ST5215: Advanced Statistical Theory

Chen Zehua

Department of Statistics & Applied Probability

Monday, October 24, 2011
The 2nd method for deriving UMVUE: conditioning

- Find an unbiased estimator of \( \vartheta \), say \( U(X) \).
- Conditioning on a sufficient and complete statistic \( T(X) \): \( E[U(X)|T] \) is the UMVUE of \( \vartheta \).
- The distribution of \( T \) is not needed. We only need to work out the conditional expectation \( E[U(X)|T] \).
- From the uniqueness of the UMVUE, it does not matter which \( U(X) \) is used. Thus, \( U(X) \) should be chosen so as to make the calculation of \( E[U(X)|T] \) as easy as possible.
Example 3.3

Let $X_1, \ldots, X_n$ be i.i.d. from the exponential distribution $E(0, \theta)$ with p.d.f. $f_\theta(x) = \frac{1}{\theta} e^{-x/\theta} I_{(0, \infty)}(x)$.

Consider the estimation of $\vartheta = 1 - F_\theta(t)$.

$\bar{X}$ is sufficient and complete for $\theta > 0$.

$I_{(t, \infty)}(X_1)$ is unbiased for $\vartheta$,

$$E[I_{(t, \infty)}(X_1)] = P(X_1 > t) = \vartheta.$$ 

Hence

$$T(X) = E[I_{(t, \infty)}(X_1)|\bar{X}] = P(X_1 > t|\bar{X})$$

is the UMVUE of $\vartheta$.

If the conditional distribution of $X_1$ given $\bar{X}$ is available, then we can calculate $P(X_1 > t|\bar{X})$ directly.

By Basu’s theorem (Theorem 2.4), $X_1/\bar{X}$ and $\bar{X}$ are independent.

By Proposition 1.10(vii),

$$P(X_1 > t|\bar{X} = \bar{x}) = P(X_1/\bar{X} > t/\bar{x}|\bar{X} = \bar{x}) = P(X_1/\bar{X} > t/\bar{x}).$$
To compute this unconditional probability, we need the distribution of
\[
X_1 / \sum_{i=1}^{n} X_i = X_1 / \left( X_1 + \sum_{i=2}^{n} X_i \right).
\]

Using the transformation technique discussed in §1.3.1 and the fact that $\sum_{i=2}^{n} X_i$ is independent of $X_1$ and has a gamma distribution, we obtain that $X_1 / \sum_{i=1}^{n} X_i$ has the Lebesgue p.d.f. $(n - 1)(1 - x)^{n-2}I_{(0,1)}(x)$. Hence

\[
P(X_1 > t|\bar{X} = \bar{x}) = (n - 1) \int_{t/(n\bar{x})}^{1} (1 - x)^{n-2} dx = \left(1 - \frac{t}{n\bar{x}}\right)^{n-1}
\]

and the UMVUE of $\vartheta$ is

\[
T(X) = \left(1 - \frac{t}{n\bar{X}}\right)^{n-1}.
\]
Example 3.4
Let $X_1, \ldots, X_n$ be i.i.d. from $N(\mu, \sigma^2)$ with unknown $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

From Example 2.18, $T = (\bar{X}, S^2)$ is sufficient and complete for $\theta = (\mu, \sigma^2)$

$\bar{X}$ and $(n-1)S^2/\sigma^2$ are independent

$\bar{X}$ has the $N(\mu, \sigma^2/n)$ distribution

$S^2$ has the chi-square distribution $\chi^2_{n-1}$.

Using the method of solving for $h$ directly, we find that

- the UMVUE for $\mu$ is $\bar{X}$;
- the UMVUE of $\mu^2$ is $\bar{X}^2 - S^2/n$;
- the UMVUE for $\sigma^r$ with $r > 1 - n$ is $k_{n-1,r}S^r$, where

\[
k_{n,r} = \frac{n^{r/2}\Gamma\left(\frac{n}{2}\right)}{2^{r/2}\Gamma\left(\frac{n+r}{2}\right)}
\]

- the UMVUE of $\mu/\sigma$ is $k_{n-1,-1}\bar{X}/S$, if $n > 2$. 
Example 3.4 (continued)

Suppose that \( \vartheta \) satisfies \( P(X_1 \leq \vartheta) = p \) with a fixed \( p \in (0, 1) \). Let \( \Phi \) be the c.d.f. of the standard normal distribution. Then

\[
\vartheta = \mu + \sigma \Phi^{-1}(p)
\]

and its UMVUE is

\[
\bar{X} + k_{n-1,1} S \Phi^{-1}(p).
\]
Example 3.4 (continued)

Suppose that $\vartheta$ satisfies $P(X_1 \leq \vartheta) = p$ with a fixed $p \in (0, 1)$. Let $\Phi$ be the c.d.f. of the standard normal distribution. Then

$$\vartheta = \mu + \sigma \Phi^{-1}(p)$$

and its UMVUE is

$$\bar{X} + k_{n-1,1}S\Phi^{-1}(p).$$

Let $c$ be a fixed constant and

$$\vartheta = P(X_1 \leq c) = \Phi \left( \frac{c - \mu}{\sigma} \right).$$

We can find the UMVUE of $\vartheta$ using the method of conditioning. Since $I_{(-\infty,c)}(X_1)$ is an unbiased estimator of $\vartheta$, the UMVUE of $\vartheta$ is

$$E[I_{(-\infty,c)}(X_1)|T] = P(X_1 \leq c|T).$$

By Basu’s theorem, the ancillary statistic $Z(X) = (X_1 - \bar{X})/S$ is independent of $T = (\bar{X}, S^2)$. 
Example 3.4 (continued)

Then, by Proposition 1.10(vii),

\[ P \left( X_1 \leq c \mid T = (\bar{x}, s^2) \right) = P \left( Z \leq \frac{c - \bar{X}}{S} \middle| T = (\bar{x}, s^2) \right) \]
\[ = P \left( Z \leq \frac{c - \bar{x}}{s} \right). \]

It can be shown that \( Z \) has the Lebesgue p.d.f.

\[ f(z) = \frac{\sqrt{n} \Gamma \left( \frac{n-1}{2} \right)}{\sqrt{\pi} (n - 1) \Gamma \left( \frac{n-2}{2} \right)} \left[ 1 - \frac{nz^2}{(n - 1)^2} \right]^{(n/2)-2} I_{(0,(n-1)/\sqrt{n})}(|z|) \]

Hence the UMVUE of \( \vartheta \) is

\[ P(X_1 \leq c \mid T) = \int_{-(n-1)/\sqrt{n}}^{(c-\bar{X})/S} f(z) \, dz \]
Example 3.4 (continued)

Suppose that we would like to estimate

\[ \vartheta = \frac{1}{\sigma} \Phi' \left( \frac{c - \mu}{\sigma} \right), \]

the Lebesgue p.d.f. of \( X_1 \) evaluated at a fixed \( c \), where \( \Phi' \) is the first-order derivative of \( \Phi \).

By the previous result, the conditional p.d.f. of \( X_1 \) given \( \bar{X} = \bar{x} \) and \( S^2 = s^2 \) is \( s^{-1} f \left( \frac{x - \bar{x}}{s} \right) \).

Let \( f_T \) be the joint p.d.f. of \( T = (\bar{X}, S^2) \).

Then

\[ \vartheta = \int \int \frac{1}{s} f \left( \frac{c - \bar{X}}{s} \right) f_T(t) dt = E \left[ \frac{1}{S} f \left( \frac{c - \bar{X}}{S} \right) \right]. \]

Hence the UMVUE of \( \vartheta \) is

\[ \frac{1}{S} f \left( \frac{c - \bar{X}}{S} \right). \]
Example

Let $X_1, \ldots, X_n$ be i.i.d. with Lebesgue p.d.f. $f_\theta(x) = \theta x^{-2} I_{(\theta,\infty)}(x)$, where $\theta > 0$ is unknown.

Suppose that $\vartheta = P(X_1 > t)$ for a constant $t > 0$. The smallest order statistic $X_{(1)}$ is sufficient and complete for $\theta$. Hence, the UMVUE of $\vartheta$ is

$$
P(X_1 > t | X_{(1)}) = P(X_1 > t | X_{(1)} = x_{(1)})$$

$$= P \left( \frac{X_1}{X_{(1)}} > \frac{t}{x_{(1)}} \middle| X_{(1)} = x_{(1)} \right)$$

$$= P \left( \frac{X_1}{X_{(1)}} > \frac{t}{x_{(1)}} \middle| X_{(1)} = x_{(1)} \right)$$

$$= P \left( \frac{X_1}{x_{(1)}} > s \right)$$

(Basu’s theorem), where $s = t/x_{(1)}$. If $s \leq 1$, this probability is 1.
Consider $s > 1$ and assume $\theta = 1$ in the calculation:

$$P\left(\frac{X_1}{X_{(1)}} > s\right) = \sum_{i=1}^{n} P\left(\frac{X_1}{X_{(1)}} > s, X_{(1)} = X_i\right)$$

$$= \sum_{i=2}^{n} P\left(\frac{X_1}{X_{(1)}} > s, X_{(1)} = X_i\right)$$

$$= (n - 1)P\left(\frac{X_1}{X_{(1)}} > s, X_{(1)} = X_n\right)$$

$$= (n - 1)P\left(X_1 > sX_n, X_2 > X_n, \ldots, X_{n-1} > X_n\right)$$

$$= (n - 1) \int_{x_1 > sx_n, x_2 > x_n, \ldots, x_{n-1} > x_n} \prod_{i=1}^{n} \frac{1}{x_i^2} dx_1 \cdots dx_n$$

$$= (n - 1) \int_{1}^{\infty} \left[ \int_{sx_n}^{\infty} \prod_{i=2}^{n-1} \left( \int_{x_n}^{\infty} \frac{1}{x_i^2} dx_i \right) \frac{1}{x_1^2} dx_1 \right] \frac{1}{x_n^2} dx_n$$

$$= (n - 1) \int_{1}^{\infty} \frac{1}{sx_n^{n+1}} dx_n = \frac{(n - 1)X_{(1)}}{nt}$$
Example (continued)

This shows that the UMVUE of $P(X_1 > t)$ is

$$h(X_{(1)}) = \begin{cases} \frac{(n-1)X_{(1)}}{nt} & X_{(1)} < t \\ 1 & X_{(1)} \geq t \end{cases}$$
Example (continued)

This shows that the UMVUE of $P(X_1 > t)$ is

$$h(X_{(1)}) = \begin{cases} \frac{(n-1)X_{(1)}}{nt} & X_{(1)} < t \\ 1 & X_{(1)} \geq t \end{cases}$$

Use the method of finding $h$:

The UMVUE must be $h(X_{(1)})$

The Lebesgue p.d.f. of $X_{(1)}$ is $\frac{n^\theta}{x^{n+1}} I(\theta, \infty)(x)$. If $\theta \geq t$, then $P(X_1 > t) = 1$ and $P(t > X_{(1)}) = 0$. Hence, if $X_{(1)} \geq t$, $h(X_{(1)})$ must be 1 a.s. $P_\theta$

The value of $h(X_{(1)})$ for $X_{(1)} < t$ is not specified.
If $\theta < t$,

$$E[h(X_{(1)})] = \int_{\theta}^{\infty} h(x) \frac{n\theta^n}{x^{n+1}} \, dx$$

$$= \int_{\theta}^{t} h(x) \frac{n\theta^n}{x^{n+1}} \, dx + \int_{t}^{\infty} \frac{n\theta^n}{x^{n+1}} \, dx = \int_{\theta}^{t} h(x) \frac{n\theta^n}{x^{n+1}} \, dx + \frac{\theta^n}{t^n}$$

Since $P(X_1 > t) = \theta/t$, we have

$$\frac{\theta}{t} = \int_{\theta}^{t} h(x) \frac{n\theta^n}{x^{n+1}} \, dx + \frac{\theta^n}{t^n} \text{ i.e. } \frac{1}{t\theta^{n-1}} = \int_{\theta}^{t} h(x) \frac{n}{x^{n+1}} \, dx + \frac{1}{t^n}$$

Differentiating both sizes w.r.t. $\theta$ leads to

$$-\frac{n-1}{t\theta^n} = -h(\theta) \frac{n}{\theta^{n+1}}$$

Hence, for any $X_{(1)} < t$,

$$h(X_{(1)}) = \frac{(n-1)X_{(1)}}{nt}.$$
A necessary and sufficient condition

Theorem 3.2
Let $\mathcal{U}$ be the set of all unbiased estimators of $\theta$ with finite variances and $T$ be an unbiased estimator of $\theta$ with $E(T^2) < \infty$.

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

(ii) Suppose that $T = h(\tilde{T})$, where $\tilde{T}$ is a sufficient statistic for $P \in \mathcal{P}$ and $h$ is a Borel function. Let $\mathcal{U}_{\tilde{T}}$ be the subset of $\mathcal{U}$ consisting of Borel functions of $\tilde{T}$. Then a necessary and sufficient condition for $T$ to be a UMVUE of $\theta$ is that $E[T(X)U(X)] = 0$ for any $U \in \mathcal{U}_{\tilde{T}}$ and any $P \in \mathcal{P}$. 
Proof of Theorem 3.2(i)

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

$$\operatorname{Var}(T_c) \geq \operatorname{Var}(T) \quad c \in \mathcal{R}, \ P \in \mathcal{P},$$

which is the same as

$$c^2 \operatorname{Var}(U) + 2c \operatorname{Cov}(T, U) \geq 0 \quad c \in \mathcal{R}, \ P \in \mathcal{P}.$$

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

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

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

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

$$\operatorname{Var}(T) = \operatorname{Cov}(T, T_0) \quad P \in \mathcal{P}.$$

Note that $[\operatorname{Cov}(T, T_0)]^2 \leq \operatorname{Var}(T) \operatorname{Var}(T_0)$. Hence $\operatorname{Var}(T) \leq \operatorname{Var}(T_0)$ for any $P \in \mathcal{P}$. 
Proof of Theorem 3.2(ii)

It suffices to show that \( E(TU) = 0 \) for any \( U \in \mathcal{U}_T \) and \( P \in \mathcal{P} \) implies that \( E(TU) = 0 \) for any \( U \in \mathcal{U} \) and \( P \in \mathcal{P} \).

Let \( U \in \mathcal{U} \).

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

\[
E(TU) = E[E(TU|\tilde{T})] = E[E(h(\tilde{T})U|\tilde{T})] = E[h(\tilde{T})E(U|\tilde{T})].
\]
Proof of Theorem 3.2(ii)

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

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

Then $E(U|\tilde{T}) \in \mathcal{U}_{\tilde{T}}$ 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.
Corollary 3.1

(i) Let $T_j$ be a UMVUE of $\vartheta_j$, $j = 1, \ldots, k$, 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 $c_1, \ldots, c_k$.

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

Example 3.7
Let $X_1, \ldots, X_n$ 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 $\theta$ 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 $\theta$. Thus, Theorem 3.1 does not apply to $X_{(n)}$. 
Example 3.7 (continued)

We now use Theorem 3.2(ii) to find a UMVUE of $\theta$. 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,

\[ E[h(X(n))U(X(n))] = 0, \quad \theta \geq 1. \]

Since \( E[h(X(n))] = \theta \), we obtain that

\[
\begin{align*}
\theta &= cP(X(n) \leq 1) + bE[X(n)I_{(1,\infty)}(X(n))] \\
&= c\theta^{-n} + \left[ b n / (n + 1) \right] (\theta - \theta^{-n}).
\end{align*}
\]

Thus, \( c = 1 \) and \( b = (n + 1)/n \). The UMVUE of \( \theta \) 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

\[ g(X(n)) = \begin{cases} 
1 & 0 \leq X(n) \leq 1 \\
X(n) & X(n) > 1. 
\end{cases} \]

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

The sufficiency follows from the fact that the joint p.d.f. of \( X_1, \ldots, X_n \) is

\[ \frac{1}{\theta^n} l_{(0,\theta)}(X(n)) = \frac{1}{\theta^n} l_{(0,\theta)}(g(X(n))). \]

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 $U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, $\theta \in \mathbb{R}$. There is no UMVUE of $\vartheta = g(\theta)$ for any nonconstant function $g$.

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

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

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 $\mathbb{R}$.

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 \( g(\theta) \),

\[ g(\theta) = \int_{\theta - \frac{1}{2}}^{\theta + \frac{1}{2}} T(x) dx \quad \text{for all } \theta \in \mathcal{R}. \]

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

\[ g'(\theta) = T(\theta + \frac{1}{2}) - T(\theta - \frac{1}{2}) = 0 \quad \text{a.e. } m. \]