2.2 Properties of Point Estimators - Efficiency

Steven Cheung
9 February 2022

Definition:

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two unbiased estimators for a parameter θ. If

$$var(\hat{\theta}_1) < var(\hat{\theta}_2)$$

we say that $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$.
The relative efficiency of $\hat{\theta}_1$ with respect to $\hat{\theta}_2$ is the ratio $\frac{var(\hat{\theta}_1)}{var(\hat{\theta}_2)}$.

Definition:

let Θ denote a set of all esimators $\hat{\theta} = h(X_1, X_2, \ldots, X_n)$ that are unbiased for the parameter θ in the continuous $p.d.f$ $f_x(x, \theta)$ (or discrete $p.m.f$ $p_x(x, \theta)$). The estimator $\hat{\theta}$ is the best (or unbiased minimum variance) estimator if $\hat{\theta^*} \in \Theta$ and

$$var(\hat{\theta^*}) \leq var(\hat{\theta}) \text{ for all } \hat{\theta} \in \Theta$$

Theorem (The Cramer-Rao Lower Bound):

Let $f_X(X, \theta)$ be a continuous *pdf* with continuous first-order and second-order derivatives.
Let $X_1, X_2, \ldots, X_n$ be a random sample from $f_X(x, \theta)$, and suppose that the set of X values, where $f_X(x, \theta) \neq 0$ does not depend on θ. Let $\hat{\theta} = g(X_1, X_2, \ldots, X_n)$ be any unbiased estitmator of θ. Then

$$var(\hat{\theta}) \geq \{nE[\frac{\partial ln(f_X(X,\theta))}{\partial \theta}]\}^{-1} = \{-nE[\frac{\partial^2 ln(f_X(X,\theta))}{\partial \theta^2}]\}^{-1}$$

If the variance of a given $\hat{\theta}$ is equal to the Cramer-Rao lower bound we say that the estimator is optimal in a sense that no unbiased $\hat{\theta}$ can estimate θ with greater precision.

The unbiased estimator $\hat{\theta}$ is said to be efficient if the variance of $\hat{\theta}$ equals to the Cramer-Rao lower bound associated with $f_x(x, \theta)$.

The efficiency of an unbiased estimator $\hat{\theta}$ is the ratio of the Cramer-Rao lower bound for $f_x(x, \theta)$ to the variance of $\hat{\theta}$.

14 February 2022

Definition:

The mean square error of the estimator $\hat{\theta}$, denoted by $MSE(\hat{\theta})$, is defined as

$$MSE(\hat{\theta}) = E(\hat{\theta} = \theta)^2$$

$$\begin{aligned} MSE(\hat{\theta}) &= E(\hat{\theta} = \theta)^2 = E[(\hat{\theta} - E(\hat{\theta})) + (E(\hat{\theta}) - \hat{\theta})]^2 \\ &= E[(\hat{\theta} - E(\hat{\theta}))^2 + (E(\hat{\theta}) - \hat{\theta})^2 + 2(\hat{\theta} - E(\hat{\theta}))(E(\hat{\theta}) - \hat{\theta})] \\ &= E(\hat{\theta} - E(\hat{\theta}))^2 + E(E(\hat{\theta}) - \hat{\theta})^2 + 2E(\hat{\theta} - E(\hat{\theta}))(E(\hat{\theta}) - \hat{\theta}) \\ &= var(\hat{\theta}) + (E(\hat{\theta}) - \theta)^2 = var(\hat{\theta}) + (Bias(\hat{\theta}, \theta))^2 \end{aligned}$$