Review of Iterated Expectation

$$E_X[X] = \int x \, f(x) \, dx$$

$$E_Y[\sim] = \int \sim f(y) \, dy$$

$$E_{X|Y}[\sim] = \int \sim f(x|y) \, dx$$

Theorem: Iterated Expectation

$$\underbrace{E}_{\substack{\text{wrt } X \\ f(x)}}[X] = \underbrace{E}_{\substack{\text{wrt } Y \\ f(y)}}\Big[\underbrace{E}_{\substack{\text{wrt. } X|Y \\ f(x|y)}}[X|Y]\Big]$$

Pf. $f(x|y) = \dfrac{f(x,y)}{f(y)}$ and $f(y|x) = \dfrac{f(x,y)}{f(x)}$

(cts) so $f(x,y) = f(x|y)f(y) = f(y|x)f(x)$

$$E[X] = \int x f(x) \, dx = \int x \underbrace{\int f(x,y) \, dy}_{f(x)} \, dx$$

$$= \iint x f(x,y) \, dy \, dx$$

$$= \iint x f(x|y) f(y) \, dx \, dy$$

$$= \int \left[\int x f(x|y) \, dx \right] f(y) \, dy$$

$$= \int \underbrace{E[X|Y=y]}_{g(y)} f(y) \, dy$$

$$= E\big[E[X|Y]\big]$$

$E[X|Y=y] = g(y)$

$E[X|Y] = g(Y)$

$$= \mathbb{E}\Big[\underbrace{\overbrace{\mathbb{E}[X|Y]}^{g(Y)}}_{g(Y)}\Big]$$

Theorem: Law of Total Variance

$$Var(X) = \mathbb{E}\Big[\underbrace{Var(X|Y)}_{\text{r.v.}}\Big] + Var\big(\underbrace{\mathbb{E}[X|Y]}_{\text{r.v.}}\big)$$

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

$Y \sim Pois(\lambda)$, $\lambda \geq 0$

$\mathbb{E}[X]$?

$$\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X|Y]\big]$$

1. $\mathbb{E}[\underbrace{X|Y=y}_{Bin(y,p)}] = yp$
2. $\boxed{\mathbb{E}[X|Y] = Yp}$
3. $\mathbb{E}\big[\mathbb{E}[X|Y]\big] = \mathbb{E}[Yp] = p\,\mathbb{E}[Y]$

$$= p\lambda$$

$$Var(X) = \underline{Var(\mathbb{E}[X|Y])} + \mathbb{E}[\underline{Var(X|Y)}]$$

$$\mathrm{Var}(X) = \underbrace{\mathrm{Var}(\mathbb{E}[X \mid Y])}_{} + \mathbb{E}[\underbrace{\mathrm{Var}(X \mid Y)}_{}]$$

(1) $\mathbb{E}[X \mid Y] = Yp$

(2) $\mathrm{Var}(\mathbb{E}[X \mid Y]) = \mathrm{Var}(Yp)$

$= p^2 \mathrm{Var}(Y)$

$\boxed{= p^2 \lambda}$

(1) $\mathrm{Var}(X \mid Y = y)$

$= yp(1-p)$

(2) $\mathrm{Var}(X \mid Y)$

$= Yp(1-p)$

(3) $\mathbb{E}[\sim]$

$= \mathbb{E}[Yp(1-p)]$

$= p(1-p)\mathbb{E}[Y]$

$\boxed{= p(1-p)\lambda}$

$$\mathrm{Var}(X) = p^2\lambda + p(1-p)\lambda = p\lambda$$

Back to Math Stats.

CRLB doesn't always work (non-exp fams).
We want a more general method for find the UMVUE.

Fact: Let $\hat{\theta}$ be unbiased for $\tau(\theta)$

so that $\mathbb{E}\hat{\theta} = \tau(\theta)$

and let $W = W(\underset{\sim}{X})$ be some function of $\underset{\sim}{X}$

(not nec. a statistic
W could depend on θ)

and let

$$\varphi = \varphi(W) = E[\hat{\theta}|W].$$

depends on W

depends on $\underline{X}$

notice

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

If φ is a statistic, then it is an unbiased est. of $\tau(\theta)$.

Fact:

$$\mathrm{Var}(\varphi) = \mathrm{Var}(E[\hat{\theta}|W])$$

Law of total variance

$$\mathrm{Var}(\hat{\theta}) = \mathrm{Var}\,E[\hat{\theta}|W] + E\,\mathrm{Var}(\hat{\theta}|W)$$

rearrange

$$\mathrm{Var}(\varphi) = \mathrm{Var}\,E[\hat{\theta}|W] = \mathrm{Var}(\hat{\theta}) - E\,\mathrm{Var}(\hat{\theta}|W)$$

$$= \mathrm{Var}(\hat{\theta}) - \text{something}$$

positive

punchline:

$$\boxed{Var(\varphi) \leq Var(\hat{\theta})}$$

Summary: (1) $\hat{\theta}$ unbiased stat. for $\tau(\theta)$
(2) W some quantity (for $\underline{X}$)
(3) $\boxed{\varphi = E[\hat{\theta} | W]}$

THEN (i) $E[\varphi] = \tau(\theta)$
(ii) $Var(\varphi) \leq Var(\hat{\theta})$

implication: if φ is a statistic
(formula for φ doesn't depend on θ)
then φ is a better estimator than $\hat{\theta}$
(we have refined $\hat{\theta}$)

Ex. let $X_n \overset{iid}{\sim} N(\theta, 1)$

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

notice: $E[\hat{\theta}] = \frac{1}{2}E[X_1] + \frac{1}{2}E[X_2] = \theta$

$\hat{\theta}$ unbiased for θ

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

let $\boxed{W = X_1}$

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

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

Aside: $\mathbb{E}[Z|Z] = ?$

$\mathbb{E}[Z | Z=5] = 5$

$\mathbb{E}[Z | Z=z] = z \Rightarrow \mathbb{E}[Z|Z] = Z$

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

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

not a statistic!

notice however

$\mathbb{E}[\varphi] = \frac{1}{2}(\overbrace{\mathbb{E}[X_1]}^{\theta} + \theta) = \theta$ ✓

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

$$\mathrm{Var}(\varphi) = \frac{1}{4}\mathrm{Var}(X_1) = 1/4 < \frac{1}{2} = \mathrm{Var}(\hat{\theta}). \quad \checkmark$$

Consider instead

$\boxed{W = \overline{X}.}$ (redo calculations)

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

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

$X_n \overset{iid}{\sim} f_\theta$ so $\mathbb{E}[X_1 \mid \overline{X}] = \mathbb{E}[X_2 \mid \overline{X}]$

$= \mathbb{E}[X_3 \mid \overline{X}] = \cdots = \mathbb{E}[X_N \mid \overline{X}]$

$$= \frac{1}{2} \, 2\,\mathbb{E}[X_1 \mid \overline{X}]$$

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

$$= \frac{1}{N} N\,\mathbb{E}[X_1 \mid \overline{X}]$$

$$= \frac{1}{N}\left(\underbrace{\mathbb{E}[X_1 \mid \overline{X}] + \mathbb{E}[X_1 \mid \overline{X}] + \cdots + \mathbb{E}[X_1 \mid \overline{X}]}_{N \text{ times}}\right)$$

$$= \frac{1}{N}\left(E[X_1|\bar{X}] + E[X_2|\bar{X}] + E[X_3|\bar{X}] + \cdots + E[X_N|\bar{X}]\right)$$

$$= E\left[\frac{X_1 + X_2 + \cdots + X_N}{N} \,\Big|\, \bar{X}\right]$$

$$= E[\bar{X}|\bar{X}] = \bar{X}.$$

So $\varphi = E[\hat{\theta}|W] = \bar{X}$

this a statistic

we know

1. $E[\varphi] = E[\bar{X}] = \theta$
2. $Var(\varphi) = Var(\bar{X}) = 1/N < \frac{1}{2} = Var(\hat{\theta})$

   $X_n \overset{iid}{\sim} N(\theta, 1)$
3. φ is a stat.

φ is a "better" stat./est. of θ.

(in fact φ is the UMVUE)

Theorem: Rao-Blackwell Theorem

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

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

① $E[\varphi] = T(\theta)$

② $Var(\varphi) \le Var(\hat{\theta})$

③ φ is a statistic.

i.e. φ is a better est. for $T(\theta)$ than $\hat{\theta}$.

Pf: we already proved ① and ②

lets show ③ : $\hat{\theta} = \theta(\underset{\sim}{X})$ and $W = W(\underset{\sim}{X})$

$$\varphi = E[\hat{\theta}(\underset{\sim}{X}) | W(\underset{\sim}{X})]$$

$$= \int \hat{\theta}(\underset{\sim}{x}) f(\underset{\sim}{x} | w) \, d\underset{\sim}{x}$$

$\left[E[X | Y = y] = \int x f(x|y) dx \right.$

$\left. E[g(X) | Y = y] = \int g(x) f(x|y) dy \right]$

however W is sufficient for θ

by defn $f(X | W)$ is free of θ

$\rightarrow \hat{\theta}$ free of θ, $f(\underset{\sim}{X} | W)$ free of θ

$\Rightarrow$ So whole integral is free of θ

$\Rightarrow$ so φ is free of θ.

i.e. φ is a Stat!