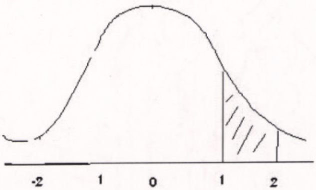

III. Estimating the Population Variance from the Sample Variance

We will show below that the sample variance is a biased estimator of the population variance and then that the correction for degrees of freedom, which transform the sample variance to s^2 yeilds an unbiased estimator of the population variance.

First we develop a useful identity.

$\sum(X_i-\mu)^2$ can be rewritten by adding and subtracting $\bar{X}$ to the term in parenthesis

$$\begin{aligned}\sum(X_i-\mu)^2 &= \sum(X_i-\bar{X}+\bar{X}-\mu)^2\\ &= \sum[(X_i-\bar{X})+(\bar{X}-\mu)]^2\\ &= \sum[(X_i-\bar{X})^2+2(X_i-\bar{X})(\bar{X}-\mu)+(\bar{X}-\mu)^2]\\ &= \sum[(X_i-\bar{X})^2+2(X_i-\bar{X})(\bar{X}-\mu)]+n(\bar{X}-\mu)^2\end{aligned}$$

Since $\bar{X}$ and μ are constant and can be factored out of the $\sum$ and since $\sum$ of a constant=n * constant. Then note $\sum(X_i-\bar{X}) = \sum X_i - n\bar{X} = n\bar{X}-n\bar{X}=0$

so the second term vanishes.

$$\sum(X-\mu)^2 = \sum(X_i-\bar{X})^2+2(\bar{X}-\mu)\sum(X_i-\bar{X})+n(\bar{X}-\mu)^2 = \sum(X_i-\bar{X})^2+n(\bar{X}-\mu)^2$$

We can also rearrange with equation so:

$\sum(X_i-\bar{X})^2 = \sum(X_i-\mu)^2 - n(\bar{X}-\mu)^2$

Now let us examine the expected value of the sample variance.