Measures of Dispersion:

① Variance

② Standard deviation

$Ages1 = \{2,2,4,4\}$ $\qquad$ $Ages2 = \{1,1,5,5\}$

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

They both have same mean
but in Ages2 5 is more far from 1 as compared to distance between 2 & 4

Ages2 is more spread

N(population) $\qquad$ n (Sample)

① Variance $\qquad$ $\sigma^2 = \sum_{i=1}^{N} \frac{(x_i - \mu)^2}{N}$ $\qquad$ $s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$

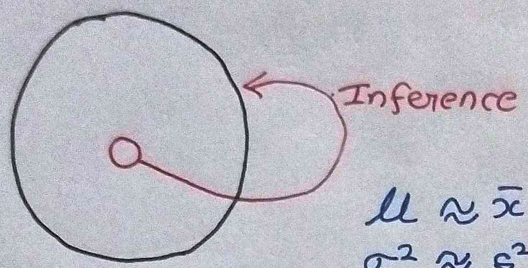
We are finding distance from mean

$$\sigma^2_{Age1} = \frac{(3-2)^2 + (3-2)^2 + (3-4)^2 + (3-4)^2}{4} = 1$$

$$\sigma^2_{Ages2} = \frac{(3-1)^2 + (3-1)^2 + (3-5)^2 + (3-5)^2}{4} = 4$$

for Age_2 data set elements are far from mean
hence $\sigma^2_{Ages_2} = 4$
When compared to $\sigma^2_{Ages1} = 1$

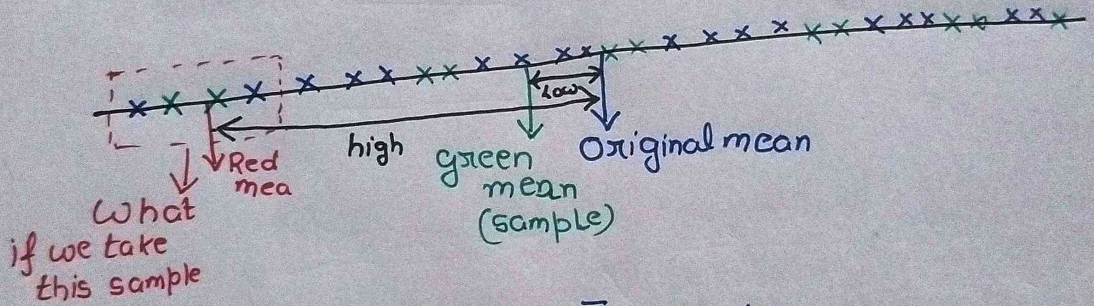
Why sample variance divided by (n-1)



Say we have
Ages = {...............}
& we want

$\mu \approx \bar{x}$
$\sigma^2 \approx s^2$

Ages →



Now for red as $\bar{x} <<< \mu$
so $s^2 <<< \sigma^2$

So to reduce the differance we put (n-1) in formula instead of n → Bessel's correction

DoF → Degree of Freedom

Standard deviation

N (population)

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

It signifies on avg. how much data point is away from mean

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

n (sample)

$$S = \sqrt{S^2}$$

$$S = \sqrt{\text{sample variance}}$$