1 Topics on Probability (MATH 289 A): Lecture 2

- Date: 10/03/2017
- Lecturer: Prof. Ruth Williams.
- Scribe: Felipe Campos.

Consider the space $\mathbb{C}^d := \mathbb{C}([0,\infty), \mathbb{R}^d) := \{f : [0,\infty) \longrightarrow \mathbb{R}^d \mid f \text{ is continuous }\}$. Then, we can endow $\mathbb{C}^d$ with the metric:

$$d(x,y) = \sum_{n=1}^{\infty} 2^{-n}(||x-y||_n \wedge 1) \tag{1.1}$$

where $||x||_n = \sup_{t\in[0,n]} |x(t)|$. We can prove that $(\mathbb{C}^d, d)$ is a Polish space (metric, separable and complete).

In a similar way, consider $\mathbb{D}^d := \mathbb{D}([0,\infty), \mathbb{R}^d)$ as the space of functions $f : [0,\infty) \longrightarrow \mathbb{R}^d$ such that f is right-continuous with left limits. Then, we endow this set with the metric m_{J_1}, which induces the Skorohod-J_1 topology.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{X^n\}_{n\geq 1}$ a sequence of stochastic processes with paths in $\mathbb{D}^d$. We consider $\mathcal{D}^d$ as the σ-algebra on $\mathbb{D}^d$ associated with the J_1-topology. This is the smallest σ-algebra containing all open sets for the J_1-topology.

The fact that X^n is measurable as a map to $\mathbb{D}^d$ ($(X^n)^{-1}(A) := \{w \in \Omega \mid X^n(\omega) \in A\} \in \mathcal{F}$ for all $A \in \mathcal{D}^d$.) is equivalent to: $\forall t > 0$ $X^n_t : \Omega \longrightarrow \mathbb{R}^d$ is measurable and $t \mapsto X^n_t(\omega)$ is in $\mathbb{D}^d$ for each $\omega \in \Omega$.

Definition 1.1 (Law of X^n) *Define $\pi_n(A) := \mathbb{P}((X^n)^{-1}(A))$ for $A \in \mathcal{D}^d$. Then, π_n is a probability measure on $(\mathbb{D}^d, \mathcal{D}^d)$ and it is called the **law** or **distribution** of X^n. In the same way, we consider $\pi(A) = \mathbb{P}(X^{-1}(A))$ for all $A \in \mathcal{D}^d$.*

1.1 Convergence in Distribution of $\{X^n\}$ to X

Consider $\{X^n\}_{n\geq 1}$ and X processes with paths in $\mathbb{D}^d$.

Definition 1.2 *We will say that $\{X^n\}$ **converges in distribution or in law** to X (denoted $X^n \Rightarrow X$) if for the corresponding laws $\{\pi_n\}_{n\geq 1}$ and π, we have that:*

$$\pi_n \rightharpoonup \pi \qquad \text{as } n \to \infty, \tag{1.2}$$

which means that: $\forall\ f : \mathbb{D}^d \longrightarrow \mathbb{R}$ continuous and bounded, we have that $\int_{\mathbb{D}^d} f d\pi_n \longrightarrow \int_{\mathbb{D}^d} f d\pi$ as $n \to \infty$.

The set of probability measures on $\mathbb{D}^d$ (or $\mathbb{C}^d$) with the topology of weak convergence is a Polish space. In other words, there exists a metric on the space that makes it separable, complete and convergence in the metric is weak convergence.

We have other types of convergence for stochastic processes.

Definition 1.3 *For d-dimensional stochastic processes* $\{X^n\}_{n\geq 1}$ *and* X *defined on* $(\Omega, \mathcal{F}, \mathbb{P})$, *we will say that* X^n ***converges to*** X ***almost surely*** *if:*

$$X^n(\omega) \longrightarrow X(\omega) \text{ in } \mathbb{D}^d \text{ as } n \to \infty \tag{1.3}$$

for $\mathbb{P}$*-a.e.* $\omega \in \Omega$. *Also, we will say that* X^n ***converges in probability to*** X *if for each* $\epsilon > 0$ *we have that:*

$$\mathbb{P}(m_{J_1}(X^n, X) > \epsilon) \longrightarrow 0 \tag{1.4}$$

as $n \to \infty$.

1.2 Criteria for Weak Convergence

Definition 1.4 *A set* $\mathcal{B}$ *of probability measures on* $(\mathbb{D}^d, \mathcal{D}^d)$ *is* ***(weakly) relatively compact*** *if each sequence* $\{\pi_n\}_{n=1}^{\infty}$ *in* $\mathcal{B}$ *has a weakly convergent subsequence with limit that is a probability measure on* $(\mathbb{D}^d, \mathcal{D}^d)$.

Definition 1.5 *A set* $\mathcal{B}$ *of probability measures on* $(\mathbb{D}^d, \mathcal{D}^d)$ *is* ***tight*** *if* $\forall \epsilon > 0$ *there is a compact set* A *in* $\mathbb{D}^d$ *such that:*

$$\pi(A) > 1 - \epsilon \qquad \forall \pi \in \mathcal{B}. \tag{1.5}$$

For a sequence of real random variables $\{X^n\}_{n\geq 1}$, their laws will be tight if $\forall \epsilon > 0,\ \exists K_\epsilon > 0$ such that:

$$\mathbb{P}(|X^n| \leq K_\epsilon) > 1 - \epsilon \tag{1.6}$$

for all $n \geq 1$.

The next theorem relates the previous concepts.

Theorem 1.6 (Prohorov's Theorem) *A set of probability measures* $\mathcal{B}$ *in* $(\mathbb{D}^d, \mathcal{D}^d)$ *is (weakly) relatively compact if and only if it is tight.*

The next theorem gives us a very useful method to prove convergence.

Theorem 1.7 *Suppose* $\{\pi_n\}_{n=1}^{\infty}$ *is a tight sequence of probability measures on* $(\mathbb{D}^d, \mathcal{D}^d)$. *Assume each weakly convergent subsequence of* $\{\pi_n\}_{n=1}^{\infty}$ *has limit* π. *Then,* $\{\pi_n\}_{n=1}^{\infty}$ *converges weakly to* π.

Proof We argue by contradiction. Suppose $\{\pi_n\}_{n=1}^{\infty}$ does not converges weakly to π. Then, there must be some $\epsilon > 0$ and some subsequence $\{\pi_{n_k}\}_{k\geq 1}$ such that:

$$d(\pi_{n_k}, \pi) > \epsilon \qquad \forall k \geq 1 \tag{1.7}$$

where d is the metric for the space of probability measures. Notice that $\{\pi_{n_k}\}_{k=1}^{\infty}$ is tight, and therefore, has a convergent subsequence $\{\pi_{n_{k_l}}\}_{l=1}^{\infty}$. But, this limit should be π, which contradicts (1.7).

$\square$

We now look at some criteria for tightness in $\mathbb{D}^d$.

Theorem 1.8 *A sequence of probability measures $\{\pi_n\}_{n=1}^{\infty}$ on $(\mathbb{D}^d, \mathcal{D}^d)$ is tight if and only if $\forall T > 0, \epsilon > 0$ there exists $K_\epsilon, \delta_\epsilon > 0$ such that:*

(i) $\limsup_{n\to\infty} \pi_n(x \in \mathbb{D}^d : ||x||_T \geq K_\epsilon) < \epsilon$

(ii) $\limsup_{n\to\infty} \pi_n(x \in \mathbb{D}^d : w'(x, \delta_\epsilon, T) \geq \epsilon) < \epsilon$

where w' is the **modified modulus of continuity**, defined as:

$$w'(x, \delta, T) = \inf_{\{t_i\}} \max_{i=1\ldots m} \sup_{s,t\in[t_{i-1},t_i)} |x(s) - x(t)| \tag{1.8}$$

where $\{t_i\}$ ranges over all partitions of the form $t_0 = 0 < t_1 < \ldots < t_m = T$ where $\min(t_i - t_{i-1}) > \delta$ for $i = 1, \ldots, m$.

Theorem 1.9 (Skorohod's Representation Theorem) *Suppose $\{\pi_n\}_{n=1}^{\infty}$ and π are probability measures on $(\mathbb{D}^d, \mathcal{D}^d)$ such that $\pi_n \rightharpoonup \pi$. Then, there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which stochastic processes $\{X^n\}_{n\geq 1}$ and X are defined in such a way that X^n has law π_n for each $n \geq 1$, X has law π, and $X^n \longrightarrow X$ almost surely as $n \to \infty$.*

Theorem 1.10 (Aldous) *Let $\{X^n\}_{n=1}^{\infty}$ be a sequence of stochastic processes with paths in $\mathbb{D}^d$. Then, the associated probability measures $\{\pi_n\}_{n=1}^{\infty}$ are tight if $\forall T > 0$:*

(i) $\lim_{K\to\infty} \limsup_{n\to\infty} \mathbb{P}(||X^n||_T \geq K) = 0$ *(compact containment).*

(ii) *$\forall \epsilon, \eta > 0$ there exists $\delta_{\epsilon,\eta}, n_{\epsilon,\eta} > 0$ such that for all $0 < \delta \leq \delta_{\epsilon,\eta}$ and $n \geq n_{\epsilon,\eta}$,*

$$\sup_{\tau_n\in\mathcal{T}^n[0,T]} \mathbb{P}(|X^n(\tau_n + \delta) - X^n(\tau_n)| \geq \epsilon) \leq \eta \tag{1.9}$$

where $\mathcal{T}^n[0, T]$ are all the finite-valued X^n-stopping times, i.e., an element of this set is a stopping time that only takes values in a finite set.

The next theorem includes the case where X has paths in $\mathbb{C}^d$.

Theorem 1.11 *Let $\{X^n\}_{n=1}^{\infty}$ be processes with paths in $\mathbb{D}^d$. The associated probability measures $\{\pi_n\}_{n=1}^{\infty}$ will be tight and any limit point will be concentrated on $\mathbb{C}^d$ if and only if $\forall T, \epsilon > 0$:*

(i) $\lim_{K\to\infty} \limsup_{n\to\infty} \mathbb{P}(||X^n||_T \geq K) = 0$ *(compact containment).*

(ii) $\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{P}(w(X^n, \delta, T) \geq \epsilon) = 0$.

where w is the **modulus of continuity**:

$$w(x, \delta, T) = \sup_{s,t \in [0,T], |s-t| < \delta} |x(s) - x(t)| \tag{1.10}$$