Lecture 09: Sufficiency, Completeness and Unbiasedness

Mathematical Statistics II, MATH 60062/70062

Thursday February 17, 2022

Reference: Casella & Berger, 7.3.3

Recap: Uniformly minimum-variance unbiased estimator

An estimator $W^* = W^*(\boldsymbol{X})$ is a **uniformly minimum-variance unbiased estimator (UMVUE)** of $\tau(\theta)$ if

1. $E_\theta(W^*) = \tau(\theta)$ for all $\theta \in \Theta$.
2. $\mathrm{Var}_\theta(W^*) \le \mathrm{Var}_\theta(W)$ for all $\theta \in \Theta$, where W is any other unbiased estimator of $\tau(\theta)$.

We will discuss two approaches to find UMVUEs:

1. (**Cramér-Rao Inequality**) Determine a **lower bound** on the variance of *any* unbiased estimator of $\tau(\theta)$. If we can find an unbiased estimator whose variance attains this lower bound, we have found the UMVUE.
2. (**Rao-Blackwell Theorem**) Relate the property of UMVUEs with the notation of **sufficiency** and **completeness**.

Rao-Blackwell Theorem

Let $W = W(\boldsymbol{X})$ be an unbiased estimator of $\tau(\theta)$, and let $T = T(\boldsymbol{X})$ be a sufficient statistic for θ. Define

$$\phi(T) = E(W \mid T).$$

Then

1. $E_\theta[\phi(T)] = \tau(\theta)$ for all $\theta \in \Theta$
2. $\mathrm{Var}_\theta(\phi(T)) \leq \mathrm{Var}_\theta(W)$ for all $\theta \in \Theta$.

That is, $\phi(T) = E(W \mid T)$ is a uniformly better unbiased estimator than W.

Rao-Blackwell Theorem

Let $W = W(\boldsymbol{X})$ be an unbiased estimator of $\tau(\theta)$, and let $T = T(\boldsymbol{X})$ be a sufficient statistic for θ. Define

$$\phi(T) = E(W \mid T).$$

Then

1. $E_\theta[\phi(T)] = \tau(\theta)$ for all $\theta \in \Theta$
2. $\mathrm{Var}_\theta(\phi(T)) \leq \mathrm{Var}_\theta(W)$ for all $\theta \in \Theta$.

That is, $\phi(T) = E(W \mid T)$ is a uniformly better unbiased estimator than W.

We can improve an unbiased estimator by conditioning on a sufficient statistic; in the search for UMVUE, we can restrict attention to those estimators that are functions of a sufficient statistic.

Proof of the Rao-Blackwell Theorem

By the law of total expectation, we know $\phi(T) = E(W \mid T)$ is an unbiased for $\tau(\theta)$,

$$E_\theta[\phi(T)] = E_\theta[E(W \mid T)] = E_\theta(W) = \tau(\theta).$$

By the law of total variance,

$$\begin{aligned} \mathrm{Var}_\theta(W) &= E_\theta[\mathrm{Var}(W \mid T)] + \mathrm{Var}_\theta[E(W \mid T)] \\ &= E_\theta[\mathrm{Var}(W \mid T)] + \mathrm{Var}_\theta[\phi(T)] \\ &\geq \mathrm{Var}_\theta[\phi(T)]. \end{aligned}$$

Proof of the Rao-Blackwell Theorem

By the law of total expectation, we know $\phi(T) = E(W \mid T)$ is an unbiased for $\tau(\theta)$,

$$E_\theta[\phi(T)] = E_\theta[E(W \mid T)] = E_\theta(W) = \tau(\theta).$$

By the law of total variance,

$$\begin{aligned}
\mathrm{Var}_\theta(W) &= E_\theta[\mathrm{Var}(W \mid T)] + \mathrm{Var}_\theta[E(W \mid T)] \\
&= E_\theta[\mathrm{Var}(W \mid T)] + \mathrm{Var}_\theta[\phi(T)] \\
&\geq \mathrm{Var}_\theta[\phi(T)].
\end{aligned}$$

There was no mention of sufficiency in the proof. It might seem that conditioning on any statistic will result in an improvement.

Proof of the Rao-Blackwell Theorem

By the law of total expectation, we know $\phi(T) = E(W \mid T)$ is an unbiased for $\tau(\theta)$,

$$E_\theta[\phi(T)] = E_\theta[E(W \mid T)] = E_\theta(W) = \tau(\theta).$$

By the law of total variance,

$$\begin{aligned}
\mathrm{Var}_\theta(W) &= E_\theta[\mathrm{Var}(W \mid T)] + \mathrm{Var}_\theta[E(W \mid T)] \\
&= E_\theta[\mathrm{Var}(W \mid T)] + \mathrm{Var}_\theta[\phi(T)] \\
&\geq \mathrm{Var}_\theta[\phi(T)].
\end{aligned}$$

There was no mention of sufficiency in the proof. It might seem that conditioning on any statistic will result in an improvement.

However, if T is not sufficient, $\phi(T) = E(W \mid T)$ may not be an estimator.

Conditioning on an insufficient statistic

Let X_1, X_2 be iid $\mathcal{N}(\theta, 1)$. The statistic $\bar{X} = (X_1 + X_2)/2$ is an unbiased estimator of θ,

$$E_\theta[\bar{X}] = \theta, \quad \mathrm{Var}_\theta[\bar{X}] = \frac{1}{2}.$$

Consider conditioning on X_1, which is not sufficient. Let $\phi(X_1) = E(\bar{X} \mid X_1)$, $E_\theta[\phi(X_1)] = \theta$ and $\mathrm{Var}_\theta[\phi(X_1)] \leq \mathrm{Var}_\theta[\bar{X}]$.

Conditioning on an insufficient statistic

Let X_1, X_2 be iid $\mathcal{N}(\theta, 1)$. The statistic $\bar{X} = (X_1 + X_2)/2$ is an unbiased estimator of θ,

$$E_\theta[\bar{X}] = \theta, \quad \mathrm{Var}_\theta[\bar{X}] = \frac{1}{2}.$$

Consider conditioning on X_1, which is not sufficient. Let $\phi(X_1) = E(\bar{X} \mid X_1)$, $E_\theta[\phi(X_1)] = \theta$ and $\mathrm{Var}_\theta[\phi(X_1)] \leq \mathrm{Var}_\theta[\bar{X}]$.

However, $\phi(X_1)$ is *not* an estimator since

$$\begin{aligned}
\phi(X_1) &= E(\bar{X} \mid X_1) \\
&= \frac{1}{2}E(X_1 \mid X_1) + \frac{1}{2}E(X_2 \mid X_1) \\
&= \frac{1}{2}X_1 + \frac{1}{2}\theta
\end{aligned}$$

depends on θ.

Uniqueness of UMVUE

If W is an UMVUE of $\tau(\theta)$, then W is unique.

Uniqueness of UMVUE

If W is an UMVUE of $\tau(\theta)$, then W is unique.

Suppose that W' is another UMVUE. It suffices to show that $W = W'$. Consider an estimator

$$W^* = \frac{1}{2}(W + W').$$

Note that W^* is an unbiased estimator

$$E_\theta(W^*) = \frac{1}{2}[E_\theta(W) + E_\theta(W')] = \tau(\theta), \quad \forall \theta \in \Theta.$$

The variance of W^* is

$$\begin{aligned}\mathrm{Var}_\theta(W^*) &= \mathrm{Var}_\theta\left[\frac{1}{2}(W + W')\right] \\ &= \frac{1}{4}\mathrm{Var}_\theta(W) + \frac{1}{4}\mathrm{Var}_\theta(W') + \frac{1}{2}\mathrm{Cov}_\theta(W, W') \\ &\leq \frac{1}{4}\mathrm{Var}_\theta(W) + \frac{1}{4}\mathrm{Var}_\theta(W') + \frac{1}{2}\left[\mathrm{Var}_\theta(W)\mathrm{Var}_\theta(W')\right]^{1/2} \\ &= \mathrm{Var}_\theta(W),\end{aligned}$$

where the inequality is an application of Cauchy-Schwartz and the last equality holds because $\mathrm{Var}_\theta(W) = \mathrm{Var}_\theta(W')$ by assumption.

The variance of W^* is

$$\begin{aligned}
\mathrm{Var}_\theta(W^*) &= \mathrm{Var}_\theta\left[\frac{1}{2}(W + W')\right] \\
&= \frac{1}{4}\mathrm{Var}_\theta(W) + \frac{1}{4}\mathrm{Var}_\theta(W') + \frac{1}{2}\mathrm{Cov}_\theta(W, W') \\
&\le \frac{1}{4}\mathrm{Var}_\theta(W) + \frac{1}{4}\mathrm{Var}_\theta(W') + \frac{1}{2}\left[\mathrm{Var}_\theta(W)\mathrm{Var}_\theta(W')\right]^{1/2} \\
&= \mathrm{Var}_\theta(W),
\end{aligned}$$

where the inequality is an application of Cauchy-Schwartz and the last equality holds because $\mathrm{Var}_\theta(W) = \mathrm{Var}_\theta(W')$ by assumption.

The inequality cannot be strict because W is UMVUE by assumption. So, it must be true that $\mathrm{Var}_\theta(W^*) = \mathrm{Var}_\theta(W)$ and (because the equality holds)

$$W' = a(\theta)W + b(\theta).$$

Using properties of covariance, we have

$$\mathrm{Cov}_\theta(W, W') = \mathrm{Cov}_\theta[W, a(\theta)W + b(\theta)] = a(\theta)\mathrm{Var}_\theta(W),$$

and we have just shown that $\mathrm{Cov}_\theta(W, W') = \mathrm{Var}_\theta(W)$. Hence, $a(\theta) = 1$.

Finally, since both W and W' are unbiased and

$$E_\theta(W') = E_\theta[a(\theta)W + b(\theta)] = E_\theta(W) + b(\theta),$$

we must have $b(\theta) = 0$.

Therefore, $W = W'$. W is unique.

Unbiased estimators of 0

Suppose $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$. W is the UMVUE of $\tau(\theta)$ if and only if W is uncorrelated with all unbiased estimators of 0.

Unbiased estimators of 0

Suppose $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$. W is the UMVUE of $\tau(\theta)$ if and only if W is uncorrelated with all unbiased estimators of 0.

Necessity: Suppose that $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$ and W is the UMVUE of $\tau(\theta)$. Suppose that U is an unbiased estimator of 0, i.e., $E_\theta(U) = 0$. *It suffices to show that* $\mathrm{Cov}_\theta(W, U) = 0$ *for all* $\theta \in \Theta$.

Unbiased estimators of 0

Suppose $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$. W is the UMVUE of $\tau(\theta)$ if and only if W is uncorrelated with all unbiased estimators of 0.

Necessity: Suppose that $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$ and W is the UMVUE of $\tau(\theta)$. Suppose that U is an unbiased estimator of 0, i.e., $E_\theta(U) = 0$. *It suffices to show that* $\mathrm{Cov}_\theta(W, U) = 0$ *for all* $\theta \in \Theta$*.* Consider an estimator

$$\phi_a = W + aU,$$

where a is a constant. The estimator is unbiased for $\tau(\theta)$ since

$$E_\theta(\phi_a) = E_\theta(W + aU) = E_\theta(W) + aE_\theta(U) = \tau(\theta).$$

The variance of ϕ_a is

$$\begin{aligned}\mathrm{Var}_\theta(\phi_a) &= \mathrm{Var}_\theta(W + aU)\\ &= \mathrm{Var}_\theta(W) + 2a\mathrm{Cov}_\theta(W, U) + a^2\mathrm{Var}_\theta(U),\end{aligned}$$

which is minimized when

$$a = -\frac{\mathrm{Cov}_\theta(W, U)}{\mathrm{Var}_\theta(U)}.$$

The best unbiased estimator in this class is

$$W - \frac{\mathrm{Cov}_\theta(W, U)}{\mathrm{Var}_\theta(U)}U,$$

which cannot be W if $\mathrm{Cov}_\theta(W, U) \neq 0$. Therefore, we must have

$$\mathrm{Cov}_\theta(W, U) = 0.$$

Sufficiency: Suppose $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$ and W is uncorrelated with any unbiased estimators of 0. That is $\mathrm{Cov}_\theta(W, U) = 0$ for any U that satisfies $E_\theta(U) = 0$, for all $\theta \in \Theta$. Let W' be any other unbiased estimator of $\tau(\theta)$. *It suffices to show that* $\mathrm{Var}_\theta(W) \leq \mathrm{Var}_\theta(W')$*.*

Sufficiency: Suppose $E_\theta(W) = \tau(\theta)$ for all $\theta \in \Theta$ and W is uncorrelated with any unbiased estimators of 0. That is $\mathrm{Cov}_\theta(W, U) = 0$ for any U that satisfies $E_\theta(U) = 0$, for all $\theta \in \Theta$. Let W' be any other unbiased estimator of $\tau(\theta)$. *It suffices to show that* $\mathrm{Var}_\theta(W) \leq \mathrm{Var}_\theta(W')$*.*

Write

$$W' = W + (W' - W),$$

and calculate

$$\mathrm{Var}_\theta(W') = \mathrm{Var}_\theta(W) + \mathrm{Var}_\theta(W' - W) + 2\mathrm{Cov}_\theta(W, W' - W),$$

where the covariance $\mathrm{Cov}_\theta(W, W' - W) = 0$ because $W' - W$ is an unbiased estimator of 0. Therefore,

$$\mathrm{Var}_\theta(W') = \mathrm{Var}_\theta(W) + \underbrace{\mathrm{Var}_\theta(W' - W)}_{\geq 0} \geq \mathrm{Var}_\theta(W).$$

Unbiased estimators of 0

Characterizing the unbiased estimators of 0 is not an easy task, and requires conditions on the PDFs/PMFs with which we are working.

If a family of PDFs/PMFs has the property that there are no unbiased estimators of 0 (other than zero function itself), then our search would be ended, since any statistic W satisfies $\mathrm{Cov}_\theta(W, 0) = 0$.

Recall: Complete statistic

Let $\{f_T(t \mid \theta); \theta \in \Theta\}$ be a family of PDFs (or PMFs) for a statistic $T = T(\boldsymbol{X})$. The family is called **complete** if the following condition holds:

$$E_\theta(g(T)) = 0, \ \forall \theta \in \Theta \implies P_\theta(g(T) = 0) = 1, \ \forall \theta \in \Theta.$$

In other words, $g(T) = 0$ almost surely for all $\theta \in \Theta$. Equivalently, $T(\boldsymbol{X})$ is called a **complete statistic**.

This means, *the only function of T that is an unbiased estimator of zero is the function that is zero itself (with probability 1)*.

Sufficiency, completeness and best unbiasedness

Let T be a complete sufficient statistic for a parameter θ, and let $\phi(T)$ be any estimator based only on T. Then $\phi(T)$ is the unique UMVUE of its expected value.

To find the UMVUE for $\tau(\theta)$.

1. Find a complete sufficient statistic T for θ
2. Find a function of T, say $\phi(T)$, that satisfies

$$E_\theta[\phi(T)] = \tau(\theta).$$

Then $\phi(T)$ is the UMVUE for $\tau(\theta)$.

Poisson UMVUE

Suppose that $X_1, \ldots, X_n$ are iid $\mathrm{Pois}(\theta)$, where $\theta > 0$.

Poisson UMVUE

Suppose that $X_1, \ldots, X_n$ are iid $\mathrm{Pois}(\theta)$, where $\theta > 0$.

The Poisson PMF is a member of the Exponential family since

$$
\begin{aligned}
f_X(x \mid \theta) &= \frac{e^{-\theta}\theta^x}{x!} I(x = 0, 1, 2, \ldots) \\
&= \frac{I(x = 0, 1, 2, \ldots)}{x!} e^{-\theta} \exp[\log \theta \cdot x] \\
&= h(x)c(\theta) \exp[w_1(\theta)t_1(x)].
\end{aligned}
$$

Therefore,

$$
T = T(\boldsymbol{X}) = \sum_{i=1}^{n} t_1(X_i) = \sum_{i=1}^{n} X_i
$$

is a sufficient statistic. Since $d = k = 1$, T is complete.

The expectation of T is

$$E_\theta(T) = E_\theta\left(\sum_{i=1}^{n} X_i\right) = n\theta.$$

So, $\bar{X} = T/n$ is an unbiased estimator of θ.

Since $\bar{X}$ unbiased and is a function of T, which is a complete and sufficient statistic, we know that $\bar{X}$ is UMVUE for θ.

Gamma UMVUE

Suppose that $X_1, \ldots, X_n$ are iid $\mathrm{Gamma}(\alpha_0, \beta)$, where α_0 is known and $\beta > 0$.

Gamma UMVUE

Suppose that $X_1, \ldots, X_n$ are iid $\text{Gamma}(\alpha_0, \beta)$, where α_0 is known and $\beta > 0$.

The $\text{Gamma}(\alpha_0, \beta)$ PDF is part of the full Exponential family

$$\begin{aligned} f_X(x \mid \beta) &= \frac{1}{\Gamma(\alpha_0)\beta^{\alpha_0}} x^{\alpha_0 - 1} e^{-x/\beta} \\ &= \frac{x^{\alpha_0 - 1} I(x > 0)}{\Gamma(\alpha_0)} \frac{1}{\beta^{\alpha_0}} \exp[(-1/\beta)x] \\ &= h(x)c(\beta)\exp[w_1(\beta)t_1(x)]. \end{aligned}$$

Therefore,

$$T = T(\boldsymbol{X}) = \sum_{i=1}^{n} t_1(X_i) = \sum_{i=1}^{n} X_i$$

is a complete and sufficient statistic.

The expectation of T is

$$E_\beta(T) = E_\beta\left(\sum_{i=1}^{n} X_i\right) = n\alpha_0\beta.$$

So,

$$\phi(T) = \frac{T}{n\alpha_0}$$

is an unbiased estimator of β.

Since $\phi(T)$ unbiased and is a function of T, which is a complete and sufficient statistic, we know that $\phi(T)$ is UMVUE for β.

The expectation of T is

$$E_\beta(T) = E_\beta\left(\sum_{i=1}^{n} X_i\right) = n\alpha_0\beta.$$

So,

$$\phi(T) = \frac{T}{n\alpha_0}$$

is an unbiased estimator of β.

Since $\phi(T)$ unbiased and is a function of T, which is a complete and sufficient statistic, we know that $\phi(T)$ is UMVUE for β.

What if we want to find UMVUE for $1/\beta$?