1. Introduction

Welcome to the series of e-learning modules on absolute and relative measures of dispersion - part 2.

In part 1 of this module we have understood the calculation of some of the absolute and relative measures of dispersion like range, quartile deviation, mean deviation and its coefficients.
In part 2, let us focus our understanding in calculating the:
- Standard deviation and coefficient of standard deviation
- Combined standard deviation
- Coefficient of variation and
- Correcting the incorrect value

Let us solve some practical problems in this session:
Problem 1: Individual Series
Calculate standard deviation by direct and shortcut methods from the following series.
Wages are given as: 40,44,54,60,62,64,70,80,90 and 96.

Table 1

Wages(Rs)	40	44	54	60	62	64	70	80	90	96

We shall first solve this using the Direct Method.

Table 2

Wages(x)	d (x-X)	d2 =(x-X) 2
40	-26	676
44	-22	484
54	-12	144
60	-6	36
62	-4	16
64	-2	4
70	4	16
80	14	196
90	24	576
96	30	900
Σx=660	Σd=0	Σd 2= 3048

Here, for calculation of the standard deviation from the actual mean let us take the deviation from the mean and then we square the deviation.
The mean is equal to summation of x divided by N which is equal to 660 divided by 10 is equal to 66 So the deviation values are (minus 26), (minus 22), (minus 12), (minus 6), (minus

4), (minus 2), 4, 14, 24 and 30, and the 'summation of d' is 'zero'.
The square of the deviation are 676, 484,144, 36, 16, 4, 16, 196, 576 and 900 and the 'summation of d square' is equal to 3048.

Now we will calculate the standard deviation by using the formula:
Standard deviation is equal to square root of summation of d square divided by N,
which is equal to square root of 3048 divided by 10,
is equal to square root of 304.8 which is equal to 17.46.
The coefficient of standard deviation is equal to standard deviation divided by mean which is equal to 17.46 divided by 66 is equal to 0.2645.

The second method we shall see here is the Shortcut Method.
In this method, we calculate the deviation from the assumed mean.

Table 3

Wages(x)	d (x-A)	$d^2 =(x-A)^2$
40	-20	400
44	-16	256
54	-6	36
60	0	0
62	2	4
64	4	16
70	10	100
80	20	400
90	30	900
96	36	1296
Σx=660	Σd=60	Σd²= 3408

Here, let us assume 60 as the mean and calculate the deviation which is equal to (minus 20), (minus 16), (minus 6), 0, 2, 4, 10, 20, 30, 36 and 'the summation of dx' is equal to 60.
Let us calculate the mean using the formula 'mean' is equal to 'A' plus 'summation dx divided by N' is equal to 60 plus '60 divided by 10'.
So mean is equal to '60 plus 6' is equal to 66.
Next we square the deviation, d. Hence we will get the values of 'd square' as 400, 256, 36, 0, 4, 16, 100, 400, 900 and 1296 and the value of the 'summation of dx square' is equal to 3408

The standard deviation is equal to square root of 'summation of square of the deviation divided by N' minus 'summation deviation by N whole square',
which is equal to 'square root of 3408 divided by 10' minus '60 divided by 10 whole square'
which is equal to 'square root of 340.8 minus 36'
is equal to square root of 304.8
which is equal to 17.46.
The coefficient of standard deviation is equal to 'standard deviation divided by mean' is equal to '17.46 divided by 66' is equal to 0.2645.

Problem 2:- Discrete Series

Calculate standard deviation from the following data:
Marks, given as: 12, 14, 16, 18, 20, 22 and 24 and the corresponding number of students given as: 6, 12, 18, 26, 16, 10 and 8.

Table 4

Marks	12	14	16	18	20	22	24
Students	6	12	18	26	16	10	8

We shall try to solve this problem using the Shortcut method.

Table 5

Marks (x)	Students (f)	dx	fdx	dx2	fdx^2
12	6	-8	-48	64	382
14	12	-6	-72	36	432
16	18	-4	-72	16	288
18	26	-2	-52	4	104
20	16	0	0	0	0
22	10	2	20	4	40
24	8	4	32	16	128
Σx=660	N=96		Σfdx= -192		Σfdx^2= 1376

First, let us calculate the deviations from the assumed mean 20.
Then we will get the deviation values as (minus 8), (minus 6), (minus 4), (minus 2), 0, 2 and 4.
Next, we will calculate summation of fdx by multiplying the deviation with the frequency we will get (minus 48), (minus 72), (minus 72), (minus 52), 0, 20 and 32. Thus, summation fdx is equal to (minus 192).
Next we will calculate the square of the deviations and multiply it with the frequency to calculate summation of fdx square.
We get the values of fdx square as 382, 432, 288, 104, 0, 40 and 128, and the summation of fdx square is equal to 1376.
We now need to calculate the mean for the given data which is equal to 'A' plus 'summation fdx divided by N' is equal to '20' plus 'minus 192 divided by 96'
is equal to 20 minus 2 is equal to 18.

The calculation of standard deviation is equal to square root of 'summation fdx square divided by N' minus 'summation of fdx divided by N whole square',
is equal to 'square root of 1376 divided by 96' minus '192 divided by 96 whole square'
is equal to 'square root of 14.33 minus 4'
is equal to square of 10.33
is equal to 3.2145.

The calculation of coefficient of standard deviation is equal to standard deviation divided by mean is equal to 3.2145 divided by 18 is equal to 0.1786.

2. Continuous Series

Problem 3:- Continuous series.
Calculate the standard deviation and its coefficient from the following data:
Age groups defined are: 10 to 20, 20-30, 30-40, 40 to 50, 50 to 60 and 60 to 70.
The corresponding number of people in each group respectively is given as: 2,4,8,10,12 and 4.

Table 6

Age	10-20	20-30	30-40	40-50	50-60	60-70
Person	2	4	8	10	12	4

Solution:
In this problem let us understand the calculation of standard deviation through all the three methods - Direct, Shortcut and Step Deviation.
In the direct method, we calculate the deviation from the actual mean.

Table 7

Age (x)	Person (f)	Mid value	fx	d	d^2	fd^2
10-20	2	15	30	-29.5	870.25	1740.5
20-30	4	25	100	-19.5	380.25	1521.0
30-40	8	35	280	-9.5	90.25	722.0
40-50	10	45	450	0.5	0.25	2.5
50-60	12	55	660	10.5	110.25	1323.0
60-70	4	65	260	20.5	420.25	1681.0
	N=40		$\Sigma fx =$ 1780			$\Sigma fd^2=$ 6990.9

So, we will calculate the actual mean by first taking the mid values of the class interval and then multiplying it by the frequency and taking its total summation fx.
Here we will get the mid values as - 15, 25, 35, 45, 55, 65 and 75.
We get the values of fx as 30, 100, 280, 450, 660 and 260, and summation fx as 1780.
Thus, the actual mean of the data is equal to summation fx by N
Which is equal to 1780 divided by 40 is equal to 44.5.
Then next, we calculate the deviations from the mean which is equal to (minus 29.5), (minus 19.5), (minus 9.5), 0.5, 10.5 and 20.5.
In the next column we square the deviations to get 870.25, 380.25, 90.25, 0.25, 110.25 and 420.25.
Now we shall multiply the deviation square with the frequency to get 'fd square' which is equal to 1740.5, 1521.0, 722.0, 2.5, 1323.0 and 1681.9 which gives a total of 6990.9 which is the summation of fd square.

Thus the standard deviation is equal to 'square root of summation fd square divided by N' is equal to 'square root of 6990.9 divided by 40' is equal to 'square root of 174.75' is equal to '13.219'.

The coefficient of standard deviation is equal to 'standard deviation divided by mean' is equal to '13.219 divide by 44.5' is equal to 0.297.

In the short cut method we will calculate the deviations from the assumed mean.
In this problem we have assumed 35 as the mean and calculated the deviations as (minus 20), (minus10), 0, 10, 20 and 30.

Table 8

Age (x)	Person (f)	Mid value	dx	fdx	d2	fd2
10-20	2	15	-20	-40	400	800
20-30	4	25	-10	-40	100	400
30-40	8	35	0	0	0	0
40-50	10	45	10	100	100	1000
50-60	12	55	20	240	400	4800
60-70	4	65	30	120	900	3600
	N=40			Σfdx = 380		Σfdx 2= 10600

Next, we square the deviations to get the values as: 400, 100, 0 100, 400 and 900.
Next, we shall multiply the deviation and its square by the frequency to get the values of fdx as (minus 40), (minus 40), 0 100, 240 and 120, and the summation of fdx calculated as 380.
In the next column we calculate fdx square to get the values: 800, 400, 0, 1000, 4800 and 3000, and the summation fdx square as 10600.
Now we will calculate the mean.
'Mean is equal to 'A' plus ;summation fdx divided by n' is equal to '35' plus '380 divided by 40' is equal to '35 plus 9.5' is equal to '44.5'.

3. Standard Deviation and Coefficient of Standard Deviation

Now, to calculate the standard deviation and coefficient of standard deviation, we use the formula for standard deviation as:

'square root of summation fdx square divided by n' minus 'summation fdx divided by N whole square' is equal to 'square root of 10600 divided by 40' minus '380 divided by 40 whole square'

is equal to 'square root of 265' minus '9.5 whole square'

is equal to 'square root of 265' minus '90.25'

is equal to square root of 174.75

is equal to 13.219.

The coefficient of standard deviation is equal to standard deviation divided by mean

is equal to 13.219 divided by 44.5

is equal to 0.297.

In the step deviation method, the deviation derived from the assumed mean is further simplified by dividing it by the common factor 'c'

Table 9

Age (x)	Person (f)	Mid value	Dx'	Fdx'	d2	Fd2
10-20	2	15	-2	-4	4	8
20-30	4	25	-1	-4	1	4
30-40	8	35	0	0	0	0
40-50	10	45	1	10	1	10
50-60	12	55	2	24	4	48
60-70	4	65	3	12	9	36
	N=40			Σfdx = 38		Σfdx 2= 106

Hence, we get the deviation dx dash which is equal to (minus 2), (minus1), 0, 1, 2, 3.

These values are squared to get 'dx dash square' which is respectively equal to 4, 1, 0, 1, 4, 9.

We now multiply the frequency with the deviations and its squares to get 'fdx dash' and 'fdx dash square'.

So, 'fdx dash' is equal to (minus 4), (minus 4), 0, 10, 24 and 12 and 'summation fdx dash' is equal to 38.

When we calculate fdx dash square we get 8, 4, 0, 10, 48 and 36 and we get 'summation of fdx dash square' as equal to 106.

Mean is equal to 'A' plus 'summation fdx dash divided by n' into c,

is equal to '35' plus '38 divided by 40 into 10'

is equal to 35 plus 9.5

is equal to 44.5.

The standard deviation is calculated by taking the square root of 'summation fdx dash square divided by n; minus 'summation fdx dash divided by N whole square' into 'c'
is equal to square root of '106 divided by 40' minus '38 divided by 40 whole square' into 10
is equal to square root of '2.65 minus .9025' into 10
is equal to square root of '1.7475 into 10'
is equal to 13.219.
The coefficient of standard deviation is equal to standard deviation divided by mean
is equal to 13.219 divided by 44.5
is equal to 0.297.

Next, we shall find the value of the mean 2.
We will use the formula combined mean is equal to 'the product of the mean and number of the first group' plus the 'product of the mean and the number of the second group' plus the' product of the mean and number of the third group' divided by 'N1 plus N2 plus N3'.
As the combined mean is known, we substitute the values and get 116 is equal to '50 into 113' plus 60(mean 2) plus '90 into 115' divided by 200
is equal to '116 into 200' is equal to 5650 plus 60 mean 2 plus 10350.
Thus, 60 mean 2 is equal to '23200 minus 5650 minus 10350'
is equal to '23200 minus 16000'
is equal to 7200.
Therefore mean 2 is equal to '7200 divided by 60' is equal to 120.

Next, to calculate the standard deviation of group three we will find the deviations of the three groups.
d1 is equal to 'mean of the first group minus combined mean'
is equal to '113 minus 116'
is equal to minus 3.
d2 is equal to 'mean of group two minus combined mean'
is equal to '120 minus 116'
is equal to 4, and
d3 is equal to 'mean of group three minus the combined mean'
is equal to '115 minus 116'
is equal to minus 1.

Now, we will substitute the values in the combined standard deviation equation which says the combined standard deviation is equal to square root of the sum of the products of the numbers and standard deviation of each group and the product of the number and deviation of each group divided by N1 plus N2 plus N3.
By taking square on both sides we get:
the square of the combined mean and standard deviation is equal to the number into the sum of standard deviation and deviation of each group respectively.
Thus, we get '200 into 7.746 square' is equal to '50 into (36 plus 9)' Plus '60 into (49 plus 16)' plus '90 into (sd+1)',
Which is equal to '200 into 60.000516' is equal to '50 into 45' plus '60 into 65' plus '90 sd plus 90',
Hence, 12000 is equal 2250 plus 3900 plus 90 plus '90 sd square',
90sd square is equal to 12000 minus 6240 is equal to 5760,

Then, sd square is equal to 5760 divided by 90 is equal to 64.
Thus the standard deviation is equal to square root of 64 is equal to 8.
Thus the unknown values are N2=60, mean 2 = 120 and standard deviation 3 is equal to 8.

4. Coefficient of Variation

Coefficient of variation:

Problem 6:

The following table shows annual profits, in crores of rupees, of two investment companies A and B during the given 5 years. If someone plans to invest Rs. 50000 which of the two companies should be preferred and why?

The years given as 1993, 1994, 1995, 1996 and 1997.

The respective values for company A are given as: 70, 60, 70, 60 and 70 and the values for company B are given as: 60, 50, 60, 80 and 80.

Table 10

	1993	1994	1995	1996	1997
A	70	60	70	60	70
B	60	50	60	80	80

Solution:

Let us now calculate the average and deviation of both the companies.

coefficient of variation helps in understanding the consistency and variability of the company.

Table 11

Year	X	Y	x	y	x^2	y^2
1993	70	60	4	-6	16	36
1994	60	50	-6	-16	36	256
1995	70	60	4	-6	16	36
1996	60	80	-6	14	36	196
1997	70	80	4	14	16	196
Total	330	330	0	0	120	720

Let us consider company A as 'X' and company B as 'Y'.

First, the average is calculated by considering the total of 'company A', which is equal to 'summation X by N'

is equal to 330 by 5 is equal to 66.

The average of 'company B' is equal to 'summation Y divided by n'

is equal to 330 by 5 is equal to 66.

As the average of both the companies is the same we shall now calculate the standard deviation for which we take the deviation of company A as small x and deviation of company B as small y.

We get the values of small x as: 4, (minus 6), 4, (minus 6) and 4.

So, 'summation small x is equal to 0'

The values of small y are calculated to get: (minus 6), (minus 16), (minus 6), 14 and 14.

So summation small y is also equal to 0.

Let us square the deviations to get X square and Y square.

X square is equal to 16, 36, 16,36,16, and so the summation x square is equal to 120,

y square is equal to 36, 256, 36, 196, 196 and so the summation y square is equal to 720.

Standard deviation of company A is equal to square root of summation x square by N

is equal to square root of 120 divided by 5

is equal to square root of 24

is equal to 4.08, and,

standard deviation of company B is equal to square root of summation y square by N

is equal to square root of 720 divided by 5,

is equal to square root of 144

is equal to 12.

Now, for deciding on the investment, we need to calculate the coefficient of variation.

Coefficient of variation of company A is equal to 'the standard deviation divided by the mean of company A'

is equal to 4.08 by 66

is equal to 0.00618, and,

the coefficient of variation of company B is equal to ;the standard deviation by mean of company B'

is equal to 12 divided by 66

is equal to 0.1818.

Thus, investment in company A will give consistent profit as the variation of Company B is more.

5. Correcting Incorrect Values in Standard Deviation

Problem 7:
The mean and the standard deviation calculated from 20 items are 15 and 10 respectively. If an additional observation 36 is left out throughout the calculation, what would be the correct mean and standard deviation?

Solution:
We know that:
Mean is equal to 'summation of x divided by N' and summation X is equal to 'N into mean'
is equal to 20 into 15
is equal to 300
But, this is an incorrect summation of X as the value 36 is not considered.
So, to get the correct summation X we add the value 36 to 300 and get 336 and the items become 21.
So, the correct mean is equal to 336 divided by 21 is equal to 16.
Next we will calculate the correct value of the standard deviation.
We know that 'standard deviation square' is equal to 'summation X square divided by N' minus 'mean square'
which means 10 square is equal to 'summation x square divided by 20' minus '15 square'
which gives 100 is equal to' summation X square divided by 20' minus 225
which gives 2000 is equal to 'summation of x square' minus 4500,
summation x square is equal to 2000 plus 4500
is equal to 6500,
which is the incorrect summation X square.
So, to get the correct value we will calculate by adding the missed value.
Thus correct summation x square is equal to '6500 plus 36 square' is equal to 7796.
Now we will calculate the correct standard deviation using the formula:
standard deviation square is equal to correct summation x square divided by N minus the correct mean
is equal to 7796 divided by 21 minus 16 square
is equal to 7796 minus 5376 divided by 21 is equal to 2420 divided by 21.
By removing the square we will get correct standard deviation which is equal to the square root of 2420 divided by 21
is equal to square root of 115.24
is equal to 10.73.
Thus we find the correct mean as 16 and the correct standard deviation as 10.73 by adding the eliminated observation 36 in the calculation.

Here's a summary of our learning in this session:
In this session we acquired an understanding of the calculation of :
- Standard deviation and coefficient of standard deviation
- Combined standard deviation
- Coefficient of variation, and
- Correcting the incorrect value