Decision Theory: Part 2

Dr. Qiuzhuang Sun

STAT3023

Minimax procedures

- Recall that for a given subset $\Theta_0 \subseteq \Theta$, a decision rule $\tilde{d}(\cdot)$ is said to be minimax (over Θ_0) if

$$\max_{\theta \in \Theta_0} E_\theta[L(\tilde{d}(\mathbf{X})|\theta)] \leq \max_{\theta \in \Theta_0} E_\theta[L(d(\mathbf{X})|\theta)]$$

  for any other d.
- Such procedures can be harder to find than Bayes procedures. However, they have a certain appeal (especially when $\Theta_0 = \Theta$) since they do not require the choice of a weight function.
- Such procedures are viewed as pessimistic since they lead to the best worst-case scenario.

Finding minimax procedures

We now present two theorems (with examples; proofs provided in the advanced workshop) which show how Bayes procedures can be used as theoretical tools to find minimax procedures.

Theorem 1. Suppose that for $k = 1, 2, \ldots, d_k(\cdot)$ is the Bayes procedure with respect to a **"proper prior"** $w_k(\cdot)$ on Θ and loss function $L(d|\theta)$. If the procedure $\tilde{d}(\cdot)$ is such that

$$\max_{\theta \in \Theta_0} E_\theta[L(\tilde{d}(\mathbf{X})|\theta)] \le \lim_{k\to\infty} B_{w_k}(d_k)$$
$$= \lim_{k\to\infty} \int_\Theta E_\theta[L(d_k(\mathbf{X})|\theta)] w_k(\theta) d\theta,$$

then $\tilde{d}$ is minimax over Θ under the loss function $L(\cdot|\theta)$.

Finding minimax procedures

Example: Suppose $X_1, \ldots, X_n$ are iid $N(\theta, 1)$. The decision space is $\mathcal{D} = \mathbb{R}$, and the loss is $(d - \theta)^2$. Write down the risks of $\bar{X}$ and $d_k(\cdot)$, the Bayes procedure using $N(\mu_0, k)$ as the weight function/prior. Hence using Theorem 1, show that $\bar{X}$ is minimax over $\mathbb{R}$.

Finding minimax procedures

(Example continued)

Finding minimax procedures

(Example continued)

Finding minimax procedures

Theorem 2. Suppose $d(\cdot)$ is a Bayes Procedure with respect to a loss function $L(\cdot|\theta)$ and a **proper prior** $w(\cdot)$ on Θ. Let Θ_0 denote the support of $w(\cdot)$. If

(i) $E_\theta[L(d(\mathbf{X})|\theta)] = c$ (constant) for all $\theta \in \Theta_0$;

(ii) $E_\theta[L(d(\mathbf{X})|\theta)] \leq c$ for all $\theta \in \Theta$,

then $d(\cdot)$ is minimax over Θ for $L(\cdot|\theta)$.

Basically, if a Bayes procedure has constant risk, it is minimax.

Finding minimax procedures

Example: Suppose $X_1, \ldots, X_n$ are independent with $X_i \sim \text{Bin}(m_i, \theta)$ with known $m_1, \ldots, m_n$. Let $M = \sum_{i=1}^n m_i$. Based on the lecture notes last week, the Bayes procedure under the squared-error loss using a Beta(α, β) density as prior is

$$d(\mathbf{X}) = \frac{\alpha + \sum_{i=1}^n X_i}{\alpha + \beta + M}.$$

Based on this Bayes procedure, using Theorem 2 to find minimax procedures over $[0, 1]$.

Finding minimax procedures

(Example continued)

Finding minimax procedures

(Example continued)

Discussion of minimax procedures

- The two minimax estimators derived in the two examples above are unfortunately quite rare; not many “globally” (over the whole parameter space) minimax procedures such as these can be explicitly determined.

- We shall see through the following example that there are cases where minimax procedures are not desirable.

Discussion of minimax procedures

Consider the minimax procedure in the binomial example above:

$$\tilde{d}(\mathbf{X}) = \frac{\sum_{i=1}^{n} X_i + \frac{\sqrt{M}}{2}}{M + \sqrt{M}}$$

(also the Bayes procedure corresponding to prior Beta$(\frac{\sqrt{M}}{2}, \frac{\sqrt{M}}{2})$).

Discussion of minimax procedures

Discussion of minimax procedures

Discussion of minimax procedures

Asymptotically minimax procedures

In general, can we find $d(\mathbf{X})$ which minimises

$$\lim_{n\to\infty} \max_{a\le\theta\le b} R_n(\theta|d)$$

for all decisions $d(\cdot)$?

Such procedures are called asymptotically minimax.

An asymptotic minimax procedure can be found with 2 steps.

(i) Determine a lower bound on

$$\lim_{n\to\infty} \max_{a\le\theta\le b} R_n(\theta|d)$$

for any procedure $d(\cdot)$.

(ii) Show that a given procedure attains the lower bound.

Asymptotically minimax procedures

Example (**interval estimation of mean of a Poisson**): Suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ are iid Poisson(θ) for some unknown $\theta \in \Theta = (0, \infty)$. The decision space is $\mathcal{D} = (0, \infty)$; the loss function is

$$L(d|\theta) = 1\{|d - \theta| > c/\sqrt{n}\}$$

for some $c > 0$. Find asymptotically minimax procedures over $[a, b]$.

(Example continued)

(Example continued)

Asymptotically minimax procedures

(Example continued)

Asymptotically minimax procedures

Recall that $d(\cdot)$ is asymptotically minimax if it minimises

$$\lim_{n\to\infty} \max_{\theta\in[a,b]} R_n(\theta|d)$$

- Sometimes R_n is rescaled
- Why maximum risk over an interval?

Asymptotic maximum risk over an interval

Example: Suppose $\mathbf{X} = (X_1, \ldots, X_n)$ consists of iid $N(\theta, 1)$ and we want to estimate θ using the squared-error loss. Consider the following 2 estimators:

- $d_1(\mathbf{X}) = \bar{X}$;
- $d_2(\mathbf{X}) = \bar{X} \cdot 1\{|\bar{X}| > n^{-\frac{1}{4}}\} = \begin{cases} \bar{X}, & |\bar{X}| > n^{-\frac{1}{4}} \\ 0, & \text{otherwise.} \end{cases}$

In words, d_2 favors the special point $\theta = 0$, if the sample mean is near 0 (i.e., $[-n^{1/4}, n^{1/4}]$). We would like to compare the two estimators under different risks.

Asymptotic maximum risk over an interval

(Example continued)

Asymptotic maximum risk over an interval

(Example continued)

Asymptotic maximum risk over an interval

(Example continued)

Asymptotic maximum risk over an interval

Lessons from the previous example:

- The limiting pointwise (rescaled) risk on its own may not be the best measure when comparing estimates. The limiting maximum (rescaled) risk over an interval is a better idea.

- We cannot always swap "$\lim$" and "$\max$".

Asymptotic minimax lower bound

The following theorem uses the (pointwise) limiting (rescaled) risk of certain Bayes procedures to provide a lower bound to the limiting **maximum** (rescaled) risk for any estimator.

Theorem: Suppose that for sequence $\{L_n(\cdot|\theta)\}$ of loss functions and any $\theta_0 < \theta_1$, the corresponding sequence of Bayes procedures $\{\tilde{d}_n(\cdot)\}$ based on the Unif(θ_0, θ_1) prior $w(\theta) = (\theta_1 - \theta_0)^{-1} 1\{\theta_0 < \theta < \theta_1\}$ is such that for each $\theta_0 < \theta < \theta_1$,

$$\lim_{n\to\infty} E_\theta[L_n(\tilde{d}_n(\mathbf{X})|\theta)] = S(\theta),$$

where $S(\cdot)$ is a continuous function. Then, for any other sequence of procedures $\{d_n(\cdot)\}$ and any $a < b$,

$$\lim_{n\to\infty} \max_{a\le\theta\le b} E_\theta[L_n(d_n(\mathbf{X})|\theta)] \geq \max_{a\le\theta\le b} S(\theta).$$

Asymptotic minimax lower bound

Example (**interval estimation of mean of a Poisson continued**): Suppose that $\mathbf{X} = (X_1, \ldots, X_n)$ are iid Poisson(θ) for some unknown $\theta \in \Theta = (0, \infty)$. The decision space is $\mathcal{D} = (0, \infty)$; the loss function is $L(d|\theta) = 1\{|d - \theta| > c/\sqrt{n}\}$ for some $c > 0$. Find asymptotically minimax procedures over $[a, b]$.

Recall we have claimed that when $d(\mathbf{X}) = \bar{X}$ (i.e., the interval estimator is $\bar{X} \pm \frac{c}{\sqrt{n}}$),

$$\lim_{n\to\infty} \max_{a\leq\theta\leq b} R_n(\theta|\bar{X}) \leq 2\left(1 - \Phi\left(\frac{c}{\sqrt{b}}\right)\right).$$

We now show that (i) the above is an equality; (ii) it is indeed a lower bound on $\lim_{n\to\infty} \max_{a\leq\theta\leq b} R_n(\theta|d)$, and hence show $d(\mathbf{X}) = \bar{X}$ is asymptotically minimax over $[a, b]$.

Asymptotic minimax lower bound

(Example continued)

Asymptotic minimax lower bound

(Example continued)

Asymptotic minimax lower bound

(Example continued)

Asymptotic minimax lower bound

(Example continued)