Chapter 4: Statistical Decision Theory *

4 Statistical Decision Theory

4.1 Basic Framework and Concepts

- To the usual statistical modeling framework from earlier

$$X, \quad \Theta, \quad \mathcal{P} = \{P_\theta : \theta \in \Theta\}$$

we add the following elements

1. some "action space" $\mathcal{A}$ with σ-algebra ϵ,
2. a suitably measurable "loss function"

$$L(\theta, a) : \Theta \times \mathcal{A} \to [0, \infty),$$

3. and (non-randomized) decision rules

$$\delta(x) : (\mathcal{X}, \mathcal{B} \to (\mathcal{A}, \epsilon))$$

For data X, $\delta(x)$ is the action taken based on X.

To identify "good" devusuib rules δ, we have to average our X, which naturally leads to expectation.

- Risk function The mapping from $\Theta \to [0, \infty)$ given by

$$R(\theta, \delta) \equiv R_\theta L(\theta, \delta(X)) = \int_{\mathcal{X}} L(\theta, \delta(x)) dP_\theta(x)$$

is call the risk function for θ.

  - δ is *at least as good as* δ' if $R(\theta, \delta) \leq R(\theta, \delta')$ for all $\theta \in \Theta$
  - δ is *better than* δ' if $R(\theta, \delta) \leq R(\theta, \delta')$ for all $\theta \in \Theta$, and $R(\theta_0, \delta) < R(\theta_0, \delta')$ for some θ_0
  - δ and δ' are *risk equivalent* if $R(\theta, \delta) = R(\theta, \delta')$ for all $\theta \in \Theta$.
  - δ is *best in a class of decision rules* Δ if $\delta \in \Delta$, and δ is at least as good as any other $\delta' \in \Delta$
  - Example: $X \sim N(\theta, 1), \theta \in \mathbb{R}$ with $\Delta =$"the class of all estimators of θ". There is no best element here. Prove by proposing two constant estimators and zero-one loss.

- If there is no best estimator,

  - Try a smaller and appropriate Δ, e.g. unbiased estimators.
  - Reduce the risk function $R(\theta, \delta)$ to a number and compare numbers for different δ's, e.g.: averaging over θ according to some distribution G on Θ is a way to make "Bayes Risk" and look for "Bayes optimal " decision rules.
  - Maximize $R(\theta, \delta)$ over θ and seek to minimize over δ's, i.e. mini-max procedures.

- Inadmissible: δ is inadmissible in Δ if there exists $\delta' \in \Delta$ that is better than δ.

- Admissible: δ is admissible in Δ if it is not inadmissible in Δ.

  Note: One may never want to use an inadmissible rule, but there are decision problems where every rule is inadmissible.

- Behavorial decision rule: If for each $x \in \mathcal{X}, \phi_x$ is a distribution on $(\mathcal{A}, \epsilon)$, then ϕ_x is called a behavorial decision rule.

*STA643: Advanced Theory of Statistical Inference. Instructed by Dr. Daniel Nordman. Arranged by Zhiling Gu

- $\mathcal{D}^* \equiv \{\phi_x\} \equiv$ the class of behavorial decision rules
- $\mathcal{D} \subset \mathcal{D}^*$ where

$$\mathcal{D} \equiv \{\delta(x)\} \equiv \text{the class of non-randomized decision rules } \delta : \mathcal{X} \to \mathcal{A}$$

- The risk function of a behaorial decision rule is defined as

$$R(\theta, \phi) = \int_{\mathcal{X}} \int_{\mathcal{A}} L(\theta, a) d\phi_x(a) dP_\theta(x)$$