Stat 543 2-28-05

Recall: 3rd piece of optimality thy for point estimation ... Cramér-Rao Inequality (based on "Fisher Information") ... Handout for $\Theta \subset \mathbb{R}^k$ and technical details ...

$L(\theta)$ likelihood

$\ell(\theta) = \log L(\theta)$ loglikelihood

score function $= \ell'(\theta) = \frac{d}{d\theta}\ell(\theta)$

Note:
$$E_{\theta_0}\ell'(\theta_0) = E_{\theta_0}\frac{d}{d\theta}\log f(X|\theta)\Big|_{\theta=\theta_0}$$

cont's case

$$\overset{\text{cont's case}}{=} \int \frac{\frac{d}{d\theta} f(x|\theta)\Big|_{\theta=\theta_0}}{f(x|\theta_0)} f(x|\theta_0)\, dx$$

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

AIP

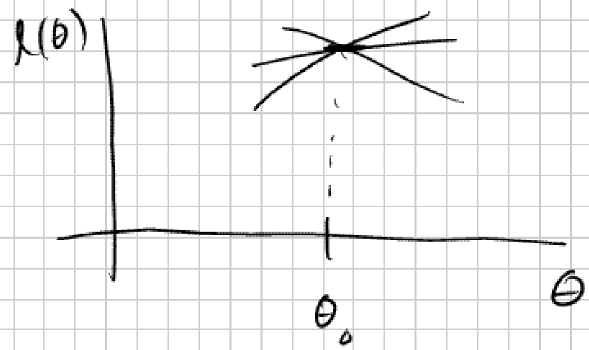
$$\overset{\text{AIP}}{=} \frac{d}{d\theta} \int f(x|\theta)\, dx \Big|_{\theta=\theta_0}$$

$$= \frac{d}{d\theta} 1 \Big|_{\theta=\theta_0}$$

$$= 0$$

The mean (θ_0) of the score function at θ_0 is 0 –

on (θ_0) average the log-likelihood is "flat" at θ_0



The happiest circumstance is that when the (random) $\ell'(\theta_0)$ is big in magnitude ... Why? because that says that the loglikelihood is climbing sharply at θ_0 or is dropping sharply at θ_0 i.e. values close to θ_0 are distinguished ... so I want big

$$E_{\theta_0}\left(\ell'(\theta_0)\right)^2$$

$$E_{\theta_0}\left(\ell'(\theta_0)\right)^2 = I_X(\theta_0) = \mathrm{Var}_{\theta_0}\, \ell'(\theta_0)$$

is the Fisher Information in X about θ at θ_0

(The multidimensional version of this is that the $k \times k$ Fisher Information matrix is the variance-covariance matrix of the score function)

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

$$L(p) = \binom{n}{X} p^X (1-p)^{n-X}$$

$$\ell(p) = \log\binom{n}{X} + X \log p + (n-X)\log(1-p)$$

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

$$E_{p_0}\, \ell'(p_0) = \frac{np_0}{p_0} - \frac{n - np_0}{1-p_0} = 0$$

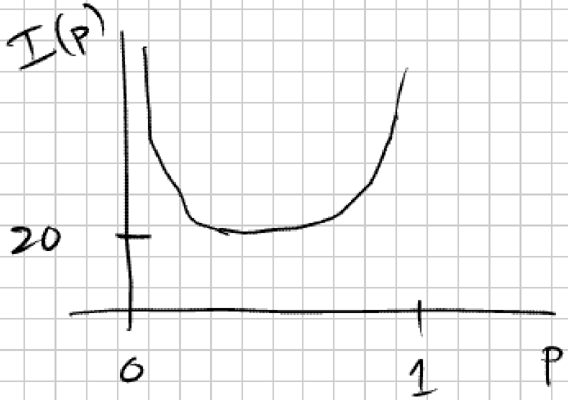
$$I(p_0) = E_{p_0}\left(\ell'(p_0)\right)^2$$

$$= \mathrm{Var}_{p_0}\left[X\left(\frac{1}{p_0} + \frac{1}{1-p_0}\right) - \frac{n}{1-p_0}\right]$$

$$= \left(\mathrm{Var}_{p_0} X\right)\left(\frac{1}{p_0(1-p_0)}\right)^2$$

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

To your collection of plots for the $Bi(5,p)$ problem you could add a plot of $I(p)$



$$I(p) = \frac{5}{p(1-p)}$$

There are a number of useful simple results about FI

Result Under appropriate regularity conditions

$$I(\theta_0) = -E_{\theta_0}\left(\ell''(\theta_0)\right)$$

The information in X about θ evaluated at θ_0 is not only the θ_0 variance of the score function evaluated at θ_0, it is <u>also</u> the θ_0 expected curvature of the log likelihood at θ_0 — The curved The loglikelihood tends to be, the more discriminating is the loglikelihood and the better is one's information about θ

"Pf" Outline for the cont's case

$$l''(\theta) = \frac{d}{d\theta}\left(\frac{\frac{d}{d\theta}f(x|\theta)}{f(x|\theta)}\right)$$

$$= \frac{f(x|\theta)\frac{d^2}{d\theta^2}f(x|\theta) - \left(\frac{d}{d\theta}f(x|\theta)\right)^2}{\left(f(x|\theta)\right)^2}$$

So $$E_{\theta_0} l''(\theta_0) = \int \frac{d^2}{d\theta^2} f(x|\theta)\Big|_{\theta=\theta_0} dx$$

$$- \int \frac{\left(\frac{d}{d\theta} f(x|\theta) \Big|_{\theta=\theta_0} \right)^2}{(f(x|\theta_0))^2} f(x|\theta_0)\, dx$$

AIP

$$\stackrel{\circ}{=} \frac{d^2}{d\theta^2} \int f(x|\theta)\, dx \Big|_{\theta=\theta_0} - \underbrace{\int \left(\frac{d}{d\theta} \log f(x|\theta) \Big|_{\theta=\theta_0} \right)^2 f(x|\theta)\, dx}_{E_{\theta_0} (l'(\theta_0))^2}$$

$$= 0 - I(\theta_0)$$