STAT 346/446 Lecture 7

Sufficiency and Unbiasedness

CB Section 7.3.3 and DS Section 7.9

Note: We skip CB Section 7.3.4 for now.

Rao-Blackwell Theorem

note. $E(X|Y)$ is a function of Y → Random var.

- Rao-Blackwell Theorem is based on iterative expectation:

$$\mathrm{E}(X) = \mathrm{E}\left(\mathrm{E}(X \mid Y)\right)$$
$$\mathrm{Var}(X) = \mathrm{Var}\left(\mathrm{E}(X \mid Y)\right) + \mathrm{E}\left(\mathrm{Var}(X \mid Y)\right)$$

Rao-Blackwell Theorem

Let W be any unbiased estimator of $\tau(\theta)$, and let T be a sufficient statistic for θ. Let

$$\phi(T) = E\left(W \mid T\right)$$

← A statistic (does not depend on θ since T is suff.)

Then $E(\phi(T)) = \tau(\theta)$ and $\mathrm{Var}(\phi(T)) \leq \mathrm{Var}(W)$.

Rao-Blackwell Theorem

Let W be any unbiased estimator of $\tau(\theta)$, and let T be a sufficient statistic for θ. Let

$$\phi(T) = E(W \mid T)$$

Then $E(\phi(T)) = \tau(\theta)$ and $\mathrm{Var}(\phi(T)) \le \mathrm{Var}(W)$. $\leftarrow *$

both W and T are functions of $X_1, X_2, \ldots, X_n$

* W is unbiased est. of $\tau(\theta)$, i.e. $E(W) = \tau(\theta)$
* T is sufficient

$\Rightarrow f(\underline{X} \mid T=t)$ does not depend on θ

W is a function of $\underline{X}$

$\Rightarrow f(w \mid T=t)$ does not depend on θ

$\Rightarrow E(W \mid T)$ — " — " — " —

$\Rightarrow \phi(T)$ is a statistic

$E(W|T)$

$= \int w\, f(w|T)\, dw$

$=$ function of T

* $\phi(T)$ is an unbiased est. of $\tau(\theta)$:

$E(\phi(T)) = E\big(E(W|T)\big) = E(W)$

$= \tau(\theta)$

Proof of $*$:

$$Var(\phi(T)) = Var(E(W|T))$$

know: $$Var(W) = Var(E(W|T)) + \underbrace{E(\underbrace{Var(W|T)}_{\geq 0})}_{\geq 0}$$

$$\geq Var(\phi(T))$$

Can always find a better (or as good) estimator by conditioning on a sufficient statistic.

Sufficiency and unbiased estimators

- The Rao-Blackwell Theorem means that in our search for best unbiased estimators we only need to consider functions of sufficient statistics!
  - Narrows the search!

Theorem 7.3.18

If W is a best unbiased estimator of $\tau(\theta)$ then W is unique.

- Any other unbiased estimator will have a larger variance.

Theorem 7.3.20

If $\mathrm{E}(W) = \tau(\theta)$, then W is a best unbiased estimators of $\tau(\theta)$ *if and only if* W is uncorrelated with all unbiased estimators of 0

- def. of a complete statistic.

Sufficiency and unbiased estimators

Theorem 7.3.23 - Lehmann-Scheffé

Let T be a complete sufficient statistic for a parameter θ, and let $\phi(T)$ be any estimator based only on T.

Then $\phi(T)$ is the unique best unbiased estimator of its $\overbrace{\text{expected value}}^{\tau(\theta)}$.

All we need is completeness!

Using Rao-Blackwell + Lehmann-Sceffé
a_{LS}

say:
- Want estimator of $\tau(\theta)$
- Have a complete suff. stat T

If $E(T) = \tau(\theta)$ then by LS, T is the UMVUE of $\tau(\theta)$

If $E(T) \neq \tau(\theta)$

- sometimes there a obvious transformation $g(T)$ that gives $E(g(T)) = \tau(\theta)$
  $\Rightarrow$ By LS, $g(T)$ is the UMVUE of $\tau(\theta)$
- If not: Can use RB to find $\phi(T)$
  $\Rightarrow$ By LS, $\phi(T)$ is the UMVUE of $\tau(\theta)$

To use RB we have to find an unbiased estimator W, i.e. where $E(W) = \tau(\theta)$

Then $\phi(T) = E(W \mid T)$

To find a W:

- some simple est. is best, e.g. based on one observation.

Example

- Say $X_1, X_2, \ldots, X_n$ are iid. $\mathrm{Poisson}(\theta)$. Show, that $\overline{X}$ is the best unbiased estimator of θ.

saw before: $T = \sum_{i=1}^{n} X_i$ is a complete statistic.

$E(T) = \sum_{i=1}^{n} E(X_i) = n\theta$

but $E(\overline{X}) = E\left(\frac{1}{n}T\right) = \theta$

$\overline{X}$ is a function of a complete stat.

$E(\overline{X}) = \theta$

$\Rightarrow \overline{X}$ is the UMVUE of θ.

↳ by Lehmann - Scheffé

Example

- Let $X_1, X_2, \ldots, X_n$ are iid. $\text{Poisson}(\overset{\theta}{\cancel{\lambda}})$. Find the UMVUE of

$$P(X_i \leq 1) = P(X_i = 0) + P(X_i = 1) = \frac{e^{-\theta}\theta^0}{0!} + \frac{e^{-\theta}\theta^1}{1!}$$
$$= e^{-\theta} + \theta e^{-\theta} = (1+\theta)e^{-\theta} = \tau(\theta)$$

$T = \sum_{i=1}^{n} X_i$ is complete

$E(\overline{X}) = \theta$

Can we find a function of T, say $\phi(T)$ such that $E(\phi(T)) = (1+\theta)e^{-\theta}$?

If so, then $\phi(T)$ is the UMVUE of $(1+\theta)e^{-\theta}$ by Lehmann-Scheffe.

One possible guess: $(1+\bar{X})e^{-\bar{X}}$ <– MLE of $\tau(\theta)$

but unfortunately $E\left((1+\bar{X})e^{-\bar{X}}\right) \neq (1+\theta)e^{-\theta}$

Instead: Use Rao-Blackwell:

Strategy:

* Have a complete suff. stat. T
* Find a simple (e.g. based on only on X_i) unbiased estimator W
* Set $\phi(T) = E(W \mid T)$ <– simplify
  - unbiased and funct. of a comp! stat.

  $\Rightarrow$ UMVUE.

so we can actually use it!

Here: $\tau(\theta) = (1+\theta)e^{-\theta} = P(X_i \le 1)$ for any $i=1,\dots,n$

Consider:
$$W = \begin{cases} 1 & \text{if } X_1 \le 1 \\ 0 & \text{o.w.} \end{cases}$$

$W \sim \text{Bernoulli}(p)$, where

$p = P(W=1) = P(X_1 \le 1) = (1+\theta)e^{-\theta} = \tau(\theta)$

and $E(W) = p = (1+\theta)e^{-\theta}$

$\Rightarrow$ W is an unbiased estimator of $\tau(\theta)$

T compl.

need to simplify i.e. write so it can be evaluated from T (or $X_1,\dots,X_n$)

$\Rightarrow$ $\phi(T) = E(W|T)$ is the

UMVUE of $(1+\theta)e^{-\theta}$

What is $\phi(T) = E(W \mid T)$?

$$\phi(t) = E(W \mid T=t)$$ $W \sim$ Bernoulli

$$= 0 \cdot P(W=0 \mid T=t) + 1 \cdot P(W=1 \mid T=t)$$

$$= P(W=1 \mid T=t)$$

$$= P(X_1 \leq 1 \mid T=t)$$ Remember: $T = \sum_{i=1}^{n} X_i$

$$= \frac{P(X_1 \leq 1, \sum_{i=1}^{n} X_i = t)}{P(T=t)}$$

① : $T \sim \text{Poisson}(n\theta) \Rightarrow P(T=t) = \dfrac{e^{-n\theta}(n\theta)^t}{t!}$

② : $P(X_1 \leq 1, \sum_{i=1}^{n} X_i = t) =$

② : $P\left(X_1 \leq 1, \sum_{i=1}^{n} X_i = t\right) =$

$$= P\left(\left[X_1 = 0, \sum_{i=2}^{n} X_i = t\right] \text{ or } \left[X_1 = 1, \sum_{i=2}^{n} X_i = t-1\right]\right)$$

$$= P\left(X_1 = 0, \sum_{i=2}^{n} X_i = t\right) + P\left(X_1 = 1, \sum_{i=2}^{n} X_i = t-1\right)$$

note: X_1 and $\sum_{i=2}^{n} X_i$ are independent

$$= P(X_1 = 0)\, P\left(\sum_{i=2}^{n} X_i = t\right) + P(X_1 = 1)\, P\left(\sum_{i=2}^{n} X_i = t-1\right)$$

note: $\sum_{i=2}^{n} X_i \sim \text{Poisson}((n-1)\theta)$

$$= \frac{e^{-\theta}\theta^{0}}{0!} * \frac{e^{-(n-1)\theta}((n-1)\theta)^{t}}{t!} + \frac{e^{-\theta}\theta^{1}}{1!} \frac{e^{-(n-1)\theta}((n-1)\theta)^{t-1}}{(t-1)!}$$

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

$$\Rightarrow \phi(t) = \frac{\dfrac{e^{-n\theta}(n-1)^t\theta^t}{t!} + \dfrac{e^{-n\theta}(n-1)^{t-1}\theta^t}{\cancel{t!}\;(t-1)!}}{\dfrac{e^{-n\theta}(n\theta)^t}{t!}}$$

$$= \frac{e^{-n\theta}\theta^t\left(\dfrac{(n-1)^t}{t!} + \dfrac{(n-1)^{t-1}}{(t-1)!}\right)}{e^{-n\theta}n^t\theta^t/t!}$$

$$= \frac{t!}{n^t}\left(\frac{(n-1)^t}{t!} + \frac{(n-1)^{t-1}}{(t-1)!}\right)$$

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

← would not have guessed!

or: $$= \left(\frac{n-1}{n}\right)^{n\bar{x}} + \bar{x}\left(\frac{n-1}{n}\right)^{n\bar{x}-1}$$

But what is $\phi(T)$?

$$E(W \mid T) = 0\, P(W=0 \mid T) + 1 \cdot P(W=1 \mid T)$$

$$= P(W=1 \mid T)$$

$$= P(X_1 \leq 1 \mid T=t) \qquad T = \sum_{i=1}^{n} X_i$$

$$= \frac{P(\overbrace{X_1 \leq 1, \sum_{i=1}^{n} X_i = t}^{A})}{P(\sum_{i=1}^{n} X_i = t)}$$

A: $X_1 \leq 1 \Rightarrow X_1 = 0$ or $X_1 = 1$, $\sum_{i=1}^{n} X_i = t$

so $\sum_{i=2}^{n} X_i = t$ or $\sum_{i=2}^{n} X_i = t-1$

Best unbiased estimators - Summary

- Let W be an *unbiased* estimator of a parameter θ
- Then $\mathrm{MSE}(W) = \mathrm{Var}(W)$
- In general, we want our unbiased estimators to have low variance (and hence low MSE)

Best unbiased estimator

= **Uniform minimum variance unbiased estimator (UMVUE)**

- The estimator that has the smallest variance in the set of all unbiased estimators
- Section 7.3 is mostly about different theoretical tricks to see if we have a best unbiased estimator

We didn't cover "equivariant" estimators so you can skip Example 7.3.6 and the text right above it

Trick 1: Cramer-Rao Lower bound

Section 7.3.2

- Puts a lower bound on the variance of all estimators - given that some "nicety" conditions hold
- Our estimators can't have a smaller variance than the Cramer-Rao lower bound (CRLB)
- So if our unbiased estimator has a variance that is equal to CRLB we know that it *is* the best unbiased estimator

Efficient estimators

An estimator W is called and **efficient estimator** if it has a variance that is equal to its CRLB.

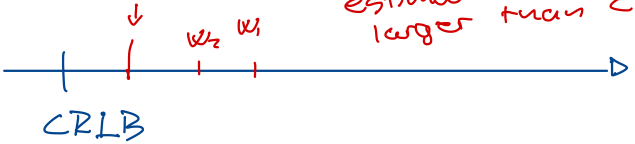
- Note that a best unbiased estimator is not necessarily efficient
- The CRLB may not be obtainable

The UMVUE is not necessarily efficient.

CRLB : lower bound for variance of all unbiased estimators

But there may not exist an unbiased estimator W with $Var(W) = CRLB$

smallest variance of all unbiased estimator may be larger than CRLB



"efficient" : $Var(W) = CRLB$

$\Downarrow$

Best

Best $\not\Rightarrow$ efficient

Trick 2: Complete statistic

Section 7.3.3.

All we need is completeness:

Lehmann-Scheffé

Let T be a complete sufficient statistic for a parameter θ, and let $\phi(T)$ be any estimator based only on T. Then $\phi(T)$ is the unique best unbiased estimator of its expected value.

- Rao-Blackwell + Lehmann-Scheffé:
  - Find a complete statistic T. If unbiased of $\tau(\theta)$, then we are done!
  - If *not* unbiased, find a simple unbiased estimator W and set

$$\phi(T) = E(W \mid T)$$

    then $\phi(T)$ is unbiased and based only on a complete sufficient statistic. Therefore $\phi(T)$ is the best unbiased estimator of $\tau(\theta)$