

Lecture 18: Complete class theorems; sufficiency; estimation viewed as a decision problem

Statistical Inference–II ■ Unit I: Decision Theory

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

M.Sc. Statistics, Semester III
Academic Session 2026–2027

Previous lecture

Lecture 17: Complete, essentially complete and minimal complete classes and their relationship.

Today

- Complete Classes of Decision Rules
 - Complete Class Theorems
- Estimation Viewed as a Decision Problem
 - The Correspondence
 - Unbiasedness as a Restriction
 - Unbiased Decision Rules

The value of the theory lies in the following results, which identify complete classes with families of Bayes rules. We state them without proof; see Ferguson (1967, Ch. 2) and Berger (1985, Ch. 8).

Theorem 18.1 (Finite Θ ; complete class theorem)

Let $\Theta = \{\theta_1, \dots, \theta_k\}$ be finite and suppose the risk set

$$S = \{(R(\theta_1, \delta), \dots, R(\theta_k, \delta)) : \delta \in \mathcal{D}^*\} \subseteq \mathbb{R}^k$$

(over the class $\mathcal{D}^$ of all randomised rules) is bounded from below and closed. Then the class of Bayes rules is complete, and the class of admissible Bayes rules is minimal complete.*

A version of this theorem for an infinite parameter space, due to Wald, is stated in Lecture Notes, Appendix A.

Theorem 18.2 (Sufficiency under convex loss)

Let $T = T(\mathbf{X})$ be sufficient for θ , let $\mathcal{A}$ be a convex subset of a vector space, and suppose that $L(\theta, \cdot)$ is convex for every θ . Then the class of non-randomised rules based only on T is essentially complete among non-randomised rules δ with $\mathbb{E}_\theta|\delta(\mathbf{X})| < \infty$ for all θ .

If the loss is strictly convex, $R(\theta, \delta) < \infty$, and a rule $\delta(\mathbf{X})$ that is not a function of T has a non-degenerate conditional distribution given T with positive probability under $\mathbb{P}_\theta$, the resulting improvement is strict at that θ .

Proof.

Given a non-randomised rule δ with $\mathbb{E}_\theta |\delta(\mathbf{X})| < \infty$ for all θ , define

$$\bar{\delta}_T(t) = \mathbb{E}[\delta(\mathbf{X}) \mid T = t].$$

Sufficiency makes this conditional expectation independent of θ , and convexity of $\mathcal{A}$ ensures that $\bar{\delta}_T(t) \in \mathcal{A}$. Conditionally on T , Jensen's inequality gives

$$L(\theta, \bar{\delta}_T(T)) = L(\theta, \mathbb{E}[\delta(\mathbf{X}) \mid T]) \leq \mathbb{E}[L(\theta, \delta(\mathbf{X})) \mid T].$$

Taking expectations yields $R(\theta, \bar{\delta}_T) \leq R(\theta, \delta)$ for every θ . Under the additional strict-convexity and non-degeneracy condition, strict Jensen inequality gives a strict risk reduction for at least one θ . □

Remark 18.3

Theorem 18.2 is the decision-theoretic form of the Rao–Blackwell theorem: under a convex loss nothing is lost by basing a rule on a sufficient statistic, and the conditional-mean rule is usually strictly better.

Without convexity the same conclusion holds provided randomised rules are allowed (Lecture Notes, Appendix A).

We now make explicit the correspondence, already used implicitly, between point estimation and decision theory.

Decision-theoretic object	Estimation counterpart
State of nature θ	Unknown parameter
Action space $\mathcal{A}$	Range of the estimand $\gamma(\theta)$
Decision rule $\delta(\mathbf{X})$	Estimator $\hat{\gamma}(\mathbf{X})$
Loss $L(\theta, a)$	Typically $(a - \gamma(\theta))^2$
Risk $R(\theta, \delta)$	Mean squared error
Admissible rule	Admissible estimator
Bayes rule	Bayes estimator (posterior mean under SELF)
Minimax rule	Minimax estimator

The classical theory of minimum-variance unbiased estimation is the decision-theoretic strategy (A) of the section on uniformly best rules: restrict $\mathcal{D}$ to the unbiased rules and seek a uniformly best rule within that class. Under SELF the risk of an unbiased estimator is exactly its variance (the example “Risk of an estimator under SELF” with $b(\theta) \equiv 0$), so the UMVUE is precisely the uniformly best rule in the restricted class.

Remark 18.4 (UMVUE need not be minimax or admissible)

The restriction to unbiasedness is imposed externally and can be costly from a decision-theoretic point of view. In the binomial problem, $\bar{p} = X/n$ is the UMVUE, whereas the estimator δ_M in Example 16.1 has a strictly smaller maximum risk. This shows that a UMVUE need not be minimax; it does *not* show that X/n is inadmissible, because the two risk functions cross and neither estimator dominates the other.

A genuine inadmissibility example is provided by normal variance estimation. For a normal sample with both the mean and variance unknown, the unbiased estimator

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

of σ^2 is inadmissible under squared error loss; see Exercise 1.11.

The notion of unbiasedness itself admits a decision-theoretic generalisation due to Lehmann (1951); see also Lehmann and Romano (2022).

Definition 18.5 (L -unbiased decision rule)

A decision rule δ is L -unbiased if

$$\mathbb{E}_{\theta}[L(\theta', \delta(\mathbf{X}))] \geq \mathbb{E}_{\theta}[L(\theta, \delta(\mathbf{X}))] \quad \text{for all } \theta, \theta' \in \Theta.$$

In words: on average, the rule comes closer (in the sense of L) to the correct state of nature than to any incorrect one.

Proposition 18.6 (L -unbiasedness under squared error loss)

Let $\Gamma = \gamma(\Theta)$ and assume that $\mathbb{E}_\theta[\delta(\mathbf{X})^2] < \infty$ for every $\theta \in \Theta$ and that $m(\theta) = \mathbb{E}_\theta[\delta(\mathbf{X})]$ belongs to Γ for every $\theta \in \Theta$. Then, under squared error loss $L(\theta, a) = \{a - \gamma(\theta)\}^2$, the rule δ is L -unbiased if and only if it is unbiased in the ordinary sense:

$$\mathbb{E}_\theta[\delta(\mathbf{X})] = \gamma(\theta) \quad \text{for every } \theta \in \Theta.$$

In particular, the conclusion holds whenever $\gamma(\Theta) = \mathbb{R}$.

Proof: L -Unbiasedness under Squared Error Loss

Proof.

Write $m(\theta) = \mathbb{E}_\theta[\delta(\mathbf{X})]$. For any θ' ,

$$\mathbb{E}_\theta \left[\{ \delta(\mathbf{X}) - \gamma(\theta') \}^2 \right] = \text{Var}_\theta(\delta(\mathbf{X})) + \{ m(\theta) - \gamma(\theta') \}^2.$$

The variance term does not depend on θ' . Hence L -unbiasedness is equivalent to

$$\{ m(\theta) - \gamma(\theta') \}^2 \geq \{ m(\theta) - \gamma(\theta) \}^2 \quad \text{for every } \theta'.$$

By assumption, $m(\theta) \in \Gamma$, so there exists a θ' satisfying $\gamma(\theta') = m(\theta)$. Substitution makes the left-hand side zero and therefore forces $m(\theta) = \gamma(\theta)$. Conversely, if $m(\theta) = \gamma(\theta)$, the required inequality is immediate. □

Remark 18.7

The range condition $m(\theta) \in \gamma(\Theta)$ cannot simply be omitted; Lecture Notes, Appendix A gives an L -unbiased estimator that is biased when it fails.

Practice (Lecture Notes, Chapter 1)

- Exercises (Section 1.13): 1.1, 1.11, 1.18, 1.19.

Next lecture

Lecture 19: Testing viewed as a decision problem; the Bayes test.