SDS 387
Linear Models

Fall 2024

Lecture 7 - Tue, Sep 17, 2024

Instructor: Prof. Ale Rinaldo

- Last time: <u>Cramer-Wold device</u>: Let $\{X_n\}$ be a sequence of r.v.'s in $\mathbb{R}^d$ and X also a r.v. in $\mathbb{R}^d$. Then

$$X_n \xrightarrow{d} X \quad \text{iff} \quad \underbrace{t^T X_n}_{\in \mathbb{R}} \xrightarrow{d} t^T X \quad \underbrace{\forall t \in \mathbb{R}^d}_{\text{deterministic}}$$

HW
we ch.f.'s.

- We also saw that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$ does <u>not</u> imply that $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ Y \end{bmatrix}$ and therefore it does not imply that $f(X_n, Y_n) \xrightarrow{d} f(X, Y)$ (not even if $f(\cdot,\cdot)$ is well-behaved, e.g. $f(x,y) = x+y$)

- <u>Exception</u>: if $X \perp\!\!\!\perp Y$ and $X_n \perp\!\!\!\perp Y_n$ $\forall n$ then $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} Y$ imply $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ Y \end{bmatrix}$

HW
Use ch.f.

• **Result**: 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 $\limsup_n \mathbb{P}(Y_n \in K) \le \mathbb{P}(X \in K)$

[Portmanteau, part (iv).]

We have, for arbitrary $\varepsilon > 0$ (ε small),

these sets are events

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

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

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

K: closed set in $\mathbb{R}^d$; x: point in $\mathbb{R}^d$

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

Therefore

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

K_ε: closed set; $\mathbb{P}(d(X_n, Y_n) > \varepsilon) \to 0$ as $n \to \infty$

$\Downarrow$

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

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

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

$\to \mathbb{P}(X \in k)$

$\hookrightarrow \limsup_n \mathbb{P}(Y_n \in k) \le \mathbb{P}(X \in k)$ all k

So, by Portmanteau (iv) $Y_n \xrightarrow{d} X$. ∎

<u>Corollary</u> $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{P} c$, c a constant,

Then $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ c \end{bmatrix}$

HW! (Use previous result!)

<u>Slutsky's Theorem</u> $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} c$. Then

$$X_n + Y_n \xrightarrow{d} X + c$$

$$Y_n X_n \xrightarrow{d} cX$$

$$\frac{X_n}{Y_n} \xrightarrow{d} \frac{X}{c}, \quad c \neq 0.$$

Analogous result holds if $\{X_n\}$ and $\{Y_n\}$ are sequences of random vectors (in fact, random matrices)

For example in $\mathbb{R}^d$

$$Y_n^T X_n \xrightarrow{d} c^T X$$

<u>Examples</u>: $X_1, X_2, \ldots$ iid $\sim (\mu, \sigma^2)$

$\vee$ first 2 moments

Then $\bar{X}_n \xrightarrow{P} \mu$ WLLN

$\frac{1}{n}\sum_{i=1}^{n} X_i$

and, as we will see, $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$

Central Limit Theorem

We want to estimate σ^2 (or σ)

We can use

$$\hat{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

How do we prove that $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$? Write

$$\hat{\sigma}_n^2 = \underbrace{\underbrace{\underbrace{\frac{n}{n-1}}_{\to 1} \left[\underbrace{\frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2}_{\xrightarrow{p} \sigma^2 \text{ by WLLN}} - \underbrace{(\bar{X}_n - \mu)^2}_{\xrightarrow{p} 0 \text{ by WLLN + CMT}} \right]}_{\xrightarrow{p} \sigma^2 \text{ by CMT}}}_{\xrightarrow{p} \sigma^2 \text{ by Slutsky}}$$

show this

joint convergence in distribution

In turn, because $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$, we can conclude that

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} = \underbrace{\frac{\sigma}{\hat{\sigma}_n}}_{\xrightarrow{p} 1} \underbrace{\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}}_{\xrightarrow{d} Z(0,1)} \xrightarrow[\text{by Slutsky}]{d} N(0,1)$$

random

↳ a $1-\alpha$ asymptotic confidence interval for μ

is $\bar{X}_n \pm \frac{\hat{\sigma}_n}{\sqrt{n}} z_{1-\alpha/2}$

↳ upper $1-\alpha/2$ quantile of $N(0,1)$

■ Uniform convergence of c.d.f.'s. Recall that

$X_n \xrightarrow{d} X$ (in $\mathbb{R}^d$) iff

$F_{X_n}(x) \to F_X(x)$ for all points $x \in \mathbb{R}^d$ at which F_X is continuous

cdf of X_n — cdf of X

pointwise convergence

If F_X is continuous at all points $x \in \mathbb{R}^d$, then

$$\sup_{x \in \mathbb{R}^d} |F_{X_n}(x) - F(x)| \to 0$$

uniform convergence !!

HW !!

van der Vaart section 2.2

■ O_p and o_p notation

big-oh p — little-oh p

Recall that a non-random sequence $\{x_n\}$ in $\mathbb{R}^d$ is $o(1)$ when $x_n \to 0$ and $O(1)$ when $\{x_n\}$ is bounded (and so it has a converging sub-sequence)

- Let $\{X_n\}$ be a sequence of random vectors and $\{r_n\}$ a sequence of positive numbers.

Then $X_n = o_p(r_n)$ means $X_n = r_n \cdot Y_n$ where $Y_n \xrightarrow{p} 0$

$$\left[\forall \varepsilon > 0 \quad \lim_n \mathbb{P}\left(\left\|\frac{X_n}{r_n}\right\| > \varepsilon \right) = 0 \right]$$

so $X_n = o_p(1)$ means $X_n \xrightarrow{p} 0$

- $X_n = O_p(r_n)$ means $\forall \varepsilon > 0$ $\exists M > 0$ and $N \in \mathbb{N}$ both depending on ε

s.t.

$$\mathbb{P}\left(\frac{\|X_n\|}{r_n} > M \right) \leq \varepsilon \quad \forall n \geq N$$

In particular

$X_n = O_p(1)$ when it is <u>bounded in probability</u>

Let $Z \sim N(0, I_d)$. Then $Z = O_p(1)$

In fact any random vector is bounded in probability

Note $X_n = O_p(1)$ we cannot conclude that

$X_n \xrightarrow{d}$