Economics

(Statistics)

Chapter 6: Measures of dispersion

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:

- Range
- Inter quartile range
- Quartile deviation or Semi-Inter-quartile range
- Mean deviation
- Standard Deviation
- Lorenz curve

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.

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

Where Q.D. = Quartile deviation

Q_3 = Third quartile or upper quartile.

Q_1 = First quartile of lower quartile.

Coefficient of quartile deviation= $Q_3 - Q_1$ $Q_3 - Q_1$ $Q_3 - Q_1$ $Q_3 - Q_1$

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

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

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

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:

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

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

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

Merits of standard donation:

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

Demerits of standard donation:

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.

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

Where C.V. = Coefficient of variation

σ = Standard deviation

X = Arithmetic mean

LORENZ CURVE:

The Lorenz curve devised by Dr. Max O. Lorenz is a graphic method of studying dispersion.

The Lorenz curve always lies-below the line of equal distribution, unless the distribution is uniform.

The Area between the line of equal distribution and the plotted curve gives the extent of inequality in the items. The larger the area, more is the inequality.

Application Lorenz Curve:

(i) Distribution of income

(ii) Distribution of wealth

(iii) Distribution of wages

(iv) Distribution of production

(v) Distribution of population

Construction of Lorenz Curve:

1. Series is converted into a cumulative frequency series. The cumulative sum of items is assumed to be 100 and different items are converted into percentage of the cumulative sum.
2. Cumulative sum of frequency is assumed to be 100 and different Frequencies are converted into percentage of sum of frequency.
3. Cumulative frequencies are plotted on x-axis and cumulative items are plotted on y-axis of graph.
4. On both axis values are plotted from 0-100.
5. A diagonal line(0 on X-axis and 100 on Y-axis are joined by a line). It is called line of equal distribution.
6. Actual data are plotted by joining different points. This is Lorenz Curve.
7. Lesser distance between the line of equal distribution and line of actual distribution shows lesser dispersion and so on.

Class : 11th Economics (Statistics)
Chapter-6 : Measures Of Dispersion

Measures Of Dispersion

Lorenz Curve

Construction
- Items of both the groups are made cumulative and then cumulative frequencies are known,
- Considering the last cumulative frequency as 100 find the cumulative percentage of all frequencies,
- Now ascertain the points of the percentage of both the groups or both the axes of the curve,
- Draw a curve by joining both the ends with free hand and compare it with the line of equal distribution,
- In continuous series, we proceed by taking the midpoint of the group.

Meaning
The Lorenz curve is a kind of cumulative frequency curve which is based on percentage and it is drawn by putting values on graph.

Relative Dispersion

- Coefficient of range $= \frac{L-S}{L-S}$
- Coefficient of quartile deviation $= \frac{Q_3 - Q_1}{Q_3 + Q_1}$
- Coefficient of mean deviation $= \frac{\text{mean deviation}}{mean}$
- Coefficient of standard deviation $= \frac{S.D.}{mean}$
- Coefficient of variation $= \frac{S.D.}{mean} \times 100$

Absolute Dispersion

Rang
L - S

Quartile Deviation

$$\text{Quartile Deviation} = \frac{Q_3 - Q_1}{2}$$

$$\text{Where } Q_1 = \frac{N^{th}}{4} term,$$

$$Q_3 = 3 \times \left(\frac{N^{th}}{4}\right) term$$

Mean Deviation

For continuous Series
$$M.D. = \frac{\Sigma f|d|}{N}$$

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

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

Standard Deviation

For Individual Series
$$S.D. = \sqrt{\frac{\Sigma d^2}{N}}$$

For Discrete & continuous Series
$$S.D. = \sqrt{\frac{\Sigma f d^2}{N}}$$

For assumed Mean Method
$$S.D. = \sqrt{\frac{\Sigma f d^2}{N} - \left[\frac{\Sigma f d}{N}\right]^2}$$

Where d = deviation from assumed mean

Important Questions

Multiple Choice Questions-

1. When it comes to comparing two or more distributions we consider:

   (a) Absolute measures of dispersion

   (b) Relative measures of dispersion

   (c) Both (a) and (b)

   (d) Either (a) or (b)

2. The most commonly used measure of dispersion is:

   (a) Range

   (b) Standard deviation

   (c) Coefficient of variation

   (d) Quartile deviation

3. The range of 15, 12, 10, 9, 17, 20 is:

   (a) 5

   (b) 12

   (c) 13

   (d) 11

4. The standard deviation of 10, 16, 10, 16, 10, 10, 16, 16 is:

   (a) 4

   (b) 6

   (c) 3

   (d) 0

5. Corresponding to first quartile, the cumulative frequency is:

   (a) N/2

   (b) N/4

   (c) 3N/4

   (d) None of these

6. What is the coefficient of range for the following wages (in `) of 8 workers?

   80, 65, 90, 60, 75, 70, 72, 85

(a) 35

(b) 25

(c) 30

(d) 20

7. ______ divide the total number of observations into 4 equal parts.

(a) median

(b) deciles

(c) quartiles

(d) percentiles

8. Above upper quartile, the frequency is equal to:

(a) N/4

(b) N/2

(c) 3N/4

(d) None of these

9. What is the value of mean deviation about mean for the following numbers?

5, 8, 6, 3, 4

(a) 5.20

(b) 7.20

(c) 1.44

(d) 2.23

10. ________ is equal to the value corresponding to cumulative frequency (N + 1)/4 from simple frequency distribution.

(a) Median

(b) 1st quartile

(c) 3rd quartile

(d) 1st decile

Very Short Questions:

1. Define dispersion.

2. What is the coefficient of dispersion?

3. Define range.

4. Explain the interquartile range.
5. What is the quartile deviation?
6. What is the formula for calculating the coefficient of quartile deviation?

Short & Long Questions:

1. Define mean deviation.
2. What is standard deviation?
3. What is a Lorenz curve?
4. Define variance.
5. A measure of dispersion is a good supplement to the central value in understanding a frequency distribution. Comment.
6. Which measure of dispersion is the best and how?
7. Some measures of dispersion depend upon the spread of values whereas some calculate the variation of values from a central value. Do you agree?
8. In town, 25% of the persons earned more than ₹ 45,000 whereas 75% earned more than 18,000. Calculate the absolute and relative values of dispersion.

ANSWER KEY

Multiple Choice Answers-

1. B
2. B
3. D
4. C
5. B
6. D
7. C
8. A
9. C
10. B

Very Short Answers:

1. Dispersion is the measure of the extent to which different items tend to dispense away from the central tendency.

2. The coefficient of dispersion shows different data percentages or relative values. It is known as a relative measure of dispersion.
3. Range is the variance between the lowest and the highest values in a series.

   Range = Highest value in the series – Lowest value in the series.
4. The difference between the first quartile (Q_1) and the third quartile (Q_3) of a series is known as the interquartile range
5. The half of the interquartile range is quartile deviation. It can also be mentioned as a semi-interquartile range..
6. For calculating the coefficient of quartile deviation, the following formula is applied:

   $Q_3 - Q_1/Q_3 + Q_1$

Short Answers:

1. The average of the deviations of all the principles taken from some average value (mean, median, mode) of the series by ignoring the signs (+ or -) of the deviation is known as the mean deviation.
2. It is the square root of the arithmetic mean of the squared deviations of the items from their mean value.
3. A Lorenz curve is a curve that shows the actual distribution deviation (of income or wealth) from the line exhibiting equal distribution.
4. Variance is another measure of dispersion. It is the square of the standard deviation.

   Stay tuned for questions papers, sample papers, syllabus, and relevant notifications on our website.

5. Dispersion is the extent to which values in a distribution differ from the avarage of the distribution. Knowledge of only average is insufficient as it does not reflect the quantum of variation in values.

   Measures of dispersion enhance the understanding of a distribution considerably by providing information about how much the actual value of items in a series deviate from the central value, e.g., per capita income gives only the average income but a measure of dispersion can tell you about income inequalities, thereby improving the understanding of the relative living standards of different sections of the society. Through value of dispersion one can better understand the distribution.

   Thus a measure of dispersion is a good supplement to the central value in understanding a frequency distribution.

6. Standard Deviation is considered to be the best measure of dispersion and is therefore the most widely used measure of dispersion.

- It is based on all values and thus provides information about the complete series. Because of this reason, a change in even one value affects the value of standard deviation.
- It is independent of origin but not of scale.
- It is us'eful in advanced statistical calculations like comparison of variability in two data sets.
- It can be used in testing of hypothesis.
- It is capable of further algebraic treatment.

7. Yes, it is true that some measures of dispersion depend upon the spread of values, whereas some calculate the variation of values from the central value. Range and Quartile Deviation measure the dispersion by calculating the spread within which the value lie. Mean Deviation and Standard Deviation calculate the extent to which the values differ from the average or the central value.

8. 25% of the persons earned more than ₹ 45,000. This implies that upper quartile Q_3 = 45,000 75% earned more than 18,000. This implies that lower quartile Q1 =18,000

Absolute Measure of Dispersion = $Q_3 - Q_1$ = 45,000 – 18,000 = 27,000

Relative Measure of Dispersion

Co-efficient of Quartile Deviation

$$= \frac{Q_3 - Q_1}{Q_3 + Q_1} \quad \text{where } Q_3 = \text{3rd Quartile}, \; Q_1 = \text{1st Quartile}$$

$$= \frac{45{,}000 - 18{,}000}{45{,}000 + 18{,}000}$$

$$= \frac{27{,}000}{63{,}000} = 0.428$$