0.6.4 The proof of Central Limit Theorem

Informal Proof without introducing Y ((the replacement Gaussian random variables))

$\mathbb{E}[X_j] = 0,$

$\mathrm{Var}(X_j) = \sigma_j^2$ (finite variances, $\sigma_j^2 < \infty$),

$\mathbb{E}[|X_j|^3] = \gamma_j$ (finite third moments, $\gamma_j < \infty$).

Define:

$S_n = \sum_{j=1}^n X_j,$

$s_n^2 = \sum_{j=1}^n \sigma_j^2$ (total variance),

$\Gamma_n = \sum_{j=1}^n \gamma_j$ (total third moment).

Statement:

If

$$\frac{\Gamma_n}{s_n^3} \to 0 \quad \text{as } n \to \infty.$$

then

$$\frac{S_n}{s_n} \xrightarrow{d} \Phi$$

where Φ denotes the standard normal distribution $\mathcal{N}(0,1)$.

Let $\boldsymbol{Z_n = S_n/s_n}$, to prove $\boldsymbol{Z_n \xrightarrow{d} \mathcal{N}(0,1)}$, it suffices to show that for any bounded, continuous test function $f \in C_b^3(\mathbb{R})$ (bounded continuous functions with three bounded derivatives):

$$\mathbb{E}[f(Z_n)] \to \mathbb{E}[f(\mathcal{N})],$$

Using Taylor's theorem, we expand $f(Z_n)$ around 0:

$$f(Z_n) = f(0) + f'(0)Z_n + \frac{f''(0)}{2}Z_n^2 + R(Z_n),$$

where $R(Z_n)$ is the remainder:

$$R(Z_n) = \frac{f^{(3)}(\xi)}{6}Z_n^3,$$

for some ξ between 0 and Z_n.

Taking expectations:

$$\mathbb{E}[f(Z_n)] = f(0) + f'(0)\mathbb{E}[Z_n] + \frac{f''(0)}{2}\mathbb{E}[Z_n^2] + \mathbb{E}[R(Z_n)].$$

Similarly, for $\mathcal{N}(0,1)$, we expand $f(\mathcal{N})$:

$$\mathbb{E}[f(\mathcal{N})] = f(0) + f'(0)\mathbb{E}[\mathcal{N}] + \frac{f''(0)}{2}\mathbb{E}[\mathcal{N}^2] + \mathbb{E}[R(\mathcal{N})].$$

Thus, the difference becomes:

$$\mathbb{E}[f(Z_n)] - \mathbb{E}[f(\mathcal{N})] = \underbrace{f'(0)(\mathbb{E}[Z_n] - \mathbb{E}[\mathcal{N}])}_{(a)} + \underbrace{\frac{f''(0)}{2}(\mathbb{E}[Z_n^2] - \mathbb{E}[\mathcal{N}^2])}_{(b)} + \underbrace{(\mathbb{E}[R(Z_n)] - \mathbb{E}[R(\mathcal{N})])}_{(c)}.$$

By analyzing each term, we have:

(a) $f'(0)(\mathbb{E}[Z_n] - \mathbb{E}[\mathcal{N}])$

- By assumption, $\mathbb{E}[X_j] = 0$, so $\mathbb{E}[S_n] = 0$, and:

$$\mathbb{E}[Z_n] = \frac{\mathbb{E}[S_n]}{s_n} = 0.$$

- For $\mathcal{N}(0,1)$, $\mathbb{E}[\mathcal{N}] = 0$.

Therefore:

$$f'(0)(\mathbb{E}[Z_n] - \mathbb{E}[\mathcal{N}]) = 0.$$

(b) $\boldsymbol{\frac{f''(0)}{2}(\mathbb{E}[Z_n^2] - \mathbb{E}[\mathcal{N}^2])}$

- The variance of S_n is $s_n^2 = \sum_{j=1}^n \sigma_j^2$, so:

$$\mathrm{Var}(Z_n) = \frac{\mathrm{Var}(S_n)}{s_n^2} = \frac{s_n^2}{s_n^2} = 1.$$

- Thus:

$$\mathbb{E}[Z_n^2] = 1.$$

- For $\mathcal{N}(0,1)$, $\mathbb{E}[\mathcal{N}^2] = 1$.

Therefore:

$$\frac{f''(0)}{2}(\mathbb{E}[Z_n^2] - \mathbb{E}[\mathcal{N}^2]) = 0.$$

(c) $\mathbb{E}[\boldsymbol{R(Z_n)}] - \mathbb{E}[\boldsymbol{R(\mathcal{N})}]$

By Taylor's theorem, We have:

$$\mathbb{E}[R(Z_n)] \leq \frac{M}{6}\mathbb{E}[Z_n^3],$$

and

$$\mathbb{E}[R(\mathcal{N})] \leq \frac{M}{6}\mathbb{E}[\mathcal{N}^3]$$

where $M = \sup_{x\in\mathbb{R}} |f^{(3)}(x)|$ is the maximum bound on the third derivative of f.

Thus,

$$|\mathbb{E}[R(Z_n)] - \mathbb{E}[R(\mathcal{N})]| \leq \frac{M}{6}\left|\mathbb{E}[Z_n^3] - \mathbb{E}[\mathcal{N}^3]\right| \tag{2}$$

For $\boldsymbol{Z_n}$:

The normalized third moment of $Z_n = S_n/s_n$ is:

$$\mathbb{E}[Z_n^3] = \frac{\mathbb{E}[|S_n|^3]}{s_n^3}. \tag{3}$$

Using Triangle inequality for sums, which states:

$$\mathbb{E}[|S_n|^3] = \mathbb{E}\left[\left|\sum_{j=1}^{n} X_j\right|^3\right] \leq \sum_{j=1}^{n}\mathbb{E}[|X_j|^3], \tag{4}$$

Combining (3) and (4) we get:

$$\mathbb{E}[|Z_n|^3] \leq \frac{\sum_{j=1}^n \gamma_j}{s_n^3},$$

where $\gamma_j = \mathbb{E}[|X_j|^3]$.

Thus:

$$\mathbb{E}[R(Z_n)] \leq \frac{M}{6}\frac{\sum_{j=1}^n \gamma_j}{s_n^3}.$$

For $\boldsymbol{\mathcal{N}(0,1)}$: The third moment of $\mathcal{N}$ is constant:

$$\mathbb{E}[\mathcal{N}^3] = 0, \tag{5}$$

To prove (5), we use moment generating function (MGF). The moment generating function (MGF) of a random variable X is defined as:

$$M_X(t) = \mathbb{E}[e^{tX}].$$

For $X \sim N(0,1)$, the MGF is known to be:

$$M_X(t) = e^{t^2/2}.$$

This follows from integrating the product of e^{tx} and the PDF of $N(0,1)$:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx = e^{t^2/2}.$$

Then, third derivative:

$$M_X^{(3)}(t) = \frac{d}{dt}\left(e^{t^2/2} + t^2 e^{t^2/2}\right).$$

Differentiate term by term:

$$M_X^{(3)}(t) = t^3 e^{t^2/2} + 3te^{t^2/2}.$$

Now evaluate $M_X^{(3)}(t)$ at $t = 0$:

$$M_X^{(3)}(0) = (0^3 + 3 \cdot 0)e^0 = 0.$$

Thus:

$$\mathbb{E}[X^3] = M_X^{(3)}(0) = 0.$$

Hence, the difference simplifies to:

$$\left|\mathbb{E}[Z_n^3] - \mathbb{E}[N^3]\right| = \left|\mathbb{E}[Z_n^3]\right| \leq \frac{\sum_{j=1}^n \gamma_j}{s_n^3} = \frac{\Gamma_n}{s_n^3}.$$

Since $\mathbb{E}[\mathcal{N}^3] = 0$, the (2) simplifies to:

$$|\mathbb{E}[R(Z_n)] - \mathbb{E}[R(\mathcal{N})]| \leq \frac{M}{6}\mathbb{E}[Z_n^3].$$

Substituting $\mathbb{E}[Z_n^3] \leq \frac{\Gamma_n}{s_n^3}$, we get:

$$\mathbb{E}[R(Z_n)] \leq \frac{M}{6}\frac{\Gamma_n}{s_n^3}.$$

Finally, by the assumption:

$$\frac{\Gamma_n}{s_n^3} \to 0 \quad \text{as } n \to \infty,$$

it follows that:

$$\mathbb{E}[f(Z_n)] \to \mathbb{E}[f(\mathcal{N})].$$