If one plays a gambling game many times, one's average winning or losses per game should be roughly the expected winning or losses in each individual game; more generally, if one plays a sequence of possibly different games, one's average winning or losses should be roughly the average of the expected winning or losses in the individual games. In symbols: if $X_1, X_2, \ldots$ are integrable and independent real random variables, then the average $n^{-1}(X_1+\cdots+X_n)$ should be close to $n^{-1}(\mathbf{E}[X_1]+\cdots+\mathbf{E}[X_n])$ when n is large.

The law of large number (LLN) is a precise formulation of this idea. It comes in several versions, depending on the hypotheses one wishes to make. The LLN is important because it "guarantees" stable long-term results for the averages of some random events.

5 Weak law of large numbers

Property 5.1. Let X be a real random variable and let $f : [0, +\infty[\to [0, +\infty[$ be monotene increasing. Then for any $\epsilon > 0$ with $f(\epsilon) > 0$, the Markov's inequality holds,

$$\mathbf{P}(|X| \geq \epsilon) \leq \frac{\mathbf{E}[f(|X|)]}{f(\epsilon)}.$$

In the special case $f(x) = x^2$, we get

$$\mathbf{P}(|X| \geq \epsilon) \leq \frac{\mathbf{E}[X^2]}{\epsilon^2}.$$

In particular, if $X \in \mathcal{L}^2(\mathbf{P})$, the Chebyshev's inequality holds:

$$\mathbf{P}(|X - \mathbf{E}[X]| \geq \epsilon) \leq \frac{\mathbf{Var}[X]}{\epsilon^2}.$$

Proof. Indeed, let $\epsilon > 0$ with $f(\epsilon) > 0$,

$$\begin{aligned}
\mathbf{E}[f(|X|)] &\geq \mathbf{E}[f(|X|)\mathbf{1}_{(f(|X|)\geq f(\epsilon))}] \\
&\geq \mathbf{E}[f(\epsilon)\mathbf{1}_{(f(|X|)\geq f(\epsilon))}] \\
&= f(\epsilon)\mathbf{P}(f(|X|) \geq f(\epsilon)) \\
&\geq f(\epsilon)\mathbf{P}(|X| \geq \epsilon).
\end{aligned}$$

$\square$

Definition 5.2. Let $X, X_1, X_2, \ldots$ be real random variables. We say that (X_n) converges to X

(i) in probability, symbolically $X_n \xrightarrow{\mathrm{P}} X$, if

$$\lim_{n\to\infty} \mathbf{P}(|X_n - X| \geq \epsilon) = 0$$

for all $\epsilon > 0$, and

(ii) almost surely, symbolically $X_n \xrightarrow{\text{a.s.}} X$, if

$$\mathbf{P}\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

Remark 5.3. Almost sure convergence implies convergence in probability.

Theorem 5.4 (Weak law of large numbers)**.** Let $X_1, X_2, \ldots$ be square integrable real random variables i.i.d. Then

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{\text{P}} \mathbf{E}[X_1].$$

More precisely, for any $\epsilon > 0$, we have

$$\mathbf{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mathbf{E}[X_1]\right| > \epsilon\right) \leq \frac{\mathbf{Var}[X_1]}{\epsilon^2 n} \qquad \text{for all } n \in \mathbb{N}. \tag{5.1}$$

Proof. Let $\epsilon > 0$ and $n \in \mathbb{N}$. Then, since $X_1, \ldots, X_n$ are independent,

$$\begin{aligned}\mathbf{Var}\left[\frac{X_1 + \cdots + X_n}{n}\right] &= \frac{\mathbf{Var}[X_1] + \cdots + \mathbf{Var}[X_n]}{n^2} \\ &= \frac{\mathbf{Var}[X_1]}{n}.\end{aligned}$$

Moreover,

$$\mathbf{E}\left[\frac{X_1 + \cdots + X_n}{n}\right] = \mathbf{E}[X_1].$$

And by Chebyshev's inequality, proposition 5.1 follows. □

Example 5.5. we present a probabilistic proof of the Weierstraß's approximation theorem

Exercise 5.1. Let S_n be the number of successes in n Bernoulli trials with probability p for success on each trial.

(i) Show, using Chebyshevs inequality, that for any $\epsilon > 0$

$$\mathbf{P}\left[\left|\frac{S_n}{n} - p\right| \geq \epsilon\right] \leq \frac{p(1-p)}{n\epsilon^2}.$$

(ii) Find the maximum possible value for $p(1-p)$ if $0 < p < 1$. Using this result show that for any $\epsilon > 0$

$$\mathbf{P}\left[\left|\frac{S_n}{n} - p\right| \geq \epsilon\right] \leq \frac{1}{4n\epsilon^2}.$$

Exercise 5.2. Let $X_1, \ldots, X_n$ be independent real random variables and let S_n be their sum. Let $M_n = \mathbf{E}[X_1] + \ldots + \mathbf{E}[X_n]$ and assume that $\mathbf{Var}[X_i] < R$ for all $i \in \mathbb{N}$. Prove that, for any $\epsilon > 0$,

$$\lim_{n\to+\infty} \mathbf{P}\left[\left|\frac{S_n}{n} - \frac{M_n}{n}\right| < \epsilon\right] = 1.$$

Exercise 5.3 (Bernstein-Chernov bound)**.** Let $n \in \mathbb{N}$ and $p_1, \ldots, p_n \in [0,1]$. Let $X_1, \ldots, X_n$ be independent random variables with $X_i \sim \mathrm{Ber}_{p_i}$ for any $i = 1, \ldots, n$. Define $S_n := X_1 + \ldots + X_n$ and $m := \mathbf{E}[S_n]$. Show that for any $\delta > 0$:

$$\mathbf{P}[S_n \geq (1+\delta)m] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \right)^m$$

and

$$\mathbf{P}[S_n \leq (1-\delta)m] \leq e^{-\frac{\delta^2 m}{2}}.$$