Measures of central Tendency

কেন্দ্রীয় প্রবণতার পরিমাপ

☐ Define central tendency and measures of central tendency. What are the different measures of central tendency?

⇒ In a representative sample, the values of a series of data have a tendency to cluster around a certain point usually at the center of the series. The tendency of clustering the values around the center of the series is usually called tendency. And its numerical measures are called the measures of central tendency.

⇒ A measure of central tendency is a typical value around which others figure congregate.

Different measures of central tendency are as follows:

1. Arithmetic Mean (A.M) (গাণিতিক গড়)
2. Geometric Mean (G.M)
3. Harmonic Mean (H.M)
4. Median (Me) (মধ্যক)
5. Mode (Mo)

Arithmetic Mean / Average Mean (গাণিতিক গড়):

Arithmetic mean of a set of observations is the sum of all observations divided by the total number of observations. It is usually denoted by A.M or $\bar{x}$.
or, sum of set data divided by the number of data.

For ungroup data (অবিন্যস্ত)

$$\bar{x} = \frac{x_1 + x_2 + \dots x_n}{n}$$

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

For group data (বিন্যস্ত)

$$\bar{x} = \frac{f_1 x_1 + f_2 x_2 + \dots + f_n x_n}{f_1 + f_2 + \dots + f_n}$$

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

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

(Arithmetic Mean for ungroup data)

Example: 1 Find the mean, from the following data.

9, 15, 11, 12, 3, 5, 10, 20, 14, 6, 8, 8, 12, 12, 18, 15, 6, 9, 18, 11.

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

$$= \frac{9+15+11+12+3+5+10+20+14+6+8+8+12+12+18+15+6+9+18+11}{20}$$

$$= 11.1$$

(Arithmetic Mean for group data)

Example: 1

Number of games	Frequency
1-5	2
6-10	7
11-15	8
16-20	3

Solution:

Number of games	Frequency (f_i)	Mid point (x_i)	$f_i x_i$
1-5	2	3	6
6-10	7	8	56
11-15	8	13	104
16-20	3	18	54
	$\sum f_i = 20$ $/n$		$\sum f_i x_i = 220$

Arithmetic Mean, $\bar{x} = \dfrac{\sum_{i=1}^{n} f_i x_i}{\sum_{i=1}^{n} n}$

$= \dfrac{220}{20}$

$= 11$.

Example-2 You grow fifty baby carrots using special soil. You dig them up and measure their length (to the nearest mm) and group the results:

Length (mm)	Frequency
150 – 155	5
155 – 160	2
160 – 165	6
165 – 170	8
170 – 175	9
175 – 180	11
180 – 185	6
185 – 190	3

(i) Calculate arithmetic mean

(ii) Draw pie diagram.

Solution: (i)

Length (mm)	Midpoint (x_i)	Frequency (f_i)	$f_i x_i$
150–155	152.5	5	762.5
155–160	157.5	2	315
160–165	162.5	6	975
165–170	167.5	8	1340
170–175	172.5	9	1552.5
175–180	177.5	11	1952.5
180–185	182.5	6	1095
185–190	187.5	3	562.5
		$\Sigma f_i = 50$	$\Sigma f_i x_i = 8550$

$$\therefore \bar{x} = \frac{\Sigma f_i x_i}{\Sigma f_i} = \frac{8550}{50} = 171.1 \text{ mm}$$

Properties of Arithmetic Mean:

The mean is widely used measure of location. It has several important properties.

1. Every set of interval level and ratio level data has a mean.
2. All the data values are used in the calculation.
3. A set of data has only one mean. that is the mean is unique.
4. The mean is a useful measure for comparing two or more populations.
5. The sum of the deviations of each value from the mean will always be zero, that is

$$\Sigma (x - \bar{x}) = 0$$

Geometric Mean

The geometric mean of a set of n non-zero positive observations is the nth root of their product.

Let $x_1, x_2, x_3, \ldots x_n$ be the series of n observations in any statistical investigation.

For ungroup data,

$$G.M = (x_1 . x_2 \ldots x_n)^{1/2}$$

Example: 1 What is the geometric mean of 2, 8 and 4?

Solution: $G.M = (2 \times 4 \times 8)^{1/3}$

$= (64)^{1/3}$

$= (4^3)^{1/3}$

$= 4.$

Uses of GM:

① To find the rate of population growth and the rate of interest.

② In the construction of index number.

GM ও HM এ শুধু
ungroup data বানাবে

Harmonic Mean

The harmonic mean of a set of n non zero observations $x_1, x_2 \ldots x_n$ is a series is the reciprocal of the arithmetic mean of the reciprocal

For ungroup data

$$\bar{x} = \frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+\cdots\frac{1}{x_n}}$$

$$= \frac{n}{\Sigma\left(\frac{1}{x}\right)}$$

For Group data

$$\bar{x} = \frac{\Sigma f}{\Sigma\left(\frac{f}{x}\right)}$$

Example-1 Find the harmonic mean of the following data. 8, 9, 6, 11, 10, 5.

Solution: Given data $\{8, 9, 6, 11, 10, 5\}$

so, Harmonic mean $= \frac{6}{\frac{1}{8}+\frac{1}{9}+\frac{1}{6}+\frac{1}{11}+\frac{1}{10}+\frac{1}{5}}$

$$= \frac{6}{0.7936}$$

$$= 7.560$$

Uses of HM: It is used for calculating speed of automobiles.

Median (Me) (মধ্যক)

The Median is defined as the middle most observation when the observation are arranged in order of magnitude.

For ungroup data,

When n is odd, Median $= \left(\frac{n+1}{2}\right)$ th observation

When n is even, $Me = \left(\frac{\frac{n}{2} + \left(\frac{n}{2}+1\right)}{2}\right)$ th observation

For group data,

$$Me = L + \frac{\frac{n}{2} - F}{f} \times c$$

where, L = Lower limit of the median class

n = Total number of observations

Fc/F = Cumulative frequency of the class just proceeding the median class (মধ্যক শ্রেণীর ক্রমযোজিত গণসংখ্যার পূর্বের সংখ্যা)

f/f_m = Frequency of the median class

c = length of the median class

Example-1

Find the median for the following data: 15, 10, 13, 5, 20, 25 and 17.

Solution: Firstly we arranged the observation by ascending order of magnitude.

5, 10, 13, 15, 17, 20, 25

Here, total number is 7 which is odd.

$\therefore$ Median $= \left(\frac{n+1}{2}\right)$ th observation

$= \left(\frac{7+1}{2}\right)$ th observation

$=$ 4th observation

$= 15.$

Example-2 Find the Median for the following data.

15, 10, 13, 5, 20, 25.

Solution: Firstly, we arranged the ~~data~~ observations by ascending order of magnitude.

5, 10, 13, 15, 20, 25.

Here, total number is 6 which is even, so,

$$Me = \frac{\frac{n}{2} + \left(\frac{n}{2}+1\right)}{2} \text{ th observation}$$

$$= \frac{\left(\frac{6}{2}\right) + \left(\frac{6}{2}+1\right)}{2} \text{ th observation}$$

$$= \frac{\text{3th observation} + \text{4th observation}}{2}$$

$$= \frac{13+15}{2}$$

$$= 9.$$

Example-3 Calculate Median from the following data.

Number of games	Frequency
150-155	5
155-160	2
160-165	6
165-170	8
170-175	9
175-180	11
180-185	6
185-190	3

Solution:

Number of games	Frequency (f)	Cumulative frequency F.C
150-155	5	5
155-160	2	7
160-165	6	13
165-170	8	21
170-175	9	30
175-180	11	41
180-185	6	47
185-190	3	50
	$\Sigma f_i = 50$	

For group data,

$$Me = L + \frac{\frac{n}{2} - Fc}{f_m} \times C$$

$\therefore n = 50$ and $n/2 = 25$ we have seen from the table median class is 170-175,

$L = 170$ $\quad f_m = 9$

$Fc = 21$ $\quad C = 5$

$$Me = L + \frac{\frac{n}{2} - F_c}{f_m} \times c$$

$$= 170 + \frac{25 - 21}{9} \times 5$$

$$= 172.22 .$$

Uses of Median:

① It is the only average to be used while dealing with qualitative data, which cannot measured but still can be arranged in ascending or descending order in magnitude. To find the average intelligence or average honesty among a group of people.

② It is used to determining the typical values in the problem concerning wages, distribution of wealth etc.

Properties of Median:

① The median is a unique value, that is like the mean, there is only one median for a set of data.

② It is not influenced by extremely large or small values.

③ It can be computed for ratio level, interval level and ordinal level data.

(ii) Fifty percent of the observations are greater than the Median and fifty percent of the observation are less than the median.

Mode (ঘনিষ্ঠক)

The mode is that observation of the variable for which the frequency is maximum.

For ungroup data

M_0 = Maximum frequency

For group data

$$M_0 = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times c$$

where, L = Lower limit of the modal class

Δ_1 = Difference between frequency of the modal class and pre modal class

Δ_2 = Difference between frequency of the modal class and post modal class.

c = length of the modal class

Example-1

Find mode for the following data, 5, 15, 10, 15, 5, 10, 10, 20, 25, 15.

Solution: There are three 10's and three 15's.

∴ Mode is 10 and 15.

Example - 2 Calculate mode from the following data.

Length (mm)	Frequency
150-154	5
155-159	2
160-164	7
165-169	10
170-174	9
175-179	11
180-184	6
185-189	3

Solution:

Length (mm)	Frequency
150-154	5
155-159	2
160-164	7
165-169	10
170-174	9
175-179	11
180-184	6
185-189	3

For group data, $M_0 = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C$

Modal class is 165-169,

$L = 165$

$C = 5$

$\Delta_1 = 10 - 7 = 3$

$\Delta_2 = 10 - 9 = 1$

$\therefore M_0 = 165 + \frac{3}{3+1} \times 5$

$= 168.75$ mm.

Properties of Mode:

① The mode can be found for all levels of data (nominal, ordinal, interval and ratio)

② The mode is not affected by extremely high or low values.

③ A set of data can have more than one mode. If it has two modes, it is said to be bimodal.

Weighted Mean

The weighted Mean is a special case of the arithmet mean. It is often useful when there are several observati of the same value.

The value of each observation is multiplied by the number of times it occurs. The sum of these products is divided by the total number of observations to give the weighted mean.

$$\text{Weighted Mean, } \bar{x} = \frac{W_1X_1 + W_2X_2 + W_3X_3 + \cdots W_nX_n}{W_1 + W_2 + W_3 + \cdots + W_n}$$

Characteristics of ideal measures of central tendency. Which is the best measure of central tendency and why?

Characteristics of ideal measures of central tendency are given below:

1. It should be rigidly defined.
2. It should easy to calculate and easy to understand.
3. It should be based upon all observations.
4. It should be suitable for further algebric treatments.
5. It should less affected by sampling fluctuation.
6. It should be least affected by extreme observation.

Arithmetic mean satisfies most of the criteria and it's considered as the best of central tendency.

Discuss merits and demerits of different Measures of central tendency:

Arithmetic mean

Merits:
1. It is rigidly defined.
2. It is easy to calculate and easy to understand.
3. It is based upon all the observations
4. It is suitable for further algebric treatment.
5. It is less affected by sampling fluctuations.

Demerits:
1. It is affected by extreme values.
2. It cannot be calculated if the extreme class is open.
3. It cannot be used when we are dealing with qualitative characteristics which cannot be measured quantitatvely.

Geometric Mean

Merits:
1. It is rigidly defined
2. It takes all the observations into account.
3. It is used for further observations.
4. It is not affected by sampling fluctuations.

Demerits:
1. It is neither easy to calculate nor easy to be understand.
2. In case of odd number of negative class value it cannot be computed at all.

Harmonic Mean:

Merits: ① It is rigidly defined

② It is based upon all the observations.

③ It is used for further algebric treatment.

④ It is not affected by sampling fluctuations.

Demerits: ① It is not easy to understand and is difficult to calculate.

② It can be calculated if the extreme class is open.

Median:

Merits: ① It is rigidly defined.

② It is easy to understand and easy to calculate.

③ It is not affected by extreme observations.

④ It can be calculated from frequency distribution, with ~~and~~ end class.

Demerits:

① It is not based on all observations.

② It is not suitable for further algebric treatment.

③ It is affected by sampling fluctuations

Mode:

Merits: ① It is easy to understand and easy to calculate.

② It is not affected by extreme value.

③ It can be calculated from frequency distribution with open class

Demerits: ① It is affected by sampling fluctuation.

② It is not based upon all the observations.

③ It is not suitable for further algebric treatment

④ It is not clearly defined in case of multimodal distribution.