Chapter 3: Unbiased Estimation
Lecture 15: UMVUE: functions of sufficient and complete statistics

Unbiased estimation

Unbiased or asymptotically unbiased estimation plays an important role in point estimation theory.
Unbiased estimators can be used as "building blocks" for the construction of better estimators.
Asymptotic unbiasedness is necessary for consistency.
Our main focus:

- How to derive unbiased estimators
- How to find the best unbiased estimators

X: a sample from an unknown population $P \in \mathscr{P}$.
ϑ: a real-valued parameter related to P.
An estimator $T(X)$ of ϑ is unbiased iff $E[T(X)] = \vartheta$ for any $P \in \mathscr{P}$.
If there exists an unbiased estimator of ϑ, then ϑ is called an *estimable* parameter.

Definition 3.1 (UMVUE)

An unbiased estimator $T(X)$ of ϑ is called the *uniformly minimum variance unbiased estimator* (UMVUE) iff $\mathrm{Var}(T(X)) \leq \mathrm{Var}(U(X))$ for any $P \in \mathscr{P}$ and any other unbiased estimator $U(X)$ of ϑ.

Remarks

- Since the mse of any unbiased estimator is its variance, a UMVUE is $\mathfrak{I}$-optimal in mse with $\mathfrak{I}$ being the class of all unbiased estimators.
- One can similarly define the uniformly minimum risk unbiased estimator in statistical decision theory when we use an arbitrary loss instead of the squared error loss that corresponds to the mse.

Sufficient and complete statistics

The derivation of a UMVUE is relatively simple if there exists a sufficient and complete statistic for $P \in \mathscr{P}$.

Theorem 3.1 (Lehmann-Scheffé theorem)

Suppose that there exists a sufficient and complete statistic $T(X)$ for $P \in \mathscr{P}$.
If ϑ is estimable, then there is a unique unbiased estimator of ϑ that is of the form $h(T)$ with a Borel function h.
Furthermore, $h(T)$ is the unique UMVUE of ϑ.
(Two estimators that are equal a.s. $\mathscr{P}$ are treated as one estimator.)

Remarks

- This theorem is a consequence of Theorem 2.5(ii) (Rao-Blackwell theorem).
- One can easily extend this theorem to the case of the uniformly minimum risk unbiased estimator under any loss function $L(P, a)$ that is strictly convex in a.
- The uniqueness of the UMVUE follows from the completeness of $T(X)$.

There are two typical ways to derive a UMVUE when a sufficient and complete statistic T is available.

The first method: Directly solving for h

- Need the distribution of T
- Try some function h to see if $E[h(T)]$ is related to ϑ
- If $E[h(T)] = \vartheta$ for all P, what should h be?

Example 3.1

Let $X_1, ..., X_n$ be i.i.d. from the uniform distribution on $(0, \theta)$, $\theta > 0$.
Consider $\vartheta = \theta$.
Since the sufficient and complete statistic $X_{(n)}$ has the Lebesgue p.d.f. $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$,

$$EX_{(n)} = n\theta^{-n}\int_0^\theta x^n dx = \frac{n}{n+1}\theta.$$

An unbiased estimator of θ is $(n+1)X_{(n)}/n$, which is the UMVUE.
Consider now $\vartheta = g(\theta)$, where g is a differentiable function on $(0, \infty)$.
An unbiased estimator $h(X_{(n)})$ of ϑ must satisfy

$$\theta^n g(\theta) = n\int_0^\theta h(x)x^{n-1}dx \qquad \text{for all } \theta > 0.$$

Example 3.1 (continued)

Differentiating both sizes of the previous equation and applying the result of differentiation of an integral (Royden (1968, §5.3)) lead to

$$n\theta^{n-1}g(\theta) + \theta^n g'(\theta) = nh(\theta)\theta^{n-1}.$$

Hence, the UMVUE of ϑ is

$$h(X_{(n)}) = g(X_{(n)}) + n^{-1}X_{(n)}g'(X_{(n)}).$$

In particular, if $\vartheta = \theta$, then the UMVUE of θ is $(1+n^{-1})X_{(n)}$.

Example 3.2

Let $X_1,...,X_n$ be i.i.d. from the Poisson distribution $P(\theta)$ with an unknown $\theta > 0$.

$T(X) = \sum_{i=1}^n X_i$ is sufficient and complete for $\theta > 0$ and has the Poisson distribution $P(n\theta)$.

Since $E(T) = n\theta$, the UMVUE of θ is T/n.

Suppose that $\vartheta = g(\theta)$, where g is a smooth function such that $g(x) = \sum_{j=0}^\infty a_j x^j$, $x > 0$.

An unbiased estimator $h(T)$ of ϑ must satisfy (for any $\theta > 0$):

Example 3.2 (continued)

$$\sum_{t=0}^{\infty} \frac{h(t)n^t}{t!}\theta^t = e^{n\theta}g(\theta)$$

$$= \sum_{k=0}^{\infty} \frac{n^k}{k!}\theta^k \sum_{j=0}^{\infty} a_j\theta^j$$

$$= \sum_{t=0}^{\infty} \left(\sum_{j,k:j+k=t} \frac{n^k a_j}{k!} \right) \theta^t.$$

Thus, a comparison of coefficients in front of θ^t leads to

$$h(t) = \frac{t!}{n^t} \sum_{j,k:j+k=t} \frac{n^k a_j}{k!},$$

i.e., $h(T)$ is the UMVUE of ϑ.
In particular, if $\vartheta = \theta^r$ for some fixed integer $r \geq 1$, then $a_r = 1$ and $a_k = 0$ if $k \neq r$ and

$$h(t) = \begin{cases} 0 & t < r \\ \frac{t!}{n^r(t-r)!} & t \geq r \end{cases}$$

Example 3.5

Let $X_1, ..., X_n$ be i.i.d. from a power series distribution (see Exercise 13 in §2.6), i.e.,

$$P(X_i = x) = \gamma(x)\theta^x / c(\theta), \qquad x = 0, 1, 2, ...,$$

with a known function $\gamma(x) \geq 0$ and an unknown parameter $\theta > 0$.

It turns out that the joint distribution of $X = (X_1, ..., X_n)$ is in an exponential family with a sufficient and complete statistic $T(X) = \sum_{i=1}^n X_i$.

Furthermore, the distribution of T is also in a power series family, i.e.,

$$P(T = t) = \gamma_n(t)\theta^t / [c(\theta)]^n, \qquad t = 0, 1, 2, ...,$$

where $\gamma_n(t)$ is the coefficient of θ^t in the power series expansion of $[c(\theta)]^n$ (Exercise 13 in §2.6).

This result can help us to find the UMVUE of $\vartheta = g(\theta)$.

Example 3.5 (continued)

For example, by comparing both sides of

$$\sum_{t=0}^{\infty} h(t)\gamma_n(t)\theta^t = [c(\theta)]^{n-p}\theta^r,$$

we conclude that the UMVUE of $\theta^r/[c(\theta)]^p$ is

$$h(T) = \begin{cases} 0 & T < r \\ \frac{\gamma_{n-p}(T-r)}{\gamma_n(T)} & T \geq r, \end{cases}$$

where r and p are nonnegative integers.

In particular, the case of $p = 1$ produces the UMVUE $\gamma(r)h(T)$ of the probability

$$P(X_1 = r) = \gamma(r)\theta^r/c(\theta)$$

for any nonnegative integer r.

Example 3.6

Let $X_1, ..., X_n$ be i.i.d. from an unknown population P in a nonparametric family $\mathscr{P}$.

We have discussed in §2.2 that in many cases the vector of order statistics, $T = (X_{(1)}, ..., X_{(n)})$, is sufficient and complete for $P \in \mathscr{P}$. (For example, $\mathscr{P}$ is the collection of all Lebesgue p.d.f.'s.)

Note that an estimator $\varphi(X_1, ..., X_n)$ is a function of T iff the function φ is symmetric in its n arguments.

Hence, if T is sufficient and complete, then a symmetric unbiased estimator of any estimable ϑ is the UMVUE.

Specific examples

- $\bar{X}$ is the UMVUE of $\vartheta = EX_1$;
- S^2 is the UMVUE of $\mathrm{Var}(X_1)$;
- $n^{-1}\sum_{i=1}^n X_i^2 - S^2$ is the UMVUE of $(EX_1)^2$;
- $F_n(t)$ is the UMVUE of $P(X_1 \leq t)$ for any fixed t.

Example 3.6 (continued)

The previous conclusions are not true if T is *not* sufficient and complete for $P \in \mathscr{P}$.

Claim

For example, if $n > 2$ and $\mathscr{P}$ contains all symmetric distributions having Lebesgue p.d.f.'s and finite means, then there is no UMVUE for $\mu = EX_1$.

Proof

Suppose that T is a UMVUE of μ.
Let $\mathscr{P}_1 = \{N(\mu,1) : \mu \in \mathscr{R}\}$.
Since the sample mean $\bar{X}$ is UMVUE when $\mathscr{P}_1$ is considered, and the Lebesgue measure is dominated by any $P \in \mathscr{P}_1$, we conclude that $T = \bar{X}$ a.e. Lebesgue measure.
Let $\mathscr{P}_2$ be the family of uniform distributions on $(\theta_1 - \theta_2, \theta_1 + \theta_2)$, $\theta_1 \in \mathscr{R}$, $\theta_2 > 0$.
Then $(X_{(1)} + X_{(n)})/2$ is the UMVUE when $\mathscr{P}_2$ is considered, where $X_{(j)}$ is the jth order statistic.

Proof (continued)

Then $\bar{X} = (X_{(1)} + X_{(n)})/2$ a.s. P for any $P \in \mathscr{P}_2$, which is impossible if $n > 2$.
Hence, there is no UMVUE of μ.

What if $n = 1$?

Consider the sub-family $\mathscr{P}_1 = \{N(\mu, 1) : \mu \in \mathscr{R}\}$.
X_1 is complete for $P \in \mathscr{P}_1$.
Since $\mathscr{P}$ is dominated by $\mathscr{P}_1$, X_1 is complete for $P \in \mathscr{P}$.
X_1 is sufficient for $P \in \mathscr{P}$.
Thus, X_1 is the UMVUE of μ.

What if $n = 2$?

$T = (X_{(1)}, X_{(2)})$ is complete for $P \in \mathscr{P}_2$.
Since $\mathscr{P}$ is dominated by $\mathscr{P}_2$, T is complete for $P \in \mathscr{P}$.
T is also sufficient for $P \in \mathscr{P}$.
Thus, $\bar{X} = (X_1 + X_2)/2 = (X_{(1)} + X_{(2)})/2$ is the UMVUE of μ.

Survey samples from a finite population

Let $\mathscr{P} = \{1, ..., N\}$ be a finite population of interest

For each $i \in \mathscr{P}$, let y_i be a value of interest associated with unit i

Let $\mathbf{s} = \{i_1, ..., i_n\}$ be a subset of distinct elements of $\mathscr{P}$, which is a sample selected with selection probability $p(\mathbf{s})$, where p is known (sampling plan or sampling design).

The value y_i is observed if and only if $i \in \mathbf{s}$

If $p(\mathbf{s})$ is constant, the sampling plan is called the simple random sampling without replacement.

Consider the estimation of $Y = \sum_{i=1}^{N} y_i$, the population total as the parameter of interest

Issues to study

- How do we find an unbiased estimator of Y? Is Y estimable?
- Is there a UMVUE of Y under some conditions?

UMVUE under simple random sampling without replacement

Let $X = (X_i, i \in \mathbf{s})$ be the vector such that

$$P(X_1 = y_{i_1}, ..., X_n = y_{i_n}) = p(\mathbf{s})/n!$$

Let $\mathscr{Y}$ be the range of y_i, $\theta = (y_1, ..., y_N)$ and $\Theta = \prod_{i=1}^{N} \mathscr{Y}$.

Under simple random sampling without replacement, the population under consideration is a parametric family indexed by $\theta \in \Theta$.

Theorem 3.13 (Watson-Royall theorem)

(i) If $p(\mathbf{s}) > 0$ for all $\mathbf{s}$, then the vector of order statistics $X_{(1)} \leq \cdots \leq X_{(n)}$ is complete for $\theta \in \Theta$.

(ii) Under simple random sampling without replacement, the vector of order statistics is sufficient for $\theta \in \Theta$.

(iii) Under simple random sampling without replacement, for any estimable function of θ, its unique UMVUE is the unbiased estimator $g(X_1, ..., X_n)$, where g is symmetric in its n arguments.

Proof.

(i) Let $h(X)$ be a function of the order statistics.
Then h is symmetric in its n arguments.
We need to show that if

$$E[h(X)] = \sum_{\mathbf{s}=\{i_1,...,i_n\}\subset\{1,...,N\}} p(\mathbf{s})h\left(y_{i_1},...,y_{i_n}\right)/n! = 0 \tag{1}$$

for all $\theta \in \Theta$, then $h(y_{i_1},...,y_{i_n}) = 0$ for all $y_{i_1},...,y_{i_n}$.
First, suppose that all N elements of θ are equal to $a \in \mathscr{Y}$.
Then (1) implies $h(a,...,a) = 0$.
Next, suppose that $N-1$ elements in θ are equal to a and one is $b > a$.
Then (1) reduces to

$$q_1 h(a,...,a) + q_2 h(a,...,a,b),$$

where q_1 and q_2 are some known numbers in $(0,1)$.
Since $h(a,...,a) = 0$ and $q_2 \neq 0$, $h(a,...,a,b) = 0$.
Using the same argument, we can show that $h(a,...,a,b,...,b) = 0$ for any k a's and $n-k$ b's.

Suppose next that elements of θ are equal to a, b, or c, $a < b < c$.
Then we can show that $h(a,...,a,b,...,b,c,...,c) = 0$ for any k a's, l b's, and $n-k-l$ c's.
Continuing inductively, we see that $h(y_1,...,y_n) = 0$ for all possible $y_1,...,y_n$.
This completes the proof of (i).

(ii) The result follows from the factorization theorem (Theorem 2.2), the fact that $p(\mathbf{s})$ is constant under simple random sampling, and

$$P(X_1 = y_{i_1},...,X_n = y_{i_n}) = P(X_{(1)} = y_{(i_1)},...,X_{(n)} = y_{(i_n)})/n!,$$

where $y_{(i_1)} \leq \cdots \leq y_{(i_n)}$ are the ordered values of $y_{i_1},...,y_{i_n}$.

(iii) The result follows directly from (i) and (ii).

Remark

- Although we have a parametric problem under simple random sampling, the sufficient and complete statistic is the same as that in a nonparametric problem (Example 2.17).
- For the completeness of the order statistics, we do not need the assumption of simple random sampling.

Example 3.19.

Under simple random sampling without replacement, we now show

$$E(\bar{X}) = Y/N = \bar{Y}$$

i.e., the sample mean $\bar{X}$ is unbiased for the population mean $\bar{Y}$.
To show this, note that every unit i appears in the same number of samples, i.e.,

$$\begin{aligned} E(X_1 + \cdots + X_n) &= \binom{N}{n}^{-1} \sum_{\text{all } \mathbf{s}} \sum_{i \in \mathbf{s}} y_i \\ &= c(y_1 + \cdots + y_N) \\ &= cY \end{aligned}$$

This holds for all y_j's, hence it holds when $y_1 = y_2 = \cdots = y_N = 1$, in which case the left hand side $= n$ and the right hand side $= N$.

Hence, we must have $c = n/N$.

Since $\bar{X}$ is symmetric in its arguments, it is the UMVUE of $\bar{Y}$.