

EE6617: Tutorial 5

- 5.1** Normally, to estimate mean value of the observation $x[n]$, you would automatically think $\hat{\theta} = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$ if the observations are i.i.d. In this problem, other powerful mean estimator will be explored.

Consider i.i.d. uniform noise

$$x[n] = w[n], \quad n = 1, \dots, N-1, \quad w[n] \sim \mathcal{U}[0, \beta],$$

where $\beta > 0$ is unknown. We will estimate the mean of $w[n]$, i.e., $\theta = \frac{\beta}{2}$ by using Neyman-Fisher factorization and RBLS theorem.

- (a) The PDF of $x[n]$ is $p(x[n]; \frac{\beta}{2}) = \begin{cases} \frac{1}{\beta} & 0 \leq x[n] \leq \beta \\ 0 & \text{otherwise} \end{cases}$. What is the PDF of $\mathbf{x} \in \mathbb{R}^{N \times 1}$, i.e., $p(\mathbf{x}; \frac{\beta}{2})$?
- (b) In order to apply Neyman-Fisher factorization to $p(\mathbf{x}; \frac{\beta}{2})$ to find a sufficient statistic $T(\mathbf{x})$, we need to have a better way of expressing $p(\mathbf{x}; \frac{\beta}{2})$. Use the unit step function $u(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$, to express $p(\mathbf{x}; \frac{\beta}{2}) = g(T(\mathbf{x}); \frac{\beta}{2})h(\mathbf{x})$.
- (c) Apply Neyman-Fisher factorization and find a sufficient statistic $T(\mathbf{x})$.
- (d) Now applying RBLS theorem to design an MVU estimator is of interest. What is the density of $p(T(\mathbf{x}); \beta/2)$? You may try to find it by using $p(T(\mathbf{x}) = T_0; \beta/2) = \frac{\partial}{\partial T_0} \Pr(T(\mathbf{x}) < T_0; \frac{\beta}{2})$.
- (e) Assume $T(\mathbf{x})$ is complete statistic. Consider a linear function of $T(\mathbf{x})$ such that $g(T(\mathbf{x})) = aT(\mathbf{x})$. Design a so that $g(T(\mathbf{x}))$ is unbiased. Is the designed $g(T(\mathbf{x})) = aT(\mathbf{x})$ the MVU estimator?
- (f) Let $\hat{\theta} = aT(\mathbf{x})$. Compute $\text{var}(\hat{\theta})$ and compare the variance performance with the sample mean.

- 5.2** The i.i.d. observations $x[n]$ for $n = 0, 1, \dots, N-1$ have the exponential PDF

$$p(x[n]; \lambda) = \begin{cases} \lambda \exp(-\lambda x[n]) & x[n] > 0 \\ 0 & x[n] \leq 0 \end{cases}. \quad (1)$$

- (a) Find a sufficient statistics for λ .
- (b) Apply RBLS to design an MVU estimator.

- 5.3** Consider the observation model $\mathbf{x} = \mathbf{1}A + \mathbf{w} = \mathbf{H}A + \mathbf{w}$ where the noise samples are zero mean, but are correlated with the covariance matrix

$$\mathbf{C} = \sigma^2 \begin{bmatrix} \mathbf{C}_B & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_B & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_B \end{bmatrix}$$

$$\mathbf{C}_B = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

The number of samples N is assumed to be even and $0 \leq \rho < 1$. The $\mathbf{C}$ is a block-diagonal matrix and so it is easily invertible. Find the BLUE and its variance. Interpret your results.

Sol. 5.1 (a) If the PDF of $x[n]$ is $p(x[n])$ for i.i.d. $x[n]$ $n = 0, \dots, N-1$, the multivariate PDF of $\mathbf{x}$ is $p^N(x[n])$. Thus,

$$\begin{aligned} p\left(\mathbf{x}; \frac{\beta}{2}\right) &= \begin{cases} \frac{1}{\beta^N} & 0 \leq x[0], x[1], \dots, x[N-1] \leq \beta \\ 0 & \text{otherwise} \end{cases} \\ &= \begin{cases} \frac{1}{\beta^N} & \min_n x[n] \geq 0, \max_n x[n] \leq \beta \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (2)$$

(b) From (2), we have

$$p\left(\mathbf{x}; \frac{\beta}{2}\right) = \underbrace{\frac{1}{\beta^N} u\left(\beta - \max_n x[n]\right)}_{g(T(\mathbf{x}); \beta/2)} \underbrace{u\left(\min_n x[n]\right)}_{h(\mathbf{x})}. \quad (3)$$

(c) From (3), a sufficient statistic is $T(\mathbf{x}) = \max_n x[n]$. This is actually the max order statistic.

(d) Notice that given $T(\mathbf{x}) = \max_n x[n]$, $\Pr(T(\mathbf{x}) < T_0; \frac{\beta}{2}) = \prod_{n=0}^{N-1} \Pr(x[n] < T_0; \frac{\beta}{2})$.

$$\begin{aligned} p_T(T_0; \frac{\beta}{2}) &= \frac{\partial}{\partial T_0} \Pr\left(T(\mathbf{x}) < T_0; \frac{\beta}{2}\right) \\ &= \frac{\partial}{\partial T_0} \prod_{n=0}^{N-1} \Pr\left(x[n] < T_0; \frac{\beta}{2}\right) \\ &= \frac{\partial}{\partial T_0} \left[\Pr\left(x[0] < T_0; \frac{\beta}{2}\right) \right]^N \\ &= N \left[\Pr\left(x[0] < T_0; \frac{\beta}{2}\right) \right]^{N-1} p_{x[0]}\left(T_0; \frac{\beta}{2}\right) \end{aligned} \quad (4)$$

where

$$\Pr\left(x[0] < T_0; \frac{\beta}{2}\right) = \begin{cases} 1 & T_0 \geq \beta \\ \frac{T_0}{\beta} & 0 \leq T_0 \leq \beta \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$p_{x[0]}\left(T_0; \frac{\beta}{2}\right) = \begin{cases} \frac{1}{\beta} & 0 \leq x[0] \leq \beta \\ 0 & \text{otherwise} \end{cases}. \quad (6)$$

Thus, plugging (5) and (6) in (4) yields

$$p_T\left(T_0; \frac{\beta}{2}\right) = \begin{cases} N \left(\frac{T_0}{\beta}\right)^{N-1} \frac{1}{\beta} & 0 \leq x[0] \leq \beta \\ 0 & \text{otherwise} \end{cases}.$$

(e) The parameter a is the design parameter such that $E[aT(\mathbf{x})] = \beta/2$.

$$\begin{aligned}
E[aT(\mathbf{x})] &= \int_0^\beta aT N \left(\frac{T}{\beta}\right)^{N-1} \left(\frac{1}{\beta}\right) dT \\
&= \frac{aN}{\beta^N} \int_0^\beta T^N dT \\
&= \frac{aN}{\beta^N} \frac{1}{N+1} T^{N+1} \Big|_0^\beta \\
&= a \frac{2N}{N+1} \frac{\beta}{2}.
\end{aligned}$$

Thus, a must be $a = \frac{N+1}{2N}$. Now, $g(T(\mathbf{x})) = \frac{N+1}{2N} T(\mathbf{x}) = \frac{N+1}{2N} \max_n x[n]$ is an unbiased estimator. Since a is unique and make the estimator $g(T(\mathbf{x}))$ unbiased, it is the MVU estimator.

(f)

$$\begin{aligned}
\text{var}(\hat{\theta}) &= E \left[\left(\frac{N+1}{2N} T(\mathbf{x}) \right)^2 \right] - \left(\frac{\beta}{2} \right)^2 \\
&= \left(\frac{N+1}{2N} \right)^2 \int_0^\beta T^2 \left(N \left(\frac{T}{\beta} \right)^{N-1} \left(\frac{1}{\beta} \right) \right) dT - \frac{\beta^2}{4} \\
&= \frac{(N+1)^2}{4N\beta^N} \int_0^\beta T^{N+1} dT - \frac{\beta^2}{4} \\
&= \frac{\beta^2}{4} \left[\frac{(N+1)^2}{4N(N+2)} - 1 \right] \\
&= \frac{\beta^2}{4N(N+2)} \rightarrow O \left(\frac{1}{N_2} \right)
\end{aligned}$$

Notice the variance of the sample mean is $\text{var} \left(\frac{1}{N} \sum_{n=1}^{N-1} x[n] \right) = \frac{1}{N^2} N \text{var}(x[0]) = \frac{1}{N} \frac{\beta^2}{12} \rightarrow O \left(\frac{1}{N} \right)$. Thus, the MVU estimator $\frac{N+1}{2N} \max_n x[n]$ outperforms the sample mean. Somewhat contrary to intuition, the sample mean is not the MVU estimator of the mean of uniformly distributed noise.

Sol. 5.2 (a) We can simplify $p(x[n]; \lambda)$ in (1) to $p(x[n]; \lambda) = \lambda \exp(-x[n]\lambda)u(x[n])$. For the vector $\mathbf{x}$ case,

$$\begin{aligned}
p(\mathbf{x}; \lambda) &= \prod_{n=0}^{N-1} \lambda \exp(-x[n]\lambda)u(\min(x[n])) \\
&= \underbrace{\lambda^N \exp \left(-\lambda \sum_{n=0}^{N-1} x[n] \right)}_{g(\sum_{n=0}^{N-1} x[n]; \lambda)} \underbrace{u(\min(x[n]))}_{h(\mathbf{x})}
\end{aligned}$$

From Neyman-Fisher factorization, $\Rightarrow T(\mathbf{x}) = \sum_{n=0}^{N-1} x[n]$

- (b) The sufficient statistic of exponential family is complete. Try to design the estimator $\hat{\theta} = g(T(\mathbf{x}))$ such that $E[g(T(\mathbf{x}))] = \lambda$. Since $E[x[n]] = \lambda$, $E\left[\sum_{n=0}^{N-1} x[n]\right] = N\lambda$. Thus, we conclude $\hat{\theta} = g(T(\mathbf{x})) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]$ is the MVU estimator.

Sol. 5.3 The inverse of $\mathbf{C}_B$ is $\mathbf{C}_B^{-1} = \frac{1}{(1-\rho^2)} \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix}$ and thus,

$$\mathbf{C}^{-1} = \frac{1}{\sigma^2} \begin{bmatrix} \mathbf{C}_B^{-1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_B^{-1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_B^{-1} \end{bmatrix}.$$

The BLUE of A is $\hat{A} = \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}$ and

$$\begin{aligned} \mathbf{1}^T \mathbf{C}^{-1} \mathbf{x} &= \frac{1}{\sigma^2(1-\rho^2)} [1-\rho \ 1-\rho \ \cdots \ 1-\rho] \mathbf{x} \\ &= \frac{1}{\sigma^2(1-\rho^2)} \sum_{n=0}^{N-1} x[n] \\ \mathbf{1}^T \mathbf{C}^{-1} \mathbf{1} &= \frac{(1-\rho)N}{\sigma^2(1-\rho^2)} = \frac{N}{\sigma^2(1+\rho)} \\ \hat{A} &= \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \\ \text{var}(\hat{A}) &= \frac{\sigma^2(1+\rho)}{N}. \end{aligned}$$

Even if the noise is correlated, the sample mean is still the best estimator for unknown PDF of $\mathbf{x}$, This indicates that the form of estimator is not affected by noise symmetry. The variance of $\hat{A}$ converges $\text{var}(\hat{A}) \rightarrow 0$ as $N \rightarrow \infty$.