Efficient Estimators

Gene Quinn

Efficiency

We have identified unbiasedness as a desirable characteristic of an estimator.

However, there may be a great many unbiased estimators for a parameter.

Efficiency

For example, if x_1 and x_2 are a random sample from a $N(\mu, \sigma^2)$ distribution, it is easy to verify that

$$h_1(x_1, x_2) = \frac{1}{2}x_1 + \frac{1}{2}x_2$$

and

$$h_2(x_1, x_2) = \frac{1}{3}x_1 + \frac{2}{3}x_2$$

are both unbiased estimators for μ (that is, $\mathsf{E}(h_1) = \mathsf{E}(h_2) = \mu$).

Efficiency

For example, if x_1 and x_2 are a random sample from a $N(\mu, \sigma^2)$ distribution, it is easy to verify that

$$h_1(x_1, x_2) = \frac{1}{2}x_1 + \frac{1}{2}x_2$$

and

$$h_2(x_1, x_2) = \frac{1}{3}x_1 + \frac{2}{3}x_2$$

are both unbiased estimators for μ (that is, $\mathrm{E}(h_1) = \mathrm{E}(h_2) = \mu$).

In fact, there are an infinite number of functions of x_1 and x_2 that are unbiased estimators for μ.

How do we choose among them? Does it make any difference which one we use?

Efficiency

In matrix notation,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad V = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Efficiency

In matrix notation,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \quad V = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Now

$$h_1(x_1, x_2) = t'\vec{x} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Efficiency

In matrix notation,

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \qquad \vec{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \qquad V = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix}$$

Now

$$h_1(x_1, x_2) = t'\vec{x} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Consequently, the variance of $h_1(x_1, x_2) = \hat{\mu}_1$ is

$$\text{Var}(\hat{\mu}_1) = t'Vt = \frac{\sigma^2}{4} + \frac{\sigma^2}{4} = \frac{\sigma^2}{2}$$

Efficiency

Second, $h_2(x_1, x_2) = \hat{\mu}_2$ is

$$h_2(x_1, x_2) = t'\vec{x} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Efficiency

Second, $h_2(x_1, x_2) = \hat{\mu}_2$ is

$$h_2(x_1, x_2) = t'\vec{x} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Consequently, the variance of $\hat{\mu}_2$ is

$$\text{Var}(\hat{\mu}_2) = t'Vt = \frac{\sigma^2}{9} + \frac{4\sigma^2}{9} = \frac{5\sigma^2}{9}$$

so

$$\text{Var}(\hat{\mu_1}) = \frac{\sigma^2}{2} < \text{Var}(\hat{\mu_2}) = \frac{5\sigma^2}{9}$$

Efficiency

So the estimator $\hat{\mu}_1$ for μ is "better" than $\hat{\mu}_2$ in the sense that, while both are unbiased, $\hat{\mu}_1$ has smaller variance than $\hat{\mu}_2$.

Of course, this raises the question of whether there is an unbiased estimator with still smaller variance.

Efficiency

So the estimator $\hat{\mu}_1$ for μ is "better" than $\hat{\mu}_2$ in the sense that, while both are unbiased, $\hat{\mu}_1$ has smaller variance than $\hat{\mu}_2$.

Of course, this raises the question of whether there is an unbiased estimator with still smaller variance.

A famous and very powerful theorem known as the *Cramer-Rao Inequality* answers this question, at least within the class of unbiased estimators.

The Cramer-Rao inequality establishes the smallest possible variance that an unbiased estimator can have.

Cramer-Rao Inequality

Theorem: (Cramer-Rao Inequality) Let $x_1, \ldots, x_n$ be a random sample from a population with continuous density function $f_X(x;\theta)$.

Suppose that $f_X(x;\theta)$ has continuous first and second order partial derivatives everywhere except possibly at a finite set of points, and the support of f does not depend on θ.

Cramer-Rao Inequality

Let

$$\hat{\theta} = h(x_1 \ldots, x_n)$$

be any unbiased estimator for θ.

Then

$$\text{Var}(\hat{\theta}) \geq \left\{ n\text{E}\left[\left(\frac{\partial \ln f_X(x;\theta)}{\partial \theta}\right)^2\right]\right\}^{-1}$$

$$= \left\{ -n\text{E}\left[\frac{\partial^2 \ln f_X(x;\theta)}{\partial \theta^2}\right]\right\}^{-1}$$

Cramer-Rao Inequality

In the application of this theorem, we are interested in finding the lower bound for the variance of the unbiased estimator $\hat{\theta}$.

Cramer-Rao Inequality

In the application of this theorem, we are interested in finding the lower bound for the variance of the unbiased estimator $\hat{\theta}$.

The theorem gives us two equivalent ways to do this: we can either evaluate

$$\left\{ n\mathsf{E}\left[\left(\frac{\partial \ln f_X(x;\theta)}{\partial \theta} \right)^2 \right] \right\}^{-1}$$

or

$$\left\{ -n\mathsf{E}\left[\frac{\partial^2 \ln f_X(x;\theta)}{\partial \theta^2} \right] \right\}^{-1}$$

Efficiency

Let $x_1, \ldots, x_n$ be a random sample from a population with continuous pdf $f_X(x; \theta)$, and let

$$\hat{\theta} \; = \; h(x_1, \ldots, x_n)$$

be an unbiased estimator for θ.

Efficiency

Let $x_1, \ldots, x_n$ be a random sample from a population with continuous pdf $f_X(x;\theta)$, and let

$$\hat{\theta} \;=\; h(x_1, \ldots, x_n)$$

be an unbiased estimator for θ.

Definition: An unbiased estimator $\hat{\theta}$ for θ is called **efficient** if Var$(\hat{\theta})$ is equal to the Cramer-Rao lower bound associated with $f(X;\theta)$.

The **efficiency** of $\hat{\theta}$ is defined to be the ratio of the Cramer-Rao lower bound associated with $f(x;\theta)$ to the variance of $\hat{\theta}$.

Best Estimators

Let $x_1, \ldots, x_n$ be a random sample from a population with continuous pdf $f_X(x;\theta)$, and let

$$\hat{\theta} \;=\; h(x_1, \ldots, x_n)$$

be an unbiased estimator for θ.

Let Θ be the set of all unbiased estimators for θ.

Best Estimators

Let $x_1, \ldots, x_n$ be a random sample from a population with continuous pdf $f_X(x;\theta)$, and let

$$\hat{\theta} \;=\; h(x_1, \ldots, x_n)$$

be an unbiased estimator for θ.

Let Θ be the set of all unbiased estimators for θ.

Definition: An unbiased estimator $\hat{\theta}^* \in \Theta$ for θ is called a **best** or **minimum variance** estimator if

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

Best Estimators

Note that an unbiased estimator $\hat{\theta}$ with variance equal to the Cramer-Rao lower bound is automatically a best or minimum variance estimator.

Best Estimators

Note that an unbiased estimator $\hat{\theta}$ with variance equal to the Cramer-Rao lower bound is automatically a best or minimum variance estimator.

However, in some cases no unbiased estimators achieve the Cramer-Rao lower bound.

In these cases it is possible for an estimator to be best or minimum variance even though it has a variance greater than the Cramer-Rao lower bound.