PROBABILITY THEORY

MATH 154

Unit 14: Weak law of large numbers

14.1. A sequence of random variables X_i is also called a **stochastic process**. We often deal with sums $S_n = X_1 + X_2 + \cdots + X_n$ and especially the time averages S_n/n. For example, if X_i is the outcome of a dice, then S_n/n is the average of all the dice outcomes. We of course know what this average should be. Experience shows that it is the average of the distribution which is $m = (1+2+3+4+5+6)/6 = 21/6 = 3.5$.

14.2. The weak law of large numbers holds for pairwise uncorrelated random variables. This is a remarkably weak assumption.

Theorem 1. *Assume $X_i \in \mathcal{L}^2$ are pairwise uncorrelated, have a common mean $\mathrm{E}[X_i] = m$ and $M = \sup_n \frac{1}{n} \sum_{i=1}^n \mathrm{Var}[X_i] < \infty$. Then $\frac{S_n}{n} \to m$ in probability.*

Proof. Since $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2 \cdot \mathrm{Cov}[X,Y]$ and X_n are pairwise uncorrelated, we get $\mathrm{Var}[X_n + X_m] = \mathrm{Var}[X_n] + \mathrm{Var}[X_m]$ and by induction $\mathrm{Var}[S_n] = \sum_{i=1}^n \mathrm{Var}[X_n]$. Using linearity, we obtain $\mathrm{E}[S_n/n] = m$ and

$$\mathrm{Var}[\frac{S_n}{n}] = \mathrm{E}[\frac{S_n^2}{n^2}] - \frac{\mathrm{E}[S_n]^2}{n^2} = \frac{\mathrm{Var}[S_n]}{n^2} = \frac{1}{n^2} \sum_{i=1}^n \mathrm{Var}[X_n] \; .$$

The right hand side converges to zero for $n \to \infty$. With Chebychev's inequality we obtain

$$\mathrm{P}[|\frac{S_n}{n} - m| \geq \epsilon] \leq \frac{\mathrm{Var}[\frac{S_n}{n}]}{\epsilon^2} \leq \frac{M}{n\epsilon^2} \; .$$

□

14.3. As an application in analysis, this leads to a constructive proof of a **theorem of Weierstrass** which states that polynomials are dense in the space $C[0,1]$ of all continuous functions on the interval $[0,1]$.

Theorem 2. *For every $f \in C[0,1]$, the* **Bernstein polynomials**

$$B_n(x) = \sum_{k=1}^n f(\frac{k}{n}) \binom{n}{k} x^k (1-x)^{n-k}$$

converge uniformly to f. If $f(x) \geq 0$, then also $B_n(x) \geq 0$.

Proof. For $x \in [0,1]$, let X_n be a sequence of independent $\{0,1\}$- valued random variables with mean value x. In other words, we take the probability space $(\{0,1\}^{\mathbb{N}}, \mathcal{A}, \mathrm{P})$

defined by $P[\omega_n = 1] = x$. Since $P[S_n = k] = \binom{n}{k} x^k (1-p)^{n-k}$, we can write $B_n(x) = E[f(\frac{S_n}{n})]$. We estimate with $||f|| = \max_{0 \leq x \leq 1} |f(x)|$

$$\begin{aligned} |B_n(x) - f(x)| &= |E[f(\frac{S_n}{n})] - f(x)| \leq E[|f(\frac{S_n}{n}) - f(x)|] \\ &\leq 2||f|| \cdot P[|\frac{S_n}{n} - x| \geq \delta] \\ &+ \sup_{|x-y| \leq \delta} |f(x) - f(y)| \cdot P[|\frac{S_n}{n} - x| < \delta] \\ &\leq 2||f|| \cdot P[|\frac{S_n}{n} - x| \geq \delta] + \sup_{|x-y| \leq \delta} |f(x) - f(y)| \, . \end{aligned}$$

The second term in the last line is called the **continuity module** of f. It converges to zero for $\delta \to 0$. By the Chebychev inequality and the proof of the weak law of large numbers, the first term can be estimated from above by

$$2||f|| \frac{\mathrm{Var}[X_i]}{n\delta^2} \, ,$$

a bound which goes to zero for $n \to \infty$ because the variance satisfies $\mathrm{Var}[X_i] = x(1-x) \leq 1/4$. □

14.4. In the weak law of large numbers, we only assumed the random variables to be uncorrelated. Under the stronger condition of independence and the moment assumption $X^4 \in \mathcal{L}^1$, the convergence can be accelerated:

Theorem 3. *Assume $X_i \in \mathcal{L}^4$ have common expectation $E[X_i] = m$ and satisfy $M = \sup_n ||X||_4 < \infty$. If X_i are independent, then $S_n/n \to m$ in probability. Even $\sum_{n=1}^{\infty} P[|\frac{S_n}{n} - m| \geq \epsilon]$ converges for all $\epsilon > 0$.*

Proof. We can assume without loss of generality that $m = 0$. Because the X_i are independent, we get

$$E[S_n^4] = \sum_{i_1,i_2,i_3,i_4=1}^{n} E[X_{i_1} X_{i_2} X_{i_3} X_{i_4}] \, .$$

Again by independence, a summand $E[X_{i_1} X_{i_2} X_{i_3} X_{i_4}]$ is zero if an index $i = i_k$ occurs alone, it is $E[X_i^4]$ if all indices are the same and $E[X_i^2]E[X_j^2]$, if there are two pairwise equal indices. Since by Jensen's inequality $E[X_i^2]^2 \leq E[X_i^4] \leq M$ we get

$$E[S_n^4] \leq nM + n(n-1)M \, .$$

Use now the Chebyshev-Markov inequality with $h(x) = x^4$ to get

$$\begin{aligned} P[|\frac{S_n}{n}| \geq \epsilon] &\leq \frac{E[(S_n/n)^4]}{\epsilon^4} \\ &\leq M \frac{n + n^2}{\epsilon^4 n^4} \leq 2M \frac{1}{\epsilon^4 n^2} \, . \end{aligned}$$

□