Kaustubh S Kabra SE-Comp-I, A, 20

* Tutorial Number - 8 (Statistics)

Q-1. Find A.M for following data.

X	1	2	3	4	5	6
F	2	5	9	11	15	8

→

x_i	f_i	$x_i f_i$
1	2	2
2	5	10
3	9	27
4	11	44
5	15	75
6	8	48

$\Sigma x_i f_i = 2+10+27+44+75+48$
$= 197$

$\Sigma f_i = 2+5+9+11+15+8$
$= 50$

$$\text{Mean} = \bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{197}{50} = \boxed{3.94}$$

Q-2. Find A.M for following data.

X	1	2	3	4	5	6	7	8	9	10
F	20	60	80	100	77	35	15	7	2	1

→

x_i	f_i	$x_i f_i$
1	20	20
2	60	120
3	80	240
4	100	400
5	77	385
6	35	210
7	15	105
8	7	56
9	2	18
10	1	10

$\Sigma x_i f_i = 20+120+240+400+385+210+105+56+18+10$
$= 1564$

$\Sigma f_i = 20+60+80+100+77+35+15+7+2+1$
$= 397$

$$\text{Mean} = \bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{1564}{397} = \boxed{3.9395}$$

Scanned by TapScanner

Q - 3. Find AM for following data.

X	-2	-1	0	1	2	3
F	5	15	25	8	4	3

→

x_i	f_i	$x_i f_i$
-2	5	-10
-1	15	-15
0	25	0
1	8	8
2	4	8
3	3	9

$\Sigma x_i f_i = -10 - 15 + 0 + 8 + 8 + 9$
$= 0$

$\Sigma f_i = 5 + 15 + 25 + 8 + 4 + 3$
$= 60$

Mean $= \bar{X} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \boxed{0}$

Q - 4. Find Mean for the following data.

Marks	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60
No. of Student	2	5	8	7	3

(A=35)

→

Marks (Range)	x_i	f_i	$\frac{x_i - A}{h} = u_i$	$f_i u_i$
10-20	15	2	-2	-4
20-30	25	5	-1	-5
30-40	35	8	0	0
40-50	45	7	1	7
50-60	55	3	2	6

$\Sigma f_i u_i = -4 - 5 + 0 + 7 + 6$
$= 4.$

$\Sigma f_i = 2 + 5 + 8 + 7 + 3$
$= 25$

Mean $= \bar{X} = A + h \times \frac{\Sigma f_i u_i}{\Sigma f_i}$

$= 35 + \cancel{10}^{2} \times \frac{4}{\cancel{25}_{5}}$

$\boxed{\bar{X} = 36.6}$

Scanned by TapScanner

Q-5. Find the mean for the following data.

Class Interval	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	6	5	15	10	5	4	3

→

Class Interval	x_i (Mid value)	f_i	$u_i = \frac{x_i - A}{h}$ (A=22.5) h	$f_i u_i$
5-10	7.5	6	-3	-18
10-15	12.5	5	-2	-10
15-20	17.5	15	-1	-15
20-25	22.5	10	0	0
25-30	27.5	5	1	5
30-35	32.5	4	2	8
35-40	37.5	3	3	9

$\Sigma f_i u_i = -18-10-15+5+8+9$
$= -21$

$\Sigma f_i = 6+5+15+10+5+4+3$
$= 48$

$\text{Mean} = \bar{X} = A + h \times \frac{\Sigma f_i u_i}{\Sigma f_i}$

$= 22.5 + 5 \times \frac{(-21)}{48}$

$= 22.5 - 2.1875$

$\boxed{\bar{X} = 20.3125}$

Q-6. Find the mean for following data.

Wages	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75
Frequency	10	13	18	21	24	28	20	11	8	7

A=52.5

→

Wages	x_i	f_i	$u_i = \frac{x_i - A}{h}$	$f_i u_i$
25-30	27.5	10	-5	-50
30-35	32.5	13	-4	-52
35-40	37.5	18	-3	-54
40-45	42.5	21	-2	-42
45-50	47.5	24	-1	-24
50-55	52.5	28	0	0
55-60	57.5	20	1	20
60-65	62.5	11	2	22
65-70	67.5	8	3	24
70-75	72.5	7	4	28

$\Sigma f_i u_i = -50-52-54-42-24+20+22+24+28$
$= -128$

$\Sigma f_i = 10+13+18+21+24+28+20+11+8+7$
$= 160$

$\text{Mean} = \bar{X} = A + h \times \frac{\Sigma f_i u_i}{\Sigma f_i} = 52.5 + 5 \times \frac{(-128)}{160}$

$\boxed{\bar{X} = 48.5}$

Scanned by TapScanner

Q-7. Find mode for the following data.

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of Student	2	5	8	14	10	6

→

Marks	No. of Student
0-10	2
10-20	5
20-30	8
30-40	14
40-50	10
50-60	6

Modal class → 30-40

$\therefore l = 30$, $f_1 = 14$, $f_2 = 10$

$h = 10$, $f_0 = 8$

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

$$= 30 + 10 \times \frac{(14-8)}{28-8-10} = 30 + \frac{10 \times 6}{10}$$

$$\boxed{\text{Mode} = 36}$$

Q-8. Find the median for the following data.

Income	100	150	80	200	250	180
No. of person	24	26	16	20	6	30

→

x	f	c.f.
80	16	16
100	24	40
150	26	66
180	30	96
200	20	116
250	6	122

$\therefore$ Total frequency $(N) = 122$

$\therefore \frac{N}{2} = 61$,

$\therefore$ The value of c.f. just greater than 61 is 66,

$\therefore$ The corresponding value of x is 150.

$$\boxed{\text{Median} = 150}$$

Q-> Obtain arithmetic mean for the following data.

x	0	1	2	3	4	5	6	7	8	9	10
f	4	5	12	12	13	16	15	13	12	5	6

->

x_i	f_i	$x_i f_i$
0	4	0
1	5	5
2	12	24
3	12	36
4	13	52
5	16	80
6	15	90
7	13	91
8	12	96
9	5	45
10	6	60

$\Sigma x_i f_i = 0+5+24+36+52+80+90+91+96+45+60$
$= 579$

$\Sigma f_i = 113$

$$\text{Mean} = \bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{579}{113} = \boxed{5.1238}$$

Q-> Obtain arithmetic mean for the following data.

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100
No of Student	8	20	14	16	20	25	13	10	5	2

$A = 55$

Marks	x_i	f_i	$u_i = \frac{x_i - A}{h}$	$f_i u_i$
0-10	5	8	-5	-40
10-20	15	20	-4	-80
20-30	25	14	-3	-42
30-40	35	16	-2	-32
40-50	45	20	-1	-20
50-60	55	25	0	0
60-70	65	13	1	13
70-80	75	10	2	20
80-90	85	5	3	15
90-100	95	2	4	8

$\Sigma f_i u_i = -158$

$\Sigma f_i = 113$

$$\text{Mean} = \bar{x} = A + h \times \frac{\Sigma f_i u_i}{\Sigma f_i}$$

$$= 55 + 10 \times \frac{(-158)}{113}$$

$$\boxed{\text{Mean} = 41.0176}$$

Scanned by TapScanner

Q-> Arithmetic mean of four numbers is 16. One item 20 is replaced by 24, what is the arithmetic mean now?

-> Lets 4 numbers be x, y, z & 20.

$\therefore \frac{x+y+z+20}{4} = 16$ ... [Given mean]

New nubers be x, y, z and 24.

$\therefore$ Mean $= \frac{x+y+z+24}{4} = \frac{x+y+z+20}{4} + \frac{4}{4} = 16+1$

$\therefore$ New Mean $= \boxed{17}$

Q-> Obtain median for the following data.

x	1	3	5	7	9	11	13	15	17
f	3	6	8	12	16	16	15	10	5

->

x	f	c.f.
1	3	3
3	6	9
5	8	17
7	12	29
9	16	45
11	16	61
13	15	76
15	10	86
17	5	91

$\therefore$ Total frequency $(N) = 91$

$\therefore \frac{N}{2} = 45.5$

$\therefore$ The value of c.f. just greater than 45.5 is 61,

the corresponding value of x is 11.

$\therefore$ Median is $\boxed{11.}$

Q-> Obtain median for the following data.

Wages in Rupees	10-20	20-30	30-40	40-50	50-60
No. of Labourers	5	8	13	10	8

->

Wages	f	c.f
10-20	5	5
20-30	8	13
30-40	13	26
40-50	10	36
50-60	8	44

Total frequency (N) = 44

$\therefore \frac{N}{2} = 22$

$\therefore$ The value of c.f just greater than 22 is 26.

$\therefore l$ = lower limit of median class = 30.
h = width of classes = 10.
f = frequency of median class = 13.
c = c.f of preceding class of median class = 13

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

$$= 30 + \frac{10}{13}(22 - 13)$$

$$= 30 + 6.9230$$

$$\boxed{\text{Median} = 36.9230}$$

Q-> Find mode of the following data:-

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Student	4	7	8	12	25	18	10

->

Marks	No. of Student
0-10	4
10-20	7
20-30	8
30-40	12
40-50	25
50-60	18
60-70	10

Model class is 40-50

$\therefore l = 40$, $f_0 = 12$, $f_1 = 25$, $f_2 = 18$

$h = 10$,

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

$$= 40 + 10 \times \frac{(25 - 12)}{50 - 12 - 18}$$

$$= 40 + 6.5$$

$$\boxed{\text{Mode} = 46.5}$$

Q-> Calculate Arithmetic mean and mean deviation for following data.

x	1	2	3	4	5	6
f	3	4	8	6	4	2

->

x_i	f_i	$x_i f_i$	$\lvert x_i - A \rvert$	$f_i \lvert x_i - A \rvert$
1	3	3	2.37	7.11
2	4	8	1.37	5.48
3	8	24	0.37	2.96
4	6	24	0.63	3.78
5	4	20	1.63	6.52
6	2	12	2.63	5.26

$$A = \frac{\Sigma x_i f_i}{\Sigma f_i} = \bar{X} = \text{Mean}$$

$$\text{Mean} = \bar{X} = A = \frac{91}{27}$$

$$\boxed{\text{Mean} = A = 3.37}$$

$$\text{Mean Deviation} = \frac{\Sigma f_i \lvert x_i - A \rvert}{\Sigma f_i} = \frac{31.11}{27} = \boxed{1.1523}$$

Q-> Calculate standard deviation for the following data.

Marks	0-10	10-20	20-30	30-40	40-50	50-60
No. of Stud	5	9	15	12	10	3

A = 35

->

x_i	f_i	$u_i = \frac{x_i - A}{h}$	$f_i u_i$	$f_i u_i^2$
5	5	-3	-15	45
15	9	-2	-18	36
25	15	-1	-15	15
35	12	0	0	0
45	10	1	10	10
55	3	2	6	12

$\Sigma f_i = 54$, $\Sigma f_i u_i = -32$, $\Sigma f_i u_i^2 = 118$

Standard Deviation $= \sigma$

$$= h \times \sqrt{\left(\frac{\Sigma f_i u_i^2}{\Sigma f_i}\right) - \left(\frac{\Sigma f_i u_i}{\Sigma f_i}\right)^2}$$

$$= 10 \times \sqrt{\left(\frac{118}{54}\right) - \left(\frac{32}{54}\right)^2}$$

$$= 10 \times 1.3542$$

$$\boxed{S.D = 13.54}$$

Q-> Standard Deviation of four numbers 9, 11, 13, 15

->

x_i	x_i^2
9	81
11	121
13	169
15	225

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

$\bar{X} = \frac{9+11+13+15}{4} = 12.$, $\Sigma x_i^2 = 81 + 121 + 169 + 225 = 596$

$\therefore S.D = \sqrt{\frac{596}{4} - (12)^2} = \boxed{\sqrt{5}}$ (D)

Q-> Standard Deviation of three number 9, 10, 11.

->

x_i	x_i^2
9	81
10	100
11	121.

$\bar{X} = \frac{9+10+11}{2} = 10$, $\Sigma x_i^2 = 81 + 100 + 121 = 302$

$S.D = \sigma = \sqrt{\frac{\Sigma x_i^2}{n} - (\bar{x})^2} = \sqrt{\frac{302}{3} - 100} = \boxed{0.8164}$ (C)

Q-> If $\Sigma x = 235$, $\Sigma x^2 = 6750$, $n = 10$ then standard deviation of distribution

→ $\sigma = \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2}$ $\therefore \sigma = \sqrt{\frac{6750}{10} - (23.5)^2}$

$\bar{x} = \frac{\Sigma x}{n} = \frac{235}{10} = 23.5$

$\sigma = \sqrt{675 - 552.25} = \boxed{11.07}$ (A)

Q-> Coefficient of variation of data - 1, 3, 5, 7, 9 is -

→

x_i	x_i^2
1	1
3	9
5	25
7	49
9	81

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

$\bar{x} = \frac{1+3+5+7+9}{5} = 5$, $\Sigma x_i^2 = 165$

$S.D = \sigma = \sqrt{\frac{165}{5} - (5)^2} = \sqrt{8} = 2.8284.$

Coefficient of variation $= C.V = \frac{\sigma}{\bar{x}} \times 100 = \boxed{56.5685}$ (B)

Q-> The Standard Deviation and AM of the distribution is 12 and 45.5 respectively. Then coefficient of the variance is

→ $C.V = \frac{S.D}{Mean} \times 100 = \frac{\sigma}{\bar{x}} \times 100 = \frac{12}{45.5} \times 100 = \boxed{26.3736}$ (A)

Q-> Which one is more stable :-

Distribution	S.D	A.M	$C.V = \frac{\sigma}{\bar{x}} \times 100$
X	18	5.4	333.33
Y	22.5	4.5	500
Z	24	6.0	400

$\therefore$ More Stable is X (A)

Q-> Which one is more stable:-

→

Distribution	S.D	A.M	$CV = \frac{SD}{AM} \times 100$
X	50	24.43	204.6663
Y	46	25.495	180.4275
Z	40	27	148.1481

∴ $\boxed{\text{Z is more stable (C)}}$

Q-> Which one is more stable:-

→

Distribution	S.D	A.M	$CV = \frac{SD}{AM} \times 100$
X	532	11	4836.36
Y	831	9	9233.33
Z	650	10	6500.00

∴ $\boxed{\text{X is more stable (A)}}$

Q-> Fluctuations in the aggregate marks obtained by two groups of student are given below. Find which of the two shows greater variability.

Group A	A^2	Group B	B^2
518	268324	825	680625
519	269361	830	688900
530	280900	830	688900
530	280900	819	670761
544	295936	814	662596
542	293764	814	662596
518	268324	844	712336
550	302500	842	708864
527	277729	842	708964
527	277729	826	682276
531	281961	832	692224
550	302500	835	697225
550	302500	835	697225
529	279841	840	705600
528	278784	840	705600

$\bar{A} = \frac{\Sigma A}{n} = \frac{7993}{15} = 532.8667$

$\Sigma A^2 = 4556989$

$\sigma_A = \sqrt{\frac{\Sigma A^2}{n} - (\bar{A})^2}$

$= \sqrt{303799.2667 - 283946.92}$

$\sigma_A \doteq 140.8983$

$CV_A = \frac{\sigma_A}{\bar{A}} \times 100 = \boxed{26.44}$

$\bar{B} = \frac{\Sigma B}{n} = \frac{12468}{15} = 831.2$

$\Sigma B^2 = 10364792$

$\sigma_B = \sqrt{\frac{\Sigma B^2}{n} - (\bar{B})^2} = 9.6277$

$CV_B = \frac{\sigma_B}{\bar{B}} \times 100 = \frac{9.6277}{831.2} \times 100 = \boxed{1.1582}$

$\boxed{\text{A shows greater variability as } CV_A > CV_B}$

Scanned by TapScanner

Q→ Calculate first four central moments about the mean for the following data.

x	2.0	2.5	3.0	3.5	4.0	4.5	5.0
f	4	35	60	90	70	40	10

→

x_i	f_i	$x_i f_i$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$	$(x_i-\bar{x})^4$	$f_i(x_i-\bar{x})$	$f_i(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^3$	$f_i(x_i-\bar{x})^4$
2	4	8	-1.56	2.43	-3.79	5.92	-6.24	9.72	-15.16	23.68
2.5	35	87.5	-1.06	1.12	-1.19	1.26	-37.1	39.2	-41.65	44.1
3	60	180	-0.56	0.31	-0.17	0.09	-33.6	18.6	-10.2	5.4
3.5	90	315	-0.06	0.0036	-0.0001	0.00	-5.4	0.32	0.01	0.00
4	70	280	0.44	0.19	0.08	0.03	30.8	13.3	5.6	2.1
4.5	40	180	0.94	0.88	0.83	0.78	37.6	35.2	33.2	31.2
5	10	50	1.44	2.07	2.98	4.29	14.4	20.7	29.8	42.9

$$\bar{X} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{1100.5}{309} = 3.56$$

$$\mu_r = \frac{[\Sigma f_i (x_i - \bar{x})^r]}{\Sigma f_i}$$

$$\mu_1 = \frac{[\Sigma f_i (x_i - \bar{x})]}{\Sigma f_i} = \frac{0.46}{309} \simeq 0$$

$$\mu_2 = \frac{[\Sigma f_i (x_i - \bar{x})^2]}{\Sigma f_i} = \frac{137.04}{309} = 0.44$$

$$\mu_3 = \frac{[\Sigma f (x_i - \bar{x})^3]}{\Sigma f_i} = \frac{10.58}{309} = 0.03$$

$$\mu_4 = \frac{[\Sigma f_i (x_i - \bar{x})^4]}{\Sigma f_i} = \frac{149.38}{309} = 0.48$$

Scanned by TapScanner

Q-> Calculate first four central moments for the following data. Hence find β_1 and β_2.

x	0	1	2	3	4	5	6
f	5	15	17	25	19	14	5

x_i	f_i	$x_i f_i$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$	$(x_i-\bar{x})^4$	$f_i(x_i-\bar{x})$	$f_i(x_i-\bar{x})^2$	$f_i(x_i-\bar{x})^3$	$f_i(x_i-\bar{x})^4$
0	5	0	-3	9	-27	81	-15	45	-135	405
1	15	15	-2	4	-8	16	-30	60	-120	240
2	17	34	-1	1	-1	1	-17	17	-17	17
3	25	75	0	0	0	0	0	0	0	0
4	19	76	1	1	1	1	19	19	19	19
5	14	70	2	4	8	16	28	56	112	224
6	5	30	3	9	27	81	15	45	135	405

$$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = \frac{300}{100} = 3$$

$$\mu_r = \frac{\left[\Sigma f_i (x_i - \bar{x})^r\right]}{\Sigma f_i}$$

$$\mu_1 = \frac{\left[\Sigma f_i (x_i - \bar{x})^1\right]}{\Sigma f_i} = \frac{0}{100} = 0$$

$$\mu_2 = \frac{\left[\Sigma f_i (x_i - \bar{x})^2\right]}{\Sigma f_i} = \frac{242}{100} = 2.42$$

$$\mu_3 = \frac{\left[\Sigma f_i (x_i - \bar{x})^3\right]}{\Sigma f_i} = \frac{-6}{100} = -0.06$$

$$\mu_4 = \frac{\left[\Sigma f_i (x_i - \bar{x})^4\right]}{\Sigma f_i} = \frac{1310}{100} = 13.10$$

Scanned by TapScanner

$$\boxed{\begin{aligned} \mu_1 &= 0 \\ \mu_2 &= 2.42 \\ \mu_3 &= -0.06 \\ \mu_4 &= 13.10 \end{aligned}}$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{3.6 \times 10^{-3}}{5.85} = \boxed{0.00061 = \beta_1}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{13.10}{5.8546} = \boxed{2.2375 = \beta_2}$$

Q→ The first four moments of a distribution about 4 of the variable are -1.5, 17, -30, and 108. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.

→ $\mu_1' = -1.5$, $\mu_2' = 17$, $\mu_3' = -30$, $\mu_4' = 108$, $A = 4$.

We know,

$$\mu_r = \mu_r' - {}^rC_1 \mu_{r-1}' \mu_1' + {}^rC_2 \mu_{r-2}' (\mu_1')^2 - \dots$$

$$\mu_1 = \mu_1' \implies \mu_1 = -1.5$$

$$\mu_2 = \mu_2' - {}^2C_1 \mu_1' . \mu_1' = 17 - 2(-1.5)^2 \implies \mu_2 = 12.5$$

$$\mu_3 = \mu_3' - {}^3C_1 \mu_2' \mu_1' + {}^3C_2 \mu_1' (\mu_1')^2 = -30 - 3(17)(-1.5) + 3(-1.5)^3$$

$$\implies \mu_3 = 36.375$$

$$\mu_4 = \mu_4' - {}^4C_1 \mu_3' \mu_1' + {}^4C_2 \mu_2' (\mu_1')^2 + {}^4C_3 \mu_1' (\mu_1')^3$$

$$= 108 - 4(-30)(-1.5) + 6(17)(-1.5)^2 + 4(-1.5)^4$$

$$\mu_4 = 177.75$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{1323.14}{1953.125} = 0.677$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{177.75}{156.25} = 1.1376 \implies \text{Platykurtosis}$$

Q-> The first four moments of a distribution about 30.2 of the variable are 0.255, 6.222, 30.211, and 400.25. Find central moments and β_1 and β_2. Also comment on skewness and kurtosis of the distribution.

→ $\mu_1' = 0.255$, $\mu_2' = 6.222$, $\mu_3' = 30.211$, $\mu_4' = 400.25$, $A = 30.2$

$\mu_r = \mu_r' - {}^rC_1 \mu_{r-1}' \mu_1' + {}^rC_2 \mu_{r-2}' (\mu_1')^2 - \dots$

$\therefore \mu_1 = \mu_1' = \boxed{0.255 = \mu_1}$

$\mu_2 = \mu_2' - {}^2C_1 \mu_1' \cdot \mu_1' = 6.222 - 2(0.255)^2$

$\boxed{\mu_2 = 6.081}$

$\mu_3 = \mu_3' - {}^3C_1 \mu_2' (\mu_1') + {}^3C_2 \mu_1' (\mu_1')^2$

$= 30.211 - 3(6.222)(0.255) + 3(0.255)^3$

$\boxed{\mu_3 = 25.500}$

$\mu_4 = \mu_4' - {}^4C_1 \mu_3' (\mu_1') + {}^4C_2 \mu_2' (\mu_1')^2 + {}^4C_1 \mu_1' (\mu_1')^3$

$= 400.25 - 4(30.211)(0.255) + 6(6.222)(0.255)^2 + 4(0.255)^4$

$\boxed{\mu_4 = 371.879}$

$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{650.25}{226.977} = 2.877$

$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{371.879}{37.100} = 10.023$ → Leptokurtic

Scanned by TapScanner

Q→ If $\beta_2 = 2.5$, then distribution is

→ If $\beta_2 = 2.5 < 3$ ∴ The distribution is Platykurtosis. (C)

Q→ The first moment of the distribution about the value 5 is 2. AM of distribution is.

→ $\mu_1' = 2$, $A = 5$

We know, $\mu_1' = \bar{X} - A$

$$\bar{X} = \mu_1' + A = 2 + 5 = 7 \Rightarrow \boxed{AM = 7} \quad (B)$$

Q→ Calculate Karl Pearson's coefficient of correlation for the following data.

x	23	28	42	17	26	35	29	37	16	46
y	25	22	38	21	27	39	24	32	18	44

x_i	x_i^2	y_i	y_i^2	$x_i y_i$
23	529	25	625	575
28	784	22	484	618
42	1764	38	1444	1596
17	289	21	441	357
26	676	27	729	702
35	1225	39	1521	1365
29	841	24	576	696
37	1369	32	1024	1184
16	256	18	324	288
46	2116	44	1936	2024

$\Sigma x_i = 299$, $\Sigma y_i = 290$

$\Sigma x_i^2 = 9849$, $\Sigma y_i^2 = 9104$

$\bar{X} = \frac{299}{10} = 29.9$, $\bar{Y} = \frac{290}{10} = 29$

$$r = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{(\Sigma x_i y_i / n) - \bar{x}\bar{y}}{\sqrt{\frac{\Sigma x_i^2}{n} - (\bar{x})^2} \cdot \sqrt{\frac{\Sigma y_i^2}{n}}}$$

$$= \frac{(9405/10) - 867.1}{\sqrt{\frac{9849}{10} - 894.01} \cdot \sqrt{\frac{9104}{10} - 841}}$$

$$= \frac{73.4}{8.53 \times 8.33} = \frac{73.4}{79.08}$$

$$\boxed{r = 0.92}$$

Q-> Find correlation coefficient between x and y, given that

$n=25, \Sigma x = 75, \Sigma y = 100, \Sigma x^2 = 250, \Sigma y^2 = 500, \Sigma xy = 325$

$$r = \frac{cov(x,y)}{\sigma_x \times \sigma_y}$$

$$cov(x,y) = \frac{\Sigma xy}{n} - \bar{X}\bar{Y},$$

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - (\bar{X})^2}, \quad \sigma_y = \sqrt{\frac{\Sigma y^2}{n} - (\bar{Y})^2}$$

$$\bar{X} = \frac{\Sigma x}{n} = \frac{75}{25} = 3, \quad \bar{Y} = \frac{\Sigma y}{n} = \frac{100}{25} = 4$$

$$\therefore r = \frac{(325/25) - 3 \times 4}{\sqrt{(250/25) - (3)^2} \cdot \sqrt{(500/25) - (4)^2}}$$

$$= \frac{13 - 12}{\sqrt{1} \cdot \sqrt{20-16}}$$

$$\boxed{r = \frac{1}{2}}$$

Q-> Find correlation coefficient between x and y, given that

$n = 100, \sigma_x = 10, \sigma_y = 12, \bar{X} = 62, \bar{y} = 53, \Sigma(x-\bar{x})(y-\bar{y}) = 8000$

$$\rightarrow r = \frac{(\Sigma xy/n) - \bar{X}\bar{Y}}{\sigma_x \sigma_y}$$

$$\Sigma(xy) - \Sigma x\bar{y} - \Sigma \bar{x} y + \Sigma \bar{x}\bar{y} = 8000$$

$$\frac{\Sigma(xy)}{n} - \bar{y}\frac{\Sigma x}{n} - \bar{X}\frac{\Sigma y}{n} + \frac{n \times \bar{X}\bar{Y}}{n} = \frac{8000}{n}$$

$$\frac{\Sigma xy}{n\sigma} = \frac{8000}{n} + \bar{x}\bar{y} + \cancel{\bar{x}\bar{y}} - \cancel{\bar{x}\bar{y}}$$

$$r = \frac{\frac{8000}{n} + \cancel{\bar{x}\bar{y}} - \cancel{\bar{x}\bar{y}}}{\sigma_x \times \sigma_y} = \frac{80}{10 \times 12} = \frac{2}{3}$$

$$\boxed{r = \frac{2}{3}}$$

Q→ Find the line regression for the following data.

x	2	3	5	7	9	10	12	15
y	2	5	8	10	12	14	15	16

→

x	x^2	y	y^2	xy
2	4	2	4	4
3	9	5	25	15
5	25	8	64	40
7	49	10	100	70
9	81	12	144	108
10	100	14	196	140
12	144	15	225	180
15	225	16	256	240

$\Sigma x = 63$ $\quad \bar{x} = \frac{63}{8} = 7.875$

$\Sigma y = 82$ $\quad \bar{Y} = \frac{82}{8} = 10.25$

$\Sigma x^2 = 637$, $\Sigma y^2 = 1014$

$\Sigma xy = 797$

$cov(x,y) = \frac{\Sigma xy}{n} - \bar{x}\bar{y} = 18.90$

$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2} = 4.19$, $\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - (\bar{Y})^2} = 4.65$

Line of regression of y on x is.

$$y - \bar{y} = \frac{cov(x,y)}{\sigma_x \cancel{\sigma_y}} \times \frac{\cancel{\sigma_y}}{\sigma_x} \times (x - \bar{x})$$

$$y = \frac{18.90}{(4.19)^2} x + \left(10.25 - \frac{18.90}{(4.19)^2} \times 7.875\right)$$

$$\boxed{y = 1.07x + 1.77} \quad \text{— (1)}$$

Line of regression of x on y is

$$x - \bar{x} = \frac{cov(x,y)}{\cancel{\sigma_x}\,\sigma_y} \times \frac{\cancel{\sigma_x}}{\sigma_y}(y - \bar{y})$$

$$x = \frac{18.90}{(4.65)^2}\,y + \left(7.875 - \frac{18.90}{(4.65)^2} \times (10.25)\right)$$

$$\boxed{x = 0.87y - 1.08} \quad \text{— (2)}$$

i) y when $x = 6$

$$y = 1.07x + 1.77 = 1.07 \times 6 + 1.77$$

$$\boxed{y = 8.19}$$

ii) x when $y = 20$

$$x = 0.87 \times 20 - 1.08$$

$$\boxed{x = 16.32}$$

Q-> By method of least squares, find the straight line that best fits the following data

x	1	2	3	4	5
y	14	27	40	55	68

→ Here, $n = 5$

To fit $y = ax + b$

∴ Normal equation to find a and b are

$$a\Sigma x + nb = \Sigma y \quad \text{— (1)}$$

$a\Sigma x^2 + b\Sigma x = \Sigma xy$ — (2)

x	x^2	y	xy
1	1	14	14
2	4	27	54
3	9	40	120
4	16	55	220
5	25	68	340

$\therefore \Sigma x = 15$, $\Sigma y = 204$, $\Sigma x^2 = 55$, $\Sigma xy = 748$.

$\therefore$ $15a + 5b = 204$ — (3)

$55a + 15b = 748$ — (4)

$a = \frac{68}{5}$, $b = 0$

$\therefore$ $\boxed{5y = 68x}$

Q→ Following is the information given.

	Variable X	Variable Y
A.M.	30	500
S.D	3	100

$r = 0.8$. Find the line of regression

→ Line of regression of y on x is

$$y - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$y = 0.8 \times \frac{100}{3}(x - 30) + \left(500 - 0.8 \times \frac{100}{3} \times 30\right)$$

$$\boxed{y = \frac{80}{3}x - 300}$$

Line of regression of x on y is

$$x - \bar{x} = r \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

$$x = 0.8 \times \frac{3}{100} \times y + 30 - 0.8 \times \frac{3}{100} \times 500$$

$$\boxed{x = \frac{24}{1000} y + 18}$$

Q-> If the two lines of regression are $9x + y - \lambda = 0$ and ~~40x + 18y = 21~~ $4x + y = \mu$ and mean of x and y are 2 and -3. Find λ and μ. Also find r.

→ $y = -9x + \lambda$, $x = -\frac{y}{4} + \frac{\mu}{4}$.

We know mean of x and y satisfy both equations :-

$$\lambda = \bar{y} + 9\bar{x} = -3 + 9 \times 2 = \boxed{15 = \lambda}$$

$$\mu = 4\bar{x} + \bar{y} = 4(2) - 3 = \boxed{5 = \mu}$$

$$b_{yx} = -9 \quad , b_{xy} = \frac{-1}{4}$$

We know,

$$b_{yx} \cdot b_{xy} = r^2$$

$$-9 \cdot \left(\frac{-1}{4}\right) = r^2$$

$$r = \pm \frac{3}{2}$$

We will consider $r = \frac{3}{2}$ as it is coefficient

Q. The regression equation are $8x - 10y + 66 = 0$ and $40x - 18y = 214$. The variance of x is 9, find ① mean of x and y
② correlation coefficient
③ Standard deviation of y.

→ ① We know, both mean of x and y satisfy equation

$\therefore$ $8\bar{x} - 10\bar{y} = -66$
$40\bar{x} - 18\bar{y} = 214$

$\therefore \boxed{\bar{x} = 13 \quad , \quad \bar{y} = 17}$

② $y = \frac{8x}{10} + \frac{66}{10}$ — ①

$x = \frac{18y}{40} + \frac{214}{40}$ — ②

$\therefore b_{yx} = \frac{8}{10}$, $b_{xy} = \frac{18}{40}$

$\therefore$ We know,

$b_{yx} \cdot b_{xy} = r^2$

$\frac{8^4}{10_5} \times \frac{18^9}{40_{10_5}} = \frac{9}{25} = r^2$

$\therefore \boxed{r = \frac{3}{5}}$

③ $b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x}$

$\frac{8^4}{10_{2_1}} = \frac{3^1}{5_1} \cdot \frac{\sigma_y}{3_3}$

$\boxed{12 = \sigma_y}$ $\therefore$ SD of y is $\underline{\underline{12}}$.

Q-> If $\Sigma xy = 2800$, $\bar{x} = 16$, $\bar{y} = 16$, $n = 10$, $\sigma_x^2 = 36$, $\sigma_y^2 = 25$, $r = ?$

→ $r = \dfrac{cov(x,y)}{\sigma_x \cdot \sigma_y} = \dfrac{(\Sigma xy/n) - \bar{x}\bar{y}}{\sigma_x \sigma_y}$

$= \dfrac{280 - 256}{6 \times 5} = \dfrac{24}{30} = \dfrac{4}{5}$

$r = \dfrac{4}{5} = 0.8$ (C)

Q-> $cov(x,y) = 25.8$, $\sigma_x = 6$, $\sigma_y = 5$, $r = ?$

→ $r = \dfrac{cov(x,y)}{\sigma_x \sigma_y} = \dfrac{26.8}{6 \times 5} = 0.86$ (D)

Q-> If the two regression coefficient are 0.16 and 4 then $r = ?$

→ If coeffient are 0.16 and 4.

We know, multiplication of coefficient equals to r^2

$\therefore r = \sqrt{0.16 \times 4} = 0.8$ (C)

Q-> $cov(x,y) = ?$ if $\Sigma xy = 2638$, $\bar{x} = 14$, $\bar{y} = 17$, $n = 10$

→ $cov(x,y) = \dfrac{\Sigma xy}{n} - \bar{X}\bar{Y}$

$= \dfrac{2638}{10} - 14 \times 17$

$= 263.8 - 238$

$cov(x,y) = 25.8$ (B)

Scanned by TapScanner

Q→ $cov(x, y) = ?$ if $\Sigma(x+y)^2 = 10623$, $\Sigma x^2 = 2219$, $\Sigma y^2 = 30$
$\bar{x} = 14.7$, $\bar{y} = 17$, $n = 10$

→ $cov(x, y) = \frac{\Sigma xy}{n} - \bar{x}\bar{y}$

$\Sigma(x+y)^2 = \Sigma x^2 + 2\Sigma xy + \Sigma y^2$
$10623 = 2219 + 2\Sigma xy + 3856$

$\Sigma xy = \frac{5348}{2} = 2674$

$cov(x, y) = \frac{2674}{10} - 14.7 \times 17$

$= 267.4 - 249.8$

$cov(x, y) = \cancel{29.4}\ 17.5$ (B)

Q→ $r(x, y) = ?$ if $\bar{x} = 5.5$, $\bar{y} = 4$, $\Sigma x^2 = 385$, $\Sigma y^2 = 192$
$\Sigma(x+y)^2 = 947$

→ $\Sigma(x+y)^2 = \Sigma x^2 + \Sigma y^2 + 2\Sigma xy$
$947 = 385 + 192 + 2\Sigma xy$
$\Sigma xy = \frac{370}{2}$

$\Sigma xy = 185$

$r = \frac{cov(x, y)}{\sigma_x \sigma_y} = \frac{\frac{\Sigma xy}{n} - \bar{x}\bar{y}}{\sqrt{\left(\frac{\Sigma x^2}{n}\right) - (\bar{x})^2} \cdot \sqrt{\left(\frac{\Sigma y^2}{n}\right) - (\bar{y})^2}}$

$= \frac{18.5 - 22}{2.87 \times 1.78} = -0.681$ (B)

Scanned by TapScanner

Q→ $n = ?$, if $r = 0.022$, $\Sigma xy = 33\,799$, $\sigma_x = 4.5$, $\sigma_y = 64.605$

$\bar{X} = 68$, $\bar{Y} = 62.125$

→ $$r = \frac{\frac{\Sigma xy}{n} - \bar{X}\bar{Y}}{\sigma_x \times \sigma_y}$$

$$\therefore 0.022 \times 4.5 \times 64.605 + 68 \times 62.125 = \frac{33799}{n}$$

$$6.3958 + 4224.5 = \frac{33799}{n}$$

$$n = \frac{33799}{4230.8858}$$

$$n = 7.98$$

$$\boxed{n = 8}$$ (D)

Q→ $\bar{X} = 53.2$, $\bar{Y} = 27.9$, $\sigma_x = 11$, $\sigma_y = 8$, $r = 0.66$,

Value of x when $y = 75$ using line of regression x on y.

$$x - \bar{x} = r \cdot \frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

$$x - 53.2 = 0.66 \times \frac{11}{8}(75 - 27.9)$$

$$x = 85.94$$

$$y - \bar{y} = r \cdot \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$75 - 27.9 = 0.66 \cdot \frac{8}{11}(x - 53.2)$$

$$x = 150.325$$

Scanned by TapScanner

Q-> For a given data, $\bar{x} = 2$, $\bar{y} = -3$, $b_{xy} = -0.11$

If $y = 10$ then $x = ?$

$$x - \bar{x} = b_{xy}(y - \bar{y})$$

$$x - 2 = -0.11(10 + 3)$$

$$\boxed{x = 0.57}$$ (D)

Q-> Line of regression y on x is $8x - 10y + 66 = 0$.
Line of regression x on y is $40x - 18y - 214 = 0$.

Var of $y = \sigma_y^2 = 16$ SD of $x = \sigma_x = ?$

->

$$y = \frac{8x}{10} + \frac{66}{10} \Rightarrow b_{yx} = \frac{8}{10} = r \cdot \frac{\sigma_x}{\sigma_x} \quad -(1)$$

$$x = \frac{18y}{40} + \frac{214}{40} \Rightarrow b_{xy} = \frac{18}{40} = r = \frac{\sigma_x}{\sigma_y} \quad -(2)$$

$$\therefore \frac{b_{yx}}{b_{xy}} = \frac{\sigma_y^2}{\sigma_x^2} = \frac{8}{10} \times \frac{40}{18}$$

$$\sigma_x^2 = \frac{\sigma_y^2 \times 18}{32} = \frac{16 \times 18}{32}$$

$$\sigma_x^2 = 9$$

$$\boxed{\sigma_x = 3}$$ (A)

Q-> The regression lines are $9x+y=15$ and $4x+y=5$. $r=?$

→ $y = -9x + 15$ $\qquad$ $x = \frac{-y}{4} + \frac{5}{4}$

$\Downarrow$ $\qquad$ $\Downarrow$

$b_{yx} = -9$ $\qquad$ $b_{xy} = \frac{-1}{4}$

We know,

$$b_{yx} \cdot b_{xy} = r^2$$

$$-9 \times \left(\frac{-1}{4}\right) = r^2$$

$$r^2 = \frac{9}{4}$$

$$\boxed{r = \frac{3}{2}}$$

Scanned by TapScanner