25 05 2022

vimp

- **Assignment-1) Sample variance (n-1) ?**

Ans ⇒ The Reason we use 'n-1' rather than 'n' is so that the sample variance will be what is called an "unbiased Estimation of population variance σ^2."

Here, "Estimator" means - It is a random variable whose underlying normal random process is choosing a sample & whose value is statistics, based on that sample, that is used to estimate a population parameter.

Also, 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. Furthermore, dividing by (n-1) makes the variance of a one-element sample undefined rather than zero.

i.e. (n-1) = Unbiased Sample Estimates.

And, In data processing, degree of freedom is the no. of independent data but always there is one dependent data which can be obtain from other data. ∴ D.O.F. = (n-1)

Consider, ⇒ $\bar{x}$ = Sample Mean; μ = population mean; σ^2 = Population variance & n = sample Data; N = population Data; s^2 = sample variance

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

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

If we take denominator = 'n'
Then, $\bar{x} <<< \mu$ & $s^2 <<< \sigma^2$

i.e. we under-estimating True population variance.

∴ when we take denominator as "(n-1)" so gap is reduced & this comes with unbiased estimation i.e. without under-estimating; we moving on with diff. sample data.

∴ ⇒ This Entire Scenario is also called as 'Bakel's Correction' or "D.O.F."