Lecture 14, Wed, 4/27/2022

Today 1. Narrow convergence of probability measures
2. Weak lower semi-conv. of $E_k[·]

Given: $X, Y$: Polish spaces, $\mu \in \mathcal{P}(X)$, $\nu \in \mathcal{P}(Y)$,
$C : X \times Y \to [0, \infty ]$: Borel measurable.

Def: $A(\mu, \nu ) = \{ \tau \in \mathcal{P}(X \times Y) : \tau \upharpoonright x = \mu , \tau \upharpoonright y = \nu \}$.

$E_k[\tau ] = \int_{x \times y} c(x, y) \, d\tau (x, y)$.

Theorem Assume $C : X \times Y \to [0, \infty ]$ is lower semi-
continuous. Then $\exists \gamma (A(\mu, \nu )) s.t. E_k[\gamma ] = \min_{\tau \in A(\mu, \nu )} E_k[\tau ]$.

Review of signed measures, narrow convergence
of probability measures, etc.

Let $(Z, d)$ be a Polish space (i.e., a complete and separable metric space).

Finite, signed Borel measures

$\mu : \mathcal{B}(Z) \to \mathbb{R}$ (not $\mathbb{R} \cup \{-\infty, +\infty \}$) is a finite signed
measure on $Z$ if $\mu$ is $\sigma$-additive: for any $E = \bigcup_{j} E_j$
disjoint, all $E_j \in \mathcal{B}(Z)$, $\sum_{j} \mu (E_j)$ converges absolutely,
and $\mu (\emptyset ) = 0$.

Denote

$\mathcal{M}(Z) = \{ \text{all finite signed measures on } Z \}$,
$\mathcal{M}_+(Z) = \{ \mu \in \mathcal{M}(Z) : \mu \geq 0 \}$,
$\mathcal{P}(Z) = \{ \mu \in \mathcal{M}_+(Z) : \mu (Z) = 1 \}$,
$\mu \in \mathcal{M}_+(Z) \iff \mu$ is a finite, Borel (positive) measure on $Z$. 
\(\mu \in \mathcal{P}(Z)\): \(\mu\) is a (Borel) probability measure on \(Z\).
\(\mathcal{P}(Z) \subseteq \mathcal{M}_+(Z) \subseteq \mathcal{M}(Z)\).
\(\mathcal{M}_+(Z)\) is a vector space. \(\mathcal{M}_+(Z)\) and \(\mathcal{P}(Z)\) are convex subsets of \(\mathcal{M}(Z)\).
\(\mathcal{M}(Z)\): \(\mu = \mu^+ - \mu^-\), \(\mu^\ast = \mu^+ + \mu^-\)
\(\mu^\ast\) is the total variation of \(\mu\).

**Proposition** \(\mathcal{M}(Z)\) is a normed vector space.
with \(\|\mu\| = \mu^\ast(1)\) if \(\mu \in \mathcal{M}(Z)\).

**Narrow convergence:** \(\mu_n, \mu \in \mathcal{P}(Z), \mu_n \rightarrow \mu\) narrowly if
\[
\int_Z \varphi \, d\mu_n \rightarrow \int_Z \varphi \, d\mu \quad \forall \varphi \in C_b(Z),
\]
where \(C_b(Z) = \{ f : Z \rightarrow \mathbb{R} \mid f\) is bounded and continuous\}.

Let \(\varphi \in C_b(Z)\). Define \(\|\varphi\|_\infty = \sup_{z \in Z} |\varphi(z)|\). Then
\((C_b(Z), \|\cdot\|_\infty)\) is a Banach space. If \(\mu \in \mathcal{M}(Z)\),
then \(\mu(\varphi) = \int_Z \varphi \, d\mu \quad (\varphi \in C_b(Z))\) defines \(\mu : C_b(Z) \rightarrow \mathbb{R}\) and \(\mu \in [C_b(Z)]^\ast\), \(\|\mu(\varphi)\| \leq \|\mu\| \|\varphi\|_\infty \quad \forall \varphi\).

If \(Z\) is locally compact, then by Riesz's Theorem,
\(\mathcal{M}(Z) = [C_0(Z)]^\ast\) (\(\varphi \in C_0(Z)\) means \(\varphi : Z \rightarrow \mathbb{R}\) is continuous, and \(\forall \varepsilon > 0, \{ z \in Z : |\varphi(z)| \geq \varepsilon \} \) is compact).
In particular, in this case, \(\mu : C_0(Z) \rightarrow \mathbb{R}\),
defined by \(\mu(\varphi) = \int \varphi \, d\mu\), satisfies \(\|\mu\|_{C_0(Z)} = \|\mu\|_M\) \((= \|\mu^\ast\|_1(Z))\).

In general, \((Z\) is a Polish space.) \(\forall \mu \in \mathcal{M}_+(Z), \mu \in C_b(Z) \rightarrow \mathbb{R}\) satisfies \(\|\mu\|_\infty = \|\mu\ast\|, \) since \(\|\mu\ast\| = \|\mu\ast\|_1\).
and choosing \( \varphi = 1 \), we get \( \| \varphi \|=1 \). Hence \( \| \mu \|=1 \). Thus, \( M_+(Z) \subseteq [C_0(Z)]^* \). In particular, \( P(Z) \) is a unit sphere of \( [C_0(Z)]^* \), which is compact w.r.t. weak-* topology (Banach-Alaoglu Thm). This weak-* topology is defined by the family of seminorms \( \{ p_\varphi : \varphi \in C_0(Z) \} \), where \( p_\varphi (f) = |f(\varphi)| \) for all \( \varphi \in [C_0(Z)]^* \). With the weak-* topology, \( [C_0(Z)]^* \) is a locally convex Hausdorff topological vector space. Call this topology on \( M_+(Z) \subseteq [C_0(Z)]^* \) the narrow topology.

Therefore, the narrow convergence of \( \mu_k \in P(Z) \) to \( \mu \in P(Z) \) is weak-* convergence w.r.t. to the weak-* topology of \( [C_0(Z)]^* \), the narrow topology. If \( \mu \in P(Z) \) and \( \mu \in M_+(Z) \) are such that
\[
\int \varphi d\mu_k \to \int \varphi d\mu \quad \forall \varphi \in C_0(Z),
\]
then, choosing \( \varphi = 1 \), we get \( \mu_k(Z) \to \mu(Z) \). Hence, \( \mu(Z) = 1 \), and \( \mu \in P(Z) \). If \( \exists \lambda \in [C_0(Z)]^* \) such that
\[
\lambda \mu_k \to \lambda \mu \text{ in the narrow topology},
\]
then, \( \lambda \mu_k \to \lambda \mu \) in the narrow topology, i.e., \( \forall \varphi \in C_0(Z), \lambda(\varphi) = \lim_{k \to \infty} \int \varphi \lambda d\mu_k \). Then, \( \lambda \mu \) may not be in \( P(Z) \).

Note: the narrow convergence is not the weak-* convergence w.r.t. \( [C_0(Z)]^* \). If \( Z \) is locally compact, then \( [C_0(Z)]^* = M_+(Z) \) (Riesz Thm). Then for \( \mu \in P(Z) \), if convergent to some \( \mu \in M_+(Z) \), then \( \mu \) may not be in \( P(Z) \).
Example. Z = \{0\}, \mu_k = \delta_{k^r} (\mathcal{P}(\{0\}) (k \in \{1, 2, \ldots\}). \mu_k \to 0 \text{ weak-* w.r.t. } [C_0(Z)]^*, \text{i.e., } \forall \varphi \in C_0(Z), \frac{\int Z \varphi d\mu_k = \varphi(k)}{\to 0}.

In the case Z is compact, \( C_0(Z) = C(Z) \) and \( \mathcal{M}(Z) = [C_0(Z)]^* \). Hence \( \mathcal{M}(Z)^* = [C_0(Z)]^{**} \supseteq C_0(Z) \). Hence, the narrow convergence is also the weak convergence.

It turns out that the weak-* topology defined by seminorms \( \{ P_\varphi : \varphi \in C_0(Z) \} \), i.e., the narrow topology, on \( \mathcal{P}(Z) \) is metrizable. The metric is the so-called Prokhorov metric, \( d_P(\cdot, \cdot) \), defined for \( \mu, \nu \in \mathcal{P}(Z) \) by

\[
d_P(\mu, \nu) = \inf \left\{ \varepsilon > 0 : \mu(A) \leq \nu(A + \varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A - \varepsilon) + \varepsilon \text{ for all } A \subset \mathcal{B}(Z) \right\},
\]

where \( A + \varepsilon = \{ x \in Z : d(x, A) < \varepsilon \} \) if \( A \neq \emptyset \), and \( \emptyset + \varepsilon = \emptyset \) for all \( \varepsilon > 0 \).

The metric is also given by the Wasserstein metric, a family of countable members in \( C_0(Z) \). See Ambrosio-Gigli-Savaré (2008): 335.1, and the metric \( \rho \) (Lecture 17).

**Theorem** If \( Z \) is a Polish space, then \( \mathcal{P}(Z) \) is a Polish space with respect to the narrow topology. The convergence \( \mu_k \rightharpoonup \mu \) (\( \mu_k, \mu \in \mathcal{P}(Z) \)) is exactly the narrow convergence.

\( \text{Q.E.D.} \)

Let us collect some results on the narrow convergence of probability measures.
Theorem Let \( Z \) be a metric space, \( \mathcal{M}(Z) \), and \( \mathcal{A}(Z) \).
Then \( \mu(A) = \inf \{ \mu(U) : U \subseteq Z : \text{open}, U \supseteq A \} \),
\[ \mu(A) = \sup \{ \mu(F) : F \subseteq Z : \text{closed}, F \supseteq A \} . \]

Proof Let \( \mathcal{R} \) be the collection of \( A \in \mathcal{B}(X) \) such that
\[ \mu(A) = \inf \{ \mu(U) : A \subseteq U : \text{open} \} \]
and
\[ \mu(A) = \sup \{ \mu(F) : A \supseteq F : \text{closed} \} . \]
Note that \( \phi \in \mathcal{R} \) and for any \( A \in \mathcal{R} \) and any \( \varepsilon > 0 \), there exist \( F \subseteq Z : \text{closed}, U \subseteq Z : \text{open} \) such that \( F \subseteq A \subseteq U \) and \( \mu(U \setminus F) < \varepsilon \).

We can verify that \( \mathcal{R} \) is a \( \sigma \)-algebra, and \( \mathcal{R} \) contains all open sets of \( Z \). Hence, \( \mathcal{R} = \mathcal{B}(X) \).

\( \Omega \) \( \Omega \)

Corollary Let \( Z \) be a metric space and \( \mu, \nu \in \mathcal{M}(Z) \).
Then \( \mu = \nu \iff \mu(U) = \nu(U) \) for all open sets \( U \subseteq Z \)
\( \iff \mu(F) = \nu(F) \) for all closed sets \( F \subseteq Z \).

The above theorem and Ulam's lemma (see next lecture) imply the following.

Theorem If \( Z \) is a Polish space then any \( \mu \in \mathcal{P}(Z) \) is regular, i.e., for any \( A \in \mathcal{B}(Z) \),
\[ \mu(A) = \inf \{ \mu(U) : U \subseteq Z : \text{open}, U \supseteq A \} , \]
\[ \mu(A) = \sup \{ \mu(K) : K \subseteq Z : \text{compact}, K \supseteq A \} . \]

The following is a useful result about convergence of probability measures on metric spaces:

Theorem Let \( Z \) be a metric space and all \( \mu_k \),
\( \mu_k \in \mathcal{P}(Z) \) \( (k=1,2,\ldots) \). The following are equivalent:
(1) \( \mu_k \to \mu \) narrowly;
(2) \( \int_Z \varphi \, d\mu_k \to \int_Z \varphi \, d\mu \) for all \( \varphi \in C_b(Z) \): bounded Lipschitz functions
(3) \( \liminf_{k \to \infty} \mu_k(F) \geq \mu(F) \) for all closed \( F \subseteq Z \);
(4) \( \limsup_{k \to \infty} \mu_k(U) \leq \mu(U) \) for all open \( U \subseteq Z \);
(5) \( \mu_k(A) \to \mu(A) \) \( \forall A \in \mathcal{B}(Z) \) with \( \mu(\partial A) = 0 \).

**Proof.** See Dudley's book. Q.E.D

**Example** Let \( Z = \mathbb{R} \), \( \mu_k = \delta_{x_k} (k = 1, 2, \ldots) \), \( \mu = \delta_0 \), \( F = (-\infty, 0] \), and \( U = (0, \infty) \). Then, \( F \) is closed and \( U \) is open.

\[
\liminf_{k \to \infty} \mu_k(F) = 0 < \mu(F) = 1.
\]

\[
\limsup_{k \to \infty} \mu_k(U) = 1 > \mu(U) = 0.
\]

**Example** \( Z = [0, 1] \), Euclidean metric. \( \mu_k = \frac{1}{k} \sum_{i=1}^{k} \delta_{x_i} \). If \( \varphi \in C_b(Z) \) then
\[
\int_Z \varphi \, d\mu_k = \frac{1}{k} \sum_{i=1}^{k} \varphi(x_i) \to \int_0^1 \varphi \, d\mu.
\]

Thus, \( \{\mu_k\} \) converges to the Lebesgue measure \( \mu \) on \([0, 1] \) (a probability measure) narrowly. But, if \( A \) is the set of all irrational numbers in \([0, 1] \), then \( \mu(A) = 1 \). But, \( \mu_k(A) = 0 \) \( \forall k \). Thus, \( \mu_k(A) \not\to \mu(A) \).

We now prove the weak-\( \epsilon \) (i.e., narrow) lower semi-continuity of the \( k \)-cost functional, a result used in the proof of the existence theorem.

**Theorem** If \( C : X \times Y \to [0, \infty] \) is lower semi-continuous, then \( E_k : \mathcal{P}(X \times Y) \to [0, \infty] \) is lower semi-continuous w.r.t. the narrow topology of \( \mathcal{P}(X \times Y) \).

**Proof** This follows from a lemma in last lecture. Here, we give a direct proof.
We construct $c_k \in C_b(X \times Y)$:

$$c_k(x, y) = \inf_{x', y', \epsilon} \left\{ c(x', y') \wedge k + k d(x, x') + k d(y, y') \right\}.$$ 

We have

1. $0 \leq c_k \leq c \wedge k = \min(c, k)$.
2. $c_k \in C_{\text{lip}}(X \times Y)$, as $\inf$ of a family of $C_{\text{lip}}$ functions with a uniform $C_{\text{lip}}$ constant (with fixed $k$).
3. $c_k \uparrow c$ (pointwise increasing, converging to $c$).

For the last point: Fix $(x, y) \in X \times Y$, assume W.l.o.g. $A(x, y) = \sup_k c_k(x, y) = \infty$. Then, $\forall k$, let $(x_k, y_k)$ be s.t.

$$c_k(x_k, y_k) = k d(x, x_k) + k d(y, y_k) \leq c_k(x, y) + \frac{1}{k}.$$ 

Hence, $k d(x, x_k) \leq A(x, y) + \frac{1}{k}$. So $d(x, x_k) \to 0$. Similarly, $d(y, y_k) \to 0$. Also,

$$c(x_k, y_k) \wedge k \leq c_k(x, y) + \frac{1}{k}.$$ 

By the lower semicont. of $c$, and the fact that $c_k \uparrow$, we have

$$c(x, y) \leq \liminf_{k \to \infty} c_k(x_k, y_k) = \liminf_{k \to \infty} c(x_k, y_k) \wedge k \leq \liminf_{k \to \infty} \left( c_k(x, y) + \frac{1}{k} \right) = \lim_{k \to \infty} c_k(x, y) = c(x, y).$$

Next, suppose $\gamma_j \in \mathcal{A}(U, V)$ (j = 1, 2, ...), $\gamma \in \mathcal{P}(X \times Y)$, satisfy $\gamma_j \to \gamma$ n. a. in $\mathcal{P}(X \times Y)$. Then,

$$\liminf_{j \to \infty} \mathbb{E}_k[\gamma_j] = \liminf_{j \to \infty} \int_{x \times y} c_k(x, y) \mathbb{E}_k^\gamma(x, y) \, dx \times dy \geq \liminf_{j \to \infty} \int_{x \times y} c_k(x, y) \mathbb{E}_k^\gamma(x, y) \, dx \times dy = \int_{x \times y} c_k dx \times dy.$$ 

Finally, by the monotone convergence theorem,

$$\liminf_{j \to \infty} \mathbb{E}_k[\gamma_j] \geq \sup_k \int_{x \times y} c_k dx \times dy = \mathbb{E}_k[\gamma]. \quad \square$$