1 Supplement: Convergence of Random Variables

Consider the standard probability space on the unit interval: $\Omega = [0, 1]$, $\mathcal{F}$ is the Borel σ-algebra, and P is the Lebesgue measure (the distribution of a $U(0, 1)$ random variable).

Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of Borel sets, i.e., $A_n \in \mathcal{F}$ for all n, and let $\{a_n\}_{n=1}^{\infty}$ be a sequence of real numbers. For each n, define a random variable X_n by

$$X_n(\omega) = a_n I_{A_n}(\omega)$$

(recall that I_A denotes the indicator function of the set A). In other words, for each $\omega \in [0, 1]$ we have

$$X_n(\omega) = \begin{cases} a_n, & \text{if } \omega \in A_n \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

We are interested in various ways the sequence $\{X_n\}$ may converge or fail to do so. Let's consider the following examples:

1. The sequence a_n converges to some $a \in \mathbb{R}$, and

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = 1, \tag{2}$$

and for any $\omega \in \bigcup_n A_n$ there is only a finite number of A_n's that do not contain it. In this case, $X_n \xrightarrow{\text{a.s.}} a$, regardless of any other details of the sets A_n.

To see this, consider what happens for each individual ω. There are two cases. First, ω may not be in any of the A_n's. But by (2) the set of all such ω's has probability zero. So we only care about those ω's that are in at least one A_n. By assumption, any such ω is contained in all but finitely many A_n's. Therefore, from (1) we see that there is some N (which may depend on ω) such that

$$X_n(\omega) = a_n, \qquad \forall n \geq N.$$

Since $\{a_n\}$ is a convergent sequence of reals, the subsequence $\{a_n\}_{n=N}^{\infty}$ converges to the same limit for any N. Consequently, $X_n(\omega) \to a$ for each $\omega \in \bigcup_n A_n$. Since the latter event has probability one, we conclude that $X_n \xrightarrow{\text{a.s.}} a$, as claimed. Because almost sure convergence implies convergence in probability, we also have $X_n \xrightarrow{\text{p.}} a$.

In the special case that $a_n \to 0$, we will have $X_n \xrightarrow{\text{a.s.}} 0$ provided (2) holds, without any other conditions on $\{A_n\}$.

2. Assume again that the sequence $\{a_n\}$ converges to some $a \in \mathbb{R}$. We are interested in seeing whether $X_n \xrightarrow{\text{p.}} a$ as well.

   Because $a_n \to a$, for any $\varepsilon > 0$ there exists some N (which may depend on ε), such that $|a_n - a| < \varepsilon$ for all $n \geq N$. Consider now an arbitrary $\omega \in [0,1]$ and some $n \geq N$. If $\omega \notin A_n$, then $X_n(\omega) = 0$ and $|X_n(\omega) - a| = |a|$. Unless $a = 0$, $\omega \notin A_n$ will imply that $|X_n(\omega) - a| \geq \varepsilon$ for all sufficiently small $\varepsilon > 0$. If $\omega \in A_n$, then $|X_n(\omega) - a| = |a_n - a| < \varepsilon$. Therefore, if $|X_n - a| \geq \varepsilon$, then $\omega \notin A_n$ (to see this, note that if $\omega \in A_n$, then it is definitely the case that $|X_n(\omega) - a| = |a_n - a| < \epsilon$). More precisely,

   $$\{\omega \in [0,1] : |X_n(\omega) - a| \geq \varepsilon\} \subseteq A_n^c, \qquad \forall n \geq N.$$

   So for all $n \geq N$ we have

   $$P\left(|X_n - a| \geq \varepsilon\right) \leq P(A_n^c) = 1 - P(A_n). \tag{3}$$

   So if $P(A_n) \to 1$, then $X_n \xrightarrow{\text{p.}} a$. It is not hard to see that X_n need not converge to a almost surely, unless we assume something extra about $\{A_n\}$. One such instance is as follows. Suppose $a_n \to a$, and the sets A_n are such that

   $$\sum_{n=1}^{\infty} P(A_n^c) < \infty.$$

   This together with (3) implies that

   $$\begin{aligned}\sum_{n=1}^{\infty} P\left(|X_n - a| \geq \varepsilon\right) &= \sum_{n=1}^{N-1} P\left(|X_n - a| \geq \varepsilon\right) + \sum_{n=N}^{\infty} P\left(|X_n - a| \geq \varepsilon\right) \\ &\leq \sum_{n=1}^{N-1} P\left(|X_n - a| \geq \varepsilon\right) + \sum_{n=N}^{\infty} P(A_n^c) \\ &\leq N - 1 + \sum_{n=1}^{\infty} P(A_n^c) \\ &< \infty.\end{aligned}$$

   Note that N will in general depend on ϵ, but this is fine. What we are after is the statement that

   $$\sum_{n=1}^{\infty} P\left(|X_n - a| \geq \varepsilon\right) < \infty$$

for any $\epsilon > 0$, and we don't need the sum to be bounded uniformly for all ϵ. By the Borel–Cantelli lemma (Lemma 1.2.2 in the lecture notes),

$$P\left(|X_n - a| \geq \varepsilon \text{ infinitely often}\right) = 0, \qquad \forall \epsilon > 0. \tag{4}$$

We claim that (4) implies that $X_n \xrightarrow{\text{a.s.}} a$. Indeed, by definition $X_n \xrightarrow{\text{a.s.}} X$ if

$$P\left(\left\{\omega : \exists k \in \mathbb{N} \text{ such that } |X_n(\omega) - X(\omega)| \geq \frac{1}{k} \text{ infinitely often}\right\}\right) = 0.$$

(can you see why?). By (4), we conclude that $X_n \xrightarrow{\text{a.s.}} a$.

Reminder: if $\{E_n\}$ is a sequence of events, then the event

$$\{E_n \text{ occurs infinitely often}\},$$

also abbreviated as $\{E_n \text{ i.o.}\}$, is defined as

$$\{E_n \text{ i.o.}\} = \bigcap_{n=1} \bigcup_{k \geq n} E_k.$$

In other words, $\{E_n \text{ i.o.}\}$ is the set of all $\omega \in \Omega$, such that ω belongs to infinitely many E_n's.

3. Let us now see whether $\{X_n\}$ converges in the mean square sense. Again, suppose that $a_n \to a$, where $a \neq 0$. Let us see whether $X_n \xrightarrow{\text{m.s.}} a$ as well. To that end, we have

$$\begin{aligned}
\mathbb{E}[|X_n - a|^2] &= \mathbb{E}\left[|a_n I_{A_n} - a|^2\right] \\
&= a_n^2\, \mathbb{E}[I_{A_n}] - 2a_n a\, \mathbb{E}[I_{A_n}] + a^2 \\
&= \left(a_n^2 - 2a_n a\right) P(A_n) + a^2 \\
&= (a_n - a)^2\, P(A_n) + a^2\left(1 - P(A_n)\right).
\end{aligned}$$

Because $\{a_n\}$ converges to a, we have

$$\lim_{n\to\infty} (a_n - a)^2 P(A_n) \leq \lim_{n\to\infty} (a_n - a)^2 = 0.$$

Since the left-hand side of the above equation is nonnegative, we have

$$\lim_{n\to\infty} \mathbb{E}[|X_n - a|^2] = a^2 - a^2 \lim_{n\to\infty} P(A_n).$$

Consequently, if $P(A_n) \to 1$, then $X_n \xrightarrow{\text{m.s.}} a$. Because convergence in mean square sense implies convergence in probability, $X_n \xrightarrow{\text{p.}} a$ as well. If $a_n \to 0$, then $X_n \xrightarrow{\text{m.s.}} 0$ as well, regardless of $\{A_n\}$, because in that case

$$\mathbb{E}[X_n^2] = a_n^2 P(A_n) \leq a_n^2.$$

The right-hand side converges to zero, so $\mathbb{E}[X_n^2]$ does too.

Another instance of m.s. convergence to zero is when $\{a_n\}$ and $\{A_n\}$ are such that the sequence $a_n^2 P(A_n)$ converges to zero. (In fact, $X_n \xrightarrow{\text{m.s.}} 0$ if and only if $a_n^2 P(A_n) \to 0$.) In this case, a_n need not be convergent. For instance, if $a_n = n^{\frac{1-\varepsilon}{2}}$ for some $\varepsilon \in (0,1)$ and $P(A_n) \le n^{-1}$ for all n, then

$$a_n^2 P(A_n) \le \frac{n^{1-\varepsilon}}{n} = n^{-\varepsilon} \xrightarrow{n\to\infty} 0.$$