## ECE 7251: Signal Detection and Estimation


Spring 2002

Prof. Aaron Lanterman

Georgia Institute of Technology


Lecture 9, 1/25/02:

Lord of the Cramér - Rao Bound

## Quest of the Cramer-Rao Bound

*One bound to beat them all,*

*and in the confusion bind them…*

*In the land of Techwood where the Engineers lie.*

- Our Quest: Find an ultimate performance bound which may not be beaten by *any* estimator

$$\mathrm{var}_q[\hat{q}(Y)] \geq \text{some unbeatable bound}$$

## The Fellowship of the CR Bound

- Supposed the estimator is unbiased. Then

$$E_q[(\hat{q}(Y)-q)] = 0$$

$$\int p(y;q)[\hat{q}(y)-q]dy = 0$$

- Take derivative w.r.t. $q$ on both sides:

$$\frac{\partial}{\partial q}\int p(y;q)[\hat{q}(Y)-q]dy = 0$$

$$\int \frac{d}{dq} p(y;q)[\hat{q}(y)-q]dy = 0$$

Should check if interchange is legit; need derivative of density to be "Absolutely integrable"

## Breakfast of the CR Bound

$$\int \frac{d}{dq} p(y;q)[\hat{q}(y)-q]dy = 0$$

$$-\underbrace{\int p(y;q)dy}_{=1} + \int \frac{\partial p(y;q)}{\partial q}[\hat{q}(y)-q]dy = 0$$

$$\int \frac{\partial \ln p(y;q)}{\partial q} p(y;q)[\hat{q}(y)-q]dy = 1$$

## Lunch of the CR Bound

$$\int \frac{\partial \ln p(y;q)}{\partial q} p(y;q)[\hat{q}(y)-q]dy = 1$$

$$\left(\int \underbrace{\left\{\frac{\partial \ln p(y;q)}{\partial q}\sqrt{p(y;q)}\right\}}_{=f(y)} \underbrace{\left\{\sqrt{p(y;q)}[\hat{q}(y)-q]\right\}}_{=g(y)} dy\right)^2 = 1^2$$

- Use the Schwarz, Lone Star!

$$\int f^2(y)dy \int g^2(y)dy \geq \left\{\int f(y)g(y)dy\right\}^2$$

$$\left\{\int\left(\frac{\partial \ln p(y;q)}{\partial q}\right)^2 p(y;q)dy\right\}\left\{\int p(y;q)[\hat{q}(y)-q]^2 dy\right\} \geq 1$$

## Dinner of the CR Bound

$$\left\{\int\left(\frac{\partial \ln p(y;q)}{\partial q}\right)^2 p(y;q)dy\right\}\left\{\int p(y;q)[\hat{q}(Y)-q]^2 dY\right\} \geq 1$$

$$E_q\left[\left(\frac{\partial \ln p(y;q)}{\partial q}\right)^2\right] E_q[(\hat{q}(Y)-q)^2] \geq 1$$

$$\mathrm{var}_q(\hat{q}) = E_q[(\hat{q}(Y)-q)^2] \geq \frac{1}{E_q\left[\left(\frac{\partial \ln p(y;q)}{\partial q}\right)^2\right]}$$

(Since estimator is unbiased)

Fisher Information

## The Two Fisher Info Towers

$$\int p(y;q)dy = 1$$

- Take derivative w.r.t. $q$ on both sides:

$$\int \frac{\partial p(y;q)}{\partial q} dy = \int \frac{\partial \ln p(y;q)}{\partial q} p(y;q) dy = 0$$

- Take derivative again (*assuming* we can!):

$$\int \frac{\partial^2 \ln p(y;q)}{\partial^2 q} p(y;q) + \frac{\partial \ln p(y;q)}{\partial q} \frac{\partial p(y;q)}{\partial q} dy = 0$$

$$\int \frac{\partial^2 \ln p(y;q)}{\partial^2 q} p(y;q) + \left[\frac{\partial \ln p(y;q)}{\partial q}\right]^2 p(y;q) dy = 0$$

## The Two Fisher Info Towers Con't

$$\int \frac{\partial^2 \ln p(y;q)}{\partial^2 q} p(y;q) + \left[\frac{\partial \ln p(y;q)}{\partial q}\right]^2 p(y;q) dy = 0$$

$$F(q) \equiv E_q\left[\left(\frac{\partial \ln p(y;q)}{\partial q}\right)^2\right] = -E_q\left[\frac{\partial^2 \ln p(y;q)}{\partial^2 q}\right]$$

- Different notations for Fisher information:

| | |
|---|---|
| Poor: | $I_q$ |
| Van Trees, Scharf, Srinath: | $J$ |
| Hero: | $F(q)$ |

## Other Termonology

- Observed or empirical Fisher information:

A specific collected data set

$$-\frac{\partial^2 \ln p(y;q)}{\partial^2 q}$$

- Stochastic Fisher information:

$$-\frac{\partial^2 \ln p(Y;q)}{\partial^2 q}$$

A random variable

- Fisher information is expected value of stochastic Fisher information

## Efficiency

- An unbiased estimator which achieves the CR Bound with equality is called efficient
- Efficient estimators are UMVE (but not necessarily the other way around)
- Efficient estimators can only exist if the density comes from an exponential family!
  - Proof uses condition for equality in the Schwarz inequality

$$\frac{\partial \ln p(y;q)}{\partial q} = [\hat{q}(y) - q]k(q)$$

## More Properties of ML Estimators

- If an efficient estimator exists, the ML est. is it!
- Suppose we have $n$ independent samples
  - ML estimators are asymptotically efficient

$$\lim_{n\to\infty} \mathrm{var}_q(\hat{q})F(q) = 1$$

  - ML estimators are asymptotically normal:

$$\hat{q}(Y) \sim N(q, F^{-1}(q))$$

Formally: $\sqrt{n}(\hat{q} - q) \underset{i.d.}{\rightarrow} z, \quad z \sim N(0, F_1^{-1}(q))$

Fisher information matrix for one data sample

## Matrix CR Bound

- Elements of the Fisher information matrix:

$$F_{rc} = E\left[\frac{\partial \ln p(y;q)}{\partial q_r} \frac{\partial \ln p(y;q)}{\partial q_c}\right] = -E\left[\frac{\partial^2 \ln p(y;q)}{\partial q_r \partial q_c}\right]$$

Average "curvature" of loglikelihood near $q$

- Matrix CR Bound for unbiased estimators:

$$\mathrm{cov}_q[\hat{q}(Y)] \geq F^{-1}(q)$$

Means $\mathrm{cov}_q[\hat{q}(Y)] - F^{-1}(q)$ is nonnegative definite

## Nonnegative Definite? Huh???

- If $A$ is nonnegative definite, then:
  - $z^T A z \geq 0$ for any real vector $z$
  - All eigenvalues of A are nonnegative
- Useful consequences: if $A \geq B$,
  - Diagonals dominated: $A_{ii} \geq B_{ii}$
  - But does not mean $A_{rc} \geq B_{rc}$ in general!
  - Total sum property:

$$\sum_{r,c} A_{rc} \geq \sum_{r,c} B_{rc}$$

## Returned of the Biased Estimator

- CR bound for biased estimator:

$$\mathrm{var}_q[\hat{q}(Y)] \geq \frac{\left(\frac{\partial}{\partial q} E_q[\hat{q}(Y)]\right)^2}{F(q)} = \frac{\left(1+\frac{\partial}{\partial q} b_q(\hat{q})\right)^2}{F(q)}$$

- Proof: Poor, pp. 169-171.
- There are extensions for multiparameter case
- Bias often difficult to compute analytically for a particular estimator; hence, the unbiased CR bound is often given anyway!

## What About Those Bayesians?

- A related bound on the MSE:

$$E[(q(Y)-\Theta)^2] \geq \frac{1}{E\left[\left(\frac{\partial \ln p(Y,q)}{\partial q}\right)^2\bigg|_{q=\Theta}\right]} = \frac{1}{-E\left[\frac{\partial^2 \ln p(Y,q)}{\partial^2 q}\bigg|_{q=\Theta}\right]}$$

- Note expectations now over Y *and* Θ
- Easily extended to multivariate case
- Not all authors cover!
  - Hero and Poor do not
  - Van Trees (pp. 72-73, 74-85) and Srinath (pp. 163) do

## Interpretations of the Bayesian Bound

- A related bound on the MSE:

Note not a func. of $q$

$$F_{Post} = -E\left[\frac{\partial^2 \ln p(Y,q)}{\partial^2 q}\bigg|_{q=\Theta}\right]$$

$$= -E\left[\frac{\partial^2 \ln p(Y|q)}{\partial^2 q}\bigg|_{q=\Theta}\right] - E\left[\frac{\partial^2 \ln p(Y|q)}{\partial^2 q}\bigg|_{q=\Theta}\right]$$

- Info from data adds with info from prior
- To be efficient (achieve bound with equality) in the Bayesian setting, posterior density *must* be Gaussian!
  - Proof uses condition of equality in Schwarz ineq.
  - Stronger requirement than in nonrandom setting