1 Relative Efficiency

- Now that we have covered some basics of Estimation, we are going to cover some basic Estimation Theory concepts
- The first step in evaluating estimators is to measure them in terms of qualities that we want
- We already covered this when we discussed:
  - Bias
  - MSE
- Now that we have a yard stick to measure estimators by, the next step is to compare estimators
- One way to compare estimators is in terms of efficiency, which is related to MSE
- Efficiency refers to the concept of how much does our estimator vary
- So, when comparing estimators it makes sense that we would want o consider their relative efficiency

Definition: Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two unbiased estimators for θ. Then the *efficiency* of $\hat{\theta}_1$ *relative* to $\hat{\theta}_2$ is defined to be

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

We would like to describe one estimator as better than another in terms of efficiency if it is more efficient than the other. In terms of the relative efficiency, we would say that $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$ iff

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

We would say that $\hat{\theta}_2$ is more efficient than $\hat{\theta}_1$ iff

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

Finally, we would say that the two are equally efficient iff

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

1.1 Examples

Let $X_1, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$.

1. Show that $\text{eff}(X_i, X_j) = 1$ for $i, j = 1, ..., n$, where X_i and X_j are estimators for μ
   Solution:

$$\begin{aligned}
X_i &\sim N(\mu, \sigma^2) \text{ for all } i \\
X_j &\sim N(\mu, \sigma^2) \text{ for all } j \\
\Rightarrow \text{eff}(X_i, X_j) &= \frac{V[X_j]}{V[X_i]} \\
&= \frac{\sigma^2}{\sigma^2} \\
&= 1
\end{aligned}$$

2. Show that $\text{eff}(\frac{X_1+X_2}{2}, X_1) > 1$
   Solution:

$$\begin{aligned}
X_1, X_2 &\quad \text{i.i.d.} \quad N(\mu, \sigma^2) \\
\Rightarrow \frac{X_1 + X_2}{2} &\sim N(\mu, \frac{\sigma^2}{2}) \\
\Rightarrow \text{eff}(\frac{X_1 + X_2}{2}, X_1) &= \frac{V[X_1]}{V[\frac{X_1+X_2}{2}]} \\
&= \frac{\sigma^2}{\sigma^2/2} \\
&= 2
\end{aligned}$$

   So, we would say that $\frac{X_1+X_2}{2}$ is *twice* as efficient as X_1 alone.

3. Based on their relative efficiency, which estimator would you prefer to use for μ, $\frac{X_1+X_2}{2}$ or $\frac{X_1+X_2+X_3}{3}$?
   Solution:

$$\begin{aligned}
X_1, X_2, X_3 &\quad \text{i.i.d.} \quad N(\mu, \sigma^2) \\
\Rightarrow \frac{X_1 + X_2}{2} &\sim N(\mu, \frac{\sigma^2}{2}) \\
\Rightarrow \frac{X_1 + X_2 + X_3}{3} &\sim N(\mu, \frac{\sigma^2}{3}) \\
\Rightarrow \text{eff}(\frac{X_1 + X_2}{2}, \frac{X_1 + X_2 + X_3}{3}) &= \frac{V[\frac{X_1+X_2+X_3}{3}]}{V[\frac{X_1+X_2}{2}]} \\
&= \frac{\sigma^2/3}{\sigma^2/2} \\
&= \frac{2}{3}
\end{aligned}$$

So, since eff$(\frac{X_1+X_2}{2}, \frac{X_1+X_2+X_3}{3}) < 1$ we would conclude that $\frac{X_1+X_2}{2}$ is worse than $\frac{X_1+X_2+X_3}{3}$

1.2 Exercises

Let X_1, X_2 be i.i.d. $U(0,\theta)$ and let $\hat{\theta}_1 = 2X_1$, $\hat{\theta}_2 = X_1 + X_2$, and $\hat{\theta}_3 = \frac{3}{2}X_{(2)}$

1. Show that all of these estimators are unbiased
2. Compare the relative efficiencies between all three (i.e. compared the relative efficency of $\hat{\theta}_1$ and $\hat{\theta}_2$, $\hat{\theta}_1$ and $\hat{\theta}_3$, and $\hat{\theta}_2$ and $\hat{\theta}_3$
3. Which estimator would you prefer to use based on the relative efficiencies?

1.3 Solutions

Let X_1, X_2 be i.i.d. $U(0,\theta)$ and let $\hat{\theta}_1 = 2X_1$, $\hat{\theta}_2 = X_1 + X_2$, and $\hat{\theta}_3 = \frac{3}{2}X_{(2)}$

1. Show that all of these estimators are unbiased
   Solution:

$$
\begin{array}{rcl}
X_1, X_2 & \text{i.i.d.} & U(0,\theta) \\
\Rightarrow X_{(2)}/\theta & \sim & \beta(2,1) \\
\rightarrow E[2X_1] & = & 2\frac{\theta}{2} \\
& = & \theta \\
\rightarrow E[X_1 + X_2] & = & EX_1 + E[X_2] \\
& = & \frac{\theta}{2} + \frac{\theta}{2} \\
& = & \theta \\
\rightarrow E[\frac{3}{2}X_{(2)}] & = & E[\frac{3}{2}(\theta)\frac{X_{(2)}}{\theta}] \\
& = & \frac{3\theta}{2}E[\frac{X_{(2)}}{\theta}] \\
& = & \frac{3\theta}{2}\frac{2}{2+1} \\
& = & \theta
\end{array}
$$

2. Compare the relative efficiencies between all three (i.e. compared the relative efficency of $\hat{\theta}_1$ and $\hat{\theta}_2$, $\hat{\theta}_1$ and $\hat{\theta}_3$, and $\hat{\theta}_2$ and $\hat{\theta}_3$

$$
\text{eff}(2X_1, X_1 + X_2) = \frac{V[X_1 + X_2]}{V[2X_1]}
$$

$$
\begin{aligned}
&= \frac{V[X_1]+V[X_2]}{4V[X_1]} \leftarrow \text{Since, } X_1 \text{ and } X_2 \text{ are i.i.d.} \\
&= \frac{\theta^2/12+\theta^2/12}{4\theta^2/12} \\
&= \frac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
\text{eff}(2X_1, \frac{3}{2}X_{(2)}) &= \frac{V[\frac{3}{2}X_{(2)}]}{V[2X_1]} \\
&= \frac{\frac{9\theta^2}{4}V[X_{(2)}/\theta]}{4V[X_1]} \\
&= \frac{\frac{9\theta^2}{4}\frac{2}{(2+1)^2(2+1+1)}}{4\theta^2/12} \\
&= \frac{3}{8} \\
\text{eff}(X_1+X_2, \frac{3}{2}X_{(2)}) &= \frac{V[\frac{3}{2}X_{(2)}]}{V[X_1+X_2]} \\
&= \frac{\frac{9\theta^2}{4}V[X_{(2)}/\theta]}{V[X_1]+V[X_2]} \\
&= \frac{\frac{9\theta^2}{4}\frac{2}{(2+1)^2(2+1+1)}}{2\theta^2/12} \\
&= \frac{3}{4}
\end{aligned}
$$

Note: $\text{eff}(X_1+X_2, \frac{3}{2}X_{(2)}) = \text{eff}(2X_1, \frac{3}{2}X_{(2)})/\text{eff}(2X_1, X_1+X_2)$

3. Which estimator would you prefer to use based on the relative efficiencies?
Solution:
Since $\text{eff}(X_1+X_2, \frac{3}{2}X_{(2)}) < 1$, $\text{eff}(2X_1, \frac{3}{2}X_{(2)}) < 1$, and $\text{eff}(2X_1, X_1+X_2) < 1$ we can rank the estimators in terms of efficiency as

   a) $\frac{3}{2}X_{(2)}$
   b) X_1+X_2
   c) $2X_1$

   Therefore, $\frac{3}{2}X_{(2)}$ is the best in terms of efficiency

2 Consistency

Another way we can evaluate estimators, is how that perform with large sample sizes. We encapsulate this idea with *Consistency*

Definition:
An estimator $\hat{\theta}_n$ is a *consistent estimator* for θ iff, for any $\varepsilon > 0$

$$\lim_{n\to\infty} P(|\hat{\theta}_n - \theta| \le \varepsilon) = 1$$

or, equivalently,

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

Basically, what this is saying is that no matter how close we ask that the estimator to be to θ there is an n large enough to guarantee that the estimator will be that close to θ.

Theorem: An unbiased estimator $\hat{\theta}_n$ is a consistent estimator of θ if

$$\lim_{n\to\infty} V[\hat{\theta}_n] = 0$$

Proof. Recall:
Remeber Tchebysheff's inequality which states:
For R.V. X with $E[X]$ and $V[X] < \infty$, then for any $k > 0$

$$P(|X - E[X]| > k\sqrt{V[X]}) \le \frac{1}{k^2}$$

So, letting $X = \hat{\theta}_n$ in the above inequality, we have

$$P(|\hat{\theta} - E[\hat{\theta}_n]| > k\sqrt{V[\hat{\theta}_n]}) \le \frac{1}{k^2}$$

for any $k > 0$.
Now, we will use this theorem to show that any unbiased estimator $\hat{\theta}_n$ that has the property $\lim_{n\to\infty} V[\hat{\theta}_n] = 0$ satisfies the defintion of a consistent estimator. So, we selct an aribitrary $\varepsilon > 0$, and let $k = \frac{\varepsilon}{\sqrt{V[\hat{\theta}_n]}}$. Now, we plug this value for k into Tchebyscheff's inequality and we get

$$\begin{aligned}
P(|\hat{\theta}_n - E[\hat{\theta}_n]| > \frac{\varepsilon}{\sqrt{V[\hat{\theta}_n]}}\sqrt{V[\hat{\theta}_n]}) &= P(|\hat{\theta}_n - E[\hat{\theta}_n]| > \varepsilon) \\
&\le \left(\frac{1}{\varepsilon/\sqrt{V[\hat{\theta}_n]}}\right)^2 \\
&= \frac{V[\hat{\theta}_n]}{\varepsilon^2} \\
\Rightarrow \lim_{n\to\infty} P(|\hat{\theta}_n - E[\hat{\theta}_n]| > \varepsilon) &\le \lim_{n\to\infty} \frac{V[\hat{\theta}_n]}{\varepsilon^2} \\
&= 0
\end{aligned}$$

Because we selected ε arbitrarily, we see that this implies that

$$\Rightarrow \lim_{n\to\infty} P(|\hat{\theta}_n - E[\hat{\theta}_n]| > \varepsilon) = 0$$

for all $\varepsilon > 0$. Thus $\hat{\theta}_n$ is a consistent estimator for θ $\square$

The definition of a consistent estimator is actually identical the a more general definition that applies to all random variables:
Definition:
A sequence of random variables X_n *converges in probability* to a value θ iff for any $\varepsilon > 0$

$$\lim_{n\to\infty} P(|X_n - \theta| \leq \varepsilon) = 1$$

or, equivalently,

$$\lim_{n\to\infty} P(|X_n - \theta| > \varepsilon) = 0$$

So, we might say that an estimator is consistent for θ if the estimator converges in probability to θ.

Sometimes we want to show that an estimator $\hat{\theta}_n$ *cannot* be consistent for a parameter θ. In these cases we can use the following theorem (whose proof will be omitted):
Theorem:
Given estimator (or sequence of Random variables) $\hat{\theta}_n$, if $\lim_{n\to\infty} V[\hat{\theta}_n] = c$, where $c > 0$ is a constant value for all n, then we can say that $\hat{\theta}$ is not a consitent estimator (or alternatively, does not converge in probability to θ).

Sometimes we are working with more than one random variable, and there are theorems that allows us to make stamtents about the convergence of these random variables:

Theorem: Let X_n converge in probability to θ and let Y_n converge in probability to δ, then

1. $X_n + Y_n$ converges in probability to $\theta + \delta$
2. $X_n \times Y_n$ converges in probability to $\theta \times \delta$
3. If $\delta \neq 0$, then X_n/Y_n converges in probability to θ/δ
4. If $g()$ is a real valued, continuous function (at θ), then $g(X_n)$ converges in probability to $g(\theta)$

2.1 Examples

Let $X_1, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$

1. We know that $\bar{X}$ is an unbiased estimator for μ. Show that $\bar{X}$ is a consistent estimator for μ

$$\begin{aligned}
X_1, ..., X_n & \quad \text{i.i.d.} \quad & N(0,1) \\
\Rightarrow \bar{X} & \quad \sim \quad & N(\mu, \frac{\sigma^2}{n}) \\
\Rightarrow V[\bar{X}] & \quad = \quad & \frac{\sigma^2}{n} \\
\Rightarrow \lim_{n\to\infty} V[\bar{X}] & \quad = \quad & \lim_{n\to\infty} \frac{\sigma^2}{n} \\
& \quad = \quad & 0
\end{aligned}$$

Thus, $\bar{X}$ is a consistent estimator for μ

2. Show that X_1 is not a consistent estimator for μ

$$\begin{aligned}
\Rightarrow X_1 & \quad \sim \quad & N(\mu, \sigma^2) \\
\Rightarrow V[X_1] & \quad = \quad & \sigma^2 \\
\Rightarrow \lim_{n\to\infty} V[\bar{X}] & \quad = \quad & \lim_{n\to\infty} \sigma^2 \\
& \quad = \quad & \sigma^2
\end{aligned}$$

3. We know that S_X^2 is an unbiased estimator for σ^2. Show that S_X^2 is a consistent estimator for σ^2

$$\begin{aligned}
X_1, ..., X_n & \quad \text{i.i.d.} \quad & N(\mu, \sigma^2) \\
\Rightarrow \frac{(n-1)}{\sigma^2} S_X^2 & \quad \sim \quad & \chi^2_{n-1} \\
\Rightarrow V[S_X^2] & \quad = \quad & \frac{\sigma^4}{(n-1)^2} V[\frac{(n-1)}{\sigma^2}] \\
& \quad = \quad & \frac{\sigma^4}{(n-1)^2} (\frac{n-1}{2})(2^2) \\
& \quad = \quad & \frac{2\sigma^4}{(n-1)} \\
\Rightarrow lim_{n\to\infty} V[S_X^2] & \quad = \quad & \lim_{n\to\infty} \frac{2\sigma^4}{(n-1)} \\
& \quad = \quad & 0
\end{aligned}$$

Thus, S_X^2 is a consistent estimator for σ^2

2.2 Exercises

Let $X_1, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$

1. Show that $\frac{(X_1-X_2)^2}{2}$ is an unbiased estimator for σ^2
2. Show that $\frac{(X_1-X_2)^2}{2}$ is not a consistent estimator for σ^2

2.3 Solutions

Let $X_1, ..., X_n$ be i.i.d. $N(\mu, \sigma^2)$

1. Show that $\frac{(X_1-X_2)^2}{2}$ is an unbiased estimator for σ^2

$$\begin{aligned} X_1, X_2 & \quad \text{i.i.d.} \quad & N(\mu, \sigma^2) \\ \Rightarrow X_1 - X_2 & \quad \sim \quad & N(0, 2\sigma^2) \\ \Rightarrow \frac{X_1 - X_2}{\sqrt{2}\sigma} & \sim N(0,1) & \\ \Rightarrow \frac{(X_1 - X_2)^2}{2\sigma^2} & \quad \sim \quad & \chi_1^2 \\ \Rightarrow E[\frac{(x_1 - X_2)^2}{2}] & \quad = \quad & \sigma^2 E[\frac{(x_1 - X_2)^2}{2\sigma^2}] \\ & \quad = \quad & \sigma^2(\frac{1}{2})(2) \\ & \quad = \quad & \sigma^2 \end{aligned}$$

2. Show that $\frac{(X_1-X_2)^2}{2}$ is not a consistent estimator for σ^2

$$\begin{aligned} X_1, X_2 & \quad \text{i.i.d.} \quad & N(\mu, \sigma^2) \\ \Rightarrow X_1 - X_2 & \quad \sim \quad & N(0, 2\sigma^2) \\ \Rightarrow \frac{X_1 - X_2}{\sqrt{2}\sigma} & \sim N(0,1) & \\ \Rightarrow \frac{(X_1 - X_2)^2}{2\sigma^2} & \quad \sim \quad & \chi_1^2 \\ \Rightarrow V[\frac{(X_1 - X_2)^2}{2}] & \quad = \quad & \sigma^4 V[\frac{(x_1 - X_2)^2}{2\sigma^2}] \\ & \quad = \quad & \sigma^2(\frac{1}{2})(2^2) \\ & \quad = \quad & 2\sigma^4 \\ \Rightarrow \lim_{n\to\infty} V[\frac{(X_1 - X_2)^2}{2}] & \quad = \quad & \lim_{n\to\infty} 2\sigma^4 \\ & \quad = \quad & 2\sigma^4 \end{aligned}$$

So, $\frac{(X_1-X_2)^2}{2}$ is not a consistent estimator for σ^2