Weak Convergence and Spaces of Bounded Functions

Lab 2, BIO251

02/10/2014

1 Weak Convergence

In this section we formalize weak convergence on metric spaces. Let $(\mathbb{D}, d)$ and $(\mathbb{E}, e)$ be metric spaces. Denote the continuous bounded functions $f : \mathbb{D} \to \mathbb{R}$ with $C_b(\mathbb{D})$. In the empirical processes applications the space $\mathbb{D} = \ell^\infty(T)$ under the sup metric.

Definition. A Borel probability measure L is *tight* if for every $\varepsilon > 0$, there exists a compact set K, such that: $L(K) \geq 1 - \varepsilon$.

Definition. Let $(\Omega_n, \mathcal{A}_n, P_n)$ are sets of probability spaces and $X_n : \Omega_n \mapsto \mathbb{D}$ are arbitrary maps. The sequence X_n converges weakly to a Borel measure L if:

$$\mathrm{E}^* f(X_n) \to \int f dL, \text{ for all } f \in C_b(\mathbb{D})$$

We denote weak convergence by $X_n \rightsquigarrow L$. If X has Borel law L, we also say that $X_n \rightsquigarrow X$. We will assume that X is measurable.

Consider the following equivalent ways of defining weak convergence in the following theorem:

Theorem (Portmanteau). TFAE:

(i) $X_n \rightsquigarrow L$

(ii) $\liminf P_*(X_n \in G) \geq L(G)$, for every open G

(iii) $\limsup P^*(X_n \in F) \leq L(F)$, for every closed F

(iv) $\liminf \mathrm{E}_* f(X_n) \geq \int f dL$, for every lower semincontinuous, bounded from below f

(v) $\limsup \mathrm{E}^* f(X_n) \leq \int f dL$, for every upper semicontinuous, bounded from above f

(vi) $\lim P^*(X_n \in B) = \lim P_*(X_n \in B) = L(B)$ for every Borel set B with $L(\delta B) = 0$.

(vii) $\liminf \mathrm{E}_* f(X_n) \geq \int f dL$ for every bounded, Lipschitz continuous, nonnegative f.

Partial Proof: Here we prove few selected parts of the theorem. The equivalence between (ii) and (iii) is obvious after taking complements. Similarly the equivalence of (iv) and (v) follows after substituting f with $-f$. The fact that (i) implies (vii) is obvious since a bounded Lipschitz continuous function is necessarily continuous, and it remains to apply the definition of inner expectation.

Another obvious implication is the fact that (iv) and (v) imply (i). Since (iv) and (v) are equivalent then (v) implies (i).

As an exercise we show that (vii) implies (ii). Other of the implications follow in a similar spirit. Take an open set G, and construct a sequence of Lipschitz functions f_m such that $f_m \uparrow \mathbb{1}_G$. An example of such a sequence is $f_m = md(x, \mathbb{D} - G) \wedge 1$. We have by (vii) that:

$$\liminf \mathrm{E}_* f_m(X_n) \geq \int f_m dL$$

Letting $m \to \infty$ gives us (ii). (note that $\mathrm{E}_* \mathbb{1}_G(X_n) \geq \mathrm{E}_* f_m(X_n)$ for all m and n)

A very important consequence of the definition of weak convergence is without a doubt the continuous mapping theorem:

Theorem (Continuous Mapping). Let $g : \mathbb{D} \mapsto \mathbb{E}$, is a continuous at every point of a set $\mathbb{D}_0 \subset \mathbb{D}$. If $X_n \rightsquigarrow X$ and X takes values in $\mathbb{D}_0$, we have $g(X_n) \rightsquigarrow g(X)$.

Continuous mapping theorem is not hard to show using the portmanteau theorem. Next by importance to the continuous mapping theorem for weak convergence is Prohorov's theorem. We need a couple of more definitions to get there.

Definition. The sequence of maps X_n is *asymptotically measurable* iff

$$\mathrm{E}^* f(X_n) - \mathrm{E}_* f(X_n) \to 0, \text{ for every } f \in C_b(\mathbb{D})$$

Definition. The sequence X_n is *asymptotically tight* if for every $\varepsilon > 0$ there exists a compact set K such thats:

$$\liminf P_*(X_n \in K^\delta) \geq 1 - \varepsilon \text{ for every } \delta > 0$$

Where $K^\delta = \{y \in \mathbb{D} : d(y, K) < \delta\}$ is the "δ-enlargement" around K

Lemma.

(i) If $X_n \rightsquigarrow X$, then X_n is asymptotically measurable

(ii) If $X_n \rightsquigarrow X$, then X is tight iff X_n is asymptotically tight.

Proof. Part (i) is straightforward because by definition of weak convergence we have $\lim \mathrm{E}^* f(X_n) \to \mathrm{E} f(X)$ and $-\lim \mathrm{E}_* f(X_n) = \lim \mathrm{E}^* -f(X_n) = \lim \mathrm{E} -f(X) = -\lim \mathrm{E} f(X)$.

For part (ii) we apply the portmanteau theorem. First let X be tight. Fix an $\varepsilon > 0$, and let K be the compact set corresponding to it such that $P(X \in K) \geq 1 - \varepsilon$. By part (ii) for any $\delta > 0$, we have that $\liminf P_*(X_n \in K^\delta) \geq P(X \in K^\delta) \geq P(X \in K) \geq 1 - \varepsilon$.

Conversely, let X_n be asymptotically tight. For a fixed $\varepsilon > 0$ take the compact set K form the definition. From part (iii) we have $P(X \in \overline{K^\delta}) \geq \limsup P^*(X_n \in K^\delta) \geq \liminf P_*(X_n \in K^\delta) \geq 1 - \varepsilon$. Let $\delta \to 0$.

Now we are ready to state Prohorov's theorem, which can be viewed as a converse to the previous lemma.

Theorem (Prohorov). If the sequence X_n is asymptotically tight and asymptotically measurable, there exists a subsequence X_{n_j} which converges weakly to a tight Borel law.

Often times, in empirical process theory some processes map into spaces that are proper subsets of $\ell^\infty(T)$, like $D(T)$ the Skorohod space. Regardless of this fact, if the metrics are kept the same (or the topology in the subspace is taken to be the relative topology) weak convergence is the same. This will be formalized in the following theorem:

Theorem. Let $\mathbb{D}_0 \subset \mathbb{D}$ be arbitrary and let X and X_n take values in $\mathbb{D}_0$. Then $X_n \rightsquigarrow X$ as maps in $\mathbb{D}_0$ iff $X_n \rightsquigarrow X$ as maps in $\mathbb{D}$. Here $\mathbb{D}_0$ and $\mathbb{D}$ are equipped with the same metric.

Proof: Since an open set in G_0 in $\mathbb{D}_0$ has the form $G_0 = G \cap \mathbb{D}_0$, for some open G in $\mathbb{D}$, by (ii) in portmanteau theorem we get:

$$\liminf P_*(X_n \in G \cap \mathbb{D}_0) \geq P(X \in G \cap \mathbb{D}_0)$$

However in the last expression $P_*(X_n \in G \cap \mathbb{D}_0) = P_*(X_n \in G)$, $P(X \in G \cap \mathbb{D}_0) = P(X \in G)$ since X_n and X take their values on the space $\mathbb{D}_0$. Since this is true for all open G in $\mathbb{D}$ this finishes the proof.

Finally there is the question whether weak convergence is to a unique Borel law or there might be more than one laws. The negative answer is provided by the following lemma:

Lemma.

(i) Let L_1 and L_2 are finite Borel measures on $\mathbb{D}$. If $\int f dL_1 = \int f dL_2$ for all $f \in C_b(\mathbb{D})$ then $L_1 \equiv L_2$.

(ii) Let L_1 and L_2 be tight Borel probability measures on $\mathbb{D}$. If $\int f dL_1 = \int f dL_2$ for every f in a vector lattice $\mathcal{F} \subset C_b(\mathbb{D})$ that contains the constant functions, and separates the points in $\mathbb{D}$, then $L_1 \equiv L_2$.

Here vector lattice means a vector space of functions $\mathcal{F}$ such that $f \in \mathcal{F}$ it follows that $f^+ = f \vee 0 \in \mathcal{F}$. $\mathcal{F}$ is said to separate the points if for every two points $x \neq y$ there exists $f \in \mathcal{F}$ such that $f(x) \neq f(y)$.

To help establishing asymptotic measurability in practice we need the following lemma:

Lemma. Let X_n be asymptotically tight, and suppose that $\mathrm{E}^* f(X_n) - \mathrm{E}_* f(X_n) \to 0$ holds for $f \in \mathcal{F}$, where $\mathcal{F}$ is a subalgebra of $C_b(\mathbb{D})$, which separates the points of $\mathbb{D}$. Then X_n is asymptotically measurable.

A subalgebra is a vector space that in addition is closed under taking products.

Slutsky's Lemma If $X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow c$ with X being separable (i.e. there exists separable, measurable set with prob 1) and c a constant, then $(X_n, Y_n) \rightsquigarrow (X, c)$.

2 Spaces of Bounded Functions

Let T be an arbitrary set. The space $\ell^\infty(T)$ is defined as the set of all uniformly bounded real functions $z : T \mapsto \mathbb{R}$ with:

$$||z||_T := \sup_{t \in T} |z(t)| < \infty$$

Definition. A *stochastic process* is simply an indexed collection of random variables $\{X(t) : t \in T\}$ defined on the same probability space: every $X(t) : \Omega \mapsto \mathbb{R}$ is a measurable map. If the *sample paths*

$t \mapsto X(t, \omega)$ are bounded, then a stochastic process yields a map $X : \Omega \mapsto \ell^\infty(T)$.

The amount of measurability given by the stochastic process definition, might already be enough for asymptotic measurability. The *marginals* $(X(t_1), X(t_2), \dots, X(t_k))$ play a special role for weak convergence, when considered as maps into $\mathbb{R}^k$, and this is obvious after the following three lemmas.

Lemma 1. Let $X_n : \Omega_n \mapsto \ell^\infty(T)$ be asymptotically tight. Then it is asymptotically measurable iff $X_n(t)$ is asymptotically measurable for every $t \in T$.

Lemma 2. Let X and Y be tight Borel measurable maps into $\ell^\infty(T)$. Then X and Y are equal in Borel law iff all corresponding marginals of X and Y are equal in law.

Theorem. Let $X_n : \Omega_n \mapsto \ell^\infty(T)$ be arbitrary. Then X_n converges weakly to a tight limit iff X_n is asymptotically tight, and the marginals $(X_n(t_1), \dots, X_n(t_k))$ converge weakly to a limit for every finite subset $t_1, \dots, t_k$ of T. If X_n is asymptotically tight and it's marginals converge weakly to the marginals $(X(t_1), \dots, X(t_k))$ of a stochastic process X then there is a version of X with uniformly bounded sample paths and $X_n \rightsquigarrow X$.

Proofs. For the first two lemmas consider the collections of functions $\mathcal{F}$, s.t. for $f \in \mathcal{F}$ we have $f : \ell^\infty(T) \mapsto \mathbb{R}$ of the form:

$$f(z) = g(z(t_1), \dots, z(t_k)), g \in C_b(\mathbb{R}^k), t_i \in T, k \in \mathbb{N}$$

Obviously, $\mathcal{F}$ forms a vector lattice and a subalgebra, and contains the constant functions, since the functions g would stay bounded under multiplication or taking the positive parts. Furthermore, it also separates the points as if two functions $z_1 \neq z_2$ differ at time $t \in T$, then the projection $f(z) = z(t)$ (note that this is a continuous function) separates the two points and $f \in \mathcal{F}$.

Lemma 1. Assume X is asymptotically measurable. Then set $f(X_n) = g(X_n(t))$ to be the projection on t and then asymptotic measurability follows by the definition. Conversely assume that $X_n(t)$ is asymptotically measurable for all t. Using the fact that asymptotic measurability of two random elements implies their joint asymptotic measurability (not trivial needs proof but we omit it!) we can see that all finite dimensional marginals are asymptotically measurable. Then the last lemma of the weak convergence section and the fact that $\mathcal{F}$ is a subalgebra give us the desired result.

Lemma 2. If $X = Y$, then the marginals ought to be the same. If the marginals are the same the fact that $\mathcal{F}$ is a lattice containing constants, and separating the points, we get that $X = Y$ from part (ii) of the next to last lemma of the weak convergence section.

Theorem Proof. If X_n is asymptotically tight and the converge marginals weakly to a tight limit, then by Lemma 1 X_n is asymptotically measurable. Now we can apply Prohorov's theorem to get that we can divide X_n into subsequences each of which is converging weakly to a tight limit. If we show that each limit is the same we are done. Because we have the marginal convergence (on all subsequences), and Lemma 2 this finishes the proof in this direction. Now if X_n converges to a tight limit we know that this is equivalent to X_n being asymptotically tight. Furthermore the marginals weak convergence comes as an implication of the continuous mapping theorem (note that the map $f(z) = (z(t_1), \dots, z(t_k))$ is continuous for $z \in \ell^\infty(T)$).

Finally note that if X_n is asymptotically tight and the marginals converge, by the first part we have that $X_n \rightsquigarrow X$, where X is tight limit. Furthermore, since X is tight, then it concentrates on a σ-compact set $K \subset \ell^\infty(T)$ with $P(X \in K) = 1$. The last implication shows the uniform boundedness of the sample paths.