UNIT 12 MODE

Structure

12.0 OBJECTIVES

After studying this unit, you should be able to:

- define mode
- compute mode for different types of data
- locate mode graphically
- appreciate the limitations and uses of mode.

12.1 INTRODUCTION

As you know, among the measures of central tendency, there are some measures which are based on all items of the data and some other measures which are positional averages. In Unit 10 you have studied about arithmetic mean which is based on all the items of data. In Unit 11 you have studied about median which is a positional average. In this unit you will study about Mode which is another positional average. You will learn the meaning, methods of computation, locating it graphically, limitations and uses of mode.

12.2 WHAT IS MODE?

Mode is also a measure of central tendency. Mode is the value of a variate which is repeated most often in the data set. The genesis of the word 'mode' lies in the French word 'le mode' that means fashion. Mode is, therefore, considered to be the most common or most fashionable value.

Mode is often considered to be that value of the variate which occurs most frequently. But it is not exactly true for every frequency distribution. Rather it is that value of the variate around which the other items tend to concentrate most heavily. It shows the centre of concentration of the frequency in and around a given value. If is not the centre of gravity like mean. It is apositional measure similar to median. It is commonly denoted by M_o.

For example, take the case of a shopkeeper who sells shoes. He is interested to know the sizes of shoes which are commonly demanded. Were in such a situation, mean would indicate a size that may not fit any person. Median may not provide a representative size because of the unevenness in the distribution. It is the mode which will help in making a choice of approximate size for which an order can be placed.

12.3 COMPUTATION OF MODE

The method of computing mode is different for grouped data and ungrouped data, Now let us study those methods separately.

12.3.1 Ungrouped Data

For an ungrouped data mode is found out simply by inspection. The value that occurs most frequently in the given distribution is taken as a mode. For example, the ages (in years) of 10 boys are as follows: 5, 6, 4, 10, 7, 6, 9, 2, 8, 6. Here the number six appeared thrice. Therefore, mode age is six years.

Mode does not exist as such in some cases. For example, take the following data set : 5, 10, 15, 20, 25, 30. In this case there is no mode because none of the numbers is repeated.

In some cases there may be more than one mode. For example, one typist typed 10 pages and the number of mistakes per page are as follows: 5, 1, 0, 1, 2, 1, 2, 3, 2, 4. In this case, both the numbers 1 and 2 appear equal number of times. Therefore, there are two modes: 1 and 2. Similarly, the distribution can be a tri-modal or even multi-modal. For such distributions, the mode as a measure of central tendency has little significance. Mode has very limited use for ungrouped data.

12.3.2 Grouped Data

The method of computing mode is different between discrete distribution and continuous distribution. Let us now study those methods in detail.

Discrete Series

For discrete distribution, i.e., when the values of individual items are known, mode can be determined just by inspection. By inspection you can find out the value of the variate around which the items are most heavily concentrated. For example, study the following frequency distribution:

Size of Item	:	20	21	22	23	24	25
Frequency	:	15	20	25	45	30	12

In this frequency distribution, 23 has the highest frequency, implying that there is a heavy concentration of items at this value. Therefore, mode is 23.

In a series like this it is easy to obtain mode. Difficulty arises when nearly equal concentrations are found in two or more neighbouring classes; i.e., there is a small difference between the maximum frequency and the frequency preceding it or succeeding it. To locate a modal class in such situations, there is a need for Grouping and Analysis.

Grouping Table : A grouping table has six columns as explained below:

Column 1 : It is of class frequencies written against each class.

Column 2 : Frequencies are grouped in this column in two's, and totals are found. Then the highest total is marked or circled.

Column 3 : Leaving first frequency from the top, the remaining frequencies are again grouped in two's and the highest total is marked.

Column 4 : Starting from the top, frequencies are grouped in three's, their totals are obtained and the highest total is marked.

Column 5 : Leaving first frequency, they are again grouped in three's. Their totals are obtained and the highest total is marked.

Column 6 : Leaving the first two frequencies from the top, remaining frequencies are grouped in three's. Their totals are calculated and the highest total is marked.

Analysis Table : After preparing a grouping table, an analysis table is prepared. It is two-fold : 1) vertical (i.e., stubs) where the column numbers, as obtained in a grouping table, are taken and 2) horizontal (i.e., captions) where the values of the variate (or the classes) are taken. Now you take the grouping table, where you have marked or circled highest frequencies in every column, Take these circled frequencies in turns along with the corresponding values of the variate. In the analysis table under these values and in the row corresponding to relevant column number, tally bars are placed, The number of bars placed in each column of an analysis table are totalled. The maximum of these totals is marked. The value of the variate corresponding to it is the mode or the modal class. Let us study the preparation of grouping and analysis tables by taking an illustration.

Illustration 1

Find the mode (M_o) for the following information on the marks obtained by the students:

Marks	55	60	61	62	63	64	65	66	68	70
No. of Students	4	6	5	10	20	22	24	6	2	1

Solution

As you notice here, the difference between the highest frequency (i.e. 24) and the two frequencies preceding it (i.e., 22 and 20) is very small. The frequency which is next to the highest frequency (i.e., 6) also is very small. Therefore, grouping has to be done to ascertain the modal class.

Grouping Table

Marks	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
55	4		×		×	×
60	6	10		15		×
61	5		11		21	
62	10	15				35
63	20		30	52		
64	22	42			66	
65	24		46			52
66	6	30		32		
68	2		8		9	
70	1	3	×	×		×

Analysis Table

Col. No.	Marks 55	60	61	62	63	64	65	66	68	70
1							I			
2					I	I				
3						I	I			
4				I	I	I				
5					I	I	I			
6						I	I	I		
Total				1	3	5	4	1		

The highest total in the analysis table is five. The item corresponding to it is 64. Therefore, the mode (M_o) is 64. It may be noted here that the highest frequency (as shown in data) is for 65, whereas grouping and analysis tables indicated concentration of frequencies around 64. Thus, the correct value of mode is 64.

Continuous Series (i.e. data with class intervals)

In the case of contiquous series, (i.e. data with class intervals) which have equal class intervals throughout, there are two major steps in computing the mode.

Step 1 : Ascertain the modal class by preparing grouping table and analysis table exactly in the same way as discrete series. The minor difference in the procedure is that different classes of the given frequency distribution are taken vertically.

Step 2 : Having located correctly a modal class mode (M_o) is obtained by interpolation by using any of the following formulas:

a) $$M_o = l + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

where l = lower limit of the modal class

i = class interval

$\Delta_1 = f_1 - f_0$

$\Delta_2 = f_1 - f_2$

f_1 **is** the frequency of the modal class

f_0 is the frequency of the class preceding the **modal** class

f_2 is the frequency of the class succeeding the modal class.

By substituting the values of A, and A, in the above formula :

$$M_o = l + \frac{f_1 - f_0}{(f_1 - f_0) + (f_1 - f_2)} \times i$$

$$= l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

Note: If $(2f_1 - f_0 - f_2)$ is zero, the **formula** becomes meaningless. If any numerator or denominator becomes negative, then the formula does not give valid result. In that case **it** should be taken as:

$$M_o = l + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times i$$

where $|f_1 - f_0|$ and $|f_1 - f_2|$ means absolute values of the difference **i.e.,** difference neglecting **signs.**

b) The mode also can be **calculated** by **using** the upper limit of the **modal** class.

$$M_o = u - \frac{f_1 - f_2}{(f_1 - f_0) + (f_1 - f_2)} \times i$$

c) Where the modal class is other than the one containing the maximum frequency, the following **formula** is more **suitable** :

$$M_o = l + \frac{f_2}{f_0 + f_2} \times i$$

Notes :

1) If the very first class of the frequency distribution is **the modal** class, the f_0 is taken as zero. If modal class is the last group, then f_2 is taken as zero.

2) These formulas hold good only for the distributions with equal class intervals. Why is it so? The reason is simple. If two class intervals of size 10 and 20 have frequencies 15 and 18 respectively, then on simple comparison it appears frequency 18 is **larger** than 15. But mode is **concerned with** concentration of items. Concentration for the first group is 15/10 or 1.5 items per unit length of class interval. While in the second case it is only 18/20 or 0.9 items per unit length of class interval. Thus, from the point of view of determining mode, **frequency 18** for class interval size 20 is less than the frequency 15 for the class **interval** size 10. Therefore, direct comparisons of frequencies can only be made when class intervals are equal.

3) For the distributions with unequal class intervals, first the class intervals are made equal assuming that frequencies are uniformly distributed or by combining groups and then apply the usual formula.

Illustration 2

For the following frequency table, calculate the mode:

Monthly Rent Paid (Rs.)	No. of Families Paying the Rent
20 - 40	6
40 - 60	9
60 - 80	11
80 - 100	14
100 - 120	20
120 - 140	15
140 - 160	10
160 - 180	8
180 - 200	7
	100

Solution

By inspection the modal class appears to be **100-120,** but let us verify by grouping.

Grouping Table

Monthly Rent (Rs.)	Col. 1	Col. 2	Col.3	Col. 4	Col.5	Col. 6
20 - 40	6		×		×	×
		15				
40 - 60	9			26		×
			20			
60 - 80	11				34	
		25				
80 - 100	14					45
			34			
100 - 120	20			49		
		35				
120 - 140	15				45	
			25			
140 - 160	10					33
		18				
160 - 180	8			25	×	
			15			
180 - 200	7	×			×	×

Analysis Table

	Monthly Rent (Rs.)								
Col. No.	20-40	40-60	60-80	80-100	100-120	120-140	140-160	160-180	180-200
1					I				
2					I	I			
3				I	I				
4				I	I	I			
5					I	I	I		
6			I	I	I				
Total			1	3	6	3	1		

The highest total being 6, the modal group is **100 - 120.**

Applying the formula :

$$M_o = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$$

$$= 100 + \frac{20 - 14}{2(20) - 14 - 15} \times 20$$

$$= 100 + \frac{6}{11} \times 20$$

$$= 100 + 10.91$$

$$= 110.91$$

$\therefore$ Mode of monthly rent is Rs. **110.91**

Illustration 3

Calculate the mode from the following data:

Size	Frequency
0 - 9	3
10 - 19	4
20 - 29	8
30 - 39	7
40 - 49	6
50 - 59	3

Solution

By inspection, it is difficult to ascertain the modal class. Therefore, we have to resort to grouping.

Grouping Table

Size	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
0 - 9	3	7	×		×	×
10 - 19	4		12	15		×
20 - 29	8	15			19	
30 - 39	7		13			21
40 - 49	6	9		16	×	
50 - 59	3		×		×	×

Analysis Table

Col. No.	0 - 9	10 - 19	20 - 29	30 - 39	40 - 49	50 - 59
1			I			
2			I	I		
3				I	I	
4				I	I	
5			I	I		
6			I	I	I	
Total			4	5	3	

Prom the analysis table, it is obvious that 30-39 is the modal class. But the maximum frequency lies in class 20-29. Therefore, a more suitable formula for calculating the mode is:

$$M_o = l + \frac{f_2}{f_0 + f_2} \times i$$

$$= 29.5 + \frac{6}{8+6} \times 10 \qquad \text{(29.5 being real limit)}$$

$$= 29.5 + \frac{60}{14}$$

$$= 29.5 + 4.29$$

$$= 33.79$$

Therefore, mode is 33.8, You may note that a different result will be obtained if mode is calculated by the following formula :

$$M_o = l + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times i$$

$$= 29.5 + \frac{|7-8|}{|7-8| + |7-6|} \times 10$$

$$= 29.5 + \frac{1}{1+1} \times 10$$

$$= 29.5 + 10/2$$

$$= 34.5$$

You should note that the mode is **34.5** under this method whereas under the earlier method it is **33.8.** If you use the formula $M_o = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$ denominator will become zero and the numerator will be negative and, therefore, this formula is not applicable. It is important to note that unlike arithmetic mean and median, the different methods of calculating mode can give different results.

12.3.3 Smooth Data

When the data shows more or less uniform movement, it is called the smooth data. For such data mode can be obtained easily without using any of the above formulas. It can be worked out by'a very simple calculation. The rules to be followed for computing mode for smooth data are as under : When $f_0 = f_2$ i.e., the frequencies neighbouring the modal class frequency are equal, the mode is the mid-point of the two limits of the modal class. Study the following illustration carefully.

Size (x) :	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequency (f) :	1	6	15	20	15	6	1

The highest frequency being **20,** the modal class here is **30-40.** Since each of the two frequencies neighbouring the maximum frequency are equal (i.e., 15), the mode is the simple mean of **30** to 40.

Therefore, $M_o = \frac{30 + 40}{2} = 35$

You may verify whether the result obtained by this formula is the same as the result obtained by the methods suggested earlier for the grouped data, Whenever $f_0 = f_2$ and both f_0 and f_2 are less than f_1 this will always happen. When $f_0 \neq f_2$ (i.e., the two frequencies neighbouring the modal frequency are not equal) and the difference between the neighbouring frequency and the modal frequency is not very large, the mode is the weighted mean of the two limits—upper (u) and the lower (l) of modal class—the weights being the neighbouring frequencies falling on either side of a modal class. Therefore $M_o = \frac{lf_0 + uf_2}{f_0 + f_2}$.
For an example, study the following illustration :

'Size :	0-10	10-20	20-30	30-40	49-50	50-60	60-70
Frequency :	500	610	740	748	745	690	500

Here the modal class is **30-40** corresponding to the highest frequency **748** (f_1). Two neighbouring frequencies are **740** (f_0) and **745** (f_2) which are not equal and they do not differ 'much from f_1. The modal class is **30-40,** 'l' is **30** and 'u' is **40**

$$\therefore M_o = \frac{30 \times 740 + 40 \times 745}{740 + 745}$$
$$= \frac{52,000}{1,485}$$
$$= 35.02$$

The result derived by this method will always be the same as obtained by using the formula : $M_o = l + \frac{f_2}{f_0 + f_2} \times i$ You may verify it.

Illustration 4
From the data given below, find the mode.

Age in Years :	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No. of Persons:	50	70	80	180	150	120	70	50

Solution
The highest frequency is in the group **35-40.** But concentration of frequency appears to be around the group 40-45. So we do grouping for ascertaining the modal class.

Grouping Table

Ages	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
20 - 25	50		x		x	x
		120				x
25 - 30	70			200		
			1 0			
30 - 35	80				330	
		260				
35 - 40	180					410
			330			
40 - 45	150			450		
		270				
45 - 50	120				340	
			190			
50 - 55	70			x		240
		120				
55-60	50		x	x	x	

We observe here that class **40-45** participates in maximum frequency in Columns **2, 3, 4, 5** and 6, **(i.e.,** 5 times out of six columns) and class 35-40 participates only 4 times. You **may** verify it by **analysis** table.

using the formula $M_o = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$

$$M_o = 40 + \frac{150 - 180}{2 \times 150 - 180 - 120} \times 5 = 40 + \frac{-30}{0} \times 5$$

So mode cannot be determined as $2f_1 - f_0 - f_2 = 2 \times 150 - 180 - 120 = 0$. Therefore, we will use the following formula :

$$M_o = l + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times i$$

$$= 40 + \frac{|150 - 180|}{|150 - 180| + |180 - 120|} \times 5$$

$$= 40 + \frac{30}{30 + 60} \times 5$$

$$= 40 + \frac{5}{3}$$

$$= 40 + 1.67$$

$$= 41.67$$

$\therefore$ **Modal** Age = 41.67

12.3.4 Empirical Method

In a symmetrical distribution (like the one taken in section 12.3.3) the values of mean, median and mode coincide. You **can verify** it. But in the case of distribution which is not symmetrical **(i.e.,** when frequencies at equal distance from central class are not equal), there are **two possibilities :**

1) **When** there is **greater** concentration in lower values, such distribution is known as **positively skewed distribution.** As shown **in** Figure 12.1, this type of distribution shall have a longer tail on right hand side. In this type of distribution, the value of the mean **is** the highest, **the** value of the mode is the lowest and median lies between mean and mode. The distance between mean and **median** is about one-third the distance between mean and mode.

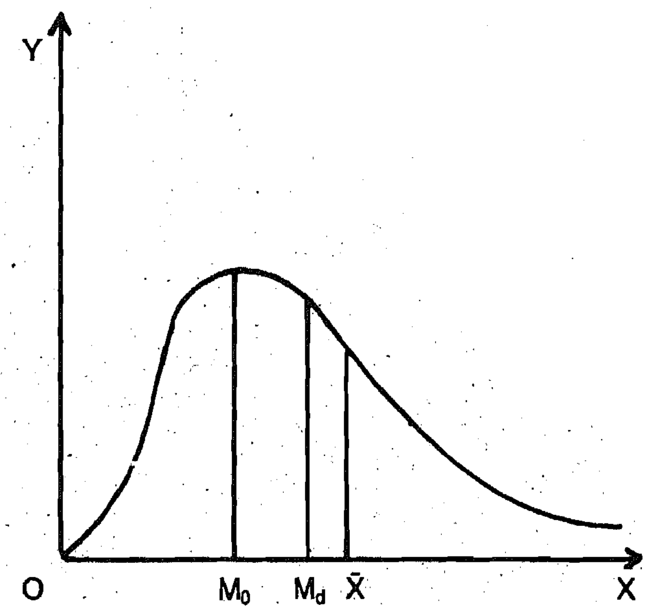

Figure 12.1 Positively Skewed Distribution

2) There may be greater concentration of the items in higher values. Such a distribution is known as negatively skewed distribution. Study Figure 12.2 carefully. You should note that this type of distribution has a longer tail on left band side. Here the mean will be the lowest, the mode will be the highest and median lies about one-third the distance from mean towards mode.

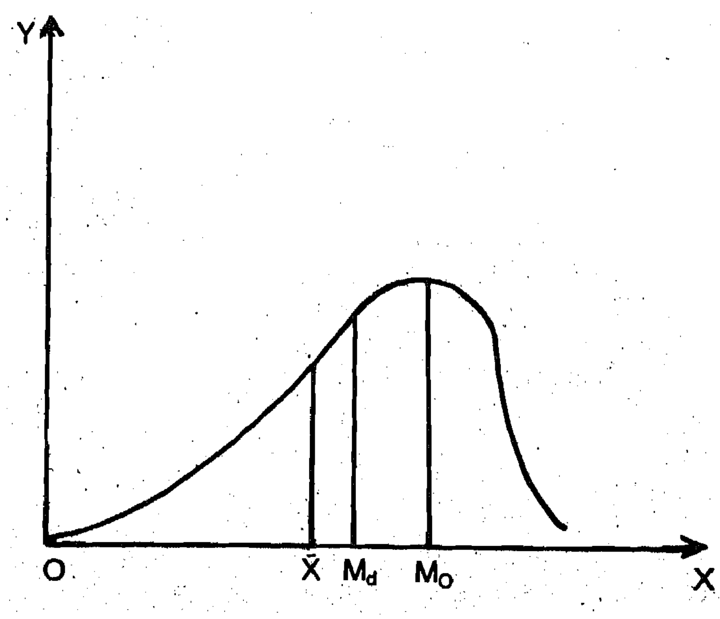

Figure 12.2 Negatively Skewed Distribution

The relationship between mean, mode and median in the two situations explained above is established by Karl Pearson using the following formula':

Mean − Mode = 3 (Mean − Median). This gives

Mode = Mean − 3 (Mean − Median)

= '3 Median − 2 Mean

or $M_o = 3M_d - 2\bar{X}$

This is an empirical relationship because this has been observed to be true in most of the moderately skewed data. It may not be true in extreme situations. A mathematical proof for it is not possible. In this situation where the mode is ill-defined, its value can be obtained empirically by using the above formula.

Illustration 5

Find the mode from the following table :

Size of the Items	Frequency
40 - 49	7
50 - 59	9
60 - 69	10
70 - 79	6
80 - 89	13
90 - 99	10
100 - 109	12
110 - 119	7

Solution

By inspection, the modal class is not clear. Hence, we have to do grouping and analysis.

Grouping Table

Size	Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
40 - 49	7	16			x	x
50 - 59	9			26		x
60 - 69	10		19			
70 - 79	6	16				29
80 - 89	13		19	29		
90 - 99	10	23			35	
100 - 109	12		22	x		29
110 - 119	7	19	x	x	x	

Analysis Table

Col. No.	60-69	70-79	80-89	90-99	100-109	110-119
1			I			
2			I	I		
3				I	I	
4		I	I	I		
5			I	I	I	
6	I	I	I	I	I	
Total	1	2	5	5	3	

In the analysis table maximum total **5** occurs twice. The mode, therefore, is ill-defined and is to be determined empirically by using the formula: $M_o = 3M_d - 2\bar{x}$. You may check yourself that here Median = **83.84** and $\bar{x}$ = **80.14.**

$\therefore M_o = 3\ (83.84) - 2\ (80.14)$

$= 251.52 - 168.28$

$= 91.24$

$\therefore$ Mode = 91.24

Check Your Progress A

1) Define mode.

....

....

....

2) State the various formulas for the computation of mode.,

....

....

....

....

....

3) What is the empirical relationship between arithmetic mean. median, and mode?

....

....

....

....

....

4) For a frequency distribution, the mean is 26.8 and the median is 27.9. Find the value of mode.

....

....

....

....

5) State whether the statements given below are True or False.

i) Mode is a unique value in a distribution.

ii) Coniputation of mode neglects the extreme values of the distribution.

iii) Mode of a data cannot be greater than arithmetic mean.

iv) Mode is always found in a class with highest frequency.

v) Even though mode can be computed from data, it is not a mathematical average.

6) Fill in the Blanks:

i) When the purpose is to know the point of the .... concentration, mode is preferred.

ii) Mode and median are .... measures.

iii) There are two modes in the following series : 1, 2, 1, 2, 3, 2, 0, 1. They are .... and ....

iv) In a series when nearly equal concentration, found in two or more neighbouring classes, we prepare .... and .... tables to find mode.

v) If $2f_1 - f_0 - f_2$ is zero, the mode is obtained by the formula ....

vi) Read the following data :

X :	0-10	10-20	20-30	30-40	40-50	50-60	60-70
F :	2	8	12	18	12	5	3

Here the mode is the simple mean .... of and ....

vii) Read the following data:

X :	10-20	20-30	30-40	40-50	50-60	60-70
F :	100	125	220	228	222	150

In this case'mode is a weighted mean of thk two limits of modal class 40 and 50, the weights being .... and ....

viii) If two values in a given data set occur more often than any others, the distribution, is said to be ....

12.4 GRAPHICAL DETERMINATION OF MODE

You have studied various methods of computing the mode. In fact, like median, mode also can be determined graphically., Determination of mode graphically involves the following procedure:

1) Draw a histogram for the given frequency distribution. A partial histogram can also be drawn by using only three classes — pre-modal, modal and post-modal. You have studied about histogram in Unit 9.

2) The top right comers of the highest rectangle (modal class rectangle) and the preceding rectangle are joined by a straight line. Similarly, the top left corners of ~~the~~ highest rectangle and the rectangle just on its right are joined by a straight line.

3) Draw a perpendicular to x-axis from the point of intersection of these two straight lines.

4) The point where it meets the x-axis gives the value of the mode.

Let us understand this proce'dure clearly through an illustration.

Illustration 6

Find the mode of the following data graphically and also check the result through calculation:

Class Interval	:	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequency	:	4	18	30	42	24	10	3

Solution

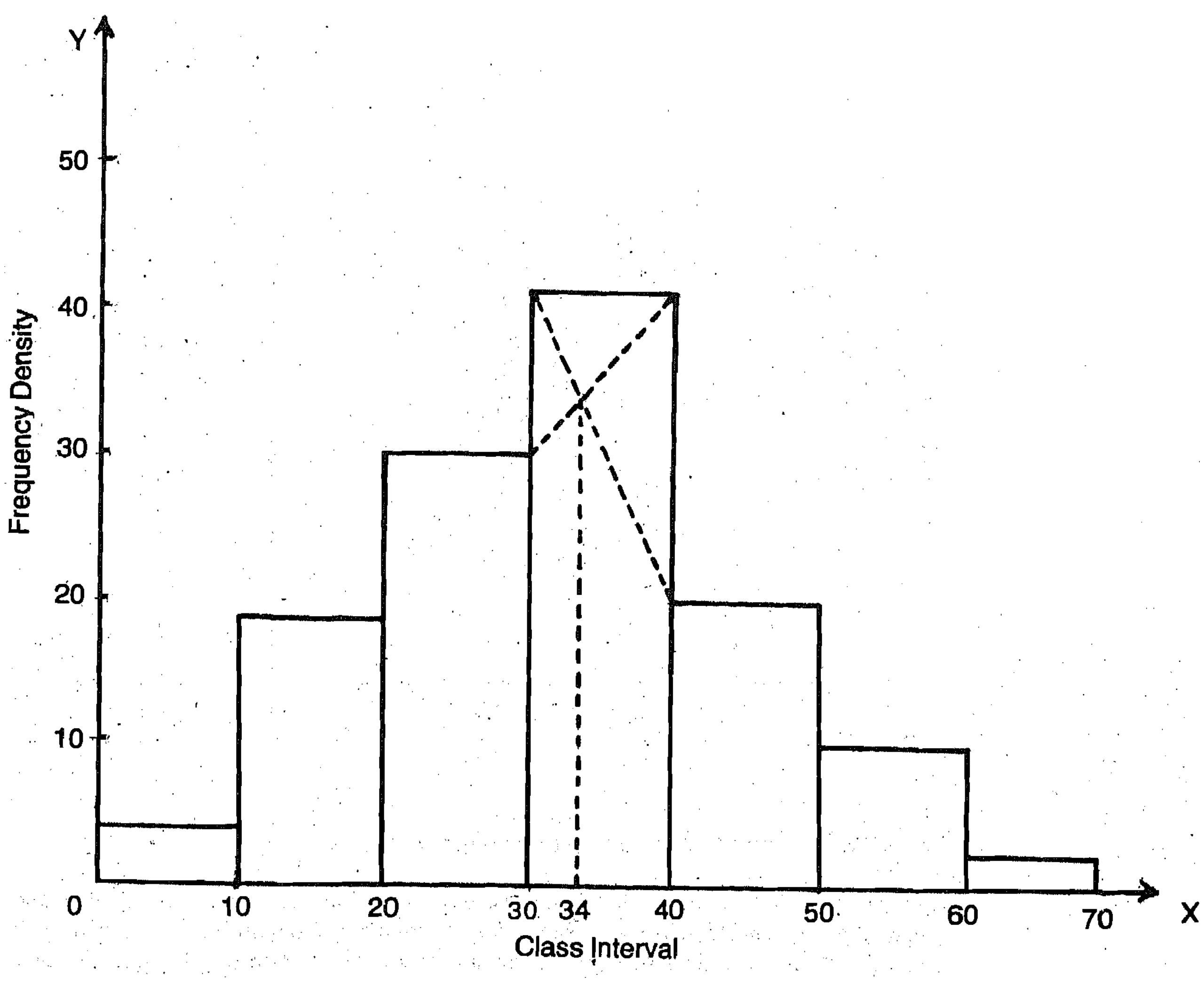

Figure 12.3. Histogram and Class Interval Determination of Mode

By using the usual formula $M_o = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$

$$M_o = 30 + \frac{42-30}{84-30-24} \times 10$$

$$= 30 + \frac{12}{30} \times 10$$

$$= 30 + 4$$

$$= 34$$

You must note that the value of mode obtained here is the same as the value obtained graphically. But, if you compute the mode by formula $M_o = l + \frac{f_2}{f_0 + f_2} \times i$ the result will not be the same as obtained by graph. The logic behind the graphic method and formula based on f_0, f_1, f_2 is same. The details of the logic are beyond the scope of this course.

There is one limitation of the graphical method of determining mode. When modal class is adjoining to the class with highest frequency, mode cannot be determined graphically in the modal class. It can only be determined from a class with highest frequency. Thus, mode calcula'ted graphically will not be a proper mode. To understand this, let **us** determine mode graphically **for the** data in Illustration **3** discussed earlier.

Solution

Here the class intervals are of inclusive type. So they have to be first converted to real limits before drawing the histogram. Now look at Figure 12.4 carefully.

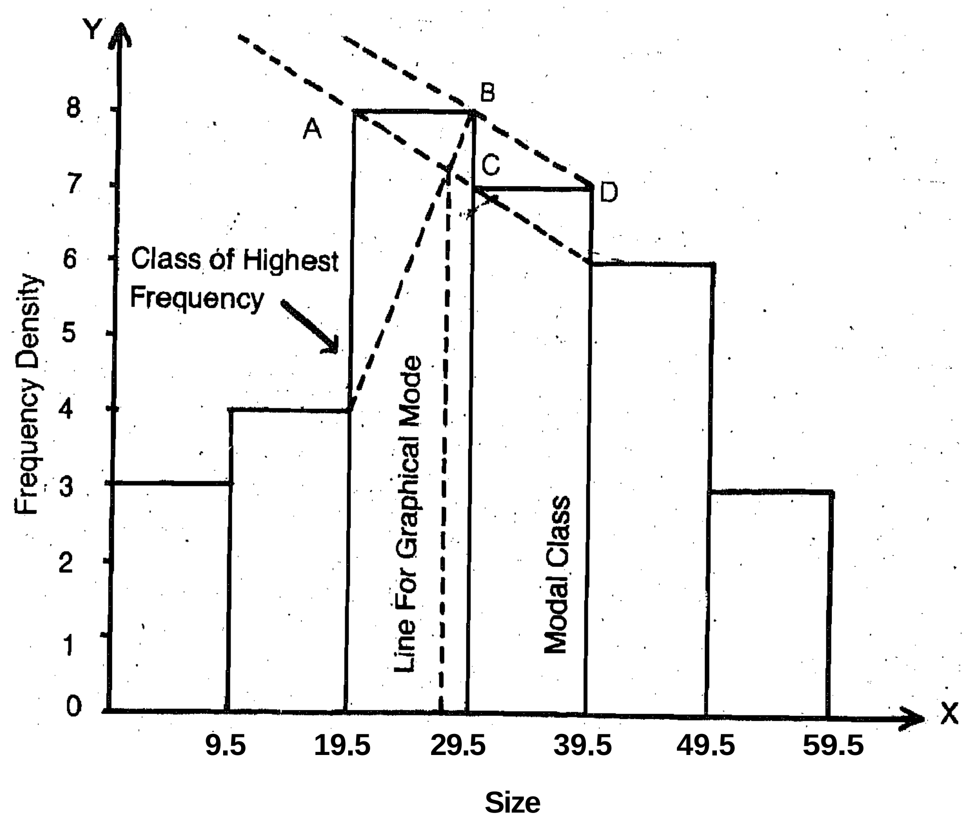

Figure 12.4. Histogram and Calculation of Mode

It may be seen that graphical value of mode can be determined from the class 19.5-29.5, and value turns out to be 27.5. This is different from the value 33.8 obtained earlier. If you try to determine the mode in the modal group, you have to join points **A** to C and B to D. The two lines **AC** and BD do not intersect in the modal class. Hence mode cannot be determined in the modal class by graphical method.

If the modal group and the group with highest frequency are not adjoining and separated by two to three groups, the mode can be determined graphically in both the groups. Out of these two modes, mode of first preference can be decided by looking to the height of the perpendicular drawn from the point of intersection to the x-axis. Such distributions can be termed as **bi-modal.** Let us take an illustration to explain this.

Illustration 7

The following distribution gives 'over time work' done by 100 employees of a company during a month. Determine the mode graphically.

Over time Hours :	10-12	12-14	14-16	16-18	18-20	20-22	22-24	24-26	26-28	28-30
No. of Employees :	3	5	16	21	17	6	4	23	3	2

Solution

By preparing grouping and analysis tables, you can easily verify that the modal class is 16-18. But the highest frequency is in the group 24-26 which is at a distance of 3 groups from modal class. Now look at Figure 12.5 for histogram and graphical determination of mode.

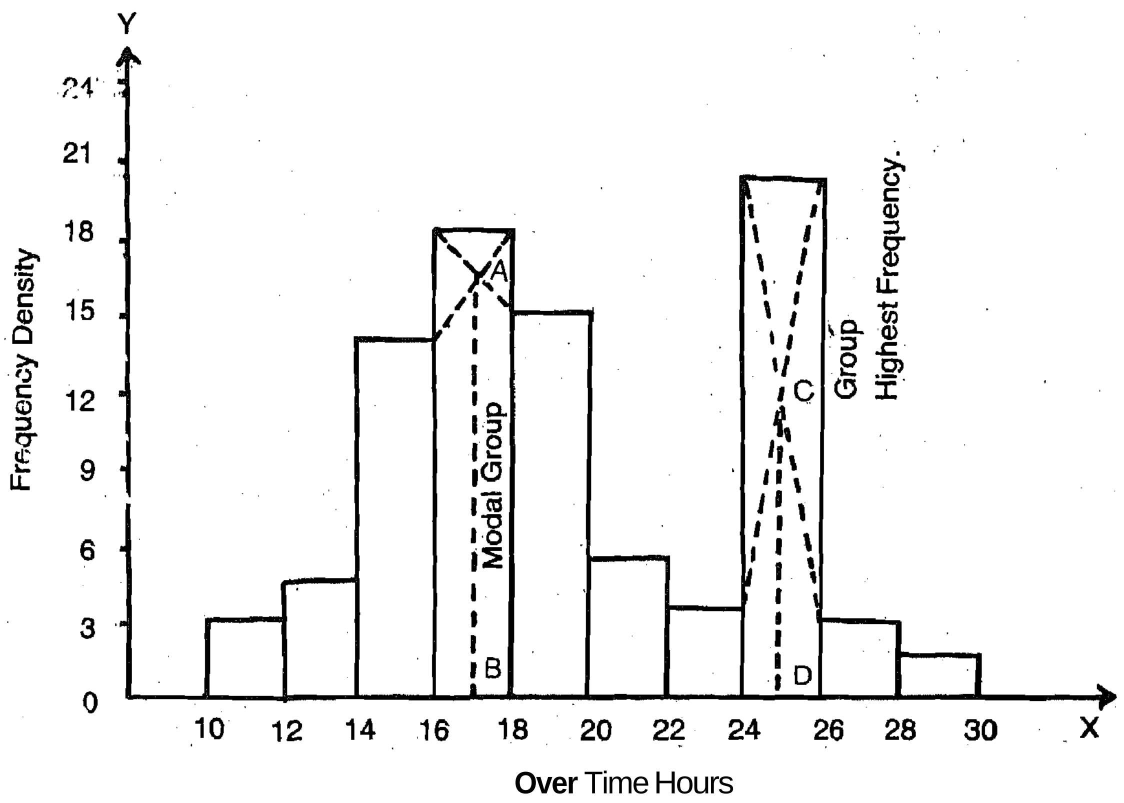

Figure 12.5. Histogram and Calculation of Mode

Point B in the modal class gives one mode. This is 16.9, approximately. Point D in the class with highest frequency gives another mode. This is 25.1, approximately. The length of the lines AB and CD gives the extent of concentration of items at two points B and D. As the length AB is greater than CD, there is a greater amount of concentration of frequencies at mode B than at mode D. Thus, the value of B (i.e. 16.9) is the mode of first preference value and D value (i.e. 25.1) will be taken as mode of second preference. By computation method also mode can be determined in both the groups. The values obtained will be approximately the same as the one determined by graph. As pointed out earlier such a data may be termed as bi-modal. However, **a perfect bi-modal data is one in which concentration of frequencies at the two modes is exactly equal.**

11.25 MERITS AND LIMITATIONS OF MODE

Merits

1) In certain situations mode is the only suitable average, e.g., modal size of garments, modal size of shoes, modal wages, modal balance of depositors in a bank, etc.

2) It is used to describe qualitative phenomena. For instance, if a printing press turns out five impressions which we rate very sharp, sharp, sharp, blurred and sharp, then the modal value is sharp.

3) For the preference of consumers' product, the modal preference is regarded. A restaurant owner who specialises in one dish may wish to know the modal preference of his potential clientele.

4) In the case of skewed distribution, mode is the indicator of the point of heaviest concentration.

5) It is very profitably used in market research.

6) Even if one or more classes are open-ended, mode can be used.

Limitations

1) Too often, there is no modal value. It is a useless measure, when there are more than one mode.
2) It is not capable of further algebraic treatment.
3) It is an ill-defined measure. Therefore, different formulas yield somewhat different answers.
4) It is not based on all the items of the data.
5) The value of the mode is affected significantly by the size of the class-intervals,
6) Although a mode is the value of a variate that occurs most frequently, its frequency does not represent a majority of the total frequencies.

Check Your Progress B

1) Why is it usually better to calculate a mode from grouped rather than ungrouped data?
   a) The ungrouped data tend to be bi-modal.
   b) The mode for the grouped data will be the same, regardless of the skewness of the distribution.
   c) Extreme values have less effect on grouped data.
   d) The chance of an unrepresentative value being chosen as the mode is reduced.

2) In which of the cases would a mode be most useful as an indicator of central tendency?
   a) Every value in a data set occurs exactly once.
   b) All but three values in a data set occur once, three values occur 100 times each.
   c) All values in a data set occur 100 times each.
   d) Every observation in a data set has the same value.

3) When bell-shaped distribution is symmetrical and has one mode, the highest point on the curve is referred to as a) Range; b) Mode; c) Median; d) Mean; e) All of these; f) b, c, and d, but not a.

4) State whether the following statements are True or False.
   i) Graphical method and computation methods of finding a mode always give identical values.
   ii) As mode can be computed from the data, it is capable of algebraic treatment.
   iii) Mode at times can be used to describe qualitative phenomena.
   iv) Mode plays an important role in checking the symmetry of the data.
   v) If it is not possible to compute mode, you cannot find it graphically also.

5) Fill in the blanks:

i) If the mean and median of a moderately asymmetrical series are 26.8 and 27.9 respectively, the most probable mode will be ....

ii) The approximate value of mode is 52 with mean 58 and median....................

iii) If the data set has only one mode and it is less than the mean, it can be concluded that the graph of the data is skewed to the ....

iv) For a moderately skewed distribution, the empirical relation is given as M_o = ...

v) The mode can be graphically determined by constructing the .... and using the .... rectangle and two .... rectangles.

vi) For preference of consumers' product, the ....preference is considered.

vii) Mode suffers from sampling....................

12.6 SOME ILLUSTRATIONS

Illustration **8**

Estimate the value of arithmetic mean if mode is 15.3 and median is 14.2

Solution

The empirical relation between mean, median and mode is:

$$M_o = 3M_d - 2\bar{x}$$

Substituting the values of M_o and M_d

$$15.3 = 3 \times 14.2 - 2\bar{x}$$

$$2\bar{x} = 42.6 - 15.3$$

$$2\bar{x} = 27.3$$

$$\bar{x} = 13.65$$

Illustration **9**

With the help of empirical relation between M_o, M_d, and $\bar{x}$ show **that**

i) $M_d = M_o + \frac{2}{3}(\bar{x} - M_o)$

ii) $\bar{x} = M_d + \frac{1}{2}(M_d - M_o)$

Solution

The empirical relation between mean, median and mode is:

$$M_o = 3M_d - 2\bar{x}$$

i) $M_o = 3M_d - 2\bar{x}$

$$M_o + 2\bar{X} = 3M_d$$

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

$$M_d = \frac{1}{3}M_o + \frac{2}{3}\bar{x}$$

$$= M_o - \frac{2}{3}M_o + \frac{2}{3}\bar{x}$$

$$= M_o + \frac{2}{3}(\bar{x} - M_o)$$

$$\therefore M_d = M_o + \frac{2}{3}(\bar{x} - M_o)$$

Median = Mode $+ \frac{2}{3}$ (Mean − Mode)

ii) $M_o = 3M_d - 2\bar{x}$

$$2\bar{x} = 3M_d - M_o$$

$$\bar{x} = \frac{3}{2}M_d - \frac{1}{2}M_o$$

$$\text{Mean} = \text{Median} + \frac{1}{2}(\text{Median} - \text{Mode})$$

Illustration 10

The following table gives the age (in years) of employees of a firm. The modal age is 32 years. Find the missing frequency.

Age in Years :	20-25	25-30	30-35	35-40	40-45
No of Employees:	5		18	9	6

Solution

Let us assume that the missing frequency is 'F'. As the mode is 32, the modal group is 30-35.

Now $M_o = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$

where $l = 30$, $f_0 = F$, $f_1 = 18$, $f_2 = 9$, $i = 5$ and $M_o = 32$

Substituting the x values:

$$32 = 30 + \frac{18 - F}{2 \times 18 - F - 9} \times 5$$

$$2 = \frac{18 - F}{27 - F} \times 5$$

$$54 - 2F = 90 - 5F$$

$$3F = 36$$

$$F = 12$$

$\therefore$ Missing frequency is 12.

Illustration 11

Calculate mode from the data given below:

Profit (Rs. in lakhs) :	0-5	5-10	10-20	20-30	30-50
No. of Companies:	4	6	15	18	20

Solution

Here the class intervals are not equal. In such cases two methods can be used:

i) Rewriting the data with equal class intervals, ii) Using empirical relationship.

i) On combining the first two groups, class intervals will become 0-10. Next two class intervals are of size 10. The last class interval is of size 20. It can be divided into two i.e. 30-40 and 40-50. Assuming frequencies as uniformly distributed, both such groups will have frequencies of 10 each. Thus, the given data can be written as:

Profit (Rs. in lakhs) :	0-10	10-20	20-30	30-40	40-50
No. of Companies :	10	15	18	10	10

It is clear that the modal class is 20-30

Now Mode $= l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times i$

Substituting the values of l, f_0, f_1, f_2, and i

$$M_o = 20 + \frac{18 - 15}{2 \times 18 - 15 - 10} \times 10$$

$$= 20 + \frac{3}{11} \times 10$$

$$= 20 + 2.7$$

$$= 22.7$$

$\therefore$ Mode of the profit is Rs. 22.7 lakhs.

ii) You may verify arithmetic mean = Rs. 24.3 lakhs and the median is Rs. 23.6 lakhs

Now mode = 3 Median − 2 Arithmetic Mean

$\therefore$ Mode $= 3 \times 23.6 - 2 \times 24.3$

$= 70.8 - 48.6$

$= 22.2$

∴ Mode of profit is Rs. 22.2 lakhs.

You may note that the modes amved at by the two methods differ. It is because, as you know, mode is not rigidly defined.

Illustration **12**

As a manager of a transport company you want to buy 100 tyres from either producer A or producer B. The price of the two tyre types is same. The following information is available about the average distance run by these two types of tyres :

Firm	Average distance run	
	Arithmetic Mean (km.)	Mode (km.)
Type A	35,000	32,000
Type B	32,000	35,000

i) which type would you buy?

ii) If you want to buy one tyre for your own car, will your decision be the same?

Solution

i) AM × No. of Items = Total value of items. So if you buy tyres from producer A, the total distance run by all 100 tyres would be 100 x 35,000 = 35,00,000 kms. If you buy from producer B, the total distance run would be 100 X 32,000 = 32,00,000. As the total distance run in first case is greater, you would prefer tyres of producer **A.**

ii) When you are buying only one tyre, it is not necessary that the tyre bought will give the same mileage as arithmetic mean. On the other hand, it is quite likely that the tyre you bought may give mileage equal to mode, the value around which you have maximum concentration of items. As the mode of producer B is higher in this case, you will prefer producer B product.

It may be noted that when large number of tyres are bought, some tyres may give mileage equal to arithmetic mean and others may give more than arithmetic mean. If the selection is done randomly, the mean of the distance run by the selected tyres would be almost same as the mean value claimed by the producer. Hence, in the case (i) arithmetic mean was used to assess which type of purchase gives greater service,

12.7 LET US SUM UP

The mode is the value of the variate around which the other items tend to concentrate most heavily. It can be computed for both ungrouped and grouped data. However, for ungrouped data it has a limited use. For a discrete distribution, mode is that value of the variate around which the items are most heavily concentrated. Where there are nearly equal concentrations in two or more neighbouring classes to a class with highest frequency, it is'difficult to determine the mode. In such cases 'grouping and analysis tables' are prepared to ascertain the modal class. For a continuous distribution, after having located a modal class, mode is calculated by using different interpolaltive formulas. For the data set displaying uniform movement, mode is obtained by using simple rules where it could be either a simple mean of two limits of a modal class or a weighted mean of them. For moderately skewed distribution, mode is obtained by an empirical relationship $M_o = 3\,M_d - 2\bar{x}$ Mode also can be graphically determined by constructing the histogram and using the highest rectangle and two of its neighbouring rectangles.

Mode is very useful in situations like finding a modal size of shoes, modal size of garments, modal wages, etc. It is also used to describe the qualitative phenomenon and to indicate modal preference of consumers for consumer products. Mode suffers from certain limitations such as incapability of further algebraic treatment, ill-defined nature, non-existence, presence of more than one mode, etc.

12.8 KEY WORDS AND SYMBOLS

Analysis Table: The table which helps to ascertain the modal class showing the maximum frequency occurring in different columns.

Bi-modal Distribution: A distribution of data in which two values occur more frequently than the rest of the values in the data set.

Empirical Relationship of Averages: The relationship that exists between averages in a moderately skewed distribution viz., $M_o = 3M_\alpha - 2\bar{x}$

Grouping Table: The table which has six columns, used for ascertaining a modal class.

Mode: The value of the variate around which the other items tend to concentrate most heavily.

Negatively Skewed Distribution: The distribution wherein there is a greater concentration in higher values with a longer tail on left hand side.

Positively **Skewed** Distribution: The distribution where there is a greater concentration in lower values with a longer tail on the right hand side.

List of symbols

In addition to the list of symbols given under Units 10 & 11, following is the list of symbols used in connection with Mode. The list is on the same lines as is Unit 10.

Difference between modal and next (higher values side) frequency	$f_1 - f_2$, $\triangle_2$, d_2. Where $\triangle_2$ and d_2 are always taken as positive.
Difference between modal and previous (lower values side) frequency	$f_1 - f_0$, $\triangle_1$, d_1. Where $\triangle_1$ and d_1 are always taken as positive.
Frequency of a group next to (higher values side) modal group.	f_2
Frequency of a group previous to (lower values side) modal group.	f_0, f_1 (when modal frequency not denoted by f_1)
Frequency of the modal group	f_1, f_m, f_{mo}, f.
Lower limit of the modal group	l, l_1, L, L_{M_o}
Mode	M_o, Z .
Upper limit of the modal group	u, l_2, U, U_m, U_{M_o}

12.9 ANSWERS TO CHECK YOUR PROGRESS

A) 4) 30.1

5) i) False; ii) True; iii) False; iv) False; v) True

6) i) highest; ii) positional; iii) 1, 2; iv) grouping analysis

v) $M_o = 1 + \frac{|f_1 - f_0|}{|f_1 - f_0| + |f_1 - f_2|} \times i$ OR $M_o = 1 + \frac{f_2}{f_0 + f_2} \times i$

vi) 30.40; vii) 220,222; viii) bi-modal;

B) 1) d

2) b

3) f

4) i) False; ii) False; iii) True; iv) True; v) True

5) i) 30.1; ii) 56; iii) right; iv) $3M_e - 2\bar{x}$; v) histogram, highest, neighbouring; vi) modal; vii) instability

12.10 TERMINAL QUESTIONS/EXERCISES

Questions

1) 'Arithmetic Mean, Median and Mode all try to give one main characteristic of the data but in their own way'. Discuss.

2) What is mode? Explain its limitations and uses as a measure of average?

Exercises

1) Find the modal age of married women at first child birth:

Age (Years) :	13	14	15	16	17	18	19	20	21	22	23	24	25
No. of Women :	37	162	343	390	256	433	161	355	'65	85	49	49	40

(Answer : 18 years)

2) From the following information regarding the wage distribution in a certain factory, determine the modal age:

Weekly Wage (Rs.)	No. of Employees
20 - 40	8
40 - 60	12
60 - 80	20
80 - 100	30
100 - 120	40
120 - 140	35
140 - 160	18
160 - 180	7
180 - 200	5

(Answer: Rs. 113.33)

3) Find the modal size of shoes from the following information:

Size of Shoes :	1	2	3	4	5	6	7	8	9	10
Frequency :	10	5	13	6	23	32	14	35	8	7

(Answer : 6)

4) The following table gives the relative distribution of sales calls made on Amar Phannaceuticals'in the past month. Find the modal calls.

No. of Sales Calls :	0	1	2	3	4	5 or more
Relative Frequency:	0.21	0.18	0.38	0.19	0:03	0.01

(Answer: 2 sales calls)

5) Calculate the mode for the following data:

Class :	10-20	20 - 30	30-40	40-50	50-60	60-70	70.-80
Frequency :	24	42	56	66	108	130	154

(Answer: 71.34)

6) Determine the most common salary graphically for the following data and verify it by using an appropriate formula:

Salary (more than Rs.) :	100	150	200	250	300	350	400	450
No. of Employees:	100	98	93	83	43	23	12	5

(Answer: 280)

7) Determine the mode graphically and also by computation.

Weight less than (kgs.) :	80	85	90	95	100	105	110	115	120	125
No. of Employees :	0	5	13	'30	55	75	85	93	120	125

(Answer : 98.1)

8) Obtain the mode for the following distributions without using the usual formulas:

i)

x :	0 - 10	10 - 20	20-30	30-40	40-50	50-60	60-70
f :	1	6	15	20	15	6	1

ii)

x :	48-52	52-56	56-60	60-64	64-68	68-72	72-76
f :	4	8	16	18	15	4	2

(Answers: i - 35, **ii** - 61.93)

9) Estimate the median when arithmetic mean is 27.9 and mode is 25.2. Give the assumptions, if any.

(Answer: 27)

10) What are the modal values of the following distributions?

a)

Hair Colour	:	Black	Brunette	Red Head	Blonde
Frequency	:	11	24	6	18

b)

Blood Group	:	AB	O	A	B
Frequency	:	4	12	35	16

(Answers: a - Brunette; b - A)

Note: These questions and exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the University. These are for your practice only.

UNIT 13 GEOMETRIC, HARMONIC AND MOVING AVERAGES

Structure

13.0 OBJECTIVES

After studying this unit, you should be able to:

- define and compute geometric mean and harmonic mean'
- enumerate the properties of geometric mean and harmonic mean
- apprec... the limitations and uses of geometric mean and harmonic mean
- explain the concept of moving average
- use moving average in determining trend of time series
- make a choice of suitable average in a given situation.

13.1 INTRODUCTION'

As you know, the averages can be classified as mathematical averages, positional averages and special averages. You have already studied about arithmetic mean which belongs to the category of mathematical averages, median and mode, which belong to the category of positional averages. In this unit you will study about the two other mathematical averages viz., Geometric Mean and Harmonic Mean. You will also study special average viz., Moving Average, and how to choose a suitable average amongst all types of averages in a given situation.

13.2 GEOMETRIC MEAN

In the situations where we deal with quantities that change over a period of time, we may be interested to know the average rate of change. In such cases the simple arithmetic mean is not suitable and we have to resort to the geometric mean.

13.2.1 Computation

Like other averages, computation procedure of geometric mean is different for grouped data and ungrouped data. Now let us study these methods.