CRLB Example

Setup: Coin tossed at times $k = 1, \ldots, n$ with heads-probability θ:

$$\Pr[Y_k = 1] = \theta \quad \& \quad \Pr[Y_k = 0] = (1-\theta).$$

Goal: Find MVUE and CRLB for estimation of $\theta \in \Lambda = [0,1]$.

For $\mathcal{Y}_q$ defined as the set of q-head sequences,

$$\Pr_\theta\{\underline{y} \in \mathcal{Y}_q\} = \begin{cases} \theta^q (1-\theta)^{n-q} & 0 \le q \le n \\ 0 & \text{else.} \end{cases}$$

Thus

$$\begin{aligned} p_\theta(\underline{y}) &= \begin{cases} \theta^{\sum_k y_k}(1-\theta)^{(n-\sum_k y_k)} & y_k \in \{0,1\} \\ 0 & \text{else.} \end{cases} \\ &= (1-\theta)^n \left(\frac{\theta}{1-\theta}\right)^{\sum_k y_k} \prod_k I_{\{0,1\}}(y_k) \\ &= \underbrace{(1-\theta)^n}_{C(\theta)} \exp\left\{ \underbrace{\log\left(\frac{\theta}{1-\theta}\right)}_{Q(\theta)} \underbrace{\sum_k y_k}_{T(\underline{y})} \right\} \underbrace{\prod_k I_{\{0,1\}}(y_k)}_{h(\underline{y})} \end{aligned}$$

Since we have an exponential family and Λ contains a rectangle, $T(\underline{Y})$ is complete sufficient statistic for θ.

Note that

$$\mathrm{E}\{T(\underline{Y})\} = \sum_{k=1}^{n} E\{Y_k\} = n\theta$$

Thus $\frac{1}{n}T(\underline{y})$ is an unbiased estimate of θ using $T(\underline{y})$.

$$\Rightarrow \quad \hat{\theta}(\underline{y}) = \frac{1}{n}T(\underline{y}) \text{ is MVUE for } \theta.$$

Furthermore

$$\begin{aligned} \mathrm{Var}_\theta(\hat{\theta}(\underline{Y})) &= \mathrm{Var}\left(\frac{1}{n}\sum_k Y_k\right) \\ &= \frac{1}{n^2}\mathrm{Var}(\textstyle\sum_k Y_k) \\ &= \frac{1}{n^2}\sum_k \mathrm{Var}(Y_k) \quad \text{since i.i.d.} \\ &= \frac{1}{n}\mathrm{Var}(Y_k) \\ &= \frac{1}{n}\left(\mathrm{E}\{Y_k^2\} - \mathrm{E}^2\{Y_k\}\right) \\ &= \frac{1}{n}(\theta - \theta^2) \end{aligned}$$

Now find CRLB...

$$
\begin{aligned}
p_\theta(\underline{y}) &= C(\theta)\exp[Q(\theta)T(\underline{y})]h(\underline{y}) \\
\log p_\theta(\underline{y}) &= \log C(\theta) + Q(\theta)T(\underline{y}) + \log h(\underline{y}) \\
&= n\log(1-\theta) + [\log(\theta) - \log(1-\theta)]T(\underline{y}) + \log h(\underline{y}) \\
\frac{\partial}{\partial\theta}\log p_\theta(\underline{y}) &= \frac{\partial}{\partial\theta}\left[T(\underline{Y})\log\theta + (n - T(\underline{Y}))\log(1-\theta)\right] \\
&= \frac{T(\underline{y})}{\theta} - \frac{n - T(\underline{y})}{1-\theta} \\
\frac{\partial^2}{\partial\theta^2}\log p_\theta(\underline{y}) &= -\frac{T(\underline{y})}{\theta^2} - \frac{n - T(\underline{y})}{(1-\theta)^2}
\end{aligned}
$$

Then

$$
\begin{aligned}
I_\theta &= -\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\log p_\theta(\underline{Y})\right] \\
&= \frac{\mathrm{E}[T(\underline{y})]}{\theta^2} + \frac{n - \mathrm{E}[T(\underline{y})]}{(1-\theta)^2} \\
&= \frac{n\theta}{\theta^2} + \frac{n - n\theta}{(1-\theta)^2} \\
&= \frac{n}{\theta(1-\theta)}
\end{aligned}
$$

The CRLB says that, for any unbiased estimator $\hat{\theta}(\underline{Y})$,

$$
\mathrm{Var}_\theta\big(\hat{\theta}(\underline{Y})\big) \;\geq\; \frac{1}{I_\theta} = \frac{\theta(1-\theta)}{n}
$$

Note that the MVUE achieves the CRLB.

Could we have predicted this?