Lecture 5 . Asymptotic Theory.

All asymptotics are fake, sometimes they are useful.

Def. ~~$a_n \to a$ if~~ Convergence for any $\varepsilon > 0$, $\exists$ an $N(\varepsilon)$ such that for all $n > N(\varepsilon)$, we have $a_n \to a$.

Def. Convergence in probability:

$$P(|X_n - X| \geq \delta) \to 0$$

Some bad event will happen.

for any δ, ε, $\exists N(\delta, \varepsilon)$ such that for all $n > N$.

$$X_n \xrightarrow{p} X$$

$$\text{or} \quad \operatorname{plim} X_n = X.$$

Different from convergence in expectation

$$E Z_n \to E Z.$$

e.g. if $Z_n = \begin{cases} 0, \text{ with prob } 1 - \frac{1}{n} \\ n, \text{ with prob } \frac{1}{n}. \end{cases}$

then $E Z_n \Rightarrow 1$ but ~~ϕ~~ $Z_n \xrightarrow{p} 0$.

LLN.

Chebyshev inequality. for any mean ~~0~~ r.v. Z_n

~~$Pr(|Z_n|>$~~ with finite variance,

$$Pr(Z_n > \delta) \leq \frac{var(Z_n)}{\delta^2}.$$

Use Chebyshev inequality to prove LLN.

if $X_1 \dots X_n$ are i.i.d, and $var(X_1) < \infty$, then

$\bar{X}_n \xrightarrow{p} EX$. applies to i.n.i.d.

$$Pr(|\bar{X}_n - EX| > \delta) \leq \frac{var(X_1)}{n\delta^2} \to 0.$$

Stronger version:

if $X_1 \dots X_n$ are i.i.d, and ~~EX exists~~ EX exists, $E|X| < \infty$, then

$\bar{X}_n \xrightarrow{p} EX$. Kolmogorov LLN.

If X is vector-valued, convergence mean convergence for each element.

Convergence in distribution.

~~#~~ $X_n \sim F_n(a)$, $\quad X \sim F$

$X_n \xrightarrow{d} X$ if $F_n(a) \longrightarrow F(a)$

at all points that $F(a)$ is continuous.

The difference between $\xrightarrow{P}$ and $\xrightarrow{d}$.

weak convergence.

<u>Lindeberg-Levy CLT</u>: if iid and ~~E~~ $\mathrm{var}(X_i) < \infty$,

then $\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} N(0, \sigma^2)$

Vector version: $\sqrt{n}(\bar{X} - \mu) \xrightarrow{d} N(0, E[(X-\mu)(X-\mu)'])$

<u>Lindeberg-Feller CLT (inid)</u>

$$\frac{\sum (y_i - \mu_i)}{\sqrt{\sum \sigma_i^2}} \xrightarrow{d} N(0,1),$$

if the Lindeberg condition is satisfied.

Sufficient condition that are easy to verify.

$\sup E(y_i - \mu_i)^3 < k < \infty$ variance 不太大

$\inf E\,\sigma_i^2 > c > 0$ 也不太小

CMT: if $Z_n \xrightarrow{P} c$, and $g(Z_n)$ is continuous at c, then $g(Z_n) \xrightarrow{P} g(c)$.

CMT for $\xrightarrow{d}$: if g is continuous almost everywhere, $g(Z_n) \xrightarrow{d} g(C)$

eg $Z_n + a \xrightarrow{P} c + a$

$aZ_n \xrightarrow{P} ac$

$\frac{a}{Z_n}$ $\xrightarrow{P}$ $\frac{a}{c}$ if $c \neq 0$.

Delta method: g is a continuous function, ~~or the discontinuity points has measure zero~~ continuously differentiable in a ngh. or μ

if $\sqrt{n}(\hat{\theta}_n - \theta_n) \xrightarrow{d} N(0, \sigma^2)$

then $\sqrt{n}\left(g(\hat{\theta}_n) - g(\theta)\right) \xrightarrow{d} N(0, g'(\theta)^2 \sigma^2)$.

Vector version

$g(\hat{\theta}) - g(\theta) \xrightarrow{d} N(0, \nabla g(\theta) \Omega \nabla g'(\theta))$

pf $\sqrt{n}\left(g(\hat{\theta}_n) - g(\theta)\right) = \sqrt{n}\, G'(\theta)(\hat{\theta} - \theta)$

$+ \sqrt{n}\, G''(\theta^*)(\hat{\theta} - \theta)^2$.

Smaller order.

Slutsky theorem

if $X_n \xrightarrow{p} X$, $Y_n \xrightarrow{d} Y$

then $X_n + Y_n \xrightarrow{d} X + Y$

$X_n Y_n \xrightarrow{d} XY$

Stochastic symbols.

if $X_n \xrightarrow{p} 0$, we say $X_n = o_p(1)$ $\frac{1}{n}\sum X_n = O_p(1)$

if $\frac{X_n}{n^r} \xrightarrow{p} 0$, we say $X_n = o_p(n^r)$. $\hat{\mu} = \mu + O_p(n^{-\frac{1}{2}})$

if for any ε, $\exists$ $M(\varepsilon)$ such that

$\limsup_{n\to\infty} Pr(|X_n| > M(\varepsilon)) \leq \varepsilon$,

then we say $X_n = O_p(1)$.

eg. normal.

~~Lecture 6.~~ Asy theory for LS.

for the projection model,

$$y = x\beta + \varepsilon, \quad E(\varepsilon x) = 0.$$

$$\hat{\beta} = (X'X)^{-1} X'\varepsilon$$

Assumption: 1. i.i.d.

2. $Ey^2 < \infty$.

3. $E\|x\|^2 < \infty$

4. $Q_{xx} = E(XX')$ is positive definite.

Thm 1. (Consistency).

$$\hat{\beta} = (X'X)^{-1} X'(X\beta + \varepsilon)$$

$$= \beta_0 + \left(\frac{X'X}{n}\right)^{-1} \left(\frac{X'\varepsilon}{n}\right).$$

$$\left(\frac{X'X}{n}\right)^{-1} \xrightarrow{P} (E\,XX')^{-1}. \quad Q_{xx} \text{ p.d. invertible}$$

$$\frac{X'\varepsilon}{n} \xrightarrow{P} E(X_i\varepsilon_i) = 0.$$

Thm 2. Asy normality.

$$\sqrt{n}(\hat{\beta}-\beta_0) = \left(\frac{X'X}{n}\right)^{-1} \frac{1}{\sqrt{n}} \sum x_i \varepsilon_i.$$

if $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \varepsilon_i \xrightarrow{d} N(0, \Omega)$

then $\sqrt{n}(\hat{\beta}-\beta_0) \xrightarrow{d} N(0, Q_{xx}^{-1} \Omega Q_{xx}^{-1})$.

What is Ω? $\Omega = E[x_i x_i' e_i^2]$

~~If~~ indep. $E(x_i x_i' e_i^2) = E(x_i x_i') E(e_i^2)$.

So second ~~ment~~ moment is sufficient,

When we allow depends between x_i, e_i, must have stronger condition.

Sufficient: $E|x_i|^4 < \infty$, $E|e_i|^4 < \infty$
for all i

Assumption 6.1.2 in the book

Cauchy Schwarz inequality to make sure $\Omega < \infty$.

if homoskedastic

$$E(x_i x_i' e_i^2) = E(x_i x_i' E(e_i^2 | x_i)) = \sigma^2 E(x_i x_i')$$

So $\sqrt{n}(\hat{\beta}-\beta_0) \xrightarrow{d} N(0, Q_{xx}^{-1} \sigma^2)$.

Estimate the error variance.

If e_i ~~can~~ were observable, $\frac{1}{n}\sum e_i^2$

Now we only have $\hat{e}_i = y_i - x_i\hat{\beta}$

$$\cancel{\frac{1}{n}\sum_{i=1}^{n}(\hat{e}_i\hat{e}_i) = \frac{1}{n}\sum_{i=1}^{n}} = x_i(\beta - \hat{\beta}) + e_i$$

$$\frac{1}{n}\sum_{i=1}^{n}\hat{e}_i^2 = \frac{1}{n}\sum_{i=1}^{n}e_i^2 + \underline{\frac{1}{n}\sum x_i e_i}(\underbrace{\hat{\beta}-\beta_0}_{\to 0})$$

$$+ \underbrace{(\hat{\beta}-\beta_0)}_{\downarrow\, 0}\underbrace{\frac{1}{n}\sum x_i x_i}_{\downarrow\, Q_{xx}}\underbrace{(\hat{\beta}-\beta_0)}_{\downarrow\, 0}$$

Under homo.

$\text{Avar}(\hat{\beta}) = \underline{\underline{Q_{xx}^{-1}}}\sigma^2$ trivial.

heteroskedastic, the derivation is complexated but the result is

$$\frac{1}{n}\sum x_i x_i \hat{e}_i^2 - \frac{1}{n}\sum x_i x_i' e_i^2 \xrightarrow{p} 0,$$

~~thk~~ $\frac{1}{n}\sum x_i x_i e_i \xrightarrow{p} \Omega$