Data Description

3.1 Measures of Central Tendency

Usually used to describe the data congestion about a central value. The measures of central tendency include:

1. **Mean**
2. **Median**
3. **Mode**

3.1.1 Mean

❖ Arithmetic Mean

Is the sum of the values, divided by the total number of values. The symbol $\bar{X}$ represents the sample mean.

✓ **For unclassified (ungrouped) data :**

$$\bar{X} = \frac{\sum x}{n} = \frac{x_1 + x_2 + x_3 \ldots\ldots + x_n}{n}$$

- Where n represents the total numbers of values in the sample.
- Is used when data have common denominator.

Example 3.1: The following data represents the splitting tensile strength of cylindrical concrete samples, in (MPa). Find the arithmetic mean.

1.8, 2.2, 3.8, 4.1, 4.7

Solution

$$\bar{X} = \frac{\sum x}{n}$$

$$\bar{X} = \frac{1.8+2.2+3.8+4.1+4.7}{5}$$

= 3.32 MPa

- ✓ **For classified (grouped) data :**

$$\bar{X} = \frac{\sum f_i * y_i}{\sum f_i}$$

Mean can be calculated based on the following steps

1. Find classes of data (c)
2. Find frequencies of data (f_i), and sum of frequencies ($\sum f_i$)
3. Find class mark (y_i)
4. Find $f_i * y_i$
5. Find sum of $f_i * y_i$ ($\sum f_i * y_i$)

Example 3.2: Find the Arithmetic Mean for frequency distribution table below:

classes	Frequency (f_i)
1-1	3
2-2	5
3-3	7
4-4	7
5-5	4
6-7	4
	$\sum f_i$= 30

Solution

- ✓ **Find class mark** (y_i)
- ✓ Find $f_i * y_i$

C	f_i	y_i	$f_i * y_i$
1-1	3	1	3
2-2	5	2	10
3-3	7	3	21
4-4	7	4	28
5-5	4	5	20
6-7	4	6.5	26
	$\sum f_i = 30$		$\sum f_i * y_i = 108$

$$\bar{X} = \frac{\sum f_i * y_i}{\sum f_i}$$

$$= \frac{108}{30}$$

$$= 3.6$$

❖ **Harmonic Mean**

Is used when data have common numerator, and computed as follows:

For unclassified (ungrouped) data:

$$\bar{H} = \frac{n}{\sum \frac{1}{x_i}}$$

Example 3.3: Find the harmonic mean for the following data

x_i= 3, 5, 6, 6, 7, 10, 12

Solution

$$\bar{H} = \frac{n}{\sum \frac{1}{x_i}}$$

$$= \frac{7}{\frac{1}{3}+\frac{1}{5}+\frac{1}{6}+\frac{1}{6}+\frac{1}{7}+\frac{1}{10}+\frac{1}{12}}$$

$$= 5.87$$

✓ **For classified (grouped) data:**

✓ $$\bar{H} = \frac{\sum f_i}{\sum \frac{f_i}{y_i}}$$

Example 3.4: The data shown in the following frequency table represent the scores of 50 students. Calculate the harmonic mean of these scores.

The scores (classes)	Number of students (frequency)
30 – 39	4
40 – 49	6
50 – 59	8
60 – 69	12
70-79	9
80-89	7
90-99	4
	$\sum f_i$= 50

Solution

$$\bar{\bar{H}} = \frac{\sum f_i}{\sum \frac{f_i}{y_i}}$$

$$= \frac{50}{\frac{4}{34.5}+\frac{6}{44.5}+\frac{8}{54.5}+\frac{12}{64.5}+\frac{9}{74.5}+\frac{7}{84.5}+\frac{2}{94.5}}$$

$$=62.5$$

Example 3.5 A tourist purchases gasoline at three filling stations where the prices are 33.333, 25, 20 cent per gallon. What is the average price.

Solution

- ✓ If gallon is taken as a fixed unit, therefor the ratios would have common denominator and the average would be the arithmetic mean

$$\bar{X} = \frac{\sum x}{n}$$

$$= \frac{33.333+25+20}{3}$$

- ✓ While when cent is considered as fixed unit, therefore the ratios would have common numerator and the mean would be a harmonic mean

$$\bar{\bar{H}} = \frac{n}{\sum \frac{1}{x_i}}$$

$$= \frac{3}{\frac{1}{33.333}+\frac{1}{25}+\frac{1}{20}}$$

- ✓ The former would the correct estimator when the tourist to buy same number of gallons from each station. While the latter would be the correct estimator when the tourist to spend the same cents.

❖ Geometric Mean

Is used in averaging values that represents a rate of change.

- ✓ **For unclassified data :**

$$\bar{x_g} = (x_1 * x_2 * x_3 * x_4 * \dots\dots\dots\dots x_n)^{\frac{1}{n}}$$

Example 3.6 A value of 100 falls to 50 and subsequently raises to 100. The ratio of change is $1/2$, and 2. Use arithmetic mean and geometric mean to determine the average rate of change. Comment on the results.

Solution

- ✓ **Arithmetic mean**

$$\bar{X} = \frac{\sum x}{n}$$

$$= \frac{1/2+2}{2}$$

$$=1.25$$

- ✓ **Geometric mean**

$$\bar{x_g} = (x_1 * x_2)^{\frac{1}{n}}$$

$$= (1/2 * 2)^{\frac{1}{2}}$$

$$= 1$$

3.1.2 Median

- ✓ **Another important measure of central tendency is called the sample median.**
- ✓ **Is the midpoint of the data, the symbol MD represents the sample median.**
- ✓ **The data should be arranged in ascending or descending order.**

1. For unclassified data

- ✓ **For odd data**
- ➢ **MD =** $\frac{n+1}{2}$
- ✓ **For even data**
- ➢ **MD = average of values (** $\frac{n}{2}$, $\frac{n}{2}+1$**)**
- ▪ **Where n is represents the total numbers of values in the sample.**

Example 3.7 suppose the data set as the following: 1.7, 2.2, 3.9, 3.11, and 14.7. Find the sample median.

Solution

MD = $\frac{n+1}{2}$

$=\frac{5+1}{2}$

= 3 (third value)

- ✓ **The sample median is 3.9**

Example 3.8 Find the median for the data listed below:

72 60 72 40 80 63.

Solution

40 60 63 72 72 80

MD = average of values ($\frac{n}{2}$, $\frac{n}{2}+1$)

1. Find $\frac{n}{2}$

- ✓ **$\frac{6}{2}$ = 3**

2. Find $\frac{n}{2}+1$

- ✓ **$\frac{6}{2}+1$ = 4**

3. Take the average of the values 3, 4 from the data

$\frac{63+72}{2}$ = 67.5

- ✓ **The median is 67.5**

2. for classified data

MD= L_M+ ($\frac{\sum f_i}{2} - AF_i$ / f_m) *L

Where:

$\frac{\sum f_i}{2}$ Represents the frequency of the median class

L_M = lower real limit of the median class

f_m = frequency of the median class

L= length of the class

Example 3.9 **According to the frequency distribution table shown below, calculate the sample median.**

c	f_i
25-29	7
30-34	19
35-39	14
40-44	7
45-49	3
	$\sum f_i = 50$

Solution

c	f_i	Ascending F_i
25-29	7	7
30-34	19	26
35-39	14	40
40-44	7	47
45-49	3	50
	$\sum f_i = 50$	

- ✓ $\frac{\sum f_i}{2} = 25$
- ✓ The median class is (30-34)
- ✓ $AF_i = 7$
- ✓ $f_m = 19$
- ✓ L=5

$$MD = L_M + \left(\frac{\frac{\sum f_i}{2} - AF_i}{f_m}\right) * L$$

$$= 29.5 + \left(\frac{25-7}{19}\right) * 5$$

3. Mode: Is the most frequent value in the data set.

- ✓ **For unclassified data**

Example 3.10 Find the mode value of the following data:

1. **10, 7, 9, 8, 10, 6**
2. **3, 5, 4, 7, 5, 6**
3. **30, 20, 45, 25, 10, 30, 25**
4. **15, 9, 24, 16, 8, 28**

Solution

1. **Mode value is 10**
2. **Mode value is 5**
3. **Mode values are 25, and 30**
4. **∅**

- ✓ **For classified data**

$$M_o = L_M + \left(\frac{D_1}{D_1 + D_2}\right) * L$$

Where:

L_M **= lower real limit of the mode class**

D_1 **= the difference between the frequency of the mode class and the previous class**

D_2 **= the difference between the frequency of the mode class and the subsequent class**

L= length of class

Example 3.11 The data shown in the following frequency table represent the scores of 50 students. Calculate the mode of these scores.

The scores	Number of students
30 – 39	4
40 – 49	6
50 – 59	8
60 – 69	12
70-79	9
80-89	7
90-99	4
	$\sum f_i$= 50

Solution

- ✓ **The most frequent class is (60-69)**

$M_o = L_M + \left(\frac{D_1}{D_1 + D_2}\right) * L$

$= 59.5 + \left(\frac{(12-8)}{(12-8)+(12-9)}\right) * 10$

=

H.W 3.3 **The data below represents the ultimate load of sixteen RC beams tested under a four point load**

121, 123, 135, 211, 222, 217, 214, 140, 147, 139, 152, 151, 235, 243, 265, 239. Find the mean, the median, and the mode.