Statistical Descriptions
Mean, Median, Mode

Mean

- Mean is the sum of all the values in the data set divided by the number of values in the data set. It is also called the Arithmetic Average.

- Mean is denoted as $\bar{x}$ and is read as **x bar**.

$$\text{Mean} (\bar{x}) = \frac{\text{Sum of Values}}{\text{Number of Values}}$$

- If $x_1, x_2, x_3, \ldots, x_n$ are the values of a data set then the mean is calculated as:

$$\bar{x} = (x_1 + x_2 + x_3 + \ldots + x_n) / n$$

Numerical

- **Example: Find the mean of data sets 10, 30, 40, 20, and 50.**

ANS:

Mean of the data 10, 30, 40, 20, 50 is

Mean = (sum of all values) / (number of values)

Mean = (10 + 30 + 40 + 20+ 50) / 5 = 30

Median

- A Median is a middle value for sorted data. The sorting of the data can be done either in ascending order or descending order. A median divides the data into two equal halves.

- Its denoted by "M".

Median (n = odd number),

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

Median (n = even number),

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

Numerical

Find the median of given data set 30, 40, 10, 20, and 50.

Solution:

Median of the data 30, 40, 10, 20, 50 is,

Step 1: Order the given data in ascending order as:

10, 20, 30, 40, 50

Step 2: Check n (number of terms of data set) is even or odd and find the median of the data with respective 'n' value.

Step 3: Here, n = 5 (odd)

Median = $[(n + 1)/2]^{th}$ term

Median = $[(5 + 1)/2]^{th}$ term

= 30

Mode

- A mode is the most frequent value or item of the data set.
- A data set can generally have one or more than one mode value.
- In statistical notation, the symbol "Z" is commonly used to represent the mode of a dataset.

Mode = Highest Frequency Term

Numerical

- **Find the mode of the given data set 1, 2, 2, 2, 3, 3, 4, 5.**
- **Solution:**

Given set is {1, 2, 2, 2, 3, 3, 4, 5}

As the above data set is arranged in ascending order.

By observing the above data set we can say that,

Mode = 2

Numerical -1

- Find the mean, median, mode, and range for the given data

190, 153, 168, 179, 194, 153, 165, 187, 190, 170, 165, 189, 185, 153, 147, 161, 127, 180

Solution 1

For Mean:

190, 153, 168, 179, 194, 153, 165, 187, 190, 170, 165, 189, 185, 153, 147, 161, 127, 180

Number of observations = 18

Mean = (Sum of observations) / (Number of observations)

      = (190+153+168+179+194+153+165+187+190+170+165+189+185+153+147 +161+127+180) / 18

      = 2871/18

      = 159.5

Therefore, the mean is 159.5

Sol 1

For Median:

The ascending order of given observations is,

127, 147, 153, 153, 153, 161, 165, 165, 168, 170, 179, 180, 185, 187, 189, 190, 190, 194

Here, n = 18

Median = 1/2 [(n/2) + (n/2 + 1)]th observation
= 1/2 [9 + 10]th observation
= 1/2 (168 + 170)
= 338/2
= 169

Thus, the median is 169

Sol 1

For Mode:

The number with the highest frequency = 153

Thus, mode = 53

For Range:

Range = Highest value – Lowest value

        = 194 – 127

        = 67

Variance(σ^2)

- Sample Variance = Sum of Squared Differences / Number of Observations - 1
- Population Variance = Sum of Squared Differences / Number of Observations
- Variance = (Standard Deviation)^2

$$S^2 = \frac{\Sigma(x_i - \overline{x})^2}{n-1}$$

X – Sample Average
x_i – Individual Population Value
n – Total Number of Sample
S^2 – Variance of Sample

$$\sigma^2 = \frac{\Sigma(x_i - \mu)^2}{n}$$

μ – Population Average
xi – Individual Population Value
n – Total Number of Population
σ^2 – Variance of Population

Standard Deviation(σ / s)

- Standard Deviation = $\sqrt{}$ Variance

Population	Sample
$\sigma = \sqrt{\dfrac{\Sigma(x_i-\mu)^2}{n}}$	$S = \sqrt{\dfrac{\Sigma(x_i-\overline{x})^2}{n-1}}$
μ - Population Average x_i - Individual Population Value n - Total Number of Population	X - Sample Average x_i - Individual Population Value n - Total Number of Sample

Numerical 2

- Example 1: Calculate the variance of the sample data: 7, 11, 15, 19, 24. as well as find the Standard deviation.

Solution 2

- We have the data, 7, 11, 15, 19, 24

mean of the data= $\bar{x}$ = (7 + 11 + 15 + 19 + 24)/5   = 76/5   = 15.2

Using the formula for variance we get,

$$\sigma2 = \sum (xi – \bar{x})2/(n – 1)$$

$$= (67.24 + 17.64 + 0.04 + 14.44 + 77.44)/(5 – 1)$$

$$= 176.8/4$$

$$= 44.2$$

Standard deviation= √ Variance = √ 44.2 = 6.64

Numerical 3

- Example 2: Calculate the number of observations if the variance of data is 12 and the sum of squared differences of data from the mean is 156.
  - Find the Standard deviation also.

Solution 3

- We have,$(x_i - \bar{x})2 = 156$

$\sigma2 = 12$

Using the formula for variance we get,$\sigma2 = \sum (x_i - \bar{x})2/n$

$12 = 156/n$

$n = 156/12$

$n = 13$

Standard deviation $= \sqrt{} \text{ Variance} = \sqrt{} 12 = 3.46$

Numerical 4

- Example: Find Standard Deviation of data set. X = {2, 3, 4, 5, 6}

Solution 4

- Given,

  n = 5

  xi = {2, 3, 4, 5, 6}

  Mean(μ) = (Sum of Observations)/(Number of Observations)

  μ = (2 + 3 + 4 + 5 + 6)/ 5 ⇒ μ = 4

  $σ2 = ∑ (xi – \bar{x})2/n$

  ⇒ σ2 = 10/5 = 2

  Thus, σ = √(2) = 1.414

Statistical Descriptions For Grouped Data

For Grouped Data

- Mean = Σ (f * m) / Σ (f)
- Median(M) = l + {h x (N/2 – cf)/f}

    Where, l = lower limit of median class.

    h = width of median class.

    f = frequency of median class,

    cf = cumulative frequency of the class preceding the median class.

    N = ∑fi

For Grouped Data

- Mode(Mo) = L + h{(fk – fk-1)/(2fk – fk-1 – fk+1)}

    L = lower limit of the modal class interval.

    fk = frequency of the modal class.

    fk-1= frequency of the class preceding the modal class.

    fk+1 = frequency of the class succeeding the modal class.

    h = width of the class interval.

Problem 1

- Find the mean of the following data.

Class Interval	0-10	10-20	20-30	30-40	40-50
Frequency	12	16	6	7	9

Sol P - 1

- Mean = $\sum(f_i.x_i)/\sum f_i$

Mean = $\sum(f_i.x_i)/\sum f_i$ = 1100/50 = 22

Class Interval	Frequency f_i	Class Mark x_i	$(f_i.x_i)$
0-10	12	5	60
10-20	16	15	240
20-30	6	25	150
30-40	7	35	245
40-50	9	45	405
	$\sum f_i=50$		$\sum f_i.x_i=1100$