Stat 543 3-7-05

Recall: Cramér-Rao

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

Example (to show the inequality can't always be attained) $X \sim Bi(n, p)$

for unbiased estimation of p^2 the C-R lower bound is

$$\frac{4p^3(1-p)}{n}$$

but $\delta(X) = \frac{X(X-1)}{n(n-1)}$ has $E_p\delta(X) = p^2$

and is the UMVUE of p^2 and has variance

Strictly larger than the Cramér-Rao lower bound

FI, Cramér-Rao and Exponential Families

In B+D form $X \sim f(x|\eta) = h(x) \exp[\eta T(x) - A(\eta)]$

one-dimensional

$$\ln f(x|\eta) = \ln h(x) + \eta T(x) - A(\eta)$$

$$\frac{d}{d\eta}(\quad\quad) = T(x) - \frac{d}{d\eta}A(\eta)$$

$$I(\eta_0) = \mathrm{Var}_{\eta_0} T(X)$$

$$\frac{d^2}{d\eta^2} \ln f(x|\eta) = -\frac{d^2}{d\eta} A(\eta)$$

So (using the 2nd form of FI)

$$- E_{\eta_0} \frac{d^2}{d\eta^2} \ln f(x|\eta) \Big|_{\eta=\eta_0} = \frac{d^2}{d\eta^2} A(\eta) \Big|_{\eta=\eta_0}$$

By Corollary 1.6.1 page 59 of B&D if $\mathcal{E}$ has a nonempty interior this is

$$\mathrm{Var}_{\eta_0} T(X) = I(\eta_0)$$

Consider what C-R says about $T(X)$

$$\mathrm{Var}_\eta T(X) \geq \frac{\left(\frac{\partial}{\partial \eta} E_\eta T(X)\right)^2}{I(\eta)} = \frac{\left(\frac{\partial}{\partial \eta}\left(\frac{\partial}{\partial \eta} A(\eta)\right)\right)^2}{I(\eta)}$$

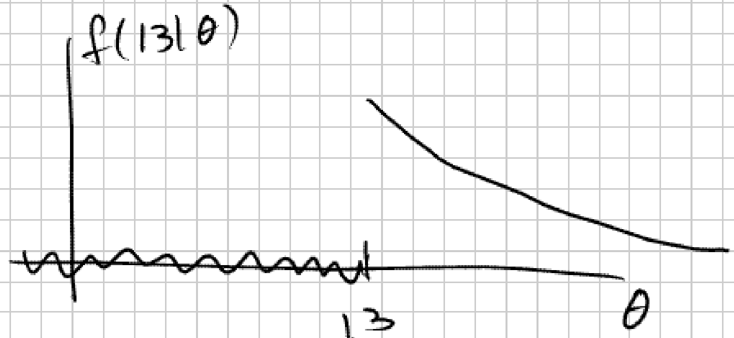
$$= \frac{\left(\frac{d^2}{d\eta^2} A(\eta)\right)^2}{I(\eta)}$$

$$= \frac{(I(\eta))^2}{I(\eta)} = I(\eta)$$

That is, treated as an unbiased estimator of $E_\eta T(X)$, $T(X)$ has variance $I(\eta)$ = Cramér-Rao lower bound ($T(X)$ is UMVUE of its mean ....) —

As it turns out, only natural sufficient statistics in exponential families end up achieving the C-R lower bound — see Thm 3.4.2 of B&D

BTW ... the regularity conditions in careful statement of this material (FI regularity in tight definition of FI) are meant to ensure that for any x, $f(x|\theta)$ changes smoothly in θ – e.g. they outlaw models like $X \sim U(0,\theta)$ ... for $x = 13$



There is a discontinuity at $\theta = 13$, besides, take any x $f(x/\theta)$ is not positive for all θ, there is no FI, $I(\theta)$

Jump now to Ch 4 of B+D and Theory for hypothesis testing - scenario

$$\Theta = \Theta_0 \sqcup \Theta_1$$

and we wish to decide between $H_0: \theta \in \Theta_0$ and $H_1: \theta \in \Theta_1$ based on $X \sim f(x|\theta)$

$a = \{0, 1\}$

0-1 loss

$$L(\theta, a) = I[\theta \in \Theta_0]\, I[a=1] + I[\theta \in \Theta_1]\, I[a=0]$$

$\phi(x)$ a decision rule / a "test"

$\phi(x) = 0$ means "accept H_0"

$\phi(x) = 1$ means "reject H_0"

$\{x \mid \phi(x) = 1\}$ = rejection region

$\{x \mid \phi(x) = 0\}$ = acceptance region

$$R(\theta, \phi) = E_\theta L(\theta, \phi(X))$$

$$= I[\theta \in \Theta_0] \, E_\theta I[\phi(X) = 1]$$

$$+ I[\theta \in \Theta_1] \, E_\theta I[\phi(X) = 0]$$

$$= I[\theta \in \Theta_0] \, P_\theta[\phi(X) = 1]$$

$$+ I[\theta \in \Theta_1] \left(1 - P_\theta[\phi(X) = 1]\right)$$

The power function for the test

i.e. define $\pi(\theta) = P_\theta[\phi(X) = 1]$ the "power function" of the test

$$R(\theta, \phi) = I[\theta \in \Theta_0]\,\pi(\theta) + I[\theta \in \Theta_1](1 - \pi(\theta))$$

I want small risk ... I want small $\pi(\theta)$ when $\theta \in \Theta_0$ and large $\pi(\theta)$ when $\theta \in \Theta_1$

How to come up with tests that do this??

There is some optimality theory and heuristics motivated by optimality thy (2nd for real/big problems)