Convergence of Random Variables[1]

The setting: probability space $(\Omega, \mathcal{F}, \mathbb{P})$, random variables $\xi_n = \xi_n(\omega),\ n \geq 1$, and $\xi = \xi(\omega)$, $\omega \in \Omega$.

Definitions.

(1) `Uniform convergence`: $\xi_n \rightrightarrows \xi$ if $\forall \varepsilon > 0\ \exists N = N(\varepsilon)\ \forall n > N\ \forall \omega \in \Omega : |\xi_n(\omega) - \xi(\omega)| < \varepsilon$.

(2) `Point-wise convergence`: $\xi_n \xrightarrow{p.w.} \xi$ if $\xi_n(\omega) \to \xi(\omega)$ for all $\omega \in \Omega$, that is, $\forall \varepsilon > 0\ \forall \omega \in \Omega\ \exists N = N(\varepsilon, \omega)\ \forall n > N : |\xi_n(\omega) - \xi(\omega)| < \varepsilon$.

(3) `Complete convergence`: $\xi_n \xrightarrow{c.c.} \xi$ if $\sum_{n=1}^{\infty} \mathbb{P}\big(|\xi_n - \xi| > \varepsilon\big) < +\infty$ for every $\varepsilon > 0$.

(4) `Almost sure convergence`: $\xi_n \xrightarrow{a.s.} \xi$ if $\mathbb{P}\Big(\omega : \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\Big) = 1$. Also known as `convergence with probability one`.

(5) `Convergence in` L_p, $p > 0$: $\xi_n \xrightarrow{L_p} \xi$ if $\lim_{n\to\infty} \mathbb{E}|\xi_n - \xi|^p = 0$.

(6) `Convergence in probability`: $\xi_n \xrightarrow{\mathbb{P}} \xi$ if $\lim_{n\to\infty} \mathbb{P}\big(|\xi_n - \xi| > \varepsilon\big) = 0$ for every $\varepsilon > 0$.

(7) `Convergence in distribution`: $\xi_n \xrightarrow{d} \xi$ if $\lim_{n\to\infty} \varphi_{\xi_n}(t) = \varphi_\xi(t)$ for every $t \in \mathbb{R}$; $\varphi_\xi(t) = \mathbb{E}e^{it\xi}$ is the characteristic function of ξ.

(8) `Uniform integrability`: the sequence $\{\xi_n,\ n \geq 1\}$ is uniformly integrable (UI) if

$$\lim_{a\to+\infty} \sup_{n\geq 1} \mathbb{E}\Big(|\xi_n|\, 1(|\xi_n| > a)\Big) = 0.$$

Each of the following is a sufficient condition for UI:

- $|\xi_n| \leq \eta$ for all n and $\mathbb{E}\eta < +\infty$;
- $\mathbb{E}\sup_n |\xi_n| < \infty$;
- $\sup_n \mathbb{E}g(|\xi_n|) < \infty$ for a non-negative increasing function $g = g(x),\ x \geq 0$, such that $\lim\limits_{x\to+\infty} \dfrac{g(x)}{x} = +\infty$. For example, $\sup_n \mathbb{E}|\xi_n|^r < \infty$ for some $r > 1$.

General implications:

$$\xrightarrow{p.v.} \Leftarrow \rightrightarrows \Rightarrow \xrightarrow{c.c.} \Rightarrow \xrightarrow{a.s.} \Rightarrow \xrightarrow{\mathbb{P}} \Rightarrow \xrightarrow{d}; \qquad \xrightarrow{L_p} \Rightarrow \xrightarrow{\mathbb{P}}.$$

The (non-trivial) reasons. $\rightrightarrows \Rightarrow \xrightarrow{c.c.}$ because $\rightrightarrows$ implies $\mathbb{P}\big(|\xi_n - \xi| > \varepsilon\big) = 0$ for all sufficiently large n; $\xrightarrow{c.c.} \Rightarrow \xrightarrow{a.s.}$ by the first Borel-Cantelli; $\xrightarrow{a.s.} \Rightarrow \xrightarrow{\mathbb{P}}$ because $\xrightarrow{a.s.}$ means $\lim_{n\to\infty} \mathbb{P}\big(\sup_{k\geq n} |\xi_k - \xi| > \varepsilon\big) = 0$; $\xrightarrow{\mathbb{P}} \Rightarrow \xrightarrow{d}$ by the Dominate Convergence Theorem; $\xrightarrow{L_p} \Rightarrow \xrightarrow{\mathbb{P}}$ by Markov's inequality.

Basic counterexamples, with $\Omega = [0, 1]$ and $\mathbb{P}$ the Lebesgue measure,

(1) **the Growing Pulse** $\xi_n(x) = n^\alpha 1(0 < x < n^{-\beta})$ or $\xi_n(x) = n^\alpha 1(0 \leq x < n^{-\beta})$ [if DO NOT want ξ_n to converge to zero point-wise], for suitable $\alpha > 0, \beta > 0$, shows that

- $\xrightarrow{p.v.}$ or $\xrightarrow{a.s.} \not\Rightarrow \xrightarrow{L_p}$ or $\xrightarrow{c.c.}$
- $\xrightarrow{c.c.} \not\Rightarrow \rightrightarrows$ or $\xrightarrow{p.w.}$

(2) **the Typewriter Sequence** $\xi_{n,k} = 1\big(k2^{-n} < x < (k+1)2^{-n}\big)$, $k = 0, \ldots, 2^n - 1$, $n = 0, 1, 2, \ldots$, shows that

$$\xrightarrow{L_p},\quad 0 < p < \infty \quad \not\Rightarrow \quad \xrightarrow{a.s.}$$

Special implications

(1) $\xrightarrow{\mathbb{P}} \Rightarrow \xrightarrow{a.s.}$ if the sequence $\{\xi_n,\ n \geq 1\}$ is monotone;

(2) $\xrightarrow{d} \Rightarrow \xrightarrow{\mathbb{P}}$ if the limit ξ is non-random: there exists a real number c such that $\mathbb{P}(\xi = c) = 1$;

(3) $\xrightarrow{d} \Rightarrow \xrightarrow{a.s.}$ for sums of independent random variables [using Cauchy criterion and zero-one law];

---

[1]Sergey Lototsky, USC

(4) $\xrightarrow{\mathbb{P}} \Rightarrow \xrightarrow{L_p}$ if the sequence $\{|\xi_n|^p,\ n \geq 1\}$ is UI.
(5) $\xrightarrow{a.s.} \Rightarrow \xrightarrow{c.c.}$ if the random variables ξ_n are independent AND the limit ξ is non-random [by the second Borel-Cantelli; if the limit ξ is random, then the random variables $\xi_n - \xi$ might be dependent and the second Borel-Cantelli will not apply];
(6) If ξ_n are iid and $\mathbb{E}|\xi_n| < \infty$, then $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \xi_k = \mathbb{E}\xi_1$ both with probability one and in L_1.
(7) If $\xi_n \xrightarrow{\mathbb{P}} \xi$, then $\xi_{n_k} \xrightarrow{c.c.} \xi$ along a suitable sub-sequence;
(8) If $\xi_n \xrightarrow{\mathbb{P}} \xi$ and, for some $p > 0$, $\lim\limits_{n\to\infty} \mathbb{E}|\xi_n|^p = \mathbb{E}|\xi|^p$, then $\{|\xi_n|^p,\ n \geq 1\}$ is UI and $\xi_n \xrightarrow{L_p} \xi$.

Working with limits

(1) If $M = a.s.$ or $M = \mathbb{P}$, and $\xi_n \xrightarrow{M} \xi$, $\eta_n \xrightarrow{M} \eta$, then $\xi_n \pm \eta_n \xrightarrow{M} \xi \pm \eta$, $\xi_n\eta_n \xrightarrow{M} \xi\,\eta$, and, if $\eta \neq 0$, then also $\xi_n/\eta_n \xrightarrow{M} \xi/\eta$.
(2) If $\xi_n \xrightarrow{M} \xi$, with $M = a.s.$ or $\mathbb{P}$ or d, and $f = f(x)$, $x \in \mathbb{R}$ is a continuous function, then $f(\xi_n) \xrightarrow{M} f(\xi)$. The result is obvious when $M = a.s.$, easy when $M = \mathbb{P}$, and non-trivial [known as the `Mann-Wald Theorem`] when $M = d$.
(3) `Slutsky's Theorem`: if $\xi_n \xrightarrow{d} \xi$ and $\eta_n \xrightarrow{d} b \in \mathbb{R}$ non-random, then $\xi_n \pm \eta_n \xrightarrow{d} \xi \pm b$, $\xi_n\eta_n \xrightarrow{d} b\xi$, and, if $b \neq 0$, then also $\xi_n/\eta_n \xrightarrow{d} \xi/b$.

Further Facts

- $\xrightarrow{p.w.}$ corresponds to a topology, but the topology is not metrizable;
- $\xrightarrow{\mathbb{P}}$ is metrizable: it is explained below how the function $\rho(\xi,\eta) = \mathbb{E}\frac{|\xi-\eta|}{1+|\xi-\eta|}$ defines a metric on the space of random variables, and $\xi_n \xrightarrow{\mathbb{P}} \xi$ if and only if $\lim_{n\to\infty} \rho(\xi_n, \xi) = 0$;
- $\xrightarrow{d}$ is metrizable using the `Lévy-Prokhorov metric`, which is a different story deserving a separate summary;
- $\xrightarrow{a.s.}$ is not metrizable and does not correspond to any topology because every sequence converging in probability has a subsequence converging with probability one, and there are sequences that converge in probability but not with probability one.

$\xrightarrow{\mathbb{P}}$ **is metrizable.** Consider the function $h(x) = x/(1+x)$, $x \geq 0$, so that $\rho(\xi,\eta) = \mathbb{E}h\big(|\xi - \eta|\big)$. The function satisfies $h(0) = 0$, and, for $x > 0$, $0 < h(x) < 1, h'(x) > 0, h''(x) < 0$. In particular, the derivative h' is decreasing so that $h(x) = xh'(x^*) > xh'(x)$, $0 < x^* < x$ (Mean Value Theorem), and then, for $0 < y < x$ and $z \in (y, x+y)$,

$$h(x+y) = h(y) + h'(z)x < h(y) + h'(x)x < h(y) + h(x).$$

Then the triangle inequality follows from

$$|x - z| \leq |x - y| + |y - z| \implies h(|x - z|) \leq h\big(|x - y| + |y - z|\big) \leq h(|x - y|) + h(|y - z|).$$

Similarly, for every $\varepsilon > 0$,

$$\rho(\xi,\eta) = \mathbb{E}\big(h(|\xi-\eta|)I(|\xi-\eta| \leq \varepsilon)\big) + \mathbb{E}\big(h(|\xi-\eta|)I(|\xi-\eta| > \varepsilon)\big)$$

and therefore

$$\frac{\varepsilon}{1+\varepsilon}\mathbb{P}(|\xi-\eta| > \varepsilon) \leq \rho(\xi,\eta) \leq \varepsilon + \mathbb{P}(|\xi-\eta| > \varepsilon).$$

The bottom line: there are many functions h that introduce a metric on the space of random variables; all these metrices are equivalent and metrize convergence in probability; the corresponding space of random variables is denoted by $L^0(\Omega, \mathcal{F}, \mathbb{P})$ (or L_0) and is a `Polish space`.