ODM ADVANCED LEARNING

13

STATISTICS

MEASURES OF CENTRAL TENDENCY

It is that single value which may be taken as the most suitable representative of the data. This single value is known as the average. Average are generally, the central part of the distribution and therefore, they are also called the measures of Central Tendency.

It can be divided into two groups :

Mathematical average :

(i) Arithmetic mean or mean

(ii) Geometric mean

(iii) Harmonic mean

Positional average :

(i) Medium

(ii) Mode or positional average

ARITHMETIC MEAN

Individual observation or unclassified data :

If $x_1, x_2 .... x_n$ be n observations, then their arithmetic mean is given by $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n}$ or

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

Example 1 :

If the heights of 5 persons are 144 cm, 153 cm, 150 cm, 158 cm and 155 cm respectively, then find mean height.

Sol. Mean Height $= \frac{144+153+150+158+155}{5} = \frac{760}{5} = 152$ cm.

Arithmetic mean of discrete frequency distribution :

Let $x_1, x_2 .... x_n$ be n observation and let $f_1, f_2, ... f_n$ be their corresponding frequencies, then their mean

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + ... + f_nx_n}{f_1 + f_2} \quad \text{or} \quad \bar{x} = \frac{\sum_{i=1}^{n} f_i\, x_i}{\sum_{i=1}^{n} f_i}$$

Example 2 :

Find the arithmetic mean of the following frequency distribution :

x:	4	7	10	13	16	19
f:	7	10	15	20	25	30

Sol. The given frequency distribution is –

x_1	f_i	$\Sigma f_i x_i$
4	7	28
7	10	70
10	15	150
13	20	260
16	25	400
19	30	570
	$\Sigma f_i = 107$	$\Sigma f_i x_i = 1478$

$$\bar{x} = \frac{\Sigma f_i\, x_i}{\Sigma f_i} = \frac{1478}{107} = 13.81$$

Short cut method : If the values of x or (and) f are large the calculation of arithmetic mean by the previous formula used, is quite tedious and time consuming. In such case we take the deviation from an arbitrary point A

$$\bar{x} = A + \frac{\Sigma f_i\, d_i}{\Sigma f_i}, \quad \text{where, } A = \text{Assumed mean}$$

$d_i = x_i - A$ = deviation for each term

Step deviation method: Sometimes during the application of shortcut method of finding the mean, the deviation d_i are divisible by a common number h (say). In such case the arithmetic is reduced to a great extent taken by

$$u_i = \frac{x_i - A}{h}, i = 1, 2, ... n \quad \therefore \quad \text{mean } \bar{x} = A + h\left(\frac{\Sigma f_i\, u_i}{\Sigma f_i}\right)$$

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

Weighted A.M.

$$= \frac{w_1x_1 + w_2x_2 + ... + w_nx_n}{w_1 + w_2 + ... + w_n} \quad \text{or} \quad \bar{x} = \frac{\sum_{i=1}^{n} w_i\, x_i}{\sum_{i=1}^{n} w_i}$$

Example 3 :

The weighted mean of first n natural number if their weight are same as the number is -

(1) $\frac{n(n+1)}{2}$ (2) $\frac{n+1}{2}$

(3) $\frac{2n+1}{3}$ (4) None of these

Sol. (3). Here the numbers are 1,2, 3,......, n and their weights also are respectively 1, 2, 3......n so weighted mean = $\frac{\Sigma wx}{\Sigma w}$

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

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

Combined mean : If $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k$ are the mean of k series of sizes $n_1, n_2, ..., n_k$ respectively then the mean $\bar{x}$ if the composite series is given by

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + ... + n_k\bar{x}_k}{n_1 + n_2 + ... + n_k}$$

Example 4 :

The mean income of a group of persons is Rs. 400. Another group of persons has mean income Rs. 480. If the mean income of all the persons in the two groups together is Rs. 430, then find the ratio of the number of persons in the groups.

Sol. $\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$

$\because \bar{x}_1 = 400, \bar{x}_2 = 480, \bar{x} = 430$

$\therefore \quad 430 = \frac{n_1(400) + n_2(480)}{n_1 + n_2} \Rightarrow 30n_1 = 50n_2$

$\Rightarrow \quad \frac{n_1}{n_2} = \frac{5}{3}$

Properties of Arithmetic mean:

1. In a statistical data, the sum of the deviation of items from A.M. is always zero.

i.e. $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$ or $\sum_{i=1}^{n} x_i - n\bar{x}$

or $n\bar{x} - n\bar{x} = 0 \quad \left(\because \bar{x} = \frac{\Sigma x_i}{n}\right)$

2. In a statistical data, the sum of squares of the deviation of items from A.M. is least i.e. $\sum_{i=1}^{n}(x_i - \bar{x})^2$ is least.

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

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

5. Arithmetic mean is independent of origin i.e. it is not effected by any change in origin.

Example 5 :

The mean of a set of number is $\bar{x}$ if each number is increased by λ, then find the mean of the new set.

Sol. $\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{\Sigma x_i}{n} \quad \therefore \quad \Sigma x_i = n\bar{x}$

New mean $= \frac{\Sigma(x_i + \lambda)}{n} = \frac{\Sigma x_i + n\lambda}{n} = \bar{x} + \lambda$

MERITS AND DEMERITS OF ARITHMETIC MEAN

Merits of Arithmetic Mean :

(i) It is rigidly defined.
(ii) It is based on all the observation taken.
(iii) It is calculated with reasonable ease and rapidity.
(iv) It is least affected by fluctuations in sampling.
(v) It is based on each observation and so it is a better representative of the data.
(vi) It is relatively reliable
(vii) Mathematical analysis of mean is possible.

Demerits of Arithmetic Mean :

(i) It is severely affected by the extreme values.
(ii) It cannot be represented in the actual data since the mean does not coincide with any of the observed value.
(iii) It cannot be computed unless all the items are known.
(iv) It cannot be calculated for qualitative data incapable of numerical measurements.
(v) It cannot be used in the study of ratios, rates etc.

GEOMETRIC MEAN

Individual data : If $x_1, x_2, x_3, ..., x_n$ are n values of a variate x, none of them being zero, then the geometric mean G is defined as $G = (x_1 x_2 x_3....x_n)^{1/n}$

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

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

Geometric Mean of grouped data :

Let $x_1, x_2, ...x_n$ be n observation and let $f_1, f_2, ..., f_n$ be their corresponding frequency then their Geometric Mean is

$$G = \left(x_1^{f_1} x_2^{f_2} ....x_n^{f_n}\right)^{1/N} \quad \text{where } N = \sum_{i=1}^{n} f_i$$

$$\therefore \quad G = \text{antilog}\left(\frac{\sum_{i=1}^{n} f_i \log x_i}{\sum_{i=1}^{n} f_i}\right)$$

Example 6 :

Find the geometric mean of numbers $7, 7^2, 7^3, \ldots\ldots, 7^n$.

Sol. Geometric mean of number $7, 7^2, 7^3, \ldots\ldots, 7^n$

$$= (7 . 7^2 . 7^3 \ldots\ldots 7^n)^{1/n} = (7^{1+2+3+\ldots+n})^{1/n}$$

$$= \left[7^{\frac{n(n+1)}{2}}\right]^{1/n} = 7^{\frac{(n+1)}{2}}$$

HARMONIC MEAN

Harmonic mean is reciprocal of mean of reciprocal.

Individual observation :

The H.M. of $x_1, x_2 \ldots\ldots, x_n$ of n observation is given by

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots\ldots + \frac{1}{x_n}} \quad \text{i.e.} \quad H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

H.M. of grouped data :

Let $x_1, x_2, \ldots\ldots, x_n$ be n observation and let $f_1, f_2 \ldots\ldots, f_n$ be their corresponding frequency then H.M. is

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

Example 7 :

Find the harmonic mean of 2, 4, 5 is ....

Sol. The harmonic mean of 2,4 and 5 is

$$= \frac{3}{\frac{1}{2} + \frac{1}{4} + \frac{1}{5}} = \frac{60}{19} = 3.16$$

Note: Relation between A.M., G.M. and H.M.

A.M. $\geq$ G.M. $\geq$ H.M.

Equality sign holds only when all the observations in the series are same.

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.

Median of an individual series :

Let n be the number of observations-

(i) arrange the data in ascending or descending order.

(ii) **(a) if n is odd then -**

$$\text{Median (M)} = \text{value of} \left(\frac{n+1}{2}\right)^{th} \text{observation}$$

(b) If n is even then

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

$$\text{i.e. } M = \frac{\left(\frac{n}{2}\right)^{th} \text{observation} + \left(\frac{n}{2}+1\right)^{th} \text{observation}}{2}$$

Example 8 :

The number of runs scored by 11 players of a cricket team of India are 5, 19, 42, 11, 50, 30, 21, 0, 52, 36, 27. Find the median.

Sol. Let us arrange the value in ascending order

0, 5, 11, 19, 21, 27, 30, 36, 42, 50, 52

$$\therefore \text{Median } M = \left(\frac{n+1}{2}\right)^{th} \text{value} = \left(\frac{11+1}{2}\right)^{th} \text{value} = 6^{th} \text{ value}$$

Now 6^{th} value in data is 27

$\therefore$ Median = 27 runs

Median of the discrete frequency distribution :

Algorithm to find the median :

Step-I : Find the cumulative frequency (C.F.)

Step-II : Find $\frac{N}{2}$, where $N = \sum_{i=1}^{n} f_1$

Step-III : See the cumulative frequency (C.F.) just greater than $\frac{N}{2}$ and determine the corresponding value of the variable.

Step-IV The value obtained in step III is the median.

Example 9 :

Find the median for the following distribution :

x:	1	2	3	4	5	6	7	8	9
f:	8	10	11	16	20	25	15	9	6

Sol.

x	f	c.f.
1	8	8
2	10	18
3	11	29
4	16	45
5	20	65
6	25	90
7	15	105
8	9	114
9	6	120

$$N = 120 = \sum f_i \qquad \therefore \frac{N}{2} = 60$$

We find that the C.F. just greater than $\frac{N}{2}$ is 65 and the value of x corresponding to 65 is 5, therefore median is 5.

Median of ground data or continuous series :

Let the no. of observations be n

(i) Prepare the cumulative frequency table

(ii) Find the median class i.e. the class in which the $\left(\frac{N}{2}\right)^{th}$ observation lies.

(iii) The median value is given by the formula

$$\text{Median (M)} = \ell + \left[\frac{\left(\frac{N}{2}\right) - F}{f}\right] \times h$$

N = total limit frequency = $\sum f_i$

ℓ = lower of median class

f = frequency of the median class

F = cumulative frequency of the class preceding the median class

h = class interval (width) of the median class

Properties of Median :

(1) The sum of the absolute value of deviations of the items from median is minimum.

(2) It is a positional average and it is not influenced by the position of the items.

MODE

Mode is that value in a series which occurs most frequently. In a frequency distribution, mode is that variable which has the maximum frequency.

Computation of Mode :

Mode for individual series : In the case of individual series, the value which is repeated maximum number of times is the mode of the series -

Example 10 :

Find the mode of the data 3, 2, 5, 2, 3, 5, 6, 6, 5, 3, 5, 2, 5.

Sol. Since 5 is repeated maximum number of times, therefore mode of the given data is 5.

Mode for grouped data (discrete frequency distribution series) : In the case of discrete frequency distribution, mode is the value of the variate corresponding to the maximum frequency.

Mode for continuous frequency distribution :

(i) First find the model class i.e. the class which has maximum frequency. The model class can be determined either by inspecting or with the help of grouping data.

(ii) The mode is given by the formula

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

where, $\ell \rightarrow$ lower limit of the model class

h $\rightarrow$ width of the model class

$f_m \rightarrow$ frequency of the model class

$f_{m-1} \rightarrow$ frequency of the class preceding model class

$f_{m+1} \rightarrow$ frequency of the class succeeding model class

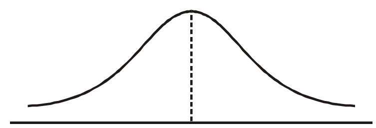
(iii) In case the model value lies in a class other than the one containing maximum frequency (model class) then we use the following formula :

$$\text{Mode} : \ell + \frac{f_{m+1}}{f_{m-1} + f_{m+1}} \times h$$

Properties of Mode : It is not effected by presence of extremely large or small items.

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.



mean = median = mode

A distribution which is not symmetric is called a skewed distribution. In a moderately asymmetric distribution, the interval between the mean and the median is approximately one-third of the interval between the mean and the mode i.e., when have the following empirical relation between them,

Mean – Mode = 3 (Mean – Median)

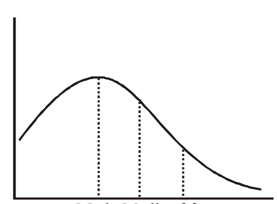
$\Rightarrow$ Mode = 3 Median – 2 Mean.

It is known as Empirical relation.

Positively skewed :

A distribution is positively skewed when is has a tail extending out to the right (larger numbers) When a distribution is positively skewed, the mean is greater than the median reflecting the fact that the mean is sensitive to each score in the distribution and is subject to large shifts when the sample is small and contains extreme scores.

Mean > Median > Mode



Mode Median Mean

Negatively skewed :

A negatively skewed distribution has an extended tail pointing to the left (smaller numbers) and reflects bunching of numbers in the upper part of the distribution with fewer scores at the lower end of the measurement scale.

Mean < Median < Mode.

In a moderately asymmetric distribution, the interval between the mean and the median is approximately one-third of the interval between the mean and the mode i.e., when have the following empirical relation between them,
Empirical formula : mode = 3 median – 2 mean

$$\text{Coefficient of skewness} = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

Limitations of central values :
An average, such as the mean or the median only locates the centre of the data and does not tell us anything about the spread of the data.

Example 11 :
If the value of mode and mean is 60 and 66 respectively, then find the value of median.

Sol. Mode = 3 Median – 2 Mean

$$\therefore \text{ Median} = \frac{1}{3}(\text{mode} + 2\text{ mean}) = \frac{1}{3}(60 + 2 \times 66) = 64$$

MEASURES OF DISPERSION

Dispersion is the measure of the variations. The degree to which numerical data tend to spread about an average value is called the dispersion of the data.
The measures of dispersion commonly used are -
(i) Range
(ii) Quartile deviation or the semi-interquartile range
(iii) Mean Deviation
(iv) Standard Deviation

Range : The difference between the greatest and the least values of variate of a distribution, is called the range of that distribution. If the distribution is continuous grouped distribution, then its
Range = upper limit of the maximum class – lower limit of the minimum class.
Also the coefficient of the range

$$= \frac{\text{difference of extreme values}}{\text{sum of extreme values}}$$

Quartile deviation : Quartile deviation $Q = \frac{Q_3 - Q_1}{2}$;

$$\text{Coefficient of Quartile deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Note : If the distribution is symmetrical, then $Q = M - Q_1 = Q_3 - M$ where M is the median.

Mean Deviation : Mean deviation is defined as the arithmetic mean of the absolute deviations of all the values taken about any central value.

(i) Mean deviation of individual observations :
If $x_1, x_2, \ldots\ldots x_n$ are n values of a variable x, then the mean deviation from an average A (median or AM) is given by

$$\text{M.D.} = \frac{1}{n}\sum_{i=1}^{n}|x_1 - A| = \frac{1}{n}\sum|d_i|, \text{ where } d_i = x_i - A$$

Example 12 :
Find the mean deviation about median from the following data : 340, 150, 210, 240, 300, 310, 320.

Sol. Arranging the observations in ascending order of magnitude, we have 150, 210, 240, 300, 310, 320, 340 clearly, the middle observation is 300. So, median = 300
Calculation of Mean deviation

x_i	$\|d_i\| = \|x_i - 300\|$
340	40
150	150
210	90
240	60
300	0
310	10
320	20

Total $\sum|d_i| = \sum|x_i - 300| = 370$

$$\text{Mean deviation} = \frac{1}{n}\sum|d_i| = \frac{1}{7}\sum|x_i - 300| = \frac{370}{7} = 52.8$$

(ii) Mean deviation of a discrete frequency distribution :
If $x_1, x_2, \ldots, x_n$ are n observation with frequencies $f_1, f_2 \ldots\ldots f_n$, then mean deviation from an average A is given by -

$$\text{Mean Deviation} = \frac{1}{n}\sum f_i|x_i - A|, \quad \text{where } N = \sum_{i=1}^{n} f_i$$

(iii) Mean deviation of a ground or continuous frequency distribution : For calculating mean deviation of a continuous frequency distribution the procedure is same as for a discrete frequency distribution. The only difference is that here we have to obtain the mid-point of the various classes and take the deviations of these mid points from the given central value (median or mean)

VARIANCE AND STANDARD DEVIATION

The variance of a variate x is the arithmetic mean of the squares of all deviations of x from the arithmetic mean of the observations and is denoted by var (x) or σ^2.
The positive square root of the variance of a variate x is known as standard deviation i.e. standard deviation $= +\sqrt{\text{var}(x)}$

The variance is a measure in squared units and has little meaning with respect to the data. Thus, the standard deviation is a measure of variability expressed in the same units as the data. The standard deviation is very much like a mean or an "average" of these deviations.

(i) Variance of individual observations :

If $x_1, x_2 .... x_n$ are n values of a variable s, then by definition

$$\text{var}(x) = \frac{1}{n}\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right] = \sigma^2 \qquad \text{....... (i)}$$

$$\text{or} \quad \text{var}(x) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \qquad \text{........(ii)}$$

If the values of variable x are large, the calculation of variance from the above formulae is quite tedious and time consuming. In that case we taken deviation from an arbitrary point A (say) then $\text{var}(x) = \frac{1}{n}\sum_{i=1}^{n} d_i^2 - \left(\frac{1}{n}\sum_{i=1}^{n} d_i\right)^2$ .... (iii)

Example 13 :

Marks of 5 students of a group are 8, 12, 13, 15, 22 then find the variance.

Sol. $\bar{x} = \frac{8+12+13+15+22}{5} = 14$

Calculation of variance

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$
8	–6	36
12	–2	4
13	–1	1
15	1	1
22	8	64

$\sum (x_i - \bar{x})^2 = 106 \qquad \because \ n = 5, \ \sum (x_i - \bar{x})^2 = 106$

$\therefore \ \text{var}(x) = \frac{1}{n}\sum (x_i - \bar{x})^2 = \frac{106}{5} = 21.2$

(ii) Variance of a discrete frequency distribution :

If $x_1, x_2, ...x_n$ are n observations with frequencies $f_1, f_2, ...f_n$

$$\text{then} \quad \text{var}(x) = \frac{1}{N}\left\{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2\right\} \qquad \text{........ (i)}$$

$$\text{or} \quad \text{var}(x) = \frac{1}{N}\sum_{i=1}^{n} f_i x_i - \bar{x}^2 \qquad \text{....... (ii)}$$

If the values of x or f are large, we take the deviations of the values of variable x from an arbitrary point A. (say)

$\therefore \ d_1 = x_i - A; \ i = 1, 2, ... n$

$$\therefore \ \text{Var}(x) = \frac{1}{N}\left(\sum_{i=1}^{n} f_i d_i^2\right) - \left(\frac{1}{N}\sum_{i=1}^{n} f_i d_i\right)^2 \qquad \text{........ (iii)}$$

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

Sometime $d_i = x_i - A$ are divisible by a common number h (say) then

$$u_i = \frac{x_i - A}{h} = \frac{d_i}{h}, \ 1, 2, ..., n$$

$$\text{then} \quad \text{var}(x) = h^2\left[\frac{1}{N}\sum_{i=1}^{n} f_i u_i^2 - \left(\frac{1}{N}\sum_{i=1}^{n} f_i u_i\right)^2\right] \qquad \text{....... (iv)}$$

(iii) Variance of a grouped or continuous frequency distribution: In a grouped or continuous frequency distribution any of the formulae discussed in discrete frequency distribution can be used .

Mathematical properties of variance :

(a) If all values of the variate in a distribution are added (subtracted) by the same quantity (say λ), then the variance of the distribution remains unchanged.
Hence $\text{Var}(X + \lambda) = \text{Var}(X)$

(b) If all values of the variate in a distribution are multiplied by a constant number k, then the variance of the distribution is multiplied by k^2.
Hence $\text{Var}(kX) = k^2 \text{ var}(X)$

(c) From above results (i) and (ii) it is obvious that
$\text{Var}(aX + b) = a^2 \text{ Var}(X)$

(d) For a continuous distribution standard deviation is not less than the mean deviation with respect to AM.

(e) Relationship between measure of dispersion are
$9\,(\text{Q.D.}) = 7.5\,(\text{M.D.}) = 6\,(\text{S.D.})$

i.e. (i) $\text{Q.D.} = \frac{5}{6}(\text{M.D.})$ (ii) $\text{QD} = \frac{2}{3}(\text{S.D.})$ (iii) $\text{M.D.} = \frac{4}{5}(\text{S.D.})$

(f) If AM's of two series containing n_1, n_2 values are $\bar{x}_1, \bar{x}_2$ and their variance's are σ_1^2, σ_2^2 respectively and combined mean is $\bar{x}$ then the variance of their combined series is given by

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

where $d_1 = \bar{x}_1 - \bar{x}, \ d_2 = \bar{x}_2 - \bar{x}$

$$\text{i.e. } \sigma^2 = \frac{n_1\sigma_1^2 + n_2\sigma_2^2}{(n_1 + n_2)} + \frac{n_1 n_2}{(n_1 + n_2)^2}(\bar{x}_1 - \bar{x}_2)^2$$

Example 14 :

Calculate the mean and standard deviation of first n natural numbers.

Sol. Here $x_i = i = i = 1, 2, ..., n$. Let $\overline{X}$ be the mean and σ be the S.D. Then,

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{n}\sum_{i=1}^{n} i = \frac{1}{n}\ (1+2+3+......+n)$$

$$\Rightarrow \quad \overline{X} = \frac{n(n+1)}{2n} = \frac{n+1}{2}$$

and $\sigma^2 = \frac{1}{n}\left(\sum_{i=1}^{n} x_i^2\right) - \left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)^2$

$$\Rightarrow \quad \sigma^2 = \frac{1}{n}(1^2 + 2^2 + ...+ n^2) - \left(\frac{n+1}{2}\right)^2$$

$$\Rightarrow \quad \sigma^2 = \frac{n\,(n+1)\,(2n+1)}{6n} - \left(\frac{n+1}{2}\right)^2$$

$$\Rightarrow \quad \sigma^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{n^2-1}{12}$$

Example 15 :

Find the variance and standard deviation for the following distribution:

Classes	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90	90 – 100
Frequency	3	7	12	15	8	3	2

Sol. Calculation of Variance and Standard Deviation

Class	Frequency (f_i)	Mid-point (x_i)	$y_i = \frac{x_i - 65}{10}$	y_i^2	$f_i y_i$	$f_i y_i^2$
30-40	3	35	–3	9	–9	27
40-50	7	45	–2	4	–14	28
50-60	12	55	–1	1	–12	12
60-70	15	65	0	0	0	0
70-80	8	75	1	1	8	8
80-90	3	85	2	4	6	12
90-100	2	95	3	9	6	18
	N = 50				–15	105

Therefore $\overline{x} = A + \frac{\sum f_i y_i}{50} \times h = 65 - \frac{15}{50} \times 10 = 62$

variance $\sigma^2 = \frac{h^2}{N^2}\left[N\sum f_i y_i^2 - \left(\sum f_i y_i\right)^2\right]$

$$= \frac{(10)^2}{(50)^2}[50 \times 105 - (-15)^2] = \frac{1}{25}\,[5250 - 225] = 201$$

and standard deviation $(\sigma) = \sqrt{201} = 14.18$

MEAN AND VARIANCE OF BINOMIAL DISTRIBUTION

If the frequencies of the values 0, 1, 2, ..., n of a variate are represented by the following coefficients of a binomial :

$q^n,\ {}^nC_1\, q^{n-1}p,\ {}^nC_2 q^{n-2}p^2,\ .....,\ p^n$

where P is the probability of the success of the experiment (variate), q is the probability of its failure and p + q = 1 i.e. distribution is a binomial distribution : then

$P\ (x = r) = {}^nC_r q^{n-r}\, p^r$; mean $\overline{x} = \sum p_i x_i = np$

Variance $\sigma^2 = npq = \overline{x}q$

Example 16 :

The mean and variance of a variate X having a binomial distribution are 6 and 4 respectively. Find the number of values of the variate in the distribution.

Sol. Here np = 6, npq = 4

$$\Rightarrow q = \frac{2}{3}, p = 1 - \frac{2}{3} = \frac{1}{3} \quad \therefore np = 6 \Rightarrow n = 18$$

Analysis of Frequency Distributions :

Measures of dispersion are unable to compare two or more series which are measured in different units even if they have the same mean. Thus, we require those measures which are independent of the units. The measure of variability which is independent of units is called coefficient of variation (C.V.). The coefficient of variation is defined as

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

where σ and $\overline{X}$ are the standard deviation and mean of the data.

For comparing the variability of two series, we calculate the coefficient of variation for each series. The series having greater C.V. is said to be more variable or conversely less consistent, less uniform less stable or less homogeneous than the other and the series having lesser C.V. is said to be more consistent (or homogeneous) than the other.

Example 17 :
The following values are calculated in respect of heights and weights of the students of a section of Class XI :

	Height	Weight
Mean	162.6 cm	52.36
Variance	$127.69\ cm^2$	$23.1361\ kg^2$

Can we say that the weights show greater variation than the heights ?

Sol. To compare the variability, we have to calculate their coefficients of variation

Given Variance of height = $127.69\ cm^2$

Therefore, Standard deviation of height

$\sqrt{127.69}$ cm = 11.3 cm

Also, Variance of weight = $23.1361\ kg^2$

Standard deviation of weight = $\sqrt{23.1361}$ kg = 4.81 kg

Now, the coefficient of variations (C.V.) are given by

$$\text{(C.V.) in heights} = \frac{\text{Standard Deviation}}{\text{Mean}} \times 100$$

$$= \frac{11.3}{162.6} \times 100 = 6.95$$

and $$\text{(C.V.) in weight} = \frac{4.81}{52.36} \times 1000 = 9.18$$

Clearly C.V. in weights is greater than the C.V. in heights Therefore, we can say that weights show more variability than heights.

TRY IT YOURSELF

Q.1 Find the mean deviation about the median for the data 13, 17, 16, 14, 11, 13, 10, 16, 11, 18, 12, 17

Q.2 Find the mean deviation about the median for the data

x_i	5	7	9	10	12	15
f_i	8	6	2	2	2	6

Q.3 Find the mean and variance for the data 6, 7, 10, 12, 13, 4, 8, 12

Q.4 Find the mean and standard deviation

x_i	60	61	62	63	64	65	66	67	68
f_i	2	1	12	29	25	12	10	4	5

Q.5 The mean and variance of 7 observations are 8 and 16, respectively. If five of the observations are 2, 4, 10, 12, 14. Find the remaining two observations.

Q.6 The standard deviation of the data 6, 5, 9, 13, 12, 8, 10 is
(A) $\sqrt{52/7}$ (B) 52/7
(C) $\sqrt{6}$ (D) 6

Q.7 Let a, b, c, d, e be the observations with mean m and standard deviation s. The standard deviation of the observations a + k, b + k, c + k, d + k, e + k is
(A) s (B) k s
(C) s + k (D) s/k

Q.8 Let $x_1, x_2, \ldots x_n$ be n observations. Let $w_i = lx_i + k$ for i = 1, 2, ...n, where l and k are constants. If the mean of x_i's is 48 and their standard deviation is 12, the mean of w_i's is 55 and standard deviation of w_i's is 15, the values of l and k should be
(A) l = 1.25, k = – 5 (B) l = – 1.25, k = 5
(C) l = 2.5, k = – 5 (D) l = 2.5, k = 5

Q.9 Consider the first 10 positive integers. If we multiply each number by –1 and then add 1 to each number, the variance of the numbers so obtained is
(A) 8.25 (B) 6.5
(C) 3.87 (D) 2.87

Q.10 The standard deviation of some temperature data in °C is 5. If the data were converted into °F, the variance would be
(A) 81 (B) 57
(C) 36 (D) 25

ANSWERS

(1) 2.33	**(2)** 3.23	**(3)** 9, 9.25
(4) 64, 1.69	**(5)** 6, 8	**(6)** (A)
(7) (A)	**(8)** (A)	**(9)** (A)
(10) (A)		

ADDITIONAL EXAMPLES

Example 1 :
Mean of 25 observations was found to be 78.4. But later on it was found that 96 was misread 69. Find the correct mean.

Sol. Mean $\bar{x} = \frac{\Sigma x}{n}$ or $\Sigma x = n\ \bar{x}$; $\Sigma x = 25 \times 78.4 = 1960$

But this Σx is incorrect as 96 was misread as 69.

$\therefore$ correct $\Sigma x = 1960 + (96 - 69) = 1987$

$\therefore$ correct mean $= \frac{1987}{25} = 79.47$

Example 2 :
Find the mean wage from the following data

Wage (in Rs.)	800	820	860	900	920	980	1000
No. of workers	7	14	19	25	20	10	5

Sol. Let the assumed mean be, A = 900. The given data can be written as under :

Wage (in Rs.) x_i	No. of workers f_i	$d_i = x_i - A = x_i - 900$	$u_i = \frac{x_i - 900}{20}$	$f_i u_i$
800	7	–100	– 5	– 35
820	14	–80	– 4	– 56
860	19	– 40	–2	– 38
900	25	0	0	0
920	20	20	1	20
980	10	80	4	40
1000	5	100	5	25

$N = \Sigma f_i = 100$ $\qquad \Sigma f_i u_i = -44$

Here A = 900, h = 20

$$\therefore \text{ Mean } = \bar{X} = A + h\left(\frac{1}{N}\Sigma f_i u_i\right) = 900 + 20\left(-\frac{44}{100}\right) = 891.2$$

Hence, mean wage = Rs. 891.2

Example 3 :

Find the median from the following distribution

Class	5–10	10 – 15	15 – 20	20 – 25	25 – 30
frequency	5	6	15	10	5
Class	30 – 35	35 – 40	40 – 45		
frequency	4	2	2		

Sol.

Class	Frequency	Cumulative frequency
5 – 10	5	5
10 – 15	6	11
15 – 20	15	26
20 – 25	10	36
25 – 30	5	41
30 – 35	4	45
35 – 40	2	47
40 – 45	2	49
		N = 49

Here N = 49. $\therefore \frac{N}{2} = \frac{49}{2} = 24.5$

The cumulative frequency just greater than N/2 is 26 and corresponding class is 15–20. Thus 15 – 20 is the median class such that $\ell = 15, f = 15, F = 11, h = 5$

$$\therefore \text{Medium} = \ell + \frac{N/2 - F}{f} \times h$$

$$= 15 + \frac{24.5 - 11}{15} \times 5$$

$$= 15 + \frac{13.5}{3} = 19.5$$

Example 4 :

Find the mean deviation about mean from the following data

x_i :	3	9	17	23	27
f_i :	8	10	12	9	5

Sol. Calculation of mean deviation about mean.

x_i	f_i	$f_i x_i$	$\lvert x_i - 15\rvert$	$f_i \lvert x_i - 15\rvert$
3	8	24	12	96
9	10	90	6	60
17	12	204	2	24
23	9	207	8	72
27	5	135	12	60

$N = \Sigma f_i = 44$ $\qquad \Sigma f_i x_i = 660$ $\qquad \Sigma f_i |x_i - 15| = 312$

$$\text{Mean} = \bar{X} = \frac{1}{N}(\Sigma f_i x_i) = \frac{660}{44} = 15$$

$$\text{Mean deviation} = \text{M. D.} = \frac{1}{N}\Sigma f_i |x_i - 15| = \frac{312}{44} = 7.09$$

Example 5 :

Find the variance of the data given below

Size of item						
3.5	4.5	5.5	6.5	7.5	8.5	9.5
frequency						
3	7	22	60	85	32	8

Sol. Let the assumed mean be A = 6.5

Calculation of variance

Size of item x_i	f_i	$d_i = x_i - 6.5$	d_i^2	$f_i d_i$	$f_i d_i^2$
3.5	3	– 3	9	–9	27
4.5	7	– 2	4	–14	28
5.5	22	– 1	1	–22	22
6.5	60	0	0	0	0
7.5	85	1	1	85	85
8.5	32	2	4	64	128
9.5	8	3	9	24	72

$N = \Sigma f_i = 217$ $\qquad \Sigma f_i d_i = 128$ $\qquad \Sigma f_i d_i^2 = 362$

$$\therefore \text{Var}(X) = \left(\frac{1}{N}\Sigma f_i d_i^2\right) - \left(\frac{1}{N}\Sigma f_i d_i\right)^2 = \frac{362}{217} - \left(\frac{128}{217}\right)^2$$

$$= 1.668 - 0.347 = 1.321$$

Example 6 :

The mean and standard deviation of 20 observations are found to be 10 and 2 respectively. On rechecking, it was found that an observation 8 was incorrect. Calculate the correct mean and standard deviation in each of the following cases: (i) If the wrong item is omitted. (ii) When 8 is omitted from the data.

Sol. (i) If the wrong item is omitted.

We have , $n = 20$, $\overline{X} = 10$ and $\sigma = 2$

$\therefore \quad \overline{X} = \frac{1}{n}\Sigma x_i \Rightarrow \quad \Sigma x_i = n\overline{X} = 20 \times 10 = 200$

$\Rightarrow$ Incorrect $\Sigma x_i = 200$

and, $\sigma = 2 \quad \Rightarrow \sigma^2 = 4 \quad \Rightarrow \frac{1}{n}\Sigma x_i^2 - (\text{Mean})^2 = 4$

$\Rightarrow \frac{1}{20}\Sigma x_i^2 - 100 = 4 \Rightarrow \Sigma x_i^2 = 104 \times 20$

$\Rightarrow$ Incorrect $\Sigma x_i^2 = 2080$

(ii)When 8 is omitted from the data.

If 8 is omitted from the data, then 19 observations are left.

Now, Incorrect $\Sigma x_i = 200 \Rightarrow$ Correct $\Sigma x_i + 8 = 200$

$\Rightarrow$ Correct $\Sigma x_i = 192$

and Incorrect $\Sigma x_i^2 = 2080$

$\Rightarrow$ Correct $\Sigma x_i^2 + 8^2 = 2080$

$\Rightarrow$ Correct $\Sigma x_i^2 = 2016$

$\therefore$ Correct mean $= \frac{192}{19} = 10.10$

$\Rightarrow$ Correct variance $= \frac{1}{19}(\text{Correct } \Sigma x_i^2) - (\text{Correct mean})^2$

$\Rightarrow$ Correct variance $= \frac{2016}{19} - \left(\frac{192}{19}\right)^2$

Correct variance $= \frac{38304 - 36864}{361} = \frac{1440}{361}$

$\therefore$ Correct standard deviation $= \sqrt{\frac{1440}{361}} = \frac{12\sqrt{10}}{19} = 1.997$

QUESTION BANK

CHAPTER 13 : STATISTICS

EXERCISE - 1 [LEVEL-1]

PART 1 : MEAN, MEDIAN AND MODE

Q.1 If the mean of the set of numbers $x_1, x_2, x_3, ..., x_n$ 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) $\bar{x}+n$

Q.2 The mean of 50 observations is 36. If two observations 30 and 42 are deleted, then the mean of the remaining observations is –
(A) 48 (B) 36
(D) 38 (D) None of these

Q.3 If the mean of n observations $1^2, 2^2, 3^2, ... n^2$ is $\frac{46n}{11}$, then n is equal to –
(A) 11 (B) 12
(C) 23 (D) 22

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

Q.5 If a variable takes values 0, 1, 2, ..., n with frequencies $1, {}^nC_1, {}^nC_2, ...{}^nC_n$, then the AM is –
(A) n (B) $\frac{2^n}{n}$
(C) n + 1 (D) $\frac{n}{2}$

Q.6 The mean of the series $x_1, x_2, ...x_n$ is $\bar{X}$. If x_2 is replaced by λ, then the new mean is –
(A) $\bar{X}-x_2+\lambda$ (B) $\frac{\bar{X}-x_2-\lambda}{n}$
(C) $\frac{(n-1)\bar{X}+\lambda}{n}$ (D) $\frac{n\bar{X}-x_2+\lambda}{n}$

Q.7 The A.M. of n observation is M. If the sum of n – 4 observations is a, then the mean of remaining 4 observation is –
(A) $\frac{nM-a}{4}$ (B) $\frac{nM+a}{2}$
(C) $\frac{nM-a}{2}$ (D) nM + a

Q.8 The mean of a set of observations is $\bar{x}$. If each observation is divided by α, α ≠ 0 and then is increased by 10 then the mean of the new set is
(A) $\frac{\bar{x}}{\alpha}$ (B) $\frac{\bar{x}+10}{\alpha}$
(C) $\frac{\bar{x}+10\alpha}{\alpha}$ (D) $a\bar{x}+10$

Q.9 The mean of n values of a distribution is $\bar{x}$. If its first value is increased by 1, second by 2, ... then the mean of the new distribution will be –
(A) $\bar{x}+n$ (B) $\bar{x}+\frac{n}{2}$
(C) $\bar{x}+\left(\frac{n+1}{2}\right)$ (D) None of these

Q.10 If the median of 21 observation is 40 and if the observations greater than the median are increased by 6 then the median of the new data will be –
(A) 40 (B) 46
(C) 46 + 40/21 (D) 46 – 40/21

Q.11 The median of 19 observations of a group is 30. If two observations with values 8 and 32 are further included, then the median of the new group of 21 observation will be–
(A) 28 (B) 30
(C) 32 (D) 34

Q.12 For a continuous series the mode is computed by the formula
(A) $\ell+\frac{f_{m-1}}{f_m-f_{m-1}-f_{m+1}}\times C$ or $\ell+\left(\frac{f_1}{f_m-f_1-f_2}\right)\times i$
(B) $\ell=\frac{f_m-f_{m-1}}{f_m-f_{m-1}-f_{m+1}}\times C$ or $\ell+\frac{f_m-f_1}{f_m-f_1-f_2}\times i$
(C) $\ell+\frac{f_m-f_{m-1}}{2f_m-f_{m-1}-f_{m+1}}\times C$ or $\ell+\frac{f_m-f_1}{2f_m-f_1-f_2}\times i$
(D) $\ell+\frac{2f_m-f_{m-1}}{f_m-f_{m-1}-f_{m+1}}\times C$ or $\ell+\frac{2f_m-f_1}{f_m-f_1-f_2}\times i$

Q.13 The mode of the following items is 0, 1, 6, 7, 2, 3, 7, 6, 6, 2, 6, 0, 5, 6, 0
(A) 0 (B) 5
(C) 6 (D) 2

Q.14 If for a slightly assymetric distribution, mean and median are 5 and 6 respectively. What is its mode
(A) 5 (B) 6
(C) 7 (D) 8

Q.15 If the mean of the distribution is 2.6, then the value of y is

Variate x	1	2	3	4	5
Frequency f of x	4	5	y	1	2

(A) 24 (B) 13
(C) 8 (D) 3

Q.16 The mode of the distribution

Marks	4	5	6	7	8
No. of students	6	7	10	8	3

(A) 5 (B) 6
(C) 8 (D) 10

Q.17 If in a moderately asymmetrical distribution mode and mean of the data are 6λ and 9λ respectively, then median is
(A) 8λ (B) 7λ
(C) 6λ (D) 5λ

PART 2 : STANDARD DEVIATION, VARIANCE, MEAN DEVIATION

Q.18 The mean deviation of the data 3, 10, 10, 4, 7, 10, 5 from the mean is
(A) 2 (B) 2.57
(C) 3 (D) 3.75

Q.19 Mean deviation from the mean for the observation –1, 0, 4 is–
(A) $\sqrt{\frac{14}{3}}$ (B) $\frac{2}{3}$
(C) 2 (D) None of these

Q.20 The variance of 2, 4, 6, 8, 10 is –
(A) 8 (B) $\sqrt{8}$
(C) 6 (D) None of these

Q.21 For the values $x_1, x_2 ...., x_{101}$ of a distribution $x_1 < x_2 < x_3 < .... < x_{100} < x_{101}$. The mean deviation of this distribution with respect to a number k will be minimum when k is equal to –
(A) x_1 (B) x_{51}
(C) x_{50} (D) $\frac{x_1 + x_2 + ... + x_{101}}{101}$

Q.22 In any discrete series (when all the value are not same) the relationship between M.D. about mean and S.D. is
(A) M.D. = S.D. (B) M.D. > S.D.
(C) M.D. < S.D. (D) M.D. ≤ S.D.

Q.23 Consider the numbers 1, 2, 3, 4, 5, 6, 7, 8, 9, 10. If 1 is added to each number, the variance of the numbers so obtained is
(A) 6.5 (B) 2.87
(C) 3.87 (D) 8.25

Q.24 If mean and SD of a binomial distribution are 3 and 3/2 respectively, then it is given by –
(A) $\left(\frac{1}{5}+\frac{4}{5}\right)^5$ (B) $\left(\frac{4}{5}+\frac{1}{5}\right)^{60}$
(C) $\left(\frac{3}{4}+\frac{1}{4}\right)^{12}$ (D) $\left(\frac{1}{2}+\frac{1}{2}\right)^{12}$

Q.25 The S.D. of the first n natural numbers is –
(A) $\frac{n+1}{2}$ (B) $\sqrt{\frac{n(n+1)}{2}}$
(C) $\sqrt{\frac{n^2-1}{12}}$ (D) None of these

Q.26 If X is a variable of a binomial distribution with mean = 3 and variance = 2, then P (X ≥ 8) is equal to –
(A) $17/3^9$ (B) $18/3^9$
(C) $19/3^9$ (D) $20/9^9$

Q.27 In an experiment with 15 observations on x, the following results were available $\Sigma x^2 = 2830$, $\Sigma x = 170$. One observation that was 20 was found to be wrong and was replaced by the correct value 30. Then the corrected variance is –
(A) 8.33 (B) 78
(C) 188.66 (D) 177.33

Q.28 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 –
(A) 2 (B) $\sqrt{2}$
(C) $\frac{1}{n}$ (D) $\frac{\sqrt{2}}{n}$

Q.29 If X is a binomial variate with n = 2 and p = 0.6, then variance of the random variable x/2 will be –
(A) 0.12 (B) 0.24
(C) 0.36 (D) 0.48

Q.30 The S.D. of 5 scores 1, 2, 3, 4, 5 is
(A) 2/5 (B) 3/5
(C) $\sqrt{2}$ (D) $\sqrt{3}$

Q.31 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) $40 + \frac{21}{25}$ (D) None of these

Q.32 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

Q.33 The S.D. of a variate x is σ. The S.D. of the variate $\frac{ax+b}{c}$ where a, b, c are constant, 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

Q.34 The mean of 5 observations is 4.4 and their variance is 8.24. If three observations are 1, 2 and 6, the other two observations are
(A) 4 and 8 (B) 4 and 9
(C) 5 and 7 (D) 5 and 9

Q.35 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

Q.36 The mean of 100 observations is 50 and their standard deviation is 5. The sum of all squares of all the observations is –
(A) 50000 (B) 250000
(C) 252500 (D) 255000

Q.37 Let x_1, x_2, x_3, x_4, x_5 be the observations with mean m and standard deviation s. The standard deviation of the observations $kx_1, kx_2, kx_3, kx_4, kx_5$ is –
(A) k + s (B) s/k
(C) k s (D) s

Q.38 Standard deviations for first 10 natural numbers is
(A) 5.5 (B) 3.87
(C) 2.97 (D) 2.87

Q.39 The mean deviation from the median is
(A) Greater than that measured from any other value
(B) Less than that measured from any other value
(C) Equal to that measured from any other value
(D) Maximum if all observations are positive

Q.40 The variance of the first n natural numbers is
(A) $\frac{n^2-1}{12}$ (B) $\frac{n^2-1}{6}$
(C) $\frac{n^2+1}{6}$ (D) $\frac{n^2+1}{12}$

Q.41 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

PART 3 : MISCELLANEOUS

Q.42 The monthly sales for the first 11 months of the year of a certain salesman were Rs. 12000 but due to his illness during the last month the average sales for the whole year came down to Rs. 11,375. The value of the sale during the last month was –
(A) Rs. 4500 (B) Rs. 6000
(C) Rs. 10000 (D) Rs. 8000

Q.43 Product of n positive numbers is unity. The sum of these numbers cannot be less than –
(A) 1 (B) n
(C) n^2 (D) None of these

Q.44 The average age of a group of men and women is 30 years. If average age of men is 32 and that of women is 27, then the percentage of women in the group is –
(A) 60 (B) 50
(C) 40 (D) 30

Q.45 When tested, the lives (in hours) of 5 bulbs were noted as follows: 1357, 1090, 1666, 1494, 1623
The mean deviations (in hours) from their mean is
(A) 178 (B) 179
(C) 220 (D) 356

Q.46 If a , b, c are any three positive numbers, then the least value of $(a + b + c)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right)$ is
(A) 3 (B) 6
(C) 9 (D) None of these

Q.47 The A.M. of the observations 1.3.5, 3.5.7, 5.7.9,....., $(2n-1)(2n+1)(2n+3)$ is –
(A) $2n^3 + 6n^2 + 7n - 2$ (B) $n^3 + 8n^2 + 7n - 2$
(C) $2n^3 + 5n^2 + 6n - 1$ (D) $2n^3 + 8n^2 + 7n - 2$

Q.48 If $\bar{X}_1$ and $\bar{X}_2$ are the means of two distribution such that $\bar{X}_1 < \bar{X}_2$ and $\bar{X}$ is the mean of the combined distribution, then –
(A) $\bar{X} < \bar{X}_1$ (B) $\bar{X} > \bar{X}_2$
(C) $\bar{X} = \frac{\bar{X}_1 + \bar{X}_2}{2}$ (D) $\bar{X}_1 < \bar{X} < \bar{X}_2$

Q.49 The mean deviation from the mean for the set of observations –1, 0, 4 is
(A) $\sqrt{14/3}$ (B) 2
(C) 2/3 (D) None of these

Q.50 The mean monthly salary of the employees in a certain factory is Rs. 500. The mean monthly salaries of male and female employees are respectively Rs. 510 and Rs. 460. The percentage of male employees in the factory is
(A) 60 (B) 70
(C) 80 (D) 90

Q.51 A car completes the first half of its journey with a velocity v_1 and the rest half with a velocity v_2. Then the average velocity of the car for the whole journey is
(A) $\frac{v_1 + v_2}{2}$ (B) $\sqrt{v_1 v_2}$
(C) $\frac{2v_1 v_2}{v_1 + v_2}$ (D) None of these

Q.52 A school has four sections of chemistry in class XII having 40, 35, 45 and 42 students. The mean marks obtained in chemistry test are 50, 60, 55 and 45 respectively for the four sections, the over all average of marks per students is
(A) 53 (B) 45
(C) 55.3 (D) 52.25

Q.53 The total expenditure incurred by an industry under different heads is best presented as a
(A) Bar diagram (B) Pie diagram
(C) Histogram (D) Frequency polygon

Q.54 The mean deviation of the numbers 3, 4, 5, 6, 7 is
(A) 0 (B) 1.2
(C) 5 (D) 25

EXERCISE - 2 (NUMERICAL VALUE BASED QUESTIONS)

NOTE : The answer to each question is a NUMERICAL VALUE.

Q.1 Find the mean deviation about the mean for the data :
4, 7, 8, 9, 10, 12, 13, 17

Q.2 Find the mean deviation about the median for the data :
36, 72, 46, 42, 60, 45, 53, 46, 51, 49

Q.3 Find the mean deviation about the mean for the data :

x_i	10	30	50	70	90
f_i	4	24	28	16	8

Q.4 The variance for the data : 6, 7, 10, 12, 13, 4, 8, 12

is $\frac{592}{X}$. Find the value of X.

Q.5 The variance for the data :

x_i	6	10	14	18	24	28	30
f_i	2	4	7	12	8	4	3

is $\frac{1736}{X}$. Find the value of X.

Q.6 The variance for the data :

x_i	92	93	97	98	102	104	109
f_i	3	2	3	2	6	3	3

is $\frac{320}{X}$. Find the value of X.

Q.7 The standard deviation for the data :

x_i	60	61	62	63	64	65	66	67	68
f_i	2	1	12	29	25	12	10	4	5

is X. Find the integer nearest to the 10 X.

Q.8 Find the variance for the following frequency distributions :

Classes	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50
Frequencies	5	8	15	16	6

Q.9 The mean and standard deviation of six observations are 8 and 4, respectively. If each observation is multiplied by 3, find the new standard deviation of the resulting observations.

EXERCISE - 3 (PREVIOUS YEARS AIEEE / JEE MAIN QUESTIONS)

Q.1 If G_1, G_2 are the geometric means of two series of observations and G is the geometric mean of the ratios of the corresponding observations, then the value of G- **[AIEEE-2002]**

(A) $\frac{G_1}{G_2}$ (B) $\frac{\log G_1}{\log G_2}$

(C) $\log(G_1 . G_2)$ (D) $\log G_1 - \log G_2$

Q.2 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 - **[AIEEE 2003]**

(A) remains the same as that of the original set

(B) is increased by 2

(C) is decreased by 2

(D) is two times the original median

Q.3 The mean and variance of a random variable X having binomial distribution are 4 and 2 respectively, then $P(X = 1)$ is – **[AIEEE 2003]**

(A) 1/4 (B) 1/32

(C) 1/16 (D) 1/8

Q.4 In an experiment with 15 observations on x, the following results were available : $\sum x^2 = 2830$, $\sum x = 170$

One observation that was 20 was found to be wrong and was replaced by the correct value 30. Then the corrected variance is **[AIEEE 2003]**

(A) 8.33 (B) 78.00

(C) 188.66 (D) 177.33

Q.5 Consider the following statements: **[AIEEE 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)

Q.6 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- **[AIEEE 2004]**

(A) 1/n (B) $\sqrt{2}$

(C) 2 (D) $\frac{\sqrt{2}}{n}$

Q.7 If in a frequency distribution, the mean and median are 21 and 22 respectively, then its mode is approximately **[AIEEE-2005]**

(A) 22.0 (B) 20.5

(C) 25.5 (D) 24.0

Q.8 Let $x_1, x_2, ... x_n$ 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 **[AIEEE-2005]**

(A) 15 (B) 18

(C) 9 (D) 12

Q.9 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 $\frac{V_A}{V_B}$ is– **[AIEEE 2006]**

(A) 9/4 (B) 4/9

(C) 2 (D) 1

Q.10 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- **[AIEEE 2007]**

(A) 40 (B) 20

(C) 80 (D) 60

Q.11 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 ? **[AIEEE 2008]**

(A) $a = 5, b = 2$ (B) $a = 1, b = 6$

(C) $a = 3, b = 4$ (D) $a = 0, b = 7$

Q.12 If the mean deviation of the numbers $1, 1 + d, 1+ 2d, ...1 +100d$ from their mean is 255, then the d is equal to - **[AIEEE 2009]**

(A) 10.0 (B) 20.0

(C) 10.1 (D) 20.2

Q.13 **Statement 1 :** The variance of first n even natural numbers is $\frac{n^2-1}{4}$. **[AIEEE 2009]**

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

(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 ture.

Q.14 For two data sets, each of 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 – **[AIEEE 2010]**

(A) 11/2 (B) 6
(C) 13/2 (D) 5/2

Q.15 If the mean deviation about the median of the numbers a, 2a, ..., 50a is 50, then | a | equals – **[AIEEE 2011]**

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

Q.16 Let x_1 , x_2 , ..., x_n be n observations, and let x be their arithmetic mean and σ^2 be the variance. **[AIEEE 2012]**

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.

Q.17 All the students of a class performed poorly in Mathematics. The teacher decided to give gracemarks of 10 to each of the students. Which of the following statistical measures will not change even after the grace marks were given ? **[JEE MAIN 2013]**

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

Q.18 The variance of first 50 even natural numbers is – **[JEE MAIN 2014]**

(A) 833/4 (B) 833
(C) 437 (D) 437/4

Q.19 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 **[JEE MAIN 2015]**

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

Q.20 If the standard deviation of the numbers 2, 3, a and 11 is 3.5, then which of the following is true : **[JEE MAIN 2016]**

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

Q.21 If $\sum_{i=1}^{9}(x_i - 5) = 9$ and $\sum_{i=1}^{9}(x_i - 5)^2 = 45$, then the standard deviation of the 9 items $x_1, x_2, \ldots, x_9$ is :

(A) 2 (B) 3 **[JEE MAIN 2018]**
(C) 9 (D) 4

Q.22 5 students of a class have an average height 150cm and variance 18 cm^2. A new student, whose height is 156 cm, joined them. The variance (in cm^2) of the height of these six students is : **[JEE MAIN 2019 (JAN)]**

(A) 22 (B) 20
(C) 16 (D) 18

Q.23 The mean and variance of seven observations are 8 and 16, respectively. If 5 of the observations are 2, 4, 10, 12, 14, then the product of the remaining two observations is : **[JEE MAIN 2019 (APRIL)]**

(A) 40 (B) 49
(C) 48 (D) 45

Q.24 A student scores the following marks in five tests : 45,54,41,57,43. His score is not known for the sixth test. If the mean score is 48 in the six tests, then the standard deviation of the marks in six tests is **[JEE MAIN 2019 (APRIL)]**

(A) $10/\sqrt{3}$ (B) $100/\sqrt{3}$
(C) $100/3$ (D) $10/3$

Q.25 If the standard deviation of the numbers –1, 0, 1, k is $\sqrt{5}$ where k > 0, then k is equal to **[JEE MAIN 2019 (APRIL)]**

(A) $2\sqrt{\frac{10}{3}}$ (B) $2\sqrt{6}$ (C) $4\sqrt{\frac{5}{3}}$ (D) $\sqrt{6}$

Q.26 The mean and the median of the following ten numbers in increasing order 10, 22, 26, 29, 34, x 42, 67, 70, y are 42 and 35 respectively, then y/x is equal to : **[JEE MAIN 2019 (APRIL)]**

(A) 7/3 (B) 9/4
(C) 7/2 (D) 8/3

Q.27 If for some $x \in R$, the frequency distribution of the marks obtained by 20 students in a test is :

Marks	2	3	5	7
Frequency	$(x+1)^2$	$2x-5$	x^2-3x	x

then the mean of the marks is : **[JEE MAIN 2019 (APRIL)]**

(A) 2.8 (B) 3.2
(C) 3.0 (D) 2.5

Q.28 If both the mean and the standard deviation of 50 observations $x_1, x_2, \ldots, x_{50}$ are equal to 16, then the mean of $(x_1-4)^2, (x_2-4)^2, \ldots, (x_{50}-4)^2$ is : **[JEE MAIN 2019 (APRIL)]**

(A) 525 (B) 380
(C) 480 (D) 400

Q.29 If the data $x_1, x_2, \ldots, x_{10}$ is such that the mean of first four of these is 11, the mean of the remaining six is 16 and the sum of squares of all of these is 2,000; then the standard deviation of this data is : **[JEE MAIN 2019 (APRIL)]**

(A) 4 (B) 2
(C) $\sqrt{2}$ (D) $2\sqrt{2}$

Q.30 If variance of first n natural numbers is 10 and variance of first m even natural numbers is 16 then the value of m + n is **[JEE MAIN 2020 (JAN)]**

Q.31 If the mean and variance of eight numbers 3, 7, 9, 12, 13, 20, x and y be 10 and 25 respectively, then x·y is equal to______ **[JEE MAIN 2020 (JAN)]**

Q.32 Mean and standard deviations of 10 observations are 20 and 2 respectively. If p ($p \neq 0$) is multiplied to each observation and then q ($q \neq 0$) is subtracted then new mean and standard deviation becomes half of original value . Then find q. **[JEE MAIN 2020 (JAN)]**

(A) –10 (B) –20
(C) –5 (D) 10

Q.33 Mean and variance of 20 observation are 10 and 4. It was found, that in place of 11, 9 was taken by mistake find correct variance. **[JEE MAIN 2020 (JAN)]**

(A) 3.99 (B) 3.98
(C) 4.01 (D) 4.02

Q.34 Let the observations x_i ($1 \le i \le 10$) satisfy the equations, $\sum_{i=1}^{10}(x_i - 5) = 10$ and $\sum_{i=1}^{10}(x_i - 5)^2 = 40$. If μ and λ are the mean and the variance of the observations, $x_1 - 3, x_2 - 3, ..., x_{10} - 3$, then the ordered pair (μ, λ) is equal to : **[JEE MAIN 2020 (JAN)]**

(A) (6, 6) (B) (3, 6)
(C) (6, 3) (D) (3, 3)

ANSWER KEY

EXERCISE - 1

Q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A	B	B	A	A	D	D	A	C	C	A	B	C	C	D	C	B	A	B	C	A
Q	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
A	B	D	D	C	C	C	B	A	A	C	A	D	B	B	B	C	C	D	B	A
Q	41	42	43	44	45	46	47	48	49	50	51	52	53	54						
A	C	A	B	C	B	C	D	D	B	C	C	D	B	B						

EXERCISE - 2

Q	1	2	3	4	5	6	7	8	9
A	3	7	16	64	40	11	17	132	12

EXERCISE - 3

Q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
A	A	A	B	B	C	C	D	B	D	C	C	C	D	A	C	D	D	B	C	A	A	B
Q	23	24	25	26	27	28	29	30	31	32	33	34										
A	C	A	B	A	A	D	B	18	54	B	A	D										

CHAPTER- 13 : STATISTICS

SOLUTIONS TO TRY IT YOURSELF

TRY IT YOURSELF

(1) Arranging the data in ascending order, we have

10, 11, 11, 12, 13, 13, 14, 16, 16, 17, 17, 18

Hence, n = 12 (which is even)

So, median is average of 6th and 7th observations

$$\therefore \quad \text{Median} = \frac{13+14}{2} = \frac{27}{2} = 13.5$$

x_i	$\lvert x_i - M \rvert$
10	3.5
11	2.5
11	2.5
12	1.5
13	0.5
13	0.5
14	0.5
16	2.5
16	2.5
17	3.5
17	3.5
18	4.5
Total	28

$$\text{M.D. about median} = \frac{1}{n}\sum_{i=1}^{n} \lvert x_i - M \rvert = \frac{1}{12} \times 28 = 2.33$$

(2)

x_i	f_i	c.f.	$\lvert x_i - 7 \rvert$	$f_i \lvert x_i - 7 \rvert$
5	8	8	2	16
7	6	14	0	0
9	2	16	2	4
10	2	18	3	6
12	2	20	5	10
15	6	26	8	48
	26			84

$$\frac{N}{2} = \frac{26}{2} = 13$$

The C.f. just greater than 13 is 14 and corresponding value of x is 7. $\therefore$ Median = 7

$$\text{M.D. about median} = \frac{1}{N}\sum f_i \lvert x_i - M \rvert = \frac{1}{26} \times 84 = 3.23$$

(3) Here x = 6, 7, 10, 12, 13, 4, 8, 12

$\therefore \quad \Sigma x = 6 + 7 + 10 + 12 + 13 + 4 + 8 + 12 = 72$

$$n = 8 \quad \therefore \quad \bar{x} = \frac{72}{8} = 9$$

$$\Sigma x^2 = (6)^2 + (7)^2 + (10)^2 + (12)^2 + (13)^2 + (4)^2 + (8)^2 + (12)^2 = 722$$

$$\text{Variance} = \sigma^2 = \frac{N\,\Sigma x^2 - (\Sigma x)^2}{N^2} = \frac{8 \times 722 - (72)^2}{(8)^2}$$

$$= \frac{5776 - 5184}{64} = \frac{592}{64} = 9.25$$

(4) We have,

x_i	f_i	$d_i = x_i - 64$	$f_i d_i$	d_i^2	$f_i d_i^2$
60	2	60 – 64 = –4	–8	16	32
61	1	61 – 64 = –3	–3	9	9
62	12	62 – 64 = –2	–24	4	48
63	29	63 – 64 = –1	–29	1	29
64	25	64 – 65 = 0	0	0	0
65	12	65 – 64 = 1	12	1	12
66	10	66 – 64 = 2	20	4	40
67	4	67 – 64 = 3	12	9	36
68	5	68 – 64 = 4	20	16	80
Total	100		0		286

$$\text{Now,} \quad \bar{x} = A + \frac{\Sigma f_i d_i}{\Sigma f_i} = 64 + \frac{0}{100} = 64$$

$$\text{Variance,} \quad \sigma^2 = \frac{1}{N}\{\Sigma f_i d_i^2\} - \left\{\frac{1}{N}\Sigma f_i d_i\right\}^2$$

$$= \left[\frac{1}{100} \times 286 - \left(\frac{1}{100} \times 0\right)^2\right] = 2.86$$

Standard deviation = $\sigma = \sqrt{2.86} = 1.69$

(5) Let x and y be remaining two observations.

$$\text{Then mean} = 8 \Rightarrow \frac{2+4+10+12+14+x+y}{7} = 8$$

$\Rightarrow \quad 42 + x + y = 56 \quad \Rightarrow x + y = 14$ .... (1)

But, variance = 16

$$\Rightarrow \quad \frac{1}{7}(2^2 + 4^2 + 10^2 + 12^2 + 14^2 + x^2 + y^2) - (\text{Mean})^2 = 16$$

$$\Rightarrow \quad \frac{1}{7}(4 + 16 + 100 + 144 + 196 + x^2 + y^2) - 64 = 16$$

[$\because$ Mean = 8]

$\Rightarrow \quad 460 + x^2 + y^2 = 7 \times 80 \Rightarrow x^2 + y^2 = 100$ .... (2)

Now, $(x+y)^2 + (x-y)^2 = 2(x^2 + y^2)$

$\Rightarrow \quad 196 + (x-y)^2 = 2 \times 100$ [Using (1) and (2)]

$\Rightarrow \quad (x-y)^2 = 4 \Rightarrow x - y = \pm 2$

If x – y = 2 then, solving x + y = 14 and x – y = 2, we have, x = 8 and y = 6

If x – y = –2 then, solving x + y = 14 and x – y = –2, x = 6 and y = 8. Remaining two observations are 6 and 8.

(6) (A) **(7)** (A) **(8)** (A)

(9) (A) **(10)** (A)

CHAPTER- 13 : STATISTICS
EXERCISE-1

(1) **(B).** $\because \frac{\Sigma x_i}{n} = \bar{x}$

Let $y_i = x_i + 2i, i = 1, 2, ...n$
Required mean

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

(2) **(B).** $\because$ Sum of the 50 observations = 36 × 50 = 1800
Two observations 30 and 42 are deleted
Sum of the remaining 48 observation
$= 1800 - [30 + 42] = 1728$
Req. mean = 1728/48 = 36

(3) **(A).** $\because \frac{1^2 + 2^2 + 3^2 + ... + n^2}{n} = \frac{46n}{11}$

$\Rightarrow \frac{n(n+1)(2n+1)}{6n} = \frac{46n}{11}$

$\Rightarrow 11(n+1)(2n+1) = 267n$
$\Rightarrow 22n^2 - 243n + 11 = 0 \Rightarrow (n-11)(22n-1) = 0$
$\therefore n = 11 \because n \neq 1/22$

(4) **(A).** On arranging the values in the ascending order

$\alpha - \frac{7}{2}, \alpha - 3, \alpha - \frac{5}{2}, \alpha - 2, \alpha - \frac{1}{2}, \alpha + \frac{1}{2}, \alpha + 4, \alpha + 5$

$(\because \alpha > 0)$

Here number of observations n = 8 (even)

$$\text{Median} = \frac{1}{2}\left[\left(\frac{n}{2}\right)\text{th obser.} + \left(\frac{n}{2}+1\right)\text{th obser.}\right]$$

$$= \frac{1}{2}\left[(\alpha - 2) + \left(\alpha - \frac{1}{2}\right)\right] = \alpha - \frac{5}{4}$$

(5) **(D).** $N = \Sigma f_i = 1 + {}^nC_1 + {}^nC_2 + ... + {}^nC_n = 2^n$
$\Sigma f_i x_i = 1 \times 0 + {}^nC_1 \times 1 + {}^nC_2 \times 2 + ... + n \times {}^nC_n$

$$= n \times 1 + \frac{n(n-1)}{2!} \times 2 + \frac{n(n-1)(n-2)}{3!} + ... + n \times 1$$
$$= n\,[{}^{n-1}C_0 + {}^{n-1}C_1 + {}^{n-1}C_2 + ... + {}^{n-1}C_{n-1}] = n \,.\, 2^{n-1}$$

$$\text{Reqd. A.M.} = \frac{\Sigma f_i x_i}{N} = \frac{n.2^{n-1}}{2^n} = \frac{n}{2}$$

(6) **(D).** $\frac{x_1 + x_2 + x_3 + ... + x_n}{n} = \bar{x}$ .... (i)

when x_2 is replaced by λ then

$$\frac{x_1 + \lambda + x_3 + ... + x_n}{n}$$

$$= \frac{(x_1 + x_2 + x_3 + ... + x_n) + \lambda - x_2}{n} = \frac{n\bar{X} + \lambda - x_2}{n}$$

by (i)

(7) **(A).** Let the mean of remaining 4 observations is $\bar{x}$ the sum of remaining 4 observations = $4\bar{x}$.

$$\therefore M = \frac{a + 4\bar{x}}{(n-4)+4} = \frac{a + 4\bar{x}}{n} \quad \therefore \bar{x} = \frac{nM - a}{4}$$

(8) **(C).** Let set of n observations $x_1, x_2, ...x_n$ and their mean $\bar{x}$.

$\therefore \frac{\Sigma x_i}{n} = \bar{x}$ .... (i)

Each observations is divided by α ($\alpha \neq 0$) and then increased by 10 then let new observations

$y_i = \frac{x_i}{\alpha} + 10$, $i = 1, 2, ...n$

Req. mean of new observations

$$= \frac{1}{n}\Sigma y_i = \frac{1}{n}\Sigma\left(\frac{x_i}{\alpha} + 10\right) = \frac{1}{\alpha}.\frac{\Sigma x_i}{n} + \frac{10n}{n} = \frac{\bar{x}}{\alpha} + 10$$

by (i)

(9) **(C).** Let n values of distribution are $x_1, x_2, ...x_n$ and

$\bar{x} = \frac{\Sigma x_i}{n}$

Let new observations of that distribution
$y_i = x_i + i, \; i = 1, 2, ...n$

$$\text{Req. mean} = \frac{1}{n}\Sigma y_i = \frac{1}{n}\Sigma (n_i + i) = \frac{\Sigma x_i}{n} + \frac{\Sigma i}{n}$$

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

(10) **(A).** n = 21 (odd), median = 40

$\text{Median} = \left(\frac{21+1}{2}\right)\text{th}$ observation = 11^{th};

observation= 40
Since observations which greater than median are increased.
But median is 11^{th} observation which remains unchanged.

(11) **(B).** n = 19 (odd), Median = 30
On arranging observations in ascending order median

$= \left(\frac{n+1}{2}\right)\text{th}$ observation = $(10)^{th}$ observation = 30

Since two observations 8, 32 are include after arranging that 21 observation in ascending order n = 21 (odd)

New median = $\left(\frac{n+1}{2}\right)\text{th}$ = $(11)^{th}$ observation = 30

$\because (8 < 30 < 32)$

(12) **(C).** For a continuous series the mode is computed by the formula :

$$\ell + \frac{f_m - f_{m-1}}{2f_m - f_{m-1} - f_{m+1}} \times C \text{ or } \ell + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times i$$

(13) **(C).** Since 6 occurs most times (i.e., 5 times) than any other 15 observations. $\therefore$ Mode = 6.

(14) **(D).** We know that,
Mode = 3Median – 2Mean = 3(6) – 2(5) = 8.

(15) **(C).** We know that, Mean $= \dfrac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} f_i}$

i.e., $2.6 = \dfrac{1\times4+2\times5+3\times y+4\times1+5\times2}{4+5+y+1+2}$

or $31.2+2.6y = 28+3y$ or $0.4y = 3.2 \Rightarrow y = 8$.

(16) **(B).** Since frequency is maximum for 6. $\therefore$ Mode = 6.

(17) **(A).** For a moderately Skewed distribution,
Mode = 3 median – 2 mean
$\Rightarrow 6\lambda = 3$ median $- 18\lambda \Rightarrow$ median $= 8\lambda$.

(18) **(B).**

(19) **(C).** $\because \ \bar{x} = \dfrac{-1+0+4}{3} = 1$; $\Sigma|x_i - \bar{x}| = 2+1+3 = 6$

Mean deviation $= \dfrac{1}{n}\Sigma|x_i-\bar{x}| = \dfrac{1}{3}\times 6 = 2$

(20) **(A).** $\because \ \bar{x} = \dfrac{30}{5} = 6$

$\because$ Variance $= (\text{S.D.})^2 = \dfrac{1}{n}\Sigma|x_i-\bar{x}|^2 = \dfrac{1}{n}\Sigma x_i^2 - \bar{x}^2$

$= \dfrac{1}{5}[2^2+4^2+6^2+8^2+10^2]-(6)^2 = \dfrac{220}{5} - 36 = 8$

(21) **(B).** Since mean deviation is minimum when it is taken by median of distribution so here K is median of given observations.
K = median $= \left(\dfrac{n+1}{2}\right)$th observation $= 51^{th}$ observation
$\therefore \ K = x_{51}$

(22) **(D).** Let x_i/f_i, i = 1, 2, ... n be a frequency distribution

then its S.D. $= \sqrt{\dfrac{1}{N}\Sigma f_i (x_i-\bar{x})^2}$ and

M.D. $= \dfrac{1}{N}\Sigma f_i |x_i - \bar{x}|$

Let $|x_i - \bar{x}| = y_i$ then,

S.D. $= \sqrt{\dfrac{1}{N}\Sigma f_i y_i^2}$ and M.D. $= \dfrac{1}{N}\Sigma f_i y_i$

Now, $(\text{S.D.})^2 - (\text{M.D.})^2 = \dfrac{1}{N}\Sigma f_i y_i^2 - \left(\dfrac{1}{N}\Sigma f_i y_i\right)^2$

$\Rightarrow \ (\text{S.D.})^2 - (\text{M.D.})^2 = \sigma^2(y) \ge 0$
$\Rightarrow \ (\text{S.D.})^2 \ge (\text{M.D.})^2 \qquad \therefore$ S.D. $\ge$ M.D.

(23) **(D).**

(24) **(C).** Mean $(\bar{x}) = 3$ and S.D. = 3/2

$\because$ Variance $= \bar{x}q \Rightarrow \dfrac{9}{4} = 3\times 2 \Rightarrow q = \dfrac{3}{4}$

$\because \ p+q = 1 \Rightarrow p = 1 - q = \dfrac{1}{4}$

Mean $(\bar{x}) = np \Rightarrow n\times\dfrac{1}{4} = 3 \Rightarrow n = 12$

Hence binomial distribution is given by

$(q+p)^n = \left(\dfrac{3}{4}+\dfrac{1}{4}\right)^{12}$

(25) **(C).** $\because$ $(\text{S.D.})^2 = \dfrac{1}{n}\Sigma(x_i-\bar{x})^2 = \dfrac{1}{n}\Sigma x_i^2 - \left(\dfrac{1}{n}\Sigma x_i\right)^2$

So S.D. of first n natural numbers

$= \sqrt{\dfrac{1}{n}\Sigma n^2 - \left(\dfrac{1}{n}\Sigma n\right)^2}$

$= \sqrt{\dfrac{1}{n}\dfrac{n(n+1)(2n+1)}{6} - \left\{\dfrac{1}{n}\dfrac{n(n+1)}{2}\right\}^2}$

$= \sqrt{\dfrac{n(n+1)(2n+1)}{6} - \dfrac{(n+1)^2}{4}} = \sqrt{\dfrac{(n+1)(n-1)}{12}} = \sqrt{\dfrac{n^2-1}{12}}$

(26) **(C).** Mean $(\bar{x}) = 3$, Variance $(\sigma^2) = 2$

$\because \ \sigma^2 = \bar{x}q \Rightarrow 2 = 3\times q \Rightarrow q = \dfrac{2}{3} \quad \therefore \ p = 1-q = \dfrac{1}{3}$

Again, $\bar{x} = np \Rightarrow 3 = n\times\dfrac{1}{3} \Rightarrow n = 9$

$\because \ P(x=r) = {}^nC_r q^{n-r} p^r$

$\therefore \ P(x\ge 8) = {}^9C_8\left(\dfrac{2}{3}\right)^{9-8}\left(\dfrac{1}{3}\right)^8 + {}^9C_9\left(\dfrac{2}{3}\right)^{9-9}\left(\dfrac{1}{3}\right)^9$

$= {}^9C_1\left(\dfrac{2}{3}\right)\left(\dfrac{1}{3}\right)^8 + {}^9C_0\left(\dfrac{2}{3}\right)^0\left(\dfrac{1}{3}\right)^9$

$= 9\times\dfrac{2}{3}\times\dfrac{1}{3^8} + 1\times1\times\dfrac{1}{3^9} = \dfrac{19}{3^9}$

(27) **(B).** Given n = 15, $\Sigma x = 170$, $\Sigma x^2 = 2830$
Since once observation 20 was found be wrong and it replaced by its correct value 30.
So, corrected $\Sigma y = \Sigma x - 20 + 30 = 180$
corrected $\Sigma y^2 = \Sigma x^2 - 20^2 + 30^2 = 3330$

The correct variance $= \dfrac{1}{n}\Sigma y^2 - \left(\dfrac{1}{n}\Sigma y\right)^2$

$= \dfrac{1}{15}\times 3330 - \left(\dfrac{1}{15}\times 180\right)^2 = 78$

(28) **(A).** Number of observations = 2n (even)
and observations are a, a, ...nand –a, –a.......n
Here $\Sigma x_i = 0 \Rightarrow \bar{x} = 0$. Now,

S.D. $= \sqrt{\dfrac{n(a-0)^2 + n(-a-0)^2}{2n}}$

$$\text{S.D.} = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}} = \sqrt{\frac{na^2 + na^2}{2n}} = \sqrt{a^2} = \pm a = |a|$$

Given S.D. = 2 $\Rightarrow |a| = 2$

(29) **(A).** Given n = 2, p = 0.6

$\because$ p + q = 1 $\Rightarrow$ q = 1 − 0.6 = 0.4

Variance of the variable $x = \sigma^2(x) = npq = 0.48$

Variance of the random variable

$$\frac{x}{2} = \sigma^2\left(\frac{x}{2}\right) = \frac{1}{4}\sigma^2(x) = 0.12$$

(30) **(C).** Mean $\bar{x} = \frac{1+2+3+4+5}{5} = 3$

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

$$= \sqrt{\frac{1}{5}(1+4+9+16+25) - 9} = \sqrt{11-9} = \sqrt{2}.$$

(31) **(A).** If each item of a data is increased or decreased by the same constant, the standard deviation of the data remains unchanged.

(32) **(D).** When each item of a data is multiplied by λ, variance is multiplied by λ^2.

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

(33) **(B).** Let $y = \frac{ax+b}{c}$ i.e., $y = \frac{a}{c}x + \frac{b}{c}$

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

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

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

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

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

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

(34) **(B).** Let the two unknown items be x and y.

Then, mean $= 4.4 \Rightarrow \frac{1+2+6+x+y}{5} = 4.4$

$\Rightarrow x + y = 13$ .....(i)

and variance = 8.24

$$\Rightarrow \frac{1^2+2^2+6^2+x^2+y^2}{5} - (\text{mean})^2 = 8.24$$

$\Rightarrow 41 + x^2 + y^2 = 5\{(4.4)^2 + 8.24\} \Rightarrow x^2 + y^2 = 97$ .....(ii)

Solving (i) and (ii) for x and y, we get

$x = 9,\ y = 4$ or $x = 4, y = 9$.

(35) **(B).** 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 $\Sigma x^2 = 365000 - 2500 + 1600 = 364100$

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

$$= \sqrt{(1820.5 - 1596)} = \sqrt{224.5} = 14.98.$$

(36) **(C).** **(37) (C).** **(38) (D).**

(39) **(B).** It is a fundamental property.

(40) **(A).** Variance $= (\text{S.D.})^2 = \frac{1}{n}\Sigma x^2 - \left(\frac{\Sigma x}{n}\right)^2, \left(\because \bar{x} = \frac{\Sigma x}{n}\right)$

$$= \frac{n(n+1)(2n+1)}{6n} - \left(\frac{n(n+1)}{2n}\right)^2 = \frac{n^2-1}{12}.$$

(41) **(C).** Let the two unknown items be x and y, then

Mean $= 4 \Rightarrow \frac{1+2+6+x+y}{5} = 4 \Rightarrow x + y = 11$ .....(i)

and variance = 5. 2

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

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

$41 + x^2 + y^2 = 106$; $x^2 + y^2 = 65$ .....(ii)

Solving (i) and (ii) for x and y, we get

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

(42) **(A).** Total sales for the first 11 months = 12000 × 11 Rs.

Average sales for the whole year = 11375 Rs.

Let value of the sale during the last month of year was x

Rs. So, $\frac{12000 \times 11 + x}{12} = 11375 \Rightarrow x = 4500$ Rs.

(43) **(B).** Let $x_1, x_2, \ldots\ldots x_n$ are n positive numbers such that

$x_1, x_2, \ldots\ldots x_n = 1$ .... (1)

$\because$ A.M. $\geq$ G.M.

So, $\frac{x_1 + x_2 + \ldots\ldots + x_n}{n} \geq (x_1, x_2, \ldots\ldots x_n)^{1/n} = 1$

$\Rightarrow x_1 + x_2 + \ldots\ldots + x_n \geq n$ by (1)

(44) **(C).** Let the number of men and women in the group are n_1 and n_2 respectively.

Total mean $= \frac{32n_1 + 27n_2}{n_1 + n_2} = 30 \Rightarrow \frac{n_1}{n_2} = \frac{3}{2}$

The % of women $= \frac{2}{5} \times 100 = 40$

(45) **(B).**

(46) **(C).** A.M. and H.M. of three positive number a, b, c.

A.M. $= \frac{a+b+c}{3}$, H.M. $= \frac{3}{\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right)}$

$$\Rightarrow \frac{a+b+c}{3} \geq \frac{3}{\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right)} \quad \because \text{A.M.} \geq \text{H.M.}$$

$\Rightarrow (a+b+c)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}\right) \geq 9$

(47) **(D).** $\because T_n = (2n-1)(2n+1)(2n+3)$

Req. A.M. $= \frac{S_n}{n} = \frac{\Sigma T_n}{n} = \frac{1}{n}[\Sigma (8n^3+12n^2-2n-3)]$

$= \frac{1}{n}[8\Sigma n^3 + 12\Sigma n^2 - 2\Sigma n - \Sigma 3] = 2n^3 + 8n^2 + 7n - 2$

(48) **(D).** Let the number of observations of two sets are n_1 and n_2 whose means $\bar{x}_1$ and $\bar{x}_2$ respectively.

Here mean of combined set is $\bar{x}$ and $\bar{x}_1 < \bar{x}_2$

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

Now, $\bar{x} - \bar{x}_1 = \frac{n_2(\bar{x}_2 - \bar{x}_1)}{n_1+n_2} > 0 \because \bar{x}_2 > \bar{x}_1$

$\Rightarrow \bar{x} > \bar{x}_1$ .... (i)

Again $\bar{x} - \bar{x}_2 = \frac{n_1}{n_1+n_2}(\bar{x}_1 - \bar{x}_2) < 0 \because \bar{x}_1 < \bar{x}_2$

$\Rightarrow \bar{x} < \bar{x}_2$ .... (ii)

By (i) & (ii), $\bar{x}_1 < \bar{x} < \bar{x}_2$

(49) **(B).** Mean $= \frac{-1+0+4}{3} = 1$.

Hence M.D. (about mean) $= \frac{|-1-1|+|0-1|+|4-1|}{3} = 2$.

(50) **(C).** The formula for combined mean is $\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1+n_2}$

Given, $\bar{x} = 500$, $\bar{x}_1 = 510$, $\bar{x}_2 = 460$

Let $n_1 + n_2 = 100$ and n_1 denotes male, n_2 denotes female for this $n_2 = 100 - n_1$

$500 = \frac{510n_1 + (100-n_1)460}{100}$

$\Rightarrow 50000 = 510n_1 + 46000 - 460n_1$

$\Rightarrow 50000 - 46000 = 50n_1 \Rightarrow 4000 = 50n_1$

$\Rightarrow n_1 = \frac{4000}{50} = 80$.

Hence, the percentage of male employees in the factory is 80.

(51) **(C).** $V_{av} = \frac{\text{Total distance}}{\text{Total time taken}}$

Time taken for first half journey is, $t_1 = (d/v_1)$ and time taken for rest half journey is, $t_2 = (d/v_2)$

$\therefore V_{av} = \frac{2d}{(d/v_1)+(d/v_2)} = \frac{2v_1v_2}{v_1+v_2}$.

(52) **(D).** Total number of students = 40 + 35 + 45 + 42 = 162

Total marks obtained

$= (40\times 50) + (35\times 60) + (45\times 55) + (42\times 45)$

$= 8465$

Overall average of marks per students $= \frac{8465}{162} = 52.25$.

(53) **(B).**

(54) **(B).** A.M. $= \frac{3+4+5+6+7}{5} = 5$

$\therefore$ Mean deviation $= \frac{\Sigma|x_i - \bar{x}|}{n}$

$= \frac{|3-5|+|4-5|+|5-5|+|6-5|+|7-5|}{5}$

$= \frac{2+1+0+1+2}{5} = \frac{6}{5} = 1.2$.

EXERCISE-2

(1) **3.** Mean of the given data is

$\bar{x} = \frac{4+7+8+9+10+13+17}{8} = \frac{80}{8} = 10$

x_i	$\lvert x_i - \bar{x}\rvert$
4	6
7	3
8	2
9	1
10	0
12	2
13	3
17	7
Total	24

M.D. about mean $= \frac{1}{n}\sum_{i=1}^{n}|x_i - \bar{x}| = \frac{1}{8}\times 24 = 3$

(2) **7.** Arranging the data in ascending order, we have

36, 42, 45, 46, 46, 49, 51, 53, 60, 72

Hence, n = 10 (which is even)

So, median is average of 5th and 6th observations

$\therefore$ Median $= \frac{46+49}{2} = \frac{95}{2} = 47.5$

x_i	$\lvert x_i - M\rvert$
36	11.5
42	5.5
45	2.5
46	1.5
46	1.5
49	1.5
51	3.5
53	5.5
60	12.5
72	24.5
Total	70

M.D. about median $= \frac{1}{n}\sum_{i=1}^{n}|x_i - M| = \frac{1}{10}\times 70 = 7$

(3) **16.**

x_i	f_i	$x_i f_i$	$\|x_i - 50\|$	$f_i\|x_i - 50\|$
10	4	40	40	160
30	24	720	20	480
50	28	1400	0	0
70	16	1120	20	320
90	8	720	40	320
	80	4000		1280

Mean $(\bar{x}) = \frac{1}{N}\sum f_i x_i = \frac{1}{80}\times 4000 = 50$

Mean deviation about mean

$$= \frac{1}{N}\sum_{i=1}^{n} f_i |x_i - \bar{x}| = \frac{1}{80}\times 1280 = 16$$

(4) **64.** Here $x = 6, 7, 10, 12, 13, 4, 8, 12$

$\Sigma x^2 = (6)^2 + (7)^2 + (10)^2 + (12)^2 + (13)^2 + (4)^2 + (8)^2 + (12)^2 = 722$

$$\text{Variance} = \sigma^2 = \frac{N\,\Sigma x^2 - (\Sigma x)^2}{N^2} = \frac{8\times 722 - (72)^2}{(8)^2}$$

$$= \frac{5776 - 5184}{64} = \frac{592}{64} = 9.25$$

(5) **40.** We have,

x_i	f_i	$f_i x_i$	$(x_i - \bar{x}) = (x_i - 19)$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
6	2	12	−13	169	338
10	4	40	−9	81	324
14	7	98	−5	25	175
18	12	216	−1	1	12
24	8	192	5	25	200
28	4	112	9	81	324
30	3	90	11	121	363
	$N = \Sigma f_i = 40$	$\Sigma f_i x_i = 760$			$\Sigma f_i(x_i - \bar{x})^2 = 1736$

Here, $N = 40$, $\Sigma f_i x_i = 760$ $\therefore\ \bar{x} = \frac{\Sigma f_i x_i}{N} = \frac{760}{40} = 19$

We have, $\bar{x} = \Sigma f_i (x_1 - \bar{x})^2 = 1736$

$\therefore\quad \sigma^2 = \frac{1}{N}\{\Sigma f_i (x_1 - \bar{x})^2\} = \frac{1736}{40} = 43.4$

(6) **11.** Let the assumed mean be 98

x_i	f_i	$d_i = x_i - 98$	d_i^2	$f_i d_i$	$f_i d_i^2$
92	3	$92 - 98 = -6$	36	−18	108
93	2	$93 - 98 = -5$	25	−10	50
97	3	$97 - 98 = -1$	1	−3	3
98	2	$98 - 98 = 0$	0	0	0
102	6	$102 - 98 = 4$	16	24	96
105	3	$104 - 98 = 6$	36	18	108
109	3	$109 - 98 = 11$	121	33	363
Total	22			44	728

$$\bar{x} = A + \frac{\Sigma f_i d_i}{\Sigma f_i} = 98 + \frac{\Sigma f_i d_i}{\Sigma f_i} = 98 + \frac{44}{22} = 98 + 2 = 100$$

Variance,

$$\sigma^2 = \frac{1}{N}\{\Sigma f_i d_i^2\} - \left\{\frac{1}{N}\Sigma f_i d_i\right\}^2 = \frac{1}{22}\{728\} - \left\{\frac{1}{22}\times 44\right\}^2$$

$$= \frac{1}{22}[728 - 22\times 2^2] = \frac{1}{22}\times 640 = \frac{320}{11} = 29.09$$

(7) **17.** We have,

x_i	f_i	$d_i = x_i - 64$	$f_i d_i$	d_i^2	$f_i d_i^2$
60	2	$60 - 64 = -4$	−8	16	32
61	1	$61 - 64 = -3$	−3	9	9
62	12	$62 - 64 = -2$	−24	4	48
63	29	$63 - 64 = -1$	−29	1	29
64	25	$64 - 65 = 0$	0	0	0
65	12	$65 - 64 = 1$	12	1	12
66	10	$66 - 64 = 2$	20	4	40
67	4	$67 - 64 = 3$	12	9	36
68	5	$68 - 64 = 4$	20	16	80
Total	100		0		286

Now, $\bar{x} = A + \frac{\Sigma f_i d_i}{\Sigma f_i} = 64 + \frac{0}{100} = 64$

Variance, $\sigma^2 = \frac{1}{N}\{\Sigma f_i d_i^2\} - \left\{\frac{1}{N}\Sigma f_i d_i\right\}^2$

$$= \left[\frac{1}{100}\times 286 - \left(\frac{1}{100}\times 0\right)^2\right] = 2.86$$

Standard deviation $= \sigma = \sqrt{2.86} = 1.69$

(8) **132.**

Classes	Mid values x_i	f_i	$u = \frac{x-25}{10}$	fu	fu^2
0-10	5	5	−2	−10	20
10-20	15	8	−1	−8	8
20-30	25	15	0	0	0
30-40	35	16	1	16	16
40-50	45	9	2	12	24
		50		10	68

Mean $(\bar{x}) = A + \frac{\Sigma fu}{N} \times h = 25 + \frac{10}{50} \times 10 = 25 + 2 = 27$

Variance $(\sigma^2) = \frac{h^2}{N^2}[N\Sigma fu^2 - (\Sigma fu)^2]$

$= \frac{(10)^2}{(50)^2}[50 \times 68 - (10)^2] = \frac{100}{2500}[3400 - 100]$

$= \frac{1}{25} \times 3300 = 132$

(9) **12.** Let the observations be x_1, x_2, x_3, x_4, x_5, and x_6, On multiplying each observation by 3, we get the new observations as $3x_1, 3x_2, 3x_3, 3x_4, 3x_5$ and $3x_6$.

Variance of new observations

$$= \frac{\sum_{i=1}^{6}(3x_i - 24)^2}{6} = \frac{3^2\sum_{i=1}^{6}(x_i - 8)^2}{6}$$

$= (9/1) \times$ Variance of old observations $= 9 \times 4^2 = 144$

Thus, standard deviation of new observations

$= \sqrt{\text{Variance}} = \sqrt{144} = 12$

EXERCISE-3

(1) **(A).** Obviously $G = \frac{G_1}{G_2}$

(2) **(A).** Median of new set remain same as that of the original set.

(3) **(B).** Mean of binomial distribution = np = 4

Variance of binomial distribution = npq = 2

from above q = 1/2

$P(X = r) = {}^nC_r\, q^{n-r} . p^r$; p + q = 1

$q = \frac{1}{2} \Rightarrow p = 1/2$; $n \times \frac{1}{2} = 4 \Rightarrow n = 8$

$P(X = 1) = {}^8C_1 . \left(\frac{1}{2}\right)^7\left(\frac{1}{2}\right)^1 = 8 \times \frac{1}{2^8} = \frac{1}{32}$

(4) **(B).** Given $N = 15,\ \Sigma x^2 = 2830\ \&\ \Sigma x = 170$

one observation 20 replaced by 30 then

$\Sigma x^2 = 2830 - 400 + 900 = 3330$

$\Sigma x = 170 - 20 + 30 = 180$

Variance,

$$\sigma^2 = \frac{\Sigma x^2}{N} - \left(\frac{\Sigma x}{N}\right)^2 = \frac{3330}{15} - \left(\frac{180}{15}\right)^2 = \frac{1170}{15} = 78$$

(5) **(C).** Only (a) and (b) are correct.

(6) **(C).** Mean of total 2n observation, $a - a = 0$

$$SD = \sqrt{\frac{\Sigma(x - \bar{x})^2}{N}},\ 2 = \sqrt{\frac{\Sigma x^2}{2n}}$$

$$4 = \sqrt{\frac{\Sigma x^2}{2n}} \Rightarrow \frac{2na^2}{2n} \Rightarrow a^2 = 4\ ;\ |a| = 2$$

(7) **(D).** Mode = 3 Median – 2 Mean

Mode = 3 (22) – 2 (21) = 66 – 42 = 24

(8) **(B).** AM of m^{th} power > m^{th} power of AM (If m > 1)

$$\frac{x_1^2 + x_2^2 + ...x_n^2}{n} > \left(\frac{x_1 + x_2 + ...x_n}{n}\right)^2$$

$$= \frac{\sum_{i=1}^{n} x_i^2}{n} > \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2 = \frac{400}{n} > \left(\frac{80}{n}\right)^2$$

$\frac{n^2}{n} > \frac{6400}{400}$; $\frac{n^2}{n} > 16$; n > 16

(9) **(D).** As no. of observations same in both cases. Also mean and difference is same hence $V_A/V_B = 1$

(10) **(C).** Taking $n_1 + n_2 = 100$. Substituting in the formula of combined mean we get the percentage of boys in the class as 80.

(11) **(C).** According to given conditions,

$$6.80 = \frac{(6-a)^2 + (6-b)^2 + (6-8)^2 + (6-5)^2 + (6-10)^2}{5}$$

$\Rightarrow 34 = (6-a)^2 + (6-b)^2 + (6-a)^2 + 4 + 1 + 16$

$\Rightarrow 13 = 9 + 4 = (6-a)^2 + (6-b)^2$

$3^2 + 2^2 = (6-a)^2 + (6-b)^2 = 3$

$6 - b = 2 \Rightarrow b = 4$

(12) Mean of given numbers,

$$\bar{x} = \frac{1+(1+d)+......(1+100d)}{101} = 1+50d$$

$$\text{Mean deviation} = \frac{\sum |x_i - \bar{x}|}{n} = \frac{\sum_{r=0}^{100} |(1+rd-(1+50d))|}{101}$$

$$\frac{\sum_{r=0}^{100} |(r-50d)|}{101} = \frac{d\ 50\times 51}{101}$$

Given that mean deviation is 255

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

(13) **(D).** Variance, $\sigma = \frac{\sum (x_i - \bar{x})^2}{n}$, where $\bar{x}$ is the mean.

Let 2, 4, 6, ... 2n be the numbers.

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

$$\sigma = \frac{\sum 4r^2 - 4n\,(n+1) + (n+1)^2}{n}$$

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

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

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

(14) **(A).** $\sigma_x^2 = 4, \sigma_y^2 = 4, \bar{x} = 2, \bar{y} = 4$

$$\frac{\Sigma x_i}{5} = 2 \;\; ; \;\; \Sigma x_i = 10, \Sigma y_i = 20$$

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

$$\Sigma x_i^2 = 40, \Sigma y_i^2 = 105$$

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

$$= \frac{1}{10}(40+105) - 9 = \frac{145-90}{10} = \frac{55}{10} = \frac{11}{2}$$

(15) **(C).** Median = 25.5 a

Mean deviation about median = 50

$$\Rightarrow \frac{\Sigma |x_i - 25.5a|}{50} = 50$$

$\Rightarrow 24.5\,a + 23.5a + ... + 0.5a + 0.5a + ... + 24.5a = 2500$

$\Rightarrow a + 3a + 5a + ... + 49a = 2500$

$$\Rightarrow \frac{25}{2}(50a) = 2500 \Rightarrow a = 4$$

(16) **(D).** A.M. of $2x_1, 2x_2.....2x_n$ is

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

So statement-2 is false

variance ($2x_i$) = 2^2 variance (x_i) = $4\sigma^2$

so statement-1 is true.

(17) **(D).** Variance is not changed by the change of origin.

If initially all marks were x_i then $\sigma_1^2 = \frac{\sum (x_i - \bar{x})^2}{N}$

Now each is increased by 10

$$\sigma_2^2 = \frac{\sum [(x_i+10)-(\bar{x}+10)]^2}{N} = \sigma_1^2$$

So variance will not change whereas mean, median and mode will increase by 10.

(18) **(B).** $\sigma^2 = \left(\frac{\Sigma x_i^2}{n}\right) - \bar{x}^2$

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

(19) **(C).** Mean = 16 ; Sum = 16 × 16 = 256

New sum = 256 – 16 + 3 + 4 + 5 = 252

Mean = 252/18 = 14

(20) **(A).** Variance

$$= \frac{\sum x_i^2}{n} - \left(\frac{\sum x_i^2}{n}\right)^2 = \frac{4+9+a^2+121}{4} - \left(\frac{16+a}{4}\right)^2$$

$= \frac{4(134+a^2)-256-a^2-32a}{16}$

$3a^2-32a+280 = 16\times\left(\frac{7}{2}\right)^2 = 4\times 49$

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

(21) **(A).** Variance $= \frac{45}{9}-(1)^2 = 5-1=4$

$\sigma = \sqrt{\text{Variance}} = 2$

(22) **(B).** Given $\bar{x} = \frac{\Sigma x_i}{5} = 150$; $\sum_{i=1}^{5} x_i = 750$ ...... (i)

$\frac{\Sigma x_i^2}{5}-(\bar{x})^2 = 18$; $\frac{\Sigma x_i^2}{5}-(150)^2 = 18$

$\Sigma x_i^2 = 112590$ ...... (ii)

Given height of new student, $x_6 = 156$

$\bar{x}_{new} = \frac{\sum_{i=1}^{6} x_i}{6} = \frac{750+156}{6} = 151$

Also, New variance $= \frac{\sum_{i=1}^{6} x_i^2}{6} - (\bar{x}_{new})^2$

$= \frac{112590+(156)^2}{6}-(151)^2 = 22821-22801 = 20.$

(23) **(C).** Let 7 observations be $x_1, x_2, x_3, x_4, x_5, x_6, x_7$

$\bar{x} = 8 \Rightarrow \sum_{i=1}^{7} x_i = 56$ ......(1)

Also, $\sigma^2 = 16$

$16 = \frac{1}{7}\left(\sum_{i=1}^{7} x_i^2\right)-(\bar{x})^2$; $16 = \frac{1}{7}\left(\sum_{i=1}^{7} x_i^2\right)-64$

$\left(\sum_{i=1}^{7} x_i^2\right) = 560$ ......(2)

Now, $x_1 = 2, x_2 = 4, x_3 = 10, x_4 = 12, x_5 = 14$

$\Rightarrow$ $x_6 + x_7 = 14$ (from (1))

$x_6{}^2 + x_7{}^2 = 100$ (from (2))

$\therefore$ $x_6{}^2 + x_7{}^2 = (x_6 + x_7)^2 - 2x_6x_7 \Rightarrow x_6x_7 = 48$

(24) **(A).** Let x be the 6^{th} observation

$\Rightarrow$ $45+54+41+57+43+x = 48\times 6 = 288 \Rightarrow x = 48$

Variance $= \left(\frac{\Sigma x_i^2}{6}-(\bar{x})^2\right) = \frac{14024}{6}-(48)^2 = \frac{100}{3}$

$\Rightarrow$ Standard deviation $= \frac{10}{\sqrt{3}}$

(25) **(B).** S.D $= \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$; $\bar{x} = \frac{\Sigma x}{4} = \frac{-1+0+1+k}{4} = \frac{k}{4}$

$\sqrt{5} = \sqrt{\frac{\left(-1-\frac{k}{4}\right)^2+\left(0-\frac{k}{4}\right)^2+\left(1-\frac{k}{4}\right)^2+\left(k-\frac{k}{4}\right)^2}{4}}$

$\Rightarrow$ $5\times 4 = 2\left(1+\frac{k}{16}\right)^2+\frac{5k^2}{8}$

$\Rightarrow$ $18 = \frac{3k^2}{4} \Rightarrow k^2 = 24 \Rightarrow k = 2\sqrt{6}$

(26) **(A).** $\frac{34+x}{2} = 35 \Rightarrow x = 36$

$42 = \frac{10+22+26+29+34+36+42+67+70+y}{10}$

$420-336 = y \Rightarrow y = 84$

$\frac{y}{x} = \frac{84}{36} = \frac{7}{3}$

(27) **(A).** $\Sigma f_i = 20 = 2x^2+2x-4 \Rightarrow x^2+2x-24 = 0$

$x = 3, -4$ (rejected)

$\bar{x} = \frac{\Sigma x_i f_i}{\Sigma f_i} = 2.8$

(28) **(D).** Mean $(\mu) = \frac{\sum x_i}{50} = 16$

Standard deviation $(\sigma) = \sqrt{\frac{\sum x_i}{50}-(\mu)^2} = 16$

$(256)\times 2 = \frac{\sum x_i^2}{50}$

New mean $= \frac{\Sigma(x_i-4)^2}{50} = \frac{\Sigma x_i^2+16\times 50-8\,\Sigma x_i}{50}$

$= (256)\times 2+16-8\times 16 = 400$

(29) **(B).** $x_1 + ... + x_4 = 44$

$x_5 + ... + x_{10} = 96$

$\bar{x} = 14, \ \Sigma x_i = 140$

$$\text{Variance} = \frac{\Sigma x_i^2}{n} - \bar{x}^2 = 4$$

Standard deviation = 2

(30) **18.** Var (1, 2, ..., n) = 10

$$\Rightarrow \frac{1^2 + 2^2 + ... + n^2}{n} - \left(\frac{1 + 2 + ... + n}{n}\right)^2 = 10$$

$$\Rightarrow \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = 10$$

$n^2 - 1 = 120 \Rightarrow n = 11$

Var (2, 4, 6, ..., 2m) = 16 $\Rightarrow$ Var (1, 2, ..., m) = 4

$\Rightarrow m^2 - 1 = 48 \Rightarrow m = 7 \Rightarrow m + n = 18$

(31) **54.** $\frac{3+7+9+12+13+20+x+y}{8} = 10$

$x + y = 16$

$$\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2 = 25$$

$3^2 + 7^2 + 9^2 + 12^2 + 13^2 + 20^2 + x^2 + y^2 = 1000$

$x^2 + y^2 = 148$

$xy = 54$

(32) **(B).** If each observation is multiplied with p & then q is subtracted. New mean $\bar{x}_1 = p\bar{x} - q$

$\Rightarrow 10 = p(20) - q$ ..... (1)

and new standard deviations

$\sigma_2 = |p|\sigma_1 \Rightarrow 1 = |p|(2) \Rightarrow |p| = 1/2$

$\Rightarrow p = \pm 1/2$

If p = 1/2 then q = 0 (from equation (1))

If p = –1/2 then q = – 20

(33) **(A).** $\frac{\Sigma x_i}{20} = 10$ ....(i)

$\frac{\Sigma x_i^2}{20} - 100 = 4$ ....(ii)

$\Sigma x_i^2 = 104 \times 20 = 2080$

$$\text{Actual mean} = \frac{200 - 9 + 11}{20} = \frac{202}{20}$$

$$\text{Variance} = \frac{2080 - 81 + 121}{20} - \left(\frac{202}{20}\right)^2$$

$$= \frac{2120}{20} - (10.1)^2 = 106 - 102.01 = 3.99$$

(34) **(D).** $\sum_{i=1}^{10}(x_i - 5) = 10$

Mean of observation $x_i - 5 = \frac{1}{10}\sum_{i=1}^{3}(x_i - 5) = 1$

$\Rightarrow \mu$ = mean of observation $(x_i - 3)$

= (mean of observation $(x_i - 5)$) + 2 = 1 + 2 = 3

Variance of observation

$$x_i - 5 = \frac{1}{10}\sum_{i=1}^{10}(x_i - 5)^2 - (\text{Mean of } (x_i - 5))^2 = 3$$

$\Rightarrow \lambda$ = variance of observation $(x_i - 3)$

= variance of observation $(x_i - 5) = 3$

$\therefore (\mu, \lambda) = (3, 3)$