STAT0008 Lecture 5
The Rao-Blackwell Theorem and Minimum Variance Unbiased Estimators

Dr. Aidan O'Keeffe

Department of Statistical Science
University College London

29th October 2018

Outline

- The Rao-Blackwell theorem
  - Proof
  - Applications
- Minimum variance unbiased estimator (MVUE)
- Completeness and its links to
  - MVUE
  - Exponential family of distributions
- Examples

The Rao-Blackwell Theorem

Let $\mathbf{X}$ be a sample with distribution $f(\mathbf{x};\theta)$ where $\theta \in \Theta$.

Suppose that $S(\mathbf{X})$ is sufficient for θ and that $V(\mathbf{X})$ is an estimator of $m(\theta)$ with $\mathbb{E}(V(\mathbf{X})^2) < \infty$ for all $\theta \in \Theta$. Then the estimator

$$T(\mathbf{X}) = \mathbb{E}\left[V(\mathbf{X}) \mid S(\mathbf{X})\right]$$

is such that $\text{MSE}(T(\mathbf{X}), m(\theta)) \leq \text{MSE}(V(\mathbf{X}), m(\theta))$ for all $\theta \in \Theta$.

Notably, if $V(\mathbf{X})$ is **unbiased** for $m(\theta)$ then $T(\mathbf{X})$ is also unbiased for $m(\theta)$ and

$$\text{Var}(T(\mathbf{X})) \leq \text{Var}(V(\mathbf{X})).$$

The Rao-Blackwell Theorem

In either case, we have equality if and only if $V(\mathbf{X})$ is a function of $\mathbf{X}$ only through $S(\mathbf{X})$.

We note that $T(\mathbf{X}) = \mathbb{E}\left[V(\mathbf{X}) \mid S(\mathbf{X})\right]$ is well-defined as an estimator of $m(\theta)$.

$T(\mathbf{X})$ does not depend on θ because the distribution of $\mathbf{X}$ (and thus the distribution of $V(\mathbf{X})$) conditional on $S(\mathbf{X})$ does not depend on θ (since $S(\mathbf{X})$ is sufficient for θ).

$T(\mathbf{X})$ is a function of $\mathbf{X}$ only through $S(\mathbf{X})$.

The Rao-Blackwell Theorem - Proof

We prove the Rao-Blackwell theorem

The Rao-Blackwell Theorem - Proof

The Rao-Blackwell Theorem - Proof

The Rao-Blackwell Theorem

In words, the Rao-Blackwell theorem may be summarised as

1. Take an estimator for $m(\theta)$, call this $V(\mathbf{X})$.
2. Take the expectation of $V(\mathbf{X})$, conditional on a sufficient statistic for θ, $S(\mathbf{X})$.
3. Obtain an estimator with mean squared error that is less than or equal to that of the original estimator.

Recall from Lecture 3 that a small mean squared error is a desirable property of any estimator. As a result, we can see that the Rao-Blackwell theorem can be useful in helping us to find desirable estimators.

The Rao-Blackwell Theorem

When $V(\mathbf{X})$ is unbiased for $m(\theta)$ then the Rao-Blackwell theorem may help us to find an **unbiased** estimator of $m(\theta)$ with smaller variance than $V(\mathbf{X})$ (if possible).

Recall, variance $=$ mean squared error for unbiased estimators of a given parameter.

Often, interest lies on unbiased estimators for θ. We have shown that, taking an unbiased estimator for θ and conditioning on a sufficient statistic, we can obtain an estimator of θ, $T(\mathbf{X})$, that is also unbiased and has variance smaller (or at worst) the same as our original estimator of θ.

The Rao-Blackwell Theorem: Example

Suppose that $X_1, \ldots, X_n$ are independent and identically distributed random variables with each $X_i \sim \mathcal{U}[0, \theta]$.

(i) Show that $T = \max\{X_1, \ldots, X_n\}$ is sufficient for θ.

(ii) Show that $U = 2X_1$ is unbiased for θ.

(iii) Use the Rao-Blackwell theorem to find an unbiased estimator of θ that is a function of T and that has variance less than or equal to that of U.

The Rao-Blackwell Theorem: Example

The Rao-Blackwell Theorem: Example

Minimum Variance Unbiased Estimators

In Lectures 3 and 4, we established that the Cramér-Rao lower bound exists as a lower bound for the variance of an estimator of some parameter, θ.

However, an unbiased estimator of θ whose variance attains the Cramér-Rao lower bound may not exist.

Typically, if we obtain an unbiased estimator of θ we compare its variance to the Cramér-Rao lower bound.

If the variance is near to (but not quite equal to) the bound then we might infer that our estimator is nearly optimal. But what if the variance is some distance from the bound?

How would we know if we can obtain a more desirable unbiased estimator (i.e. an estimator with smaller variance)?

Minimum Variance Unbiased Estimator

An unbiased estimator of θ with the smallest possible variance is known as a **minimum variance unbiased estimator** (MVUE).

Definition: Minimum Variance Unbiased Estimator (MVUE)

An estimator T^* is a **minimum variance unbiased estimator** of $m(\theta)$ if T^* satisfies $\mathbb{E}(T^*) = m(\theta)$ for all θ and, for any other estimator T with $\mathbb{E}(T) = m(\theta)$, we have

$$\text{Var}(T^*) \leq \text{Var}(T) \quad \text{for all } \theta.$$

We note that, if such an estimator attains the Cramér-Rao lower bound, then the estimator is a minimum variance bound unbiased estimator (MVBUE) (see Lecture 4 notes).

Efficiency

If $T(\mathbf{X})$ is an unbiased estimator of θ then the **efficiency** of $T(\mathbf{X})$ as an estimator of θ is defined

$$\mathsf{Eff}(T(\mathbf{X})) = \frac{1}{\mathcal{I}(\theta)\mathsf{Var}(T(\mathbf{X}))}.$$

This is the ratio of the minimum possible variance (i.e. the Cramér-Rao lower bound) to the variance of $T(\mathbf{X})$

We say that an estimator is **efficient** if the variance of the estimator attains the Cramér-Rao lower bound.

Minimum Variance Unbiased Estimator - Uniqueness

We prove that a MVUE must be unique.

Minimum Variance Unbiased Estimator - Uniqueness

Minimum Variance Unbiased Estimator

The Rao-Blackwell theorem tells us that we can take an unbiased estimator of $m(\theta)$ and condition on a sufficient statistic to find another unbiased estimator for $m(\theta)$ with variance less than or equal to that of our original estimator.

But the new estimator, obtained using the Rao-Blackwell theorem, will not necessarily be a MVUE.

However, we do know that our new estimator is a function of a sufficient statistic. So, when attempting to find a MVUE, we can restrict our attention to estimators that are functions of a sufficient statistic for the parameter of interest.

Minimum Variance Unbiased Estimator

Suppose we have found an unbiased estimator of $m(\theta)$, say $T(\mathbf{X})$. Then, to establish whether or not $T(\mathbf{X})$ is a MVUE, we need to consider whether or not any other unbiased estimator of $m(\theta)$ exists whose variance is less than that of $T(\mathbf{X})$.

We define

$$T'(\mathbf{X}) = T(\mathbf{X}) + aV(\mathbf{X})$$

where a is some real-valued constant and $V(\mathbf{X})$ is an unbiased estimator of 0.

Clearly, $T'(\mathbf{X})$ is unbiased for $m(\theta)$.

Now, we shall examine the variance of $T'(\mathbf{X})\ldots$

Minimum Variance Unbiased Estimator

$$\begin{aligned}\text{Var}(T'(\mathbf{X})) &= \text{Var}\left(T(\mathbf{X}) + aV(\mathbf{X})\right) \\ &= \text{Var}(T(\mathbf{X})) + a^2\text{Var}(V(\mathbf{X})) + 2a\text{Cov}(T(\mathbf{X}), V(\mathbf{X}))\end{aligned}$$

We see that $\text{Var}(T'(\mathbf{X})) < \text{Var}(T(\mathbf{X}))$ (i.e. T' is a better estimator than T if

$$a^2\text{Var}(V(\mathbf{X})) + 2a\text{Cov}(T(\mathbf{X}), V(\mathbf{X})) < 0.$$

In fact, the only situation in which the above cannot occur, for arbitrary a, is where $\text{Cov}(T(\mathbf{X}), V(\mathbf{X})) = 0$.

We can deduce that $T(\mathbf{X})$ is the MVUE of $m(\theta)$ if and only if $T(\mathbf{X})$ is uncorrelated with all unbiased estimators of 0.

Minimum Variance Unbiased Estimator

Let us briefly verify this. Suppose that we have an estimator T (where $\mathbb{E}(T) = m(\theta)$) that is uncorrelated with all unbiased estimators of 0.

Then let T' be any other estimator such that $\mathbb{E}(T') = m(\theta)$. Consider

$$T' = T + (T' - T)$$

then, the variance of T' is

$$\begin{aligned}
\mathsf{Var}(T') &= \mathsf{Var}\left(T + (T' - T)\right) \\
&= \mathsf{Var}(T) + \mathsf{Var}(T' - T) + 2\mathsf{Cov}(T, (T' - T)) \\
&= \mathsf{Var}(T) + \mathsf{Var}(T' - T)
\end{aligned}$$

since $T' - T$ is an unbiased estimator of 0 and, hence, T and $T' - T$ are uncorrelated.

Minimum Variance Unbiased Estimator

We have

$$\mathsf{Var}(T') = \mathsf{Var}(T) + \mathsf{Var}(T' - T)$$

We know that $\mathsf{Var}(T' - T) \geq 0$ and this implies that

$$\mathsf{Var}(T') \geq \mathsf{Var}(T).$$

Since T' is arbitrary, it follows that T is the minimum variance unbiased estimator of $m(\theta)$.

Minimum Variance Unbiased Estimator

We note here that an unbiased estimator of zero may be regarded as *random noise*.

Therefore, it follows that the addition of an unbiased estimator of zero to an unbiased estimator of $m(\theta)$ is unlikely to make sense.

Indeed, if an estimator could be improved (in terms of a reduction in its variance) by adding random noise then we'd be inclined to think that such an estimator is defective.

Minimum Variance Unbiased Estimator

We have established that a MVUE must be uncorrelated with all unbiased estimators of zero. But how useful is this property in application?

How could we verify that an estimator is uncorrelated with **all** unbiased estimators of zero?

We can characterise all unbiased estimators of zero if we place some conditions on the pdf/pmf with which we are working.

If the family of pdfs/pmfs with which we work, $f(x \mid \theta)$, is such that there are no unbiased estimators of zero for this family of pdfs/pmfs (other than zero itself!) then our work would be done, because

$$\text{Cov}(T, 0) = 0$$

for all estimators T.

Complete Statistic

We shall make use of a property known as **completeness**.

Completeness/Complete Statistic

Suppose $f(t \mid \theta)$ is a family of pdfs/pmfs for a statistic $T(\mathbf{X})$. The family of probability distributions is called **complete** (with respect to θ) if

$$\mathbb{E}\left(g(T(\mathbf{X}))\right) = 0 \text{ for all } \theta \implies \mathbb{P}(g(T(\mathbf{X})) = 0) = 1 \text{ for all } \theta.$$

If this property holds, then $T(\mathbf{X})$ is known as a **complete statistic**.

Completeness - Exponential Family

If $T(\mathbf{X})$ is sufficient for θ then, as we shall see, the property of completeness for the family of distributions of $T(\mathbf{X})$ will be important.

In other words, once we've found a sufficient statistic for θ, we want to determine whether or not this sufficient statistic is complete.

In particular, if a statistic has a distribution that belongs to the exponential family of distributions then the statistic is complete with respect to the unknown distributional parameters.

Hence, if the **sampling distribution** of a **sufficient statistic** $T(\mathbf{X})$ is a member of the **exponential family**, parameterised by θ, then the statistic $T(\mathbf{X})$ is **complete** and **sufficient** for θ.

Minimum Variance Unbiased Estimator

Suppose $\mathbf{X} = \{X_1, \ldots, X_n\}$ is a random sample and $T(\mathbf{X})$ is a complete, sufficient statistic for θ. Then, if $U(\mathbf{X})$ is an unbiased estimator for θ, the estimator

$$S(\mathbf{X}) = \mathbb{E}\left[U(\mathbf{X}) \mid T(\mathbf{X})\right]$$

is the unique MVUE for θ.

Minimum Variance Unbiased Estimator

Moreover...

If T is a complete sufficient statistic for θ and $\mathbb{E}(g(T)) = \tau(\theta)$. Then $g(T)$ is the unique minimum variance unbiased estimator of $\tau(\theta)$.

This result is known as the Lehmann-Scheffé theorem.

Under mild regularity conditions, a complete sufficient statistic is also minimal sufficient.

Minimum Variance Unbiased Estimator

$$S(\mathbf{X}) = \mathbb{E}\left[U(\mathbf{X}) \mid T(\mathbf{X})\right]$$

Here, we have

1. Taken a complete, sufficient statistic $T(\mathbf{X})$.
2. Taken an unbiased estimator of θ, $U(\mathbf{X})$ (no matter how inefficient, so long as $U(\mathbf{X})$ is unbiased for θ).
3. Use the **Rao-Blackwell theorem**, conditioning on the complete, sufficient statistic, to determine the unique MVUE for θ.

Minimum Variance Unbiased Estimator

Suppose that $\tilde{T}$ is the MVUE for θ and S is complete sufficient for θ. Then, by the Rao-Blackwell theorem, we set

$$T = \mathbb{E}(\tilde{T} \mid S).$$

We know that T is such that

- T is unbiased for θ.
- T is a function of S.
- $\mathrm{Var}(T) \leq \mathrm{Var}(\tilde{T})$.

But since $\tilde{T}$ is a MVUE then $\mathrm{Var}(T) = \mathrm{Var}(\tilde{T})$ and the uniqueness of the MVUE implies that $T = \tilde{T}$.

Minimum Variance Unbiased Estimator

What have we learned. . . ?

Well, we see that if we have an unbiased estimator for an unknown parameter θ (however inefficient) then. . .

If we can find a **complete**, **sufficient** statistic for θ. . .

Then we can use the Rao-Blackwell theorem to find the MVUE for θ.

Complete Statistic: Example

Suppose that $X \sim \text{Bin}(n, \theta)$ where n is known and $0 < \theta < 1$. Verify that X is complete for θ.

Complete Statistic: Example

MVUE: Example

Suppose that $X_1, \ldots, X_n$ are iid random variables with each $X_i \sim$ Bern(θ). Find the MVUE for θ^2.

MVUE: Example

MVUE: Example

Estimation of a Distribution Function

Suppose that we wish to find the MVUE of the cumulative distribution function of a random variable, X, at a point a.

In other words, we seek an estimator of

$$F(a; \theta) = \mathbb{P}(X \leq a; \theta)$$

and we have a random sample $X_1, \ldots, X_n$ with each $X_i \sim \mathcal{D}(\theta)$.

Suppose that $S(\mathbf{X})$ is a complete, sufficient statistic for θ.

Estimation of a Distribution Function

Consider the statistic $T(\mathbf{X})$, where $T(\mathbf{X})$ is defined

$$T(\mathbf{X}) = \begin{cases} 1 & \text{if } X_1 \leq a; \\ 0 & \text{otherwise.} \end{cases}$$

Then, clearly, $\mathbb{E}(T) = \mathbb{P}(X_1 \leq a) = F(a;\theta)$ so T is unbiased for $F(a;\theta)$

But $T(\mathbf{X})$ is obviously an inefficient estimator since $T(\mathbf{X})$ ignores all of the data apart from the first observation!

However, we know that the MVUE of $F(a;\theta)$ is given by

$$\mathbb{E}(T(\mathbf{X}) \mid S(\mathbf{X})) = \mathbb{P}(X_1 \leq a \mid S(\mathbf{X})).$$

Example: Estimation of a CDF

Suppose that $X_1, \ldots, X_n$ are iid Exp(θ) random variables. Find the MVUE of $\mathbb{P}(X_i \leq a)$ where a is some real-valued constant such that $a > 0$.

Example: Estimation of a CDF

Example: Estimation of a CDF

Learning Outcomes

- Understand the Rao-Blackwell theorem, how to prove this theorem and its importance in finding estimators.
- Understand the concept of a minimum variance unbiased estimator (MVUE) and how to find a MVUE.
- Understand the definition of an efficient estimator.
- Be able to show that the MVUE for a given parameter is unique.
- Understand the definition of completeness and its importance in establishing a MVUE.