ECE531 Screencast 3.6: A Complete Example on Using the RBLS Theorem

D. Richard Brown III

Worcester Polytechnic Institute

Problem Statement

Suppose $Y_k \overset{\text{i.i.d.}}{\sim} p_Y(y\,;\,\theta)$ for $k = 0, \dots, n-1$ where

$$p_Y(y\,;\,\theta) = \begin{cases} \frac{y}{\sigma^2} \exp\left(\frac{-y^2}{2\sigma^2}\right) & y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

1. Find a complete sufficient statistic for estimating the scalar parameter $\theta = \sigma^2$.
2. Use the RBLS theorem to find the MVU estimator.

Solution Step 1

First we should write out the joint density. Since the observations are i.i.d., the joint density is just the product of the marginal densities and we have

$$p_Y(y\,;\,\theta) = \frac{1}{\theta}\left(\prod_{k=1}^{n} y_k\right)\exp\left\{-\frac{\sum_{k=1}^{n} y_k^2}{2\theta}\right\} I(y>0)$$

where

$$I(y>0) = \begin{cases} 1 & y_k > 0 \text{ for all } k \\ 0 & \text{otherwise.} \end{cases}$$

Now we need to find a sufficient statistic. We could guess and check (by conditioning on the statistic to see if the conditional pdf no longer depends on θ) or we could use the Neyman-Fisher Factorization Theorem.

Solution Step 2

Let's use N-F. We want

$$p_Y(y\,;\,\theta) = \frac{1}{\theta}\left(\prod_{k=1}^{n} y_k\right)\exp\left\{-\frac{\sum_{k=1}^{n} y_k^2}{2\theta}\right\} I(y>0)$$

to look like $g_\theta(T(y))h(y)$. Lets try

$$T(y) = \sum_{k=1}^{n} y_k^2,$$

Then, rearranging a bit, we have

$$p_Y(y\,;\,\theta) = \underbrace{\frac{1}{\theta}\exp\left\{-\frac{\sum_{k=1}^{n} y_k^2}{2\theta}\right\}}_{g_\theta(T(y))}\underbrace{\left(\prod_{k=1}^{n} y_k\right) I(y>0)}_{h(y)}$$

Note that this is a valid factorization because $g_\theta(T(y))$ is only a function of $T(y)$ and θ and also because $h(y)$ is only a function of y. So the N-F factorization theorem tells us that $T(y)$ is a sufficient statistic for the estimation of $\theta = \sigma^2$.

Solution Step 3

To check for completeness, we'll try to use the completeness theorem for exponential families. We have a scalar parameter here, so we want to check if

$$p_Y(y\,;\,\theta) \quad = \quad a(\theta)\exp\left\{q(\theta)T(y)\right\}h(y)$$

We have

$$p_Y(y\,;\,\theta) = \underbrace{\frac{1}{\theta}}_{a(\theta)} \exp\left\{ -\underbrace{\frac{1}{2\theta}}_{q(\theta)} \underbrace{\sum_{k=1}^{n} y_k^2}_{T(y)} \right\} \underbrace{\left(\prod_{k=1}^{n} y_k\right) I(y>0)}_{h(y)}$$

Hence, the Completeness Theorem for Exponential Families implies that $T(y)$ is a complete sufficient statistic.

Solution Step 4

We have a complete sufficient statistic. Now we need to apply the RBLS theorem. Since the observations Y_k are i.i.d. Rayleigh distributed, we know

$$\mathrm{E}[Y_k] = \sigma\sqrt{\frac{\pi}{2}} = \sqrt{\frac{\theta\pi}{2}}$$
$$\mathrm{var}(Y_k) = \frac{4-\pi}{2}\sigma^2 = \frac{4-\pi}{2}\theta$$

for $k = 0, \dots, n-1$. Hence we can compute

$$\mathrm{E}\left[\sum_{k=0}^{n-1} Y_k^2\right] = n\mathrm{var}(Y_k) + n(\mathrm{E}[Y_k])^2 = \frac{n(4-\pi)}{2}\theta + \frac{n\pi}{2}\theta = 2n\theta.$$

This gives a good idea for an unbiased estimator: $\hat{g}(y) = \frac{1}{2n}\sum_{k=0}^{n-1} y_k^2$.

This is not necessarily the MVU estimator, however. One more step...

Solution Step 5

We have a complete sufficient statistic $T(y)$ and an unbiased estimator $\hat{g}(y)$. The last step is to compute the conditional expectation

$$\tilde{g}[T(y)] \quad := \quad \mathrm{E}\left[\hat{g}(Y) \,|\, T(Y) = T(y)\right]$$

then we know $\tilde{g}[T(y)]$ is an MVU estimator of of θ.

Note that $\hat{g}(Y) = \frac{1}{2n}T(Y)$. Then

$$\begin{aligned}
\tilde{g}[T(y)] &= \mathrm{E}\left[\hat{g}(Y) \,|\, T(Y) = T(y)\right] \\
&= \mathrm{E}\left[\frac{1}{2n}T(Y) \,|\, T(Y) = T(y)\right] \\
&= \frac{1}{2n}\mathrm{E}\left[T(Y) \,|\, T(Y) = T(y)\right] \\
&= \frac{1}{2n}T(y) = \frac{1}{2n}\sum_{k=0}^{n-1} y_k^2
\end{aligned}$$

Hence $\hat{\theta}(y) = \frac{1}{2n}\sum_{k=0}^{n-1} y_k^2$ is MVU.