(*). GEOMETRIC MEAN :→

<a> Geometric mean (G.M.) of n individual observations $x_1, x_2, \ldots x_n$ ($x_i \neq 0$) is the n^{th} root of their product.

Thus : $$G = (x_1 x_2 \ldots x_n)^{\frac{1}{n}}$$

Taking logarithms of both sides.

$$\log G = \frac{1}{n}\left[\log x_1 + \log x_2 + \ldots + \log x_n\right]$$

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

$$\Rightarrow \quad G = \text{antilog}\left[\frac{1}{n}\sum_{i=1}^{n} \log x_i\right]$$

<b> If $x_1, x_2, \ldots x_n$ occur $f_1, f_2, \ldots f_n$ times respectively and : $N = \sum_{1}^{n} f_i$; then G.M is given by

$$G = \left(x_1^{f_1} x_2^{f_2} x_3^{f_3} \ldots x_n^{f_n}\right)^{\frac{1}{N}}$$

Taking logarithms of both sides.

$$\log G = \frac{1}{N}\left[f_1 \log x_1 + f_2 \log x_2 + \ldots + f_n \log x_n\right]$$

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

$$\Rightarrow G = \text{antilog}\left[\frac{1}{N}\sum_{i=1}^{n} f_i \log x_i\right]$$

<c> In the case of continuous frequency distribution; x is taken to be the value corresponding to the mid-points of the class-intervals.

(*) Merits and Demerits

Merits:→

<1> It is rigidly defined.
<2> It is based upon all the observations.
<3> It is suitable for further mathematical treatment.
<4> It is not much affected by fluctuations of sampling.

Demerits:→

<1> For a non-mathematical student, it is not easy to understand and calculate.
<2> If any one of the observations is zero, G.M. is zero and if an odd number of observations are negative, it cannot be calculated.

Ex-1: Find the Geometric mean of the series $1, 2, 4, 8, \ldots, 2^n$.

Solⁿ:

$x : 1, 2^1, 2^2, 2^3, \ldots, 2^n$

Number of observations $= n+1$

$$G.M. = \left(1 \cdot 2^1 \cdot 2^2 \cdot 2^3 \cdots 2^n\right)^{\frac{1}{n+1}}$$

$$= \left(2^{1+2+\cdots+n}\right)^{\frac{1}{n+1}}$$

$$= \left[2^{\frac{n(n+1)}{2}}\right]^{\frac{1}{n+1}}$$

$$= 2^{\frac{n}{2}}$$

Ex-2: Compute the Geometric Mean from the following data:

10, 110, 120, 50, 80, 60, 52, 37.

Solⁿ:

Size of items (x)	$\log x$
10	1.0000
110	2.0414
120	2.0792
50	1.6990
80	1.9031
60	1.7782
52	1.7160
37	1.5682
No. of items (n) = 8	13.7851

$$\log G = \frac{1}{n} \Sigma \log x_i$$

$$= \frac{1}{8} (13.7857)$$

$$= 1.723$$

$\therefore \quad G = \text{antilog}(1.723) = 52.84.$

EX-3 The marks obtained by seven students are 5, 10, 15, 20, 25, 30, 35. Find the geometric-mean

Solⁿ:-

size of items (x)	log x
5	0.69897
10	1
15	1.17609
20	1.30103
25	1.39794
30	1.47712
35	1.54406
No. of items (n) = 7	$\Sigma \log x_i = 8.5952$

$$\log G = \frac{1}{n} [\Sigma \log x_i]$$

$$= \frac{1}{7} [8.59521]$$

$$= 1.227887 = 1.2279$$

$\therefore \quad G = \text{antilog}(1.2279) = 16.90$

3.7 HARMONIC MEAN

The *harmonic mean* of a number of observations, none of which is zero, is the reciprocal of the arithmetic mean of the reciprocals of the given values.

The harmonic mean of n observations $x_1, x_2, \ldots, x_n$ is given by

$$\text{HM} = \frac{1}{\frac{1}{n}\sum\left(\frac{1}{x}\right)}$$

$$= \frac{1}{\frac{1}{n}\left(\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}\right)}$$

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

For example, the harmonic mean of 2, 4 and 5 is

$$\text{HM} = \frac{3}{\frac{1}{2}+\frac{1}{4}+\frac{1}{5}} = 3.16$$

In case of a frequency distribution consisting of n observations $x_1, x_2, \ldots, x_n$ with respective frequencies $f_1, f_2, \ldots, f_n$, the harmonic mean is given by

$$\text{HM} = \frac{f_1+f_2+\cdots+f_n}{\frac{f_1}{x_1}+\frac{f_2}{x_2}+\cdots+\frac{f_n}{x_n}}$$

$$\text{H.M.} = \frac{\sum f}{\sum\left(\frac{f}{x}\right)}$$

If $x_1, x_2, \ldots, x_n$ are n observations with weights $w_1, w_2, \ldots, w_n$ respectively, their weighted harmonic mean is given by

$$\text{HM} = \frac{\sum w}{\sum\left(\frac{w}{x}\right)}$$

Example 1

16/7 *Calculate the harmonic mean of the following data:*

x	20	21	22	23	24	25
f	4	2	7	1	3	1

Solution

x	f	$\frac{f}{x}$
20	4	0.2
21	2	0.095
22	7	0.318
23	1	0.043
24	3	0.125
25	1	0.04
	$\Sigma f = 18$	$\Sigma\left(\frac{f}{x}\right) = 0.821$

$$\text{HM} = \frac{\Sigma f}{\Sigma\left(\frac{f}{x}\right)} = \frac{18}{0.821} = 21.924$$

Example 2

Find the harmonic mean of the following distribution:

Class interval	0–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80
Frequency	5	8	11	21	35	30	22	18

Solution

Class Interval	Frequency f	Midvalue x	$\frac{f}{x}$
0–10	5	5	1
10–20	8	15	0.533
20–30	11	25	0.44
30–40	21	35	0.6
40–50	35	45	0.778
50–60	30	55	0.545
60–70	22	65	0.338
70–80	18	75	0.24
	$\Sigma f = 150$		$\Sigma\left(\frac{f}{x}\right) = 4.474$

$$\text{HM} = \frac{\sum f}{\sum\left(\frac{f}{x}\right)} = \frac{150}{4.474} = 33.527$$

Relation between Arithmetic Mean, Geometric Mean, and Harmonic Mean

The arithmetic mean (AM), geometric mean (GM), and harmonic mean (HM) for a given set of observations of a series are related as

$$\text{AM} \geq \text{GM} \geq \text{HM}$$

For two observations x_1 and x_2 of a series,

$$\text{AM} = \frac{x_1 + x_2}{2}$$

$$\text{GM} = \sqrt{x_1\, x_2}$$

$$\text{HM} = \frac{2}{\frac{1}{x_1} + \frac{1}{x_2}} = \frac{2x_1\, x_2}{x_1 + x_2}$$

$$\text{AM} \cdot \text{HM} = \left(\frac{x_1 + x_2}{2}\right)\left(\frac{2x_1\, x_2}{x_1 + x_2}\right) = x_1\, x_2 = (\text{GM})^2$$

$$\therefore \quad \text{GM} = \sqrt{\text{AM} \cdot \text{HM}}$$

Example 1

If the AM of two observations is 15 and their GM is 9, find their HM and the two observations.

Solution

$$\text{GM} = \sqrt{\text{AM} \cdot \text{HM}}$$

$$9 = \sqrt{15 \times \text{HM}}$$

$$\therefore \quad \text{HM} = 5.4$$

Let the two observations be x_1 and x_2.

$$\text{AM} = \frac{x_1 + x_2}{2} = 15$$

$$x_1 + x_2 = 30 \qquad \text{...(1)}$$

$$\text{GM} = \sqrt{x_1\, x_2} = 9$$

$$x_1\, x_2 = 81 \qquad \text{...(2)}$$

3.8 STANDARD DEVIATION

Standard deviation is the positive square root of the arithmetic mean of the squares of the deviations of the given values from their arithmetic mean. It is denoted by the Greek letter σ. Let X be a random variable which takes on values, viz., $x_1, x_2, \ldots, x_n$. The standard deviation of these n observations is given by

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$$

where $\bar{x} = \dfrac{\sum x}{n}$ is the arithmetic mean of these observations.

This equation can be modified further.

$$\sigma = \sqrt{\frac{\sum (x^2 - 2x\bar{x} + \bar{x})^2}{n}}$$

$$= \sqrt{\frac{\sum x^2 - 2\bar{x}\sum x + \bar{x}^2 \sum 1}{n}}$$

$$= \sqrt{\frac{\sum x^2}{n} - 2\frac{\sum x}{n}\frac{\sum x}{n} + \left(\frac{\sum x}{n}\right)^2 \cdot \frac{n}{n}} \qquad \left[\because \sum 1 = n\right]$$

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\text{Mean of squares} - \text{Square of mean}}$$

In case of a frequency distribution consisting of n observations $x_1, x_2, \ldots, x_n$ with respective frequencies $f_1, f_2, \ldots, f_n$, the standard deviation is given by

$$\sigma = \sqrt{\frac{\sum f(x-\bar{x})^2}{N}}$$

This equation can also be modified.

$$\sigma = \sqrt{\frac{\sum f\,(x^2 - 2x\bar{x} + \bar{x}^2)}{N}}$$

$$= \sqrt{\frac{\sum fx^2}{N} - \frac{2\bar{x}\sum fx}{N} + \bar{x}^2 \frac{\sum f}{N}}$$

$$= \sqrt{\frac{\sum fx^2}{N} - 2\frac{\sum fx}{N}\frac{\sum fx}{N} + \left(\frac{\sum fx}{N}\right)^2} \qquad \left[\because \sum f = N \text{ and } \bar{x} = \frac{\sum fx}{N}\right]$$

$$\sigma = \sqrt{\frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N}\right)^2}$$

3.8.1 Variance

The *variance* is the square of the standard deviation and is denoted by σ^2. The method for calculating variance is same as that given for the standard deviation.

Example 1

Calculate the standard deviation of the weights of ten persons.

Weight (in kg)	45	49	55	50	41	44	60	58	53	55

Solution

$$n = 10$$

$$\sum x = 45 + 49 + 55 + 50 + 41 + 44 + 60 + 58 + 53 + 55 = 510$$

$$\sum x^2 = 45^2 + 49^2 + 55^2 + 50^2 + 41^2 + 44^2 + 60^2 + 58^2 + 53^2 + 55^2 = 26366$$

$$\sigma = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{26366}{10} - \left(\frac{510}{10}\right)^2}$$

$$= 5.967$$

Aliter:

$$\bar{x} = \frac{\sum x}{n} = \frac{510}{10} = 51$$

x	$x - \bar{x}$	$(x - \bar{x})^2$
45	–6	36
49	–2	4
55	4	16
50	–1	1
41	–10	100
44	–7	49
60	9	81
58	7	49
53	2	4
55	4	16
		$\Sigma(x - \bar{x})^2 = 356$

$$\sigma = \sqrt{\frac{\sum (x-\bar{x})^2}{n}}$$

$$= \sqrt{\frac{356}{10}}$$

$$= 5.967$$

Example 2

Calculate the standard deviation of the following data:

x	10	11	12	13	14	15	16	17	18
f	2	7	10	12	15	11	10	6	3

Solution

x	f	fx	x^2	fx^2
10	2	20	100	200
11	7	77	121	847
12	10	120	144	1440
13	12	156	169	2028
14	15	210	196	2940
15	11	165	225	2475
16	10	160	256	2560
17	6	102	289	1734
18	3	54	324	972
	$\sum f = 76$	$\sum fx = 1064$		$\sum fx^2 = 15196$

$$N = \sum f = 76$$

$$\sigma = \sqrt{\frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N}\right)^2}$$

$$= \sqrt{\frac{15196}{76} - \left(\frac{1064}{76}\right)^2}$$

$$= 1.987$$

Aliter:

$$N = \sum f = 76$$

$$\bar{x} = \frac{\sum fx}{N} = \frac{1064}{76} = 14$$

Example 4

The number of matches played and goals scored by two teams A and B in World Cup Football 2002 were as follows:

Matches played by Team A	27	9	8	5	4
Matches played by Team B	17	9	6	5	3
No. of goals scored in a match	0	1	2	3	4

Find which team may be considered more consistent.

Solution

For Team *A*,

$$N_A = 27+9+8+5+4 = 53$$

$$\sum fx_A = (27\times 0)+(9\times 1)+(8\times 2)+(5\times 3)+(4\times 4) = 56$$

$$\sum fx_A^2 = (27\times 0^2)+(9\times 1^2)+(8\times 2^2)+(5\times 3^2)+(4\times 4^2) = 150$$

$$\sigma_A = \sqrt{\frac{\sum fx_A^2}{N_A} - \left(\frac{\sum fx_A}{N_A}\right)^2}$$

$$= \sqrt{\frac{150}{53} - \left(\frac{56}{53}\right)^2}$$

$$= 1.31$$

$$\bar{x}_A = \frac{\sum fx_A}{N_A} = \frac{56}{53} = 1.06$$

$$\text{CV}_A = \frac{\sigma_A}{\bar{x}_A}\times 100$$

$$= \frac{1.31}{1.06}\times 100$$

$$= 123.58\%$$

For Team *B*,

$$N_B = 17+9+6+5+3 = 40$$

$$\sum fx_B = (17\times 0)+(9\times 1)+(6\times 2)+(5\times 3)+(3\times 4) = 48$$

$$\sum fx_B^2 = (17\times 0^2)+(9\times 1^2)+(6\times 2^2)+(5\times 3^2)+(3\times 4^2) = 126$$

$$\sigma_B = \sqrt{\frac{\sum fx_B^2}{N_B} - \left(\frac{\sum fx_B}{N_B}\right)^2}$$

$$= \sqrt{\frac{126}{40} - \left(\frac{48}{40}\right)^2}$$

$$= 1.31$$

$$\bar{x}_B = \frac{\sum fx_B}{N_B} = \frac{48}{40} = 1.2$$

$$\text{CV}_B = \frac{\sigma_B}{\bar{x}_B}\times 100$$

$$= \frac{1.31}{1.2}\times 100$$

$$= 109.17\%$$

since $CV_B < CV_A$,
$\therefore$ Team B is more consistent in performance