M378K: October 25th, 2024.

Mean Squared Errors.

Def'n. Let $\hat{\theta}$ be an estimator for the parameter θ.

1. the error of $\hat{\theta}$ is $\hat{\theta} - \theta$
2. the absolute error of $\hat{\theta}$ is $|\hat{\theta} - \theta|$
3. the squared error of $\hat{\theta}$ is $(\hat{\theta} - \theta)^2$
4. the mean squared error of $\hat{\theta}$ is $\boxed{\mathbb{E}[(\hat{\theta} - \theta)^2] = MSE(\hat{\theta})}$

Proposition.

$$\boxed{MSE(\hat{\theta}) = (bias(\hat{\theta}))^2 + Var[\hat{\theta}]}$$

$\rightarrow$:

$$MSE(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2]$$

$$= \mathbb{E}[((\hat{\theta} - \mathbb{E}[\hat{\theta}]) + (\mathbb{E}[\hat{\theta}] - \theta))^2]$$

$$= \mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2 + 2(\hat{\theta} - \mathbb{E}[\hat{\theta}])(\mathbb{E}[\hat{\theta}] - \theta) + \underbrace{\boxed{(\mathbb{E}[\hat{\theta}] - \theta)^2}}_{constant}]$$

(linearity of $\mathbb{E}$)
$$= \underbrace{\boxed{\mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])^2]}}_{= Var[\hat{\theta}]} + 2\boxed{\mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])(\mathbb{E}[\hat{\theta}] - \theta)]} + \underbrace{\boxed{(\mathbb{E}[\hat{\theta}] - \theta)^2}}_{(bias(\hat{\theta}))^2}$$

Focus on:

$$\mathbb{E}[(\hat{\theta} - \mathbb{E}[\hat{\theta}])\underbrace{(\mathbb{E}[\hat{\theta}] - \theta)}_{constant}] = (\mathbb{E}[\hat{\theta}] - \theta)\underbrace{(\mathbb{E}[\hat{\theta}] - \mathbb{E}[\hat{\theta}])}_{=0}$$

$\square$

Def'n. The standard error of $\hat{\theta}$ is

$$\boxed{SE(\hat{\theta}) = \sqrt{Var[\hat{\theta}]}}$$

Def'n. An estimator $\hat{\theta}$ for θ is said to be **uniformly minimum variance unbiased estimator (UMVUE)** if:

1. $\hat{\theta}$ is unbiased, i.e., $\mathbb{E}[\hat{\theta}]=\theta$
2. $MSE(\hat{\theta}) \leq MSE(\hat{\theta}')$ where $\hat{\theta}'$ is **ANY** unbiased for θ

M378K Introduction to Mathematical Statistics
Problem Set #13
Bias. MSE.

Problem 13.1. Source: "Mathematical Statistics with Applications" by Wackerly, Mendenhall, Scheaffer.
Let $Y_1, , Y_2, Y_3$ be a random sample from $E(\tau)$. Consider the following five estimators of τ:

$$\hat{\theta}_1 = Y_1, \quad \hat{\theta}_2 = \frac{Y_1 + Y_2}{2}, \quad \hat{\theta}_3 = \frac{Y_1 + 2Y_2}{3}, \quad \hat{\theta}_4 = 3Y_{(1)}, \quad \hat{\theta}_5 = \bar{Y}.$$

Which ones of these estimators are unbiased? Among the unbiased ones which one has the smallest variance? ✓

$\mathbb{E}[\hat{\theta}_1] = \mathbb{E}[Y_1] = \tau$ ✓

$\mathbb{E}[\hat{\theta}_2] = \mathbb{E}\left[\frac{Y_1+Y_2}{2}\right] = \tau$ ✓

$\mathbb{E}[\hat{\theta}_3] = \mathbb{E}\left[\frac{Y_1+2Y_2}{3}\right] = \tau$ ✓

$\mathbb{E}[\hat{\theta}_4] = \mathbb{E}[3Y_{(1)}] = 3\,\mathbb{E}[Y_{(1)}] = 3 \cdot \frac{\tau}{3} = \tau$ ✓

$\mathbb{E}[\hat{\theta}_5] = \mathbb{E}[\bar{Y}] = \tau$ ✓

$Var[\hat{\theta}_1] = Var[Y_1] = \tau^2$

$Var[\hat{\theta}_2] = Var\left[\frac{Y_1+Y_2}{2}\right] = \frac{\tau^2}{2}$ ✗

$Var[\hat{\theta}_3] = Var\left[\frac{Y_1+2Y_2}{3}\right] = \frac{1}{9}\left(Var[Y_1] + 4Var[Y_2]\right) = \frac{5}{9}\tau^2$

$Var[\hat{\theta}_4] = Var[3Y_{(1)}] = 9 \cdot \left(\frac{\tau}{3}\right)^2 = \tau^2$

$Var[\hat{\theta}_5] = Var[\bar{Y}] = \frac{Var[Y_1] + Var[Y_2] + Var[Y_3]}{9} = \frac{3\tau^2}{9} = \frac{\tau^2}{3}$