

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

Department of Chemical Engineering

CH5115: Parameter and State Estimation**Assignment 1****Due:** Thursday, November 12, 2020

1. Suppose a random sample of N observations drawn from the following PDF are given:

$$f(y) = \begin{cases} e^{-(y-\theta)}, & x > \theta, \quad -\infty < \theta < \infty \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

- (a) Consider the statistic $T_N = 2 \min(\mathbf{y}_N)$. Derive the PDF of T_N .
- (b) If T_N is used as an estimator of θ , verify theoretically if it is unbiased. If T_N is found to be biased, apply a correction factor to make it unbiased. Denote the corrected statistic as T'_N .
- (c) Prove that T'_N converges to θ in probability using your answer in (1a).
2. Given an N -length realization of a random process, $v[k]$, compute the DFT coefficients and an estimator of the PSD as:

$$V[n] \triangleq V(f_n) = \sum_{k=0}^{N-1} v[k] \exp(-j2\pi f_n k), \quad f_n = \frac{n}{N}, \quad n = 0, 1, \dots, N-1 \quad (2)$$

$$\mathbb{P}[n] \triangleq \mathbb{P}(f_n) = \frac{|V(f_n)|^2}{N} = \frac{a_n^2 + b_n^2}{N} \quad a = \Re(V[n]), \quad b = \Im(V[n]) \quad (3)$$

For analysis, take up the zero-mean, unit variance GWN process. Recall that the true PSD is

$$\gamma_{vv}(f) = \sum_{l=-\infty}^{\infty} \sigma_{vv}[l] \exp(-j2\pi fl) \quad (4)$$

where $\sigma_{vv}[l]$ is the ACVF.

For MC simulations, wherever required, use $N = 1000$ and $R = 300$ (no. of realizations). Choose a frequency f_n corresponding to $n =$ last two digits of your Roll No. + 1.

- (a) Show (theoretically) that the DFT coefficients, a_n and b_n are Gaussian distributed. Verify your answer through MC simulations.
- (b) Determine the correlation between a_n and b_n .
- (c) Next, determine the distribution of $\zeta_n = \frac{2\mathbb{P}_{vv}(f_n)}{\gamma_{vv}(f_n)}$. From this distribution, determine the mean and variance of ζ_n .
- (d) Finally, verify if $\mathbb{P}(f_n)$ is a consistent estimator of $\gamma_{vv}(f_n)$ in the mean-square sense both theoretically and by MC simulations.

3. (a) In the frequency detection example discussed in the lecture, assume both amplitude and frequency are unknown. Given N observations, determine the CRLB for amplitude and frequency estimates. Compare the variance bounds with those obtained when each of the parameters is estimated assuming the other is known.
- (b) Design a BLUE that estimates variance of a zero-mean GWN process $y[k] = e[k] \sim \mathcal{N}(0, \sigma^2)$ from N observations of $y[k]$. If such an estimator does not exist, design a BLUE for σ^2 using transformed data.
4. (a) A random collection of 100 samples of a polymer manufactured by a leading chemical industry revealed a mean molecular weight of 14578 with a standard deviation of 1845. With what degree of confidence can we assert that the average molecular weight of that polymer is in between 12000 and 16000?
- (b) Suppose that we want to investigate whether on the average, students from an elite institution A perform better in an exam than students from another elite institution B. If 60 students of A demonstrated on the average $\bar{x}_1 = 85.2$ marks (out of 100) with $s_1 = 6.8$ marks, while 55 students from institute B produced an average $\bar{x}_1 = 87.2$ marks (out of 100) with $s_1 = 8.8$ marks, what can we conclude?
5. (a) Prove that for Q.1, T_N is a sufficient statistic for θ . Design the MVUE for θ .
- (b) Compare the efficiency of estimator obtained in (5a) with that of $\hat{\theta}_2 = \theta - 1$ and $\hat{\theta}_2 = 1/\hat{\sigma}$, where $\hat{\sigma}$ is the sample standard deviation using MC simulations.