II) Measure of Dispersion

two imp. topics to understand in measure of dispers-ion :-

1. Varience
2. Standard Deviation.

2.1) Variance :-

Population Variance

$$\boxed{\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - \mu)^2}{N}}$$

, x_i = Data points
μ = Population mean
N = Population size

Sample Variance

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

WHY DOES THE SAMPLE VARIANCE HAVE N-1 IN THE DENOMINATOR? The reason we use n-1 rather than n is so that the sample variance will be what is called an unbiased estimator of the population variance

Proved from: "Bessel Correction"

, x_i → Data points
$\bar{x}$ → Sample mean
n → Sample size

Ex 1: Sample :- {1, 2, 3, 4, 5}

Soln:- using $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

x_i	$\bar{x}$	$(x_i - \bar{x})^2$
1	3	4
2	3	1
3	3	0
4	3	1
5	3	4

$\sum_{i=1}^{n} (x_i - \bar{x})^2 = 10$ $\Rightarrow$ $s^2 = \frac{10}{4} = \frac{5}{2} = 2.5$

Q1) What does Varience Value Signifies?

Solⁿ- $s^2 = 2.5$ $\qquad\qquad$ $s^2 = 6.5$

So; Varience specifies the spread of the distribution as varience value ↑ed spread also ↑es.

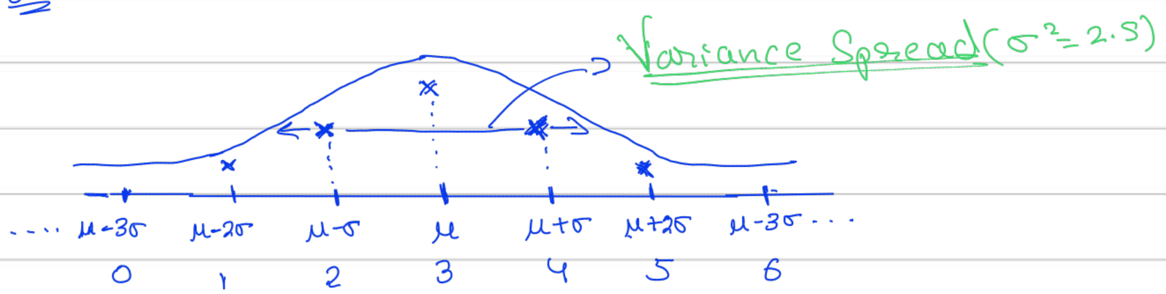
2.2) Standard Deviation:-

Population Std $\qquad\qquad$ **Sample Std.**

$\boxed{\sigma = \sqrt{\text{varience}}}$ $\qquad\qquad$ $\boxed{s = \sqrt{\text{Sample Varience}}}$

Ex: Suppose for a given data set : $\{1, 2, 3, 4, 5\}$, $\mu = 3$ and $\sigma = 1$ are the mean & std. deviation. Then Explain the std. deviation.

Solⁿ:-



Most of the element in a dataset are always present b/w 1st 3 std. deviations ($\mu-\sigma, \mu, \mu+\sigma$).