There are 2 ways to use this

1) if I have an unbiased $\delta(X)$ and can see how to do Rao-Blackwellization, I can find the UMVUE
2) if I simply recognize a function of $T(X)$ that is unbiased, it must be UMVU

Example $X_1, X_2, \ldots, X_n$ iid Poisson λ, $\lambda > 0$
$\gamma(\lambda) = \lambda$ This is an exponential family and it's easy to see that the natural parameter space contains an open rectangle — any unbiased function of $\sum X_i$ is UMVU

$$\bar{X} = \frac{1}{n}\sum X_i \quad \text{has} \quad E_\lambda \bar{X} = \lambda$$

so $\bar{X}$ is the UMVUE of λ

day 24

Example $X_1, X_2, \ldots, X_n$ iid $N(\mu, 1)$

$$\gamma(\mu) = P_\mu[X_1 < c] = \Phi(c - \mu)$$

$$\delta(X) = I[X_1 < c] \text{ is unbiased for } \gamma(\mu)$$

$$T(X) = \sum X_i \text{ is the natural sufficient statistic}$$

here, so

$$E[\delta(X) \mid T(X)] \text{ is the UMVUE of } \gamma(\mu)$$

It's slightly more convenient but equivalent to work with $T'(X) = \bar{X} = \frac{1}{n}T(X)$ and

$$E\left[I[X_1 < c] \mid T'(X)\right] = P\left[X_1 < c \mid T'(X)\right]$$

$$\begin{pmatrix} X_1 \\ \bar{X} \end{pmatrix} \sim MVN_2 \left(\begin{pmatrix} \mu \\ \mu \end{pmatrix}, \begin{pmatrix} 1 & \frac{1}{n} \\ \frac{1}{n} & \frac{1}{n} \end{pmatrix} \right)$$

and given $\bar{X}$, $X_1 \sim N(\bar{X}, \frac{n-1}{n})$

$\left(\text{given } T(X) \quad X_1 \sim N(\frac{1}{n}T(X), \frac{n-1}{n}) \right)$

So $$P[X_1 < c \mid \bar{X}] = \Phi\left(\frac{c - \bar{X}}{\sqrt{\frac{n-1}{n}}}\right)$$

is the UMVUE of $\gamma(\mu)$

Lest we get too carried away with enthusiasm

Example $X \sim \text{Poisson}(\lambda)$

$\gamma(\lambda) = e^{-2\lambda} = P[\text{2 independent Poisson}(\lambda) \text{ r.v.'s are both } 0]$

X is the obvious natural sufficient statistic — one can check that

$$\delta(X) = (-1)^X = \begin{cases} -1 & \text{if } X \text{ is odd} \\ 1 & \text{if } X \text{ is even} \end{cases}$$

is unbiased — it's a function of X and $\therefore$ the UMVUE — it is, in fact, the only unbiased estimator of $\gamma(\lambda)$ — but it's clearly silly — even

$$\delta^*(X) = I[X \text{ is even}]$$

is better (since $e^{-2\lambda} > 0$) — Meeden used to call such estimators RUBEs ☺

(ridiculous unbiased estimator(s))

The Third bit of optimality Theory for non-Bayes point estimation is Cramér-Rao - It provides a lower bound on the variance of an estimator

Thm (Cramér-Rao) If the model for X specified by $f(x|\theta)$ is FI regular at $\theta_0 \in R^1$ (regularity/differentiability etc. conditions hold at $\theta = \theta_0$) and $0 < I(\theta_0) < \infty$ and

$$E_\theta \delta(X) = \int \delta(x) f(x|\theta)\,dx \text{ or } \sum_x \delta(x) f(x|\theta)$$

can be differentiated under the integral (sum) sign at θ_0 $\left(\frac{d}{d\theta} E_\theta \delta(X) \Big|_{\theta=\theta_0} = \int \delta(x) \frac{d}{d\theta} f(x|\theta) \Big|_{\theta=\theta_0} dx \right)$

Then

$$\mathrm{Var}_{\theta_0} \delta(X) \geq \frac{\left(\frac{d}{d\theta} E_\theta \delta(X) \Big|_{\theta=\theta_0} \right)^2}{I(\theta_0)}$$

Notice that in the case that $E_\theta \delta(X) = \theta$ and thus that $\frac{d}{d\theta} E_\theta \delta(X) = 1 \; \forall \theta$ This promises that

$$\mathrm{Var}_\theta \, \delta(X) \geq \frac{1}{I(\theta)}$$

and we have a lower bound on the variance of an unbiased estimator of θ That is simply the reciprocal of the FI in X about θ

Pf (of Cramér–Rao)

$$\frac{d}{d\theta} E_\theta \delta(X) \Big|_{\theta=\theta_0} = \int \delta(x) \frac{d}{d\theta} f(x|\theta) \Big|_{\theta=\theta_0} dx$$

$$= \int \delta(x) \frac{\frac{d}{d\theta} f(x|\theta)\big|_{\theta=\theta_0}}{f(x|\theta_0)} f(x|\theta_0)\, dx$$

$$= E_{\theta_0}\left(\delta(X) \frac{d}{d\theta} \log f(X|\theta) \Big|_{\theta=\theta_0} \right)$$

$$= \mathrm{Cov}_{\theta_0}\left(\delta(X), \frac{d}{d\theta} \log f(X|\theta) \Big|_{\theta=\theta_0} \right)$$

since $0 = E_0\left(\frac{d}{d\theta} \log f(X|\theta) \Big|_{\theta=\theta_0} \right)$

Then, since $(\mathrm{Cov}(U,V))^2 \le \mathrm{Var}\, U \;\mathrm{Var}\, V$
(absolute correlations are not larger than 1)

So

$$\left(\frac{d}{d\theta} E_\theta \delta(X) \Big|_{\theta=\theta_0} \right)^2 \le \mathrm{Var}_{\theta_0} \delta(X) \cdot \underbrace{\mathrm{Var}_{\theta_0} \frac{d}{d\theta} \log f(X|\theta) \Big|_{\theta=\theta_0}}_{I(\theta_0)}$$

■

Example $X \sim Bi(n,p)$

C-R says that any unbiased estimator of p has

$$\mathrm{Var}_p\, \delta(X) \ge \frac{1}{n\left(\frac{1}{p(1-p)}\right)} = \frac{p(1-p)}{n}$$

Since $\mathrm{Var}_p \frac{X}{n} = \frac{p(1-p)}{n}$ provides another proof
That $\frac{X}{n}$ is the UMVUE of p

<u>Example</u> showing the CR lower bound may not be attained (actually it is usually <u>not</u> attained)

Suppose $X \sim Bi(n, p)$ and $\gamma(p) = p^2$ is of interest – the CR lower bound is

$$Var_p\, \delta(X) \geq \frac{(2p)^2}{\frac{n}{p(1-p)}} = \frac{4p^3(1-p)}{n}$$

As a matter of fact

$$\delta(X) = \frac{X(X-1)}{n(n-1)}$$

has
$$E_p\, \delta(X) = \frac{1}{n(n-1)}\left[np(1-p) + (np)^2 - np \right] = p^2$$

and is (as it turns out) the UMVUE of p^2 (using Lehmann-Scheffé) — but

$$Var_p\, \delta(X) > \frac{4p^3(1-p)}{n} \quad \forall\, p$$

When is CR going to be attained? Well basically when

$$\delta(X) \text{ is perfectly correlated with } \left.\frac{d}{d\theta} \log f(X|\theta)\right|_{\theta=\theta_0}$$

under $f(x|\theta_0)$ – i.e. $\delta(X)$ is essentially the derivative of a log likelihood – when does that seem plausible? in exponential families!