Supplemental material (writing on the white board during lecture in week 5)

1) Example that sometime unbiased estimator may not be very useful:

$f(x,\theta) = \theta(1-\theta)^{x-1}$, $x = 1, 2, \dots$. $T(x)$ based on $n=1$ observation, $\tau(\theta)=\theta$. The unbiasedness means: $\sum_{x=1}^{\infty} T(x)\theta(1-\theta)^{x-1} = \theta \ \forall\theta$. Cancel θ, set $\eta = 1-\theta \in (0,1)$. Then:

$$T(1) + \eta T(2) + \eta^2 T(3) + \dots = 1 \quad \forall \eta \in (0,1).$$

Hence $T(1) = 1$ and $T(2) = T(3) = T(4) = \dots = 0$

Note: estimator $\tilde{T}(x) = \frac{1}{x}$ (which happens to be the MLE in this example) is much more useful.

2) Example illustrating that when condition (*) is violated we could have estimators which are unbiased and have a variance < CR Bound:

$X_1, X_2, \dots, X_n$ i.i.d Uniform $[0,\theta)$ (so the support depends on θ and (*) is violated)

$$f_{X_1}(x) = \begin{cases} 1/\theta, & 0 < x \theta \\ 0 & \text{else} \end{cases} \rightarrow F_{X_1}(x) = \begin{cases} 0 & x \le 0 \\ \frac{x}{\theta} & 0 < x < \theta \\ 1 & x \ge \theta \end{cases}$$

Take $X_{(n)}$. $F_{X_{(n)}}(y) = P(X_{(n)} < y) = P(X_1 < y \cap X_2 < y \cap \dots X_n < y)$

$= (P(X_1 < y))^n = (y/\theta)^n$ when $0 < y < \theta$

Hence $$f_{X_{(n)}}(y) = \begin{cases} \frac{ny^{n-1}}{\theta^n}, & 0<y<\theta \\ 0 & \text{else} \end{cases}$$

Take $EX_{(n)} = \int_0^\theta y \frac{ny^{n-1}}{\theta^n} dy = \frac{n}{n+1}\theta \neq \theta$, ②

i.e. $X_{(n)}$ is biased for estimating θ. <u>BUT:</u>

$T = \frac{n+1}{n} X_{(n)}$ is <u>unbiased</u> for estimating θ.

$$Var\,T = E(T^2) - \theta^2 = \left(\frac{n+1}{n}\right)^2 \int_0^\theta y^2 \frac{ny^{n-1}}{\theta^n} dy - \theta^2 = \dots = \frac{\theta^2}{n(n+1)}$$

But for $f_{X_1}(\theta) = \frac{1}{\theta}, 0 < x < \theta$ we have $\ln f_{X_1}(\theta) = -\ln\theta$

$\frac{\partial}{\partial\theta}\ln f_{X_1}(\theta) = -\frac{1}{\theta}$ $\quad E\left(\frac{\partial}{\partial\theta}\ln f_{X_1}(\theta)\right)^2 = \frac{1}{\theta^2}$ and

reckless application of CR Bound would imply

CR Bound $= \frac{\theta^2}{n}$. As we see now:

$$Var(T) = \frac{\theta^2}{n(n+1)} < \frac{\theta^2}{n}$$

Reason for this seeming contradiction: the condition $(*)$ was <u>violated</u> in this example.

3) Score function for the Poisson(θ) example:

$X_1, X_2, \dots, X_n$ i.i.d. Poisson(θ)

$$P(X_i = x) = \frac{e^{-\theta}\theta^x}{x!}, x = 0, 1, 2, \dots$$

$$L(\mathbf{X}, \theta) = \frac{e^{-n\theta}\theta^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$ Hence

$$V(\mathbf{X}, \theta) = \frac{\partial}{\partial\theta}\log L(\mathbf{X}, \theta) = -n + \frac{\sum_{i=1}^n X_i}{\theta}$$

If $\tau(\theta) = \theta \to V(\mathbf{X}, \theta) = \frac{n}{\theta}(\bar{X} - \theta) \to$ factorization possible and $\bar{X}$ is the UMVUE of θ that attains the CR Bound

However, if $\tau(\theta) = e^{-\theta} = P(X_1 = 0)$

then $V(\mathbf{X}, \theta) = n e^{\theta}\left(\frac{1}{\theta} e^{-\theta} \bar{x} - e^{-\theta}\right)$

This is <u>NOT</u> a statistic (depends explicitly on the unknown θ (!)) hence CR Bound is <u>not</u> attainable by any unbiased estimator of $e^{-\theta}$.

<u>However</u>, as will be seen in week 6, $\left(1 - \frac{1}{n}\right)^{n\bar{x}}$ is an unbiased estimator of $\tau(\theta) = e^{-\theta}$ that <u>is</u> the <u>UMVUE</u> (just that its variance, even if the smallest possible, is $>$ than the quantity $\frac{(\tau'(\theta))^2}{I_{\mathbf{X}}(\theta)}$ given by the bound).

4) Rao - Blackwell Theorem - proof:

i) $E\hat{\tau}(T) = E_T(E(W|T)) \underset{\text{iterative property of expected value}}{=} EW = \tau(\theta)$

(hence $\hat{\tau}(T)$ is unbiased for $\tau(\theta)$).

ii) We show that "always" $Var(Y|X) \le Var\, Y$ (i.e. the variance after conditioning is <u>never</u> increased):

Let $a(X) = E(Y|X)$. Then:

$$Var\, Y = E(Y - EY \pm a(X))^2 = E(Y - a(X))^2 + E(a(X) - EY)^2 + 2E[(Y - a(X))(a(X) - EY)]$$

Next:

$$E(Y-a(X))(a(X)-EY) \overset{\text{iterative property}}{=} E_X\{E((Y-a(X))(a(X)-EY)|X\}$$

$$= E_X\{(a(X)-EY)\, E(Y-a(X)|X)\} =$$

$$= E_X\{(a(X)-EY)(\underbrace{E(Y|X)}_{=a(X)} - a(X))\} = 0$$

Hence $\operatorname{Var} Y = E(Y-a(X))^2 + E(a(X)-EY)^2 \geq$

$$\geq E(a(X)-EY)^2 = E(a(X)-E(a(X)))^2 =$$

$= \operatorname{Var}(a(X))$

i.e. $\operatorname{Var}(Y|X) \leq \operatorname{Var} Y$