Lecture 10 - Rao-Blackwell

Review : Conditional Expectation

$$E[X \mid Y = y] = \int x\, f(x|y)\, dx = g(y)$$

random version of this

Iterated Expectation

$$\rightarrow \quad \mathbb{E}[X] = \mathbb{E}\Big[\overbrace{\mathbb{E}[X|Y]}^{\text{a RV}}\Big]$$

Theorem: Law of Total Variance

$$\rightarrow \mathrm{Var}(X) = \mathbb{E}\big[\mathrm{Var}(X|Y)\big] + \mathrm{Var}\, \mathbb{E}[X|Y]$$

Ex. $X|Y=y \sim \mathrm{Bin}(y, p)$, $p \in [0,1]$

$Y \sim \mathrm{Pois}(\lambda)$, $\lambda > 0$

$\mathbb{E}[X]$

1. $\mathbb{E}[X|Y=y] = yp$
2. $\mathbb{E}[X|Y] = Yp$
3. $\mathbb{E}\big[\mathbb{E}[X|Y]\big] = \mathbb{E}[Yp] = p\,\mathbb{E}Y = p\lambda$

$\mathrm{Var}(X)$?

$\rightarrow \mathrm{Var}(X|Y=y) = yp(1-p)$

$\rightarrow \mathrm{Var}(X|Y) = Yp(1-p)$

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{Var}\, \mathbb{E}[X|Y] + \mathbb{E}\, \mathrm{Var}(X|Y) \\
&= \mathrm{Var}(Yp) + \mathbb{E}\big[Yp(1-p)\big] \\
&= p^2\lambda + p(1-p)\lambda \\
&= p\lambda
\end{aligned}$$

$$= p^2 \lambda + p(1-p)\lambda$$
$$= \underline{p\lambda}$$

Back to math. stats.

CRLB doesn't always allow us to find UMVUE.

Let's look at how to refine bad estimators into better estimators using conditioning

---

Some facts

① Let $\hat{\theta}$ be unbiased for $\tau(\theta)$
so that $\mathbb{E}\hat{\theta} = \tau(\theta)$

Let $W = W(\underset{\sim}{X})$ to be some fn of $\underset{\sim}{X}$
(maybe a stat - maybe not could depend on θ)

Let $\varphi(W) = \varphi = \mathbb{E}[\hat{\theta} \mid W]$

a fn of $\underset{\sim}{X}$

could be a stat - or not

In any case

$$\mathbb{E}\varphi = \mathbb{E}\big[\mathbb{E}[\hat{\theta} \mid W]\big] = \mathbb{E}\hat{\theta} = \tau(\theta)$$

$$\boxed{\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}X}$$

If φ is a stat, then it is unbiased for $\tau(\theta)$.

② $\underbrace{Var(\varphi)}_{\varphi = E[\hat\theta | W]} \le Var(\hat\theta)$.

Law of total var says

$$Var\,\hat\theta = \underbrace{Var\,E[\hat\theta | W]}_{Var(\varphi)} + \underbrace{E\,Var(\hat\theta | W)}_{\geq 0}$$

So $Var(\hat\theta) \geq Var(\varphi)$.

<u>Summary:</u> Start w/ $\hat\theta$ where $E\hat\theta = \tau(\theta)$

If I define $\varphi = E[\hat\theta | W]$

then ① $E\varphi = \tau(\theta) \leftarrow$

② $Var(\varphi) \le Var(\hat\theta) \leftarrow$

Main issue : <u>can't guarantee φ is a stat.</u>

<u>Ex.</u> $X_n \overset{iid}{\sim} N(\theta, 1)$

define $\hat\theta = \frac{1}{2}(X_1 + X_2)$.

<u>notice:</u> $E\hat\theta = E\left[\frac{1}{2}(X_1 + X_2)\right] = \frac{1}{2}(EX_1 + EX_2)$

$$= \frac{1}{2}(\theta + \theta) = \theta$$

$$Var\,\hat\theta = \frac{1}{4}\left(Var(X_1) + Var(X_2)\right) = \frac{1}{2}$$

$W = X_1$

$$\varphi = E[\hat\theta | W] = E\left[\frac{1}{2}(X_1 + X_2) \,\middle|\, X_1\right]$$

$$\varphi = \mathbb{E}[\hat{\theta} \mid W] = \mathbb{E}\left[\tfrac{1}{2}(X_1 + X_2) \mid X_1\right]$$

$$= \tfrac{1}{2}\left(\mathbb{E}[X_1 \mid X_1] + \mathbb{E}[X_2 \mid X_1]\right)$$

Aside: $\mathbb{E}[Z \mid Z=5] = 5$

$\mathbb{E}[Z \mid Z=3] = 3$

$\mathbb{E}[Z \mid Z] = Z$

$$= \tfrac{1}{2}\left(X_1 + \mathbb{E}[X_2]\right)$$

$$\boxed{= \tfrac{1}{2}(X_1 + \theta) = \varphi}$$

problem: not a stat

① $\mathbb{E}[\varphi] = \mathbb{E}\left[\tfrac{1}{2}(X_1 + \theta)\right] = \tfrac{1}{2}\left(\mathbb{E}[X_1] + \theta\right)$

$$= \tfrac{1}{2}(\theta + \theta) = \theta$$

② $\mathrm{Var}(\varphi) = \frac{1}{4}\mathrm{Var}(X_1) = \frac{1}{4} < \mathrm{Var}\,\hat{\theta} = \frac{1}{2}$

Let $W = \bar{X}$ and re-do

$$\varphi = \mathbb{E}[\hat{\theta} \mid W] = \mathbb{E}\left[\tfrac{1}{2}(X_1 + X_2) \mid \bar{X}\right]$$

$$= \tfrac{1}{2}\left(\mathbb{E}[X_1 \mid \bar{X}] + \mathbb{E}[X_2 \mid \bar{X}]\right)$$

$X_n \overset{iid}{\sim} N(\theta, 1)$ so $\mathbb{E}[X_1 \mid \bar{X}] = \mathbb{E}[X_2 \mid \bar{X}] = \cdots$

$$= \tfrac{1}{2}\, 2\, \mathbb{E}[X_1 \mid \bar{X}]$$

$$= \mathbb{E}[X_1 \mid \bar{X}]$$

$$= \tfrac{1}{N} N\, \mathbb{E}[X_1 \mid \bar{X}]$$

$$= \tfrac{1}{N} \sum_{i=1}^{N} \mathbb{E}[X_i \mid \bar{X}]$$

$$= \mathbb{E}\left[\tfrac{1}{N}\sum_{i=1}^{N} X_i \mid \bar{X}\right]$$

$$= \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} X_i \,\middle|\, \bar{X}\right]$$

$$= \mathbb{E}[\bar{X} \mid \bar{X}]$$

$$\boxed{\varphi = \bar{X}} \quad (*)$$

this is a stat.

We also know:

① $\mathbb{E}\varphi = \theta = \mathbb{E}\bar{X}$

② $\underbrace{\mathrm{Var}\,\varphi}_{1/N} \le \underbrace{\mathrm{Var}\,\hat{\theta}}_{1/2}$

Theorem: Rao-Blackwell Theorem

Let $\hat{\theta}$ is unbiased for $\tau(\theta)$ and W is sufficient for θ. Then if

$$\varphi = \mathbb{E}[\hat{\theta} \mid W]$$

① $\mathbb{E}[\varphi] = \tau(\theta)$

② $\mathrm{Var}\,\varphi \le \mathrm{Var}\,\hat{\theta}$

③ φ is a statistic.

Pf: Fact ③

$$\varphi = \mathbb{E}[\underbrace{\hat{\theta}(\underset{\sim}{X})}_{g} \mid \overset{\text{sufficient}}{W}]$$

$$= \int \underbrace{\hat{\theta}(\underset{\sim}{x})}_{\text{a stat, no } \theta} \underbrace{f_{\underset{\sim}{X}|W}(\underset{\sim}{x})}_{W \text{ sufficient, no } \theta \text{ in this PDF}} \, d\underset{\sim}{x}$$

a stat

no θ

W sufficient,

no θ in this PDF

$\mathbb{E}[g(X)] = \int g(x) f(x)\,dx$

$\mathbb{E}[g(X) \mid Y = y]$

$= \int g(x) f(x|y)\,dx$

Theorem: Lehmann - Scheffé

tech. condition.

Let W be a (complete) sufficient statistic for θ and let $\hat{\theta}$ be an unbiased est. for $\tau(\theta)$ that depends on $\underset{\sim}{X}$ s only through W,

$$\hat{\theta} = \hat{\theta}(W) = \hat{\theta}(W(\underset{\sim}{X}))$$

then $\hat{\theta}$ is the UMVUE for $\tau(\theta)$.

Basically: If I can form an unbiased est. for $\tau(\theta)$ from a S.S. it is the UMVUE.

Ex. $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$

Known

don't know

Q: what is the UMVUE for μ?

Use Lehmann - Scheffé

1. Find a SS for μ : $\bar{X}$

2. Guess a fn of $\bar{X}$ that is unbiased for μ

   $\hat{\mu} = \bar{X}$ is unbiased for μ, $E\bar{X} = \mu$.

3. So $\bar{X}$ is the UMVUE.

Ex. Let $\tau(\mu) = \mu^2$.

① SS for μ : $\bar{X}$

② Find / guess a fn of $\bar{X}$ that is unbiased for $\mu^2 = \tau(\mu)$.

Consider: $\bar{X}^2$, $\mathbb{E}[\bar{X}^2] = Var(\bar{X}) + (\mathbb{E}\bar{X})^2$

$\boxed{\bar{X} \sim N(\mu, \sigma^2/N)}$ $\quad = \sigma^2/N + \mu^2$

Let $\hat{\mu^2} = \bar{X}^2 - \sigma^2/N$

$\mathbb{E}[\hat{\mu^2}] = \mathbb{E}[\bar{X}^2] - \sigma^2/N = \mu^2$

So $\bar{X}^2 - \sigma^2/N$ is the UMVUE for μ^2.

Takeaway:

How to find UMVUE for $\tau(\theta)$ w/ Lehmann-Scheffé

① Find SS for θ — call it W

② Find a fn of W that is unbiased for $\tau(\theta)$

(i) Guess fn $\hat{\theta}(W)$ so that $\mathbb{E}[\hat{\theta}(W)] = \tau(\theta)$

(ii) Use Rao-Blackwell and condition using W,

$\rightarrow$ Find any unbiased est. of $\tau(\theta)$ call it V

$\rightarrow \hat{\theta} = E[V/W]$ is the UMVUE.

$\Rightarrow \theta = \mathbb{E}[V|W]$ is the UMVUE.