Lecture 6 : Convergence of random variables

STAT205 *Lecturer: Jim Pitman* Scribe: *Tye Lidman* `<tlid@berkeley.edu>`

(These notes are a revision of the work of Jin Kim, 2002.)

6.1 Convergence of random variables

First significant example: the *weak law of large numbers (WLLN)*. We want to state that with a general notion of convergence in probability.

Definition 6.1 *Given a sequence of r.v's X_n defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, say X_n* converges in probability *to X, $X_n \xrightarrow{\mathbb{P}} X$, if X is a r.v. on $(\Omega, \mathcal{F})$, and for all $\epsilon > 0$,*

$$\lim_{m \to \infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$$

Theorem 6.2 (Weak Law of Large Numbers) *Let $X, X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}|X| < \infty$. Then*

$$\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{\mathbb{P}} \mathbb{E}(X).$$

Other notions of convergence of r.v.'s:

Simplest: (discussed in previous lectures) is $\to$.

Pointwise Convergence: $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega$. This is a very strong notion: too strong for many purposes.

Almost Sure Convergence: We say $X_n \xrightarrow{a.s.} X$ if $X_n(\omega) \to X(\omega)$ for all $\omega \notin N$, with $\mathbb{P}(N) = 0$, or equivalently $\mathbb{P}(\omega : X_n(\omega) \to X(\omega) \text{ as } n \to \infty) = 1$.

Convergence in L^p $(p \geq 1)$: We say $X_n \xrightarrow{\mathbf{L}^p} X$ if $\|X_n - X\|_p \to 0$, i.e. $\lim_{n\to\infty} \mathbb{E}|X_n - X|^p = 0$.

Convergence in Distribution: (Not really a notion of convergence of r.v.) A notion of convergence of a probability distribution on $\mathbb{R}$ (or more general space). We say $X_n \xrightarrow{d} X$ if $\mathbb{P}(X_n \leq x) \to \mathbb{P}(X \leq x)$ for all x at which the RHS is continuous.

This weak convergence appears in the central limit theorem.

Fact 6.3 *(See text)* $X_n \xrightarrow{d} X \iff \mathbb{E}f(X_n) \longrightarrow \mathbb{E}f(X)$ *for all bounded and continuous function* f.

Properties in Common for $\xrightarrow{\mathbb{P}}, \xrightarrow{p.w.}, \xrightarrow{a.s.}, \xrightarrow{\mathbf{L}^p}$:

a) $X_n \to X$, $Y_n \to Y \implies X_n + Y_n \to X + Y$, $X_nY_n \to XY$.

b) $X_n \to X \iff (X_n - X) \to 0$ (useful and common reduction).

c) For all of $\xrightarrow{\mathbb{P}}$, $\xrightarrow{a.s.}$, and $\xrightarrow{\mathbf{L}^p}$ the limit X is unique up to a.s. equivalence.

d) Cauchy sequences are convergent (completeness). (Need a metric to metrize $\xrightarrow{\mathbb{P}}$, but that is easily provided. See text.)

Theorem 6.4 *The following property holds among the types of convergence.*

$$
\begin{tikzcd}
\boxed{X_n \xrightarrow{a.s.} X} \arrow[dr, Rightarrow] & & \boxed{X_n \xrightarrow{\mathbf{L}^p} X} \arrow[dl, Rightarrow, "(*)"] \\
& \boxed{X_n \xrightarrow{\mathbb{P}} X} \arrow[d, Rightarrow, "(**)"] & \\
& \boxed{X_n \xrightarrow{d} X} &
\end{tikzcd}
$$

Proof: $(*)$ can be proved by Chebyshev's inequality (with usually $p = 2$):

$$\mathbb{P}(|X_n - X| > \epsilon) \leq \frac{\mathbb{E}(|X_n - X|^p)}{\epsilon^p}$$

$(**)$ is proved in the text. ■

Example 6.5 (Moving blip) *(An example showing that almost sure convergence is a stronger condition than convergence in probability.) On* $[0, 1]$ *with Lebesgue measure, define* $X_n = 1_{(x_n, x_n+1)}$ *where the addition is interpreted as modulo* 1 *and* x_n *is any sequence with:* $x_{n+1} - x_n \to 0$ *as* $x_n \uparrow \infty$ *(e.g.* $x_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ *or* $x_n = \log n$*).* $\mathbb{P}(|X_n| > \epsilon) = X_{n+1} - X_n \to 0$ *for all* $0 < \epsilon < 1 \implies X_n \xrightarrow{\mathbb{P}} 0$*, but* X_n *does not converge almost surely to* 0.

Example 6.6 *Suppose that $X_1, X_2, \ldots$ are r.v.'s that have mean 0, have finite variances, and are uncorrelated. Let $S_n = X_1 + \cdots + X_n$. If $\sum_{k=1}^{\infty} \mathbb{E}(X_k^2) < \infty$, then S_n converges in L^2 to a limit S_∞, hence $S_n \xrightarrow{\mathbb{P}} S_\infty$, i.e. $\lim_{n\to\infty} \mathbb{P}(|S_n - S_\infty| > \epsilon) = 0$ for all $\epsilon > 0$.*

Proof: Look at the Cauchy criterion. Take $m > n$:

$$\mathbb{E}(S_m - S_n)^2 = \mathbb{E}\left(\sum_{k=n+1}^{m} X_k\right)^2 = \sum_{k=n+1}^{m} \mathbb{E}(X_k^2) \to 0$$

as $m, n \to \infty$. Therefore,

$$\sum_{k=1}^{\infty} \mathbb{E}(X_k^2) < \infty.$$

■

Fact 6.7 *If the X_n are independent (or more generally, martingale distributions), then $S_n \xrightarrow{a.s.} S_\infty$.*

The proof of this fact is deferred.

Fact 6.8 (Stout's Almost Sure Convergence) *There are examples of uncorrelated sequences with $\sum_n X_n^2 < \infty$ where a.s. convergence fails.*

6.2 Preliminaries for Study of a.s. Convergence

Definition 6.9 *Let q_n be some statement, true or false for each n. We say q_n* infinitely often *or (q_n i.o.) if for all n there is $m \geq n$ such that q_m is true, and (q_n ev.) if there exists n such that for all $m \geq n$, q_m is true. Now let q_n depend on ω, giving events*

$$A_n = \{\omega : q_n(\omega) \text{ is true}\}.$$

We now have new events,

$$\{A_n \text{ i.o.}\} = \{\omega : \omega \in A_n \text{ i.o.}\} = \bigcap_n \bigcup_{m \geq n} A_m,$$

and

$$\{A_n \text{ ev.}\} = \bigcup_n \bigcap_{m \geq n} A_m.$$

In analysis, $1_{(A_n \text{ i.o.})} = \lim_{n\to\infty} \sup_{m\geq n} 1_{A_m}$ and $1_{(A_n \text{ ev.})} = \lim_{n\to\infty} \inf_{m\geq n} 1_{A_m}$.

Given a sequence of events A_n for each $\omega \in \Omega$, consider $1_{A_n(\omega)}$ as a function of n, $\omega \longmapsto (1, 0, 0, 1, \dots)$.

Notice (de Morgan) that $\{A_n \text{ i.o.}\}^c = \{A_n^c \text{ ev.}\}$ and $\{A_n \text{ ev.}\}^c = \{A_n^c \text{ i.o.}\}$

Observe $X_n \xrightarrow{a.s.} X \iff \forall\epsilon > 0, \mathbb{P}(|X_n - X| > \epsilon \text{ i.o.}) = 0$.

Argue this (Facts about convergence) $X_n \to X \iff \forall\epsilon > 0, |X_n - X| < \epsilon$ ev., so

$$\begin{aligned} X_n \xrightarrow{a.s.} X &\iff \forall\epsilon > 0, \mathbb{P}(|X_n - X| \leq \epsilon \text{ ev.}) = 1 \\ &\iff \forall\epsilon > 0, \mathbb{P}(|X_n - X| > \epsilon \text{ i.o.}) = 0. \end{aligned}$$