Probability Theory: Overview

A probability space (S, Q, P) triple.

S: Sample space

Q: σ-algebra

P: probability Measure

P:

① $1 \geq P(A) \geq 0$

② $P(S) = 1$

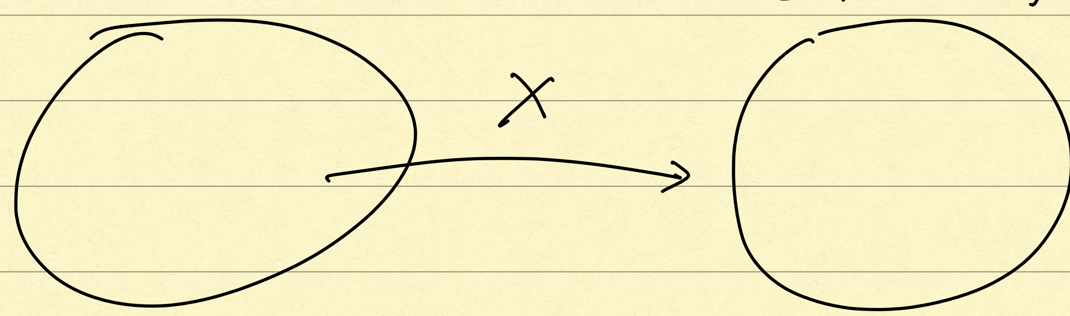
③ $A_1, \cdots, A_n$ are pairwise disjoint $\quad P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$

random variable X is a function

$X: S \to \mathbb{R}$

with the property that $X^{-1}(B) \in Q$ for all Borel set $B \subseteq \mathbb{R}$

Picture $\quad (S, Q, P) \qquad\qquad (\mathbb{R}, B, P_X)$



a sequence of r.v.'s $X_1, \cdots, X_n, \cdots$

Types of convergence:

Tools: 1. almost surely convergence. (convergence with prob. 1)

$$X_n \xrightarrow{a.s.} X$$

$$P(\lim_{n\to\infty} X_n = X) = 1 \qquad P(\lim_{n\to\infty} (X_n - X) = 0) = 1$$

Y_n real number

2. convergence in probability

$$X_n \xrightarrow{P} X$$

$\forall \varepsilon > 0.$ $\quad \lim_{n\to\infty} P(|X_n - X| < \varepsilon) = 1$

3. convergence in distribution

$$\lim_{n\to\infty} P(X_n \le x) = P(X \le x) = F(x)$$

for all x where F is continuous

Tools for a.s.

1. Strong Law of Large Numbers (SLLN)

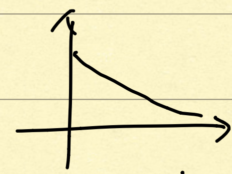
$X_1, \cdots, X_n$ iid $\quad E(X_i) = \mu \quad \mathrm{Var}\, X_i = \sigma^2 < +\infty$ or $E|X_i| < +\infty$

Then $\quad \bar{X}_n = \dfrac{\sum_{i=1}^{n} X_i}{n} \xrightarrow{a.s} \mu$

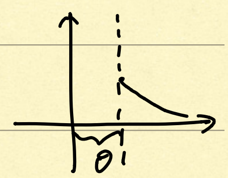
2. Borel – Cantelli Lemma

$\forall \varepsilon > 0.$ If $\sum_{n=1}^{\infty} P(|X_n - \theta| > \varepsilon)$ is finite

Then $\quad X_n \xrightarrow{a.s} \theta$



Example. $\quad X_1, \cdots, X_n$ iid $\quad f_X(x|\theta) = e^{-(x-\theta)} I(x>\theta)$



MLE: $\quad \hat{\theta}_{MLE} = X_{(1)} \leftarrow$ minimum of n sample

MOM: $\quad \hat{\theta}_{MOM} = \bar{X}_n - 1 \xrightarrow[a.s.]{S.LLN} \theta \qquad E(X) = 1 + \theta$

$E(X) - 1 = \theta$

$\forall \varepsilon > 0.$ $\quad P(|X_n - \theta| > \varepsilon) \qquad X_n - \theta \sim \mathrm{Exp}(1)$

$= P(X_n - \theta > \varepsilon)$

$= e^{-\varepsilon} < 1$

$$\hat{\theta}_{MLE} = X_{(1:n)} = Y_n$$

$$X_{(1:n)} \xrightarrow{a.s} \theta$$

$$\forall \varepsilon > 0. \quad \sum_{n=1}^{\infty} P(\ |X_{(1:n)} - \theta| > \varepsilon) = \sum_{n=1}^{\infty} (e^{-\varepsilon})^n < +\infty$$

$$= P(\ X_{(1:n)} - \theta > \varepsilon) \qquad 0 < e^{-\varepsilon} < 1$$

$$= P(\ X_{(1:n)} > \theta + \varepsilon)$$

$$= \prod_{i=1}^{n} P(X_i > \theta + \varepsilon)$$

$$= (e^{-\varepsilon})^n$$

Convergence in Probability.

Eg: Suppose $X_n = \begin{cases} \frac{1}{2^n} & \text{with prob. } P_n \\ 100 & \text{w.p. } 1 - P_n \end{cases}$

$\forall \varepsilon > 0$. for $n > -\frac{\log \varepsilon}{\log 2}$ $\quad (\frac{1}{2^n} < \varepsilon)$

$$P(|X_n - 0| > \varepsilon) = P(X_n = 100) = 1 - P_n$$

as long as $P_n \to 1$. we have $X_n \xrightarrow{P} 0$

Remark: the beauty of convergen in Prob. is that a small & diminishing amount of "bad behavior" is allowed.

Consistency: estimator $\hat{\theta}_n = \hat{\theta}_n(X_1, \cdots, X_n)$ is consistent for θ

$\forall \varepsilon > 0$. $\forall \theta \in \Theta$

we have $\lim_{n \to \infty} P(|\hat{\theta}_n - \theta| \geq \varepsilon) = 0$ $\qquad \hat{\theta}_n \xrightarrow{P} \theta$ for all $\theta \in \Theta$.

Tools: ① WLLN: $\bar{X}_n \xrightarrow{P} \mu$ (same as SLLN)

② Chebyshev's inequality

$$P(\ |X - E(X)| > \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$$

- $\bar{X}_n \xrightarrow{P} \mu$ (iid) $\forall \varepsilon > 0$, $P(|\bar{X}_n - \mu| > \varepsilon) \le \frac{Var(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \to 0$ as $n \to +\infty$

- $X_1, \cdots, X_n$ are r.v. $E(X_i) = \mu_i$, $Var(X_i) = \sigma_i^2$

$$Cov(X_i, X_j) = \sigma_{ij}$$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i, \quad \bar{\mu}_n = \frac{1}{n}\sum_{i=1}^n \mu_i$$

$$P(|\bar{X}_n - \bar{\mu}_n| > \varepsilon) \le \frac{Var(\bar{X}_n)}{\varepsilon^2}$$

as long as $Var(\bar{X}_n) = \frac{1}{n^2}\left[\sum_{i=1}^n \sigma_i^2 + 2\sum_{i<j}\sigma_{ij}\right] \to 0$ as $(n \to +\infty)$

$$\bar{X}_n - \bar{\mu}_n \xrightarrow{P} 0$$

1. if independence holds, $\sigma_{ij} = 0$ then $\frac{1}{n^2}\sum_{i=1}^n \sigma_i^2 \to 0$ is needed
2. if identical $\frac{1}{n^2}\left[n\sigma^2 + 2\sum_{i<j}\sigma_{ij}\right] \to 0$ is needed

   or $\frac{1}{n^2}\sum_{i<j}\sigma_{ij} \to 0$ is needed

Eg: AR(1) $\{X_n\}_{n=-\infty}^{+\infty}$

Stationary

$X_t = \phi X_{t-1} + W_t$ for all t

- W_t's iid mean 0, variance σ^2.
- W_t's are uncorrelated with X_t's.
- $|\phi| < 1$

$$\begin{cases} E[X_t] = \mu_x, \quad Var[X_t] = \sigma_x^2 \\ Cov(X_t, X_{t+h}) = \gamma_x(h) \end{cases}$$

$\{X_1, \cdots, X_n\}$ from AR(1) how to estimate μ_x?

$$E(\bar{X}_n) = \frac{1}{n}\sum_{i=1}^n E(X_i) = \frac{1}{n}\sum_{i=1}^n \mu_x = \mu_x$$

$$Var(\bar{X}_n) \to 0 \Rightarrow \bar{X}_n \xrightarrow{P} \mu_x \,!$$

$$Var(\bar{X}_n) = \frac{1}{n^2}\left[n\sigma_X^2 + 2\sum_{i<j} Cov(X_i, X_j)\right] \qquad \sum_{i<j} = \sum_{i=1}^{n-1}\sum_{j=i+1}^{n}$$

$$= \frac{1}{n^2}\left[n\sigma_X^2 + 2\sum_{i<j} \gamma_X(j-i)\right]$$

$$\sigma_X^2 = Var(X_i) = Cov(X_i, X_i)$$
$$= Cov(\phi X_{i-1} + W_i,\ \phi X_{i-1} + W_i)$$
$$= \phi^2 \underbrace{Cov(X_{i-1}, X_{i-1})} + 2\phi\, Cov(X_{i-1}, \underbrace{W_i}) + Cov(W_i, W_i)$$
$$= \phi^2\sigma_X^2 + 0 + \sigma^2$$
$$\sigma_X^2 = \frac{\sigma^2}{1-\phi^2}$$

$$\gamma_X(h) = Cov(X_t, X_{t-h})$$
$$= Cov(\phi X_{t-1} + W_t, X_{t-h})$$
$$= \phi\, Cov(X_{t-1}, X_{t-h}) + \cancel{Cov(W_t, X_{t-h})}$$
$$= \phi\, \gamma_X(h-1)$$
$$= \phi^2 \gamma_X(h-2) = \cdots = \phi^h \gamma_X(0)$$
$$\gamma_X(h) = \frac{\sigma^2}{1-\phi^2}\phi^h$$

$$Cov(X_i, X_j) = \frac{\sigma^2}{1-\phi^2}\phi^{|j-i|}$$

$$Var(\bar{X}_n) = \frac{1}{n^2}\left[n\sigma_X^2 + 2\sum_{i<j}\frac{\sigma^2}{1-\phi^2}\phi^{j-i}\right]$$
$$= \frac{\sigma^2}{1-\phi^2} \times \underline{\frac{1}{n^2} \times \left[n\sigma^2\right.} + \left.2\sum_{i<j}\phi^{j-i}\right]$$
$$= \frac{\sigma^2}{1-\phi^2}\underset{\to 0}{\frac{\sigma^2}{n}} + \frac{\sigma^2}{1-\phi^2}\underset{\to 0}{\frac{1}{n^2}\, 2\sum_{i<j}\phi^{j-i}} \checkmark$$

$$
\begin{array}{l}
i=1 \\
i=2 \\
\vdots \\
i=n-1
\end{array}
\quad
\begin{array}{ccccc}
\phi & \phi^2 & \phi^3 & \cdots & \phi^{n-1} \\
\hline
 & \phi & \phi^2 & \cdots & \phi^{n-2} \\
 & & \ddots & & \vdots \\
 & & & & \phi
\end{array}
\quad \le \sum_{i=1}^{\infty} |\phi|^i
$$

---

big O_p small o_p notation,

we write $X_n = o_p(Y_n)$ iff. $\frac{X_n}{Y_n} \xrightarrow{P} 0$ as $n \to +\infty$

write $X_n = O_p(Y_n)$

$\forall \varepsilon > 0, \ \exists M > 0$ $\qquad \frac{X_n}{Y_n}$ is bounded in probability

$$P\left(\left| \frac{X_n}{Y_n} \right| > M \right) < \varepsilon \quad \forall n$$

$X_n = o_p(1) \iff \frac{X_n}{1} \xrightarrow{P} 0 \iff X_n \xrightarrow{P} 0$ $\qquad o_p(1) \le O_p(1)$

- $o_p(1) + o_p(1) = o_p(1) \leftarrow$
- $o_p(1) + O_p(1) = O_p(1)$
- $O_p(1) \times O_p(1) = O_p(1)$
- $O_p(1) \times o_p(1) = o_p(1)$
- $(1 + o_p(1))^{-1} = O_p(1)$
- $o_p(R_n) = R_n \, o_p(1)$
- $O_p(R_n) = R_n \, O_p(1)$
- $o_p(O_p(1)) = o_p(1)$

- Multivariate case. $\quad \underset{\sim}{X}_1 = \begin{pmatrix} X_{11} \\ \vdots \\ X_{1p} \end{pmatrix} \ \cdots \ \underset{\sim}{X}_n = \begin{pmatrix} X_{n1} \\ \vdots \\ X_{np} \end{pmatrix}$

$$\underset{\sim}{X}_n \xrightarrow{P} \underset{\sim}{X} \qquad \underset{\sim}{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}$$

$$\forall \varepsilon > 0. \quad \lim_{n\to\infty} P(\|\underset{\sim}{X}_n - \underset{\sim}{X}\| > \varepsilon) = 0 \qquad \|\underset{\sim}{a}\| = \sqrt{a_1^2 + \cdots + a_p^2}$$

$$\Leftrightarrow \quad \forall \varepsilon > 0. \quad \lim_{n\to\infty} P(|X_{nk} - X_k| > \varepsilon) = 0$$

for all $k = 1, \cdots, p$

$$\|\underset{\sim}{a}\| > |a_i|$$

$$\underset{\sim}{X}_n - \underset{\sim}{X} \xrightarrow{P} 0 \iff X_{nk} - X_k \xrightarrow{P} 0 \text{ for all } k = 1, \cdots, p$$

Convergence in Distribution.

Vector version: $\underset{\sim}{X}_n \xrightarrow{d} \underset{\sim}{X}$

if $F_{\underset{\sim}{X}_n}(\underset{\sim}{x}) \to F_{\underset{\sim}{X}}(\underset{\sim}{x})$ at all $\underset{\sim}{x}$ where $F_{\underset{\sim}{X}}$ is continuous

$$\underset{\sim}{X}_n \xrightarrow{d} \underset{\sim}{X} \Rightarrow X_{nk} \xrightarrow{d} X_k \text{ for all } k = 1, \cdots p \quad ? \quad \checkmark \text{ correct)}$$

$P(X_{nk} \le x) \to P(X_k \le x)$ for all x wher F_{X_k} is continuous

$$\underset{\sim}{X}_n \xrightarrow{d} \underset{\sim}{X} \quad \cancel{\Leftarrow} \quad X_{nk} \xrightarrow{d} X_k \text{ for all } k = 1, \cdots p$$

$$\underset{\sim}{X} = \begin{pmatrix} Z_1 \\ Z_2 \end{pmatrix} \qquad \underset{\sim}{X}_n = \begin{pmatrix} Z \\ -Z \end{pmatrix}$$

$$X_1 + X_2 = Z_1 + Z_2 \sim N(0, 2) \not\leftarrow X_{n1} + X_{n2} = 0$$

Example: M-dependence sequence.

AR(1). $X_t = \phi X_{t-1} + W_t \qquad Cov(X_t, X_{t+h}) = \phi^h \times \frac{\sigma^2}{1-\phi^2}$

Moving Average. $\{U_1 \cdots U_n \cdots\}$ are iid $E(U) = 0$, $Var(U) = \sigma^2$ (white noise)

$Cov(U_i, U_j) = 0$

$$X_1 = \frac{U_1 + \cdots + U_m}{m} \qquad X_2 = \frac{U_2 + \cdots + U_{m+1}}{m}, \quad \cdots$$

$$X_m = \frac{U_m + \cdots + U_{2m-1}}{m} \qquad X_{m+1} = \frac{U_{m+1} + \cdots + U_{2m}}{m}$$

$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} 0$ ✓

$$Var(\bar{X}_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j} cov(X_i, X_j)\right)$$

$$= \frac{1}{n^2}\left(n \times \frac{\sigma^2}{m} + 2\sum_{k=1}^{m}\sum_{j-i=1}^{k} cov(X_i, X_j)\right) \to 0 \quad (\text{as } n \to +\infty)$$

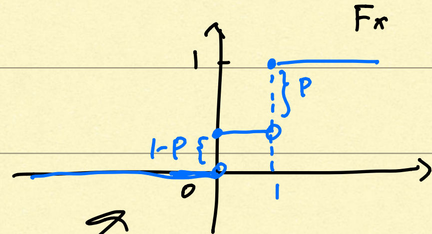
Convergence in Distribution (in Law, weakly convergence)

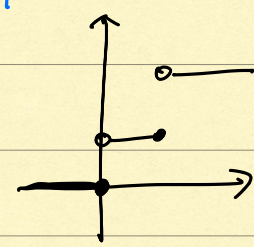
$X_n \xrightarrow{d} X$

~~$P(|X_n - X| > \varepsilon) \to 0$?~~

" $F_{X_n}(x) \to F(x)$ as $n \to +\infty$ for any x where $F(x)$ is continuous "

Example $X_n = X + \frac{1}{n}$ X is Bernulli (p) $\Rightarrow$ $P(X=1) = p$, $P(X=0) = 1-p$

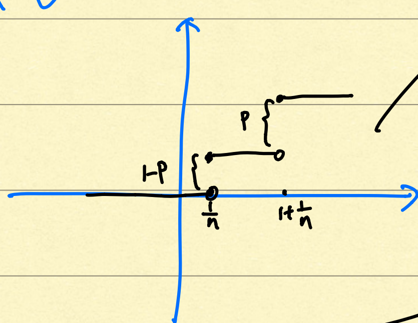




$X_n \xrightarrow{a.s.} X$

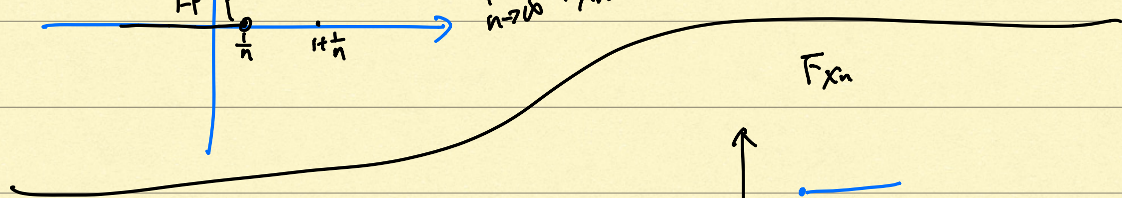
$X_n \xrightarrow{d} X$ ✓

F_{X_n}



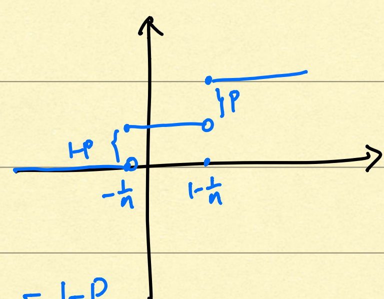
$F_{X_n}(0) = P(X_n \le 0) = P(X \le -\frac{1}{n}) = 0$

$\lim_{n\to\infty} F_{X_n}(0) = 0$



F_{X_n}

$X_n = X - \frac{1}{n}$ $X_n \xrightarrow{d} X$



$F_{X_n}(0) = P(X_n \le 0) = P(X \le \frac{1}{n}) = 1-p$

Example. $X_n \sim N(0, \sigma_n^2)$

$$Y_n = \frac{X_n}{\sigma_n}$$

Does $Y_n \xrightarrow{d} N(0,1)$? if $\sigma_n \to +\infty$ as $n \to +\infty$

$X_n \xrightarrow{d}$?

$\forall x$: $P(X_n \le x) = P\left(\frac{X_n}{\sigma_n} \le \frac{x}{\sigma_n}\right) = P\left(\underset{N(0,1)}{Z} \le \frac{x}{\sigma_n}\right) \to \Phi(0) = \frac{1}{2}$

$P(X_n > x) = \frac{1}{2}$

$$X = \begin{cases} +\infty & \frac{1}{2} \text{ prob} \\ -\infty & \frac{1}{2} \text{ prob} \end{cases}$$

$x \to -\infty$ $\quad \lim_{n \to +\infty} F_{X_n}(x) \not\to 0$

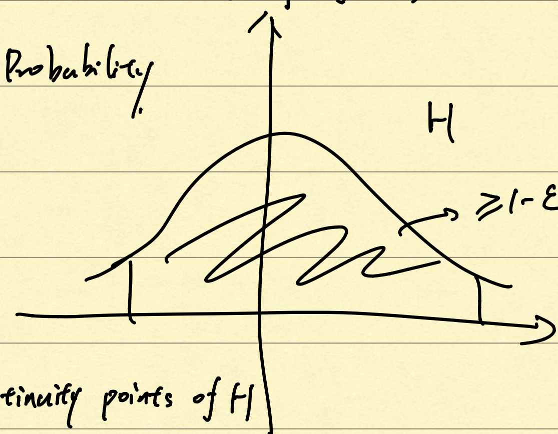
$x \to +\infty$ $\quad \lim_{n \to +\infty} F_{X_n}(x) \not\to 1$

$O_p(1)$ bounded in Probability $\quad X = O_p(1)$

$\forall \varepsilon > 0, \ \exists M \quad P(|X| < M) > 1 - \varepsilon$

Theorem suppose $X_n \xrightarrow{d} X$ following distribution H (cdf of X)

Then. X_n is bounded in Probability



Proof: $\forall \varepsilon > 0$. find M, n_0 st. goal.

$P(|X_n| < M) > 1 - \varepsilon$ for $n > n_0$

because H is a cdf $\quad H(-\infty) = 0$, $H(+\infty) = 1$

$\forall \varepsilon > 0$, we can find M_1, M_2. which are continuity points of H

$H(M_1) > 1 - \frac{\varepsilon}{4}$, $\quad H(-M_2) < \frac{\varepsilon}{4}$

$X_n \xrightarrow{d} X \Rightarrow H_n(M_1) \to H(M_1) \quad H_n(-M_2) \to H(-M_2)$ as $n \to +\infty$

$\exists n_0 > 0. \quad \forall n > n_0$

$$H_n(M_1) > H(M_1) - \frac{\varepsilon}{4} > 1 - \frac{\varepsilon}{2}, \quad H_n(-M_2) < H(-M_2) + \frac{\varepsilon}{4} < \frac{\varepsilon}{2}$$

Take $M = \max(|M_1|, |M_2|)$

$$P(|X_n| < M) \geq P(-M_2 < X_n < M_1)$$

$$= H_n(M_1) - H_n(-M_2)$$

$$> 1 - \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = 1 - \varepsilon \quad \#$$

Lemma If $X_n \xrightarrow{d} X$. a, b are constant.

Then $bX_n + a \xrightarrow{d} bX + a$

Theorem. If $X_n \xrightarrow{d} X$. $A_n \xrightarrow{P} a$, $B_n \xrightarrow{P} b$. where a, b are constants

then $A_n + B_n X_n \xrightarrow{d} a + bX$.

$$X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

represented version

(skorokhod's representative theorem)

If $X_n \xrightarrow{d} X$. there exists a probability space $(\tilde{S}, \tilde{Q}, \tilde{P})$

and $\tilde{X}_n, \tilde{X}$ defined on

such that. $\tilde{X}_n \stackrel{d}{=} X_n$, $\tilde{X} \stackrel{d}{=} X$

and $\tilde{X}_n \xrightarrow{a.s.} \tilde{X}$

Equivelent definitions of $X_n \xrightarrow{d} X$

① $\underline{E f(X_n) \to E f(X)}$ for all bounded and $\underline{\text{continuous}}$ functions f.

Example: $Y_n = \begin{cases} Y & \text{with prob. } 1-p_n \\ n & \text{with prob } p_n \end{cases}$

$\forall y$ $P(Y_n \le y) = (1-p_n) P(Y \le y) + p_n P(n \le y)$

$= (1-p_n) P(Y \le y)$ when $n > y$

$= P(Y \le y)$ if $p_n \to 0$

$Y_n \xrightarrow{d} Y$

$f(y) = y$ for all y

$E f(Y_n) = E Y_n = \underbrace{(1-p_n) E Y}_{\downarrow\ E(Y)} + \underbrace{p_n n}_{?}$

If $p_n = \frac{1}{n}$. $E f(Y_n) \to E(Y) + 1$

Example. Suppose $Y_n \xrightarrow{d} Y$ the cdf of Y is H

H has a $\underline{\text{discontinuity}}$ at a.

$f(y) = \begin{cases} 1 & \text{if } y \le a \\ 0 & \text{if } y > a. \end{cases}$

$E f(Y_n) = P(Y_n \le a)$ $\downarrow$? may not happen.

$E f(Y) = P(Y \le a)$

- $Ef(X_n) \to Ef(X)$ for all bounded, Lipschitz continuous functions f.

  $|f(x)-f(x')| < L\,|x-x'|$ for all x, x'.

- $\liminf Ef(X_n) \geq Ef(X)$ for all nonnegative continuous functions f.
- $\liminf P(X_n \in G) \geq P(X \in G)$ for all open set G
- $\limsup P(X_n \in F) \leq P(X \in F)$ for all closed set F
- $P(X_n \in B) \to P(X \in B)$ for all Borel set B s.t. $P(X \in \text{boundary of } B) = 0$

- Characteristic function (cf) of a r.v. X

  is $\phi_X(t) = E[e^{itX}]$; $\phi_X(0) = 1$

  $|\phi_X(t)| \leq E[|e^{itx}|] = E(1) = 1$

- $\underline{X_n \xrightarrow{d} X} \iff \phi_{X_n}(t) \to \phi_X(t)$ pointwisely at t.

  $\hookrightarrow$ $\lim_{n\to\infty} F_{X_n}(x)$ is a valid cdf.

  Moreover: if $\phi_{X_n}(t) \xrightarrow{\text{pointwisely}} \phi(t)$ (which is continuous at $t=0$)

  then $\phi(t)$ is a valid cf of some r.v. X.

  and $X_n \xrightarrow{d} X$

- $X_1, \cdots, X_n \overset{iid}{\sim}$ cf. $\phi_X(t)$

  then $\bar{X}_n \xrightarrow{P} \theta$ iff ϕ_X is differentiable at 0

  and $i\theta = \phi_X'(0)$

Central Limit theorem: $X_1, \cdots, X_n \overset{iid}{\sim}$ $E(X)=\mu$, $Var(X)=\sigma^2$ are finite

$$\frac{\bar{X}_n - E(X)}{\sqrt{Var(X)/n}} \xrightarrow{d} N(0,1) \quad \text{as } n \to +\infty$$

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

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

Proof: WOLG: assume $\mu = 0$. $\sigma^2 = 1 = E(X^2) - \mu^2$

$$\phi_X(t) = E[e^{itX}]$$

$$\phi_X'(t) = \left(E[e^{itX}]\right)' = E[(e^{itX})']$$

$$= E[ixe^{itx}]$$

$$\to \phi_X'(0) = E[ixe^{i0x}] = E[ix] = 0$$

$$\phi_X''(t) = E[(ixe^{itx})']$$

$$= E[(-1)x^2 e^{itx}]$$

$$\to \phi_X''(0) = -E(X^2) = -1$$

$\phi_X(0) = E[e^{i0x}] = 1$

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

Goal: show $\sqrt{n}\,\bar{X}_n \xrightarrow{d} N(0,1)$

$$\phi_{\sqrt{n}\bar{X}_n}(t) = E[e^{it\sqrt{n}\bar{X}_n}]$$

$$= E\left[e^{it\sqrt{n} \times \frac{1}{n}\sum_{i=1}^{n} X_i}\right]$$

$$= E\left[e^{it\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i}\right]$$

$$\overset{iid}{=} E\left[\prod_{i=1}^{n} e^{it\frac{1}{\sqrt{n}}X_i}\right]$$

$$\overset{iid}{=} \prod_{i=1}^{n} E\left[e^{it\frac{1}{\sqrt{n}}X_i}\right]$$

$$= \left[\underbrace{\phi_X\left(\frac{t}{\sqrt{n}}\right)}\right]^n \qquad \phi_X(t) = E[e^{itX_i}]$$

Taylor's expance

$$= \left(\phi_X(0) + \left(\frac{t}{\sqrt{n}}\right)\phi_X'(0) + \left(\frac{t}{\sqrt{n}}\right)^2 \frac{1}{2}\phi_X''(0) + o\left(\frac{1}{n}\right)\right)^n$$

$$= \left(1 + \frac{t}{\sqrt{n}} \times 0 + \frac{t^2}{2n}(-1) + o\left(\frac{1}{n}\right)\right)^n$$

$$= \left(1 - \frac{1}{2}\frac{t^2}{n} + o\left(\frac{1}{n}\right)\right)^n \qquad \left(1+\frac{a}{n}\right)^n \to e^a$$

$(n \to +\infty)$ $\phi_{\sqrt{n}\bar{X}_n}(t) \longrightarrow e^{-\frac{1}{2}t^2}$ for all t (pointwisely)

$\qquad\qquad ||$

characteristic function of $N(0,1)$

Thus: $\sqrt{n}\bar{X}_n \xrightarrow{d} N(0,1)$ #

- Example: $X_1, \cdots, X_n \overset{iid}{\sim}$ Bernulli (p). $p \in (0,1)$

by CLT: $\dfrac{\sqrt{n}(\bar{X}_n - p)}{\sqrt{pq}} \xrightarrow{d} N(0,1)$ as $n \to +\infty$

Berry-Essean Theorem: $X_1, \cdots, X_n$ $\overset{iid}{\sim}$ distribution F, which has a finite third moment $E|X-E(X)|^3 < +\infty$

then, there exists a constant C (independent of F) such that for all x,

$$\left| P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x \right) - \Phi(x) \right| \leq \frac{C}{\sqrt{n}} \cdot \frac{E|X-E(X)|^3}{\sigma^3}$$

$\Phi(x)$: cdf of $N(0,1)$

not stochastically

remark: holds for $C = .7975$ does not hold for $C < .4097$

- Let $X_1, \cdots, X_n$ $\overset{iid}{\sim}$ distribution F which is not a lattic distribution and has a finite third moment. Then.

$$P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \leq x \right) = \Phi(x) + \frac{E(X-E(X))^3}{6\sigma^3\sqrt{n}}(1-x^2)\phi(x) + o\left(\frac{1}{\sqrt{n}}\right)$$

$\phi(x)$: density of $N(0,1)$

Example: $X_1 \sim \text{Bernulli}(P_1)$

$X_1, X_2 \sim \text{Bernulli}(P_2)$

$X_1 \cdots X_3 \sim \text{Bernoulli}(P_3)$

$\vdots$

$X_1, \cdots, X_n \overset{iid}{\sim} \text{Bernulli}(P_n)$

$$\frac{\sqrt{n}(\bar{X}_n - P_n)}{\sqrt{P_n q_n}} \xrightarrow{d} N(0,1) ?$$

Corollary: $X_1, \cdots, X_n \overset{iid}{\sim} F_n$ with a finite third moment

$$P\left(\frac{\sqrt{n}(\bar{X}_n - \mu_n)}{\sigma_n} \leq x \right) \longrightarrow \Phi(x) \checkmark$$

for all sequence F_n with mean μ_n and variance σ_n^2

for which $E_n|X-\mu_n|^3$

$$\frac{\quad}{\sigma_n^3} = o(\sqrt{n})$$

Continued:

$$\frac{E_n|X-\mu_n|^3}{\sigma_n^3} = o(\sqrt{n}) \qquad X \sim \text{Bernoulli}(P_n)$$

$\sigma_n = \sqrt{P_n q_n}$ $\qquad q_n = 1-P_n$

$\mu_n = P_n$

$$E_n|X-\mu_n|^3 = E|X-P_n|^3 = P_n^3 \times P(X=0) + (1-P_n)^3 \times P(X=1)$$

$$= P_n^3 q_n + q_n^3 P_n$$

$$= P_n q_n (P_n^2 + q_n^2)$$

$$\frac{P_n q_n (P_n^2+q_n^2)}{(\sqrt{P_n q_n})^3} = \frac{P_n^2+q_n^2}{\sqrt{P_n q_n}} = o(\sqrt{n})$$

If $P_n \to P$ as $n \to +\infty$ where $0<p<1$

$$\frac{P_n^2+q_n^2}{\sqrt{P_n q_n}} \to \frac{p^2+q^2}{\sqrt{pq}} = o(\sqrt{n}) \qquad q=(1-p)$$

If $P_n \to 0$, as $n \to +\infty$ $\quad \frac{P_n^2+q_n^2}{\sqrt{P_n q_n}}$ is of order $\left(\frac{1}{\sqrt{P_n}}\right) = o(\sqrt{n})$

$\longrightarrow$ ~~CLT~~ $\left\{ \begin{array}{l} \text{If } P_n = \frac{\lambda}{n} \text{ for } \lambda>0 \\ \left(\frac{1}{\sqrt{P_n}}\right) = \frac{1}{\sqrt{\lambda}} \times \sqrt{n} \neq o(\sqrt{n}) \end{array} \right.$

If $P_n = \frac{\lambda}{n^\alpha}$, $0<\alpha<1$

Then $\left(\frac{1}{\sqrt{P_n}}\right) = \frac{1}{\sqrt{\lambda}} \cdot (\sqrt{n})^\alpha = o(\sqrt{n})$

CLT ✓

$X_1, \dots, X_n \overset{iid}{\sim} \text{Bernoulli}\left(\frac{\lambda}{n}\right)$,

$$\frac{\sqrt{n}(\bar{X}_n - \frac{\lambda}{n})}{\sqrt{\frac{\lambda}{n}(1-\frac{\lambda}{n})}} \overset{d}{\not\to} N(0,1)$$

$X_1, \cdots, X_n \overset{iid}{\sim} \text{Bernoulli}\left(\frac{\lambda}{\sqrt{n}}\right)$

$$\frac{\sqrt{n}\left(\bar{X}_n - \frac{\lambda}{\sqrt{n}}\right)}{\sqrt{\frac{\lambda}{\sqrt{n}}\left(1-\frac{\lambda}{\sqrt{n}}\right)}} \xrightarrow{d} N(0,1)$$

Example. (Sample Median) $X_1, \cdots, X_n \overset{iid}{\sim}$ the distribution

$P(X \le x) = F(x-\theta)$ where $F(0) = \frac{1}{2}$.

when $x = \theta$. $F(x-\theta) = F(0) = \frac{1}{2} = P(X \le \theta)$

θ is the population median of F

$n = 2m-1$

$X_1;\ X_2, \cdots;\ X_{m-1}, X_m, X_{m+1}, \cdots;\ X_{2m-1}$

$X_{(1)} \le X_{(2)} \le \cdots \le X_{(m-1)} \le X_{(m)} \le \cdots \le X_{(2m-1)}$

$\uparrow$

Sample median

$\hat{\theta} = X_{(m)}$

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} ?$$

$$P\left(\sqrt{n}(\hat{\theta}-\theta) \le x\right)$$

$$= P\left(\sqrt{n}\,\hat{\theta} \le x + \sqrt{n}\,\theta\right)$$

$$= P\left(\hat{\theta} \le \frac{x+\sqrt{n}\,\theta}{\sqrt{n}}\right)$$

$$= P\left(\hat{\theta} \le \frac{x}{\sqrt{n}} + \theta\right)$$

Define

$$Y_i = \begin{cases} 1 & \text{if } X_i \le \frac{x}{\sqrt{n}} + \theta \\ 0 & \text{otherwise} \end{cases}$$

Y_i iid Bernoulli (P_n) this is because X_i's are iid.

$$P_n = P(Y_i = 1) = P\left(X_i \le \frac{x}{\sqrt{n}} + \theta\right)$$
$$= F\left(\frac{x}{\sqrt{n}} + \theta - \theta\right)$$
$$= F\left(\frac{x}{\sqrt{n}}\right)$$

as $n \to +\infty$, $P_n \to F(0) = \frac{1}{2}$ ✓

Our goal is to study $P\left(\hat{\theta} \le \frac{x}{\sqrt{n}} + \theta\right) = P\left(X_{(m)} \le \frac{x}{\sqrt{n}} + \theta\right)$

$X_{(m)} \le \frac{x}{\sqrt{n}} + \theta$ is equivallent to that

there are at least m X_i's less than $\frac{x}{\sqrt{n}} + \theta$

i.e. $S_n = \sum_{i=1}^{n} Y_i \ge m$

Thue $P\left(\hat{\theta} \le \frac{x}{\sqrt{n}} + \theta\right) = P(S_n \ge m) = 1 - P(S_n \le m-1)$

$$= 1 - P\left(S_n \le \frac{n-1}{2}\right)$$

From the previous example and the last corollary, we know that

$$P\left(\frac{\sqrt{n}\left(\frac{S_n}{n} - P_n\right)}{\sqrt{P_n q_n}} \le x\right) - \underset{\text{cdf of } N(0,1)}{\Phi}(x) \to 0 \quad \text{for all } x. \quad \text{as } n \to +\infty$$

known $P\left(S_n \le \frac{n-1}{2}\right) = P\left(\frac{\sqrt{n}\left(\frac{S_n}{n} - P_n\right)}{\sqrt{P_n q_n}} \le \frac{\sqrt{n}\left(\frac{n-1}{2n} - P_n\right)}{\sqrt{P_n q_n}}\right)$,

thus $P\left(S_n \le \frac{n-1}{2}\right) - \Phi\left(\underbrace{\frac{\sqrt{n}\left(\frac{n-1}{2n} - P_n\right)}{\sqrt{P_n q_n}}}_{x_n}\right) \to 0$ as $n \to +\infty$

$$x_n = \sqrt{n}\left(\frac{1}{2} - P_n - \frac{1}{2n}\right) \Big/ \sqrt{P_n q_n}$$

$p_n = F\left(\frac{x}{\sqrt{n}}\right) \quad \frac{1}{2} = F(0)$, and $p_n \to \frac{1}{2}$ as $n \to +\infty$

$$x_n = \frac{\sqrt{n}\left(F(0) - F\left(\frac{x}{\sqrt{n}}\right) - \frac{1}{2n} \right)}{\sqrt{p_n q_n}}$$

$$= \frac{\frac{F(0) - F\left(\frac{x}{\sqrt{n}}\right)}{x/\sqrt{n}} \times x}{\sqrt{p_n q_n}} + o(1)$$

$$\to -2xf(0) \quad \text{as } n \to +\infty$$

Thus, $P\left(S_n \leq \frac{n-1}{2}\right) \to \Phi\left(-2xf(0)\right)$

i.e, $P\left(\hat{\theta} \leq \frac{x}{\sqrt{n}} + \theta\right) = P(S_n > m) = 1 - P(S_n \leq m-1) = 1 - P\left(S_n \leq \frac{n-1}{2}\right)$

$$\to 1 - \Phi\left(-2xf(0)\right)$$

$$= \Phi\left(2xf(0)\right)$$

$$= P\left(Z \leq 2xf(0)\right)$$

$$= P\left(\frac{Z}{2f(0)} \leq x\right)$$

$$= P\left(N\left(0, \frac{1}{4f^2(0)}\right) \leq x\right) \quad \forall x$$

Thus $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N\left(0, \frac{1}{4f^2(0)}\right)$

In general. Let $\hat{\theta}_\tau$ be the sample τ-th quantile

θ_τ be the population τ-th quantile; i.e. $F^{-1}(\tau) = \theta_\tau$

Then $\sqrt{n}(\hat{\theta}_\tau - \theta_\tau) \xrightarrow{d} N\left(0, \tau(1-\tau) \cdot \frac{1}{f^2(\theta_\tau)}\right)$

provided $f(\theta_\tau) > 0$.

Delta Method

Delta Method Suppose $\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \tau^2)$ as $n \to +\infty$ If $f'(\theta) \neq 0$ exists

then $\sqrt{n}(f(T_n) - f(\theta)) \xrightarrow{d} N(0, \tau^2 [f'(\theta)]^2)$ as $n \to +\infty$

what if $f'(\theta) = 0$, $f''(\theta)$ exists $f''(\theta) \neq 0$

$n(f(T_n) - f(\theta)) \xrightarrow{d} \frac{1}{2} \tau^2 \chi_1^2 f''(\theta)$ ✓ as $n \to +\infty$

$f(T_n) \approx f(\theta) + f'(\theta)(T_n - \theta) + \frac{1}{2} f''(\theta)(T_n - \theta)^2 + o(\quad)$

$f(T_n) - f(\theta) \approx \frac{1}{2} f''(\theta)(T_n - \theta)^2 + \ldots$

$\sqrt{n}(T_n - \theta) \xrightarrow{d} N(0, \tau^2) = \tau Z$

$n(T_n - \theta)^2 \xrightarrow{d} \tau^2 \chi_1^2$

$n[f(T_n) - f(\theta)] \xrightarrow{d} \frac{1}{2} f''(\theta) \tau^2 \chi_1^2$ as

$f'(\theta)=0$

$f(T_n)-f(\theta)$

Variance Stablization

extra approximation

$\tau(\theta) \leftarrow \tau(\hat{\theta})$

$\sqrt{n}(T_n-\theta) \xrightarrow{d} N(0,\tau^2)$

$\left[T_n \pm z_{\alpha/2}\cdot\frac{\tau}{\sqrt{n}}\right]$

Find f such that

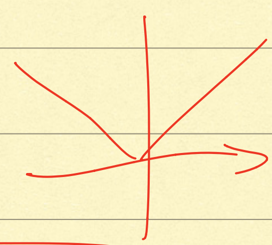
$\sqrt{n}\left(f(T_n)-f(\theta)\right) \xrightarrow{d} N\left(0, \tau^2(\theta)[f'(\theta)]^2\right)$

where $\tau^2(\theta)\left(f'(\theta)\right)^2 = c$, a constant

---

Example Absolute Vale. $\sqrt{n}(T_n-\theta) \xrightarrow{d} N(0,\tau^2)$

$f(x)=|x|$



$\sqrt{n}(|T_n|-|\theta|) \xrightarrow{d}$?

If $\theta>0$. $f'(\theta)=1$ $\quad \sqrt{n}(|T_n|-|\theta|) \xrightarrow{d} N(0,\tau^2)$

If $\theta<0$ $f'(\theta)=-1$ $\quad \sqrt{n}(|T_n|-|\theta|) \xrightarrow{d} N(0,\tau^2)$

If $\theta = 0$

$$P\left(\sqrt{n}\left(|T_n| - |\theta|\right) \le x\right)$$

$$\sqrt{n}(T_n - \theta) = \sqrt{n}(T_n) \xrightarrow{d} N(0, \tau^2)$$

$$= P\left(\sqrt{n}\,|T_n| \le x\right)$$

$$\frac{\sqrt{n}\,T_n}{\tau} \xrightarrow{d} N(0,1)$$

$$= P\left(\sqrt{n}\left|\frac{T_n}{\tau}\right| \le \frac{x}{\tau}\right)$$

$$= P\left(|Z| \le \frac{x}{\tau}\right) \quad \#$$

$\sqrt{n}\,|T_n| \xrightarrow{d} \tau \cdot |Z|$. where $Z \sim N(0,1)$

- CLT for non iid case

① independent but not identical

$X_1 \sim \text{Bernoulli}(P_1)$, $\cdots$, $X_n \sim \text{Bernoulli}(P_n)$ independent

$E(X_i) = P_i$

$Var(X_i) = P_i q_i$ where $q_i = 1 - P_i$

Define $Y_i = X_i - q_i$

$T_n = \sum_{i=1}^{n} Y_i$ $\quad E(T_n) = 0.$ $\quad Var(T_n) = \sum_{i=1}^{n} P_i q_i = S_n^2$

Does $\frac{T_n}{S_n} \xrightarrow{d} N(0,1)$?

No, but under what condition, it does?

(R codes)

Two conditions:

Notation: $X_1, \cdots X_n$ are independent random variables

such that $E(X_i) = \mu_i$, $Var(X_i) = \sigma_i^2 < +\infty$

Define $Y_i = X_i - \mu_i$

$T_n = \sum_{i=1}^{n} Y_i$ $\quad E(T_n) = 0$

$S_n^2 = Var(T_n) = \sum_{i=1}^{n} \sigma_i^2$

If either

Lindeberg Condition:

$\forall \varepsilon > 0$, $\quad \frac{1}{S_n^2} \sum_{i=1}^{n} E\left(Y_i^2 I\{|Y_i| > \varepsilon S_n\}\right) \to 0$ as $n \to +\infty$ (LinC)

or Lyapunov Condition,

$\exists \delta > 0$. s.t. $\quad \frac{1}{S_n^{2+\delta}} \sum_{i=1}^{n} E\left(|Y_i|^{2+\delta}\right) \to 0$ as $n \to +\infty$ (Lya C)

holds, $\quad \frac{T_n}{S_n} \xrightarrow{d} N(0,1)$ as $n \to +\infty$

Thm: The Lyapunov condition (LyaC) implies the Lindeberg condition (LinC)

Proof:

$\forall \varepsilon > 0$, if $|Y_i| > \varepsilon S_n$

then $\left| \frac{Y_i}{\varepsilon S_n} \right| > 1$

for $\delta > 0$, $\left| \frac{Y_i}{\varepsilon S_n} \right|^\delta > 1$

Then
$$\frac{1}{S_n^2} \sum_{i=1}^{n} E\left(Y_i^2 \, I\{|Y_i| \geq \varepsilon S_n\} \right)$$
$$\leq \frac{1}{S_n^2} \sum_{i=1}^{n} E\left(Y_i^2 \, I\{|Y_i| \geq \varepsilon S_n\} \times \frac{|Y_i|^\delta}{|\varepsilon S_n|^\delta} \right)$$
$$= \frac{1}{\varepsilon^\delta S_n^{2+\delta}} \cdot \sum_{i=1}^{n} E\left(|Y_i|^{2+\delta} \cdot I\{|Y_i| \geq \varepsilon S_n\} \right)$$
$$\leq \underbrace{\frac{1}{\varepsilon^\delta S_n^{2+\delta}} \sum_{i=1}^{n} E|Y_i|^{2+\delta}}_{(LypC)} \to 0 \quad \text{as } n \to +\infty$$

(Special Case) take $\delta = 1$ in (LyP C)

If $\sum_{i=1}^{n} E|Y_i|^3 = o(S_n^3)$

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

Corollary. X_i independent with $E(X_i)=\mu_i$, $Var(X_i)=\sigma_i^2$.

if X_i's are uniformly bounded.

i.e, there exists a cons, A, $|X_i|\le A$ for all i

Then $\sum_{i=1}^{n} E|Y_i|^3 = o(S_n^3)$ provided $S_n^2=\sum_{i=1}^{n}\sigma_i^2 \to \infty$

Proof: $\sum_{i=1}^{n}|Y_i|^3 = \sum_{i=1}^{n}|X_i-\mu_i|^3 = \sum_{i=1}^{n}|X_i-\mu_i|(X_i-\mu_i)^2$

$(|X_i-\mu_i|\le|X_i|+|\mu_i|\le 2A)$ $\le 2A\sum_{i=1}^{n}(X_i-\mu_i)^2$

Take expectation

$$\sum_{i=1}^{n}E|Y_i|^3 \le 2A\cdot S_n^2 = o(S_n^3) \text{ if } S_n\to+\infty$$ #

Back to the Bernoulli example $X_i\sim$ Bernoulli (P_i) independently

$|X_i|\le 1$ for all i

Then $\frac{Y_n}{S_n}\xrightarrow{d} N(0,1)$ provided $S_n\to+\infty$

$$S_n=\sum_{i=1}^{n}P_iq_i=\sum_{i=1}^{n}(P_i-P_i^2)$$

If $P_i=\frac{1}{i^k}$, $k>1$, Then $S_n\not\to+\infty$

If $P_i=\frac{1}{i}$, Then $S_n\to+\infty$

or if $P_i\to 0<a<1$, as $i\to+\infty$, then $S_n\to+\infty$

meaning that when n is large X_i's are like iid.

Connection with iid case.

If $X_1, \cdots, X_n$ are iid, we know $\frac{T_n}{S_n} \xrightarrow{d} N(0,1)$

$$\frac{T_n}{S_n} = \frac{\sum_{i=1}^{n}(X_i - \mu_i)}{\sqrt{n\sigma^2}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sqrt{\sigma^2}} \xrightarrow{d} N(0,1)$$

Does (Lin C) hold? Yes.

$\forall \varepsilon > 0.$

$$\frac{1}{S_n^2}\sum_{i=1}^{n} E\left(Y_i^2 I(|Y_i| \geq \varepsilon \sigma_n)\right)$$

$$= \frac{1}{n\sigma^2}\sum_{i=1}^{n} E\left(Y_i^2 I(|Y_i| \geq \varepsilon \sigma\sqrt{n})\right)$$

$$= \frac{1}{\sigma^2} E\left(Y_1^2 I(|Y_1| \geq \varepsilon \sigma\sqrt{n})\right)$$

$$n \to +\infty. \quad Y_1^2 I(|Y_1| \geq \varepsilon \sigma\sqrt{n}) \downarrow 0$$

by monotone convergence theorem or DCT.

$$E(Y_1^2 I(|Y_1| \geq \varepsilon \sigma\sqrt{n})) \to 0 \text{ as } n \to +\infty$$

In general, $X_i \sim \mu_i, \sigma_i^2$ independently

$\frac{T_n}{S_n} \xrightarrow{d} N(0,1)$ does not imply (Lin C)!

Motivating Example

$X_i \sim E(X_i) = \mu_i, Var(X_i) = \sigma_i^2$ where $\sigma_i^2 < +\infty$ is know.

How to estimate μ?

$\frac{2(X_i - \mu)}{\sigma_i^2}$

weighted least square

$\hat{\mu}$ minimizes $\sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma_i^2}$ $\quad \hat{\mu} = \frac{\sum_{i=1}^{n} \frac{X_i}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$

So $\hat{\mu} = \sum_{i=1}^{n} W_{in} X_i \qquad W_{in} = \dfrac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$ for $i=1,\cdots,n$.

$$E(\hat{\mu}) = \mu$$

$$Var(\hat{\mu}) = \sum_{i=1}^{n} W_{in}^2 \, Var(X_i) = \sum_{i=1}^{n} \frac{\frac{1}{\sigma_i^4}}{\left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)^2} \sigma_i^2 = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}$$

does $\dfrac{\hat{\mu} - \mu}{\sqrt{\frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}}} = \sqrt{\sum_{i=1}^{n} \frac{1}{\sigma_i^2}}\,(\hat{\mu} - \mu) \xrightarrow{d} N(0,1)$ as $n \to +\infty$?

Define
$$T_n = \sum_{i=1}^{n} W_{ni}(X_i - \mu) = \sum_{i=1}^{n} Y_{ni}$$

where $Y_{ni} = W_{ni}(X_i - \mu)$

$n=1. \quad Y_{11} = W_{11}(X_1 - \mu) \qquad \sim F_{11}$

$n=2. \quad Y_{21} = W_{21}(X_1 - \mu) \quad Y_{22} = W_{22}(X_2 - \mu) \qquad \sim F_{21}, F_{22}$ (indp)

$\vdots \qquad \ddots$

$n=n \quad Y_{n1} = W_{n1}(X_1 - \mu) \; - - - - \; Y_{nn} = W_{nn}(X_n - \mu)$

$\sim F_{n1}, \cdots, F_{nn}$ (indp)

How to handle this type of structure for CLT?

Another Motivating Example. (simple regression.)

$$Y_i = \alpha + \beta x_i + \varepsilon_i$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y}_n)(X_i - \bar{X}_n)}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2} = \sum_{i=1}^{n} \frac{(X_i - \bar{X}_n)}{\sum_{i=1}^{n}(X_i - \bar{X}_n)^2} Y_i = \sum_{i=1}^{n} w_{ni} Y_i$$

General Setting:

$$\begin{array}{llllll} X_{11} & & & & \sim F_{11} & \\ X_{21} & X_{22} & & & \sim F_{21}, F_{22} & \text{independently} \\ X_{31}, & X_{32} & X_{33} & & \sim F_{31}, F_{32}, F_{33}, & \text{indep...} \\ \vdots & & & & & \\ X_{n1} & - - & - - & - X_{nn} & \sim F_{n1}, F_{n2}, \cdots, F_{n3} & \text{indep...} \end{array}$$

$E(X_{ni}) = \mu_{ni}$, $\quad Y_{ni} = X_{ni} - \mu_{ni}$ $\quad Var(X_{ni}) = Var(Y_{ni}) = \sigma_{ni}^2$

$$T_n = \sum_{i=1}^{n} Y_{ni}$$

$$S_n^2 = Var(T_n) = \sum_{i=1}^{n} \sigma_{ni}^2$$

The Lindeberg-Feller theorem: $\quad \frac{T_n}{S_n} \xrightarrow{d} N(0,1)$

provided the Lindeberg Condition holds: $\quad \forall \varepsilon > 0$

$$\frac{1}{S_n^2} \sum_{i=1}^{n} E\left\{ Y_{ni}^2 I(|Y_{ni}| \geq \varepsilon S_n) \right\} \to 0 \quad \text{as } n \to +\infty \qquad (Lin\,C)$$

Conversely, if $\frac{\max_{i \leq n} \sigma_{ni}^2}{S_n^2} \to 0$ as $n \to +\infty$ and if $\frac{T_n}{S_n} \xrightarrow{d} N(0,1)$

then the Lindeberg Condition $(Lin\,C)$ holds.

We know that the Lyapunov condition implies the Lindeberg Condition in the independent but not indentically distributed case.

Same here, the Lyapunov condition now becomes

$$\exists \delta > 0. \text{ s.t. } \frac{1}{S_n^{2+\delta}} \sum_{i=1}^{n} E\left(|Y_{ni}|^{2+\delta} \right) \to 0 \quad \text{as } n \to +\infty$$

if the Lyapunov condition holds, the $\frac{T_n}{S_n} \xrightarrow{d} N(0,1)$ as $n \to +\infty$

In general, we take $\delta = 1$. i.e., we need

$$\frac{1}{S_n^3} \sum_{i=1}^{n} E(|Y_{ni}|^3) \to 0$$

Theorem: (weighted averge) Let $Y_1, \cdots, Y_n$ be iid with $E(Y_i)=0$, $Var(Y_i)=\sigma^2>0$ and $E|Y_i^3| = \gamma < +\infty$. Then

$$\frac{\sum_{i=1}^{n} W_{ni} Y_i}{\sigma \sqrt{\sum_{i=1}^{n} W_{ni}^2}} \xrightarrow{d} N(0,1)$$

provided that $\left(\sum_{i=1}^{n} |W_{ni}|^3 \right)^2 = o\left(\sum_{i=1}^{n} W_{ni}^2 \right)^3$. $\quad (\ast)$

Proof: $T_n = \sum_{i=1}^{n} Y_{ni}^{*}$ where $Y_{ni}^{*} = W_{ni} Y_i$

$S_n^2 = \sum_{i=1}^{n} Var(Y_{ni}^{*}) = \sigma^2 \sum_{i=1}^{n} W_{ni}^2$

by taking $\delta=1$, Lyapunov Condition asks for

$$\sum_{i=1}^{n} E(|Y_{ni}^{*}|^3) = o(S_n^3)$$

or $\left[\sum_{i=1}^{n} E(|Y_{ni}^{*}|^3) \right]^2 = o\left((S_n^2)^3 \right) \quad (\ast)$

because $(S_n^2)^3 = (\sigma^2)^3 \left(\sum_{i=1}^{n} W_{ni}^2 \right)^3$

$$E|Y_{ni}^{*}|^3 = E|W_{ni}|^3 \times |Y_i|^3$$
$$= |W_{ni}|^3 \gamma,$$

we have $(\ast)$ equivalent to

$$\left(\sum_{i=1}^{n} |W_{ni}|^3 \right)^2 = o\left(\sum_{i=1}^{n} W_{ni}^2 \right)^3 \quad \#$$

Thm. the sufficient condition $(\ast)$ is equivalent to

$(\#1)$ $\max\limits_{i=1,\cdots,n} (W_{ni}^2) = o\left(\sum_{i=1}^{n} W_{ni}^2 \right)$

Proof: If (*) holds

Then $\left(\max(w_{ni}^2)\right)^3 = \left(\max(|w_{ni}|^3)\right)^2 \le \left(\sum |w_{ni}|^3\right)^2$

$= o\left(\sum_{i=1}^n w_{ni}^2\right)^3$

If (*1) holds

Then $\sum_{i=1}^n |w_{ni}^3| \le \left(\max |w_{ni}|\right) \sum_{i=1}^n w_{ni}^2$

$= o\left(\sqrt{\sum_{i=1}^n w_{ni}^2}\right) \sum_{i=1}^n w_{ni}^2$

$= o\left(\sum_{i=1}^n w_{ni}^2\right)^{3/2}$ #

Back to the weighted least square estimate

$X_i \sim E(X_i) = \mu$, $Var(X_i) = \sigma_i^2$

$\hat{\mu} = \sum_{i=1}^n w_{ni}^* X_i$. $w_{ni}^* = \dfrac{\frac{1}{\sigma_i^2}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$

$\hat{\mu} - \mu = \sum_{i=1}^n w_{ni}^* (X_i - \mu)$

$= \sum_{i=1}^n \dfrac{\frac{1}{\sigma_i}\left(\frac{X_i - \mu}{\sigma_i}\right)}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$

$= \sum_{i=1}^n w_{ni} \cdot Y_i$ where $w_{ni} = \dfrac{\frac{1}{\sigma_i}}{\sum_{i=1}^n \frac{1}{\sigma_i^2}}$

$E(Y_i) = 0$ $Var(Y_i) = 1$

Suppose that $E|X_i|^3 < +\infty$ Then $E(|Y_i|^3) < +\infty$

Thus $\dfrac{\sum_{i=1}^n w_{ni} Y_i}{Var(\sum_{i=1}^n w_{ni} Y_i)} = \dfrac{\hat{\mu} - \mu}{\sqrt{\sum_{i=1}^n 1/\sigma_i^2}} \xrightarrow{d} N(0,1)$

provided that $\max w_{ni}^2 = o\left(\sum_{i=1}^n w_{ni}^2\right)$

$$\Longleftrightarrow \quad \max \frac{1}{\sigma_i^2} = o\left(\sum_{i=1}^{n} \frac{1}{\sigma_i^2}\right)$$

- Similarly, back to the simple linear regression case.

$$Y_i = \alpha + \beta x_i + \varepsilon_i, \qquad \varepsilon_i \overset{ind}{\sim} E(\varepsilon_i)=0, \ Var(\varepsilon_i) = \sigma^2 < +\infty$$

Suppose $E(|\varepsilon_i|^3) < +\infty$ ← can we remove this?

Yes, check Lindeberg Condition.

Then
$$\hat{\beta} = \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)}{\sum_{i=1}^{n}(x_i - \bar{x}_n)^2} Y_i$$

$$= \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)}{\sum(x_i - \bar{x}_n)^2} (\alpha + \beta x_i + \varepsilon_i)$$

$$= \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)x_i}{\sum(x_i - \bar{x}_n)^2} \beta + \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)}{\sum(x_i - \bar{x}_n)^2} \varepsilon_i$$

$$= \beta + \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)}{\sum(x_i - \bar{x}_n)^2} \varepsilon_i$$

$$\hat{\beta} - \beta = \sum_{i=1}^{n} \frac{(x_i - \bar{x}_n)}{\sum(x_i - \bar{x}_n)^2} \varepsilon_i \qquad Var\left(\sum_{i=1}^{n} W_{ni} Y_i^*\right) = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_n)^2 \cdot \sigma^2}{\left(\sum_{i=1}^{n} (x_i - \bar{x}_n)^2\right)^2} = \frac{\sigma^2}{\sum(x_i - \bar{x}_n)^2}$$

$$= \sum_{i=1}^{n} W_{ni} Y_i^*$$

$$\frac{\hat{\beta} - \beta}{Var(\hat{\beta})} = \frac{\hat{\beta} - \beta}{\sqrt{\sigma^2 / \sum(x_i - \bar{x}_n)^2}} \xrightarrow{d} N(0,1)$$

provided $\max W_{ni}^2 = o\left(\sum_{i=1}^{n} W_{ni}^2\right)$

$$\Longleftrightarrow \quad \max (x_i - \bar{x}_n)^2 = o\left(\sum_{i=1}^{n} (x_i - \bar{x}_n)^2\right)$$

Now, we assume

(a) ε_i's are identically distributed with $E(\varepsilon_i) = 0$ $Var(\varepsilon_i) = \sigma^2$

(b) $\max (x_i - \bar{x}_n)^2 = o\left(\sum_{i=1}^{n} (x_i - \bar{x}_n)^2\right)$

we note that $\hat{\beta} - \beta = \sum_{i=1}^{n} \frac{(X_i - \bar{X}_n)}{\sum (X_i - \bar{X}_n)^2} \varepsilon_i$

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

Let $Y_{ni}^* = (X_i - \bar{X}_n)\varepsilon_i$ $\quad T_n = \sum_{i=1}^{n} Y_{ni}^*$

$$= \sum_{i=1}^{n} (X_i - \bar{X}_n)\varepsilon_i$$

$$S_n^2 = Var(T_n) = \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \sigma^2$$

Thus If $\frac{1}{S_n^2} \sum_{i=1}^{n} E\left(Y_{ni}^{*2} I\{|Y_{ni}^*| \geq \varepsilon S_n\}\right) \to 0$ as $n \to +\infty$

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

$$\frac{T_n}{S_n} = \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)\varepsilon_i}{\sqrt{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \sigma^2}} = \frac{\hat{\beta} - \beta}{Var(\hat{\beta})} \xrightarrow{d} N(0,1)$$

now we verify this $\frac{1}{S_n^2} \sum_{i=1}^{n} E\left(Y_{ni}^{*2} I\{|Y_{ni}^*| \geq \varepsilon S_n\}\right)$

$$= \frac{1}{S_n^2} \sum_{i=1}^{n} E\left\{\varepsilon_i^2 (X_i - \bar{X}_n)^2 I\left(|\varepsilon_i (X_i - \bar{X}_n)| \geq \varepsilon S_n\right)\right\} \quad (*)$$

Let $\gamma_n^2 = \frac{\max (X_i - \bar{X}_n)^2}{\sum (X_i - \bar{X}_n)^2}$ (we know $\gamma_n^2 \to 0$ as $n \to +\infty$)

$$|\varepsilon_i (X_i - \bar{X}_n)| \geq \varepsilon S_n$$

$$\Leftrightarrow \quad \varepsilon_i^2 (X_i - \bar{X}_n)^2 \geq \varepsilon^2 S_n^2 \quad \Leftrightarrow \quad \varepsilon_i^2 \geq \varepsilon^2 \cdot \frac{S_n^2}{(X_i - \bar{X}_n)^2} = \varepsilon^2 \sigma^2 \cdot \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}{(X_i - \bar{X}_n)^2}$$

$$= \varepsilon^2 \sigma^2 \cdot \frac{1}{\left[\frac{(X_i - \bar{X}_n)^2}{\sum_{i=1}^{n} (X_i - \bar{X}_n)^2}\right]}$$

$$\Rightarrow \quad \varepsilon_i^2 \geq \varepsilon^2 \cdot \sigma^2 \frac{1}{\gamma_n^2}$$

$$\Rightarrow \quad |\varepsilon_i| \geq \varepsilon \sigma \frac{1}{\gamma_n}$$

Thus $I\left(|\varepsilon_i (X_i - \bar{X}_n)| \geq \varepsilon S_n\right) \leq I\left(|\varepsilon_i| \geq \varepsilon \sigma \frac{1}{\gamma_n}\right)$

Then $(*) \le \dfrac{1}{S_n^2} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 \underbrace{E\left\{ \varepsilon_i^2 I\left(|\varepsilon_i| \ge \frac{\varepsilon \sigma}{r_n} \right) \right\}}_{\text{same for } i}$

$$= \frac{1}{\sigma^2} E\left\{ \varepsilon_i^2 I\left(|\varepsilon_i| \ge \frac{\varepsilon \delta}{r_n} \right) \right\}$$

$\to 0$ Because $r_n \to 0$, $\frac{\varepsilon \delta}{r_n} \to +\infty$ as $n \to +\infty$ #

- Lastly. Let $X_1, \cdots, X_n \overset{iid}{\sim} E(X)=0$. $Var(X)=\sigma^2<+\infty$

Let $\{N(t)\}, t \ge 0$. be a family of nonnegative integer-valued random variables all defined on a common probability space.

Assume that $\frac{N(t)}{t} \xrightarrow{P} C$, $0<C<+\infty$, as $t \to +\infty$.

Then $$\frac{S_{N(t)}}{\sigma \sqrt{N(t)}} \xrightarrow{d} N(0,1) \quad \text{as } t \to +\infty$$

where $S_{N(t)} = X_1 + X_2 + \cdots + X_{N(t)}$

- I~~ID~~ ✓ ←
- ~~I~~ID No independence

→ Stationary m-dependence.

- A sequence of r.v. $Y_1, Y_2, \cdots$ is said to be m-dependent.

  if for every integer $s \geq 1$,

  the sets of r.v.'s $\{Y_1, \cdots, Y_s\}$ and $\{Y_{m+s+1}, Y_{m+s+2}, \cdots\}$

  are independent.

  If $m=0$, the $\{Y_1, \cdots\}$ is a sequence of independent r.v.'s.

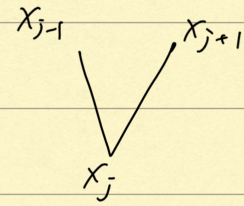
- A sequence of r.v. $Y_1, Y_2, \cdots$ is said to be (strictly) stationary.

  if any positive integers s, and t, the joint distribution

  of vector. $(Y_t, Y_{t+1}, \cdots, Y_{t+s})$ does not depend on t.

Examples:
- $X_0, X_1, X_2, \cdots$ iid r.v.

  Run of increasing values.

  $$Y_i = I\{X_{j-1} > X_i < X_{j+1}\}$$

  $$\{Y_i\}_{i=1}^{+\infty} \qquad \sum Y_i$$

  a stationary 3-dependent sequence

- Success runs. $X_1, \cdots, X_n \cdots \overset{iid}{\sim}$ Bernoulli (p)

  $$Y_i = (1 - X_{i-1}) X_i \qquad m=2.$$

$Y_i = 1$ "0" "1"

- Runs of length r. $Y_i = (1 - X_{i-1}) X_i X_{i+1} \cdots X_{i+r-1} (1 - X_{i+r})$

$$\text{"0"} \; \underbrace{\text{"1"} \; \text{"1"} \cdots \text{"1"}}_{r} \; \text{"0"}$$

$$m = r + 2.$$

$$Y_i = g(X_i, \cdots, X_{i+m})$$

Say $Y_1, \cdots, Y_n \cdots$ is a stationary m-dependent sequence.

$$T_n = \sum_{i=1}^{n} Y_i$$

$$S_n^2 = Var(T_n) = Var\left(\sum_{i=1}^{n} Y_i\right) \qquad n \geq m$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov(Y_i, Y_j)$$

$$= \sum_{i=1}^{n} Var(Y_i) + 2 \sum_{i<j} Cov(Y_i, Y_j)$$

by stationary. $\sigma_0^2 = Var(Y_i)$ for i

$\sigma_1^2 = Cov(Y_i, Y_{i+1})$ for i

$\sigma_2^2 = Cov(Y_i, Y_{i+2})$

$\vdots$

$\sigma_m^2 = Cov(Y_i, Y_{i+m})$ for i

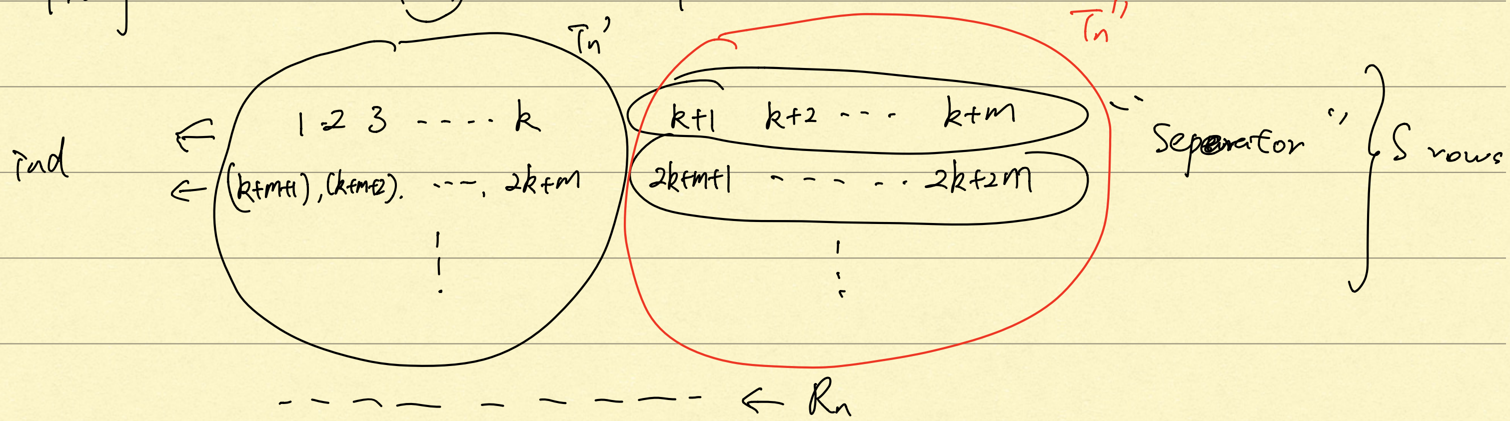
$$\underline{S_n^2 = n\sigma_0^2 + 2(n-1)\sigma_1^2 + 2(n-2)\sigma_2^2 + \cdots + 2(n-m)\sigma_m^2}$$

$$\frac{T_n - E(T_n)}{\sqrt{Var(T_n)}} = \frac{T_n - E(T_n)}{S_n} \xrightarrow{d} N(0,1)$$ If $\sigma_j^2 < +\infty$, for $j = 0, 1, \cdots, m$

Let $E(Y_i) = \mu$. $\sqrt{n}(\bar{Y}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$

where $\sigma^2 = \sigma_0^2 + 2\sigma_1^2 + 2\sigma_2^2 + \cdots + 2\sigma_m^2$

Proof: $T_n = \sum_{i=1}^{n} Y_i$ pick $k > m$



$T_n = T_n' + T_n'' + R_n$ say $n = s(k+m) + r$. where $r < (k+m)$

where $T_n' = \sum_{j=0}^{s-1} V_{kj}$ $V_{kj} = \sum_{i=1}^{k} Y_{j(k+m)+i}$

$T_n'' = \sum_{j=0}^{s-1} W_{kj}$ $W_{kj} = \sum_{i=k+1}^{k+m} Y_{j(k+m)+i}$

$$\frac{T_n - E(T_n)}{\sqrt{Var(T_n)}} \xrightarrow{d} N(0,1) \text{ as } n \to +\infty$$

Notation is self-contained in this Lemma.

$\boxed{\text{Lemma}}$ $T_n = Z_{nk} + X_{nk}$ for $n = 1, 2, \cdots$ and $k = 1, 2, \cdots$.

if (1) $X_{nk} \xrightarrow{P} 0$ uniformly in n as $k \to +\infty$

(2) $Z_{nk} \xrightarrow{d} Z_k$ as $n \to +\infty$ for each k.

(3) $Z_k \xrightarrow{d} Z$ as $k \to +\infty$

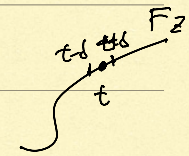
then $T_n \xrightarrow{d} Z$ as $n \to +\infty$, $k \to +\infty$

Proof: "as $n \to +\infty$, $F_{T_n}(t) \to F_Z(t)$ for $t \in C(F_Z)$ continuity set of F_Z"

Let $\varepsilon > 0$. and $t \in C(F_Z)$.

Find $\delta > 0$ such that $P(|Z - t| < \delta) < \varepsilon$

such that $t+\delta$ and $t-\delta$ are in the continuity sets $C(F_Z)$ and $C(F_{Z_k})$ for all k.



From condition (1), we may find K such that

$$P(|X_{nk}| \geq \delta) < \varepsilon \text{ for all } k > K \text{ and all } n.$$

From condition (3), we may find $K' > K$, such that for $k \geq K'$,

$$|P(Z_k \leq t+\delta) - P(Z \leq t+\delta)| < \varepsilon$$

and

$$|P(Z_k \leq t-\delta) - P(Z \leq t-\delta)| < \varepsilon.$$

Now, fix $k > K'$,

$$F_{T_n}(t) = P(T_n \leq t) = P(Z_{nk} + X_{nk} \leq t)$$

$$\text{"}Z_{nk}+X_{nk} \leq t\text{"} \iff \text{"}Z_{nk}+X_{nk} \leq t,\ X_{nk} \leq -\delta\text{"} \cup \text{"}Z_{nk}+X_{nk} \leq t,\ X_{nk} > -\delta\text{"}$$

$$\begin{aligned}
P(Z_{nk}+X_{nk} \leq t) &= P(Z_{nk}+X_{nk} \leq t,\ X_{nk} \leq -\delta) + P(Z_{nk}+X_{nk} \leq t,\ X_{nk} > -\delta) \\
&\leq P(X_{nk} \leq -\delta) + P(Z_{nk} \leq t - X_{nk},\ X_{nk} > -\delta) \\
&\leq P(X_{nk} \leq -\delta) + P(Z_{nk} \leq t+\delta) \\
&\leq P(|X_{nk}| \geq \delta) + P(Z_{nk} \leq t+\delta)
\end{aligned}$$

$$\begin{aligned}
\limsup_{n\to\infty} P(T_n \leq t) &= \limsup_{n\to\infty} P(Z_{nk}+X_{nk} \leq t) \\
&\leq \limsup_{n\to\infty} \{P(|X_{nk}| \geq \delta) + P(Z_{nk} \leq t+\delta)\} \\
&\leq \limsup_{n\to\infty} P(Z_{nk} \leq t+\delta) \\
&\leq P(Z_k \leq t+\delta) + \varepsilon \\
&\leq P(Z \leq t+\delta) + 2\varepsilon \\
&\leq P(Z \leq t) + 3\varepsilon
\end{aligned}$$

1

Similarly, $P(Z_{nk}+X_{nk} \leq t) + P(|X_{nk}| \geq \delta) \geq P(Z_{nk} \leq t-\delta)$

$$= P(Z_{nk} \le t-\delta, X_{nk} \ge \delta) + P(Z_{nk} \le t-\delta, X_{nk} < \delta)$$

$$\le P(X_{nk} \ge \delta) + P(Z_{nk} \le t-\delta, X_{nk} < \delta)$$

$$\le P(|X_{nk}| \ge \delta) + P(Z_{nk} + \delta \le t, X_{nk} < \delta)$$

$$\le P(|X_{nk}| \ge \delta) + P(Z_{nk} + X_{nk} \le t)$$

$$\liminf_{n\to\infty} \{P(Z_{nk} + X_{nk} \le t) + P(|X_{nk}| \ge \delta)\} \ge \liminf P(Z_{nk} \le t-\delta)$$

similarly $\Rightarrow$ $\liminf_{n\to\infty} P(T_n \le t) \ge P(Z \le t) - 3\varepsilon$ 2

$$P(Z\le t) - 3\varepsilon \le \liminf_{n\to\infty} P(T_n \le t) \le \limsup_{n\to\infty} P(T_n \le t) \le P(Z \le t) + 3\varepsilon$$

Let $\varepsilon \downarrow 0$. $\Rightarrow$ $\lim_{n\to\infty} P(T_n \le t) = P(Z \le t)$ where $t \in C(F_Z)$

$$F_{T_n}(t) \to F_Z(t) \quad \#$$

---

back to CLT.

$$T_n = T_n' + T_n'' + R_n$$

$$= (T_n' + R_n) + T_n''$$

$E(T_n) = n\mu$ w.L.O.G set $\mu = 0$

$S_n^2 = n\sigma_0^2 + 2(n-1)\sigma_{01}^2 + 2(n-2)\sigma_{02}^2 + \cdots + 2(n-m)\sigma_{0m}^2$

$$\Rightarrow \frac{T_n - E(T_n)}{S_n} \xrightarrow{d} N(0,1)$$

$$\frac{T_n}{\sqrt{n}} = \underbrace{\frac{T_n' + R_n}{\sqrt{n}}}_{Z_{nk}} + \underbrace{\frac{T_n''}{\sqrt{n}}}_{X_{nk}}$$

① " $X_{nk} \xrightarrow{P} 0$ " as $k \to +\infty$ uniformly in n ✓

$$Var(\text{"}X_{nk}\text{"}) = Var\left(\frac{\text{"}T_n'\text{"}}{\sqrt{n}}\right) = \frac{1}{n}\sum_{j=0}^{s-1} Var(W_{kj})$$

$$= \frac{s}{n} Var(W_{kj})$$

$$= \frac{s}{n} \underbrace{Var(Y_{k+1} + Y_{k+2} \cdots + Y_{k+m})}$$

$n = s(k+m)+r$

$$= \frac{Var(W_{kj})}{(k+m)+r/s} \to 0 \text{ as } k \to +\infty$$

② " $Z_{nk} \to Z_k$ " as $n \to +\infty$ for a fixed k.

$$Z_{nk} = \frac{T_n' + R_n}{\sqrt{n}} = \frac{\sum_{j=0}^{s-1} V_{kj} + R_n}{\sqrt{n}}$$

(i) $\underline{\frac{R_n}{\sqrt{n}} \xrightarrow{P} 0}$

$$Var\left(\frac{R_n}{\sqrt{n}}\right) = \frac{1}{n} Var(R_n)$$

$$= \frac{1}{n} Var(\underbrace{Y_1 + \cdots + Y_r}) \text{ where } r < k+m$$

$$\leq \frac{r^2}{n} Var(Y)$$

$$\leq \frac{(k+m)^2}{n} \cdot Var(Y)$$

$\to 0$ as $\underline{n \to +\infty \text{ for a fixed } k}$

(ii) $\frac{\sum_{j=0}^{s-1} V_{kj}}{\sqrt{n}}$ apply CLT

$$\frac{\sum_{j=0}^{s-1} V_{kj}}{\sqrt{s}} \xrightarrow{d} N(0, Var(V_{kj}))$$

$$\frac{s}{n} = \frac{s}{s(k+m)+r}$$

$\|$

for a fixed k
$n \to \infty$?

$$\sqrt{\frac{s}{n}} \times \frac{\sum_{j=0}^{s-1} V_{kj}}{\sqrt{s}} \xrightarrow{d} \sqrt{\frac{1}{k+m}} \times N(0, Var(V_{kj}))$$

at a fixed k, $n \to +\infty$, $Z_{nk} \xrightarrow{d} Z_k$

$$Z_k = \sqrt{\frac{1}{k+m}} \times N(0, Var(V_{kj}))$$

$$Var(Z_k) = \frac{Var(V_{kj})}{k+m}$$

$V_{kj} = Y_1 + Y_2 + \cdots + Y_k$ where $k > m$

$$\frac{Var(V_{kj})}{k+m} = \frac{k\sigma_0^2 + 2(k-1)\sigma_{01}^2 + \cdots + 2(k-m)\sigma_{0m}^2}{k+m} \quad k \to +\infty$$

$$\to \sigma_0^2 + 2\sigma_1^2 + \cdots + 2\sigma_m^2 = \sigma^2$$

$Z_k \xrightarrow{d} N(0, \sigma^2)$ as $k \to +\infty$ #

$X_1 \cdots X_n \cdots$ iid $E(X) = \mu$, $Var(X) = \sigma_X^2$

$Y_i = X_i X_{i+m}$ $\quad E(Y_i) = E(X_i X_{i+m}) = E(X_i)E(X_{i+m}) = \mu^2$

$Y_1 \cdots Y_n \cdots$ is a stationary m-dependent sequence.

$$\sigma^2 = 2\sigma_0^2 + 2\sigma_1^2 + 2\sigma_2^2 + \cdots + 2\sigma_m^2$$

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

$$\sigma_0^2 = Var(Y_i) = Var(X_i X_{i+m}) = \cancel{Var(X_i)Var(X_{i+m})}?$$

$$= E(X_i^2 X_{i+m}^2) - [E(X_i X_{i+m})]^2$$

$$= E(X_i^2)E(X_{i+m}^2) - \mu^4$$

$$= (\mu^2+\sigma^2)^2 - \mu^4$$

$$\sigma_j^2 = \mathrm{Cov}(Y_1, Y_{1+j}) = \mathrm{Cov}(X_1 X_{1+m},\ X_{1+j} X_{1+m+j})$$

$$= 0 \qquad \text{if } j \neq 0,\ j \neq m$$

$$\sigma_m^2 = \mathrm{Cov}(Y_1, Y_{1+m}) = \mathrm{Cov}(X_1 X_{1+m},\ X_{1+m} X_{1+2m})$$

$$= E(X_1 X_{1+m}^2 X_{1+2m}) - \mu^4$$

$$= \mu^2(\mu^2+\sigma^2) - \mu^4 = \mu^2\sigma^2$$