

More formal statistical calculations:

$E(\cdot)$: Expected value \rightarrow think average of everything in population

$$E(x_i) = \mu = \frac{1}{N} \sum_{i=1}^N x_i$$

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

$$\begin{aligned} \hookrightarrow E[(x_i - \mu)^2] &= E[(x_i^2 - 2x_i\mu + \mu^2)] \\ &= E(x_i^2) - 2\mu E(x_i) + \mu^2 \\ &= E(x_i^2) - 2\mu^2 + \mu^2 \\ &= E(x_i^2) - \mu^2 = \sigma^2 \quad \text{OR} \quad E(x_i^2) = \sigma^2 + \mu^2 \end{aligned}$$

What if we had $\hat{\sigma}^2 = \frac{1}{n} \sum (x_i - \bar{x})^2$ NOT $s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$??

$$\begin{aligned} E(\hat{\sigma}^2) &= \frac{1}{n} \sum_{i=1}^n E[(x_i - \bar{x})^2] = n \cdot \frac{1}{n} E[(x_i - \bar{x})^2] \\ &= E[(x_i - \bar{x})^2] \\ &= E(x_i^2) - 2E(x_i \bar{x}) + E(\bar{x}^2) \end{aligned}$$

BY FUN ALGEBRA

$$\begin{aligned} &= (\sigma^2 + \mu^2) - 2\left(\frac{n-1}{n} \mu^2 + \frac{1}{n} (\sigma^2 + \mu^2)\right) \\ &\quad + \left(\frac{n^2 - n}{n^2} \mu^2 + \frac{n}{n^2} (\sigma^2 + \mu^2)\right) \quad \text{EWW!!} \end{aligned}$$

BUT IT SIMPLIFIES

$$= \frac{n-1}{n} \sigma^2 \neq \sigma^2 \quad \text{uH OH!}$$

$$\text{Luckily!} \quad s^2 = \frac{n}{n-1} \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

$$E(s^2) = \frac{n}{n-1} E(\hat{\sigma}^2) = \frac{\cancel{n}}{\cancel{n-1}} \times \frac{n-1}{\cancel{n}} \times \sigma^2 = \sigma^2$$

YAY !!!