Lecture 20: Best Unbiased Estimator

22 March 2016

1 Best unbiased estimation, sufficient statistics and Rao-Blackwell theorem

"Conditioning an estimator on a sufficient statistic preserves bias and reduces variance."

1.1 Uniqueness of Best Unbiased Estimator (BUE)

Theorem 1.1. *If W is a BUE of $g(\theta)$, then W is unique.*

Proof. Suppose that there exists W^1, another BUE of $g(\theta)$. Consider,

$$W^* = \frac{1}{2}(W + W^1). \tag{1}$$

We have,

$$E_\theta[W^*] = g(\theta), \text{ and} \tag{2}$$

$$Var_\theta[W^*] = \frac{1}{4}Var_\theta[W] + \frac{1}{4}Var_\theta[W^1] + \frac{1}{2}Cov_\theta[W, W^1]. \tag{3}$$

Using Cauchy-Schwartz inequality, we get,

$$Var_\theta[W^*] \leq \frac{1}{4}Var_\theta[W] + \frac{1}{4}Var_\theta[W^1] + \frac{1}{2}\sqrt{(Var_\theta[W])^2}$$

$$Var_\theta[W^*] \leq Var_\theta[W]. \tag{4}$$

Since W is a BUE of $g(\theta)$, $Var[W^*]$ cannot be less than $Var[W]$. So eqn. (4) must hold with equality.

Let $W^1 = a(\theta)W + b(\theta)$, then

$$Cov_\theta[W, W^1] = Cov_\theta[W, (a(\theta)W + b(\theta))],$$

$$Cov_\theta[W, W^1] = a(\theta)Var_\theta(W). \tag{5}$$

From eqn. (4) using the definition W^*, we get

$$Cov_\theta[W, W^1] = Var_\theta(W). \tag{6}$$

This gives,

$$a(\theta) = 1, \ \forall \theta. \tag{7}$$

As W and W^1 are unbiased estimators it follows that, $b(\theta) = 0$. Hence,

$$W = W^1,$$

which shows that W is unique if it is a BUE of some function $g(\theta)$. □

1.2 Characterization of BUE

Theorem 1.2. *If $\mathbb{E}_\theta[W] = g(\theta)$ for all $\theta \in \Theta$, then W is a BUE of $g(\theta)$ if and only if W is uncorrelated with all unbiased estimator of 0 (could be any function $W(x)$ with zero expectation).*

Proof. To prove the "if" statement, assume that W is the BUE of $g(\theta)$. Let U be an unbiased estimator of 0, that is,

$$\mathbb{E}_\theta[U] = 0 \ \forall \, \theta.$$

Then, the estimator

$$W_a := W + aU, \ a \in \mathbb{R}$$

is an unbiased estimator of $g(\theta)$.

$$\begin{aligned} Var_\theta[W_a] &= Var_\theta[W + aU], \\ &= Var_\theta[W] + 2aCov_\theta[W, U] + a^2 Var_\theta[U]. \end{aligned} \tag{8}$$

<u>Case 1:</u> Let $Cov_\theta[W, U] < 0$ for some θ. Which makes,

$$2aCov_\theta[W, U] + a^2 Var_\theta[U] < 0 \ \text{ if } \ a \in \left(0, \frac{-2Cov_\theta[W, U]}{Var_\theta[U]}\right) \tag{9}$$

<u>Case 2:</u> $Cov_\theta[W, U] > 0$. This gives $Var_\theta[W_a] < Var_\theta[W]$ for a suitable choice of a.

Both cases 1 and 2 lead to contradictions. Therefore, $Cov_\theta[W, U] = 0$.

To prove the "only if" statement, let us assume that $Cov_\theta[W, U] = 0, \ \forall \, \theta$, whenever $\mathbb{E}_\theta[U] = 0 \ , \forall \, \theta$. Let W' be any other unbiased estimator of $g(\theta)$.

$$\mathbb{E}_\theta[W] = \mathbb{E}_\theta[W'] = g(\theta).$$

Consider $W' = W + (W' - W)$, (note that $W - W'$ is the unbiased estimator of 0),

$$\begin{aligned} Var[W'] &= Var[W] + 2Cov(W, W' - W) + Var(W' - W) \\ &\geq Var[W] \end{aligned}$$

Therefore, W is a BUE of $g(\theta)$. □

Corollary 1.3. *If the only unbiased estimator of* 0 *is* 0 *itself, then* W *is the BUE of* $\mathbb{E}_\theta[W]$.

Definition 1.4. A family of pdfs $\{f(x|\theta), \theta \in \Theta\}$ over $\mathbb{R}^d$ is said to be complete, if

$$\mathbb{E}_\theta[g(x)] = 0, \ \forall \ \theta \in \Theta, \ \forall \, g : \mathbb{R}^d \to \mathbb{R} \ \text{ implies } \mathbb{P}[g(x) = 0] = 1.$$

Example 1.5. The binomial family (with known number of trails),

$$\{Bin(n, \theta) : \theta \in [01]\}$$

is complete if,

$$\sum_{m=0}^{n} \binom{n}{m} \theta^m (1-\theta)^{n-m} g(m) = 0.$$

This requires,

$$g(m) = 0, \quad m = 1, 2, \ldots, n$$

Definition 1.6. Let $X \sim f(x|\theta)$. A statistic $T(X)$ said to be a "complete statistic" if the family of pdfs (or pmfs) of $T(X)$ induced by the pdf (or pmf) of X is complete.

1.3 BUE and complete sufficient statistic

Theorem 1.7. *Let* T *be a complete and sufficient statistic for* θ. *Let* $\Phi(T)$ *be any estimator based on* T. *Then* $\Phi(T)$ *is the BUE of its expectation,* $\mathbb{E}_\theta[\Phi(T)]$.

Proof. The family of distributions $\{Q_\theta : \theta \in \Theta\}$ of $T(X)$, included by $X \sim f(x|\theta)$ is complete, which implies that no unbiased estimator of 0 based on T other-than 0 exists. It follows that $T(X)$ is uncorrelated with all unbiased estimators of 0 based on T.

Therefore, T is the BUE of $g(\theta) := \mathbb{E}_\theta[T]$. □

Example 1.8. Uniform BUE:
Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} Unif[0 \ \theta], \ \theta \in (0 \ \infty)$, and let $Y := \max_{i=1,\ldots,n} X_i$,. We have,

$$\frac{Y}{\theta} \sim Beta(n, 1) \tag{10}$$

and $\mathbb{E}[Y] = (\frac{n}{n+1})\theta$.

Remark 1. If T is a finite dimensional complete and sufficient statistic then T is a minimal sufficient statistic.

1.4 Sufficiency, Minimal Sufficiency and Completeness for Exponential families

Theorem 1.9. *Suppose* $X_1, X_2, \ldots, X_n \overset{iid}{\sim} f(x|\theta)$, *where* $f(x|\theta) = h(x)C(\theta)\exp\{\sum_{i=1}^{k} \theta_i t_i(x)\}$, *and* $\theta \equiv (\theta_1, \theta_2, \ldots, \theta_k) \subseteq \Theta \subseteq \mathbb{R}^k,\ t_i : \Gamma \to \mathbb{R}$, *with the joint pdf (or pmf) being*

$$f(x_1, x_2, \ldots, x_n|\theta) = \left(\prod_{i=1}^{n} h(x_i)\right) (C(\theta))^n \exp\left(\sum_{i=1}^{k} \theta_i \sum_{j=1}^{n} t_i(x_j)\right).$$

Consider the statistic

$$T(X_1, X_2, \ldots, X_n) = \left(\sum_{j=1}^{n} t_1(X_j)\right) \cdot \left(\sum_{j=1}^{n} t_2(X_j)\right) \cdots \left(\sum_{j=1}^{n} t_k(X_j)\right),$$

$T(X)$ *is,*

- *A sufficient statistic for* θ.
- *A minimal sufficient statistic for* θ, *when* $\theta_1, \ldots, \theta_k$ *do not satisfy a linear relation. (That is* Θ *is not to be conditioned in a vector space of dimension less than* k*).*
- *A complete sufficient statistic of* Θ *contains a* k*-dimensional rectangle i.e., a set of the form* $[a_1,\ b_1] \times [a_2,\ b_2] \times \cdots \times [a_k,\ b_k]$.