

Test Yourself 1

The following data represents Annual wages of two Factories X and Y for the given information

- I. Determine range and coefficient of range. (in '000 Tk)
- II. Determine the quartile deviation and coefficient of Co-efficient of quartile deviation.

Table 1: Annual wages of Factory X workers (in '000 Tk)

91	70	74	79	86	93
60	71	76	79	87	96
112	72	127	79	87	62
68	72	77	79	90	76
69	73	77	85	48	157

Table 2: Annual wages of Factory Y workers (in '000 Tk)

97	78	85	92	97	105
72	79	85	92	97	107
112	79	87	92	97	72
113	80	90	96	68	75
78	82	90	97	100	

Hints:

1. Find the lowest and highest value for each table.
2. Calculate the quartiles (Q_1 , Q_2 , Q_3) for each table.

I. Determine range and coefficient of range. (in '000 Tk)

For table 1:

$$R = 157 - 48 = 109$$

$$\text{Coefficient of Range} = \frac{157 - 48}{157 + 48} = 0.5317$$

For table 2:

$$R = 113 - 68 = 45$$

$$\text{Coefficient of Range} = \frac{113 - 68}{113 + 68} = 0.2486$$

II. Determine the quartile deviation and coefficient of Co-efficient of quartile deviation.

For table 1:

First Quartile Q_1 :

$$\frac{in}{4} = \frac{1 \times 30}{4} = 7.5, \text{ not an integer}$$

$$Q_1 = 72$$

Second Quartile Q_2 :

$$\frac{in}{4} = \frac{2 \times 30}{4} = 15, \text{ an integer}$$

$$Q_2 = \frac{1}{2} [77 + 79] = 78$$

Third Quartile Q_3 :

$$\frac{in}{4} = \frac{3 \times 30}{4} = 22.5, \text{ not an integer}$$

$$Q_3 = 87$$

$$QD = \frac{Q_3 - Q_1}{2} = \frac{87 - 72}{2} = 7.5$$

$$\text{Coefficient of QD} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{87 - 72}{87 + 72} = 0.09434$$

For table 2:

First Quartile Q_1 :

$$\frac{in}{4} = \frac{1 \times 29}{4} = 7.25, \text{ not an integer}$$

$$Q_1 = 79$$

Second Quartile Q_2 :

$$\frac{in}{4} = \frac{2*29}{4} = 14.5, \text{ not an integer}$$

$$Q_2 = 90$$

Third Quartile Q_3 :

$$\frac{in}{4} = \frac{3*29}{4} = 21.75, \text{ not an integer}$$

$$Q_3 = 97$$

$$QD = \frac{Q_3 - Q_1}{2} = \frac{97 - 79}{2} = 9$$

$$\text{Coefficient of QD} = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{97 - 79}{97 + 79} = 0.10227$$

Test Yourself 2

A population of 10 students got the marks in the examination as given in the table below. Find the variance and Standard Deviation of the given data.

13 15 14 16 2 8 9 23 28 12

Answer:

x_i	x_i^2
13	169
15	225
14	196
16	256
2	4
8	64
9	81
23	529
28	784
12	144
$\sum x_i = 140$	$\sum x_i^2 = 2452$

$$\text{Variance, } \sigma^2 = \frac{1}{N} \left[\sum_{i=1}^N x_i^2 - \frac{(\sum_{i=1}^N x_i)^2}{N} \right] = \frac{1}{10} \left[2452 - \frac{(140)^2}{10} \right] = 49.2$$

$$\text{SD} = 7.01427$$

Test Yourself 3

A population of 40 students got marks in the examination as given in the table below. Find the variance and Standard Deviation of the given data.

X_i	15	20	25	30	35
f_i	6	8	15	7	4

x_i	f_i	$f_i * x_i^2$	$f_i * x_i$
15	6	1350	90
20	8	3200	160
25	15	9375	375
30	7	6300	210
35	4	4900	140
		$\sum f_i * x_i^2 = 25125$	$\sum f_i * x_i = 975$

$$\sigma^2 = \frac{1}{N} \left[\sum_{i=1}^N f_i x_i^2 - \frac{(\sum_{i=1}^N f_i x_i)^2}{N} \right] = \frac{1}{40} \left[25125 - \frac{(975)^2}{40} \right] = 33.9844$$

$$SD = 5.8296$$

Test Yourself 4

x_i	x_i^2
13	169
15	225
14	196
16	256
2	4
8	64
9	81
23	529
28	784
12	144
$\sum x_i = 140$	$\sum x_i^2 = 2452$

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^N x_i^2 - \frac{(\sum_{i=1}^N x_i)^2}{n} \right] = \frac{1}{9} \left[2452 - \frac{(140)^2}{10} \right] = 54.66667$$

$$SD = 7.39369$$

Test Yourself 5

A sample of 40 students got marks in the examination as given in the table below. Find the variance and Standard Deviation of the given data.

X_i	15	20	25	30	35
f_i	6	8	15	7	4

x_i	f_i	$f_i * x_i^2$	$f_i * x_i$
15	6	1350	90
20	8	3200	160
25	15	9375	375
30	7	6300	210
35	4	4900	140
		$\sum f_i * x_i^2 = 25125$	$\sum f_i * x_i = 975$

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^N f_i x_i^2 - \frac{(\sum_{i=1}^N f_i x_i)^2}{n} \right] = \frac{1}{39} \left[25125 - \frac{(975)^2}{40} \right] = 34.8558$$

$$SD = 5.9039$$

Test Yourself 6

An Advertising company is looking for a group of extras to shoot a sequence for a movie. The ages of the first 20 candidates to be interviewed are

50	56	44	49	52	57	56	57	56	59
54	55	61	60	51	59	62	52	54	49

x_i	x_i^2
50	2500
56	3136
44	1936
49	2401
52	2704
57	3249
56	3136
57	3249
56	3136
59	3481
54	2916
55	3025
61	3721
60	3600
51	2601
59	3481
62	3844
52	2704
54	2916
49	2401
$\sum x_i = 1093$	$\sum x_i^2 = 60137$

$$s^2 = \frac{1}{n-1} \left[\sum_{i=1}^N x_i^2 - \frac{(\sum_{i=1}^N x_i)^2}{n} \right] = \frac{1}{19} \left[60137 - \frac{(1093)^2}{20} \right] = 21.29211$$

$$SD = 4.6143$$

As the director suggested that a standard deviation of 3 years would be accepted.

So this group of extras will not qualify.

Test Yourself 7

In a University, students can take any number of courses per semester. For two samples of 30 students each. The data of how many courses one takes is given below:

Course Number	2	3	4	5	6
Sample 1	2	5	10	12	1
Sample 2	1	6	8	13	2

For which sample of students, the relative variability of course numbers is higher?

Sample 1:

x_i	f_i	$f_i * x_i^2$	$f_i * x_i$
2	2	8	4
3	5	45	15
4	10	160	40
5	12	300	60
6	1	36	6
		$\sum f_i * x_i^2 = 549$	$\sum f_i * x_i = 125$

$$s_1^2 = \frac{1}{n-1} \left[\sum_{i=1}^N f_i x_i^2 - \frac{(\sum_{i=1}^N f_i x_i)^2}{n-1} \right] = \frac{1}{29} \left[549 - \frac{(125)^2}{30} \right] = 0.971264$$

$$SD_1 = 0.985527$$

$$AM_1 = 4.1667$$

$$CV_1 = \frac{SD}{AM} = 0.2365$$

Sample 2:

x_i	f_i	$f_i * x_i^2$	$f_i * x_i$
2	1	4	2
3	6	54	18
4	8	128	32
5	13	325	65
6	2	72	12
		$\sum f_i * x_i^2 = 583$	$\sum f_i * x_i = 129$

$$s_2^2 = \frac{1}{n-1} \left[\sum_{i=1}^N f_i x_i^2 - \frac{(\sum_{i=1}^N f_i x_i)^2}{n-1} \right] = \frac{1}{29} \left[25125 - \frac{(975)^2}{30} \right] = 0.9759$$

$$SD_2 = 0.9879$$

$$AM_2 = 4.3$$

$$CV_2 = \frac{SD}{AM} = 0.2297$$

Here,

$$CV_1 > CV_2$$

So, the relative variability is higher in sample 1.

Test Yourself 8

From the analysis of monthly wages paid to employees in two service organizations X and Y, the following results were obtained:

	Organization X	Organization Y
Number of wage-earners	550	650
Average monthly wages	5000	4500
Variance of the distribution of wages	900	1600

- a. Which organization pays a larger amount as monthly wages?

Organization X pays: $550 * 5000 = 2750000$

Organization Y pays: $650 * 4500 = 2925000$

Organization Y pays larger amount as monthly wages.

- b. Determine the combined variance of all the employees taken together?

For organization X:

$$\mu_1 = 5000$$

$$\text{Variance, } \sigma_1^2 = 900$$

$$SD, \sigma_1 = 30$$

$$n_1 = 550$$

For organization Y:

$$\mu_2 = 4500$$

$$\text{Variance, } \sigma_2^2 = 1600$$

$$SD, \sigma_2 = 40$$

$$n_2 = 650$$

Now,

$$\mu_{12} = \frac{n_1\mu_1 + n_2\mu_2}{n_1 + n_2} = \frac{550*5000 + 650*4500}{550 + 650} = 4729.16667$$

$$d_1 = \mu_{12} - \mu_1 = -270.83333$$

$$d_2 = \mu_{12} - \mu_2 = 229.16667$$

$$\text{Combined Variance, } \sigma_{12}^2 = \frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2} = 63345.13144$$

Test Yourself 9

For a group of 50 male workers, the mean and standard deviation of their monthly wages are tk. 6300 and tk. 600 respectively. For a group of 40 female workers, these are tk. 5400 and tk. 600 respectively. Find the standard deviation of monthly wages for the combined group of workers.

For Group 1:

$$\mu_1 = 6300$$

$$SD, \sigma_1 = 600$$

$$n_1 = 50$$

For Group 2:

$$\mu_2 = 5400$$

$$SD, \sigma_2 = 600$$

$$n_2 = 40$$

Now,

$$\mu_{12} = \frac{n_1\mu_1 + n_2\mu_2}{n_1 + n_2} = 5900$$

$$d_1 = \mu_{12} - \mu_1 = -400$$

$$d_2 = \mu_{12} - \mu_2 = 500$$

$$\text{Combined Standard deviation, } \sigma_{12} = \sqrt{\frac{n_1(\sigma_1^2 + d_1^2) + n_2(\sigma_2^2 + d_2^2)}{n_1 + n_2}} = 748.3315$$

Test Yourself 10

1. If for a distribution Mean = 18, Median = 32, and Mode = 36, the distribution is negatively skewed. [Mean < Median < Mode]

2. If for a distribution Mean = 20, Median = 26.4, and SD = 3.3, the distribution is negatively skewed.

$$Sk_p = \frac{3(\text{Mean} - \text{Median})}{SD} = -5.818 < 0 \text{ [Negatively skewed]}$$

3. If for a distribution, Mean = 35.6, Mode = 24, and SD = 5.2, what is the skewness coefficient of the distribution?

$$Sk_p = \frac{\text{Mean} - \text{Mode}}{SD} = 2.2307 > 0. \text{ The distribution is positively skewed.}$$

Test Yourself 11

Construct a box plot for the data given below and hence comment on the skewness of the distribution:

99	75	84	33	45	66	97	69	55	61
72	91	74	93	54	76	62	91	77	68

Minimum value = 33

Maximum value = 99

Median = 73

First Quartile Q_1 :

$$\frac{in}{4} = \frac{1 \times 20}{4} = 5, \text{ an integer}$$

$$Q_1 = \frac{1}{2}[61 + 62] = 61.5$$

Second Quartile Q_2 :

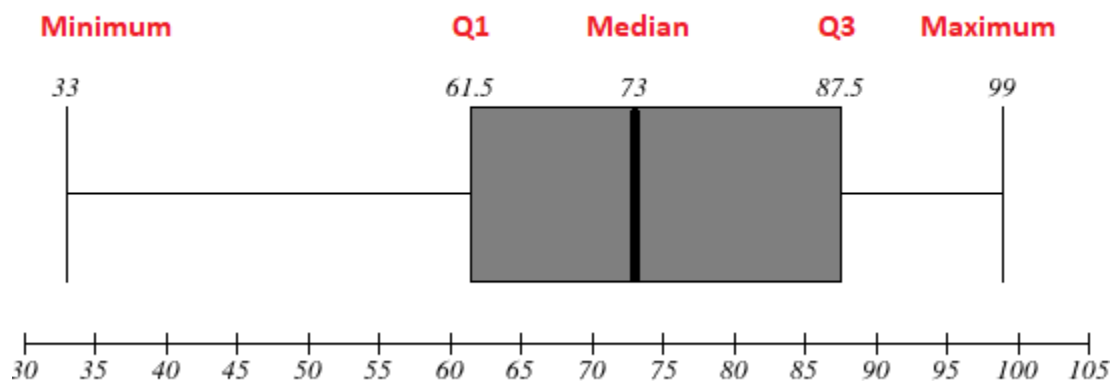
$$\frac{in}{4} = \frac{2 \times 20}{4} = 10, \text{ an integer}$$

$$Q_2 = \frac{1}{2}[72 + 74] = 73$$

Third Quartile Q_3 :

$$\frac{in}{4} = \frac{3 \times 20}{4} = 15, \text{ an integer}$$

$$Q_3 = \frac{1}{2}[84 + 91] = 87.5$$



Using Bowley's coefficient for Skewness we get

$$Sk_b = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = 0.1154 > 0, \text{ the distribution is positively skewed.}$$