StudyByU

STATISTICS

1. INTRODUCTION

1.1 What is Statistics?

Statistics is the branch of mathematics that deals with:

- **Collection of data**
- **Organization of data**
- **Presentation of data**
- **Analysis of data**
- **Interpretation of data**

1.2 Measures of Central Tendency

The three main measures are:

1. **Mean - Average value**
2. **Median - Middle value**
3. **Mode - Most frequent value**

Why do we need these?

- **To summarize large data**
- **To understand the "center" of data**
- **To compare different data sets**

1.3 Types of Data

Ungrouped Data: Individual observations

Example: 5, 7, 9, 12, 15

Grouped Data: Data organized in class intervals

Example:

Class Interval	Frequency
0-10	5
10-20	8
20-30	12

This chapter focuses on GROUPED DATA

2. BASIC TERMINOLOGY

2.1 Class Interval

A class interval is a group of values.

Example: 10-20

- Lower limit (l) = 10

- Upper limit (u) = 20

2.2 Class Mark (Mid-value)

Class Mark = (Lower limit + Upper limit)/2

Formula: $x_i = (l + u)/2$

Example: For class 10-20

Class mark = (10 + 20)/2 = 15

2.3 Class Size (Width)

Class Size = Upper limit - Lower limit

Formula: $h = u - l$

Example: For class 10-20

Class size = 20 - 10 = 10

2.4 Frequency (f)

Frequency is the number of observations in each class.

2.5 Cumulative Frequency (cf)

Cumulative Frequency is the sum of frequencies up to that class.

Example:

Class	Frequency	Cumulative Frequency
0-10	5	5
10-20	8	5 + 8 = 13

20-30 | 12 | 13 + 12 = 25

Two Types:

1. **Less than cumulative frequency (sum from above)**
2. **More than cumulative frequency (sum from below)**

3. MEAN OF GROUPED DATA

⭐ **MOST IMPORTANT TOPIC** ⭐

3.1 What is Mean?

Mean is the average of all observations.

For grouped data:

Mean = (Sum of all observations) / (Total number of observations)

3.2 Method 1: Direct Method

Formula:

Mean ($\bar{x}$) = $\Sigma f_i x_i$ / Σf_i

Where:

- **x_i = Class mark of ith class**
- **f_i = Frequency of ith class**
- **Σf_i = Total frequency = n**

Steps:

1. **Find class mark (x_i) for each class**
2. **Multiply x_i by f_i to get $f_i x_i$**
3. **Find $\Sigma f_i x_i$ (sum of $f_i x_i$ column)**
4. **Find Σf_i (sum of frequencies)**
5. **Apply formula: Mean = $\Sigma f_i x_i$ / Σf_i**

Example:

Class	Frequency (fi)	Class Mark (xi)	fixi
0-10	5	5	25
10-20	8	15	120
20-30	12	25	300
30-40	6	35	210
	Σfi = 31		Σfixi = 655

Mean = Σfixi / Σfi = 655 / 31 = 21.13

3.3 Method 2: Assumed Mean Method

⭐ MOST USED METHOD IN EXAMS ⭐

Formula:

Mean ($\bar{x}$) = a + (Σfidi / Σfi)

Where:

- a = Assumed mean (any class mark, usually middle one)
- $d_i = x_i - a$ (deviation from assumed mean)
- fi = Frequency

Steps:

1. Choose assumed mean (a) - usually central class mark
2. Find deviation: $d_i = x_i - a$ for each class
3. Calculate fidi for each class
4. Find Σfidi
5. Apply formula: Mean = a + (Σfidi / Σfi)

Example:

Class	fi	xi	di = xi - 25	fidi
0-10	5	5	-20	-100
10-20	8	15	-10	-80
20-30	12	25	0	0
30-40	6	35	10	60
	31			-120

a = 25 (assumed mean)

Σfidi = -120

Σfi = 31

Mean = a + (Σfidi / Σfi)

   = 25 + (-120/31)

   = 25 - 3.87

   = 21.13

When to use: When class marks are large numbers

3.4 Method 3: Step Deviation Method

Formula:

Mean ($\bar{x}$) = a + [(Σfiui / Σfi) × h]

Where:

- a = Assumed mean
- ui = (xi - a) / h (deviation units)
- h = Class size

Steps:

1. Choose assumed mean (a)
2. Find h (class size)
3. Calculate ui = (xi - a)/h for each class

4. **Calculate fiui**
5. **Apply formula**

Example:

Class	fi	xi	ui = (xi-25)/10	fiui
0-10	5	5	-2	-10
10-20	8	15	-1	-8
20-30	12	25	0	0
30-40	6	35	1	6
	31			-12

h = 10, a = 25

$\Sigma fiui$ = -12, Σfi = 31

Mean = a + [($\Sigma fiui$ / Σfi) × h]

= 25 + [(-12/31) × 10]

= 25 + (-3.87)

= 21.13

When to use: **When class marks are large AND class size is constant**

⭐ **EXAM TIP:** **Step deviation method is FASTEST when applicable!**

4. MEDIAN OF GROUPED DATA

⭐ **VERY IMPORTANT** ⭐

4.1 What is Median?

Median is the middle-most value when data is arranged in order.

It divides data into two equal parts:

- **50% observations below median**
- **50% observations above median**

4.2 Finding Median for Grouped Data

Formula:

Median = l + [(n/2 - cf)/f] × h

Where:

- **l = Lower limit of median class**
- **n = Total frequency (Σfi)**
- **cf = Cumulative frequency of class preceding median class**
- **f = Frequency of median class**
- **h = Class size**

Steps:

Step 1: Find n (total frequency)

Step 2: Calculate n/2

Step 3: Prepare cumulative frequency table

Step 4: Find median class

- **The class whose cumulative frequency is just greater than or equal to n/2**

Step 5: Identify:

- **l (lower limit of median class)**
- **cf (cumulative frequency before median class)**
- **f (frequency of median class)**
- **h (class size)**

Step 6: Apply formula

4.3 Example

Data:

Class	Frequency (fi)	Cumulative Frequency
0-10	5	5
10-20	8	13
20-30	12	25
30-40	6	31

Step 1: n = Σfi = 31

Step 2: n/2 = 31/2 = 15.5

Step 3: Cumulative frequency column already prepared

Step 4: Median class = 20-30

    (cf = 25 is just greater than 15.5)

Step 5: For median class 20-30:

    l = 20

    cf = 13 (cumulative frequency before 20-30)

    f = 12 (frequency of median class)

    h = 10 (class size)

Step 6: Median = l + [(n/2 - cf)/f] × h

    = 20 + [(15.5 - 13)/12] × 10

    = 20 + (2.5/12) × 10

    = 20 + 2.08

    = 22.08

Answer: Median = 22.08

5. MODE OF GROUPED DATA

⭐ IMPORTANT ⭐

5.1 What is Mode?

Mode is the value that appears most frequently in the data.

For grouped data, modal class is the class with highest frequency.

5.2 Formula for Mode

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

Where:

- l = Lower limit of modal class
- f_1 = Frequency of modal class
- f_0 = Frequency of class preceding modal class
- f_2 = Frequency of class succeeding modal class
- h = Class size

Steps:

Step 1: Find modal class (class with highest frequency)

Step 2: Identify:

- l (lower limit)
- f_1 (frequency of modal class)
- f_0 (frequency before modal class)
- f_2 (frequency after modal class)
- h (class size)

Step 3: Apply formula

5.3 Example

Data:

Class	Frequency
0-10	5
10-20	8
20-30	12
30-40	6

Step 1: Modal class = 20-30 (frequency 12 is highest)

Step 2: For modal class 20-30:

$l = 20$

$f_1 = 12$ (frequency of 20-30)

$f_0 = 8$ (frequency of 10-20)

$f_2 = 6$ (frequency of 30-40)

$h = 10$

Step 3: Mode = $l + [(f_1 - f_0)/(2f_1 - f_0 - f_2)] \times h$

$= 20 + [(12 - 8)/(2 \times 12 - 8 - 6)] \times 10$

$= 20 + [4/(24 - 14)] \times 10$

$= 20 + (4/10) \times 10$

$= 20 + 4$

$= 24$

Answer: Mode = 24

6. RELATIONSHIP BETWEEN MEAN, MEDIAN, AND MODE

6.1 Empirical Relationship

For moderately skewed distributions:

3 Median = Mode + 2 Mean

OR

Mode = 3 Median - 2 Mean

This relationship is VERY IMPORTANT for exam questions!

6.2 Example

If Mean = 21 and Median = 22, find Mode.

Mode = 3 Median - 2 Mean

$= 3(22) - 2(21)$

$= 66 - 42$

= 24

Answer: **Mode = 24**

7. CUMULATIVE FREQUENCY CURVES (OGIVES)

7.1 What is an Ogive?

An ogive is a graphical representation of cumulative frequency distribution.

Two types:

1. **Less than ogive**
2. **More than ogive**

7.2 Less Than Ogive

Steps:

1. **Convert data to "less than" cumulative frequency**
2. **Plot: Upper limit vs Cumulative frequency**
3. **Join points with smooth curve**

Example:

Class	Frequency	Upper Limit	Less than cf
0-10	5	10	5
10-20	8	20	13
20-30	12	30	25
30-40	6	40	31

Plot: **(10, 5), (20, 13), (30, 25), (40, 31)**

Shape: **Rising curve (increasing)**

7.3 More Than Ogive

Steps:

1. **Convert data to "more than" cumulative frequency**
2. **Plot: Lower limit vs More than cf**
3. **Join points with smooth curve**

Example:

Class	Frequency	Lower Limit	More than cf
0-10	5	0	31
10-20	8	10	26
20-30	12	20	18
30-40	6	30	6

Plot: (0, 31), (10, 26), (20, 18), (30, 6)

Shape: Falling curve (decreasing)

7.4 Finding Median from Ogive

Method:

1. **Draw both ogives on same graph**
2. **The point where they intersect**
3. **Draw perpendicular to x-axis**
4. **The x-coordinate is the median**

8. IMPORTANT FORMULAS (QUICK REFERENCE)

Mean

Direct Method:

Mean = $\Sigma f_i x_i / \Sigma f_i$

Assumed Mean Method:

Mean = $a + (\Sigma f_i d_i / \Sigma f_i)$

where **$d_i = x_i - a$**

Step Deviation Method:

Mean = $a + [(\Sigma f_i u_i / \Sigma f_i) \times h]$

where **$u_i = (x_i - a)/h$**

Median

Median = $l + [(n/2 - cf)/f] \times h$

Mode

Mode = $l + [(f_1 - f_0)/(2f_1 - f_0 - f_2)] \times h$

Empirical Relationship

Mode = 3 Median - 2 Mean

9. PROBLEM-SOLVING STRATEGY

For Mean:

1. **Choose appropriate method (step deviation is fastest)**
2. **Make proper table**
3. **Calculate step by step**
4. **Verify calculation**

For Median:

1. **Always make cumulative frequency column**
2. **Find n/2 first**
3. **Identify median class carefully**
4. **Apply formula with correct values**

For Mode:

1. **Identify modal class (highest frequency)**
2. **Note f_0, f_1, f_2 carefully**
3. **Don't confuse with median class**

10. QUICK REVISION POINTS

Mean:

- **Average of all values**
- **Three methods: Direct, Assumed, Step Deviation**
- **Use step deviation for large values with constant class size**

Median:

- **Middle value**
- **Need cumulative frequency**
- **Median class: cf just > n/2**
- **Formula: $l + [(n/2 - cf)/f] \times h$**

Mode:

- **Most frequent value**
- **Modal class: highest frequency**
- **Formula: $l + [(f_1 - f_0)/(2f_1 - f_0 - f_2)] \times h$**

Relationship:

- **Mode = 3 Median - 2 Mean**

Ogive:

- **Less than: upper limit, rising curve**
- **More than: lower limit, falling curve**
- **Intersection gives median**

11. PRACTICE CHECKLIST

Before Exam:

☐ **Practice mean by all 3 methods (10 problems each)**

☐ **Practice median problems (15 problems)**

☐ **Practice mode problems (10 problems)**

☐ **Learn empirical relationship**

☐ **Practice ogive drawing (5 graphs)**

☐ **Solve all NCERT Exercise 13.1 (mean)**

☐ **Solve all NCERT Exercise 13.2 (median)**

☐ **Solve all NCERT Exercise 13.3 (mode)**

☐ **Practice missing frequency problems (5)**

☐ **Revise all formulas**

☐ **Solve previous year questions (last 5 years)**

☐ **Time yourself on 5 complete problems**

All the Best for Board Exams 2026! 🌟

StudyByU