SMART NOTES

A game of chess is a classic example of the topic Permutation.

It involves numerous permutations as a game unfolds.

Std. XI

Mathematics and Statistics Commerce (Part - 2)

Target Publications® Pvt. Ltd.

Written as per the latest textbook prescribed by the Maharashtra State Bureau of Textbook Production and Curriculum Research, Pune.

SMART NOTES

MATHEMATICS & STATISTICS – COMMERCE (PART - II)

Std. XI

Salient Features

- Written as per the latest textbook
- Exhaustive coverage of entire syllabus
- Precise theory for every topic
- Covers answers to all exercises and miscellaneous exercises given in the textbook.
- All derivations and theorems covered
- Includes additional problems for practice
- Contains One mark questions in the form of MCQs, True or False and Fill in the blanks
- Illustrative examples for selective problems
- Recap of important formulae at the end of the book
- Activity Based Questions covered in every chapter
- Smart Check to enable easy rechecking of solutions

Printed at: **India Printing Works,** Mumbai

Balbharati Registration No.: 2018MH0022

TEID: 3176

P.O. No. 8712

PREFACE

"The only way to learn Mathematics is to do Mathematics" – Paul Halmos

"Mathematics & Statistics – Commerce (Part – II): Std. XI" forms a part of **'Smart Notes'** prepared as per the **Latest Textbook**. It is a complete and thorough guide critically analysed and extensively drafted to boost the students' confidence.

The book provides **answers to all textbook questions** included in exercises as well as miscellaneous exercises. Apart from these questions, we have provided **ample questions for additional practice** to students based on every exercise of the textbook. Only the final answer has been provided for such additional practice questions.

Precise theory has been provided at the required places for better understanding of concepts. Further, all **derivations and theorems have been covered** wherever required. A **recap of all important formulae** has been provided at the end of the book for quick revision. This book features **activity-based questions** and **one-mark questions** in each chapter. **'Smart Check'** has been incorporated to assist students in comprehending the process of verifying the accuracy of their answers.

The journey to create a complete book is strewn with triumphs, failures and near misses. If you think we've nearly missed something or want to applaud us for our triumphs, we'd love to hear from you. Please write to us on: mail@targetpublications.org

A book affects eternity; one can never tell where its influence stops.

Best of luck to all the aspirants!

Publisher

Edition: Third

KEY FEATURES

Illustrative Example

Illustrative Example provides a detailed approach towards solving a problem.

One Mark Questions

In this section we have provided multiple choice questions, True or False and Fill in the blanks for practice which will help students test their concept clarity.

Additional Problems for Practice

In this section, we have provided additional problems which will help students to test their understanding of the chapter.

Activities for Practice

In this section we have provided multiple activities for practice which will help students understand the concepts.

Important formulae

Important Formulae given at the end of the book includes all the key formulae in the chapter.
It offers students a handy tool to solve problems and ace the last minute revision.

INDEX

Note: Solved examples from textbook are indicated by "+".
Smart check is indicated by ✓ symbol.

5 Correlation

Syllabus

- Concept of Correlation
- Methods of computing correlation
- Properties of Covariance
- Karl Pearson's coefficient of correlation
- Scatter diagram
- Interpretation

Let's Study

Introduction

Statistical methods are commonly used to analyse data related to single variable. However, we come across certain situations where each individual or item possesses two criteria or characters represented by two variables. For e.g. income same and expenditure, demand and supply of a product, rainfall and agricultural production etc. Such statistical data related to observation of two variables is known as Bivariate data.

The variables are denoted by X and Y respectively and can also be represented by n ordered pairs (x_1, y_1), (x_2, y_2),…, (x_n, y_n).

The pair (x_i, y_i) denotes the values of the variables for i^{th} object.

Correlation

In bivariate data, variables are said to be correlated if changes in value of one variable tend to cause changes in value of the other variable.

Correlation indicates such association or mutual relationship between two variables.

The correlated variables move in sympathy in the same direction or in the opposite direction.

Some examples of correlated variables are as follows:

i. Price and demand of a commodity
ii. Income and expenditure of an individual.

Covariance

Covariance is a measure of joint variation between two variables X and Y.

Let the variable X take the values x_1, x_2, …., x_n and Y take the values y_1, y_2, …., y_n. Then the n ordered pairs for values of X and Y are (x_1, y_1), (x_2, y_2), …, (x_n, y_n).

Now, the covariance between these two variables, denoted by Cov (X, Y) is defined as

$$\text{Cov}(X, Y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\cdot\bar{y}$$

Another form of Cov (X, Y) is

$$\text{Cov}(X, Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}), \text{ where } \bar{x} = \frac{\sum x_i}{n}, \quad \bar{y} = \frac{\sum y_i}{n}$$

The above formula can be further simplified as follows:

$$\begin{aligned}
\text{Cov}(X, Y) &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n}(x_i y_i - x_i\bar{y} - \bar{x}y_i + \bar{x}\bar{y}) \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{y}\frac{1}{n}\sum_{i=1}^{n} x_i - \bar{x}\frac{1}{n}\sum_{i=1}^{n} y_i + \frac{\bar{x}\bar{y}}{n}\sum_{i=1}^{n} 1 \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\cdot\bar{y} - \bar{x}\cdot\bar{y} + \bar{x}\cdot\bar{y} \\
&= \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\cdot\bar{y}
\end{aligned}$$

Properties of Covariance:

1. Cov (X, Y) = Cov (Y, X)
2. Cov (X, C) = 0, where C is a constant
3. Cov (X, X) = Var (X)
4. Covariance may be positive, negative or zero.
5. Covariance is not affected by change of origin but is affected by change of scale.

   i.e., If $U = \frac{X - a}{h}$ and $V = \frac{Y - b}{k}$, where a, b, h, k are constants and h, k ≠ 0, then,

   $\text{Cov}(U, V) = \frac{1}{hk}\text{Cov}(X, Y)$

∴ Cov (X, Y) = hk Cov (U,V)

Methods of measuring correlation

The following methods are used for measuring correlation between quantitative variables:

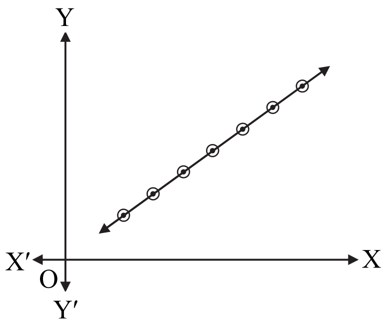
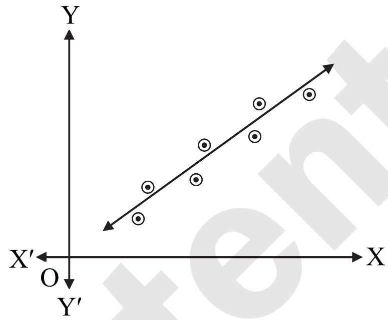
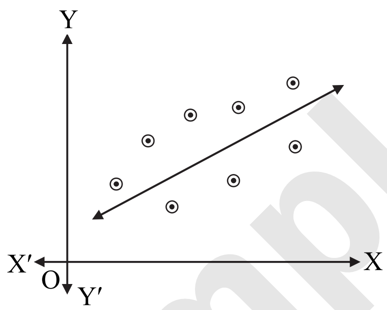
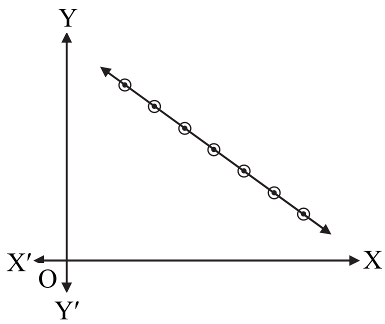
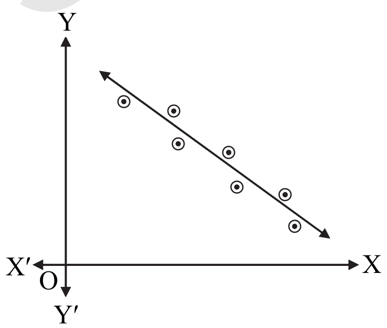
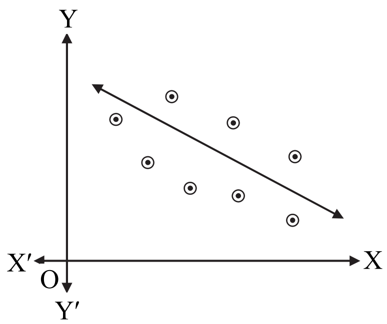
i. Scatter Diagram
ii. Karl Pearson's Method

Scatter diagram:

Scatter diagram is a graphical picture of the sample data to study correlation between two variables.

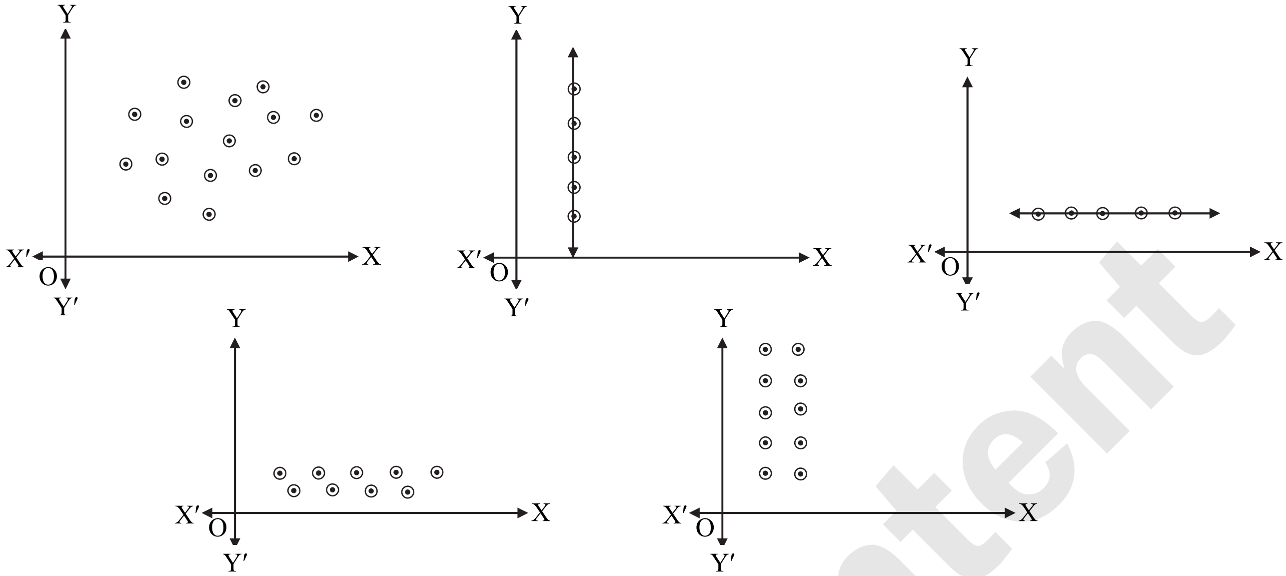
Let (x_1, y_1), (x_2, y_2), …, (x_n, y_n) be n pairs of values of two variables X and Y.

A graph in which these points are plotted by taking the values of X along the X-axis and values of Y-axis is called a scatter diagram. The scatter diagram so obtained can be analysed as follows:

i. Perfect positive correlation: If all the plotted points lie on a straight line rising from lower left point to upper right then the scatter diagram is said to indicate a perfect positive correlation. 	**ii. High degree positive correlation:** If all the plotted points do not lie on a straight line but show a tendency to rise from left lower point to upper right point and points form a band that is smaller in width then the correlation indicated is high degree positive correlation. 
iii. Low degree positive correlation: If all the plotted points do not lie on a line but show a tendency to rise from left lower point to right upper point and the points form a band that is larger in width then the correlation indicated is low degree positive correlation. 	**iv. Perfect negative correlation:** If all the plotted points lie on a line falling from left upper point to right lower point then the scatter diagram is said to indicate a perfect negative correlation. 
v. High degree negative correlation: If all the plotted points do not lie on a line but show a tendency to fall from left upper point to right lower point and the points form a band that is smaller in width then the correlation indicated is high degree negative correlation. 	**vi. Low degree negative correlation:** If all the plotted points do not lie on a line but show a tendency to fall from left upper point to right lower point and the points form a band that is larger in width then the scattered diagram indicated is low degree negative correlation. 

vii. No correlation (Zero correlation)

If all the plotted points are completely scattered all over, with no particular tendency, then there is no correlation or zero correlation.



Although the scatter diagram is the simplest method of measuring correlation, the method is very subjective in nature and the degree of correlation cannot be numerically expressed. This method is ideal for simply detecting correlation.

Karl Pearson's method:

The correlation between two variables is measured by using a mathematical method suggested by British Statistician Karl Pearson. In this method the degree of correlation between two variables is measured using a Correlation Co-efficient. It is denoted by r_{xy} or simply r.

$$r_{xy} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

where σ_X = standard deviation of series X.

σ_Y = standard deviation of series Y.

i.e. $\sigma_X = \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\bar{x}\right)^2}$

$\sigma_Y = \sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \left(\bar{y}\right)^2}$

Other forms of correlation coefficient:

i. $r_{xy} = \dfrac{\frac{1}{n}\sum_{i=1}^{n} x_i\, y_i - \bar{x} \cdot \bar{y}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2 - \left(\bar{x}\right)^2}\sqrt{\frac{1}{n}\sum_{i=1}^{n} y_i^2 - \left(\bar{y}\right)^2}}$ when $\bar{x}$, $\bar{y}$ are integers and small numbers.

ii. $r = \dfrac{n\sum_{i=1}^{n} x_i\, y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{\sqrt{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\sqrt{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2}}$

when $\bar{x}$, $\bar{y}$ are decimals and $\sum_{i=1}^{n} x_i$ and $\sum_{i=1}^{n} y_i$ are comparatively smaller numbers.

Properties of correlation co-efficient (r):

1. $-1 \le r \le 1$
2. Coefficient of correlation measures only linear correlation between X and Y.
3. If two variables X and Y are independent then r = 0
4. Coefficient of correlation between X and X is 1.
5. Coefficient of correlation is independent of shift of origin and change of scale.

6. If $u = \dfrac{x-a}{h}$ and $v = \dfrac{y-b}{k}$, where, a, b, h, k are constants and $h \neq 0$, $k \neq 0$

Then, $\text{Corr}(U, V) = \text{Corr}\left(\dfrac{x-a}{h}, \dfrac{y-b}{k}\right)$

$= \text{Corr}(X, Y)$ ….(if h and k have same algebraic sign)

$= -\text{Corr}(X, Y)$ ….(if h and k have opposite algebraic sign)

Interpretation of value of correlation coefficient:

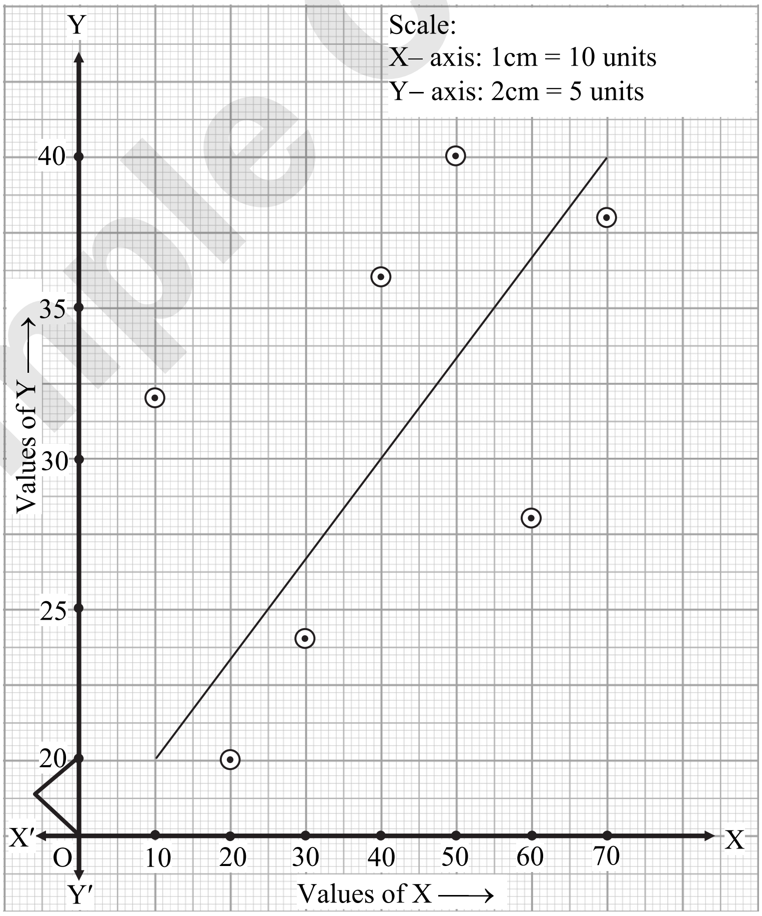
Value of r	Type of correlation between two variables
$r = 1$	Perfect positive correlation
$r = -1$	Perfect negative correlation
$r = 0$	No linear relation
$r > 0$	Positive correlation
$r < 0$	Negative correlation
$\lvert r \rvert > 0.8$	High correlation
$0.3 < \lvert r \rvert < 0.8$	Moderate correlation
$\lvert r \rvert < 0.3$	Insignificant or Poor correlation

Exercise 5.1

1. Draw scatter diagram for the data given below and interpret it.

x	10	20	30	40	50	60	70
y	32	20	24	36	40	28	38

***Solution*:**



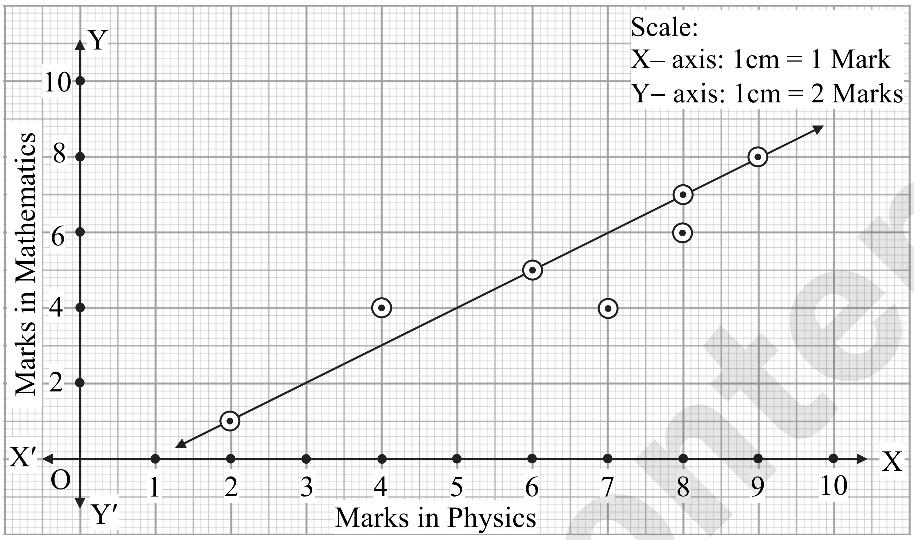
Since, all the points lie in a band rising from left to right.
Therefore, there is a positive correlation between the values of X and Y respectively.

2. **For the following data of marks of 7 students in Physics (x) and Mathematics (y), draw scatter diagram and state the type of correlation.**

x	8	6	2	4	7	8	9
y	6	5	1	4	4	7	8

Solution:

We take marks in Physics on X-axis and marks in Mathematics on Y-axis and plot the points as below.

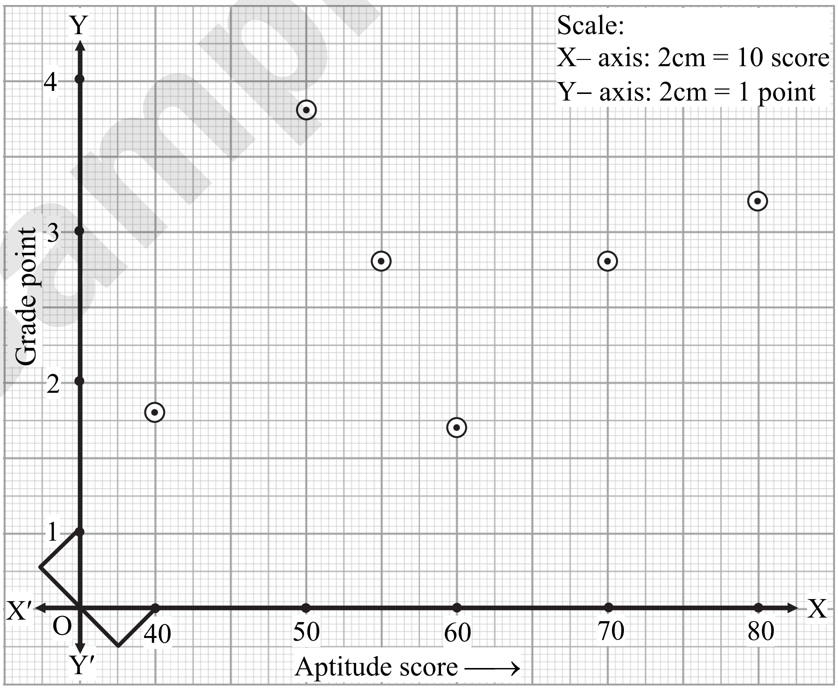


We get a band of points rising from left to right. This indicates the positive correlation between marks in Physics and marks in Mathematics.

3. **Draw scatter diagram for the data given below. Is there any correlation between Aptitude score and Grade points?**

Aptitude Score	40	50	55	60	70	80
Grade Points	1.8	3.8	2.8	1.7	2.8	3.2

Solution:



The points are completely scattered i.e., no trend is observed.

∴ there is no correlation between Aptitude score (X) and Grade point (Y).

4. Find correlation coefficient between x and y series for the following data:
$n = 15$, $\bar{x} = 25$, $\bar{y} = 18$, $\sigma_X = 3.01$, $\sigma_Y = 3.03$, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 122$.

Solution:

Here, $n = 15$, $\bar{x} = 25$, $\bar{y} = 18$, $\sigma_X = 3.01$, $\sigma_Y = 3.03$ and $\sum(x_i - \bar{x})(y_i - \bar{y}) = 122$

Since, $\text{Cov}(X, Y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$

$\therefore$ $\text{Cov}(X, Y) = \frac{1}{15} \times 122$

$= 8.13$

Since, $r = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$

$\therefore$ $r = \frac{8.13}{3.01 \times 3.03}$

$= \frac{8.13}{9.1203}$

$\therefore$ $r = 0.89$

---

5. The correlation coefficient between two variables x and y is 0.48. The covariance is 36 and the variance of x is 16. Find the standard deviation of y.

Solution:

Given, $r = 0.48$, $\text{Cov}(X, Y) = 36$

Since $\sigma_X^2 = 16$

$\therefore$ $\sigma_X = 4$

Since, $r = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$

$\therefore$ $0.48 = \frac{36}{4 \times \sigma_Y}$

$\therefore$ $\sigma_Y = \frac{36}{0.48 \times 4} = \frac{9}{0.48}$

$= \frac{900}{48} = 18.75$

$\therefore$ standard deviation of y is 18.75.

---

6. In the following data one of the value of y is missing. Arithmetic means of x and y series are 6 and 8 respectively. $(\sqrt{2} = 1.4142)$

x	6	2	10	4	8
y	9	11	?	8	7

i. Estimate missing observation.
ii. Calculate correlation coefficient.

Solution:

i. Let $X = x_i$, $Y = y_i$ and missing observation be 'a'.

Given, $\bar{x} = 6$, $\bar{y} = 8$, $n = 5$

$\bar{y} = \frac{\Sigma y_i}{n}$

$\therefore$ $8 = \frac{35 + a}{5}$

$\therefore$ $40 = 35 + a$

$\therefore$ $a = 5$

ii. We construct the following table:

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	6	9	36	81	54
	2	11	4	121	22
	10	a = 5	100	25	50
	4	8	16	64	32
	8	7	64	49	56
Total	**30**	**40**	**220**	**340**	**214**

From the table, we have

$\sum x_i = 30,\ \sum y_i = 40,\ \sum x_i^2 = 220,\ \sum y_i^2 = 340,\ \sum x_i\, y_i = 214$

Since, Cov (X, Y) $= \frac{1}{n}\sum x_i\, y_i - \bar{x}\,\bar{y}$

$\therefore$ Cov (X, Y) $= \frac{1}{5} \times 214 - 6 \times 8$

$= 42.8 - 48$

$= -5.2$

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

$= \frac{220}{5} - (6)^2 = 44 - 36$

$\therefore \sigma_X^2 = 8$

$\therefore \sigma_X = \sqrt{8} = 2\sqrt{2} = 2(1.4142) = 2.83$

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

$= \frac{340}{5} - (8)^2 = 68 - 64$

$\therefore \sigma_Y^2 = 4$

$\therefore \sigma_Y = \sqrt{4} = 2$

Thus, the correlation coefficient between X and Y is

$r = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$

$= \frac{-5.2}{2.83 \times 2}$

$= \frac{-2.6}{2.83}$

$= -0.92$

---

7. Find correlation coefficient from the following data. $\left[\textbf{Given: } \sqrt{3} = 1.732\right]$

x	3	6	2	9	5
y	4	5	8	6	7

Solution:

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	3	4	9	16	12
	6	5	36	25	30
	2	8	4	64	16
	9	6	81	36	54
	5	7	25	49	35
Total	**25**	**30**	**155**	**190**	**147**

From the table, we have

n = 5, $\sum x_i = 25$, $\sum y_i = 30$, $\sum x_i^2 = 155$, $\sum y_i^2 = 190$, $\sum x_i y_i = 147$

$\bar{x} = \dfrac{\sum x_i}{n} = \dfrac{25}{5}$

$= 5$

$\bar{y} = \dfrac{\sum y_i}{n} = \dfrac{30}{5}$

$= 6$

Since, Cov (X, Y) = $\dfrac{1}{n}\sum x_i y_i - \bar{x}\bar{y}$

$\therefore$ Cov (X, Y) = $\dfrac{1}{5} \times 147 - (5 \times 6)$

$= 29.4 - 30$

$= -0.6$

$\sigma_X^2 = \dfrac{\sum x_i^2}{n} - (\bar{x})^2 = \dfrac{155}{5} - (5)^2$

$= 31 - 25$

$\therefore \quad \sigma_X^2 = 6$

$\therefore \quad \sigma_X = \sqrt{6}$

$\sigma_Y^2 = \dfrac{\sum y_i^2}{n} - (\bar{y})^2$

$= \dfrac{190}{5} - (6)^2 = 38 - 36$

$\therefore \quad \sigma_Y^2 = 2$

$\therefore \quad \sigma_Y = \sqrt{2}$

$\therefore \quad \sigma_X\sigma_Y = \sqrt{6}\sqrt{2} = \sqrt{12}$

$= 2\sqrt{3}$

$= 2(1.732) = 3.464$

Thus, the correlation coefficient between X and Y is

$r = \dfrac{\text{Cov (X, Y)}}{\sigma_X\ \sigma_Y} = \dfrac{-0.6}{3.464}$

$= -0.1732$

---

8. Correlation coefficient between x and y is 0.3 and their covariance is 12. The variance of x is 9, find the standard deviation of y.

Solution:

Given, r = 0.3, Cov (X, Y) = 12,

$\sigma_X^2 = 9$

$\therefore \quad \sigma_X = 3$

Since, $r = \dfrac{\text{Cov}(X, Y)}{\sigma_X\ \sigma_Y}$

$\therefore \quad 0.3 = \dfrac{12}{(3)(\sigma_Y)}$

$\therefore \quad \sigma_Y = \dfrac{12}{(3)(0.3)}$

$\therefore \quad \sigma_Y = \dfrac{4}{0.3}$

$\therefore \quad \sigma_Y = 13.33$

$\therefore$ the standard deviation of y is 13.33.

Miscellaneous Exercise – 5

1. Two series of x and y with 50 items each have standard deviations 4.8 and 3.5 respectively. If the sum of products of deviations of x and y series from respective arithmetic means is 420, then find the correlation coefficient between x and y.

Solution:

Given, $n = 50,\ \sigma_X = 4.8,\ \sigma_Y = 3.5,\ \sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 420$

$$\text{Cov}(X, Y) = \frac{1}{n}\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$$

$$= \frac{1}{50} \times 420$$

$\therefore$ Cov (X, Y) = 8.4

$$r = \frac{\text{Cov}\left(X, Y\right)}{\sigma_X\ \sigma_Y} = \frac{8.4}{(4.8)\ (3.5)} = \frac{84 \times 10}{48 \times 35} = \frac{1}{2} = 0.5$$

---

2. Find the number of pairs of observations from the following data,

$$\mathbf{r = 0.15,\ \sigma_y = 4,\ \Sigma\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 12,\ \Sigma\left(x_i - \bar{x}\right)^2 = 40.}$$

Solution:

Given, $r = 0.15,\ \sigma_Y = 4,\ \sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right) = 12,\ \sum\left(x_i - \bar{x}\right)^2 = 40$

Since, $\sigma_X = \sqrt{\frac{1}{n}\sum\left(x_i - \bar{x}\right)^2} = \sqrt{\frac{40}{n}}$

$$\text{Cov}(X, Y) = \frac{1}{n}\sum\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)$$

$$= \frac{1}{n} \times 12$$

$\therefore$ $\text{Cov}(X, Y) = \frac{12}{n}$

Since, $r = \frac{\text{Cov}\left(X, Y\right)}{\sigma_X\ \sigma_Y}$

$\therefore$ $0.15 = \dfrac{\frac{12}{n}}{\sqrt{\frac{40}{n}} \times 4}$

$\therefore$ $0.15 = \dfrac{3}{n \times \sqrt{\frac{40}{n}}}$

$\therefore$ $0.05 = \dfrac{1}{\sqrt{n} \times \sqrt{40}}$

Squaring on both the sides, we get

$$0.0025 = \frac{1}{n \times 40}$$

$\therefore$ $n = \dfrac{1}{0.0025 \times 40}$

$$= \frac{10000}{25 \times 40}$$

$$= \frac{10000}{1000}$$

$\therefore$ n = 10

3. Given that r = 0.4, σ_y = 3, $\sum(x_i - \bar{x})(y_i - \bar{y})$ = 108, $\sum(x_i - \bar{x})^2$ = 900. Find the number of pairs of observations.

Solution:

Given, r = 0.4, $\sigma_Y = 3$, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 108$, $\sum(x_i - \bar{x})^2 = 900$

$$\text{Cov (X, Y)} = \frac{1}{n}\sum(x_i - \bar{x})(y_i - \bar{y})$$

$$= \frac{1}{n} \times 108$$

$$\therefore \quad \text{Cov (X, Y)} = \frac{108}{n}$$

$$\sigma_X = \sqrt{\frac{1}{n} \times \sum(x_i - \bar{x})^2}$$

$$= \sqrt{\frac{1}{n} \times 900}$$

$$= \sqrt{\frac{900}{n}} = \frac{30}{\sqrt{n}}$$

$$\text{Since, } r = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

$$\therefore \quad 0.4 = \frac{\frac{108}{n}}{\frac{30}{\sqrt{n}} \times 3}$$

$$\therefore \quad 0.4 = \frac{108}{n} \times \frac{\sqrt{n}}{30 \times 3}$$

$$\therefore \quad 0.4 = \frac{12}{10\sqrt{n}}$$

$$\therefore \quad \sqrt{n} = \frac{12}{4} = 3$$

$$\therefore \quad n = 9$$

---

4. Given the following information, $\sum x_i^2$ = 90, $\sum x_i y_i$ = 60, r = 0.8, σ_y = 2.5, where x_i and y_i are the deviations from their respective means, find the number of items.

Solution:

Here, r = 0.8, $\sum x_i y_i = 60$, $\sigma_Y = 2.5$, $\sum x_i^2 = 90$

Here, x_i and y_i are the deviations from their respective means.

$\therefore$ If X_i, Y_i are elements of x and y series respectively, then

$X_i - \bar{x} = x_i$ and $Y_i - \bar{y} = y_i$

$$\therefore \quad \sum x_i y_i = \sum(X_i - \bar{x})(Y_i - \bar{y}) = 60, \ \sum x_i^2 = \sum(X_i - \bar{x})^2 = 90$$

$$\text{Now, } \sigma_X^2 = \frac{\sum(X_i - \bar{x})^2}{n}$$

$$\therefore \quad \sigma_X^2 = \frac{90}{n}$$

$$\therefore \quad \sigma_X = \sqrt{\frac{90}{n}}$$

Also, $\text{Cov}(X, Y) = \frac{1}{n}\sum\left(X_i - \bar{x}\right)\left(Y_i - \bar{y}\right)$

$\therefore$ $\text{Cov}(X, Y) = \frac{60}{n}$

$r = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$

$\therefore$ $0.8 = \frac{\frac{60}{n}}{\sqrt{\frac{90}{n}} \times 2.5}$

$\therefore$ $0.8 \times 2.5 \times \sqrt{\frac{90}{n}} = \frac{60}{n}$

$\therefore$ $2 \times \frac{\sqrt{90}}{\sqrt{n}} = \frac{60}{n}$

$\therefore$ $\frac{n}{\sqrt{n}} = \frac{60}{2 \times \sqrt{90}}$

$\therefore$ $\frac{\sqrt{n} \times \sqrt{n}}{\sqrt{n}} = \frac{30}{\sqrt{90}} = \frac{\sqrt{30} \times \sqrt{30}}{\sqrt{3}\sqrt{30}}$

$\therefore$ $\sqrt{n} = \sqrt{10}$

$\therefore$ $n = 10$

---

5. A sample of 5 items is taken from the production of a firm. Length and weight of 5 items are given below. $\left[\textbf{Given}: \sqrt{0.8823} = 0.9393\right]$

Length (cm)	3	4	6	7	10
Weight (gm.)	9	11	14	15	16

Calculate correlation coefficient between length and weight and interpret the result.

Solution:

Let length = x_i (in cm), Weight = y_i (in gm)

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	3	9	9	81	27
	4	11	16	121	44
	6	14	36	196	84
	7	15	49	225	105
	10	16	100	256	160
Total	**30**	**65**	**210**	**879**	**420**

From the table, we have

$n = 5$, $\sum x_i = 30$, $\sum y_i = 65$, $\sum x_i^2 = 210$, $\sum y_i^2 = 879$, $\sum x_i y_i = 420$

$\bar{x} = \frac{\sum x_i}{n} = \frac{30}{5} = 6$, $\bar{y} = \frac{\sum y_i}{n} = \frac{65}{5} = 13$

$\text{Cov}(X, Y) = \frac{1}{n}\sum x_i y_i - \bar{x}\,\bar{y}$

$= \frac{1}{5} \times 420 - 6 \times 13$

$= 84 - 78 = 6$

$\sigma_X^2 = \frac{\sum x_i^2}{n} - \left(\bar{x}\right)^2 = \frac{210}{5} - (6)^2 = 42 - 36$

$\therefore$ $\sigma_X^2 = 6$

$\therefore$ $\sigma_X = \sqrt{6}$

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

$= \frac{879}{5} - (13)^2$

$= 175.8 - 169$

$\sigma_Y^2 = 6.8$

$\therefore \quad \sigma_Y = \sqrt{6.8}$

Thus, the coefficient of correlation between X and Y is

$r = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{6}{\sqrt{6}\sqrt{6.8}} = \sqrt{\frac{6}{6.8}} = \sqrt{\frac{60}{68}} = \sqrt{\frac{15}{17}} = \sqrt{0.8823}$

$\therefore \quad r = 0.9393 \approx 0.94$

$\therefore$ the value of r indicates high degree of positive correlation between length and weight of items.

---

6. Calculate correlation coefficient from the following data and interpret it.

x	1	3	5	7	9	11	13
y	12	10	8	6	4	2	0

Solution:

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	1	12	1	144	12
	3	10	9	100	30
	5	8	25	64	40
	7	6	49	36	42
	9	4	81	16	36
	11	2	121	4	22
	13	0	169	0	0
Total	**49**	**42**	**455**	**364**	**182**

From the table, we have

$n = 7,\ \sum x_i = 49,\ \sum y_i = 42,\ \sum x_i^2 = 455,\ \sum y_i^2 = 364,\ \sum x_i y_i = 182.$

$\therefore \quad \bar{x} = \frac{\sum x_i}{n} = \frac{49}{7} = 7,$

$\bar{y} = \frac{\sum y_i}{n} = \frac{42}{7} = 6$

$\text{Cov}(X, Y) = \frac{1}{n}\sum x_i y_i - \bar{x}\,\bar{y}$

$= \frac{1}{7} \times 182 - (7 \times 6)$

$= 26 - 42$

$\therefore \quad \text{Cov}(X, Y) = -16$

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

$= \frac{455}{7} - (7)^2 = 65 - 49$

$\therefore \quad \sigma_X^2 = 16$

$\therefore \quad \sigma_X = 4$

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

$= \frac{364}{7} - (6)^2 = 52 - 36$

$\sigma_Y^2 = 16$

$\therefore \quad \sigma_Y = 4$

Thus, the coefficient of correlation between X and Y is

$$r = \frac{\text{Cov}(X, Y)}{\sigma_X\ \sigma_Y} = \frac{-16}{4 \times 4} = -1$$

$\therefore$ the value of r indicates perfect negative correlation between x and y.

---

7. Calculate correlation coefficient from the following data and interpret it.

x	9	7	6	8	9	6	7
y	19	17	16	18	19	16	17

Solution:

	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
	9	19	81	361	171
	7	17	49	289	119
	6	16	36	256	96
	8	18	64	324	144
	9	19	81	361	171
	6	16	36	256	96
	7	17	49	289	119
Total	**52**	**122**	**396**	**2136**	**916**

From the table, we have

$n = 7,\ \sum x_i = 52,\ \sum y_i = 122,\ \sum x_i^2 = 396,\ \sum y_i^2 = 2136,\ \sum x_i y_i = 916$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{52}{7}$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{122}{7}$$

$$\therefore \quad \bar{x}\,\bar{y} = \frac{52 \times 122}{49} = \frac{6344}{49}$$

$$\text{Cov}(X, Y) = \frac{1}{n}\sum x_i y_i - \bar{x}\,\bar{y}$$

$$= \frac{916}{7} - \frac{6344}{49}$$

$$= \frac{6412 - 6344}{49} = \frac{68}{49}$$

$$\sigma_X^2 = \frac{\sum x_i^2}{n} - \left(\bar{x}\right)^2$$

$$= \frac{396}{7} - \left(\frac{52}{7}\right)^2 = \frac{2772 - 2704}{49} = \frac{68}{49}$$

$$\sigma_Y^2 = \frac{\sum y_i^2}{n} - \left(\bar{y}\right)^2$$

$$= \frac{2136}{7} - \left(\frac{122}{7}\right)^2$$

$$= \frac{14952 - 14884}{49} = \frac{68}{49}$$

$$\therefore \quad \sigma_X \sigma_Y = \sqrt{\sigma_X^2\ \sigma_Y^2}$$

$$= \sqrt{\frac{68}{49} \times \frac{68}{49}} = \frac{68}{49}$$

$$r = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\left(\frac{68}{49}\right)}{\frac{68}{49}} = 1$$

$\therefore$ the value of r indicates perfect positive correlation between x and y.

8. If the correlation coefficient between X and Y is 0.8, what is the correlation coefficient between

i. **2X and Y** **ii.** $\frac{X}{2}$ **and Y** **iii.** **X and 3Y**

iv. **X – 5 and Y – 3** **v.** **X + 7 and Y + 9** **vi.** $\frac{X-5}{7}$ **and** $\frac{Y-3}{8}$**?**

Solution:

Correlation coefficient remains unaffected by the change of origin and scale.

i.e., if $u_i = \frac{x_i - a}{h}$ and $v_i = \frac{y_i - b}{k}$, then Corr (U, V) = ± Corr (X, Y), according to the same or opposite signs of h and k.

i. $u_i = \frac{2(x_i - 0)}{1}$, $v_i = \frac{y_i - 0}{1}$

Here, h = 1 and k = 1 are of the same signs.

∴ Corr (U, V) = Corr (X, Y) = 0.8

---

ii. $u_i = \frac{x_i - 0}{2}$, $v_i = \frac{y_i - 0}{1}$

Here, h = 2 and k = 1 are of the same signs.

∴ Corr (U, V) = Corr (X, Y) = 0.8

---

iii. Corr (X, 3Y) = Corr (X, Y) = 0.8

---

iv. Corr (X – 5, Y – 3) = Corr (X, Y) = 0.8

---

v. Corr (X + 7, Y + 9) = Corr (X, Y) = 0.8

---

vi. Corr $\left(\frac{X-5}{7}, \frac{Y-3}{8}\right)$ = Corr (X, Y) = 0.8

---

9. In the calculation of the correlation coefficient between the height and weight of a group of students of a college, one investigator took the measurements in inches and pounds while the other investigator took the measurements in cm. and kg. Will they get the same value of the correlation coefficient or different values? Justify your answer.

Solution:

Coefficient of correlation is a ratio of covariance and standard deviations.

Since, covariance and standard deviations are independent of units of measurement.

∴ coefficient of correlation is also independent of units of measurement.

∴ values of coefficient of correlation obtained by first and second investigators are same.

Activities for Practice

1. Complete the following activity to find Cov(X, Y) from the following data $\sum x_i = 15$, $\sum y_i = 36$, $\sum x_i y_i = 110$, n = 5.

Solution:

$$\text{Cov}(X, Y) = \frac{1}{n}\left(\sum x_i y_i - \frac{1}{n}\left(\sum x_i\right)\left(\sum y_i\right)\right)$$

$$= \frac{1}{5}\left(110 - \frac{1}{\square}(15)(\square)\right)$$

$$= \frac{1}{5}\left(\frac{\square - 540}{\square}\right)$$

$$= \frac{2}{\square}$$

$$= \square$$

2. Find the coefficient of correlation r when Cov (X, Y) = 16.5, Var (X) = 8.25, Var (Y) = 33 by completing the following activity.

Solution:

$$r = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)}\sqrt{\text{Var}(Y)}}$$

$$= \frac{\boxed{}}{\sqrt{8.25}\sqrt{\boxed{}}}$$

$$= \frac{\boxed{}}{\sqrt{272.25}}$$

$$= \boxed{}$$

---

3. Measure weight and height of at least 20 students in your class/college.
Prepare a bivariate frequency table and calculate the correlation coefficient between weight and height.
(Textbook page no. 64)

*[**Note:** Students can attempt this activity on their own.]*

---

4. Calculate the correlation coefficient between age (in years) and blood pressure from Q.4. of Miscellaneous Exercise 4. *(Textbook page no. 64)*

---

5. Using the given data plot the points and draw the scatter diagram and identify the type of correlation.

X	8	12	16	20	24	28	32
Y	2	3	4	5	6	7	8

(Textbook page no. 64)

---

6. Select any 2 stocks and record the share prices for 10 days. Draw the scatter diagram of them.
(Textbook page no. 64)

*[**Note:** Students can attempt this activity on their own.]*

One Mark Questions

Multiple Choice Questions

1. Correlation analysis aims at
(A) Predicting one variable for a given value of the other variable.
(B) Establishing relation between two variables.
(C) measuring the extent of relation between two variables.
(D) Both (B) and (C).

2. Covariance is affected by ________.
(A) change of origin (B) change of scale
(C) both (A) and (B) (D) change of properties

3. If the correlation coefficient between two variables is unity, then there is
(A) perfect correlation (B) perfect positive correlation
(C) perfect negative correlation (D) no correlation

4. If the plotted points in a scatter diagram are evenly distributed, then the correlation is
(A) positive (B) zero (C) negative (D) none of these

5. If all the plotted points in a scatter diagram lie on a single line, then the correlation is
(A) perfect positive (B) perfect negative (C) Both (A) and (B) (D) Either (A) or (B)

6. If the plotted points in a scatter diagram lie on a single line, from upper left to lower right, then the correlation is
(A) positive (B) zero (C) negative (D) none of these

7. The correlation between shoe size and intelligence is
(A) zero (B) positive (C) negative (D) none of these

8. Scatter diagram helps us to
(A) find the nature of correlation between two variables.
(B) compute the extent of correlation between two variables.
(C) obtain the mathematical relationship between two variables.
(D) Both (A) and (C).

9. Pearson's correlation coefficient is used for finding
(A) correlation for any type of relation. (B) correlation for linear relation only.
(C) correlation for curvilinear relation only (D) Both (B) and (C).

10. What is the range of correlation coefficient?
(A) ϕ (B) $\{-1, 1\}$ (C) $[0, 1]$ (D) $[-1, 1]$

11. If for two variables x and y, the Cov (X, Y) is 40, $\sigma_X^2 = 16$ and $\sigma_Y^2 = 256$. What is the correlation coefficient?
(A) 0.01 (B) 0.625 (C) 0.4 (D) 0.5

12. What is the value of correlation coefficient due to Pearson on the basis of the following data:

x	– 5	– 4	– 3	– 2	– 1	0	1	2	3	4	5
y	27	18	11	6	3	2	3	6	11	18	27

(A) 1 (B) – 1 (C) 0 (D) – 0.5

13. The correlation between the speed of an automobile and the distance travelled by it after applying the breaks is
(A) negative (B) zero (C) positive (D) None

14. If the value of correlation coefficient is positive, then the points in a scatter diagram tend to cluster
(A) from lower left corner to upper right corner.
(B) from lower left corner to lower right corner.
(C) from lower right corner to upper left corner.
(D) from lower right corner to upper right corner.

15. When r = –1, all the points in a scatter diagram would be
(A) on a straight line directed from lower left to upper right.
(B) on a straight line from upper left to lower right.
(C) on a straight line.
(D) Both (A) and (B)

True or False

1. If r is a correlation co-efficient, then $0 \le r \le 1$.
2. If two variables X and Y are independent, then r = 0.
3. Coefficient of correlation between Y and Y is 0.
4. coefficient of correlation gets affected by the shift of origin and change of scale.
5. Coefficient of correlation measures only linear correlation between X and Y.

Fill in the blanks

1. For a perfect positive correlation, r = _______
2. For a perfect negative correlation, r = _______
3. r = _______ in case of no correlation.
4. If 0.3 < r < 0.8, the correlation is _______
5. If r > 0.8, there is _______ positive correlation.

Additional Problems for Practice

Based on Exercise 5.1

+1. A train travelled between two stations and distance and time were recorded as below,

Distance(km)	80	120	160	200	240
Time(Hr)	2	3	4	5	6

Draw scatter diagram and identify the type of correlation.

+2. Draw scatter diagram for the following data and identify the type of correlation.

Capital (in crores ₹)	2	3	4	5	6	8	9
Profit (in lakh ₹)	3	2	4	4	5	6	7

3. Draw scatter diagram for the data given below and interpret it.

Capital (in crores ₹)	2	3	4	5	6	8	9
Profit (in lakh ₹)	6	5	7	7	8	12	11

4. Find Cov (X, Y) between x and y from the following data:
$\sum x = 15, \sum y = 36, \sum xy = 110$, n = 5

+5. Compute correlation coefficient for the following data,
n = 100, $\bar{x} = 62$, $\bar{y} = 53$, $\sigma_X = 10$,
$\sigma_Y = 12$, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 8000$.

6. For a bivariate data r = 0.75, $\sigma_X = 12$, $\sigma_Y = 15$, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 2700$, find n.

+7. Find correlation coefficient between x and y for the following data.
n = 25, $\sum x_i = 75$, $\sum y_i = 100$, $\sum x_i^2 = 250$, $\sum y_i^2 = 500$, $\sum x_i y_i = 325$.

8. Find the correlation coefficient if n = 10, $\sum x_i y_i = 1807$, $\bar{x} = 15.3$, $\bar{y} = 9.4$, $\sigma_X = 16$, $\sigma_Y = 5$.

9. Find the correlation coefficient if n = 10, $\sum(x_i - \bar{x})(y_i - \bar{y}) = 519$, $\sigma_X = 16.4$, $\sigma_Y = 12$

+10. Find correlation coefficient between x and y for the following data and interpret it.

x	1	2	3	4	5	6	7	8	9
y	12	11	13	15	14	17	16	19	18

$\left(\sqrt{666} = 25.80\right)$

11. Find correlation coefficient for the following data.

x	3	5	1	4	2
y	5	6	3	7	2

12. Compute correlation coefficient for the following data:

x	9	7	6	8	15
y	19	17	16	18	25

+13. Calculate correlation coefficient from the following data,

$n = 10$, $\sum x_i = 140$, $\sum y_i = 150$, $\sum(x_i - 10)^2 = 180$, $\sum(y_i - 15)^2 = 500$, and $\sum(x_i - 10)(y_i - 15) = 60$.

+14. Calculate correlation coefficient between age of husbands and age of wives.

Age of husbands	23	27	28	29	30	31	33	35	36	39
Age of wives	18	22	23	24	25	26	28	30	31	34

Based on Miscellaneous Exercise - 5

1. Two series of x and y with 50 items each have standard deviations 4.5 and 3.5 respectively. If the sum of product of deviations of x and y series from respective arithmetic means is 420, then find the coefficient of correlation between x and y.

2. Find Cov (X, Y) between x and y if $\sum x_i = 55$, $\sum y_i = 74$, $\sum x_i y_i = 441$, $n = 10$

3. Calculate coefficient of correlation for the following data:

x	10	6	9	10	12	13	11	9
y	9	4	6	9	11	13	8	4

4. Calculate the coefficient of correlation between x and y.

x	1	3	4	5	7	8	10
y	2	6	8	10	14	16	20

5. Compute the correlation coefficient by Karl Pearson's method between x and y and interpret its value.

x	11	12	14	16	12	17	18	19	20	17
y	8	9	10	12	9	13	14	13	15	12

6. Compute correlation coefficient between x and y for the following data:

x	20	30	25	65	70	80	95
y	9	10	7	11	30	40	45

7. Calculate the Karl Pearson's coefficient of correlation from the following data:

Marks in Ist term	75	81	70	76	77	81	84	75
Marks in IInd term	62	68	65	60	69	72	76	72

Answers

Activities for Practice

1. i. 5 ii. 36 iii. 550 iv. 5
   v. 5 vi. 0.4

2. i. 16.5 ii. 33 iii. 16.5 iv. 1

4.

<table>
<tr><th>Age in years (X)
Blood Pressure (Y)</th><th>35 – 45</th><th>45 – 55</th><th>55 – 65</th><th>65 – 75</th><th>Total (f_y)</th></tr>
<tr><td>115 – 130</td><td>|||| (4)</td><td>–</td><td>–</td><td></td><td>4</td></tr>
<tr><td>130 – 145</td><td>|| (2)</td><td>| (1)</td><td>||| (3)</td><td>–</td><td>6</td></tr>
<tr><td>145 – 160</td><td>| (1)</td><td>| (1)</td><td>|| (2)</td><td>||| (3)</td><td>7</td></tr>
<tr><td>160 – 175</td><td>| (1)</td><td>|| (2)</td><td>|| (2)</td><td>|| (2)</td><td>7</td></tr>
<tr><td>Total (f_x)</td><td>8</td><td>4</td><td>7</td><td>5</td><td>24</td></tr>
</table>

Blood Pressure (Y)	Mv_y	Age in years (X) Class	35 – 45	45 – 55	55 – 65	65 – 75					
		Mv_x	40	50	60	70					
		v \ u	– 1	0	1	2	fy	v	fv	fv^2	fuv
115 – 130	122.5	– 1	4 $\boxed{4}$	0 $\boxed{0}$	0 $\boxed{0}$	0 $\boxed{0}$	4	– 1	– 4	4	4
130 – 145	137.5	0	2 $\boxed{0}$	1 $\boxed{0}$	3 $\boxed{0}$	0 $\boxed{0}$	6	0	0	0	0
145 – 160	152.5	1	1 $\boxed{-1}$	1 $\boxed{0}$	2 $\boxed{2}$	3 $\boxed{6}$	7	1	7	7	7
160 – 175	167.5	2	1 $\boxed{-2}$	2 $\boxed{0}$	2 $\boxed{4}$	2 $\boxed{8}$	7	2	14	28	10
		fx	8	4	7	5	24		17	39	21
		u	– 1	0	1	2					
		fu	– 8	0	7	10	9				
		fu^2	8	0	7	20	35				
		f_{UV}	1	0	6	14	21				

$$u = \frac{Mv_x - 50}{10} \; ; \; v = \frac{Mv_y - 137.5}{15}$$

$$r = \frac{n\sum fuv - \left(\sum fu\right)\left(\sum fv\right)}{\sqrt{n\sum fu^2 - \left(\sum fu\right)^2}\sqrt{n\sum fv^2 - \left(\sum fv\right)^2}}$$

$$= \frac{(24)\,(21) - (9)\,(17)}{\sqrt{(24)\,(35) - (9)^2} \cdot \sqrt{(24)\,(39) - (17)^2}}$$

$$= \frac{504 - 153}{\sqrt{840 - 81}\,\sqrt{936 - 289}}$$

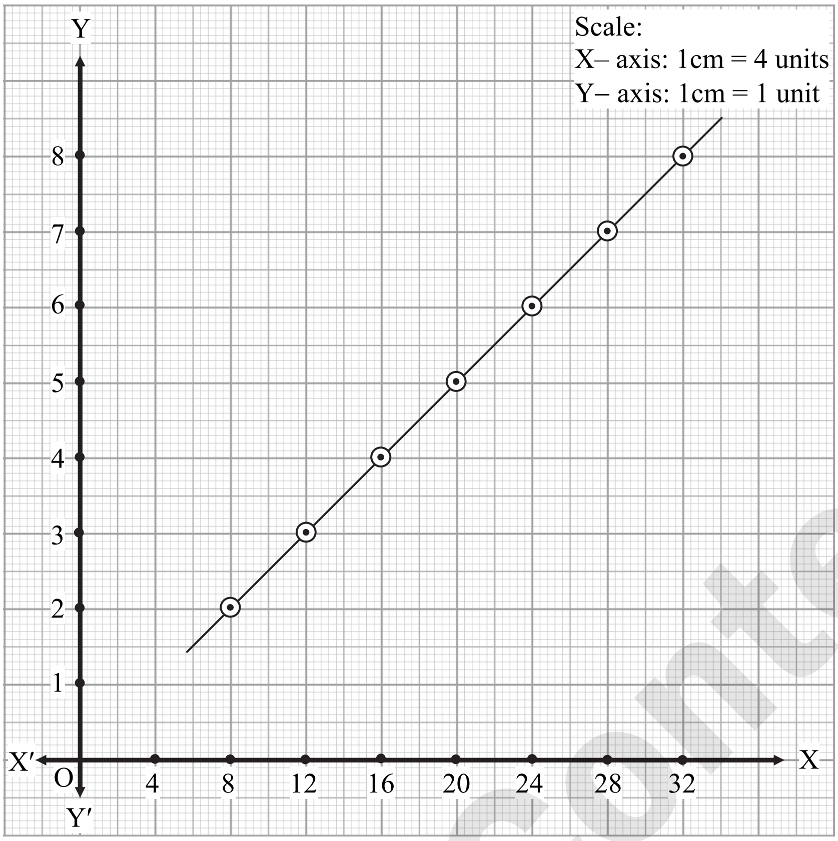
$$= \frac{351}{\sqrt{759}\,\sqrt{647}}$$

$$= \frac{351}{(27.5500)\,(25.436)}$$

$$= \frac{351}{700.7618}$$

$r = 0.5009$

5.



Graph given above gives perfectly positive correlation.

One Mark Questions

Multiple Choice Questions

1. (D) 2. (B) 3. (B) 4. (B) 5. (D) 6. (C) 7. (A) 8. (A) 9. (B) 10. (D)
11. (B) 12. (C) 13. (A) 14. (A) 15. (B)

True or False

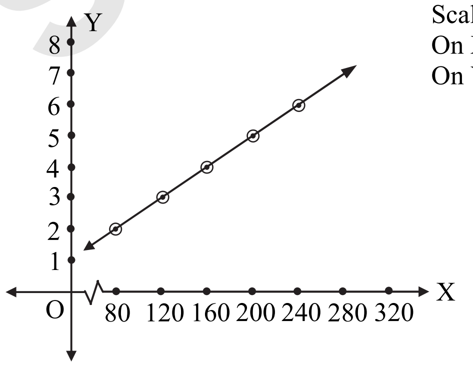
1. False 2. True 3. False 4. False 5. True

Fill in the blanks

1. 1 2. –1 3. 0 4. moderate positive 5. high

Additional Problems for Practice

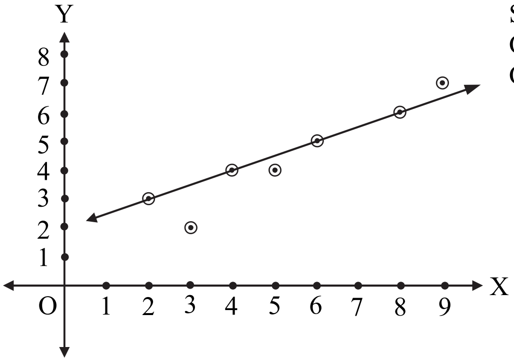
1.



Scale:
On X-axis: 1cm = 40 units
On Y-axis: 1cm = 1 unit

Perfect positive correlation

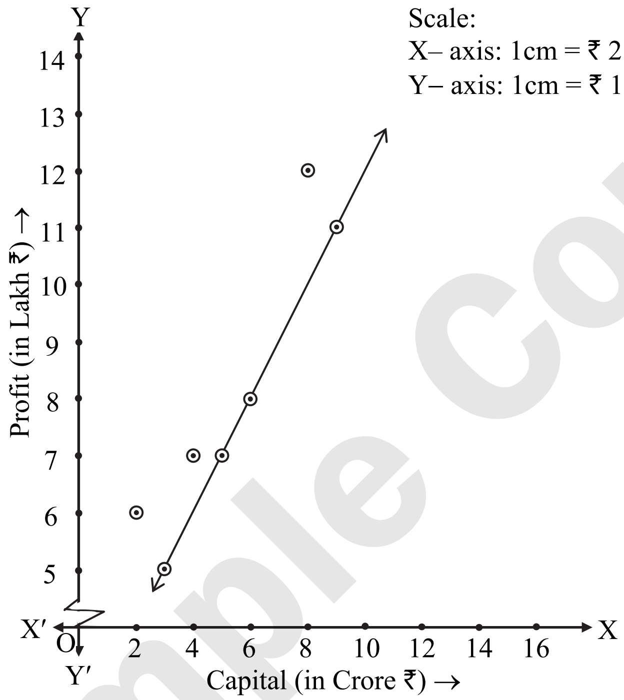
2.



Scale:
On X-axis: 1cm = 1 Crore
On Y-axis: 1cm = 1 Lakh

Positive correlation

3. We take Capital (in Crore ₹) on X-axis and Profit (in Lakh ₹) on Y-axis and plot the points as shown in the figure below



Scale:
X– axis: 1cm = ₹ 2 Crore capital
Y– axis: 1cm = ₹ 1 Lakh profit

Positive Correlation

4.	0.4	5.	0.67	6.	20	7.	0.5
8.	0.461	9.	0.264	10.	0.93; high degree positive correlation		
11.	0.839	12.	1	13.	0.6	14.	1

Based on Miscellaneous Exercise - 5

1.	0.53	2.	3.4	3.	0.8958	4.	1

5. r = 0.98; *x* and *y* have high positive correlation.
6. 0.89
7. 0.62