Probability Theory

VII Central Limit Theorem

Seongho, Joo

August 7, 2025

Seoul National University

Characteristic Functions

Definition 1

Let $\mu \in \mathrm{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$. For $\xi \in \mathbb{R}^d$, define

$$\hat{\mu(\xi)} := \int_{\mathbb{R}^d} e^{i\xi \cdot x} \mu(\,\mathrm{d}x)$$

the *Fourier transform* of μ.

If $\bar{X} : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$ is a random vector, its *chracteristic function* $\varphi_{\bar{X}} : \mathbb{R}^d \to \mathbb{C}$ is

$$\varphi_{\underline{\overline{\mathbf{X}}}}(\xi) = \hat{\mu}_{\underline{\overline{\mathbf{X}}}}(\xi) = \mathbb{E}\, e^{i\xi \cdot \bar{X}}$$

Proposition 1

$\mu \mapsto \hat{\mu}$ is injective: if $\hat{\mu}(\xi) = \hat{\nu}(\xi)\ \forall \xi \in \mathbb{R}^d$, then $\mu = \nu$.

Thus, in principle, we can recover μ from $\hat{\mu}$.

Characteristic Functions

Theorem 1

If $\mu \in \text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, then for $a < b$ in $\mathbb{R}$,

$$\lim_{R \to \infty} \frac{1}{2\pi} \int_{-R}^{R} \hat{\mu}(\xi) \left(\frac{e^{-ia\xi} - e^{-ib\xi}}{i\xi} \right) \mathrm{d}\xi = \mu((a,b)) + \frac{1}{2}\mu(\{a,b\})$$

In addition,

$$\mu(\{a\}) = \lim_{R \to \infty} \frac{1}{2R} \int_{-R}^{R} e^{-ia\xi} \hat{\mu}(\xi) \, \mathrm{d}\xi$$

Corollary 1

If $\mu \in \text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $\hat{\mu} \in L^1(\lambda)$, then $\mu << \lambda$ and its density $\rho = \frac{\mathrm{d}\mu}{\mathrm{d}\lambda}$ is

$$\rho(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{\mu}(\xi) e^{-I\xi x} \, \mathrm{d}\xi$$

Characteristic Functions

Proposition 2 (Properties of the Fourier Transform $\hat{\mu}$)

1. $\hat{\mu}(0) = 1$ and $|\hat{\mu}(\xi)| \leq 1 \; \forall \xi \in \mathbb{R}^d$
2. $\hat{\mu} \in C_{\mathbb{C}}(\mathbb{R}^d)$
3. $\overline{\hat{\mu}(\xi)} = \hat{\mu}(-\xi) \; \forall \xi \in \mathbb{R}^d$. In particular, $\hat{\xi}$ is $\mathbb{R}$-valued if and only if μ is symmetric $(\mu(B) = \mu(-B) \; \forall B \in \mathcal{B}\left(\mathbb{R}^d\right))$
4. If $\int_{\mathbb{R}^d} |x|^k \mu(\mathrm{d}x) < \infty$ then $\hat{\mu} \in C_{\mathbb{C}}^k$ and
$$\frac{\partial}{\partial \xi_{j_1}} \cdots \frac{\partial}{\partial \xi_{j_k}} = \int_{\mathbb{R}^d} (ix_{j_1}) \ldots (ix_{j_k}) e^{i\xi \cdot x} \mu(\mathrm{d}x)$$

Proof.

Proposition 3

If $\mu, \nu \in \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$, then

$$\widehat{\mu * \nu} = \hat{\mu} \cdot \hat{\nu}$$

I.e. If $\underline{\overline{\mathbf{X}}}, \underline{\overline{\mathbf{Y}}}$ are independent random vectors in $\mathbb{R}^d$, then

$$\varphi_{\underline{\overline{\mathbf{X}}}+\underline{\overline{\mathbf{Y}}}}(\xi) = \varphi_{\underline{\overline{\mathbf{X}}}}(\xi) \cdot \varphi_{\underline{\overline{\mathbf{Y}}}}(\xi) \quad \forall \xi \in \mathbb{R}^d$$

Moreover, if $a \in \mathbb{R}, v \in \mathbb{R}^d$, then $\varphi_{a\underline{\overline{\mathbf{X}}}+v} = e^{i\xi \cdot v} \varphi_{\underline{\overline{\mathbf{X}}}}(a\xi)$

Example.

- $N \overset{d}{=}$ Poisson(λ). $\varphi_N(\xi) = \mathbb{E}\, e^{i\xi \cdot N} = \sum_{n=0}^{\infty} e^{i\xi \cdot n} e^{-\lambda} \frac{\lambda^n}{n!} = e^{\lambda(e^{i\xi}-1)}$

- $Y \overset{d}{=}$ Rademacher : $\mathbb{P}(Y \pm 1) = \frac{1}{2}$. $\varphi_Y(\xi) = \mathbb{E}\, e^{i\xi \cdot Y} = \frac{1}{2} e^{i\xi 1} + \frac{1}{2} e^{i\xi(-1)} = \cos \xi$

So, if $Y_1, \ldots, Y_N$ and iid Rademachers, $S_n = Y_1 + \ldots Y_n$,

$$\varphi_{S_n}(\xi) = \varphi_{Y_1}(\xi) \ldots \varphi_{Y_n}(\xi) = (\cos \xi)^n$$

By the taylor theorem, for some $\eta \in (0, t)$

$$\log \varphi_{S_n/b_n}(\xi) = n \cdot (-\sec^2(n/b_n))\xi^2/b_n^2$$

Take $b_n = \sqrt{n}, \log \varphi_{S_n/\sqrt{n}} \to -\frac{1}{2}\xi^2$, and $\varphi_{S_n/\sqrt{n}}(\xi) \to e^{-\frac{1}{2}\xi^2}$ which the characteristic function of $\mathcal{N}(0,1)$.

Riemann-Lebesgue

If μ admits a density ρ w.r.t Lebesgue measure, we denote $\hat{\mu} = \hat{\rho}$.

Lemma 1 (Riemann-Lebesgue)

If $\rho \in L^1$, then $\hat{\rho} \in C_0$, i.e. $\hat{\rho}(\xi) \to 0$ as $\to \infty$

Proof.

Step 1: Show that the result holds for $\rho \in C_c^\infty(\mathbb{R}^d)$
Step 2: For general $\rho \in L^1(\mathbb{R}^d, \lambda)$, approximate by C_c^∞ functions.
Step 3: Combine. Let $\varepsilon > 0$, and $\psi \in C_c^\infty(\mathbb{R}^d)$ s.t. $\|\rho - \psi\|_{L^1} < \varepsilon/2$

Continuity Theorem

If $\mu_n \xrightarrow{w} \mu$, then $\hat{\mu}(\xi) \to \hat{\mu}(\xi)$ $\forall \xi \in \mathbb{R}^d$. The converse also holds!

Theorem 2 (Continuity Theorem)

Let $\{\mu_n\}_{n=1}^{\infty} \subset \mathrm{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$. Suppose that $\varphi(\xi) := \lim_{n\to\infty} \hat{\mu}_n(\xi)$ exists $\forall \xi \in \mathbb{R}^d$. If φ is continuous at 0, then $\exists \mu \in \mathrm{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$ such that $\varphi = \hat{\mu}$, and $\mu_n \xrightarrow{w} \mu$.

Example. The Scaled sum of iid Rademacher random variables converges to uniform normal distribution: $\frac{X_1+\cdots+X_n}{\sqrt{n}} \xrightarrow{w} \mathcal{N}(0,1)$.

Continuity Theorem

If $\mu_n \xrightarrow{w} \mu$, then $\hat{\mu}(\xi) \to \hat{\mu}(\xi)$ $\forall \xi \in \mathbb{R}^d$. The converse also holds!

Theorem 2 (Continuity Theorem)

Let $\{\mu_n\}_{n=1}^{\infty} \subset \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$. Suppose that $\varphi(\xi) := \lim_{n\to\infty} \hat{\mu}_n(\xi)$ exists $\forall \xi \in \mathbb{R}^d$. If φ is continuous at 0, then $\exists \mu \in \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$ such that $\varphi = \hat{\mu}$, and $\mu_n \xrightarrow{w} \mu$.

Example. The Scaled sum of iid Rademacher random variables converges to uniform normal distribution: $\frac{X_1+\cdots+X_n}{\sqrt{n}} \xrightarrow{w} \mathcal{N}(0,1)$.

Proposition 4 (Characteristic tail estimate)

Let ρ be a probability density on $\mathbb{R}^d$, supported in $\bar{B}_1$. Let $M > 0$ be such that $|\hat{\rho}(\xi)| \leq \frac{1}{2}$ for all $|\xi| \geq M$.
Then $\forall \mu \in \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$, and $a > 0$,

$$\mu(\left\{x \in \mathbb{R}^d : |x| \geq a\right\}) \leq 2 \int_{B_1} \left[1 - \Re\hat{\mu}(\frac{M}{a}x)\right] \rho(x)\, \mathrm{d}x$$

Continuity Theorem

Corollary 2

If $\{\mu_n\}_{n=1}^{\infty} \subset \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)$ are such that $\varphi(\xi) := \lim_{n\to\infty} \hat{\mu}_n(\xi)$ exists $\forall \xi \in \mathbb{R}^d$ and φ is continuous at 0, then $\{\mu_n\}_{n=1}^{\infty}$ is tight.

Continuity Theorem

Corollary 2

If $\{\mu_n\}_{n=1}^{\infty} \subset \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)$ are such that $\varphi(\xi) := \lim_{n\to\infty} \hat{\mu}_n(\xi)$ exists $\forall \xi \in \mathbb{R}^d$ and φ is continuous at 0, then $\{\mu_n\}_{n=1}^{\infty}$ is tight.

Proof. Fix ρ, M as in the tail estimate proposition:

$$\mu_n\left\{x \in \mathbb{R}^d : |x| \geq a\right\} \leq 2\int_{B_1}\left[1 - \Re\hat{\mu}_n(\frac{M}{a}x)\right]\rho(x)\,\mathrm{d}x \leq 2\delta(a) + \frac{\varepsilon}{2}\ldots(\dagger)$$

where $\delta(a) = \sup_{|x|\leq x}|1 - \Re\varphi(\frac{M}{a}x)|$, since φ is continuous at 0, $\lim_{a\to\infty}\delta(a) = 0$
Fix $\varepsilon > 0$, choose a large enough so that $\delta(a) < \frac{\varepsilon}{4}$. Choose N such that $\forall n \geq N$
† holds,

$$\mu_n(\mathbb{R}^d \setminus \bar{B}_a) \leq 2\delta(a) + \frac{\varepsilon}{2} < \varepsilon$$

Proof for continuity Theorem

Theorem 3 (Continuity Theorem)

Let $\{\mu\}_{n=1}^{\infty} \subset \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)$. Suppose that $\varphi(\xi) := \lim_{n\to\infty} \hat{\mu}_n(\xi)$ exists $\forall \xi \in \mathbb{R}^d$. If φ is continuous at 0, then $\exists \mu \in \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$ such that $\varphi = \hat{\mu}$, and $\mu_n \xrightarrow{w} \mu$.

Proof. By the preceding corollary, $\{\mu\}_{n=1}^{\infty}$ is tight. By Prokhorov, $\exists$ subsequence s.t. $\mu_{n_k} \xrightarrow{w} \mu$ for some $\mu \in \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}\right))$. Therefore
$\hat{\mu}_{n_k}(\xi) \to \hat{\mu}(\xi) = \varphi(\xi) \quad \forall \xi \in \mathbb{R}^d$.

Claim: $\mu_n \xrightarrow{w} \mu$

If not, $\exists g \in C_b(\mathbb{R}^d)$ s.t. $\int g \, \mathrm{d}\mu_n \not\to \int g \, \mathrm{d}\mu$

I.e. $\exists \varepsilon > 0, \exists n_k'$ s.t. $\left| \int g \, \mathrm{d}\mu_{n_k'} - \int g \, \mathrm{d}\mu \right| \geq \varepsilon \quad \forall k$

By Prokhorov, $\exists$ further subsequence $\{n_k''\}_{k=1}^{\infty} \subseteq \{n_k'\}_{k=1}^{\infty}$ s.t. $\mu_{n_k''} \xrightarrow{w} \nu$ for some $\nu \in \text{Prob}(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)) \implies \hat{\nu} = \hat{\mu}$. By the injectivity of Fourier transform, $\nu = \mu \implies \mu_{n_k''} \to \mu$. However, this contradicts the assumption that n_k' s.t. $\left| \int g \, \mathrm{d}\mu_{n_k'} - \int g \, \mathrm{d}\mu \right| \geq \varepsilon \quad \forall k$.

Basic Central Limit Theorem

Theorem 4

Let $\{x_n\}_{n=1}^{\infty}$ be i.i.d. L^2 random variables with common mean $\mathbb{E}\, x_n = t$ and variance $\mathrm{Var} x_n = \sigma^2$. Let $S_n = X_1 + \cdots + X_n$. $\frac{\mathring{S}_n}{\sigma\sqrt{n}} = \frac{S_n - nt}{\sigma\sqrt{n}} \xrightarrow{w} Z \overset{d}{=} \mathcal{N}(0,1)$

Proof. By Levy's continuity theorem, it suffices to show that

$$\varphi_{\mathring{S}_n/\sigma\sqrt{n}}(\xi) \to e^{-\xi^2/2} \quad \forall \xi \in \mathbb{R}$$

$$\varphi_{\mathring{S}_n/\sigma\sqrt{n}} = \varphi_{\mathring{S}_n}(\xi/\sigma\sqrt{n}) = \varphi_{\mathring{X}_1+\cdots+\mathring{X}_n}(\xi/\sigma\sqrt{n}) = \varphi_{\mathring{X}_1}(\xi/\sigma\sqrt{n})^n$$

Note that $X_1 \in L^2$, so $\mathbb{E}\, \mathring{X}_1^2 = \mathrm{Var} X_1 = \sigma^2 < \infty$, $\therefore \varphi_{\mathring{X}_1} \in C^2$

By Taylor's theorem,

$$\varphi_{\mathring{X}_1}(x) = \varphi_{\mathring{X}_1}(0) + \varphi'_{\mathring{X}_1}(0)x + \frac{1}{2}\varphi''_{\mathring{X}_1}(r(x))x^2, \qquad \text{form some } r(x) \text{ between } 0 \text{ and } x$$

$$= 1 + \frac{1}{2}\varphi''(r(x))x^2$$

$$\therefore (\varphi_{\mathring{X}_1}(\xi/\sigma\sqrt{n}))^n = \left(1 + \frac{1}{2}\varphi''_{\mathring{X}_1}(r(\xi/\sigma\sqrt{n})(\frac{\xi}{\sigma\sqrt{n}})^2\right)^n$$

$$\lim_{n\to\infty} \varphi_{\mathring{X}_1}(\xi/\sigma\sqrt{n})^n = \lim_{n\to\infty} \left(1 + \tfrac{1}{2}(-\sigma^2)(\tfrac{\xi^2}{\sigma^2 n})\right)^n = e^{-\xi^2/2}$$

Central Limit Theorem

There is a similar CLT for iid random *vectors*, with any given (common) covariance of entries.

Definition 2

Let Q be a positive definite $d \times d$ matrix i.e. $Q = AA^\top$ for some $d \times d$ matrix A. The centered normal distribution of covariacne Q is the unique measure $\mu \in$ Prob$(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right))$ with $\hat{\mu} = e^{-\frac{1}{2}Q\xi\cdot\xi} = e^{-\frac{1}{2}|A\xi|^2}$. Denote is as $\mathcal{N}(0, Q)$

- If $\underline{\overline{\mathbf{X}}} \overset{d}{=} \mathcal{N}(0, Q)$, then $\mathrm{Cov} X_i X_j = Q_{ij}$, and $X_i \overset{d}{=} \mathcal{N}(0, Q_{ii})$

Central Limit Theorem

There is a similar CLT for iid random *vectors*, with any given (common) covariance of entries.

Definition 2

Let Q be a positive definite $d \times d$ matrix i.e. $Q = AA^\top$ for some $d \times d$ matrix A. The centered normal distribution of covariacne Q is the unique measure $\mu \in$ Prob$(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with $\hat{\mu} = e^{-\frac{1}{2}Q\xi\cdot\xi} = e^{-\frac{1}{2}|A\xi|^2}$. Denote is as $\mathcal{N}(0, Q)$

- If $\underline{\overline{\mathbf{X}}} \overset{d}{=} \mathcal{N}(0, Q)$, then $\mathrm{Cov} X_i X_j = Q_{ij}$, and $X_i \overset{d}{=} \mathcal{N}(0, Q_{ii})$

Theorem 5 (Multivariate CLT)

If $\{\underline{\overline{\mathbf{X}}}_n\}\, n$ are i.i.d random vectros in $\mathbb{R}^d$ with L^2 entries, and $Q = \underline{\overline{\mathring{\mathbf{X}}}}_1 \underline{\overline{\mathring{\mathbf{X}}}}_1^\top$, then

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \underline{\overline{\mathring{\mathbf{X}}}}_j \xrightarrow{w} Z_1 \overset{d}{=} \mathcal{N}(0, Q)$$

Multivariate CLT

Lemma 2 (Cramer-Wold Device)

Let $\{\underline{\overline{\mathbf{X}}}_n\}_{n=1}^{\infty}$ and $\underline{\overline{\mathbf{X}}}$ be random vectors in $\mathbb{R}^d$. Then $\underline{\overline{\mathbf{X}}}_n \xrightarrow{w} \underline{\overline{\mathbf{X}}}$ if and only $\xi \cdot \underline{\overline{\mathbf{X}}}_n \xrightarrow{w} \xi \cdot \underline{\overline{\mathbf{X}}} \quad \forall \xi \in \mathbb{R}^d$.

Theorem 5 Proof. Fix $\xi \in \mathbb{R}^d$. Let $X_n^{\xi} := \xi \cdot \underline{\overline{\mathbf{X}}}$. Then $\{X_n^{\xi}\}_{n=1}^{\infty}$ are independent, and

$$\varphi_{X_n^{\xi}}(u) = \mathbb{R}e^{iu\xi \cdot \underline{\overline{\mathbf{X}}}_n} = \varphi_{\underline{\overline{\mathbf{X}}}_n}(u\xi) = \varphi_{\underline{\overline{\mathbf{X}}}_1}(u\xi)$$

$\therefore \{X_n^{\xi}\}_{n=1}^{\infty}$ are i.i.d. They are in L^2.

$$\begin{aligned}
\mathbb{E}\, X_n^{\xi} &= \mathbb{E}\, \xi \cdot \underline{\overline{\mathbf{X}}}_n = \xi \cdot \mathbb{E}\, \underline{\overline{\mathbf{X}}}_n = \xi \cdot \mathbb{E}\, \underline{\overline{\mathbf{X}}}_1 \\
\mathrm{Var} X_n^{\xi} &= \mathbb{E}\, (\xi \cdot \underline{\overline{\mathbf{X}}}_n)^2 - (\xi \cdot \mathbb{E}\, \underline{\overline{\mathbf{X}}}_n)^2 \\
&= \mathbb{E}\, \xi \cdot \underline{\overline{\mathbf{X}}}_n \underline{\overline{\mathbf{X}}}_n^{\top} - \xi \cdot \mathbb{E}\, \vec{x}\, \mathbb{E}\, \vec{x}^{\top} \xi \\
&= \xi \cdot (\underbrace{\mathbb{E}\, \underline{\overline{\mathbf{X}}\overline{\mathbf{X}}}^{\top} - \mathbb{E}\, \underline{\overline{\mathbf{X}}}_n\, \mathbb{E}\, \underline{\overline{\mathbf{X}}}_n^{\top}}_{= \mathbb{E}\, (\underline{\overline{\mathbf{X}}}_1 - \mathbb{E}\, \underline{\overline{\mathbf{X}}}_1)(\underline{\overline{\mathbf{X}}}_1 - \mathbb{E}\, \underline{\overline{\mathbf{X}}}_1)^{\top} = Q} \xi = \xi \cdot Q\xi
\end{aligned}$$

By basic CLT,

$$\frac{1}{\sqrt{Q\xi \cdot \xi}} \frac{1}{\sqrt{n}} \sum_{j-1}^{n} \left(X_j^{\xi} - \xi \cdot \mathbb{E}\, \underline{\overline{\mathbf{X}}}_1 \right) \xrightarrow{w} \mathcal{N}(0, 1)$$

Infinite divisibility

Definition 3

A probability measure $\mu \in \text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is *infinitely divisible* if, for each $n \in \mathbb{N}$ $\exists \mu_n \in \text{Prov}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\mu = \mu_n^{*n} = \mu_n * \mu_n * \cdots * \mu_n$
I.e. $\exists \{X_{n,k}\}_{k=1}^{\infty}$ iid s.t. $S_n = X_{n,1} + \ldots X_{n,n} \overset{d}{=} \mu$
I.e. $\exists$ non-constant characteristic function φ_n such that $\hat{\varphi}(\xi) = \varphi_n(\xi)^n \quad \forall \xi \in \mathbb{R}^d$

Example. If $X_{n,k} \overset{d}{=} \mathcal{N}(0, \sigma^2/n)$ are independent, then
$S_n = X_{n,1} + \ldots X_{n,n} \overset{d}{=} \mathcal{N}(0, \sigma^2)$
Note: If μ, ν are infinitely divisible, so is $\mu * \nu = (\mu_n * \nu_n)^*$

When a measure is infinite divisible?

Theorem 6

A probability measure $\mu \in \text{Prob}(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is infinitely divisible if and only if $\exists$ a triangular array $\{X_{n,k}\}_{k=1}^{m_n} \quad m_n \uparrow \infty, n \in \mathbb{N}$ of random variables such that for each $n, \{X_{n,k}\}_{k=1}^{m_n}$ are iid, and

$$S_n = \sum_{k=1}^{m_n} X_{n,k} \xrightarrow{w} X \stackrel{d}{=} \mu$$

Uniformity

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

Triangular Arrays

$\{X_{n,k}\}_{k=1}^{n}$ independent, centered L^2 random variables such that $\mathbb{E}\,X_{n,k} = 0, \mathbb{E}\,X_{n,k}^2 = \mathrm{Var}X_{n,k} = \sigma_{n,k}^2 < \infty$. We may assume $\sum_{k=1}^{n} \sigma_{n,k}^2 = 1$

Uniformity

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

Triangular Arrays

$\{X_{n,k}\}_{k=1}^{n}$ independent, centered L^2 random variables such that
$\mathbb{E}\, X_{n,k} = 0, \mathbb{E}\, X_{n,k}^2 = \mathrm{Var} X_{n,k} = \sigma_{n,k}^2 < \infty$. We may assume $\sum_{k=1}^{n} \sigma_{n,k}^2 = 1$

Definition 4 (Average Uniformity Conditions)

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

- **DV**: The Decaying Variance condition:

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

- **UAN**: The uniform Asymptotic Negligibility condition:

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

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

Note: DV condition implies UAN condition:

$$\max_{1 \le k \le n} \mathbb{P}(|X_{n,k}| > \varepsilon) \le \max_{1 \le k \le n} \frac{\mathrm{Var} X_{n,k}}{\varepsilon^2} \le \frac{1}{\varepsilon^2} \max_{k \le n} \sigma_{n,k}^2 \to 0$$

Uniformity

We'd like to prove a CLT for triangular arrays assuming something like (DV). Actually, slightly stronger conditions:

• (Lind) The Lindberge condition:

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

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

$$\sum_{k=1}^{n}\mathbb{E}\,X_{n,k}^2 : |X_{n,k}| > \varepsilon = \frac{1}{b^2 n}\sum_{k=1}^{n}\mathbb{E}\,\mathring{X}_k^2 : |\mathring{X}_k| > |b|\sqrt{n}\varepsilon = \frac{1}{b^2}\,\mathbb{E}\,\mathring{X}_1^2\mathbf{1}_{|\mathring{X}_1|>|b|\sqrt{n}\varepsilon} \to 0$$

Uniformity

We'd like to prove a CLT for triangular arrays assuming something like (DV). Actually, slightly stronger conditions:

- (Lind) The Lindberge condition:

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

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

$$\sum_{k=1}^{n}\mathbb{E}\,X_{n,k}^2 : |X_{n,k}| > \varepsilon = \frac{1}{b^2 n}\sum_{k=1}^{n}\mathbb{E}\,\mathring{X}_k^2 : |\mathring{X}_k| > |b|\sqrt{n}\varepsilon = \frac{1}{b^2}\,\mathbb{E}\,\mathring{X}_1^2\mathbf{1}_{|\mathring{X}_1|>|b|\sqrt{n}\varepsilon} \to 0$$

Proposition 5

Lindberg condition $\implies$ DV condition

Proof.

Lindberg CLT

Lindberg CLT

If $\{X_{n,k}\}_{1\le k\le n}^{n\in\mathbb{N}}$ is a standard triangular array satisfying Lindberg condition, then $S_n \xrightarrow{w} \mathcal{N}(0,1)$.

Lemma 3

If $a_j, b_j \in \mathbb{C}$ with $|a_j|, |b_j| \le 1$, then

$$|a_1 a_2 \dots a_N - b_1 b_2 \dots b_n| \le \sum_{j=1}^{n} |a_j - b_j|$$

Proof. Proceed by induction.

Lemma 4

If $X \in L^2$, $|\varphi_X(\xi) - \left(1 + i\,\mathbb{E}\,X\xi - \frac{1}{2}\,\mathbb{E}\,X^2\xi^2\right)| \le \xi^2\mathcal{E}(\xi)$

where $\mathcal{E}(\xi) = \mathbb{E}\,X^2 \wedge \frac{|X|^3}{3!}|\xi| \downarrow 0$ as $|\xi| \to 0$ by DCT

Proof. Taylor's theorem: $|e^{it} - (1 + it - \frac{1}{2}t^2)| \le \frac{|it|^3}{3!}$

Lindberg CLT

Lindberg CLT

If $\{X_{n,k}\}_{1\le k\le n}^{n\in\mathbb{N}}$ is a standard triangular array satisfying Lindberg condition, then $S_n \xrightarrow{w} \mathcal{N}(0,1)$.

Proof. Suffice to show $\varphi_{S_n}(\xi) \to e^{-\frac{\xi^2}{2}}\ \forall \xi \in \mathbb{R}$.

LHS $= \varphi_{X_{n,1}} \dots \varphi_{X_{n,n}}(\xi)$, RHS $= e^{-\frac{\xi^2}{2}\sigma_{n,1}^2} \dots e^{-\frac{\xi^2}{2}\sigma_{n,n}^2}$

By Lemma 3, $\left|\varphi_{S_n}(\xi) - e^{-\frac{\xi^2}{2}}\right| \le \sum_{k=1}^n \left|\varphi_{X_{n,k}}(\xi) - e^{-\frac{\xi^2}{2}\sigma_{n,k}^2}\right|$.

Note that $\varphi_{X_{n,}}(\xi) \approx 1 + i\,\mathbb{E}\,X_{n,k}\xi - \frac{1}{2}\,\mathbb{E}\,X_{n,k}^2\xi^2 = 1 - \frac{1}{2}\sigma_{n,k}^2\xi^2$

$$\left|\varphi_{X_{n,k}}(\xi) - e^{-\frac{\xi^2}{2}\sigma_{n,k}^2}\right| \le \underbrace{\left|\varphi_{X_{n,k}}(\xi) - \left(1 - \frac{1}{2}\sigma_{n,k}^2\xi^2\right)\right|}_{A_{n,k}} + \underbrace{\left|\left(1 - \frac{1}{2}\sigma_{n,k}^2\xi^2\right) - e^{-\frac{\xi^2}{2}\sigma_{n,k}^2}\right|}_{B_{n,k}}$$

Suffices to show $\sum_{k=1}^n (A_{n,k} + B_{n,k}) \to 0$ as $n \to \infty$.

Lindberg CLT

$$A_{n,k} = \left|\varphi_{X_{n,k}} - \left(1 - \frac{1}{2}\sigma_{n,k}^2\xi^2\right)\right| \underset{Lemma 17}{\leq} \xi^2\, \mathbb{E}\, X_{n,k}^2 \wedge |\xi| \frac{|X_{n,k}|^3}{3!}$$

$$\leq \xi^2 \left(\mathbb{E}\, X_{n,k}^2 \wedge \frac{|\xi|}{3!}|X_{n,k}|^3 : |X_{n,k}| \leq \varepsilon + \mathbb{E}\, X_{n,k}^2 \wedge \frac{|\xi|}{3!}|X_{n,k}|^3 : |X_{n,k}| > \varepsilon\right)$$

$$\leq \frac{|\xi|^3}{3!}\varepsilon\sigma_{n,k}^2 + \varepsilon\, \mathbb{E}\, X_{n,k}^2 : |X_{n,k}^2| > \varepsilon$$

$$\therefore \sum_{k=1}^n A_{n,k} \leq \frac{|\xi|^3}{3!}\varepsilon \sum_{k=1}^n \sigma_{n,k}^2 + \xi^2 \sum_{k=1}^n \mathbb{E}\, X_{n,k}^2 : |X_{n,k}| > \varepsilon$$

The second term goes to 0 as $n \to \infty$ by the Lindberg condition.

$\therefore \limsup_{n\to\infty} \sum_{k=1}^n A_{n,k} \leq \frac{|\xi|^3}{6}\varepsilon\ \forall\varepsilon > 0$

Lindberg CLT

$$B_{n,k} = \left| e^{-\frac{\xi^2}{2}\sigma_{n,k}^2} - \left(1 - \frac{1}{2}\sigma_{n,k}^2\xi^2\right) \right|$$

Note that: $|e^{-u} - (1-u)| \leq \frac{u^2}{2}\ \forall u \geq 0$

$$\therefore \sum_{k=1}^{n} B_{n,k} \leq \sum_{k=1}^{n} \frac{1}{2}\left(\frac{1}{2}\sigma_{n,k}^2\xi^2\right)^2 = \frac{1}{8}\xi^4 \sum_{k=1}^{n} \sigma_{n,k}^4$$

Note that $\sigma_{n,k}^4 = \sigma_{n,k}^2 \cdot \sigma_{n,k}^2 \leq \max_{1\leq j\leq n} \sigma_{n,j}^2 \sigma_{n,k}^2$

$$\therefore \frac{1}{8}\xi^4 \sum_{k=1}^{n} \sigma_{n,k}^4 \leq \frac{1}{8}\xi^4 \underbrace{\max_{1\leq j\leq n} \sigma_{n,j}^2}_{\to 0 \text{ as } n\to\infty} \cdot \sum_{k=1}^{n} \sigma_{n,k}^2$$