Lecture 4: UMVUE and unbiased estimators of 0

Problems of approach 1.

Approach 1 has the following two main shortcomings.

- Even if Theorem 7.3.9 or its extension is applicable, there is no guarantee that the bound is sharp, i.e., there may be a UMVUE but it still cannot achieve the Cramér-Rao lower bound.
- The conditions for Theorem 7.3.9 is somewhat strong.

Example 7.3.13 (a case where Theorem 7.3.9 is not applicable)

Let $X_1,...,X_n$ be iid with from *uniform*$(0,\theta)$, where $\theta > 0$ is unknown.
The pdf of X_i is $f_\theta(x_i) = \theta^{-1} I(0 < x_i < \theta)$.
Since $P_\theta(0 < X_i < \theta) = 1$, we can focus on $0 < x_i < \theta$:

$$\log f_\theta(x_i) = -\log\theta, \qquad \frac{\partial}{\partial\theta}\log f_\theta(x_i) = -\frac{1}{\theta}, \qquad 0 < x_i < \theta$$

Then

$$E_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(X_i)\right]^2 = \frac{1}{\theta^2}$$

Consider the estimation of $g(\theta) = \theta$, $g'(\theta) = 1$.
According to Theorem 7.3.9 (if it holds), for any unbiased estimator $T(X)$ of θ, we should have

$$\mathrm{Var}_\theta(T) \geq \frac{[g'(\theta)]^2}{nE_\theta\left[\frac{\partial}{\partial\theta}\log f_\theta(X_1)\right]^2} = \frac{\theta^2}{n}$$

Let $X_{(n)}$ be the largest order statistic.
From the result in Chapter 5, the pdf of $X_{(n)}$ is

$$\frac{ny^{n-1}}{\theta^n}, \qquad 0 < y < \theta$$

Thus,

$$E_\theta(X_{(n)}) = \int_0^\theta \frac{ny^n}{\theta^n}dy = \frac{n}{\theta^n}\frac{y^{n+1}}{n+1}\bigg|_0^\theta = \frac{n\theta}{n+1}$$

showing that $\frac{n+1}{n}X_{(n)}$ is an unbiased estimator of θ.

$$E_\theta(X_{(n)}^2) = \int_0^\theta \frac{ny^{n+1}}{\theta^n}dy = \frac{n}{\theta^n}\frac{y^{n+2}}{n+2}\bigg|_0^\theta = \frac{n\theta^2}{n+2}$$

Then

$$\begin{aligned}
\mathrm{Var}_\theta\left(\frac{n+1}{n}X_{(n)}\right) &= \frac{(n+1)^2}{n^2}\mathrm{Var}_\theta(X_{(n)}) \\
&= \frac{(n+1)^2}{n^2}\left[E_\theta(X_{(n)}^2) - \{E_\theta(X_{(n)})\}^2\right] \\
&= \frac{(n+1)^2}{n^2}\left[\frac{n\theta^2}{n+2} - \left(\frac{n\theta}{n+1}\right)^2\right] \\
&= \frac{\theta^2}{n(n+2)} < \frac{\theta^2}{n}
\end{aligned}$$

Hence, Theorem 7.3.9 does not apply.

What is wrong?

Note that the key condition for Theorem 7.3.9 is that

$$\frac{\partial}{\partial\theta}E_\theta(T) = \int_{\mathscr{X}} T(x)\frac{\partial f_\theta(x)}{\partial\theta}dx \qquad \theta \in \Theta$$

The right hand side is

$$\int_0^\theta \cdots \int_0^\theta T(x)\frac{\partial}{\partial\theta}\left(\frac{1}{\theta^n}\right)dx_1\cdots dx_n = -\frac{n}{\theta^{n+1}}\int_0^\theta \cdots \int_0^\theta T(x)dx_1\cdots dx_n$$

The left hand side is

$$\frac{\partial}{\partial\theta}\left[\frac{1}{\theta^n}\int_0^\theta \cdots \int_0^\theta T(x)dx_1\cdots dx_n\right] = \frac{T(\theta,...,\theta)}{\theta^n} - \frac{n}{\theta^{n+1}}\int_0^\theta \cdots \int_0^\theta T(x)dx_1\cdots dx_n$$

which is not the same as the right hand side unless $T(\theta,...,\theta) = 0$ for all θ.

The problem is the support of $f_\theta(x)$, $\{x : f_\theta(x) > 0\}$, depends on θ.

When this occurs, Theorem 7.3.9 is typically not applicable.

Approach 2: using unbiased estimators of 0

To see when an unbiased estimator is best unbiased, we might ask how could we improve upon a given unbiased estimator?

Suppose that $T(X)$ is unbiased for $g(\theta)$ and $U(X)$ is a statistic satisfying $E_\theta(U) = 0$ for all θ, i.e., U is unbiased for 0.

Then, for any constant a,

$$T(X) + aU(X)$$

is unbiased for $g(\theta)$.

Can it be better than $T(X)$?

$$\mathrm{Var}_\theta(T + aU) = \mathrm{Var}_\theta(T) + 2a\mathrm{Cov}_\theta(T, U) + a^2\mathrm{Var}_\theta(U)$$

If for some θ_0, $\mathrm{Cov}_{\theta_0}(T, U) < 0$, then we can make

$$2a\mathrm{Cov}_{\theta_0}(T, U) + a^2\mathrm{Var}_{\theta_0}(U) < 0$$

by choosing $0 < a - 2\mathrm{Cov}_{\theta_0}(T, U)/\mathrm{Var}_{\theta_0}(U)$.

Hence, $T(X) + aU(X)$ is better than $T(X)$ at least when $\theta = \theta_0$ and $T(X)$ cannot be UMVUE.

Similarly, if $\mathrm{Cov}_{\theta_0}(T, U) > 0$ for some θ_0, then $T(X)$ cannot be UMVUE either.

Thus, $\mathrm{Cov}_\theta(T, U) = 0$ is necessary for $T(X)$ to be a UMVUE, for all unbiased estimators of 0.

It turns out that $\mathrm{Cov}_\theta(T, U) = 0$ for all $U(X)$ unbiased for 0 is also sufficient for $T(X)$ being a UMVUE.

Theorem 7.3.20.

An unbiased estimator $T(X)$ of $g(\theta)$ is UMVUE iff $T(X)$ is uncorrelated with all unbiased estimators of 0.

Proof.

We have shown the necessity (the only if part).
We now show the sufficiency (the if part).
Let $W(X)$ be another unbiased estimator of $g(\theta)$.
Then $W(X) - T(X)$ is unbiased for 0 and

$$\begin{aligned}
\mathrm{Var}_\theta(W) &= \mathrm{Var}_\theta(T + (W - T)) \\
&= \mathrm{Var}_\theta(T) + \mathrm{Var}_\theta(W - T) + 2\mathrm{Cov}_\theta(T, W - T) \\
&= \mathrm{Var}_\theta(T) + \mathrm{Var}_\theta(W - T) \\
&\geq \mathrm{Var}_\theta(T)
\end{aligned}$$

The result follows since W is arbitrary.

An unbiased estimator of 0 is a first-order ancillary statistic and can be treated as a random noise.
It makes sense that the most sensible way to estimate 0 is with 0, not with a random noise.
Theorem 7.3.20 says that an estimator is correlated with a random noise, then it can be improved.

The following two useful results are consequence of Theorem 7.3.20.

Corollary

Let T_j be a UMVUE of $g_j(\theta)$, $j = 1, ..., m$, where m is a fixed positive integer.

Then $T = \sum_{j=1}^m c_j T_j$ is a UMVUE of $g(\theta) = \sum_{j=1}^m c_j g_j(\theta)$ for any constants $c_1, ..., c_m$.

Proof.

Let $U(X)$ be any unbiased estimator of 0.
First, the unbiasedness of T_j's imply that T is unbiased for $g(\theta)$:

$$E_\theta(T) = E_\theta\left(\sum_{j=1}^m c_j T_j\right) = \sum_{j=1}^m E_\theta(T_j) = \sum_{j=1}^m c_j g_j(\theta) = g(\theta)$$

By Theorem 7.3.20, $\mathrm{Cov}_\theta(T_j, U) = 0$ for all $j = 1, ..., m$, and $\theta \in \Theta$.

$$\mathrm{Cov}_\theta(T, U) = \mathrm{Cov}_\theta\left(\sum_{j=1}^m c_j T_j, U\right) = \sum_{j=1}^m c_j \mathrm{Cov}_\theta(T_j, U) = 0$$

Again, by Theorem 7.3.20, this shows that T is a UMVUE of $g(\theta)$.

Theorem 7.3.19.

If $T(X)$ is a UMVUE of $g(\theta)$, then $T(X)$ is unique in the sense that if $T_1(X)$ is another UMVUE of $g(\theta)$, then

$$P_\theta(T(X) = T_1(X)) = 1 \qquad \text{for any } \theta \in \Theta$$

Proof.

Let $T(X)$ and $T_1(X)$ be both UMVUE of $g(\theta)$.
Since both $T(X)$ and $T_1(X)$ are unbiased, $T(X) - T_1(X)$ is an unbiased estimator of 0.
Since both T and T_1 are UMVUE, by Theorem 7.3.20,

$$\mathrm{Cov}_\theta(T, T - T_1) = 0, \qquad \mathrm{Cov}_\theta(T_1, T - T_1) = 0 \qquad \theta \in \Theta$$

Then

$$\begin{aligned} E_\theta(T - T_1)^2 &= E_\theta[(T - T_1)(T - T_1)] \\ &= E_\theta[T(T - T_1)] - E_\theta[T_1(T - T_1)] \\ &= \mathrm{Cov}_\theta(T, T - T_1) - \mathrm{Cov}_\theta(T_1, T - T_1) \\ &= 0 \end{aligned}$$

Hence, $P_\theta(T(X) = T_1(X)) = 1$ for any $\theta \in \Theta$.

Although Theorem 7.3.20 provides an interesting characterization of best unbiased estimators and two useful results, its usefulness of checking whether a particular unbiased estimator T is a UMVUE is limited since it asks to check whether T is uncorrelated with all unbiased estimator of 0.

However, Theorem 7.3.20 is sometimes useful in determining that an estimator is not best unbiased or showing that no UMVUE exists.

Example 7.3.21.

Let X be an observation from *uniform*$(\theta, \theta+1)$, $\theta \in \mathscr{R}$.

We want to show that there is no UMVUE of $g(\theta)$ for any nonconstant differentiable function g.

Note that an unbiased estimator $U(X)$ of 0 must satisfy

$$\int_{\theta}^{\theta+1} U(x)dx = 0 \qquad \text{for all } \theta \in \mathscr{R}$$

Differentiating both sizes of the previous equation, $U(\theta) = U(\theta+1)$, $\theta \in \mathscr{R}$, i.e.,

$$U(x) = U(x+1), \qquad x \in \mathscr{R}$$

If T is a UMVUE of $g(\theta)$, then $T(X)U(X)$ is unbiased for 0 and, hence,

$$T(x)U(x) = T(x+1)U(x+1), \qquad x \in \mathscr{R}$$

where $U(X)$ is any unbiased estimator of 0.
Since this is true for all U,

$$T(x) = T(x+1), \qquad x \in \mathscr{R}$$

Since T is unbiased for $g(\theta)$,

$$g(\theta) = \int_{\theta}^{\theta+1} T(x)dx \qquad \text{for all } \theta \in \mathscr{R}$$

Differentiating both sides of the previous equation we obtain that

$$g'(\theta) = T(\theta+1) - T(\theta) = 0, \qquad \theta \in \mathscr{R}$$

Thus, g is a constant function.

We consider more for $g(\theta) = \theta$.
Since $E_\theta(X) = \theta + \frac{1}{2}$, $X - \frac{1}{2}$ is unbiased for θ, and its variance is $\frac{1}{12}$.
One unbiased estimator of 0 is $U = \sin(2\pi X)$, i.e.,

$$E_\theta(U) = \int_{\theta}^{\theta+1} \sin(2\pi x)dx = 0 \qquad \theta \in \mathscr{R}$$

$$\begin{aligned}
\mathrm{Cov}_\theta(X - \tfrac{1}{2}, U) &= \mathrm{Cov}_\theta(X, \sin(2\pi X)) = E_\theta[X \sin(2\pi X)] \\
&= \int_\theta^{\theta+1} x \sin(2\pi x) dx \\
&= -\frac{x\cos(2\pi x)}{2\pi}\Big|_\theta^{\theta+1} + \int_\theta^{\theta+1} \frac{\cos(2\pi x)}{2\pi} dx \\
&= -\frac{\cos(2\pi\theta)}{2\pi}
\end{aligned}$$

Hence $X - \frac{1}{2}$ is correlated with an unbiased estimator of 0, and cannot be a UMVUE of θ.

In fact, $X - \frac{1}{2} + \frac{\sin(2\pi X)}{2\pi}$ is unbiased for θ and has variance $0.071 < \frac{1}{12}$.

Sufficient statistics

If there is a sufficient statistic T, then the Rao-Blackwell theorem says that $E(W|T)$ is a better unbiased estimator than an unbiased estimator W.

The following is a Rao-Blackwell theorem for unbiased estimators, a special case of what we have shown in Chapter 6.

Theorem 7.3.17.

Let W be any unbiased estimator of $g(\theta)$ and T be a sufficient statistic for θ. Define $\psi(T) = E(W|T)$, which does not depend on θ since T is sufficient. Then $\psi(T)$ is an unbiased estimator of $g(\theta)$ and $\mathrm{Var}_\theta(T) \leq \mathrm{Var}_\theta(W)$.

Thus, conditioning any unbiased estimator on a sufficient statistic will result in a uniform improvement.
This and the next result indicate that we need consider only estimators that are functions of a sufficient statistic in our search for the UMVUE.

An alternative to Theorem 7.3.20.

Suppose that T is a sufficient statistic for θ. An unbiased estimator $\psi(T)$ of $g(\theta)$ is a UMVUE iff $\mathrm{Cov}_\theta(\psi(T), h(T)) = 0$ for any $\theta \in \Theta$ and any $h(T)$ that is an unbiased estimator of 0.

Proof.

The “only if" part follows from the “only if" part of Theorem 7.3.20.

We now prove the “if" part.

Suppose that $\mathrm{Cov}_\theta(\psi(T), h(T)) = 0$ for any $\theta \in \Theta$ and any $h(T)$ that is an unbiased estimator of 0.
For any unbiased estimator of 0 U and any θ,

$$\mathrm{Cov}_\theta(\psi(T), U) = E_\theta[\psi(T)U] = E_\theta\{E[\psi(T)U|T]\} = E_\theta\{\psi(T)E(U|T)\} = 0$$

since $E(U|T)$ is a function of T and $E_\theta[E(U|T)] = E_\theta(U) = 0$.

If we have another sufficient statistic S, should we consider $E[E(W|T)|S]$?
If there is a function h such that $S = h(T)$, then by the properties of conditional expectation,

$$E[E(W|T)|S] = E(W|S) = E[E(W|S)|T]$$

That is, we should always conditioning on a simpler sufficient statistic, such as a minimal sufficient statistic or a complete sufficient statistic.
If we do have a sufficient and complete statistic T, then by the completeness, 0 is essentially the only unbiased estimator of 0 that is a function of T.

Thus, by the alternative to Theorem 7.3.20, we have actually proved the following result useful for finding the UMVUE.

Theorem 7.3.23 (Lehmann-Scheffé Theorem)

Let T be a complete sufficient statistic for θ. If $\psi(T)$ is an unbiased estimator of $g(\theta)$, then it is the unique UMVUE.

By Example 7.3.21, Theorem 7.3.23 does not hold if T is only minimal sufficient, since X is minimal sufficient for θ in Example 7.3.21.
This is because functions of minimal sufficient statistic can be first-order ancillary and can still be useful.

Example 7.3.13.

Let $X_1,...,X_n$ be iid from *uniform*$(0,\theta)$ with unknown $\theta > 0$.
Consider the estimation of θ.
Previously we show that approach 1 is not applicable in this example.
Since $T = X_{(n)}$ is the sufficient and complete statistic for θ and

$$E(X_{(n)}) = n\theta/(n+1)$$

The UMVUE of θ is $(1+n^{-1})X_{(n)}$.

Suppose now that $\Theta = [1, \infty)$.
Then $X_{(n)}$ is not complete, although it is still sufficient for θ.
Thus, Theorem 7.3.23 does not apply to $X_{(n)}$.
We now illustrate how to use the alternative Theorem 7.3.20 to find a UMVUE of θ.
Let $U(X_{(n)})$ be an unbiased estimator of 0.
Since $X_{(n)}$ has pdf $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$,

$$0 = \int_0^1 U(x)x^{n-1}dx + \int_1^\theta U(x)x^{n-1}dx \quad \text{for all } \theta \geq 1.$$

This implies that $U(x) = 0$ on $[1, \infty)$ and

$$\int_0^1 U(x)x^{n-1}dx = 0.$$

Consider $T = h(X_{(n)})$.
To have $E(TU) = 0$, we must have

$$\int_0^1 h(x)U(x)x^{n-1}dx = 0.$$

Thus, we may consider the following function:

$$h(x) = \begin{cases} c & 0 \le x \le 1 \\ bx & x > 1, \end{cases}$$

where c and b are some constants.
From the previous discussion,

$$E[h(X_{(n)})U(X_{(n)})] = 0, \qquad \theta \ge 1.$$

Since $E[h(X_{(n)})] = \theta$, we obtain that

$$\begin{aligned} \theta &= cP(X_{(n)} \le 1) + bE[X_{(n)} I_{(1,\infty)}(X_{(n)})] \\ &= c\theta^{-n} + [bn/(n+1)](\theta - \theta^{-n}). \end{aligned}$$

Thus, $c = 1$ and $b = (n+1)/n$.
The UMVUE of θ is then

$$h(X_{(n)}) = \begin{cases} 1 & 0 \le X_{(n)} \le 1 \\ (1+n^{-1})X_{(n)} & X_{(n)} > 1. \end{cases}$$

This estimator is better than $(1+n^{-1})X_{(n)}$, which is the UMVUE when $\Theta = (0,\infty)$ and does not make use of the information about $\theta \ge 1$.
When $\Theta = (0,\infty)$, this estimator is not unbiased.