- <u>Result</u> : if $X_n \xrightarrow{d} X$ and $Y_n - X_n \xrightarrow{P} 0$, then $Y_n \xrightarrow{d} X$.

Pf/ Let K be a closed set. Want to show that $\limsup_n \mathbb{P}(Y_n \in K) \leq \mathbb{P}(X \in K)$ [Portmanteau, part (iv).]

We have, for arbitrary $\underset{\text{small}}{\varepsilon > 0}$,

these sets are events

$$\{Y_n \in K\} = \left(\{Y_n \in K\} \cap \{d(X_n, Y_n) \leq \varepsilon\}\right) \cup \left(\{Y_n \in K\} \cap \{d(X_n, Y_n) > \varepsilon\}\right)$$

$$\subseteq \{d(X_n, K) \leq \varepsilon\} \cup \{d(X_n, Y_n) > \varepsilon\}$$

where $d(K, x) = \inf_{y \in K} d(x, y)$

closed set in $\mathbb{R}^d$ — point in $\mathbb{R}^d$

$\{x : d(K, x) \leq \varepsilon\}$

Therefore

$$\mathbb{P}(Y_n \in K) \leq \mathbb{P}(X_n \in K_\varepsilon) + \mathbb{P}(d(X_n, Y_n) > \varepsilon)$$

closed set — $\to 0$ as $n \to \infty$

$\Downarrow$

$$\limsup_n \mathbb{P}(Y_n \in K) \leq \limsup_n \mathbb{P}(X_n \in K_\varepsilon)$$

$$\leq \mathbb{P}(X \in K_\varepsilon) \quad \text{by Portmanteau part (iv)}$$

We now let $\varepsilon \downarrow 0$ so that $\mathbb{P}(X \in K_\varepsilon)$

SDS 387
Linear Models

Fall 2024

Lecture 8 - Thu, Sep 19, 2024

Instructor: Prof. Ale Rinaldo

- Last time: O_p and o_p notation. Useful notation for asymptotic calculations.

- Example: $X_1, X_2, \ldots$ iid $\sim (\mu, \sigma^2)$. Then

$$\frac{\sqrt{n}\,(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1) \qquad \text{CLT}$$

$$\hookrightarrow \quad \sqrt{n}\,(\bar{X}_n - \mu) = O_p(1)$$

$$\hookrightarrow \quad (\bar{X}_n - \mu) = O_p\left(\frac{1}{\sqrt{n}}\right)$$

The last statement implies $\bar{X}_n - \mu = o_p(1)$ HW! which the WLLN. But we also get additional information, namely that, by inflating $\bar{X}_n - \mu$ by a factor of $\sqrt{n}$, we obtain a sequence that is bounded in probability.

- Rules for O_p / o_p calculus,

  - $o_p(1) + o_p(1) = o_p(1)$ → Sum of a fixed # of terms that are $o_p(1)$ is also $o_p(1)$
  - $O_p(1) + O_p(1) = O_p(1)$
  - $\underbrace{O_p(1)\, o_p(1)}$ $= o_p(1)$

    $O_p(o_p(1)) = o_p(1) = o_p(O(1))$
  - $(1 + o_p(1))^{-1} = O_p(1)$
  - $\dfrac{1}{O_p(1)}$: what can you say about this ?

- Last time I remarked that if $X_n = O_p(1)$ we cannot conclude that $X_n \xrightarrow{d}$ anything !
  It only happens along subsequences.

<u>Prokhorov's Thm:</u> i) If $X_n \xrightarrow{d} X$ then $X_n = O_p(1)$

$\Downarrow$ bounded in prob. or tight

ii) If $X_n = O_p(1)$ then
$\exists$ $\{n_k\}_{k=1,2,\dots}$ s.t.

$$X_{n_k} \xrightarrow{d} Y \quad \text{some } Y \quad \text{as } k \to \infty$$

CENTRAL LIMIT THEOREM

Basic Form: Let $X_1, X_2, \ldots \overset{\text{iid}}{\sim} (\mu, \Sigma)$ in $\mathbb{R}^d$

[$\mathbb{E}[X_i] = \mu \in \mathbb{R}^d$ and $\mathbb{E}\left[(X_i - \mu)(X_i - \mu)^T\right] = \underset{d \times d}{\Sigma} > 0$ all i]

Then

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) \xrightarrow{d} N_d(0, \Sigma)$$

or

$$\sqrt{n}\,(\bar{X}_n - \mu)$$

Pf/ Use ch.f.'s. Let $X \overset{d}{=} X_i$ all i, and let

$t \in \mathbb{R}^d \longmapsto \varphi(t) = \mathbb{E} \exp\{ i\, t^T (X - \mu) \}$ the ch. f. of $X - \mu$

Then,

$$\varphi_{\sqrt{n}(\bar{X}_n - \mu)}(t) = \varphi_{\sum_{i=1}^{n} (X_i - \mu)}\left(t/\sqrt{n}\right)$$

$$\overset{\text{by indep}}{=} \prod_{i=1}^{n} \varphi_{X_i - \mu}\left(t/\sqrt{n}\right)$$

$$\overset{\text{because } X_i\text{'s are identically distributed}}{=} \left(\varphi\left(t/\sqrt{n}\right)\right)^n \qquad (*)$$

Next recall that $\varphi(0) = 1$, $\nabla \varphi(0) = i\, \mathbb{E}[X - \mu] = 0$

Hessian $\leftarrow \nabla^2 \varphi(0) = i^2 \Sigma = -\Sigma$

By Taylor serie expansion of $\varphi(t/\sqrt{n})$ around 0.

$$(*) = \left(1 + \underbrace{\frac{i\, t^T \nabla\varphi(0)}{\sqrt{n}}}_{=0} + \frac{1}{2}\,\frac{i^2}{n}\, t^T \left[\int_0^1 \nabla\varphi\left(u\frac{t}{\sqrt{n}}\right) du\right] t\right)^n$$

$$= \left(1 + \frac{1}{n}\underbrace{\frac{1}{2}\, i^2\, t^T \int_0^1 \nabla\varphi(u\, t/\sqrt{n})\, du\; t}_{a_n}\right)^n$$

Next

$$a_n \longrightarrow -\frac{t^T}{2}\Sigma \underbrace{\int_0^1 du}_{1}\, t = -\frac{t^T}{2}\Sigma\, t$$

it is ok to bring limit inside the integral

Because $(1 + c_n)^n = \exp\{\lim_n n c_n\}$ if limit exist

here $c_n = \frac{a_n}{n}$

$\hookrightarrow$ $\varphi_{\sqrt{n}(\bar{X}_n - \mu)}(t) \longrightarrow \exp\{-t^T\Sigma t\}$ as $n\to\infty$

$\Downarrow$

ch. f. of $N_d(0,\Sigma)$

By Continuity Theorem for ch.f's

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0,\Sigma)$$

- CLT: triangular array version. A triangular array is an infinite collection of r.v.'s $\{X_{i,n},\ i \le n\}$ organized in this manner:

$$\begin{array}{lllll}
X_{1,1} & & & & \\
X_{1,2} & X_{22} & & & \\
X_{1,3} & X_{2,3} & X_{3,3} & & \\
\vdots & & & & \\
X_{1,n} & X_{2,n} & X_{3,n} & \cdots & X_{n,n} \\
\vdots & & & &
\end{array}$$

The rows of the array consist of independent r.v.'s.

The Lindeberg-Feller CLT:

Let $\{X_{i,n}\}$ be a triangular array of r.v.'s in $\mathbb{R}$ s.t. $\mathbb{E}[X_{i,n}] = 0$ $\forall i, n$. Let

$S_n = \sum_{i=1}^{n} X_{i,n}$ and $B_n^2 = \sum_{i=1}^{n} \sigma_{i,n}^2$ where

$\sigma_{i,n}^2 = \mathrm{Var}[X_{i,n}]$. Then

$$\frac{S_n}{B_n} \xrightarrow{d} N(0,1)$$

if the LF (Lindeberg-Feller) condition

$$(LF) \quad \forall \underbrace{\varepsilon > 0}_{\text{small}} \quad \frac{1}{B_n^2} \sum_{i=1}^{n} \mathbb{E}\left[X_{i,n}^2 \, 1\{|X_{i,n}| > \varepsilon B_n\}\right] \xrightarrow{0} \text{ as } n \to \infty$$

is met.

Conversely, if $\frac{S_n}{B_n} \xrightarrow{d} N(0,1)$ and if

asymptotic uniform negligibility $\Leftarrow$ $\frac{\max_{i=1,\dots,n} \sigma_{i,n}^2}{B_n^2} \to 0$ as $n \to \infty$

then (LF) holds.

- Often, it is easier to establish a CLT via Lyapunov's condition

How! $\leftarrow$ $$\frac{1}{B_n^{2+\delta}} \sum_{i=1}^{n} \mathbb{E}\left[|X_{i,n}|^{2+\delta}\right] \to 0$$

some $\delta > 0$.

This implies LF.

$\downarrow$ it requires existence of moments higher than 2.

The multivariate case.

Consider a triangular array of centered random vectors in $\mathbb{R}^d$ with 2 moments ($\mathrm{Cov}[X_{i,n}]$ exists $\forall i,n$).

Let
$$Y_{i,n} = \left(\sum_{i=1}^{n} \mathrm{Cov}[X_{i,n}] \right)^{-1/2} X_{i,n}$$

Then, if

$$\text{(LF)} \quad \forall \varepsilon > 0, \quad \lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{E}\left[\|Y_{i,n}\|^2 \, 1\{\|Y_{i,n}\| > \varepsilon\} \right] = 0$$

$$\sum_{i=1}^{n} Y_{i,n} \xrightarrow{d} N(0, I_d)$$