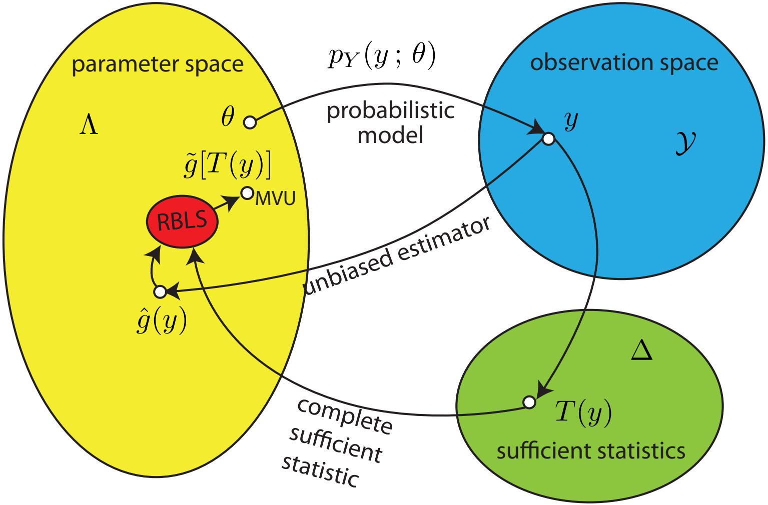
ECE531 Screencast 3.3: Introduction to the Rao-Blackwell-Lehmann-Sheffe Theorem

D. Richard Brown III

Worcester Polytechnic Institute

Conceptual Model



RBLS: A Procedure for Finding MVU Estimators

To find an MVU estimator for the non-random parameter θ, we can follow a three-step procedure:

1. Find a **complete sufficient statistic** $T : \mathcal{Y} \to \Delta$ for the family of pdfs $\{p_Y(y\,;\,\theta)\,;\,\theta \in \Lambda\}$ parameterized by θ.
2. Find *any* unbiased estimator $\hat{g}(y)$ of θ.
3. Compute $\tilde{g}[T(y)] = \mathrm{E}[\hat{g}(Y)\,|\,T(Y) = T(y)]$.

The **Rao-Blackwell-Lehmann-Sheffe Theorem** says that $\tilde{g}[T(y)]$ will be a MVU estimator of the non-random parameter θ.

Some Intuition About Sufficient Statistics

Suppose we have observations given by

$$Y_k = \theta + W_k \qquad k = 0, \dots, n-1$$

where $W_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$. We already know that

$$\hat{\theta}(y) = \frac{1}{n} \sum_{k=0}^{n-1} y_k$$

is a MVU estimator.

- The set of observations is $\mathcal{S} = \{y_0, \dots, y_{n-1}\}$.
- If we threw away some of the observations, would we still be able to compute the MVU estimator?
- What about these observation sets?

$$\mathcal{S}' = \{y_0\}$$
$$\mathcal{S}'' = \{y_0 + y_1, y_2, \dots, y_{n-1}\}$$
$$\mathcal{S}''' = \{y_0 + y_1 + \cdots + y_{n-1}\}$$

Some Intuition About Sufficient Statistics

- The original set of observations is always a "sufficient statistic". But there are often smaller sets that contain the **relevant** information.
- A sufficient statistic **summarizes** all of the information in the original set of observations so that, if you condition the pdf of the observations on the sufficient statistic, the pdf no longer depends on the unknown parameter.
- If you are given the sufficient statistic, you can ignore the observations. They contain no additional relevant information.
- There may be many different sufficient statistics.
- The sufficient statistic is not allowed to be a function of the unknown parameter θ. It may only be a function of the observations (and other known parameters, e.g. noise variance, number of observations, ...).