Chapter 5: Measures of Dispersions

Part-2

Learning Outcomes

After Completing the chapter ,you will able to :

- Necessity of measures of Dispersion.
- What is measures of Dispersion?
  and it’s purpose.
- Different types of measures of Dispersion and their application.
- Their uses and Limitations.

Contents

From this lecture, you are going to learn…

- Discussion on sample variance and sample standard deviation with maths.
- Coefficient of Variation (C.V.)
- Exercises

Sample Variance and Standard Deviation

Sample variance, $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$

$$= \frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \cdots \ldots + (X_n-\bar{X})^2}{n-1}$$

$\therefore$ **Sample Standard deviation,**

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

$$= \sqrt{\frac{(X_1-\bar{X})^2 + (X_2-\bar{X})^2 + \cdots \ldots + (X_n-\bar{X})^2}{n-1}}$$

Where $X_1,\ X_2, \ldots..\ X_N$ are Population observation
N= Population size.
$\bar{X}$= Sample Mean.

Why Sample standard deviation?
When entire information of the population will not available and need to draw conclusion based on sample we use sample standard deviation .

Example of Sample Variance and Standard Deviation

Example: Calculate the Sample variance and standard deviation for the data set 1, 2, 2, 3, 4, 5.

Solution:

Sample variance, $S^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$

Where, Sample Mean, $\bar{X}=\frac{\sum_{i=1}^{n} X_i}{n}$

$=\frac{1+2+2+3+4+5}{6}$

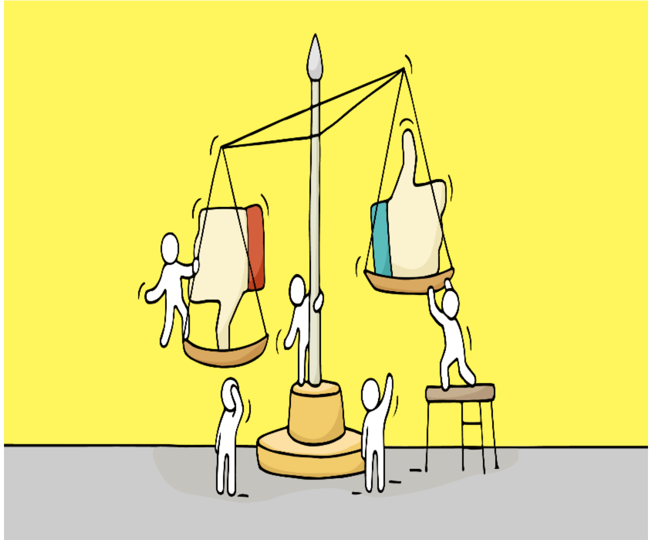
$= 2.83$

$\therefore\ S^2=\frac{(1-2.83)^2+(2-2.83)^2+\cdots\ldots+(5-2.83)^2}{6-1}$

$=\frac{10.84}{6-1}= 2.87$

$\therefore$Sample standard deviation, $\sigma =\sqrt{2.87}=1.69$

Merits and limitations of Standard Deviation

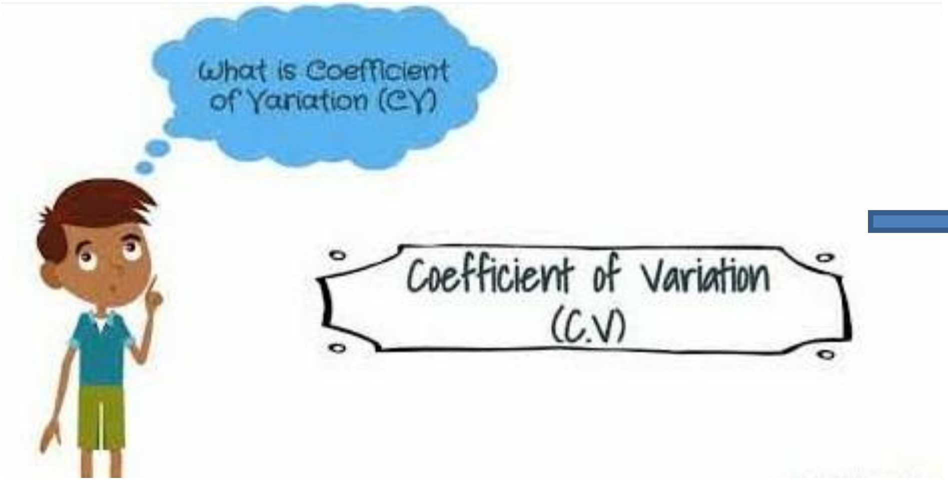


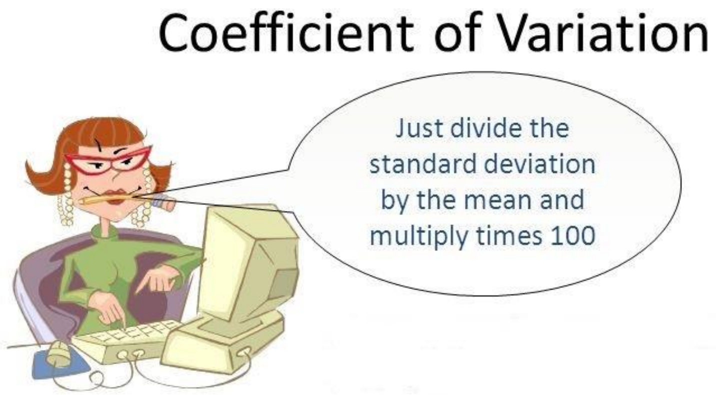
Merits of Standard Deviation:
Among all measures of dispersion Standard Deviation is considered superior because it possesses almost all the requisite characteristics of a good measure of dispersion.

1) It is based on all the observations of the data set.
3) It is amenable to further mathematical calculation.

Limitations:
1) It is more affected by extreme items.

Relative measure: Coefficient of Variation(c.v.)





Coefficient of Variation expresses the standard deviation as a percentage of the mean.

Relative measure: Coefficient of Variation(c.v.)

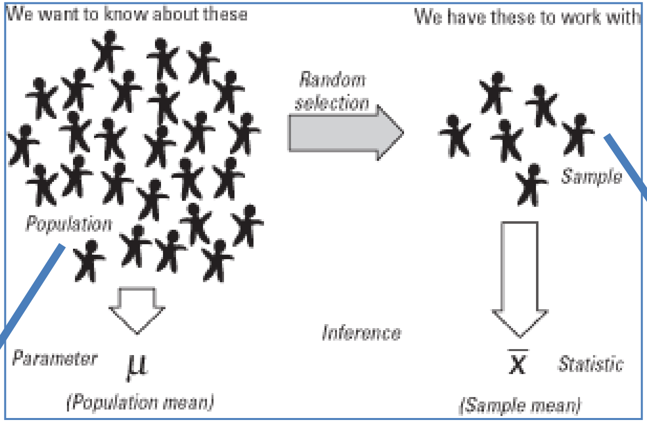
Purposes of studying coefficient of variation(C.V):

- To compare the variability of two or more data sets.
- To compare the Consistency, Uniformity, Homogeneity, Stability of two or more data sets.
- Since c.v. is a unit free measure, we can compare two or more data sets having different units.

High c.v. → High variability

Low c.v. → High consistency, High uniformity, High homogeneity, High stability

Relative measure: Coefficient of Variation(c.v.)



The Coefficient of Variation from population is,

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

Here, σ = population standard deviation

And μ = Population mean

The Coefficient of Variation from sample is,

C.V. = $\frac{S}{\bar{X}} \times 100$

Here, S = Sample standard deviation

And $\bar{X}$ = Population mean

Relative measure: Coefficient of Variation(c.v.)

Example: The mean final exam marks of Section A is 30 out of 40 with standard deviation 4 and the mean final exam marks of Section B is 25 out of 40 with standard deviation 6.
Which Section is more consistent in getting final exam marks?

Solution:

For the students of section A,

Here,

$$Mean = 30$$

$$Standard\ Deviation = 4$$

$$Coefficient\ of\ Variation = \frac{\sigma}{\mu} \times 100$$

$$Coefficient\ of\ Variation = \frac{4}{30} \times 100$$

$$Coefficient\ of\ Variation = 13.33\%$$

For the students of section B,

Here,

$$Mean = 25$$

$$Standard\ Deviation = 6$$

$$Coefficient\ of\ Variation = \frac{\sigma}{\mu} \times 100$$

$$Coefficient\ of\ Variation = \frac{6}{25} \times 100$$

$$Coefficient\ of\ Variation = 24\%$$

Since c.v. of section A is less than c.v. of section B, so section A is more consistent in getting their final exam marks.

Example

Example: Comments on Children in a community

	Height	weight
Mean	40 inch	10 kg
SD	5 inch	2 kg
CV	12.5	20

Sol:
Since the coefficient of variation for weight is greater than that of height, we would tend to conclude that weight has more variability than height in the population.

For better understanding of c.v. you can watch this link:

- https://www.youtube.com/watch?v=gUkeVA6iD8s

Exercises to solve

- The thickness of printed circuit board is very important characteristics. A sample of 10 boards had the following thickness (in thousands of an inch) :

  61 65 62 61 64 60 66 68
  61 63

Find the mean, standard deviation and coefficient of variation for the given data.

Lives of two models of computers in a recent survey were found as follows:

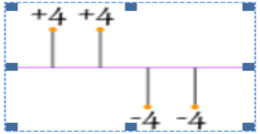
Life (average years)	1	3	5	7	9
Model A	5	16	13	7	5
Model B	2	7	12	19	9

i. What is the average life of each model of these computers?
ii. Which of the two models shows more uniformity?

Footnote: why standard deviation is better compare with mean deviation?

If we just add up the differences from the mean … the negatives cancel the positives: suppose we have some data 4, 4,-4,-4

The mean is



$4 + 4 - 4 - 4 = 0$

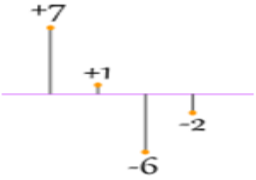
By using mean deviation, the value is

So that won't work. How about we use absolute values?

$|4| + |4| + |-4| + |-4| = 4 + 4 + 4 + 4 = 4$

That looks good (and is the Mean Deviation), but what about this case:

Suppose we have another pair of data: 7,1,-6,-2.



$|7| + |1| + |-6| + |-2| = 7 + 1 + 6 + 2 = 4$

Oh No! It also gives a value of 4, Even though the differences are more spread out.


Thank you