PROBABILITY THEORY

MATH 154

Unit 15: Strong law of large numbers

15.1. Laws of large numbers make statements about the stochastic convergence of sums $\frac{S_n}{n} = \frac{X_1+\cdots+X_n}{n}$ of random variables X_n. The weak law dealt with convergence in probability. The strong laws of large numbers make analog statements about almost everywhere convergence.

15.2. The first version of the strong law does not assume the random variables to have the same distribution. They are assumed to have the same expectation and have to be bounded in $\mathcal{L}^4$. We only need decorrelation.

Theorem 1. *Assume X_n are pairwise decorrelated random variables in $\mathcal{L}^4$ with common expectation $\mathrm{E}[X_n] = m$ and for which $M = \sup_n ||X_n||_4^4 < \infty$. Then $S_n/n \to m$ almost everywhere.*

Proof. In the proof of the weak law of large numbers, we derived from these assumptions $\mathrm{P}[|\frac{S_n}{n} - m| \geq \epsilon] \leq 2M\frac{1}{\epsilon^4 n^2}$. This means that $S_n/n \to m$ converges completely. But complete convergence implies almost everywhere convergence. □

15.3. The strong law for IID $\mathcal{L}^1$ random variables was first proven by Kolmogorov in 1930 under the assumption of independence. Much later, in 1981, it has been observed that the weaker notion of **pairwise independence** is sufficient.

Theorem 2 (Etemadi)**.** *Assume $X_n \in \mathcal{L}^1$ are pairwise independent and identically distributed random variables with mean m. Then $S_n/n \to m$ almost everywhere.*

Proof. (0) We can assume without loss of generality that $\boxed{X_n \geq 0}$ because we can split $X_n = X_n^+ + X_n^-$ into its positive $X_n^+ = X_n \vee 0 = \max(X_n, 0)$ and negative part $X^- = -X \vee 0 = \max(-X, 0)$ and because knowing the result for $X_n^\pm$ implies it for X_n. Define $f_R(t) = t \cdot 1_{[-R,R]}$, the random variables $X_n^{(R)} = f_R(X_n)$ and $Y_n = X_n^{(n)}$ as well as $S_n = \frac{1}{n}\sum_{i=1}^n X_i$, $T_n = \frac{1}{n}\sum_{i=1}^n Y_i$.
(i) $\boxed{T_n - \mathrm{E}[T_n] \to 0 \text{ implies } S_n - \mathrm{E}[S_n] \to 0}$.
Proof. Since $\mathrm{E}[Y_n] \to m$, we have $\mathrm{E}[T_n] \to m$. Because of (0) $\sum_{n\geq 1} \mathrm{P}[Y_n \neq X_n] \leq \sum_{n\geq 1} \mathrm{P}[X_n \geq n] = \sum_{n\geq 1} \mathrm{P}[X_1 \geq n] = \sum_{n\geq 1}\sum_{k\geq n} \mathrm{P}[X_n \in [k, k+1]]$
$= \sum_{k\geq 1} k \cdot \mathrm{P}[X_1 \in [k, k+1]] \leq \mathrm{E}[X_1] = m < \infty$, we get by the first Borel-Cantelli lemma that $\mathrm{P}[Y_n \neq X_n, \text{ infinitely often}] = 0$. This means $T_n - S_n \to 0$ almost everywhere, proving $\mathrm{E}[S_n] \to m$ if $\mathrm{E}[T_n] \to m$.
(ii) $\boxed{\text{Complete convergence along subsequence.}}$ Fix a real number $\alpha > 1$ and define an exponentially growing subsequence $k_n = [\alpha^n]$ which is the **integer part** of α^n. Denote

by μ the law of the random variables X_n. For every $\epsilon > 0$, we get (using Chebyshev's inequality using pairwise independence) for $k_n = [\alpha^n]$ some constants C which can vary in each step: $\sum_{n=1}^\infty \mathrm{P}[|T_{k_n} - \mathrm{E}[T_{k_n}]| \geq \epsilon] \leq \sum_{n=1}^\infty \frac{\mathrm{Var}[T_{k_n}]}{\epsilon^2} = \sum_{n=1}^\infty \frac{1}{\epsilon^2 k_n^2} \sum_{m=1}^{k_n} \mathrm{Var}[Y_m]$ $= \frac{1}{\epsilon^2} \sum_{m=1}^\infty \mathrm{Var}[Y_m] \sum_{n, k_n \geq m} \frac{1}{k_n^2} \leq^{(1)} \frac{1}{\epsilon^2} \sum_{m=1}^\infty \mathrm{Var}[Y_m] \, \frac{C}{m^2} \leq C \sum_{m=1}^\infty \frac{1}{m^2} \mathrm{E}[Y_m^2]$. In (1) we used that with $k_n = [\alpha^n]$ one has $\sum_{n: k_n \geq m} k_n^{-2} \leq C \cdot m^{-2}$. In the last step we used that $\mathrm{Var}[Y_m] = \mathrm{E}[Y_m^2] - \mathrm{E}[Y_m]^2 \leq \mathrm{E}[Y_m^2]$. After catching breath, we continue:

$$\begin{aligned}
\sum_{n=1}^\infty \mathrm{P}[|T_{k_n} - \mathrm{E}[T_{k_n}]| \geq \epsilon] &\leq C \sum_{m=1}^\infty \frac{1}{m^2} \mathrm{E}[Y_m^2] \\
&\leq C \sum_{m=1}^\infty \frac{1}{m^2} \sum_{l=0}^{m-1} \int_l^{l+1} x^2 \, d\mu(x) \\
&= C \sum_{l=0}^\infty \sum_{m=l+1}^\infty \frac{1}{m^2} \int_l^{l+1} x^2 \, d\mu(x) \\
&\leq C \sum_{l=0}^\infty \sum_{m=l+1}^\infty \frac{(l+1)}{m^2} \int_l^{l+1} x \, d\mu(x) \\
&\leq^{(2)} C \sum_{l=0}^\infty \int_l^{l+1} x \, d\mu(x) \\
&\leq C \cdot \mathrm{E}[X_1] < \infty \, .
\end{aligned}$$

In (2) we used that $\sum_{m=l+1}^n m^{-2} \leq C \cdot (l+1)^{-1}$.
We have now proved complete convergence. As before, this implies the almost everywhere convergence of $T_{k_n} - \mathrm{E}[T_{k_n}] \to 0$.
(iii) $\boxed{\text{General case.}}$ Convergence has been verified along a subsequence k_n. Because we assumed $X_n \geq 0$, the sequence $U_n = \sum_{i=1}^n Y_n = nT_n$ is monotonically increasing. For $n \in [k_m, k_{m+1}]$, we get therefore $\frac{k_m}{k_{m+1}} \frac{U_{k_m}}{k_m} = \frac{U_{k_m}}{k_{m+1}} \leq \frac{U_n}{n} \leq \frac{U_{k_{m+1}}}{k_m} = \frac{k_{m+1}}{k_m} \frac{U_{k_{m+1}}}{k_{m+1}}$ and from $\lim_{n\to\infty} T_{k_m} = \mathrm{E}[X_1]$ almost everywhere, the statement $\frac{1}{\alpha}\mathrm{E}[X_1] \leq \liminf_n T_n \leq \limsup_n T_n \leq \alpha \mathrm{E}[X_1]$ follows. □

15.4. The strong law of large numbers can be interpreted as a statement about the growth of the sequence $\sum_{k=1}^n X_n$. For $\mathrm{E}[X_1] = 0$, the convergence $\frac{1}{n}\sum_{k=1}^n X_n \to 0$ means that for all $\epsilon > 0$ there exists m such that for $n > m$

$$|\sum_{k=1}^n X_n| \leq \epsilon n \, .$$

This means that the trajectory $\sum_{k=1}^n X_n$ is finally contained in any arbitrary narrow cone. In other words, it grows slower than linear. The exact description for the growth of $\sum_{k=1}^n X_n$ is given by the **law of the iterated logarithm of Khinchin** which says that a sequence of IID random variables X_n with $\mathrm{E}[X_n] = m$ and $\sigma(X_n) = \sigma \neq 0$ satisfies with $\Lambda_n = \sqrt{2\sigma^2 n \log\log n}$:

$$\limsup_{n\to\infty} \frac{S_n}{\Lambda_n} = +1, \liminf_{n\to\infty} \frac{S_n}{\Lambda_n} = -1 \, .$$