FACULTY OF INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING
DEGREE PROGRAMME IN ELECTRONICS AND COMMUNICATIONS ENGINEERING

Statistical Signal Processing I
MATLAB Homework

Athmajan Vivekananthan	Hathurusingha Anjalee Anuththara De Silva
avivekan23@student.oulu.fi	hdesilva23@student.oulu.fi
2305573	2303245

1 TASK

The data $x[n] = \theta n + w[n]$ where $n = 0, 1, 2, \ldots, N-1$ are observed values, and $w[n]$ represents zero-mean white Gaussian noise with variance σ^2. The objective is to estimate the parameter θ. It is known that the Cramer-Rao lower bound is $CRLB = \frac{\sigma^2}{\sum_{n=0}^{N-1} n^2}$.

1.1 Question 1

Show that the estimator $\hat{\theta} = \frac{\sum_{n=0}^{N-1} x[n]n}{\sum_{n=0}^{N-1} n^2}$ is **efficient** by Monte Carlo (MC) simulation.

In order to show that $\hat{\theta}$ is an efficient estimator, first it needs to be shown to be **an unbiased** estimator and secondly that it **attains the Cramer-Rao Lower Bound**. 1.1 shows in from Line 31 to Line 36 that $\hat{\theta}$ is unbiased. Line 41 to Line 48 show the calculation of simulated CRLB value compared with the simulated variance of $\hat{\theta}$ showing that they are approximately equal for three significant bits. Thus, $\hat{\theta}$ being **an efficient estimator for** θ.

```
2   % Setting up parameters of Monte Carlo simulation
3   % ----------------------------------------
4   clc
5   clear all
6   clf
7
8   numMonteCLoops = 1E4; % Number of Monte Carlo Loops
9   N = 10; % Number of observations
10  noise_var = 2; % Variance of WGN
11  theta = 10; % True value of the parameter
12  % ----------------------------------------
13
14  % Initializing variable for stability
15  estimate = ones(1,numMonteCLoops)*NaN;
16
17  for MC=1:numMonteCLoops
18      %%% defining the Numerator in the estimator equation
19      % x[n]*n = theta*n*n + n*w[n]
20      noisePart = randn(1,N)*sqrt(noise_var).*(1:N); % the WGN part
21      numSignal = theta*(1:N).^2 + noisePart; % the signal parr
22
23      %% defining the denominator in the estimator equation
24      denomSignal = (1:N).^2;
25
26      %% defining the estimate as the divisions of summations of numerator
27      %% and denominator
28      estimate(MC) = sum(numSignal) / sum(denomSignal);
29  end
30
31  %% Checking whether the mean of the estimate is closely approximated to the
32  %% true value of the parameter
33  fprintf('Expected value of the estimate is: %s\n', num2str(mean(estimate)));
34  if round(mean(estimate))== theta
35      disp("The estimator is performing unbiased")
36  end
37
38  %% Calculating the CRLB for simulation
39  CRLB_theta = (noise_var) / (sum((1:N).^2));
40
41  %% Check whether CRLB simulated value and the variance simulated value
42  %% are approximately the same or not.
43  fprintf('Cramer-Rao Lower Bound is : %s\n', num2str(CRLB_theta));
44  fprintf('Variance of the estimate is: %s\n', num2str(var(estimate)));
45
46  if abs(round(var(estimate),3) - round(CRLB_theta,3))== 0
47      disp("The estimator is performing efficient")
48  end
```

Figure 1.1. Matlab code to generate MC simulation for estimation of θ

```
Expected value of the estimate is: 10.0003

The estimator is performing unbiased


Cramer-Rao Lower Bound is : 0.0051948
Variance of the estimate is: 0.0050057

The estimator is performing efficient
```

Figure 1.2. Matlab output for the MC simulations

1.2 Question 2

Generate figure that shows that simulated and theoretical probability density functions agree for the estimator in (1). Make sure that your figures includes labels and legends.

In order to find the theoretical PDF, the theoretical mean and theoretical variance must be derived.

$$\hat{\theta} = \frac{\sum_{n=0}^{N-1} x[n]n}{\sum_{n=0}^{N-1} n^2}. \tag{1.1}$$

$$\begin{aligned}
\mathbb{E}[\hat{\theta}] &= \frac{\mathbb{E}\left[\sum_{n=0}^{N-1} x[n]n\right]}{\mathbb{E}\left[\sum_{n=0}^{N-1} n^2\right]} \\
&= \frac{\sum_{n=0}^{N-1} \mathbb{E}\left[x[n]n\right]}{\sum_{n=0}^{N-1} \mathbb{E}\left[n^2\right]} \\
&= \frac{\sum_{n=0}^{N-1} \mathbb{E}\left[\theta n^2 + w[n]n\right]}{\sum_{n=0}^{N-1} \mathbb{E}\left[n^2\right]} \\
&= \frac{\sum_{n=0}^{N-1} \mathbb{E}\left[\theta n^2\right] + \mathbb{E}\left[w[n]n\right]}{\sum_{n=0}^{N-1} \mathbb{E}\left[n^2\right]}
\end{aligned}$$

Since θ is deterministic and $\mathbb{E}[w[n]] = 0$,

$$\begin{aligned}
&= \theta \times \frac{\sum_{n=0}^{N-1} \mathbb{E}\left[n^2\right]}{\sum_{n=0}^{N-1} \mathbb{E}\left[n^2\right]} \\
&= \theta
\end{aligned}$$

Since in Part 1.1 we have already proved that this estimate is efficient, the variance of $\hat{\theta}$ is the same as the CRLB.

Therefore it can be concluded that $\mu_{\hat{\theta}_theo} = \theta$ and $var_{\hat{\theta}_theo} = \frac{\sigma^2}{\sum_{n=0}^{N-1} n^2} = \frac{\sigma^2}{\frac{(N-1)\times N\times(2N-1)}{6}}$

Using the above results the theoretical PDF of the estimate $\hat{\theta}$ can be written as :

$$PDF_{the0} = \frac{1}{\sqrt{2\times\pi\times\mathrm{var}_{\hat{\theta}_{\mathrm{theo}}}}} \times \exp\left(-\frac{(x-\mu_{\hat{\theta}_{\mathrm{theo}}})^2}{2\times\mathrm{var}_{\hat{\theta}_{\mathrm{theo}}}}\right) \tag{1.2}$$

```
% Initiating figure
figure(1)
clf
% Drawing the Probability Density Function of the simulated estimates
H = histogram(estimate, 'Normalization', 'pdf');
hold on

% Taking a copy of x values for the theoritical PDF
x = linspace(H.BinLimits(1),H.BinLimits(2),100);

%% Calculation of theoritical mean and variance included in the report
theo_mean = theta;
CRLB = noise_var/ (N*(N+1)*((2*N)+1)/6);
theo_var = CRLB

%% Defining the Theoritical PDF using theoritical mean and variance
theo_pdf = 1 / (sqrt(2 * pi * theo_var)) * exp(-((x - theo_mean).^2) / (2 * theo_var));

%% Ploting the Theoritical PDF on top of the PDF of simulated values
plot(x, theo_pdf, 'r', 'LineWidth', 2);
legend('Simulated','Theoritical');
xlabel('x');
ylabel('pdf')
```

Figure 1.3. Matlab code to generate the theoretical and simulated PDFs

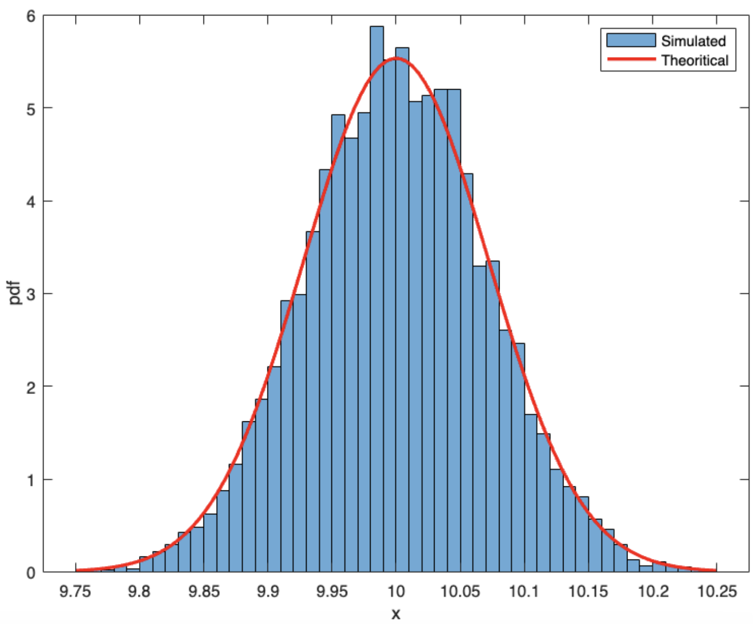

Figure 1.4. Theoretical and Simulated PDFs for N=10

1.3 Question 3

What happens when $N = 1$, and $N \rightarrow \infty$? Explain why.

When $N = 1$, the equation $x[n] = \theta n + w[n]$ becomes more dependent on the WGN because we only have one observation to support our estimator. Therefore the estimator shows the characteristics of the noise vector rather than the observations.

Since the variance of the estimate is inversely proportional to N, when N grows the variance of the estimate becomes smaller and smaller. $var_{\hat{\theta}_theo} = \frac{\sigma^2}{\frac{(N-1)\times N\times(2N-1)}{6}}$

Therefore it is evident that the samples are narrowing down towards the mean value when more and more observations are made to the estimate. This also can be seen by checking the results starting from N=1 and increase the value of N and plotting the resulting PDFs.

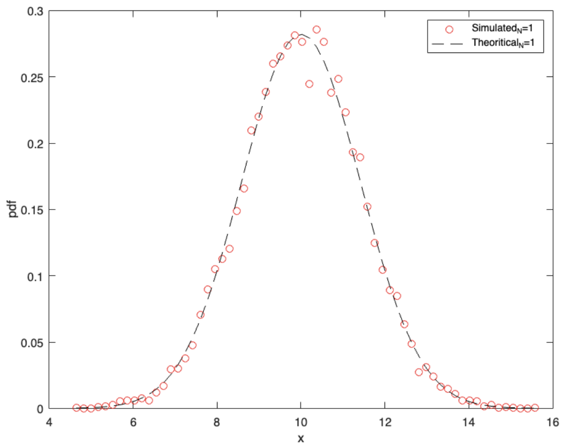

Figure 1.5. Simulated and Theoretical PDFs for N=1

N = 2

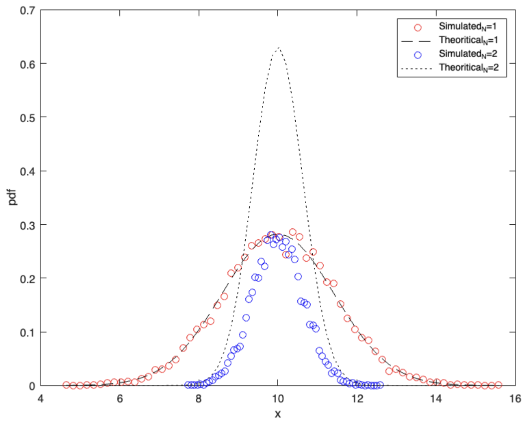

Figure 1.6. Simulated and Theoretical PDFs for N=2

N = 5

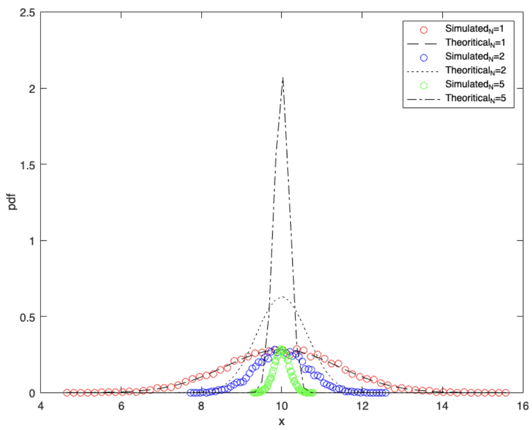

Figure 1.7. Simulated and Theoretical PDFs for N=5

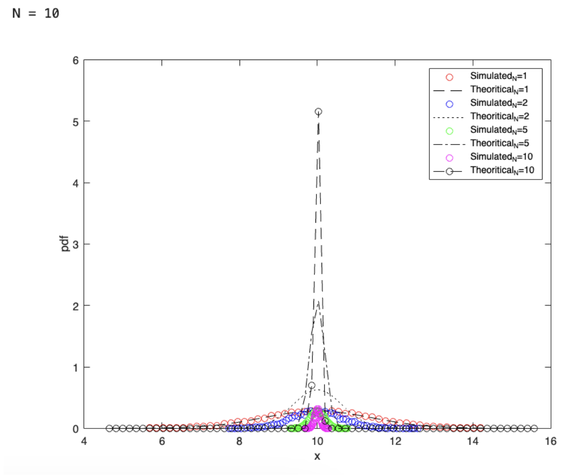

Figure 1.8. Simulated and Theoretical PDFs for N=10