PSet 1 posted: Due Mon 13, November 2:30pm

PSet 2 (Empirical) will be posted Friday, 10 Nov, due Monday 20 Nov.

No school Friday, 10 Nov.

Uniform Laws of Large Numbers

Lemma: Suppose $\hat{Q}(\theta) = E_n[q(z_i, \theta)]$, and suppose $z_1, \dots, z_n$ is stationary and mixing and

i) $q(z, \theta)$ is continuous at θ with probability one

ii) Θ is compact

iii) $E[\sup_{\theta \in \Theta} |q(z, \theta)|] < +\infty$.

Then $Q(\theta) = E[q(z, \theta)]$ is continuous on Θ and

$$\sup_{\theta \in \Theta} |\hat{Q}(\theta) - Q(\theta)| \xrightarrow{P} 0$$

MLE Consistency

Thm: If z_i is iid with pdf $f(z|\theta)$ and

i) $f(z|\theta) \neq f(z|\theta_0)$ $\forall \theta \neq \theta_0$ (i.e. $\Pr[f(z|\theta) \neq f(z|\theta_0) | \theta_0] > 0$)

ii) Θ is compact

iii) $f(z|\theta)$ is continuous at all θ with prob. 1.

iv) $E[\sup_{\theta \in \Theta} |\ln f(z|\theta)|] < +\infty$

Then $\hat{\theta} \xrightarrow{P} \theta_0$.

Lemma (Information inequality): If $\int |\ln f(z|\theta)| f(z|\theta_0) dz < +\infty$, for each $\theta \in \Theta$, then if $f(z|\theta) \neq f(z|\theta_0)$, we have

$$E_{\theta_0}[\ln f(Z|\theta)] < E_{\theta_0}[\ln f(Z|\theta_0)]$$

Pf: By Jensen's inequality,

$$E_{\theta_0}\left[\ln\left(\frac{f(z|\theta)}{f(z|\theta_0)}\right)\right] < \ln\left[E_{\theta_0}\left[\frac{f(z|\theta)}{f(z|\theta_0)}\right]\right]$$

$$= \ln\left[\int \frac{f(z|\theta)}{f(z|\theta_0)} f(z|\theta_0) dz\right]$$

$$= \ln \int f(z|\theta) dz = \ln 1 = 0$$

$$\Rightarrow E[\ln f(z|\theta)] < E[\ln f(z|\theta_0)]$$

Pf of theorem: We need only check that the conditions of theorem 1 hold.

i) Let $Q(\theta) = E[\ln f(z|\theta)]$. Then
$$Q(\theta) - Q(\theta_0) = E\left[\ln \frac{f(z|\theta)}{f(z|\theta_0)}\right] < 0$$

ii) holds by compactness

iii) and iv) hold by ULLN.

When can we drop compactness?

Theorem: If $\hat{Q}(\theta)$ is convex and i) $Q(\theta)$ is continuous and uniquely minimized at θ_0, ii) $\Theta \subseteq \mathbb{R}^k$ is convex, and iii) $\hat{Q}(\theta) \xrightarrow{p} Q(\theta)$ for each $\theta \in \Theta$.

Then, $\hat{\theta} \xrightarrow{p} \theta_0$.

Probit: The likelihood function for a single observation under probit is

$$f(z|\theta) = \Phi(x'\theta)^y [1 - \Phi(x'\theta)]^{1-y}$$, where

$\Phi(v)$ is the standard normal cdf

Thm: If $E[xx']$ exists and is nonsingular, then the probit MLE $\hat{\theta}$ satisfies $\hat{\theta} \xrightarrow{p} \theta_0$

Pf: Fairly technical.

Consistency of GMM.

Our identification condition requires that

$E[g(z,\theta)] = 0$ iff $\theta = \theta_0$.

Thm: If z_i is iid and

i) $E[g(z,\theta)] = 0$ iff $\theta = \theta_0$ and $\hat{A} \xrightarrow{p} A$ with A positive definite

ii) Θ is compact

iii) $g(z,\theta)$ is continuous at each θ with probability one

iv) $E[\sup_{\theta \in \Theta} \|g(z,\theta)\|] < +\infty$.

Then $\hat{\theta} \xrightarrow{p} \theta$

Here, iii) and iv) guarantee that $\sup_{\theta \in \Theta} |\hat{Q}(\theta) - Q(\theta)| \xrightarrow{p} 0$,
where $\hat{Q}(\theta) = E_n[g(Z,\theta)]' \hat{A} E_n[g(Z,\theta)]$ and
$Q(\theta) = E[g(Z,\theta)]' A E[g(Z,\theta)]$, by ULLN.

Extremum Estimators

- Define $\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} Q_n(\theta)$
- assume conditions of thm 1 hold. Then $\hat{\theta} \xrightarrow{p} \theta_0$
  where $\theta_0 = \underset{\theta \in \Theta}{\operatorname{argmin}} Q(\theta)$

If Q_n is smooth and $\hat{\theta} \in \operatorname{int} \Theta$, then $\hat{\theta}$ solves
$\nabla Q_n(\hat{\theta}) = 0$ with probability one. By Taylor's theorem, (around θ_0):

θ^* differs from eqn to eqn if we have multiple eqns.

$$0 = \nabla Q_n(\theta_0) + \nabla^2 Q_n(\theta^*)(\hat{\theta} - \theta_0), \quad \theta^* \in \operatorname{conv}(\theta_0, \hat{\theta}),$$

which implies $\sqrt{n}(\hat{\theta} - \theta_0) = -[\nabla^2 Q_n(\theta^*)]^{-1} \sqrt{n} \nabla Q_n(\theta_0)$

- Need to assume $\nabla^2 Q_n(\theta^*) \xrightarrow{p} G$ (need some LLN)
- also, need some CLT to get $\sqrt{n} \nabla Q_n(\theta_0) \xrightarrow{d} N(0, \Omega)$

Then $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{d} -G^{-1} N(0, \Omega) \overset{d}{=} N(0, G^{-1} \Omega G^{-1})$
by Slutsky's theorem.

Thm: If $\hat{\theta} \xrightarrow{p} \theta_0$ and

i) $\theta_0 \in \operatorname{int}(\Theta)$,

ii) $Q_n(\theta)$ is twice continuously differentiable in some nbhd N around θ_0,

iii) $\sqrt{n}\, \nabla Q_n(\theta_0) \xrightarrow{d} N(0, \Omega)$

iv) $\exists\, G(\theta)$ that is continuous at θ_0 and

$\sup_{\theta \in N} \| \nabla^2 Q_n(\theta) - G(\theta) \| \xrightarrow{p} 0$, and

v) $G = G(\theta_0)$ is nonsingular.

Then $\sqrt{n}\,(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, G^{-1} \Omega G^{-1})$

Example: MLE with iid data:

$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmax}}\ E_n[\ln f(z, \theta)]$. Denote $l_i(\theta) \equiv \nabla \ln f(z_i, \theta)$

$= \dfrac{\nabla f(z_i, \theta)}{f(z_i, \theta)}$. Then $E_n[l_i(\hat{\theta})] = 0$, so

$0 = E_n[l_i(\theta_0)] + E_n[\nabla l_i(\theta^*)](\hat{\theta} - \theta)$

$\Rightarrow \sqrt{n}\,(\hat{\theta} - \theta_0) = -[E_n[\nabla l_i(\theta^*)]]^{-1} \sqrt{n}\, E_n[l_i(\theta_0)]$.

By LLN, $E_n[\nabla l_i(\theta^*)] \xrightarrow{p} G$ and by CLT,

$\sqrt{n}\, E_n[l_i(\theta_0)] \xrightarrow{d} N(0, \Omega)$, where $\Omega = E[l_i(\theta_0) l_i(\theta_0)']$
$= -G$

Thus, by Slutsky,

$\sqrt{n}\,(\hat{\theta} - \theta_0) \xrightarrow{d} -G^{-1} N(0, \Omega) \overset{d}{=} N(0, G^{-1} \Omega G^{-1})$
$\overset{d}{=} N(0, G^{-1})$