Asymptotic Statistics

Lecture 10

Minimax lower bounds

Goal

Find lower bound for the minimax risk

$$\inf_{\hat{\theta}} \sup_{\theta \in \Theta} E_\theta d^p(\hat{\theta}, \theta).$$

Gives a lower bound for the best possible rate that any estimator can achieve, uniformly over Θ.

Assouad's lemma

Let $X \sim p_\theta$, $\theta \in \{0,1\}^r$.

Theorem.
We have

$$\inf_{\hat{\theta}} \max_{\theta \in \Theta} E_\theta d_{ham}(\hat{\theta}, \theta) \geq \frac{r}{2} \min_{d_{ham}(\theta_0,\theta_1)=1} \int (p_{\theta_0} \wedge p_{\theta_1}),$$

where the infimum is over all estimators $\hat{\theta} = \hat{\theta}(X)$.

I.i.d. data

Let $X_1, \ldots, X_n$ be a sample from p_θ, $\theta \in \{0,1\}^r$.

Corollary.

We have

$$\inf_{\hat{\theta}} \max_{\theta \in \Theta} E_\theta d_{ham}(\hat{\theta}, \theta) \geq \frac{r}{4} \min_{d_{ham}(\theta_0, \theta_1)=1} \left(1 - \frac{1}{2} \int (\sqrt{p}_{\theta_0} - \sqrt{p}_{\theta_1})^2\right)^{2n},$$

where the infimum is over all estimators $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$.

Useful lemma

Let $X \sim p_\theta$, $\theta \in \{0,1\}^r$. Let $g : \Theta \to A$, where A is a space with pseudometric d. Suppose that for $p, C > 0$,

$$d^p(g(\theta_1), g(\theta_2)) \geq C d_{ham}(\theta_1, \theta_2)$$

for all $\theta_1, \theta_2 \in \Theta$.

Lemma.

$$\inf_T \max_{\theta \in \Theta} E_\theta d^p(T, g(\theta)) \geq \frac{C}{2^p} \inf_{\hat{\theta}} \max_{\theta \in \Theta} E_\theta d_{ham}(\hat{\theta}, \theta).$$

Lower bound for smooth parametric models

Smooth parametric models

Let $X_1, \ldots, X_n$ be a sample from a density p_θ, with $\theta \in \Theta \subset \mathbb{R}^k$. Suppose Θ contains at least one interior point and that the model is smooth in the sense that for some constant $C > 0$,

$$\int (\sqrt{p}_\theta - \sqrt{p}_\psi)^2 \leq C^2 \|\theta - \psi\|^2$$

for all $\theta, \psi \in \Theta$ close enough.

Lower bound

Theorem.
there exists a constant $K > 0$ such that

$$\inf_{\hat{\theta}} \sup_{\theta \in \Theta} P_\theta(\|\hat{\theta} - \theta_0\| > 1/\sqrt{n}) \geq K,$$

and for all $p > 0$,

$$\inf_{\hat{\theta}} \sup_{\theta \in \Theta} E_\theta \|\hat{\theta} - \theta_0\|^p \geq K n^{-p/2},$$

where the infima are over all estimators $\hat{\theta} = \hat{\theta}(X_1, \ldots, X_n)$.

Lower bound for density estimation

Lower bound

Let $\mathcal{F}_{m.M}$ be the class of densities f that are m times continuously differentiable and for which $\int |f^{(m)}(x)|^2 \, dx \leq M$.

Theorem.

There exists a constant $C_{m,M} > 0$ such that

$$\inf_{\hat{f}} \sup_{f \in \mathcal{F}_{m,M}} \int (\hat{f}(x) - f(x))^2 \, dx \geq C_{m,M} \, n^{-2m/(1+2m)},$$

where the infimum is over all density estimators $\hat{f} = \hat{f}(X_1, \ldots, X_n)$.