

1

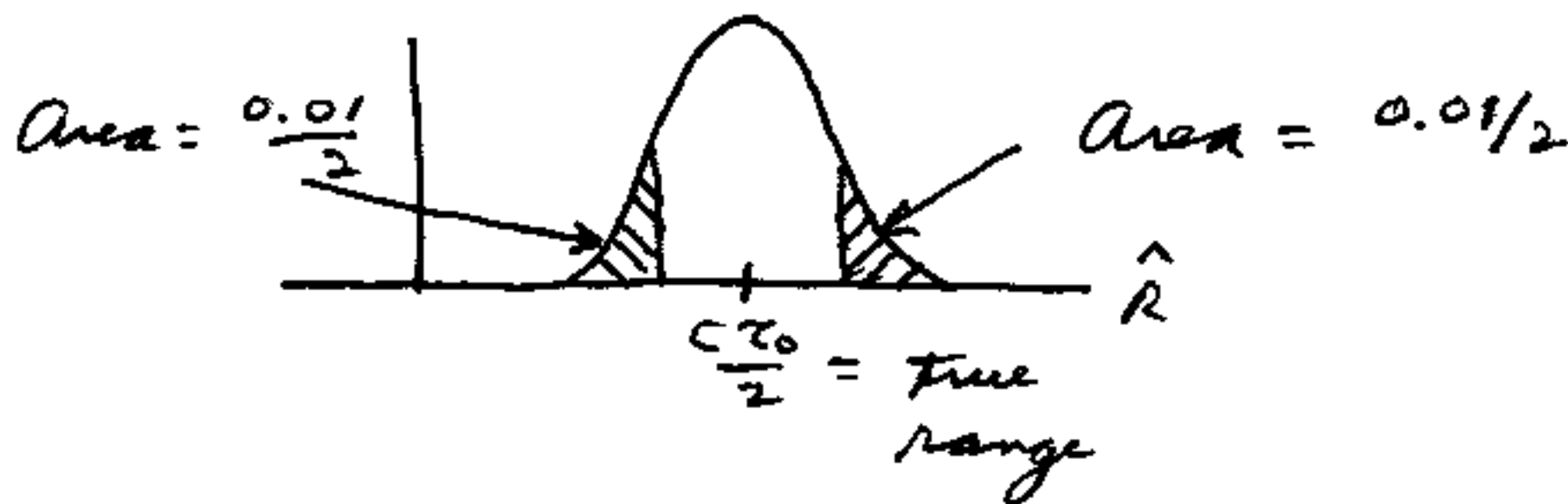
PROBLEM SOLUTIONS

FUNDAMENTALS OF STATISTICAL
SIGNAL PROCESSING: ESTIMATION
THEORY

BY STEVEN KAY

Chapter 1

- 1) Since $R = c\tau_0/2$, we use $\hat{R} = c\hat{\tau}_0/2$.
 The PDF is from $\hat{\tau}_0 \sim N(\tau_0, \sigma_{\hat{\tau}_0}^2)$,
 $\hat{R} \sim N(c\tau_0/2, \frac{c^2}{4} \sigma_{\hat{\tau}_0}^2)$



To be within 100 m we must have

$$Pr \{ |\hat{R} - c\tau_0/2| < 100 \} = 0.99$$

$$\Rightarrow Pr \left\{ \underbrace{\left| \frac{\hat{R} - c\tau_0/2}{\frac{c}{2} \sigma_{\hat{\tau}_0}} \right|}_{N(0,1)} < \frac{100}{\frac{c}{2} \sigma_{\hat{\tau}_0}} \right\} = 0.99$$

$$\Rightarrow \frac{100}{\frac{c}{2} \sigma_{\hat{\tau}_0}} = 2.58 \quad \text{or} \quad \sigma_{\hat{\tau}_0} = 2.6 \times 10^{-7} \text{ sec} \\ = 0.26 \mu\text{sec}$$

- 2) No, in fact θ could have been any value.

If θ were indeed 100, then

$$p(x; \theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-100)^2}$$

and the probability of x being in the interval $[-97, 103]$ or $\mu \pm 3\sigma$ is 0.999.

Hence, his assertion is likely to be correct. However, we cannot be

certain, since if $\theta = 99$, then the probability of x being in the observed interval (for a single experiment) is

$$\int_{-97}^{103} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-99)^2} dx =$$

$$\int_{-2}^4 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du = 0.977.$$

Thus, $\theta = 99$ is also highly likely.

3) $x = \theta + w$

$$p(x; \theta) = p_w(x - \theta)$$

For θ a random variable independent of w ,

$$p(x|\theta) = \frac{p_{x\theta}(x, \theta)}{p(\theta)} = \frac{p_{w\theta}(x - \theta, \theta)}{p(\theta)}$$

$$= \frac{p_w(x - \theta) p(\theta)}{p(\theta)} = p_w(x - \theta)$$

which is the same as before. If w and θ are not independent, then

$$p(x|\theta) = \frac{p_{w|\theta}(x - \theta) p(\theta)}{p(\theta)} = p_{w|\theta}(x - \theta)$$

which will be different than $p_w(x - \theta)$.

In general, $p(x; \theta) \neq p(x|\theta)$.

4) As shown in text $E(\hat{A}) = A$. Also,

$$E(\check{A}) = \frac{1}{N+2} (2A + (N-2)A + 2A) = A$$

Also, we know that $\text{var}(\hat{A}) = \sigma^2/N = 1/N$ and

$$\begin{aligned} \text{var}(\check{A}) &= \frac{1}{(N+2)^2} \left[4\sigma^2 + \sum_{n=1}^{N-2} \sigma^2 + 4\sigma^2 \right] \\ &= \frac{N+6}{(N+2)^2} \sigma^2 = \frac{N+6}{(N+2)^2} \end{aligned}$$

$$\begin{aligned} \text{var}(\check{A}) - \text{var}(\hat{A}) &= \frac{N+6}{(N+2)^2} - \frac{1}{N} \\ &= \frac{N(N+6) - (N+2)^2}{N(N+2)^2} \\ &= \frac{2N-4}{N(N+2)^2} > 0 \text{ for } N > 2 \end{aligned}$$

Hence, both estimators yield the correct value on the average but \hat{A} has less variance. Conclusion is the same for any value of A .

5) \hat{A} is not an estimator since to implement it requires knowledge of A (to determine the SNR).

Chapter 2

$$\begin{aligned}
 1) \quad E(\hat{\sigma}^2) &= E\left(\frac{1}{N} \sum_{n=0}^{N-1} x^2[n]\right) = \frac{1}{N} \sum_{n=0}^{N-1} E(x^2[n]) \\
 &= \sigma^2 \quad \text{for all } \sigma^2 > 0 \text{ (allowable values)} \\
 &\Rightarrow \text{unbiased}
 \end{aligned}$$

$$\begin{aligned}
 \text{var}(\hat{\sigma}^2) &= \frac{1}{N^2} \text{var}\left(\sum_n x^2[n]\right) \\
 &= \frac{1}{N^2} N \text{var}(x^2[n]) \quad (x[n]'s \text{ are IID} \\
 &\quad \Rightarrow x^2[n] \text{ are IID}) \\
 &= \frac{1}{N} \text{var}(x^2[n])
 \end{aligned}$$

$$\begin{aligned}
 \text{var}(x^2[n]) &= E(x^4[n]) - E(x^2[n])^2 \\
 &= 3\sigma^4 - \sigma^4 = 2\sigma^4
 \end{aligned}$$

$$\Rightarrow \text{var}(\hat{\sigma}^2) = 2\sigma^4/N \rightarrow 0 \text{ as } N \rightarrow \infty$$

Hence, the PDF of $\hat{\sigma}^2$ collapses about the true value as $N \rightarrow \infty$.

$$2) \quad \text{Let } \hat{\theta} = 2 \frac{1}{N} \sum_{n=0}^{N-1} x[n] \quad \text{since } E(x[n]) = \theta/2$$

$$E(\hat{\theta}) = \frac{2}{N} \sum_{n=0}^{N-1} \theta/2 = \theta$$

$$3) \quad \hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \quad \hat{A} \text{ is Gaussian since it is a } \underline{\text{linear}} \text{ function of independent Gaussian random variables. The mean was found to be } A \text{ and the variance is}$$

$$\begin{aligned}
 \text{var}(\hat{A}) &= \text{var}\left(\frac{1}{N} \sum_{n=0}^{N-1} x[n]\right) \\
 &= \frac{1}{N^2} \text{var}\left(\sum_{n=0}^{N-1} x[n]\right) \\
 &= \frac{1}{N^2} N \text{var}(x[n])
 \end{aligned}$$

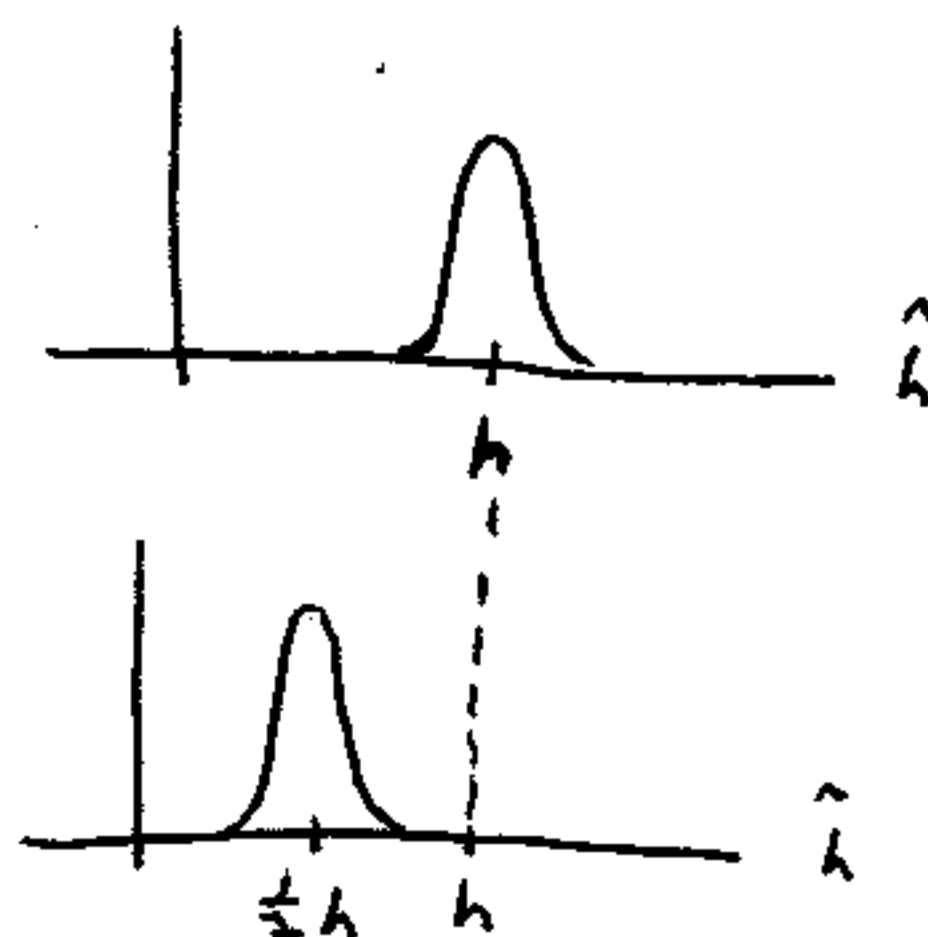
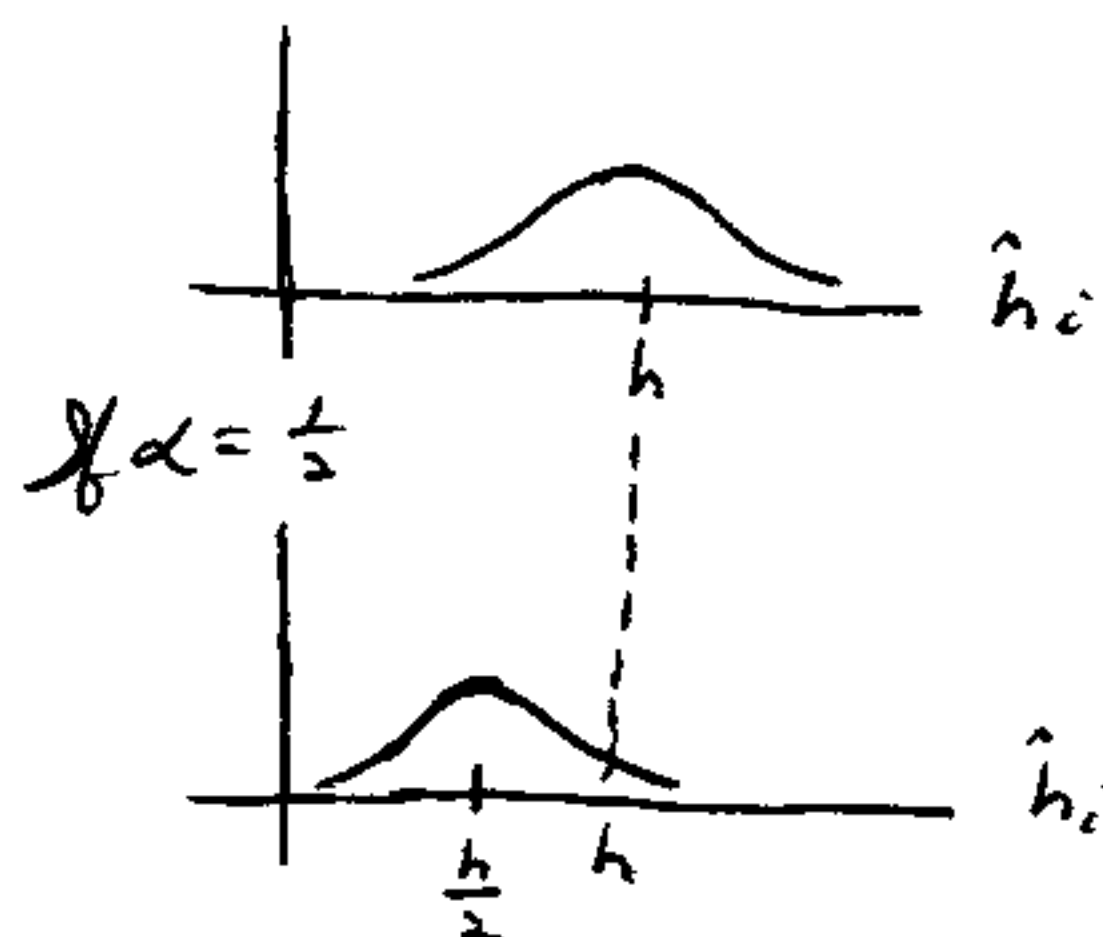
Since the $x[n]$'s are IID and thus uncorrelated

$$\Rightarrow \text{var}(\hat{A}) = \sigma^2/N.$$

$$4) \quad \hat{h} = \frac{1}{10} \sum_{i=1}^{10} h_i \quad E(\hat{h}) = \frac{1}{10} \sum_{i=1}^{10} E(h_i) = \alpha h$$

$$\text{var}(\hat{h}) = \text{var}(\hat{h}_i) / 10 = 1/10$$

If $\alpha = 1$, we have



In second case ($\alpha = \frac{1}{2}$), averaging causes the PDF to be more heavily concentrated about the wrong value of h . The probability

of \hat{h} being close to h actually decreases due to averaging. For $\alpha=1$, averaging, of course, is beneficial.

$$5) \quad X[N] \sim N(0, \sigma^2) \quad \text{or} \quad \frac{X[N]}{\sigma} \sim N(0, 1)$$

$$\Rightarrow \left(\frac{X[N]}{\sigma}\right)^2 \sim \chi^2_1 \quad \text{and}$$

$$y = \left(\frac{X[0]}{\sigma}\right)^2 + \left(\frac{X[1]}{\sigma}\right)^2 \sim \chi^2_2$$

$$\text{or } p(y) = \begin{cases} \frac{1}{2} e^{-y/2} & y > 0 \\ 0 & y < 0 \end{cases}$$

Transforming, we have $\hat{\sigma}^2 = \frac{\sigma^2}{2} y$
so that

$$p(\hat{\sigma}^2) = \frac{p_y(y(\hat{\sigma}^2))}{|d\hat{\sigma}^2/dy|}$$

$$= \frac{\frac{1}{2} e^{-\frac{1}{2}(2\hat{\sigma}^2/\sigma^2)}}{\sigma^2/2} \quad \begin{matrix} \hat{\sigma}^2 > 0 \\ 0 & \hat{\sigma}^2 < 0 \end{matrix}$$

$$= \frac{1}{\sigma^2} e^{-\hat{\sigma}^2/\sigma^2} \quad \begin{matrix} \hat{\sigma}^2 > 0 \\ 0 & \hat{\sigma}^2 < 0 \end{matrix}$$



Clearly, not symmetric.

$$\begin{aligned}
 \text{But } E(\hat{\sigma}^2) &= \int_0^\infty \hat{\sigma}^2 \frac{1}{\sigma^2} e^{-\hat{\sigma}^2/\sigma^2} d\hat{\sigma}^2 \\
 &= \sigma^2 \int_0^\infty u e^{-u} du \\
 &= \sigma^2 \left[-u e^{-u} - e^{-u} \right] \Big|_0^\infty \\
 &= \sigma^2
 \end{aligned}$$

$\hat{\sigma}^2$ is unbiased but PDF is not symmetric about σ^2 .

$$b) \quad E(\hat{A}) = \sum_{n=0}^{N-1} a_n A = A \Rightarrow \sum_{n=0}^{N-1} a_n = 1$$

$$\text{var}(\hat{A}) = \sum_{n=0}^{N-1} a_n^2 \text{var}(x[n]) = \sum_{n=0}^{N-1} a_n^2 \sigma^2$$

$$\text{Let } F = \sigma^2 \sum_{n=0}^{N-1} a_n^2 + \lambda \left(\sum_{n=0}^{N-1} a_n - 1 \right)$$

$$\frac{\partial F}{\partial a_i} = 2\sigma^2 a_i + \lambda = 0 \quad i = 0, 1, \dots, N-1$$

$$\Rightarrow a_i = -\lambda / 2\sigma^2 \quad \text{for all } i$$

Thus, the a_i 's must be equal. But $\sum_{n=0}^{N-1} a_n = 1 \Rightarrow N a_i = 1$ or $a_i = 1/N$ or

\hat{A} is just the sample mean estimator or as shown in Example 3.3, the MVU estimator.

$$7) \quad \frac{\hat{\theta} - \theta}{\sqrt{\text{var}(\hat{\theta})}} \sim N(0, 1) \quad \frac{\check{\theta} - \theta}{\sqrt{\text{var}(\check{\theta})}} \sim N(0, 1)$$

$$P_r \{ |\hat{\theta} - \theta| > \epsilon \} = P_r \left\{ \left| \frac{\hat{\theta} - \theta}{\sqrt{\text{var}(\hat{\theta})}} \right| > \frac{\epsilon}{\sqrt{\text{var}(\hat{\theta})}} \right\}$$

$$\text{Let } \Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} dt$$

= cumulative distribution
function for $N(0, 1)$

$$\Rightarrow P_r \{ |\hat{\theta} - \theta| > \epsilon \} = 2\Phi\left(\frac{-\epsilon}{\sqrt{\text{var}(\hat{\theta})}}\right)$$

$$\text{If } \text{var}(\hat{\theta}) < \text{var}(\check{\theta})$$

$$\Rightarrow \Phi\left(\frac{-\epsilon}{\sqrt{\text{var}(\hat{\theta})}}\right) < \Phi\left(\frac{-\epsilon}{\sqrt{\text{var}(\check{\theta})}}\right)$$

$$\text{or } P_r \{ |\hat{\theta} - \theta| > \epsilon \} < P_r \{ |\check{\theta} - \theta| > \epsilon \}$$

$$8) \text{ From Prob. 2.3 } \hat{A} \sim N(A, \sigma^2/N)$$

$$P_r \{ |\hat{A} - A| > \epsilon \} = P_r \left\{ \left| \frac{\hat{A} - A}{\sqrt{\sigma^2/N}} \right| > \frac{\epsilon}{\sqrt{\sigma^2/N}} \right\}$$

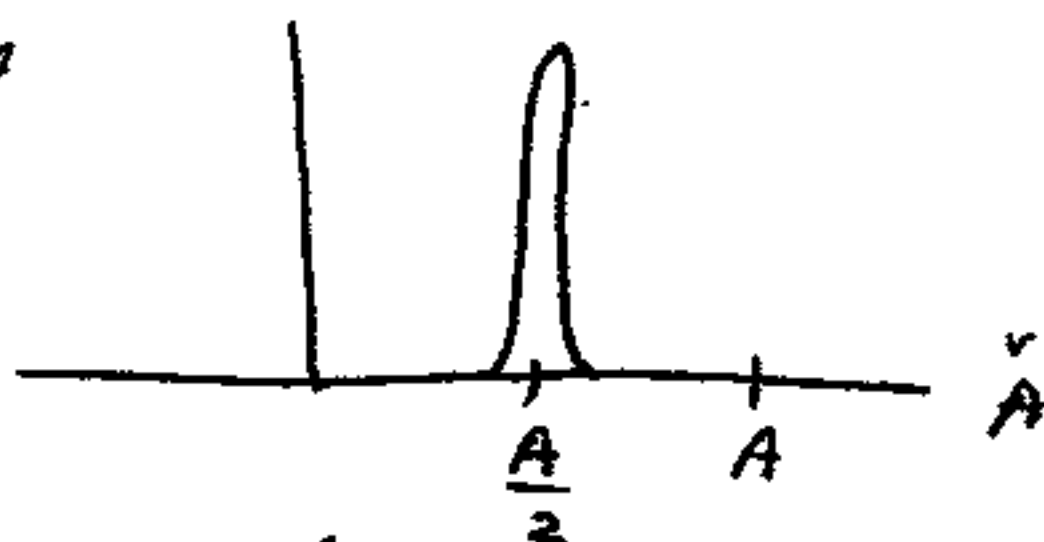
$$= 2\Phi\left(\frac{-\epsilon}{\sqrt{\sigma^2/N}}\right) \rightarrow 0$$

$$\text{Since } \frac{-\epsilon}{\sqrt{\sigma^2/N}} \rightarrow -\infty \text{ as } N \rightarrow \infty$$

$$\text{If } \hat{A} = \frac{1}{2N} \sum_{n=0}^{N-1} x[n] \text{ is used,}$$

$$\hat{A} \sim N(A/2, \sigma^2/4N)$$

As $N \rightarrow \infty$, the variance $\rightarrow 0$ and the PDF approaches



Thus, $\lim_{N \rightarrow \infty} P\{|\hat{A} - A| > \epsilon\} = 0$ for any $\epsilon > 0$

\hat{A} is consistent while \check{A} is inconsistent.

9) $\hat{\theta} = (\hat{A})^2$ where $\hat{A} \sim N(A, \sigma^2/N)$
 $E(\hat{\theta}) = E\{\hat{A}^2\} = \text{var}(\hat{A}) + E(\hat{A})^2$
 $= \sigma^2/N + A^2 = \theta + \sigma^2/N \neq \theta$

$\hat{\theta}$ is biased but as $N \rightarrow \infty$, it is unbiased. $\hat{\theta}$ is said to be asymptotically unbiased (for large data records).

10) Clearly, $E(\hat{A}) = A$. To find $E(\hat{\theta}^2)$

$$E(\hat{\theta}^2) = \frac{1}{N-1} \sum_{n=0}^{N-1} E[(x[n] - \hat{A})^2]$$

$$\text{But } E[(x[n] - \hat{A})^2] = E\left[\left(x[n] - \frac{1}{N} \sum_{m=0}^{N-1} x[m]\right)^2\right]$$

$$= E\left[\left(x[n]\left(1 - \frac{1}{N}\right) - \frac{1}{N} \sum_{\substack{m=0 \\ m \neq n}}^{N-1} x[m]\right)^2\right]$$

$$= \left(\frac{N-1}{N}\right)^2 E[x^2[n]] - 2 \frac{(N-1)}{N^2} E\left[x[n] \sum_{\substack{m=0 \\ m \neq n}}^{N-1} x[m]\right]$$

$$+ \frac{1}{N^2} E \left[\left(\sum_{\substack{m=0 \\ m \neq n}}^{N-1} x[m] \right)^2 \right]$$

$$= \left(\frac{N-1}{N} \right)^2 (\sigma^2 + A^2) - \frac{2(N-1)}{N^2} E[x[n]] E \left(\sum_{\substack{m=0 \\ m \neq n}}^{N-1} x[m] \right) \\ + \frac{1}{N^2} \left[\text{var} \left(\sum_{\substack{m=0 \\ m \neq n}}^{N-1} x[m] \right) + E \left(\sum_{\substack{m=0 \\ m \neq n}}^{N-1} x[m] \right)^2 \right]$$

$$= \left(\frac{N-1}{N} \right)^2 (\sigma^2 + A^2) - 2 \frac{(N-1)}{N^2} A (N-1) A \\ + \frac{1}{N^2} (N-1) \sigma^2 + \frac{1}{N^2} [(N-1) A]^2$$

$$= \sigma^2 \left[\frac{N^2 - 2N + 1 + N - 1}{N^2} \right] = \sigma^2 \frac{N-1}{N}$$

$$\Rightarrow E(\hat{\sigma}^2) = \frac{1}{N-1} \sum_{n=0}^{N-1} \sigma^2 \frac{N-1}{N} = \sigma^2$$

$\Rightarrow \hat{\sigma}^2$ is unbiased.

11) $\hat{\theta} = g(x[0])$

$$E(\hat{\theta}) = \theta \Rightarrow \int g(x[0]) p(x[0]) dx[0] = \theta$$

$$\text{or } \int_0^{1/\theta} g(x[0]) \theta dx[0] = \theta$$

$$\int_0^{1/\theta} g(u) du = 1 \quad \text{for all } \theta > 0$$

Now suppose a g could be found. Then

for any $\theta_2 < \theta_1$, we would have

$$\int_0^{\theta_2} g(u) du = 1$$

$$\int_0^{\theta_1} g(u) du = 1$$

and subtracting the two gives

$$\int_{\theta_2}^{\theta_1} g(u) du = 0 \quad \text{for any } \theta_2 < \theta_1$$

Clearly, we must have $g(u) = 0$ for all u , which produces a biased estimator.

Chapter 3

$$1) \quad p(x(n); \theta) = \frac{1}{\theta} (u(x(n)) - u(x(n) - \theta))$$

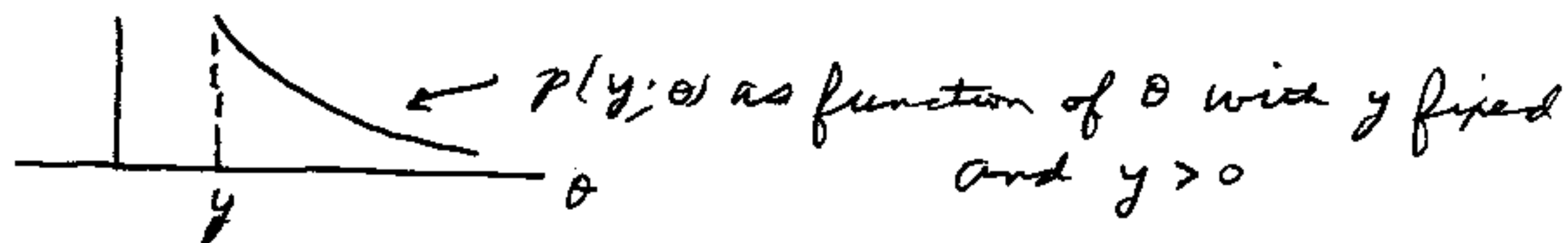
$$\text{where } u(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Since $p(x; \theta) = \prod_{n=0}^{N-1} p(x(n); \theta)$, it is enough to show that

$$E \left[\frac{\partial \text{LN} p(x(n); \theta)}{\partial \theta} \right] \neq 0$$

(The expectation will be independent of n).
Let $y = x(n)$ so that

$$p(y; \theta) = \frac{1}{\theta} (u(y) - u(y - \theta))$$



For $\theta > y$

$$\begin{aligned} E \left[\frac{\partial \text{LN} p(y; \theta)}{\partial \theta} \right] &= E \left[\frac{\partial \text{LN} 1/\theta}{\partial \theta} \right] \\ &= -1/\theta \neq 0 \end{aligned}$$

$$2) \quad X(0) = A + W(0)$$

$$p_x(X(0); A) = p_w(X(0) - A) = p(X(0) - A)$$

$$I(A) = E \left[\left(\frac{\partial \text{LN} p(X(0) - A)}{\partial A} \right)^2 \right]$$

$$\begin{aligned}
&= E \left[\left(\frac{\partial \ln p(x|0)-A}{\partial (x|0)-A} (-1) \right)^2 \right] \\
&= E \left[\left(\frac{1}{p(x|0)-A} \frac{\partial p(x|0)-A}{\partial (x|0)-A} \right)^2 \right] \\
&= \int_{-\infty}^{\infty} \left(\frac{\partial p(x|0)-A}{\partial (x|0)-A} \right)^2 \frac{1}{p^2(x|0)-A} p(x|0); A) dx|0 \\
&= \int_{-\infty}^{\infty} \left(\frac{\partial p(x|0)-A}{\partial (x|0)-A} \right)^2 \frac{1}{p^2(x|0)-A} p(x|0)-A) dx|0
\end{aligned}$$

Letting $u \equiv x|0)-A$

$$I(A) = \int_{-\infty}^{\infty} \frac{\left(\frac{dp(u)}{du} \right)^2}{p(u)} du$$

For $p(u) = \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|u|/\sigma}$ which is even in u

$$dp/du = -\frac{\sqrt{2}}{\sigma} \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}u/\sigma} \quad u > 0$$

$$\begin{aligned}
I(A) &= 2 \int_0^{\infty} \frac{\frac{1}{\sigma^4} e^{-2\sqrt{2}u/\sigma}}{\frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}u/\sigma}} du \\
&= \frac{2\sqrt{2}\sigma}{\sigma^4} \int_0^{\infty} e^{-\sqrt{2}u/\sigma} du = \frac{2\sqrt{2}\sigma}{\sigma^4} \frac{\sqrt{2}\sigma}{2} \\
&= 2/\sigma^2
\end{aligned}$$

$\Rightarrow \text{var}(\hat{A}) \geq \sigma^2/2$ The CRLB is half of that for the Gaussian case. In fact,

it can be shown that the Gaussian PDF produces the largest CRLB.

$$3) \quad p(x; A) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x(n) - Ar^n)^2}$$

$$\begin{aligned} \frac{\partial \text{LNP}}{\partial A} &= -\frac{2}{2\sigma^2} \sum_n (x(n) - Ar^n) (-r^n) \\ &= \frac{1}{\sigma^2} \sum_n (x(n) - Ar^n) r^n \end{aligned}$$

$$\frac{\partial^2 \text{LNP}}{\partial A^2} = -\frac{1}{\sigma^2} \sum_n r^{2n}$$

$$-E\left[\frac{\partial^2 \text{LNP}}{\partial A^2}\right] = \frac{1}{\sigma^2} \sum_n r^{2n}$$

$$\text{or } \text{var}(\hat{A}) \geq \frac{\sigma^2}{\sum_{n=0}^{N-1} r^{2n}} \quad (\text{or use (3.14)})$$

To show that an efficient estimator exists

$$\begin{aligned} \frac{\partial \text{LNP}}{\partial A} &= \frac{1}{\sigma^2} \left(\sum_n x(n) r^n - A \sum_n r^{2n} \right) \\ &= \underbrace{\frac{\sum_n r^{2n}}{\sigma^2}}_{I(A)} \left(\underbrace{\frac{\sum_n x(n) r^n}{\sum_n r^{2n}}}_{\hat{A}} - A \right) \end{aligned}$$

\hat{A} is efficient and $1/I(A)$ is its variance.

$$\text{var}(\hat{A}) \rightarrow \sigma^2(1-r^2) \quad 0 < r < 1$$

$$\rightarrow 0 \quad r \geq 1$$

as $N \rightarrow \infty$.

$$5) \quad p(\underline{x}; r) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - r^n)^2}$$

$$\frac{\partial \text{LNP}}{\partial r} = \frac{1}{\sigma^2} \sum_n (x[n] - r^n) n r^{n-1}$$

Can't put into form $I(r)(\hat{r} - r) \Rightarrow$
no efficient estimator.

$$\frac{\partial^2 \text{LNP}}{\partial r^2} = \frac{1}{\sigma^2} \sum_n \left[x[n] n(n-1) r^{n-2} - n(2n-1) r^{2n-2} \right]$$

$$\begin{aligned} E \left[\frac{\partial^2 \text{LNP}}{\partial r^2} \right] &= \frac{1}{\sigma^2} \sum_n \left[n(n-1) r^{2n-2} - n(2n-1) r^{2n-2} \right] \\ &= -\frac{1}{\sigma^2} \sum_n n^2 r^{2n-2} \end{aligned}$$

$$\text{var}(\hat{r}) \geq \frac{\sigma^2}{\sum_{n=0}^{N-1} n^2 r^{2n-2}} \quad (\text{could also use (3.14)})$$

$$5) \quad p(\underline{x}; A) = \frac{1}{(2\pi)^{N/2} \det(\underline{C})^{1/2}} e^{-\frac{1}{2} (\underline{x} - A\underline{1})^T \underline{C}^{-1} (\underline{x} - A\underline{1})}$$

where $\underline{1} = [1 \ 1 \ \dots \ 1]^T$

$$\frac{\partial \text{LNP}}{\partial A} = -\frac{1}{2} \frac{\partial}{\partial A} \left[\underline{x}^T \underline{C}^{-1} \underline{x} - 2 \underline{1}^T \underline{C}^{-1} \underline{x} A + \underline{1}^T \underline{C}^{-1} \underline{1} A^2 \right]$$

$$\begin{aligned}
 &= \underline{1}^T \underline{C}^{-1} \underline{x} - \underline{1}^T \underline{C}^{-1} \underline{1} A \\
 &= \underbrace{\underline{1}^T \underline{C}^{-1} \underline{1}}_{I(A)} \left(\underbrace{\frac{\underline{1}^T \underline{C}^{-1} \underline{x}}{\underline{1}^T \underline{C}^{-1} \underline{1}}}_{\hat{A}} - A \right)
 \end{aligned}$$

\hat{A} is efficient and has variance $1/I(A)$.

6)

$$\begin{aligned}
 X(0) &\sim N(\theta, 1) \\
 X(1) &\sim N(\theta, 1) \quad \theta \geq 0 \\
 &\quad N(\theta, 2) \quad \theta < 0
 \end{aligned}$$

$$\begin{aligned}
 p(\underline{x}; \theta) &= \frac{1}{2\pi} e^{-\frac{1}{2} [(x(0)-\theta)^2 + (x(1)-\theta)^2]} \quad \theta \geq 0 \\
 &\quad \frac{1}{2\pi\sqrt{2}} e^{-\frac{1}{2} [(x(0)-\theta)^2 + \frac{1}{2}(x(1)-\theta)^2]} \quad \theta < 0
 \end{aligned}$$

For $\theta \geq 0$

$$\begin{aligned}
 \frac{\partial \ln p}{\partial \theta} &= -\frac{1}{2} [2(x(0)-\theta)(-1) + 2(x(1)-\theta)(-1)] \\
 &= (x(0)-\theta) + (x(1)-\theta)
 \end{aligned}$$

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

$$\text{var}(\hat{\theta}) \geq \frac{1}{2}$$

For $\theta < 0$

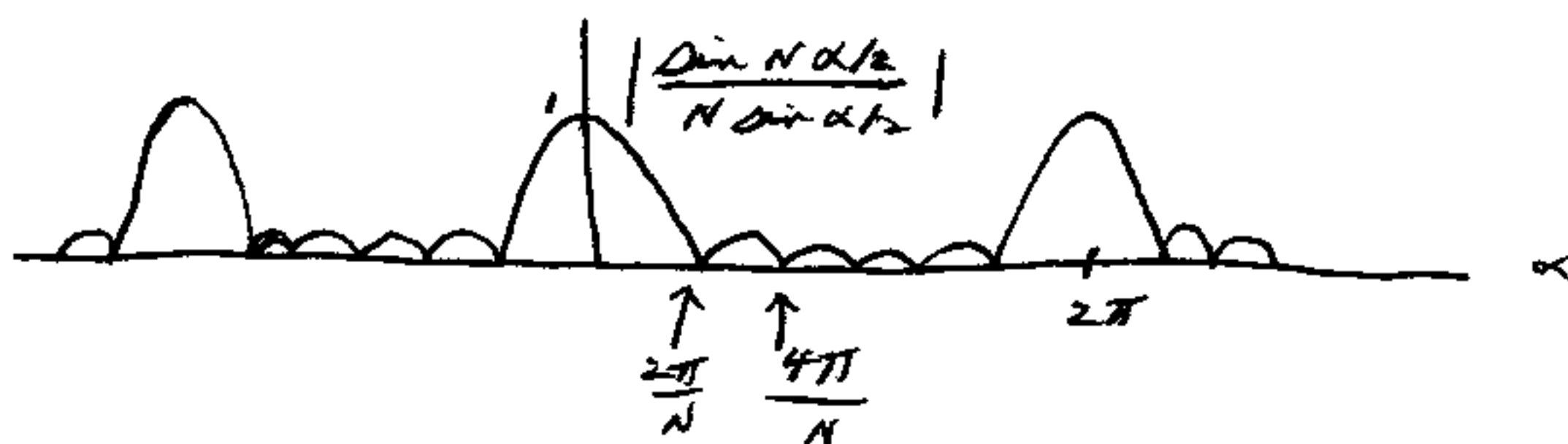
$$\begin{aligned}
 \frac{\partial \ln p}{\partial \theta} &= -\frac{1}{2} [-2(x(0)-\theta) - (x(1)-\theta)] \\
 &= (x(0)-\theta) + \frac{1}{2}(x(1)-\theta)
 \end{aligned}$$

$$\frac{\partial^2 L_{NP}}{\partial \theta^2} = -3/2 \Rightarrow E\left[-\frac{\partial^2 L_{NP}}{\partial \theta^2}\right] = 3/2$$

$$\text{var}(\hat{\theta}) \geq 2/3$$

7) Let $\alpha = 4\pi f_0$, $\beta = 2\phi$

$$\begin{aligned} \frac{1}{N} \sum_{n=0}^{N-1} \cos(\alpha n + \beta) &= \frac{1}{N} \text{Re} \left[\sum_n e^{j(\alpha n + \beta)} \right] \\ &= \frac{1}{N} \text{Re} \left[e^{j\beta} \frac{1 - e^{j\alpha N}}{1 - e^{j\alpha}} \right] \\ &= \frac{1}{N} \text{Re} \left[e^{j\beta} \frac{e^{j\alpha N/2}}{e^{j\alpha/2}} \frac{e^{-j\alpha N/2} - e^{j\alpha N/2}}{e^{-j\alpha/2} - e^{j\alpha/2}} \right] \\ &= \frac{1}{N} \text{Re} \left[e^{j\beta} e^{j\alpha(N-1)/2} \frac{\sin N\alpha/2}{\sin \alpha/2} \right] \\ &= \frac{\sin N\alpha/2}{N \sin \alpha/2} \cos \left[\alpha \left(\frac{N-1}{2} \right) + \beta \right] \end{aligned}$$



As long as α is not near 0 or 2π , this term is approximately zero. But $\alpha = 4\pi f_0 \Rightarrow f_0$ cannot be near 0 or $1/2$.

$$8) \quad p(\underline{x}; A) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2}$$

$$\frac{\partial \ln p}{\partial A} = \frac{1}{\sigma^2} \sum_n (x[n] - A)$$

$$\left(\frac{\partial \ln p}{\partial A} \right)^2 = \frac{1}{\sigma^4} \sum_m \sum_n (x[m] - A)(x[n] - A)$$

$$\begin{aligned} E \left[\left(\frac{\partial \ln p}{\partial A} \right)^2 \right] &= \frac{1}{\sigma^4} \sum_m \sum_n \underbrace{E[(x[m] - A)(x[n] - A)]}_{\sigma^2 \delta_{mn}} \\ &= \frac{1}{\sigma^4} \sum_n \sigma^2 = N/\sigma^2 \end{aligned}$$

$$\text{var}(\hat{A}) \geq \sigma^2/N$$

9) Using (3.32)

$$\mathbf{I}(A) = \left(\frac{\partial \underline{\mu}(A)}{\partial A} \right)^T \mathbf{C}^{-1} \frac{\partial \underline{\mu}(A)}{\partial A}$$

$$\underline{\mu}(A) = [A \ A]^T \Rightarrow \frac{\partial \underline{\mu}(A)}{\partial A} = \underline{1}$$

$$\text{var}(\hat{A}) \geq \frac{1}{\underline{1}^T \mathbf{C}^{-1} \underline{1}} \quad (\text{or use approach of Prob 3.5})$$

$$\mathbf{C}^{-1} = \frac{1}{\sigma^2} \frac{\begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}}{1 - \rho^2}$$

$$\Rightarrow \text{var}(\hat{A}) \geq \frac{\sigma^2(1 - \rho^2)}{2 - 2\rho} = \frac{\sigma^2}{2} (1 + \rho)$$

If $\rho = 0$, $\text{var}(\hat{A}) \geq \sigma^2/2$, as expected.

If $\rho \rightarrow 1$, $\text{var}(\hat{A}) \geq \sigma^2$. This is the same bound as for one sample and occurs because as $\rho \rightarrow 1$, $W[0]$ and $W[1]$ will be equal. Hence, we have only one independent data sample. If $\rho \rightarrow -1$, $\text{var}(\hat{A}) \geq 0$ and in fact, in this case $W[0] = -W[1]$. Thus,

$$\begin{aligned}\hat{A} &= \frac{1}{2}(X[0] + X[1]) \\ &= \frac{1}{2}(A + W[0] + A - W[0]) = A\end{aligned}$$

for any realization of the noise samples. Additivity property of Fisher information only holds for independent samples. In this example we could have

$$i(A) \leq \mathcal{I}(A) < \infty i(A)$$

where $i(A) = 1/\sigma^2$.

$$10) \quad [\mathcal{I}(\underline{\theta})]_{ij} = E \left[\frac{\partial \text{LNP}}{\partial \theta_i} \frac{\partial \text{LNP}}{\partial \theta_j} \right]$$

$$\mathcal{I}(\underline{\theta}) = E \left[\frac{\partial \text{LNP}}{\partial \underline{\theta}} \frac{\partial \text{LNP}^T}{\partial \underline{\theta}} \right]$$

$$\begin{aligned}\underline{a}^T \mathcal{I}(\underline{\theta}) \underline{a} &= E \left[\underline{a}^T \frac{\partial \text{LNP}}{\partial \underline{\theta}} \frac{\partial \text{LNP}^T}{\partial \underline{\theta}} \underline{a} \right] \\ &= E \left[\left(\underline{a}^T \frac{\partial \text{LNP}}{\partial \underline{\theta}} \right)^2 \right] \geq 0\end{aligned}$$

for all $\underline{a} \Rightarrow \mathcal{I}(\underline{\theta})$ is positive semidefinite

for all $\underline{\theta}$.

From Prob 3.3 $\underline{\theta} = \begin{bmatrix} A \\ r \end{bmatrix}$ and using (3.31)

$$[\underline{I}(\underline{\theta})]_{ij} = \frac{1}{\sigma^2} \frac{\partial \underline{\mu}(\underline{\theta})}{\partial \theta_i} \frac{\partial \underline{\mu}(\underline{\theta})}{\partial \theta_j}$$

$$\text{where } \underline{\mu}(\underline{\theta}) = \begin{bmatrix} A \\ Ar \\ \vdots \\ Ar^{N-1} \end{bmatrix}$$

$$\frac{\partial \underline{\mu}(\underline{\theta})}{\partial A} = [1 \ r \ \dots \ r^{N-1}]^T$$

$$\frac{\partial \underline{\mu}(\underline{\theta})}{\partial r} = A [0 \ 1 \ \dots \ (N-1) \ r^{N-2}]^T$$

$$\underline{I}(\underline{\theta}) = \frac{1}{\sigma^2} \begin{bmatrix} \sum_{n=0}^{N-1} r^n & A \sum_{n=0}^{N-1} n r^{2n-1} \\ A \sum_{n=0}^{N-1} n r^{2n-1} & A^2 \sum_{n=0}^{N-1} n^2 r^{2n-2} \end{bmatrix}$$

If $A = 0$, $\underline{I}(\underline{\theta})$ is not positive definite since its determinant is zero. Clearly, in this case there is no information in the data about r .

- 11) Since $\underline{I}(\underline{\theta})$ is positive definite, $a > 0$, $c > 0$ and $\det(\underline{I}(\underline{\theta})) > 0$ or $ac - b^2 > 0$. But

$$(\mathbf{I}^{-1}(\underline{\theta}))_{ii} = \frac{c}{ac - b^2} = \frac{1}{a - b^2/c} \geq \frac{1}{a}$$

Thus, the CRLB is almost always increased when we estimate additional parameters.

Equality holds if and only if $b = 0$ or the Fisher information matrix is "decoupled", i.e., it is diagonal. In this case the additional parameter does not affect the CRLB.

$$12) \quad 1^2 = (\underline{e}_i^T \sqrt{\mathbf{I}(\underline{\theta})} \sqrt{\mathbf{I}^{-1}(\underline{\theta})} \underline{e}_i)^2$$

$$\text{Since } \sqrt{\mathbf{I}^{-1}(\underline{\theta})} = (\sqrt{\mathbf{I}(\underline{\theta})})^{-1}$$

$$1^2 \leq \underline{e}_i^T \mathbf{I}(\underline{\theta}) \underline{e}_i \quad \underline{e}_i^T \mathbf{I}^{-1}(\underline{\theta}) \underline{e}_i$$

$$1 \leq [\mathbf{I}(\underline{\theta})]_{ii} [\mathbf{I}^{-1}(\underline{\theta})]_{ii}$$

$$\Rightarrow [\mathbf{I}^{-1}(\underline{\theta})]_{ii} \geq \frac{1}{[\mathbf{I}(\underline{\theta})]_{ii}}$$

New bound achieved when an efficient estimator exists and $\mathbf{I}(\underline{\theta})$ is diagonal.

$$13) \text{ From (3.33) with } s(n; \underline{\theta}) = \sum_{k=0}^{p-1} A_k n^k$$

$$[I(\theta)]_{ij} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \left(\frac{\partial}{\partial \theta_i} \sum_{k=0}^{p-1} A_k n^k \right) \left(\frac{\partial}{\partial \theta_j} \sum_{k=0}^{p-1} A_k n^k \right)$$

$$= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} n^{i-1} n^{j-1} = \frac{1}{\sigma^2} \sum_{n=0}^{N-1} n^{i+j-2}$$

$$\text{where } \theta_i = A_{i-1} \quad i=1, 2, \dots, p$$

$$14) \quad \hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x(n)$$

If we condition the mean and variance on A , then we can regard A as the observed value A_0 . Hence, from Example 3.3

$$E(\hat{A} | A = A_0) = A_0$$

$$\text{var}(\hat{A} | A = A_0) = \sigma^2/N$$

and since $\sigma^2/N \rightarrow 0$ as $N \rightarrow \infty$, $\hat{A} \rightarrow A_0$.

Now consider A as a random variable

as $N \rightarrow \infty$

$$\begin{aligned} \text{var}(\hat{\sigma}_A^2) &= \text{var}(\hat{A}^2) \\ &\rightarrow \text{var}(A^2) \end{aligned}$$

$$\text{But } \text{var}(A^2) = E(A^4) - E(A^2)^2$$

$$= 3\sigma_A^4 - \sigma_A^4 = 2\sigma_A^4$$

so that $\text{var}(\hat{\sigma}_A^2) \rightarrow 2\sigma_A^4$, which is just the CRLB as $N \rightarrow \infty$. $\hat{\sigma}_A^2$ cannot be estimated without error since even as

$N \rightarrow \infty$, although we can nullify the noise effects by averaging ($\hat{A} \rightarrow A_0$), we cannot reduce the random nature of A . This is because we have only one realization of A . Since $\hat{\sigma}_A^2$ is the square of \hat{A} , it will also exhibit the same variability.

15) Because the $X(W)$'s are independent

$$I(p) = N i(p)$$

where $i(p)$ is the Fisher information for a single vector sample. Using (3.32)

$$i(p) = \frac{1}{2} \pi \left[\left(\underline{C}^{-1}(p), \frac{\partial \underline{C}(p)}{\partial p} \right)^2 \right]$$

$$\underline{C}^{-1}(p) = \frac{\begin{bmatrix} 1 & -p \\ -p & 1 \end{bmatrix}}{1-p^2} \quad \frac{\partial \underline{C}(p)}{\partial p} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\underline{D} = \underline{C}^{-1}(p) \frac{\partial \underline{C}(p)}{\partial p} = \frac{\begin{bmatrix} -p & 1 \\ 1 & -p \end{bmatrix}}{1-p^2}$$

$$\underline{D}^2 = \frac{\begin{bmatrix} -p & 1 \\ 1 & -p \end{bmatrix} \begin{bmatrix} -p & 1 \\ 1 & -p \end{bmatrix}}{(1-p^2)^2} = \frac{\begin{bmatrix} 1+p^2 & - \\ - & 1+p^2 \end{bmatrix}}{(1-p^2)^2}$$

$$i(p) = \frac{1}{2} \frac{2+2p^2}{(1-p^2)^2} = \frac{1+p^2}{(1-p^2)^2}$$

$$\text{var}(\hat{\rho}) \geq \frac{(1-\rho^2)^2}{N(1+\rho^2)}$$

$$16) \quad \mathbf{I}(\rho_0) = \frac{1}{2} \text{tr} \left[\left(\underline{\mathbf{C}}^{-1}(\rho_0) \frac{\partial \underline{\mathbf{C}}(\rho_0)}{\partial \rho_0} \right)^2 \right]$$

$$\begin{aligned} \text{Let } r_{xx}[k] &= \mathcal{F}^{-1} \{ P_{xx}(f) \} \\ &= \mathcal{F}^{-1} \{ P_0 Q(f) \} \\ &= P_0 \mathcal{F}^{-1} \{ Q(f) \} \end{aligned}$$

Let $\mathcal{F}^{-1} \{ Q(f) \} = g[k]$ and construct the Toeplitz autocorrelation matrix $\underline{\mathbf{C}}_g$ (of dimension $N \times N$), where

$$(\underline{\mathbf{C}}_g)_{ij} = g[i-j]$$

Then, $\underline{\mathbf{C}}(\rho_0) = P_0 \underline{\mathbf{C}}_g$ and

$$\underline{\mathbf{C}}^{-1}(\rho_0) \frac{\partial \underline{\mathbf{C}}(\rho_0)}{\partial \rho_0} = \frac{1}{P_0} \underline{\mathbf{C}}_g^{-1} \underline{\mathbf{C}}_g = \frac{1}{P_0} \mathbf{I}$$

$$\mathbf{I}(\rho_0) = \frac{1}{2} \text{tr} \left[\frac{1}{P_0^2} \mathbf{I}^2 \right] = \frac{N}{2P_0^2}$$

$$\text{var}(\hat{\rho}_0) \geq 2P_0^2/N$$

Using the asymptotic form

$$\begin{aligned} \mathbf{I}(\rho_0) &= \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\partial \ln P_{xx}(f; \rho_0)}{\partial \rho_0} \right)^2 df \\ &= \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\partial \ln P_0 Q(f)}{\partial \rho_0} \right)^2 df \end{aligned}$$

$$= \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{P_0^2} df = \frac{N}{2P_0^2}$$

The two CRLB's are identical but in general this will not be true.

- 17) All elements of $\underline{I}(\underline{\theta})$ are the same except the sums run from $n = -M$ to $n = M$. Thus, now $(\underline{I}(\underline{\theta}))_{23} = 0$ since

$$\sum_{n=-M}^M n = 0$$

This makes $\underline{I}(\underline{\theta})$ diagonal. Letting $N = 2M+1$

$$\text{var}(\hat{A}) \geq 2\sigma^2/N \quad \text{same as before}$$

$$\text{var}(\hat{\phi}) \geq \frac{2\sigma^2}{NA^2} = \frac{1}{N\eta} \quad \text{less than before}$$

$$\text{var}(\hat{f}_0) \geq \frac{\sigma^2}{2A^2\pi^2 \sum_{n=-M}^M n^2}$$

$$\text{But } \sum_{n=-M}^M n^2 = 2 \sum_{n=1}^M n^2 = \frac{2M(M+1)(2M+1)}{6}$$

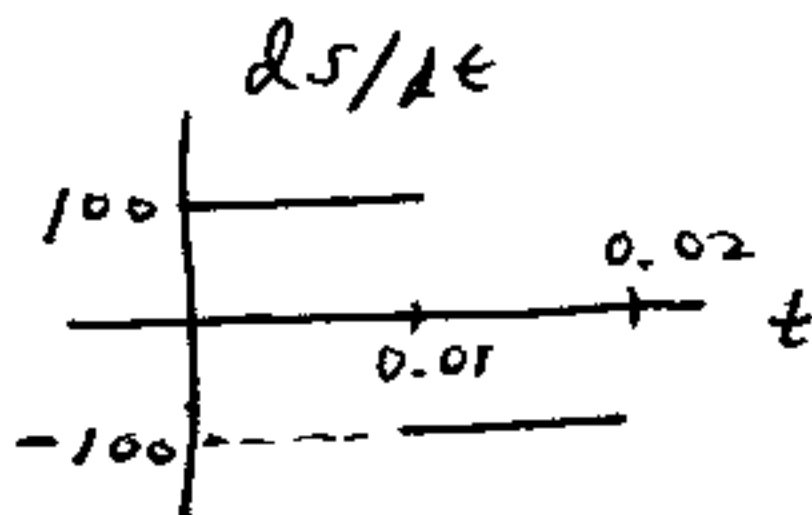
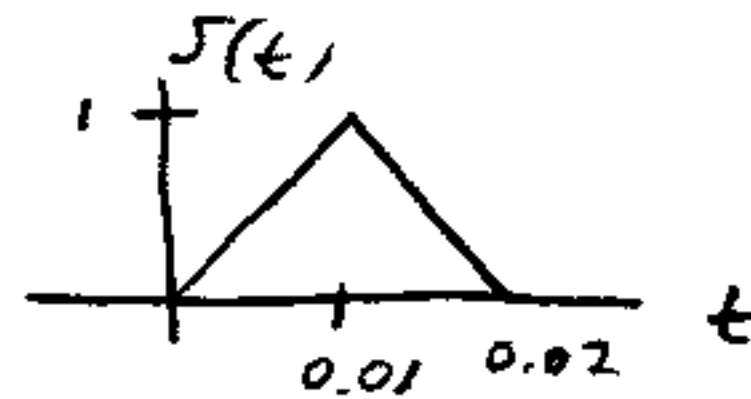
$$= \frac{1}{3} \left(\frac{N-1}{2} \right) \left(\frac{N+1}{2} \right) N = \frac{N(N^2-1)}{12}$$

$$\text{var}(\hat{f}_0) \geq \frac{6\sigma^2}{A^2\pi^2 N(N^2-1)} = \frac{3}{\eta\pi^2 N(N^2-1)}$$

$$= \frac{12}{(2\pi)^2 \eta N(N^2-1)} \quad \text{Same as before}$$

$$18) \quad \text{var}(\hat{R}) \geq \frac{C^2/4}{\frac{\epsilon}{N_0/2} \bar{F}^2}$$

$$\text{where } \bar{F}^2 = \frac{\int_0^{T_s} \left(\frac{ds}{dt}\right)^2 dt}{\int_0^{T_s} s^2(t) dt}$$



$$\bar{F}^2 = \frac{\int_0^{0.02} (100)^2 dt}{\epsilon} = \frac{200}{\epsilon}$$

$$\text{var}(\hat{R}) \geq \frac{C^2/4}{\frac{1}{N_0/2} 200} = \frac{(1500)^2/4}{10^6 \cdot 200}$$

$$= 0.00281$$

$$\text{or } \sqrt{\text{var}(\hat{R})} \geq 0.05 \text{ m}$$

19) From Example 3.15

$$\text{var}(\hat{\beta}) \geq \frac{12}{(2\pi)^2 M \eta \frac{M+1}{M-1} \left(\frac{L}{\lambda}\right)^2 \sin^2 \beta}$$

$$\text{For } \beta = 90^\circ, \eta = 1, F_0 = 10, L = (M-1)d \\ = (M-1)\lambda/2$$

$$\text{var}(\hat{\beta}) \geq \frac{12}{(2\pi)^2 M \frac{M+1}{M-1} \left(\frac{M-1}{2}\right)^2}$$

$$= \frac{12}{(2\pi)^2 \frac{M}{4} (M^2-1)}$$

$$M(M^2-1) \geq \frac{48}{(2\pi)^2 (5\pi/180)^2} = 159.7$$

or $M \geq 6$. But then

$$L = (M-1)\lambda/2 = (M-1) \frac{c}{2F_0} = \frac{5(3 \times 10^8)}{2 \times 10^6} \\ = 750 \text{ m}$$

This is clearly impossible.

$$\begin{aligned} 20) \quad \text{var}(\hat{P}_{xx}(f)) &\geq \frac{\left(\frac{\partial P_{xx}(f)}{\partial a(l)}\right)^2}{I(a(l))} \\ &= \frac{\left(\frac{\partial P_{xx}(f)}{\partial a(l)}\right)^2}{N/(1-a(l))^2} \end{aligned}$$

using results from Example 3.16.

$$P_{xx}(f) = \frac{\sigma_u^2}{|A(f)|^2} \quad \text{where } A(f) = 1 + a(1) e^{-j2\pi f}$$

$$\frac{\partial P_{xx}(f)}{\partial a(1)} = \sigma_u^2 \frac{\partial}{\partial a(1)} \left(\frac{1}{A(f)A^*(f)} \right)$$

$$= - \frac{\sigma_u^2}{|A(f)|^4} \frac{\partial}{\partial a(1)} A(f)A^*(f)$$

$$= - \frac{\sigma_u^2}{|A(f)|^4} (A(f) e^{j2\pi f} + A^*(f) e^{-j2\pi f})$$

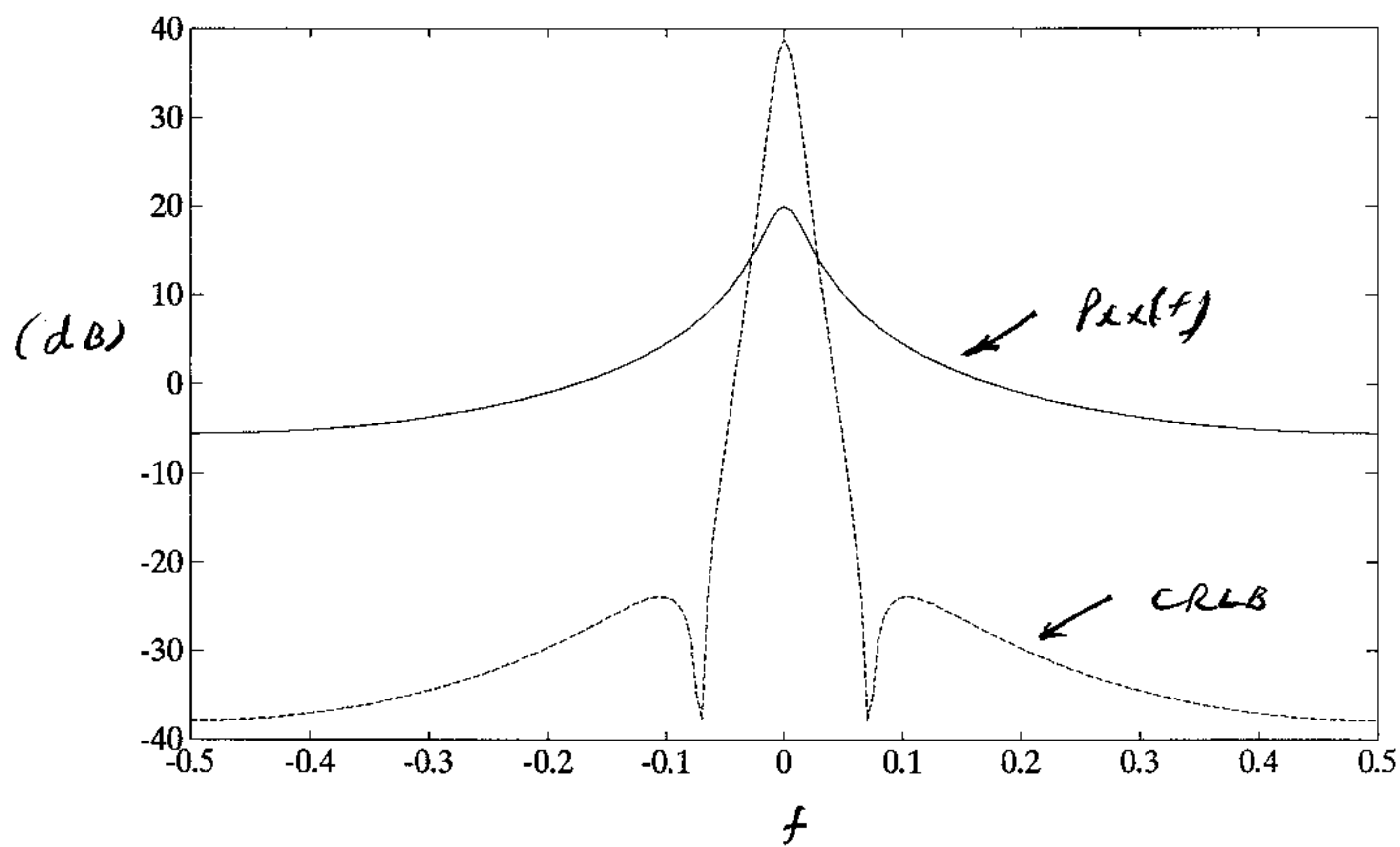
$$= \frac{-\sigma_u^2}{|A(f)|^4} \underbrace{2 \operatorname{Re}(A(f) e^{j2\pi f})}_{a(1) + \cos 2\pi f}$$

$$\operatorname{var}(\hat{P}_{xx}(f)) \geq \frac{\frac{4\sigma_u^4}{N} (1 - a^2(1)) (a(1) + \cos 2\pi f)^2}{|A(f)|^8}$$

For the given values

$$\begin{aligned} \operatorname{var}(\hat{P}_{xx}(f)) &\geq 0.0076 \frac{(a(1) + \cos 2\pi f)^2}{|A(f)|^8} \\ &= \frac{0.0076 (-0.9 + \cos 2\pi f)^2}{|1 - 0.9 e^{-j2\pi f}|^8} \end{aligned}$$

See Figure.



Prob. 3.20

Because of the sensitivity of the PSD to small changes in $a(\omega)$ for ω near zero, the variance is highest at DC. Note that

$$\begin{aligned} \left. \frac{\partial P_{xx}(\omega)}{\partial a(\omega)} \right|_{\omega=0} &= \frac{-\sigma_u^2 \cdot 2(1+a(\omega))}{(1+a(\omega))^4} \\ &= \frac{-2\sigma_u^2}{(1+a(\omega))^3} \end{aligned}$$

and for this example

$$\left[\left. \frac{\partial P_{xx}(\omega)}{\partial a(\omega)} \right|_{\omega=0} \right]^2 = 4 \times 10^6$$

Chapter 4

1) This fits linear model form.

$$\underline{X} = \underline{H} \underline{\theta} + \underline{W}$$

$$\underline{H} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ r_1 & r_2 & \dots & r_p \\ \vdots & \vdots & & \vdots \\ r_1^{N-1} & r_2^{N-1} & \dots & r_p^{N-1} \end{bmatrix} \quad \underline{\theta} = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_p \end{bmatrix}$$

$$\hat{\underline{\theta}} = (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{X} \quad C_{\hat{\underline{\theta}}} = \sigma^2 (\underline{H}^T \underline{H})^{-1}$$

For $p=2$, $r_1=1$, $r_2=-1$ and N even

$$\underline{H} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ \vdots & \vdots \\ 1 & -1 \end{bmatrix} \Rightarrow \underline{H}^T \underline{H} = \begin{bmatrix} N & 0 \\ 0 & N \end{bmatrix} = N \underline{I}$$

since columns are orthogonal

$$\hat{\underline{\theta}} = \frac{1}{N} \underline{H}^T \underline{X} = \begin{bmatrix} \frac{1}{N} \sum_{n=0}^{N-1} X[n] \\ \frac{1}{N} \sum_{n=0}^{N-1} (-1)^n X[n] \end{bmatrix}$$

$$C_{\hat{\underline{\theta}}} = \frac{\sigma^2}{N} \underline{I}$$

2) First assume columns of \underline{H} are linearly independent
Then, $\underline{H} \underline{x} = \sum_{i=1}^p x_i \underline{h}_i \neq 0$ if $\underline{x} \neq 0$

$$\Rightarrow \underline{x}^T \underline{H}^T \underline{H} \underline{x} = \|\underline{H} \underline{x}\|^2 > 0 \text{ for } \underline{x} \neq 0$$

$\Rightarrow \underline{H}^T \underline{H}$ is positive definite

Now assume $\underline{H}^T \underline{H}$ is positive definite or

$$\underline{x}^T \underline{H}^T \underline{H} \underline{x} > 0 \quad \text{for all } \underline{x} \neq \underline{0}$$

$$\text{or} \quad \|\underline{H} \underline{x}\|^2 > 0 \quad \text{for all } \underline{x} \neq \underline{0}$$

$$\Rightarrow \underline{H} \underline{x} \neq \underline{0} \quad \text{for all } \underline{x} \neq \underline{0}$$

\Rightarrow columns of \underline{H} are linearly independent

It can further be shown that for matrices of the form $\underline{H}^T \underline{H}$, invertibility is equivalent to being positive definite.

$$3) \quad \underline{H}^T \underline{H} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1+\epsilon \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1+\epsilon \end{bmatrix} = \begin{bmatrix} 3 & 3+\epsilon \\ 3+\epsilon & 2+(1+\epsilon)^2 \end{bmatrix}$$

$$(\underline{H}^T \underline{H})^{-1} = \frac{1}{\begin{vmatrix} 2+(1+\epsilon)^2 & -(3+\epsilon) \\ -(3+\epsilon) & 3 \end{vmatrix}} \begin{bmatrix} 2+(1+\epsilon)^2 & -(3+\epsilon) \\ -(3+\epsilon) & 3 \end{bmatrix}$$

$$= \frac{1}{3[2+(1+\epsilon)^2] - (3+\epsilon)^2}$$

$$= \frac{1}{2\epsilon^2} \begin{bmatrix} 3+2\epsilon+\epsilon^2 & -(3+\epsilon) \\ -(3+\epsilon) & 3 \end{bmatrix}$$

$$2\epsilon^2$$

As $\epsilon \rightarrow 0$, all elements $\rightarrow \infty$

$$\begin{aligned}
\hat{\underline{\theta}} &= (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{x} \\
&= \frac{1}{2\epsilon^2} \begin{bmatrix} 3+2\epsilon+\epsilon^2 & -(3+\epsilon) \\ -(3+\epsilon) & 3 \end{bmatrix} \begin{bmatrix} 6 \\ 6+2\epsilon \end{bmatrix} \\
&= \frac{1}{2\epsilon^2} \begin{bmatrix} 18+12\epsilon+6\epsilon^2-18-6\epsilon-6\epsilon-2\epsilon^2 \\ -18-6\epsilon+18+6\epsilon \end{bmatrix} \\
&= \begin{bmatrix} 2 \\ 0 \end{bmatrix}
\end{aligned}$$

Hence, even as $\epsilon \rightarrow 0$, $\hat{\underline{\theta}} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$. This is because \underline{x} lies in the subspace spanned by the first column of \underline{H} , which does not depend on ϵ .

$$4) \quad \hat{\underline{\theta}} \sim N(\underline{\theta}, \sigma^2 (\underline{H}^T \underline{H})^{-1})$$

$$\hat{\underline{y}} = \underline{H} \hat{\underline{\theta}} \sim N(\underline{H} \underline{\theta}, \sigma^2 \underline{H} (\underline{H}^T \underline{H})^{-1} \underline{H}^T)$$

Note that the covariance matrix is singular since $\underline{H} (\underline{H}^T \underline{H})^{-1} \underline{H}^T$ is a projection matrix of rank p . (See Chapter 8).

For Example 4.2

$$\hat{y}(n) = \sum_{k=1}^M \hat{a}_k \cos 2\pi k \frac{n}{N} + \sum_{k=1}^M \hat{b}_k \sin 2\pi k \frac{n}{N}$$

where $\hat{a}_k = \frac{2}{N} \sum_{n=0}^{N-1} x(n) \cos 2\pi k \frac{n}{N}$

$$\hat{b}_k = \frac{2}{N} \sum_{n=0}^{N-1} x(n) \sin 2\pi k \frac{n}{N}$$

Also, $(\underline{H}^T \underline{H})^{-1} = \frac{2}{N} \underline{I} \Rightarrow$

$$\underline{\hat{y}} \sim N(\underline{y}, \frac{2\sigma^2}{N} \underline{H} \underline{H}^T) \quad \text{where } \underline{y} = \underline{H} \underline{\theta}.$$

5) Let $\omega_k = \frac{2\pi}{N} k n$

$$\sum_{n=0}^{N-1} \cos \omega_k \cos \omega_l = \frac{1}{2} \sum_n [\cos(\omega_k + \omega_l) + \cos(\omega_k - \omega_l)]$$

$$= \frac{1}{2} \operatorname{Re} \sum_n e^{j(\omega_k + \omega_l)} + \frac{1}{2} \operatorname{Re} \sum_n e^{j(\omega_k - \omega_l)}$$

But $\sum_{n=0}^{N-1} e^{j(\omega_k + \omega_l)} = \sum_{n=0}^{N-1} e^{j \frac{2\pi}{N} (k+l)n}$

$$= \frac{1 - e^{j \frac{2\pi}{N} (k+l)N}}{1 - e^{j \frac{2\pi}{N} (k+l)}} = \frac{1 - e^{j 2\pi (k+l)}}{1 - e^{j \frac{2\pi}{N} (k+l)}} = 0$$

And similarly for $\sum_n e^{j(\omega_k - \omega_l)}$.

6) From Example 4.2

$$\hat{a}_k \sim N(a_k, 2\sigma^2/N) \quad \hat{b}_k \sim N(b_k, 2\sigma^2/N)$$

and \hat{a}_k, \hat{b}_k are independent

$$\begin{aligned}
 E(\hat{P}) &= \frac{E(\hat{a}_k^2) + E(\hat{b}_k^2)}{2} \\
 &= \frac{\text{var}(\hat{a}_k) + E(\hat{a}_k)^2 + \text{var}(\hat{b}_k) + E(\hat{b}_k)^2}{2} \\
 &= \frac{a_k^2 + b_k^2 + 4\sigma^2/N}{2} = P + \frac{2\sigma^2}{N}
 \end{aligned}$$

$$\text{var}(\hat{P}) = \frac{\text{var}(\hat{a}_k^2) + \text{var}(\hat{b}_k^2)}{4}$$

since \hat{a}_k, \hat{b}_k are independent

But $\text{var}(\hat{a}_k^2)$ can be found as follows:

$$\text{If } Z \sim N(\mu, \sigma^2)$$

$$\text{var}(Z^2) = 4\mu^2\sigma^2 + 2\sigma^4$$

(see development preceding (3.19))

$$\Rightarrow \text{var}(\hat{a}_k^2) = 4a_k^2 \frac{2\sigma^2}{N} + 2\left(\frac{2\sigma^2}{N}\right)^2$$

$$\text{var}(\hat{b}_k^2) = 4b_k^2 \frac{2\sigma^2}{N} + 2\left(\frac{2\sigma^2}{N}\right)^2$$

$$\text{var}(\hat{P}) = 2P \frac{2\sigma^2}{N} + \left(\frac{2\sigma^2}{N}\right)^2 = \frac{2\sigma^2}{N} \left(2P + \frac{2\sigma^2}{N}\right)$$

$$\frac{E(\hat{P})^2}{\text{var}(\hat{P})} = \frac{(P + 2\sigma^2/N)^2}{\frac{2\sigma^2}{N} (2P + \frac{2\sigma^2}{N})}$$

When no signal is present or $P = 0$, the measure is one.

For example, if $p \gg 4\sigma^2/N$ or $NP \gg 4\sigma^2$,

$$\frac{E(\hat{P})^2}{\text{var}(\hat{P})} = \frac{p^2}{4p\sigma^2/N} = \frac{p}{\frac{4\sigma^2}{N}} \gg 1$$

and the sinusoid will be easily detectable.

$$7) \quad [\underline{H}^T \underline{H}]_{ij} = \sum_{n=1}^N u[n-i]u[n-j]$$

For large N this is $\sum_{n=-\infty}^{\infty} u[n-i]u[n-j]$ since for $u[n]=0$, $n \leq 0$ and $n \geq N-1$, this will add only about $2p$ additional terms. If $N \gg p$, these terms will be negligible.

Now assume $i \geq j$ and let $m = n - i$

$$\begin{aligned} [\underline{H}^T \underline{H}]_{ij} &= \sum_{m=-\infty}^{\infty} u[m]u[m+i-j] \\ &= \sum_{m=0}^{N-1-(i-j)} u[m]u[m+i-j] \end{aligned}$$

Similarly for $i < j$

$$[\underline{H}^T \underline{H}]_{ij} = \sum_{m=0}^{N-1-(j-i)} u[m]u[m+j-i]$$

$$\Rightarrow [\underline{H}^T \underline{H}]_{ij} = \sum_{m=0}^{N-1-|i-j|} u[m]u[m+|i-j|]$$

$$8) \quad x[n] = \sum_{l=0}^{\infty} h[l] u[n-l]$$

$$\begin{aligned} r_{ux}[k] &= E[u[n] x[n+k]] \\ &= E\left[u[n] \sum_{l=0}^{\infty} h[l] u[n+k-l]\right] \\ &= \sum_{l=0}^{\infty} h[l] E[u[n] u[n+k-l]] \\ &= \sum_{l=0}^{\infty} h[l] r_{uu}[k-l] \\ &= h[k] \star r_{uu}[k] \end{aligned}$$

$$\Rightarrow P_{ux}(f) = H(f) P_{uu}(f)$$

If $P_{uu}(f) = \sigma^2$, then $H(f) = P_{ux}(f) / \sigma^2$
or

$$\hat{H}(f) = \frac{\hat{P}_{ux}(f)}{r_{uu}[0]}$$

If $P_{uu}(f) = 0$ over a band of frequencies, it would be impossible to estimate $H(f)$ over that band. This is because there would be no power over that band at the output, leading to $\hat{P}_{ux}(f) = 0$, independent of $H(f)$. The TDL estimator of (4.23) is just the inverse Fourier transform of $\hat{H}(f)$.

$$9) \quad p(\underline{x}; \underline{\theta}) = \frac{1}{(2\pi)^{N/2} \det(\underline{C})^{1/2}} e^{-\frac{1}{2}(\underline{x} - \underline{H}\underline{\theta})^T \underline{C}^{-1}(\underline{x} - \underline{H}\underline{\theta})}$$

$$\begin{aligned} \frac{\partial \ln p}{\partial \underline{\theta}} &= -\frac{1}{2} \frac{\partial}{\partial \underline{\theta}} \left[\underline{x}^T \underline{C}^{-1} \underline{x} - 2 \underline{\theta}^T \underline{H}^T \underline{C}^{-1} \underline{x} + \underline{\theta}^T \underline{H}^T \underline{C}^{-1} \underline{H} \underline{\theta} \right] \\ &= -\frac{1}{2} \left[-2 \underline{H}^T \underline{C}^{-1} \underline{x} + 2 \underline{H}^T \underline{C}^{-1} \underline{H} \underline{\theta} \right] \text{ using (4.3)} \\ &= \underline{H}^T \underline{C}^{-1} \underline{x} - \underline{H}^T \underline{C}^{-1} \underline{H} \underline{\theta} \\ &= \underbrace{\underline{H}^T \underline{C}^{-1} \underline{H}}_{\underline{I}(\underline{\theta})} \underbrace{(\underline{H}^T \underline{C}^{-1} \underline{H})^{-1} \underline{H}^T \underline{C}^{-1} \underline{x} - \underline{\theta}}_{\hat{\underline{\theta}}} \end{aligned}$$

$\therefore \hat{\underline{\theta}} = (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1} \underline{H}^T \underline{C}^{-1} \underline{x}$ is MVU estimator (and efficient)

$$\underline{C}_{\hat{\underline{\theta}}} = (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1}$$

$$10) \quad \underline{C}^{-1} = \underline{D}^T \underline{D}$$

$$\text{Since } \underline{C} = \text{diag}(\sigma_0^2, \sigma_1^2, \dots, \sigma_{N-1}^2)$$

$$\underline{C}^{-1} = \text{diag}(1/\sigma_0^2, 1/\sigma_1^2, \dots, 1/\sigma_{N-1}^2)$$

$$\Rightarrow \underline{D} = \text{diag}(1/\sigma_0, 1/\sigma_1, \dots, 1/\sigma_{N-1})$$

$$\hat{A} = \sum_{n=0}^{N-1} d_n x[n]$$

$$\text{where } d_n = \frac{[\underline{D}^{-1}]_n}{\underline{1}^T \underline{D}^{-1} \underline{1}} = \frac{1/\sigma_n}{\sum_{m=0}^{N-1} 1/\sigma_m^2}$$

Since \underline{C} is already diagonal or the

components of \underline{w} are uncorrelated, we need only form $\underline{x}' = \underline{D}\underline{x}$ or $x'(n) = x(n)/\sigma_n$ so that all variances are one. Then, we "average" the $x'(n)$ samples. Actually, we weight the samples since the DC level has been changed to a non-DC signal due to the prewhitening stage.

If a $\sigma_n^2 = 0$, say σ_m^2 , then we cannot prewhiten the data. In this case, however, as $\sigma_m^2 \rightarrow 0$, $d_n \rightarrow 0$ for $n \neq m$ and $d_n \rightarrow \sigma_m$ for $n = m$. Thus, $\hat{A} \rightarrow d_m x'(m) = x(m)$, as expected.

$$11) \quad \hat{A} = (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1} \underline{H}^T \underline{C}^{-1} \underline{x} \quad \text{var}(\hat{A}) = (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1}$$

$$\underline{C} = \sigma^2 \begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix} \quad \underline{H} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\underline{C}^{-1} = \frac{1}{\sigma^2(1-p^2)} \begin{pmatrix} 1 & -p \\ -p & 1 \end{pmatrix}$$

$$\underline{H}^T \underline{C}^{-1} \underline{H} = \frac{2-2p}{\sigma^2(1-p^2)} = \frac{2(1-p)}{\sigma^2(1-p^2)} = \frac{2}{\sigma^2(1+p)}$$

$$\begin{aligned} \underline{H}^T \underline{C}^{-1} \underline{x} &= \frac{1^T}{\sigma^2(1-p^2)} \begin{pmatrix} 1 & -p \\ -p & 1 \end{pmatrix} \begin{pmatrix} x(0) \\ x(1) \end{pmatrix} \\ &= \frac{x(0)(1-p) + x(1)(1-p)}{\sigma^2(1-p^2)} \end{aligned}$$

$$= \frac{x(0) + x(1)}{\sigma^2(1+p)}$$

$$\hat{A} = \frac{\sigma^2(1+p)}{2} \frac{x(0) + x(1)}{\sigma^2(1+p)} = \frac{1}{2}(x(0) + x(1))$$

$$\text{var}(\hat{A}) = \frac{\sigma^2(1+p)}{2}$$

We don't need prewhitening here because \underline{H} is an eigenvector of \underline{C} . Hence,

$$\begin{aligned} (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1} \underline{H}^T \underline{C}^{-1} &= (\underline{H}^T \frac{1}{\lambda} \underline{H})^{-1} \frac{1}{\lambda} \underline{H}^T \\ &= (\underline{H}^T \underline{H})^{-1} \underline{H}^T \end{aligned}$$

As $\rho \rightarrow 1$, $\text{var}(\hat{A}) \rightarrow \sigma^2$

As $\rho \rightarrow -1$, $\text{var}(\hat{A}) \rightarrow 0$.

See Prob 3.9 for explanation.

$$12) \quad \text{If } \underline{x} = \underline{H}\underline{\theta} + \underline{N}, \quad \underline{x}' = \underline{A}(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{x}$$

$$\begin{aligned} \underline{x}' &= \underline{A}(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{H} \underline{\theta} + \underline{A}(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{N} \\ \uparrow &= \underline{A} \underline{\theta} + \underline{N}' = \underline{H}' \underline{\alpha} + \underline{N}' \quad \text{where } \underline{H}' = \underline{I} \\ r \times 1 & \quad \quad \quad \uparrow \\ & \quad \quad \quad r \times 1 \end{aligned}$$

$$\begin{aligned} \underline{C}' &= E(\underline{N}' \underline{N}'^T) = E(\underline{A}(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{N} \underline{N}^T \underline{H}(\underline{H}^T \underline{H})^{-1} \underline{A}^T) \\ &= \sigma^2 \underline{A}(\underline{H}^T \underline{H})^{-1} \underline{A}^T \end{aligned}$$

Since A is full rank, C^{-1} is positive definite and C^{-1} exists.

$$\Rightarrow \hat{\underline{\alpha}} = (\underline{H}^T \underline{C}^{-1} \underline{H})^{-1} \underline{H}^T \underline{C}^{-1} \underline{x}$$

$$= \underline{C}^{-1} \underline{C}^{-1} \underline{x} = \underline{A} (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{x} = \underline{A} \hat{\underline{\theta}}$$

$$\begin{aligned} 13) \quad E(\hat{\underline{\theta}}) &= E\left\{(\underline{H}^T \underline{H})^{-1} \underline{H}^T (\underline{H} \underline{\theta} + \underline{w})\right\} \\ &= \underline{\theta} + E\left\{(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{w}\right\} \\ &= \underline{\theta} + E\left\{(\underline{H}^T \underline{H})^{-1} \underline{H}^T\right\} E(\underline{w}) \\ &= \underline{\theta} \quad \text{since } E(\underline{w}) = \underline{0}. \end{aligned}$$

$$\begin{aligned} \underline{C}_{\hat{\theta}} &= E\left\{(\hat{\underline{\theta}} - \underline{\theta})(\hat{\underline{\theta}} - \underline{\theta})^T\right\} \\ &= E\left\{(\underline{H}^T \underline{H})^{-1} \underline{H}^T (\underline{x} - \underline{H} \underline{\theta})\right. \\ &\quad \left.(\underline{H}^T \underline{H})^{-1} \underline{H}^T (\underline{x} - \underline{H} \underline{\theta})^T\right\} \\ &= E\left\{(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{w} \underline{w}^T \underline{H} (\underline{H}^T \underline{H})^{-1}\right\} \\ &= E_{H|W} E_W \left\{(\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{w} \underline{w}^T \underline{H} (\underline{H}^T \underline{H})^{-1}\right\} \\ &= E_{H|W} \left[(\underline{H}^T \underline{H})^{-1} \underline{H}^T \sigma^2 \underline{I} \underline{H} (\underline{H}^T \underline{H})^{-1}\right] \\ &= E_{H|W} \left[\sigma^2 (\underline{H}^T \underline{H})^{-1}\right] = \sigma^2 E_H \left[(\underline{H}^T \underline{H})^{-1}\right] \end{aligned}$$

Since \underline{H} and \underline{w} are independent. If \underline{H} and \underline{w} are not independent $\hat{\underline{\theta}}$ may be biased.

$$14) \quad \underline{H} = \underline{1} \quad \text{with probability } 1 - \epsilon$$

$$\underline{H} = \left[\underbrace{11 \dots 1}_M \underbrace{00 \dots 0}_{N-M} \right]^T \quad \text{with probability } \epsilon$$

$$\hat{\underline{A}} = (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{x}$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} x[n] \quad \text{no fade}$$

$$\frac{1}{M} \sum_{n=0}^{M-1} x[n] \quad \text{fade}$$

$$\text{var}(\hat{\underline{A}}) = \sigma^2 E_H [(\underline{H}^T \underline{H})^{-1}]$$

$$= \sigma^2 \left[\frac{1}{N} (1 - \epsilon) + \frac{1}{M} \epsilon \right]$$

$$= \frac{\sigma^2}{N} \left[1 - \epsilon + \frac{N}{M} \epsilon \right]$$

$$= \frac{\sigma^2}{N} \left[1 + \left(\frac{N}{M} - 1 \right) \epsilon \right] > \sigma^2 / N$$

Clearly, the variances are the same only if $M = N$ or $\epsilon = 0$. Otherwise, it is increased.

Chapter 5

$$\begin{aligned}
 1) \quad p(\underline{x} \mid T(\underline{x}) = T_0; \sigma^2) &= \frac{p(\underline{x}; \sigma^2) \delta(T(\underline{x}) - T_0)}{p(T(\underline{x}) = T_0; \sigma^2)} \\
 &= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_n x^2(n)}}{\delta(T(\underline{x}) - T_0)} \\
 &\quad p\left(\sum_n x^2(n) = T_0; \sigma^2\right)
 \end{aligned}$$

But $\sum_n x^2(n) = T_0$

$$p\left(\sum_n x^2(n)\right) = \frac{1}{\sigma^2} p_r\left(\sum_n x^2(n) \mid \sigma^2\right)$$

$$\begin{aligned}
 p(\underline{x} \mid T(\underline{x}) = T_0; \sigma^2) &= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} T_0} \delta(T(\underline{x}) - T_0)}{\frac{1}{\sigma^2} \frac{1}{2^{N/2} \Gamma(N/2)} e^{-T_0/2\sigma^2} \left(\frac{T_0}{\sigma^2}\right)^{N/2-1}} \\
 &= \frac{\frac{1}{\pi^{N/2}} \delta(T(\underline{x}) - T_0)}{\frac{1}{\Gamma(N/2)} T_0^{N/2-1}}
 \end{aligned}$$

$$\begin{aligned}
 2) \quad p(\underline{x}; \sigma^2) &= \prod_{n=0}^{N-1} \frac{x(n)}{\sigma^2} e^{-\frac{1}{2} x^2(n)/\sigma^2} \quad \text{all } x(n) > 0 \\
 &= \underbrace{u(\min x(n)) \prod_{n=0}^{N-1} x(n)}_{h(\underline{x})} \underbrace{\frac{1}{\sigma^{2N}} e^{-\frac{1}{2} \sum_n x^2(n)/\sigma^2}}_{g(T(\underline{x}), \sigma^2)}
 \end{aligned}$$

where $u(x)$ is the unit step function

$T(\underline{x}) = \sum_{n=0}^{N-1} x^2[n]$ is a sufficient statistic

$$3) \quad p(\underline{x}; \lambda) = \prod_{n=0}^{N-1} \lambda e^{-\lambda x[n]} \quad \text{all } x[n] > 0$$

$$= \underbrace{\lambda^N e^{-\lambda \sum_n x[n]}}_{g(T(\underline{x}), \lambda)} \cdot \underbrace{u(\min x[n])}_{h(\underline{x})}$$

$T(\underline{x}) = \sum_{n=0}^{N-1} x[n]$ is a sufficient statistic

$$4) \quad p(\underline{x}; \theta) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{1}{2\theta}(x[n]-\theta)^2}$$

$$= \frac{1}{(2\pi\theta)^{N/2}} e^{-\frac{1}{2\theta} \sum_n (x[n]-\theta)^2}$$

But $\sum_n (x[n]-\theta)^2 = \sum_n x^2[n] - 2\theta \sum_n x[n] + N\theta^2$

$$p(\underline{x}; \theta) = \underbrace{\frac{1}{(2\pi\theta)^{N/2}} e^{-\frac{1}{2\theta} \sum_n x^2[n] - \frac{1}{2} N\theta}}_{g(T(\underline{x}), \theta)} \underbrace{e^{\sum_n x[n]}}_{h(\underline{x})}$$

$T(\underline{x}) = \sum_{n=0}^{N-1} x^2[n]$ is a sufficient statistic

$$5) \quad p(x[n]; \theta) = \frac{1}{2\theta} (u(x[n]+\theta) - u(x[n]-\theta))$$

$$p(\underline{x}; \theta) = \frac{1}{(2\theta)^N} \prod_{n=0}^{N-1} [u(x[n]+\theta) - u(x[n]-\theta)]$$

But the product is zero unless $-\theta \leq x[n] \leq \theta$
for all $x[n]$ or $\min x[n] \geq -\theta$, $\max x[n] \leq \theta$

or $\max |x[n]| \leq \theta$ so that

$$p(\underline{x}; \theta) = \underbrace{\frac{1}{(2\theta)^N} u(\theta - \max |x[n]|)}_{g(T(\underline{x}), \theta)} \cdot \underbrace{1}_{h(\underline{x})}$$

and $T(\underline{x}) = \max |x[n]|$ is the sufficient statistic.

$$b) \quad p(\underline{x}; \sigma^2) = \underbrace{\frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2}}_{g(T(\underline{x}), \sigma^2)} \cdot \underbrace{1}_{h(\underline{x})}$$

$T(\underline{x}) = \sum_{n=0}^{N-1} (x[n] - A)^2$ is a sufficient statistic

To make it unbiased, divide by N . Thus,

$\hat{\sigma}^2 = \frac{1}{N} \sum_{n=0}^{N-1} (x[n] - A)^2$ is the MVU estimator.

$$7) \quad p(\underline{x}; f_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - \cos 2\pi f_0 n)^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left[\sum_n x^2[n] - 2 \sum_n x[n] \cos 2\pi f_0 n + \sum_n \cos^2 2\pi f_0 n \right]}$$

Because of the $\sum_n x[n] \cos 2\pi f_0 n$ term, which cannot be separated into a statistic, there does not appear to be a sufficient statistic.

Note that $\sum_n x[n] \cos 2\pi f_0 n$ is not a statistic

since it depends on f_0 .

$$8) \quad p(\underline{x}; r) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - r^n)^2}$$

$$\text{But } \sum_n (x[n] - r^n)^2 = \sum_n x^2[n] - 2 \sum_n x[n] r^n + \sum_n r^{2n}$$

Again, the term $\sum_n x[n] r^n$ cannot be separated into a single sufficient statistic and a function of r .

$$9) \quad p_r \{x[n]\} = \theta^{x[n]} (1-\theta)^{1-x[n]} \quad x[n] = 0, 1$$

$$\begin{aligned} p_r \{x\} &= \prod_{n=0}^{N-1} \theta^{x[n]} (1-\theta)^{1-x[n]} \\ &= \underbrace{\theta^{\sum_n x[n]} (1-\theta)^{N - \sum_n x[n]}}_{g(T(x), \theta)} \cdot \underbrace{1}_{h(x)} \end{aligned}$$

where $T(x) = \sum_{n=0}^{N-1} x[n]$ is a sufficient statistic. To make it unbiased divide by N so that

$$\hat{\theta} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \text{ is the MVU estimator.}$$

10) At high SNR we ignore the noise so that

$$T_1(x) \approx \sum_n A \cos(2\pi f_0 n + \phi) \cos 2\pi f_0 n$$

$$= \sum_n \frac{A}{2} [\cos \phi + \cos(4\pi f_0 n + \phi)]$$

$$\hat{\phi} \approx NA/2 \cos \phi \quad \text{using results of Prob 3.7}$$

Also,

$$T_2(\underline{x}) \approx \sum_n A \cos(2\pi f_0 n + \phi) \sin 2\pi f_0 n$$

$$= \sum_n \frac{A}{2} [-\sin \phi + \sin(4\pi f_0 n + \phi)]$$

$$\hat{\phi} \approx -\frac{NA}{2} \sin \phi$$

$$\text{Thus, } \hat{\phi} = -\arctan \frac{T_2(\underline{x})}{T_1(\underline{x})} = -\arctan \frac{-\frac{NA}{2} \sin \phi}{\frac{NA}{2} \cos \phi} \\ = \phi$$

This is not the MVU estimator since it is not unbiased. To see this note that even if we could assume $E(T_1(\underline{x})) = \frac{NA}{2} \cos \phi$ and

$$E(T_2(\underline{x})) = -\frac{NA}{2} \sin \phi$$

(which will not be true in general - only at high SNR), it is not true that

$$E(\hat{\phi}) = E\left[-\arctan \frac{T_2(\underline{x})}{T_1(\underline{x})}\right] \\ = -\arctan \frac{E\{T_2(\underline{x})\}}{E\{T_1(\underline{x})\}} \\ \uparrow \\ \text{incorrect}$$

$$11) \quad \theta = 2A + 1 \Rightarrow A = \frac{\theta - 1}{2}$$

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

$$p'(\underline{x}; \theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \sum_n (x(n) - \frac{\theta-1}{2})^2}$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left[\sum_n x^2(n) - (\theta-1) \sum_n x(n) + N \left(\frac{\theta-1}{2} \right)^2 \right]}$$

$$= \underbrace{\frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2} \left[N \left(\frac{\theta-1}{2} \right)^2 - (\theta-1) \sum_n x(n) \right]}}_{g(T(\underline{x}), \theta)} \underbrace{e^{-\frac{1}{2\sigma^2} \sum_n x^2(n)}}_{h(\underline{x})}$$

Thus, $\sum_{n=0}^{N-1} x(n)$ is a sufficient statistic for θ .
To make it unbiased

$$E\left(\sum_n x(n)\right) = NA = N\left(\frac{\theta-1}{2}\right) = N\frac{\theta}{2} - \frac{N}{2}$$

$$\text{Let } \hat{\theta} = \frac{2}{N} \left(\sum_n x(n) + \frac{N}{2} \right) = \frac{2}{N} \sum_n x(n) + 1$$

This is the MVU estimator. Note that $\hat{\theta} = 2\hat{A} + 1$, where \hat{A} is MVU estimator for A .

For $\theta = A^3$ the sufficient statistic is again $\sum x(n)$. To make it unbiased we need a function g such that

$$E\left(g\left(\sum_n x(n)\right)\right) = A^3$$

$$\text{or } E(h(\bar{X})) = A^3 \quad \text{where } \bar{X} \sim N(A, \sigma^2/N)$$

since \bar{X} is also a sufficient statistic for θ .

Examining \bar{X}^3 we have

$$E[(\bar{X} - A)^3] = E(\bar{X}^3) - 3AE(\bar{X}^2) + 3A^2E(\bar{X}) - A^3 = 0$$

since all odd-order moments are zero.

$$E(\bar{X}^3) = 3A(A^2 + \sigma^2/N) - 3A^3 + A^3 = A^3 + 3A\sigma^2/N$$

Try $\hat{\theta} = \bar{X}^3 - 3\bar{X}\sigma^2/N$. This is unbiased and is a function of the sufficient statistic. Thus, it is the MVUE estimator.

$$12) \quad g_1(u) = \frac{1}{N}u \Rightarrow E\left[\frac{1}{N}\sum_n X(n)\right] = A$$

$$g_2(u) = \frac{1}{N}u^{1/3} \Rightarrow E\left[\frac{1}{N}\left(\sum_n X(n)\right)^3\right]^{1/3} = A$$

Also, note that $T_2 = T_1^3$, which is a one-to-one transformation. For T_3 there is no function that would make it unbiased. Also, T_3 is not a one-to-one transformation of T_1 .

$$13) \quad p(\underline{x}; \theta) = e^{-\left(\sum_n X(n) - \theta\right)} u(\min X(n) - \theta)$$

$$= \underbrace{e^{-\sum_n X(n)}}_{h(\underline{x})} \underbrace{e^{N\theta} u(\min X(n) - \theta)}_{g(T(\underline{x}), \theta)}$$

where $T(x) = \min x(n)$ is the sufficient statistic. To find the MVU we proceed as in Example 5.8.

$$\begin{aligned}
 \Pr\{T \leq z\} &= 1 - \Pr\{T \geq z\} \\
 &= 1 - \Pr\{x(0) \geq z, \dots, x(N-1) \geq z\} \\
 &= 1 - \prod_{n=0}^{N-1} \Pr\{x(n) \geq z\} \\
 &= 1 - \Pr^N\{x(n) \geq z\}
 \end{aligned}$$

$$\begin{aligned}
 p_T(z) &= \frac{d \Pr\{T \leq z\}}{dz} = -N \Pr\{x(n) \geq z\}^{N-1} \\
 &\quad \cdot \frac{d \Pr\{x(n) \geq z\}}{dz}
 \end{aligned}$$

$$\begin{aligned}
 \text{But } \frac{d \Pr\{x(n) \geq z\}}{dz} &= \frac{d [1 - \Pr\{x(n) \leq z\}]}{dz} \\
 &= - \frac{d \Pr\{x(n) \leq z\}}{dz} = \begin{cases} -e^{-(z-\theta)} & z > \theta \\ 0 & z < \theta \end{cases}
 \end{aligned}$$

and

$$\Pr\{x(n) \geq z\} = \begin{cases} 1 - \int_0^z e^{-(x-\theta)} dx & \text{for } z > \theta \\ 1 & \text{for } z < \theta \end{cases}$$

$$\begin{aligned}
 \text{For } z > \theta &= 1 + e^0 e^{-x} \Big|_0^z = 1 + e^{-(z-\theta)} - 1 \\
 &= e^{-(z-\theta)}
 \end{aligned}$$

$$\begin{aligned}
 p_T(z) &= \begin{cases} 0 & z < \theta \\ -N [e^{-(z-\theta)^N}]' [-e^{-(z-\theta)^N}] & z > \theta \end{cases} \\
 &= \begin{cases} N e^{-N(z-\theta)} & z > \theta \\ 0 & \text{otherwise} \end{cases}
 \end{aligned}$$

$$\begin{aligned}
 E(T) &= \int_{\theta}^{\infty} z N e^{-N(z-\theta)} dz \\
 &= N e^{N\theta} \int_{\theta}^{\infty} z e^{-Nz} dz \\
 &= N e^{N\theta} \left(-\frac{z}{N} e^{-Nz} - \frac{1}{N^2} e^{-Nz} \right) \Big|_{\theta}^{\infty} \\
 &= N e^{N\theta} \left(\frac{\theta}{N} e^{-N\theta} + \frac{1}{N^2} e^{-N\theta} \right) = \theta + \frac{1}{N}
 \end{aligned}$$

$\Rightarrow \hat{\theta} = T - 1/N = \min x_i - \frac{1}{N}$ is the MVU estimator.

$$\begin{aligned}
 14) \ a) \ p(x; \mu) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2 + x\mu - \frac{1}{2}\mu^2} \\
 &= e^{\underset{\substack{\uparrow \\ A(\mu) B(x)}}{x\mu} - \frac{1}{2}x^2 + \left(-\frac{1}{2}\mu^2 + \ln \frac{1}{\sqrt{2\pi}} \right)} \\
 &\quad \quad \quad \uparrow \quad \quad \quad \uparrow \quad \quad \quad \uparrow \\
 &\quad \quad \quad A(\mu) B(x) \quad C(x) \quad D(\mu)
 \end{aligned}$$

$$\begin{aligned}
 b) \ p(x; \sigma^2) &= \frac{x}{\sigma^2} e^{-\frac{1}{2}x^2/\sigma^2} u(x) \\
 &= e^{\ln x/\sigma^2 - \frac{1}{2}x^2/\sigma^2 + \ln u(x)} \\
 &= e^{\underbrace{-\frac{1}{2}x^2/\sigma^2}_{A(\sigma^2) B(x)} + \underbrace{\ln x u(x)}_{C(x)} - \underbrace{\ln \sigma^2}_{D(\sigma^2)}}
 \end{aligned}$$

$$\begin{aligned}
 c) \quad p(x; \lambda) &= \lambda e^{-\lambda x} u(x) \\
 &= \underbrace{e^{-\lambda x}}_{A(\lambda)B(x)} \underbrace{u(x)}_{C(x)} \underbrace{\lambda}_{D(\lambda)}
 \end{aligned}$$

$$\begin{aligned}
 15) \quad p(\underline{x}; \theta) &= \prod_{n=0}^{N-1} e^{A(\theta)B(x[n]) + C(x[n]) + D(\theta)} \\
 &= e^{A(\theta) \sum_n B(x[n]) + \sum_n C(x[n]) + N D(\theta)} \\
 &= \underbrace{e^{A(\theta) \sum_n B(x[n]) + N D(\theta)}}_{g(T(\underline{x}), \theta)} \underbrace{e^{\sum_n C(x[n])}}_{h(\underline{x})}
 \end{aligned}$$

$$\text{where } T(\underline{x}) = \sum_{n=0}^{N-1} B(x[n])$$

$$a) \quad B(x) = x \Rightarrow T(\underline{x}) = \sum_n x[n]$$

$$b) \quad B(x) = x^2 \Rightarrow T(\underline{x}) = \sum_n x^2[n]$$

$$c) \quad B(x) = x \Rightarrow T(\underline{x}) = \sum_n x[n]$$

Need only make $T(\underline{x})$ unbiased

$$a) \quad \hat{\mu} = 1/N \sum_n x[n]$$

$$\begin{aligned}
 b) \quad E(x^2) &= \int_0^{\infty} \frac{x^2}{\sigma^2} e^{-\frac{1}{2} x^2 / \sigma^2} dx \\
 &= 2\sigma^2
 \end{aligned}$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{2N} \sum_n x^2 \ln x$$

$$c) E(x) = \int_0^\infty x \lambda e^{-\lambda x} dx = 1/\lambda$$

It is not obvious how to make T unbiased for this PDF. However, if we reparameterize the PDF by $\theta = 1/\lambda$, the MVU of θ is easily found.

$$1b) p(\underline{x}; \underline{\theta}) = \frac{1}{(2\pi)^{N/2} \det^{1/2}(\underline{C})} e^{-\frac{1}{2}(\underline{x}-\underline{\mu})^T \underline{C}^{-1}(\underline{x}-\underline{\mu})}$$

$$\text{If we have } \underline{C} = \begin{pmatrix} a & \underline{b}^T \\ \underline{b} & \underline{I} \end{pmatrix} \quad \begin{aligned} a &= N-1 + \sigma^2 \\ \underline{b} &= -\underline{1} \end{aligned}$$

$$\begin{aligned} \det(\underline{C}) &= \det(\underline{I}) \det(a - \underline{b}^T \underline{I}^{-1} \underline{b}) \\ &= a - \underline{b}^T \underline{b} = N-1 + \sigma^2 - (-\underline{1})^T(-\underline{1}) \\ &= N-1 + \sigma^2 - (N-1) = \sigma^2 \end{aligned}$$

$$\begin{aligned} \underline{C}^{-1} &= \begin{bmatrix} (a - \underline{b}^T \underline{b})^{-1} & - (a - \underline{b}^T \underline{b})^{-1} \underline{b}^T \\ - \underline{b} (a - \underline{b}^T \underline{b})^{-1} & (\underline{I} - \frac{\underline{b} \underline{b}^T}{a})^{-1} \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sigma^2} & \frac{1}{\sigma^2} \underline{1}^T \\ \frac{1}{\sigma^2} \underline{1} & (\underline{I} - \frac{\underline{1} \underline{1}^T}{N-1 + \sigma^2})^{-1} \end{bmatrix} \end{aligned}$$

$$\text{But } (\underline{I} - \frac{\underline{1} \underline{1}^T}{N-1 + \sigma^2})^{-1} = \underline{I} + \frac{\frac{\underline{1} \underline{1}^T}{N-1 + \sigma^2}}{1 - \frac{\underline{1}^T \underline{1}}{N-1 + \sigma^2}}$$

$$= \underline{I} + \frac{\underline{1}\underline{1}^T}{\sigma^2}$$

$$\underline{C}^{-1} = \frac{1}{\sigma^2} \begin{pmatrix} 1 & \underline{1}^T \\ \underline{1} & \sigma^2 \underline{I} + \underline{1}\underline{1}^T \end{pmatrix}$$

$$(\underline{x} - \underline{\mu})^T \underline{C}^{-1} (\underline{x} - \underline{\mu}) =$$

$$\frac{1}{\sigma^2} \left[\underbrace{(x[0] - N\mu \quad x[1] \dots x[N-1])}_{\underline{x}'^T} \begin{pmatrix} 1 & \underline{1}^T \\ \underline{1} & \sigma^2 \underline{I} + \underline{1}\underline{1}^T \end{pmatrix} \begin{pmatrix} x[0] - N\mu \\ \underline{x}' \end{pmatrix} \right]$$

$$= \frac{1}{\sigma^2} \left\{ (x[0] - N\mu)^2 + (x[0] - N\mu) \underline{1}^T \underline{x}' + (x[0] - N\mu) \underline{x}'^T \underline{1} + \sigma^2 \underline{x}'^T \underline{x}' + (\underline{x}'^T \underline{1})^2 \right\}$$

$$= \frac{1}{\sigma^2} \left\{ (x[0] - N\mu + \underline{1}^T \underline{x}')^2 + \sigma^2 \underline{x}'^T \underline{x}' \right\}$$

$$= \frac{1}{\sigma^2} \left\{ \left(\sum_{n=0}^{N-1} x[n] - N\mu \right)^2 + \sigma^2 \sum_{n=1}^{N-1} x^2[n] \right\}$$

$$= \frac{N^2}{\sigma^2} \left(\bar{x} - \mu \right)^2 + \sum_{n=1}^{N-1} x^2[n]$$

$$p(\underline{x}; \underline{\theta}) = \underbrace{\frac{1}{(2\pi)^{N/2} \sigma}}_{g(T(\underline{x}), \underline{\theta})} e^{-\frac{N^2}{2\sigma^2} (\bar{x} - \mu)^2} \underbrace{e^{-\frac{1}{2} \sum_{n=1}^{N-1} x^2[n]}}_{h(\underline{x})}$$

$\Rightarrow \bar{x}$ is a sufficient statistic for $\underline{\theta}$.