EE640 Lecture X: Unbiased estimate of Variance

Consider the identity

$$\sum_{i=1}^{n} (x_i - \mu)^2 = \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 + n(\bar{x}_n - \mu)^2 \quad (1)$$

where $\bar{x}_n = \frac{1}{n} \sum_{j=1}^{n} x_j$

We claim that $s_0^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$ is biased

Let us prove the identity in equation (1) then we will show s_0^2 is biased.

$$\sum_{i=1}^{n} (x_i^2 - 2\mu x_i + \mu^2) = \sum_{i=1}^{n} (x_i - \bar{x}_n)^2 + n(\bar{x}_n - \mu)^2$$

$$= \sum_{i=1}^{n} (x_i^2 - 2x_i \bar{x}_n + \bar{x}_n^2) + n(\bar{x}_n^2 - 2\mu \bar{x}_n + \mu^2)$$

②

$$= \sum_{i=1}^{n} x_i^2 - 2\bar{x}_n \sum_{i=1}^{n} x_i + \underbrace{n\bar{x}_n^2 + n\bar{x}_n^2}_{2n\bar{x}_n^2} - 2n\bar{x}_n\mu + n\mu^2$$

$$= \sum_{i=1}^{n} x_i^2 - 2\frac{1}{n}\sum_{j=1}^{n} x_j \sum_{i=1}^{n} x_i + 2\frac{\cancel{n}}{n^{\cancel{2}}}\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j - 2\frac{\cancel{n}\mu}{\cancel{n}}\sum_{i=1}^{n} x_i + n\mu^2$$

$$\rightarrow 0 \leftarrow$$

$$= \sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2 \qquad \text{QED}$$

So we know the Eq. (1) identity is valid

then $E\{S_0^2\} = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2\right\}$

$$= E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2\right\} - E\{(\bar{x}_n - \mu)^2\}$$

We know $E\{(x_i-\mu)^2\} = \sigma^2$

so $E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2\right\} = \frac{1}{n}\sum_{i=1}^{n}\sigma^2 = \sigma^2$

and $E\{(\bar{x}_n-\mu)^2\} = \frac{\sigma^2}{n}$

so $E\{s_0^2\} = \sigma^2 - \frac{1}{n}\sigma^2 = \frac{n-1}{n}\sigma^2$

so $s_1^2 = \frac{n}{n-1}s_0^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x}_n)^2$ is unbiased

such that $E\{s_1^2\} = \sigma^2$