CHAPTER -14

STATISTICS

Prepared by, Puja Kumari
TGT (MATHS)
KV Patratu

- **Statistics** is one of the parts of mathematics in which we study about the collecting, organizing, analyzing, interpreting and presenting data
- **Ungrouped data** - Ungrouped data is data in its original or raw form. The observations are not classified into groups.
- **Grouped data** - In grouped data, observations are organized in groups.

For example, a class of students got different marks in periodic test. The data is tabulated as below

Marks interval	0-10	10-20	20-30	30-40
No. of students	4	8	12	16

- **Frequency** (f) -Frequency is the number of times a particular observation occurs in data.
- **Class Interval** - Data can be grouped into class intervals such that all observations in that range belong to that class.
- **Class width/Class Size** (h) = upper class limit – lower class limit

Three measures of central tendency

A) Mean ($\bar{x}$) **B) Median** **C) Mode**

A) METHODS OF FINDING MEAN

i) Direct Method:

$$\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$$

x_i = Class mark

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

f i = frequency

Example: Find Mean by Direct method:

Class interval	10 - 25	25 - 40	40 - 55	55 - 70	70 - 85	85 - 100
Number of students	2	3	7	6	6	6

Solution:

Class interval	Number of students (f_i)	Class mark (x_i)	$f_i x_i$
10 - 25	2	17.5	35.0
25 - 40	3	32.5	97.5
40 - 55	7	47.5	332.5
55 - 70	6	62.5	375.0
70 - 85	6	77.5	465.0
85 - 100	6	92.5	555.0
Total	$\Sigma f_i = 30$		$\Sigma f_i x_i = 1860.0$

$$\bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{1860.0}{30} = 62$$

This method of finding the mean is known as the **Direct Method.**

ii) Assumed Mean Method

$$\text{Mean } (\bar{x}) = a + \frac{\sum_{i=1}^{n} fidi}{\sum_{i=1}^{n} fi}$$

where a = assumed mean

$d_i = x_i - a$

Example: Find Mean by Assumed Mean method:

Class interval	10 - 25	25 - 40	40 - 55	55 - 70	70 - 85	85 - 100
Number of students	2	3	7	6	6	6

Class interval	Number of students (f_i)	Class mark (x_i)	$d_i = x_i - 47.5$	$f_i d_i$
10 - 25	2	17.5	–30	–60
25 - 40	3	32.5	–15	–45
40 - 55	7	47.5	0	0
55 - 70	6	62.5	15	90
70 - 85	6	77.5	30	180
85 - 100	6	92.5	45	270
Total	$\Sigma f_i = 30$			$\Sigma f_i d_i = 435$

the mean of the deviations, $\bar{d} = \frac{\Sigma f_i d_i}{\Sigma f_i}$

Mean of deviations, $\bar{d} = \frac{\Sigma f_i d_i}{\Sigma f_i}$

So, $\bar{d} = \frac{\Sigma f_i (x_i - a)}{\Sigma f_i}$

$= \frac{\Sigma f_i x_i}{\Sigma f_i} - \frac{\Sigma f_i a}{\Sigma f_i}$

$= \bar{x} - a \frac{\Sigma f_i}{\Sigma f_i}$

$= \bar{x} - a$

So, $\bar{x} = a + \bar{d}$

i.e., $\bar{x} = a + \frac{\Sigma f_i d_i}{\Sigma f_i}$

Substituting the values of a, $\Sigma f_i d_i$ and Σf_i from Table 14.4, we get

$$\bar{x} = 47.5 + \frac{435}{30} = 47.5 + 14.5 = 62.$$

Therefore, the mean of the marks obtained by the students is 62.

The method discussed above is called the **Assumed Mean Method**.

B) Mode

Mode of grouped data can be found as

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

where l = lower limit of the modal class,

h = size of the class interval (assuming all class sizes to be equal),

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.

Example:

A survey conducted on 20 households in a locality by a group of students resulted in the following frequency table for the number of family members in a household:

Family size	1 - 3	3 - 5	5 - 7	7 - 9	9 - 11
Number of families	7	8	2	2	1

Find the mode of this data.

Solution :

The maximum class frequency is 8, and the class corresponding to this frequency is 3 – 5. So, the modal class is 3 – 5.

Now modal class = 3 – 5,

lower limit (l) of modal class = 3,

class size (h) = 2

frequency (f_1) of the modal class = 8,

frequency (f_0) of class preceding the modal class = 7,

frequency (f_2) of class succeeding the modal class = 2.

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

$$= 3 + \left(\frac{8 - 7}{2 \times 8 - 7 - 2}\right) \times 2 = 3 + \frac{2}{7} = 3.286$$

Therefore, the mode of the data above is 3.286.

C) Median

The median for grouped data can be found by using the formula

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

Where l = lower limit of the median class

n = number of observations

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

h = class width

Example:

The following table gives the distribution of the life time of 400 neon lamps :

Life time (in hours)	Number of lamps
1500 - 2000	14
2000 - 2500	56
2500 - 3000	60
3000 - 3500	86
3500 - 4000	74
4000 - 4500	62
4500 - 5000	48

Find the median life time of a lamp

Solution:

Lifetime (in hours)	Number of lamps	*cf*
1500 – 2000	14	14
2000 – 2500	56	70
2500 – 3000	60	130
3000 – 3500	86	216
3500 – 4000	74	290
4000 – 4500	62	352
4500 – 5000	48	400
Total	400	

Here, $\frac{n}{2} = \frac{400}{2} = 200$

$\therefore$ Median class = 3000 – 3500

So, $f = 86, cf = 130, h = 500$

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

$= 3000 + \left(\frac{200 - 130}{86}\right) \times 500$

$= 3000 + \frac{35000}{86} = 3000 + 406.98 = 3406.98$ hours

- **There is a empirical relationship between the three measures of central tendency:**

3 Median = Mode + 2 Mean

Practice Questions

I. Multiple Choice Questions (1marks each)

*i) If the mode of a distribution is 8 and its mean is also 8, then its median is

a) 10 b) 8 c) 7 d) 6

ii)Consider the following distribution:

Marks obtained	0 or more	10 or more	20 or more	30 or more	40 or more	50 or more
No. of students	63	58	55	51	48	42

The frequency of the class 30-40 is

a) 3 b) 4 c) 48 d) 5

**iii) Consider the following frequency distribution of the heights of 60 students of a class

Height in cm	150-155	155-160	160-165	165-170	170-175	175-180
No. of students	15	13	10	8	9	5

The sum of the lower limit of the modal class and upper limit of the median class is.

(a) 310 (b) 315 (c) 320 (d) 330

iv)The modal class of the following distribution is.

class	0-10	10-20	20-30	30-40	40-50	50-60
frequency	3	9	15	30	18	5

a) 40-50 b)20-30 c) 30-40 d) 50-60

v) The median class of the following frequency distribution is

class	0-10	10-20	20-30	30-40	40-50	50-60
frequency	8	10	12	22	30	18

a) 20-30 b)30-40 c) 40-50 d) none of these

*vi) For the following distribution

Class	10-15	15-20	20-25	25-30	30-35
frequency	25	30	27	35	21

The sum of the lower limit of the median class and the lower limit of the modal class is

a) 45 b) 50 c) 55 d) 62

*vii)The times, in seconds, taken by 50 athletes to run a 110 m hurdle race are tabulated below.

Time in seconds	13.8-14	14-14.2	14.2-14.4	14.4-14.6
No. of athletes	2	14	16	18

The number of athletes who completed the race in less than 14.4 seconds is :

a) 2 b)32 c) 16 d)50

***viii) Relationship between mean median and mode is

a)3 Median = 2 Mode + Mean b) 3Mode = 3 Mean + Median

c) 3Median = 2Mean +Mode d) 3Mean = 2Mode + Median

ix) Consider the following frequency distribution

Class interval	0-6	6-12	12-18	18-24	24-30
Frequency	12	10	15	8	11

The median class is

a) 6-12 b) 12-18 c)18-24 d) 24-30

**x) If the mean of the following distribution is 2.6, then the value of y is

Variable (x)	1	2	3	4	5
Frequency(f)	4	5	Y	1	2

a) 3 b) 13 c) 24 d) 8

II.VERY SHORT ANSWER TYPE QUESTIONS (2 marks each)

1. Calculate the median from the following data

Marks	0-10	10-20	20-30	30-40	40-50
No. of students	5	15	30	8	2

2.Find the mode of the following frequency distribution

Class interval	0-6	6-12	12-18	18-24	24-30
Frequency	7	5	10	12	6

**3. Find the value of p, if the arithmetic mean of the following distribution is 25:

Class interval	0-10	10-20	20-30	30-40	40-50
Frequency	5	8	15	p	6

4. Find median class of the following distribution

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70
frequency	4	4	8	10	12	8	4

***5. Find x and y from the following frequency distribution

Class	frequency	Cumulative frequency
0-8	15	15
8-16	X	28
16-24	15	43
24-32	18	Y
32-40	9	70

6. Find mean of the following distribution

class	3-5	5-7	7-9	9-11	11-13
frequency	5	10	10	7	8

*7. For the following distribution find the modal class

Marks	Below 10	Below 20	Below 30	Below 40	Below 50	Below 60
No of students	3	12	27	57	75	80

**8. Find the sum of lower limit of median class and upper limit of modal class

class	10-20	20-30	30-40	40-50	50-60	60-70
frequency	1	3	5	9	7	3

III. SHORT ANSWER TYPE QUESTIONS (3 marks each)

**1. The length of 40 leaves of a plant are measured correct to nearest millimetre, and the data obtained is represented in the following table.

Length (in mm)	118-126	127-135	136-144	145-153	154-162	163-171	172-180
No of leaves	3	5	9	12	5	4	9

Find the average length of the leaves

2. Find mean of the following distribution

class	20-30	30-40	40-50	50-60	60-70
frequency	25	40	42	43	10

3.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 modal age of the participants.

*4.The marks obtained by 110 students in an examination are given below

Class	30-35	35-40	40-45	45-50	50-55	55-60	60-65
frequency	14	16	28	23	18	8	3

Find the mean marks of the students.

**5.If the mean of the following frequency distribution is 18. Find the missing frequency

Class	11-13	13-15	15-17	17-19	19-21	21-23	23-25
frequency	3	6	9	13	-------	5	4

**6.If the mean of the following data is 14.7, find the values of p and q

Class	0-6	6-12	12-18	18-24	24-30	30-36	36-42	Total
frequency	10	P	4	7	Q	4	1	40

IV. Long Answer Type Questions (5 marks each)

***1. 250 apples of a box were weighed and distribution of the masses of the apples is given below in the following table.

Mass (in grams)	80-100	100-120	120-140	140-160	160-180
Number of apples	20	60	70	X	60

a) find the value of x and the mean mass of the apples

b) find the modal mass of apples

**2. The mode of the following frequency distribution is 55.Find the missing frequencies ‘a’ and ‘b’

Class interval	0-15	15-30	30-45	45-60	60-75	75-90	Total
frequency	6	7	a	15	10	b	51

***3.The median of the following data is 50. Find the values of ‘p’ and ‘q’, if the sum of all frequencies is 90. Also find the mode of the data

Marks obtained	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No of students	P	15	25	20	q	8	10

***4.The following table gives the distribution of the life time of 400 neon lamps:

Lifetime (in hours)	1500-2000	2000-2500	2500-3000	3000-3500	3500-4000	4000-4500	4500-5000
No of lamps	14	56	60	86	74	62	48

Find the average lifetime of a lamp

*5.The following distribution shows the daily pocket allowance of children of a locality. The mean pocket allowance is 18. Find the missing frequency.

Daily pocket allowance (in Rs)	11-13	13-15	15-17	17-19	19-21	21-23	23-25
No of children	7	6	9	13	f	5	4

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

Class interval	0-100	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900	900-1000
frequency	2	5	x	12	17	20	y	9	7	4

**7. A life insurance agent found the following data for distribution of ages of 100 policyholders. Calculate the median age, if policies are given only to persons having age 18years onwards but less than 60 years.

Age (in years)	Below 20	20-25	25-30	30-35	35-40	40-45	45-50	50-55	55-60
No of policy holders	2	5	x	12	17	20	y	9	7

***8. The mean of the following distribution is 18. Find frequency of the class 19-21.

Class interval	11-13	13-15	15-17	17-19	19-21	21-23	23-25
Frequency	3	6	9	13	f	5	4

**9. The distribution given below shows the number of wickets taken by bowlers in one day cricket matches. Find the mean and median of the number of wickets taken.

No. of wickets	20-60	60-100	100-140	140-180	180-220	220-260
No. of bowlers	7	5	16	12	2	3

V. CASE STUDY BASED QUESTIONS

*1.A group of students decided to make a project on Statistics. They are collecting the heights (in cm) of their 51 girls of Class X A, B and C of their school. After collecting the data, they arranged the data in the following less than cumulative frequency distribution table form.

Height (in cm)	Number of girls
Less than 140	4
Less than 145	11
Less than 150	29
Less than 155	40
Less than 160	46
Less than 165	51

Class intervals	Frequency	Cumulative frequency
Below 140	4	4
140 - 145	7	11
145 - 150	18	29
150 - 155	11	40
155 - 160	6	46
160 - 165	5	51

i) What is the lower limit of the median class?

a)145 b)150 c)155 d)160 145

ii) What is the upper limit of the modal class?

a)145 b)150 c)155 d)160 150

iii) What is the mean of the lower limits of the median and modal class?

a)145 b)150 c)155 d)160 145

iv) What is the width of the class ?

a)10 b)15 c) 5 d) none of these

**2. Overweight and obesity may increase the risk of many health problems including diabetes, heart disease and certain cancers. The basic reason behind is the laziness, eating more junk foods and less physical exercise. The school management give instructions to the school to collect the weight data of each student. During medical check of 35 students from class X A, their weight was recorded as follows:

Weight (in Kg)	No of students
Less than 38	0
Less than 40	3
Less than 42	5
Less than 44	9
Less than 46	14
Less than 48	28
Less than 50	32
Less than 52	35

i) Find the median class of the above data?

a) 44-46 b) 46-48 c) 48-50 d) 0-52

ii)What is the median weight of the data?

a) 46 b) 46.5 c) 47 d) 47.5

iii) what is the mean of the above data?

a) 45.8 b) 46.2 c) 45.2 d) 46.5

iv) How many students have weight in the range of 44-46 kg ?

a) 2 b) 3 c) 5 d) 5

***3. A group of students went to another city to collect the data of monthly consumptions(in units) to complete their Statistics project. They prepare the following frequency distribution table from the collected data which gives monthly consumers of a locality.

Monthly consumption (in units)	No. of consumers
65-85	4
85-105	5
105-125	13
125-145	20
145-165	14
165-185	8
185-205	4

i) What is the lower limit of the median class?

a) 145 b) 165 c)105 d)125

ii)What is the lower limit of the modal class?

a) 105 b) 125 c)145 d) 165

iii)What is the width of the class?

a) 5 b) 10 c) 25 d) 20

iv) How many consumers' monthly consumption is more than 145 units?

a) 22 b) 14 c) 26 d) 8

***4. The COVID-19 pandemic, also known as coronavirus pandemic, is an ongoing pandemic of coronavirus disease caused by the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) among humans.

The following tables shows the age distribution of case admitted during a day in two different hospitals.

Table 1

Age (in years)	5-15	15-25	25-35	35-45	45-55	55-65
No. of cases	6	11	21	23	14	5

Table 2

Age(in years)	5-15	15-25	25-35	35-45	45-55	55-65
No. of cases	8	16	10	42	24	12

Based on the above information answer the following questions.

i)The average age for which maximum cases occurred is (refer to table 1)

a)32.24 b)34.36 c)36.82 d)42.24

ii)The upper limit of the modal class is (refer to table 1)

a)15 b)25 c)35 d)45

iii)The mode of the given data is(refer to table2)

a)41.4 b)48.2 c)55.3 d)64.6

iv)The median of the given data is (refer to table 2)

a)32.7 b)40.2 c)42.3 d)48.6

CHAPTER 14 ANSWER KEY I MCQ

i).b	ii).a	iii)b	iv) c	v) b	vi) a	vii) b	viii) c	ix) b	x) d

II SHORT ANSWER TYPE QUESTIONS (2 marks each)

1) Median= 23.3 2) Mode =19.5 3) p=6

4. 30-40 5. x = 13 , y =61 6. Mean 8.15 7. 30-40 8. 90

III SHORT ANSWER TYPE QUESTION (3 marks each)

1) 146.75 2) Approx. 42.5 3)58.205 4)44.81 5)8

6) p=11, q=3

IV LONG ANSWER TYPE QUESTIONS (5 marks each)

1) x=40, Mean=134.8, Median132.85 2) a= 5, b=8 3) p =5, q=7, Mode= 46.67

4) 34105) f=20 6) x=9, y=15

7) 35.76 8) f=8 9) Median= 126.25; Mean= 125.33

V.CASE STUDY BASED QUESTIONS

1.

i)a	ii)b	iii)a	iv)c

2.

i) b	ii)b	iii)a	iv)c

3.

i) d	ii) b	iii) d	iv) c

4.

i) c	ii) d	iii) a	iv) b