LECTURE - 13

Agenda:

1. Relative efficiency
2. Examples

We will formalize our understanding of one estimator being "better" than another estimator, through the notion of relative efficiency. Suppose we have two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ for the same parameter θ. We want to understand which one of them is better.

RULE: If $MSE(\hat{\theta}_1) < MSE(\hat{\theta}_2)$, then $\hat{\theta}_1$ is a better estimator than $\hat{\theta}_2$.

Definition: Given two estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ of a parameter θ, the relative efficiency of $\hat{\theta}_1$ with respect to $\hat{\theta}_2$ is defined as

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{MSE(\hat{\theta}_2)}{MSE(\hat{\theta}_1)}.$$

If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased, then

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)}.$$

Clearly, if

$$\text{eff}(\hat{\theta}_1, \hat{\theta}_2) > 1,$$

then $\hat{\theta}_1$ is a better estimator of θ than $\hat{\theta}_2$.

Example 1: Let $Y_1, Y_2, ---, Y_n$ denote a random sample from the uniform distribution on the interval $(0, \theta)$. Two unbiased estimators for θ are

$$\hat{\theta}_1 = 2\bar{Y} \quad \text{and} \quad \hat{\theta}_2 = \left(\frac{n+1}{n}\right) \underbrace{\max(Y_1, Y_2, ---, Y_n)}_{\text{denoted as } Y_{(n)}}$$

TASK: Find the relative efficiency of $\hat{\theta}_1$ with respect to $\hat{\theta}_2$.

Note that $E[2\bar{Y}] = \frac{2}{n}\sum_{i=1}^{n} E[Y_i] = \frac{2}{n}\sum_{i=1}^{n}\frac{\theta}{2} = \theta.$

$$V(2\bar{Y}) = 4V(\bar{Y}) = \frac{4}{n}V(Y_1) = \frac{4(\theta-0)^2}{n \times 12} = \frac{\theta^2}{3n}.$$

Let us now derive the MSE of $\hat{\theta}_2 = \left(\frac{n+1}{n}\right) Y_{(n)}$.

Note that the density of $Y_{(n)}$ is given by

$$f_{Y_{(n)}}(y) = n\left(F_y(y)\right)^{n-1} f_y(y) = \begin{cases} n\left(\frac{y}{\theta}\right)^{n-1}\left(\frac{1}{\theta}\right) \text{ if } 0 \le y \le \theta, \\ 0 \text{ otherwise.} \end{cases}$$

Thus,

$$E[Y_{(n)}] = \frac{n}{\theta^n}\int_0^\theta y \cdot y^{n-1}\,dy = \frac{n}{(n+1)\theta^n}\left[y^{n+1}\right]_0^\theta$$

$$= \frac{n}{n+1}\theta.$$

$$\Rightarrow E\left[\left(\frac{n+1}{n}\right)Y_{(n)}\right] = \left(\frac{n+1}{n}\right)\left(\frac{n}{n+1}\right)\theta = \theta.$$

$$MSE(\hat{\theta}_2) = V(\hat{\theta}_2) = V\left(\left(\frac{n+1}{n}\right)Y_{(n)}\right)$$

$$= \frac{(n+1)^2}{n^2} V(Y_{(n)})$$

$$= \frac{(n+1)^2}{n^2}\left\{E[Y_{(n)}^2] - \left(E[Y_{(n)}]\right)^2\right\}$$

$$= \frac{(n+1)^2}{n^2}E[Y_{(n)}^2] - \frac{(n+1)^2}{n^2}\left(\frac{n}{n+1}\theta\right)^2$$

$$= \frac{(n+1)^2}{n^2}E[Y_{(n)}^2] - \theta^2.$$

Note that

$$E[Y_{(n)}^2] = \int_{-\infty}^{\infty} y^2 \cdot f_{Y_{(n)}}(y)\,dy$$

$$= n\int_0^{\theta} y^2 \frac{y^{n-1}}{\theta^n}\,dy$$

$$= \frac{n}{n+2}\theta^2.$$

Hence, $\quad MSE(\hat{\theta}_2) = \frac{(n+1)^2}{n^2}\cdot\frac{n}{n+2}\theta^2 - \theta^2$

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

$$= \frac{1}{n(n+2)}\theta^2.$$

Hence $\quad eff(\hat{\theta}_1, \hat{\theta}_2) = \frac{V(\hat{\theta}_2)}{V(\hat{\theta}_1)} = \frac{\frac{1}{n(n+2)}\theta^2}{\frac{1}{3n}\theta^2}$

$$= \frac{3}{n+2}.$$

If $n \geq 1$, $eff(\hat{\theta}_1, \hat{\theta}_2) < 1$, in which case $\hat{\theta}_2$ is a better estimator of θ than $\hat{\theta}_1$.

Example 2: Suppose that $Y_1, Y_2, ---, Y_n$ is a random sample from a normal distribution with mean μ and variance σ^2. Two unbiased estimators of σ^2 are

$$\hat{\sigma}_1^2 = S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 \text{ and } \hat{\sigma}_2^2 = \frac{1}{2}(Y_1 - Y_2)^2 .$$

TASK: Find the relative efficiency of $\hat{\sigma}_1^2$ with respect to $\hat{\sigma}_2^2$.

We already have derived that

$$E[\hat{\sigma}_1^2] = E[S^2] = \sigma^2,$$

and,

$$MSE[\hat{\sigma}_1^2] = V(S^2) = \frac{2\sigma^4}{n-1}.$$

Let us derive the expectation and mean squared error for $\hat{\sigma}_2^2$.

$$E[\hat{\sigma}_2^2] = E\left[\frac{1}{2}(Y_1 - Y_2)^2\right]$$

$$= \frac{1}{2}\left\{E[Y_1^2] + E[Y_2^2] - 2E[Y_1 Y_2]\right\}$$

$$= \frac{1}{2}\left\{ V(Y_1) + (E[Y_1])^2 + V(Y_2) + (E[Y_2])^2 - 2E[Y_1]E[Y_2]\right\}$$

$$= \frac{1}{2}\left\{\sigma^2+\mu^2+\sigma^2+\mu^2-2\mu^2\right\}$$

$$= \sigma^2.$$

$$MSE(\hat{\sigma}_2^2) = V\left(\frac{(Y_1-Y_2)^2}{2}\right) = E\left[\frac{(Y_1-Y_2)^4}{4}\right] - \left(E\left(\frac{(Y_1-Y_2)^2}{2}\right)\right)^2.$$

$$E\left[\frac{(Y_1-Y_2)^4}{4}\right] = \frac{1}{4}\left\{E[Y_1^4]+E[Y_2]^4-4E[Y_1^3Y_2] - 4E[Y_1Y_2^3]+6E[Y_1^2Y_2^2]\right\}$$

We can simplify this, but a much easier approach is the following. Let us pretend that we only have a sample Y_1, Y_2 of size 2. Then, the adjusted variance estimate based on this sample of size 2 is

$$S_2^2 = \frac{1}{2-1}\sum_{i=1}^{2}\left(Y_i-\left(\frac{Y_1+Y_2}{2}\right)\right)^2$$

$$= \left(Y_1-\frac{(Y_1+Y_2)}{2}\right)^2+\left(Y_2-\left(\frac{Y_1+Y_2}{2}\right)\right)^2$$

$$= \left(\frac{Y_1-Y_2}{2}\right)^2+\left(\frac{Y_2-Y_1}{2}\right)^2$$

$$= 2\left(\frac{Y_1-Y_2}{2}\right)^2$$

$$= \frac{1}{2}(Y_1-Y_2)^2.$$

We know that if we have a sample of size n, then the MSE of the adjusted sample variance is $\frac{2\sigma^4}{n-1}$. Hence with a sample of size 2,

$$MSE(S_2^2) = MSE\left(\frac{1}{2}(Y_1 - Y_2)^2\right) = \frac{2\sigma^4}{2-1} = 2\sigma^4.$$

Hence, the relative efficiency of $\hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ is given by

$$eff(\hat{\sigma}_1^2, \hat{\sigma}_2^2) = \frac{2\sigma^4}{\frac{2\sigma^4}{n-1}} = n-1.$$

If $n > 2$, $eff(\hat{\sigma}_1^2, \hat{\sigma}_2^2) > 1$, which means that $\hat{\sigma}_1^2$ is a better estimator of σ^2 than $\hat{\sigma}_2^2$.