1 # Correlation Regression.

Date 07/06

If x & y are two variables related in such a way that an increasing one is accompanied by increase or decrease in a other such a relationship is called Correlation or Co-variation.

Note: * If x & y increases or decreases togeather then x & y are positively correlated.

* If x increases as y decreases or vice-versa then x & y are ~~the~~ negatively correlated.
* The correlation is measured by Co-efficient of Correlation.
* The Co-efficient of Correlation 'r' between x & y is given by

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n \sigma_x \sigma_y}$$

$$= \frac{\sum X_i Y_i}{n \sigma_x \sigma_y} \quad \text{where } X_i = x_i - \bar{x}, \; Y_i = y_i - \bar{y}$$

furthermore $r = \dfrac{\sigma_{xy}}{\sigma_x \sigma_y}$ where $\sigma_{xy} = \dfrac{\sum X_i Y_i}{n}$

σ_{xy} is called co-variance.

Also, we know that

$$\sigma_x^2 = \frac{1}{n} \sum (x_i - \bar{x})^2 = \frac{1}{n} \sum X_i^2$$

$$\sigma_y^2 = \frac{1}{n} \sum (y_i - \bar{y})^2 = \frac{1}{n} \sum Y_i^2$$

r can also be written as.

$$r = \frac{\sum X_i Y_i}{\sqrt{\sum X_i^2 \sum Y_i^2}}$$

$$= \frac{\sum X_i Y_i}{n \sqrt{\frac{1}{n} \sum X_i^2 \cdot \frac{1}{n} \sum Y_i^2}}$$

Note:- i) $-1 \le r \le 1$

2) If ~~x &~~ $r = +1$ or -1 then x & y are perfectly correlated.

3) If $r = 0$ x & y are not correlated.

3) Co-efficient r is also referred as karl Pearson's Coefficient of correlation

* Problems:

1) The following tables gives the edges of 10 Couples, Calculate the Co-variance & the Co-efficient of Co-relation between these ages.

Age of husband	23	27	28	29	30	31	33	35	36	39
Age of wife (y)	18	22	23	24	25	26	28	29	30	32

→ $\sigma_{xy} = ?$ $r = ?$

$$\sigma_{xy} = \frac{\Sigma X_i Y_i}{n} \qquad \frac{\Sigma x_i}{n} = 31.1$$

$$= \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{n}, \qquad \frac{\Sigma y_i}{n} = 25.7$$

x_i	y_i	$X_i = x_i - \bar{x}$	$Y_i = y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y}) = X_i Y_i$	X_i^2	Y_i^2
23	18	-8.1	-7.7	62.37	65.61	59.29
27	22	-4.1	-3.7	151.17	16.81	13.69
28	23	-3.1	-2.7	8.37	9.61	7.29
29	24	-2.1	-1.7	3.57	4.41	2.89
30	25	-1.1	-0.7	0.77	1.21	0.44
31	26	-0.1	+0.3	-0.03	0.01	0.09
33	28	1.9	2.3	4.37	3.61	5.24
35	29	3.9	3.3	12.87	15.21	10.89
36	30	4.9	4.3	21.07	24.01	18.49
39	32	7.9	6.3	49.77	62.41	39.69

Co-variance: $\sigma_{xy} = \frac{\Sigma X_i Y_i}{n} = \frac{178.30}{10} = 17.83.$

ii) $r = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ $\quad \sigma^2 x = \frac{1}{n} \Sigma X_i^2 \quad \sigma_x = \sqrt{\frac{1}{n} \Sigma X_i^2}$

$$\sigma_x = \sqrt{\frac{1}{10} \times 202.9} = \sqrt{20.29} = 4.504.$$

$$\sigma_y = \sqrt{\frac{1}{10} \times 158.1} = 3.9761$$

$$r = \frac{17.83}{4.504 \times 3.9761} = \underline{\underline{0.995}}$$

$\bar{x} = 31.1$
$\bar{y} = 25.7$
$\sigma_x = 4.5044$
$\sigma_y = 3.9761$
$r = 0.9955$

2) Psychological test of intelligence & Computational ability were applied to adults, following is the record showing intelligence ratio IR & Ability ratio AR Calculate the Coefficient Co-relation

$r = 0.5177$

IR (x) :	105	104	102	101	100	99	98	96	95	94
AR (y)	101	103	100	98	95	96	104	97	97	96

x_i	y_i	$X_i = x_i - \bar{x}$	$Y_i = y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	X_i^2	Y_i^2
105	101	5.6	2.3	12.88	31.36	5.29
104	103	4.6	4.3	19.78	21.16	18.49
102	100	2.6	1.3	3.38	6.76	1.69
101	98	1.6	-0.7	-1.12	2.56	0.49
100	95	0.6	-3.7	-2.22	0.36	13.69
99	96	-0.4	-2.7	1.08	0.16	7.29
98	104	-1.4	5.3	-7.42	1.96	28.09
96	97	-3.4	-1.7	5.78	11.56	2.89
95	97	-4.4	-1.7	7.48	19.36	2.89
94	96	-5.4	-2.7	14.58	29.16	7.29

$\bar{x} = 58.57$

$\bar{y} = 40.28$

Find the co-efficient of co-relation between industrial production & export using following data.

Production	55	56	58	59	60	60	62
Export	35	38	38	39	44	43	45

x_i	y_i	$X_i = x_i - \bar{x}$	$Y_i = y_i - \bar{y}$	$X_i Y_i$	X_i^2	Y_i^2
55	35	-3.57	-5.28	18.84	12.74	27.87
56	38	-2.57	-2.28	5.85	6.60	5.19
58	38	-0.57	-2.28	1.29	0.32	5.19
59	39	0.43	-1.28	-0.55	0.18	1.63
60	44	1.43	3.72	5.31	2.04	13.83
60	43	1.43	2.72	3.88	2.04	7.39
62	45	3.43	4.72	16.18	11.76	22.27
				$\Sigma X_iY_i = 50.8$	~~$\Sigma X_i^2 = 5.09$~~ ΣX_i^2 35.63	$\Sigma Y_i^2 =$ 83.37

$$r = \frac{\Sigma X_i Y_i}{\sqrt{\Sigma X_i^2 \times \Sigma Y_i^2}}$$

$$= \frac{50.85}{\sqrt{35.63 \times 83.37}} = \underline{\underline{0.93}}$$

If $z = ax + by$ & r is correlation co-efficient between x & y show that

$$r = \frac{\sigma^2 z - (a^2\sigma^2 x + b^2\sigma^2 y)}{2ab\,\sigma x\,\sigma y}$$

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

$$\sigma^2 y = \frac{\Sigma Y_i^2}{n} = \frac{\Sigma (y_i - \bar{y})^2}{n}$$

$$\sigma^2 z = \frac{\Sigma Z^2}{n} = \frac{\Sigma (z_i - \bar{z})^2}{n}$$

$$r = \frac{\Sigma X_i Y_i}{n\,\sigma x\,\sigma y} = \frac{\Sigma (x_i - \bar{x})(y_i - \bar{y})}{n\,\sigma x\,\sigma y}$$

Given: $z = ax + by$

$\bar{z} = a\bar{x} + b\bar{y}$

$z_i = ax_i + by_i$

$(z_i - \bar{z}) = a(x_i - \bar{x}) + b(y_i - \bar{y})$

Squaring these,

$$(z_i - \bar{z})^2 = a^2(x_i - \bar{x})^2 + b^2(y_i - \bar{y})^2 + 2ab(x_i - \bar{x})(y_i - \bar{y})$$

By summing these.

$$\Sigma (z_i - \bar{z})^2 = a^2 \Sigma (x_i - \bar{x})^2 + b^2 \Sigma (y_i - \bar{y})^2 + 2ab + \Sigma (x_i - \bar{x})(y_i - \bar{y})$$

$$n\sigma^2 z = a^2 n\sigma^2 x + b^2 n\sigma^2 y + 2ab\,n\,r\,\sigma x\,\sigma y$$

$$\sigma^2 z = a^2\sigma^2 x + b^2\sigma^2 y + 2ab\,r\,\sigma x\,\sigma y$$

$$\frac{\sigma^2 z - (a^2\sigma^2 x + b^2\sigma^2 y)}{2ab\,\sigma x\,\sigma y} = r$$

$$\Rightarrow \quad r = \frac{\sigma^2 z - (a^2\sigma^2 x + b^2\sigma^2 y)}{2ab\,\sigma x\,\sigma y}$$

Note:

1) if $z = ax + by$ & if $a = 1, b = 1$, $r = \frac{\sigma^2 x{+}y - (\sigma^2 x + \sigma^2 y)}{2\sigma x\,\sigma y}$

2) if $a = 1, b = -1$; $r = \frac{\sigma^2 x{-}y - (\sigma^2 x - \sigma^2 y)}{-2\sigma x\,\sigma y}$

If std. deviation of x & y are 2 & 3 respectively & the if the co-relation co-efficient between x & y is 0.4. Find the std. deviatn of $x+y$ & $x-y$.

→ Given: $\sigma x = 2$, $\sigma y = 3$, $r = 0.4$.

$\sigma x{+}y = ?$ $\quad \sigma x{-}y = ?$

$z = ax + by$.

a) $a=1$, $b=1$

$$\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y + 2r\sigma_x\sigma_y$$
$$= 4+9+2\times 0.9\times 2\times 3$$

$\sigma^2_{x+y} = 17.8$ $\quad\quad \sigma_{x+y} = 4.21$

b) $a=1$, $b=-1$

$$\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y - 2\times r\sigma_x\times\sigma_y$$
$$= 4+9-2\times 0.4\times 2\times 3$$

$\sigma^2_{x-y} = 8.2$ $\quad\quad \sigma_{x-y} = 2.86$

2) If the variables x & y are such that
1) $x+y$ has variance 15
2) $x-y$ has variance 11
3) $2x+y$ has " 29

find σ_x, σ_y & r

→
$$r = \frac{\sigma^2_z - (a^2\sigma^2_x + b^2\sigma^2_y)}{2ab\sigma_x\sigma_y}$$

1) $a=1$ $b=1$ $\quad\quad$ 2) $a=1$ $b=-1$

$\sigma^2_{x+y} = 15$ $\quad\quad$ $\sigma^2_{x-y} = 11$

3) $a=2$ $b=1$

$\sigma^2_{2x+y} = 29$

$$\sigma^2_{x+y} = \sigma^2_x + \sigma^2_y + 2\times r\sigma_x\sigma_y = 15 \quad -(1)$$
$$\sigma^2_{x-y} = \sigma^2_x + \sigma^2_y - 2r\sigma_x\sigma_y = 11 \quad -(2)$$
$$\cancel{\sigma^2_{2x+y} = \sigma^2_{2x} + \sigma^2_y + 4\times r\sigma_x\sigma_y = 29 \quad -(3)}$$
$$\sigma^2_{2x+y} = 4\sigma^2_x + \sigma^2_y + 4r\sigma_x\sigma_y = 29 \quad -(3)$$

(1) + (2) $\quad\quad\quad\quad$ (1) − (2)

$2\sigma^2_x + 2\sigma^2_y = 26$ $\quad\quad$ $4r\sigma_x\sigma_y = 4$ $\quad -(a)$

$\sigma^2_x + \sigma^2_y = 26/2$ $\quad\quad$ $r\sigma_x\sigma_y = 1$

$\sigma^2_x + \sigma^2_y = 13$ $\quad -(4)$

sub (a) in (3)

$$4\sigma^2_x + \sigma^2_y + 4 = 29$$
$$4\sigma^2_x + \sigma^2_y = 25 \quad -(5)$$

(5) − (4), $3\sigma^2_x = 25 - 13 = 12$

$$\sigma^2_x = 12/3$$
$$\sigma^2_x = 4$$
$$\boxed{\sigma_x = 2}$$

Sub in (4)

$$4 + \sigma^2_y = 13$$
$$\sigma^2_y = 13-4$$
$$\sigma^2_y = 9$$
$$\boxed{\sigma_y = 3}$$

put σ_x & σ_y in eq^n (a)

$$\cancel{4}r\times 2\times 3 = \cancel{4}$$
$$r = \frac{1}{6} = 0.1666 = 0.167$$

✱ Regression Analysis

It is the statistical tool which is employed for the purpose of making estimates.

In this analysis there are 2 types of variables. the variable whose value is influenced is called dependent variable & the variable which influences is called independent variable.

Suppose n pairs of values (x_1, y_1) (x_2, y_2) .... (x_n, y_n) are given. a straight line is fit to this data taking x as independent & y as dependent variable. The straight line obtained is called Line of Regression of y on x. It's slope is called.

Regression Coefficient of y on x: If the straight line fit by taking y as independent & x as dependent thus line obtained is Line of Regression of x on y. & the reciprocal of slope is called as regression coefficient of x on y.

* Equations of Regression Lines.

$$(y-\bar{y}) = r\frac{\sigma y}{\sigma x}(x-\bar{x}) \qquad \left\{ b_{yx} = r\frac{\sigma y}{\sigma x} \right.$$

This eq^n is line of regression of y on x & its slope $\left(r\frac{\sigma y}{\sigma x}\right)$ which is denoted by b_{yx} is called regression coefficient of y on x.

* $$(x-\bar{x}) = r\frac{\sigma x}{\sigma y}(y-\bar{y})$$

This eq^n is line of regression of x on y & its slope is $\frac{1}{r}\frac{\sigma y}{\sigma x}$ & hence reciprocal of this slope denoted by $b_{xy} = r\frac{\sigma x}{\sigma y}$ is regression coefficient of x on y.

*. Remarks:

1) Since σx & σy are positive $r, \sigma x, \sigma y$ have same signs.
2) $|r| = \sqrt{b_{yx}\, b_{xy}}$
3) Lines of Regression always passes through $(\bar{x}, \bar{y})$

$r = 0.95 \qquad \bar{x} = 5$
$\sigma x = 2.58 \qquad \bar{y} = 12$
$\sigma y = 2.58$.

1) Calculate the coefficient of correlation & obtain lines of Regression for the following data.

x:	1	2	3	4	5	6	7	8	9
y:	9	8	10	12	11	13	14	16	15

Obtain an estimate of y which corresponds to $x = 6.2$

→ i) $r = ?$ ii) $(y-\bar{y}) = b_{yx}(x-\bar{x})$ iii) $y = ?$ $x = 6.2$
$(y-\bar{y}) = r\left(\frac{\sigma y}{\sigma x}\right)(x-\bar{x})$

i) $r = 0.95$

ii) $(y-12) = 0.95 \times \left(\frac{2.58}{2.58}\right)(x-5)$.

iii) $(y-12) = 0.95 \times (6.2-5) = 1.14$
$y = 1.14 + 12$
$y = \underline{13.14}$.

ii) $y - 12 = (x-5) \times 0.95 \qquad y = 0.95x + 7.25$.
$x - 5 = (y-12) \times 0.95 \qquad x = 0.95y - 6.4$.

Short-cut for iii) $y = 0.95 \times 6.2 + 7.25 = 13.14$.

x_i	y_i	$X_i = (x-\bar{x})$	Y_i	X_iY_i	X_i^2	Y_i^2
1	9	−4	−3	12	16	9
2	8	−3	−4	12	9	16
3	10	−2	−2	4	4	4
4	12	−1	0	0	1	0
5	11	0	−1	0	0	1
6	13	1	1	1	1	1
7	14	2	2	4	4	4
8	16	3	4	12	9	16
9	15	4	3	12	16	9

$\Sigma X_iY_i = 57 \quad \Sigma X_i^2 = 60 \quad \Sigma Y_i^2 = 60$

$$r = \frac{\Sigma x_i y_i}{\sqrt{\Sigma x_i^2 \times \Sigma y_i^2}} = \frac{57}{\sqrt{60 \times 60}} = \frac{57}{60} = 0.95$$

$$\sigma x = \sqrt{\frac{\Sigma x_i^2}{n}} = \sqrt{\frac{60}{9}} = \sqrt{\frac{60}{9}} = 2.58$$

$$\sigma y = \sqrt{\frac{\Sigma y_i^2}{n}} = \sqrt{\frac{60}{9}} = 2.58$$

2) The following data is found in respect to the prices of a certain consumer item in belagavi & bengaluru. Average price at belagavi is 65 rs. & bengaluru is 67 rs. Std. deviation at belagavi is 2.5 & 3.5 in bengaluru. The co-efficient of corelation betn the prices in two cities is 0.8. Find most likely price in Bengaluru corresponding to price of rupees 70 in belagavi.

→ x: rate of an item in belagavi

y: rate of an item in bengaluru.

$\bar{x} = 65$ $\quad \bar{y} = 67$

$\sigma x = 2.5$, $\sigma y = 3.5$ $\quad r = 0.8$.

$x = 70$ $\quad y = ?$

Line of regression

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

$$y - 67 = 0.8 \frac{3.5}{2.5}(70 - 65)$$

$$y = \underline{72.6}$$

5) In a partially destroyed laboratory record of corelation data the following results only are available. i) variance of x is 9 ii) regression equations are

$8x - 10y = -66$; $40x - 18y = 214$.

find, i) mean of x & y ii) Std deviation y. iii) Corelation coefficient between x & y.

→ Given:- $\sigma^2 x = 9$ $\qquad \underline{\sigma x = 3}$ $\qquad$ i) $\bar{x}$ & $\bar{y}$

$8x - 10y = -66$ $\qquad$ ii) σy.

$40x - 18y = 214$. $\qquad$ iii) r

from third remark: line of regression always passes through ($\bar{x}$ & $\bar{y}$)

$$\left.\begin{aligned} 8\bar{x} - 10\bar{y} &= -66 \\ 40\bar{x} - 18\bar{y} &= 214. \end{aligned}\right\} \text{∵ lines of regression always passes through } (\bar{x}, \bar{y}).$$

$\bar{x} = 13$, $\bar{y} = 17$.

$$\left.\begin{aligned} b_{yx} &= r\frac{\sigma y}{\sigma x} \\ b_{xy} &= r\frac{\sigma x}{\sigma y} \end{aligned}\right\}$$

$$y = \frac{8x}{10} + \frac{66}{10} \Rightarrow b_{yx} = \frac{8}{10}.$$

$$x = \frac{18}{40}y + \frac{214}{40}. \Rightarrow b_{xy} = \frac{18}{40} = \frac{9}{20}.$$

$$r = \sqrt{b_{yx}\, b_{xy}} = \sqrt{8/10 \times 9/20} = 0.6.$$

$$b_{yx} = r\frac{\sigma y}{\sigma x} \Rightarrow \frac{8}{10} = 0.6 \times \frac{\sigma y}{3}.$$

$$\sigma y = 4.$$

* Multiple & Partial Corelation.

When data involves two variable the corelation between the variables is called simple corelation. But when the data involves more than two variable then multiple & partial corelations are studied.

* Multiple Corelation

It is used to study the cumulative effect all independent variables on dependent variable.

(1 is considered to be dependent 2 & 3 are independent)

$$R_{1,23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2 r_{12} r_{13} r_{23}}{1 - r_{23}^2}}$$

(2 dependent) $$R_{2,13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2 r_{21} r_{23} r_{13}}{1 - r_{13}^2}}$$

(3 dependent) $$R_{3,12} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2 r_{31} r_{32} r_{12}}{1 - r_{12}^2}}$$

Note:

1) The magnitude of multiple corelation coefficient is always between 0 & 1
2) If the value of coefficient is zero then there is an absence of linear relationship & when it is equal to 1 it indicates there is a perfect relationship

IA-I

Here the effect of only 1 independent variable at a time while others are held constant. On dependent variable, is of study.

$$r_{12,3} = \frac{r_{12} - r_{13} \times r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$

$$r_{13.2} = \frac{r_{13} - r_{12} \times r_{23}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{23}^2}}$$

$$r_{23.1} = \frac{r_{23} - r_{21} \times r_{31}}{\sqrt{1 - r_{21}^2}\sqrt{1 - r_{31}^2}}$$

$\sqrt{\Sigma x_i \Sigma y_i^2}$

Date ____ Page ____

1) * Calculate the coefficient of multiple corelation $R_{1,23}$ on the basis of set. data relating 3 variables. It is found that $r_{12} = -0.7$ $r_{13} = 0.65$ $r_{23} = -0.5$

→

$$R_{1,23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2 r_{12} r_{13} r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(-0.7)^2 + (0.65)^2 - 2 \times (-0.7) \times 0.65 \times -0.5}{1 - (-0.5)^2}}$$

$$= 0.781$$

2) In a trivariate distribution it is found that $r_{12} = 0.7$ $r_{13} = 0.61$ $r_{23} = 0.40$ find the values i) $r_{12.3}$ ii) $r_{13.2}$ iii) $r_{23.1}$

→ $$r_{12.3} = \frac{r_{12} - r_{13} \times r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}} = 0.6278$$

$$r_{13.2} = \frac{r_{13} - r_{12} \times r_{23}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{23}^2}} = 0.504$$

$$r_{23.1} = \frac{r_{23} - r_{21} \times r_{31}}{\sqrt{1 - r_{21}^2}\sqrt{1 - r_{31}^2}} = -0.0477$$

3) Obtain the multiple & partial corelation of the following.

x_1	x_2	x_3	$X_1 = x_1 - \bar{x}_1$	$X_2 = x_2 - \bar{x}_2$	$X_3 = x_3 - \bar{x}_3$	$X_1 X_2$	X_1^2	X_2^2
28	74	5	-3.66	-5.5	-2.66	20.13	13.39	30.25
33	87	11	1.34	7.5	3.34	10.05	1.79	56.29
21	69	4	-10.66	-10.5	-3.66	111.43	113.6	110.25
24	69	9	-7.66	-10.5	1.34	80.43	58.67	110.25
38	81	7	6.34	1.5	-0.66	9.51	40.19	2.25
46	97	10	14.34	17.5	2.34	250.95	205.63	306.25

$r_{12} = 0.935$

$r_{13} = 0.554$ $\Sigma X_1 X_2 = 483$. ~~$\Sigma X_1^2 = 104$~~ ~~$\Sigma X_2^2 = 433.34$~~

$r_{23} = 0.668$. $\Sigma X_1^2 = 433.34$ $\Sigma X_2^2 = 615.5$

X_2X_3	X_3^2	X_1X_3
14.63	7.07	9.73
25.05	11.15	4.47
38.43	13.39	39.01
-14.07	1.79	-10.26
-0.99	0.43	-4.18
40.95	5.47	33.55

$\Sigma X_3^2 = 39.33 \quad \Sigma X_1X_3 = 72.33 \quad \Sigma X_2X_3 = 104$

$$r_{12} = \frac{\Sigma X_1X_2}{\sqrt{\Sigma X_1^2 \Sigma X_2^2}} = \frac{483}{\sqrt{433.34 \times 615.5}} = 0.935$$

$$r_{13} = \frac{\Sigma X_1X_3}{\sqrt{\Sigma X_1^2 \Sigma X_3^2}} = \frac{72.33}{\sqrt{433.34 \times 39.33}} = 0.5540.$$

$$r_{23} = \frac{\Sigma X_2X_3}{\sqrt{\Sigma X_2^2 \Sigma X_3^2}} = \frac{104}{\sqrt{615.5 \times 39.33}} = 0.6684.$$

Multiple Correlation

$$R_{1,23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}r_{13}r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(0.935)^2 + (0.554)^2 - 2 \times 0.935 \times 0.554 \times 0.668}{1 - (0.668)^2}}$$

$$= \underline{\underline{0.939}}.$$

$$R_{2,13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2r_{21}r_{23}r_{13}}{1 - r_{13}^2}} = \underline{\underline{0.952}}$$

$$R_{3,12} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2r_{31} \times r_{32}r_{12}}{1 - r_{12}^2}} = \underline{\underline{0.697}}$$

* Partial Corelation:

$$r_{12.3} = \frac{r_{12} - r_{13} \times r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}} = 0.911$$

$$r_{13.2} = \frac{r_{13} - r_{12} \times r_{32}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{32}^2}} = -0.267.$$

$$r_{23.1} = \frac{r_{23} - r_{21} \times r_{31}}{\sqrt{1 - r_{21}^2}\sqrt{1 - r_{31}^2}} = 0.508.$$

* Obtain multiple & partial corelation of the following tables.

x_1	x_2	x_3	X_1	X_2	X_3	X_1X_2	X_2X_3	X_1X_3	X_1^2	X_2^2
193	108	851	39	-20.95	-215.34	-817.05				
172	2.2	1058	18	-20.35	-8.34	-366.8				
113	33	1357	-41	+0.45	290.66					
230	15.5	816	76	-7.05	-250.34					
91	43	1201	-63	20.45	134.66					
125	40	1115	-29	17.45	48.66					

$\bar{x}_1 = 154 \qquad \Sigma X_1X_2 = -3942 \qquad \Sigma X_1^2 = 14112$

$\bar{x}_2 = 22.55 \qquad \Sigma X_2X_3 = 13086.3 \qquad \Sigma X_2^2 = 1734.62$

$\bar{x}_3 = 1066.34 \qquad \Sigma X_1X_3 = -49386 \qquad \Sigma X_3^2 = 214095.33$

$r_{12} = -0.796 \qquad r_{13} = -0.898 \qquad r_{23} = 0.679.$

$R_{1,23} = 0.933 \qquad R_{2,13} = 0.8 \qquad R_{3,12} = 0.899.$

$r_{12.3} = -0.576 \qquad r_{13.2} = -0.804 \qquad r_{23.1} = -0.134.$

* Regression Analysis on Multi-variates:

If x & y & z are three variables then the line of regression of z on x and y is given by

$$z = a + bx + cy$$

y on x & z $\quad y = a + bx + cz$

x on y & z $\quad x = a + by + cz$.

Here the parameters a, b & c can be obtained using the normal equations

* 3 equations

$z = a + bx + cy$
$$\Sigma z = na + b\Sigma x + c\Sigma y$$
$$\Sigma xz = a\Sigma x + b\Sigma x^2 + c\Sigma xy.$$
$$\Sigma yz = a\Sigma y + b\Sigma xy + c\Sigma y^2.$$

$y = a + bx + cz$
$$\Sigma y = na + b\Sigma x + c\Sigma z.$$
$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma xz.$$
$$\Sigma yz = a\Sigma z + b\Sigma xz + c\Sigma z^2.$$

$x = a + by + cz$
$$\Sigma x = na + b\Sigma y + c\Sigma z.$$
$$\Sigma xy = a\Sigma y + b\Sigma y^2 + c\Sigma yz$$
$$\Sigma zx = a\Sigma z + b\Sigma yz + c\Sigma z^2$$

* Obtain the lines of regression for the following trivariate.

x	57	59	49	62	51	50	55	48	52	42	61	57
y	8	10	6	11	8	7	10	9	10	6	12	9
	64	71	53	67	55	58	~~77~~	57	56	51	76	68

x	y	~~z~~	z	x^2	y^2	z^2	xy	yz	xz.
57	8	~~64~~	64	3249					
59	10	~~53~~	71	3481					
49	6	~~67~~	53	2401					
62	11	~~55~~	67	3844					
51	8	~~48~~	55						
50	7	~~47~~	58.						
55	10	~~57~~	77.						
48	9	~~56~~	57						
52	10	~~51~~	56						
42	6	~~76~~	51						
61	12		76						
57	9		68.						

$\Sigma X = 643 \quad \Sigma Y = 106 \quad \Sigma Z = 753.$

$\Sigma X^2 = 34843 \quad \Sigma Y^2 = 976 \quad \Sigma Z^2 = 48139$

$\Sigma XY = 5779. \quad \Sigma YZ = 6796. \quad \Sigma xz = 40830.$

$$z = a + bx + cy.$$

$$\Sigma z = na + b\Sigma x + c\Sigma y.$$
$$753 = 12a + b(643) + c(106). \quad -(1)$$

$$\Sigma xz = a\Sigma x + b\Sigma x^2 + c\Sigma xy$$
$$40830 = a(643) + b(34843) + c(5779) \quad -(2)$$

$$\Sigma yz = a\Sigma y + b\Sigma xy + c\Sigma y^2$$
$$6796 = a(106) + b(5779) + c(976) \quad -(3)$$

$a = 3.651 \quad b = 0.854 \quad c = 1.5063.$

b) line of regression of y on x & z.

5

$\Sigma y = na + b\Sigma x + c\Sigma z$
$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma xz$
$\Sigma yz = a\Sigma z + b\Sigma xz + c\Sigma z^2$

$106 = 12 \times a + b(643) + c(753)$
$5779 = a(643) + b(34843) + c(40830)$
$6796 = a(753) + b(40830) + c(48139)$

$a = -4.560 \quad b = 0.162 \quad c = 0.0743$

c) line of regression of x on y & z.

$\Sigma x = na + b\Sigma y + c\Sigma z$.
$\Sigma xy = a\Sigma y + b\Sigma y^2 + c\Sigma zy$.
$\Sigma xz = a\Sigma z + b\Sigma yz + c\Sigma z^2$

$643 = 12 \times a + b(106) + c(753)$
$5779 = a(106) + b(976) + c(6796)$.
$40830 = a(753) + b(6796) + c(48139)$.

$a = 21.319 \quad \cdot \quad b = 1.286 \qquad c = 0.333$,

$z = 3.651 + 0.854x + 1.506y$
$y = 4.560 + 0.162x + 0.0743z$
$x = 21.319 + 1.286y + 0.333z$.