ذ

Chemical Engineering Department **Subject: Engineering Statistics**
Class: Third **Lecturer: Dr.Mohammad N. Fares**
Date : 2018-2019 **Chapter: Three**

MEASURES OF CENTRAL TENDENCY

3.2 Measures of Central Tendency:

In the study of a population with respect to one in which we are interested we may get a large number of observations. It is not possible to grasp any idea about the characteristic when we look at all the observations. So it is better to get one number for one group.

That number must be a good representative one for all the observations to give a clear picture of that characteristic. Such representative number can be a central value for all these observations. This central value is called a measure of central tendency or an average or a measure of locations. There are five averages. Among them the mean , geometric mean and harmonic mean are called special averages , median and mode are called simple averages .

1- Mean :

Arithmetic mean or simply the mean of a variable is defined as the sum of the observations divided by the number of observations. If the variable x assumes n values x_1, x_2, x_3 then the mean, x, is given by

$$\overline{x} = \frac{x_1 + x_1 + x_1 + \ldots\ldots + x_n}{n}$$

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

This formula is for the ungrouped or raw data.

Example 1 :

Calculate the mean for 2, 4, 6, 8, 10

Solution:

$$\bar{x} = \frac{2+4+6+8+10}{5} = 6$$

2- Harmonic mean (H.M) :

Harmonic mean of a set of observations is defined as the reciprocal of the arithmetic average of the reciprocal of the given values. If x_1, x_2,…, x_n are n observations,

$$H.M = \frac{n}{\sum_{i=1}^{n}\left(\frac{1}{x_i}\right)}$$

For a frequency distribution

$$H.M = \frac{N}{\sum_{i=1}^{n} f.\left(\frac{1}{x_i}\right)}$$

Example 2:

From the given data calculate H.M 5,10,17,24,30

X	$\frac{1}{x}$
5	0.2000
10	0.1000
17	0.0588
24	0.0417
30	0.0333
Total	0.4338

$$H.M = \frac{n}{\Sigma\left[\frac{1}{x}\right]}$$

$$= \frac{5}{0.4338} = 11.526$$

Example 3:

The marks secured by some students of a class are given below. Calculate the harmonic mean.

Marks	20	21	22	23	24	25
Number of Students	4	2	7	1	3	1

Solution:

Marks X	No of students f	$\frac{1}{x}$	$f(\frac{1}{x})$
20	4	0.0500	0.2000
21	2	0.0476	0.0952
22	7	0.0454	0.3178
23	1	0.0435	0.0435
24	3	0.0417	0.1251
25	1	0.0400	0.0400
	18		0.8216

$$H.M = \frac{N}{\Sigma f\left[\frac{1}{x}\right]}$$

$$= \frac{18}{0.1968} = 21.91$$

3- Geometric mean :

The geometric mean of a series containing n observations is the n^{th} root of the product of the values. If $x_1, x_2, \ldots, x_n$ are observations then

$$G.M = \sqrt[n]{x_1 . x_2 \ldots x_n}$$

$$= (x_1 . x_2 \ldots x_n)^{1/n}$$

$$\log GM = \frac{1}{n} \log(x_1 . x_2 \ldots x_n)$$

$$= \frac{1}{n} (\log x_1 + \log x_2 + \ldots + \log x_n$$

$$= \frac{\sum \log x_i}{n}$$

$$GM = \text{Antilog} \frac{\sum \log x_i}{n}$$

For grouped data

$$GM = \text{Antilog} \left[\frac{\sum f \log x_i}{N}\right]$$

Example 4:

Calculate the geometric mean of the following series of monthly income of a batch of families 180,250,490,1400,1050.

x	logx
180	2.2553
250	2.3979
490	2.6902
1400	3.1461
1050	3.0212
	13.5107

$$GM = \text{Antilog}\left[\frac{\sum \log x}{n}\right]$$

$$= \text{Antilog}\ \frac{13.5107}{5}$$

$$= \text{Antilog}\ 2.7021 \quad = 503.6$$

Example 5:

Calculate the average income per head from the data given below .Use geometric mean.

Class of people	Number of families	Monthly income per head (Rs)
Landlords	2	5000
Cultivators	100	400
Landless – labours	50	200
Money – lenders	4	3750
Office Assistants	6	3000
Shop keepers	8	750
Carpenters	6	600
Weavers	10	300

Solution:

Class of people	Annual income (Rs) X	Number of families (f)	Log x	f logx
Landlords	5000	2	3.6990	7.398
Cultivators	400	100	2.6021	260.210
Landless – labours	200	50	2.3010	115.050
Money – lenders	3750	4	3.5740	14.296
Office Assistants	3000	6	3.4771	20.863
Shop keepers	750	8	2.8751	23.2008
Carpenters	600	6	2.7782	16.669
Weavers	300	10	2.4771	24.771
		186		482.257

$$GM = \text{Antilog}\left[\frac{\sum f \log x}{N}\right]$$

$$= \text{Antilog}\left[\frac{482.257}{186}\right]$$

$$= \text{Antilog}\ (2.5928)$$

$$= \text{Rs } 391.50$$

4- Median :

The median is that value of the variable which divides the group into two equal parts, one part comprising all values greater, and the other, all values less than median. By formula

$$\text{Median} = \text{Md} = \left(\frac{n+1}{2}\right)^{th} \text{item.}$$

Example 6 :

When odd number of values are given. Find median for the following data

25, 18, 27, 10, 8, 30, 42, 20, 53

Solution:

Arranging the data in the increasing order 8, 10, 18, 20, 25, 27, 30, 42, 53

The middle value is the 5^{th} item i.e., 25 is the median

Using formula

$$\text{Md} = \left(\frac{n+1}{2}\right)^{th} \text{item.} = \left(\frac{9+1}{2}\right)^{th} \text{item.}$$

$$= \left(\frac{10}{2}\right)^{th} \text{item} = 5^{th} \text{item} = 25$$

<u>Example 7:</u>

When even number of values are given. Find median for the following data

5, 8, 12, 30, 18, 10, 2, 22

Solution:

Arranging the data in the increasing order 2, 5, 8, 10, 12, 18, 22, 30

Here median is the mean of the middle two items (ie) mean of (10,12) ie

$$= \left(\frac{10+12}{2}\right) = 11$$

$$\therefore \text{median} = 11.$$

ذ

Using the formula

$$\text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{item.} = \left(\frac{8+1}{2}\right)^{th} \text{item.} = \left(\frac{9}{2}\right)^{th} \text{item} = 4.5^{th} \text{ item}$$

$$= 4^{th} \text{ item} + \left(\frac{1}{2}\right)(5^{th} \text{ item} - 4^{th} \text{ item}) = 10 + \left(\frac{1}{2}\right)[12\text{-}10]$$

$$= 10 + \left(\frac{1}{2}\right) \times 2 = 10 + 1 = 11$$

Example 8:

The following table represents the marks obtained by a batch of 10 students in certain class tests in statistics and Accountancy.

Serial No	1	2	3	4	5	6	7	8	9	10
Marks (Statistics)	53	55	52	32	30	60	47	46	35	28
Marks (Accountancy)	57	45	24	31	25	84	43	80	32	72

Indicate in which subject is the level of knowledge higher ?

Solution:

For such question, median is the most suitable measure of central tendency. The mark in the two subjects are first arranged in increasing order as follows:

Serial No	1	2	3	4	5	6	7	8	9	10
Marks in Statistics	28	30	32	35	46	47	52	53	55	60
Marks in Accountancy	24	25	31	32	43	45	57	72	80	84

$$\text{Median} = \left(\frac{n+1}{2}\right)^{th} \text{item} = \left(\frac{10+1}{2}\right)^{th} \text{item} = 5.5^{th} \text{ item}$$

$$= \frac{Value\ of\ 5^{th}\ item + value\ of\ 6^{th}\ item}{2}$$

$$\text{Md (Statistics)} = \frac{46+47}{2} = 46.5$$

$$\text{Md (Accountancy)} = \frac{43+45}{2} = 44$$

There fore the level of knowledge in Statistics is higher than that in Accountancy.

5- Mode:

The mode refers to that value in a distribution, which occur most frequently. It is an actual value, which has the highest concentration of items in and around it.

Example 9:

2 , 7, 10, 15, 10, 17, 8, 10, 2

Mode = M_0=10

In some cases the mode may be absent while in some cases there may be more than one mode.

Example 10:

1- 12, 10, 15, 24, 30 (**no mode**)

2- 7, 10, 15, 12, 7, 14, 24, 10, 7, 20, 10

the modes are **7** and **10**

3.3 Mean vs Median vs Mode which measures the centre best?

Choosing which of these three measures to use in practice can sometimes seem like a difficult task. However if we understand a little about the relative merits of each we should at least be able to make an informed decision.

If the distribution is symmetric then

Mean = Median

If the distribution is Positively Skewed (to the right) then

Mean > Median

If the distribution is Negatively Skewed (to the left) then

Median > Mean

So the difference between the mean and median can be used to measure the skewness of a dataset.