STA 3000 Lecture 16.

$$\Pi\left(B(\theta^*,\varepsilon) \mid X_1^n\right) \xrightarrow{P} 1 \quad \forall \varepsilon > 0.$$

Thm (Bernstein-von-Mises).

Assume QMD, $I(\theta^*) \succ 0$, stoch Lip.

$\exists$ sequence of tests ϕ_n

$$\boxed{\begin{aligned} \text{s.t. } & \mathbb{E}_{\theta^*}[\phi_n(X)] \to 0 \\ & \sup_{\|\theta-\theta^*\|\geq\varepsilon} \mathbb{E}_\theta[1-\phi_n(X)] \to 0. \end{aligned}} \quad (\ast)$$

Then, for $\theta \sim \Pi_n(\cdot \mid X_1^n)$.

$$d_{TV}\left(\sqrt{n}(\theta-\theta^*) \mid X_1^n, \ N\left(\Delta_n, I(\theta^*)^{-1}\right)\right) \xrightarrow{P} 0$$

$$\Longleftrightarrow \ d_{TV}\left(\Pi(\cdot \mid X_1^n), \ N\left(\theta^* + \frac{\Delta_n}{\sqrt{n}}, \frac{I(\theta^*)^{-1}}{n}\right)\right) \xrightarrow{P} 0$$

$$\Delta_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n I(\theta^*)^{-1} \dot{l}_{\theta^*}(X_i)$$

$\nabla \log p_{\theta^*}(X_i)$.

Proof Idea:

$$\tilde{\Pi}_n(h) = \Pi\left(\theta^* + \frac{h}{\sqrt{n}} \mid X_1^n\right)$$

$$\log \tilde{\pi}_n(h) = \log \pi\left(\theta^* + \frac{h}{\sqrt{n}}\right) + \sum_{i=1}^n \log p_{\theta^* + \frac{h}{\sqrt{n}}}(X_i) + C$$

$$\approx \log \pi(\theta^*) + \sum_{i=1}^n \log p_{\theta^*}(X_i) + \frac{h^T}{\sqrt{n}} \sum_{i=1}^n \nabla \log p_{\theta^*}(X_i)$$

$$+ \frac{1}{2} h^T \underbrace{\frac{1}{n}\sum_{i=1}^n \nabla^2 \log p_{\theta^*}(X_i)}_{\downarrow \; -I(\theta^*)} h + C$$

$$\log \tilde{\pi}_n(h) \approx C' - \frac{1}{2} h^T I(\theta^*) h + \frac{h^T}{\sqrt{n}} \sum_1^n \nabla \log p_{\theta^*}(X_i)$$

Local Asymptotic Minimax.

Corollary : L loss function. satisfying
- Symmetric. $L(-x) = L(x)$
- Quasi-convex $\{x : L(x) \le a\}$ is convex $\forall a$.

$$\lim_{c \to +\infty} \limsup_{n \to +\infty} \inf_{\hat{\theta}_n} \sup_{\|\theta - \theta^*\|_2 \le \frac{c}{\sqrt{n}}} \mathbb{E}\, L\left(\sqrt{n}\left(\hat{\theta}_n - \theta^*\right)\right)$$

$$\ge \mathbb{E}\left[L(Z)\right]$$

$$Z \sim N\left(0, I(\theta^*)^{-1}\right).$$

$$\boxed{\text{Lemma} \quad \underset{h \in \mathbb{R}^d}{\text{minimize}} \; \mathbb{E}\left[L(h + Z)\right] \qquad h^* = 0.}$$

Nonparametrics.

1. Density $\quad X_1 \cdots X_n \overset{iid}{\sim} p^* \in \mathcal{P}$.

2. Nonpara regression. $(X_i)_{i=1}^n$ deterministic / random.

$$Y_i = f^*(X_i) + \varepsilon_i \qquad \mathbb{E}[\varepsilon_i | X_i] = 0$$

$$\varepsilon_i \text{ indep} \qquad f^* \in \mathcal{F}.$$

Popular classes on K. compact subset of $\mathbb{R}^k$

— Hölder class. $\quad \Sigma(\beta, L) \quad \begin{matrix} \beta > 0 \\ L > 0 \end{matrix}$

$\beta \in \mathbb{N}$, $\Sigma(\beta, L) := \left\{ f: K \to \mathbb{R} \;\middle|\; |\partial^m f(x)| \le L, \ \forall\, m \le \beta,\ x \in K \right\}$

$$\left|\partial_{x_1}^{m_1} \partial_{x_2}^{m_2} \cdots \partial_{x_k}^{m_k} f(x)\right| \le L \qquad \|m\|_1 \le \beta.$$

$$\beta = b + \gamma \qquad \begin{matrix} b \in \mathbb{N} \\ \gamma \in [0,1). \end{matrix}$$

$$\Sigma(\beta, L) := \left\{ f: K \to \mathbb{R} \;\middle|\; \begin{matrix} \forall\, x, y \quad \forall\, \|m\|_1 \le b \\ |\partial^m f(x) - \partial^m f(y)| \le L \|x-y\|_2^\gamma \end{matrix} \right\}$$

Goal: $\displaystyle \inf_{\hat{f}_n} \sup_{f^* \in \Sigma(\beta, L)} \mathbb{E}\left[\int_K |f^*(x) - \hat{f}_n(x)|^2\, dx\right].$

or $\mathbb{E}\left[|f^*(x_0) - \hat{f}_n(x_0)|^2\right]$

— Sobolev

(integer β)

$$S(\beta, L) := \left\{ f : \begin{array}{l} \forall m \quad \|m\|_1 \le \beta \\ \int_K |\partial^m f(x)|^2 \, dx \le L^2 \end{array} \right\}$$

— $K \subseteq \mathbb{R}$.

$$\mathcal{F}_{\text{increasing}} = \{ f : f(x) \le f(y) \quad \forall x \le y \}.$$