Convergence: Weak, Almost Uniform, and in Probability

Consider the relationships between the convergence concepts introduced in the previous section and weak convergence. First we shall be a bit formal and note that convergence in probability to a constant can be defined for maps with different domains \((\Omega, A, P)\) too, so that it is not covered by Definition 1.9.1 in the preceding section.

1.10.1 Definition. Let \(X_n : \Omega_n \to \mathbb{D}\) be an arbitrary net of maps and \(c \in \mathbb{D}\). Then \(X_n\) converges in outer probability to \(c\) if \(P^*(d(X_n, c) > \varepsilon) \to 0\), for every \(\varepsilon > 0\). This is denoted \(X_n \xrightarrow{P^*} c\).

In general, convergence in outer probability is stronger than weak convergence, though they are equivalent if the limit is constant.

1.10.2 Lemma. Let \(X, Y\) be arbitrary maps and \(X\) Borel measurable.

(i) If \(X \xrightarrow{P^*} X\) and \(d(X_n, Y) \xrightarrow{P^*} 0\), then \(Y \xrightarrow{X}\).

(ii) If \(X_n \xrightarrow{P^*} X\), then \(X_n \xrightarrow{X}\).

(iii) \(X_n \xrightarrow{P^*} c\) if and only if \(X_n \xrightarrow{c}\).

Proof. (i). Let \(F\) be closed. Then, for every fixed \(\varepsilon > 0\), it holds that \(\limsup P^*(Y_n \in F) = \limsup P^*(Y_n \in F \land d(X_n, Y_n) \leq \varepsilon) \leq \limsup P^*(X_n \in F) \leq P(X \in F)\). Letting \(\varepsilon \downarrow 0\) completes the proof.

(ii). Clearly, \(X \xrightarrow{c} X\) and \(d(X, X_n) \xrightarrow{P^*} 0\). Apply (i).

(iii). One direction follows from (ii). For the other, note that \(P^*(d(X_n, c) > \varepsilon) = P^*(X_n \notin B(c, \varepsilon))\), where \(B(c, \varepsilon)\) is the ball of radius \(\varepsilon\).
around \( c \). By the portmanteau theorem, the limsup of this is smaller then or equal to \( P(c \notin B(c, \varepsilon)) = 0. \)

The second assertion of the previous lemma can certainly not be inverted: weakly convergent maps need not even be defined on the same probability space. However, according to the almost sure representation theorem, for every weakly convergent net there is an almost surely convergent net (defined on some probability space) that is the “same” as far as laws are concerned. The nonmeasurable version of this result sounds somewhat more complicated than the measurable version; so it is worth considering the case of measurable maps first. For every \( \alpha \), let \( \mathcal{D}_\alpha \) be a \( \sigma \)-field on \( \mathbb{D} \), not larger than the Borel \( \sigma \)-field. For convenience of notation, the limit variable will be written as \( X_\infty \) rather than \( X \) and a statement that is valid “for every \( \alpha \)” will be understood to apply to \( \alpha = \infty \) too.

1.10.3 Theorem (a.s. representations). Let \( X_\alpha : \Omega_\alpha \rightarrow \mathbb{D} \) be \( \mathcal{D}_\alpha \)-measurable maps. If \( X_\alpha \rightsquigarrow X_\infty \) and \( X_\infty \) is separable, then there exist \( \mathcal{D}_\alpha \)-measurable maps \( \tilde{X}_\alpha : \tilde{\Omega} \rightarrow \mathbb{D} \) defined on some probability space \((\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})\) with

\[(i) \ \tilde{X}_\alpha \text{ a.s. } X_\infty;
(ii) \ \tilde{X}_\alpha \text{ and } X_\alpha \text{ are equal in law on } \mathcal{D}_\alpha \text{ for every } \alpha.\]

Usually this theorem will be applied with every \( \mathcal{D}_\alpha \) equal to the Borel or ball \( \sigma \)-field. In any case, the smaller the \( \mathcal{D}_\alpha \) are, the weaker the result. In the extreme case that the \( \mathcal{D}_\alpha \) are the trivial \( \sigma \)-fields, the theorem is still true, but it yields “representations” \( \tilde{X}_\alpha \) with no relationship to the original \( X_\alpha \) whatsoever. Thus it is worthwhile to pursue a stronger formulation for nonmeasurable maps. The problem is to generalize the statement that every pair \( \tilde{X}_\alpha \) and \( X_\alpha \) are equal in law. In the initial formulation, equality in law will be interpreted in the sense that

\[E^*f(\tilde{X}_\alpha) = E^*f(X_\alpha), \quad \text{for every bounded } f : \mathbb{D} \rightarrow \mathbb{R}.\]

In particular, \( P^*(\tilde{X}_\alpha \in B) = P^*(X_\alpha \in B) \) for every set \( B \), and the same for inner probabilities. If \((\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})\) is complete (which may be assumed without loss of generality), this implies that the laws of \( \tilde{X}_\alpha \) and \( X_\alpha \) are the same on the maximal \( \sigma \)-field \( \{B \subset \mathbb{D} : X_\alpha^{-1}(B) \in \mathcal{A}_\alpha\} \) for which they are defined, that is, for which \( X_\alpha \) is measurable (Problem 1.2.10). However, in general, equality of the outer expectations says a lot more than just equality in law.

The following nonmeasurable representation theorem holds for sequences but not nets in general. Call a directed set nontrivial if it permits a net of strictly positive numbers \( \delta_\alpha \) with \( \delta_\alpha \rightarrow 0 \). Of course, the set of natural numbers with the usual ordering is nontrivial.†

† There do exist directed sets that are trivial; for each of them there is a net \( X_\alpha \) for which the theorem fails (Problem 1.10.7). On the other hand, for such a directed set \( X_\alpha \rightsquigarrow X_\infty \) with...