PHYSICAL EDUCATION

Subject : Physical Education
Semester : 4^{th}
Paper No. and Title : (401) Test, Measurement and Evaluation in Physical Education

Unit-4: Measures of Dispersion

4.0 Objectives:

Measures of central tendency give you one single figure to represent the entire series of data. But, it alone is not sufficient to describe a data set. Measure of dispersion is the next important measure to describe a data set. After reading this unit you will be able to:

- understand importance of dispersion
- define measures of dispersion
- calculation of different measures of dispersion

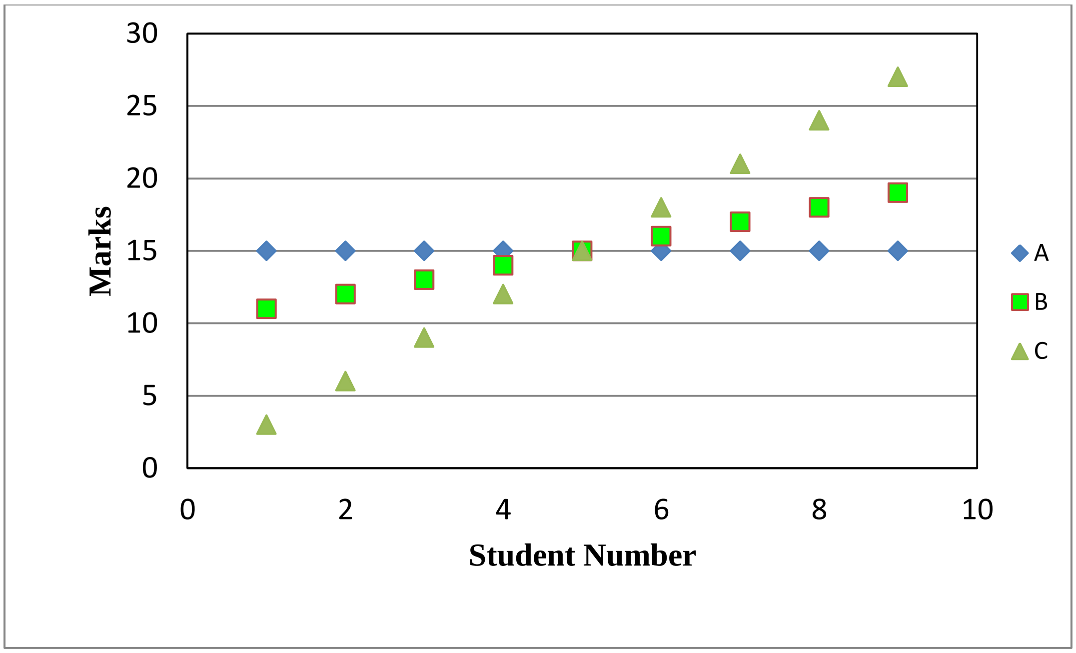
4.1 Measures of Dispersion- Why?

The measures of central tendency give us the concentration of the observations about the central part of the distribution. Sometimes, using these measures it is difficult to form a complete idea about the distribution. For example, suppose, we have three series A, B and C as given in the following table

Series										Total	Mean
A	15	15	15	15	15	15	15	15	15	135	15
B	11	12	13	14	15	16	17	18	19	135	15
C	3	6	9	12	15	18	21	24	27	135	15

Here, sample size= 9 and Mean =15 for each of these series.

The graph obtained from these series is



Thus, if, we are given that mean of a series of 9 observations is 15, then we cannot determine if we are talking of the series A, B or C, as all the three series have same mean 15. So, measures of central tendency are inadequate here. To overcome this, it must be supported and supplemented by some other measures. One of them is “Dispersion”. The literal meaning of dispersion is “scatteredness”.The term “dispersion” means the deviation or difference or spread of certain values from their central value. Dispersion is necessary to have an idea about homogeneity or heterogeneity of the distribution.

To quantify the extent of the variation/ difference/ deviation, there are certain measures namely:

(i) Range

(ii) Quartile Deviation

(iii) Mean Deviation

(iv) Standard Deviation

Range and Quartile Deviation measure the dispersion within which the values of a series lie, whereas, Mean Deviation and Standard Deviation calculate the extent to which the values of the series differ from the average. Here, calculation of three measures- range, quartile deviation and standard deviation have been considered as mean deviation is rarely used in real life situations.

(i) Range:

Range (R) is the difference between the largest (L) and the smallest value (S) in a series of data. Thus, R = L – S. Higher value of Range implies higher dispersion and vice-versa. For example: The range(s) of the three series are:

Series										**L**	**S**	**(R)**
A	**15**	**15**	**15**	**15**	**15**	**15**	**15**	**15**	**15**	15	15	0

B	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**19**	**11**	**8**
C	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27**	**27**	**3**	**24**

This is an example of ungrouped data.

In case of grouped data, the range is the difference between the upper boundary of the highest class and the lower boundary of the lowest class. It is also calculated by using the difference between the mid points of the highest class and the lowest class.

(ii) Quartile Deviation:

<u>Quartiles:</u>

The measures which divide the data into four equal parts after arranging the data in ascending or descending order of magnitude is termed as quartiles, where, each portion contains equal numbers of observations. The first quartile (denoted by Q_1) or lower quartile has 25% of items of the distribution below it and 75% of the items are greater than it. Th0e second quartile (denoted by Q_2) or median has 50% of items below it. The third quartile (denoted by Q_3) or upper quartile has 75% of the items of the distribution below it. Thus, Q_1 and Q_3 denoted the two limits within which central 50% of the data lies.

<u>Calculation of Quartiles and Quartile Deviations (Q.D.):</u>

The method for locating the quartile is same as that of the median in case of individual and discrete series. The values of quartiles of an ordered series can be obtained by the following formula where N is the number of observations.

$$Q_1 = \frac{(N+1)th}{4} observation$$

$$Q_2 = Median = \frac{2(N+1)th}{4} observation$$

$$Q_3 = \frac{3(N+1)th}{4} observation$$

Example:

Calculate the value of lower quartile from the data of the marks obtained by ten students in an examination.

22, 26, 14, 30, 18, 11, 35, 41, 12, 31

Arrange the data in an ascending order: 11,12,14,18, 22, 26, 30, 32, 35, 41

$$Q_1 = size\ of\ \frac{(N+1)th}{4} observation = size\ of\ \frac{(10+1)th}{4} observation$$

$$= size\ of\ 2.75th\ observation$$

$$= 2nd\ observation + .75(3rd\ observation - 2nd\ observation) = 12 + .75(14 - 12)$$

$$= 13.5$$

$Q_2 = Median = \frac{2(N+1)th}{4} observation = \frac{22+26}{2} = 24$

$Q_3 = size\ of\ \frac{3(N+1)th}{4} observation$ = size of $\frac{3(10+1)}{4} = \frac{33}{4}$ = size of 8.25th $observation$

= 8th $observation$ + 0.25(9th $observation$ – 8th $observation$) = 32 + 0.25(35-32) = 32.75

Now, the quartile deviation or semi interquartile range Q.D. is given by

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

For the above example the quartile deviation is = $\frac{(Q_3 - Q_1)}{2} = \frac{32.75 - 13.5}{2} = 19.25$

The above example is a case for ungrouped data. Let us consider the case of grouped data.

Calculate the quartile deviation for the following data:

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of Students	6	5	8	15	7	6	3

Calculation:

Marks	Mid value X	No. of Students (f)	$d = \frac{x-35}{10}$	fd	$\lvert x - \bar{x} \rvert = \lvert x - 33.4 \rvert$	$f \lvert x - \bar{x} \rvert$	Less than c.f
0-10	5	6	-3	-18	28.4	170.4	6
10-20	15	5	-2	-10	18.4	92.0	11
20-30	25	8	-1	-8	8.4	67.2	19
30-40	35	15	0	0	1.6	24.0	34
40-50	45	7	1	7	11.6	81.2	41
50-60	55	6	2	12	21.6	129.6	47
60-70	65	3	3	9	31.6	94.8	50

Total	----	50	----	-8		659.2	

Here N=50; $\frac{1}{4}N = 12.75$; $\frac{3}{4}N = 37.25$

The c.f just greater than 12.75 is 19. Hence, the corresponding class 20-30 contains Q_1.

Therefore, $Q_1 = 20 + \frac{10}{8}(12.75 - 11) = 22.19$

The c.f. just greater than 37.25 is 41. Hence, the corresponding class 40-50 contains Q_3.

Therefore, $Q_3 = 40 + \frac{10}{7}(37.25 - 34) = 44.64$

Hence, $Q.D. = \frac{1}{2}(Q_3 - Q_1) = \frac{1}{2}(44.64 - 22.19) = 11.23$

Like quartiles two other partition values are deciles and percentiles. Let us define these statistics in brief.

Deciles and Percentiles:

The descriptive statistics, deciles divide a set of data into ten equal parts after arranging the data in order (ascending or descending). So, there are nine decile values. These are generally denoted by- D_1 , D_2 , D_3 , D_4 , D_5 , D_6 ,D_7 , D_8 and D_9. Similarly, there are ninety nine partition values- P_1 , P_2 , P_3 etc., which divide whole set of data into hundred equal parts after arranging the data in order (ascending or descending). The same calculation procedure with slight modification (e.g. N/10 in case of deciles and N/100 in case of percentiles instead of N/2 as in case of median) can be applied in calculation of these measures.

(iii) Standard Deviation

*Standard deviation is the most important and widely used measure of dispersion. Standard Deviation*of a set of values is the positive square root of the A.M. of squared deviation when deviations are taken from mean. It is denoted by σ(Greek small letter sigma). It is the positive square root of variance. Thus *SD* = √variance.

Ungrouped data: If $x_1, x_2, \ldots, x_n$ are values of x then-

$$\sigma = \sqrt{\frac{1}{n}\sum (x_i - \bar{x})^2}$$

$$= \sqrt{\frac{1}{n}\sum x_i{}^2 - \bar{x}^2}$$

Grouped data: If $x_1, x_2, \ldots, x_n$are the values of x with respective frequencies $f_1, f_2, \ldots, f_n$ then

$$\sigma = \sqrt{\frac{1}{N}\sum f_i (x_i - \bar{x})^2}$$

$$= \sqrt{\frac{1}{N}\sum f_i x_i^2 - \bar{x}^2}$$

Example of standard deviation for ungrouped data: Consider the series A, B and C again.

Then the standard deviations are

Series										Total	Mean	Standard deviation
A	**15**	**15**	**15**	**15**	**15**	**15**	**15**	**15**	**15**	135	15	**0**
B	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**135**	**15**	**2.57**
C	**3**	**6**	**9**	**12**	**15**	**18**	**21**	**24**	**27**	**135**	**15**	**7.75**

For Group Data: Find the standard deviation from the following frequency distribution.

Value	12-17	17-22	22-27	27-32	32-37	37-42	42-47
frequency	2	22	19	14	3	4	6

Solution:

First we construct the following table:

Class Interval	**Mid-value** x_i	**Frequency** f_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
12-17	14.5	2	-15	225	450
17-22	19.5	22	-10	100	2200
22-27	24.5	19	-5	25	475

27-32	29.5	14	0	0	0
32-37	34.5	3	5	25	75
37-42	39.5	4	10	100	400
42-47	44.5	6	15	225	1350
	$\bar{x}$ =29.5	$\sum f_i$=70			$\sum f_i(x_i - \bar{x})^2$ =4950

Now,

$$\text{Standard deviation} = \sigma = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i=n}}^{\mathrm{n}}(\mathrm{x_i} - \bar{\mathrm{x}})^2}$$

$$\sigma = \sqrt{\frac{4950}{70}} = 8.40$$

The required standard deviation is 8.40