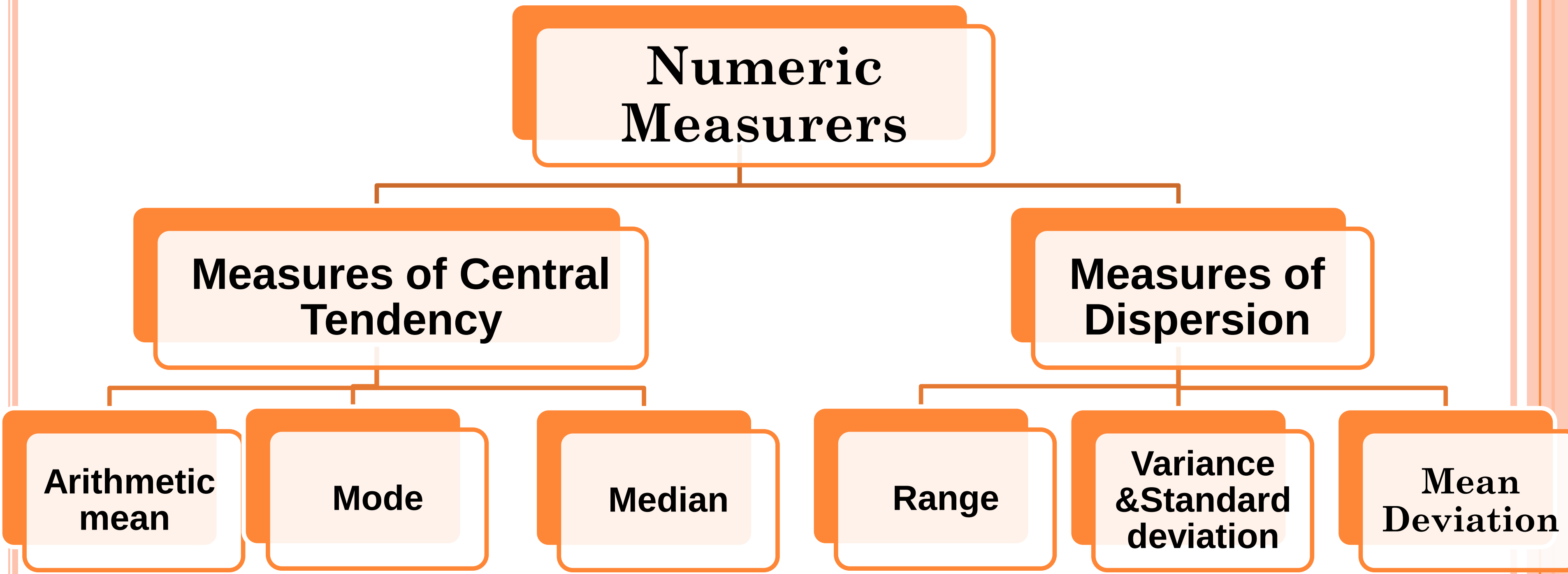
Chapter (3)

Describing Data Numerical Measures

Examples


Numeric Measurers
Measures of Central Tendency
Measures of Dispersion
Arithmetic mean
Mode
Median
Range
Variance &Standard deviation
Mean Deviation

Example (1)

If King Saud University (girls section) has 10 coffee shops on the campus and if their purchases (in thousands) per month are as follows:

10, 20, 30, 40, 40, 41, 45, 50, 52, 60

Find the Population mean.

Solution:

$$\mu = \frac{\sum_{i=1}^{10} X_i}{10} = \frac{10 + 20 + \cdots + 60}{10} = \frac{388}{10} = 38.8$$

Example (2)

The following are ages (in years) for a sample of students from school A:

15, 10, 18, 17, 18, 11, 11, 15, 14,13,11,10

Find the sample mean.

Solution:

$$\bar{X} = \frac{\sum_{i=1}^{12} X_i}{12} = \frac{15+10+\cdots.+10}{12} = \frac{136}{12} = \mathbf{13.58}$$

Example (3)

The average age of 5 women in a group is 27 years. If two other women aged 40 and 28 years join the group, find, in years, the new average age of the group of women.

Solution:

The average age of 5 women = $\bar{X} = \mathbf{27}$

$$\sum X = 135$$

The average age of 7 women = $\frac{135+40+28}{7}$=29

Mean for grouped data

1- The Discrete data

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i}$$

X: represents any particular value .

f : is the frequency in each value.

Example (6)

Table shows the distribution of households by the number of children

Number of children(X)	0	1	2	3	4	5
Number of households(F)	5	7	8	5	3	4

Calculate the arithmetic mean of the number of children

Solution:

							Total
Number of children(X)	0	1	2	3	4	5	
Number of households(F)	5	7	8	5	3	4	32
FX	0	7	16	15	12	20	70

$$\bar{X} = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i} = \frac{70}{32} = 2.19 = 3$$

2-The Arithmetic Mean of Grouped Data (**Continuous data)**

Example (7)

The following table gives the marks of a sample of 30 students.

marks	**(2 -4)**	**(4-6)**	**[6-8)**	**(8-10)**	**(10-12)**	**(12-14)**
frequency	**3**	**6**	**8**	**7**	**4**	**2**

Find the mean

Solution:

marks	Frequency (f)	Midpoint (m)	fm
[2-4)	3	3	9
[4-6)	6	5	30
[6-8)	8	7	56
[8-10)	7	9	63
[10-12)	4	11	44
[12-14]	2	13	26
Total	$\sum f = 30$		$\sum mf$ =228

$$\bar{X} = \frac{\sum fm}{n} = \frac{228}{30} = 7.6$$

The Median

The median is the midpoint of the values after they have been ordered from the smallest to the largest or from the largest to the smallest.

How do you find the median when (n) is odd?

1- Arrange all values (N) from smallest to largest

2- Find it by counting {(n+1) /2} observations up from the bottom

3- The median is the center of the list

Example (8)

Find the median for the sample values:

3, 5, 1, 6, 7, 4, 8, 7, 5, 10, 4, 7, 8, 9, 2

Solution: Arranging the data in ascending order gives:

1, 2 , 3 , 4 , 4 , 5 , 5 , 6 , 7 , 7 , 7 , 8 , 8 , 9 ,10

n =15 "odd"

The rank of median = (15+1)/2 =8

Median : the value of the observation of order 8 = 6

How do you find the median when (n) is even?

1- Arrange all values (N) from smallest to largest

2- Find it by counting $\{R_1 = (n/2) \quad R_2 = (n/2) + 1)\}$ observations up from the bottom

3- The median is the average of the center two values

Example (9)

Find the median for the sample values:

3, 5, 1, 6, 7, 4, 8, 7, 5, 4, 7, 8, 9, 2

Solution: Arranging the data in ascending order gives:

1, 2 , 3 , 4 , 4 , 5 , 5 , 6 , 7 , 7 , 7 , 8 , 8 , 9

n =14 "even"

$R_1 = n/2 = 14/2 = 7$, $R_2 = (n/2) + 1 = 7+1=8$

The value of the observation of ordered 7 is 5

The value of the observation of ordered 8 is 6

Then the median is the average of these two values:

Median= = (5+6)/2 = 5.5

The Mode

- The mode is the value of the observation that appears most frequently.
- If all values are different or have the same frequency, there is no mode.
- A set of values may have more than one mode.

Example (10):

Find the mode of the following data:

Group (1): 12, 15, 18, 17, 15, 14, 13, 15

Group (2): 12, 13, 18, 17, 15, 14, 13, 15

Group (3): 12, 10, 18, 17, 11, 14, 13, 15

Group (4): 12, 12, 18, 18, 11, 11, 13, 13

Solution:

The mode of group (1) is **15**

The mode of group (2) is **13 and 15**

There is no mode in group **(3)**

There is no mode in group **(4)**

A comparison of the properties of measures of central tendency

Mode

- It can be computed for open-ended tables
- It can found for both quantitative and qualitative variables.
- A set of values may have more than one mode.
- Mode depends on the value of the most frequent.
- The mode cannot be distorted by extreme values

Median

- It can be computed for an open-ended table.
- The median can only found for quantitative and qualitative (ordinal) variables.
- A set of data has a unique median
- median in his account depends only on the value that mediates data
- The median cannot be distorted by extreme values

Mean

- It is difficult to compute the mean from an open-ended table.
- The mean can only be found for quantitative variables
- A set of data has a unique mean
- All the values are included in computing the mean.
- The mean can be distorted by extreme values

The Relative Positions of the Mean, Median and the Mode

Example (11)

The following are the grades a professor gave on the first test in a statistics class: 52,61,74,75,82,83,86,87,88 and 90 . Distribution of grades is :

(A) Negatively skewed	(B) Bimodal
(C) Normally distributed	(D) Positively skewed

Mean = 77.8

Median =82.5

Mode = None

Median > Mean

Range

The different between the largest and the smallest values

Range = largest value - smallest value

Example (12)

Find the range for the following data:

20, 40, 45, 70, 99, 50, 30, 31, 60, 34

Solution:

20, 30, 31, 34, 40, 45, 50, 60, 70, 99

Largest value = 99

Smallest value = 20

Range = 99 – 20 = 79

Example (13)

Find the range for the following table:

Age (in years)	(15-20)	(20-25)	(25-30)	(30-35)	(35-40)	(40-45)
Frequency	3	6	10	7	6	2

Solution:

The midpoint of the last class =42.5

The midpoint of the first class =17.5

R= 42.5 -17.5=25

Mean Deviation :

The mean of the absolute deviation of a set of data about the data's mean .

$$MD = \frac{\sum |X - \bar{X}|}{n}$$

Example (14)

If the sample is removed from the factory workers of foodstuffs size 5 workers, and record the number of years of experience, and were as follows :

9 5 10 13 8

Calculate the mean deviation

Solution :

$$\bar{X} = 9$$

						Total
X_i	9	5	10	13	8	45
$(X_i - \bar{X})$	0	-4	1	4	-1	0
$\lvert X_i - \bar{X} \rvert$	0	4	1	4	1	10

Mean Deviation : $MD = \frac{\sum |X - \bar{X}|}{n} = \frac{10}{5} = 2$

Variance and Standard Deviation (for ungrouped data)

Variance:The arithmetic mean of the squared deviation from the mean.

Population variance

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$$

Sample variance

$$S^2 = \frac{\sum(X-\bar{X})^2}{n-1}$$

Standard Deviation :The square root of the variance

Population Standard Deviation :

$$\sigma = \sqrt{\frac{\sum(X-\mu)^2}{N}}$$

Sample Standard Deviation:

$$S = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$$

Example (15)

The following are age (in years) of a sample of babies from clinic A:

2, 3, 4, 5, 4, 5, 6, 3

Find the variance and Standard Deviation

Solution:

$$\bar{X} = 4$$

									Total
X_i	2	3	4	5	4	5	6	3	32
$(X_i - \bar{X})$	-2	-1	0	1	0	1	2	-1	0
$(X_i - \bar{X})^2$	4	1	0	1	0	1	4	1	12

Sample variance

$$S^2 = \frac{\sum(X-\bar{X})^2}{n-1} = \frac{12}{7} = 1.71$$

Sample Standard Deviation:

$$S = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}} = \sqrt{1.71} = 1.31$$

Variance and Standard Deviation

for Grouped Data

$$variance\ S^2 = \frac{\sum f(M-\bar{X})^2}{n-1}$$

$$Standard\ deviation \quad S = \sqrt{\frac{\sum f(M-\bar{X})^2}{n-1}}$$

M : is the midpoint of the class.

f : is the class frequency.

n : is the of observations in the sample .

$\bar{X}$: is the designation for the sample mean

Example (16)

The following table gives the marks of a sample of students.

marks	(4-6)	(6-8)	(8-10)	(10-12)	(12-14)	(14-16)
frequency	1	3	7	4	3	2

Find the variance.

Marks	f	Midpoint (mi)	$f_i m_i$	$(m_i - \overline{X})$	$(m_i - \overline{X})^2$	$f(m_i - \overline{X})^2$
(4-6)	1	5	5	-5.1	26.01	26.01
(6-8)	3	7	21	-3.1	9.61	28.83
(8-10)	7	9	63	-1.1	1.21	8.47
(10-12)	4	11	44	0.9	0.81	3.24
(12-14)	3	13	39	2.9	8.41	25.23
(14-16)	2	15	30	4.9	24.01	48.02
Total	20		202			139.8

$$\bar{X} = \frac{202}{20} = 10.1$$

$$variance = S^2 = \frac{\sum f(M - \bar{X})^2}{n - 1} = \frac{139.8}{20 - 1} = 7.36$$

$$Standard\ deviation = S = \sqrt{\frac{\sum f(M - \bar{X})^2}{n - 1}}$$

$$= \sqrt{\frac{139.8}{20 - 1}} = \sqrt{7.36} = 2.71$$

- **Other solution :**

$$S^2 = \frac{\sum m^2 f - \frac{(\sum mf)^2}{n}}{n-1}$$

$$S = \sqrt{\frac{\sum m^2 f - \frac{(\sum mf)^2}{n}}{n-1}}$$

Marks	f	Midpoint (m)	fm	F(m^2)
(4-6)	1	5	5	25
(6-8)	3	7	21	147
(8-10)	7	9	63	567
(10-12)	4	11	44	484
(12-14)	3	13	39	507
(14-16)	2	15	30	450
Total	20		202	2180

$$S^2 = \frac{2180 - 2040.2}{19} = 7.36$$

$$S = \sqrt{7.36} = 2.71$$