efficient / consistent =>

Deviation :

$$S^2 = \frac{\Sigma (x - \bar{x})^2}{n-1}$$

" $X_1 \Rightarrow X_2$; low measure of depression.

16 Jan' 2020

The Variance & Standard Deviation :

→ variance → "change" → variation, not exact. Example : refilling of bottles (250ml)
↳ "deviation of observed from mean data"
→ Error ↑ Estimate (not realiable)
→ The Answer of variance and take "square - root" of it $\sqrt{\sigma^2} = \sqrt{-4}$
→ constant values ⇒ zero variation e.g., 5, 5, 5, 5 $\sigma = 2i$

⇒ The variance of a set of observation is defined as :
"The mean of the squares of deviations of all the observations from their mean is called variance"
when it is calculated from the entire population, the variance is called '**population variance**' traditionally denoted by "σ^2"
If instead the data from the sample are used to calculate the variance it is refered to as '**sample variance**' & is denoted by S^2. In order to distinguish between the two the symbolic definition for variance is

$$\boxed{\sigma^2 = \frac{\Sigma (x_i - \mu)^2}{N}} \rightarrow \text{for population}$$

Scanned with CamScanner

$$S^2 = \frac{\Sigma (x_i - \bar{x})^2}{n} \Rightarrow \text{for sample}$$

⇒ The variance is also denoted by '<u>Var(X)</u>'.

⇒ The term variance was first introduced by '<u>R.A. Fisher</u>'

Standard Deviation:

"The positive square root of the variance is called standard deviation."

Symbolically written as;

$$\sigma = \sqrt{\frac{\Sigma (x_i - \mu)^2}{N}} \Rightarrow \text{for population}$$

$$S = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n}} \Rightarrow \text{for sample}$$

x_i observa

$\bar{x}$ mean

⇒ The standard deviation is expressed in the same units as the observations themselves and is a measure of the average spread about the mean.

The concept of standard deviation was first given by '<u>Carl Pearson</u>' who is called the "founder of the science of statistics"

Question 8

A population of $N = 10$ has the observations

x_i	$x_i - \mu$	$\lvert x_i - \mu \rvert$	$(x_i - \mu)^2$
7	$7-17 = -10$	10	100
8	$8-17 = -9$	9	81
10	$10-17 = -7$	7	49
13	$13-17 = -4$	4	16
14	$14-17 = -3$	3	9
19	$19-17 = 2$	2	4
20	$20-17 = 3$	3	9
25	$25-17 = 8$	8	64
26	$26-17 = 9$	9	81
28	$28-17 = 11$	11	121
$\Sigma x_i = 170$			$\Sigma (x_i - \mu)^2 = 534$

find its variance & standard deviation.

$$\sigma^2 = \frac{\Sigma (x_i - \mu)^2}{N}$$

$$\sigma^2 = \frac{534}{10}$$

$$\sigma^2 = 53.4$$

Scanned with CamScanner

$\therefore \sqrt{\sigma^2} = \sqrt{53.4}$

$\sigma = 7.30$

Calculate the variance and standard deviation from the sample mark obtained by 9 students are :
45 , 32 , 37 , 46 , 39 , 36 , 41 , 48 , 36

x_i	$(x_i - \bar{x})$	$\lvert x_i - \bar{x} \rvert$	$(x_i - \bar{x})^2$
45	45 - 40 = 5	5	25
32	32 - 40 = -8	8	64
37	37 - 40 = -3	3	9
46	46 - 40 = 6	6	36
39	39 - 40 = -1	1	1
36	36 - 40 = -4	4	16
41	41 - 40 = -1	1	1
48	48 - 40 = 8	8	64
36	36 - 40 = -4	4	16
$\Sigma = 360$			$\Sigma(x_i - \bar{x})^2 = 232$

$x = \frac{\Sigma x}{N}$

$\frac{360}{9} = 40$

$S^2 = \frac{\Sigma (x_i - \bar{x})^2}{n}$

$S^2 = \frac{232}{9}$

$S^2 = 25.7$

$\sqrt{25.7} = \boxed{5}$ Ans.

Scanned with CamScanner

variation b/w one series ⇒ variation
variation b/w ~~two~~ or more series = CV

⇒ small variability is prefe[rred] because high variability mean[s] high variation causes wrong res[ults]

Co-efficient of Variation:

The variability of two or more Than two sets of data can not be compared unless we have a relative measure. For this purpose 'Carl Pearson' introduced a relative measure of variation known as the coefficient of variation abbreviated at 'CV' which express the standard deviation as a percentage of the arithmetic mean of a data set.

Symbolically it is defined as:

$$C.V = \left(\frac{S}{\bar{X}}\right) 100 \Rightarrow \text{for sample}$$

$$C.V = \frac{\sigma}{\mu} \times 100 \Rightarrow \text{for population.}$$

As the coefficient of variation (CV) is a pure number without units, it is therefore used to compare the variation in two or more data sets or distributions that are measured in different units. A large value of CV indicates that the variability is great and a small value of CV indicates that there is less variability.

The co-efficient of variation is also used to compare the performance of two candidates or two players given their scores in various papers or games. The smaller the co-efficient of variation, the more consistent is the performance of the candidates.

$$S^2 = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N} \quad \text{(OR)} \quad S^2 = \frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N}\right)^2$$

It should be noted that this co-efficient is unrealistic when arithmetic mean is very small.

Question:

Using coefficient of variation determine whether or not there is greater variation among the prices of certain similar co. Than among the life in hours under test

X	X^2	Y	Y^2
8	64	130	16900
13	169	150	22500
18	324	180	32400
23	529	250	62500
30	900	345	119025
$\Sigma X = 92$	$\Sigma X^2 = 1986$	$\Sigma Y = 1055$	$\Sigma Y^2 = 253325$

$\bar{X} = \frac{92}{5} = \text{Rs } 18.4$

$S_x = \sqrt{\frac{1986}{5} - \left(\frac{92}{5}\right)^2}$

$S_x = \sqrt{58.44}$

$S_x = 7.66 \text{ Rs}$

So

$CV = \frac{7.66}{18.4} \times 100$

$CV = 41.63\%$

$\bar{Y} = \frac{1055}{5} = 211 \text{ hours}$

$S_y = \sqrt{\frac{253325}{5} - \left(\frac{1055}{5}\right)^2}$

$S_y = \sqrt{6144} = 78.38 \text{ hours.}$

so

$CV = \frac{78.38}{211} \times 100$

$CV = 37.15\%$

→ the constant of mean is always constant.
→ variance can never be zero.

23 Jan' 2020

Properties of Variance and Standard Deviation:

The variance and standard deviation have the following useful and interesting properties:

1. The variance of a constant is equal to zero. If 'a' is any constant than;

$$Var(a) = \frac{1}{N} \Sigma [a - a]^2$$

2. The variance is independent of the origin i.e, it remains unchanged when a constant is added or subtracted from each observation of the variable. Symbolically, we can write it as;

$$Var(X \pm a) = Var(X)$$

where 'a' is any arbitrary constant.

3. The variance is multiplied or divided by the square of the constant when each observation of the variable X is either multiplied or divided by a constant. Symbolically, we can written it as;

$$Var(aX) = \frac{1}{N} \Sigma (aX_i - a\mu)^2$$

... is applied to constant without any random
when variance is applied to constant along with random variable then
is applied only to random variable and constant become square

$$Var(cX) = c^2 \, Var(x) \quad ; \quad Var(X+c) = Var(X)$$

$$Var(cx+b) = c^2 \, Var(x)$$

$$Var(aX) = \frac{1}{N} \Sigma (aX_i - a\mu)^2$$

$$= a^2 \frac{1}{N} \Sigma (X_i - \mu)^2$$

$$Var(aX) = a^2 \, Var(x)$$

4 The variance of the sum or difference of two independant variables is equal to the sum of their respective variance. Symbolically it can be written as;

$$Var(X \pm Y) = Var(X) + Var(Y)$$

It is relevant to note that all these properties are valid for standard deviation which is the positive square root of the variance. In other words,

(i) $S.D(a) = 0$

(ii) $SD(X+a) = SD(X)$

(iii) $SD(aX) = |a| \, SD(X)$

(iv) $SD(X \pm Y) = \sqrt{Var(X) + Var(Y)}$