If \( \xi' = \begin{bmatrix} 1 & \frac{1}{2} & 0 \\ & \frac{1}{3} & \vdots \\ & 0 & \frac{1}{n} \end{bmatrix} \),

\[ D = \begin{bmatrix} 1 & \sqrt{\frac{1}{n}} & 0 \\ & \sqrt{\frac{1}{n}} & \vdots \\ & 0 & \sqrt{\frac{1}{n}} \end{bmatrix} = D^T \]

\( x' = Dx = \begin{bmatrix} x(0) \\ x(1)/\sqrt{n} \\ \vdots \\ x(n-1)/\sqrt{n} \end{bmatrix} \)

\[ d_n = \sqrt{n/(n+1)} \]

\[ \frac{\sum_{n=0}^{n-1} s(n)(n+1)}{\sum_{n=0}^{n-1} s(n)} \]

Correlator weights

\[ s'(n) = \frac{A}{\sqrt{n+1}} \]

\[ \hat{A} = \frac{\sum_{n=0}^{n-1} x'(n)s(n)}{\sum_{n=0}^{n-1} \sqrt{n+1}} \]

Summary

If \( x = H_0 + w \), where \( H \) is \( n \times p \) (known), \( w \sim N(0, \Sigma) \) and \( \hat{\theta} \) is to be estimated

\[ \hat{\theta} = (H^T\Sigma^{-1}H)^{-1}H^T\Sigma^{-1}x \]
IS THE MVU ESTIMATOR (AND ALSO EFFICIENT) AND HAS COVARIANCE

\[ C_\theta = (H^T \hat{E}^{-1} H)^{-1}. \]

THIS IS THE GENERAL LINEAR MODEL.

GENERAL MVU ESTIMATION

NOW ASSUME CALIBRATED NOT SATISFIED WITH EQUALITY. THERE IS NO EFFICIENT ESTIMATOR.

HOW DO WE FIND THE MVU ESTIMATOR (IF IT EXISTS)?

SUFFICIENT STATISTICS

THESE LEAD TO MVU ESTIMATORS.

RECALL DC LEVEL IN WGN EXAMPLE

\[ \hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \] IS MVU ESTIMATOR WITH VARIANCE \( \sigma^2 / N \)

CONSIDER \( \hat{A} = x[0] \) AS AN ESTIMATOR

\[ E(\hat{A}) = E(x[0]) = A \rightarrow \text{UNBIASED} \]

\[ \text{VAR}(\hat{A}) = \text{VAR}(x[0]) = E(w^2 \text{LO}) = \sigma^2 \]
VARIANCE OF A MUCH LARGER DECREASE
\{x(1), x(2), ..., x(n-1)\} INFORMATION NOT USED

WITH REGARDS TO ESTIMATION WE ASK WHICH DATA ARE IMPORTANT OR SUFFICIENT FOR ESTIMATION?

CONSIDER

\[ S_1 = \{ x(0), x(1), ..., x(n-1) \} \]
\[ S_2 = \{ x(0), x(1), x(2), ..., x(n-1) \} \]
\[ S_3 = \{ \sum_{n=0}^{n-1} x(n) \} \]

ALL SETS ARE SUFFICIENT SINCE \( \hat{\theta} \) MAY BE FOUND. \( S_3 \), HOWEVER, IS THE MINIMAL ONE.

ONCE THE SUFFICIENT STATISTICS ARE KNOWN WE CAN DISCARD DATA. ALL INFORMATION OF DATA IS SUMMARIZED.

TO QUANTIFY THESE IDEAS:

CONSIDER \( \sum_{n=0}^{n-1} x(n) \) AS THE SS

AFTER OBSERVING \( \sum_{n=0}^{n-1} x(n) = T \), THE DATA WILL TELL US NOTHING ABOUT \( \theta \) OR
\[ p \left( x \mid \sum_{n=0}^{N-1} x_n = T_0; A \right) \text{ will not depend on } A. \text{ Otherwise, data will provide additional information about } A. \]

To show that \[ \sum_{n=0}^{N-1} x_n \] is a sufficient statistic for the conditional PDF.

\[ p \left( x \mid T(x) = T_0; A \right) = \frac{p \left( x, T(x) = T_0; A \right)}{p \left( T(x) = T_0; A \right)} \]

where \[ T(x) = \sum_{n=0}^{N-1} x_n \]

But \[ p \left( x, T(x) = T_0; A \right) = 0 \] if \( T(x) \neq T_0 \)

\[ p(x; A) \] if \( T(x) = T_0 \)

Since \( T(x) \) depends on \( x \).

\[ p \left( x \mid T(x) = T_0; A \right) = \frac{p(x; A) \quad T(x) = T_0}{p(T(x) = T_0; A)} \]

But \[ T(x) = \sum_{n=0}^{N-1} x_n \]

\[ \sim \mathcal{N}(N \mu, N \sigma^2) \]
\[ P(x; \alpha) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left( \sum_{n=0}^{N-1} (x_n - \alpha)^2 \right)} \]

When \( T(x) = T_0 \)

\[ P(x; \alpha) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left( \sum_{n} x_n^2 - 2\alpha \sum_{n} x_n + N\alpha^2 \right)} \]

\[ = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left( \sum_{n} x_n^2 - 2\alpha T_0 + N\alpha^2 \right)} \]

\[ P(x\mid T(x) = T_0; \alpha) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left( \sum_{n} x_n^2 \right) - 2\alpha T_0 + N\alpha^2} \]

\[ \frac{1}{\sqrt{2\pi N\sigma^2}} e^{-\frac{1}{2\sigma^2} \left( T_0 - N\alpha \right)^2} \]

\[ = \frac{\sqrt{N}}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n} x_n^2} \frac{T_0^2}{2N\sigma^2} \]

For \( T(x) = T_0 \)

\[ = 0 \quad \text{for} \ T(x) \neq T_0 \]

is not a function of \( \alpha \).

\[ \sum_{n=0}^{N-1} x_n \] is so for \( \alpha \).
Not easy to verify if $T(x)$ is SS
(LET ALONE FIND $T(x)$)
Finding SS

Neyman-Fisher Factorization Theorem -
If we can factor PDF as

$$p(x; \theta) = g(T(x), \theta) h(x)$$

Then $T(x)$ is SS. Otherwise a SS does not exist.

Example: DC level in WGN

$$x(n) = A + w(n)$$

$$p(x; A) = \frac{1}{(2\pi \sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x(n) - A)^2}$$

To factor as above

$$\sum (x(n) - A)^2 = \sum x^2(n) - 2A \sum x(n) + NA^2$$

$$p(x; A) = \frac{1}{(2\pi \sigma^2)^{1/2}} e^{-\frac{1}{2\sigma^2} (NA^2 - 2A \sum x(n))}$$

$$g(\sum x(n), A)$$
\[ T(x) = \sum_{n=0}^{N-1} x[n] \text{ is SS for } A \]

So also, is \( T'(x) = 2 \sum_{n=0}^{N-1} x[n] \) or any 1-1 function of SS.

**Example:** Phase of sinusoid

\[ x[n] = A \cos(2\pi fn + \phi) + w[n], \quad n = 0, 1, \ldots, N-1 \]

\[ \rho(x; \phi) = \frac{1}{2\pi \sigma_x^2} e^{-\frac{1}{2\sigma_x^2} \sum_{n=0}^{N-1} (x[n] - A \cos(2\pi fn + \phi))^2} \]

As before we expand exponent

\[ \sum x^2[n] - 2A \sum x[n] \cos(2\pi fn + \phi) + \sum A^2 \cos^2 \]

\[ = \sum x^2[n] - 2A \left( \sum x[n] \cos 2\pi fn \right) \cos \phi \]

\[ + 2A \left( \sum x[n] \sin 2\pi fn \right) \sin \phi \]

\[ + \sum A^2 \cos^2(2\pi fn + \phi)\]
\[ p(x; \theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} x_n^2} \frac{1}{\sigma} \left( \frac{\sigma^2}{2\pi} \right)^{N/2} \left[ \sum_{n=0}^{N-1} A_n^2 \cos^2 (\pi x_n \cos \phi) - 2AT_1(x) \cos \phi + 2AT_2(x) \sin \phi \right] \]

\[ f(T_1(x), T_2(x), \phi) \]

\[ e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} x_n^2} \cdot h(x) \]

WHERE \[ T_1(x) = \sum_{n=0}^{N-1} x_n \cos \pi \sin \phi \]
\[ T_2(x) = \sum_{n=0}^{N-1} x_n \sin \pi \sin \phi \]

GENERALIZED NEYMAN-FISHER SAYS:
\[ T_1(x), T_2(x) \text{ ARE JOINTLY SUFFICIENT.} \]

NO SCALAR SS EXISTS!

GENERALIZED NEYMAN-FISHER:

IF \[ p(x; \theta) = g(T_1(x), ..., T_r(x), \theta) h(x) \]

\[ \Rightarrow \{T_1, T_2, ..., T_r\} \text{ ARE SS FOR } \theta. \]
BY SS WE NOW MEAN THAT

\[ p(x_1, x_2, \ldots, x_n; \theta) \text{ does not depend on } \theta. \]

**Example:** Original data are SS since if \( r = n \) and

\[ T_{an}(x) = x_1 \alpha \quad \alpha = 0, 1, \ldots, N-1 \]
\[ g = p \]
\[ h = 1 \]

**Factorization holds or**

\[ p(x; \theta) = p(x; \alpha_0, \alpha_1, \ldots, \alpha_{N-1}; \theta) \cdot \frac{1}{h(x)} \]
\[ g(T_1, \ldots, T_n, \theta) \]

**Finding MVU Estimator**

Relies on SS and also Rao-Blackwell-Lehmann-Scheffe theorem.

**Best to illustrate with example.**
EXAMPLE: DC LEVEL IN WGN

TWO METHODS - BOTH BASED ON SS

\[ T = \frac{\chi^2 - 1}{\lambda} x E_{\infty} \]

1) FIND UNBIASED ESTIMATOR OF
\[ \hat{A} = x(0) \text{ AS EXAMPLE,} \]
\[ \Rightarrow \hat{A} = E(\hat{A} | T) \]

2) FIND A FUNCTION \( g \) SO THAT
\[ \hat{A} = g(T) \text{ IS UNBIASED ESTIMATOR} \]
\[ \Rightarrow \hat{A} \text{ IS MVU ESTIMATOR} \]

APPROACH 1: \[ \hat{A} = E(x(0) | \sum_{n} x(n)) \]

ASIDE: NEED \( E(x(1)y) \) FOR \( x, y \) JOINTLY GAUSSIAN

ASSUME \[ \begin{bmatrix} y \\ x \end{bmatrix} \sim N \left( \begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix} \right) \]

FIND \( p(x|y) \)

\[ p(x|y) = \frac{p(x,y)}{p(y)} = \frac{\frac{1}{\sqrt{2\pi \sigma_x^2}} e^{-\frac{1}{2}(x-m_x)^2/(2\sigma_x^2)}}{\sqrt{2\pi \sigma_y^2}} e^{-\frac{1}{2}(y-m_y)^2/(2\sigma_y^2)} \]
Consider exponent

\[
\begin{align*}
(x - mx)^2 \left( e^{-} \right)_{11} + 2(x - mx)(y - my) \left( e^{-} \right)_{12} \\
+ (y - my)^2 \left( e^{-} \right)_{22} - \frac{1}{\sigma_y^2} (y - my)^2
\end{align*}
\]

\[
C^{-1} = \begin{bmatrix}
\sigma_y^2 - \sigma_{xy}^2 \\
-\sigma_{xy}^2 \sigma_x^2
\end{bmatrix}
\]

\[
\frac{\sigma_x^2 \sigma_y^2 - \sigma_{xy}^4}{\sigma_x^2 \sigma_y^2}
\]

\[
\frac{\sigma_y^2 (x - mx)^2}{\alpha} - \frac{2(x - mx)(y - my) \sigma_{xy}^2}{\alpha} + \frac{(y - my)^2 \sigma_x^2}{\alpha}
\]

\[
- \frac{1}{\sigma_y^2} (y - my)^2
\]

\[\Rightarrow \text{QUADRATIC IN } x \Rightarrow p(x|y) \text{ IS GAUSSIAN} \]

\[p(x|y) = \frac{1}{\sqrt{2\pi \sigma_{xy}^2}} e^{- \frac{\alpha}{2} (x - mx_{xy})^2} \]

\[\Rightarrow \sigma_{xy}^2 = \frac{\alpha}{\sigma_y^2} \]

\[-2 \frac{mx \sigma_y^2}{\alpha} - 2 \frac{\sigma_{xy}^2}{\alpha} (y - my) = \frac{-2 mx \sigma_y^2}{\alpha} \]

\[= -2 \frac{mx_{xy} \sigma_y^2}{\alpha} \]

\[\Rightarrow mx_{xy} = mx + \frac{\sigma_{xy}^2}{\sigma_y^2} (y - my) \]
Now let \( x = x_0 \)
\[ y = \sum_{n=1}^{N} x(n) \]

\[
\begin{pmatrix}
 x \\
 y
\end{pmatrix} = \begin{pmatrix}
 x(0) \\
 \sum_{n=1}^{N} x(n)
\end{pmatrix} = \frac{1}{\sqrt{N}} \begin{pmatrix}
 1 & 0 & \ldots & 0
\end{pmatrix} \begin{pmatrix}
 x(0) \\
 \sqrt{N} x(1) \\
 \vdots \\
 \sqrt{N} x(N-1)
\end{pmatrix}
\]

Since \( x(0), \ldots, x(N-1) \) is jointly Gaussian,
\[ \begin{pmatrix}
 x \\
 y
\end{pmatrix} \text{ is jointly Gaussian (linear transformation)} \]

\[ \begin{pmatrix}
 x \\
 y
\end{pmatrix} \sim N(\mu, \Sigma) \]

\[ \mu = \mathbb{E}(x) = \frac{1}{N} A \begin{pmatrix}
 A
\end{pmatrix} \]

\[ \Sigma = \mathbb{E}(x^2) = \frac{1}{N} \begin{pmatrix}
 1 & 1 \\
 1 & 1
\end{pmatrix} \begin{pmatrix}
 1 & 1
\end{pmatrix} \]

\[ \Rightarrow \hat{A} = \mathbb{E}(x|y) = \mu + \frac{\text{cov}(x, y)}{\text{var}(y)} (y - \mu y) \]

\[ = A + \frac{\sigma_x}{\sigma_y^2} \left( \frac{1}{N} \sum_{n=0}^{N-1} x(n) - \hat{A} \right) \]

\[ = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \]

**Approach 2:** \[ \hat{A} = g \left( \sum_{n=0}^{N-1} x(n) \right) \]

Choose \( g \) so that \( \hat{A} \) is unbiased.
By inspection \( g(x) = \frac{1}{n} x \)

\[\hat{\theta} = \frac{1}{n} \sum_{i=0}^{n} x_i \]

**Rao-Blackwell-Lehmann-Scheffe Theorem:**

If \( \theta \) is unbiased estimator of \( \theta \), and

\( T(x) \) is a su for \( \theta \), then \( \hat{\theta} = E(\theta | T(x)) \) is

1) A valid estimator (doesn't depend on \( \theta \))
2) Unbiased
3) Of less variance than \( \theta \),

And if \( T(x) \) is complete, \( \hat{\theta} \) is the

MVE estimator.

**Example:** Previous example

\[ \theta = x(0) \]
\[ T(x) = \sum_{i} x_i \]

\( \bar{x} \) is unbiased; \( T(x) \) is su \( \Rightarrow \)

\[ \hat{\theta} = E(\theta | T(x)) = \frac{1}{n} \sum_{i} x_i \]

1) Does not depend on \( a \)
2) Is unbiased
3) Has less variance \((\sigma^2/n)\) than \( \theta \) \((\sigma^2)\)
Furthermore, we know $\frac{1}{N} \sum_{i=1}^{N} X_i$ is MVU estimator. This follows since $T(x)$ is complete.

Complete means there is only one function of the SS that results in unbiased estimator.

Assume this to be true. Why is $\hat{\theta}$ now MVU estimator?

1) If $T(x)$ complete $\Rightarrow$ $\hat{\theta}$ is unique (only one function of $T(x)$ which is unbiased must be $E(\hat{\theta} | T(x))$.

2) All $\theta$ produce same $\hat{\theta}$.

3) But variance must be decreased.
NOTE THAT SINCE THERE IS ONLY ONE FUNCTION $g(t)$ THAT PRODUCES AN UNBIASED ESTIMATOR, THIS MUST PRODUCE $\hat{\sigma}_{MVU}$.

VERIFYING COMPLETENESS IS HARD IN GENERAL. FOR THIS EXAMPLE WE CAN DO SO.

Let $T = \sum_{n=0}^{N-1} X(n)$

WISH TO SHOW THAT IF

$E[g(t)] = A$ FOR ALL $A$

THEN THERE IS BUT ONE SOLUTION FOR $g$.

PROOF: ASSUME THAT THERE EXISTS ANOTHER FUNCTION $h$ SUCH THAT

$E[h(t)] = A$ FOR ALL $A$

$\Rightarrow E[g(t) - h(t)] = 0$ FOR ALL $A$

LET $\nu(t) = g(t) - h(t)$

$E[\nu(t)] = \int_{-\infty}^{\infty} \nu(t) \phi(t) dt$

$= \int_{-\infty}^{\infty} \frac{\nu(t)}{(2\pi \sigma^2)^{1/2}} e^{-\frac{t^2}{2\sigma^2}} dt$
Let $\tau = T/N$, $N'(\tau) = N(N\tau)$

$$E[N(T)] = \int_{-\infty}^{\infty} \frac{N(T)}{(2\pi N \sigma^2)^{1/2}} e^{-\frac{N}{2\sigma^2}(A - T/N)^2} dT$$

$$= \int_{-\infty}^{\infty} \frac{N N'(\tau)}{(2\pi N \sigma^2)^{1/2}} e^{-\frac{N}{2\sigma^2}(A - \tau)^2} d\tau$$

$$\Rightarrow \int_{-\infty}^{\infty} N'\tau \left[ \frac{N}{(2\pi N \sigma^2)^{1/2}} e^{-\frac{N}{2\sigma^2}(A - \tau)^2} \right] d\tau = 0$$

But this is a convolution $\Rightarrow$

$$N'(\tau) \ast W(\tau) = 0 \text{ for all } \tau$$

$$\Rightarrow \quad V'(\tau) W(\tau) = 0 \quad \text{all } \tau$$

But $W(\tau) = \mathcal{F} \{ \text{Gaussian pulse} \} > 0$ all $\tau$

$$\Rightarrow \quad V'(\tau) = 0 \quad \text{all } \tau$$

$$\Rightarrow \quad N'(\tau) = 0 \quad \text{all } \tau$$

$$\Rightarrow \quad g = b$$

$\Rightarrow$ only one function of $T$ which is
UNBIASED.

See also Example 5.7 for noncomplete ss.

Note that to prove ss is complete, we must show

\[ E(n(T)) = 0 \quad \text{for all } \theta \]

\[ \Rightarrow n(T) = 0 \quad \text{for all } T \]

or

\[ \int_{-\infty}^{\infty} n(T) \, \varphi(T; \theta) \, dT = 0 \quad \text{for all } \theta \]

\[ \Rightarrow n(T) = 0 \]

Thus, completeness is a property of \( \varphi(T; \theta) \) or of \( \varphi(x; \theta) \) and \( T \).

**Summary**

To determine MVU estimator (when CRLB not satisfied) if it exists:

1. Find ss for \( \theta \) or \( T(x) \) using Neyman-Fisher factorization theorem
2) **Determine if SS is complete and if so proceed (if not stop and ?)**
3) **Find function \( g \) of the SS that yields unbiased estimator or**

\[
\hat{\theta} = g(T(x))
\]

**Where** \( E(\hat{\theta}) = \theta \)

\[ \Rightarrow \hat{\theta} \text{ is MVU estimator} \]

**How do we find \( g \)?**

**Alternatively, replace 3) by**

3') **Evaluate** \( \hat{\theta} = E(\hat{\theta} | T(x)) \) **for any unbiased estimator** \( \hat{\theta} \).

**Example:** Mean of uniform noise

\[
x[n] = w[n] \quad n = 0, 1, \ldots, N-1
\]

\( w[n] \sim U[0, A] \quad \text{IID (independent \& \ identically distributed)} \)
ESTIMATE MEAN \( \theta = \beta/2 \).

MIGHT GUESS \( \hat{\theta} = \frac{1}{N} \sum_{n=0}^{N-1} x(n) \)

THIS IS UNBIASED AND HAS VARIANCE

\[
\text{VAR}(\hat{\theta}) = \text{VAR}\left(\frac{1}{N} \sum_{n=0}^{N-1} x(n)\right)
\]

\[
= \frac{1}{N} \text{VAR}(x(n)) = \frac{\beta^2}{12N}
\]

IS THIS MVU ESTIMATOR?

WE NOW FIND MVU ESTIMATOR USING RBLS THEOREM.

\[
p_x(x_0; \theta) = \frac{1}{\beta} \left( \mu(x_0) - \mu(x_0 - \beta) \right)
\]

WHERE \( \mu(x) = \begin{cases} 
1 & x > 0 \\
0 & x < 0 
\end{cases} \) \text{ UNIT STEP}

\[
p(x; \beta) = \frac{1}{\beta^n} \prod_{n=0}^{N-1} \left[ \mu(x(n)) - \mu(x(n) - \beta) \right]
\]

\[= 0 \text{ UNLESS } 0 < x(n) < \beta \text{ FOR ALL } n \]
\[
= \begin{cases} \frac{1}{\beta^n} & \text{if } 0 < x(n) < \beta, \; \min x(n) > 0 \\ 0 & \text{otherwise} \end{cases}
\]

\[
= \frac{1}{\beta^n} \underbrace{\mathbb{U}(\beta - \max x(n)) \mathbb{U}(\min x(n))}_{g(t(x, \theta))}
\]

\[T(x) = \max x(n) \text{ is } \text{SF} \text{ for } \theta = \beta/2\]

Next we must find \(g\) so that \(g(\max x(n))\) is unbiased.

\[\Rightarrow \text{need PDF of } \max x(n)\]

This involves order statistics (standard problem in probability).

\[P_r \{T \leq t\} = P_r \{x(0) \leq t, \ldots, x(N-1) \leq t\} = \prod_{n=0}^{N-1} P_r \{x(n) \leq t\} \quad \text{(independent)}\]