The CLT arises from independence.
Identical distribution is not strictly required, but some kind of "average uniformity" is needed.

E.g. $X_n \overset{d}{=} \mathcal{N}(0, b_n^2)$ independent. Let $S_n = X_1 + \cdots + X_n$.

$X_n \overset{d}{=} b_n Z \qquad Z \overset{d}{=} \mathcal{N}(0,1)$

$$\varphi_{X_n}(\xi) = \varphi_{b_n Z}(\xi) = \varphi_Z(b_n \xi) = e^{-(b_n\xi)^2/2}$$

$$\therefore \varphi_{S_n/\sqrt{n}}(\xi) = \varphi_{X_1+\cdots+X_n}(\xi/\sqrt{n}) = \varphi_{X_1}(\xi/\sqrt{n}) \cdots \varphi_{X_n}(\xi/\sqrt{n})$$
$$= e^{-b_1^2\xi^2/2n} \cdots e^{-b_n^2\xi^2/2n}$$
$$= e^{-\frac{1}{n}(b_1^2 + \cdots + b_n^2)\xi^2/2}$$
$$\to e^{-b^2\xi^2/2} = \mathcal{N}(0, b^2)^\wedge \quad \checkmark$$

If all $b_n = b$, or even if $b_n \to b$,

But what if, e.g., $b_n = \sqrt{n}$? Then $\sum_{j=1}^{n} b_n^2 = \frac{1}{2}n(n+1)$.

$$\varphi_{S_n/\sqrt{n}}(\xi) = e^{-\frac{1}{\cancel{n}}\cdot\frac{1}{2}\cancel{n}(n+1)\frac{\xi^2}{2}} = e^{-\frac{1}{4}(n+1)\xi^2} \longrightarrow \begin{cases} 1, & \xi = 0 \\ 0, & \xi \neq 0. \end{cases}$$

Triangular Arrays

$\{X_{n,k}\}_{k=1}^{n}$ independent, centered L^2 random variables
$\mathbb{E}[X_{n,k}]=0$, $\mathbb{E}[X_{n,k}^2] = Var[X_{n,k}] =: \sigma_{n,k}^2 < \infty$

$$\begin{array}{lllll} X_{1,1} & & & & \\ X_{2,1} & X_{2,2} & & & \\ X_{3,1} & X_{3,2} & X_{3,3} & & \\ \vdots & \vdots & \vdots & \ddots & \\ X_{n,1} & X_{n,2} & X_{n,3} & \cdots & X_{n,n} \\ \vdots & \vdots & \vdots & & \ddots \end{array} \Bigg\} \; S_n = X_{n,1} + \cdots + X_{n,n}.$$

Better to build in standardization.

Assume $\sum_{k=1}^{n} \sigma_{n,k}^2 = 1$.

Eg. If $\{X_n\}_{n=1}^{\infty}$ are iid with $Var[X_n] = b^2$, the CLT deals with

$$\frac{1}{b\sqrt{n}} \sum_{k=1}^{n} \mathring{X}_k .$$

$$\sqrt{Var \sum_{k=1}^{n} X_k}$$

Why? $Var\left[\sum_{k=1}^{n} X_k\right] = \sum_{k=1}^{n} Var[X_k] = nb^2$

$$X_{n,k} = \frac{\mathring{X}_k}{b\sqrt{n}}$$

Average Uniformity Conditions

$\{X_{n,k}\}_{k=1}^{n}$ centered L^2 random variables

$\mathbb{E}[X_{n,k}]=0$, $\mathbb{E}[X_{n,k}^2] = Var[X_{n,k}] =: \sigma_{n,k}^2$, $\sum_{k=1}^{n} \sigma_{n,k}^2 = 1$.

Below are two conditions that precisely interpret the requirement that "the terms are small and comparable in size"

- (DV) The Decaying Variance condition:

$$\max_{1\le k\le n} \sigma_{n,k}^2 \to 0 \text{ as } n\to\infty$$

If $\sigma_{n,k} = \frac{b_n}{\sqrt{n}}$

$\max_{1\le k\le n} \sigma_{n,k}^2 = \frac{b_n^2}{n}$

i.e. $b_n = o(\sqrt{n})$

- (UAN) The Uniform Asymptotic Negligibility condition:

$$\forall \varepsilon>0, \lim_{n\to\infty} \max_{1\le k\le n} \mathbb{P}(|X_{n,k}|>\varepsilon) = 0.$$

Lemma: (DV) $\Rightarrow$ (UAN).

Pf. By Chebyshev, $\underline{\max_{k\le n} \mathbb{P}(|X_{n,k}|>\varepsilon)} \le \max_{k\le n} \frac{Var[X_{n,k}]}{\varepsilon^2} = \frac{1}{\varepsilon^2} \max_{k\le n} \sigma_{n,k}^2 \to 0.$

$\to 0$. ///

We'd like to prove a CLT for triangular arrays, assuming something like (DV). That's not quite correct. We need slightly stronger conditions.

- (Lind) The Lindberg condition: $\sum_{k=1}^{n} \sigma_{n,k}^2 = 1$

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}[X_{n,k}^2 : |X_{n,k}| > \varepsilon] = 0 \quad \forall \varepsilon > 0.$$

$$\mathbb{E}[X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| > \varepsilon\}}]$$

$$\sum_{k=1}^{n} \left(\sigma_{n,k}^2 - \mathbb{E}[X_{n,k}^2 : |X_{n,k}| \le \varepsilon]\right) \to 0 \text{ as } n \to \infty.$$

This is a decreasing function of $\varepsilon > 0$, so suffices for this to hold for small $\varepsilon > 0$.

Eg. $X_{n,k} = \frac{1}{b\sqrt{n}} \mathring{X}_k$ where $\{X_k\}_{k=1}^{\infty}$ are iid w $\mathrm{Var}[X_k] = b^2$.

$$\sum_{k=1}^{n} \mathbb{E}[X_{n,k}^2 : |X_{n,k}| > \varepsilon] = \frac{1}{b^2 n} \sum_{k=1}^{n} \mathbb{E}[\mathring{X}_k^2 : |\mathring{X}_k| > |b|\sqrt{n}\,\varepsilon]$$

$$= \frac{1}{b^2} \mathbb{E}[\mathring{X}_1^2 \mathbb{1}_{|\mathring{X}_1| > |b|\sqrt{n}\,\varepsilon}] \to 0 \text{ by DCT.}$$

Prop: (Lind) $\Rightarrow$ (DV).

Pf. $\sigma_{n,k}^2 = \mathbb{E}[X_{n,k}^2] = \mathbb{E}[X_{n,k}^2(\mathbb{1}_{|X_{n,k}|\le\varepsilon} + \mathbb{1}_{|X_{n,k}|>\varepsilon})]$

$$= \mathbb{E}[X_{n,k}^2 : |X_{n,k}|\le\varepsilon] + \mathbb{E}[X_{n,k}^2 : |X_{n,k}|>\varepsilon]$$

$$\le \varepsilon^2 + \underbrace{\sum_{j=1}^n \mathbb{E}[X_{n,j}^2 : |X_{n,j}|>\varepsilon]}_{\to 0 \text{ as } n\to\infty \text{ by (Lind)}}$$

$\therefore$ $\max_{1\le k\le n} \sigma_{n,k}^2$

$\therefore \limsup_{n\to\infty} \max_{1\le k\le n} \sigma_{n,k}^2 \le \varepsilon^2 + 0$. $\therefore \lim_{n\to\infty} \max_{1\le k\le n} \sigma_{n,k}^2 \to 0$. ///

Lindberg-Feller CLT

$\{X_{n,k}\}_{k=1}^n$ centered independent L^2 random variables, $S_n = X_{n,1} + \cdots + X_{n,n}$

$\mathbb{E}[X_{n,k}]=0$, $\mathbb{E}[X_{n,k}^2] = \mathrm{Var}[X_{n,k}] =: \sigma_{n,k}^2$, $\sum_{k=1}^n \sigma_{n,k}^2 = 1$. (i.e. $\mathrm{Var}[S_n]=1$)

If the Lindberg condition holds for $\{X_{n,k}\}$, then $S_n \to_w \mathcal{N}(0,1)$.

Conversely: if $\{X_{n,k}\}$ satisfy the decaying variance condition (DV) and if $S_n \to_w \mathcal{N}(0,1)$, then the Lindberg condition holds for $\{X_{n,k}\}$.

Historical note: before Lindberg-Feller, a "triangular CLT" was proved by Liapunov, with an even stronger hypothesis.

- (Liap) $\exists\, \alpha > 2$ s.t. $\lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}[\,|X_{n,k}|^{\alpha}] = 0$

$\phi(t) = t^{\alpha-2}$.

or a bit weaker:

$\exists\ \phi : [0,\infty) \to [0,\infty)$ non-decreasing, $\phi(t) > 0\ \forall t > 0$, s.t.

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}[\,X_{n,k}^{2}\, \phi(X_{n,k})] = 0$$

Lemma: (Liap) $\Rightarrow$ (Lind)

Pf.

$$\sum_{k=1}^{n} \mathbb{E}[\,X_{n,k}^{2} : |X_{n,k}| > \varepsilon]$$

$t > \varepsilon \Rightarrow \phi(t) \geq \phi(\varepsilon)$

$\dfrac{\phi(t)}{\phi(\varepsilon)} \geq 1$

$$\leq \mathbb{E}\left[X_{n,k}^{2}\, \frac{\phi(|X_{n,k}|)}{\phi(\varepsilon)} : |X_{n,k}| > \varepsilon\right]$$

$$\leq \frac{1}{\phi(\varepsilon)}\, \mathbb{E}\left[X_{n,k}^{2}\, \phi(|X_{n,k}|)\right]$$

///