MODULE - IV

CORRELATION AND REGRESSION

PART - A

(1) we write the values in the following table:

	x	y	xy	x^2	y^2
	10	13	130	100	269
	12	18	216	144	324
	18	12	216	324	144
	24	25	600	476	625
	23	30	690	529	900
	27	10	270	729	100
Σ	144	108	2122	2402	2362

$$r = \frac{Cov(x,y)}{\sigma_x \sigma_y} = \frac{\frac{1}{n}\Sigma xy - \bar{x}\bar{y}}{\sigma_x \sigma_y}$$

$$\bar{x} = \frac{\Sigma x}{n} = \frac{144}{6} = 19$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{108}{6} = 18$$

$$\sigma_x^2 = \frac{\Sigma x^2}{n} - \bar{x}^2$$

$$\Rightarrow \frac{2402}{6} - 361 = 400.33 - 361 = 39.33$$

$$\sigma_y^2 = \frac{\Sigma y^2}{n} - \bar{y}^2 = \frac{2362}{6} - 324 = 393.67 - 324$$

$$= 69.67$$

$$Cov(x,y) = \frac{1}{6}(2122) - (19 \times 18) = 353.67 - 342$$

$$= 11.67$$

$$\therefore r = \frac{11.67}{\sqrt{39.33}\sqrt{69.67}} = 0.223$$

---

② Here $N = 10$

Ranks by A (x)	Ranks by B (y)	Ranks by C (z)	$D_1 = x-y$	$D_2 = x-z$	$D_3 = y-z$	D_1^2	D_2^2	D_3^2
1	3	6	−2	−5	−3	4	25	9
6	5	4	1	2	1	1	4	1
5	8	9	−3	−4	−1	9	16	1
10	4	8	6	2	−4	36	4	16
3	7	1	−4	2	6	16	4	36
2	10	2	−8	0	8	64	0	64
4	2	3	2	1	−1	4	1	1
9	1	10	8	−1	−9	64	1	81
7	6	5	1	2	1	1	4	1
8	9	7	−1	1	2	1	1	4
Total : (Σ)			0	0	0	200	60	214

$$\rho_1(x, y) = 1 - \frac{6\Sigma D_1^2}{N(N^2-1)} = 1 - \frac{6(200)}{10(99)} = -\frac{7}{33}$$

$$\rho_2(x, z) = 1 - \frac{6\Sigma D_2^2}{N(N^2-1)} = 1 - \frac{6(60)}{10(99)} = \frac{7}{11}$$

$$\rho_3(Y,Z) = 1 - \frac{6 \Sigma D_3^2}{N(N^2-1)} = 1 - \frac{6(214)}{10(99)} = \frac{-49}{165}$$

Since $\rho_2(X,Z)$ is maximum, we conclude that the Pair of judges A and C has the nearest approach to common likings in music.

③ Properties of rank correlation Coefficient:

① The value of ρ lies between +1 and -1.

② If $\rho = 1$, there is a complete agreement in the order of the ranks and the direction of the rank is Same.

③ If $\rho = -1$, there is a complete disagreement in the order of the ranks and they are in opposite directions.

X	Y	Rank X (x)	Rank Y (y)	D = x - y	D^2
68	62	4	5	-1	1
64	58	6	7	-1	1
75	68	2.5	3.5	-1	1
50	45	9	10	-1	1
64	81	6	1	5	25
80	60	1	6	-5	25
75	68	2.5	3.5	-1	1
40	48	10	9	1	1
55	50	8	8	0	0
64	20	6	2	4	16

$\Sigma D = 0$; $\Sigma D^2 = 72$

In X-series, 75 occurs two times, the ranks should be alloted are 2 and 3. Therefore, the average of 2 and 3 i.e, 2.5 is given to both.

Next value 68, so rank = 4

64 occurs 3 times. The ranks should be 5, 6 and 7. So rank = $\frac{5+6+7}{3} = 6$

To ΣD^2, we add $\frac{m(m^2-1)}{12}$ for each value repeated, so for 75, m=2, for 64, m=3

So, for X series, correlation is $\frac{2(4-1)}{12} + \frac{3(9-1)}{12}$

$= 5/2$

In Y-series, 68 occurs twice, so rank $= \frac{3+4}{2} = 3.5$

68 occurs twice, so m = 2

So for Y series, correlation is $\frac{2(4-1)}{12} = \frac{1}{2}$

i.e. $\rho = \frac{1 - 6\left(\Sigma D^2 + \frac{5}{2} + \frac{1}{2}\right)}{N(N^2-1)} = \frac{1-6(72+3)}{10 \times 99}$

$= 0.545$

(4) TO prove: The coefficient of correlation lies between -1 and 1.

Proof: Let x and y be deviations of X and Y series from their mean.

Let σ_x and σ_y be their respective standard deviations.

Let $\sum\left(\frac{x}{\sigma_x}+\frac{y}{\sigma_y}\right)^2 = \sum\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}+\frac{2xy}{\sigma_x\sigma_y}\right)$

$$= \frac{\sum x^2}{\sigma_x^2}+\frac{\sum y^2}{\sigma_y^2}+\frac{2\sum xy}{\sigma_x\sigma_y} \quad \text{— (1)}$$

But $\frac{\sum x^2}{\sigma_x^2} = N$. Similarly, $\frac{\sum y^2}{\sigma_y^2} = N$ — (2)

Again $r = \frac{\sum xy}{N\sigma_x\sigma_y} \Rightarrow Nr = \frac{\sum xy}{\sigma_x\sigma_y}$ — (3)

From (1), (2) and (3)

$$\Rightarrow \sum\left[\frac{x}{\sigma_x}+\frac{y}{\sigma_y}\right]^2 = N+N+2Nr$$

$$\Rightarrow 2N(1+r)$$

$$\Rightarrow 1+r \geq 0$$

$$r \geq -1.$$

But $\left(\frac{x}{\sigma_x}+\frac{y}{\sigma_y}\right)^2$ is the sum of squares of real quantities and as such it cannot be negative; at the most it can be zero $\Rightarrow$

$$2N(1+r) \geq 0$$

Hence r cannot be less than -1; at most it can be -1.

Similarly, by expanding $\sum\left(\frac{x}{\sigma_x}-\frac{y}{\sigma_y}\right)^2$ it can be shown that $\sum\left(\frac{x}{\sigma_x}-\frac{y}{\sigma_y}\right)^2 = 2N(1-r)$

This again cannot be negative; at most it can be zero

$$\Rightarrow 2N(1-r) \geq 0$$

$$\Rightarrow 1-r \geq 0$$

$$\Rightarrow 1 \geq r$$

$$\Rightarrow r \leq 1$$

Hence r cannot be greater than $+1$; at the most it can be $+1$.

Hence $-1 \leq r \leq 1$

Notes: $r=1 \rightarrow$ Correlation is Perfect & Positive

$r=-1 \rightarrow$ Correlation is perfect & Negative

$r=0 \rightarrow$ No relationship between the variables.

(5)

Subject A	Subject B	D = A − B	D^2
1	10	−9	81
2	7	−5	25
3	2	1	1
4	6	−2	4
5	4	1	1
6	8	−2	4
7	3	4	16
8	1	7	49
9	11	−2	4
10	15	−5	25
11	9	2	4
12	5	7	49
13	14	−1	1
14	12	2	4
15	13	2	4
		$\Sigma D = 0$	$\Sigma D^2 = 272$

Rank Correlation Coefficient, $\rho = 1 - \frac{6 \Sigma D^2}{N(N^2-1)}$.

$\rho = 1 - \frac{6.(272)}{\underbrace{15(224)}_{0.48}} = 0.52$

$\cancel{\rho = 1 - \frac{6(136)}{16(255)}}$

(6) Let the lines of regression of x on y and y on x are respectively given by

$$x - \bar{x} = r \frac{\sigma x}{\sigma y} (y - \bar{y}) \quad \text{--- (1)}$$

$$y - \bar{y} = r \frac{\sigma y}{\sigma x} (x - \bar{x}) \quad \text{--- (2)}$$

Slope of line (1) $\rightarrow m_1 = \frac{1}{r} \frac{\sigma y}{\sigma x}$

Slope of line (2) $\rightarrow m_2 = r \frac{\sigma y}{\sigma x}$.

Let θ be the angle between two regression lines x on y and y on x. Then.

$$\tan\theta = \frac{m_1 - m_2}{1 + m_1 m_2} = \frac{\frac{1}{r}\frac{\sigma y}{\sigma x} - r\frac{\sigma y}{\sigma x}}{1 + \left(\frac{1}{r}\frac{\sigma y}{\sigma x}\right)\left(r\frac{\sigma y}{\sigma x}\right)}$$

$$\Rightarrow \frac{\frac{\sigma y}{\sigma x}\left(\frac{1}{r} - r\right)}{1 + \frac{\sigma y^2}{\sigma x^2}} = \frac{\sigma y}{\sigma x}\left(\frac{1-r^2}{r}\right)$$

$$\frac{\sigma_x^2}{\sigma_x^2 + \sigma_y^2}$$

$$= \frac{(1-r^2)}{r} \cdot \frac{\sigma x \, \sigma y}{\sigma_x^2 + \sigma_y^2}$$

(2) $\tan\theta = \left(\frac{1-r^2}{r}\right) \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$

$$\theta = \tan^{-1}\left[\frac{1-r^2}{r} \cdot \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}\right]$$

Here, $\sigma_x = \sigma_y = \sigma$

$$\theta = \tan^{-1}\left[\frac{1-r^2}{r} \cdot \frac{\cancel{\sigma^2}}{2\cancel{\sigma^2}}\right] = \tan^{-1}\left(\frac{1-r^2}{2r}\right) \quad -(1)$$

By data; $\theta = \tan^{-1}(3) \quad -(2)$

compare (1) & (2)

$$\frac{1-r^2}{2r} = 3$$

$$1 - r^2 = 6r$$

$$r^2 + 6r - 1 = 0$$

$r_1 = -3 + \sqrt{10}$ and $r_2 = -3 - \sqrt{10}$

$r_1 = 0.16$ and $r_2 = -6.16$

$|r| \leq 1$, $r = -3 + \sqrt{10}$ $\quad (\because -1 \leq r \leq 1)$

$\underline{\underline{r = 0.16}}$

(8) Given, $\sigma_y = 2\sigma_x$

$r = 0.25$

If 'θ' is the angle between two regression lines, then

$$\tan\theta = \left(\frac{1-r^2}{r}\right) \cdot \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$$

$$= \left(\frac{1-(0.25)^2}{0.25}\right) \frac{\sigma_x (2\sigma_x)}{\sigma_x^2 + 4\sigma_x^2}$$

$$\Rightarrow \frac{1-0.0625}{0.25} \cdot \frac{2}{5} = 3.75 \cdot \left(\frac{2}{5}\right) = 1.5$$

(9.) We have to calculate the expected value of Y when X is 120 so, we have to find our regression equation of Y on X.

Mean of X series $(\bar{X}) = 7.6$ | Mean of Y series $(\bar{Y}) = 14.8$

σ of X series $= 3.6$ | σ of Y series $= 2.5$

Coefficient of correlation $(r) = 0.99$

Regression of Y on X is $(Y-\bar{Y}) = r\frac{\sigma_y}{\sigma_x}(X-\bar{X})$

$$\Rightarrow Y - 14.8 = 0.99 \cdot \frac{2.5}{3.6}(X-7.6)$$

$$\cancel{Y-14.8} = 0.688X + 9.57$$

When; $X = 12 \Rightarrow Y = 0.688(12) + 9.57 = 17.826$

Hence, the expected value of Y is 17.83

(10)

x	$X=(x-13)$	X^2	y	$Y=(y-41)$	Y^2	XY
10	-3	9	40	-1	1	3
12	-1	1	38	-3	9	3
13	0	0	43	2	4	0
12	-1	1	45	4	16	-4
16	3	9	37	-4	16	-12
15	2	4	43	2	4	4

Regression equation of Y on X is

$$Y-\bar{Y} = r\frac{\sigma_y}{\sigma_x}(X-\bar{X}) \quad \text{— (1)}$$

$$r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma XY}{\Sigma X^2} = -0.25$$

From (1) $\quad Y-41 = -0.25(X-13)$.

$$Y = -0.25X + 44.25$$

When X is 20 $\quad Y = 39.25$

When the price is Rs. 20, the likely demand is 39.25.

PART - B

(1)

Mathematics Marks (x)	Ranks x	Statistics Marks (y)	Ranks y	D = x - y	D^2
85	2	93	1	1	1
60	4	75	3	1	1
73	3	65	4	-1	1
40	5	50	5	0	0
90	1	80	2	-1	1
-	-	-	-	-	$\Sigma D^2 = 4$

Here $N = 5$

$\Sigma D^2 = 4$

$\therefore$ Rank Correlation, $\rho = 1 - \frac{6 \Sigma D^2}{N(N^2-1)}$

$$\rho = 1 - \frac{6(4)}{5(24)}$$

$$\rho = 1 - \frac{1}{5}$$

$$\rho = 1 - 0.2$$

$$\rho = 0.8$$

(2) In both the series items are in small numbers so there is no need to take deviations.

we use the formula, $r = \frac{cov(x,y)}{\sigma x \cdot \sigma y}$

n = No. of observations = 7

X	Y	X^2	Y^2	XY
12	14	144	196	168
9	8	81	64	72
8	6	64	36	48
10	9	100	81	90
11	11	121	121	121
13	12	169	144	156
7	13	49	169	91
$\Sigma x = 70$	$\Sigma y = 73$	$\Sigma x^2 = 728$	$\Sigma y^2 = 811$	$\Sigma xy = 746$

$$r = \frac{(\Sigma xy \times N) - (\Sigma x \times \Sigma y)}{\sqrt{\Sigma x^2 \times N - (\Sigma x)^2 \cdot (\Sigma y^2 \times N - (\Sigma y)^2)}}$$

$$r = \frac{(746 \times 7) - (70 \times 73)}{\sqrt{(728 \times 7 - (70)^2) \cdot (811 \times 7 - (73)^2)}}$$

0.4288

$r = 0.43$

(3)

R	A	S	D_1 R − A	D_2 R − S	D_3 A − S	D_1^2	D_2^2	D_3^2
1	45	35	−44	−34	10	1936	1156	100
2	70	90	−68	−88	−20	4624	7744	400
3	65	70	−62	−67	−5	3844	4489	25
4	30	40	−26	−36	−10	676	1296	100
5	90	95	−85	−90	−5	7225	8100	25
6	40	40	−36	−34	0	1296	1156	0
7	50	80	−43	−73	−30	1849	5329	900
8	75	80	−67	−72	−5	4489	5184	25
9	85	80	−76	−71	5	5776	5041	25
10	60	50	−50	−40	10	2500	1600	100
			$\Sigma D_1^2 = -557$	$\Sigma D_2^2 = -605$	$\Sigma D_3 - 50$	34215	41095	1675

$$\rho_1(R,A) = 1 - \frac{6\Sigma D_1^2}{N(N^2-1)} = 1 - \frac{6(34215)}{10\times 99} = -\frac{2270}{11}$$

$$\rho_2(R,S) = 1 - \frac{6\Sigma D_2^2}{N(N^2-1)} = 1 - \frac{6(41095)}{10\times 99} = -\frac{8186}{33}$$

$$\rho_3(A,S) = 1 - \frac{6\Sigma D_3^2}{N(N^2-1)} = 1 - \frac{6(1675)}{10\times 99} = -\frac{302}{33}$$

$\Rightarrow$

(4)

Wages (x)	$X = x - \bar{x}$, $\bar{x} = 99$	X^2	Cost of living (y)	$Y = y - \bar{y}$, $\bar{y} = 95$	Y^2	XY
100	1	1	98	3	9	3
101	2	4	99	4	16	8
102	3	9	99	4	16	12
102	3	9	97	2	4	6
100	1	1	95	0	0	0
99	0	0	92	-3	9	0
97	-2	4	95	0	0	0
98	-1	1	94	-1	1	1
96	-3	9	90	-5	25	15
Σx 895	$\Sigma X = 4$	ΣX^2 38	Σy 859	$\Sigma Y = 4$	ΣY^2 80	ΣXY 45

$$r = \frac{\Sigma XY}{\sqrt{\Sigma X^2 \cdot \Sigma Y^2}} = \frac{45}{\sqrt{38 \times 80}}$$

$$r = +0.816$$

(5) Here $N = 8$; $\bar{x} = 29$ and $\bar{y} = 119$.

Fertiliser used x			Productivity y			
Fertiliser used x	$X = x - \bar{x}$ $= x - 29$	X^2	Productivity y	$Y = y - \bar{y}$ $= y - 119$	Y^2	XY
15	-14	196	85	-34	1156	476
18	-11	121	93	-26	676	286
20	-9	81	95	-24	576	216
24	-5	25	105	-14	196	70
30	1	1	120	1	1	1
35	6	36	130	11	121	66
40	11	121	150	31	961	341
50	21	441	160	41	1681	861
	$\Sigma X = 0$	$\Sigma X^2 = 1022$		$\Sigma Y = -14$	$\Sigma Y^2 = 5368$	$\Sigma XY = 2317$

$$r = \frac{\Sigma XY (N) - (\Sigma X)(\Sigma Y)}{\sqrt{(\Sigma X)^2 \times N - (\Sigma X)^2 \times (\Sigma Y)^2 \times N + (\Sigma Y)^2}}$$

$$r = \frac{2317}{\sqrt{1022 \times 5343.5}} = \frac{2317}{2336.89} = 0.99$$

(6) In order to make the data comparable it is necessary to find out the number of blinds out of a fixed number. we have to find out the number of blind persons corresponding to one lakh, in each age group.

The 1st figure : $\frac{55}{100000} \times 100000 = 55$

The 2nd figure : $\frac{40}{600000} \times 100000 = 67$

The 3rd figure : $\frac{40}{400000} \times 100000 = 100$ ... so on

We can get the other values continuing in the same number.

Age x	Mid value m	$X = \frac{m-45}{10}$	X^2	Blind Per lakh y	Deviation from assumed mean (180) Y	Y^2	XY
0-10	5	-4	16	55	-125	15625	500
10-20	15	-3	9	67	-113	12769	339
20-30	25	-2	4	100	-80	6400	160
30-40	35	-1	1	111	-69	4761	69
40-50	45	0	0	150	-30	900	0
50-60	55	1	1	200	20	400	20
60-70	65	2	4	300	120	14400	240
70-80	75	3	9	500	320	102400	960
		$\Sigma X - 4$	ΣX^2 44		ΣY 43	$\Sigma Y^2 = 157655$	$\Sigma XY = 2288$

∴ Coefficient of correlation,

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\left\{\Sigma X^2 - \frac{(\Sigma X)^2}{N}\right\}\left\{\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}\right\}}}$$

$$r = \frac{2288 - \left(\frac{-4 \times 43}{8}\right)}{\sqrt{\left\{44 - \frac{(-4)^2}{8}\right\}\left\{157655 - \frac{(43)^2}{8}\right\}}}$$

$$r = \frac{2309.5}{2571.34} = 0.898$$

Ranks in statistics (x)	Ranks in Mathematics (y)	$D = (x - y)$	D^2
1	2	-1	1
2	4	-2	4
3	1	+2	4
4	5	-1	1
5	3	+2	4
6	9	-3	9
7	7	0	0
8	10	-2	4
9	6	+3	9
10	8	+2	4

$\boxed{\Sigma D^2 = 40}$

$$\rho = 1 - \frac{6 \Sigma D^2}{N(N^2-1)}$$

$$\rho = 1 - \frac{6(40)}{10 \times 99}$$

$$\rho = 0.76$$

(8)

Ranks in maths (x)	Ranks in statistics (y)	$D = x - y$	D^2
1	1	0	0
2	10	-8	64
3	3	0	0
4	4	0	0
5	5	0	0
6	7	-1	1
7	2	5	25
8	6	2	4
9	8	1	1
10	11	-1	1
11	15	-4	16
12	9	3	9
13	14	-1	1
14	12	2	4
15	16	-1	1
16	13	3	9
		$\Sigma D = 0$	$\Sigma D^2 = 136$

$$\rho = 1 - \frac{6\Sigma D^2}{N(N^2-1)} = 1 - \frac{6(136)}{16(255)} = \frac{4}{5}$$

$$= 0.8$$

(1)

Fathers (X)	Sons (Y)	Rank of X	Rank of Y	D = X − Y	D^2
65	68	9	5.5	3.5	12.25
63	66	11.5	8	3.5	12.25
67	68	6.5	5.5	1	1
64	65	10.5	9.5	1	1
68	69	4.5	3	1.5	2.25
62	66	12.5	8	4.5	20.25
70	68	2	5.5	−3.5	12.25
66	65	7.5	9.5	−2	4
68	71	4.5	1	3.5	12.25
69	68	3	5.5	−2.5	6.25
71	70	1	2	−1	1

$\Sigma D^2 = 87.75$

In X series, 68 repeated 2 times

$$C.F = \frac{m(m^2-1)}{12} = \frac{2(2^2-1)}{12} = 0.5$$

In Y series, 68 repeated 4 times

$$C.F = \frac{m(m^2-1)}{12} = \frac{4(4^2-1)}{12} = 5$$

In Y series, 66 repeated 2 times

$$C.F = \frac{m(m^2-1)}{12} = \frac{2(2^2-1)}{12} = 0.5$$

In Y. Series, 65 repeated 2 times,

$$C.F = \frac{m(m^2-1)}{12} = \frac{2(2^2-1)}{12} = 0.5$$

Rank correlation coefficient

$$\rho = 1 - \frac{6(\Sigma d^2 + C.F's)}{n(n^2-1)}$$

$$\rho = 1 - \frac{6(84.75 + 0.5 + 5 + 0.5)}{11(120)}$$

$$\rho = 1 - \frac{6(90.75)}{11(120)}$$

$$\rho = \frac{47}{80}$$

$$\rho = 0.5875$$

(10)

X	Rank (x)	Y	Rank (y)	$D = x - y$	D^2
48	8	13	5.5	2.5	6.25
33	6	13	5.5	0.5	0.25
40	7	24	10	-3	9.00
9	1	6	2.5	-1.5	2.25
16	3	15	7	4	16.00
16	~~10~~ 3	4	1	2	4.00
65	10	20	9	1	1.00
24	5	9	4	1	1.00
16	3	6	2.5	5	0.25
57	9	19	8	1	1.00
					$\Sigma D^2 = 41$

In x series, 16 repeated 3 times

$$C.F = \frac{m(m^2-1)}{12} = \frac{3(3^2-1)}{12} = 2$$

In y series 13 & 6 repeated 2 times

$$C.F = \frac{m(m^2-1)}{12} \Rightarrow \frac{2(2^2-1)}{12} = 0.5$$

Rank correlation coefficient,

$$\rho = 1 - \frac{(6\Sigma d^2 + C.F's)}{n(n^2-1)} = 1 - \frac{(6(41) + 2 + 0.5 + 0.5)}{10 \times 99}$$

$$\rho = 0.733$$

(11) Straight line is $Y = a + bx$

The two normal equations are

$$\Sigma Y = b\Sigma X + Na.$$

$$\Sigma XY = b\Sigma X^2 + a\Sigma X$$

X	X^2	Y	XY
10	100	10	100
12	144	22	264
13	169	24	312
16	256	27	432
17	289	29	493
20	400	33	660
25	625	37	925
$\Sigma X = 113$	$\Sigma X^2 = 1938$	$\Sigma Y = 182$	$\Sigma XY = 3186$

Substitute the values in the above normal equations, we get

$$113b + 7a = 182 \quad -(1)$$

$$1983b + 113a = 3186 \quad -(2)$$

Solve (1) & (2), we get $a = 0.82$

$b = 1.56$

Thus the equation of straight line is

$$Y = a + bX$$

$$Y = 0.82 + 1.56X.$$

$\downarrow$

Regression equation of Y on X.

T	$d_T = \frac{T-40}{10}$	d_T^2	S	$d_S = \frac{S-75}{10}$	d_S^2	$d_T d_S$
0	-4	16	54	-2.1	4.41	8.4
20	-2	4	65	-1.0	1.00	2.0
40	0	0	75	0	0.00	3.0
60	+2	4	85	1.0	1.00	2.0
80	+4	16	96	2.1	4.41	8.4
200	0	40	375	0	10.82	20.8

Now; $m = \frac{\Sigma d_T d_S}{\Sigma d_T^2} = \frac{20.8}{40} = 0.52$

and 'b' is given by the equation

$\Sigma S = m \Sigma T + Nb$

$375 = (0.52)(200) + 5b$

$b = 54.2$

When $T = 50°C$

$S = 0.52 \times 50 + 54.2$

$= 80.2$

(13). equation $Y = a + bX$

$\Sigma Y = na + b\Sigma X$

$\Sigma XY = a\Sigma X + a\Sigma X^2$

Given; $\Sigma X = 11.34$, $\Sigma Y = 20.78$, $\Sigma X^2 = 12.16$,

$\Sigma Y^2 = 84.96$, $\Sigma XY = 22.13$

Substituting the values,

$20.72 = 200a + 11.34b$

$a = \dfrac{20.72 - 11.34b}{200}$

$= 0.1036 - 0.0567b$

$= 0.1036 - 0.0567(1.82)$

$\boxed{\therefore a = 0.0005}$

$22.13 = 11.13a + 12.16$

$22.13 = 11.34(0.1036 - 0.0567b) + 12.16b$

$20.955 = 11.517$

$\boxed{b = 1.82}$

(14) $\tan\theta = \left(\dfrac{1-r^2}{r}\right)\dfrac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2} \Rightarrow \theta = \tan^{-1}\left(\dfrac{1-r^2}{r}\cdot\dfrac{\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}\right)$

Here $\sigma_x = \sigma_y = \sigma$

$\theta = \tan^{-1}\left(\dfrac{1-r^2}{r}\cdot\dfrac{\sigma^2}{2\sigma^2}\right) = \tan^{-1}\left(\dfrac{1-r^2}{2r}\right)$ — (1)

Given, $\theta = \tan^{-1}\left(\dfrac{4}{3}\right)$ — (2)

Compare (1) & (2)

$\Rightarrow \dfrac{1-r^2}{2r} = \dfrac{4}{3}$

$\Rightarrow 3r^2 + 8r - 3 = 0$

$r = 1/3$ or -3

$\boxed{r = 1/3 \text{ only}}$

(6) The line of regression of y on x is

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

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

The line of regression of x on y is

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

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(y-\bar{y})^2}$$

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
2	4	−6	−2	36	4	12
4	2	−4	−4	16	16	16
6	5	−2	−1	4	1	2
8	10	0	4	0	16	0
10	4	2	−2	4	4	−4
12	11	4	5	16	25	20
14	12	6	6	36	36	36
56	48			112	102	82

$\bar{x} = 8$ $\bar{y} = 6.8$

$$b_{yx} = \frac{82}{112} = 0.73$$

$$b_{xy} = \frac{82}{102} = 0.80$$

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

$y - 6 = 0.73(x - 8)$

Given; $\boxed{x = 13}$, $y = ?$

$y - 6 = 0.73(13 - 8)$

$y - 6 = 0.73(5)$

$\boxed{y = 9.65}$

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

$x - 8 = 0.80(y - 6)$

Given; $\boxed{y = 11.5}$ and $x = ?$

$x - 8 = 0.80(11.5 - 6)$

$x - 8 = 0.80(5.5)$

$\boxed{x = 12.4}$

(16) Given; $y = 0.399x + 6.394$

$x = 1.212y + 2.461$

Let $\bar{x}$ and $\bar{y}$ be the means. Then we have

$\bar{y} = 0.399\bar{x} + 6.394$ and $\bar{x} = 1.212\bar{y} + 2.461$

They can be written as

$$\bar{y} - 0.399\bar{x} = 6.394 \quad -(1)$$

$$-1.212\bar{y} + \bar{x} = 2.461 \quad -(2)$$

(1) × 1.212 gives

$$1.212\bar{y} - 0.483588\bar{x} = 7.7495$$

$$-1.212\bar{y} + \bar{x} = 2.461$$

$$0.516412\bar{x} = 10.2105$$

$$\bar{x} = 197720.03$$

$$\therefore \bar{y} = 78890.28$$

$b_{yx} = 0.399$

$b_{xy} = 1.212$

$$r = \sqrt{b_{yx} \cdot b_{xy}} = \sqrt{0.399 \times 1.212}$$

$$= \sqrt{0.483588}$$

$$= 0.6953$$

(17) We have to Calculate the value of Y when $X = 40$. So we have to find the regression equation of Y on X.

$\bar{x} = 30$; $\bar{y} = 500$; $\sigma_x = 5$; $\sigma_y = 100$.

$r = 0.8$

Regression of Y on X is

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

$$Y - 500 = (0.8) \frac{(100)}{5} (x - 30)$$

Given $x = 40$ $\Rightarrow$ $Y = 500 + \frac{4}{10} = 500.4$

Hence, the expected value of Y is 500.4 kg

---

(18) Let x = Height of the mother and Y = Height of the daughter.

Let $dx = X - 65$ and $dy = Y - 67$. Then,

$\Sigma x = 522$, $\Sigma dx = 2$, $\Sigma dx^2 = 50$, $\Sigma Y = 530$

$\Sigma dy = -6$, $\Sigma dy^2 = 74$, $\Sigma dx\,dy = 20$

Now; $\bar{x} = \frac{\Sigma x}{N} = 66.25$ & $\bar{y} = \frac{\Sigma Y}{N} = 65.25$

$$b_{yx} = \frac{\Sigma dx dy - \frac{\Sigma dx \cdot \Sigma dy}{N}}{\Sigma dx^2 - \frac{(\Sigma dx)^2}{N}} = 0.434$$

Hence, the regression equation of Y on X is

$$Y - \bar{Y} = b_{yx}(X - \bar{X})$$

$$Y = 37.43 + 0.434\bar{X}$$

when $X = 64.5$

$$\boxed{Y = 65.423}$$

(9)

P	X	$x = X - \bar{X}$	x^2	Y	$y = Y - \bar{Y}$	xy
1	46	+3	9	40	2	6
2	42	−1	1	38	0	0
3	44	+1	1	36	−2	2
4	40	−3	9	35	−3	9
5	43	0	0	39	1	0
6	41	−2	4	37	−1	2
7	45	+2	4	41	3	6
	301		28	266		21

$\bar{X} = 43$; $\bar{Y} = 38$;

Regression equation of Y on X is $Y - \bar{Y} = b_{yx}(X - \bar{X})$ — (1)

$$b_{yx} = \frac{\Sigma xy}{\Sigma x^2} = \frac{21}{28} = 0.75$$

from (1)

$$Y - 38 = 0.75(X - 43)$$

$$Y = 0.75(X) + 5.75$$

If $X = 37$, then $Y = 0.75(37) + 5.75 = 33.5$

Hence, if the judge Q would have been present, he would have awarded 33.5 marks to the eight performance.

(20)

x	y	$x-\bar{x}$	$y-\bar{y}$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1	6	-2	4	4	16	-8
5	1	2	-1	4	1	-2
3	0	0	-2	0	24	0
2	0	-1	-2	1	24	2
1	1	-2	-1	4	1	2
1	2	-2	0	4	0	0
7	1	4	-1	16	1	-4
3	5	0	3	0	9	0
23	16			33	76	-10

3

$\bar{x}$ 2.87 $\bar{y} = 2$

$$b_{yx} = r\frac{\sigma_y}{\sigma_x} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}$$

$$b_{yx} = \frac{-10}{33} = -0.303$$

$$b_{xy} = r\frac{\sigma_x}{\sigma_y} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(y-\bar{y})^2}$$

$$b_{xy} = \frac{-10}{76} = -0.131.$$

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

$y - 2 = -0.303\,(x - 3)$

when $\boxed{x = 10}$, $y = ?$

$y - 2 = -0.303\,(10 - 3)$

$y - 2 = -2.121$ (-2.121)

$\boxed{y = -0.121} \rightarrow -0.121$

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

$x - \overset{3}{\cancel{2.8}} = -0.131\,(y - 2)$

when; $\boxed{y = 2.5}$; $x = ?$

$x - \overset{3}{\cancel{2.8}} = -0.131\,(2.5 - 2)$

$\boxed{x = 2.93}$ (when $\bar{x} = 3$)

(or)

$\boxed{x = 2.736}$ (when $\bar{x} = 2.8$)

Values may change upon $\bar{x}$ & $\bar{y}$ values and approximation

Scrapped!

Prepared by:
K. Sai Chandu
AIML-C