1

Partial and Multiple Correlation and Regression

INTRODUCTION

Partial and multiple correlation and regression are extension of the technique of simple correlation and regression under which we study the interrelationship between three or more variables.

(1) Multiple Correlation

Multiple correlation is the study of the relationship among three or more variables. Multiple correlation measures the combined influence of two or more independent variables on a single dependent variable. For example, if we study the combined influence of amount of fertiliser (x_2) and rainfall (x_3) on the yield of wheat (x_1), then it is called the problem of multiple correlation. We shall denote the multiple correlation coefficient between x_1, the dependent variables x_2 and x_3 independent variables by $R_{1.23}$. Similarly, we shall denote the other multiple correlation coefficients by $R_{2.13}$ and $R_{3.12}$.

Calculation of Coefficient of Multiple Correlation: The formulae for calculating the multiple correlation coefficients $R_{1.23}$, $R_{2.13}$ and $R_{3.12}$ are as follows :

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

Where, $R_{1.23}$ = Multiple correlation coefficient

r_{12}, r_{13}, r_{23} = Simple or zero order correlation coefficient.

$$R_{2.13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2r_{21} \cdot r_{23} \cdot r_{13}}{1 - r_{13}^2}}$$

$$R_{3.12} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2r_{31} \cdot r_{32} \cdot r_{12}}{1 - r_{12}^2}}$$

Limits of Multiple Correlation Coefficients: The value of multiple correlation coefficient $(R_{1.23})$ lies between 0 and 1. It can never be negative.

$$0 \le R_{1.23} \le 1$$

- The following examples illustrate the calculations of multiple correlation coefficients:

Example 1. **Calculate $R_{1.23}$, $R_{3.12}$ and $R_{2.13}$ for the following data:**

$$r_{12} = 0.60,\ r_{13} = 0.70,\ r_{23} = 0.65$$

Solution. Given, $r_{12} = 0.60, r_{13} = 0.70, r_{23} = 0.65$

(i)
$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(0.6)^2 + (0.7)^2 - 2\,(0.6)\,(0.7)\,(0.65)}{1 - (0.65)^2}}$$

$$= \sqrt{\frac{0.36 + 0.49 - 0.546}{0.5775}}$$

$$= \sqrt{0.526} = 0.725$$

(ii)
$$R_{2.13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2r_{21} \cdot r_{23} \cdot r_{13}}{1 - r_{13}^2}}$$

$$= \sqrt{\frac{(0.6)^2 + (0.65)^2 - 2\,(0.6)\,(0.65)\,(0.7)}{1 - (0.70)^2}}$$

$$= \sqrt{\frac{0.36 + 0.4225 - 0.546}{1 - 0.49}}$$

$$= \sqrt{0.4638} = 0.6809$$

(iii)
$$R_{3.12} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2r_{31} \cdot r_{32} \cdot r_{12}}{1 - r_{12}^2}}$$

$$= \sqrt{\frac{(0.70)^2 + (0.65)^2 - 2\,(0.70)\,(0.65)\,(0.60)}{1 - (0.60)^2}}$$

$$= \sqrt{\frac{0.49 + 0.4225 - 0.546}{1 - 0.36}} = \sqrt{0.5726} = 0.756.$$

Example 2. **For a large group of students x_1 = Score in Economics, x_2 = Score in Maths, ... Score in Statistics, $r_{12} = 0.69$, $r_{13} = 0.45$, $r_{23} = 0.58$. Determine the coeffici... multiple correlation $R_{3.12}$.**

Solution.
$$R_{3.12} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2r_{31} \cdot r_{32} \cdot r_{12}}{1 - r_{12}^2}}$$

$$= \sqrt{\frac{(0.45)^2 + (0.58)^2 - 2\,(0.45)\,(0.58)\,(0.69)}{1 - (0.69)^2}}$$

$$= \sqrt{\frac{0.2025 + 0.3364 - 0.3601}{1 - 0.4761}} = \sqrt{0.3412} = 0.584$$

Example 3. **The following zero order correlation coefficient are given:**

$$r_{12} = 0.98,\ r_{13} = 0.44 \text{ and } r_{23} = 0.54$$

Calculate multiple correlation coefficient treating the first variable as dependent and second and third variables as independent.

Solution. We have to calculate the multiple correlation coefficient treating first variable as dependent and second and third variable as independent *i.e.*, we have to find $R_{1.23}$.

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

Substituting the given values,

$$R_{1.23} = \sqrt{\frac{(0.98)^2 + (0.44)^2 - 2\,(0.98)\,(0.44)\,(0.54)}{1-(0.54)^2}}$$

$$= \sqrt{\frac{0.9604 + 0.1936 - 0.4657}{1 - 0.2916}} = \sqrt{\frac{.6883}{.7084}}$$

$$= \sqrt{0.9716} = 0.985$$

Example 4. **If $R_{1.23} = 1$, prove that $R_{2.13} = 1$.**

Solution.

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

and

$$R_{2.13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2r_{21} \cdot r_{23} \cdot r_{13}}{1 - r_{23}^2}}$$

putting $R_{1.23} = 1$ and squaring both sides,

$$1 = \frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}$$

$$\Rightarrow \quad r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23} = 1 - r_{23}^2$$

$$\Rightarrow \quad r_{12}^2 + r_{23}^2 - 2r_{12} \cdot r_{13} \cdot r_{23} = 1 - r_{13}^2$$

$$\Rightarrow \quad \frac{r_{12}^2 + r_{23}^2 - 2r_{12} \cdot r_{23} \cdot r_{23}}{1 - r_{13}^2} = 1$$

$$\Rightarrow \quad R_{2.13}^2 = 1 \text{ or } R_{2.13} = 1$$

Since, the coefficient of multiple correlation is considered non-negative.

(2) Partial Correlation

Partial Correlation is the simple correlation between two variables after eliminating the influence of the third variable from them. For example, if we measure the relationship between yield of wheat (x_1) and the amount of fertiliser (x_2), eliminating the effect of climate (x_3) from both (having the same climate), then it is called the problem of partial correlation. For three variables

(x_1, x_2 and x_3), there are three partial correlation coefficients. They are denoted by $r_{12.3}$, $r_{13.2}$ and $r_{23.1}$. The partial correlation coefficient $r_{12.3}$ indicates the relationship between x_1 and x_2 when the effect of x_3 is eliminated from both.

Calculation of Partial Correlation Coefficients : The formulae for calculating the partial correlation coefficients $r_{12.3}$, $r_{13.2}$ and $r_{23.1}$ are as follows :

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

Where, $r_{12.3}$ = Partial correlation between x_1 and x_2

r_{12}, r_{13} and r_{23} = Simple or zero order correlation coefficient.

Similarly, we have

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

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

Limits of Partial Correlation Coefficient : The value of $r_{12.3}$ lies betwen -1 and $+1$.

$$-1 \le r_{12.3} \le 1$$

The following examples illustrate the calculations of partial correlation coefficient.

Example 5. **Given that $r_{12} = 0.7$, $r_{13} = 0.61$, $r_{23} = 0.4$. Find the values of $r_{12.3}$, $r_{13.2}$, $r_{23.1}$.**

Solution.

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

Substituting the values,

$$r_{12.3} = \frac{0.7 - (0.61)(0.4)}{\sqrt{1-(0.61)^2}\sqrt{1-(0.4)^2}}$$

$$= \frac{0.456}{0.792 \times 0.916} = 0.629$$

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

$$= \frac{0.61 - (0.7)(0.4)}{\sqrt{1-(0.7)^2}\sqrt{1-(0.4)^2}}$$

$$= \frac{0.61 - 0.28}{\sqrt{1-.49}\sqrt{1-.16}}$$

$$= \frac{0.33}{0.714 \times 0.916} = \frac{0.33}{0.654} = 0.505$$

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

$$=\frac{0.4-(0.7)(0.61)}{\sqrt{1-(0.7)^2}\sqrt{1-(0.61)^2}}$$

$$=\frac{0.4-0.427}{\sqrt{1-(0.49)}\sqrt{1-0.3721}}=\frac{-0.027}{0.714\times 0.792}=-0.048$$

Example 6. **On the basis of observations made on 30 cotton plants the total correlation of yield of cotton (x_1) the number of balls *i.e.*, seed vessels (x_2) and height (x_3) are found to be:**

$$r_{12}=0.8,\ r_{13}=0.65,\ r_{23}=0.7$$

Compute the partial correlation between yield of cotton and number of balls, eliminating the effect of height.

Solution. We have to find the partial correlation between yield of cotton (x_1) and the number of balls (x_2), eliminating the effect of height (x_3) *i.e.*, we have to find $r_{12.3}$.

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

Substituting the given values,

$$r_{12.3}=\frac{0.8-(0.65)(0.7)}{\sqrt{1-(.65)^2}\sqrt{1-(0.7)^2}}$$

$$=\frac{0.8-0.455}{\sqrt{1-0.4225}\sqrt{1-0.49}}$$

$$=\frac{0.345}{0.76\times 0.714}=\frac{0.345}{0.543}=0.635$$

Example 7. **For a large group of students x_1 = Score in theory, x_2 = Score in method, x_3 = Score in field work. The following results were found :**

$$r_{12}=0.69,\ r_{13}=0.45,\ r_{23}=0.58$$

Determine the partial correlation coefficient between score in field work and score in theory keeping the score in method constant and interpret the result.

Solution. We have to find partial correlation coefficient between score in field (x_3) and score in theory (x_1) keeping the scores in method constant *i.e.*, we have to find $r_{31.2}$.

$$r_{31.2}=\frac{r_{31}-r_{32}\cdot r_{12}}{\sqrt{1-r_{32}^2}\sqrt{1-r_{12}^2}}$$

$$=\frac{0.45-(0.58)(0.69)}{\sqrt{1-(0.58)^2}\sqrt{1-(0.69)^2}}$$

$$=\frac{0.45-0.4002}{\sqrt{1-0.3364}\sqrt{1-0.4761}}$$

$$= \frac{0.0498}{\sqrt{0.6636}\,\sqrt{0.5239}}$$

$$= \frac{0.0498}{0.81 \times 0.72} = \frac{0.0498}{0.5832} = 0.085$$

Thus, there is low degree of correlation between score in field work and score in theory.

Example 8. **Is it possible to have the following set of experimental data:**

$$r_{12} = 0.6,\ r_{23} = 0.8,\ r_{31} = -0.5.$$

Solution. In order to see whether there is inconsistency in the given data, we should calculate $r_{12.3}$. If the value of $r_{12.3}$ exceeds one, there is inconsistency, otherwise not.

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

Substituting the given values,

$$= \frac{0.6-(-0.5)(0.8)}{\sqrt{1-(0.5)^2}\,\sqrt{1-(0.8)^2}}$$

$$= \frac{0.6+0.4}{\sqrt{1-0.25}\,\sqrt{1-0.64}}$$

$$= \frac{1}{\sqrt{0.75}\,\sqrt{0.36}} = \frac{1}{0.866 \times 0.6} = \frac{1}{0.52} = 1.92$$

Since, the value of $r_{12.3}$ is greater than one, there is some inconsistency in the given data.

Aliter : We can also check the inconsistency in the data by calculating $R_{1.23}$. If the value of $R_{1.23}$ exceeds 1, there is some inconsistency otherwise not

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1-r_{23}^2}}$$

$$= \sqrt{\frac{(0.6)^2 + (-0.5)^2 - 2(0.6)(-0.5)(0.8)}{1-(0.8)^2}}$$

$$= \sqrt{\frac{0.36+0.25+0.48}{1-.64}} = \sqrt{\frac{1.09}{0.36}} = \sqrt{3.0277} = 1.74$$

Since, the value of $R_{1.23}$ is greater than one, there is some inconsistency in the data.

9. **Suppose a computer has found, for a given set of values of x_1, x_2 and x_3, $r_{12} = 0.96$, $r_{13} = 0.36$ and $r_{23} = 0.78$.**

Explain whether these computations may be said to be free from errors.

For determining whether the given computed values are free from errors or not, we compute the value of $r_{12.3}$. If $r_{12.3}$ comes out to be greater than one, the computed values cannot be regarded as free from errors.

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

Substituting the given values,

$$r_{12.3}=\frac{0.96-(0.36)(0.78)}{\sqrt{1-(0.36)^2}\sqrt{1-(0.78)^2}}$$

$$=\frac{0.96-0.2808}{\sqrt{0.8704}\sqrt{0.3916}}=\frac{0.6792}{0.9329\times 0.6258}=\frac{0.6792}{0.5838}=1.163$$

Since, $r_{12.3}$ is greater than one, the given computed values do contain some errors.

Relationship between Simple, Partial and Multiple Correlation Coefficients

There exists relationship between simple, partial and multiple correlation coefficients which is clear from the following equation:

(i) $1-R_{1.23}^2=(1-r_{12}^2)(1-r_{13.2}^2)$

(ii) $1-R_{2.13}^2=(1-r_{21}^2)(1-r_{23.1}^2)$ and

(iii) $1-R_{3.12}^2=(1-r_{31}^2)(1-r_{32.1}^2)$

Example 10. **In a trivariate distribution, $r_{12}=0.60$, $r_{13}=0.70$, $r_{23}=0.65$, find $R_{1.23}^2$ from r_{12} and $r_{13.2}$.**

Solution. Given : $r_{12}=0.60, r_{13}=0.70, r_{23}=0.65$

Multiple, Simple and Partial Correlation coefficients are related as :

$$R_{1.23}^2=1-(1-r_{12}^2)(1-r_{13.2}^2)$$

$$r_{13.2}=\frac{r_{13}-r_{12}\cdot r_{32}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{32}^2}}=\frac{0.70-0.60\times 0.65}{\sqrt{1-(0.60)^2}\sqrt{1-(0.65)^2}}$$

$$=\frac{0.70-0.39}{0.8\times 0.760}=0.509$$

$$\therefore \quad r_{13.2}^2=0.259, r_{12}^2=0.36$$

Substituting values r_{12}^2 and $r_{13.2}^2$ for $R_{1.23}^2$, we have

$$R_{1.23}^2=1-(1-0.36)(1-0.259)$$

$$=1(0.64)(0.74)=0.526.$$

MISCELLANEOUS SOLVED EXAMPLES

Exmaple 11. **x_1, x_2 and x_3 are measured from their means with :**

$N=10$, $\Sigma x_1^2=90$, $\Sigma x_2^2=160$, $\Sigma x_3^2=40$

$\Sigma x_1x_2=60$, $\Sigma x_2x_3=60$, $\Sigma x_3x_1=40$

Calculate $r_{12.3}$ and $R_{2.31}$.

Solution.

$$r_{12}=\frac{\Sigma x_1x_2}{\sqrt{\Sigma x_1^2}\sqrt{\Sigma x_2^2}}=\frac{60}{\sqrt{90\times 160}}=\frac{60}{120}=0.5$$

$$r_{13}=\frac{\Sigma x_1 x_3}{\sqrt{\Sigma x_1^2 \times \Sigma x_3^2}}=\frac{40}{\sqrt{90 \times 40}}=\frac{40}{60}=0.67$$

$$r_{23}=\frac{\Sigma x_2 x_3}{\sqrt{\Sigma x_2^2 \times \Sigma x_3^2}}=\frac{60}{\sqrt{160 \times 40}}=\frac{60}{80}=0.75$$

Now,

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

Substituting the values, we have

$$r_{12.3}=\frac{0.5-0.67 \times 0.75}{\sqrt{1-(0.67)^2}\sqrt{1-(0.75)^2}}=-\frac{0.0025}{0.4910}=-0.0051$$

$$R_{2.31}=\sqrt{\frac{r_{23}^2+r_{21}^2-2r_{23} \cdot r_{21} \cdot r_{31}}{1-r_{31}^2}}$$

$$=\sqrt{\frac{(0.75)^2+(0.5)^2-2\,(0.75)\,(0.5)\,(0.67)}{1-(.67)^2}}$$

$$=\sqrt{\frac{0.5625+0.25-0.5025}{0.5511}}=\sqrt{\frac{0.31}{0.5511}}=0.75$$

Example 12. **Calculate $r_{12.3}$ and $R_{1.23}$ from the following data :**

X :	3	4	5	6	7	8	9
Y :	2	5	6	4	3	2	4
Z :	5	6	4	5	6	5	8

Solution.

Calculation of $r_{12.3}$ and $R_{1.23}$

X	X^2	Y	Y^2	Z	Z^2	XY	XZ	YZ
3	9	2	4	5	25	6	15	10
4	16	5	25	6	36	20	24	30
5	25	6	36	4	16	30	20	24
6	36	4	16	5	25	24	30	20
7	49	3	9	6	36	21	42	18
8	64	2	4	5	25	16	40	10
9	81	4	16	8	64	36	72	32
$N=7$ $\Sigma X=42$	$\Sigma X^2=280$	$\Sigma Y=26$	$\Sigma Y^2=110$	$\Sigma Z=39$	$\Sigma Z^2=227$	$\Sigma XY=153$	$\Sigma XZ=243$	$\Sigma YZ=14$

$$r_{12}=\frac{N.\Sigma XY-\Sigma X.\Sigma Y}{\sqrt{[\Sigma X^2.N-(\Sigma X)^2][\Sigma Y^2.N-(\Sigma Y)^2]}}=\frac{7\times 153-(42\times 26)}{\sqrt{[280\times 7-(42)^2][110\times 7-(26)^2]}}=-0.155$$

$$r_{13}=\frac{N.\Sigma XZ-\Sigma X.\Sigma Z}{\sqrt{[\Sigma X^2.N-(\Sigma X)^2][\Sigma Z^2.N-(\Sigma Z)^2]}}=\frac{7\times 243-(42\times 39)}{\sqrt{[280\times 7-(42)^2][227\times 7-(39)^2]}}=0.546$$

$$r_{23}=\frac{N.\Sigma YZ-\Sigma Y.\Sigma Z}{\sqrt{[\Sigma Y^2.N-(\Sigma Y)^2][\Sigma Z^2.N-(\Sigma Z)^2]}}=\frac{144\times 7-26\times 39}{\sqrt{[110\times 7-(26)^2][227\times 7-(39)^2]}}=-0.075$$

Partial Correlation Coefficient

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

$$=\frac{-0.155-(0.546\times -0.075)}{\sqrt{1-(0.546)^2}\sqrt{1-(-0.075)^2}}=-0.1366$$

Multiple Correlation Coefficient

$$R_{1.23}=\sqrt{\frac{r_{12}^2+r_{13}^2-2r_{12}.r_{13}.r_{23}}{1-r_{23}^2}}$$

$$=\sqrt{\frac{(-0.155)^2+(0.546)^2-(2\times -0.155\times 0.546\times -0.075)}{1-(-0.075)^2}}$$

$$=\sqrt{\frac{0.024+0.298-(.01269)}{1-.006}}$$

$$=\sqrt{\frac{0.1951}{0.994}}$$

$$=\sqrt{0.1962}=0.443$$

Example 12 A. **In a trivariate distribution, $r_{12}=0.80$, $r_{23}=-0.56$, $r_{31}=-0.40$, compute $r_{23.1}$ and $R_{1.23}$.**

Solution. (i)

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

$$=\frac{-0.56-(0.8).(-0.40)}{\sqrt{1-(0.8)^2}\sqrt{1-(-0.4)^2}}$$

$$=\frac{-0.56+0.32}{\sqrt{1-0.64}\sqrt{1-0.16}}$$

$$=\frac{-0.24}{\sqrt{0.36\times 0.84}}=\frac{-0.24}{0.5499}=-0.436$$

(ii)

$$R_{1.23}=\sqrt{\frac{r_{12}^2+r_{13}^2-2r_{12}.r_{13}.r_{23}}{1-r_{23}^2}}$$

$$=\sqrt{\frac{(0.8)^2+(-0.4)^2-2(0.8)(-0.4)(-0.56)}{1-(-0.56)^2}}$$

$$=\sqrt{\frac{0.64+0.16-0.3584}{1-0.3136}}=\sqrt{\frac{0.4416}{0.6864}}=0.802$$

Scanned by TapScanner

Example 13. **The linear correlation coefficient between x_1 (Yield), x_2 (Irrigation) ... (Fertiliser) are as follows :**

$$r_{12} = 0.81,\ r_{13} = 0.90,\ r_{23} = 0.65$$

Calculate the partial correlation coefficient of:

(i) yield with irrigation

(ii) yield with fertiliser.

Solution. (i) We have to find $r_{12.3}$

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

Substituting the given values,

$$r_{12.3} = \frac{(0.81)-(0.90)\,(0.65)}{\sqrt{1-(0.90)^2}\sqrt{1-(0.65)^2}}$$

$$= \frac{0.81-0.585}{0.4358 \times 0.7599} = \frac{0.225}{0.3311} = 0.679$$

(ii) We have to find $r_{13.2}$

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

$$= \frac{(0.90)-(0.81)\,(0.65)}{\sqrt{1-(0.81)^2}\sqrt{1-(0.65)^2}}$$

$$= \frac{0.3735}{\sqrt{0.3439}\sqrt{0.5775}}$$

$$= \frac{0.3735}{0.5864 \times 0.7599} = \frac{0.3735}{0.4456} = 0.838$$

Example 14. **Given the following zero order correlation coefficient, find (i) partial corre... coefficient between x_2 and x_3 and (ii) multiple correlation taking ... dependent on x_2 and x_3.**

$$r_{12} = 0.98,\ r_{13} = 0.44,\ r_{23} = 0.54$$

Solution. (i)

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

$$= \frac{0.54-(0.98)\,(0.44)}{\sqrt{1-(0.98)^2}\sqrt{1-(0.44)^2}}$$

$$= \frac{0.54-0.4312}{\sqrt{1-0.9604}\sqrt{1-0.1936}}$$

$$= \frac{0.1088}{\sqrt{0.0396}\sqrt{0.8064}} = \frac{0.1088}{0.1786} = 0.6091$$

(ii)
$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$
$$= \sqrt{\frac{(0.98)^2 + (0.44)^2 - 2(0.98)(0.44)(0.54)}{1 - (0.54)^2}}$$
$$= \sqrt{\frac{0.9604 + 0.1936 - 0.4656}{1 - (0.2916)}} = \sqrt{\frac{0.6884}{0.7084}}$$
$$= \sqrt{0.9717} = 0.985.$$

Example 15. Is it possible to get the following from a set of experimental data:
(i) $r_{23} = 0.8, r_{31} = 0.5, r_{12} = 0.6$
(ii) $r_{23} = 0.7, r_{31} = -0.4, r_{12} = 0.6$

Solution. (i) In order to see whether there is any inconsistency, we should calculate $r_{12.3}$, If its value exceed one, there is inconsistency, otherwise not.

$$r_{12.3} = \frac{r_{12} - r_{13} \cdot r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$
$$= \frac{0.6 - (0.5)(0.8)}{\sqrt{1 - (0.5)^2}\sqrt{1 - (0.8)^2}}$$
$$= \frac{0.20}{\sqrt{0.75}\sqrt{0.36}} = \frac{0.20}{0.52} = 0.384$$

Since, the value of $r_{12.3}$ is less than one, the data is consistent.

(ii)
$$r_{12.3} = \frac{r_{12} - r_{13} \cdot r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}} = \frac{(0.6) - (-0.4)(0.7)}{\sqrt{1 - (-0.4)^2}\sqrt{1 - (0.7)^2}}$$
$$= \frac{0.6 + 0.28}{\sqrt{0.84}\sqrt{0.51}} = \frac{0.88}{0.655} = 1.344$$

Since, the value or $r_{12.3}$ is greater than 1 there is some inconsistency in the given data.

Example 16. Test the consistency of the following data : $r_{12} = 0.8, r_{13} = 0.4, r_{23} = -0.56$.

Solution. For testing whether the given computations are consistent or not, we compute the value of $r_{13.2}$. If $r_{13.2}$ comes out to be greater than one, the computations cannot regarded as consistent.

$$r_{12.3} = \frac{r_{12} - r_{13} \cdot r_{23}}{\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}}$$
$$= \frac{(0.8) - (0.4)(-0.56)}{\sqrt{1 - (0.4)^2}\sqrt{1 - (-0.56)^2}}$$
$$= \frac{0.8 + 0.224}{\sqrt{0.84}\sqrt{0.6864}} = \frac{1.024}{0.7593} = 1.349$$

Since, $r_{12.3}$ is greater than one, the given computations of r_{12}, r_{13} and r_{23} are not consistent.

Example 17. **If $r_{12} = 0.77$, $r_{13} = 0.72$ and $r_{23} = 0.52$, find the partial correlation coefficient $r_{12.3}$ and multiple correlation coefficient $R_{1.23}$.**

Solution. Given : $r_{12} = 0.77$, $r_{13} = 0.72$ and $r_{23} = 0.52$

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

$$= \frac{.77 - .72 \times .52}{\sqrt{1-(.72)^2}\,\sqrt{1-(.52)^2}}$$

$$= \frac{.77 - .37}{\sqrt{1-.5184}\,\sqrt{1-.2704}}$$

$$= \frac{.40}{\sqrt{.4816} \times \sqrt{.7296}} = \frac{.4}{.593} = 0.6745$$

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1-r_{23}^2}}$$

$$= \sqrt{\frac{(.77)^2 + (.72)^2 - 2\,(.77)\,(.72)\,(.52)}{1-(.52)^2}}$$

$$= \sqrt{\frac{.5929 + .5184 - .5766}{1-.2704}} = \sqrt{\frac{.5347}{.7296}} = 0.856\,.$$

Example 18. **If $r_{12} = 0.60$, $r_{13} = 0.70$, $r_{23} = 0.65$, find partial correlation between x_1 and x_2 and multiple correlation between x_1 dependent on x_2 and x_3.**

Solution. Given : $r_{12} = 0.60$, $r_{13} = 0.70$, $r_{23} = 0.65$

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

$$= \frac{0.6 - 0.7 \times 0.65}{\sqrt{1-(0.7)^2}\,\sqrt{1-(0.65)^2}}$$

$$= \frac{0.6 - 0.455}{\sqrt{0.51 \times 0.5775}} = \frac{0.145}{0.543} = 0.2670$$

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1-r_{23}^2}}$$

$$= \sqrt{\frac{(.6)^2 + (.7)^2 - 2(.6)\,(.7)\,(.65)}{1-(.65)^2}} = \sqrt{\frac{.304}{.5775}} = 0.726\,.$$

Example 19. **Following table shows the correlation matrix of three variables x_1 (Height), x_2 (Weight) and x_3 (Diameter of Chest) of 10 randomly selected players :**

$$\begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 1.0000 & 0.8630 & 0.6480 \\ & 1.0000 & 0.7090 \\ & & 1.0000 \end{bmatrix} \end{array}$$

Calculate $r_{12.3}$ and $R_{1.23}$.

Solution. Given : $r_{12} = 0.863, r_{13} = 0.648, r_{23} = 0.709$

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

$$= \frac{.863 - .648 \times .709}{\sqrt{1-(.648)^2}\sqrt{1-(.709)^2}}$$

$$= \frac{.863 - .4594}{\sqrt{.580} \times \sqrt{.497}} = \frac{.4036}{\sqrt{.2883}} = \frac{.4036}{.537} = 0.752$$

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} \cdot r_{13} \cdot r_{23}}{1-r_{23}^2}}$$

$$= \sqrt{\frac{(.863)^2 + (.648)^2 - 2\,(.863)\,(.648)\,(.709)}{1-(.709)^2}}$$

$$= \sqrt{\frac{.745 + .42 - .793}{.497}} = 0.865$$

EXERCISE – 1

1. In a trivariate distribution, it is found that
$$r_{12} = 0.41, r_{13} = 0.71, r_{23} = 0.5$$
Find the value of $r_{23.1}$ and $r_{13.2}$. **[Ans.** $r_{23.1} = 0.325, r_{13.2} = 0.639$]
2. If $r_{12} = 0.7, r_{13} = 0.61$ and $r_{23} = 0.4$, find the value of $r_{12.3}, r_{13.2}$ and $r_{23.1}$.
[Ans. $r_{12.3} = 0.629, r_{13.2} = 0.505, r_{23.1} = -0.048$]
3. Is it possible to have the following experimental data:
$r_{12} = 0.6, r_{23} = 0.8, r_{31} = -0.5$ **[Ans.** $r_{12.3} = 1.92$, Inconsistency]
4. In a trivariate distribution, $r_{23} = .2, r_{13} = .5, r_{12} = .6$. Compute $r_{12.3}$ and $R_{1.23}$.
[Ans. $r_{12.3} = 0.47, R_{1.23} = 0.714$]
5. Suppose a computer has found for a given set of values of $x_1, x_2, x_3 : r_{12} = 0.91, r_{13} = 0.33$ and $r_{23} = 0.81$. Explain whether these computations may be said to be free from errors.
[Ans. $r_{12.3} = 1.161$; Not free from errors]
6. The following zero order correlation coefficients are given:
$$r_{12} = 0.98, r_{13} = 0.44, r_{23} = 0.54$$

Calculate :

(i) the partial correlation coefficient between first (x_1) and third (x_3) variables; and

(ii) multiple correlation coefficient treating first variable (x_1) as dependent and se and third variable as independent. [Ans. $r_{13.2} = -0.53$, $R_{1.23} = 0$

7. If $r_{12} = 0.9, r_{13} = 0.75, r_{23} = 0.7$, find the $R_{1.23}$. [Ans. $R_{1.23} = 0$

8. Test the consistency of the data ::

$$r_{12} = 0.6, r_{13} = 0.5 \text{ and } r_{23} = 0.2.$$

Compute $r_{12.3}$, and $R_{1.23}$. [Ans. $r_{12.3} = 0.589$, Consis

9. Given the following values:

$$r_{12} = 0.6, r_{23} = r_{31} = 0.8,$$

find : $r_{23.1}$ and $R_{1.23}$. [Ans. $r_{23.1} = 0.667$, $R_{1.23} = 0$

10. For a large group of students, x_1 = Score in Economics, x_2 = Score in Maths, x_3 = Scor Statistics, $r_{12} = 0.69, r_{13} = 0.45, r_{23} = 0.58$. Determine the coefficient of multiple correla $R_{3.12}$. [Ans. $R_{3.12} = 0$

11. The simple correlation coefficient between temperature (x_1), corp yield (x_2) and rainfall (x_3) are :

$r_{12} = 0.59, r_{13} = 0.46$ and $r_{23} = 0.77$. Calculate $r_{12.3}$ and $R_{1.23}$.

[Ans. $r_{12.3} = 0.416$, $R_{1.23} = 5$

12. x_1, x_2 and x_3 are measured from their means with:

$N = 6, \Sigma x_1^2 = 90, \Sigma x_2^2 = 140, \Sigma x_3^2 = 4008$

$\Sigma x_1 x_2 = -100, \Sigma x_1 x_3 = -582, \Sigma x_2 x_3 = 720$

Calculate $r_{12.3}$ and $R_{1.23}$

[Ans. $r_{12} = -0.891, r_{13} = -.969, r_{23} = 0.961, r_{12.3} = 0.605, R_{1.23} = 0.9$

(3) MULTIPLE REGRESSION

In multiple regression, we study three variables and we consider one variable as depende variable and the other two as independent variables. Multiple regression analysis is used estimate the most probable value of the dependent variable for given values of the independe variables.

Methods to obtain Multiple Regression Equations

Multiple regression equations can be worked out by two methods, which are as follows:

(1) Multiple Regression Equations using Normal Equations

(2) Multiple Regression Eqautions in terms of Simple Correlation Coefficients

Let us discuss them.

(1) Multiple Regression Equations using Normal Equations : This method is also called **Least Square Method.** Under this method computation of regression equations is done by solvi three normal equations. This method becomes clear by the following :

Multiple regression equation of X_1 on X_2 and X_3 is given by :

$$X_1 = a_{1.23} + b_{12.3} X_2 + b_{13.2} X_3$$

Where, X_1 = Dependent variable, X_2 and X_3 = Independent variables.

$b_{12.3}$ and $b_{13.2}$ = Partial regression coefficients.

Using least square method, the values of constants $a_{1.23}$, $b_{12.3}$ and $b_{13.2}$ are obtained by solving the following three normal equations:

$$\Sigma X_1 = N \cdot a_{1.23} + b_{12.3}\ \Sigma X_2 + b_{13.2}\ \Sigma X_3 \quad ...(1)$$
$$\Sigma X_1 X_2 = a_{1.23}\ \Sigma X_2 + b_{12.3}\ \Sigma X_2^2 + b_{13.2}\ \Sigma X_2 X_3 \quad ...(2)$$
$$\Sigma X_1 X_3 = a_{1.23}\ \Sigma X_3 + b_{12.3}\ \Sigma X_2 X_3 + b_{13.2}\ \Sigma X_3^2 \quad ...(3)$$

Similarly, the multiple regression equations of X_2 on X_1 and X_3 and X_3 ; on X_1 and X_2 and their normal equations can also be written as:

Multiple Regression Equation of X_2 on X_1 and X_3 is given by:

$$X_2 = a_{2.13} + b_{21.3} X_1 + b_{23.1} X_3$$

Three Normal Equations are :

$$\Sigma X_2 = N a_{2.13} + b_{21.3}\ \Sigma X_1 + b_{23.1}\ \Sigma X_3$$
$$\Sigma X_2 X_1 = a_{2.13}\ \Sigma X_1 + b_{21.3} \Sigma X_1^2 + b_{23.1}\ \Sigma X_3 X_1$$
$$\Sigma X_2 X_3 = a_{2.13}\ \Sigma X_3 + b_{21.3}\ \Sigma X_1 X_3 + b_{23.1}\ \Sigma X_3^2$$

Multiple Regression Equation of X_3 on X_1 and X_2 is given by:

$$X_3 = a_{3.12} + b_{31.2}\ X_1 + b_{32.1}\ X_2$$

Three Normal Equations are:

$$\Sigma X_3 = N a_{3.12} + b_{31.2}\ \Sigma X_1 + b_{32.1}\ \Sigma X_2$$
$$\Sigma X_3 X_1 = a_{3.12}\ \Sigma X_1 + b_{31.2}\ \Sigma X_1^2 + b_{32.1}\ \Sigma X_2 X_1$$
$$\Sigma X_3 X_2 = a_{3.12}\ \Sigma X_2 + b_{31.2}\ \Sigma X_1 X_2 + b_{32.1}\ \Sigma X_2^2$$

The following example illustrate the procedure of fitting multiple regression equations:

Example 1. **For the following set of data, calculate multiple regression equation of X_1 on X_2 and X_3 :**

X_1 :	**4**	**6**	**7**	**9**	**13**	**15**
X_2 :	**15**	**12**	**8**	**6**	**4**	**3**
X_3 :	**30**	**24**	**20**	**14**	**10**	**4**

Solution. The regression equation of X_1 on X_2 and X_3 is

$$X_1 = a_{1.23} + b_{12.3}\ X_2 + b_{13.2}\ X_3$$

The three normal equations are :

$$\Sigma X_1 = N a_{1.23} + b_{12.3}\ \Sigma X_2 + b_{13.2}\ \Sigma X_3$$
$$\Sigma X_1 X_2 = a_{1.23}\ \Sigma X_2 + b_{12.3}\ \Sigma X_2^2 + b_{13.2}\ \Sigma X_3 X_2$$
$$\Sigma X_1 X_3 = a_{1.23}\ \Sigma X_3 + b_{12.3}\ \Sigma X_2 X_3 + b_{13.2}\ \Sigma X_3^2$$

X_1	X_2	X_3	X_1X_2	X_1X_3	X_2X_3	X_2^2	X_3^2
4	15	30	60	120	450	225	900
6	12	24	72	144	288	144	576
7	8	20	56	140	160	64	400
9	6	14	54	126	84	36	196
13	4	10	52	130	40	16	100
15	3	4	45	60	12	9	16
$\Sigma X_1 = 54$	$\Sigma X_2 = 48$	$\Sigma X_3 = 102$	$\Sigma X_1X_2 = 339$	$\Sigma X_1X_3 = 720$	$\Sigma X_2X_3 = 1034$	$\Sigma X_2^2 = 494$	$\Sigma X_3^2 = 2188$

Substituting the values in the normal equations :

$$54 = 6a_{1.23} + 48b_{12.3} + 102b_{13.2}$$
$$339 = 48a_{1.23} + 494b_{12.3} + 1034b_{13.2}$$
$$720 = 102a_{1.33} + 1034b_{12.3} + 2188b_{13.2}$$

Multiplying (i) by 8, we get

$$432 = 48a_{1.23} + 384b_{12.3} + 816b_{13.2}$$

Subtracting (ii) from (iv), we get

$$-93 = 110b_{12.3} + 218b_{13.2}$$

Multiplying (i) by 17, we get

$$918 = 102a_{1.23} + 816b_{12.3} + 1734b_{13.2}$$

Subtracting (iii) from (vi), we get

$$-198 = 218b_{12.3} + 454b_{13.2}$$

Multiplying (v) by 109, we obtain

$$-10137 = 11990b_{12.3} + 23762b_{13.2}$$

Multiplying (vii) by 55, we get

$$-10890 = 11990b_{12.3} + 24970b_{13.2}$$

Subtracting (viii) from (ix), we get

$$753 = -1208b_{13.2}$$

$$\therefore \quad b_{13.2} = -\frac{753}{1208} = -0.623$$

Substituting the value of $b_{13.2}$ in equation (v), we get

$$-93 = 110b_{12.3} + 218(-0.623)$$
$$135.814 - 93 = 110b_{12.3}$$
$$\therefore \quad b_{12.3} = \frac{42.814}{110} = 0.389$$

Substituting the values of $b_{12.3}$ and $b_{13.2}$ in equation (i), we get

$$6a_{1.23} + 48(0.389) + 102(-0.623) = 54$$
$$6a_{1.23} + 18.672 - 63.546 = 54$$
$$6a_{1.23} = 54 - 18.672 + 63.546$$
$$6a_{1.23} = 98.874$$
$$\therefore \quad a_{1.23} = \frac{98.874}{6} = 16.479$$

Hence, the required equation is

$$X_1 = 16.479 + 0.389X_2 - 0.623X_3$$

EXERCISE – 2

1. From the following data, find the least square regression of X_1, X_2 and X_3 and estimate value of X_1 for given values of $X_2 = 16$ and $X_3 = 4$:

X_1 :	10	5	10	4	8
X_2 :	16	13	21	10	13
X_3 :	3	6	4	5	3

[Ans. $X_1 = 4.753 + 0.502X_2 - 1.115X_3$, 8

2. Compute the values of b_0, b_1 and b_2 for the equation $Y = b_0 + b_1X_1 + b_2X_2$ from the following data:

Y:	3	5	6	8	12	14
X_1:	16	10	7	4	3	2
X_2:	90	72	54	42	30	12

[**Ans.** $Y = 16.1067 + .426X_1 - 0.221X_2$]

3. Obtain the parameters of the multiple linear regression model: $Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3$ from the following data:

$$N = 6, \Sigma Y = 54, \Sigma X_2 = 48, \Sigma X_3 = 102$$
$$\Sigma YX_2 = 339, \Sigma YX_3 = 720, \Sigma X_2X_3 = 1034, \Sigma X_2^2 = 494, \Sigma X_3^2 = 2188$$

[**Ans.** $Y = 16.479 + 0.389X_2 - 0.623X_3$]

Short-Cut Method : When the size of the values of the variables are very large, then the above system of solving normal equations becomes a very tedious procedure. In such a case, in place of actual values, deviations from the means of the variables are used to simplify the computation procedure.

Multiple Regression Eqaution of X_1 on X_2 and X_3 in deviation form is given by :

$$X_1 - \overline{X}_1 = b_{12.3}(X_2 - \overline{X}_2) + b_{13.2}(X_3 - \overline{X}_3)$$

or $\quad x_1 = b_{12.3}x_2 + b_{13.2}x_3 \quad$ where $x_1 = X_1 - \overline{X}_1, x_2 = X_2 - \overline{X}_2, x_3 = X_3 - \overline{X}_3$

The values of the partial regression coefficients ($b_{12.3}$ and $b_{13.2}$) can be obtained by solving the following two normal equations :

$$\Sigma x_1 x_2 = b_{12.3}\Sigma x_2^2 + b_{13.2}\Sigma x_3x_2$$
$$\Sigma x_1x_3 = b_{12.3}\Sigma x_2x_3 + b_{13.2}\Sigma x_3^2$$

Further solved, we have

$$b_{12.3} = \frac{(\Sigma x_1x_2)(\Sigma x_3^2) - (\Sigma x_1x_3)(\Sigma x_2x_3)}{(\Sigma x_2^2)(\Sigma x_3^2) - (\Sigma x_2x_3)^2}$$

$$b_{13.2} = \frac{(\Sigma x_1x_3)(\Sigma x_2^2) - (\Sigma x_1x_2)(\Sigma x_3x_2)}{(\Sigma x_3^2)(\Sigma x_2^2) - (\Sigma x_3x_2)^2}$$

Similarly, the multiple regression equation of X_2 on X_1 and X_3 ; and X_3 on X_1 and X_2 and their normal equations can also be written.

Multiple Regression Equation of X_2 on X_1 and X_3 in deviation form is given by:

$$X_2 - \overline{X}_2 = b_{21.3}(X_1 - \overline{X}_1) + b_{23.1}(X_3 - \overline{X}_3)$$

or $\quad x_2 = b_{21.3}x_1 + b_{23.1}x_3$

Two normal equations are :

$$\Sigma x_2x_1 = b_{21.3}\Sigma x_1^2 + b_{23.1}\Sigma x_1x_3$$
$$\Sigma x_2x_3 = b_{21.3}\Sigma x_1x_3 + b_{23.1}\Sigma x_3^2$$

Further solved, we have

$$b_{21.3} = \frac{(\Sigma x_2x_1)(\Sigma x_3^2) - (\Sigma x_2x_3)(\Sigma x_1x_3)}{(\Sigma x_1^2)(\Sigma x_3^2) - (\Sigma x_1x_3)^2}$$

$$b_{23.1} = \frac{(\Sigma x_2 x_3)(\Sigma x_1^2) - (\Sigma x_2 x_1)(\Sigma x_3 x_1)}{(\Sigma x_3^2)(\Sigma x_1^2) - (\Sigma x_3 x_1)^2}$$

Multiple Regression Equation of X_3 on X_1 and X_2 in deviation form is given by:

$$X_3 - \overline{X}_3 = b_{31.2}(X_1 - \overline{X}_1) + b_{32.1}(X_2 - \overline{X}_2)$$

or

$$x_3 = b_{31.2}\, x_1 + b_{32.1} x_2$$

Two normal equations are :

$$\Sigma x_1 x_3 = b_{31.2}\, \Sigma x_1^2 + b_{23.1}\, \Sigma x_1 x_2$$

$$\Sigma x_2 x_3 = b_{31.2}\, \Sigma x_1 x_2 + b_{32.1}\, \Sigma x_2^2$$

Further solved, we have

$$b_{31.2} = \frac{(\Sigma x_3 x_1)(\Sigma x_2^2) - (\Sigma x_3 x_2)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

$$b_{32.1} = \frac{(\Sigma x_3 x_2)(\Sigma x_1^2) - (\Sigma x_3 x_1)(\Sigma x_2 x_1)}{(\Sigma x_2^2)(\Sigma x_1^2) - (\Sigma x_2 x_1)^2}$$

The following examples would clarify the method:

Example 1. **From the following data, find the least square regression of X_3 on X_1 and using actual mean method. Also estimate X_3 when $X_1 = 10$ and $X_2 = 6$.**

X_1 :	3	5	6	8	12	14
X_2 :	16	10	7	4	3	2
X_3 :	90	72	54	42	30	12

Solution.

X_1	$x_1 = (X_1 - \overline{X}_1)$	x_1^2	X_2	$x_2 = (X_2 - \overline{X}_2)$	x_2^2	X_3	$x_3 = (X_3 - \overline{X}_3)$	x_3^2	$x_1 x_2$	$x_1 x_3$	[illegible]
3	−5	25	16	+9	81	90	+40	1600	−45	−200	[illegible]
5	−3	9	10	+3	9	72	+22	484	−9	−66	[illegible]
6	−2	4	7	0	0	54	+4	16	0	−8	[illegible]
8	0	0	4	−3	9	42	−8	64	0	0	[illegible]
12	+4	16	3	−4	16	30	−20	400	−16	−80	[illegible]
14	+6	36	2	−5	25	12	−38	1444	−30	−228	[illegible]
$\Sigma X_1 = 48$	$\Sigma x_1 = 0$	$\Sigma x_1^2 = 90$	$\Sigma X_2 = 42$	$\Sigma x_2 = 0$	$\Sigma x_2^2 = 140$	$\Sigma X_3 = 300$	$\Sigma x_3 = 0$	$\Sigma x_3^2 = 4008$	$\Sigma x_1 x_2 = -100$	$\Sigma x_1 x_3 = -582$	[illegible]

$$\overline{X}_1 = \frac{48}{6} = 8, \overline{X}_2 = \frac{42}{6} = 7, \overline{X}_3 = \frac{300}{6} = 50,$$

Regression Equation of X_3 on X_1 and X_2 is:

$$X_3 - \overline{X}_3 = b_{31.2}(X_1 - \overline{X}_1) + b_{32.1}(X_2 - \overline{X}_2)$$

$$b_{31.2} = \frac{(\Sigma x_3 x_1)(\Sigma x_2^2) - (\Sigma x_3 x_2)(\Sigma x_1 x_2)}{(\Sigma x_1^2)(\Sigma x_2^2) - (\Sigma x_1 x_2)^2}$$

$$= \frac{(-582)(140) - (720)(-100)}{(90)(140) - (-100)^2}$$

$$= \frac{-81480+72000}{12600-10000} = \frac{-9480}{2600} = -3.646$$

$$b_{32.1} = \frac{(\Sigma x_3 x_2)(\Sigma x_1^2) - (\Sigma x_3 x_1)(\Sigma x_2 x_1)}{(\Sigma x_2^2)(\Sigma x_1^2) - (\Sigma x_2 x_1)^2}$$

$$= \frac{(720)(90) - (-582)(-100)}{(90)(140) - (-100)^2}$$

$$= \frac{64800-58200}{12600-10000} = \frac{6600}{2600} = 2.538$$

Substituting the values in the above equations, we get

$$X_3 - 50 = -3.646(X_1 - 8) + 2.538(X_2 - 7)$$
$$X_3 - 50 = -3.646X_1 + 29.168 + 2.538X_2 - 17.766$$
$$X_3 = -3.646X_1 + 2.538X_2 + 61.402$$

When $X_1 = 10$ and $X_2 = 6$, So, $X_3 = -3.646(10) + 2.538(6) + 61.402$

$$= -36.46 + 15.228 + 61.402 = 40.17 \text{ or } 40.$$

Example 3. Given the following information (variables are measured from their respective means):

$$\Sigma x_1 x_2 = 720,\ \Sigma x_2 x_3 = -582,\ \Sigma x_1 x_3 = -100$$
$$\Sigma x_2^2 = 4008,\ \Sigma x_3^2 = 90,\ \Sigma x_1^2 = 140$$
$$\overline{X}_1 = 7,\ \overline{X}_2 = 50,\ \overline{X}_3 = 8$$

Find the multiple regression equation of X_1 on X_2 and X_3. Estimate X_1 when $X_2 = 10$ and $X_3 = 95$.

Solution. Regression Equation of X_1 on X_2 and X_3 is given by :

$$X_1 - \overline{X}_1 = b_{12.3}(X_2 - \overline{X}_2) + b_{13.2}(X_3 - \overline{X}_3)$$

$$b_{12.3} = \frac{(\Sigma x_1 x_2)(\Sigma x_3^2) - (\Sigma x_1 x_3)(\Sigma x_2 x_3)}{(\Sigma x_2^2)(\Sigma x_3^2) - (\Sigma x_2 x_3)^2}$$

$$= \frac{(720)(90) - (-100)(-582)}{(4008)(90) - (-582)^2}$$

$$= \frac{64800-58200}{360720-338724} = \frac{6600}{21996} = 0.30$$

$$b_{13.2} = \frac{(\Sigma x_1 x_3)(\Sigma x_2^2) - (\Sigma x_1 x_2)(\Sigma x_3 x_2)}{(\Sigma x_2^2)(\Sigma x_3^2) - (\Sigma x_2 x_3)^2}$$

$$= \frac{(-100)(4008) - (720)(-582)}{(4008)(90) - (-582)^2}$$

$$= \frac{-400800+419040}{360720-338724} = \frac{18240}{21996} = 0.829 = 0.83$$

We are given : $\overline{X}_1 = 7, \overline{X}_2 = 50, \overline{X}_3 = 8$

Substituting the values in the above equation, we get

$$X_1 - 7 = 0.30\,(X_2 - 50) + 0.83\,(X_3 - 8)$$

or $$X_1 - 7 = 0.30\,X_2 - 15 + 0.83\,X_3 - 6.64$$

$\therefore X_1 = 0.30\,X_2 + 0.83\,X_3 - 14.64$ is the required equation

When $X_2 = 20$ and $X_3 = 30$

$$X_1 = 0.30(20) + 0.83(30) - 14.64 = 6 + 24.9 - 14.64 = 16.26$$

Example 4. **The following data for three variables X_1, X_2 and X_3 are given below**

$$\Sigma x_1 x_2 = 218, \quad \Sigma x_1 x_3 = -198, \quad \Sigma x_2 x_3 = -93$$

$$\Sigma x_1^2 = 454, \quad \Sigma x_2^2 = 110, \quad \Sigma x_3^2 = 90$$

x_1, x_2 and x_3 are measured from their means. Find the two partial r… coefficients ($b_{12.3}$ and $b_{13.2}$).

Solution.

$$b_{12.3} = \frac{(\Sigma x_1 x_2)(\Sigma x_3^2) - (\Sigma x_1 x_3)(\Sigma x_2 x_3)}{(\Sigma x_2^2)(\Sigma x_3^2) - (\Sigma x_2 x_3)^2}$$

$$= \frac{(218)(90) - (-198)(-93)}{(90)(110) - (-93)^2}$$

$$= \frac{19620 - 18414}{9900 - 8649}$$

$$= \frac{1206}{1251} = 0.964$$

$$b_{13.2} = \frac{(\Sigma x_1 x_3)(\Sigma x_2^2) - (\Sigma x_1 x_2)(\Sigma x_3 x_2)}{(\Sigma x_3^2)(\Sigma x_2^2) - (\Sigma x_3 x_2)^2}$$

$$= \frac{(-198)(110) - (218)(-93)}{(90)(110) - (-93)^2}$$

$$= \frac{-21780 + 20274}{9900 - 8649} = \frac{-1506}{1251} = -1.203$$

EXERCISE – 3

1. For the following set of data, find the multiple regression of X_1 on X_2 and X_3 using mean method. Also predict the value of X_1 when $X_2 = 5$ and $X_3 = 7$:

X_1:	12	24	32	28
X_2:	6	12	16	22
X_3:	4	6	12	18

[Ans. $X_1 = 2.577 + 1.661\,X_2 + 0.0169\,X_3$, X_1

2. From the data given below, find the multiple regression equation of X_1 on X_2 and X_3 actual mean method :

X_1:	4	6	7	9	13	15
X_2:	15	12	8	6	4	3
X_3:	30	24	20	14	10	4

[Ans. $X_1 = 16.479 + 0.389\,X_2 + 0.6$

3. From the data given below, find the multiple linear regression of X_1 on X_2 and X_3 using actual mean method :

X_1 :	18	20	17	14	21
X_2 :	38	40	25	28	44
X_3 :	20	15	5	12	18

[Ans. $X_1 = 0.5X_2 - 0.36X_3 + 5.54$]

4. Given the following information (variables are measured from their respective means):

$$\Sigma x_1 x_2 = 1900, \quad \Sigma x_1 x_3 = -20, \quad \Sigma x_2 x_3 = -50,$$
$$\Sigma x_1^2 = 1350, \quad \Sigma x_2^2 = 2800, \quad \Sigma x_3^2 = 24,$$
$$\overline{X}_1 = 65, \quad \overline{X}_2 = 55, \quad \overline{X}_3 = 30,$$

Obtain the partial regression coefficients ($b_{12.3}$ and $b_{13.2}$)
Also estimate the value of X_1 when $X_2 = 60$ and $X_3 = 25$.

[Ans. $b_{12.3} = 0.689$, $b_{13.2} = 0.603$, $X_1 = 65.43$]

5. Given the following information (variates are measured from their respective means) :

$$\Sigma x_1^2 = 1350, \ \Sigma x_2^2 = 2800, \ \Sigma x_3^2 = 24$$
$$\Sigma x_1 x_2 = 1900, \ \Sigma x_1 x_3 = -20, \ \Sigma x_2 x_3 = -50$$

Determine the regression equation of X_1 on X_2 and X_3. [Ans. $X_1 = 0.689X_2 + .603X_3$]

(2) Multiple Regression Equations in terms of Simple Correlation Coefficients

When the values of $\overline{X}_1, \overline{X}_2$ and $\overline{X}_3, \sigma_1, \sigma_2$ and σ_3 and r_{12}, r_{13} and r_{23} are given, then the multiple regression equations are expressed in the following manner :

(1) Multiple Regression Equation of X_1 on X_2 and X_3

$$X_1 - \overline{X}_1 = b_{12.3}(X_2 - \overline{X}_2) + b_{13.2}(X_3 - \overline{X}_3)$$

or $$x_1 = b_{12.3}\,x_2 + b_{13.2}\,x_3$$ Where, $x_1 = X_1 - \overline{X}_1$, $x_2 = X_2 - \overline{X}_2$, $x_3 = X_3 - \overline{X}_3$

The values of partial regression coefficients $b_{12.3}$ and $b_{13.2}$ are determined by using the following formulae :

$$b_{12.3} = \left[\frac{\sigma_1}{\sigma_2}\right] \cdot \left[\frac{r_{12} - r_{13} \cdot r_{23}}{1 - r_{23}^2}\right]$$

$$b_{13.2} = \left[\frac{\sigma_1}{\sigma_3}\right] \cdot \left[\frac{r_{13} - r_{12} \cdot r_{32}}{1 - r_{32}^2}\right]$$

Multiple Regression Equation of X_1 on X_2 and X_3 can also be written as :

$$x_1 = \left[\frac{\sigma_1}{\sigma_2}\right]\left[\frac{r_{12} - r_{13} \cdot r_{23}}{1 - r_{23}^2}\right] x_2 + \left[\frac{\sigma_1}{\sigma_3}\right]\left[\frac{r_{13} - r_{12} \cdot r_{32}}{1 - r_{32}^2}\right] x_3$$

or $$X_1 - \overline{X}_1 = \left[\frac{\sigma_1}{\sigma_2}\right]\left[\frac{r_{12} - r_{13} \cdot r_{23}}{1 - r_{23}^2}\right](X_2 - \overline{X}_2) + \left[\frac{\sigma_1}{\sigma_3}\right]\left[\frac{r_{13} - r_{12} \cdot r_{32}}{1 - r_{23}^2}\right](X_3 - \overline{X}_3)$$

(2) Multiple Regression Equation of X_2 on X_1 and X_3 :

$$X_2 - \overline{X}_2 = b_{21.3}(X_1 - \overline{X}_1) + b_{23.1}(X_3 - \overline{X}_3)$$

or $$x_2 = b_{21.3}x_1 + b_{23.1}x_3$$

Where,

$$b_{21.3} = \left[\frac{\sigma_2}{\sigma_1}\right]\left[\frac{r_{21} - r_{23} \cdot r_{13}}{1 - r_{13}^2}\right]$$

$$b_{23.1} = \left[\frac{\sigma_2}{\sigma_3}\right]\left[\frac{r_{23} - r_{21} \cdot r_{31}}{1 - r_{31}^2}\right]$$

Multiple Regression Equation of X_2 on X_1 and X_3 can also be written :

$$x_2 = \left[\frac{\sigma_2}{\sigma_1}\right]\left[\frac{r_{21} - r_{23} \cdot r_{13}}{1 - r_{13}^2}\right]x_1 + \left[\frac{\sigma_2}{\sigma_3}\right]\left[\frac{r_{23} - r_{21} \cdot r_{31}}{1 - r_{31}^2}\right]x_3$$

or $$X_2 - \overline{X}_1 = \left[\frac{\sigma_2}{\sigma_1}\right]\left[\frac{r_{21} - r_{23} \cdot r_{13}}{1 - r_{13}^2}\right](X_1 - \overline{X}_1) + \left[\frac{\sigma_2}{\sigma_3}\right]\left[\frac{r_{23} - r_{21} \cdot r_{31}}{1 - r_{31}^2}\right](X_3 - \overline{X}_3)$$

(3) Multiple Regression Equation of X_3 on X_1 and X_2 :

$$X_3 - \overline{X}_3 = b_{31.2}(X_1 - \overline{X}_1) + b_{32.1}(X_2 - \overline{X}_2)$$

or $$x_3 = b_{31.2}\, x_1 + b_{32.1}\, x_2$$

Where $$b_{31.2} = \left[\frac{\sigma_3}{\sigma_1}\right]\left[\frac{r_{31} - r_{32} \cdot r_{12}}{1 - r_{12}^2}\right]$$

$$b_{32.1} = \left[\frac{\sigma_3}{\sigma_2}\right]\left[\frac{r_{32} - r_{31} \cdot r_{21}}{1 - r_{21}^2}\right]$$

Multiple regression on X_3 on X_1 and X_2 can also be written as :

$$x_3 = \left[\frac{\sigma_3}{\sigma_1}\right]\left[\frac{r_{31} - r_{32} \cdot r_{12}}{1 - r_{12}^2}\right]x_1 + \left[\frac{\sigma_3}{\sigma_2}\right]\left[\frac{r_{32} - r_{31} \cdot r_{21}}{1 - r_{21}^2}\right]x_2$$

or $$X_3 - \overline{X}_3 = \left[\frac{\sigma_3}{\sigma_1}\right]\left[\frac{r_{31} - r_{32} \cdot r_{12}}{1 - r_{12}^2}\right](X_1 - \overline{X}_1) + \left[\frac{\sigma_3}{\sigma_2}\right]\left[\frac{r_{32} - r_{31} \cdot r_{21}}{1 - r_{21}^2}\right](X_2 - \overline{X}_2)$$

Note : $r_{12} = r_{21}$, $r_{23} = r_{32}$, $r_{13} = r_{31}$.

The following examples would clarify the procedure :

Example 5. ***A teacher in mathematics wishes to determine the relationship of marks in examination to those in two tests given during the semester. Calling X_1 and X_3, the marks of a student on Ist, 2nd and final examination respectively he made the following computations from a total of 120 students :***

$$\overline{X}_1 = 6.8 \quad \overline{X}_2 = 7.0 \quad \overline{X}_3 = 74$$

$$\sigma_1 = 1.0 \quad \sigma_2 = 0.80 \quad \sigma_3 = 9.0$$

$$r_{12} = 0.60 \quad r_{13} = 0.70 \quad r_{23} = 0.65$$

(i) Find the relevant regression equation.

(ii) Estimate the final marks of two students who secured respectively 9 and 8 on the two tests.

Solution. The relevant least square regression equation will be X_3 on X_1 and X_2 which is given by :

$$X_3 - \overline{X}_3 = b_{31.2}(X_1 - \overline{X}_1) + b_{32.1}(X_2 - \overline{X}_2)$$

$$b_{31.2} = \frac{\sigma_3}{\sigma_1}\cdot\left[\frac{r_{31} - r_{32}\cdot r_{12}}{1 - r_{12}^2}\right]$$

$$= \frac{9}{1}\times\left[\frac{(.70)-(.65)(.60)}{1-(.60)^2}\right]$$

$$= 9\times\left[\frac{(.70)-(.39)}{1-(.60)^2}\right] = 9\times\left[\frac{.31}{.64}\right] = \frac{2.79}{.64} = 4.36$$

$$b_{32.1} = \frac{\sigma_3}{\sigma_2}\cdot\left[\frac{r_{32} - r_{31}\cdot r_{21}}{1 - r_{21}^2}\right]$$

$$= \frac{9}{.80}\times\left[\frac{(.65)-(.70)(.60)}{1-(.60)^2}\right]$$

$$= \frac{9}{.80}\times\left[\frac{.65-.42}{0.64}\right] = 4.04$$

Thus, the regression equation of X_3 on X_1 and X_2 is

$$X_3 - 74 = 4.36(X_1 - 6.8) + 4.04(X_2 - 7)$$

$$\therefore \quad X_3 = 16.07 + 4.36X_1 + 4.04X_2$$

Final marks of students who scored 9 and 7 marks :

When $X_1 = 9$ and $X_2 = 7$

$$X_3 = 16.07 + 4.36(9) + 4.04(7)$$

$$= 16.07 + 39.24 + 28.28 = 83.59 \text{ or } 84$$

Final marks of students who scored 4 and 8 marks

When $X_1 = 4$ and $X_2 = 8$

$$X_3 = 16.07 + 4.36(4) + 4.04(8)$$

$$= 16.07 + 17.44 + 32.32 = 65.8 \text{ or } 66$$

Example 6. **Given the following, determine the regression equations of :**

(i) x_1 on x_2 and x_3 and

(ii) x_2 on x_1 and x_3 when the variates are measured from their means :

$$r_{12} = 0\cdot 8 \quad r_{13} = 0\cdot 6 \quad r_{23} = 0\cdot 5$$

$$\sigma_1 = 10, \sigma_2 = 8, \sigma_3 = 5$$

Solution. (i) The regression equation of x_1 on x_2 and x_3 when variates are measured from means is given by :

$$x_1 = b_{12\cdot 3}x_2 + b_{13\cdot 2}x_2 \quad \text{where , } x_1 = X_1 - \overline{X}_1, x_2 = X_2 - \overline{X}_2, x_3 = X_3 - \overline{X}_3$$

Scanned by TapScanner

$$b_{12\cdot3}=\left[\frac{\sigma_1}{\sigma_2}\right]\times\left[\frac{r_{12}-r_{13}\cdot r_{23}}{1-r_{23}^2}\right]$$

$$=\left[\frac{10}{8}\right]\times\left[\frac{(0\cdot8)-(0\cdot6)(0\cdot5)}{1-(0\cdot5)^2}\right]=0\cdot833$$

$$b_{13\cdot2}=\left[\frac{\sigma_1}{\sigma_3}\right]\times\left[\frac{r_{13}-r_{12}\cdot r_{32}}{1-r_{32}^2}\right]$$

$$=\left[\frac{10}{5}\right]\times\left[\frac{(0\cdot6)-(0\cdot8)(0\cdot5)}{1-(0\cdot5)^2}\right]=0\cdot533$$

∴ The required regression equation is :

$$x_1=\cdot833\,x_2+\cdot533\,x_3$$

(ii) The regression equation of x_2 on x_1 and x_3 when variates are measured from means is given by :

$$x_2=b_{21\cdot3}\,x_1+b_{23\cdot1}\,x_3$$

$$b_{21\cdot3}=\left[\frac{\sigma_2}{\sigma_1}\right]\times\left[\frac{r_{21}-r_{23}\cdot r_{13}}{1-r_{13}^2}\right]$$

$$=\left[\frac{8}{10}\right]\times\left[\frac{(0\cdot8)-(\cdot5)(\cdot6)}{1-(\cdot6)^2}\right]=\cdot625$$

$$b_{23\cdot1}=\left[\frac{\sigma_2}{\sigma_3}\right]\times\left[\frac{r_{23}-r_{21}\cdot r_{31}}{1-r_{31}^2}\right]$$

$$=\left[\frac{8}{5}\right]\times\left[\frac{(\cdot5)-(0\cdot8)(0\cdot6)}{1-(\cdot6)^2}\right]=0\cdot05$$

∴ The required regression equation is :

$$x_2=\cdot625\,x_1+\cdot05\,x_3$$

Example 7. ***A random sample of 15 students of Basic Statistics course when observed for weights (X_1), age (X_2) and height (X_3) offered the following information :***

$$r_{12}=0\cdot8,\ r_{23}=0\cdot3,\ r_{13}=0\cdot5,\ S_1=8\cdot5,\ S_2=4\cdot5,\ S_3=2\cdot1$$

$$\overline{X}_1=70\text{ kg},\ \overline{X}_2=22\text{ yrs and }\overline{X}_3=160\text{ cms.}$$

Obtain :

(i) Multiple and partial correlation coefficients $R_{1\cdot23}$ and $r_{13\cdot2}$.

(ii) Multiple regression of X_1 on X_2 and X_3 and estimate the value of X_1 for $X_2=25$ yrs and $X_3=140$ cms.

Solution. (i) $$R_{1\cdot23}=\sqrt{\frac{r_{12}^2+r_{13}^2-2r_{12}\cdot r_{13}\cdot r_{23}}{1-r_{23}^2}}$$

$$= \sqrt{\frac{(0.8)^2 + (0.5)^2 - 2(0.8)(0.5)(0.3)}{1-(0.3)^2}}$$

$$= \sqrt{\frac{0.64 + 0.25 - 0.24}{0.91}} = \sqrt{\frac{0.65}{0.91}} = 0.8452$$

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

$$= \frac{(0.8) - (0.5)(0.3)}{\sqrt{1-(0.5)^2}\sqrt{1-(0.3)^2}}$$

$$= \frac{0.8 - 0.15}{\sqrt{.75}\sqrt{0.91}} = \frac{0.65}{0.8261} = 0.7868$$

(ii) Multiple Regression on X_1 on X_2 and X_3 is given by :

$$X_1 - \overline{X}_1 = b_{12.3}(X_2 - \overline{X}_2) + b_{13.2}(X_3 - \overline{X}_3)$$

$$b_{12.3} = \left[\frac{S_1}{S_2}\right] \times \left[\frac{r_{12} - r_{13} . r_{23}}{1 - r_{23}^2}\right]$$

$$= \left[\frac{8.5}{4.5}\right] \times \left[\frac{0.8 - (0.5)(0.3)}{1-(.3)^2}\right] = 1.349$$

$$b_{13.2} = \left[\frac{S_1}{S_3}\right] \times \left[\frac{r_{13} - r_{12} . r_{32}}{1 - r_{32}^2}\right]$$

$$= \left[\frac{8.5}{2.1}\right]\left[\frac{(0.5) - (0.8)(0.3)}{1-(0.3)^2}\right] = 1.156$$

Substituting the values in the equation, we get

$$X_1 - 70 = 1.349(X_2 - 22) + 1.156(X_3 - 160)$$
$$X_1 - 70 = 1.349X_2 - 29.678 + 1.156X_3 - 184.96$$
$$\therefore \quad X_1 = 1.349X_2 + 1.156X_3 - 144.638$$

Estimation of X_1 for $X_2 = 25$ and $X_3 = 140$:

When $X_2 = 25$ and $X_3 = 140$, $X_1 = 1.349(25) + 1.156(140) - 144.638$

$$= 33.725 + 161.84 - 144.638 = 50.927$$

Example 8. **In a trivariate distribution :**

$$\sigma_1 = 3, \sigma_2 = 4, \sigma_3 = 5$$
$$r_{23} = 0.4, r_{31} = 0.6, r_{12} = 0.7$$

(i) Compute $r_{23.1}$ and $R_{1.23}$

(ii) Determine the regression equation of x_1 on x_2 and x_3 if the variates are measured from their means :

Solution. (i)

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

$$= \frac{(0\cdot4)-(0\cdot7)(0\cdot6)}{\sqrt{1-(0\cdot7)^2}\,\sqrt{1-(0\cdot6)^2}}$$

$$= \frac{0\cdot4-(\cdot42)}{\sqrt{0\cdot51}\,\sqrt{0\cdot64}} = \frac{-0\cdot02}{0\cdot5713} = -0\cdot035$$

$$R_{1\cdot23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}\cdot r_{13}\cdot r_{23}}{1-r_{23}^2}}$$

$$= \sqrt{\frac{(0\cdot7)^2+(0\cdot6)^2-2\,(0\cdot7)\,(0\cdot6)\,(0\cdot4)}{1-(0\cdot4)^2}}$$

$$= \sqrt{\frac{0\cdot49+0\cdot36-0\cdot336}{0\cdot84}} = \sqrt{\frac{0\cdot514}{0\cdot84}} = 0\cdot782$$

(ii) The regression equation of x_1 on x_2 and x_3 when variates are measured mean is given by :

$$x_1 = b_{12\cdot3}x_2 + b_{13\cdot2}x_3 \qquad \text{where, } x_1 = X_1 - \overline{X}_1,\ x_2 = X_2 - \overline{X}_2,\ x_3 = X_3$$

$$b_{12\cdot3} = \left[\frac{\sigma_1}{\sigma_2}\right] \times \left[\frac{r_{12} - r_{13}\cdot r_{23}}{1-r_{23}^2}\right]$$

$$= \left[\frac{3}{4}\right]\left[\frac{(0\cdot7)-(0\cdot6)\,(0\cdot4)}{1-(0\cdot4)^2}\right]$$

$$= \frac{0\cdot75\times0\cdot46}{0\cdot84} = \frac{0\cdot345}{0\cdot84} = 0\cdot41$$

$$b_{13\cdot2} = \left[\frac{\sigma_1}{\sigma_3}\right] \times \left[\frac{r_{13} - r_{12}\cdot r_{32}}{1-r_{32}^2}\right]$$

$$= \left[\frac{3}{5}\right] \times \left[\frac{(0\cdot6)-(0\cdot7)\,(0\cdot4)}{1-(0\cdot4)^2}\right] = \frac{0\cdot6\times0\cdot32}{0\cdot84} = \frac{0\cdot192}{0\cdot84} = 0\cdot2$$

Thus, the required regression equation is :

$$x_1 = 0\cdot41x_2 + 0\cdot229x_3.$$

STANDARD ERROR OF ESTIMATE (OR RELIABILITY OF ESTIMATES) FOR MULTIPLE REGRESSION

The standard error of estimate measures the realiability of the estimates given by the mul regression equation. It shows to what extent the estimated valued given by the regres equations are closer to the actual values.

For three regression equations, there are three standard error or estimates :

(1) Standard Error of Estimate of X_1 on X_2 and X_3 $(S_{1\cdot23})$

(2) Standard Error of Estimate of X_2 on X_1 and X_3 $(S_{2\cdot13})$

(3) Standard Error of Estimate of X_3 on X_1 and X_2 $(S_{3\cdot12})$

Scanned by TapScanner

The formulae for calculating the standard error of estimates are given as follows :

$$S_{1\cdot23} = \sigma_1 \cdot \sqrt{\frac{1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

$$S_{2\cdot13} = \sigma_2 \cdot \sqrt{\frac{1 - r_{21}^2 - r_{23}^2 - r_{13}^2 + 2r_{21} \cdot r_{23} \cdot r_{13}}{1 - r_{13}^2}}$$

$$S_{3\cdot12} = \sigma_3 \cdot \sqrt{\frac{1 - r_{31}^2 - r_{32}^2 - r_{13}^2 + 2r_{31} \cdot r_{32} \cdot r_{12}}{1 - r_{12}^2}}$$

Example 9. **If $r_{12} = 0\cdot8$, $r_{13} = 0\cdot5$, $r_{23} = 0\cdot3$ and $S_1 = 8\cdot5$, compute the standard error of estimate of X_1 on X_2 and X_3.**

Solution. Standard Error of Estimate of X_1 on X_2 and X_3 is given by :

$$S_{1\cdot23} = \sigma_1 \cdot \sqrt{\frac{1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

$$= 8\cdot5 \sqrt{\frac{1 - (0\cdot8)^2 - (0\cdot5)^2 - (0\cdot3)^2 + 2\,(0\cdot8)\,(0\cdot5)\,(0\cdot3)}{1 - (0\cdot3)^2}} = 4\cdot543$$

Coefficient of Multiple Determination (R^2)

The coefficient of determination in multiple regression denoted by $R_{1\cdot23}^2$ is similar to the coefficient of determination r^2 in the simple linear regression. It represents the proportion (fraction) of the total variation in the dependent variable X_1 that has been explained by the independent variables (X_2 and X_3) in the multiple regression equation.

For example, if $R_{1\cdot23} = 0\cdot7252$, then $R_{1\cdot23}^2 = 0\cdot5259 = 0\cdot526$

The value of $R_{1\cdot23}^2 = 0\cdot526$ indicates that $52\cdot6\%$ variation in the dependent variable (X_1) are explained by the independent variables X_2 and X_3 in the multiple regression equation of X_1 on X_2 and X_3.

Example 10. **A random sample of 15 students of advanced course in statistics when observed for weight (X_1), age (X_2) and height (X_3) offered the following information :**

$r_{12} = 0\cdot8$, $r_{13} = 0\cdot5$, $r_{23} = 0\cdot3$

$S_1 = 8\cdot5$, $S_2 = 4\cdot5$ and $S_3 = 2\cdot1$

Find the following :

(a) Partial regression coefficient $b_{1\cdot23}$ and $b_{13\cdot2}$.

(b) Standard error of estimate $S_{1\cdot23}$.

(c) Correlation Coefficients $R_{1\cdot23}$ and $r_{12\cdot3}$.

(d) Multiple regression of X_1 on X_2 and X_3 when $\overline{X}_1 = 70$ kg, $\overline{X}_2 = 22$ years and $\overline{X}_3 = 150$ cm.

(e) Weight of a student (X_1) of 25 years of age and 140 cm in height.

Solution. Given : $r_{12} = 0\cdot8, r_{13} = 0\cdot5, r_{23} = 0\cdot3$

$$S_1 = 8\cdot5,\ S_2 = 4\cdot5,\ S_3 = 2\cdot1$$

(a)
$$b_{12\cdot3} = \left[\frac{S_1}{S_2}\right]\cdot\left[\frac{r_{12} - r_{13}\cdot r_{23}}{1 - r_{23}^2}\right]$$

$$= \left[\frac{8\cdot5}{4\cdot5}\right]\times\left[\frac{0\cdot8 - (0\cdot5)(0\cdot3)}{1 - (0\cdot3)^2}\right] = 1\cdot349$$

$$b_{13\cdot2} = \left[\frac{S_1}{S_3}\right]\times\left[\frac{r_{13} - r_{12}\cdot r_{32}}{1 - r_{32}^2}\right]$$

$$= \left[\frac{8\cdot5}{2\cdot1}\right]\times\left[\frac{(0\cdot5) - (0\cdot8)(0\cdot3)}{1 - (0\cdot3)^2}\right] = 1\cdot156$$

(b)
$$S_{1\cdot23} = S_1\sqrt{\frac{1 - r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12}\cdot r_{13}\cdot r_{23}}{1 - r_{23}^2}}$$

$$= 8\cdot5\times\sqrt{\frac{1 - (0\cdot8)^2 - (0\cdot5)^2 - (0\cdot3)^2 + 2(0\cdot8)(0\cdot5)(0\cdot3)}{1 - (0\cdot3)^2}}$$

$$= 8\cdot5\times\sqrt{\frac{1 - \cdot64 - \cdot25 - 0\cdot09 + 0\cdot24}{0\cdot91}} = 4\cdot543$$

(c)
$$R_{1\cdot23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12}\cdot r_{13}\cdot r_{23}}{1 - r_{23}^2}}$$

$$= \sqrt{\frac{(0\cdot8)^2 + (0\cdot5)^2 - 2(0\cdot8)(0\cdot5)(0\cdot3)}{1 - (0\cdot3)^2}}$$

$$= \sqrt{\frac{0\cdot64 + \cdot25 - 0\cdot24}{0\cdot91}} = \sqrt{\frac{0\cdot65}{0\cdot91}} = 0\cdot8452$$

$$r_{12\cdot3} = \frac{r_{12} - r_{13}\cdot r_{23}}{\sqrt{1 - r_{13}^2}\ \sqrt{1 - r_{23}^2}}$$

$$= \frac{(0\cdot8) - (0\cdot5)\cdot(0\cdot3)}{\sqrt{1 - (0\cdot5)^2}\ \sqrt{1 - (0\cdot3)^2}} = \frac{0\cdot65}{0\cdot8261} = 0\cdot7868$$

(d) Multiple Regression Equation of X_1 on X_2 and X_3 :

$$X_1 - \overline{X}_1 = b_{12\cdot3}(X_2 - \overline{X}_2) + b_{13\cdot2}(X_3 - \overline{X}_3)$$

Substituting the values, we have

$$X_1 - 70 = 1\cdot349(X_2 - 22) + 1\cdot156(X_3 - 150)$$

$$X_1 = -133\cdot078 + 1\cdot349X_2 + 1\cdot156X_3$$

(e) For $X_2 = 25$ and $X_3 = 140$,

$$X_1 = -133\cdot078 + 1\cdot348(25) + 1\cdot156(140)$$

$$= -133\cdot078 + 33\cdot725 + 161\cdot84 = 62\cdot487$$

Scanned by TapScanner

Example 11. **Given the following data, determine the regression equation of x_1 on x_2 and x_3 if the variates are measured from their means :**

$$r_{12}=0\cdot 8, \quad r_{13}=0\cdot 6, \quad r_{23}=0\cdot 5$$
$$\sigma_1=10, \quad \sigma_2=8, \quad \sigma_3=15$$

Also find the standard error of the estimate of x_1 on x_2 and x_3.

Solution. The regression equation of x_1 on x_2 and x_3 is :

$$x_1=b_{12\cdot 3}\,x_2+b_{13\cdot 2}\,x_3 \quad \text{where, } x_1=X_1-\overline{X}_1,\ x_2=X_2-\overline{X}_2 \text{ and } x_3=X_3-\overline{X}_3$$

Here,
$$b_{12\cdot 3}=\left[\frac{\sigma_1}{\sigma_2}\right]\cdot\left[\frac{r_{12}-r_{13}\cdot r_{23}}{1-r_{23}^2}\right]$$
$$=\left[\frac{10}{8}\right]\times\left[\frac{0\cdot 8-(0\cdot 6)(0\cdot 5)}{1-(0\cdot 5)^2}\right]$$
$$=\left[\frac{10}{8}\right]\times\left[\frac{0\cdot 8-\cdot 30}{1-\cdot 25}\right]=\frac{10}{8}\times\frac{0\cdot 50}{0\cdot 75}=0\cdot 833$$

$$b_{13\cdot 2}=\frac{\sigma_1}{\sigma_3}\cdot\left[\frac{r_{13}-r_{12}\cdot r_{32}}{1-r_{32}^2}\right]$$
$$=\frac{10}{5}\times\left[\frac{(0\cdot 6)-(0\cdot 8)(0\cdot 5)}{1-(0\cdot 5)^2}\right]$$
$$=\frac{10}{5}\times\frac{0\cdot 20}{0\cdot 75}=\frac{2}{3\cdot 75}=0\cdot 53$$

Thus, regression equation of x_1 on x_2 and x_3 is :

$$x_1=0\cdot 833\,x_2+0\cdot 53\,x_3$$

Standard Error of Estimate of X_1 on X_2 and X_3

$$S_{1\cdot 23}=\sigma_1\sqrt{\frac{1-r_{12}^2-r_{13}^2-r_{23}^2+2r_{12}\cdot r_{13}\cdot r_{23}}{1-r_{23}^2}}$$
$$=10\,.\sqrt{\frac{1-(0\cdot 8)^2-(0\cdot 6)^2-(0\cdot 5)^2-2(0\cdot 8)(0\cdot 6)(0\cdot 5)}{1-(0\cdot 5)^2}}$$
$$=10\,.\sqrt{\frac{1-\cdot 64-\cdot 36-\cdot 25+\cdot 48}{0\cdot 75}}$$
$$=10\,.\sqrt{\frac{0\cdot 23}{0\cdot 75}}=10\times\cdot 5537=5\cdot 537$$

Example 11A. **The following values have been obtained from the measurement of three variables x_1, x_2 and x_3 :**

$$\overline{X}_1=6\cdot 8 \quad \overline{X}_2=7\cdot 0 \quad \overline{X}_3=7\cdot 4$$
$$S_1=1\cdot 0 \quad S_2=0\cdot 80 \quad S_3=0\cdot 90$$
$$r_{12}=0\cdot 60 \quad r_{13}=0\cdot 70 \quad r_{23}=0\cdot 65$$

(i) Obtain regression eqaution of X_1 on X_2 and X_3.
(ii) Estimate the value of X_1 for $X_2 = 10$ and $X_3 = 9$.
(iii) Find the coefficient of multiple determination $R^2_{1.23}$ from r_{12} and r_{13}...

Solution. The regression equation of X_1 on X_2 and X_3 is given by :

$$X_1 - \bar{X}_1 = b_{12.3}(X_2 - \bar{X}_2) + b_{13.2}(X_3 - \bar{X}_3) \quad \text{...(i)}$$

where,

$$b_{12.3} = \frac{s_1}{s_2}\left[\frac{r_{12} - r_{13} \cdot r_{23}}{1 - r_{23}^2}\right] = \frac{1}{0.80}\left[\frac{0.60 - 0.70 \times 0.65}{1 - (0.65)^2}\right]$$

or

$$b_{12.3} = (1.25)\left[\frac{0.60 - 0.455}{0.578}\right] = 0.313$$

$$b_{13.2} = \frac{s_1}{s_2} \cdot \left[\frac{r_{13} - r_{12} \cdot r_{32}}{1 - r_{32}^2}\right] = \frac{1}{0.90}\left[\frac{0.70 - 0.60 \times 0.65}{1 - (0.65)^2}\right]$$

$$= (1.111)\left[\frac{0.70 - 0.39}{0.578}\right] = 0.595$$

Substituting the values in equation (i), we have,

$$X_1 - 6.8 = 0.313(X_2 - 7.0) + 0.595(X_3 - 7.4)$$

or

$$X_1 = 0.206 + 0.313 X_2 + 0.595 X_3$$

(ii) Substituting for $X_2 = 10$ and $X_3 = 9$ in the above regression and solving fo...

$$X_1 = 0.206 + 0.313(10) + 0.595(9) = 8.691$$

(iii) Multiple and partial correlation coefficients are related as :

$$R^2_{12.3} = 1 - (1 - r_{12}^2)(1 - r_{13.2}^2)$$

$$r_{13.2} = \frac{r_{13} - r_{12} \cdot r_{32}}{\sqrt{1 - r_{12}^2}\sqrt{1 - r_{32}^2}} = \frac{0.70 - 0.60 \times 0.65}{\sqrt{1 - (0.60)^2}\sqrt{1 - (0.65)^2}}$$

$$= \frac{0.70 - 0.39}{0.8 \times 0.760} = 0.509$$

or,

$$r_{13.2}^2 = 0.259$$

Substituting the values of r_{12}^2 and $r_{13.2}^2$ for $R^2_{1.23}$ we have

$$R^2_{1.23} = 1 - (1 - 0.36)(1 - 0.259) = 0.526$$

EXERCISE – 4

1. The following constants are obtained from measurements of length in mm (x_1), volu... c.c. (x_2) and weight in gm (x_3) of 300 eggs :

$\bar{X}_1 = 55.95$	$S_1 = 2.26$	$r_{12} = 0.578$
$\bar{X}_2 = 51.48$	$S_2 = 4.39$	$r_{13} = 0.581$
$\bar{X}_3 = 56.03$	$S_3 = 4.41$	$r_{23} = 0.974$

Obtain the linear regression equation of egg weight on egg length and egg volume. Hence estimate the weight of an egg whose length is 58 mm and volume is 52·5 c.c.

[Ans. $X_3 = 3{\cdot}54 + 0{\cdot}052X_1 + 0{\cdot}963X_2, X_3 = 57{\cdot}11$ gms.]

2. In a trivarlate distribution :

$\sigma_1 = 2{\cdot}7,\quad \sigma_2 = 2{\cdot}4,\quad \sigma_3 = 2{\cdot}7$
$r_{12} = 0{\cdot}28,\quad r_{23} = 0{\cdot}49,\quad r_{31} = 0{\cdot}51$

Determine the regression equation of x_3 on x_1 and x_2 if the variates are measured from their means.

[Ans. $x_3 = 0{\cdot}405x_1 + 0{\cdot}424x_2$]

3. Given the following data :

$\overline{X}_1 = 6,\quad \overline{X}_2 = 7,\quad \overline{X}_3 = 8$
$\sigma_1 = 1,\quad \sigma_2 = 2,\quad \sigma_3 = 3$
$r_{12} = 0{\cdot}6,\quad r_{13} = 0{\cdot}7,\quad r_{23} = 0{\cdot}8$

Obtain the linear regression equation of X_3 on X_1 and X_3. Hence estimate X_3 when $X_1 = 4$ and $X_2 = 5$.

[Ans. $x_3 = -4{\cdot}41 + 1{\cdot}03x_1 + 0{\cdot}89x_2, x_3 = 4{\cdot}16$]

4. The following results were obtained in the analysis of a trivariate distribution :

$S_1 = 3, S_2 = S_3 = 5, r_{12} = 0{\cdot}7, r_{23} = r_{31} = 0{\cdot}6$

Find (i) partial correlation coefficient $r_{12\cdot3}$ (ii) Multiple correlation coefficient $R_{1\cdot23}$ and (iii) the partial regression coefficients ($b_{12\cdot3}$ and $b_{13\cdot2}$).

[Ans. $r_{12\cdot3} = 0{\cdot}531, R_{1\cdot23} = 0{\cdot}735, b_{12\cdot3} = 0{\cdot}319, b_{13\cdot2} = 0{\cdot}169$]

5. If $r_{12} = 0{\cdot}926, r_{13} = 0{\cdot}891, r_{23} = 0{\cdot}955$ and $S_1 = 1{\cdot}51$, compute the standard error of estiamte of X_1 on X_2 and X_3 ($S_{1\cdot23}$).

[Ans. $S_{1\cdot23} = 0{\cdot}5702$]

6. A random sample of 50 students of M. Com. when observed for weight (x_1), age (x_2), and height (x_3) offered the following information :

$r_{12} = 0{\cdot}7,\quad r_{13} = 0{\cdot}8,\quad r_{23} = 0{\cdot}5$
$S_1 = 5{\cdot}6,\quad S_2 = 4{\cdot}5,\quad S_3 = 3{\cdot}5$

Obtain the following :

(a) Partial regression coefficient $b_{12\cdot3}$ and $b_{13\cdot2}$.
(b) Standard error of estimate $S_{1\cdot23}$.
(c) Coefficient of multiple correlation ($R_{1\cdot23}$).
(d) Coefficient of partial correlation ($r_{12\cdot3}$).

[Ans. $b_{12\cdot3} = 0{\cdot}496, b_{13\cdot2} = 0{\cdot}96, S_{1\cdot23} = 2{\cdot}74, R_{1\cdot23} = 0{\cdot}87, r_{12\cdot3} = 0{\cdot}516$]

7. Given the following data, determine the regression equation of x_1 on x_2 and x_3 if the variates measured from their means :

$r_{12} = 0{\cdot}8,\quad r_{13} = 0{\cdot}6,\quad r_{23} = 0{\cdot}5$
$\sigma_1 = 10,\quad \sigma_2 = 8,\quad \sigma_3 = 5$

Also find the standard error of estimate of x_1 on x_2 and x_3.

[Ans. $x_1 = 0{\cdot}833x_2 + 0{\cdot}53x_2, S_{1\cdot23} = 5{\cdot}537$]

8. Given the following data, calculate the estimated value of X_1 when $X_2 = 20$ and $X_3 = 25$.

$\overline{X}_1 = 30\quad S_1 = 5\quad r_{12} = -0{\cdot}4$
$\overline{X}_2 = 35\quad S_2 = 10\quad r_{13} = -0{\cdot}5$
$\overline{X}_3 = 40\quad S_3 = 15\quad r_{23} = 0{\cdot}6$

Also calculate the standard error of estimate of X_1 on X_2 and X_3.

[Ans. $X_3 = 38{\cdot}225 - 0{\cdot}075X_2 - 0{\cdot}14X_3, X_1 = 33, S_{1\cdot23} = 1{\cdot}1$]

9. Given the following data :

$$\overline{X}_1 = 55 \quad S_1 = 5 \quad r_{12} = 0.57$$
$$\overline{X}_2 = 51 \quad S_2 = 7 \quad r_{13} = 0.58$$
$$\overline{X}_3 = 56 \quad S_3 = 9 \quad r_{23} = 0.97$$

Calculate :

(i) Multiple Regression of X_3 on X_1 and X_2.

(ii) Multiple Correlation coefficient $R_{1.23}$

[Ans. $X_3 = 0.08X_1 + 1.21X_2 - 10.11$, $R_{1.23}$ ...

10. From the following data,

$$r_{12} = 0.8, r_{13} = 0.5, r_{23} = 0.3$$
$$S_1 = 8.5, S_2 = 4.5 \text{ and } S_3 = 2.1$$

(i) Obtain the regression equation of X_1 on X_2 and X_3 with $\overline{X}_1 = 70$ kg, $\overline{X}_2 = 22$ years, $\overline{X}_3 = 150$ cm.

(ii) Estimate the value of X_1 on $X_2 = 25$ years and $X_3 = 140$ cm, and

(iii) Find the coefficient of multiple determination $R^2_{1.23}$ from r_{12} and $r_{13.2}$. What does it indicate ?

[Ans. $X_1 = 1.349X_2 + 1.156X_3 - 133.078$, $X_1 = 62.487$ $R^2_{1.23} = 0.7443$, R^2 indicates the 74.43% variation in X_1 are explained by the multiple regression equation]

MISCELLANEOUS SOLVED EXAMPLE

Example 12. In a trivariate distribution :

$$\overline{X}_1 = 28.02, \overline{X}_2 = 4.91, \overline{X}_3 = 594, S_1 = 4.4, S_2 = 1.1, S_3 = 80$$
$$r_{12} = 0.80, r_{23} = -0.56, r_{31} = -0.40$$

(i) Find the correlation coefficient $r_{23.1}$ and $R_{1.23}$.

(ii) Estimate the value of X_1 when $X_2 = 6.0$ and $X_3 = 650$.

Solution.

(i)
$$r_{23.1} = \frac{r_{23} - r_{21} \cdot r_{31}}{\sqrt{1 - r_{21}^2}\sqrt{1 - r_{31}^2}}$$

Substituting the values, we get

$$r_{23.1} = \frac{-0.56 - (0.80).(-0.40)}{\sqrt{1-(0.80)^2}\sqrt{1-(-0.40)^2}}$$
$$= \frac{-0.56 + .32}{\sqrt{1-.64}\sqrt{1-.16}} = \frac{-0.24}{0.6 \times 0.916} = -0.436$$

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 + 2r_{12} \cdot r_{13} \cdot r_{23}}{1 - r_{23}^2}}$$

Substituting values, we get

$$R_{1.23} = \sqrt{\frac{(0.80)^2 + (-0.40)^2 - 2(0.80)(-0.40)(-0.56)}{1-(-.56)^2}}$$

$$= \sqrt{\frac{0.64 + .16 - .3584}{1 - .3136}} = \sqrt{\frac{0.4416}{0.6864}} = 0.802$$

(ii) The linear regression equation of X_1 on X_2 and X_3 is :

$$X_1 - \overline{X}_1 = b_{12.3}(X_2 - \overline{X}_2) + b_{13.2}(X_3 - \overline{X}_3)$$

$$b_{12.3} = \frac{S_1}{S_2} \times \left[\frac{r_{12} - r_{13} . r_{23}}{1 - r_{23}^2}\right]$$

$$= \frac{4.4}{1.1} . \left[\frac{0.80 - 0.224}{0.6864}\right] = \frac{4.4}{1.1} . \left[\frac{0.576}{0.6864}\right] = 3.357$$

$$b_{13.2} = \frac{S_1}{S_3} . \left[\frac{r_{13} - r_{12} . r_{32}}{1 - r_{32}^2}\right] = \frac{4.4}{80} . \left[\frac{-0.40 - (0.80)(-0.56)}{1 - (-.56)^2}\right]$$

$$= \frac{4.4}{80} . \left[\frac{-0.40 + 0.448}{0.6884}\right] = \frac{4.4}{80} . \left[\frac{0.048}{0.6864}\right] = 0.0038 = 0.004$$

Substituting the values in the equation, we get

$$X_1 - 28.02 = 3.357(X_2 - 4.91) + .004(X_3 - 594)$$

or $$X_1 - 28.02 = 3.357 X_2 - 16.4828 + .004 X_3 - 2.376$$

$$X_1 = 3.357 X_2 + .004 X_3 + 9.1612$$

(iv) Estimation of X_1 for X_2 and X_3 :

When $X_2 = 6.00$, $X_3 = 650$, $X_1 = 3.357(6.00) + .004(650) + 9.1612$

$$= 20.142 + 2.6 + 9.1612$$

$$= 31.9032$$

IMPORTANT FORMULAE

Multiple Correlation Coefficients :

$$R_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2r_{12} . r_{13} . r_{23}}{1 - r_{23}^2}}$$

$$R_{2.13} = \sqrt{\frac{r_{21}^2 + r_{23}^2 - 2r_{21} . r_{23} . r_{13}}{1 - r_{13}^2}}$$

$$R_{3.12} = \sqrt{\frac{r_{31}^2 + r_{32}^2 - 2r_{31} . r_{32} . r_{12}}{1 - r_{12}^2}}$$

Partial Correlation Coefficients :

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

$$r_{23\cdot1} = \frac{r_{23} - r_{21} \cdot r_{31}}{\sqrt{1-r_{21}^2}\sqrt{1-r_{31}^2}}$$

$$r_{13\cdot2} = \frac{r_{13} - r_{12} \cdot r_{32}}{\sqrt{1-r_{12}^2}\sqrt{1-r_{32}^2}}$$

Relationship between Simple, Partial and Multiple Correlation Coefficients

$$1 - R_{1\cdot23}^2 = (1-r_{12}^2)(1-r_{13\cdot2}^2)$$ and

$$1 - R_{2\cdot13}^2 = (1-r_{21}^2)(1-r_{23\cdot1}^2)$$

$$1 - R_{3\cdot12}^2 = (1-r_{21}^2)(1-r_{32\cdot1}^2)$$

Multiple Regression of X_1 on X_2 and X_3 :

$$X_1 = a_{1\cdot23} + b_{12\cdot3} X_2 + b_{13\cdot2} X_3$$

Where, $$b_{12\cdot3} = \left[\frac{\sigma_1}{\sigma_2}\right] \times \left[\frac{r_{12} - r_{13} \cdot r_{23}}{1-r_{23}^2}\right]$$

Standard Error of Estimate :

$$S_{1\cdot23} = \sigma_1 \cdot \sqrt{\frac{1-r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12} \cdot r_{13} \cdot r_{23}}{1-r_{23}^2}}$$

QUESTIONS

1. Distinguish between partial correlation and multiple correlation.
2. Write down the expression for $r_{12\cdot3}$ and $R_{1\cdot23}$ in terms of r_{12}, r_{13} and r_{23}. Also state limits within which $r_{12\cdot3}$ and $R_{1\cdot23}$ must lie.
3. Explain the concept of multiple regression and discuss its utility in business.
4. How will you fit a multiple regression equation of X_1 on X_2 and X_3 ?
5. Write the normal equations in case of multiple linear regression of X_1 on X_2 and X_3.
6. Write a short note on "Standard Error of Estiamte" for multiple regression.
7. Define simple, partial and multiple correlation coefficients and find their relationship with one another.
8. Write the equations/formulae to calculate the followings :
   (i) Regression equation for X_2 on X_1 and X_3.
   (ii) Partial regression coefficients ($b_{12\cdot3}$ and $b_{13\cdot2}$).
   (iii) Standard error of estimate of X_1 on X_2 and X_3 ($S_{1\cdot23}$)
   (iv) Multiple correlation coefficient ($R_{1\cdot23}$)
   (v) Partial correlation coefficient ($r_{12\cdot3}$)
9. How will you interpret the value of R^2 in a multiple regression equation ?

❀❀❀

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

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

Relationship between Simple, Partial and Multiple Correlation Coefficients

$$1 - R_{1.23}^2 = (1-r_{12}^2)(1-r_{13.2}^2)$$ and

$$1 - R_{2.13}^2 = (1-r_{21}^2)(1-r_{23.1}^2)$$

$$1 - R_{3.12}^2 = (1-r_{21}^2)(1-r_{32.1}^2)$$

...ltiple Regression of X_1 on X_2 and X_3 :

$$X_1 = a_{1.23} + b_{12.3} X_2 + b_{13.2} X_3$$

...re, $$b_{12.3} = \left[\frac{\sigma_1}{\sigma_2}\right] \times \left[\frac{r_{12} - r_{13} \cdot r_{23}}{1-r_{23}^2}\right]$$

...rror of Estimate :

$$S_{1.23} = \sigma_1 \cdot \sqrt{\frac{1-r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12} \cdot r_{13} \cdot r_{23}}{1-r_{23}^2}}$$

QUESTIONS

...ween partial correlation and multiple correlation.

...expression for $r_{12.3}$ and $R_{1.23}$ in terms of r_{12}, r_{13} and r_{23}. Als...

...h $r_{12.3}$ and $R_{1.23}$ must lie.

...of multiple regression and discuss its utility in business.

...ltiple regression equation of X_1 on X_2 and X_3 ?

...ions in case of multiple linear regression of X_1 on X_2 and...

...andard Error of Estiamte" for multiple regression.

...multiple correlation coefficients and find their relatio...

... to calculate the followings :

... on X_1 and X_3.

... ($b_{12.3}$ and $b_{13.2}$).

... on X_2 and X_3 ($S_{1.23}$)

... ($R_{1.23}$)

...)

...in a multiple regression equation ?

2

Sampling & Sampling Distribution

INTRODUCTION

In all the spheres of life (such as Economic, Social and Business) the need for statistical investigation and data analysis is rising day by day. There are two methods of collection of statistical data : (i) **Census Method**, and (ii) **Sample Method.** Under census method, information relating to the entire field of investigation or units of population is collected; whereas under sample method, rather than collecting information about all the units of population, information relating to only selected units is collected. Before we make a detailed study of both the methods, we will explain some basic concepts related to them.

SOME BASIC CONCEPTS

(1) Universe or Population : In statistics, universe or population means an aggregate of items about which we obtain information. A universe or population means the entire field unde[r] investigation about which knowledge is sought. For example, if we want to collect informatio[n] about the average monthly expenditure of all the 2,000 students of a college, then the enti[re] aggregate of 2,000 students will be termed as Universe or Population. A population can be of tw[o] kinds (i) Finite and (ii) Infinite. In a finite population, number of items is definite such as, numb[er] of students or teachers in a college. On the other hand, an infinite population has infinite number items *e.g.*, number of stars in the sky, number of water drops in an ocean, number of leaves on a t[ree] or number of hairs on the head.

(2) Sample : A part of population is called sample. In other words, selected or sorted units fr[om] the population is known as a sample. In fact, a sample is that part of the population which we se[lect] for the purpose of investigation. For example, if an investigator selects 200 students from ... students of a college who represent all of them, then these 200 students will be termed as a sam[ple]. Thus, sample means some units selected out of a population which represent it.

CENSUS AND SAMPLE METHODS

There are two methods to collect statistical data :

(1) Census Method

(2) Sample Method

(1) Census Method

Census method is that method in which information or data is collected from each an[d every] unit of the population relating to the problem under investigation and conclusions are d[rawn on] their basis. This method is also called as Complete Enumeration Method. For example,