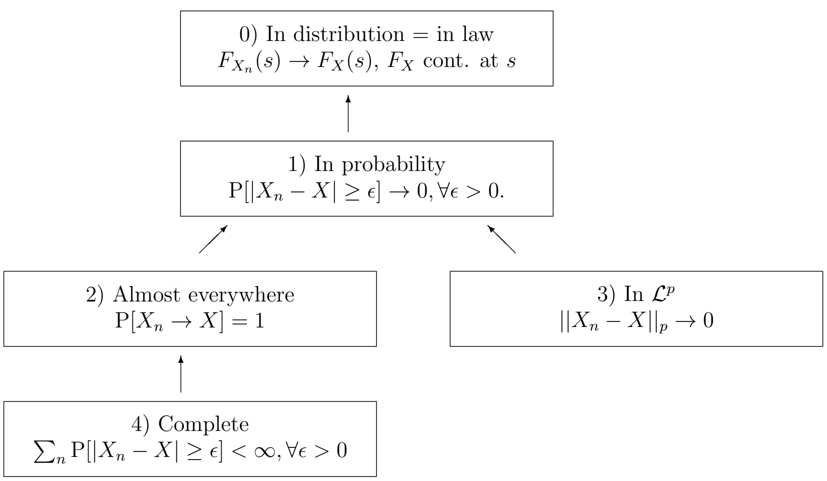
PROBABILITY THEORY

MATH 154

Unit 13: Stochastic convergence

13.1. A sequence of random variables X_n on $(\Omega, \mathcal{A}, \mathrm{P})$ is called a **stochastic process**. We say X_n converges **in probability** to X, if $\mathrm{P}[|X_n - X| \geq \epsilon] \to 0$ for all $\epsilon > 0$. We say X_n converges **almost everywhere** or **almost surely** to X if $\mathrm{P}[X_n \to X] = 1$. We say X_n **converges in** $\mathcal{L}^p$ to X, if $||X_n - X||_p \to 0$ for $n \to \infty$. Finally, X_n converges **completely** if $\sum_n \mathrm{P}[|X_n - X| \geq \epsilon] < \infty$ for all $\epsilon > 0$.

13.2. A sequence X_n of random variables X_n on $(\Omega_n, \mathcal{A}_n, \mathrm{P}_n)$ converges **in distribution**, if $F_{X_n}(s) \to F_X(s)$ at all continuity points s of F_X. We say X_n converges **in law** to X, if the laws μ_n of X_n converge weakly to the law μ of X meaning that for every continuous function f on $\mathbb{R}$ of compact support, one has $\int f(x)\, d\mu_n(x) \to \int f(x)\, d\mu(x)$.



Proof. $\boxed{2) \Rightarrow 1)}$: $\{X_n \to X\} = \bigcap_k \bigcup_m \bigcap_{n \geq m} \{|X_n - X| \leq 1/k\}$ is equivalent to $1 = \mathrm{P}[\bigcup_m \bigcap_{n\geq m}\{|X_n - X| \leq \frac{1}{k}\}] = \lim_{m\to\infty} \mathrm{P}[\bigcap_{n\geq m}\{|X_n - X| \leq \frac{1}{k}\}]$ for all k and so $0 = \lim_{n\to\infty} \mathrm{P}[\bigcup_{n\geq m}\{|X_n - X| \geq \frac{1}{k}\}]$ for all k. Therefore $\mathrm{P}[|X_m - X| \geq \epsilon] \leq \mathrm{P}[\bigcup_{n\geq m}\{|X_n - X| \geq \epsilon\}] \to 0$ for all $\epsilon > 0$. $\boxed{4) \Rightarrow 2)}$: The first Borel-Cantelli lemma implies that for all $\epsilon > 0$ $\mathrm{P}[\{|X_n - X| \geq \epsilon$, infinitely often$\}] = 0$. We get so for $\epsilon_k \to 0$ the relation $\mathrm{P}[B_k] = \mathrm{P}[\bigcup_n \{|X_n - X| \geq \epsilon_k$, infinitely often$\}] \leq \sum_n \mathrm{P}[\{|X_n - X| \geq \epsilon_k$, infinitely often$\}] = 0$ and $\bigcup_k B_k$ has measure zero. Now $\mathrm{P}[X_n \to X] = 1 - \mathrm{P}[\bigcup_k B_k] = 1 - 0 = 1$. $\boxed{3) \Rightarrow 1)}$: Chebychev-Markov implies $\mathrm{P}[|X_n - X| \geq \epsilon] \leq \frac{\mathrm{E}[|X_n - X|^p]}{\epsilon^p}$. $\boxed{1) \Rightarrow 0)}$: $\mathrm{P}[X_n \leq c] \leq \mathrm{P}[X \leq c + \epsilon] + \mathrm{P}[|X_n - X| > \epsilon]$. □

Theorem 1. *Given $X_n \in \mathcal{L}^\infty$ with $||X_n||_\infty \leq K$ for all n, then $X_n \to X$ in probability if and only if $X_n \to X$ in $\mathcal{L}^1$.*

Proof. (i) For $k \in \mathbb{N}$, $\mathrm{P}[|X| > K + \frac{1}{k}] \leq \mathrm{P}[|X - X_n| > \frac{1}{k}] \to 0, n \to \infty$ so that $\mathrm{P}[|X| > K + \frac{1}{k}] = 0$. Therefore $\mathrm{P}[|X| > K] = \mathrm{P}[\bigcup_k \{|X| > K + \frac{1}{k}\ \}] = 0$. (ii) Given $\epsilon > 0$. Choose m such that $\mathrm{P}[|X_n - X| > \frac{\epsilon}{3}] < \frac{\epsilon}{3K}$ for all $n > m$. Use the notation $\mathrm{E}[X; A] = \mathrm{E}[X \cdot 1_A]$. By (i) we have $\mathrm{E}[|X_n - X|] = \mathrm{E}[(|X_n - X|; |X_n - X| > \frac{\epsilon}{3}] + \mathrm{E}[(|X_n - X|; |X_n - X| \leq \frac{\epsilon}{3}] \leq 2K\mathrm{P}[|X_n - X| > \frac{\epsilon}{3}] + \frac{\epsilon}{3} \leq \epsilon$. □

13.3. A family $\mathcal{C} \subset \mathcal{L}^1$ of random variables is called **uniformly integrable**, if

$$\lim_{R \to \infty} \sup_{X \in \mathcal{C}} \mathrm{E}[X; |X| > R] = 0$$

for all $X \in \mathcal{C}$ still using notation $\mathrm{E}[X; A] = \mathrm{E}[X 1_A]$.

Theorem 2. *Given $X \in \mathcal{L}^1$ and $\epsilon > 0$. There exists $K \geq 0$ with* $\mathrm{E}[|X|; |X| > K] < \epsilon$.

Proof. Assume we are given $\epsilon > 0$. If $X \in \mathcal{L}^1$, we can find $\delta > 0$ such that if $\mathrm{P}[A] < \delta$, then $\mathrm{E}[|X|; A] < \epsilon$. Since $K\mathrm{P}[|X| > K] \leq \mathrm{E}[|X|]$, we can choose K such that $\mathrm{P}[|X| > K] < \delta$. Therefore $\mathrm{E}[|X|; |X| > K] < \epsilon$. □

13.4. The next proposition gives a necessary and sufficient condition for $\mathcal{L}^1$ convergence.

Theorem 3. *Given $X_n \in \mathcal{L}^1$ and $X \in \mathcal{L}^1$. The following is equivalent:*
a) X_n converges in probability to X and $\{X_n\}_{n \in \mathbb{N}}$ is uniformly integrable.
b) X_n converges in $\mathcal{L}^1$ to X.

Proof. $a) \Rightarrow b)$. For any random variable X and $K \geq 0$ define the bounded variable $X^{(K)} = X \cdot 1_{\{-K \leq X \leq K\}} + K \cdot 1_{\{X > K\}} - K \cdot 1_{\{X < -K\}}$. By the uniform integrability condition and the previous theorem applied to $X^{(K)}$ and X, we can choose K such that for all n, $\mathrm{E}[|X_n^{(K)} - X_n|] < \frac{\epsilon}{3}$, $\mathrm{E}[|X^{(K)} - X|] < \frac{\epsilon}{3}$. Since $|X_n^{(K)} - X^{(K)}| \leq |X_n - X|$, we have $X_n^{(K)} \to X^{(K)}$ in probability. By the last theorem we know $X_n^{(K)} \to X^{(K)}$ in $\mathcal{L}^1$ so that for $n > m$ $\mathrm{E}[|X_n^{(K)} - X^{(K)}|] \leq \epsilon/3$. Therefore, for $n > m$ also

$$\mathrm{E}[|X_n - X|] \leq \mathrm{E}[|X_n - X_n^{(K)}|] + \mathrm{E}[|X_n^{(K)} - X^{(K)}|] + \mathrm{E}[|X^{(K)} - X|] \leq \epsilon\ .$$

$b) \Rightarrow a)$. We have seen already that $X_n \to X$ in probability if $||X_n - X||_1 \to 0$. We have to show that $X_n \to X$ in $\mathcal{L}^1$ implies that X_n is uniformly integrable.
Given $\epsilon > 0$. There exists m such that $\mathrm{E}[|X_n - X|] < \epsilon/2$ for $n > m$. By the absolutely continuity property, we can choose $\delta > 0$ such that $\mathrm{P}[A] < \delta$ implies

$$\mathrm{E}[|X_n|; A] < \epsilon, 1 \leq n \leq m, \mathrm{E}[|X|; A] < \epsilon/2\ .$$

Because X_n is bounded in $\mathcal{L}^1$, we can choose K such that $K^{-1} \sup_n \mathrm{E}[|X_n|] < \delta$ which implies $\mathrm{P}[|X_n| > K] < \delta$. For $n \geq m$, we have therefore, using the notation $\mathrm{E}[X; A] = \mathrm{E}[X \cdot 1_A]$

$$\mathrm{E}[|X_n|; |X_n| > K] \leq \mathrm{E}[|X|; |X_n| > K] + \mathrm{E}[|X - X_n|] < \epsilon\ .$$

□

Oliver Knill, knill@math.harvard.edu, Math 154, Spring, 2026