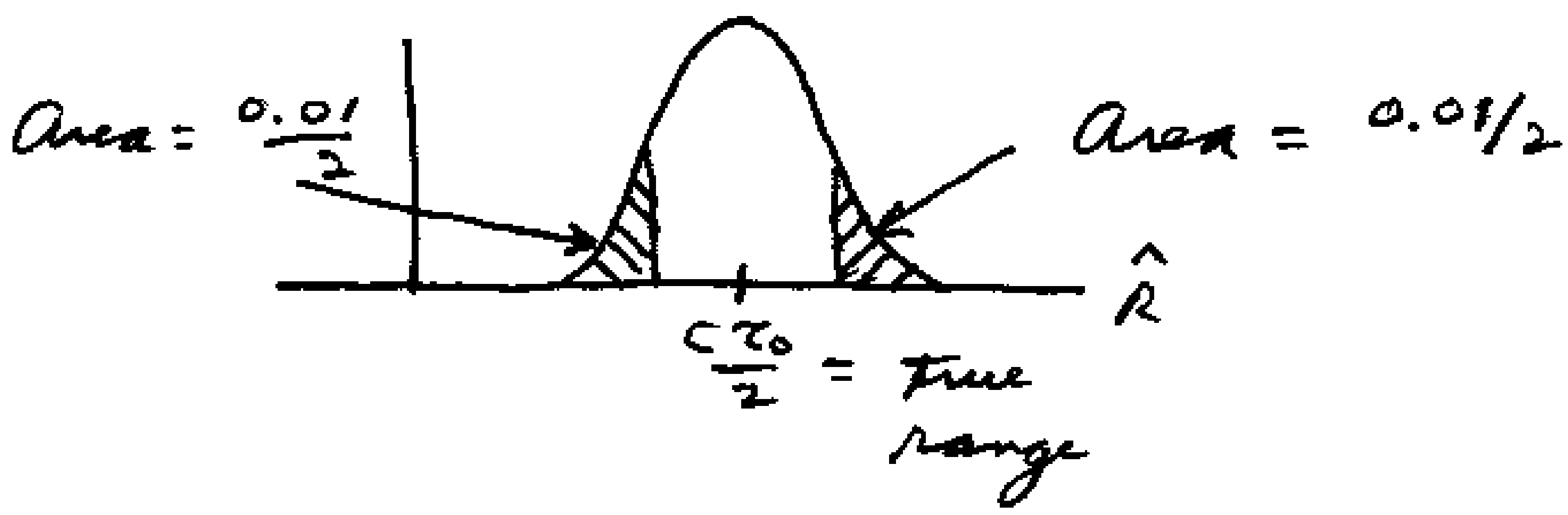
PROBLEM SOLUTIONS

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 $\hat{\tau}_0 \sim N(\tau_0, \sigma^2_{\hat{\tau}_0})$,
$$\hat{R} \sim N\left(c\tau_0/2,\ \frac{c^2}{4}\sigma^2_{\hat{\tau}_0}\right)$$



To be within 100 m we must have
$$P_r\{|\hat{R} - c\tau_0/2| < 100\} = 0.99$$
$$\Rightarrow P_r\left\{\underbrace{\left|\frac{\hat{R} - c\tau_0/2}{\frac{c}{2}\sigma_{\hat{\tau}_0}}\right|}_{N(0,1)} < \frac{100}{\frac{c}{2}\sigma_{\hat{\tau}_0}}\right\} = 0.99$$

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

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 <u>likely</u> 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,

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

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

$$\mathrm{var}(\check{A}) = \frac{1}{(N+2)^2}\left[4\sigma^2 + \sum_{n=1}^{N-2}\sigma^2 + 4\sigma^2\right]$$

$$= \frac{N+6}{(N+2)^2}\sigma^2 = \frac{N+6}{(N+2)^2}$$

$$\mathrm{var}(\check{A}) - \mathrm{var}(\hat{A}) = \frac{N+6}{(N+2)^2} - \frac{1}{N}$$

$$= \frac{N(N+6) - (N+2)^2}{N(N+2)^2}$$

$$= \frac{2N-4}{N(N+2)^2} > 0 \quad \text{for } N > 2$$

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

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

$$\text{var}(\hat{\sigma}^2) = \frac{1}{N^2}\text{var}\left(\sum_n x^2[n]\right)$$

$$= \frac{1}{N^2} N \,\text{var}(x^2[n])$$ ($x[n]$'s are IID $\Rightarrow x^2[n]$ are IID)

$$= \frac{1}{N}\text{var}(x^2[n])$$

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

$\Rightarrow \text{var}(\hat{\sigma}^2) = 2\sigma^4/N \to 0$ as $N \to \infty$
Hence, the PDF of $\hat{\sigma}^2$ collapses about the true value as $N \to \infty$.

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

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

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

$$\mathrm{var}(\hat{A}) = \mathrm{var}\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right)$$

$$= \frac{1}{N^2}\,\mathrm{var}\left(\sum_{n=0}^{N-1} x[n]\right)$$

$$= \frac{1}{N^2}\, N\, \mathrm{var}(x[n])$$

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

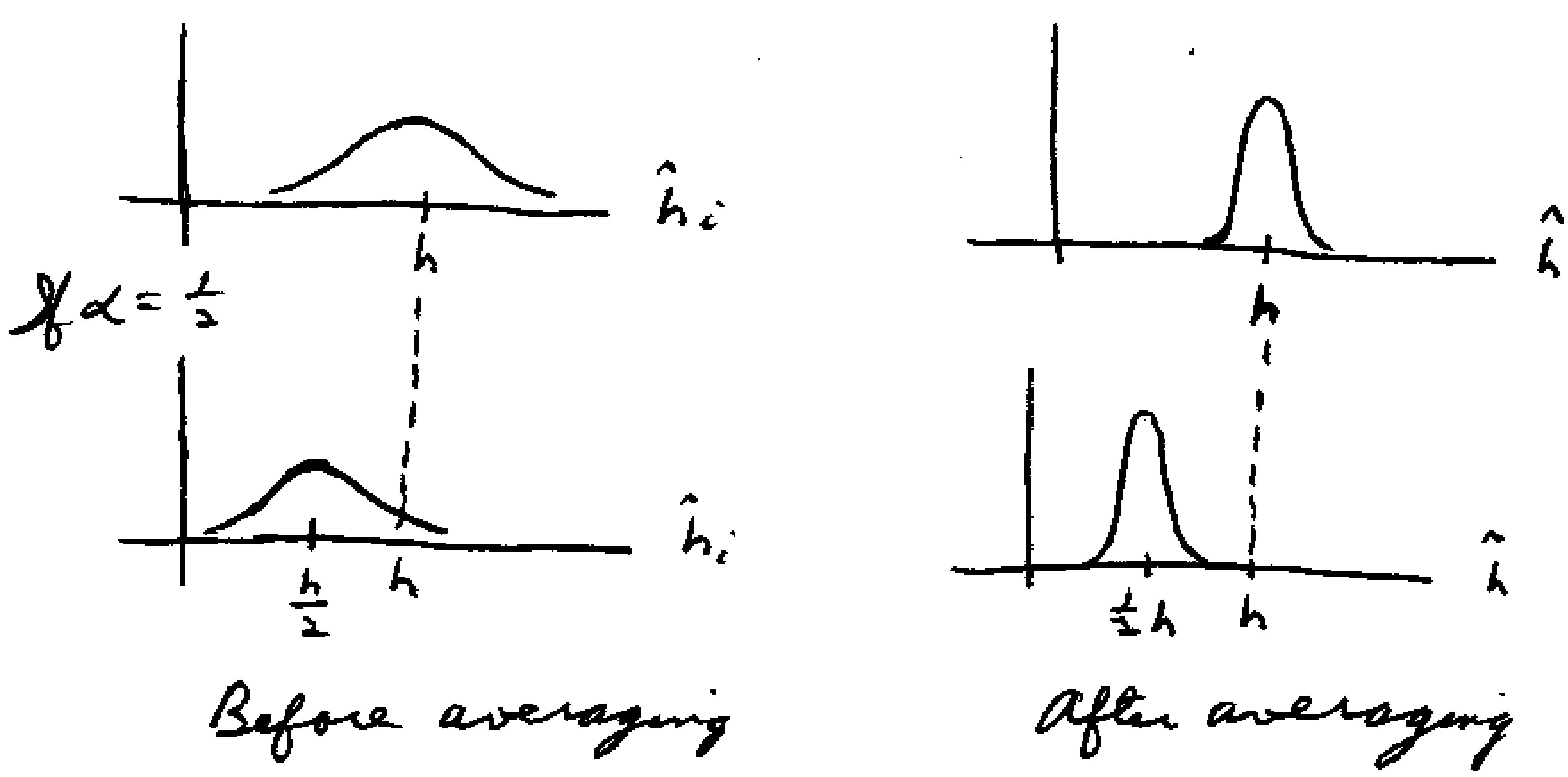
$$\Rightarrow \mathrm{var}(\hat{A}) = \sigma^2/N.$$

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

$$= \alpha h$$

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

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) $x[n] \sim \mathcal{N}(0, \sigma^2)$ or $\frac{x[n]}{\sigma} \sim \mathcal{N}(0,1)$

$\Rightarrow \left(\frac{x[n]}{\sigma}\right)^2 \sim \chi_1^2$ and

$$y = \left(\frac{x[0]}{\sigma}\right)^2 + \left(\frac{x[1]}{\sigma}\right)^2 \sim \chi_2^2$$

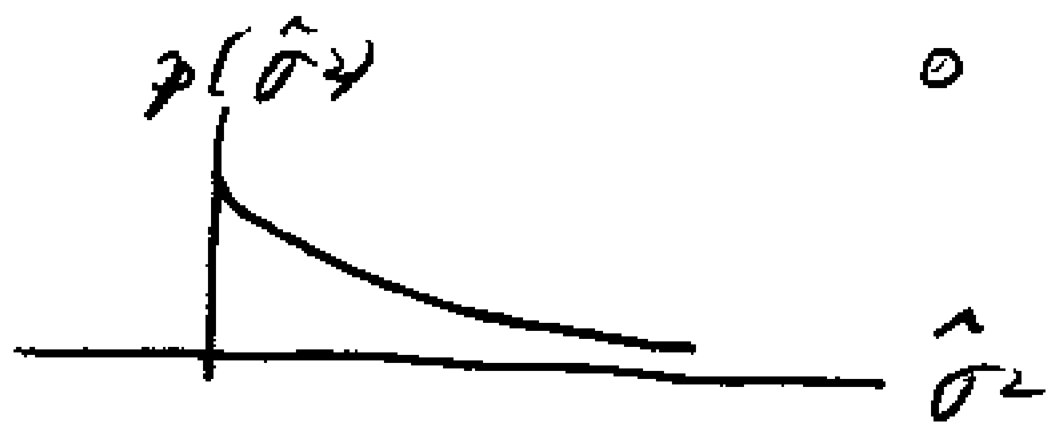
or $p(y) = \begin{cases} \frac{1}{2} e^{-y/2} & y > 0 \\ 0 & y < 0 \end{cases}$

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

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

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

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



Clearly, not symmetric.

But $E(\hat{\sigma}^2) = \int_0^\infty \hat{\sigma}^2 \frac{1}{\sigma^2} e^{-\hat{\sigma}^2/\sigma^2} d\hat{\sigma}^2$

$$= \sigma^2 \int_0^\infty u e^{-u} du$$
$$= \sigma^2 \left[-u e^{-u} - e^{-u} \right] \Big|_0^\infty$$
$$= \sigma^2$$

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

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

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

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

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

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

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) $$\frac{\hat{\theta}-\theta}{\sqrt{\mathrm{var}(\hat{\theta})}} \sim N(0,1) \qquad \frac{\check{\theta}-\theta}{\sqrt{\mathrm{var}(\check{\theta})}} \sim N(0,1)$$

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

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

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

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

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

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

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

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

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

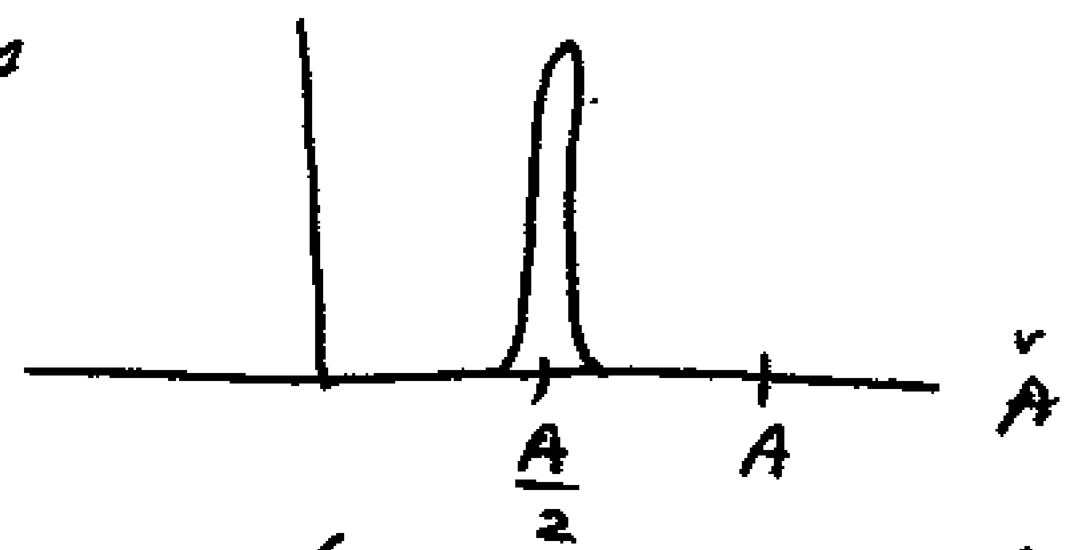
$$= 2\Phi\left(\frac{-\epsilon}{\sqrt{\sigma^2/N}}\right) \to 0$$

since $\frac{-\epsilon}{\sqrt{\sigma^2/N}} \to -\infty$ as $N \to \infty$

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

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

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



Thus, $\lim_{N \to \infty} P_r\{|\check{A} - A| > \epsilon\} = 1$ 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 $N \to \infty$ it is unbiased. $\hat{\theta}$ is said to be asymptotically unbiased (for large data records).

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

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

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

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

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

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

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

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

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

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

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

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

$\Rightarrow \hat{\underline{\theta}}$ is unbiased.

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

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

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

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

Now suppose a g could be found. Then

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

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

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

and subtracting the two gives

$$\int_{1/\theta_1}^{1/\theta_2} g(u)\,du = 0 \quad \text{for any } \theta_2 < \theta_1$$

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

Chapter 3

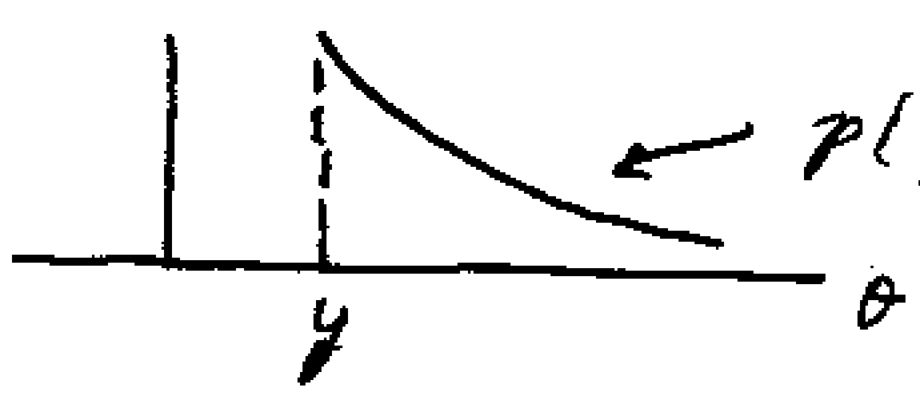
1) $$p(x[n];\theta) = \frac{1}{\theta}\left(u(x[n]) - u(x[n]-\theta)\right)$$

where $u(x) = 1 \quad x > 0$
$\qquad\qquad\quad 0 \quad x < 0$

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

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

(The expectation will be independent of n).
Let $y = x[n]$ so that
$$p(y;\theta) = \frac{1}{\theta}\left(u(y) - u(y-\theta)\right)$$



← $p(y;\theta)$ as function of θ with y fixed and $y>0$

For $\theta > y$
$$E\left[\frac{\partial \ln p(y;\theta)}{\partial\theta}\right] = E\left[\frac{\partial \ln 1/\theta}{\partial\theta}\right]$$
$$= -1/\theta \neq 0$$

2) $x[0] = A + w[0]$

$$p_x(x[0];A) = p_w(x[0]-A) = p(x[0]-A)$$

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

$$= E\left[\left(\frac{\partial \ln p(x[0]-A)}{\partial (x[0]-A)}(-1)\right)^2\right]$$

$$= E\left[\left(\frac{1}{p(x[0]-A)}\frac{\partial p(x[0]-A)}{\partial (x[0]-A)}\right)^2\right]$$

$$= \int_{-\infty}^{\infty}\left(\frac{\partial p(x[0]-A)}{\partial (x[0]-A)}\right)^2 \frac{1}{p^2(x[0]-A)} p_x(x[0];A)\, dx[0]$$

$$= \int_{-\infty}^{\infty}\left(\frac{\partial p(x[0]-A)}{\partial (x[0]-A)}\right)^2 \frac{1}{p^2(x[0]-A)} p(x[0]-A)\, dx[0]$$

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

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

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

$$= \frac{2\sqrt{2}\sigma}{\sigma^4}\int_0^{\infty} e^{-\sqrt{2}u/\sigma}\, du = \frac{2\sqrt{2}\sigma}{\sigma^4}\frac{\sqrt{2}\sigma}{2}$$

$$= 2/\sigma^2$$

$\Rightarrow$ $\mathrm{var}(\hat{A}) \geq \sigma^2/2$ 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) $$p(\underline{x};A) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x(n)-Ar^n)^2}$$

$$\frac{\partial \ln p}{\partial A} = -\frac{2}{2\sigma^2}\sum_n (x(n)-Ar^n)(-r^n)$$

$$= \frac{1}{\sigma^2}\sum_n (x(n)-Ar^n)r^n$$

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

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

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

To show that an efficient estimator exists

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

$$= \underbrace{\frac{\sum_n r^{2n}}{\sigma^2}}_{I(A)}\left(\underbrace{\frac{\sum_n x(n)r^n}{\sum_n r^{2n}}}_{\hat{A}} - A\right)$$

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

$$\operatorname{var}(\hat{A}) \to \sigma^2(1-r^2) \qquad 0 < r < 1$$

$\to 0 \qquad r \geq 1$

as $N \to \infty$.

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

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

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

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

$$E\left[\frac{\partial^2 \ln p}{\partial r^2}\right] = \frac{1}{\sigma^2}\sum_n \left[n(n-1) r^{2n-2} - n(2n-1) r^{2n-2}\right]$$

$$= -\frac{1}{\sigma^2}\sum_n n^2 r^{2n-2}$$

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

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

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

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

$$= \underline{1}^T \underline{C}^{-1} \underline{x} - \underline{1}^T \underline{C}^{-1} \underline{1} A$$

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

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

6)
$$x[0] \sim N(\theta, 1)$$
$$x[1] \sim N(\theta, 1) \quad \theta \geq 0$$
$$N(\theta, 2) \quad \theta < 0$$

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

For $\theta \geq 0$

$$\frac{\partial \ln p}{\partial \theta} = -\frac{1}{2}\left[2(x[0]-\theta)(-1) + 2(x[1]-\theta)(-1)\right]$$
$$= (x[0]-\theta) + (x[1]-\theta)$$

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

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

For $\theta < 0$

$$\frac{\partial \ln p}{\partial \theta} = -\frac{1}{2}\left(-2(x[0]-\theta) - (x[1]-\theta)\right)$$

$$= (x[0]-\theta) + \frac{1}{2}(x[1]-\theta)$$

$$\frac{\partial^2 \ln p}{\partial \theta^2} = -3/2 \Rightarrow E\left(-\frac{\partial^2 \ln p}{\partial \theta^2}\right) = 3/2$$

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

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

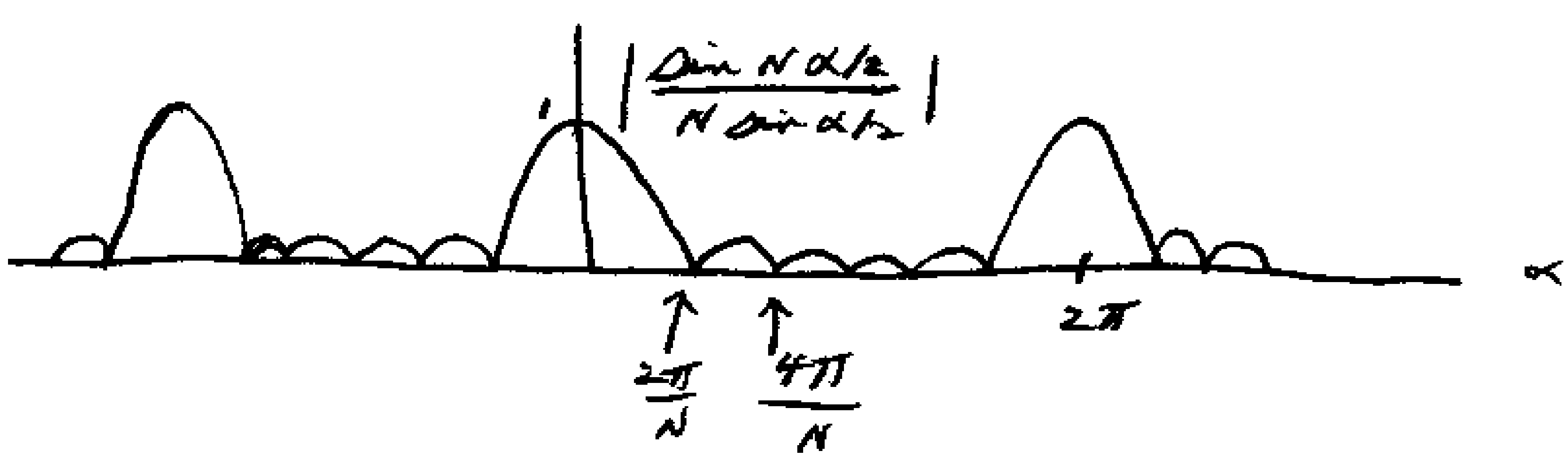
$$\frac{1}{N}\sum_{n=0}^{N-1} \cos(\alpha n + \beta) = \frac{1}{N} \mathrm{Re}\left(\sum_n e^{j(\alpha n + \beta)} \right)$$

$$= \frac{1}{N} \mathrm{Re}\left[e^{j\beta} \frac{1 - e^{j\alpha N}}{1 - e^{j\alpha}} \right]$$

$$= \frac{1}{N} \mathrm{Re}\left(e^{j\beta} \frac{e^{j\alpha N/2}}{e^{j\alpha/2}} \frac{e^{-j\alpha N/2} - e^{j\alpha N/2}}{e^{-j\alpha/2} - e^{j\alpha/2}} \right)$$

$$= \frac{1}{N} \mathrm{Re}\left[e^{j\beta} e^{j\alpha\left(\frac{N-1}{2}\right)} \frac{\sin N\alpha/2}{\sin \alpha/2} \right]$$

$$= \frac{\sin N\alpha/2}{N \sin \alpha/2} \cos\left[\alpha\left(\frac{N-1}{2}\right) + \beta\right)$$



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

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

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

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

$$= \frac{1}{\sigma^4}\sum_n \sigma^2 = N/\sigma^2$$

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

9) Using (3.32)

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

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

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

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

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

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

$$\hat{A} = \frac{1}{2}(x[0] + x[1])$$
$$= \frac{1}{2}(A + w[0] + A - w[0]) = A$$

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

$$i(A) \leq I(A) < \infty \, i(A)$$

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

10) $$[\underline{I}(\underline{\theta})]_{ij} = E\left[\frac{\partial \ln p}{\partial \theta_i} \frac{\partial \ln p}{\partial \theta_j}\right]$$

$$\underline{I}(\underline{\theta}) = E\left[\frac{\partial \ln p}{\partial \underline{\theta}} \frac{\partial \ln p}{\partial \underline{\theta}}^T\right]$$

$$\underline{a}^T \underline{I}(\underline{\theta}) \underline{a} = E\left[\underline{a}^T \frac{\partial \ln p}{\partial \underline{\theta}} \frac{\partial \ln p}{\partial \underline{\theta}}^T \underline{a}\right]$$

$$= E\left[\left(\underline{a}^T \frac{\partial \ln p}{\partial \underline{\theta}}\right)^2\right] \geq 0$$

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

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

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

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

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

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

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

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

$$\left(\underline{I}^{-1}(\underline{\theta})\right)_{11} = \frac{c}{ac-b^2} = \frac{1}{a-b^2/c} \geq \frac{1}{a}$$

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) $$1^2 = \left(\underline{e}_i^T \sqrt{\underline{I}(\underline{\theta})}\sqrt{\underline{I}^{-1}(\underline{\theta})}\,\underline{e}_i\right)^2$$

since $\sqrt{\underline{I}^{-1}(\underline{\theta})} = \left(\sqrt{\underline{I}(\underline{\theta})}\right)^{-1}$

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

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

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

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

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

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

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

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

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

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

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

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

and since $\sigma^2/N \to 0$ as $N \to \infty$, $\hat{A} \to A_0$.
Now consider A as a random variable
As $N \to \infty$

$$\mathrm{var}(\hat{\sigma}_A^2) = \mathrm{var}(\hat{A}^2)$$
$$\to \mathrm{var}(A^2)$$

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

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

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

$N \to \infty$, although we can nullify the noise effects by averaging ($\hat{A} \to A_0$), 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 $x(n)$'s are independent

$$I(\rho) = N\, i(\rho)$$

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

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

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

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

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

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

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

16) $$I(P_0) = \frac{1}{2} \mathrm{tr}\left[\left(\underline{C}^{-1}(P_0) \frac{\partial \underline{C}(P_0)}{\partial P_0}\right)^2\right]$$

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

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

$$(\underline{C}_q)_{ij} = q[i-j]$$

Then, $\underline{C}(P_0) = P_0 \underline{C}_q$ and

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

$$I(P_0) = \frac{1}{2} \mathrm{tr}\left(1/P_0^2 \, \underline{I}^2\right) = \frac{N}{2P_0^2}$$

$$\mathrm{var}(\hat{P}_0) \geq 2P_0^2/N$$

Using the asymptotic form

$$I(P_0) = \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\partial \ln P_{xx}(f;P_0)}{\partial P_0}\right)^2 df$$

$$= \frac{N}{2} \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\frac{\partial \ln P_0 Q(f)}{\partial P_0}\right)^2 df$$

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

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 $N = 2M+1$

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

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

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

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

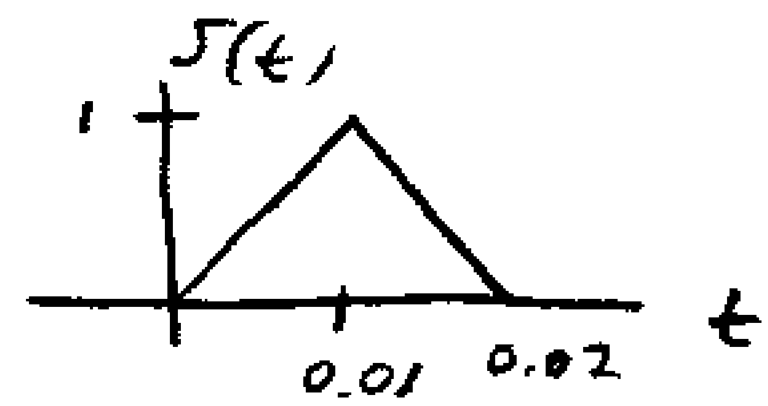
$$= \frac{1}{3}\left(\frac{N-1}{2}\right)\left(\frac{N+1}{2}\right)N = \frac{N(N^2-1)}{12}$$

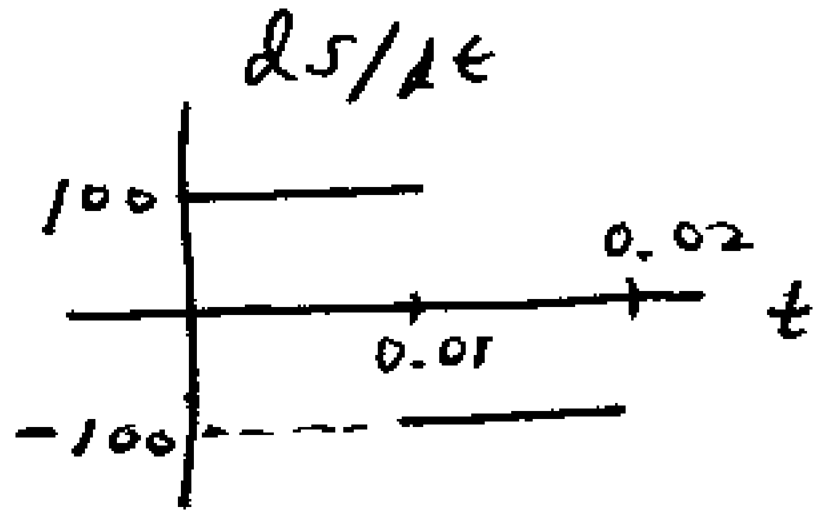
$$\text{var}(\hat{f}_0) \geq \frac{6\sigma^2}{A^2\pi^2 N(N^2-1)} = \frac{3}{\eta\pi^2 N(N^2-1)}$$

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

18) $$\mathrm{var}(\hat{R}) \geq \frac{c^2/4}{\frac{\mathcal{E}}{N_0/2}\,\overline{F^2}}$$

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





$$\overline{F^2} = \frac{\int_0^{0.02} (100)^2 dt}{\mathcal{E}} = \frac{200}{\mathcal{E}}$$

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

$$= 0.00281$$

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

19) From Example 3.15

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

For $\beta = 90°$, $\eta = 1$, $F_0 = 10$, $L = (M-1)d$
$= (M-1)\lambda/2$

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

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

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

or $M \geq 6$. But then

$$L = (M-1)\lambda/2 = (M-1)\frac{c}{2F_0} = \frac{5(3\times10^8)}{2\times10^6}$$

$= 750$ m

This is clearly impossible.

20) $$\mathrm{var}(\hat{P}_{xx}(f)) \geq \frac{\left(\frac{\partial P_{xx}(f)}{\partial a(1)}\right)^2}{I(a(1))}$$

$$= \frac{\left(\partial P_{xx}(f)/\partial a(1)\right)^2}{N/(1-a(1))^2}$$

using results from Example 3.16.

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

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

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

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

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

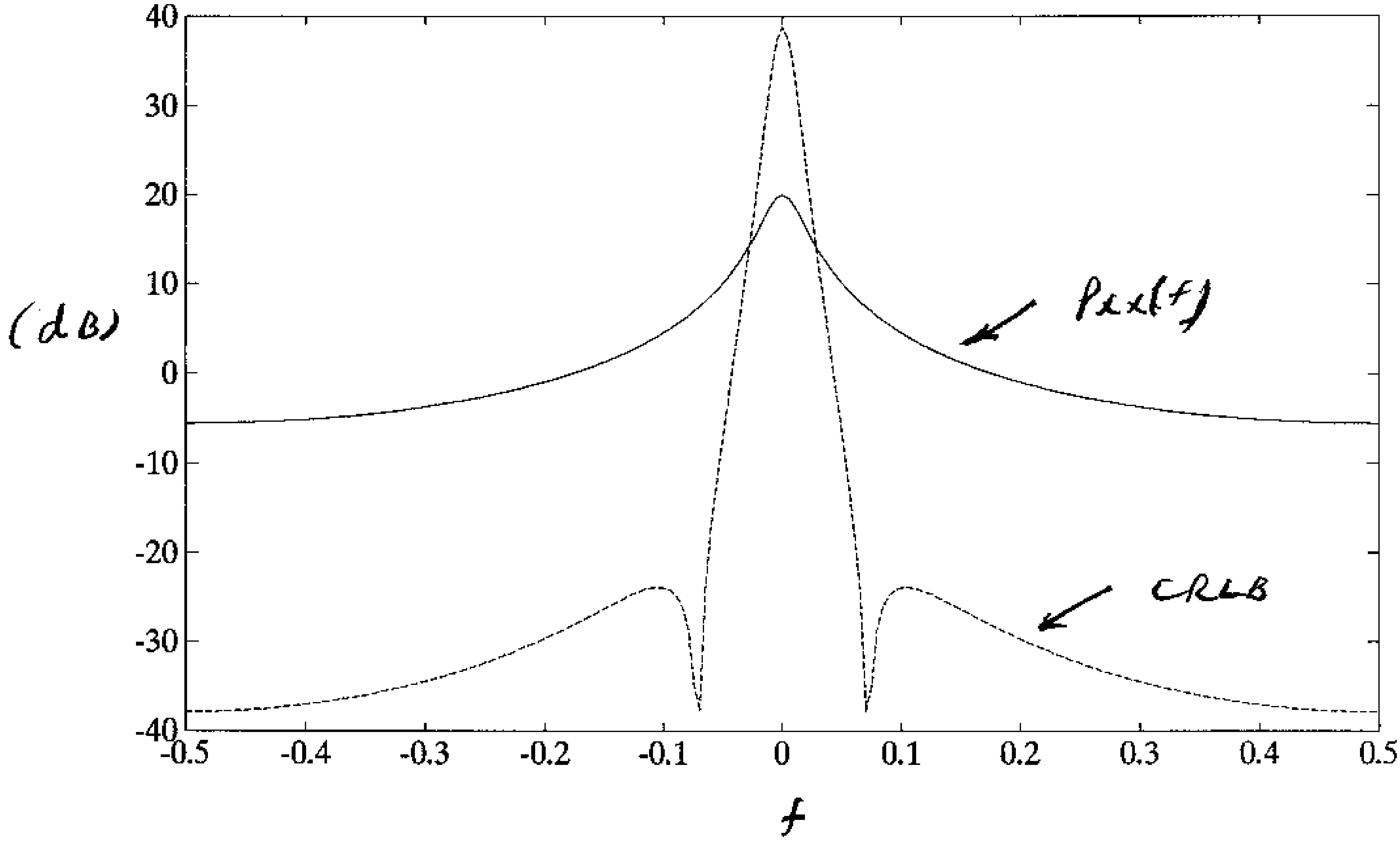
$$\operatorname{var}(\hat{P}_{xx}(f)) \geq \frac{4\sigma_u^4}{N} (1 - a^2[1]) \frac{(a[1] + \cos 2\pi f)^2}{|A(f)|^8}$$

For the given values

$$\operatorname{var}(\hat{P}_{xx}(f)) \geq 0.0076 \frac{(a[1] + \cos 2\pi f)^2}{|A(f)|^8}$$

$$= \frac{0.0076 \, (-0.9 + \cos 2\pi f)^2}{|1 - 0.9 e^{-j2\pi f}|^8}$$

See Figure.



Prob. 3.20

Because of the sensitivity of the PSD to small changes in $a[1]$ for f near zero, the variance is highest at DC. Note that

$$\frac{\partial P_{xx}(f)}{\partial a[1]}\Big|_{f=0} = -\sigma_u^2 \frac{2(1+a[1])}{(1+a[1])^4}$$

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

and for this example

$$\left[\partial P_{xx}(f)/\partial a[1]\Big|_{f=0}\right]^2 = 4\times 10^6$$

Chapter 4

1) This fits linear model form.

$$\underline{x} = \underline{H}\,\underline{\theta} + \underline{w}$$

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

$$\hat{\underline{\theta}} = (\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{x} \qquad \underline{C}_{\hat{\theta}} = \sigma^2(\underline{H}^T\underline{H})^{-1}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= \frac{1}{2\epsilon^2}\begin{bmatrix} 3+2\epsilon+\epsilon^2 & -(3+\epsilon) \\ -(3+\epsilon) & 3 \end{bmatrix}\begin{bmatrix} 6 \\ 6+2\epsilon \end{bmatrix}$$

$$= \frac{1}{2\epsilon^2}\begin{bmatrix} 18+12\epsilon+6\epsilon^2-18-6\epsilon-6\epsilon-2\epsilon^2 \\ -18-6\epsilon+18+6\epsilon \end{bmatrix}$$

$$= \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

$$= \frac{1 - e^{j\frac{2\pi}{N}(k+\ell)N}}{1 - e^{j\frac{2\pi}{N}(k+\ell)}} = \frac{1 - e^{j2\pi(k+\ell)}}{1 - e^{j\frac{2\pi}{N}(k+\ell)}} = 0$$

and similarly for $\sum_n e^{j(\omega_k - \omega_\ell)}$.

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

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

$$= \frac{\mathrm{var}(\hat{a}_k) + E(\hat{a}_k)^2 + \mathrm{var}(\hat{b}_k) + E(\hat{b}_k)^2}{2}$$

$$= \frac{a_k^2 + b_k^2 + 4\sigma^2/N}{2} = P + \frac{2\sigma^2}{N}$$

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

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

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

If $z \sim N(\mu, \sigma^2)$

$$\mathrm{var}(z^2) = 4\mu^2\sigma^2 + 2\sigma^4$$

(see development preceding (3.19))

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

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

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

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

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

For example, if $P >> 4\sigma^2/N$ or $NP >> 4\sigma^2$,

$$\frac{E(\hat{P})^2}{\mathrm{var}(\hat{P})} = \frac{P^2}{4P\sigma^2/N} = \frac{P}{\frac{4\sigma^2}{N}} >> 1$$

and the sinusoid will be easily detectable.

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

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

$$[\underline{H}^T\underline{H}]_{ij} = \sum_{m=-\infty}^{\infty} u[m]u[m+i-j]$$

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

Similarly for $i < j$

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

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

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

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

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

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

$$\frac{\partial \ln p}{\partial \underline{\theta}} = -\frac{1}{2}\frac{\partial}{\partial \underline{\theta}}\left[\underline{x}^T\underline{C}^{-1}\underline{x} - 2\underline{\theta}^T\underline{H}^T\underline{C}^{-1}\underline{x} + \underline{\theta}^T\underline{H}^T\underline{C}^{-1}\underline{H}\underline{\theta}\right]$$

$$= -\frac{1}{2}\left[-2\underline{H}^T\underline{C}^{-1}\underline{x} + 2\underline{H}^T\underline{C}^{-1}\underline{H}\underline{\theta}\right) \text{ using } (4.3)$$

$$= \underline{H}^T\underline{C}^{-1}\underline{x} - \underline{H}^T\underline{C}^{-1}\underline{H}\underline{\theta}$$

$$= \underbrace{\underline{H}^T\underline{C}^{-1}\underline{H}}_{\underline{I}(\underline{\theta})}\left(\underbrace{(\underline{H}^T\underline{C}^{-1}\underline{H})^{-1}\underline{H}^T\underline{C}^{-1}\underline{x}}_{\hat{\underline{\theta}}} - \underline{\theta}\right)$$

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

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

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

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

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

$$\Rightarrow \quad \underline{D} = \mathrm{diag}(1/\sigma_0, 1/\sigma_1, \ldots, 1/\sigma_{N-1})$$

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

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

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 the 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 $\sigma_m^2 \to 0$, $d_n \to 0$ for $n \neq m$ and $d_n \to \sigma_m$ for $n = m$. Thus, $\hat{A} \to d_m x'[m] = x[m]$, as expected.

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

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

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

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

$$\underline{H}^T \underline{C}^{-1} \underline{x} = \frac{\underline{1}^T}{\sigma^2(1-\rho^2)} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \end{bmatrix}$$

$$= \frac{x[0](1-\rho) + x[1](1-\rho)}{\sigma^2(1-\rho^2)}$$

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

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

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

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

$$(\underline{H}^T\underline{C}^{-1}\underline{H})^{-1}\underline{H}^T\underline{C}^{-1} = (\underline{H}^T\tfrac{1}{\lambda}\underline{H})^{-1}\tfrac{1}{\lambda}\underline{H}^T$$

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

As $\rho \to 1$, $\mathrm{var}(\hat{A}) \to \sigma^2$
As $\rho \to -1$, $\mathrm{var}(\hat{A}) \to 0$.
See Prob 3.9 for explanation.

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

$$\underset{r\times 1}{\underline{x}'} = \underline{A}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{H}\underline{\theta} + \underline{A}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{w}$$

$$= \underline{A}\underline{\theta} + \underline{w}' = \underline{H}'\underset{r\times 1}{\underline{\alpha}} + \underline{w}' \qquad \text{where } \underline{H}' = \underline{I}$$

$$\underline{C}' = E(\underline{w}'\underline{w}'^T) = E(\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{H}^T\underline{w}\underline{w}^T\underline{H}(\underline{H}^T\underline{H})^{-1}\underline{A}^T)$$

$$= \sigma^2\underline{A}(\underline{H}^T\underline{H})^{-1}\underline{A}^T$$

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

$$\Rightarrow \quad \hat{\alpha} = (H'^T C'^{-1} H')^{-1} H'^T C'^{-1} x'$$

$$= C' C'^{-1} x' = A (H^T H)^{-1} H^T x = A \hat{\theta}$$

13) $$E(\hat{\theta}) = E\left[(H^T H)^{-1} H^T (H\theta + w) \right]$$
$$= \theta + E\left[(H^T H)^{-1} H^T w \right]$$
$$= \theta + E\left[(H^T H)^{-1} H^T \right] E(w)$$
$$= \theta \quad \text{since } E(w) = 0.$$

$$C_{\hat{\theta}} = E\left[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T \right]$$
$$= E\left[\left((H^T H)^{-1} H^T (x - H\theta) \right) \left((H^T H)^{-1} H^T (x - H\theta) \right)^T \right]$$
$$= E\left[(H^T H)^{-1} H^T w w^T H (H^T H)^{-1} \right]$$
$$= E_{H|w} E_w \left[(H^T H)^{-1} H^T w w^T H (H^T H)^{-1} \right]$$

$$= E_{H|w} \left[(H^T H)^{-1} H^T \sigma^2 I H (H^T H)^{-1} \right]$$
$$= E_{H|w} \left[\sigma^2 (H^T H)^{-1} \right] = \sigma^2 E_H \left[(H^T H)^{-1} \right]$$

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

14) $H = 1$ with probability $1-\epsilon$

$$H = [\underbrace{1 1 \dots 1}_{M} \underbrace{0 0 \dots 0}_{N-M}]^T \text{ with probability } \epsilon$$

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

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

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

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

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

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

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

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

Chapter 5

1) $$p(\underline{x} \mid T(\underline{x}) = T_0; \sigma^2) = \frac{p(\underline{x}; \sigma^2)\,\delta(T(\underline{x}) - T_0)}{p(T(\underline{x}) = T_0; \sigma^2)}$$

$$= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} e^{-\frac{1}{2\sigma^2}\sum_n x^2[n]} \,\delta(T(\underline{x}) - T_0)}{p\left(\sum_n x^2[n] = T_0; \sigma^2\right)}$$

But $\sum_n x^2[n] = s\sigma^2$

$$p\left(\sum_n x^2[n]\right) = \frac{1}{\sigma^2} p_s\left(\sum_n x^2[n] / \sigma^2\right)$$

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

$$= \frac{\frac{1}{\pi^{N/2}}\,\delta(T(\underline{x}) - T_0)}{\frac{1}{\Gamma(N/2)} T_0^{\frac{N}{2}-1}}$$

2) $$p(\underline{x}; \sigma^2) = \prod_{n=0}^{N-1} \frac{x[n]}{\sigma^2} e^{-\frac{1}{2}x^2[n]/\sigma^2} \quad \text{all } x[n] > 0$$

$$= \underbrace{u(\min x[n]) \prod_{n=0}^{N-1} x[n]}_{h(\underline{x})} \; \underbrace{\frac{1}{\sigma^{2N}} e^{-\frac{1}{2}\frac{\sum_n x^2[n]}{\sigma^2}}}_{g(T(\underline{x}), \sigma^2)}$$

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

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

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

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

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

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

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

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

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

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

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

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

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

or $\max |x[n]| \le \theta$ 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.

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

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

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

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

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) $P_r\{x[n]\} = \theta^{x[n]}(1-\theta)^{1-x[n]} \qquad x[n] = 0, 1$

$$P_r\{\underline{x}\} = \prod_{n=0}^{N-1} \theta^{x[n]}(1-\theta)^{1-x[n]}$$

$$= \underbrace{\theta^{\sum_n x[n]}(1-\theta)^{N-\sum_n x[n]}}_{g(T(\underline{x}),\theta)} \cdot \underbrace{1}_{h(\underline{x})}$$

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

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

10) At high SNR we ignore the noise so that

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

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

$\approx NA/2 \cos\phi$ using results of Prob 3.7

Also,

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

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

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

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

$= \phi$

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

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

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

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

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

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

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

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

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

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

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 $\sum_n x(n)$. To make it unbiased we need a function g such that

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

or $E(h(\bar{x})) = A^3$ 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 MVU estimator.

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

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

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

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

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

$$Pr\{T \le \zeta\} = 1 - Pr\{T \ge \zeta\}$$
$$= 1 - Pr\{x[0] \ge \zeta, \ldots, x[N-1] \ge \zeta\}$$
$$= 1 - \prod_{n=0}^{N-1} Pr\{x[n] \ge \zeta\}$$
$$= 1 - Pr^N\{x[n] \ge \zeta\}$$

$$p_T(\zeta) = \frac{d\, Pr\{T \le \zeta\}}{d\zeta} = -N\, Pr\{x[n] \ge \zeta\}^{N-1} \cdot \frac{d Pr\{x[n] \ge \zeta\}}{d\zeta}$$

But $\dfrac{d\, Pr\{x[n] \ge \zeta\}}{d\zeta} = \dfrac{d\,[1 - Pr\{x[n] \le \zeta\}]}{d\zeta}$

$$= -\frac{d\, Pr\{x[n] \le \zeta\}}{d\zeta} = -\begin{cases} e^{-(\zeta-\theta)} & \zeta > \theta \\ 0 & \zeta < \theta \end{cases}$$

and

$$Pr\{x[n] \ge \zeta\} = \begin{cases} 1 - \int_\theta^\zeta e^{-(x-\theta)}\,dx & \text{for } \zeta > \theta \\ 1 & \text{for } \zeta < \theta \end{cases}$$

For $\zeta > \theta$ $\quad = 1 + e^{\theta} e^{-x}\Big|_\theta^\zeta = 1 + e^{-(\zeta-\theta)} - 1$

$$= e^{-(\zeta-\theta)}$$

$$p_T(\mathfrak{z}) = \begin{cases} 0 & \mathfrak{z} < \theta \\ -N\left[e^{-(\mathfrak{z}-\theta)}\right]^{N-1}\left[-e^{-(\mathfrak{z}-\theta)}\right] & \mathfrak{z} > \theta \end{cases}$$

$$= \begin{cases} N e^{-N(\mathfrak{z}-\theta)} & \mathfrak{z} > \theta \\ 0 & \text{otherwise} \end{cases}$$

$$E(T) = \int_\theta^\infty \mathfrak{z}\, N e^{-N(\mathfrak{z}-\theta)} d\mathfrak{z}$$
$$= N e^{N\theta} \int_\theta^\infty \mathfrak{z}\, e^{-N\mathfrak{z}} d\mathfrak{z}$$
$$= N e^{N\theta}\left(-\frac{\mathfrak{z}}{N} e^{-N\mathfrak{z}} - \frac{1}{N^2} e^{-N\mathfrak{z}}\right)\Big|_\theta^\infty$$
$$= N e^{N\theta}\left(\frac{\theta}{N} e^{-N\theta} + \frac{1}{N^2} e^{-N\theta}\right) = \theta + \frac{1}{N}$$

$\Rightarrow \quad \hat{\theta} = T - 1/N = \min x[n] - \frac{1}{N}$ is the MVU estimator.

14) a) $$p(x;\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2 + x\mu - \frac{1}{2}\mu^2}$$

$$= e^{\underbrace{x\mu}_{A(\mu)B(x)} \underbrace{- \frac{1}{2}x^2}_{C(x)} + \underbrace{(-\frac{1}{2}\mu^2 + \ln 1/\sqrt{2\pi})}_{D(\mu)}}$$

b) $$p(x;\sigma^2) = \frac{x}{\sigma^2} e^{-\frac{1}{2}x^2/\sigma^2} u(x)$$

$$= e^{\ln x/\sigma^2 - \frac{1}{2}x^2/\sigma^2 + \ln u(x)}$$

$$= e^{\underbrace{-\frac{1}{2}x^2/\sigma^2}_{A(\sigma^2)B(x)} + \underbrace{\ln x u(x)}_{C(x)} \underbrace{- \ln\sigma^2}_{D(\sigma^2)}}$$

c) $p(x;\lambda) = \lambda e^{-\lambda x} u(x)$

$$= e^{\underbrace{-\lambda x}_{A(\lambda)B(x)} + \underbrace{\ln u(x)}_{C(x)} + \underbrace{\ln \lambda}_{D(\lambda)}}$$

15) $p(\underline{x};\theta) = \prod_{n=0}^{N-1} e^{A(\theta)B(x[n]) + C(x[n]) + D(\theta)}$

$$= e^{A(\theta)\sum_n B(x[n]) + \sum_n C(x[n]) + N D(\theta)}$$

$$= \underbrace{e^{A(\theta)\sum_n B(x[n]) + N D(\theta)}}_{g(T(\underline{x}),\theta)} \underbrace{e^{\sum_n C(x[n])}}_{h(\underline{x})}$$

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

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

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

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

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

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

b) $E(x^2) = \int_0^\infty \frac{x^3}{\sigma^2} e^{-\frac{1}{2}x^2/\sigma^2} dx$

$$= 2\sigma^2$$

$$\Rightarrow \quad \hat{\sigma}^2 = \frac{1}{2N} \sum_n x^2(n)$$

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.

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

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

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

$$\underline{C}^{-1} = \begin{bmatrix} (a - \underline{b}^T\underline{b})^{-1} & -(a-\underline{b}^T\underline{b})^{-1}\underline{b}^T \\ -\underline{b}(a - \underline{b}^T\underline{b})^{-1} & \left(\underline{I} - \frac{\underline{b}\,\underline{b}^T}{a}\right)^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{\sigma^2} & \frac{1}{\sigma^2}\underline{1}^T \\ \frac{1}{\sigma^2}\underline{1} & \left(\underline{I} - \frac{\underline{1}\,\underline{1}^T}{N-1+\sigma^2}\right)^{-1} \end{bmatrix}$$

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

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

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

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

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

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

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

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

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

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

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