Lab 10

Laws of Large Numbers & Central Limit Theorems

10.1. Strong Law of Large Numbers

We present three versions of the SLLN, each successively requiring fewer conditions (existence of the fourth moment, to existence of the second moment, to existence of only the first moment). Throughout, let $S_n = \sum_{k=1}^n X_k$ and $\overline{X}_n = \frac{1}{n} S_n$.

Theorem 10.1.1. (Cantelli) Let $\{X_k\}$ be independent with $\mathbb{E}X_k^4 < \infty$ and $\mathbb{E}|X_k - \mathbb{E}X_k|^4 \leq C < \infty$. Then $\overline{X}_n \overset{\text{a.s.}}{\to} \mathbb{E}\overline{X}_n$.

Proof. WLOG, assume $\mathbb{E}X_k = 0$. We have

$$S_n^4 = \underbrace{\sum_{i=1}^n X_i^4}_{A_n} + \underbrace{\binom{4}{2} \sum_{i<j} X_i^2 X_j^2}_{B_n} + \underbrace{\binom{4}{2,1,1} \sum_{\substack{i \neq j \\ i \neq k \\ j<k}} X_i^2 X_j X_k}_{C_n} + \underbrace{\binom{4}{3} \sum_{i \neq k} X_i^3 X_j}_{D_n} + \underbrace{4! \sum_{i<j<k<l} X_i X_j X_k X_l}_{E_n}$$

Since $\mathbb{E}X_k = 0$ and each of the C_n, D_n, E_n summands have at least one term of X_k (to the first power), their expected values are all 0, but A_n, B_n remain:

$$\begin{aligned}
\mathbb{E}S_n^4 &= \sum_{i=1}^n \mathbb{E}X_i^4 + 6 \sum_{i<j} \mathbb{E}X_i^2 \mathbb{E}X_j^2 \\
&\leq nC + 6 \sum_{i<j} \sqrt{\mathbb{E}X_i^4 \mathbb{E}X_j^4} && \text{Bounded condition \& Cauchy Schwarz} \\
&\leq nC + 6 \frac{n(n-1)}{2} C && \text{Bounded condition} \\
&= (3n^2 - 2n)C < 3n^2 C
\end{aligned}$$

Hence

$$\sum_{n=1}^\infty \mathbb{P}(|\overline{X}_n| \geq \varepsilon) \leq \frac{1}{\varepsilon^4} \sum_{n=1}^\infty \mathbb{E}|\overline{X}_n|^4 < \frac{3C}{\varepsilon^4} \sum_{n=1}^\infty \frac{1}{n^2} < \infty$$

for any $\varepsilon > 0$. Hence showing complete convergence to 0, which implies almost-sure convergence to 0. □

The conditions for Cantelli's SLLN can be considerably weakened by the use of more precise methods. In this way we can prove a stronger law of large numbers (Kolmogorov's SLLN from Class/Lecture 19). We will take a different approach from lecture that is more intuitive. We need the following (intuitive) lemmas.

Lemma 10.1.1. (Toeplitz) Let $\{a_k\}$ satisfy $a_k \geq 0$ and $\sum_{k=1}^n a_k \to \infty$. Let $x_k \to x$. Then

$$\frac{\sum_{k=1}^n a_k x_k}{\sum_{k=1}^n a_k} \to x$$

In particular, if $a_k = 1$, then $\frac{x_1 + \cdots + x_n}{n} \to x$.

Lemma 10.1.2. (Kronecker) Let $\{b_k\}$ be positive and increasing with $b_k \to \infty$. Let $\{x_k\}$ be such that $\sum_{k=1}^n x_k$ converges. Then

$$\frac{1}{b_n} \sum_{k=1}^n b_k x_k \to 0$$

In particular, if $b_n = n, x_k = \frac{y_k}{k}$ so that $\sum_{k=1}^n \frac{y_k}{k}$ converges, then $\frac{y_1 + \cdots + y_n}{n} \to 0$.

Theorem 10.1.2. (Kolmogorov Version I) Let $\{X_k\}$ be independent with $\mathbb{E}X_k^2 < \infty$ and $\{b_n\}$ being positive, increasing, $b_n \to \infty$, and

$$\sum_{k=1}^\infty \frac{\operatorname{Var}(X_k)}{b_k^2} < \infty \tag{$\star$}$$

Then $(S_n - \mathbb{E}S_n)/b_n \overset{\text{a.s.}}{\to} 0$. In particular, for $b_n = n$, we have $\overline{X}_n \overset{\text{a.s.}}{\to} \mathbb{E}\overline{X}_n$.

Proof. Since

$$\frac{S_n - \mathbb{E}S_n}{b_n} = \frac{1}{b_n} \sum_{k=1}^n b_k \left(\frac{X_k - \mathbb{E}X_k}{b_k} \right)$$

A sufficient condition, by Kronecker's lemma, is that $\xi_n \overset{\text{def}}{=} \sum_{k=1}^n \left(\frac{X_k - \mathbb{E}X_k}{b_k} \right)$ converges a.s. This is guaranteed by applying Kolmogorov's inequality with ($\star$). □

Example 10.1.1. Let X_k iid with $\mathbb{P}(X_k = \pm 1) = \frac{1}{2}$. Let $b_n = \sqrt{n}\log(n)$. By Kolmogorov Version I, we have $\frac{S_n}{\sqrt{n}\log(n)} \overset{\text{a.s.}}{\to} 0$.

Now let's prove a version of SLLN where we don't even require the existence of the second moment. Since there is no free lunch, we now require our random variables to be iid (not just independent). We need another lemma.

Lemma 10.1.3. Let W be a non-negative random variable. Then $\sum_{n=1}^\infty \mathbb{P}(W \geq n) \leq \mathbb{E}W \leq 1 + \sum_{n=1}^\infty \mathbb{P}(W \geq n)$.

Theorem 10.1.3. (Kolmogorov Version II) Let $\{X_k\}$ iid. The following statements are equivalent:

- $\mathbb{E}|X_k| < \infty$ and common mean μ

- $\overline{X}_n \overset{\text{a.s.}}{\to} \mu$

Remark: I believe Marcello calls this Khintchine's SLLN, but I learned it as another one of Kolmogorov's theorems. Note that this version is if and only if!

Proof. Because we aren't guaranteed a bounded second moment for X_n, the trick is to establish asymptotic equivalence with the locally bounded RV $\widetilde{X}_n \overset{\text{def}}{=} X_n \mathbb{I}(|X_n| < n)$ and show the local boundedness has variance that does not "grow too quickly". WLOG, again assume $\mu = 0$. By the previous lemma and Borel-Cantelli,

$$\mathbb{E}|X_k| < \infty \iff \sum_{n=1}^{\infty} \mathbb{P}(|X_k| \geq n) < \infty \iff \sum_{n=1}^{\infty} \mathbb{P}(|X_n| \geq n) < \infty \iff \mathbb{P}(|X_n| \geq n \text{ i.o.}) = 0$$

which implies that there exists an N such that $|X_n| < n$ with probability 1 for all $n > N$. This gives us $\mathbb{E}\widetilde{X}_n \to \mathbb{E}X_n = 0$. Hence by Toeplitz's lemma,

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\widetilde{X}_k \to 0$$

and consequently $\overline{X}_n \overset{\text{a.s.}}{\to} 0$ if and only if

$$\frac{1}{n}\sum_{k=1}^{n} (\widetilde{X}_k - \mathbb{E}\widetilde{X}_k) \overset{\text{a.s.}}{\to} 0 \tag{$\clubsuit$}$$

Hence, we establish asymptotic equivalence between $\widetilde{X}_n$ and X_n. Now we will show the not "grow too quickly" part. Denote $\widehat{X}_k = \widetilde{X}_k - \mathbb{E}\widetilde{X}_k$. By Kronecker's lemma, it suffices to prove $\sum_{k=1}^{n} \frac{\widehat{X}_k}{k}$ converges almost surely to establish ($\clubsuit$). In turn, it suffices to prove $\sum_{k=1}^{\infty} \text{Var}(\widehat{X}_k)/k^2 < \infty$ by Kolmogorov's inequality. Through some tedious inequality chasing (this is why you should get good with inequalities!), we can establish

$$\sum_{k=1}^{\infty} \frac{\text{Var}(\widehat{X}_k)}{k^2} \leq 2\mathbb{E}|X_1| < \infty$$

Hence we have the forward direction of our result. The converse is left as an (easy) exercise. □

10.2. Central Limit Theorem

Theorem 10.2.1. (CLT for iid RV) Let $\{X_i\}$ iid with mean μ and variance σ^2. Then

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma^2} \overset{\mathcal{D}}{\to} N(0,1)$$

Proof. WLOG, assume X_i has mean 0 and variance 1. Then

$$\begin{aligned}\varphi_{\sqrt{n}\overline{X}_n}(t) &= \mathbb{E}e^{it\sqrt{n}\overline{X}_n} = (\mathbb{E}e^{itX_1/\sqrt{n}})^n = \varphi_X(t/\sqrt{n})^n \\ &= \left(\varphi_X(0) + \varphi_X'(0)\frac{t}{\sqrt{n}} + \frac{1}{2}\varphi_X''(0)\frac{t^2}{n} + o\left(\frac{t^2}{n}\right)\right)^n\end{aligned}$$

$$= \left(1 - \frac{t^2}{2n} + o(n^{-1})\right)^n$$
$$\to e^{-t^2/2}$$

which is the CF of N(0, 1). □

Theorem 10.2.2. (Lindeberg CLT) Let X_k be independent, but each with mean μ_k, variance σ_k^2, and distribution function F_k. Let $D_n^2 = \sum_{k=1}^n \sigma_k^2$. If the *Lindeberg condition* is satisfied: for all $\varepsilon > 0$,

$$\frac{1}{D_n^2} \sum_{k=1}^n \text{Var}(X_k \mathbb{I}(|X_k - \mu_k| \geq \varepsilon D_n)) \to 0$$

Then $(S_n - \mathbb{E}S_n)/D_n \to N(0, 1)$.

Proof. Too long. Let's have a beer and discuss this. □

Theorem 10.2.3. (Lyapunov CLT) Replace Lindeburg's condition with

$$\frac{1}{D_n^{2+\delta}} \sum_{k=1}^n \mathbb{E}|X_k - \mu_k|^{2+\delta} \to 0$$

for some $\delta > 0$.

Proof. WLOG let $\mu_k = 0$. Clearly $\mathbb{E}|X_k \mathbb{I}(|X_k| \geq \varepsilon D_n)|^2 \leq (\varepsilon^\delta D_n^\delta)^{-1} \mathbb{E}|X_k|^{2+\delta}$, hence

$$\frac{1}{D_n^2} \sum_{k=1}^n \mathbb{E}|X_k \mathbb{I}(|X_k| \geq \varepsilon D_n)|^2 \leq \frac{1}{\varepsilon^\delta} \frac{1}{D_n^{2+\delta}} \sum_{k=1}^n \mathbb{E}|X_k|^{2+\delta} \to 0$$

Since Lindeburg's condition is satisfied, Lyapunov's CLT is confirmed. □

I would love to include some CLT theory for *dependent* sequences, but involves a ton of theory regarding martingales. Learn more about this in Prob II.

10.3. Additional Problems

1. Let $f : [0, \infty) \to \mathbb{R}$ be a bounded, continuous function. Compute

$$\lim_{n\to\infty} 2^n \int_0^\infty \cdots \int_0^\infty f\left(\frac{x_1 + \cdots + x_n}{n}\right) e^{-2(x_1+\cdots+x_n)} dx_1 \cdots dx_n$$

[Hint: Apply the strong law of large numbers to a suitably chosen iid sequence.]

2. Suppose $\{X_k\}$ independent with

$$\mathbb{P}(X_k = x) = \begin{cases} 1 - \frac{1}{k\log(1+k)} & x = 0 \\ \frac{1}{2k\log(1+k)} & x = -k, k \\ 0 & \text{otherwise} \end{cases}$$

Show that $\overline{X}_n \xrightarrow{\mathcal{P}} 0$, but $\overline{X}_n \overset{\text{a.s.}}{\not\to} 0$.

3. Show that Lindeberg's condition is satisfied when $\{X_k\}$ iid with mean μ and variance σ^2.

4. Let X, Y independent with $X \sim \text{Pois}(\lambda)$ and $Y \sim \text{Pois}(\mu)$. Show that

$$\frac{(X - \lambda) - (Y - \mu)}{\sqrt{X + Y}} \xrightarrow{\mathcal{D}} N(0, 1) \quad \text{as } \lambda, \mu \to \infty$$

[Hint: Rewrite $\frac{1}{\sqrt{X+Y}} = \frac{1}{\sqrt{\lambda+\mu}} \frac{\sqrt{\lambda+\mu}}{\sqrt{X+Y}}$ and apply Slutsky's.]