The Weak Law of Large Numbers

Tong Zhou

JHU

1. L^2 WEAK LAW OF LARGE NUMBERS

This section presents a stronger version of weak law of large numbers where a finite second moment condition is imposed. The following theorem summarizes this conclusion.

Theorem 1 (L^2 Weak Law of Large Numbers). *Consider $S_n = \sum_{i=1}^n X_i$ for uncorrelated random variables $X_1, \cdots, X_n, \cdots$. Suppose that $Var(X_i) \leqslant C$ and $\mathbb{E}[X_i] = \bar{x}$ for some finite constants $C, \bar{x}$, and all $i = 1, 2, \cdots$. Then, $n^{-1}S_n \xrightarrow{L^2} \bar{x}$ as $n \to \infty$, and hence also $n^{-1}S_n \xrightarrow{P} \bar{x}$.*

Exercise 1 (Amir Exercise 2.1.5). *Show that the conclusion of the L^2 weak law of large numbers holds even for correlated X_i, provided $\mathbb{E}[X_i] = \bar{x}$ and $Cov(X_i, X_j) \leqslant r(|i-j|)$ for all i, j, and some bounded sequence $r(k) \to 0$ as $k \to \infty$.*

Proof. It suffices to show $Var\left(\frac{S_n}{n}\right) \to 0$.

$$\begin{aligned}
\mathsf{Var}(n^{-1}S_n) &= \frac{1}{n^2}\left\{\sum_{i=1}^n \mathsf{Var}(X_i) + 2\sum_{i>j}\mathsf{Cov}(X_i, X_j)\right\} \\
&= \frac{1}{n^2}\left[nr(0) + 2(r(n-1) + 2r(n-2) + \cdots + (n-1)r(1))\right] \\
&= \frac{r(0)}{n} + 2\left(\frac{r(n-1)}{n^2} + \frac{2r(n-2)}{n^2} + \cdots + \frac{(n-1)r(1)}{n^2}\right) \to 0 \text{ as } n \to \infty.
\end{aligned}$$

□

2. TRUNCATION METHOD

Weak Law for Triangular Arrays; identical dist. NOT required

Theorem 2 (Weak Law for Triangular Arrays). *Suppose that for each n, the random variables $X_{n,k}, k = 1, \cdots, n$ are pairwise independent. Let $\bar{X}_{n,k} = X_{n,k}, \mathbb{1}_{|X_{n,k} \leqslant b_n|}$ for non-random $b_n > 0$ such that as $n \to \infty$ both*

(1) $\sum_{k=1}^n \mathbb{P}(|X_{n,k}| > b_n) \to 0$,
and

(2)

$$\frac{\sum_{k=1}^n \mathsf{Var}(\bar{X}_{n,k})}{b_n^2} \to 0.$$

Then, $b_n^{-1}(S_n - a_n) \xrightarrow{P} 0$ as $n \to \infty$, where $S_n = \sum_{k=1}^n X_{n,k}$ and $a_n = \sum_{k=1}^n \mathbb{E}[\bar{X}_{n,k}]$.

PROOF. Define

$$\bar{S}_n = \sum_{k=1}^{n} \bar{X}_{n,k}.$$

Observe that for any $\varepsilon > 0$,

$$\left\{ \frac{S_n - a_n}{b_n} > \varepsilon \right\} \subseteq \left\{ S_n \neq \bar{S}_n \right\} \cup \left\{ \frac{\bar{S}_n - a_n}{b_n} > \varepsilon \right\}.$$

Then

$$\mathbb{P}\left(\frac{S_n - a_n}{b_n} > \varepsilon \right) \leqslant \mathbb{P}(S_n \neq \bar{S}_n) + \mathbb{P}\left(\frac{\bar{S}_n - a_n}{b_n} > \varepsilon \right).$$

Now it suffices to show the RHS is $o_P(1)$.

For the first term:

$$\mathbb{P}(S_n \neq \bar{S}_n) \leqslant \sum_{k=1}^{n} \mathbb{P}(\left|X_{n,k}\right| > b_n) \longrightarrow 0 \quad \text{(by condition (1))}.$$

So $\mathbb{P}(S_n \neq \bar{S}_n) \longrightarrow 0$ as $n \longrightarrow \infty$.

For the second term:

$$\begin{aligned} \mathbb{P}\left(\frac{\bar{S}_n - a_n}{b_n} > \varepsilon \right) &\leqslant \frac{\mathsf{Var}(\bar{S}_n)}{\varepsilon^2 b_n^2} \\ &= \frac{\sum_{k=1}^{n} \mathsf{Var}(\bar{X}_{n,k})}{\varepsilon^2 b_n^2} \longrightarrow 0 \quad \text{(by condition (2))}. \end{aligned}$$

So $\mathbb{P}\left(\frac{\bar{S}_n - a_n}{b_n} > \varepsilon \right) \longrightarrow 0$ as $n \longrightarrow \infty$.

Therefore, we have

$$\mathbb{P}\left(\frac{S_n - a_n}{b_n} > \varepsilon \right) \longrightarrow 0, \text{ as } n \longrightarrow \infty.$$

$\square$

A special case of Theorem 2 for a single sequence yields the following.

Weak Law for a Single Sequence; i.i.d. required

PROPOSITION 1. *Consider i.i.d. random variables $\{X_i\}$, such that $x\mathbb{P}(|X_1| > x) \longrightarrow 0$ as $x \longrightarrow \infty$. Then, $n^{-1}S_n - \mu_n \xrightarrow{P} 0$, where $S_n = \sum_{i=1}^{n} X_i$ and $\mu_n = \mathbb{E}[X_1 \mathbb{1}_{\{|X_1| \leqslant n\}}]$.*

PROOF. $\{X_i\}$ is a special case where $n = 1$, and $i = 1, \cdots, n$.
Define $\bar{X}_i = X_i \mathbb{1}_{|X_i| \leqslant n}$. Check the two conditions:
For the first condition: $\sum_{i=1}^{n} \mathbb{P}(X_i \neq \bar{X}_i) = \sum_{i=1}^{n} \mathbb{P}(|X_i| > n) = n\mathbb{P}(|X_1| > n) \longrightarrow 0$ as $n \longrightarrow \infty$.
For the second condition, $n^{-2} \sum_{i=1}^{n} \mathsf{Var}(\bar{X}_i) = n^{-1}\mathsf{Var}(\bar{X}_1) = n^{-1}\left(\mathbb{E}[\bar{X}_1^2] - (\mathbb{E}[\bar{X}_1])^2\right) \leqslant n^{-1}\mathbb{E}[\bar{X}_1^2]$.
Observe that $\mathbb{P}\left(\left|\bar{X}_1\right| > y\right) = \mathbb{P}(|X_1| > y) - \mathbb{P}(|X_1| > n) \leqslant \mathbb{P}(|X_1| > y)$, where $y \in (0, n)$.

Therefore,

$$\begin{aligned} n^{-1}\mathbb{E}[\bar{X}_1^2] &= n^{-1}\int_0^n 2y\mathbb{P}\left(\left|\bar{X}_1\right| > y\right)\,\mathrm{d}y \\ &\leqslant n^{-1}\int_0^n 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y \end{aligned}$$

Then we want to show $n^{-1}\int_0^n 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y \to 0$ as $n \to \infty$. Since by assumption $2y\mathbb{P}(|X_1| > y) \to 0$ as $y \to \infty$, for any $\varepsilon > 0$, $\exists \bar{y} \in \mathbb{R}$ such that $2y\mathbb{P}(|X_1| > y) < \varepsilon$ whenever $y > \bar{y}$. Since n will go to ∞ eventually, we assume $\bar{y} < n$. Hence,

$$\begin{aligned} n^{-1}\int_0^n 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y &= n^{-1}\int_0^{\bar{y}} 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y + n^{-1}\int_{\bar{y}}^n 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y \\ &\leqslant \underbrace{n^{-1}\int_0^{\bar{y}} 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y}_{\to 0,\text{as } n\to 0} + \varepsilon\left(\frac{n-\bar{y}}{n}\right) \\ &\leqslant \frac{\bar{y}^2}{n} + \varepsilon\left(\frac{n-\bar{y}}{n}\right) \\ &\to \varepsilon, \quad \text{as } n \to \infty. \end{aligned}$$

Thus,

$$\limsup_{n\to\infty} \frac{1}{n}\int_0^n 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y \leqslant \varepsilon.$$

Let $\varepsilon \downarrow 0$, then we have

$$n^{-1}\int_0^n 2y\mathbb{P}(|X_1| > y)\,\mathrm{d}y \to 0, \quad \text{as } n \to 0.$$

Thus, the second condition holds:

$$n^{-2}\sum_{i=1}^n \mathsf{Var}(\bar{X}_1) \to 0 \quad \text{as} \quad n \to \infty.$$

□

REMARK 1. *The two conditions are quite weak. Finite expectation is actually not assumed. See* **??** *2.*

- why

$$\mathbb{E}[Z^2] = \int_0^\infty 2y\mathbb{P}(|Z| > y)\,\mathrm{d}y$$

LEMMA 1 (Lemma 1.4.31 of Amir Dembo).

1. *For any $r > p > 0$ and any random variable $Y \geqslant 0$,*

$$\mathbb{E}[Y^p] = \int_0^\infty py^{p-1}\mathbb{P}(Y > y)\,\mathrm{d}y = \int_0^\infty py^{p-1}\mathbb{P}(Y \geqslant y)\,\mathrm{d}y$$
$$= (1 - \frac{p}{r})\int_0^\infty py^{p-1}\mathbb{E}[\min(Y/y, 1)^r]\,\mathrm{d}y.$$

PROOF.

1. Define $h_p(y) := py^{p-1}\mathbb{1}_{y>0}$, then $H_p(x) := \int_{-\infty}^x h_p(y)\,\mathrm{d}y = \int_0^x py^{p-1}\,\mathrm{d}y$. By the change of variables formula,

$$\mathbb{E}[X^p] = \mathbb{E}[H_p(X)] = \int_\Omega H_p(X)\,\mathrm{d}\mathbb{P} = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} py^{p-1}\mathbb{1}_{(0,x)}\,\mathrm{d}\lambda(y)\right)\mathrm{d}\mathbb{P}_X(x)$$
$$= \int_{\mathbb{R}} py^{p-1}\mathbb{1}_{y>0}\left(\int_{\mathbb{R}} \mathbb{1}_{y<x}\,\mathrm{d}\mathbb{P}_X(x)\right)\mathrm{d}\lambda(y)$$
$$= \int_0^\infty py^{p-1}\mathbb{P}(X > y)\,\mathrm{d}\lambda(y)$$
$$= \int_0^\infty py^{p-1}\mathbb{P}(X \geqslant y)\,\mathrm{d}\lambda(y)$$

□

$Var(X_1)$ not exist ; $\mathbb{E}(|X_1|) \equiv \infty$ ($X_1 \notin L^1$); i.i.d. required

EXERCISE 2 (Amir: Exercise 2.1.13). *Let $\{X_i\}$ be i.i.d. with $\mathbb{P}(X_1 = (-1)^k k) = 1/(ck^2\log k)$ for integers $k \geqslant 2$ and a normalization constant $c = \sum_k 1/(k^2\log k)$. Show that $\mathbb{E}[X_1] = \infty$, but there is a non-random $\mu < \infty$ such that $\frac{S_n}{n} \xrightarrow{P} \mu$.*

PROOF. •

$$\begin{aligned}\mathbb{E}[|X_1|] &= \sum_{k=2}^{\infty} k\frac{1}{ck^2\log k}\\ &= \frac{1}{c}\sum_{k=2}^{\infty}\frac{1}{k\log k}\end{aligned}$$

By the integral test, consider the function $f(x) = \frac{1}{x\log x}$ for $x \geqslant 2$. Then

$$\int_2^{\infty}\frac{1}{x\log x}\,\mathrm{d}x = \log(\log x)\Big|_2^{\infty} = \infty.$$

Thus the log-series $\mathbb{E}[|X_1|] = \frac{1}{c}\sum_{k=2}^{\infty}\frac{1}{k\log k} = \infty$.
Also notice that $\mathbb{E}[X_1] < \infty$ and $\mathsf{Var}(X_1) = \infty$. So the L^2-Weak Law of Large Numbers is not applicable.
To apply Proposition 1, we first check

$$\begin{aligned}n\mathbb{P}(|X_1| > n) &\leqslant n\sum_{k\geqslant n}\frac{1}{ck^2\log k}\\ &= \frac{1}{cn\log n} + \sum_{k>n}\frac{1}{ck^2\log k} \longrightarrow 0.\end{aligned}$$

Therefore, by Proposition 1, we have

$$\frac{S_n}{n} - \mu_n \xrightarrow{\mathsf{P}} 0, n \longrightarrow \infty.$$

where $\mu_n = \mathbb{E}[X_1\mathbb{1}_{|X_1|\leqslant n}]$.
Let $\mu = \mathbb{E}[X_1]$. We next show that $\mu_n - \mu \longrightarrow 0$:

$$\begin{aligned}\mu - \mu_n &= \mathbb{E}[X_1\mathbb{1}_{|X_1|>n}]\\ &= \sum_{k>n}(-1)^k\frac{1}{ck\log k}\\ &\longrightarrow 0,\end{aligned}$$

since the series $\sum_{k\geqslant 2}(-1)^k\frac{1}{ck\log k}$ is an alternating series with $\sum_{k\geqslant 2}\frac{1}{ck\log k} < \infty$.
By the Slutsky theorem, then

$$\left(\frac{S_n}{n} - \mu\right) + (\mu - \mu_n) \xrightarrow{\mathsf{P}} 0, \quad \text{as } n \longrightarrow \infty,$$

or equivalently,

$$\frac{S_n}{n} \xrightarrow{\mathsf{P}} \mu,$$

where $\mu = \mathbb{E}[X_1]$ is non-random.

□

i.i.d. $X_i \in L^1$

COROLLARY 1. *Consider* $S_n = \sum_{k=1}^n X_k$ *for i.i.d. random variables* $\{X_i\}$ *such that* $\mathbb{E}[|X_1|] < \infty$. *Then,* $\frac{S_n}{n} \xrightarrow{P} \mathbb{E}[X_1]$ *as* $n \to \infty$.

PROOF. It is easy to show $n\mathbb{P}(|X_1| > n) \to 0$. Thus, apply Proposition 1

$$\frac{S_n}{n} - \mu_n \xrightarrow{\mathsf{P}} 0.$$

Also, it is easy to show

$$\mu - \mu_n \to 0$$

where $\mu = \mathbb{E}[X_1]$ and $\mu = \mathbb{E}[X_1 \mathbb{1}_{|X_1| \leqslant n}]$. $\square$