Stat 709: Mathematical Statistics
Lecture 31

Jun Shao

Department of Statistics
University of Wisconsin
Madison, WI 53706, USA

Lecture 31: UMVUE: a necessary and sufficient condition

When a complete and sufficient statistic is not available, it is usually very difficult to derive a UMVUE.
In some cases, the following result can be applied, if we have enough knowledge about unbiased estimators of 0.

Theorem 3.2

Let $\mathscr{U}$ be the set of all unbiased estimators of 0 with finite variances and T be an unbiased estimator of ϑ with $E(T^2) < \infty$.

(i) A necessary and sufficient condition for $T(X)$ to be a UMVUE of ϑ is that $E[T(X)U(X)] = 0$ for any $U \in \mathscr{U}$ and any $P \in \mathscr{P}$.

(ii) Suppose that $T = h(\tilde{T})$, where $\tilde{T}$ is a sufficient statistic for $P \in \mathscr{P}$ and h is a Borel function.
Let $\mathscr{U}_{\tilde{T}}$ be the subset of $\mathscr{U}$ consisting of Borel functions of $\tilde{T}$.
Then a necessary and sufficient condition for T to be a UMVUE of ϑ is that $E[T(X)U(X)] = 0$ for any $U \in \mathscr{U}_{\tilde{T}}$ and any $P \in \mathscr{P}$.

Lecture 31: UMVUE: a necessary and sufficient condition

When a complete and sufficient statistic is not available, it is usually very difficult to derive a UMVUE.
In some cases, the following result can be applied, if we have enough knowledge about unbiased estimators of 0.

Theorem 3.2

Let $\mathscr{U}$ be the set of all unbiased estimators of 0 with finite variances and T be an unbiased estimator of ϑ with $E(T^2) < \infty$.

(i) A necessary and sufficient condition for $T(X)$ to be a UMVUE of ϑ is that $E[T(X)U(X)] = 0$ for any $U \in \mathscr{U}$ and any $P \in \mathscr{P}$.

(ii) Suppose that $T = h(\tilde{T})$, where $\tilde{T}$ is a sufficient statistic for $P \in \mathscr{P}$ and h is a Borel function.
Let $\mathscr{U}_{\tilde{T}}$ be the subset of $\mathscr{U}$ consisting of Borel functions of $\tilde{T}$.
Then a necessary and sufficient condition for T to be a UMVUE of ϑ is that $E[T(X)U(X)] = 0$ for any $U \in \mathscr{U}_{\tilde{T}}$ and any $P \in \mathscr{P}$.

Proof of Theorem 3.2(i)

Suppose that T is a UMVUE of ϑ.
Then $T_c = T + cU$, where $U \in \mathscr{U}$ and c is a fixed constant, is also unbiased for ϑ and, thus,

$$\mathrm{Var}(T_c) \geq \mathrm{Var}(T) \qquad c \in \mathscr{R},\ P \in \mathscr{P},$$

which is the same as

$$c^2\,\mathrm{Var}(U) + 2c\,\mathrm{Cov}(T,U) \geq 0 \qquad c \in \mathscr{R},\ P \in \mathscr{P}.$$

This is impossible unless $\mathrm{Cov}(T,U) = E(TU) = 0$ for any $P \in \mathscr{P}$.

Suppose now $E(TU) = 0$ for any $U \in \mathscr{U}$ and $P \in \mathscr{P}$.
Let T_0 be another unbiased estimator of ϑ with $\mathrm{Var}(T_0) < \infty$.
Then $T - T_0 \in \mathscr{U}$ and, hence,

$$E[T(T - T_0)] = 0 \qquad P \in \mathscr{P},$$

which with the fact that $ET = ET_0$ implies that

$$\mathrm{Var}(T) = \mathrm{Cov}(T,T_0) \qquad P \in \mathscr{P}.$$

Note that $[\mathrm{Cov}(T,T_0)]^2 \leq \mathrm{Var}(T)\,\mathrm{Var}(T_0)$.
Hence $\mathrm{Var}(T) \leq \mathrm{Var}(T_0)$ for any $P \in \mathscr{P}$.

Proof of Theorem 3.2(i)

Suppose that T is a UMVUE of ϑ.
Then $T_c = T + cU$, where $U \in \mathscr{U}$ and c is a fixed constant, is also unbiased for ϑ and, thus,

$$\mathrm{Var}(T_c) \geq \mathrm{Var}(T) \qquad c \in \mathscr{R},\ P \in \mathscr{P},$$

which is the same as

$$c^2\,\mathrm{Var}(U) + 2c\,\mathrm{Cov}(T,U) \geq 0 \qquad c \in \mathscr{R},\ P \in \mathscr{P}.$$

This is impossible unless $\mathrm{Cov}(T,U) = E(TU) = 0$ for any $P \in \mathscr{P}$.
Suppose now $E(TU) = 0$ for any $U \in \mathscr{U}$ and $P \in \mathscr{P}$.
Let T_0 be another unbiased estimator of ϑ with $\mathrm{Var}(T_0) < \infty$.
Then $T - T_0 \in \mathscr{U}$ and, hence,

$$E[T(T - T_0)] = 0 \qquad P \in \mathscr{P},$$

which with the fact that $ET = ET_0$ implies that

$$\mathrm{Var}(T) = \mathrm{Cov}(T,T_0) \qquad P \in \mathscr{P}.$$

Note that $[\mathrm{Cov}(T,T_0)]^2 \leq \mathrm{Var}(T)\,\mathrm{Var}(T_0)$.
Hence $\mathrm{Var}(T) \leq \mathrm{Var}(T_0)$ for any $P \in \mathscr{P}$.

Proof of Theorem 3.2(ii)

It suffices to show that $E(TU) = 0$ for any $U \in \mathscr{U}_{\tilde{T}}$ and $P \in \mathscr{P}$ implies that $E(TU) = 0$ for any $U \in \mathscr{U}$ and $P \in \mathscr{P}$

Let $U \in \mathscr{U}$.

Then $E(U|\tilde{T}) \in \mathscr{U}_{\tilde{T}}$ and the result follows from the fact that $T = h(\tilde{T})$ and

$$E(TU) = E[E(TU|\tilde{T})] = E[E(h(\tilde{T})U|\tilde{T})] = E[h(\tilde{T})E(U|\tilde{T})].$$

Theorem 3.2 can be used

- to find a UMVUE,
- to check whether a particular estimator is a UMVUE, and
- to show the nonexistence of any UMVUE.

If there is a sufficient statistic, then by Rao-Blackwell's theorem, we only need to focus on functions of the sufficient statistic and, hence, Theorem 3.2(ii) is more convenient to use.

Proof of Theorem 3.2(ii)

It suffices to show that $E(TU) = 0$ for any $U \in \mathscr{U}_{\tilde{T}}$ and $P \in \mathscr{P}$ implies that $E(TU) = 0$ for any $U \in \mathscr{U}$ and $P \in \mathscr{P}$

Let $U \in \mathscr{U}$.

Then $E(U|\tilde{T}) \in \mathscr{U}_{\tilde{T}}$ and the result follows from the fact that $T = h(\tilde{T})$ and

$$E(TU) = E[E(TU|\tilde{T})] = E[E(h(\tilde{T})U|\tilde{T})] = E[h(\tilde{T})E(U|\tilde{T})].$$

Theorem 3.2 can be used

- to find a UMVUE,
- to check whether a particular estimator is a UMVUE, and
- to show the nonexistence of any UMVUE.

If there is a sufficient statistic, then by Rao-Blackwell's theorem, we only need to focus on functions of the sufficient statistic and, hence, Theorem 3.2(ii) is more convenient to use.

Proof of Theorem 3.2(ii)

It suffices to show that $E(TU) = 0$ for any $U \in \mathscr{U}_{\tilde{T}}$ and $P \in \mathscr{P}$ implies that $E(TU) = 0$ for any $U \in \mathscr{U}$ and $P \in \mathscr{P}$
Let $U \in \mathscr{U}$.
Then $E(U|\tilde{T}) \in \mathscr{U}_{\tilde{T}}$ and the result follows from the fact that $T = h(\tilde{T})$ and

$$E(TU) = E[E(TU|\tilde{T})] = E[E(h(\tilde{T})U|\tilde{T})] = E[h(\tilde{T})E(U|\tilde{T})].$$

Theorem 3.2 can be used

- to find a UMVUE,
- to check whether a particular estimator is a UMVUE, and
- to show the nonexistence of any UMVUE.

If there is a sufficient statistic, then by Rao-Blackwell's theorem, we only need to focus on functions of the sufficient statistic and, hence, Theorem 3.2(ii) is more convenient to use.

As a consequence of Theorem 3.2, we have the following useful result.

Corollary 3.1

(i) Let T_j be a UMVUE of ϑ_j, $j = 1, ..., k$, where k is a fixed positive integer.
Then $\sum_{j=1}^{k} c_j T_j$ is a UMVUE of $\vartheta = \sum_{j=1}^{k} c_j \vartheta_j$ for any constants $c_1, ..., c_k$.

(ii) Let T_1 and T_2 be two UMVUE's of ϑ.
Then $T_1 = T_2$ a.s. P for any $P \in \mathscr{P}$.

Example 3.7

Let $X_1, ..., X_n$ be i.i.d. from the uniform distribution on the interval $(0, \theta)$.
In Example 3.1, $(1 + n^{-1})X_{(n)}$ is shown to be the UMVUE for θ when the parameter space is $\Theta = (0, \infty)$.
Suppose now that $\Theta = [1, \infty)$.
Then $X_{(n)}$ is not complete, although it is still sufficient for θ.
Thus, Theorem 3.1 does not apply to $X_{(n)}$.

As a consequence of Theorem 3.2, we have the following useful result.

Corollary 3.1

(i) Let T_j be a UMVUE of ϑ_j, $j = 1,...,k$, where k is a fixed positive integer.
Then $\sum_{j=1}^k c_j T_j$ is a UMVUE of $\vartheta = \sum_{j=1}^k c_j \vartheta_j$ for any constants $c_1,...,c_k$.

(ii) Let T_1 and T_2 be two UMVUE's of ϑ.
Then $T_1 = T_2$ a.s. P for any $P \in \mathscr{P}$.

Example 3.7

Let $X_1,...,X_n$ be i.i.d. from the uniform distribution on the interval $(0,\theta)$.
In Example 3.1, $(1+n^{-1})X_{(n)}$ is shown to be the UMVUE for θ when the parameter space is $\Theta = (0,\infty)$.
Suppose now that $\Theta = [1,\infty)$.
Then $X_{(n)}$ is not complete, although it is still sufficient for θ.
Thus, Theorem 3.1 does not apply to $X_{(n)}$.

As a consequence of Theorem 3.2, we have the following useful result.

Corollary 3.1

(i) Let T_j be a UMVUE of ϑ_j, $j = 1,...,k$, where k is a fixed positive integer.
Then $\sum_{j=1}^{k} c_j T_j$ is a UMVUE of $\vartheta = \sum_{j=1}^{k} c_j \vartheta_j$ for any constants $c_1,...,c_k$.

(ii) Let T_1 and T_2 be two UMVUE's of ϑ.
Then $T_1 = T_2$ a.s. P for any $P \in \mathscr{P}$.

Example 3.7

Let $X_1,...,X_n$ be i.i.d. from the uniform distribution on the interval $(0,\theta)$.
In Example 3.1, $(1+n^{-1})X_{(n)}$ is shown to be the UMVUE for θ when the parameter space is $\Theta = (0,\infty)$.
Suppose now that $\Theta = [1,\infty)$.
Then $X_{(n)}$ is not complete, although it is still sufficient for θ.
Thus, Theorem 3.1 does not apply to $X_{(n)}$.

Example 3.7 (continued)

We now illustrate how to use Theorem 3.2(ii) to find a UMVUE of θ.

Let $U(X_{(n)})$ be an unbiased estimator of 0.

Since $X_{(n)}$ has the Lebesgue p.d.f. $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$,

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

This implies that $U(x) = 0$ a.e. Lebesgue measure 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:

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

where c and b are some constants.

Example 3.7 (continued)

We now illustrate how to use Theorem 3.2(ii) to find a UMVUE of θ.
Let $U(X_{(n)})$ be an unbiased estimator of 0.
Since $X_{(n)}$ has the Lebesgue p.d.f. $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$,

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

This implies that $U(x) = 0$ a.e. Lebesgue measure 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:

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

where c and b are some constants.

Example 3.7 (continued)

We now illustrate how to use Theorem 3.2(ii) to find a UMVUE of θ.
Let $U(X_{(n)})$ be an unbiased estimator of 0.
Since $X_{(n)}$ has the Lebesgue p.d.f. $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$,

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

This implies that $U(x) = 0$ a.e. Lebesgue measure 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:

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

where c and b are some constants.

Example 3.7 (continued)

From the previous discussion,

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

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

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

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

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

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 $\theta \geq 1$.
When $\Theta = (0,\infty)$, this estimator is not unbiased.

In fact, $h(X_{(n)})$ is complete and sufficient for $\theta \in [1,\infty)$.

Example 3.7 (continued)

From the previous discussion,

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

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

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

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

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

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 $\theta \geq 1$.
When $\Theta = (0,\infty)$, this estimator is not unbiased.

In fact, $h(X_{(n)})$ is complete and sufficient for $\theta \in [1,\infty)$.

Example 3.7 (continued)

From the previous discussion,

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

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

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

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

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

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 $\theta \geq 1$.
When $\Theta = (0,\infty)$, this estimator is not unbiased.

In fact, $h(X_{(n)})$ is complete and sufficient for $\theta \in [1,\infty)$.

Example 3.7 (continued)

From the previous discussion,

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

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

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

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

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

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 $\theta \geq 1$.
When $\Theta = (0,\infty)$, this estimator is not unbiased.

In fact, $h(X_{(n)})$ is complete and sufficient for $\theta \in [1,\infty)$.

Example 3.7 (continued)

It suffices to show that

$$g(X_{(n)}) = \begin{cases} 1 & 0 \le X_{(n)} \le 1 \\ X_{(n)} & X_{(n)} > 1. \end{cases}$$

is complete and sufficient for $\theta \in [1, \infty)$.

The sufficiency follows from the fact that the joint p.d.f. of $X_1, ..., X_n$ is

$$\frac{1}{\theta^n} I_{(0,\theta)}(X_{(n)}) = \frac{1}{\theta^n} I_{(0,\theta)}(g(X_{(n)})).$$

If $E[f(g(X_{(n)})] = 0$ for all $\theta > 1$, then

$$0 = \int_0^\theta f(g(x)) x^{n-1} dx = \int_0^1 f(1) x^{n-1} dx + \int_1^\theta f(x) x^{n-1} dx$$

for all $\theta > 1$.

Letting $\theta \to 1$ we obtain that $f(1) = 0$.

Then

$$0 = \int_1^\theta f(x) x^{n-1} dx$$

for all $\theta > 1$, which implies $f(x) = 0$ a.e. for $x > 1$.

Hence, $g(X_{(n)})$ is complete.

Example 3.8

Let X be a sample (of size 1) from the uniform distribution $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, $\theta \in \mathscr{R}$.
We now apply Theorem 3.2 to show that there is no UMVUE of $\vartheta = g(\theta)$ for any nonconstant function g.

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

$$\int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} U(x)dx = 0 \qquad \text{for all } \theta \in \mathscr{R}.$$

Differentiating both sides of the previous equation and applying the result of differentiation of an integral lead to

$$U(x) = U(x+1) \quad \text{a.e. } m,$$

where m is the Lebesgue measure on $\mathscr{R}$.
If T is a UMVUE of $g(\theta)$, then $T(X)U(X)$ is unbiased for 0 and, hence,

$$T(x)U(x) = T(x+1)U(x+1) \quad \text{a.e. } m,$$

where $U(X)$ is any unbiased estimator of 0.

Example 3.8

Let X be a sample (of size 1) from the uniform distribution $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, $\theta \in \mathscr{R}$.
We now apply Theorem 3.2 to show that there is no UMVUE of $\vartheta = g(\theta)$ for any nonconstant function g.

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

$$\int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} U(x)dx = 0 \qquad \text{for all } \theta \in \mathscr{R}.$$

Differentiating both sides of the previous equation and applying the result of differentiation of an integral lead to

$$U(x) = U(x+1) \quad \text{a.e. } m,$$

where m is the Lebesgue measure on $\mathscr{R}$.
If T is a UMVUE of $g(\theta)$, then $T(X)U(X)$ is unbiased for 0 and, hence,

$$T(x)U(x) = T(x+1)U(x+1) \quad \text{a.e. } m,$$

where $U(X)$ is any unbiased estimator of 0.

Example 3.8 (continued)

Since this is true for all U,

$$T(x) = T(x+1) \quad \text{a.e. } m.$$

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

$$g(\theta) = \int_{\theta-\frac{1}{2}}^{\theta+\frac{1}{2}} T(x)dx \qquad \text{for all } \theta \in \mathscr{R}.$$

Differentiating both sides of the previous equation and applying the result of differentiation of an integral, we obtain that

$$g'(\theta) = T\left(\theta + \tfrac{1}{2}\right) - T\left(\theta - \tfrac{1}{2}\right) = 0 \quad \text{a.e. } m.$$