§ 1.7. Strong law of large numbers.

Strong law of large numbers (SLLN) is a central result in classical probability theory. The convergence of series estabalished in Section 1.6 paves a way towards proving SLLN using the Kronecker lemma.

(i). *Kronecker lemma and Kolmogorov's criterion of SLLN.*

Kronecker Lemma. Suppose $a_n > 0$ and $a_n \uparrow \infty$. Then $\sum_n x_n/a_n < \infty$ implies $\sum_{j=1}^n x_j/a_n \to 0$.

Proof. Set $b_n = \sum_{i=1}^n x_i/a_i$ and $a_0 = b_0 = 0$. Then, $b_n \to b_\infty < \infty$ and $x_n = a_n(b_n - b_{n-1})$. Write

$$
\begin{aligned}
\frac{1}{a_n}\sum_{j=1}^n x_j = \frac{1}{a_n}\sum_{j=1}^n a_j(b_j - b_{j-1}) &= \frac{1}{a_n}\Big[\sum_{j=1}^n a_j b_j - \sum_{j=1}^n a_j b_{j-1}\Big] \\
= \quad & b_n + \frac{1}{a_n}\Big[\sum_{j=1}^{n-1} a_j b_j - \sum_{j=1}^n a_j b_{j-1}\Big] = b_n + \frac{1}{a_n}\Big[\sum_{j=1}^n a_{j-1} b_{j-1} - \sum_{j=1}^n a_j b_{j-1}\Big] \\
= \quad & b_n - \frac{1}{a_n}\sum_{j=1}^n b_{j-1}(a_j - a_{j-1}) \\
\to \quad & b_\infty - b_\infty = 0.
\end{aligned}
$$

□

The following proposition is an immediate application of the Kronecker lemma and the Khintchine-Kolmogorov convergence of series.

Proposition (Kolmogorov's criterion of SLLN). Suppose $X_1, X_2 ...,$ are independent such that $E(X_n) = 0$ and $\sum_n \mathrm{var}(X_n)/n^2 < \infty$. Then, $S_n/n \to 0$ a.e..

Proof. Consider the series $\sum_{i=1}^n X_i/i < \infty$, $n \geq 1$. Then Theorem 1.6 implies $\sum_n X_n/n < \infty$ a.s.. And the above Kronecker Lemma ensures $S_n/n \to 0$ a.s.. □

Obviously, if $X, X_1, X_2, ...$ are iid with finite variance, the above proposition implies the SLLN: $S_n/n \to E(X)$ a.s.. In fact, a stronger result than the above SLLN is also straightforward:

Corollary. If $X_1, X_2, ...$ are iid with mean μ and finite variance. Then,

$$
\frac{S_n - n\mu}{\sqrt{n(\log n)^\delta}} \to 0 \qquad .a.s.
$$

for any $\delta > 1$.

We leave the proof as an exercise.

The corollary gives a rate of a.s. convergence of sample mean S_n/n to population mean μ at a rate $n^{-1/2}(\log n)^\delta$ with $\delta > 1/2$. This is, although not the sharpest rate, close to the sharpest rate of a.s. convergence at $n^{-1/2}(\log\log n)^{1/2}$ given in *Kolmogorov's law of iterated logarithm*:

$$
\begin{cases}
\limsup \frac{S_n - n\mu}{\sqrt{2\sigma^2 n \log\log n}} = 1 \qquad a.s.. \\
\liminf \frac{S_n - n\mu}{\sqrt{2\sigma^2 n \log\log n}} = -1 \qquad a.s..
\end{cases}
$$

for iid r.v.s with mean μ and finite variance σ^2. We do not intend to cover the proofs of Kolmogorov's law of iterated logarithm.

(ii) *Kolmogorov's strong law of large numbers.*

The above SLLN requires finite moments of the series. The most standard classical SLLN, established by Kolmogorov, for iid r.v.s. holds as long as the population mean exist. In statistical view, the sample mean shall always converge to the population mean as long as the population mean exists, without any further moment condition. In fact, the sample mean converges to a finite limit if and only if the population mean is finite, in which case, the limit is the population mean.

Theorem 1.7. *Kolmogorov's strong law of large numbers.* Suppose $X, X_1, X_2, ...$ are iid and $E(X)$ exists. Then,

$$S_n/n \to E(X), \quad a.s..$$

Conversely, if $S_n/n \to \mu$ which is finite, then $\mu = E(X)$.

Proof. Suppose first $E(X_1) = 0$. We shall utilize the above proposition of Kolmogorov's criterion of SLLN. Consider

$$Y_n = X_n 1_{\{|X_n| \le n\}} - E(X_n 1_{\{|X_n| \le n\}}).$$

Write

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{\text{var}(Y_n)}{n^2} &\le \sum_{n=1}^{\infty} \frac{1}{n^2} E(X^2 1_{\{|X| \le n\}}) = E\Big(X^2 \sum_{n=1}^{\infty} \frac{1}{n^2} 1_{\{|X| \le n\}}\Big) \\ \le \quad & E\Big(X^2 \sum_{n \ge |X| \vee 1}^{\infty} \frac{2}{n(n+1)}\Big) \le 2E(|X| + 1) < \infty \end{aligned}$$

It then follows from Kolmogorov's criterion of SLLN that

$$\frac{1}{n} \sum_{i=1}^{n} Y_i \to 0 \qquad a.s..$$

Next, since $E(X_n 1_{\{|X_n| \le n\}}) \to E(X) = 0$. $\sum_{i=1}^n E(X_i 1_{\{|X_i| \le i\}})/n \to 0$. Hence,

$$\frac{1}{n} \sum_{i=1}^{n} X_i 1_{\{|X_i| \le i\}} \to 0 \qquad a.s..$$

Observe that $E(X) = 0$ implies $E|X| < \infty$, and

$$E|X| < \infty \Longleftrightarrow \sum_n P(|X| > n) < \infty \Longleftrightarrow P(|X_n| > n, i.o.) = 0 \Longleftrightarrow X_n/n \to 0 \quad a.s..$$

Therefore, $\sum_{i=1}^n X_i 1_{\{|X_i| > i\}}/n \to 0$ a.s.. As a result, the SLLN holds.

Suppose $E(X) < \infty$. the SLLN holds by considering $X_i - E(X)$, which is mean 0.

Suppose $E(X) = \infty$. Then, $(1/n)\sum_{i=1}^n X_i \wedge C \to E(X_1 \wedge C)$ a.s., which $\uparrow \infty$ when $C \uparrow \infty$. Since $S_n \ge \sum_{i=1}^n X_i \wedge C$, the SLLN holds. Likewise for the case $E(X) = -\infty$.

Conversely, if $S_n/n \to \mu$ a.s. where μ is finite, $X_n/n \to 0$ a.s.. Hence, $E|X| < \infty$ and $\mu = E(X)$ by the SLLN just proved. □

Remark Kolmogorov's SLLN also holds for r.v.s that are *pairwise independent* following the same distribution, which is slightly more general. We have chosen to follow the historic development of the classical probability theory.

(iii). *Strong law of large numbers when $E(X)$ does not exist.*

Kolmogorov's SLLN in Theorem 1.7 already shows that the classical SLLN does not hold if $E(X)$ does not exist, i.e., $E(X^+) = E(X^-) = \infty$. The SLLN becomes quite complicated. We introduce the theorem proved by W. Feller:

Proposition Suppose $X, X_1, ...$ are iid with $E|X| = \infty$. Suppose $a_n > 0$ and a_n/n is nondecreasing. Then,

$$\begin{cases} \limsup |S_n|/a_n = 0 & \text{if } \sum_n P(|X| \ge a_n) < \infty \\ \limsup |S_n|/a_n = \infty & \text{if } \sum_n P(|X| \ge a_n) = \infty. \end{cases}$$

The proof is somewhat technical but still along the same line as the that of Kolmogorov's SLLN. Interested students may refer to the textbook (page 67). We omit the details.

EXAMPLE 1.10. (THE ST. PETERSBURG PARADOX) See Example 1.5 in which we have shown

$$\frac{S_n}{n \log n} \to \frac{1}{\log 2} \qquad \text{in probability}$$

Analogous to the calculation therein,

$$\sum_{n=2}^{\infty} P(X \geq n \log n) = \sum_{n=2}^{\infty} P(X \geq 2^{\log(n \log n)/\log 2}) \geq \sum_{n=2}^{\infty} 2^{-\log(n \log n)/\log 2} = \sum_{n=2}^{\infty} 1/(n \log n) = \infty$$

By the above proposition,

$$\limsup \frac{S_n}{n \log n} = \infty \qquad a.s..$$

On the other hand, one can also show with same calculation that, for $\delta > 1$,

$$\limsup \frac{S_n}{n (\log n)^{\delta}} = 0 \qquad a.s..$$

□

The following Marcinkiewicz-Zygmund SLLN is useful in connecting the rate of convergence with the moments of the iid r.v.s.

Theorem 1.8. (MARCINKIEWICZ-ZYGMUND STRONG LAW OF LARGE NUMBERS). *Suppose $X, X_1, X_2, ...$ are iid and $E(|X|^p) < \infty$ for some $0 < p < 2$. Then,*

$$\begin{cases} \frac{S_n - nE(X)}{n^{1/p}} \to 0, \quad a.s. \qquad \text{for } 1 \leq p < 2 \\ \frac{S_n}{n^{1/p}} \to 0 \quad a.s. \qquad \text{for } 0 < p < 1. \end{cases}$$

Proof. The case with $p = 1$ is Kolmogorov's SLLN. The cases with $0 < p < 1$ and $1 < p < 2$ are consequences of the corollary following Theorem 1.6 and the Kronecker lemma. □

EXAMPLE 1.11 Suppose $X, X_1, X_2, ...$ are iid and X is symmetric with $P(X > t) = t^{-\alpha}$ for some $\alpha > 0$ and all large t.

(1). $\alpha > 2$: Then, $E(X^2) < \infty$, $S_n/n \to 0$ a.s. and, moreover, Kolmogorov's law of iterated logarithm gives the sharp rate of the a.s. convergence.

(2). $1 < \alpha \leq 2$: for any $0 < p < \alpha$

$$\frac{S_n}{n^{1/p}} \to 0, \quad a.s.$$

It implies that S_n/n converges to 0 a.s. at a rate faster than $n^{-1+1/p}$, but not at the rate of $n^{-1+1/\alpha}$. In particular, if $\alpha = 2$, S_n/n converges to $E(X)$ a.s. at a rate faster than $n^{-\beta}$ with any $0 < \beta < 1/2$, but not at the rate of $n^{-1/2}$.

(3). $0 < \alpha \leq 1$: $E(X)$ does not exist. For any $0 < p < \alpha$,

$$\frac{S_n}{n^{1/p}} \to 0, \quad a.s.$$

Moreover, the above proposition implies

$$\limsup \frac{|S_n|}{n^{1/\alpha}} = \infty \quad a.s. \quad \text{and} \quad \frac{S_n}{n^{1/\alpha} (\log n)^{\delta/\alpha}} \to 0 \qquad a.s.$$

for any $\delta > 0$.

Remark. In the above example, for $0 < \alpha < 2$, $S_n/n^{1/\alpha}$ converges *in distribution* to a nondegenerate distribution called stable law. In particular, if $\alpha = 1$, S_n/n converges in distribution to a Cauchy distribution. For $\alpha = 2$, $S_n/(n \log n)^{1/2}$ converges to a normal distribution, and for $\alpha > 2$, $S_n/n^{1/2}$ converges to a normal distribution,

□

DIY Exercises.

Exercise 1.25 If $X_1, ...$ are iid with mean μ and finite variance. Then,

$$\frac{S_n - n\mu}{\sqrt{n(\log n)^\delta}} \to 0 \qquad a.s.$$

for any $\delta > 1$.

Exercise 1.26 Suppose $X, X_1, ...$ are iid. Then, $(S_n - C_n)/n \to 0$ a.s. if and only if $E(|X|) < \infty$.

Exercise 1.27 Suppose $X, X_1, ...$ are iid with $E(|X|^p) = \infty$ for some $0 < p < \infty$. Then, $\limsup |S_n|/n^{1/p} = \infty$ a.s..

Exercise 1.28 Suppose $X_n, n \geq 1$ are independent with mean μ_n and variance σ_n^2 such that $\mu_n \to 0$ and $\sum_{j=1}^n \sigma_j^2 \to \infty$. show that

$$\frac{\sum_{j=1}^n X_j/\sigma_j^2}{\sum_{j=1}^n \sigma_j^{-2}} \to 0 \qquad a.s.$$

Hint: Consider the series $\sum_{j=1}^n (X_j - \mu_j)/(\sigma_j^2 \sum_{k=1}^j \sigma_k^{-2})$.