M378K: March 7th, 2025.

Estimators.

Def'n. The bias of an estimator $\hat{\theta}$ of the parameter θ is defined as:

$$\text{bias}(\hat{\theta}) := \mathbb{E}(\hat{\theta} - \theta)$$

Notation from book: " $\mathbb{E}_{\theta}(\cdot)$, $\mathbb{E}^{\theta}(\cdot)$, $\mathbb{E}[\cdots/\theta]$ "

We say that an estimator $\hat{\theta}$ is unbiased for the parameter θ if

$$\boxed{\mathbb{E}[\hat{\theta}] = \theta \iff \text{bias}(\hat{\theta}) = 0}$$

for all possible values of θ.

Example. Consider a random sample $Y_1, Y_2, \dots, Y_n$ from $N(\mu, \sigma)$ w/ both $\mu \in \mathbb{R}$ and $\sigma > 0$ unknown

$$\hat{\mu} = \underbrace{\bar{Y}}_{\text{sample mean}} := \frac{Y_1 + Y_2 + \dots + Y_n}{n}$$

Then, $\boxed{\mathbb{E}[\hat{\mu}] = \mu}$, i.e., $\hat{\mu} = \bar{Y}$ is unbiased for μ.

Example. Let $Y_1, \dots, Y_n$ be a random sample from $N(\mu_0, \sigma)$ w/ μ_0 known and $\sigma > 0$ unknown

We propose this estimator for the variance σ^2:

$$S^2 := \frac{1}{n} \sum_{i=1}^{n} (Y_i - \mu_0)^2$$

Then,

$$\mathbb{E}[S^2] = \frac{1}{n} \sum_{i=1}^{n} \underbrace{\boxed{\mathbb{E}[(Y_i - \mu_0)^2]}}_{\sigma^2} = \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \sigma^2 = \sigma^2$$

$\Rightarrow S^2$ is unbiased for σ^2.

Example. Let $Y_1, Y_2, \ldots, Y_n$ be a random sample from $N(\mu, \sigma)$ with both μ and σ unknown.

Goal: Find a "good" estimator for σ^2!

Propose:

$$(S')^2 := \frac{1}{n} \boxed{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

Q: Is S' unbiased for σ^2?

$$\mathbb{E}[(S')^2] = \frac{1}{n} \sum_{i=1}^{n} \boxed{\mathbb{E}[(Y_i - \bar{Y})^2]}$$

$$\mathbb{E}[Y_i^2 - 2Y_i \cdot \bar{Y} + \bar{Y}^2]$$

$$= \frac{1}{n}\left(\sum_{i=1}^{n} \mathbb{E}[Y_i^2] - 2 \sum_{i=1}^{n} \mathbb{E}[Y_i \cdot \bar{Y}] + \sum_{i=1}^{n} \mathbb{E}[(\bar{Y})^2] \right)$$

$$= \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \mathbb{E}[Y_1^2] - 2\,\mathbb{E}\left[\boxed{\frac{1}{n}\sum_{i=1}^{n} Y_i} \cdot \bar{Y}\right] + \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \mathbb{E}[(\bar{Y})^2]$$

$$= \mathbb{E}[Y_1^2] - 2 \cdot \mathbb{E}[(\bar{Y})^2] + \mathbb{E}[(\bar{Y})^2]$$

$$= \underbrace{\mathbb{E}[Y_1^2]}_{\mathrm{Var}[Y_1] + (\mathbb{E}[Y_1])^2 \,=\, \sigma^2 + \mu^2} - \boxed{\mathbb{E}[(\bar{Y})^2]}$$

$$\mathbb{E}[(\bar{Y})^2] = \mathrm{Var}[\bar{Y}] + (\mathbb{E}[\bar{Y}])^2 = \frac{\sigma^2}{n} + \mu^2$$

$$\mathbb{E}[(S')^2] = \sigma^2 + \cancel{\mu^2} - \left(\frac{\sigma^2}{n} + \cancel{\mu^2} \right) = \left(1 - \frac{1}{n}\right)\sigma^2 = \boxed{\left(\frac{n-1}{n}\right)}\sigma^2$$

$$\mathbb{E}\left[(S')^2 \cdot \frac{n}{n-1}\right] = \sigma^2$$

$$\mathbb{E}\left[\frac{\cancel{n}}{n-1} \cdot \frac{1}{\cancel{n}} \sum (Y_i - \bar{Y})^2\right]$$

$$\Rightarrow \boxed{\mathrm{bias}((S')^2) = \mathbb{E}[(S')^2 - \sigma^2] = -\frac{\sigma^2}{n}}$$

So, the UNBIASED estimator for σ^2 is:

$$\boxed{S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$