STATISTICS

TOPIC 1

STATISTICS

Statistics is the discipline that concerns the collection, organisation, analysis, interpretation and presentation of data.

In the previous chapter, we have already studied the three measures of central tendency, *i.e.*, mean, median and mode to find the average middle term and most frequency term, respectively, of an ungrouped data. Now, in this chapter, we will study mean, median and mode of grouped data. Further, we will extend our discussion to concept of cumulative frequency and drawing its ogives.

Here, we will discuss about the mean, median and mode of the ungrouped and grouped data.

Ungrouped Data

A data that is obtained in original form, *i.e.* a list of numbers is called an ungrouped data.

Grouped Data

A data that is organised into several groups is called grouped data.

A data can be grouped in two forms:

(1) Exclusive Form (Overlapping/Continuous Form) : In this type of grouped data, the upper limit is excluded from each class interval and the class intervals are made in such a manner that the upper limit of one class is the lower limit of the next class.

For example, 0 – 10 and 10 – 20. Here, 10 is included in 10 – 20 and with 0 – 10.

(2) Inclusive Form (Non Overlapping/Discontinuous Form): In this type of grouped data, the upper limit is also included in the class itself and so, it is not included in the next higher class.

For example 0 – 19 and 20 – 29.

Important

➥ *To convert inclusive form to exclusive form, we substract* $\frac{h}{2}$ *from the lower and add* $\frac{h}{2}$ *to the upper limit of each class interval where, h is the difference between upper of a class and lower limit of the preceding class.*

➥ *In the exclusive form, the lower limit and the upper limit are called the True Lower Limit and the True Upper Limit respectively.*

➥ *The difference between the true upper limit and true lower limit of a grouped frequency distribution is called the class-size or class width. It is denoted by h.*

➥ *The middle-value of a class-interval which lies between true lower limit and true upper limit is called the class mark. It is denoted by x.*

TOPIC 2

MEAN

The mean of a data, or a set of observation is the average value of the data, or observation, and is denoted by $\bar{x}$, read as *x* bar. Generally, it is obtained by dividing the sum of all the observations by the total number of the observations.

Mean of Ungrouped Data

Then, Mean $(\bar{x}) = \frac{x_1 + x_2 + \ldots + x_n}{x}$

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

If x_1, x_2, x_n are the n observation with frequencies $f_1, f_2, \ldots, f_n$ respectively i.e. the observation *x*, is repated f_1 time, x_2 is repeated f_2 times and so on, then

sum of all the observations

$$= f_1x_1 + f_2x_2 + \ldots + f_n x_n$$

$$= \sum_{i-1}^{n} f_i x_i$$

Also, Total number of observations

$$= f_1 + f_2 + \ldots + f_n$$

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

$$\therefore \text{Mean } (\bar{x}) = \frac{\sum_{i=1}^{n} f x_i}{\sum_{i=1}^{n} f}$$

Important

➥ *If each observation of a given set of data is increased by number* p, *then mean of the data will also be increased by p.*

➥ *If each observation of a given set of data is decreased by a number p then mean of the data will also be decreased by p.*

➥ *If each observation of given set of data is multiplied by a number k, then mean of the data will also be multiplied by k.*

➥ *If each observation of a given set of data is divided by a number k provided* $k \neq (0)$, *then mean of the data will also be divided be k.*

Mean of Grouped Data

To calculate the mean of grouped data, following three methods can be used.

(1) Direct method
(2) Assumed mean method
(3) Step deviation method

Direct Method

To find the mean of grouped data by direct method, following stpes are followed.

Step 1. Find the class mark of each class interval by the formula.

$$\text{class mark} = \frac{\text{Upper class limit} + \text{lower class limit}}{2}$$

It is represented by x_i.

Step 2. Multiply x_i with the corresponding frequency f_i.

Step 3. Calculate the sum of all $f_i x_i$ to get $\Sigma f_i x_i$.

Step 4. Calculate the sum of all the frequencies to get Σf_i.

$$\text{Thus, mean } (\bar{x}) = \frac{\Sigma f_i x_i}{\Sigma f_i}$$

Assumed Mean Method

When the numerical values of x_i and f_i are large and computation of product of x_i and f_i is time consuming then we use this method.

To calculate mean of a grouped data by assumed mean method following steps are followed.

Step 1. Find the class mark (x_i) of each class interval.

Step 2. Among all x_i's, choose an assumed mean and denote it by 'a'.

Step 3. Calculate the difference (d_i) between a and each of the x_i's i.e. $d_i = x_i - a$

Step 4. Multiply d_i with the corresponding f_i.

Step 5. Calculate $\Sigma f_i d_i$ and Σf_i.

$$\text{Thus, Mean } (\bar{x}) = a + \frac{\Sigma f_i d_i}{\Sigma f_i}$$

Important

➥ *Try to take the value of assumed mean i.e. a such that it lies in the centre of $x_1, x_2, x_3, \ldots x_n$. This will make calculation easier.*

MOST LIKELY Questions

Very Short Answer Type Questions

[1 mark]

1. Find the class marks of the classes 20-50 and 35-60.

Ans. Class mark of 20 – 50 = $\frac{20+50}{2}$, *i.e.*, 35

Class mark of 35 – 60 = $\frac{35+60}{2}$, *i.e.*, 47.5

2. The mean of 20 observations is 12. If each observation is increased by 5, then find the new mean.

Ans. Given: Mean of 20 observations = 12

Here,each observation is increased by 5.

So, new mean will also increased by 5.

Thus, new mean = 12 + 5 = 17.

3. What is the arithemetic of first n natural numbers?

Ans. Arithmetic mean = $\frac{n(n+1)}{2n}$ *i.e.*, $\frac{n+1}{2}$

Explanation: First n natural numbers are :

1, 2, 3, 4, 5, ..., n

We know,

$$\text{Mean} = \frac{\text{Sum of all the observations}}{\text{Total number of observation}}$$

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

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

$$\left[\because 1+2+3+\ldots+n = n\left(\frac{n+1}{2}\right)\right]$$

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

Hence the arithmetic mean of first n natural numbers is $\frac{n+1}{2}$.

4. If the mean of 5 observations x, $x + 2$, $x + 4$, $x + 6$ and $x + 8$ is 11, then find the value of x.

Ans. We know,

$$\text{Mean} = \frac{\text{Sum of observations}}{\text{Total number of observations}}$$

$$\Rightarrow 11 = \frac{x+(x+2)+(x+4)+(x+6)+(x+8)}{5}$$

$$\Rightarrow 55 = 5x + 20$$

$$\Rightarrow 5x = 35$$

$$\Rightarrow x = 7$$

Hence the value of x is 7.

Short Answer Type-I Questions (SA-I)

[2 marks]

5. Find the mean of the following distribution:

Class	1-3	3-5	5-7	7-10
Frequency	9	22	27	17

Ans. We first find the class marks (x_i) of each class

Class	Class mark (x_i)	Frequency (f_i)	f_ix_i
1-3	2	9	18
3-5	4	22	88
5-7	6	27	162
7-10	8.5	17	144.5
		$\Sigma f_i = 75$	$\Sigma f_ix_i = 412.5$

Thus, Mean, $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{412.5}{75} = 5.5$

Hence, the mean of the given distribution is 5.5.

6. Find the value of p, if the mean of the following distribution is 7.5.

Classes	2–4	4–6	6–8	8–10	10–12	12–14
Frequency	6	8	15	p	8	4

Ans.

Classses	2-4	4-6	6-8	8-10	10-12	12-14
Class – mark (x_i)	3	5	7	9	11	13

Frequency (f_i)	6	8	15	p	8	4
f_ix_i	18	40	105	$9p$	88	52

Here, $\Sigma f_i = 41 + p$

and $\Sigma f_ix_i = 303 + 9p$

We know, Mean $= \frac{\Sigma f_i x_i}{\Sigma f_i}$

But, Mean $= 7.5$ (Given)

$\therefore \quad 7.5 = \frac{303+9p}{41+p}$

$\Rightarrow \quad 303 + 9p = 307.5 + 7.5\,p$

$\Rightarrow \quad 1.5\,p = 4.5$

$\Rightarrow \quad p = 3$

Hence, the value of p is 3.

Short Answer Type-II Questions (SA-II)

[3 marks]

7. The following table gives the number of pages written by Sarika for completing her own book for 30 days:

Number of pages written per day	16-18	19-21	22-24	25-27	28-30
Number of days	1	3	4	9	13

Calculate the average number of pages written in 30 days.

Ans. We need to convert the data into continuous classes by subtracting 0.5 from the lower limit and adding 0.5 to the upper limit of each class.

Class	Class mark (x_i)	Number of days (f_i)	f_ix_i
15.5-18.5	17	1	17
18.5-21.5	20	3	60
21.5-24.5	23	4	92
24.5-27.5	26	9	234
27.5-30.5	29	13	377
		$\Sigma f_i = 30$	$\Sigma f_ix_i = 780$

$\therefore \quad$ Mean, $\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{780}{30} = 26.$

Hence, she wrote an average of 26 pages per day.

8. **The daily income of a sample of 50 employees are tabulated as follows:**

Income (in ₹)	1-200	201-400	401-600	601-800
Number of employees	14	15	14	7

Find the mean daily income of the employees.

Ans. The given data is not continuous, so we need to convert the data into continuous form, by subtracting 0.5 from the lower limit and adding 0.5 to the upper limit of each class:

Income	Class mark (x_i)	$d_i = (x_i - a)$	f_i	$f_i d_i$
0.5-200.5	100.5	−200	14	−2800
200.5-400.5	300.5 = a	0	15	0
400.5-600.5	500.5	200	14	2800
600.5-800.5	700.5	400	7	2800
			$\Sigma f_i = 50$	$\Sigma f_i d_i = 2800$

Here, a = assumed mean = 300.5

We know,

$$\text{Mean } (\bar{x}) = a + \frac{\Sigma f_i d_i}{\Sigma f_i}$$

$$= 300.5 + \frac{2800}{50}$$

$$= 300.5 + 56$$

$$\bar{x} = 356.5$$

Thus, the mean daily income of the employees is ₹ 356.50.

9. **An aircraft has 120 passenger seats. The number of seats occupied during 100 flights is given in the following table:**

Numbe of seats	100-104	104-108	108-112	112-116	116-120
Frequency	15	20	32	18	15

Determine the mean number of seats occupied during the flights.

Ans.

Class Interval (No. of seats)	Class mark (x_i)	Deviation $d_i = (x_i - a)$	Frequency (f_i)	$f_i d_i$
100-104	102	−8	15	−120
104-108	106	−4	20	−80
108-112	a = 110	0	32	0
112-116	114	4	18	72
116-120	118	8	15	120
			$\Sigma f_i = 100$	$\Sigma f_i d_i = -8$

Here, assumed mean, a = 110

$$\therefore \quad (\bar{x}) = a + \frac{\Sigma f_i d_i}{\Sigma f_i}$$

$$= 110 + \frac{(-8)}{100} = 110 - 0.08$$

$$= 109.92$$

Thus, the mean number of seats occupied over the flights is 109.92 or 110 approx.

10. **The mean of the following frequency distribution is 62.8 and the sum of all the frequencies is 50. Compute the missing frequencies f_1 and f_2.**

Classes	0-20	20-40	40-60	60-80	80-100	100-120
Frequency	5	f_1	10	f_2	7	8

Ans. The frequency distribution for calculating the mean, for the given data is:

Classes	Frequency (f_i)	Class mark (x_i)	$d_i = x_i - a$	$f_i d_i$
0–20	5	10	–40	–200
20–40	f_1	30	–20	$-20f_1$
40–60	10	$50 = a$	0	0
60–80	f_2	70	20	$20f_2$
80–100	7	90	40	280
100–120	8	110	60	480
	$\Sigma f_i = 30 + f_1 + f_2$			$\Sigma f_i d_i = 560 - 20f_1 + 20f_2$

Since, sum of all the frequencies is 50.

$\therefore \quad 30 + f_1 + f_2 = 50$

$\Rightarrow \quad f_1 + f_2 = 20$

$\Rightarrow \quad f_1 = 20 - f_2 \quad ...(i)$

Now, mean = 62.8

We know,

$$\text{Mean} = a + \left(\frac{\Sigma f_i d_i}{\Sigma f_i}\right)$$

$$\Rightarrow \quad 62.8 = 50 + \frac{560 - 20f_1 + 20f_2}{50}$$

$$\Rightarrow \quad 12.8 = \frac{560 - 20f_1 + 20f_2}{50}$$

$\Rightarrow \quad 640 = 560 - 20f_1 + 20f_2$ [Using (i)]

$\Rightarrow \quad 80 = -20f_1 + 20f_2$

$\Rightarrow \quad 4 = f_1 + f_2$

$\Rightarrow \quad 4 = 20 + f_2 + f_2$ [From (i)]

So, using (i), $f_1 = 20 - 12 = 8$

$\therefore f_1 = 8$ and $f_2 = 12$

11. Find the mean marks of the students for the following distribution:

Marks	Number of Students
0 and above	80
10 and above	77
20 and above	72
30 and above	65
40 and above	55
50 and above	43
60 and above	28
70 and above	16
80 and above	10
90 and above	8
100 and above	0

Ans.

Marks	Class mark (x_i)	Number of students (Cumulative frequency)	Frequency (f_i)	$f_i x_i$
0-10	5	80	80 – 77 = 3	15
10-20	15	77	77 – 72 = 5	75
20-30	25	72	72 – 65 = 7	175
30-40	35	65	65 – 55 = 10	350
40-50	45	55	55 – 43 = 12	540
50-60	55	43	43 – 28 = 15	825
60-70	65	28	28 – 16 = 12	780
70-80	75	16	16 – 10 = 6	450
80-90	85	10	10 – 8 = 2	170
90-100	95	8	8 – 0 = 8	760
100-110	105	0	0	0
			$\Sigma f_i = 80$	$\Sigma f_i x_i = 4140$

$$\therefore \text{Mean} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{4140}{80} = 51.7$$

TOPIC 3

MODE

The observation which occurs most frequently among the given set of observations *i.e.* the value of the observation having the maximum frequency is the mode of the given data.

Sometimes it is possible that in a given set of data, more than one value may have the same maximum frequency. Such type of data in said to be multimodal.

Mode of Ungrouped Data

An observation with the maximum frequency is the mode of the given set of ungrouped data.

Mode of Grouped Data

To find the mode of a grouped data, modal class is determined which is defined as the class with the maximum frequency. The mode lies in the modal class.

Mode of a grouped data is given by the formula:

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times h$$

Where, l = Lower limit of the modal class,

h = Size of class intervals,

f_i = Frequency of the modal class,

f_0 = Frequency of the class preceding the modal class, and

f_2 = Frequency of the class succeeding the modal class.

Important

➥ *Mode can also be calculated for grouped data for unequal class sizes.*

12. Find the mode of the following frequency distribution:

Class	15-20	20-25	25-30	30-35	35-40	40-45
Frequency	3	8	9	10	3	2

Ans. The modal class is 30-35, as this class has the maximum frequency of 10.

For this class,

$l = 30$, $h = 5$, $f_1 = 10$, $f_0 = 9$ and $f_2 = 3$.

We know, Mode $= l + \dfrac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$

$= 30 + \dfrac{10 - 9}{20 - 9 - 3} \times 5$

$= 30 + \dfrac{5}{8}$

$= 30 + 0.625$

$= 30.625$

13. Find the mode of the following data:

Class	0-20	20-40	40-60	60-80	80-100	100-120	120-140
Frequency	6	8	10	12	6	5	3

Ans. As the maximum frequency is 12 for the class 60-80.

So, the modal class is 60-80

For this class,

$l = 60$, $h = 20$, $f_1 = 12$, $f_0 = 10$ and $f_2 = 6$.

We know, Mode $= l + \dfrac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$

$= 60 + \dfrac{12 - 10}{24 - 10 - 6} \times 20$

$= 60 + \dfrac{40}{8}$

$= 60 + 5$

$= 65$

Thus, the mode is 65.

14. The following table gives the number of participants in a yoga camp:

Age (in years)	20-30	30-40	40-50	50-60	60-70
No. of Participants	8	40	58	90	83

Find the modal age of the participants.

Ans. In the given frequency distribution, the modal class is 50-60 having maximum frequency 90.

So, $l = 50$, $f_1 = 90$, $f_0 = 58$, $f_2 = 83$ and $h = 10$,

We know,

Mode, $= l + \left(\dfrac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \times h$

$= 50 + \left(\dfrac{90 - 58}{180 - 58 - 83}\right) \times 10$

$= 50 + \dfrac{32}{180 - 141} \times 10$

$= 50 + \dfrac{32}{39} \times 10$

$= 50 + 8.2$

$= 58.2$

Hence, the modal age of the participants is 58.2 years.

15. The mode of the following data is 67. Find the missing frequency x.

Class	40-50	50-60	60-70	70-80	80-90
Frequency	5	x	15	12	7

Ans. Mode of given data, $M_0 = 67$

Then, modal class is 60 – 70

We know, $M_0 = l + \left(\frac{f_1 - f_0}{2f_1 - f_2 - f_0}\right) \times h$

Here, $l = 60$, $f_1 = 15$, $f_0 = x$; $f_2 = 12$ and $h = 10$

Then, $67 = 60 + \left(\frac{15 - x}{30 - 12 - x}\right) \times 10$

$\Rightarrow \quad 7(18 - x) = 150 - 10x$

$\Rightarrow \quad 126 - 7x = 150 - 10x$

$\Rightarrow \quad 3x = 150 - 126$

$\Rightarrow \quad x = \frac{24}{3} = 8$

Hence, the value of x is 8.

16. Daily wages of 110 workers, obtained in a survey, are tabulated below:

Daily Wages (in ₹)	100-120	120-140	140-160	160-180	180-200	200-220	220-240
Number of Workers	10	15	20	22	18	12	13

Compute the mean daily wages and modal daily wages of these workers.

Ans.

Daily Wages (in ₹)	Number of Workers (f_i)	x_i	$d_i = x_i - a$	$f_i d_i$
100-120	10	110	–60	–600
120-140	15	130	–40	–600
140-160	20	150	–20	–400
160-180	22	$170 = a$	0	0
180-200	18	190	20	360
200-220	12	210	40	480
220-240	13	230	60	780
Total	110			20

Mean daily wages $= 170 + \frac{1}{110} \times 20/110 =$ ₹ 170.19 (approx.)

Mode $= 160 + \frac{22 - 20}{44 - 20 - 18} \times 20 =$ ₹ 166.67 (approx.)

TOPIC 4

MEDIAN

Median is defined as the middle most or the central observation, when the given set of observations are arranged in an ascending order of their magnitudes.

Median of Ungrouped Data

Let n be the total number of observations arranged in an ascending order.

Case I : If 'n' is odd, then

$$\text{Median} = \text{value of } \frac{(n+1)}{2} \text{ observations.}$$

Case II : If 'n' is even, then

$$\text{Median} = \text{Mean value of } \left(\frac{n}{2}\right)^{\text{th}} \text{ and } \left(\frac{n}{2} + 1\right)^{\text{th}} \text{ observations.}$$

Median of Grouped Data

To find the median of a grouped data, we need to determine median class and cumulative frequencies of all class intervals.

Before heading to calculation of median, let us discuss median class and cumulative frequency.

Cumulative Frequency

Cumulative frequency of a class interval is determined by adding all the frequencies upto that class interval.

Median Class

A class having cumulative frequency just greater than half of sum of all the frequencies, is called median class.

Thus, to find the median class, we need to make a cumulative frequency column in the distribution table.

Now, the median of a grouped data is computed, using the formula:

$$\text{Median} = l + \frac{\left(\frac{n}{2} - cf\right) \times h}{f}$$

Where, l = Lower limit of the median class,

n = Total number of observations,

f = Frequency of the median class,

cf = Frequency of the class preceding the median class, and

h = Size of the median class.

MOST LIKELY Questions

Short Answer Type-II Questions (SA-II)

[3 marks]

17. From the following distribution, find the median :

Classes	500 – 600	600 – 700	700 – 800	800 – 900	900 – 1000
Frequency	36	32	32	20	30

Ans.

Classes	500 – 600	600 – 700	700 – 800	800 – 900	900 – 1000
Frequency	36	32	32	20	30
Cumulative Frequency	36	68	100	120	150

Here, $\frac{N}{2} = \frac{150}{2} = 75$

Cumulative frequency just greater than 75 is 100, which belongs to class 700 - 800.

So, the median class is 700 – 800

For this class,

$l = 700$, $h = 100$, $\frac{N}{2} = 75$, $cf = 68$, $f = 32$.

So, $\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$

$= 700 + \frac{75 - 68}{32} \times 100$

$= 700 + \frac{700}{32}$

$= 700 + 21.875 = 721.875$

18. The distribution given below shows the number of wickets taken by bowlers in one-day cricket matches. Find the mean and the median for the numbers of wickets taken.

Number of wickets	20 – 60	60 – 100	100 – 140	140 – 180	180 – 220	220 – 260
Number of Persons	7	5	16	12	2	3

Ans. Calculation of Mean:

Number of wickets (class)	Class-mark (x_i)	Frequency (f_i)	$d_i = x_i - a$ where $a = 120$	$f_i d_i$
20-60	40	7	-80	-560
60-100	80	5	-40	-200
100-140	120 = a	16	0	0
140-180	160	12	40	480
180-220	200	2	80	160
220-260	240	3	120	360
		$\Sigma f_i = 45$		$\Sigma f_i u_i = 240$

We know, Mean $= a + \frac{\Sigma f_i u_i}{\Sigma f_i}$

$= 120 + \frac{240}{45}$

$= 120 + 5.33 = 125.33$

Calculation of Median:

Classes	Frequency	Cumulative Frequency
20 – 60	7	7
60 – 100	5	12
100 – 140	16	28
140 – 180	12	40
180 – 220	2	42
220 – 260	3	45

Here, N = 45, $\Rightarrow \frac{N}{2} = 22.5$

Cumulative frequency just greater than 22.5 is 28, which belongs to class 100-140.

So, median class is 100 – 140

For this class,

$l = 100, h = 40, cf = 12, \frac{N}{2} = 22.5, f = 16$

So, Median $= l + \left(\frac{\frac{N}{2} - cf}{f}\right) \times h$

$= 100 + \left(\frac{22.5 - 12}{16}\right) \times 40$

$= 100 + 26.25$

$= 126.25$

Hence, the mean number of wickets taken is 125.33 and the median number of wickets taken is 126.25.

19. The median of the following data is 16. Find the missing frequencies *a* and *b*, if the total of the frequencies is 70.

Class	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	12	*a*	12	15	*b*	6	6	4

Ans.

Class	Frequency	Cumulative frequency
0-5	12	12
5-10	*a*	12 + *a*
10-15	12	24 + *a*
15-20	15	39 + *a*
20-25	b	39 + *a* + *b*
25-30	6	45 + *a* + *b*
30-35	6	51 + *a* + *b*
35-40	4	55 + *a* + *b*
Total	70	

$\text{median} = l + \frac{\frac{N}{2} - cf}{f} \times h$

$16 = 15 + \frac{35 - 24 - a}{15} \times 5$

$1 = \frac{(11 - a)}{3}$

$a = 8$

$55 + a + b = 70$

$55 + 8 + b = 70$

$b = 7$

$55 + a + b = 70$

$a + b = 15$

Class	Frequency	Cumulative frequency
0-5	12	12
5-10	*a*	12 + *a*
10-15	12	24 + *a*
15-20	15	39 + *a*
20-25	*b*	39 + *a* + *b*
25-30	6	45 + *a* + *b*
30-35	6	51 + *a* + *b*
35-40	4	55 + *a* + *b*
Total	70	

Then, $55 + a + b = 70$

$a + b = 15$ ...(i)

Given, median = 16

Then median class is 15 – 20

$\therefore$ $\text{Median} = l + \frac{\frac{N}{2} - cf}{f} \times h$

Here, $l = 15$

$h = 5$

$cf = 24 + a$

$f = 15$

$N = 70$

$\therefore$ $16 = 15 + \frac{35 - 24 - a}{15} \times 5$

$\Rightarrow$ $1 = \frac{11 - a}{15} \times 5$

$\Rightarrow$ $3 = 11 - a$

$\Rightarrow$ $a = 8$

Put the value 'a' in equation (I), we get

$$b = 15 - 8 = 7$$

Hence, the values of 'a' and 'b' are 8 and 7.

20. If the median of the following frequency distribution is 32.5. Find the values of f_1 and f_2.

Class	Frequency
0-10	f_1
10-20	5
20-30	9
30-40	12
40-50	f_2
50-60	3
60-70	2
Total	40

Ans.

Class	Frequency	Cumulative frequency
0–10	f_1	f_1
10–20	5	$5 + f_1$
20–30	9	$14 + f_1$
30–40	12	$26 + f_1$
40–50	f_2	$26 + f_1 + f_2$
50–60	3	$29 + f_1 + f_2$
60–70	2	$31 + f_1 + f_2$
Total	40	

Given: Median = 32.5

Then, the median class is 30–40.

For this class,

$l = 30$, $f = 12$, $cf = 14 + f_1$, $\frac{N}{2} = 20$ and $h = 10$.

We know,

$$\text{Median} = l + \left(\frac{\frac{N}{2} - cf}{f}\right) \times h$$

$$\Rightarrow \quad 32.5 = 30 + \left(\frac{20 - (14 + f_1)}{12}\right) \times 10$$

$$\Rightarrow \quad 2.5 = \frac{10}{12}(6 - f_1)$$

$$\Rightarrow \quad \frac{12 \times 2.5}{10} = 6 - f_1$$

$$\Rightarrow \quad f_1 = 6 - 3 = 3$$

Also, $31 + f_1 + f_2 = 40$

$$\Rightarrow \quad f_1 + f_2 = 9$$

$$\Rightarrow \quad f_2 = 9 - f_1$$

$$\Rightarrow \quad = 9 - 3$$

$$= 6$$

Hence, the values of f_1 and f_2 are 3 and 6, respectively.

21. Find the unknown entries *a, b, c, d, e, f* in the following distribution of the heights of the students in a class:

Height (in cm)	Frequency	Cumulative Frequency
150-155	12	*a*
155-160	*b*	25
160-165	10	*c*
165-170	*d*	43
170-175	*e*	48
175-180	2	*f*
Total	50	

Ans.

Height (in cm)	Frequency (f_i)	Cumulative Frequency (Given)	Cumulative frequency (Calculated)
150-155	12	*a*	12
155-160	*b*	25	$12 + b$
160-165	10	*c*	$22 + b$
165-170	*d*	43	$22 + b + d$
170-175	*e*	48	$22 + b + d + e$
175-180	2	*f*	$24 + b + d + e$
Total	50		

Comparing the last two columns,

i.e., cumulative frequency calculated and cumulative frequency given, we get:

$$a = 12$$

Also, $12 + b = 25$

$\Rightarrow b = 25 - 12 = 13$

$22 + b = c$

$\Rightarrow c = 22 + 13 = 35$

$22 + b + d = 43$

$\Rightarrow 22 + 13 + d = 43$

$\Rightarrow d = 43 - 35$

$d = 8$

$22 + b + d + e = 48$

$\Rightarrow 22 + 13 + 8 + e = 48$

$\Rightarrow 43 + e = 48$

$\Rightarrow e = 5$

$24 + b + d + e = f$

$\Rightarrow 24 + 13 + 8 + 5 = f$

$\Rightarrow f = 50$

Hence, $a = 12$, $b = 13$, $c = 35$, $d = 8$, $e = 5$ and $f = 50$

Case Based Question (VSA Type)

[4 & 5 marks]

22. *A survey was conducted by an NGO to know the monthly expenditure of families living in slums in Delhi. A total of 200 families were interviewed and it was found that their minimum monthly expenditure was ₹ 1000. The result is tabulated as given below:*

Monthly expenditure (in ₹)	Number of families
Less than 2000	10
Less than 3000	27
Less than 4000	46
Less than 5000	75
Less than 6000	115
Less than 7000	140
Less than 8000	163
Less than 9000	182
Less than 10000	200

(A) Find the number of families whose monthly expenditure is more than or equal to ₹ 8000.

(B) Find the number of families whose monthly expenditure is in the range ₹ (6000 – 7000).

(C) Find the lower limit of the median class.

(D) Find the median monthly expenditure of the families as per the frequency distribution table.

Ans. (A) We can see from the given frequency distribution table that the number of families having monthly expenditure less than ₹ 8000 is 163 out of a total of 200 families. Therefore, number of families whose monthly expenditure is more than or equal to ₹ 8000 = 200 – 163 = 37.

(B) Let us prepare a frequency distribution table from the given cumulative frequency table as below:

Monthly expenditure (in Rs.)	Number of families	Cumulative frequency
1000 – 2000	10	10
2000 – 3000	27 – 10 = 17	27
3000 – 4000	46 – 27 = 19	46
4000 – 5000	75 – 46 = 29	75
5000 – 6000	115 – 75 = 40	115
6000 – 7000	140 – 115 = 25	140
7000 – 8000	163 – 140 = 23	163
8000 – 9000	182 – 163 = 19	182
9000 - 10000	200 – 182 = 18	200
Total	200	

Therefore, number of families whose monthly expenditure is in the range ₹ (6000 – 7000) is 25.

(C) Median class is the class whose cumulative frequency is just greater than half of sum of all the frequencies. Here, $\frac{N}{2} = 100$. As the cumulative frequency of the class 5000 – 6000 is 115, which is just greater than 100, therefore the median class is 5000 – 6000 and thus the lower limit of the median class is 5000.

(D) The median class is 5000 – 6000. The formula for calculating the median is:

$$\text{Median} = l + \left(\frac{\frac{n}{2} - cf}{f}\right) \times h$$

Where, l = lower limit of the median class,

h = size of the class interval,

n = number of observations,

cf = cumulative frequency of class preceding the median class,

f = frequency of the median class,

Using the table drawn in part (B), we have $l = 5000, h = 1000, cf = 75, f = 40$ and $n = 200$.

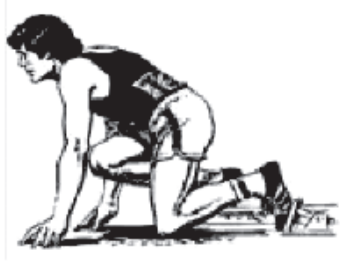
Putting these values in the formula for calculating median, we get

$$\text{Median} = 5000 + \left(\frac{\frac{200}{2} - 75}{40}\right) \times 1000$$

$$= 5000 + \left(\frac{25}{40}\right) \times 1000$$

$$= 5000 + 625 = 5625$$

23. *100m RACE A stopwatch was used to find the time that it took a group of students to run 100 m.*



Time (in sec)	No. of students
0 - 20	8
20 - 40	10
40 - 60	13
60 - 80	6
80 - 100	3

(A) Estimate the mean time taken by a student to finish the race.

(B) What will be the upper limit of the modal class ?

(C) Find the sum of lower limits of median class and modal class ?

(D) How many students finished the race within 1 minute?

Ans. (A)

$$\text{Mean} = \frac{\sum f_i x_i}{\sum f_i} = \frac{1720}{40}$$

$$= 43$$

(B) Since, modal class is 40 – 60, with maximum frequency 13.

Then upper limit is 60.

(C)

Time (In sec)	No. of Students (f)	Cumulative frequency (cf)
0 – 20	8	8
20 – 40	10	18
40 – 60	13	31
60 – 80	6	37
80 – 100	3	40

$$\frac{N}{2} = \frac{40}{2} = 20$$

Class corresponding to cumulative frequency just greater than 20 is 40 - 60.

So, Median class is 40 – 60.

$\therefore$ Lower limit of median class = 40

Also, lower limit of modal class = 40

[From part (B)]

$\therefore$ Required sum = 40 + 40 = 80

(D) 1 min = 60 sec

Then, number of students who finished race within 1 minute = 8 + 10 + 13

= 31

TOPIC 5

RELATIONSHIP BETWEEN MEAN, MEDIAN AND MODE

There is an empirical reltaiton between the measures of central tendency which is given by.

Mode = 3 Median – 2 Mean

or, $\text{Mean} = \frac{3\,\text{Median} - \text{Mode}}{2}$

or, $\text{Median} = \frac{\text{Mode} + 2\,\text{Mean}}{3}$

MOST LIKELY Questions

Very Short Answer Type Questions
[1 mark]

24. The mean and median of a distribution are both equal to 635.97. Find the mode.

Ans. Mean= Median = 635.97 [Given]

We know,

Mode = 3 Median – 2 Mean

= 3 Mean – 2 Mean

[∵ Mean = Median]

= Mean

= 635.97.

25. Write the empirical relationship among the three measures of central tendency mean, mode and median.

Ans. The required relationship is:

3 Median = Mode + 2Mean

26. Using the empirical formula, find the mode of a distribution whose mean is 8.32 and the median is 8.05.

Ans. The empirical formula is:

3 Median = Mode + 2 Mean

So, Mode = 3 Median – 2 Mean

= 3 × 8.05 – 2 × 8.32

= 24.15 – 16.64

= 7.51

Hence, the mode of the distribution is 7.51.

Short Answer Type-I Questions (SA-I)
[2 marks]

27. If X, M and Z are denoting mean, median and mode of a data and X : M = 9 : 8, then the ratio M : Z is?

Ans. We know,

Mode = 3 Median – 2 Mean

$\Rightarrow$ Z = 3M – 2X ...(i)

Given X : M = 9 : 8

$\Rightarrow \frac{X}{M} = \frac{9}{8}$

$\Rightarrow X = \frac{9M}{8}$

On putting the value of X in eq (i), we get

$Z = 3M - 2 \times \frac{9M}{8}$

$= 3M - \frac{9M}{4}$

$= \frac{3M}{4}$

$\frac{M}{Z} = \frac{4}{3}$

M : Z = 4 : 3

Hence, the required probability is $\frac{13}{18}$.