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 $f(\epsilon) > 0$,

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

□

Definition 5.2. Let $X_1, X_2, \ldots$ be integrable real random variables and let $\tilde{S}_n := \sum_{i=1}^n (X_i - \mathbf{E}[X_i])$. We say that $X_1, X_2, \ldots$ fulfills the weak law of large numbers if

$$\lim_{n\to+\infty} \mathbf{P}\left(\left|\frac{1}{n}\tilde{S}_n\right| > \epsilon\right) = 0 \qquad \text{for any } \epsilon > 0.$$

And we say that $X_1, X_2, \ldots$ fulfills the strong law of large numbers if

$$\mathbf{P}\left(\limsup_{n\to+\infty}\left|\frac{1}{n}\tilde{S}_n\right| = 0\right) = 1.$$

Remark 5.3. The strong law of large numbers implies the weak law. Indeed, let us suppose that $X_1, X_2, \ldots$ fulfills the strong law of large numbers and let

$$A := \left(\limsup_{n\to+\infty}\left|\frac{1}{n}\tilde{S}_n\right| > 0\right) \quad \text{and} \quad A_n^\epsilon := \left(\left|\frac{1}{n}\tilde{S}_n\right| > \epsilon\right)$$

for any $n \in \mathbb{N}$ and $\epsilon > 0$. Then, since $\mathbf{P}(A) = 0$ and $\limsup_{n\to+\infty} A_n^{1/k} \uparrow A$,

$$\mathbf{P}\left(\limsup_{n\to+\infty} A_n^\epsilon\right) = 0$$

for any $\epsilon > 0$. Finally, let $\epsilon > 0$; then, by the Fatou's lemma,

$$\begin{aligned}
\limsup_{n\to+\infty} \mathbf{P}(A_n^\epsilon) &= 1 - \liminf_{n\to+\infty} \mathbf{E}[\mathbf{1}_{(A_n^\epsilon)^c}] \\
&\leq 1 - \mathbf{E}\left[\liminf_{n\to+\infty} \mathbf{1}_{(A_n^\epsilon)^c}\right] \\
&= \mathbf{E}\left[\limsup_{n\to+\infty} \mathbf{1}_{A_n^\epsilon}\right] \\
&= \mathbf{E}\left[\mathbf{1}_{\limsup A_n^\epsilon}\right] \\
&= 0.
\end{aligned}$$

Theorem 5.4. Let $X_1, X_2, \ldots$ be uncorrelated square integrable real random variables with $V := \sup_{n\in\mathbb{N}} \mathbf{Var}[X_n] < +\infty$. Then $X_1, X_2, \ldots$ fulfills the weak law of large numbers. More precisely, for any $\epsilon > 0$, we have

$$\mathbf{P}\left[\left|\frac{1}{n}\tilde{S}_n\right| > \epsilon\right] \leq \frac{V}{\epsilon^2 n} \qquad \text{for all } n \in \mathbb{N}.$$

Proof. Let $\epsilon > 0$ and $n \in \mathbb{N}$. Then, since $X_1 - \mathbf{E}[X_1], \ldots, X_n - \mathbf{E}[X_n]$ are also uncorrelated,

$$\begin{aligned}
\mathbf{Var}\left[\frac{1}{n}\tilde{S}_n\right] &= \frac{1}{n^2}\left(\mathbf{Var}[X_1 - \mathbf{E}[X_1]] + \cdots + \mathbf{Var}[X_n - \mathbf{E}[X_n]]\right) \\
&= \frac{1}{n^2}\left(\mathbf{Var}[X_1] + \cdots + \mathbf{Var}[X_n]\right) \\
&\leq \frac{V}{n}.
\end{aligned}$$

And, by Chebyshev's inequality,

$$\mathbf{P}\left[\left|\frac{1}{n}\tilde{S}_n\right| > \epsilon\right] \leq \frac{\mathbf{Var}\left[\frac{1}{n}\tilde{S}_n\right]}{\epsilon^2} \leq \frac{V}{\epsilon^2 n}.$$

□

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

Proof. Exercise. □

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

1. Show, using Chebyshevs inequality, that for any $\epsilon > 0$

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

2. 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 $X_1, \dots, X_n$ be independent real random variables and let S_n be their sum. Let $M_n = \mathbf{E}[X_1] + \dots + \mathbf{E}[X_n]$ 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 $p_1, \dots, p_n \in [0,1]$. Let $X_1, \dots, X_n$ be independent random variables with $X_i \sim \mathrm{Ber}_{p_i}$ for any $i = 1, \dots, n$. Define $S_n := X_1 + \dots + X_n$ 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
$$\mathbf{P}[S_n \leq (1-\delta)m] \leq e^{-\frac{\delta^2 m}{2}}.$$