Unit-2

* Measures of Dispersion

- **Range:-**

The range of a set of data is defind as the difference, between the largest and the smallest value in the set.

Symbolically, Range = L-S.

where, L = largest value &
S = Smallest value.

- **Coefficient of Range:**

It is defined by,

$$\text{Co-efficient of range} = \frac{\text{Range}}{\text{Sum of the largest \& smallest values}}$$

$$= \frac{L-S}{L+S}.$$

* Quartile Deviation:-

⇒ How to find quartiles Q_1 & Q_3, for...

i.) Individual Observations,

$$Q_1 = \text{size of } \left(\frac{n+1}{4}\right)^{th} \text{ item.}$$

$$Q_3 = \text{size of } 3\left(\frac{n+1}{4}\right)^{th} \text{ item.}$$

ii.) Descrete Series,

$$Q_1 = \text{size of } \left(\frac{N+1}{4}\right)^{th} \text{ item.}$$

$$Q_3 = \text{size of } \left(\frac{3(N+1)}{4}\right)^{th} \text{ item}$$

iii.) Continuous Series;

$$Q_1 = l + \frac{N/4 - c}{f} \times h.$$

$$Q_3 = l + \frac{3N/4 - c}{f} \times h.$$

$\Rightarrow$ Inter-quartile range = $Q_3 - Q_1$.

$\Rightarrow$ Quartile Deviation (Q.D.) $= \frac{Q_3 - Q_1}{2}$.

$\Rightarrow$ Co-efficient of Quartile Deviation

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1}.$$

* Mean Deviation:-

- **Computation of Mean Deviation:-**

(i) Individual observation:-

For a given set of n observations $X_1, X_2, \ldots, X_n$, the mean deviation (M.D.) about average, say A, is given by,

Mean Deviation (about an average A)

$$= \frac{\Sigma|X-A|}{n} = \frac{\Sigma|D|}{n}$$

where $|D| = |X-A|$.

(ii) Discrete Series:-

In case of discrete series where the variable X takes the values $X_1, X_2, \ldots X_n$ with respective frequencies $f_1, f_2, \ldots, f_n$, the mean deviation about an average A is given by,

Mean Deviation (about an average A)

$$= \frac{\Sigma f|X-A|}{N}$$

$$= \frac{\Sigma f|D|}{N}.$$

where $N = \Sigma f$ is total of frequency &

$D = X - A$.

iii) Continuous Series:

→ The computation of the mean deviation in the case of continuous series is exactly the same as discussed above for discrete series. The only difference is that there we have to obtain the mid values of various classes and take absolute devation of these values from the average A. Thus, if $X_1, X_2, \ldots, X_n$ are the ~~class~~ mid values of a set of grouped data with corresponding class frequencies $f_1, f_2, \ldots, f_n$, then mean deviation about an average A is given by,

Mean Deviation (about an average A)

$$= \frac{\Sigma f|X-A|}{N} = \frac{\Sigma f|D|}{N}$$

where $N = \Sigma f$ is total frequency.

→ Co-efficient of Mean Deviation:-

Cofficient of M.D. = $\frac{\text{Mean Deviation}}{\text{Average about which it is calculated}}$

∴ co-efficient of M.D. about mean = $\frac{M.D}{Mean}$

& ∴ co-efficient of M.D. about median = $\frac{M.D.}{Median}$

Imp NOTE:- The mean deviations gives the best results when deviations are taken from median. Thus when nothing is specified in the question, that about which the deviation is calculated, we shall take deviation from median.

Standard Deviation:-

• Calculation of Standard Deviation:

(i.) Individual observations:

- The standard deviation of a set of n observations $X_1, X_2, \ldots, X_n$ is given as,

$$(S.D.) = \sigma = \sqrt{\frac{(X_1-\bar{X})^2+(X_1-\bar{X})^2+\ldots+(X_n-\bar{X})^2}{n}}$$

$$= \sqrt{\frac{\Sigma(X-\bar{X})^2}{n}}$$

where $\bar{X} = \frac{\Sigma X}{n}$ is the arithmetic mean of the given observations.

(ii.) Discrete Series:

- In the case of discrete series, we have n observations $X_1, X_2, \ldots, X_n$ and we are also given the corresponding frequency as $f_1, f_2, \ldots, f_n$ then standard deviation is given as,

$$(S.D.)\ \sigma = \sqrt{\frac{f_1(X_1-\bar{X})^2 + f_2(X_2-\bar{X})^2 + \ldots + f_n(X_n-\bar{X})^2}{f_1+f_2+\ldots+f_n}}$$

$$= \sqrt{\frac{\Sigma f(X-\bar{X})^2}{N}}$$

where $N = \Sigma f$ is the total frequency. &

$\bar{X} = \frac{\Sigma f X}{N}$ is the arithmetic mean.

(iii.) Continuous Series:-

In the case of ~~conti~~ finding S.D. (σ) of continuous series we will have $X_1, X_2, \ldots, X_n$ as mid-values along with the corresponding frequency, $f_1, f_2, \ldots, f_n$, then the S.D. (σ) is given as,

standard deviation (σ).

$$= \sqrt{\frac{f_1(X_1-\bar{X})^2 + f_2(X_2-\bar{X})^2 + \ldots + f_n(X_n-\bar{X})^2}{f_1+f_2+\ldots+f_n}}$$

$$= \sqrt{\frac{\Sigma f(X-\bar{X})^2}{N}}$$

where $N = \Sigma f$ is the total frequency &

$\bar{x} = \frac{\Sigma fx}{N}$ is the arithmetic mean.

Variance:-

The variance of given set of observations is defined as the square of its standard deviation and it is denoted by σ^2.

Thus,

→ In case of individual observation,

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

→ And In case of Descrete series & continuous series,

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

* Different Methods of Calculating Standard Deviation:-

(i.) For Individual Observation:

⟹ By def^n of S.D it is given by,

$$\sigma = \sqrt{\frac{\Sigma\cancel{f}(X-\bar{x})^2}{n}} \quad \text{where } \bar{x} = \frac{\Sigma X}{n}.$$

- formula-1 (Short-cut method).

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

- formula-2. (Assumed Mean method).

$$\sigma = \sqrt{\frac{\Sigma d^2}{n} - \left(\frac{\Sigma d}{n}\right)^2}, \text{ where } d = X - A.$$

- formula-3. (Step-deviation Method).

$$\sigma = \sqrt{\frac{\Sigma u^2}{n} - \left(\frac{\Sigma u}{n}\right)^2} \times C.$$

where $u = \frac{X-A}{C}$ and C = common factor.

(ii.) for Discrete & Continuous Series:

⟹ By defⁿ of S.D, it is given as,

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

where $N = \Sigma f$ & $\bar{x} = \frac{\Sigma fx}{N}$.

- formula-1 (Short-cut Method)

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

- formula-2 (Assumed Mean Method)

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2}, \text{ where } d = x - A.$$

- formula-3 (Step-deviation Method)

$$\sigma = \sqrt{\frac{\Sigma fu^2}{N} - \left(\frac{\Sigma fu}{N}\right)^2} \times h.$$

where $u = \frac{x-A}{h}$.

* Combined Standard Deviation:

→ If $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ be the means, $\sigma_1, \sigma_2, \ldots, \sigma_k$ be the standard deviations and $n_1, n_2, \ldots, n_k$ be the number of observations in each set, then the standard deviations of combined data with $n_1+n_2+\ldots+n_k$ observation is given by,

$$\sigma = \sqrt{\frac{n_1(\sigma_1^2+d_1)^2+n_2(\sigma_2^2+d_2^2)+\ldots+n_k(\sigma_k^2+d_k^2)}{n_1+n_2+\ldots+n_k}}$$

where $d_1 = \bar{x}_1-\bar{x}$, $d_2 = \bar{x}_2-\bar{x}$, ..., $d_k = \bar{x}_k-\bar{x}$

and $\bar{x} = \dfrac{n_1\bar{x}_1+n_2\bar{x}_2+\ldots+n_k\bar{x}_k}{n_1+n_2+\ldots+n_k}$

is combined mean.

* Coefficient of S.D. :-

It is given as,

co-efficient of Standard Deviation $= \dfrac{\text{S.D.}}{\text{Mean}}$

$= \dfrac{\sigma}{\bar{x}}$.

* Co-efficien of Variation :-

It is given as,

$$\text{Co-efficient of variation} = \frac{S.D.}{Mean} \times 100$$

$$= \frac{\sigma}{\bar{x}} \times 100.$$

NOTE:-

→ Co-efficient of Variation can be used to compare the variability of two or more sets of data.

– A distribution for which the coefficient of variation is smaller is considered as more consistent, more stable data.