Quadrant II – Transcript and Related Materials

Programme: Bachelor of Arts (Third year)

Subject: Economics

Paper Code: ECD 114 (Honours)

Paper Title: Research Methodology II

Unit II : Measures of Central Tendency and Dispersion

Module Name: Dispersion

Module No: 11

Name of the Presenter: Ms. Yogita Eknath Rao

Notes

MEASURES OF DISPERSION

An average can represent a series only as best as a single figure can, but it certainly cannot reveal the entire story of any phenomenon under study. Specially it fails to give any idea about the scatter of the values of items of a variable in the series around the true value of average. In order to measure this scatter, statistical devices called measures of dispersion are calculated. Important measures of dispersion are

a) Range
b) Inter-quartile Range
c) Quartile Deviation

A) RANGE

Range is the simplest possible measure of dispersion and is defined as the difference between the values of extreme items of series.

Thus,

Range = H (Highest value of an item in a series) – L (Lowest value of an item in a series)

The utility of range is that it gives an idea of the variability very quickly, but the drawback is that range is affected very greatly by fluctuations of sampling. Its value is never stable, being based on only two values of the variable. This measure requires ordering of scores according to size. It can thus he applied in cases where the distribution is at least on an ordinary level of measurement.

For example,

Problem 1 : In a set of values 3, 5, 7, 9, 12, the range is difference between 12 and 3

Range = H (Highest value of an item in a series) – L (Lowest value of an item in a series)

= 12 – 3

= 9

Problem 2 : Calculate the range of A's monthly earnings for a year

Table No. 1 : A's monthly earnings for a year

Month	Monthly earnings (in '00 Rs.)
1	139
2	150
3	151
4	151
5	157
6	158
7	160
8	161
9	162
10	162
11	173
12	175

Solutions:

Highest earnings (H) = Rs. 17,500

Lowest earnings (L) = 13,900

Range = H – L

= 17500 – 13900

= 3,600

Problem 3: The following table gives the age distribution of a group of 50 individuals.Calculate the Range.

Table Number 2: Age distribution of a group of 50 individuals

Age (in years)	16 – 20	21 - 25	26 – 30	31 - 36
No. of persons	10	15	17	8

Solution: Since age is a continuous variable we should first covert the given classes into the continuous classes. The first class will then become 15.5 - 20.5 and the last class will become 30.5 – 35.5.

Highest Value = 35.5

Lowest Value = 15.5

Range = H – L

Range = 35.5 -15.5

= 20 years

Merits:

1. The range is easy to compute and understand. It is a useful device for gaining a quick impression of dispersion.
2. It is useful for computation of standard deviation, the most important measure of deviations.

Demerits:

1. Range is not based on the entire set of data. It is based on the two extreme values only, even if one of the extreme value changes, the range will fluctuate sharply. It is thus, sensitive to extreme values.
2. The ranges cannot be compared with each other unless they are based on a similar number of observations.

3. The range does not give us any information about the pattern of variation.

Uses

1. In spite of the above limitations and shortcomings range, as a measure of dispersion, has its applications in a number of fields where the data have small variations like the stock market fluctuations, the variations in money rates and the rate of exchange.
2. Range is used in industry for the statistical quality control of the manufactured product by the construction of R-chart i.e., the control chart range.
3. Range is by far most widely used measure of variability in our day-to-day life. For example the answer to the problems like, 'daily sales in a departmental store', monthly wages of workers in a factory, is usually provided by the probable limits.

QUARTILES

Quartiles are measures of location which divide a distribution into four equal parts. The first quartile, Q1 is the point which has 25% of the cases below it and 75% above it. The second quartile is of course the median. The third quartile Q3, has 75% of the cases below it and 25% of the cases above it. The following formulas are used for computing Q1 and Q2.

$$Q_1 = 1 + \left[\frac{n/4 - cf}{f} X(i) \right]$$

$$Q_3 = 1 + \left[\frac{3n/4 - cf}{f} X(i) \right]$$

Where '1' is the lower limit of the first and the third quartile classes respectively;

F = frequency of the first and third quartile classes respectively.

Example

Consider the data from table number 3 and find out Quartile 1 and Quartile 2 respectively.

Table No. 3: Distribution of Respondents by Age

Age Group	Frequency	Cumulative Frequency
1-20	15	15
21-40	32	47
41-60	54	101
61-80	30	131
81-100	19	150
Total	**150**	

Where: Q1 = Quartile 1; Lower limit of Q1 = 21; n = 150; c.f Q1 = 15, i = 20; f= 32

$$Q_1 = 21 + \left[\frac{37.5 - 15}{32} X\, 20 \right] = 21 + 14 = 35$$

Thus,

Where: Q2 = Quartile 2; Lower limit of Q2 = 61; n = 150; c.f Q2 = 101, i = 20; f= 30

$$Q_3 = 61 + \left[\frac{112.5 - 101}{30} X\, 20 \right] = 61 + 7.6 = 68.6$$

Thus,

Thus the value of Quartile 1 (Q1) and Quartile 2 (Q2) is 35 and 68.6 respectively.

B. INTERQUARTILE RANGE

The inter-quartile range is the difference between the third and first quartiles (Q3 - Q1). It measures the spread in the middle half of the distribution, and is less affected by extreme values. It is thus a more stable measure of dispersion that is the range.

Problem 1. Consider the data from the table number 3.

Solution : Q3 and Q1 for those data are 68.6 and 35.0 respectively. Hence the inter-quartile range for them is (Q3) 68.6 and (Q1) is 35.0

Interquartile Range = (Q3 – Q1)

= (68.6 – 35.0)

$$= 33.6$$

The principle of the inter-quartile range can be applied to any parts of the distribution.

Merits : it has an advantage over range, in as much as, it is not affected by the values of the extreme items. In fact 50% of the values of a variable are between the two quartiles and as such the inter-quartile range gives a fair measure of variability. But inter-quartile range is easy to calculate and is readily understood.

Demerits : The inter-quartile range suffers from the same defects from which range suffers. It is a measure of location, and its value is not very stable.

QUARTILE DEVIATION OR SEMI-INTER-QUARTILE RANGE

It is one-half of the difference between the third quartile and first quartile.

Quartile Deviation = Q3- Q1/2

Where Q3 and Q1 stand for the upper and lower quartile respectively.

Quartile deviation is an absolute measure of dispersion. If it is divided by the average value of two quartiles, a relative measure of dispersion is obtained. It is called the Co-efficient of Quartile Deviation.

Symbolically,

$$\textbf{Quartile Deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

This is usually used in measuring Arc elasticity of demand in economics.

Consider the data from the table number 3.

Q3 and Q1 for those data are 68.6 and 35.0 respectively. Hence the inter-quartile range for them is (Q3) 68.6 and (Q1) is 35.0.

$$\textbf{Quartile Deviation} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$=\frac{68.6-35.0}{68.6+35.0}$$

$$=\frac{33.6}{103.6}$$

$$=0.3243$$

Merits of Quartile Deviation

Quartile deviation is quite easy to understand and calculate. It has a number of advantages over range as a measurement of dispersion.

Merits:

1. As against range which was based on two observations only. Q.D. makes use of 50% of the data and as such is obviously a better measure than range.
2. Since Q.D. ignores 25% of the data from the beginning of the distribution and another 25% of the data from the beginning of the distribution and another 25% of the data from the top end, it is not affected at all by extreme observations.
3. Q.D. can be computed from the frequency distribution with open end classes. In fact, Q.D. is the only measure of dispersion which can be obtained while dealing with a distribution having open end classes.

Demerits:

1. Q.D. is not based on all the observations since it ignores 25% of the data at the lower end and 25% of the data at the upper end of the distribution. Hence, it cannot be regarded as relative measures of variability.
2. Q.D. is affected considerably by fluctuations of sampling.
3. Q.D. is not suitable for further mathematical treatment.

Thus quartile deviation is not a reliable measure of variability, particularly for distribution in which the variation is considerable.