Lecture 3: Randomised decision rules; convexity of the risk set

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

This lecture

Previous lecture

Lecture 2: Non-randomised decision rules; the risk function; risk of an estimator and of a test.

Today

- Decision Rules and the Risk Function
  - Randomised Decision Rules

Randomised decision rules

In some problems it is advantageous — and in some it is unavoidable — to allow the action to be chosen by an auxiliary random mechanism after $\mathbf{x}$ has been observed. This is familiar from the theory of testing, where randomised tests are needed to attain an exact size α in discrete problems.

Definition 3.1 (Randomised decision rule)

A *randomised decision rule* δ^* is a function which assigns to each $\mathbf{x} \in \mathcal{X}$ a probability distribution $\delta^*(\cdot \mid \mathbf{x})$ on the action space $\mathcal{A}$. When $\mathbf{x}$ is observed, the action actually taken is drawn from $\delta^*(\cdot \mid \mathbf{x})$ by an auxiliary randomisation whose law does not depend on θ.

Randomised rules: finite action space; test functions

When $\mathcal{A} = \{a_1, a_2, \ldots, a_k\}$ is finite, δ^* is described by k functions $p_1(\mathbf{x}), \ldots, p_k(\mathbf{x})$ with

$$p_j(\mathbf{x}) \geq 0 \quad \text{and} \quad \sum_j p_j(\mathbf{x}) = 1,$$

where $p_j(\mathbf{x})$ is the probability of taking action a_j when $\mathbf{x}$ is observed.

- A non-randomised rule is the special case in which, for each $\mathbf{x}$, some $p_j(\mathbf{x}) = 1$.

For the two-action problem it is customary to write a randomised rule as a single *test function*

$$\phi(\mathbf{x}) = \mathbb{P}\left(\text{take action } a_1 \mid \mathbf{X} = \mathbf{x}\right) \in [0, 1].$$

Definition 3.2 (Risk of a randomised rule)

The loss incurred by a randomised rule at $(\theta, \mathbf{x})$ is itself an expected value over the randomisation:

$$L\left(\theta, \delta^*(\cdot \mid \mathbf{x})\right) = \int_{\mathcal{A}} L(\theta, a)\, \delta^*(da \mid \mathbf{x}).$$

The risk function is therefore

$$R(\theta, \delta^*) = \int_{\mathcal{X}} \left[\int_{\mathcal{A}} L(\theta, a)\, \delta^*(da \mid \mathbf{x}) \right] f(\mathbf{x} \mid \theta)\, d\mathbf{x}.$$

In the finite-action case this reduces to

$$R(\theta, \delta^*) = \mathbb{E}_\theta \left[\sum_{j=1}^{k} L(\theta, a_j)\, p_j(\mathbf{X}) \right].$$

Why randomisation matters

Proposition 3.3 (Randomisation convexifies the risk set)

Let $\delta_1, \delta_2 \in \mathcal{D}$ and let δ^* be the randomised rule which uses δ_1 with probability λ and δ_2 with probability $1-\lambda$, $0 \le \lambda \le 1$, the choice being made independently of $\mathbf{X}$. Then

$$R(\theta, \delta^*) = \lambda R(\theta, \delta_1) + (1-\lambda) R(\theta, \delta_2), \qquad \forall\, \theta \in \Theta.$$

Proof: randomisation convexifies the risk set

Proof.

Let U be a Bernoulli(λ) random variable independent of $\mathbf{X}$, and let δ^* use δ_1 if $U = 1$ and δ_2 if $U = 0$. Conditioning on U,

$$\begin{aligned} R(\theta, \delta^*) &= \mathbb{E}_\theta[L(\theta, \delta^*(\mathbf{X}, U))] \\ &= \mathbb{E}\{\mathbb{E}_\theta[L(\theta, \delta^*(\mathbf{X}, U)) \mid U]\} \\ &= \mathbb{P}(U = 1)\, \mathbb{E}_\theta[L(\theta, \delta_1(\mathbf{X}))] + \mathbb{P}(U = 0)\, \mathbb{E}_\theta[L(\theta, \delta_2(\mathbf{X}))] \\ &= \lambda R(\theta, \delta_1) + (1 - \lambda) R(\theta, \delta_2). \end{aligned}$$

□

Consequences of convexity

The proposition says that the set of attainable risk functions becomes *convex* once randomised rules are admitted. This has two important consequences.

(i) **Better minimax rules.** Mixing two rules can lower the maximum risk: if δ_1 is good where δ_2 is bad and vice versa, a suitable mixture balances the two.

- Hence the smallest attainable maximum risk over randomised rules can be strictly smaller than over non-randomised rules.
- This is illustrated in "Randomised Minimax Rules in Finite Problems" and in end-of-chapter Exercise 1.20, where the minimax rule is necessarily randomised.

(The game-theoretic background, the *minimax theorem*, is outlined in Lecture Notes, Appendix A.)

Consequences of convexity (cont.)

(ii) **Randomisation never helps under a convex loss.** If $\mathcal{A}$ is a convex subset of $\mathbb{R}$ and $L(\theta, \cdot)$ is convex for each θ, then by Jensen's inequality the non-randomised rule

$$\bar{\delta}(\mathbf{x}) = \int a \, \delta^*(da \mid \mathbf{x})$$

satisfies

$$R(\theta, \bar{\delta}) \leq R(\theta, \delta^*) \quad \text{for all } \theta.$$

Hence in estimation under squared error loss one may restrict attention to non-randomised rules without loss. This is the content of the Rao–Blackwell type argument and explains why randomisation is prominent in testing but invisible in estimation.

After this lecture

Practice (Lecture Notes, Chapter 1)

- Short-answer questions (Section 1.12): Q2, Q15.

Next lecture

Lecture 4: Uniformly best rules; admissibility; linear estimators of a normal mean.