1 February 24th, 2022

1.1 Relative Compactness and Tightness of $\{\mu_n\}$

Definition 1.1 (**relative compactness**)**.** We say $\{\mu_n\}$ is relatively compact if for any subsequence of μ_n, say μ_{n_k}, one can find a further subsequence $\mu_{n_{k_i}}$ which converges weakly.

Proposition 1.2

If $\mu_n \implies \mu$ in the finite-dimensional sense and if $\{\mu_n\}$ is relatively compact, then $\mu_n \implies \mu$ ($\mu_n \to \mu$ weakly).

Proof. Finite-dimensional convergence means that $\mu_n \circ \Pi^{-1}_{t_1,\ldots,t_k} \implies \mu \circ \Pi^{-1}_{t_1,\ldots,t_k}$ for any $k \in \mathbb{N}$ and any given $t_1, \ldots, t_k \in [0,1]$. Relative compactness means that for any $\{\mu_{n_k}\}_k$, we have $\mu_{n_{k_i}} \implies \nu_{k_i}$ which is a probability measure. But we already know that:

$$\mu_{n_{k_i}} \circ \Pi^{-1}_{t_1,\ldots,t_n} \implies \mu \circ \Pi^{-1}_{t_1,\ldots,t_k}$$

On the other hand, the finite-dimensional distribution determines the measure, meaning that: $\nu_{k_i} = \mu$. This further means that for any subsequence $\{\mu_{n_k}\}_k$ there exists a further $\{\mu_{n_k}\}_i$, s.t. $\mu_{n_{k_i}} \implies \mu$, $i \to \infty$. meaning that $\mu_n \implies \mu$. □

Definition 1.3 (**tightness**)**.** We say $\{\mu_n\}$ is tight if for any $\epsilon > 0$, there exists a compact set $K = K_\epsilon \subset C$, s.t.:

$$\mu_n(K) \geq 1 - \epsilon, \quad n \geq n_0$$

Remark 1.4 — Roughly speaking $\{\mu_m\}_n$ are almost supported on a compact set of C.

Theorem 1.5 (**Prokhorov's Theorem**)

tightness $\iff$ relative compactness.

The heuristic is that the space of a measure on a compact space is also compact. The proof is very lengthy and won't be introduced in this course.

Remark 1.6 — If we are considering $\mathbb{R}$ instead of C, the relation between relative compactness and tightness is easy to understand. Tightness means that it is supported by a bounded and closed set. For example if we consider $\mu_n = \frac{1}{3}\delta_0 + \frac{2}{3}\delta_n$. This is not tight, since the mass will escape to ∞, it is also not relatively compact, since it's limit does not go to a probability measure.

Now, we will see how to show tightness. We will first use the

Theorem 1.7 (**Arzela-Ascoli Theorem**)
We consider a set $A \subset C[0,1]$ is relatively compact if and only if:

1. $\sup_{x\in A} |x(0)| < \infty$ (uniform boundedness)
2. $\lim_{\delta\to 0} \sup_{x\in A} \omega_x(\delta) = 0$, where $\omega_x(\delta) = \sup_{|s-t|\le\delta} |x(s)-x(t)|$, $0<\delta\le 1$ which is known as the **modulus of continuity**. (uniform equicontinuous)

Definition 1.8 (**equicontinuous** at a point t_0)**.** For all $x \in A, \forall\epsilon > 0, \exists\delta$ s.t.:

$$\sup_{x\in A} |x(t)-x(t_0)| \le \epsilon \quad \forall |t-t_0| \le \delta$$

is called equicontinuous.

Definition 1.9 (**uniformly equicontinuous**)**.** $\forall\epsilon > 0, \exists\delta$ s.t.:

$$\sup_{x\in A} \sup_{|t-s|\le\delta} |x(t)-x(t_0)| \le \epsilon.$$

Theorem 1.10
A sequence of probability measures $\{\mu_n\}$ on C is tight if and only if:

1. For any $\eta > 0$, there exists an a and n_0 s.t.:
$$\mu_n\{(x : |x(0)| \le a)\} \ge 1-\eta, \quad n \ge n_0.$$
2. For each $\epsilon > 0$ and $\eta > 0$, there exists a δ and n_0 s.t. :
$$\mu_n(\{x : \omega_x(\delta) \le \epsilon\}) \ge 1-\eta, \quad n \ge n_0.$$

We can take the intersection and closure of 1 and 2 to get a compact set by Arzela-Ascoli theorem. Proof is omitted.

We can write 1 and 2 as:

$$\mathbb{P}(|X_n(0)| > a) \le \eta \text{ and } \mathbb{P}(\omega_{X_n}(\delta) > \epsilon) \le \eta.$$

and this is how we can typically prove tightness, for example using Markov inequality. However, we will stop the discussion of tightness here.

1.2 Examples of Convergence to Gaussian Process

We will consider two examples, the first of the weak convergence of random walk to Brownian motion, and then briefly introduce a second example of the convergence of empirical processes.

Definition 1.11 (**Brownian Motion on $[0,1]$**)**.** A 1D Brownian motion is a real valued random function $X(t), t \in [0,1]$ s.t. $X(0) = 0$ and:

1. If $0 = t_0 < t_1 < \ldots < t_k$, then $X(t_1) - X(t_0), \ldots, X(t_k) - X(t_{k-1})$ are independent.
2. If $s, t \geq 0$, then: $X(s+t) - X(s) \sim N(0, t)$
3. With probability 1, $X(t)$ is continuous.

Note that 1 and 2 give you the finite-dimensional distribution. Fixing on a rectangle, we have:

$$\begin{aligned}\mathbb{P}((X(t_1), \ldots, X(t_k)) \in A_1 \times \ldots A_k) &= \mu_x \circ \Pi^{-1}_{t_1, \ldots t_k}(A_1 \times \ldots \times A_k) \\ &= \int_{A_1} dx_1 \ldots \int_{A_k} dx_k \prod_{m=1}^{k} \beta_{t_m - t_{m-1}}(x_{m-1}, x_m),\end{aligned}$$

where:

$$\beta_t(a, b) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(b-a)^2}{2t}\right)$$

Example 1.12

If $k = 2$, then:

$$f_{X(t_1), X(t_2)}(x_1, x_2) \propto \exp\left(-\frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix} \Gamma^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right)$$

with $\Gamma = \begin{bmatrix} t_1 & t_1 \\ t_1 & t_2 \end{bmatrix}$ $t_1 < t_2$.

Definition 1.13 (Gaussian process). A random function whose finite-dimensional distributions are all multivariate Gaussian.

Note that Brownian motion is a Gaussian process with covariance. Also, once we identify the covariance, the process is determined.

Theorem 1.14 (Donsker's invariance principle (functional CLT))

Let $\xi, \ldots \xi_n$ be i.i.d. with $\mathbf{E}\xi_i = 0$, $\mathbf{Var}\xi_i = 1$. Let $S_n = \sum_{i=1}^{n} \xi_i$, $S_0 = 0$. Define a random function in $C[0, 1]$:

$$X_n(t) = \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor} + (nt - \lfloor nt \rfloor) \frac{1}{\sqrt{n}} \xi_{n+1}$$

note that this is rescaling the trajectory. Under the above assumption:

$$X_n(t) \implies X(t) \leftarrow \text{Brownian motion.}$$

Proof. We need to prove the finite-dimensional convergence and tightness. Note that tightness will be omitted, since it is case by case basis. For the finite-dim convergence, let us denote $\varphi_{nt} = (nt - \lfloor nt \rfloor)\xi_{\lfloor nt \rfloor + 1} / \frac{1}{\sqrt{n}} \to 0$. We have:

$$
\begin{aligned}
(X_n(s), X_n(t) - X_n(s)) =& \frac{1}{\sqrt{n}}(S_{\lfloor ns \rfloor, S_{\lfloor nt \rfloor} - S_{\lfloor ns \rfloor}})) + (\varphi_{ns}, \varphi_{nt} - \varphi_{ns}) \\
&\Longrightarrow (\underbrace{N_1}_{N(0,s)}, \underbrace{N_2}_{N(0,t-s)}) \\
\Longrightarrow \ (X_n(s), X_n(t)) &\Longrightarrow (N_1, N_1 + N_2)
\end{aligned}
$$

The extension to k components is straightforward. $\square$