STAT 5010 Tutorial 4*

Nov. 2022

Definitions

1. Let $\Im$ be a class of decision rules (randomized or nonrandomized). A decision rule $T \in \Im$ is called *$\Im$-admissible* (or *admissible* when $\Im$ contains all possible rules) if and only if there does not exist any $S \in \Im$ that is better than T (in terms of the risk).

2. Now we consider an average of $R_T(P)$ over $P \in \mathcal{P}$:

$$r_T(\Pi) = \int_{\mathcal{P}} R_T(P) d\Pi(P),$$

where Π is a known probability measure on $(\mathcal{P}, \mathcal{F}_{\mathcal{P}})$ with an appropriate σ-field $\mathcal{F}_{\mathcal{P}}$. $r_T(\Pi)$ is called the *Bayes risk* of T w.r.t. Π. If $T_* \in \Im$ and $r_{T_*}(\Pi) \leq r_T(\Pi)$ for any $T \in \Im$, then T_* is called a *$\Im$-Bayes rule* (or *Bayes rule* when $\Im$ contains all possible rules) w.r.t. Π. The second method is to consider the worst situation, i.e., $\sup_{P \in \mathcal{P}} R_T(P)$. If $T_* \in \Im$ and $\sup_{P \in \mathcal{P}} R_{T_*}(P) \leq \sup_{P \in \mathcal{P}} R_T(P)$ for any $T \in \Im$, then T_* is called a *$\Im$-minimax* rule (or *minimax* rule when $\Im$ contains all possible rules).

3. To test the hypotheses H_0 versus H_1, there are only two types of statistical errors we may commit: rejecting H_0 when H_0 is true (called the *type I error*) and accepting H_0 when H_0 is wrong (called the *type II error*). In statistical inference, a test T, which is a statistic from $\mathcal{X}$ to $\{0, 1\}$, is assessed by the probabilities of making two types of errors:

$$\alpha_T(P) = P(T(X) = 1) \quad P \in \mathcal{P}_0$$

and

$$1 - \alpha_T(P) = P(T(X) = 0) \quad P \in \mathcal{P}_1,$$

which are denoted by $\alpha_T(\theta)$ and $1 - \alpha_T(\theta)$ if P is in a parametric family indexed by θ. Note that these are risks of T under the $0 - 1$ loss in statistical decision theory. However, an optimal decision rule (test) does not exist even for a very simple problem with a very simple class of tests (Example).

4. A common approach to finding an "optimal" test is to assign a small bound α to one of the error probabilities, say $\alpha_T(P), P \in \mathcal{P}_0$, and then to attempt to minimize the other error probability $1 - \alpha_T(P), P \in \mathcal{P}_1$, subject to

$$\sup_{P \in \mathcal{P}_0} \alpha_T(P) \leq \alpha.$$

The bound α is called the *level of significance*. The left-hand side above is called the *size* of the test T. Note that the level of significance should be positive, otherwise no test satisfies the inequality above except the silly test $T(X) \equiv 0$ a.s. $\mathcal{P}$.

5. For most tests satisfying the inequality above, a small α leads to a "small" rejection region. It is good practice to determine not only whether H_0 is rejected or accepted for a given α and a chosen test T_α, but also the smallest possible level of significance at which H_0 would be rejected for the computed $T_\alpha(x)$, i.e., $\hat{\alpha} = \inf\{\alpha \in (0,1) : T_\alpha(x) = 1\}$. Such an $\hat{\alpha}$, which depends on x and the chosen test and is a statistic, is called the *p-value* for the test T_α.

Propositions and Theorems

1. (TPE Lemma 5.1.13) If δ_Λ is the Bayes estimator of $g(\theta)$ with respect to Λ and if

$$r_\Lambda = E\left[\delta_\Lambda(\mathbf{X}) - g(\Theta)\right]^2$$

is its Bayes risk, then

$$r_\Lambda = \int \operatorname{var}[g(\Theta) \mid \mathbf{x}] dP(\mathbf{x}).$$

In particular, if the posterior variance of $g(\Theta) \mid \mathbf{x}$ is independent of $\mathbf{x}$, then

$$r_\Lambda = \operatorname{var}[g(\Theta) \mid \mathbf{x}].$$

*Dept. of Stat., CUHK. TA: YX

2. (TPE Theorem 5.1.12) Suppose that $\{\Lambda_n\}$ is a sequence of prior distributions with Bayes risks r_n satisfying $r_n \leq r = \lim_{n\to\infty} r_n$ and that δ is an estimator for which

$$\sup_{\theta} R(\theta, \delta) = r.$$

Then (i) δ is minimax and (ii) the sequence $\{\Lambda_n\}$ is least favorable.

3. For $X \sim N(\theta, I_p)$ $(p \geq 3)$ with θ unknown. Under loss $L(\theta, a) = ||a - \theta||^2 = \sum_{i=1}^{p} (a_i - \theta_i)^2$, X is a minimax estimator. But X is inadmissible and dominated by the James-Stein estimator $\delta_c = X - \frac{p-2}{||X-c||^2}(X - c)$

4. Following the previous proposition, the James-Stein estimator with any c is still inadmissible and dominated by $\delta_c^+ = X - min\{1, \frac{p-2}{||X-c||^2}\}(X - c)$.

Question 1

1. Let X be an observation from the Poisson distribution with unknown mean $\theta > 0$. Consider the estimation of θ under the squared error loss.

   (i) Show that $\sup_\theta R_T(\theta) = \infty$ for any estimator $T = T(X)$, where $R_T(\theta)$ is the risk of T.

   (ii) Let $\Im = \{aX + b : a \in \mathcal{R}, b \in \mathcal{R}\}$. Show that 0 is an admissible estimator of θ within $\Im$.

Solution:

(i) When the gamma distribution with shape parameter α and scale parameter γ is used as the prior for θ, the Bayes estimator is $\delta(X) = \gamma(X + \alpha)/(\gamma + 1)$ with Bayes risk $r_\delta = \alpha\gamma^2/(\gamma + 1)$. Then, for any estimator T,

$$\sup_{\theta>0} R_T(\theta) \geq r_\delta = \frac{\alpha\gamma^2}{\gamma + 1} \to \infty$$

as $\gamma \to \infty$.

(ii) The risk of 0 is θ^2. The risk of $aX + b$ is

$$a^2 \operatorname{Var}(X) + [aE(X) + b - \theta]^2 = (a - 1)^2\theta^2 + \left[2(a - 1)b + a^2\right]\theta + b^2.$$

If 0 is inadmissible, then there are a and b such that

$$\theta^2 \geq (a - 1)^2\theta^2 + \left[2(a - 1)b + a^2\right]\theta + b^2$$

for all $\theta > 0$. Letting $\theta \to 0$, we obtain that $b = 0$. Then

$$\theta \geq (a - 1)^2\theta + a^2$$

for all $\theta > 0$. Letting $\theta \to 0$ again, we conclude that $a = 0$. This shows that 0 is admissible within the class $\Im$.