CHAPTER 3: CENTRAL LIMIT THEOREM

September 7, 2021

Contents

1 Convergence in Distribution **2**

2 Characteristic Functions **7**

3 Central Limit Theorem **10**

4 Related Topics **12**

1 Convergence in Distribution

Definition. F_n is a sequence of distribution functions (d.f.). F is a d.f.. F_n is said to **converge weakly** to F, denoted by $F_n \Rightarrow F$, if $F_n(y) \to F(y)$ for every continuity point of F.

Definition. X_n is a sequence of random variables. X is a random variable. X_n converges to X **weakly (or in distribution)**, denoted by $X_n \Rightarrow X$ (or $X_n \xrightarrow{d} X$), if $F_{X_n} \Rightarrow F_X$.

Fact. If $X_n \to X$ in probability, then $X_n \Rightarrow X$. If $X_n \Rightarrow c$ where c is a constant, then $X_n \to c$ in probability.

Note that unlike the convergence concepts we have learnt so far, convergence in distribution does not require X_n and X to be defined on the same probability space. However, the following theorem suggests that we can if we want.

Theorem 3.2.2. (Skorokhod's Theorem). If $F_n \Rightarrow F_\infty$, then we can construct random variables $Y_n, 1 \le n \le \infty$ on the same probability space such that Y_n has d.f. F_n for $1 \le n \le \infty$ and $Y_n \to Y_\infty$ $a.s.$

Proof. Let $\Omega = (0,1)$, $\mathcal{F} = \{\text{Borel sets}\}$, $P =$Lebesgue measure. Define $Y_n : \Omega \to \mathbb{R}$ to be

$$Y_n(\omega) = F_n^{-1}(\omega),$$

where

$$F_n^{-1}(\omega) := \inf\{y : F_n(y) \ge \omega\} = \sup\{y : F_n(y) < \omega\}.$$

Recall from Chapter 1 that the d.f. of the above constructed Y_n is F_n. Let Ω_0 consists of those $x \in (0,1)$ such that the preimage of F at x is either empty or is a unique real number. Then Ω_0^c is countable; hence Ω_0 has probability 1. Moreover, it follows from simple calculus that $Y_n(x) \to Y_\infty(x)$ for all $x \in \Omega_0$. □

Corollary. A version of Fatou's lemma: Let $g \ge 0$ be a continuous function. If $X_n \Rightarrow X_\infty$, then $\liminf_{n\to\infty} Eg(X_n) \ge Eg(X_\infty)$.

Slutsky's Theorem. If $X_n \Rightarrow X$ and $Y_n \to c$ in probability, where c is a constant. Suppose the random variables are defined on the same probability space. Then

$$X_n + Y_n \Rightarrow X + c, \quad X_n Y_n \Rightarrow cX.$$

Proof. Let $x - c$ be a continuity point of F_X (so that x is a continuity point of F_{X+c}). Choose a decreasing sequence of $\epsilon \downarrow 0$ such that $x - c + \epsilon$ is a continuity point of F_x for every ϵ in the sequence (we can do this because continuity points are dense on $\mathbb{R}$). We have

$$\begin{aligned} P(X_n + Y_n \le x) \le & P(|Y_n - c| \ge \epsilon) + P(X_n + Y_n \le x, |Y_n - c| < \epsilon) \\ \le & P(|Y_n - c| \ge \epsilon) + P(X_n \le x - c + \epsilon) \\ \to & P(X + c \le x + \epsilon) \quad \text{as } n \to \infty, \end{aligned}$$

where in the last step, we used the conditions $Y_n \to c$ in probability and $X_n \Rightarrow X$. Letting $\epsilon \downarrow 0$ and using the continuity from the right property of distribution functions, we have

$$P(X_n + Y_n \leq x) \leq P(X + c \leq x).$$

Following similar arguments, we can prove the lower bound

$$P(X_n + Y_n \leq x) \geq P(X + c \leq x).$$

This shows $X_n + Y_n \Rightarrow X + c$.

To prove $X_n Y_n \Rightarrow cX$ for the case $c > 0$, we use instead, for $0 < \epsilon < c$,

$$P(X_n Y_n \leq x) \leq P(|Y_n - c| > \epsilon) + P(X_n \leq \frac{x}{c - \epsilon})$$

to prove the upper bound. The lower bound and the other two cases $c < 0$ and $c = 0$ are proved similarly. $\square$

The next result gives a sufficient and necessary condition for convergence in distribution.

Theorem 3.2.3. $X_n \Rightarrow X$ if and only if $\forall\ g : \mathbb{R} \to \mathbb{R}$ bounded and continuous, we have $Eg(X_n) \to Eg(X)$.

Proof. "$\Longleftarrow$": Let F_n (F resp.) be the distribution function of X_n (X resp.). For any continuity point x of F and $\epsilon > 0$, define a continuous function

$$g(y) = \begin{cases} 1 & \text{if } y \leq x \\ 0 & \text{if } y \geq x + \epsilon \\ \text{linear} & \text{if } x \leq y \leq x + \epsilon. \end{cases}$$

We have

$$\begin{aligned} F_n(x) &= E1_{\{X_n \leq x\}} \leq Eg(X_n) \\ &\xrightarrow{\textit{condition}} Eg(X) \leq E1_{\{X \leq x+\epsilon\}} = F(x+\epsilon) \downarrow F(x), \text{ as } \epsilon \downarrow 0. \end{aligned}$$

Therefore, $\limsup_{n\to\infty} F_n(x) = F(x)$. Lower bound is proved similarly.

"$\Longrightarrow$": By Skorokhod's theorem, we can construct $Y_n, n \geq 1$ and Y on the same probability space such that Y_n has d.f. F_n for all n, Y has d.f. F and $Y_n \to Y$ a.s. We have

$$Eg(X_n) = Eg(Y_n) \xrightarrow{BCT} Eg(Y) = Eg(X).$$

$\square$

Remark. By modifying the above proof, we can change the equivalent statement to: $\forall\ g : \mathbb{R} \to \mathbb{R}$ bounded, continuous and having bounded and continuous derivatives of sufficiently large order, we have $Eg(X_n) \to Eg(X)$.

Now we are ready to prove the central limit theorem (CLT) assuming in addition finiteness of third moments. The proof is by **Lindeberg's swapping argument.**

Theorem (CLT assuming finite third moment). Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables such that $EX_i = \mu, Var(X_i) = \sigma^2$ and $E|X_i|^3 < \infty$. Let

$$W_n = \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma\sqrt{n}}.$$

Then

$$W_n \xrightarrow{d} Z \sim N(0,1).$$

Proof. Let $g : \mathbb{R} \to \mathbb{R}$ be any bounded continuous function with bounded and continuous derivatives up to the third order. By the above remark, it suffices to prove $Eg(W_n) \to Eg(Z)$. For each $i = 1, \dots, n$, let

$$\xi_i = \frac{X_i - \mu}{\sigma\sqrt{n}}.$$

Define $\eta_1, \dots, \eta_n$ on the same probability space such that $\{\xi_1, \cdots, \xi_n, \eta_1, \dots, \eta_n\}$ are independent and $\eta_i \sim N(0, \frac{1}{n})$ for all i. Note that

$$E\xi_i = E\eta_i = 0,\ E\xi_i^2 = E\eta_i^2 = \frac{1}{n},\ E|\xi_i|^3, E|\eta_i|^3 \le \frac{C}{n^{3/2}}. \tag{1.1}$$

Note also that $\sum_{i=1}^{n} \eta_i \sim N(0,1)$. By using the telescoping sums of Taylor's expansion, we have

$$\begin{aligned}
&Eg(W_n) - Eg(Z) \\
=&\sum_{k=1}^{n} \{Eg(\xi_1 + \cdots + \xi_{k-1} + \xi_k + \eta_{k+1} + \cdots + \eta_n) - Eg(\xi_1 + \cdots + \xi_{k-1} + \eta_k + \eta_{k+1} + \cdots + \eta_n)\} \\
=&E\sum_{k=1}^{n} \left\{ \left[g(V_k) + \xi_k g'(V_k) + \frac{\xi_k^2}{2} g''(V_k) + O(|\xi_k|^3) \right] - \left[g(V_k) + \eta_k g'(V_k) + \frac{\eta_k^2}{2} g''(V_k) + O(|\eta_k|^3) \right] \right\},
\end{aligned}$$

where $V_k := \xi_1 + \cdots + \xi_{k-1} + \eta_{k+1} + \cdots + \eta_n$. Using independence and (1.1) for cancellation, we obtain

$$|Eg(W_n) - Eg(Z)| \le Cn \frac{1}{n^{3/2}} \to 0.$$

$\square$

Now we return to the general discussion of weak convergence and state some related results.

Theorem 3.2.4 (Continuous Mapping Theorem). Let $g : \mathbb{R} \to \mathbb{R}$ be a measurable function and let $D_g := \{x : g \text{ is discontinuous at } x\}$. If $X_n \Rightarrow X$ and $P(X \in D_g) = 0$, then

$$g(X_n) \Rightarrow g(X).$$

Proof. By Skorokhod's theorem, we can construct $Y_n \to Y$ a.s. and Y_n (Y resp.) has the same distribution as X_n (X resp.) Because $P(Y \in D_g) = P(X \in D_g) = 0$, $g(Y_n) \to g(Y)$ a.s.. This implies $g(Y_n) \Rightarrow g(Y)$; hence $g(X_n) \Rightarrow g(X)$. □

Theorem 3.2.5 (Portmanteau Theorem). The following are equivalent:

(i) $X_n \Rightarrow X_\infty$.
(ii) $\forall$ open sets G, $\liminf_{n\to\infty} P(X_n \in G) \geq P(X_\infty \in G)$.
(iii) $\forall$ closed sets K, $\limsup_{n\to\infty} P(X_n \in K) \leq P(X_\infty \in K)$.
(iv) $\forall$ A with $P(X_\infty \in \partial A) = 0$, $\lim_{n\to\infty} P(X_n \in A) = P(X_\infty \in A)$.

Proof. "(i)$\Longrightarrow$ (ii)": By Skorokhod's theorem, we can construct $Y_n \to Y_\infty$ a.s. and Y_n (Y_∞ resp.) has the same distribution as X_n (X_∞ resp.) We have

$$\begin{aligned} &\liminf_{n\to\infty} P(X_n \in G) = \liminf_{n\to\infty} P(Y_n \in G) = \liminf_{n\to\infty} E1_{\{Y_n\in G\}} \\ \geq &E \liminf_{n\to\infty} 1_{\{Y_n \in G\}} && \text{(Fatou)} \\ \geq &E1_{\{Y_\infty \in G\}} && \text{(by } Y_n \to Y_\infty \text{ a.s. and } G \text{ is open)} \\ = &P(Y_\infty = G) = P(X_\infty \in G). \end{aligned}$$

"(ii)$\Longleftrightarrow$ (iii)": Consider $K = G^c$.
"(ii)+(iii)$\Longrightarrow$ (iv)": Use $A^o \subset A \subset \bar{A}$.
"(iv)$\Longrightarrow$ (i)": Consider $A = (-\infty, x]$. □

We state the following two theorem without proof.

Theorem 3.2.6 (Helly's Selection Theorem). Every sequence of distribution functions F_n has a subsequence $F_{n(k)}$ converging **vaguely** to a right continuous nondecreasing function F (not necessarily a distribution function), meaning

$$\lim_{k\to\infty} F_{n(k)}(y) = F(y), \ \forall \text{ continuity point } y \text{ of } F.$$

Theorem 3.2.7. Every subsequential limit above is a distribution function if and only if $\{F_n\}$ is **tight**, meaning that $\forall\ \epsilon > 0$, $\exists M_\epsilon$ such that $P(|X_n| \geq M_\epsilon) \leq \epsilon, \forall\ n$, where X_n has the distribution function F_n.

Example. A sequence of random variables X_n (or their distributions) is tight if $E|X_n| \leq C < \infty$ for all n.

Next, we define a few distances between distribution functions.

Definition. The **total variation distance** between two probability measures μ_1 and μ_2 is defined as

$$d_{TV}(\mu_1, \mu_2) := \sup_{A \in \mathcal{F}} |\mu_1(A) - \mu_2(A)|.$$

Theorem. Let f_n be a sequence of probability density functions (pdf), f be a pdf. Suppose $X_n \sim f_n$, $X \sim f$ (continuous random variables with the corresponding pdf). If $f_n(x) \to f(x)$ for all $x \in \mathbb{R}$, then

$$d_{TV}(\mathcal{L}(X_n), \mathcal{L}(X)) \to 0.$$

Proof. N.T.S. for any Borel set $B \in \mathcal{B}$,

$$|\int_B f_n(x)dx - \int_B f(x)dx| \to 0.$$

We have

$$\begin{aligned} &|\int_B f_n(x)dx - \int_B f(x)dx| \leq \int_{-\infty}^{\infty} |f_n(x) - f(x)|dx \\ =&2\int_{-\infty}^{\infty} (f_n(x) - f(x))^+ dx && \text{(uses } \textstyle\int_{-\infty}^{\infty} f_n(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1) \\ \to&0. && \text{(from DCT)} \end{aligned}$$

□

Definition. The **Kolmogorov distance** between two probability measures μ_1 and μ_2 (with d.f. F_1 and F_2 resp.) is defined as

$$d_K(\mu_1, \mu_2) := \sup_{x \in \mathbb{R}} |F_1(x) - F_2(x)|.$$

Besides convergence, it is of interest to provide explicit error bounds in the approximation for finite samples. The most famous result in this direction is the following Berry-Esseen bound. Its proof is beyond the scope of this course.

Theorem 3.4.9 (Berry-Esseen Theorem). Let $X_1, X_2, \ldots$ be i.i.d. with $EX_i = 0$, $EX_i^2 = \sigma^2$, $E|X_i|^3 = \gamma < \infty$. Let $W_n = \frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$. Let Φ be the standard normal distribution function. Then

$$\sup_{x \in \mathbb{R}} |P(W_n \leq x) - \Phi(x)| \leq \frac{\gamma}{\sigma^3\sqrt{n}}.$$

2 Characteristic Functions

In this section, we introduce the classical tool of proving distributional approximations via characteristic functions.

Definition. The **characteristic function (ch.f.)** of a random variable X is defined to be

$$\varphi_X(t) := Ee^{itX} = E\cos(tX) + i \cdot E\sin(tX).$$

Properties. 1. $\varphi_X(0) = 1$, $|\varphi_X(t)| \le 1$.

2. $\varphi_X(-t) = \overline{\varphi_X(t)}$. (conjugate)

3. $|\varphi_X(t+h) - \varphi_X(t)| \le E|e^{ihX} - 1| \to 0$, as $h \to 0$. (by DCT). That is, $\varphi_X(t)$ is uniformly continuous.

4. $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$.

5. If X_1 is independent of X_2, then $\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)$.

Theorem 3.3.8. If $E(X^2) < \infty$, then

$$\varphi_X(t) = 1 + i \cdot tE(X) - \frac{t^2}{2}E(X^2) + o(t^2), \text{ as } t \to 0.$$

Proof. By Taylor's expansion,

$$\varphi_X(t) = Ee^{itX} = 1 + EiXe^{itX}|_{t=0} \cdot t + E\frac{(iX)^2}{2}e^{itX}|_{t=0} \cdot t^2 + \text{error},$$

where

$$\boxed{\begin{aligned}|\text{error}| \le & CE\left[(t^3|X|^3) \wedge (t^2X^2)\right] \\ = & Ct^2E\left[(t|X|^3) \wedge (X^2)\right] \\ = & o(t^2), \text{ as } t \to 0\end{aligned}}$$

by DCT. □

Example. 1. $X \sim N(0,1)$: $\varphi_X(t) = e^{-\frac{t^2}{2}}$.

2. $X \sim N(\mu, \sigma^2)$: $\varphi_X(t) = e^{it\mu - \frac{\sigma^2 t^2}{2}}$.

3. $X \sim Poi(\lambda)$: $\varphi_X(t) = e^{\lambda(e^{it}-1)}$.

4. X has pdf $\frac{1-\cos x}{\pi x^2}$: $\varphi_X(t) = (1-|t|)^+$.

5. X has the α stable distribution, $0 < \alpha \le 2$: $\varphi_X(t) = e^{-|t|^\alpha}$.

The first main result in this part is the following **inversion formula**, which recovers the distribution function from the corresponding characteristic function.

Theorem 3.3.4 (The Inversion Formula). Let $\varphi_X(t) = E(e^{itX})$ be the ch.f. of a random variable X. Then for any $a < b$, we have

$$\lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)dt = P(a < X < b) + \frac{1}{2}P(X = a) + \frac{1}{2}P(X = b).$$

The existence of the limit is part of the statement of the theorem.

Proof. First note that

$$|\frac{e^{-ita} - e^{-itb}}{it}| = |\int_a^b e^{-ity}dy| \leq b - a.$$

By Fubini's theorem, for any T,

$$\begin{aligned}
&\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t)dt \\
=&E\frac{1}{2\pi} \int_{-T}^{T} \frac{e^{it(X-a)} - e^{it(X-b)}}{it} dt \\
=&E\frac{1}{2\pi} \int_{-T}^{T} \left[\frac{\cos(it(X-a)) - \cos(it(X-b))}{it} + \frac{\sin(t(X-a)) - \sin(t(X-b))}{t}\right] dt \\
=&E\frac{1}{\pi} \int_{0}^{T} \left[\frac{\sin(t(X-a)) - \sin(t(X-b))}{t}\right] dt,
\end{aligned} \tag{2.1}$$

where we used (anti-) symmetry in the last step.

Using Exercise 1.7.5, we have the following facts:

$$\boxed{\begin{gathered}\lim_{T\to\infty} \int_0^T \frac{\sin(tc)}{t} dt = \begin{cases} \frac{\pi}{2} & c > 0, \\ 0 & c = 0, \\ -\frac{\pi}{2} & c < 0, \end{cases} \\ \text{and} \quad |\int_0^T \frac{\sin(tc)}{t} dt| \leq 4.\end{gathered}}$$

Applying these facts to (2.1) and from a case-by-case discussion, we obtain the desired result. □

The next two results are special cases (which require some additional effort, though) of the previous theorem.

Exercise 3.3.2. If X is integer-valued, then

$$P(X = x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itx} \varphi_X(t)dt, \ \forall\ x \in \mathbb{Z}.$$

Theorem 3.3.5. If $\int_{-\infty}^{\infty} |\varphi_X(t)|dt < \infty$, then X is a continuous random variable with bounded and continuous density

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_X(t) dt.$$

The second main result in this part is the following criterion of weak convergence using characteristic functions.

Theorem 3.3.6. (i) If $X_n \xrightarrow{d} X$, then $\varphi_{X_n}(t) \to \varphi_X(t)$, $\forall\, t \in \mathbb{R}$.

(ii) If $\varphi_{X_n}(t) \to \varphi(t)$, $\forall\, t \in \mathbb{R}$ and φ is continuous at 0, then φ is a ch.f. of some random variable X and $X_n \xrightarrow{d} X$.

Proof sketch. (i) follows from the Skorokhod's theorem and BCT.

For (ii), the continuity of φ at 0 implies tightness of the sequence X_n. Therefore, every subsequential vague limit is a distribution function. By the condition $\varphi_{X_n}(t) \to \varphi(t)$, every subsequential weak limit is the same distribution function having the ch.f. $\varphi(t)$. This implies the whole sequence converges weakly to the limit. $\square$

3 Central Limit Theorem

Using characteristic functions, we now give the second proof of the CLT.

Theorem. Let $X_1, X_2, \dots$ be a sequence of i.i.d. random variables such that $EX_i = \mu, Var(X_i) = \sigma^2$. Let

$$W_n = \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma\sqrt{n}}.$$

Then

$$W_n \xrightarrow{d} Z \sim N(0,1).$$

Proof. N.T.S. $Ee^{itW_n} \to e^{-\frac{t^2}{2}}$ for all $t \in \mathbb{R}$. We have, by the expression of W_n and independence,

$$\begin{aligned}
Ee^{itW_n} &= E\exp(it(\frac{X_1-\mu}{\sigma\sqrt{n}} + \cdots + \frac{X_n-\mu}{\sigma\sqrt{n}})) = \prod_{j=1}^{n} Ee^{i(\frac{t}{\sqrt{n}})(X_j-\mu)} \\
&= \prod_{j=1}^{n}\left[1 + i\frac{t}{\sigma\sqrt{n}}E(X_j-\mu) - \frac{t^2}{2\sigma^2 n}E(X_j-\mu)^2 + o(\frac{t^2}{n})\right] \qquad \text{(from Theorem 3.3.8)} \\
&= \left[1 - \frac{t^2}{2n} + o(\frac{t^2}{n})\right]^n \\
&\to e^{-t^2/2}.
\end{aligned}$$

□

Next, we give a version of CLT for sums of independent but not necessarily identically distributed random variables.

Theorem 3.4.5 (The Lindeberg-Feller Theorem). Assume for each n, $\xi_{n1}, \xi_{n2}, \dots, \xi_{nn}$ are independent with $E\xi_{ni} = 0$ for all i and $E\sum_{i=1}^{n}\xi_{ni}^2 = 1$. If

$$\forall\ \epsilon > 0, \quad \sum_{i=1}^{n} E\xi_{ni}^2 1_{\{|\xi_{ni}|>\epsilon\}} \to 0, \qquad \text{(Lindeberg's Condition)}$$

then

$$\sum_{i=1}^{n} \xi_{ni} \xrightarrow{d} N(0,1).$$

Remarks. (1) CLT for i.i.d. sequence is a corollary of the above theorem: For $X_1, X_2, \dots,$ i.i.d. with $EX_i = \mu, \mathrm{Var}(X_i) = \sigma^2$. Consider $\xi_{ni} := \frac{X_i - \mu}{\sigma\sqrt{n}}$ and $W_n := \sum_{i=1}^{n}\xi_{ni}$. It can be checked by DCT that the Lindeberg condition is satisfied and hence CLT.

(2) A sufficient condition for Lindeberg's condition is $\sum_{i=1}^n E|\xi_{ni}|^p \to 0$ for some $p > 2$. (proof by inserting a factor of $\frac{|\xi_{ni}|^{p-2}}{\epsilon^{p-2}}$ inside of the expectation in Lindeberg's condition).

(3) Lindeberg's condition implies

$$\max_{1\le i\le n} E\xi_{ni}^2 \to 0, \quad \text{as } n \to \infty.$$

(proof using $\sum_{i=1}^n E\xi_{ni}^2 1_{\{|\xi_{ni}|>\epsilon\}} \ge \max_{1\le i\le n} E\xi_{ni}^2 1_{\{|\xi_{ni}|>\epsilon\}} \ge \max_{1\le i\le n} E\xi_{ni}^2 - \epsilon^2.$)

(4) $\max_{1\le i\le n} E\xi_{ni}^2 \to 0$ is a necessary condition for CLT. (Otherwise, there is a non-negaligible component in the summation which can be taken to be an arbitrary distribution, and the sum may not be Gaussian.)

Proof of the Lindeberg-Feller theorem. Let

$$\varphi_n(t) = Ee^{it\sum_{i=1}^n \xi_{ni}}.$$

From the proof of Theorem 3.3.8, we have

$$\begin{aligned}
\varphi_n(t) &= \prod_{i=1}^n Ee^{it\xi_{ni}} \\
&= \prod_{i=1}^n E\left[1 + it\xi_{ni} - \frac{t^2}{2}\xi_{ni}^2 + O(t^2\xi_{ni}^2 1_{\{|\xi_{ni}|>\epsilon\}}) + O(t^3|\xi_{ni}|^3 1_{\{|\xi_{ni}|\le\epsilon\}})\right] \\
&= \prod_{i=1}^n \left[1 - \frac{t^2}{2}E\xi_{ni}^2 + O(t^2E\xi_{ni}^2 1_{\{|\xi_{ni}|>\epsilon\}}) + O(t^3\epsilon E|\xi_{ni}|^2)\right] \\
&\to \prod_{i=1}^n e^{-\frac{t^2}{2}E\xi_{ni}^2} = e^{-t^2/2},
\end{aligned}$$

where we used Lemma 3.4.3: Let $z_1, \dots, z_n$ and $w_1, \dots, w_n$ be complex numbers with $|z_i| \le 1, |w_i| \le 1$ for all i. Then

$$\left|\prod_{i=1}^n z_i - \prod_{i=1}^n w_i\right| \le \sum_{i=1}^n |z_i - w_i|.$$

□

4 Related Topics

Depending on time availability, we will discuss (without proofs):

- Improvements over CLT: Berry-Esseen theorem; Self-normalized CLT; Discrete normal approximation; Local limit theorem; Cramér's moderate deviation; Edgeworth expansion; ...

- Other limit theorems: Poisson limit theorem; Poisson process; Limit theorem on $\mathbb{R}^d$; ...

- Open problems: Random assignment problem; large-dimensional CLT; ...