Example.44 Find combined standard deviation from the following information.

	Group A	Group B
No. of observation	20	10
Mean	22	16
S.D.	$\sqrt{6}$	$\sqrt{2}$

Solution :

Here $n_1 = 20$, $n_2 = 10$, $\bar{x}_1 = 22$, $\bar{x}_2 = 16$, $S_1 = \sqrt{6}$, $S_2 = \sqrt{2}$

First, we find the combined mean $\bar{x}_{12}$

$$\bar{x}_{12} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$$

$$= \frac{20(22) + 10(16)}{20 + 10}$$

$$= \frac{440 + 160}{30}$$

$$= \frac{600}{30} = 20$$

Now $d_1 = \bar{x}_1 - \bar{x}_{12}$ $\qquad d_2 = \bar{x}_2 - \bar{x}_{12}$

$= 22 - 20$ $\qquad = 16 - 20$

$d_1 = 2$ $\qquad d_2 = -4$

$d_1^2 = 4$ $\qquad d_2^2 = 16$

$$S_{12} = \sqrt{\frac{n_1(S_1^2 + d_1^2) + n_2(S_2^2 + d_2^2)}{n_1 + n_2}}$$

$$= \sqrt{\frac{20(6+4) + 10(2+16)}{(20+10)}}$$

$$= \sqrt{\frac{200 + 180}{30}} = \sqrt{\frac{380}{30}} = \sqrt{12.67}$$

$S_{12} = 3.56$ is the combined standard deviation

EXERCISE

1. The Monthly wages received by 50 labourers were as follows :

40 42 50 55 60 68 50 59 60 70
70 60 55 50 44 70 63 56 50 44
70 64 56 50 45 45 52 58 65 72
75 65 58 53 45 75 65 59 [illegible] 46
48 55 60 65 [illegible] 60 [illegible] 48

Prepare a frequency distribution of wages.

2. The following figures give the height(in inches) of 80 [illegible]. Prepare a frequency distribution with suitable class-[illegible]

62.1 64.6 60.4 63.9 59.7 60.7 62.5 59.4 60.5 63.6
61.5 64.6 63.6 63.4 62.2 67.6 60.8 65.6 67.3 65.5
63.0 64.5 59.1 63.0 63.5 64.1 66.3 64.4 61.7 65.1
62.2 66.4 64.8 64.3 67.4 59.6 59.9 62.0 66.5 62.8
64.7 64.2 61.9 61.2 63.6 64.5 63.5 63.7 65.3 61.8
63.1 62.4 62.6 62.7 68.5 61.1 61.8 66.3 63.4 64.0
65.8 63.3 67.0 64.6 60.0 65.7 65.4 63.8 65.2 68.7
60.3 64.0 68.1 64.9 61.3 60.2 67.3 66.7 66.2 66.6

Soln.–1 :

Classes	38–44	45–51	52–58	59–65	66–72	73–79	80–86
Frequency	04	11	11	13	07	03	01

Soln.–2 :

Classes	59–1–60.4	60–4–61.7	61.7–63.0	63.0–64.3	64.3–
Frequency	3	8	12	19	15
Classes	65.6–66.9	66.9–68.2	68.2–69.5		
Frequency	10	6	2		

3. Draw frequency polygon, frequency curve and histogram for the frequency distribution, obtained for exercise 1.

4. The following table gives the male population of in a [illegible] of West–Bengal in 1931. Draw, frequency polygon, [illegible] curve and hitogram.

Age-group [years]	Male population [in '000]
1–6	9
7–12	8
13–18	8
19–24	7
25–30	15
31–36	12
37–42	9
43–48	6
48–54	4
	78

5. The weekly wages in rupees of 5 labourers are 240, 260, 236, 225. Calculate their A.M.

6. Calculate the A.M., G.M. and H.M. of four values 27, 72, 108 and 144.

7. Find the A.M. of the following distribution.

Weights (in pounds)	100	110	120	130	140
No. of Men	15	20	25	30	10

8. The following distribution of age is obtained for 100 students. Find mean age of the students.

Age (in yrs)	20	19	18	17	16	15	14	13
No. of students	3	10	21	25	19	15	5	2

9. Compute mean of the following distribution :

x_i	59	60	61	62	63	64
f_i	6	7	9	12	18	8

10. 100 patients took treatment on a certain day in a hospital. The distribution of their ages are given below. Find the mean age of the patients.

Age in years	0–10	10–20	20–30	30–40	40–50	50–60
No. of patients	12	8	7	10	30	33

11. Find the arithmetic mean of the weekly income from the following frequency distribution :

Weekly income (in Rs)	20–25	25–30	30–35	35–40	40–45	45–50
No. of wokers	200	700	900	800	600	400

12. Find mean from the following distribution

Observation	2	3	4	5–7	7–10	10–15	15–20	20–25
Frequency	1	2	3	3	5	12	9	4

13. Find the mean for the following frequency distribution

Marks	0–10	0–20	0–30	0–40	0–50	0–60
No. of students	1	13	37	73	93	100

14. Find the Quartiles and Mode of the values
15, 45, 25, 24, 23, 32, 40, 10, 20, 25, 30, 27

15. Find the Quartiles and mode of the following data

Daily wages (Rs)	22	24.5	28	31.5	34	36.5
No. of workers	10	23	32	28	2	5

16. Find the Quartiles and mode of the following data

Observation	5	10	15	20	25	30	35	40
Frequency	2	3	9	11	20	16	4	2

17. Find the Quartiles and mode of the following data

Class	0–6	6–12	12–18	18–24	24–30	30–36	36–42	42–48	48–54
Frequency	13	25	57	79	105	79	57	25	13

18. The earnings of workers are given as follows

Daily wages (Rs)	12.5–17.5	17.5–22.5	22.5–27.5	27.5–32.5	32.5–37.5	37.5–42.5	42.5–47.5	47.5–52.5	52.5–57.5
No. of workers	2	22	10	14	3	4	5	1	1

Calculate the median and mode

19. Find Quartiles and Mode of the following frequency distribution.

Class	3–5	5–10	10–20	20–50	50–80	80–100
Frequency	8	12	40	70	15	5

20. Following is the frequency distribution of no. of students in 85 colleges of commerce faculty. Find the median and first quartile.

No. of students	less than 20	200–500	500–800	800–1000	1000–1500
No. of colleges	12	8	20	30	15

21. Form an ordinary frequency table from the following cumulative distribution of marks obtained by 22 students. Calculate second quartile and mode.

Marks (Below)	10	20	30	40	50
No. of students	3	8	17	20	22

22. You are given below a certain frequency distribution :

Value	less than 100	100–200	200–300	300–400	400 and above
Frequency	40	89	148	64	39

Calculate the most suitable average and the firsth quantile.

[Hint : Here some classes are open-ended, so median is the most suitable average].

23. Find median and mode from the following frequency distribution.

Observation	1	4	7	8–10	10–15	15–25	25–40	40–50
Frequency	1	3	7	10	15	8	4	2

24. Find median and Q_3 from the following data.

Marks (More than)	0	10	20	30	40	50	60	70	80
No. of students	200	192	151	91	69	32	16	8	3

25. Calculate Median from the following data :

mid value	115	125	135	145	155	165	175	185	195
frequency	6	25	48	72	116	60	38	22	3

26. Find the missing frequency in the following frequency distribution if mean = 37.

Marks	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
No. of students	4	4	9	?	12	6	3	2

27. Following is the frequency distribution of weights (in lbs) of certain no. of persons. If mean of the frequency distribution is 120 lbs calculate the missing frequency.

Weight (lbs)	100–110	110–120	120–130	130–140	140–150
No. of persons	100	130	72	20	?

28. Obtain the missing frequencies from the following distribution if mean is 16.5.

Class	0–5	5–10	10–15	15–20	20–25	25–30	30–35
f	1	7	11	?	?	4	1

29. If mean of a frequency distribution is 74, obtain the frequencies. If Total frequency = 100

Marks :	0–20	20–40	40–60	60–80	80–100	100–120
No. of : students	5	13	?	15	?	7

30. If Median of the following frequency distribution is 46, find the missing frequency.

Observation	less than 100	100–150	150–200	200–250	250–300
Frequency	35	60	?	40	25

31. If mode of the frequency distribution is 27. find the missing frequency.

Class:	0–6	6–12	12–18	18–24	24–30	30–36	36–42	42–48
Freq.:	13	25	57	79	105	79	57	42

32. The third quartile of the following frequency distribution is 290. Find the missing frequencies.

Value	less than 50	50–100	100–150	150–200	200–250	250–300	300–350
freq.:	2	4	7	21	?	?	21

33. The expenditure of 1000 families is given as under

Expenditure (in Rs)	40–59	60–79	80–99	100–119	120–139
No. of families	50	?	500	?	50

The median and mean for the distribution are both equal to Rs 87.50. Calculate the missing frequencies.

34. For the following frequency distribution, two class frequencies are missing

Intelligence Quotient	55–64	65–74	75–84	85–94	95–104	105–114	115–124	125–134	135–144
No. of students	2	19	78	?	301	?	92	14	4

It is however known that the total frequency is 900 and the median is 100.048. Find the two missing frequencies.

35. You are given the following incomplete frequency distribution. It is known that the total frequency is 360 and mode is 1376. Find the missing frequencies.

Class	0–400	400–800	800–1200	1200–1600	1600–2000	2000–2400	2400–2800	2800–3200
f	14	22	?	124	?	32	15	5

36. Mean of the following classified continuous data is 19.57 and assumed mean is 19.75. The following is the distribution of variable d. Find the original distribution.

d	−3	−2	−1	0	1	2
f	2	6	12	20	8	2

37. The form of a frequency distribution after change of origin and scale is as follows.

d	3	2	1	0	−1	−2	−3	−4
f	5	10	18	35	23	15	10	4

If mean = 4.7 and class-interval is 6, find original distribution.

38. The height (in cm) of 10 students is as follows. Find the range, relative range, quartile deviation and coefficient of Q.D.

161, 152, 167, 158, 165, 153, 168, 170, 150, 156.

39. The monthly pocket money expenses (in Rs) of 15 students of a hostel are 45, 30, 50, 60, 36, 48, 40, 66, 57, 72, 60, 30, 27, 39. Find the range, relative range, quartile deviation and coefficient of Q.D.

40. Find range, coefficient of range, quartile deviation, coefficient of quartile deviation

Observation:	3	5	8	12	17	20	24	30	35
Frequency:	1	3	7	15	20	13	10	7	4

41. Find range, coefficient of range, quartile deviation, coefficient of quartile deviation

Height (inches)	58	59	60	61	62	63	64	65	66
No. of students	15	20	32	35	33	22	20	10	8

42. Compute Range, Coefficient of R. and Q.D. and coefficient of Q.D. from the following data

Profits (Rs. lakhs)	4–8	8–12	12–16	16–20	20–24	24–28	28–32	32–36	36–40
No. of companies	6	10	18	30	15	12	10	6	2

43. Find Range, coefficient of range, quartile deviation, coefficient of Q.D.

Class	5–10	10–15	15–20	20–25	25–30	30–35	35–40	40–45	45–50
Frequency	2	7	10	28	20	18	10	4	1

44. Find the appropriate measure of dispersion and the relative measure from the following frequency distribution

Monthly income (Rs)	Less than 500	500–900	900–1300	1300–1700	1700–2100	2100 & more
No. of families	20	20	45	70	28	17

45. Find the appropriate measure of dispersion and the relati measure from the following frequency distribution

Class	less than 2	2–5	5–9	9–13	13–20	20–30	30–40	40–60	More than 60
freq	2	6	9	30	32	12	9	3	2

46. Marks out of 25 obtained by 9 students are given below.
7, 4, 10, 9, 15, 12, 7, 9, 7
Calculate mean deviation from mean, median and mode.

47. Find the mean deviation about the median of the following :
8, 15, 53, 49, 19, 62, 7, 15, 95, 77

48. Calculate coefficient of mean deviation from mean and median

Marks	10	15	20	30	40	50
frequency	8	12	15	10	3	2

49. Compute coefficient of mean deviation from the following data :

Class	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–100
frequency	10	22	35	50	78	40	30	20	10	5

50. Find the mean deviation about the median from the following distribution

Class :	2–6	6–10	10–14	14–18
Frequency :	6	8	4	2

51. Find mean deviation

observation:	50–100	28–50	18–28	10–18	7	3	1
frequency:	2	13	15	6	6	5	3

52. Calculate standard deviation using the following data regarding income in Rs. of 10 persons.
600, 620, 640, 620, 680, 670, 680, 640, 700, 650.

53. Runs scored by a cricketer in 10 innings are given below.
42, 17, 83, 59, 72, 76, 64, 45, 40, 32.
Calculate coefficient of variation.

54. Compare the variation of the following two groups.

Group A	28	15	43	9	30	
Group B	12	38	21	7	25	47

Hint : The coefficient of variation of group B is greater than that of group A. Hence, there is more variation in group B

Calculate standard deviation from the following

Marks	10	20	30	40	50
No. of students	8	12	20	10	7

56. Find out the standard deviation from the following the weights of 200 persons

Weight in kg	50	55	60	65	
No. of Men	30	40	65	50	

57. The frequency distribution of the age of 50 wome of their wedding is as follows. Find the coefficient

Age (yrs)	16–20	21–25	26–30	31–35	36–
No. of women.	5	42	33	12	

58. A consignment of 180 articles is classified accor size of the article as under. Find the standard dev its coefficient of variation.

Measure –ment	0–10	10–20	20–30	30–40	40–50	50–60	60–
No. of articles	4	6	20	40	45	31	

59. Lives of two models of refrigerators in a recent

Life (yrs)	No. of refrigerators	
	Model A	Model B
0–2	2	5
2–4	7	16
4–6	12	13
6–8	19	7
8–10	9	5
10–12	1	4

Which model has greater Variability ?

60. If the mean and the standard deviation of the distrib 52.4 and 6.68 resp., find the original frequency dis

d_i	–3	–2	–1	0	1	2	3
f	7	12	21	24	35	35	13

278

61. The mean and standard deviation of a continuous distribution are 135.3 and 9.6. The distribution of d_i obtained after changing origin and scale is given below. Findthe original distribution

d_i	-4	-3	-2	-1	0	1	2	3
f_i	2	5	8	18	22	13	8	4

62. The following are some of the particulars of the distribution of weights of boys and girls in a class

	Boys	Girls
No.	100	50
Mean weight	60 kgs	45 kgs
Variance	9	4

Find the standard deviation of the combined data

63. Given below is the distribution of marks obtained by 140 students

X (Marks more than)	10	20	30	40	50	60	70	80	90	100
f (students)	140	133	118	100	75	45	25	9	2	0

Calculate Mean, Median and Mode.

64. 1500 workers are working in an industrial establishment. Their age is classified as follows :

Age (in Yrs)	18–22	22–26	26–30	30–34	34–38
No. of Workers	120	125	180	260	155
Age (in Yrs)	38–42	42–46	46–50	50–54	54–58
No. of Workers	184	162	86	75	53

Find out Mean, Median and Mode.

65. Calculate the standard deviation and mean from the following data

Class	0–30	30–60	60–90	90–120	120–150	150–180	180–210
frequency	9	17	43	82	81	44	24

66. Given the following frequency distribution with some missing frequencies.

Class	10–20	20–30	30–40	40–50	50–60	60–70	70–80
frequency	185	-	34	180	136	-	50

If the total frequency is 685 and Median is 42.6, find the missing frequencies

Probability

5.1 Introduction

Probability has became a part
and managerial decisions, we face
theory to measure uncertainty. Whe
a seventy percent chance of rain, w
to a pool game. While playing bri
estimate before attempting a fine
inventories of highly styled women'
chances that sales will reach or exce
practice to add, words, "PROBABL
in routine conversation, for exampl
there are chances that sachin will g
reflect the situation of "UNCERT
phenomena.

Gamblers often use "ODDS"
a french gamblar named, Antoine G
the mathematical basis for success
question was "What are the odds
in twenty four rolls of a pair of dic
mathematician, solved the problem
ideas with famous mathematician P
these three together laid the foundatio
foundation. other scholars of prob
1705), Abraham de Moivre (1667
(1702–1761) and Joseph Kagra
sophisticated formulas and techniqu
games of chance. In 19th Century,
(1749–1827) unified all these conce
of probability.

(11) If two variables are changing in same direction and in same proportion, then give the value of r.

(12) If two variables are changing in the opposite direction and in same proportion, then give the value of r.

(13) Find the coefficient of correlation between X and Y = aX where a is real number.

Answers :

(1) High degree of negative correlation

(2) Low degree of negative correlation

(3) $r = +1$

(4) $r = -1$

(5) $r = -1$

(6) $r = +1$

(7) $\Sigma d^2 = 0$, $r = 1$

(8) $r = -1$

(9) $r = 0$

(10) (i) 0.5 (ii) 0.5 (iii) 0.5 (iv) − 0.5 (v) 0.5

(11) $r = 1$

(12) $r = -1$.

(13) $r = \frac{a}{|a|}$ where $|a|$ is the absolute value of a. i.f. $a > 0$ then $r = 1$ and if a is negative $r = -1$.

7. Draw a scatter diagram for the following data and from it interprete value of r.

X	15	25	35	45	55
Y	10	15	20	25	30

[Ans : $r = 1$]

8. Draw a scatter diagram for the following data.

X	1	2	3	4	5	6
Y	6	4	3	5	4	2

[Ans : Partial negative correlation]

9. Calculate Pearson's coefficient of correlation between advertisement cost and sales as per the data given below.

Advertisement cost ('000 Rs)	139	165	162	190	182	175	125	198	136	178
Sales (lakh Rs)	147	153	158	186	162	168	160	191	151	184

[Ans : $r = 0.78$]

10. Find the correlation coefficient between the sales and expenses of the following 10 companies

Company	1	2	3	4	5	6	7	8	9	10
Sales	70	70	75	80	85	85	85	80	80	70
Expenses	11	13	14	16	16	15	15	14	13	13

[Ans : $r = 0.79$]

11. Find the coefficient of correlation by to the method due to Karl Pearson.

X	5	4	0	−3	−1	2	7
Y	8	10	12	15	12	10	5

[Ans : $r = -0.97$]

12. Calculate correlation coefficient for the following data and also calculate coefficient of determination.

X	8	7	6	1	2	3	9	4	5
Y	16	14	13	9	8	10	15	12	11

[Ans. : $r = 0.95$, $r^2 = R^2 = 0.90$]

13. Find product moment correlation coefficient r and coefficient of deremination write your interpretation.

X	75	80	86	94	95	98	100	105	110	112
Y	29	38	42	40	39	37	45	34	38	31

[Ans. : $r = 0.03$, $r^2 = R^2 = 0.0009$]

14. Calculate correlation coefficient from the following data.

X	73	74	76	70	74	72	75	72	78	72	87
Y	79	81	69	68	69	77	79	91	80	76	60

[Ans : $r = -0.73$]

15. Find Pearson's r from following bivariate data.

Price (in Rs)	15	25	35	45	55	65	75	85	95
Demand (in Kg)	1800	1700	1700	1600	1200	1100	900	500	200

[Ans : $r = -0.97$]

16. Calculate coefficient of correlation from the following data :

X	3000	5000	6000	8000	10000	11000	13000
Y	0.15	0.25	0.35	0.45	0.55	0.65	0.75

[Ans : $r = 1.00$]

17. Calculate correlation coefficient using the following data : Covariance between X and Y is 101.1, and variance of Y is 170.5 and varience of X is 106.6.

[Ans : $r = 0.75$].

18. Calculate correlation coefficient if $\Sigma xy = 413$, $n = 25$, $S_x^2 = 884$, $S_y^2 = 669$ where x and y are deviations of X and Y from $\overline{X}$ and $\overline{Y}$

[$r = 0.54$]

19. In two sets of values of variables X and Y with 50 observations each, the following information is available :

$\Sigma X = 490$, $\Sigma Y = 294$, Standard deviation of $X = 21.21$, variance of $Y = 200$, $\Sigma XY = 3030$. Find $r(X, Y)$.[$r = 0.3$]

20. Find the number of observations if $r = 0.89$, $S_y = 3$, $\Sigma xy = 122$, $\Sigma x^2 = 136$ and x any y are deviations of X and Y taken from their means. [Ans : $n = 15$]

21. For 10 pairs of observations on X and Y following data is recorded

Sum of squares of deviations of X and Y taken from their means is 5389 and 2224 respectively.

Sum of products of deviations taken from their means = 2704

Calculate correlation coefficient. [Ans : $r = 0.78$]

22. Calculate correlation coefficient using following data

$n = 10$, $\Sigma X = 650$, $\Sigma Y = 660$, $\Sigma(X - 65)^2 = 15398$

$\Sigma(Y - 66)^2 = 12224$, $\Sigma(X - 65)(Y - 66) = 12704$

[Ans : $r = 0.93$]

23. Calculate correlation coefficient using following data

$n = 12$, $\Sigma X = 1296$, $\Sigma Y = 684$, $\Sigma(X - 100)^2 = 96$,

$\Sigma(Y - 50)^2 = 1380$

$\Sigma(X - 100)(Y - 50) = 312$ [Ans : $r = 0.67$]

24. In order to find the correlation coefficient between two variables X and Y from 12 pairs of observations the following results are available:

$\Sigma X = 30$, $\Sigma Y = 5$, $\Sigma X^2 = 670$, $\Sigma Y^2 = 285$, $\Sigma XY = 344$.

Later it was found that one particular pair of observation (10, 14) was wrongly taken as (11, 4). Find the correct value of the correlation coefficient. [Ans : $r = 0.78$]

25. For 50 pairs of observations, $\overline{X} = 10$, $\overline{Y} = 6$, $S_x = 3$, $S_y = 2$, $r = 0.3$. Later on one pair (10, 6) was omitted. Calculate the correlation coefficient between X and Y. [Ans : $r = 0.3$]

26. Calculate coefficient of correlation between the marks obtained by a batch of 100 students in Accountancy and Statistics as given in the following table. Also calculate coefficient of determination.

Marks in statistics	Marks in Accountancy				
	20–30	30–40	40–50	50–60	60–70
15–25	5	9	3	–	–
25–35	–	10	25	2	–
35–45	–	1	12	2	–
45–55	–	–	4	16	5
55–65	–	–	–	4	2

[Ans. : $r = 0.80$, $R^2 = 0.64$]

27. The following table shows the frequency distribution of the final grades of 100 students in Mathematics and Accountancy. Determine coefficient of correlation.

Maks in Maths	Maks in Accountancy					
	40–49	50–59	60–69	70–79	80–89	90–99
90–99	–	–	–	2	4	4
80–89	–	–	1	4	6	5
70–79	–	–	5	10	8	1
60–69	1	4	9	5	2	–
50–59	3	6	6	2	–	–
40–49	2	5	4	–	–	–

[Ans. : $r = 0.77$]

28. From the bivariate table given below calculate coefficient of correlation between the ages of fathers and sons.

Age of Fathers	Age of sons							
	2	6	10	14	18	22	26	30
20–25	27	10	2	–	–	–	–	–
25–30	15	26	20	1	–	–	–	–
30–35	6	12	25	16	–	–	–	–
35–40	–	–	15	20	8	–	–	–
40–45	–	–	–	14	18	3	–	–
45–50	–	–	–	2	13	8	4	–
50–55	–	–	–	–	8	10	6	2
55–60	–	–	–	–	–	6	5	3

[Ans. : $r = 0.98$]

29. Ten competitors were given marks in a beauty contest by 3 judges as follows. Use rank correlation coefficient to determine which of the 2 Judges have similar approach, in terms of common tastes and liking for beauty.

Participants :	1	2	3	4	5	6	7	8	9	10
Judge A :	45	15	51	18	30	20	5	40	10	28
Judge B :	22	5	40	25	17	9	32	16	37	2
Judge C :	19	31	21	42	30	50	11	10	60	27

[Ans. : $r_{AB} = -0.01$, $r_{AC} = -0.03$, $r_{BC} = -0.43$
Judge A and B have nearest approach]

30. Find the rank correlation coefficient from the following data.

Marks in statistics :	22	27	28	15	32	14	17	09	05	06
Marks in Mathematics :	32	39	40	20	21	22	16	35	48	16

[Ans : $r = 0.03$]

31. Calculate coefficient of correlation between advertisement cost and sales from the data given below, and find coefficient of determination.

Advertisement cost	39	65	62	90	82	75	25	98	36	78
Sales	47	53	58	86	62	68	60	91	51	84

[Ans : $r = 0.82$, $R^2 = 0.67$]

32. Calculate the Spearman's coefficient of correlation from the following bivariate data.

X :	48	33	40	9	16	16	65	24	16	57
Y :	13	13	24	6	15	4	20	9	6	19

[Ans. : $r = 0.73$]

33. Find rank correlation coefficient from the following data.

X :	–1.03	–1.03	1.17	2.4	–2.04	1.3	0	0.3	–1.03	–1.3
Y :	–1.02	0	0.8	2.1	1.02	1.2	2.3	0	–0.08	–0.8

[Ans : $r = 0.55$]

34. Rank correlation coefficient between 10 pairs of X and Y was obtained as 0.54. Later on it noticed that one of the differences of ranks was taken as (–2) instead of (–3). Find correct rank correlation coefficient. [Ans. : $r = 0.51$]

35. Rank correlation coefficient between 12 pairs of X and Y was obtained as 0.525. Later on it was noticed that one of the differences of ranks was wrongly taken as 5 instead of 4. Find correct rank correlation coefficient. [Ans. : $r = 0.56$]

36. The coefficient of rank correlation is 2/3 and the sum of the squares of the differences in ranks is 55. Find the number of observations. [Ans. : $n = 10$]

37. The coefficient of rank correlation between X and Y was obtained as – 0.05 and the sum of the squares of the differences in ranks is 126. Find the number of observations.

[Ans. : $n = 9$]

$$\therefore \quad a_1a_2X + a_2b_1Y + a_2c_1 = 0$$
$$a_1a_2X + a_1b_2Y + a_1c_2 = 0$$
$$- \quad - \quad -$$
$$\therefore \quad a_2b_1X - a_1b_2Y + a_2c_1 - a_1c_2 = 0$$
$$\therefore \quad Y(a_2b_1 - a_1b_2) = a_1c_2 - a_2c_1$$

$$\therefore \quad Y = \frac{a_1c_2 - a_2c_1}{a_2b_1 - a_1b_2} \quad \text{or} \quad \overline{Y} = \frac{a_1c_2 - a_2c_1}{a_2b_1 - a_1b_2}$$

Substitute value of Y in $a_1X + b_1Y + c_1 = 0$

$$\therefore \quad a_1X + b_1\left[\frac{a_1c_2 - a_2c_1}{a_2b_1 - a_1b_2}\right] + c_1 = 0$$

$$\therefore \quad a_1X + \frac{a_1b_1c_2 - a_2b_1c_1}{a_2b_1 - a_1b_2} + c_1 = 0$$

$$\therefore \quad a_1X + \frac{(a_1b_1c_2 - a_2b_1c_1) + c_1(a_2b_1 - a_1b_2)}{a_2b_1 - a_1b_2} = 0$$

$$\therefore \quad a_1X + \frac{a_1b_1c_2 - a_2b_1c_1 + a_2b_1c_1 - a_1b_2c_1}{a_2b_1 - a_1b_2} = 0$$

$$\therefore \quad a_1X + \frac{a_1b_1c_2 - a_1b_2c_1}{a_2b_1 - a_1b_2} = 0$$

$$\therefore \quad a_1X = -\frac{(a_1b_1c_2 - a_1b_2c_1)}{a_2b_1 - a_1b_2}$$

$$\therefore \quad X = -\frac{(a_1b_1c_2 - a_1b_2c_1)}{a_1(a_2b_1 - a_1b_2)}$$

$$\therefore \quad X = \frac{-a_1(b_1c_2 - b_2c_1)}{a_1(a_2b_1 - a_1b_2)}$$

$$\therefore \quad X = \frac{b_2c_1 - b_1c_2}{a_2b_1 - a_1b_2} \quad \text{or} \quad \overline{X} = \frac{b_2c_1 - b_1c_2}{a_2b_1 - a_1b_2}$$

EXERCISE

1. What is regression ? Explain different types of regression.
2. Explain regression lines. Why there are two regression lines ?
3. State the utility of regression analysis.
4. Clarify difference between correlation and regression analysis.
5. State the properties of regression coefficients.
6. Show that correlation coefficient is a geometric mean of regression coefficients.
7. Show that $\frac{b_{yx} + b_{xy}}{2} \geq r$.
8. Fill up the blanks :
   (1) The methods of fitting of regression line are _____ and _____.
   (2) If both regression lines coincide then r = _____.
   (3) If regression lines are perpendicular to each other then r = _____.
   (4) Correlation coefficient is the _____ mean of regression coefficient.
   (5) The two regression lines intersect each other at the point (_____, _____)
   (6) The possible change in the value of y with unit change in the value of x can be given by _____.
   (7) If two regression coefficients are – 0.9 and – 0.4 then correlation coefficient is _____.
   (8) Two variables are perfectly correlated. If $b_{yx} = \frac{1}{25}$ then b_{xy} = ...
   (9) Greater the angle between the regression lines, ... the correlation between the variables.
   (10) Regression coefficients are independent of change of ..... but not .....

Ans. : (1) scatter diagram, method of least square (2) ± 1 (3) 0 (4) geometric (5) $\bar{x}$, $\bar{y}$ (6) $b_{..}$ (7) − 0.6 (8) b_{yx} = 25 (9) lesser (10) origin, scale

9. The following data relate to the age of husbands and wifes.

Age of husbands	25	28	30	32	35	36	38	39	42	45
Age of wifes	20	26	29	30	25	18	26	34	35	46

Obtain the two regression equations and determine the most likely age of husbands for age of wife 25 yrs. and most likely age of wife for age of husband 35 yrs.

[Ans. : X = 19.28 + 0.542Y, X = 32.83 for Y = 25, Y = 2.675 + 0.905 X, Y = 24.475 for X = 30]

10. Obtain the regression line of cost on percent activity using the following data

Percent activity (X)	60	20	100	40	80
Cost (Y)	900	1100	500	800	700

[Ans. : Y = 1190 − 6.5 x]

11. From the following data, find the two regression equations

X	4	5	6	7	1	2	3
Y	6	5	6	5	2	4	7

[Ans. : Y = 3.44 + 0.39 X, X = 0.55 + 0.69 Y]

12. Obtain the regression line of blood-pressure on Age in year and estimate the blood pressure (B.P.) when the age is 50.

Age	56	42	72	36	63	47	60	68	42	38	49	55
B.P.	147	125	160	118	149	128	155	152	140	111	145	150

[Ans. : Y = 83.29 + 1.09X, Y = 137.79 when X = 50]

13. The following data give the monthly income and expenditure on food of 10 families.

Income	120	90	80	150	130	140	110	95	75	105
Expenditure	40	36	40	45	40	44	45	38	50	35

Calculate the linear regression of expenditure on food on income.

[Ans. : Y = 40.10 + 0.011X]

14. From the following data, obtain the two regression equations :

Sales (X)	91	97	108	121	67	124	51	73	111	57
Purchase (Y)	71	75	69	97	70	91	39	61	80	47

[Ans. : Y = 15.1 + 0.61 X, X = − 3.7 + 1.36Y]

15. Obtain regression line of profit on recurring expenses.

Recurring expenses (X) (lakh Rs)	10	40	50	110	80
Profit (Y) (1000 Rs)	53	54	52	3	41

[Ans. : Y = 5.32 − 0.01 X]

16. The following are the marks obtained by 132 students in test X and test Y.

Y/X	30–40	40–50	50–60	60–70	70–80
20–30	2	5	3	–	–
30–40	1	8	12	6	–
40–50	–	5	22	14	1
50–60	–	2	16	9	2
60–70	–	1	8	6	1
70–80	–	–	2	4	2

Obtain both linear regression equations.

[Ans. : Y = 8.11 + 0.7X, X = 42.93 + 0.29Y]

17. The following table gives the distribution of total cultivable area (X) in hectare and the area under cultivation of rice (Y) in hectare in a distict of 70 villages.

Area under Rice (Y)	Total cultivable area (X)				
	0–50	50–100	100–150	150–200	200–250
0–20	12	16	–	–	–
20–40	2	18	4	2	1
40–60	–	4	7	–	–
60–80	–	1	–	2	3
80–100	–	–	1	2	3

Obtain the Y on X regression line.

[Ans. : Y = 47.98 + 0.32X]