Stat 543 3-4-05

Recall : Fisher Information

$L(\theta)$
$\ell(\theta) = \ln L(\theta)$

Def : $I(\theta_0) = E_{\theta_0}(\ell'(\theta_0))^2 = Var_{\theta_0}(\ell'(\theta_0))$

Thm : Often $I(\theta_0) = -E_{\theta_0}(\ell''(\theta_0))$

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

$I(p) = \frac{n}{p(1-p)}$ from definition

$\ell''(p) = \frac{X}{p^2} - \frac{(n-X)}{(1-p)^2}$

$$E_{p_0}\, \ell''(p_0) = -\frac{n p_0}{p_0^2} - \frac{n(1-p_0)}{(1-p_0)^2}$$

$$= -n\left(\frac{1}{p_0} + \frac{1}{1-p_0}\right)$$

$$= -\frac{n}{p_0(1-p_0)} = -I(p_0)$$

A second useful property of FI is that for models involving independent it's <u>additive</u> — i.e. if

$$X = (X_1, \ldots, X_n)$$

has a dsn depending upon θ that makes $X_1, \ldots, X_n$ independent

$$I_X(\theta) = \sum_{i=1}^{n} I_{X_i}(\theta)$$

Why? Take, for example, an iid problem with marginal pdf or pmf $f(x|\theta)$

$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$

$$\ell(\theta) = \sum_{i=1}^{n} \log f(X_i|\theta)$$

$$\ell'(\theta) = \sum_{i=1}^{n} \frac{\frac{d}{d\theta} f(X_i|\theta)}{f(X_i|\theta)}$$

for any θ this is a sum of iid r.v.'s under any θ_0

$$\underset{\substack{\| \\ I_X(\theta_0)}}{\mathrm{Var}_{\theta_0}\, \ell'(\theta_0)} = \sum_{i=1}^{n} \mathrm{Var}_{\theta_0} \frac{\frac{d}{d\theta} f(X_i|\theta)\Big|_{\theta=\theta_0}}{f(X_i|\theta_0)}$$

$$= \sum_{i=1}^{n} I_{X_i}(\theta_0) = n I_{X_1}(\theta_0)$$

<u>Example</u> $X_1, X_2, \ldots, X_n$ iid Bernulli(p)

$$I_{X_1}(p) = \frac{1}{p(1-p)}$$

$$I_X(p) = \frac{n}{p(1-p)}$$

FI in X

FI in $\sum X_i$

Bin(n, p)

<u>Thm</u> (Cramér-Rao) If the model for X specified by $f(x|\theta)$ is FI regular at $\theta_0 \in \mathbb{R}^1$ and $0 < I(\theta_0) < \infty$ and

$$E_\theta \delta(X) \underset{\text{cont}^{\underline{s}}}{=} \int \delta(x) f(x|\theta)\, dx$$

can be differentiated under the integral sign at θ_0

i.e.

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

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 what this say about unbiased estimators of θ

$\delta(X)$ unbiased for θ means

$$E_\theta \delta(X) = \theta \qquad \left(\frac{d}{d\theta} E_\theta \delta(X) = 1 \right)$$

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

and this is potentially another way to argue that an estimator is UMVU

<u>pf</u>:

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

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

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

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

remember that $\left(\mathrm{Cov}(U,V)\right)^2 \le \mathrm{Var}\,U \;\mathrm{Var}\,V$

because a correlation can be no larger that 1 in absolute value — so

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

QED

<u>Example</u> $X \sim Bi(n,p)$ C-R says that any unbiased estimator of p has variance at least

$$\frac{1}{I_X(p)} = \frac{1}{\frac{n}{p(1-p)}} = \frac{p(1-p)}{n}$$

Note that $\frac{X}{n}$ achieves this lower bound and is unbiased for p $\therefore$ UMVU for p

<u>Example</u> (to show that C-R can not always be achieved) (actually "usually" it can not be achieved)

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

consider unbiased estimation of p^2 – C-R promises than any unbiased estimator of p^2 has variance at least

$$\frac{(2p)^2}{\frac{n}{p(1-p)}} = \frac{4p^3(1-p)}{n}$$