Lecture 23. M,Z cont

Sunday, December 3, 2017 10:57 PM

$\left.\begin{array}{l} M_n \\ \\ \Psi_n \end{array}\right\}$ random functions that depend on $(\underbrace{\{X_i\}}_{\text{Sample}} ; \underbrace{\theta}_{\text{parameter}})$

$\max M_n(\theta) \Rightarrow \hat{\theta}_n$

solve $\Psi_n(\theta) = 0 \Uparrow$

Consistency $\hat{\theta}_n \xrightarrow[p]{} \theta$

TH uniform convergence of Ψ_n or M_n to their large-sample limits $\underbrace{\Psi, M}_{\text{deterministic}}$

Separation of the root

Approximate maximizer (M_n), approx asymptotic zero (Ψ_n)
asymptotic

$\Downarrow$ Consistency

TH relaxes the unif conv. requirement

- pointwise conv.

$$\Psi_n(\theta) \xrightarrow[p]{} \Psi(\theta)$$

- or
  - (a) Ψ_n is continuous, has unique zero $\hat{\theta}_n$
  - (b) Ψ_n is non-decreasing, and $\hat{\theta}_n$ is its asymptotic zero

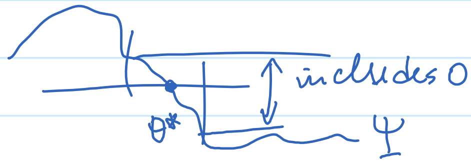
$$\Psi_n(\hat{\theta}_n) = o_p(1)$$

$$\left(\Psi_n(\hat{\theta}_n) \xrightarrow[p]{} 0\right)$$

- θ^* that's the root of Ψ is such that

$$0 \in [\Psi(\theta^* \pm \varepsilon)], \quad \forall \varepsilon > 0$$

$$\searrow [\Psi(\theta^* - \varepsilon), \Psi(\theta^* + \varepsilon)]$$



Then $\hat{\theta}_n \xrightarrow[p]{} \theta$.

Proof.

Let's do (a)-version 1st

$$\{0 \in [\Psi_n(\theta^* \pm \varepsilon)]\} \Rightarrow \hat{\theta}_n \in [\theta^* \pm \varepsilon] \qquad (1)$$

$$\underset{\Psi(\theta^*)}{\overset{\shortparallel}{0}}$$

Ψ_n is cont. wrt θ; it has unique root $\hat{\theta}_n$

$\Rightarrow \{\ldots\}$ means Ψ_n achieves its zero in $[\theta^* \pm \varepsilon]$

Ψ_n is cont. $\Rightarrow [\Psi_n(\theta^* \pm \varepsilon)] = [\Psi_n(\theta^*) \pm \tilde{\varepsilon}(\varepsilon)]$ (meaning $\begin{matrix} + \tilde{\varepsilon}_1(\varepsilon) \\ - \tilde{\varepsilon}_2(\varepsilon) \end{matrix}$)

$$\tilde{\varepsilon} \xrightarrow[\varepsilon \to 0]{} 0$$

$\Downarrow$

$$0 \in [\Psi_n(\theta^* \pm \varepsilon)] \iff \Psi_n(\theta^*) - \tilde{\varepsilon}_1 \le 0 \le \Psi_n(\theta^*) + \tilde{\varepsilon}_2 \iff$$

$$-\tilde{\varepsilon}_1 \le -\Psi_n(\theta^*) + \underbrace{\Psi(\theta^*)}_{=0} \le \tilde{\varepsilon}_2$$

$$P(-\tilde{\varepsilon}_1 \le \Psi(\theta^*) - \Psi_n(\theta^*) \le \tilde{\varepsilon}_2) \xrightarrow[p]{} 1$$

b/c $\Psi_n \xrightarrow[p]{} \Psi$ $\forall$ fixed θ

$\Downarrow$

$$P\{0 \in [\Psi_n(\theta^* \pm \varepsilon)]\} \xrightarrow[n \to \infty]{} 1$$

Note: Ψ_n has to be monotonic in some small area

Note: Ψ_n has to be monotonic in some small area around θ^* $(\theta^* \pm \varepsilon)$ eventually for ε small enough

$\Uparrow$

continuity of Ψ_n & unique zero

$(1) \Rightarrow$

$$\underset{\substack{P\downarrow \\ 1}}{P}\left\{0 \in [\Psi_n(\theta^* \pm \varepsilon)]\right\} \leq P\left\{\hat{\theta}_n \in [\theta^* \pm \varepsilon]\right\}$$

$$\Downarrow$$

$$P\left(\theta^* - \varepsilon \leq \hat{\theta}_n \leq \theta^* + \varepsilon\right) \xrightarrow[n\to\infty]{} 1$$

$$P\left(|\hat{\theta}_n - \theta^*| \leq \varepsilon\right) \to 1 \Rightarrow$$

$$\hat{\theta}_n \xrightarrow[P]{} \theta^*$$

Let's prove (b)

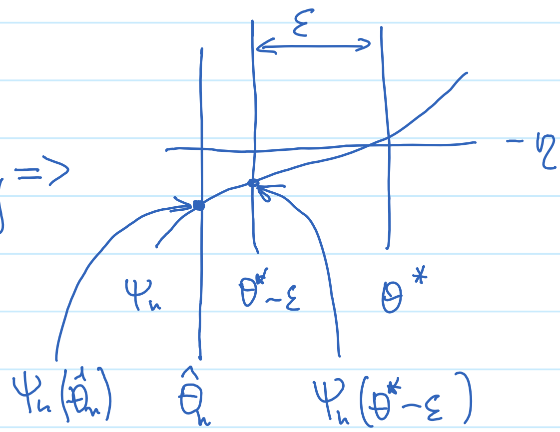
Note: If $\hat{\theta}_n$ is an exact zero of Ψ_n then proof is similar (a)

We need to carefully work around the fact that $\hat{\theta}_n$ is an asymptotic zero.

Take some $\eta > 0$

Consider event

$$\left.\begin{array}{l} \Psi_n(\theta^* - \varepsilon) < -\eta \\ \text{and} \\ \hat{\theta}_n \leq \theta^* - \varepsilon \end{array}\right\} \Rightarrow$$



$$\Rightarrow \Psi_n(\hat{\theta}_n) < -\eta$$

Similar argument gives

$$\Psi_n(\theta^* + \varepsilon) > \eta$$

$$\left.\begin{array}{l}\Psi_n(\theta^*+\varepsilon) > \eta \\ \hat{\theta}_n \geq \theta^*+\varepsilon\end{array}\right\} \Rightarrow \Psi_n(\hat{\theta}_n) > \eta$$

$$P\Big(\overbrace{|\Psi_n(\theta^*\pm\varepsilon)|>\eta}^{A_n} \cap \overbrace{|\hat{\theta}_n-\theta^*|\geq\varepsilon}^{B_n}\Big) \leq P\left(|\Psi_n(\hat{\theta}_n)|>\eta\right) \to 0$$

b/c $\Psi_n(\hat{\theta}_n)=o_p(1)$

$$\Rightarrow P(A_n B_n) \to 0$$

$|\Psi_n(\theta^*\pm\varepsilon)| \xrightarrow[p]{} C>0,\ \forall\varepsilon$ b/c θ^* is unique root of Ψ and Ψ_n is non-decr.

$$\Rightarrow \text{for } \eta<C,\ P(A_n)\to 1 \Rightarrow P(\bar{A}_n)\to 0$$

$$P(B_n) = P(A_nB_n)+P(\bar{A}_nB_n) \leq \underbrace{P(A_nB_n)}_{\to 0} + \underbrace{P(\bar{A}_n)}_{\to 0} \Rightarrow$$

$$\Rightarrow P(|\hat{\theta}_n-\theta^*|\geq\varepsilon) = P(B_n) \to 0 \Rightarrow$$

$$\Rightarrow \hat{\theta}_n \xrightarrow[p]{} \theta^* \qquad \square$$

Example: median

Let $\hat{\theta}_n$ be the sample median

$$\Psi_n(\theta) = \frac{1}{n}\sum_{i=1}^{n} \text{sign}(\theta - X_i) \xrightarrow[n\to\infty]{} \Psi(\theta)$$

$$\Psi_n(\theta) \xrightarrow[p]{} \mathbb{E}\,\text{sign}(\theta-X) \quad \text{by LLN}$$

$$\text{sign}(x) = \mathbb{1}(x) - \mathbb{1}(-x) \Rightarrow$$

$$\mathbb{E}\,\text{sign}(\theta-X) = P(X<\theta) - P(X>\theta) := \Psi(\theta)$$

$\hat{\theta}_n$ is a solution to $\Psi_n(\theta)=0$ by construction

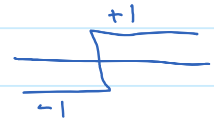
θ^* to $\Psi(\theta)=0 \Leftrightarrow$

$$P(X>\theta) = P(X<\theta)$$

Ψ_n is not continuous $\Rightarrow$ can't use (a)

$\Rightarrow$ using (b), Ψ_n is non-decreasing

$\Psi_n(\hat{\theta}_n) = 0$ by def. of median above



] θ^* is unique (making an assumption here) $\Rightarrow$

$$\Rightarrow 0 = \Psi(\theta^*) \in [\Psi(\theta^* \mp \varepsilon)]$$

$\Rightarrow$ by the previous Th $\Rightarrow$

median is a consistent estimator

□

Normality of M and Z estimators

Consider Z-estimators

$$] \quad \Psi_n(\theta) = \mathbb{P}_n \psi_\theta = \frac{1}{n}\sum_{i=1}^{n} \psi_\theta(X_i)$$

$$\Psi(\theta) = P\psi_\theta = \mathbb{E}\{\psi_\theta(X)\}$$

] $\hat{\theta}_n$ is a consistent Z-estimator

$$\hat{\theta}_n \xrightarrow[p]{} \theta^*$$

$$\Psi_n(\hat{\theta}_n) = 0, \quad \Psi(\theta^*) = 0 = P\psi_\theta$$

BTW: $\sqrt{n}(\mathbb{P}_n - P)f$ is an empirical process acting on f

Taylor series

$$0 = \Psi_n(\hat{\theta}_n) = \Psi_n(\theta^*) + \Psi_n'(\theta^*)\cdot(\hat{\theta}_n - \theta^*) + \frac{1}{2}\Psi_n''(\tilde{\theta})(\hat{\theta}_n - \theta^*)^2$$

$\tilde{\theta}$ is between θ^* and $\hat{\theta}_n$

$$\tilde{\theta} = \theta^* + t(\hat{\theta}_n - \theta^*), \quad t \in [0,1]$$

$$\Rightarrow \sqrt{n}(\hat{\theta}_n - \theta^*) = \frac{-\sqrt{n}\,\Psi_n(\theta^*)}{\Psi_n'(\theta^*) + \frac{1}{2}\Psi_n''(\tilde{\theta})(\hat{\theta}_n - \theta^*)}$$

$$P\psi_\theta^2 < \infty \Rightarrow \text{CLT} \quad -\sqrt{n}\,\Psi_n(\theta^*) \rightsquigarrow \mathcal{N}(P\psi_{\theta^*}, P\psi_{\theta^*}^2)$$

$P\psi_\theta^2 < \infty \Rightarrow$ CLT $- \sqrt{n}\,\psi_n(\theta^*) \rightsquigarrow \mathbb{N}(\underbrace{P\psi_{\theta^*}}_{\overset{\shortparallel}{0}}, P\psi_{\theta^*}^2)$

$\psi_n' = \mathbb{P}_n \psi_\theta' \xrightarrow[p]{} P\psi_\theta'$ by LLN

$\psi_n'' = \mathbb{P}_n \psi_{\tilde\theta}'' \xrightarrow[p]{} P\psi_{\tilde\theta}''$ by some LLN (informal statement)

Note: $\tilde\theta = \tilde\theta(X_1, \dots, X_n) \Rightarrow \psi_{\tilde\theta}''(X_i)$ are not i.i.d.

By Slutsky

$$\sqrt{n}\,(\hat\theta_n - \theta^*) \rightsquigarrow \mathbb{N}\left(0, \frac{P\psi_{\theta^*}^2}{(P\psi_{\theta^*}')^2}\right)$$

In the multivariate form

$$\sqrt{n}\,(\hat\theta_n - \theta^*) \rightsquigarrow \mathbb{N}\left(0, [P\psi_{\theta^*}']^{-1} P\psi_{\theta^*}\psi_{\theta^*}^T [P\psi_{\theta^*}'^T]^{-1}\right)$$