quad f = 43 \quad c.f = 36.$

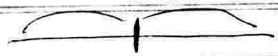
$$P_{60} = l + \frac{\left(\frac{kN}{100} - c.f\right)}{F} \times h$$

$$= 70 + \frac{(72 - 36)}{43} \times 10$$

$$= 78.37 \; \#$$

3(b)

Mark (x)	No of students (f)	c.f
0-10	45	45
10-20	85	130
20-30	160	290
30-40	75	365
40-50	35	400
	N=400	



P_{50}. 50

For Q_1; $\frac{iN}{4}$ class $= \frac{1\times \cancel{400}\,100}{\cancel{4}} = 100$ is 130 so

class is 10 - 20

$l = 10 \quad h = 10 \quad f = 85 \quad c.f = 45.$

$$Q_1 = l + \frac{\left(\frac{iN}{4} - c.f\right)}{F} \times h$$

$$= 10 + \frac{(100 - 45)}{85} \times 10$$

$$= 16.47$$

For $Q_3 = \frac{iN}{4} = \frac{3\times \cancel{400}\,100}{\cancel{4}} = 300$ is 365 so 30-40

$l = 30 \quad F = 75 \quad c.f = 290 \quad h = 10$

$$Q_3 = l + \frac{\left(\frac{iN}{4} - c.f\right)}{F} \times h$$

$$= 30 + \frac{(300 - 290)}{75} \times 10$$

$$= 31.33 \; \#$$

Note Cotnd.

(v) D_i is the foot of perpendicular of a point on less than ogive whose respective ordinate is $\frac{iN}{10}$. $(i = 1, 2, 3, 4 --- 9)$

(vi) P_i is the foot of perpendicular of a point on less than ogive whose respective ordinate is $\frac{iN}{100}$ $(i = 1, 2, 3 --- 99)$

(vii) $(M_e, \frac{N}{2})$ is the point of intersection of less than and more than type ogives. Example:- If $(50, 60)$ is the point of intersection of less than and more than type ogives then find the value of median and total number of observations.

$M_e = 50 \quad N = 140$.

(viii) Median is the foot of perpendicular of a point of intersection of less than and more than type ogives.

Mode

- The most repeted value
- The most frequent value
- The most usual value
- The most common value
- The most concentrate value.

Mode is the most repeated value and the value has the maximum concentration of observations around it. Example:-

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

X :	1	2	3	4	5	6
F :	2	3	4	2	2	1

Example 2

10, 15, 20, 15, 11, 12, 15, 20, 20, 25, 26

X :	10	11	12	15	20	25	26
F :	1	1	1	3	3	1	1

Mode = 15 & 20

Ex: 3 105, 100, 108, 120, 150, 180.

Mode doesn't exist.

Notes:-

(1) Unimodal distribution

A frequency distribution having only one mode. It is also known as single ~~peak~~ peaked distribution.

(2) Bimodal distribution

A frequency distribution having two modes.

(3) Multimodal distribution

Multimodal distribution having more than one mode.

- Mode is illdefined.
- Mode is uniquely defined for single picked regular distribution.
- Frequency distribution is needed for the calculation of mode.
- It is usually better to calculate mode from grouped data rather than the ungrouped data. because the chance of an unrepresentative value beign chosen as mode is reduced.

Date
Page No.

Calculation of mode for unimodal regular distribution.

(A) For ungrouped data
Mode = a value, which has max. frequency

(B) For grouped data
Modal class = A class which has max. frequency.

$$M_o = l + \frac{d_1}{d_1 + d_2} \times h$$

where,
h = class size
$d_1 = f_1 - f_0$
$d_2 = f_1 - f_2$
f_1 = maximum frequency / modal class frequency
f_0 = frequency preceding to f_1
f_2 = frequency sucessing to f_1

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

Note :- It is applicable under the following cases.
1. Classes are continous with exclusive type.
2. Classes are equal in size.

44 (b)

C.I	0-10	10-20	20-40	40-50	50-70
F	5	15	40	32	28

d = difference of
or, d = 800 - 600

Date
Page No.

For the calculation of mode, class size of the given frequency distribution are made equal by regrouping the given distribution.
Regrouping the given frequency distribution as:-

Class :	0-10	10-20	20-30	30-40	40-50	50-60	60-70
F :	5	15	20	20	32	14	14

Modal class = 40 - 50

$l = 40$ $h = 10$ $f_1 = 32$ $f_0 = 20$ $f_2 = 14$

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

$$= 40 + \frac{(32 - 20)}{2 \times 32 - 20 - 14} \times 10$$

$$= 40 + \frac{12}{64 - 20 - 14} \times 10$$

$$= 40 + \frac{\cancel{12}^{4}}{\cancel{30}} \times \cancel{10}$$

$$= 44$$

(38)

Income (central value)	600	800	1000	1200	1400
No of school	2	3	5	4	2
Average no of teachers	20	26	21	21	9

~~Income~~ ~~No of schools~~ ~~No of teacher~~

Income (central value) (m)	No of school	No of teachers	Class	Total no of teachers (F)	$d' = \frac{x - 1000}{200}$	Fd'	C.F
600	2	20×2=40	500-700	40	-2	-80	40
800	3	26 78	700-900	78	-1	-78	118
1000	5	21 105	900-1100	105	0	0	223
1200	4	21 84	1100-1300	84	1	84	307
1400	2	9 18	1300-1500	18	2	36	325.

N = 325

d = difference of two consecutive mid values
or $d = 800 - 600 = 200$

$\therefore$ Correction factor

$$C.F = \frac{d}{2} = \frac{200}{2} = 100$$

$$LCL = \text{Mid value} - C.F$$

$$UCL = \text{Mid value} + C.F$$

For 1^{st} class interval

$$LCL = 600 - 100 = 500$$

$$UCL = 600 + 100 = 700$$

$$\text{Class} = 500 - 700 \;\&\; 800$$

Calculation of mean

$A = 1000$, $h = 200$, $N = 35$, $\Sigma fd' = -38$

$$\text{Mean} = A + \frac{\Sigma fd'}{N} \times h$$

$$= 1000 + \frac{-38}{35} \times 200$$

$$= 1000 - 217.14$$

$$= 782.8$$

Method of grouping

If a given frequency distribution is not regularly distributed i.e frequencies are irregular then mode can be obtained by using method of grouping. It consists of two tables.

(i) Grouping table.
(ii) Anyalysis table

(i) Grouping table

Grouping table consists of 6 column of frequencies. In first column, given frequencies are taken. In second column frequencies are added 2 by 2 from first. In the third column frequencies are added 2 by 2 leaving first. In fourth column frequencies are added 3 by 3 ~~leaving~~ adding first.
In fifth column frequencies are added 3 by 3 leaving first.
6th column - 3 by 3 leaving first two.

(ii) Anyalysis table.

Concentration of observations at each value is calculated and mode is obtained as a value which has the maximum number of concentration.

(42)

x_i : 20 25 30 35 40 45 50 55 60 70
f_i : 5 6 8 10 12 3 2 6 1 4.

x	Col I	Col II	Col III	Col IV	Col V	Col VI
20	5					
25	6	11		19		
30	8		14		24	
35	10	18				30
40	12		22	25	17	
45	3	15				11
50	2		5			
55	6	8		9	11	
60	1		7			
70	4	5				

Analysis table

X / Column	20	25	30	35	40	45	50	55	60	70
I					1					
II			1	1						
III			~~1~~	1						
IV				1	1					
V				1	1	1				
VI		1	1	1						
VII			1	1	1					
	–	1	3	5	4	1				

Mode = 35.

Empirical formula for mode

For multimodal frequency distribution, a mode value can be obtained by using the empirical formula which is given by :-

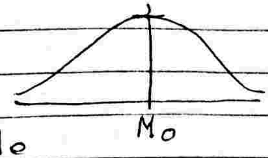
$$\text{Mode} = 3\,\text{Median} - 2\,\text{Mean} \quad \text{—(i)}$$

holds for a moderately asymmetrical distribution.

It is an empirical relationship between mean, median and mode. It doesn't satisfy mathematical properties.

Note :- Moderately asymmetrical distribution is not a very skewed distribution or not a very ~~skewed dis~~ asymmetrical distribution.

Symmetrical distribution



In a symmetrical distribution, elongation of curve on either side about mode are identical.

In a symmetrical distribution,

$$\text{Mean} = \text{Mode} = \text{Median}$$

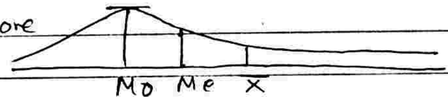
i.e Mean, Median and mode coincide at a point.

Skewness :- (Lack of symmetry).

It is of two types

(i) Right skewness or positive skewness.

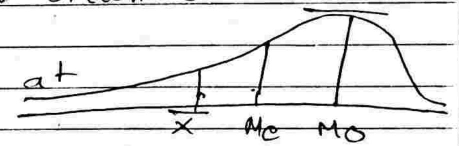
– Elongation of curve is more at right than that of left.



– Extreme values lies at right side of the distribution in right skewed distribution. $Mo < Me < \bar{X}$

(ii) Left skewness or negative skewness.

– Elongation of curve is more at left than that of right side.



– Extreme value lies at left side.

– $\bar{X} < Me < Mo$.

Note :-

For moderately assymetric distribution.

1. Median always lies between mean and mode.

Note 2:- Median divides difference between mean and mode in the ratio 1:2 being closer to mean.

Note 3:- $\frac{\bar{X} - Me}{Me - Mo} = \frac{1}{2}$

$$\frac{\bar{X} - Me}{\bar{X} - Mo} = \frac{1}{3}$$

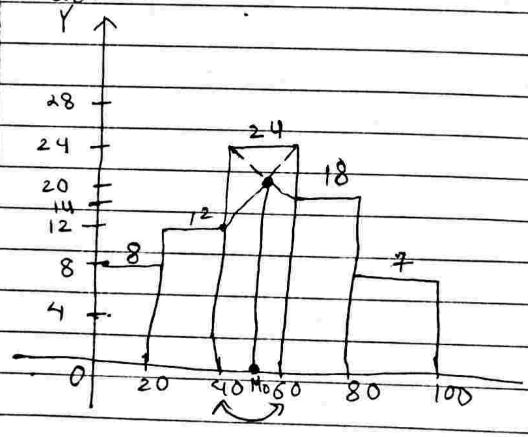
$$\frac{Me - Mo}{\bar{X} - Mo} = \frac{2}{3}$$

Graphical calculation of mode :-

Example :- Obtain mode graphically.

Class :-	0-20	20-40	40-60	60-80	80-100
f :-	8	12	25	18	7

Soln :-



(1) Note :- Mode of the grouped frequency distribution can be obtained graphically using histogram.

(2) In histogram, mode is the foot of perpendicular of a point of intersection of diagonals obtained by joining upper corners of the highest rectangle to the points on it's sides where upper corner of adjacent rectangles touch.

Properties of mode

(1) If all the observations assumed by a variable are constant then their mode is also the same constant. Ex :- 3, 3, 3, 3, 3 $Mo = 3$.

(2) It is dependent on change of origin and scale.

(3) If x and y are linearly related by $y = a + bx$ then; $y_{mo} = a + b\,x_{mo}$ where a and b are constants.

Example :- Find mode of y if mode of x is 10 for $X - Y = 2$

Soln :- For, $X - Y = 2$

$$X_{mo} - Y_{mo} = 2$$

or, $10 - Y_{mo} = 2$

$$\therefore Y_{mo} = 8.$$

35) Size Below 10 10-12 12-14 14-16 16-18 18-20

Size	Below 10	10-12	12-14	14-16	16-18	18-20
Demand	3	15	27	20	3	2
	3	18	45	65	68	70

$N = 70$

Class = 12-14

$l = 12 \quad h = 2 \quad f_1 = 27 \quad f_0 = 15 \quad f_2 = 20$

$$M_0 = l + \frac{(f_1 - f_0)}{2f_1 - f_0 - f_2} \times h$$

$$= 12 + \frac{27-15}{2\times 27 - 15 - 20} \times 2$$

$$= 12 + \frac{12}{19} \times 2$$

= 13.26 inches

For Md,

$N = 70 \quad \frac{N}{2} = \frac{70}{2} = 35$

so class 12-14

$$Md = l + \frac{(\frac{N}{2} - c.f)}{F} \times h$$

So, Md lies in the same class.

39.

Sales	No of dealers
0 - 500	40
500 - 1000	48
1000 - 1500	60
1500 - 2000	52
2000 - 2500	35
2500 & more	22

So the modal calass is 1000 - 1500

$l = 1000 \quad h = 500 \quad f_1 = 60 \quad f_0 = 48 \quad f_2 = 52$

$$M_0 = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$= 1000 + \frac{60-48}{2\times 60 - 48 - 52} \times 500$$

$$= 1000 + \frac{12}{20} \times 500$$

$= 1300$

300 200

1000 1300 1500

$= \frac{\cancel{200}}{\cancel{300}} \times 60 \quad 20 = 40.$

~~Class 1000-1500 = 60+52+35+22 = 169.~~

Here,

Number of dealer who sells between 1300 to 1500 bags $= \frac{200}{500} \times 60$

$= 24$

So, total number of dealers getting awards

$= 24 + 52 + 35 + 22$

$= 133$

And,

total awards = Rs 5000 x 133

= Rs 665,000

40) Frequcy distribution density $= \frac{F}{h} = \frac{60}{500} \times (250)$ class width

2

Date
Page No.

(40) →

x	F	(X)	FX	
0-20	14	10	140	
20-40	x	30	30x	690
40-60	27	50	1350	
60-80	y	70	70y	1470
80-100	15	90	1350.	
	N = 100.		~~2940xy~~	

$\overline{X} = \frac{\Sigma FX}{N}$

$14 + x + 27 + y + 15 = 100$

$x + y = 44$ ∴ $y = 44 - x$

Here the given mode = 48 so the group lies in 40-60.

Here, $l = 40$ $h = 20$ $f_1 = 27$ $f_0 = x$ $f_2 = y$

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

$$48 = 40 + \frac{(27 - x)}{2 \times 27 - x - y} \times 20$$

$$48 = 40 + \frac{(27 - x)}{54 - x - y} \times 20$$

$$8 = \frac{27 - x}{54 - x - y} \times 20$$

$8(54 - x - y) = 20(27 - x)$

$432 - 8x - 8y = 540 - 20x$

$432 - 8y = 540 - 12x$

$432 - 8(44 - x) = 540 - 12x$

$432 - 352 + 8x = 540 - 12x$

Date
Page No.

or, $20x = 460$

$23 = x$

$y = 44 - 23$

$y = 21.$

Now, $30 \times x = 30 \times 23 = 690$

$70 \times y = 70 \times 21 = 1470$

So, $\Sigma FX = 5000$

$N = 100$

So for mean,

$$\overline{X} = \frac{\Sigma FX}{N}$$

$$\overline{X} = \frac{5000}{100}$$

$\overline{X} = 50$ #.

Date
Page No.

G.M

geometric mean is n^{th} root of product of 'n' non-zero and non-negative observations

if $x_1, x_2, x_3 ---- x_n$ are 'n' non-zero and non-negative observations, then,

$$GM = \sqrt[n]{x_1 x_2, x_3 --- x_n}$$

i.e $G.M = (x_1 x_2, x_3 --- x_n)^{\frac{1}{n}}$

then we can derive,

$$GM = \text{Antilog}\left(\frac{\Sigma \log x}{n}\right)$$

Calculation of G.M

(A) Individual series.

(I) $G = (m_1 \cdot m_2 \cdot m_3 ---- m_n)^{\frac{1}{n}}$

(II) $G = \text{Antilog}\left(\frac{\Sigma \log X}{n}\right)$

(B) Discrete or continuous series:-

(I) $G = (m_1^{F_1} \cdot m_2^{F_2} \cdot m_3^{F_3} --- m_n^{F_n})^{\frac{1}{N}}$

(II) $G = \text{Antilog}\left(\frac{\Sigma F \log X}{N}\right)$

Example I :- find G.M of the following data.

X	2	3	4	6
F	2	3	3	2

Soln:-

$G = (2^2 . 3^3 . 4^3 . 6^2)^{\frac{1}{10}}$

$= (2^2 3^3 . 2^6 . 2^2 . 3^2)^{\frac{1}{10}}$

$= (2^{10} 3^5)^{\frac{1}{10}}$

$= 2^{10 \times \frac{1}{10}} 3^{5 \times \frac{1}{10} 2}$

Number ✗ √ 19 times then -1 ani ✗ 2.27695 =

$19\sqrt{} \times 2.27695$

Date
Page No.

$= 2\sqrt{3}$

$= 2.\sqrt{3}$ #.

45 (a)

(i) Variable value (X) 5 7 9 12 18.

X	log X
5	0.6989
7	0.8451
9	0.9542
12	1.0792
18	1.2553
	4.8327

$G = \text{Antilog}\left(\frac{\Sigma \log X}{n}\right)$

$= \text{Antilog}\left(\frac{4.8327}{5}\right)$

$= \text{Antilog}(0.96654)$

$= 9.26$ #

45 a (iii)

C.I	C.F	Mid values (X)	log x	F log X
5-14	5	9.5	0.97	4.8885
15-24	7	19.5	1.2900	9.03
25-34	10	29.5	1.4968	14.698
35-44	4	39.5	1.5965	6.3860
45-54	2	49.5.	1.6946	3.3892
	N=28			38.3917

Antilog $\left(\frac{\Sigma F \log X}{N}\right)$

$N = 28$; $\Sigma F . \log X = 38.3917$

$\therefore G = \text{Antilog}\left(\frac{\Sigma F . \log}{N}\right)$

$= \text{Antilog}\left(\frac{38.3917}{28}\right)$

$1.3711 \div 227695 + 1 \ldots$

Date
Page No.

$= \text{Antilog}\left(\frac{38.3917}{28}\right)$

$= \text{Antilog}(1.3711) = 23.51$

Properties of G.M.

(1) If all the observation assumed by a variable are constant then their G.M is also the same constant.

Example:- 2, 2, 2, 2, 2

$G.M = (2\times2\times2\times2\times2)^{1/5}$

$= 2^{5\times\frac{1}{5}}$

$= 2$ #

(2) G.M of all the observations in G.P is given by $\sqrt{a \cdot l}$. For example:-

* Find G.M of $X: 1, 3, 3^2, 3^3, 3^4 \ldots 3^n$

$= \sqrt{1\times3^n}$

$= 3^{n/2}$ #

(3) Product of n observations is equal to nth power of their G.M.

(4) If each observations is multiplied or divided by a constant value then their G.M is also multiplied or divided by the same constant value.

$GM(cX) = c \cdot GM(X)$

where c is any constant.

(5) If X and Y are any two variables. G.M mean of X·Y is,

(a) $GM(X \cdot Y) = Gm(X) \cdot Gm(Y)$

Date
Page No.

b) $Gm\left(\frac{X}{Y}\right) = \frac{Gm(X)}{Gm(Y)}$

c) $Gm\left(\frac{1}{X}\right) = \frac{1}{Gm(X)}$

Harmonic Mean:-

Example:- X: 2 4 5 8

$\frac{1}{X}: \frac{1}{2}\ \frac{1}{4}\ \frac{1}{5}\ \frac{1}{8}$

am

$\frac{\Sigma \frac{1}{x}}{n}: \frac{\frac{1}{2}+\frac{1}{4}+\frac{1}{5}+\frac{1}{8}}{4} = \frac{1.075}{4}$

reciprocal

$H = \frac{n}{\Sigma\frac{1}{x}} = \frac{4}{\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{5}+\frac{1}{8}\right)} = \frac{4}{1.075} = 3.7209.$

$= \frac{4\times40}{20+10+8+5}$

$= \frac{160}{43}.$

Harmonic mean is the reciprocal of A.M of reciprocal of all non-zero observations.

If $x_1, x_2, x_3 \ldots x_n$ are 'n' non zero observations then Hm is

$$H = \frac{n}{\left(\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\cdots\frac{1}{x_n}\right)}$$

or, $H = \frac{n}{\Sigma\frac{1}{x}}$

Example:-

* Find harmonic mean of $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4} \ldots \frac{1}{n}$.

$H = \frac{n}{\Sigma\frac{1}{n}} = \frac{n}{1+2+3+4 \ldots n}$

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

$$= \frac{2}{n+1} \#$$

Calculation of harmonic mean.

A # Individual series.

$$H = \frac{n}{\Sigma \frac{1}{X}}$$

B # Discrete or Continuous series

$$H = \frac{N}{\Sigma \frac{f}{X}}$$

41) $\frac{Me - Mo}{\bar{X} - Mo} = \frac{2}{3}$

or, $\frac{28 - Mo}{25 - Mo} = \frac{2}{3}$

or, $84 - 3Mo = 50 - 2Mo$

or, $34 = Mo$

$Mode = 3Me - 2M\bar{X}$

$= 3 \times 25 - 2 \times 28$

~~$= 84 - 50$~~

$= 75 - 56$

$= 19$ #

42)

44b)

M	F
0-10	5
10-20	15
20-30	20
30-40	20
40-50	32
50-60	14
60-70	14

Since 32 has the highest frequency the class is 40-50.

$l = 40 \quad h = 10 \quad F_1 = 32 \quad F_0 = 20 \quad F_2 = 14$

$$Mo = l + \frac{F_1 - F_0}{2F_1 - F_0 - F_2} \times h$$

$$= 40 + \frac{32 - 20}{2 \times 32 - 20 - 14} \times 10$$

$$= 40 + \frac{\cancel{12}\ 4}{\cancel{30}} \times \cancel{10}$$

$= 44.$

44a)

Profit		F.		C.F	mid (X)	FX
~~0-10~~ Below 10.	~~0~~	0	~~0~~	0	5	0
10-20	~~5~~	5	~~5~~	5	15	75
20-30	~~9~~	9	~~5~~	14	25	225
30-40	~~21~~	13	~~14~~	27	35	455
40-50	~~20~~	21	~~27~~	48	45	945
50-60	~~15~~	20	~~48~~	68	55	1100
60-70	~~8~~	15	~~68~~	83	65	975
70-80	~~3~~	8	~~8~~	91	75	600
80-90		3		94	85	255
		N = 94				$\Sigma fx =$ 4630

$= \frac{\Sigma fX}{N}$

$= \frac{4630}{94} =$

$= 49.2$

For median,

$\frac{N}{2} = \frac{94}{2} = 47.$

C.F. greater than 47 is 48 so class 40-50

$= l + \frac{\left(\frac{N}{2} - C.F\right)}{F} \times h$

$= 40 + \frac{47 - 27}{21} \times 10$

$= 40 + \frac{20}{21} \times 10$

$= 49.52$

So Mode $= 3Me - 2\times\overline{X}$

$= 3\times 49.52 - 2\times 49.2$

$= 148.56 - 98.4$

$= 50.1$

us(ii)

X	F	log X	F·log X
10	2	1.0000	2.0000
20	4	1.3010	5.204
30	8	1.4771	11.8168
40	3	1.6828	5.0484
55	2	1.7403	3.4806
	N=19	7.2012	27.5498

$G.M = \text{Antilog}\left(\frac{\Sigma F \log X}{N}\right).$

$= \text{Antilog}\left(\frac{27.5498}{19}\right)$

$= \text{Antilog}(1.44)$

Date
Page No.

Find H.M of

x : 2 3 4 12
f : 2 3 3 2

$$H = \frac{N}{\sum \frac{f}{x}} = \frac{10}{\frac{2}{3}+\frac{3}{3}+\frac{3}{4}+\frac{2}{12}} = \frac{10}{2.916}$$

$$= \frac{10}{2.916} = 3.428 = 3.43$$

Properties of H.M

- If all the observations assumed by a variable are constant then their H.M is also the same constant. i.e

i.e X : 3, 3, 3, 3, 3

$$H = \frac{n}{\sum \frac{1}{x}} = \frac{5}{\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}+\frac{1}{3}}$$

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

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

$$= 3$$

- If each observations is multiplied or divided by a constant value then their H.M is also multiplied or divided by the same constant value.

$$H(c \cdot X) = c \cdot H(x)$$

where c is a constant.

$$\text{Sum of reciprocal} = \frac{n}{H}$$

Date
Page No.

- If n_1 and n_2 be the number of observations of two groups and H_1 and H_2 are their respective H.M's then combined harmonic mean of the two groups taken together is given by

$$H = \frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}}$$

Relationship between AM, GM & HM

Example :- X : 2, 6, 8, 10

$$AM = \frac{2+6+8+10}{4} = 6.5$$

$$GM = (2 \times 6 \times 8 \times 10)^{\frac{1}{4}} = 5.56$$

$$H.M = \frac{4}{\frac{1}{2}+\frac{1}{6}+\frac{1}{8}+\frac{1}{10}}$$

$$= \frac{4}{0.8416} = 4.48$$

$$\boxed{\therefore HM < GM < AM}$$

Example :

X : 2, 4, 8, 16

$$A.M : \frac{2+4+8+16}{4} = 7.5$$

$$G.M : \sqrt{2 \times 16} = 4\sqrt{2}$$

$$HM : \frac{4}{\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}}$$

$$= \frac{4 \times 16}{16\left(\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\frac{1}{16}\right)}$$

$$= \frac{64}{15}.$$

$$\sqrt{AM \times HM} = \sqrt{7.5 \times \frac{64}{15}} = \sqrt{32} = G.M.$$

Note no 1:- For distinct observations, ~~AM < GM < HM~~ $HM < GM < AM$

2:- For any observations, $HM \le GM \le AM$

3. For observations in G.P, $\sqrt{A.M \times HM} = GM$

or, $H.M \times A.M = (G.M)^2$

4. For any two observations a and b

(a) $A.M = \frac{a+b}{2}$

(b) $G.M = \sqrt{ab}$

(c) $HM = \frac{2ab}{a+b}$

(d) $\sqrt{AM \times HM} = G.M.$

Criteria for good average

1. It should be ridgidly defined.
2. It should be easy to compute and simple to understand.
3. It should be based on all the observations.
4. It satisfieds mathematical properties. algebrical treatment
5. It should be least affected by extreme observations.
6. It should be least affected by fluctuation of sampling.

Merits / Demerits	Mean	Median	Mode	G.M	H.M
			✗		
1. It is ridgily defined.	✓	✓	ill defin	✓	✓
2. It is easy to compute and simple to understand.	✓ (3) easy	✓ (2) easier	✓ (1) easiest	✗ (5)	✗ (4)
3. It is based ~~of~~ on all the observations.	✓	✗ approx central sv.	✗	✓	✓
4. It satisfies mathematical properties.	✓	✗	✗	✓	✓
5. It is least affected by extreme observations.	✗ Highly affected	✓ Not affected at all (1)	✓ Not affected (1)	✓ (3)	✓ (2)
6. It is least affected by fluctuation of sampling.	✓ (1)	✗ affected significantly	✗ highly affected	✓ (2)	✓ (3)
7. It can be calculated for the distribution with open type classes.	✗	✓	✓	✗	✗
8. It satisfies liniarity porperty.	✓	✓	✓	✗	✗

Application

Application of harmonic mean :-

- It is suitable for finding average of data of rates. For example :- Speed, price etc.

* For example :- A vehicle moves from A to B which speed sixty kilometer per hour. and then it returns from B to A with speed 45 kilometer per hour. Find the average speed of it in entire trip.

$\therefore$ Average speed is given by $= \frac{2ab}{a+b} = \frac{2 \times 60 \times 45}{60+45}$

$= 51.43$ km/hr

Date

Page No.

- It has limited use.

Application of geometric mean
- It is the most suitable average for the data of (rate, ratio, percentage).

Note:- GM is the best average for constructing index number.

Mode (Application)
- It is commonly used average for the data related to sales and business. For example. To find average size of shoes for sale, mode is appropriate average here.

Median (Application)
- It is the most suitable average under the following conditions.

(i) In missing some extreme observations

Example:-

* Find average mark of 15 students. Marks obtained by 10 pass students are available which are given below.

80, 45, 40, 65, 70
90, 95, 85, 42, 56

Arranging the 15 observation in ascending order.

-, -, -, -, -, 40, 42, 45, 56, 70, 80, 85, 90, 95.

Here,

$n = 15$

$Me = \left(\frac{n+1}{2}\right)^{th}$ value $= \frac{16}{2}^{th}$ value $= 8^{th}$

$= 45$

$\therefore$ Average 45 marks.

AM - stable average
Mode - unstable average.

Date

Page No.

(II) Frequency distribution consist of open type classes.

(III) Frequency distribution consists of varying level of class size. (skewed)

(IV) In assemetrical distribution.

(V) In the distribution, consisting of extreme values.

(VI) To find average of qualitative data - such as average itellegent, average honest, average beauty.

Application of mean.
- It is the best average and is commonly used.

Date
Page No.

Measures of dispersion:-

For a given set of observations, dispersion may be defined as the amount of deviation of observations measured from an appropriate measure of central tendency.

Importance of measures of dispersion:-

1. It is used to examine the relaibility of average for the given data.
2. It helps in the comparision of variability of two or more data.
3. It is suitable for further stastistical treatment. ~~for~~ / anyalysis.

Types of measures of dispersion

Measure of dispersion is broadly classified into two types:
1. Absolute measures of dispersion.
2. Relative measures of dispersion.

Dispersion

Graphical nature Lorons curve.

(A) Absolute measures
(I) Range
(II) Quartile deviation
(III) Mean deviation
(IV) Standard deviation

(B) Relative measures.
(I) Coefficient range
(II) Coefficient of Q.D
(III) Coefficient of M.D
(IV) Coefficient of variation (CV)

6

Date
Page No.

Differentiate between absolute and relative measures of dispersion.

Absolute measure	Relative measure.
1. Absolute measures are dependent on units of the variables i.e it is unit associated measures.	1. Relative measures are independent on units of variables i.e it is unit free measure.
2. For the comparision of variability of two or more data, absolute measures are not considered.	2. For the comparision of variability of two or more data, ~~absolute~~ relative measures are considered.
3. Absolute measures are easy to compute & simple to understand.	3. Relative measures are difficult to compute & complex to understand as per ~~relati~~ absolute measure.

(A) Range

For a given set of observations range is difference between largest and smallest value of the given observation i.e

Range = Largest value − Smallest Value
$= L - S$

and Coefficient of range $= \frac{L-S}{L+S} \times 100$

i.e Degree of range / (Relative range) $= \frac{\text{Absolute range}}{\text{Sum of extreme values}} \times 100$

Note :-

① Range is independent on frequencies of the given distribution.

② For a grouped frequency distribution, range is the difference between class boundaries of two extreme classes.

Example :-

* Find range.

X :	10 - 19	20 - 29	30 - 39	40 - 49
F :	8	12	16	5

$S = 9.5$

$L = 49.5$

$R = L - S$
$= 49.5 - 9.5$
$= 40$ #

③ Range cannot be calculated for the distribution with open type classes.

④ For a symmetrical distribution (i) smallest value =
→ (i) $\bar{x} - \frac{R}{2}$

(ii) Largest value $= \bar{x} + \frac{R}{2}$.

⑤ As sample size increases range also tends to increase but not proportionate.

(B) Quartile deviation

For a given set of observations quartile deviation is half of difference of two extreme quartiles.

i.e. $QD = \frac{Q_3 - Q_1}{2}$

and coefficient of Q.D $= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$

$= \frac{Q_3 - Q_1}{2Me} \times 100$

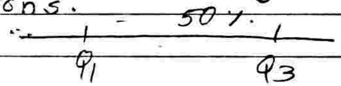
$= \frac{QD}{Me} \times 100$

Notes :-

① It is the most appropriate major for the distribution with open end classes.

② Interquartile range :-

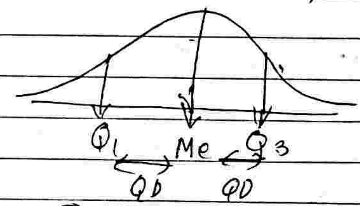
It is a range of central 50% observations.



Interquartile range $= Q_3 - Q_1$

∴ Interquartile range $= 2 \times QD$.

③ For symmetrical distribution,



(a) $Q_1 = Me - QD$

(b) $Q_3 = Me + QD$

(c) $Me = \frac{Q_3 + Q_1}{2}$

(d) Coefficient of $QD = \frac{QD}{Me} \times 100$.

(C) Mean deviation

For a given set of observation mean deviation is an A.M of absolute deviations of all observations taken from either mean, median or mode.

i.e MD from (m) = $\frac{\Sigma|X - m|}{n}$

where 'm' may be $\bar{X}$ or Me or Mo

Calculation of mean deviation.

(A) Absolute measure :-

(I) Individual series:-

(a) M.D of $\bar{X} = \frac{\Sigma|X - \bar{X}|}{n}$

(b) M.D from Me = $\frac{\Sigma|X - Me|}{n}$

(c) M.D from Mo = $\frac{\Sigma|X - Mo|}{n}$

(II) Discrete or continuous series

(a) MD from $\bar{X} = \frac{\Sigma f|X - \bar{X}|}{N}$

(b) MD from Me = $\frac{\Sigma f|X - Me|}{N}$

(c) MD from Mo = $\frac{\Sigma f|X - Mo|}{n}$

(B) Relative Measure

Coeff of M.D from $\bar{X}$ = $\frac{\text{MD from } \bar{X}}{\bar{X}} \times 100$;

Coeff of M.D from Me = $\frac{\text{MD from Me}}{\text{Me}} \times 100$;

Coefficient of M.D from Mo = $\frac{\text{MD from Mo}}{\text{Mo}} \times 100$.

Exercise

(1) a) 2000 − 500 = 1500

(b) 100 − 50 = 50

(2a)

X	$X - \bar{X}$
12	2.5
6	3.5
7	2.5
3	6.5
15	5.5
10	0.5
18	8.5
5	4.5
$\Sigma X = 76$	Σ 34

Here, n = 8, $\Sigma X = 76$

$\therefore \bar{X} = \frac{\Sigma X}{n} = \frac{76}{8} = 9.5$

$\therefore$ MD from $\bar{X} = \frac{\Sigma|X - \bar{X}|}{n}$

$= \frac{34}{8}$

$= 4.25$ #

Calculate Mean = Mean i.e $\frac{\text{Total}}{\text{number}}$

Shortcut

$\bar{X} = 9.5$, $\Sigma X a$

M.D = $\frac{(10+12+15+18) - (6+7+3+5)}{n}$

n → number of observations.

n_a = number of above mean
a = above mean value

M = mean
when equal

Date
Page No.

Note (1)

(1) $M.D = \dfrac{(\Sigma X_a - \Sigma X_b) - M(n_a - n_b)}{n}$

2(b) $\bar{X} = \dfrac{98 - (27) - 9(1-4)}{8}$ (8+8+8+3)

$= \dfrac{-9+27}{8}$

$= \dfrac{18}{8} = 2.25$.

Note no 2

* For any two observations a and b

$M.D = \dfrac{|a-b|}{2}$ = Half of range

Eg:- X : 8, 12

$\bar{X} = \dfrac{8+12}{2} = 10$

$M.D = \dfrac{\Sigma|X-\bar{X}|}{N}$

don't

$= \dfrac{|8-10| + |10-12|}{2}$

$= \dfrac{2+2}{2} = 2$#

Or alternatively

$M.D = \dfrac{|8-12|}{2} = 2$

For first n natural number

(i) $M.D = \dfrac{n^2-1}{4n}$ if n is odd

(ii) $M.D = \dfrac{n}{4}$ if n is even

Date
Page No.

* Find mean deviation of first 7 natural number.

X : 1, 2, 3, 4, 5, 6, 7

n = 7

$M.D = \dfrac{(7)^2-1}{4\times7}$

$= \dfrac{49-1}{28}$

$= \dfrac{48}{28} = \dfrac{12}{7}$ #

Find coefficient of mean deviation of first 10 natural number.

X : 1, 2, 3, ——, 10

n = 10 (even)

$M.D = \dfrac{n}{4} = \dfrac{10}{4} = 2.5$

$\bar{X} = \dfrac{n+1}{2} = \dfrac{10+1}{2} = 5.5$.

It's coeff = $\dfrac{M.D}{\bar{X}} \times 100$

$= \dfrac{2.5}{5.5} \times 100$

$= 45.45$.

For n observations in A.P

(i) $M.D = \dfrac{n^2-1}{4n} \times d$ if n is odd

(ii) $M.D = \dfrac{n}{4} \times d$ if n is even.

Find deviation from mean of the following data.

X: 13, 17, 21, 25, 29, 33

∴ Are in A.P

∴ n = 6 (even) d = 4

$$\therefore M.D = \frac{n}{4} \times d$$

Standard deviation

For a given set of observations standard deviation is square root of A.M of square of deviations of all observations from their mean. It is usually denoted by σ and standard deviation.

$$\sigma = \sqrt{\frac{\Sigma (x-\bar{x})^2}{n}}$$

where,

$\Sigma (X-\bar{X})^2$ = Sum of square of deviations taken from mean ($\bar{X}$)

$$\text{or } \sigma = \sqrt{\frac{\Sigma X^2}{n} - \left(\frac{\Sigma X}{n}\right)^2}$$

ΣX^2 = Sum of square of observations.

ΣX = Sum of observations.

$$\sigma^2 = \frac{\Sigma X^2}{n} - \bar{X}^2$$

$$\sigma^2 + \bar{x}^2 = \frac{\Sigma \bar{x}^2}{n}$$

$$\boxed{\therefore \Sigma X^2 = n(\sigma^2 + \bar{X}^2)}$$

Variance :- Mean of square of deviations of all observations, from mean, is variance

$$\text{Variance} = \frac{\Sigma (X-\bar{X})^2}{n}$$

$$\text{But } \sigma = \sqrt{\frac{\Sigma (X-\bar{X})^2}{n}}$$

$$\boxed{\therefore Var = \sigma^2}$$

$$\therefore \sigma = \sqrt{Var.}$$

Calculation:-

(A) Absolute measure :-

(I) $\sigma = \sqrt{\frac{\Sigma (X-\bar{X})^2}{n}}$ (Actual mean deviation method)

where, $\Sigma (X-\bar{X})^2$: Sum of square of deviations from mean.

(II) $\sigma = \sqrt{\frac{\Sigma X^2}{n} - \left(\frac{\Sigma X}{n}\right)^2}$ (Direct Method)

where, ΣX^2 = Sum of square of observations

ΣX = Sum of all observations.

(III) $\sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$ (short cut method)

where $d = X - A$: Deviations taken from assumed mean.

Σd = Sum of deviations.

Σd^2 = Sum of square of deviations.

(IV) $\sigma = \sqrt{\frac{\Sigma d'^2}{n} - \left(\frac{\Sigma d'}{n}\right)^2} \times h$

Date
Page No.

where $d' = \frac{X - A}{h}$

h : Common factor of X.

(B) For discrete or continuous series.

(I) $\sigma = \sqrt{\frac{\Sigma f(X - \bar{X})^2}{N}}$ (Actual mean deviation method)

(II) $\sigma = \sqrt{\frac{\Sigma f x^2}{N} - \left(\frac{\Sigma f x}{N}\right)^2}$ (Direct Method)

(III) $\sigma = \sqrt{\frac{\Sigma f d^2}{N} - \left(\frac{\Sigma f d}{N}\right)^2}$ (Short cut Method)

(IV) $\sigma = \sqrt{\frac{\Sigma f d'^2}{N} - \left(\frac{\Sigma f d'}{N}\right)^2} \times h$ where $d' = \frac{X - A}{h}$

h: Common factor of X or Class size.

Relative measure

Coefficient of variation is given by
$C.V = \frac{\sigma}{\bar{X}} \times 100\%$

it is expressed in percentage.
C.V is a standard deviation if average of observation is 100.

Date
Page No.

It is used to ~~measure~~ compare the variability of two or more data.

A distribution having less C.V is considered
(I) Less variation
(II) Less inequality
(III) more homogenous
(IV) more equitable
(V) more consistency
(VI) more stable
(VII) more uniform

$\left.\begin{matrix} \\ \\ \\ \\ \end{matrix}\right\}$ less

Notes:-

(1) For any two observations a and b, S.D is given by $\sigma = \frac{|a-b|}{2}$ = half of range.
Example :- Find S.D of 10 and 6.
$\sigma = \frac{10-6}{2} = 2$

(2) For 1^{st} n natural number S.D is given by
$\sigma = \sqrt{\frac{n^2-1}{12}}$ i.e $n = \sqrt{12\sigma^2 + 1}$

Example :- Find S.D of first seven natural number. $n = 7$
$\sigma = \sqrt{\frac{n^2-1}{12}} = \sqrt{\frac{7^2-1}{12}} = \sqrt{\frac{48}{12}}$
$= \sqrt{4}$
$= 2$

(3) For n observations in A.P, SD is given by
$\sigma = \sqrt{\frac{n^2-1}{12}} \times d$

$\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100\%$

Date
Page No.

Example :- Find variance of given observation.
X : 23, 27, 31, 35, 39 are in A.P.

Soln :-

$n = 5$

$d = 4$

$$\sigma = \sqrt{\frac{n^2-1}{12}} \times d = \sqrt{\frac{5^2-1}{12}} \times 4$$

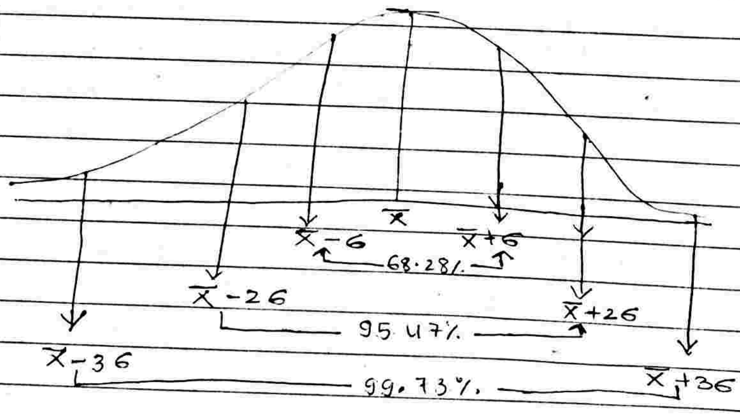
$$= \sqrt{\frac{24}{12}} \times 4$$

$$= 4\sqrt{2} \#$$

$\therefore$ Variance $= \sigma^2$

$$= (4\sqrt{2})^2$$

$$= 16 \times 2 = 32 \#$$



(I) $[\bar{x} \pm \sigma]$ includes central 68.28% observation (aprox)
(II) $[\bar{x} \pm 2\sigma]$ includes central 95.47% observation (approx)
(III) $[\bar{x} \pm 3\sigma]$ includes " 99.73% o " "

⑤ Relationship between QD, M.D, SD and Range.

Date
Page No.

For symmetrical distribution or moderately asymmetrical distribution.

(I) $6QD = 5MD = 4SD$

a) $QD : MD : SD = 10 : 12 : 15$

b) $QD < MD < SD$

c) $QD = \frac{2}{3}\sigma$

d) $MD = \frac{4}{5}\sigma$

(II) Range $= 6\sigma$ (approx)

Qno 5) Calculate SD
20, 15, 19, 24, 16, 14.

x	$d = x-19$	d^2
20	1	1
15	-4	16
19	0	0
24	5	25
16	-3	9
14	-5	25
	-6	76

$n = 6$, $\Sigma d = -6$, $\Sigma d^2 = 76$

$$\therefore \sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$$

$$= \sqrt{\frac{76}{6} - \left(\frac{-6}{6}\right)^2}$$

$$= \sqrt{12.6 - 1}$$

$$= \sqrt{11.6}$$

$$= 3.40 \#$$

Date
Page No.

(17)

Class	F	Midvalues (x)	$d' = \frac{X-23}{5}$	Fd'	Fd'^2
1-5	7	3	-4	-28	112
6-10	10	8	-3	-30	90
11-15	16	13	-2	-32	64
16-20	32	18	-1	-32	32
21-25	24	23	0	0	0
26-30	18	28	1	18	16
31-35	10	33	2	20	40
36-40	5	38	3	15	45
41-45	1	43	4	4	16
	N=123			-65	415

$A = 23, \quad h = 5, \quad N = 123$

$\Sigma Fd' = -65$

$$\bar{X} = A + \frac{\Sigma Fd'}{N} \times h$$

$$= 23 + \frac{-65}{\cancel{125}\ 25} \times \cancel{5}$$

$$= 20.4$$

$$\sigma = \sqrt{\frac{\Sigma Fd'^2}{N} - \left(\frac{\Sigma Fd'}{N}\right)^2} \times h$$

$$= \sqrt{\frac{415}{123} - \left(\frac{-65}{123}\right)^2} \times 5 \qquad \frac{4225}{15129}$$

$$= \sqrt{3.37 - 0.27} \times 5$$

$$= 9$$

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

Date
Page No.

$$= \frac{9}{20.4} \times 100$$

$$= 44.1\%$$

(1) (a)

X	500	700	900	1200	1500	1800	2000
F	40	90	111	125	75	60	20

Range $= 2000 - 500$
$= 1500$

Relate coefficient of range = $\frac{\text{Absolute range}}{\text{Sum of extreme values}}$

$$= \frac{1500}{2500}$$

$$= 0.6$$

(b) Range $= \cancel{100}/\cancel{50}$ $\quad S = 49.5$
$L = 100.5$

$R = L - S$

$\cancel{= 100.5 - 49.5}$ $= 100 - 50$
$= 50$.

Coeff $= \frac{50}{100+50}$

$$= \frac{50}{150} = 0.2$$

Date
Page No.

2b)

X	$\lvert X-\bar{X}\rvert$
9	0
3	6
8	1
8	1
9	0
8	1
9	0
18	9
72	Σ 18

$\therefore \bar{X} = \frac{\Sigma X}{n} = \frac{72}{8} = 9$

M.D from $\bar{X} = \frac{\Sigma\lvert X-\bar{X}\rvert}{n}$

$= \frac{18}{8}$

$= 2.25$ #

3)

X	F	c.f	FX	$\lvert X-\bar{X}\rvert$	$f\lvert X-\bar{X}\rvert$	$\lvert X-Md\rvert$	$f\lvert X-Md\rvert$
5	8	8	40	24	192	20	160
15	12	20	180	14	168	10	120
25	10	30	250	4	40	0	0
35	8	38	280	6	48	10	80
45	3	41	135	16	48	20	60
55	2	43	110	26	52	30	60
65	7	50	455	36	252	40	280
	N=50		$\Sigma fX=$ 1450	$\Sigma\lvert X-\bar{X}\rvert$ =126	$f\lvert X-\bar{X}\rvert$ =800		$\Sigma f\lvert X-Md\rvert$ = 760

$A.M = \frac{\Sigma fX}{\Sigma f} = \frac{1450}{50} = 29$

Date
Page No.

For median, $\frac{N+1}{2} = \frac{50+1}{2} = \frac{51}{2} = 25.5$

The c.f just greater than 25.5 is 30 so, the median $M_d = 25$

Mean deviation taken from median $= \frac{1}{N}\Sigma f\lvert X-M_d\rvert$

$= \frac{1}{50}\times 760$

$= 15.2$

M.D about mean $= \frac{1}{N}\Sigma f\lvert X-\bar{X}\rvert = \frac{1}{50}\times 800$

$= 16$ #

4)

C.I	F	X	FX	$\lvert X-\bar{X}\rvert$	$F\lvert X-\bar{X}\rvert$
2-4	20	3	60	2.6	52
4-6	40	5	200	0.6	24
8-6	30	7	210	1.4	42
8-10	10	9	90	3.4	34
	N=100		$\Sigma fX=560$	$\Sigma 8$	$\Sigma=152$

$A.M = \bar{X} = \frac{\Sigma fX}{\Sigma f} = \frac{560}{100} = 5.6$

M.D from mean $= \frac{1}{N}\Sigma f\lvert X-\bar{X}\rvert$

$= \frac{1}{100}\times 152$

$= 1.52$ #

(6)

X	d = X − 140	d^2
120	−20	400
180	40	1600
140	0	0
150	10	100
160	20	400

$\Sigma d = 50 \quad \Sigma d^2 = 2500$

$n = 5 \quad \Sigma d = 50 \quad \Sigma d^2 = 2500$

$$\therefore \sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}$$
$$= \sqrt{\frac{2500}{5} - \left(\frac{50}{5}\right)^2}$$
$$= \sqrt{500 - 100}$$
$$= \sqrt{400}$$
$$= 20$$

If equal amount of the is added or substracted from the data then S.D doesn't get affected, it remains same. i.e.

(b) S.D = 20 (c) SD = 20

If the values are divided by 10 then $SD = \frac{20}{10} = 2$ #

(7)

X	F	d = X − 2	Fd	Fd^2
0	1	−2	−2	4
1	7	−1	−7	7
2	10	0	0	0
3	6	1	6	6
4	1	2	2	4
	$\Sigma f = 25$		$\Sigma Fd = -1$	$\Sigma Fd^2 = 21$

Now, standard deviation $SD = \sqrt{\frac{1}{N}\Sigma fd^2 - \left(\frac{\Sigma fd}{N}\right)^2}$

$$= \sqrt{\frac{1}{25} \times 21 - \left(\frac{-1}{25}\right)^2}$$
$$= \sqrt{0.84 - (0.04)^2}$$
$$= \sqrt{0.84 - 0.0016}$$
$$= \sqrt{0.8384}$$
$$= 0.91$$

Variance $= \sigma^2 = (0.91)^2$
$= 0.83$ #

Calculation of S.D

(8)

Duration (in sec)	No of calls	Mid point (X)	$d' = \frac{X-105}{30}$	fd'	fd'^2
0 − 30	6	15	−3	−18	54
30 − 60	12	45	−2	−24	48
60 − 90	18	75	−1	−18	18
90 − 120	20	105	0	0	0
120 − 150	15	135	1	15	15
150 − 180	20	165	2	40	80
180 − 210	9	195	3	27	81
	N = 100			$\Sigma fd' = 22$	$\Sigma fd'^2 = 296$

$$\sigma = \sqrt{\frac{1}{N}\Sigma fd'^2 - \left(\frac{\Sigma fd'}{N}\right)^2} \times h$$
$$= \sqrt{\frac{1}{100} \times 296 - \left(\frac{22}{100}\right)^2} \times 30$$
$$= \sqrt{2.96 - 0.0484} \times 30$$
$$= \sqrt{2.9116} \times 30$$
$$= 1.706 \times 30$$
$$= 51.19 \text{ #}$$

Since C.V of regular MBS is higher so regular MBS.

Date
Page No.

(10) For regular MBS

X	$d = x - \bar{x}$	d^2
23	1.8	~~3.24~~
29	4.2	~~3.24~~ 17.64
27	2.2	4.84
22	−2.8	7.84
24	−0.8	0.64
21	−3.8	14.44
25	0.2	0.04
26	1.2	1.44
27	2.2	4.84
24	0.8	0.64
$\Sigma X = 248$	$\Sigma d = 3.4$	$\Sigma d^2 = 37.96$ 52.36

We have $\bar{X} = \frac{\Sigma X}{n} = \frac{248}{10} = 24.8$

$$\sigma = \sqrt{\frac{d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2} = \sqrt{\frac{\cancel{37.96}\ 52.36}{10} - \left(\frac{\cancel{52.36}\ 3.4}{10}\right)^2}$$

$$= \sqrt{5.236 - 0.1156}$$

$$= 2.26$$

$$C.V = \frac{\sigma}{\bar{x}} \times 100 = \frac{2.26}{24.8} \times 100 = 9.12$$

For evening MBS

X	X^2
27	729
34	1156
30	900
29	841
28	748
30	900
34	1156
35	1225
28	748
29	841
$\Sigma X = 304$	$\Sigma X^2 = 9224$

$\bar{X} = \frac{\Sigma X}{n} = 30.4$

$$S.D\ \sigma = \sqrt{\frac{1}{n}\Sigma X^2 - (\bar{X})^2}$$

$$= \sqrt{\frac{9224}{10} - (30.4)^2}$$

$$= 924.4 - 924.16$$

$$= 0.48$$

$$C.V = \frac{\sigma}{\bar{X}} \times 100 = 8.2$$

Date
Page No.

(9)

Bonus (Rs) (X)	Salary group	Tally Bars	No of employees (f)	$d' = \frac{X - A}{5}$ (A = 20)	fd'	fd'^{2}
10	61 − 75	III	3	−2	−6	12
15	76 − 90	IIII	4	−1	−4	4
20	91 − 105	~~IIII~~	5	0	0	0
25	106 − 120	IIII	5	1	5	5
30	121 − 135	~~IIII~~ II	7	2	14	28
35	136 − 150	~~IIII~~ I	6	3	18	54
			N = 30		27	103

Here,
A = 20, h = 5, N = 30,
$\Sigma fd = 27$ $\Sigma fd'^2 = 103$

Average bonus:-

$$\bar{X} = A + \frac{\Sigma fd'}{N} \times h$$

$$= 20 + \frac{27}{30} \times 5$$

$$= 20 + 4.5$$

$$= 24.5$$ Ans

Standard deviation is,

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times h$$

$$= \sqrt{\frac{103}{30} - \left(\frac{27}{30}\right)^2} \times 5$$

$$= \sqrt{3.43 - 0.81} \times 5$$

$$= \sqrt{2.62} \times 5$$

$$= 1.61 \times 5$$

$$= 8.09$$ Ans

Date
Page No.

(10) Given,

$n = 200$; $\bar{X} = 40$, $\sigma = 12$

Incorrect observation : {23, 15}

Correct observation = {43, 18}

$\Sigma X = n\bar{X} = 200 \times 40 = 8{,}000$

and $\Sigma X^2 = n(\sigma^2 + \bar{X}^2)$

$= 200\,(12^2 + 40^2)$

$= 348800$

After Correction:-

$\Sigma X_c = 8000 - 23 - 15 + 43 + 18 = 8023$

$\Sigma X_c^2 = 348800 - 23^2 - 15^2 + 43^2 + 18^2$

$= 348800 - 529 - 225 + 1849 + 324$

$= \cancel{346521}\ \# = 350219$

$n_c = 200 - 2 + 2 = 200$

$$\sigma = \sqrt{\frac{\Sigma X_c^2}{n_c} - \left(\frac{\Sigma X}{n}\right)^2}$$

$$= \sqrt{\frac{348800}{200} - \left(\frac{8023}{200}\right)^2}$$

$$= \sqrt{1744 - 1609.21}$$

$= 11.91$ #

Date
Page No.

Common properties	Range	Q.D	M.D	S.D
1. If all the observation assumed by variable are constant then it is zero.	✓	✓	✓	✓
2. It is independent on change of origin.	✓	✓	✓	✓
3. It is dependent due to change of scale in a constant porportion.	✓	✓	✓	✓
4. If each observations is multiplied or divided by a constant value then is is also multiplied or divided by absoute value of that constant	✓	✓	✓	✓
5. If X & Y are linerally related by $Y = a + bX$ where a and b are constant then ~~$R_y = \|b\| R_x$~~	$R_y = \|b\| R_x$	$QD_y = \|b\| QD_x$	$MD_y = \|b\| MD_x$	$SD_y = \|b\| SD_x$

Mean Deviation (Additional Property)

· Mean Deviation from median is minimum.

i.e M.D from Me $= \dfrac{\Sigma\,|X - Me|}{n}$ is min.

Date
Page No.

~~7~~6th Property (Standard Deviation)

If n_1 and n_2 be the number of observations of 2 groups. $\bar{X}_1$ & $\bar{X}_2$ are their respective mean. σ_1 and σ_2 are their ~~σ_1~~ respective standard deviation then combined standard deviation of the two groups taken together is given by.

$$\text{Combined } \sigma = \sqrt{\frac{n_1(\sigma_1^2+d_1^2)+n_2(\sigma_2^2+d_2^2)}{n_1+n_2}} \quad \text{--- (iii)}$$

where :- $\left.\begin{array}{l} d_1^2 = (\bar{X}_1 - \bar{X})^2 \\ \text{and } d_2^2 = (\bar{X}_2 - \bar{X})^2 \end{array}\right\} \longrightarrow \text{(ii)}$

$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2} \quad \text{--- (i)}$$

Q no 13 Here,

$n = 100$

$\bar{X} = 8$

$\sigma = \sqrt{10.5}$

$n_1 = 50$; $\bar{X}_1 = 10$, $\sigma_1 = 2$

$n_2 = 50$, $\bar{X}_2 = ?$ $\sigma_2^2 = ?$

$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1+n_2} \quad \text{--- (i)}$$

$$8 = \frac{50\times10 + (50\times\bar{X}_2)}{50+50}$$

$800 = 500 + 50\bar{X}_2$

$300 = 50\bar{X}_2$

$6 = \bar{X}_2$

Date
Page No.

Now,

$d_1^2 = (\bar{X}_1 - \bar{X})^2 = (10-8)^2 = 4$

$d_2^2 = (\bar{X}_2 - \bar{X})^2 = (6-8)^2 = 4$

$$\text{Combined } \sigma = \sqrt{\frac{n_1(\sigma_1^2+d_1^2)+n_2(\sigma_2^2+d_2^2)}{n_1+n_2}}$$

$$\text{or, } \sqrt{10.5} = \sqrt{\frac{50(2^2+4)+50(\sigma_2^2+4)}{100}}$$

Squaring on both sides,

$$10.5 = \frac{400 + 50\sigma_2^2 + 200}{100}$$

or, $1050 - 600 = 50\sigma_2^2$

$\therefore \sigma_2^2 = \frac{450}{50} = 9$

$\therefore$ Variance $= 9$ //.

Properties of C.V

1. If each observations is increased ~~or decreased~~ by a constant (positive) then their C.V gets reduced.
2. If each observations is decreased by a positive constant then their C.V gets increased.
3. If each observation is multiplied or divided by a positive constant then their C.V remains unchanged.
4. $\boxed{CV_a(\bar{X}+a) = CV_b(\bar{X}-b) = 100\sigma}$

Date
Page No.

~~Not~~

Example :- If each observations is increased by a constant value 10 then their C.V will be 30%. and if each observations is decreased by a constant 15, then their CV will be 60%. Find ~~origin~~ mean and S.D of original observations.

Soln:-

Let $\bar{X}$ and σ be the mean & S.D of original observations respectively.

$a = 10$, $CV_a = 30\%$

$b = 15$, $CV_b = 60\%$

$\therefore CV_a(\bar{X}+a) = CV_b(\bar{X}-b) = 100\sigma$

$30\%.(\bar{X}+10) = 60\%(\bar{X}-15)$

$30\bar{X}+300 = 60\bar{X}-750$

$\bar{X} = 40$

$CV_a \times (\bar{X}-a) = 100\sigma$

$30(\bar{X}+10) = 100\sigma$

$\sigma = \frac{30(40+10)}{100}$

$\sigma = 15.$

Notes :-

(1) (i) $R(a+bX) = |b| R_X$ — origin, scale

(ii) $QD(a+bX) = |b| QD_X$

(iii) $MD(a+bX) = |b| MD_X$

(iv) $SD(a+bX) = |b| SD_X$

Exampe :- Find SD of $\frac{ax+b}{c}$ where a, b, c constant.

$SD\left(\frac{ax+b}{c}\right) = SD\left(\frac{ax}{c}+\frac{b}{c}\right)$

$= \left|\frac{a}{c}\right| \cdot \sigma_x$

Date
Page No.

* Find the variance of $\frac{2-3x}{5}$ if SD of x is 5.

$V = (\sigma)^2 \left[SD\left(\frac{2-3x}{5}\right)\right]^2 = \left[SD\left(\frac{2}{5}-\frac{3x}{5}\right)\right]^2$

$= \left(\frac{3}{5} SD x\right)^2$

$= \left(\frac{3}{\cancel{5}}\times\cancel{5}\right)^2$

$= 9$ #.

(2) For $ax+by+c=0$ / $ax = by+c$ / $ax-by=c$ — Take this

(i) $|a| R_x = |b| R_y$

(ii) $|a| QD_x = |b| QD_y$

(iii) $|a| MD_y = |b| MD_y$

(iv) $|a| SD_x = |b| SD_y$

Example :- Find QD of y if QD of x is 10 for $2x+5y=11$

$2x = 5y+11$

$|2| QD_x = |5| \cdot QD_y$

~~$\therefore QD_x =$~~

$2\times 10 = 5\, QD_y$

$QD_y = 4$ #.

Example no 2 :- Find coefficient of mean deviation from mean of y if mean and mean deviation (M.D) of x are 15 and 0.3 for $5x-3y=6$

Now, ~~$5x = 3$~~

Date
Page No.

$5x - 3y = 6$

$\bar{x} = 15$, $M \cdot D = 0.3$ Coeff of $MD_y = ?$

$= \frac{MD_y}{\bar{Y}} \times 100$

For mean

$5\bar{X} - 3\bar{Y} = 6$ Also,

$5 \times 15 - 3\bar{Y} = 6$

$75 - 6 = 3\bar{Y}$

$\bar{Y} = \frac{69}{3}$

$\bar{Y} = 23$

Also,

$5\,MD_x = 3\,MD_y$

$5 \times \cancel{0.3}^{0.1} = 3 \times MD_y$

$MD_y = 0.5$

For coeff of $M \cdot D$ =

$= \frac{0.5}{23} \times 100$

$= 2.17$

Example no 3

If $X: 101, 102, 103, ---, ---, 150$ $\therefore \sigma_x = k$ (say)

$Y: 151, 152, 153, ---, ---, 200$ $\therefore \sigma_y = k$

Since X and Y added by same constant k then,

Find $\frac{V(X)}{V(Y)} = ? = \left(\frac{\sigma_x^2}{\sigma_y^2}\right) = \left(\frac{\sigma_x}{\sigma_y}\right)^2 = \left(\frac{k}{k}\right)^2 = 1$

Since in A.P $\therefore \sigma_x = \sqrt{\frac{n^2-1}{12}} \times d$

$= \sqrt{\frac{50^2-1}{12}} \times 1$

Combined mean - Imp

Date
Page No.

$X: 25, 30, 42, 50, 70$

$Y: 54, 64, 88, 104, 144$

Find $\sigma_x : \sigma_y = 1:2$

Soln:-

$y = a + \textcircled{b}x$

slope $\frac{y_2 - y_1}{x_2 - x_1}$

$m = \frac{64 - 54}{30 - 25} = \frac{10}{5} = 2$

$Y = 4 + 2x$ $\quad Y = (2x + 4)$ $(2x + 4)$

$\frac{\sigma_x}{\sigma_y} = \frac{\sigma_x}{2\sigma_x} = \frac{1}{2}$

$\frac{\sigma_x}{\sigma_y} = \frac{\sigma_x}{2\sigma_x} = \frac{1}{2}$

Note

* Combined S.D is given by,

(i) $\sigma = \sqrt{\frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2} + \frac{n_1 n_2}{n^2}(\bar{X}_1 - \bar{X}_2)^2}$

(ii) If $\bar{X}_1 = \bar{X}_2$

$\sigma = \sqrt{\frac{n_1\sigma_1^2 + n_2\sigma_2^2}{n_1 + n_2}}$

Question no 14 (200)

$n = 250$

Date
Page No.

(15)

BBS, CMAT	Consituents
$n_1 = 200$	$n_2 = 300$
$\bar{X}_1 = 70$	$\bar{X}_2 = 60$
$\sigma_1 = 30$	$\sigma_2 = 30$

$$\sigma = \sqrt{\frac{200 \times (30)^2 + 300 \times (30)^2}{500} + \frac{200 \times 300\,(70-60)^2}{(500)^2}}$$

$$= \sqrt{\frac{180000 + 27000}{500} + \frac{60000}{250000} \times 100}$$

$$= \sqrt{\frac{4500\cancel{00}}{5\cancel{00}} + 24} = \sqrt{924} = 30.39 \#$$

Characteristics of a good measure of dispersion

1. It should be rigidly defined.
2. It should be easy to compute & simple to understand.
3. It should be based on all the observations
4. It should satisfy mathematical property or it should be suitable for further mathematical treatment.
5. It should be least affected by extreme observations.
6. It should be least affected by sampling fluctation.
7.

Date
Page No.

Merits / Demerits	Range	QD	M.D	S.D
(1) It is ridgidly defined.	✓	✓	✓	✓
(2) It is easy to compute and simple to understand.	(1) Easiest	(2) ✓	(3)	Difficult to compute & complex to understand
(3) It is based on all the observation.	✗ only two extreme values	✗ extreme value of central 50%.	✓	✓
(4) It satisfies mathematical property / treatment.	✗	✗	✗	✓
(5) It is least affected by extreme observations.	✗ (4) Highly affected	✓ (1) Not affeted at all	✓ (2)	✓ (3)
(6) It is least affected by sampling fluctation	✗ Highly affected	✗ affted significantly	(2)	(1)
(7) It can be calculated for the distribution with open-type classes.	✗	✓	✗	✗

Application for measure of disperlion.

Application of Range
- In explaining metrological information like temperature, rainfall e.t.c.
- In stastistical quality control.

Application of Quartile Deviation
* Under the following cases, quartile deviation is the most appropriate measure of dispersion.

(i) In missing some extreme observation.

(ii) In the distribution with open type classes.
(iii) In highly asymetrical distribution.
(iv) In the distribution with varying class size.

Mean deviation (Application)

(i) It is an improved measure of dispersion than range and quartile deviation because mean deviation is based on all the given observations, but range and QD are not.
(ii) Mean deviation has also a serious demirits that it doesn't satisfy the matematical property and can't be suitable for further mathematical anyalysis because it directly igoners negative sign of deviation.
As compared to mean deviation standard deviation is improved measure and it's all application is replaced by standard deviation because standard deviation can be treated algebrically.

Application of Standard Deviation

- It is the best measure of dispersion and is commonly used measure.
-

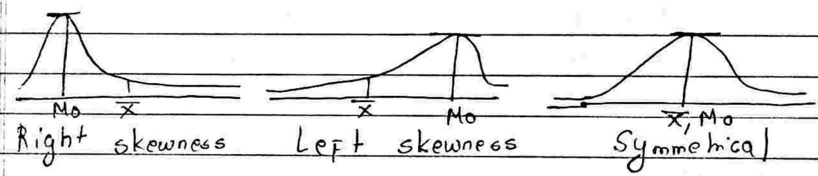
Skewness & kurtosis

Measure of skewness

Literal meaning of "skewness" is "Lack of symmetric".

Skewness

A stastistical measure which is used to measure the elongation of frequency curve about mode wheather it is elongated at right side or left side of the frequency curve, is known as skewness.



Methods of measuring skewness.

(A) karl pearson's measure of skewness (Best)
(B) Bowley's measure of skewness. (Best for open type classified)
(C) kelly's measure of skewness
(D) Measure of skewness based on moments

Date
Page No.

(19) Market X — Market Y

X_1	$d_1 = X_1 - 9.25$	d_1^2	X_2	$d_2 = X_2 - 9.50$	d_2^2
8.00	-1.25	1.5625	8.50	-1	1
8.50	-0.75	0.5625	8.50	-1	1
9.00	-0.25	0.0625	9.25	-0.25	0.0625
9.25	0	0	9.50	0	0
9.25	0	0	9.50	0	0
10.00	~~8.25~~ 0.75	0.5625	9.50	0	0
10.00	0.75	0.5625	10.75	1.25	1.5625
	$\Sigma d_1 = -0.75$	$\Sigma d_1^2 = 3.3125$		$\Sigma d_2 = -1$	$\Sigma d_2^2 = 3.625$

For Mr. A,

$$\text{Mean } \overline{X_1} = A + \frac{\Sigma d_1}{n} = 9.25 + \frac{(-0.75)}{7}$$

$$= 9.14$$

$$S.D\ \sigma = \sqrt{\frac{1}{n}\Sigma d_1^2 - \left(\frac{\Sigma d_1}{n}\right)^2}$$

$$= \sqrt{\frac{1}{7}3.3125 - \left(\frac{-0.75}{7}\right)^2}$$

$$= \sqrt{\cancel{1.5675}\ 0.47 - 0.0018}$$

$$= \cancel{1.251} = 0.68$$

$$(CV)_A = \frac{\sigma_A}{\overline{X}_A} \times 100$$

$$= \frac{\cancel{1.251}\ 0.68}{9.14} \times 100\ \%$$

$$= \cancel{13.69\%} = 7.48$$

For Marke Y

$$\text{Mean } \overline{X_1} = A + \frac{\Sigma d_2}{n} = 9.50 + \left(\frac{-1}{7}\right)$$

$$= 9.22$$

$$S D\ \sigma = \sqrt{\frac{1}{n}\Sigma d_2^2 - \left(\frac{\Sigma d_2}{n}\right)^2}$$

$$= \sqrt{\frac{1}{7}(3.625)^2 - \left(\frac{3.625}{7}\right)^2}$$

$$= \sqrt{0.51 - 0.019}$$

$$= 0.70$$

$$(CV)_A = \frac{\sigma_A}{\overline{X}_A} \times 100\ \%$$

$$= \frac{0.70}{9.36} \times 100\ \%$$

$$= 7.59$$

Since $(CV_Y) > (CV_X)$ variation of price in market Y is higher.

(20) Here, $n_1 = 100$ Number of wages earners by Factory A
$n_2 = 80$ " " B.
Average weekly wages $(\overline{X}_1) = 460$
Average weekly $(\overline{X}_2) = 490$
$\sigma_A = 50$
$\sigma_B = 40$
Now,

$$\overline{X} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}$$

$$= \frac{100 \times 460 + 80 \times 490}{100 + 80} = 473.33.$$

Date
Page No.

The combined mean is 473.33.

Again $\sigma = \sqrt{\dfrac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}}$

$d_1 = \bar{x}_1 - \bar{x} = 460 - 473.33 = -13.33$
$d_2 = \bar{x}_2 - \bar{x} = 490 - 473.33 = 16.67$

Now,

$\sigma = \sqrt{\dfrac{100(\cancel{50}^{2500} + 177.68) + 80(\cancel{40}^{1600} + 277.88)}{100 + 80}}$

$= \sqrt{\dfrac{\cancel{22768.0}^{267768} + 150230.4\cancel{25431.11}2}{180}}$

$= \sqrt{\dfrac{25299.112}{180}}$

$= \cancel{11.85}$ 48.18% #

(b) $C.V_A = \dfrac{\sigma}{\bar{x}} \times 100\%$

$= \dfrac{50}{460} \times 100\%$

$= 10.86\%$

$CV_B = \dfrac{80}{490} \times 100\%$

$= 16.32\%$

Since $CV_B > CV_A$ we can say that variability of B is higher.

(a) $460 \times 100 = 46000$ and $490 \times 80 = 39200$ #
Since average weekly wages of Factory A is higher it pays larger amount of weekly wages.

Date
Page No.

(21) Here,

$CV_1 = 55\%$ $\qquad$ $CV_2 = 70\%$
$\sigma_1 = 22$ $\qquad$ $\sigma_2 = 15.40$
$n = 80\%$ $\qquad$ $n_2 = 20\%$

Now,
$CV_1 = \dfrac{\sigma_1}{\bar{x}} \times 100\%$

$55\% = \dfrac{22}{\bar{x}} \times 100\%$

$\bar{x} = 40$

Now,
$CV_2 = \dfrac{\sigma_2}{\bar{x}_2} \times 100\%$

$70\% = \dfrac{15.40}{\bar{x}_2} \times 100\%$

$\bar{x}_2 = 22$

Now,
$\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

$\bar{x} = \dfrac{80 \times 40 + 20 \times 22}{100}$

$\therefore \bar{x} = 36.4$

$d_1^2 = (\bar{x} - \bar{x})^2 = (36.4 - 40)^2$
$= 12.96$
$d_1^2 = (\bar{x}_2 - \bar{x})^2 = (36.4 - ...)$
$= 207.36$

Now,
$(\sigma)^2 = \dfrac{n_1(\sigma_1^2 + \sigma_1^2) + n_2(\sigma_2^2 + \sigma_2^2)}{n_1 + n_2}$

$= \dfrac{80 \times (22^2 + 12.96) + 20(15.40^2 + 207.36)}{100}$

$(\sigma)^2 = 486.47$ #

$\Sigma x = \bar{x}$.

(12) Here,

$n = 20$

$\bar{X} = 20$ cm

$\sigma = 5$ cm

or, $\Sigma X^2 = n(\sigma^2 + \bar{X}^2)$

$\Sigma X^2 = 20(25 + 400)$

$\Sigma X^2 = 8500$

$X_c = 400 + 13 - 30$

$= 383$

$X_c^2 = 8500 + (13)^2 - (30)^2$

$= 8500 + 169 - 900$

$= 7769$

$N = 20 + \cancel{1} - \cancel{1}$

$= 20.$

$\sigma = \sqrt{\frac{X_c^2}{n} - \left(\frac{X_c}{n}\right)^2}$

$= \sqrt{\frac{7769}{20} - \left(\frac{383}{20}\right)^2}$

$= \sqrt{388.45 - 366.7}$

$= \sqrt{21.72}$

$= 4.6613$

$\bar{X} = \frac{X_c}{N}$

$= \frac{383}{20}$

$= .$

(16) Here,

$\Sigma f = N = 25$

$\Sigma fx = 230$

$\Sigma fx^2 = 2260$

$\bar{x} = \frac{\Sigma fx}{N}$

or, $\bar{x} = \frac{230}{25}$

$\therefore \bar{x} = 9.2$

Now,

$\sigma = \sqrt{\frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx}{N}\right)^2}$

$\sigma = \sqrt{\frac{2260}{25} - (9.2)^2}$

$\sigma = 4.6648$

Now,

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

$= 50.70$ //

(14) Here,

$\bar{X} = 560$

$CV = 54.97$

$N = 250$

$\sigma = 74.14$

For city A

$N_1 = 100$

$\bar{x}_1 = 650$

$CV_1 = 14$

$\sigma = 11$

For city B,

$N_2 = 150$

$\bar{x}_2 = ?$

$CV_2 = ?$

And,

$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2}$

$560 = \frac{100 \times 650 + 150 \times \bar{X}_2}{250}$

or, $140000 - 65000 = 150 \times \bar{X}_2$

$\bar{X}_2 = 500$

Now,

$d_1^2 = (\bar{X}_1 - \bar{X})^2$

$= (650 - 560)^2$

$= (90)^2$

$d_1^2 = 8100$

$d_2^2 = (\bar{X}_2 - \bar{X})^2$

$= (500 - 560)^2$

$= (-60)^2$

$d_2^2 = 3600$

Finally,

$\sigma = \sqrt{\frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}}$

$74.14 = \sqrt{\frac{100(11^2 + 8100) + 150(\sigma_2^2 + 3600)}{250}}$

$\sigma_2^2 = 80.57.$

$\therefore$ The variance of salary is 80.57 //.

(11) Given

Mean $(\bar{x}) = 80$

Standard deviation $(\sigma) = 12$

$n = 200$ ~~$n = 100$~~

Additional item $= 90$

$\therefore \Sigma n = n \cdot \bar{X} = 200 \times 80 = 16,000$

and

$\Sigma x^2 = n(\sigma^2 + \bar{X}^2)^2$

$= 200(144 + 10000)^2$

Date
Page No.

$= 200\,(6544)$
$= 1308800$

After addition of observation,
$\Sigma X_c = 16000 + 70 = 16070$

$\Sigma X_c^2 = 1308800 + 70^2$
$= 1308800 + 4900$
$= 1313700$
$n = 200 + 1 = 201.$

Correct value of SD is,
$$\sigma_c = \sqrt{\frac{\Sigma X_c^2}{n_c} - \left(\frac{X_c}{n_c}\right)^2}$$
$$= \sqrt{\frac{1313700}{201} - \left(\frac{16070}{201}\right)^2}$$
$$= \sqrt{6535.82 \cdot 6392.04}$$
$$= \sqrt{143.77}$$
$$= 11.99$$

$$\bar{X} = \frac{\Sigma X_c}{M} = \frac{16070}{201} = 79.95 \#$$

$$\bar{X} = \frac{\Sigma X c}{N.}$$
$=$

~~Excess~~

$\Sigma X = n\,\bar{X}$
$= 100 \times 80 = 8{,}000$
$\bar{X}_c = 8{,}000 + 70 = 8070.$
$N = 100 \quad N_c = 100 + 1 = 101.$
$$\bar{X}_c = \frac{\Sigma X c}{N} = \frac{8070}{101} = 79.90 \#.$$

Date
Page No.

Skewness & kurtosis

(A) Karl pearson's measure of skewness

(a) Absolute measure:-
Skewness $= \bar{X} - M_o$

(b) Relative measure:-
Coefficient of skewness is,
$$Sk_p = \frac{\bar{X} - M_o}{\sigma}$$

If mode is ill-defined
(A) Absolute measure:-
Skewness $= \bar{X} - M_e$

(B) Relative measure
Coefficient of skewness is,
$$Sk_p = \frac{3(\bar{X} - M_e)}{\sigma}$$

Note:-
(i) $-3 \leq Sk_p \leq 3$
(ii) ~~For a symm~~
Conclusion:-
(i) If $Sk_p = 0 \Rightarrow$ Symmetrical
(ii) If $Sk_p > 0 \Rightarrow$ Right skewed
(ii) If $Sk_p < 0 \Rightarrow$ Left skewed

(4) Hints
Simple frequency distribution is

Date
Page No.

Class marks	No of students
0-10	6
10-20	10
20-30	25
30-40	9
40-50	12
50-60	25
60-70	8
70-80	5

→ 25, 25 (two) so ill defined.
Here mode is ill defined so, $Sk_p = \frac{3(\bar{X} - Me)}{\sigma}$

Date
Page No.

(B) Bowley's measure of skewness :-
or
Measure of skewness based on quartiles :-

(i) Absolute measure

$Q_1 \longleftrightarrow Me \longleftrightarrow Q_3$

$$\text{Skewness} = (Q_3 - Me) - (Me - Q_1)$$

(ii) Relative measure

Coefficient of skewness is

$$Sk_B = \frac{(Q_3 - Me) - (Me - Q_1)}{(Q_3 - Me) + (Me - Q_1)}$$

$[-1 \leq Sk_B \leq 1]$

Conclusion :-
(i) $Sk_B = 0 \Rightarrow$ Symmetrical
(ii) $Sk_B > 0 \Rightarrow$ Right skewed
(iii) $Sk_B < 0 \Rightarrow$ Left skewed.

(C) kelly's measure of skewness :-
or
Measure of skewness based on percentiles (Deciles)

(i) Absolute measure :-

$$\text{Skewness} = (P_{90} - P_{50}) - (P_{50} - P_{10})$$

P_{10} P_{50} P_{90}
$\downarrow$ $\downarrow$ $\downarrow$
D_1 D_5 D_9

(ii) Relative measure :

Coefficient of skewness is

$$Sk_K = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{(P_{90} - P_{50}) + (P_{50} - P_{10})}$$

$[-1 \leq Sk_K \leq 1]$

Conclusion
(i) $Sk_K = 0 \Rightarrow$ Symmetrical
(ii) $Sk_K > 0 \Rightarrow$ Right skewed
(iii) $Sk_K < 0 \Rightarrow$ Left skewed

Date
Page No.

Notes

Q no 1) For a symmetrical distribution:-

(a) $Q_3 + Q_1 = 2Me$

(b) $P_{90} + P_{10} = 2P_{50}$

2) For a positively (Right) Skewed

(a) $Q_3^2 + Q_1 > 2Me$

(b) $P_{90} + P_{10} > 2P_{50}$

3) For a negatively skewed distribution.

(a) $Q_3 + Q_1 < 2Me$

(b) $P_{90} + P_{10} < 2P_{30}$.

Measure of kurtosis :-

A stastical device which is used to measure the flatness or peakedness of a frequency curve, is known as kurtosis.
It is useful to test the normality of the curve.

Kinds of kurtosis:-

kurtosis can be classified into three types:-

(a) Mesokurtic (Normal curve)
(b) Leptokurtic (Peaked curve)
(c) Platykurtic (Flaten curve)

(a) Mesokurtic

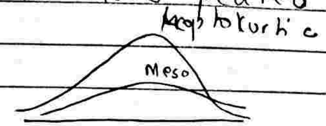
A frequency curve is known to be mesokurtic if the curve is neither so flat nor so peaked. It is a normal curve

Mesokurtic.

Date
Page No.

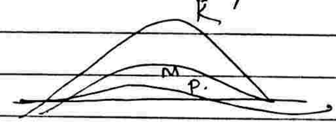
(b) Leptokurtic

A frequency curve is known to be leptokurtic if the curve is more peaked than the mesokurtic curve.



(c) Platykurtic

A frequency curve is known to be platykurtic if the curve is more flatkened than mesokurtic



Methods of measuring kurtosis.

(i) Percentile measure of kurtosis.
(ii) Measure of kurtosis based on moments

(i) Percentile measure of kurtosis

Coefficient of kurtosis is

$$k = \frac{QD}{\text{Percentile range}}$$

$$\text{i.e } k = \frac{Q_3 - Q_1}{2(P_{90} - P_{10})}$$

Conclusion :-

(i) $k = 0.263 \Rightarrow$ mesokurtic
(ii) $k > 0.263 \Rightarrow$ leptokurtic
(iii) $k < 0.263 \Rightarrow$ Platykurtic

Date
Page No.

(1)

X	f	FX	FX^2
1	1	1	1
3	2	6	18
5	6	30	150
7	3	21	147
9	1	9	81

$N = 13 \quad \Sigma fX = 67 \quad \Sigma fX^2 = 397$

Mean, $\bar{X} = \frac{\Sigma fX}{N} = \frac{67}{13} = 5.15$

SD $\sigma = \sqrt{\frac{1}{n}\Sigma fX^2 - \left(\frac{\Sigma fX}{N}\right)^2}$

$= \sqrt{\frac{1}{13} \times 397 - (5.15)^2}$

$= \sqrt{30.5 - 26.52}$

$= 1.99$

$\frac{\bar{x} - Mo}{\sigma}$

Since the maximum frequency is 6, mode = 5

Now, Person's coefficient of skewness $S_k(P) = \frac{5.15 - 5}{1.9}$

$S_k(P) = 0.078$

(2)

X	F	Mid point(m)	XF	FX^2
5-10	2	7.5	15	112.5
10-15	5	12.5	62.5	781.25
15-20	10	17.5	175	3062.5
20-25	4	22.5	90	2025
25-30	1	27.5	27.5	756.25
N=22			$\Sigma FX = 370$	6737.5 $\Sigma FX^2 = \cancel{4712.5}$

$\bar{X} = \frac{\Sigma FX}{N} = \frac{370}{22} = 16.81$

Date
Page No.

$SD = \sigma = \sqrt{\frac{1}{n}\Sigma fX^2 - \left(\frac{\Sigma FX}{N}\right)^2}$

$= \sqrt{\frac{4712.5}{22} - (16.81)^2}$

$= \sqrt{\cancel{214.20}\ 306.2 - 282.51}$

$= \sqrt{236}$

$= 4.86$

Mode = 10 i.e 15 - 20

$= A + \frac{10 - 5}{2 \times 10 - 5 - 4}$

$\times 10$

$= 15 + \frac{5}{20 - 5 - 4}$

$= 15 + 0.45$

$= 15.45$

(3) 45 5 5 150 10 15 25 35 20 100

$\bar{x} = 46 \quad Mo =$

(14) Given,

$S_{kp} = 0.4$

C.V = 30%, $Mo = 88$

$\bar{X} = ?$ $Me = ?$

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

$S_{kp} = \frac{\bar{X} - Mo}{\sigma}$

$0.4 = \frac{\bar{X} - 88}{\sigma}$

or, $\bar{X} - 88 = 0.4 \times \sigma$

or, $\bar{X} = 88 + 0.4\sigma$ —— (1)

Date
Page No.

$\therefore CV = \frac{6}{\bar{x}} \times 100$

$30 = \frac{6 \times 100}{88 + 0.46}$

$2640 + 126 = 6 \times 100$

$2640 = 100 6 - 126$

$2640 = 88 6$

$30 = 6$ #

$\therefore M_0 = 3M_e - 2\bar{x}$

$88 \cancel{M_0} = 3 \times M_e - 2 \times 88 + 0.4$

$88 = 3M_e - \cancel{188}\ 200$

$2\cancel{88} = 3M_e$

$96 = M_e$ #

(13) $\bar{x} = 65$ $M_e = 70$ $S_K = -0.6$

(15) $M_e - Q_1 = 5(Q_3 - M_e)$

$S_{KB} = ?$

$\therefore S_{KB} = \frac{(Q_3 - M_e) - (M_e - Q_1)}{(Q_3 - M_e) + (M_e - Q_1)}$

$= \frac{(Q_3 - M_e) - 5(Q_3 - M_e)}{(Q_3 - M_e) + 5(Q_3 - M_e)}$

$= \frac{-4\cancel{(Q_3 - M_e)}}{6\cancel{(Q_3 - M_e)}}$

$= \frac{-2}{3}$ #

Date
Page No.

Correlation analysis:-

A stastistical device which is used to measu degree of association (relationship) between two or more variables, is known as co-relation.

Types of co-relation

(1) Simple co-relation

It measures degree of association between two variables only.

(2) Partial co-relation

It measures the degree of relationship between one dependent variable and another independent variable keeping others independent variable constant.

(3) Multiple co-relation

It measures the joint effect of a number of idependent variables on a dependent variable.

(1)(i) → Linear and non-linear co-relation

Two variables are known to be linearly co-related if the values of the two variables are porportional to each other or the rate of change remains constant, otherwise they are known to be non-linearly co-related.

Here, Example:- X 10 15 8 24 30

Y 23 33 19 51 63

Rate of change of Y per unit change of X is

Date
Page No.

2.

$\therefore$ X & Y are linearly correlated. $Y = 2x + 3$.

Positive or direct co-relation

Two variables are known to be positively co-related if there values move together in the same direction.

For example :- ① Temperature and sales of cold drink.

② - Age and Premium of isurance.

③ Distance required to stop and speed of a vehicle after applying breaks.

④ Pig iron and soot content

Negative co-relation / Inverse co-relation

Two variables are known to be negatively co-relative if their value move together in opposite direction. Example:-

1. Unempolyment index and purchasing power of money.
2. Price and demand.
3. Distance travelled and speed of a vehicle after applying breaks.

Methods of studying co-relation.

① Scattered diagram method. (Graphical method)

② Kalpearson's corellation coefficient / ~~Kalpearson~~ Product moment co-relation coefficient.

③ Spearman's rank corellation coefficient

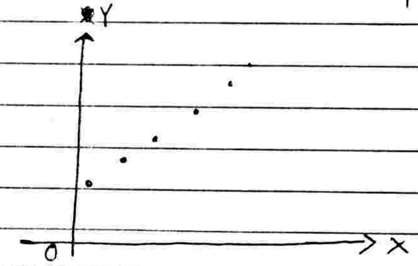
① Scattered diagram method

It is a simple graphic method of studing co-relation between two variables. Under this method values of one variable are taken in x-axis and the another is taken along Y-axis.

Date
Page No.

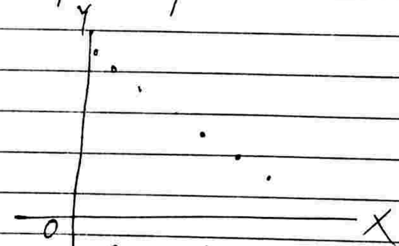
A number of pair variables are plotted on the graphs and we get a number of dots or points in a certain pattern on the graph calle scattered diagram by studying or inspecting pattern of scatterdness of the points. Nature and direction of the two variables can be stated which are explained below with diagram.

(i)



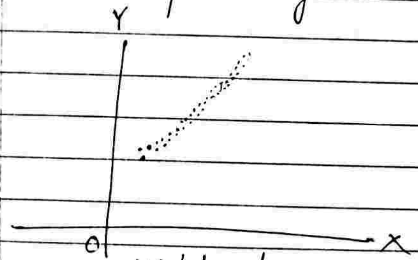
Perfect positive co-relation.

(ii)



Perfect negative co-relation.

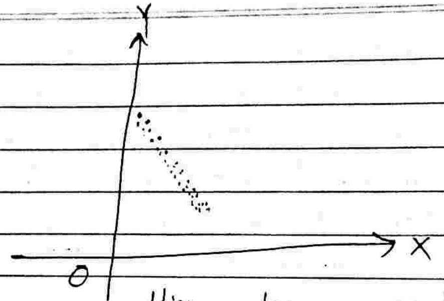
(iii)



Hight degree positive co-relation.

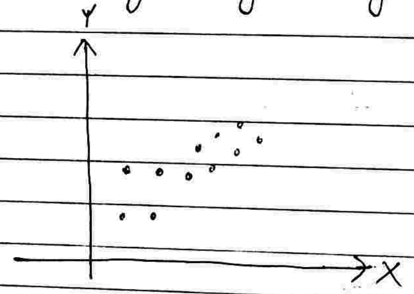
Date
Page No.

(iv)



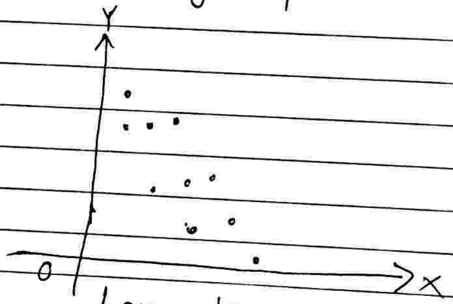
High degree negative co-relation.

(v)



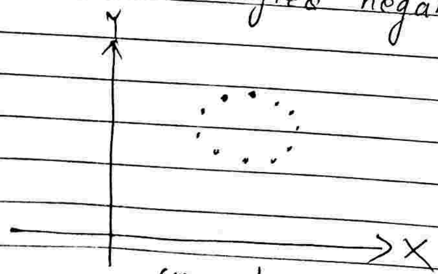
Low degree positive corelation.

(vi)



Low degree negative corellation.

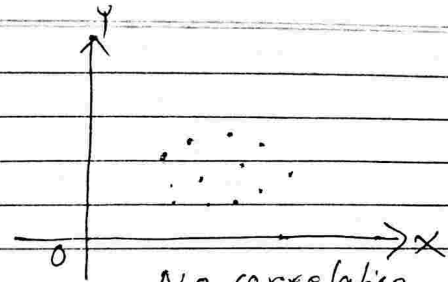
(vii)



Curve linear co-realation

Date
Page No.

(vii)



No correlation

(3) 45, 55, 5, 150, 10, 15, 25, 35, 20, 100

X_1	X_1	$d_1 = X_1 - 25$	d_1^2
45	5	−20	400
55	10	−15	225
5	15	−10	100
150	20	−5	25
10	25	0	0
15	35	10	100
25	45	20	400
35	55	30	900
20	100	75	5625
100	150	125	15625.
		$d_1 = 210$	$\Sigma d_1^2 = 23400$

For X

$$\text{Mean } \overline{X}_1 = A + \frac{\Sigma d_1}{n} = 25 + \frac{210}{10}$$

$$= 25 + 21$$

$$= 46$$

$$\text{Median } M_d = \text{Value of } \left(\frac{n+1}{2}\right)^{th} \text{ item}$$

$$= \left(\frac{10+1}{2}\right)^{th} \text{ item}$$

$$= 5.5 \text{ item} = \frac{25+35}{2} = 30$$

Date
Page No.

$$S.D\ \sigma = \sqrt{\frac{1}{n}\Sigma d_1^2 - \left(\frac{\Sigma d_1}{n}\right)^2}$$

$$= \sqrt{\frac{1}{10}\times 23400 - \left(\frac{210}{10}\right)^2}$$

$$= \sqrt{2340 - 441}$$

$$= \sqrt{1899}$$

$$= 43.5\ \#$$

$$S_{KP} = \frac{3(\bar{X} - M_o)}{\sigma}$$

$$= \frac{3(46 - 30)}{43.5}$$

$$= 1.10\ \#$$

4

X	F	FX	FX²	C.F
0	100	0	0	100
10	94	940	9400	194
20	84	1680	33600	278
30	59	1770	53100	337
40	50	2000	80,000	387
50	38	1900	95000	425
60	13	780	46800	438
70	5	350	24500	443

772.91 −

$F = 443$ $\Sigma FX = \cancel{7908}$ 9420 $\Sigma FX^2 = \cancel{294600}$ 3424.00

$$\text{Mean } \bar{X} = \frac{\Sigma FX}{N} = \frac{\cancel{7908}\ 9420}{443} = \cancel{988.5}\ 21.26\ \ 1177.5$$

42800

$$SD,\ \sigma = \sqrt{\frac{1}{N}\Sigma FX^2 - (\bar{X})^2}$$

$$= \sqrt{\frac{1}{443}\,294600 - (\cancel{988.5}\ 21.26)^2}$$

Date
Page No.

$$= \sqrt{\cancel{36825}\ 42800 - \cancel{1386506}\ 451.98}$$

$$\cancel{= \sqrt{665.01 - 451.98}}$$

$$= \sqrt{772.91 - 451.98}$$

$$= 17.91$$

Since the distribution is

3

$$\text{Median} = \frac{N+1}{2} = \frac{443+1}{2} = 222$$

c.f just greater than 222 is 278 so 20.

5 Since Mean is greater than Median it is right skewed.

6 Since symmetrical so Mean = Median = Mode
20.33 #

7

X	F	C.F	X	$d' = \frac{X-350}{100}$	Fd'	Fd'²
~~Below~~ 0−100	10	10	50	−3	−30	90
100−200	5	15	150	−2	−10	20
200−300	15	30	250	−1	−15	15
300−400	12	42	350	0	0	0
400−500	15	57	450	1	15	15
500−600	7	64	550	2	14	28
600−700	4	68	650	3	12	36

N = 68 $\Sigma Fd' = 44$ $\Sigma Fd'^2 = \cancel{324}$ 204

Since (15, 15) ill defined,

$$Sk(p) = \frac{3(\bar{X} - M_d)}{\sigma}$$

$$\text{Mean } \bar{X} = A + \frac{\Sigma Fd'}{N}\times h = 350 + \frac{44}{68}\times 100$$

$$\cancel{= 350.6} = 414.70$$

(0.26)

Date
Page No.

$S.D,\ \sigma = \sqrt{\frac{1}{N} \Sigma fd^2 - \left(\frac{\Sigma fd'}{N}\right)^2} \times h$

$= \sqrt{\frac{1}{68} \times 324 - \left(\frac{44}{68}\right)^2} \times 100$

$= 4.76 - 41.41$

$= 6.053$

Median $= \frac{N}{2} = \frac{68}{2}\ 34 = 34$

The cumulative frequency just greater than 34 is 42 300 - 400

$Md = l + \frac{h}{f}\left(\frac{N}{2} - c.f\right)$

$= 300 + \frac{100}{68}(34 - 42)$

Since bimodal

Class $\left(\frac{N}{2}\right)^{th}$ value

$= 34^{th}$ value

c.f greater than 34 is 42, median lies in between (300-400)

Median $= L + \frac{\frac{N}{2} - c.f}{f} \times h$

$= 300 + \frac{34 - 30}{12} \times 100$

$= 333.33$

Mean $(\bar{x}) = A + \frac{\Sigma fd'}{N} \times h$

$= 350 + \frac{-14}{68} \times 100$

$= 329.41$

$S.D = \sigma = \sqrt{\frac{\Sigma fd'^2}{N} - \left(\frac{\Sigma fd'}{N}\right)^2} \times h$

$= \sqrt{\frac{204}{68} - \left(\frac{-14}{68}\right)^2} \times 100$

$= 20.59$

Now, coefficient of skewness

$Ksp = \frac{3(\bar{X} - Me)}{\sigma}$

$= \frac{3(329.41 - 333.33)}{20.59} = -0.57$ #

Date
Page No.

8. $N = 10$

$\Sigma X = 452$

$\Sigma X^2 = 24,270$

$M_0 = 43.7$

We have, $CV = \frac{\sigma}{\bar{X}} \times 100$

$\bar{X} = \frac{\Sigma X}{N} = \frac{452}{10} = 45.2$

$S.D = \sigma = \sqrt{\frac{1}{n} \Sigma X^2 - \left(\frac{\Sigma X}{n}\right)^2}$

$= \sqrt{\frac{1}{10}\ 24,270 - \left(\frac{452}{10}\right)^2}$

$= \sqrt{2427 - 2043.04}$

$= \sqrt{383.96}$

$= 19.59$

Now, $Sk(P) = \frac{452 - 43.7}{19.59}$

$= 20.84$ #

9.

Value in Rs	F	c.f
less than 50	40	40
50-100	80	120
100-150	130	250
150-200	60	310
200 & above	30	340
Now	N = 340	

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

$= \left(\frac{3 \times 340}{4}\right)^{th}$ term

$= (255)^{th}$ term

Date
Page No.

Since c.f just greater than 255 is 310, Q_3 lies in (150 - 200)

$$Q_3 = L + \frac{\frac{3N}{4} - c.f}{F} \times h$$

$$= 150 + \frac{255 - 250}{60} \times 50$$

$$= 154.16.$$

(Q1) = $\left(\frac{N}{4}\right)$th term i.e. $\left(\frac{340}{4}\right)$th term = 85

Since C.F just greater than 85 is 120 so class is 50-100.

$$Q_1 = L + \frac{h}{F}\left(\frac{N}{4} - c.f\right)$$

$$= 50 + \frac{\overset{85}{\cancel{340}} - \overset{40}{\cancel{200}}}{80} \times 50$$

$$\cancel{= 237.5.} \quad 78.125$$

For median: let $\frac{N}{2} = \frac{340}{2} = 170$

Since C.F just greater than 170 is 250 so class lies between 100 - 150

Then $Md = L + \frac{h}{F}\left(\frac{N}{2} - c.f\right)$

$$= 100 + \frac{50}{130}(170 - 120)$$

$$= 100 + 19.2$$

$$= 119.23.$$

Now,

$$S_K(B) = \frac{Q_3 + Q_1 - 2Md}{Q_3 - Q_1}$$

$$= \frac{154.16 + \cancel{237.5}\ 78.125 - 2\times119.23}{154.16 - \cancel{237.5}\ 78.125}$$

$$= \frac{\cancel{391.66}\ 232.125 - 238.46}{-83.34} = \frac{153.2}{-83.34}$$

-6.355

$\cancel{80}.76.035$

$= -0.08$

Date
Page No.

(10) Dis^n A: Mean 100; Median 90, SD = 10
Dis^n B: Mean 90, Median 80, SD = 10

Here, we know,

For A

$$Mode = 3Me - 4\bar{X}$$
$$= 3\times90 - 4\times100$$
$$= 270 - 400$$
$$= -130$$

For B

$$Mo = 3Me - 4\bar{X}$$
$$= 3\times80 - 4\times90$$
$$= 240 - 360$$
$$= 120.$$

$$Sk(A) = \frac{3(\bar{X} - Md)}{\sigma} = \frac{3(100 - 90)}{10} = 3$$

$$Sk(B) = \frac{3(\bar{X} - Md)}{\sigma} = \frac{3(90 - 80)}{10} = 3$$

We can conclude that measurement of skewness is same.

$$C.V_1 = \frac{\sigma}{\bar{X}} \times 100 = \frac{10}{\cancel{100}} \times \cancel{100} = 10$$

$$C.V_2 = \frac{\sigma}{\bar{X}} \times 100 = \frac{\cancel{10}}{\cancel{90}} \times 100 = 11.1$$

Also, degree of variation isn't same.

(11) $SK(B) = 0.5$ Also $Q_3 + Q_1 = 28$

$Me = 11$

$Q_1 = 28 - Q_3$ — (i)

or $SkB = \frac{Q_3 + Q_1 - 2Md}{Q_3 - Q_1}$

or $0.5 = \frac{28 - 2\times11}{Q_3 - Q_1}$

or $(Q_3 - Q_1) = \frac{6}{0.5}$

or, $0.5Q_3 - 0.5Q_1 = 6$

$Q_3 - Q_1 = 12$

$-Q_1 = 12 - Q_3$ — (ii)

Date
Page No.

$Q_1 = 28 - Q_3$ — (i)
$-Q_1 = 12 - Q_3$ — (ii)
$40 - 2Q_3 = 0$
$2Q_3 = 40$
$Q_3 = 20$

For $Q_1 = 28 - Q_3$
$= 28 - 20$
$= 8$ #

(12) $Sk(p) = 0.5$
$C.V = 5\%$
$\sigma = 2.$

$Skp = \frac{Mean - Mode}{SD}$

$0.5 = \frac{Mean - Mode}{\sigma}$

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

$5 = \frac{2}{\bar{X}} \times 100$

$\therefore 5\bar{X} = 200$

$\therefore \bar{X} = \frac{200}{5} = 40$

So, $0.5 = \frac{40 - Mo}{2}$

or, $1 = 40 - Mo$ $\therefore$ or, $39 = Mo$ #

(13). $\bar{X} = 65$ $Mo = 70$ $Skp = -0.6$

$Skp = \frac{3(\bar{X} - Me)}{\sigma}$

$-0.6 = \frac{3(65-70)}{\sigma}$

$0.6\sigma = 15$

$\sigma = 25$

~~Mode = 3Me − 4X̄ = 3×70 − 4×65 = 210 − 260~~ =

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

$= \frac{25}{65} \times 100 = 38.46\%$

$Skp = \frac{\bar{X} - Mo}{\sigma}$

$-0.6 = \frac{65 - Mo}{\sigma}$

$-15 - 65 = -Mo$

Date
Page No.

(4)

x	f	M	fM	M^2	fM^2	C.F
0-10	6	5	30	25	150	6
10-20	10	15	150	225	2250	16
20-30	25	25	625	625	15625	41
30-40	9	35	315	1225	11025	50
40-50	12	45	540	2025	24300	62
50-60	25	55	1375	3025	75625	87
60-70	8	65	520	4225	33800	95
70-80	5	75	375	5625	28125	100
	100		3930		190900	

$\bar{X} = \frac{\Sigma fM}{N} = \frac{3930}{100} = 39.30$

$\sigma = \sqrt{\frac{\Sigma fM^2}{N} - \left(\frac{\Sigma fM}{N}\right)^2}$

$= \sqrt{\frac{190900}{100} - (39.30)^2}$

$= \sqrt{1909 - 1544.49}$

$= \sqrt{364.51}$

$= 19.09$

$Me = \left(\frac{N}{2}\right)^{th}$ term

$= \left(\frac{100}{2}\right)^{th}$ term

$= (50)^{th}$ term

In c.f column, 50 is just More on 50 in c.f column & it's corresponding class is 30-40.

Median class $= l + \frac{\frac{N}{2} - c.f}{f} \times i$

$= 30 + \frac{50-41}{9} \times 10$

Date
Page No.

$$= 30 + \frac{9}{9} \times 10$$

$$= 30 + 10$$

$$= 40$$

Hence,

$$\text{Pearson's Coeff of skewness} = \frac{3(\bar{X} - Me)}{6}$$

$$= \frac{3(39.30 - 40)}{19.09}$$

$$= \frac{-2.1}{19.09}$$

$$= -0.11$$

The maximum frequency is repeated twice which is ill-defined.

(5)

$$6 = \sqrt{\frac{6_1 d_1^2 + 6_2 d_2^2}{n_1 + n_2}} \qquad d = (x_1 - x)^2$$

Date
Page No.

Points obtained on graph are:-

(a) lies on a single line, shows perfect co-relat
(b) Are very close to each other in a certain direct indicates high degree co-relation
are apart from each other indicates low degree co-relation.
(d) lies on a certain curve shows curve linear co-relation are evenly distributed or distributed without any pattern.
(c) Moved from lower left corner to upper right indicates positive co-relation.
(d) Moved from upper left corner to lower right indicates negative co-relation.

Notes

1. It measures both linear as well as curve linear co-relation between two variables
2. It can't find amount or extent or strength of corelationship between the two variables.
3. It finds nature and directions ~~and~~ of relationship between two variables.

(B) Karl pearson's correlation coefficient or product moment correlation coefficient

Karl pearson's correlation coefficient is defined as ratio of covariance to the product of SDs of a pair of variables say X and Y. → It is denoted by r or r_{XY} or $r(X, y)$

$$r = \frac{\text{Covariance } (X, Y)}{\text{Product of SD's of } X \,\&\, Y}$$

Date
Page No.

i.e $r = \frac{Cov(X,Y)}{\sigma_x, \sigma_y}$

Note

- It measures degree of dissociation between two variables in linear only.
- It measures amount/extent or strength of co-relation between two variables and the value helps to find nature and direction of the relationship between the two variables.
- It doesn't measure the curve linear relationship between the variables.

Properties

1- It's value lie from -1 to $+1$

2- It is independent on the units of the variables i.e it is a pure number. In other word it is unit free measure.

3- Two independent variables are uncorelated but it's converse may not be always true.

4- It is independent on change of origin as well as change of scale but depending on the sign of scale factors

if $U = \frac{X-A}{h}$ $V = \frac{Y-B}{K}$

then,

$$r(U,V) = \frac{|h|\,|k|}{h \cdot K} \, r(X,Y)$$

-0.76
-0.77
-0.78
-0.2
0.
-0.75
-0.25

Date
Page No.

Notes

1. $r = -1$ —— -0.75 —— -0.25 —— $r=0$ —— 0.25 —— 0.75 —— $r=1$

a) If $r = -1 \Rightarrow$ perfect negative co-relation.
b) If $r = 1 \Rightarrow$ perfect positive co-relation.
c) If $r = 0 \Rightarrow$ no linear co-relation.
d) r is close to 1 (i.e $0.75 \le r \le 1$)
$\Rightarrow$ High degree posithe co-rellation.
e) $0.25 \le r < 0.75 \Rightarrow$ Moderately posithe co-relation
e) r is close to '0' ($0 < r < 0.25$)
$\Rightarrow$ low degree positive co-relation.
F) r is close to -1 ($-1 < r \le -0.75$)
$\Rightarrow$ High degree negative co-relation.
g) $-0.75 < r \le -0.25 \Rightarrow$ Moderate negative co-reation.
h) r is close to '0' (i.e $-0.25 < r < 0$)
$\Rightarrow$ low degree negative correation.

If x and y are linearly related by $Y = a + bX$ then
a) $r = 1$; if $b > 0$
b) $r = -1$; if $b < 0$
where b = slope of the line.

Example :- Find co-relation coefficient of x and y if the two variables x and y are related as $2x + 5y = 11$

$r = -1$ (Slope of the line $-\frac{2}{5}$ which is negative)

3. If $U = a + bx$, $V = c + dy$ then
$r(U,V) = r(a+bX, c+dy) = r(x,y)$
if b & d have the same sign.

Date
Page No.

$= -r(X,Y)$ if b and d have opposite sign.

Example:-
Find a co-relation coefficient between U and V if correlation coefficient of X and Y is $\boxed{0.98}$ for $2U + 5x = 11$ and $5V - 10Y = 21$.

$2U = 11 - 5X$ and $5V = 10Y + 21$

$U = 5.5 - 2.5X$ $V = 2Y + 4.2$

$r(U,V) = r(5.5 - 2.5X, 2Y + 4.2)$

since − and + (opposite sign)

$= -r(x,y)$

$= -0.98$ #

Find the correlation coefficient of U & Y if the co-relation coefficient of X and Y is −0.6 for $5x + 2U = 24$

$r(U,y) = r(12 - 2.5X, 1Y)$

$= -r(X,Y)$

$= -(-0.6)$

$= 0.6$ Ans.

4. If two variables x and y are independent then $r = 0$ (i.e $cov(x,y) = 0$)

Date
Page No.

Covariance of two variables x & y

Variance (X)	Covariance (X,Y)
− It measures variation in x.	− It measures joint variation in x and y.
− $V(X) = \frac{\Sigma(X-\bar{X})^2}{n}$	− $Cov(X,Y) = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{n}$
− $V(X) = \frac{\Sigma X^2}{n} - \bar{X}^2$	− $Cov(X,Y) = \Sigma XY$
− $V(X) = \frac{1}{n^2}\{n\Sigma X^2 - (\Sigma X)^2\}$	− $Cov(X,Y) = \frac{1}{n^2}\{n\Sigma XY - \Sigma X \Sigma Y$

Notes:-

× ÷

Affected by scale but no origin + −

(1) $Cov(a + bX, c + dy) = b \cdot d \; Cov(X,Y)$

not affected

(2) Covariance may be either positive, negative or zero.

(3) (a) If $Cov(X,Y) > 0$ then $r > 0$
(b) if $Cov(X,Y) < 0$ then $r < 0$
(c) If $Cov(X,Y) = 0$ then $r = 0$

(4) absolute value of co-variance is ~~less or~~ equal to the product of standard deviation
i.e $|Cov(X,Y)| \leq \sigma_x \sigma_y$

(5) For a pair of variables X, Y
$\{Cov(X,Y)\}^2 \leq V(X) \cdot V(Y)$

Date
Page No.

(6) Sum of product of a pair of observations is given by $\Sigma XY = n[Cov(X,Y) - \bar{X}\bar{Y}]$

Calculation of corelation coefficient

(i) $r = \dfrac{Cov(X,Y)}{\sigma_x \cdot \sigma_y}$ [From defination]

(ii) $r = \dfrac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2}\sqrt{\Sigma(Y-\bar{Y})^2}}$

i.e $r = \dfrac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$

Actual mean deviation method

where $x = X - \bar{X}$, $y = Y - \bar{Y}$ } i.e x & y are deviations, taken from the respetive means.

(iii) $r = \dfrac{n\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$

(iv) $r = \dfrac{n\Sigma UV - \Sigma U \cdot \Sigma V}{\sqrt{n\Sigma U^2 - (\Sigma V^2)^2}\sqrt{n\Sigma V^2 - (\Sigma V)^2}}$

(Shortcut method)

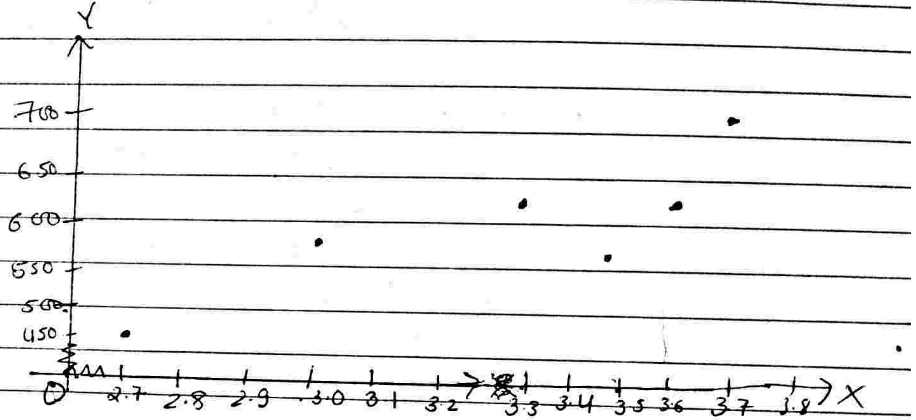
$U = \dfrac{X-A}{h}$ and $V = \dfrac{Y-B}{K}$.

Date
Page No.

Numerical Problems

(1) Let X: GPA Math test score
Y: SAT " " "

(a)



(b) The dots obtained on the scattered diagram (graph) move from lower left corner to upper right which indicates that there is a positive corellation between GPA and SAT. Mathematics test scores.

(c)

X	Y	$U=10(X-3.3)$	$V=\frac{Y-570}{10}$	UV	U^2	V^2
2.7	450	−6	−12	72	36	144
3.5	560	2	−1	−2	4	1
3.7	700	4	13	52	16	169
3.3	620	0	5	0	0	25
3.6	640	3	7	21	9	49
3.0	570	−3	0	0	9	0
		0	12	143	74	388

Here,
$n=6$, $\Sigma U = 0$, $\Sigma V = 12$
$\Sigma UV = 143$, $\Sigma U^2 = 74$

Date
Page No.

$$\therefore r = \frac{n\,\Sigma UV - \Sigma U . \Sigma V}{\sqrt{n\,\Sigma U^2 - (\Sigma U)^2}\sqrt{n\,\Sigma V^2 - (\Sigma V)^2}}$$

$$\text{or, } r = \frac{(6\times 143) - (0\times 12)}{\sqrt{(6\times 74) - (0)^2}\sqrt{6\times 388 - (12)^2}}$$

$$\text{or, } r = \frac{858}{\sqrt{444}\,\sqrt{2328 - 144}}$$

$$= \frac{858}{21.07 \times 46.73}$$

$$r = 0.8714 \#$$

Here,
$r = 0.87$ indicates that there is a high degree positive correlation between GPA and SAT Math test scores.

Test of consistency of r :-
or
Test of significance of r.

$$(P.E)(r) = \frac{0.6745\,(1 - r^2)}{\sqrt{n}}$$

Case I: If $|r| < PE(r)$, then r is not significant.
Case II: If $|r| \geq 6\,PE(r)$, then r is significant.
otherwise, conclusion cannot be made using this test.

Estimation of population co-relation coefficient:

Limits of population co-relation coefficient are

$$[r - PE(r),\ r + PE(r)]$$

Example:-
- If $r = 0.18$, $n = 10$; test the significance of r.
- Find limits of population co-relation coefficient.

$$P.E(r) = \frac{0.6745(1 - r^2)}{\sqrt{n}} = \frac{0.6745(1 - 0.18^2)}{\sqrt{10}}$$

$$= \frac{0.65313}{3.16}$$

$$= 0.206$$

Here probable error $> |r|$
$\therefore$ r is not significant

Limit of popn correlation coefficient.
$[r - PE(r),\ r + PE(r)]$
i.e $[0.18 - 0.206,\ 0.18 + 0.206]$
i.e $[-0.026,\ 0.386]$.

Coefficient of determination

It measures the porportion (percentage) of variations on a dependent variable which can be explained by the variation of values of independent variable. It is denoted by r^2, i.e

Coefficient of determination $= r^2$

$$\text{i.e } r^2 = \frac{\text{Explained / Accounted variation}}{\text{Total variation}}$$

and the coefficient of non-determination $= 1 - r^2$

$$\text{i.e } 1 - r^2 = \frac{\text{Unexplained / Unaccounted variation}}{\text{Total variation}}$$

Example:-

If $r = 0.9$ then $r^2 = 0.81$ % = 81%

$r^2 = 81\% \rightarrow$ It means that 81% of total variation on a dependent variable is explained by an independent variable and the remaining 90% of the variation is unexplained.

Q no 6

$n = 15$, $\bar{X} = 25$, $\bar{Y} = 18$

$\sigma_x = 3.01$, $\sigma_Y = 3.03$

$\Sigma(X-\bar{X})^2 = 136$; $\Sigma(X-\bar{Y})^2 = 138$

$\Sigma(X-\bar{X})(Y-\bar{Y}) = 122$

Method :- 1

$$Cov(X,Y) = \frac{1}{n}\Sigma(X-\bar{X})(Y-\bar{Y})$$

$$= \frac{1}{15} \times 122$$

$$= 8.13$$

$$\therefore r = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_Y}$$

$$= \frac{8.13}{3.01 \times 3.03}$$

$\therefore r = 0.89$ #

Method :- 2

$$r = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2}\sqrt{\Sigma(Y-\bar{Y})^2}}$$

$$= \frac{122}{\sqrt{136}\sqrt{138}}$$

$$= \frac{122}{136.99}$$

$= 0.89$ #

Q no 7

$\Sigma U = -14$, $\Sigma V = 18$

$\Sigma U^2 = 4034$ $\Sigma V^2 = 6308$

$\Sigma UV = 1510$; $n = 12$

$$r = \frac{n\Sigma UV - \Sigma U \Sigma V}{\sqrt{n\Sigma U^2 - (\Sigma U)^2}\sqrt{n\Sigma V^2 - (\Sigma V)^2}}$$

$$= \frac{(12 \times 1510) - (-14) \times (18)}{\sqrt{(12 \times 4304) - (-14)^2}\sqrt{(12 \times 6308) - (18)^2}}$$

$$= \frac{18120 + 252}{\sqrt{51648 - 196}\sqrt{75696 - 324}}$$

$$= \frac{18372}{\sqrt{51452}\sqrt{75372}}$$

$$= \frac{18372}{226.83 \times 274.53}$$

$$= \frac{18372}{62273.82}$$

$= 0.2950$ #

Kurtosis Page 225

(17) Daily wages (in Rs) 50-60 60-70 70-80 80-90 90-100 100-110 110-120 Total 100

Daily wages	F	C.F
50-60	10	10
60-70	14	24
70-80	18	42
80-90	24	66
90-100	16	82
100-110	12	94
110-120	6	100
	N=100	

For Q_1, let $\frac{N}{4} = \frac{100}{4} = 25$ i.e. Q lies 42 i.e. 70 - 80

Then $Q_1 = L + \frac{\frac{N}{4} - c.f}{f} \times h$

$$= 70 + \frac{25 - 24}{18} \times 10$$

$$= 70 + \frac{1}{18} \times 10 = 70.5$$

For Q_3, let $\frac{3N}{4} = \frac{3 \times 100}{4} = 75$ just greater 82 class 90-100

Then $Q_3 = L + \frac{\frac{N}{4} - c.f}{f} \times h$

$$= 90 + \frac{75 - 66}{16} \times 10$$

$$= 90 + \frac{9}{16} \times 10$$

$$= 95.625$$

For P_{10}, let $\frac{10N}{100} = \frac{10 \times 100}{100} = 10$ i.e. P_{10} lies in CI 50-60

Here $P_{10} = L + \frac{h}{f}\left(\frac{10N}{100} - cf\right)$

$$= 50 + \frac{10}{10}(10 - 0)$$

$$= 50 + 1$$

$$= 51$$

For, P_{90} let $\frac{90N}{100} = \frac{90\times 100}{100} = 90$ i.e P_{90} lies in the CI 100-110

$$P_{90} = L + \frac{h}{f}\left(\frac{90N}{100} - c.f\right)$$
$$= 100 + \frac{10}{12}(90 - 82)$$
$$= 100 + \frac{10}{12}\times 8$$
$$= 100 + 6.66$$
$$= 106.66$$

Now,

$$k = \frac{\frac{1}{2}(Q_3 - Q_1)}{P_{90} - P_{10}}$$
$$= \frac{\frac{1}{2}(95.625 - 70.5)}{106.66 - 51}$$
$$= \frac{\frac{1}{2}\times 25.125}{55.66}$$
$$= \frac{12.56}{55.66}$$
$$= 0.2252.$$

This distribution is close to normal since $k = 0.263$ is normal / mesokurtic.

(18)

Marks	X	$d' = \frac{x-85}{10}$	f	c.f	fd'	fd'^{2}
40-50	45	-4	10	10	-40	160
50-60	55	-3	15	25	-45	135
60-70	65	-2	20	45	-40	80
70-80	75	-1	28	73	-28	28
80-90	85	0	35	108	0	0
90-100	95	1	25	133	25	25
100-110	105	2	18	151	36	72
110-120	115	3	10	161	30	90
120-130	125	4	8	169	32	128
			N=169		$\Sigma fd' = -30$	$\Sigma fd'^2 = \cancel{637}$ 718

Mean $\bar{X} = A + \frac{\Sigma fd}{N}\times h$

$$= 85 + \frac{-30}{169}\times 10$$
$$= 83.225$$

Since max F (35) (80-90) ∴ Mode $= L + \frac{f_m - f_1}{2f_m - f_1 - f_2}\times h$

$$= 80 + \frac{35-28}{2\times 35 - 28 - 25}\times 10$$
$$= 80 + \frac{7}{17}\times 10$$
$$= 84.11$$

$\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\times h$

S.D, $\sigma = \sqrt{\frac{1}{N}\Sigma fd'^2 - \left(\frac{\Sigma fd'}{N}\right)^2}\times h$

$$= \sqrt{\frac{1}{169}\times 718 - \left(\frac{-30}{169}\right)^2}\times 10$$
$$= \sqrt{4.248 - 0.0315}\times 10$$
$$= 2.05\times 10$$
$$= 20.53$$

Pearson's coefficient of skewness, $Sk(p) = \frac{\text{Mean} - \text{Mode}}{SD}$

$$= \frac{83.225 - 84.11}{20}$$
$$= -0.005$$

= Approximately symmetrical

Date
Page No.

(19)

	f	$c.f$
0 - 5	3	3
5 - 10	7	10
10 - 15	18	28
15 - 20	35	63
20 - 25	25	88
25 - 30	8	96
30 - 35	4	100

$N = 100$ (Same).

Q_1 Let $\frac{N}{4} = \frac{100}{4} = 25$ i.e 28 10-15

$$Q_1 = L + \frac{\frac{N}{4} - c.F}{f} \times h$$

$$= 10 + \frac{25 - 10}{7} \times 5$$

$$= 20.71$$

Q_3. Let $\frac{3N}{4} = \frac{3 \times 100}{4} = 75$ .i.e 88 so 20-25

$$Q_3 = L + \frac{\frac{N}{4} - c.F}{f} \times h$$

$$= 20 + \frac{75 - 63}{25} \times 10$$

$$= 20 + 10$$

$$= 30$$

For P_{10} Let $\frac{10N}{100} = \frac{10 \times 100}{100} = 10 \rightarrow 5\text{-}10$.

$$P_{10} = L + \frac{h}{f}\left(\frac{10N}{100} - c.f\right)$$

$$= 5 + \frac{10}{3}(10 - 0)$$

$$= 5 + 33.3$$

$$= 38.3$$

Date
Page No.

P_{90} is $\frac{90N}{100} = \frac{90 \times 100}{100} = 90$ is 96 25 - 30.

$$P_{90} = L + \frac{h}{f}\left(\frac{90N}{100} - c.f\right)$$

$$= 25 + \frac{10}{8}(90 - 88)$$

$$= 25 + 2.5$$

$$= 27.5$$

Now,

$$K = \frac{\frac{1}{2}(Q_3 - Q_1)}{P_{90} - P_{10}}$$

$$= \frac{\frac{1}{2}(30 - 20.71)}{27.5 - 38.3}$$

Date
Page No.

Date
Page No.

Geometric Progression (GP)

F_1

F_2

F_3

F_4

(F5) Geometric mean (G.M)

G.M betn a & b $= \sqrt{a \cdot b}$

Know it :- If $y\sqrt{x3}$ (or $y^2 = x3$)

(Read as y is GM between x & 3)

Then we write $\boxed{x, y, 3 \text{ are in GP}}$

(F6)

To insert 4 GMs between a & b

Here 1st term $(t_1) = a$

No of terms $(n) = 4 + 2 = 6$

last term i.e 6th term $= (t_6) = b$

Now Common ratio $(r) = \left(\frac{t_6}{t_1}\right)^{\frac{1}{6-1}} = \left(\frac{b}{a}\right)^{\frac{1}{5}}$

Then 1st G.M = 2nd term $= t_2 = a \cdot r$

2nd G.M = 3rd term $= t_3 = ar^2$

3rd GM = 4th term $= t_4 = a \cdot r^3$.

4th GM = 5th term $= t_5 = a \cdot r^4$.

Practise

* Insert 4 G.Ms between 5 and 1215.

We know, $a = 5$

$t_6 = 1215$.

$n = 4$.

$r = \left(\frac{t_6}{t_1}\right)^{\frac{1}{6-1}}$

$= \left(\frac{1215}{5}\right)^{\frac{1}{5}}$

$= (1243)^{\frac{1}{5}}$

$= (3)^{5 \times \frac{1}{5}}$

$= 3$

Date
Page No.

Then,

1^{st} G.M $= ar = 5 \times 3 = 15$

2^{nd} G.M $= ar^2 = 5 \times 9 = 45$

3^{rd} G.M $= ar^3 = 5 \times 27 = 135$

4^{th} G.M $= ar^4 = 5 \times 81 = 405$.

F7 General and Special Notation of terms in G.P.

No of terms	General notation	Special Notation
3	a, ar, ar^2	$\frac{a}{r}, a, ar$
5	a, ar, ar^2, ar^3, ar^4	$\frac{a}{r}, a, ar,$
4	a, ar, ar^2, ar^3	$\frac{a}{r^2}, \frac{a}{r}, a, ar, ar^2$

Note:- If product of terms ~~are~~ is given we solve the problem by using special notation.
If product is not given then we use general notation.

F8 Properties of G.P

If a, b, c are in G.P then,

(i) ak, bk, ck are also in G.P.

(ii) $\frac{a}{k}, \frac{b}{k}, \frac{c}{k}$ (where $k \neq 0$) are also in G.P

(iii) a^k, b^k, c^k are also in G.P

(iv) $\frac{1}{a}, \frac{1}{b}, \frac{1}{c}$ are also in G.P.

Date
Page No.

Ex:- 2(B)

1) In a G.P,

$(p+q)^{th}$ term i.e $t_{p+q} = m$ ——— (i)

$(p-q)^{th}$ term i.e $t_{p-q} = n$ ——— (ii)

To prove: p^{th} term i.e $t_p = \sqrt{m.n}$

Multiplying eq^n (i) & (ii)

$t_{p+q} \times t_{p-q} = m.n$

or, $a.r^{p+q-1} \times a.r^{p-q-1} = mn$

or, $a^2.r^{p+\cancel{q}-1+p-\cancel{q}-1} = mn$

or, $a^2 r^{2p-2} = mn$

or, $a^2. r^{2(p-1)} = mn$

or, $(a.r^{(p-1)})^2 = n.m$

or, $(t_p)^2 = n.m$ ~~proved~~

Taking sq. root,

$\therefore t_p = \sqrt{m.n}$ proved

2)

v) Given,

Not G.P ← $\frac{4}{7} - \frac{5}{7^2} + \frac{4}{7^3} - \frac{5}{7^4} + \frac{4}{7^5}$ ---- to ∞

Above series can be written as,

$= \left(+\frac{4}{7} + \frac{4}{7^3} + \frac{4}{7^5}\right)$ — to ∞ $- \left(\frac{5}{7^2} + \frac{5}{7^4} + \frac{5}{7^6} \text{---- to } \infty\right)$

$= \left(\frac{a}{1-r}\right) - \left(\frac{a}{1-r}\right)$

$= \left(\frac{\frac{4}{7}}{1-\frac{1}{7^2}}\right) - \left(\frac{\frac{5}{7^2}}{1-\frac{1}{7^2}}\right)$

$= \left(\frac{\frac{4}{\cancel{7}}}{\frac{49-1}{\cancel{7^2}}}\right) - \left(\frac{\frac{5}{\cancel{7^2}}}{\frac{49-1}{\cancel{7^2}}}\right)$

Date
Page No.

$\frac{\frac{4}{7}\times 49}{48} - \frac{\frac{5}{49}\times 49}{48}$

$= \frac{28}{48} - \frac{5}{48}$

$= \frac{23}{48}$ #

(4) $n = ?$

Sum of the given series $\sqrt{3}+3+3\sqrt{3}+9+\ldots$
is $39+13\sqrt{3}$ → i.e sum

Here,

$t_1 = \sqrt{3}$

$r = \frac{t_2}{t_1} = \frac{3}{\sqrt{3}} = \sqrt{3}$

$S_n = 39+13\sqrt{3}$

We know,

$S_n = \frac{a(r^n-1)}{r-1}$

or, $39+13\sqrt{3} = \frac{\sqrt{3}(\sqrt{3}^n-1)}{\sqrt{3}-1}$

or, $(39+13\sqrt{3})(\sqrt{3}-1) = \sqrt{3}(\sqrt{3}^n-1)$

or, $39\sqrt{3} - \cancel{39} + \cancel{39} - 13\sqrt{3} = \sqrt{3}(\sqrt{3}^n-1)$

or, $39\sqrt{3} - 13\sqrt{3} = \sqrt{3}(\sqrt{3}^n-1)$

or, $26\cancel{\sqrt{3}} = \cancel{\sqrt{3}}(\sqrt{3}^n-1)$

or, $26 = \sqrt{3}^n - 1$

or, $27 = \sqrt{3}^n$

or, $3^3 = (3)^{n/2}$

or, $3 = \frac{n}{2}$

or, $n = 6$ #

Date
Page No.

(6) In an infinite G.P,
Sum to infinity $= 2$
Sum of their squares $= 4/3$
i.e Sum of series $(a+ar+ar^2+\ldots\ldots$ to $\infty) = 2$
And
Sum of the series $[a^2+[ar]^2+[ar^2]^2+\ldots$ to $\infty] = \frac{4}{3}$

i.e, $\frac{a^2}{1-r^2} = \frac{4}{3}$ i.e $a^2 = \frac{4(1-r^2)}{3}$ —— (ii)

And, Sum of the series $[a+ar+ar^2+\ldots$ to $\infty) = 2$

i.e $\frac{a}{1-r} = 2$ i.e $a = 2(1-r)$ —— (i)

Using the value of eqn (i) in eq (ii), then,

$[2(1-r)]^2 = \frac{4(1-r^2)}{3}$

or, $4(1-r)^2 = \frac{4(1+r)(1-r)}{3}$

or, $1-r = \frac{1+r}{3}$

or, $3-3r = 1+r$

or, $2 = 4r$

$\therefore r = \frac{2}{4} = \frac{1}{2}$ $(|r| = \frac{1}{2} < 1)$

Putting value in eq (i) we get.

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

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

Now, $= 1$.

Series

$a = 1$

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

$ar^2 = 1\times\frac{1}{2^2} = \frac{1}{4}$ --- Hence required series is $1+\frac{1}{2}+\frac{1}{4}+$ —— ∞

Date

Page No.

8(ii) Let three terms in G.P be a, ar, and ar^2

It is given that,

Their sum is $\frac{21}{4}$ i.e $a+ar+ar^2=\frac{21}{4}$

It is given that,

~~Their sum~~

or, $a+ar+ar^2=21$

or, $a(1+r+r^2)=\frac{21}{4}$

or, $a=\frac{21}{4(1+r+r^2)}$ ---- (i)

And sum of their squares $=\frac{273}{16}$

i.e $a^2+(ar)^2+(ar^2)^2=\frac{273}{16}$

or, $a^2(1+r^2+r^4)=\frac{273}{16}$ ---- (ii)

Using the value of a from eq^n (i) to eq^n (ii)

$$\left[\frac{21}{4(1+r+r^2)}\right]^2\cdot(1+r^2+r^4)=\frac{273}{16}$$

$$\frac{21\ \cancel{441}}{\cancel{16}(1+r+r^2)^{\cancel{2}}}\cdot(\cancel{1+r+r^2})(1-r+r^2)=\cancel{\frac{273}{16}}\ 13$$

$$\frac{21}{1+r+r^2}\times 1-r+r^2=13$$

$21-21r+21r^2=13+13r+13r^2$

$8-34r+8r^2=0$

$8r^2-34r+8=0$

$8(r^2-3r+8)=0$

$r^2-(4-1)r+8=0$

$r^2-4r+r+8=0$

$r(r-4)-1(r$

$\frac{21}{13}$ / 8

1 / 3 / 34

Calculations

$1+r^2+r^4=$

$=1+2r^2+r^4-r^2$

$=(1+r^2)^2-(r)^2$

$=(1+r^2+r)(1+r^2-r)$

$(1-2r+r^2)(1+2r+r^2)$

1-

Date

Page No.

Put $r=4$ in eq^n (i)

$$a=\frac{21}{4(1+r+r^2)}=\frac{21}{4(1+4+16)}=\frac{21}{4\times 21}=\frac{1}{4}$$

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

$$a=\frac{21}{4(1+r+r^2)}=\frac{21}{4\left(1+\frac{1}{4}+\frac{1}{16}\right)}$$

$$=\frac{21}{4\left(\frac{16+4+1}{16}\right)}$$

$$=\frac{21}{4\times\frac{21}{16}}=\frac{2\times 16}{4\times 21}$$

$=4.$

Now,

When $r=4$ & $a=\frac{1}{4}$

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

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

$ar^2=\frac{1}{4}\times(4)^2=4$

i.e 3 term in G.P are $\frac{1}{4}$, 1 & 4 #

When $r=\frac{1}{4}$ & $a=4$

$a=4$

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

$ar^2=4\cdot\frac{1}{16}=\frac{1}{4}$

i.e 3 term in G.P are $4, 1, \frac{1}{4}$ #

Date
Page No.

(32)

Aritho-Geometric Series (A.G Series)

Example :-

(1) $2\cdot5 + 5\cdot5^2 + 8\cdot5^3 + 11\cdot5^4 + \text{----}$

A.P $= 2, 5, 8, 11$ ---- $(d=3)$

G.P $= 5, 5^2, 5^3, 5^4,$ --- $(r=5)$

(2) $\frac{4}{7} + \frac{8}{7^2} + \frac{12}{7^3} + \frac{16}{7^4} + \text{----}$ to ∞

A.P $= 4, 8, 12, 16$ --- $(d=4)$

G.P $= \frac{1}{7}, \frac{1}{7^2}, \frac{1}{7^3}$ --- $\left(r=\frac{1}{7}\right)$

Formula :- To find sum to infinity of A.G series)

$$S_\infty = \frac{1}{1-r}\left[t_1 + d\left(\frac{a}{1-r}\right)\right]$$

where r = common ratio of G.P

t_1 = first term of A.G series

d = common difference of A.P

a = 2nd term of G.P.

Date

Page No.

Date

Page No.

Co-rrelation

(18) $n = 15$, $\sigma_x = 3.01$, $\sigma_y = 3.03$

$\sum xy = 122$, where, $x = X - \bar{X}$

$y = Y - \bar{Y}$

$\therefore Cov(X, Y) = \frac{1}{n} \sum (X - \bar{X})(Y - \bar{Y}) = \frac{122}{15}$

$= 8.133.$

$\therefore r = \frac{Cov(X, Y)}{\sigma_x \cdot \sigma_y} = \frac{8.133}{3.01 \times 3.03} = \frac{8.133}{9.1203}$

$= 0.8917$ #

Also, $PE(r) = \frac{0.6745(1 - r^2)}{\sqrt{n}}$

$= \frac{0.6745(1 - (0.891)^2)}{\sqrt{15}}$

$= \frac{0.6745(1 - 0.79388)}{3.87}$

$= \frac{0.6745 \times 0.20612}{3.87}$

$= 0.035$ #

Again,

limits of population correlation coeff are:-

$[r - PE(r), r + PE(r)]$

i.e. $[0.891 - 0.0356, 0.891 + 0.0356]$.

$[0.8554, 0.9266]$.

(19) $r = 0.81$, $n = 10$

$P.E(r) = \frac{0.6745(1 - r^2)}{\sqrt{n}}$

$= \frac{0.6745(1 - 0.6561)}{\sqrt{10}}$

$= \frac{0.6745 \times 0.3439}{3.16} = 0.0734$

Now,

$6 \times PE(r) = 6 \times 0.073 = 0.44$

$r > 6\,PE(r)$

$\therefore r$ is significant.

(c) Spearman's rank correlation Coefficient:-

- It is especially utilized for measuring co-relation between two qualitative characterstics (attributes).
- It can also be applied for measuring co-relation between two variables while measuring corelation between two variables, it is easy but less accurate than that of pearson's measure.
- It can also be applied for measuring a level of agreement or disagreement between two judges while measuring co-relation between two attributes.

Defination

Spearman's rank co-relation coefficient is given by

(1) $r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$ (If rank is not repeted this formula applied).

(2) $r_s = 1 - \frac{6\left[\Sigma d^2 + \frac{m_1(m_1^2-1)}{12} + \frac{m_2(m_2^2-1)}{12} + \cdots\right]}{n(n^2-1)}$

(If ranks are repeated)

where $d = R_1 - R_2$: Difference of ranks.

M_i = No of times, a rank is repeated in a series.

[Note :- $\Sigma d^2 = \frac{(1-r_s)\{n(n^2-1)\}}{6}$]

Property

(1) It's values lies from -1 to +1. i.e $-1 \le r_s \le 1$

(2) Sum of differences of rank is always zero.

(25) ⇒ (b)

X	Y	R_x	R_y	$d = R_x - R_y$	d^2
57	19	9	7.5	1.5	2.25
16	6	3	2.5	0.5	0.25
24	9	5	4	1	1
65	20	10	9	1	1
16	4	3	1	2	4
16	15	3	6	-3	9
9	6	1	2.5	-1.5	2.25
40	24	7	10	-3	9
48	19	8	7.5	0.5	0.25
33	13	6	5	1	1
				0	$\Sigma d^2 = 30$

$P.E = \frac{0.6745(1-r^2)}{\sqrt{n}}$

$m_1 = 3,\ m_2 = 2,\ m_3 = 2,\ \Sigma d^2 = 30,\ n = 10$

$\therefore r_s = 1 - \frac{6\left\{\Sigma d^2 + \frac{m_1(m_1^2-1)}{12} + \frac{m_2(m_2^2-1)}{12} + \frac{m_3(m_3^2-1)}{12}\right\}}{n(n^2-1)}$

$= 1 - \frac{6\left[30 + \frac{3\times 8}{12} + \frac{2\times 3}{12} + \frac{2\times 3}{12}\right]}{10(10^2-1)}$

$= \frac{6(30 + 2 + 0.5 + 0.5)}{10(99) = 990.}$

Date
Page No.

$= 1 - \frac{6 \times 33}{990}$

$= 1 - 0.2$

$= 0.8$ #

② $(\bar{X} = 4)$

X	Y	$x = X - \bar{X}$	x^2	$y = Y - \bar{Y}$ $(\bar{Y} = 4)$	y^2	xy
1	7	-3	9	3	9	-9
2	6	-2	4	2	4	-4
3	5	-1	1	1	1	-1
4	4	0	0	0	0	0
5	3	1	1	-1	1	-1
6	2	2	4	-2	4	-4
7	1	3	9	-3	9	-9
$\Sigma X = 28$	$\Sigma Y = 28$	$\Sigma x = 0$	$\Sigma x^2 = 28$	$\Sigma y = 0$	$\Sigma y^2 = 28$	$\Sigma xy = -28$

We have $\bar{X} = \frac{\Sigma X}{n} = \frac{28}{7} = 4$ and $\bar{Y} = \frac{\Sigma Y}{n} = \frac{28}{7} = 4$

Now, correlation coefficient, $r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\sqrt{\Sigma y^2}}$

$= \frac{-28}{\sqrt{28} \cdot \sqrt{28}}$

$= \frac{-\cancel{28}}{\cancel{28}}$

$= -1$ #

(b)

X	Y	X^2	Y^2	XY
9	6	81	36	~~2916~~ 54
11	2	121	4	~~484~~ 22
x	10	x^2	100	~~$100x^2$~~ $10x$
8	y	64	y^2	~~$64y^2$~~ $8y$
7	8	49	64	56
$\Sigma X = 35 + x$	$\Sigma Y = 26 + y$	$\Sigma X^2 = 315 + x^2$	$\Sigma Y^2 = 204 + y^2$	$\Sigma XY = 132 + 10x + 8y$

Date
Page No.

$r = \frac{n \Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{n \Sigma X^2 - (\Sigma X)^2}\sqrt{n \Sigma Y^2 - (\Sigma Y)^2}}$

$=$

②

X	Y	$X = X -$
9	6	
11	2	
x	10	
8	y	
7	8	
$\Sigma X = 35 + x$	$\Sigma Y = 26 + y$	

$n = 5$

$\bar{X} = 8$

$\frac{9 + 11 + x + 8 + 7}{n} = 8$

$9 + 11 + x + 8 + 7 = 8 \times 5$

$9 + 11 + x + 8 + 7 = 40$

$x = 5$ #

$\frac{6 + 2 + 10 + y + 8}{5} = 6$

$6 + 2 + 10 + y + 8 = 6 \times 5$

$y = 4$ #

Date

Page No.

(3)

X	Y	X	Y	$U=X-A$ $(A=30)$	$V=Y-B$ $B=25$	U^2	V^2
23	18	23	18	-7	-7	49	49
27	26	27	20	-3	-5	9	25
28	22	28	22	-2	-3	4	9
28	27	28	27	-2	2	4	4
30	21	28	21	-2	-3	4	9
30	29	30	29	0	4	0	16
33	27	30	27	0	2	0	4
35	29	33	29	3	4	9	16
38	29	35	28	5	3	25	9
		38	29	8	4	64	16
		$\Sigma X=300$		$\Sigma U=-8$	$\Sigma V=1$	$U^2=168$	$V^2=157$

Now, coefficient correlation, r =

$$= \frac{n\,\Sigma UV - (\Sigma U).(\Sigma V)}{\sqrt{n\,\Sigma U^2 - (\Sigma U)^2}\sqrt{n\,\Sigma V^2 - (\Sigma V)^2}}$$

$$= \frac{10.(-8\times1) - (-8)(1)}{\sqrt{10.168 - (-8)^2}\ \sqrt{10.157 - (1)^2}}$$

$$= \frac{-80 - 8}{\sqrt{1680 - 64}\cdot\sqrt{1570 - 1}}$$

$$= \frac{-88}{40.19 - 39.61}$$

$$= \frac{-88}{0.58}$$

$$= 151.72$$

Date

Page No.

(3b)

X	Y	X^2	Y^2	XY
78	125	6084	15625	9750
89	137	7921	18769	12193
97	156	9409	24336	15132
69	112	4761	12544	7728
59	107	3481	11449	6313
79	136	6241	18496	10744
68	123	4624	15129	8364
61	108	3721	11664	6588
600	1004	46242	128012	76812

Now, correlation coefficient,

$$r = \frac{n\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{n\,\Sigma X^2 - (\Sigma X)^2}\ \sqrt{n\,\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{8\times76812 - 600\times1004}{\sqrt{8\times46242 - (600)^2}\cdot\sqrt{8\times128012 - (1004)^2}}$$

$$= \frac{614496 - 602400}{\sqrt{369936 - 360000}\ \sqrt{1024096 - 1008016}}$$

$$= \frac{\cancel{12069}\ 12096}{99.67\times126.80}$$

$$= \frac{12096}{12638.84}$$

$$= 0.957$$

$\therefore r^2 = 0.915$

$$PE = 0.6745\times\frac{1-r^2}{\sqrt{N}}$$

$$= 0.6745\times\frac{1-0.915}{\sqrt{8}}$$

$$= 0.6745\times\frac{0.085}{2.828}$$

$= 0.6745 \times 0.0300$

$= 0.0202$ #

(4)

X	Y	X^2	Y^2	XY
10	3	100	9	30
20	2	400	4	40
30	1	900	1	30
40	5	1600	25	200
50	4	2500	16	200
150	15	5500	55	500

$$r = \frac{n\,\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\,\Sigma X^2 - (\Sigma X)^2}\,\sqrt{n\,\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{5 \times 500 - 150 \cdot 15}{\sqrt{5 \times 5500 - (150)^2}\,\sqrt{5 \times 55 - (15)^2}}$$

$$= \frac{2500 - 2250}{\sqrt{27500 - 22500} \times \sqrt{275 - 225}}$$

$$= \frac{250}{70.71 \times 7.07}$$

$$= \frac{250}{499.9}$$

$$= 0.507.$$

There will be no change because because it is not affected by origin so if we add 40 to each level the value of r will still be 0.507. #

(9) $r = 0.38$

$Cov(X, Y) = 10.2$

$Var(X) = \cancel{61}\ 1.6$

We have, Correlation coefficient, $r = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$

$$\sigma_Y \cdot 0.38 = \frac{10.2}{\sqrt{\cancel{61}\ \cancel{10.2}\ \cancel{61}\ 1.6}}$$

$$\sigma_Y = \frac{10.2}{0.38 \times \sqrt{\cancel{61}\ \cancel{10.2}\ \cancel{61}\ 1.6}}$$

$$\sigma_Y = \frac{10.2}{\cancel{0.38}\ 1.52}$$

$$= 6.71 \ \#$$

(10) $\sigma_X = 2.45 \quad \sigma_Y = 2.61 \quad n = 5$

$Cov(X, Y) = 6$

Correlation coefficient, $r = \frac{Cov(X, Y)}{\sigma_X, \sigma_Y}$

$$= \frac{6}{2.45 \times 2.61}$$

$$= \frac{6}{6.3945}$$

$$= 0.9383 \ \#$$

(11) $n=6 \quad \Sigma X = 160 \quad \Sigma X^2 = 4480 \quad \Sigma Y = 116 \quad \Sigma Y^2 = 2462$

$\Sigma XY = 3275$

$$r = \frac{n\,\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\,\Sigma X^2 - (\Sigma X)^2}\,\sqrt{n\,\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{6 \times 3275 - 160 \times 116}{\sqrt{6 \times 4480 - (160)^2}\,\sqrt{6 \times 2462 - (116)^2}}$$

$$= \frac{19650 - 18560}{\sqrt{26880 - 25600}\,\sqrt{14772 - 13456}}$$

$$= \frac{1090}{35.78 \times 36.27} = \frac{1090}{1297.98} = 0.839$$

Date
Page No.

r = 0.84 indicates that there is high degree posite correlation.

(13)

Age (X)	No of students	Regular players (Y)	X^2	Y^2	XY
15	250	200	225	40,000	3000
16	200	150	256	22500	2400
17	150	90	289	8100	1530
18	120	48	324	2304	864
19	100	30	361	900	570
20	80	12	400	144	240
$\Sigma X = 105$		$\Sigma Y = 530$	$\Sigma X^2 = 1855$	$\Sigma Y^2 = 73948$	$\Sigma XY = 8604$

$$= \frac{n\,\Sigma XY - \Sigma X\,\Sigma Y}{\sqrt{n\,\Sigma X^2 - (\Sigma X)^2}\,\sqrt{n\,\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{6 \times 8604 - 105 \times 530}{\sqrt{6 \times 1855 - (105)^2}\,\sqrt{6 \times 73948 - (530)^2}}$$

$$= \frac{51624 - 55650}{\sqrt{11130 - 11025} \times \sqrt{443688 - 280900}}$$

$$= \frac{-4026}{\sqrt{105} \times \sqrt{162788}}$$

$$= \frac{-4026}{10.24 \times 403.46}$$

$$= \frac{-4026}{4131.43}$$

$r = -0.9\cancel{8}2$

∴ r = −0.982 indicates that there is high degree negative correlation.

Date
Page No.

(15)

Size group (X)	No of items	No of defective items (Y)	Mid value (X)	X^2	Y^2	XY
15-16	400	300	15.5	240.25	90,000	4650
16-17	540	320	16.5	272.25	102400 ~~74120.06~~	5280
17-18	680	340	17.5	306.25	115600	5950
18-19	720	360	18.5	342.25	129600	6660
19-20	800	360	19.5	380.25	129600	7020
20-21	600	240	20.5	420.25	57600	4920
		$\Sigma Y = 1920$	$\Sigma X = 108$	$\Sigma X^2 = 1861.5$	$\Sigma Y^2 = 624800$	$\Sigma XY = 28180$

$$= \frac{n\,\Sigma XY - \Sigma X\,\Sigma Y}{\sqrt{n\,\Sigma X^2 - (\Sigma X)^2}\,\sqrt{n\,\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{6 \times 28180 - 108 \times 1920}{\sqrt{6 \times 1861.5 - (108)^2} \times \sqrt{6 \times 624800 - (1920)^2}}$$

$$= \frac{169080 - 207360}{\sqrt{\cancel{1169}\ 11766 - 11664} \times \sqrt{3748800 - 3686400}}$$

$$= \frac{-38280}{\sqrt{10.09} \times 249.79}$$

$$= \frac{-38280}{\cancel{5545.5}\ 2522.75}$$

$r^2 =$ 82% of the variation of the dependent variable is defined on the independent variable and rest 92% is not defined.

Date
Page No.

(12)

R_1	R_2	R_3	$d_{12}=R_1-R_2$	$d_{13}=R_1-R_3$	d_{12}^2	d_{13}^2	d_{23}^2	$d_{23}=R_2-R_3$
2	4	5	-2	-3	4	9	1	-1
1	3	8	-2	-7	4	49	25	-5
4	2	4	2	0	4	0	4	-2
6	5	7	1	-1	1	1	4	-2
5	1	10	4	-5	16	25	81	-9
8	6	2	2	6	4	36	16	4
9	8	1	1	8	1	64	49	7
10	9	6	1	4	1	16	9	3
7	10	9	-3	2	9	4	1	1
3	7	3	-4	0	16	0	16	4
			0	0	60	204	206	0

$n=10$ $\Sigma d_{12}^2 = 60$, $\Sigma d_{13}^2 = 204$, $\Sigma d_{23}^2 = 206$

$$r_{12} = 1 - \frac{6\Sigma d_{12}^2}{n(n^2-1)}$$

$$r_{13} = 1 - \frac{6\Sigma d_{13}^2}{n(n^2-1)} = 1 - \frac{6 \times 204}{10(100-1)} = -0.236$$

$$r_{23} = 1 - \frac{6\Sigma d_{23}^2}{n(n^2-1)} = 1 - \frac{6 \times 206}{10(100-1)} = -0.2484$$

$$r_{12} = 1 - \frac{6 \times 60}{10(100-1)} = 0.63$$

~~Here, $r_{12} > r_{13}$ and $r_{12} > r_{23}$~~
~~Then, there is a high degree agreement between judge 1st and 2nd than others.~~
~~∴ Judge 1st and 2nd have the nearest approach to music~~

Pair of first and second judges has the nearest approach to common tastes of music. But r_{13} and r_{23} are negative ie they have opposite taste of music.

Date
Page No.

(29) $r = 0.2$, $n = 10$,
incorrect $d = 9$
correct $d = 7$

$$\Sigma d^2 = \frac{(1-r_s)\, n(n^2-1)}{6}$$

$$= \frac{(1-0.2) \times 10 \times 99}{6}$$

$\Sigma d^2 = 132$

Now,

$\Sigma d^2 = 132 - 9^2 + 7^2 = 100$

$n_c = 10 - 1 + 1 = 10$

Correct rank corelation coefficient is,

$$r = 1 - \frac{6\Sigma d_c^2}{n_c(n_c^2-1)}$$

$$= 1 - \frac{6 \times 100}{100 \times 99}$$

$= 0.94$ #

Date
Page No.

Regression Analysis :-

Regression analysis

Literal meaning of Regression is "stepping back" to an average.

A stastistical device which is used to measure an average mathematical (algebrical) relationship between two or more variables in terms of original units of measurement of data, is known as regretion.

In simple regration anyalsis involves only two variables say x and y where 1 is dependent variable and another is independent variable. If a variable y depends on another variable x then the variable y is known as

(i) Dependent variable or regressed variable or explained variable or effect variable.

(ii) The variable x is known as independent variable or regressor or explanator variable or predictor.

In Simple linear regression anyalysis is carried out by establishing linear relationship between two variable x and y of the form $y = a + bx$, called regression line.

There are two types of regression line.

(i) Regression equation of y on x.

(ii) Regression equation of x on y.

Date
Page No.

(i) Regression equation of y on x.

[To estimate Y, for a given x; we consider it]

Regression eqn of Y on X is,

$$Y = a + bX \quad \text{--- (1)}$$

where 'a' and 'b' are constants called regression parameters.

\# Regression coefficient of Y on X = b_{yx}

\# Slope of regression line of Y on X is b_{yx}

To determine a and b

Date
Page No.

Correlation

(17) $n = 30$ Incorrect observations $= (8,10)\ (12,7)$

$\Sigma X = 120$ Correct " $= (8,12)\ (10,8)$

$\Sigma Y = 90$

$\Sigma X^2 = 600$

$\Sigma Y^2 = 250$

$\Sigma XY = 356$

$n_c = 30 - 2 + 2 = 30$

$\Sigma X_c = 120 - 8 - 12 + 8 + 10 = 118$

$\Sigma X_c^2 = 600 - 8^2 - 12^2 + 8^2 + 10^2 = 556$

$\Sigma Y_c = 90 - 10 - 7 + 12 + 8 = 93$

$\Sigma Y^2_c = 250 - 10^2 - 7^2 + 12^2 + 8^2 = 309$

$\Sigma XY_c = 356 - (8 \times 10) - (12 \times 7) + (8 \times 12) + (10 \times 8)$

$= 356 - 80 - 84 + 96 + 80$

$= 368$

Correct value of r is,

$$r_c = \frac{n_c \Sigma XY_c - \Sigma X_c . \Sigma Y_c}{\sqrt{n_c \Sigma X^2_c - (\Sigma X_c)^2} . \sqrt{n_c \Sigma Y_c^2 - (\Sigma Y_c)^2}}$$

$$= \frac{30 \times 368 - 118 \times 93}{\sqrt{30 \times \cancel{309}\ 556 - (118)^2}\ \sqrt{30 \times 309 - (93)^2}}$$

$$= \frac{11040 - 10974}{\sqrt{16680 - 13924}\ \sqrt{9270 - 8649}}$$

$$= \frac{66}{\sqrt{52.4976} \times \sqrt{24.9199}}$$

$$= \cancel{\frac{66}{130804}} = \frac{66}{1308.235}$$

$= 0.05$ #

Date
Page No.

(18) No of pair of observation $(X, Y) = 15$

σ of X series $= \cancel{3.01}\ 3.01$

σ of Y series $= 3.03$

$\Sigma xy = 122$, where $x = X - \bar{X}$ and $y = Y - \bar{Y}$

$$r = \frac{\Sigma xy}{n \sigma_x . \sigma_y}$$

$$r = \frac{122}{15 \times 3.01 \times 3.03}$$

$= 0.891$

$$P.E(r) = \frac{0.6745(1 - (0.891)^2)}{\sqrt{15}}$$

$$= \frac{0.206}{3.872}$$

$= 0.036.$

Limit of pop^n correlation coefficient.

$[r - PE(r),\ r + PE(r)]$

~~r_c, $[0.036$~~

ie, $[0.891 - 0.036,\ 0.89 + 0.036]$

or, $[0.854,\ 1.926]$ #

(19) $N = 10$

$r = 0.81$

Yes, $PE = \dfrac{0.6745(1 - (0.81)^2)}{\sqrt{10}}$

$= \dfrac{0.6745(0.02)}{\sqrt{10}} = \dfrac{0.0134}{3.16} = 0.004$

Since $\cancel{|r|} \geq 6 \times PE = 0.02$

Since $|r| \geq 6PE$ then r is significant in conclusion.

Date
Page No.

(20) $r = 0.7$

$PE = 0.0344$

$$PE = \frac{0.6745(1-(0.7)^2)}{\sqrt{n}}$$

$$0.0344 = \frac{0.343995}{\sqrt{n}}$$

$$(\sqrt{n})^2 = (9.99)^2$$

$$n = 100 \#$$

(21)

X	Y	X^2	Y^2	XY
21	20	441	400	420
28	35	784	1225	980
28	30	784	900	840
35	28	1225	784	980
35	45	1225	2025	1575
42	40	1764	1600	1680
42	42	1764	1764	1764
$\Sigma X = 231$	$\Sigma Y = 240$	$\Sigma X^2 = 7987$	$\Sigma Y^2 = 8698$	$\Sigma XY = 8239$

$$r = \frac{n\Sigma XY - \Sigma X \cdot \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$= \frac{7 \times 8239 - 231 \times 240}{\sqrt{7 \times 7987 - (231)^2}\sqrt{7 \times 8698 - (240)^2}}$$

$$= \frac{57673 - 55440}{\sqrt{55909 - 53361}\sqrt{60886 - 57600}}$$

$$= \frac{2233}{50.47 \times 57.32}$$

$$= \frac{2233}{2893.12}$$

$$= 0.7717$$

Date
Page No.

$$\frac{0.6745(1-r^2)}{\sqrt{n}}$$

$$PE(r) = \frac{0.6745(1-r^2)}{\sqrt{n}}$$

$$= \frac{0.6745(1-(0.7717)^2)}{\sqrt{7}}$$

$$= \frac{0.6745 \times \cancel{61} = 0.4045}{\sqrt{7}}$$

$$\frac{0.6745(1-r^2)}{\sqrt{n}}$$

$$= \frac{0.2728}{2.64}$$

$$= 0.1033.$$

$|r|$ i.e. $6 \times P.E = 6 \times 0.1033 = 0.619$

Since $|r| \geq 6PE$, then r is significant.

Hoe Rank are not given so we have to arrange.

Ranks of X

(22)

R_1	R_2	$d = R_1 - R_2$	d^2
4	2	2	4
6	3	3	9
3	4	-1	1
9	9	0	0
1	5	-4	16
5	7	-2	4
2	1	1	1
7	10	-3	9
10	8	2	4
8	6	2	4
		$\Sigma d = 0$	$\Sigma d^2 = 52$

Now, Spearman rank

$$r_s = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 52}{10(100-1)}$$

$$= 1 - \frac{312}{990}$$

$$= 0.685.$$

$r_s = 0.685$

23

X	Y	R_x	R_y	$d = R_x - R_y$	d^2
57	83	3	7	-4	16
45	37	1	1	0	0
72	41	6	2	4	16
78	84	7	8	-1	1
53	56	2	3	-1	1
63	85	5	10	-5	25
86	77	9	6	3	9
98	87	10	9	1	1
59	70	4	5	-1	1
71	59	8	4	4	16

$\sum d = 0$; $\sum d^2 = 86$

$$= 1 - \frac{6 \sum d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 86}{10(10^2-1)}$$

$$= 1 - \frac{\cancel{6 \times 7396}\ 516}{990}$$

$$= 0.564$$ #

25

X	Y	R_x	R_y	$d = R_x - R_y$	d^2
48	13	8	5.5	2.5	6.25
33	13	5	5.5	0.5	0.25
40	24	7	10	-3	9
9	6	1	2.5	-1.5	2.25
16	15	3	7	-4	16
16	4	3	1	2	4
65	20	10	9	1	1
24	9	5	4	1	1
16	6	3	2.5	0.5	0.25
57	19	9	8	1	1

$\sum d = 0$

16, 16, 16 3 ta ma divide game

2, 3, 4 $\quad \frac{2+3+4}{3} = 3$

36864

Date
Page No.

Overhead	Units	$U = X-178$	$V = Y-42$	$U.V.$
191	40	13	-2	26
170	42	-8	0	0
272	53	94	11	1034
155	35	23	-7	-161
280	56	107	14	1498
173	39	-5	-3	15
234	48	56	6	336
116	30	-62	-12	744
153	37	-25	-5	125
178	40	0	-2	0
		$\Sigma U = 192$	$\Sigma V = -11$	$\Sigma U.V = 3617$

$$\bar{X} = A + \frac{\Sigma U}{N} \qquad \bar{Y} = A + \frac{\Sigma V}{N.}$$

$$= 178 + \frac{192}{10} \qquad = 42 + \frac{-11}{10}$$

$$= 197.2 \qquad = 40.9$$

$$byx = \frac{n \Sigma U . \Sigma V - \Sigma U . \Sigma V}{n \Sigma U^2 - (\Sigma U)^2}$$

$$\cancel{byx = (10 \cdot 192 \times -11) - (192 \times -11)}$$

$$byx = \frac{n \Sigma U.V - \Sigma U . \Sigma V}{n \Sigma U^2 - (\Sigma U)^2}$$

$$= \frac{10 \times 3617 - (192 \times -11)}{10 \times 28677 - (192)^2}$$

$$\frac{36170 - 2112}{286770 - 36864} = \frac{34058}{249906}.$$

Date
Page No.

Regression

(A) Regression equation of Y on X:
[To estimate Y, we consider it]
Regression eq^n of Y on X, is

$$Y = a + bX \quad \text{(i)}$$

where, a & b are constants, called regression parameters.

\# Regression coefficient of Y on X = b_{yx}
\# Slope of regression line of Y on X = b_{yx}

To determine 'a' and 'b'
The values of 'a' and 'b' are estimated by using least square method,
Using LSM, we get two normal eq^{n}s, which are.

$$\Sigma Y = na + b\Sigma X$$
$$\Sigma XY = a\Sigma X + b\Sigma X^2$$

On solving these two normal equations, we get

$$b_{yx} = \frac{n\Sigma XY - \Sigma X.\Sigma Y}{n\Sigma X^2 - (\Sigma X)^2} = \hat{b}$$

$$a = \bar{Y} - b_{yx}\bar{X} = \hat{a}$$

From eq^n (i), we get,

$\hat{Y} = \hat{a} + \hat{b}X$; which is best fit or estimated or predicted regression line of Y on X.

As an alternate of it → Standardize regression eq^n of Y on X is,

$$Y - \bar{Y} = b_{yx}(X - \bar{X}) \quad \text{(ii)}$$

Date
Page No.

(b) Regression eq^n of X on Y.
[To estimate X, we consider it]
Regression eq^n of X on Y is,

$$X = a + bY \quad \text{(i)}$$

where a and b are constants, called regression parameters

\# Regression coefficient of X on Y = b_{xy}
\# Slope of regression line of X on Y = $\frac{1}{b_{xy}}$

To determine 'a' and 'b'
The value of 'a' and 'b' are estimated by using least square method;
Using LSM, we get two normal eq^{n}s which are,

$$\Sigma X = na + b\Sigma Y$$
$$\Sigma XY = a\Sigma Y + b\Sigma Y^2$$

On solving these two normal eq^{n}s, we get

$$b_{xy} = \frac{n\Sigma XY - \Sigma X.\Sigma Y}{n\Sigma Y^2 - (\Sigma Y)^2}$$

$$a = \bar{X} - b_{xy}\bar{Y} = \hat{a}$$

From eq^n (1), we get;

$\hat{X} = \hat{a} + \hat{b}Y$, which is the 'best fit' or estimated or predicted regression line of X on Y.

\# Standardise regression eq^n of X on Y is,

$$X - \bar{X} = b_{xy}(Y - \bar{Y}) \quad \text{(ii)}$$

Date
Page No.

Regression coefficient: (Calculation)

(A) $$b_{xy} = \frac{n\Sigma XY - \Sigma X\cdot\Sigma Y}{n\Sigma X^2 - (\Sigma X)^2}$$

and $$b_{xy} = \frac{n\Sigma XY - \Sigma X\cdot\Sigma Y}{n\Sigma Y^2 - (\Sigma Y)^2}$$

(B) $$\cancel{b_{xy} = \frac{n\Sigma XY - \Sigma X\cdot\Sigma Y}{n\Sigma Y^2 - (\Sigma Y)^2}}$$

$$b_{yx} = \frac{Cov(X,Y)}{V(X)}$$

and $$b_{xy} = \frac{Cov(X,Y)}{V(Y)}$$

(C) $$b_{yx} = \frac{r\sigma_y}{\sigma_x} \quad \text{and} \quad b_{xy} = \frac{r\sigma_x}{\sigma_y}$$

$r = \frac{Cov(X,Y)}{\sigma_x \cdot \sigma_y}$

$Cov(X,Y) = r\cdot\sigma_x\cdot\sigma_y$

(D) $$b_{yx} = \frac{n\Sigma UV - \Sigma U\cdot\Sigma V}{n\Sigma U^2 - (\Sigma U)^2}$$

and $$b_{xy} = \frac{n\Sigma UV - \Sigma U\Sigma V}{n\Sigma V^2 - (\Sigma V)^2}$$

where $U = X - A$ and $V = Y - B$.

if $U' = \frac{X-A}{h}$ and $V' = \frac{Y-B}{K}$

$$b_{yx} = b_{v'u'} \times \frac{k}{h}$$

and $$b_{xy} = b_{u'v'} \times \frac{h}{K}$$

Date
Page No.

Example:-

Q. Following are two normal eq^{n}s obtained for deriving the regression eq^n of Y on X. ~~Find~~ $5a+10b = 40$; $10a+25b = 95$ then find regression line of Y on X.

$5\ (5a+10b = 40)$
$10a+25b = 95.$
$a=3 \quad b=2$
$y = 2 + 3x$

Example:-

If regression coefficient of Y on X is 2.4 and $U = 2x+5$ and $V = \frac{y}{2} - 8$. Then find regression coefficient of V on U.

$U = 2x+5$ $\qquad \frac{k}{h} = 2 \qquad 2x = \left(\frac{x}{\frac{1}{2}}\right)$

$V = \frac{y}{2} - 8.$

$b_{yx} = b_{vu} \times \frac{k}{h}$

$b_{yx} = b_{v\cdot u} \times \frac{2}{\left(\frac{1}{2}\right)}$

$2.4 = b_{v\cdot u}\cdot 4$

$b_{v\cdot u} = 0.6$

Properties of regression

- Correlation coefficient is G.M of the two regression coefficients for a pair of variables say X and Y that is r is the G.M of b_{yx} and b_{xy}
  i.e $r = \pm\sqrt{b_{yx}\cdot b_{xy}}$

- Correlation coefficient and the two regression coefficient have the same sign i.e r b_{yx}, b_{xy} have the same sign.

3 - Product of the two regression coefficients for a pair of variables is always less or equal to unity.
i.e $b_{yx} \cdot b_{xy} \leq 1$

4 - If one of the regression coefficient is greater than unity then the other must be less than unity i.e if $b_{yx} > 1$ then $b_{xy} < 1$. and if $b_{xy} > 1$ then $b_{yx} < 1$.

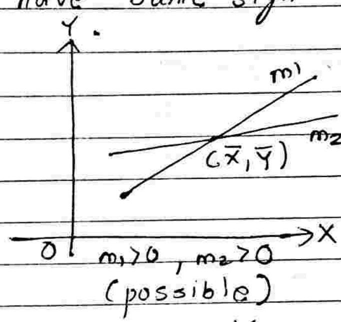
5 - A.M of the two regression coefficients is always greater or equal to the co-relation coefficient i.e

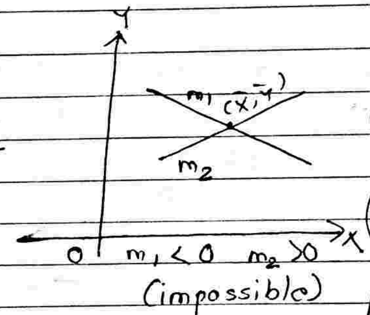
$$\frac{b_{yx} + b_{xy}}{2} \geq r$$

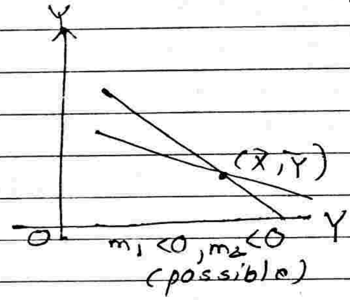
6 - Regression coefficients are independent on change of origin but not of the scales.
i.e $U = X - A$ & $V = Y - B$ then
(i) $b_{yx} = b_{vu}$ & $b_{xy} = b_{uv}$
and if $V = a + bX$ and $V = c + dY$
(i) $b_{yx} = b_{vu} \times \frac{1/d}{1/b}$
(ii) $b_{xy} = b_{uv} \times \frac{1/b}{1/d}$

7 - The two regression lines always pass through $(\bar{X}, \bar{Y})$. In other word $(\bar{X}, \bar{Y})$ is the point of intersection of the two regression lines.

8 - ~~The two regression line have~~
8 - Slopes of the two regression line have same sign.







$a_1 x + b_1 y = c_1$
$a_2 x + b_2 y = c_2$

(A) $\frac{a_1}{a_2} \neq \frac{b_1}{b_2}$ (Intersecting)

(B) $\frac{a_1}{a_2} = \frac{b_1}{b_2} = \frac{c_1}{c_2}$ (Coincident or identical)

(C) $\frac{a_1}{b_1} = \frac{a_2}{b_2} \neq \frac{c_1}{c_2}$ (Parallel)

Slope (m) = $-\frac{x \text{ coeff}}{y \text{ coeff}}$

(1) Let x = mark in Math
y = ~~mark~~ in Stat

X	Y	U = X − 32	V = X − 36	UV	U^2	V^2
25	43	−7	7	−49	49	49
28	46	−4	10	−40	16	100
35	49	−3	13	39	9	169
32	41	0	5	0	0	25
31	36	−1	0	0	1	0
36	32	4	−4	−16	16	16
29	31	−3	−5	+15	9	25
38	30	6	−6	−36	36	36
34	33	2	−3	−6	4	9
32	39	0	3	0	0	9

$\Sigma U = 6$ $V = 20$ $UV = -123$ $U^2 = 140$ $V^2 = 436$.

Date
Page No.

Here, $n = 10$, $A = 32$, $B = 36$,
$\Sigma U = 0$: $\Sigma V = 20$, $\Sigma UV = -\cancel{123}\ 93$
$\Sigma U^2 = 140$, $\Sigma V^2 = 438$.

$\bar{X} = A + \frac{\Sigma U}{n}$ | $\bar{Y} = B + \frac{\Sigma V}{n}$

$= 32 + \frac{0}{10}$ | $= 36 + \frac{20}{10}$

$= 32$ | $= 38$

$$b_{yx} = \frac{n\Sigma UV - \Sigma U \cdot \Sigma V}{n\Sigma U^2 - (\Sigma U)^2}$$
$$= \frac{10 \times (-\cancel{123}\ 93) - 0 \times 20}{10 \times 140 - (0)^2}$$
$$= \frac{-930}{140} = -0.6643$$

$$b_{xy} = \frac{n\Sigma UV - \Sigma U \cdot \Sigma V}{n\Sigma V^2 - (\Sigma V)^2}$$
$$= \frac{10 \times (-93) - (0 \times 20)}{(10 \times 438) - (20)^2}$$
$$= \frac{-930}{3980}$$
$$= -0.2337$$

(a) Regression eq^n of Y on X is,
$Y - \bar{Y} = b_{yx}(X - \bar{X})$
or, $Y - 38 = -0.664(X - 32)$
or, $Y - 38 = -0.6643X + 21.2576$
or, $Y = 59.2576 - 0.6643X$.
~~or, Y.~~

③ Y on X
⑥ take 15.
⑦ X = 90

Date
Page No.

(b) Also, regression eq^n of X on Y is,
$X - \bar{X} = b_{xy}(Y - \bar{Y})$
or, $X - 32 = -0.2337(Y - 38)$
or, $X - 32 = -0.2337Y + 8.8806$
or, $X = 40.8806 - 0.22337Y$ #

(b) The coefficient of correlation between marks in math and sat is given by:
$r = \pm\sqrt{b_{yx} \cdot b_{xy}}$
Here, $b_{yx} < 0$, $b_{xy} < 0$ then,
r is also negative.
$\therefore r = -\sqrt{(-0.6643) \times (-0.2337)}$
$= -0.3940$

(c) For $X = 30$, $Y = ?$
To estimate a value of Y we have to consider the regression equation of Y on X which is
$Y = 59.2576 - 0.6643X$.
For $X = 30$
$Y = 59.2576 - 0.6643 \times 30$
$= 59.2576 - 19.929$
$= 39.327$ #
$= 39$ marks (approx)

$2x - 5y = 10$
$5x - 7y = 11$
$b_{yx} = 2/5$ (Reciprocal of slope)
$b_{xy} = 7/5$

$\begin{pmatrix} 2 & 5 \\ 5 & 7 \end{pmatrix} = 2 \times 7 - (5 \times 5)$
$= -ve$

$5x + y = 10$, X on Y
$x + y = -3$

$(5\ \ -1)$
$(1\ \ 1)$
Y on X.

$r^2 = \frac{1}{5}$, Same True
r is less because ... trend.

Date

Page No.

Note no 1 $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$ are two regression equation.

a) These two eqns are consistant if they are non-parallel and sign of their slopes are same.

b) $a_1x + b_1y + c_1 = 0$ is the regression eqn of Y on X if $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} < 0$

Note no 2 If $a_1x + b_1y + c_1 = 0$ and $a_2x + b_2y + c_2 = 0$ be the regression eqn of Y on X and X on Y respectively.

(a) $r^2 = \frac{a_1b_2}{a_2b_1}$ if $a_1b_2 < a_2b_1$

(b) $r = +\sqrt{\frac{a_1b_2}{a_2b_1}}$ if slopes of the lines are +ve

$= -\sqrt{\frac{a_1b_2}{a_2b}}$ if slopes of the lines are -ve

(c) $b_{yx} = -\frac{a_1}{b_1}$

$b_{xy} = -\frac{b_2}{a_2}$

Date

Page No.

(18) $n = 30$, $\Sigma x = 120$, $\Sigma x^2 = 600$

$\Sigma y = 90$, $\Sigma y^2 = 250$, $\Sigma xy = 356$

wrong pair of observations $= (8,10)(12,7)$

Correct pair of observations $= (8,12)(10,8)$

Corrections of calculated information.

$n_c = 30 - 2 + 2 = 30$

$\Sigma X_c = 120 - 8 - 12 + 8 + 10 = 18$
$= 118$

$\Sigma X^2c = 600 - 8^2 - 12^2 + 8^2 + 10^2 = 556$

$\Sigma y_c = 90 - 10 - 7 + 12 + 8 = 93$

Σy^2_c ~~$= 356 - (8\times10) - (12\times7) + (8\times12) + (10\times8)$~~
~~$= 309$~~.
$= 250 - 10^2 - 7^2 + 12^2 + 8^2 = 309$.

$\Sigma Xy_c = 356 - (8\times10) - (12\times7) + (8\times12) + (10\times8)$
$= 368$.

(1) Correct value of correlation coefficient is,

$$r_c = \frac{n_c \Sigma XY_c - \Sigma X_c \cdot \Sigma Y_c}{\sqrt{n_c \Sigma X_c^2 - (\Sigma X_c)^2}\sqrt{n\Sigma Y_c^2 - (\Sigma Y_c)^2}}$$

$$= \frac{(30\times368) - (118\times93)}{\sqrt{(30\times556) - (118)^2}\sqrt{(30\times38) - (93)^2}}$$

$$= \frac{11040 - \cancel{11564}\ 10974}{\sqrt{(16680) - (13924)}\sqrt{1140 - 8849}}$$

$$= \frac{\cancel{-524}\ 66}{\sqrt{2756}\sqrt{-7709}}$$

$= 0.05$.

$\bar{X}_c = \frac{\Sigma X_c}{n_c} = \frac{118}{30} = 3.93$.

$\bar{Y}_c = \frac{\Sigma Y_c}{n_c} = \frac{93}{30} = 3.1$.

Date
Page No.

$$b_{yx_c} = \frac{n_c \Sigma XY_c - \Sigma X_c \cdot \Sigma Y_c}{n_c \Sigma X_c^2 - (\Sigma X_c)^2}$$

$$= \frac{30 \times 368 - 118 \times 309}{30 \times 556 - (118)^2}$$

$$= 0.0234.$$

\# Regression eqⁿ of Y on X is,

$$Y_c - \bar{Y}_c = b_{yx_c}(X_c - \bar{X}_c)$$

or, $Y_c - 3.1 = 0.0234\,(X_c - 3.93)$

or, $Y_c = 0.0234X_c - (0.0234 \times 3.93) + 3.1$

or, $Y_c = 0.0234\,X_c - 3.006$ #

(iii) For $X = 25$

$Y = 0.0234\,X_c + 3.006$

$= (0.023 \times 25) + 3.006$

$= \cancel{2.431} = 3.591$ #

(10) The two regression eqⁿs are,

$3x + 2y - 26 = 0$ —— (i)

$6x + y - 31 = 0$ —— (ii)

(i) Since the two regression eqⁿ always pass through $(\bar{X}, \bar{Y})$

i.e $(3\bar{X} + 2\bar{Y}) - 26 = 0$

and $6\bar{x} + \bar{y} - 31 = 0$

$\cancel{3\bar{x} + 2\bar{y} = 26}$

$\bar{y} = -10.$

Date
Page No.

$\bar{X} = 4.$

(i) Since, the two regression eqⁿs always pass through $(\bar{X}, \bar{Y})$

i.e $3\bar{x} + 2\bar{y} - 26 = 0$

$6\bar{x} + \bar{y} - 31 = 0$

Also,

(ii) Suppose eqⁿ (i) & (ii) are the regression eqⁿ of Y on X and X on Y respectively.

from (i)

$2y = 26 - 3x$

or, $y = 13 - 1.5x \quad \therefore b_{yx} = -1.5.$

Also, from (ii)

$6x = 31 - y$

$\therefore x = 5.16 - \frac{1}{6}y$

$\therefore b_{xy} = -\frac{1}{6}.$

Now,

$b_{yx} \cdot b_{xy} = -1.5 \times -\frac{1}{6} = 0.25$

Here,

$b_{yx} \cdot b_{xy} < 1$

$\therefore$ Our assuption is correction.

Hence It is illustrated that Eqⁿ (i) & (ii) are the regression eqⁿ of Y on X and X on Y respectively.

& ~~bxy=~~ $b_{yx} = -1.5$ and $b_{xy} = -\frac{1}{6}.$

(c) Coefficient of co-rellation is,

$$r = \pm\sqrt{b_{ym} \cdot b_{my}}$$

$$\text{or, } r = -\sqrt{0.25} \quad [\text{Since; } b_{xy}<0 \;\&\; b_{xy}<0]$$

$$= -0.5$$

(d) To estimate a value of Y.
We have to consider the regression eqn of Y on X, which is given by,

$$3x+2y-26=0.$$

For $X=5$.

$$3\times5+2y-26=0$$
$$15+2y-26=0$$
$$\therefore y = 5.5.$$

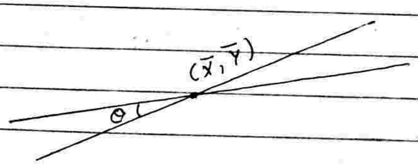
(e) σY if $\sigma x^2 = 25$ $\therefore \sigma x = 5$.

$$\cancel{b_{xy}} = b_{yx} = \frac{r\sigma_y}{\sigma_X}$$

$$-1.5 = \frac{-0.5\times\sigma_y}{5}$$

$$15 = \sigma_y$$

Angle between two regression lines:-



Let,

M_1: Slope of regression eqn of Y on X.
M_2: Slope " " " of X on Y.

$$M_1 = b_{yx} = \frac{r\sigma_y}{\sigma_X}$$

$$M_1 = b_{xx} = \frac{r\sigma_y}{\sigma_x}.$$

$$M_2 = \frac{1}{b_{xy}} = \frac{1}{\frac{r\sigma_x}{\sigma_y}} = \frac{\sigma_y}{r\sigma_y}$$

Case I

If the two regression lines are coincident (identical)
(i.e $\theta = 0°$)
then,

$$r^2 = 1$$

i.e $r = \pm1$ (Perfect co-relation)

Case II

If the two regression lines are perpendicular to each other (i.e $\theta=90°$).
~~i.e~~
then $r=0$ (No linear corellation).

$m_1 \cdot m_2 = -1$
$\frac{r\sigma_y}{\sigma_x} \cdot \frac{\sigma_y}{r\sigma_x} = -1$
$\frac{r\sigma_y^2}{r\sigma_x^2} = 1$
$\frac{x}{x} = -1, x \neq 0$ ...

Regression (Numerical Problem)

(2)

X	Y	$U = X - 3$	$V = Y - 8$	U^2	V^2	UV
0	5	-3	-3	9	9	9
2	7	-1	-1	1	1	1
3	8	0	0	0	0	0
5	10	2	2	4	4	4
6	12	3	4	9	16	12
		$\Sigma U = 1$	$\Sigma V = 2$	$\Sigma U^2 = 23$	$\Sigma V^2 = 30$	$\Sigma UV = 26$

Here, $n = 5$, $A = 3$ $B = 8$

$\Sigma U = 1$ $\Sigma V = 2$ $\Sigma U^2 = 28$ $\Sigma V^2 = 30$ $\Sigma UV = 26$

$$\bar{X} = A + \frac{\Sigma U}{n} = \frac{3 + 1}{5} = \frac{16}{5} = 3.2$$

$$\bar{Y} = B + \frac{\Sigma V}{n} = 8 + \frac{2}{5} = \frac{42}{5} = 8.4$$

$$b_{yx} = \frac{n \Sigma UV - \Sigma U \cdot \Sigma V}{n \Sigma U^2 - (\Sigma U)^2} = \frac{5 \times 26 - 1 \times 2}{5 \times 28 - (1)^2} = \frac{130 - 2}{140 - 1} = \frac{128}{139} = 0.9209$$

Regression eqn of Y on X is,

$$Y - \bar{Y} = b_{yx} (X - \bar{X})$$

$$Y - 8.4 = 0.9209 (X - 3.2)$$

$$Y - 8.4 = 0.9209 X - 2.94688$$

$$Y = 0.9209 X + 5.45312$$

Here, for $X = 4$, $Y = ?$

$$Y = 0.9209 X + 5.45312$$

$$Y = 0.9209 \times 4 + 5.45312$$

$$Y = 3.6836 + 5.45312$$

$$Y = 9.13672 \text{ lbs.}$$

(b)

X	Y	$U = X - 5$	$V = Y - 3$	U^2	V^2	UV
1	1	-4	-2	16	4	8
3	2	-2	-1	4	1	2
5	3	0	0	0	0	0
6	4	1	1	1	1	1
5	5	0	2	0	4	0
		$\Sigma U = -5$	$\Sigma V = 0$	$\Sigma U^2 = 21$	$\Sigma V^2 = 10$	$\Sigma UV = 11$

Here,

$n = 5$, $A = 5$ $B = 3$ $\Sigma U = -5$ $\Sigma V = 0$

$\Sigma UV = 11$ $\Sigma U^2 = 21$ $\Sigma V^2 = 10$.

$$\bar{X} = A + \frac{\Sigma U}{n} = 5 + \frac{-5}{5} = 5 - 1 = 4$$

$$\bar{Y} = B + \frac{\Sigma V}{n} = 3 + \frac{0}{5} = 3 + 0 = 3.$$

$$b_{xy} = \frac{n \Sigma UV - \Sigma U \cdot \Sigma V}{n \Sigma V^2 - (\Sigma V)^2} = \frac{5 \times 11 - (-5) \cdot 0}{5 \times 10 - (0)^2} = \frac{55 + 0}{50 - 0} = 1.1$$

Date
Page No.

Regression eq^n of X on Y.

$$(X-\bar{X}) = b_{xy}\,(Y-\bar{Y})$$
$$(X-4) = 1.1\,(Y-3)$$
$$X-4 = 1.1Y - 3.3$$
$$X = 1.1Y + 0.7 \;\#$$

(4)

Year	X	Y	U = X − 720	V = Y − 1300	U^2	V^2	UV
1	600	1250	−120	−50	14400	2500	6000
2	630	1100	−90	−200	8100	40,000	18000
3	720	1300	0	0	0	0	0
4	750	1350	30	50	900	2500	1500
5	800	1500	80	200	6400	40,000	16000
			$\Sigma U = -100$	$\Sigma Y = 0$	$\Sigma U^2 =$ 29800	85000	41500

$n = 5$, $\Sigma U = 100$, $\Sigma V = 0$, $\Sigma U^2 = 29800$, $\Sigma V^2 = 85{,}000$, $UV = 41500$

$$b_{yx} = \frac{n\,\Sigma UV - \Sigma U\cdot\Sigma V}{n\,\Sigma U^2 - (\Sigma U)^2}$$
$$= \frac{5 \times 41500 - 100 \times 0}{5 \times \cancel{100}\,29800 - (100)^2}$$
$$= \frac{207500}{139000} = 1.4928$$

Now,

$$(Y-\bar{Y}) = b_{yx}\,(X-\bar{X})$$
$$(Y-700) = 1.4928\,(X-1300)$$
$$Y-700 = 1.4928X - 1940.64$$
$$Y$$

Date
Page No.

$$\bar{Y} = A + \frac{\Sigma U}{n} \qquad \bar{X} = A + \frac{\Sigma V}{n}$$
$$= 720 - \frac{100}{5} \qquad \bar{X} = 1300 + 0$$
$$= 700 \qquad \bar{X} = 1300$$

$$b_{xy} = \frac{n\,\Sigma UV - \Sigma U\cdot\Sigma V}{n\,\Sigma V^2 - (\Sigma V)^2}$$
$$= \frac{5 \times 41500 - 100 \times 0}{5 \times 85{,}000 - (0)^2}$$
$$= \frac{207500}{425000}$$
$$= 0.488235$$

$$(X-\bar{X}) = b_{xy}\,(Y-\bar{Y})$$
$$(X-1300) = 0.488235\,(Y-700)$$
$$(X-1300) = 0.4882Y - 341.76$$
$$X = 0.4882Y + 958.24$$

Now,

$$Y = 850$$
$$X = 0.4882 \times 850 + 958.24$$
$$= 1372.97 \;\#$$

Date
Page No.

(7) (8)

Year	X	Y	$U=X-72$	$V=Y-130$	U^2	V^2	UV
1	60	125	-12	-5	144	25	60
2	63	110	-9	-20	81	400	180
3	72	130	0	0	0	0	0
4	75	135	3	5	9	25	15
5	80	150	8	20	64	400	160
			$\Sigma U=-10$	$\Sigma V=0$	$\Sigma U^2=298$	$\Sigma V^2=850$	$\Sigma UV=415$

$n=5 \quad \Sigma U=-10 \quad \Sigma V=0 \quad \Sigma U^2=298$

$\Sigma V^2=850 \quad \Sigma UV=415$.

$\bar{Y} = A + \frac{\Sigma U}{n}$ $\qquad$ $\bar{X} = A + \frac{\Sigma V}{n}$

$= 72 + \frac{-10}{5}$ $\qquad$ $= 130 + \frac{0}{5}$

$= 72 - 2$ $\qquad$ $= 130 + 0$

$= 70$ $\qquad$ $= 130$.

$$byx = \frac{n\Sigma UV - \Sigma U \cdot \Sigma V}{n\Sigma U^2 - (\Sigma U)^2}$$

$$= \frac{5 \times 415 - (-10) \times 0}{5 \times 298 - (-10)^2}$$

$$= \frac{2075}{1490} = \cancel{1.49280} = 1.49280 \quad \cancel{=1.39261}$$

Regn on eqn Y on X

$Y - \bar{Y} = byx(X - \bar{X})$

$Y - 70 = 1.49280 (X - 130)$

$Y - 70 = 1.49280X - 194.06$

$Y = 1.49280X - 124.06$

$$bxy = \frac{n\Sigma UV - \Sigma U \cdot \Sigma V}{n\Sigma V^2 - (\Sigma V)^2}$$

$$= \frac{5 \times 415 - (-10) \times 0}{5 \times 850 - (0)^2} = \frac{2075}{4250} = 0.48823$$

Date
Page No.

~~X = 15,000~~ ~~X = 90,000~~

~~i.e 15~~

~~Y = 1.49280 × 90 − 124.06~~

$Y - \bar{Y} = byx (X - \bar{X})$

$Y - 70 = byx (X - 130)$ 1.39261

$Y - 70 = \cancel{1.49280} (X - 130)$.

~~Y − 70 = 1.49280X − 194.064~~

$Y = \cancel{1.49280}X - 124.064$

~~Y = 1.49280 × 90 − 124.064~~

~~= 134.352~~ − 124.064.

$*$ $Y - 70 = 1.39261X - 181.0393$.

$Y - 70 = 1.39261 \times 90 - 181.0393$

$Y = 125.3349 - 111.0393$.

Regression eqn Y on X.

$(Y - \bar{Y}) = byx (X - \bar{X})$

$Y - 70 = byx (X - 130)$.

$Y - 70 = 1.49280 (X - 130)$.

$Y - 70 = \cancel{194.064X}$

$Y - 70 = 1.49280X - 194.064$

$Y = 1.49280X - 124.064$.

when $X = 90$ then since 90,000.

$Y = 1.49280X - 124.064$.

$Y = 1.49280 \times 90 - 124.064$.

$= 10.288$.

$r = \pm\sqrt{byx \cdot bxy}$

$r = \pm\sqrt{1.49280 \times 0.48823}$

$= \pm\sqrt{0.4183}$

$= 0.86$ #.

Date
Page No.

(10) $\bar{X} = 53.2$

$\bar{Y} = 27.9$

Regression coefficient of Y on X (byx) = −1.5

Regression coefficient of X on Y (bxy) = −0.2

The most likely value of Y when X = 60 is

$$Y - \bar{Y} = byx(X - \bar{X})$$

$$(Y - 27.9) = -1.5(\cancel{53.2\ X}\ X - 53.2)$$

$$Y - 27.9 = -1.5X + 79.8$$

$$Y = -1.5X + 107.7$$

When $X = 60$

$$Y = -1.5 \times 60 + 107.7$$

$$Y = -90 + 107.7$$

$$Y = 17.7.$$

The coefficient co-relation between X and Y is,

$$r = \pm\sqrt{byx \cdot bxy}$$

Here, bxy < 0 and byx < 0 then r is also negative.

$$r = -\sqrt{(-1.5) \times (-0.2)}$$

$$r = -\sqrt{0.3}$$

$$r = -0.547722$$

(11) $\bar{X}$ no of workers on strike = 800 (x)

$\bar{X}$ loss of daily production = 35 (Y).

σ of daily production 000 = ~~100~~ 2

σ of daily no of workers on strike = 100

Now,

Date
Page No.

Coefficient of correlation betn no of workers on strk and daily productn (r) = 0.8

(Assume 6 working days in a week).

~~Here,~~

$$b_{yx} = r \times \frac{\sigma y}{\sigma x}$$

$$= 0.8 \times \frac{2}{100}$$

$$= \cancel{0.0016}\ 0.016.$$

Now,

The regression eqn on Y on X is

$$Y - \bar{Y} = byx(X - \bar{X}).$$

$$Y - 35 = 0.016(x - 800).$$

~~Y~~ When $x = 1800$,

$$Y = 0.016 \times 1800 - 0.016 \times 800 + 35.$$

∴ Y = 51 [000] × 6 [working] = 6 d.

= 306,000

~~12.8 + 35.~~

−22.2 +22.2.

51 × 6

306 × 000

306,000

$b_{yx} = r \times \frac{\sigma x}{\sigma y}$

Date
Page No.

Differentiate between correlation and regression

Correlation	Regression
(1) It measures the degree to which the variables are linearly related.	(1) It is a measure an average mathematical relationship between two variable whether they are linerally or non-linerally related.
(2) Correlation coefficients are symmetrical regarding their variables. i.e $r_{xy} = r_{yx}$.	(2) Regression coefficients are non-symmetrical regarding their variables i.e $b_{yx} \neq b_{xy}$.
(3) Corelation need not imply cause and effect relationship between the variable.	(3) Regression clearly indicates cause and effect relationship between the variables.
(4) Correlation coefficient is independent on units of pair of data.	(4) Regression coefficients is dependent on units of the variables.
(5) It is independent on change of origin or scale but depends of the sign of scale factor	(5) It is independent on change of orgin but dependent on change of scale.

Date
Page No.

Least square method

Least square method involves minimizing "sum of square of errors" i.e Σe^2 is minimized.

Errors (or Residue):-

Errors may be defined as the difference between observed and estimated value of a variable ~~from~~ for a given for a given value of another variable.

e = Observed value − Estimated value.

There are two types of errors commited in regression anyalysis

(1) Verticle errors
$e = Y - \hat{Y}$; which is ∥el to Y-axis.

(2) Horizontal error
$e = x - \hat{x}$; which is parallel to X-axis.

Notes:- Regression line of y on x is obtained by minimizing sum of square of verticle errors. Regression line of x on y is obtained by minimizing sum of square of horizontal errors.

Date
Page No.

Time series analysis:-

The arrangment of statistical data accrading to the ~~acco~~ occurance of time (chronological order) is known as time series analysis. And the stastistical anyalysis of the variation of time series due to the chronological ordering is known as time series analysis.

(#) Importance of time series analysis.

- It helps in forcasting the value of variables.
- It helps to know the past behaviour of data during certain time interval. It helps to estimate the rate of change of time series value per unit time.
- It entirely helps in business planning, policy making and decision making on the basis of scientific point of view.
-

Components of time series.

- The various forces which affect the smooth flow of time series are known as components of time series.
- The four basis component of time series are:-

1. Secular trend (long term fluctuation)
2. Cyclical variation (short term variation) or Business cycle.

Date 19-26
Page No.

3. Seasonal variation (Short term fluctuation)
   - Natural S.V
   - Artificial S.V.
4. Random or irregular variation.

(9)

X	Y	U = X − 22	V = X − 45	U^2	V^2	UV
14	31	−8	−14	64	196	112
19	36	−3	−9	9	81	27
24	48	2	3	4	9	6
21	37	−1	−8	1	64	8
26	50	4	5	16	25	20
22	45	0	0	0	0	0
15	33	−7	−12	49	144	84
20	41	−2	−4	4	16	8
19	39	−3	−6	9	36	18
		$\Sigma U = -17$	$\Sigma V = -45$	156	571	283

$n = 9$ $\quad \Sigma U = -17$ $\quad \Sigma V = -45$ $\quad \Sigma U^2 = 156$ $\quad \Sigma V^2 = 571$

$\Sigma UV = 283$

$$\bar{Y} = A + \frac{\Sigma U}{n} = 22 + \frac{(-17)}{9} = 20.1111$$

$$\bar{X} = A + \frac{\Sigma V}{n} = 45 + \frac{(-45)}{9} = 40$$

$$byx = \frac{n\Sigma UV - \Sigma U . \Sigma V}{n\Sigma U^2 - (\Sigma U)^2} = \frac{9 \times 283 - (-17) \cdot (-45)}{9 \times 156 - (-17)^2} = \frac{2547 - 765}{1404 - 289} = \frac{1782}{1115} = 1.59820$$

$$bxy = \frac{n\Sigma UV - \Sigma U \cdot \Sigma V}{n\Sigma V^2 - (\Sigma V)^2} = \frac{9 \times 283 - (-17) \times (-45)}{9 \cdot 571 - (-45)^2} = \frac{2547 - 765}{5139 - 2025} = \frac{1782}{3114} = 0.57225.$$

Date

Page No.

$r = \pm\sqrt{b_{xy} \cdot b_{yx}}$

$= \sqrt{1.59820 \times 0.57225}$

$= \sqrt{0.914}$

$= 0.95633.$

$\therefore PE = 0.6745\,\dfrac{(1 - (0.9472)^2)}{\sqrt{9}}$

$= \dfrac{0.6745 \times \cancel{0.0528}\ 0.1029}{3}$

$= \dfrac{0.06940605}{3}$

$= 0.023$ #

$r > 6\,PE(r)$ then significant $PE(r)$

0.95633 $\quad = 6 \times 0.023$

$= 0.1314$

$0.95633 > 0.1314$

So, Yes it indicates the termination of services of low-test sorvouies is justified.

When

$(X - \bar{X}) = b_{xy}(Y - \bar{Y})$

$(X - 20.12) = 0.57225\,(Y - 40)$

$X - 20.12 = 0.57225\,(Y - 40)$

$X - 20.12 = 0.57225Y - 22.89$

$X = 0.57225Y - 2.77$

When $Y = 30$ i.e 300

$X = 0.57225 \times 30 - 2.77$!

$= 17.1675 - 2.77$

$= 14.39$ #

Date

Page No.

(12) $n = 50$ Let X = Expendite on accomodation

Y = Expendite on food & entertainment.

$\Sigma X = 8500$ $\quad \sigma_x = 60$ $\quad r = 0.6$

$\Sigma Y = 9600$ $\quad \sigma_y = 20$

$b_{yx} = r \times \dfrac{\sigma_y}{\sigma_x}$

$= 0.6 \times \dfrac{20}{60}$

$= 0.1999$ i.e 0.2.

Now,

The regression eqn of Y on X is,

$(Y - \bar{Y}) = b_{yx}(X - \bar{X})$

$Y -$

Here,

$\Sigma X = 8500$ $\quad \bar{X} = \dfrac{8500}{n} = \dfrac{8500}{50} = 170$

$\Sigma Y = 9600$ $\quad \bar{Y} = \dfrac{9600}{n} = \dfrac{9600}{50} = 192$

$(Y - \bar{Y}) = b_{yx}(X - \bar{X})$

$(Y - 192) = 0.2(X - 170)$

$Y - 192 = 0.2X - 34$

$Y = 0.2X + 158$

When $X = 200$ then,

$Y = 0.2 \times 200 + 158$

$Y = 198$ #.

(13) $\Sigma X = 60$, $\Sigma Y = 40$, $\Sigma XY = 1150$
$\Sigma X^2 = 4160$, $\Sigma Y^2 = 1720$, $N = 10$.

$\bar{X} = \frac{60}{10} = 6$

$\bar{Y} = \frac{40}{10} = 4$.

We have, regression coefficient of Y on X is,

$b_{yx} = r \frac{\sigma_y}{\sigma_x} = \frac{\sigma_{xy}}{\sigma_x^2}$

or, $b_{yx} = \frac{\text{cov}(X,Y)}{\sigma_x^2}$

$= \frac{\frac{1}{n}\Sigma(X-\bar{X})(Y-\bar{Y})}{\frac{1}{n}\Sigma(X-\bar{X})^2}$

$= \cancel{\frac{1}{10}} \cancel{\Sigma} = \frac{\Sigma xy}{\Sigma x^2}$

$\therefore b_{yx} = \frac{1150}{4160} = 0.27644$.

The regression coefficient of X on Y is,

$b_{xy} = \frac{\Sigma xy}{\Sigma y^2} = \frac{1150}{1720} = 0.66860$

The regression eqn of Y on X is,

$(Y-\bar{Y}) = b_{yx}(X-\bar{X})$

$\cancel{Y - 4 = 0.66860(X-6)}$

$\cancel{Y - 4 = 0.66860X - 4.0116}$

$\cancel{Y = 0.66860X - 0.0116}$

when $(Y-4) = 0.27644(X-6)$

$Y - 4 = 0.27X - 1.648$

$Y = 0.27X + 2.5636$ #

The regression eqn of X on Y is,

$(X-\bar{X}) = b_{xy}(Y-\bar{Y})$

$(X-6) = b_{xy}(Y-4)$

$(X-6) = 0.66860(Y-4)$

$X - 6 = 0.66860Y - 2.6744$

$X = 0.66860Y + 3.3256$ #

(14) Average price in market A $\bar{X}$ = Rs 67
Average price in market B $\bar{Y}$ = Rs 65.
Coefficient of variance at market A ~~$b_{xy} = 5.22$~~
~~$b_{yx} = 3.85$.~~
~~$r = 0.82$~~

(A) $CV = \frac{\sigma_x}{\bar{X}} \times 100$

$5.22 = \frac{\sigma_x}{67} \times 100$ $\boxed{\sigma_x = 3.49774}$

~~The regression eqn of X of Y is,~~

~~$(X-\bar{X}) = b_{xy}(Y-\bar{Y})$~~

~~$(X-67) = 5.22(Y-65)$~~

~~$X - 67 = 5.22Y - 339.3$~~

~~$X = 5.22Y - 332.3$~~

~~when $Y = 75$.~~

~~$X = 5.22 \times 75 - 332.3$ 265.3~~

~~$= 391.5 - 332.3$~~

(B) $CV = \frac{\sigma_y}{\bar{Y}} \times 100$ = $3.85 = \frac{\sigma_y}{65} \times 100$

$\boxed{\sigma_y = 2.5025}$

The $b_{xy} = \frac{\sigma_y}{\sigma_x} \times 0.82$

$b_{xy} = \frac{2.5025}{3.49774} \times 0.82$

$= 0.58667$. $(Y = 75)$

$(X - \bar{X}) = b_{xy}(Y - \bar{Y})$.

or, $(X - 67) = 0.58667(Y - 65)$

or, $X - 67 = 0.58667 \times 75 - 38.134123$

or, $X = 44.00025 + 28.865877$

$= 72$ #

Date
Page No.

(15) $\bar{X} = 80$ $\bar{Y} = 50$ $\sigma_x = 15$ $\sigma_y = 10$
$r = -0.40$
(X) Mathematics (Y) = English.

$b_{xy} = r \times \frac{\sigma_x}{\sigma_y}$
$= -0.40 \times \frac{15}{10}$
$= -0.40 \times 1.5$
$= -0.60$

$(X - \bar{X}) = b_{xy} (Y - \bar{Y})$.
$(X - 80) = -0.60 (Y - 50)$
$(X - 80) = -0.60Y - 30$
$X = -0.60Y + 110$
$Y = 60$
$X = -0.60 \times 60 + 110$
$= -36 + 110$
$= 74$

Date
Page No.

(17) $\Sigma X = 50$ $\bar{X} = 5$, $\Sigma Y = 60$ $\bar{Y} = 6$ $\Sigma XY = 350$
Variance of $X = 4$, Variance of $Y = 9$.
Variance of X i.e $\sigma_x^2 = \frac{1}{n} \Sigma X^2$

$4 = \frac{1}{n} \Sigma X^2$

$\Sigma(X - \bar{X})(Y - \bar{Y}) = \Sigma xy = 350$
$350 = (X - \bar{X})(Y - \bar{Y})$
$350 = (50 - \bar{x})$

Variance of $y = \sigma_y^2 = 9$

$\frac{\Sigma X}{\bar{X}} = \frac{50}{5} = 10$.

$b_{yx} = \frac{\text{covariance of } (x, y)}{\text{Variance of } x}$
$= \frac{\frac{\Sigma xy}{n} - \frac{\Sigma x}{n} \cdot \frac{\Sigma y}{n}}{4}$
$= \frac{\frac{350}{10} - \frac{50}{10} \cdot \frac{60}{10}}{4}$
$= \frac{35 - 5 \times 6}{4}$
$= \frac{5}{4} = 1.25$.

$b_{xy} = \frac{\text{Covariance of } (X, Y)}{\text{Variance of } y}$
$= \frac{\frac{1}{n} \Sigma(X - \bar{X})(Y - \bar{Y})}{\text{Variance of } y}$
$= \frac{\frac{\Sigma XY}{n} - \frac{\Sigma X}{n} \cdot \frac{\Sigma Y}{n}}{9}$
$=$

Date
Page No.

(18) $X = 12$
$Y = ?$

$\bar{X} = 8 \quad \bar{Y} = 15; \quad \sigma_x = 4 \quad \sigma_y = 3.$
$r = 0.99$

Now,

$$b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x}$$
$$= 0.99 \times \frac{3}{4}$$
$$= 0.7425.$$

$$(Y - \bar{Y}) = b_{yx}(X - \bar{X})$$
$$(Y - 15) = 0.7425(12 - 8)$$
$$Y - 15 = 2.97$$
$$Y = 17.97$$

(20) The eqn are :-
$$2X - 3Y = 0$$
$$4Y - 5X = 8.$$

(i) Since the two regression line always passes through $\bar{x}, \bar{y}$.

$$2\bar{X} - 3\bar{Y} = 0$$
$$4Y - 5X = 8.$$
$$2\bar{X} - 3\bar{Y} = 0) \; 5.$$
$$-5\bar{X} + 4\bar{Y} = 8.) \; 2$$
$$10\bar{X} - 15\bar{Y} = 0$$
$$\cancel{-10\bar{X}} + 8\bar{Y} = 16$$
$$-7\bar{Y} = 16$$
$$\bar{Y} = -2.2857.$$

Date
Page No.

$$2\bar{X} - 3\bar{Y} = 0.$$
$$2\bar{X} - 3 \times 2.2857 = 0$$
$$2\bar{X} = -6.8571$$
$$\bar{X} = -3.42855.$$

$$2\bar{x} - 3\bar{y} = 0 \quad \text{—(i)}$$
$$4\bar{y} - 5\bar{x} = 8. \quad \text{—(ii)}$$

(ii) Suppose eqn (i) & (ii) are the regression eqn of Y on X and X on Y respectively,

$$3\bar{y} = 2\bar{x}$$
$$\bar{y} = 0.6\bar{x}$$
$$b_{yx} = 0.6.$$

Also from (ii)
$$4\bar{y} - 5\bar{x} = 8.$$
$$4 \times 0.6 - 5\bar{x} = 8.$$
$$2.4 - 5\bar{x} = 8.$$
$$-5\bar{x} = 5.6.$$
$$b_{xy} = -1.12.$$

Now,
$$b_{yx} \cdot b_{xy} = -1.12 \times 0.6$$
$$= -0.672.$$

Here
$$b_{yx} \cdot b_{xy} < 1.$$

Our assumption is correct.
Hence, It is illustrated that Eqn (i) (ii) are the regression eqn of Y on X and X on Y respectively.

Regression eqn of Y on X =
$$(Y - \bar{Y}) = -1.12(X - \bar{X})$$
$$(Y - 2.2857) = -1.12(X - 3.42855).$$

Regression eqn of X on Y is,
$$(X - \bar{X}) = 0.6(Y - \bar{Y})$$
$$(X + 3.43) = 0.6(Y + 2.29).$$
$$(X + 3.43) = 0.6Y + 1.374.$$

Date
Page No.

(25) The regression eq^n are:-

$$3X + 12Y = 19$$
$$9X + 3Y = 46$$

$$3\bar{x} + 12\bar{y} = 19 \quad \text{--- (i)}$$
$$9\bar{x} + 3\bar{y} = 46 \quad \text{--- (ii)}$$

We suppose the eq^n (i) as eq^n of Y on X and eq^n (ii) as X on Y then,

$$3\bar{x} + 12\bar{y} = 19$$
$$12y = 19 - 3x$$
$$y = 1.59 - 0.25x$$
$$y = -0.25x. \quad b_{yx} = -0.25 = -\frac{1}{4}$$

Also from (ii)

$$9\bar{x} + 3\bar{y} = 46.$$
$$\therefore 9\bar{x} = 46 - 3\bar{y}$$
$$\bar{x} = 5.12 - 0.34\bar{y}$$
$$b_{xy} = -0.34 \; \cancel{8} = -\frac{1}{3}$$

Now,

$$b_{yx} \cdot b_{xy} = -0.25 \times -0.34$$
$$= 0.085$$

$\therefore$ Our assuption is ~~co~~ correct.

~~Hence it is~~ Also,

$$r = \pm\sqrt{b_{yx} \cdot b_{xy}}$$
$$= -\sqrt{0.085}.$$
$$= -0.29.$$

To estimate to value of $\bar{x}$.

$\cancel{3\bar{x} + 12\bar{y} = 19}$ $(3\bar{x} + 12y = 19) \times 3$

$3x$

$$\cancel{9\bar{x}} + 36y = 57$$
$$\cancel{9\bar{x}} + 3\bar{y} = 46$$
$$- \quad - \quad -$$
$$33y = 11$$
$$y = \cancel{11} \frac{1}{3}$$

Date
Page No.

Also, for x,

$$3x + 12y = 19$$
$$3 \times x + 12 \times \frac{1}{3} = 19$$

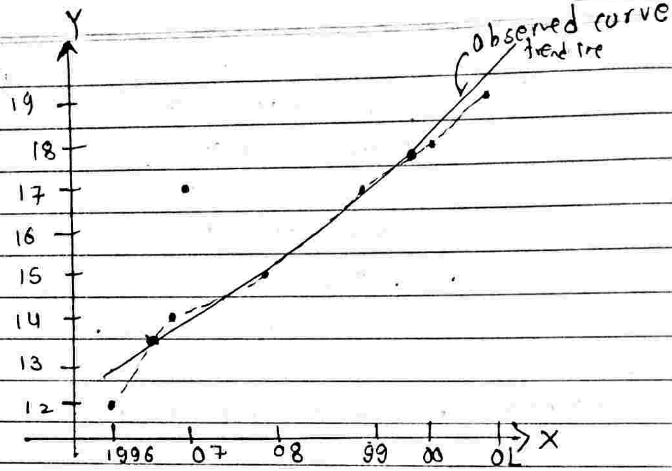
$$3x + 4 = 19$$
$$3x = 15$$
$$\therefore x = 5 \; \cancel{//}.$$

Methods of measuring trend

1. Free hand curve method
2. Semi average method
3. Moving average method
4. Least square method

(2) Semi average method

Year	Method	Semi-average
1996	12	
1997	14	13.67
1998	15	
1999	17	
2000	18	18
2001	19	



(3) Moving average method

Year	Value (Y)	4 yearly moving total	4 yearly moving average	Centered moving average (Y_t).	Short term Fluctuation $(Y-Y_t)$
1992	2				
1993	6				
1994	1	14	3.5	3.625.	-2.625.
		15	3.75		
1995	5			3.875.	1.125
		16	4		
1996	3			4.125	-1.125.
		17	4.25		
1997	7			4.375.	2.625.
		18	4.5		
1998	2			4.625	-2.625
		15	4.75		
1999	6			4.875	1.125.
		12	5		
2000	4			5.125.	-1.125.
		16	5.25		
2001	8				
2002	3				

value = 2+6+1+5 = 14
6+1+5+3 = 15
1+5+3+7 = 17

Moving average = ~~14+15+16+17~~ / ~~4~~
=
$= \frac{2+6+1+5}{4}$
= 3.5
$= \frac{6+1+5+3}{4}$
= 3.75

Moving average $= \frac{3.5+3.75}{2}$
= 3.625
$= \frac{3.75+4}{2}$
= 3.875

Date
Page No.

Mathematical model for time series

Basically there are two types of mathematical model:

1. Additive model for time series.
   Under this model, time series value is given by $Y = T + C + S + R$
   $\therefore Y - T =$ Short term fluctuation.

This model is considered when the four components are independent on each other.

2. Multiplicative model for time series :-
   Under this model, time series value is given by $Y = T \times C \times S \times R$
   Trend elliminated value $= \frac{Y}{T} = C \times S \times R$

   If annual data is given
   $TEV = \frac{Y}{T} = C \times R.$

This model is consider when the four components are interdependent on each other. It is commonly used mathematical model for time series.

(4) Least square method.
Let a trend line be $y = a + bx$ —(i)
where x : time variable
y : value (time series).
The values of a and b are determined by using least square method

Date
Page No.

Using LSM, we get two normal eqns which are :-

$$\Sigma y = na + b\,\Sigma x$$
$$\Sigma xy = a\,\Sigma x + b\,\Sigma x^2$$

(A) if n is odd
$x = t -$ middlemost time.

(B) If n is even.
$x = 2(t -$ A.M of two middle time$)$

(6) .

t	Y	$x = t-1998$	xy	x^2	yt.	TEV.
1995	60	-3	-180	9	60.142	97.674
1996	72	-2	-144	4	66.286	108.62
1997	75	-1	-75	1	71.143	105.43
1998	65	0	0	0	76	85.52
1999	80	1	80	1	80.857	98.936
2000	85	2	170	4	85.714	99.16
2001	95	3	285	9	90.571	104.89
			136	28		

(1) Let a trend line be
$y = a + bx$ —(i)
where $x = t - 1998$.
a & b are determined by using LSM. Using LSM,
$\Sigma y = na + b\,\Sigma x$ —(i)
$\Sigma xy = a\,\Sigma x + b\,\Sigma x^2$ —(ii)
from (i)
$\Sigma y = na + b\,\Sigma x$
$532 = 7a + b \times 0 \quad \therefore a = 76$

Date
Page No.

From (iii) (i) $\Sigma xy = a\Sigma x + b\Sigma x^2$

$136 = a \times 0 + b \times 28$

$\therefore b = \frac{136}{28} = 4.86$

From (i)

$y = 76 + 4.86x$.

Trend values are,

For 1995, $x = -3$; then $y_t = 76 + 4.857 \times -3$
$= 61.429$.

For 1996, $x = -2$, then $y_t = 76 + 4.857 \times -2$
$= 66.286$.

and so on as shown in table.

For 1997 = $y_t = 71.143$.

For 1998 = $y_t = 76$

For 19

(iii) For year 2002.

$x = 2002 - 1998$
$= 4$

$y_t = 76 + 4.857x$
$= 76 + (4.857 \times 4)$
$= 76 + 19.428$
$= 95.428$ lakh.

(iv) Here, $b = 4.857 > 0$ (which indicates that these figure show an increasing trend.

Annual increment in profit = 4.857 lakh

Monthly increment in profit $= \frac{4.857}{12} =$ ~~404~~

$= 0.40475 \times 100$

Rs 40475.

Date
Page No.

Yearly

$\boxed{Y_t = 76 + 4.857x}$ Month $Y_t = \frac{76}{12} + \frac{4.857}{12} \times \frac{x}{12}$

Trend eliminated values are given by,

$TEV = \frac{Y}{Y_t} \times 100$

=

For 1996, $TEV = \frac{60}{61.429} \times 100$
$= 97.674$

For 1996, $TEV = \frac{72}{66.286} \times 100 = 108.62$

and so on in table

For 1997 = $TEV = \frac{75}{71.143} \times 100 = 105.42$

1998 = $\frac{65}{76} \times 100 = 85.52$

$\frac{80}{80.857} \times 100 = 98.936$

(vi) Under the assumption of multiplicative modle.

$Y = T \times C \times S \times R$.

if trend is eliminated.

$\frac{Y}{T} = CSR$.

Since annual data is given,

$\therefore TEV = C \times R$

Only cyclical and random variations are assumed to be left, if trend is elliminated.

Date
Page No.

Methods of measuring seasonal variation

(1) Method of simple average
(Best)(2) Ratio to moving average method (common)
(3) Ratio to trend method
(4) Link relative method.

(1) Method of simple averages.

Year	Spring	Summer	Fall	Winter
1998	8	10	7	5
1989	9	10	7	6
1990	10	11	7	6
1991	10	12	8	7
1992	11	13	9	8
Seasonal total	48	56	38	32
Seasonal average ($\bar{x}_i$)	9.6	11.2	7.6	6.4
S.I	110.34	128.73	87.35	73.56.

Here,
$\bar{x}_1 = 9.6 \quad \bar{x}_2 = 11.2 \quad \bar{x}_3 = 7.6 \quad \bar{x}_4 = 6.4.$

Average of averages is
$$\bar{X} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4}{4}$$
$$= \frac{9.6 + 11.2 + 7.6 + 6.4}{4}$$
$$= 8.7.$$

Seasonal index for spring $= \frac{\bar{x}_1}{\bar{X}} \times 100$
$$= \frac{9.6}{8.7} \times 100$$
$$= 110.34 \#$$

Date
Page No.

Seasonal index for summer $= \frac{\bar{x}_2}{\bar{X}} \times 100$
$$= \frac{11.2}{8.7} \times 100$$
$$= 128.73.$$

Seasonal index for fall $= \frac{\bar{x}_3}{\bar{X}} \times 100$
$$= \frac{7.6}{8.7} \times 100 = 87.35.$$

Seasonal index for winer $= \frac{6.4}{8.7} \times 100 = 73.56.$

(ii) Change*/ His working capital need change between summer and winter is given by,
$$128.73 - 73.56$$
$$= 55.17$$
% change decrease in working capital:
$$= \frac{55.17}{128.73} \times 100$$
$$= 42.857 \#.$$

(18) The sales in the moth of August = Rs. 60,000
And sales in the month of September = Rs 69,000
Also, season incides for August = 105.
S I for September = 140
S.I $\propto$ sales (Worth of sales).
i.e. $\frac{\text{SI for August}}{\text{S.I for September}} = \frac{\text{Sales in August}}{\text{Sales in Sept}}$

or, $\frac{105}{140} = \frac{60,000}{x}$

$\therefore x = 80,000 \#$

Seasonal effect

Note No 2 :- Sum of seasonal indices is equal to 100 times the number of season.

- Semi average method doesn't ensure the elimination of short-term variation.
- Moving average method eliminates short term flutuation.

Probability (Chance of occurance of an event)

Some important terms related to probability.

(A) Experiment :-

Any activity that generate a certain result or a set of results.

(B) Random experiment:-

An experiment which generate results based on chance only

(C) Trial

Performing a random experiment.

(D) Sample space (S)

A set of all possible outcome of a random experiment under the study.

Example 1

(1) If two coins are tossed

$S = \{TT, TH, HT, HH\}$.

$n(S) = 4$

$n(\text{No of } H) = 1$

$n(\text{one } H) = 2$

$n(\text{two } H) = 2$

Example 2

If 3 coins are tossed.

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

$n(S) = 8$

$n(\text{No } H) = 1$

$n(\text{One } H) = 3$

$n(\text{Two } H) = 3$

$n(\text{Three } H) = 1$

Total = 8

Notes

If n coins are tossed

$n(S) = 2^n$

$n(x \text{ heads}) = {}^nC_x$

$P(x \text{ heads}) = \frac{{}^nC_x}{2^n}$

(2) If two dice is rolled.

$S = \{$(1,1) (2,1) (3,1) (4,1) (5,1) (6,1)
(1,2) (2,2) (3,2) (4,2) (5,2) (6,2)
(1,3) (2,3) (3,3) (4,3) (5,3) (6,3)
(1,4) (2,4) (3,4) (4,4) (5,4) (6,4)
(1,5) (2,5) (3,5) (4,5) (5,5) (6,5)
(1,6) (2,6) (3,6) (4,6) (5,6) (6,6).

$n(S) = 36 = 6^2$

(E) Exhaustive events.

All possible results of a random experiment are known as exhaustive event.

Ex:- If a single die is rolled.

A ; Even no $= \{2, 4, 6\}$.

B: odd no = {1, 3, 5}

Here, A and B are two exhaustive event.

In other word a number of events are known to be exhaustive if they all contain all possible results. i.e if

$A \cup B = S$

then A and B are two exhaustive event.

(F) Event

A result or a set of result which is subset of sample space. It is always denoted by capital alphabet letter (A, B, C, D).

It is of two types:-

1. Simple or elementary event
2. Compound or composite event.

Example:- if two coins are tossed.

A = at least one head {TH, HT, HH}

B = exactly two head {HH}.

Here A is a compound event and B is a simple event.

Simple event / Elementary

An event which contains only one sample point and can't be further decomposed into two or more events.

Compound event.

An event which contain more than one sample point and can be further decomposed into two or more event.

Equally likely event

A number of events are known to be equally likely events if their chance of occurance are the same.

In other word, events are known to be equally likely if there is no reason to expect one in preferance to other from available prior information about the events.

Mutually exclusive events

Two events are known to be mutually exclusive if the occurance of an event implies non occurance of another event.

In other word two events A and B are known to be mutually exclusive if

$A \cap B = \emptyset$.

A: A horse A wins the race

B: Another horse B wins the same race.

Here, A and B are known to be mutually exclusive events.

Independent events

Two events are know to be independent if the occurance of an event doesn't affect the occurance or non-occurance of another event. A

Date
Page No.

Ex:- A:- A person hits a target
B:- Another person hits the same target
Here, A and B are two independent events.

Notes:-
1. If two events are mutually exclusive then they cannot be independent.
2. If two events are independent then they cannot be mutually exclusive.

Defination of probability
Defination of probability is broadly classified into two types.
(i) Subjective approach
(ii) Objective approach

(i) Subjective approach
The value of probability is assigned under this approach is completely based on personal belief, opinion or judgment is known as subjective approach of probability. Different person may assign different probability of an event. It is commonly used in management by high level authorities in solving their problems

(ii) Objective approach
Under this approach the value of probability is assigned on the basis of numerial informations (data).

$2(x-0)$

Date
Page No.

(lakh tons)

(7)

Year(T)	Sales(Y)	$x=t-1988$	xy	x^2
1986	100	-2	-200	4
1987	120	-1	-120	1
1988	110	0	0	0
1989	140	1	140	1
1990	80	2	160	4
	$\Sigma Y=550$	$\Sigma x=0$	$\Sigma xy=-20$	$\Sigma x^2=10$

Let a trend line be,
$$y=a+bx \quad \text{--- (i)}$$
where $x=t-1988$

a and b are determined by using least square method. Using LSM
$$\Sigma y = na + b\Sigma x$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad \text{--- (ii)}$$

From (i)
$$\Sigma y = na + b\Sigma x$$
$$550 = 5a + b\times 0$$
$$550 = 5a$$
$$110 = a$$

From (ii) (i) $\Sigma xy = a\Sigma x + b\Sigma x^2$
$$\Sigma xy = 110\times 0 + b\times 10$$
$$-20 = 0 + 10b$$
$$-2 = b$$
$$y = 110 - 2x$$

~~For 1993~~ For 1998 =, $x=0$ then $yt = 110 - 2\times 0$
$= 110$

For year ~~20~~ 1993
$x = 1988 - 1993$
$= 5$

$yt = 110$ ~~EE~~ $- 2\times 5$
$=$

Date

Page No.

(8)

Year(t)	y	$x = t-1982$	xy	x^2
1980	80	-2	-160	4
1981	85	-1	-85	1
1982	87	0	0	0
1983	93	1	93	1
1984	100	2	200	4
	$\Sigma y = 445$	$\Sigma x = 0$	$\Sigma xy = 48$	$\Sigma x^2 = 10$

We know,

Straight line trend eq^n $y = a + bx$.

a & b are determined by using LSM and method

Using LSM. we get two normal eq^n :-

$$\Sigma y = na + b\Sigma x \quad \text{(i)}$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad \text{(ii)}$$

from eq^n (i)

$$\Sigma y = na + b\Sigma x$$

$$445 = 5 \times a + b \times 0$$

$$445 = 5a$$

$$89 = a.$$

from eq^n (ii)

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

$$48 = 89 \times 0 + b \times 10$$

$$48 = 10b$$

$$4.8 = b.$$

Now, from eq^n (i) $y = 89 + 4.8x$

for the year 1986.

$= 1986 - 1982$

$= 4.$

Now

$= 89 + 4.8 \times 4$

$= 108.2$

Date

Page No.

(9)

t	y	$x = t-4$	xy	x^2
1	15	-3	-45	9
2	15	-2	-30	4
3	26	-1	-26	1
4	27	0	0	0
5	33	1	33	1
6	41	2	82	4
7	51	3	153	9
	$\Sigma y = 208$	$\Sigma x = 0$	$\Sigma xy = 167$	$\Sigma x^2 = 28$

we know,

$$\Sigma y = na + b\Sigma x \quad \text{(i)}$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad \text{(ii)}$$

from eq^n (i)

$$\Sigma y = na + b\Sigma x$$

$$208 = 7 \times a + b \times 0.$$

$$208 = 7a$$

$$a = 29.7142$$

from eq^n (ii)

$$\Sigma xy = a\Sigma x + b\Sigma x^2$$

$$167 = 29 \times 0 + b \times 28$$

$$167 = 28b$$

$$5.9642 = b.$$

Now, for eq^n (i)

$$y = 29.714 + 5.9642b$$

from the year of March.

$= 8 - 4$

$= 4.$

$y = 29.714 + 5.964 \times 4$

$= 29.714 + 23.856.$

$= 53.57 \approx 54$

Date
Page No.

(11)

t	y	$x = t - 1999$	xy	x^2
1997	7	-2	-14	4
1998	10	-1	-10	1
1999	12	0	0	0
2001	17	2	34	4
2002	24	3	72	9
	$\Sigma y = 70$	2	$\Sigma xy = 82$	$x^2 = 18$

$\Sigma y = na + b\Sigma x$ —(i)

$\Sigma xy = a\Sigma x + b\Sigma x^2$ —(ii)

From (i)

$\Sigma y = na + b\Sigma x$

$70 = na + b \times 2$

$70 = 5a + 2b$

$\cancel{\frac{70}{2b} = a}$

$\frac{70 - 2b}{5} = a$

From (ii)

$\Sigma xy = a\Sigma x + b\Sigma x^2$

$82 = a \times 2 + b \times 18$

$82 = 2a + 18b$

$82 = \frac{70}{2b} \times 2 + 18b$

$82 = \frac{70}{b} + 18b$

$82 = \frac{70 + 18b^2}{b}$

$82b = 70 + 18b^2$

$18b^2 - 82b +$

$\Sigma xy = a\Sigma x + b\Sigma x^2$

$82 = a \times 2 + 18b$

$82 = \frac{70 - 2b}{5} \times 2 + 18b$

$82 = \frac{140 - 4b + 90b}{5}$

$410 = 140 - 4b + 90b$

$270 = 86b$

$3.13 = b$

$\frac{70 - 2b}{5} = a$

$\frac{70 - 2 \times 3.13}{5} = a \quad \frac{63.7}{5} = a$

Date
Page No.

~~Therefore~~

$y = 12.74 + 3.13x$.

$y = 2003$

$= 2003 - 1999$

$= 4$.

$\therefore y = 12.74 + 3.13 \times 4$

$= 25.3$ ('000' Rs) #.

(13)

Year	I	II	III	IV
1999	78	66	84	80
2000	76	74	82	78
2001	72	68	80	70
2002	74	70	84	74
2003	70	74	86	82
Total	370	352	416	384
Average	74	70.4	83.2	76.8

Average of averages is)

$\bar{x} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4}{4}$

$= \frac{74 + 70.4 + 83.2 + 76.8}{4}$

$= 76.1$.

For, ~~1999~~ $I = \frac{\bar{x}_1}{\bar{x}} \times 100$

$= \frac{74}{76.1} \times 100$

$= 97.24$ #

Date
Page No.

For II = $\frac{x}{\bar{x}} \times 100$

$= \frac{70.4}{76.1} \times 100 = 92.5$

For III = $\frac{83.2}{76.1} \times 100 = 109.32$

For IV = $\frac{76.8}{76.1} \times 100 = 100.91$ #

(16)

Year	Summer	Monsoon	Autumn	Winter
1993	30	82	62	119
1994	33	104	86	171
1995	42	153	99	221
1996	56	172	129	235
1997	67	201	136	302
	228	711	512	1048
	45.6	142.2	102.4	209.6

Average of averages = $\frac{45.6 + 142.2 + 102.4 + 209.6}{4}$

$= 124.95$

Summer = $\frac{45.6}{124.95} \times 100 = 36.49$

Monsoon = $\frac{142.2}{124.95} \times 100 = 113.80$

Autum = $\frac{102.4}{124.95} \times 100 = 81.95$

Winner = $\frac{209.6}{124.95} \times 100 = 167.75$ #

Date
Page No.

(18)

	T	S.I.
August (1)	60,000	105
September (2)	69,000	140.

Increase in sales betn Feb to March

= Rs (69,000 − 60,000)

= Rs 9,000 September

The expected sale in the month of ~~March~~ with respect of August keeping in view of seasonal indices = $\frac{60,000}{105} \times 140$

= 80,000

Thus the owner could have expected a sale of 80,000 in the month of March.

Expected increase in the sale = Rs (80,000 − 60,000)

= 20,000

But actual increase in sale in only given Rs 9,000. which is less than the expectn of sales. That is why the owner of the company wasn't satisfied.

(6)

(2)

add $\frac{4}{}$

$y = na + bx$

$\Sigma y = na + b\Sigma x$

$\therefore a =$

$\Sigma xy = a\Sigma x + b\Sigma x^2$

$\therefore b =$

$\boxed{y = a + bx}$

$y = a + bx$

Probability

Objective approach is also divided into 3 types:-

1. Classical approach
2. Empirical approach
3. Axiomatic approach

Defⁿ of probability

- Sub approach
- Objective approach
  - (III) Axiomatic approach or Modern approach
  - (I) Classical approach or Mathematical approach or prior approach or posterior approach or Relative frequency approach
  - (II) Empirical approach

(I) Classical approach

Let us consider a random experiment generates n finite results which are assumed to be equally likely, mutually exclusive, exclusive. Out of these n, m events are favourable to event A then probability of the event A is defined as the ratio of m to n i.e.

$$P(A) = \frac{m}{n} = \frac{n(A)}{n(S)}$$

56-64 65-70 32 51-55

(II) Empirical approach

Let us consider a random experiment is repeated a large number of times say n times under an identical set of conditions. We next assume that an event A occurs 'm' times out of these 'n'. Then probability of the event A is defined as the limiting value of the ratio of m to n as n tends to infinity

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

(III) Axiomatic approach

Let us consider a sample space (S) define under a random experiment. We next assume an event A such that $A \subseteq S$. then a real valued function $P(A)$ is defined as the probability of A if B satisfies the following ~~property~~ axioms:-

(i) Axiom of non-negativity:

$$P(A) \geq 0 \text{ for all } A \leq S.$$

(ii) Axiom of certainity

$$P(S) = 1.$$

(iii) Axiom of addative

For a sequence of mutually exclusive

$$P(A_1 \cup A_2 \cup A_3 \cup \dots) = P(A_1) + P(A_2) + P(A_3)$$

Date
Page No.

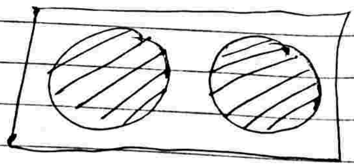
(1) (i) $P(\text{impossible event}) = 0$
(ii) $P(\text{sure event}) = 1$
(iii) $0 \le P(A) \le 1$

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

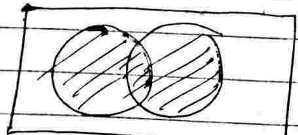
(3) Odds ratio :-
(i) odds in favour of event A is a : b.
$$P(A) = \frac{a}{a+b}$$

(ii) odds against event B is c : d
$$P(B) = \frac{d}{c+d}.$$

Case I
If A and B are two mutually exclusive events then
$$P(A \cup B) = P(A) + P(B).$$

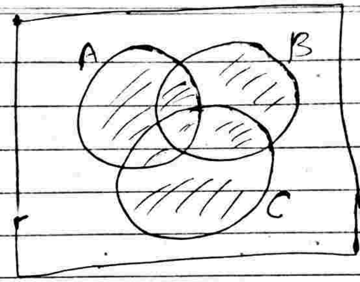


Case II
For any two events A and B
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Notes :-
(1) For any 3 events A, B and C
$$P(A \cup B \cup C) = P(A) + P(B) - P(C) - P(A \cdot B) - P(A \cdot C) - P(B \cdot C) + P(A \cdot B \cdot C).$$

Date
Page No.



Compound / Multiplication theorem of probability.

Case I : If A and B are two independent events then
$$P(A \cap B) = P(A) \times P(B)$$

Case II :
If A and B are two dependent events then
$$P(A \cap B) = P(A) \times P(B/A)$$
or,
$$= P(B) \times P(A|B).$$

where,
$P(B/A)$: Conditional probability of B given that A.
and $P(A/B)$ " " " A given that B.

Conditional probabilities :

$P(A|B)$: Finds the prob of A when event B has already known/occured or given.

Ex 1

Ex 1:

If two king cards are selected at random from 52 to well soffeled card in sucession. without replacement the find $P(K_2/K_1)$

$$P(K_2|K_1) = \frac{3}{51}.$$

Ex 2:- In a class,

	pass	Fails	Total
Boys	8	16	24
Girls	7	1	8
	15	17	32

(i) $P(B \cap F) = \frac{16}{32}$ (iii) $P(B|F) = \frac{16}{17}$

(ii) $P(G \cap P) = \frac{7}{32}$ (iv) $P(P|G) = \frac{7}{8}$

Ex 2:

Two dice are rolled

A : Total is '5'

B : 1st die turns up 3.

Find $P(A|B)$

$P(B/A)$

sol^n:-

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

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

$P(A|B) = \frac{1}{6}$. $\therefore P(B/A) = \frac{1}{4}$

(A) $P(A/B) = \frac{P(A \cap B)}{P(B)}$; provided $P(B) \neq 0$. i.e B isn't impossible even.

(B) $P(B/A) = \frac{P(A \cap B)}{P(A)}$; provided $P(A) \neq 0$ i.e A isn't impossible event.

Notes :-

(1) set Notations of event.

(i) $\bar{A}$: Not occurance of A $P(\bar{A}) = 1 - P(A)$.

(ii) $A \cup B$: Ocurance of A or B.

or

occurance of at least one.

or

" " one or more.

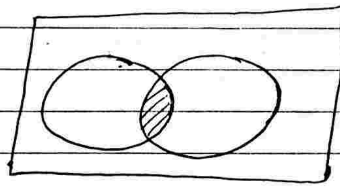
$P(A \cup B) = 1 - P(\overline{A \cup B})$

i.e P(at least one) $= 1 - P$(None)

(iii) $A \cap B$: Occurance of A and B.

or

Occurance of both.

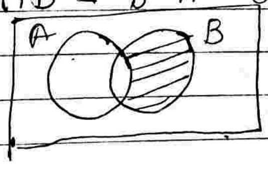


$P(A \cap B) = P(A) + P(B) - P(A \cup B)$.

(iv) $\bar{A} \cap B = B - A$: occurrence of only B

or

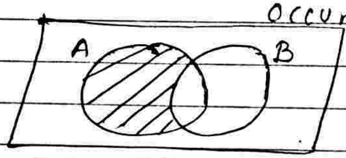
occurance of exactly B.



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

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

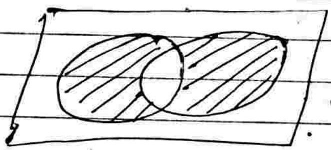
(v) $A\cap\bar{B} = A-B$: Occurance of only A
or
Occurance of exactly B



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

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

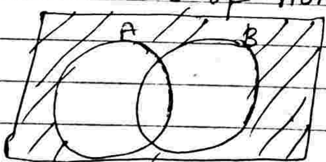
(vi) $(A\cap\bar{B})\cup(\bar{A}\cap B)$: Occurance of exactly one
or Occurance of only one.



Contradiction
↓

$P(\text{exactly one}) = P(A\cap\bar{B}) + P(\bar{A}\cap B)$
$= P(\text{only } A) + P(\text{only } B)$.

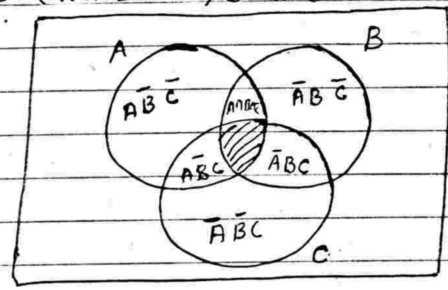
(vii) $\overline{A\cup B}$: Occurance of none or neigther A nor B.



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

(viii) $P(\overline{A\cup B}) = P(\bar{A}\cap\bar{B})$ } De-morgans law.
(ix) $P(\bar{A}\cup\bar{B}) = P(\overline{A\cap B})$

For any 3 events A, B and C



(i) P(Occurance of all) = $P(A\cap B\cap C)$

(ii) P(Occurance of excatly one) = $P(A\bar{B}\bar{C}) + P(\bar{A}B\bar{C}) + P(\bar{A}\bar{B}C)$

(iii) P(Occurance of exactly/only two) = $P(\bar{A}BC) + P(A\bar{B}C) + P(AB\bar{C})$

(iv) P(Occurance of none) = $P(\bar{A}\bar{B}\bar{C}) = P(\overline{A\cup B\cup C})$

(v) P(Occurance of at least one) = $P(A\cup B\cup C) = 1 - P(\bar{A}\bar{B}\bar{C})$

(vi) P(at most one) = $P(A\bar{B}\bar{C}) + P(\bar{A}B\bar{C}) + P(\bar{A}\bar{B}C) + P(\bar{A}\bar{B}\bar{C})$

(vii) P(at least two) = $P(\bar{A}BC) + P(A\bar{B}C) + P(AB\bar{C}) + P(ABC)$.

Note no 3

1. If A, B and C are three exhaustive events then $P(A\cup B\cup C) = 1$.
2. If A, B and C are mutually exclusive events then $P(A\cup B\cup C) = P(A) + P(B) + P(C)$

12 408 / 2.81

(3) If A, B and C are three mutually exclusive and exhaustive events then,

$$P(A) + P(B) + P(C) = 1.$$

(4) If A, B and C are three mutually exclusive, exhaustive and equally likely events.

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

(5) If A and B are two independent event i.e $P(A \cdot B) = P(A) \cdot P(B)$. then

(i) A and $\bar{B}$ are also independent.

i.e $P(A \cdot \bar{B}) = P(A) \times P(\bar{B})$

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

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

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

(ii) $\bar{A}$ and B are also independent.

(iii) $\bar{A}$ and $\bar{B}$ are also indep.

If A, B and C are mutual independent then total number of condition may arise = 4

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

(ii) $P(BC) = P(B) \cdot P(C)$

(iii) $P(AC = P(A) \cdot P(C)$

(iv) $P(ABC) = P(A) \cdot P(B) \cdot P(C)$.

If A and B are two dependent events.

(i) $P(A/B)$, defined if B isn't an impossible event.

(ii) $P(A/\bar{B})$ defined if $\bar{B}$ isn't an impossible event.
i.e B isn't sure event.

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

$$\frac{P(A \cap B)}{P(B)} \le \frac{P(A)}{P(B)}$$

$\hookrightarrow P(A/B) \le \frac{P(A}{P(B)}$ & $P(B/A) \le \frac{P(B)}{P(A)}$

(iv) $P(A/B) = \frac{P(A)}{P(B)} \times P(B/A)$

and $P(B/A) = \frac{P(B)}{P(A)} \times P(A/B)$

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

Note

If A and B are two independent event then,

$P(B/A) = P(B)$

& $P(A/B) = P(A)$.

Ex.

(1) $P(0) + P(1) + P(2) + P(3) + P(4) = 1$.

Here,

$0 \le P \le 1$ & $\sum P = 1$

$\therefore$ Hence these are valid probability assignment.

$P(A) = P(2)$ or $P(1) + P(0)$

$= P(2) + P(1) + P(0)$

$= 0.58$

$P(B) = P(4$ or more$)$

$= P(4)$.

$= 0.12$.

Date
Page No.

(2) (a) → Yes
(b) → Impossible because −ve.
Should be +ve and less than 1.

(3) $P(0) = \frac{8}{50}$, $P(1) = \frac{20}{50}$, $P(2) = \frac{12}{50}$, $P(3) = \frac{6}{50}$

$P(4) = \frac{3}{50}$; $P(5) = \frac{1}{50}$.

(a) $P(A) = P(\text{at least one}) = P(1) + P(2) + P(3) + P(4) + P(5)$

$= 1 - \cancel{P(0)}\ \frac{8}{50}$ $\boxed{1 - P(0)}$

$= \frac{42}{50} = 0.84.$

(b) $P(B) = P(3 \text{ or more}) = P(3) + P(4) + P(5)$

$= \frac{5}{50} + \frac{3}{50} + \frac{1}{50}$

$= \frac{10}{50}$

$= 0.2$

(c) $P(C) = P(\text{exactly } 2) = P(2)$

$= \frac{12}{50}.$

(4) (a) A non-leap year consist of 365 days out of which 52 weeks and 1 day additional.
P(52 Sunday) = sure (1)
P(54 Sunday) = Impossible 0.

Occurrence of P(53 Sunday) in non-leap year is based on one additional day.

Date
Page No.

The additional day may be.
{S, M, T, W, Th, F, Sa}

$P(53 \text{ Sunday}) = \frac{1}{7}.$

(b) A leap year.
In a leap year consists of 366 days out of which 52 weeks and two days additional occurence of 53 Sunday in a leap year on the two two additional days
The additional days may be.

{(S, M) (M, T) (T, W) (W, Th) (Th, F) (F, Sa) (Sa, S)}

$P(53 \text{ Sunday}) = \frac{2}{7}.$

(c) What is the chances that a leap year should have 53 Sunday or 53 Monday or 53 Tuesday?

A leap year consists of ~~364 days~~ 52 weeks and 2 additional day.

{(S, M), (M, T), (T, W) (W, Th) (Th, F) (~~Th, S~~ F, S) (Sa, S)}

$P(53\ \text{'S'}, 53\ \text{'M'}, 53\ \text{'T'}) = \frac{1+1+1+1}{7}$

$= \frac{4}{7}.$

(8) $\frac{1}{7}$

Date
Page No.

8\. $P(A)$ = Chance of attending seminar by Mr. A = 0.6
$P(B)$ = Chance of attending seminar by Mr. B = 0.3
$P(A \cap B)$ = Chance of attending seminar by both M $(A \cap B)$ =

Since A and B are two independent A.

$$P(A \cap B) = P(A) \times P(B) = 0.6 \times 0.3 = 0.18$$

$P(A \cup B)$ = chance that at least one of them attending the seminar

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$= 0.6 + 0.3 - 0.18$$
$$= 0.72 \#.$$

9\. $P(E) = \frac{3}{4}$ $\quad$ $P(E \cap M) = \frac{12}{40}$

$$P(E \cup M) = \frac{5}{6}.$$

$$P(E \cup M) = P(E) + P(M) - P(E \cap M).$$

11\. A bag contains = 7 Red
12 white
4 green
Total 23 balls

Since 3 balls are drawn randomly from the bag containing 23 balls.

$$n(S) = {}^{23}C_3 = \frac{23 \times \cancel{22}^{11} \times \cancel{21}^{7}}{\cancel{3 \times 2} \times 1} = 1771.$$
$$= 23 \times 11 \times 7.$$

and = ×
or = +

Date
Page No.

Let $P(A)$ = Prob that the three balls are all white = ?

$$n(A) = {}^{12}C_3$$
$$= \frac{\cancel{12}^{2} \times 11 \times 10}{\cancel{3 \times 2} \times 1}$$
$$= 2 \times 11 \times 10$$

$$\therefore P(A) = \frac{n(A)}{n(S)} = \frac{2 \times \cancel{11} \times 10}{23 \times \cancel{11} \times 7}$$
$$= \frac{20}{161}.$$

(b) Let $P(B)$ (Prob that 3 balls are one of each colours).

$$n(B) = {}^{7}C_1 \times {}^{12}C_1 \times {}^{4}C_1$$
$$= 7 \times 12 \times 4$$

$$\text{Probability} = \frac{n(B)}{n(S)} = \frac{\cancel{7} \times 12 \times 4}{23 \times 11 \times \cancel{7}} = \frac{48}{253}.$$

12\. A bag contains 7 white
9 black
Total 16 balls.

Since two balls are drawn in succession at random with replacement.

Let $P(W_1)$ = Prob of selecting of white ball in 1st drawn = 7/16
$P(B_1)$ = Prob of selecting a black ball in 1st drawn = 9/16
$P(B_2)$ = Prob of " " white ball in 1st drawn. = 7/16
$P(B_2)$ = Prob of selecting a black ball in 2nd drawn = 9/16.

Date
Page No.

Let $P(D)$ = Prob that one of them is white and the other one is black

$= P(W_1 \cap B_2) + P(B_1 \cap W_2)$

Since events are independent

$= P(W_1) \times P(B_2) + P(B_1) \times P(W_2)$

$= \frac{7}{16} \times \frac{9}{16} + \frac{9}{16} \times \frac{7}{16}$

$= \frac{63}{256} + \frac{63}{256}$

$= \frac{\cancel{126}}{\cancel{256}} \frac{63}{128} = \frac{63}{128}$

(14) One tickets is drawn at random from 20 tickets

$n(S) = {}^{20}C_1 = 20$

Let A : Event that the selected ticket is multiple of 3

B : Event that the selected ticket is multiple of 5

$n(A) = 10$; $\therefore P(A) = \frac{n(A)}{n(S)} = \frac{10}{20}$

(5, 10, 15, 20) $n(B) = 4$; $P(B) = \frac{n(B)}{n(S)} = \frac{4}{20}$

$A \cap B$: Event that the selected ticket is multiple of 2 & 5.

multiple of 10

$n(A \cap B) = 2$

$P(A \cap B) = \frac{n(A \cap B)}{n(S)} = \frac{2}{20}$

(i) $P(A \text{ or } B) = P(A) + P(B) - P(A \cap B)$.

$= \frac{10}{20} + \frac{4}{20} - \frac{2}{20} = \frac{\cancel{12}}{\cancel{20}} \frac{3}{5}$

Date
Page No.

15. Let A : Event that a person who is 50 yrs living till 70 yrs.

Let B : Event that a person who is 60 yrs living till 80 yrs.

Given,

The odds against A is 9:5

and the odds against B is 8:6

$P(A) = \frac{5}{14}$ and $P(B) = \frac{6}{14}$.

Since, even A and B are two independent events

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

$= \frac{5}{14} \times \frac{6}{14} = \frac{30}{14 \times 14}$.

$P(A \cup B)$ = P that at least one of them alive after twenty years.

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

$= \frac{5}{14} + \frac{6}{14} - \frac{30}{14 \times 14}$

$= \frac{5 \times 14 + 6 \times 14 - 30}{14 \times 14}$

$= \frac{124}{196}$

$= 0.632$

(19) Hints :-

P(target being hitted) = P(at least one hit the target)

= $P(A \cup B)$.

(a) $P(A \cap B)$ (b) $P(A \cap \bar{B})$ (c) $P(\bar{A} \cap B)$ (d) $P(A \cup B)$

(e) $P(\bar{A} \cup \bar{B})$ (f) P(only one).

Date

Page No.

(21) A box contains 6 Brown
8 blue
1 black
Total 15 balls

and another box contains 3 brown
7 blue
5 black
Total 15 balls

Since one ball is drawn from each box
Let $P(A_1)$: Probability of selecting a brown ball from box first $= \frac{6}{15}$

$P(B_1)$: Probability of selecting blue ball from I $= \frac{8}{15}$

$P(C_1)$: Probability of selecting black ball from I $= \frac{1}{15}$

$P(A_2)$: Probability of selecting brown ball from box II $= \frac{3}{15}$

$P(B_2)$ " " " blue " " " II $= \frac{7}{15}$

$P(C_2)$ " " " black " " " " $= \frac{5}{15}$.

Let $P(D)$ = Probability of selecting both balls of the same colours from diff. boxes.

$P(D) = P(A_1 \cap A_2) + P(B_1 \cap B_2) + P(C_1 \cap C_2)$

Since the events are independent

$= P(A_1) \cdot P(A_2) + P(B_1) \cdot P(B_2) + P(C_1) \cdot P(C_2)$

$= \frac{6}{15} \cdot \frac{3}{15} + \frac{8}{15} \cdot \frac{7}{15} + \frac{3}{15} \cdot \frac{5}{15}$.

$=$ 0.35 #

Date

Page No.

22. Let A, B and C be the event that a book will be reviewed by 3 independent critics respectively.

Given,

Odds in favour of A is $5:2 : \frac{5}{7}$.

Odds in favour of B is $4:3 : \frac{4}{7}$.

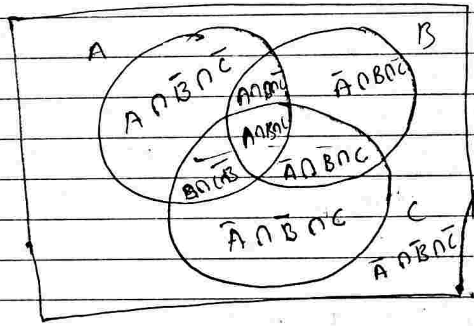
Odds " " of C is $3:4 : \frac{3}{7}$

Odds in against of

$P(\bar{A})$ $2:7$

$P(\bar{B})$ $3:7$

$P(\bar{C})$ $4:7$



Let $P(D)$ = Prob that a majority will be favourable

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

Since events A, B and C are independent

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

$= \frac{5}{2} \cdot \frac{4}{3} \cdot \frac{2}{7} + \frac{5}{2} \times \frac{4}{7} \times \frac{3}{4} + \frac{5}{7} \times \frac{4}{3} \times \frac{3}{4} +$

$\frac{5}{2} \times \frac{4}{3} \times \frac{3}{4}$

$=$

Date
Page No.

$= \frac{5}{7}\times\frac{4}{7}\times\frac{4}{7}+\frac{5}{7}\times\frac{3}{7}\times\frac{3}{7}+\frac{2}{7}\times\frac{4}{7}\times\frac{3}{7}+\frac{5}{7}\times\frac{4}{7}\times\frac{3}{7}$

$= \frac{80}{343}+\frac{45}{343}+\frac{24}{343}+\frac{60}{343}$

$= \frac{80+45+24+60}{343}$

$= \frac{209}{343}$

$= 0.6093.$

23. Probability that a trainee will remain with company
$P(A) = 0.8$
Prob that an employee earn more than 20,000 per year
$P(B) = 0.4$
$P(A \cup B) = P(A \text{ or } B) = 0.9.$
$P(A\cap B) = P(A) + P(B) - P(A \cup B)$
$= 0.8 + 0.4 - 0.9$
$= 0.3.$

$P(B/A) = \frac{P(A\cap B)}{P(A)}$
$= \frac{0.3}{0.8} = \frac{3}{8}$
$= 0.375.$

25 A bag contains = 10 gold coins
8 silver coins
Total 18 coins
Since, Two successive drawings of 3 coins are made. [Question No 3+9]

Date
Page No.

Let A : Event that the first drawing will give 3 gold coins
B : The second drawing will give 3 silver coin.

(i) The coins are replaced before the second drawing
$P(A\cap B) = P(A) \times P(B)$ [$\because$ A & B are two independent event]
$= \frac{^{10}C_3}{^{18}C_3} \times \frac{^{8}C_3}{^{18}C_3}$
$= \frac{\frac{10\times9\times8}{6}}{\frac{18\times17\times16}{6}} \times \frac{\frac{8\times7\times6}{6}}{\frac{18\times17\times16}{6}}$
$= \frac{10\times9\times8\times8\times7\times6}{18\times17\times16\times18\times17\times16}$ $= 272160$
$=$

(ii) The coins are not replaced before the second drawing
$P(A\cup B) = P(A) \times P(B/A)$ [$\because$ A & B are two dependent event]
$= \frac{^{10}C_3}{^{18}C_3} \times \frac{^{8}C_3}{^{15}C_3}.$ $10-3=7$ 8
$=$

Date
Page No.

Price relative:-

It is a ratio of current to base prices of a commodity, expressed in percentage.

$$P = \frac{P_1}{P_0} \times 100$$

Example :

$P_0 = n$

increased by 15%

$P_1 = (n + 15\% \text{ of } n) = 1.15n$

$P = \frac{P_1}{P_0} \times 100 = \frac{1.15n}{n} \times 100$

$\therefore P = 115$

Notes :-

1. Price relative is also an index number of an individual.
2. If the price of a commodity increase by $n\%$ then price relative $(P) = 100 + n$
3. If the price of a commodity decreased by $n\%$ then price relative $(P) = 100 - n$

Construction of price index

Price index
- Simple / unweighted PI
  - (I) Aggregative method
  - II) Average of price relative
    - Using AM
    - Using G.M
- (A) Weighted P.I
  - (iii) Aggregative method
    - Laspeyre's method
    - Paasche's "
    - Fishers "
    - Dorbish-Basley's "
    - Marshell-Edgrowrth
  - Average of price relative
    - Using Am
    - " G.M

Date
Page No.

(I) Simple aggregative method

$$P_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$$

① P_0 : Price in 2001
P_1 : Price in 2002

$P_{01} = \frac{\Sigma P_1}{\Sigma P_0} \times 100$

$= \frac{\cancel{2875}\ 4600}{2875} \times 100$

$= 160.$

(II) Simple average of price relative method.

(A) Using A.M

$$P_{01} = \frac{\Sigma P}{n}$$

(B) Using G.M

or, $P_{01} = \text{Antilog}\left(\frac{\Sigma \log P}{n}\right)$

Q 3

	P_0	P_1	~~Year~~ $P = \frac{P_1}{P_0} \times 100$	Log
A	45	55	122.22	~~4.1743~~ 2.0871
B	60	70	116.6	2.06670
C	20	30	150	2.1760
D	50	75	150	2.1760
E	85	90	105.88	2.0250
F	120	130	108.33	2.03475
			$\Sigma P = 753.1$	$\Sigma p =$ ~~10.3791~~ 12.562

Here,

$\Sigma P = 753.1$; $n = 6$. $= \frac{\Sigma P}{n} = \frac{753.1}{6} = 125.516$

Date
Page No.

For G.M

$$P_{01} = \text{Antilog}\left(\frac{\Sigma \log p}{n}\right)$$
$$= \text{Antilog}\left(\frac{12.56}{6}\right)$$
$$= \text{Antilog}(2.094)$$
$$= 124.21$$

(III) Weighted aggregative method

(A) Laspeyre's method (Base year weighting index)

$$P_{01}(L) = \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100.$$

$$q_{01}(L) = \frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times 100$$

Notes :-

1. For constructing Laspeyre's price index base year quantity q_0 is taken as weight.
2. For constructing Laspeyre's quantity index base year price P_0 is taken as weight.
3. Laspeyre's formula is also termed as base year weighting index number.
4. It is upward biased index number which overestimates the value of index number.

(B) Paasche's Method (Current year weighting index)

$$P_{01}(Pa) = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$$

$$q_{01}(Pa) = \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1} \times 100$$

Date
Page No.

Note

1. For constructing price index, current year quantity Q_1 is taken as weight.
2. For constructing quantity index current year price P_1 is taken as weight.

This method is also known as current year weighting index number. It is downward biased index number which underestimates the value of index number.

(C) Fishers Method

$P_{01}(F)$ = GM of $P_{01}(L)$ and $P_{01}(Pa)$

i.e $P_{01}(F) = \sqrt{P_{01}(L) \times P_{01}(Pa)}$

i.e $P_{01}(F) = \sqrt{\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times \frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}} \times 100$

$$q_{01}(F) = \sqrt{\frac{\Sigma q_1 p_0}{\Sigma q_0 p_0} \times \frac{\Sigma q_1 p_1}{\Sigma q_0 p_1}} \times 100$$

Note :- 1. It is the best formula of index number.
2. It is a geometric mean of Laspeyre's and Paasche's index number.

(D) Dorbish - Bowley's Method

It is an A.M of Laspeyre's and Paasche's index number.

$$P_{01}(DB) = \frac{P_{01}(L) + P_{01}(Pa)}{2}$$

(E) Marshall - Edgeworth method

$$P_{01}(ME) = \frac{\Sigma p_1 \left(\frac{q_0 + q_1}{2}\right)}{\Sigma p_0 \left(\frac{q_0 + q_1}{2}\right)} \times 100$$

$$P_{01}(ME) = \frac{\Sigma p_1 q_0 + \Sigma p_1 q_1}{\Sigma p_0 q_0 + \Sigma p_0 q_1} \times 100$$

Date
Page No.

- For constructing price index A.M of base and current year, quantity $\frac{(q_1+q_0)}{2}$ is taken as weight.
- It is the best alternative after Fisher's formula.

Commodities	p_0	$V_0=p_0q_0$	p_1	$V_1=p_1q_1$	$q_1=\frac{V_1}{p_1}$	$q_0=\frac{V_0}{p_0}$	p_0q_1	p_1q_0
A	2	40	5	75	15	20	30	100
B	4	16	8	40	5	4	20	32
C	1	20	2	24	12	10	12	20
D	5	25	10	60	6	5	30	50
		91					92	202

$\Sigma p_0q_0 = 91$ $\Sigma p_1q_1 = 199$

$\Sigma p_0q_1 = 92$ $\Sigma p_1q_0 = 202$

Fisher's Price index $P_0(F) = \sqrt{\frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times \frac{\Sigma p_1q_1}{\Sigma p_0q_1}} \times 100$

$= \sqrt{\frac{202}{91} \times \frac{199}{92}} \times 100$

$= \sqrt{2.219780 \times 2.16304}$

~~= 322.270~~ = 219.1186

(IV) Weighted average of price method.

(A) Using A.M

$$P_{01} = \frac{\Sigma PW}{\Sigma W}$$

(B) Using G.M

$$P_{01} = \text{Antilog}\left(\frac{\Sigma W \cdot \log P}{\Sigma W}\right)$$

Date
Page No.

where,

W: Weight

W = p_0q_0 / % expenses / expenditure.

Q.no 13

(B)

Item	q_0	p_0	p_1	$W=p_0q_0$	$P=\frac{p_1}{p_0}\times 100$	PW
A	20	200	320	4000	160	640000
B	14	400	420	5600	105	588000
C	15	100	120	1500	120	180000
D	18	40	60	720	150	108000
E	10	20	28	200	140	28000
				$\Sigma W = 12020$		1544000 ΣPW

log (P)
2.204126
2.021194
2.07918
2.176094
2.146132
10.4462

Now, $\left(\frac{\Sigma PW}{\Sigma W}\right)$

$= \frac{1544000}{12020}$

$= 128.45$ #

Test of Consistency of index number

There are 4 tests for test of consistency of index number.

1. Test of consistency

(i) Unit test

This test satisfy if the formula of index number is independent on units in which prices and quantities are quoted.

$\frac{p_1q_0}{p_0q_0}\times 100$ $\frac{p_1q_1}{q_1}$

Note:-

1. All the formula of index number satisfy this test except simple aggregrative formula.

(ii) Time reversal test

This test is stasisfied if the product of two price indices is equal to unity where one price index is current to base whereas another is base to current.

$$P_{01} \times P_{10} = 1$$

In other word price index of current ~~two~~ to base is reciprocal to price index of base to current.

Note :- This test is satisfied by the following formula :-

1. Simple aggregative formula
2. Simple G.M of price relative method
3. Fisher's formula
4. Marsllwei worth method
5. Weighted aggegative formula with fixed weight.

(iii) Factor reversal test

This test satisfies if the product of price and quantity indicies should be equal to the corresponding value index.

$$P_{01} \times q_{01} = V_{01}$$

Note 1:- This test is satisfied by Fisher's formula only.

2 :- Fisher's price index is also known as an ideal index number because it satisfies both time reversal test and factor reversal test.

(iv) Circular test

It is an extented form of time reversal test. According to this test

$$P_{01} \times P_{12} \times P_{20} = 1.$$

This test is satisfied by the following formula

(i) Simple aggregative formula
(ii) Simple G.M of price relative
(iii) Weighted aggregative with fixed weight.

(16) (c)

Items	p_0	q_0	p_1	q_1	p_0q_0	p_1q_1	p_0q_1	p_1q_0
A	5	50	8	40	250	320	200	400
B	7	25	12	30	175	360	210	300
C	9	10	15	25	90	375	225	150
D	12	5	20	18	60	360	216	100
					Σp_0q_0 575	$\Sigma p_1q_1=$ 1415	$\Sigma p_0q_1=$ 851	$\Sigma p_1q_0=$ 950

$\Sigma p_0q_1 = 851$ $\quad$ $\Sigma p_1q_1 = 1415.$

$\Sigma p_1q_0 = 950$ $\quad$ $\Sigma p_0q_0 = 575.$

$$P_{01}(F) = \sqrt{\frac{\Sigma p_1q_0}{\Sigma p_0q_0} \times \frac{\Sigma p_1q_1}{p_0q_1}} \times \cancel{100}$$

$$= \sqrt{\frac{950}{575} \times \frac{1415}{851}} \times \cancel{100}$$

$$= \sqrt{1.6521 \times 1.66274} \; \cancel{\times 100}$$

$$= \cancel{165.7416}.$$

$P_{01} \times 1$

$$P_{10}(F) = \sqrt{\frac{\Sigma p_0q_1}{\Sigma p_1q_1} \times \frac{\Sigma p_0q_0}{p_1q_0}}$$

$$= \sqrt{\frac{851}{1415} \times \frac{575}{950}}$$

$V_{01} = \frac{p_1q_1}{p_0q_0}$

$$q_{01}(F) = \sqrt{\frac{\Sigma q_1p_0}{\Sigma q_0p_0} \times \frac{\Sigma q_1p_1}{\Sigma q_0p_1}}$$

$$= \sqrt{\frac{851}{575} \times \frac{1415}{950}}$$

Date
Page No.

$$V_{01} = \frac{\Sigma p_1 q_1}{\Sigma p_0 q_0} = \frac{1415}{575}.$$

(i) Time reversal test :-

$$P_{01}(F) \times P_{10}(F)$$

$$= \sqrt{\frac{\cancel{950}}{\cancel{575}} \times \frac{\cancel{1415}}{\cancel{851}}} \times \sqrt{\frac{\cancel{851}}{\cancel{1415}} \times \frac{\cancel{575}}{\cancel{950}}}$$

$$= 1.$$

(ii) Fishers reversal test :

$$P_{01}(F) \times q_{01}(F)$$

$$= \sqrt{\frac{\cancel{950}}{575} \times \frac{1415}{\cancel{851}}} \times \sqrt{\frac{\cancel{851}}{\cancel{577}} \times \frac{1415}{\cancel{950}}}$$

$$= \frac{1415}{575}$$

$$= V_{01} \#$$

Base shifting

Example :- If the 1970 index with base 1965 is 200 and 1965 index with base 1960 is 150 then find the index of 1970 with base 1960.

Given,

1970 index with base 1965 = 200

1965 index with base 1960 = 150

1970 index with base 1960 = ?

Let '0' stand for 1960

'1' stand for 1965

'2' stand for 1970

Then, $P_{12} = 200$; $P_{01} = 150$

Date
Page No.

$P_{02} = ?$

$$P_{01} \times P_{12} \times P_{20} = 1$$

$$P_{01} \times P_{12} = \frac{1}{P_{20}}$$

$$P_{02} = P_{01} \times P_{12}.$$

$$= 1.5 \times 2$$

$$= 3.$$

$\therefore P_{01} = 300 \#.$

$$\text{index of 'a' year with base 'b'} = \frac{\text{index of 'a' with base 'c'}}{\text{index of 'b' with base 'c'}} \times 100$$

(18)

Year ~~1990~~	index (Base 1990) ~~100~~	ind $\frac{\text{old index}}{\text{index of new base 1993}} \times 100$
1990	100	$\frac{100}{\cancel{140}\ 160} \times 100 = 62.5$
1991	104	$\frac{104}{160} \times 100 = 65$
1992	115	$\frac{115}{160} \times 100 = 71.875.$
		$\frac{\cancel{160}}{\cancel{160}} \times 100 = 10$
		$\frac{195}{160} \times 100 = 121.875.$
		$\frac{232}{160} \times 100 = 145.$

Date
Page No.

Real wage or Deflated value

Purchasing power of money

Ex $\quad \overset{2060}{P_0 = x} \quad \overset{2070}{P_1 = 2x}$.

Purchasing power of money $= \frac{x}{2x} = \frac{1}{2} = \frac{P_0}{P_1}$

Price relative (P) $= \frac{P_1}{P_0} = \frac{2x}{x} = 2$

↓

Price index.

Real wage = Current wage × P.P of money.

i.e Real wage = Current wage × $\frac{1}{\text{Price Index}} \times 100$

$$R.W = \frac{\text{Current wage}}{\text{Current P.I}} \times 100$$

Notes :-

- Purchasing power of money is the reciprocal of it's price index.
- Real wage $= \frac{\text{Current wage}}{\text{Current P.I}} \times 100$.
- Real wage index $= \frac{\text{Current real wage}}{\text{Base real wage}} \times 100$.
- To maintain the same level of standard of living
  Real wage of ~~current~~ 'a' year = Real wage of 'b' year.
  $\frac{\text{Wage of 'a' year}}{\text{Price index of 'a' year}} \times 100 = \frac{\text{Wages of 'b' year}}{\text{P.I of 'b' year}} \times 100$.

or, $\frac{\text{Wage of 'a' year}}{\text{Wage of 'b' year}} = \frac{\text{PI of 'a' year}}{\text{PI of 'b' year}}$

or, Wage $\propto$ PI.

Date
Page No.

Cost of living index : or Consumer price index or Retail price index

A price index number which is designed to measure an average relative change in price of goods or services utilized by a particular class of people with respect to time or places, is known as cost of living index number. It is also termed as cosumer price index number or retaid price index number.

Methods of constructing ~~m~~ cost of living index.

(A) Aggregative expenditure method.

Cost of living index $= \frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$

which is same as Lasper's price index.

(B) Family budget method.

Cost of living index $= \frac{\Sigma PW}{\Sigma W}$; which is same as weighted A.M of price relative.

$$= \frac{\Sigma IW}{\Sigma W}$$

where I : Group indicies.
W : Weights.
W : $p_0 q_0$ / %-expenses / expenditure.

Date
Page No.

(19)

(21)

Year	Wages	Price index	(i) $RW = \frac{Wage}{P.I} \times 100$	(ii) $R.W\ index = \frac{Current\ RW}{Base\ RW} \times 100$
1980	1800	100	$\frac{1800}{100} \times 100 = 1800$	$\frac{1800}{1800} \times 100 = 100$
1981	2200	170	$\frac{2200}{170} \times 100 = 1294.12$	$\frac{1294.12}{1800} \times 100 = 71.90$
1982	3400	340	$\frac{3400}{340} \times 100 = 1000$	$\frac{1000}{1800} \times 100 = 55.55$
1983	3600	320	$\frac{3600}{320} \times 100 = 1125$	$\frac{1125}{1800} \times 100 = 62.5.$

(23)

Commodities	Weight (w)	% increase in price	P	PW
Food	15	32	132	1980
Clothing	3	54	154	462
Rent	4	47	147	588
Fuel and light	2	78	178	358
Miscellaneous	1	58	158	158
	$\Sigma W = 25$			$\Sigma PW\ 3544$

Cost of living index $= \frac{\Sigma PW}{\Sigma W} = \frac{3544}{25} = 141.76.$

Again, Base period Current period

Income :- Rs 2050 Rs x

PI :- 100 141.76

Let Rs x be, be his earning in current period.
To maintain the same level of standard of living

Date
Page No.

Real income of base year = Real income of current year

$$\frac{\text{Income of base year}}{\text{P.I of base year}} \times 100 = \frac{\text{Income of current year}}{\text{PI of current year}} \times 100$$

$$\frac{2050}{100} = \frac{x}{141.76}$$

$\therefore x = 2906.08$ #.

24. Let x and y be the expenditure on clothing and house rent respectively

Given,

Total earning = Rs 3,500 per month.

Cost of living index = 136.

Group	Expenditure (W)	Group index (I)	IW
Food	1400	180	252000
Clothing	x	150	$150x$
House rent	y	100	$100y$
Fuel & lighting	560	110	61600
Miscellaneous	630	80	50400
	$\Sigma W = 2590 + x + y.$		$\Sigma IW = 364000 + 150x + 100y.$

$$C.O.L.I = \frac{\Sigma PW}{\Sigma W}.$$

Total earning = Total expenditure

$3500 = \Sigma W.$

$\Sigma W = 3500.$

or, $2590 + x + y = 3500$

or, $x + y = 910$

Also,

Date
Page No.

Cost of living index $= \dfrac{\Sigma IW}{\Sigma W}$

$136 = \dfrac{\cancel{3}364000 + 150x + 100y}{3500}$

$476000 - 364000 = 150x + 100y$

$112000 = 150x + 100y$

$2240 = 3x + 2y.$

$2240 = 3x + 2(910 - x)$

$\cancel{2x + 2y = 1820}$

$2240 = 3x + 1820 - 2x$

$420 = x$

and For $y = 910 - x$

$= 910 - 420$

$= 490$

(25) Given

Cost of living index = 200.

Group	% increase in price	Weights (w)	p.	pw
Rent	60	16	160	2560
Clothing	250	12	350	4200
Fuel and light	150	8	250	2000
Miscellaneous	~~150~~ 120	4	220	880 8760
Food	~~120~~ m	60	100 + m	6000 + 60m
		$\Sigma W = 100$		15640 + 60m.

$C.I = \dfrac{15640 + 60m}{100}$

$200 = \dfrac{15640 + 60m}{100}$

$20,000$

$\dfrac{4360}{60} = m$

$72.67 = m$

Date
Page No.

Sampling theory

Some basic terminology used in sampling theory:-

(A) Population / Universe.

The aggregate of all the units under the study is known as popolation / uniserve.

(B) Sample

A part or some part of population, selected with a view to get maximum information about the charactersitics of population, is known as sample.

(C) Census or Complete innumeration.

~~Complete~~ The study of all the units including in the popolation, is known as census.

(D) Survey

The study of all the units included in a sample. is known as survey

(E) Parameter:

Characterstic of the entire population is known as parameter.

In other word a summary value of all the units involved in a population is known as parametre.

Some example of parameter are:-

(1) Population mean

(2) Pop^n standard deviation, pop^n porportion. etc.

Date
Page No.

Statistics

Characteristic of sample obtained from sample observations is known as a statistic.
Example of statistics are :- sample mean, sample S.D, sample porportion e.t.c.

Notation

(A) Pop^n size : N and sample size : N

(B) Pop^n mean : μ Sample mean : $\bar{x}$

$\mu = \frac{\Sigma Y}{N}$ and i.e $\bar{x} = \frac{\Sigma x}{n}$

(C) Pop^n S.D = σ and Sample S.D = s

$\sigma = \sqrt{\frac{\Sigma (Y-\mu)^2}{N}}$. $s = \sqrt{\frac{\Sigma (x-\bar{x})}{n}}$

(D) Pop^n porportion : P Sample purportion : p

$P = \frac{Y}{N}$ $p = \frac{x}{n}$

where Y: No of one of attribute in a pop^n. x = No of units of one of attribute in a sample

Sampling and it's type

A method or a skim which is used to select some part of pop^n to get maximum information about the parametres of the pop^n is know as sampling.
Basically it is of two types :-

(i) Probability sampling / Random sampling
(ii) ~~So~~ Non-probability sampling / non-random sampling
(iii)

Date
Page No.

(i) ~~Simple random~~ Probability sampling
A sampling skim in which sample observation are selected such that each and every member of the population has a fixed or pre asigned chance to belong to the sample, is known as probability sampling.
In this sampling skim, sampling units are selected on the basis of probalistic law.
Some important probability sampling are :-

1. Simple random sampling
2. Stratified sampling
3. Cluster sampling
4. Multistage sampling
5. Systematic sampling

(ii) Non-probability sampling
In this sampling skim sampling units are selected on the basis of personal judgment or interest or opinion of a sampler. It is pure subjective sampling and uses non-probabilistic law. It is commonly used when a sample of very small size has to be choosen from a large population. Some examples of non-probability sampling are Judgemental or purposive sampling.
- Convience sampling.
- Quota sampling

(1) Simple random sampling
A ~~simple random~~ sampling scheme in which each member of pop^n has an equal chance to belong to the sample,

Date
Page No.

is known as simple random sampling. Under this sampling scheme there are two methods of selecting sampling unit.

(i) Simple random sampling with replacement (SRSWR)

Example :- Pop^n unit, {a, b, c}, Sample size 2

Here $N = 3$, $n = 2$.

Possible sample = {(a, b) (a, a) (a, c)
(b, a) (b, b) (b, c)
(c, a) (c, b) (c, c)}

Note :-

$$\text{Possible number of sample} = N^n = 3^2 = 9$$

(ii) Simple random sampling without replacement (SRSWOR)

Example :- Pop^n unit, {a, b, c}, Sample size 2

Here, $N = 3$, $n = 2$

Possible samples = {(a, b), (a, c) (b, c)}

Note :-

$$\text{Possible number of sample} = {}^N C_n$$

Condition

- It is a simple and effective method of sampling under the following condition :-
  - Pop^n isn't very large in size.
  - Pop^n isn't heterogenous in characterstics.

Note :- It is also known as lottery sampling.

(2) Startified sampling

It is an effective method under the following cases :-

Date
Page No.

- Population is very large in size.
- Population is heterogenous in characterstics.
- Some prior information about the population are available.

Under this sampling skim total population is divided into a number of groups (called strata) in such a way that units in a stratum are homogenous in characterstics and samples are selected at random in porportionate term each and every strata.

$N = 1000$

$n = 100$

$N_1 = 150$	$N_2 = 550$	$N_3 = 300$
$\Downarrow$	$\Downarrow$	$\Downarrow$
$n_1 = 15$	n_2	n_3

$n_i \propto N_i$

i.e $\dfrac{n_i}{N_i} = k$

i.e $\dfrac{n_i}{N_i} = \dfrac{n}{N}$

$$\boxed{n_i = \frac{n}{N} \times N_i} \quad \text{(i)}$$

$$n_1 = \frac{n}{N} \times N_i = \frac{100}{1000} \times 150 = 15$$

$$n_2 = \frac{n}{N} \times N_2 = \frac{100}{1000} \times 550 = 55$$

$$n_3 = \frac{n}{N} \times N_3 = \frac{100}{1000} \times 300 = 30$$

Cluster sampling

It is an effective method under the following case.

- Pop^n in large in size.
- Pop^n is heterogenous is str.
- In this sampling scheme total Population is divided into a number of group called cluster in such a way that there is a heterogenity of the units in a cluster and some clusters are selected at random, considered as an ultimate sample.

Multi-stage sampling scheme

It is an extension of cluster sampling.

((Condition same as cluster sampling)

- In this sampling scheme first of all population is supposed to compose of a number of first stage units. (FSU)
- Each FSU is supposed to composed a number of second stage sampling unit and again each second stage sampling unit suppose to compose a number of third stage sampling unit. It is continued upto ultimate sampling unit.

This sampling is carried out at each stages of units ~~sampling~~. ~~FSU~~ FSSU (First stage sampling unit) are selected from pop^n at random. From each and every selected FSSU's a number of second stage sample units are selected at random. And again third stage sampling units are selected from each first stage sampling unit. This process is continued till we reach ultimate sample units.

Notes

- It is cost effective as well as time effective method of sampling as compared to stratified sampling.
- It provides flexibility in the sampling procedure.
- It is likely to be less accurate than that of stratified sampling.

Systematic sampling

It is also know as mixed sampling because in this sampling scheme first sampling unit is selected at random satisfying probabilistic law and the remaining all the sampling units are selected according to predecided pattern.

Steps involved in this sampling scheme

Step 1:- Population units are numbered from 1 to N.

Step 2:- Find uniform sample interval as

$$K = \frac{N}{n} = \frac{1000}{50} = 20$$

Step 3:- First sampling unit is selected at random first K pop^n units.

Let us consider the selected unit is ith pop^n unit i.e 1^{st} sampling unit $= i^{th}$ pop^n unit.

Step 4:- And remaining all the sampling units are selected in the interval of 20 units i.e second sampling unit $= (i+k)^{th}$ pop^n unit.

Date
Page No.

3rd sampling unit = $(i+2k)^{th}$ popn unit
and so on
n^{th} sampling unit = $\{i+(n-1)k\}^{th}$ popn unit.

$\therefore$ 48th sampling unit = $[7+(48-1)20]^{th}$ popn unit
= 947th units

Ex:- N = 100 1st sampling unit is 7th popn unit
n = 50 then,
Uniform sample interval as,
$k = \frac{N}{n} = \frac{1000}{50} = 20$

$\therefore$ 48th sampling unit = $[7+(48-1)20]^{th}$ popn unit
= 947th popn unit.

Sampling distribution and standard error:-

Popn unit = (1, 3, 5).
Sample size = 2
Here; N = 3, n = 2
Popn mean $(\mu) = \frac{\Sigma Y}{N} = \frac{1+3+5}{3} = 3$

$\therefore \mu = 3$
and popn S.D
$$\sigma = \sqrt{\frac{\Sigma(Y-\mu)^2}{N}} = \sqrt{\frac{(1-3)^2+(3-3)^2+(5-3)^2}{3}}$$
$$\therefore \sigma = \sqrt{\frac{8}{3}}$$

Date
Page No.

Sampling distribution and standard error:-

(A) For SRSWR
No of possible samples = $N^n = 3^2 = 9$
Possible samples : {(1,1) (1,3) (1,3) (3,1) (3,3)
(3,5) (5,1) (5,3) (5,5) }.

~~No of possible samples = $N^n = 3^2 = 9 = k$ (Say)~~

~~Possible samples =~~
Sample mean $(\bar{x})$ = 1, 2, 3, 2, 3, 4, 3, 4, 5.
Mean of sample mean is,
$$\bar{\bar{x}} = \frac{\Sigma \bar{x}}{k} = \frac{27}{9} = 3$$
$$\therefore \bar{\bar{x}} = \mu$$

$$S.D(\bar{x}) = \sqrt{\frac{\Sigma(\bar{x}-\bar{\bar{x}})^2}{k}}$$
$$= \sqrt{\frac{4+1+0+1+0+1+0+1+4}{9}}$$
$$= \sqrt{\frac{12}{9}}$$
$$= \sqrt{\frac{4}{3}} = \frac{2}{\sqrt{3}}.$$

$$SD(\bar{x}) = SE(\bar{x}) = \frac{2}{\sqrt{3}}.$$

Here, $\bar{x}$:	1	2	3	4	5	: Total
$P(\bar{x})$:	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{4}{9}$	$\frac{5}{9}$	: 1

is known as sampling distribution $\bar{x}$

Date
Page No.

(13) For SRSWOR

No of possible samples $= {}^N C_n = {}^3C_2 = 3 = k$ (Say)

Possible samples (1,3), (1,5) (3,5)

Sample mean $(\bar{x})$: 2, 3, 4

Mean of sample means is,

$$\bar{\bar{x}} = \frac{\Sigma \bar{x}}{K} = \frac{2+3+4}{3} = 3$$

$$\therefore \bar{\bar{x}} = \mu$$

$$SD(\bar{x}) = \sqrt{\frac{\Sigma(\bar{x}-\bar{\bar{x}})^2}{K}} = \sqrt{\frac{1+0+1}{3}} = \frac{2}{\sqrt{3}}$$

Notes :-

- Mean of the sample mean is always equal to their population mean i.e $\bar{\bar{x}} = \mu$.
- A table consisting of values taken by a stastictic and the corresponding probabilities is known as sampling distribution of that stastictic.
- Standard deviation of a statistic is known as standard error of that statistic.
- Standard error measures sampling fluctuation of a sampling distribution.

$SE(\bar{x}) = \frac{\sigma}{\sqrt{n}}$ [For SRSWR]

$SE(\bar{x}) = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$ [for SRSWOR]

where $\sqrt{\frac{N-n}{N-1}}$ is known as FPC or FPM.

FPC = Finite popn correction

FPM = Finite popn multiplier

Date
Page No.

Standard error is inversly porportional to the square root of sample size.

(6)

(7) $N = 520$

$n = 25$.

Sigma $\sigma = 105$

(8)

ICAN

(16) $P(A) = \frac{1}{3}$ $P(B) = \frac{1}{4}$, $P(C) = \frac{1}{2}$

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

$P(\bar{B}) = 1 - \frac{1}{4} = \frac{3}{4}$

$P(\bar{C}) = 1 - \frac{1}{2} = \frac{1}{2}$

$P(A \cup B \cup C) = 1 - P(\bar{A} \cup \bar{B} \cup \bar{C})$

$= 1 - P(\bar{A}) \cdot P(\bar{B}) \cdot P(\bar{C})$

$= 1 - \frac{2}{3} \times \frac{3}{4} \times \frac{1}{2}$

$\frac{\cancel{6}}{\cancel{24}}\ \frac{1}{4}$

$= 1 - \frac{\cancel{6}}{\cancel{24}}\ 4$

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

$=$

1	2	3
Urn A	Urn B	Urn C
4 Red	5 Red	4 Red
3 Black	4 Black	4 Black
7 balls	9 balls	8 balls
1	1	1

(19)

P(one Red and two Black) $= P(R_1 \cap B_2 \cap B_3) + P(B_1 \cap R_2 \cap B_3) + P(B_1 \cap B_2 \cap R_3)$

$= P(R_1) \times P(B_2) \times P(B_3) + P(B_1) \times P(R_2) \times P(B_3) + P(B_1) \times P(B_2) \times P(R_3)$

$= \left(\frac{4}{7} \times \frac{4}{9} \times \frac{4}{8}\right) + \left(\frac{3}{7} \times \frac{5}{9} \times \frac{4}{8}\right) + \left(\frac{3}{7} \times \frac{4}{9} \times \frac{4}{8}\right)$

$= \frac{64}{504} + \frac{60}{504} + \frac{48}{504}$

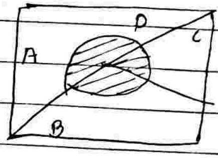
$= \frac{172}{504} = 0.34126$

(22) $P(A) = 0.5$, $P(B) = 0.3$, $P(C) = 0.2$

$P(D/A) = 0.03$

$P(D/B) = 0.04$

$P(D/C) = 0.05$



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

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

$= 0.5 \times 0.03 + 0.3 \times 0.04 + 0.2 \times 0.05$

$= 0.037$

$1 - \frac{4}{6}$ $\frac{6-4}{6}$ $\frac{7}{6}$ $1 - \frac{3}{6}$ $\frac{6-3}{6}$ $\frac{3}{6}$

(25) Winning by A : '6' → '6' → '6' or '5' or '4' → 6, 5, 4, 3, 2

Winning by A : A_1 or $(\bar{A}_1 \cap \bar{B}_2 \cap A_3)$ or $(\bar{A}_1 \cap \bar{B}_2 \cap \bar{A}_3 \cap \bar{B}_4 \cap A_5)$

Winning by B : $(\bar{A}_1 \cap B_2)$ or $(\bar{A}_1 \cap \bar{B}_2 \cap \bar{A}_3 \cap B_4)$ or $(\bar{A}_1 \cap \bar{B}_2 \cap \bar{A}_3 \cap \bar{B}_4 \cap \bar{A}_5 \cap B_6)$

↓ '6' or '5' ↓ 6, 5, 4, 3 ↓ 6, 5, 4, 3, 2, 1

$P(A \text{ wins}) = P(A_1) + P(\bar{A}_1 \cap \bar{B}_2 \cap A_3) + P(\bar{A}_1 \cap \bar{B}_2 \cap \bar{A}_3 \cap \bar{B}_4 \cap A_5)$

$= \frac{1}{6} + \left(\frac{5}{6} \times \frac{4}{6} \times \frac{3}{6}\right) + \left(\frac{5}{6} \times \frac{4}{6} \times \frac{3}{6} \times \frac{2}{6} \times \frac{5}{6}\right)$

$= \frac{1}{6} + \left(\frac{60}{216}\right) + \left(\frac{600}{7776}\right)$

$= \frac{1296 + 2160 + 600}{7776}$

$= \left(\frac{4056}{7776}\right)$ $\frac{~~676~~}{~~1296~~}$ $\frac{~~338~~}{~~648~~}$ $\frac{169}{324}$

$= \frac{169}{324}$

Probability $P(A) + P(B)$

⇒ [HH], [HT] [TH] [TT]

(5) Probability of getting one head (A) $= \frac{1}{4}$

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

- - at least one $= \frac{3}{4}$

Date
Page No.

(6) $n = 6 \times 6 = 36$ Total outcome

Faces in second die

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

Faces in first die

Total of 8 = (2,6) (3,5) (5,3) (4,4) (6,2)

$m = 5$

Total of 8 $= \frac{m}{n} = \frac{5}{36}$

Total of 5 or 7 = (1,4) (1,6) (2,3) (2,5) - - - -

$= \frac{\cancel{10}\ 5}{\cancel{36}\ 18}$

$= \frac{5}{18}$ #

(7) Coin is flipped thrice

$n = \cancel{2 \times 3 = 6}$ $(2)^{3} = 8$ (3 = flipped)

~~[HH] [HT] [H] [T]~~

[HHH] [HHT] [HTH] [THH]

[TTT] ~~[HHH]~~ [TTH] [THT] [HTT]

All head $= \frac{1}{8}$

At least one $= \frac{7}{8}$ #

Date
Page No.

(8) $p(A) = 0.6$

$p(B) = 0.3$

$P(A \cap B) = P(B) \cdot P(B) = 0.18$

$= 0.6 \times 0.3$

$= 0.18$

$P(A \cup B)$ at least one

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

$= 0.6 + 0.3 - 0.18$

$= 0.72$

(10) $p(A) = 0.12$

$p(B) = 0.29$

$p(A \cap B) = 0.07$

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

$= 0.12 + 0.29 - 0.07$

$= 0.34$ #

$p(\overline{A \cup B}) = 1 - 0.34$

$= 0.66$ #

(13) 3 E 6 0

$P(A) = \frac{3}{10}$ $P(B) = \frac{4}{10}$ $P(C) = \frac{2}{10}$ $P(D) = \frac{1}{10}$

(i) ~~One of each kind~~ =

$n = {}^{10}C_4 \Rightarrow 210$

(i) m = no. of favourable ev

$= {}^3C_1 \times {}^4C_1 \times {}^2C_1 \times {}^1C_1 \Rightarrow 24$

$P(A) = m/n \Rightarrow 24/210 \Rightarrow \frac{8}{70} = 4/35$

$\frac{24}{10000}$

Date
Page No.

(ii) has at least one economist

$^3C_1 = 3$

$\Rightarrow {}^3C_3 \times {}^7C_1 + {}^3C_2 \times {}^7C_2 + {}^3C_1 \times {}^7C_3$

$= 1 \times 7 + 3 \times 21 + 3 \times 35$

$= 7 + 63 + 105$

$= 175$

$= \frac{\cancel{175}\ 5}{\cancel{210}\ 6} = \frac{5}{6}$

(iii) has ~~at least~~ the doctor and three others

$1 \times {}^9C_3$

$= 84$

$= 85$

$= \frac{\cancel{84}\ 2}{\cancel{210}\ 5} = \frac{2}{5}$

(14) 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20.

$n = 20$

(i) 2 or 5 → 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 5, 15

$= \frac{12}{20}$

$= 0.6$

(ii) 3 or 5 → 3, 6, 9, 12, 15, 18, 5, 10, 20

$= \frac{9}{20}$

$= 0.45$

Date
Page No.

(15) $n = 56$.

(i) either diamond or king

$\frac{13}{52} \times \frac{4}{52} =$

$m_1 = 13$

$P(A) = 13/52$ $\quad P(A \cap B) = \frac{1}{52}$

$P(B) = 4/52$

$P(A \cup B) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52}$

$= \frac{13 + 4 - 1}{52}$

$= \frac{\cancel{16}\ 4}{\cancel{52}\ 13}$

$= \frac{4}{13}$.

(ii) either black or queen

$P(A) = \frac{26}{52}$ $\quad P(B) = \frac{4}{52}$

$P(A \cap B) = \frac{\cancel{26}}{\cancel{52}\ 2} \times \frac{\cancel{4}\ 2}{\cancel{52}\ 26}\ \frac{1}{} = \frac{1}{26}$.

$P(A \cup B) = \frac{26}{52} + \frac{4}{52} - \frac{1}{26}$.

$= \frac{26 + 4 - 2}{52}$

$= \frac{\cancel{28}\ 4}{\cancel{52}\ 13} = \frac{4}{13}$

$a \to (a:b) \quad \frac{b}{a+b}$

$\frac{a}{a+b} \quad c:d$

$\frac{d}{c+d}$

Odds.

17   9 - 5   8 - 6

Years   50 - 70   60 → 80

A

$P(A) = \frac{5}{13}$

$P(B) = \frac{6}{14}$

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

$= \frac{5}{13} \times \frac{6}{14}$

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

$= \frac{5}{13} + \frac{6}{14} - \frac{30}{182}$

$= \frac{70 + 78 - 30}{182}$

$\Rightarrow \frac{118}{182} \quad 0.648$

(18) 60%

$P(A) = 0.6$

$P(B) = 0.6$

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

$= 0.6 \times 0.6$

$= 0.36$

$P(A \cup B) = 0.6 + 0.6 - 0.36$

$= 0.84$ #.

(11)

$\frac{6}{14} + \frac{14}{30} - \frac{4}{15}$

$\frac{45 + 49 - 28}{105}$

$\frac{66}{105}$

(20) $P(A) = \frac{3}{5}$

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

$P(A) \times P(B) = \frac{3}{5} \times \frac{2}{3} = \frac{2}{5}$

$P(A \cap B)$

At least one ~~Both~~ will be alive $= P(A) + P(B) - P(A \cap B)$

$P(A \cup B) \quad = \frac{3}{5} + \frac{2}{3} - \frac{2}{5}$

$= \frac{9 + 10 - 6}{15}$

$= \frac{13}{15}$

Only the man will be alive $= \frac{3}{5} - \frac{2}{5}$

$= \frac{1}{5}$

Only wife $= \frac{2}{3} - \frac{2}{5}$

$= \frac{10 - 6}{15} = \frac{4}{15}$.

(e) ~~None will be alive $= 1 - \frac{3}{5} - \frac{2}{3}$~~

~~$= \frac{15 - 9 - 10}{15}$~~

$1 - P(A \cup B)$

$= 1 - \frac{13}{15}$

$= \frac{2}{15}$.

(f) only one

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

$= \frac{3}{5} - \frac{2}{5}$

$= \frac{1}{5}$

$P(B) - P(A \cap B)$.

$= \frac{2}{3} - \frac{2}{5}$

$= \frac{10 - 6}{15}$

$= \frac{4}{15}$.

~~$= \frac{4}{60}$~~ $= \frac{1}{5} + \frac{4}{15}$

$= \frac{3 + 4}{15} = \frac{7}{15}$ #

Date

Page No.

(24) $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{3}$ $P(A\cap B)=\frac{1}{4}$

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

$\frac{1}{2}+\frac{1}{3}-\frac{1}{4}$

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

(1) Two dices together

Two number equal (1,1) (2,2) (3,3) (4,4) (5,5) (6,6)

$=\frac{6}{36}=0.166$ #

(2) $king=\frac{4}{52}$ $n={}^{52}C_3\times 50$

$m={}^4C_3$

$\frac{m}{n}=\frac{{}^4C_3}{{}^{52}C_3}$

$=\frac{\frac{4\times3\times2}{\cancel{6}}}{\frac{52\times51\times50}{\cancel{6}}}$

$=\frac{\cancel{24}}{\cancel{132600}}\quad\frac{1}{5525}$

(3)

2 ${}^{9}C_2=\frac{\cancel{9\times8}}{\cancel{2}}$

$n-1C_n$

$10-1C_{10}$

${}^9C_{10}$

(3)

(m) $={}^4C_2\times{}^5C_2$

$=\frac{4\times3}{2}\times\frac{5\times4.2}{2}$

$=\cancel{30}\ 60$

$n={}^9C_4$

$=9\times8\times7\times6$

$\Rightarrow 21\times6$

$=126$

$\therefore P(A)=m/n=60/126\Rightarrow\frac{20}{42}\Rightarrow\frac{10}{21}$

(4) $m={}^4C_1\times{}^{48}C_3$

$n={}^{52}C_4$

$\frac{m}{n}=\frac{{}^4C_1\times{}^{48}C_3}{{}^{52}C_4}$

$P(A)=\frac{25}{100}$ $\frac{24}{100}$

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

$P(A\cap B)=\frac{4}{100}$

(5) $P(A)\ \frac{\cancel{20}}{\cancel{100}}\ 5\times\frac{\cancel{25}}{\cancel{100}}\ 4$ $P(B)$

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

$=\frac{4+5}{20}$

$=\frac{9}{20}$

$\frac{24}{100}\times\frac{20}{100}$

$\frac{1}{5}+\frac{1}{4}-\frac{9}{20}$

$\frac{4+5-9}{20}$

$P(A\cup B)$

$\frac{25}{100}\times\frac{20}{100}$

$\frac{25}{100}+\frac{20}{100}-\frac{1}{4}$

$\frac{25+20-25}{100}$

$\frac{20}{100}$

Date
Page No.

(Q) (20) $\frac{1}{2}, \frac{3}{4}, \frac{1}{4}$

$P(A) = \frac{1}{2} \quad P(B) = \frac{3}{4} \quad P(C) = \frac{1}{4}$

$\Rightarrow \boxed{P(A) \cap P(B) \cap P(C)}$

$= \frac{1}{2} \times \frac{3}{4} \times \frac{1}{4}$

$= \frac{3}{32}$

Now,

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

$\cancel{= \frac{1}{2} + \frac{3}{4} + \frac{1}{4} - \frac{3}{32}}$ $= \frac{2+3+1}{4} \quad \frac{6}{4} \frac{3}{2} \quad \frac{3}{32} \quad \frac{}{32}$

$= \frac{16+24+8-3}{32} = \frac{45}{32}$

$\overline{P(A \cap B \cap C)}$ $1 - \frac{3}{32}$

$= \frac{32-3}{32} = \frac{29}{32}$

(21) $P(A) = 0.8$ remain.

720,000 $P(B) = 0.4 \rightarrow$ employee

~~$P(C \cap B) = 0.9$~~

$P(B \cap C) = 0.9 \rightarrow$ trainee and 20,000

$P(A \cap B) = 0.8 \times 0.4$

$= 0.32$

$P(A) = 0.8$ $0.32 = 0.8 + 0.4$

Date
Page No.

(23) $P(A) = 0.6 \quad = P(\bar{A}) = 0.4$

(*) $P(B) = 0.9 \quad = P(\bar{B}) = 0.1$

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

$P(A \cap B) = 0.6 \times 0.9$

$= 0.54$

$P(A \cup B) = 0.6 + 0.9 - 0.54$

$= 0.96$

~~$P(\overline{A \cup B})$~~

The probability they won't speak truth $= 0.4 \times 0.1$

$= 0.04$

The probability that both won't have same truth opinion $= P(\overline{A \cap B})$

$= 1 - 0.54$

$= 0.46$

The probability that they will contradict $= 0.46 - 0.04$

$= 0.42$ #

(8) Pass in paper I $= 60\%$ $P(A) = 0.6$ Pass in paper I

Paper II $= 50\%$ $P(B) = 0.5$ Pass in paper II

$n(A \cap B) = 0.6 \times 0.5$ (Bo

$= 0.3$

$n(A \cup B) = 0.6 + 0.5 - 0.3$

$= 0.8$

Donot Pass in paper I ~~only~~ $= 1 - 0.6 = 0.4$

Donot Pass in paper II $= 1 - 0.5 = 0.5$

Now, Probability that pass in paper I only $= 0.4 \times 0.6$

$= 0.24$

Again Probability that pass in paper II only $= 0.5 \times 0.5$

$= 0.25$

Now, only one $= 0.24 + 0.25 = 0.49 \approx 0.5$ #

$\frac{8}{15}+\frac{}{7}$

$\frac{280+150+180}{525}$ $\frac{610}{525}$

Date

Page No.

(9) $P(A)$ = Usual function $\frac{1}{4}$

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

(a) both machine

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

$= \frac{1}{12}$

(b) at least one

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

$= \frac{1}{4} + \frac{1}{3} - \frac{1}{12}$

$= \frac{3+4-1}{12} = \frac{\cancel{6}\,1}{\cancel{12}\,2} = \frac{1}{2}$ #

(10) $P(A) = \frac{4}{5}$ $P(A \cap B)$ $P(B \cap C)$ $P(C \cap A)$

$P(B) = \frac{2}{3}$ $\frac{4}{5} \times \frac{2}{3}$ | $\frac{2}{\cancel{3}} \times \frac{\cancel{3}}{7}$ | $\frac{3}{7} \times \frac{4}{5}$.

$P(C) = \frac{3}{7}$ $= \frac{8}{15}$ | $= \frac{2}{7}$ | $= \frac{12}{35}$.

$\frac{8}{15}+\frac{2}{7}+\frac{12}{35}+\frac{8}{35}$ Probability that problem will be solved by all $= P(A \cap B \cap C)$

$= \frac{4}{5} \times \frac{2}{3} \times \frac{3}{7}$

$\frac{20}{35}$ $\frac{\cancel{35}}{105}$

$= \frac{\cancel{24}}{\cancel{105}} \frac{8}{35} = \frac{8}{35}$.

Now, $P(A) + P(B) + P(C) - P(A \cap B \cap C) - P(A \cap B) +$

$= \frac{4}{5} + \frac{2}{3} + \frac{3}{7} - \frac{8}{35}$

$\frac{175}{105}$

$= \frac{84+70+45-24}{105} = \frac{175}{105} - \frac{610}{525}$.

~~$\frac{4}{5}+\frac{2}{3}+\frac{3}{7}-\frac{4}{5}\frac{2}{3}\frac{3}{7}$~~

(11) $P(X) = \frac{6}{14}$ $P(Y) = \frac{14}{30}$

$P(X \cap Y) = \frac{6}{\cancel{14}} \times \frac{\cancel{14}}{30} = \frac{1}{5}$

(a) $P(X \cup Y) = \frac{6}{14} + \frac{14}{30} - \frac{1}{5}$

$= \frac{45+49-21}{105} = \frac{73}{105}$ #.

(b) $P(\bar{X}) = 1 - \frac{6}{14} = \frac{8}{14}$ $P(\bar{Y}) = \frac{\cancel{14}}{\cancel{30}} \cdot 1 - \frac{14}{30} = \frac{16}{30}$

Now, $\frac{8}{14} \times \frac{16}{30} = \frac{\cancel{128}}{\cancel{420}} \frac{32}{195} = \frac{32}{195}$ #.

(12) $P(A) = 0.9$

$P(B) = 0.8$

$P(A \cap B) = 0.9 \times 0.8$

$= 0.72$

at least $P(A) + P(B) - P(A \cap B)$

$= 0.9 + 0.8 - 0.72$

$= 0.98$.

(14) $P(A) = 0.15$ $P(B) = 0.05$ $P(C) = 0.10$

$P(A \cap B \cap C) = 0.15 \times 0.05 \times 0.10$

$= \cancel{0.02} = 0.00075$

$P(A \cup B \cup C) = 0.15 + 0.05 - 0.00075$

$= 0.29925$.

$\frac{875-610}{525}$ $\frac{265}{525}$