Measure of Dispersion.

1) Variance :

population variance

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

$x_i \rightarrow$ Data points

$\mu \rightarrow$ Population mean

$N \rightarrow$ population size

sample variance

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

Why sample variance has $n-1$ in denominator?

To create an unbiased estimator of the population variance.

Bessel correction

$x_i \rightarrow$ Data pts

$\bar{x} \rightarrow$ sample mean

$n \rightarrow$ sample size

eg $\rightarrow$ $\{1, 2, 3, 4, 5\}$ $\Rightarrow$ sample

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

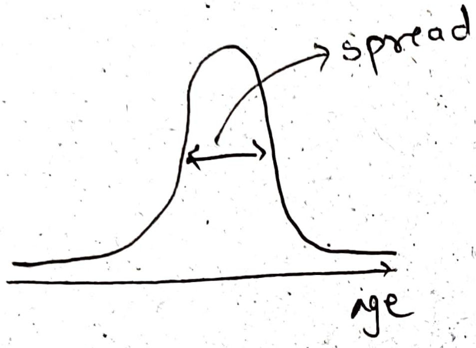
x_i	$\bar{x}$	$(x_i - \bar{x})^2$
1	3	4
2	3	1
3	3	0
4	3	1
5	3	4

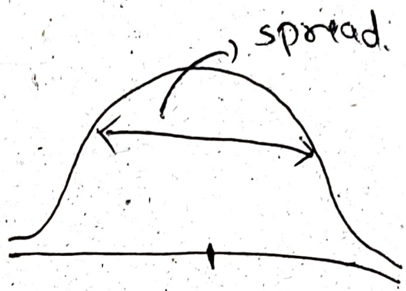
$\Sigma (x_i - \bar{x})^2 = 10$

Variance : spread of the data

$$\boxed{s^2 = 2.5}$$

$s^2 = 2.5$ $s^2 = 6.5$





Larger variance → Larger spread.

2) Standard deviation :

population std dev

$\sigma = \sqrt{\text{variance}}$

sample std dev.

$S = \sqrt{\text{Sample var}}$

$s^2 = 2.5$

$\sqrt{s^2}$ = sample std

consider

$\mu = 3$ $\sigma = 1$