Measures of Dispersion

Points to remember

Dispersion is a measure of the variation of the items from central value.
The measures of dispersion is important to compare uniformity, consistency and reliability amongst variables / series.

Absolute measures of dispersion are expressed in terms of original unit of series.

Relative measures are expressed in ratios or percentage of average, also known as coefficients of dispersion.

Measures of Dispersion:
(i) Range
(ii) Inter quartile range
(iii) Quartile deviation or Semi-Inter-quartile range
(iv) Mean deviation
(v) Standard Deviation
(vi) Lorenz curve (Not in syllabus)

Range: Range is defined as the difference between two extreme observations i.e. the largest and the smallest value.

Symbolically, R = L-S
Where R = Range
L = Largest Value
S = Smallest value

Coefficient of range = $\frac{L-S}{L+S}$

Inter Quartile Range:
Inter quartile range is the difference between upper quartile and lower quartile.

Inter-quartile range = $Q_3 - Q_1$

Where Q_3 = Third quartile or upper quartile.

Q_1 = First quartile or lower quartile

Quartile Deviation:

Quartile deviation is known as half of difference of third quartile (Q_3) and first quartile (Q_1). It is also known as semi inter quartile range.

$$\boxed{Q.D. = \frac{Q_3 - Q_1}{2}}$$

Where Q.D. = Quartile deviation

Q_3 = Third quartile or upper quartile.

Q_1 = First quartile of lower quartile.

Coefficient of quartile deviation=

Mean Deviation:
Mean deviation/average deviation is the arithmetic mean of the deviations of various items from their average (mean, median or mode) generally from the median.

Calculation of mean deviation

Individual Series $\boxed{M.D. = \frac{\Sigma|D|}{N}}$

Discrete Series $\boxed{M.D. = \frac{\Sigma f|D|}{N}}$

Continuous Series $\boxed{\frac{\Sigma f|D|}{N}}$

Where,
MD = Mean deviation
| D | = Deviations from mean or median ignoring + Signs
N = Number of item (Individual Series)
N = Total number of Frequencies (Discrete and continuous series)
F = Number of frequencies.
Coefficient of mean deviation

Merit of Mean deviation:

1. As in case of X, every term is taken in account hence, it is certainly a better measure than other measures of dispersion i.e. Range, Percentile Range or Quartile Range.
2. Mean deviation is extensively used in other fields such as Economics, Business, Commerce or any other field of such type.
3. It has least sampling fluctuations as compared to Range, Percentile Range and Quartile Deviation.
4. When comparison is needed this is perhaps the best measure between two or more series.
5. This calculation has its base upon measurement than an estimate.
6. Mean Deviation is rigidly defined; one of the main focus point of any measure used for statistical Analysis.
7. It we calculate it from median it is less affected by extreme terms.
8. As it is based on the deviations about an average, it gives us better measure for comparison.

Demerits of Mean Deviation:

1. If average is in fractions, it is difficult to compile M.D.

2. Main property is absent, It is not capable of further Algebraic Treatment.
3. Not so easy to calculate to calculate X, M or Z first and then to go for other measures.
4. If it is calculated from Z it is not much reliable as Mode (Z) is not the true representative of the series.
5. M.D. and its co-efficient taken from X, M and Z often differ.
6. As +,- signs are ignored which is not possible mathematically. Algebraically we have to proceed for Standard Deviation; or another measure of dispersion.
7. As for mean, open and series cannot be taken for the true result.
8. If Range increases in case the sample increases, Average deviation also increases but not in the same ratio.
9. For Sociological studies, it is almost not used.

Standard Deviation:

Standard deviation is the best and widely used measure of dispersion. Standard deviation is the square root of the arithmetic mean of the squares of deviation of its items from their arithmetic mean. Calculation of standard deviation in individual series.
Actual mean method.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

Where σ = Standard Deviation

Σx^2 = Sum total of square of Deviation taken from Mean
N = Number of items

Shortcut Method or assumed mean method:

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

Where d^2 = Square of deviation taken from assumed mean.

Calculation of standard deviation in discrete series:
Actual mean method or direct method

$$\sigma = \sqrt{\frac{\Sigma x^2}{N}}$$

Where σ = S.D.

Σx^2 = Sum total of the squared deviations multiplied by frequency
N = Number of pair of observations.

Shortcut method or assumed method:

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

σ = S.D.

Σfd^2 = Sum total of the squared deviations Multiplied by frequency

Σfd^2 = Sum total of deviations multiplied by frequency.

N = Number of pair of observations.

Step deviation method:

$$\sigma = \sqrt{\frac{\Sigma fd^2}{\Sigma f} - \left(\frac{\Sigma fd}{\Sigma f}\right)^2} \times C$$

σ = Standard Deviation

Σfd^2 = Sum total of the squared step deviations multiplied by frequency.

Σfd^2 = Sum total of step deviations multiplied by frequency

C = Common factor

N = Number of pair of observation

Individual Series:

1. Actual Mean Method

$$\sigma = \sqrt{\frac{\Sigma x^2}{N}}$$

$$x = X - \overline{X}$$

2. Assumed Mean Method

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

Discrete/Continuous Series:

1. Actual Mean Method

$$\sigma = \sqrt{\frac{\Sigma fx^2}{\Sigma f}}$$

$$x = X - \overline{X}$$

2. Assumed Mean Method

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

3. Step Deviation Method

$$\sigma = \sqrt{\frac{\Sigma fd^{12}}{\Sigma f} - \left(\frac{\Sigma fd^{1}}{\Sigma f}\right)^{2}} \text{ X C}$$

Merits of standard deviation:

1. Based on all values
2. Rigidly defined
3. Less effect of fluctuations
4. Capable of algebraic treatment

Demerits of standard deviation:

1. Difficult to compute
2. More stress on extreme items
3. Dependent on unit of measurement.

Coefficient of variation:

When two or more groups of similar data are to be compared with respect to stability (or uniformly or consistency or homogeneity). Coefficient of variation is the most appropriate measures. It is the ratio of the standard deviation to the mean.

$$\text{CV} = \frac{\sigma}{\overline{\text{X}}} \times 100$$

Where C.V. = Coefficient of variation

σ = Standard deviation

X = Arithmetic mean

Next content is Practice Set