CHAPTER – 11

MEASURES OF CENTRAL TENDENCY AND DISPERSION

LEARNING OBJECTIVES

After reading this Chapter , a student will be able to understand different measures of central tendency, i.e. Arithmetic Mean, Median, Mode, Geometric Mean and Harmonic Mean, and computational techniques of these measures.

They will also learn comparative advantages and disadvantages of these measures and therefore which measures to use in which circumstance.

However, to understand a set of observation, it is equally important to have knowledge of dispersion which indicates the volatility. In advanced stage of chartered accountancy course, volatility measures will be useful in understanding risk involved in financial decision making. This chapter will also guide the students to know details about various measures of dispersion.

11.1 DEFINITION OF CENTRAL TENDENCY

In many a case, like the distributions of height, weight, marks, profit, wage and so on, it has been noted that starting with rather low frequency, the class frequency gradually increases till it reaches its maximum somewhere near the central part of the distribution and after which the class frequency steadily falls to its minimum value towards the end. Thus, central tendency may be defined as the tendency of a given set of observations to cluster around a single central or middle value and the single value that represents the given set of observations is described as a measure of central tendency or, location, or average. Hence, it is possible to condense a vast mass of data by a single representative value. The computation of a measure of central tendency plays a very important part in many a sphere. A company is recognized by its high average profit, an educational institution is judged on the basis of average marks obtained by its students and so on. Furthermore, the central tendency also facilitates us in providing a basis for comparison between different distribution. Following are the different measures of central tendency:

(i) Arithmetic Mean (AM)

(ii) Median (Me)

(iii) Mode (Mo)

(iv) Geometric Mean (GM)

(v) Harmonic Mean (HM)

11.2 CRITERIA FOR AN IDEAL MEASURE OF CENTRAL TENDENCY

Following are the criteria for an ideal measure of central tendency:

(i) It should be properly and unambiguously defined.

(ii) It should be easy to comprehend.

(iii) It should be simple to compute.

(iv) It should be based on all the observations.

(v) It should have certain desirable mathematical properties.

(vi) It should be least affected by the presence of extreme observations.

11.3 ARITHMETIC MEAN

For a given set of observations, the AM may be defined as the sum of all the observations divided by the number of observations. Thus, if a variable x assumes n values $x_1, x_2, x_3, \ldots\ldots\ldots x_n$, then the AM of x, to be denoted by $\overline{x}$, is given by,

$$\overline{X} = \frac{x_1 + x_2 + x_3 + \ldots\ldots\ldots\ldots + x_n}{n}$$

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

$$\overline{X} = \frac{\sum x_i}{n} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11.1)$$

In case of a simple frequency distribution relating to an attribute, we have

$$\overline{x} = \frac{f_1x_1 + f_2x_2 + f_3x_3 + \ldots\ldots\ldots\ldots + f_nx_n}{f_1 + f_2 + f_3 + \ldots\ldots\ldots\ldots + f_n}$$

$$= \frac{\sum f_i x_i}{\sum f_i}$$

$$\overline{X} = \frac{\sum f_i x_i}{N} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11.2)$$

assuming the observation x_i occurs f_i times, i=1,2,3,……..n and N=≤f_i.
In case of grouped frequency distribution also we may use formula (11.2) with x_i as the mid value of the i-th class interval, on the assumption that all the values belonging to the i-th class interval are equal to x_i.
However, in most cases, if the classification is uniform, we consider the following formula for the computation of AM from grouped frequency distribution:

$$\overline{x} = A + \frac{\sum f_i d_i}{N} \times C \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11.3)$$

Where, $d_i = \frac{x_i - A}{C}$

A = Assumed Mean
C = Class Length

Illustrations

Example 11.1: Following are the daily wages in rupees of a sample of 9 workers: 58, 62, 48, 53, 70, 52, 60, 84, 75. Compute the mean wage.

Solution: Let x denote the daily wage in rupees.

Then as given, $x_1=58$, $x_2=62$, $x_3= 48$, $x_4=53$, $x_5=70$, $x_6=52$, $x_7=60$, $x_8=84$ and $x_9=75$.

Applying (11.1) the mean wage is given by,

$$\bar{x} = \frac{\sum_{i=1}^{9} x_i}{9}$$

$$= \text{Rs. } \frac{(58+62+48+53+70+52+60+84+75)}{9}$$

$$= \text{Rs. } \frac{562}{9}$$

$$= \text{Rs. } 62.44.$$

Example. 11.2: Compute the mean weight of a group of BBA students of St. Xavier's College from the following data :

Weight in kgs.	44 – 48	49 – 53	54 – 58	59 – 63	64 – 68	69 – 73
No. of Students	3	4	5	7	9	8

Solution: Computation of mean weight of 36 BBA students

Weight in kgs. (1)	No. of Student (f1) (2)	Mid-Value (x_i) (3)	$f_i x_i$ (4) = (2) x (3)
44 – 48	3	46	138
49 – 53	4	51	204
54 – 58	5	56	280
59 – 63	7	61	427
64 – 68	9	66	594
69 – 73	8	71	568
Total	36	–	2211

Applying (11.2), we get the average weight as

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{2211}{36} \text{ kgs.}$$

$$= 61.42 \text{ kgs.}$$

Example. 11.3: Find the AM for the following distribution:

Class Interval	350 – 369	370 – 389	390 – 409	410 – 429	430 – 449	450 – 469	470 – 489
Frequency	23	38	58	82	65	31	11

Solution: We apply formula (11.3) since the amount of computation involved in finding the AM is much more compared to **Example 11.2**. Any mid value can be taken as A. However, usually A is taken as the middle most mid-value for an odd number of class intervals and any one of the two middle most mid-values for an even number of class intervals. The class length is taken as C.

Table 11.2 Computation of AM

Class Interval (1)	Frequency(f_i) (2)	Mid-Value(x_i) (3)	$d_i = \frac{x_i - A}{c}$ $= \frac{x_i - 419.50}{20}$ (4)	$f_i d_i$ (5) = (2)X(4)
350 – 369	23	359.50	– 3	– 69
370 – 389	38	379.50	– 2	– 76
390 – 409	58	399.50	– 1	– 58
410 – 429	82	419.50 (A)	0	0
430 – 449	65	439.50	1	65
450 – 469	31	459.50	2	62
470 – 489	11	479.50	3	33
Total	308	–	–	– 43

The required AM is given by

$$\bar{x} = A + \frac{\sum f_i d_i}{N} \times C$$

$$= 419.50 + \frac{(-43)}{308} \times 20$$

$$= 419.50 - 2.79$$

$$= 416.71$$

Example. 11.4: Given that the mean height of a group of students is 67.45 inches. Find the missing frequencies for the following incomplete distribution of height of 100 students.

Height in inches	60 – 62	63 – 65	66 – 68	69 – 71	72 – 74
No. of Students	5	18	–	–	8

Solution : Let x denote the height and f_3 and f_4 as the two missing frequencies.

Table 11.3

Estimation of missing frequencies.

CI (1)	Frequency (f_i) (2)	Mid - Value (x_i) (3)	$\frac{x_i - 67}{3}$ (4)	$f_i d_i$ (5) = (2) x (4)
60-62	5	61	-2	-10
63 – 65	18	64	– 1	– 18
66 – 68	f_3	67 (A)	0	0
69 – 71	f_4	70	1	f_4
72 – 74	8	73	2	16
Total	$31 + f_3 + f_4$	–	–	$-12 + f_4$

As given, we have

$$31 + f_3 + f_4 = 100$$

$$\Rightarrow \quad f_3 + f_4 = 69 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

and

$$\bar{x} = 67.45$$

$$\Rightarrow \quad A + \frac{\sum f_i d_i}{N} \times C = 67.45$$

$$\Rightarrow \quad 67 + \frac{(-12 + f_4)}{100} \times 3 = 67.45$$

$$\Rightarrow \quad (-12 + f_4) \times 3 = (67.45 - 67) \times 100$$

$$\Rightarrow \quad -12 + f_4 = 15$$

$$\Rightarrow \quad f_4 = 27$$

On substituting 27 for f_4 in (1), we get

$$f_3 + 27 = 69 \quad \Rightarrow \quad f_3 = 42$$

Thus, the missing frequencies would be 42 and 27.

Properties of AM

(i) **If all the observations assumed by a variable are constants, say k, then the AM is also k.** For example, if the height of every student in a group of 10 students is 170 cm, then the mean height is, of course, 170 cm.

(ii) the algebraic sum of deviations of a set of observations from their AM is zero

i.e. for unclassified data , $\sum(x_i - \bar{x}) = 0$

and for grouped frequency distribution, $\sum f_i(x_i - \bar{x}) = 0$ ....(11.4)

For example, if a variable x assumes five observations, say 58,63,37,45,29, then =46.4. Hence, the deviations of the observations from the AM i.e. are 11.60, 16.60, –9.40, –1.40 and –17.40, then 11.60 + 16.60 + (–9.40) + (–1.40) + (–17.40) = 0 .

(iii) AM is affected due to a change of origin and/or scale which implies that if the original variable x is changed to another variable y by effecting a change of origin, say a, and scale say b, of x i.e. y=a+bx, then the AM of y is given by $\bar{y} = a + b\bar{x}$.

For example, if it is known that two variables x and y are related by 2x+3y+7=0 and $\bar{x} = 15$, then the AM of y is given by

$$= \frac{-7 - 2\times 15}{3} = \frac{-37}{3} = -12.33 .$$

(iv) If there are two groups containing n_1 and n_2 observations and $\bar{x}_1$ and $\bar{x}_2$ as the respective arithmetic means, then the combined AM is given by

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2} \quad \text{.....................................(11.5)}$$

This property could be extended to k>2 groups and we may write

$$\bar{x} = \frac{\sum n_i\bar{x}_i}{\sum n_i} \quad i = 1, 2, \ldots\ldots n. \quad \text{.....................................(11.6)}$$

Example 11.5 : The mean salary for a group of 40 female workers is Rs.5200 per month and that for a group of 60 male workers is Rs.6800 per month. What is the combined mean salary?

Solution : As given n_1 = 40, n_2 = 60, $\bar{x}_1$ = Rs.5200 and $\bar{x}_2$ = Rs.6800 hence, the combined mean salary per month is

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

$$= \frac{40\times \text{Rs. } 5200 + 60\times \text{Rs. } 6800}{40 + 60} = \text{Rs.}6160.$$

11.4 MEDIAN – PARTITION VALUES

As compared to AM, median is a positional average which means that the value of the median is dependent upon the position of the given set of observations for which the median is wanted. Median, for a given set of observations, may be defined as the middle-most value when the observations are arranged either in an ascending order or a descending order of magnitude.

As for example, if the marks of the 7 students are 72, 85,56,80,65,52 and 68, then in order to find the median mark, we arrange these observations in the following ascending order of magnitude: 52, 56, 65, 68, 72, 80, 85.

Since the 4^{th} term i.e. 68 in this new arrangement is the middle most value, the median mark is 68 i.e. Median (Me) = 68.

As a second example, if the wages of 8 workers, expressed in rupees are

56, 82, 96, 120, 110, 82, 106, 100 then arranging the wages as before, in an ascending order of magnitude, we get Rs.56, Rs.82, Rs.82, Rs.96, Rs.100, Rs.106, Rs.110, Rs.120. Since there are two middle-most values, namely, Rs.96, and Rs.100 any value between Rs.96 and Rs.100 may be, theoretically, regarded as median wage. However, to bring uniqueness, we take the arithmetic mean of the two middle-most values, whenever the number of the observations is an even number. Thus, the median wage in this example, would be

$$M = \frac{\text{Rs. } 96 + \text{Rs.} 100}{2} = \text{Rs. } 98$$

In case of a grouped frequency distribution, we find median from the cumulative frequency distribution of the variable under consideration. We may consider the following formula, which can be derived from the basic definition of median.

$$M = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l} \right) \times C \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (11.7)$$

Where,

l_1 = lower class boundary of the median class i.e. the class containing median.

N = total frequency.

N_l = less than cumulative frequency corresponding to l_1. (Pre median class)

N_u = less than cumulative frequency corresponding to l_2. (Post median class)

l_2 being the upper class boundary of the median class.

C = $l_2 - l_1$ = length of the median class.

Example 11.6 : Compute the median for the distribution as given in **Example 11.3.**

Solution: First, we find the cumulative frequency distribution which is exhibited in **Table 11.4.**

Table 11.4

Computation of Median

Class boundary	Less than cumulative frequency
349.50	0
369.50	23
389.50	61
409.50 (l_1)	119 (N_l)
429.50 (l_2)	201(N_u)
449.50	266
469.50	297
489.50	308

We find, from the **Table 11.4**, $\frac{N}{2} = \frac{308}{2}$ = 154 lies between the two cumulative frequencies 119 and 201 i.e. 119 < 154 < 201 . Thus, we have N_l = 119, N_u = 201 l_1 = 409.50 and l_2 = 429.50. Hence C = 429.50 – 409.50 =20.
Substituting these values in (11.7), we get,

$$M = 409.50 + \frac{154-119}{201-119} \times 20$$

$$= 409.50+8.54$$

$$= 418.04.$$

Example 11.7: Find the missing frequency from the following data, given that the median mark is 23.

Mark :	0 – 10	10 – 20	20 – 30	30 – 40	40 – 50
No. of students :	5	8	?	6	3

Solution : Let us denote the missing frequency by f_3. Table 11.5 shows the relevant computation.

Table 11.5
(Estimation of missing frequency)

Mark	Less than cumulative frequency
0	0
10	5
20(l_1)	13(N_l)
30(l_2)	13+f_3(N_u)
40	19+f_3
50	22+f_3

Going through the mark column , we find that 20<23<30. Hence l_1=20, l_2 =30 and accordingly N_l=13, N_u=13+f_3. Also the total frequency i.e. N is 22+f_3. Thus,

$$M = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l}\right) \times C$$

$$\Rightarrow \quad 23 = 20 + \frac{\left(\frac{22+f_3}{2}\right) - 13}{(13+f_3) - 13} \times 10$$

$$\Rightarrow \quad 3 = \frac{22+f_3-26}{f_3} \times 5$$

$$\Rightarrow \quad 3f_3 = 5f_3 - 20$$

$$\Rightarrow \quad 2f_3 = 20$$

$$\Rightarrow \quad f_3 = 10$$

So, the missing frequency is 10.

Properties of median

We cannot treat median mathematically, the way we can do with arithmetic mean. We consider below two important features of median.

(i) If x and y are two variables, to be related by y=a+bx for any two constants a and b, then the median of y is given by
$y_{me} = a + bx_{me}$

For example, if the relationship between x and y is given by 2x – 5y = 10 and if x_{me} i.e. the median of x is known to be 16.
Then 2x – 5y = 10

$\Rightarrow \quad y = -2 + 0.40x$

$\Rightarrow \quad y_{me} = -2 + 0.40\, x_{me}$

$\Rightarrow \quad y_{me} = -2 + 0.40 \times 16$

$\Rightarrow \quad y_{me} = 4.40.$

(ii) For a set of observations, the sum of absolute deviations is minimum when the deviations are taken from the median. This property states that $\sum|x_i - A|$ is minimum if we choose A as the median.

PARTITION VALUES OR QUARTILES OR FRACTILES

These may be defined as values dividing a given set of observations into a number of equal parts. When we want to divide the given set of observations into two equal parts, we consider median. Similarly, quartiles are values dividing a given set of observations into four equal parts. So there are three quartiles – first quartile or lower quartile denoted by Q_1, second quartile or median to be denoted by Q_2 or Me and third quartile or upper quartile denoted by Q_3. First quartile is the value for which one fourth of the observations are less than or equal to Q_1 and the remaining three – fourths observations are more than or equal to Q_1. In a similar manner, we may define Q_2 and Q_3.

Deciles are the values dividing a given set of observation into ten equal parts. Thus, there are nine deciles to be denoted by D_1, D_2, D_3,…..D_9. D_1 is the value for which one-tenth of the given observations are less than or equal to D_1 and the remaining nine-tenth observations are greater than or equal to D_1 when the observations are arranged in an ascending order of magnitude.

Lastly, we talk about the percentiles or centiles that divide a given set of observations into 100 equal parts. The points of sub-divisions being P_1, P_2,………..P_{99}. P_1 is the value for which one hundredth of the observations are less than or equal to P_1 and the remaining ninety-nine hundredths observations are greater than or equal to P_1 once the observations are arranged in an ascending order of magnitude.

For unclassified data, the p^{th} quartile is given by the $(n+1)p^{th}$ value, where n denotes the total number of observations. p = 1/4, 2/4, 3/4 for Q_1, Q_2 and Q_3 respectively. p=1/10, 2/10,…………9/10. For D_1, D_2,……,D_9 respectively and lastly p=1/100, 2/100,….,99/100 for P_1, P_2, P_3….P_{99} respectively.

In case of a grouped frequency distribution, we consider the following formula for the computation of quartiles.

$$Q = l_1 + \left(\frac{Np - N_l}{N_u - N_l}\right) \times C \quad \text{.. (11.8)}$$

The symbols, except p, have their usual interpretation which we have already discussed while computing median and just like the unclassified data, we assign different values to p depending on the quartile.

Another way to find quartiles for a grouped frequency distribution is to draw the ogive (less than type) for the given distribution. In order to find a particular quartile, we draw a line parallel to the horizontal axis through the point Np. We draw perpendicular from the point of intersection of this parallel line and the ogive. The x-value of this perpendicular line gives us the value of the quartile under discussion.

Example 11.8: Following are the wages of the labourers: Rs.82, Rs.56, Rs.90, Rs.50, Rs.120, Rs.75, Rs.75, Rs.80, Rs.130, Rs.65. Find Q_1, D_6 and P_{82}.

Solution: Arranging the wages in an ascending order, we get Rs.50, Rs.56, Rs.65, Rs.75, Rs.75, Rs.80, Rs.82, Rs.90, Rs.120, Rs.130.
Hence, we have

$$Q_1 = \frac{(n+1)}{4}\text{th value}$$

$$= \frac{(10+1)}{4}\text{th value}$$

$$= 2.75^{th}\text{ value}$$

$$= 2^{nd}\text{ value} + 0.75 \times \text{difference between the third and the } 2^{nd}\text{ values.}$$

$$= \text{Rs. } [56 + 0.75 \times (65 - 56)]$$

$$= \text{Rs. } 62.75$$

$$D_6 = (10 + 1) \times \frac{6}{10}\text{ th value}$$

$$= 6.60^{th}\text{ value}$$

$$= 6^{th}\text{ value} + 0.60 \times \text{difference between the } 7^{th}\text{ and the } 6^{th}\text{ values.}$$

$$= \text{Rs. } (80 + 0.60 \times 2)$$

$$= \text{Rs. } 81.20$$

$$P_{82} = (10+1) \times \frac{82}{100}\text{ th value}$$

$$= 9.02^{th}\text{ value}$$

$$= 9^{th}\text{ value} + 0.02 \times \text{difference between the } 10^{th}\text{ and the } 9^{th}\text{ values}$$

$$= \text{Rs. } (120 + 0.02 \times 10)$$

$$= \text{Rs.}120.20$$

Next, let us consider one problem relating to the grouped frequency distribution.

Example 11.9: Following distribution relates to the distribution of monthly wages of 100 workers.

Wages in Rs. :	less than 500	500–699	700–899	900–1099	1100–1499	more than 1500
No. of workers :	5	23	29	27	10	6

Compute Q_3 , D_7 and P_{23} .

Solution: This is a typical example of an open end unequal classification as we find the lower class limit of the first class interval and the upper class limit of the last class interval are not stated, and theoretically, they can assume any value between 0 and 500 and 1500 to any number respectively. The ideal measure of the central tendency in such a situation is median as the median or second quartile is based on the fifty percent central values. Denoting the first LCB and the last UCB by the L and U respectively, we construct the following cumulative frequency distribution:

Table 11.7
Computation of quartiles

Wages in rupees (CB)	No. of workers (less than cumulative frequency)
L	0
499.50	5
699.50	28
899.50	57
1099.50	84
1499.50	94
U	100

For Q_3, $\frac{3N}{4} = \frac{3 \times 100}{4} = 75$

since, 57<75 <84, we take N_l = 57, N_u=84, l_1=899.50, l_2=1099.50, c = $l_2 – l_1$ = 200 in the formula (11.8) for computing Q_3.

Therefore, Q_3 = Rs. $\left[899.50 + \frac{75 - 57}{84 - 57} \times 200\right]$ =Rs.1032.83

Similarly, for D_7, $\frac{7N}{10} = \frac{7 \times 100}{10} = 70$ which also lies between 57 and 84.

Thus, $D_7 = \text{Rs.}\left[899.50 + \frac{70 - 57}{84 - 57} \times 200\right]$ = Rs.995.80

Lastly for P_{23}, $\frac{23N}{100} = \frac{23}{100}$ × 100 = 23 and as 5 < 23 < 28, we have

P_{23} = Rs. [499.50 + $\frac{23 - 5}{28 - 5}$ × 200]

= Rs. 656.02

11.5 MODE

For a given set of observations, mode may be defined as the value that occurs the maximum number of times. Thus, mode is that value which has the maximum concentration of the observations around it. This can also be described as the most common value with which, even, a layman may be familiar with.

Thus, if the observations are 5, 3, 8, 9, 5 and 6, then Mode (Mo) = 5 as it occurs twice and all the other observations occur just once. The definition for mode also leaves scope for more than one mode. Thus sometimes we may come across a distribution having more than one mode. Such a distribution is known as a multi-modal distribution. Bi-modal distribution is one having two modes.

Furthermore, it also appears from the definition that mode is not always defined. As an example, if the marks of 5 students are 50, 60, 35, 40, 56, there is no modal mark as all the observations occur once i.e. the same number of times.

We may consider the following formula for computing mode from a grouped frequency distribution:

$$\text{Mode} = l_1 + \left(\frac{f_0 - f_{-1}}{2f_0 - f_{-1} - f_1}\right) \times c \qquad \text{..........................(11.9)}$$

where,

l_1 = LCB of the modal class.
i.e. the class containing mode.
f_0 = frequency of the modal class
f_{-1} = frequency of the pre – modal class
f_1 = frequency of the post modal class
C = class length of the modal class

Example 11.10: Compute mode for the distribution as described in Example. 11.3

Solution : The frequency distribution is shown below

Table 11.8
Computation of mode

Class Interval	Frequency
350 - 369	23
370 - 389	38
390 - 409	58 (f_{-1})
410 - 429	82 (f_0)
430 - 449	65 (f_1)
450 - 469	31
470 - 489	11

Going through the frequency column, we note that the highest frequency i.e. f_0 is 82. Hence, f_{-1}

= 58 and f_1 = 65. Also the modal class i.e. the class against the highest frequency is 410 – 429. Thus l_1 = LCB=409.50 and c=429.50 – 409.50 = 20

Hence, applying formulas (11.9), we get

$$Mo = 409.5 + \frac{82-58}{2\times 82-58-65}\times 20$$

= 421.21 which belongs to the modal class. (410 – 429)

When it is difficult to compute mode from a grouped frequency distribution, we may consider the following empirical relationship between mean, median and mode:

Mean – Mode = 3(Mean – Median) …………………….(11.9A)

or Mode = 3 Median – 2 Mean

(11.9A) holds for a moderately skewed distribution. We also note that if y = a+bx, then $y_{mo}=a+bx_{mo}$ ……………………………………(11.10)

Example 11.11: For a moderately skewed distribution of marks in statistics for a group of 200 students, the mean mark and median mark were found to be 55.60 and 52.40. What is the modal mark?

Solution: Since in this case, mean = 55.60 and median = 52.40, applying (11.9A), we get the modal mark as

Mode = 3 × Median – 2 × Mean
= 3 × 52.40 – 2 × 55.60
= 46.

Example 11.12: If y = 2 + 1.50x and mode of x is 15, what is the mode of y?

Solution:

By virtue of (11.10), we have

y_{mo} = 2 + 1.50 × 15
= 24.50.

11.6 GEOMETRIC MEAN AND HARMONIC MEAN

For a given set of n positive observations, the geometric mean is defined as the n-th root of the product of the observations. Thus if a variable x assumes n values $x_1, x_2, x_3, \ldots, x_n$, all the values being positive, then the GM of x is given by

$G= (x_1 \times x_2 \times x_3 \ldots \times x_n)^{1/n}$ .... (11.11)

For a grouped frequency distribution, the GM is given by

$G= (x_1^{f_1} \times x_2^{f_2} \times x_3^{f_3} \ldots \times x_n^{f_n})^{1/N}$ .... (11.12)

Where $N = \sum f_i$

In connection with GM, we may note the following properties :

(i) Logarithm of G for a set of observations is the Am of the logarithm of the observations; i.e.

$$\log G = \frac{1}{r} \Sigma \log x \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11.13)$$

(ii) if all the observations assumed by a variable are constants, say K > 0, then the GM of the observations is also K.

(iii) GM of the product of two variables is the product of their GM's i.e. if $z = xy$, then

GM of z = (GM of x) × (GM of y) ……………………………………..(11.14)

(iv) GM of the ratio of two variables is the ratio of the GM's of the two variables i.e. if $z = x/y$ then

$$\text{GM of } z = \frac{\text{GM of x}}{\text{GM of y}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11.15)$$

Example 11.13: Find the GM of 3, 6 and 12.

Solution: As given $x_1=3$, $x_2=6$, $x_3=12$ and n=3.

Applying (11.11), we have $G= (3\times6\times12)^{1/3} = (6^3)^{1/3}=6$.

Example. 11.14: Find the GM for the following distribution:

x :	2	4	8	16
f :	2	3	3	2

Solution : According to (11.12) , the GM is given by

$$G = (x_1^{f_1} \times x_2^{f_2} \times x_3^{f_3} \times x_4^{f_4})^{1/N}$$

$$= (2^2 \times 4^3 \times 8^3 \times 16^2)^{1/10}$$

$$= (2)^{2.50}$$

$$= 4\sqrt{2}$$

$$= 5.66$$

Harmonic Mean

For a given set of non-zero observations, harmonic mean is defined as the reciprocal of the AM of the reciprocals of the observation. So, if a variable x assumes n non-zero values x_1, x_2, x_3,……………,x_n, then the HM of x is given by

$$H=\frac{n}{\sum(1/x_i)}$$

For a grouped frequency distribution, we have

$$H = \frac{N}{\sum\left[\frac{f_i}{x_i}\right]}$$

Properties of HM

(i) If all the observations taken by a variable are constants, say k, then the HM of the observations is also k.

(ii) If there are two groups with n_1 and n_2 observations and H_1 and H_2 as respective HM's than the combined HM is given by

$$\frac{n_1 + n_2}{\frac{n_1}{H_1} + \frac{n_2}{H_2}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots(11.18)$$

Example 11.15: Find the HM for 4, 6 and 10.

Solution: Applying (11.16), we have

$$H = \frac{3}{\frac{1}{4} + \frac{1}{6} + \frac{1}{10}}$$

$$= \frac{3}{0.25 + 0.17 + 0.10}$$

$$= 5.77$$

Example 11.16: Find the HM for the following data:

X:	2	4	8	16
f:	2	3	3	2

Solution: Using (11.17), we get

$$H = \frac{10}{\frac{2}{2} + \frac{3}{4} + \frac{3}{8} + \frac{2}{16}}$$

$$= 4.44$$

Relation between AM, GM, and HM

For any set of positive observations, we have the following inequality:

AM ≥ GM ≥ HM …………………….. (11.19)

The equality sign occurs, as we have already seen, when all the observations are equal.

Example 11.17: compute AM, GM, and HM for the numbers 6, 8, 12, 36.

Solution: In accordance with the definition, we have

$$AM = \frac{6+8+12+36}{4} = 15.5\ 0$$

$$GM = (6 \times 8 \times 12 \times 36)^{1/4}$$
$$= (2^8 \times 3^4)^{1/4} = 12$$

$$HM = \frac{4}{\frac{1}{6}+\frac{1}{8}+\frac{1}{12}+\frac{1}{36}} = 9.93$$

The computed values of AM, GM, and HM establish (11.19).

Weighted average

When the observations under consideration have a hierarchical order of importance, we take recourse to computing weighted average, which could be either weighted AM or weighted GM or weighted HM.

$$\text{Weighted AM} = \frac{\sum w_i x_i}{\sum w_i} \quad \ldots\ldots\ldots\ldots(11.20)$$

$$\text{Weighted GM} = \text{Ante log}\left(\frac{\sum w_i \log x_i}{\sum w_i}\right) \quad \ldots\ldots\ldots\ldots(11.21)$$

$$\text{Weighted HM} = \frac{\sum w_i}{\sum\left(\frac{w_i}{x_i}\right)} \quad \ldots\ldots\ldots\ldots(11.22)$$

Example 11.18: Find the weighted AM and weighted HM of first n natural numbers, the weights being equal to the squares of the corresponding numbers.

Solution: As given,

x	1	2	3	….	n
w	1^2	2^2	3^2	….	n^2

$$\text{Weighted AM} = \frac{\sum w_i x_i}{\sum w_i}$$

$$= \frac{1\times1^2 + 2\times2^2 + 3\times3^2 + ...n\times n^2}{1^2 + 2^2 + 3^2 + ... + n^2}$$

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

$$= \frac{\left[\frac{n(n+1)}{2}\right]^2}{\frac{n(n+1)(2n+1)}{6}}$$

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

$$\text{Weighted HM} = \frac{\sum w_i}{\sum\left(\frac{w_i}{x_i}\right)}$$

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

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

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

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

A General review of the different measures of central tendency

After discussing the different measures of central tendency, now we are in a position to have a review of these measures of central tendency so far as the relative merits and demerits are concerned on the basis of the requisites of an ideal measure of central tendency which we have already mentioned in section 11.2. The best measure of central tendency, usually, is the AM. It is rigidly defined, based on all the observations, easy to comprehend, simple to calculate and amenable to mathematical properties. However, AM has one drawback in the sense that it is very much affected by sampling fluctuations. In case of frequency distribution, mean cannot be advocated for open-end classification.

Like AM, median is also rigidly defined and easy to comprehend and compute. But median is not based on all the observation and does not allow itself to mathematical treatment. However, median is not much affected by sampling fluctuation and it is the most appropriate measure of central tendency for an open-end classification.

Although mode is the most popular measure of central tendency, there are cases when mode remains undefined. Unlike mean, it has no mathematical property. Mode is also affected by sampling fluctuations.

GM and HM, like AM, possess some mathematical properties. They are rigidly defined and based on all the observations. But they are difficult to comprehend and compute and, as such, have limited applications for the computation of average rates and ratios and such like things.

Example 11.19 : Given two positive numbers a and b, prove that **$AH=G^2$**. Does the result hold for any set of observations?

Solution: For two positive numbers a and b, we have,

$$A = \frac{a+b}{2}$$

$$G = \sqrt{ab}$$

And $$H = \frac{2}{\frac{1}{a}+\frac{1}{b}}$$

$$= \frac{2ab}{a+b}$$

Thus $$AH = \frac{a+b}{2} \times \frac{2ab}{a+b}$$

$$= ab = G^2$$

This result holds for only two positive observations and not for any set of observations.

Example 11.20: The AM and GM for two observations are 5 and 4 respectively. Find the two observations.

Solution: If a and b are two positive observations then as given

$$\frac{a+b}{2} = 5$$

$\Rightarrow$ $a+b = 10$ ....(1)

and $\sqrt{ab} = 4$

$\Rightarrow$ $ab = 16$ ....(2)

$$\therefore (a-b)^2 = (a+b)^2 - 4ab$$

$$= 10^2 - 4\times16$$

$$= 36$$

$\Rightarrow \quad a - b = 6$ (ignoring the negative sign)……………………….(3)

Adding (1) and (3) We get,

$$2a = 16$$

$$\Rightarrow \quad a = 8$$

From (1), we get $b = 10 - a = 2$

Thus, the two observations are 8 and 2.

Example 11.21: Find the mean and median from the following data:

Marks	:	less than 10	less than 20	less than 30
No. of Students	:	5	13	23
Marks	:	less than 40	less than 50	
No. of Students	:	27	30	

Also compute the mode using the approximate relationship between mean, median and mode.

Solution: What we are given in this problem is less than cumulative frequency distribution. We need to convert this cumulative frequency distribution to the corresponding frequency distribution and thereby compute the mean and median.

Table 11.9

Computation of Mean Marks for 30 students

Marks Class Interval (1)	No. of Students (f_i) (2)	Mid - Value (x_i) (3)	$f_i x_i$ (4)= (2)×(3)
0 – 10	5	5	25
10 – 20	13 – 5 = 8	15	120
20 – 30	23 – 13 = 10	25	250
30 – 40	27 – 23 = 4	35	140
40 – 50	30 – 27 = 3	45	135
Total	30	–	670

Hence the mean mark is given by

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{670}{30}$$

$$= 22.33$$

Table 11.10

Computation of Median Marks

Marks (Class Boundary)	No.of Students (Less than cumulative Frequency)
0	0
10	5
20	13
30	23
40	27
50	30

Since $\frac{N}{2} = \frac{30}{2} = 15$ lies between 13 and 23,
we have l_1 = 20, N_l = 13, N_u= 23
and C = $l_2 - l_1$ = 30 – 20 = 10
Thus,

$$\text{Median} = 20 + \frac{15-13}{23-13} \times 10$$

$$= 22$$

Since Mode = 3 Median – 2 Mean approximately, we find that

Mode = 3x22 – 2x22.33

$= 21.34$

Example 11.22: Following are the salaries of 20 workers of a firm expressed in thousand rupees: 5, 17, 12, 23, 7, 15, 4, 18, 10, 6, 15, 9, 8, 13, 12, 2, 12, 3, 15, 14. The firm gave bonus amounting to Rs. 2000, Rs. 3000, Rs. 4000, Rs.5000 and Rs. 6000 to the workers belonging to the salary groups 1000 – 5000, 6000 – 10000 and so on and lastly 21000 – 25000. Find the average bonus paid per employee.

Solution: We first construct frequency distribution of salaries paid to the 20 employees. The average bonus paid per employee is given by $\frac{\sum f_i x_i}{N}$ Where x_i represents the amount of bonus paid to the i^{th} salary group and f_i, the number of employees belonging to that group which would be obtained on the basis of frequency distribution of salaries.

Table 11.11
Computation of Average bonus

Salary in thousand Rs. (Class Interval) (1)	Tally Mark (2)	No of workers (f_i) (3)	Bonus in Rupees x_i (4)	$f_i x_i$ (5) = (3) × (4)
1-5	||||	4	2000	8000
6-10	~~||||~~	5	3000	15000
11-15	~~||||~~ |||	8	4000	32000
16-20	||	2	5000	10000
21-25	|	1	6000	6000
TOTAL	–	20	–	71000

Hence, the average bonus paid per employee

$$= \text{Rs.}\frac{71000}{20}$$

Rs. = 3550

11.7 EXERCISE

Set A

Write down the correct answers. Each question carries 1 mark.

1. Measures of central tendency for a given set of observations measures
   (i) The scatterness of the observations (ii) The central location of the observations
   (iii) Both (i) and (ii) (iv) None of these.
2. While computing the AM from a grouped frequency distribution, we assume that
   (i) The classes are of equal length (ii) The classes have equal frequency
   (iii) All the values of a class are equal to the mid-value of that class
   (iv) None of these.
3. Which of the following statements is wrong?
   (i) Mean is rigidly defined
   (ii) Mean is not affected due to sampling fluctuations

(iii) Mean has some mathematical properties

(iv) All these

4. Which of the following statements is true?

(i) Usually mean is the best measure of central tendency

(ii) Usually median is the best measure of central tendency

(iii) Usually mode is the best measure of central tendency

(iv) Normally, GM is the best measure of central tendency

5. For open-end classification, which of the following is the best measure of central tendency?

(i) AM (ii) GM (iii) Median (iv) Mode

6. The presence of extreme observations does not affect

(i) AM (ii) Median (iii) Mode (iv) Any of these.

7. In case of an even number of observations which of the following is median ?

(i) Any of the two middle-most value

(ii) The simple average of these two middle values

(iii) The weighted average of these two middle values

(iv) Any of these

8. The most commonly used measure of central tendency is

(i) AM (ii) Median (iii) Mode (iv) Both GM and HM.

9. Which one of the following is not uniquely defined?

(i) Mean (ii) Median (iii) Mode (iv) All of these measures

10. Which of the following measure of the central tendency is difficult to compute?

(i) Mean (ii) Median (iii) Mode (iv) GM

11. Which measure(s) of central tendency is(are) considered for finding the average rates?

(i) AM (ii) GM (iii) HM (iv) Both (ii) and(iii)

12. For a moderately skewed distribution, which of he following relationship holds?

(i) Mean – Mode = 3 (Mean – Median) (ii) Median – Mode = 3 (Mean – Median)

(iii) Mean – Median = 3 (Mean – Mode) (iv) Mean – Median = 3 (Median – Mode)

13. Weighted averages are considered when

(i) The data are not classified

(ii) The data are put in the form of grouped frequency distribution

(iii) All the observations are not of equal importance

(iv) Both (i) and (iii).

14. Which of the following results hold for a set of distinct positive observations?

(i) AM ≥ GM ≥ HM (ii) HM ≥ GM ≥ AM

(iii) AM > GM > HM (iv) GM > AM > HM

15. When a firm registers both profits and losses, which of the following measure of central tendency cannot be considered?

(i) AM (ii) GM (iii) Median (iv) Mode

16. Quartiles are the values dividing a given set of observations into

(i) Two equal parts (ii) Four equal parts (iii) Five equal parts (iv) None of these.

17. Quartiles can be determined graphically using

(i) Histogram (ii) Frequency Polygon (iii) Ogive (iv) Pie chart.

18. Which of the following measure(s) possesses (possess) mathematical properties?

(i) AM (ii) GM (iii) HM (iv) All of these

19. Which of the following measure(s) satisfies (satisfy) a linear relationship between two variables?

(i) Mean (ii) Median (iii) Mode (iv) All of these

20. Which of he following measures of central tendency is based on only fifty percent of the central values?

(i) Mean (ii) Median (iii) Mode (iv) Both (i) and(ii)

Set B

Write down the correct answers. Each question carries 2 marks.

1. If there are 3 observations 15, 20, 25 then the sum of deviation of the observations from their AM is

(i) 0 (ii) 5 (iii) –5 (iv) None of these.

2. What is the median for the following observations?

5, 8, 6, 9, 11, 4.

(i) 6 (ii) 7 (iii) 8 (iv) None of these

3. What is the modal value for the numbers 5, 8, 6, 4, 10, 15, 18, 10?

(i) 18 (ii) 10 (iii) 14 (iv) None of these

4. What is the GM for the numbers 8, 24 and 40?

(i) 24 (ii) 12 (iii) $8 \cdot \sqrt[3]{15}$ (iv) 10

5. The harmonic mean for the numbers 2, 3, 5 is

(i) 2.00 (ii) 3.33 (iii) 2.90 (iv) $-\sqrt[3]{30}$.

6. If the AM and GM for two numbers are 6.50 and 6 respectively then the two numbers are

   (i) 6 and 7 (ii) 9 and 4 (iii) 10 and 3 (iii) 8 and 5.

7. If the AM and HM for two numbers are 5 and 3.2 respectively then the GM will be

   (i) 16.00 (ii) 4.10 (iii) 4.05 (iv) 4.00.

8. What is the value of the first quartile for observations 15, 18, 10, 20, 23, 28, 12, 16?

   (i) 17 (ii) 16 (iii) 12.75 (iv) 12

9. The third decile for the numbers 15, 10, 20, 25, 18, 11, 9, 12 is

   (i) 13 (ii) 10.70 (iii) 11 (iv) 11.50

10. If there are two groups containing 30 and 20 observations and having 50 and 60 as arithmetic means, then the combined arithmetic mean is

    (i) 55 (ii) 56 (iii) 54 (iv) 52.

11. The average salary of a group of unskilled workers is Rs.10000 and that of a group of skilled workers is Rs.15,000. If the combined salary is Rs.12000, then what is the percentage of skilled workers?

    (i) 40% (ii) 50% (iii) 60% (iv) none of these

12. If there are two groups with 75 and 65 as harmonic means and containing 15 and 13 observation then the combined HM is given by

    (i) 65 (ii) 70.36 (iii) 70 (iv) 71.

13. What is the HM of 1,½, 1/3,…………….1/n?

    (i) n (ii) 2n (iii) $\frac{2}{(n+1)}$ (iv) $\frac{n(n+1)}{2}$

14. An aeroplane flies from A to B at the rate of 500 km/hour and comes back from B to A at the rate of 700 km/hour. The average speed of the aeroplane is

    (i) 600 km. per hour (ii) 583.33 km. per hour

    (iii) $100\sqrt{35}$ km. per hour (iv) 620 km. per hour.

15. If a variable assumes the values 1, 2, 3…5 with frequencies as 1, 2, 3…5, then what is the AM?

    (i) $\frac{11}{3}$ (ii) 5 (iii) 4 (iv) 4.50

16. Two variables x and y are given by y= 2x – 3. If the median of x is 20, what is the median of y?

    (i) 20 (ii) 40 (iii) 37 (iv) 35

17. If the relationship between two variables u and v are given by 2u + v + 7 = 0 and if the AM of u is 10, then the AM of v is

(i) 17 (ii) –17 (iii) –27 (iv) 27.

18. If x and y are related by x–y–10 = 0 and mode of x is known to be 23, then the mode of y is

(i) 20 (ii) 13 (iii) 3 (iv) 23.

19. If GM of x is 10 and GM of y is 15, then the GM of xy is

(i) 150 (ii) Log 10 × Log 15 (iii) Log 150 (iv) None of these.

20. If the AM and GM for 10 observations are both 15, then the value of HM is

(i) Less than 15 (ii) More than 15 (iii) 15 (iv) Can not be determined.

Set C

Write down the correct answers. Each question carries 5 marks.

1. What is the value of mean and median for the following data:

Marks :	5–14	15–24	25–34	35–44	45–54	55–64
No. of Student :	10	18	32	26	14	10

(i) 30 and 28 (ii) 29 and 30 (iii) 33.68 and 32.94 (iv) 34.21 and 33.18

2. The mean and mode for the following frequency distribution

Class interval :	350–369	370–389	390–409	410–429	430–449	450–469
Frequency :	15	27	31	19	13	6

are

(i) 400 and 390 (ii) 400.58 and 390 (iii) 400.58 and 394.50 (iv) 400 and 394.

3. The median and modal profits for the following data

Profit in '000 Rs.:	below 5	below 10	below 15	below 20	below 25	below 30
No. of firms:	10	25	45	55	62	65

are

(i) 11.60 and 11.50 (ii) Rs.11556 and Rs.11267

(iii) Rs.11875 and Rs.11667 (iv) 11.50 and 11.67.

4. Following is an incomplete distribution having modal mark as 44

Marks :	0–20	20–40	40–60	60–80	80–100
No. of Students :	5	18	?	12	5

What would be the mean marks?

(i) 45 (ii) 46 (iii) 47 (iv) 48

5. The data relating to the daily wage of 20 workers are shown below:

   Rs.50, Rs.55, Rs.60, Rs.58, Rs.59, Rs.72, Rs.65, Rs.68, Rs.53, Rs.50, Rs.67, Rs.58, Rs.63, Rs.69, Rs.74, Rs.63, Rs.61, Rs.57, Rs.62, Rs.64.

   The employer pays bonus amounting to Rs.100, Rs.200, Rs.300, Rs.400 and Rs.500 to the wage earners in the wage groups Rs. 50 and not more than Rs. 55 Rs. 55 and not more than Rs. 60 and so on and lastly Rs. 70 and not more than Rs. 75, during the festive month of October.

   What is the average bonus paid per wage earner?

   (i) Rs.200 (ii) Rs.250 (iii) Rs.285 (iv) Rs.300

6. The third quartile and 65th percentile for the following data

Profits in '000 Rs.:	les than 10	10–19	20–29	30–39	40–49	50–59
No. of firms :	5	18	38	20	9	2

   are

   (i) Rs.33500 and Rs.29184 (ii) Rs.33000 and Rs.28680

   (iii) Rs.33600 and Rs.29000 (iv) Rs.33250 and Rs.29250.

7. For the following incomplete distribution of marks of 100 pupils, median mark is known to be 32.

Marks :	0–10	10–20	20–30	30–40	40–50	50–60
No. of Students :	10	–	25	30	–	10

   What is the mean mark?

   (i) 32 (ii) 31 (iii) 31.30 (iv) 31.50

8. The mode of the following distribution is Rs. 66. What would be the median wage?

Daily wages (Rs.) :	30–40	40–50	50–60	60–70	70–80	80–90
No of workers :	8	16	22	28 –	12	

   (i) Rs.64.00 (ii) Rs.64.56 (iii) Rs.62.32 (iv) Rs.64.25

ANSWERS

Set A					
1 (ii)	2 (iii)	3 (ii)	4 (i)	5 (iii)	6 (ii)
7 (ii)	8 (i)	9 (iii)	10 (iv)	11 (iv)	12 (i)
13 (iii)	14 (iii)	15 (ii)	16 (ii)	17 (iii)	18 (iv)
19 (iv)	20 (ii)				
Set B					
1 (i)	2 (ii)	3 (ii)	4 (iii)	5 (iii)	6 (ii)
7 (iv)	8 (iii)	9 (ii)	10 (iii)	11 (i)	12 (iii)
13 (iii)	14 (ii)	15 (i)	16 (iii)	17 (iii)	18 (ii)
19 (i)	20 (iii)				
Set C					
1 (iii)	2 (iii)	3 (iii)	4 (iv)	5 (iv)	6 (i)
7 (iii)	8 (iii)				

11.8 DEFINITION OF DISPERSION

The second important characteristic of a distribution is given by dispersion. Two distributions may be identical in respect of its first important characteristic i.e. central tendency and yet they may differ on account of scatterness. The following figure shows a number of distributions having identical measure of central tendency and yet varying measure of scatterness. Obviously, distribution is having the maximum amount of dispersion.

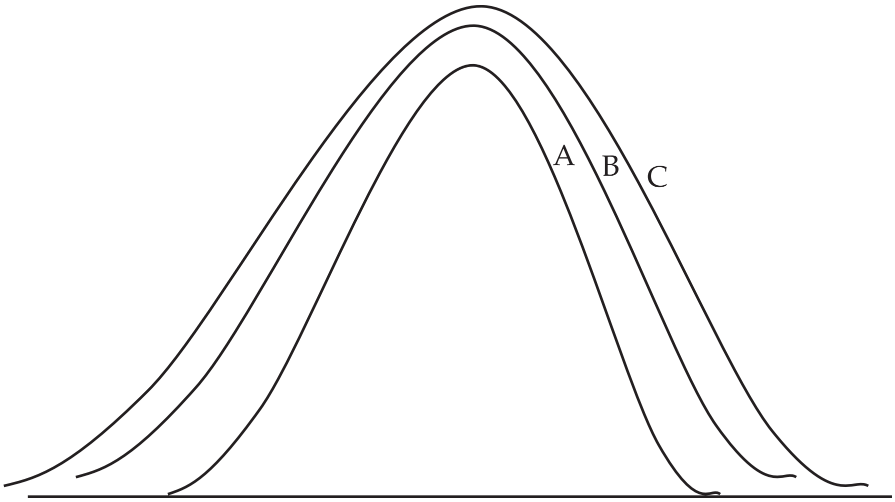

Figure 11.1

Showing distributions with identical measure of central tendency and varying amount of dispersion.

Dispersion for a given set of observations may be defined as the amount of deviation of the observations, usually, from an appropriate measure of central tendency. Measures of dispersion may be broadly classified into

1. Absolute measures of dispersion.
2. Relative measures of dispersion.

Absolute measures of dispersion are classified into

(i) Range
(ii) Mean Deviation
(iii) Standard Deviation
(iv) Quartile Deviation

Likewise, we have the following relative measures of dispersion :

(i) Coefficient of range.
(ii) Coefficient of Mean Deviation
(iii) Coefficient of Variation
(iv) Coefficient of Quartile Deviation.

We may note the following points of distinction between the absolute and relative measures of dispersion :

I Absolute measures are dependent on the unit of the variable under consideration whereas the relative measures of dispersion are unit free.

II For comparing two or more distributions, relative measures and not absolute measures of dispersion are considered.

III Compared to absolute measures of dispersion, relative measures of dispersion are difficult to compute and comprehend.

Characteristics for an ideal measure of dispersion

As discussed in section 11.2 an ideal measure of dispersion should be properly defined, easy to comprehend, simple to compute, based on all the observations, unaffected by sampling fluctuations and amenable to some desirable mathematical treatment.

11.9 RANGE

For a given set of observations, range may be defined as the difference between the largest and smallest of observations. Thus if L and S denote the largest and smallest observations respectively then we have

Range = L – S

The corresponding relative measure of dispersion, known as coefficient of range, is given by

$$\text{Coefficient of range} = \frac{L-S}{L+S}\times 100$$

For a grouped frequency distribution, range is defined as the difference between the two extreme class boundaries. The corresponding relative measure of dispersion is given by the ratio of the difference between the two extreme class boundaries to the total of these class boundaries, expressed as a percentage.

We may note the following important result in connection with range:

Result:

Range remains unaffected due to a change of origin but affected in the same ratio due to a change in scale i.e., if for any two constants a and b, two variables x and y are related by y = a + bx,

Then the range of y is given by

$$R_y = |b| \times R_x \quad \text{...} \quad (11.23)$$

Example 11.23: Following are the wages of 8 workers expressed in rupees:
82, 96, 52, 75, 70, 65, 50, 70. Find the range and also it's coefficient.

Solution : The largest and the smallest wages are L = Rs.96 and S= Rs.50
Thus range = Rs.96 – Rs.50 = Rs.46

$$\text{Coefficient of range} = \frac{96-50}{96+50}\times 100$$

$$= 31.51$$

Example 11.24 : What is the range and its coefficient for the following distribution of weights?

Weights in kgs. :	50 – 54	55 – 59	60 – 64	65 – 69	70 – 74
No. of Students :	12	18	23	10	3

Solution : The lowest class boundary is 49.50 kgs. and the highest class boundary is 74.50 kgs. Thus we have

Range = 74.50 kgs. – 49.50 kgs.

= 25 kgs.

Also, coefficient of range = $\frac{74.50 - 49.50}{74.50 + 49.50} \times 100$

$= \frac{25}{124} \times 100$

$= 20.16$

Example 11.25 : If the relationship between x and y is given by 2x+3y=10 and the range of x is Rs. 15, what would be the range of y?

Solution: Since 2x+3y=10

Therefore, $y = \frac{10}{3} - \frac{2}{3}x$

Applying (11.23) , the range of y is given by

$$R_y = |b| \times R_x$$
$$= 2/3 \times \text{Rs. } 15$$
$$= \text{Rs.}10.$$

11.10 MEAN DEVIATION

Since range is based on only two observations, it is not regarded as an ideal measure of dispersion. A better measure of dispersion is provided by mean deviation which, unlike range, is based on all the observations. For a given set of observation, mean deviation is defined as the arithmetic mean of the absolute deviations of the observations from an appropriate measure of central tendency. Hence if a variable x assumes n values $x_1, x_2, x_3 \ldots x_n$, then the mean deviation of x about an average A is given by

$$MD_A = \frac{1}{n}\sum|x_i - A| \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(11.24)$$

For a grouped frequency distribution, mean deviation about A is given by

$$MD_A = \frac{1}{n}\sum|x_i - A|f_i \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots....(11.25)$$

Where x_i and f_i denote the mid value and frequency of the i-th class interval and

$$N = \sum f_i$$

In most cases we take A as mean or median and accordingly, we get mean deviation about mean or mean deviation about median.

A relative measure of dispersion applying mean deviation is given by

$$\text{Coefficient of mean deviation} = \frac{\text{Mean deviation about A}}{\text{A}} \times 100 \quad \ldots\ldots\ldots\ldots\ldots.(11.26)$$

Mean deviation takes its minimum value when the deviations are taken from the median. Also mean deviation remains unchanged due to a change of origin but changes in the same ratio due to a change in scale i.e. if y = a + bx, a and b being constants,

then MD of y = |b| × MD of x ………………………(11.27)

Example. 11.26 : What is the mean deviation about mean for the following numbers?

5, 8, 10, 10, 12, 9.

Solution:

The mean is given by

$$\overline{X} = \frac{5+8+10+10+12+9}{6} = 9$$

Table 11.12

Computation of MD about AM	
x_i	$\|x_i - \overline{x}\|$
5	4
8	1
10	1
10	1
12	3
9	0
Total	10

Thus mean deviation about mean is given by

$$\frac{\sum|x_i - \bar{x}|}{n} = \frac{10}{6} = 1.67$$

Example. 11.27: Find mean deviations about median and also the corresponding coefficient for the following profits ('000 Rs.) of a firm during a week.

82, 56, 75, 70, 52, 80, 68.

Solution:

The profits in thousand rupees is denoted by x. Arranging the values of x in an ascending order, we get

52, 56, 68, 70, 75, 80, 82.

Therefore, Me = 70. Thus, Median profit = Rs. 70,000.

Table 11.13

Computation of Mean deviation about median

x_i	$\lvert x_i - Me \rvert$
52	18
56	14
68	2
70	0
75	5
80	10
82	12
Total	61

Thus mean deviation about median = $\frac{\sum|x_i - \text{Median}|}{n}$

$$= \text{Rs.}\frac{61}{7}\times 1000$$

$$= \text{Rs.}8714.28$$

$$\text{Coefficient of mean deviation} = \frac{\text{MD about median}}{\text{Median}} \times 100$$

$$= \frac{8714.28}{70000} \times 100$$

$$= 12.45$$

Example 11.28 : Compute the mean deviation about the arithmetic mean for the following data:

x :	1	3	5	7	9
f :	5	8	9	2	1

lso find the coefficient of the mean deviation about the AM.

Solution: We are to apply formula (11.25) as these data refer to a grouped frequency distribution the AM is given by

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$$= \frac{5\times1+8\times3+9\times5+2\times7+1\times9}{5+8+9+2+1} = 3.88$$

Table 11.14

Computation of MD about the AM

x (1)	f (2)	$\lvert x-\bar{x}\rvert$ (3)	$f\lvert x-\bar{x}\rvert$ (4) = (2) × (3)
1	5	2.88	14.40
3	8	0.88	7.04
5	9	1.12	10.08
7	2	3.12	6.24
9	1	5.12	5.12
Total	25	–	42.88

Thus, MD about AM is given by

$$\frac{\sum f \lvert x-\bar{x}\rvert}{N}$$

$= \frac{42.88}{25}$

$= 1.72$

$$\text{Coefficient of MD about its AM} = \frac{\text{MD about AM}}{\text{AM}} \times 100$$

$$= \frac{1.72}{3.88} \times 100$$

$$= 44.33$$

Example 11.29 : Compute the coefficient of mean deviation about median for the following distribution:

Weight in kgs.	:	40-50	50-60	60-70	70-80
No. of persons	:	8	12	20	10

Solution: We need to compute the median weight in the first stage

Table 11. 15

Computation of median weight

Weight in kg (CB)	No. of Persons (Cumulative Frequency)
40	0
50	8
60	20
70	40
80	50

Hence, $$M = l_1 + \left(\frac{\frac{N}{2} - N_l}{N_u - N_l}\right) \times C$$

$$= \left[60 + \frac{25-20}{40-20} \times 10\right] \text{kg.} = 62.50\text{kg.}$$

Table 11.16

Computation of mean deviation of weight about median

weight (kgs.) (1)	mid-value (x_i) kgs. (2)	No. of persons (f_i) (3)	$\lvert x_i - Me \rvert$ (kgs.) (4)	$f_i \lvert x_i - Me \rvert$ (kgs.) (5)=(3)×(4)
40–50	45	8	17.50	140
50–60	55	12	7.50	90
60–70	65	20	2.50	50
70–80	75	10	12.50	125
Total	–	50	–	405

Mean deviation about median = $\frac{\sum f_i \lvert x_i - \text{Median} \rvert}{N}$

$$= \frac{405}{50} \text{kg.}$$

$$= 8.10 \text{ kg.}$$

Coefficient of mean deviation about median = $\frac{\text{Mean deviation about median}}{\text{Median}} \times 100$

$$= \frac{8.10}{62.50} \times 100$$

$$= 12.96$$

Example 11.30: If x and y are related as 4x+3y+11 = 0 and mean deviation of x is 5.40, what is the mean deviation of y?

Solution: Since 4x + 3y + 11 = 0

Therefore, $y = \left(\frac{-11}{3}\right) + \left(\frac{-4}{3}\right)x$

Hence MD of y= |b| × MD of x

$$= \frac{4}{3} \times 5.40$$
$$= 7.20$$

11.11 STANDARD DEVIATION

Although mean deviation is an improvement over range so far as a measure of dispersion is concerned, mean deviation is difficult to compute and further more, it cannot be treated mathematically. The best measure of dispersion is, usually, standard deviation which does not possess the demerits of range and mean deviation.

Standard deviation for a given set of observations is defined as the root mean square deviation when the deviations are taken from the AM of the observations. If a variable x assumes n values x_1, x_2, x_3 ....x_n then its standard deviation(s) is given by

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}} \quad \text{..........................(11.28)}$$

For a grouped frequency distribution, the standard deviation is given by

$$s = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{N}} \quad \text{.......................... (11.29)}$$

(11.28) and (11.29) can be simplified to the following forms

$$s = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2} \quad \text{for unclassified data}$$

$$= \sqrt{\frac{\sum f_i x_i^2}{N} - \bar{x}^2} \quad \text{for a grouped frequency distribution.}$$

.... (11.30)

Sometimes the square of standard deviation, known as variance, is regarded as a measure of dispersion. We have, then,

$$\text{Variance} = s^2 = \frac{\sum (x_i - \bar{x})^2}{n} \quad \text{for unclassified data}$$

$$= \frac{\sum f_i (x_i - \bar{x})^2}{N} \quad \text{for a grouped frequency distribution ....(11.31)}$$

A relative measure of dispersion using standard deviation is given by coefficient of variation (cv) which is defined as the ratio of standard deviation to the corresponding arithmetic mean, expressed as a percentage.

$$\text{Coefficient of Variation (CV)} = \frac{SD}{AM} \times 100 \quad \text{.......................... (11.32)}$$

Illustration

Example 11.31: Find the standard deviation and the coefficient of variation for the following numbers: 5, 8, 9, 2, 6

Solution: We present the computation in the following table.

Table 11.17
Computation of standard deviation

x_i	x_i^2
5	25
8	64
9	81
2	4
6	36
30	$\sum x_i^2 = 210$

Applying (11.30), we get the standard deviation as

$$s = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$

$$= \sqrt{\frac{210}{5} - \left(\frac{30}{5}\right)^2} \qquad \left(\text{since } \bar{x} = \frac{\Sigma x_i}{n}\right)$$

$$= \sqrt{42 - 36}$$

$$= \sqrt{6}$$

$$= 2.45$$

The coefficient of variation is

$$CV = 100 \times \frac{SD}{AM}$$

$$= 100 \times \frac{2.45}{6}$$

$$= 40.83$$

Example 11.32: Show that for any two numbers a and b, standard deviation is given by $\frac{|a-b|}{2}$.

Solution: For two numbers a and b, AM is given by $\bar{x} = \frac{a+b}{2}$

The variance is

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

$$= \frac{\left(a - \frac{a+b}{2}\right)^2 + \left(b - \frac{a+b}{2}\right)^2}{2}$$

$$= \frac{\frac{(a-b)^2}{4} + \frac{(a-b)^2}{4}}{2}$$

$$= \frac{(a-b)^2}{4}$$

$$\Rightarrow \quad s = \frac{|a-b|}{2}$$

(The absolute sign is taken, as SD cannot be negative).

Example 11.33: Prove that for the first n natural numbers, SD is $\sqrt{\frac{n^2-1}{12}}$.

Solution: for the first n natural numbers AM is given by

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

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

$$= \frac{n+1}{2}$$

$$\therefore \text{SD} = \sqrt{\frac{\sum x_i^2}{n} - \bar{x}^2}$$

$$= \sqrt{\frac{1^2+2^2+3^2\ldots\ldots\ldots\ldots\ldots\ldots+n^2}{n} - \left(\frac{n+1}{2}\right)^2}$$

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

$$= \sqrt{\frac{(n+1)(4n+2-3n-3)}{12}} = \sqrt{\frac{n^2-1}{12}}$$

Thus, SD of first n natural numbers is SD = $\sqrt{\frac{n^2-1}{12}}$

We consider the following formula for computing standard deviation from grouped frequency distribution with a view to saving time and computational labour:

$$S = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \qquad \text{……………………………..(11.33)}$$

Where $d_i = \frac{x_i - A}{C}$

Example 11.34: Find the SD of the following distribution:

Weight (kgs.)	:	50-52	52-54	54-56	56-58	58-60
No. of Students	:	17	35	28	15	5

Solution:

Table 11.17
Computation of SD

Weight (kgs.) (1)	No. of Students (f_i) (2)	Mid-value (x_i) (3)	$d_i = \frac{x_i - 55}{2}$ (4)	$f_i d_i$ (5)=(2)×(4)	$f_i d_i^2$ (6)=(5)×(4)
50-52	17	51	–2	–34	68
52-54	35	53	–1	–35	35
54-56	28	55	0	0	0
56-58	15	57	1	15	15
58-60	5	59	2	10	20
Total	100	–	–	– 44	138

Applying (11.33), we get the SD of weight as

$$= \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times C$$

$$= \sqrt{\frac{138}{100} - \frac{(-44)^2}{100}} \times 2\text{kgs.}$$

$$= \sqrt{1.38 - 0.1936} \times 2 \text{ kgs.}$$

$$= 2.18 \text{ kgs.}$$

Properties of standard deviation

I. If all the observations assumed by a variable are constant i.e. equal, then the SD is zero. This means that if all the values taken by a variable x is k, say , then s = 0. This result applies to range as well as mean deviation.

II. SD remains unaffected due to a change of origin but is affected in the same ratio due to a change of scale i.e., if there are two variables x and y related as y = a+bx for any two constants a and b, then SD of y is given by

…………………………..(11.34)

III. If there are two groups containing n_1 and n_2 observations, $_1$ and $_2$ as respective AM's, s_1 and s_2 as respective SD's , then the combined SD is given by

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}} \quad \text{…………………………..(11.35)}$$

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

$d_2 = \bar{x}_2 - \bar{x}$

and $\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}$ = combined AM

This result can be extended to more than 2 groups. For x(7[2]) groups, we have

$$s = \sqrt{\frac{\sum n_i s_i^2 + \sum n_i d_i^2}{\sum n_i}} \quad \text{………………………….. (11.36)}$$

With $d_i = x_i - \bar{x}$

and $\bar{x} = \frac{\sum n_i \bar{x}_i}{\sum n_i}$

Where $\bar{x}_1 = \bar{x}_2$ (11.35) is reduced to

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}}$$

Example 11.35: If AM and coefficient of variation of x are 10 and 40 respectively, what is the variance of (15–2x)?

Solution: let y = 15 – 2x

Then applying (11.34), we get,

$s_y = 2 \times s_x$ ………………………………… (1)

As given cv_x = coefficient of variation of x = 40 and $\bar{x}$ = 10

This $cv_x = \frac{s_x}{x} \times 100$

$$\Rightarrow \quad 40 = \frac{S_x}{10} \times 100$$

$$\Rightarrow \quad S_x = 4$$

From (1), $S_y = 2 \times 4 = 8$

Therefore, variance of $(15-2x) = {S_y}^2 = 64$

Example 11.36: Compute the SD of 9, 5, 8, 6, 2.
Without any more computation, obtain the SD of

Sample I	–1,	–5,	–2,	–4,	–8,
Sample II	90,	50,	80,	60,	20,
Sample III	23,	15,	21,	17,	9.

Solution:

Table 11.18
Computation of SD

x_i	${x_i}^2$
9	81
5	25
8	64
6	36
2	4
30	210

The SD of the original set of observations is given by

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

$$= \sqrt{\frac{210}{5} - \left(\frac{30}{5}\right)^2}$$

$$= \sqrt{42 - 36}$$

$$= \sqrt{6}$$

$$= 2.45$$

If we denote the original observations by x and the observations of sample I by y, then we have

$y = -10 + x$

$y = (-10) + (1)\, x$

$\therefore S_y = |1| \times S_x$

$= 1 \times 2.45$

$= 2.45$

In case of sample II, x and y are related as

$Y = 10x$

$= 0 + (10)x$

$\therefore s_y = |10| \times s_x$

$= 10 \times 2.45$

$= 24.50$

And lastly, $y = (5) + (2)x$

$\Rightarrow s = 2 \times 2.45$

$= 4.90$

Example 11.37: For a group of 60 boy students, the mean and SD of stats. marks are 45 and 2 respectively. The same figures for a group of 40 girl students are 55 and 3 respectively. What is the mean and SD of marks if the two groups are pooled together?

Solution: As given $n_1 = 60$, $\bar{x}_1 = 45$, $s_1 = 2$ $n_2 = 40$, $\bar{x}_2 = 55$, $s_2 = 3$

Thus the combined mean is given by

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

$$= \frac{60 \times 45 + 40 \times 55}{60 + 40}$$

$= 49$

Thus $d_1 = \bar{x}_1 - \bar{x} = 45 - 49 = -4$

$d_2 = \bar{x}_2 - \bar{x} = 55 - 49 = 6$

Applying (11.35), we get the combined SD as

$$s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2 + n_1 d_1^2 + n_2 d_2^2}{n_1 + n_2}}$$

$$s = \sqrt{\frac{60 \times 2^2 + 40 \times 3^2 + 60 \times (-4)^2 + 40 \times 6^2}{60 + 40}}$$

$= \sqrt{30}$

$= 5.48$

Example 11.38: The mean and standard deviation of the salaries of the two factories are provided below :

Factory	No. of Employees	Mean Salary	SD of Salary
A	30	Rs.4800	Rs.10
B	20	Rs. 5000	Rs.12

i) Find the combined mean salary and standard deviation of salary.

ii) Examine which factory has more consistent structure so far as satisfying its employees are concerned.

Solution: Here we are given

$n_1 = 30$, = Rs.4800, s_1= Rs.10,

$n_2 = 20$, $\bar{x}_2$= Rs.5000, s_2= Rs.12

i) $$\frac{30 \times \text{Rs.}4800 + 20 \times \text{Rs.}5000}{30+20} = \text{Rs.}4800$$

$d_1 = \bar{x}_1 - \bar{x}$ = Rs.4,800 - Rs.4880 = - Rs.80

$d_2 = \bar{x}_2 - \bar{x}$ = Rs.5,000 - Rs.4880 = Rs.120

hence, the combined SD in rupees is given by

$$s = \sqrt{\frac{30\times10^2 + 20\times12^2 + 30\times(-80)^2 + 20\times120^2}{30+20}}$$

$= \sqrt{9717.60}$

$= 98.58$

thus the combined mean salary and the combined standard deviation of salary are Rs.4880 and Rs.98.58 respectively.

ii) In order to find the more consistent structure, we compare the coefficients of variation of the two factories. Letting $CV_A = 100 \times \frac{S_A}{\bar{x}_A}$ and $CV_B = 100 \times \frac{S_B}{\bar{x}_B}$

We would say factory A is more consistent

if $CV_A < CV_B$. Otherwise factory B would be more consistent.

$$\text{Now } CV_A = 100\times\frac{s_A}{\bar{x}_A} = 100\times\frac{s_1}{\bar{x}_1} = \frac{100\times10}{4800} = 0.21$$

and $CV_B = 100 \times \frac{s_B}{\bar{x}_B} = 100 \times \frac{s_2}{\bar{x}_2} = \frac{100 \times 12}{5000} = 0.24$

Thus we conclude that factory A has more consistent structure.

Example 11.39: A student computes the AM and SD for a set of 100 observations as 50 and 5 respectively. Later on, she discovers that she has made a mistake in taking one observation as 60 instead of 50. What would be the correct mean and SD if

i) The wrong observation is left out?
ii) The wrong observation is replaced by the correct observation?

Solution: As given, n = 100, $\bar{x} = 50$, S = 5

Wrong observation = 60(x), correct observation = 50(V)

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

$$\Rightarrow \quad \sum x_i = n\bar{x} = 100 \times 50 = 5000$$

$$\text{and} \quad s^2 = \frac{\sum x_i^2}{n} - \bar{x}^2$$

$$\Rightarrow \sum x_i^2 = n(\bar{x}^2 + s^2) = 100(50^2 + 5^2) = 252500$$

i) Sum of the 99 observations = 5000 – 60 = 4940
AM after leaving the wrong observation = 4940/99 = 49.90
Sum of squares of the observation after leaving the wrong observation
= 252500 – 60^2 = 248900
Variance of the 99 observations = 248900/99 – $(49.90)^2$
= 2514.14 – 2490.01
= 24.13
∴ SD of 99 observations = 4.91

ii) Sum of the 100 observations after replacing the wrong observation by the correct observation = 5000 – 60 + 50 =4990

$$AM = \frac{4990}{100} = 49.90$$

Corrected sum of squares = $252500 + 50^2 - 60^2 = 251400$

Corrected SD $= \sqrt{\frac{251400}{100} - (49.90)^2}$

$= \sqrt{45.99}$

$= 6.78$

11.12 QUARTILE DEVIATION

Another measure of dispersion is provided by **quartile deviation** or **semi - inter –quartile** range which is given by

$$Q_d = \frac{Q_3 - Q_1}{2} \quad \text{.................................(11.37)}$$

A relative measure of dispersion using quartiles is given by coefficient of quartile deviation which is

$$\text{Coefficient of quartile deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 \quad \text{.................................(11.38)}$$

Quartile deviation provides the best measure of dispersion for open-end classification. It is also less affected due to sampling fluctuations. Like other measures of dispersion, quartile deviation remains unaffected due to a change of origin but is affected in the same ratio due to change in scale.

Example 11.40 : Following are the marks of the 10 students : 56, 48, 65, 35, 42, 75, 82, 60, 55, 50. Find quartile deviation and also its coefficient.

Solution:
After arranging the marks in an ascending order of magnitude, we get 35, 42, 48, 50, 55, 56, 60, 65, 75, 82

$$\text{First quartile } (Q_1) = \frac{(n+1)}{4} \text{th observation}$$

$= \frac{(10+1)}{4} th$ observation

$= 2.75^{th}$ observation

$= 2^{nd}$ observation + 0.75 × difference between the third and the 2^{nd} observation.

= 42 + 0.75 × (48 – 42)

= 46.50

$$\text{Third quartile } (Q_3) = \frac{3(n+1)}{4} \text{th observation}$$

$= 8.25^{th}$ observation

= 65 + 0.25 × 10

= 67.50

Thus applying (11.37), we get the quartile deviation as

$$\frac{Q_3 - Q_1}{2} = \frac{67.50 - 46.50}{2} = 10.50$$

Also, using (11.38), the coefficient of quartile deviation

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

$$= \frac{67.50 - 46.50}{67.50 + 46.50} \times 100$$

$$= 18.42$$

Example 11.41 : If the quartile deviation of x is 6 and 3x + 6y = 20, what is the quartile deviation of y?

Solution: 3x + 6y = 20

$$y = \left(\frac{20}{6}\right) + \left(\frac{-3}{6}\right)x$$

Therefore, quartile deviation of $y = \frac{|-3|}{6} \times \text{quartile deviation of } x$

$$= \frac{1}{2} \times 6$$

$$= 3.$$

Example 11.42: Find an appropriate measures of dispersion from the following data:

Daily wages (Rs.)	:	upto 20	20-40	40-60	60-80	80-100
No. of workers	:	5	11	14	7	3

Solution: Since this is an open-end classification, the appropriate measure of dispersion would be quartile deviation as quartile deviation does not taken into account the first twenty five percent and the last twenty five per cent of the observations.

Table 11.19
Computation of Quartile

Daily wages in Rs. (Class boundary)	No. of workers (less than cumulative frequency)
a	0
20	5
40	16
60	30
80	37
100	40

Here a denotes the first Class Boundary

$Q_1 = \text{Rs. } \left[20 + \frac{10-5}{16-5} \times 20\right] = \text{Rs. } 29.09$

$Q_3 = \text{Rs. } 60$

Thus quartile deviation of wages is given by

$$\frac{Q_3 - Q_1}{2}$$

$$= \frac{\text{Rs. } 60 - \text{Rs. } 29.09}{2}$$

$= \text{Rs. } 15.46$

Example 11.43: The mean and variance of 5 observations are 4.80 and 6.16 respectively. If three of the observations are 2,3 and 6, what are the remaining observations?

Solution: Let the remaining two observations be a and b, then as given

$$\frac{2+3+6+a+b}{5} = 4.80$$

$\Rightarrow \quad 11+a+b = 24$

$\Rightarrow \quad a+b = 13 \quad ....(1)$

and $\frac{2^2 + a^2 + b^2 + 3^2 + 6^2}{5} - (4.80)^2$

$\Rightarrow \quad \frac{49 + a^2 + b^2}{5} - 23.04 = 6.16$

$\Rightarrow \quad 49 + a^2 + b^2 = 146$

$\Rightarrow \quad a^2 + b^2 = 97 \quad ....(2)$

From (1), we get $a = 13 - b$ ....(3)

Eliminating a from (2) and (3), we get

$(13 - b)^2 + b^2 = 97$

$\Rightarrow \quad 169 - 26b + 2b^2 = 97$

$\Rightarrow \quad b^2 - 13b + 36 = 0$

$\Rightarrow \quad (b-4)(b-9) = 0$

$\Rightarrow \quad b = 4 \text{ or } 9$

From (3), $a = 9 \text{ or } 4$

Thus the remaining observations are 4 and 9.

Example 11.44 : After shift of origin and change of scale, a frequency distribution of a continuous variable with equal class length takes the following form of the changed variable (d):

d	:	–2	–1	0	1	2
frequency	:	17	35	28	15	5

If the mean and standard deviation of the original frequency distribution are 54.12 and 2.1784 respectively, find the original frequency distribution.

Solution: we need find out the origin A and scale C from the given conditions.

Since $d_i = \frac{x_i - A}{C}$

$\Rightarrow \quad x_i = A + Cd_i$

once A and C are known, the mid- values x_i's would be known. Finally, we convert the mid-values to the corresponding class boundaries by using the formula:

$LCB = x_i - C/2$

and $UCB = x_i + C/2$

On the basis of the given data, we find that

$\Sigma f_i d_i = -44$, $\Sigma f_i d_i^2 = 138$ and $N = 100$

$$\text{Hence } s = \sqrt{\frac{\Sigma f_i d_i^2}{N} - \left(\frac{\Sigma f_i d_i}{N}\right)^2} \times C$$

$$\Rightarrow \quad 2.1784 = \sqrt{\frac{138}{100} - \left(\frac{-44}{100}\right)^2} \times C$$

$$\Rightarrow \quad 2.1784 = \sqrt{1.38 - 0.1936} \times C$$

$$\Rightarrow \quad 2.1784 = 1.0892 \times C$$

$$\Rightarrow \quad C = 2$$

$$\text{Further, } \bar{x} = A + \frac{\Sigma f_i d_i}{N} \times C$$

$$\Rightarrow \quad 54.12 = A + \frac{-44}{100} \times 2$$

$$\Rightarrow \quad 54.12 = A - 0.88$$

$$\Rightarrow \quad A = 55$$

$$\text{Thus } x_i = A + Cd_i$$

$$\Rightarrow \quad x_i = 55 + 2d_i$$

Table 11.20

Computation of the Original Frequency Distribution

d_i	f_i	$x_i = 55 + 2d_i$	class interval $x_i \pm \frac{C}{2}$
–2	17	51	50-52
–1	35	53	52-54
0	28	55	54-56
1	15	57	56-58
2	5	59	58-60

Example 11.45: Compute coefficient of variation from the following data:

Age	:	under 10	under 20	under 30	under 40	under 50	under 60
No. of persons Dying	:	10	18	30	45	60	80

Solution: What is given in this problem is less than cumulative frequency distribution. We need first convert it to a frequency distribution and then compute the coefficient of variation.

Table 11.21

Computation of coefficient of variation

Age in years class Interval	No. of persons dying (f_i)	Mid-value (x_i)	$d_i = \frac{x_i - 25}{10}$	$f_i d_i$	$f_i d_i^2$
0-10	10	5	–2	–20	40
10-20	18–10= 8	15	–1	–8	8
20-30	30–18=12	25	0	0	0
30-40	45–30=15	35	1	15	15
40-50	60–45=15	45	2	30	60
50-60	80–60=20	55	3	60	180
Total	80	–	–	77	303

The AM is given by:

$$\bar{x} = A + \frac{\sum f_i d_i}{N} \times C$$

$$= \left(25 + \frac{77}{80} \times 10\right) \text{ years}$$

$$= 34.63 \text{ years}$$

The standard deviation is

$$s = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times C$$

$$= \sqrt{\frac{303}{80} - \left(\frac{77}{80}\right)^2} \times 10 \text{ years}$$

$$= \sqrt{3.79 - 0.93} \times 10 \text{ years}$$

$$= 16.91 \text{ years}$$

Thus the coefficient of variation is given by

$$CV = \frac{S}{X} \times 100$$

$$= \frac{16.91}{34.63} \times 100$$

$$= 48.83$$

Example 11.46 : you are given the distribution of wages in two factors A and B

Wages in Rs.	:	100-200	200-300	300-400	400-500	500-600	600-700
No. of workers in A	:	8	12	17	10	2	1
No. of workers in B	:	6	18	25	12	2	2

State in which factory, the wages are more variable.

Solution :

As explained in example 11.36, we need compare the coefficient of variation of A(i.e. v_A) and of B (i.e v_B).

If $v_A > v_B$, then the wages of factory A woyld be more variable. Otherwise, the wages of factory B would be more variable where

$$V_A = 100 \times \frac{s_A}{\bar{x}_A} \quad \text{and} \quad V_B = 100 \times \frac{s_B}{\bar{x}_B}$$

Table 11.22

Computation of coefficient of variation of wages of Two Factories A and B

Wages in rupees (1)	Mid-value x (2)	d= (3)	No. of workers of A f_A (4)	No. of workers of B f_B (5)	f_Ad (6)=(3)×(4)	f_Ad^2 (7)=(3)×(6)	f_Bd (8)=(3)×(5)	f_Bd^2 (9)=(3)×(8)
100-200	150	–2	8	6	–16	32	–12	24
200-300	250	–1	12	18	–12	12	–18	18
300-400	350	0	17	25	0	0	0	0
400-500	450	1	10	12	10	10	12	12
500-600	550	2	2	2	4	8	4	8
600-700	650	3	1	2	3	9	6	18
Total	–	–	50	65	–11	71	– 8	80

For Factory A

$$\bar{x}_A = \text{Rs.}\left(350 + \frac{-11}{50} \times 100\right) = \text{Rs.}328$$

$$S_A = \text{Rs.}\sqrt{\frac{71}{50} - \left(\frac{-11}{50}\right)^2} \times 100 = \text{Rs. } 117.12$$

$$\therefore V_A = \frac{S_A}{\bar{x}_A} \times 100 = 35.71$$

For Factory B

$$\bar{x}_B = \text{Rs.}\left(350 + \frac{-8}{65} \times 100\right) = \text{Rs.}337.69$$

$$S_B = \text{Rs.}\sqrt{\frac{80}{65} - \left(\frac{-8}{65}\right)^2} \times 100$$

$$= \text{Rs.}110.25$$

$$\therefore V_B = \frac{110.25}{337.69} \times 100 = 32.65$$

As $V_A > V_B$, the wages for factory A is more variable.

Comparison between different measures of dispersion

We may now have a review of the different measures of dispersion on the basis of their relative merits and demerits. Standard deviation, like AM, is the best measure of dispersion. It is rigidly defined, based on all the observations, not too difficult to compute, not much affected by sampling fluctuations and moreover it has some desirable mathematical properties. All these merits of standard deviation make SD as the most widely and commonly used measure of dispersion.

Range is the quickest to compute and as such, has its application in statistical quality control. However, range is based on only two observations and affected too much by the presence of extreme observation(s).

Mean deviation is rigidly defined, based on all the observations and not much affected by sampling fluctuations. However, mean deviation is difficult to comprehend and its computation is also time consuming and laborious. Furthermore, unlike SD, mean deviation does not possess mathematical properties.

Quartile deviation is also rigidly defined, easy to compute and not much affected by sampling fluctuations. The presence of extreme observations has no impact on quartile deviation since quartile deviation is based on the central fifty-percent of the observations. However, quartile deviation is not based on all the observations and it has no desirable mathematical properties. Nevertheless, quartile deviation is the best measure of dispersion for open-end classifications.

11.13 EXERCISE

Set A

Write down the correct answers. Each question carries one mark.

1. Which of the following statements is correct?
   (a) Two distributions may have identical measures of central tendency and dispersion.
   (b) Two distributions may have the identical measures of central tendency but different measures of dispersion.
   (c) Two distributions may have the different measures of central tendency but identical measures of dispersion.
   (d) All the statements (a), (b) and (c).
2. Dispersion measures
   (a) The scatterness of a set of observations
   (b) The concentration of a set of observations
   (c) Both a) and b)
   (d) Neither a) and b).

3. When it comes to comparing two or more distributions we consider
   (a) Absolute measures of dispersion (b) Relative measures of dispersion
   (c) Both a) and b) (d) Either (a) or (b).
4. Which one is easier to compute?
   (a) Relative measures of dispersion (b) Absolute measures of dispersion
   (c) Both a) and b) (d) Range
5. Which one is an absolute measure of dispersion?
   (a) Range (b) Mean Deviation
   (c) Standard Deviation (d) All these measures
6. Which measure of dispersion is most usefull?
   (a) Standard deviation (b) Quartile deviation
   (c) Mean deviation (d) Range
7. Which measures of dispersions is not affected by the presence of extreme observations?
   (a) Range (b) Mean deviation
   (c) Standard deviation (d) Quartile deviation
8. Which measure of dispersion is based on the absolute deviations only?
   (a) Standard deviation (b) Mean deviation
   (c) Quartile deviation (d) Range
9. Which measure is based on only the central fifty percent of the observations?
   (a) Standard deviation (b) Quartile deviation
   (c) Mean deviation (d) All these measures
10. Which measure of dispersion is based on all the observations?
   (a) Mean deviation (b) Standard deviation
   (c) Quartile deviation (d) (a) and (b) but not (c)
11. The appropriate measure of dispersion for open – end classification is
   (a) Standard deviation (b) Mean deviation
   (c) Quartile deviation (d) All these measures.
12. The most commonly used measure of dispersion is
   (a) Range (b) Standard deviation
   (c) Coefficient of variation (d) Quartile deviation.

13. Which measure of dispersion has some desirable mathematical properties?
    (a) Standard deviation (b) Mean deviation
    (c) Quartile deviation (d) All these measures
14. If the profits of a company remains the same for the last ten months, then the standard deviation of profits for these ten months would be ?
    (a) Positive (b) Negative (c) Zero (d) (a) or (c)
15. Which measure of dispersion is considered for finding a pooled measure of dispersion after combining several groups?
    (a) Mean deviation (b) Standard deviation
    (c) Quartile deviation (d) Any of these
16. A shift of origin has no impact on
    (a) Range (b) Mean deviation
    (c) Standard deviation (d) All these and quartile deviation.
17. The range of 15, 12, 10, 9, 17, 20 is
    (a) 5 (b) 12 (c) 13 (d) 11.
18. The standard deviation of, 10, 16, 10, 16, 10, 10, 16, 16 is
    (a) 4 (b) 6 (c) 3 (d) 0.
19. For any two numbers SD is always
    (a) Twice the range (b) Half of the range
    (c) Square of the range (d) None of these.
20. If all the observations are increased by 10, then
    (a) SD would be increased by 10
    (b) Mean deviation would be increased by 10
    (c) Quartile deviation would be increased by 10
    (d) All these three remain unchanged.
21. If all the observations are multiplied by 2, then
    (a) New SD would be also multiplied by 2
    (b) New SD would be half of the previous SD
    (c) New SD would be increased by 2
    (d) New SD would be decreased by 2.

Set B

Write down the correct answers. Each question carries two marks.

1. What is the coefficient of range for the following wages of 8 workers?

   Rs.80, Rs.65, Rs.90, Rs.60, Rs.75, Rs.70, Rs.72, Rs.85.

   (a) Rs.30 (b) Rs.20 (c) 30 (d) 20

2. If R_x and R_y denote ranges of x and y respectively where x and y are related by 3x+2y+10=0, what would be the relation between x and y?

   (a) $R_x = R_y$ (b) $2 R_x = 3 R_y$ (c) $3 R_x = 2 R_y$ (d) $R_x = 2 R_y$

3. What is the coefficient of range for the following distribution?

Class Interval :	10-19	20-29	30-39	40-49	50-59
Frequency :	11	25	16	7	3

   (a) 22 (b) 50 (c) 72.46 (d) 75.82

4. If the range of x is 2, what would be the range of –3x +50 ?

   (a) 2 (b) 6 (c) –6 (d) 44

5. What is the value of mean deviation about mean for the following numbers?
   5, 8, 6, 3, 4.

   (a) 5.20 (b) 7.20 (c) 1.44 (d) 2.23

6. What is the value of mean deviation about mean for the following observations?
   50, 60, 50, 50, 60, 60, 60, 50, 50, 50, 60, 60, 60, 50.

   (a) 5 (b) 7 (c) 35 (d) 10

7. The coefficient of mean deviation about mean for the first 9 natural numbers is

   (a) 200/9 (b) 80 (c) 400/9 (d) 50.

8. If the relation between x and y is 5y–3x = 10 and the mean deviation about mean for x is 12, then the mean deviation of y about mean is

   (a) 7.20 (b) 6.80 (c) 20 (d) 18.80.

9. If two variables x and y are related by 2x + 3y –7 =0 and the mean and mean deviation about mean of x are 1 and 0.3 respectively, then the coefficient of mean deviation of y about its mean is

   (a) –5 (b) 12 (c) 50 (d) 4.

10. The mean deviation about mode for the numbers 4/11, 6/11, 8/11, 9/11, 12/11, 8/11 is
    (a) 8/11 (b) 1/11 (c) 6/11 (d) 5/11.

11. What is the standard deviation of 5, 5, 9, 9, 9, 10, 5, 10, 10?

    (a) $\sqrt{14}$ (b) $\frac{\sqrt{42}}{3}$ (c) 4.50 (d) 8

12. If the mean and SD of x are a and b respectively, then the SD of $\frac{x-a}{b}$ is

(a) –1 (b) 1 (c) ab (d) a/b.

13. What is the coefficient of variation of the following numbers?
53, 52, 61, 60, 64.
(a) 8.09 (b) 18.08 (c) 20.23 (d) 20.45

14. If the SD of x is 3, what us the variance of (5–2x)?
(a) 36 (b) 6 (c) 1 (d) 9

15. If x and y are related by 2x+3y+4 = 0 and SD of x is 6, then SD of y is

(a) 22 (b) 4 (c) $\sqrt{5}$ (d) 9.

16. The quartiles of a variable are 45, 52 and 65 respectively. Its quartile deviation is

(a) 10 (b) 20 (c) 25 (d) 8.30.

17. If x and y are related as 3x+4y = 20 and the quartile deviation of x is 12, then the quartile deviation of y is

(a) 16 (b) 14 (c) 10 (d) 9.

18. If the SD of the 1st n natural numbers is 2, then the value of n must be

(a) 2 (b) 7 (c) 6 (d) 5.

19. If x and y are related by y = 2x+ 5 and the SD and AM of x are known to be 5 and 10 respectively, then the coefficient of variation is

(a) 25 (b) 30 (c) 40 (d) 20.

20. The mean and SD for a,b and 2 are 3 and $\frac{2}{\sqrt{3}}$ respectively, The value of ab would be

(a) 5 (b) 6 (c) 12 (d) 3.

Set C

Write down the correct answer. Each question carries 5 marks.

1. What is the mean deviation about mean for the following distribution?

Variable :	5	10	15	20	25	30
Frequency:	3	4	6	5	3	2

(a) 6.00 (b) 5.93 (c) 6.07 (d) 7.20

2. What is the mean deviation about median for the following data?

X :	3	5	7	9	11	13	15
F :	2	8	9	16	14	7	4

(a) 2.50 (b) 2.46 (c) 2.43 (d) 2.37

3. What is the coefficient of mean deviation for the following distribution of heights? Take deviation from AM.

Height in inches:	60-62	63-65	66-68	69-71	72-74
No. of students:	5	22	28	17	3

(a) 2.30 inches (b) 3.45 (c) 3.82 (d) 2.48 inches

4. The mean deviation of weights about median for the following data:

Weight (lb) :	131-140	141-150	151-160	161-170	171-180	181-190
No. of persons :	3	8	13	15	6	5

Is given by

(a) 10.97 (b) 8.23 (c) 9.63 (d) 11.45.

5. What is the standard deviation from the following data relating to the age distribution of 200 persons?

Age (year) :	20	30	40	50	60	70	80
No. of people:	13	28	31	46	39	23	20

(a) 15.29 (b) 16.87 (c) 18.00 (d) 17.52

6. What is the coefficient of variation for the following distribution of wages?

Daily Wages (Rs.)	30 – 40	40 – 50	50 – 60	60 – 70	70 – 80	80 – 90
No. of workers	17	28	21	15	13	6

(a) Rs.14.73 (b) 14.73 (c) 26.93 (d) 20.82

7. Which of the following companies A and B is more consistent so far as the payment of dividend is concerned ?

Dividend paid by A :	5	9	6	12	15	10	8	10
Dividend paid by B :	4	8	7	15	18	9	6	6

(a) A (b) B (c) Both (a) and (b) (d) Neither (a) nor (b)

8. The mean and SD for a group of 100 observations are 65 and 7.03 respectively. If 60 of these observations have mean and SD as 70 and 3 respectively, what is the SD for the group comprising 40 observations?

(a) 16 (b) 25 (c) 4 (d) 2

9. If two samples of sizes 30 and 20 have means as 55 and 60 and variances as 16 and 25 respectively, then what would be the SD of the combined sample of size 50?

(a) 5.00 (b) 5.06 (c) 5.23 (d) 5.35

10. The mean and SD of a sample of 100 observations were calculated as 40 and 5.1 respectively by a CA student who took one of the observations as 50 instead of 40 by mistake. The current value of SD would be

(a) 4.90 (b) 5.00 (c) 4.88 (d) 4.85.

11. The value of appropriate measure of dispersion for the following distribution of daily wages

Wages (Rs.) :	Below 30	30-39	40-49	50-59	60-79	Above 80
No. of workers	5	7	18	32	28	10

is given by

(a) Rs.11.03 (b) Rs.10.50 (c) 11.68 (d) Rs.11.68.

ANSWERS

Set A					
1 (d)	2 (a)	3 (b)	4 (d)	5 (d)	6 (a)
7 (d)	8 (b)	9 (b)	10 (d)	11 (c)	12 (b)
13 (a)	14 (c)	15 (b)	16 (d)	17 (d)	18 (c)
19 (b)	20 (d)	21 (a)			
Set B					
1 (d)	2 (c)	3 (c)	4 (b)	5 (c)	6 (a)
7 (c)	8 (a)	9 (b)	10 (b)	11 (b)	12 (b)
13 (a)	14 (a)	15 (b)	16 (a)	17 (d)	18 (b)
19 (c)	20 (a)				
Set C					
1 (c)	2 (d)	3 (b)	4 (a)	5 (b)	6 (c)
7 (a)	8 (c)	9 (b)	10 (b)	11 (a)	

ADDITIONAL QUESTION BANK

1. The number of measures of central tendency is

   (a) two (b) three (c) four (d) five

2. The words "mean" or "average" only refer to

   (a) A.M (b) G.M (c) H.M (d) none

3. ——————— is the most stable of all the measures of central tendency.

   (a) G.M (b) H.M (c) A.M (d) none.

4. Mean is of ———— types.

   (a) 3 (b) 4 (c) 8 (d) 5

5. Weighted A.M is related to

   (a) G.M (b) frequency (c) H.M (d) none.

6. Frequencies are also called weights.

   (a) True (b) false (c) both (d) none

7. The algebraic sum of deviations of observations from their A.M is

   (a) 2 (b) -1 (c) 1 (d) 0

8. G.M of a set of n observations is the ———— root of their product.

   (a) n/2 th (b) (n+1)th (c) nth (d) (n -1)th

9. The algebraic sum of deviations of 8,1,6 from the A.M viz.5 is

   (a) -1 (b) 0 (c) 1 (d) none

10. G.M of 8, 4,2 is

    (a) 4 (b) 2 (c) 8 (d) none

11. ———————— is the reciprocal of the A.M of reciprocal of observations.

    (a) H.M (b) G.M (c) both (d) none

12. A.M is never less than G.M

    (a). True (b) false (c) both (d) none

13. G.M is less than H.M

    (a) true (b) false (c) both (d) none

14. The value of the middlemost item when they are arranged in order of magnitude is called

    (a) standard deviation (b) mean (c) mode (d) median

15. Median is unaffected by extreme values.

    (a) true (b) false (c) both (d) none

16. Median of 2,5,8,4,9,6,71 is

(a) 9 (b) 8 (c) 5 (d) 6

17. The value which occurs with the maximum frequency is called

(a) median (b) mode (c) mean (d) none

18. In the formula Mode = $L_1 + (d_1 \text{ X } c)/ (d_1 + d_2)$

d_1 is the difference of frequencies in the modal class & the ———— class.

(a) preceding (b) following (c) both (d) none

19. In the formula Mode = $L_1 + (d_1 \text{ X } c)/ (d_1 + d_2)$

d_2 is the difference of frequencies in the modal class & the ———— class.

(a) preceding (b) succeeding (c) both (d) none

20. In formula of median for grouped frequency distribution N is

(a) total frequency (b) frequency density
(c) frequency (d) cumulative frequency

21. When all observations occur with equal frequency ———— does not exit.

(a) median (b) mode (c) mean (d) none

22. Mode of the observations 2,5,8,4,3,4,4,5,2,4,4 is

(a) 3 (b) 2 (c) 5 (d) 4

23. For the observations 5,3,6,3,5,10,7,2 , there are ———— modes.

(a) 2 (b) 3 (c) 4 (d) 5

24. ———— of a set of observations is defined to be their sum, divided by the no. of observations.

(a) H.M (b) G.M (c) A.M (d) none

25. Simple average is sometimes called

(a) weighted average (b) unweighted average
(c) relative average (d) none

26. When a frequency distribution is given, the frequencies themselves treated as weights.

(a) True (b) false (c) both (d) none

27. Each value is considered only once for

(a) simple average (b) weighted average
(c) both (d) none

28. Each value is considered as many times as it occurs for

(a) simple average (b) weighted average
(c) both (d) none

29. Multiplying the values of the variable by the corresponding weights and then dividing the sum of products by the sum of weights is

(a) simple average (b) weighted average
(c) both (d) none

30. Simple & weighted average are equal only when all the weights are equal.

(a) True (b) false (c) both (d) none

31. The word " average " used in "simple average " and "weighted average " generally refers to

(a) median (b) mode (c) A.M , G.M or H.M (d) none

32. ———— average is obtained on dividing the total of a set of observations by their number

(a) simple (b) weighted (c) both (d) none

33. Frequencies are generally used as

(a) range (b) weights (c) mean (d) none

34. The total of a set of observations is equal to the product of their number and the

(a) A.M (b) G.M (c) H.M (d) none

35. The total of the deviations of a set of observations from their A.M is always

(a) 0 (b) 1 (c) -1 (d) none

36. Deviation may be positive or negative or zero

(a) true (b) false (c) both (d) none

37. The sum of the squares of deviations of a set of observations has the smallest value, when the deviations are taken from their

(a) A.M (b) H.M (c) G.M (d) none

38. For a given set of positive observations H.M is less than G.M

(a) true (b) false (c) both (d) none

39. For a given set of positive observations A.M is greater than G.M

(a) true (b) false (c) both (d) none

40. Calculation of G.M is more difficult than

(a) A.M (b) H.M (c) median (d) none

41. ————— has a limited use

(a) A.M (b) G.M (c) H.M (d) none

42. A.M of 8,1,6 is

(a) 5 (b) 6 (c) 4 (d) none

43. ———— can be calculated from a frequency distribution with open end intervals

(a) Median (b) Mean (c) Mode (d) none

44. The values of all items are taken into consideration in the calculation of

(a) median (b) mean (c) mode (d) none

45. The values of extreme items do not influence the average in case of

(a) median (b) mean (c) mode (d) none

46. In a distribution with a single peak and moderate skewness to the right, it is closer to the concentration of the distribution in case of

(a) mean (b) median (c) both (d) none

47. If the variables x & z are so related that z = ax + b for each $x = x_i$ where a & b are constants, then $\bar{z} = a\bar{x} + b$

(a) true (b) false (c) both (d) none

48. G.M is defined only when

(a) all observations have the same sign and none is zero

(b) all observations have the different sign and none is zero

(c) all observations have the same sign and one is zero

(d) all observations have the different sign and one is zero

49. ———— is useful in averaging ratios, rates and percentages.

(a) A.M (b) G.M (c) H.M (d) none

50. G.M is useful in construction of index number.

(a) true (b) false (c) both (d) none

51. More laborious numerical calculations involves in G.M than A.M

(a) True (b) false (c) both (d) none

52. H.M is defined when no observation is

(a) 3 (b) 2 (c) 1 (d) 0

53. When all values occur with equal frequency, there is no

(a) mode (b) mean (c) median (d) none

54. ———— cannot be treated algebraically

(a) mode (b) mean (c) median (d) none

55. For the calculation of ———— , the data must be arranged in the form of a frequency distribution.

(a) median (b) mode (c) mean (d) none

56. ———— is equal to the value corresponding to cumulative frequency

(a) mode (b) mean (c) median (d) none

57. ————— is the value of the variable corresponding to the highest frequency

(a) mode (b) mean (c) median (d) none

58. The class in which mode belongs is known as

(a) median class (b) mean class (c) modal class (d) none

59. The formula of mode is applicable if classes are of ————— width.

(a) equal (b) unequal (c) both (d) none

60. For calculation of ——— we have to construct cumulative frequency distribution

(a) mode (b) median (c) mean (d) none

61. For calculation of ——— we have to construct a grouped frequency distribution

(a) median (b) mode (c) mean (d) none

62. Relation between mean, median & mode is

(a) mean - mode = 2 (mean—median) (b) mean - median = 3 (mean—mode)
(c) mean - median = 2 (mean—mode) (d) mean - mode = 3 (mean—median)

63. When the distribution is symmetrical, mean, median and mode

(a) coincide (b) do not coincide (c) both (d) none

64. Mean, median & mode are equal for the

(a) Binomial distribution (b) Normal distribution
(c) both (d) none

65. In most frequency distributions, it has been observed that the three measures of central tendency viz.mean , median & mode ,obey the approximate relation , provided the distribution is

(a) very skew (b) not very skew (c) both (d) none

66. —————— divides the total number of observations into two equal parts.

(a) mode (b) mean (c) median (d) none

67. Measures which are used to divide or partition the observations into a fixed number of parts are collectively known as

(a) partition values (b) quartiles (c) both (d) none

68. The middle most value of a set of observations is

(a) median (b) mode (c) mean (d) none

69. The number of observations smaller than ——— is the same as the number larger than it.

(a) median (b) mode (c) mean (d) none

70. ———— is the value of the variable corresponding to cumulative frequency N /2

(a) mode (b) mean (c) median (d) none

71. ———————— divide the total no. observations into 4 equal parts.

(a) median (b) deciles (c) quartiles (d) percentiles

72. ———————— quartile is known as Upper quartile

(a) First (b) Second (c) Third (d) none

73. Lower quartile is

(a) first quartile (b) second quartile (c) upper quartile (d) none

74. The number of observations smaller than lower quartile is the same as the no. lying between lower and middle quartile.

(a) true (b) false (c) both (d) none

75. ———— are used for measuring central tendency , dispersion & skewness.

(a) Median (b) Deciles (c) Percentiles (d) Quartiles.

76. The second quartile is known as

(a) median (b) lower quartile (c) upper quartile (d) none

77. The lower & upper quartiles are used to define

(a) standard deviation (b) quartile deviation
(c) both (d) none

78. Three quartiles are used in

(a) Pearson"s formula (b) Bowley"s formula
(c) both (d) none

79. Less than First quartile , the frequency is equal to

(a) N /4 (b) 3N /4 (c) N /2 (d) none

80. Between first & second quartile, the frequency is equal to

(a) 3N/4 (b) N /2 (c) N /4 (d) none

81. Between second & upper quartile, the frequency is equal to

a) 3N/4 (b) N /4 (c) N /2 (d) none

82. Above upper quartile, the frequency is equal to

(a) N /4 (b) N /2 (c) 3N /4 (d) none

83. Corresponding to first quartile, the cumulative frequency is

(a) N /2 (b) N / 4 (c) 3N /4 (d) none

84. Corresponding to second quartile, the cumulative frequency is

(a) N /4 (b) 2 N / 4 (c) 3N /4 (d) none

85. Corresponding to upper quartile, the cumulative frequency is

(a) 3N/4 (b) N / 4 (c) 2N /4 (d) none

86. The values which divide the total number of observations into 10 equal parts are

(a) quartiles (b) percentiles (c) deciles (d) none

87. There are ————— deciles.

(a) 7 (b) 8 (c) 9 (d) 10

88. Corresponding to first decile, the cumulative frequency is

(a) N/10 (b) 2N /10 (c) 9N /10 (d) none

89. Fifth decile is equal to

(a) mode (b) median (c) mean (d) none

90. The values which divide the total number of observations into 100 equal parts is

(a) percentiles (b) quartiles (c) deciles (d) none

91. Corresponding to second decile, the cumulative frequency is

(a) N /10 (b) 2N /10 (c) 5N /10 (d) none

92. There are ———— percentiles.

(a) 100 (b) 98 (c) 97 (d) 99

93. 10^{th} percentile is equal to

(a) 1^{st} decile (b) 10^{th} decile (c) 9^{th} decile (d) none

94. 50^{th} percentile is known as

(a) 50^{th} decile (b) 50^{th} quartile (c) mode (d) median

95. 20^{th} percentile is equal to

(a) 19^{th} decile (b) 20^{th} decile (c) 2^{nd} decile (d) none

96. (3^{rd} quartile —— 1^{st} quartile)/2 is

(a) skewness (b) median (c) quartile deviation (d) none

97. 1^{st} percentile is less than 2^{nd} percentile.

(a) true (b) false (c) both (d) none

98. 25^{th} percentile is equal to

(a) 1^{st} quartile (b) 25^{th}quartile (c) 24^{th} quartile (d) none

99. 90^{th} percentile is equal to

(a) 9^{th} quartile (b) 90^{th} decile (c) 9^{th} decile (d) none

100. 1^{st} decile is greater than 2^{nd} decile

(a) True (b) false (c) both (d) none

101. Quartile deviation is a measure of dispersion.

(a) true (b) false (c) both (d) none

102. To define quartile deviation we use

(a) lower & middle quartiles (b) lower & upper quartiles
(c) upper & middle quartiles (d) none

102. Calculation of quartiles, deciles ,percentiles may be obtained graphically from

(a) Frequency Polygon (b) Histogram (c) Ogive (d) none

103. 7^{th} decile is the abscissa of that point on the Ogive whose ordinate is

(a) 7N/10 (b) 8N /10 (c) 6N /10 (d) none

104. Rank of median is

(a) (n+ 1)/2 (b) (n+ 1)/4 (c) 3(n + 1)/4 (d) none

105. Rank of 1^{st} quartile is

(a) (n+ 1)/2 (b) (n+ 1)/4 (c) 3(n + 1)/4 (d) none

106. Rank of 3rd quartile is

(a) 3(n+ 1)/4 (b) (n+ 1)/4 (c) (n + 1)/2 (d) none

107. Rank of k th decile is

(a) (n+ 1)/2 (b) (n+ 1)/4 (c) (n + 1)/10 (d) k(n +1)/10

108. Rank of k th percentile is

(a) (n+ 1)/100 (b) k(n+ 1)/10 (c) k(n + 1)/100 (d) none

109. ———————— is equal to value corresponding to cumulative frequency (N + 1)/2 from simple frequency distribution

(a) Median (b) 1^{st} quartile (c) 3^{rd} quartile (d) 4^{th} quartile

110. ———— is equal to the value corresponding to cumulative frequency (N + 1)/4 from simple frequency distribution

(a) Median (b) 1^{st} quartile (c) 3^{rd} quartile (d) 1^{st} decile

111. ———— is equal to the value corresponding to cumulative frequency 3 (N + 1)/4 from simple frequency distribution

(a) Median (b) 1^{st} quartile (c) 3^{rd} quartile (d) 1^{st} decile

112. ———— is equal to the value corresponding to cumulative frequency k (N + 1)/10 from simple frequency distribution

(a) Median (b) kth decile (c) kth percentile (d) none

113. ———— is equal to the value corresponding to cumulative frequency k(N + 1)/100 from simple frequency distribution

(a) kth decile (b) kth percentile (c) both (d) none

114. For grouped frequency distribution ———— is equal to the value corresponding to cumulative frequency N /2

(a) median (b) 1st quartile (c) 3rd quartile (d) none

115. For grouped frequency distribution ———— is equal to the value corresponding to cumulative frequency N /4

(a) median (b) 1st quartile (c) 3rd quartile (d) none

116. For grouped frequency distribution ———— is equal to the value corresponding to cumulative frequency 3N /4

(a) median (b) 1st quartile (c) 3rd quartile (d) none

117. For grouped frequency distribution ———— is equal to the value corresponding to cumulative frequency kN/10

(a) median (b) kth percentile (c) kth decile (d) none

118. For grouped frequency distribution ———— is equal to the value corresponding to cumulative frequency kN /100

(a) kth quartile (b) kth percentile (c) kth decile (d) none

119. In Ogive, abscissa corresponding to ordinate N/2 is

(a) median (b) 1st quartile (c) 3rd quartile (d) none

120. In Ogive, abscissa corresponding to ordinate N/4 is

(a) median (b) 1st quartile (c) 3rd quartile (d) none

121. In Ogive, abscissa corresponding to ordinate 3N/4 is

(a) median (b) 3rd quartile (c) 1st quartile (d) none

122. In Ogive, abscissa corresponding to ordinate ———— is kth decile.

(a) kN/10 (b) kN/100 (c) kN/50 (d) none

123. In Ogive , abscissa corresponding to ordinate ———— is kth percentile.

(a) kN/10 (b) kN/100 (c) kN/50 (d) none

124. For 899 999 391 384 590 480 485 760 111 240
Rank of median is

(a) 2.75 (b) 5.5 (c) 8.25 (d) none

125. For 333 999 888 777 666 555 444
Rank of 1st quartile is

(a) 3 (b) 1 (c) 2 (d) 7

126. For 333 999 888 777 1000 321 133
Rank of 3rd quartile is

(a) 7 (b) 4 (c) 5 (d) 6

127. Price per kg.(Rs.) : 45 50 35; Kgs.Purchased : 100 40 60 Total frequency is

(a) 300 (b) 100 (c) 150 (d) 200

128. The length of a rod is measured by a tape 10 times . You are to estimate the length of the rod by averaging these 10 determinations.

What is the suitable form of average in this case——

(a) A.M (b) G.M (c) H.M (d) none

129. A person purchases 5 rupees worth of eggs from 10 different markets.You are to find the average no. of eggs per rupee for all the markets taken together. What is the suitable form of average in this case——

(a) A.M (b) G.M (c) H.M (d) none

130. You are given the population of India for the courses of 1981 & 1991. You are to find the population of India at the middle of the period by averaging these population figures, assuming a constant rate of increase of population.

What is the suitable form of average in this case—

(a) A.M (b) G.M (c) H.M (d) none

131. ———————— is least affected by sampling fluctions.

(a) Standard deviation (b) Quartile deviation
(c) both (d) none

132. "Root –Mean Square Deviation from Mean" is

(a) Standard deviation (b) Quartile deviation

(c) both (d) none

133. Standard Deviation is

(a) absolute measure (b) relative measure (c) both (d) none

134. Coefficient of variation is

(a) absolute measure (b) relative measure (c) both (d) none

135. ———————— deviation is called semi-interquartile range.

(a) Percentile (b) Standard (c) Quartile (d) none

136. —————————— Deviation is defined as half the difference between the lower & upper quartiles.

(a) Quartile (b) Standard (c)both (d) none

137. Quartile Deviation for the data 1,3,4,5,6,6,10 is

(a) 3 (b) 1 (c) 6 (d)1.5

138. Coefficient of Quartile Deviation is

(a) (Quartile Deviation x 100)/Median (b) (Quartile Deviation x 100)/Mean
(c) (Quartile Deviation x 100) /Mode (d) none

139. Mean for the data 6,4,1,6,5,10,3 is

(a) 7 (b) 5 (c) 6 (d) none

140. Coefficient of variation = (Standard Deviation x 100)/Mean

(a) true (b) false (c) both (d) none

141. If mean = 5, Standard deviation = 2.6 then the coefficient of variation is

(a) 49 (b) 51 (c) 50 (d) 52

142. If median = 5, Quartile deviation = 1. 5 then the coefficient of quartile deviation is

(a) 33 (b) 35 (c) 30 (d) 20

143. A.M of 2,6,4,1,8,5,2 is

(a) 4 (b) 3 (c) 4 (d) none

144. Most useful among all measures of dispersion is

(a) S.D (b) Q.D (c) Mean deviation (d) none

145. For the observations 6,4,1,6,5,10,4,8 Range is

(a) 10 (b) 9 (c) 8 (d) none

146. A measure of central tendency tries to estimate the

(a) central value (b) lower value (c) upper value (d) none

147. Measures of central tendency are known as

(a) differences (b) averages (c) both (d) none

148. Mean is influenced by extreme values.

(a) true (b) false (c) both (d) none

149. Mean of 6,7,11,8 is

(a) 11 (b) 6 (c) 7 (d) 8

150. The sum of differences between the actual values and the arithmetic mean is

(a) 2 (b) -1 (c) 0 (d) 1

151. When the algebraic sum of deviations from the arithmetic mean is not equal to zero, the figure of arithmetic mean ———— correct.

(a) is (b) is not (c) both (d) none

152. In the problem

No. of shirts :	30—32	33—35	36—38	39—41	42—44
No. of persons :	15	14	42	27	18

The assumed mean is

(a) 34 (b) 37 (c) 40 (d) 43

153. In the problem

Size of items :	1—3	3—8	8—15	15—26
Frequency :	5	10	16	15

The assumed mean is

(a) 20.5 (b) 2 (c) 11.5 (d) 5.5

154. The average of a series of over-lapping averages, each of which is based on a certain number of item within a series is known as

(a) moving average (b) weighted average
(c) simple average (d) none

155. ———— averages is used for smoothening a time series.

(a) moving average (b) weighted average
(c) simple average (d) none

156. Pooled Mean is also called

(a) Mean (b) Geometric Mean (c) Grouped Mean (d) none

157. Half of the numbers in an ordered set have values less than the ———— and half will have values greater than the ———— .

(a) mean, median (b)median, median (c) mode ,mean (d) none.

158. The median of 27,30,26,44,42,51,37 is

(a) 30 (b) 42 (c) 44 (d) 37

159. For an even number of values the median is the

(a) average of two middle values (b) middle value
(c) both (d) none

160. In the case of a continuous frequency distribution , the size of the ———— item indicates class interval in which the median lies.

(a) (n-1)/2 th (b) (n+ 1)/2 th (c) n/2th (d) none

161. The deviations from median are ———— if negative signs are ignored as compared to other measures of central tendency.

(a) minimum (b) maximum (c) same (d) none

162. Ninth Decile lies in the class interval of the item

(a) n/9 (b) 9n/10 (c) 9n/20 (d) none item.

163. Ninety Ninth Percentile lies in the class interval of the item

(a) 99n/100 (b) 99n/10 (c) 99n/200 (d) none item.

164. ———— is the value of the variable at which the concentration of observation is the densest.

(a) mean (b) median (c) mode (d) none

165. Height in cms : 60—62 63—65 66—68 69—71 72—74

No. of students : 15 118 142 127 18

Modal group is

(a) 66—68 (b) 69—71 (c) 63—65 (d) none

166. A distribution is said to be symmetrical when the frequency rises & falls from the highest value in the ———— proportion.

(a) unequal (b) equal (c) both (d) none

167. ———— always lies in between the arithmetic mean & mode.

(a) G.M (b) H.M (c) Median (d) none

168. Logarithm of G.M is the ———— of logarithms of the different values.

(a) weighted mean (b) simple mean (c) both (d) none

169. ———— is not much affected by fluctuations of sampling.

(a) A.M (b) G.M (c) H.M (d) none

170. The data 1,2,4,8,16 are in

(a) Arithmetic progression (b) Geometric progression

(c) Harmonic progression (d) none

171. ———— & ———— can not be calculated if any observation is zero.

(a) G.M & A.M (b) H.M & A.M (c) H.M & G. M (d) None.

172. ———— & ———— are called ratio averages.

(a) H.M & G.M (b) H. M & A.M (c) A.M & G.M (d) none

173. ———— is a good substitute to a weighted average.

(a) A.M (b) G.M (c) H.M (d) none

174. For ordering shoes of various sizes for resale, a ———— size will be more appropriate.

(a) median (b) modal (c) mean (d) none

175. ——————— is called a positional measure.

(a) mean (b) mode (c) median (d) none

176. 50% of actual values will be below & 50% of will be above ——————

(a) mode (b) median (c) mean (d) none

177. Extreme values have ———— effect on mode.

(a) high (b) low (c) no (d) none

178. Extreme values have ———— effect on median.

(a) high (b) low (c) no (d) none

179. Extreme values have ———— effect on A.M.

(a) greatest (b) least (c) some (d) none

180. Extreme values have ———— effect on H.M.

(a) least (b) greatest (c) medium (d) none

181. ——————— is used when representation value is required & distribution is asymmetric.

(a) mode (b) mean (c) median (d) none

182. ——————— is used when most frequently occurring value is required (discrete variables).

(a) mode (b) mean (c) median (d) none

183. ——————— is used when rate of growth or decline required.

(a) mode (b) A.M (c) G.M (d) none

184. In finding ————, the distribution has open-end classes.

(a) median (b) mean (c) standard deviation (d) none

185. The cumulative frequency distribution is used for

(a) median (b) mode (c) mean (d) none

186. In ——— the quantities are in ratios.

(a) A.M (b) G.M (c) H.M (d) none

187. —————— is used when variability has also to be calculated.

(a) A.M (b) G.M (c) H.M (d) none

188. —————— is used when the sum of absolute deviations from the average should be least.

(a) Mean (b) Mode (c) Median (d) None

189. —————— is used when sampling variability should be least.

(a) Mode (b) Median (c) Mean (d) none

190. —————— is used when distribution pattern has to be studied at varying levels.

(a) A.M (b) Median (c) G.M (d) none

191. The average discovers

(a) uniformity in variability (b) variability in uniformity of distribution
(c) both (d) none

192. The average has relevance for

(a) homogeneous population (b) heterogeneous population
(c) both (d) none

193. The correction factor is applied in

(a) inclusive type of distribution (b) exclusive type of distribution
(c) both (d) none

194. "Mean has the least sampling variability" prove the mathematical property of mean

(a) True (b) false (c) both (d) none

195. "The sum of deviations from the mean is zero" —— is the mathematical property of mean

(a) True (b) false (c) both (d) none

196. "The mean of the two samples can be combined" — is the mathematical property of mean

(a) True (b) false (c) both (d) none

197. "Choices of assumed mean does not affect the actual mean"— prove the mathematical property of mean

(a) True (b) false (c) both (d) none

198. "In a moderately asymmetric distribution mean can be found out from the given values of median & mode"— is the mathematical property of mean

(a) True (b) false (c) both (d) none

199. The mean wages of two companies are equal. It signifies that the workers of both the companies are equally well-off.

(a) True (b) false (c) both (d) none

200. The mean wage in factory A is Rs.6000 whereas in factory B it is Rs.5500. It signifies that factory A pays more to all its workers than factory B.

(a) True (b) false (c) both (d) none

201. Mean of 0,3,5,6,7 , 9,12,0,2 is

(a) 4.9 (b) 5.7 (c) 5.6 (d) none

202. Median of 15,12,6,13,12,15,8,9 is

(a) 13 (b) 8 (c) 12 (d) 9

203. Median of 0.3,5,6,7,9,12,0,2 is

(a) 7 (b) 6 (c) 3 (d) 5

204. Mode of 0,3,5,6,7,9,12,0,2 is

(a) 6 (b) 0 (c) 3 (d) 5

205. Mode 0f 15,12,5,13,12,15,8,8,9,9,10,15 is

(a)15 (b) 12 (c) 8 (d) 9

206. Median of 40,50,30,20,25,35,30,30,20,30 is

(a) 25 (b) 30 (c) 35 (d) none

207. Mode of 40,50,30,20,25,35,30,30,20,30 is

(a) 25 (b) 30 (c) 35 (d) none

208. ———————— in particular helps in finding out the variability of the data.

(a) Dispersion (b) Median (c) Mode (d) None

209. Measures of central tendency are called averages of the ———order.

(a) 1^{st} (b) 2^{nd} (c) 3^{rd} (d) none

210. Measures of dispersion are called averages of the ———order.

(a) 1^{st} (b) 2^{nd} (c) 3^{rd} (d) none

211. In measuring dispersion, it is necessary to know the amount of ———— & the degree of —————.

(a) variation, variation (b) variation, median
(c) median, variation (d) none

212. The amount of variation is designated as ———————— measure of dispersion.

(a) relative (b) absolute (c) both (d) none

213. The degree of variation is designated as ———————— measure of dispersion.

(a) relative (b) absolute (c) both (d) none

214. For purposes of comparison between two or more series with varying size or no. of items, varying central values or units of calculation, only ———————— measures can be used.

(a) absolute (b) relative (c) both (d) none

215. The relation Relative range = Absolute range/Sum of the two extremes. is

(a) True (b) false (c) both (d) none

216. The relation Absolute range = Relative range/Sum of the two extremes is

(a) True (b) false (c) both (d) none

217. In quality control ———— is used as a substitute for standard deviation.

(a) mean deviation (b) median (c) range (d) none

218. ———————— factor helps to know the value of standard deviation.

(a) Correction (b) Range (c) both (d) none

219. ———————— is extremely sensitive to the size of the sample

(a) Range (b) Mean (c) Median (d) Mode

220. As the sample size increases, —————— also tends to increase.

(a) Range (b) Mean (c) Median (d) Mode

221. As the sample size increases, range also tends to increase though not proportionately.

(a) true (b) false (c) both (d) none.

222. As the sample size increases, range also tends to

(a) decrease (b) increase (c) same (d) none

223. The dependence of range on extreme items can be avoided by adopting

(a) standard deviation (b) mean deviation (c) quartile deviation (d) none

224. Quartile deviation is called

(a) semi inter quartile range (b) quartile range (c) both (d) none

225. When 1st quartile = 20, 3rd quartile = 30, the value of quartile deviation is

(a) 7 (b) 4 (c) -5 (d) 5

226. $(Q_3 - Q_1)/(Q_3 + Q_1)$ is

(a) coefficient of Quartile Deviation (b) coefficient of Mean Deviation
(c) coefficient of Standard deviation (d) none

227. Standard deviation is denoted by
(a) σ^2 (b) σ (c) $\sqrt{\sigma}$ (d) none

228. The square of standard deviation is known as

(a) variance (b) standard deviation
(c) mean deviation (d) none

229. Mean of 25,32,43,53,62,59,48,31,24,33 is

(a) 44 (b) 43 (c) 42 (d) 41

230. For the following frequency distribution

Class interval :	10—20	20—30	30—40	40—50	50—60	60—70
Frequency :	20	9	31	18	10	9

assumed mean can be taken as

(a) 55 (b) 45 (c) 35 (d) none

231. The value of the standard deviation does not depend upon the choice of the origin.

(a) True (b) false (c) both (d) none

232. Coefficient of standard deviation is

(a) S.D/Median (b) S.D/Mean (c) S.D/Mode (d) none

233. The value of the standard deviation will change if any one of the observations is changed.

(a). True (b) false (c) both (d) none

234. When all the values are equal then variance & standard deviation would be

(a) 2 (b) -1 (c) 1 (d) 0

235. For values lie close to the mean, the standard deviations are

(a) big (b) small (c) moderate (d) none

236. If the same amount is added to or subtracted from all the values, variance & standard deviation shall

(a) changed (b) unchanged (c) both (d) none

237. If the same amount is added to or subtracted from all the values, the mean shall increase or decrease by the ———— amount

(a) big (b) small (c) same (d) none

238. If all the values are multiplied by the same quantity, the ————— & ———— also would be multiple of the same quantity.

(a) mean, standard deviation (b) mean , median
(c) mean, mode (d) median , deviations

239. For a moderately non-symmetrical distribution, Mean deviation = 4/5 of standard deviation

(a) True (b) false (c) both (d) none

240. For a moderately non-symmetrical distribution, Quartile deviation = Standard deviation/3

(a) True (b) false (c) both (d) none

241. For a moderately non-symmetrical distribution, Probable error of standard deviation = Standard deviation/3

(a) True (b) false (c) both (d) none

242. Quartile deviation = Probable error of Standard deviation.

(a) True (b) false (c) both (d) none

243. Coefficient of Mean Deviation is

(a) Mean deviation x 100/Mean or mode (b) Standard deviation x 100/Mean or median

(c) Mean deviation x 100/Mean or median (d) none

244. Coefficient of Quartile Deviation = Quartile Deviation x 100/Median

(a) True (b) false (c) both (d) none

245. Karl Pearson's measure gives

(a) coefficient of Mean Variation (b) coefficient of Standard deviation
(c) coefficient of variation (d) none

246. In ——— range has the greatest use.

(a) Time series (b) quality control (c) both (d) none

247. Mean is an absolute measure & standard deviation is based upon it. Therefore standard deviation is a relative measure.

(a) True (b) false (c) both (d) none

248. Semi—quartile range is one-fourth of the range in a normal symmetrical distribution.

(a) Yes (b) No (c) both (d) none

249. Whole frequency table is needed for the calculation of

(a) range (b) variance (c) both (d) none

250. Relative measures of dispersion make deviations in similar units comparable.

(a) True (b) false (c) both (d) none

251. Quartile deviation is based on the

(a) highest 50 % (b) lowest 25 %
(c) highest 25 % (d) middle 50% of the item.

252. S.D is less than Mean deviation

(a) True (b) false (c) both (d) none

253. Coefficient of variation is independent of the unit of measurement.

(a) True (b) false (c) both (d) none

254. Coefficient of variation is a relative measure of

(a) mean (b) deviation (c) range (d) dispersion.

255. Coefficient of variation is equal to

(a) Standard deviation x 100 / median (b) Standard deviation x 100 / mode
(c) Standard deviation x 100 / mean (d) none

256. Coefficient of Quartile Deviation is equal to

(a) Quartile deviation x 100 / median (b) Quartile deviation x 100 / mean
(c) Quartile deviation x 100 / mode (d) none

257. If each item is reduced by 15 A.M is

(a) reduced by 15 (b) increased by 15 (c) reduced by 10 (d) none

258. If each item is reduced by 10, the range is

(a) increased by 10 (b) decreased by 10 (c) unchanged (d) none

259. If each item is reduced by 20, the standard deviation

(a) increased (b) decreased (c) unchanged (d) none

260. If the variables are increased or decreased by the same amount the standard deviation is

(a) decreased (b) increased (c) unchanged (d) none

261. If the variables are increased or decreased by the same proportion, the standard deviation changes by

(a) same proportion (b) different proportion (c) both (d) none

262. The mean of the 1st n natural no. is

(a) n/2 (b) (n-1)/2 (c) (n+1)/2 (d) none

263. If the class interval is open-end then it is difficult to find

(a) frequency (b) A.M (c) both (d) none

264. Which one is true—

(a) A.M = assumed mean + arithmetic mean of deviations of terms

(b) G.M = assumed mean + arithmetic mean of deviations of terms

(c) Both (d) none

265. If the A.M of any distribution be 25 & one term is 18. Then the deviation of 18 from A.M is

(a) 7 (b) -7 (c) 43 (d) none

266. For finding A.M in Step—deviation method, the class intervals should be of

(a) equal lengths (b) unequal lengths (c) maximum lengths (d) none

267. The sum of the squares of the deviations of the variable is ———— when taken about A.M

(a) maximum (b) zero (c) minimum (d) none

268. The A.M of 1,3,5,6,x,10 is 6 . The value of x is

(a) 10 (b) 11 (c) 12 (d) none

269. The G.M of 2 & 8 is

(a) 2 (b) 4 (c) 8 (d) none

270. (n+1)/2 th term is median if n is

(a) odd (b) even (c) both (d) none

271. For the values of a variable 5,2,8,3,7,4, the median is

(a) 4 (b) 4.5 (c) 5 (d) none

272. The abscissa of the maximum frequency in the frequency curve is the

(a) mean (b) median (c) mode (d) none

273. variable : 2 3 4 5 6 7
no. of men : 5 6 8 13 7 4

variable :	2	3	4	5	6	7
no. of men :	5	6	8	13	7	4

Mode is

(a) 6 (b) 4 (c) 5 (d) none

274. The class having maximum frequency is called

(a) modal class (b) median class (c) mean class (d) none

275. For determination of mode, the class intervals should be

(a) overlapping (b) maximum (c) minimum (d) none

276. First Quartile lies in the class interval of the

(a) n/2th item (b) $n/4^{th}$ item (c) $3n/4^{th}$ item (d) $n/10^{th}$ item

277. The value of a variate that occur most often is called

(a) median (b) mean (c) mode (d) none

278. For the values of a variable 3,1,5,2,6,8,4 the median is

(a) 3 (b) 5 (c) 4 (d) none

279. If y = 5 x - 20 & $\overline{x}$ = 30 then the value of $\overline{y}$ is

(a) 130 (b) 140 (c) 30 (d) none

280. If y = 3 x - 100 and $\overline{x}$ = 50 then the value of $\overline{y}$ is

(a) 60 (b) 30 (c) 100 (d) 50

281. The median of the numbers 11,10,12,13,9 is

(a) 12.5 (b) 12 (c) 10.5 (d) 11

282. The mode of the numbers 7,7,7,9,10,11,11,11,12 is

(a) 11 (b) 12 (c) 7 (d) 7 & 11

283. In a symmetrical distribution when the 3^{rd} quartile plus 1^{st} quartile is halved, the value would give

(a) mean (b) mode (c) median (d) none

284. In Zoology, ———————— is used.

(a) median (b) mean (c) mode (d) none

285. For calculation of Speed & Velocity

(a) G.M (b) A.M (c) H.M (d) none is used.

286. The S.D is always taken from

(a) median (b) mode (c) mean (d) none

287. Coefficient of Standard deviation is equal to

(a) S.D/A.M (b) A.M/S.D (c) S.D/GM (d) none

288. The distribution , for which the coefficient of variation is less, is ——— consistent.

(a) less (b) more (c) moderate (d) none

ANSWERS

1	(b)	2	(a)	3	(c)	4	(a)	5	(b)
6	(a)	7	(d)	8	(c)	9	(b)	10	(a)
11	(a)	12	(a)	13	(b)	14	(d)	15	(a)
16	(d)	17	(b)	18	(a)	19	(b)	20	(a)
21	(b)	22	(d)	23	(a)	24	(c)	25	(b)
26	(a)	27	(a)	28	(b)	29	(b)	30	(a)
31	(c)	32	(a)	33	(b)	34	(a)	35	(a)
36	(a)	37	(a)	38	(a)	39	(a)	40	(a)
41	(c)	42	(a)	43	(a)	44	(b)	45	(a)
46	(b)	47	(a)	48	(a)	49	(b)	50	(a)
51	(a)	52	(d)	53	(a)	54	(a)	55	(b)
56	(c)	57	(a)	58	(c)	59	(c)	60	(b)
61	(b)	62	(d)	63	(a)	64	(b)	65	(b)
66	(c)	67	(c)	68	(a)	69	(a)	70	(c)
71	(c)	72	(c)	73	(a)	74	(a)	75	(d)
76	(a)	77	(b)	78	(b)	79	(a)	80	(c)
81	(b)	82	(a)	83	(b)	84	(b)	85	(a)
86	(c)	87	(c)	88	(b)	89	(b)	90	(a)
91	(b)	92	(d)	93	(a)	94	(d)	95	(c)
96	(c)	97	(a)	98	(a)	99	(c)	100	(b)
101	(a)	102	(c)	103	(c)	104	(a)	105	(b)
106	(a)	107	(d)	108	(c)	109	(a)	110	(b)
111	(c)	112	(b)	113	(b)	114	(a)	115	(b)
116	(c)	117	(c)	118	(b)	119	(a)	120	(b)
121	(b)	121	(a)	122	(a)	123	(b)	124	(b)
125	(c)	126	(d)	127	(d)	128	(a)	129	(c)
130	(b)	131	(a)	132	(a)	133	(a)	134	(b)
135	(c)	136	(a)	137	(d)	138	(a)		

139 (b)	140 (a)	141 (d)	142 (c)	143 (c)
144 (a)	145 (b)	146 (a)	147 (b)	148 (a)
149 (d)	150 (c)	151 (b)	152 (b)	153 (c)
154 (a)	155 (a)	156 (c)	157 (b)	158 (d)
159 (a)	160 (c)	161 (a)	162 (b)	163 (a)
164 (c)	165 (a)	166 (b)	167 (c)	168 (a)
169 (b)	170 (b)	171 (c)	172 (a)	173 (c)
174 (b)	175 (c)	176 (b)	177 (d)	178 (c)
179 (c)	180 (b)	181 (c)	182 (a)	183 (c)
184 (a)	185 (a)	186 (b)	187 (a)	188 (c)
189 (c)	190 (b)	191 (a)	192 (b)	193 (b)
194 (b)	195 (a)	196 (a)	197 (a)	198 (b)
199 (b)	200 (b)	201 (a)	202 (c)	203 (d)
204 (b)	205 (a)	206 (b)	207 (b)	208 (a)
209 (a)	210 (b)	211 (a)	212 (b)	213 (a)
214 (b)	215 (a)	216 (b)	217 (c)	218 (a)
219 (a)	220 (a)	221 (a)	222 (b)	223 (c)
224 (a)	225 (d)	226 (a)	227 (b)	228 (a)
229 (d)	230 (c)	231 (a)	232 (b)	233 (a)
234 (d)	235 (b)	236 (b)	237 (c)	238 (a)
239 (b)	240 (b)	241 (b)	242 (a)	243 (c)
244 (a)	245 (c)	246 (b)	247 (b)	248 (a)
249 (c)	250 (b)	251 (d)	252 (b)	253 (a)
254 (d)	255 (c)	256 (a)	257 (a)	258 (c)
259 (c)	260 (c)	261 (a)	262 (c)	263 (b)
264 (a)	265 (b)	266 (a)	267 (c)	268 (b)
269 (b)	270 (b)	271 (b)	272 (c)	273 (c)
274 (a)	275 (a)	276 (b)	277 (c)	278 (c)
279 (a)	280 (d)	281 (d)	282 (d)	283 (c)
284 (c)	285 (c)	286 (c)	287 (a)	288 (b)