CHAPTER 23

Measures of Central Tendency and Dispersion

Chapter Highlights

Measures of central tendency, Arithmetic mean, Geometric mean, Harmonic mean, Median, Quartiles, deciles and percentiles, Mode, Symmetric distribution.

MEASURES OF CENTRAL TENDENCY

For a given data, a single value of the variable which describes its characteristics is identified. This single value is known as the *average*. An average value generally lies in the central part of the distribution and therefore, such values are called the *measures of central tendency*.

The commonly used measures of central tendency are:

1. Arithmetic Mean
2. Geometric Mean
3. Harmonic Mean
4. Median
5. Mode

ARITHMETIC MEAN

Mean of Unclassified Data

Let $x_1, x_2, ..., x_n$ be n observations, then their arithmetic mean is given by,

$$\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Mean of Grouped Data

Let $x_1, x_2, x_3, ..., x_n$ be n observations and let $f_1, f_2, \ldots, f_n$ be their corresponding frequencies, then their arithmetic mean is given by

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + ... + f_nx_n}{f_1 + f_2 + ... + f_n} = \frac{\sum_{i=1}^{n} f_ix_i}{\sum_{i=1}^{n} f_i}$$

Short Cut Method

For the given data, we suitably choose a term, usually the middle term and call it the assumed mean, to be denoted by A.

We find the deviation, $d_i = (x_i - A)$ for each term. Then the arithmetic mean is given by

$$\bar{x} = A + \frac{\sum f_id_i}{\sum f_i}$$

Step Deviation Method

When the class intervals in a grouped data are equal, then the calculations can be simplified further by taking out the common factor from the deviations. This common factor is equal to the width of the class interval. In such cases, the deviation $d_i = x_i - A$, of variates x_i from the assumed mean A are divided by the common factor. The A.M. is then obtained by the following formula:

$$\bar{x} = A + \frac{\sum f_id_i}{N} \times h; N = \sum f_i$$

where A = assumed mean,

$d_i = \frac{x_i - A}{h}$ = the deviation of any variate from A,

h = the width of the class - interval.

Weighted Arithmetic Mean

If $w_1, w_2, w_3, ..., w_n$ are the weights assigned to the values $x_1, x_2, x_3, ..., x_n$ respectively, then the weighted average is defined as:

$$\text{Weighted A.M.} = \frac{w_1x_1 + w_2x_2 + ... + w_n\, x_n}{w_1 + w_2 + ... + w_n}$$

Combined Mean

If we are given the A.M. of two data sets and their sizes, then the combined A.M. of two data sets can be obtained by the formula:

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

where, $\bar{x}_{12}$ = Combined mean of the two data sets 1 and 2
$\bar{x}_1$ = Mean of the first data
$\bar{x}_2$ = Mean of the second data
n_1 = Size of the first data
n_2 = Size of the second data

Properties of A. M.

1. In a statistical data, the sum of the deviations of individual values from A.M. is always zero, i.e.,

$$\sum_{i=1}^{n} f_i\,(x_i - \bar{x}) = 0$$

where f_i is the frequency of x_i $(1 \le i \le n)$

2. In a statistical data, the sum of squares of the deviations of individual values from A.M. is least, i.e.,

$$\sum_{i=1}^{n} f_i\,(x_i - \bar{x})^2 \text{ is least.}$$

3. If each of the n given observations is doubled, then their mean is doubled.

4. If $\bar{x}$ is the mean of $x_1, x_2, \ldots, x_n$, then the mean of $ax_1, ax_2, \ldots, ax_n$ where a is any number different from zero, is $a\bar{x}$.

Some Points About Arithmetic Mean

- Of all types of averages, the arithmetic mean is most commonly used average.
- It is based upon all observations.
- If the number of observations is very large, it is more accurate and more reliable basis for comparison.

GEOMETRIC MEAN

If $x_1, x_2, x_3, \ldots, x_n$ are n observations, none of them being zero, then their geometric mean is defined as

$$\text{G.M.} = (x_1 \cdot x_2 \cdot x_3 \ldots x_n)^{\frac{1}{n}}$$

$$\text{G.M.} = \text{antilog}\left(\frac{\log x_1 + \log x_2 + \ldots\ldots + \log x_n}{n}\right)$$

In the case of a grouped data, geometric mean of n observations $x_1, x_2, \ldots\ldots, x_n$, is given by

$$\text{G.M.} = \left(x_1^{f_1} \cdot x_2^{f_2} \ldots\ldots x_n^{f_n}\right)^{1/N},$$

where $$N = \sum_{i=1}^{n} f_i$$

or $$\text{G.M.} = \text{antilog}\left(\frac{\sum_{i=1}^{n} f_i \log x_i}{N}\right)$$

IMPORTANT POINTS

In the case of continuous or grouped frequency distribution, the values of the variate x are taken to be the values corresponding to the mid-points of the class intervals.

Some Points About Geometric Mean

- It is based on all items of the series.
- It is most suitable for constructing index number, average ratios, percentages etc.
- G.M. cannot be calculated if the size of any of the items is zero or negative.

HARMONIC MEAN

The harmonic mean of n observations $x_1, x_2, \ldots, x_n$ is defined as:

$$\text{H.M.} = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots\ldots + \frac{1}{x_n}}$$

If $x_1, x_2, x_3, \ldots, x_n$ are n observations which occur with frequencies $f_1, f_2, \ldots, f_n$ respectively, then, their H.M. is given by

$$\text{H.M.} = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n}\left(\frac{f_i}{x_i}\right)}$$

Some Points About H.M.

- It is based on all item of the series.
- This is useful in problems related with rates, ratios, time etc.
- A.M. ≥ G.M. ≥ H.M. and also $(\text{G.M.})^2 = (\text{A.M.})(\text{H.M.})$.
- A.M. gives more weightage to larger values whereas G.M. and H.M. give more weightage to smaller values.

Relation Between AM, GM and HM

The arithmetic mean (AM), geometric mean (GM) and harmonic mean (HM) for a given set of observations are related as under:

$$AM \geq GM \geq HM$$

Equality sign holds only when all the observations are equal.

SOLVED EXAMPLES

1. Mean of 25 observations was found to be 78.4. But later on it was found that 96 was misread as 69. The correct mean is

(A) 79.48 (B) 76.54
(C) 81.32 (D) 78.4

Solution: (A)

We know that the mean is given by

$$\bar{x} = \frac{\sum x}{n} \text{ or } \sum x = n\bar{x}$$

Here $\bar{x} = 78.4, n = 25$

$\therefore$ $\Sigma x = 25 \times 78.4 = 1960$

But this Σx is incorrect as 96 was misread as 69.
Correct $\Sigma x = 1960 - 69 + 96 = 1987$

$$\therefore \quad \text{Correct mean} = \frac{1987}{25} = 79.48.$$

2. The mean height of 15 students is 154 cm. It is discovered later on that while calculating the mean the reading 175 cm was wrongly read as 145 cm. The correct mean height is

(A) 145 cm (B) 170 cm
(C) 156 cm (D) None of these

Solution: (C)

Total height of 15 students = $\Sigma x = 154 \times 15 = 2310$ cm.
It was found that 175 cm was wrongly read as 145
Correct sum = 2310 – 145 + 175 = 2340 cm.

$$\text{Correct mean} = \frac{2340}{15} = 156 \text{ cm.}$$

3. A firm of readymade garments make both men's and women's shirts. Its profit average is 6% of sales. Its profits in men's shirts average 8% of sales and women's shirts comprise 60% of output. The average profit per sales rupee in women's shirts is

(A) 0.0466 (B) 0.0166
(C) 0.0666 (D) None of these

Solution: (A)

Here $\bar{x} = 6$, $\bar{x}_1 = 8$, $n_1 = 40$, $n_2 = 60$. Assuming that the total output is 100, we are required to find out $\bar{x}_2$, we know that

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} = \frac{40 \times 8 + 60 \times \bar{x}_2}{40 + 60}$$

$$\Rightarrow \quad 6 = \frac{320 + 60\bar{x}_2}{100}$$

$$\Rightarrow \quad \bar{x}_2 = \frac{600 - 320}{60} = \frac{280}{60} = \frac{14}{3} = 4.66.$$

Thus, the average profit in womens shirt is 4.66 % of sales or Re 0.0466 per sale rupee.

4. The weighted mean of the first n natural numbers if their weights are the same as the numbers, is

(A) $\frac{n+1}{3}$ (B) $\frac{2n+1}{3}$
(C) $\frac{2n-1}{3}$ (D) None of these

Solution: (B)

Here the numbers are 1, 2, 3, ... , n and their weights also are respectively 1, 2, 3, ... , n.

So weighted

$$A_w = \frac{\sum wx}{\sum w} = \frac{1.1 + 2.2 + 3.3 + \ldots\ldots + n.n}{1 + 2 + 3 + \ldots\ldots + n}$$

$$= \frac{1^2 + 2^2 + 3^2 + \ldots.. + n^2}{1 + 2 + 3 + \ldots.. + n}$$

$$= \frac{n(n+1)(2n+1)}{6 \cdot \frac{n(n+1)}{2}} = \frac{2n+1}{3}$$

5. If the frequencies of first four numbers out of 1, 2, 4, 6, 8 are 2, 3, 3, 2 respectively, then the frequency of 8 if their A.M. is 5, is

(A) 4 (B) 5
(C) 6 (D) None of these

Solution: (C)

Here mean $A = 5$.

Let the frequency of 8 be x. Then by the formula

$$A = \frac{\sum xf}{\sum f}$$

$$5 = \frac{1.2 + 2.3 + 4.3 + 6.2 + 8.x}{2 + 3 + 3 + 2 + x} = \frac{32 + 8x}{10 + x}$$

or $18 = 3x$; $\therefore x = 6$.

6. The mean weight of 120 students in the second year class of a college is 56 kg. If the mean weights of the boys and that of the girls in the class are 60 kg and 50 kg respectively, then the number of boys and girls separately in the class are

(A) 72, 64 (B) 38, 64
(C) 72, 48 (D) None of these

Solution: (C)

We know that the combined mean

$$A = \frac{n_1 A_1 + n_2 A_2}{n_1 + n_2} \quad (1)$$

Here A_1 = mean weight of boys = 60 kg.

A_2 = mean weight of girls = 50 kg.

A = combined mean = 56 kg.

and $n_1 + n_2 = 120 \quad (2)$

So, from (1) and (2),

$$56 = \frac{n_1 \cdot 60 + n_2 \cdot 50}{120};$$

$\therefore$ $56 \times 120 = n_1 \cdot 60 + (120 - n_1)\, 50;$

$\therefore$ $120\,(56 - 50) = 10n_1;$

$\therefore$ $n_1 = 72,\ n_2 = 48.$

Thus, the number of boys = 72 and the number of girls = 48.

7. The mean of 10 numbers is 12.5; the mean of the first six is 15 and the last five is 10. The sixth number is

(A) 15 (B) 12
(C) 18 (D) None of these

Solution: (A)

Let the mean of the last four be A_2. Then by the formula for combined mean,

$$12.5 = \frac{6 \times 15 + 4 \times A_2}{6 + 4};$$

or $125 = 90 + 4\,A_2;$

$\therefore$ $A_2 = \dfrac{35}{4}$

Let the sixth number $= x$; then taking the sixth number as a collection, the combined mean of this collection and the collection of the last four is 10, by question.

$\therefore$ By definition of combined mean

$$10 = \frac{1 \times x + 4 \times \frac{35}{4}}{1 + 4};$$

$\therefore$ $50 = x + 35;\ \therefore\ x = 15.$

$\therefore$ Sixth number = 15.

8. In a family, there are 8 men, 7 women and 5 children whose mean ages separately are respectively 24, 20 and 6 years. The mean age of the family is

(A) 17.1 years (B) 18.1 years
(C) 19.1 years (D) None of these

Solution: (B)

Here we have three collections for which $A_1 = 24$, $n_1 = 8$; $A_2 = 20$, $n_2 = 7$ and $A_3 = 6$, $n_3 = 5$. Their combined mean is the required mean.

By the formula

$$A = \frac{n_1 A_1 + n_2 A_2 + n_3 A_3}{n_1 + n_2 + n_3}$$

$$\therefore \quad A = \frac{8 \times 24 + 7 \times 20 + 5 \times 6}{8 + 7 + 3}$$

$$= \frac{192 + 140 + 30}{20} = \frac{362}{20} = 18.1$$

$\therefore$ The mean age of the family = 18.1 years.

9. The mean of 100 items is 50 and their S.D. is 4. The sum of all the items and also the sum of the squares of the items is

(A) 5000, 251600 (B) 4000, 251600
(C) 5000, 261600 (D) None of these

Solution: (A)

Here $n = 100$, $A = 50$, $\sigma = 4$.

Now, $A = \dfrac{\sum x}{n}$; $\therefore \sum x = nA = 100 \times 50 = 5{,}000.$

Again, from the formula,

$$\sigma^2 + A^2 = \frac{\sum x^2}{n}, \text{ we get } \sum x^2 = n(\sigma^2 + A^2)$$

$$\therefore \sum x^2 = n(\sigma^2 + A^2) = 100(16 + 2500) = 2{,}51{,}600$$

10. If the mean of the set of number $x_1, x_2, \ldots\ldots, x_n$ is $\bar{x}$, then the mean of the numbers $x_i + 2i$, $1 \le i \le n$ is

(A) $\bar{x} + 2n$ (B) $\bar{x} + n + 1$
(C) $\bar{x} + 2$ (D) $\bar{x} + n$.

Solution: (B)

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\Rightarrow \quad \sum_{i=1}^{n} x_i = n\bar{x}$$

$$\therefore \frac{\sum_{i=1}^{n}(x_i + 2i)}{n} = \frac{\sum_{i=1}^{n} x_i + 2(1+2+\ldots.+n)}{n}$$

$$= \frac{n\bar{x} + 2\dfrac{n(n+1)}{2}}{n} = \bar{x} + (n+1).$$

11. The A.M. of n observations is M. If the sum of $n - 4$ observations is a, then the mean of remaining 4 observations is

(A) $\frac{nM - a}{4}$ (B) $\frac{nM - a}{2}$

(C) $\frac{nM - a}{4}$ (D) $nM + a$

Solution: (A)

Let the mean of the remaining 4 observations be $\bar{X}_1$.

$$\text{Then, } M = \frac{a + 4\bar{X}_1}{(n-4)+4} \Rightarrow \bar{X}_1 = \frac{nM - a}{4}.$$

12. The weighted mean of first n natural numbers whose weights are equal to the squares of corresponding numbers is

(A) $\frac{n+1}{2}$ (B) $\frac{3n(n+1)}{2(2n+1)}$

(C) $\frac{(n+1)(2n+1)}{6}$ (D) $\frac{n(n+1)}{2}$

Solution: (B)

$$\text{Weighted Mean} = \frac{1.1^2 + 2.2^2 + \ldots\ldots + n.n^2}{1^2 + 2^2 + \ldots\ldots. + n^2}$$

$$= \frac{\Sigma n^3}{\Sigma n^2} = \frac{\dfrac{n(n+1)}{2}\,\dfrac{n(n+1)}{2}}{\dfrac{n(n+1)(2n+1)}{6}}$$

$$= \frac{3n(n+1)}{2(2n+1)}.$$

13. Mean of 100 items is 49. It was discovered that three items which should have been 60, 70, 80 were wrongly read as 40, 20, 50 respectively. The correct mean is

(A) 48 (B) $82\frac{1}{2}$

(C) 50 (D) 80

Solution: (C)

Sum of 100 items = 49 × 100 = 4900
Sum of items added = 60 + 70 + 80 = 210
Sum of items replaced = 40 + 20 + 50 = 110
New Sum = 4900 – 110 + 210 = 5000

$$\therefore \text{New mean} = \frac{5000}{100} = 50$$

14. The mean of n items is $\bar{X}$. If the first item is increased by 1, second by 2 and so on, then the new mean is

(A) $\bar{X} + n$ (B) $\bar{X} + \frac{n}{2}$

(C) $\bar{X} + \frac{n+1}{2}$ (D) None of these

Solution: (C)

Let $x_1, x_2, \ldots x_n$ be n items.

Then $\bar{X} = \frac{1}{n}\Sigma x_i$.

Let $y_1 = x_1 + 1, y_2 = x_2 + 2$

$y_3 = x_3 + 3, \ldots.., y_n = x_n + n$

Then the mean of the new series is

$$\frac{1}{n}\Sigma y_i = \frac{1}{n}\Sigma(x_i + i) = \frac{1}{n}\Sigma x_i + \frac{1}{n}(1+2+3+\ldots..+n)$$

$$= \bar{X} + \frac{1}{n}\cdot\frac{n(n+1)}{2} = \bar{X} + \frac{n+1}{2}.$$

15. The number of observations in a group is 40. If the average of first 10 is 4.5 and that of the remaining 30 is 3.5, then the average of the whole group is

(A) $\frac{15}{4}$ (B) $\frac{1}{5}$

(C) 8 (D) 4

Solution: (A)

$$\frac{x_1 + x_2 + \ldots.. + x_{10}}{10} = 4.5$$

$$\Rightarrow x_1 + x_2 + \ldots.. + x_{10} = 45$$

$$\text{Also, } \frac{x_{11} + x_{12} + \ldots.. + x_{40}}{30} = 3.5$$

$$\Rightarrow x_{11} + x_{12} + \ldots.. + x_{40} = 105$$

$$\therefore x_1 + x_2 + \ldots.. + x_{40} = 150$$

$$\therefore \frac{x_1 + x_2 + \ldots.. + x_{40}}{40} = \frac{150}{40} = \frac{15}{4}$$

16. A person purchases one kg of tomatoes from each of the 4 places at the rate of 1 kg, 2 kg, 3 kg, 4 kg per rupee respectively. On the average he has purchased x kg of tomatoes per rupee, then the value of x is

(A) 2 (B) 2.5

(C) 1.92 (D) None of these

Solution: (C)

Since we are given rate per rupee, harmonic mean will give the correct answer

$$\text{H.M.} = \frac{4}{\frac{1}{1}+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}} = \frac{4\times 12}{12+6+4+3} = \frac{48}{25}$$

$= 1.92$ kg per rupee.

17. The A.M. of ${}^{2n}C_0, {}^{2n}C_2, {}^{2n}C_4, \ldots, {}^{2n}C_{2n}$ is

(A) $\dfrac{2^n}{(n+1)}$ (B) $\dfrac{2^{2n}}{(n+1)}$

(C) $\dfrac{2^{2n-1}}{(n+1)}$ (D) $\dfrac{2^{n-1}}{(n+1)}$

Solution: (C)

$$(1+x)^{2n} = {}^{2n}C_0 + {}^{2n}C_1 x + {}^{2n}C_2 x^2 + {}^{2n}C_3 x^3 +, \ldots, + {}^{2n}C_{2n} x^{2n}$$

Put $x = -1$

$$ {}^{2n}C_0 - {}^{2n}C_1 + {}^{2n}C_2 - {}^{2n}C_3 +, \ldots, + {}^{2n}C_{2n} = 0 \quad (1)$$

Now put $x = 1$

$$ {}^{2n}C_0 + {}^{2n}C_1 + {}^{2n}C_2 + {}^{2n}C_3 +, \ldots, + {}^{2n}C_{2n} = 2^{2n} \quad (2)$$

Adding (1) and (2), we get

$$ {}^{2n}C_0 + {}^{2n}C_2 +, \ldots, + {}^{2n}C_{2n} = 2^{2n-1}$$

$$\text{A.M. of } {}^{2n}C_0, {}^{2n}C_2 +, \ldots, + {}^{2n}C_{2n} = \frac{2^{2n-1}}{(n+1)}$$

18. The A.M. of ${}^{2n+1}C_0, {}^{2n+1}C_1, {}^{2n+1}C_2, \ldots, {}^{2n+1}C_n$ is

(A) $\dfrac{2^n}{n}$ (B) $\dfrac{2^n}{n+1}$

(C) $\dfrac{2^{2n}}{n}$ (D) $\dfrac{2^{2n}}{(n+1)}$

Solution: (D)

$$ {}^{2n+1}C_0 + {}^{2n+1}C_1 + {}^{2n+1}C_2 +, \ldots, + {}^{2n+1}C_{2n} + {}^{2n+1}C_{2n+1} = 2^{2n+1}$$

Now $\quad {}^{2n+1}C_0 = {}^{2n+1}C_{2n+1}, {}^{2n+1}C_1 = {}^{2n+1}C_{2n}, \ldots$

$\quad {}^{2n+1}C_r = {}^{2n+1}C_{2n-r+1}$

So sum of first $(n+1)$ terms = Sun of last $(n+1)$ terms

or $\quad {}^{2n+1}C_0 + {}^{2n+1}C_1 + {}^{2n+1}C_2 +, \ldots, + {}^{2n+1}C_n = 2^{2n}$

or $$\frac{{}^{2n+1}C_0 + {}^{2n+1}C_1 + {}^{2n+1}C_2 +, \ldots, + {}^{2n+1}C_n}{n+1} = \frac{2^{2n}}{(n+1)}$$

19. The mean of the values 0, 1, 2, 3, ... , n with the corresponding weights ${}^nC_0, {}^nC_1, \ldots, {}^nC_n$ respectively is

(A) $\dfrac{2^n}{(n+1)}$ (B) $\dfrac{2^{n+1}}{n(n+1)}$

(C) $\dfrac{n+1}{2}$ (D) $\dfrac{n}{2}$

Solution: (D)

$$\frac{{}^nC_1 + 2\cdot{}^nC_2 + 3\,{}^nC_3 + \ldots + n\,{}^nC_n}{{}^nC_0 + {}^nC_1 + {}^nC_2 + \ldots + {}^nC_n} = \frac{n\cdot 2^{n-1}}{2^n} = \frac{n}{2}$$

20. In a factory, workers work in three shifts say shift 1, shift 2 and shift 3 and they get wages in the ratio 4 : 5 : 6 depending on the shift 1, 2 and 3 respectively. Number of workers in the shifts are in the ratio 3 : 2 : 1. If total number of workers working is 1500 and wages per worker in Ist shift is Rs 400. Then mean wage of a worker is

(A) Rs. 467 (B) Rs. 500

(C) Rs. 600 (D) Rs. 400

Solution: (A)

Workers in Ist shift = 750

Wages in Ist shift = Rs. 400

Workers is IInd shift = 500

Wages in IInd shift = Rs. 500

Workers in IIIrd shift = 250

Wages in IIIrd shift = Rs. 600

$$\therefore \text{ Mean} = \frac{750\times 400 + 500\times 500 + 250\times 600}{1500}$$

Rs. 467 per worker

21. If a variable takes values 1, 2, 3, ... , n with frequencies $1^2, 2^2, \ldots, n^2$, then the mean is

(A) Σn (B) $\dfrac{\Sigma n^3}{\Sigma n^2}$

(C) $\dfrac{\Sigma n^3}{\Sigma n}$ (D) None of these

Solution: (B)

$$\frac{1.1^2 + 2.2^2 + \ldots + n.n^2}{1^2 + 2^2 + \ldots + n^2} = \frac{\Sigma n^3}{\Sigma n^2}$$

22. Mean of n items is x. If these n items are increased by $1^2, 2^2, 3^2, \ldots, n^2$ successively, then mean gets increased by

(A) $\dfrac{(n+1)(2n+1)}{6}$ (B) $\dfrac{n(n+1)(2n+1)}{6}$

(C) $\dfrac{n^2}{2}$ (D) Remains same

Solution: (A)

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + \ldots + x_n}{n}$$

$$\bar{x}' = \frac{x_1 + 1^2 + x_2 + 2^2 + x_3 + 3^2 + \ldots + x_n + n^2}{n}$$

$$\Rightarrow \quad \bar{x}' = \frac{x_1 + x_2 + \ldots + x_n}{n} + \frac{1^2 + 2^2 + \ldots + n^2}{n}$$

$$\Rightarrow \quad \bar{x}' = \bar{x} + \frac{(n+1)(2n+1)}{6}$$

MEDIAN

Median is the middle most or the central value of the variate in a set of observations, when the observations are arranged either in ascending or in descending order of their magnitudes. It divides the arranged series in two equal parts.

Calculation of Median

Median of an Individual Series

Let n be the number of observations.

1. Arrange the data in ascending or descending order.
2. (A) If n is odd, then

   Median = value of the $\frac{1}{2}$ $(n + 1)$th observation

   (B) If n is even, then

   Median = mean of the $\left(\frac{n}{2}\right)$th and $\left(\frac{n}{2}+1\right)$th observation.

Median of a Discrete Series

1. Arrange the values of the variate in ascending or descending order.
2. Prepare a cumulative frequency table.
3. (A) If n is odd, then

   Median = size of the $\left(\frac{n+1}{2}\right)$ th term.

   (B) If n is even, then

   Median = size of the $\left(\frac{\left(\frac{n}{2}\right)+\left(\frac{n}{2}+1\right)}{2}\right)$ th term.

Median of a Continuous Series

1. Prepare the cumulative frequency table.
2. Find the median class, i.e., the class in which the $\left(\frac{n}{2}\right)$ th observation lies.
3. The median value is given by the formula

$$\text{Median} = l + \left(\frac{\left(\frac{n}{2}\right) - c_f}{f}\right) \times h, \text{ where}$$

l = lower limit of the median class
n = total frequency
f = frequency of the median class
h = width of the median class
c_f = cumulative frequency of the class preceding the median class

Some Points About Median

- It is an appropriate average in dealing with qualitative data, like intelligence, wealth etc.
- The sum of the deviations of the items from median, ignoring algebraic signs, is less than the sum from any other point.

QUARTILES, DECILES AND PERCENTILES

Quartile

Just as the median divides a set of observations (when arranged in ascending or descending order of magnitudes), into two equal parts, similarly *Quartile* divides the observations into four equal parts. The value of the item midway, between the first item and the median is known as *first or lower quartile* and is denoted be Q_1. The value of the item midway between the last item and the median is known *Third or Upper Quartile* and is denoted Q_3. The median is known as the *Second Quartile* and is denoted by Q_2. The methods for finding the values of Q_1 and Q_3 are similar to that of the median. In the case of ungrouped data, when arranged in ascending or descending order of magnitudes Q_1, Q_3 can be obtained as follows:

$$Q_1 = \frac{n+1}{4} \text{ th item, } Q_3 = \frac{3(n+1)}{4} \text{ th item.}$$

For a frequency distribution, Q_1 and Q_3 are given by,

$$Q_1 = l + \frac{[(n/4) - C_f]}{f} \times h,$$

$$Q_3 = l + \frac{[(3n/4) - C_f]}{f} \times h,$$

where l = lower limit of the class in which a particular quartile lies,
f = frequency of the class-interval in which a particular quartile lies,

i = class-interval of the class in which a particular quartile lies,

c_f = cumulatively frequency of the class preceding the class in which the particular quartile lies.

In general, $Q_i = l + \dfrac{[(nh/4) - c_f]}{f} \times h$, $i = 1, 2, 3, 4$

Decile

The value of the variable which divides the series, when arranged in ascending or descending order, into 10 equal parts is called decile. There are 9 deciles denoted by D_1, D_2 ... D_9. When the series is ungrouped the deciles are calculated as follows:

$$D_i = \frac{n \times h}{10},\ i = 1, 2, ..., 9$$

When the data is classified or grouped,

$$D_i = l + \frac{[(nh/10) - c_f]}{f} \times h$$

where symbols have their usual meaning.

Percentile

The value of the variable which divides the series, when arranged in ascending or descending order, into 100 equal parts is called *percentile*. There are 99 percentiles denoted by P_1, P_2, P_3, P_4, ..., P_{99} respectively. When the series is ungrouped the percentiles are calculated by the following formula:

$$P_i = \frac{n \times h}{100},\ h = 1, 2, ..., 99.$$

When the data is classified or grouped, the percentiles are calculated by the formula

$$P_i = l + \frac{[(nh/100) - c_f]}{f} \times h,\ i = 1, 2, ..., 99.$$

where symbols have their usual meanings.

SOLVED EXAMPLES

23. If a variable takes the discrete values $\alpha + 4$, $\alpha - \frac{7}{2}$, $\alpha \frac{5}{2} - \leftarrow$, $\alpha - 3$, $\alpha - 2$, $\alpha + \frac{1}{2}$, $\alpha - \frac{1}{2}$, $\alpha + 5$ $(\alpha > 0)$, then the median is

(A) $\alpha - \frac{5}{4}$ (B) $\alpha - \frac{1}{2}$

(C) $\alpha - 2$ (D) $\alpha + \frac{5}{4}$

Solution: (A)

Arrange the data as $\alpha - \frac{7}{2}$, $\alpha - 3$, $\alpha - \frac{5}{2}$, $\alpha - 2$, $\alpha - \frac{1}{2}$, $\alpha + \frac{1}{2}$, $\alpha + 4$, $\alpha + 5$.

$$\text{Median} = \frac{\alpha - 2 + \alpha - \frac{1}{2}}{2} = \frac{2\alpha - \frac{5}{2}}{2} = \alpha - \frac{5}{4}.$$

24. Median of ${}^{2n}C_0$, ${}^{2n}C_1$, ${}^{2n}C_2$, ${}^{2n}C_3$,..., ${}^{2n}C_n$ (when n is even) is

(A) ${}^{2n}C_{n/2}$ (B) ${}^{2n}C_{(n+1)/2}$

(C) ${}^{2n}C_{(n-1)/2}$ (D) None of these

Solution: (A)

${}^{2n}C_0$, ${}^{2n}C_1$, ${}^{2n}C_2$, ... , ${}^{2n}C_n$ is odd number of binomial coefficients (when n is even) and middle binomial coefficient is ${}^{2n}C_{n/2}$.

25. Median of ${}^{2n}C_0$, ${}^{2n}C_1$, ${}^{2n}C_2$, ${}^{2n}C_n$ (when n is odd) is

(A) $\frac{1}{2}\left({}^{2n}C_{(n-1)/2} + {}^{2n}C_{(n+1)/2}\right)$

(B) ${}^{2n}C_{n/2}$

(C) ${}^{2n}C_n$

(D) None of these

Solution: (A)

${}^{2n}C_0$, ${}^{2n}C_1$, ${}^{2n}C_2$, ... , ${}^{2n}C_n$ is even number of binomial coefficients (when n is odd), and then middle terms are ${}^{2n}C_{n-1/2}$ and ${}^{2n}C_{n+1/2}$. So median is $\dfrac{{}^{2n}C_{n-1/2} + {}^{2n}C_{n+1/2}}{2}$.

26. If a variable takes the discrete values $\alpha + 4$, $\alpha - \frac{5}{2}$ $\alpha - \frac{7}{2}$, $\alpha - 3$, $\alpha - 2$, $\alpha + \frac{1}{2}$, $\alpha - \frac{1}{2}$, $\alpha + 5$ $(\alpha > 0)$, then the median is

(A) $\alpha - \frac{1}{2}$ (B) $\alpha - \frac{5}{4}$

(C) $\alpha - 2$ (D) $\alpha + \frac{5}{4}$

Solution: (B)

$$\text{Median} = \frac{1}{2}\left(\alpha - 2 + \alpha - \frac{1}{2}\right)$$

$$= \alpha - \frac{5}{4}$$

MODE

Mode is that value in a series which occurs most frequently.

In a frequency distribution, mode is that variate which has the maximum frequency.

Computation of Mode

Mode of Individual Series

In the case of individual series, the value which is repeated maximum number of times is the mode of the series.

Mode of Discrete Series

In the case of discrete frequency distribution, mode is the value of the variate corresponding to the maximum frequency.

Mode of Continuous Series

1. Find the modal class, i.e., the class which has maximum frequency. The modal class can be determined either by inspection or with the help of grouping table.
2. The mode is given by the formula

$$\text{Mode} = l + \frac{f_m - f_{m-1}}{2f_m - f_{m-1} - f_{m+1}} \times h,$$

where l = the lower limit of the modal class
h = the width of the modal class
f_{m-1} = the frequency of the class preceding modal class
f_m = the frequency of the modal class
f_{m+1} = the frequency of the class succeeding modal class

In case, the modal value lies in a class other than the one containing maximum frequency, we take the help of the following formula;

$$\text{Mode} = l = \frac{f_{m+1}}{f_{m-1} + f_{m+1}} \times h,$$

where symbols have usual meaning.

Some Points about Mode

- It is not based on all items of the series
- It is not necessary that a distribution has unique mode.
- As compared to other averages mode is affected to a large extent by fluctuations of sampling.
- It is not suitable in a case where the relative importance of items have to be considered.

SYMMETRIC DISTRIBUTION

A distribution is a symmetric distribution if the values of mean, mode and median coincide. In a symmetric distribution frequencies are symmetrically distributed on both sides of the centre point of the frequency curve.

Measures of Dispersion

The degree to which numerical values in the set of values tend to spread about an average value is called the dispersion or variation.

The commonly used measures of dispersion are:

1. Range
2. Quartile Deviation or Semi-interquartile range
3. Mean Deviation
4. Standard Deviation

Range

It is the difference between the greatest and the smallest observations of the distribution.

If L is the largest and S is the smallest observation in a distribution, then its Range = $L - S$. Also,

$$\text{Coefficient of range} = \frac{L - S}{L + S}.$$

Quartile deviation

Quartile deviation or semi-interquartile range is given by

$$\text{Q.D.} = \frac{1}{2}(Q_3 - Q_1)$$

$$\text{Coefficient of Q.D.} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Mean deviation

For a frequency distribution, the mean deviation from an average (median, or arithmetic mean) is given by,

$$\text{M.D.} = \frac{\sum_{i=1}^{n} f_i \left| x_i - \bar{x} \right|}{\sum_{i=1}^{n} f_i}$$

$$\text{Coefficient of M.D.} = \frac{\text{Mean deviation}}{\text{Corresponding average}}$$

Standard deviation

The *standard deviation* of a statistical data is defined as the positive square root of the squared deviations of observations from the A.M. of the series under consideration.

1. Standard deviation (also denoted by σ) for ungrouped set of observations is given by

$$\sigma = \frac{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}}{n}$$

2. Standard deviation for frequency distribution is given by,

$$\text{S.D.} = \frac{\sqrt{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}}{N}$$

where f_i is the frequency of x_i $(1 \le i \le n)$.

When the values of the variable are given in the form of classes, then their respective mid-points are taken as the values of the variable.

Standard Deviation of n Natural Numbers

$$\sigma = \left(\frac{1}{12}(n^2 - 1)\right)^{1/2}$$

Standard deviation shows the limits of variability by which the individual observation in a distribution will vary from the mean. For a symmetrical distribution with mean $\bar{x}$, the following area relationship holds good:

$\bar{x} \pm \sigma$ covers 68.27 % observations.

$\bar{x} \pm 2\sigma$ covers 95.45 % observations.

$\bar{x} \pm 3\sigma$ covers 99.73 % observations.

These limits are illustrated by the following curve known as *Normal Curve*.

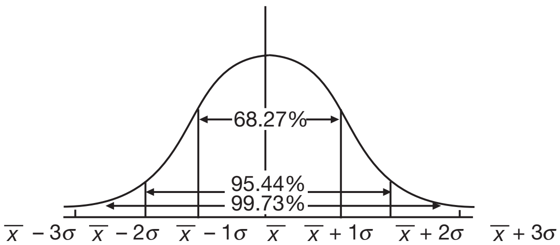

Fig. 23.1

Empirical relationships

If the data is moderately non-symmetrical, then the following empirical relationships hold:

Mean deviation = $\frac{4}{5}\sigma$

Semi-Inter-quartile range = $\frac{2}{3}\sigma$

Probable error of standard deviation = $\frac{2}{3}\sigma$

= Semi-inter-quartile range.

Quartile deviation = $\frac{5}{6}$ M.D.

From these relationships, we have

4 S.D. = 5 M.D. = 6 Q.D

Coefficient of S.D. (C.V.)

For comparing two or more series for variability, the relative measure, called coefficient of variation (C.V.) is used. This measure is defined as

$$\text{C.V.} = \frac{\sigma}{\bar{x}} \times 100$$

The coefficient of variation is also represented as percentage.

The square of *S.D.* is called the variance of the distribution and is denoted by σ^2.

Computation of Standard Deviation

Direct Method

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

Short-cut Method

$$\sigma = \sqrt{\frac{\sum d^2}{n} - \left(\frac{\sum d}{n}\right)^2}, \text{ for ungrouped data}$$

where A is assumed mean and $d = x - A$.

$$\sigma = \sqrt{\frac{\sum fd^2}{N} - \left(\frac{\sum fd}{N}\right)^2}, \text{ for grouped data}$$

where $N = \sum f$.

Step-deviation Method

$$\sigma = h\sqrt{\frac{\sum fd'^2}{N} - \left(\frac{\sum fd'}{N}\right)^2}\ ;\ d' = \frac{x - A}{h}$$

Combined Standard Deviation

Let A_1 and A_2 be two series having n_1 and n_2 observations respectively. Let their A.M. be $\bar{x}_1$ and $\bar{x}_2$, and standard deviations be σ_1 and σ_2. Then the combined standard deviation σ or σ_{12} of A_1 and A_2 is given by

$$\sigma_{12} \text{ or } \sigma = \sqrt{\frac{n_1\sigma_1^2 + n_2\sigma_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}}$$

$$= \sqrt{\frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}}$$

where $d_1 = (\bar{x}_1 - \bar{x}_{12}),\ d_2 = (\bar{x}_2 - \bar{x}_{12})$,

and $\bar{x}_{12} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$ is the combined mean.

IMPORTANT POINTS

- Quartile deviation is less affected by extreme values of the series.
- Mean deviation is based on all the items of series. It is therefore more representative than the range or quartile deviation.
- Mean deviation from the median is less than that measured from any other mean.
- Standard deviation ≤ Range i.e., variance ≤ (Range)^2.
- S.D. of first n natural numbers is $\sqrt{\dfrac{n^2-1}{12}}$.

SOLVED EXAMPLES

27. The coefficient of variation of two series are 58% and 69%. If their standard deviations are 21.2 and 15.6, then their A.Ms are

(A) 36.6, 22.6 (B) 34.8, 22.6
(C) 36.6, 24.4 (D) None of these

Solution: (A)

We know that

$$\text{C.V.} = \frac{\sigma \times 100}{\overline{x}}$$

or $$\overline{x} = \frac{\sigma}{\text{C.V.}} \times 100$$

$$\therefore \text{ Mean of first series} = \frac{21.2 \times 100}{58} = 36.6$$

$$\text{Mean of second series} = \frac{15.6 \times 100}{69} = 22.6$$

28. Mean deviation of the series a, $a + d$, $a + 2d$, $a + 2nd$ from its mean is

(A) $\dfrac{(n+1)\,d}{(2n+1)}$ (B) $\dfrac{nd}{2n+1}$

(C) $\dfrac{n\,(n+1)\,d}{(2n+1)}$ (D) $\dfrac{(2n+1)\,d}{n\,(n+1)}$

Solution: (C)

$$\overline{x} = \frac{\dfrac{2n+1}{2}.(a+a+2nd)}{(2n+1)} = a + nd$$

$$\Sigma\left|x - \overline{x}\right| = 2d\,(1 + 2 + \ldots\ldots + n) = n\,(n+1)\,d$$

$$\therefore \quad \text{M.D.} = \frac{n(n+1)d}{2n+1}$$

29. The mean deviation from the mean for the set of observations – 1, 0, 4 is

(A) less than 3 (B) less than 4
(C) greater than 2.5 (D) greater than 4.9

Solution: (A and B)

$$\overline{x} = \frac{-1+0+4}{3} = 1.$$

$$\therefore \text{ Mean Deviation} = \frac{1}{3}\,(|-1-1| + |0-1| + |4-1|)$$
$$= 2$$

30. If the S.D of a set of observations is 4 and if each observation is divided by 4, the S.D of the new set of observations will be

(A) 4 (B) 3
(C) 2 (D) 1

Solution: (D)

We know that if $y = x/h$ when $\sigma y = \sigma_x/|h|$.
∴ The S.D. of new set of observations will be 4/4 = 1.

31. A sample of 35 observations has the mean 80 and S.D. as 4. A second sample of 65 observations from the same population has mean 70 and S.D. 3. The S.D. of the combined sample is

(A) 5.85 (B) 5.58
(C) 3.42 (D) None of these

Solution: (A)

Here $n_1 = 35,\ \overline{x}_1 = 80,\ \sigma_1 = 4,$

$n_2 = 65,\ \overline{x}_2 = 70,\ \sigma_2 = 3.$

$$\therefore \quad \overline{x}_{12} = \frac{35 \times 80 + 65 \times 70}{35 + 65} = 73.5.$$

$$\sigma_{12} = \sqrt{\left[\frac{35(16 + 42 \times 25) + 65(9 + 12 \times 25)}{100}\right]}$$

$$= \sqrt{34.21} = 5.85$$

32. If μ is the mean of a distribution, then $\sum f_i\,(y_i - \mu)$ is equal to

(A) M.D. (B) S.D.
(C) 0 (D) None of these

Solution: (C)

We have, $$\mu = \frac{\Sigma f_i y_i}{\Sigma f_i}$$

$$\Rightarrow \quad \Sigma f_i y_i - \Sigma f_i\,\mu = 0$$

$$\Rightarrow \quad \Sigma f_i\,(y_i - \mu) = 0$$

33. The means of five observations is 4 and their variance is 5.2. If three of these observations are 1, 2, and 6, then the other two are

(A) 2 and 9 (B) 3 and 8
(C) 4 and 7 (D) 5 and 6

Solution: (C)

$$\bar{x}_1 = 4, N = 5$$

and $$\frac{\Sigma(x-\bar{x})^2}{N} = 5.2$$

$\Rightarrow$ $$\Sigma(x-\bar{x})^2 = (5.2)\,5$$

$\therefore$ $$\Sigma(x-\bar{x})^2 = 26$$

$\therefore$ $$(1-4)^2 + (2-4)^2 + (6-4)^2 + (\alpha-4)^2 + (\beta-4)^2 = 26$$

where α, β are the other two observatios.

$\therefore$ $9 + 4 + 4 + (\alpha-4)^2 + (\beta-4)^2 = 26$

$\therefore$ $$(\alpha-4)^2 + (\beta-4)^4 = 9$$

Also, $$\frac{1+2+6+\alpha+\beta}{5} = 4$$

$\therefore$ $$\alpha + \beta = 20 - 9 = 11$$

Clearly 4, 7 only satisfy the above equation in α, β.
Hence reqd. numbers are 4, 7.

34. If 25 % of the items are less than 20 and 25 % are more then 40, the quartile deviation is

(A) 20 (B) 30
(C) 40 (D) 10

Solution: (D)

$$\text{Q.D.} = \frac{40-20}{2} = 10.$$

35. The sum of squares of deviations for 10 observations taken from mean 50 is 250. The coefficient of variation is

(A) 10 % (B) 40 %
(C) 50 % (D) None of these

Solution: (A)

Co-efficient of variation

$$= \frac{\sigma}{\bar{x}} \times 100 = \frac{\sigma}{50} \times 100 \quad (\because \bar{x} = 50) = 2\sigma$$

Also, $$\sigma = \sqrt{\frac{\Sigma(x_i-50)^2}{n}} = \sqrt{\frac{250}{10}} = \sqrt{25} = 5$$

$\therefore$ Co.efficient of variation $= 2 \times 5 = 10\%$.

EXERCISES

Single Option Correct Type

1. The average of n numbers $x_1, x_2, x_3, \ldots, x_n$ is M. If x_n is replaced by x', then new average is

(A) $M - x_n + x'$ (B) $\dfrac{nM - x_n + x'}{n}$

(C) $\dfrac{(n-1)M + x'}{m}$ (D) $\dfrac{M - x_n + x'}{n}$

2. The standard deviation of 25 numbers is 40. If each of the numbers is increased by 5, then the new standard deviation will be

(A) 40 (B) 45
(C) $40 + \dfrac{21}{25}$ (D) None of these

3. If M.D. is 12, the value of S.D. will be

(A) 15 (B) 12
(C) 24 (D) None of these

4. The mean weight of 9 items is 51. If one more item is added to the series the mean becomes 16. The value of the 10th item is

(A) 35 (B) 30
(C) 25 (D) 20

5. The mean and S.D. of the marks of 200 candidates were found to be 40 and 15 respectively. Later, it was discovered that a score of 40 was wrongly read as 50. The correct mean and S.D. respectively are

(A) 14.98, 39.95 (B) 39.95, 14.98
(C) 39. 95, 224.5 (D) None of these

6. If Q.D. is 16, the most likely value of S.D. will be

(A) 24 (B) 42
(C) 10 (D) None of these

7. If a variable x takes values 0, 1, 2, ..., n with frequencies proportional to the binomial coefficients ${}^nC_0$, ${}^nC_1$, ${}^nC_2$,..., ${}^nC_n$, then the Var (x) is

(A) $\frac{n^2-1}{12}$ (B) $\frac{n}{2}$

(C) $\frac{n}{4}$ (D) None of these

8. The sum of squares of deviations for 10 observations taken from mean 50 is 250. The coefficient of variation is

(A) 50% (B) 10%

(C) 40% (D) None of these

9. If the standard deviation of n observations x_1, x_2,..., x_n is 4 and another set of n observations y_1, y_2,..., y_n is 3. The standard deviation of n observations $x_1 - y_1$, $x_2 - y_2$,..., $x_n - y_n$ is

(A) 1 (B) $\frac{2}{\sqrt{3}}$

(C) 5 (D) Data insufficient

10. Let r be the range and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$ be the S.D. of a set of observations x_1, x_2,..., x_n, then

(A) $S \le r\sqrt{\frac{n}{n-1}}$ (B) $S = r\sqrt{\frac{n}{n-1}}$

(C) $S \ge r\sqrt{\frac{n}{n-1}}$ (D) None of these

11. The A.M. of n numbers of a series is $\bar{x}$. If the sum of the first (n – 1) term is k, them the nth number is

(A) $\bar{x} - k$ (B) $n\bar{x} - k$

(B) $\bar{x} - nk$ (D) $n\bar{x} - nk$

12. If a variable takes values 0, 1, 2,..., n with frequencies q^n, $\frac{n}{1}q^{n-1}p$, $\frac{n(n-1)}{1.2}q^{n-2}p^2$,..., p^n, where $p + q = 1$, then the mean is

(A) np (B) nq

(C) $n(p + q)$ (D) None of these

13. The S.D. of a variate x is σ. The S.D. of the variate $\frac{ax+b}{c}$ where a, b, c are constants, is

(A) $\left(\frac{a}{c}\right)\sigma$ (B) $\left|\frac{a}{c}\right|\sigma$

(C) $\left(\frac{a^2}{c^2}\right)\sigma$ (D) None of these

14. Consider any set of observations x_1, x_2, x_3,..., x_{101}; it being given that $x_1 < x_2 < x_3 < ... < x_{100} < x_{101}$; then the mean deviation of this set of observations about a point k is minimum when k equals

(A) x_1 (B) x_{51}

(C) $\frac{x_1 + x_2 + ... + x_{101}}{101}$ (D) x_{50}

15. The mean of the numbers

$\frac{{}^{50}C_0}{1}, \frac{{}^{50}C_2}{3}, \frac{{}^{50}C_4}{5} ..., \frac{{}^{50}C_{50}}{51}$ equals

(A) $\frac{2^{50}}{51}$ (B) $\frac{2^{49}}{51}$

(C) $\frac{2^{49}}{39 \times 17}$ (D) None of these

16. The standard deviation of a distribution is 30 and each item is raised by 3, then new S.D. is

(A) 32 (B) 28

(C) 27 (D) None of these

17. For three numbers a, b, c product of the average of the numbers a^2, b^2, c^2 and $\frac{1}{a^2}, \frac{1}{b^2}, \frac{1}{c^2}$ cannot be less than

(A) 1 (B) 3

(C) 9 (D) None of these

18. The variance of α, β and γ is 9, then variance of 5α, 5β and 5γ is

(A) 45 (B) 9/5 (C) 5/9 (D) 225

19. Mean of the numbers 1, 2, 3,..., n with respective weights $1^2 + 1$, $2^2 + 2$, $3^2 + 3$,..., $n^2 + n$ is

(A) $\frac{3n(n+1)}{2(2n+1)}$ (B) $\frac{2n+1}{3}$

(C) $\frac{3n+1}{4}$ (D) $\frac{3n+1}{2}$

20. The G.M. of the number 3, 3^2, 3^3,..., 3^{3n} is

(A) $3^{\frac{n}{2}}$ (B) $3^{\frac{3n}{2}}$

(C) $3^{\frac{3n+1}{2}}$ (D) $3^{\frac{n+1}{2}}$

21. The reciprocal of the weighted mean of first n natural numbers whose weights are equal to the squares of the corresponding numbers is

(A) $\frac{2(2n+1)}{3n(n+1)}$ (B) $\frac{3n(n+1)}{n(2n+1)}$

(C) $\frac{3n(n+1)}{2n+1}$ (D) None of these

22. The A.M. of a set of 50 numbers is 38. If two numbers of the set, namely 55 and 45 are discarded, the A.M. of the remaining set of numbers is

(A) 38.5 (B) 37.5
(C) 36.5 (D) 36

23. The mean weight per student in a group of seven students is 55 kg If the individual weights of 6 students are 52, 58, 55, 53, 56 and 54; then weight of the seventh student is

(A) 55 kg (B) 60 kg
(C) 57 kg (D) 50 kg

24. If the mean of a set of observations $x_1, x_2, x_3,..., x_n$ is $\bar{x}$, then mean of observations $x_i + 3i \ \forall \ i = 1, 2, 3,... n$ equals

(A) $\bar{x} + 3(n+1)$ (B) $\bar{x} + \frac{3(n+1)}{2}$
(C) $\bar{x} + \frac{n+1}{2n}$ (D) None of these

25. The weighted mean of the square of 1st n natural numbers whose weights are corresponding numbers, equals

(A) $\frac{(n+1)(2n+1)}{2}$ (B) $\frac{n(n+1)}{2}$
(C) $\frac{n+1}{2}$ (D) None of these

26. If the variate of a distribution takes the values 1, 2, 3, ...n with frequencies $n, n-1, n-2, ...$ 3, 2, 1, then mean value of the distribution is

(A) $\frac{n(n+2)}{3}$ (B) $\frac{n(n+1)(n+2)}{6}$
(C) $\frac{n+2}{3}$ (D) $\frac{(n+1))(n+2)}{6}$

27. The means of five observations is 4 and their variance is 5.2. If three of these observations are 1, 2 and 6, then the other two are

(A) 2 and 9 (B) 3 and 8
(C) 4 and 7 (D) 5 and 6

28. If the variate takes the values 0, 2, 4, 8, ... 2^n with frequencies ${}^nC_0, {}^nC_1, {}^nC_n$ and if the mean is $\frac{91 \times 8}{2^n}$, then n equals

(A) 4 (B) 6
(C) 5 (D) None of these

29. The mean of n items is $\bar{x}$. If each item is successively increased by $3, 3^2, 3^3,... 3^n$, then new mean equals

(A) $\bar{x} + \frac{3^{n+1}}{n}$ (B) $\bar{x} + 3\frac{(3^n - 1)}{2n}$
(C) $\bar{x} + \frac{3^n}{n}$ (D) $\bar{x} + 3\frac{(3^n - 1)}{2n}$

30. A sequence of odd positive integers is written as

1
3 5 7 9
11, 13, 15, 17, 19, 21, 23, 25, 27

The mean of the n^{th} row is

(A) $\frac{n^3(2n^2+1)}{3}$ (B) $\frac{n^3(4n^2+2)}{6}$
(C) $\frac{n(n-1)(2n-1)}{6}$ (D) $\frac{n(2n^2+1)}{3}$

31. The arithmetic mean of a set of observation is . If each observation is divided by α and then is increased by 10, the means of the new series is

(A) $\frac{\bar{x}}{\alpha}$ (B) $\frac{\bar{x}+10}{\alpha}$
(C) $\frac{\bar{x}+10\alpha}{\alpha}$ (D) $\alpha\bar{x}+10$

32. The average salary of male employees in a firm was Rs. 520 and that of females was Rs. 420. The mean salary of all the employees was Rs. 500. The percentage of male employees is

(A) 80 (B) 60 (B) 40 (D) 20

33. The average weight of students in a class of 35 students is 40 kg. If the weight of the teacher be included, the average rises by $\frac{1}{2}$ kg; the weight of the teacher is

(A) 40.5 kg (B) 50 kg
(C) 41 kg (D) 58 kg

34. An automobile driver travels from plane to a hill station 120 km distant at an average speed of 30 km per hour. He then makes the return trip at an average speed of 25 km per hour. He covers another 120 km distance on plane at an average speed of 50 km per hour. His average speed over the entire distance of 360 km will be

(A) $\frac{30+25+50}{3}$ km/h (B) $(30 \cdot 25 \cdot 50)^{\frac{1}{3}}$
(C) $\frac{3}{\frac{1}{30}+\frac{1}{25}+\frac{1}{50}}$ km/h (D) None of these

35. If the mean deviation about the median of the numbers $a, 2a, ..., 50a$ is 50, then $|a|$ equals

(A) 5 (B) 2 (C) 3 (D) 4

36. Let $x_1, x_2, ..., x_n$ be n observations, and let $\bar{x}$ be their arithmetic mean and σ^2 be the variance

Statement-1: Variance of $2x_1, 2x_2, ..., 2x_n$ is $4\sigma^2$.

Statement-2: Arithmetic mean $2x_1, 2x_2, ..., 2x_n$ is $4\bar{x}$.

(A) Statement-1 is false, Statement-2 is true.
(B) Statement-1 is true, statement-2 is true; statement-2 is a correct explanation for Statement-1.
(C) Statement-1 is true, statement-2 is true; statement-2 is **not** a correct explanation for Statement-1.
(D) Statement-1 is true, statement-2 is false.

37. If the mean of a set of observations $x_1, x_2, ..., x_{10}$ is 20 then the mean of $x_1 + 4, x_2 + 8, x_3 + 12, ..., x_{10} + 40$ is

(A) 34 (B) 42 (C) 38 (D) 40

38. The mean of the numbers a, b, 8, 5, 10 is 6 and the variance is 6.80. Then which one of the following gives possible values of a and b?

(A) $a = 0, b = 7$ (B) $a = 5, b = 2$
(C) $a = 1, b = 6$ (D) $a = 3, b = 4$

39. If the mean deviation of number $1, 1 + d, 1 + 2d, ..., 1 + 100d$ from their mean is 255, then the d is equal to

(A) 10.0 (B) 20.0
(C) 10.1 (D) 20.2

40. **Statement-1:** The variance of first n even natural numbers is $\dfrac{n^2 - 1}{4}$

Statement-2: The sum of first n natural numbers is $\dfrac{n(n+1)}{2}$ and the sum of squares of first n natural numbers is $\dfrac{n(n+1)(2n+1)}{6}$

(A) Statement-1 is true, Statement-2 is true; Statement-2 is a correct explanation for Statement-1.
(B) Statement-1 is true, Statement-2 is true; Statement-2 is **not** a correct explanation for Statement-1.
(C) Statement-1 is true, Statement-2 is false.
(D) Statement-1 is false, Statement-2 is true.

Previous Year's Questions

41. The median of a set of 9 distinct observations is 20.5. If each of the largest 4 observations of the set is increased by 2, then the median of the new set **[2003]**

(A) is increased by 2
(B) is decreased by 2
(C) is two times the original median
(D) remains the same as that of the original set

42. Let two numbers have arithmetic mean 9 and geometric mean 4. Then these numbers are the roots of the quadratic equation **[2004]**

(A) $x^2 + 18x + 16 = 0$
(B) $x^2 - 18x - 16 = 0$
(C) $x^2 + 18x - 16 = 0$
(D) $x^2 - 18x + 16 = 0$

43. Consider the following statements **[2004]**

(A) Mode can be computed from histogram
(B) Median is not independent of change of scale
(C) Variance is independent of change of origin and scale.

Which of these is/are correct?

(A) only (A) (B) only (B)
(C) only (A) and (B) (D) (A), (B) and (C)

44. In a series of $2n$ observations, half of them equal a and remaining half equal $-a$. If the standard deviation of the observations is 2, then $|a|$ equals **[2004]**

(A) $\dfrac{1}{n}$ (B) $\sqrt{2}$
(C) 2 (D) $\dfrac{\sqrt{2}}{n}$

45. If in a frequency distribution, the mean and median are 21 and 22 respectively, then its mode is approximately **[2005]**

(A) 22.0 (B) 20.5
(C) 25.5 (D) 24.0

46. Let x_1, $x2$, …, x_n be n observations such that $\sum x_i^2 = 400$ and $\sum x_i = 80$. Then a possible value of n among the following is **[2005]**

(A) 15 (B) 18
(C) 9 (D) 12

47. Suppose a population A has 100 observations 101, 102, … , 200, and another population B has 100 observations 151, 152, … , 250. If V_A and V_B represent the variances of the two populations, respectively, then $\dfrac{V_A}{V_B}$ is **[2006]**

(A) 1 (B) 9/4
(C) 4/9 (D) 2/3

48. The average marks of boys in a class is 52 and that of girls is 42. The average marks of boys and girls combined is 50. The percentage of boys in the class is **[2007]**

(A) 40 (B) 20
(C) 80 (D) 60

49. The mean of the numbers a, b, 8, 5, 10 is 6 and the variance is 6.80. Then which one of the following gives possible values of a and b? **[2008]**

(A) $a = 0, b = 7$
(B) $a = 5, b = 2$
(C) $a = 1, b = 6$
(D) $a = 3, b = 4$

50. If the mean deviation of number 1, $1 + d$, $1 + 2d$, ….. , 1 + 100d from their mean is 255, then the d is equal to **[2009]**

(A) 10.0 (B) 20.0
(C) 10.1 (D) 20.2

51. For two data sets, each with size 5, the variances are given to be 4 and 5 and the corresponding means are given to be 2 and 4, respectively. The variance of the combined data set is **[2010]**

(A) $\frac{11}{2}$ (B) 6
(C) $\frac{13}{2}$ (D) $\frac{5}{2}$

52. If the mean deviation about the median of the numbers $a, 2a \ldots 50a$ is 50, then $|a|$ equals **[2011]**

(A) 3 (B) 4
(C) 5 (D) 2

53. Let $x_1, x_2 \ldots x_n$ be n observations, and let $\bar{x}$ be their arithmetic mean and σ^2 be their variance.

Statement-1: Variance of $2x_1, 2x_2 \ldots 2x_n$ is $4\sigma^2$.
Statement-2: Arithmetic mean of $2x_1, 2x_2 \ldots 2x_n$ is $4\bar{x}$. **[2012]**

(A) Statement-1 is false, statement-2 is true
(B) Statement-1 is true, statement-2 is true; statement 2 is a correct explanation for statement 1
(C) Statement-1 is true, statement-2 is true; statement 2 is not a correct explanation for statement 1
(D) Statement-1 is true, statement-2 is false

54. All the students of a class performed poorly in Mathematics. The teacher decided to give grace marks of 10 to entire class. Which of the following statistical measures will not change even after the grace marks were given? **[2013]**

(A) median (B) mode
(C) variance (D) mean

55. The variance of the first 50 even natural numbers is **[2014]**

(A) $\frac{833}{4}$ (B) 833
(C) 437 (D) $\frac{437}{4}$

56. The mean of the data set comprising of 16 observations is 16. If one of the observation valued 16 is deleted and three new observations valued 3, 4 and 5 are added to the data, then the mean of the resultant data, is **[2015]**

(A) 16.0 (B) 15.8
(C) 14.0 (D) 16.8

57. If the standard deviation of the numbers 2, 3, a and 11 is 3.5, then which of the following is true? **[2016]**

(A) $3a^2 - 23a + 44 = 0$ (B) $3a^2 - 26a + 55 = 0$
(C) $3a^2 - 32a + 84 = 0$ (D) $3a^2 - 34a + 91 = 0$

ANSWER KEYS

Single Option Correct Type

1. (B)	**2.** (A)	**3.** (A)	**4.** (C)	**5.** (B)	**6.** (A)	**7.** (C)	**8.** (B)	**9.** (D)	**10.** (A)
11. (B)	**12.** (A)	**13.** (B)	**14.** (B)	**15.** (C)	**16.** (D)	**17.** (A)	**18.** (D)	**19.** (C)	**20.** (C)
21. (A)	**22.** (B)	**23.** (C)	**24.** (B)	**25.** (B)	**26.** (C)	**27.** (C)	**28.** (B)	**29.** (B)	**30.** (D)
31. (C)	**32.** (A)	**33.** (D)	**34.** (C)	**35.** (D)	**36.** (D)	**37.** (B)	**38.** (D)	**39.** (C)	**40.** (D)

Previous Year's Questions

41. (C)	**42.** (D)	**43.** (C)	**44.** (C)	**45.** (D)	**46.** (B)	**47.** (A)	**48.** (C)	**49.** (D)	**40.** (C)
51. (A)	**52.** (B)	**53.** (D)	**54.** (C)	**55.** (B)	**56.** (C)	**57.** (C)			

HINTS AND SOLUTIONS

Single Option Correct Type

1. $M = \dfrac{x_1 + x_2 + x_3 \ldots x_n}{n}$

$nM = x_1 + x_2 + x_3 + \ldots x_{n-1} + x_{n0}$

i.e., $nM - x_n = x_1 + x_2 + x_3 + \ldots x_{n-1}$

$$\frac{nM - x_n + x'}{n} = \frac{x_1 + x_2 + x_3 + \ldots x_{n-1} + x'}{n}$$

$\therefore$ New average $= \dfrac{nM - x_n + x'}{n}$

The correct option is (B)

2. If each item of a data is increased or decreased by the same constant, the standard deviation of the data remains unchanged.

The correct option is (A)

3. We know that

Q.D. $= \dfrac{5}{6} \times$ M.D. $= \dfrac{5}{6} \times 12 = 10$

$\therefore$ S.D. $= \dfrac{3}{2} \times$ Q.D. $= \dfrac{3}{2} \times 10 \Rightarrow$ S.D. $= 15$

The correct option is (A)

4. Let the values of 9 items be $x_1, x_2, \ldots, x_9$

Therefore, mean of $x_1, x_2, \ldots, x_9$ is

$15 = \dfrac{x_1 + x_2 + \ldots + x_9}{9}$

$\Rightarrow x_1 + x_2 + \ldots + x_9 = 15 \times 9 = 135$

Let x_{10} be the 10th item. The mean of $x_1, x_2, \ldots x_9, x_{10}$ is 16.

$\Rightarrow \dfrac{x_1 + x_2 + \ldots + x_9 + x_{10}}{10} = 16$

$\Rightarrow x_1 + x_2 + \ldots + x_9 + x_{10} = 160 \Rightarrow 135 + x_{10} = 160$

$\therefore x_{10} = 160 - 135 = 25$

The correct option is (C)

5. Corrected $\Sigma x = 40 \times 200 - 50 + 40 = 7990$

$\therefore$ Corrected $\bar{x} = 7990 / 200 = 39.95$

Incorrect $\Sigma x^2 = n[\sigma^2 + \bar{x}^2] = 200[15^2 + 40^2] = 365000$

Correct $\Sigma x^2 = 365000 - 2500 + 1600 = 364100$

$\therefore$ Corrected $\sigma = \sqrt{\dfrac{364100}{200} - (39.95)^2}$

$= \sqrt{(1820.5 - 1596)}$

$= \sqrt{224.5} = 14.98$

The correct option is (B)

6. We know that, S.D. $= \dfrac{3}{2}$ Q.D.

$\therefore$ S.D. $= \dfrac{3}{2} \times 16 = 24$

The correct option is (A)

7. We have,

$$\bar{x} = \frac{0 \cdot {}^nC_0 + 1 \cdot {}^nC_1 + 2 \cdot {}^nC_2 + \ldots + n \cdot {}^nC_n}{{}^nC_0 + {}^nC_1 + {}^nC_2 + \ldots + {}^nC_n}$$

$$\Rightarrow \bar{x} = \frac{\sum_{r=0}^{n} r \cdot {}^nC_r}{\sum_{r=0}^{n} r \cdot {}^nC_r} \Rightarrow \bar{x} = \frac{1}{2^n} \sum_{r=0}^{n} \not{r} \cdot \frac{n}{\not{r}} {}^{n-1}C_{r-1}$$

$$\Rightarrow \bar{x} = \frac{n}{2^n} \sum_{r=1}^{n} {}^{n-1}C_{r-1} = \frac{n}{2^n} 2^{n-1} = \frac{n}{2}$$

$$\left[\because \sum_{r=1}^{n} {}^{n-1}C_{r-1} = 2^{n-1}\right]$$

and $$\frac{1}{n} \sum f_i x_i^2 = \frac{\sum f_i x_i^2}{\sum f_i} = \frac{\sum_{r=0}^{n} r^{2n}C_r}{\sum_{r=0}^{n} {}^nC_r}$$

$$\Rightarrow \frac{1}{n} \sum f_i x_i^2 = \frac{\sum_{r=0}^{n} r^{\not{2}} \frac{n}{\not{r}} {}^{n-1}C_{r-1}}{2^n}$$

$$\Rightarrow \frac{1}{n} \sum f_i x_i^2 = \frac{n}{2^n} \sum_{r=0}^{n} (r - 1 + 1)\ {}^{n-1}C_{r-1}$$

$$\Rightarrow \frac{1}{n} \sum f_i x_i^2 = \frac{n}{2^n} \left[\sum_{r=0}^{n} (r-1)^{n-1}C_{r-1} + \sum_{r=0}^{n} {}^{n-1}C_{r-1}\right]$$

$$= \frac{n}{2^n} \left[\sum_{r=0}^{n} (r-1) \frac{(n-1)}{(r-1)} {}^{n-2}C_{r-2} + 2^{n-1}\right]$$

$$\Rightarrow \frac{1}{n} \sum f_i x_i^2 = \frac{n}{2^n} [(n-1)2^{n-2} + 2^{n-1}]$$

$$= \frac{n}{2^n} [(n - 1 + 2)2^{n-2}] = \frac{n}{2^n} (n+1)$$

Now, Var $(x) = \dfrac{1}{n} \sum f_i x_i^2 - \bar{x}^2$

$\therefore$ Var $(x) = \dfrac{n(n+1)}{4} - \dfrac{n^2}{4} = \dfrac{n}{4}$

The correct option is (C)

8. S.D. $(\sigma) = \sqrt{\dfrac{250}{10}} = \sqrt{25} = 5$

Hence, coefficient of variation $= \dfrac{\sigma}{\text{mean}} \times 100$

$$= \frac{5}{50} \times 100 = 10\%.$$

The correct option is (B)

9. S.D $= \sigma^2_{x-y}$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i - \bar{x} + \bar{y})^2$$

$$= \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 + \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y})$$

$$- \frac{2}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

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

As cov (x, y) is not known, therefore we cannot find σ^2_{x-y} or σ_{x-y}. Hence data is insufficient.

The correct option is (D)

10. We have $r = \max\limits_{i \neq j} |x_i - x_j|$

and $\quad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Now, $(x_i - \bar{x})^2 = \left(x_i - \frac{x_1 + x_2 + \ldots + x_n}{n}\right)^2$

$$= \frac{1}{n^2}[(x_i - x_1) + (x_i - x_2) + \ldots + (x_i - x_i - 1) + (x_i - x_i + 1) + \ldots + (x_i - x_n)]$$

$$\leq \frac{1}{n^2}[(n-1)r]^2 \qquad (\because |x_i - x_j| \leq r)$$

$$\Rightarrow (x_i - \bar{x}) \leq r^2 \Rightarrow \sum_{i=1}^{n}(x_i - \bar{x})^2 \leq nr^2$$

$$\Rightarrow \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \frac{nr^2}{(n-1)} \Rightarrow S^2 \leq \frac{nr^2}{(n-1)}$$

$$\Rightarrow S \leq r\sqrt{\frac{n}{n-1}}$$

The correct option is (A)

11. Let the numbers be $x_1, x_2, \ldots, x_n$. Then,

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$\Rightarrow \bar{x} = \frac{x_1 + x_2 + \ldots + x_{n-1} + x_n}{n}$$

$$\Rightarrow \bar{x} = \frac{k + x_n}{n} \qquad [\because x_1 + x_2 + \ldots + x_{n-1} = k]$$

$$\therefore \; x_n = n\bar{x} - k$$

The correct option is (B)

12. The required mean is

$$\Rightarrow \bar{x} = \frac{0 \cdot q^n + 1 \cdot \frac{n}{1} q^{n-1} p + 2 \cdot \frac{n(n-1)}{2!} q^{n-2} p^2 + \ldots + n \cdot p^n}{q^n + \frac{n}{1} q^{n-1} p + \frac{n(n-1)}{2!} q^{n-2} p^2 + \ldots + p^n}$$

$$\Rightarrow \bar{x} = \frac{0 \cdot {}^nC_0 q^n p^0 + 1 \cdot {}^nC_1 q^{n-1} p + \ldots + n \cdot {}^nC_n q^0 p^n}{{}^nC_0 q^n p^0 + {}^nC_1 q^{n-1} p^1 + \ldots + {}^nC_n q^{n-n} p^n}$$

$$\Rightarrow \bar{x} = \frac{\sum_{r=0}^{n} r \cdot {}^nC_r q^{n-r} p^r}{\sum_{r=0}^{n} {}^nC_r q^{n-r} p^r}$$

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

$$\Rightarrow \bar{x} = \frac{np\left(\sum_{r=1}^{n} {}^{n-1}C_{r-1}\, p^{r-1} q^{(n-1)-(r-1)}\right)}{\sum_{r=0}^{n} {}^nC_r q^{n-r} p^r}$$

$$\Rightarrow \bar{x} = \frac{np(q+p)^{n-1}}{(q+p)^n}$$

$$\therefore \; \bar{x} = np \quad (\because q + p = 1)$$

The correct option is (A)

13. Let $y = \frac{ax+b}{c}$ i.e., $y = \frac{a}{c}x + \frac{b}{c}$

i.e., $y = Ax + B$, where $A = \frac{a}{c}$, $B = \frac{b}{c}$

$$\therefore \; \bar{y} = A\bar{x} + B$$

$$\therefore \; y - \bar{y} = A(x - \bar{x}) \Rightarrow (y - \bar{y})^2 = A^2(x - \bar{x})^2$$

$$\Rightarrow \Sigma(y - \bar{y})^2 = A^2\Sigma(x - \bar{x})^2$$

$$\Rightarrow n \cdot \sigma_y^2 = A^2 \cdot n\sigma_x^2$$

$$\Rightarrow \sigma_y^2 = A^2\sigma_x^2$$

$$\Rightarrow \sigma_y = |A|\sigma_x \Rightarrow \sigma_y = \left|\frac{a}{c}\right|\sigma_x$$

Thus, new S.D. $= \left|\frac{a}{c}\right|\sigma.$

The correct option is (B)

14. Mean deviation is minimum when it is considered about the item, equidistant from the beginning and the end i.e., the median. In this case median is $\frac{101+1}{2}$th i.e., 51st item i.e., x_{51}.

The correct option is (B)

15. Consider $(1 + x)^{50} = {}^{50}C_0 + {}^{50}C_1 x^1 + \ldots + {}^{50}C_{50} x^{50}$ (1)

and $(1-x)^{50} = {}^{50}C_0 - {}^{50}C_1x^1 + \ldots + {}^{50}C_{50}x^{50}$ (2)

Adding (1) and (2), we get on integrating with limits 0 to 1,

$$\left.{}^{50}C_0 x + \frac{{}^{50}C_2}{3}x^3 + \frac{{}^{50}C_4}{5}x^5 + \ldots + \frac{{}^{50}C_{50}x^{51}}{51}\right]_0^1$$

$$= \frac{1}{2}\left[\frac{(1+x)^{51}}{51} - \frac{(1-x)^{51}}{51}\right]_0^1$$

$$\therefore \quad {}^{50}C_0 + \frac{{}^{50}C_2}{3} + \frac{{}^{50}C_4}{5} + \ldots + \frac{{}^{50}C_{50}}{51} = \frac{1}{2}\frac{2^{51}}{51} = \frac{2^{50}}{51}$$

Now number of terms from ${}^{50}C_0$ to ${}^{50}C_{50}$ are 26 (items)

$\therefore$ Required mean

$$\frac{\frac{{}^{50}C_0}{1} + \frac{{}^{50}C_2}{3} + \frac{{}^{50}C_4}{5} + \ldots + \frac{{}^{50}C_{50}}{51}}{26}$$

$$= \frac{2^{50}}{51} \times \frac{1}{26} = \frac{2^{49}}{39 \times 17}$$

The correct option is (C)

16. S.D. of a series is unaltered if each item is raised (reduced) by the same scalar quantity, S.D. is independent of change of origin.

Hence S.D. will be same as it was already.

$\therefore$ S.D. = 30

The correct option is (D)

17. $\dfrac{a^2+b^2+c^2}{3} \geq (a^2 b^2 c^2)^{1/3}$ ($\because$ A.M $\geq$ G.M.)

and $\dfrac{\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}}{3} \geq \left(\frac{1}{a^2}\frac{1}{b^2}\frac{1}{c^2}\right)^{1/3}$

On multiplying, we get

$$\left(\frac{a^2+b^2+c^2}{3}\right) \cdot \left(\frac{\frac{1}{a^2}+\frac{1}{b^2}+\frac{1}{c^2}}{3}\right) \geq 1$$

$\therefore$ Product of the averages of a^2, b^2, c^2 and $\frac{1}{a^2}, \frac{1}{b^2}, \frac{1}{c^2}$ cannot be less than 1.

The correct option is (A)

18. When each item of a data is multiplied by λ, variance is multiplied by λ^2.

Hence, new variance $= 5^2 \times 9 = 225$.

The correct option is (D)

19. Here, for each $x_i = i$, weight $w_i = i^2 + i$

Hence, the required mean $= \dfrac{\Sigma w_i x_i}{\Sigma w_i} = \dfrac{\sum_{i=1}^{n} i(i^2+i)}{\sum_{i=1}^{n} (i^2+i)}$

$$= \frac{\sum_{i=1}^{n} i^3 + \sum_{i=1}^{n} i^2}{\sum_{i=1}^{n} i^2 + \sum_{i=1}^{n} i} = \frac{\frac{n^2(n+1)^2}{4} + \frac{n(n+1)(2n+1)}{6}}{\frac{n(n+1)(2n+1)}{6} + \frac{n(n+1)}{2}}$$

$$= \frac{\frac{n(n+1)}{2}\left[\frac{n(n+1)}{2} + \frac{2n+1}{3}\right]}{\frac{n(n+1)}{2}\left[\frac{2n+1}{3} + 1\right]}$$

$$= \frac{3n^2+7n+2}{2(2n+4)} = \frac{(3n+1)(n+2)}{4(n+2)} = \frac{3n+1}{4}$$

The correct option is (C)

20. G.M. of $3^1, 3^2, 3^3, \ldots, 3^{3n}$

$$= (3.3^2 \cdot 3^3 \ldots 3^{3n})^{\frac{1}{3n}} = (3^{1+2+3+\ldots 3n})^{\frac{1}{3n}}$$

$$= 3^{\frac{3n(3n+1)}{3n \cdot 2}} = 3^{\frac{3n+1}{2}}$$

The correct option is (C)

21. Given

x	1	2	3	4	$\ldots n$
f_w	1	2^2	3^2	4^2	$\ldots n^2$

Weighted mean $= \dfrac{w_1x_1 + w_2x_2 + \ldots + w_ix_i}{w_1 + w_2 + \ldots + w_i} = \dfrac{\Sigma w_i x_i}{\Sigma w_i}$

$w(\bar{x})$ = weighted mean

$$= \frac{1.1 + 2.2^2 + 3.3^2 + \ldots + n.n^2}{1^2 + 2^2 + 3^2 + \ldots + n^2}$$

$$= \frac{1^3 + 2^3 + 3^3 + \ldots + n^3}{1^2 + 2^2 + 3^2 + \ldots + n^2}$$

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

$\therefore$ the reciprocal of the weighted mean is $\dfrac{2(2n+1)}{3n(n+1)}$

The correct option is (C)

22. Given, $\dfrac{\Sigma x_i}{50} = 38$, $\therefore \Sigma x_i = 1900$

New value of $\Sigma x_i = 1900 - 55 - 45 = 1800$, $n = 48$

$\therefore$ New mean $= \dfrac{1800}{48} = 37.5$

The correct option is (B)

23. Total weight of 7 students is $= 55 \times 7 = 385$ kg

Sum of weights of 6 students

$= 52 + 58 + 55 + 53 + 56 + 54 = 328$ kg

$\therefore$ Weight of seventh student $= 385 - 328 = 57$ kg.

The correct option is (C)

24. Given: $n\ \bar{x} = x_1 + x_2 + x_3 + \ldots + x_n$

Now new observation are $x_1 + 3, x_2 + 3.2, \ldots x_n + 3.n$

$\therefore$ new mean

$$= \frac{(x_1+3)+(x_2+3.2)+(x_3+3.3)+...+(x_n+3.n)}{n}$$

$$= \frac{x_1+x_2+x_3+...+x_n}{n} + \frac{3(1+2+3+...+n)}{n}$$

$$= \bar{x} + \frac{3(n+1)}{2}$$

The correct option is (B)

25.

x	1	2	3	4	$...n$
x^2	1	4	9	16	$...n^2$
f_i	1	2	3	4	$...n$

$$\bar{x} = \frac{\Sigma f_i x_i^2}{\Sigma f_i} = \frac{1.1+2.2^2+3.3^2+...+n.n^2}{1+2+3+...+n}$$

$$= \frac{n^2(n+1)^2}{4} \times \frac{2}{n(n+1)} = \cdot\frac{n(n+1)}{2}$$

The correct option is (B)

26.

x_i :	1	2	3	4	...	$n-1$	n
f_i :	n	$n-1$	$n-2$	$n-3$	...	2	1

$$\therefore \sum_{i=1}^{n} x_i f_i = 1\cdot n + 2(n-1) + 3(n-2) +, ..., + (n-2)\,2 + n.1$$

$$= \sum_{r=1}^{n} (n+1)r - r^2$$

$$= (n+1)\sum_{r=1}^{n} r - \sum_{r=1}^{n} r^2$$

$$= \frac{(n+1)\,(n)\,(n+1)}{2} - \frac{n(n+1)\,(2n+1)}{6}$$

$$= \frac{n(n+1)\,(n+2)}{6}$$

Also, $\Sigma f_i = 1 + 2 +, ..., + n = \dfrac{n(n+1)}{2}$

Mean $= \dfrac{\Sigma f_i x_i}{\Sigma f_i} = \dfrac{n(n+1)\,(n+2)}{6} \times \dfrac{2}{n(n+1)} = \dfrac{n+2}{3}$

The correct option is (C)

27. Let the two unknown items be x and y, then Mean = 4

$$\Rightarrow \frac{1+2+6+x+y}{5} = 4$$

$$\Rightarrow \quad x + y = 11 \qquad (1)$$

and variance = 5.2

$$\Rightarrow \frac{1^2+2^2+6^2+x^2+y^2}{5} - (\text{mean})^2 = 5.2$$

$\Rightarrow 41 + x^2 + y^2 = 5[5.2 + (4)^2]$

$\Rightarrow 41 + x^2 + y^2 = 106$

or $x^2 + y^2 = 65$ (2)

Solving (1) and (2) for x and y, we get

$x = 4, y = 7$ or $x = 7, y = 4$

The correct option is (C)

28. Given mean $= 91 \times 2^{3-n}$

x_i	f_i	$f_i x_i$
0	${}^nC_0$	0
2^1	${}^nC_1$	$2\cdot{}^nC_1$
2^2	${}^nC_2$	$2^2\cdot{}^nC_2$
2^3	${}^nC_3$	$2^3\cdot{}^nC_3$
2^n	${}^nC_n$	$2^n\,{}^nC_n$
	$\Sigma f_i = 2^n$	$\Sigma f_i x_i = 3^n - 1$

Now mean $= \dfrac{\Sigma f_i x_i}{\Sigma f_i}$

$$\Rightarrow \frac{91\times 2^3}{2^n} = \frac{3^n-1}{2^n}$$

$\therefore \quad 3^n = 3^6$

or $\quad n = 6$

The correct option is (B)

29. Let n items be denoted by $x_1, x_2, x_3,... x_n$

$\therefore$ new items are $x_1 + 3, x_2 + 3^2, x_3 + 3^3,... x_n + 3^n$

$\therefore$ new mean

$$= \frac{(x_1+3)+(x_2+3^2)+(x_3+3^3)+...+(x_n+3^n)}{n}$$

$$= \frac{(x_1+x_2...+x_n)}{n} + \frac{3^1+3^2+...+3^n}{n}$$

$$= \bar{x} + \frac{3(3^n-1)}{2n}$$

The correct option is (B)

30. The number of numbers in the n^{th} row $= n^2$

Sequence of first terms in different row is

1, 3, 11, 29, 61, ...

$\therefore$ T_n of 1, 3, 11, 29, 61,... $= \dfrac{1}{3}(2n^3 - 3n^2 + n + 3)$ = first element of n^{th} row.

Similarly, sequence of last terms of each row = 1, 9, 27, 59,...

$$\therefore \quad t_n = \frac{1}{3}[2n^3 + 3n^2 + n - 3]$$

= last element of the n^{th} row.

Hence, in the n^{th} row elements can be written as

$$\frac{1}{3}(2n^3 - 3n^2 + n + 3), ... \frac{1}{3}(2n^3 + 3n^2 + n - 3)$$

(**Note**: adding 2 in the preceding number to get the succeeding number)

$\therefore$ sum of the elements of nth row (using sum of n terms of A.P.)

$$= \frac{N}{2}(A+L) = \frac{n^2}{2}\left(\frac{4n^3+2n}{3}\right) = \frac{n^3}{3}(2n^2+1)$$

$\therefore$ mean of the numbers in the n^{th} row

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

Here $\left\{\begin{aligned} & N = n^2, A = T_n, L = t_n \\ & \therefore \quad A + L = T_n + t_n = \frac{4n^3+2n}{3} \end{aligned}\right\}$

The correct option is (D)

31. Let $x_1, x_2, \ldots x_n$ be n observations.

Then, $\bar{x} = \dfrac{1}{n}\Sigma x_i$

Let $y_i = \dfrac{x_i}{\alpha} + 10$

Then, $\bar{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i = \dfrac{1}{\alpha}\left(\dfrac{1}{n}\sum_{i=1}^{n} x_i\right) + \dfrac{1}{n}(10n)$

$\therefore \ \bar{y} = \dfrac{1}{\alpha}\bar{x} + 10 = \dfrac{\bar{x} + 10\alpha}{\alpha}$

The correct option is (B)

32. $\bar{x}_1 = 520,\ \bar{x}_2 = 420$ and $\bar{x} = 500$

Also, we know $\bar{x} = \dfrac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

$\Rightarrow \ 500\,(n_1 + n_2) = 520n_1 + 420n_2 \Rightarrow 20n_1 = 80n_2$

$\Rightarrow \ n_1 : n_2 = 4 : 1$

Hence, the percentage of male employees in the firm

$= \left(\dfrac{4}{4+1}\right) \times 100 = 80\%$

The correct option is (A)

33. Let the weight of the teacher be w kg, then

$40 + \dfrac{1}{2} = \dfrac{30 \times 40 + w}{35+1}$

$\Rightarrow \ 36 \times 40 + 36 \times \dfrac{1}{2} = 35 \times 40 + w \Rightarrow w = 58$

$\therefore$ Weight of the teacher = 58 kg.

The correct option is (D)

34. Average speed $= \dfrac{120+120+120}{\dfrac{120}{30} + \dfrac{120}{25} + \dfrac{120}{50}}$

$= \dfrac{3}{\dfrac{1}{30} + \dfrac{1}{25} + \dfrac{1}{50}}$ km/h.

The correct option is (C)

35. From the given data, median $= \dfrac{25+26}{2}a = 25.5a$

Required mean deviation about median

$= \dfrac{2\,|0.5 + 1.5 + 2.5 + \ldots + 24.5|}{50}|a| = 50$

$\Rightarrow \ |a| = 4$

The correct option is (D)

36. A.M. of $2x_1, 2x_2, \ldots, 2x_n$ is $\dfrac{2x_1 + 2x_2 + \ldots + 2x_n}{n}$

$= 2\left(\dfrac{x_1 + x_2 + \ldots + x_n}{n}\right) = 2\bar{x}$

So statement-2 is false. Variance $(2x_i) = 2^2$ variance (x_i) $= 4\sigma^2$ so statement-1 is true.

The correct option is (D)

37. Mean $= \dfrac{(x_1+4) + (x_2+8) + (x_3+12) + \ldots + (x_{10}+40)}{10}$

$= \dfrac{x_1 + x_2 + \ldots + x_{10}}{10} + \dfrac{4(1+2+3+\ldots+10)}{10}$

$= 20 + 22 = 42$

The correct option is (B)

38. Mean of $a, b, 8, 5, 10$ is 6

$\Rightarrow \ \dfrac{a+b+8+5+10}{5} = 6$

$\Rightarrow \ a + b = 7$ (1)

Given that Variance is 6.8

$\therefore$ Variance $= \dfrac{\Sigma(X_i - A)^2}{n}$

$= \dfrac{(a-6)^2 + (b-6)^2 + 4 + 1 + 16}{5} = 6.8$

$\Rightarrow \quad a^2 + b^2 = 25$

$a^2 + (7-a)^2 = 25$ [Using (1)]

$\Rightarrow \ a^2 - 7a + 12 = 0$

$\therefore \quad a = 4, 3$ and $b = 3, 4$

The correct option is (D)

39. Mean $(\bar{x}) = \dfrac{\text{sum of quantities}}{n}$

$= \dfrac{\frac{\cancel{n}}{2}(a+1)}{\cancel{n}} = \dfrac{1}{2}[1 + 1 + 100d] = 1 + 50d$

M.D. $= \dfrac{1}{n}\sum|x_i - \bar{x}|$

$\Rightarrow 255 = \dfrac{1}{101}(50d + 49d + 48d + \ldots + d + 0 + d + \ldots + 50d)$

$= \dfrac{2d}{101}\left(\dfrac{50 \times 51}{\cancel{2}}\right)$

$\Rightarrow \ d = \dfrac{255 \times 101}{50 \times 51} = 10.1$

The correct option is (C)

40. Statement-2 is true

Statement-1: Sum of n even natural numbers $= n\,(n+1)$

Mean $(\bar{x}) = \dfrac{n(n+1)}{n} = n + 1$

$$\text{Variance} = \left[\frac{1}{n}\sum (x_i)^2\right] - (\bar{x})^2$$

$$= \frac{1}{n}[2^2 + 4^2 + \ldots + (2n)^2] - (n+1)^2$$

$$= \frac{1}{n}2^2(1^2 + 2^2 + \ldots + n^2) - (n+1)^2$$

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

$$= \frac{(n+1)[2(2n+1) - 3(n+1)]}{3}$$

$$= \frac{(n+1)[4n + 2 - 3n - 3]}{3}$$

$$= \frac{(n+1)(n-1)}{3} = \frac{n^2-1}{3}$$

∴ Statement 1 is false.

The correct option is (D)

Previous Year's Questions

41. $0 \le \frac{3x+1}{3} + \frac{1-x}{4} + \frac{1-2x}{2} \le 1$

$\Rightarrow 12x + 4 + 3 - 3x + 6 - 12x \le 1$

$\Rightarrow 0 \le 13 - 3x \le 12$

$\Rightarrow 3x \le 13$

$\Rightarrow x \ge \frac{1}{13}$.

The correct option is (C)

42. Let the numbers be a and b then $a + b = 18$ and $\sqrt{ab} = 4 \Rightarrow ab = 16$

So, a and b are roots of the equation

$x^2 - 18x + 16 = 0$.

The correct option is (D)

43. Mode can be computed from histogram and median is dependent on the scale. Hence statements (a) and (b) are correct.

The correct option is (C)

44. $x_i = a$ for $i = 1, 2, \ldots, n$ and $x_i = -a$ for $i = n+1, \ldots, 2n$

$$\text{S.D} = \sqrt{\frac{1}{2n}\sum_{i=1}^{2n}(x_i - \bar{x})^2}$$

$$\Rightarrow 2 = \sqrt{\frac{1}{2n}\sum_{i=1}^{2n}x_i^2} \quad \left(\text{Since} \sum_{i=1}^{2n} x_i = 0\right)$$

$$\Rightarrow 2 = \sqrt{\frac{1}{2n}\cdot 2na^2} \Rightarrow |a| = 2$$

The correct option is (C)

45. We have Mode + 2 Mean = 3 Median

$\Rightarrow$ Mode $= 3 \times 22 - 2 \times 21 = 66 - 42 = 24$

The correct option is (D)

46. We have that $\frac{\sum x_i^2}{n} \ge \left(\frac{\sum x_i}{n}\right)^2$

$\Rightarrow n \ge 16$.

The correct option is (B)

47. The variance is given by $\sigma_x^2 = \frac{\Sigma d_i^2}{n}$ (Here deviations are taken from the mean)

Since A and B both has 100 consecutive integers, therefore both have same standard deviation and hence the variance.

$\therefore \frac{V_A}{V_B} = 1$ (As $\sum d_i^2$ is same in both the cases).

The correct option is (A)

48. According to question

$52x + 42y = 50(x + y)$

$\Rightarrow 2x = 8y$

$\Rightarrow \frac{x}{y} = \frac{4}{1}$ and $\frac{x}{x+y} = \frac{4}{5}$

∴ % of boys = 80.

The correct option is (C)

49. Mean of $a, b, 8, 5, 10$ is 6

$$\Rightarrow \frac{a+b+8+5+10}{5} = 6$$

$\Rightarrow a + b = 7$ (1)

Given that Variance is 6.8

$$\therefore \text{Variance} = \frac{\sum (X_i - A)^2}{n}$$

$$= \frac{(a-6)^2 + (b-6)^2 + 4 + 4 + 16}{5} = 6.8$$

$\Rightarrow a^2 + b^2 = 25$

$a^2 + (7 - a)^2 = 25$ (Using (1))

$\Rightarrow a^2 - 7a + 12 = 0$

$\therefore a = 4, 3$ and $b = 3, 4$.

The correct option is (D)

50. $$\text{Mean } (\bar{x}) = \frac{\text{sum of quantities}}{n} = \frac{\frac{\cancel{n}}{2}(a+l)}{\cancel{n}}$$

$$= \frac{1}{2}[1 + 1 + 100d] = 1 + 50d$$

Now, since M.D. about mean is 255, we have

$$\text{M.D.} = \frac{1}{n}\sum\left|x_i - \bar{x}\right|$$

$$\Rightarrow 255 = \frac{1}{101}[50d + 49d + 48d + ... + d + 0 + d + ... + 50d]$$

$$= \frac{2d}{101}\left[\frac{50\times 51}{2}\right]$$

$$\Rightarrow d = \frac{255\times 101}{50\times 51} = 10.1$$

The correct option is (C)

51. $\sigma_x^2 = 4$

$\sigma_y^2 = 5$

$\bar{x} = 2$

$\bar{y} = 4$

$$\frac{\sum x_i}{5} = 2 \qquad \sum x_i = 10; \sum y_i = 20$$

$$\sigma_x^2 = \left(\frac{1}{2}\sum x_i^2\right) - (\bar{x})^2 = \frac{1}{5}\left(\sum y_i^2\right) - 16$$

$$\sum x_i^2 = 40$$

$$\sum y_i^2 = 105$$

$$\sigma_z^2 = \frac{1}{10}\left(\sum x_i^2 + \sum y_i^2\right) - \left(\frac{\bar{x}+\bar{y}}{2}\right)^2$$

$$= \frac{1}{10}(40 + 105) - 9 = \frac{145 - 90}{10}$$

$$= \frac{55}{10} = \frac{11}{2}$$

The correct option is (A)

52. $\frac{1}{n}\sum\left|x_i - A\right|$

$$\text{A} = \text{Median} = \frac{25a + 26a}{2} = 25 - 5a$$

$$\text{Mean deviation } = \frac{1}{50}\left\{\left|a - 25.5a\right| + \left|2a - 25.5a\right|\right\}$$

$$= \frac{2}{50}\left\{(24.5a + 23.5a) + ...(0.5a)\right\}$$

$$= \frac{2}{50}\{312.5a\} = 50 \text{ Given}$$

$$\Rightarrow 625a = 2500 \Rightarrow a = 4$$

The correct option is (B)

53. $\sigma^2 = \sum\frac{x_i^2}{n}\left(\sum\frac{x_i}{n}\right)^2$

$$\text{Variance of } 2x_1, 2x_2, ..., 2x_n = \sum\frac{(2x_i)^2}{n} - \left(\sum\frac{2x_i}{n}\right)^2$$

$$= 4\left[\sum\frac{x_i^2}{n} - \left(\sum\frac{x_i}{n}\right)^2\right] = 4\sigma^2$$

Statement-1 is true.

$$\text{A.M. of } 2x_1, 2x_2, ..., 2x_n = \frac{2x_1 + 2x_2 + ... + 2x_n}{n}$$

$$= 2\left(\frac{x_1 + x_2 + ... + x_n}{n}\right) = 2\bar{x}$$

Statement-2 is false.

The correct option is (D)

54. Variance is not changed by the change of origin.

$$\sigma = \sqrt{\frac{\sum\left|x - \bar{x}\right|^2}{n}}$$

Therefore, $y = x + 10 \Rightarrow \bar{y} + \bar{x} + 10$

$$\sigma_1 = \sqrt{\frac{\sum\left|y + 10 - \bar{y} - 10\right|^2}{n}} = \sqrt{\frac{\sum\left|y - \bar{y}\right|^2}{n}} = \sigma\ .$$

The correct option is (C)

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

$$\bar{x} = \frac{\sum_{r=1}^{50} 2r}{50} = 51$$

$$\sigma^2 = \frac{\sum_{r-1}^{50} 4r^2}{50} - (51)^2 = 833$$

The correct option is (B)

56. New sum $\sum y_i = (16\times 16 - 16) + (3 + 4 + 5) = 252$

Number of observation = 18

$\Rightarrow$ New mean

$$\bar{y} = \frac{252}{18} = 14\ .$$

The correct option is (C)

57. $\text{S.D.} = \sqrt{\frac{\Sigma x_i^2}{n} - \left(\frac{\Sigma x_i}{n}\right)^2}$

$$\therefore\ \frac{49}{4} = \frac{4 + 9 + a^2 + 121}{4} - \left(\frac{16 + a}{4}\right)^2$$

$$\Rightarrow\ 3a^2 - 32a + 84 = 0$$

The correct option is (C)