Series 12: central limit theorem **Solutions**

Exercise 1
Let μ be the law of X. Since X and Y have the same characteristic function, μ is also the law of Y. Since X and Y are assumed independent, their joint law is the product measure $\mu \otimes \mu$. In particular, we have

$$\mathbb{P}[X = Y] = \int \mathbb{1}_{\{x=y\}}\, d\mu(x)\, d\mu(y). \tag{1}$$

We now compute

$$\begin{aligned}
\frac{1}{2T}\int_{-T}^{T} |\varphi(t)|^2\, dt &= \frac{1}{2T}\int_{-T}^{T} \varphi(t)\overline{\varphi(t)}\, dt \\
&= \frac{1}{2T}\int_{-T}^{T} \int e^{itx}\, d\mu(x) \int e^{-ity}\, d\mu(y)\, dt \\
&= \int \frac{1}{2T}\int_{-T}^{T} e^{it(x-y)}\, dt\, d\mu(x)\, d\mu(y),
\end{aligned}$$

by Fubini's theorem. Now,

$$\frac{1}{2T}\int_{-T}^{T} e^{itz}\, dt = \operatorname{sinc}(Tz),$$

where

$$\operatorname{sinc}(z) = \left| \begin{array}{ll} \sin(z)/z & \text{if } z \neq 0, \\ 1 & \text{if } z = 0. \end{array} \right.$$

Hence, we get

$$\frac{1}{2T}\int_{-T}^{T} |\varphi(t)|^2\, dt = \int \operatorname{sinc}(T(x-y))\, d\mu(x)\, d\mu(y).$$

We check that sinc is a bounded function, and that $\operatorname{sinc}(Tz) \to 0$ as $T \to \infty$ if $z \neq 0$, and otherwise $\operatorname{sinc}(Tz) = 1$. By the dominated convergence theorem, we thus get

$$\frac{1}{2T}\int_{-T}^{T} |\varphi(t)|^2\, dt \xrightarrow[T\to+\infty]{} \int \mathbb{1}_{\{x=y\}}\, d\mu(x)\, d\mu(y),$$

which, when combined with (1), gives the result.

Exercise 2
Substracting the expectation out of each X_i, one can see that it suffices to prove the result for mean zero random variables. We thus assume that $\mathbb{E}X_i = 0\ \forall i$. Let

$$Y_{m,n} = \frac{X_m}{\alpha_n}, \quad n \geq 1, 1 \leq m \leq n.$$

Since $S_n/\alpha_n = Y_{1,n} + \ldots + Y_{n,n}$, it suffices to check that the $Y_{m,n}$ satisfy the Lindeberg-Feller conditions. Indeed,

- $\mathbb{E}(Y_{m,n}) = 0\ \forall m, n.$
- For any n,
$$\sum_{m=1}^{n} \mathbb{E}(Y_{m,n}^2) = \frac{\operatorname{Var}(X_1 + \ldots + X_n)}{\alpha_n^2} = 1.$$

- For $\epsilon > 0$ and $n \geq 1$,

$$\sum_{m=1}^{n} \mathbb{E}(Y_{m,n}^2 \cdot I_{(Y_{m,n}>\epsilon)}) \leq \sum_{m=1}^{n} \mathbb{E}\left(Y_{m,n}^2 \cdot \left(\frac{Y_{m,n}}{\epsilon}\right)^{\delta}\right) =$$

$$\frac{1}{\epsilon^{\delta}} \sum_{m=1}^{n} \mathbb{E}(Y_{m,n}^{2+\delta}) = \frac{1}{\epsilon^{\delta}\alpha_n^{2+\delta}} \sum_{m=1}^{n} \mathbb{E}(X_m^{2+\delta}),$$

so $\sum_{m=1}^{n} \mathbb{E}(Y_{m,n}^2 \cdot I_{(Y_{m,n}>\epsilon)}) \to 0$ when $n \to \infty$.

Exercise 3
Following the hint, we write

$$\begin{aligned} \mathbb{E}[S_n^2] &\geqslant \sum_{k=1}^{n} \mathbb{E}[(S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2)\, \mathbb{1}_{A_k}] \\ &\geqslant \sum_{k=1}^{n} \mathbb{E}[S_k^2\, \mathbb{1}_{A_k}] + 2\mathbb{E}[S_k(S_n - S_k)\, \mathbb{1}_{A_k}]. \end{aligned}$$

Concerning the last expectation, note that $S_k \mathbb{1}_{A_k}$ is $\sigma(X_1, \ldots, X_k)$-measurable, while $(S_n - S_k)$ is $\sigma(S_{k+1}, \ldots, S_n)$-measurable. These random variables are thus independent, and since $S_n - S_k$ is centred, we get

$$\mathbb{E}[S_k(S_n - S_k)\, \mathbb{1}_{A_k}] = 0.$$

On the other hand, on the event A_k, we have $S_k^2 \geqslant x^2$, and the events (A_k) are disjoint, thus

$$\mathbb{E}[S_n^2] \geqslant \sum_{k=1}^{n} \mathbb{E}[S_k^2\, \mathbb{1}_{A_k}] \geqslant x^2 \sum_{k=1}^{n} \mathbb{E}[\mathbb{1}_{A_k}] \geqslant x^2 \mathbb{P}\left[\max_{1\leqslant k\leqslant n} |S_k| \geqslant x\right].$$

Exercise 4
Let

$$Y_n = \frac{S_{N_n}}{\sigma\sqrt{a_n}}, \quad Z_n = \frac{S_{a_n}}{\sigma\sqrt{a_n}}.$$

By the central limit theorem, $Z_n \xrightarrow{(L)} N(0,1)$. We will show that $Y_n - Z_n \to 0$ in probability; we will then be able to conclude that

$$Y_n = Z_n + (Y_n - Z_n) \xrightarrow{(L)} N(0,1).$$

Let $\epsilon > 0$ and $\delta < \epsilon^3/8$. We have

$$\mathbb{P}(|Y_n - Z_n| > \epsilon) = \mathbb{P}\left(\left|\sum_{i=1}^{N_n} X_i - \sum_{i=1}^{a_n} X_i\right| > \epsilon\sigma\sqrt{a_n}\right) \leq$$

$$\mathbb{P}\left(\left|\frac{N_n}{a_n} - 1\right| > \delta\right) + \mathbb{P}\left(N_n \in [a_n(1-\delta), a_n], \left|\sum_{i=N_n+1}^{a_n} X_i\right| > \epsilon\sigma\sqrt{a_n}\right) + \mathbb{P}\left(N_n \in [a_n, a_n(1+\delta)], \left|\sum_{i=a_n+1}^{N_n} X_i\right| > \epsilon\sigma\sqrt{a_n}\right)$$

$$\leq \mathbb{P}\left(\left|\frac{N_n}{a_n} - 1\right| > \delta\right) + \mathbb{P}\left(\max_{a_n(1-\delta)\leq j\leq a_n} \left|\sum_{i=j}^{a_n} X_i\right| > \epsilon\sigma\sqrt{a_n}\right) + \mathbb{P}\left(\max_{a_n\leq j\leq a_n(1+\delta)} \left|\sum_{i=a_n}^{j} X_i\right| > \epsilon\sigma\sqrt{a_n}\right) =$$

$$\leq \mathbb{P}\left(\left|\frac{N_n}{a_n} - 1\right| > \delta\right) + 2\mathbb{P}\left(\max_{a_n\leq j\leq a_n(1+\delta)} \left|\sum_{i=a_n}^{j} X_i\right| > \epsilon\sigma\sqrt{a_n}\right),$$

since the X_i are i.i.d. We now use Kolmogorov's inequality:

$$\mathbb{P}\left(\max_{a_n \le j \le a_n(1+\delta)} \left|\sum_{i=a_n}^{j} X_i\right| > \epsilon\sigma\sqrt{a_n}\right) \le \frac{\mathrm{Var}(X_{a_n}) + \ldots + \mathrm{Var}(X_{\lceil a_n(1+\delta)\rceil})}{\epsilon^2\sigma^2 a_n} \le \frac{\lceil \delta a_n \rceil \sigma^2}{\epsilon^2\sigma^2 a_n} \le 2\frac{\delta}{\epsilon^2} < \frac{\epsilon}{4}.$$

Since $N_n/a_n \to 1$ in probability, we can choose n_0 such that $n \ge n_0$ implies $\mathbb{P}\left(\left|\frac{N_n}{a_n} - 1\right| > \delta\right) < \epsilon/2$. For $n \ge n_0$, we thus have $\mathbb{P}(|Y_n - Z_n| > \epsilon) < \frac{\epsilon}{2} + 2\frac{\epsilon}{4} = \epsilon$.