

CLT for non-iid r.v's

Lindeberg's CLT. For each $n=1,2,\dots$ let $X_{n1}, X_{n2}, \dots, X_{nn}$ be independent, mean zero r.v's. Assume that, as $n \rightarrow \infty$,

$$(i) \sum_{k=1}^n \mathbb{E} X_{nk}^2 \rightarrow \sigma^2 \in (0, \infty)$$

$$(ii) \sum_{k=1}^n \mathbb{E} X_{nk}^2 \mathbb{1}_{\{|X_{nk}| > \varepsilon\}} \rightarrow 0 \quad \forall \varepsilon > 0$$

"the Lindeberg condition"

Then $\sum_{k=1}^n X_{nk} \xrightarrow{d} Z \sim N(0, \sigma^2)$

!!
 S_n

Remarks

① Lind CLT $\Rightarrow$ classical CLT (take $X_{nk} = \frac{X_k}{\sqrt{n}}$; (ii) $= \mathbb{E} X_1^2 \mathbb{1}_{\{|X_1| > \varepsilon \sqrt{n}\}} \rightarrow 0$ by DCT)

② The variances $\sigma_{nk}^2 := \mathbb{E} X_{nk}^2$ satisfy $\max_k \sigma_{nk}^2 \rightarrow 0$ as $n \rightarrow \infty$

$\leftarrow$ "the Feller condition"

$$\mathbb{E} X_{nk}^2 = \underbrace{\mathbb{E} X_{nk}^2 \mathbb{1}_{\{|X_{nk}| \leq \varepsilon\}}}_{\uparrow \varepsilon} + \mathbb{E} X_{nk}^2 \mathbb{1}_{\{|X_{nk}| > \varepsilon\}}$$

$$\Rightarrow \max_k \mathbb{E} X_{nk}^2 \leq \varepsilon^2 + o(1) \text{ as } n \rightarrow \infty \text{ by (ii). } \varepsilon \text{ is arbitrary } \Rightarrow \square$$



Show by example that Lindeberg's condition $\neq$ Feller's condition

(hint: sparsity. Think of Poisson paradigm)

Proof same idea as in Classical CLT.

Lévy's Continuity Thm (p.124) $\Rightarrow$ enough to prove

$$\varphi_{S_n}(t) \xrightarrow{?} \varphi_Z(t) = e^{-t^2 \sigma^2 / 2} \quad \forall t \in \mathbb{R}$$

$$\mathbb{E} e^{itS_n} = \prod_{k=1}^n \mathbb{E} e^{itX_{nk}}$$

- To approximate each φ_{nk} , use Taylor's approx. of e^{itz} (p.126):

$$|e^{itz} - (1 + ix - \frac{x^2}{2})| \leq 4 \min(\underbrace{|x|^3}_{4ax^2 \text{ if } |x| \leq a}, x^2) \quad \forall x \in \mathbb{R}.$$

$$\leq 4ax^2 + 4x^2 \mathbb{1}_{\{|x| \geq a\}} \quad \forall a > 0$$

- Let X be a r.v. with mean 0, variance σ^2

Substitute $z = tX$, take $\mathbb{E}$, use Jensen. In RHS, choose $a = t\varepsilon$:

$$|\mathbb{E} e^{itX} - (1 - \frac{t^2 \sigma^2}{2})| \leq 4t^3 \varepsilon \sigma^2 + 4t^2 \mathbb{E}[X^2 \mathbb{1}_{\{|X| > \varepsilon\}}]$$

$$\stackrel{\parallel}{=} \exp(-\frac{t^2 \sigma^2}{2}) - (1 - \frac{t^2 \sigma^2}{2}) \quad (\text{use } |e^{-x} - (1-x)| \leq x^2 \quad \forall x \geq 0)$$

Use for $X = X_{nk}$: mean 0, variance $=: \sigma_{nk}^2 \Rightarrow$

$$|\underbrace{\mathbb{E} e^{itX_{nk}}}_{\parallel a_k} - \underbrace{\exp(-\frac{t^2 \sigma_{nk}^2}{2})}_{\parallel b_k}| \leq 4t^3 \varepsilon \sigma_{nk}^2 + 4t^2 \mathbb{E}[X_{nk}^2 \mathbb{1}_{\{|X_{nk}| > \varepsilon\}}] + t^4 \sigma_{nk}^4$$

This inequality holds $\forall$ complex numbers whose moduli are ≤ 1

By induction, it suffices to prove it for $n=2$.

$$|a_1 a_2 - b_1 b_2| = |(a_1 - b_1)a_2 + b_1(a_2 - b_2)|$$

$$\leq \underbrace{|a_1 - b_1|}_{\uparrow} \underbrace{|a_2|}_{\uparrow} + \underbrace{|b_1|}_{\uparrow} \underbrace{|a_2 - b_2|}_{\uparrow}$$

$$|\prod_{k=1}^n a_k - \prod_{k=1}^n b_k| \leq \sum_{k=1}^n |a_k - b_k|$$

$$|\varphi_{S_n}(t) - \exp(-\frac{t^2}{2} \sum_{k=1}^n \sigma_{nk}^2)| \leq 4t^3 \varepsilon \underbrace{\sum_{k=1}^n \sigma_{nk}^2}_{\downarrow \sigma^2} + 4t^2 \underbrace{\sum_{k=1}^n \mathbb{E}[X_{nk}^2 \mathbb{1}_{\{|X_{nk}| > \varepsilon\}}]}_{\downarrow 0 \text{ (assumption ii)}} + t^4 \underbrace{\max_k (\sigma_{nk}^2)}_{\downarrow 0 \text{ (Rem @ p.128)}} \cdot \underbrace{\sum_{k=1}^n \sigma_{nk}^2}_{\uparrow \sigma^2}$$

$$\limsup_{n \rightarrow \infty} |\varphi_{S_n}(t) - \exp(-\frac{t^2 \sigma^2}{2})| \leq 4t^3 \varepsilon \sigma^2 \quad \forall \varepsilon > 0$$

$$\Rightarrow = 0. \quad \text{QED.}$$

Example (De Moivre-Laplace CLT)

$$Y_{nk} \sim \text{Ber}(p_n) \text{ indep.} \Rightarrow T_n = \sum_{k=1}^n Y_{nk} \sim \text{Binom}(n, p_n)$$

$$\text{Assume } \text{Var}(T_n) = np_n(1-p_n) \rightarrow \infty \text{ as } n \rightarrow \infty$$

$$\text{Apply Lind CLT for } X_{nk} := \frac{Y_{nk} - p_n}{\sqrt{\text{Var}(T_n)}}$$

$$\begin{cases} \text{(i)} & \sum_{k=1}^n EX_{nk}^2 = 1 \quad \checkmark \\ \text{(ii)} & |X_{nk}| \leq \frac{1}{\sqrt{\text{Var}(T_n)}} \rightarrow 0 \Rightarrow \text{the sum in (ii)} = 0 \text{ for large } n. \quad \checkmark \end{cases}$$

$$\Rightarrow S_n = \frac{T_n - np_n}{\sqrt{\text{Var}(T_n)}} \xrightarrow{d} Z \sim N(0,1) \Rightarrow \text{renaming } T_n \text{ to } S_n:$$

THM let $S_n \sim \text{Binom}(n, p_n)$. If $\text{Var}(S_n) \rightarrow \infty$ then

("De Moivre-Laplace CLT")

$$\frac{S_n - ES_n}{\sqrt{\text{Var}(S_n)}} \xrightarrow{d} N(0,1)$$

What happens if $\text{Var}(S_n) \rightarrow \text{const}$?
 $\downarrow$

Poisson Limit Theorem

- let $X_{nk} \sim \text{Ber}(p_n)$, $S_n = \sum_{k=1}^n X_{nk} \sim \text{Binom}(n, p_n)$

Assume $p_n \leq 1/2$ (by symmetry) and $\mathbb{E} S_n = n p_n = \lambda = \text{const}$ as $n \rightarrow \infty$

i.e. const # of expected successes

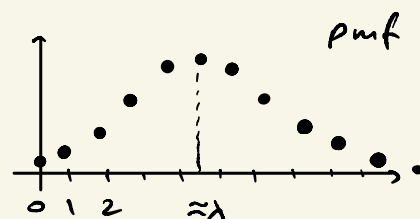
- $\forall$ fixed $k=0,1,\dots,n$:

$$P\{S_n=k\} = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n(n-1)\dots(n-k+1)}{k!} \cdot \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} \rightarrow e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

$\xrightarrow{(n \rightarrow \infty, k \text{ fixed})}$
 $\downarrow e^{-\lambda}$

Def (Poisson distribution) $X \sim \text{Poisson}(\lambda)$ if

$$P\{X=k\} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad k=0,1,2,\dots$$


 $\mathbb{E} X = \lambda, \quad \text{Var}(X) = \lambda \quad (\text{Dir})$

- We showed (by summing $k = \text{fixed}$ # of terms) that
- $$P\{S_n \leq k\} \rightarrow P\{X \leq k\} \quad \forall \text{ fixed } k$$

$$\Rightarrow S_n \xrightarrow{d} X$$

• Alternative, Fourier-theoretic proof:

Poisson Limit Theorem let $S_n \sim \text{Binom}(n, p_n)$, $p_n \leq 1/2$.

If $np_n \rightarrow \lambda \neq 0$ as $n \rightarrow \infty$, then

$$S_n \xrightarrow{d} X \sim \text{Poisson}(\lambda).$$

Proof By Levy Continuity Thm, enough to show that

$$\phi_{S_n} \rightarrow \phi_X \text{ pointwise.}$$

$$\begin{aligned} \textcircled{1} \phi_X(t) &= \mathbb{E} e^{itX} = \sum_{k=0}^{\infty} e^{itk} \cdot e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda e^{it})^k}{k!} \\ &= e^{-\lambda} \exp(\lambda e^{it}) = \exp[\lambda(e^{it} - 1)] \end{aligned}$$

$$\begin{aligned} \textcircled{2} \phi_{S_n}(t) &= \mathbb{E} e^{it(X_{n1} + \dots + X_{nn})} = \left(\mathbb{E} e^{itX_{n1}} \right)^n \\ &= \left[e^{it \cdot 1} p_n + e^{it \cdot 0} (1-p_n) \right]^n = \left[1 + \underbrace{p_n}_{\parallel \frac{\lambda}{n} (1+o(1))} (e^{it} - 1) \right]^n \\ &\rightarrow \exp[\lambda(e^{it} - 1)] = \phi_X(t). \quad \square \end{aligned}$$

SUMMARY:

THM (Gaussian - Poisson dichotomy)

let $S_n \sim \text{Binom}(n, p_n)$ with $p_n \leq 1/2$.

(i) If $\mathbb{E} S_n = np_n \rightarrow \infty$ then $\frac{S_n - np_n}{\sqrt{np_n(1-p_n)}} \xrightarrow{d} N(0,1)$ ← frequent successes

(ii) If $\mathbb{E} S_n \rightarrow \lambda \neq 0 = \text{const}$ then $S_n \xrightarrow{d} \text{Poisson}(\lambda)$ ← rare successes.

HW: $\text{Poisson}(\lambda) \rightarrow \text{normal}$ if $\lambda \rightarrow \infty$

$$\text{Poisson}(\lambda) + \text{Poisson}(\mu) = \text{Poisson}(\lambda + \mu)$$

↑ indep ↑

⤴ (HW) $\mathbb{E} S_n \rightarrow 0 \Rightarrow S_n \xrightarrow{d} 0$