Measures of Central Tendency

Geometric Mean: Geometric mean of a set of n non-zero positive observations is the n-th root of their product. The GM of n non-zero positive values $x_1, x_2, \ldots\ldots\ldots\ldots, x_n$ of a variable x is given by,

$$GM = \sqrt[n]{x_1 . x_2 . \ldots\ldots\ldots\ldots . x_n} = (x_1 . x_2 . \ldots\ldots\ldots\ldots . x_n)^{\frac{1}{n}}$$

$$\Rightarrow \log(GM) = \log(x_1 . x_2 . \ldots\ldots\ldots\ldots . x_n)^{\frac{1}{n}} = \frac{1}{n}\sum_{i=1}^{n} \log x_i$$

$$\therefore GM = \text{Anti}\log\left\{\frac{1}{n}\sum_{i=1}^{n} \log x_i\right\}$$

In case of frequency distribution, when $f_1, f_2, \ldots\ldots\ldots\ldots, f_n$ be the frequencies of $x_1, x_2, \ldots\ldots\ldots\ldots, x_n$ respectively then,

$$GM = \text{Anti}\log\left\{\frac{1}{n}\sum_{i=1}^{n} f_i \ \log x_i\right\}$$

Advantages of Geometric Mean:

- It is rigidly defined.
- It is based upon all observations.
- It is not affected much by sampling fluctuations.
- In measuring rate of change it is the most suitable average.

Disadvantages of Geometric Mean:

- It cannot be computed where there is any negative or zero values in the series.
- It is not easy to understand and to calculate for persons having very weak mathematical skills.
- The value of GM may not be found in the series.

Uses of Geometric Mean:

- Geometric Mean is used to find the average of ratios, rate of population growth, rate of interest, average of percentages.

Ex.1) Rate of increase of yield of a new wheat variety compared with a local variety in 10 selected agricultural farms are given below:

Rate of increase of yield (%)	0-5	5-10	10-15	15-20	20-25
Number of Farm	1	2	4	2	1

Solution: For computation of GM, we construct the following table:

Rate of increase of yield (%)	**Frequency f_i**	**Mid Value x_i**	**$\log x_i$**	**$f_i \log x_i$**
0-5	1	2.5	0.39794	0.39794
5-10	2	7.5	0.87506	1.75012
10-15	4	12.5	1.09691	4.38764
15-20	2	17.5	1.24304	2.48068
20-25	1	22.5	1.35218	1.35218
	$\sum f_i = N = 10$			$\sum f_i \log x_i = 10.37396$

$$GM = \text{Anti}\log\left\{\frac{1}{n}\sum_{i=1}^{n} f_i \log x_i\right\} = \text{Anti}\log\left\{\frac{1}{10} \times 10.37396\right\} \approx 10.9$$

$\therefore$ The average rate of change of yield of the new variety of wheat is 10.9%.

Harmonic Mean: The harmonic mean is a type of numerical average. It is calculated by dividing the number of observations by the reciprocal of each number in the series. Thus, the harmonic mean is the reciprocal of the arithmetic mean of the reciprocals.

The HM of n non-zero positive values $x_1, x_2, \dots\dots\dots\dots, x_n$ of a variable x is given by,

$$HM = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + \cdots\dots\dots + \frac{1}{x_n}\right)} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$$

In case of frequency distribution, when $f_1, f_2, \dots\dots\dots\dots, f_n$ be the frequencies of $x_1, x_2, \dots\dots\dots\dots, x_n$ respectively then,

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

Advantages of Harmonic Mean:

- It is rigidly defined.
- It is based upon all observations.
- Sampling fluctuation is less.
- It is not affected much by extreme values.

Disadvantages of Harmonic Mean:

- It cannot be computed where there are any zero values in the series.
- It is not easy to understand and difficult to compute.
- It is very complex for further algebraic operations.

Uses of Harmonic Mean:

- Harmonic Mean is used when observations are made in terms of work done per hour, speeds (kilometers covered per hour), quantity of things purchased per taka etc.

Ex.2) The frequency distribution of profit per share of 10 companies are given below:

Profit per share (Tk)	0-5	5-10	10-15	15-20	20-25
Number of Companies	1	2	4	2	1

Solution: For computation of HM, we construct the following table:

Profit per share (Tk)	**Frequency f_i**	**Mid Value x_i**	$\frac{f_i}{x_i}$
0-5	1	2.5	0.4000
5-10	2	7.5	0.2667
10-15	4	12.5	0.3200
15-20	2	17.5	0.1143
20-25	1	22.5	0.0444
	$\sum f_i = N = 10$		$\sum \frac{f_i}{x_i} = 1.1454$

Therefore,

$$HM = \frac{N}{\sum_{i=1}^{n} \frac{f_i}{x_i}} = \frac{10}{1.1454} = 8.73$$

∴ The average profit per share is 8.73 Tk.

- **Relationship among AM, GM and HM: AM ≥ GM ≥ HM**

(1)Quartiles: Q_i (1,2,3); devide the whole distribution into 4 equal parts.

$$Q_i = L_i + \frac{\frac{i \times N}{4} - p.c.f}{f_i} \times h; \qquad i = 1,2,3$$

(2) Deciles: D_j (1,2, … … … ,9); devide the whole distribution into 10 equal parts.

$$D_j = L_j + \frac{\frac{j \times N}{10} - p.c.f}{f_j} \times h; \qquad j = 1,2, \dots \dots \dots \dots ,9$$

(2)Percentiles: P_k (1,2, … … … ,99); devide the whole distribution into 100 equal parts.

$$P_k = L_k + \frac{\frac{k \times N}{100} - p.c.f}{f_k} \times h; \qquad k = 1,2, \dots \dots \dots \dots ,99$$

H.W:

1) The rice yield (in kg) from a number of small plots are grouped with common class interval of 2 kg in the table below; the x values are mid values of the classes.

Yield (x)	2.8	3.0	3.2	3.4	3.6	3.8	4.0	4.2	4.4	4.6	4.8	5.0
Number of Farm	4	15	20	47	63	88	59	35	15	18	10	5

Compute AM, GM and HM and verify the relationship $\text{AM} \geq \text{GM} \geq \text{HM}$.

2) The frequency distribution below gives the cost of production of sugar in different holdings:

Cost (Tk)	10-14	14-18	18-22	22-26	26-30	30-34	34-38	38-42
Number of Holdings	11	27	42	45	35	30	20	15

Compute:

(i) AM, GM, HM

(ii) Verify the relationship $\text{AM} \geq \text{GM} \geq \text{HM}$

(iii) $\text{Q}_1, \text{D}_4, \text{P}_{80}$