CHAPTER 31

Statistics

Exercise

1. The mean of n observations is $\bar{x}$. If one observation x_{n+1} is added, then the mean continues to be the same. The value of x_{n+1} is
(a) 0 (b) 1
(c) n (d) $\bar{x}$

2. In an examination the standard of passing was 40%. Out of 9 students who appeared, 4 failed and the remaining got 80%, 57%, 51%, 68% and 79% marks. The median of the percentage marks is equal to
(a) 51% (b) 57%
(c) 68% (d) 79%

3. When the number of classes is increased indefinitely and the width of the classes is decreased indefinitely, then the frequency polygon becomes
(a) histogram (b) frequency curve
(c) pie-chart (d) line graph

4. Consider the following frequency distribution.

Class Interval	0–10	10–20	20–30	30–40	40–50
Frequency	14	f_1	28	f_2	15

If the sum of the frequencies is 100 and median be 25, then f_1 and f_2 will be
(a) (15, 28) (b) (20, 23)
(c) (22, 21) (d) (21, 22)

5. The statistical data regarding production of food grains in India, available in a Government of India Publication, is
(a) primary
(b) secondary
(c) primary as well as secondary
(d) neither primary nor secondary

6. The arithmetic mean of a set of observations is $\bar{X}$. If each observation is divided by α and then is increased by 10, then the mean of the new series is
(a) $\dfrac{\bar{X}}{\alpha}$ (b) $\dfrac{\bar{X}+10}{\alpha}$
(c) $\dfrac{\bar{X}+10\alpha}{\alpha}$ (d) $\alpha\bar{X}+10$

7. A student obtains 75%, 80% and 85% marks in three subjects. If the marks of any other subject are added, then their average cannot be less than
(a) 60% (b) 65%
(c) 70% (d) 80%

8. When 10 is substracted from each of the given observations, the mean is reduced to 60%. If 5 is added to all the given observations, the mean will be
(a) 25 (b) 30
(c) 60 (d) 65

9. If the values $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, ..., \frac{1}{n}$ occur at frequencies 1, 2, 3, 4, 5, ..., n respectively, in a frequency distribution, then the mean is
(a) 1 (b) n
(c) $\dfrac{1}{n}$ (d) $\dfrac{2}{n+1}$

10. In a class of 50 students, 10 have failed and their average marks are 28. The total marks obtained by the entire class are 2800. The average marks of those who have passed are
(a) 43 (b) 53
(c) 63 (d) 70

11. The weighted means of first n natural numbers whose weights are equal to the squares of corresponding numbers is
(a) $\dfrac{n+1}{2}$ (b) $\dfrac{3n(n+1)}{2(2n+1)}$

(c) $\frac{(n+1)(2n+1)}{6}$ (d) $\frac{n(n+1)}{2}$

12. The mean of n items is $\overline{X}$. If the first item is increased by 1, second by 2 and so on, then the new mean is
(a) $\overline{X}+n$ (b) $\overline{X}+\frac{n}{2}$
(c) $\overline{X}+\frac{n+1}{2}$ (d) None of these

13. The standard deviation of 7, 9, 11, 13, 15 is
(a) 2.42 (b) 2.50
(c) 2.75 (d) 2.82

14. The expenditure of 100 families are given below.

Expenditure	Number of families
0 – 10	14
10 – 20	x_1
20 – 30	27
30 – 40	x_2
40 – 50	15

The mode of the distribution is 25, then the missing frequencies are
(a) 22 and 22 (b) 23 and 21
(c) 24 and 20 (d) None of these

15. If mean and SD of a distribution be 10 and 8 respectively, what will be the coefficient of variation of that distribution?
(a) 0.8 (b) 8
(c) 1.25 (d) 80

16. A student obtained 80% in English, 88% in Mathematics, 50% in Chemistry and 79% in Physics. If Mathematics and Physics are considered twice as important as English and Chemistry, then the weighted mean of marks obtained by the student is
(a) 70 (b) 78.33
(c) 74.33 (d) None of these

17. In a frequency distribution, class marks are 37, 47, 57, 67. Its class boundaries are given by
(a) 34.5–44.5, 44.5–54.5, 54.5–64.5
(b) 31.5–41.5, 41.5–51.5, 51.5–61.5
(c) 32–42, 42–52, 52–62, 62–72
(d) 25–45, 45–55, 55–65, 65–75

18. If the mean of the following data is 13.5, then the value of p is

x	5	10	15	20	25
f	10	10	p	2	8

(a) 10 (b) 15
(c) 18 (d) 21

19. The mean of the following distribution is 50, but the frequencies of 20-40 and 60-80 are unknown. They are represented by f_2 and f_4 respectively.

Group	Frequency
0 – 20	17
20 – 40	f_2
40 – 60	32
60 – 80	f_4
80 – 100	19
Total	120

f_2 and f_4 are equal to
(a) 18, 14 (b) 28, 24
(c) 20, 24 (d) 22, 28

20. What is the mode of the data 20, 20, 20, 21, 21, 21, 21, 21, 22, 22, 22, 22, 22, 22, 22, 23, 23, 23, 23, 23, 24, 24 and 25?
(a) 20 (b) 21
(c) 22 (d) 25

21. Match List-I with List-II and select the correct answer using the codes given below the lists.

List-I		List-II	
A.	Correlation	1.	Cumulative frequency curves
B.	Arithmetic mean	2.	Continuous distribution
C.	Normal Distribution	3.	Central tendency based on all the observations in the data
D.	Ogives	4.	Independent of the unit of measurements

Codes:

	A	B	C	D		A	B	C	D
(a)	1	2	3	4	(b)	1	3	2	4
(c)	4	2	3	1	(d)	4	3	2	1

22. For dealing with qualitative data, the best average is
(a) AM (b) GM
(c) HM (d) Median

23. The positional average of central tendency is
(a) GM (b) HM
(c) AM (d) Median

24. The mean and standard deviation of a set of values are 5 and 2, respectively. If 5 is added to each value, then what is the coefficient of variation for the new set of values ?
(a) 10 (b) 20
(c) 40 (d) 70

25. In a frequency distribution, the frequencies of two classes are missing. However, the total frequency and the median of the complete data are known. The data is as follows.

Class Intervals	Frequency
0 – 10	6
10 – 20	—
20 – 30	30
30 – 40	28
40 – 50	—

The total frequency is 80 and the median is 28. The frequency of the class 40–50 is

(a) 9 (b) 6
(c) 7 (d) 8

26. Expenditure of a family in different items is as follows

Food	Rent of house	Cloth	Education	Others
₹ 560	₹ 420	₹ 180	₹ 160	₹ 120

If these data are represented by pie-chart, what will be angle of sector corresponding to the expenditure of cloth?

(a) 45° (b) 90°
(c) 54° (d) 180°

27. If standard deviation of 5 numbers be 2.5, what will be their variance?

(a) 2.25 (b) 6.25
(c) $\sqrt{2.5}$ (d) 9.25

28. What is the mean deviation from the mean of the number 10, 9, 21, 16, 24? **[NDA-I 2016]**

(a) 5.2 (b) 5.0
(c) 4.5 (d) 4.0

29. If the total number of observations is 20, $\Sigma x_i = 1000$ and $\Sigma x_i^2 = 84000$, then what is the variance of the distribution? **[NDA-I 2016]**

(a) 1500 (b) 1600
(c) 1700 (d) 1800

30. If m is the geometric mean of

$$\left(\frac{y}{z}\right)^{\log(yz)}, \left(\frac{z}{x}\right)^{\log(zx)}, \left(\frac{x}{y}\right)^{\log(xy)}$$, then what is the value of m? **[NDA-I 2016]**

(a) 1 (b) 3
(c) 6 (d) 9

31. The mean of the series $x_1, x_2, \ldots\ldots, x_n$ is $\bar{x}$. If x_2 is replaced by λ, then what is the new mean? **[NDA-I 2016]**

(a) $\bar{x} - x_2 + \lambda$ (b) $\dfrac{\bar{x} - x_2 - \lambda}{n}$

(c) $\dfrac{\bar{x} - x_2 + \lambda}{n}$ (d) $\dfrac{n\bar{x} - x_2 + \lambda}{n}$

32. For the data 3, 5, 1, 6, 5, 9, 5, 2, 8, 6 the mean, median and mode are x, y and z respectively. Which one of the following is correct? **[NDA-I 2016]**

(a) $x = y \neq z$ (b) $x \neq y = z$
(c) $x \neq y \neq z$ (d) $x = y = z$

33. Consider the following statements in respect of a histogram **[NDA-I 2016]**

1. The total area of the rectangles in a histogram is equal to the total area bounded by the corresponding frequency polygon and the X-axis.
2. When class intervals are unequal in a frequency distribution, the area of the rectangle is proportion to the frequency.

Which of the above statements is/are correct?

(a) Only 1 (b) Only 2
(c) Both 1 and 2 (d) Neither 1 nor 2

34. The scores of 15 students in an examination were recorded as 10, 5, 8, 16, 20, 18, 8, 10, 16, 20, 18, 11, 16, 14 and 12. After calculating the mean, median and mode, an error is found. One of the value is wrongly written as 16 instead of 18. Which of the following measures of central tendency will change? **[NDA-II 2016]**

(a) Mean and median (b) Median and mode
(c) Only Mode (d) Mean and Mode

35. Two variates, x and y are uncorrelated and have standard deviations σ_x and σ_y respectively, what is the correlation coefficient between $x + y$ and $x - y$? **[NDA-II 2016]**

(a) $\dfrac{\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2}$ (b) $\dfrac{\sigma_x + \sigma_y}{2\sigma_x\sigma_y}$

(c) $\dfrac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$ (d) $\dfrac{\sigma_y - \sigma_x}{\sigma_x\sigma_y}$

36. For 10 observations price (x) and supply (y), the following data was obtained **[NDA-II 2016]**

$\Sigma x = 130$, $\Sigma y = 220$, $\Sigma x^2 = 2288$

$\Sigma y^2 = 5506$ and $\Sigma xy = 3467$

What is the line of regression of y on x?

(a) $y = 0.91\,x + 8.74$ (b) $y = 1.02\,x + 8.74$
(c) $y = 1.02\,x - 7.02$ (d) $y = 0.91\,x - 7.02$

37. In a study of two groups, the following results were obtained. **[NDA-II 2016]**

	Group A	Group B
Sample Size	20	25
Sample mean	22	23
Sample standard deviation	10	12

Which of the following statements is correct ?

(a) Group A is less variable than Group B, because group A's standard deviation is smaller.
(b) Group A is less variable than Group B, because Group B's sample size is smaller.
(c) Group A is less variable than Group B, because Group A's sample mean is smaller.

(d) Group A is less variable than Group B, because Group A's coefficient of variation is smaller.

38. Consider the following statements in respect of class intervals of grouped frequency distribution. **[NDA-II 2016]**
    1. Class intervals need not be mutually exclusive.
    2. Class intervals should be exhaustive.
    3. Class intervals need not be of equal width.

    Which of the above statements are correct?
    (a) 1 and 2 (b) 2 and 3
    (c) 1 and 3 (d) 1, 2 and 3

39. If the covariance between x and y is 30, variance of x is 25 and variance of y is 144, then what is correlation coefficient? **[NDA-II 2016]**
    (a) 0.4 (b) 0.5
    (c) 0.6 (d) 0.7

40. A random sample of 20 people is classified in the following table according to their ages **[NDA-II 2016]**

Age	Frequency
15 – 25	2
25 – 35	4
35 – 45	6
45 – 55	5
55 – 65	3

What is the mean of this group of people?
(a) 41.0 (b) 41.5
(c) 42.0 (d) 42.5

41. The mean weight of 150 students in a certain class is 60 kg. The mean weight of boys in the class is 70 kg and that of girls is 55 kg. What is the number of boys in the class? **[NDA-I 2017] [NDA-II 2019]**
    (a) 50 (b) 55
    (c) 60 (d) 100

42. Data can be represented in which of the following forms?
    1. Textual forms 2. Tabular form
    3. Graphical form

    Select the correct answer using the code given below. **[NDA-I 2017]**
    (a) 1 and 2 only (b) 2 and 3 only
    (c) 1 and 3 only (d) 1, 2 and 3

43. If the data are moderately non-symmetrical, then which one of the following empirical relationships is correct? **[NDA-I 2017]**
    (a) 2 × standard deviation = 5 × mean deviation
    (b) 5 × standard deviation = 2 × mean deviation
    (c) 4 × standard deviation = 5 × mean deviation
    (d) 5 × standard deviation = 4 × mean deviation

44. For given statistical data, the graphs for less than ogive and more than ogive are drawn. If the point at which the two curves intersect is P, then abscissa of point P gives the value of which one of the following measures of central tendency? **[NDA-I 2017]**
    (a) Median (b) Mean
    (c) Mode (d) Geometric mean

45. If two regression lines between height (x) and weight (y) are $4y - 15x + 410 = 0$ and $30x - 2y - 825 = 0$, then what will be the correlation coefficient between height and weight? **[NDA-I 2017]**
    (a) 1/3 (b) 1/2
    (c) 2/3 (d) 3/4

46. The mean of a group of 100 observations was found to be 20. Later it was found that four observations were incorrect, which were recorded as 21, 21, 18 and 20. What is the mean if the incorrect observations are omitted? **[NDA-I 2017]**
    (a) 18 (b) 20
    (c) 21 (d) 22

47. If coefficient of regressions of x on y and y on x are $-\frac{1}{2}$ and $-\frac{1}{8}$ respectively, then what is the correlation coefficient between x and y? **[NDA-I 2017]**
    (a) $-\frac{1}{4}$ (b) $-\frac{1}{16}$
    (c) $\frac{1}{16}$ (d) $\frac{1}{4}$

48. The variance of 20 observations is 5. If each observation is multiplied by 3, then what is the new variance of the resulting observations? **[NDA-I 2017]**
    (a) 5 (b) 10
    (c) 15 (d) 45

49. In an examination, 40% of candidates, got second class. When the data are represented by a pie-chart, what is the angle corresponding to second class? **[NDA-I 2017]**
    (a) 40° (b) 90°
    (c) 144° (d) 320°

50. Consider the following statements. **[NDA-I 2017]**
    Statement 1: Range is not a good measure of disperison.
    Statement 2: Range is highly affected by the existence of extreme values.
    Which one of the following is correct in respect of the above statement?
    (a) Both Statement 1 and Statement 2 are correct and Statement 2 is the correct explanation of Statement 1.
    (b) Both Statement 1 and Statement 2 are correct but Statement 2 is not the correct explanation of Statement 1.
    (c) Statement 1 is correct but Statement 2 is not correct.
    (d) Statement 2 is correct but Statement 1 is not correct.

51. Compute the mode for the following frequency distribution **[NDA-II 2017]**

x	95	105	115	125	135	145	155	165	175
f	4	2	18	22	21	19	10	3	2

(a) 115 (b) 125
(c) 22 (d) 120

52. The following table gives the monthly expenditure of two families.

Expenditure (in ₹)		
Items	**Family A**	**Family B**
Food	3,500	2,700
Clothing	500	800
Rent	1,500	1,000
Education	2,000	1,800
Miscellaneous	2,500	1,800

In constructing a pie diagram to the above data, the radii of the circles are to be chosen by which one of the following ratios? **[NDA-II 2017]**
(a) 1 : 1 (b) 10 : 9
(c) 100 : 91 (d) 5 : 4

53. The coefficient of correlation when coefficients of regression are 0.2 and 1.8 is **[NDA-II 2017]**
(a) 0.36 (b) 0.2
(c) 0.6 (d) 0.9

54. If a variable takes value 0, 1, 2, ..., n with frequencies 1, ${}^nC_1$, ${}^nC_2$, ..., ${}^nC_n$ then the AM is **[NDA-II 2017]**
(a) n (b) $\frac{2^n}{n}$
(c) $n+1$ (d) $\frac{n}{2}$

55. Which one of the following can be considered as appropriate pair of values of regression coefficient of y on x and regression coefficient of x on y? **[NDA-II 2017]**
(a) (1, 1) (b) (– 1, 1)
(c) $\left(-\frac{1}{2}, 2\right)$ (d) $\left(\frac{1}{3}, \frac{10}{3}\right)$

56. Consider the following statements **[NDA-II 2017]**
1. Coefficient of variation depends on the unit of measurement of the variable.
2. Range is a measure of dispersion.
3. Mean deviation is least when measured about median.

Which of the above statements are correct?
(a) 1 and 2 only (b) 2 and 3 only
(c) 1 and 3 only (d) 1, 2 and 3

57. It is given that $\bar{x} = 10, \bar{y} = 90, \sigma_x = 3, \sigma_y = 12$ and $r_{xy} = 0.8$. Then, regression equation of x and y is **[NDA-II 2017]**
(a) $y = 3.2\,x + 58$ (b) $x = 3.2\,y + 58$
(c) $x = -8 + 0.2\,y$ (d) $y = -8 + 0.2\,x$

58. Consider the following statements **[NDA-II 2017]**
1. Variance is unaffected by change of origin and change of scale.
2. Coefficient of variance is independent of the unit of observations

Which of the following statements given above is/are correct?
(a) Only 1 (b) Only 2
(c) Both 1 and 2 (d) Neither 1 nor 2

59. If the correlation coefficient between x and y is 0.6, covariance is 27 and variance of y is 25, then what is the variance of x ? **[NDA-I 2018]**
(a) 9/5 (b) 81/25
(c) 9 (d) 81

60. Let $\bar{x}$ be the mean of x_1, x_2, x_3 ....., x_n. If $x_i = a + cy_i$ for some constants a and c, then what will be the mean of y_1, y_2, y_3 ....., y_n? **[NDA-I 2018]**
(a) $a + c\bar{x}$ (b) $a - \frac{1}{c}\bar{x}$
(c) $\frac{1}{c}\bar{x} - a$ (d) $\frac{\bar{x} - a}{c}$

61. Consider the following statements.
I. If the correlation coefficient $r_{xy} = 0$, then the two lines of regression are parallel to each other.
II. If the correlation coefficient $r_{xy} = +1$, then the two lines of regression are perpendicular to each other.

Which of the above statement(s) is are correct ? **[NDA-I 2018]**
(a) Only I (b) Only II
(c) Both I and II (d) Neither I nor II

62. If $4x - 5y + 33 = 0$ and $20x - 9y = 107$ are two lines of regression, then what are the value of $\bar{x}$ and $\bar{y}$, respectively ? **[NDA-I 2018]**
(a) 12 and 18 (b) 18 and 12
(c) 13 and 17 (d) 17 and 13

63. Consider the following statements.
I. Mean is independent of change in scale and change in origin.
II. Variance is independent of change in scale but not in origin.

Which of the above statement(s) is/are correct ? **[NDA-I 2018]**
(a) Only I (b) Only II
(c) Both I and II (d) Neither I nor II

64. Consider the following statements.
I. The sum of deviations from mean is always zero.
II. The sum of absolute deviations is minimum when taken around median.

Which of the above statement(s) is/are correct? **[NDA-I 2018]**

(a) Only I (b) Only II
(c) Both I and II (d) Neither I nor II

65. What is the median of the numbers 4.6, 0, 9.3, –4.8, 7.6, 2.3, 12.7, 3.5, 8.2, 6.1, 3.9, 5.2 ? **[NDA-I 2018]**

(a) 3.8 (b) 4.9
(c) 5.7 (d) 6.0

66. In a test of Mathematics, 20% of the students obtained first class. If the data are represented by a pie-chart, what is the central angle corresponding to first class ? **[NDA-I 2018]**

(a) 20° (b) 36°
(c) 72° (d) 144°

67. A train covers the first 5 km of its journey at a speed of 30 km/h and the next 15 km at a speed of 45 km/h. What is the average speed of the train? **[NDA-I 2018]**

(a) 35 km/h (b) 37.5 km/h
(c) 39.5 km/h (d) 40 km/h

68. The standard deviation σ of the first N natural numbers can be obtained using which one of the following formulae? **[NDA-I 2018]**

(a) $\sigma = \frac{N^2-1}{12}$ (b) $\sigma = \sqrt{\frac{N^2-1}{12}}$

(c) $\sigma = \sqrt{\frac{N-1}{12}}$ (d) $\sigma = \sqrt{\frac{N^2-1}{6N}}$

69. Consider the following statements:
1. If 10 is added to each entry on a list, then the average increases by 10.
2. If 10 is added to each entry on a list, then the standard deviation increases by 10.
3. If each entry on a list is doubled, then the average doubles.

Which of the above statements are correct? **[NDA-II 2018]**

(a) 1, 2 and 3 (b) 1 and 2 only
(c) 1 and 3 only (d) 2 and 3 only

70. The variance of 25 observations is 4. If 2 is added to each observation, then the new variance of the resulting observations is **[NDA-II 2018]**

(a) 2 (b) 4
(c) 6 (d) 8

ANSWERS

1.	(d)	**2.**	(a)	**3.**	(b)	**4.**	(c)	**5.**	(c)	**6.**	(c)	**7.**	(a)	**8.**	(b)	**9.**	(d)	**10.**	(c)
11.	(b)	**12.**	(c)	**13.**	(d)	**14.**	(a)	**15.**	(d)	**16.**	(c)	**17.**	(c)	**18.**	(a)	**19.**	(b)	**20.**	(c)
21.	(d)	**22.**	(d)	**23.**	(d)	**24.**	(b)	**25.**	(b)	**26.**	(a)	**27.**	(b)	**28.**	(a)	**29.**	(c)	**30.**	(a)
31.	(d)	**32.**	(d)	**33.**	(c)	**34.**	(d)	**35.**	(c)	**36.**	(b)	**37.**	(d)	**38.**	(a)	**39.**	(b)	**40.**	(b)
41.	(a)	**42.**	(b)	**43.**	(c)	**44.**	(a)	**45.**	(b)	**46.**	(b)	**47.**	(a)	**48.**	(d)	**49.**	(c)	**50.**	(a)
51.	(b)	**52.**	(b)	**53.**	(c)	**54.**	(d)	**55.**	(a)	**56.**	(b)	**57.**	(d)	**58.**	(b)	**59.**	(d)	**60.**	(d)
61.	(d)	**62.**	(c)	**63.**	(d)	**64.**	(c)	**65.**	(b)	**66.**	(c)	**67.**	(d)	**68.**	(b)	**69.**	(c)	**70.**	(b)

Explanations

1. (d) Given $\frac{x_1+x_2+\ldots+x_n}{n} = \overline{x}$

When one observation x_{n+1} is added, mean remains same.

i.e., $\frac{x_1+x_2+\ldots+x_n+x_{n+1}}{n+1} = \overline{x}$

$\Rightarrow x_1+x_2+\ldots+x_n+x_{n+1} = (n+1)\overline{x}$

$\Rightarrow n\overline{x}+x_{n+1} = n\overline{x}+\overline{x}$

$\Rightarrow x_{n+1} = \overline{x}$

2. (a) Marks of the four failed students would have been less than 51%.

Let their marks be x_1, x_2, x_3 and x_4.

So, arranging the marks in ascending order

x_1, x_2, x_3, x_4, 51, 57, 68, 79, 80.

Number of terms = 9 (odd)

So, median $= \frac{n+1}{2}$ th term = 5th term = 51%

4. (c) Median = 25

Median class = 20 – 30

Let $f_1 = x$ then $f_2 = 43 - x$

Class Interval	f	$c.f$
0 – 10	14	14
10 – 20	x	$14 + x$
20 – 30	28	$42 + x$
30 – 40	$43 - x$	85
40 – 50	15	100

Median $= L_1 + \dfrac{(L_2 - L_1)}{f} \times \left(\dfrac{N}{2} - C\right)$

$\Rightarrow \quad 25 = 20 + \dfrac{10}{28}[50 - (14 + x)]$

$\Rightarrow \quad x = 22 = f_1$ and $f_2 = 43 - 22 = 21$

6. (c) Given $\overline{X} = \dfrac{x_1 + x_2 + ... + x_n}{n}$

Since, each observation is divided by α and then increased by 10. Therefore, new mean

$= \dfrac{1}{\alpha}\overline{X} + 10 = \dfrac{\overline{X} + 10\alpha}{\alpha}$

7. (a) Total marks in 3 subjects = 75 + 80 + 85 = 240 out of 300

Let the minimum marks in fourth subject = 0

So, average marks of 4 subjects $= \dfrac{240}{4}$

$\Rightarrow$ Average cannot be lesser than 60%.

8. (b) Let the mean of $x_1, x_2, ..., x_n$ be $\overline{x}$.

If 10 is subtracted from each observation, then mean $= \overline{x} - 10$

$\overline{x} - 10 = 60\%$ of $\overline{x}$

$\overline{x} - 10 = \dfrac{3}{5}\overline{x} \quad \Rightarrow \quad \dfrac{2}{5}\overline{x} = 10$

$\overline{x} = 25$

So, on adding 5 to each observation, mean will be 30.

9. (d) Mean $= \dfrac{\sum fx}{\sum f} = \dfrac{1 + 1 + 1 + ... + 1}{n\dfrac{(n+1)}{2}} = \dfrac{2}{n+1}$

x	f
1	1
$\dfrac{1}{2}$	2
$\dfrac{1}{3}$	3
....	....
$\dfrac{1}{n}$	n

10. (c) Passed student = 40

Total marks of class = Marks of Passed Students + Marks of failed students

$2800 = x \times 40 + 28 \times 10 \;\Rightarrow\; x = 63$

11. (b) Weighted Mean $= \dfrac{1 \cdot 1^2 + 2 \cdot 2^2 + 3 \cdot 3^2 + ... + n \cdot n^2}{1^2 + 2^2 + ... + n^2}$

$= \dfrac{\sum n^3}{\sum n^2} = \dfrac{n^2(n+1)^2 6}{4 \times n(n+1)(2n+1)} = \dfrac{3n(n+1)}{2(2n+1)}$

12. (c) Given $\overline{X} = \dfrac{x_1 + x_2 + ... + x_n}{n}$

New Mean $= \dfrac{(x_1 + 1) + (x_2 + 1) + ... + (x_n + n)}{n}$

$= \dfrac{x_1 + x_2 + ... + x_n}{n} + \dfrac{(1 + 2 + ... + n)}{n}$

$= \overline{X} + \dfrac{n(n+1)}{2n} = \overline{X} + \dfrac{n+1}{2}$

13. (d) $\overline{X} = \dfrac{7 + 9 + 11 + 13 + 15}{5} = 11$

SD $= \sqrt{\dfrac{\sum |x_i - \overline{x}|^2}{n}} = \sqrt{\dfrac{16 + 4 + 0 + 4 + 16}{5}}$

$= 2\sqrt{2} = 2 \times 1.414 = 2.82$

14. (a) Total families = 100

$\Rightarrow \; x_1 + x_2 = 44$

$\because$ Mode is 25.

So, modal class = 20 – 30

Mode $= L_1 + \dfrac{f_m - f_1}{2f_m - f_1 - f_2} \times (L_2 - L_1)$

$25 = 20 + \dfrac{27 - x_1}{54 - (x_1 + x_2)}(10)$

$5 = \dfrac{27 - x_1}{54 - 44} \times 10$

$5 = 27 - x_1 \Rightarrow x_1 = 22 \;\Rightarrow x_2 = 22$

15. (d) Coefficient of variation

$= \dfrac{\text{SD}}{\text{Mean}} \times 100 = \dfrac{8}{10} \times 100 = 80$

16. (c) Weighted Mean $= \dfrac{80 + 2 \times 88 + 50 + 2 \times 70}{6}$

$= \dfrac{80 + 176 + 50 + 140}{6} = \dfrac{446}{6} = 74.33$

17. (c) Class Marks are the average of class boundaries.

So, class boundaries

= 32 – 42, 42 – 52, 52 – 62, ...

18. (a) Mean $=\dfrac{\sum fx}{\sum f}$

$$13.5=\frac{50+100+15p+40+200}{30+p}$$

$405 + 13.5p = 390 + 15p$

$\Rightarrow\ 1.5p = 15\ \Rightarrow\ p = 10$

19. (b) Let $f_2 = x$, then $f_4 = 52 - x$.

Group	x	f
0 – 20	10	17
20 – 40	30	x
40 – 60	50	32
60 – 80	70	$52 - x$
80 – 100	90	19
Total		120

Mean $=\dfrac{\sum fx}{\sum f}$

$\Rightarrow\ 120\times 50 = 170+30x+1600+3640-70x+1710$

$\Rightarrow\ 6000 = 7120-40x$

$\Rightarrow\ 40x = 1120\ \Rightarrow\ x = 28$

Hence, $f_2 = 28$ and $f_4 = 52 - x = 52 - 28 = 24$

20. (c) In the given observations, 20 occurs 3 times, 21 occurs 5 times, 22 occurs 7 times, 23 occurs 5 times, 24 occurs 2 times and 25 occurs 1 time.

$\because$ Mode = Max times repeating number

So, mode = 22 (7 times)

22. (d) Best average dealing with qualitative data = Median

23. (d) Positional average of central tendency is median.

24. (b) Mean = 5; S.D. = 2

If 5 is added to each value, then new mean = 5 + 5 = 10

While SD will remain same.

So, coefficient of variation

$$=\frac{\text{SD}}{\text{Mean}}\times 100=\frac{2}{10}\times 100=20$$

25. (b) Let the missing frequency of class 10 – 20 be x.

Then,

Class Interval	f	$c.f$
0 – 10	6	6
10 – 20	x	$6 + x$
20 – 30	30	$36 + x$
30 – 40	28	$64 + x$
40 – 50	$16 - x$	80

Median class = 20 – 30

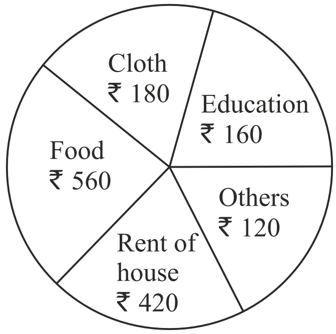
$$\text{Median} = L_1+\frac{(L_2-L_1)}{f}\left(\frac{N}{2}-C\right)$$

$$28 = 20+\frac{10}{30}(40-6-x)$$

$$8=\frac{1}{3}(34-x)\ \Rightarrow x=10$$

So, frequency of class 40 – 50 = 16 – 10 = 6

26. (a)



Angle of sector corresponding to the expenditure of cloth $=\dfrac{180}{1440}\times 360^\circ = 45^\circ$

27. (b) Variance = $(\text{S.D.})^2 = 2.5 \times 2.5 = 6.25$

28. (a) For 10, 9, 21, 16, 24

$$\text{Mean, } \bar{x}=\frac{10+9+21+16+24}{5}=16$$

$$\text{Mean deviation, } \delta=\frac{\Sigma(x_i-\bar{x})}{n}$$

$$=\frac{|-6|+|-7|+|5|+|0|+|8|}{5}=\frac{26}{5}=5.2$$

29. (c) Variance $=\dfrac{\Sigma(x_i-\bar{x})^2}{n}$

$$=\frac{\Sigma x_i^2}{n}-\left(\frac{\Sigma x_i}{n}\right)^2=\frac{84000}{20}-\left(\frac{1000}{20}\right)^2$$

$= 4200 - 2500 = 1700$

30. (a) Let GM

$$=(m)=\left\{\left(\frac{y}{z}\right)^{\log yz}\times\left(\frac{z}{x}\right)^{\log zx}\times\left(\frac{x}{y}\right)^{\log xy}\right\}^{1/3}$$

$$\log m=\frac{1}{3}\left[\log(yz)\log\left(\frac{y}{z}\right)+\log(zx)\log\left(\frac{z}{x}\right)+\log(xy)\log\left(\frac{x}{y}\right)\right]$$

$$=\frac{1}{2}[(\log y+\log z)(\log y-\log z)+(\log z+\log x)(\log z-\log x)+(\log x+\log y)(\log x-\log y)]$$

$$= \frac{1}{3}[(\log y)^2 - (\log z)^2 + (\log z)^2 - (\log x)^2 + (\log x)^2 - (\log z)^2]$$

$\log m = 0 \Rightarrow m = 1$

31. (d) $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$

$\Rightarrow x_1 + x_2 + ... + x_n = n\bar{x}$

Now, x_2 is replaced by λ

So, new mean $= \frac{n\bar{x} - x_2 + \lambda}{n}$

32. (d) Arranging the data in ascending order, 1, 2, 3, 5, 5, 5, 6, 6, 8, 9

Mean $= \frac{1+2+3+5+5+5+6+6+8+9}{10} = 5 = x$

Median $= \frac{\left(\frac{10}{2}\right)\text{th term} + \left(\frac{10}{2}+1\right)\text{th term}}{2}$

$= \frac{\text{5th term} + \text{6th term}}{2} = \frac{5+5}{2} = 5 = y$

Mode = maximum repeating number

So, $z = 5 \Rightarrow x = y = z$

33. (c) Both 1 and 2 are correct.

34. (d) Scores are 10, 5, 8, 16, 20, 18, 8, 10, 16, 20, 18, 11, 16, 14, 12

Mean $= \frac{\sum x_i}{n} = \frac{202}{15} = 13.46$

Total scores are 15.

So, median will be $\frac{15+1}{2}$th term, i.e., 8th score when the data is arranged in ascending order

{5, 8, 8, 10, 10, 11, 12, 14, 16, 16, 16, 18, 18, 20, 20}

So, median = 14

and mode = maximum repeating score = 16

Now, the score changes to 16 then new mean

$= \frac{202+18+18+18-16-16-16}{15}$

$= \frac{208}{15} = 13.87$

Median = 14 and Mode = 18

So, mean and mode will be changed.

35. (c) $\text{var}(x+y) = \text{var}(x) + \text{var}(y) = \sigma_x^2 + \sigma_y^2$

$\text{var}(x-y) = \text{var}(x) + \text{var}(-y) = \text{var}(x) + \text{var}(y)$

$= \sigma_x^2 + \sigma_y^2$

$\text{cov}(x+y, x-y) = \text{cov}(x, x) - \text{cov}(x, y) + \text{cov}(y, x) - \text{cov}(y, y)$

$\text{cov}(x, x) - \text{cov}(y, y) = \text{var}(x) - \text{var}(y) = \sigma_x^2 - \sigma_y^2$

Correlation coefficient

$\rho = \frac{\text{cov}(x+y, x-y)}{\sqrt{\text{var}(x+y)}\sqrt{\text{var}(x-y)}}$

$= \frac{\sigma_x^2 - \sigma_y^2}{\sigma_x^2 + \sigma_y^2}$

36. (b) Regression line of y on x is $y = a + bx$.

$\Rightarrow \sum y = na + b\sum x$ and $\sum xy = a\sum x + b\sum x^2$

Put the given values then,

$220 = 10a + 130b$

$a + 13b = 22$ ...(i)

$3467 = 130a + 2288b$

$130a + 2288b = 3467$ ...(ii)

On solving eq. (i) and (ii), we get

$a = 8.74$ and $b = 1.02$

So, equation of regression line of y on x is

$y = 8.74 + 1.02x$

37. (d) Coefficient of variation $= \frac{\text{SD}}{\text{Mean}} \times 100$

For group A; coefficient of variation

$= \frac{10}{22} \times 100 = 45.45$

for group B; coefficient of variation

$= \frac{12}{23} \times 100 = 52.17$

So, A is less variable than group B; since, group A has less coefficient of variation.

38. (a) Statement 1 and 2 are correct.

39. (b) Correlation coefficient

$\rho = \frac{\text{Cov}(x, y)}{\sqrt{\text{var}(x)}\sqrt{\text{var}(y)}} = \frac{30}{\sqrt{25}\sqrt{144}} = \frac{30}{5 \times 12} = 0.5$

40. (b) Mean age $= \frac{\sum fx}{\sum f}$

Age (C.I.)	x	f
15 – 25	20	2
25 – 35	30	4
35 – 45	40	6
45 – 55	50	5
55 – 66	60	3

$= \frac{20 \times 2 + 30 \times 4 + 40 \times 6 + 50 \times 5 + 60 \times 3}{2+4+6+5+3}$

$= \frac{830}{20} = 41.5$

41\. (a) Let the number of boys = x
Then, the number of girls = $150 - x$
Now, combined mean = 60

$$\Rightarrow \quad 60 = \frac{70 \times x + (150 - x) \times 55}{150}$$

$$\Rightarrow \quad 15x + 150 \times 55 = 150 \times 60$$

$$\Rightarrow \quad 15x = 750 \Rightarrow x = 50$$

So, number of boys = 50

42\. (b) Data can be represented in any of the forms of tabular form and graphical form.

43\. (c) For non-symmetrical data, empirical relationship between standard deviation and mean deviation is
$4 \times \text{SD} = 5 \times \text{MD}$

44\. (a) Less than ogive and more than ogive curves intersect at a point which is the median of the data.

45\. (b) $4y - 15x + 410 = 0$

$$\Rightarrow \quad y = \frac{15}{4}x - \frac{410}{4} \Rightarrow b_{yx} = \frac{15}{4}$$

and $30x - 2y - 825 = 0$

$$\Rightarrow \quad x = \frac{2}{30}y + \frac{825}{30} \text{ and } b_{xy} = \frac{2}{30}$$

Now, correlation coefficient

$$\rho = \sqrt{b_{xy} \cdot b_{yx}} = \sqrt{\frac{15}{4} \times \frac{2}{30}} = \frac{1}{2}$$

46\. (b) Sum of 100 observations = $20 \times 100 = 2000$
When four incorrect observations are omitted, then

$$\text{mean} = \frac{2000 - 21 - 21 - 18 - 20}{96} = \frac{1920}{96} = 20$$

47\. (a) Correlation coefficient,

$$\rho = \sqrt{b_{xy} \cdot b_{xy}} = \sqrt{-\frac{1}{2} \times -\frac{1}{8}} = -\frac{1}{4}$$

48\. (d) Variance $= \sqrt{\dfrac{(x_i - \overline{x})^2}{n}}$

Since, each observation is multiplied by 3, so mean will also be multiplied by 3 and hence, variance will also be multiplied by 3.
So, variance = $5 \times 9 = 45$

49\. (c) Angle corresponding to second class = 40% of 360°

$$= \frac{40}{100} \times 360^\circ = 144^\circ$$

50\. (a) Both statements are correct and Statement 2 is the correct explanation of Statement 1.

51\. (b) The value 125 has the highest frequency 22. So, mode = 125

52\. (b) Total expenditure of family A = 10,000
Total expenditure of family B = 8100

$$\frac{\pi r_A^2}{\pi r_B^2} = \frac{10000}{8100} \Rightarrow \frac{r_A}{r_B} = \frac{10}{9}$$

53\. (c) $b_{xy} = 0.2$ and $b_{yx} = 1.8$
So, coefficient of correlation

$$r = \sqrt{b_{xy} \times b_{yx}} = \sqrt{0.2 \times 1.8} = 0.6$$

54\. (d) $\text{AM} = \dfrac{0.1 + 1\,{}^nC_1 + 2\,{}^nC_2 + \ldots + n \cdot {}^nC_n}{1 + {}^nC_1 + {}^nC_2 + \ldots + {}^nC_n}$

$$= \frac{\sum_{r=0}^{n} r\,{}^nC_r}{2^n} = \frac{n \cdot 2^{n-1}}{2^n} = \frac{n}{2}$$

55\. (a) $\because$ Coefficient of correlation

$$\rho = \sqrt{b_{xy} \cdot b_{xy}} \text{ and } -1 \le \rho \le 1$$

So, b_{xy} and b_{yx} can have the values (1, 1).

56\. (b) Only 2 and 3 statement are correct.

57\. (d) $\overline{x} = 10, \overline{y} = 90, \sigma_x = 3, \sigma_y = 12, r_{xy} = 0.8$

$$r_{xy} = \frac{\text{cov}(x, y)}{\sqrt{\text{var } x}\sqrt{\text{var } y}} = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y}$$

$$0.8 = \frac{\text{cov}(x, y)}{3 \times 12}$$

$\text{cov}(x, y) = 28.8$

$$b_{xy} = \frac{\text{cov}(x, y)}{\sigma_{xy}} = \frac{28.8}{12 \times 12} = 0.2$$

So, regression equation of x on y is

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

$$x - 10 = 0.2(y - 90)$$

$$x = 0.2y - 8 \quad \text{or} \quad x = -8 + 0.2\,y$$

58\. (b) Variance is always independent of the change of a origin but not change of scale.
While coefficient of variance is independent of the unit of observation. So, only Statement 2 is true.

59\. (d) Given, $\rho = 0.6$
Cov $(x, y) = 27$ and Var $((y) = 25$

$$\text{Then, } \rho = \frac{\text{Cov}(x, y)}{\sqrt{\text{Var}(x)}\sqrt{\text{Var}(y)}}$$

$$\Rightarrow 0.6 = \frac{27}{\sqrt{\text{Var}(x)} \cdot \sqrt{25}} \Rightarrow \sqrt{\text{Var}(x)} = \frac{27}{0.6 \times 5} = 9$$

$\therefore$ Var $(x) = 81$

60\. (d) Given, $\overline{x} = \dfrac{x_1 + x_2 + \ldots + x_n}{n}$ i.e., $\overline{x} = \dfrac{\Sigma x_i}{n}$

Given, $x_i = a + cy_i \Rightarrow y_i = \dfrac{x_i - a}{c}$

Mean of $y_1, y_2, y_3, \ldots, y_n$ is

$$\overline{y} = \frac{\Sigma y_i}{n} = \frac{\Sigma \frac{(x_i - a)}{c}}{n} = \frac{1}{c}\left(\frac{\Sigma x_i}{n}\right) - \frac{an}{cn}$$

$$= \frac{\overline{x}}{c} - \frac{a}{c} = \frac{\overline{x} - a}{c}$$

61. (d) If the correlation coefficient $r_{xy} = 0$, then two lines of regression are perpendicular.
If the correlation coefficient $r_{xy} = 1$, then two lines of regression are parallel.
Hence, both the given statements are incorrect.

62. (c) Given, two lines of regression are
$4x - 5y + 33 = 0$ and $20x - 9y - 107 = 0$
These equations can be written as $4\bar{x} - 5\bar{y} + 33 = 0$
and $20\bar{x} - 9\bar{y} - 107 = 0$
On solving these equations, we get
$\bar{x} = 13$ and $\bar{y} = 17$

63. (d) Mean is independent of change of origin and change of scale. While variance is independent of changes of origin but depend on change of scale.
Hence, both statements are wrong.

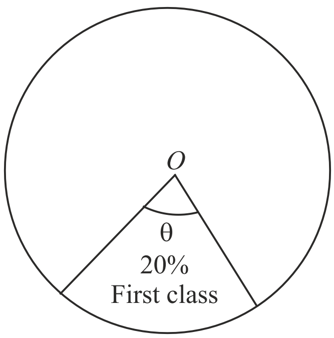
64. (c) Deviation $(\delta) = x_i - \bar{x}$
Hence, sum of deviations from mean is always zero and the sum of absolute deviations is minimum when taken around median.
Hence, both the statements are correct.

65. (b) Arranging the given data in ascending order.
– 4.8, 0, 2.3, 3.5, 3.9, 4.6, 5.2, 6.1, 7.6, 8.2, 9.3, 12.7
Number of terms = 12

$$\text{So, median} = \frac{\left(\frac{12}{2}\text{th}\right)\text{term} + \left(\frac{12}{2}\text{th} + 1\right)\text{term}}{2}$$

$$= \frac{6\text{th term} + 7\text{th term}}{2} = \frac{4.6 + 52}{2} = \frac{9.8}{2} = 4.9$$

66. (c) Let θ be the central angle corresponding to first class.



Then, required central angle of first class of 360°

$$= \frac{20}{100} \times 360^\circ = \frac{1}{5} \times 360^\circ = 72^\circ$$

67. (d) $d_1 = 5$ km, $s_1 = 30$ km/s
$d_2 = 15$ km, $s_2 = 45$ km/h

$$\text{Total time} = \frac{5}{30} + \frac{15}{45} = \frac{1}{6} + \frac{1}{3} = \frac{3}{6} = \frac{1}{2}\text{h}$$

$$\text{So, average speed} = \frac{\text{Total distance}}{\text{Total time}}$$

$$= \frac{20}{1/2} = 40 \text{ km/h}$$

68. (b)
$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{n} x_i^2 - \left(\frac{1}{N}\sum_{i=1}^{n} x_i\right)^2$$

$$= \frac{1}{N}\{\Sigma N^2\} - \left\{\frac{2}{N}\Sigma N\right\}^2$$

$$= \frac{1}{N}\left\{\frac{N(N+1)(2N+1)}{6}\right\} - \left\{\frac{1}{N}\cdot\frac{N(N+1)}{2}\right\}^2$$

$$= \frac{(N+1)(2N+1)}{6} - \frac{N^2 + 1 + 2N}{4}$$

$$= \frac{4N^2 + 6N + 2 - 3N^2 - 3 - 6N}{12} = \frac{N^2 - 1}{12}$$

$$\therefore \ \sigma = \sqrt{\frac{N^2 - 1}{12}}$$

69. (c) Let the n numbers be $x_1, x_2, \ldots, x_n$

$$\text{Then average } \bar{x} = \frac{\Sigma x_i}{n} \text{ and SD } = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n}}$$

Hence, 1 and 3 are correct.

70. (b) $\text{Variance} = \dfrac{\Sigma(x_i - \bar{x})^2}{n}$
If 2 is added to each observation, then mean $\bar{x}$ will also be increased by 2.
Hence there is no change in variance.
So, new variance = 4