MATH 6605 LECTURE 15

JOE TRAN

Definition 1. Given Borel probability distributions $\{\mu, \mu_n\}_{n=1}^{\infty}$ on $\mathbb{R}$, we denote $\mu_n \rightharpoonup \mu$ and say that $\{\mu_n\}_{n=1}^{\infty}$ *converges weakly* to μ if

$$\lim_{n\to\infty} \int_{\mathbb{R}} f \, \mathrm{d}\mu_n = \int_{\mathbb{R}} f \, \mathrm{d}\mu$$

for all bounded continuous function $f : \mathbb{R} \to \mathbb{R}$.

Theorem 1. *Let $\{\mu, \mu_n\}_{n=1}^{\infty}$ be probability measures on $(\mathbb{R}, \mathfrak{B}(\mathbb{R})$. The following are equivalent definitions of the statement that $\{\mu_n\}_{n=1}^{\infty}$ converges weakly to μ.*

1. $\lim_{n\to\infty} E_{\mu_n}(f) = E_{\mu}(f)$ *for every bounded continuous function $f : \mathbb{R} \to \mathbb{R}$.*
2. $\lim_{n\to\infty} \mu_n(A) = \mu(A)$ *for every $A \in \mathfrak{B}(\mathbb{R})$ such that $\mu(\partial A) = 0$.*
3. $\lim_{n\to\infty} \mu_n((-\infty, x]) = \mu((-\infty, x])$ *for all x such that $\mu(\{x\}) = 0$.*
4. *(Skorohod's Theorem) There exists a probability space $(\Omega, \mathcal{F}, P)$ with random variables $\{Y, Y_n\}_{n=1}^{\infty}$ such that $\mathscr{L}(Y) = \mu$, $\mathscr{L}(Y_n) = \mu_n$ for all $n \in \mathbb{N}$ and $Y_n \to Y$ almost surely.*
5. $\lim_{n\to\infty} E_{\mu_n}(f) = E_{\mu}(f)$ *for every bounded Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $\mu(D_f) = 0$, where D_f is the set of all discontinuities of f.*

Proof. (5) $\Rightarrow$ (1): Immediate.

(5) $\Rightarrow$ (2): This follows by setting $f = \chi_A$ so that $D_f = \partial A$ and $\mu(D_f) = \mu(\partial A) = 0$. Then

$$\lim_{n\to\infty} \mu_n(A) = \lim_{n\to\infty} \int_A f \, \mathrm{d}\mu_n = \int_A f \, \mathrm{d}\mu = \mu(A)$$

(2) $\Rightarrow$ (3): Immediate, since the boundary of $(-\infty, x]$ is $\{x\}$.

(1) $\Rightarrow$ (3): Let $\varepsilon > 0$ be arbitrary, and let

$$f(t) = \begin{cases} 1 & \text{if } t \leq x \\ 0 & \text{if } t \geq x + \varepsilon \end{cases}$$

with f linear on the interval $(x, x+\varepsilon)$. Then f is continuous with

$$\chi_{(-\infty,x]} \leq f \leq \chi_{(-\infty,x+\varepsilon]}$$

and so

$$\limsup_{n\to\infty} \mu_n((-\infty, x]) \leq \limsup_{n\to\infty} \int_{(-\infty,x]} f \, \mathrm{d}\mu_n = \int_{(-\infty,x]} f \, \mathrm{d}\mu \leq \mu((-\infty, x+\varepsilon])$$

Date: March 7, 2025.

As $\varepsilon > 0$ was arbitrary, we have

$$\limsup_{n\to\infty} \mu_n((-\infty, x]) \leq \mu((-\infty, x])$$

Similarly, if g is the function defined by

$$g(x) = \begin{cases} 1 & \text{if } t \leq x - \varepsilon \\ 0 & \text{if } t \geq x \end{cases}$$

with g linear on the interval $(x-\varepsilon, x)$, then $\chi_{(-\infty, x-\varepsilon]} \leq g \leq \chi_{(-\infty, x]}$, and so

$$\liminf_{n\to\infty} \mu_n((-\infty, x]) \geq \liminf_{n\to\infty} \int_{(-\infty,x]} f \, \mathrm{d}\mu_n \geq \mu((-\infty, x-\varepsilon])$$

As $\varepsilon > 0$ was arbitrary, we have $\liminf\limits_{n\to\infty} \mu_n((-\infty, x]) \geq \mu((-\infty, x))$.

On the other hand, if $\mu(\{x\}) = 0$, then $\mu((-\infty, x]) = \mu((-\infty, x])$, so we have

$$\limsup_{n\to\infty} \mu_n((-\infty, x]) = \liminf_{n\to\infty} \mu_n((-\infty, x]) = \mu((-\infty, x])$$

(3) $\Rightarrow$ (4): We first define the cumulative distribution functions by $F_n(x) = \mu_n((-\infty, x])$ and $F(x) = \mu((-\infty, x])$. Then if we let $(\Omega, \mathcal{F}, P)$ be the Lebesgue measure on $[0,1]$, and let $Y_n(\omega) = \inf\{x : F_n(x) \geq \omega\}$, and $Y(\omega) = \inf\{x : F(x) \geq \omega\}$, then as in Lemma 7.1.2, we have $\mathscr{L}(Y_n) = \mu_n$ and $\mathscr{L}(Y) = \mu$. Note that if $F(z) < a$ then $Y(a) \geq z$, while if $F(\omega) \geq b$, then $Y(b) \leq z$.

Since $\{F_n\} \to F$ almost everywhere, it seems reasonable that $\{Y_n\} \to Y$ almost everywhere as well. We will show that $\{Y_n\} \to Y$ at points of continuity of Y. Then, since Y is non-decreasing, it can have a countable number of discontinuitis. Indeed, it has at most $m(Y(n+1) - Y(n)) < \infty$ discontinuities of size at least $1/m$ within the interval $(n, n+1]$, then take countable union over m and n. Since countable sets have Lebesgue measure 0, this implies that $\{Y_n\} \to Y$ with probability 1, proving (4).

Now suppose that Y is continuous at ω, let $y = Y(\omega)$. Then for any $\varepsilon > 0$, we claim that $F(y-\varepsilon) < \omega < F(y+\varepsilon)$. Indeed, if we had $F(y-\varepsilon) = \omega$, thens etting $\omega = y - \varepsilon$ and $b = \omega$ as above, this would imply that $Y(\omega) \leq y - \varepsilon = Y(\omega) - \varepsilon$, which is a contradiction. Similar when we consider $\omega = F(y+\varepsilon)$.

Next, let $\varepsilon > 0$ be arbitrary. Find a $\delta > 0$ such that $0 < \delta < \varepsilon$ such that $\mu(\{y - \delta\}) = \mu(\{y+\delta\}) = 0$. Then $F_n(y-\delta) \to F(y-\delta)$ and $F_n(y+\delta) \to F(y+\delta)$, so $F_n(y-\delta) < \omega < F_n(y+\delta)$ for sufficiently large n. Hence, $Y_n(\omega) \to Y(\omega)$.

(4) $\Rightarrow$ (5): If $\{Y_n\} \to Y$ and $Y \notin D_f$, then $\{f(Y_n)\} \to f(Y)$. It follows that $P(\{f(Y_n)\} \to f(Y)) \geq P(\{Y_n\} \to Y \text{ and } Y \notin D_f)$. But by assumption, $P(\{f(Y_n)\} \to Y) = 1$ and $P(Y \notin D_f) = \mu(D_f^c) = 1$, so $P(\{f(Y_n)\} \to f(Y)) = 1$. If f is bounded, then from the bounded convergence theorem, $E(f(Y_n)) \to E(f(Y))$. $\square$

Example 1. Let $\{U_n\}_{n=1}^\infty$ be iid random variables each with Uniform$[0,1]$ distribution and let $M_n = \min\{U_1, ..., U_n\}$. Its cdf is

$$F_{M_n}(t) = P(M_n \leq t) = \begin{cases} 1 & \text{if } t \geq 1 \\ 0 & \text{if } t \leq 0 \end{cases}$$

For $t \in (0,1)$, we have $F_{M_n}(t) = 1 - P(M_n > t) = 1 - (1-t)^n$, so $F_{M_n}(t) \to 1$ if $t > 0$ and $F_{M_n}(t) \to 0$ if $t \le 0$.

Proposition 1. *If $\{X_n\}_{n=1}^{\infty} \to X$ in probability, then $\mathscr{L}(X_n) \rightharpoonup \mathscr{L}(X)$.*

Proof. Let $\varepsilon > 0$ be arbitrary. If $X > z - \varepsilon$ and $|X_n - X| < \varepsilon$, then $X_n > z$. That is,

$$\{X \le z + \varepsilon\} \cup \{|X_n - X| \ge \varepsilon\} \supseteq \{X_n \le z\}$$

Hence, by the order-preserving property and subadditivity, $P(X_n \le z) \le P(X \le z + \varepsilon) + P(|X_n - X| \ge \varepsilon)$. Since $\{X_n\} \to X$ in probability, we have

$$\limsup_{n\to\infty} P(X_n \le z) \le P(X \le z + \varepsilon)$$

Letting $\varepsilon \to 0$, we have

$$\limsup_{n\to\infty} P(X_n \le z) \le P(X \le z)$$

Similarly, it can be shown that

$$\liminf_{n\to\infty} P(X_n \le z) \ge P(X < z)$$

Otherwise, if $P(X = z) = 0$, then $P(X < z) = P(X \le z)$, so we have

$$\liminf_{n\to\infty} P(X_n \le z) = \limsup_{n\to\infty} P(X_n \le z) = P(X \le z)$$

as claimed. $\square$