Weak Convergence in Higher Dimensions
January 30 and February 4, 2003

Metric Spaces. Recall that a metric space is a pair \((X, d)\), where \(X \neq \emptyset\) and \(d : X \times X \to [0, \infty)\) satisfies:

a) \(d(x, y) = 0\) iff \(x = y\).

b) \(d(x, y) = d(y, x)\).

c) \(d(x, z) \leq d(x, y) + d(y, z)\)

for all \(x, y, z \in X\). Then \(d(x, y)\) is called the distance from \(x\) to \(y\).

Example: Let \(X = \mathbb{R}^p\) and \(d(x, y) = \|x - y\|\) for \(x, y \in \mathbb{R}^p\).

The reader is assumed to be familiar with elementary properties of metric spaces—for example, open and closed sets, convergent sequences, etc.

In the sequel \(X\) denotes a metric space. Recall that if \(Y\) is a second metric space, with metric \(\rho\) say, then a function \(f : X \to Y\) is said to be continuous at \(x \in X\) iff \(\forall \epsilon > 0\) there is a \(\delta > 0\) for which \(\rho[f(x), f(y)] \leq \epsilon\) when \(d(x, y) \leq \delta\); and \(f\) is continuous if it is continuous at every \(x \in X\). The function \(f\) is said to be Lipschitz continuous if there is a constant \(K\) for which \(\rho[f(x), f(y)] \leq Kd(x, y)\) for all \(x, y \in X\). Clearly, any Lipschitz continuous function is continuous. The reader is assumed to familiar with the following

Proposition 1. A function \(f : X \to Y\) is continuous iff \(f^{-1}(U)\) is open in \(X\) for every open \(U \subseteq Y\); equivalently, iff \(f^{-1}(C)\) is closed in \(X\) for every closed \(C \subseteq Y\).

If \(B \subseteq X\), let

\[
d(B, x) = \inf\{\|x - y\| : y \in B\}.
\]

Then \(d(B, x)\) is called the distance from \(x\) to \(B\). If \(B \subseteq X\), then the closure of \(B\) is the smallest closed set containing \(B\) and the interior of \(B\) is the largest open set contained in \(B\). These are denoted by \(\bar{B}\) and \(B^o\). The boundary of \(B\) is \(\partial B = \bar{B} - B^o\).

Lemma 1. If \(B \subseteq X\) and \(x, y \in X\), then \(|d(B, x) - d(B, y)| \leq d(x, y)|\).

Proof. If \(\epsilon > 0\), then there is a \(z \in B\) for which \(d(y, z) \leq d(y, B) + \epsilon\). Then \(d(x, B) \leq d(x, z) \leq d(x, y) + d(y, z) \leq d(x, y) + d(y, B) + \epsilon\). So, \(d(x, B) \leq d(y, B) + d(x, y)\), since \(\epsilon > 0\) was arbitrary. The lemma now follows by reversing the roles of \(x\) and \(y\). \(\diamond\)

Lemma 2. If \(B \subseteq X\), then \(d(B, x) = 0\) iff \(x \in \bar{B}\), the closure of \(B\).

Proof. \(d(x, B) = 0\) iff there are \(x_n \in B\) for which \(d(x, x_n) \leq 1/n\) for all \(n\) iff \(x \in \bar{B}\). \(\diamond\)

Now let \(B\) denote the Borel sigma-algebra in \(X\), the smallest sigma-algebra containing all open sets. Observe that if \(f : X \to \mathbb{R}\) is Lipschitz continuous and \(g : \mathbb{R} \to \mathbb{R}\) is also Lipschitz continuous, then \(g \circ f\) is Lipschitz continuous.
Lemma 3. If $F$ and $G$ are finite measures (on $\mathcal{B}$) for which

$$\int_{\mathcal{X}} hdF = \int_{\mathcal{X}} hdG$$

for all bounded Lipschitz continuous $h : \mathbb{R}^k \to \mathbb{R}$, then $F = G$.

**Proof.** Given a closed set $C$

$$\psi_m(t) = \begin{cases} 1 & \text{if } t \leq 0 \\ 1 - mt & \text{if } 0 \leq mt \leq 1 \\ 0 & \text{if } mt > 1 \end{cases}$$

and

$$h_m(x) = \psi_m[d(C, x)], \ x \in \mathbb{R}^k.$$ Then $h_m$ are bounded Lipschitz continuous functions for which $\lim_{m \to \infty} h_m(x) = 1_C$. So,

$$F\{C\} = \lim_{m \to \infty} \int h_m dF = \lim_{m \to \infty} \int h_m dG = G\{C\}.$$ That $F = G$ then follows from the $\pi$-$\lambda$ Theorem. \hfill \Diamond

Lemma 4. If $g : \mathcal{X} \to \mathbb{R}$ is continuous, then $\partial\{x : g(x) \leq c\} \subseteq \{x : g(x) = c\}$.

**Proof.** Let $C = \{x : g(x) \leq c\}$ and $U = \{x : g(x) < c\}$. Then $C$ is closed, $U$ is open, and $U \subseteq C$. So, $\partial C = C - C^o \subseteq C - U = \{x : g(x) = c\}$. \hfill \Diamond

If $F$ is a finite measure then $B \in \mathcal{R}^p$ is a **continuity set for $F$** iff $F(\partial B) = 0$, where $\partial B$ is the boundary of $B$.

**Theorem 1.** The Portmanteau Theorem. If $F_n$ and $F$ are probability measures on $\mathcal{R}^p$, then the following are equivalent:

a) $\lim_{n \to \infty} \int_{\mathcal{X}} gdF_n = \int_{\mathcal{X}} gdF$ for all bounded continuous $g$;

b) $\lim_{n \to \infty} \int_{\mathcal{X}} gdF_n = \int_{\mathcal{X}} gdF$ for all bounded Lipschitz continuous $g$;

c) $\limsup_{n \to \infty} F_n\{C\} \leq F\{C\}$ for all closed $C$;

$\liminf_{n \to \infty} F_n\{U\} \leq F\{U\}$ for all open $U$;

e) $\lim_{n \to \infty} F_n\{B\} = F\{B\}$ for all $F$-continuity sets $B$.

**Proof.** It is clear that a) implies b). If b) holds and $C$ is a closed set, construct $h_m$ as in Lemma 3. Then

$$F_n\{C\} \leq \int_{\mathcal{X}} h_m dF_n \to \int_{\mathcal{X}} h_m dF$$
as \( n \to \infty \) for each \( m \), since each \( h_m \) is bounded and Lipschitz continuous. So,

\[
\limsup_{n \to \infty} \int_{\mathcal{X}} h_m dF_n \leq \inf_{m \geq 1} \int_{\mathcal{X}} h_m dF = F\{C\},
\]

as in the proof of Lemma 3. So, b) implies c), and it is clear that c) and d) are equivalent. If c) and d) holds and \( B \) is a continuity set of \( F \), then

\[
\limsup_{n \to \infty} F_n\{B\} \leq \limsup_{n \to \infty} F_n\{\bar{B}\} \leq F\{\bar{B}^o\} \leq \liminf_{n \to \infty} F_n\{B\},
\]

so that e) holds.

Finally, suppose that e) holds. Let \( g \) be a bounded continuous function and let

\[
H_n(y) = F_n\{x \in \mathcal{X} : g(x) \leq y\},
\]

\[
H(y) = F\{x \in \mathcal{X} : g(x) \leq y\}.\]

Then \( H_n \Rightarrow H \), by e). Let \( b \) be an upper bound for \( g \) and let \( e(y) = -b, y, \) or \( b \), accordingly as \( y \leq -b, -b \leq y \leq b, \) or \( y > b \). Then

\[
\int_{\mathcal{X}} g dF_n = \int_{\mathbb{R}} y H_n\{dy\} = \int_{\mathbb{R}} e dH_n \to \int_{\mathbb{R}} e dH = \int_{\mathcal{X}} g dF,
\]

establishing a).

\( F_n, \; n \geq 1, \) is said to converge weakly to \( F \) if any of the five equivalent conditions holds; and if \( X_n \) and \( X \) are random vectors with distributions \( F_n \) and \( F \), then \( X_n \) is said to converge in distribution to \( F \) iff \( F_n \) converges weakly to \( F \). These relations are denoted by \( F_n \Rightarrow F \) and \( X_n \Rightarrow X \). Thus, \( X_n \Rightarrow X \) iff

\[
\lim_{n \to \infty} E_n[g(x_n)] = E[g(X)]
\]

for all bounded (Lipschitz) continuous functions \( g : \mathcal{X} \to \mathbb{R} \), since \( E_n[g(X_n)] = \int_{\mathcal{X}} g dF_n \).

By the Lemma, weak limits are unique. Limits in distribution are not.

**Corollary.** If \( X_n \Rightarrow X \) and \( d(X_n, Y_n) \to 0 \) in probability, then \( Y_n \Rightarrow X \).

**Proof.** It \( g \) is any bounded Lipschitz continuous function, say \( |g| \leq c \) and \( |g(y) - g(x)| \leq K d(x, y) \), then \( E_n[g(X_n)] \to E[g(x)] \) by assumption, and

\[
|E_n[g(Y_n)] - E_n[g(X_n)]| \leq E_n|g(Y_n) - g(X_n)| \leq E_n|2c \wedge K d(X_n, Y_n)| \to 0,
\]

by the Dominated Convergence Theorem. So, \( E[g(Y_n)] \to E[g(X)] \), as asserted. \( \Box \)

Now let \( \mathcal{Y} \) be a second metric space, with metric \( \rho \) say; \( h : \mathcal{X} \to \mathcal{Y} \) be a Borel measurable function and let \( D_h = \{ x \in \mathcal{X} : h \text{ is not continuous at } x \} \}. \) Thus, if \( X \) has distribution \( F \), then \( Y = h(X) \) has distribution \( F \circ h^{-1} \).
Theorem 2. The Continuous Mapping Theorem. Let $h : \mathcal{X} \to \mathcal{Y}$ be Borel measurable. If $F_n \Rightarrow F$ and $F\{D_h\} = 0$, then $F_n \circ h^{-1} \Rightarrow F \circ h^{-1}$.

Proof. If $C \subseteq \mathcal{Y}$ is closed, then $h^{-1}(C)^- \subseteq h^{-1}(C) \cup D_h$. For if $x_n \in h^{-1}(C)$ and $x_n \to x \in \mathcal{X}$, then either $h(x_n) \to h(x) \in C$, so that $x \in h^{-1}(C)$, or $x \in D_h$. So,

$$\limsup_{n \to \infty} F_n \circ h^{-1} \{ C \} = \limsup_{n \to \infty} F_n \{ h^{-1}(C) \}$$

$$\leq \limsup_{n \to \infty} F_n \{ h^{-1}(C)^- \}$$

$$\leq F \{ h^{-1}(C)^- \}$$

$$\leq F \{ h^{-1}(C) \} + F \{ D_h \}$$

$$= F_n \circ h^{-1} \{ C \}$$

and, therefore, $F_n \circ h^{-1} \Rightarrow F \circ h^{-1}$. \hfill \Box

Remark. In fact, $D_h$ is a Borel set, so that the outer measure is not needed above.

Example. If $X_n \in IR^p$ are random vectors for which $X_n \Rightarrow X$, then $\| X_n \|^2 \Rightarrow \| X \|^2$. A direct proof of this would be difficult. \hfill \Box

Addendum to Portmanteau. A function $g : \mathcal{X} \to IR$ is said to be lower semi-continuous if

$$\liminf_{y \to x} g(y) \geq g(x)$$

for all $x \in \mathcal{X}$. That is, $f$ is l.s.c. if for each $x \in \mathcal{X}$ and $\epsilon > 0$, there is a $\delta > 0$ for which $g(y) \geq g(x) - \epsilon$ whenever $d(x, y) \leq \delta$.

Example. The indicator function of any open set is l.s.c.. To see this, let $U \subseteq \mathcal{X}$ be open and $g = 1_U$. If $x \in U'$, then $g(x) = 0$, so that (*) is clear; and if $x \in U$, then there is $\delta > 0$ for which $y \in U$ whenever $d(x, y) \leq \delta$, so that $1_U(y) = 1$ for such $y$. \hfill \Box

Lemma 5. $f$ is lower semi-continuous iff $f^{-1}(-\infty, c]$ is closed in $\mathcal{X}$ for every $c \in IR$.

Proof. Exercise

Lemma 6. If $g$ is l.s.c. and bounded below, then there are bounded continuous $g_n$ for which $g_n(x) \uparrow g(x)$ for all $x \in \mathcal{X}$.

Proof: The Idea. Let

$$h_n(x) = \inf \{ g(y) + nd(x, y) : y \in \mathcal{X} \}.$$ 

Then $g_n = h_n \wedge n$ has the desired properties. The details are left as an exercise. \hfill \Box

Proposition 1. If $F_n$ and $F$ are probability distributions, then $F_n \Rightarrow F$ iff

$$\liminf_{n \to \infty} \int_{\mathcal{X}} gdF_n \geq \int_{\mathcal{X}} gdF$$
for all l.s.c. $g$ that are bounded below.

**Proof.** Suppose first that the condition holds. If $g$ is any bounded continuous function, then $\pm g$ are bounded continuous functions that are bounded below. So,

$$\liminf_{n \to \infty} \int_\chi \pm gdF_n \geq \int_\chi \pm gdF$$

and, therefore,

$$\lim_{n \to \infty} \int_\chi gdF_n = \int_\chi gdF.$$

Conversely, suppose that $F_n \Rightarrow F$. If $g$ is bounded and lower semi-continuous, then there are bounded continuous $g_k$ for which $g_k \uparrow g$. Then

$$\liminf_{n \to \infty} \int_\chi gdF_n \geq \lim_{n \to \infty} \int_\chi g_kdF_n = \lim_{n \to \infty} \int_\chi g_kdF$$

for all $k$. So,

$$\liminf_{n \to \infty} \int_\chi gdF_n \geq \sup_{k \geq 1} \int_\chi g_kdF = \int_\chi gdF.$$  

**Problem 1.** Show that $D_h$ is a Borel set.

**Problem 2.** Show that the $h_n$ in Lemma 3 are continuous functions for which $h_n(x) \uparrow g(x)$ for all $x$. 

5