Minimum Variance Unbiased Estimator (MVUE)

Raj Thilak Rajan

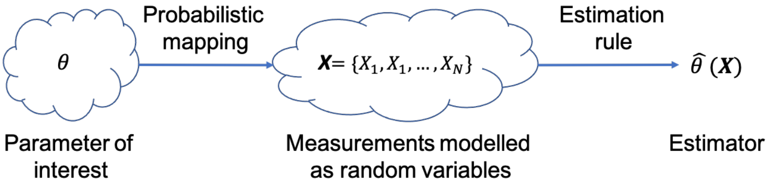
Summary

- Estimates and Estimators

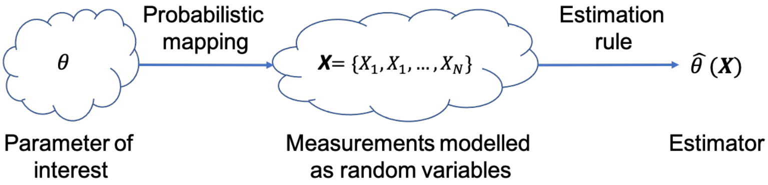
- Minimum variance

- MVUE

Philosophy

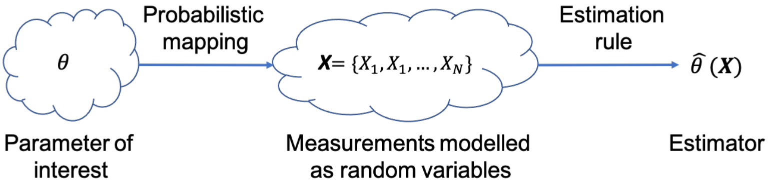

Probabilistic mapping
Estimation rule
θ
$\boldsymbol{X} = \{X_1, X_1, \ldots, X_N\}$
$\hat{\theta}\ (\boldsymbol{X})$
Parameter of interest
Measurements modelled as random variables
Estimator

Philosophy



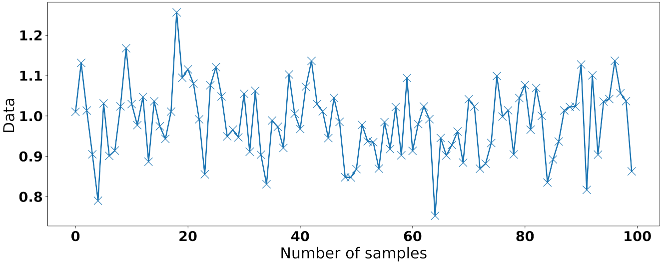
- Let $X = \{X_1, X_2, \ldots, X_N\}$ be a set of random samples drawn from probability distributions $p_{X_n}(x_n; \theta) \; \forall \; 1 \leq n \leq N$, where θ is the parameter of interest

Philosophy



- Let $X = \{X_1, X_2, \ldots, X_N\}$ be a set of random samples drawn from probability distributions $p_{X_n}(x_n; \theta)\ \forall\ 1 \leq n \leq N$, where θ is the parameter of interest
- We aim to
  (a) recover the unknown θ from the measurements X,
  (b) provide a performance measure of the estimated θ, and
  (c) discuss its' statistical *optimality*.

Example: Constant in Noise



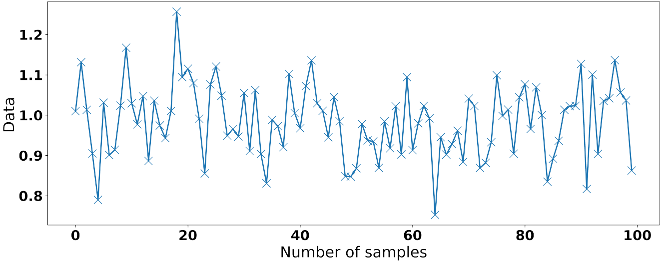
Consider the following measurement process

$$x[n] = \theta + w[n], \qquad n = 0, \ldots, N-1,$$

where, we assume

- θ is deterministic and *unknown*,
- $w[n]$ is a zero-mean IID Gaussian random process with variance σ^2,
- $x[n]$ is the measured data, which is an instance of a random variable.

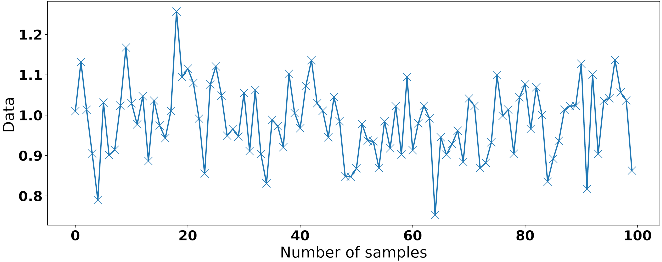
Example: Constant in Noise



Potential estimators for θ

- $\hat{\theta}_1 = x[0]$
- $\hat{\theta}_2 = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$
- $\hat{\theta}_3 = \frac{a}{N}\sum_{n=0}^{N-1} x[n]$, for some constant a
- $\ldots$

Example: Constant in Noise



Potential estimators for θ

- $\hat{\theta}_1 = x[0]$
- $\hat{\theta}_2 = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$
- $\hat{\theta}_3 = \frac{a}{N}\sum_{n=0}^{N-1} x[n]$, for some constant a
- $\ldots$

Which estimator is an *optimal* estimator ?

Unbiasedness

- An unbiased estimator "on the average" yields the true value, i.e.,

$$\mathbb{E}(\hat{\theta}) = \theta \quad \text{or} \quad bias(\theta) = \mathbb{E}(\hat{\theta}) - \theta = 0.$$

Unbiasedness

- An unbiased estimator "on the average" yields the true value, i.e.,

$$\mathbb{E}(\hat{\theta}) = \theta \quad \text{or} \quad bias(\theta) = \mathbb{E}(\hat{\theta}) - \theta = 0.$$

- For the aforementioned potential estimators of θ, we have
  - $\mathbb{E}(\hat{\theta}_1) = \mathbb{E}(x[0]) = \theta$
  - $\mathbb{E}(\hat{\theta}_2) = \mathbb{E}\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}(x[n]) = \theta$
  - $\mathbb{E}(\hat{\theta}_3) = \mathbb{E}\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a}{N}\sum_{n=0}^{N-1} \mathbb{E}(x[n]) = a\theta$

Unbiasedness

- An unbiased estimator "on the average" yields the true value, i.e.,

$$\mathbb{E}(\hat{\theta}) = \theta \quad \text{or} \quad bias(\theta) = \mathbb{E}(\hat{\theta}) - \theta = 0.$$

- For the aforementioned potential estimators of θ, we have
  - $\mathbb{E}(\hat{\theta}_1) = \mathbb{E}(x[0]) = \theta$
  - $\mathbb{E}(\hat{\theta}_2) = \mathbb{E}\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}(x[n]) = \theta$
  - $\mathbb{E}(\hat{\theta}_3) = \mathbb{E}\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a}{N}\sum_{n=0}^{N-1} \mathbb{E}(x[n]) = a\theta$
- Note: $\hat{\theta}_1, \hat{\theta}_2$ are unbiased estimators, $\hat{\theta}_3$ is a biased estimator.

Unbiasedness

- An unbiased estimator "on the average" yields the true value, i.e.,

$$\mathbb{E}(\hat{\theta}) = \theta \quad \text{or} \quad bias(\theta) = \mathbb{E}(\hat{\theta}) - \theta = 0.$$

- For the aforementioned potential estimators of θ, we have
  - $\mathbb{E}(\hat{\theta}_1) = \mathbb{E}(x[0]) = \theta$
  - $\mathbb{E}(\hat{\theta}_2) = \mathbb{E}\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}(x[n]) = \theta$
  - $\mathbb{E}(\hat{\theta}_3) = \mathbb{E}\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a}{N}\sum_{n=0}^{N-1}\mathbb{E}(x[n]) = a\theta$
- Note: $\hat{\theta}_1, \hat{\theta}_2$ are unbiased estimators, $\hat{\theta}_3$ is a biased estimator.
- Caution: An unbiased estimator does not mean an optimal estimator !

Variance

For the potential estimators of A, we have

- $var(\hat{\theta}_1) = \sigma^2$
- $var(\hat{\theta}_2) = var\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{\sigma^2}{N}$
- $var(\hat{\theta}_3) = var\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a^2}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{a^2\sigma^2}{N}$

Variance

For the potential estimators of A, we have

- $var(\hat{\theta}_1) = \sigma^2$
- $var(\hat{\theta}_2) = var\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{\sigma^2}{N}$
- $var(\hat{\theta}_3) = var\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a^2}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{a^2\sigma^2}{N}$

Note:

- As $N \rightarrow \infty$, $var(\hat{\theta}_2) \rightarrow 0$, and $var(\hat{\theta}_3) \rightarrow 0$

Variance

For the potential estimators of A, we have

- $var(\hat{\theta}_1) = \sigma^2$
- $var(\hat{\theta}_2) = var\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{\sigma^2}{N}$
- $var(\hat{\theta}_3) = var\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a^2}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{a^2\sigma^2}{N}$

Note:

- As $N \to \infty$, $var(\hat{\theta}_2) \to 0$, and $var(\hat{\theta}_3) \to 0$
- $\hat{\theta}_2$ is an unbiased estimator and $var(\hat{\theta}_2) < var(\hat{\theta}_1)$,

Variance

For the potential estimators of A, we have

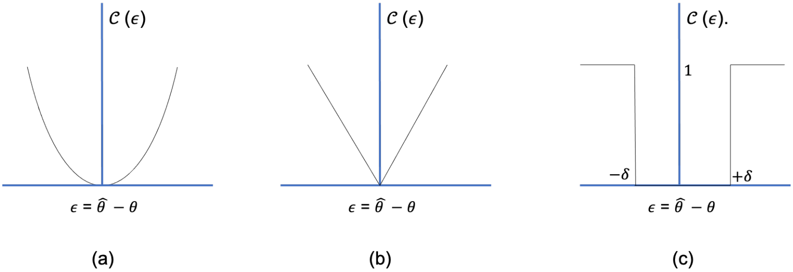
- $var(\hat{\theta}_1) = \sigma^2$
- $var(\hat{\theta}_2) = var\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{1}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{\sigma^2}{N}$
- $var(\hat{\theta}_3) = var\left(\frac{a}{N}\sum_{n=0}^{N-1} x[n]\right) = \frac{a^2}{N}\sum_{n=0}^{N-1} var(x[n]) = \frac{a^2\sigma^2}{N}$

Note:

- As $N \to \infty$, $var(\hat{\theta}_2) \to 0$, and $var(\hat{\theta}_3) \to 0$
- $\hat{\theta}_2$ is an unbiased estimator and $var(\hat{\theta}_2) < var(\hat{\theta}_1)$,
- $var(\hat{\theta}_3)$ is a function of constant a

Is $\hat{\theta}_2$ an optimal estimator ? What is the *error* on $\hat{\theta}_2$?

Cost functions



(a) $\mathcal{C}(\epsilon) = (\hat{\theta} - \theta)^2$

(b) $\mathcal{C}(\epsilon) = |\epsilon|$

(c) $\mathcal{C}(\epsilon) = 0$ if $|\epsilon| < \delta$ or $\mathcal{C}(\epsilon) = 1$

Optimality criterion

Mean square error (MSE)

$$
\begin{aligned}
mse(\hat{\theta}) &= \mathbb{E}\left[(\hat{\theta}-\theta)^2\right] = \mathbb{E}\left\{\left[(\hat{\theta}-\mathbb{E}(\hat{\theta})) + (\mathbb{E}(\hat{\theta})-\theta)\right]^2\right\} \\
&= \mathbb{E}\left[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2\right] + (\mathbb{E}(\hat{\theta})-\theta)^2 = \underbrace{var(\hat{\theta})}_{\text{variance}} + \underbrace{(\mathbb{E}(\hat{\theta})-\theta)}_{\text{bias}}{}^2,
\end{aligned}
$$

which consists of errors due to

- variance of the estimator
- bias of the estimator, which is a function of the unknown parameter.

Optimality criterion

Mean square error (MSE)

$$
\begin{aligned}
mse(\hat{\theta}) &= \mathbb{E}\left[(\hat{\theta}-\theta)^2\right] = \mathbb{E}\left\{\left[(\hat{\theta}-\mathbb{E}(\hat{\theta})) + (\mathbb{E}(\hat{\theta})-\theta)\right]^2\right\} \\
&= \mathbb{E}\left[(\hat{\theta}-\mathbb{E}(\hat{\theta}))^2\right] + (\mathbb{E}(\hat{\theta})-\theta)^2 = \underbrace{var(\hat{\theta})}_{\text{variance}} + \underbrace{(\mathbb{E}(\hat{\theta})-\theta)}_{\text{bias}}{}^2,
\end{aligned}
$$

which consists of errors due to

- variance of the estimator
- bias of the estimator, which is a function of the unknown parameter.

Note for the unbiased estimators $\hat{\theta}_1$, $\hat{\theta}_2$

- $mse(\hat{\theta}_1) = var(\hat{\theta}_1)$, $mse(\hat{\theta}_2) = var(\hat{\theta}_2)$

Example 1

- Consider the estimator $\hat{\theta}_3 = \frac{a}{N}\sum_{n=0}^{N-1} x[n]$ with

$$\mathbb{E}[\hat{\theta}_3] = a\theta, \qquad var[\hat{\theta}_3] = \frac{a^2\sigma^2}{N}$$

$$\mathrm{MSE}(\hat{\theta}_3) = \frac{a^2\sigma^2}{N} + (a-1)^2\theta^2$$

Example 1

- Consider the estimator $\hat{\theta}_3 = \frac{a}{N}\sum_{n=0}^{N-1} x[n]$ with

$$\begin{aligned} \mathbb{E}[\hat{\theta}_3] &= a\theta, \qquad var[\hat{\theta}_3] = \frac{a^2\sigma^2}{N} \\ \mathrm{MSE}(\hat{\theta}_3) &= \frac{a^2\sigma^2}{N} + (a-1)^2\theta^2 \end{aligned}$$

- Solve for $d\, mse(\hat{\theta}_3)/da$ and setting to zero yields,

$$a_{opt} = \frac{\theta^2}{\theta^2 + \sigma^2/N},$$

Example 1

- Consider the estimator $\hat{\theta}_3 = \frac{a}{N}\sum_{n=0}^{N-1} x[n]$ with

$$\begin{aligned}\mathbb{E}[\hat{\theta}_3] &= a\theta, \qquad var[\hat{\theta}_3] = \frac{a^2\sigma^2}{N} \\ \text{MSE}(\hat{\theta}_3) &= \frac{a^2\sigma^2}{N} + (a-1)^2\theta^2\end{aligned}$$

- Solve for $d\ mse(\hat{\theta}_3)/da$ and setting to zero yields,

$$a_{opt} = \frac{\theta^2}{\theta^2 + \sigma^2/N},$$

and subsequently, the optimal estimator is

$$\hat{\theta}_3 = \frac{\theta^2}{N\theta^2 + \sigma^2}\sum_{n=0}^{N-1} x[n]$$

which depends on the unknown parameter and thus not *realizable*.

Minimum Variance Unbiased Estimator (MVU)

- In practice, we cannot always compute the MSE estimators.

Minimum Variance Unbiased Estimator (MVU)

- In practice, we cannot always compute the MSE estimators.
- Solution: Constrain the bias of the MSE to zero, then

$$mse(\hat{\theta}) = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] + (\mathbb{E}(\hat{\theta}) - \theta)^2 = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] = var(\hat{\theta})$$

where $\hat{\theta}$ is an unbiased estimator.

Minimum Variance Unbiased Estimator (MVU)

- In practice, we cannot always compute the MSE estimators.
- Solution: Constrain the bias of the MSE to zero, then

$$mse(\hat{\theta}) = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] + (\mathbb{E}(\hat{\theta}) - \theta)^2 = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] = var(\hat{\theta})$$

  where $\hat{\theta}$ is an unbiased estimator.
- For any other unbiased estimator $\tilde{\theta}$, if

$$var(\hat{\theta}) \leq var(\tilde{\theta})$$

  then $\hat{\theta}$ is the Minimum Variance Unbiased estimator (MVU) for all θ.

Minimum Variance Unbiased Estimator (MVU)

- In practice, we cannot always compute the MSE estimators.
- Solution: Constrain the bias of the MSE to zero, then

$$mse(\hat{\theta}) = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] + (\mathbb{E}(\hat{\theta}) - \theta)^2 = \mathbb{E}\left[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2\right] = var(\hat{\theta})$$

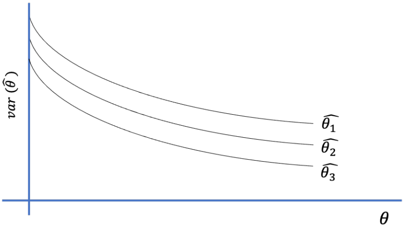
  where $\hat{\theta}$ is an unbiased estimator.
- For any other unbiased estimator $\tilde{\theta}$, if

$$var(\hat{\theta}) \leq var(\tilde{\theta})$$

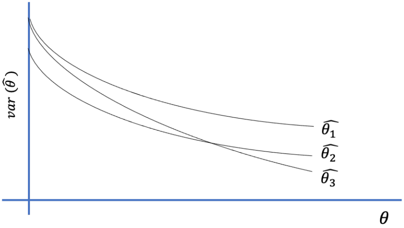
  then $\hat{\theta}$ is the Minimum Variance Unbiased estimator (MVU) for all θ.
- Does a MVU always exist i.e., an unbiased estimator with minimum variance for all θ ?

Existence of MVU

Consider a set of unbiased estimators $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3$,



(a)



(b)

$\hat{\theta}_3$ is the MVU

MVU does not exist

Example 3

- Consider two independent random processes x and y, defined as

$$x \sim \mathcal{N}(\theta, 1) \qquad y \sim \begin{cases} \mathcal{N}(\theta, 1), & \theta \geq 0 \\ \mathcal{N}(\theta, 2), & \theta < 0 \end{cases}$$

and let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two potential unbiased estimators for θ i.e.,

$$\hat{\theta}_1 = \frac{1}{2}(x + y), \qquad \hat{\theta}_2 = \frac{2}{3}x + \frac{1}{3}y$$

Example 4

- Consider two independent random processes x and y, defined as

$$x \sim \mathcal{N}(\theta, 1) \qquad y \sim \begin{cases} \mathcal{N}(\theta, 1)\,, & \theta \geq 0 \\ \mathcal{N}(\theta, 2)\,, & \theta < 0 \end{cases}$$

and let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two potential unbiased estimators for θ i.e.,

$$\hat{\theta}_1 = \frac{1}{2}(x + y)\,, \qquad \hat{\theta}_2 = \frac{2}{3}x + \frac{1}{3}y$$

- The variances of the the estimators are

$$var(\hat{\theta}_1) = \frac{1}{4}(var(x) + var(y)) \quad = \quad \begin{cases} \frac{18}{36}\,, & \theta \geq 0 \\ \frac{27}{36}\,, & \theta < 0 \end{cases}$$

$$var(\hat{\theta}_2) = \frac{4}{9}var(x) + \frac{1}{9}var(y) \quad = \quad \begin{cases} \frac{20}{36}\,, & \theta \geq 0 \\ \frac{24}{36}\,, & \theta < 0 \end{cases}$$

Example 5

- Consider two independent random processes x and y, defined as

$$x \sim \mathcal{N}(\theta, 1) \qquad y \sim \begin{cases} \mathcal{N}(\theta, 1), & \theta \geq 0 \\ \mathcal{N}(\theta, 2), & \theta < 0 \end{cases}$$

  and let $\hat{\theta}_1$ and $\hat{\theta}_2$ be two potential unbiased estimators for θ i.e.,

$$\hat{\theta}_1 = \frac{1}{2}(x + y), \qquad \hat{\theta}_2 = \frac{2}{3}x + \frac{1}{3}y$$

- The variances of the the estimators are

$$var(\hat{\theta}_1) = \frac{1}{4}(var(x) + var(y)) = \begin{cases} \frac{18}{36}, & \theta \geq 0 \\ \frac{27}{36}, & \theta < 0 \end{cases}$$

$$var(\hat{\theta}_2) = \frac{4}{9}var(x) + \frac{1}{9}var(y) = \begin{cases} \frac{20}{36}, & \theta \geq 0 \\ \frac{24}{36}, & \theta < 0 \end{cases}$$

- Neither $\hat{\theta}_1$, nor $\hat{\theta}_2$ are MVU estimators.

Finding the MVU

Even if the MVU exists, there is no standard "recipe" to find it

Some directions:

- Determine Cramér-Rao Lower Bound (Ch. 3)
- Apply Rao-Blackwell-Lehmann-Scheffe theorem (will not be discussed)
- Restrict estimators to be both unbiased AND linear (Ch. 6)

Summary

Key points:

- An unbiased estimator has zero bias i.e., $\mathbb{E}(\hat{\theta}) = \theta$
- MSE is composed of the variance and the bias^2 of the estimator
- MVU estimator is unbiased, with the lowest variance for all possible values of the unknown parameters
- MVU does not always exist, but can be found for some problems, under certain conditions

Next session:

- Estimator accuracy and the Cramér-Rao Lower Bound (CRLB)