Universidad de Guanajuato
CIMAT

Proyecto Final

Teorema de Donsker

Autores:

Roberto Vásquez Martínez

Profesor:

Dr. Ehyter M. Martín
González

8 de junio de 2021

Índice

1. Nota Histórica

2. Introducción

3. El espacio C

Empezaremos presentando al espacio de funciones continuas y algunas propiedades que utilizaremos para describir las medidas de probabilidad en ese espacio y en consecuencia los procesos estocásticos a tiempo continuo.

A lo largo de este escrito consideraremos $C = C[0,1]$, al ser $[0,1]$ compacto y separable en $\mathbb{R}$ entonces $C[0,1]$ es un espacio métrico separable y completo con la métrica

$$d(f,g) = \|f-g\|_\infty = \sup_{t\in[0,1]} |f(t)-g(t)|$$

Al ser C un espacio métrico sea $(C,\mathcal{C})$ es el espacio medible correspondiente donde $\mathcal{C}$ es la σ-álgebra de Borel inducida por le métrica anterior, además denotaremos por $\mathcal{P}(\mathcal{C})$ el conjunto de medidas de probabilidad en el espacio medible $(C,\mathcal{C})$.

Para cada $f \in E$ y $\delta > 0$ definimos el módulo de continuidad como

$$\omega_f(\delta) = \omega(f,\delta) = \sup_{\substack{s,t\in[0,1] \\ |s-t|\le\delta}} |f(s)-f(t)|,$$

esta medida nos dice el grado de continuidad uniforme de f.

Por otro lado, para $0 \le t_1 < \ldots < t_k \le 1$ definimos las proyecciones naturales $\pi_{t_1,\ldots,t_k} : \mathrm{C} \to \mathbb{R}^k$ como

$$\pi_{t_1,\ldots,t_k}(g) = (g(t_1),\ldots,g(t_k)),$$

lo que nos permite definir la clase de conjuntos finito dimensionales en el espacio C

$$\mathcal{C}_f = \{\pi^{-1}_{t_1,\ldots,t_k}(H) \ : \ H \in \mathcal{B}(\mathbb{R}^k)\}$$

A continuación enunciamos un resultado que relaciona a $\mathcal{C}$ y $\mathcal{C}_f$.

Teorema 3.1

Con la notación anterior se cumple que

$$\mathcal{C} = \sigma(\mathcal{C}_f),$$

además $\mathcal{C}_f$ es clase separante en $\mathcal{P}(\mathcal{C})$.

El Teorema 3.1 nos dice que $\mathcal{C}_f$ es clase separante en $\mathcal{P}(\mathcal{C})$ pero no es clase determinante de convergencia, a continuación investigaremos una condición suficiente para que $\mathcal{C}_f$ sea clase determinante de convergencia, lo que nos ayudará a probar el Teorema de Donsker.

La condición suficiente que estamos buscando para que $\mathcal{C}_f$ sea clase determinante de convergencia para la sucesión de medidas de probabilidad $\{\mathbb{P}_n : n \in \mathbb{N}\}$ es la compacidad de relativa y como (C, d) es un espacio separable completo, por el Teorema de Prohorov la compacidad relativa de $\{\mathbb{P}_n : n \in \mathbb{N}\}$ en $\mathcal{P}(\mathcal{C})$ es equivalente a que $\{\mathbb{P}_n : n \in \mathbb{N}\}$ sea una familia tensa de medidas de probabilidad.

Bajo esta directriz, enunciamos el Teorema de Arzelà-Ascoli, que caracteriza la compacidad relativa en el espacio C.

Teorema 3.2 (*Arzelà-Ascoli*)

Un conjunto $A \subset C$ es relativamente compacto si y sólo si se cumple las siguientes dos condiciones

(i)
$$\sup_{f \in A} |f(0)| < \infty,$$

(ii)
$$\lim_{\delta \to 0} \sup_{f \in A} \omega_f(\delta) = 0.$$

Con ayuda de este teorema podemos caracterizar la compacidad relativa, lo que es lo mismo la tensión, en $\mathcal{P}(\mathcal{C})$ con la topología débil a través de este resultado

Teorema 3.3

La sucesión $\{\mathbb{P}_n : n \in \mathbb{N}\} \subset \mathcal{P}(\mathcal{C})$ es tensa si y sólo si se satisfacen las siguientes condiciones

(i) Para todo $\eta > 0$, existen $a > 0$ y $n_0 \in \mathbb{N}$ tal que
$$\mathbb{P}_n[f \in C : |f(0)| \geq a] \leq \eta, \quad n \geq n_0.$$

(ii) Para cada $\varepsilon > 0$ y $\eta > 0$, existe δ con $0 < \delta < 1$ y $n_0 \in \mathbb{N}$ tal que
$$\mathbb{P}_n\left[f \in C : \omega_f(\delta) \geq \varepsilon\right] \leq \eta, \quad n \geq n_0.$$

Demostración. Supongamos $\{\mathbb{P}_n : n \in \mathbb{N}\}$ es tensa. Entonces para todo $\eta > 0$ existe $K \subseteq C$ compacto tal que

$$\mathbb{P}_n[K] \geq 1 - \eta. \tag{3.1}$$

Como K es compacto en particular es relativamente compacto en C, entonces por el Teorema de Arzelà-Ascoli existe a suficientemente grande tal que

$$K \subset \{f \in C \,:\, |f(0)| < a\},$$

y para todo $\varepsilon, \eta > 0$ existe $\delta \in (0,1)$ tal que

$$K \subset \{f \in C \,:\, w_f(\delta) < \varepsilon\}.$$

De la monotonía de la medida de probabilidad y (3.1) se sigue

$$\mathbb{P}_n[\{f \in C \,:\, |f(0)| \geq a\}] \leq \eta,$$

y

$$\mathbb{P}_n[\{f \in C \,:\, w_f(\delta) \geq \varepsilon\}] \leq \eta,$$

para todo $n \in \mathbb{N}$, por lo que se cumple (i) y (ii).

Supongamos que se cumple (i) y (ii) dados $\eta, \varepsilon > 0$.

Observamos que

$$A_n = \{f \in C \,:\, |f(0)| \geq n\},$$

es una sucesión de eventos decreciente tal que $\bigcap_{n=1}^{\infty} A_n = \emptyset$, por lo que para $j = 1, \ldots, n_0 - 1$ existe $N_j \in \mathbb{N}$ tal que

$$\mathbb{P}_j[A_{N_j}] \leq \eta.$$

Consideremos $N := \text{máx}\{a, N_1, \ldots, N_j\}$, luego para toda $n \in \mathbb{N}$ se cumple que

$$\mathbb{P}_n[\{f \in C \,:\, |f(0)| \geq N\}] \leq \eta. \tag{3.2}$$

De manera análoga, notando que la sucesión eventos $\{\{f \in C \,:\, w_f(1/n) \geq \varepsilon\} : n \in \mathbb{N}\}$ es decreciente a $\emptyset$ se tiene que para todo $\varepsilon > 0$ y $\eta > 0$ existe $\delta \in (0,1)$ suficientemente pequeño tal que

$$\mathbb{P}_n[\{f \in C \,:\, w_f(\delta) \geq \varepsilon\}] \leq \eta.$$

Por lo tanto, para cada $k \in \mathbb{N}$ existe $\delta_k > 0$ tal que si $B_k = \{f \in C \,:\, w_f(\delta_k) < 1/k\}$ entonces

$$\mathbb{P}_n[B_k] \geq 1 - \frac{\eta}{2^{k+1}}. \tag{3.3}$$

Por (3.2) existe $N > 0$ tal que si $B = \{f \in C \ : \ |f(0)| < N\}$ entonces

$$\mathbb{P}_n[B] \geq 1 - \frac{\eta}{2}. \tag{3.4}$$

Consideremos

$$A = B \cap \left(\bigcap_{k=1}^{\infty} B_k \right),$$

y $K = \overline{A}$, entones $K^c \subseteq A^c$, luego

$$K^c \subseteq B^c \cup \left(\bigcup_{k=1}^{\infty} B_k^c \right).$$

De la monotonía y subaditividad de las medidas de probabilidad obtenemos

$$\mathbb{P}_n[K^c] \leq \mathbb{P}_n[B^c] + \sum_{k=1}^{\infty} \mathbb{P}_n[B_k^c].$$

Por (3.3) y (3.4) tenemos

$$\mathbb{P}_n[K^c] \leq \frac{\eta}{2} + \sum_{k=1}^{\infty} \frac{\eta}{2^{k+1}} = \eta.$$

Por lo tanto

$$\mathbb{P}_n[K] \geq 1 - \eta. \tag{3.5}$$

Observamos que

$$\sup_{f \in A} |f(0)| < N < \infty.$$

Por otro lado, por la propiedad arquimediana para todo $\varepsilon > 0$ existe $k \in \mathbb{N}$ con $1/k < \varepsilon$ y a su vez existe δ_k tal que

$$w_f(\delta_k) < \frac{1}{k} \quad \text{para cada } f \in A,$$

luego

$$\sup_{f \in A} w_f(\delta_k) \leq \frac{1}{k},$$

por lo que

$$\lim_{\delta \to 0} \sup_{f \in A} w_f(\delta) = 0.$$

De esto y mediante el Teorema de Arzelà-Ascoli llegamos a que A es relativamente compacto por lo que K es un conjunto compacto en C y por (3.5) concluimos $\{\mathbb{P}_n : n \in \mathbb{N}\}$ es una sucesión de medidas de probabilidad tensa, que es lo que queríamos probar. □

Observación 3.1

Observamos que para δ fijo $\omega(\cdot, \delta)$ es medible por ser una función continua. La condición (ii) en el Teorema 3.3 es equivalente a que para todo $\varepsilon > 0$

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}_n \left[\{ f \in C : \omega_f(\delta) \geq \varepsilon \} \right] = 0.$$

Ahora enunciaremos una desigualdad que nos servirá más adelante

Lema 3.1

Sea $\mathbb{P} \in \mathcal{P}(\mathcal{C})$ y $\delta > 0$, si $0 = t_0 < t_1 < \ldots < t_k = 1$ y

$$\min_{1 < i < k} (t_i - t_{i-1}) \geq \delta,$$

entonces para cualquier $\varepsilon > 0$

$$\mathbb{P}[\{f \in C : \omega_f(\delta)\} \geq 3\varepsilon] \leq \mathbb{P}[\{f \in C : \sup_{t_{i-1} \leq s \leq t_i} |f(s) - f(t_{i-1})| \geq \varepsilon\}].$$

Hemos especificado algunas propiedades del espacio C que nos servirán para tratar la convergencia débil de medidas de probabilidad en este espacio, ahora detallaremos como se ven los elementos aleatorios con un dominio un espacio de probabilidad abstracto $(\Omega, \mathcal{F}, \mathbb{P})$ y rango C.

Ejemplificando la idea ya hemos estudiado variables aleatorias como funciones medibles de la forma $X : \Omega \to \mathbb{R}$, vectores aleatorios como funciones medibles $X : \Omega \to \mathbb{R}^k$ y procesos estocásticos a tiempo discreto como una función medible $X : \Omega \to \mathbb{R}^{\mathbb{N}}$, en este último caso se cumple que la clase de conjuntos finito dimensionales, que se define de manera similar a como lo hemos hecho, es clase determinante de convergencia.

Generalizando los procesos estocásticos a tiempo discreto están los procesos estocásticos a tiempo continuo, un ejemplo de esto sería el Proceso Poisson. A un proceso estocástico a tiempo continuo lo podemos ver como una función medible de la forma $X : \Omega \to C$, este tipo de función induce una medida de probabilidad que estará en $\mathcal{P}(\mathcal{C})$ que será la distribución del proceso, con esta idea en mente podemos estudiar la convergencia débil de procesos estocásticos a tiempo continuo considerando sus la convergencia débil de sus distribuciones en $\mathcal{P}(\mathcal{C})$.

3.1. Elementos Aleatorios en C

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $(C, \mathcal{C})$ el espacio medible asociado al espacio métrico C, consideraremos un elemento aleatorio sobre C a una función $X : \Omega \to C$ que es $\mathcal{F}\backslash\mathcal{C}$-medible, esto es para cada $\omega \in \Omega$ se tiene que $X(\cdot, \omega) : [0,1] \to \mathbb{R}$ es un elemento de C, es decir, una función continua.

Simplificando la notación, para $t \in [0,1]$ fijo sea $X_t := X(t, \cdot) : \Omega \to \mathbb{R}$, que es una función real con dominio Ω, observamos que

$$X_t = \pi_t \circ X,$$

luego X_t es la proyección natural de X en t.

Podemos considerar proyecciones más generales, para $0 \leq t_1 < \ldots < t_k \leq 1$ sea

$$(X_{t_1}, \ldots, X_{t_k}) = \pi_{t_1,\ldots,t_k} \circ X.$$

Como X es $\mathcal{F}\backslash\mathcal{C}$-medible y por el Teorema 3.1 tenemos que $\pi_{t_1,\ldots,t_k}$ es $\mathcal{C}\backslash\mathcal{B}(\mathbb{R}^k)$-medible entonces $\pi_{t_1,\ldots,t_k} \circ X$ es $\mathcal{F}\backslash\mathcal{B}(\mathbb{R}^k)$-medible, este hecho nos permite definir la familia de distribuciones finito-dimensionales del proceso X.

Definición 3.1

Sea $X : \Omega \to C$ un elemento aleatorio, la distribución de X es

$$\mathbb{P}_X[B] := \mathbb{P}\left[X^{-1}(B)\right] \text{ con } B \in \mathcal{C},$$

luego $\mathbb{P}_X \in \mathcal{P}(\mathcal{C})$ y sea

$$\mathbb{P}_X \circ \pi^{-1}_{t_1,\ldots,t_k}(H) := \mathbb{P}[(X_{t_1}, \ldots, X_{t_k}) \in H] = \mathbb{P}_X\left[\pi^{-1}_{t_1,\ldots,t_k}(H)\right] \text{ con } H \in \mathcal{B}(\mathbb{R}^k).$$

A la familia de medidas de probabilidad

$$\left\{\mathbb{P}_X \circ \pi^{-1}_{t_1,\ldots,t_k} : 0 \leq t_1 < \ldots, t_k \leq 1\right\},$$

la denominamos como la familia de distribuciones finito dimensionales de X.

La definición anterior nos permite relacionar la convergencia de procesos en C con la convergencia débil de medidas de probabilidad en $\mathcal{P}(\mathcal{C})$.

Definición 3.2

Decimos que una sucesión de elementos aleatorios $\{X^n : n \in \mathbb{N}\}$ converge a un elemento aleatorio en X, que escribimos como $X^n \Rightarrow X$ si $\mathbb{P}_{X^n} \Rightarrow \mathbb{P}_X$ cuando $n \to \infty$.

En base a la caracterización de tensión de una familia de medidas de probabilidad que hemos descrito en esta sección y la definición de convergencia débil de procesos podemos relacionar la convergencia débil de procesos en C con la convergencia de los procesos definidos por las proyecciones canónicas mediante este resultado

Teorema 3.4

Sean $\{X^n : n \in \mathbb{N}\}, X$ elementos aleatorios en C con dominio un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Si

(i)
$$(X^n_{t_1}, \ldots, X^n_{t_k}) \Rightarrow (X_{t_1}, \ldots, X_{t_k}) \text{ para } 0 \leq t_1 < \ldots < t_k \leq 1.$$

(ii) Para todo $\varepsilon > 0$
$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}[\omega(X^n, \delta) \geq \varepsilon] = 0.$$

entonces $X^n \Rightarrow X$ cuando $n \to \infty$.

Demostración. Sean $\mathbb{P}_X$ y $\{\mathbb{P}_n : n \in \mathbb{N}\}$ las distribuciones de X y $\{X^n : n \in \mathbb{N}\}$, respectivamente. Notamos que la condición (i) es equivalente a

$$\mathbb{P}_n \circ \pi^{-1}_{t_1,\ldots,t_k} \Rightarrow \mathbb{P}_X \circ \pi^{-1}_{t_1,\ldots,t_k}.$$

Queremos demostrar que $\mathbb{P}_n \Rightarrow \mathbb{P}_X$. Como la clase de conjuntos finito dimensionales es separante por el Teorema 3.1, entonces por la Proposición 2.4 de Martín González, 2021 basta demostrar que $\{\mathbb{P}_n : n \in \mathbb{N}\}$ es relativamente compacto que en el espacio C es equivalente a tensión.

Sabemos que $\mathbb{P}_n \circ \pi_0^{-1} \Rightarrow \mathbb{P}_X \circ \pi_0^{-1}$, luego $\{\mathbb{P}_n \circ \pi_0^{-1} : n \in \mathbb{N}\}$ es tensa en el espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, por lo que para todo $\eta > 0$ existe $K \subset \mathbb{R}$ compacto tal que

$$\mathbb{P}_n \circ \pi_0^{-1}[K] \geq 1 - \eta.$$

Al ser $K \subset \mathbb{R}$ compacto por el Teorema de Heine-Borel existe $a > 0$ tal que $K \subset (-a, a)$

$$\mathbb{P}_n \circ \pi_0^{-1}[x \in \mathbb{R} : |x| \geq a] \leq \eta,$$

que equivale a

$$\mathbb{P}_n[f \in C : |f(0)| \geq a] \leq \eta,$$

lo que verifica la condición (i) del Teorema 3.3 para la familia de medidas de probabilidad $\{\mathbb{P}_n : n \in \mathbb{N}\}$.

Observamos que para todo ε.

$$\mathbb{P}_n\big[f \in C : \omega_f(\delta) \geq \varepsilon\big] = \mathbb{P}[\omega(X^n, \delta) \geq \varepsilon],$$

de esto y la Observación 3.1 se tiene que la condición (ii) de este Teorema es equivalente a la condición (ii) del Teorema 3.3.

Por el Teorema 3.3 concluimos $\{\mathbb{P}_n : n \in \mathbb{N}\}$ es una familia de medidas de probabilidad tensa, que es lo que queríamos probar. $\square$

4. Medida de Wiener

En C podemos considerar un proceso basado en las proyecciones naturales, recordemos que para $t \in [0,1]$ se tiene que $\pi_t : C \to \mathbb{R}$ es una función $\mathcal{C}\backslash\mathcal{B}(\mathbb{R})$-medible, luego si definimos $P_t : C \to \mathbb{R}$ como

$$P_t(f) = \pi_t(f) = f(t),$$

tenemos que P_t es $\mathcal{C}\backslash\mathcal{B}(\mathbb{R})$-medible, en consecuencia es una variable aleatoria. Podemos considerar el proceso $P := \{P_t : t \in [0,1]\}$, además $P_t(f)$ es continua en t pues f es una función continua, por lo que P es formalmente un elemento aleatorio en C.

Para demostrar la versión funcional del TLC o Teorema de Donsker describiremos el proceso límite, la medida de probabilidad correspondiente inducida por la distribución del proceso límite es conocida como la Medida de Wiener, que definiremos a continuación

Definición 4.1

La medida de Wiener, que denotaremos por $\mathbb{W}$, es una medida de probabilidad en $(C, \mathcal{C})$ que tiene las siguientes propiedades.

(i)
$$\mathbb{W}[P_0 = 0] = 1$$

(ii) P_t tiene distribución normal con media 0 y varianza t bajo $\mathbb{W}$, es decir,

$$\mathbb{W}[P_t \leq \alpha] = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\alpha} e^{-\frac{u^2}{2t}} du.$$

(iii) Si $0 \leq t_0 \leq t_1 < \ldots < t_k = 1$ entonces la colección de variables aleatorias $\{P_{t_i} - P_{t_{i-1}} : i = 1, \ldots, n\}$ son independientes respecto a la medida $\mathbb{W}$.

Para que la definición anterior sea apropiada, debemos demostrar que en efecto dicha medida en $(C, \mathcal{C})$ existe, para probar la existencia de la medida de Wiener utilizaremos el siguiente resultado

Teorema 4.1 (*Continuidad de Kolmogorov*)

Sea $\xi = \{\xi_t : t \in [0,1]\}$ un proceso estocástico y $\mathbb{P}_{t_0,\ldots,t_k}$ la distribución del vector $(\xi_{t_0}, \ldots, \xi_{t_k})$ en $\mathbb{R}^k$. Si existen constantes $\alpha, \delta, K > 0$ tal que

$$\mathbb{E}[|\xi_t - \xi_s|^\alpha] \leq K|t-s|^{1+\delta},$$

para todo $t, s \in [0,1]$, entonces existe una única medida μ en $(C, \mathcal{C})$ tal que

$$\mathbb{P}_{t_0,\ldots,t_k} = \mu \circ \pi^{-1}_{t_1,\ldots,t_k},$$

para todo $k \in \mathbb{N}$ y $t_0, \ldots, t_k \in [0,1]$.

A continuación demostraremos la existencia de la medida de Wiener.

Teorema 4.2

Existe una medida $\mathbb{W}$ en el espacio medible $(C, \mathcal{C})$ que cumple con las condiciones de la Definición 4.1.

Demostración. Por el Teorema de Consistencia de Kolmogorov existe un proceso estocástico $\xi := \{\xi_t : t \in [0,1]\}$ sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ tal que se cumple lo siguiente

(i) La variable $\xi_t - \xi_s$ tiene distribución normal con media 0 y varianza $t - s$ con $0 \leq s < t \leq 1$.
(ii) Si $0 \leq t_0 < \ldots < t_k \leq 1$ las variables aleatorias $\{\xi_{t_i} - \xi_{t_{i-1}} : i = 1, \ldots, k\}$ son independientes.

Notamos que bajo estas condiciones la distribución del vector $(\xi_{t_0}, \xi_{t_1} - \xi_{t_0}, \ldots, \xi_{t_k} - \xi_{t_{k-1}})$ considerando la función de distribución normal centrada con varianza σ^2 como Φ_{σ^2} es

$$\mathbb{P}_{t_0, t_1 - t_0, \ldots, t_k - t_{k-1}}\left[\xi_{t_0} \leq \alpha_0, \xi_{t_1} - \xi_{t_0} \leq \alpha_1, \ldots, \xi_{t_k} - \xi_{t_{k-1}} \leq \alpha_k\right] = \Phi_{t_0}(\alpha_0) \cdot \prod_{i=1}^{k} \Phi_{t_i - t_{i-1}}(\alpha_i),$$

por lo que la distribución del vector $(\xi_{t_0}, \xi_{t_1} - \xi_{t_0}, \ldots, \xi_{t_k} - \xi_{t_{k-1}})$ es una normal multivariada con media 0 y matriz de covarianzas $\operatorname{diag}(t_0, t_1 - t_0, \ldots, t_k - t_{k-1})$.

Observamos que esto determina la distribución del vector $(\xi_{t_0}, \xi_{t_1}, \ldots, \xi_{t_k})$, que es normal multivariada pues este vector aleatorio es una transformación lineal del vector $(\xi_{t_0}, \xi_{t_1} - \xi_{t_0}, \ldots, \xi_{t_k} - \xi_{t_{k-1}})$, nos quedamos con la distribución de este último vector aleatorio por simplicidad.

Ahora por el Teorema de Isserlis tenemos que

$$\mathbb{E}\left[|\xi_t - \xi_s|^4\right] = 3(\mathbb{E}[\xi_t - \xi_s]^2)^2 = 3|t-s|^2,$$

por lo que se satisfacen las condiciones del Teorema 4.1, luego existe una única medida $\mathbb{W}$ en $(C, \mathcal{C})$, tal que

$$\mathbb{P}_{t_0, t_1 - t_0, \ldots, t_k - t_{k-1}}\left[\xi_{t_0} \leq \alpha_0, \xi_{t_1} - \xi_{t_0} \leq \alpha_1, \ldots, \xi_{t_k} - \xi_{t_{k-1}} \leq \alpha_k\right] = W \circ \pi^{-1}_{t_0, t_1 - t_0, \ldots, t_k - t_{k-1}}[(-\infty, \boldsymbol{\alpha}]],$$

donde $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_k)$ y

$$(-\infty, \boldsymbol{\alpha}] = \left\{ x := (x_0, \ldots, x_k) \in \mathbb{R}^{k+1} : x_i \leq \alpha_i \text{ para } i = 0, 1, \ldots, k \right\}.$$

Por lo tanto $\mathbb{W}$ satisface las condiciones (ii) y (iii) de la Definición 4.1.

Para verificar la condición (i) basta notar que $\xi_{1/n^2} \xrightarrow{P} 0$ respecto a la medida $\mathbb{W}$ cuando $n \to \infty$, luego para todo $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{W}\left[|P_{1(n^2)}| > \varepsilon\right] = \lim_{n\to\infty} \mathbb{P}\left[|\xi_{1/n^2}| > \varepsilon\right] = 0,$$

y por la continuidad del proceso P se tiene que

$$\mathbb{W}[|P_0| > \varepsilon] = 0,$$

por lo que $\mathbb{W}[P_0 = 0] = 1$, lo que concluye la prueba. □

Dado que en el Teorema de Donsker nos interesa el límite de una convergencia débil de procesos es útil obtener a partir de la medida de Wiener una proceso con distribución dicha medida.

Observación 4.1

El proceso $P = \{P_t : t \in [0,1]\}$ es un elemento aleatorio en C cuya distribución es la medida W, al proceso P se le conoce como Movimiento Browniano.

A partir de este punto ya tenemos las herramientas necesarias para por fin enunciar el Teorema de Donsker y darle una demostración, tema central de la siguiente sección.

5. Teorema de Donsker

Empezamos considerando $\{\xi_n : n \in \mathbb{N}\}$ un conjunto de variables aleatorias sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ independientes e idénticamente distribuidas con media 0 y varianza σ^2.

Definimos $S_0 = 0$ y para $n \in \mathbb{N}$

$$S_n = \xi_1 + \cdots + \xi_n.$$

Para $n \in \mathbb{N}$ y $i = 1, \ldots, n$ consideramos los puntos $i/n \in [0, 1]$ y definimos

$$X_n(i/n, \omega) = \frac{1}{\sigma\sqrt{n}} S_i(\omega) \quad \text{para } \omega \in \Omega.$$

Se define tambien $X^n(t, \omega)$ para $t \in [0, 1]$ a través de una interpolación lineal, esto es para $t \in [(i-1)/n, i/n]$

$$\begin{aligned} X_n(t, \omega) =& \frac{(i/n) - t}{1/n} X_n((i-1)/n, \omega) + \frac{t - (i-1)/n}{1/n} X_n(i/n, \omega) \\ =& \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega) + n\left(t - \frac{i-1}{n}\right) \frac{1}{\sigma\sqrt{n}} \xi_i(\omega). \end{aligned}$$

Observamos que si $t \in [(i-1)/n, i/n)$ entonces $\lfloor nt \rfloor = i - 1$, entonces podemos definir $X^n(t, \omega)$ de forma más concisa

$$X_n(t, \omega) = \frac{1}{\sigma\sqrt{n}} S_{\lfloor nt \rfloor}(\omega) + (nt - \lfloor nt \rfloor) \frac{1}{\sigma\sqrt{n}} \xi_{\lfloor nt \rfloor + 1}(\omega). \tag{5.1}$$

A partir de la definición anterior se tiene que para t fijo $X_n(t, \cdot) : \Omega \to \mathbb{R}$ es una variable pues $S_{\lfloor nt \rfloor}$ es una variable aleatoria, además para $\omega \in \Omega$ fijo, por ser $X_n(\cdot, \omega) : [0, 1] \to \mathbb{R}$ una interpolación lineal entonces esta función es continua, por lo que $X_n : \Omega \to C$ es un elemento aleatorio en C definido de la siguiente forma

$$X_n(\omega) = X_n(\cdot, \omega),$$

y abusando un poco de la notación usaremos X_n como función o elemento aleatorio según el contexto, pero la explicación anterior debería liberar de cualquier ambigüedad.

Notamos que el proceso X_n definido en (5.1) es similar a la suma de variables aleatorias centrada y escalada en el caso continuo.

Antes de enunciar y demostrar el Teorema de Donsker enunciaremos los siguientes lemas:

Lema 5.1

Para todo $\varepsilon > 0$ y $\delta \in (0,1)$ si $\lambda = \frac{\varepsilon}{\sqrt{2\delta}}$ existe $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$ existe $m := m(n) \in \mathbb{N}$ de modo que

$$\mathbb{P}[\omega(X_n, \delta) \geq 3\varepsilon] \leq \frac{4\lambda^2}{\varepsilon^2} \cdot \mathbb{P}\left[\max_{j \leq m} |S_j| \geq \lambda\sigma\sqrt{m}\right].$$

Demostración. Consideremos $0 = t_0 < t_1 < \ldots < t_k = 1$ de la forma

$$t_i = \frac{m_i}{n} \quad \text{para } i = 1, 2, \ldots, k.$$

con $m_i \in \mathbb{N}$.

Si $\min_{1<i<k}(t_i - t_{i-1}) \geq \delta$ entonces por el Lema 3.1 se tiene que

$$\mathbb{P}[\omega(X_n, \delta) \geq 3\varepsilon] \leq \sum_{i=1}^{k} \mathbb{P}\left[\sup_{t_{i-1} \leq s \leq t_i} |X_n(s) - X_n(t_{i-1})| \geq \varepsilon\right] \tag{5.2}$$

Sea $m = \lceil n\delta \rceil$, si tomamos

$$m_i = i \cdot m \quad \text{para } i = 0, 1, \ldots, k-1,$$

y

$$m_k = n,$$

entonces para $1 \leq i < k$

$$m_i - m_{i-1} = m \geq n\delta,$$

y si a partir de esta definición de m_i consideramos t_i como antes tenemos

$$t_i - t_{i-1} \geq \delta \quad \text{para } i = 1, \ldots, k-1.$$

Con esta elección de la partición $\{t_i : i = 0, 1, \ldots, k\}$ se sigue cumpliendo la desigualdad en (5.2).

Por la definición de X_n y la elección de la partición $\{t_i : i = 0, 1, \ldots, k\}$ de $[0,1]$ se tiene que

$$\sup_{t_{i-1} \leq s \leq t_i} |X_n(s) - X_n(t_{i-1})| = \max_{m_{i-1} \leq j \leq m_i} \frac{1}{\sigma\sqrt{n}} \cdot |S_j - S_{m_{i-1}}| \quad \text{para } i = 1, 2, \ldots, k.$$

Al ser $\{\xi_n : n \in \mathbb{N}\}$ variables aleatorias independientes e idénticamente distribuidas se sigue que $S_j - S_{m_{i-1}}$ tiene la misma distribución que $S_{j-m_{i-1}}$ y haciendo un cambio de variable llegamos a que

$$\mathbb{P}\left[\max_{m_{i-1} \leq j \leq m_i} \frac{1}{\sigma\sqrt{n}} |S_j - S_{m_{i-1}}| \geq \varepsilon\right] = \mathbb{P}\left[\max_{j \leq m_i - m_{i-1}} |S_j| \geq \varepsilon\sigma\sqrt{n}\right] \quad \text{para } i = 1, 2, \ldots, k.$$

Por otro lado, elegimo $k = \lceil n/m \rceil$, como $m = n\delta$ y $\delta \in (0,1)$ entonces $k \geq 1$, pues la función techo es monótona creciente.

Con esta elección de k se tiene que

$$k - 1 \leq \frac{n}{m} < k,$$

de lo que se sigue

$$m_{k-1} \leq m_k = n < km,$$

por lo que en este caso

$$m_k - m_{k-1} < m,$$

y se tiene así que

$$m_i - m_{i-1} \leq m \quad \text{para } i = 1, 2, \ldots, k,$$

por lo cual tenemos la siguiente desigualdad

$$\mathbb{P}\left[\max_{m_{i-1} \leq j \leq m_i} \frac{1}{\sigma\sqrt{n}} |S_j - S_{m_{i-1}}| \geq \varepsilon\right] \leq \mathbb{P}\left[\max_{j \leq m} |S_j| \geq \varepsilon\sigma\sqrt{n}\right] \quad \text{para } i = 1, 2, \ldots, k.$$

De esta última desigualdad y (5.2) obtenemos que

$$\mathbb{P}[\omega(X_n, \delta) \geq 3\varepsilon] \leq k \cdot \mathbb{P}\left[\max_{j \leq m} |S_j| \geq \varepsilon\sigma\sqrt{n}\right]. \tag{5.3}$$

Ahora como $m = \lceil n\delta \rceil$ entonces

$$\lim_{n \to \infty} \lceil n/m \rceil = \frac{1}{\delta} > \frac{2}{\delta},$$

y

$$\lim_{n \to \infty} n/m = \frac{1}{\delta} > \frac{1}{2\delta}.$$

De los límites anteriores podemos concluir que existe $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$ se tiene que

$$k < \frac{2}{\delta} \quad \text{y} \quad n > \frac{m}{2\delta}.$$

Por lo tanto, por monotonía de la medida de probabilidad y la desigualdad en (5.3) llegamos a que

$$\mathbb{P}[\omega(X_n, \delta) \geq 3\varepsilon] \leq \frac{2}{\delta} \cdot \mathbb{P}\left[\max_{j \leq m} |S_j| \geq \varepsilon\sigma\sqrt{\frac{m}{2\delta}}\right] \quad \text{para } n \geq n_0.$$

Con el cambio de variable $\lambda = \varepsilon/\sqrt{2\delta}$ realizando los cálculos correspondientes en la desigualdad anterior obtenemos

$$\mathbb{P}[\omega(X_n, \delta) \geq 3\varepsilon] \leq \frac{4\lambda^2}{\varepsilon^2} \cdot \mathbb{P}\left[\max_{j\leq m}|S_j| \geq \lambda\sigma\sqrt{m}\right] \quad \text{para } n \geq n_0,$$

y esta es la desigualdad que queríamos probar. □

Ahora enunciamos la siguiente desigualdad que será útil para nuestros propósitos.

Lema 5.2

Con las mismas consideraciones para $\{\xi_n : n \in \mathbb{N}\}$ y las sumas parciales como al principio de esta sección se cumple que

$$\mathbb{P}\left[\max_{j\leq m}|S_j| \geq \lambda\sigma\sqrt{m}\right] \leq 2 \cdot \mathbb{P}[|S_m| \geq (\lambda - \sqrt{2})\sigma\sqrt{m}].$$

Con estos resultados estamos listos para por fin enunciar formalmente y demostrar el Teorema de Donsker.

Teorema 5.1 (*Teorema de Donsker*)

Sea W el elemento aleatorio en C con distribución la medida de Wiener $\mathbb{W}$ y $\{X_n : n \in \mathbb{N}\}$ la familia de elementos aleatorios en C definidos en (5.1). Entonces se cumple que

$$X_n \Rightarrow W,$$

cuando $n \to \infty$.

Demostración. En primer lugar, veamos que las distribuciones finito dimensionales de $\{X_n : n \in \mathbb{N}\}$ convergen a las distribuciones finito dimensionales de W.

Primero veamos que para $t \in [0,1]$ se cumple que

$$X_n(t) \xrightarrow{d} W_t.$$

Por (5.1) se tiene

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}}S_{\lfloor nt\rfloor}| \leq \frac{1}{\sigma\sqrt{n}}|\xi_{\lfloor nt\rfloor}|.$$

Mediante un argumento usando la desigualdad de Chebyshev llegamos a que

$$\frac{1}{\sigma\sqrt{n}}\xi_{\lfloor nt\rfloor} \xrightarrow{P} 0.$$

luego
$$\left| X_n(t) - \frac{1}{\sigma\sqrt{n}} S_{\lfloor nt \rfloor} \right| \xrightarrow{P} 0. \tag{5.4}$$

Por otro lado, se ve que $\lfloor nt \rfloor / n \to t$ cuando $n \to \infty$ y
$$\frac{1}{\sigma\sqrt{n}} S_{\lfloor nt \rfloor} = \sqrt{\frac{\lfloor nt \rfloor}{n}} \frac{1}{\sigma\sqrt{\lfloor nt \rfloor}} S_{\lfloor nt \rfloor}.$$

Del Teorema 4.2 se tiene que W_t es normal con media 0 y varianza t, con esta observación y por el TLC se tiene que
$$\frac{1}{\sigma\sqrt{n}} S_{\lfloor nt \rfloor} \xrightarrow{d} W_t,$$
de esto y (5.4) se sigue a través del Teorema de Slutsky que
$$X_n(t) \xrightarrow{d} W_t.$$

Ahora para probar la condición (i) del Teorema 3.4 basta con probar para $t, s \in [0, 1]$ con $t < s$ que
$$(X_n(t), X_n(s)) \xrightarrow{d} (W_t, W_s). \tag{5.5}$$

Observamos que si probamos
$$(X_n(t), X_n(s) - X_n(t)) \xrightarrow{d} (W_t, W_s - W_t),$$
tenemos la convergencia en (5.5) como consecuencia del Teorema de Mapeo Continuo ya que $(X_n(t), X_n(s))$ es una transformación lineal del vector aleatorio $(X_n(t), X_n(s) - X_n(t))$.

Por (5.4) y el Teorema de Slutsky es suficiente con demostrar
$$\left(\frac{1}{\sigma\sqrt{n}} S_{\lfloor nt \rfloor}, \frac{1}{\sigma\sqrt{n}} (S_{\lfloor ns \rfloor} - S_{\lfloor nt \rfloor}) \right) \xrightarrow{d} (W_t, W_s - W_t).$$

Ya sabemos que se tiene la convergencia en distribución por coordenadas y como las variables aleatorias en cada entrada son independientes también se tiene la convergencia en distribución de la conjunta, de lo que se concluye la convergencia de las distribuciones finito dimensionales de $\{X_n : n \in \mathbb{N}\}$ a las finito dimensionales de W, lo que verifica la condición (i) del Teorema 3.4.

Ahora verifiquemos la condición (ii) del Teorema 3.4.

Sabemos por el Lema 5.1 que para todo $\varepsilon > 0$ y $\delta \in (0,1)$ existe $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$ y con $m = \lceil n\delta \rceil$ y $\lambda = \frac{\varepsilon}{\sqrt{2\delta}}$ se cumple que

$$\mathbb{P}[\omega(X_n, \delta) > 3\varepsilon] \leq \frac{4\lambda^2}{\varepsilon^2} \cdot \mathbb{P}\left[\max_{j \leq m} |S_j| \geq \lambda\sigma\sqrt{m}\right],$$

y por el Lema 5.2 tenemos que

$$\mathbb{P}[\omega(X_n, \delta) > 3\varepsilon] \leq \frac{8\lambda^2}{\varepsilon^2} \cdot \mathbb{P}\left[\frac{1}{\sigma\sqrt{m}}|S_m| \geq \lambda - \sqrt{2}\right] \quad \text{para } n \geq n_0.$$

Si $n \to \infty$ y $m = \lceil n\delta \rceil$ para δ fijo $m \to \infty$, si N es una variable aleatoria normal estándar con dominio el espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, entonces por el TLC se tiene que

$$\limsup_{n \to \infty} \mathbb{P}[\omega(X_n, \delta) > 3\varepsilon] \leq \frac{8\lambda^2}{\varepsilon^2} \mathbb{P}[|N| \geq \lambda - \sqrt{2}].$$

Como tomaremos $\delta \to 0^+$ entonces existe δ tal que $\lambda > 2\sqrt{2}$ para ese δ y todos los menores a el, por lo que

$$\frac{\lambda}{2} > \sqrt{2},$$

luego

$$\lambda - \sqrt{2} > \frac{\lambda}{2},$$

y de la monotónia de la medida de probabilidad se sigue que

$$\limsup_{n \to \infty} \mathbb{P}[\omega(X_n, \delta) > 3\varepsilon] \leq \frac{8\lambda^2}{\varepsilon^2} \mathbb{P}\left[|N| \geq \frac{\lambda}{2}\right],$$

por la desigualdad de Chebyshev tenemos que

$$\begin{aligned} \limsup_{n \to \infty} \mathbb{P}[\omega(X_n, \delta) > 3\varepsilon] \leq & \frac{8\lambda^2}{\varepsilon^2} \cdot \frac{8}{\lambda^3} \cdot \mathbb{E}[|N|^3] \\ = & \left(\frac{8}{\varepsilon}\right)^2 \cdot \frac{\sqrt{2\delta}}{\varepsilon} \cdot \mathbb{E}[|N|^3], \end{aligned}$$

y de esta desigualdad tomando $\delta \to 0$ llegamos a que

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}[\omega(X_n, \delta) > 3\varepsilon] = 0,$$

por lo que se cumple la condición (ii) del Teorema 3.4.

Al cumplirse (i) y (ii) del Teorema 3.4 con $\{X_n : n \in \mathbb{N}\}$ y el proceso W con distribución la medida de Wiener concluimos $X_n \Rightarrow W$, que es lo que queríamos probar. $\square$

Referencias

Bhattacharya, R. & Waymire, E. C. (2016). *A Basic Course in Probability Theory* (2.ª ed.). Cham, Springer International Publishing : Imprint: Springer.

Billingsley, P. (1999). *Convergence of probability measures* (2.ª ed.). New York, Wiley.

Martín González, E. M. (2021). *Probabilidad Avanzada* (Notas de clase). México, Guanajuato.

Parthasarathy, K. R. (2005). *Probability measures on metric spaces*. Providence, R.I, AMS Chelsea Pub.