If a variable x assumes n values $x_1, x_2, \ldots, x_n$, then its geometric mean, denoted by G, is

$$G = (x_1 x_2 \cdots x_n)^{1/n} = \left(\prod_{i=1}^{n} x_i\right)^{1/n} \cdots \quad (4k)$$

For a frequency distribution,

$$G = \left(x_1^{f_1} x_2^{f_2} \cdots x_r^{f_r}\right)^{1/n} = \left(\prod_{i=1}^{r} x_i^{f_i}\right)^{1/n}, \quad n = \sum_{i=1}^{r} f_i$$

Advantages of G.M :

(i) It is rigidly defined.

(ii) It is directly based on all observations

(iii) It possesses some properties which enables the measure to be readily applicable in theoretical work.

(iv) Generally, the presence of a few ~~extremly~~ extremely small or large values has no considerable ~~eff~~ effect on G.M.

Disadvantage

(i) It is abstract in nature

(ii) It is difficult to compute

(iii) If a single ~~posi~~ value of a variable is zero, then the geometric mean becomes zero. Also, it may be imaginary if some values are negative. Generally, we use geometric mean for positive values.

3.4.2 Harmonic mean (H.M.)

The harmonic mean of a set of non-zero values of a variable is the reciprocal of the arithmetic mean of the reciprocals of the values.

Thus, H.M. of n non-zero values $x_1, x_2, \ldots, x_n$

of a variable x is $H = \dfrac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\cdots+\frac{1}{x_n}}$ ... 4(l)

For a frequency distribution, $H = \dfrac{n}{\sum_{i=1}^{r} \frac{f_i}{x_i}}$ where $n = \sum_{i=1}^{r} f_i$

Advantages of H.M

(i) It is rigidly defined.

(ii) It is directly based on all observations.

Disadvantages

(i) It is undefined if a single value is zero

(ii) It is abstract in nature

(iii) It involves a lot of computational labour

(iv) It is not amenable to algebraic treatment.

Example 4.12 If A.M and G.M of two positive real numbers are 25 and 15 respectively, then find their H.M

Solution: Let x_1, x_2 be two positive real numbers

So, we have $\dfrac{x_1+x_2}{2} = 25$ or, $x_1+x_2 = 50$

Also $\sqrt{x_1x_2} = 15$ or, $x_1x_2 = 225$

$\therefore$ Required $H.M = \dfrac{2}{\frac{1}{x_1}+\frac{1}{x_2}} = \dfrac{2x_1x_2}{x_1+x_2} = \dfrac{\cancel{2}\times \cancel{225}^{9}}{\cancel{50}_{25}} = 9$

3.4.3 Relationship between A.M, G.M. and H.M.

Let us consider n positive values, $x_1, x_2, \cdots x_n$ of a variable. If A, G, and H be the A.M, G.M and H.M then $A \geq G \geq H$ (Prove it)

5. Measures of Dispersion

The values of a variable are generally not equal. In some case, the values are close to one another; again, in some cases, they are markedly different from one another.. In order to get a proper idea about the overall nature of a given set of values, it is necessary to know, besides average, the extent to which the given values differ among themselves or equivalently, how they are scattered about the average. This feature of a frequency distribution which represents the variability of the given values or reflects how scattered the values are, is called its dispersion. A device which is used to measure this characteristic, namely scatter, is referred to as a measure of dispersion.

5.1 Different measures of Dispersion

(a) Range

The range of a variable is the simplest measure of its dispersion and it is defined as the difference between the greatest and the least of its given set of values. It should be noted that if the data are given in a grouped frequency distribution, the range can be considered as the difference between the largest upper boundary and smallest lower boundary.

Example 5.1. Suppose a variable takes the values 3, 5, -1, 8, 4, calculate the range.

Solution: Here the greatest value = 8 and the least value = -1

Thus, the range of the variable = 8 - (-1) = 9.

If $y = a + bx$ be the relation between the two variables x and y, then

$$\text{Range of } y = |b| \times \text{Range of } x \quad (\text{prove it})$$

Example 5.2

If two variables x and y are related as $3y + 4x = 9$ and range of x is 3, then find the range of y.

Solution: Here we have $3y + 4x = 9$ or $y = 3 - \frac{4}{3}x$

$\therefore$ Range of $y = \left|-\frac{4}{3}\right| \times$ Range of $x = \frac{4}{3} \times 3 = 4$

(3) Mean deviation

The mean deviation is actually the mean of absolute deviation of the given values of the variable from some ~~average~~ average. Suppose $x_1, x_2, \ldots, x_n$ are the given values of ~~the~~ a variable x and c is a chosen average. Then first we consider the deviation $x_i - c$ $(i = 1, 2, \ldots, n)$ of the values from c. The greater the deviation, more is the dispersion. To get a suitable measure it is necessary to combine these deviations. The simple arithmetic mean of the deviation will not serve the purpose, since the sum of the deviation may be small