Measures of Dispersion

Basuki Nath Jha (B.Ed)

Sem– 2, TC- 204

Godda College Godda

An average gives an idea of central tendency of the given distribution but it is necessary to know how the variates are clustered around or scattered away from the average. To explain it more clearly we consider the works of two typists who typed the following number of pages in 6 working days of a week:

	Mon	Tue	Wed	Thurs	Fri	Sat	Total Pages
1 Typist:	**15**	**20**	**25**	**25**	**30**	**35**	**150**
2 Typist:	**10**	**20**	**25**	**25**	**30**	**40**	**150**

We see that each of the typists 1 and 2 typed 150 pages in 6 working days and so the average in both the cases is 25. Thus there is no difference in the average, but we know that in the first case the number of pages varies from 15 to 35 while in the second case the number of pages varies from 10 to 40. This denotes that the greatest deviation from the mean in the first case is 10 and in the second case it is 15 i.e., there is a difference between the two series. The variation of this type is termed scatter or dispersion or spread. Thus we say, "The degree to which numerical data tend to spread about an average value is called variation or dispersion or spread of the data."

Various measures of dispersion are available, the most common are following:

The Range

It is the simplest possible measure of dispersion. The range of a set of numbers (data) is the difference between the largest and the least numbers in the set i.e. values of the variable. If this difference is small then the series of numbers is supposed regular and if this difference is large then the series is supposed to be irregular.

Example: *Compute the range for the following observation*

15 20 25 25 30 35

Solution: Range = Largest – Smallest

= 35 - 15

= 20

Quartile Deviation

The quartile deviation of a set of data is defined by

Quartile Deviation (Q) = $(Q_3 – Q_1)/2$

Where Q1 and Q3 are respectively the first and third quartiles for the data and Quartile deviation (or Semi-inter quartile range) is denoted by Q .

The Quartile Deviation is a better measure of dispersion than the range and is easily computed. Its drawback is that it does not take into account all the items.

Example 2: Compute the Quartile Deviation of the marks of 63 students in Mathematics given below:

Marks Group	No. of Students	Marks Group	No. of students
0-10	5	50-60	7
10-20	7	60-70	3
20-30	10	70-80	2
30-40	16	80-90	2
40-50	11	90-100	0

Solution:

Marks Group	Frequency f	Cumulative Frequency c.f.
0-10	5	5
10-20	7	12
20-30	10	22
30-40	16	38
40-50	11	49
50-60	7	56

60-70	3	59
70-80	2	61
80-90	2	63
90-100	0	63
	$\sum f = 63$	

To calculate first Quartile Q1. Here N = 60. So ¼ (N+1)th i.e., 16th students lies in the marks group 20-30. Thus lower quartile class is 20-30.

$$\textbf{Q1} = l + \frac{\frac{1}{4}N - F}{f} \times i = 20 + \frac{15.75 - 12}{10} \times 10 = 23.75.$$

Similarly, $\textbf{Q3} = 40 + \frac{47.25 - 38}{11} \times 10 = 48.4$

Quartile Deviation (**Q**) = ½ (Q3 – Q1)

= ½ (48.4-23.75) = 12.32 (Marks)

Mean Deviation

The average (or mean) deviation about any point M, of a set of N numbers $x_1, x_2, \ldots, x_N$ is defined by

$$\textbf{Mean Deviation (M. D.)} = \boldsymbol{\delta_m} = \frac{1}{N}\sum_{i=1}^{n} x_i - M$$

where M is the mean or median or mode according as the mean deviation from the mean or median or mode is to be computed, $| x_i - M |$ represents the absolute (or numerical) value.

If $x_1, x_2, \ldots, x_k$ occur with frequencies $f_1, f_2, \ldots, f_k$ respectively, then the mean deviation (δ_m) is defined by

$$\boldsymbol{\delta_m} = \frac{1}{N}\sum_{j=1}^{k} f_j (x_j - M) = \frac{1}{N}\sum f (x - M)$$

Mean deviation depends on all the values of the variables and therefore it is a better measure of dispersion than the range or the quartile deviation. Since signs of the deviations are ignored (because all deviations are taken positive), some artificiality is created.

In case of grouped frequency distribution the mid-values are taken as x.

Example 1. Find the mean deviation from the arithmetic mean of the following distribution:

Marks	:	**0 - 10**	**10 - 20**	**20 - 30**	**30 - 40**	**40-50**
No. of students	**:**	**5**	**8**	**15**	**16**	**6**

Solution: Let assumed mean A = 25 and i = 10

Class	Mid value X	Frequency f	$u = \frac{x - A}{i}$	fu	x -M	f l x - M l
0-10	5	5	-2	-10	-22	110
10-20	15	8	-1	-8	-12	96
20-30	25	15	0	0	-2	30
30-40	35	16	1	16	8	128
40-50	45	6	2	12	18	108
Total		$\sum f = 50$		$\sum fu = 10$		$\sum f \mid x - M \mid = 472$

Arithmetic mean $\mathbf{M} = u + \frac{\sum fu}{N} \times i = 25 + 10/50 \times 10 = 27.$

Mean deviation from arithmetic mean

$$\delta_m = \frac{\sum f(x - M)}{N} = 472 / 50 = 9.44$$

Standard Deviation:

The standard deviation or SD is the most stable index of variability and is customarily employed in experimental work in research studies. The SD differs from Mean deviation in several respects. In computing the MD, we disregard signs and treat all deviations as positive, whereas in finding the SD we avoid the difficulty of signs by squaring the separate deviations.

The formula for Standard Deviation (SD):

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

The End