Wednesday April 14 – Lecture 36

$Y_k = N_k + \mu S_k \qquad k = 1, 2 \ldots n \qquad N_k \sim N(0, \sigma^2)$ iid

$s = [s_1\, s_2 \ldots s_n]^T$

$\mu \in \mathbb{R} \qquad \sigma^2 > 0.$

Situation to Start … μ unknown, to be estimated; σ^2 Known.

$\mu \qquad \theta = \mu$

$$f_\mu(y) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^{n} (y_k - \mu s_k)^2 \right\}$$

$\log f_\mu(y)$ …. $\Longrightarrow$ and take $\frac{\partial}{\partial\mu} \log f_\mu(y) \overset{\text{Set}}{=} 0$ Solve $\mu = \hat{\mu}_{ML}(y)$

$$\frac{\partial}{\partial\mu}\left\{ \text{above} \right\}\Bigg|_{\mu = \hat{\mu}_{ML}} = \frac{1}{\sigma^2} \sum_{k=1}^{n} s_k (y_k - \hat{\mu}_{ML} s_k) = 0$$

$$\Longrightarrow \hat{\mu}_{ML}(y) = \frac{1}{n\,\overline{s^2}} \sum_{k=1}^{n} s_k y_k \qquad \overline{s^2} = \frac{1}{n} \sum_{k=1}^{n} s_k^2$$

Compute

$$\frac{\partial^2}{\partial \mu^2} \log f_\mu(y) = -\frac{1}{\sigma^2} \sum_{k=1}^{n} s_k^2 = \frac{-n\overline{s^2}}{\sigma^2} < 0$$

$\Rightarrow$ $\log f_\mu(y)$ is concave in μ so soln. to the likelihood eqn. gives a global max.

Recalling the prev. example: See that $\hat{\mu}_{ML}$ is same as the MVUE of μ whence

$$E_\mu\left\{\hat{\mu}_{ML}(Y)\right\} = \mu \qquad Var_\mu\left\{\hat{\mu}_{ML}(Y)\right\} = \frac{\sigma^2}{n\overline{s^2}}$$

$$= I_\mu^{-1} = CRLB.$$

See this from computing

$$-E_\mu\left\{\frac{\partial^2}{\partial \mu^2} \log f_\mu(Y)\right\}.$$

Also note

$$\frac{\partial}{\partial \mu} \log f_\mu(y) = \frac{1}{\sigma^2}\Big\{\underbrace{n \frac{1}{n}\sum_{k=1}^{n} s_k y_k}_{\overline{s^2}\,\hat{\mu}_{ML}(y)} - \mu n \overline{s^2}\Big\}$$

$$= \left[\frac{n\overline{s^2}}{\sigma^2}\right]\left[\hat{\mu}_{ML}(y) - \mu\right]$$ --- form needed for equality in CRLB.

Continue Example: μ is known; σ^2 unknown

For this case the likelihood is

$$\frac{\partial}{\partial \sigma^2} \log f_{\sigma^2}(y) \Big|_{\sigma^2 = \hat{\sigma}^2_{ML}(y)} = 0$$

$$= -\frac{n}{2\hat{\sigma}^2_{ML}} + \frac{1}{2(\hat{\sigma}^2_{ML})^2} \sum_{k=1}^{n} (y_k - \mu s_k)^2 = 0$$

$$= \frac{1}{2\hat{\sigma}^2_{ML}} \left[-n + \frac{1}{\hat{\sigma}^2_{ML}} \sum_{k=1}^{n} (y_k - \mu s_k)^2 \right] = 0$$

$\Rightarrow$ Unique soln (that depends on observation) is

$$\hat{\sigma}^2_{ML}(y) = \frac{1}{n} \sum_{k=1}^{n} (y_k - \mu s_k)^2$$

Also note:

$$\frac{\partial}{\partial \sigma^2} \log f_{\sigma^2}(y) = \frac{n}{2\sigma^4}\left[\hat{\sigma}^2_{ML}(y) - \sigma^2\right] \quad (*)$$

From which see ...

① derivative is < 0 for $\hat{\sigma}^2_{ML} < \sigma^2$
" " > 0 " $\hat{\sigma}^2_{ML} > \sigma^2$

$\Longrightarrow$ viewed as a funct of σ^2 see

$\log f_{\sigma^2}(y)$ is increasing for $\sigma^2 < \hat{\sigma}^2_{ML}$
and decreasing for $\sigma^2 > \hat{\sigma}^2_{ML}$

$\therefore$ abs max of $\log f_{\sigma^2}(y)$ is achieved @ $\sigma^2 = \hat{\sigma}^2_{ML}$

② From form $(*)$ see $\hat{\sigma}^2_{ML}$ is unbiased & and achieves the CRLB and that it is MVUE

$$I_{\sigma^2} = n/2\sigma^4 \quad \text{i.e.} \quad CRLB = \frac{2\sigma^4}{n}.$$

Continuing: Both μ and σ^2 are unknown. $\Theta = [\mu, \sigma^2]$ and search for max. of

$$\log f_\Theta(y) \quad \text{over} \quad \Theta.$$

Note that $\hat{\mu}_{ML}(y)$ which maximizes over μ does not depend on σ^2. Therefore

$$\max_{(\mu,\sigma^2)} \log f_\Theta(y) = \max_{\sigma^2} \left\{ \max_{\mu} \log f_\Theta(y) \right\}$$

$$= \max_{\sigma^2} \left\{ -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^{n} \left(y_k - \hat{\mu}_{ML}(y) s_k\right)^2 \right\}$$

But above is same max prob. as was solved in case of known μ (only with μ set to $\hat{\mu}_{ML}$). Thus, the MLE for unknown $\Theta = [\mu, \sigma^2]$ is

$$\hat{\Theta}_{ML}(y) = \begin{bmatrix} \hat{\mu}_{ML}(y) \\ \hat{\sigma}^2_{ML}(y) \end{bmatrix} = \begin{bmatrix} \frac{1}{n} \sum_{k=1}^{n} s_k y_k \Big/ \overline{s^2} \\ \frac{1}{n} \sum_{k=1}^{n} \left(y_k - \hat{\mu}_{ML}(y) s_k\right)^2 \end{bmatrix}$$

Also comparing to MVUE case ...

$$\hat{\mu}_{MV}(y) = \hat{\mu}_{ML}(y)$$

$$\hat{\sigma}^2_{MV}(y) = \left(\frac{n}{n-1}\right) \hat{\sigma}^2_{ML}(y)$$

Also

$$E_\theta\left\{\hat{\sigma}^2_{ML}(Y)\right\} = \frac{n-1}{n}\sigma^2$$

ie MLE of σ^2 is biased ($n \to \infty$ unbiased)

$$\mathrm{Var}_\theta\left\{\hat{\sigma}^2_{ML}(Y)\right\} = \left(\frac{n-1}{n}\right)^2 \mathrm{Var}_\theta\left\{\hat{\sigma}^2_{MV}(Y)\right\}$$

$$= \left(\frac{n-1}{n}\right)^2 \cdot \frac{2\sigma^4}{n-1}$$

Thus MLE actually has a lower variance than that of the MVUE. Explanation ... MLE is biased.

<u>Compare MSEs:</u>

$$\mathrm{MSE}_\theta\left\{\hat{\sigma}^2_{MV}(Y)\right\} \triangleq E_\theta\left\{\left(\hat{\sigma}^2_{MV}(Y) - \sigma^2\right)^2\right\}$$

$$= \mathrm{Var}_\theta\left\{\hat{\sigma}^2_{MV}(Y)\right\} = \frac{2\sigma^4}{n-1}$$

$$\mathrm{MSE}_\theta\left\{\hat{\sigma}^2_{ML}(Y)\right\} = E_\theta\left\{\left(\hat{\sigma}^2_{ML}(Y) - \sigma^2\right)^2\right\}$$

$$= \mathrm{Var}_\theta\left\{\hat{\sigma}^2_{ML}(Y)\right\} + \left(E_\theta\left\{\hat{\sigma}^2_{ML}(Y)\right\} - \sigma^2\right)^2$$

$$= \left(\frac{n-1}{n}\right)^2 \frac{2\sigma^4}{n-1} + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2$$

$$= \sigma^4 \, \frac{2n-1}{n^2}$$

$$\frac{MSE_{\theta}\left\{\hat{\sigma}^2_{MV}(Y)\right\}}{MSE_{\theta}\left\{\hat{\sigma}^2_{ML}(Y)\right\}} = \left(\frac{n}{n-1}\right)\left(\frac{2n}{2n-1}\right) > 1$$

In this case the MLE has a uniformly lower MSE than does the MVUE.

Here the increase in MSE of the MLE due to bias is more than off-set by the increase in variance of MVUE which was needed to achieve 0 bias.