Lecture 29: UMVUE and the method of using the distribution of a sufficient and complete statistic

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

UMVUE

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

Definition 3.1. An unbiased estimator $T(X)$ of ϑ is called the *uniformly minimum variance unbiased estimator* (UMVUE) if and only if $\mathrm{Var}(\mathrm{T}(\mathrm{X})) \leq \mathrm{Var}(\mathrm{U}(\mathrm{X}))$ for any $P \in \mathcal{P}$ and any other unbiased estimator $U(X)$ of ϑ.

Since the mse of any unbiased estimator is its variance, a UMVUE is $\Im$-optimal in mse with $\Im$ 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 $P \in \mathcal{P}$.

Theorem 3.1 (Lehmann-Scheffé theorem). Suppose that there exists a sufficient and complete statistic $T(X)$ for $P \in \mathcal{P}$. 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. (Two estimators that are equal a.s. $\mathcal{P}$ are treated as one estimator.) Furthermore, $h(T)$ is the unique UMVUE of ϑ.

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 $T(X)$.

Two typical ways to derive a UMVUE when a sufficient and complete statistic T is available.

The 1st 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 $(0, \theta)$, $\theta > 0$.
Suppose that $\vartheta = \theta$.
Since the sufficient and complete statistic $X_{(n)}$ has the Lebesgue p.d.f. $n\theta^{-n}x^{n-1}I_{(0,\theta)}(x)$,

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

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

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

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 $\theta > 0$.
Then $T(X) = \sum_{i=1}^n X_i$ 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

$$\begin{aligned}\sum_{t=0}^\infty \frac{h(t)n^t}{t!}\theta^t &= e^{n\theta}g(\theta) \\ &= \sum_{k=0}^\infty \frac{n^k}{k!}\theta^k \sum_{j=0}^\infty a_j\theta^j \\ &= \sum_{t=0}^\infty \left(\sum_{j,k:j+k=t} \frac{n^k a_j}{k!} \right) \theta^t\end{aligned}$$

for any $\theta > 0$.
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 $\vartheta = g(\theta)$.
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 $\mathcal{P}$.
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 $P \in \mathcal{P}$.
(For example, $\mathcal{P}$ is the collection of all Lebesgue p.d.f.'s.) Note that an estimator $\varphi(X_1, ..., X_n)$ is a function of T if and only if 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.
For example,
$\bar{X}$ is the UMVUE of $\vartheta = EX_1$;
S^2 is the UMVUE of 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.

These conclusions are not true if T is *not* sufficient and complete for $P \in \mathcal{P}$.
For example, if $n > 1$ and $\mathcal{P}$ contains all symmetric distributions having Lebesgue p.d.f.'s and finite means, then there is no UMVUE for $\vartheta = EX_1$.
Suppose that T is a UMVUE of μ.

Let $\mathcal{P}_1 = \{N(\mu, 1) : \mu \in \mathcal{R}\}$.
Since the sample mean $\bar{X}$ is UMVUE when $\mathcal{P}_1$ is considered, and the Lebesgue measure is dominated by any $P \in \mathcal{P}_1$, we conclude that $T = \bar{X}$ a.e. Lebesgue measure.
Let $\mathcal{P}_2$ be the family of uniform distributions on $(\theta_1 - \theta_2, \theta_1 + \theta_2)$, $\theta_1 \in \mathcal{R}$, $\theta_2 > 0$.
Then $(X_{(1)} + X_{(n)})/2$ is the UMVUE when $\mathcal{P}_2$ is considered, where $X_{(j)}$ is the jth order statistic.
Then $\bar{X} = (X_{(1)} + X_{(n)})/2$ a.s. P for any $P \in \mathcal{P}_2$, which is impossible if $n > 1$.
Hence, there is no UMVUE of μ.

What if $n = 1$?
Consider the sub-family $\mathcal{P}_1 = \{N(\mu, 1) : \mu \in \mathcal{R}\}$.
Then X_1 is complete for $P \in \mathcal{P}_1$.
Hence, $E[h(X_1)] = 0$ for any $P \in \mathcal{P}$ implies that $E[h(X_1)] = 0$ for any $P \in \mathcal{P}_1$ and, thus, $h = 0$ a.e. Lebesgue measure.
This shows that X_1 is complete when the family $\mathcal{P}$ is considered.
Since $EX_1 = \mu$, X_1 is the UMVUE of μ.