Lecture 20

Instructor: Quan Zhou

For more details about the materials covered in this note, see Chapters 9.7 and 9.8 of Resnick [2] and Chapter 3.4 of Durrett [1].

20.1 Auxiliary lemmas

Lemma 20.1. *Let $z_1, \dots, z_n$ and $w_1, \dots, w_n$ be complex numbers with modulus bound (from above) by θ. Then,*

$$\left| \prod_{k=1}^n z_k - \prod_{k=1}^n w_k \right| \leq \theta^{n-1} \sum_{k=1}^n |z_k - w_k|.$$

Proof. By triangle inequality,

$$\begin{aligned} \left| \prod_{k=1}^n z_k - \prod_{k=1}^n w_k \right| &\leq \left| \prod_{k=1}^n z_k - z_1 \prod_{k=2}^n w_k \right| + \left| z_1 \prod_{k=2}^n w_k - \prod_{k=1}^n w_k \right| \\ &\leq \theta \left| \prod_{k=2}^n z_k - \prod_{k=2}^n w_k \right| + \theta^{n-1} |z_1 - w_1|. \end{aligned}$$

Hence, if the lemma is true for some fixed $n-1$, we have

$$\left| \prod_{k=1}^n z_k - \prod_{k=1}^n w_k \right| \leq \theta^{n-1} \left(\sum_{k=2}^n |z_k - w_k| \right) + \theta^{n-1} |z_1 - w_1|.$$

That is, the lemma is also true for n. Since the case $n = 1$ is obvious, we can use induction to conclude the proof. □

Lemma 20.2. *If $c_n \to c \in \mathbb{C}$, then $(1 + c_n/n)^n \to e^c$.*

Proof. For any complex number c with $|c| \leq 1$, we have $e^c = \sum_{k=0}^\infty b^k/k!$. Hence,

$$|e^c - 1 - c| \leq \sum_{k=2}^\infty \frac{|c|^k}{k!} \leq |c|^2 \sum_{k=2}^\infty \frac{1}{k!} \leq M|c|^2$$

for some constant $M \in (0, \infty)$. In particular, it holds for $M = 1$. Since $c_n \to c$, we can assume that for sufficiently large n, $|c_n| \le K$ for some $K < \infty$. By Lemma 20.1 (let $z_k = 1 + c_n/n$ and $w_k = e^{c_n/n}$),

$$\left|\left(1 + \frac{c_n}{n}\right)^n - e^{c_n}\right| \le e^{K/n} \sum_{k=1}^n |1 + n^{-1}c_n - e^{c_n/n}| \le \frac{e^{K(n-1)/n}|c_n|^2}{n} \le \frac{K^2 e^K}{n},$$

which vanishes as $n \to \infty$. The result then follows by using $e^{c_n} \to e^c$ (exp is continuous in the complex plane). □

Lemma 20.3. *For a triangular array $\{c_{n,k} \in \mathbb{R}\}$, if $\max_{1 \le k \le n} |c_{n,k}| \to 0$, $\sum_{k=1}^n c_{n,k} \to \lambda$ and $\sup_n \sum_{k=1}^n |c_{n,k}| < \infty$, then $\prod_{k=1}^n (1 + c_{n,k}) \to e^\lambda$.*

Proof. Since $\max_{1 \le k \le n} |c_{n,k}| \to 0$, we may assume that $|c_{n,k}| < 1/2$ for every n and k without loss of generality. For $|c| < 1/2$, one can easily verify $c - c^2 \le \log(1 + c) \le c$. Hence,

$$\limsup_{n\to\infty} \sum_{k=1}^n \log(1 + c_{n,k}) \le \limsup_{n\to\infty} \sum_{k=1}^n c_{n,k} = \lambda.$$

For the other direction,

$$\begin{aligned}\liminf_{n\to\infty} \sum_{k=1}^n \log(1 + c_{n,k}) \ge \liminf_{n\to\infty} \sum_{k=1}^n (c_{n,k} - c_{n,k}^2) &= \lambda - \limsup_{n\to\infty} \sum_{k=1}^n c_{n,k}^2 \\ &\ge \lambda - \limsup_{n\to\infty} \max_{1\le k\le n} |c_{n,k}| \sum_{k=1}^n |c_{n,k}| = \lambda.\end{aligned}$$

Note that all the three assumptions have been used. □

Remark 20.1. Consider the sub-array $\{c_{2n,k}\}$ (that is, we only consider even rows). Let $c_{2n,k} = 1/\sqrt{n}$ if k is odd, and $c_{2n,k} = -1/\sqrt{n}$ if k is even. Then, clearly $\lambda = \sum_{k=1}^{2n} c_{2n,k} = 0$, and $\max_{1\le k\le 2n} |c_{2n,k}| \to 0$. However, $\sup_{2n} \sum_{k=1}^{2n} |c_{2n,k}| = \infty$ and $\prod_{k=1}^{2n}(1 + c_{2n,k}) \to e^{-1}$.

20.2 Central limit theorem for i.i.d. sequences

Theorem 20.1. *Let $X_1, X_2, \dots$ be i.i.d. random variables with mean μ and variance $\sigma^2 < \infty$. Let $S_n = X_1 + \cdots + X_n$. Then,*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{D} Z,$$

where Z is a normal random variable with $E[Z] = 0$ and $\mathrm{Var}(Z) = 1$.

Proof. Without loss of generality, we can assume $\mu = 0$. By Theorem 19.2 and Jensen's inequality,

$$\left| E(e^{itX_1}) - E\left(1 + itX - \frac{t^2 X_1^2}{2}\right) \right| \leq E \min\left\{ \frac{|tX_1|^3}{6}, |tX_1|^2 \right\} =: h(t).$$

Hence, letting ϕ_X denote the characteristic function of X_1, we get

$$\phi_X(t) = 1 - \frac{\sigma^2 t^2}{2} + r(t), \qquad |r(t)| \leq h(t).$$

By Proposition 19.2,

$$E[e^{itS_n/(\sigma\sqrt{n})}] = \phi_X\left(\frac{t}{\sigma\sqrt{n}}\right)^n = \left(1 - \frac{t^2}{2n} + r(t/\sigma\sqrt{n})\right)^n$$

We claim that for every $t \in \mathbb{R}$,

$$\lim_{n\to\infty} n\, h\left(\frac{t}{\sigma\sqrt{n}}\right) = 0. \tag{1}$$

This implies that $E[e^{itS_n/(\sigma\sqrt{n})}] \to e^{-t^2/2}$ by Lemma 20.2. The CLT then follows from the continuity and uniqueness theorems of characteristic functions. To verify (1), note that

$$\frac{n}{t^2} h\left(\frac{t}{\sigma\sqrt{n}}\right) = E \min\left\{ \frac{|t||X_1|^3}{6\sigma^3 n^{1/2}}, \frac{|X_1|^2}{\sigma^2} \right\} \leq 1,$$

since $EX_1^2 = \sigma^2$. Hence, by the dominated convergence theorem,

$$\lim_{n\to\infty} \frac{n}{t^2} h\left(\frac{t}{\sigma\sqrt{n}}\right) \leq E\left[\lim_{n\to\infty} \frac{|t||X_1|^3}{6\sigma^3 n^{1/2}}\right] = 0,$$

for any $t \in \mathbb{R}$. □

Remark 20.2. Consider the error term $h(t)$. Note that because $h(t) \leq t^2 EX_1^2 < \infty$, we can apply DCT to show that $h(t) = o(t^2)$; that is, $h(t)/t^2 \to 0$ as $t \to 0$. Thus, we may write $\phi_X(t) = 1 - \sigma^2 t^2/2 + o(t^2)$.

Example 20.1. Let $X_1, X_2, \dots$ be i.i.d. with $\mathsf{P}(X_1 = 1) = \mathsf{P}(X_1 = -1) = 1/2$ and let $S_n = X_1 + \cdots + X_n$. The De Moivre-Laplace theorem states that

$$\lim_{n\to\infty} \mathsf{P}(a \leq S_n/\sqrt{n} \leq b) = \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx,$$

for any $a < b$, i.e. $S_n/\sqrt{n}$ converges in distribution to a standard normal random variable. Of course this is just a special case of the above central limit theorem, but it can be proven by straightforward calculations of the probability mass function of S_n using Stirling's formula:

$$n! \sim n^n e^{-n}\sqrt{2\pi n}, \qquad \text{as } n \to \infty.$$

Example 20.2. Let $\xi_1, \xi_2, \dots$ be i.i.d. with $\mathsf{P}(\xi_i = 1) = \mathsf{P}(\xi_i = -1) = 1/2$. Define $\{X_n\}_{n\geq 1}$ as

$$X_1 = \xi_1, \quad X_2 = \xi_1\xi_2, \quad X_3 = \xi_1\xi_3, \quad X_4 = \xi_1\xi_2\xi_3,$$
$$X_5 = X_1\xi_4, \quad X_6 = X_2\xi_4, \quad X_7 = X_3\xi_4, \quad X_8 = X_4\xi_4, \quad \dots$$

That is, for $m = 2^{n-1}+j$, where $0 < j \leq 2^{n-1}$ and $n \geq 1$, we let $X_m = X_j\xi_{n+1}$. By construction, it is easy to check that all X_i's are pairwise independent, which yields that $E[S_n] = 0$ and $\mathrm{Var}(S_n) = n$. But the central limit theorem fails since

$$S_{2^n} = \xi_1(1+\xi_2)(1+\xi_3)\cdots(1+\xi_{n+1}),$$

which satisfies $\mathsf{P}(S_{2^n} = 0) = 1 - 2^{-n} \to 1$.

20.3 Lindeberg-Feller central limit theorem

Theorem 20.2 (Lindeberg-Feller CLT)**.** *Consider a triangular array of random variables $\{X_{n,k}\colon n \geq 1,\ 1 \leq k \leq n\}$ where for each n, $X_{n,1}, \dots, X_{n,n}$ are independent with mean zero and finite variance. Suppose*

(i) $\sum_{k=1}^n EX_{n,k}^2 \to \sigma^2 \in (0,\infty)$,

(ii) for all $\epsilon > 0$, $\lim_{n\to\infty} \sum_{k=1}^n E\left(|X_{n,k}|^2 \mathbb{1}_{\{|X_{n,k}|>\epsilon\}}\right) = 0$.

Then $S_n/\sigma \xrightarrow{D} Z$ where Z has standard normal distribution and $S_n = X_{n,1} + \cdots + X_{n,n}$.

Proof. Let $\phi_{n,k}$ denote the characteristic function of $X_{n,k}$ and $\sigma_{n,k}^2 = EX_{n,k}^2$. By Theorem 19.2,

$$\begin{aligned}\left|\phi_{n,k}(t) - 1 + \frac{t^2\sigma_{n,k}^2}{2}\right| &\leq E\min\left\{\frac{|tX_{n,k}|^3}{6}, |tX_{n,k}|^2\right\} \\ &\leq E\left(\frac{|tX_{n,k}|^3}{6}\mathbb{1}_{\{|X_{n,k}|\leq\epsilon\}}\right) + E\left(t^2X_{n,k}^2\mathbb{1}_{\{|X_{n,k}|>\epsilon\}}\right) \\ &\leq \frac{\epsilon|t|^3}{6}E\left(X_{n,k}^2\mathbb{1}_{\{|X_{n,k}|\leq\epsilon\}}\right) + t^2E\left(X_{n,k}^2\mathbb{1}_{\{|X_{n,k}|>\epsilon\}}\right).\end{aligned}$$

Summing over $k = 1, \dots, n$, we get

$$\sum_{k=1}^{n} \left| \phi_{n,k}(t) - 1 + \frac{t^2 \sigma_{n,k}^2}{2} \right|$$
$$\leq \frac{\epsilon |t|^3}{6} \sum_{k=1}^{n} E\left(X_{n,k}^2\right) + t^2 \sum_{k=1}^{n} E\left(X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| > \epsilon\}}\right) \to \frac{\sigma^2 \epsilon |t|^3}{6}$$

by conditions (i) and (ii). Letting $\epsilon \downarrow 0$, we obtain that the left-hand side goes to zero as $n \to \infty$. Next, we notice that

$$\max_{1 \leq k \leq n} \sigma_{n,k}^2 \leq \epsilon^2 + \max_{1 \leq k \leq n} E\left(X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| > \epsilon\}}\right) \leq \epsilon^2 + \sum_{k=1}^{n} E\left(X_{n,k}^2 \mathbb{1}_{\{|X_{n,k}| > \epsilon\}}\right).$$

Condition (ii) implies that $\limsup_{n \to \infty} \max_k \sigma_{n,k}^2 \leq \epsilon^2$. Letting $\epsilon \downarrow 0$, we get $\max_k \sigma_{n,k}^2 \to 0$. Hence, for any fixed t, there exists N such that for all $n \geq N$ and all $1 \leq k \leq n$, we have $1 - t^2 \sigma_{n,k}^2 / 2 \geq -1$. So, we may apply Lemma 20.1 with $\theta = 1$ to obtain

$$\left| \prod_{k=1}^{n} \phi_{n,k}(t) - \prod_{k=1}^{n} \left(1 - \frac{t^2 \sigma_{n,k}^2}{2}\right) \right| \to 0.$$

By Lemma 20.3, for every t, as $n \to \infty$,

$$\prod_{k=1}^{n} \left(1 - \frac{t^2 \sigma_{n,k}^2}{2}\right) \to e^{-t^2 \sigma^2 / 2},$$

which yields the result. □

Example 20.3. The CLT for i.i.d. sequences is just a special case. To see this, let $Y_1, Y_2, \dots$ be i.i.d. with mean zero and let $X_{n,k} = Y_k / \sqrt{n}$. Then $\sum_{k=1}^{n} E X_{n,k}^2 = \sigma^2$ and Lindeberg's condition (i.e. the second condition in Theorem 20.2) can be verified by DCT.

Example 20.4. Define independent random variables $\{Y_n\}_{n \geq 1}$ by

$$\mathsf{P}(Y_n = \pm 1) = \frac{1 - c}{2}, \quad \mathsf{P}(Y_n = \pm n) = \frac{c}{2n^2}, \quad \mathsf{P}(Y_n = 0) = \frac{c(n^2 - 1)}{n^2},$$

for some $c \in (0,1)$. It is easy to check that $E(Y_n) = 0$ and $\mathrm{Var}(Y_n) = 1$ for every n. Now define $X_{n,k} = Y_k/\sqrt{n}$. Then $\sum_{k=1}^n EX_{n,k}^2 = 1$. However, Lindeberg's condition is not satisfied since for any fixed $\epsilon > 0$ and $n > \epsilon^2$,

$$\sum_{k=1}^{n} E\left(|X_{n,k}|^2 \mathbb{1}_{\{|X_{n,k}|>\epsilon\}}\right) \geq \sum_{k=\lceil \epsilon\sqrt{n}\rceil}^{n} E\left(|X_{n,k}|^2 \mathbb{1}_{\{|X_{n,k}|>\epsilon\}}\right)$$

$$\geq \sum_{k=\lceil \epsilon\sqrt{n}\rceil}^{n} E\left(|X_{n,k}|^2 \mathbb{1}_{\{|X_{n,k}|=k/\sqrt{n}\}}\right) = c\frac{n - \lceil \epsilon\sqrt{n}\rceil}{n} \to c > 0.$$

Further, though we do not prove here, $S_n = X_{n,1} + \cdots + X_{n,n}$ does not converge in distribution to a standard normal.

References

[1] Rick Durrett. *Probability: Theory and Examples*, volume 49. Cambridge university press, 2019.

[2] Sidney Resnick. *A Probability Path*. Springer, 2019.

[3] Jordan M Stoyanov. *Counterexamples in probability*. Courier Corporation, 2013.