Statistics

39.1 STATISTICS. Statistics is a branch of science dealing with the collection of data, organising, summarising, presenting and analysing data and drawing valid conclusions and thereafter making reasonable decisions on the basis of such analysis.

39.2 FREQUENCY DISTRIBUTION. Frequency distribution is the arranged data, summarised by distributing it into classes or categories with their frequencies.

Wages of 100 workers

Wages in Rs.	0-10	10-20	20-30	30-40	40-50
Numbers of workers	12	23	35	20	10

39.4 AVERAGE OR MEASURES OF CENTRAL TENDENCY

An average is a value which is representative of a set of data. Average value may also be termed as measures of central tendency. There are five types of averages in common.

(*i*) Arithmetic average or mean (*ii*) Median (*iii*) Mode

(*iv*) Geometric Mean (*v*) Harmonic Mean
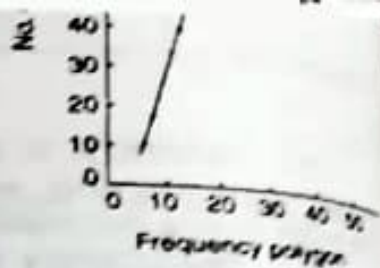
39.5 ARITHMETIC MEAN If $x_1, x_2, x_3, \ldots\ldots x_n$ are n numbers, then their arithmetic mean (A.M.) is defined by

$$A.M. = \frac{x_1 + x_2 + x_3 + \ldots\ldots + x_n}{n} = \frac{\Sigma x}{n}$$

If the number x_1 occurs f_1 times, x_2 occurs f_2 times and so on, then

$$A.M. = \frac{f_1 x_1 + f_2 x_2 + \ldots\ldots + f_n x_n}{f_1 + f_2 + \ldots\ldots + f_n} = \frac{\Sigma f x}{\Sigma f}$$

This is known as direct method.

Example 1. *Find the mean of 20, 22, 25, 28, 30.*

Solution. $A.M. = \dfrac{20 + 22 + 25 + 28 + 30}{5} = \dfrac{125}{5} = 25$ Ans.

Example 2. *Find the mean of the following :*

Numbers	8	10	15	20
Frequency	5	8	8	4

Solution.

$$\Sigma fx = 8 \times 5 + 10 \times 8 + 15 \times 8 + 20 \times 4$$
$$= 40 + 80 + 120 + 80 = 320$$
$$\Sigma f = 5 + 8 + 8 + 4 = 25$$
$$A.M. = \frac{\Sigma fx}{\Sigma f} = \frac{320}{25} = 12.8.$$

(b) Short cut method

Let a be the assumed mean, d the deviation of the variate x from a. Then

$$\frac{\Sigma fd}{\Sigma f} = \frac{\Sigma f(x-a)}{\Sigma f} = \frac{\Sigma fx}{\Sigma f} - \frac{\Sigma fa}{\Sigma f} = A.M. - \frac{a\Sigma f}{\Sigma f}$$

$$= A.M. - a$$

$$\therefore \quad A.M. = a + \frac{\Sigma fd}{\Sigma f}$$

Example 3. *Find the arithmetic mean for the following distribution :*

Class	0–10	10–20	20–30	30–40	40–50
Frequency	7	8	20	10	5

Solution. Let assumed mean $(a) = 25$.

Class	Mid-value x	Frequency f	$x - 25 = d$	fd
0–10	5	7	– 20	– 140
10–20	15	8	– 10	– 80
20–30	25	20	0	0
30–40	35	10	+ 10	+ 100
40–50	45	5	+ 20	+ 100
Total		50		– 20

$$A.M. = a + \frac{\Sigma fd}{\Sigma f} = 25 + \frac{-20}{50} = 24.6 \qquad \text{Ans.}$$

(c) Step deviation method

Let a be the assumed mean, i the width of the class interval and

$$D = \frac{x-a}{i}, \quad A.M. = a + \frac{\Sigma fD}{\Sigma f} \cdot i$$

Example 4. *Find the arithmetic mean of the data given in example 3 by step deviation method.*

Solution. $a = 25$

Class	Mid-value x	frequency f	$D = \frac{x-a}{i}$	$f.D$
0–10	5	7	– 2	– 14
10–20	15	8	– 1	– 8
20–30	25	20	0	0
30–40	35	10	+ 1	+ 10
40–50	45	5	+ 2	+ 10
Total		50		– 2

$$A.M. = a + \frac{\Sigma fD}{\Sigma f} \cdot i = 25 + \frac{-2}{50} \times 10 = 24.6 \qquad \text{Ans.}$$

39.6 MEDIAN

Median is defined as the measure of the central item when they are arranged in ascending or descending order of magnitude.

When the total number of the items is odd and equal to say n, then the value of $\frac{1}{2}(n+1)$th item gives the median.

When the total number of the frequencies is even, say n, then there are two middle items, and so the mean of the values of $\frac{1}{2}n$th and $\left(\frac{1}{2}n+1\right)$th items is the median.

Example 5. *Find the median of 6, 8, 9, 10, 11, 12, 13.*

Solution. Total number of items = 7

$$\text{The middle item} = \frac{1}{2}(7+1)^{th} = 4th$$

$$\therefore \quad \text{Median} = \text{Value of the 4th item} = 10 \qquad \textbf{Ans.}$$

For grouped data,

$$\text{Median} = l + \frac{\frac{1}{2}N - F}{f} \cdot i$$

where l is the lower limit of the median class, f is the frequency of the class, i is the width of the class-interval, F is the total of all the preceeding frequencies of the median-class and N is total frequency of the data.

Example 6. Find the value of Median from the following data:

No. of days for which absent (less than)	5	10	15	20	25	30	35	40	45
No. of students	29	224	465	582	634	644	650	653	655

Solution. The given cumulative frequency distribution will first be converted into ordinary frequency as under

Class Interval	Cumulative frequency	Ordinary frequency
0-5	29	29 = 29
5-10	224	224 - 29 = 195
10-15	465	465 - 224 = 241
15-20	582	582 - 465 = 117
20-25	634	634 - 582 = 52
25-30	644	644 - 634 = 10
30-35	650	650 - 644 = 6
35-40	653	653 - 650 = 3
40-45	655	655 - 653 = 2

$\text{Median} = \text{size of } \dfrac{655}{2} \text{ or } 327.5\text{th item}$

327.5th item lies in 10–15 which is the median class.

$$M = l + \frac{\frac{N}{2} - C}{f} i$$

where l stands for lower limit of median class,

    N stands for the total frequency,

    C stands for the cumulative frequency just proceeding the median class,

    i stands for class interval

    f stands for frequency for the median class.

$$\text{Median} = 10 + \frac{\frac{655}{2} - 224}{241} \times 5$$

$$= 10 + \frac{103.5 \times 5}{241} = 10 + 2.15 = 12.15$$

39.7 MODE

Mode is defined to be the size of the variable which occurs most frequently.