EQUIVALENT DEF'S OF CONVERGENCE IN DISTRIBUTION:

- By def, $X_n \xrightarrow{d} X \iff \mathbb{P}\{X_n \in B\} \to \mathbb{P}\{X \in B\}$ $\forall$ interval $B = (-\infty, x]$ s.t. $\mathbb{P}\{X = x\} = 0$

$$\iff \mathbb{E}h(X_n) \to \mathbb{E}h(X) \quad \text{where } h = \mathbb{1}_{(-\infty, x]}$$

- We can extend this to more general "test sets" B & "test functions" h:

Lem (Portmanteau lemma) TFAE:

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

(ii) $\mathbb{P}\{X_n \in [a,b]\} \to \mathbb{P}\{X \in [a,b]\}$ whenever $\mathbb{P}\{X=a\} = \mathbb{P}\{X=b\} = 0$

(iii) $\mathbb{P}\{X_n \in B\} \to \mathbb{P}\{X \in B\}$ $\forall$ continuity set $B \subset \mathbb{R}$ of X

(iv) $\mathbb{E}h(X_n) \to \mathbb{E}h(X)$ $\forall$ bounded, continuous $h: \mathbb{R} \to \mathbb{R}$

(v) $\mathbb{E}h(X_n) \to \mathbb{E}h(X)$ whenever h and all its derivatives $h, h', h'', \dots$ are bounded & compactly supported.

i.e. $\forall$ set B such that $\mathbb{P}\{X \in \partial B\} = 0$ where $\partial B = \bar{B} \setminus B^\circ$ is the boundary of B

Proof

(i) $\Rightarrow$ (iii) By Skorokhod's representation thm, we can assume wLOG that $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{p} X$. Fix $\forall \varepsilon > 0$

$$\mathbb{P}\{X_n \in B\} \le \mathbb{P}\underbrace{\{X_n \in B \;\&\; |X_n - X| < \varepsilon\}}_{\Downarrow \Delta} + \underbrace{\mathbb{P}\{|X_n - X| \ge \varepsilon\}}_{\to 0}$$

$X \in B_\varepsilon := \{y \in \mathbb{R} : \exists x \in B \text{ s.t. } |x - y| < \varepsilon\}$ (ε-neighborhood)

$$\Rightarrow \limsup_n \mathbb{P}\{X_n \in B\} \le \underbrace{\mathbb{P}\{X \in B_\varepsilon\}}_{\downarrow \text{(continuity of probability)}}.$$

- Let $\varepsilon \downarrow 0 \Rightarrow B_\varepsilon \downarrow \bar{B} \Rightarrow \mathbb{P}\{X \in \bar{B}\}$

$$\Rightarrow \limsup_n \mathbb{P}\{X_n \in B\} \le \mathbb{P}\{X \in \underset{B^\circ \sqcup \partial B}{\overset{\shortparallel}{\bar{B}}}\} = \mathbb{P}\{X \in B^\circ\} \le \mathbb{P}\{X \in B\}.$$

(B is a continuity set)

- Apply this argument for B^c instead of B, which must also be a continuity set $(\partial(B^c) = \partial B) \Rightarrow \limsup_n \mathbb{P}\{X_n \in B^c\} \le \mathbb{P}\{X \in B^c\}$

$\Rightarrow \liminf_n \mathbb{P}\{X_n \in B\} \ge \mathbb{P}\{X \in B\}$. Combine $\Rightarrow \mathbb{P}\{X_n \in B\} \to \mathbb{P}\{X \in B\}$. $\square$

(iii) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (i) Since $P\{|X|>M\} \to 0$ as $M \to \infty$ (continuity of measure),
$\forall \varepsilon>0 \ \exists M$ s.t. $P\{|X|>M\}<\varepsilon$ and $P\{|X|=M\}=0 \leftarrow$ (set of discontinuities of F_X is countable)

$$\Rightarrow P\{X_n \le x\} = \underbrace{P\{X_n \le x, |X_n| \le M\}}_{\text{closed interval}} + \underbrace{P\{|X_n|>M\}}_{\text{complement of a closed interval}}$$

$$\xrightarrow{\text{assm.}} P\{X \le x, |X| \le M\} + P\{|X|>M\} \quad \text{whenever } P\{X=x\}=0 \ \text{(assm.)}$$

$$\le P\{X \le x\} + \varepsilon.$$

$$\Rightarrow \limsup_n P\{X_n \le x\} \le P\{X \le x\} + \varepsilon.$$

- A lower bound follows by the same argument with $X_n \leftrightarrow X$. $\square$

✓ (i) $\Rightarrow$ (iv) By Skorokhod's a.s. representation thm (p. 111), we can assume WLOG

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

Continuity of h $\Rightarrow$ $\qquad h(X_n) \xrightarrow{a.s.} h(X)$

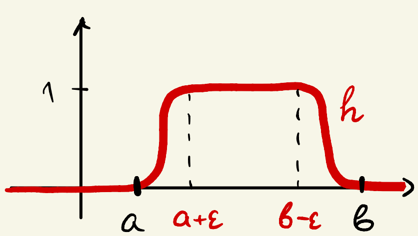
Boundedness of h $\overset{DCT}{\Longrightarrow}$ $\qquad \mathbb{E}h(X_n) \to \mathbb{E}h(X)$. $\square$

✓ (iv) $\Rightarrow$ (v) is trivial.

✓ (v) $\Rightarrow$ (ii) Smoothing: Fix any $a<b$ satisfying $P\{X=a\}=P\{X=b\}=0$ (*)

Fix $\varepsilon>0$. $\exists$ function h as in the conclusion satisfying

$$\mathbb{1}_{[a+\varepsilon, b-\varepsilon]} \le h \le \mathbb{1}_{[a,b]} \text{ pointwise.}$$



Evaluate on X_n, take $\mathbb{E}$ $\Rightarrow$

$$P\{X_n \in [a,b]\} \ge \mathbb{E}h(X_n)$$
$$\to \mathbb{E}h(X) \quad \text{(assumption)}$$
$$\ge P\{X \in [a+\varepsilon, b-\varepsilon]\}$$

$$\Rightarrow \liminf_n P\{X_n \in [a,b]\} \ge P\{X \in [a+\varepsilon, b-\varepsilon]\} \xrightarrow[\text{continuity of meas}]{\varepsilon \downarrow 0} P\{X \in (a,b)\} \overset{(*)}{=} P\{X \in [a,b]\}$$

The upper bd is similar (DIY) $\square$.

- REMARK | For random vectors X_n taking values in $\mathbb{R}^n$, convergence of distr. can be defined via P.L. (iv).

HW (a) $X_n \xrightarrow{w} X \not\Rightarrow \mathbb{E}X_n \to \mathbb{E}X$ in general

(b) $X_n \xrightarrow{w} X$ & $\mathbb{E}|X_n| \to \mathbb{E}|X| \Rightarrow X_n \overset{w}{\Rightarrow} X$

THE CENTRAL LIMIT THEOREM

We will prove:

Classical CLT. Let $X_1, X_2, \dots$ be iid r.v's with mean μ and variance σ^2.

Then $S_n = X_1 + \dots + X_n$ satisfies

$$\underbrace{\frac{S_n - n\mu}{\sigma\sqrt{n}}}_{\bar{S}_n \, !!} \xrightarrow{w} Z \sim N(0,1)$$

- By Portmanteau lemma, it is enough to show that

$$\mathbb{E}h(\bar{S}_n) \to \mathbb{E}h(Z)$$

$\forall$ "test function" $h \in \mathcal{S}$ (Schwartz class)

Toward that goal:

THM (Nonasymptotic CLT)
Let $X_1, X_2, \dots$ be independent mean zero r.v's with $\mathbb{E}|X_k|^3 < \infty$.
Then $S_n := X_1 + \dots + X_n$ satisfies

$$\left|\mathbb{E}h(S_n) - \mathbb{E}h(Z)\right| \le C \|h'''\|_\infty \cdot \sum_{k=1}^n \mathbb{E}|X_k|^3 \qquad \forall \text{ 3 times diff'ble } h,$$

where $Z \sim N(0, \mathrm{Var}(S_n))$ and C is an absolute constant.

Proof (Lyapunov-Lindeberg replacement method)

- $S_n = S_{n-1} + X_n$. WLOG $\|h'''\|_\infty = 1$. Taylor's approximation (w/ Lagrange form of the remainder) $\Rightarrow$

$$\left| h(S_n) - h(S_{n-1}) - X_n h'(S_{n-1}) - \frac{X_n^2}{2} h''(S_{n-1}) \right| \le \frac{|X_n|^3}{3!}$$

independent $\Rightarrow$ mean $= 0$ independent

Take $\mathbb{E}$ on both sides, use $|\mathbb{E}\cdots| \le \mathbb{E}|\cdots|$ (Jensen) $\Rightarrow$

$$\left| \mathbb{E}h(S_n) - \mathbb{E}h(S_{n-1}) - \frac{\sigma_n^2}{2}\mathbb{E}h''(S_{n-1}) \right| \le \frac{\mathbb{E}|X_n|^3}{6} \qquad \text{where } \boxed{\sigma_n^2 := \mathbb{E}X_n^2}. \quad (*)$$

- Let $Z_n \sim N(0, \sigma_n^2)$ be independent of $X_1, \dots, X_n$.

Arguing like in $(*)$ but replacing X_n by Z_n, we get

$$\left| \mathbb{E}h(S_{n-1} + Z_n) - \mathbb{E}h(S_{n-1}) - \frac{\sigma_n^2}{2}\mathbb{E}h''(S_{n-1}) \right| \le \frac{\mathbb{E}|Z_n|^3}{6} \qquad (**)$$

Subtract $(*) - (**) \overset{\Delta}{\Rightarrow}$

$$\left|\mathbb{E}h(S_n) - \mathbb{E}h(S_{n-1}+Z_n)\right| \le \frac{1}{6}\Big(\underbrace{\mathbb{E}|X_n|^3 + \mathbb{E}|Z_n|^3}_{\gtrsim}\Big) \le C\cdot\mathbb{E}|X_n|^3.$$

$\Big(Z_n/\sigma_n \sim N(0,1) \Rightarrow \|Z_n/\sigma_n\|_{L^3} \lesssim 1$ ↖(abs. const factor) $\Rightarrow \|Z_n\|_{L^3} \lesssim \sigma_n = \|X_n\|_{L^2} \le \|X\|_{L^3}\Big)$

- Continue the replacement process. $X_{n-1} \mapsto Z_{n-1} \Rightarrow$

$$\left|\mathbb{E}h(S_n) - \mathbb{E}h(S_{n-2}+Z_{n-1}+Z_n)\right| \le C\,\mathbb{E}|X_n|^3 + C\,\mathbb{E}|X_{n-1}|^3$$

... (n times) ... $\Rightarrow$

$$\left|\mathbb{E}h(S_n) - \mathbb{E}\Big(\sum_{k=1}^n Z_k\Big)\right| \le C\sum_{k=1}^n \mathbb{E}|X_k|^3$$

- $\sum_{k=1}^n Z_k$ is normal, mean $=0$, Variance $= \sum_{k=1}^n \sigma_k^2 = \mathrm{Var}(S_n)$ $\square$

Proof of CLT assuming $\mathbb{E}|X_k|^3 < \infty$:

WLOG $\mu = 0$, $\sigma = 1$. Apply the nonasymptotic CLT for $\frac{S_n}{\sqrt{n}} = \sum_{k=1}^n \frac{X_k}{\sqrt{n}}$ ← mean 0, var 1. $\Rightarrow$

$$\left|\mathbb{E}h\Big(\frac{S_n}{\sqrt{n}}\Big) - \mathbb{E}h(\overset{N(0,1)}{Z})\right| \le C\|h'''\|_\infty \cdot \sum_{k=1}^n \mathbb{E}\left|\frac{X_k}{\sqrt{n}}\right|^3 \overset{\text{iid}}{=} \frac{C\|h'''\|_\infty \cdot \mathbb{E}|X_1|^3}{\sqrt{n}} \to 0 \text{ as } n\to\infty.$$

Portmanteau lemma $\Rightarrow$ $\square$.

Proof of CLT in full generality (Truncation):

- WLOG $\mu=0$, $\sigma=1$, $\|h'\|_\infty \le 1$, $\|h'''\|_\infty \le 1$. Fix $\forall \varepsilon>0$. $\forall n, \forall k \in \{1,\dots,n\}$:

$$Y_{n,k} := X_k \mathbb{1}_{\{|X_k| \le \varepsilon\sqrt{n}\}} \;;\quad T_n := Y_{n,1}+\dots+Y_{n,n} \;;\quad \mu_n := \mathbb{E}Y_{n,k} \;;\quad \sigma_n^2 := \mathrm{Var}(Y_{n,k})$$

DCT $\Rightarrow$ $\mu_n \to \mathbb{E}X_k = 0$, $\sigma_n^2 \to \mathrm{Var}(X_n) = 1$ as $n\to\infty$.

- Apply Non-asymp. CLT for $\frac{T_n - n\mu_n}{\sqrt{n}} = \sum_{i=1}^n \frac{Y_{n,k}-\mu_n}{\sqrt{n}}$ $\leftarrow$ Mean 0, Variance σ_n^2 $\Rightarrow$

$$\left|\mathbb{E}h\left(\frac{T_n - n\mu_n}{\sqrt{n}}\right) - \mathbb{E}h(\underbrace{Z_n}_{N(0,\sigma_n^2)})\right| \le C\sum_{k=1}^n \mathbb{E}\left|\frac{Y_{n,k}-\mu_n}{\sqrt{n}}\right|^3 \overset{\text{iid}}{=} \frac{C\,\mathbb{E}|Y_{n,1}-\mathbb{E}Y_{n,1}|^3}{\sqrt{n}}$$

$$\le \frac{8C\,\mathbb{E}|Y_{n,1}|^3}{\sqrt{n}} \qquad \left(\|Y-\mathbb{E}Y\|_{L^3} \overset{\Delta}{\le} \|Y\|_{L^3} + \underbrace{|\mathbb{E}Y|}_{\le \mathbb{E}|Y| = \|Y\|_{L^1} \le \|Y\|_{L^3}} \le 2\|Y\|_{L^3}\right)$$

$$\le 8C\varepsilon. \qquad \left(\mathbb{E}|Y_{n,1}|^3 \le \|Y_{n,1}\|_\infty \cdot \mathbb{E}Y_{n,1}^2 \le \varepsilon\sqrt{n}\cdot \mathbb{E}X_1^2 = \varepsilon\sqrt{n}\right) \qquad (*)$$

- Let's get rid of truncation. ① $\mathbb{E}h\left(\frac{T_n-n\mu_n}{\sqrt{n}}\right) \overset{?}{\approx} \mathbb{E}h\left(\frac{S_n}{\sqrt{n}}\right)$ ② $\mathbb{E}h(Z_n) \overset{?}{\approx} \mathbb{E}h(Z)$

$$\left|\mathbb{E}h\left(\frac{T_n-n\mu_n}{\sqrt{n}}\right) - \mathbb{E}h\left(\frac{S_n}{\sqrt{n}}\right)\right| \le \mathbb{E}\left|h\left(\frac{T_n-n\mu_n}{\sqrt{n}}\right) - h\left(\frac{S_n}{\sqrt{n}}\right)\right|$$

$$\le \mathbb{E}\left|\frac{T_n - S_n - n\mu_n}{\sqrt{n}}\right| \qquad (\text{by assumption, } \|h'\|_\infty \le 1)$$

$$\overset{\text{Jensen}}{\le} \sqrt{\mathbb{E}\left(\frac{\overbrace{T_n}^{\text{mean } n\mu_n} - \overbrace{S_n}^{\text{mean } 0} - n\mu_n}{\sqrt{n}}\right)^2} = \sqrt{\frac{\mathrm{Var}(T_n - S_n)}{n}}$$

$$= \sqrt{\mathrm{Var}\left(X_n \mathbb{1}_{\{X_n > \varepsilon\sqrt{n}\}}\right)} \qquad \left(\text{since } S_n - T_n = \sum_{k=1}^n (X_n - Y_{n,k}) = \sum_{k=1}^n X_k \mathbb{1}_{\{|X_k| > \varepsilon\sqrt{n}\}}\right)$$

$$\le \sqrt{\mathbb{E}X_n^2 \mathbb{1}_{\{|X_n| > \varepsilon\sqrt{n}\}}} \longrightarrow 0 \text{ as } n\to\infty \text{ by DCT.}$$

☺

② $\underbrace{Z_n}_{N(0,\sigma_n^2)} \overset{d}{=} \underbrace{\sigma_n}_{\to 1} \underbrace{Z}_{N(0,1)} \xrightarrow{\text{a.s.}} Z \Rightarrow Z_n \xrightarrow{d} Z \overset{\text{P.L.}}{\Rightarrow} \mathbb{E}h(Z_n) \to \mathbb{E}h(Z)$ ☺

- $(*)$ & ① & ② $\overset{\Delta}{\Longrightarrow} \limsup_n \left|\mathbb{E}h\left(\frac{S_n}{\sqrt{n}}\right) - \mathbb{E}h(Z)\right| \le 8C\varepsilon \leftarrow$ arbitrary $\Rightarrow$ □

- Remarks on the nonasymptotic CLT:
  - Rate of convergence $= O(1/\sqrt{n})$
  - $\|f'''\|_\infty$ can be replaced by $\|f'\|_\infty$ (later - Stein's method)
  - Third moment can be replaced by $(2+\varepsilon)$-moment.
  - But the bound gets weaker (Khoshnevisan, Probability. Problem 7.44]

- Remark on the replacement method:
  - It works in wider generality than sums of r.v's
    e.g. in RMT, see e.g. [Chatterjee], [Van-Vu 4th moment thm]

- Application for mean estimation: $\hat{\mu} := \frac{X_1 + \dots + X_n}{n} \overset{?}{\approx} \mu$. $\quad Var(\hat{\mu}) = \frac{\sigma}{\sqrt{n}}$

$$P\Big\{ |\underset{\overset{\|}{S_n/n}}{\hat{\mu}} - \mu| \le \frac{t\sigma}{\sqrt{n}} \Big\} = P\Big\{ \Big| \frac{S_n - n\mu}{\sigma\sqrt{n}} \Big| \le t \Big\} \xrightarrow[\text{CLT}]{n\to\infty} P\{ |\underset{\underset{N(0,1)}{\wr}}{Z}| \le t \} = \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} e^{-x^2} dx$$

Confidence interval.