Continue Example

A reasonable estimator ....

$$\hat{\theta}(y) = y$$

$$\Rightarrow E_\theta\{\hat{\theta}(Y)\} = \theta, \quad Var_\theta\{\hat{\theta}(Y)\} = \sigma^2$$

To connect to CRLB

$$\ln f_\theta(y) = -\ln\sqrt{2\pi\sigma^2} - \frac{1}{2\sigma^2}(y-\theta)^2$$

$$\frac{\partial}{\partial\theta}\ln f_\theta(y) = \frac{1}{\sigma^2}(y-\theta)$$

$$\frac{\partial^2}{\partial\theta^2}\ln f_\theta(y) = -\frac{1}{\sigma^2} \Rightarrow \frac{1}{\sigma^2} = \underbrace{-\frac{\partial^2}{\partial\theta^2}\ln f_\theta(y)}_{\text{curvature.}}$$

Larger curvature $\Longleftrightarrow$ sharper $\Longleftrightarrow$ smaller σ^2

Thus, in this case

$$\mathrm{Var}_\theta\left\{\hat{\theta}(Y)\right\} = \frac{1}{-\frac{\partial^2}{\partial\theta^2}\ln f_\theta(y)} = \sigma^2$$

formula

Note: In this case this does not depend. But in general, it would, and since the Var. of an estimator cannot dep on $Y=y$ .... we should compute an average.

ie $\quad -E_\theta\left\{\frac{\partial^2}{\partial\theta^2}\ln f_\theta(Y)\right\}$

Information Inequality / CRLB Theorem

$\hat{\theta}$ an estimator of parameter θ in a family $\{P_\theta : \theta \in \Lambda\}$ modeling observations $Y \in \Gamma$. Suppose:

① Λ is an open interval

② $\{P_\theta\}$ has corresp. pdfs $\{f_\theta : \theta \in \Lambda\}$ all of which have the same support.

③ $\frac{\partial f_\theta(y)}{\partial \theta}$ exists and is finite $\forall\, \theta \in \Lambda$ and all y in support of f_θ

④ $\frac{\partial}{\partial \theta} \int_\Gamma h(y) f_\theta(y)\, dy$ exists and equals $\int_\Gamma h(y) \left[\frac{\partial f_\theta(y)}{\partial \theta} \right] dy$

$\forall\, \theta$ and for $h(y) = \hat{\theta}(y)$ and $h(y) \equiv 1$.

Then: $$\mathrm{Var}_\theta\{\hat{\theta}(Y)\} \geq \frac{\left(\frac{\partial}{\partial \theta} E_\theta\{\hat{\theta}(Y)\} \right)^2}{I_\theta}$$

where $I_\theta \triangleq E_\theta\left\{ \left(\frac{\partial}{\partial \theta} \log f_\theta(Y) \right)^2 \right\}$.

I. I. / CRLB (cont'd.)

Furthermore if also have

(5) $\dfrac{\partial^2 f_\theta(y)}{\partial \theta^2}$ exists $\forall\, \theta \in \Lambda$ and $y \in$ support of f_θ

and
$$\int_\Gamma \left(\frac{\partial^2}{\partial \theta^2} f_\theta(y) \right) dy = \frac{\partial^2}{\partial \theta^2} \int_\Gamma f_\theta(y)\, dy$$

Then I_θ can be computed via

$$I_\theta = -E_\theta \left\{ \frac{\partial^2}{\partial \theta^2} \log f_\theta(Y) \right\}.$$

Proof Start by differentiating

$$\frac{\partial}{\partial\theta}\left\{ E_\theta\{\hat{\theta}(Y)\} = \int_\Gamma \hat{\theta}(y) f_\theta(y)\, dy \right\}$$

so using ④ ...

$$\text{above} = \int_\Gamma \hat{\theta}(y) \frac{\partial}{\partial\theta} f_\theta(y)\, dy$$

apply ④ for $h(y) = \hat{\theta}(y)$.

Also apply ④ for $h(y) = 1$ ....

$$\frac{\partial}{\partial\theta}\int_\Gamma f_\theta(y)\, dy = \frac{\partial}{\partial\theta} 1 = 0 = \int_\Gamma \frac{\partial}{\partial\theta} f_\theta(y)\, dy$$

Therefore

$$\frac{\partial}{\partial \theta} E_\theta\{\hat{\theta}(Y)\} = \int_\Gamma \left[\hat{\theta}(y) - E_\theta\{\hat{\theta}(Y)\}\right] \frac{\partial}{\partial \theta} f_\theta(y)\, dy$$

To get the form needed ....

$$\frac{\partial}{\partial \theta} \log f_\theta(y) = \frac{1}{f_\theta(y)} \frac{\partial}{\partial \theta} f_\theta(y)$$

$$\Rightarrow \frac{\partial}{\partial \theta} f_\theta(y) = \left(\frac{\partial}{\partial \theta} \log f_\theta(y)\right) \cdot f_\theta(y)$$

Plugging in as indicated ....

$$\frac{\partial}{\partial \theta} E_\theta\{\hat{\theta}(Y)\} = E_\theta\Big\{\underbrace{\left[\hat{\theta}(Y) - E_\theta\{\hat{\theta}(Y)\}\right]}_{U} \underbrace{\frac{\partial}{\partial \theta} \log f_\theta(Y)}_{V}\Big\}$$

We will apply Cauchy-Schwartz ...

U, V are random variables.

$$\left(E\{UV\}\right)^2 \le E\{U^2\} \cdot E\{V^2\}.$$

Now ...

$$\left[\frac{\partial}{\partial\theta} E_\theta\{\hat{\theta}(Y)\}\right]^2 \le \underbrace{E_\theta\left\{\left[\hat{\theta}(Y) - E_\theta\{\hat{\theta}(Y)\}\right]^2\right\}}_{Var_\theta\{\hat{\theta}(Y)\}} \cdot \underbrace{E_\theta\left\{\left[\frac{\partial}{\partial\theta}\log f_\theta(Y)\right]^2\right\}}_{I_\theta}$$

$$\Rightarrow Var_\theta\{\hat{\theta}(Y)\} \ge I_\theta^{-1}\left[\frac{\partial}{\partial\theta} E_\theta\{\hat{\theta}(Y)\}\right]^2$$

If unbiased ie $E_\theta\{\hat{\theta}(Y)\} = \theta$ then

$$\mathrm{Var}\{\hat{\theta}(Y)\} \geq \frac{1}{I_\theta}$$

This is the usual formulation of CRLB.

The final statement ie

$$I_\theta = -E_\theta\left\{\frac{\partial^2}{\partial\theta^2} \log f_\theta(Y)\right\}.$$

is easy by taking another deriv. and rearranging.

$I_\theta \triangleq$ Fisher's Information.

It is <u>not</u> always possible to have equality in the CRLB.

* More information (ie larger I_θ) provided on average by an observation Y tends to mean that a good estimator performs better.

* I_θ has another property expected by an info measure ....

$Y = [Y_1\, Y_2 \cdots Y_n]$ $\quad Y_k$ iid with $f^*_\theta(\cdot)$

$$\Rightarrow \quad f_\theta(y) = f_\theta(y_1\, y_2 \cdots y_n)$$
$$= f^*_\theta(y_1)\, f^*_\theta(y_2) \cdots f^*_\theta(y_n).$$

$$\Rightarrow \log f_\theta(y) = \sum_{k=1}^{n} \log f^*_\theta(y_k).$$

Putting Y back in place of y ...

$$\log f_\theta(Y) = \sum_{k=1}^{n} \log f^*_\theta(Y_k)$$

a sum of iid rvs. each of zero mean. (for same reason true before).

$$I_\theta = E_\theta\left\{\left(\frac{\partial}{\partial\theta}\log f_\theta(Y)\right)^2\right\}$$

$$= \mathrm{Var}_\theta\left\{\frac{\partial}{\partial\theta}\log f_\theta(Y)\right\}$$

$$= n\,\mathrm{Var}_\theta\left\{\frac{\partial}{\partial\theta}\log f^*_\theta(Y_1)\right\}.$$

$$= n\,E_\theta\left\{\left(\frac{\partial}{\partial\theta}\log f^*_\theta(Y_1)\right)^2\right\}$$

$$= n\,i_\theta$$

Overall Fisher info is sum of indiv. parts. "Additivity prop" .... Common to all meas. of info.