

6

STATISTICS

6.1 Introduction

6.2 Mathematical

6.3 Arithmetic mean of Grouped data

6.4 Arithmetic mean for continuous frequency distribution

6.5 Mode of Grouped Data

6.6 Empirical

6.7 Graphical Representation

6.1 INTRODUCTION

The word 'statistics' appears to have derived from the latin word 'status' meaning 'a (political) state'. In its origin, statistics was simply the collection of data on different aspects of the life of people, useful to the state.

Statistics deals with collection, organisation, analysis and interpretation of data. The word 'statistics' has different meanings in different contexts.

In the second sentence, the word 'statistics' is used as a singular noun, meaning the subject analysis of data as well as drawing of meaningful conclusions from the data.

In this chapter we shall extend the study of these three measures, i.e. mean, median and mode from ungrouped dates to that of grouped data and also we shall discuss the concept of cumulative frequency and cumulative frequency distribution.

Statistics: The word statistics is used in two senses singular & plural.

"Statistics is the science which deal with collection, analysis & interrelation of numerical data".

Data : Collection of facts & figures.

Primary data : Data collected by investigator himself are called primary data. eg. notes, lists.

Secondary data : When investigator doesnot himself collect the data but he collects the data from other sources. eg. Published reports official statistics collected by the Government on various facts.

Frequency Distribution : A tabular arrangement of data by classes together with the corresponding frequencies is called a frequency distribution or a frequency table. e.g.

Marks	35-45	45-55	55-65	65-75	75-85	85-95
Frequency	3	4	8	4	5	1

Frequency : The number of observations in each class is called frequency of that class. In Table the frequency of class 55-65 is 8 and that of the class 85-95 is 1.

Class-Intervals and Class Limits : In the frequency Table 35-45 is called “class-interval” and the end numbers, 35 and 45 are called “class limits”, the smaller number 35 is the lower class limit and the larger number 45 is the upper class limit.

Class Boundaries : In the above table 34.5, 44.5 are called class boundaries or true class limits.

The size or width of a Class Interval : The size or width of a class-interval is the difference between the lower and upper class boundaries.

e.g., size = $44.5 - 34.5 = 10$

Class Mark : The class-mark is the mid-point of the class-interval. e.g. the class mark for the interval

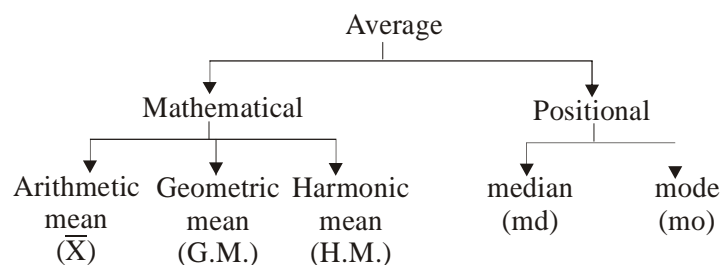
$$35-45 \text{ is } \frac{35 + 45}{2} = \frac{80}{2} = 40.$$

Cumulative Frequency Table : The total of frequencies of all the previous and the given class is called the cumulative frequency of the class e.g.

Class-intervals	Frequency	Cumulative frequency
35-45	3	3
45-55	4	7
55-65	8	15
65-75	4	19
75-85	5	24
85-95	1	25
Total	25	

Central tendency : A no. or quantity which is typical or representative of a set of data is called central tendency measure of this kind is known as averages.

Measures of central tendency or average are usually of the following types



pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

6.2 MATHEMATICAL

- (a) Mean (Arithmetic mean of individual observations) or ungrouped data
 (i) Mean of 'n' numbers $x_1, x_2, x_3, \dots, x_n$ is given by

$$\bar{X} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

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

Illustration 1

It the heights of 5 persons are 144 cm, 152 cm, 151 cm, 158 cm and 155 cm respectively find the mean height.

Solution

$$\text{Mean height} = \frac{144+152+151+158+155}{5} = \frac{760}{5} \text{ cm} = 152 \text{ cm}$$

6.3 ARITHMETIC MEAN OF GROUPED DATA OR DISCRETE FREQUENCY DISTRIBUTION

- (i) Direct method
 (ii) Short cut method
 (iii) Step deviation method

6.3.1 Direct method

Mean of 'n' observations $x_1, x_2, x_3, x_4, \dots, x_n$ with frequencies $f_1, f_2, f_3, \dots, f_n$ is given by

$$\bar{X} = \frac{f_1 x_1 + f_2 x_2 + f_3 x_3 + \dots + f_n x_n}{f_1 + f_2 + f_3 + \dots + f_n}$$

$$\bar{X} = \frac{\sum_{i=1}^n f_i x_i}{N} \quad \text{where}$$

$$N = \sum_{i=1}^n f_i = f_1 + f_2 + \dots + f_n$$

Illustration 2

Find the mean of the following distribution:

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

X:	4	6	9	10	15
f:	5	10	10	7	8

Solution

x_i	f_i	$f_i x_i$
4	5	20
6	10	60
9	10	90
10	7	70
15	8	120
	$N = \sum f_i$ = 40	$\sum f_i x_i = 360$

$$\begin{aligned} \text{Mean} \\ \bar{X} &= \frac{\sum f_i x_i}{\sum f_i} \\ &= \frac{360}{40} = 9 \\ \bar{X} &= 9 \end{aligned}$$

6.3.2 Short cut method of finding mean

Let $x_1, x_2, x_3, \dots, x_n$ be values of a variables 'x' with corresponding frequencies $f_1, f_2, f_3, \dots, f_n$ respectively taking deviations about on arbitrary point 'A'

$$\bar{X} = A + \frac{1}{N} \left(\sum_{i=1}^n f_i d_i \right)$$

$$\text{Where } N = \sum_{i=1}^n f_i$$

The number 'A' is generally known as the assumed mean and the method is also called assumed mean method:

Illustration 3

The following table shows the weights of 12 students:

Weight (kg)	67	70	72	73	75
Number of students	4	3	2	2	1

Find the mean weight

Solution

Let the assumed mean be $A = 72$

Weight (in kg) x_i	Number of students f_i	$d_i = x_i - A$ $= x_i - 72$	$f_i d_i$
67	4	-5	-20
70	3	-2	-6
72	2	0	0
73	2	1	2
75	1	3	3
	$N = \sum f_i = 12$		$\sum f_i d_i = -21$

We have $N = 12$ $\sum f_i d_i = -21$ and $A = 72$

$$\therefore \text{Mean} = A + \frac{1}{N} \sum f_i d_i$$

$$72 + \left(\frac{-21}{12} \right) = 72 - \frac{7}{4} = \frac{288 - 7}{4} = \frac{281}{4}$$

$$= 70.25 \text{ kg}$$

6.3.3 Step deviation method

$$\left[\bar{X} = A + h \left(\frac{1}{N} \sum_{i=1}^n f_i u_i \right) \right]$$

Where h = common number
 A = Assumed mean

Illustration 4

Apply step-deviation method to find the AM of the distribution.

Variate (x) :	5	10	15	20	25	30	35	40	45	50
Frequency (f) :	20	43	75	67	72	45	39	9	8	6

Solution

Let the assumed mean be $A = 25$ & $h = 5$

Variate x_i	Frequency f_i	Deviations $d_i = x_i - 25$	$u_i = \frac{x_i - 25}{5}$	$f_i u_i$
5	20	-20	-4	-80
10	43	-15	-3	-129
15	75	-10	-2	-150
20	67	-5	-1	-67
25	72	0	0	0
30	45	5	1	45
35	39	10	2	78
40	9	15	3	27
45	8	20	4	32
50	6	25	5	30
	$N = \sum f_i$ $= 384$			$\sum f_i u_i = -214$

We have $N = 384$, $A = 25$, $h = 5$, and $\sum f_i u_i = -214$

$$\therefore \text{Mean} = \bar{X} = A + h \left(\frac{1}{N} \sum f_i u_i \right)$$

$$\text{Mean} = 25 + 5 \times \left(\frac{-214}{384} \right)$$

$$25 - 2 : 786$$

$$[\bar{X} = 22 : 214]$$

6.4 ARITHMETIC MEAN OF CONTINUOUS FREQUENCY DISTRIBUTION (FREQUENCY DISTRIBUTION WITH CLASS INTERVAL)

(i) The frequency of each class-interval is centred around its mid-point (or class mark)

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

Illustration 5

Find the mean of the following frequency distribution.

Class Interval	0-10	10-20	20-30	30-40	40-50
Frequency	7	10	15	8	10

Solution

Class – Interval	Mid – Values (x_i)	Frequency f_i	$d_i = x_i - 25$	$u_i = \frac{x_i - 25}{10}$	$f_i u_i$
0-10	5	7	-20	-2	-14
10-20	15	10	-10	-1	-10
20-30	25	15	0	0	0
30-40	35	8	10	1	8
40-50	45	10	20	2	20
		$N = \sum f_i$ $= 50$			$\sum f_i u_i$ $= 4$

We have $A = 25$, $h = 10$, $N = 50$ and

$$\sum f_i u_i = 4$$

$$\text{Mean} = A + h \left(\frac{1}{n} \sum f_i u_i \right)$$

$$= 25 + 10 \times \frac{4}{50}$$

$$= 25.8$$

Note: We can solve this question by assumed mean method.

Mean for an Inclusive Series:

- (i) **Median:** Median is the value of middle item of a series arranged in ascending or descending order of magnitudes.

(i) Median for ungrouped data

(a) If n is odd the median = $\left(\frac{n+1}{2}\right)^{\text{th}}$ term

(b) If n is even, there are two middle terms.

i.e. $\left(\frac{n}{2}\right)^{\text{th}}$ term & $\left(\frac{n}{2} + 1\right)^{\text{th}}$ term

$$\text{median} = \frac{\left(\frac{n}{2}\right)^{\text{th}} \text{ term} + \left(\frac{n}{2} + 1\right)^{\text{th}} \text{ term}}{2}$$

- (ii) In a continuous frequency distribution the value of the median would be in class-interval

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

l = lower limit of median class

n = no. of observations

c.f = cumulative frequency of class preceding the median class.

h = class size (assuming class size to be equal)

Median for Grouped data

Method for finding the median for grouped data

Step-1: For the given frequency distribution, prepare the cumulative frequency table and obtain $N = \sum f_i$

Step-2: Find $(N/2)$

Step-3: Look at the cumulative frequency just greater than $(N/2)$ and find the corresponding class, known as median class.

Step-4: Use the formula

$$\text{Median, (Me)} = l + \left\{ h \times \frac{(N/2 - C.f.)}{f} \right\}$$

l = Lower limit of median class

h = Width of median class

f = Frequency of median class

C.f. = Cumulative frequency of the class preceding the median class

$$N = \sum f_i$$

Illustration 6

Calculate the median for the following data.

Marks obtained	Number of students
Below 10	6
Below 20	15
Below 30	29
Below 40	41
Below 50	60
Below 60	70

Class – Interval	Frequency (f_i)	Comulative frequency
0-10	6	6
10-20	9	15
20-30	14	29
30-40	12	41
40-50	19	60
50-60	10	70
	$N = \sum f_i = 70$	

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

$$N = 70 \Rightarrow (N/2) = 35$$

The cumulative frequency just greater than 35 is 41. and the corresponding class is 30-40.

Thus the median class is 30-40

$$\therefore l = 30, \quad h = 10, \quad f = 12, \quad \text{C.f.} = \text{Cumulative frequency of preceding class} = 29$$

$$\begin{aligned} \text{Median } Me &= l + \left\{ h \times \frac{(N/2 - C.f.)}{F} \right\} \\ &= 30 + \left\{ 10 \times \frac{(35 - 29)}{12} \right\} \end{aligned}$$

$$\text{Ans.} = 35$$

6.5 MODE OF A GROUPED DATA

Mode : It is value of variate which occurs most often.

More precisely mode is that value of variable at which the concentration of data is maximum.

It is not necessary that in a series there must be only one mode. A distribution having only one mode is called unimodal, having two, bimodal and more than two multi modal.

Modal class: In a frequency distribution the class having maximum frequency is called modal class.

- Ex.** (i) The set of numbers 2, 3, 4, 7, 4, 5, 4, 9, 4 has mode 4 as it occurs the maximum number of times.
(ii) The set of numbers 5, 7, 6, 9, 1, 2 has no mode, as no number occurs more number of times than the other numbers.

Modal Class: In a frequency distribution, the class having maximum frequency is called the modal class.

Formula for calculating mode:

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

l = Lower limit of the modal class interval

f_1 = Frequency of the modal class

f_0 = Frequency of the class preceding the modal class

f_2 = Frequency of the class succeeding the modal class

h = Width of the class Interval

Illustration 7

Height (In Cm)	160-162	163-165	166-168	169-171	172-174
No. of Students	15	118	142	127	18

Find the average height of maximum number of students.

Solution

We have to find the mode of the data. The given data is an inclusive series. So we convert it to an exclusive form.

Class	Frequency
159.5 - 162.5	15
162.5 - 165.5	118
165.5 - 168.5	142
168.5 - 171.5	127
171.5 - 174.5	18

The class 165.5 - 168.5 has maximum frequency. So it is the modal class.

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

$$\begin{aligned}\Rightarrow \text{Mode} &= l + \left\{ \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right\} \times h \\ &= 165.5 + \left\{ 3 \times \frac{(142 - 118)}{(2 \times 142 - 181 - 127)} \right\} \\ &= 165.5 + 1.85 = 167.35\end{aligned}$$

Mode = 67.35 Cm.

Hence the average height of maximum number of is 167.35 Cm.

6.6 EMPIRICAL

Relationship between mean, median & mode

Mode = 3 Median - 2 Mean

6.7 GRAPHICAL REPRESENTATION

Graphical Representation : Many types of graphs are employed in statistics, depending upon the nature of the data involved. Among these are

- | | |
|--|---|
| (i) Bar chart (or Bar graph) | (ii) Histograms |
| (iii) Frequency Polygon | (iv) Cumulative Frequency Curve (Ogive) |
| (v) Pie chart (or Pie graph or Pie Diagrams) | |

Cumulative frequency Curve (Ogive) : An ogive is the graphical representation of cumulative frequency distribution. We can construct two types of ogives. The first form is “less than ogive” and the second is “more than ogive”.

In the “less than” method we start with the upper limit of the classes and go on adding the frequencies. When these are plotted, we get a rising curve.

In the “more than” method we start with the lower limit of the classes and from the total frequencies we subtract the frequency of each class. When these are plotted, we get declining curve, e.g.

The less than method: In this method the Ogive is cumulated upward. Scale the cumulative frequencies along the y-axis, and exact upper limits along the x-axis. The scale along the y-axis should be such as may accommodate the total frequency.

Procedure:

Step-1: Form the cumulative frequency table.

Step-2: Mark the actual upper class limits along the x-axis.

Step-3: Mark the cumulative frequencies of respective classes along the y-axis.

Step-4: Plot the points (upper limits, corresponding cumulative frequency).

To complete the ogive we also plot the point (lower limit of the lowest class, 0).

Step-5: Join these points by a smooth curve.

The curve so obtained is the required ogive.

Illustration 8

Draw a cumulative frequency curve (Ogive) for the following data.

Age (in years)	0-10	10-20	20-30	30-40
Number of patients	15	12	8	20

Solution

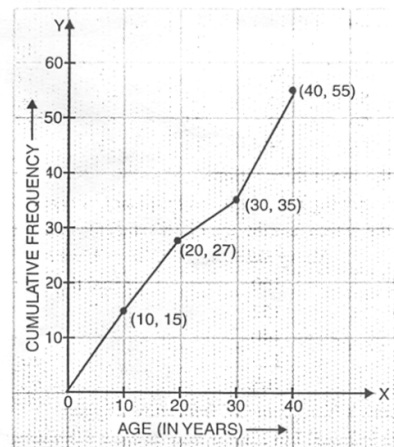
When nothing is mentioned then we will draw the 'less than curve'. The cumulative frequency table is as given below.

Age (in years)	Number of patients	Cumulative frequency
0-10	15	15
10-20	12	27
20-30	8	35
30-40	20	55

Upper class limits: 10 20 30 40

Cumulative frequency: 15 27 35 55

Plotting the points (10, 15), (20, 27), (30, 35), (40, 55) and joining them by a free hand curve we get the ogive as shown in figure. To complete it, we join the curve to the point (lower limit of the lowest class, 0), i.e (0,0).



Note: If we join the points by straight lines instead of by means of a free-hand, we get what is called a cumulative Frequency polygon.

“The more than” method. In this method the ogive is cumulated downward. Scale the cumulative frequencies along the y-axis and the exact lower limits along the x-axis.

Procedure:

Step-1: Scale the cumulative frequencies along the Y-axis and the actual lower limits along the X-axis.

Step-2: Plot the ordered pairs (lower limit, corresponding cumulative frequency).

To complete the ogive we also plot the ordered pair (upper limit of the highest class, 0)

Step-3: Join these plotted points by a smooth curve.

The curve so obtained is the required ogive.

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

“Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's” A.Sarras - USA

Illustration 9

Draw a 'more than ogive' from the following distribution.

Marks obtained	0-9	10-19	20-29	30-39	40-49	50-59
No. of candidates	4	6	7	5	10	3

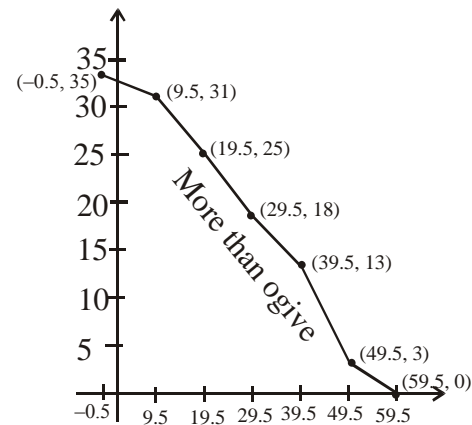
Solution

We write the given distribution using actual limits as under:

Marks obtained	-0.5-9.5	9.5-19.5	19.5-29.5	29.5-39.5	39.5-49.5	49.5-59.5
No. of candidates	4	6	7	5	10	3

For "more than ogive" we convert it to cumulative frequency distribution as under

Marks	Frequency	Plotting Points
- 0.5 or more	35	(- 0.5, 35)
9.5 or more	31	(9.5, 31)
19.5 or more	25	(19.5, 25)
29.5 or more	18	(29.5, 18)
39.5 or more	13	(39.5, 13)
49.5 or more	3	(49.5, 3)
more than 59.5	0	(59.5, 0)



Solved Examples

Example 1

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

x :	3	5	7	9	11	13
f :	6	8	15	p	8	4

Solution

Calculation of Mean

x_i	f_i	$f_i x_i$
3	6	18
5	8	40
7	15	105
9	p	$9p$
11	8	88
13	4	52
	$N = \sum f_i = 41 + p$	$\sum f_i x_i = 303 + 9p$

We have $\Sigma f_i = 41 + p$, $\Sigma f_i x_i = 303 + 9p$

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

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

$$\Rightarrow 7.5 \times (41 + p) = 303 + 9p$$

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

$$\Rightarrow 9p - 7.5p = 307.5 - 303 \Rightarrow 1.5p = 4.5 \Rightarrow p = 3$$

Example 2

The following tables gives weekly wages in rupees of workers in a certain commercial organization. The frequency of class 49–52 is missing. It is known that the mean of the frequency distribution is 47.2. Find the missing frequency

Weekly wages (Rs.)	40-43	43-46	46-49	49-52	52-55
Number of workers	31	58	60	?	27

Solution

Let the missing frequency be f , the assumed mean by $A = 47$ and $h = 3$.

Calculation of Mean

Class-intervals	Mid-values (x_i)	(f_i)	$d_i = x_i - 47.5$	$u_i = \frac{x_i - 47.5}{3}$	$f_i u_i$
40-43	41.5	31	-6	-2	-62
43-46	44.5	58	-3	-1	-58
46-49	47.5	60	0	0	0
49-52	50.5	f	3	1	f
52-55	53.5	27	6	2	54
	$N = \Sigma f_i = 176 + f$			$\Sigma f_i u_i = f - 66$	

We have,

$$\bar{X} = 47.2, A = 47.5 \text{ and } h = 3$$

$$\therefore \bar{X} = A + h \left\{ \frac{1}{N} \Sigma f_i u_i \right\}$$

$$\Rightarrow 47.2 = 47.5 + 3 \times \left\{ \frac{f - 66}{176 + f} \right\}$$

$$\Rightarrow -0.3 = 3 \times \left\{ \frac{f - 66}{176 + f} \right\}$$

$$\Rightarrow -0.3 = 3 \times \left\{ \frac{f - 66}{176 + f} \right\}$$

$$\Rightarrow \frac{-1}{10} = \frac{f - 66}{176 + f}$$

$$\Rightarrow -176 - f = 10f - 660$$

$$\Rightarrow 11f = 484 \Rightarrow f = 44$$

Hence, the missing frequency is 44.

Example 3

The median of the following data is 525. Find the values of x and y , if the total frequency is 100.

Class intervals	Frequency(f)
0-100	2
100-200	5
200-300	x
300-400	12
400-500	17
500-600	20
600-700	Y
700-800	9
800-900	7
900-1000	4

Solution**Computation of Median**

Class intervals	Frequency(f)	Cumulative frequency (f)
0-100	2	0
100-200	5	7
200-300	x	$7 + x$
300-400	12	$19 + x$
400-500	17	$36 + x$
500-600	20	$56 + x$
600-700	Y	$56 + x + y$
700-800	9	$65 + x + y$
800-900	7	$72 + x + y$
900-1000	4	$76 + x + y$
		Total = 100

We have

$$N = \sum f_i = 100$$

$$\Rightarrow 76 + x + y = 100 \Rightarrow x + y = 24$$

It is given that the median is 525. Clearly, it lies in the class 500 – 600.

$$\therefore l = 500, h = 100, f = 20, F = 36 + x \text{ and } N = 100$$

Now,

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

$$\Rightarrow 525 = 500 + \frac{50 - (36 + x)}{20} \times 100$$

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

EXERCISE-I

1. If the mean of the following distribution is 6, find the value of p .

x	2	4	6	10	$p + 5$
f	3	2	3	1	2

2. The weekly observations on cost of living index in a certain city for the year 2005-2006 are given below. Compute the mean weekly cost of living index.

Cost of living Index	No. of weeks
140–150	5
150–160	10
160–170	20
170–180	9
180–190	6
190–200	2
Total	52

3. Find the median for the following frequency distribution:

Class interval	0–8	8–16	16–24	24–32	32–40	40–48
Frequency	8	10	16	24	15	7

4. Find the missing frequencies in the following frequency distribution table, if $N = 100$ and median is 32.

Marks	0–10	10–20	20–30	30–40	40–50	50–60	Total
Number of students	10	?	25	30	?	10	100

Mode

5. Calculate the mode for the following frequency distribution:

Class	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
Frequency	5	8	7	12	28	20	10	10

6. The mode of the following series is 36. Find the missing frequency in it.

Class interval	0–10	10–20	20–30	30–40	40–50	50–60	60–70
Frequency	8	10	...	16	12	6	7

7. Following is the age distribution of a group of students. Draw the cumulative frequency curve of less than type and hence obtain the median value.

Age (in years)	Frequency
4–5	36
5–6	42
6–7	52
7–8	60
8–9	68
9–10	84
10–11	96
11–12	82
12–13	66
13–14	48
14–15	50
15–16	16

8. For the following frequency distribution, draw a cumulative frequency curve of more than type and hence obtain the median value.

Class interval	0–10	10–20	20–30	30–40	40–50	50–60	60–70
Frequency	5	15	20	23	17	11	9

EXERCISE-II

1. Find the mean marks of students from the following cumulative frequency table:

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

Marks	No 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

2. Find the mean marks of students from the following data:-

Marks	f
Below 10	5
Below 20	9
Below 30	17
Below 40	29
Below 50	45
Below 60	60
Below 70	70
Below 80	78
Below 90	83
Below 100	85

3. Compare the modal ages of two groups of students appearing for an entrance examination.

Age (in years)	16–18	18–20	20–22	22–24	24–26
Group A	50	78	46	28	23
Group B	54	89	40	25	17

4. The median of the following data is 26. Find the value of x and y if the total frequency is 80

Class Interval	0-8	8-16	16-24	24-32	32-40	40-48
Frequency	8	10	x	24	y	7

5. A survey regarding the heights (In cm) of 50 boys of a class was conducted and the following data was obtained

Height (In cm)	140-145	145-150	150-155	155-160	160-165	165-170	170-175
No. of boys	12	10	8	9	6	7	1

Draw both the ogives for the above data and hence obtain the median height of the students.

6. The table below shows the daily expenditure on food of 25 households in a locality

Daily expenditure (In Rs.)	100-150	150-200	200-250	250-300	300-350
No. of households	4	5	12	2	2

Find the mode and the mean of the data give above compute and interpret the two measures of central tendency.

7. Find the mean of the following data

Marks obtained	Less than 10	Less than 20	Less than 30	Less than 40	Less than 50
No. of students	7	19	32	42	50

8. Using a graph paper, draw an Ogive for the following distribution which shows a record of the weight in kilogram of 200 students.

Weight	40-45	45-50	50-55	55-60	60-65	65-70	70-75	75-80
Frequency	5	17	22	45	51	31	20	9

Use the ogive to estimate the following:

- The percentage of students weighing 55 kg or more;
 - The weight above which the heaviest 30% of the students fall;
 - The number of students who are;
 - Under weight (2) over weight, if 55.70 kg is considered as the standard weight.
9. Find the missing frequencies in the following distribution if it is known that the mean of the distribution is 1.46.

Number of accidents(x):	0	1	2	3	4	5	Total
Frequency(f):	46	?	?	25	10	5	200

EXERCISE-III

SECTION-A

- Multiple choice question with one correct answers

- The cumulative frequency distribution is represented by:
(A) Ogive (B) Bar diagram (C) Logistic curve (D) Histogram
- If a set of data has zero as an observation, then which one of the following is NOT an appropriate measure of central tendency?
(A) Arithmetic mean (B) Geometric mean (C) Median (D) Mode
- A, B, C are three sets of values of x:
A: 2, 3, 7, 1, 3, 2, 3 B: 7, 5, 9, 12, 5, 3, 8 C: 4, 4, 11, 7, 2, 3, 4
Select the correct statement from among the following
(A) Mean of A is equal to Mode of C (B) Mean of C is equal to Median of B
(C) Median of B is equal to Mode of A (D) Mean, Median and Mode of A are same

4. The mean of the values of 1, 2, 3,.....n with respective frequencies $x, 2x, 3x, \dots, nx$ is:
- (A) $\frac{n+1}{2}$ (B) $\frac{2n+1}{6}$ (C) $\frac{n}{2}$ (D) $\frac{2n+1}{3}$
5. If the variate is of discrete type, then the frequency distribution can be represented by:
- (A) A scatter diagram (B) A bar diagram (C) A histogram (D) A pie diagram
6. Find the simple and weighted mean of the first 'n' natural numbers, the weights being the corresponding numbers:
- (A) $\frac{n-1}{2}, \frac{2n+1}{3}$ (B) $\frac{n+1}{2}, \frac{2n-1}{3}$ (C) $\left\{ \frac{n+1}{2}, \frac{2n+1}{3} \right\}$ (D) $\frac{n-1}{2}, \frac{2n-1}{3}$
7. The number of observation in a group is 40. If the average of first 10 is 4.5 and that of the remaining 30 is 3.5, then the average of the whole group is:
- (A) $\frac{1}{5}$ (B) $\frac{15}{4}$ (C) 4 (D) 8
8. The median of a set of 9 distinct observations is 20.5. If each of the largest 4 observations of the set is increased by 2, then the median of the new set:
- (A) Is increased by 2 (B) Is decreased by 2
(C) Is two times the original median (D) Remains the same as that of the original set
9. If the mean of x_1 and x_2 is M_1 , and that of x_1, x_2, x_3, x_4 is M_2 , the mean of $ax_1, ax_2, x_3/a, x_4/a$ is
- (A) $\frac{M_1 + M_2}{2}$ (B) $\frac{aM_1 + (M_2/a)}{2}$
(C) $\frac{1}{2a} \left\{ (a^2 - 1)M_1 + 2M_2 \right\}$ (D) $\frac{1}{2a} \left\{ 2(a^2 - 1)M_1 + 2M_2 \right\}$
10. If a frequency distribution for the number of persons x in a household is prepared with class intervals as (1-4), (5-8), (9-12), etc., then the number of persons x belonging to class interval (5-8) satisfies.
- (A) $5 < x < 8$ (B) $5 \leq x < 8$ (C) $5 < x \leq 8$ (D) $5 \leq x \leq 8$
11. In a histogram heights of rectangles are:
- (A) Always proportional to the frequencies of the classes
(B) Proportional to the frequency densities if the classes of distributions are not of equal size
(C) Always proportional to the cumulative frequencies of the classes
(D) Proportional to the cumulative frequencies only when classes of distribution are of equal size
12. Arithmetic mean of n observations is m . If two observations 0 and m are added, then the new mean will be:
- (A) m (B) $\frac{m}{n+1}$ (C) $\frac{nm}{n+1}$ (D) $\frac{(n+1)m}{n+2}$
13. For a symmetric distribution, the empirical relationship between mean, median and mode is:
- (A) Mean > Median > Mode (B) Mean = Median + Mode
(C) Mode - Mean = 3 (Median - Mean) (D) Mean - Mode = 3 (Mean - Median)

14. In a histogram with equal class intervals, heights of bars are proportional to:
 (A) Mid - value of the classes (B) Frequencies of respective classes
 (C) Cumulative frequency of the classes (D) Class interval of the classes
15. The median of the following incomplete frequency distribution is 4

x	1	2	3	4	5	6	7	8
Frequency	2	3	4	1	2	4	2	-

The frequency of 8 is:

- (A) 1 (B) 2 (C) 3 (D) 4
16. The mid value of a class interval is 42. If the class size is 10, then the upper and lower limits of the class are:
 (A) 47 and 37 (B) 37 and 47 (C) 37.5 and 47.5 (D) 47.5 and 37.5
17. The mode of the given distribution is:

Weight (in kg)	40	43	46	49	52	55
Number of children	5	8	16	9	7	3

- (A) 40 (B) 46 (C) 55 (D) None of these

SECTION-B

- Assertion & Reason**

Instructions: In the following questions as Assertion (A) is given followed by a Reason (R). Mark your responses from the following options.

- (A) Both Assertion and Reason are true and Reason is the correct explanation of 'Assertion'
 (B) Both Assertion and Reason are true and Reason is not the correct explanation of 'Assertion'
 (C) Assertion is true but Reason is false
 (D) Assertion is false but Reason is true

1. **Assertion:** The monthly salaries of four persons are Rs. 10,000, Rs. 10,500, Rs. 11,000 and Rs. 12,000. The arithmetic mean is typical of their salaries.
Reason: Arithmetic mean is strongly affected by extreme values.
2. **Assertion:** The most accurate graphical representation of statistical data is by means of ogives.
Reason: Because cumulative frequencies are represented in it.
3. **Assertion:** The mean of $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8,$ and x_9 which are in an arithmetic progression is x_5 .
Reason: Mean is always the middle most observation if the data are in an

SECTION-C

- Match the following (one to one)**

Column-I and **column-II** contains **four** entries each. Entries of column-I are to be matched with some entries of column-II. Only One entries of column-I may have the matching with the same entries of column-II and one entry of column-II Only one matching with entries of column-I

1. **Column I**

(A) The direct method

(B) Step deviation method

(C) Mode

(D) Median

Column II

(P)
$$\bar{X} = a + \frac{\sum f_i u_i}{\sum f_i} h$$

(Q)
$$l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2} \right) xh$$

(R)
$$l + \left(\frac{\frac{N}{2} - c.f}{f} \right) \times h$$

(S)
$$\bar{X} = \frac{\sum f_i x_i}{\sum f_i}$$

2.

Class Interval	Frequency
30-34	7
35-39	10
40-44	12
45-49	13
50-54	8
55-59	4

Column I

(A) The actual class limits of the fourth class

(B) The class boundaries of the sixth class

(C) The class mark of the third class

(D) The size of the third class

Column II

(P) 5

(Q) 44.5 - 49.5

(R) 54.5 - 59.5

(S) 42

Answers

EXERCISE-I

1. P = 7

3. 26

4. $b_1 = 9, b_2 = 16$

6. 10

2. 166.35

7. 35

5. 46.67

pdfMachine - is a pdf writer that produces quality PDF files with ease!

Get yours now!

"Thank you very much! I can use Acrobat Distiller or the Acrobat PDFWriter but I consider your product a lot easier to use and much preferable to Adobe's" A.Sarras - USA

EXERCISE-II

1. 51.75 marks
2. 48.41 marks
3. Modal A group = 18.93, B group = 18.83 (A group > B group)
4. $x = 16, y = 15$
5. 151.88 cm.
6. Mode = Rs. 220.59, Mean = Rs. 211.00
Mean and mode use very close to each other. Maximum number of households spend Rs. 220.59 per day while average per day expenditure of a household of Rs. 211.00
7. 25 marks.
8. (i) = 78% (ii) 45, 154
9. $p = 11$

EXERCISE-III**SECTION-A**

- | | | | | |
|---------|---------|---------|---------|---------|
| 1. (A) | 2. (D) | 3. (D) | 4. (D) | 5. (B) |
| 6. (C) | 7. (B) | 8. (D) | 9. (B) | 10. (B) |
| 11. (A) | 12. (A) | 13. (C) | 14. (B) | 15. (A) |
| 16. (A) | 17. (B) | | | |

SECTION-B

1. (C)
2. (A)
3. (C)

SECTION-C

1. (A) - (S), (B) - (P), (C) - (Q), (D) - (R)
2. (A) - (Q), (B) - (R), (C) - (S), (D) - (P)