Correlation & Regression Analysis

1. Correlation is the process of studying / establishing relationship / association between 2 or more variables although they are not in proportion.

EXAMPLES :
a. Sale of cold-drinks & temperature has positive correlation.
b. No. of trees & Rainfall has positive correlation.
c. No. of accident & claims, profit of insurance company has negative correlation
d. No. of accidents in Pune & USA's GDP, has No correlation

2. Studying correlation between

2 Variables

↓

Bi - variate correlation

positive correlation ($\uparrow\uparrow$, $\downarrow\downarrow$)

Negative correlation ($\uparrow\downarrow$, $\downarrow\uparrow$)

i) small/weak degree
ii) Moderate degree
iii) High degree
iv) perfect degree

3 or More Variables

↓

Multi - variate correlation

↓

Not in CA Foundation syllabus

3. Whether correlation between 2 variables exists OR Not?

Yes / No

Yes:

→ What is type / Nature of correlation?
- → positive correlation
- → Negative correlation

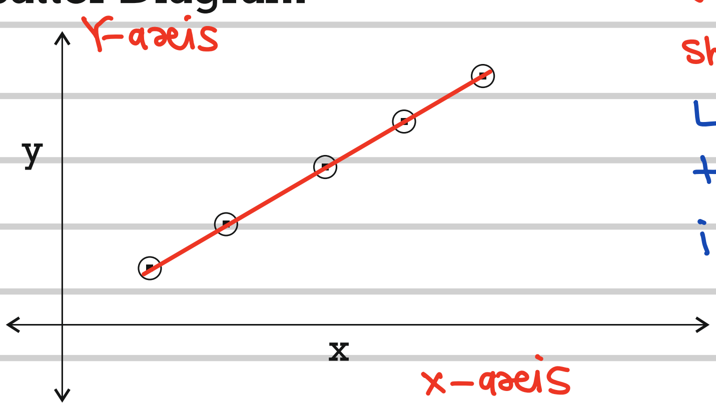
→ What is Degree of correlation?
- → Low/small/weak degree
- → Moderate degree
- → High/strong degree
- → perfect degree

4. There are 4 methods to measure correlation between 2 variables :

1. Scatter diagram
2. Spearman's rank correlation coefficient (γ)
3. Coefficient of concurrent deviation (γ)
4. Karl pearson's product moment correlation coefficient (γ)
(The Best method to obtain correlation betn 2 variables)

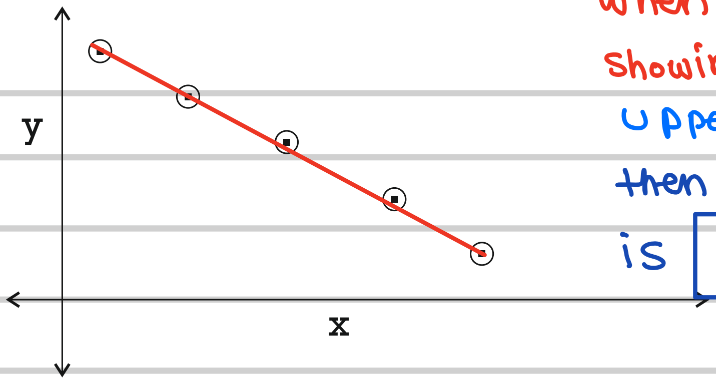
5. Scatter Diagram

a.



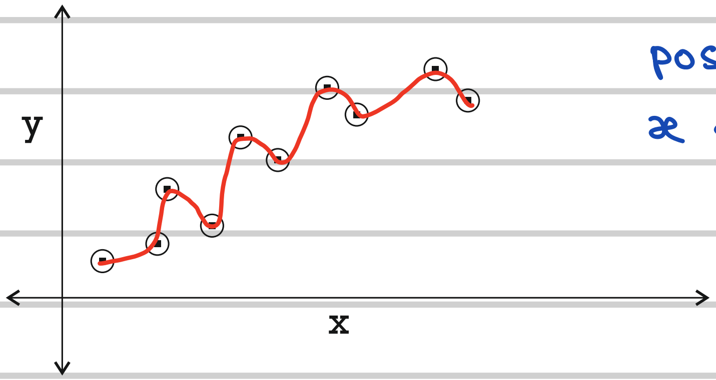
When scatter diagram is showing a straight line from Lower left to upper right then it is said that there is $\boxed{\text{perfect positive correlation}}$ between x & y

b.



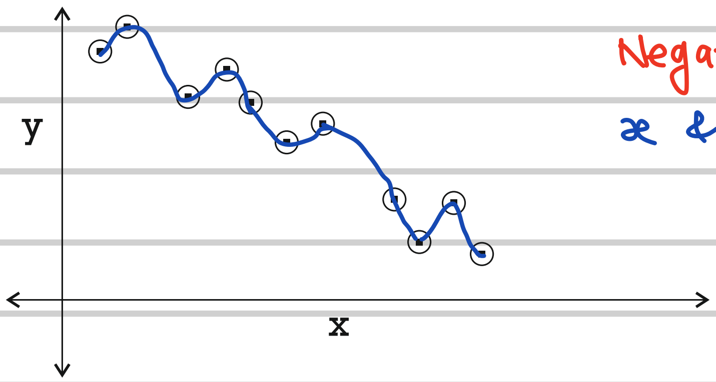
when scatter diagram is showing a straight line from upper left to Lower right then it is said that there is **perfect Negative correlation** between x & y

c.



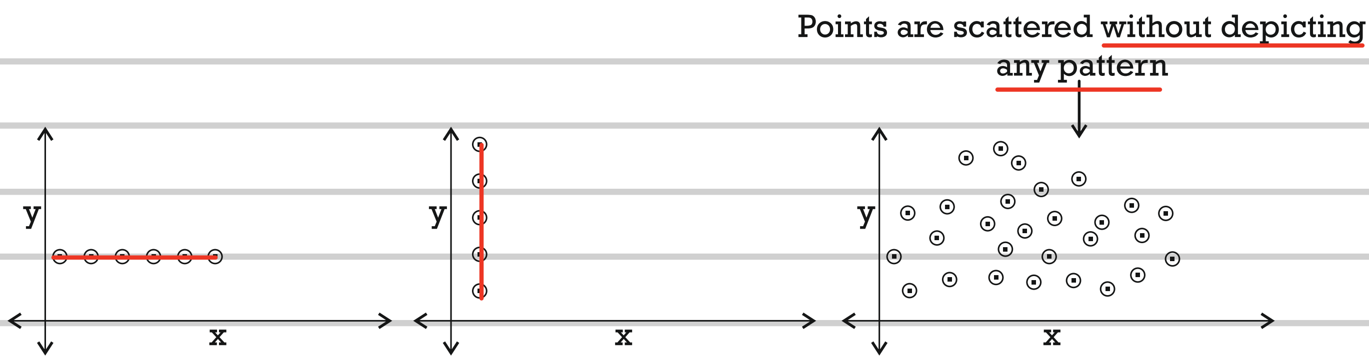
Scatter diagram is showing positive correlation between x & y.

d.



Scatter diagram is showing Negative correlation between x & y.

e.



Scatter diagrams are showing No correlation.

6. Scatter diagram can give an idea about Type / Nature of correlation but it can not give the exact degree of correlation.

To know type of correlation as well as degree of correlation we have following three methods :

1. Spearman's rank correlation coefficient (r)
2. Coefficient of concurrent deviation (r)
3. Karl pearson's product moment correlation coefficient (r) (The Best method to obtain correlation betn 2 variables)

7. Find Spearman's rank correlation coefficient for

x	56	59	73	81	79	52	48	88	93	2
y	280	730	831	631	789	123	666	581	983	281

→ n = No. of pairs of observations = 10

x	y	Rank of x	Rank of y	d^2
56	280	7	9	4
59	730	6	4	4
73	831	5	2	9
81	631	3	6	9
79	789	4	3	1
52	123	8	10	4
48	666	9	5	16
88	581	2	7	25
93	983	1	1	0
2	281	10	8	4
			Σd^2	$= 76$

Spearman's rank correlation coefficient

$$r = 1 - \left[\frac{6\,\Sigma d^2}{n(n^2-1)}\right]$$

$$r = 1 - \frac{6 \times 76}{10 \times (10^2 - 1)}$$

$$r = 1 - \frac{456}{990}$$

$$r = \frac{534}{990} = 0.5394$$

Moderate degree of positive correlation betn x & y

$-1.00 \leq r \leq 1.00$

8. $-1.00 \leq r \leq 1.00$

r = 1.00	perfect positive correlation
0.80 < r < 1.00	High/strong degree } of positive correlation
0.30 < r < 0.80	Moderate degree } of positive correlation
0 < r < 0.30	weak/small degree } of positive correlation
r = 0	$\boxed{\text{No correlation}}$
-0.30 < r < 0	weak/small degree } of Negative correlation
-0.80 < r < -0.30	Moderate degree } of Negative correlation
-1.00 < r < -0.80	High/strong degree } of Negative correlation
r = -1.00	perfect Negative correlation

Value of 'r' helps us to find type of correlation as well as Degree of Correlation

9. Find Spearman's rank correlation coefficient for

x	123	236	111	886	781	336	893
y	1013	986	993	781	286	583	960

Spearman's rank correlation coefficient

x	y	Rank of x	Rank of y	d^2
123	1013	6	1	25
236	986	5	3	4
111	993	7	2	25
886	781	2	5	9
781	286	3	7	16
336	583	4	6	4
893	960	1	4	9
			Σd^2	92

$$r = 1 - \left[\frac{6 \Sigma d^2}{n(n^2-1)}\right]$$

$$= 1 - \left(\frac{6 \times 92}{7 \times 48}\right)$$

$$= 1 - \frac{552}{336}$$

$$= 1 - 1.642857$$

$$= -0.642857$$

Moderate degree of Negative correlation.

math $a^2+b^2=c^2$

Question : Find spear man's coefficient

x	38	93	81	81	46	38	81	91
y	20	36	59	80	59	59	19	36

$\Rightarrow$ n = no. of pairs of obs^{n}s = 8 & t = (No. of obs^n involved in a tie) = 3, 2, 3, 2

x	Rank of x	y	Rank of y	d^2
38	7.50	20	7	0.25
93	1	36	5.50	20.25
81	4	59	3	1
81	4	80	1	9
46	6	59	3	9
38	7.50	59	3	20.25
81	4	19	8	16
91	2	36	5.50	12.25
				88

$$\sum\left(\frac{t^3-t}{12}\right) = \left(\frac{3^3-3}{12}\right) + \left(\frac{2^3-2}{12}\right) + \left(\frac{3^3-3}{12}\right) + \left(\frac{2^3-2}{12}\right)$$

$$= 2 + 0.50 + 2 + 0.50 = 5$$

Spear man's rank corr. coeffi. (with tie)

$$= 1 - \left[\frac{6\left(\sum d^2 + \sum \frac{t^3-t}{12}\right)}{n(n^2-1)}\right]$$

$$= 1 - \frac{6(88+5)}{8(64-1)} = 1 - \frac{558}{504}$$

$$= -0.1071$$

10. If scatter diagram is showing a straight **line** from

	Correlation is
Lower left to upper right	perfect positive
Upper left to lower right	Perfect Negative
Lower right to upper left	Perfect Negative
Upper right to lower left	perfect positive

$0 < r < 0.30$: small degree of positive correlation

$0.80 < r < 1.00$: High degree of positive correlation

$-0.80 < r < -0.30$: Moderate degree of negative correlation

$-0.30 < r < 0$: small degree of negative correlation

11. Find Spearman's rank correlation coefficient for

x	11	13	9	14	16	25	31	38
y	46	49	51	89	68	78	33	25

→

x	Rank of x	y	Rank of y	d^2
11	7	46	6	1
13	6	49	5	1
9	8	51	4	16
14	5	89	1	16
16	4	68	3	1
25	3	78	2	1
31	2	33	7	25
38	1	25	8	49
			Σd^2	110

Spearman's rank correlation coeffi. (without tie)

$$= 1 - \left[\frac{6\Sigma d^2}{n(n^2-1)}\right]$$

$$= 1 - \left(\frac{6 \times 110}{8 \times 63}\right)$$

$$= 1 - \frac{660}{504}$$

$$= -0.3095238095\overline{2}$$

12. r = 1.00 : Represents perfect positive correlation

r = -1.00 : Represents perfect negative correlation

r = 0 : Represents No correlation

Maximum Value of r = 1.00

Minimum Value of r = −1.00

$0 < r \leq 1.00$: positive correlation

$-1.00 \leq r < 0$: Negative correlation

Maximum value of $r^2 = 1.00$

Minimum value of $r^2 = 0.00$

math

13. Which of the following is correct

a. $0 < r \leq 1.00$

b. $-1.00 \leq r < 0$

c. $-1.00 < r < 1.00$

d. $-1.00 \leq r \leq 1.00$

CORRELATION COEFFICIENT WILL ALWAYS LIE BETWEEN -1.00 AND 1.00 INCLUDING BOTH LIMITS.

14. Find Spearman's rank correlation coefficient for

x	33	381	231	583
y	1081	2356	5523	1234

x	y	Rank of x	Rank of y	d^2
33	1081	4	4	0
381	2356	2	2	0
231	5523	3	1	4
583	1234	1	3	4

$\Sigma d^2 = 8$

Spearman's rank correlation coeffi. (without tie) $= 1 - \left[\frac{6 \Sigma d^2}{n(n^2-1)}\right] = 1 - \frac{6 \times 8}{4 \times (4^2-1)} = 1 - \frac{48}{60}$

$= 0.20$

Spearman's coefficient

without tie

$= 1 - \left[\frac{6 \Sigma d^2}{n(n^2-1)}\right]$

with tie

$= 1 - \left[\frac{6\left(\Sigma d^2 + \Sigma \frac{t^3-t}{12}\right)}{n(n^2-1)}\right]$

IMP

15. Find Spearman's rank correlation coefficient for

x	11	15	16	19	15	13	19	21	15	15	28
y	123	831	583	236	583	781	123	123	281	560	281

→

x	Rank of x	y	Rank of y	d^2
11	11	123	10	1
15	7.50	831	1	42.25
16	5	583	3.50	2.25
19	3.50	236	8	20.25
15	7.50	583	3.50	16
13	10	781	2	64
19	3.50	123	10	42.25
21	2	123	10	64
15	7.50	281	6.50	1
15	7.50	560	5	6.25
28	1	281	6.50	30.25
				289.50

$t = 2, 4, 2, 2, 3$ = No. of obs^n involved in a tie

$$\Sigma\left(\frac{t^3-t}{12}\right) = \frac{2^3-2}{12} + \frac{4^3-4}{12} + \frac{2^3-2}{12} + \frac{2^3-2}{12} + \frac{3^3-3}{12}$$

$$= 0.50 + 5 + 0.50 + 0.50 + 2 = 8.50$$

Spearman's rank correlation coeffi. (with tie) $= 1 - \left[\dfrac{6\left(\Sigma d^2 + \Sigma \frac{t^3-t}{12}\right)}{n(n^2-1)}\right]$

$$= 1 - \left[\frac{6(289.50 + 8.50)}{11(11^2-1)}\right] = 1 - \frac{1788}{1320} = -0.35454545$$

16. SPEARMAN'S RANK CORRELATION COEFFICIENT

Without tie

$$= 1 - \left[\frac{6\sum d^2}{n(n^2-1)}\right]$$

With tie

$$= 1 - \left[\frac{6\left(\sum d^2 + \sum \frac{t^3-t}{12}\right)}{n(n^2-1)}\right]$$

where

t = No. of obs^{n}s involved in a tie.

n = No. of pairs of observations

d = Diff of rank

17. Find Spearman's rank correlation coefficient

x	50	80	50	50	45
y	81	93	56	56	28

x	Rank of x	y	Rank of y	d^2
50	3	81	2	1
80	1	93	1	0
50	3	56	3.50	0.25
50	3	56	3.50	0.25
45	5	28	5	0
				1.50

$t = 3, 2$

$$\sum\left(\frac{t^3-t}{12}\right) = \left(\frac{3^3-3}{12}\right) + \left(\frac{2^3-2}{12}\right)$$

$$= 2 + 0.50$$

$$= 2.50$$

$$r = 1 - \left[\frac{6\left(\sum d^2 + \sum \frac{t^3-t}{12}\right)}{n(n^2-1)}\right]$$

$$r = 1 - \left[\frac{6(1.50+2.50)}{5\times(5^2-1)}\right] = 1 - \frac{24}{5\times 24} = \frac{4}{5}$$

$$= 0.80$$

18. Find Coefficient of Concurrent Deviation for :

x	58	63	28	93	58	26	33	38	59	83
y	123	633	833	589	289	333	533	888	231	123

x	y	Deviation of x	Deviation of y	product
58	123			
63	633	+	+	+
28	833	−	+	−
93	589	+	−	−
58	289	−	−	+
26	333	−	+	−
33	533	+	+	+
38	888	+	+	+
59	231	+	−	−
83	123	+	−	−

$n =$ No. of pairs of obs^{n}s $= 10$

$m =$ No. of deviation $= n-1$

$m = 10 - 1 = 9$

$c =$ No. of concurrent deviations

$=$ No. of '+' signs in product columns

$= 4$

coeffi. of concurrent deviation

$$r = \pm\sqrt{\pm\left(\frac{2c-m}{m}\right)} = -\sqrt{-\left(\frac{2(4)-9}{9}\right)}$$

$$r = -\sqrt{+\frac{1}{9}} = -\frac{1}{3} = -0.33333333$$

19. If $\left(\frac{2c-m}{m}\right) < 0$ then $-1 \leq r < 0$

If $\left(\frac{2c-m}{m}\right) > 0$ then $0 < r \leq 1.00$

If $\left(\frac{2c-m}{m}\right) = 0$ then $r = 0$

20. No. of positive signs in product = x column

No. of Negative signs in product = y column

and x > y	$1.00 \geq r > 0$
x < y	$-1.00 \leq r < 0$
x = y	$r = 0$

21. Find Coefficient of Concurrent Deviation & Spearman's Rank Correlation Coeeficient :

x	10	18	10	26	11
y	53	91	98	53	28

x	y	Devi. of x	Devi. of y	product
10	53			
18	91	+	+	+
10	98	−	+	−
26	53	+	−	−
11	28	−	−	+

$r = \pm\sqrt{\pm\left(\frac{2c-m}{m}\right)}$

$= \sqrt{\frac{2(2)-4}{4}} = 0$

Here $c = 2$, $m = 4$

x	y	Rank of x	Rank of y	d^2
10	53	4.50	3.50	1
18	91	2	2	0
10	98	4.50	1	12.25
26	53	1	3.50	6.25
11	28	3	5	4.00

$\Sigma d^2 = 23.50$

$t = 2, 2$

$$\Sigma\left(\frac{t^3-t}{12}\right) = \frac{2^3-2}{12} + \frac{2^3-2}{12}$$

$= 0.50 + 0.50$

$= 1$

Spearman's coefficient

$$= 1 - \frac{6\,(23.50+1)}{5\,(5^2-1)}$$

$$= 1 - \left(\frac{147}{120}\right) = -0.225$$

22. 1. Coefficient of Concurrent Deviation Method ignores the amount or magnitude of change but it considers only direction of change We can say that, it is a casual to obtain correlation coefficient.

2. In order to know degree of agreement of 2 Judges in a Beauty contest we can use :

→ Spearman's rank correlation coefficient

	Judge-1 PS	Judge-2 SB	d^2
A	5	3	4
B	1	7	36
partici. C	4	5	1
D	3	6	9
E	6	4	4
F	2	8	36
G	8	1	49
H	7	2	25

$$r = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 164}{8 \times 63}$$

$$= 1 - \frac{984}{504}$$

$= -0.95238$

There is strong disagreement beth opinion of 2 Judges.

23. Find Karl Pearson's Product Moment Correlation Coefficient for :

x	20	30	28	72
y	82	70	38	40

→

x	y	x^2	y^2	xy
20	82			
30	70			
28	38			
72	40			
150	230	7268	14668	7684

$$\bar{x} = \frac{\Sigma x}{n} = \frac{150}{4} = 37.50$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{230}{4} = 57.50$$

$$SD_x = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2} = \sqrt{\frac{7268}{4} - 37.50^2} = 20.2669681995$$

$$SD_y = \sqrt{\frac{\Sigma y^2}{n} - \bar{y}^2} = \sqrt{\frac{14668}{4} - 57.50^2} = 18.9934199132$$

$$\text{co-variance of } x,y = \left[\frac{\Sigma x.y}{n} - \bar{x}.\bar{y}\right] = \frac{7684}{4} - (37.50 \times 57.50)$$

$$= -235.25$$

$$\text{karl-pearson's product moment correlation coefficient} = \frac{cov(x,y)}{SD_x \times SD_y}$$

$$= \left(\frac{-235.25}{20.2669681995 \times 18.9934199132}\right) = -0.611136$$

24. Karl Pearson's Product Moment Correlation Coefficient =

$$r = \boxed{\frac{\operatorname{cov}(x,y)}{\sigma_x \cdot \sigma_y}}$$

$$r = \left[\frac{\frac{\sum xy}{n} - \bar{x}\cdot\bar{y}}{\sqrt{\frac{\sum x^2}{n} - \bar{x}^2} \times \sqrt{\frac{\sum y^2}{n} - \bar{y}^2}}\right]$$

$$r = \left[\frac{\frac{\sum xy}{n} - \left(\frac{\sum x}{n}\cdot\frac{\sum y}{n}\right)}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \times \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}}\right]$$

$$r = \left[\frac{\sum\left[(x-\bar{x}) \times (y-\bar{y})\right]}{\sqrt{\sum(x-\bar{x})^2} \times \sqrt{\sum(y-\bar{y})^2}}\right]$$

a. Karl Pearson's Product Moment Correlation Coefficient

$$= \left[\frac{\frac{\sum xy}{n} - \left(\frac{\sum x}{n} \times \frac{\sum y}{n}\right)}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \times \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}} \right]$$

$$= \left(\frac{\text{Covariance of } (x,y)}{\sigma x \,.\, \sigma y}\right)$$

b. Coefficient of concurrent deviation

$$= \pm \sqrt{\pm\left(\frac{2c - m}{m}\right)}$$

c. Spearman's Rank Correlation Coefficient

Without tie

$$= 1 - \left[\frac{6\,\Sigma d^2}{n(n^2-1)}\right]$$

With tie

$$= 1 - \left[\frac{6\left(\Sigma d^2 + \Sigma \frac{t^3 - t}{12}\right)}{n(n^2-1)}\right]$$

25. In case of Qualitative data / measurements which of the following is suitable

a. Scatter Diagram
~~b.~~ Spearman's Coefficient
c. Karl Pearson's Coefficient
d. Coefficient of Concurrent Deviation

26.

x	10	18	12	20
y	30	20	25	15

Find Correlation Coefficient (Karl Pearson's)

$\bar{x} = 15 \qquad \Sigma x^2 = 968 \qquad \Sigma xy = 1260$

$\bar{y} = 22.50 \qquad \Sigma y^2 = 2150$

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

$$r = \frac{\frac{1260}{4} - (15 \times 22.50)}{\sqrt{\frac{968}{4} - 15^2} \times \sqrt{\frac{2150}{4} - 22.50^2}} = \frac{-22.50}{\sqrt{17} \times \sqrt{31.25}} = \frac{-22.50}{\sqrt{17 \times 31.25}}$$

$$= -0.9762$$

$$\text{Co-covariance of } (x,y) = \frac{\Sigma x.y}{n} - (\bar{x}.\bar{y})$$

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

$$\sigma_y = \sqrt{\frac{\Sigma y^2}{n} - (\bar{y})^2}$$

$$r = \left[\frac{(\text{Covariance of } x,y)}{SD_x \times SD_y}\right]$$

Karl Pearson's Method

27. Find Karl Pearson's Coefficient

x	50	60	48	72	30
y	112	118	138	92	60

$\bar{x} = 52$, $\bar{y} = 104$, $\Sigma x^2 = 14488$, $\Sigma y^2 = 57576$

$\Sigma xy = 27728$

$$\text{cov}(x,y) = \frac{\Sigma xy}{n} - \bar{x}.\bar{y} = \frac{27728}{5} - (52 \times 104) = 137.60$$

$$SD_x = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2} = \sqrt{\frac{14488}{5} - 52^2} = \sqrt{193.60}$$

$$SD_y = \sqrt{\frac{\Sigma y^2}{n} - \bar{y}^2} = \sqrt{\frac{57576}{5} - 104^2} = \sqrt{699.20}$$

$$r = \frac{\text{cov}(x,y)}{SD_x . SD_y} = \frac{137.60}{\sqrt{193.60 \times 699.20}} = 0.374$$

28. Find Correlation Coefficient by all 3 methods

x	11	12	9	18	10
y	15	12	33	8	22

x	y	Devi. of x	Devi. of y	Product	Rank of x	Rank of y	d^2	x^2	y^2	xy
11	15				3	3	0			
12	12	+	−	−	2	4	4			
9	33	−	+	−	5	1	16			
18	8	+	−	−	1	5	16			
10	22	−	+	−	4	2	4			
							40	770	2006	970

$\bar{x} = 12$

$\bar{y} = 18$

① Karl pearson's coeffi. $= \dfrac{\frac{970}{5} - (12 \times 18)}{\sqrt{\frac{770}{5} - 12^2} \times \sqrt{\frac{2006}{5} - 18^2}} = \dfrac{-22}{\sqrt{10} \times \sqrt{77.20}}$

$= -0.7918$

② spearman's coeffi. $= 1 - \left(\dfrac{6 \times 40}{5 \times 24}\right) = 1 - 2 = -1.00$

③ coeffi. of concurrent deviation $= \pm\sqrt{\pm\left(\dfrac{2c-m}{m}\right)} = -\sqrt{-\left(\dfrac{2(0)-4}{4}\right)} = -1.00$

29. Find Correlation Coefficient by all 3 methods (Home-work)

x	315	833	292	300
y	400	282	188	150

$\bar{x} = 435$ $\Sigma x^2 = 968378$ $\Sigma xy = 460802$

$\bar{y} = 255$ $\Sigma y^2 = 297368$

1. Karl Pearson's coefficient = $\dfrac{4275.50}{\sqrt{52869.50 \times 9317}} = 0.19264$

2. Spearman's rank correlation coefficient =

$= 1 - \dfrac{6 \times 4}{4 \times 15} = 1 - \dfrac{24}{60} = 0.60$

3. Coefficient of Concurrent Deviation =

$= \pm\sqrt{\pm\dfrac{2(1)-3}{3}} = -0.57735$

30. Find Correlation Coefficient by all 3 methods

x	10	30
y	20	100

$\boxed{\text{when } n=2 : \text{Either } r=-1.00 \text{ OR } r=1.00}$

x	y	Rank of x	Rank of y	d^2	Devi. of x	Devi. of y	product	xy
10	20	2	2	0				200
30	100	1	1	0	+	+	+	3000

Spearman's rank correlation coefficient = $1 - \frac{6\Sigma d^2}{n(n^2-1)} = 1 - \frac{6 \times 0}{2 \times (3)}$

$= 1 - 0 = 1.00$

Coefficient of Concurrent Deviation = $\pm\sqrt{\pm\frac{2c-m}{m}} = \sqrt{\frac{2(1)-1}{1}} = 1.00$

Karl Pearson's coefficient = $\frac{\frac{3200}{2} - 20 \times 60}{10 \times 40} = \frac{400}{400} = 1.00$

31. Find Correlation Coefficient by all 3 methods

x	60	90
y	20	10

x	y	Rank of x	Rank of y	d^2	Devi. of x	Devi. of y	product	xy
60	20	2	1	1				1200
90	10	1	2	1	+	−	−	900

$\Sigma d^2 = 2$ 2100

$\bar{x} = 75$

$\bar{y} = 15$ $SD_x = 15$, $SD_y = 5$

Spearman's rank correlation coefficient = $1 - \frac{6 \times 2}{2 \times 3} = 1 - \frac{12}{6} = -1.00$

Coefficient of Concurrent Deviation = $\pm\sqrt{\pm\frac{2(0)-1}{1}} = -1.00$

Karl Pearson's coefficient = $\frac{\frac{2100}{2} - 75 \times 15}{15 \times 5} = \frac{-75}{75} = -1.00$

32. When n = 2 then value of 'r' Must be either 1.00 OR −1.00

x	y
10	30
20	80

r = 1.00

x	y
50	800
15	600

r = 1.00

x	y
100	8500
150	6800

r = −1.00

x	y
1000	2800
850	3500

r = −1.00

33.

x	80	60	70	100
y	35	65	75	25

Find cov (x,y)

→

x	y	xy
80	35	
60	65	
70	75	
100	25	
310	200	14450

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

$$= \frac{14450}{4} - (77.50 \times 50)$$

$$= -262.50$$

$$cov(x,y) = \frac{\sum[(x-\bar{x})(y-\bar{y})]}{n} = \frac{-1050}{4}$$

$$= -262.50$$

OR

x	$(x - \bar{x})$	y	$(y - \bar{y})$	$(x - \bar{x})(y - \bar{y})$
80	2.50	35	−15	−37.50
60	−17.50	65	15	−262.50
70	−7.50	75	25	−187.50
100	22.50	25	−25	−562.50
				$\sum[(x-\bar{x})(y-\bar{y})] = -1050$

34. $[\sum(x - \bar{x})(y - \bar{y})] = 1200$

$n = 22$

Find cov (x,y)

$$cov(x,y) = \left[\frac{\sum[(x-\bar{x})(y-\bar{y})]}{n}\right] = \frac{1200}{22}$$

$$= 54.5454545454$$

$$\boxed{cov(x,y) = \left[\frac{\sum xy}{n} - \bar{x}\cdot\bar{y}\right] \text{ OR } \left[\frac{\sum[(x-\bar{x})(y-\bar{y})]}{n}\right]}$$

35. If cov (x,y) = 125, $\sigma_x = 10.50$, $\sigma_y = 13.80$. Find r.

~~a.~~ 0.8626 b. -0.8626 c. a or b d. Can't say

$$r = \left[125 / 10.50 \times 13.80\right] = 0.8626$$

36. If cov (x,y) = 138.50, $\sigma_x = 8.53$, $\sigma_y = 9.80$. Find r.

a. -1.6568 b. 1.6568 c. a or b ~~d.~~ Wrong data

$$r = \left[138.50 / 8.53 \times 9.80\right] = 1.6568$$ as r can't be greater than 1.00

37.

x	y
30	80
40	78

Find r.

$$\boxed{\text{when } n = 2 \text{ then either } r = 1.00 \text{ OR } r = -1.00}$$

a. 1.00 ~~b.~~ -1.00 c. 0.00 d. Can't find

38. If cov(x,y) > 0 then

~~a.~~ $1.00 \geq r > 0$

b. $1.00 \geq r \geq -1$

c. $0 \geq r > -1$

d. None of these

39. If cov(x,y) < 0 then

$-1 \leq r < 0$

40. If cov (x,y) = 0 then r is also Zero → $r = 0$

cov (x,y) = Positive then r is also Positive → $0 < r \leq 1.00$

cov (x,y) = Negative then r is also Negative → $-1 \leq r < 0$

$$\boxed{\text{Cov}(x,y) = \frac{\Sigma x.y}{n} - (\bar{x}.\bar{y}) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n}}$$

41. Karl Pearson's Product Moment Correlation Coefficient $= \left[\frac{\text{Cov}(x,y)}{SD_x \times SD_y}\right]$

$$r = \frac{\frac{\Sigma x.y}{n} - (\bar{x}.\bar{y})}{\sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2} \times \sqrt{\frac{\Sigma y^2}{n} - (\bar{y})^2}}$$

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

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

$$r = \left[\frac{\Sigma X.Y}{\sqrt{\Sigma X^2} \times \sqrt{\Sigma Y^2}}\right]$$

where $X = x - \bar{x}$

$Y = y - \bar{y}$

42. $\Sigma(x-\bar{x})(y-\bar{y}) = 2500$

$\Sigma(x-\bar{x})^2 = 10,000$

$\Sigma(y-\bar{y})^2 = 8,250$

Find 'r'

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2 \times \Sigma(y-\bar{y})^2}} = \frac{2500}{\sqrt{10{,}000 \times 8250}}$$

$r = 0.275241$

43.

x	10	22	63	35
y	5	3	2	10

Find r_{xy}

$\bar{x} = 32.50$

$\bar{y} = 5.00$

$$SD_x = \sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$$
$$= \sqrt{\frac{5778}{4} - (32.50)^2}$$
$$= 19.7041$$

$$SD_y = \sqrt{\frac{\sum y^2}{n} - (\bar{y})^2}$$
$$= \sqrt{\frac{138}{4} - (5)^2}$$
$$= 3.08221$$

$$Cov(x,y) = \frac{\sum x.y}{n} - (\bar{x}.\bar{y})$$
$$= \frac{592}{4} - (32.50 \times 5)$$
$$= -14.50$$

$$r_{xy} = \left(\frac{-14.50}{19.7041 \times 3.08221}\right)$$

$$r_{xy} = -0.2388$$

u = 2x+10	30	54	136	80
v = 5y-8	17	7	2	42

Find r_{uv}

$\bar{u} = 75$

$\bar{v} = 17$

$$SD_u = \sqrt{\frac{\sum u^2}{n} - (\bar{u})^2}$$
$$= \sqrt{\frac{28712}{4} - (75)^2}$$
$$= 39.40812$$

$$SD_v = \sqrt{\frac{\sum v^2}{n} - (\bar{v})^2}$$
$$= \sqrt{\frac{2106}{4} - (17)^2}$$
$$= 15.4110$$

$$Cov(u,v) = \frac{\sum uv}{n} - (\bar{u}.\bar{v})$$
$$= \frac{4520}{4} - (75 \times 17)$$
$$= -145$$

$$r_{uv} = \left(\frac{-145}{39.40812 \times 15.4110}\right)$$

$$r_{uv} = -0.2388$$

'r' is not affected by Change of Origin as well as by Change in Scale

Covariance is affected only by Change in Scale & not by Change of Origin.

44. Value of 'r' helps us to know :

a. Type / Nature of correlation

b. Degree of correlation

c. Both of these

d. None of these

45. Scatter diagram can help us to know :

a. Type / Nature of correlation

b. Degree of correlation

c. Both of these

d. None of these

46.

x	30	35	39	40	38	60	63
y	40	38	28	39	86	80	98

To find 'r' by concurrent devi. method. Find c, m. Also find 'r'

→

x	y	Devi. of x	Devi. of y	product
30	40			
35	38	+	−	−
39	28	+	−	−
40	39	+	+	+
38	86	−	+	−
60	80	+	−	−
63	98	+	+	+

c = No. of concurrent deviations

= No. of '+' signs in product column

= 2

m = n − 1 = 7 − 1 = 6

$$r = \pm\sqrt{\pm\left(\frac{2c-m}{m}\right)} = -\sqrt{-\left[\frac{2(2)-6}{6}\right]} = -0.57735$$

math

47. If n = 4, $\sum d^2$ = 49. Find Spearman's Coefficient

a. -3.90   b. 3.90   c. a or b   ~~d.~~ Wrong data

$r = 1 - \frac{6 \times 49}{4 \times (4^2-1)} = 1 - \frac{294}{60} = 1 - 4.90 = -3.90$

as $\boxed{\text{As } -1.00 \le r \le 1.00}$

48. $SD_x = 3$, If u = 8x + 20

$SD_y = 5$, If v = -3y + 33

Find SD_u, SD_v

$SD_u = 8 \times SD_x = 8 \times 3 = 24$

$SD_v = |-3| \times SD_y = 3 \times 5 = 15$

$cov(u,v) = 8 \times -3 \times cov(x,y)$

$cov(u,v) = -24 \times cov(x,y)$

$(r_{uv}) = \frac{cov(u,v)}{\sigma_u \cdot \sigma_v} = \frac{-\cancel{24} \times cov(x,y)}{\cancel{8} \times SD_x \times \cancel{3} \times SD_y} = (-r_{xy})$

49. u = mx + 33

v = ky - 88

then cov (u,v) = m x k x cov (x,y)

50. u = mx - 86 $SD_u = |m| \times SD_x$

v = jy + 200   then $SD_v = |j| \times Sd_y$

cov (u,v) = m x j x cov (x,y)

51. $u = 3x + 18$	$u = -3x + 33$	$u = 81x + 55$
$v = 8y - 33$	$v = -18y - 33$	$v = -33y + 21$
$r_{xy} = -0.80$	$r_{xy} = -0.57$	$r_{xy} = -0.80$
$r_{uv} = -0.80$	$r_{uv} = -0.57$	$r_{uv} = 0.80$
$u = -3x + 21$	$u = \frac{3}{2}x + 55$	$u = 188x + 22$
$v = 88y - 33$	$v = -\frac{5}{7}y + 23$	$v = 33y + 56$
$r_{xy} = 0.56$	$r_{xy} = -0.583$	$r_{xy} = 0.81812$
$r_{uv} = -0.56$	$r_{uv} = 0.583$	$r_{uv} = 0.81812$
$u = 13x - 212$	$u = -15y - 21$	$u = -1.50x + 21$
$v = -18y + 63$	$v = -18x + 33$	$v = 81y - 33$
$r_{uv} = -0.63$	$r_{uv} = 0.2121$	$r_{uv} = -0.8651$
$r_{xy} = 0.63$	$r_{xy} = 0.2121$	$r_{xy} = 0.8651$

52. If $\sum d^2 = 2, n = 4$, . Find Spearman's Coefficient

$$r = 1 - \frac{6 \Sigma d^2}{n(n^2-1)}$$

$$= 1 - \frac{6 \times 2}{4 \times 15} = 1 - \frac{12}{60} = 1 - 0.20 = 0.80$$

53. Vinod Reddy - C.E.O
Rishit - Marketing Director

x - Marketing Exp. for the year
y - Turnover for the year
$r_{xy} = 0.8123$

Question from VR to Rishit

x = 580 crores y = ?

y = 7800 crores x = ?

After studying correlation between 2 variables process of estimating / predicting / determining value of one variable on the basis of other is known as **'Regression Analysis'**

54. Regression Analysis is used to find

a. Relation between 2 variables

~~b.~~ One variable on the basis of other

c. Both of these

d. None of these

55. Presence of correlation between 2 variables is the PRE-REQUISITE for studying Regression Analysis

⇓

If 2 variables are un-correlated then there is no question of using Regression analysis

56. $\boxed{\text{IMP}}$ (concept of Least squares)

Determining value of one variable on the basis of other is known as

Regression Analysis

x = Given

y = ?

y is to be estimated on the basis of given value of x

↓

Regression of y on x

↓

To estimate y on the basis of x, we use Eqn of Reg. line of y on x

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

where b_{yx} = Regression coeffi. of y on x

↓

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

$$b_{yx} = \frac{cov(x,y)}{\sigma_x \cdot \cancel{\sigma_y}} \times \frac{\cancel{\sigma_y}}{\sigma_x}$$

$$b_{yx} = \left(\frac{cov(x,y)}{\text{variance of } x}\right)$$

y = Given

x = ?

x is to be estimated on the basis of given value of y

↓

Regression of x on y

↓

To estimate x on the basis of y we use Eqn of Reg. line of x on y

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

where b_{xy} = Regression coeffi. of x on y

$$b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y}$$

$$b_{xy} = \frac{cov(x,y)}{\cancel{\sigma_x} \cdot \sigma_y} \times \frac{\cancel{\sigma_x}}{\sigma_y}$$

$$b_{xy} = \frac{cov(x,y)}{\text{variance of } y}$$

57. $r = -0.80, \sigma_y = 8, \sigma_x = 5$. Find b_{yx}, b_{xy}

$$b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x} = -0.80 \times \frac{8}{5} = -1.28$$

$$b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y} = -0.80 \times \frac{5}{8} = -0.50$$

58. If r is positive then b_{yx} & b_{xy} are also positive & vice versa.

If r is negative then b_{yx} & b_{xy} are also negative & vice versa.

$$b_{yx} = r \cdot \left(\frac{\sigma_y}{\sigma_x}\right)$$

In short b_{yx}, b_{xy}, r either

All these 3 will be Positive

OR

All these 3 will be Negative

OR

All these are Zero

$\boxed{\text{$b_{yx}$, b_{xy} \& r all these 3 coeffi. will always have same sign}}$

59. $r = -0.8136, \sigma_x = +1.2381, \sigma_y = +2.8133$. Find b_{yx} & b_{xy}

$$b_{yx} = \left[r \cdot \frac{\sigma_y}{\sigma_x}\right] = -0.8136 \times \frac{2.8133}{1.2381}$$

$$= -1.848720\ 52338$$

$$b_{xy} = \left[r \cdot \frac{\sigma_x}{\sigma_y}\right] = -0.8136 \times \frac{1.2381}{2.8133}$$

$$= -0.3580557 2103$$

$$(b_{yx} \times b_{xy}) = r \cdot \frac{\cancel{\sigma_y}}{\cancel{\sigma_x}} \times r \cdot \frac{\cancel{\sigma_x}}{\cancel{\sigma_y}} = r^2 \qquad \therefore r^2 = b_{yx} \cdot b_{xy}$$

60. If $\bar{x} = 50, \bar{y} = 80, \sigma_x = 2, \sigma_y = 5, r = 0.90$

Find a. Reg. line of y on x

b. Reg. line of x on y

c. $x = 58, y = ?$

d. $y = 81.50, x = ?$

① Eqn of Regression Line of y on x

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

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

$$y - 80 = 0.90 \times \frac{5}{2} \times (x - 50)$$

$$y - 80 = 2.25(x - 50)$$

$$y - 80 = 2.25x - 112.50$$

$$\boxed{y = -32.50 + 2.25x}$$

② Eqn of Regression Line of x on y

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

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

$$x - 50 = 0.90 \times \frac{2}{5} \times (y - 80)$$

$$x - 50 = 0.36(y - 80)$$

$$x - 50 = 0.36y - 28.80$$

$$\boxed{x = 21.20 + 0.36y}$$

③ $x = 58, y = ?$

$$y = -32.50 + 2.25x = -32.50 + 2.25(58)$$

$$y = 98$$

④ $y = 81.50, x = ?$

$$x = 21.20 + 0.36y = 21.20 + 0.36(81.50) = 50.54$$

61. $\bar{x} = 280, \bar{y} = 350, \sigma_x = 20, \sigma_y = 60, r = 0.80$

If $x = 295.80, y = ?$

$\Rightarrow$ $y - \bar{y} = b_{yx}(x - \bar{x})$ ..... Reg line of y on x

$y - 350 = 0.80 \times \frac{60}{20} \times (295.80 - 280)$

$\boxed{y = 387.92}$

62. $\bar{x} = 81.80, \bar{y} = 36.52, \sigma_x = 5.82, \sigma_y = 3.53, r = -0.75$, If $y = 50, x = ?$

$\Rightarrow$ $(x - \bar{x}) = b_{xy}(y - \bar{y})$ .... Reg. line of x on y

$x - 81.80 = -0.75 \times \frac{5.82}{3.53} \times (50 - 36.52)$

$\boxed{x = 65.13138810 2}$

If $\bar{x} = 92, \bar{y} = 80, \sigma_x = 4, \sigma_y = 8, r = 0.85$

① If $x = 98.20, y = ?$ ② If $y = 75.80, x = ?$

$\Rightarrow$ ① $(y - 80) = 0.85 \times \frac{8}{4} \times (98.20 - 92)$

$\boxed{y = 90.54}$

② $(x - 92) = 0.85 \times \frac{4}{8} \times (75.80 - 80)$

$\boxed{x = 90.215}$

63. $\bar{x} = 135.22, \bar{y} = 1083.96, \sigma_x = 12.81, \sigma_y = 83.56, r = -0.8122$,

If $x = 148.53, y = ?$

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

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

$y - 1083.96 = -0.8122 \times \frac{83.56}{12.81} \times (148.53 - 135.22)$

$\boxed{y = 1013.44356598}$

64. $b_{yx} = r.\frac{\sigma_y}{\sigma_x}$, $b_{xy} = r.\frac{\sigma_x}{\sigma_y}$

$b_{yx}.b_{xy} = r.\frac{\cancel{\sigma_y}}{\cancel{\sigma_x}} \times r.\frac{\cancel{\sigma_x}}{\cancel{\sigma_y}}$

$\boxed{b_{yx}.b_{xy} = r^2}$

$\therefore r^2 = b_{yx}.b_{xy}$

$\boxed{r = \sqrt{b_{yx}.b_{xy}}}$

$\therefore$ 'r' is the G.M. of b_{yx} & b_{xy}

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

'r' is GM of b_{yx} & b_{xy}

$\therefore$ correlation coeffi. 'r' is the GM of 2 regression coefficients b_{yx} & b_{xy}

Correlation coefficient is the Geometric Mean of 2 regression coefficients

65. If $b_{yx} = 1.53$, $b_{xy} = 0.3136$. Find 'r'

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

$= 1.53 \times 0.3136$

$r^2 = 0.479808$

$r = \sqrt{0.479808}$

$\boxed{r = +0.6926817451}$

$-1.00 \leq r \leq 1.00$

$0.00 \leq r^2 \leq 1.00$

$0.00 \leq b_{yx} \cdot b_{xy} \leq 1.00$

Min. value of $r = -1.00$

Max. value of $r = +1.00$

Min. value of $r^2 = 0.00$

Max. value of $r^2 = 1.00$

66. If $b_{yx} = -5.8188$, $b_{xy} = -0.125833$. Find 'r'

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

$= -5.8188 \times -0.125833 = 0.7321970604$

$r = -0.8556851409 2$

$b_{yx} = 1.8592$, $b_{xy} = 0.6592$, $r = ?$

(a) 1.10706 (b) -1.10706 (c) 1.22558 ~~(d)~~ wrong data

Pls. Note that $(b_{yx} \cdot b_{xy}) = r^2$ can't exceed 1.00

67. If $u = \frac{3x - 19}{8}$ & $v = \frac{-18y + 63}{5}$ then

a. $r_{xy} = r_{uv}$ ~~b.~~ $r_{xy} = -r_{uv}$ c. Both d. None

$u = \left(\frac{3}{8}\right)x - \frac{19}{8}$, $v = \left(-\frac{18}{5}\right)y + \frac{63}{5}$

68.

u	80	90
v	63	61

then Find r_{uv}

$r_{uv} = -1.00$

69. If $\bar{x} = 60, \bar{y} = 100, \sigma_x = 10, \sigma_y = 25, r_{xy} = -0.60$

Find 1. Reg. line of y on x.

2. Reg. line of x on y.

① Reg. line of y on x

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

$y - 100 = -0.60 \times \frac{25}{10}(x - 60)$

$y - 100 = -1.50(x - 60)$

$y - 100 = -1.50x + 90$

$\boxed{y = 190 - 1.50x}$

② Reg. line of x on y

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

$x - 60 = -0.60 \times \frac{10}{25}(y - 100)$

$x - 60 = -0.24y + 24$

$\boxed{x = 84 - 0.24y}$

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

$= -0.60 \times \frac{25}{10}$

$= -1.50$

$b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y} = -0.60 \times \frac{10}{25}$

$= -0.24$

If Reg. line of y on x is written in the form of $y = a + bx$ then $b = b_{yx}$

If Reg. line of x on y is written in the form of $x = a + by$ then $b = b_{xy}$

70. If $\bar{x} = 90, \bar{y} = 200, \sigma_x = 10, \sigma_y = 40, r = 0.95$

Find 1. If $x = 80, y = ?$

2. If $y = 208, x = ?$

① Reg. line of y on x : $(y - 200) = 0.95 \times \frac{40}{10} \times (80 - 90)$

$\boxed{y = 162}$

② Reg. line of x on y : $(x - 90) = 0.95 \times \frac{10}{40} \times (208 - 200)$

$\boxed{x = 91.90}$

71. True / False

$r_{xy} = r_{yx}$ $\left(\begin{matrix}\text{corr. coeffi.} \\ \text{betn } x \,\&\, y\end{matrix}\right) = \left(\begin{matrix}\text{corr. coeffi.} \\ \text{betn } y \,\&\, x\end{matrix}\right)$	True
$b_{yx} = b_{xy}$	False
$r^2 = b_{yx} . b_{xy}$	True
r is G.M. of b_{yx} & b_{xy}	True
If r = 0 then $b_{yx}, b_{xy} = 0$	True
If cov(x,y) > 0 then r > 0	True
If cov(x,y) = 0 then r = 0	True
$b_{yx} = r . \frac{\sigma_y}{\sigma_x} = \left[\frac{cov(x,y)}{\text{variance of } x}\right] = \left[\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(x-\bar{x})^2}\right]$	True
$b_{xy} = r . \frac{\sigma_x}{\sigma_y} = \left[\frac{cov(x,y)}{\text{variance of } y}\right] = \left[\frac{\Sigma(x-\bar{x})(y-\bar{y})}{\Sigma(y-\bar{y})^2}\right]$	True

72. If $\bar{x} = 50, \bar{y} = 90, \sigma_x = 3, \sigma_y = 6, r = 0.75$

Find 1. Reg. line of y on x.

2. Reg. line of x on y.

3. Point of intersection of 2 Reg. lines.

→ ① Reg. line of y on x

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

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

$$y - 90 = 0.75 \times \frac{6}{3} \times (x - 50)$$

$$y - 90 = 1.50x - 75$$

$$\boxed{y = 15 + 1.50x} \text{ ------ ①}$$

② Reg. line of x on y

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

$$x - 50 = 0.75 \times \frac{3}{6} \times (y - 90)$$

$$x - 50 = 0.375(y - 90)$$

$$x - 50 = 0.375y - 33.75$$

$$\boxed{x = 16.25 + 0.375y} \text{ ---- ②}$$

By solving 2 Linear eq^ns ① & ② simultaneously we will get point of intersection of 2 lines

$$y = 15 + 1.50(16.25 + 0.375y)$$

$$y = 15 + 24.375 + 0.5625y$$

$$0.4375y = 39.375$$

$$\boxed{y = 90}$$

$$\therefore \ x = 16.25 + 0.375(90) = 50 \quad \boxed{\therefore x = 50}$$

∴ point of intersection of 2 reg. lines is (50, 90)

2 Regression Lines will always intersect at point $(\bar{x}, \bar{y})$

math $a^2+b^2=c^2$

73. If $\bar{x} = 85, \bar{y} = 20, \sigma_x = 10, \sigma_y = 2, r = -0.80$

Find 2 Reg. lines & point of intersection.

→

2 Regression Lines will always intersect at point $(\bar{x}, \bar{y})$

① Reg. line of y on x

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

$$y - 20 = -0.80 \times \frac{2}{10} \times (x - 85)$$

$$y - 20 = -0.16(x - 85)$$

$$y - 20 = -0.16x + 13.60$$

$$\boxed{y = 33.60 - 0.16x}$$

② Reg. Line of x on y

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

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

$$x - 85 = -0.80 \times \frac{10}{2} \times (y - 20)$$

$$x - 85 = -4(y - 20)$$

$$x - 85 = -4y + 80$$

$$\boxed{x = 165 - 4y}$$

③ Let's solve 2 reg. eqns simultaneously to find point of intersection of 2 lines

$$y = 33.60 - 0.16(165 - 4y)$$

$$y = 33.60 - 26.40 + 0.64y$$

$$0.36y = 7.20 \qquad \boxed{y = 20}$$

$$x = 165 - 4(20)$$

$$\boxed{x = 85}$$

point of intersection of 2 reg. Lines $\equiv (\bar{x}, \bar{y}) \equiv (85, 20)$

74. If $\bar{x} = 20$, $\bar{y} = 150$, $r = 0.70$, $\sigma_x = 5$, $\sigma_y = 20$.

Find b_{yx}, Reg. line of y on x, b_{xy}, Reg. line of x on y.

① $b_{yx} = r \cdot \dfrac{\sigma_y}{\sigma_x} = 0.70 \times \dfrac{20}{5} = 2.80$

② $b_{xy} = r \cdot \dfrac{\sigma_x}{\sigma_y} = 0.70 \times \dfrac{5}{20} = 0.175$

③ Reg. line of y on x

$y - 150 = 2.80(x - 20)$

$y - 150 = 2.80x - 56$

$\boxed{y = 94 + 2.80x}$

If Reg. line of y on x is written in the form of $y = a + bx$ then $b = b_{yx}$

④ Reg. line of x on y

$x - 20 = 0.175(y - 150)$

$x - 20 = 0.175y - 26.25$

$\boxed{x = -6.25 + 0.175y}$

If Reg. line of x on y is written in the form of $x = a + by$ then $b = b_{xy}$

If Reg. line of y on x is written in the form of y = a + bx then 'b' represents 'b_{yx}'

If Reg. line of x on y is written in the form of x = a + by then 'b' represents 'b_{xy}'

75. If Reg line of y on x is 8x + 13y = 28. Find b_{yx}

$$8x + 13y = 28$$
$$13y = 28 - 8x$$
$$y = \frac{28}{13} - \frac{8}{13}x \qquad \therefore\ b_{yx} = -\frac{8}{13}$$
$$b_{yx} = -0.6153846153 8$$

76. If Reg line of x on y is 18x - 55y = 20. Find b_{xy}

$$18x - 55y = 20$$
$$18x = 20 + 55y$$
$$x = \frac{20}{18} + \frac{55}{18}y$$
$$\therefore\ b_{xy} = \frac{55}{18} = 3.0555555$$

77. If Reg line of y on x is 3x - 2y - 30 + 55x - 28y = 100x - 20y + 28.

Find b_{yx}

$$3x - 2y - 30 + 55x - 28y = 100x - 20y + 28$$
$$-2y - 28y + 20y = 28 + 100x - 55x - 3x + 30$$
$$-10y = 58 + 42x$$
$$y = -\frac{58}{10} - \frac{42}{10}x$$
$$\therefore\ b_{yx} = -42/10 = -4.20$$

b_{yx}	b_{xy}	r
-1.3838	-0.2828	-0.62557065148
-0.13889127042	-5.8319	-0.90
0.07925282931	0.908125	0.268275
1.22675	0.58925	0.85021317179
-3.369825	-0.2081925	-0.83759912327

78. $b_{yx} = 20, b_{xy} = -0.003521$, Find 'r'

a. 0.26537 b. -0.26537 c. 0.07042 ~~d.~~ Wrong data

$r^2 = b_{yx} \cdot b_{xy} = 20 \times -0.003521$

$r^2 = -0.07042$

As r^2 can't be negative, it is a wrong data.

79. $b_{yx} = -1.8, b_{xy} = -0.38$, Find 'r'

$b_{yx} \cdot b_{xy} = -1.80 \times -0.38 = 0.684$

$\therefore r = -\sqrt{0.684} = -0.8270429512$

80. $r = -0.8136, b_{xy} = -2.8056, b_{yx} = ?$

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

$(-0.8136)^2 = b_{yx} \times -2.8056$

$\boxed{b_{yx} = -0.2359370402}$

81. $r = 0.228675, b_{yx} = 0.583175, b_{xy} = ?$

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

$(0.228675)^2 = 0.583175 \times b_{xy}$

$\therefore b_{xy} = 0.0896682052\text{9}$

82. $\frac{b_{yx} + b_{xy}}{2} \geq r$ .... True / False

$\left(\frac{b_{yx} + b_{xy}}{2}\right) \geq r$

$\left(\frac{b_{yx} + b_{xy}}{2}\right) \geq \sqrt{b_{yx} \cdot b_{xy}}$

$\therefore$ This statement is True.

(AM of b_{yx} & b_{xy}) $\geq$ (GM of b_{yx} & b_{xy})

83. If $b_{yx} = 1.53$, $b_{xy} = 0.80$. Find r

a. 1.1063 b. -1.1063 c. 1.2240 ~~d.~~ Wrong data

$b_{yx} \cdot b_{xy} = 1.53 \times 0.80 = 1.224$

which is impossible as max. value of r^2 is 1.00

84. $-1.00 \leq r \leq 1.00$

$0 \leq r^2 \leq 1.00$

$\boxed{0 \leq b_{yx} \cdot b_{xy} \leq 1.00}$

Note:

Product of 2 regression coefficients must lie between 0 & 1.00 including both limits.

Minimum value of r = -1.00

Maximum value of r = 1.00

Minimum value of r^2 = 0.00

Maximum value of r^2 = 1.00

85. If reg. line of x on y is $33x - 21y - 28 = 15x - 101y + 58$. Find b_{xy}

$33x - 21y - 28 = 15x - 101y + 58$

$33x - 15x = -101y + 58 + 28 + 21y$

$18x = 86 - 80y$

$x = \frac{86}{18} - \frac{80}{18}y$

$\therefore b_{xy} = -\frac{80}{18} = -\frac{40}{9}$

$b_{xy} = -4.4444444$

86. If $b_{yx} > 1$ then b_{xy} must be less than 1................ True / False

True $\quad 0 \leq (b_{yx} \times b_{xy}) \leq 1.00$

b_{yx} & b_{xy} both can't be greater than 1

87. If $b_{xy} < 1$ then b_{yx} must be greater than 1................ True / False

False

88. If Reg. line of y on x is y = 88 + 0.20x & Reg. line of x on y is

x = 100 + 1.20 y. Find $(\bar{x}, \bar{y})$

$y = 88 + 0.20x$ .... ①

$x = 100 + 1.20y$ .... ②

Let's solve 2 Reg. Equations simultaneously

$y = 88 + 0.20(100 + 1.20y)$

$y = 88 + 20 + 0.24y$

$0.76y = 108$

$\boxed{y = 142.105263157}$

$x = 100 + 1.20(142.105263157)$

$\boxed{x = 270.526315788}$

$\therefore$ point of intersection of 2 reg. lines

$= (\bar{x}, \bar{y}) = (270.5263, 142.1053)$

89. b_{yx} = Reg Coefficient of y on x.

$$= r \cdot \frac{\sigma_y}{\sigma_x}$$

$$= \frac{\text{Cov}(x,y)}{\sigma_x . \cancel{\sigma_y}} \times \frac{\cancel{\sigma_y}}{\sigma_x}$$

$$= \frac{\text{Cov}(x,y)}{\text{Variance of } x}$$

$$b_{yx} = \frac{\frac{\sum(x-\bar{x})(y-\bar{y})}{n}}{\frac{\sum(x-\bar{x})^2}{n}}$$

$$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$$

b_{xy} = Reg Coefficient of x on y.

$$= r \cdot \frac{\sigma_x}{\sigma_y}$$

$$= \frac{\text{Cov}(x,y)}{\cancel{\sigma_x} . \sigma_y} \times \frac{\cancel{\sigma_x}}{\sigma_y}$$

$$= \frac{\text{Cov}(x,y)}{\text{Variance of } y}$$

$$b_{xy} = \frac{\frac{\sum(x-\bar{x})(y-\bar{y})}{n}}{\frac{\sum(y-\bar{y})^2}{n}}$$

$$b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(y-\bar{y})^2}$$

$$\boxed{b_{yx} = r \cdot \frac{\sigma_y}{\sigma_x} = \frac{\text{Cov}(x,y)}{\text{Variance of } x}}$$

$$\boxed{b_{yx} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}}$$

$$\boxed{b_{xy} = r \cdot \frac{\sigma_x}{\sigma_y} = \frac{\text{Cov}(x,y)}{\text{Variance of } y}}$$

$$\boxed{b_{xy} = \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(y-\bar{y})^2}}$$

math π $a^2+b^2=c^2$

90. If r = 0 then cov (x,y) = ?

~~a.~~ 0 b. 1 c. Between 0 & 1 d. Can't say

$$\boxed{\begin{array}{lll} \text{If } r = 0 & \text{then} & cov(x,y) = 0 \\ \text{If } 0 < r \le 1.00 & \text{then} & cov(x,y) > 0 \\ \text{If } -1 \le r < 0 & \text{then} & cov(x,y) < 0 \end{array}}$$

91. If covariance of (x,y) = 100

Variance of x = 8000

Variance of y = 500

Find 'r'

$$r = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y} = \frac{100}{\sqrt{8000} \times \sqrt{500}} = 0.05$$

92. $\Sigma(x-\bar{x})(y-\bar{y}) = k$

$\Sigma(x-\bar{x})^2 = m$

$\Sigma(y-\bar{y})^2 = j$

Find 'r'

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

$$r = \frac{k}{\sqrt{m} \times \sqrt{j}} = \frac{k}{\sqrt{mj}}$$

KARLPEARSON'S COEFFICIENT

$$= \frac{\Sigma X.Y}{\sqrt{\Sigma X^2 \cdot \Sigma Y^2}}$$

Where,

$X = (x-\bar{x})$

$Y = (y-\bar{y})$

93. Variance of data can be Zero OR positive	SD of data can be Zero OR positive	cov(x,y) of data can be Zero OR positive OR Negative

a. Zero

b. Positive

c. Negative

94. If cov (x,y) = 50, σ_x = 10 then

~~a.~~ $\sigma_y \geq 5$

b. $\sigma_y \leq 5$

c. $\sigma_y < 5$

d. Can't say

$$r = \frac{cov(x,y)}{SD_x \cdot SD_y} = \left(\frac{50}{10 \times SD_y}\right)$$

$$\therefore SD_y \geq 5.00$$

95. If scatter diagram is showing a line | | to Y-axis then there is ________ correlation

a. Positive

b. Negative

c. Spurious

~~d.~~ No

96. If scatter diagram is showing points without depicting any pattern then there is No correlation

97. Process of establishing relation / association between 2 or more variable is known as correlation analysis and process of ascertaining value of one variable on the basis of other is known as Regression analysis.

98. x = Rainfall in city $r_{xy} = 0.60$

y = No. of tress in city

Coefficient of Determination = $r^2 = 0.60^2$
(Explained Variance) $= 0.36 = \boxed{36\%}$

Coefficient of Non-determination = $1 - r^2 = 1 - 0.60^2$
(Unexplained Variance) $= 1 - 0.36 = 0.64 = \boxed{64\%}$

99.

a. Coefficient of Determination $= \left(\frac{\text{Explained Variance}}{\text{Total Variance}}\right) = r^2$

b. Coefficient of Non-determination $= \left(\frac{\text{Un-explained Variance}}{\text{Total Variance}}\right) = 1 - r^2$

= 1.00

(Coefficient of Determination + Coefficient of Non-determination)

100. If r = 0.85. Find

Coefficient of Determination = $r^2 = 0.85^2 = 0.7225$

Coefficient of Non-determination = $1 - r^2 = 1 - 0.7225$
$= 0.2775$

Explained Variance = 72.25% = (Explained vari / Total vari)

Unexplained Variance = 27.75%
= (unexplained vari / Total variance)

101. a. 2 Regression lines will coincide when

There is perfect positive correlation OR
perfect Negative correlation

b. Two Regression lines will coincide when there is $r = 1$ OR
$r = -1$

c. If r = 0 then Two Regression lines are generally $\perp$ to each other

d. When 2 regression lines are $\perp$ to each other then there is No correlation

102. If c = 5, m = 11. Find coefficient of concurrent deviation.

➡ c = No. of concurrent deviations
= No. of '+' signs in product column

$m = n - 1 = 11$

positive signs = 5   Negative signs = 6

$r = -\sqrt{-\frac{2c-m}{m}} = -\sqrt{-\left(\frac{2(5)-11}{11}\right)} = -\sqrt{1/11} = -0.301511$

103. If c = 8, m = 12. Find coefficient of concurrent deviation.

$$r = \pm\sqrt{\pm\frac{2c-m}{m}}$$

$$r = \sqrt{\frac{2(8)-12}{12}}$$

$$r = \sqrt{\frac{1}{3}} = 0.57735$$

104.

	Maths (x)	Physics (y)
AM	85	92
SD	8	11
r = 0.89		

What expected score of maths if a student scored 90 marks in physics?

$x = ?$, $y = 90$

Let's use Reg. line of x on y

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

$$x - 85 = 0.89 \times \frac{8}{11} \times (90 - 92)$$

$$\boxed{x = 83.70545}$$

105. Marks of maths, stat of 25 students

(5,16), (18,19), (6,8), (7,13), (2,18)
(12,13), (5,7), (2,8), (19,2), (16,5)
(13,18), (12,15), (6,8), (9,5), (6,12)
(13,15), (7,12), (13,6), (2,12), (13,3)
(8,5), (9,6), (12,5), (3,16), (5,8)

Prepare Bi-variate Frequency Distribution Table

→

Bi-variate Frequency Distribution Table

Marks of Stat (y) / Marks of Maths (y)	0 - 5	5 - 10	10 - 15	15 - 20	Total
0 - 5	= 0	| = 1	| = 1	|| = 2	4
5 - 10	= 0	$\cancel{||||}$ || = 7	||| = 3	| = 1	11
10 - 15	| = 1	|| = 2	| = 1	||| = 3	7
15 - 20	| = 1	| = 1	= 0	| = 1	3
Total	2	11	5	7	25

a. Find Marginal Distribution of x-marks of Maths

x	0 - 5	5 - 10	10 - 15	15 - 20
f	4	11	7	3

b. Find Marginal Distribution of y-marks of Stats

y	0 - 5	5 - 10	10 - 15	15 - 20
f	2	11	5	7

c. Find Conditional Distribution of y when x is 5 - 10.

y	0 - 5	5 - 10	10 - 15	15 - 20
f	0	7	3	1

d. Find Conditional Distribution of x when y is 15 - 20.

x	0 - 5	5 - 10	10 - 15	15 - 20
f	2	1	3	1 -

106.

y / x	10 - 30	30 - 90	90 - 110	Total
10 - 20	53	28	98	179
20 - 30	113	167	283	563
30 - 40	669	813	822	2304
40 - 50	1083	786	555	2424
Total	1918	1794	1758	547

a. Find Marginal Distribution of x

x	10-20	20-30	30-40	40-50
f	179	563	2304	2424

b. Find Marginal Distribution of y

y	10-30	30-90	90-110
f	1918	1794	1758

c. Find Conditional Distribution of x when y is 90 - 110.

x	10-20	20-30	30-40	40-50
f	98	283	822	555

d. Find Conditional Distribution of y when x is 30 - 40.

y	10-30	30-90	90-110
f	669	813	822

107. $r = 0.50, \sum d^2 = 82.50, n = ?$

r = spearman's rank correlation coefficient $= 1 - \frac{6\sum d^2}{n(n^2-1)}$

$0.50 = 1 - \frac{6 \times 82.50}{n(n^2-1)}$

$\left[\frac{495}{n(n^2-1)}\right] = 1 - 0.50 = 0.50$

$\therefore n(n^2-1) = \frac{495}{0.50} = 990$

$n(n^2-1) = 10 \times (10^2-1)$

$\boxed{\therefore n = 10}$

$u = 5 + 3x,\ v = 19 + 8y$

$\Rightarrow r_{uv} = r_{xy}$

$\sigma_u = 3 \times \sigma_x$

$\sigma_v = 8 \times \sigma_y$

$b_{vu} = r_{uv} \times \frac{\sigma_v}{\sigma_u}$

$= r_{xy} \times \frac{\sigma_y \times 8}{\sigma_x \times 3}$

$b_{vu} = \frac{8}{3} \times b_{yx}$

108. **Regression coefficients are affected by Change in Scale but Not affected by Change in Origin**

109. Theory of Regression has been derived from

Method of Least Squares

110. There are some cases when we find correlation between 2 variables although 2 variables are not casually related. This is due to existence of Third variable which is related to both the variables under consideration, such type of correlation is known as

Spurious correlation OR Non-sense correlation

111. Regression coefficient are unaffected due to ____________

~~a.~~ Shift of origin b. Change in scale

c. Both d. None

112. Lungs damage & cigarette smoking, ____________ correlation may exist.

~~a.~~ Positive b. Negative c. Zero d. None

113. b_{yx} and b_{xy} are always same. True / False

→ False

114. r_{xy} and r_{yx} are always same. True / False

→ True. (corr. coeffi. between x & y) = (corr. coeffi. betn y & x)

115. Co-variance measures __Joint__ variation between 2 variables.

a. Joint b. Common c. Relative d. None

$u = 8 - 15x$, $v = -63y + 21$ then

$\Rightarrow$

① $r_{uv} = r_{xy}$

② $b_{vu} = r_{uv} \cdot \dfrac{\sigma_v}{\sigma_u} = r_{xy} \times \dfrac{63 \times \sigma_y}{15 \times \sigma_x} = \left[\dfrac{63}{15} \times r_{xy} \times \dfrac{\sigma_y}{\sigma_x}\right]$

$\therefore b_{vu} = \dfrac{63}{15} \times b_{yx}$

③ $b_{uv} = \dfrac{15}{63} \times b_{xy}$

116. 1. If $u = 3 + x$
$v = -8 + y$ then

~~a.~~ $r_{xy} = r_{uv}$ b. $r_{xy} = -r_{uv}$

c. $r_{xy} = \frac{1}{r_{uv}}$ d. $\frac{r_{xy}}{r_{uv}} = 2$

2. $u = 18 - 3x$
$v = -63 + 55y$
then
$r_{uv} = -r_{xy}$

$b_{vu} = -\frac{55}{3} \times b_{yx}$

$b_{uv} = \frac{3}{-55} \times b_{xy}$

117. If $u = 13x + 9$
$v = -18y + 33$ then
Find relation between (b_{uv} & b_{xy}) and (b_{vu} & b_{yx}) & (r_{uv} & r_{xy})

→ (1) $r_{uv} = -r_{xy}$

(2) $b_{vu} = r_{uv} \times \frac{\sigma_v}{\sigma_u} = -r_{xy} \times \frac{18 \times \sigma_y}{13 \times \sigma_x}$

$b_{vu} = -\frac{18}{13} \times \left(r_{xy} \times \frac{\sigma_y}{\sigma_x}\right) = -\frac{18}{13} \times b_{yx}$

(3) $b_{uv} = -\frac{13}{18} b_{xy}$

118. If $u = 15x + 93$
$v = -18 + 61y$
then
$r_{uv} = r_{xy}$
$b_{uv} = b_{xy} \times \frac{15}{61}$
$b_{vu} = b_{yx} \times \frac{61}{15}$

119. If $u = 59 - 68x$
$v = -51y + 33$
then
$r_{uv} = r_{xy}$
$b_{uv} = b_{xy} \times \frac{-68}{-51}$
$= \frac{4}{3} \times b_{xy}$
$b_{vu} = \frac{3}{4} \times b_{yx}$

math

120. $u = \frac{13 + 5x}{19}$ $v = \frac{-16 + 22y}{101}$

then $u = \frac{13}{19} + \frac{5}{19}x$ $v = \frac{-16}{101} + \frac{22}{101}y$

$b_{uv} = b_{xy} \times \frac{5/19}{22/101} = \frac{505}{418} \times b_{xy}$

$b_{vu} = b_{yx} \times \frac{22/101}{5/19} = \frac{418}{505} \times b_{yx}$

$r_{uv} = r_{xy}$

121. Spearman's rank correlation coefficient = 0.80 for 10 pairs of observations.

Later on it is observed that one value of 'd' was taken as 7 instead of 6. Find correct 'r'.

$$r = 1 - \left[\frac{6\sum d^2}{n(n^2-1)}\right]$$

$$0.80 = 1 - \frac{6\sum d^2}{10 \times 99}$$

$$\frac{6\sum d^2}{990} = 1 - 0.80 = 0.20$$

wrong $\sum d^2 = 33$

correct $\sum d^2 = 33 - 7^2 + 6^2$

$= 33 - 49 + 36$

$= 20$

$$r = 1 - \frac{6 \times 20}{990}$$

$$r = 1 - \frac{120}{990}$$

$$\boxed{r = 0.87878787}$$

122.

x	y
↑	↑
↓	↓

Type of Correlation : positive

x	y
↓	↑
↑	↓

Type of Correlation : Negative

123. **(Product of 2 regression coefficients) = (Correlation coefficient)^2**

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

$\therefore \; r = \sqrt{b_{yx} \cdot b_{xy}}$

sign of b_{yx}, b_{xy} & r will always be same

124. n = No. of pairs of observations in given data (sample)

N = Population size

As we draw conclusion on the basis of sample & not on the basis of population there can be some error in our judgement

Probable Error

$= 0.674 \times$ standard error

$= 0.674 \times \left(\frac{1-r^2}{\sqrt{N}}\right)$

Standard Error

$= \left(\frac{1-r^2}{\sqrt{N}}\right)$

N = population size

125. 1. **Marginal Distribution** is the frequency distribution of one variable (x or y) across the other variable's full range of values.

2. **Conditional Distribution** is the frequency distribution of one variable across the particular sub-population of other variable.

126. Bi-variate data is collected for :

a. 2 variables

b. 3 or more variables

~~c.~~ 2 variables at same point of time

d. 2 variables at different point of time

127. The diff between actual value and estimated value is known as

Error or Residue

128. If r = 0.40 then

% of known or accounted variation = $0.40^2 = 0.16 = 16\%$

% of unknown or un-accounted variation = $1 - 0.40^2 = 1 - 0.16$
$= 84\%$

Coefficient of determination = $r^2 = 0.16$

Coefficient of non-determination = $1 - r^2 = 0.84$

129.

Variables	Nature of Correlation
1. No. of claims & profit of Insurance company	Negative
2. Demand for Giffen goods & Price of Giffen goods.	Positive
3. Sale of Woolen garments & temp.	Negative
4. Marketing expenses & Turnover	positive
5. No. of trees & Rainfall	positive
6. Cigarette Smoking & Lungs damage	positive
7. Rainfall & Crop Yield	positive
8. Years of education & Income	positive
9. Temp & Sale of Tea, Coffee	Negative
10. No. of hrs on social media, marks in exam.	Negative

130. For the Bi-variate data
(20,5), (21,4), (22,3) r = ?

a. 1.00 ~~b.~~ -1.00 c. Can't say d. r = 2.00

→

x	y	xy	x^2	y^2
20	5			
21	4			
22	3			
		250	1325	50

$$r = \frac{\frac{250}{3} - (21 \times 4)}{\sqrt{\frac{1325}{3} - 21^2} \times \sqrt{\frac{50}{3} - 4^2}}$$

$$r = \frac{-0.666666}{\sqrt{0.666666 \times 0.666666}} = -1.00$$

131. Simple correlation is known as

→ Linear correlation

132. Slope of Reg. line of y on x is :

→ Reg. line of y on x is

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

$y - \bar{y} = b_{yx} \cdot x - b_{yx} \cdot \bar{x}$

$b_{yx} \cdot \bar{x} - \bar{y} = b_{yx} \cdot x - y$

$b_{yx} \cdot x - y = b_{yx} \cdot \bar{x} - \bar{y}$

slope of the line $= -\frac{b_{yx}}{-1} = b_{yx}$

$$\boxed{\text{slope of the line } ax + by + c = 0 \text{ is } -a/b}$$

133. Slope of Reg. line of x on y is :

→ Reg. line of x on y is

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

$x - \bar{x} = b_{xy} \cdot y - b_{xy} \cdot \bar{y}$

$x - b_{xy} \cdot y = \bar{x} - b_{xy} \cdot \bar{y}$

Slope of this Line $= -1/-b_{xy} = (1/b_{xy})$

134. The Best method to measure correlation is ____________

→ Karl perason's product moment corr. coefficient

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

135. 2 regression lines

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

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

intersect at point $(\bar{x}, \bar{y})$

136. 2 Reg. lines are $2x - 3y = 10$ & $5x - 8y = 33$. Find b_{yx} & b_{xy} & r.

→ Trial & error method

x on y

$2x - 3y = 10$

$2x = 10 + 3y$

$x = \frac{10}{2} + \frac{3}{2}y$

$$\boxed{b_{xy} = \frac{3}{2} = 1.50}$$

y on x

$5x - 8y = 33$

$-8y = 33 - 5x$

$y = -\frac{33}{8} + \frac{5}{8}x$

$$\boxed{\therefore b_{yx} = \frac{5}{8} = 0.6250}$$

$r^2 = b_{yx} \times b_{xy} = 0.6250 \times 1.50 = 0.9375$

$$\boxed{\therefore r = 0.9682458365}$$

NEVER DREAM ABOUT SUCCESS,
WORK FOR IT !

① Karl pearson's coefficient (r)

$$r = \left[\frac{\text{Covariance of } (x, y)}{SD_x \cdot SD_y}\right]$$

$$r = \left[\frac{\frac{\sum xy}{n} - \left(\frac{\sum x}{n} \times \frac{\sum y}{n}\right)}{\sqrt{\frac{\sum x^2}{n} - \bar{x}^2} \times \sqrt{\frac{\sum y^2}{n} - \bar{y}^2}}\right]$$

$$r = \left[\frac{\frac{\sum\left[(x-\bar{x})(y-\bar{y})\right]}{n}}{\frac{\sqrt{\sum(x-\bar{x})^2} \times \sqrt{\sum(y-\bar{y})^2}}{n}}\right]$$

$$r = \left[\frac{\sum\left[(x-\bar{x})(y-\bar{y})\right]}{\sqrt{\sum(x-\bar{x})^2} \times \sqrt{\sum(y-\bar{y})^2}}\right] = \left[\frac{\sum X \cdot Y}{\sqrt{\sum X^2} \cdot \sqrt{\sum Y^2}}\right]$$

where $X = x - \bar{x}$
$Y = y - \bar{y}$

math

(2)

x	y
80	35
90	25
55	60
75	100

$\bar{x} = 75$

$\bar{y} = 55$

$$\text{cov}(x,y) = \frac{\sum xy}{n} - \bar{x}\cdot\bar{y}$$

$$= \frac{15{,}850}{4} - 75 \times 55$$

$$= -162.50$$

$$SD_x = \sqrt{\frac{23150}{4} - 75^2}$$

$$= \boxed{12.7475487839}$$

$$SD_y = \sqrt{\frac{15{,}450}{4} - 55^2}$$

$$= 28.9395922569$$

$$r_{xy} = \frac{-162.50}{12.7475487839 \times 28.939592256}$$

$$= -0.44048819592$$

r_{xy} = (corr. coeffi. betn x & y)

$u = 2x+10$	$V = 5y-20$
170	155
190	105
120	280
160	480

$u = 2x + 10$

$\bar{u} = 2\bar{x} + 10 = 2 \times 75 + 10$

$= 160$

$V = 5y - 20$

$\therefore \bar{V} = 5\bar{y} - 20 = 5(55) - 20 = 255$

$SD_u = 12.7475487839 \times 2$

$= 25.495097 5678$

$SD_v = 28.9395922569 \times 5$

$= 144.697961284$

$$\text{cov}(u,v) = \frac{\sum uv}{n} - \bar{u}\cdot\bar{v}$$

$$= \text{cov}(x,y) \times 2 \times 5$$

$$= \left(\frac{1{,}56{,}700}{4} - 160 \times 255\right)$$

$$= -1625$$

$$r_{uv} = \frac{-1625}{(25.4950975678 \times 144.697961284)}$$

$$r_{uv} = -0.44048819592$$

$\therefore$ corr. coeffi. (r) is unaffected by change of origin as well as by change in scale

(3) If $u = 10x - 25$
$v = 16y + 300$

then $SD_u = 10 \times SD_x$

$SD_v = 16 \times SD_y$

$cov(u,v) = 10 \times 16 \times cov(x,y)$

$$r_{uv} = \left[\frac{cov(u,v)}{SD_u \times SD_v}\right] = \frac{\cancel{10} \times \cancel{16} \times cov(x,y)}{\cancel{10} \times SD_x \times \cancel{16} \times SD_y} = \left[\frac{cov(x,y)}{SD_x \cdot SD_y}\right]$$

$r_{uv} = r_{xy}$

$\therefore$ 'r' is not affected by change of origin as well as by change in scale

(4)

x	y
10	25
30	65

$u = 2x + 20$	$v = -3y + 300$
40	225
80	105

$cov(x,y) = \frac{2200}{2} - 20 \times 45$
$= 200$

$cov(u,v) = \frac{17,400}{2} - 60 \times 165$
$= -1200$

$r_{xy} = \frac{200}{10 \times 20} = 1.00$

$r_{uv} = \frac{-1200}{20 \times 60} = -1.00$

Here $r_{xy} = -r_{uv}$
OR $r_{uv} = -r_{xy}$

(5) $u = 18 - 33x$ $\quad v = -26y + 300$

$r_{xy} = -0.5625$ Find r_{uv}

$\Longrightarrow r_{uv} = r_{xy} = -0.5625$

(6) $u = 22 + 15x$, $\quad v = 600 - 3y$

$r_{xy} = 0.8691$, $r_{uv} = ?$

$\Longrightarrow r_{uv} = -r_{xy} = -0.8691$

(7) $u = -19 + 12x$ $\quad v = -2y + 930$

$r_{xy} = 0.2293$ $\quad r_{uv} = ?$

$\Longrightarrow r_{uv} = -0.2293$