CONVERGENCE FOR RANDOM VARIABLES

domenica 24 febbraio 2019 10:42

Most of the results in Probability theory deal with LIMIT THEOREMS, i.e., the asymptotic behaviour of random processes. Therefore studying the convergence of random variables becomes necessary. We begin by recalling some definitions pertaining to convergence of random variables.

NOTIONS OF CONVERGENCE FOR R.V.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{X_n\}_{n\in\mathbb{N}}$ be a sequence of r.v. defined on this probability space.

Def: (Almost sure convergence)

A sequence of r.v. $\{X_n\}_{n\in\mathbb{N}}$ is said to converge almost surely (or with probability 1) to X if

$$\mathbb{P}\left(X_n \xrightarrow[n\to\infty]{} X\right) = \mathbb{P}\left(\left\{\omega\in\Omega \mid X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)\right\}\right) = 1$$

In this case we write $X_n \xrightarrow{a.s.} X$.

Remark: This def. gives the random variables "freedom" not to converge on a set of zero measure!

Remark 2: Assume that $X_n \xrightarrow{a.s.} X$ and A is the set of prob. 1 s.t. $\forall \omega\in A$, $X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega)$. Then $\forall \varepsilon>0$, $\forall \omega\in\Omega$, $\exists$ $N(\omega)>0$ s.t.

dependend on ω!!!

$$\forall n\geq N(\omega), \quad |X_n(\omega) - X(\omega)| \leq \varepsilon$$

⚠ **ATTENTION:** $N(\omega)$ depends on ω and in general it's not true that $\forall\varepsilon>0$, $\exists$ a universal N s.t. $\forall \omega\in\Omega$ and $\forall n>N$ $|X_n(\omega)-X(\omega)|<\varepsilon$.

Def: (Convergence in probability)

A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ is said to converge in

A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ is said to <u>converge in probability</u> to X if $\forall \varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(|X_n - X| > \varepsilon \right) = 0.$$

In this case we write $X_n \xrightarrow{\mathbb{P}} X$.

Remark 3: At the first glance, it may seem that these two first notions of convergence are very similar, but the two definitions actually tell very different stories:

- For "a.s. convergence" we have a condition on the probability of a single event,
- For "convergence in probability" we have a condition on a sequence of probabilities.

Def: (convergence in L^p, $p\geq 1$)

A sequence of random variables $\{X_n\}_{n\in\mathbb{N}}$ is said to converge to X in L^p if

$$\lim_{n\to\infty} \mathbb{E}\left[|X_n - X|^p\right] = 0.$$

In this case we write $X_n \xrightarrow{L^p} X$.

Remark 4: L^p-convergence is a convergence for <u>integrals</u>!!

For the last notion of convergence we first consider a sequence $\{\mu_n\}_{n\in\mathbb{N}}$ of <u>probability measures</u> on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Def: (weak convergence)

A sequence of probability measures $\{\mu_n\}_{n\in\mathbb{N}}$ is said to <u>weak-converge</u> to a probability measure μ if

$$\lim_{n\to\infty} \int f \, d\mu_n = \int f \, d\mu \quad , \text{ for all } f: \mathbb{R} \to \mathbb{R} \text{ continuous \& bounded}$$

In this case we write $\mu_n \xrightarrow{w} \mu$.

Let now $\{Y_n\}_{n\in\mathbb{N}}$, Y be real-valued random variables not necessarily defined on the same probability space. Let $\{\mathcal{L}(Y_n)\}_{n\in\mathbb{N}}$, $\mathcal{L}(Y)$ denote the laws of the random variables $\{Y_n\}_{n\in\mathbb{N}}$, Y.

Def: (convergence in distribution)

A sequence of random variables $\{Y_n\}_{n\in\mathbb{N}}$ is said to <u>converge in distribution</u> to Y if $\mathcal{L}(Y_n) \xrightarrow{w} \mathcal{L}(Y)$, that is, if

$$\lim_{n\to\infty} \mathbb{E}[f(Y_n)] = \mathbb{E}[f(Y)], \quad \text{for all } f:\mathbb{R}\to\mathbb{R} \text{ } \underline{\text{continuous}} \text{ \& } \underline{\text{bounded}}.$$

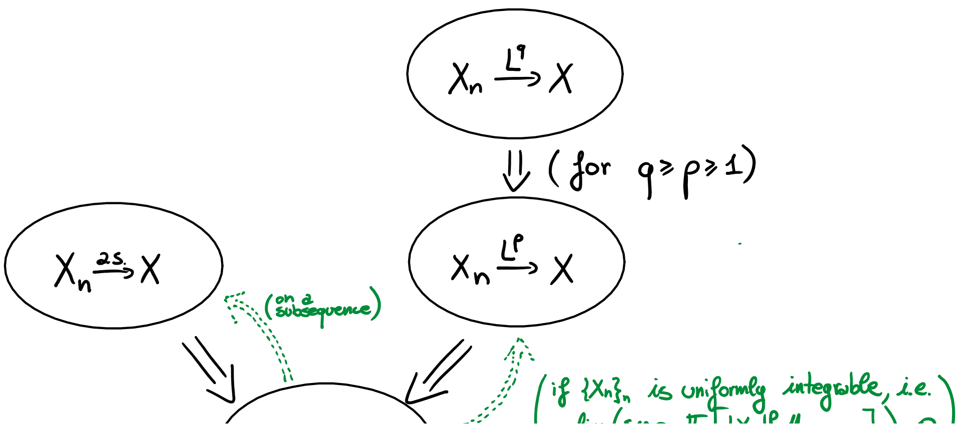
$$\mathbb{E}[f(Y_n)] = \int f \, d\mathcal{L}(Y_n) \qquad \mathbb{E}[f(Y)] = \int f \, d\mathcal{L}(Y)$$

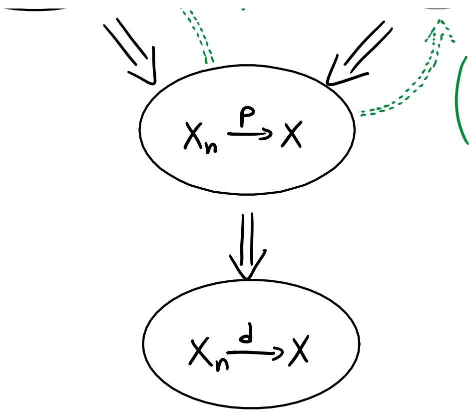
In this case we write $Y_n \xrightarrow{d} Y$.

Remark 5: The notion of convergence in distribution involves <u>only</u> the <u>laws</u> of the random variables. Recall that given a real-r.v. X then $\mathcal{L}(X)$ is a probability measure <u>defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$</u> defined as $\mathcal{L}(X)(A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}(\{\omega\in\Omega \mid X(\omega)\in A\})$.

Therefore we DO <u>NOT</u> HAVE TO ASSUME that the considered random variables $\{Y_n\}_{n\in\mathbb{N}}$, Y leaves in the same probability space!

RELATIONS AMONG VARIOUS NOTIONS OF CONVERGENCE





(if $\{X_n\}_n$ is uniformly integrable, i.e. $\lim_{a \to \infty} \left(\sup_n \mathbb{E}\left[|X_n|^p \mathbb{1}_{|X_n| > a} \right] \right) = 0$)

(⊛ The proof for the green implications is omitted.)

Before proving the "black implications", we recall the following important three results.

Theorem:

(a) <u>MONOTONE CONVERGENCE</u>: Let $X_n \in L^1$, $X_n \leq X_{n+1}$ a.s. and $X_n \xrightarrow{a.s.} X \in L^1$. then $\mathbb{E}[X_n] \xrightarrow[n \to \infty]{} \mathbb{E}[X]$.

(b) <u>FATOU'S LEMMA</u>: $\{X_n\}_{n \in \mathbb{N}}$ non-negative r.v. $X = \liminf_{n \to \infty} X_n$ then $\mathbb{E}[X] \leq \liminf_{n \to \infty} \mathbb{E}[X_n]$

(c) <u>DOMINATED CONVERGENCE</u>: $X_n \in L^1$, $X_n \xrightarrow{a.s.} X$ and $\exists\ Y \in L^1$ s.t. $|X_n| \leq Y$ a.s. then $X \in L^1$ and $\mathbb{E}[X_n] \xrightarrow[n \to \infty]{} \mathbb{E}[X]$.

Proof: Omitted. $\square$

We can now prove the four implications.

Theorem: $X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{P} X$

Proof: Fix $\varepsilon > 0$. Let $Y_n = \mathbb{1}_{\{|X_n - X| > \varepsilon\}}$. Since $X_n \xrightarrow{a.s.} X$ then $Y_n \xrightarrow{a.s.} 0$ and $|Y_n| \leq 1$. By dominated convergence $\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{E}[Y_n] \xrightarrow[n \to \infty]{} 0$. $\square$

Theorem: $X_n \xrightarrow{L^p} X \implies X_n \xrightarrow{P} X$

Proof: By Markov's inequality

$$\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{P}\left(|X_n - X|^p > \varepsilon^p \right) \leq \frac{\mathbb{E}\left[|X_n - X|^p \right]}{\varepsilon^p} \xrightarrow[n \to \infty]{} 0 \quad \left(\text{since } X_n \xrightarrow{L^p} X\right). \quad \square$$

Theorem: $q \geq p \geq 1$. $X_n \xrightarrow{L^q} X \Rightarrow X_n \xrightarrow{L^p} X$.

Proof: Simple application of Lyapunov's inequality, i.e.,

$$\mathbb{E}\left[|X_n - X|^p\right]^{1/p} \leq \mathbb{E}\left[|X_n - X|^q\right]^{1/q}$$ $\square$

Theorem: $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$.

Proof: Recall that for a real sequence $\{a_n\}_{n\in\mathbb{N}}$, $a_n \xrightarrow[n\to\infty]{} a \in \mathbb{R}$ if and only if $|a_n|$ is bounded and every convergent subsequence $\{a_{n_k}\}_k$ has limit a.

Let $f: \mathbb{R} \to \mathbb{R}$ cont. & bounded. $a_n = \mathbb{E}[f(X_n)]$ and $a = \mathbb{E}[f(X)]$.

1. a_n is bounded: $|a_n| \leq \|f\|_\infty$
2. Assume that $a_{n_k} \to b$. Since $X_{n_k} \xrightarrow{P} X$ then there exists a subsequence $X_{n_{k_\ell}} \xrightarrow{a.s.} X$ and so $f(X_{n_{k_\ell}}) \xrightarrow{a.s.} f(X)$ since f is continuous.

   Since $|f(X_{n_{k_\ell}})| \leq \|f\|_\infty$, by dominated convergence:

   $$a_{n_{k_\ell}} = \mathbb{E}\left[f(X_{n_{k_\ell}})\right] \xrightarrow[\ell\to\infty]{} \mathbb{E}\left[f(X)\right] = a$$

   and so $b = a$. $\square$

We conclude this section with a list of counterexample.

- For $\left(X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{L^p} X\right)$ take $X_n = \begin{cases} n^3 & \text{w.p. } 1/n^2 \\ 0 & \text{w.p. } 1 - 1/n^2 \end{cases}$ independent
- For $\left(X_n \xrightarrow{d} X \not\Rightarrow X_n \xrightarrow{P} X\right)$ take $Y \sim \text{Bern}(\frac{1}{2})$, $X_n = Y \ \forall n$, $X = 1 - Y$.
- For $\left(X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{a.s.} X\right)$ take $X_n = \begin{cases} 1 & \text{w.p. } 1/n \\ 0 & \text{w.p. } 1 - 1/n \end{cases}$ independent.
- For $\left(X_n \xrightarrow{a.s.} X \not\Rightarrow X_n \xrightarrow{L^2} X\right)$ take $(\Omega, \mathcal{A}, \mathbb{P}) = ([0,1], \mathcal{B}([0,1]), \underbrace{\lambda}_{\text{Lebesgue measure}})$ & $X_n(\omega) = \begin{cases} n, & \text{if } \omega \in [0, \frac{1}{n}] \\ 0, & \text{otherwise.} \end{cases}$

Exercise: Prove the previous results. (Hint for 3: use Borel-Cantelli)
Can you generalize the 4^{th} counter-example $\forall p \geq 1$ (instead of just $p=2$)?

Exercise: Let $c \in \mathbb{R}$, $X_n \xrightarrow{d} c$, X_n defined on the same prob. space.
Then $X_n \xrightarrow{P} X$.

Exercise: An important tool to prove a.s.-convergence is the Borel-Cantelli theorem. Can you prove the following?

Theorem: $X_n : \Omega \to \mathbb{R}$, $\forall n \in \mathbb{N}$, random variables and X another real-valued r.v.

$$A_n(\varepsilon) := \{ \omega \in \Omega \mid |X_n(\omega) - X(\omega)| > \varepsilon \}.$$

Then:

1. If $\sum_{n=1}^{\infty} \mathbb{P}(A_n(\varepsilon)) < \infty \Rightarrow X_n \xrightarrow{a.s.} X$

2. If $X_1, \dots, X_n, \dots$ are independent and $\sum_{n=1}^{\infty} \mathbb{P}(A_n(\varepsilon)) = \infty \Rightarrow X_n \not\xrightarrow{a.s.} X$.

SOME MORE RESULTS FOR CONVERGENCE IN DISTRIBUTION
(without proofs)

Theorem: Let $\{X_n\}_{n \in \mathbb{N}}$, X be real-valued random variables. Set

$$F_n(x) = \mathbb{P}(X_n \leq x),$$

$$F(x) = \mathbb{P}(X \leq x).$$

Then TFAE:

(a) $X_n \xrightarrow{d} X$

(b) $F_n(x) \longrightarrow F(x)$ $\forall x$ where F is continuous.

Observation: In general, it is NOT true that $X_n \xrightarrow{d} X \Rightarrow F_n(x) \to F(x)$ $\forall x \in \mathbb{R}$.

Exercise: Let $\{z_n\}_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ s.t. $z_n \downarrow z \in \mathbb{R}$. Set $X_n = z_n$ and $X = z$.
Show that $X_n \xrightarrow{d} X$ but $F_n(z) \not\longrightarrow F(z)$.

Before stating the next theorem, we recall the definition.

Def: (Characteristic function)

Def: (Characteristic function)

Let X be a real-valued r.v. Its characteristic function is defined as
$$\begin{cases} \varphi_X : \mathbb{R} \longrightarrow \mathbb{C} \\ \varphi_X(t) = \mathbb{E}\left[e^{itX}\right] \end{cases}$$

Theorem: Let X be a real-valued r.v. Then $\varphi_X(t)$ is continuous and if $\mathbb{E}[|X|^N] < \infty$ then $\varphi_X(t) \in \mathcal{C}^N$ and for all $\alpha \leq N$,
$$\frac{\partial^\alpha}{\partial t^\alpha} \varphi_X(t) = i^\alpha \, \mathbb{E}\left[X^\alpha e^{itX}\right]$$

Theorem: Let X, Y two real-valued r.v. Assume that $\varphi_X(t) = \varphi_Y(t)$, $\forall t \in \mathbb{R}$, then $X \overset{d}{=} Y$, i.e., $\mathcal{L}(X) = \mathcal{L}(Y)$.

Theorem (P. Levy)

(a) If $X_n \xrightarrow{d} X$ then $\varphi_{X_n}(t) \xrightarrow[n\to\infty]{} \varphi_X(t)$ $\forall t \in \mathbb{R}$.

(b) Assume that $\varphi_{X_n}(t) \longrightarrow f(t)$ $\forall t \in \mathbb{R}$ and that f is continuous in $t=0$. Then $f = \varphi_X$ for some r.v. X and $X_n \xrightarrow{d} X$.