Alternate Form for CRLB

\[ \text{var}(\hat{\theta}) \geq \frac{1}{\theta^2 E \left[ \left( \frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right]} \]

See Appendix 3A for Derivation

Sometimes it is easier to find the CRLB this way.

This also gives a new viewpoint of the CRLB:

From Gardner’s Paper (IEEE Trans. on Info Theory, July 1979)

Consider the Normalized version of this form of CRLB

\[ \frac{\text{var}(\hat{\theta})}{\theta^2} \geq \frac{1}{\theta^2 E \left[ \left( \frac{\partial \ln p(x; \theta)}{\partial \theta} \right)^2 \right]} \]

We’ll “derive” this in a way that will re-interpret the CRLB.
Consider the “Incremental Sensitivity” of \( p(x; \theta) \) to changes in \( \theta \):

If \( \theta \rightarrow \theta + \Delta \theta \), then it causes \( p(x; \theta) \rightarrow p(x; \theta + \Delta \theta) \)

How sensitive is \( p(x; \theta) \) to that change??

\[
\tilde{S}_\theta^p (x) \triangleq \left[ \frac{\Delta p(x; \theta)}{\Delta \theta} \right] = \% \text{ change in } p(x; \theta) = \left[ \frac{\Delta p(x; \theta)}{\Delta \theta} \right] \left[ \frac{\theta}{p(x; \theta)} \right]
\]

Now let \( \Delta \theta \rightarrow 0 \): \( S_\theta^p (x) = \lim_{\Delta \theta \rightarrow 0} \tilde{S}_\theta^p (x) = \left[ \frac{\partial p(x; \theta)}{\partial \theta} \right] \left[ \frac{\theta}{p(x; \theta)} \right] = \theta \frac{\partial \ln p(x; \theta)}{\partial \theta} \)

Recall from Calculus:
\[
\frac{\partial \ln f(x)}{\partial x} = \frac{1}{f(x)} \frac{\partial f(x)}{\partial x}
\]

**Interpretation**
Norm. CRLB = Inverse Mean Square Sensitivity
Definition of Fisher Information

The denominator in CRLB is called the Fisher Information $I(\theta)$

It is a measure of the “expected goodness” of the data for the purpose of making an estimate

$$I(\theta) = -E\left\{ \frac{\partial^2 \ln p(x; \theta)}{\partial \theta^2} \right\}$$

Has the needed properties for “info” (as does “Shannon Info”):
1. $I(\theta) \geq 0$ (easy to see using the alternate form of CRLB)
2. $I(\theta)$ is additive for independent observations

follows from:

$$\ln p(x; \theta) = \ln \left[ \prod_n p(x[n]; \theta) \right] = \sum_n \ln[p(x[n]; \theta)]$$

If each $I_n(\theta)$ is the same: $I(\theta) = N \times I(\theta)$
3.5 CRLB for Signals in AWGN

When we have the case that our data is “signal + AWGN” then we get a simple form for the CRLB:

Signal Model: \( x[n] = s[n; \theta] + w[n], \quad n = 0, 1, 2, \ldots, N-1 \)

Q: What is the CRLB?

First write the likelihood function:

\[
p(x; \theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - s[n; \theta])^2 \right\}
\]

Differentiate Log LF twice to get:

\[
\frac{\partial^2}{\partial \theta^2} \ln p(x; \theta) = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left( (x[n] - s[n; \theta]) \frac{\partial^2 s[n; \theta]}{\partial \theta^2} - \left[ \frac{\partial s[n; \theta]}{\partial \theta} \right]^2 \right)
\]
E\left\{ \frac{\partial^2}{\partial \theta^2} \ln p(x; \theta) \right\} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left\{ \frac{E\{x[n]\} - s[n; \theta]}{s[n; \theta]} \right\} \frac{\partial^2 s[n; \theta]}{\partial \theta^2} - \left[ \frac{\partial s[n; \theta]}{\partial \theta} \right]^2
\]

\[
= \frac{- \sum_{n=0}^{N-1} \left[ \frac{\partial s[n; \theta]}{\partial \theta} \right]^2}{\sigma^2}
\]

Then using this we get the **CRLB for Signal in AWGN**: 

\[
\text{var}(\hat{\theta}) \geq \frac{\sigma^2}{\sum_{n=0}^{N-1} \left[ \frac{\partial s[n; \theta]}{\partial \theta} \right]^2}
\]

Note: \( \left[ \frac{\partial s[n; \theta]}{\partial \theta} \right]^2 \) tells how sensitive signal is to parameter 

If signal is very sensitive to parameter change… then CRLB is small 
… can get very accurate estimate!
Ex. 3.5: CRLB of Frequency of Sinusoid

Signal Model: \( x[n] = A\cos(2\pi f_o n + \phi) + w[n] \) \( 0 < f_o < \frac{1}{2} \) \( n = 0, 1, 2, \ldots, N - 1 \)

\[ \text{var}(\hat{\theta}) \geq \frac{1}{SNR \times \sum_{n=0}^{N-1} [2\pi n \sin(2\pi f_o n + \phi)]^2} \]

Error in Book

Bound on Variance

Bound on Std. Dev.

Signal is less sensitive if \( f_o \) near 0 or \( \frac{1}{2} \)
3.6 Transformation of Parameters

Say there is a parameter $\theta$ with known CRLB$_\theta$

But imagine that we instead are interested in estimating some other parameter $\alpha$ that is a function of $\theta$:

$$\alpha = g(\theta)$$

Q: What is CRLB$_\alpha$?

$$\text{var}(\alpha) \geq \text{CRLB}_{\alpha} = \left( \frac{\partial g(\theta)}{\partial \theta} \right)^2 \text{CRLB}_\theta$$

Captures the sensitivity of $\alpha$ to $\theta$

Large $\partial g/\partial \theta$ $\rightarrow$ small error in $\theta$ gives larger error in $\alpha$
$\rightarrow$ increases CRLB (i.e., worsens accuracy)

Proved in Appendix 3B
Example: Speed of Vehicle From Elapsed Time

![Diagram of speed measurement setup]

But... really want to measure speed $V = d/T$

Find the CRLB$_V$:

$$CRLB_V = \left[ \frac{\partial}{\partial T} \left( \frac{D}{T} \right) \right] ^2 \times CRLB_T$$

$$= \left( -\frac{D}{T} \right) ^2 \times CRLB_T$$

$$= \frac{V^4}{D^2} \times CRLB_T$$

**Accuracy Bound**

$$\sigma_V \geq \frac{V^2}{D} \sqrt{CRLB_T} \quad (m/s)$$

- Less accurate at High Speeds (quadratic)
- More accurate over large distances
Effect of Transformation on Efficiency

Suppose you have an efficient estimator of $\theta: \hat{\theta}$
But... you are really interested in estimating $\alpha = g(\theta)$

Suppose you plan to use $\hat{\alpha} = g(\hat{\theta})$

Q: Is this an efficient estimator of $\alpha$???
A: **Theorem**: If $g(\theta)$ has form $g(\theta) = a\theta + b$, then $\hat{\alpha} = g(\hat{\theta})$ is efficient.

**Proof:**
First: $\text{var}(\hat{\alpha}) = \text{var}(a\hat{\theta} + b) = a^2 \text{var}(\hat{\theta}) = a^2 \text{CRLB}_\theta$

Now, what is CRB $\alpha$? Using transformation result:

$$CRLB_\alpha = \left[\frac{\partial(a\theta + b)}{\partial \theta}\right]^2 = a^2 \text{CRLB}_\theta \quad \rightarrow \quad \text{var}(\hat{\alpha}) = \text{CRLB}_\alpha$$

Efficient!
Asymptotic Efficiency Under Transformation

If the mapping $\alpha = g(\theta)$ is not affine… this result does NOT hold

But… if the number of data samples used is large, then the estimator is approximately efficient ("Asymptotically Efficient")