# 7

# MEASURES OF CENTRAL TENDENCY

## 7.1 INTRODUCTION

We are familiar with the collection, classification and tabulation of statistical data. After the collection, classification and tabulation of statistical data, the next essential step in statistical enquiry is the *analysis* of data. The important objective of statistical analysis is to determine various numerical measures which describe the inherent features of a frequency distribution. The most important of such measures is *average*. The averages are the measures which condense the numerical facts into single numerical value which truly represents the whole set of data. Averages are those numerical constants or values which provide us the summary and give a bird's eye view of the huge mass of numerical data.

## 7.2 AVERAGE–MEANING AND DEFINITIONS

An average is also called the *measure of central tendency* because individual values of the variable usually cluster around it. According to **Simpson and Kafka**, "***A measure of central tendency is a typical value around which other figures congregate***". To make the data more comparable and to make it free from other abnormalities it becomes essential that the various phenomena which are being compared are reduced to a single most representative figure. **A.L. Bowley** states, "***Averages are statistical constants which enable us to comprehend in a single effort the significance of the whole***". According to **A.E. Waugh**, "***An average stands for the whole group of which it forms a part yet represents the whole.***" In the words of **C.T. Clark** and **L.L. Schkade**, "***Average is an attempt to find one single figure to describe whole of figures.***"

## 7.3 OBJECTIVES, IMPORTANCE AND FUNCTIONS OF AVERAGES

The importance of average is nicely given by **K.A. Yeomans** who states that "***primarily we need some means of describing the situations with which we are confronted. A concise numerical description is often preferable to a lengthy tabulation, and this form of description also enables us to form a mental image of the data and interpret its significance.***"

The following points highlight the importance and functions of averages :

**1. To get a single value as an indicator of whole of the data :** Although tabulation, diagrammatic and graphic presentation of data present the mass of data in a precise and condensed form yet we need a simple and single value which can truly reflect the whole of the characteristics of the problem under study. As such an average reduces a mass of data into a single typical figure which enables us to draw a general conclusion about its characteristics.

**2. To facilitate comparison :** Another important objective of measures of central tendency is to facilitate comparison. Figures become more comparable when they are represented by some most common and representative figure or an average. For example, comparison of per capita income of two countries can easily describe the standard of living of people in two countries.

**3. Act as basis for other statistical measures :** All other statistical measures such as measures of dispersion, skewness, correlation, regression analysis, time series, index numbers etc. depend upon the various measures of central tendencies. Thus, measures of central tendency form the basis for other statistical measures.

**4. To draw conclusions about the universe from a sample :** The most important way of studying the characteristics of the universe is the sampling technique. The average of a sample gives a true picture of the universe. As such average is useful in drawing conclusions about universe from the sample.

**5. To help in decision making :** In the modern business world most of the decisions are taken on the basis of average figures. The knowledge of an average value of a variable is significant in decision making.

## 7.4 CHARACTERISTICS OR ESSENTIALS OF A GOOD AVERAGE

An average is a single value which represents a group of values. There are various methods to determine an average. Almost all the statisticians have emphasised the following properties of a good measure of central tendency :

**1. It should be rigidly defined.** An average should be such that it has only one interpretation and there should be no confusion regarding its meaning and description *i.e.* it should be rigidly defined. It should be defined in terms of an algebraic formula.

**2. It should be easy to understand.** An average should not be based upon typical and complex mathematical formulation. The average should be simple and easy to understand.

**3. It should be easy to compute.** An average should be easy to understand but also simple to calculate. An average which involves complex calculations can not be regarded as a good average or measure of central tendency.

Pg 143

**4. It should be based on all the observations.** The average will be a true representative of all the items if it is based upon all the items in the series. Some of the measures of central tendency are not based on all the observations. That is why they are not treated as satisfactory.

**5. It should not be unduly affected by fluctuations in sampling.** A good average will be least affected by sampling fluctuations. If a few samples are taken from the same universe the average used should be such that there is least variation in averages derived from the individual samples. Only then the results obtained will be considered to be the true representative of the universe.

**6. It should be capable of further algebraic treatment.** The use of an average becomes restricted if it is not capable for further algebraic operations. For example, if we are given the averages of a few groups separately one should be capable of finding out the combined average.

**7. It should not be unduly affected by extreme values of the items.** An average value should not be unduly affected by the presence of extreme values. If in the data a few very small or large values unduly affect the average, in that case it will not be a true representative value of the entire group. For example, if we have three figures 10, 20, 30, then its mean is $\frac{10+20+30}{3} = 20$. In we add one more value *i.e.* 340 then average value become $\frac{10+20+30+340}{4} = \frac{400}{4} = 100$. It means the average value has increased from 20 to 100 just because of introduction of an extreme value *i.e.* 340.

**8. It should be located graphically.** An ideal average should be able to be located graphically.

**9. It should not be affected by skewness.** A good average is that which would not be affected presence of skewness. Skewness means slanting or tilt of values towards a particular value. An avera value which is affected by presence of skewness in data is not a true representative measure of cent tendency.

## 7.5 KINDS OF STATISTICAL AVERAGES

The following are the main types of averages used in statistical analysis :

(*i*) Arithmetic Mean or Simply 'Mean'

(*ii*) Geometric Mean

(*iii*) Harmonic Mean

(*iv*) Median

(*v*) Mode

Out of these measures, first three measures are called *mathematical averages* and last two measure are called *positional averages*.

In the following sections, we shall discuss each of these methods in detail.

## 7.6 ARITHMETIC MEAN

Arithmetic mean is the only average which qualifies almost all the essential requirements of a goo measure of central tendency. Whenever the term 'mean' is used it always refers to the arithmetic mean. It i also called '*common average*'. Each and every items has equal share in determining the arithmetic mea Arithmetic mean is the number which is obtained by adding all the values of a series and dividing the tota by the number of values. According to **H. Secrist,** ***"The arithmetic mean is the amount secured b dividing the sum of value of the items in a series by their numbers."***

Mathematically, if $X_1, X_2, ......, X_n$ are given $n$ observations then

$$\text{Arithmetic Mean } \overline{X} = \frac{X_1 + X_2 + ......+ X_n}{n} = \frac{\Sigma X}{n} \quad ...(7.1$$

Further, for a discrete frequency distribution as in table 7.1,

| X | $X_1$ | $X_2$ | ............. | $X_n$ |
|---|---|---|---|---|
| $f$ | $f_1$ | $f_2$ | ............. | $f_n$ |

Table 7.1

Arithmetic mean $\overline{X}$

$$= \frac{(X_1 + X_1 + ... f_1 \text{ times}) + (X_2 + X_2 + .... f_2 \text{ times}) + ....+ (X_n + X_n + .... f_n \text{ times})}{f_1 + f_2 + .....+ f_n}$$

$$= \frac{f_1X_1 + f_2X_2 + ......+ f_nX_n}{f_1 + f_2 + .....+ f_n} = \frac{\Sigma fX}{N} \text{ where } N = \Sigma f = f_1 + f_2 + .....+ f_n$$

*i.e.* Arithmetic mean $\overline{X} = \dfrac{\Sigma f X}{N}$ ...(7.2

where $N = \Sigma f$

For a continuous frequency distribution, we take mid values of the class-intervals as X and arithmetic mean is given by the formula (7.2).

**Properties of Arithmetic Mean**

The arithmetic mean possesses the following important mathematical properties :

**Property I.** The algebraic sum of deviations of the items from the arithmetic mean is zero.

*i.e.* $\Sigma(X - \overline{X}) = 0$ (for individual series)

or $\Sigma f(X - \overline{X}) = 0$ (for a frequency distribution)

**Proof :** Consider

$$\Sigma(X - \overline{X}) = \Sigma X - \Sigma \overline{X}$$

$$= n\overline{X} - \overline{X}\Sigma 1 \qquad \left(\because \overline{X} = \frac{\Sigma X}{n}\right)$$

$$= n\overline{X} - n\overline{X} = 0$$

Thus $\Sigma(X - \overline{X}) = 0$ for an individual series

Similarly $\Sigma f(X - \overline{X}) = 0$ for a frequency distribution.

**Property II.** The sum of squares of deviations of the given set of observations is least when deviations are taken from arithmetic mean.

**Proof :** Let $S = \Sigma (X - A)^2$ where A is any arbitrary value.

For S to be least, $\frac{dS}{dA} = 0$ and $\frac{d^2S}{dA^2} > 0$

Now $\frac{dS}{dA} = 0 \Rightarrow \Sigma 2(X - A)(-1) = 0 \Rightarrow \Sigma X - \Sigma A = 0$

$$\Rightarrow \qquad n\overline{X} - nA = 0 \qquad \left(\because \overline{X} = \frac{\Sigma X}{n}\right)$$

$$\Rightarrow \qquad A = \overline{X}$$

Further $$\frac{d^2S}{dA^2} = \Sigma - 2(0 - 1) = \Sigma 2 = 2n > 0$$

Thus S will be least when $A = \overline{X}$

Similarly for a frequency distribution, $\Sigma f(X - A)^2$ will be least if $A = \overline{X}$.

**Property III.** If the arithmetic mean of two or more series with the number of items in each series is given then we can calculate the arithmetic mean of the combined series. (See Section 7.6.4)

**Property IV.** Arithmetic mean is not independent of change of origin and scale. In other words, if the variable X is transformed to a new variable U by using the transformation $U = \frac{X - A}{h}$ then $\overline{X} = A + h\overline{U}$

*i.e.* $\overline{X}$ depends upon A and $h$.

**Proof :** Since $U = \frac{X - A}{h} \Rightarrow X = A + hU$

For an individual series having n observations,

$$\overline{X} = \frac{\Sigma X}{n} = \frac{\Sigma(A + hU)}{n} = \frac{\Sigma A + \Sigma hU}{n} = \frac{nA}{n} + h\frac{\Sigma U}{n}$$

*i.e.* $$\overline{X} = A + h\overline{U} \qquad ...(7.$$

Similarly for a frequency distribution, $\overline{X} = A + h\overline{U}$

*i.e.* $\overline{X}$ is not independent of A and $h$.

Notes (*a*) Addition or subtraction of a constant from a variable is termed as *change of origin* a multiplication or division of a variable by a non-zero constant is called *change of scale*.

(*b*) From equation (7.3), it is clear that

$$\overline{X} = A + h\frac{\Sigma U}{n} \quad \text{(for an individual series)} \qquad ...(7.4$$

and $$\overline{X} = A + \frac{h\Sigma fU}{N} \quad \text{(for a frequency distribution)} \qquad ...(7.5$$

where $$U = \frac{X - A}{h}$$

(*c*) If $U = X - A$ then by substituting $h = 1$ in formula (7.4) and (7.5), we have

$$\overline{X} = A + \frac{\Sigma U}{n} \quad \text{(for individual series)} \qquad ...(7.6$$

and $$\overline{X} = A + \frac{\Sigma fU}{N} \quad \text{(for frequency distribution)} \qquad ...(7.7$$

**Property V.** If each item of a series is replaced by arithmetic mean then the sum of the series will remai unchanged.

*i.e.* $$\Sigma X = \overline{X} + \overline{X} + \overline{X} + ....... \text{ n times}$$

or $$\Sigma X = n\overline{X}$$

**Property VI.** If each item of a series is added, subtracted, multiplied or divided by some non-zero constant then the mean also gets added, subtracted, multiplied or divided by the same constant.

**Merits and Demerits of Arithmetic Mean**

**Merits :** The following are the merits of arithmetic mean :

(*i*) The calculations of arithmetic mean are based on rigid mathematical properties and as such there is no scope for any misinterpretation.

(*ii*) It is easily understandable and easy to calculate. Because of this reason it is the most common measure of central tendency in day to day use in all fields.

(*iii*) Its results are based on all the items in the series. It truly represents all the items in the series.

(iv) Because of its rigid mathematical properties it acts as base for all the further statistical devices such as measures of dispersion, skewness, correlation, regression analysis, analysis of time series etc.

(v) Arithmetic mean is capable of further algebraic treatment. *e.g.* if the arithmetic means and the number of items contained in them are given for two or more series we can find out their combined mean.

(vi) It gives weightage to all the items directly proportional to the size of the items.

(vii) If different samples are taken from the same universe and various measures of central tendencies are calculated, the arithmetic mean will be the only measure having least variation. As such, we can say that arithmetic mean is the least affected sampling fluctuations. Because of this property we say that the arithmetic average is the most stable one.

(viii) Amongst all the measures only arithmetic mean is considered to be best for the purpose of comparison. Because of its rigidity and stability it is best for comparing the various issues from one place to another and from one time to the other.

**Demerits :** The following are the demerits of arithmetic mean :

(i) When the data is not expressible in the form of numerical measurements *i.e.,* in the quantitative form, we have no scope to calculate its arithmetic mean. Qualitative aspects such as intelligence, poverty etc. are beyond the limits of this measure of central tendency.

(ii) Arithmetic mean cannot be located by inspection after just having a glance on the data as is possible in the case of median or mode.

(iii) The most serious demerit of this measure of central tendency is that it is too much affected by the extreme items in the data. Extraordinary large or small items as compared to other items in the data will pull the value of the mean on one side and it will not remain the true representative of the items under consideration.

(iv) Arithmetic mean cannot be calculated if we have open ends frequency distribution. Mean can not be determined even if a single value is missing or not readable.

(v) Arithmetic average calculated on the basis of some given data may not belong to the series. We will name such an average as a *fictitious average.*

(vi) Arithmetic mean cannot be determined graphically whereas mode or median can be calculated graphically.

## 7.6.1 CALCULATION OF ARITHMETIC MEAN IN CASE OF INDIVIDUAL SERIES

The arithmetic mean in the case of individual series is obtained by adding the values of all the items and dividing the sum by the number of items. We have following three methods to calculate arithmetic mean :

**(i) Direct Method :** If $X_1, X_2, X_3 ...... X_n$ are the $n$ observations then the arithmetic mean by this method is obtained by using the formula (7.1)

*i.e.*
$$\overline{X} = \frac{X_1 + X_2 + X_3 + ...... + X_n}{n} = \frac{\Sigma X}{n}$$

where $\overline{X}$ = Arithmetic Mean, $\Sigma X$ = Sum of the observations and $n$ = Number of observations *(ii)* **Short Cut Method :** The short cut method is used when the size of items is very large. Short-cut method reduces the complexity and makes calculations easier. This method involves following steps :

**Step 1 :** For finding actual mean any figure in the given data is taken as the assumed mean. As far as possible the assumed figure should almost form the middle of data. The assumed mean is usually denoted by 'A'.

**Step 2 :** After the selection of 'A' deviations from the assumed mean *i.e.*, X − A are calculated. Deviations so obtained are usually denoted by '$dx$'.

**Step 3 :** These deviations are added and the sum is denoted by $\Sigma (X - A)$ or $\Sigma dx$.

**Step 4 :** After getting the sum of deviations *i.e.* $\Sigma dx$, use the formula of the form (7.6) to calculate mean value

*i.e.*
$$\text{Mean } \overline{X} = A + \frac{\Sigma dx}{n} \quad \ldots(7.\,)$$

where $\overline{X}$ = mean, A = Assumed Mean, $n$ = number of items on observation in the series and

$\Sigma\, dx$ = Sum of deviations

*(iii)* **Step Deviation Method :** Step deviation method is used when some common factor can be taken from values of $dx$. Common factor is the highest common value which divides all the values of $dx$.

*i.e.*
$$\text{Mean } \overline{X} = A + \frac{\Sigma d'x}{n} \times c \quad \ldots(7.\,)$$

where $c$ = Common factor from $dx$, A = Assumed Mean and $d'x = \dfrac{dx}{c}$

## CHECKPOINTS

1. Define Central Tendency. What are various methods of Central Tendency ?
   (G.N.D.U. B.Sc. C.Sc. April 2003, B.C.A. April 2007, Sept. 2007, P.U. B.C.A. April 2007)
2. What should be the characteristics of an ideal measure of Central Tendency ?
   (P.U. B.C.A. April 2005, Sept. 2006, 2007)
3. Write a short note on arithmetic mean.
   (P.U. B.C.A. April 2008)
4. Explain the purpose or significance of arithmetic mean.
   (G.N.D.U. B.Sc. C.Sc. April 2005, 2006, 2007)
5. Is Arithmetic mean independent of origin and scale ?
6. What are the merits and demerits of A.M. ?

# ILLUSTRATIVE EXAMPLES

**Example 1.** The daily income of 10 families in Rupees in a certain locality are given below :

| Family : | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| Income : | 85 | 70 | 10 | 75 | 44 | 8 | 42 | 250 | 40 | 36 |

Calculate the arithmetic average by *(a)* Direct Method *(b)* Short-cut Method.

Sol. (a) **Direct Method**

$$\text{Arithmetic Average} = \frac{\Sigma X}{n}$$

$$= \frac{85 + 70 + 10 + 75 + 44 + 8 + 42 + 250 + 40 + 36}{10} = \frac{660}{10} = 66$$

(b) **Short-cut Method**

Let assumed mean be 65. Now we prepare the following table :

| Family | Income (X) | A = 65 $dx = X - A$ |
|---|---|---|
| A | 85 | 20 |
| B | 70 | 05 |
| C | 10 | – 55 |
| D | 75 | 10 |
| E | 44 | – 21 |
| F | 08 | – 57 |
| G | 42 | – 23 |
| H | 250 | 185 |
| I | 40 | – 25 |
| J | 36 | – 29 |
| | | $\Sigma dx = 10$ |

c = 5

$\therefore$ Arithmetic Average $= A + \frac{\Sigma dx}{n} = 65 + \frac{10}{10} = 66$

Example 2. Calculate the arithmetic mean from following data using step deviation method :

30, 40, 50, 55, 60, 70, 80, 90, 100

Sol. Let assumed mean be 60. Now we prepare the following table :

| X | A = 60 $dx = X - A$ | $c = 5$ $d'x = dx/c$ |
|---|---|---|
| 30 | –30 | – 6 |
| 40 | – 20 | – 4 |
| 50 | –10 | –2 |
| 55 | –5 | – 1 |
| 60 | 0 | 0 |
| 70 | 10 | 2 |
| 80 | 20 | 4 |
| 90 | 30 | 6 |
| 100 | 40 | 8 |
| | | $\Sigma d'x = 7$ |

$\therefore$ Arithmetic Mean $\overline{X} = A + \frac{\Sigma d'x}{n} \times c = 60 + \frac{7}{9} \times 5 = 60 + \frac{35}{9} = 60 + 3.89 = 63.89$

## EXERCISE 7.1

1. Following figures give the marks of students in an examination. Calculate Arithmetic mean
   12, 8, 17, 13, 15, 9, 18, 11, 6, 1
2. The following table gives the monthly income of 12 families in the town . Calculate arithmetic mean.

| Sr. No. | Monthly income (Rs.) | Sr. No. | Monthly income (Rs.) |
|---|---|---|---|
| 1 | 280 | 7 | 80 |
| 2 | 180 | 8 | 94 |
| 3 | 96 | 9 | 100 |
| 4 | 98 | 10 | 75 |
| 5 | 104 | 11 | 600 |
| 6 | 75 | 12 | 200 |

3. The monthly income of 8 families in hundred of rupees in a certain locality are given below. Calculate arithmetic mean by short cut method

| Family : | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| Income : | 70 | 10 | 500 | 75 | 8 | 250 | 8 | 42 |

4. Calculate arithmetic mean of the items using short cut method :

| Roll No : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Marks : | 5 | 10 | 20 | 25 | 45 | 50 | 45 | 35 | 30 | 55 |

5. Calculate arithmetic mean from following data using short cut and step deviation method.

| Roll No : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Marks : | 25 | 35 | 50 | 40 | 55 | 60 | 75 | 65 | 90 | 85 |

## ANSWERS

1. 11  2. 165.17  3. Rs. 120.375 hundred  4. 32  5. 58

### 7.6.2 CALCULATION OF ARITHMETIC MEAN IN CASE OF DISCRETE SERIES

In discrete series the items are given with their respective frequencies and the arithmetic mean of the items is calculated by the following methods :

**(i) Direct Method :** The direct method involves the following steps :

**Step 1 :** First of all multiply the various frequencies with the corresponding values of the variable and then take their sum. It is denoted by $\Sigma fX$.

**Step 2 :** Now, we apply the formula (7.2) to calculate mean *i.e.* The sum so obtained in step I is divided by the sum of frequencies (denoted by $\Sigma f = N$) to obtain the arithmetic mean.

*i.e.* $$\text{Mean } \overline{X} = \frac{\Sigma fX}{N} \text{ or } \frac{\Sigma fX}{\Sigma f}$$

where $\overline{X}$ = Arithmetic mean, $f$ = Frequency and $N = \Sigma f$ = Sum of frequencies.

*(ii)* **Short Cut Method :** For calculation of arithmetic mean by short cut method the following steps are followed :

**Step 1 :** Like the individual series, we take assumed mean 'A' from the middle of values of the variable.

**Step 2 :** Take deviations of all the values from 'A' and denote it by $dx$.

**Step 3 :** Multiply each value of $dx$ with the corresponding value of $f$.

**Step 4 :** The sum of all the values obtained in step 3 will be denoted by $\Sigma f\,dx$.

**Step 5 :** Apply the following formula of the form (7.7) to calculate $\overline{X}$

*i.e.* $$\overline{X} = A + \frac{\Sigma f \cdot dx}{N} \qquad ...(7.10)$$

where A = Assumed Mean and $N = \Sigma f$.

*(iii)* **Step Deviation Method :** This method is applied when some common factor can be taken from values of $dx$. In this method, we apply the formula of the form (7.5) to calculate $\overline{X}$

*i.e.* $$\text{Mean } \overline{X} = A + \frac{\Sigma f d'x}{N} \times c \qquad ...(7.11)$$

where $c$ = Common factor from $dx$ and $d'x = \frac{dx}{c}$.

# ILLUSTRATIVE EXAMPLES

**Example 1.** The numbers of telephone calls received in 245 successive one minute intervals at an exchange are shown in the following frequency distribution. Calculate the mean.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Number of Calls : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Frequency : | 14 | 21 | 25 | 43 | 51 | 40 | 39 | 12 |

**Sol.** First, we prepare the following table :

| Number of Calls X | Frequency $f$ | $fX$ |
|---|---|---|
| 0 | 14 | 0 |
| 1 | 21 | 21 |
| 2 | 25 | 50 |
| 3 | 43 | 129 |
| 4 | 51 | 204 |
| 5 | 40 | 200 |
| 6 | 39 | 234 |
| 7 | 12 | 84 |
| | $\Sigma f = N = 245$ | $\Sigma fX = 922$ |

$$\therefore \quad \text{Mean } \overline{X} = \frac{\Sigma f X}{N} = \frac{922}{245} = 3.76.$$

**Example 2.** Find the Arithmetic mean of the following frequency distribution using short cut and step deviation method.

| Weight in kgs : | 50 | 55 | 60 | 65 | 70 | Total |
|---|---|---|---|---|---|---|
| Number of Men : | 15 | 20 | 25 | 30 | 10 | 100 |

**Sol.** First, we prepare the following table :

| Weight in kg. (X) | Number of Men ($f$) | A = 60 $dx = X - A$ | $c = 5$ $d'x = dx/c$ | $fd'x$ |
|---|---|---|---|---|
| 50 | 15 | −10 | −2 | −30 |
| 55 | 20 | −5 | −1 | −20 |
| 60 | 25 | 0 | 0 | 0 |
| 65 | 30 | 5 | 1 | 30 |
| 70 | 10 | 10 | 2 | 20 |
| | N = 100 | $\Sigma dx = 0$ | | $\Sigma fd'x = 0$ |

**Short Cut Method :**

$$\text{Arithmetic Mean } \overline{X} = A + \frac{\Sigma f\,dx}{N} = 60 + \frac{0}{100} = 60$$

**Step Deviation Method :**

$$\text{Arithmetic Mean } \overline{X} = A + \frac{\Sigma f d'x}{N} \times c = 60 + \frac{0}{100} \times c = 60$$

**Example 3.** The following are the monthly salaries in rupees of 20 employees in a firm

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 130 | 62 | 145 | 118 | 125 | 76 | 151 | 142 | 110 | 98 |
| 65 | 116 | 100 | 103 | 71 | 85 | 80 | 122 | 132 | 95 |

The firm gave bonus of Rs. 10, 15, 20, 25 and 30 for the individuals in the respective salary groups exceeding 60 but not exceeding Rs. 80, exceeding Rs. 80 but not exceeding Rs. 100 and so on upto exceeding Rs. 140 but not exceeding Rs. 160. Find average bonus paid per employee.

(P.U. B.C.A. Sept. 2005)

**Sol.** First, we prepare the following table :

| Salary | Tally Bar | Frequency $f$ | Bonus X | A = 20 $dx = X - A$ | C = 5 $d'x$ | $fd'x$ |
|---|---|---|---|---|---|---|
| Exceeding Rs. 60 but not Rs. 80 | ~~\|\|\|\|~~ | 5 | 10 | −10 | −2 | −10 |
| Exceeding Rs. 80 but not Rs. 100 | \|\|\|\| | 4 | 15 | −5 | −1 | −4 |
| Exceeding Rs. 100 but not Rs. 120 | \|\|\|\| | 4 | 20 | 0 | 0 | 0 |
| Exceeding Rs. 120 but not Rs. 140 | \|\|\|\| | 4 | 25 | 5 | 1 | 4 |
| Exceeding Rs. 140 but not Rs. 160 | \|\|\| | 3 | 30 | 10 | 2 | 6 |
| | | N = 20 | | | | $\Sigma fd'x = -4$ |

$\therefore$ Average Bonus paid $= \overline{X} = A + \frac{\Sigma fd'x}{N} \times c = 20 - \frac{4}{20} \times 5 = 20 - 1 =$ Rs. 19

**Example 31.** For a certain frequency table with total frequency 30, the arithmetic mean was found to be 18 but while copying out the table, a typist left out two of the class frequencies say $p$ and $q$ so that the table is given in the form :

| | | | | | | |
|---|---|---|---|---|---|---|
| X : | 5 | 10 | 15 | 20 | 25 | 30 |
| $f$ : | 2 | $p$ | 7 | $q$ | 5 | 2 |

Determine $p$ and $q$. (P.U. B.C.A. April 2005)

**Sol.** First we prepare the following table :

| X | $f$ | $f$X |
|---|---|---|
| 5 | 2 | 10 |
| 10 | $p$ | $10\,p$ |
| 15 | 7 | 105 |
| 20 | $q$ | $20\,q$ |
| 25 | 5 | 125 |
| 30 | 2 | 60 |
| | $N = 16 + p + q$ | $\Sigma f X = 300 + 10\,p + 20\,q$ |

Now, $N = 30 \Rightarrow \quad 16 + p + q = 30$

$\Rightarrow \quad p + q = 14 \quad \ldots(i)$

Also $\overline{X} = 18$

$\Rightarrow \quad \frac{\Sigma f X}{N} = 18 \Rightarrow \frac{300 + 10\,p + 20\,q}{30} = 18$

$\Rightarrow \quad 10\,p + 20\,q = 240$

or $\quad p + 2\,q = 24 \quad \ldots(ii)$

Solving equations (*i*) and (*ii*) for $p$ and $q$, we get

$p = 4, \; q = 10$

# EXERCISE 7.2

1. Calculate simple arithmetic mean from the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Marks : | 10 | 20 | 30 | 40 | 50 | 60 |
| No. of Students : | 10 | 20 | 25 | 20 | 15 | 10 |

2. The following data relates to the distance travelled by 520 villagers to buy their weekly requirements :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Kms. travelled : | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 |
| Number of Villagers : | 38 | 104 | 140 | 78 | 48 | 42 | 28 | 24 | 16 | 2 |

Calculate the arithmetic mean.

3. From the following data calculate the mean by short cut method :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Land Holdings (Acres) : | 12 | 14 | 16 | 18 | 20 | 22 | 24 | 26 |
| No. of Farmers : | 3 | 4 | 8 | 12 | 9 | 6 | 5 | 3 |

4. Calculate Arithmetic Mean for the following data using step deviation method :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Wages (Rs.) : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |
| Workers : | 4 | 5 | 3 | 2 | 5 | 2 | 3 | 1 | 2 |

5. The pass result of 40 students in an examination is given below :

| | | | | | | |
|---|---|---|---|---|---|---|
| Marks : | 4 | 5 | 6 | 7 | 8 | 9 |
| Number of Students : | 8 | 10 | 9 | 6 | 4 | 3 |

If the average marks for all the 50 students in class were 5.16 then calculate the average marks of the students who failed.

6. The following are the salaries of workers in Rupees :

130, 110, 80, 100, 76, 103, 62, 98, 71, 151, 122, 145, 65, 85, 142, 132, 118, 95, 116, 125

The firm gives bonus of Rs. 20, 30, 40, 50 and 70 to individuals in the respective salary groups exceeding Rs. 60 but not Rs. 80, exceeding Rs. 80 but not Rs. 100 and so on upto exceeding Rs. 140 but not Rs. 160. Find total and average bonus paid per employee.

7. From the following data calculate the missing value given mean = 13.913

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| X : | 5 | 6 | 8 | 10 | ? | 12 | 15 | 20 |
| $f$ : | 4 | 6 | 5 | 10 | 9 | 18 | 2 | 1 |

## ANSWERS

| | | | |
|---|---|---|---|
| 1. 34 | 2. 7.7769 | 3. 18.92 acres | 4. 42.593 |
| 5. 2.1 | 6. Rs. 790, Rs. 39.50 | 7. 33.69 | |

### 7.6.3 CALCULATION OF ARITHMETIC MEAN IN CASE OF CONTINUOUS SERIES

In continuous series the values of the variable are given in the form of class intervals with corresponding frequency. For calculation of mean in case of continuous series, following methods are used:

**(i) Direct method :** Let '$m$' represents the mid values of class intervals and '$f$' the corresponding frequencies. The arithmetic mean by this method is calculated by the formula of the form (7.2)

*i.e.*

$$\text{Mean } \overline{X} = \frac{\Sigma fm}{\Sigma f} = \frac{\Sigma fm}{N} \qquad ...(7.12)$$

**(ii) Short Cut Method :** After computing the mid values of various class intervals, deviations are taken from assumed mean 'A'. The deviations are represented by $dx$. Arithmetic mean is represented by the formula of the form (7.7)

*i.e.*

$$\text{Mean } \overline{X} = A + \frac{\Sigma fdx}{N} \qquad ...(7.13)$$

where $dx = m - A$.

*(iii)* **Step Deviation Method :** In a continuous series the calculations can be further reduced if the values of '$dx$' obtained in the short-cut method are divided by some common factor. Values so obtained are represented by '$d'x$'. The arithmetic mean is calculated by the formula of the form (7.5)

*i.e.*

$$\text{Mean } \overline{X} = A + \frac{\Sigma fd'x}{N} \times c \qquad ....(7.14)$$

where $c$ = Common Factor from $dx$ and $d'x = \frac{dx}{c}$.

**Key notes for calculating Arithmetic Mean**

1. Arrangement of data in ascending or descending order is not necessary to calculate arithmetic mean.
2. Conversion of inclusive series into exclusive series is not necessary to calculate A.M.
3. Unequal class intervals need not be converted into equal class intervals to find means.
4. In case of open end class intervals, class intervals are completed by the class intervals of adjoining class.
5. If data is in 'More than' or 'Less than' cumulative frequency form, convert it into simple frequency form before calculating A.M.
6. If in question it is not given which method is to be used, students are advised to use short cut method. If some common factor exists in $dx$, use step deviation method.

# ILLUSTRATIVE EXAMPLES

**Example 1.** Find the arithmetic mean for the following data :

| | | | | |
|---|---|---|---|---|
| Class : | 3 – 5 | 5 – 7 | 7–9 | 9– 11 |
| Frequency : | 10 | 30 | 20 | 40 |

(G.N.D.U. B.Sc. C.Sc. April 2004)

**Sol.** First we prepare the following table :

| Class (X) | Frequency ($f$) | Mid value ($m$) | $fm$ |
|---|---|---|---|
| 3 – 5 | 10 | 4 | 40 |
| 5 – 7 | 30. | 6 | 180 |
| 7 – 9 | 20 | 8 | 160 |
| 9 – 11 | 40 | 10 | 400 |
| | N = 100 | | $\Sigma fm = 780$ |

Now, arithmetic mean $\overline{X} = \frac{\Sigma fm}{N} = \frac{780}{100} = 7.8$

**Example 2.** From the data below, find Arithmetic Mean :

| | | | | |
|---|---|---|---|---|
| Class– interval | 50 – 54 | 55 – 59 | 60 – 64 | 65 – 69 |
| Frequency | 1 | 5 | 17 | 36 |
| Class– interval | 70–74 | 75–79 | 80–84 | 85–89 |
| Frequency | 25 | 11 | 4 | 1 |

**Sol.** First we prepare the following table :

| Class Interval | Frequency $f$ | Mid Value $(m)$ | A = 67 $dx = m - A$ | $c = 5$ $d'x$ | $fd'x$ |
|---|---|---|---|---|---|
| 50 – 54 | 1 | 52 | – 15 | – 3 | – 3 |
| 55 – 59 | 5 | 57 | – 10 | – 2 | – 10 |
| 60 – 64 | 17 | 62 | – 5 | – 1 | – 17 |
| 65 – 69 | 36 | 67 | 0 | 0 | 0 |
| 70 – 74 | 25 | 72 | 5 | 1 | 25 |
| 75 – 79 | 11 | 77 | 10 | 2 | 22 |
| 80 – 84 | 4 | 82 | 15 | 3 | 12 |
| 85 – 89 | 1 | 87 | 20 | 4 | 04 |
| | N = 100 | | | | $\Sigma fd'x = 33$ |

$\therefore$ arithmetic mean $\overline{X} = A + \frac{\Sigma fdx'}{N} \times c = 67 + \frac{33 \times 5}{100} = 67 + 1.65 = 68.65$

**Example 3.** Compute the arithmetic average if the changes in the prices of securities in a certain stock exchange are as follows :

| % Changes in Price | Frequency $(f)$ |
|---|---|
| 10 and over | 54 |
| 8 to 10 | 30 |
| 6 to 8 | 50 |
| 4 to 6 | 71 |
| 2 to 4 | 119 |
| 0 to 2 | 346 |
| 0 to – 2 | 149 |
| – 2 to – 4 | 100 |
| – 4 to – 6 | 74 |
| – 6 to – 8 | 31 |
| – 8 and under | 51 |
| | N = 1075 |

Sol. Since the given data is in the form of open ends therefore write the data in the proper form as under :

| Price of Securities (X) | Frequency ($f$) | Mid Value ($m$) | A = 1 $dx = m - A$ | $c = 2$ $d'x$ | $fd'x$ |
|---|---|---|---|---|---|
| 10 to 12 | 54 | 11 | 10 | 5 | 270 |
| 8 to 10 | 30 | 9 | 8 | 4 | 120 |
| 6 to 8 | 50 | 7 | 6 | 3 | 150 |
| 4 to 6 | 71 | 5 | 4 | 2 | 142 |
| 2 to 4 | 119 | 3 | 2 | 1 | 119 |
| 0 to 2 | 346 | 1 | 0 | 0 | 0 |
| 0 to – 2 | 149 | – 1 | – 2 | – 1 | – 149 |
| – 2 to – 4 | 100 | – 3 | – 4 | – 2 | – 200 |
| – 4 to – 6 | 74 | – 5 | – 6 | – 3 | – 222 |
| – 6 to – 8 | 31 | – 7 | – 8 | – 4 | – 124 |
| – 8 to – 10 | 51 | – 9 | – 10 | – 5 | – 255 |
| | N = 1075 | | | | $\Sigma fd'x = -149$ |

$\therefore$ Arithmetic average $= A + \frac{\Sigma fd'x}{N} \times c = 1 - \frac{149}{1075} \times 2 = 1 - \frac{298}{1075} = \frac{777}{1075} = 0.723$

Example 4. Calculate mean age from the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Age (Yrs.) less than : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| No. of Persons : | 18 | 34 | 49 | 61 | 71 | 76 | 78 | 79 |

Sol. First, we convert cumulative series into simple frequency series as follows :

| Age (Yrs.) (X) | No. of Persons ($f$) | Mid Value ($m$) | A = 35 $dx = m - A$ | $c = 10$ $d'x$ | $fd'x$ |
|---|---|---|---|---|---|
| 0–10 | 18 | 5 | −30 | −3 | −54 |
| 10−20 | 34−18 = 16 | 15 | −20 | −2 | −32 |
| 20−30 | 49−34 = 15 | 25 | −10 | −1 | −15 |
| 30−40 | 61−49 = 12 | 35 | 0 | 0 | 0 |
| 40−50 | 71−61 = 10 | 45 | 10 | 1 | 10 |
| 50−60 | 76−71 = 5 | 55 | 20 | 2 | 10 |
| 60−70 | 78−76 = 2 | 65 | 30 | 3 | 6 |
| 70−80 | 79−78 = 1 | 75 | 40 | 4 | 4 |
| | N = 79 | | | | $\Sigma fd'x = -71$ |

$\therefore$ Mean age $= A + \frac{\Sigma fd'x}{N} \times c = 35 - \frac{71}{79} \times 10 = 35 - \frac{710}{79} = 35 - 8.99 = 26.01 \approx 26$ years

**Example 5.** A market with 168 operating firms has the following distribution of average number of workers in various income groups.

| | | | | | | |
|---|---|---|---|---|---|---|
| Income | : | 150 – 300 | 300 – 500 | 500 – 800 | 800 – 1200 | 1200 – 1800 |
| Number of Firms | : | 40 | 32 | 26 | 28 | 42 |
| Avg. No. of workers per firm | : | 8 | 12 | 7.5 | 8.5 | 4 |

Find the average salary paid per worker in the whole market.

**Sol.** First, we prepare the following table :

| Income (X) | No. of Firms | Avg. No. of Workers | Total Workers ($f$) | Mid Value ($m$) | $fm$ |
|---|---|---|---|---|---|
| 150 – 300 | 40 | 8 | 320 | 225 | 72000 |
| 300 – 500 | 32 | 12 | 384 | 400 | 153600 |
| 500 – 800 | 26 | 7.5 | 195 | 650 | 126750 |
| 800 – 1200 | 28 | 8.5 | 238 | 1000 | 238000 |
| 1200 – 1800 | 42 | 4 | 168 | 1500 | 252000 |
| | | | N = 1305 | | $\Sigma fm = 842350$ |

$\therefore$ Average Salary paid per worker $= \dfrac{\Sigma fm}{N} = \dfrac{842350}{1305} = 645.48$

**Example 6.** From following data find out missing frequencies given Mean = 35 and $N = \Sigma f = 68$.

| | | | | | | |
|---|---|---|---|---|---|---|
| Marks : | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 |
| No. of Students : | 4 | 10 | 12 | ? | 20 | ? |

**Sol.** Let the missing frequencies be $f_1$ and $f_2$. Now we prepare the following table :

| X | $f$ | Mid Value $m$ | $fm$ |
|---|---|---|---|
| 0–10 | 4 | 5 | 20 |
| 10–20 | 10 | 15 | 150 |
| 20–30 | 12 | 25 | 300 |
| 30–40 | $f_1$ | 35 | $35 f_1$ |
| 40–50 | 20 | 45 | 900 |
| 50–60 | $f_2$ | 55 | $55 f_2$ |
| | $N = 46 + f_1 + f_2$ | | $\Sigma fm = 1370 + 35 f_1 + 55 f_2$ |

Now $\qquad \overline{X} = \dfrac{\Sigma fm}{N}$

$$35 = \frac{1370+35 f_1+55 f_2}{68}$$

$\therefore$

$$2380 = 1370 + 35 f_1 + 55 f_2$$

$$35 f_1 + 55 f_2 = 1010 \quad \ldots(i)$$

$$N = 68$$

Also

$$46 + f_1 + f_2 = 68$$

$\Rightarrow$

$$f_1 + f_2 = 22 \quad \ldots(ii)$$

$\Rightarrow$

Solving equations (*i*) and (*ii*) for $f_1$ and $f_2$, we get

$$f_1 = 10 \text{ and } f_2 = 12$$

# EXERCISE 7.3

1. Calculate A.M. from the following table :

| Marks : | 0 – 10 | 10 – 30 | 30 – 40 | 40 – 50 | 50 – 80 | 80 – 100 |
|---|---|---|---|---|---|---|
| No. of Students : | 5 | 7 | 15 | 8 | 3 | 2 |

(G.N.D.U. B.C.A. April 2009)

2. Calculate mean using following data using short cut method :

| Marks : | 5–15 | 15–25 | 25–35 | 35–45 | 45–55 | 55–65 | 65–75 |
|---|---|---|---|---|---|---|---|
| No. of Students : | 8 | 10 | 14 | 20 | 16 | 12 | 6 |

3. Calculate mean temperature using following data :

| Temp. C : | – 40 to – 30 | – 30 to – 20 | – 20 to – 10 | – 10 to – 0 | 0 to 10 | 10 to 20 | 20 to 30 |
|---|---|---|---|---|---|---|---|
| No. of Days : | 10 | 28 | 30 | 42 | 65 | 180 | 10 |

4. Population of males in different age-groups according to 1981 census of India is given below :

| Age Group (yrs.) | No. of Males (Population in lakhs) | Age Group (yrs.) | No. of Males (Population in lakhs) |
|---|---|---|---|
| 5 – 14 | 447 | 45 – 54 | 157 |
| 15 – 24 | 307 | 55 – 64 | 91 |
| 25 – 34 | 279 | 65 – 74 | 39 |
| 35 – 44 | 220 | | |

Compute the average age.

5. Find average salary from the following data :

| Salary (Rs.) : | 150–300 | 300– 500 | 500– 800 | 800 – 1200 | 1200 – 1800 |
|---|---|---|---|---|---|
| No. of Workers : | 40 | 32 | 26 | 28 | 42 |

6. Calculate arithmetic mean from the following series :

| X : | Below 10 | 10–20 | 20–30 | 30–40 | 40–50 | Above 50 |
|---|---|---|---|---|---|---|
| $f$ : | 5 | 8 | 6 | 7 | 4 | 3 |

7. Find the mean from the following data :

| Marks above : | 0 | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|---|
| No. of Students : | 36 | 26 | 21 | 14 | 10 | 0 |

8. Find the mean marks from the following data :

| Marks above : | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of Students : | 80 | 77 | 72 | 65 | 55 | 43 | 28 | 16 | 10 | 8 | 0 |

9. Calculate Arithmetic Mean from the following data :

| Mid. Values : | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
|---|---|---|---|---|---|---|---|---|
| Frequency : | 7 | 20 | 19 | 32 | 38 | 17 | 4 | 1 |

10. Following data relates to wages and total hours of work and average hours of work of workers.

| Wages (Rs.) : | 50-70 | 70-90 | 90-110 | 110-130 | 130-150 | 150-170 |
|---|---|---|---|---|---|---|
| Total hrs. of work : | 72 | 200 | 255 | 154 | 78 | 38 |
| Avg. hrs. of work : | 9 | 8 | 8.5 | 7 | 7.8 | 7.6 |

Calculate average wage paid per worker.

11. In a frequency distribution the value of the arithmetic mean is 18. Find the missing frequency.

| Class Intervals : | 11–13 | 13–15 | 15–17 | 17–19 | 19–21 | 21–23 | 23–25 |
|---|---|---|---|---|---|---|---|
| Frequency : | 3 | 6 | 9 | 13 | ? | 5 | 4 |

12. A certain number of salesmen were appointed in different territories and the following data were compiled from their sales reports.

| Sales ('000 Rs.) : | 4 – 8 | 8 – 12 | 12 – 16 | 16 – 20 | 20 – 24 | 24 – 28 | 28 – 32 | 32 – 36 | 36 – 40 |
|---|---|---|---|---|---|---|---|---|---|
| No. of salesmen : | 11 | 13 | 16 | 14 | ? | 9 | 17 | 6 | 4 |

If the average sale is believed to be Rs. 19,920, find the missing value.

13. Find the missing frequency of class interval 15 – 20 if the mean is 19 :

| $x$ : | 5 – 10 | 10 – 15 | 15 – 20 | 20 – 25 | 25 – 30 |
|---|---|---|---|---|---|
| $f$ : | 2 | 2 | ? | 4 | 4 |

## ANSWERS

1. 35.625
2. 40
3. 4.288
4. 27.95 years
5. Rs. 772.02
6. 21.818
7. 24.72
8. 51.75
9. 30.29
10. Rs. 103.20
11. 8
12. 10
13. 8

### 7.6.4 COMBINED ARITHMETIC MEAN

**Procedure**

Suppose that we have two groups of size $n_1$ and $n_2$ and mean $\overline{X}_1$ and $\overline{X}_2$.

Now, sum of $n_1$ observations of first group = $n_1\overline{X}_1$ $\left[\because \overline{X} = \frac{\Sigma X}{n} \Rightarrow \Sigma X = n\overline{X}\right]$

and sum of $n_2$ observations of first group = $n_2\overline{X}_2$

So, if we combine two groups then

sum of $(n_1 + n_2)$ observations of the combined group = $n_1\overline{X}_1 + n_2\overline{X}_2$

Hence mean of the combined group having $n_1 + n_2$ observations is given by

$$\overline{X}_{12} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2} \qquad ...(7.15)$$

In general if we have $k$ groups of size $n_1, n_2, ......, n_k$ and mean $\overline{X}_1, \overline{X}_2, ......, \overline{X}_k$ then mean of the combined group is given by

$$\overline{X}_{12...k} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2 + ......+ n_k\overline{X}_k}{n_1 + n_2 + ......+ n_k} \qquad ...(7.16)$$

## ILLUSTRATIVE EXAMPLES

**Example 1.** A set of 8 observations has mean 4 and another set of 10 observations has mean 6. Find the mean of the combined set.

(P.U. B.C.A. April 2007)

**Sol.** Given $n_1 = 8,\ \overline{X}_1 = 4,\ n_2 = 10,\ \overline{X}_2 = 6$

$$\therefore \qquad \overline{X}_{12} = \frac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2} = \frac{8 \times 4 + 10 \times 6}{8 + 10} = 5.11$$

**Example 2.** The average monthly wages of all the workers in a factory is Rs. 444. If the average wages paid to male and female workers are Rs. 480 and Rs. 360 respectively, find the percentage of male and female workers employed by the factory.

(G.N.D.U. B.C.A. Sept. 2008)

**Sol.** Let the total number of workers in the factory be 100, the number of male and female workers be $n_1$ and $n_2$ respectively, the average wages paid to the male and female workers be $\overline{X}_1$ and $\overline{X}_2$ respectively and the average monthly wages of all the workers be $\overline{X}_{12}$.

$$\therefore \qquad \overline{X}_1 = 480,\ \overline{X}_2 = 360 \text{ and } \overline{X}_{12} = 444.$$

Also $\qquad n_1 + n_2 = 100 \qquad \ldots(i)$

Now $\qquad \overline{X}_{12} = \dfrac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}$

$\Rightarrow \qquad 444 = \dfrac{480\, n_1 + 360\, n_2}{100}$

$\Rightarrow \qquad 480\, n_1 + 360\, n_2 = 44400$

or $\quad 12 \div \quad 4\, n_1 + 3\, n_2 = 370 \qquad \ldots(ii)$

Solving equations (*i*) and (*ii*) for $n_1$ and $n_2$, we get

$$n_1 = 70 \text{ and } n_2 = 30$$

Hence percentage of male workers = 70%

and $\quad$ percentage of female workers = 30%

**Example 3.** The mean height of 50 students in a college is 5′ – 8″. The height of 30 of these is given in the distribution below. Find the arithmetic mean of the height of the remaining 20 students.

| Height in inches : | 5′ – 4″ | 5′ – 6″ | 5′ – 8″ | 5′ – 10″ | 6′ – 0″ |
|---|---|---|---|---|---|
| Frequency : | 4 | 12 | 4 | 8 | 2 |

**Sol.** Mean height of 30 students given in the data

$$\overline{X}_1 = \frac{4\times 64 + 12\times 66 + 4\times 68 + 8\times 70 + 2\times 72}{30} = \frac{2024}{30} \text{ inches} \qquad (\text{Using } 1' = 12'')$$

Let $\overline{X}_2$ be the mean height of remaining 20 students.

Now we have $\qquad \overline{X}_1 = \dfrac{2024}{30}$ inches, $\overline{X}_{12} = 5' - 8'' = 68$ inches, $\overline{X}_2 = ?$, $n_1 = 30$, $n_2 = 20$

Now $\qquad \overline{X}_{12} = \dfrac{n_1\overline{X}_1 + n_2\overline{X}_2}{n_1 + n_2}$

$\therefore \qquad 68 = \dfrac{30 \times \dfrac{2024}{30} + 20\overline{X}_2}{30 + 20}$

$\Rightarrow \qquad 68 \times 50 = 2024 + 20\,\overline{X}_2$

$\Rightarrow \qquad 20\,\overline{X}_2 = 3400 - 2024 = 1376$

$\Rightarrow \qquad \overline{X}_2 = 68.8'' = 5' - 8.8''$

## EXERCISE 7.4

1. The mean height of 25 male workers in a factory is 61 cms and the mean height of 35 female workers in the same factory is 58 cms. Find the combined mean height of 60 workers in the factory ?

2. Four groups consists of 15, 20, 18 and 10 persons. Their respective Arithmetic means are 1.62, 1.48, 1.40 and 1.53. Determine combined mean of the four groups taken together.

3. The mean actual salaries paid to all employees was Rs. 5000. The mean actual salaries paid to male and female employees were Rs. 5200 and Rs. 4200 respectively. Determine the percentage of males and females employed by the company. (G.N.D.U. B.C.A. Sept. 2006)

4. Fifty students appeared in a competitive examination. The result of those who secured less than first class marks (60%) is tabulated below :

| Marks | No. of Students |
|---|---|
| 0–10 | 8 |
| 10–20 | 12 |
| 20–30 | 20 |

If the average for all the students was 25 marks, find out the average marks for those who secured 60% or more.

5. The mean wages of a group of 100 workers is Rs. 350. The mean wages of 20 lowest paid workers is Rs. 50 and 10 highest paid is Rs. 500. Find the average wage of the other 70 workers.

6. The first group of students consists of 10 third divisioners with mean marks 36, 15 second divisioners with mean marks 56 and 25 first divisioners with 64 as their mean marks. Second group of students consists of 13 third divisioners, 17 second divisioners and 20 first divisioners with mean marks 45, 55 and 74 respectively. Calculate combined mean of two groups separately and then the combined mean of the groups taken together.

## ANSWERS

| | | | |
|---|---|---|---|
| 1. 59.25 cm | 2. 1.498 | 3. 80%, 20% | 4. 53 |
| 5. Rs. 414 | 6. 58 | | |

### 7.6.5 CORRECTING INCORRECT ARITHMETIC MEAN

Suppose that the arithmetic mean $\overline{X}$ of $n$ observations is given . Later on, if it is found that one or more observations in calculating mean were copied wrongly then in order to calculate the correct value of means we follow the following steps :

1. Find incorrect $\Sigma X$ by using the formula $\Sigma X = n\overline{X}$

2. Calculate correct $\Sigma X$ by using the formula

   correct $\Sigma X$ = Incorrect $\Sigma X$ – Incorrect values + Correct Values

3. Find correct Mean by using the formula

$$\text{correct } \overline{X} = \frac{\text{Correct } \Sigma X}{n}.$$

# ILLUSTRATIVE EXAMPLES

**Example 1.** Mean of 100 observations is found to be 40. If at the time of computation two items were wrongly taken as 30 and 27 instead of 3 and 72, find the correct mean. (G.N.D.U. B.C.A. Sept. 2006)

**Sol.** Given $n = 100$ and $\overline{X} = 40$.

Since $\overline{X} = \dfrac{\Sigma X}{n}$ so $\Sigma X = n\overline{X}$

$\therefore$ incorrect $\Sigma X = 100 \times 40 = 4000$

So correct $\Sigma X = 4000 - 30 - 27 + 3 + 72 = 4018$

Hence correct mean $= \dfrac{4018}{100} = 40.18$

**Example 2.** The mean marks of 100 students were found to be 40. Later on it was discovered that a score of 53 was misread as 83. Find the correct mean corresponding to correct score in the following cases :

(*i*) When it is an individual series.

(*ii*) When the data is grouped in the class intervals 0– 5, 5 – 10, 10 – 15,........etc.

(*iii*) When the data is grouped in the class intervals 0 – 10, 10– 20, 20– 30,.....etc.

**Sol.** Since $\overline{X} = \dfrac{\Sigma X}{N}$

$\therefore$ $40 = \dfrac{\Sigma X}{10}$

$\therefore$ incorrect $\Sigma X = 4000$

(*i*) When the data happens to be an individual series :

Correct value of $\Sigma X = 4000 - 83 + 53 = 3970$

$\therefore$ Correct $\overline{X} = \dfrac{\Sigma X}{N} = \dfrac{3970}{100} = 39.7$

(*ii*) When the data is grouped in the class intervals 0– 5, 5 –10, 10–15 .........etc. we shall subtract 82.5 in place of 83 and add 52.5 in place of 53 because 83 and 53 belong to the class intervals 80 – 85 and 50 – 55 respectively whose mid-values are 82.5 and 52.5.

Thus correct $\Sigma X = 4000 - 82.5 + 52.5 = 3970$

$\therefore$ Correct $\overline{X} = \dfrac{\Sigma X}{N} = \dfrac{3970}{100} = 39.7$

(*iii*) When the data is grouped in the class intervals 0–10 , 10–20, 20–30.........etc. we shall subtract 85 in place 83 and add 55 in place of 53 because 83 and 53 belong to the class intervals 80–90 and 50–60 respectively whose mid values are 85 and 55.

Thus correct value of $\Sigma X = 4000 - 85 + 55 = 3970$

$\therefore$ Correct $\overline{X} = \dfrac{\Sigma X}{N} = \dfrac{3970}{100} = 39.7$

# EXERCISE 7.5

1. The mean marks of 100 students were found to be 40. Later on it was discovered that a score of 53 was misread as 83. Find the correct mean corresponding to correct score.
2. The average weight of a group of 25 boys was calculated to be 78.4 lbs. It was later found that weight of one boy was misread as 69 lbs instead of the correct value 96 lbs. Calculate the correct average.
3. Mean wage of 100 workers per day was found to be Rs. 90. But later on, it was found that wages of two workers Rs. 93 and Rs 59 were misread as Rs. 39 and Rs. 95. Find out the correct mean wage.
4. Average height of 90 students was found to be 165 cms. Later on it was discovered that height of 5 students was measured wrongly as 158, 156, 160, 150 and 161. It was decided to ignore these heights. Find out the correct average height.

## ANSWERS

1. 39.7 2. 79.48 lbs 3. Rs. 90.18 4. 165.47 cm

### 7.6.6 WEIGHTED ARITHMETIC MEAN

While calculating the simple arithmetic mean we give equal importance to all the items. When *weighted arithmetic mean* is to be calculated we assign weights to the items according to their importance in the particular field to which they belongs. Thus, the weighted arithmetic mean is calculated by the formula

$$\overline{X}_w = \frac{\Sigma WX}{\Sigma W} \qquad ...(7.17)$$

where $\overline{X}_w$ = weighted arithmetic mean and W represent the weights assigned to the items

**Relation between Weighted and Simple Arithmetic Mean**

Following characteristics of weighted arithmetic mean as compared to simple arithmetic must be born in mind :

(*i*) Weighted average is equal to simple average *i.e.* $\overline{X}_w = \overline{X}$ if and only if equal weights are assigned to all the items in series or distribution i.e. when equal importance is given to all the items.

(*ii*) Weighted average is greater than simple average *i.e.* $\overline{X}_w > \overline{X}$ when larger weights are assigned to larger items and smaller weights are assigned to smaller items.

(*iii*) Weighted average is less than the simple average *i.e.* $\overline{X}_w < \overline{X}$ when smaller weights are assigned to larger items and greater weights are assigned to smaller items.

**CHECKPOINTS**

1. What is the relationship between simple arithmetic mean and weighted arithmetic mean ?

# ILLUSTRATIVE EXAMPLES

**Example 1.** Comment on the performance of the students of three universities given below using (*i*) simple arithmetic mean and (*ii*) weighted arithmetic mean.

| Courses of Study | Bombay University | | Calcutta University | | Madras University | |
|---|---|---|---|---|---|---|
| | Pass % | No. of students (in hundreds) | Pass % | No. of students (in hundreds) | Pass % | No. of students (in hundreds) |
| M.A. | 71 | 3 | 82 | 2 | 81 | 2 |
| M.Com. | 83 | 4 | 76 | 3 | 76 | 3.5 |
| B.A. | 73 | 5 | 73 | 6 | 74 | 4.5 |
| B.Com. | 74 | 2 | 76 | 7 | 58 | 2 |
| B.Sc. | 65 | 3 | 65 | 3 | 70 | 7 |
| M.Sc. | 66 | 3 | 60 | 7 | 73 | 2 |

**Sol.** (*i*) Calculation of simple mean in case of each university.

(*a*) Arithmetic mean for Bombay University $= \frac{\Sigma X}{N} = \frac{71+83+73+74+65+66}{6} = 72$

(*b*) Arithmetic mean for Calcutta University $= \frac{\Sigma X}{N} = \frac{82+76+73+76+65+60}{6} = 72$

(*c*) Arithmetic mean for Madras University $= \frac{\Sigma X}{N} = \frac{81+76+74+58+70+73}{6} = 72$

Since the simple arithmetic mean for all the universities is the same therefore we cannot come to any conclusion regarding the performance of the universities as such.

(*ii*) Calculation of weighted arithmetic mean :

| Courses of Study | Bombay Uni. | | | Calcutta Uni. | | | Madras Uni. | | |
|---|---|---|---|---|---|---|---|---|---|
| | (X) Pass % | W | WX | (X) Pass % | W | WX | (X) Pass % | W | WX |
| M.A. | 71 | 3 | 213 | 82 | 2 | 164 | 81 | 2 | 162 |
| M.Com. | 83 | 4 | 332 | 76 | 3 | 228 | 76 | 3.5 | 266 |
| B.A. | 73 | 5 | 365 | 73 | 6 | 438 | 74 | 4.5 | 333 |
| B.Com. | 74 | 2 | 148 | 76 | 7 | 532 | 58 | 2 | 116 |
| B.Sc. | 65 | 3 | 195 | 65 | 3 | 195 | 70 | 7 | 490 |
| M.Sc. | 66 | 3 | 198 | 60 | 7 | 420 | 73 | 2 | 146 |
| Total | | $\Sigma W = 20$ | $\Sigma WX = 1451$ | | $\Sigma W = 28$ | $\Sigma WX = 1977$ | | $\Sigma WX = 21$ | $\Sigma WX = 1513$ |

$\therefore$ weighted mean for Bombay University $= \dfrac{\Sigma WX}{\Sigma W} = \dfrac{1451}{20} = 72.55$

$\therefore$ weighted mean for Calcutta University $= \dfrac{\Sigma WX}{\Sigma W} = \dfrac{1977}{28} = 70.60$

$\therefore$ weighted mean for Madras University $= \dfrac{\Sigma WX}{\Sigma W} = \dfrac{1513}{21} = 72.00$

Thus, the performance of Bombay University is the best since it has the highest weighted mean.

## EXERCISE 7.6

1. A candidate obtained the following percentage of marks in various subjects as follow. Find the weighted mean if the weights are given against each ?

| Subject : | English | Hindi | Maths | Physics | Chemistry |
|---|---|---|---|---|---|
| % Marks : | 60 | 75 | 63 | 60 | 55 |
| Weights : | 1 | 1 | 2 | 3 | 3 |

2. A candidate obtains following percentages in an examination : English 60, Hindi 75, Mathematics 63, Physics 59, and Chemistry 55. Find the weighted mean if weights 1, 2, 1, 3 and 3 respectively are allotted to the subjects. Also calculate simple mean.

3. Compute weighted arithmetic mean of the Index Number from the data given below :

| Group : | Food | Clothing | Fuel and Lighting | House Rent | Misc. |
|---|---|---|---|---|---|
| Index No. : | 125 | 133 | 141 | 173 | 182 |
| Weight : | 7 | 5 | 4 | 1 | 3 |

4. Illustrate with the help of some suitable examples when weighted average is

(*i*) equal to (*ii*) greater than (*iii*) less than the simple average.

## ANSWERS

1. 60.6 2. 61.5, 62.4 3. 141.5

### 7.7 GEOMETRIC MEAN

Geometric mean is a mathematical measure of central tendency. The $n$th root of the product of $n$ items is known as *Geometric Mean. i.e.* if $X_1, X_2, X_3, \ldots X_n$ are the individual observations then,

$$\text{Geometric Mean G.M.} = (X_1 . X_2 . X_3 \ldots\ldots X_n)^{\frac{1}{n}} \qquad \ldots(7.18)$$

Since the calculation of $n$th root of product of numbers is difficult, therefore, we transform it in the logarithmic form by taking log on both sides for formula (7.18).

$$\therefore \qquad \log \text{G.M.} = \frac{1}{n}\{\log X_1 + \log X_2 + \log X_3 + \ldots\ldots + \log X_n\} = \frac{\Sigma \log X}{n}$$

$$\therefore \qquad \text{G.M.} = \text{A.L.}\left(\frac{\Sigma \log X}{n}\right) \qquad \ldots(7.19)$$

where A.L. stands for antilog.

Further, for a discrete frequency distribution as in table 7.2,

| | | | | |
|---|---|---|---|---|
| X | $X_1$ | $X_2$ | ............... | $X_n$ |
| $f$ | $f_1$ | $f_2$ | ............... | $f_n$ |

Table 7.2

Geometric Mean G.M.

$$= [(X_1 \times X_1 \times \ldots f_1 \text{ times}) \times (X_2 \times X_2 \times \ldots. f_2 \text{ times}) \times \ldots. \times (X_n \times X_n \times \ldots. f_n \text{ times})]^{1/N}$$

$$= \left[X_1^{f_1} X_2^{f_2} \ldots\ldots\ldots X_n^{f_n}\right]^{1/N} \quad \text{where } N = f_1 + f_2 + \ldots\ldots + f_n = \Sigma f$$

$$\therefore \qquad \log \text{G.M.} = \frac{1}{N}\log\left(X_1^{f_1} X_2^{f_2} \ldots\ldots\ldots X_n^{f_n}\right)$$

$$= \frac{1}{N}[f_1 \log X_1 + f_2 \log X_2 + \ldots\ldots\ldots + f_n \log X_n]$$

$$= \frac{1}{N}\Sigma f \log X$$

So

$$\text{G.M.} = \text{A.L.}\left(\frac{1}{N}\Sigma f \log X\right) \qquad \ldots(7.20)$$

where A.L. stands for antilog.

For a continuous frequency distribution we take mid values of the class-intervals as X and geometric mean is given by the formula (7.20).

## Properties of Geometric Mean

The geometric mean possesses the following properties :

**Property I :** By definition of geometric mean it is $n$th root of the product of $n$ items

*i.e.*

$$\text{G.M.} = (X_1 X_2 X_3 \ldots\ldots X_n)^{\frac{1}{n}}$$

This leads us to the fact that

$$(\text{G.M.})^n = X_1\, X_2\, X_3 \ldots\ldots X_n$$

*i.e.* the $n$th power of the G.M. is always equal to the product of all the $n$ items.

**Property II :** If the geometric means of a few groups of items is given along with the number of items in each group we can find the combined geometric mean. (See Section 7.7.5)

**Property III :** If two or more series have the equal number of items then G.M. of the product of the corresponding items in the series is the same as the product of the G.M.s of the individual series.

**Property IV :** The product of the ratios of the values less than or equal to the G.M. to the G.M. is equal to the product of the ratios of the values greater than or equal to the G.M. to the G.M.

**Property V :** The sum of the deviations of the log of the values from the log of the G.M. is always zero.

**Property VI :** The geometric mean of the ratios of the corresponding items of two series with equal number of items is equal to the ratio of geometric means.

**Merits and Demerits of Geometric Mean**

**Merits :** The following are merits of geometric mean :

(*i*) Geometric mean is the only average which is most suitable for measuring the average of ratios and percentages.

(*ii*) In social and economic problems where large weights are assigned to small items and small weights to large items, it is the most suitable average.

(*iii*) This is less affected by extreme values in the data.

(*iv*) Because of its mathematical characteristics it is capable of further algebraic treatment.

(*v*) It is based upon all the items in the series.

**Demerits :** The following are demerits of geometric mean :

(*i*) Geometric mean cannot be determined if any of its item has zero value. Even negative value of any item will leave the value of G.M. undefined.

(*ii*) Although G.M. is based upon mathematical formulae but even then its computation is not easy as compared to other measures of central tendency.

(*iii*) Because of its tedious calculation it is not a widely known and commonly used average.

## 7.7.1 CALCULATION OF GEOMETRIC MEAN FOR INDIVIDUAL SERIES

In case of individual series the geometric mean is calculated by using formula (7.19)

*i.e.*
$$\text{G.M.} = \text{A.L.}\left(\frac{\Sigma \log X}{n}\right)$$

where A.L. stands for antilog.

### 7.7.2. CALCULATION OF GEOMETRIC MEAN IN DISCRETE SERIES

In case of discrete series the geometric mean is calculated by using the formula (7.20)

*i.e.*
$$\text{G.M.} = \text{A.L.}\left(\frac{\Sigma f \log X}{N}\right)$$

where A.L. stands for antilog.

### 7.7.3 CALCULATION OF GEOMETRIC MEAN IN CONTINUOUS SERIES

In case of continuous series, we take the mid value of each class interval and apply the formula of the form (7.20)

*i.e.*
$$\text{G.M.} = \text{A.L.}\left(\frac{\Sigma f \log m}{N}\right) \quad ...(7.21)$$

where $m$ stands for mid value of the class intervals and A.L. stands for antilog.

### 7.7.4 CALCULATION OF WEIGHTED GEOMETRIC MEAN

The weighted geometric mean is calculated by using the formula

$$\text{G.M.}_W = \text{A.L.}\left[\frac{\Sigma W \log X}{\Sigma W}\right] \quad ...(7.22)$$

where $\text{G.M.}_W$ = weighted geometric mean, W represents weight assigned to the various items and A.L. stands for antilog.

### 7.7.5 COMBINED GEOMETRIC MEAN

If $G_1$ and $G_2$ represent the geometric means of two groups of size $n_1$ and $n_2$ respectively then the geometric mean of the group formed by combining these two groups is given by

$$G_{12} = \text{A.L.}\left[\frac{n_1 \log G_1 + n_2 \log G_2}{n_1 + n_2}\right] \quad ...(7.23)$$

where A.L. stands for antilog.

The formula (7.23) for calculating combined geometric mean can be extended for the group formed by combining more than two groups.

### 7.7.6 COMPUTATION OF RATE OF GROWTH

The geometric mean can be used for calculating rate of growth of a phenomenon which grow at geometric progression rate *i.e.* compound rate.

If $P_0$ is the initial value of the variable, $P_n$ is the final value of the variable at the end of the period '$n$' and $r$ is the rate of growth per unit time then

$$P_n = P_0(1+r)^n \quad ...(7.24)$$

In equation (7.24), we have four quantities $P_0, P_n, r$ and $n$. If we are given the values of any three quantities then the value of fourth quantity can be calculated using the formula (7.24).

**Note** If the variable increases at different rates in different intervals of time *i.e.* if $P_0$ is the initial value of the variable, $P_n$ is the final value of the variable at the end of period '$n$' and rate of growth for 1st, 2nd, 3rd, ....., $n$th period is $r_1, r_2, r_3, \ldots\ldots, r_n$ then

$$P_n = P_0(1+r_1)(1+r_2)\ldots\ldots(1+r_n)$$

If $r$ is assumed to be constant rate of growth per unit time then

$$P_n = P_0(1+r)^n \qquad \text{(by using formula (7.24))}$$

$$\therefore \quad P_0(1+r)^n = P_0(1+r_1)(1+r_2)\ldots\ldots(1+r_n)$$

$$\Rightarrow \quad 1+r = [(1+r_1)(1+r_2)\ldots\ldots(1+r_n)]^{1/n} \qquad \ldots(7.25)$$

Further if $r_1, r_2, r_3, \ldots\ldots, r_n$ denote the percentage growth for 1st, 2nd, 3rd, ....., $n$th period of time respectively then from formula (7.25), we have

$$1+\frac{r}{100} = \left[\left(1+\frac{r_1}{100}\right)\left(1+\frac{r_2}{100}\right)\ldots\ldots\left(1+\frac{r_n}{100}\right)\right]^{1/n} \qquad \ldots(7.26)$$

where $r$ is average percentage growth rate for $n$ periods.

Formula (7.26) can also be written in the form

$$r = [(100+r_1)(100+r_2)\ldots\ldots(100+r_n)]^{1/n} - 100$$

In other words if $r_1, r_2, r_3, \ldots\ldots, r_n$ are percentage rate of growth for 1st, 2nd, ................., $n$th period of time then average percentage growth can be obtained by subtracting 100 from the G.M. of $(100 + r_1)$, $(100 + r_2)$, ......, $(100 + r_n)$.

## CHECKPOINTS

1. What do you understand by Geometric Mean ? (P.U. B.C.A. April 2008)
2. What are the merits and demerits of Geometric Mean ?
3. Explain the purpose or significance of Geometric Mean.

(G.N.D.U. B.Sc. C.Sc. April 2005, 2006, 2007)

# ILLUSTRATIVE EXAMPLES

**Example 1. Calculate G.M. for the following individual series :**

95, 70, 15, 75, 500, 8, 45, 250, 40, 36

(G.N.D.U. B.Sc. I.T. 2007)

**Sol.** First, we prepare the following table :

| Income (X) | log X |
|---|---|
| 95 | 1.9777 |
| 70 | 1.8451 |
| 15 | 1.1761 |
| 75 | 1.8751 |
| 500 | 2.6990 |
| 8 | 0.9031 |
| 45 | 1.6532 |
| 250 | 2.3979 |
| 40 | 1.6021 |
| 36 | 1.5563 |
| | $\Sigma \log X = 17.6856$ |

$$\therefore \quad \text{G.M.} = \text{A.L.} \frac{\Sigma \log X}{n} = \text{A.L.} \frac{(17.6856)}{10}$$

$$= \text{A.L.}\ (1.7856) = 61.04$$

**Example 2.** Find the average rate of increase in population which in first decade has increased by 20%, in second decade by 30% and in third decade by 40%. (G.N.D.U. B.C.A. Sept. 2006)

**Sol.** First, we prepare the following table :

| Decade | % Increase | Increased Population (X) | log X |
|---|---|---|---|
| I | 20 | 120 | 2.0792 |
| II | 30 | 130 | 2.1139 |
| III | 40 | 140 | 2.1461 |
| | | | $\Sigma \log X = 6.3392$ |

Now $$\text{G.M.} = \text{A.L.} \frac{\Sigma \log X}{n} = \text{A.L.} \left(\frac{6.3392}{3}\right) = \text{A.L.}\ (2.1131) = 129.75$$

$\therefore$ average rate of increase = 29.75%

**Example 3.** Calculate geometric mean from the following data :

| Yield | Number of Farms | Yield | Number of farms |
|---|---|---|---|
| 7.5 | 5 | 23.0 | 7 |
| 13.0 | 8 | 24.0 | 3 |
| 18.5 | 10 | 25.0 | 4 |
| 20.5 | 14 | 26.0 | 2 |
| 22.0 | 6 | 28.0 | 1 |

Sol. First, we prepare the following table :

| Yield (X) | Number of Farms ($f$) | log X | $f$ log X |
|---|---|---|---|
| 7.5 | 5 | 0.8751 | 4.3755 |
| 13.0 | 8 | 1.1139 | 8.9112 |
| 18.5 | 10 | 1.2672 | 12.6720 |
| 20.5 | 14 | 1.3118 | 18.3652 |
| 22.0 | 6 | 1.3424 | 8.0544 |
| 23.0 | 7 | 1.3617 | 9.5319 |
| 24.0 | 3 | 1.3802 | 4.1406 |
| 25.0 | 4 | 1.3979 | 5.5916 |
| 26.0 | 2 | 1·4150 | 2.8300 |
| 28.0 | 1 | 1.4472 | 1.4472 |
| | 60 | | $\Sigma f \log X = 75.9196$ |

$$\therefore \quad \text{G.M.} = \text{A.L.}\ \frac{\Sigma f \log X}{N} = \text{A.L.}\ \frac{75.9196}{60} = \text{A.L.}\ (1.2653) = 18.42$$

Example 4. The following table gives marks obtained by 70 students in Mathematics. Calculate simple mean and geometric mean of the series.

| | | | | | | |
|---|---|---|---|---|---|---|
| Marks (more than) : | 70 | 60 | 50 | 40 | 30 | 20 |
| No. of Students : | 7 | 18 | 40 | 40 | 63 | 70 |

Sol. First, we convert the cumulative frequency distribution into simple frequency distribution and prepare the following table :

| Marks | Mid Value ($m$) | No. of Student ($f$) | $fm$ | log $m$ | $f$ log $m$ |
|---|---|---|---|---|---|
| 20 – 30 | 25 | 7 = (70 – 63) | 175 | 1.3979 | 9.7853 |
| 30 – 40 | 35 | 23 = (63 – 40) | 805 | 1.5441 | 35.5143 |
| 40 – 50 | 45 | 0 = (40 – 40) | 0 | 1.6532 | 0.0000 |
| 50 – 60 | 55 | 22 = (40 – 18) | 1210 | 1.7404 | 38.2888 |
| 60 – 70 | 65 | 11 = (18 – 7) | 715 | 1.8129 | 19.9419 |
| 70 – 80 | 75 | 7 | 525 | 1.8751 | 13.1257 |
| | | N = 70 | $\Sigma fm = 3430$ | | $\Sigma f \log m = 116.656$ |

$$\therefore \quad \text{G.M.} = \text{A.L.}\ \frac{\Sigma f \log m}{N} = \text{A.L.}\ \frac{116.656}{70} = \text{A.L.}\ (1.6665) = 46.39$$

$$\text{and} \quad \text{Arithmetic Mean} = \frac{\Sigma fm}{N} = \frac{3430}{70} = 49$$

**Example 5.** The population of a town has doubled itself in twenty years. Is it correct to say the rate of growth has been 5% per annum ? If yes why ? If not, give the correct average % increase.

**Sol.** Let the initial population = 100 = $P_0$

Population after 20 years = 200 = $P_n$

Let $r$ be the rate of growth per unit

Thus $$P_n = P_0 (1 + r)^n$$

$$\Rightarrow \quad 200 = 100 (1 + r)^{20}$$

$$\Rightarrow \quad 2 = (1+r)^{20}$$

$$\Rightarrow \quad \log 2 = 20 \log (1 + r)$$

$$\Rightarrow \quad \log (1 + r) = \frac{0.3010}{20} = 0.0151$$

$$\Rightarrow \quad 1 + r = \text{A.L.} (0.0151) = 1.035$$

$$\Rightarrow \quad r = 0.035 \text{ per annum} = 3.5\%$$

5% rate of growth have been obtained by dividing 100 by 20 (number of years), which cannot be so because average rate of growth, when given in percentages is obtained by G.M. only.

**Example 6.** The value of the machine when new is Rs. 20000. It depreciates in its value at the rate of 3% per annum in the first 4 years and then at the rate of 5% per annum in the next six years. What will be its value after 10 years ?

**Sol.** Let $P_n$ be the price of the machine after 10 years and $r_1$ and $r_2$ be the rates of depreciation per annum in first 4 years and next 6 years respectively.

Also $P_0 = 20000$.

Since $$P_n = P_0(1-r_1)^{n_1}(1-r_2)^{n_2}$$

$$\therefore \quad P_n = 20000 (1 - 0.03)^4 (1 - 0.05)^6 = 20000 (0.97)^4 (0.95)^6 = 13015.43 \approx \text{Rs. } 13015$$

**Example 7.** A firm of ready-made garments make both men's and women's shirts. Its profits average 6% of sales, its profits in men's shirts average 8% of sales and women's shirts comprise 60% of output. What is the average percentage profit in women's shirts ?

**Sol.** Total return in men's shirts = $G_1$ = 108

Total return in all shirts = $G_{12}$ = 106

Let the total return in women's shirts be $G_2$

Further it is given that $n_1 = 40$ and $n_2 = 60$

Now $$G_{12} = \text{A.L.}\left(\frac{n_1 \log G_1 + n_2 \log G_2}{n_1 + n_2}\right)$$

$$\therefore \quad 106 = \text{A.L.}\left(\frac{40 \times \log 108 + 60 \times \log G_2}{40 + 60}\right)$$

$$\Rightarrow \quad \log(106) = \frac{4 \times 2.0334 + 6 \log G_2}{10}$$

$$\Rightarrow \quad 10 \times 2.0253 = 4 \times 2.0334 + 6 \log G_2$$

$$\Rightarrow \quad 6 \log G_3 = 12.1194$$

$$\Rightarrow \quad \log G_3 = 2.0199$$

$$\therefore \quad G_3 = \text{A.L. } 2.0199 = 104.7$$

$\therefore$ Average percentage profit on sales from women's shirts = 104.7–100 = 4.7%

# EXERCISE 7.7

1. Calculate the geometric mean of the following price relatives :

| Commodity : | Wheat | Rice | Pulses | Sugar | Salt | Oils |
|---|---|---|---|---|---|---|
| Price Relatives : | 207 | 198 | 156 | 124 | 107 | 196 |

2. Calculate G.M. of the following data :

   0.9842, 0.3154, 0.0252, 0.0068, 0.0200, 0.0002, 0.5444, 0.4010

3. Calculate the geometric mean of the following two series :

   Series $a$ : 2574, 475, 75, 5, 0.8, 0.08, 0.005, 0.0009

   Series $b$ : 0.8974, 0.0570, 0.0081, 0.5677, 0.0002, 0.0984, 0.0854, 0.5672

4. The annual rate of growth of output of a factory in 5 years are 5.0, 7.5, 2.5, 5.0 and 5.0 percent respectively. What is the compound rate of growth of output per annum for the period ?

5. An assesse depreciated the machinery of his factory by 10% each in the first two years and by 40% in the third year and there by claimed 21% average depreciation relief from taxation department but the I.T.O. objected and allowed only 20%. Confirm which of the two is right ?

6. Find the Geometric Mean from the following data :

| X : | 10 | 20 | 30 | 40 | 50 | 60 |
|---|---|---|---|---|---|---|
| $f$ : | 12 | 15 | 25 | 10 | 6 | 2 |

7. Compute Geometric Mean from the following data relating to the weight of 100 individuals :

| Weight (lbs.) | No. of Persons | Weight (lbs.) | No. of Persons |
|---|---|---|---|
| 100 – 110 | 14 | 130 – 140 | 20 |
| 110 – 120 | 16 | 140 – 150 | 15 |
| 120 – 130 | 30 | 150 – 160 | 5 |

8. Find geometric mean of the following :

| Marks : | 0 – 10 | 10 – 20 | 20 – 30 | 30 – 40 | 40 – 50 |
|---|---|---|---|---|---|
| No. of Students : | 5 | 7 | 15 | 25 | 8 |

9. The bacterial count in a certain culture increased from 1000 to 4000 in three days. What was the average percentage increase per day ?

10. A machine reduces to half its original value in a period of 20 years. What is the average percentage depreciation per year ?

11. The value of a machine decreases at a constant rate from the cost price of Rs. 10000 to the scrap value of Rs. 1000 in a period of 10 years. Find the average percentage rate of decrease per year.

12. Initially the population of town was 5000. At first it rose at the rate of 5% per annum for first 3 years and then it decreased at the rate of 2% p.a. for two years. What will be the population of the town after 5 years ?

13. A group consisting of five items has the geometric mean as 6 and another group of three items has G.M. as 7. Find out the G.M. of all the items taken together.

14. Three groups of observations contains 8, 7 and 5 observations. Their geometric means are 8.52, 10.12 and 7.75 respectively. Find the geometric mean of the 20 observation in the single group formed by pooling the three groups.

## ANSWERS

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. | 159.8 | 2. | 0.0511 | 3. | 1.841, 0.0622 | 4. | 5% |
| 5. | assesse | 6. | 25.29 | 7. | 126.3 lbs | 8. | 25.63 |
| 9. | 58.7% | 10. | 3.42% | 11. | 20.57% | 12. | 5559 |
| 13. | 6.357 | 14. | 8.836 | | | | |

## 7.8 HARMONIC MEAN

*Harmonic mean* is defined as the reciprocal of the arithmetic mean of the reciprocal of the items. *i.e.* $X_1, X_2, \ldots\ldots, X_n$ are $n$ observations in an individual series then

$$\text{Harmonic Mean H.M.} = \text{Reciprocal of } \frac{\frac{1}{X_1} + \frac{1}{X_2} + \ldots\ldots + \frac{1}{X_n}}{n}$$

*i.e.*
$$\text{H.M.} = \frac{n}{\Sigma\left(\frac{1}{X}\right)} \quad \ldots(7.27)$$

Further for a discrete frequency distribution of the form as in table 7.3,

| X | $X_1$ | $X_2$ | ............... | $X_n$ |
|---|---|---|---|---|
| $f$ | $f_1$ | $f_2$ | ............... | $f_n$ |

Table 7.3

$$\text{Harmonic Mean H.M} = \frac{N}{\frac{f_1}{X_1} + \frac{f_2}{X_2} + \ldots\ldots + \frac{f_n}{X_n}}$$

i.e. $$\text{H.M.} = \frac{N}{\Sigma\left(\frac{f}{X}\right)} \quad ...(7.28)$$

where $N = \Sigma f$.

For a continuous frequency distribution, we take the mid values of the class intervals as $X$ so that harmonic mean is given by formula (7.28)

## Merits and Demerits of Harmonic Mean

**Merits :** The following are the merits of harmonic mean :

(*i*) Harmonic mean is based on all the items in the series.

(*ii*) Since Harmonic mean involves the reciprocal of the items of the series, so it gives greater weightage to smaller items and less weightage to larger items. Thus it is not very much affected by one or two large items present in the series.

(*iii*) It is not very much affected by fluctuations of sampling.

(*iv*) This average is most suitable for the problems of averaging rates and ratios where time factor is variable and the act being performed (*e.g.* distance) remains constant. (refer to sec. 7.8.4 note (*b*))

(*v*) It is capable of further algebraic treatment.

For example, if $H_1$ and $H_2$ are harmonic means of two groups of size $n_1$ and $n_2$ respectively then harmonic mean of the combined group can be calculated by the formula

$$\frac{1}{H} = \frac{1}{n_1 + n_2}\left[\frac{n_1}{H_1} + \frac{n_2}{H_2}\right] \quad ...(7.29)$$

**Demerits :** The following are the demerits of harmonic mean :

(*i*) This average is difficult to compute and not easily understandable.

(*ii*) Harmonic mean can be computed only when all the items are known.

(*iii*) Harmonic mean cannot be calculated if any one item has zero value.

(*iv*) Harmonic mean is the value which usually does not belong to the series.

## Relationship Between Arithmetic Mean, Geometric Mean and Harmonic Mean

If A.M., G.M. and H.M. are the arithmetic mean, geometric mean and harmonic mean of a series with $n$ observations then there are related to each other in the following relations :

(*i*) $$\text{A.M.} \geq \text{G.M.} \geq \text{H.M.} \quad ...(7.30)$$

where the sign of equality holds if and only if all the $n$ observations are equal.

(*ii*) $$(\text{G.M.})^2 = \text{A.M.} \times \text{H.M.} \quad ...(7.31)$$

**Proof :** To prove this relationship, we assume the series with only two observations $a$ and $b$ (where $a$ and $b$ are positive real numbers) for sake of simplicity and clarity although the relationship holds for any number of observations.

Now, by definition

$$\text{A.M.} = \frac{a+b}{2}$$

$$\text{G.M.} = \sqrt{ab}$$

$$\text{H.M.} = \frac{2}{\frac{1}{a}+\frac{1}{b}} \text{ or } \frac{2}{\frac{b+a}{ab}} = \frac{2ab}{a+b}$$

We know that $(\sqrt{a}-\sqrt{b})^2 \geq 0$ as square of a real number can not be negative.

$$\Rightarrow \quad a+b-2\sqrt{ab} \geq 0$$

$$\Rightarrow \quad a+b \geq 2\sqrt{ab}$$

$$\Rightarrow \quad \frac{a+b}{2} \geq \sqrt{ab}$$

or $$\text{A.M.} \geq \text{G.M.} \qquad \text{...(7.32)}$$

Again, $$(\sqrt{a}-\sqrt{b})^2 \geq 0$$

$$\Rightarrow \quad a+b-2\sqrt{ab} \geq 0$$

$$\Rightarrow \quad a+b \geq 2\sqrt{ab}$$

Dividing both sides by $a\,b$, we get

$$\frac{a+b}{ab} \geq \frac{2}{\sqrt{ab}}.$$

$$\sqrt{ab} \geq \frac{2ab}{a+b}$$

or $$\text{G.M.} \geq \text{H.M.} \qquad \text{...(7.33)}$$

From (7.32) and (7.33), we have

$$\text{A.M.} \geq \text{G.M.} \geq \text{H.M.}$$

Further, if $a = b$

then $$\text{A.M.} = \text{G.M.} = \text{H.M.}$$

(*ii*) $$\text{A.M.} \times \text{H.M.} = \frac{a+b}{2} \times \frac{2ab}{a+b} = ab = (\text{G.M.})^2$$

Hence $$(\text{G.M.})^2 = (\text{A.M.}) \times (\text{H.M.})$$

### 7.8.1 CALCULATION OF HARMONIC MEAN FOR INDIVIDUAL SERIES

Harmonic mean for an individual series is calculated by using the formula (7.27)

*i.e.* Harmonic Mean $\text{H.M.} = \dfrac{n}{\Sigma\left(\dfrac{1}{X}\right)}$

## 7.8.2 CALCULATION OF HARMONIC MEAN FOR DISCRETE SERIES

Harmonic mean for a discrete series is calculated by using the formula (7.28)

*i.e.* Harmonic Mean H.M. $= \dfrac{N}{\Sigma\left(\dfrac{f}{X}\right)}$ where $N = \Sigma f$.

## 7.8.3 CALCULATION OF HARMONIC MEAN FOR CONTINUOUS SERIES

For a continuous series, we take mid values of $m$ of the class-intervals and apply the formula of the form (7.28).

*i.e.* Harmonic Mean H.M. $= \dfrac{N}{\Sigma\left(\dfrac{f}{m}\right)}$

where $N = \Sigma f$ and $m$ represents the mid values of class intervals.

## 7.8.4 CALCULATION OF WEIGHTED HARMONIC MEAN

The weighted harmonic mean is calculated by the formula

$$\text{Weighted Harmonic Mean H.M.}_W = \frac{\Sigma W}{\Sigma\left(\dfrac{W}{X}\right)} \qquad ...(7.35)$$

where W represents the weights attached to different items.

Notes (*a*) If $W_1, W_2, ......, W_n$ are the distances covered by a body with speed $X_1, X_2, ...., X_n$ respectively then average speed of the body is given by weighted harmonic mean of speeds $X_1, X_2, ...., X_n$ with weights $W_1, W_2, ......, W_n$ respectively.

*i.e.* $$\text{average speed} = \frac{\Sigma W}{\Sigma\left(\dfrac{W}{X}\right)}.$$

Proof : Here, total distance covered by body $= W_1 + W_2 + ...... + W_n$.

Also, total time taken by body to cover the total distance $= \dfrac{W_1}{X_1} + \dfrac{W_2}{X_2} + ..... + \dfrac{W_n}{X_n}$

$$\therefore \quad \text{average speed} = \frac{\text{Total distance covered}}{\text{Total time taken}}$$

$$= \frac{W_1 + W_2 + ..... + W_n}{\dfrac{W_1}{X_1} + \dfrac{W_2}{X_2} + ..... + \dfrac{W_n}{X_n}} = \frac{\Sigma W}{\Sigma \dfrac{W}{X}}$$

(*b*) If a body covers same distance say W with different speeds say $X_1, X_2, ....., X_n$ successively then average speed of body is given by harmonic mean of speeds $X_1, X_2, ....., X_n$

*i.e.* $$\text{average speed} = \frac{n}{\Sigma \frac{1}{X}}$$

**Proof :** As discussed above,

$$\text{average speed} = \frac{W_1 + W_2 + ..... + W_n}{\frac{W_1}{X_1} + \frac{W_2}{X_2} + ..... + \frac{W_n}{X_n}}$$

$$= \frac{W + W + ..... + W}{\frac{W}{X_1} + \frac{W}{X_2} + ..... + \frac{W}{X_n}} \qquad (\because W_1 = W_2 = ...... = W_n = W)$$

$$= \frac{nW}{W\left(\frac{1}{X_1} + \frac{1}{X_2} + ..... + \frac{1}{X_n}\right)}$$

$$= \frac{n}{\Sigma \frac{1}{X}}$$

## CHECKPOINTS

1. What do you understand by Harmonic Mean ?

(P.U. B.C.A. April 2008, Sept. 2008)

2. What are the merits and demerits of Harmonic Mean ?
3. Explain the purpose or significance of Harmonic Mean.

(G.N.D.U. B.Sc. C.Sc. April 2005, 2006, 2007)

4. How A.M., G.M. and H.M.s are related ?

# ILLUSTRATIVE EXAMPLES

**Example 1.** Find the H.M. of the following data :

15, 20, 21, 22, 26, 29

(G.N.D.U. B.C.A. April 2005; B.Sc. I.T. 2007)

Sol. First, we construct the following table :

| X | 1/X |
|---|---|
| 15 | 0.0667 |
| 20 | 0.0500 |
| 21 | 0.0476 |
| 22 | 0.0454 |
| 26 | 0.0385 |
| 29 | 0.0345 |
| | $\Sigma$ 1/X = 0.2827 |

Now, Harmonic Mean $\text{H.M.} = \dfrac{n}{\Sigma \dfrac{1}{X}} = \dfrac{6}{0.2827} = 21.22$

Example 2. Find the harmonic mean for the following frequency distribution :

| Wages in Rs. : | 40 – 50 | 50 – 60 | 60 – 70 | 70 – 80 | 80 – 90 | 90 – 100 |
|---|---|---|---|---|---|---|
| No. of workers : | 12 | 10 | 15 | 17 | 8 | 3 |

Sol. First, we construct the following table :

| Wages in Rs. | M.V. ($m$) | No. of Workers ($f$) | $f/m$ |
|---|---|---|---|
| 40 – 50 | 45 | 12 | 12/45 = 0.2667 |
| 50 – 60 | 55 | 10 | 10/55 = 0.1818 |
| 60 – 70 | 65 | 15 | 15/65 = 0.2308 |
| 70 – 80 | 75 | 17 | 17/75 = 0.2267 |
| 80 – 90 | 85 | 8 | 8/85 = 0.0941 |
| 90 – 100 | 95 | 3 | 3/95 = 0.0316 |
| | | N = 65 | $\Sigma f/m$ = 1.0317 |

$\therefore$ Harmonic mean $\text{H.M.} = \dfrac{N}{\Sigma\left(\dfrac{f}{m}\right)} = \dfrac{65}{1.0317} = 63$

Example 3. A taxi-car drives from a plain town on a hill station, 60 kms. distant, at a mileage rate of 10 kms. per gallon of petrol and on return trip at 15 kms. per gallon. Find the average distance per gallon. Verify your result.

Sol. Let us take the mileage rate (km./gallon) as X-axis.

Since distance covered by taxi car is same so average distance per gallon *i.e.* average mileage rate is given by harmonic mean of X-series.

$$\therefore \quad \text{average mileage rate} = \frac{n}{\Sigma \frac{1}{X}} = \frac{2}{\frac{1}{10}+\frac{1}{15}} = \frac{2\times 150}{15+10} = \frac{300}{25} = 12 \text{ km/gallon.}$$

**Verification**

$$\text{Total petrol consumed while going} = \frac{60}{10} = 6 \text{ gallons}$$

$$\text{Total petrol consumed coming back} = \frac{60}{15} = 4 \text{ gallons}$$

Total petrol consumed both ways = 10 gallons

Total distance covered both ways = 120 kms

$$\therefore \quad \text{Distance per gallon} = \frac{120}{10} = 12 \text{ kms/gallon.}$$

**Example 4.** A man travels first 900 kms. of his journey by train at an average speed of 80 kms. per hour, next 2000 kms. by plane at an average speed of 300 kms. per hour and finally 20 kms. by taxi at an average speed of 30 kms. per hour. What is his average speed for the entire journey ?

**Sol.** Let us take the speed as 'X' series and distance covered as 'W' series. Now we construct the following table :

| Speed (kms./hr) (X) | Distance Covered (W) | W/X |
|---|---|---|
| 80 | 900 | 900/80 = 11.25 |
| 300 | 2000 | 2000/300 = 6.67 |
| 30 | 20 | 20/30 = 0.67 |
| | ΣW = 2920 | ΣW /X = 18.59 |

$$\therefore \quad \text{Average speed} = \frac{\Sigma W}{\Sigma\left(\frac{W}{X}\right)} = \frac{2920}{18.59} = 157.07 \text{ km./hr.}$$

## EXERCISE 7.8

1. Calculate the H.M. of the following individual series :

   4, 7, 10, 12, 19

   (G.N.D.U. B.C.A. April 2004)

2. The following table gives the weights of 31 persons in a sample enquiry. Calculate Harmonic mean.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Weight (lbs) : | 130 | 135 | 140 | 145 | 146 | 148 | 149 | 150 | 157 |
| No. of Persons : | 3 | 4 | 6 | 6 | 3 | 5 | 2 | 1 | 1 |

3. Find the Harmonic Mean from the following data

| | | | | | |
|---|---|---|---|---|---|
| X : | 10 | 20 | 25 | 40 | 50 |
| $f$ : | 20 | 30 | 50 | 15 | 5 |

4. Calculate Harmonic Mean of the following data :

| | | | | |
|---|---|---|---|---|
| Class : | 2 – 4 | 4 – 6 | 6 – 8 | 8 – 10 |
| Frequency : | 20 | 40 | 30 | 10 |

5. Obtain consumer price index numbers from the following groups using (*i*) Arithmetic mean (*ii*) Geometric mean (*iii*) Harmonic mean and confirm that : $A.M_w > G.M_w > H.M_w$

| Group | Index Number (X) | Weights (W) |
|---|---|---|
| Food | 352 | 48 |
| Fuel | 220 | 10 |
| Cloth | 230 | 8 |
| House Rent | 160 | 12 |
| Miscellaneous | 190 | 15 |

6. An aeroplane flies around a square. The aeroplane covers the first side at a speed of 100 km./hr., second side at an average speed of 200 km./hr., the third at an average speed of 300 km./hr. and the fourth at an average speed of 400 km./hr. What is the average speed around the square ?

7. A train starts from rest and travels successive quarters of a km. at average speeds of 12, 16, 24 and 48 km./hr. The average speed over the whole journey is 19·2 km./hr. and not 25 km./hr. Explain and verify.

8. A man travels from Delhi to Mathura at the speed of 60 km/hr and returns at the speed of 40 km/hr. Find his average speed.

9. If the interest paid on each of three different sums of money yielding 5%, 6% and 7% simple interest per annum respectively is the same, what is the average yield percent on the total sum invested ?

10. A railway train runs for 30 minutes at speed of 40 km/hr., then because of repair of the track runs for 10 minutes at a speed of 8 km/hr., after which it resumes its previous speed and runs for 20 minutes excep. for a period of 2 minutes when it has to run over a bridge with a speed of 30 km/hr. What is the average speed ?

## ANSWERS

1. 7.95
2. 142.33
3. 20.08
4. 4.98
5. (*i*) 276.41 (*ii*) 263.8 (*iii*) 250.94
6. 192 km/hr
8. 48 km/hr
9. 5.89%
10. 34.33 km/hr

## 7.9 MEDIAN

The next measure of central tendency is *median*. It is a positional measure of central tendency. Median means middle value. It divides the series into two equal parts when data are arranged in ascending or descending order. It is generally denoted by 'M'.

According to **Prof. L.R. Connor**, "***The median is that value of the variable which divides the group into two equal parts, one comprising all values greater than and other, all values less than the median***". Because of this characteristic median is known as the positional average. Median is not based on the values of all the items instead it is concerned with the number of items in series only.

In order to locate the median value, all the items of the series are arranged either in ascending order or descending order. The calculation of the median value in case of different series is explained in the following subsections.

### 7.9.1 MEDIAN IN CASE OF INDIVIDUAL SERIES

In case of individual series median is calculated by the following formula :

$$\text{Median M} = \text{Size or value of } \left(\frac{n+1}{2}\right)\text{th item} \qquad ...(7.36)$$

where $n$ is number of observations

Two cases arise in individual series :

***(a)* Odd Number Series :** Calculation of median in case of series having odd number items involves the following steps :

**Step 1 :** Arrange the terms in ascending or descending order.

**Step 2 :** Count the number of items and put it equal to $n$.

**Step 3 :** The size or value of $\left(\frac{n+1}{2}\right)$th item gives the median of the series.

***(b)* Even Number Series :** If number of items *i.e.* '$n$' happens to be an even number then the quantity $\frac{n+1}{2}$ will be in fractions. Therefore, in this case median would be equal to mean of two middle values. For example, if $\frac{n+1}{2} = 3.5$, then median $= \frac{\text{size of 3rd item} + \text{size of 4th item}}{2}$. $= \dfrac{\frac{n}{2}\text{th item} + \left(\frac{n}{2}+1\right)\text{th}}{2}$

### 7.9.2 MEDIAN IN CASE OF DISCRETE SERIES

In the case of discrete series median is determined by the following formula :

$$\text{Median} = \text{Size or value of } \frac{N+1}{2}\text{ th item} \qquad ...(7.37)$$

Where $N = \Sigma f$ or sum of frequencies.

Calculation of median in case of discrete series involves following steps :

**Step 1 :** Arrange the data in ascending or descending order.

**Step 2 :** Calculate cumulative frequencies.

**Step 3 :** Calculate $\frac{N+1}{2}$

**Step 4 :** Locate the cumulative frequency ($c.f.$) which is equal to or next higher to $\frac{N+1}{2}$ th item.

**Step 5 :** The value of variable corresponding to this cumulative frequency is median value.

## 7.9.3 MEDIAN IN CASE OF CONTINUOUS SERIES

In continuous series median is calculated by following formula :

$$\text{Median} = \text{Size or value of } \frac{N}{2}\text{th item} \quad ...(7.38)$$

where $N = \Sigma f$

Calculation of median in case of continuous series involves following steps :

**Step 1 :** Data is first arranged in the ascending order of magnitude.

**Step 2 :** Calculate cumulative frequencies.

**Step 3 :** Calculate $\frac{N}{2}$.

**Step 4 :** Locate the cumulative frequency which is equal to or next higher to $\frac{N}{2}$. The class interval corresponding to this cumulative frequency contains the median and this class interval is called *median class*.

**Step 5 :** Apply the following formula :

$$\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i \quad ...(7.39)$$

where L is the lower limit of the median class interval

$i$ = width of class interval of median class

$f$ = frequency of the median class

$c.f.$ = cumulative frequency of the class before the median class

Median can also be calculated by the following formula if in data is arranged in the descending order

$$\text{Median} = U - \frac{\frac{N}{2} - c.f.}{f} \times i \quad ...(7.40)$$

Where U is the upper limit of median class and the other notations are same as given above.

**Merits and Demerits of Median**

**Merits :** The following are the merits of median :

(*i*) Median is simply the size of the middle item after the observations are arranged in the ascending or in the descending order of magnitude. Thus, it is easy to understand and simple to explain.

(*ii*) Its calculations are not complex. One can to simply get the value of middle item after their proper arrangement.

(*iii*) Median can be calculated even if we have the open ends frequency distribution or the class intervals are unequal.

(*iv*) Median is least affected by the presence of extreme values in the series. Median is merely the value of the middle item.

(*v*) Median can be calculated even when the qualitative characteristics of a problem for example playing habits, beauty etc. are given.

(*vi*) Median is used in a number of other statistical devices such as skewness, mean deviation, sampling etc.

(*vii*) Median can be calculated graphically by using ogive (refer to Sec. 7.9.5).

**Demerits :** The following are the demerits or limitations of median :

(*i*) Median is not based on all the items of the series and as such it is not a true representative of the data.

(*ii*) To arrange the data in the ascending or descending order of magnitude is an additional hazard. Even these efforts go futile if a few items are added or removed from the series.

(*iii*) Because its calculations are not based on concrete mathematical formulae, therefore, one can question its reliability and use in further statistical measures.

(*iv*) If the median is known we have no scope to find the sum of the observations whereas it is possible in case of arithmetic mean by using the formula $\sum X = n\overline{X}$ .

(*v*) Although median is the value of middle item in the series but if the number of items is even we have to take the help of arithmetic mean to calculate the median by taking the A.M. of the two middle values.

(*vi*) A very small change in the series may have considerable effect on the value of the median. For example the observations 5, 15, 23, 45, 50 have the median as 23 because it is the value of the middle *i.e.*, the third item in the series. If only one item say 54 is introduced in the series the median will change to the A.M. of the two middle items *i.e.*, its revised value will be $\frac{23+45}{2}$ *i.e.* 34, which is quite large as compared to the previous value of the median.

(*vii*) If the median of two or more series with the number of items are given, we have no technique to find the combined median of these series.

## CHECKPOINTS

1. What do you understand by Median ? (P.U. B.C.A. April 2008)
2. Explain the purpose or significance of Median. (G.N.D.U. B.Sc. C.Sc. April 2005, 2006, 2007)
3. Compare mean and median in a tabular manner. (P.U. B.C.A. April 2005, Sept. 2006; G.N.D.U. B.Sc. I.T. April 2009)
4. Give algorithm to find median for a set of observations. (G.N.D.U. B.C.A. April 2007)
5. What are the merits and demerits of median ?

# ILLUSTRATIVE EXAMPLES

**Example 1.** Calculate Median from the following values : odd

30, 45, 75, 65, 50, 52, 28, 40, 49, 35, 52.

**Sol.** Arrange the given values in the ascending order as under :

28, 30, 35, 40, 45, 49, 50, 52, 52, 65, 75

Now Median = Value of $\frac{n+1}{2}$th item = Value of $\frac{11+1}{2}$ th item = Value of 6th item

$\therefore$ Median = 49

**Example 2.** Calculate Median from the following data : — even value

36, 32, 28, 22, 26, 20, 18, 40.

**Sol.** Arrange the given values in the ascending order as under :

18, 20, 22, 26, 28, 32, 36, 40.

Median = Value of $\frac{n+1}{2}$th item = Value of $\frac{8+1}{2}$ th item = Value of 4.5th item

$\therefore$ Median is the A.M. of 4th and 5th items

$\therefore$ Median = $\frac{26+28}{2} = 27$

**Example 3.** Calculate the value of Median from the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Income (in Rs.) : | 100 | 150 | 80 | 200 | 250 | 180 |
| No. of persons : | 12 | 13 | 8 | 10 | 3 | 15 |

**Sol.** Arranging the data in the ascending order of magnitude, we have :

| Income (in Rs.) | No. of Persons ($f$) | c.f. |
|---|---|---|
| 80 | 8 | 8 |
| 100 | 12 | 20 |
| 150 | 13 | 33 |
| 180 | 15 | 48 |
| 200 | 10 | 58 |
| 250 | 03 | 61 |
| | N = 61 | |

Median = Value of $\frac{N+1}{2}$ th item = Value of $\frac{61+1}{2}$ th item = Value of 31st item.

The c.f. next higher to 31 is 33. Therefore, median = 150

**Example 4.** Find median for the following data :

| | | | | |
|---|---|---|---|---|
| Class : | 1 – 3 | 3 – 5 | 5 – 7 | 7 – 9 |
| Frequency : | 40 | 30 | 20 | 10 |

(G.N.D.U. B.Sc. C.Sc. April 2004)

**Sol.** First, we prepare the following table :

| Class (X) | Frequency ($f$) | c. f. |
|---|---|---|
| 1 – 3 | 40 | 40 |
| 3 – 5 | 30 | 70 |
| 5 – 7 | 20 | 90 |
| 7 – 9 | 10 | 100 |
| | N = 100 | |

Now $\quad$ Median $= \text{Size of } \frac{N}{2} \text{ th item} = \text{Size of } \frac{100}{2} \text{ th item} = \text{Size of 50th item}$

The $c.f.$ next higher to 50 is 70.

∴ Median lies in the class interval 3–5

So $\quad \boxed{\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i} = 3 + \frac{50-40}{30} \times 2 = 3.67$

**Example 5.** The frequency distribution of weight (in grams) of mangoes are given below. Calculate the arithmetic mean and the median.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Weight (in grams) : | 410–419 | 420–429 | 430–439 | 440–449 | 450–459 | 460–469 | 470–479 |
| No. of Mangoes : | 14 | 20 | 42 | 54 | 45 | 18 | 7 |

(G.N.D.U. B.Sc. I.T. April 2009)

**Sol.** First write the given data in exclusive form of class-intervals for finding median and prepare the following table :

| Weight (in gms.) (X) | No. of Mangoes ($f$) | $c.f.$ | Mid Value ($m$) | A = 444.5 $dx = m - A$ | $c = 10$ $d'x$ | $fd'x$ |
|---|---|---|---|---|---|---|
| 409.5–419.5 | 14 | 14 | 414.5 | – 30 | – 3 | – 42 |
| 419.5–429.5 | 20 | 34 | 424.5 | – 20 | – 2 | – 40 |
| 429.5–439.5 | 42 | 76 | 434.5 | – 10 | –1 | – 42 |
| 439.5–449.5 | 54 | 130 | 444.5 | 0 | 0 | 0 |
| 449.5–459.5 | 45 | 175 | 454.5 | 10 | 1 | 45 |
| 459.5–469.5 | 18 | 193 | 464.5 | 20 | 2 | 36 |
| 469.5–479.5 | 7 | 200 | 474.5 | 30 | 3 | 21 |
| | N = 200 | | | | | $\Sigma fd'x = -22$ |

Now, Arithmetic Mean $\overline{X} = A + \frac{\Sigma fd'x}{N} \times c = 444.5 + \frac{-22}{200} \times 10 = 443.4$

Also $\quad \text{Median} = \text{Value of } \frac{N}{2}\text{th item} = \text{Value of } \frac{200}{2}\text{th item} = \text{Value of 100th item}$

The *c.f.* next higher to 100 is 130.

Therefore, Median lies in class interval 439.5 – 449.5

$$\therefore \quad \text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i = 439.5 + \frac{100-76}{54} \times 10 = 439.5 + \frac{240}{54} = 443.94$$

**Example 6.** Determine the median for the following distribution of monthly income of 580 lower-class people :

| | | | | | |
|---|---|---|---|---|---|
| Income (Rs.) : | less than 300 | 300–350 | 350–400 | 400–450 | 450–500 |
| Frequency : | 53 | 81 | 114 | 195 | 63 |
| Income (Rs.) : | 500–550 | 550–600 | 600–650 | 650–700 | more than or equal to 700 |
| Frequency : | 32 | 20 | 11 | 8 | 3 |

(G.N.D.U. B.Sc. C.Sc. April 2005)

**Sol.** First we prepare the following table :

| Income (Rs.) (X) | Frequency ($f$) | *c. f.* |
|---|---|---|
| Less than 300 | 53 | 53 |
| 300 – 350 | 81 | 134 |
| 350 – 400 | 114 | 248 |
| 400 – 450 | 195 | 443 |
| 450 – 500 | 63 | 506 |
| 500 – 550 | 32 | 538 |
| 550 – 600 | 20 | 558 |
| 600 – 650 | 11 | 569 |
| 650 – 700 | 8 | 577 |
| More than or equal to 700 | 3 | 580 |
| | N = 580 | |

Now $\quad \text{Median} = \text{Size of } \frac{N}{2}\text{th item} = \text{Size of } \frac{580}{2}\text{th item} = \text{Size of 290th item}.$

The *c.f.* next higher to 290 is 443.

$\therefore$ median lies in class interval 400 – 450.

$$\text{So} \quad \text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i = 400 + \frac{290-248}{195} \times 50 = 410.77$$

**Example 7.** Calculate Median of the following series :

| Wages less than (Rs.) : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|---|---|
| No. of Workers : | 15 | 35 | 60 | 84 | 96 | 127 | 198 | 250 |

**Sol.** To find median we convert cumulative series into simple frequency series as follows :

| Wages (X) | No. of Workers ($f$) | $c.f$ |
|---|---|---|
| 0 – 10 | 15 | 15 |
| 10 – 20 | 20 ( = 35 – 15 ) | 35 |
| 20 – 30 | 25 ( = 60 – 35 ) | 60 |
| 30 – 40 | 24 ( = 84 – 60) | 84 |
| 40 – 50 | 12 ( = 96 – 84 ) | 96 |
| 50 – 60 | 31 ( = 127 – 96 ) | 127 |
| 60 – 70 | 71 ( = 198 – 127) | 198 |
| 70 – 80 | 52 ( = 250 – 198 ) | 250 |
| | N = 250 | |

Now Median = Value of $\frac{N}{2}$ th item = Value of $\frac{250}{2}$ th item = Value of 125th item

The $c.f.$ next higher to 125 is 127.

$\therefore$ Median lies in the class interval 50 – 60

$$\therefore \quad \text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i$$

$$= 50 + \frac{125-96}{31} \times 10 = 50 + \frac{290}{31} = 50 + 9.35 = 59.35$$

**Example 8.** Find the median of the following data :

| Mid Points : | 5 | 15 | 25 | 35 | 45 |
|---|---|---|---|---|---|
| Frequency : | 3 | 9 | 8 | 5 | 3 |

**Sol.** Since the uniform gap between the mid values is 10 therefore corresponding class intervals will be (5–5) to (5 + 5), (15 – 5) to (15 + 5) etc. Thus the given data can be written in the exclusive form as under :

| Classes (X) | Frequency ($f$) | $c.f.$ |
|---|---|---|
| 0 – 10 | 3 | 3 |
| 10 – 20 | 9 | 12 |
| 20 – 30 | 8 | 20 |
| 30 – 40 | 5 | 25 |
| 40 – 50 | 3 | 28 |
| | N = 28 | |

Now Median = Value of $\frac{N}{2}$ th item = Value of $\frac{28}{2}$ th item = Value of 14th item

The $c.f.$ next higher to 14 is 20.

Therefore, Median lies in class interval 20 – 30

Now $$\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i = 20 + (14-12)\frac{10}{8} = 20 + \frac{20}{8} = 22.5$$

Example 9. In a factory employing 3000 persons in a day, 5% work less than 3 hours, 580 works from 3.01 to 4.50 hours, 30% work from 4.51 to 6.00 hours, 500 work from 6.01 to 7.50 hours, 20% work from 7.51 to 9.00 hours and the rest work 9.01 or more hours. What is the median hours of work ?

(P.U. B.C.A. Sept. 2004)

Sol. First we prepare the following table :

| Working hours | No. of Workers |
|---|---|
| less than 3 hours | $3000 \times \frac{5}{100} = 150$ |
| 3.01 – 4.50 | 580 |
| 4.51 – 6.00 | $3000 \times \frac{30}{100} = 900$ |
| 6.01 – 7.50 | 500 |
| 7.51 – 9.00 | $3000 \times \frac{20}{100} = 600$ |
| 9.01 or more | 3000 – 150 – 580 – 900 – 500 – 600 = 270 |

Now we complete the class-intervals and convert inclusive series into exclusive series as follows :

| Working hours (X) | No. of Workers ($f$) | $c.f.$ |
|---|---|---|
| 1.505 – 3.005 | 150 | 150 |
| 3.005 – 4.505 | 580 | 730 |
| 4.505 – 6.005 | 900 | 1630 |
| 6.005 – 7.505 | 500 | 2130 |
| 7.505 – 9.005 | 600 | 2730 |
| 9.005 – 10.505 | 270 | 3000 |
| | N = 3000 | |

Now Median = Size of $\frac{N}{2}$ th item = Size of $\frac{3000}{2}$ th item = Size of 1500th item

The $c.f.$ next higher to 1500 is 1630.

∴ Median lies in the class-interval 4.505 – 6.005.

So $$\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i = 4.505 + \frac{1500-730}{900} \times 1.5 = 5.79 \text{ hours}$$

**Example 10.** The median of the following data is 20.75. Find the missing frequencies $x$ and $y$ if the total frequencies is 100.

| Class interval : | 0–5 | 5–10 | 10–15 | 15–20 | 20–25 | 25–30 | 30–35 | 35–40 |
|---|---|---|---|---|---|---|---|---|
| Frequency : | 7 | 10 | $x$ | 13 | $y$ | 10 | 14 | 9 |

(P.U. B.C.A. Sept. 2008)

**Sol.** First, we prepare the following table :

| Class Interval (X) | Frequency ($f$) | $c.f.$ |
|---|---|---|
| 0 – 5 | 7 | 7 |
| 5 – 10 | 10 | 17 |
| 10 – 15 | $x$ | $17+x$ |
| 15 – 20 | 13 | $30+x$ |
| 20 – 25 | $y$ | $30+x+y$ |
| 25 – 30 | 10 | $40+x+y$ |
| 30 – 35 | 14 | $54+x+y$ |
| 35 – 40 | 9 | $63+x+y$ |
| | $N = 63+x+y$ | |

Given $N = 100$

$$\therefore \quad 63 + x + y = 100$$

$$\Rightarrow \quad x + y = 37 \quad \text{...(i)}$$

Also median is 20.75 which lies in the class-interval 20 – 25

Since
$$\text{median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i$$

160/2

$$\therefore \quad 20.75 = 20 + \frac{50 - (30 + x)}{y} \times 5$$

$$\Rightarrow \quad 20.75\,y = 20\,y + 100 - 5\,x$$

$$\Rightarrow \quad 5\,x + 0.75\,y = 100$$

$$\Rightarrow \quad x + 0.15\,y = 20 \quad \text{...(ii)}$$

Subtracting equation (*ii*) from (*i*), we have,

$$0.85\,y = 17 \quad \text{or} \quad y = \frac{17}{0.85} = 20$$

So from equation (*i*), $x = 37 - 20 = 17$

Hence the missing frequencies are $x = 17$ and $y = 20$.

# EXERCISE 7.9

1. Calculate Median from the following data :

96, 98, 75, 20, 102, 100, 94, 75.

2. A class of 16 boys and 16 girls were given a common intelligence test and the following marks were obtained :

| Sr. No.: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Boys : | 15 | 35 | 43 | 46 | 48 | 48 | 49 | 50 | 55 | 56 | 60 | 64 | 71 | 75 | 80 | 85 |
| Girls : | 10 | 30 | 45 | 52 | 55 | 58 | 61 | 61 | 63 | 69 | 70 | 72 | 74 | 75 | 75 | 90 |

Calculate the median marks of boys and girls.

3. The following marks have been obtained by three students in an examination in a paper. Calculate median marks and show which student has the higher general level of knowledge ?

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A : | 36 | 56 | 41 | 46 | 54 | 59 | 55 | 51 | 52 | 44 | 37 | 59 |
| B : | 58 | 54 | 21 | 51 | 59 | 46 | 65 | 31 | 68 | 41 | 70 | 36 |
| C : | 65 | 55 | 26 | 40 | 30 | 74 | 45 | 29 | 85 | 32 | 80 | 39 |

4. Given below the marks obtained by 20 students in a certain class test in History and Geography.

| Roll No. | Marks in History | Marks in Geography | Roll No. | Marks in History | Marks in Geography |
|---|---|---|---|---|---|
| 1 | 53 | 58 | 11 | 25 | 10 |
| 2 | 54 | 55 | 12 | 42 | 42 |
| 3 | 52 | 25 | 13 | 33 | 15 |
| 4 | 32 | 32 | 14 | 48 | 46 |
| 5 | 30 | 26 | 15 | 72 | 50 |
| 6 | 60 | 85 | 16 | 51 | 64 |
| 7 | 47 | 44 | 17 | 45 | 39 |
| 8 | 46 | 80 | 18 | 33 | 38 |
| 9 | 35 | 33 | 19 | 65 | 30 |
| 10 | 28 | 72 | 20 | 29 | 36 |

Using median show in which subject the level of knowledge of the students is higher ?

5. Find the median from the following frequency distribution :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| X : | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
| F : | 5 | 9 | 10 | 12 | 6 | 4 | 2 |

(G.N.D.U. B.C.A. April 2004)

6. Calculate mean, median and mode from the following frequency distribution of marks at a test in statistics :

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Marks : | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
| No. of Students : | 20 | 43 | 75 | 76 | 72 | 45 | 36 | 9 | 8 | 6 |

7. Find the Median size from the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Size : | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| Frequency : | 2 | 6 | 10 | 20 | 15 | 8 | 7 | 3 |

8. Find out Median from the following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Variable : | 17 | 25 | 45 | 20 | 24 | 32 | 10 |
| Frequency : | 100 | 125 | 128 | 111 | 127 | 118 | 115 |

9. Find median from the following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Group : | 0–5 | 5–10 | 10–15 | 15–20 | 20–25 | 25–30 | 30–35 |
| Frequency : | 25 | 125 | 100 | 300 | 250 | 125 | 20 |

10. Find out percentage median marks from following data :

| | | | | | |
|---|---|---|---|---|---|
| Marks (%) : | 0 – 20 | 20 – 40 | 40 – 60 | 60 – 80 | 80 – 100 |
| No. of Students : | 0 | 5 | 22 | 25 | 16 |

11. Below is given population of females in the different age groups. Compute the median age from the data.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Age Group (Yrs.) : | 5 – 14 | 15 – 24 | 25 – 34 | 35 – 44 | 45 – 54 | 55– 64 | 65 – 74 |
| No. of Females : | 420 | 300 | 206 | 205 | 133 | 86 | 40 |

12. Calculate median from following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Size : | 60–69 | 50–59 | 40–49 | 30–39 | 20–29 | 10–19 |
| Frequency : | 13 | 15 | 21 | 20 | 19 | 12 |

13. Find the median from the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Class : | 0 – 3 | 3 – 6 | 6 – 10 | 10 – 12 | 12 – 15 | 15 – 18 |
| Frequency : | 4 | 8 | 10 | 14 | 16 | 20 |
| Class : | 18 – 20 | 20 – 24 | 24 – 28 | 28 – 30 | 30 – 36 | |
| Frequency : | 24 | 14 | 11 | 11 | 6 | |

14. Calculate Median of the following data :

| | | | | | |
|---|---|---|---|---|---|
| Class Boundaries : | 349.5–449.5 | 449.5–499.5 | 549.5–599.5 | 649.5–699.5 | 749.5–799.5 |
| Frequency : | 1 | 3 | 6 | 18 | 13 |
| Class Boundaries : | 849.5–899.5 | 949.5–999.5 | | | |
| Frequency : | 6 | 4 | | | |

(P.U. B.C.A. Sept. 2007)

15. Find an appropriate measure of central tendency for the following distribution :

| Monthly Income (in Rs. ) in locality X | No. of Persons |
|---|---|
| Below 100 | 50 |
| 100 – 200 | 500 |
| 200 – 300 | 555 |
| 300 – 400 | 100 |
| 400 – 500 | 3 |
| 500 and above | 2 |

16. Calculate Median from the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Value less than : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
| Frequency : | 4 | 16 | 40 | 76 | 96 | 112 | 120 | 125 |

17. Calculate median marks from following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Marks more than or equal to: | 10 | 25 | 40 | 55 | 70 | 85 |
| No. of Students : | 100 | 94 | 74 | 30 | 4 | 1 |

18. Calculate median from the following data :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mid values : | 115 | 125 | 135 | 145 | 155 | 165 | 175 | 185 | 195 |
| Frequency : | 6 | 25 | 48 | 72 | 116 | 60 | 38 | 22 | 3 |

19. Find the missing frequency from the following distribution of sales of shops, given that median sale of shops is Rs. 24.00.

| | | | | | |
|---|---|---|---|---|---|
| Sale (In hundred of Rs. ) : | 0 – 10 | 10 – 20 | 20 – 30 | 30 – 40 | 40 – 50 |
| Number of Shops : | 5 | 25 | ? | 18 | 7 |

20. An incomplete distribution is given as follows :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Class : | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 | 60–70 | 70–80 |
| Frequency : | 12 | 30 | – | 65 | – | 25 | 18 |

Find the missing frequencies if N = 229 and the median value is 46. (P.U. B.C.A. April 2003)

21. The age distribution of the members of a certain children's club is as follow :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Age as on last Birthday (Years ) : | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Number of Children : | 5 | 9 | 18 | 35 | 42 | 32 | 15 | 7 | 3 |

There is a member 'A' such that there are twice as many members older than 'A' as there are members younger than A. Estimate his age in years. (P.U. B.C.A. Sept. 2005)

## ANSWERS

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1. | 95 | 2. | 52.5, 62 | 3. | 51.5, 52.5, 42.5, B | 4. | History |
| 5. | 3 | 6. | 20 | 7. | 7 | 8. | 24 |
| 9. | 18.71 | 10. | 65.6 | 11. | 23.67 | 12. | 38.5 |
| 13. | 17.55 | 14. | 692.56 | 15. | 209.91 | 16. | 36.25 |
| 17. | 48.18 | 18. | 153.79 | 19. | 25 | 20. | 33.5, 45.5 |
| 21. | 7.68 years | | | | | | |

## 7.9.4 OTHER POSITIONAL MEASURES

We know that median divides the given data into exactly two equal parts when the data is arranged in the ascending or in the descending order of magnitude. Apart from median there are other positional measures namely *Quartiles, Quintiles, Octiles, Deciles* and *Percentiles*. These are also known as *partition values* or *fractiles*. For the sake of clarity let the straight line AB represents the whole of the data. Now we shall explain each of these positional measures as follows :

**(i) Quartiles :** *Quartiles* divide the data into four equal parts. Thus, there will be three points of division and these will correspond to *lower, middle* and *upper quartiles*. Middle quartile is always equal to median. 25% of the items lie below the lower quartile ($Q_1$) and 25% of the items are above the upper quartile ($Q_3$). Middle 50% of the items lie between $Q_1$ and $Q_3$.

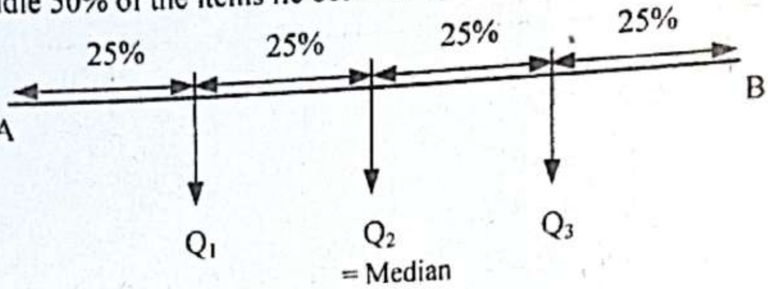



**Fig. 7.1**

**(ii) Quintiles :** *Quintiles* divide the data into five equal parts, each equivalent to 20 % of the items. The points of division are four which are represented by $Q_{u1}$, $Q_{u2}$, $Q_{u3}$, $Q_{u4}$. $Q_{u1}$ is at $\frac{1}{5}$ th of the total items, $Q_{u2}$ at $\frac{2}{5}$ th of the items and so on.

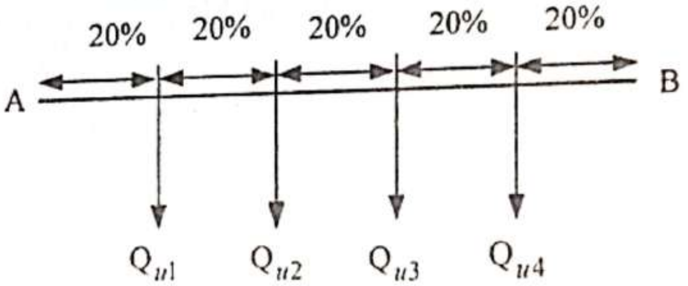



**Fig. 7.2**

**(iii) Octiles :** *Octiles* divide the data into eight equal parts. The points of division are seven which are represented by $O_1$ , $O_2$ , $O_3$ ......., $O_7$ respectively. $O_1$ is at $\frac{1}{8}$ th of the total items, $O_2$ is at $\frac{2}{8}$ th of the total items and so on. $O_4$ is always equal to median.

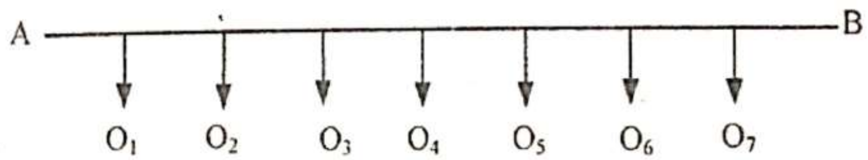



**Fig. 7.3**

**(iv) Deciles :** *Deciles* divide the data into ten equal parts. The points dividing the whole data will be nine and are represented by $D_1$, $D_2$ , ......., $D_9$. Each part is equivalent to 10% of the total number of items. $D_5$ is equal to median.

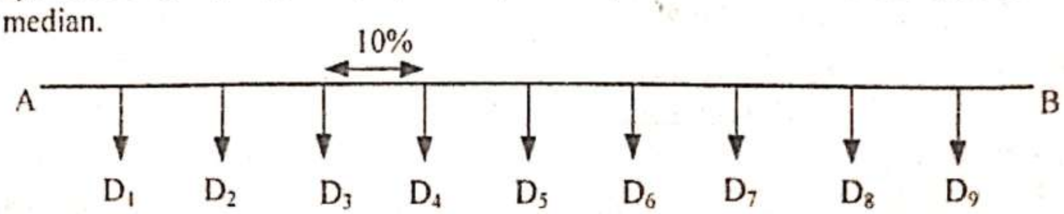



**Fig. 7.4**

(v) **Percentiles :** If data is divided into one hundred equal parts, the points of division are known as *percentiles*. These are represented by $P_1$, $P_2$, $P_3$.........$P_{99}$. Each part is equivalent to 1% of the total number of items. $P_{50}$ is equal to Median

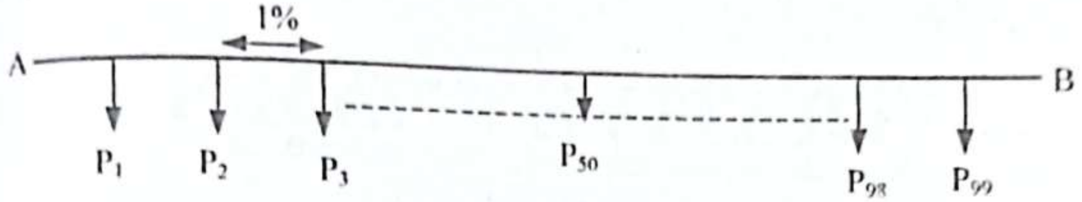



**Fig. 7.5**

The method of calculation for different positional measures is same as like median. The algorithm for these positional measures can be developed as that of median. The formulae to calculate different positional measures (or partition values) are given in following table :

**Table 7.4**

| Partitions Value | Individual Series | Discrete Series | Continuous Series | Formula to be used in continuous series |
|---|---|---|---|---|
| $Q_1$ | Value of $\frac{n+1}{4}$ th item | Value of $\frac{N+1}{4}$ th item | Value of $\frac{N}{4}$ th item | $Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times i$ |
| $Q_3$ | Value of $\frac{3(n+1)}{4}$ th item | Value of $\frac{3(N+1)}{4}$ th item | Value of $\frac{3N}{4}$ th item | $Q_3 = L + \frac{\frac{3N}{4} - c.f.}{f} \times i$ |
| $Q_{u1}$ | Value of $\frac{n+1}{5}$ th item | Value of $\frac{N+1}{5}$ th item | Value of $\frac{N}{5}$ th item | $Q_{u1} = L + \frac{\frac{N}{5} - c.f.}{f} \times i$ |
| $Q_{u4}$ | Value of $\frac{4(n+1)}{5}$ th item | Value of $\frac{4(N+1)}{5}$ th item | Value of $\frac{4N}{5}$ th item | $Q_{u4} = L + \frac{\frac{4N}{5} - c.f.}{f} \times i$ |
| $D_1$ | Value of $\frac{n+1}{10}$ th item | Value of $\frac{(N+1)}{10}$ th item | Value of $\frac{N}{10}$ th item | $D_1 = L + \frac{\frac{N}{10} - c.f.}{f} \times i$ |
| $D_6$ | Value of $\frac{6(n+1)}{10}$ th item | Value of $\frac{6(N+1)}{10}$ th item | Value of $\frac{6N}{10}$ th item | $D_6 = L + \frac{\frac{6N}{10} - c.f.}{f} \times i$ |
| $P_1$ | Value of $\frac{n+1}{100}$ th item | Value of $\frac{N+1}{100}$ th item | Value of $\frac{N}{100}$ th item | $P_1 = L + \frac{\frac{N}{100} - c.f.}{f} \times i$ |
| $P_{30}$ | Value of $\frac{30(n+1)}{100}$ th item | Value of $\frac{30(N+1)}{100}$ item | Value of $\frac{30N}{100}$ th item | $P_{30} = L + \frac{\frac{30N}{100} - c.f.}{f} \times i$ |

**Key notes for calculating median and other positional measures**

1. Always prefer to arrange data in ascending order.
2. Always Convert inclusive series into exclusive series to calculate median and other positional averages.

3. Conversion of unequal class intervals into equal class intervals is not necessary.
4. If data are in 'more than' or 'less than' cumulative frequency form, convert these into simple frequency form to calculate median and other positional measures.

# ILLUSTRATIVE EXAMPLES

**Example 1.** Calculate the median, two quartiles, 3rd quintile, 5th octile, 8th decile, 61st percentile of the following series of marks obtained by 10 candidates in an examination.

22, 26, 14, 30, 18, 11, 35, 41, 12, 32

**Sol.** Arrange the marks in the ascending order as under :

11, 12, 14, 18, 22, 26, 30, 32, 35, 41

$$\text{Median} = \text{Value of } \frac{n+1}{2} \text{ th item} = \text{Value of } \frac{10+1}{2} \text{ th item} = \text{Value of 5.5th item}$$

$\therefore$ Median is the A.M. of 5th and 6th item

*i.e.,* $$\text{Median} = \frac{22+26}{2} = \frac{48}{2} = 24$$

$$Q_1 = \text{Value of } \frac{n+1}{4} \text{ th item} = \text{Value of } \frac{10+1}{4} \text{ th item} = \text{Value of 2.75th item}$$

$$\therefore \quad Q_1 = \text{2nd item} + 0.75\ (\text{3rd item} - \text{2nd item}) = 12 + 0.75\ (14 - 12) = 12 + 0.75\ (2) = 13.5$$

$$Q_3 = \text{Value of } \frac{3(n+1)}{4} \text{ th item} = \text{Value of } \frac{3(10+1)}{4} \text{ th item} = \text{Value of 8.25th item}$$

$$\therefore \quad Q_3 = \text{8th item} + 0.25\ (\text{9th item} - \text{8th item}) = 32 + 0.25\ (35 - 32) = 32 + 0.25\ (3) = 32.75$$

$$Q_{u3} = \text{Value of } \frac{3(n+1)}{5} \text{ th item} = \text{Value of } \frac{3(10+1)}{5} \text{ th item} = \text{Value of 6.6th item}$$

$$\therefore \quad Q_{u3} = \text{6th item} + 0.6\ (\text{7th item} - \text{6th item}) = 26 + 0.6\ (30 - 26) = 26 + 0.6\ (4)$$
$$= 26 + 2.4 = 28.4$$

$$O_5 = \text{Value of } \frac{5(n+1)}{8} \text{ th item} = \text{Value of } \frac{5(10+1)}{8} \text{ th item} = \text{Value of 6.875th item}$$

$$\therefore \quad O_5 = \text{6th item} + 0.875\ (\text{7th item} - \text{6th item}) = 26 + 0.875\ (30 - 26)$$
$$= 26 + 0.875\ (4) = 26 + 3{\cdot}5 = 29.5$$

$$D_8 = \text{Value of } \frac{8(n+1)}{10} \text{ th item} = \text{Value of } \frac{8(10+1)}{10} \text{ th item} = \text{Value of 8.8th item}$$

$$\therefore \quad D_8 = \text{8th item} + 0.8\ (\text{9th item} - \text{8th item}) = 32 + 0.8\ (35 - 32) = 32 + 2.4 = 34.4$$

$$P_{61} = \text{Value of } \frac{61(n+1)}{100} \text{ th item} = \text{Value of } \frac{61(10+1)}{100} \text{ th item} = \text{Value of 6.71th item}$$

$$\therefore \quad P_{61} = \text{6th item} + 0.71\ (\text{7th item} - \text{6th item}) = 26 + 0.71\ (30 - 26) = 26 + 0.71\ (4) = 28.84$$

Example 2. Following are the monthly salaries (in hundred of Rs.) of 44 employees of a firm.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 130 | 125 | 121 | 100 | 96 | 128 | 76 | 78 | 103 | 100 | 55 |
| 62 | 151 | 142 | 71 | 74 | 66 | 142 | 136 | 85 | 81 | 105 |
| 145 | 110 | 107 | 80 | 76 | 144 | 98 | 101 | 122 | 123 | 150 |
| 118 | 95 | 60 | 132 | 134 | 114 | 116 | 124 | 95 | 90 | 96 |

Classify the data into class intervals of class width 10 each and calculate $Q_1$, $Q_3$, $D_7$, $P_{35}$, $P_{78}$.

Sol. First, we prepare the following table :

| Class Interval | Tally Bar | Frequency | c. f. |
|---|---|---|---|
| 50 – 60 | \| | 1 | 1 |
| 60 – 70 | \|\|\|\| | 4 | 5 |
| 70 – 80 | $\overline{\|\|\|\|}$ | 5 | 10 |
| 80 – 90 | \|\|\| | 3 | 13 |
| 90 – 100 | $\overline{\|\|\|\|}$ | 5 | 18 |
| 100 – 110 | $\overline{\|\|\|\|}$ \| | 6 | 24 |
| 110 – 120 | \|\|\|\| | 4 | 28 |
| 120 – 130 | $\overline{\|\|\|\|}$ \| | 6 | 34 |
| 130 – 140 | \|\|\|\| | 4 | 38 |
| 140 – 150 | \|\|\|\| | 4 | 42 |
| 150 – 160 | \|\| | 2 | 44 |
| | | N = 44 | |

Now $Q_1$ = Value of $\frac{N}{4}$th item = Value of $\frac{44}{4}$th item = Value of 11th item.

The *c.f.* next higher to 11 is 13

∴ $Q_1$ lies between 80 – 90 class interval

$$\therefore \quad Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times i = 80 + \frac{11-10}{3} \times 10 = 80 + \frac{10}{3} = 83.33$$

$Q_3$ = Value of $\frac{3N}{4}$th item = Value of $\frac{3 \times 44}{4}$th item = Value of 33rd item.

The *c.f.* next higher to 33 is 34

∴ $Q_3$ lies between 120 – 130 class interval

$$\therefore \quad Q_3 = L + \frac{\frac{3N}{4} - c.f.}{f} \times i = 120 + \frac{33-28}{6} \times 10 = 120 + \frac{50}{6} = 128.33$$

$$D_7 = \text{Value of } \frac{7N}{10} \text{th item} = \text{Value of } \frac{7 \times 44}{10} \text{th item} = \text{Value of } 30.8\text{th item.}$$

The *c.f.* next higher to 30.8 is 34

$\therefore$ $D_7$ lies between 120 – 130 class interval

$$\therefore \quad D_7 = L + \frac{\frac{7N}{10} - c.f.}{f} \times i = 120 + \frac{30.8 - 28}{6} \times 10 = 120 + 4.67 = 124.67$$

$$P_{35} = \text{Value of } \frac{35N}{100} \text{th item} = \text{Value of } \frac{35 \times 44}{100} \text{th item} = \text{Value of } 15.4\text{th item.}$$

The *c.f.* next higher to 15.4 is 18

$\therefore$ $P_{35}$ lies between 90 – 100 class interval

$$\therefore \quad P_{35} = L + \frac{\frac{35N}{100} - c.f.}{f} \times i = 90 + \frac{15.4 - 13}{5} \times 10 = 90 + 2\,(2.4) = 94.8$$

$$P_{78} = \text{Value of } \frac{78N}{100} \text{th item} = \text{Value of } \frac{78 \times 44}{100} \text{th item} = \text{Value of } 34.32\text{nd item.}$$

The *c.f.* next higher to 34.32 is 38

$\therefore$ $P_{78}$ lies between 130 – 140 class interval

$$\therefore \quad P_{78} = L + \frac{\frac{78N}{100} - c.f.}{f} \times i = 130 + \frac{(34.32 - 34)}{4} \times 10 = 130 + \frac{10}{4}\,(0.32) = 130.8$$

**Example 3.** Calculate the arithmetic mean, median and quartiles from the following distribution of 1[illegible] persons by age.

| | | | | | | |
|---|---|---|---|---|---|---|
| Age : | 15 – 19 | 20 – 24 | 25 – 29 | 30 – 34 | 35 – 39 | 40 – 44 |
| Number : | 4 | 20 | 38 | 24 | 10 | 4 |

**Sol.** First we convert the data in the exclusive form of class intervals (Although it is not necessary f[illegible] calculating A.M., but it is one of requisites for determining the positional averages).

| Age (X) | Mid Value ($m$) | A = 27 $dx = m - A$ | $d'x$ $c = 5$ | $f$ | $c.f.$ | $fd'x$ |
|---|---|---|---|---|---|---|
| 14.5 – 19.5 | 17 | – 10 | – 2 | 4 | 4 | – 8 |
| 19.5 – 24.5 | 22 | – 5 | – 1 | 20 | 24 | – 20 |
| 24.5 – 29.5 | 27 | 0 | 0 | 38 | 62 | 00 |
| 29.5 – 34.5 | 32 | 5 | 1 | 24 | 86 | 24 |
| 34.5 – 39.5 | 37 | 10 | 2 | 10 | 96 | 20 |
| 39.5 – 44.5 | 42 | 15 | 3 | 4 | 100 | 12 |
| | | | | N = 100 | | $\Sigma fd'x = 28$ |

Now Arithmetic mean $= A + \frac{\Sigma f dx'}{N} \times c = 27 + \frac{28}{100}(5) = 27 + 1.4 = 28.4$

Median = Value of $\frac{N}{2}$ th item = Value of $\frac{100}{2}$ th item = Value of 50th item

The *c.f.* next higher to 50 is 62

$\therefore$ Median lies in $24.5 - 29.5$

Now $$\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i = 24.5 + (50 - 24)\frac{5}{38} = 24.5 + 3.42 = 27.92$$

$Q_1$ = Value of $\frac{N}{4}$ th item = Value of $\frac{100}{4}$ th item = Value of 25th item

The *c.f.* next higher to 25 is 62

$\therefore$ $Q_1$ lies in class interval $24.5 - 29.5$

Now $$Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times i = 24.5 + (25 - 24)\frac{5}{38} = 24.5 + \frac{5}{38} = 24.5 + 0.13 = 24.63$$

$$Q_3 = \text{Value of } \frac{3N}{4} \text{ th item} = \text{Value of } \frac{300}{4} \text{ th item} = \text{Value of 75th item}$$

The *c.f.* next higher to 75 is 86

$\therefore$ $Q_3$ lies in class interval $29.5 - 34.5$

Now $$Q_3 = L + \frac{\frac{3N}{4} - c.f.}{f} \times i = 29.5 + (75 - 62)\frac{5}{24} = 29.5 + \frac{5 \times 13}{24} = 29.5 + 2.71 = 32.21$$

**Example 4.** Calculate the Median, Quartiles and Sixth decile from the data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Marks less than : | 80 | 70 | 60 | 50 | 40 | 30 | 20 | 10 |
| No. of Candidates : | 100 | 90 | 80 | 60 | 32 | 20 | 13 | 5 |

**Sol.** First we convert the given cumulative frequency distribution into simple frequency distribution and arrange the data in ascending order as follows :

| Marks | No. of Candidates ($f$) | *c.f.* |
|---|---|---|
| 0 – 10 | 5 | 5 |
| 10 – 20 | 13 – 5 = 8 | 13 |
| 20 – 30 | 20 – 13 = 7 | 20 |
| 30 – 40 | 32 – 20 = 12 | 32 |
| 40 – 50 | 60 – 32 = 28 | 60 |
| 50 – 60 | 80 – 60 = 20 | 80 |
| 60 – 70 | 90 – 80 = 10 | 90 |
| 70 – 80 | 100 – 90 = 10 | 100 |
| | N = 100 | |

Now, Median = Value of $\frac{N}{2}$ th item = Value of $\frac{100}{2}$ th item = Value of 50th item

The *c.f.* next higher to 50 is 60.

Therefore median lies in class interval 40 – 50

Now
$$\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i$$
$$= 40 + (50 - 32)\frac{10}{28} = 40 + \frac{10 \times 18}{28}$$
$$= 40 + 6.43 = 46.43$$

$$Q_1 = \text{Value of } \frac{N}{4} \text{ th item} = \text{Value of } \frac{100}{4} \text{ th item} = \text{Value of 25th item}$$

The *c.f.* next higher to 25 is 32.

Therefore $Q_1$ lies in class interval 30 – 40

Now
$$Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times i = 30 + (25 - 20)\frac{10}{12} = 30 + \frac{10 \times 5}{12}$$
$$= 30 + 4.17 = 34.17$$

$$Q_3 = \text{Value of } \frac{3N}{4} \text{ th item} = \text{Value of } \frac{300}{4} \text{ th item} = \text{Value of 75th item}$$

The *c.f.* next higher to 75 is 80.

Therefore $Q_3$ lies in class interval 50 – 60

Now
$$Q_3 = L + \frac{\frac{3N}{4} - c.f.}{f} \times i$$
$$= 50 + (75 - 60)\frac{10}{20} = 50 + \frac{15}{2} = 57.5$$

$$D_6 = \text{Value of } \frac{6N}{10} \text{ th item} = \frac{600}{10} = 60\text{th item}$$

Therefore $D_6$ lies in class interval 40 – 50

Now
$$D_6 = L + \frac{\frac{6N}{10} - c.f.}{f} \times i = 40 + (60 - 32)\frac{10}{28} = 50$$

**Example 5.** The first and the third quartiles of the following data are given to be 12.5 marks and 25 marks respectively :

| Marks : | 0 – 5 | 5 – 10 | 10 – 15 | 15 – 20 | 20 – 25 | 25 – 30 | 30 – 35 | 35 – 40 | Total |
|---|---|---|---|---|---|---|---|---|---|
| Frequency : | 4 | 8 | ? | 19 | ? | 10 | 5 | ? | 72 |

Find the missing frequencies ?

Let the missing frequencies be $x$, $y$ and $z$. Now we prepare the following table :

| Marks (X) | Frequency ($f$) | c.f. |
|---|---|---|
| 0 – 5 | 4 | 4 |
| 5 – 10 | 8 | 12 |
| 10 – 15 | $x$ | $12+x$ |
| 15 – 20 | 19 | $31+x$ |
| 20 – 25 | $y$ | $31+x+y$ |
| 25 – 30 | 10 | $41+x+y$ |
| 30 – 35 | 5 | $46+x+y$ |
| 35 – 40 | $z$ | $46+x+y+z$ |
| | $N = x+y+z+46$ | |

Since $N = 72$ so $x+y+z+46 = 72$

$\Rightarrow$ $x+y+z = 26$ ...(I)

Also $Q_1$ = Value of $\frac{N}{4}$th item = Value of $\frac{72}{4}$th item = Value of 18th item.

Since $Q_1 = 12.5$ which lies in the class interval 10 – 15

$$\therefore \quad 12.5 = 10 + (18-12)\frac{5}{x}$$

$$\Rightarrow \quad \frac{30}{x} = 2.5 \qquad \Rightarrow x = \frac{30}{2.5} = 12 \qquad ...(II)$$

Further $Q_3$ = Value of $\frac{3N}{4}$ th item = Value of $\frac{3\times 72}{4}$ th item = Value of 54th item

Because of the reason stated above, $Q_3$ too cannot be located from *c.f.* column. But $Q_3$ is given to be 25, which lies between 25 – 30.

Since $Q_3 = 25$, which lies in class-interval 20–25

$$\therefore \quad 25 = 20 + \frac{(54-(31+x))}{y}\times 5 \qquad \left[\text{Using } Q_3 = \frac{L+\frac{3N}{4}-c.f.}{f}\times i\right]$$

$$\Rightarrow \quad y = 54 - 31 - x$$

$$\Rightarrow \quad y = 23 - 12 \qquad \text{[Using (II)]}$$

$$\Rightarrow \quad y = 11$$

From (I), (II) and (III), we have

$12 + 11 + z = 26 \quad \Rightarrow \quad z = 3$ Thus $x = 12$, $y = 11$, $z = 3$

Now, Median = Value of $\frac{N}{2}$ th item = Value of $\frac{100}{2}$ th item = Value of 50th item

The $c.f.$ next higher to 50 is 60.

Therefore median lies in class interval 40 – 50

Now $$\text{Median} = L + \frac{\frac{N}{2} - c.f.}{f} \times i$$

$$= 40 + (50 - 32)\frac{10}{28} = 40 + \frac{10 \times 18}{28}$$

$$= 40 + 6.43 = 46.43$$

$$Q_1 = \text{Value of } \frac{N}{4} \text{ th item} = \text{Value of } \frac{100}{4} \text{ th item} = \text{Value of 25th item}$$

The $c.f.$ next higher to 25 is 32.

Therefore $Q_1$ lies in class interval 30 – 40

Now $$Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times i = 30 + (25 - 20)\frac{10}{12} = 30 + \frac{10 \times 5}{12}$$

$$= 30 + 4.17 = 34.17$$

$$Q_3 = \text{Value of } \frac{3N}{4} \text{ th item} = \text{Value of } \frac{300}{4} \text{ th item} = \text{Value of 75th item}$$

The $c.f.$ next higher to 75 is 80.

Therefore $Q_3$ lies in class interval 50 – 60

Now $$Q_3 = L + \frac{\frac{3N}{4} - c.f.}{f} \times i$$

$$= 50 + (75 - 60)\frac{10}{20} = 50 + \frac{15}{2} = 57.5$$

$$D_6 = \text{Value of } \frac{6N}{10} \text{ th item} = \frac{600}{10} = 60\text{th item}$$

Therefore $D_6$ lies in class interval 40 – 50

Now $$D_6 = L + \frac{\frac{6N}{10} - c.f.}{f} \times i = 40 + (60 - 32)\frac{10}{28} = 50$$

**Example 5.** The first and the third quartiles of the following data are given to be 12.5 marks and 25 marks respectively :

| Marks : | 0 – 5 | 5 – 10 | 10 – 15 | 15 – 20 | 20 – 25 | 25 – 30 | 30 – 35 | 35 – 40 | Total |
|---|---|---|---|---|---|---|---|---|---|
| Frequency : | 4 | 8 | ? | 19 | ? | 10 | 5 | ? | 72 |

Find the missing frequencies ?

Let the missing frequencies be $x$, $y$ and $z$. Now we prepare the following table :

| Marks (X) | Frequency ($f$) | c. f. |
|---|---|---|
| 0 – 5 | 4 | 4 |
| 5 – 10 | 8 | 12 |
| 10 – 15 | $x$ | $12 + x$ |
| 15 – 20 | 19 | $31 + x$ |
| 20 – 25 | $y$ | $31 + x + y$ |
| 25 – 30 | 10 | $41 + x + y$ |
| 30 – 35 | 5 | $46 + x + y$ |
| 35 – 40 | $z$ | $46 + x + y + z$ |
| | $N = x + y + z + 46$ | |

Since N = 72 so $x + y + z + 46 = 72$

$$x + y + z = 26 \qquad \text{...(I)}$$

$\Rightarrow$

$$Q_1 = \text{Value of } \frac{N}{4}\text{th item} = \text{Value of } \frac{72}{4}\text{th item} = \text{Value of 18th item.}$$

Also

Since $Q_1 = 12.5$ which lies in the class interval 10 – 15

$$12.5 = 10 + (18 - 12)\,\frac{5}{x}$$

$\therefore$

$$\frac{30}{x} = 2.5 \qquad \Rightarrow \quad x = \frac{30}{2.5} = 12 \qquad \text{...(II)}$$

$\Rightarrow$

Further $Q_3 = \text{Value of } \frac{3N}{4}\text{th item} = \text{Value of } \frac{3 \times 72}{4}\text{th item} = \text{Value of 54th item}$

Because of the reason stated above, $Q_3$ too cannot be located from *c.f.* column. But $Q_3$ is given to be 25, which lies between 25 – 30.

Since $Q_3 = 25$, which lies in class-interval 20-25

$\therefore$

$$25 = 20 + \frac{(54 - (31 + x))}{y} \times 5 \qquad \left[\text{Using } Q_3 = \frac{L + \frac{3N}{4} - c.f.}{f} \times i\right]$$

$$\Rightarrow \quad y = 54 - 31 - x$$

$$\Rightarrow \quad y = 23 - 12 \qquad \text{[Using (II)]}$$

$$\Rightarrow \quad y = 11$$

From (I), (II) and (III), we have

$12 + 11 + z = 26 \quad \Rightarrow \quad z = 3$ Thus $x = 12, \quad y = 11, \quad z = 3$

# EXERCISE 7.10

1. From the following data find $Q_1$, $Q_3$, $D_5$, $P_{25}$, $P_{67}$.
   Marks : 37, 39, 45, 53, 41, 57, 43, 47, 51, 49, 55
2. Find out $Q_3$, $D_4$ and $P_{40}$ from the following weights (in Kg.) :
   19, 27, 24, 39, 57, 44, 56, 50, 59, 67, 62, 42, 47, 60, 26, 34, 57, 51, 59, 45
3. From the data given below compute the value of $Q_1$, $Q_3$, $D_7$ and $P_{46}$

| Size of Shoes : | 4.5 | 5.0 | 5.5 | 6.0 | 6.5 | 7.0 | 7.5 | 8.0 | 8.5 | 9.0 | 9.5 | 10.0 | 10.5 | 11.0 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Frequency : | 1 | 2 | 4 | 5 | 15 | 30 | 60 | 95 | 82 | 75 | 44 | 25 | 15 | 4 |

4. From the following data find out Median (M), First Quartile ($Q_1$), Third Quartile ($Q_3$) and Percentile ($P_{60}$).

| Daily Wages (Rs.) : | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No. of Workers : | 3 | 6 | 10 | 15 | 24 | 42 | 75 | 90 | 79 | 55 | 36 | 26 | 19 | 13 |

5. Calculate median, first quartile and 85th percentile of the following data :

| Income ('00 Rs.) : | 0 – 5 | 5 – 10 | 10 – 15 | 15 – 20 | 20 – 25 | 25 – 30 | 30 – 35 | 35 – 4 |
|---|---|---|---|---|---|---|---|---|
| No. of Families : | 75 | 250 | 350 | 192 | 68 | 35 | 24 | 6 |

6. Find the value $Q_1$ and $Q_3$ from following data :

| Mid Value : | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 |
|---|---|---|---|---|---|---|---|---|
| Frequency : | 12 | 8 | 13 | 11 | 9 | 6 | 3 | 10 |

7. Find out $D_2$ and $P_{25}$ from the following :

| Wages below (Rs.) : | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 |
|---|---|---|---|---|---|---|---|---|
| Workers : | 5 | 12 | 20 | 30 | 39 | 50 | 56 | 70 |

8. Calculate Median, 3rd decile and 60th percentile for the data :

| Marks Above : | 0 | 10 | 20 | 30 | 40 | 50 |
|---|---|---|---|---|---|---|
| No. of Students : | 150 | 140 | 100 | 90 | 80 | 70 |

9. For a certain group of saree-weavers of varanasi, the median and quartile earnings per week are Rs. 44.3, Rs. 43.0 and 45.9 respectively. The earnings for the group range between Rs. 40 and Rs. 50. Ten percent of the group earn under Rs. 42, 13% earn Rs. 47 and over and 6% Rs. 48 and over. Put these data in the form of frequency distribution and obtain the value of the mean wage.

## ANSWERS

1. $Q_1 = 41$, $Q_3 = 53$, $D_5 = 47$, $P_{25} = 41$, $P_{67} = 51.08$
2. $Q_3 = 58.5$, $D_4 = 44.4$, $P_{40} = 44.4$
3. $Q_1 = 7.5$, $Q_3 = 9.0$, $D_7 = 9.0$, $P_{46} = 8.0$
4. $M = 16$, $Q_1 = 15$, $Q_3 = 18$, $P_{60} = 17$
5. $M = 12.5$, $Q_1 = 8.5$, $P_{85} = 19.56$
6. $Q_1 = 22.50$, $Q_3 = 56.67$
7. $D_2 = 21.25$, $P_{25} = 23.44$
8. $M = 45$, $D_3 = 18.75$, $P_{60} = 51.43$
9. Rs. 44.50

## CALCULATION OF MEDIAN AND ALLIED POSITIONAL AVERAGES GRAPHICALLY

Median, Quartiles, Deciles can also be determined graphically. For calculation of these measures by graph we have to make use of graph paper and follow all instructions and precautions which we generally use while drawing a graph. *Ogive* is an important graph for the calculation of median and other fractiles.

The procedure for sketching ogive curve is as follows :

**Less than ogive**

(*i*) Represented the given distribution in less than cumulative frequency distribution form.

(*ii*) Along the *x*-axis, we take the values of the variable (in case of discrete series) or mid values of the class intervals (in case of continuous series). Along the *y*-axis, take cumulative frequencies.

(*iii*) Plot the cumulative frequency corresponding to value of variable (in case of discrete series) or upper limit of the class interval (in case of continuous series).

(*iv*) Draw a smooth curve by joining the points obtained in step (*iii*). The curve so formed is called less than ogive.

**More than ogive**

The more than ogive can be drawn in the same way as the less than ogive curve except that we represent the given distribution in more than cumulative frequency distribution form.

Notes. (*a*) To get the median from less than ogive, we draw line parallel to the *x*-axis from the point corresponding to cumulative frequency $\frac{N+1}{2}$ (in case of discrete series) or $\frac{N}{2}$ (in case of continuous series) on *y*-axis. The value of *x* corresponding to the point where this line intersects the ogive gives the median.

(*b*) Another method to obtain median is that we draw less than ogive and more than ogive. The value of *x* corresponding to the point of intersection of these two ogives gives the value of median.

(*c*) To find any other positional measure say $Q_3$ from less than ogives we draw a straight line parallel to *x*-axis from the point corresponding to cumulative frequency $\frac{3(N+1)}{4}$ (in case of discrete series) or $\frac{3N}{4}$ (in case of continuous series) on *y*-axis. The value of *x* corresponding to the point where this line intersects its the ogive gives $Q_3$.

Now, we shall illustrate this method with the help of following example :

Consider the following data relating to monthly wages in rupees and number of workers working in a factory :

| Monthly Wages Rs. : | 50–55 | 55–60 | 60–65 | 65–70 | 70–75 | 75–80 | 80–100 |
|---|---|---|---|---|---|---|---|
| Number of Workers : | 6 | 10 | 22 | 30 | 16 | 12 | 15 |

To draw less than ogive, we prepare the following table :

| Wages Less than (Rs.) | No. of Workers |
|---|---|
| 55 | 6 |
| 60 | 16 ( = 10 + 6 ) |
| 65 | 38 ( = 16 + 22 ) |
| 70 | 68 ( = 38 + 30 ) |
| 75 | 84 ( = 68 + 16 ) |
| 80 | 96 ( = 84 + 12 ) |
| 100 | 111 ( = 96 + 15) |

On the graph paper, wages (in Rs.) will be taken along X-axis and number of workers along Y-axis, according to some suitable scale. We plot the points corresponding to the above table and join the points. The sketch so obtained is termed as less than ogive (See Fig. 7.6).

Now, from this less than ogive, we can obtain positional averages.

For example,

$$\text{Median} = \text{Value of } \frac{N}{2}\text{th item}$$

$$= \text{Value of 55.5th item.}$$

We draw a line parallel to $x$-axis corresponding to $y = 55.5$. The value of $x$ corresponding to the point where this line intersects the curve gives median. Hence, median = 68.

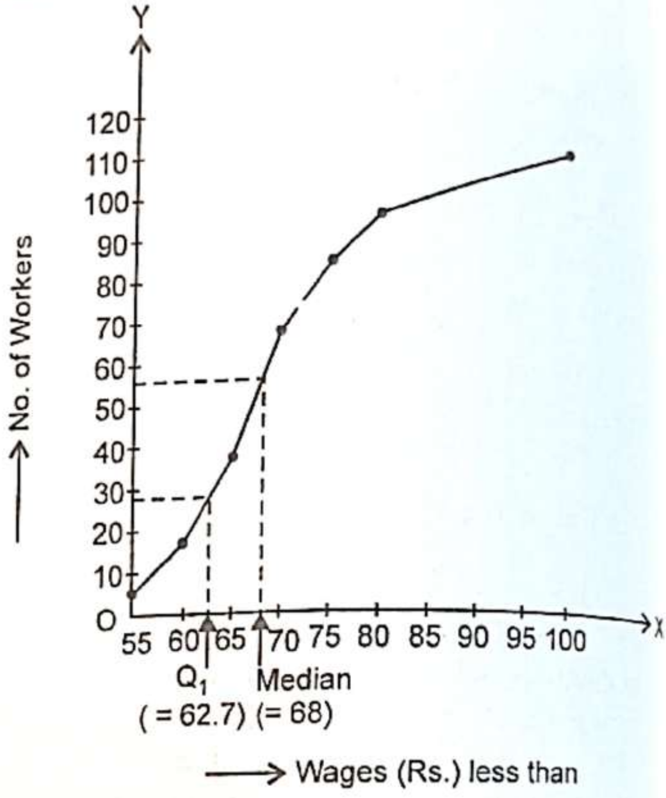



**Fig. 7.6 (less than ogive)**

$$\text{Similarly Lower Quartile } (Q_1) = \text{Value of } \frac{N}{4}\text{th item}$$

$$= \text{Value of } \frac{111}{4}\text{th item} = \text{Value of 27.75th item}$$

The value of $x$ corresponding to the point where the line which is drawn parallel to $x$-axis at $y = 27.75$ intersects the curve gives first quartile. Here $Q_1 = 62.7$.

Similarly, we can obtain other positional averages graphically.

In order to verify the above calculated values, we prepare the following table :

| Monthly Wages | No. of Workers | c.f. |
|---|---|---|
| 50 – 55 | 6 | 6 |
| 55 – 60 | 10 | 16 |
| 60 – 65 | 22 | 38 |
| 65 – 70 | 30 | 68 |
| 70 – 75 | 16 | 84 |
| 75 – 80 | 12 | 96 |
| 80 – 100 | 15 | 111 |
| | N = 111 | |

(*i*) Median is the value of $\frac{N}{2}$ th item = Value of 55.5th item

The *c.f.* next higher to 55.5 is 68

$\therefore$ Median lies in 65 – 70.

$$\text{Now Median } = L + \frac{\frac{N}{2} - c.f.}{f} \times i = 65 + (55.5 - 38)\,\frac{5}{30} = 65 + 2.9 = 68 \text{ (approx.)}$$

(*ii*) Lower Quartile ($Q_1$) = Value of $\frac{N}{4}$ th item = Value of 27.75th item.

The *c.f.* next higher to 27.75 is 38

$\therefore$ $Q_1$ lies in 60 – 65.

$$Q_1 = L + \frac{\frac{N}{4} - c.f.}{f} \times i = 60 + (27.75 - 16)\,\frac{5}{22} = 60 + 2.67 = 62.7 \text{ (approx)}$$

Note As discussed earlier, we can also obtained median by converting the data in the 'less than' and 'more than' form and plot both the curves on the same graph paper and on the same scale. The point where the two curves intersect give us the value of median. But it should be noted that we can calculate only median by this method. For this, we prepare the following table :

| Wages Less then ( Rs. ) | No. of Workers | Wages More than ( Rs. ) | No. of Workers |
|---|---|---|---|
| 55 | 6 | 50 | 111 |
| 60 | 16 ( = 10 + 6 ) | 55 | 105 ( = 111 – 6 ) |
| 65 | 38 ( = 16 + 22 ) | 60 | 95 ( = 105 – 10 ) |
| 70 | 68 ( = 38 + 30 ) | 65 | 73 ( = 95 – 22 ) |
| 75 | 84 ( = 68 + 16 ) | 70 | 43 ( = 73 – 30 ) |
| 80 | 96 ( = 84 + 12 ) | 75 | 27 ( = 43 – 16 ) |
| 100 | 111 ( = 96 + 15) | 80 | 15 ( = 27 – 12 ) |

On the graph paper take the wages in both the cases along X-axis and number of workers along Y-axis according to some suitable scale. The less than ogive and more than ogive are drawn in fig. 7.7.

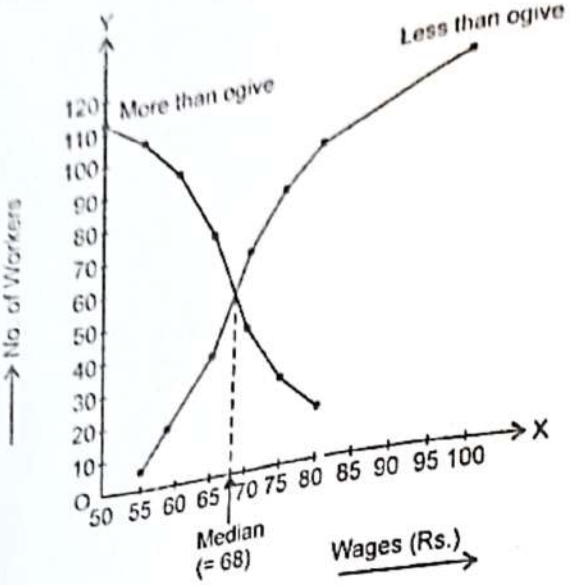



**Fig. 7.7**

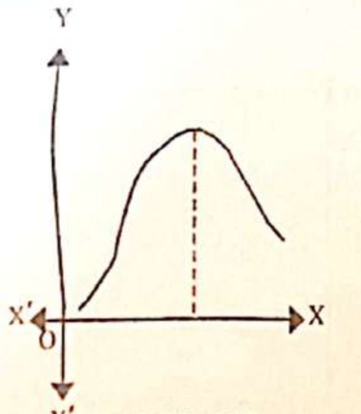
From fig. 7.7, it is clear that median = 68.

## 7.10 MODE

*Mode* is another positional measure of central tendency. Mode is that value which occurs most frequently in a statistical distribution. *i.e.* Mode is that value of the variable around which there is highest concentration of values. Mode is generally denoted by Z.

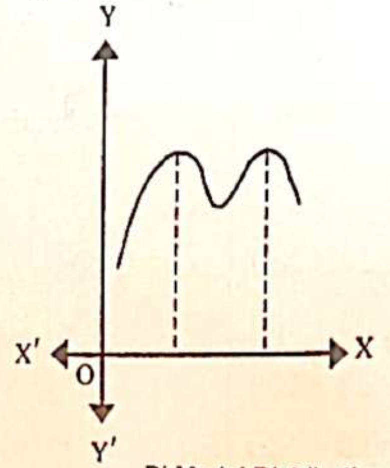
According to **Croxton and Cowden,** ***" The mode of a distribution is the value at the point around which the items tend to be most heavily concentrated. It may be regarded as the most typical of a series of values".*** For example, a particular size of pair of shoes or the size of the shirts which has the largest demand in the market will determine the value of the mode.

There will be no mode if each item occurs equal number of times in a distribution. When there is only one mode, the distribution is known as *uni-modal.* If in a series two values occur equal number of times but with largest frequencies as compared to other items in the distribution, it is known as a *bi-modal* series. We can extend our definition to *tri-modal* or *multi–modal* distribution in a similar manner. Mode is shown diagrammatically by frequency curves (fig. 7.8, 7.9 and 7.10) as follows :



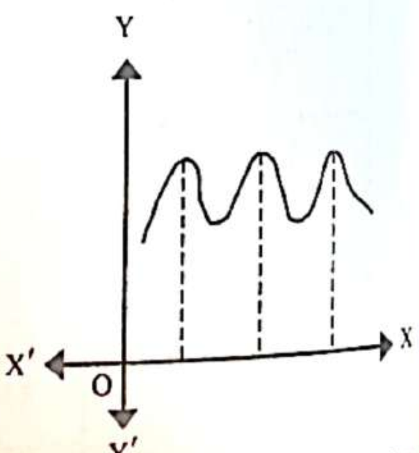

Uni-Modal Distribution

**Fig. 7.8**




Bi-Modal Distribution

**Fig. 7.9**




Tri-Modal Distribution

**Fig. 7.10**

### Merits and Demerits of Mode

**Merits :** The following are the merits of mode :

(i) It is simplest to measure. We can guess its value by just having a glance over the data.

(ii) It has importance in the scientific and commercial fields in particular. It is highly useful where we are concerned with the production of items depending upon its size such as shoes, shirts etc.

(iii) Since the number of items corresponding to the value of mode in a series is the highest therefore, while selecting some items at random, the item with mode as its value has the greatest chance for selection in the sample.

(iv) Mode can be calculated in series which are classified on qualitative basis.

(v) Like median it is little affected by the extreme values in the data.

**Demerits :** The following are the demerits of mode :

(i) Mode is not rigidly defined.

(ii) If the number of items is too large only then mode can be said to be the representative of the series. Otherwise the value is not much dependable.

(iii) Since the mode is not based on any algebraic calculations, therefore it is not capable for any further statistical treatment.

## 7.10.1 CALCULATION OF MODE IN CASE OF INDIVIDUAL SERIES

The following steps are involved in the calculated of mode in case of individual series :

**Step 1 :** Arrange the items in the ascending or descending order of magnitude.

**Step 2 :** The value with the highest repetition is known as Mode.

## 7.10.2 CALCULATION OF MODE IN CASE OF DISCRETE SERIES

The following two methods are used for determining mode in case of discrete series :

***(i)* Inspection Method :** According to this method we simply find the item with highest frequency and the value of the variable corresponding to the highest frequency is the mode.

***(ii)* Grouping Method :** When the distribution is such that the maximum frequency is repeated in the distribution or maximum frequency occur either in the beginning or at the end of the distribution or the frequencies are not regular in value *i.e.,* there is rise and fall in the values of the frequencies, we fail to locate the value of the mode by inspection properly. In this case grouping method is used to find mode. The procedure for grouping method is as follows :

**Formation of Grouping Table :** A grouping table has six columns. Following steps are followed to form grouping table :

1. Write given frequencies in Ist column.
2. Write total of frequencies taking two frequencies at a time in IInd column.
3. Ignore first frequency and take total of frequencies taking two frequencies at a time and write in IIIrd column.
4. Write total of frequencies taking three frequencies at a time in IVth column.
5. Ignore first frequency and take total of frequencies taking three frequencies at a time and write in Vth column.

6. Ignore first two frequencies and take total of frequencies taking three frequencies at a time and write in VIth column.

7. Put a circle around the maximum values of all the columns.

**Formation of Analysis Table :** In the analysis table, six rows (I – VI) are drawn corresponding to each column in the grouping table. In this table, columns are made for values of the variable whose frequencies account for giving maximum totals in the columns of the grouping table. In this table, tick or cross marks are given to the value of the variable as often as their frequencies are added to make the total maximum in the columns of the grouping table. The value of the variable which gets the maximum marks is declared to be the mode of the distribution.

## 7.10.3 CALCULATION OF MODE IN CASE OF CONTINUOUS SERIES

Like discrete series, the group having maximum frequency is first located either by inspection or by grouping method, depending upon the nature of the problem. The foliowing formula is used to determine mode :

$$\text{Mode} = L + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i \quad ...(7.41)$$

where L = lower limit of the modal class

$f_1$ = frequency of the modal group

$f_0$ = frequency preceding the modal class

$f_2$ = frequency following the modal class

$i$ = width of class interval of the modal class

Some times the above formula gives us misleading results. Particularly, the mode cannot be calculated in special circumstances when $2f_1 - f_0 - f_2 = 0$. Then we use the following formula :

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i \quad ...(7.42)$$

where $\Delta_1 = |f_1 - f_0|$ *i.e.*, $f_1 - f_0$ with +ve sign only

and $\Delta_2 = |f_1 - f_2|$ *i.e.*, $f_1 - f_2$ with +ve sign only

We may use the upper limit of the modal group to find the mode by the following formula :

$$\text{Mode} = U - \frac{f_1 - f_2}{2f_1 - f_0 - f_2} \times i \quad ...(7.43)$$

or

$$\text{Mode} = U - \frac{\Delta_2}{\Delta_1 + \Delta_2} \times i \quad ...(7.44)$$

where U = upper limit of modal class.

**Key notes for calculating mode**

1. Always prefer to arrange data in ascending order to find Mode.
2. Always convert inclusive series into exclusive series to find mode.
3. Always convert unequal class intervals into equal class intervals to find Mode.
4. If data is in 'More than' Or 'Less than' cumulative frequency form then convert it into simple frequency distribution.

## 7.10.4 EMPIRICAL RELATION BETWEEN MEAN, MEDIAN AND MODE

In case of symmetrical distribution, the values mean, median and mode are same. But if the distribution is moderately asymmetrical then the values of mean, median and mode are not same. Prof. Karl Pearson gave the following important empirical relationship between mean, median and mode :

$$\text{Mode} = \text{Mean} - 3\,(\text{Mean} - \text{Median}) \qquad ...(7.45)$$

or

$$\text{Mode} = 3\ \text{Median} - 2\ \text{Mean} \qquad ...(7.46)$$

Notes (a) Rewriting formula (7.46) as

$$\text{Mean} - \text{Median} = \frac{1}{3}\,(\text{Mean} - \text{Mode}) \qquad ...(7.47)$$

From formula (7.47), it is clear that the difference between mean and median is one third of difference between mean and mode. *i.e.* mean is closer to median that mode.

(b) As discussed earlier, there may arise situations in which mode can not be evaluated merely by inspection. In these cases, we apply grouping method. But in some cases, the grouping method may give two values as mode. In these cases mode is said to be *ill-defined*. An estimate value of mode in such cases can be calculated by the empirical relation (7.46) between mean, median and mode.

### CHECKPOINTS

1. What do you understand by Mode ?
2. What are the merits and demerits of Mode ? (P.U. B.C.A. April 2008)
3. Explain the purpose or significance of Mode. (G.N.D.U. B.Sc. C.Sc. April 2005, 2006, 2007)
4. What is the difference between Mean, Median and Mode ? How these are related ? (G.N.D.U. B.C.A. April 2008)
5. What is the relation among Mean, Median and Mode ? (Pbi. U. B.C.A. Sept. 2006)

# ILLUSTRATIVE EXAMPLES

**Example 1.** Age of 15 students is given as follows. Calculate modal age.

22, 24, 17, 18, 19, 18, 21, 20, 21, 20, 23, 22, 22, 22, 22

**Sol.** Arrange the given values in the ascending order of magnitude as follows :

17, 18, 18, 19, 20, 20, 21, 21, 22, 22, 22, 22, 22, 23, 24

Numbers 22 occur maximum times.

$\therefore$ modal age = 22

**Example 2.** Calculate modal marks from the data given below :

Marks : 25, 32, 59, 37, 17, 22, 26, 28, 33, 40, 45, 58, 67.

**Sol.** Arrange the given values in the ascending order of magnitude as follows :

17, 22, 25, 26, 28, 32, 33, 37, 40, 45, 58, 59, 67

Since each value occurs only once therefore mode is ill-defined. So we have to find Mode by the following empirical relation :

$$\boxed{\text{Mode} = 3\,(\text{Median}) - 2\,(\text{Mean})}$$

Now Mean of the given numbers $= \dfrac{25+32+59+37+17+22+26+28+33+40+45+58+\ldots}{13}$

$$= \frac{489}{13} = 37.615$$

Number of items $n = 13$

$\therefore$ Median = value of $\dfrac{n+1}{2}$ th item = value of $\dfrac{13+1}{2}$ th item = value of 7th item

$\therefore$ Median = 33

$\therefore$ Mode by the empirical relation $= 3\,(33) - 2\,(37.615) = 99 - 75.23 = 23.77$

**Example 3.** The following is the distribution of wages per thousand employees in a certain factory

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Daily Wages in Rs. : | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 22 | 24 | Total |
| No. of Employees : | 3 | 13 | 43 | 102 | 175 | 220 | 204 | 139 | 69 | 25 | 6 | 1 | 1000 |

Calculate the modal wage.

**Sol.** Because highest number of employers are with daily wages of Rs. 12.

$\therefore$ Mode = 12

**Example 4.** The frequency distribution of weights of 60 students of a class is given below :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Weight (in kg.) : | 40 | 41 | 42 | 43 | 44 | 45 | 46 |
| Frequency : | 8 | 10 | 12 | 15 | 8 | 5 | 2 |

(P.U. B.C.A. Sept. 200...)

Find the mode of the weights.

**Sol.** Let us denote the weights of the students by X and frequency by $f$. Now, we prepare the grouping table as follows :

| X | ($f$) Col. I | Col. II | Col. III | Col. IV | Col. V | Col. VI |
|---|---|---|---|---|---|---|
| 40 | 8 | 18 | | 30 | | |
| 41 | 10 | | 22 | | 37 | |
| 42 | 12 | 27 | | | | 35 |
| 43 | 15 | | 23 | 28 | | |
| 44 | 8 | 13 | | | 15 | |
| 45 | 5 | | 7 | | | |
| 46 | 2 | | | | | |

Now we prepare analysis table as follows :

| Col. No. \ X → | 40 | 41 | 42 | 43 | 44 | 45 | 46 |
|---|---|---|---|---|---|---|---|
| I | | | √ | √ | | | |
| II | | | | √ | | | |
| III | √ | √ | √ | √ | √ | | |
| IV | | √ | √ | | | | |
| V | | | √ | √ | | | |
| VI | | | | √ | √ | | |
| Total | 1 | 2 | 4 | 5 | 2 | – | – |

From analysis table, it is clear that mode of the weights is 43.

Example 5. Find the value of mode for the following distributions :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| X : | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| f : | 15 | 18 | 12 | 30 | 27 | 40 | 20 | 20 | 12 |

(G.N.D.U. B.C.A. April 2005, 2009)

Sol. First, we construct grouping table as follows :

| X | (f) Col. I | Col. II | Col. III | Col. IV | Col. V | Col. VI |
|---|---|---|---|---|---|---|
| 4 | 15 | 33 | | 45 | | |
| 5 | 18 | | 30 | | 60 | |
| 6 | 12 | 42 | | | | 69 |
| 7 | 30 | | 57 | 97 | | |
| 8 | 27 | 67 | | | 87 | |
| 9 | 40 | | 60 | | | 80 |
| 10 | 20 | 40 | | 52 | | |
| 11 | 20 | | 32 | | | |
| 12 | 12 | | | | | |

250
75
175

Now we prepare analysis table as follows :

| X → / Col. No. | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|
| I | | | | | | √ | | | |
| II | | | | | √ | √ | | | |
| III | | | | | | √ | √ | | |
| IV | | | | √ | √ | √ | | | |
| V | | | | | √ | √ | √ | | |
| VI | | | | | | √ | √ | √ | |
| Total | – | – | – | 1 | 3 | 6 | 3 | 1 | – |

From analysis table, it is clear that mode of the given frequency distribution is 9.

**Example 6.** Calculate the modal value from the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| X : | 0 – 5 | 5 – 10 | 10 – 15 | 15 – 20 | 20 – 25 | 25 – 30 | 30 – 35 | 35 – 40 |
| $f$ : | 29 | 195 | 241 | 117 | 52 | 10 | 6 | 3 |

(G.N.D.U. B.Sc. C.Sc. April 2003)

**Sol.** By inspection, it is clear that mode lies in the interval 10–15.

Now $$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where $$\Delta_1 = |f_1 - f_0|, \quad \Delta_2 = |f_1 - f_2|$$

Here $$f_0 = 195, \; f_1 = 241, \; f_2 = 117, \; L = 10 \text{ and } i = 5$$

$$\therefore \quad \Delta_1 = |241 - 195| = 46, \; \Delta_2 = |241 - 117| = 124$$

So $$\text{Mode} = 10 + \frac{46}{46 + 124} \times 5$$

$$= 11.35$$

**Example 7.** Calculate mode from the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Weights (kg) : | 40 – 44 | 45 – 49 | 50 – 54 | 55 – 59 | 60 – 64 | 65 – 69 | 70 – 74 | 75 – 79 |
| No. of Students : | 3 | 6 | 13 | 18 | 15 | 7 | 3 | 2 |

Sol. First, we change the given class interval from inclusive form to exclusive form and prepare the grouping table as follows :

| Weights (Kg.) (X) | No. of Students Col. I | Col. II | Col. III | Col. IV | Col. V | Col. VI |
|---|---|---|---|---|---|---|
| 39.5 – 44.5 | 3 | 9 | | | | |
| 44.5 – 49.5 | 6 | | 19 | 22 | | |
| 49.5 – 54.5 | 13 | 31 | | | 37 | |
| 54.5 – 59.5 | 18 | | 33 | | | 46 |
| 59.5 – 64.5 | 15 | 22 | | 40 | | |
| 64.5 – 69.5 | 7 | | 10 | | 25 | |
| 69.5 – 74.5 | 3 | 5 | | | | 12 |
| 74.5 – 79.5 | 2 | | | | | |

Now we prepare analysis table as follows :

| X → / Col. No. ↓ | 39.5 – 44.5 | 44.5 – 49.5 | 49.5 – 54.5 | 54.5 – 59.5 | 59.5 – 64.5 | 64.5 – 69.5 | 69.5–74.5 |
|---|---|---|---|---|---|---|---|
| I | | | | √ | | | |
| II | | | √ | √ | | | |
| III | | | | √ | √ | | |
| IV | | | | √ | √ | √ | |
| V | | √ | √ | √ | | | |
| VI | | | √ | √ | √ | | |
| Total | | 1 | 3 | 6 | 3 | 1 | |

From the analysis table it is clear that 54.5 – 59.5 is the modal class interval

Now
$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where
$$\Delta_1 = |f_1 - f_0|, \quad \Delta_2 = |f_1 - f_2|$$

Here
$$L = 54.5, \; f_0 = 13, \; f_1 = 18, \; f_2 = 15$$

$\therefore$ $$\Delta_1 = |18 - 13| = 5,\ \Delta_2 = |18 - 15| = 3$$

$\therefore$ $$\text{Mode} = 54.5 + \frac{5}{5+3} \times 5 = 54.5 + \frac{25}{8} = 57.625$$

**Example 8.** Determine the mode for the following distribution of monthly income of 580 lower-class people :

| Income (Rs.) : | less than 300 | 300 – 350 | 350 – 400 | 400 – 450 | 450 – 500 |
|---|---|---|---|---|---|
| Frequency : | 53 | 81 | 114 | 195 | 63 |
| Income (Rs.) : | 500 – 550 | 550 – 600 | 600 – 650 | 650 – 700 | more than or equal to 700 |
| Frequency : | 32 | 20 | 11 | 8 | 3 |

(G.N.D.U. B.Sc. C.Sc. April 2005)

**Sol.** First, we complete the class intervals as follows :

| Income (Rs.) (X) | Frequency ($f$) |
|---|---|
| 250 – 300 | 53 |
| 300 – 350 | 81 |
| 350 – 400 | 114 |
| 400–450 | 195 |
| 450–500 | 63 |
| 500 – 550 | 32 |
| 550 – 600 | 20 |
| 600 – 650 | 11 |
| 650 – 700 | 8 |
| 700 – 750 | 3 |

By inspection, mode lies in the class-interval 400 – 450.

Now, $$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where $$\Delta_1 = |f_1 - f_0|,\ \ \Delta_2 = |f_1 - f_2|$$

Here $L = 400,\ f_0 = 114,\ f_1 = 195,\ f_2 = 63$ and $i = 50$

$\therefore$ $$\Delta_1 = |195 - 114| = 81,\ \Delta_2 = |195 - 63| = 132$$

So $$\text{Mode} = 400 + \frac{81}{81 + 132} \times 50 = 419.01$$

**Example 9.** The following table shows the marks obtained by 100 students. Find the mode of this distribution :

| Marks : | < 10 | < 20 | < 30 | < 40 | < 50 | < 60 | < 70 | < 80 |
|---|---|---|---|---|---|---|---|---|
| No. of Students : | 7 | 21 | 34 | 46 | 66 | 77 | 92 | 100 |

(P.U. B.C.A. Sept. 2008)

Sol. First, we convert the cumulative frequency distribution into simple frequency distribution as follows :

| Marks (X) | Number of Students ($f$) |
|---|---|
| 0 – 10 | 7 |
| 10 – 20 | 14 |
| 20 – 30 | 13 |
| 30 – 40 | 12 |
| 40 – 50 | 20 |
| 50 – 60 | 11 |
| 60 – 70 | 15 |
| 70 – 80 | 8 |

By inspection, mode lies in the interval 40 – 50.

Now, $$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where $$\Delta_1 = |f_1 - f_0|, \quad \Delta_2 = |f_2 - f_1|$$

Here $$L = 40,\ f_0 = 12,\ f_1 = 20,\ f_2 = 11 \text{ and } i = 10$$

$\therefore$ $$\Delta_1 = |12 - 20| = 8,\ \Delta_2 = |20 - 11| = 9$$

So $$\text{Mode} = 40 + \frac{8}{8+9} \times 10 = 44.70$$

**Example 10.** In the following wage distribution the median and mode are Rs. 33.5 and Rs. 34 respectively, but three class frequencies are missing. Find out missing frequencies.

| Wages (Rs.) : | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 | 60–70 | Total |
|---|---|---|---|---|---|---|---|---|
| Frequency : | 4 | 16 | ? | ? | ? | 6 | 4 | 230 |

(P.U. B.C.A. Sept. 2003)

Sol. Let the missing frequencies be $x$, $y$ and $z$.

Now, we prepare the following table :

| Wages (X) | $f$ | $c.f.$ |
|---|---|---|
| 0–10 | 4 | 4 |
| 10–20 | 16 | 20 |
| 20–30 | $x$ | $20 + x$ |
| 30–40 | $y$ | $20 + x + y$ |
| 40–50 | $z$ | $20 + x + y + z$ |
| 50–60 | 6 | $26 + x + y + z$ |
| 60–70 | 4 | $30 + x + y + z$ |
| | $N = 30 + x + y + z$ | |

Now $N = 30 + x + y + z = 230$ (Given)

$\therefore \quad x + y + z = 200 \quad \ldots(i)$

Now $\quad \text{Median} = \text{value of } \frac{N}{2} \text{ th item} = \text{value of } \frac{230}{2} \text{ th item} = \text{value of 115th item.}$

Also $\quad \text{Median} = 33.5$

$\therefore$ It lies in the class interval 30 − 40

We know $\quad \text{Median} = L + \dfrac{\frac{N}{2} - c.f.}{f} \times i$

$$\Rightarrow \quad 33.5 = 30 + \frac{115 - 20 - x}{y} \times 10$$

$$\Rightarrow \quad 3.5 = \frac{95 - x}{y} \times 10$$

$$\Rightarrow \quad 3.5\,y = 950 - 10\,x$$

$$10\,x + 3.5\,y = 950 \quad \ldots(ii)$$

Given Mode = 34 which lies in the class interval 30 − 40.

Also $\quad \text{Mode} = L + \dfrac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$

Here $L = 30$, $f_0 = x$, $f_1 = y$, $f_2 = z$ and $i = 10$

$$\Rightarrow \quad 34 = 30 + \frac{y - x}{2y - x - z} \times 10$$

$$\Rightarrow \quad \frac{34 - 30}{10} = \frac{y - x}{2y - x - z}$$

$$\Rightarrow \quad 8y - 4x - 4z = 10y - 10x$$

$$\Rightarrow \quad 6x - 2y - 4z = 0 \quad \ldots(iii)$$

On solving equations (*i*), (*ii*) and (*iii*) for $x$, $y$ and $z$, we get

$$x = 60, \; y = 100 \text{ and } z = 40$$

# EXERCISE 7.11

1. Calculate Mode from the following set of numbers :

   16, 18, 22, 16, 15, 16, 22, 16, 14, 10, 11, 16

2. Calculate mode from the following data :

| Sr. No. : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Marks : | 15 | 32 | 35 | 33 | 15 | 21 | 41 | 32 | 11 | 18 | 20 | 22 | 11 | 15 | 35 | 23 | 38 | 12 |

3. Calculate mode from the following data of the heights in inches of a group of students :

61, 62, 63, 61, 63, 64, 64, 60, 65, 63, 64, 65, 66, 64

Now suppose that a group of students whose heights are 60, 66, 59, 68, 67 and 70 inches, is added to the original group. Find mode of the combined group.

4. Find the mode of the following frequency distribution :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Size : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| Frequency : | 9 | 8 | 10 | 11 | 15 | 18 | 18 | 9 | 8 |

(P.U. B.C.A. April 2003)

5. Find the mode of the following frequency distribution :

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| *f*: | 3 | 8 | 15 | 23 | 35 | 40 | 32 | 28 | 20 | 45 | 14 | 6 |

(P.U. B.C.A. April 2004)

6. Calculate mean and mode from the following data :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| X : | 5 | 10 | 15 | 20 | 25 | 30 | 35 | 40 |
| *f* : | 5 | 9 | 13 | 21 | 20 | 15 | 8 | 3 |

7. Calculate mode for the following data :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Wages (X) : | 0 – 5 | 5 – 10 | 10 – 15 | 15 – 20 | 20 – 25 | 25 – 30 | 30 – 35 | 35 – 40 | 40 – 45 |
| No. of Workers (*f*) : | 20 | 24 | 32 | 28 | 20 | 16 | 3 | 10 | 8 |

(P.U. B.C.A. April 2002)

8. Calculate the mode of the following data :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Wages (Rs.) (X) : | 25 – 35 | 35 – 45 | 45 – 55 | 55 – 65 | 65 – 75 | 75 – 85 | 85–95 | 95–105 | 105–115 |
| No. of Workers (*f*) : | 4 | 44 | 38 | 28 | 6 | 8 | 12 | 2 | 2 |

(P.U. B.C.A. Sept. 2005)

9. Calculate the value of the mode from the following data :

| | | | | |
|---|---|---|---|---|
| Class Interval : | 100 – 110 | 110 – 120 | 120 – 130 | 130 – 140 |
| Frequency : | 4 | 6 | 20 | 32 |
| Class Interval : | 140 – 150 | 150 – 160 | 160 – 170 | 170 – 180 |
| Frequency : | 33 | 17 | 8 | 2 |

10. Find the modal wage from the following :

| | | | | |
|---|---|---|---|---|
| Wages (Rs.) : | 50 – 59.99 | 60 – 69.99 | 70 – 79.99 | 80 – 89.99 |
| No. of Employees : | 8 | 10 | 16 | 14 |
| Wages (Rs.) : | 90 – 99.99 | 100 – 109.99 | 110 – 119.99 | |
| No. of Employees : | 10 | 5 | 2 | |

11. Calculate mode from following :

| | | | | | |
|---|---|---|---|---|---|
| X : | Less than 50 | 50 – 100 | 100 – 150 | 150 – 200 | More than or equal to 200 |
| $f$ : | 4 | 15 | 8 | 5 | 2 |

12. Find the modal wage group from following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Wages Above (Rs.) : | 330 | 340 | 350 | 360 | 370 | 380 | 390 |
| Workers : | 520 | 470 | 399 | 210 | 105 | 45 | 7 |

13. Calculate mode from following distribution

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mid Value : | 15 | 25 | 35 | 45 | 55 | 65 | 75 | 85 |
| Frequency : | 5 | 8 | 12 | 16 | 28 | 15 | 3 | 2 |

14. Calculate median and mode from the following data :

| | | | | | |
|---|---|---|---|---|---|
| Income between Rs. : | 100 – 200 | 100 – 300 | 100 – 400 | 100 – 500 | 100 – 600 |
| No. of Persons : | 15 | 33 | 63 | 83 | 100 |

15. Calculate missing frequency from the following data, given mode of distribution is 72.5

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| X : | 50 – 55 | 55 – 60 | 60 – 65 | 65 – 70 | 70 – 75 | 75 – 80 | 80 – 85 | 85 – 90 | 90 – 95 | 95 – 100 |
| $f$ : | 10 | 11 | 12 | ? | 25 | 15 | 10 | 15 | 9 | 8 |

16. Find out the missing frequency in the following distribution, if mode = 24 and number of families is equal to 100.

| | | | | | |
|---|---|---|---|---|---|
| Expenditure (Rs.) : | 0 – 10 | 10 – 20 | 20 – 30 | 30 – 40 | 40 – 50 |
| Families : | 14 | ? | 27 | ? | 15 |

## ANSWERS

| | | | |
|---|---|---|---|
| 1. 16 | 2. 15 | 3. 64, 64 | 4. 6 |
| 5. 6 | 6. 22.13, 15.74 | 7. 13.33 | 8. 48.75 |
| 9. 140.05 | 10. 77.495 | 11. 80.56 | 12. 355.84 |
| 13. 54.8 | 14. 356.67, 354.54 | 15. 15 | 16. 23, 21 |

### 7.10.5 GRAPHICAL PRESENTATION OF MODE

The following steps are involved in the calculation of mode graphically :

**Step 1.** Prepare histogram on the basis of given information.

**Step 2.** The rectangle with the maximum height will corresponds to the modal class.

**Step 3.** The left hand corner of the highest rectangle is joined with the left hand corner of the next rectangle and the right hand corner of the highest rectangle is joined with the right hand corner of the preceding rectangle. (See fig. 7.9)

**Step 4.** From the point where these two lines meet, a line is drawn perpendicular to X-axis.

**Step 5.** The distance of the foot of the perpendicular from the origin measured along X-axis will give us the value of the mode.

Now we shall illustrate this method with the help of following example :

Consider the following data :

| | | | | | | |
|---|---|---|---|---|---|---|
| Income less than (Rs.) : | 100 | 200 | 300 | 400 | 500 | 600 |
| No. of Persons : | 8 | 22 | 35 | 60 | 67 | 70 |

To find mode, first convert the cumulative frequency distribution into simple frequency distribution as follows :

| Income (Rs.) | No. of Persons |
|---|---|
| 0 – 100 | 08 |
| 100 – 200 | (22 – 8 ) = 14 |
| 200 – 300 | (35–22) = 13 |
| 300 – 400 | (60–35) = 25 |
| 400 – 500 | (67–60) = 07 |
| 500 – 600 | (70–67) = 03 |
| | N = 70 |

Using the given data, we prepare a histogram as given below :

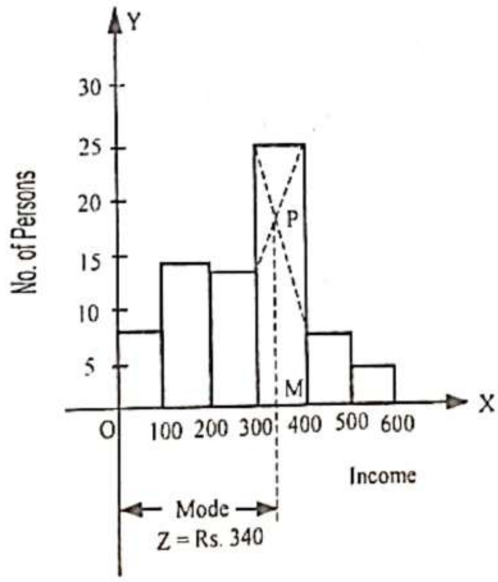



**Fig. 7.11**

From figure 7.11, Mode = OM = Rs. 340

**Verification :**

By inspection mode lies in the class interval 300 – 400.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where $\Delta_1 = |25 - 13| = 12$ and $\Delta_2 = |25 - 7| = 18$

$$\therefore \quad \text{Mode} = 300 + \frac{12}{12+18} \times 100 = 300 + \frac{1200}{30} = 340$$

# MISCELLANEOUS EXERCISE

1. The element tin consists of a mixture of ten isotopes which have atomic weights ranging from 112 to 124. The proportions in which the isotopes occur in the elements are given in the following table. Calculate the mean atomic weight of the mixture.

| Isotopes (A.W.) : | 112 | 114 | 115 | 116 | 117 | 118 | 119 | 120 | 122 | 124 |
|---|---|---|---|---|---|---|---|---|---|---|
| Percentage : | 1.1 | 0.8 | 0.4 | 15.5 | 9.1 | 22.5 | 9.8 | 28.5 | 5.5 | 6.8 |

2. The following table gives the distribution of 100 accidents during seven days of the week of a given month. During the particular month there are 5 Mondays, Tuesdays and Wednesdays and only 4 of each of the other days. Calculate the number of accidents per day ?

| Day : | Sun. | Mon. | Tue. | Wed. | Thurs. | Fri. | Sat |
|---|---|---|---|---|---|---|---|
| No. of Accidents : | 26 | 16 | 12 | 10 | 8 | 10 | 18 |

3. From the following data of income distribution of 100 persons calculate A.M. It is given that (*i*) total income of the persons in the highest group is Rs. 435 and (*ii*) None is earning less than Rs. 20

| Income (Rs.) : | below 30 | 40 | 50 | 60 | 70 | 80 | Above 80 |
|---|---|---|---|---|---|---|---|
| No. of Persons : | 16 | 36 | 61 | 76 | 87 | 95 | 5 |

4. Find the class intervals if the arithmetic mean of the following distribution is 33 and assumed mean is 35.

| Step Deviation : | 2 | 1 | 0 | − 1 | − 2 | − 3 |
|---|---|---|---|---|---|---|
| Frequency : | 10 | 20 | 30 | 25 | 10 | 5 |

5. Calculate A.M. from the following data :

| $X$ : | 1 | 2 | 3 | 4 − 6 | 7− 9 | 10 − 12 | 13 − 20 | 21 − 28 | 29− 36 |
|---|---|---|---|---|---|---|---|---|---|
| $f$: | 10 | 5 | 3 | 9 | 6 | 2 | 1 | 10 | 15 |

6. From the following data calculate average profits :

| Profits (Rs.) | Number of firms | Loss (Rs.) | Number of firms |
|---|---|---|---|
| 5000−6000 | 15 | 0−1000 | 15 |
| 4000−5000 | 20 | 1000−2000 | 16 |
| 3000−4000 | 35 | 2000−3000 | 18 |
| 2000−3000 | 15 | 3000−4000 | 13 |
| 1000−2000 | 5 | 4000−5000 | 12 |
| 0−1000 | 6 | | |
| | | | Total =170 |

7. Geometric mean of 10 items is 16.2. If an item 21.9 is misread as 12.9, what will be the corrected G.M. ?

8. Given the following data :

| | | | | | |
|---|---|---|---|---|---|
| Marks in Law : | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 |
| No. of Students : | 20 | 25 | 50 | 40 | 15 |

Calculate the median marks. If 60% of the students pass test find the minimum pass marks obtained by a student.

9. Find the median from the following distribution :

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| X : | 1 | 2 | 3 | 4 – 6 | 7 – 9 | 10 – 12 | 13 – 20 | 21 – 28 | 29 – 36 |
| Frequency : | 10 | 5 | 3 | 9 | 6 | 2 | 1 | 10 | 15 |

10. The following are the marks obtained by 50 students in Statistics.

| | | | | | | |
|---|---|---|---|---|---|---|
| Marks less than : | 10 | 20 | 30 | 40 | 50 | 60 |
| No. of Students : | 4 | 10 | 30 | 40 | 47 | 50 |

Calculate median and give the ranges within which middle 50% and middle 80% students fall.

11. Given below are the weekly wages in Rs. of 60 workers in a factory manufacturing plastic products :

23, 48, 51, 64, 72, 82, 56, 37, 50, 42, 35, 88, 77, 65, 39, 52, 48, 64, 49, 57, 41, 72, 62, 49, 32, 57, 67, 46, 55, 50, 82, 44, 75, 56, 51, 63, 59, 69, 53, 42, 75, 85, 68, 55, 52, 45, 42, 57, 20, 57, 46, 51, 50, 16, 62, 56, 54, 40, 55, 71.

Form a frequency distribution taking the lowest interval as 10 – 20 and compute mode and median.

12. From the following data find out the missing frequency :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Marks : | 55 | 65 | 75 | 85 | 95 | 105 | 115 |
| Frequency : | 8 | 10 | 16 | ? | 10 | 5 | 2 |

Arithmetic mean being 79.77.

13. Calculate mean, median and mode for the following frequency distribution :

| | | | | | |
|---|---|---|---|---|---|
| Marks : | 0 – 20 | 20 – 40 | 40 – 60 | 60 – 80 | 80 – 100 |
| Frequency : | 3 | 17 | 27 | 20 | 9 |

(P.U. B.C.A. April 2006)

14. The following table shows the number of telephone calls received at an exchange per interval for 245 successive one minute intervals :

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| No. of Calls : | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Frequency : | 14 | 21 | 25 | 43 | 51 | 40 | 39 | 12 |

Estimate mean, median and mode.

(G.N.D.U. B.Sc. C.Sc. Sept. 2006)

15. Compute the mean, median and mode for the following distribution of I.Q. for 300 six year old children.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| I. Q. : | 160–169 | 150–159 | 140–149 | 130–139 | 120–129 | 110–119 | 100–109 |
| Frequency : | 2 | 3 | 7 | 19 | 37 | 79 | 60 |
| I.Q. : | 90–99 | 80–89 | 70–79 | 60–69 | 50–59 | 40–49 | |
| Frequency : | 65 | 17 | 5 | 3 | 2 | 1 | |

(G.N.D.U. B.Sc. C.Sc. April 2006)

16. Obtain the value of mean, median and mode for the following data :

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Class : | 0–10 | 10–20 | 20–30 | 30–40 | 40–50 | 50–60 | 60–70 |
| Frequency : | 4 | 9 | 17 | 25 | 11 | 8 | 6 |

(G.N.D.U. B.Sc. C.Sc. Sept. 2007)

17. The arithmetic mean, mode and median of a group of 75 observations were calculated to be 27, 34 and 29 respectively. It was later discovered that one observation was wrongly read as 43 instead of correct value 53. Examine to what extent the calculated values of the three averages will be affected by the error.

(G.N.D.U. B.C.A. Sept. 2007)

## ANSWERS

**1.** 118.785 **2.** 14 **3.** Rs. 48

**4.** 0–10, 10–20, ....., 50–60 **5.** 14.64 **6.** Rs. 982.35 **7.** 17.08

**8.** 26, 23 **9.** 8.25 **10.** 27.5, 16.25, 35.47 **11.** 54, 54.76

**12.** 14 **13.** 53.95, 53.33, 51.76 **14.** 3.76, 4, 4

**15.** 108.6, 109, 101.58 **16.** 34.75, 34, 33.64

**17.** Correct mean = 27.13, median and mode remain unchanged