10.1. **Strong law of large numbers for finite fourth moment.** The result we obtain in this section will be largely superseded in the next. We include it because its proof is much more elementary than that needed for the definitive version of the strong law which follows.

Theorem 10.1.1. *Let $(X_n : n \in \mathbb{N})$ be a sequence of independent random variables such that, for some constants $\mu \in \mathbb{R}$ and $M < \infty$,*

$$\mathbb{E}(X_n) = \mu, \quad \mathbb{E}(X_n^4) \leq M \quad \text{for all } n.$$

Set $S_n = X_1 + \cdots + X_n$. Then

$$S_n/n \to \mu \quad \text{a.s., as } n \to \infty.$$

Proof. Consider $Y_n = X_n - \mu$. Then $Y_n^4 \leq 2^4(X_n^4 + \mu^4)$, so

$$\mathbb{E}(Y_n^4) \leq 16(M + \mu^4)$$

and it suffices to show that $(Y_1 + \cdots + Y_n)/n \to 0$ a.s.. So we are reduced to the case where $\mu = 0$.

Note that X_n, X_n^2, X_n^3 are all integrable since X_n^4 is. Since $\mu = 0$, by independence,

$$\mathbb{E}(X_i X_j^3) = \mathbb{E}(X_i X_j X_k^2) = \mathbb{E}(X_i X_j X_k X_l) = 0$$

for distinct indices i, j, k, l. Hence

$$\mathbb{E}(S_n^4) = \mathbb{E}\left(\sum_{1 \leq i \leq n} X_k^4 + 6 \sum_{1 \leq i < j \leq n} X_i^2 X_j^2 \right).$$

Now for $i < j$, by independence and the Cauchy–Schwarz inequality

$$\mathbb{E}(X_i^2 X_j^2) = \mathbb{E}(X_i^2)\mathbb{E}(X_j^2) \leq \mathbb{E}(X_i^4)^{1/2}\mathbb{E}(X_j^4)^{1/2} \leq M.$$

So we get the bound

$$\mathbb{E}(S_n^4) \leq nM + 3n(n-1)M \leq 3n^2 M.$$

Thus

$$\mathbb{E} \sum_n (S_n/n)^4 \leq 3M \sum_n 1/n^2 < \infty$$

which implies

$$\sum_n (S_n/n)^4 < \infty \quad \text{a.s.}$$

and hence $S_n/n \to 0$ a.s.. $\square$

10.2. **Strong law of large numbers.**

Theorem 10.2.1. *Let m be a probability measure on $\mathbb{R}$, with*

$$\int_{\mathbb{R}} |x| m(dx) < \infty, \quad \int_{\mathbb{R}} x m(dx) = \nu.$$

Let $(E, \mathcal{E}, \mu)$ be the canonical model for a sequence of independent random variables with law m. Then

$$\mu(\{x : (x_1 + \cdots + x_n)/n \to \nu \text{ as } n \to \infty\}) = 1.$$

Proof. By Theorem 9.2.1, the shift map θ on E is measure-preserving and ergodic. The coordinate function $f = X_1$ is integrable and $S_n(f) = f + f \circ \theta + \cdots + f \circ \theta^{n-1} = X_1 + \cdots + X_n$. So $(X_1 + \cdots + X_n)/n \to \bar{f}$ a.e. and in L^1, for some invariant function $\bar{f}$, by Birkhoff's theorem. Since θ is ergodic, $\bar{f} = c$ a.e., for some constant c and then $c = \mu(\bar{f}) = \lim_n \mu(S_n/n) = \nu$. □

Theorem 10.2.2 (Strong law of large numbers)**.** *Let $(Y_n : n \in \mathbb{N})$ be a sequence of independent, identically distributed, integrable random variables with mean ν. Set $S_n = Y_1 + \cdots + Y_n$. Then*

$$S_n/n \to \nu \quad \text{a.s., as } n \to \infty.$$

Proof. In the notation of Theorem 10.2.1, take m to be the law of the random variables Y_n. Then $\mu = \mathbb{P} \circ Y^{-1}$, where $Y : \Omega \to E$ is given by $Y(\omega) = (Y_n(\omega) : n \in \mathbb{N})$. Hence

$$\mathbb{P}(S_n/n \to \nu \text{ as } n \to \infty) = \mu(\{x : (x_1 + \cdots + x_n)/n \to \nu \text{ as } n \to \infty\}) = 1.$$

□