# Numerical Descriptive Measures

# Summary Definitions

- The **central tendency** is the extent to which all the **data values group around a typical or central value**.

- The **variation** is the amount of **dispersion**, or **scattering**, of values

- The **shape** is the pattern of the distribution of values from the **lowest value to the highest** value.

# Measures of Central Tendency: The Mean

- The arithmetic mean (often just called "mean") is the most common measure of central tendency
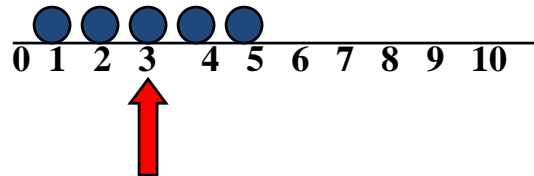
- For a sample of size n:

Pronounced x-bar

The i$^{th}$ value

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Sample size

Observed values

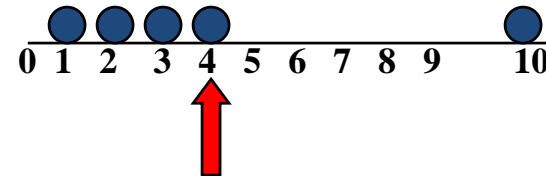# Measures of Central Tendency: The Mean

*(continued)*

- The most common measure of central tendency

- Mean = sum of values divided by the number of values

- Affected by extreme values (outliers)



**Mean = 3**

$$\frac{1+2+3+4+5}{5} = \frac{15}{5} = 3$$

**Mean = 4**

$$\frac{1+2+3+4+10}{5} = \frac{20}{5} = 4$$

# Mean for Grouped Data

Formula for Mean is given by $\overline{X} = \dfrac{\sum f(X)}{n}$

Where

$\overline{X}$ = Mean

$\sum f(X)$ = Sum of cross products of frequency in each class with midpoint X of each class

n = Total number of observations (Total frequency) = $\sum f$

# Mean for Grouped Data Example

Find the arithmetic mean for the following continuous frequency distribution:

| Class | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 |
|-------|-----|-----|-----|-----|-----|-----|
| Frequency | 1 | 4 | 8 | 7 | 3 | 2 |

# Solution for the Example

|   | A | B | C | D |
|---|---|---|---|---|
| 1 | Class | X (mid pt) | f | fX |
| 2 | 0-1 | 0.5 | 1 | 0.5 |
| 3 | 1-2 | 1.5 | 4 | 6.0 |
| 4 | 2-3 | 2.5 | 8 | 20.0 |
| 5 | 3-4 | 3.5 | 7 | 24.5 |
| 6 | 4-5 | 4.5 | 3 | 13.5 |
| 7 | 5-6 | 5.5 | 2 | 11.0 |
| 8 | Totals | | 25 | 75.5 |
| 9 | Mean | | | 3.02 |

Applying the formula $\bar{X} = \dfrac{\sum f(X)}{n}$ $= 75.5/25 = 3.02$

# Mean

| Class interval | | f |
|---|---|---|
| 0 | 49.99 | 78 |
| 50 | 99.99 | 123 |
| 100 | 149.99 | 187 |
| 150 | 199.99 | 82 |
| 200 | 249.99 | 51 |
| 250 | 299.99 | 47 |
| 300 | 349.99 | 13 |
| 350 | 399.99 | 9 |
| 400 | 449.99 | 6 |
| 450 | 499.99 | 4 |
| | | 600 |

IIT ROORKEE    NPTEL ONLINE CERTIFICATION COURSE

| Class interval | | f |
|---|---|---|
| 0 | 49.99 | 78 |
| 50 | 99.99 | 123 |
| 100 | 149.99 | 187 |
| 150 | 199.99 | 82 |
| 200 | 249.99 | 51 |
| 250 | 299.99 | 47 |
| 300 | 349.99 | 13 |
| 350 | 399.99 | 9 |
| 400 | 449.99 | 6 |
| 450 | 499.99 | 4 |
| | | 600 |

By taking mid values as 25, 75,… 475.

$$\overline{X} = \frac{\sum f(X)}{n}$$

Mean: 142.25

$f(X) = 85350$

$n = 600$

Mean using coding:

| Class | f |
|-------|---|
| 0-7 | 2 |
| 8-15 | 6 |
| 16-23 | 3 |
| 24-31 | 5 |
| 32-39 | 2 |
| 40-47 | 2 |

**Mean=x0 + w \* (Summation of u\*f))**
                              **n**

w=numerical width of class interval
X0=value of midpoint assigned code 0

Mean using coding:

| Class | mid | f | Code (u) | u*f |
|-------|-----|---|----------|-----|
| 0-7 | 3.5 | 2 | -2 | -4 |
| 8-15 | 11.5 | 6 | -1 | -6 |
| 16-23 | 19.5 | 3 | 0 | |
| 24-31 | ?? | 5 | 1 | ?? |
| 32-39 | ?? | 2 | 2 | 4 |
| 40-47 | 43.5 | 2 | 3 | 6 |
| | | 20 | | 5 |

**Mean=x0 + w * (Summation of u*f))**

$$\frac{}{n}$$

$= 19.5 + 8 * (5) / (20) = 21.5$

w=numerical width of class interval
X0=value of midpoint assigned code 0