Lecture On Central Limit Theorem

Lemma 1 *For any t in $\mathbb{R}$, $n \geq 1$ and δ in $[0,1]$, one has*

$$e^{it} - \sum_{k=0}^{n} \frac{(it)^k}{k!} = \frac{(it)^{n+1}}{n!} \int_0^1 e^{itu}(1-u)^n d\,u \tag{1}$$

$$= i^{n+1} \int_0^t d\,t_{n+1} \int_0^{t_{n+1}} d\,t_n \cdots \int_0^{t_2} e^{it_1} dt_1. \tag{2}$$

and

$$|e^{it} - \sum_{k=0}^{n} \frac{(it)^k}{k!}| \leq \frac{2^{1-\delta}|t|^{n+\delta}}{\prod_{k=1}^{n}(k+\delta)}. \tag{3}$$

Proof. Let $t \in \mathbb{R}, n \geq 1, 0 \leq \delta \leq 1$ be fixed. For any $1 \leq k \leq n$, set $a_1 = b_1 = e^{it} - 1$ and

$$\begin{aligned} a_{k+1} &= \frac{(it)^{k+1}}{k!} \int_0^1 e^{itu}(1-u)^k d\,u \\ b_{k+1} &= i^{k+1} \int_0^t d\,t_{k+1} \int_0^{t_{k+1}} d\,t_k \cdots \int_0^{t_2} e^{it_1} d\,t_1. \end{aligned}$$

Through integrating by parts we obtain

$$\begin{aligned} a_{k+1} &= a_k - \frac{(it)^k}{k!} \\ b_{k+1} &= i^k \int_0^t d\,s_{k+1} \int_0^{s_{k+1}} d\,s_k \cdots \int_0^{s_3} (e^{is_2} - 1) d\,s_2 \\ &= i^k \int_0^t d\,s_{k+1} \int_0^{s_{k+1}} d\,s_k \cdots \int_0^{s_3} e^{is_2} d\,s_2 - i^k \int_0^t d\,s_{k+1} \int_0^{s_{k+1}} d\,s_k \cdots \int_0^{s_3} d\,s_2 \\ &= b_k - \frac{(it)^k}{k!} \end{aligned}$$

where the last equality follows from a change of variable $s_{k+1} = t_k$. By induction we have $a_k = b_k$ for all k and thus (2). The equation (1) follows from the fact that

$$\begin{aligned} e^{it} &= 1 + a_1 \\ &= 1 + \sum_{k=1}^{n} (a_k - a_{k+1}) + a_{n+1}. \end{aligned}$$

It is clear that for any x with $|x| \geq 1$, $|\sin(x)| \leq |x|^\delta$. If $|x| < 1$, then by mean value theorem we have

$$|\sin x| \leq |x| \leq |x|^\delta.$$

Thus for any x we have

$$|\sin x| \leq |x|^\delta. \tag{4}$$

By direct calculation, we have

$$\begin{aligned} |e^{it}-1| &= \sqrt{2(1-\cos t)} \\ &= 2|\sin\frac{t}{2}| \le 2|\frac{t}{2}|^{\delta} \\ &= 2^{1-\delta}|t|^{\delta}. \end{aligned}$$

The result (3) can now be derived as follows.

$$\begin{aligned} |e^{it}-\sum_{k=0}^{n}\frac{(it)^k}{k!}| &\le \int_0^{|t|} d\,t_{n+1}\cdots\int_0^{t_3}|\int_0^{t_2}e^{it_1}d\,t_1|d\,t_2 \\ &= \int_0^{|t|} d\,t_{n+1}\cdots\int_0^{t_3}|e^{it_2}-1|d\,t_2 \\ &\le 2^{1-\delta}\int_0^{|t|} d\,t_{n+1}\cdots\int_0^{t_3}t_2^{\delta}d\,t_2 \\ &= \frac{2^{1-\delta}|t|^{n+\delta}}{\prod_{k=1}^{n}(k+\delta)}. \end{aligned}$$

□

Lemma 2 *For any complex number z with $|z|\le 1$, we have*

$$|e^z-1-z|\le |z|^2. \qquad (5)$$

If $z=x$ is non-positive real number, then

$$|e^x-1-x|\le \frac{x^2}{2}. \qquad (6)$$

Proof.

$$\begin{aligned} |e^z-1-z| &= |\sum_{k=2}^{\infty}\frac{z^k}{k!}| \\ &\le \sum_{k=2}^{\infty}\frac{|z|^k}{k!} \\ &\le |z|^2\sum_{k=2}^{\infty}\frac{1}{2^{k-1}}=|z|^2. \end{aligned}$$

If $x\le 0$, then

$$e^x-(1+x)=e^{\xi}\frac{x^2}{2}$$

for some ξ between x and 0.

□

Theorem 3 (Lindeberg-Feller CLT) *Let $\{X_n : n\ge 1\}$ be a sequence of independent random variables with mean zero and variances $\{\sigma_n^2 : n\ge 1\}$. Set*

$$S_n=\sum_{k=1}^{n}X_k,\quad B_n^2=\sum_{k=1}^{n}\sigma_k^2.$$

1 *If the Lindeberg condition holds, i.e., for any* $\varepsilon > 0$

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} \mathbb{E}[X_k^2 I_{\{|X_k|>\varepsilon B_n\}}]}{B_n^2} = 0, \tag{7}$$

then

$$\frac{S_n}{B_n} \longrightarrow Z \sim N(0,1) \tag{8}$$

in distribution.

2 *Assume that Feller's conditions hold, i.e.,*

$$\lim_{n\to\infty} B_n = \infty$$
$$\lim_{n\to\infty} \frac{\sigma_n}{B_n} = 0.$$

If (8) *holds, then the Lindeberg condition holds.*

Proof. **Proof of 1:** Let $S_0 = 0, B_0 = 0$. Then we have

$$\begin{aligned}
&|\mathbb{E}[e^{it\frac{S_n}{B_n}}] - e^{-\frac{t^2}{2}}| \\
&= \left|\mathbb{E}[\exp[it\frac{S_n}{B_n}]] - e^{-\frac{t^2}{2}}\mathbb{E}[\exp[it\frac{S_0}{B_n}]]\right| \\
&= e^{-\frac{t^2}{2}}\left|\mathbb{E}\left[\exp[it\frac{S_n}{B_n} + \frac{B_n^2 t^2}{2B_n^2}]\right] - \mathbb{E}\left[\exp[it\frac{S_0}{B_n} + \frac{B_0^2 t^2}{2B_n^2}]\right]\right| \\
&= e^{-\frac{t^2}{2}}|\sum_{k=1}^{n}\left(\mathbb{E}[\exp[it\frac{S_k}{B_n} + \frac{B_k^2 t^2}{2B_n^2}]] - \mathbb{E}[\exp[it\frac{S_{k-1}}{B_n} + \frac{B_{k-1}^2 t^2}{2B_n^2}]]\right)| \\
&\le e^{-\frac{t^2}{2}}\sum_{k=1}^{n}\left|\left(\mathbb{E}[\exp[it\frac{S_k}{B_n} + \frac{B_k^2 t^2}{2B_n^2}]] - \mathbb{E}[\exp[it\frac{S_{k-1}}{B_n} + \frac{B_{k-1}^2 t^2}{2B_n^2}]]\right)\right| \\
&\le e^{-\frac{t^2}{2}}\sum_{k=1}^{n}\left|\mathbb{E}[\exp[it\frac{S_{k-1}}{B_n} + \frac{B_k^2 t^2}{2B_n^2}]]\left(\mathbb{E}[\exp[it\frac{X_k}{B_n}]] - e^{-\frac{\sigma_k^2 t^2}{2B_n^2}}\right)\right| \\
&\le e^{-\frac{t^2}{2}} e^{\frac{t^2}{2}}\sum_{k=1}^{n}\left|\mathbb{E}[e^{it\frac{X_k}{B_n}}] - e^{-\frac{\sigma_k^2 t^2}{2B_n^2}}\right|.
\end{aligned} \tag{9}$$

Set

$$Y_k(t) = e^{itX_k} - 1 - itX_k + \frac{t^2 X_k^2}{2}$$

and

$$h_k(t) = e^{-\frac{\sigma_k^2}{2}t^2} - 1 + \frac{\sigma_k^2 t^2}{2}.$$

By the Lemma 1,

$$\begin{aligned}
|Y_k(t)| &\le \min\{|a_2| + \frac{t^2 X_k^2}{2}, |a_3|\} \\
&\le \min\{t^2 X_k^2, \frac{|tX_k|^3}{6}\}
\end{aligned}$$

which implies that

$$
\begin{aligned}
\left|\mathbb{E}[e^{it\frac{X_k}{B_n}}] - e^{-\frac{\sigma_k^2 t^2}{2B_n^2}}\right| &= \left|\mathbb{E}[Y_k(\frac{t}{B_n})] - h_k(\frac{t}{B_n})\right| \qquad (10)\\
&\leq \mathbb{E}\left[\frac{t^2}{B_n^2}X_k^2 I_{\{|X_k|>\varepsilon B_n\}} + \varepsilon\frac{|t|^3|X_k|^2}{B_n^2}I_{\{|X_k|\leq\varepsilon B_n\}}\right]\\
&\quad +\frac{\sigma_k^4 t^4}{8B_n^4}
\end{aligned}
$$

where (6) is used in the last inequality to $h_k(\frac{t}{B_n})$,i.e.,

$$
|h_k(\frac{t}{B_n})| \leq \frac{(\sigma_k^2 t^2/2B_n^2)^2}{2} = \frac{\sigma_k^4 t^4}{8B_n^4}.
$$

Putting (9) and (10) together one obtains

$$
\begin{aligned}
&|\mathbb{E}[e^{it\frac{S_n}{B_n}}] - e^{-\frac{t^2}{2}}|\\
&\leq \sum_{k=1}^{n}\mathbb{E}[\frac{t^2}{B_n^2}X_k^2 I_{\{|X_k|>\varepsilon B_n\}}] + \varepsilon|t|^3 \qquad (11)\\
&\quad + \max_{1\leq k\leq n}\frac{\sigma_k^2}{B_n^2}\frac{t^4}{8}.
\end{aligned}
$$

Noting that

$$
\begin{aligned}
\max_{1\leq k\leq n}\frac{\sigma_k^2}{B_n^2} &= \max_{1\leq k\leq n}B_n^{-2}\left(\mathbb{E}[X_k^2 I_{\{|X_k|>\varepsilon B_n\}}] + \mathbb{E}[X_k^2 I_{\{|X_k|\leq\varepsilon B_n\}}]\right)\\
&\leq \varepsilon^2 + \frac{1}{B_n^2}\sum_{k=1}^{n}\mathbb{E}[X_k^2 I_{\{|X_k|>\varepsilon B_n\}}],
\end{aligned}
$$

it follows from the Lindeberg condition that

$$
\lim_{n\to\infty}\mathbb{E}[e^{it\frac{S_n}{B_n}}] - e^{-\frac{t^2}{2}}| = 0. \qquad (12)
$$

Proof of 2: Assume that Feller's conditions hold. Noting that for any $1 \leq m \leq n$

$$
\max_{1\leq k\leq n}\frac{\sigma_k^2}{B_n^2} \leq \max_{1\leq k\leq m}\frac{\sigma_k^2}{B_n^2} + \max_{m\leq k\leq n}\frac{\sigma_k^2}{B_k^2},
$$

it follows by letting n going to infinity followed by m going to infinity that

$$
lim_{n\to\infty}\max_{1\leq k\leq n}\frac{\sigma_k^2}{B_n^2} = 0. \qquad (13)
$$

By Lemma 1, we have that for any $1 \leq k \leq n$

$$
\max_{1\leq k\leq n}|\mathbb{E}[e^{it\frac{X_k}{B_n}}] - 1| \leq \frac{t^2}{2}\max_{1\leq k\leq n}\frac{\sigma_k^2}{B_n^2} = o(1) \qquad (14)
$$

and

$$
\sum_{k=1}^{n}|\mathbb{E}[e^{it\frac{X_k}{B_n}}] - 1| \leq \frac{t^2}{2}.
$$

For complex number z satisfying $|z| < 1$ one has

$$\log(1+z) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{z^k}{k}.$$

If $|z| < \frac{1}{2}$, then

$$|\log(1+z) - z| \leq \frac{|z|^2}{2} \sum_{k=2}^{\infty} |z|^{k-2} \leq |z|^2$$

This combined with (14) implies that for n large enough we have

$$\begin{aligned} \sum_{k=1}^{n} |\log \mathbb{E}[e^{it\frac{X_k}{B_n}}] - (\mathbb{E}[e^{it\frac{X_k}{B_n}}] - 1)| &\leq \sum_{k=1}^{n} |\mathbb{E}[e^{it\frac{X_k}{B_n}}] - 1|^2 \\ &\leq \frac{t^2}{2} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{B_n^2} = o(1) \end{aligned}$$

Applying the central limit theorem we get

$$\lim_{n \to \infty} \left| \frac{t^2}{2} - \sum_{k=1}^{n} (1 - \mathbb{E}[e^{it\frac{X_k}{B_n}}]) \right| = 0.$$

Focusing on the real part, we obtain

$$\frac{t^2}{2} - \sum_{k=1}^{n} \mathbb{E}[1 - \cos \frac{tX_k}{B_n}] = o(1).$$

Since $1 - \cos \frac{tX_k}{B_n} \leq \frac{t^2 X_k^2}{2B_n^2}$, it follows that

$$\begin{aligned} &\frac{t^2}{2} \frac{1}{B_n^2} \sum_{k=1}^{n} \mathbb{E}[X_k^2 I_{\{|X_k| > \varepsilon B_n\}}] \\ &\leq \frac{t^2}{2} - \sum_{k=1}^{n} \mathbb{E}[(1 - \cos \frac{tX_k}{B_n}) I_{\{|X_k| \leq \varepsilon B_n\}}] \\ &= \sum_{k=1}^{n} \mathbb{E}[(1 - \cos \frac{tX_k}{B_n}) I_{\{|X_k| > \varepsilon B_n\}}] + o(1) \\ &\leq 2 \sum_{k=1}^{n} \mathbb{E}[I_{\{|X_k| > \varepsilon B_n\}}] + o(1) \\ &\leq \frac{2}{\varepsilon^2 B_n^2} \sum_{k=1}^{n} \mathbb{E}[X_k^2] + o(1) \\ &= \frac{2}{\varepsilon^2} + o(1). \end{aligned}$$

Letting n go to infinity followed by t going to infinity we obtain the Lindeberg condition. □