6. Laws of large numbers

6.1. **Strong law of large numbers.** Consider a sequence of random variables with the same expectation, say μ. The strong law of large numbers, or briefly SLLN, means

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \mu \quad \text{a.s.}$$

Theorem 6.1 (Kolmogorov's SLLN)**.** *If $(X_n)_{n=1}^\infty$ are independent and $\sum_{n=1}^\infty \mathrm{Var}(X_n)/n^2 < \infty$, then the strong law of large numbers holds.*

Proof. It loses no generality to assume that $\mu = 0$. For $n \geq 1$, set $S_n = X_1 + \cdots + X_n$ and let $b_n = \mathrm{Var}(X_n) = \mathbb{E}X_n^2$. Note that

$$\frac{S_n}{n} \to 0 \quad \text{a.s.} \quad \Leftrightarrow \quad \mathbb{P}\left(\frac{|S_n|}{n} > \epsilon \text{ i.o.}\right) = 0, \quad \forall \epsilon > 0 \quad \Leftrightarrow \quad \mathbb{P}(A_k(\epsilon) \text{ i.o.}) = 0, \quad \forall \epsilon > 0$$

where

$$A_k(\epsilon) = \{|S_n| \geq \epsilon n \text{ for some } n \in [2^k, 2^{k+1}]\}.$$

By the Borel-Cantelli lemma, it suffices to show that $\sum_k \mathbb{P}(A_k(\epsilon)) < \infty$ for all $\epsilon > 0$. To see this, observe that

$$A_k(\epsilon) \subset \{|S_n| \geq \epsilon 2^k \text{ for some } n \leq 2^{k+1}\} = \left\{\max_{1 \leq n \leq 2^{k+1}} |S_n| \geq \epsilon 2^k\right\}$$

By Kolmogorov's inequality, this implies

$$\mathbb{P}(A_k(\epsilon)) \leq \frac{1}{\epsilon^2 2^{2k}} \mathbb{E}S_{2^{k+1}}^2 = \frac{1}{\epsilon^2 2^{2k}} \sum_{n=1}^{2^{k+1}} b_n.$$

Thus, we have

$$\begin{aligned} \epsilon^2 \sum_{k=1}^\infty \mathbb{P}(A_k(\epsilon)) &\leq \sum_{k=1}^\infty \sum_{n=1}^{2^{k+1}} b_n 2^{-2k} = \sum_{k=1}^\infty \sum_{n=1}^\infty b_n 2^{-2k} \mathbf{1}_{[1, 2^{k+1}]}(n) \\ &= \sum_{n=1}^\infty b_n \sum_{k=1}^\infty 2^{-2k} \mathbf{1}_{[\log_2 n - 1, \infty)}(k) \leq \sum_{n=1}^\infty b_n \times \frac{n^{-2} 2^2}{1 - 1/4} < \infty \end{aligned}$$

□

Exercise 6.1. Let $X_1, X_2, ...$ be independent random variables satisfying

$$\mathbb{P}(X_n = n) = \mathbb{P}(X_n = -n) = \frac{p_n}{2}, \quad \mathbb{P}(X_n = 0) = 1 - p_n.$$

Prove that the SLLN holds if and only if $\sum_{n=1}^\infty \mathbb{E}X_n^2/n^2 < \infty$.

Corollary 6.2. *Let $X_1, X_2, ...$ be independent random variables with $\mu_n = \mathbb{E}X_n$. If the sequence $\frac{1}{n}(\mu_1 + \mu_2 + \cdots + \mu_n)$ converges to μ and $\sum_{n=1}^\infty \frac{\mathrm{Var}(X_n)}{n^2} < \infty$, then $\frac{1}{n}\sum_{i=1}^n X_n$ converges to μ a.s..*

Corollary 6.3. *Let $X_1, X_2, ...$ be independent with the same mean. If $\mathrm{Var}(X_n)$ is bounded, then the SLLN holds.*

Theorem 6.4 (Khinchin's SLLN)**.** *If $X_1, X_2, ...$ are i.i.d. with $\mathbb{E}|X_1| < \infty$, then the SLLN holds.*

Proof. For $n \geq 1$, let $Y_n = X_n \mathbf{1}_{\{|X_n| \leq n\}}$ and set

$$S_n = \sum_{i=1}^{n} X_i, \quad S'_n = \sum_{i=1}^{n} Y_n.$$

Consider the following two steps.

Step 1: $\frac{1}{n}(S_n - S'_n) \to 0$ a.s. To see this, observe that

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n \neq Y_n) = \sum_{n=1}^{\infty} \mathbb{E}\mathbf{1}_{\{|X_1|>n\}} = \mathbb{E}\left(\sum_{n=1}^{\infty} \mathbf{1}_{\{|X_1|>n\}}\right) \leq \mathbb{E}|X_1| < \infty$$

By the Borel-Cantelli lemma, $\mathbb{P}(X_n \neq Y_n \text{ i.o.}) = 0$ and this implies $(S_n - S'_n)/n \to 0$ almost surely.

Step 2: We will prove in the following that $\sum_{n=1}^{\infty} \mathrm{Var}(Y_n)/n^2 < \infty$. Note that

$$\sum_{n=1}^{\infty} \frac{\mathrm{Var}(Y_n)}{n^2} \leq \sum_{n=1}^{\infty} \frac{1}{n^2}\mathbb{E}[X_n^2 \mathbf{1}_{\{|X_n| \leq n\}}] = \sum_{n=1}^{\infty} \frac{1}{n^2}\mathbb{E}[X_1^2 \mathbf{1}_{\{|X_1| \leq n\}}] = \mathbb{E}\left(X_1^2 \sum_{n=1}^{\infty} \frac{\mathbf{1}_{\{|X_1| \leq n\}}}{n^2}\right)$$

Further, if $X_1(\omega) \neq 0$, then

$$\sum_{n=1}^{\infty} \frac{\mathbf{1}_{\{|X_1| \leq n\}}(\omega)}{n^2} = \sum_{n=\lceil |X_1(\omega)| \rceil}^{\infty} \frac{1}{n^2} \leq \frac{2}{|X_1(\omega)|}.$$

This leads to

$$\sum_{n=1}^{\infty} \frac{1}{n^2}\mathrm{Var}(Y_n) \leq 2\mathbb{E}|X_1| < \infty.$$

Note that $\mathbb{E}Y_n \to \mathbb{E}X_1$. By Corollary 6.2, S'_n/n converges almost surely to $\mathbb{E}X_1$. By Step 1, $S_n/n \to \mathbb{E}X_1$ a.s.. $\square$

Exercise 6.2. Let $X_n \geq 0$ be i.i.d. random variables with $\mathbb{E}X_1 = \infty$. Prove that

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to \infty \text{ a.s.}$$

Example 6.1. Let $\Omega = [0, 1)$ and $\mathbb{P}$ be the Lebesgue measure on Ω. For $\omega \in \Omega$, consider the binary expression of ω, $0.\omega_1\omega_2...$, with infinitely many 0. For $n \in \mathbb{N}$, let $X_n : \Omega \to \{0, 1\}$ be defined by $X_n(\omega) = \omega_n$. Note that, for $n \in \mathbb{N}$ and $x_1, ..., x_n \in \{0, 1\}$,

$$\{X_i = x_i, \forall 1 \leq i \leq n\} = \{\omega | 0.x_1x_2...x_n \leq \omega < 0.x_1x_2...x_n + 2^{-n}\}.$$

This implies that $(X_n)_{n=1}^{\infty}$ are random variables with $\mathbb{P}(X_i = x_i, \forall 1 \leq i \leq n) = 2^{-n}$ and, hence, are i.i.d. Bernoulli random variables with parameter $1/2$. By the strong law of large numbers, we have

$$\mathbb{P}\left(\omega : \frac{1}{n}\sum_{i=1}^{n} \omega_i \to \frac{1}{2}\right) = 1.$$

6.2. **Weak law of large numbers.** Recall that a sequence of random variables $(X_n)_{n=1}^{\infty}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to converges to X in probability if

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0 \quad \forall \epsilon > 0.$$

Theorem 6.5 (The weak law of large numbers)**.** *Let $X_1, X_2, ...$ be a sequence of independent random variables with $\mathbb{E}X_n = a_n$. Assume that $\frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) \to 0$. Then,*

$$\frac{X_1 + \cdots + X_n}{n} - \frac{a_1 + \cdots + a_n}{n} \to 0 \quad \text{in probability.}$$

Proof. By setting $Y_k = X_k - a_k$ and $S_n = Y_1 + \cdots + Y_n$, one has

$$\mathbb{E}\left(\frac{S_n^2}{n^2}\right) = \frac{1}{n^2}\sum_{k=1}^{n} \mathrm{Var}(X_k) \to 0.$$

This implies that S_n/n converges to 0 in $L^2(\mathbb{P})$ and, hence, in probability. □

Example 6.2. Consider Bernstein's polynomials.

$$B_n(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k(1-x)^{n-k} \quad \forall x \in (0,1).$$

It is obvious that $B_n(0) = f(0)$ and $B_n(1) = f(1)$ for $n \geq 1$. We shall prove in the following that $B_n(x) \to f(x)$ for $x \in (0,1)$.

Let $X_1, X_2, ...$ be independent Bernoulli random variables with parameters $p \in (0,1)$. Observe that $\mathbb{E}X_n = p$ and

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathrm{Var}(X_k) = \frac{p(1-p)}{n} \to 0 \quad \text{as } n \to \infty.$$

Set $S_n = X_1 + \cdots + X_n$. By the WLLN, $S_n/n \to p$ in probability. This implies that if f is continuous on $[0,1]$, then

$$f\left(\frac{S_n}{n}\right) \to f(p) \quad \text{in probability.}$$

Note that

$$B_n(p) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\mathbb{P}(S_n = k) = \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right]$$

For $\epsilon > 0$ and $n \geq 1$, one has

$$|B_n(p) - f(p)| \leq \mathbb{E}\left|f\left(\frac{S_n}{n}\right) - f(p)\right| \leq 2c\mathbb{P}\left(\left|f\left(\frac{S_n}{n}\right) - f(p)\right| > \epsilon\right) + \epsilon,$$

where $c = \max\{|f(x)| : 0 \leq x \leq 1\}$. This implies $B_n(p) \to f(p)$ for all $p \in (0,1)$.

Exercise 6.3. Prove that $\max_{x\in[0,1]} |B_n(x) - f(x)| \to 0$ as $n \to \infty$. (Hint: The chebyshev inequality.)

Theorem 6.6 (The WLLN for arrays)**.** *For each n, let $X_{n,k}$ with $1 \leq k \leq n$ be independent random variables. Let $b_n > 0$ satisfying $b_n \to \infty$ and set $\overline{X}_{n,k} = X_{n,k}1_{\{|X_{n,k}| \leq b_n\}}$. Suppose that*

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

(2) $b_n^{-2}\sum_{k=1}^{n} \mathbb{E}\overline{X}_{n,k}^2 \to 0$ *as* $n \to \infty$.

Set $S_n = X_{n,1} + \cdots + X_{n,n}$ and $s_n = \sum_{k=1}^{n} \mathbb{E}\overline{X}_{n,k}$. Then,

$$\frac{S_n - s_n}{b_n} \to 0 \quad \textit{in probability.}$$

Proof. Let $\overline{S}_n = \overline{X}_{n,1} + \cdots + \overline{X}_{n,n}$. Note that

$$\mathbb{P}\left(\left|\frac{S_n - s_n}{b_n}\right| > \epsilon\right) \leq \mathbb{P}(S_n \neq \overline{S}_n) + \mathbb{P}\left(\left|\frac{\overline{S}_n - s_n}{b_n}\right| > \epsilon\right).$$

The first term of the right side is bounded by

$$\mathbb{P}(S_n \neq \overline{S}_n) \leq \mathbb{P}\left(\bigcup_{k=1}^{n}\{X_{n,k} \neq \overline{X}_{n,k}\}\right) \leq \sum_{k=1}^{n} \mathbb{P}(|X_{n,k}| > b_n) \to 0.$$

By Chebyshev's inequality, the second terms is bounded by

$$\mathbb{P}\left(\left|\frac{\overline{S}_n - s_n}{b_n}\right| > \epsilon\right) \le \epsilon^{-2}\mathbb{E}\left|\frac{\overline{S}_n - s_n}{b_n}\right|^2 = \frac{1}{\epsilon^2 b_n^2}\sum_{k=1}^n \mathrm{Var}(\overline{X}_{n,k}) \le \frac{1}{\epsilon^2 b_n^2}\sum_{k=1}^n \mathbb{E}|\overline{X}_{n,k}|^2 \to 0$$

□

Corollary 6.7. *Let $X_1, X_2, \ldots$ be i.i.d. random variables satisfying*

$$x\mathbb{P}(|X_1| > x) \to 0 \quad \text{as } x \to \infty. \tag{6.1}$$

Let $S_n = X_1 + \cdots + X_n$ and $\mu_n = \mathbb{E}(X_1 \mathbf{1}_{\{|X_1| \le n\}})$. Then $S_n/n - \mu_n \to 0$ in probability. In particular, if $X_1, X_2, \ldots$ are i.i.d. with $E|X_1| < \infty$, then $(X_1 + \cdots + X_n)/n \to \mathbb{E}X_1$ in probability.

Proof. Observe that for $x \ge 1$ with $x \in [n, n+1]$,

$$n\mathbb{P}(|X_1| > n+1) \le x\mathbb{P}(|X_1| > x) \le (n+1)\mathbb{P}(|X_1| > n).$$

By (6.1), one has $n\mathbb{P}(|X_1| > n) \to 0$. Let $X_{n,k} = X_k$ for $1 \le k \le n$ and $n \ge 1$. For such an array, we have

$$\sum_{k=1}^n \mathbb{P}(|X_{n,k}| > n) = n\mathbb{P}(|X_1| > n) \to 0$$

and

$$\begin{aligned}
&\frac{1}{n^2}\sum_{k=1}^n \mathbb{E}\left(X_{n,k}^2 \mathbf{1}_{\{|X_{n,k}| \le n\}}\right) = \frac{\mathbb{E}(X_1^2 \mathbf{1}_{\{|X_1| \le n\}})}{n} \le \frac{1}{n}\int_0^{n^2} \mathbb{P}(X_1^2 > y)dy \\
=&\frac{2}{n}\int_0^n z\mathbb{P}(|X_1| > z)dz = \frac{2}{n}\left(\int_0^{\epsilon n} z\mathbb{P}(|X_1| > z)dz + \int_{\epsilon n}^n z\mathbb{P}(|X_1| > z)dz\right) \\
\le& 2(c\epsilon + n\mathbb{P}(|X_1| > \epsilon n)),
\end{aligned}$$

where $c = \sup\{x\mathbb{P}(|X_1| > x) : x > 0\}$. Letting $n \to \infty$ and then $\epsilon \to 0$ implies the desired property. □