Stochastic Convergence

Kosorok: 7.2.2 - 7.3

June 30, 2021

Outline

Subsection 7.2.2: Spaces of Bounded Functions

We turn our attention away from weak convergence in general metric spaces to the specific case $X_n \in \ell^\infty(T)$, the metric space of all uniformly bounded functions on arbitrary index set T.

This is a useful restriction, since most statistical applications of empirical process theory will occur within $\ell^\infty(T)$.

A useful property in this setting is that asymptotic measurability of X_n follows from the asymptotic measurability of $X_n(t)$ for each $t \in T$:

Lemma 7.16

Let the sequence of maps X_n in $\ell^\infty(T)$ be asymptotically tight. Then X_n is asymptotically measurable if and only if $X_n(t)$ is asymptotically measurable for each $t \in T$.

Proof of Lemma 7.16

$\Longrightarrow$

Let $f_t : \ell^\infty(T) \mapsto \mathbb{R}$ be the marginal projection at $t \in T$, i.e., $f_t(x) = x(t)$ for any $x \in \ell^\infty(T)$. Since f_t is continuous, asymptotic measurability of X_n implies asymptotic measurability of $X_n(t)$ for each $t \in T$.

$\Longleftarrow$

Now, assume that $X_n(t)$ is asymptotically measurable for each $t \in T$. Lemma 7.14 implies asymptotic measurability for all finite-dimensional joint sequences of marginals, $(X(t_1), \ldots, X(t_k))$.

$\Longleftarrow$ (cont.)
Consequently, all functions $f \in \mathcal{F} \subset C_b(\ell^\infty(T))$ of the form $f(x) = g(x(t_1), \dots, x(t_k))$ with $g \in C_b(\mathbb{R}^k)$ are asymptotically measurable. Since $\mathcal{F}$ is a valid subalgebra that separates points in $\ell^\infty(T)$, asymptotic measurability of X_n follows readily from lemma 7.9.

Theorem 7.17 yields a convenient result, that convergence of finite dimensional distributions together with asymptotic tightness is equivalent to weak convergence in $\ell^\infty(T)$:

Theorem 7.17

The sequence X_n converges to a tight limit in $\ell^\infty(T)$ iff X_n is asymptotically tight, and all finite-dimensional marginals converge weakly to limits.
Moreover, if X_n is asymptotically tight and all its finite dimensional marginals $(X_n(t_1), \ldots, X_n(t_k))$ converge weakly to the marginals of a process $(X(t_1), \ldots, X(t_k))$, then there is a version of X such that $X_n \rightsquigarrow X$ and X resides in $UC(T, \rho)$ for some semimetric ρ making T totally bounded.

Proof of Theorem 7.17

We begin with the first statement. A sketch of the proof follows:

$\Longleftarrow$

Define a vector lattice on a subset of continuous bounded functions as in lemma 7.3. Applying lemma 7.9 yields asymptotic measurability of X_n. Applying Prohorov's theorem yields that X_n is relatively compact, or that there is a weakly convergent subsequence $X_{n'}$ which converges to a tight Borel law. By convergence of finite-dimensional marginals, all of the finite-dimensional marginals of $X_{n'}$ and X_n must converge weakly to the same limits, and by consequence of lemma 7.3, the limiting law of X_n is the limiting tight law of $X_{n'}$.

The implication is simpler:

$\Longrightarrow$

Assume that X_n converges to a tight limit. Lemma 7.12 yields that X_n is asymptotically tight. The continuous mapping theorem yields that the finite-dimensional marginals of X_n converge to those of X.

Proof of Theorem 7.17

The final implication:

$\implies$

Assume that X_n is asymptotically tight and all of its finite-dimensional marginals converge weakly to the marginals of a process X. By the asymptotic tightness of X_n, X is tight, and there is a version of X that lies in some σ-compact $K \subset \ell^\infty(T)$ with probability one. Application of of Arzelá-Ascoli gives that $K \subset UC(T, \rho)$ for some ρ making T totally bounded.

Looking back to Theorem 2.1

We refer back to Theorem 2.1 from chapter 2.

Theorem 2.1

X_n converges weakly to a tight X in $\ell^\infty(T)$ if and only if:

(i) For all finite $\{t_1, \ldots, t_k\} \subset T$, the multivariate density of $\{X_n(t_1), \ldots, X_n(t_k)\}$ converges to that of $\{X(t_1), \ldots, X(t_k)\}$.

(ii) There exists a semimetric ρ for which T is totally bounded and:

$$\inf_{\delta \downarrow 0} \limsup_{n \to \infty} P^* \left\{ \sup_{s,t \in T \text{ s.t. } \rho(s,t) < \delta} |X_n(s) - X_n(t)| > \epsilon \right\} = 0$$

When the second condition holds for all $\epsilon > 0$, we say the sequence X_n is *asymptotically uniformly ρ-equicontinuous in probability*.

Looking back to Theorem 2.1

Theorem 2.1 is slightly informal in that the conditions imply $X_n \rightsquigarrow X'$ in $\ell^\infty(T)$ for some tight version X' of X rather than $X_n \rightsquigarrow X$ directly.

With the prior developments, we are adequately prepared to prove Theorem 2.1.

First, recall that:

$$\|x\|_T = \sup_{t \in T} |x(t)|$$

$\Longrightarrow$

Begin by assuming $X_n \rightsquigarrow X$ in $\ell^\infty(T)$ for tight X. Convergence of all finite-dimensional distributions follows from CMT. Since X is tight, theorem 7.2 guarantees that $P(X \in UC(T,\rho)) = 1$ for some semimetric ρ making T totally bounded. Thus, for every $\eta > 0$, there exists some compact subset K of $UC(T,\rho)$ such that:

$$\limsup_{n\to\infty} P_*(X_n \in K^\delta) \geq 1 - \eta, \; \forall \delta > 0$$

Fix $\eta > 0$ and let the compact set K satisfy the above. For an arbitrary fixed $\epsilon > 0$, the first propostion of theorem 6.2 provides the existence of a $\delta_0 > 0$ such that:

$$\sup_{x \in K} \sup_{s,t:\rho(s,t)<\delta_0} |x(s) - x(t)| \leq \epsilon/3$$

$\Longrightarrow$ (*cont.*)
We now have:

$$P^*\left[\sup_{s,t:\rho(s,t)<\delta_0}|X_n(s)-X_n(t)|>\epsilon\right]$$
$$\leq P^*\left[\sup_{s,t:\rho(s,t)<\delta_0}|X_n(s)-X_n(t)|>\epsilon, X_n\in K^{\epsilon/3}\right]+P^*(X_n\notin K^{\epsilon/3})$$
$$\equiv E_n$$

which satisfies $\limsup_{n\to\infty}E_n\leq\eta$, since if $x\in K^{\epsilon/3}$ then $\sup_{s,t:\rho(s,t)<\delta_0}|x(s)-x(t)|<\epsilon$. Since η and ϵ were arbitrary, X_n is asymptotically uniformly ρ-continuous in probability.

$\Longleftarrow$

To prove this direction, we rely on lemma 7.18. We will first state lemma 7.18, then return to finish this direction of theorem 2.1:

Theorem 7.18

Assume that conditions (i) and (ii) of theorem 2.1 hold. Then X_n is asymptotically tight.

Proof found in section 7.4 of the book.

Return to Theorem 2.1

$\Longleftarrow$ (*cont.*)

With lemma 7.18, the remainder of the proof of theorem 2.1 is quite simple.

Assuming conditions (i) and (ii), asymptotic tightness of X_n holds readily by lemma 7.18. Now, asymptotic tightness together with convergence of all finite-dimensional marginals satisfies the premise of lemma 7.17, and thus X_n converges weakly to a tight limit, as required.

Note on Theorem 2.1

So far, we've shown that for tight X and $X_n \rightsquigarrow X$, any semimetric ρ which defines a σ-compact $UC(T, \rho)$ such that $P(X \in UC(T, \rho) = 1$ will also result in X_n being uniformly ρ-equicontinuous in probability.

How about the converse? I.e., can any semimetric, say ρ_*, which enables uniform asymptotic equicontinuity of X_n also be used to define a σ-compact $UC(T, \rho_*)$ wherein X resides with probability 1? Theorem 7.19 shows that the two statements are interchangeable when considering $\ell^\infty(T)$.

Theorem 7.19

Assume $X_n \rightsquigarrow X$ in $\ell^\infty(T)$, and let ρ be a semimetric making (T, ρ) totally bounded. Then the following are equivalent:

(i) X_n is asymptotically uniformly ρ-equicontinuous in probability

(ii) $P(X \in UC(T, \rho)) = 1.$

We prove Theorem 7.19 in the following slides.

Proof of Theorem 7.19

$(ii) \implies (i)$
As mentioned, this direction is readily proven by the arguments in our proof of 2.1.

$(i) \implies (ii)$ Assuming (i). For any $x \in \ell^\infty(T)$, for $\delta > 0$, define the function $M_\delta(x) \equiv \sup_{s,t:\rho(s,t)<\delta} |x(s) - x(t)|$.

Restricting $\delta \in (0,1)$ yields that $x \mapsto M_{(\cdot)}(x)$, is a map from $\ell^\infty(T)$ to $\ell^\infty((0,1))$ which is continuous since $|M_\delta(x) - M_\delta(y)| \leq 2||x - y||_T$, $\forall \delta \in (0,1)$.

Proof of Theorem 7.19

$(i) \implies (ii)(cont.)$
Since this map is continuous for $\delta \in (0,1)$, we have that $M_{(\cdot)}(X_n) \rightsquigarrow M_{(\cdot)}(X)$ in $\ell^\infty((0,1))$

Since X_n is asymptotically uniformly ρ-equicontinuous in probability, there exists a positive sequence $\delta_n \downarrow 0$ such that $P(M_{\delta_n}(X_n) > \epsilon) \to 0$ for every $\epsilon > 0$. Thus, $M_{\delta_n}(X) \rightsquigarrow 0$.

A application of theorem 2.1 to X yields that X is tight, and thus the desired result.

Discussion of Theorem 7.19

Taking theorems 2.1 and 7.19 together with lemma 7.4 results in an interesting consequence when X_n converges weakly in $\ell^\infty(T)$ to a tight Gaussian process X.

Consider the semimetric $\rho_p(s,t) = (E|X(s) - X(t)|^p)^{1/(p \vee 1)}$ for any $p \in (0, \infty)$. Then for any $p \in (0, \infty)$, (T, ρ_p) is totally bounded and the sample paths of X are ρ_p-equicontinuous, and X_n is asymptotically uniformly ρ_p-equicontinuous in probability.

A special, convenient case, is found by taking $p = 2$, the "standard deviation" metric.

Lemma 7.20

We conclude this section with an equivalent condition to X_n being asymptotically uniformly ρ-equicontinuous in probability. This condition, stated in lemma 7.20, is sometimes easier to verify in certain settings.

Lemma 7.20

Let X_n be a sequence of stochastic processes indexed by T. Then the following are equivalent:

(i) There exists a semimetric ρ making T totally bounded and for which X_n is uniformly ρ-equicontinuous in probability.

(ii) For every $\epsilon, \eta > 0$, there exists a finite partition $T = \cup_{i=1}^{k} T_i$ such that:

$$\limsup_{n\to\infty} P^* \left(\sup_{1\le i\le k} \sup_{s,t\in T} |X_n(s) - X_n(t)| > \epsilon \right) < \eta$$

The proof is omitted here, but can be found in section 7.4.

Section 7.3: Other Modes of Convergence

Other Modes of Convergence

We begin by recalling the definitions of convergence in probability and convergence outer almost surely.

We say that X_n converges to X in probability (denoted $X_n \xrightarrow{P} X$ if $P\left\{d(X_n, X)^* > \epsilon\right\} \to 0$ for all $\epsilon > 0$.

We say that X_n converges outer almost surely to X (denoted $X_n \xrightarrow{as*} X$) if there exists a sequence of measurable random variables Δ_n, such that $d(X_n, X) \leq \Delta_n$ for all n and $P\left\{\limsup_{n\to\infty} \Delta_n = 0\right\} = 1$.

Other Modes of Convergence

We additionally define two other modes of convergence which can be useful:

We say that X_n *converges almost uniformly* to X if for every $\epsilon > 0$, there exists a measurable set A such that $P(A) \geq 1 - \epsilon$ and $d(X_n, X) \to 0$ uniformly on A.

We say that X_n *converges almost surely* to X if $P_* \left(\mathrm{lim}_{n \to \infty} d(X_n, X) = 0\right) = 1$

Note that the definitions of convergence almost surely and convergence outer almost surely differ only in that for the latter, $d(X_n, X)$ is required to be bounded above by a measurable random variable which converges to 0.

This distinction is not trivial. It can be shown that almost sure convergence does not, in general, imply convergence in probability when $d(X_n, X)$ is not measurable.

Lemma 7.21 on the following slide describes the relationships between the modes.

Lemma 7.21

Let $X_n, X : \Omega \mapsto \mathbb{D}$ be maps with X Borel measurable. Then:

(i) $X_n \overset{as*}{\to} X \implies X_n \overset{P}{\to} X$

(ii) $X_n \overset{P}{\to} X$ if and only if every subsequence $X_{n'}$ has a further subsequence $X_{n''}$ such that $X_{n''} \overset{as*}{\to} X$

(iii) $X_n \overset{as*}{\to} X$ if and only if X_n converges almost uniformly to X if and only if $\sup_{m \geq n} d(X_m, X) \overset{P}{\to} 0$.

Note that for sequences of maps, almost uniform convergence and outer almost sure convergence are equivalent. This is not true for nets.

Extending probability convergence

Thus far we have restricted ourselves to sequences X_n defined on a fixed probability space Ω.

To allow for probability spaces which change in n, we need to extend the definition of convergence in probability to the convergence of a stochastic process to a constant.

This extended convergence mode is simply denoted $X_n \xrightarrow{P} c$, for a constant c, and will be distinguished only by context.

Proposition 7.22

The following proposition gives the connection between convergence almost surely and convergence in probability. We sketch the proof below.

Proposition 7.22

Let $X_n, Y_n : \Omega \mapsto \mathbb{D}$ be maps with Y_n measurable. Suppose every subsequence n' has a a further subsequence n'' such that $X_{n''} \to 0$ almost surely. Suppose also that $d(X_n, Y_n) \xrightarrow{P} 0$. Then $X_n \xrightarrow{P} 0$.

The idea of the proof is recognizing that any the further subsequence of any arbitrary subsequence of Y_n, $Y_{n''}$, converges to 0 almost surely. Measurability of Y_n then implies that $Y_{n''} \xrightarrow{as*} 0$. This gives $Y_n \xrightarrow{P} 0$, and $X_n \xrightarrow{P} 0$ follows directly.

Lemma 7.23 describes important relationships between weak convergence and convergence in probability.

Lemma 7.23

Let $X_n, Y_n : \Omega_n \mapsto \mathbb{D}$ be maps, $X\Omega \mapsto \mathbb{D}$ be Borel measurable, and $c \in \mathbb{D}$ be a constant. Then:

(i) If $X_n \rightsquigarrow X$ and $d(X_n, Y_n) \stackrel{P}{\to} 0$, then $Y_n \rightsquigarrow X$

(ii) $X_n \stackrel{P}{\to} X$ implies $X_n \rightsquigarrow X$.

(iii) $X_n \stackrel{P}{\to} c$ if and only if $X_n \rightsquigarrow c$.

Proof to follow.

Proof of Lemma 7.23

Begin with (i). Let $F \subset \mathbb{D}$ be closed, and fix some $\epsilon > 0$. Then

$$\begin{aligned}
&\limsup_{n\to\infty} P^*(Y_n \in F) \\
&\qquad = \limsup_{n\to\infty} P^*(Y_n \in F, d(X_n, Y_n)^* \le \epsilon) \\
&\qquad \le \limsup_{n\to\infty} P^*(X_n \in \overline{F^\epsilon}) \\
&\qquad \le P(X \in \overline{F^\epsilon})
\end{aligned}$$

Letting $\epsilon \downarrow 0$ yields the result by the portmanteau theorem. For (ii), assume that $X_n \xrightarrow{P} X$, thus $d(X, X_n) \xrightarrow{P} 0$. Since $X \rightsquigarrow X$, direct application of (i) yields that $X_n \rightsquigarrow X$.

Proof of Lemma 7.23

For (*iii*), the implication is simple. $X_n \xrightarrow{P} c$ implies that $X_n \rightsquigarrow c$ by (*ii*).

For the converse, assume $X_n \rightsquigarrow c$ and fix some $\epsilon > 0$. It's clear that $P^*(d(X_n, c) \geq \epsilon) = P^*(X_n \notin B(c, \epsilon))$ where $B(c, \epsilon)$ is an open ϵ-ball around c in $\mathbb{D}$. By the portmanteau theorem, $\limsup_{n\to\infty} P^*(X_n \notin B(c, \epsilon)) \leq P(X \notin B(c, \epsilon)) = 0$. Thus $X_n \xrightarrow{P} c$ since ϵ is arbitrary.

Theorem 7.24: Extended continuous mapping

Let $\mathbb{D}_n \subset \mathbb{D}$ and $g_n : \mathbb{D}_n \mapsto \mathbb{E}$ satisfy the following. If $x_n \to x$ with $x_n \in \mathbb{D}_n$ for all $n \geq 1$ and $x \in \mathbb{D}_0$, then $g_n(x_n) \to g(x)$, where $\mathbb{D}_0 \subset \mathbb{D}$ and $g : \mathbb{D}_0 \mapsto \mathbb{E}$. Let X_n be maps taking values in $\mathbb{D}_n$, and let X be Borel measurable and separable. Then:

(i) $X_n \rightsquigarrow X$ implies $g_n(X_n) \rightsquigarrow g(X)$
(ii) $X_n \xrightarrow{P} X$ implies $g_n(X_n) \xrightarrow{P} g(X)$
(iii) $X_n \xrightarrow{as*}$ implies $g_n(X_n) \xrightarrow{as*} g(X)$.

Note that we've generalized in that we are interested in the convergence of a function g_n which is dependent on n. Following the book, we omit the proof here.

Theorem 7.25

Theorem 7.25 gives another continuous mapping result for convergence in probability and outer almost surely. Note that theorem 7.25, unlike theorem 7.24, does not require X to be separable.

Theorem 7.25

Let $g : \mathbb{D} \mapsto \mathbb{E}$ be continuous at all points in $\mathbb{D}_0 \subset \mathbb{D}$, and let X be Borel measurable with $P_*(X \in \mathbb{D}_0) = 1$. Then:

(i) $X_n \stackrel{P}{\to} X$ implies $g(X_n) \stackrel{P}{\to} g(X)$

(ii) $X_n \stackrel{as*}{\to}$ implies $g(X_n) \stackrel{as*}{\to} g(X)$.

Theorem 7.26

Theorem 7.26 covers a outer almost sure representation result for weak convergence. This allows certain weak convergence problems to be represented as ones of convergence of fixed sequences.

Theorem 7.26

Let $X_n : \Omega_n \mapsto \mathbb{D}$ be a sequence of maps, and let X_∞ be Borel measurable and separable. If $X_n \rightsquigarrow X_\infty$, then there exists a probability space $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$ with maps $X_n : \tilde{\Omega} \mapsto \mathbb{D}$ with:

(i) $\tilde{X}_n \overset{as*}{\to} \tilde{X}_\infty$
(ii) $E^* f(\tilde{X}_n) = E^* f(X_n)$, for every bounded $f : \mathbb{D} \mapsto \mathbb{R}$ and all $1 \leq n \leq \infty$.

Moreover, $\tilde{X}_n$ can be chosen such that is equal to $X_n \circ \phi_n$ for all $1 \leq n \leq \infty$, where the $\phi_n : \tilde{\Omega} \mapsto \Omega_n$ are measurable and perfect maps, and $P_n = \tilde{P} \circ \phi_n$

Proposition 7.27 relies directly on theorem 7.26, and provides a method for studying the weak convergence of certain statistics which can be expressed as stochastic integrals, such as the Wilcoxon statistic.

Proposition 7.27

Let $X_n, G_n \in D[a, b]$ be stochastic processes with $X_n \rightsquigarrow X$ and $G_n \xrightarrow{P} G$ in $D[a, b]$, where X is bounded with continuous sample paths, G is fixed, and G_n and G have total variation bounded by some $K < \infty$. Then $\int_a^{(\cdot)} X_n(s) dG_n(s) \rightsquigarrow \int_a^{(\cdot)} X(s) dG(s)$ in $D[a, b]$.

Proof of Proposition 7.27

Slutsky's theorem and lemma 7.23 provide that $(X_n, G_n) \rightsquigarrow (X, G)$. Next, we rely on theorem 7.26, which provides existence of a new probability space with processes $\tilde{X}_n$, $\tilde{X}$, $\tilde{G}_n$, and $\tilde{G}$ for which the outer integrals are the same for all bounded functions as their original counterparts, and also satisfy that $(\tilde{X}_n, \tilde{G}_n) \overset{as*}{\to} (\tilde{X}, \tilde{G})$

For each integer $m \geq 1$, define $t_j = a + (b-a)j/m$, $j = 0, \ldots, m$, and let:

$$M_m \equiv \max_{1 \leq j \leq m} \sup_{s,t \in (t_{j-1}, t_j]} |\tilde{X}(s) - \tilde{X}(t)|$$

now, define $\tilde{X}_m \in D[a,b]$ such that $\tilde{X}_m(a) = \tilde{X}(a)$, and $\tilde{X}_m(t) \equiv \sum_{j=1}^{m} 1\{t_{j-1} < t \leq t_j\} \tilde{X}(t_j)$ for $t \in (a, b]$.

Proof of Proposition 7.27

For integrals of the range $(a, t]$ for $t = a$, we will define the value over the integral to be 0. For any $t \in [a, b]$, we have:

$$
\begin{aligned}
&\left| \int_a^t \tilde{X}_n(s) d\tilde{G}_n(s) - \int_a^t \tilde{X}(s) d\tilde{G}(s) \right| \\
&\leq \int_a^b |\tilde{X}_n(s) - \tilde{X}(s)| \times |d\tilde{G}_n(s)| + \int_a^b |\tilde{X}_m(s) - \tilde{X}(s)| \times |d\tilde{G}_n(s)| \\
&+ \left| \int_a^t \tilde{X}_m(s) \left\{ d\tilde{G}_n(s) - d\tilde{G}(s) \right\} \right| \\
&\leq K \left(\|\tilde{X}_n - \tilde{X}\|_{[a,b]} + M_m \right) \\
&+ \left| \sum_{j=1}^{m} \tilde{X}(t_j) \int_{(t_{j-1}, t_j] \cap (a,t])} \left\{ d\tilde{G}_n(s) - d\tilde{G}(s) \right\} \right|
\end{aligned}
$$

Proof of Proposition 7.27

(Continued from last slide)

$$
\begin{aligned}
& K\left(\|\tilde{X}_n - \tilde{X}\|_{[a,b]} + M_m\right) + \left|\sum_{j=1}^{m} \tilde{X}(t_j) \int_{(t_{j-1},t_j]\cap(a,t])} \left\{ d\tilde{G}_n(s) - d\tilde{G}(s) \right\}\right| \\
& \le K\left(\|\tilde{X}_n - \tilde{X}\|_{[a,b]} + M_m\right) + m\left(\|\tilde{X}\| \times \|\tilde{G}_n - \tilde{G}\|^*_{[a,b]}\right) \\
& \equiv E_n(m)
\end{aligned}
$$

Note that $E_n(m)$ is measurable and converges to 0 almost surely. Define D_n to be the infimum of $E_n(m)$ over m. Since $D_n \overset{as*}{\to} 0$ and D_n is measurable, we have that:
$\int_a^{(\cdot)} \tilde{X}_n(s) d\tilde{G}_n(s) \overset{as*}{\to} \int_a^{(\cdot)} \tilde{X}(s) d\tilde{G}(s)$

Proof of Proposition 7.27

Now, note that for any $f \in C_b(D[a,b])$, the map

$$(x, y) \mapsto f\left(\int_a^{(\cdot)} x(s)dy(s)\right)$$

for $x, y \in D[a,b]$ is bounded when the total variation of y is bounded. Thus, by (ii) of theorem 7.26:

$$\begin{aligned}
\mathrm{E}^* f\left(\int_a^{(\cdot)} X_n(s)dG_n(s)\right) &= \mathrm{E}^* f\left(\int_a^{(\cdot)} \tilde{X}_n(s)d\tilde{G}_n(s)\right) \\
&\to \mathrm{E} f\left(\int_a^{(\cdot)} \tilde{X}(s)d\tilde{G}(s)\right) \\
&= \mathrm{E} f\left(\int_a^{(\cdot)} X(s)dG(s)\right)
\end{aligned}$$

which completes the proof, since f is arbitrary.

Lemma 7.28

The final result of the section is useful when certain questions about weakly convergent sequences are easier to answer for measurable maps. The lemma shows that a nonmeasurable, weakly convergent sequence X_n is usually close to a measurable sequence Y_n.

Proposition 7.28

Let $X_n : \Omega_n \mapsto \mathbb{D}$ be a sequence of maps.If $X_n \rightsquigarrow X$, where X is Borel measurable and separable, then there exists a Borel measurable sequence $Y_n : \Omega_n \mapsto \mathbb{D}$ with $d(X_n, Y_n) \xrightarrow{P} 0$.