eMasters in Communication Systems

Prof. Aditya Jagannatham

Elective Module:

Estimation for Wireless Communication

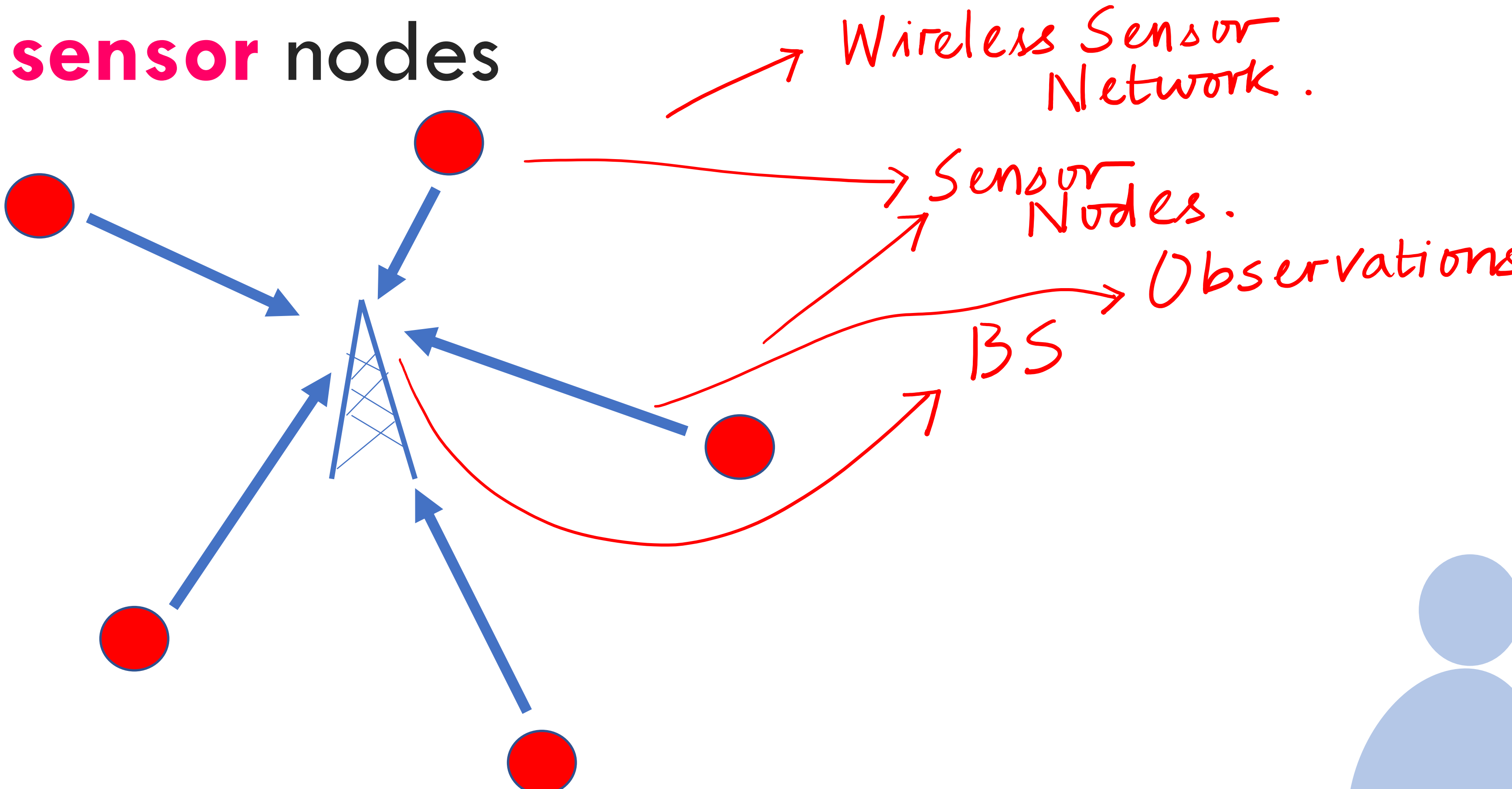
Chapter 1

(ML)

Maximum Likelihood Estimation

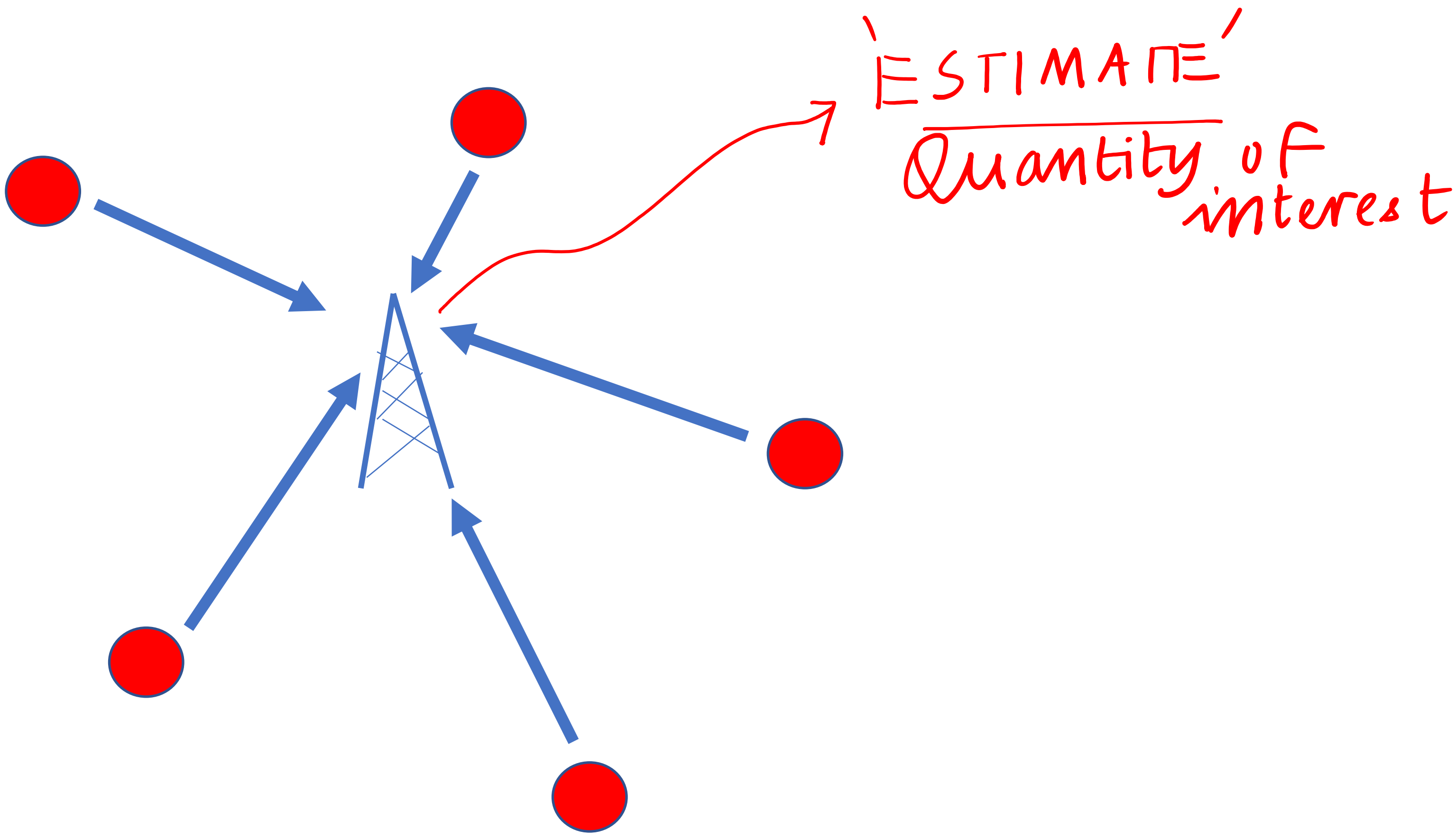
Introduction to Estimation

- Where do we use estimation?
- Consider large number of **wireless sensor** nodes



Introduction to Estimation

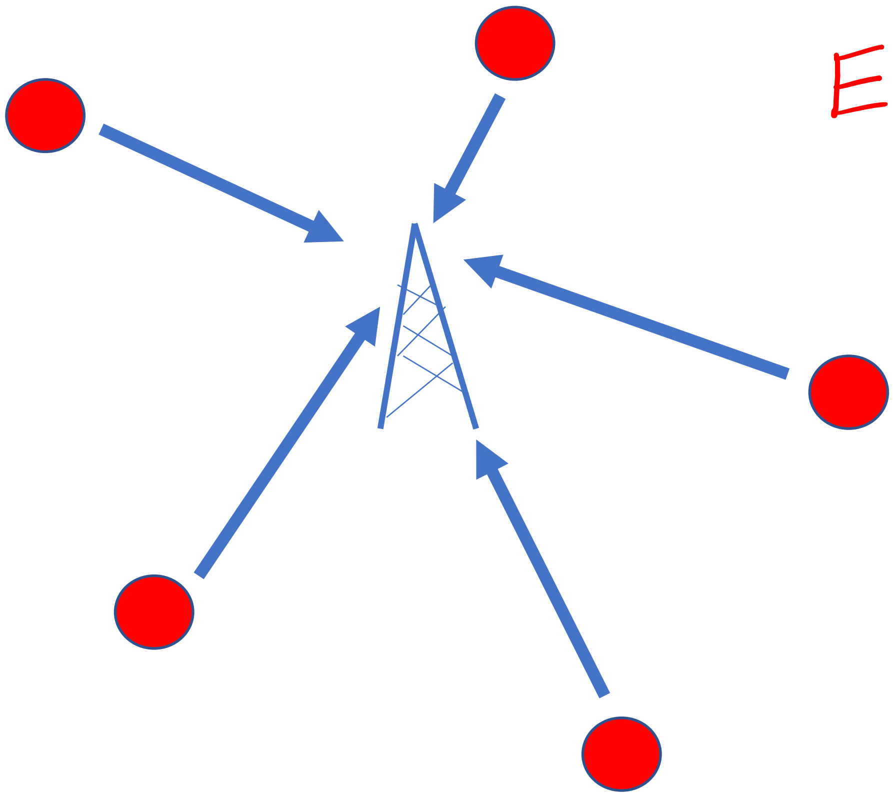
- Each sensor node transmits an ***observation***
  - Ex: Temperature, pressure, moisture level etc



Introduction to Estimation

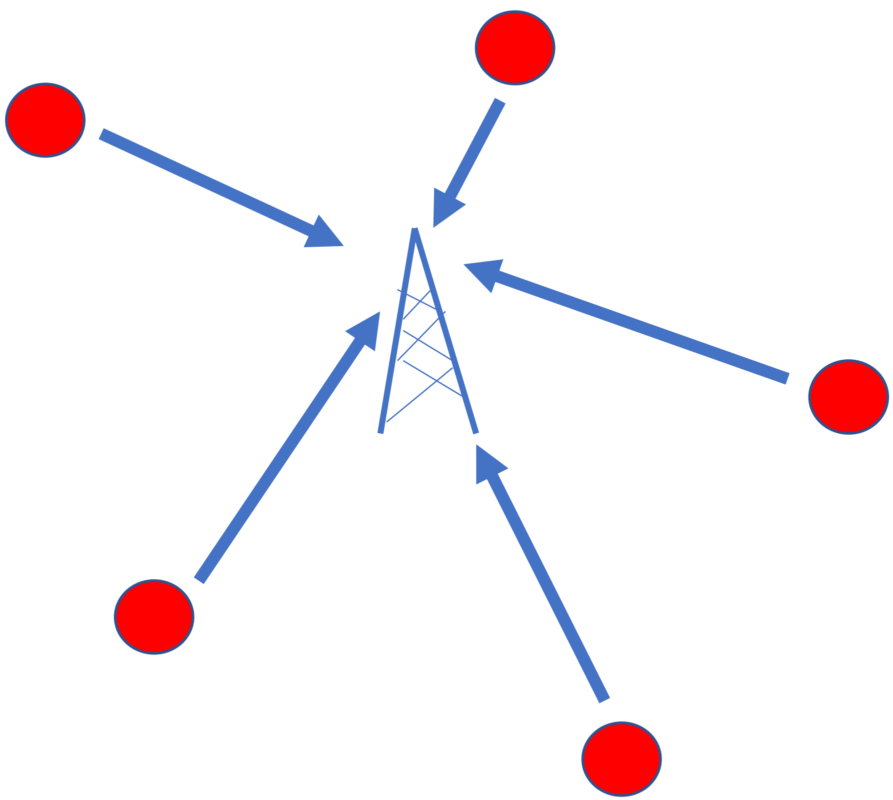
- How to estimate the ***unknown quantity*** at the BS?
  - This is termed as a PARAMETER.

Estimate
Parameter of interest



Introduction to Estimation

- How to estimate the ***unknown quantity*** at the BS?
  - This is termed as a **parameter**.



Estimation Model

- This can be modeled as follows

Measurement/ Observation

Parameter

$$y = h + v$$

Noise

Noisy Observation

- **Observation = Parameter + Noise**

Estimation Model

Gaussian Noise

- Let noise v be **Gaussian**
  - Mean $=$ 0 and variance $= \sigma^2$

$\mu = 0$

$F_V(v) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{v^2}{2\sigma^2}}$

Gaussian
Mean = h
Var = σ^2

$$y = h + v$$

unknown
constant
fixed.

- **What is PDF of y?**

Estimation Model

- PDF of y is given as

$$y \sim \mathcal{N}(h, \sigma^2)$$

$$\underbrace{f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-h)^2/2\sigma^2}}_{\substack{\text{Gaussian PDF} \\ \text{mean} = h \\ \text{var} = \sigma^2}}$$

Estimation Model

- PDF of y is given as
- y is **Gaussian** with mean h and variance σ^2

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-h)^2}{2\sigma^2}}$$

Estimation Model

N observations/measurements.

- Consider now N measurements

$$
\begin{cases}
y(1) = h + v(1) \\
y(2) = h + v(2) \\
\vdots \\
y(N) = h + v(N)
\end{cases}
$$

N = # measurements.

Estimation Model

$\underbrace{\bar{y} = h \cdot \bar{1} + \bar{v}}_{\text{vector model.}}$

- Consider now N measurements

$$\underbrace{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}_{\bar{y}} = \underbrace{h}_{h \cdot \bar{1}} + \underbrace{\begin{bmatrix} v(1) \\ v(2) \\ \vdots \\ v(N) \end{bmatrix}}_{\bar{v}}$$

$$\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Estimation Model

- PDF of each $y(k)$ is given as

Gaussian
mean = h
var = σ^2

Mean = 0
var = σ^2

$$y(k) = h + v(k)$$

$$\mathcal{N}(0, \sigma^2)$$

$$\underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-(y(k)-h)^2 / 2\sigma^2} = f_{Y(k)}(y(k))}_{\text{PDF of } y(k).}$$

Estimation Model

- PDF of each $y(k)$ is given as

$$f_{Y(k)}\big(y(k)\big) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y(k)-h)^2}{2\sigma^2}}$$

Estimation Model

Joint PDF of observations $y(1) \dots y(N)$.

- Considering IID noise samples

$$f_{\bar{\mathbf{Y}}}(\bar{\mathbf{y}}) = \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y(1)-h)^2}{2\sigma^2}}}_{\text{PDF } y(1)} \times \cdots \times \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y(N)-h)^2}{2\sigma^2}}}_{\text{PDF } y(N)}$$

$$= \underbrace{\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(y(k)-h)^2}}_{\text{Joint PDF}}$$

Likelihood Function

- This is termed as a ***likelihood function*** wrto h

$$p(\bar{y};h) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} \cdot e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(y(k)-h)^2}$$

Function of h
Termed Likelihood of h

How well does h
Explain observations.

Likelihood Function

- This is termed as a ***likelihood function*** wrto h

semi-colon

parameter is deterministic unknown

$$p(\bar{\mathbf{y}}; h) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(y(k)-h)^2}$$

Function of h.

Likelihood Function

Value of h that best explains observations.

- The estimate of h is obtained by maximizing the likelihood

$$\hat{h} = \max p(\bar{\mathbf{y}}; h)$$

max. likelihood function

$$= \max \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(y(k)-h)^2}$$

ML Estimate

MLE

Maximum Likelihood Estimate.

Likelihood Function

- The likelihood can be maximized as

constant

$$\max\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{N}{2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{N}(y(k)-h)^2}$$

–ve sign

maximize

minimize

constant

$$\equiv \min \sum_{k=1}^{N}(y(k)-h)^2$$

EQUIVALENT

minimization of above Qty.

Likelihood Function

- This can be minimized by differentiating wrto h and set to 0

$$\underbrace{\frac{d}{dh}\sum_{k=1}^{N}\left(y(k)-h\right)^2 = 0}_{\text{minimization}}$$

$$\sum_{k=1}^{N}\cancel{2}\left(y(k)-h\right)\cancel{(-1)} = 0$$

$$\sum_{k=1}^{N} \left(y(k) - h \right) = 0$$

$$\Rightarrow \quad N h = \sum_{k=1}^{N} y(k)$$

Sample Mean

$$\underbrace{\Rightarrow \quad \hat{h} = \frac{1}{N} \sum_{k=1}^{N} y(k)\,.}_{\text{ML Estimate}}$$

Likelihood Function

- This can be **minimized** by differentiating wrto h and set to 0

$$\frac{d}{dh}\sum_{k=1}^{N}(y(k)-h)^2 = 0$$

$$\Rightarrow \sum_{k=1}^{N}-2(y(k)-h) = 0$$

Likelihood Function

$$\Rightarrow \hat{h} = \frac{1}{N}\sum_{k=1}^{N} y(k)$$

ML ESTIMATE

- Sample Mean
- Average of observations.

Maximum Likelihood

- This is termed the ***Maximum Likelihood*** (ML) Estimate

  MLE

- It is also the sample mean of the observations

Properties of MLE

ESTIMATE

- We now explore properties of the ML Estimate

Random Qty.

- What is the distribution of $\hat{h}$

$$\hat{h} = \frac{1}{N} \sum_{k=1}^{N} y(k)$$

Properties of MLE

Gaussian

Gaussian

$$\hat{h} = \frac{1}{N}\sum_{k=1}^{N} y(k) = \underbrace{\frac{1}{N}\left(y(1)+\cdots+y(N)\right)}_{\text{Linear Transform}}$$

- $\hat{h}$ is a linear combination of **Gaussian RVs**

Average of Observations.

  - Hence, it is **Gaussian**

$$\hat{h} \sim \text{Gaussian}$$

$$\mathcal{N}$$

Properties of MLE

- What is mean of $\hat{h}$ $\quad \frac{1}{N}\sum_{k=1}^{N} y(k)$.

$$E\{\hat{h}\} =$$

$$E\{\hat{h}\} = ?$$

$$E\{\hat{h}\} = E\left\{\frac{1}{N}\sum_{k=1}^{N} y(k)\right\}$$

$$= \frac{1}{N}\sum_{k=1}^{N} E\{y(k)\} = \frac{1}{N}\sum_{k=1}^{N} E\{h + v(k)\}$$

$$= \frac{1}{N}\sum_{k=1}^{N}\Big(h + \underbrace{E\{v(k)\}}_{=0}\Big)$$

zero mean Gaussian

Properties of MLE

Mean of Estimate = True Parameter

$$E\{\hat{u}\} = \frac{1}{N}\sum_{k=1}^{N} u = \frac{1}{N} \times N u$$

$$= u$$

$$\boxed{E\{\hat{u}\} = u}$$

True Parameter

Very interesting!!! Property.

'UNBIASED' Estimator

Properties of MLE

- We now explore properties of the ML Estimate

$$E\{\hat{h}\} = E\left\{\frac{1}{N}\sum_{k=1}^{N} y(k)\right\} = \frac{1}{N}\sum_{k=1}^{N} E\{y(k)\}$$

$$= \frac{1}{N}\sum_{k=1}^{N} E\{h + v(k)\} = \frac{1}{N}\sum_{k=1}^{N} h = h$$

Properties of MLE

- Therefore

$$E\{\hat{h}\} = h$$

- This is termed an **unbiased estimate**

Properties of MLE

$$\underbrace{E\left\{(\hat{h}-h)^2\right\}}_{\text{Mean Square Error}}$$

- What about MSE?

$$E\left\{\left(\hat{h}-h\right)^2\right\}=?$$

Mean.

- This is also variance
- This can be found as follows

Properties of MLE

$$E\{|\hat{h}-h|^2\}.$$

$$E\{(\hat{h}-h)^2\}$$

$$= E\left\{\left(\left(\frac{1}{N}\sum_{k=1}^{N} y(k)\right) - h\right)^2\right\}$$

$$= E\left\{\left(\frac{1}{N}\sum_{k=1}^{N}\left(y(k)-h\right)\right)^2\right\}$$

$$= \frac{1}{N^2} E\left\{\left(\sum_{k=1}^{N} v(k)\right)^2\right\}.$$

$$y(k) = h + v(k)$$

$$\Rightarrow y(k) - h = v(k).$$

Properties of MLE

MSE

$$= \frac{1}{N^2} \cdot E\left\{ \left(\sum_{k=1}^{N} v(k) \right)^2 \right\}$$

$$= \frac{1}{N^2} \cdot E\left\{ \left(\sum_{k=1}^{N} v(k) \right) \left(\sum_{l=1}^{N} v(l) \right) \right\}.$$

$$= \frac{1}{N^2} \cdot E\left\{ \sum_{k=1}^{N} \sum_{l=1}^{N} v(k)\, v(l) \right\}$$

$$= \frac{1}{N^2} \cdot \sum_{k=1}^{N} \sum_{l=1}^{N} E\left\{ v(k)\, v(l) \right\}$$

Properties of MLE

$V(k)$ — i.i.d.

independent identically distributed.

$$E\{V(k)\,V(\ell)\} = \sigma^2\,\delta(k-\ell)$$

if $k \neq \ell$

$$E\{V(k)\}\,E\{V(\ell)\} = 0 \times 0 = 0.$$

if $k = \ell$

$$E\{V^2(k)\} = \sigma^2$$

Properties of MLE

$$E\{(\hat{h}-h)^2\} = \frac{\sigma^2}{N}..$$

$$MSE = \frac{1}{N^2} \cdot \sum_{k=1}^{N} \sum_{l=1}^{N} E\{v(k)\,v(l)\}$$

$$= \frac{1}{N^2} \cdot \sum_{k=1}^{N} \sum_{l=1}^{N} \sigma^2 \, \delta(k-l)$$

Mean Square Error

$$= \frac{1}{N^2} \cdot \sum_{k=1}^{N} \sigma^2 = \frac{1}{N^2} \cdot N\sigma^2$$

$$= \frac{\sigma^2}{N} = MSE = \text{Variance.}$$

$$\hat{h} \sim \mathcal{N}\left(h, \frac{\sigma^2}{N}\right)$$

Properties of MLE

$$E\left\{\left(\hat{h}-h\right)^2\right\} = E\left\{\left(\frac{1}{N}\sum_{k=1}^{N} y(k) - h\right)^2\right\}$$

$$= E\left\{\left(\frac{1}{N}\sum_{k=1}^{N}\big(y(k) - h\big)\right)^2\right\} = E\left\{\left(\frac{1}{N}\sum_{k=1}^{N} v(k)\right)^2\right\}$$

Properties of MLE

$$E\left\{\left(\frac{1}{N}\sum_{k=1}^{N} v(k)\right)^2\right\}$$

$$= \frac{1}{N^2} E\left\{\left(\sum_{k=1}^{N} v(k)\right)\left(\sum_{l=1}^{N} v(l)\right)\right\}$$

$$= \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N} E\{v(l)v(k)\}$$

Properties of MLE

$$\frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}E\{v(l)v(k)\} = \frac{1}{N^2}\sum_{k=1}^{N}\sum_{l=1}^{N}\sigma^2\delta(k-l)$$

$$= \frac{1}{N^2}\sum_{k=1}^{N}\sigma^2 = \frac{\sigma^2}{N}$$

MSE decreases as $\frac{1}{N}$.

Properties of MLE

- **Therefore, MSE decreases as**

$$MSE = \frac{\sigma^2}{N} \propto \frac{1}{N}$$

$MSE \to 0$ as $N \to \infty$.

Properties of MLE

- Therefore, $\hat{h}$ is Gaussian with mean 0 and variance $\dfrac{\sigma^2}{N}$

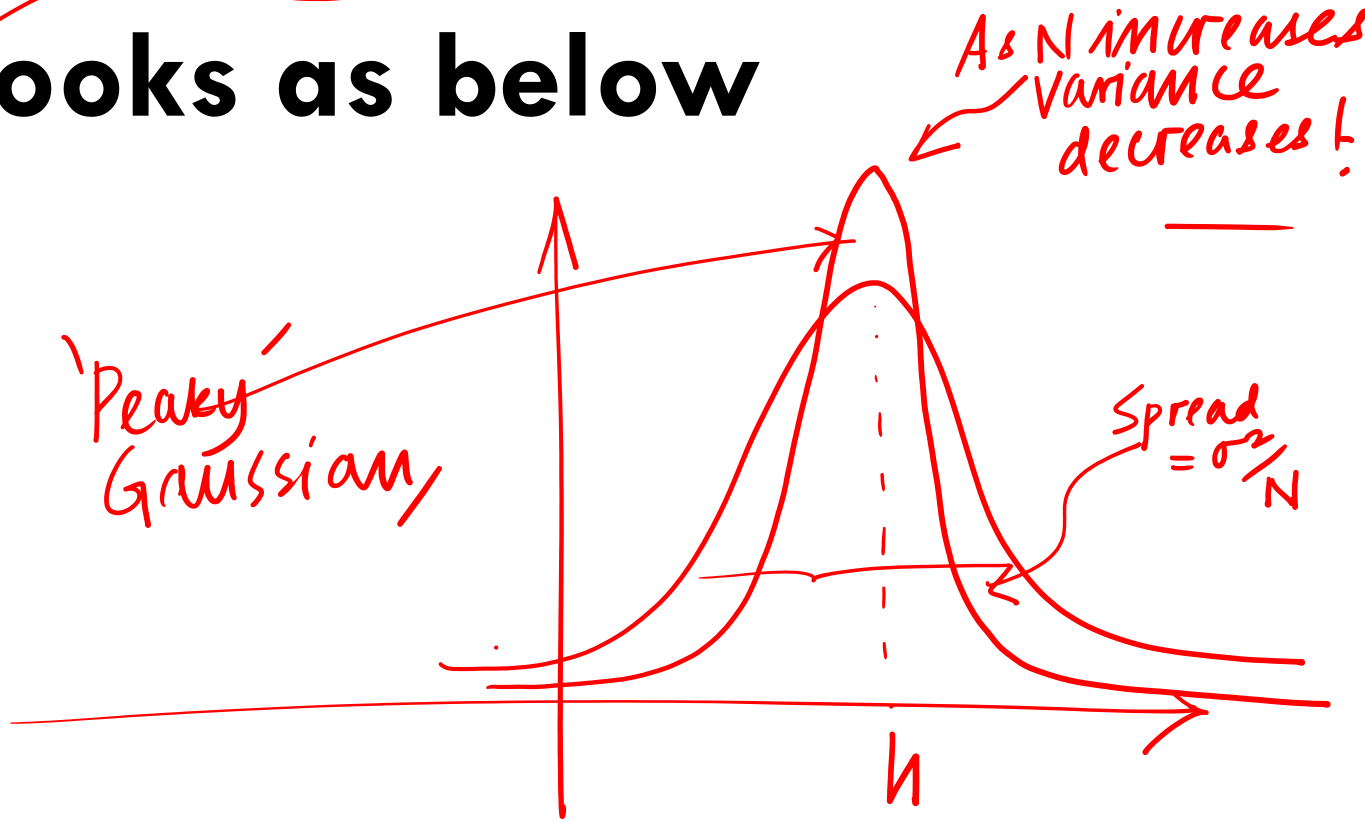
Mean

$$\hat{h} \sim \mathcal{N}\left(h, \frac{\sigma^2}{N}\right)$$

Variance.

Properties of MLE of $\hat{\mu}$

- PDF looks as below



MLE Example

$N = 4.$

- Consider the observations be $y(1) = 1$, $y(2) = -3$, $y(3) = 2$, $y(4) = -1$.
- What is the **maximum likelihood** estimate $\hat{h}$ of the unknown parameter h ?

$$\frac{1}{4}\{y(1) + y(2) + y(3) + y(4)\}$$

$$= \frac{1}{4}\{1 + (-3) + 2 + (-1)\} = -\frac{1}{4}.$$

MLE Example

- The ML estimate is given by the sample mean

MLE Example

- The ML estimate is given by the sample mean

$$\hat{h} = \frac{1}{N}\sum_{k=1}^{N} y(k) = \frac{1-3+2-1}{4} = -\frac{1}{4}$$

$$\hat{h} = \frac{-1}{4}.$$

MLE Example

- Consider the observations be $y(1) = 1$, $y(2) = -3, y(3) = 2, y(4) = -1$.
- IID Gaussian noise samples of variance $\sigma^2 = \frac{1}{4}$. What is the variance of the ML estimate ?

$$MSE = Variance = \frac{\sigma^2}{4} = \frac{\sigma^2}{N}$$

$$= \frac{1/4}{4} = \frac{1}{16}.$$

MLE Example

- Given $\sigma^2 = \frac{1}{4}$. The variance of the sample mean is $\frac{\sigma^2}{N}$.
- Given $N = 4$, the variance of the sample mean is

MLE Example

- Given $\sigma^2 = \frac{1}{4}$. The variance of the sample mean is $\frac{\sigma^2}{N}$.
- Given $N = 4$, the variance of the sample mean is

$$\frac{\sigma^2}{N} = \frac{\frac{1}{4}}{4} = \frac{1}{16}$$

MSE/Variance