Chapter 3 - Organisation Of Data

Question 1(i):

The class midpoint is equal to:

(a) The average of the upper class limit and the lower class limit.

(b) The product of upper class limit and the lower class limit.

(c) The ratio of the upper class limit and the lower class limit.

(d) None of the above.

Answer:

The option (a) is correct.

The class midpoint is equal to the average of the upper class limit and the lower class limit. It is known by adding the values of upper and lower limits and dividing the total by 2.

Question 1(ii):

The frequency distribution of two variables is known as

(a) Univariate Distribution

(b) Bivariate Distribution

(c) Multivariate Distribution

(d) None of the above

Answer:

The option (b) is correct.

The frequency distribution of two variables is known as Bivariate Frequency Distribution. In other words, Bivariate Frequency Distribution shows the series of statistical data having frequencies of two variables such as the data on income and expenditure of the households.

Question 1(iii):

Statistical calculations in classified data are based on

(a) the actual values of observations

(b) the upper class limits

(c) the lower class limits

(d) the class midpoints

Answer:

The option (d) is correct.

The calculations in classified data or continuous series are based on the class midpoints. The items in a continuous series cannot be exactly measured. Consequently, the class midpoints are calculated.

Question 1(iv):

Under Exclusive method,

(a) the upper class limit of a class is excluded in the class interval

(b) the upper class limit of a class is included in the class interval

(c) the lower class limit of a class is excluded in the class interval

(d) the lower class limit of a class is included in the class interval

Answer:

The option (a) is correct.

A series in which upper limit of one class becomes the lower limit of the succeeding class interval is called exclusive series. In such series, the frequencies of the lower limit are included in that particular class whereas the frequencies of the upper limit are excluded.

Question 1(v):

Range is the

(a) difference between the largest and the smallest observations

(b) difference between the smallest and the largest observations

(c) average of the largest and the smallest observations

(d) ratio of the largest to the smallest observation

Answer:

The option (a) is correct.

Range is defined as the difference between the largest and the smallest observations.

Algebraically,

$R = H - L$

Where,

R denotes range

H is the highest value

L is the lowest value

Question 2:

Can there be any advantage in classifying things? Explain with an example from your daily life.

Answer:

Yes, there are many advantages of classifying things. The following are the advantages associated with classification:

1. **Saves Time and Energy-** Classification of things not only saves our time but also our energy which would otherwise be utilised in searching from entire lot of things.

2. **Quick Information-** Information can be easily collected from the classified things.

3. **Easy Classification-** Classification facilitates comparisons and helps in drawing fast conclusions or inferences.

The advantage of classification can be better understood with the help of a daily life example. A post office on the regular basis sorts letters and then classifies them according to various attributes. Letters are classified first according to the states, then according to the cities and streets. Thus, this process of classification helps the postman to deliver posts quickly, efficiently and in a non-haphazard manner.

Question 3:

What is a variable? Distinguish between a discrete and a continuous variable.

Answer:

A measurable characteristic whose value changes overtime is called variable. It refers to that quantity which keeps on changing and which can be measured by some unit. For example, if we measure the height of students of a class, then height is regarded as a variable. A variable can be either discrete or continuous.

Discrete Variable	**Continuous Variable**
A variable that takes only whole number as its value is called discrete variable.	A variable that can take any value, within a reasonable limit is called a continuous variable.
These variables increase in jumps or in complete numbers.	These variables assume a range of values or increase in fractions and not in jumps.

For example- Number of people in a family, number of students in a class, etc.	For example- age, height, weight, etc.

Question 4:

Explain the 'exclusive' and 'inclusive' methods used in classification of data.

Answer:

Exclusive Method– This method is used for those series in which the upper limit of one class becomes the lower limit of the next class. It is called as exclusive series because the frequencies of the upper limit of a class interval are not included in that particular class. In such type of series, the upper limit of one class becomes the lower limit of the next class, for example, 0–10, 10–20, 20–30 and so on. The upper limit is excluded but the lower limit is included in the class interval. This method is most appropriate for data of continuous variables.

Inclusive Method– Under this method of classification of data, the classes are formed in such a manner that the upper limit of a class interval does not repeat itself as the lower limit of the next class interval. In such a series, both the upper limit and the lower limit are included in the particular class interval, for example, 1–5, 6–10, 11–15 and so on. The interval 1–5 includes both the limits i.e. 1 and 5.

Question 5(i):

Use the data in Table 3.2 that relate to monthly household expenditure (in Rs) on food of 50 households and obtain the range of monthly household expenditure on food.

Answer:

Calculation of Range

Range = Highest Value – Lowest Value
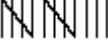
Highest Value = 5090

Lowest Value = 1007

So, Range = 5090 – 1007 = 4083

Question 5(ii):

Divide the range into appropriate number of class intervals and obtain the frequency distribution of expenditure.

Answer:

(ii) *Preparing Tally Marks*

Class Intervals	Tally Marks	Frequency
1000 – 1500	𝍷𝍷𝍷𝍷 𝍷𝍷𝍷𝍷 𝍷𝍷𝍷𝍷 𝍷𝍷𝍷𝍷	20
1500 – 2000	𝍷𝍷𝍷𝍷 𝍷𝍷𝍷𝍷 |||	13
2000 – 2500	𝍷𝍷𝍷𝍷		06
2500 – 3000	𝍷𝍷𝍷𝍷	05
3000 – 3500	||	02
3500 – 4000			01
4000 – 4500	||	02
4500 – 5000	–	00
5000 – 5500			01
Total		**50**

Question 5(iii):

Find the number of households whose monthly expenditure on food is

(a) less than Rs 2000

(b) more than Rs 3000

(c) between Rs 1500 and Rs 2500

Answer:

a) Number of households whose monthly expenditure on food is less than Rs 2000

= 20 + 13 = 33

b) Number of households whose monthly expenditure on food is more than Rs 3000

= 2+1+2+0+1 = 6

c) Number of households whose monthly expenditure on food is between Rs 1500 and Rs 2500

= 13 + 6 = 19

Question 6:

In a city 45 families were surveyed for the number of domestic appliances they used. Prepare a frequency array based on their replies as recorded below.

1	3	2	2	2	2	1	2	1	2	2	3	3	3	3
3	3	2	3	2	2	6	1	6	2	1	5	1	5	3
2	4	2	7	4	2	4	3	4	2	0	3	1	4	3

Answer:

Frequency Array of appliances being used by households

No. of Domestic Appliances	No. of Households
0	1
1	7
2	15
3	12
4	5
5	2
6	2
7	1
Total	**45**

Question 7:

What is 'loss of information' in classified data?

Answer:

'Loss of information' is a major drawback of the classified data. The classification or grouping of raw data into classes makes it more concise and understandable. But simultaneously there exists loss of information. The calculations involved in the classified data or the continuous series are based on the class midpoints. The items in such series cannot be exactly measured and consequently, an individual observation loses its importance during the statistical calculations. Further, the statistical calculations are based on the values of the class marks, ignoring the exact observations of the data leading to the problem of loss of information.

Question 8:

Do you agree that classified data is better than raw data?

Answer:

The classified data has following advantages over the raw data.

1. **Comprehensive-**Raw data are large and entangled, whereas classified data are comprehensive and easily manageable.

2. **Quick Information-** It is troublesome to pick up information from unclassified data. Information can be easily collected from the classified data.

3. **Conclusions -** Classification facilitates comparisons and helps in drawing fast conclusions or inferences.

4. **Saves Time and Energy-** Classified data not only save our time but also our energy, which would otherwise be utilised in searching from entire lot of things.

Question 9:

Distinguish between univariate and bivariate frequency distribution.

Answer:

Univariate Frequency Distribution	Bivariate Frequency Distribution
The word '*Uni*' means one. A series of statistical data showing the frequency of only one variable is called Univariate Frequency Distribution. In other words, the frequency distribution of single variable is called Univariate Frequency Distribution. For example- income of people, marks scored by students, etc.	The word '*Bi*' means two. A series of statistical data showing the frequency of two variables simultaneously is called Bivariate Frequency Distribution. In other words, the frequency distribution of two variables is called Bivariate Frequency Distribution. For example- sales and advertisement expenditure, weight and height of individuals, etc.

Question 10:

Prepare a frequency distribution by inclusive method taking class interval of 7 from the following data:

28	17	15	22	29	21	23	27	18	12	7	2	9	4
1	8	3	10	5	20	16	12	8	4	33	27	21	15
3	36	27	18	9	2	4	6	32	31	29	18	14	13
15	11	9	7	1	5	37	32	28	26	24	20	19	25

19   20   6   9

Answer:

Class Interval	Tally marks	Frequency
0 – 7	|||| |||| ||||	15
8 – 15	|||| |||| ||||	15
16 – 23	|||| |||| ||||	14
24 – 31	|||| ||||		11
32 – 39	||||	5
Total		**60**