

Sample variance why  $n-1$ ?

The relation we use ' $n-1$ ' rather than ' $n$ ' is so that the sample variance will be what is called as 'unbiased estimation population variance ( $\sigma^2$ )'.

The variance estimator makes use of sample mean & as a consequence under-estimate the true variance of the population. Dividing by  $(n-1)$  instead of  $(n)$  corrects for that bias. Further more, dividing by  $(n-1)$  makes the variance of a one-element sample undefined rather than zero.

i.e.  $(n-1)$  = unbiased sample Estimates

In data processing, degree of freedom is the no. of independent data but always there is one dependent data which can be obtained from data. As the observer takes the values  $(n-2)$ ,  $(n-3)$ ,  $(n-4)$  or so on. finally it evaluates the degree of freedom =  $(n-1)$

We know the population data has the more data as compare to sample data.

consider  $\bar{x}$  = sample mean,  $\mu$  = population mean

$n$  = sample data

$N$  = population data

For the value  $s^2 \approx \sigma^2$  &  $\bar{x} \approx \mu$

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

If we take only  $(n)$  instead of  $(n-1)$  then  $\bar{x} \ll \mu$   
&  $s^2 \ll \sigma^2$

When we take  $n-1$  the the the value are approximately near. i.e.  $s^2 \approx \sigma^2$ .

This is also called as 'Bessel correction'.