Mr. Mohammad Manjur Alam (Manju)
Assistant Professor
Department of Computer Science and Engineering
International Islamic University Chittagong.
Mobile: 01715966144
Email: manjuralam44@yahoo.com

CHAPTER THREE

Measure of Dispersion Or Measure of Variability

Define measures of dispersion.

Ans: Dispersion is an important characteristic of frequency distributions. It tells how completely the individual values are distributed around the average.
According to Dr. A. L. Bowley, "Dispersion is the measure of variation of the items."

Spiegel defined, "The degree to which the numerical data tends to spread about an average value is called dispersion or variation of data."

Importance of measure of dispersion

The central value like mean is generally used to convey the general behavior of a data set. For example, the average score of the class in a math test hints at the general comfort level of the class in the topic which is tested. But, if two classes have the same average score, can the teacher conclude that the understanding level is same for the two classes? The arithmetic mean does not convey the variations displayed in the individual marks of the students. The teacher needs to have some idea about the spread of the marks of both the classes. Teacher needs some numerical measure of dispersion which would convey how the marks are spread about the central value, the mean in this case.

Q : A measure of dispersion is a good supplement to the central value in understanding a frequency distribution. Comment.

The study of the averages is only one sided distribution story. In order to understand the frequency distribution fully, it is essential to study the variability of the observations. The average measures center of the data whereas the quantum of the variation is measured by the measures of dispersion like range, quartile deviation, mean deviation and Standard Deviation. For example, if a country has very high income group people and very low income group people, then we can say that the country has large income disparity.

Types of measures dispersion.

There are two types of dispersions:

(I) Absolute measure of dispersion
(II) Relative measures of dispersion

Absolute measure of dispersion:

These measures give us an idea about the amount of dispersion in a set of observations. They give the answers in the same units as the units of the original observations.

Different types of absolute measures are:

(i) Range
(ii) **Semi-interquartile range** or Quartile deviation
(iii) Mean deviation
(iv) Variance and standard deviation

Relative measures of dispersion

(i) Coefficient of Range
(ii) Coefficient of quartile deviation
(iii) Coefficient of mean deviation
(iv) Coefficient of variation

Range:

Range is the difference of the highest and the lowest observations of the distributions.

Range = Maximum value- Minimum value.

Example:

12, 25, 27, 29, 36, 38, 40, 43, 50, 54, 62

Range = 62 - 12 = 50

For grouped data, range is the difference between the highest observation in the last class interval and the lowest observation in the first class interval. Range can never be zero.

Test-1: Find the range of the following data set:
$\{1;\ 4;\ 5;\ 8;\ 6;\ 7;\ 5;\ 6;\ 7;\ 4;\ 10;\ 9;\ 10\}$

Advantages of Range:

(i) It is very easy to understand and easy to calculate
(ii) It gives a quick idea about the variability of a set of data
(iii) It is the simplest of all measures of dispersion

Disadvantages of Range:

(i) It is very much affected by extreme values;
(ii) It provides us with an idea of only two extreme values in a set of data;
(iii) It cannot be computed for data set having open ended class interval

Quartile deviation/: Semi-interquartile range:

If Q_1 & Q_3 denote the first and third quartiles respectively of a frequency distribution, the quartile deviation of the data is defined by the relation -
Inter quartile range $=Q_3-Q_1$ and

Quartile deviation/: Semi-interquartile range, $Q.D = \dfrac{Q_3 - Q_1}{2}$

Advantages of Quartile deviation:
(i). It is easily understandable
(ii) The quartile deviation is not affected by extreme values.
(iii)It is not possible to compute-quartile deviation for open ended class

ROSE
2019
' Š á M ĭ '

Disadvantages of Quartile deviation:
(i). It is not based upon all the observations.
(ii) It is more affected by sampling fluctuations
(iii) It is not satiable for further algebraic treatment

Mean deviation: The average deviation is the arithmetic mean of the absolute values of the deviations from the mean/median/mode.

For ungrouped data, mean deviation –

a) about mean: $M.D.(\bar{x}) = \frac{\sum_{i=1}^{n}|x_i - \bar{x}|}{n}$ b) about median: $M.D.(M_e) = \frac{\sum_{i=1}^{n}|x_i - M_e|}{n}$

c) about mode: $M.D.(M_o) = \frac{\sum_{i=1}^{n}|x_i - M_o|}{n}$

For grouped data or frequency distributions,

Mean deviation –about mean: $M.D.(\bar{x}) = \frac{\sum_{i=1}^{n} f_i|x_i - \bar{x}|}{n}$ b) about median

$M.D.(M_e) = \frac{\sum_{i=1}^{n} f_i|x_i - M_e|}{n}$ c) about mode: $M.D.(M_o) = \frac{\sum_{i=1}^{n} f_i|x_i - M_o|}{n}$

Learn by heart: Mean deviation

Ungroup data	Group data
$M.D.(\bar{x}) = \frac{\sum_{i=1}^{n}\lvert x_i - \bar{x}\rvert}{n}$	$M.D.(\bar{x}) = \frac{\sum_{i=1}^{n} f_i\lvert x_i - \bar{x}\rvert}{n}$

Example-1: Mean Deviation of 3, 6, 6, 7, 8, 11, 15, 16

$$\bar{x} = \frac{3+6+6+7+8+11+15+16}{8} = \frac{72}{8} = 9$$

x	$\lvert x - \bar{x}\rvert$
3	6
6	3
6	3

7	2
8	1
11	2
15	6
16	7
	$\Sigma\lvert x - \bar{x}\rvert = 30$

Mean Deviation $= \dfrac{\Sigma\lvert x - \bar{x}\rvert}{n} = \dfrac{30}{8} = 3.75$

Variance: The variance is the arithmetic mean of the squared deviations from the mean. It is denoted by σ^2 Standard deviation: The positive square root of the variance is called standard deviation. It is denoted by σ Let $x_1, x_2, x_3, \ldots\ldots\ldots, x_n$ be the series of n observations, then variance, $\sigma^2 = \dfrac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$ And SD is,

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

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}(x_i^2 - 2x_i.\bar{x} + \bar{x}^2)}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}x_i^2 - 2\bar{x}\sum_{i=1}^{n}x_i + n\bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}x_i^2 - 2\bar{x}.n\bar{x} + n\bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}x_i^2 - 2n\bar{x}^2 + n\bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}x_i^2 - n\bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n}x_i^2}{n} - \bar{x}^2 = \frac{\sum_{i=1}^{n}x_i^2}{n} - \left(\frac{\sum_{i=1}^{n}x_i}{n}\right)^2$$

$$\boxed{\Rightarrow \sigma = \sqrt{\frac{\sum_{i=1}^{n}x_i^2}{n} - \left(\frac{\sum_{i=1}^{n}x_i}{n}\right)^2}} \ldots\ldots\ldots\ldots(1)$$

Equation (1) is called the working formula of standard deviation for ungrouped data.

Again, for grouped data or frequency distributions, variance is –

$$\sigma^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n}.$$

$\therefore$ **Standard deviation is –** $\sigma = \sqrt{\dfrac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n}}$ **Now,**

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

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i (x_i - \bar{x})^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i (x_i^2 - 2x_i.\bar{x} + \bar{x}^2)}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2 - 2\bar{x}.\sum_{i=1}^{n} f_i x_i + \sum_{i=1}^{n} f_i \bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2 - 2\bar{x}.n\bar{x} + n\bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2 - 2n\bar{x}^2 + n\bar{x}^2}{n}$$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2 - n\bar{x}^2}{n}$$

Here, $\bar{x} = \dfrac{\sum_{i=1}^{n} f_i x_i}{n} \Rightarrow n\bar{x} = \sum_{i=1}^{n} f_i x_i$

$$\Rightarrow \sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2}{n} - \bar{x}^2 = \frac{\sum_{i=1}^{n} f_i x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{n}\right)^2$$

$$\Rightarrow \boxed{\sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{n}\right)^2}} \text{............(2)}$$

Equation (2) is called the working formula for frequency distribution of grouped data.

Learn by heart: **Variance**

Ungroup data	Group data
$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$	$\sigma^2 = \frac{\sum_{i=1}^{n} f_i(x_i - \bar{x})^2}{n}$

Learn by heart: Standard deviation

Ungroup data	Group data
$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$	$\sigma = \sqrt{\frac{\sum_{i=1}^{n} f_i(x_i - \bar{x})^2}{n}}$

Table 1. Comparison of Sample Statistics and Population Parameters

	Sample Statistic	Population Parameter
Mean	$\bar{x}$	μ
Standard deviation	s	sigma
Variance	s²	sigma²

Example 1 – Variance and standard deviation (For ungroup data):

**Find variance and standard deviation for the following data set.

5, 10, 8, 12,20,24,25,15,16,22

Solution: we know,

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$= \frac{5+10+8+12+20+24+25+15+16+22}{10} = \frac{157}{10} = 15.7$$

Table for calculating variance and standard deviation:

X_i	$(X_i - \bar{x})$ $\bar{X} = 15.7$	$(X_i - \bar{x})^2$
5	-10.7	114.49
10	-5.7	32.49

8	-7.7	59.29
12	-3.7	13.69
20	4.3	18.49
24	8.3	68.89
25	9.3	86.49
15	-0.7	0.49
16	0.3	0.09
22	6.3	36.69
$\sum_{i=1}^{n} x_i = 157$		$\sum_{i=1}^{10}(x_i - \bar{x})^2 = 434.1$

We know, variance

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

$$= \frac{434.1}{10} = 43.41$$

And standard deviation

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

$$\sigma = \sqrt{\frac{434.1}{10}} = \sqrt{43.41} = 6.59$$

Alternative method of Variance and standard deviation

We know, Variance, $\sigma^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} x_i}{n}\right)^2$

X_i	x_i^2
5	25
10	100
8	64
12	144
20	400

24	576
25	625
15	225
16	256
22	484
$\sum_{i=1}^{n} x_i = 157$	$\sum_{i=1}^{n} x_i^2 = 2899$

Therefore, $\sigma^2 = \dfrac{\sum_{i=1}^{n} x_i^2}{n} - \left(\dfrac{\sum_{i=1}^{n} x_i}{n}\right)^2 = \dfrac{2899}{10} - \left(\dfrac{157}{10}\right)^2 = 289.9 - 246.49 = 43.41$

And standard deviation

$$\sigma = \sqrt{43.41}$$

$$= 6.59$$

Example 2 – Standard deviation (For ungroup data)

A hen lays eight eggs. Each egg was weighed and recorded as follows:

60 g, 56 g, 61 g, 68 g, 51 g, 53 g, 69 g, 54 g.

1. First, calculate the mean:

$$\bar{x} = \frac{\sum x}{n}$$

$$= \frac{472}{8}$$

$$= 59$$

Now, find the standard deviation.

Table 1. Weight of eggs, in grams		
Weight (x)	**$(x - \bar{x})$**	**$(x - \bar{x})^2$**
60	1	1
56	-3	9
61	2	4
68	9	81
51	-8	64
53	-6	36
69	10	100

54	-5	25
472		320

2. Using the information from the above table, we can see that

$\sum (x - \bar{x})^2 = 320$

In order to calculate the standard deviation, we must use the following formula:

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

$$= \sqrt{\frac{320}{8}}$$

= 6 32 grams

Exercises 1–6, find variance and standard deviation of the following set of measurements.

(1) 5, 12, 7, 14, 8, 9, 7
(2) 30, 37, 32, 39, 33, 34, 32
(3) 5, 12, 7, 24, 8, 9, 7
(4) 20, 37, 32, 39, 33, 34, 32
(5) 5, 12, 7, 14, 9, 7
(6) 30, 37, 32, 39, 34, 32

Variance and standard deviation (For group data):

Example-1: Find variance and standard deviation for the following data set.

Class interval	5-10	10-15	15-20	20-25
Frequency	2	5	8	3

Solution: Table for calculating variance and standard deviation:

Class interval	**Frequency(f)**	**Mid value (x_i)**	$f_i x_i$	x_i^2	$f_i x_i^2$
5-10	2	7.5	15	56.25	112.5
10-15	5	12.5	62.5	156.25	781.25
15-20	8	17.5	140	306.25	2450
20-25	3	22.5	67.5	506.25	1518.75
			$\sum_{i=1}^{n} f_i x_i = 285$		$\sum_{i=1}^{n} f_i x_i^2 = 4862.5$

We know, Variance, $\sigma^2 = \frac{\sum_{i=1}^{n} f_i x_i^2}{n} - \left(\frac{\sum_{i=1}^{n} f_i x_i}{n}\right)^2$

$= \frac{4862.5}{18} - \left(\frac{285}{18}\right)^2 = 270.89 - (15.8)^2 = 270.1389 - 249.64$ And standard deviation

$= 20.49889$

$\sigma = \sqrt{20.49889} = 4.53$

Exercise-1: Find variance and standard deviation for the following data set.

Class interval	10-20	20-30	30-40	40-50	50-60
Frequency	5	8	9	5	3

Properties of standard deviation

1. The standard deviation is zero if all the observations under study are same
2. Standard deviation is independent of change of origin but not of scale
3. For two observations, standard deviation is the half of the range $\left|\frac{x_1 - x_2}{2}\right|$
4. It is suitable for further algebraic treatment
5. Standard deviation of n natural numbers is $\sqrt{\frac{n^2 - 1}{12}}$

If x= mean, **S** = standard deviation and **x** = a value in the data set, then

- About 68% of the data lie in the interval: $\bar{x}$- **S** < **x** < $\bar{x}$+ S.
- About 95% of the data lie in the interval: x- **2S** < **x** <x + 2S.
- About 99% of the data lie in the interval: x- **3S** < **x** < x+ 3S.

Relative measure of dispersion:

These measures are calculated for the comparison of dispersion in two or more than two sets of observations. These measures are free of the units in which the original data is measured. These measures are a sort of ratio and are called coefficients.

Different types of relative measures are:

i) **Co efficient of range:**

The co-efficient of range is defined as $C.R. = \frac{x_{max} - x_{min}}{x_{max} + x_{min}} \times 100$ where, x_{max} and x_{min} is the highest and lowest observation respectively

ii) Co-efficient of quartile deviation:

The co-efficient of quartile deviation is defined as $C.Q.D. = \frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$, where Q_3 and Q_1 is the third and first quartile respectively.

iii)Co-efficient of mean deviation:

The co-efficient of mean deviations are:

a) about mean, defined as $C.M.D.(\bar{x}) = \frac{M.D.(\bar{x})}{\bar{x}} \times 100$

b) about median, defined as $C.M.D.(M_e) = \frac{M.D.(M_e)}{M_e} \times 100$

c) about mod, as defined as $C.M.D.(M_o) = \frac{M.D.(M_o)}{M_o} \times 100$

Learn by heart: iv) Co-efficient of variation:

The **coefficient of variation** is the **ratio** between the **standard deviation** of a sample and its **mean**. i.e, $C.V = \frac{\sigma}{\bar{x}}$

The coefficient of variation is usually expressed in percentages:

$$C.V = \frac{\sigma}{\bar{x}} \times 100$$

It allows us to compare the dispersions of two different distributions if their means are positive.

Example to explain the coefficient of variation

Problem-1: A distribution is $\bar{x} = 140$ and $\sigma = 28.28$ and the other is $\bar{x} = 150$ and $\sigma = 24$

Which of the two has a greater dispersion?

Example-1: Suppose the mean height of 200 students was 150 Cm with a standard deviation of 12 cm and mean weight was 60 kg with a standard deviation of 7 kg.

Then Coefficient of variation of height = `12/150` x 100 = 8 %

Coefficient of variation of weight = `7/60` x 100 = 11.67 %

Thus weight has a higher variability than the height. The coefficient of variation is used to compare the variability, homogeneity, stability, uniformity and consistency of two series. Smaller the coefficient of variation, higher is the consistency, homogeneity, stability, uniformity etc.

SHOT ON REDMI NOTE 5
MI DUAL CAMERA

2019/7/2 15:06

Example-2: If the laboratory technician A completes 40 analyses with a standard deviation of 5 and technician B completes 160 analyses per day with standard deviation of 15, find which employee shows less variability.

Solution: Compute the coefficient of variation for both technicians

C. V for technician A = `5/40` x 100 = 12.5 %

C. V for technician B = `15/60` x 100 = 9.4 %

So, we find that technician B less variation than technician A.

Uses of Co-efficient of variation:

Without an understanding of the relative size of the standard deviation compared to the original data, the standard deviation is somewhat meaningless for use with the comparison of data sets. To address this problem the coefficient of variation is used.

Problem-1: Two cricketers scored the following runs in randomly selected 10 one day matches:

Player-A	42	32	40	45	17	83	59	64	76	72
Player-B	95	3	28	70	31	14	82	0	59	108

(i) Who is the better run-getter?
(ii) Who is the consistent player?
(iii) A prize is given to the best player. Who will get the prize?

Solution: Table for calculating coefficient of variation:

Player-A (x_i)	Player-B (y_i)	x_i^2	y_i^2
42	95	1764	9025
32	3	1024	9
40	28	1600	784
45	70	2025	4900
17	31	289	961
83	14	6889	196
59	82	3481	6724
64	0	4096	0
76	59	5776	3481
72	108	5184	11664
$\sum_{i=1}^{n} x_i = 530$	$\sum_{i=1}^{n} y_i = 490$	$\sum_{i=1}^{n} x_i^2 = 32127$	$\sum_{i=1}^{n} y_i^2 = 37744$

(i) Cricketer-A: $\bar{X}_A = \frac{\sum_{i=1}^{n} x_i}{n}$

Scanned by CamScanner

$$= \frac{530}{10}$$
$$= 53$$

Cricketer: $\overline{Y_B} = \frac{\sum_{i=1}^{n} y_i}{n}$

$$= \frac{490}{10}$$
$$= 49$$

Since the average score for the player A is higher than player B, hence A is a better run-getter.

(ii) We know that,

$$\sigma_A^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\overline{x}\right)^2$$

$$= \frac{32127}{10} - (53)^2 = 3212.7 - 2809 = 403.7$$

Therefore $\sigma_A = 20.09$

Similarly,

$$\sigma_B^2 = \frac{\sum_{i=1}^{n} y_i^2}{n} - \left(\overline{y}\right)^2$$

$$= \frac{37744}{10} - (49)^2 = 3774.4 - 2401 = 1373.4$$

Therefore $\sigma_A = 37.05$

C.V (A) = $\frac{\sigma_A}{\overline{x_A}} \times 100 = = \frac{20.09}{53} \times 100 = 37.90\%$

And C.V (B) = $\frac{\sigma_B}{\overline{x_B}} \times 100 = = \frac{37.05}{49} \times 100 = 75.61\%$

The coefficient variation for player A is less than player B; hence player A is more consistent.

(iii) Player A will get the prize

Similar problem-1: Run of two batsmen of last 6 ODI cricket matches is given below:

Batsman A	65	68	25	78	85	51
Batsman B	40	100	35	18	28	25

(i) Which batsman has higher average run?
(ii) Which batsman is more consistent?

Similar problem-2: CGPA of two sections of students, each section containing 8 students, which are given below:

CGPA of Section A	3.90	4.00	3.75	3.50	2.90	3.45	3.80	3.95
CGPA of Section B	4.00	2.95	3.50	3.25	3.59	3.80	3.60	3.90

i) Which section students CGPA is more than other section?
ii) In which section students results is more consistent than other section?

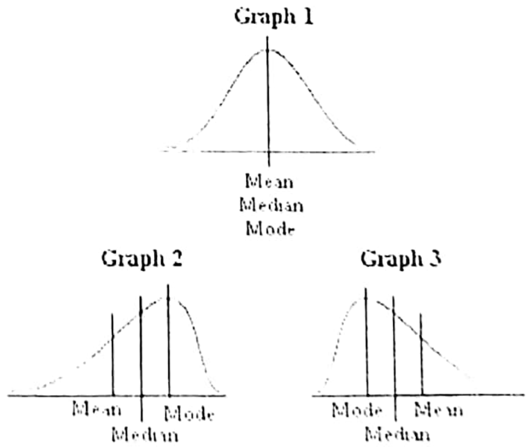
Similar problem-3: Daily wages of workers of two factories A& B are as follows:

Daily wages (Tk.)	90-99	100-109	110-119	120-129	130-139
No. of workers of factory A	6	13	20	6	5
No. of workers of factory B	4	8	29	4	5

i) Which factory has less variability wages?
ii) Find combined mean and standard deviation of wages of the factories.

Skewness: Skewness means "lack of symmetry" i.e. departed from symmetry of a distribution.
A distribution is said to be symmetric if –

a) If mean, median and mode give different values.
b) If Q_1 and Q_2 are not equidistance from the median.



Skewness

Skewness are two types –

i) **Positive skewness:** If the curve has a long tail in the positive side of the graph (Graph 3, here, $\bar{x} > M_e > M_o$).
ii) **Negative skewness:** If the curve has a long tail in the negative side of the graph (Graph 2, here $\bar{x} < M_e < M_o$).

Coefficient of skewness:

Co-efficient of skewness is denoted by β_1 where, $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$, $\gamma_1 = \sqrt{\beta_1}$.

- If $\beta_1 = 0$ and $\gamma_1 = 0$, the distribution is symmetric.
- If $\beta_1 > 0$ and $\gamma_1 > 0$, the distribution is positive skewed.

- If $\beta_1 < 0$ and $\gamma_1 < 0$, the distribution is negative skewed.

Kurtosis:

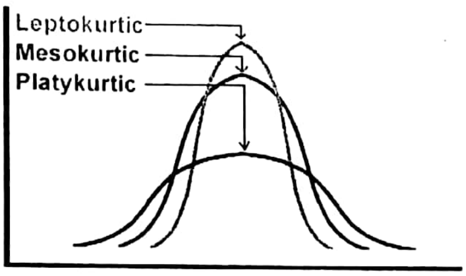
The degree of peakness or flatness of a distribution relative to a symmetrical distribution is called kurtosis.

Types of kurtosis

a) Leptokurtic: If the curve is more peaked than the normal curve.

b) Platykurtic: If the curve is more flat toped than the normal curve.

c) Mesokurtic: The normal curve itself.



Kurtosis

Coefficient of kurtosis:

Co-efficient of kurtosis is denoted by β_2 and $\beta_2 = \frac{\mu_4}{\mu_2^2}$, $\gamma_2 = \beta_2 - 3$

- If $\beta_2 = 3$ the distribution is mesokurtic.
- If $\beta_2 > 3$ the distribution is leptokurtic.
- If $\beta_2 < 3$ the distribution is platykurtic.

Examples:

No.	Coefficients	Comments
1.	$\beta_1 = 0$ & $\beta_2 = 1.5$	Symmetric and platykurtic curve.
2.	$\beta_1 = 1$ & $\beta_2 = 3$	Positive skewed and mesokurtic curve.
3.	$\beta_1 = -1.5$ & $\beta_2 = 4.5$	Negative skewed and leptokurtic curve.
4.	$\beta_1 = 0.40$ & $\beta_2 = 3$	Positive skewed and mesokurtic curve.
5.	$\beta_1 = -0.40$ & $\beta_2 = 2$	Negative skewed and platykurtic curve.
6.	$\beta_1 = 0$ & $\beta_2 = 3$	Symmetric and mesokurtic curve.

Previous semester question:

Spring-14(CSE):

1.(a) What do you mean by dispersion? What are the absolute measures of dispersion? Explain the superiority of coefficient of variation over standard deviation.

(b) Particulars regarding the weekly wages of workers of two factories A and B:

	Factory A	Factory B
Numbers of workers	600	500
Averages wages (in TK.)	1750	1850
81 Standard deviation (in TK.)	100	81

(i) Find coefficient of variation of the weekly wages of the two factories.
(ii) Which factory has better salary structure?

2.(a) Discuss skewness graphically.

(b) For the following values 5, 7, 9, 11, 13; Show that AM $\geq$GM $\geq$HM

(c) calculate range, mean deviation, standard deviation and C.V for the following samples:
4, 2, 1, 0, 1, 4, 2.3, 1.52

Autumn-12(CSE):

(a) Find mean and variance of first n natural numbers.

(b) Define Variance and coefficient of variation.

(c) Run of two batsmen of last four T-20 cricket matches is given below:

Batsman- X	35	28	25	48
Batsman-Y	40	10	35	18

(i) Which batsman has higher average run?
(ii) Which batsman is more consistent?

Define skewness and kurtosis. Find arithmetic mean and range of the following sets:
(i) 5,3,8,4,7,6,12,4,3.
(ii) 8.772, 6.453, 10.624, 8.628, 9.434, 6.351

Autumn-12(EEE):

1.(a) Define skewness and kurtosis. Explain skewness graphically.

(b)) i. Is primary source more reliable than secondary source?
ii. Is unstructured question are open-ended?
iii. Is Statistics science of averages?

(c) Find range of the following sets: (i) 7, 3, 8, 4, 7, 6, 12, 4, 5.
(ii) 6.772, 6.453, 10.624, 8.628, 9.434, 5.351

2. (a) Is variance the square of standard deviation? Define mean deviation and standard deviation.

(b) For n non-negative observations, show that, $CV \leq 100\sqrt{n-1}$

3.(c) CGPA of two sections of students, each section containing 8 students, which are given Below:

CGPA of Section 2A	3.90	4.00	3.75	3.50	2.90	3.45	3.80	3.95
CGPA of Section 2B	4.00	2.95	3.50	3.25	3.59	3.80	3.60	3.90

i) Which section students CGPA is more than other section?
ii) In which section students results is more consistent than other section?

Spring-12(ETE)

(a) What is dispersion?
(b) Indicate the different types of measures of dispersion. Discuss standard deviation and coefficient of variation.
(c) Find mean deviation and standard deviation of the following observations:
-2b, -b, 0, b, 2b

Autumn-11(BBA):
(a) What are the objectives of measuring measure of variation? Define standard deviation and mension its important properties.
(b) Daily wages of workers of two factories A& B are as follows:

Daily wages (Tk.)	90-99	100-109	110-119	120-129	130-139
No. of workers of factory A	6	13	20	6	5
No. of workers of factory B	4	8	29	4	5

i) Which factory has less variability wages?
ii) Find combined mean and standard deviation of wages of the factories.

Autumn-09(BBA):
Define average deviation.
(b)For the following values 5, 7, 9, 11, 13; Show that AM $\geq$GM $\geq$HM
(c) CGPA of two sections of students, each section containing 8 students, which are given Below:

CGPA of Section A	3.90	4.00	3.75	3.50	2.90	3.45	3.80	3.95
CGPA of Section B	4.00	2.95	3.50	3.25	3.59	3.80	3.60	3.90

i) Which section students CGPA is more than other section?
ii) In which section students results is more consistent than other section?

Autumn-10 (BBA):

(a) Distinguish between absolute and relative measures of dispersion. What are the properties of standard deviation?
(b) Run of two batsmen of last 6 ODI cricket matches is given below:

Batsman A	65	68	25	78	85	51
Batsman B	40	100	35	18	28	25

(i) Which batsman has higher average run?
(ii) Which batsman is more consistent?