4 The Central Limit Theorem

4.1 CLT for the impatient

In order to get an intuition and to motivate our more detailed study in the sequel, we will sketch the proof of a preliminary version of the CLT. [1] Obviously,

$$e^{\imath x} = 1 + \imath x - x^2/2 + O(x^3) \;, \qquad x \in \mathbb{R} \;;$$

hence, if $(X_n)_{n\in\mathbb{N}}$ is an i.i.d. sequence with $\mathrm{E}\,X_1 = 0$, $\mathrm{Var}(X_1) = 1$ and $\mathrm{E}\,|X_1|^3 < \infty$, we have

$$\varphi_{X_1}(\gamma) = \mathrm{E}\,e^{\imath\gamma X_1} = 1 - \gamma^2/2 + \mathrm{E}\,|X_1|^3 \cdot O(\gamma^3) \;.$$

Consequently,

$$\begin{aligned} \varphi_{\frac{1}{\sqrt{n}}\sum_{i\le n} X_i}(\lambda) &= (\varphi_{X_1}(\lambda/\sqrt{n}))^n \\ &= \Big[1 - \frac{\lambda^2}{2n} + \mathrm{E}\,|X_1|^3 O\Big(\frac{\lambda^3}{n^{3/2}}\Big)\Big]^n \\ &\overset{n\to\infty}{\longrightarrow} e^{-\lambda^2/2} = \widehat{\mathcal{N}(0,1)}(\lambda) \;. \end{aligned}$$

By Lévy's continuity theorem, it follows that

$$\frac{1}{\sqrt{n}} \sum_{i\le n} X_i \overset{d}{\longrightarrow} \mathcal{N}(0,1) \;.$$

Although this is already a nice result, a couple of questions deserve further investigation:

1. It seems odd to study one fixed sequence (X_n) if what matters is just the distribution of $\sum_{i\le n} X_i$ as $n \to \infty$; this should be formalized and studied.
2. While independence seems a natural enough assumption, one surely cannot guarantee identical distribution in most applications of interest. (Even in highly idealized situations like molecular bombardement of a body, where X_n is the impulse induced by the nth collision, there are molecules of different size.) Hence, it is of interest to ask for generalizations to the non–i.i.d. case.
3. Is the assumption on $\mathrm{E}\,|X_1|^3$ optimal, or can it be improved?

4.2 CLT in depth

Let now a triangular array of random variables $X_{n,k}$ be given, where $n \in \mathbb{N}$ and $k \in \{1, \dots, r_n\}$ with $r_n \in \mathbb{N}$. We assume throughout this section:

(i) $\mathrm{E}\,X_{nk} = 0$ and $\sigma_{nk} = \mathrm{Var}(X_{n,k}) < \infty$ for every $n \in \mathbb{N}$ and $k \in \{1, \dots, r_n\}$,

[1] The following reasoning is easily made rigorous, but we will have rigorous generalizations anyway, hence we sacrifice rigorousity for ease of access.

(ii) $(X_{n1}, \dots, X_{nr_n})$ independent for every $n \in \mathbb{N}$, and $s_n^2 = \sum_{k=1}^{r_n} \sigma_{n,k}^2 = 1$.

Put

$$S_n^* = \sum_{k=1}^{r_n} X_{nk} \; ;$$

then $\mathrm{E}(S_n^*) = 0$ and $\mathrm{Var}(S_n^*) = 1$. Question: convergence in distribution of $(S_n^*)_{n\in\mathbb{N}}$? For notational convenience: all random variables X_{nk} are defined on a common probability space $(\Omega, \mathfrak{A}, P)$.

Example 1. $(X_n)_{n\in\mathbb{N}}$ i.i.d. with $X_1 \in \mathfrak{L}^2$ and $\mathrm{Var}(X_1) = \sigma^2 > 0$. Put $m = \mathrm{E}(X_1)$, take

$$r_n = n, \qquad X_{nk} = (X_k - m)/\sqrt{n\sigma^2}.$$

Then

$$S_n^* = \frac{\sum_{k=1}^n X_k - n \cdot m}{\sqrt{n} \cdot \sigma}.$$

Definition 1.

(i) The *Lyapunov condition* holds iff

$$\exists\, \delta > 0 : \lim_{n\to\infty} \sum_{k=1}^{r_n} \mathrm{E}(|X_{nk}|^{2+\delta}) = 0.$$

(ii) The *Lindeberg condition* holds iff

$$\forall\, \varepsilon > 0 : \lim_{n\to\infty} \sum_{k=1}^{r_n} \int_{\{|X_{nk}|\geq\varepsilon\}} X_{nk}^2 \, dP = 0.$$

Lemma 1. The Lyapunov condition entails the Lindeberg condition; further, from the Lindeberg condition it follows that $r_n \to \infty$ and that $\max_{k\leq r_n} \sigma_{n,k}^2 \to 0$.

Proof. If we have the Lyapunov condition, then

$$\int_{\{|X_{nk}|\geq\varepsilon\}} X_{nk}^2 \, dP \leq \frac{1}{\varepsilon^\delta} \cdot \int_{\{|X_{nk}|\geq\varepsilon\}} |X_{nk}|^{2+\delta} \, dP \leq \frac{1}{\varepsilon^\delta} \cdot \mathrm{E}(|X_{nk}|^{2+\delta}) \; .$$

This entails the Lindeberg condition. Further, if the Lindeberg condition holds, for every $\varepsilon > 0$ there is n_0 such that for $n \geq n_0$ we have

$$1 = \sum_{k\leq r_n} \mathrm{E}\, X_{nk}^2 \leq 1/2 + \sum_{k\leq r_n} \int_{|X_{nk}|\leq\varepsilon} X_{nk}^2 \leq 1/2 + \varepsilon^2 r_n \; ;$$

this shows $r_n \to \infty$. Finally, for any $\varepsilon > 0$,

$$\overline{\lim_{n\to\infty}} \max_{k\leq r_n} \sigma_{nk}^2 \leq \varepsilon + \overline{\lim_{n\to\infty}} \sum_{k\leq r_n} \int_{|X_{n,k}|\geq\varepsilon} X_{n,k}^2 = \varepsilon \; .$$

□

Example 2. In Example 1,

$$\sum_{k=1}^{n} \int_{\{|X_{nk}| \geq \varepsilon\}} X_{nk}^2 \, dP = \frac{1}{\sigma^2} \cdot \int_{\{|X_1 - m| \geq \varepsilon \cdot \sqrt{n} \cdot \sigma\}} (X_1 - m)^2 \, dP.$$

Hence the Lindeberg condition is satisfied.

In the sequel

$$\varphi_{nk} = \varphi_{X_{nk}}$$

denotes the characteristic function of X_{nk}.

Lemma 2. For $y \in \mathbb{R}$ and $\varepsilon > 0$

$$\left|\varphi_{nk}(y) - (1 - \sigma_{nk}^2/2 \cdot y^2)\right| \leq y^2 \cdot \Big(\varepsilon \cdot |y| \cdot \sigma_{nk}^2 + \int_{\{|X_{nk}| \geq \varepsilon\}} X_{nk}^2 \, dP\Big).$$

Proof. For $u \in \mathbb{R}$

$$\left|\exp(\imath u) - (1 + \imath u - u^2/2)\right| \leq \min(u^2,\ |u|^3/6),$$

see Billingsley (1979, Eqn. (26.4)). Hence

$$\begin{aligned}
&\left|\varphi_{nk}(y) - (1 - \sigma_{nk}^2/2 \cdot y^2)\right| \\
&\quad = \left|\mathrm{E}(\exp(\imath \cdot X_{nk} \cdot y)) - \mathrm{E}\Big(1 + \imath \cdot X_{nk} \cdot y - X_{nk}^2 \cdot y^2/2)\right| \\
&\quad \leq \mathrm{E}\big(\min(y^2 \cdot X_{nk}^2\,,\ |y|^3 \cdot |X_{nk}|^3)\big) \\
&\quad \leq |y|^3 \cdot \int_{\{|X_{nk}| < \varepsilon\}} \varepsilon \cdot X_{nk}^2 \, dP + y^2 \cdot \int_{\{|X_{nk}| \geq \varepsilon\}} X_{nk}^2 \, dP \\
&\quad \leq \varepsilon \cdot |y|^3 \cdot \sigma_{nk}^2 + y^2 \cdot \int_{\{|X_{nk}| \geq \varepsilon\}} X_{nk}^2 \, dP.
\end{aligned}$$

□

Lemma 3. Put

$$\Delta_n(y) = \prod_{k=1}^{r_n} \varphi_{nk}(y) - \exp(-y^2/2), \qquad y \in \mathbb{R}.$$

If the Lindeberg condition is satisfied, then

$$\forall\, y \in \mathbb{R} : \lim_{n \to \infty} \Delta_n(y) = 0.$$

Proof. From the triangle inequality one has for any complex x_i, y_i,

$$\Big|\prod_{i \leq m} x_i - \prod_{i \leq m} y_i\Big| \leq \Big|\prod_{i \leq m-1} x_i\Big| \cdot |y_m - x_m| + |y_m| \cdot \Big|\prod_{i \leq m-1} x_i - \prod_{i \leq m-1} y_i\Big|\,;$$

hence, if $|x_i|, |y_i| \leq 1$, we get by induction

$$\Big|\prod_{i \leq m} x_i - \prod_{i \leq m} y_i\Big| \leq \sum_{i \leq m} |x_i - y_i|\,.$$

Since $|\varphi_{nk}(y)| \leq 1$ and $|\exp(-y^2\sigma_{nk}^2/2)| \leq 1$, this applies, and

$$|\Delta_n(y)| = \Big|\prod_{k=1}^{r_n} \varphi_{nk}(y) - \prod_{k=1}^{r_n} \exp(-y^2\sigma_{nk}^2/2)\Big| \leq \sum_{k=1}^{r_n} \big|\varphi_{nk}(y) - \exp(-y^2\sigma_{nk}^2/2)\big| \; .$$

We assume

$$\max_{1 \leq k \leq r_n} \sigma_{nk}^2 \cdot y^2 \leq 1 \, ,$$

which holds for fixed $y \in \mathbb{R}$ if n is sufficiently large, see Lemma 1. Using

$$0 \leq u \leq 1/2 \quad \Rightarrow \quad |\exp(-u) - (1-u)| \leq u^2$$

and Lemma 2 we obtain

$$\begin{aligned}
|\Delta_n(y)| &\leq \sum_{k=1}^{r_n} |\varphi_{nk}(y) - (1 - y^2\sigma_{nk}^2/2)| + \sum_{k=1}^{r_n} y^4\sigma_{nk}^4/4 \\
&\leq y^2 \cdot \Big(\varepsilon \cdot |y| + \sum_{k=1}^{r_n} \int_{\{|X_{nk}| \geq \varepsilon\}} X_{nk}^2 \, dP\Big) + y^4/4 \cdot \max_{1 \leq k \leq r_n} \sigma_{nk}^2
\end{aligned}$$

for every $\varepsilon > 0$. Thus Lemma 1 yields

$$\limsup_{n \to \infty} |\Delta_n(y)| \leq |y|^3 \cdot \varepsilon.$$

□

Theorem 1 (Central Limit Theorem). If $(X_{nk})_{n,k}$ satisfies the Lindeberg condition, then $P_{S_n^*} \xrightarrow{\text{w}} N(0,1)$.

Proof. Recall that $\widehat{\mu}(y) = \exp(-y^2/2)$ for the standard normal distribution μ. Consider the characteristic function $\varphi_n = \varphi_{S_n^*}$ of S_n^*. By Theorem 3.2.(ii)

$$\varphi_n = \prod_{k=1}^{r_n} \varphi_{nk},$$

and therefore Lemma 3 implies

$$\forall\, y \in \mathbb{R} : \lim_{n \to \infty} \varphi_n(y) = \widehat{\mu}(y).$$

It remains to apply Corollary 3.2. □

Corollary 1. Let $(X_n)_{n \in \mathbb{N}}$ be i.i.d. with $X_1 \in \mathfrak{L}^2$ and $\sigma^2 = \operatorname{Var}(X_1) > 0$. Then

$$\frac{\sum_{k=1}^n X_k - n \cdot \operatorname{E}(X_1)}{\sqrt{n} \cdot \sigma} \xrightarrow{\text{d}} Z$$

where $Z \sim N(0,1)$.

Proof. Theorem 1 and Example 2. □

Example 3. Example 2 continued, and Corollary 1 reformulated. Let

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{x} \exp(-u^2/2)\, du, \qquad x \in \mathbb{R},$$

denote the distribution function of the standard normal distribution. Due to the Central Limit Theorem and Theorem III.3.2

$$\sup_{x\in\mathbb{R}} \left| P(\{S_n \le x \cdot \sqrt{n} \cdot \sigma\}) - \Phi(x) \right| \; = \; \sup_{x\in\mathbb{R}} \left| P(\{S_n \le x\}) - \Phi(x/(\sqrt{n} \cdot \sigma)) \right| \; \to \; 0 \; .$$

The speed of this convergence can be further quantified (Berry–Essén Theorem).

Let now

$$B_c = \{\overline{\lim_{n\to\infty}}\, S_n/\sqrt{n} \ge c\} \supset \overline{\lim_{n\to\infty}}\, \{S_n/\sqrt{n} > c\}, \qquad c > 0.$$

Using Remark 1.2.(ii) we get

$$P(B_c) \ge P(\overline{\lim_{n\to\infty}}\, \{S_n/\sqrt{n} > c\}) \ge \overline{\lim_{n\to\infty}}\, P(\{S_n/\sqrt{n} > c\}) = 1 - \Phi(c/\sigma) > 0.$$

Kolmogorov's Zero-One Law yields

$$P(B_c) = 1,$$

and therefore

$$P(\{\overline{\lim_{n\to\infty}}\, S_n/\sqrt{n} = \infty\}) = P\left(\bigcap_{c\in\mathbb{N}} B_c\right) = 1.$$

By symmetry

$$P(\{\underline{\lim_{n\to\infty}}\, S_n/\sqrt{n} = -\infty\}) = 1.$$

In particular, for $P_{X_1} = 1/2(\delta_1 + \delta_{-1})$, we have

$$P(\overline{\lim_{n\to\infty}}\, \{S_n = 0\}) = 1 \; ;$$

this is the simplest *recurrence* result: Almost surely, the random walk S_n returns to 0 infinitely often.