Consistent Estimators:

Let $\{\hat{\theta}_n\}$ be a sequence of estimators, where $\hat{\theta}_n$ is based on a sample of size n.

The sequence is <u>consistent</u> for θ if $\hat{\theta}_n$ converges in probability to θ as $n \to \infty$:

For any $\varepsilon > 0$, $\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| > \varepsilon) = 0$

- Intuition: If n large enough, probability $\uparrow$ that $\hat{\theta}_n$ is very close to θ.
- Example: Law of Large Numbers says that $\bar{X}$ is a consistent estimator of μ.

<u>Thm:</u> Suppose $X_1, \dots, X_n$ are iid r.v.'s with pdf $f_X(x|\theta)$. Under smoothness conditions on f, $\hat{\theta}^{MLE}$ is a consistent estimator of θ.

Cramér-Rao Lower Bound (CRLB):

Let $X_1, \dots, X_n$ be i.i.d. r.v.'s with pdf $f_X(x|\theta)$, and let $T = t(X_1, \dots, X_n)$ be an unbiased estimator of θ. Then, under smoothness conditions on f,

$$Var(T) \geq \frac{1}{I(\theta)}$$

Define efficiency: Unbiased, $Var(\hat{\theta}) = \frac{1}{I(\theta)}$

Example: We showed that if $X_1, \dots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$, $I(\lambda) = \frac{n}{\lambda}$

Also, $\hat{\lambda}^{MLE} = \bar{X}$. $Var(\hat{\lambda}^{MLE}) = Var(\bar{X}) = Var(\frac{1}{n}\sum_{i=1}^{n} X_i) = \frac{1}{n^2} \cdot n\, Var(X_i) = \frac{1}{n}\lambda = \frac{1}{I(\lambda)}$.

$\hat{\lambda}^{MLE}$ achieves the CRLB for the Poisson distribution. No unbiased estimator could have lower variance. <u>BUT</u>, a biased estimator could have lower variance & MSE...

Example: Estimators of variance of a normal distribution

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$

Define $\hat{\sigma}_1^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$

- We have shown in HW that $E[\hat{\sigma}_1^2] = \sigma^2$:
  $\hat{\sigma}_1^2$ is an unbiased estimator of σ^2.
- Also, can show that $Var(\hat{\sigma}_1^2) = \frac{2\sigma^4}{n-1}$

$$\therefore MSE(\hat{\sigma}_1^2) = \{\overset{0}{\cancel{Bias(\hat{\sigma}_1^2)}}\}^2 + Var(\hat{\sigma}_1^2) = \frac{2\sigma^4}{n-1}$$

Now consider $\hat{\sigma}^2_{MLE} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2 = \frac{n-1}{n} S^2$

$$E[\hat{\sigma}^2_{MLE}] = E\left(\frac{n-1}{n} S^2\right) = \frac{n-1}{n} E(S^2) = \frac{n-1}{n} \sigma^2$$

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

$$\therefore MSE(\hat{\sigma}^2_{MLE}) = \left\{\frac{n-1}{n}\sigma^2 - \sigma^2\right\}^2 + \frac{n-1}{n^2} \cdot 2\sigma^4$$

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

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

$\cancel{MSE(\hat{\sigma}^2) = \frac{2\sigma^4}{n-1}}$

$$MSE(\hat{\sigma}^2_{MLE}) = \left(\frac{2n-1}{n^2}\right)\sigma^4 < \frac{2n}{n^2}\sigma^4 = \frac{2}{n}\sigma^4 < \frac{2\sigma^4}{n-1} = MSE(\hat{\sigma}_1^2)$$

S^2 is unbiased, but has larger MSE than $\hat{\sigma}^2_{MLE}$.

It can be shown that among all estimators of the form
$\hat{\sigma}_c^2 = c \cdot \sum_{i=1}^{n} (X_i - \bar{X})^2$, $c = \frac{1}{n+1}$ results in lowest MSE.

Our first example of a shrinkage estimator: Shrink the estimate towards 0, introducing some bias in exchange for a greater reduction in variance, and MSE.