SDS 387
Linear Models

Fall 2024

Lecture 6 - Thu, Sep 12, 2024

Instructor: Prof. Ale Rinaldo

Portmanteau Theorem

The following conditions are equivalent (we are in $\mathbb{R}^d$)

i) $X_n \xrightarrow{d} X$

ii) $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ for all bounded and continuous functions $f: \mathbb{R}^d \to \mathbb{R}$

iii) $\liminf \mathbb{P}(X_n \in G) \geq \mathbb{P}(X \in G)$ for all open sets G

iv) $\limsup \mathbb{P}(X_n \in K) \leq \mathbb{P}(X \in K)$ for all closed sets K

v) $\mathbb{P}(X_n \in A) \to \mathbb{P}(X \in A)$ for all (Borel) sets A s.t. $\mathbb{P}(X \in \partial A) = 0$

$\partial A = \bar{A} \setminus A^\circ$

Continuous mapping theorem: CMT

Let $f: \mathbb{R}^d \to \mathbb{R}$ and let X be a random vector in $\mathbb{R}^d$ s.t. $\mathbb{P}(X \in C) = 1$ where C is the set of continuity points of f. Then

$$X_n \xrightarrow{*} X \implies f(X_n) \xrightarrow{*} f(X)$$

where $*$ means wp 1 or in probability or in distribution.

<u>Note</u>: if f is continuous (meaning continuous at all points in its domain, which contain the image of X), the CMT automatically holds.

Pf/ We will prove that $X_n \xrightarrow{d} X$ implies that $f(X_n) \xrightarrow{d} f(X)$ [if $\mathbb{P}(X \in C) = 1$, C the set of continuity points of f]

see van der Vaart, chapter 2.

[x is a continuity point of f if

$\forall \{x_n\}$ s.t. $x_n \to x$, $f(x_n) \to f(x)$

or $\forall \overset{\to \text{small}}{\varepsilon} > 0$ $\exists \delta = \delta(\varepsilon)$ s.t. $|x - y| < \delta$

$\to |f(x) - f(y)| < \varepsilon$.]

Use Portmanteau Thm. Let K be a closed set in $\mathbb{R}$.

$$\{f(X_n) \in K\} = \{X_n \in \underbrace{f^{-1}(K)}_{\{x \in \mathbb{R}^d : f(x) \in K\}}\}$$

Then, I claim $\rightarrow \bar{A}$ is the closure of A

$$f^{-1}(K) \subseteq \overline{f^{-1}(K)} \subseteq f^{-1}(K) \cup C^c$$

To see this, take $x \in \overline{f^{-1}(K)}$. Then $\{x_n\} \subset f^{-1}(K)$ s.t. $x_n \to x$ and $f(x_n) \in K$ all n. If $x \in C$ then $f(x_n) \to f(x) \in K$ because K is closed.

otherwise $x \in C^c$. So

$$\limsup_n \mathbb{P}(f(X_n) \in K) \leq \limsup_n \mathbb{P}\left(X_n \in \overline{f^{-1}(K)}\right)$$

by Portmanteau Thm part (iv) $\leftarrow$ $\leq \mathbb{P}\left(X \in \overline{f^{-1}(K)}\right)$

$$\leq \mathbb{P}(X \in f^{-1}(K)) + \underbrace{\mathbb{P}(X \in C^c)}_{=0 \text{ by assumption}}$$

$$\| $$

$$= \mathbb{P}(f(X) \in K).$$

So $\limsup_n \mathbb{P}(f(X_n) \in K) \leq \mathbb{P}(f(X) \in K)$

all closed sets K. So by part (iv) of Portmanteau Thm,

$$f(X_n) \xrightarrow{d} f(X)$$

■

Example: we need f to be continuous (wp 1 wrt to distribution of X).

Let
$$X_n = \begin{cases} 0 + 1/n & \text{with prob } 1/2 \\ 0 - 1/n & \text{"} \end{cases}$$

↳ $X_n \xrightarrow{d} 0$ Let $f(x) = \mathbb{1}\{x \geq 0\}$ → bounded but not continuous

but
$$f(X_n) \overset{d}{\not\to} 1 \qquad f(X) \text{ is degenerate at } 1$$

Example of CMT: $X_1, X_2, \ldots \overset{iid}{\sim} (\mu, \sigma^2)$. Then

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

Then $(T_n)^2 \xrightarrow{d} \chi^2_1$

■ Characteristic functions

Powerful analytic approach to demonstrate convergence in distribution (and prove WLLN) and to prove CLT.

See Ferguson Chapter 3 and or van der Vaart Section 2.3

Definition For a r.v. $X \in \mathbb{R}^d$, its characteristic function is:
$$t \in \mathbb{R}^d \longmapsto \varphi_X(t) = \mathbb{E}\left[\exp\{i\, t^T X\}\right]$$
$$= \mathbb{E}\left[\cos(t^T X) + i \sin(t^T X)\right]$$

The ch.f. of a random variable encodes the properties of its distribution

Continuity Theorem (Ferguson Thm 3e)

i) $X_n \xrightarrow{d} X$ iff $\varphi_{X_n}(t) \to \varphi_X(t)$ $\forall t \in \mathbb{R}^d$

ii) Moreover if $\varphi_{X_n}(t)$ converges pointwise to a function, say φ, continuous at 0, then φ is the ch.f. of a random variable, say X, s.t. $X_n \xrightarrow{d} X$

iii) $X \stackrel{d}{=} Y$ iff $\varphi_X(t) = \varphi_Y(t)$ $\forall t \in \mathbb{R}^d$

Remark: if $Z \sim N_d(\mu, \Sigma)$ then

$$\varphi_Z(t) = \exp\left\{ i\, t^T \mu - \frac{t^T \Sigma t}{2} \right\}$$

Let's use ch.f. to prove WLLN. Recall Taylor series expansion formula.

Let $f: \mathbb{R}^d \to \mathbb{R}$ and assume that f has $(k+1)$ continuous partial derivatives at all points in an open set U. Let $x, x_0 \in U$ s.t. the line segment $\overline{x x_0} \subset U$. Then

$$f(x) = f(x_0) + \sum_{i=1}^{k} \frac{1}{k!} D^{(k)} f(x_0, x - x_0) + \underbrace{\text{Rem}}_{o(\|x - x_0\|^k)}$$

where $h = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_d \end{bmatrix} \in \mathbb{R}^d$

$$D^{(k)} f(x_0, h) = \sum_{i_1 \le i_2 \le i_3 \le \cdots \le i_k} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(x_0)\, h_{i_1} h_{i_2} \cdots h_{i_k}$$

[when $k=1$ $\quad D^{(k)} f(x, h) = h^T \nabla f(x_0)$

$k=2$ $\quad$ " $\quad = h^T \nabla^2 f(x_0) h$

$k=3$ $\quad$ " $\quad \langle \nabla^3 f(x_0), h \otimes h \otimes h \rangle$

$\vdots$]

and Rem can be expressed as

Lagrangian $\quad$ i) $\displaystyle \frac{1}{(k+1)!} D^{(k+1)} f(z, x - x_0)$ $\quad$ some z on $\overline{x x_0}$

Integral $\quad$ ii) $\displaystyle \frac{1}{k!} \int_0^1 D^{(k)} f(\tau x + (1-\tau) x_0, x - x_0)\, d\tau$

Back to ch. f. to prove WLLN: $X_1, X_2, \ldots \overset{\text{iid}}{\sim}$ from a distribution with mean $\mu = \mathbb{E}[X_1]$. Then

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

Pf/

$$\varphi_{\bar{X}_n}(t) = \varphi_{X_1 + X_2 + \cdots + X_n}(t/n)$$

by indep. $\swarrow = \prod_{i=1}^{n} \varphi_{X_i}(t/n)$

X_i's have same distribution because $= \left(\varphi_X(t/n)\right)^n$ $\quad X \overset{d}{=} X_i$ all i

(Notice that $\nabla \varphi_X(0) = i\mu$ Exercise)

$$= \Big(\underbrace{\varphi_X(0)}_{=1} + \int_0^1 \frac{t^T}{n} \nabla \varphi_X(\tau t/n)\, d\tau\Big)^n$$

So

$$\lim_{n\to\infty} \varphi_{\bar{X}_n}(t) = \lim_n \left(1 + \int_0^1 \frac{t^T}{n} \nabla \varphi_X(\tau t/n)\, d\tau\right)^n$$

[fact $\lim_n (1+a_n)^n = \exp\{\lim_{n\to\infty} n a_n\}$ if $\lim_n n a_n$ exists]

$$= \exp\left\{\lim_{n\to\infty} \int_0^1 t^T \nabla \varphi_X(\tau t/n)\, d\tau\right\}$$

Dominated convergence Thm $\swarrow$

$$= \exp\{t^T \nabla \varphi_X(0)\} = \exp\{i\, t^T \mu\}$$

$\Downarrow$

ch f. of a degenerate r.v. at μ.

- By Continuity Thm $\bar{X}_n \xrightarrow{d} \mu$. Therefore $\bar{X}_n \xrightarrow{P} \mu$

Thm (Cramer-Wold device) Let $\{X_n\}$ be a sequence of r.v.'s in $\mathbb{R}^d$. Then

$$X_n \xrightarrow{d} X \quad \text{iff} \quad t^T X_n \xrightarrow{d} t^T X$$

$$\forall t \in \mathbb{R}^d$$

X: some random vector

Recall the example from last time

$$X_n = U \sim \text{Uniform}(0,1)$$

$$Y_n = \begin{cases} U & n \text{ even} \\ 1-U & \text{odd} \end{cases}$$

Then $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \not\xrightarrow{d}$ even though

$$X_n \xrightarrow{d} U$$

$$Y_n \xrightarrow{d} U$$

Use Cramer-Wold with $t = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. They

$$t^T \begin{bmatrix} X_n \\ Y_n \end{bmatrix} = \begin{cases} 2U & n \text{ even} \\ 1 & \text{odd} \end{cases}$$

which does not converge.