Comparing estimators

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Uniform[0, \theta]$. Some possible estimates:

MLE: $\hat{\theta} = X_{(n)}$ $\quad f_{X_{(n)}}(x) = \dfrac{n x^{n-1}}{\theta^n}$

$\mathbb{E}[X_{(n)}] = \dfrac{n}{n+1}\theta$ (tends to underestimate)

$Var(X_{(n)}) = \dfrac{\theta^2 n}{(n+1)^2(n+2)}$

$P(|X_{(n)} - \theta| > \varepsilon) = P(X_{(n)} \leq \theta - \varepsilon)$

$= \left(\dfrac{\theta - \varepsilon}{\theta}\right)^n \to 0$

MOM: $\hat{\theta} = 2\bar{X}$

$\mathbb{E}[2\bar{X}] = 2\mathbb{E}[\bar{X}] = 2\dfrac{\theta}{2} = \theta$

$Var(2\bar{X}) = 4Var(\bar{X}) = \dfrac{4}{n}Var(X_i)$

$= \dfrac{4\theta^2}{12n} = \dfrac{\theta^2}{3n}$

$\hat{\theta} \xrightarrow{P} \theta$

$(\bar{X} \xrightarrow{P} \frac{\theta}{2} \Rightarrow 2\bar{X} \xrightarrow{P} \theta)$

> What properties might I want an estimator $\hat{\theta}$ to possess?

$\mathbb{E}[\hat{\theta}] = \theta$ $\qquad$ $Var(\hat{\theta})$ is small

Identifiability $\qquad$ $\hat{\theta} \xrightarrow{P} \theta$

$\hat{\theta} \approx$ Normal?

Bias, Variance and MSE

Mean squared error (MSE): Let $\hat{\theta}$ be an estimator of θ.

The MSE of $\hat{\theta}$ is $\mathbb{E}_\theta[(\hat{\theta}-\theta)^2]$

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

$$= \underbrace{\mathbb{E}_\theta[(\hat{\theta}-\mathbb{E}_\theta[\hat{\theta}])^2]}_{Var(\hat{\theta})} + \mathbb{E}_\theta[(\mathbb{E}_\theta[\hat{\theta}]-\theta)^2] + \underbrace{2\mathbb{E}_\theta[(\hat{\theta}-\mathbb{E}_\theta[\hat{\theta}])(\mathbb{E}_\theta[\hat{\theta}]-\theta)]}_{0}$$

$$Var(\hat{\theta}) + \underbrace{(\mathbb{E}_\theta[\hat{\theta}]-\theta)^2}_{\text{Bias}} \qquad 0$$

$$\text{Bias} = \mathbb{E}_\theta[\hat{\theta}] - \theta$$

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + \text{Bias}^2(\hat{\theta})$$

One approach is to try and minimize MSE

$X_i \overset{iid}{\sim}$ Uniform $[0, \theta]$

$$\text{Bias}(X_{(n)}) = \mathbb{E}[X_{(n)}] - \theta = \frac{\theta n}{n+1} - \theta = \frac{-\theta}{n+1}$$

$$\text{Var}(X_{(n)}) = \frac{\theta^2 n}{(n+1)^2(n+2)}$$

$$\text{MSE}(X_{(n)}) = \frac{\theta^2}{(n+1)^2} + \frac{\theta^2 n}{(n+1)^2(n+2)} = \frac{2\theta^2}{(n+1)(n+2)}$$

$$\text{MSE}(2\bar{X}) = \underset{\substack{\uparrow \\ \text{Bias}^2}}{0} + \frac{\theta^2}{3n} = \frac{\theta^2}{3n} > \text{MSE}(X_{(n)})$$

Try unbiasing $X_{(n)}$:

$$\hat{\theta} = \left(\frac{n+1}{n}\right) X_{(n)} \Rightarrow \mathbb{E}[\hat{\theta}] = \theta$$

$$\text{Var}(\hat{\theta}) = \left(\frac{n+1}{n}\right)^2 \text{Var}(X_{(n)})$$

$$= \frac{\theta^2}{n(n+2)} < \frac{2\theta^2}{(n+1)(n+2)} = \text{MSE}(X_{(n)})$$

Example

$\mathbb{E}[\frac{1}{n}\sum_i (X_i-\mu)^2] = \sigma^2$

$\mathbb{E}[\frac{1}{n}\sum_i (X_i-\bar{X})^2] = (\frac{n-1}{n})\sigma^2$

intuition: $\bar{X}$ minimizes $\sum_i (X_i-a)^2$

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$. On homework, we considered

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 \qquad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

and we showed that $\mathbb{E}\widehat{\sigma}^2 = \frac{n-1}{n}\sigma^2, \mathbb{E}(s^2) = \sigma^2$, and $\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$.

> Calculate the MSE of both $\widehat{\sigma}^2$ and s^2. It may help that if $V \sim \chi^2_\nu$, then $E[V] = \nu$ and $Var(V) = 2\nu$.

$$\mathbb{E}[S^2] = \sigma^2$$

$$\Rightarrow \text{Bias}(S^2) = \sigma^2 - \sigma^2 = 0$$

$$\text{Var}(S^2) = \text{Var}\left(\frac{\sigma^2}{n-1} \cdot \frac{n-1}{\sigma^2} \cdot S^2\right)$$

$$= \left(\frac{\sigma^2}{n-1}\right)^2 \text{Var}\left(\frac{n-1}{\sigma^2} \cdot S^2\right)$$

$$= \frac{\sigma^4}{(n-1)^2} \cdot 2(n-1) = \frac{2\sigma^4}{n-1}$$

$$\text{MSE}[S^2] = \frac{2\sigma^4}{n-1}$$

$$\mathbb{E}[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2 \Rightarrow \text{Bias}(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2 - \sigma^2 = -\frac{\sigma^2}{n}$$

$$\text{Var}(\hat{\sigma}^2) = \text{Var}\left(\frac{n-1}{n} S^2\right) = \left(\frac{n-1}{n}\right)^2 \text{Var}(S^2) = \frac{2\sigma^4(n-1)}{n^2}$$

$$\text{MSE}(\hat{\sigma}^2) = \left(\frac{\sigma^2}{n}\right)^2 + \frac{2\sigma^4(n-1)}{n^2} = \frac{(2n-1)\sigma^4}{n^2} < \frac{2\sigma^4}{n-1}$$

$$\text{MSE}(\hat{\sigma}^2) < \text{MSE}(S^2)$$

MSE and consistency

Best unbiased estimators

Cramer-Rao lower bound