TTT4275 EDC
Suggested Solutions for Problem Set Estimation 1

Problem 1

(a) The object is an unknown distance D away from the observer. Thus the pulse round-trip time is $T = 2D/C \;\Rightarrow\; D = CT/2$
The models for running time t and distance d are respectively
$t = T + w$ and $d = D + Cw/2$ where $p(w) = N(0, \sigma_w^2)$

Thus the distribution of d is given by $p(d) = N(D, (C\sigma_w)^2/4)$ i.e.
$\sigma_d = C\sigma_w/2$

A single measurement means that our estimator is $\hat{D} = d$. We now want minimum 99% of the area under $p(\hat{D}) = p(d)$ to be max 1 meter away from D. Using tables for the normalized Gaussian $p(x) = N(0,1)$ the 99% span is $x = \pm 2.5758$. The corresponding value for d is found by multiplying with $\sigma_d = C\sigma_w/2$. Thus we have $2.5758 C\sigma_w/2 = 1$ meter; which gives $\sigma_w = 2.59 * 10^{-9} = 2.59$ nanoseconds

Using N observations means that we use the sample mean estimator $\hat{D} = \sum_n d(n)/N$ which has a distribution $p(\hat{D}) = N(D, \sigma_d^2/N)$. Thus we have to scale the standard deviation with $1/\sqrt{N}$ and use the new given noise standard deviation $\sigma_w = 10^{-8} = 10$ nanoseconds. This results in $2.5758 C\sigma_w/(2\sqrt{N}) = 1$ Rearranging and squaring we get $N = 14.92$, i.e. minimum $N = 15$ observations.

(b) Given $p(x) = N(0, \sigma^2)$, N independent observations $x = [x(0), x(1), ..., x(N-1)]$ and the estimator $\hat{\sigma^2} = \hat{v} = \sum x^2(n)/N$ where we for convenience use the notation $v = \sigma^2$.

- Checking for bias :
Now $E\{\hat{v}\} = E\{\sum x^2(n)/N\} = \sum E\{x^2(n)\}/N$. However $E\{x^2(n)\} = var(x) = v = \sigma^2$ as $E\{x(n)\} = 0$. Thus $E\{\hat{v}\} = \sum v/N = v = \sigma^2 \;\;\Rightarrow\;$ the estimator is unbiased.

- Checking variance :
$var(\hat{v}) = E\{[\hat{v} - v]^2\}$ Since the estimator is unbiased the two cross terms in the quadratic expression are zero and we end up with $var(\hat{v}) = E\{[\hat{v}]^2\} - v^2$
The first term written as a function of $x(n)$ becomes :
$var(\hat{v}) + v^2 = E\{(\sum_n x^2(n)/N)(\sum_m x^2(m)\}/N)\} =$
$E\{\sum_n \sum_m x^2(n)x^2(m)\}/N^2\} = \sum_n \sum_m E\{[x(n)x(m)]^2\}/N^2\}$.

  We now has two cases for m :
a) $m = n$ where $E\{[x(n)x(n)]^2\} = E\{x^4(n)\} = 3v^2/4$ (see hint in task). This happens N times in the double sum.
b) $m \neq n$ which leads to
$E\{[x(n)x(m)]^2\} = E\{x^2(n)\}E\{x(m)^2\} = v * v = v^2$. This happens $N^2 - N$ times
Thus we end up with $var(\hat{v}) = (N3v^2 + (N^2 - N)v^2 - v^2)/N^2 =$
$3v^2/N + v^2 - v^2/N - v^2 = 2v^2/N$.

  Referring to example 7 in the compendium we see that the variance is equal to the CRLB. Thus we have an efficient MVU estimator!

(c) The estimator is given by $\hat{\theta} = [\sum x(n)/N]^2 = \hat{A}^2$.
We can write $\hat{A} = \sum(A + w(n))/N = A + \sum w(n)/N = A + q(n)$ where $p(q) = N(0, \sigma^2/N)$.
Now $E\{\hat{\theta}\} = E\{\hat{A}^2\} = E\{(A + q(n))^2\} = A^2 + 2AE\{q(n)\} + E\{q^2(n)\} =$
$A^2 + 2A * 0 + \sigma^2/N = A^2 + \sigma^2/N \;\Rightarrow\;$ Biased! However, as $N \rightarrow \infty$ the last term goes towards zero, i.e. aymptotically unbiased.

Problem 2

(a) Given N waiting times observations $\Delta = [\delta_0, \delta_1, .., \delta_{N-1}]$ where $\delta_i = t_i - t_{i-1}$ and t_i is packet time arrivals. The distribution in δ is assumed to be$p(\delta, \beta) = \frac{1}{\beta} \exp -(\delta/\beta)$ and β is the unknown parameter to estimate.

Thus for N observations we have
$p(\Delta; \beta) = \prod_i p(\delta_i, \beta) = \prod_i \frac{1}{\beta} exp(-\delta_i/\beta)$ and
$logp(\Delta; \beta) = \sum_i -log\beta - \delta_i/\beta$

The CRLB is found by taking the derivative twice and then take the expectation:
a) $d(logp)/d\beta = \sum_i(-1/\beta + \delta_i/\beta^2)$
b) $d^2(logp)/d2\beta = \sum_i(1/\beta^2 - 2\delta_i/\beta^3)$
c) $E\{d^2(logp)/d^2\beta\} = E\{\sum_i(1/\beta^2 - 2\delta_i/\beta^3)\}) = N/\beta^2 - 2NE\{\delta_i\}/\beta^3 = -N/\beta^2$ since $E\{\delta\} = \beta$
(see https://en.wikipedia.org/wiki/Exponential_distribution)
Thus $CRLB = \beta^2/N$

Using the sample mean estimator the best way of testing for MVU is to try to write $d(logp)/d\beta = I(\beta)(\hat{(\beta)} - \beta)$. Thus we know that the estimator is MVU and efficient.
We start from a) :
$d(logp)/d\beta = \sum_i(-1/\beta + \delta_i/\beta^2) = -\frac{N}{\beta} + \frac{1}{\beta^2}\sum_i \delta_i = \frac{N}{\beta^2}(\sum_i \delta_i/N - \beta)$.

Thus $I(\beta) = \frac{N}{\beta^2}$ and $\hat{\beta} = \sum_i \delta_i/N$ (sample mean) is an efficient estimator with variance $1/I(\beta) = \beta^2/N$!

Now let us assume a distribution very similar to the first one, but with a slightly different way of defining the unknown parameter :
$p(\delta, \lambda) = \lambda \exp(-\lambda\delta)$. Thus we have $\lambda = 1/\beta$

Doing the calculations :
a) $log(p(\Delta, \lambda)) = \prod_i log[\lambda \exp(-\lambda\delta_i) = Nlog\lambda - \sum_i \lambda\delta_i$
b) $d[log(p)]/d\lambda = \sum_i(1/\lambda - \delta_i)$

We see that we can not reformulate the derivative in b) into the necessary form $I(\lambda)[\hat{\lambda} - \lambda]$!! Hence not efficient.
taking the second derivative (derivate of b) we easily find that the CRLB is given by $var(\hat{\lambda}) \geq \lambda^2/N$. Thus for our estimator we will have an inequality.

(b) A communication problem is given by that the channel amplitudes /gain has a Rayleigh distribution $p(r, \sigma_2) = (r, v) = \frac{r}{v} \exp -(r^2/2v)$ where we use the notation $\sigma_2 = v$.

Taking the logarithm : $log(p) = \sum_n [log r(n) - log(v) - r^2/2v]$

If we can write $d(logp)/dv = I(v)(\hat{v} - v)$ we have an efficient MVU estimator and the variance is given by $1/I(v)$:

$d(logp)/dv = \sum_n [0 - 1/v + r(n)/2v^2] = \frac{1}{v^2}[(\sum_n r^2(n)/2N) - v]$

Thus $var(\hat{v}) = 1/I(v) = v^2/N = \sigma^4/N$ and $\hat{v} = \hat{\sigma 2} = \frac{1}{2N} \sum_n r^2(n)$

corresponds to an efficient MVU estimator

(c) A script for the computer assignment can be found in Blackboard under the subfolder Exercizes