Chapter-4

Standard deviation & other Measures of Dispersion :-

Dispersion, or variation :- The degree to which numerical data tend to spread about an average value is called the dispersion, or variation, of the data. Examples:- Range, Mean deviation, semi-interquartile range, and standard deviation.

Range :- The difference between the largest and smallest numbers of a data set.

see → $\boxed{\text{Example-1}}$, 4.1 & 4.2 → Page → 95 & 102 & 103.

Mean deviation :-

$$\text{Mean deviation (MD)} = \frac{\sum_{j=1}^{N} |X_j - \bar{X}|}{N} \rightarrow \text{Ungrouped data}$$

Examples: 2 & 4.3 → Page → 96 & 103.

If $X_1, X_2, \dots, X_K$ occur with frequencies $f_1, f_2, \dots, f_K$, respectively,

$$MD = \frac{\sum_{j=1}^{K} f_j |X_j - \bar{X}|}{N = \sum_{j=1}^{K} f_j}$$

Example: 4.4 & 4.5 → Page → 102 & 103

In 4.5 → 3 = class size.

Inter quartile Range = $Q_3 - Q_1$

semi- inter quartile Range = $\frac{1}{2}(Q_3 - Q_1)$

Examples:- 4.6 & 4.7 → page → 104 & 105.

Standard deviation (S.D.):-

$$S = \sqrt{\frac{\sum_{j=1}^{N}(X_j - \bar{X})^2}{N}}$$ → Ungrouped data. — ①

S is the root mean square (RMS) of the deviation from the mean. Example → 4.9 → page → 105

If $X_1, X_2, \ldots, X_k$ occur with frequencies $f_1, f_2, \ldots, f_k$, respectively,

Example → 4.11 → page → 106

$$S = \sqrt{\frac{\sum_{j=1}^{k} f_j (X_j - \bar{X})^2}{N}}$$ — ②

For a better estimation of S of a population, replacing N by (N-1) in the denominators of ① & ②.

For large values of N (certainly N>30), there is ~~part~~ practically no difference between the two definitions.

<u>variance</u> :- is the square of the standard deviation. that is, s^2.

s^2 and σ^2 represents the sample variance and population variance, respectively.

<u>Short Methods for computing the S.D :</u>

$$S = \sqrt{\frac{\sum (x_j - \bar{x})^2}{N}} = \sqrt{\frac{\sum (x_j^2 - 2x_j \cdot \bar{x} + \bar{x}^2)}{N}}$$

$$= \sqrt{\left(\frac{\sum x_j^2}{N} - 2\bar{x} \cdot \frac{\sum x_j}{N} + \frac{N\bar{x}^2}{N}\right)}$$

$$= \sqrt{\left(\frac{\sum x_j^2}{N} - 2\bar{x}^2 + \bar{x}^2\right)}$$

$$= \sqrt{\frac{\sum x_j^2}{N} - \bar{x}^2}$$

$$= \sqrt{\frac{\sum x_j^2}{N} - \left(\frac{\sum x_j}{N}\right)^2} \quad \text{——— (3)}$$

→ Ungrouped data

$$S = \sqrt{\frac{\sum_{j=1}^{k} f_j x_j^2}{N} - \left(\frac{\sum_{j=1}^{k} f_j x_j}{N}\right)^2} \quad \text{——— (4)}$$

Example → 4.14 (page-108)

If $d_j = x_j - A$ are the deviations of x_j from some arbitrary constant A, then (3) & (4) become, respectively,

$$s = \sqrt{\frac{\sum d^2}{N} - \left(\frac{\sum d}{N}\right)^2} \quad \text{——— (5)}$$

$$s = \sqrt{\frac{\sum f d^2}{N} - \left(\frac{\sum f d}{N}\right)^2} \quad \text{——— (6)}$$

Example - 4.17(a) page → 109

c = Equal size of class intervals,

Then, $d_j = c u_j$ where, $u_j = \dfrac{x_j - A}{c}$

$\Rightarrow$ (6) becomes,

$$s = c\sqrt{\frac{\Sigma f u^2}{N} - \left(\frac{\Sigma f u}{N}\right)^2}$$

$\boxed{\text{Example- 4.17 (b) Page-110}}$

$\boxed{\text{Properties of S.D}}$ $\rightarrow$ see page- 98.

Empirical relations between Measures of dispersion :-

Mean deviation $= \frac{4}{5}$ (standard deviation)

semi- interquartile range $= \frac{2}{3}$ (standard deviation).

See $\rightarrow$ $\boxed{\text{Example - 4.18}}$