15
24

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

T E S I S

CONVERGENCIA DEBIL Y TEOREMA CENTRAL DE LIMITE EN C(0,1)

QUE PARA OBTENER EL TITULO DE

M A T E M A T I C O

PRESENTA :

JORGE GARCIA VILLEDA

MEXICO D.F. ABRIL DE 1991

TESIS CON FALLA DE ORIGEN

INTRODUCCION

El objetvo de los primeros 6 capítulos es presentar los conocimientos básicos en la teoría de la convergencia debil, los cuales se utilizarán en el desarrollo del trabajo.

En el capítulo 1, se dan las definiciones básicas así como algunos ejemplos interesantes de tensión (ejemplos 1 y 3) y se encontró un ejemplo donde no hay tensión (ejemplo 2) .

El capítulo 2 contiene el importante teorema de *Pormanteau*. En el libro de Billingsley se prueba el teorema de Pormanteau con 5 equivalencias, en el libro de Kurtz, sin embargo, se agrega una sexta equivalencia al introducir la métrica de *Prohorov* (bajo la hipótesis de separabilidad del espacio) , en este capítulo se retoma esta sexta equivalencia pero hay midificaciones y simplificaciones en la demostración.

Cabe notar la importancia de tal teorema, pues gracias a el, se obtiene resultados muy importantes en convergencia debil.

En los capítulos 3,4 y 5 se dan propiedades diversas sobre convergencia debil, el estudio en *espacios producto*, el teorema de *mapeo continuo* , etc,

En el capítulo 10 se estudia el *teorema central de límite* para variables aleatorias en $\mathbb{R}$ y en $\mathbb{R}^n$.

En el capítulo 6 se estudia el teorema de *Prohorov* que relaciona la tensión de una familia con el concepto de compacidad relativa de dicha familia y también aquí se detalla la demostración que aparece en Billingsley.

Dentro del capítulo 7 se estudia la convergencia debil de medidas sobre el espacio $(C, \mathcal{C})$, (donde C es el espacio métrico de las funciones continuas en $[\ 0,1\ ]$, con la métrica del supremo, y $\mathcal{C}$ su σ-álgebra boreliana.) y se dan condiciones para la tensión ahí. Esto es de gran importancia porque se traduce a procesos continuos, que de hecho es lo que se hace en el capítulo 8.

En el capítulo 11 se estudia debido a su importancia propia, el teorema de *Donsker*, dado que es una generalización del teorema central de límite en $\mathbb{R}$.

El capítulo 9 da nuevos teoremas para la tensión de procesos continuos, y se llega a un resultado importante (teorema 11.3) que se utiliza con mucha frecuencia.

En el capítulo 12 se aborda el tema principal de este trabajo que consiste en plantear el *teorema central de límite* en $C\ [\ 0,1\ ]$.

Finalemente en el capítulo 13 se estudia los trabajos de *Hahn y Delporte.* de 1975-1978, en donde se dan condiciones sobre un proceso para que satisfaga el teorema central de límite en $C[\ 0,1\ ]$.

En Hahn (1975) utiliza todo lo relacionado a tensión, que aparecen el el capítulo 9 de esta tesis.

En resumen : este trabajo consiste en presentar con detalle y buscando simplificar la expocisión, los temas básicos de convergencia debil para medidas definidas en un espacio métrico arbitrario (capítulos 1-6) . Después se especializa en el estudio de la convergencia debil en el espacio $C[\,0,1\,]$. Lo cual permite trabajar con procesos continuos y entonces se aplica este hecho al estudio del teorema central de límite para procesos continuos.

I N D I C E

CAPITULO PAGINA

1. CONVERGENCIA DEBIL

Se tiene un espacio métrico S y $\mathcal{S}$ su σ-álgebra de Borel (la σ-álgebra generada por los abiertos de la topología inducida por la métrica) sobre S. Llamemos $C(S)$ al conjunto de funciones continuas y acotadas sobre S.

Definición 1.1

Decimos que P_n converge debilmente a P si sucede que :

$$\int_S f dP_n \to \int_S f dP \qquad \forall f \in C(S).$$

Definición 1.2

Se dice que P (medida de probabilidad definida en $\mathcal{S}$) es regular si :

$\forall A \in \mathcal{S}, \forall \epsilon > 0, \ \exists\, G$ abierto y $\exists\, F$ cerrado tales que :

$$(*) \qquad F \subseteq A \subseteq G \qquad \text{y} \qquad P(G - F) < \epsilon.$$

Teorema 1.1

Cualquier medida de probabilidad sobre S es regular.

Demostración

Sea d la métrica en S y $A \in \mathcal{S}$ y sea $\epsilon > 0$. Supongamos que A es cerrado. Sea $G_{1/n} = \{x \in S \mid d(x, A) < 1/n\}$. Es claro que $G_{1/n} \downarrow A$ si $n \to \infty$ por tanto

$$\lim_{n\to\infty} P(G_{1/n}) = P(A)$$

y por tanto $P(G_{1/n} - A) < \epsilon$ si $n \geq N_0$, por ser P medida de probabilidad en $\mathcal{S}$. Así, la propiedad (*) se cumple para todos los cerrados. Sea

$$\mathcal{G} = \{B \in \mathcal{S} \mid (*) \text{ se satisface para } B\}$$

Observemos $\mathcal{G}$ contiene a los cerrados.

Dado que $\mathcal{G} \subseteq \mathcal{S}$, basta probar que $\mathcal{G}$ es σ-álgebra.

1.- $\emptyset \in \mathcal{G}$.

2.- $A \in \mathcal{G} \Rightarrow A^c \in \mathcal{G}$ (basta tomar complementos en $F \subseteq A \subseteq G$).

3.- Sea $A_n \in \mathcal{G}$. Para cada $n \in \mathbf{N}$ $\exists\, F_n$ cerrado y $\exists\, G_n$ abierto con

$$F_n \subseteq A_n \subseteq G_n \quad \text{y} \quad P(G_n - F_n) < \epsilon/2^{n+1}.$$

Sea $G = \bigcup_{n=1}^{\infty} G_n, G$ es abierto y $\bigcup_{n=1}^{\infty} A_n \subseteq G$.

Dado que

$$\bigcup_{i=1}^{n} F_i \uparrow \bigcup_{i=1}^{\infty} F_i \quad \text{si} \quad n \to \infty$$

entonces

$$\lim_{n\to\infty} P(\bigcup_{i=1}^{n} F_i) = P(\bigcup_{i=1}^{\infty} F_i)$$

por tanto $\exists\, N_0 \in \mathbf{N}$ tal que

$$P(\bigcup_{i=1}^{\infty} F_i - \bigcup_{i=1}^{n} F_i) < \epsilon/2 \quad \text{si } n \geq N_0.$$

Sea $F = \bigcup_{n=1}^{N_0} F_n$ entonces F es cerrado y $F \subseteq \bigcup_{n=1}^{\infty} A_n \subseteq G$. Ahora

$$\begin{aligned}
P(G - F) &= P\left(\bigcup_{n=1}^{\infty} G_n - \bigcup_{n=1}^{N_0} F_n\right) \\
&= P\left(\left(\bigcup_{n=1}^{\infty} G_n - \bigcup_{n=1}^{\infty} F_n\right) \bigcup \left(\bigcup_{n=1}^{\infty} F_n - \bigcup_{n=1}^{N_0} F_n\right)\right) \\
&\leq P\left(\bigcup_{n=1}^{\infty} G_n - \bigcup_{n=1}^{\infty} F_n\right) + P\left(\bigcup_{n=1}^{\infty} F_n - \bigcup_{n=1}^{N_0} F_n\right) \\
&< P\left(\bigcup_{n=1}^{\infty} (G_n - F_n)\right) + \epsilon/2 \\
&\leq \sum_{n=1}^{\infty} P(G_n - F_n) + \epsilon/2 \\
&< \epsilon/2 + \epsilon/2 \\
&= \epsilon.
\end{aligned}$$

Por tanto $\bigcup_{n=1}^{\infty} A_n \in \mathcal{G}$ y entonces $\mathcal{G}$ es σ-álgebra.

Observación 1.4

$$\text{Si } A \in S, \text{ entonces } \quad P(A) = \sup_{H \subseteq A, H \text{ cerrado}} P(H).$$

Es interesante ver cuando 2 medidas de probabilidad coinciden, supongamos que tenemos P, Q dos medidas de probabilidad sobre S y si :

$$\int_S f dP = \int_S f dQ \quad \forall f \in C(S)$$

entonces $P \equiv Q$.

Para ver que esto es cierto, empleemos un lema que nos servirá de ayuda.

Lema (de Uryson) 1.5

Si F es un cerrado y $\epsilon > 0$ entonces $\exists\ f \in C(S)$ tal que $f(x) = 1 \quad \forall x \in F$ y $f(x) = 0$ si $d(x, F) \geq \epsilon$ y $f(x) \in [0,1] \quad \forall x \in S$ y f es uniformemente continua en S.

Demostración

Sea $\epsilon > 0$ y sea $\phi: \mathbf{R} \to [0,1]$ definida como

$$\phi(t) = \begin{cases} 1 & \text{si } t \leq 0; \\ 1 - t & \text{si } 0 \leq t \leq 1; \\ 0 & \text{si } 1 \leq t. \end{cases}$$

y $f: S \to [0,1]$ definida como

$$f(x) = \phi(\frac{1}{\epsilon} d(x, F)).$$

a.- Si $x \in F \Rightarrow d(x, F) = 0$ y por tanto $f(x) = \phi(0) = 1$

b.- Si $x \notin F$ y $d(x, F) < \epsilon$ entonces $0 < d(x, F) < \epsilon$ por tanto

$$f(x) = \phi(\frac{1}{\epsilon} d(x, F)) \in [0,1]$$

c.- Si $d(x, F) \geq \epsilon$ se tiene $\frac{1}{\epsilon} d(x, F) \geq 1$ por tanto $f(x) = \phi(\frac{1}{\epsilon} d(x, F)) = 0$.

d.- Veamos ahora que f es uniformemente continua. Sea $\lambda > 0$, consideremos $\delta = \min\{\frac{\epsilon}{N}, \frac{\lambda\epsilon}{N}\}$ donde $\frac{1}{N} < \frac{1}{\lambda}$ y $\frac{1}{N} < \lambda$.

Mostraremos que si $d(x, y) < \delta$ entonces $| f(x) - f(y) | < \lambda$.

(1) Si $x \in F, y \in F$ y $d(x,F) < \delta$ entonces

$$| f(x) - f(y) | = | 1 - 1 | = 0 < \lambda.$$

(2) Si $d(x,F) \geq \epsilon$ y $d(y,F) \geq \epsilon$ y $d(x,y) < \delta$ entonces

$$| f(x) - f(y) | = | 0 - 0 | = 0 < \lambda.$$

(3) Si $d(x,F) \geq \epsilon$ y $0 < d(y,F) < \epsilon$ y $d(x,y) < \delta$ entonces

$$\epsilon \leq d(x,F) \leq d(x,y) + d(y,F) \leq \frac{\lambda\epsilon}{N} + d(y,F)$$

por tanto $\epsilon - \frac{\lambda\epsilon}{N} \leq d(y,F)$ luego $1 - \frac{\lambda}{N} \leq d(y,F)/\epsilon$ por tanto

$$0 < 1 - \frac{d(y,F)}{\epsilon} \leq \frac{\lambda}{N} < \lambda.$$

y

$$| f(x) - f(y) | = | 0 - (1 - \frac{d(y,F)}{\epsilon}) | = | 1 - \frac{d(y,F)}{\epsilon} | < \lambda.$$

(4) Si $x \in F$ y $0 < d(y,F) < \epsilon$ y $d(x,y) < \delta$ entonces

$$d(y,F) \leq d(y,x) + d(x,F) \leq \frac{\epsilon}{N} + 0$$

por tanto $\frac{d(y,F)}{\epsilon} \leq \frac{1}{N} < \lambda$ y entonces

$$| f(x) - f(y) | = | 1 - (1 - \frac{d(y,F)}{\epsilon}) | = | \frac{d(y,F)}{\epsilon} | \leq \frac{1}{N} < \lambda.$$

y entonces f es uniformemente continua en S.

Teorema 1.6

Sean P, Q dos probabilidades tales que :

$$\int_S f dP = \int_S f dQ \quad \forall f \in C(S)$$

entonces $P \equiv Q$.

Demostración

Sea F cerrado y sea

$$\phi_n(t) = \phi(nt)$$

con ϕ como en el lema anterior. y

$$f_n(x) = \phi(nd(x,F))$$

$f_n \in C(S)$ $\forall n \in \mathbb{N}$ (pues f_n es continua $\forall n$ y $f_n(x) \in [0,1]$ $\forall x \in S$) y $\{f_n\}$ es decreciente.

1.- Si $x \in F$ entonces $f_n(x) = \phi(0) = 1$ por tanto $f_n(x) \to 1$ si $n \to \infty$

2.- Si $x \notin F$ como F es cerrado $d(x,F) \geq \epsilon$ para algún $\epsilon > 0$ por tanto

$$nd(x,F) \geq n\epsilon \geq 1 \qquad \text{si } n \geq n_0.$$

Por tanto $f_n(x) = \phi(nd(x,F)) = 0$ y $f_n(x) \to I_F(x)$ $\forall x \in S$.

Por el teorema de convergencia monótona :

$$\int_S f_n dP \to \int_S I_F dP$$

y

$$\int_S f_n dQ \to \int_S I_F dQ \qquad \text{si } n \to \infty$$

así $P(F) = Q(F)$, en virtud de que

$$P(A) = \sup_{H \subseteq A, H \text{cerrado}} P(H)$$

se tiene que $P \equiv Q$.

Observación 1.7

Por el resultado anterior, si se conocen $\int_S f dP$ $\forall f \in C(S)$ se conoce a la medida, es decir, los valores de $\int_S f dP$ determinan de manera única a P.

Definición 1.8

Un soporte de una medida de probabilidad P sobre S, es un conjunto $A \in \mathcal{S}$ tal que $P(A) = 1$.

Teorema 1.12

Si S es un espacio métrico separable y completo, entonces cada medida de probabilidad sobre S es tensa.

Demostración

Dado que S es separable para cada $n \in \mathbb{N}$, $\exists\, A_{n_1}, A_{n_2}, \ldots$ bolas abiertas de radio $1/n$ tales que cubren a S. Como $\bigcup_{j=1}^{k} A_{n_j} \uparrow S$ si $k \to \infty$, entonces

$$P\left(\bigcup_{j=1}^{k} A_{n_j}\right) > 1 - \epsilon/2^n \qquad \text{si} \quad k \geq \lambda_n.$$

Veamos que el conjunto $A = \bigcap_{n=1}^{\infty} \bigcup_{j=1}^{\lambda_n} A_{n_j}$ es precompacto.* Sea $\alpha > 0$, entonces $\exists\, N \in \mathbb{N}$, tal que $\frac{1}{N} < \frac{\alpha}{2}$.

Como $A \subseteq \bigcup_{j=1}^{\lambda_n} A_{n_j}$ y dado que las bolas A_{n_j} tienen diámetro $\frac{2}{N} < \alpha$ se tiene que A es precompacto y por tanto tiene cerradura compacta. Entonces :

$$\begin{aligned}
P(\overline{A}) &\geq P\left(\bigcap_{n=1}^{\infty} \bigcup_{j=1}^{\lambda_n} A_{n_j}\right) \\
&= 1 - P\left(\bigcup_{n=1}^{\infty} \bigcap_{j=1}^{\lambda_n} A^c_{n_j}\right) \\
&\geq 1 - \sum_{n=1}^{\infty} P\left(\bigcap_{j=1}^{\lambda_n} A^c_{n_j}\right) \\
&= 1 - \sum_{n=1}^{\infty} \left(1 - P\left(\bigcup_{j=1}^{\lambda_n} A_{n_j}\right)\right) \\
&> 1 - \sum_{n=1}^{\infty} \epsilon/2^n = 1 - \epsilon.
\end{aligned}$$

Ejemplos 1.13

1. Como un ejemplo particular del teorema 1.12, en un espacio de Hilbert (un espacio normado, el cual es completo y separable), cualquier medida de probabilidad que ahí se defina, resulta ser tensa.

---

* ver apéndice para definición

2. Si consideramos un subconjunto de $[0,1]$ con la métrica discreta y no medible, por ejemplo el conjunto de Vitali, y le asociamos una medida definida como la medida exterior, entonces en efecto tal medida es una probabilidad sobre ese espacio con la σ-álgebra generada por la métrica, ya que dicho conjunto tiene medida exterior 1 (y medida interior 0), a pesar de esto tomando un compacto contenido, por ser medible, coinciden las medidas exterior con la interior y ya que tiene medida interior cero (por estar contenido), su medida es cero, de tal suerte que la medida así definida no es tensa.

3. Sea S un espacio de Hilbert, con una base ortonormal $x_1, x_2, x_3 \ldots$ numerable ya que S es separable y completo cualquier medida de probabilidad sobre S es tensa. Sin embargo ningún conjunto con interior no vacio es compacto y por lo tanto S no es compacto, ni localmente compacto y tampoco σ-compacto. Sea P una medida de probabilidad sobre S que asigna masa positiva a cada y_n $\forall n$ con $y_n \in D$ y con D denso en S. Entonces P no tiene un soporte localmente compacto en la topología relativa.

En efecto :

(a) Ningún conjunto con interior no vacio es compacto.

Supongamos que hay un conjunto compacto $A \in S$ tal que $A^\circ \neq \emptyset$. Sea $x_0 \in A^\circ$ entonces $\exists\, \epsilon > 0$ tal que $B_\epsilon(x_0) \subseteq A^\circ$. Por tanto

$$x_0 + \frac{\epsilon}{2}x_n \in B_\epsilon(x_0) \quad \forall n,$$

pues $d(x_0 + \frac{\epsilon}{2}x_n, x_0) = \frac{\epsilon}{2} < \epsilon$, por tanto $x_0 + \frac{\epsilon}{2}x_n \in A$ $\forall n$. Por otra parte $A \subseteq \bigcup_{x\in A} B_{\epsilon/4}(x)$ como A es compacto $\exists\, F$ finito $F \subseteq A$ tal que

$$A \subseteq \bigcup_{x\in F} B_{\epsilon/4}(x)$$

Ahora dado que A contiene una infinidad de puntos de la forma

$$x_0 + \frac{\epsilon}{2}x_n$$

entonces, al menos dos de ellos pertenecen a una misma bola, digamos $B_{\epsilon/4}(x'), x' \in F$ por tanto

$$d(x_0 + \frac{\epsilon}{2}x_n, x_0 + \frac{\epsilon}{2}x_m) < \epsilon/2$$

pero

$$d(x_0 + \frac{\epsilon}{2}x_n, x_0 + \frac{\epsilon}{2}x_m) = \frac{\epsilon}{2} \parallel x_n - x_m \parallel = \frac{\epsilon}{2}\sqrt{2} = \frac{\epsilon}{\sqrt{2}} > \frac{\epsilon}{2}$$

2. TEOREMA DE PORMANTEAU

El problema de determinar si una sucesión de medidas de probabilidad converge debilmente o no puede atacarse mediante algunas caracterizaciones de convergencia débil, las cuales se darán en este capítulo, de ahí su gran importancia en el resto de este trabajo y la tracendencia en el desarrollo de la probabilidad.

Observación 2.1

$d(x,A) \leq d(x,y) + d(y,A) \quad \forall y \in S$

En efecto, sea $y \in S$ entonces $d(x,z) \leq d(x,y) + d(y,z) \quad \forall z \in A$ tomando infimos sobre z se obtiene el resultado.

Sea $\epsilon > 0$ y A un subconjunto de S, denotemos por A^{ϵ} a el conjunto

$$\{x \in S \mid d(x,A) < \epsilon\}.$$

Observación 2.2 .

A^{ϵ} es abierto.

Demostración

Sea $y \in A^{\epsilon} \Rightarrow d(y,A) < \epsilon$ sea $\epsilon^* = \epsilon - d(x,A) > 0$, probaremos que

$$B_{\epsilon^*}(y) \subseteq A^{\epsilon}.$$

Si $x \in B_{\epsilon^*}(y)$ entonces $d(x,y) < \epsilon^*$, luego $d(x,y) + d(y,A) < \epsilon$ por la observación 2.1 $d(x,A) \leq d(x,y) + d(y,A) < \epsilon$, lo cual prueba la observación.

Proposición 2.3

Sea $B_A = \{x \in S - A \mid \exists\, y \in A,\ d(x,y) < \epsilon\}$ entonces $A^{\epsilon} = A \bigcup B_A$.

Demostración

$\supseteq$) Si $x \in A$ entonces $x \in A^{\epsilon}$. Si $x \in B_A$ entonces $\exists\, y \in A, d(x,y) < \epsilon$, por tanto $d(x,A) = \inf\{d(x,z) | z \in A\} \leq d(x,y) < \epsilon$ y $x \in A^{\epsilon}$.

$\subseteq$) Sea $x \in A^{\epsilon}$.

1) Si $x \in A$ es inmediato.

2) Si $x \notin A$ como

$$d(x,A) = \inf\{d(x,z) | z \in A\}$$

para $d(x,A) < r < \epsilon$ $\exists\, z \in A$ tal que

$$d(x,A) \leq d(x,z) < r < \epsilon,$$

por tanto $x \in B_A$.

Observación 2.4

$(S - A^\epsilon) \bigcap A = \emptyset$

Demostración

$x \in S - A^\epsilon \Rightarrow d(x,y) \geq \epsilon$ por tanto $x \notin A$.

Observación 2.5

$(S - A^\epsilon)^\epsilon \subseteq S - A.$

Demostración

Por la proposicíon 2.3 $(S - A^\epsilon)^\epsilon = (S - A^\epsilon) \bigcup B_{(S-A^\epsilon)}$. Nótese que :

$$B_{S-A^\epsilon} = \{x \in S - (S - A^\epsilon) |\ \exists\, y \in S - A^\epsilon, d(x,y) < \epsilon\}.$$

1) Si $x \in S - A^\epsilon$ entonces por la observacíon 2.4 $x \in S - A$.

2) Si $x \notin S - A^\epsilon$ entonces $x \in B_{(S-A^\epsilon)}$, es decir

$$x \in A^\epsilon \text{ y } \ \exists\, y \in S - A^\epsilon \text{ tal que } d(x,y) < \epsilon.$$

Suponiendo que $x \in A$ y dado que $\epsilon \leq d(y,A)$, pues $y \in S - A^\epsilon$ llegamos a que

$$\epsilon \leq d(y,A) \leq d(y,x) < \epsilon$$

lo cual es un absurdo, por lo tanto $x \in S - A$.

Nótese que la contencíon inversa en la observación 2.5 no siempre es verdadera, ya que $(S - A^\epsilon)^\epsilon$ es siempre abierto y $S - A$ no siempre lo es.

Observación 2.6

$(A^\alpha)^\beta \subseteq A^{\alpha+\beta}.$

Demostración

Sea $x \in (A^\alpha)^\beta$ como $d(x, A^\alpha) < \beta$ entonces $\exists\, z \in A^\alpha, d(x,A^\alpha) \leq d(x,z) < r$. Ahora $d(x,A) \leq d(x,z) + d(z,A) < \alpha + \beta$.

Sea $\hat{P}(S)$ el conjunto de medidas de probabilidad definidas sobre S. Sea $\lambda: \hat{P}(S) \times \hat{P}(S) \to \mathbb{R}$ definida como:

$$\lambda(P,Q) = \inf\{\epsilon > 0 | P(A) < Q(A^{\epsilon}) + \epsilon, \quad \forall A \in S\}$$

Teorema 2.7

λ es una metrica en $\hat{P}(S)$ llamada métrica de Prohorov.

Demostración

(i) λ es reflexiva. En efecto, sea $\epsilon > 0$ tal que $P(A) < Q(A^{\epsilon}) + \epsilon \quad \forall A \in S$ en particular para $S - A^{\epsilon}$, se tiene $P(S - A^{\epsilon}) < Q((S - A^{\epsilon})^{\epsilon}) + \epsilon$ y por la observación 2.4

$$Q((S - A^{\epsilon})^{\epsilon}) \leq Q(S - A)$$

por tanto

$$1 - P(A^{\epsilon}) = P(S - A^{\epsilon}) < Q(S - A^{\epsilon}) + \epsilon = 1 - Q(A) + \epsilon$$

luego $Q(A) < P(A^{\epsilon}) + \epsilon$ como A fue arbitrario entonces

$$Q(A) < P(A^{\epsilon}) + \epsilon \quad \forall A \in S$$

lo cual quiere decir que

$$\{\epsilon > 0 | \ P(A) < Q(A^{\epsilon}) + \epsilon \quad \forall A \in S\} = \{\epsilon > 0 | \ Q(A) < P(A^{\epsilon}) + \epsilon \quad \forall A \in S\}$$

por tanto $\lambda(P,Q) = \lambda(Q,P)$.

(ii) λ es simétrica.

En efecto, si $P \equiv Q$ y si $\delta > 0$ y $A \in S$ entonces $Q(A) \leq Q(A^{\delta})$ por tanto $Q(A) + \delta \leq Q(A^{\delta}) + \delta$ y por tanto

$$P(A) \leq Q(A) < Q(A^{\delta}) + \delta \quad \forall A \in S,$$

así que $\lambda(P,Q) = 0$.

Súpongase ahora que que $\lambda(P,Q) = 0$. Sea $A \in S$ con A cerrado, entonces :

$$P(A) < Q(A^{1/n}) + 1/n \quad \forall A \in S \quad \forall n \in \mathbb{N}.$$

Nótese que $A^{1/n} \downarrow A$ ya que A es cerrado, por tanto

$$P(A) = \lim_{n\to\infty} P(A) \leq \lim_{n\to\infty}(Q(A^{1/n}) + 1/n) = Q(A).$$

aplicando el mismo razonamiento, y en virtud de que $\lambda(Q,P) = 0$ (por (i)), se tiene que para $S - A$ se satisface

$$P(S - A) \leq Q(S - A)$$

es decir $Q(A) \leq P(A)$ y ya que las medidas coindciden en los cerrados se tiene $P \equiv Q$.

(iii) Sean $P, Q, R \in \hat{P}(S)$, probaremos que si

$$B_1 = \{\epsilon > 0 | P(A) < Q(A^\epsilon) + \epsilon \quad \forall A \in S\}$$
$$B_2 = \{\epsilon > 0 | Q(A) < R(A^\epsilon) + \epsilon \quad \forall A \in S\}$$
$$B_3 = \{\epsilon > 0 | P(A) < R(A^\epsilon) + \epsilon \quad \forall A \in S\}$$

entonces $B_3 \supseteq B_1 + B_2$.

Sea $\epsilon \in B_1, \delta \in B_2$, entonces se cumple que

$$P(A) < Q(A^\epsilon) + \epsilon < R((A^\epsilon)^\delta) + \delta + \epsilon \leq R(A^{\epsilon+\delta}) + \delta + \epsilon$$

por tanto $\delta + \epsilon \in B_3$ como queriamos probar.

En virtud de esto $\inf(B_1 + B_2) \geq \inf(B_3)$ es decir

$$\lambda(P, Q) \leq \lambda(P, R) + \lambda(R, Q).$$

Proposición 2.8

Si $\mathfrak{L}$ es una familia de subconjuntos que contiene a los cerrados de S, contenida en S y

$$\hat{A} = \{\epsilon > 0 | P(A) < Q(A^\epsilon) + \epsilon, \forall A \in S\}$$
$$\hat{B} = \{\epsilon > 0 | P(A) < Q(A^\epsilon) + \epsilon, \forall A \in L\}$$

entonces $\inf \hat{A} = \inf \hat{B}$, es decir da lo mismo tomar la métrica de Prohorov en S que en cualquier subfamilia que contenga a los cerrados de S.

Demostración

De la demostración de que λ es una medida simétrica se sigue que

$$\inf\{\epsilon > 0 | P(A) < Q(A^\epsilon) + \epsilon, \forall A \in L\} = \inf\{\epsilon > 0 | Q(A) < P(A^\epsilon) + \epsilon, \forall A \in L\}$$

Sea $\rho_A = \inf \hat{A}$ y $\rho_B = \inf \hat{B}$. Es claro que como $\hat{A} \subseteq \hat{B}$ entonces $\inf \hat{A} \geq \inf \hat{B}$, es decir $\rho_A \geq \rho_B$. Súpongase que $\rho_A > \rho_B$ entonces $\exists\ \epsilon_0 \in \hat{B}$ tal que

$$\rho_A > \epsilon_0 \geq \rho_B,$$

como $\epsilon_0 \notin \hat{A}$ se tiene que $\exists\ A \in S$ con

$$Q(A) \geq P(A^{\epsilon_0}) + \epsilon_0.$$

$S - A$ es cerrado y como $\epsilon_0 \in \hat{B}$ se tiene que

$$\begin{aligned}
1 - P(A^\epsilon) &= P(S - A^{\epsilon_0}) \\
&< Q((S - A^{\epsilon_0})^{\epsilon_0}) + \epsilon_0 \\
&\leq Q(S - A) + \epsilon_0 \\
&= 1 - Q(A) + \epsilon_0
\end{aligned}$$

por tanto $Q(A) < P(A^{\epsilon_0}) + \epsilon_0$ lo cual es un absurdo. Así que $\inf \hat{A} = \inf \hat{B}$

Definición 2.9

Si $A \in S$ y $P(\partial A) = 0$ diremos que A es un conjunto de P-continuidad , o que es un conjunto P-continuo.

Denotemos por $\|f\|$ a el número $\sup_{x \in S} |f(x)|$ siempre que $f \in C(S)$.

Teorema 2.10(de Pormanteau) .

Supongase que S es separable, entonces las siguientes afirmaciones son equivalentes:

(1) $P_n \Rightarrow P$

(2) $\lim_{n\to\infty} \int_S f dP_n = \int_S f dP \quad \forall f$ acotada y uniformemente continua en S

(3) $\limsup P_n(F) \leq P(F) \quad \forall F \in S \quad F$ cerrado.

(4) $\liminf P_n(G) \geq P(G) \quad \forall G \in S \quad G$ abierto.

(5) $\lim_{n\to\infty} P_n(A) = P(A) \quad \forall A \in S$ con A P-continuo.

(6) $\lambda(P_n, P) \to 0$ si $n \to \infty$.

Demostración

(1) $\Rightarrow$ (2) Sea f acotada y uniformemente continua, entonces $f \in C(S)$ y como $P_n \Rightarrow P$ se tiene que $\lim_{n\to\infty} \int_S f dP_n = \int_S f dP$

(2) $\Rightarrow$ (3) Sea F cerrado, $F \in S$ y sea $\delta > 0$, dado que

$$G_\epsilon = \{x | d(x, F) < \epsilon\} \downarrow F \quad \text{si } \epsilon \to 0,$$

entonces $\exists\, \epsilon > 0$ tal que $P(G_\epsilon) < P(F) + \delta$ sabemos por el lema de Uryson que $\exists\, f \in C(S)$ tal que

$$f(x) = 1 \quad \forall x \in F \quad \text{y } f(x) = 0 \text{ si } x \in G_\epsilon$$

con $0 \leq f(x) \leq 1$ y f es uniformemente continua. Por (2)

$$\lim_{n\to\infty} \int_S f dP_n = \int_S f dP.$$

Ahora

$$P_n(F) = \int_F f dP_n \leq \int_S f dP_n$$

por tanto

$$\begin{aligned} \limsup P_n(F) &\leq \limsup \int_S f dP_n \\ &= \int_S f dP \\ &\leq \int_{G_\epsilon} f dP \\ &\leq P(G_\epsilon) < P(F) + \delta \end{aligned}$$

haciendo $\delta \to \infty$ se tiene el resultado.

(3) $\Rightarrow$ (4) Sea G abierto, $G \in S$ entonces para $(S - G)$ cerrado se tiene por hipótesis que

$$\begin{aligned} 1 - \liminf P_n(G) &= \limsup(1 - P_n(G)) \\ &= \limsup[P_n(S) - P_n(G)] \\ &= \limsup[P_n(S - G)] \\ &\leq P(S - G) \\ &= 1 - P(G) \end{aligned}$$

por tanto $P(G) \leq \liminf P_n(G)$.

(4) $\Rightarrow$ (5) Sea $A \in S$ tal que $P(\partial A) = 0$ como $\partial A = \bar{A} - A^\circ$ entonces $0 = P(\bar{\partial} A) - P(A^\circ)$ por tanto :

$$\begin{aligned} P(\bar{A}) &= P(A^\circ) \\ &\leq \liminf P_n(A^\circ) \\ &\leq \liminf P_n(A) \\ &\leq \limsup P_n(A) \\ &\leq \limsup P_n(\bar{A}) \\ &\leq P(\bar{A}) = P(A^\circ), \end{aligned}$$

entonces $\lim_{n\to\infty} P_n(A) = P(\bar{A}) = P(A^\circ) = P(A)$. Análogamente (5) $\Rightarrow$ (4).

(5) $\Rightarrow$ (1) Sea $f \in C(S)$ con $f \geq 0$ entonces $\partial\{f \geq t\} \subseteq \{f = t\}$, por tanto $\{f \geq t\}$ es un conjunto $P-$continuo para casi todos los t excepto quizá, una cantidad a lo mas numerable de $t \geq 0$. Entonces

$$\begin{aligned} \lim_{n\to\infty} \int_S f dP_n &= \lim_{n\to\infty} \int_0^{\|f\|} P_n\{f \geq t\} dt \\ &= \int_0^{\|f\|} P\{f \geq t\} dt = \int_S f dP \end{aligned}$$

†

Aplicando esto a $(\|f\| - f)$ y $(\|f\| + f)$ obtenemos el resultado.

Nótese que hasta aquí las 5 condiciones anteriores son equivalentes y no se ha usado la separabilidad del espacio.

Mostraremos que (4) $\Rightarrow$ (6) y (6) $\Rightarrow$ (1) lo cual terminará la prueba.

(4) $\Rightarrow$ (6) Sea $\epsilon > 0$ ya que S es separable, $\exists\, E_1, E_2, \ldots \in S$ tales que cubren a S y tienen diámetro menor que $\epsilon/2$. Sea $F_n = \bigcup_{i=1}^n E_i$, es claro que $F_n \uparrow S$, por

---

† Para la primera igualdad, ver notas de seminario de probabilidad, Ma.E.Caballero, B. Fernandez.

continuidad de P se tiene que $P(F_n) \uparrow 1$, por tanto $\exists\, n$ tal que

$$P(\bigcup_{i=1}^{n} E_i) = P(F_n) \geq 1 - \epsilon/2,$$

sea N el mas pequeño de ellos. Y sea

$$\mathcal{L} = \left\{ (\bigcup_{i \in I} E_i)^{\epsilon/2} \mid I \subseteq \{1, 2, \ldots, N\} \right\}.$$

Ya que para cada $G \in \mathcal{L}$, $\liminf P_n(G) \geq P(G)$, pues G es abierto, para $\epsilon > 0$ $\exists M$, tal que

$$P(G) \leq P_n(G) + \epsilon/2 \quad \forall n \geq M$$

y dado que $\mathcal{L}$ es finita, podemos elegir n_0 tal que

$$P(G) \leq P_n(G) + \epsilon/2 \quad \forall n \geq n_0 \quad \forall G \in \mathcal{L}\,.$$

Sea F cerrado, y sea

$$F_0 = \bigcup \{E_i | 1 \leq i \leq N, E_i \cap F \neq \emptyset\}$$

Entonces $F_0^{\epsilon/2} \in \mathcal{L}$ y además

$$\begin{aligned} P(F) &= P(F \cap F_0) + P(F \cap (S - F_0)) \\ &\leq P(F_0) + P(S - F_0) \\ &\leq P(F_0^{\epsilon/2}) + 1 - P(F_0) \\ &\leq P(F_0^{\epsilon/2}) + 1 - P\left(\bigcup_{i=1}^{N} E_i\right) \\ &\leq P(F_0^{\epsilon/2}) + \epsilon/2\,. \end{aligned}$$

Notemos que $F_0^{\epsilon/2} \subseteq F^{\epsilon}$ pues los E_i tienen diámetro menor que epsilon. Por tanto :

$$P(F) \leq P({F_0}^{\epsilon/2}) + \epsilon/2 \leq P_n({F_0}^{\epsilon/2}) + \epsilon \leq P_n(F^{\epsilon}) + \epsilon \quad \forall n \geq n_0\,.$$

Gracias a la proposición 2.8, y dado que ϵ fue arbitraria se tiene que :

$$\lambda(P_n, P) \to 0 \text{ si } n \to \infty.$$

(6) $\Rightarrow$ (1) Para cada $n \in \mathbb{N}$ consideremos $\epsilon_n = \lambda(P_n, P) + 1/n$. Sea $f \in C(S)$ supongamos que $f \geq 0$. Ahora

$$\int_S f\,dP_n = \int_0^{\|f\|} P_n\{f \geq t\}dt \leq \int_0^{\|f\|} P(\{f \geq t\}^{\epsilon_n})dt + \epsilon_n \|f\|.$$

Ya que esto pasa para toda n tomando límite superior obtenemos :

$$\begin{aligned}\limsup_{n\to\infty}\int_S f\,dP_n &\leq \lim_{n\to\infty}\int_0^{\|f\|} P(\{f\geq t\}^{\epsilon_n})dt\\ &= \int_0^{\|f\|} P(\{f\geq t\})dt = \int_S f\,dP.\end{aligned}$$

Nótese que en la penúltima desigualdad se utilizo el hecho de que $\{f\geq t\}$ es cerrado y por tanto

$$\{f\geq t\}^{\epsilon_n}\downarrow\{f\geq t\} \text{ si } n\to\infty$$

Aplicando esto a $\|f\|-f$ y $\|f\|+f$ obtenemos :

(*) $$\limsup_{n\to\infty}\int_S(\|f\|-f)dP_n\leq\int_S(\|f\|-f)dP.$$

(**) $$\limsup_{n\to\infty}\int_S(\|f\|+f)dP_n\leq\int_S(\|f\|+f)dP.$$

y entonces de (*) se tiene

$$\begin{aligned}\int_S\|f\|dP_n-\liminf_{n\to\infty}\int_S f\,dP_n &= \limsup_{n\to\infty}\int_S\|f\|dP_n-\liminf_{n\to\infty}\int_S f\,dP_n\\ &\leq\int_S\|f\|dP-\int_S f\,dP\end{aligned}$$

Por tanto

$$\liminf_{n\to\infty}\int_S f\,dP_n\geq\int_S f\,dP$$

análogamente por (**) se tiene

$$\limsup_{n\to\infty}\int_S f\,dP_n\leq\int_S f\,dP$$

De lo cual se deduce que :

$$\lim_{n\to\infty}\int_S f\,dP_n=\int_S f\,dP.$$

Nótese que la condición 6 implica cualquiera de las 5 primeras, aunque el espacio no sea separable.

El espacio $\hat{P}(S)$ con la métrica de Prohorov es completo, y si S es separable también lo es $\hat{P}(S)$.

A manera de información, existe otra métrica que así como la de Prohorov, bajo condiciones de separabilidad, metriza a la topología de la convergencia débil, tal es llamada la métrica dual acotada de Lipschitz. En el caso de $S=\mathbb{R}$ existe otra métrica llamada métrica de Levy, la cual metriza a la convergencia débil, pero no es equivalente a la de Prohorov.†

---

†ver Araujo and Gine

Teorema 2.11

Sea U una subclase de S tal que

(i) U es cerrada bajo intersecciones finitas.

(ii) Cada abierto de S es una unión finita o numerable de elementos de U Y si $P_n(A) \to P(A) \quad \forall A \in U$

Entonces $P_n \Rightarrow P$.

Demostración

Probaremos por inducción sobre m que

$$(*) \qquad P_n(\bigcup_{i=1}^{m} A_i) \to P(\bigcup_{i=1}^{m} A_i), \quad A_i \in U$$

1) m=1 es inmediato.

2) Supóngase cierto (*) para $A_1, A_2, \ldots A_m \in U$

3)

$$\begin{aligned}
P_n(\bigcup_{i=1}^{m+1} A_i) &= P_n(\bigcup_{i=1}^{m} A_i \bigcup A_{m+1}) \\
&= P_n(\bigcup_{i=1}^{m} A_i) + P_n(A_{m+1}) - P_n\left((\bigcup_{i=1}^{m} A_i) \bigcap A_{m+1}\right) \\
&= P_n(\bigcup_{i=1}^{m} A_i) + P_n(A_{m+1}) - P_n\left(\bigcup_{i=1}^{m} (A_i \bigcap A_{m+1})\right) \\
&\to P(\bigcup_{i=1}^{m} A_i) + P(A_{m+1}) - P\left(\bigcup_{i=1}^{m} (A_i \bigcap A_{m+1})\right) \\
&= P(\bigcup_{i=1}^{m+1} A_i)
\end{aligned}$$

Sea G abierto entonces $G = \bigcup_{i=1}^{\infty} A_i$ para alguna sucesión $\{A_i\}$ en S.

Como $\bigcup_{i=1}^{m} A_i \uparrow G$ si $m \to \infty$ por continuidad de P

para $\epsilon > 0.\ \exists\, m \in \mathbb{N}$ tal que $P(G) - P(\bigcup_{i=1}^{m} A_i) < \epsilon$. Entonces

$$\begin{aligned}
P(G) - \epsilon &< P((\bigcup_{i=1}^{m} A_i) \\
&= \lim_{n \to \infty} P_n(\bigcup_{i=1}^{m} A_i) \\
&= \liminf_{n \to \infty} P_n(\bigcup_{i=1}^{m} A_i) \\
&\le \liminf_{n \to \infty} P_n(G)\ .
\end{aligned}$$

Como ϵ fué arbitraria se tiene $P(G) \leq \liminf_{n\to\infty} P_n(G)$ y por el teorema 2.10 se tiene el resultado.

Corolario 2.12

Sea U una subclase de S con S separable tal que

(i) U es cerrada bajo intersecciones finitas,

(ii) $\forall x \in S$ y $\epsilon > 0$ $\exists\ A \in U$ tal que $x \in A^\circ \subseteq A \subseteq B_\epsilon(x)$

si $P_n(A) \to P(A) \quad \forall A \in U$

Entonces $P_n \Rightarrow P$.

Demostración

Sea G un abierto en S y D el denso numerable en S .

Probaremos que G es unión de elementos de U y aplicando el teoremea 2.11 se tendrá el resultado.

Sea $x \in G$ por ser G abierto $\exists\ \epsilon_x > 0$ tal que $B_{\epsilon_x}(x) \subseteq G$ en virtud de (ii) sabemos que $\exists\ A_{\epsilon_x} \in U$ tal que

$$x \in A^\circ_{\epsilon_x} \subseteq A_{\epsilon_x} \subseteq B_{\epsilon_x}(x)$$

por lo tanto $\bigcup_{x\in G} A^\circ_{\epsilon_x} \subseteq G$ y ya que $x \in G \Rightarrow x \in A^\circ_{\epsilon_x}$ se tiene que $G = \bigcup_{x\in G} A^\circ_{\epsilon_x}$. Por ser S separable

$$G = \bigcup_{i=1}^{\infty} A_i, \quad A_i \in U.$$

Corolario 2.13

Supóngase que para cada intersección finita A de esferas abiertas se tiene que $P_n(A) \to P(A)$ con A P-continuo. Si S es separable entonces $P_n \Rightarrow P$.

Demostración

Sea $x \in S$ ya que $\partial(B_r(x)) \bigcap \partial(B_{r'}(x)) = \emptyset$ si $r < r'$, entonces hay a lo mas una cantidad numerable de esferas alrededor de x tales que su frontera tiene medida no cero. Sea $\mathcal{A}$ la clase de esferas en S P-continuas y U la clase de intersecciones finitas de $\mathcal{A}$, entonces U satisface la condición (i) del corolario 1, para cada $x \in S$ y $\epsilon > 0$, $\exists\ \delta$ tal que $0 < \delta \leq \epsilon$ y $B_\delta(x) \in U \subseteq \mathcal{A}$ por tanto U satisface las hipótesis del corolario 1, y entonces se tiene el resultado.

El siguiente teorema puede parecer a simple vista facil, sin embargo será muy util en el desarrollo del siguiente capítulo.

Teorema 2.14

$P_n \Rightarrow P$ si y solo si cada subsucesión $\{P_{n'}\}$ contiene una subsucesión $\{P_{n''}\}$ tal que $P_{n''} \Rightarrow P$

Demostración

$\Rightarrow$) Sea $\{P_{n'}\}$ una subsucesión de $\{P_n\}$ por el teorema 2.10 se tiene que

$$\lim_{n\to\infty} P_n(A) = P(A) \quad \text{si } A \text{ es } P-\text{continuo.}$$

Así que $\lim_{n\to\infty} P_{n'}(A) = P(A)$ para todo A $P-$continuo empleando el mismo teorema se tiene que $P_{n'} \Rightarrow P$ por tanto $\{P_{n'}\}$ contiene una subsucesión que converge, a saber, ella misma.

$\Leftarrow$) Sea A un conjunto P-continuo y sea $a_n = P_n(A)$ para cada número natural n, y $a = P(A)$, consideremos $\{a_{n'}\}$ una subsucesión de a_n.

Ya que la sucesion correspondiente $\{P_{n'}\}$ contiene una subsucesión $\{P_{n''}\}$ que converge a P, entonces por el teorema 2.10 se tiene que $\lim_{n\to\infty} P_{n''}(A) = P(A)$ por tanto, la respectiva subsucesión $\{a_{n''}\}$ de $\{a_{n'}\}$ converge a a.

Luego, dado que cada subsucesión $\{a_{n'}\}$ de $\{a_n\}$ contiene una subsucesión $\{a_{n''}\}$ que converge a a, entonces $\{a_n\}$ converge a a, aplicando el teorema 2.10 se tiene que $P_n \Rightarrow P$.

3. CONVERGENCIA DEBIL EN ESPACIOS PRODUCTO.

Durante esta sección se verá como generalizar la convergencia débil en espacios productos, para que así, sabiendo la convergencia entrada por entrada, y adicionando alguna hipótesis se tenga el resultado.

Sea $S = S' \times S''$ con S', S'' espacios métricos.

Teorema 3.1

Si S es separable entonces una condición necesaria y suficiente para que $P_n \Rightarrow P$, es que

$$P_n(A' \times A'') \to P(A' \times A'')$$

$\forall A' \subseteq S'$ P' − continuo y $\forall A'' \subseteq S''$ P'' − continuo.

Donde P_n, P son medidas de probabilidad en $S = S' \times S''$ y

$$P'(A') = P(A' \times S''), \quad P''(A'') = P(A'' \times S')$$

las distribuciones marginales de P .

Demostración

$\Rightarrow$) Supóngase que A' es P' − continuo y que A'' es P'' − continuo entonces

$$\partial(A' \times A'') \subseteq (\partial A' \times S'') \bigcup (S' \times \partial A'').*$$

por tanto

$$\begin{aligned} P(\partial(A' \times A'')) &\leq P(\partial A' \times S'') + P(S' \times \partial A'') \\ &= P'(\partial A') + P''(\partial A'') \\ &= 0 + 0 \\ &= 0. \end{aligned}$$

Y entonces $A' \times A''$ es P − continuo, como $P_n \Rightarrow P$ por el teorema 2.10

$$P_n(A' \times A'') \to P(A' \times A'').$$

---

* ver apéndice

$\Leftarrow$) Sea U la clase de conjuntos de la forma $A' \times A''$ con A' P' – continuo y A'' P'' – continuo. Observemos que si $A' \times A''$, $B' \times B'' \in U$ entonces

$$(A' \times A'')\bigcap(B' \times B'') = (A'\bigcap B') \times (A''\bigcap B'') \in U.$$

Por tanto U es cerrada bajo intersecciones finitas. Dado que S es separable, en virtud del corolario 2.12; bastaria ver que $\forall \hat{x} \in S$, $\forall \epsilon > 0$ $\exists A \in U$ tal que

$$\hat{x} \in A^{\circ} \subseteq A \subseteq B_{\epsilon}(\hat{x}),$$

para que se tenga $P_n \Rightarrow P$ ya que las otras 2 condiciones se satisfacen.
Sea $(x', x'') \in S$, $\epsilon > 0$ y

$$A_{\delta} = \{x \in S' \mid d'(x, x') < \delta\} \times \{y \in S'' \mid d''(y, x'') < \delta\}$$

La métrica d en S está dada por

$$d((x', x''), (y', y'')) = \max\{d'(x', y'), d''(x'', y'')\}$$

entonces A_{δ} es la esfera con centro en (x', x'') y radio δ. Ahora si $\delta_1 < \delta_2$ entonces A_{δ_1} y A_{δ_2} tienen fronteras ajenas, por tanto $\exists$ $\epsilon > 0$ tal que $0 < \delta < \epsilon$ y $A_{\delta} \in B_{\epsilon}(x', x'')$ Por tanto

$$(x, x') \in A_{\delta}^{\circ} \subseteq A_{\delta} \subseteq B_{\epsilon}(x', x'')$$

y se tiene lo que se quería.

Teorema 3.2

Si S es separable entonces $P'_n \times P''_n \Rightarrow P' \times P''$ si y solo si $P'_n \Rightarrow P'$ y $P''_n \Rightarrow P''$

donde $P'_n \times P''_n(A' \times A'') = P'_n(A') \cdot P''_n(A'')$

Demostración

$\Rightarrow$) Supóngase que $P'_n \times P''_n \Rightarrow P' \times P''$ y que S es separable. Por el teorema 3.1 $P'_n \times P''_n(A' \times A'') \to P' \times P''(A' \times A'')$ $\forall A'$ P' – continuo $\forall A''$ P'' – continuo.

Ahora :

$$\begin{aligned} P'_n(A') \cdot P''_n(A'') &= P'_n \times P''_n(A' \times A'') \\ &\to P' \times P''(A' \times A'') = P'(A') \cdot P''(A'') \end{aligned}$$

en particular si A' es P'-continuo como S'' es P''-continuo entonces

$$P'_n(A') \cdot 1 = P'_n(A') \cdot P''_n(S'') \to P' \times P''(A' \times S'') = P'(A')$$

Por tanto $P'_n \Rightarrow P'$ análogamente $P''_n \Rightarrow P''$.

$\Leftarrow$) Supóngase que $P_n' \Rightarrow P'$ y $P_n'' \Rightarrow P''$.

Sean A' P' – continuo y A'' P'' – continuo. Entonces $P_n'(A') \to P'(A')$ y $P_n''(A'') \to P''(A'')$ por tanto

$$\begin{aligned} P_n' \times P_n''(A' \times A'') &= P_n'(A') \cdot P_n''(A'') \\ &\to P'(A') \cdot P''(A'') = P' \times P''(A' \times A'') \end{aligned}$$

Como S es separable por el teorema 3.1 se tiene el resultado.

4. CONVERGENCIA DEBIL DE VARIABLES ALEATORIAS.

Aplicando la sección anterior podemos ahora entrar en lo que se llama : convergencia en distribución de variables aleatorias.

Sean $(\Omega, \beta, \mathrm{P})$ y $(\Omega_n, \beta_n, \mathrm{P}_n)$ espacios de probabilidad y X, X_n variables aleatorias definidas en $(\Omega, \beta), (\Omega_n, \beta_n, \mathrm{P}_n)$ respectivamente y con valores en $(S, \mathcal{S})$ con S espacio métrico.

La distribución de la variable aleatoria X es una medida de probabilidad P_X en $(S, \mathcal{S})$ definida por

$$P_X(A) = \mathrm{P}(X \in (A)) \quad A \in \mathcal{S}.$$

☞ *Definición 4.1*

Se dice que $\{X_n\}$ converge en distribución a X (y lo denotamos por $X_n \xrightarrow{D} X$), si las correspondictes funciones de distribución convergen debilmente a la distribución de X. Es decir :

$$X_n \xrightarrow{D} X \iff P_{X_n} \Rightarrow P_X \quad ,$$

☞ *Definición 4.2*

Se dice que X_n converge en probabilidad a una constante a, $a \in S$ (y lo denotamos por $X_n \xrightarrow{\mathrm{P}} a$), si $\forall \epsilon > 0$

$$\mathrm{P}_n\{w \in \Omega_n \mid d(X_n(w), a) \geq \epsilon\} \to 0 \quad \text{si } n \to \infty$$

☞ *Definición 4.3*

Si $X_n : \Omega \to S$ son variables aleatorias y S es un espacio métrico separable y si para cada $\epsilon > 0$

$$\mathrm{P}_n\{w \in \Omega \mid d(X_n(w), X(w)) \geq \epsilon\} \to 0 \quad \text{si } n \to \infty$$

Entonces decimos que X_n converge en probabilidad a X y lo denotaremos por $X_n \xrightarrow{\mathrm{P}} X$.

Definición 4.4

Decimos además que $A \in S$ es un conjunto de X-continuidad ó que es P_X-continuo si

$$P_X(\partial A) = \mathrm{P}\{X \in \partial A\} = 0.$$

Observación 4.5

Supóngase que para cada $n, X_n, Y_n: (\Omega_n, \beta_n) \to (S, \mathcal{S})$ (es decir son medibles), entonces tiene setido hablar de $Z_n = d(X_n, Y_n): \Omega_n \to \mathbb{R}$, en virtud de que S es separable Z_n es medible.* Por tanto se puede hablar de la convergencia en probabilidad de Z_n a la constante $0 \in \mathbb{R}$.

Teorema 4.6

Si S es separable y $X_n \xrightarrow{D} X$ y $d(X_n, Y_n) \xrightarrow{P} 0$ entonces $Y_n \xrightarrow{D} X$.

Demostración

Sea $F_\epsilon = \{x \in S \mid d(x, F) \leq \epsilon\}$, F_ϵ es cerrado. Probaremos primero que

$$(*) \qquad \{Y_n \in F\} \subseteq \{d(X_n, Y_n) \geq \epsilon\} \bigcup \{X_n \in F_\epsilon\}.$$

Sea $w \in \{Y_n \in F\}$. Si $X_n(w) \in F_\epsilon$ es clara la contención. Si $X_n(w) \notin F_\epsilon$ entonces

$$\epsilon < d(X_n(w), F) \leq d(X_n(w), Y_n(w)).$$

Por tanto $w \in \{d(X_n, Y_n) \geq \epsilon\}$. De (*) se sigue que

$$P_{Y_n}(F) \leq \mathrm{P}_n\{d(X_n, Y_n) \geq \epsilon\} + P_{X_n}(F_\epsilon)$$

Ahora :

$$P_{Y_n}(F) \leq \limsup \mathrm{P}\{d(X_n, Y_n) \geq \epsilon\} + P_{X_n}(F_\epsilon)$$

Como $d(X_n, Y_n) \xrightarrow{P} 0$, entonces

$$\mathrm{P}_n\{|d(X_n(w), Y_n(w)) - 0| \geq \epsilon\} = \mathrm{P}_n\{d(X_n(w), Y_n(w)) \geq \epsilon\} \to 0.$$

Por tanto

$$\limsup P_{Y_n}(F) \leq \limsup P_{X_n}(F_\epsilon) \leq P_X(F_\epsilon).$$

Tomando F cerrado, se tiene que $F_\epsilon \downarrow F$ y en tal caso $P_X(F_\epsilon) \downarrow P_X(F)$, haciendo epsilon tender a cero, se tiene que:

$$\limsup P_{Y_n}(F) \leq P_X(F)$$

y por el teorema 2.10 obtenemos que $Y_n \xrightarrow{D} X$.

---

*ver apéndice

Teorema 4.7

Sea (Ω,β) un espacio de probabilidad y S un espacio métrico separable. Sean $X, X_k, Y_n, X_n^1, X_n^2, \ldots : (\Omega,\beta) \to (S,\mathcal{S})$ $k,n \in \mathbb{N}$ tales que :

(i) $\forall k \in \mathbb{N}$ $X_n^k \xrightarrow{D} X_k$ si $n \to \infty$,

(ii) $X_k \xrightarrow{D} X$ si $k \to \infty$

(iii) $\lim_{k\to\infty} \limsup \mathrm{P}\{d(X_n^k, Y_n) \geq \epsilon\} = 0 \quad \forall \epsilon > 0$,

Entonces $Y_n \xrightarrow{D} X$ si $n \to \infty$.

Demostración

Sea $F_\epsilon = \{x \in S \mid d(x,F) \leq \epsilon\}$. con F cerrado.

Se tiene análogamente $\{Y_n \in F\} \subseteq \{d(X_n^k, Y_n) \geq \epsilon\} \bigcup \{X_n^k \in F_\epsilon\}$. entonces

$$P_{Y_n}(F) \leq \mathrm{P}\{d(X_n^k, Y_n) \geq \epsilon\} + P_{X_n^k}(F_\epsilon)$$

por tanto

$$\limsup P_{Y_n}(F) \leq \limsup \mathrm{P}(\{d(X_{k_n}, Y_n) \geq \epsilon\}) + \limsup P_{X_n^k}(F_\epsilon)$$

Pero $X_n^k \xrightarrow{D} X_k$ si $n \to \infty$ por el teorema 2.10

$$\limsup_n P_{X_n^k}(F_\epsilon) =\leq P_{X_k}(F_\epsilon)$$

por tanto

$$\limsup_n P_{Y_n}(F) \leq \limsup_n \mathrm{P}(\{d(X_{k_n}, Y_n) \geq \epsilon\}) + \limsup_n P_{X_k}(F_\epsilon)$$

tomando límites cuando $k \to \infty$ obtenemos:

$$\lim_{k\to\infty} \limsup_n P_{Y_n}(F) \leq \lim_{k\to\infty} \limsup_n \mathrm{P}(\{d(X_n, Y_n) \geq \epsilon\}) + \lim_{k\to\infty} P_{X_k}(F_\epsilon)$$

y de la hipótesis nos queda que

$$\limsup P_{Y_n}(F) \leq P_X(F_\epsilon)$$

haciendo $\epsilon \to 0$ y aplicando el teorema 2.10 se tiene el resultado.

Teorema 4.8

Si $X_n \xrightarrow{P} X$ entonces

$$\mathrm{P}(\{X_n \in A\} \triangle \{X \in A\}) \to 0 \quad \forall A \ P_X - \text{continuo}$$

donde $B \triangle C = (B - C) \bigcup (C - B) = (B \bigcup C) - (B \bigcap C)$.

Demostración

Probaremos primero que

$$\{X_n \in A, X \notin A\} \subseteq \{d(X_n, X) \geq \epsilon\} \bigcap \{d(X, A) < \epsilon, X \notin A\} \; .$$

Sea $w \in \{X_n \in A, X \notin A\}$

1) Si $w \in \{d(X_n, X) \geq \epsilon\}$ es clara la contención.

2) Si $w \notin \{d(X_n, X) \geq \epsilon\}$ entonces

$$X_n(w) \in A, \; X(w) \notin A \text{ y } d(X_n(w), X(w)) < \epsilon$$

esto implica que:

$$d(X(w), A) \leq d(X_n(w), X(w)) < \epsilon$$

por tanto $w \in \{d(A, X) < \epsilon, X \notin A\}$.

Análogamente se tiene el resultado para $S - A$. Por tanto :

$$\mathrm{P}\{X_n \in A, X \notin A\} \leq \mathrm{P}\{d(X_n, X) \geq \epsilon\} + \mathrm{P}\{d(X, A) < \epsilon, X \notin A\}$$

y

$$\mathrm{P}\{X_n \notin A, X \in A\} \leq \mathrm{P}\{d(X_n, X) \geq \epsilon\} + \mathrm{P}\{d(X, S - A) < \epsilon, X \in A\} \; .$$

Como $X_n \xrightarrow{\mathrm{P}} X$ se tiene que $\mathrm{P}\{d(X_n, X) \geq \epsilon\} \to 0$ si $n \to \infty$ entonces

$$\begin{aligned}
&\limsup \mathrm{P}(\{X_n \in A\} \triangle \{X \in A\}) \\
&\qquad \leq \limsup \mathrm{P}(\{X_n \in A\} - \{X \in A\}) + \limsup \mathrm{P}(\{X \in A\} - \{X_n \in A\}) \\
&\qquad \leq \mathrm{P}\{d(X, A) < \epsilon, X \notin A\} + \mathrm{P}\{d(X, S - A) < \epsilon, X \in A\}
\end{aligned}$$

Es claro que : $\{d(X, A) < \epsilon\} \downarrow \bar{A}$ si $\epsilon \to 0$. Entonces

$$\{d(X, A) < \epsilon, X \notin A\} \downarrow \bar{A} - A \text{ si } \epsilon \to 0$$

y análogamente

$$\{d(X, S - A) < \epsilon, X \in A\} \downarrow \overline{(S - A)} - (S - A) \text{ si } \epsilon \to 0$$

Como $\overline{(S - A)} - (S - A) = A \bigcap \overline{(S - A)} = A - A^\circ$ entonces

$$\begin{aligned}
\mathrm{P}\{d(X, A) < \epsilon, X \notin A\} + \mathrm{P}\{d(X, S - A) < \epsilon, X \in A\} &\to \mathrm{P}(\bar{A} - A) + \mathrm{P}(A - A^\circ) \\
&= \mathrm{P}((\bar{A} - A) \bigcup (A - A^\circ)) \\
&= \mathrm{P}(\partial A) = 0
\end{aligned}$$

por tanto

$$0 \leq \liminf \mathrm{P}(\{X_n \in A\} \triangle \{X \in A\}) \leq \limsup \mathrm{P}(\{X_n \in A\} + \{X \in A\}) = 0$$

es decir $\mathrm{P}(\{X_n \in A\} \triangle \{X \in A\}) \to 0$.

Nótese que en general

$$B \subseteq (B - C) \bigcup (C - B) \bigcup C = (B \triangle C) \bigcup C$$

y

$$C \subseteq (B \triangle C) \bigcup B$$

entonces

$$\mathrm{P}(B) - \mathrm{P}(C) \leq \mathrm{P}(B \triangle C) \quad \text{y} \quad \mathrm{P}(C) - \mathrm{P}(B) \leq \mathrm{P}(B \triangle C)$$

por tanto $|\mathrm{P}(B) - \mathrm{P}(C)| \leq \mathrm{P}(B \triangle C)$.

Corolario 4.9

Si $X_n \xrightarrow{D} X$ entonces $X_n \xrightarrow{\mathrm{P}} X$.

Demostración

Como:

$$|P_{X_n}(A) - P_X(A)| = |\mathrm{P}\{X_n \in A\} - \mathrm{P}\{X \in A\}| \leq \mathrm{P}\left(\{X_n \in A\} \triangle \{X \in A\}\right)$$

y dado que $\mathrm{P}\left(\{X_n \in A\} \triangle \{X \in A\}\right) \to 0$ se tiene que $P_{X_n}(A) \to P_X(A)$ $\forall A$ $P_X -$ continuo y por el teorema 2.10 tenemos el resultado.

Teorema 4.10

Si $X_n{}' \xrightarrow{D} X'$ y $X_n{}'' \xrightarrow{\mathrm{P}} a''$ entonces $(X_n{}', X_n{}'') \xrightarrow{D} (X_n{}', a'')$ donde $X_n{}' : \Omega_n{}' \to S'$, $X_n{}'' : \Omega_n{}'' \to S''$, $S = S' \times S''$

Demostración

Por el teorema 3.1 $(X_n{}', X_n{}'') \xrightarrow{D} (X', X'')$ si y solo si

$$\mathrm{P}\left\{\left(X_n{}', X_n{}''\right) \in A' \times A''\right\} \to \mathrm{P}\left\{\left(X', X''\right) \in A' \times A''\right\}$$

$\forall A'$, $P_{X'}$ − continuo y $\forall A''$ $P_{X''}$ − continuo.

Por tanto basta probar que si A'' es $P_{X''}$ − continuo y A'' es $P_{X''}$ − continuo. entonces :

$$\mathrm{P}\left\{X_n{}' \in A', X_n{}'' \in A''\right\} \to \mathrm{P}\left\{X' \in A', X'' \in A''\right\}$$

para $X'' \equiv a''$

Sean pues A'', $P_{X''}$ − continuo. y A'' $P_{X''}$ − continuo.

La distribución de X'' es

$$P_{X''}(B) = \begin{cases} 0 & \text{si } a'' \notin B; \\ 1 & \text{si } a'' \in B. \end{cases}$$

1) Supongamos que $a'' \in A''$, como $X_n'' \xrightarrow{P} a''$ entonces $X_n'' \xrightarrow{D} a''$ en particular $P_{X_n''}(S'' - A'') \to P_{X''}(S'' - A'') = 0$

   Dado que

$$\begin{aligned} \mathrm{P}\left\{X_n' \in A'\right\} - \mathrm{P}\left\{X_n'' \notin A''\right\} &\leq \mathrm{P}\left\{X_n' \in A', X_n'' \in A''\right\} \\ &\leq P_{X_n'}(A') \\ &\to P_{X'}(A') \end{aligned}$$

   entonces

$$\mathrm{P}\left\{X_n' \in A', X_n'' \in A''\right\} \to P_{X'}(A') = \mathrm{P}\left\{X' \in A', a'' \in A''\right\}$$

2) Supongamos que $a'' \notin A''$ entonces

$$0 \leq \mathrm{P}\left(\left\{X_n' \in A', X_n'' \in A''\right\}\right) \leq P_{X_n''}(S'' - A'') \to P_{X''}(S'' - A'') = 0$$

Por tanto $(X_n', X_n'') \xrightarrow{D} (X_n', a'')$.

5. CONVERGENCIA DEBIL Y MAPEOS.

Sin duda alguna resultados muy hermosos son los expuestos en este capítulo, pues basta :

1) la convergencia débil en un espacio,

2) una función de este a otro espacio,

3) una hipótesis,

y se tiene, oh maravilla, la convergencia débil en este otro espacio.

Sea h una función, $h: S \to S'$ y sea $D_h = \{ x \in S \mid h \text{ es discontinua} \}$. Sean P_n, P medidas de probabildad sobre S', $n \in \mathbb{N}$, es sabido que $P_n h^{-1}$, Ph^{-1} son medidas de probabilidad sobre S', con $(P_n h^{-1})(A) = P_n(h^{-1}(A))$ $A \in S$.

Teorema 5.1

Si $P_n \Rightarrow P$ y $P(D_h) = 0$ entonces $P_n h^{-1} \Rightarrow Ph^{-1}$

Demostración

Mostraremos primero que $D_h \in S$. Sea d' la métrica en S' y sea

$$A_{\delta,\epsilon} = \{x \in S \mid \exists\, y, z \in S, d(x,y) < \delta, d(x,z) < \delta, d'(h(y), h(z)) \geq \epsilon\}$$

entonces $A_{\delta,\epsilon}$ es abierto, pues dada $x \in A_{\delta,\epsilon}$ y tomando

$$r = \min\{\delta - d(x,y), \delta - d(x,z)\}$$

y si $w \in B_r(x)$ se tiene que $\exists y, z \in S$ tales que :

$$d(w,y) \leq d(w,x) + d(x,y) < \delta - d(x,y) + d(x,y) = \delta$$

y

$$d(w,z) \leq d(w,x) + d(x,z) < \delta - d(x,y) + d(x,y) = \delta$$

y $d'(h(y), h(z)) \geq \epsilon$ por tanto $B_r(x) \subseteq A_{\delta,\epsilon}$.

Probaremos ahora que $D_h = \bigcup_{\epsilon>0} \bigcap_{\delta>0} A_{\delta,\epsilon}$ $\epsilon, \delta \in \mathbb{Q}$ y con esto se tendra que $D_h \in S$ pues es unión numerable de elementos de S.

$\subseteq$) $x \in D_h$ como h es discontinua en x, $\exists\, \epsilon > 0, \forall \delta > 0$ $\exists\, y, z \in B_\delta(x)$ con $d'(h(y), h(z)) \geq \epsilon$, como $\exists\, \epsilon^* \in \mathbb{Q}$, $\epsilon^* \leq \epsilon$ entonces $x \in A_{\delta,\epsilon^*}$ $\forall \delta > 0$ por tanto

$$x \in \bigcup_{\epsilon>0} \bigcap_{\delta>0} A_{\delta,\epsilon} \quad \epsilon, \delta \in \mathbb{Q}.$$

$\supseteq$) $x \in \bigcup_{\epsilon>0} \bigcap_{\delta>0} A_{\delta,\epsilon}$ $\epsilon, \delta \in \mathbb{Q}$. Sea $\delta^* > 0$, como $\exists\, \delta \in \mathbb{Q}^+$, $\delta \leq \delta^*$, y como $\exists\, \epsilon > 0$ tal que

$$\exists\, y, z \in B_\delta(x) \subseteq B_{\delta^*}(x) \text{ con } d'(h(y), h(z)) \geq \epsilon$$

entonces h es discontinua en x y por tanto $x \in D_h$.

Nótese que no se pidió la medibilidad de h, esto es importante pues dada cualquier función, su conjunto de discontinudades es medible.

Sea F cerrado en S, ya que $P_n \Rightarrow P$ entonces

$$\limsup P_n(h^{-1}(F)) \leq \limsup P_n(\overline{h^{-1}(F)}) \leq P(\overline{h^{-1}(F)})$$

En virtud de que $P(D_h) = 0$ bastaria ver que $\overline{h^{-1}(F)} \subseteq D_h \bigcup h^{-1}(F)$ para que (aplicando el teorema de Pormanteu) se tenga el resultado.

Sea $w \in \overline{h^{-1}(F)}$, si $w \in h^{-1}(F)$ se tiene la contencion, si $w \notin h^{-1}(F)$ entonces $h(w) \notin F$ y por ser F cerrado tenemos que $\exists \epsilon > 0$ tal que $d'(h(w), F) \geq \epsilon$, por tanto $\forall \delta > 0$ $\exists\, y \in B_\delta(w)$ tal que $y \in h^{-1}(F)$ (pues $w \in \overline{h^{-1}(F)}$) y tal que $h(y) \notin B_\epsilon(h(w))$ entonces $w \in D_h$.

Corolario 5.2

Si $X_n \xrightarrow{D} X$ y $\mathrm{P}(\{X \in D_h\}) = 0$ entonces $h(X_n) \xrightarrow{D} h(X)$ donde $h: (S, \mathcal{S}) \to (S', \mathcal{S}')$.

Demostración

Sean P_{X_n} las distribuciones de X_n y P_X la de X, sabemos que

$$P_{X_n} \Rightarrow P_X, \quad P_X(D_h) = \mathrm{P}\{X \in D_h\} = 0$$

Por el teorema 5.1 $P_{X_n}h^{-1} \Rightarrow P_X h^{-1}$ y como las distribuciones de $h(X_n)$ son $P_{X_n}h^{-1}$ y la de $h(X)$ es $P_X h^{-1}$ se tiene el resultado.

Corolario 5.3

Si $X_n \xrightarrow{\mathrm{P}} X \equiv a$ y h es continua entonces $h(X_n) \xrightarrow{\mathrm{P}} h(a)$.

Demostración

Como $X_n \xrightarrow{\mathrm{P}} a$ entonces $X_n \xrightarrow{D} a$ y como $D_h = \emptyset$ entonces por el corolario 5.1 $h(X_n) \xrightarrow{D} h(a)$, y por tanto $h(X_n) \xrightarrow{\mathrm{P}} h(a)$.

Teorema 5.4

En el caso $S' = \mathbb{R}$, $\mathcal{S}' = \mathcal{B}(\mathbb{R})$.

(i) Si $P_n \Rightarrow P$ entonces $P_n h^{-1} \Rightarrow P h^{-1}$ $\forall h: S \to \mathbb{R}$ medible con $P(D_h) = 0$

(ii) Si $P_n h^{-1} \Rightarrow P h^{-1}$ $\forall h$ real,continua y acotada, entonces $P_n \Rightarrow P$.

(iii) Si $P_n \Rightarrow P$ y h es real, medible y acotada con $P(D_h) = 0$ entonces $\int_S h\, dP_n \to \int_S h\, dP$.

Demostración

(i) Es inmediata aplicando el teorema 5.1

(ii) Sabemos que $P_n h^{-1} \Rightarrow P h^{-1}$ significa, por definición

$$\int_S f\, d(P_n h^{-1}) \to \int_S f\, d(P h^{-1}) \quad \forall f \in C(S)$$

Por el teorema de cambio de variable tenemos que

$$\int_{\mathbb{R}} f \circ h\, dP_n \to \int_{\mathbb{R}} f \circ h\, dP.$$

Ya que h es acotada $(|h| \leq M)$ consideremos:

$$f(x) = \begin{cases} M & x > M; \\ x & |x| \leq M; \\ -M & x < -M. \end{cases}$$

entonces $f \in C(\mathbb{R})$, y por tanto

$$\int_{\mathbb{R}} h\, dP_n \to \int_{\mathbb{R}} h\, dP$$

como h fue arbitraria en $C(S)$ se cumple que $P_n \Rightarrow P$.

(iii) Sea h real medible y acotada con $P(D_h) = 0$ como $P_n \Rightarrow P$ por (i) $P_n h^{-1} \Rightarrow P h^{-1}$ y por definición

$$\int f \circ h\, dP_n \to \int f \circ h\, dP \quad f \in C(\mathbb{R}),$$

en particular es válido para f como en (ii) y entonces se tiene que

$$\int_S h\, dP_n \to \int_S h\, dP.$$

Teorema 5.5

Si $X_n \xrightarrow{D} X$ entonces $E|X| \leq \liminf E|X_n|$

Demostración

Sea

$$h(x) = \begin{cases} |x| & |x| \leq \alpha; \\ 0 & |x| > \alpha. \end{cases}$$

Si $P_X(D_h) = \mathrm{P}\{|X| = \alpha\} = 0$ por el teorema 5.2 (iii) dado que $P_{X_n} \Rightarrow P_X$(P_{X_n} distribución de X_n y P_X distribución de X) se tiene que

$$\int_S h\, dP_{X_n} \to \int_S h\, dP_{X_n}$$

por tanto

$$\int_S |X|\, dP_n = \int_S |X_n|\, dP \to \int_S |X|\, dP$$

y entonces

$$\int_{\{|X|\leq\alpha\}} |X|\, dP = \lim_{n\to\infty} \int_{\{|X_n|\leq\alpha\}} |X_n|\, dP \leq \liminf E|X_n|,$$

ya que

$$L = \{\alpha \mid P\{|X| = \alpha\} > 0\}$$

es a lo mas numerable, entonces haciendo $\alpha \to \infty, \alpha \notin L$ se tiene

$$\int_S |X|\, dP = E|X| \leq \liminf E|X_n|.$$

Teorema 5.6

Supóngase que $X_n \xrightarrow{D} X$. Si $\{X_n\}$ es uniformemente integrable, es decir :

$$\lim_{\alpha\to\infty} \sup_n \int_{\{|X_n|\geq\alpha\}} |X_n|\, dP = 0$$

entonces $E(X_n) \to E(X)$. Si $X, X_n \geq 0$ son integrables y $E(X_n) \to E(X)$ entonces $\{X_n\}$ es uniformemente integrable.

Demostración

Como $\{X_n\}$ es uniformemente integrable entonces

$$\sup E(X_n) < \infty$$

y

$$E|X| \leq \liminf E|X_n| \leq \sup E(X_n) < \infty,$$

por tanto $|X|$ es integrable.

Ya que $X_n \xrightarrow{D} X$ para

$$h_\alpha(x) = \begin{cases} |x| & |x| \leq \alpha; \\ 0 & |x| > \alpha. \end{cases}$$

y si $P(\{|X| = \alpha\}) = 0$ por el teorema 5.2 (iii) $\int_S h_\alpha\, dP_n \to \int_S h_\alpha\, dP$ Por un teorema de cambio de variable

(1)
$$\int_S h_\alpha \circ X_n\, dP \to \int_S h_\alpha \circ X\, dP$$

Dado que:

(2)
$$E(X_n) - E(h_\alpha(X_n)) = \int_{\{|X_n|\geq\alpha\}} |X_n|\, dP$$

y

$$(3) \qquad E(X) - E(h_\alpha(X)) = \int_{\{|X| \geq \alpha\}} |X| \, dP$$

tenemos :

$$\begin{aligned}
&\limsup |E(X_n) - E(X)| \\
&= \limsup \left| \int_{\{|X_n| \geq \alpha\}} X_n \, dP - \int_{\{|X| \geq \alpha\}} X \, dP \right| + \limsup |E(h_\alpha(X_n)) - E(h_\alpha(X))| \\
&= \limsup \left| \int_{\{|X_n| \geq \alpha\}} X_n \, dP - \int_{\{|X| \geq \alpha\}} X \, dP \right| \\
&\leq \limsup \int_{\{|X_n| \geq \alpha\}} |X_n| \, dP + \limsup \int_{\{|X| \geq \alpha\}} |X| \, dP \\
&\leq \sup_n \int_{\{|X_n| \geq \alpha\}} |X_n| \, dP + \int_{\{|X| \geq \alpha\}} |X| \, dP,
\end{aligned}$$

haciendo $\alpha \to \infty$ se tiene que $\lim_{n \to \infty} |E(X_n) - E(X)| = 0$.

Ahora supongamos que $X_n, X \geq 0$ y que $E(X_n) \to E(X)$.

De (1),(2) y (3) se tiene que :

$$\int_{\{X_n \geq \alpha\}} X_n \, dP \to \int_{\{X \geq \alpha\}} X \, dP$$

y como $|X| = X$ es integrable entonces

$$\int_{\{X \geq \alpha\}} X \, dP < \epsilon$$

si α es suficientemente grande y por tanto

$$\int_{\{X_n \geq \alpha\}} X_n \, dP < \epsilon$$

si n es suficientemente grande y ya que cada X_n es integrable entonces $\{X_n\}$ es uniformemente integrable.

Teorema 5.7

Sean $h_n, h : (S, \mathcal{S}) \to (S', \mathcal{S}')$ y sea

$$E = \{x \in S \mid h_n(x_n) \to h(x) \text{ es falso para alguna subsucesion en } S \text{ con } x_n \to x\}$$

Si $P_n \Rightarrow P$ y si $P(E) = 0$ entonces $P_n h_n^{-1} \Rightarrow P h^{-1}$.

Demostración

$E \in \mathcal{S}$†. Sea G abierto, probaremos que $\liminf P_n h_n^{-1}(G) \geq P h^{-1}(G)$.

---

†ver apéndice

Antes que nada observemos que si $T_k = \bigcap_{n\geq k} h_n^{-1}(G)$ entonces

$$(*) \qquad h^{-1}(G) \subseteq E \bigcup \left(\bigcup_{k=1}^{\infty} T_k^{\circ} \right).$$

En efecto, tomemos $x \in h^{-1}(G)$

Si $x \in E$ es inmediata la contención.

Si $x \notin E$ y si x no está en el segundo uniendo de el lado derecho de (*), entonces $\forall k > 0 \ \exists \ i \geq k$ tal que

$$\forall \delta > 0, \ B_\delta(h_i^{-1}) \not\subseteq G$$

En particular para $k = n \ \exists \ i_n \geq k$ tal que para $\delta = 1/n > 0 \ \exists \ y_{i_n}$ con $d(x, y_{i_n}) < \delta$ y $h_{i_n}(y_{i_n}) \notin G$. Sea

$$x_m = \begin{cases} y_{i_n} & m = i_n; \\ x & m \neq i_n. \end{cases}$$

entonces $x_m \to x$ dado que $x \notin E$ por tanto $h_n(x_n) \to h(x)$ en particular $h_{i_n}(y_{i_n}) \to h(x)$ como G es abierto $\exists \ N_0 \in \mathbb{N}$ tal que $n \geq N_0 \Rightarrow h_{i_n}(y_{i_n}) \in G$ lo cual es un absurdo.

Dado que $T_k^{\circ} \subseteq T_{k+1}^{\circ}$ se tiene que

$$P\left(\bigcup_{k=1}^{\infty} T_k^{\circ} \right) = \lim_{n\to\infty} P(\left(\bigcup_{k=1}^{n} T_k^{\circ} \right) = \lim_{n\to\infty} P(T_n^{\circ})$$

y entonces para $r > 0 \ \exists \ N \in \mathbb{N}$ tal que

$$P\left(\bigcup_{k=1}^{\infty} T_k^{\circ} \right) < P(T_n^{\circ}) + r \qquad \text{si } n \geq N$$

y ya que $P_n \Rightarrow P$ se tiene que

$$P(T_n^{\circ}) \leq \liminf P_n(T_n^{\circ}) \leq \liminf P_n(h_n^{-1}(G)),$$

así que

$$\begin{aligned} P(h^{-1}(G)) &\leq P(E) + P\left(\bigcup_{k=1}^{\infty} T_k^{\circ} \right) \\ &= P\left(\bigcup_{k=1}^{\infty} T_k^{\circ} \right) \\ &< P(T_n^{\circ}) + r \\ &\leq \liminf P_n(T_n^{\circ}) + r \\ &\leq \liminf P_n(h_n^{-1}(G)) + r \end{aligned}$$

Como r fue arbitraria se tiene lo que se quería.

6. TEOREMA DE PROHOROV.

Consideremos el espacio de las funciones continuas sobre $[0,1]$ démosle la métrica del supremo a tales funciones, y llamémosle $C([0,1])$ a dicho espacio. A la σ-álgebra determinada por los abiertos inducidos por la métrica, llamémosla : $\mathcal{C}$ se sabe que tal espacio con esta métrica es completo.

Analizaremos ahora un teorema importante para la convergencia débil. Tal teorema da condiciones bajo las cuales basta la convergencia de las sucesiones finito dimensionales, para que la sucesión converja debilmente. Antes definiremos algunos conceptos que nos servirán de ayuda.

☞ *Definición 6.1*

Sea Π una familia de medidas de probabilidad sobre un espacio $(S, \mathcal{S})$. Decimos que la familia Π es relativamente campacta (compacta por sucesiones) si para cualquier sucesión existe una subsucesión y existe una medida de probabilidad Q tal que la subsucesión converge debilmente a Q.

Cabe notar que la medida Q de la definición no necesariamente pertenece a la familia Π.

Observación 6.2

Sea P_n una sucesión de medidas de probabilidad sobre $(C, \mathcal{C})$ tales que sus distribuciones finito dimensionales convergen debilmente a las distribuciones finito-dimensionales de P. Supóngase además que $\{P_n\}$ es relativamente compacto. Entonces $P_n \Rightarrow P$.

Demostración

Sabemos que cada subsucesión $\{P_{n'}\}$ de $\{P_n\}$ contiene una subsucesión $\{P_{n''}\}$ que converge debilmente a un límite Q. En particular $P_{n''}\pi^{-1}_{t_1\cdots t_n} \Rightarrow Q\pi^{-1}_{t_1\cdots t_n}$, y dado que $P_{n''}\pi^{-1}_{t_1\cdots t_n} \Rightarrow P\pi^{-1}_{t_1\cdots t_n}$, se tiene que $P\pi^{-1}_{t_1\cdots t_n} = Q\pi^{-1}_{t_1\cdots t_n}$.

Dado que las distribuciones finito dimensionales determinan de manera única a la medida, se tiene que tal medida es única. Por tanto cada sucesión $\{P_{n'}\}$ de $\{P_n\}$ contiene una subsucesión $\{P_{n''}\}$ que converge debilmente a P; por el teorema 2.14 $P_n \Rightarrow P$.

Teorema 6.3

Supongamos que P_n es relativamente compacto, supongamos también que sus distribuciones finito dimensionales convergen debilmente a alguna medida de probabilidad $\mu_{t_1\cdots t_n}$ sobre $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. Entonces $P_n \Rightarrow P$ para algún P.

Demostración

Sea $\{P_{n'}\}$ una subsucesón de $\{P_n\}$, entonces $\{P_{n'}\}$ contiene una subsucesión $\{P_{n''}\}$ que converge debilmente a un límite Q. En particular las distribuciones finito dimensionales de Q son las $\mu_{t_1\cdots t_n}$.

Sea $\{P_{m_k}\}$ una subsucesión de $\{P_n\}$, entonces $\exists\{P_{m'_k}\}$ subsucesión de $\{P_{m_k}\}$ tal que

$$P_{m'_k}\pi^{-1}_{t_1\cdots t_n} \Rightarrow \mu_{t_1\cdots t_n} = Q\pi^{-1}_{t_1\cdots t_n}.$$

Entonces $P_{m'_k} \Rightarrow Q$. Lo cual prueba el teorema.

Teorema 6.4

Sea Π una familia de medidas de probabilidades sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Π es relativamente compacto $\iff$

$$(*) \qquad \forall\epsilon > 0 \ \ \exists\, a, b \in \mathbb{R}, \ P(a, b] > 1 - \epsilon \qquad \forall P \in \Pi.$$

Demostración

$\Leftarrow$) Sea $\{P_n\}$ una sucesión de Π para cada $n \in \mathbb{N}$ y F_n la función de distribución de P_n $(F_n(x) = P_n\{y \le x\} = P_n[-\infty, x))$.

El teorema de Helly † asegura que existe una subsucesión $\{F_{n'}\}$ de $\{F_n\}$ y $\exists\, F$ tal que si $C_F = \{x |\ F \text{ es continua en } x\}$ entonces

$$(**) \qquad F_{n'}(x) \to F(x) \quad \forall x \in C_F$$

(además, F se puede elegir continua por la derecha).

Sea μ una medida asociada a F en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$$\mu(a, b] = F(b) - F(a).$$

Si $\mu(\mathbb{R}) = 1$ entonces μ es una medida de probabilidad sobre $\mathbb{R}$ y $(**)$ implica que $P_{n'} \Rightarrow \mu$ y entonces Π es tensa. Así que basta probar que $\mu(\mathbb{R}) = 1$.

Por hipótesis $\forall\epsilon > 0, \exists\, a, b \in \mathbb{R}, \ P(a, b] > 1 - \epsilon \quad \forall P \in \Pi$, ya que eligiendo F creciente, C_F es a lo mas numerable, y por tanto podemos tomar para cada $\epsilon > 0$, $a' < a$, $b < b'$ con a, b puntos de continuidad de F tal que se satisface $(**)$. Tomemos $\epsilon = \frac{1}{m}$ por tanto $\exists\, a_m, b_m \in \mathbb{R}$ tales que

$$P(a_m, b_m] > 1 - \epsilon \qquad \forall P \in \Pi, \ a_m, b_m \in C_F$$

además podemos elegir $a_{m+1} \le a_m < b_m \le b_{m+1} \quad \forall m$.

---

† Ver Notas del curso M.E.Caballero y B.Fernandez, Sem.Prob.

Entonces

$$\mu(a_m, b_m] = F(b_m) - F(a_m) = \lim_{n\to\infty} P_{n'}(a_m, b_m] \geq \epsilon$$

Por tanto $\mu(\mathbb{R}) \geq \mu(a', b'] \geq 1 - \frac{1}{m}$ es decir :

$$\mu(\mathbb{R}) \geq 1 - \frac{1}{m} \quad \forall m$$

por tanto $\mu(\mathbb{R}) \geq 1$. Dado que $\mu(\mathbb{R}) > 1$, es imposible tenemos que $\mu(\mathbb{R}) = 1$.

$\Rightarrow$) Supongamos que es falsa la conclusión. Entonces

$$\exists\ \epsilon > 0\ \forall a, b \in \mathbb{R}\ \ \exists\ P \in \Pi P(a, b] \leq 1 - \epsilon.$$

Para

$$a = -n,\ b = n\ \ \exists\ P_n \in \Pi,\ P_n(a, b] \leq 1 - \epsilon.$$

Dado que Π es relativamente compacto $\exists$ $\{P_{n'}\}$ subsucesión de $\{P_n\}$ tal que $\{P_{n'}\} \to Q$ para alguna Q medida de probabilidad sobre $\mathbb{R}$. Si $x \in \mathbb{R}$ entonces

$$Q(-x, x) \leq \liminf P_{n'}(-x, x)$$

esto, debido al teorema de Pormanteau y que $(-x, x)$ es abierto.

Dado que

$$\liminf_{n'} P_{n'}(-x, x) \leq \liminf_{n'} P_{n'}(-n', n') \leq 1 - \epsilon\ .$$

Entonces $Q(-x, x) \leq 1 - \epsilon$. Por tanto

$$1 = Q(\mathbb{R}) = \lim_{x\to\infty} Q(-x, x) \leq 1 - \epsilon.$$

Lo cual es un absurdo. Y se tiene el teorema.

Ejemplo.

Sea

$$P_n = \begin{cases} 1 & \text{si } n \in A \\ 0 & \text{si no}\ . \end{cases}$$

La distribución de P_n es la siguiente :

$$F_n(x) = \begin{cases} 1 & \text{si } n \leq x \\ 0 & \text{si } n > x \end{cases}\ .$$

Dada cualquier subsucesión $F_{n'}$ de F_n y dado x se tiene

$$F_{n'}(x) \to 0,$$

por tanto la medida asociada a F es $\mu \equiv 0$. Además

$$\mu(\mathbb{R}) = \lim_{n\to\infty}(F(x) - F(-x)) = 0$$

Por tanto P_n no satisface el teorema 6.4 y esto se debe a que la familia no satisface la necesidad del teorema, pues $\exists\ \epsilon > 0\ \left(\epsilon = \frac{1}{2}\right)$, tal que $\forall a, b \in \mathbb{R}\ \exists\ P \in \Pi\quad (P = P_m),\ P(a,b] \leq 1 - \epsilon$, con $m = \max\{|a|, |b|\} + 1$.

Corolario 6.5

Π es relativamente compacto si y solo si :

(6.5) $$\forall \epsilon > 0\ \ \exists\ K, \text{ compacto, tal que } P(K) \geq 1 - \epsilon \quad \forall P \in \Pi.$$

Demostración

$\Rightarrow$) Por el teorema (6.4) $\exists\ a, b \in \mathbb{R},\ P(a,b] \geq 1 - \epsilon \quad \forall P \in \Pi$. Sea $K = [a,b]$, entonces $P(K) > 1 - \epsilon \quad \forall P \in \Pi$ con K compacto.

$\Leftarrow$) Sea $\epsilon > 0$, sabemos que $\exists\ K$ compacto tal que

(6.6) $$P(K) \geq 1 - \epsilon \quad \forall P \in \Pi.$$

Dado que $\exists\ M > 0$ tal que $|x| < M,\ \forall x \in K$ tomando $a = -M, b = M,\ P \in \Pi$ se tiene que $P(a,b] \geq P(K) > 1 - \epsilon$.

los resultado anteriores, inducen una definición la cual nos permite generalizar estos resultados.

Definición 6.6

Una familia Π de medidas de probabilidad sobre un espacio métrico $(S, \mathcal{S})$, se dice que Π ***es tensa*** si

$$\forall \epsilon > 0\ \ \exists\ K, P(K) \geq 1 - \epsilon \quad \forall P \in \Pi.$$

con K compacto. Así el corolario anterior, podemos reformularlo de la siguiente forma

Corolario 6.7

Π es relativamente compacto si y solo si es tensa.

Este resultado admite una generalización a espacios métricos, arbitrarios :

Teorema 6.8 PROHOROV

Sea Π una familia de medidas de probabilidad sobre un espacio métrico $(S, \mathcal{S})$.

(a) Si Π es tensa entonces Π es relativamente compacto.

(b) Si S es un espacio métrico completo y separable y si Π es relativamente compacto, entonces Π es tensa.

Demostración (a) Se hará por casos.

CASO 1 $(S = \mathbb{R}^k)$. Consideremos aquí solamente $k = 2$, para k arbitraria la demostración será análoga.

Sea $\{P_n\}$ una sucesión de elementos de Π. Por el teorema de Helly $\exists\, F_{n'}$, $\exists\, F$ tal que

$$F_{n'}(x) \to F(x) \qquad \forall x \in C_F.$$

Sea μ la medida sociada a F $(\mu \colon \mathcal{B}(\mathbb{R}^2) \to [\,0,1\,])$; Nuevamente es claro que si $\mu(\mathbb{R}^2) = 1$ entonces $P_{n'} \Rightarrow \mu$. Basta probar entonces que $\mu(\mathbb{R}^2) = 1$.

Sea $\epsilon > 0$. Como Π es tensa $\exists\, K \subseteq \mathbb{R}^2$, K compacto tal que

$$P_{n'}(K) > 1 - \epsilon \quad \forall n'.$$

Sea $(a, b] = \{(x, y) \in \mathbb{R}^2 \mid a_1 < x \leq a_2, b_1 < y \leq b_2\}$ una caja que contenga a K, con $a = (a_1, b_1), b = (a_2, b_2)$.

A lo mas hay una cantidad numerable de rectas en el plano que tengan μ-medida positiva. Por tanto podemos considerar que los lados de la caja tienen μ-medida cero. Asíque cada vértice de la caja es un punto de continuidad de F. Entonces

$$\begin{aligned}
P_{n'}(a,b] &= F_{n'}((a_2,b_2)) - F_{n'}((a_2,b_1)) - F_{n'}((a_1,b_2)) + F_{n'}((a_1,b_1)) \\
&\to F((a_2,b_2)) - F((a_2,b_1)) - F((a_1,b_2)) + F((a_1,b_1)) \\
&= \mu(a,b] \\
&\geq 1 - \epsilon
\end{aligned}$$

figurita uno

Por tanto $\mu(a,b]) \geq 1-\epsilon$ y entonces $\mu(\mathbb{R}^2) \geq 1-\epsilon$, lo cual prueba el resultado.

CASO 2 $(S = \mathbb{R}^\infty)$. Para este caso emplearemos un lema con el cual la demostración es mas sencilla.

Lema (*)

Si Π es una familia tensa sobre $(S, \mathcal{S})$ y si h es una función contínua de S en S' entonces : $\{Ph^{-1} \mid P \in \Pi\}$ es una familia tensa sobre $(S', \mathcal{S}')$.

Demostración (lema (*))

Sea $\epsilon > 0$, elijamos un compacto K en S tal que $P(K) > 1 - \epsilon \quad \forall P \in \Pi$.

Si $K' = h(K)$ entonces K es compacto, ya que las funciones continuas mandan compactos en compactos. Además $h^{-1}(K') \supseteq K$ por tanto

$$P(h^{-1}(K')) \geq P(K) > 1 - \epsilon \quad \forall P \in \Pi.$$

Sea pues Π una familia de probabilidades tensa en $\mathbb{R}^\infty$, por el lema (*) la familia $\{P\pi_k^{-1}\}$ para cada k es tensa sobre $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$.

Por el caso anterior dada una sucesión $\{P_n\}$ de Π podemos elegir para $k = 1$ una subsucesión $\{P_{n^1}\}$ de $\{P_n\}$ tal que $P_{n^1}\pi_1^{-1} \Rightarrow \mu_1\pi_1^{-1}$, para $k = 2$ podemos elegir una subsucesión $\{P_{n^2}\}$ de $\{P_{n^1}\}$ tal que $P_{n^2}\pi_2^{-1} \Rightarrow \mu_1\pi_2^{-1}$ con este procedimiento vamos hallando una sucesión de sucesiones que, con el clásico método de tomar la diagonal, podemos elegir una sucesion $\{P_{n'}\}$ tal que

$$P_{n'}\pi_k^{-1} \Rightarrow \mu_1\pi_k^{-1} \;\forall k,$$

Dado que tales medidas satisfacen las condiciones de consistencia del teorema de existencia de Kolmogorov † se tiene que $\exists\, Q$ medida de probabilidad sobre $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ tal que

$$Q\pi_k^{-1} = \mu\pi_k^{-1} \quad \forall k\ .$$

Por tanto

$$P_{n'}\pi_k^{-1} \Rightarrow Q\pi_k^{-1} \quad \forall k.$$

En $\mathbb{R}^\infty$ esto implica que $P_{n'} \Rightarrow Q$. Y por tanto $\mathbb{R}^\infty$ es relativamente compacto.

Antes de proseguir con los dos casos restantes veamos algunas observaciones y algunos lemas.

Sea $S_0 \subseteq S$ tal que $S_0 \in \mathcal{S}$ consideremos ahora la sigma-álgebra $\mathcal{S}_0$ restringida a S_0. Notemos que

$$\mathcal{S}_0 = \{A \subseteq S_0 \mid A \in \mathcal{S}\}.\quad ††$$

Asíque $\mathcal{S}_0 \subseteq \mathcal{S}$. Consideremos ahora una medida P de probabilidad sobre $(S, \mathcal{S})$, definamos la restricción P^r de P a $(S_0, \mathcal{S}_0)$ como:

$$P^r(A) = P(A),\ \text{con } A \in \mathcal{S}_0$$

---

† ver Notas de M. E. Caballero y B. Fernandez

†† ver apéndice

Análogamente podemos considerar una medida de probabilidad sobre $(S_0, \mathcal{S}_0)$ y podemos definir la extensión P^e de P a S como:

$$P^e(A) = P(A \cap S_0), \text{ con } A \in \mathcal{S}$$

Notemos que, de la última definición que $P^e(S_0) = 1$.

Una observación importante es que si P es una medida de probabilidad sobre $(S, \mathcal{S})$ con $P(S_0) = 1$ entonces $P^r(A) = P(A)$, por tanto P^r es una medida de probabilidad sobre $(S_0, \mathcal{S}_0)$ y entonces podemos definir $(P^r)^e$ sobre $(S, \mathcal{S})$. Se tiene pues que :

$$(P^r)^e(A) = P^r(A \cap S_0) = P(A \cap S_0) = P(A)$$

es decir $(P^r)^e = P$.

Análogamente $(P^e)^r = P$ siempre que P sea una medida de probabilidad sobre $(S_0, \mathcal{S}_0)$.

Lema (**)

Si Π es una familia de medidas de probabilidad tensa sobre $(S_0, \mathcal{S}_0)$ entonces $\Pi^e = \{P^e \mid P \in \Pi\}$ es una familia tensa sobre $(S, \mathcal{S})$. Además si $P_n \Rightarrow P$ en $(S_0, \mathcal{S}_0)$ entonces $P_n^e \Rightarrow P^e$ en $(S, \mathcal{S})$.

Demostración (lema (**))

Sea

$$h: S_0 \to S, \quad h(s) = s, \qquad \forall s \in S_0.$$

Para h es continua, con $D_h = \{\text{discontinuidades de h}\}$ se tine $P(D_h) = 0$ y nótese también que si $A \subseteq S$ entonces $h^{-1}(A) = S_0 \cap A$. Por el lema (*)

$$\{Ph^{-1} \mid P \in \Pi\}$$

es una familia tensa sobre $(S, \mathcal{S})$, pero $Ph^{-1}(A) = P(S_0 \cap A) = P^e(A)$ por tanto

$$\Pi^e = \{P^e \mid P \in \Pi\} = \{Ph^{-1} \mid P \in \Pi\}$$

es tensa. Si $P_n \Rightarrow P$, por el teorema 5.1, $P_n h^{-1} \Rightarrow Ph^{-1}$, es decir $P^e \Rightarrow P^e$ y se tiene el resultado.

Lema (***)

Si $P_n \Rightarrow P$ en $(S, \mathcal{S})$ y si $P_n(S_0) = 1 = P(S_0)$ $\forall n$ entonces $P_n^r \Rightarrow P^r$ en $(S_0, \mathcal{S}_0)$.

Demostración (lema (***))

Sea G_0 abierto en S_0 entonces $G_0 = S_0 \cap G$ para algún G abierto en S dado que $P_n \Rightarrow P$, por el teorema de Pormanteau se tiene que

$$\liminf P_n(G) \geq P(G)$$

es decir

$$\liminf P_n^r(G_0) \geq P^r(G_0).$$

Nótese que

$$P_n^r(G_0) = P_n(G_0) = P_n(G \cap S_0) = P_n(G)$$

y que

$$P^r(G_0) = P(G_0) = P(G \cap S_0) = P(G).$$

CASO 3 (σ-compacto)

Sea S σ-compacto, entonces S es separable †,por tanto :

$$\exists\, \phi : S \to \mathbb{R}^\infty$$

tal que ϕ es un inyectiva y bicontinua. Ahora $\phi(S) \in \mathcal{B}(\mathbb{R}^\infty)$ por lo tanto S es homeomorfo †† a $\phi(S)$.

Esto debido a que

$$S = \bigcup_{i=1}^{\infty} K_i$$

con K_i compacto, por tanto $\phi(K_i)$ es compacto, y entonces

$$\phi(S) = \bigcup_{i=1}^{\infty} \phi(K_i)$$

por lo tanto $\phi(S)$ es σ-compacto.

Sea

$$S^* = \phi(S) = \phi(\bigcup_{i=1}^{\infty} K_i) = \bigcup_{i=1}^{\infty} K_i$$

con $\hat{K}_i = \phi(K_i)$.

S^* es σ-compacto, y $S^* \subseteq \mathbb{R}^\infty$, además $S^* \in \mathcal{B}(\mathbb{R}^\infty)$.

Sea Π una familia de medidas de probabilidad sobre S tensa;

Por el lema (**), $\{P\phi^{-1} \mid P \in \Pi\} = \Pi^*$ es tensa sobre S^*;

Por el lema (***), $\Pi^{*e} = \{(P\phi^{-1})^e \mid P \in \Pi\}$ es tensa sobre $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$.

Por el caso anterior Π^{*e} es relativamente compacto.

Sea $\{P_n\phi^{-1}\}$ una sucesión en Π^* entonces la sucesión correspondiente $\{(P_n\phi^{-1})^e\}$ contiene una subsucesión $\{(P_{n'}\phi^{-1})^e\}$ que converge debilmente en $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ a algún Q.

Como Π^* es tensa sobre $(S^*, \mathcal{S}^*)$, para $\epsilon > 0$ $\exists\, K \subseteq S^*$ tal que

$$(P\phi^{-1})(K) > 1 - \epsilon \quad \forall P \in \Pi$$

† ver Int. a la Top. A. G. Maynez

por tanto

$$(P_{n'}\phi^{-1})^e(K) = (P_{n'}\phi^{-1})(K) > 1 - \epsilon \quad \forall n'.$$

Por lo tanto

$$Q(S^*) \geq Q(K) \geq \limsup (P_{n'}\phi^{-1})^e(K) \geq 1 - \epsilon$$

por lo tanto, $Q(S^*) = 1$ y $P_{n'}\phi^{-1}(S^*) = 1$. Por el lema $(^{***})$, se tiene que

$$((P_{n'}\phi^{-1})^e)^r \Rightarrow Q^r$$

en $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$, es decir

$$P_{n'}\phi^{-1} \Rightarrow Q^r$$

en $(\mathbb{R}^\infty, \mathcal{B}(\mathbb{R}^\infty))$ por tanto Π^* es relativamente compacto.

Dada $\{P_n\}$ sucesión en Π, consideramos $\phi^{-1} : S^* \to S$, por el lema $(^{**})$, $\{P_n\phi^{-1}\}$ contiene una subsucesión que converge debilmente a Q, por tanto $P_{n'}\phi^{-1} \Rightarrow Q$, como ϕ es continua, entonces

$$\left(P_{n'}\phi^{-1}\right)\phi \Rightarrow Q\phi,$$

es decir

$$P_{n'} \Rightarrow Q\phi.$$

CASO (4) (Caso general).

Sea S un espacio métrico. Π una familia tensa sobre $(S, \mathcal{S})$.

Para $\epsilon_i = 1/i > 0$, $\exists\, K_i$ compacto con $P(K_i) > 1 - 1/i \quad \forall P \in \Pi$. Sea

$$S_0 = \bigcup_{i=1}^{\infty} K_i$$

entonces $P(S_0) = 1$, $\forall P \in \Pi$ por tanto $\Pi^r = \{P^r \mid P \in \Pi\}$ es una familia tensa sobre $(S_0, \mathcal{S}_0)$, pues para $\epsilon > 0$ $\exists\, K_i$ compacto en S_0 $(K_i \subseteq S_0)$ tal que

$$P(K_i) > 1 - 1/i > 1 - \epsilon \quad \forall P \in \Pi$$

pero $P(K_i) = P^e(K_i)$, por tanto Π^r es relativamente compacto. Sea $\{P_n\}$ un subsucesión en Π, entonces $\{P_n^r\}$ tiene una subsucesión convergente $\{P_{n'}^r\}$ a Q, por el lema $(^{***})$

$$(P_{n'}^r)^e \Rightarrow Q^e$$

es decir $P_{n'} \Rightarrow Q$, por tanto Π es relativamente compacto, lo cual prueba el teorema.

Ahora podemos analizar el recíproco, para lo cual emplearemos un lema que facilitará los cálculos.

Lema (#)

Si $\forall \epsilon > 0, \forall \delta > 0, \quad \exists\ A_1, A_2, \ldots, A_n$ δ -esferas tales que

$$P\left(\bigcup_{i=1}^{k(i)} A_i\right) > 1 - \epsilon \quad \forall P \in \Pi$$

entonces Π es tensa.

Demostración (lema (#))

Sea $\epsilon > 0$ consideremos para cada $\delta_k = 1/k > 0$ $A_1, A_2, \ldots, A_{n_k} 1/k$ -esferas con $P(\left(\bigcup_{i=1}^{n_k} A_i\right) > 1 - 1/k$.

De la demostración del teorema 1.10 se tiene que si K es la cerradura de el conjunto totalmente acotado $\bigcap_{k\geq 1} \bigcup_{i\leq n_k} A_i$, entonces $P(K) > 1 - \epsilon$, y ya que S es completo, K es compacto. Por tanto Π es tensa.

Usando este hecho probaremos (b) por contradicción.

Demostración (de (b)).

Supongamos que Π no es tensa, y Π es relativamente compacto. Por el lema (#) $\exists\ \epsilon > 0, \quad \exists\ \delta > 0,$ tales que $\forall A_1, A_2, \ldots, A_{n_k} \delta$ -esferas, se tiene que

$$P\left(\bigcup_{i=1}^{n_k} A_i\right) \leq 1 - \epsilon$$

para algún $P \in \Pi$.

Dado que S es separable, por el mismo argumento del teorema (1.10)

$$S = \bigcup_{i=1}^{\infty} A_i$$

con A_i δ -esferas. Sea $B_n = \bigcup_{i=1}^{n} A_i$, ya que $\exists\ P_n \in \Pi$ con $P_n(B_n) \leq 1 - \epsilon$ y Π es relativamente compacto, $\exists\ \{P_{n'}\}$ subsucesión de $\{P_n\}$ tal que

$$P_{n'} \Rightarrow Q$$

para algún Q, ahora B_n es abierto para toda n, por lo tanto

$$P(B_m) \leq \liminf P_{n'}(B_m) \leq \liminf P_{n'}(B_{n'}) \leq 1 - \epsilon$$

Pues $B_m \subseteq B_{n'}$ para n' suficientemente grande. Así que $P(B_m) \leq 1 - \epsilon$, es decir

$$1 = P(S) = \lim_{m\to\infty} P(B_m) \leq 1 - \epsilon.$$

Lo cual es un absurdo, entonces Π es tensa.

(ii) $\forall\epsilon > 0$, $\forall\eta > 0$, $\exists\,\delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que:

$$P_n\{x \in C \mid w_x(\delta) \geq \epsilon\} \leq \eta, \qquad n \geq N_0 \tag{II}$$

Observemos que el hecho de que $\{P_n\}$ satisfaga (I), significa que dado cualquier número positivo, podemos elegir a suficientemente grande, tal que casi todas las funciones continuas, valuadas en cero, quedan por debajo de a, en el sentido de que al conjunto de las otras funciones (las que quedan por encima de a), cualquier P_n le asocia medida menor que el número positivo dado.

Y el hecho de que $\{P_n\}$ satisfaga (II), significa que dado cualquier número positivo ϵ se puede elegir δ suficientemente pequeño, tal que el conjunto de funciones con δ-módulo de continuidad mayor que ϵ, tienen probabilidad P_n muy pequeña, a partir de cierto rango. De ahí que la demostración de este teorema tenga que ver con el teorema de Arzela-Ascoli.

Demostración

$\Rightarrow$) Supongamos que $\{P_n\}$ es tensa. Dados $\epsilon, \delta > 0$ elegimos un compacto K tal que $P_n(K) > 1 - \eta\ \forall n$. Por el teorema de Arzelà-Azcoli * sabemos que

$$K \subseteq \{x \in C \mid |x(0)| \leq a\}$$

para algún a suficientemente grande y además

$$K \subseteq \{x \in C \mid w_x(\delta) < \epsilon\}$$

para δ suficientemente pequeño. Por tanto :

$$\begin{aligned} P_n\left(C - \{x \in C \mid |x(0)| \leq a\}\right) &= P_n\{x \in C \mid |x(0)| > a\} \\ &\leq P_n(C - K) < \eta, \end{aligned}$$

y también

$$\begin{aligned} P_n\left(C - \{x \in C \mid w_x(\delta) < \epsilon\}\right) &= P_n\{x \in C \mid w_x(\delta) \geq \epsilon\} \\ &\leq P_n(C - K) < \eta \end{aligned}$$

y dado que esto se satisface para toda n, tenemos entonces probada la primera parte del teorema.

$\Leftarrow$) Supongamos ahora que se satisface (i), (ii).

Sean $\epsilon, \eta > 0$ sabemos por (ii) que existe una $\delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que se satisface (II). Para cada $k \leq N_0$, $\{P\}$ es tensa, y aplicando la primera parte de la demostración se tiene que $\exists\, \delta_k$ que satisface las propiedades antes mencionadas.

Sea $\hat{\delta} = \min\{\delta_1, \delta_2, \ldots, \delta_{N_0}, \delta\}$ dado que $w_x(\cdot)$ es decreciente entonces :

$$\{x \in C \mid w_x(\hat{\delta}) \geq \epsilon\} \subseteq \{x \in C \mid w_x(\delta_k) \geq \epsilon\} \quad \forall k \leq N_0$$

---

* ver Dieudonné, Int. al Análisis.

y

$$\{x \in C \mid w_x(\hat{\delta}) \geq \epsilon\} \subseteq \{x \in C \mid w_x(\delta) \geq \epsilon\} .$$

Entonces (II) se satisface para toda n. (es decir podemos suponer, para fines de probar el teorema, que (II) se satisface para toda n.)

Sea $\eta > 0$ para $\frac{\eta}{2}$ por (i) $\exists a$ tal que

$$P_n(B) = \{x \in C \mid |x(0)| > a\} \leq \frac{\eta}{2} \quad \forall n$$

Por tanto $P_n(C - B = A) > 1 - \frac{\eta}{2} \quad \forall n.$

Elijamos ahora $\delta_k > 0$ tal que si

$$A_k = \{x \in C \mid w_x(\delta_k) < \frac{1}{k}\}$$

entonces

$$P_n(A_k) \geq 1 - \frac{\eta}{2^{k+1}}, \qquad \forall n.$$

Sea K la cerradura del conjunto $A \bigcap (\bigcup_{k=1}^{\infty} A_k)$, entonces

$$\begin{aligned}
P_n(C - K) &\leq P_n\left((C - A) \bigcap (\bigcup_{k=1}^{\infty} A_k\right) \\
&\leq P_n\left((C - A) \bigcup (\bigcup_{k=1}^{\infty} (C - A_k)\right) \\
&\leq P_n(C - A) + P\left(\bigcup_{k=1}^{\infty} (C - A_k)\right) \\
&\leq P_n(C - A) + \sum_{k=1}^{\infty} P(C - A_k) \\
&\leq \frac{\eta}{2} + \sum_{k=1}^{\infty} \frac{\eta}{2^{k+1}} \\
&= \sum_{k=1}^{\infty} \frac{\eta}{2^k} = \eta \ .
\end{aligned}$$

Por tanto $P_n(K) \geq 1 - \eta \cdot \ \forall n.$

Basta entonces probar que K es compacto para que $\{P_n\}$ sea tensa.

Probaremos que K satisface las hipótesis del teorema de Arzela-Ascoli y se tendrá entonces que K es compacto.

Sabemos que B satisface :

$$|x(0)| \leq a \qquad \forall x \in B, \tag{7.4}$$

y que si $x \in \bigcup_{k=1}^{\infty} A_k$ entonces $\quad \forall \alpha > 0, \ \exists \ \beta > 0$ tal que

$$|x(s) - x(t)| \leq w_x(\delta) < \alpha \qquad \text{si } |s - t| < \beta \tag{7.5}$$

Nótese que la condición 7.5 implica que B es equicontinuo.

Por otra parte dada $t \in [0,1]$, $x \in B$, $\exists\, k \in \mathbb{N}$ tal que $\frac{1}{k} < t$. Ya que $x \in A_k$, entonces $\exists\, \delta_k > 0$ tal que

$$|x(s) - x(u)| \leq w_x(\delta_k) \leq \frac{1}{k} \quad \text{si } |s-u| < \delta_k \ .$$

Sabemos que $\exists\, N \in \mathbb{N}$ tal que $\frac{N\delta_k}{2} \leq t < \frac{(N+1)\delta_k}{2}$. Ahora

$$|x(t) - x(0)| \leq \sum_{i=0}^{N} |x(\frac{(i+1)\delta_k}{2}) - x(\frac{i\delta_k}{2})| + |x(t) - x(\frac{N\delta_k}{2})|$$

Por tanto

$$|x(t)| \leq \underbrace{\frac{1}{k} + \frac{1}{k} + \cdots + \frac{1}{k}}_{(N+1)-\text{veces}} + |x(0)|$$

es decir

$$|x(t)| \leq |x(0)| + \frac{(N+1)}{k}$$

Por tanto $\{x(t)\}$ está acotada, y dado que en $\mathbb{R}$ basta acotado para ser relativamente compacto, se tienen las hipótesis del teorema de Arzela-Ascoli.

Veamos ahora una modificación de la suficiencia del teorema 7.2

Teorema 7.4

Una sucesión $\{P_n\}$ de medidas de probabilidad sobre $(C, \mathcal{C})$ es tensa si las siguientes 2 condiciones se satisfacen:

(i) $\forall \eta > 0$, $\exists\, a \in \mathbb{R}$ tal que:

$$P_n\{x \in C \mid |x(0)| > a\} \leq \eta, \qquad n \geq 1 \tag{III}$$

(ii) $\forall \epsilon > 0$, $\forall \eta > 0$, $\exists\, \delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que:

$$\frac{1}{\delta} P_n\{x \in C \mid \sup_{t \leq s \leq t+\delta} |x(s) - x(t)| \geq \epsilon\} \leq \eta, \qquad n \geq N_0 \tag{IV}$$

Nótese que para cada $t \in [0,1]$ se toma el supremo sobre los s para los cuales tiene sentido $x(s)$ es decir los s que satisfagan $t \leq s \leq t+\delta$ y $s \in [0,1]$.

Demostración

Basta probar que la condición (ii) del teorema implica la condición (ii) del teorema 7.2, dado que la condición (i) es la misma. Sea $\epsilon > 0, \eta > 0$

Para $\frac{\epsilon}{4} > 0, \frac{\eta}{2} > 0$ por (ii) $\exists \delta, 0 < \delta < 1$, $\exists\, N_0$, tal que 7.4 se satisface con $\frac{\epsilon}{4}, \frac{\eta}{2}$. Sea

$$A_t = \{x \in C \mid \sup_{t \leq s \leq t+\delta} |x(s) - x(t)| \geq \epsilon\}.$$

Probemos primero que

$$\{x \in C \mid w_x(\delta) \geq \epsilon\} \subseteq \bigcup_{i<\frac{1}{\delta}} A_{i\delta}.$$

Sea x tal que $\sup_{|t-s|<\delta} |x(s) - x(t)| \geq \epsilon$ y supongamos que $x \notin A_{i\delta}\ \forall i < \frac{1}{\delta}$ es decir

$$\sup_{i\delta \leq s \leq (i+1)\delta} |x(s) - x(i\delta)| < \epsilon \quad \forall i < \frac{1}{\delta}$$

Sean $t, s \in [0,1]$, $|t-s| < \delta$, supongamos que $s \leq t$ entonces $\exists\, i$ tal que

$$i\delta \leq s \leq (i+1)\delta$$

como en la figurita uno.

figurita uno

Puede suceder que $t \leq (i+1)\delta$, o que $t \leq (i+2)\delta$ pero no puede pasar que $t > (i+2)\delta$ dado que $|t-s| < \delta$, en el primer caso se tiene

$$|x(s) - x(t)| \leq |x(s) - x(i\delta)| + |x(t) - x(i\delta)| \leq \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{2}{4\epsilon}$$

en el segundo caso

$$\begin{aligned} |x(s) - x(t)| \leq{} & |x(s) - x(i\delta)| \\ & + |x(i\delta) - x((i+1)\delta)| + |x(t) - x((i+1)\delta)| \\ < {} & \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{3}{4\epsilon} \end{aligned}$$

En ambos casos $|x(s) - x(t)| < \frac{3}{4\epsilon}$ por tanto

$$\sup_{|t-s|<\delta} |x(s) - x(t)| \leq \frac{3}{4\epsilon} < \epsilon$$

es decir $w_x(\delta) < \epsilon$, lo cual es un absurdo.

De (*) se sigue que $P_n\{x \in C \mid w_x(\delta) \geq \epsilon\} \leq P_n\left(\bigcup_{i<\frac{1}{\delta}} A_{i\delta}\right)$ aplicando ahora la hipótesis de que $P_n(A_{i\delta}) \leq \frac{\delta\eta}{2}$ tenemos que

$$P_n\{x \in C \mid w_x(\delta) \geq \epsilon\} \leq \sum_{i<\frac{1}{\delta}} P_n(A_{i\delta}) \leq (\delta\frac{\eta}{2})([\frac{1}{\delta}] + 1) < \eta$$

lo cual es debido a que $(\delta\frac{\eta}{2})([\frac{1}{\delta}] + 1) < \frac{\eta}{2}(\delta[\frac{1}{\delta}] + \delta) < \frac{\eta}{2}(1+1) = \eta$, por tanto se tiene lo que se quería.

Corolario 7.5

Si $0 = t_0 < t_1 < \cdots < t_n = 1$ y si $t_i - t_{i-1} \geq \delta, \quad 2 \leq i \leq r-1$ entonces :

$$P\{x \in C |\ w_x(\delta) \geq \epsilon\} \leq \sum_{i=1}^{r} P\{x \in C |\ \sup_{t_{i-1} \leq s \leq t_i} |x(s) - x(t_{i-1})| \geq \epsilon\} .$$

Demostración

Está contenida en la demostración de 7.4.

8. CONVERGENCIA DEBIL DE PROCESOS SOBRE C[0,1].

En este capítulo introduciremos algunos conceptos que son de gran importancia en la teoría general de procesos, los cuales utilizaremos durante los siguientes 4 capítulos. También se establecen los teoremas del capítulo anterior, relativos a tensión, solo que ahora con procesos, además se construye un proceso a través de una suma de variables aleatorias, el cual es muy utilizado en la teoria antes mencionada.

☞ *Definición 8.1*

Un proceso estocástico sobre C es una función X de un espacio $(\Omega, \mathcal{B}, P)$ a el espacio C la cual es medible.

Si consideramos $X' : \Omega \times [0,1] \to \mathbb{R}$ definida como

$$X'(\omega, t) = (X(\omega))(t)$$

resulta que :

(i) $X'(\cdot, t)$ es una función medible de Ω en $\mathbb{R}$.

(ii) $X'(\omega, \cdot)$ es continua de $[0,1]$ en $\mathbb{R}$ y se le llama la ω-trayectoria de X.

Y a veces llamaremos a el proceso X una función aleatoria, con valores en C.

✎ ***Proposición 8.2***

X es un proceso estocástico si y solo si cada función $X'(t, \cdot)$ es una variable aleatoria medible.

Demostración (proposición)

$\Rightarrow$) Supóngase que X es $\mathcal{C}$-medible, de $(\Omega, \mathcal{B})$ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Sea $t \in [0,1]$ fijo, y $X'(t, \cdot) : \Omega \to \mathbb{R}$ sea $A = (-\infty, a]$, $B = \{x \in C | x(t) \leq a\}$ entonces

$$\begin{aligned} X'(t, \cdot)^{-1}(A) &= \{\omega \in \Omega | \; X'(t, \omega) \leq a\} \\ &= \{\omega \in \Omega | \; X'(\omega) \in B\} \\ &= X'^{-1}(B) \end{aligned}$$

por tanto $X'^{-1}(B) = X'^{-1}(t, \cdot)(A)$ dado que B es cerrado, $B \in \mathcal{C}$ por tanto $X'^{-1}(B) \in \mathcal{B}$ y entonces

$$X'^{-1} : (\Omega, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R})) .$$

es medible.

$\Leftarrow$) Notemos primero que si $B = \{y \in C \mid \|y - x\| \leq \epsilon\}$ entonces

$$X^{-1}(B) = \bigcap_{r \in \mathbb{Q}} \left\{ \omega \in \Omega \mid X'(r, \omega) \in [x(r) - \epsilon, x(r) + \epsilon] \right\}$$

En efecto

$\subseteq$) si $\omega \in X^{-1}(B)$ entonces $\|X(\omega) - x\| \leq \epsilon$, por tanto $\forall t \in [0, 1]$

$$|X'(t, \omega) - x(t)| \leq \epsilon$$

es decir,

$$X'(t, \omega) \in [x(t) - \epsilon, x(t) + \epsilon],$$

en particular esto es válido para los racionales.

$\supseteq$) Sea

$$\omega \in \bigcap_{r \in \mathbb{Q}} \{\omega \in \Omega \mid X(r, \omega) \in [x(r) - \epsilon, x(r) + \epsilon]\}$$

y supongamos que $\|X(\omega) - x(r)\| > \epsilon$ entonces $\exists\, t \in [0, 1]$, $|X(t, \omega) - x(t)| > \epsilon$.

Para $\epsilon^* = \frac{|X(t,\omega) - x(t)| - \epsilon}{2} > 0$

$$\exists\, \delta_1,\ |u - t| < \delta_1 \Rightarrow |x(t) - x(u)| < \epsilon^* \quad \text{por ser } x(t) \text{ continua}$$

y

$$\exists\, \delta_2,\ |u - t| < \delta_2 \Rightarrow |X'(t, \omega) - X'(u, \omega)| < \epsilon^* \quad \text{por ser } X'(\cdot, \omega) \text{ continua.}$$

Sea $\delta = \min\{\delta_1, \delta_2\} > 0$, entonces $\exists\, r \in \mathbb{Q}, |r - t| < \delta$. Entonces

$$\begin{aligned}
|X'(r,\omega) - x(r)| &= |X'(r,\omega) - X'(t,\omega) + X'(t,\omega) - x(t) + x(t) - x(r)| \\
&\geq |X'(t,\omega) - x(t)| - |X'(r,\omega) - X'(t,\omega)| - |x(t) - x(r)| \\
&\geq |X'(t,\omega) - x(t)| - \epsilon^* - \epsilon^* \\
&= |X'(t,\omega) - x(t)| + \epsilon - |X'(t,\omega) - x(t)| = \epsilon
\end{aligned}$$

Ya que $\epsilon^* + \epsilon^* = |X'(t,\omega) - x(t)| + \epsilon$.

Como $X'(r, \cdot)$ es medible

$$\Big\{\omega \in \Omega \mid X'(r,\omega) \in [x(r) - \epsilon, x(r) + \epsilon]\Big\} \in \mathcal{B} .$$

Por tanto $X^{-1}(B) \in \mathcal{B}$.

Dado que C es separable, cualquier abierto de C es unión numerable de esferas cerradas, y por tanto X es $\mathcal{C}$-medible.

Cuando hablemos de la función aleatoria X, C-valuada, la identificaremos con su respectiva X', por tal motivo aunque esto sea un abuso de lenguaje escribiremos indistintamente X como un proceso, o como una variable aleatoria C-valuada.

Nótese que, dado que la función X es medible, X tiene una distribución P la cual es una medida de probabilidad sobre C.

☞ *Definición 8.3*

Se dice que una sucesión de procesos estocásticos en C *es tensa*, si la correspondiente sucesión de distribuciones lo es.

Así que el teorema 7.3 se puede formular de la siguiente manera :

✎ *Teorema 8.4*

Una sucesión $\{X_n\}$ de procesos estocásticos continuos es tensa si y solo si las siguientes 2 condiciones se satisfacen.

(i) $\forall \eta > 0, \ \exists\, a \in \mathbb{R}$ tal que:

$$\mathrm{P}\{\omega \in \Omega |\ |X_n(0,\omega)| > a\} \leq \eta, \qquad n \geq 1 \tag{V}$$

(ii) $\forall \epsilon > 0, \ \forall \eta > 0 \ \exists\, \delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que:

$$\mathrm{P}\{\omega \in \Omega |\ w(X_n(\omega), \delta) \geq \epsilon\} \leq \eta, \qquad n \geq N_0 \tag{VI}$$

Nótese que la condición (i) es equivalente a que la sucesión de variables aleatorias $X_n(0)$ sea tensa. De forma análoga podemos formular el teorema 7.4 :

✎ *Teorema 8.5*

Una sucesión $\{X_n\}$ de procesos estocásticos continuos es tensa si las siguientes 2 condiciones se satisfacen.

(i) la sucesión $X_n(0)$ es tensa.

(ii) $\forall \epsilon > 0, \ \forall \eta > 0 \ \exists\, \delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que:

$$\frac{1}{\delta}\mathrm{P}\left\{\sup_{t \leq s \leq t+\delta} |X_n(s) - X_n(t)| \geq \epsilon\right\} \leq \eta, \qquad n \geq N_0 \tag{VII}$$

Construyamos un proceso estocástico X de la siguiente forma :

Sean $\xi_1, \xi_2, \ldots,$ variables aleatorias definidas sobre un espacio de probabilidad $(\Omega, \mathcal{B}, P)$. Definamos

$$S_i = \xi_1 + \xi_2 + \cdots + \xi_i, \quad S_0 = 0,$$

y

$$X_n(\frac{i}{n}, \omega) = \frac{1}{\sigma\sqrt{n}} S_i(\omega).$$

Definamos a X lineal y continuamente en cada intervalo $[\frac{i}{n}, \frac{i+1}{n}]$, tenemos que :

$$\begin{aligned}
X_n(t,\omega) &= \frac{(\frac{i}{n}) - t}{\frac{1}{n}} X_n(\frac{i-1}{n}, \omega) + \frac{(t - \frac{i-1}{n}) - t}{\frac{1}{n}} X_n(\frac{i}{n}, \omega) \\
&= \frac{1}{\sigma\sqrt{n}} S_{i-1}(\omega) + n\left(t - \frac{i-1}{n}\right) \frac{1}{\sigma\sqrt{n}} \xi_i(\omega) \\
&= \frac{1}{\sigma\sqrt{n}} S_{[nt]}(\omega) + (nt - n[t]) \frac{1}{\sigma\sqrt{n}} \xi_{[nt]+1}(\omega) \qquad (8.6)
\end{aligned}$$

Veamos un ejemplo de X_3 (figurita dos).

$$\begin{aligned}
X_n(\frac{0}{3}, \omega) &= \frac{1}{\sigma\sqrt{3}} S_0(\omega) = 0 \\
X_n(\frac{1}{3}, \omega) &= \frac{1}{\sigma\sqrt{3}} \xi_1(\omega) \\
X_n(\frac{2}{3}, \omega) &= \frac{1}{\sigma\sqrt{3}} (\xi_1(\omega) + \xi_2(\omega)) \\
X_n(\frac{3}{3}, \omega) &= \frac{1}{\sigma\sqrt{3}} (\xi_1(\omega) + \xi_2(\omega) + \xi_3(\omega))
\end{aligned}$$

figurita dos

Como las ξ_i, S_i son variables aleatorias, se tiene que para cada t la función $X(t,\cdot)$ es medible, y por la proposición 8.2 X es un proceso estocástico sobre C.

Es claro que como $X_n(0) = 0$ entonces, es en este caso $\{X_n(0)\}$ es tensa.

Observemos que si $t = \frac{k}{n}$, $t + \delta = \frac{j}{n}$, con k, j enteros y

$$s \in [\frac{r}{n}, \frac{r+1}{n}] \quad \text{con} \quad k \le r \le j$$

entonces para cada ω, (ver figurita 3)

$$
\begin{aligned}
|X_n(s,\omega) - X_n(t,\omega)| &= |X_n(s,\omega) - X_n(\frac{k}{n},\omega)| \\
&\leq \max\left\{ |X_n(\frac{r}{n},\omega) - X_n(\frac{k}{n},\omega)|\ ,\ |X_n(\frac{r+1}{n},\omega) - X_n(\frac{k}{n},\omega)| \right\} \\
&\leq \max_{\{k\leq r\leq j\}} |X_n(\frac{r}{n},\omega) - X_n(\frac{k}{n},\omega)| \\
&= \max_{\{k\leq r\leq j\}} \left(\frac{1}{\sigma\sqrt{n}} |S_r(\omega) - S_k(\omega)| \right) \\
&= \max_{i\leq j-k} \left(\frac{1}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \right) \\
&= \max_{i\leq n\delta} \left(\frac{1}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \right) \qquad (\#)
\end{aligned}
$$

Para la primera desigualdad se considera

$$d\colon [\frac{r}{n}, \frac{r+1}{n}] \to \mathbb{R}, \quad d(u) = |X_n(t,\omega) - X_n(u,\omega)|, \quad d,$$

la cual es monótona.

figurita tres

Por tanto

$$\sup_{t\leq s\leq t+\delta} |X_n(s,\omega) - X_n(t,\omega)| \leq \max_{i\leq j-k} \left(\frac{1}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \right)$$

Y por tanto

$$\mathrm{P}\left\{ \sup_{t\leq s\leq t+\delta} |X_n(s,\omega) - X_n(t,\omega)| \geq \epsilon \right\} \leq \mathrm{P}\left\{ \max_{i\leq j-k} \left(\frac{1}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \right) \geq \epsilon \right\}$$

Entonces (VII) se transforma en

$$\frac{1}{\delta}\mathrm{P}\left\{ \max_{i\leq n\delta} \left(\frac{1}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \right) \geq \epsilon \right\} \leq \eta$$

Proposición 8.6

Supóngase que $\forall\epsilon > 0$, $\forall\eta > 0$ $\exists\,\delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que:

$$\frac{1}{\delta}\mathrm{P}\left\{\max_{i\leq n\delta}\left(\frac{1}{\sigma\sqrt{n}}|S_{k+i}(\omega) - S_k(\omega)|\right) \geq \epsilon\right\} \leq \eta \quad \forall n \geq N_0,\ \forall k. \qquad (VIII)$$

Entonces $\forall\epsilon > 0$, $\forall\eta > 0$ $\exists\,\delta > 0$ con $0 < \delta < 1$, y $\exists\, N_0 \in \mathbb{N}$ tal que:

$$\frac{1}{\delta}\mathrm{P}\left\{\sup_{t\leq s\leq t+\delta}|X_n(s) - X_n(t)| \geq \epsilon\right\} \leq \eta \quad n \geq N_0 \quad \forall t \in [\,0,1\,] \qquad (IX)$$

Demostración

Sean $\epsilon > 0$, $\eta > 0$, para $\epsilon^* = \frac{\epsilon}{2} > 0$, $\eta^* = \frac{\eta}{2} > 0$, $\exists\,\delta^*$ con $0 < \delta < 1$ y $\exists\, N_0^*$ tal que :

$$(8.11) \qquad \frac{1}{\delta^*}\mathrm{P}\left\{\max_{i\leq n\delta^*}\left(\frac{1}{\sigma\sqrt{n}}|S_{k+i}(\omega) - S_k(\omega)|\right) \geq \epsilon^*\right\} \leq \eta^*$$

$\forall n \geq N_0^*$, $\forall k$. Sea $\delta = \frac{\delta^*}{2}$, $N_0 = \max\{N_0^*, \frac{4}{\delta}\}$, $n \geq N_0$ y $t \in [\,0,1\,]$.

Para $t, t + \frac{\delta^*}{2}$ $\exists\, k$, $\exists\, j$ tales que

$$\frac{k}{n} \leq t < \frac{k+1}{n}, \frac{j-1}{n} \leq t + \frac{\delta^*}{2} < \frac{j}{n}.$$

Sea $s \in [t, t + \frac{\delta^*}{2}]$, $\omega \in \Omega$, supongamos ademas que

$$\frac{k+\lambda}{n} \leq s < \frac{k+\lambda+1}{n} \quad .$$

Lo que se hace a continuación es acotar las diferencias $|X_n(\frac{k}{n},\omega) - X_n(\frac{k+\lambda}{n},\omega)|$ aprovechando la definición de X en los puntos $\frac{k}{n}$ y ayudándonos de (#), así que

$$\begin{aligned}
|X_n(s,\omega) - X_n(t)| &\leq |X_n(s,\omega) - X_n(\frac{k}{n},\omega)| + |X_n(\frac{k}{n},\omega) - X_n(t,\omega)| \\
&\leq \max_{k\leq m\leq j}\left\{|X_n(\frac{k}{n},\omega) - X_n(\frac{m}{n},\omega)|\right\} + \max_{k\leq m\leq j}\left\{|X_n(\frac{k}{n},\omega) - X_n(\frac{m}{n},\omega)|\right\} \\
&= 2\max_{0\leq i\leq j-k}\left\{|X_n(\frac{k+i}{n},\omega) - X_n(\frac{k}{n},\omega)|\right\} \\
&= 2\max_{0\leq i\leq j-k}\left(\frac{1}{\sigma\sqrt{n}}|S_{k+i}(\omega) - S_k(\omega)|\right) .
\end{aligned}$$

Por tanto

$$\sup_{t\leq s\leq t+\frac{\delta^*}{2}} |X_n(s,\omega) - X_n(t,\omega)| \leq 2\max_{i\leq j-k}\left(\frac{1}{\sigma\sqrt{n}}|S_{k+i}(\omega) - S_k(\omega)|\right)$$

dado que $n \geq N_0 \geq \frac{4}{\delta^*}$ entonces

$$\frac{j}{n} - \frac{k}{n} = \left(\frac{j}{n} - (t + \frac{\delta^*}{2})\right) + \left(\frac{\delta^*}{2} + t - t\right) + \left(t - \frac{k}{n}\right) \leq \frac{\delta^*}{2} + \frac{2}{n} \leq \delta^*$$

por tanto $j - k < \delta^* n$ y entonces

$$\sup_{t \leq s \leq t + \frac{\delta^*}{2}} |X_n(s,\omega) - X_n(t,\omega)| \leq 2 \max_{i \leq \delta^* n} \left(\frac{1}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)|\right) \qquad (*)$$

Dado que se satisface (8.11) se tiene entonces que

$$\begin{aligned}
\frac{1}{\delta} \mathrm{P}\left\{\sup_{t \leq s \leq t+\delta} |X_n(s) - X_n(t)| \geq \epsilon\right\} &= \frac{2}{\delta^*} \mathrm{P}\left\{\sup_{t \leq s \leq t + \frac{\delta^*}{2}} |X_n(s) - X_n(t)| \geq \epsilon\right\} \\
&\leq \frac{2}{\delta^*} \mathrm{P}\left\{\max_{i \leq n\delta^*} \frac{2}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \geq \epsilon\right\} \\
&= \frac{2}{\delta^*} \mathrm{P}\left\{\max_{i \leq n\delta^*} \frac{2}{\sigma\sqrt{n}} |S_{k+i}(\omega) - S_k(\omega)| \geq \epsilon^*\right\} \\
&= 2\eta^* = \eta \quad \forall n \geq N_0^*
\end{aligned}$$

Y por tanto también para $n \geq N_0$.

El esquema que se tenemos para las X_n definidas en (8.6) es el siguiente :

(i) $\{X_n(0)\}$ siempre es tensa.

(ii) Para verificar la condición VII, basta verificar la VIII, (porque esta implica IX).

(iii) Pero la condición II (que es para procesos en general) se traduce en la condición VI, (que es para variables aleatorias C-valuadas), que a su vez se traduce en la condición VII (para procesos estocásticos).

En resumen tenemos el siguiente :

Corolario 8.7

$\{X_n\}$ es tensa si se satisface VIII.

Teorema 8.8

Supóngase que $\{X_n\}$ está definida por (8.6). La sucesión $\{X_n\}$ es tensa si $\forall \epsilon > 0$, $\exists\ \lambda > 1$ y $\exists N_0$ tal que

$$\mathrm{P}\left\{\max_{i \leq n} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{n}\right\} \leq \frac{\epsilon}{\lambda^2}$$

$\forall n \geq N_0$ y $\forall k$,.

Demostración

Sea $\epsilon > 0$, $\eta > 0$, $k \geq 1$. Mostraremos que $\exists\ \delta > 0$ y $\exists\ N_0 \in \mathbb{N}$ tal que se satisface la condición VIII y en vista del corolarío 8.7 se tendrá que $\{X_n\}$ es tensa.

Tomemos

$$\eta^* = \min\{\eta, 1\},\ \epsilon_1 = \min\{\epsilon, 1\},\ \epsilon^* = \eta^* \epsilon_1^2 .$$

Nótese que $\epsilon^* \leq \epsilon_1^2 \leq \epsilon_1 < \epsilon$, para $\epsilon^* > 0$, $\exists\ \lambda > 1$ y $\exists N_1$ tal que

$$\mathrm{P}\left\{\max_{i \leq n} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{n}\right\} \leq \frac{\eta^*\epsilon_1^2}{\lambda^2}$$

$\forall n \geq N_1$ y $\forall k$, sea $\delta = \frac{\epsilon_1^2}{\lambda^2}$ como $\lambda > 1$ entonces $0 < \delta < 1$. Sea N_0 tal que $N_0 \geq \frac{N_1}{\delta}$, sea $n \geq N_0$, entonces $[n\delta] \geq N_1$ por tanto

$$\mathrm{P}\left\{\max_{i \leq [n\delta]} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{[n\delta]}\right\} \leq \frac{\eta^*\epsilon_1^2}{\lambda^2} = \eta^*\delta .$$

Notemos que

$$\lambda\sqrt{[n\delta]} = \lambda\sqrt{[n\epsilon^{*2}\delta]} \leq \frac{\lambda\sqrt{n}\epsilon^*}{\lambda} = \sqrt{n}\epsilon^* \qquad (*)$$

y que

$$\delta = \frac{\epsilon^{*2}}{\lambda^2} < 1 \Rightarrow \left[\frac{n\epsilon^{*2}}{\lambda^2}\right] \leq \frac{n\epsilon^{*2}}{\lambda^2} . \qquad (**)$$

Tenemos entonces que :

$$\begin{aligned}
\frac{1}{\delta}\mathrm{P}\left\{\max_{i \leq n\delta} \frac{1}{\sigma\sqrt{n}}|S_{k+i} - S_k| \geq \epsilon\right\} &= \frac{1}{\delta}\mathrm{P}\left\{\max_{i \leq [n\delta]} \frac{1}{\sigma\sqrt{n}}|S_{k+i} - S_k| \geq \epsilon\right\} \\
&= \frac{1}{\delta}\mathrm{P}\left\{\max_{i \leq [n\delta]} |S_{k+i} - S_k| \geq \epsilon\sigma\sqrt{n}\right\} \\
&= \frac{1}{\delta}\mathrm{P}\left\{\max_{i \leq [n\delta]} |S_{k+i} - S_k| \geq \epsilon^*\sigma\sqrt{n}\right\} \qquad \text{por } * \\
&= \frac{1}{\delta}\mathrm{P}\left\{\max_{i \leq [n\delta]} |S_{k+i} - S_k| \geq \lambda\sigma\sqrt{[n\delta]}\right\} \qquad \text{por } ** \\
&\leq \frac{1}{\delta}\eta^*\delta = \eta^* < \eta ,
\end{aligned}$$

Lo cual prueba el teorema.

9. TENSION EN C[0,1].

Estudiaremos en este capítulo, algunas condiciones bajo las cuales una sucesión de procesos es tensa, ya que esto acompañado de la convergencia de las distribuciones finito-dimensionales, hace que la convergencia de procesos estocásticos (teorema (7.1)) sea débil.

Sean $\xi_1, \xi_2, \ldots, \xi_m$ variables aleatorias, no necesariamente independientes ni tampoco necesariamente idénticamente distribuidas.

Sea $S_n = \xi_1 + \xi_2 + \cdots + \xi_n$ $(S_0 = 0)$ y

$$M_m = \max_{0 \leq k \leq m} |S_k|.$$

Nuestro propósito es lograr cotas para $\mathrm{P}\{M_m \geq \lambda\}$. Definamos también

$$M'_m = \max_{0 \leq k \leq m} \min\{|S_k|, |S_m - S_k|\}.$$

Dado que : $\min\{|S_k|, |S_m - S_k|\} \leq |S_k|$ entonces

$$M'_m \leq M_m.$$

Observemos que :

$$|S_k| \leq |S_m| + |S_m - S_k| \text{ y } |S_k| \leq |S_k| + |S_m|$$

lo cual implica que

$$|S_k| \leq \min\{(|S_k| + |S_m - S_k|)\ ,\ (|S_k| + |S_m|)\}$$

Dado que

$$\min\{(|S_k| + |S_m - S_k|)\ ,\ (|S_k| + |S_m|)\} = |S_m| + \min\{|S_m - S_k|,\ |S_k|\}$$

entonces

$$|S_k| \leq |S_m| + \min\{|S_m - S_k|\ , |S_k|\}$$

Lo que implica

$$M_m \leq M'_n + |S_m|.$$

y por lo tanto

$$\mathrm{P}\{M_m \geq \lambda\} \leq \mathrm{P}\left\{M'_m \geq \frac{\lambda}{2}\right\} + \mathrm{P}\left\{S_m \geq \frac{\lambda}{2}\right\} \qquad (I)$$

El siguiente teorema nos da una cota para el primer sumando del lado derecho de la desigualdad anterior.

Teorema 9.1

Sean $\gamma > 0, \alpha > \frac{1}{2}$.

Supóngase que $\forall \lambda > 0$ y $\forall i, \forall j, \forall k, 0 \leq i \leq j \leq k \leq m$ se satisface que :

$$\mathrm{P}\left\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\right\} \leq \frac{1}{\lambda^{2\gamma}} \left(\sum_{i<l\leq k} u_l\right)^{\frac{2}{\alpha}} \qquad (II)$$

Para algunos numeros no negativos $u_1, u_2, \ldots, u_m$. Entonces $\forall \lambda > 0$ se tiene que :

$$\mathrm{P}\left\{M'_m \geq \lambda\right\} \leq \frac{K_{\gamma,\alpha}}{\lambda^{2\gamma}}(u_1 + u_2 + \cdots + u_m)^{\frac{2}{\alpha}} \qquad (III)$$

Donde $K_{\gamma,\alpha}$ es una constante que depende solamente de γ y α.

De hecho se puede calcular explicitamente la constante y es :

$$K_{\gamma,\alpha} = \left[\frac{1}{2^{\frac{1}{2\gamma+1}}} - \left(\frac{1}{2^{\frac{1}{(2\gamma+1)}}}\right)^{\frac{2}{\alpha}}\right]^{-(2\gamma+1)}$$

No haremos la demostración de este teorema ya que esta es muy técnica. Sin embargo haremos una observación y esbozaremos solo la idea.

Nótese que

$$\{M'_m \geq \lambda\} \subseteq \bigcup_{0\leq i\leq j\leq k\leq m} \{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\} \ .$$

Pues si tomamos un ω en el primer conjunto podemos suponer que

$$\max_{0\leq k\leq m} (\min\{|S_k|, |S_m - S_k|\}) = \min\{|S_{k_0}|, |S_m - S_{k_0}|\}$$

Y por tanto ω está en $\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\}$ con $i = 0,\ j = k_0,\ k = m$.

Así que

$$\begin{aligned}
\mathrm{P}\{M'_m \geq \lambda\} &\leq \sum_{0\leq i\leq j\leq k\leq m} \mathrm{P}\left\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\right\} \\
&\leq \sum_{0\leq i\leq j\leq k\leq m} \frac{1}{\lambda^{2\gamma}} \left(\sum_{i<l\leq k} u_l\right)^{\frac{2}{\alpha}} \\
&\leq \sum_{0\leq i\leq j\leq k\leq m} \frac{1}{\lambda^{2\gamma}} \left(\sum_{i=1}^{m} u_i\right)^{\frac{2}{\alpha}} \\
&= \frac{K}{\lambda^{2\gamma}} \left(\sum_{i=1}^{m} u_i\right)^{\frac{2}{\alpha}}
\end{aligned}$$

Nótese que aquí la constante K depende de m.

La idea es demostrar que el resultado vale para una constante K mayor que uno que no dependa de m, y la idea es hacerlo por inducción, sobre m, haremos el caso de $m = 1, m = 2$.

Si $m = 1$ entonces $M_1' = |S_1|$ y por tanto

$$\mathrm{P}\left\{M_m' \geq \lambda\right\} = \mathrm{P}\left\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\right\}$$

con $i = 0 = j = 1, k = 1$ por tanto se tiene el resulado para $m = 1$.

Si $m = 2$ entonces $M_2' = \min\{|S_2 - S_1|, |S_1|\}$; Dado que

$$\{\min\{|S_2 - S_1|, |S_1|\} \geq \lambda\} \subseteq \{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\}$$

con $i = 0 = j = 1, k = 2$ entonces por hipótesis se tiene que :

$$\mathrm{P}\left\{M_m' \geq \lambda\right\} \leq \frac{1}{\lambda^{2\gamma}} \left(\sum_{0<l\leq 2} u_l\right)^{\frac{2}{\alpha}} \leq \frac{K}{\lambda^{2\gamma}}(u_1 + u_2).$$

Los pasos subsecuentes de la demostración consisten en suponer que vale para los enteros menores o iguales que m, y escribir $u = u_1 + u_2 + \cdots + u_m$ encontrar un entero h tal que

$$\frac{u_1 + u_2 + \cdots + u_{h-1}}{u} \leq \frac{1}{2} \leq \frac{u_1 + u_2 + \cdots + u_h}{u}$$

y aplicar la hipótesis de inducción a las variables aleatorias : $\xi_1, \xi_2, \ldots, \xi_{h-1}$ y después aplicar nuevamente inducción a las variables aleatorias : $\xi_{h+1}, \xi_{h+2}, \ldots, \xi_m$.

Nótese que dado un m fijo podemos calcular la constante mínima entre la dada en el teorema y la de la observación.

Mostremos ahora algunas consecuencias del teorema anteiror.

Teorema 9.2

Sean $\gamma > 0, \alpha > 1$.

Supóngase que $\forall\lambda > 0$ y $\forall i, \forall j,\ 0 \leq i \leq j \leq m$ se satisface que :

$$\mathrm{P}\left\{|S_j - S_i| \geq \lambda\right\} \leq \frac{1}{\lambda^{\gamma}} \left(\sum_{i<l\leq j} u_l\right)^{\alpha} \qquad (IV)$$

Para algunos numeros no negativos $u_1, u_2, \ldots, u_m$. Entonces $\forall\lambda > 0$ se tiene que :

$$\mathrm{P}\left\{M_m \geq \lambda\right\} \leq \frac{K_{\gamma,\alpha}'}{\lambda^{2\gamma}}(u_1 + u_2 + \cdots + u_m)^{\alpha} \qquad (V)$$

Donde $K_{\gamma,\alpha}'$ es una constante que depende solamente de γ y α.

Demostración

Ya que

$$0 \leq \mathrm{P}(E_1 \cap E_2) \leq \mathrm{P}(E_1)$$

y dado que

$$0 \leq \mathrm{P}(E_1 \cap E_2) \leq \mathrm{P}(E_2)$$

entonces

$$\mathrm{P}(E_1 \cap E_2) \leq \mathrm{P}(E_1)^{\frac{1}{2}}\mathrm{P}(E_2)^{\frac{1}{2}} \qquad (*1).$$

Por otra parte $x = \left(\sum_{i<l\leq j} u_l\right)^{\frac{\alpha}{2}}$, y $y = \left(\sum_{j<l\leq k} u_l\right)^{\frac{\alpha}{2}}$ en $xy \leq (x+y)^2$ obtenemos que :

$$\left(\sum_{i<l\leq j} u_l\right)^{\frac{\alpha}{2}} \left(\sum_{j<l\leq k} u_l\right)^{\frac{\alpha}{2}} \leq \left(\sum_{i<l\leq k} u_l\right)^{\alpha} \qquad (*2)$$

De (*1) y de (*2) obtengo que

$$\begin{aligned} \mathrm{P}\left\{|S_j - S_i| \geq \lambda, |S_k - S_j| \geq \lambda\right\} &\leq \frac{1}{\lambda^{\frac{\gamma}{2}}}\left(\sum_{i<l\leq j} u_l\right)^{\frac{\alpha}{2}} \frac{1}{\lambda^{\frac{\gamma}{2}}}\left(\sum_{j<l\leq k} u_l\right)^{\frac{\alpha}{2}} \\ &\leq \frac{1}{\lambda^{\gamma}}\left(\sum_{i<l\leq k} u_l\right)^{\alpha} \end{aligned}$$

Se puede aplicar el teorema (9.1), pues se cumple (II) y concluir que

$$\mathrm{P}\left\{M'_m \geq \lambda\right\} \leq \frac{K_{\gamma,\alpha}}{\lambda^{\gamma}}(u_1 + u_2 + \cdots + u_m)^{\alpha} . \qquad (*3)$$

Además de la hipótesis tomando $i = 0$, $j = m$ se tiene que

$$\mathrm{P}\left\{|S_m| \geq \lambda\right\} \leq \frac{1}{\lambda^{\gamma}}(u_1 + u_2 + \cdots + u_m)^{\alpha} . \qquad (*4)$$

De (I) se sigue, aplicando (*3) y (*4) que :

$$\begin{aligned} \mathrm{P}\left\{M_m \geq \lambda\right\} &\leq \frac{K_{\frac{\gamma}{2},\alpha}}{\lambda^{\frac{\gamma}{2}}}(u_1 + u_2 + \cdots + u_m)^{\alpha} + \frac{1}{\lambda^{\frac{\gamma}{2}}}(u_1 + u_2 + \cdots + u_m)^{\alpha} \\ &= \frac{K_{\frac{\gamma}{2},\alpha} + 1}{\lambda^{\frac{\gamma}{2}}}(u_1 + u_2 + \cdots + u_m)^{\alpha} . \end{aligned}$$

Con lo que se tiene el resultado.

Teorema 9.3

La sucesión $\{X_n\}$ es tensa si satisface las siguentes dos condiciones :

(i) La sucesión $\{X_n(0)\}$ es tensa.

(ii) Existen constantes $\gamma \geq 0$ y $\alpha > 1$ y existe F función no decreciente continua sobre [0,1] tal que :

$$\mathrm{P}\{|X_n(t_2) - X_n(t_1)| \geq \lambda\} \leq \frac{1}{\lambda^\gamma}|F(t_2) - F(t_1)|^\alpha \qquad (VI)$$

$\forall t_1, \forall t_2, \forall n, \forall \lambda > 0$.

Demostración

Por el teorema (7.4) solo basta probar que $\forall \epsilon > 0$, $\forall \eta > 0$, $\exists\, \delta$, $0 < \delta < 1$ tal que

$$\mathrm{P}\{w(X_n, \delta) \geq \epsilon\} \leq \eta \qquad \forall n .$$

Pero por el corolario (7.5) esto es cierto si : $\delta^{-1} \in \mathbb{Z}$ y además :

$$\sum_{j<\delta^{-1}} \mathrm{P}\left\{ \sup_{j\delta \leq s \leq (j+1)\delta} |X_n(s) - X_n(j\delta)| \geq \epsilon \right\} \leq \eta . \qquad (VII)$$

Sea $n \in \mathbb{N}$, $\epsilon > 0$, $\eta > 0$. Consideremos $\delta \in]\,0,1\,[$ tal que

$$\frac{K}{\epsilon^\gamma}|F(1) - F(0)| \left[\max_{j<\delta^{-1}} |F((j+1)\delta) - F(j\delta)| \right]^{\alpha-1} < \eta.$$

Esto puede hacerse ya que F es uniformemente continua y $\alpha > 1$.

Para cada entero positivo m consideremos las variables aleatorias :

$$\xi_i = X_n(j\delta + \frac{i}{m}\delta) - X_n(j\delta + \frac{i-1}{m}\delta), \quad i = 1, 2, \ldots, m.$$

Si llamamos $u_i = F(j\delta + i\delta m^{-1}) - F(j\delta + (i-1)\delta m^{-1})$, probaremos que las variables aleatorias ξ_i satisfacen la condición del teorema 9.2, para tales u_i $i = 1, 2, \ldots$ Nótese que como F es no decreciente, los números u_i, $i = 1, 2, \ldots$ son no negativos.

Sean i, j tales que $i \leq j$, entonces :

$$\begin{aligned}
|S_j - S_i| &= |\xi_{i+1} + \xi_{i+2} + \cdots + \xi_j| \\
&= |X_n(j\delta + \frac{i+1}{m}\delta) - X_n(j\delta + \frac{i}{m}\delta) + \\
&\quad X_n(j\delta + \frac{i+2}{m}\delta) - X_n(j\delta + \frac{i+1}{m}\delta) + \\
&\quad + \cdots + \\
&\quad X_n(j\delta + \frac{j}{m}\delta) - X_n(j\delta + \frac{j-1}{m}\delta)| \\
&= |X_n(j\delta + \frac{j}{m}\delta) - X_n(j\delta + \frac{i}{m}\delta)| ,
\end{aligned}$$

Por tanto :

$$\begin{aligned}
\mathrm{P}\left\{|S_j - S_i| \geq \epsilon\right\} &= \mathrm{P}\left\{|X_n(j\delta + \frac{j}{m}\delta) - X_n(j\delta + \frac{i}{m}\delta)| \geq \epsilon\right\} \\
&\leq \frac{1}{\epsilon^\gamma}|F(j\delta + \frac{j}{m}\delta) - F(j\delta + \frac{i}{m}\delta)|^\alpha \quad \text{por (VI)} \\
&= \frac{1}{\epsilon^\gamma}|\sum_{k>i}^{j}\left(F(j\delta + \frac{k}{m}\delta) - F(j\delta + \frac{k-1}{m}\delta)\right)|^\alpha \\
&\leq \frac{1}{\epsilon^\gamma}\sum_{k>i}^{j}|F(j\delta + \frac{k}{m}\delta) - F(j\delta + \frac{k-1}{m}\delta)|^\alpha \\
&= \frac{1}{\epsilon^\gamma}\sum_{k>i}^{j} u_k^\alpha \quad .
\end{aligned}$$

Así que se cumple (VI) del teorema 9.2, por tanto existe una constante K, que satisface

$$\mathrm{P}\left\{\max_{0\leq i\leq m}|X_n(j\delta + \frac{i}{m}\delta) - X_n(j\delta)| \geq \epsilon\right\} \leq \frac{K}{\epsilon^\gamma}|F\left((j+1)\delta\right) - F(j\delta)|^\alpha. \qquad (*5)$$

Sea $0 < \beta < \epsilon$ probaremos que :

$$\mathrm{P}\left\{\sup_{j\delta\leq s\leq(j+1)\delta}|X_n(s) - X_n(j\delta)| \geq \epsilon\right\} \leq \frac{K}{(\epsilon-\beta)^\gamma}|F\left((j+1)\delta\right) - F(j\delta)|^\alpha. \qquad (*6)$$

Sea

$$\omega \in \{\alpha = \sup_{j\delta\leq s\leq(j+1)\delta}|X_n(s) - X_n(j\delta)| \geq \epsilon\}.$$

Por definición de supremo se tiene que $\exists\, s \in [j\delta, (j+1)\delta]$ tal que

$$|X_n(s) - X_n(j\delta)| \in (\alpha - \frac{\beta}{2}, \alpha]. \qquad (*7)$$

Ya que X_n es continua en s, para $\frac{\beta}{2}$, $\exists\, r > 0$ tal que si $|x - s| < r$ entonces :

$$|X_n(s) - X_n(x)| < \frac{\beta}{2} \quad . \qquad (*8)$$

Para $r > 0$, $\exists\, m$ tal que $\frac{1}{m} < r$. Supóngase que

$$s \in [j\delta + \frac{k}{m}\delta, j\delta + \frac{k+1}{m}\delta],$$

entonces :

$$\begin{aligned}
\max_{0\leq i\leq m}|X_n(j\delta + \frac{i}{m}\delta) - X_n(j\delta)| &\geq |X_n(j\delta + \frac{k}{m}\delta) - X_n(j\delta)| \\
&\geq |X_n(j\delta) - X_n(s)| - |X_n(s) - X_n(j\delta + \frac{k}{m}\delta)| \\
&\geq \alpha - \frac{\beta}{2} - \frac{\beta}{2}, \ \text{por (*7) y (*8)} \\
&= \alpha - \beta \geq \epsilon - \beta, \quad \text{ya que } \alpha \geq \epsilon.
\end{aligned}$$

Así que

$$\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X_n(s)-X_n(j\delta)|\geq \epsilon\right\}\subseteq\left\{\max_{0\leq i\leq m}|X_n(j\delta+\frac{i}{m}\delta)-X_n(j\delta)|\geq \epsilon-\beta\right\}$$

Y por lo anterior tenemos el resultado.

Dado que (*6) se cumple para toda $\beta\in(0,\epsilon)$, tenemos :

$$\mathrm{P}\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X_n(s)-X_n(j\delta)|\geq \epsilon\right\}\leq \lim_{\beta\to 0}\frac{K}{(\epsilon-\beta)^\gamma}[F((j+1)\delta)-F(j\delta)]^\alpha .$$

Por tanto

$$\mathrm{P}\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X_n(s)-X_n(j\delta)|\geq \epsilon\right\}\leq \frac{K}{\epsilon^\gamma}[F((j+1)\delta)-F(j\delta)]^\alpha .$$

Ahora tenemos que :

$$\begin{aligned}
&\sum_{j<\delta^{-1}}\mathrm{P}\left\{\sup_{j\delta\leq s\leq (j+1)\delta}|X_n(s)-X_n(j\delta)|\geq \epsilon\right\}\\
&\leq \frac{K}{\epsilon^\gamma}\sum_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)]^\alpha\\
&=\frac{K}{\epsilon^\gamma}\sum_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)][F((j+1)\delta)-F(j\delta)]^{\alpha-1}\\
&\leq \frac{K}{\epsilon^\gamma}\sum_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)]\max_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)]^{\alpha-1}\\
&=\frac{K}{\epsilon^\gamma}\max_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)]^{\alpha-1}\sum_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)]\\
&=\frac{K}{\epsilon^\gamma}\max_{j<\delta^{-1}}[F((j+1)\delta)-F(j\delta)]^{\alpha-1}[F(1)-F(0)] \ .\\
&<\eta
\end{aligned}$$

Esta última desigualdad es debida a la elección de δ, lo cual prueba el teorema.

10. TEOREMA CENTRAL DE LIMITE EN $\mathbb{R}^n$.

Sea $\phi: \mathbb{R} \to \mathbb{R}$ definida como:

$$\phi(t) = \begin{cases} 1 & t \leq 0 \\ 0 & 1 \leq t \\ \alpha^{-1} \int_t^1 \exp(\frac{-1}{s(1-s)})\ ds & t \in [0,1] \end{cases}$$

con $\alpha = \int_0^1 \exp(\frac{-1}{s(1-s)})\ ds$.

Observación 10.1.

ϕ es C^∞ y sus derivadas son acotadas.

Teorema 10.2

Sean P_n, P medidas de probabilidad sobre $\mathbb{R}$ tales que :

$$\int f\ dP_n = \int f\ dP \qquad \forall f \in C^\infty \text{acotada y con derivadas acotadas}$$

entonces $P_n \Rightarrow P$.

Demostración

Sean F_n, F las distribuciones de P_n, P respectivamente $n \in \mathbb{N}$. Sea x tal que f es continua en x y $u \in \mathbb{N}$ entonces

$$\begin{aligned} F_n(x) = P_n\{y \leq x\} &= \int_{\{y \leq x\}} dP_n(y) \\ &= \int_{\{(y-x) \leq 0\}} \phi(u(y-x))\ dP_n(y) \leq \int_{\mathbb{R}} \phi(u(y-x))\ dP_n(y). \end{aligned}$$

Por tanto

$$\begin{aligned}
\limsup F_n(x) &\leq \lim_{n\to\infty} \int_{\mathbb{R}} \phi(u(y-x))\ dP_n(y) \\
&= \int_{\mathbb{R}} \phi(u(y-x))\ dP(y) \\
&= \int_{\{u(y-x)\leq 1\}} \phi(u(y-x))\ dP(y) \\
&\leq \int_{\{u(y-x)\leq 1\}} dP(y) \\
&= \int_{\{y\leq x+\frac{1}{u}} dP(y) \\
&= P\left\{y \leq x + \frac{1}{u}\right\} \\
&= F(x+\frac{1}{u}).
\end{aligned}$$

Por otra parte

$$\begin{aligned}
F(x-\frac{1}{u}) &\leq \int_{\{y\leq x-\frac{1}{u}\}} 1\ dP(y) \\
&= \int_{\{u(y-x)\leq -1\}} 1\ dP(y) \\
&= \int_{\{u(y-x)\leq -1\}} \phi(u(y-x)+1)\ dP(y) \\
&\leq \int_{\mathbb{R}} \phi(u(y-x)+1)\ dP(y) \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} \phi(u(y-x)+1)\ dP_n(y) \\
&= \lim_{n\to\infty} \int_{\{u(y-x)\leq 0\}} \phi(u(y-x)+1)\ dP_n(y) \\
&\quad + \lim_{n\to\infty} \int_{\{u(y-x)>0\}} \phi(u(y-x)+1)\ dP_n(y) \\
&= \lim_{n\to\infty} \int_{\{u(y-x)\leq 0\}} \phi(u(y-x)+1)\ dP_n(y) \\
&\quad + \lim_{n\to\infty} \int_{\{u(y-x)>0\}} 0\ dP_n(y) \\
&= \lim_{n\to\infty} \int_{\{u(y-x)\leq 0\}} \phi(u(y-x)+1)\ dP_n(y) \\
&\leq \liminf \int_{\{u(y-x)\leq 0\}} 1\ dP_n \\
&= \liminf F_n(x).
\end{aligned}$$

Haciendo $u \to \infty$, se tiene

$$F(x) = \lim_{u\to\infty} F(x-\frac{1}{u}) \leq \liminf F_n(x) \leq \limsup F_n(x) \leq \lim_{u\to\infty} F(x+\frac{1}{u}) = F(x)$$

Por tanto $F_n(x) \to F(x)$, lo cual prueba que $P_n \Rightarrow P$.

Teorema 10.3

Sean $\xi_{n_1}, \xi_{n_2}, \ldots, \xi_{n_{k(n)}}$ para cada n, variables aleatorias independientes idénticamente distribuidas con media cero y tales que para cada $j \leq k(n)$, ξ_{n_j} tiene varianza finita $\sigma^2_{n_j}$. Sea

$$S_n = \xi_{n_1} + \xi_{n_2} + \cdots + \xi_{n_{k(n)}},$$

con varianza $s_n^2 = \sigma^2_{n_1} + \sigma^2_{n_2} + \cdots + \sigma^2_{n_{k(n)}} > 0$. Sea N una variable aleatoria normalmente distribuida con media cero y varianza 1. Si $\forall \epsilon > 0$:

$$(**) \qquad \frac{1}{s_n^2} \sum_{k=1}^{k(n)} \int_{\{|\xi_{n_k}| \geq \epsilon s_n\}} \xi_{n_k}{}^2 \ dP \to 0 \quad \text{si } n \to \infty$$

entonces

$$\frac{S_n}{s_n} \xrightarrow{D} N.$$

En base al teorema 10.2 basta probar que $\forall f \in C^\infty$ con derivadas acotadas se tiene que

$$\int f(\frac{S_n}{s_n}) \ dP = \int f(N) \ dP \ .$$

Sea $f \in C^\infty$ con derivadas acotadas y sea

$$g(h) = \sup_{x \in \mathbb{R}} |f(x+h) - f(x) - f'(x)h - \frac{1}{2} f''(x)h^2|.$$

Observación 10.4. g es borel-medible.(de hecho es continua)

Demostración (observación 10.4).

Sea $\alpha > 0$, $h_0 \in \mathbb{R}$ y sean

$$i(x) = |f(x+h_0) - f(x) - f'(x)h_0 - \frac{1}{2} f''(x)h_0^2| \ ,$$

$$j(x) = |f(x+h) - f(x) - f'(x)h - \frac{1}{2} f''(x)h^2|.$$

Sean M_1, M_2 las cotas para f', f'' respectivamente, dado que $k(x) := x^2$ es continua, para $2\alpha/(3M_2) > 0$, $\exists \ \delta_1 > 0$ tal que $|h - h_0| < \delta_1 \Rightarrow |h^2 - h_0^2| < \frac{2\alpha}{(3M_2)}$, sea ahora $\delta_2 = \frac{\alpha}{3M_1}$, tomemos $\delta = \min\{\delta_1; \delta_2\}$, si h cumple con $|h - h_0| < \delta$, y si $x \in \mathbb{R}$ entonces

$$\begin{aligned} |i(x) - j(x)| &\leq |f(x+h_0) - f(x+h)| + |f'(x)||h_0 - h| + |\frac{f''(x)}{2}||h_0^2 - h^2| \\ &= |f'(\xi_{x,h_0})||h_0 - h| + |f'(x)||h_0 - h| + |\frac{f''(x)}{2}||h_0^2 - h^2| \\ &\leq \alpha/3 + \alpha/3 + \frac{M_2}{2} \frac{2\alpha}{(3M_2)} = \alpha \ . \end{aligned}$$

Por tanto $i(x) < j(x) + \alpha$, tomndo supremos obtenemos :

$$\sup_{x\in\mathbb{R}} i(x) \le \alpha + \sup_{x\in\mathbb{R}} j(x)$$

es decir $g(h_0) \le g(h) + \alpha$, analogamente $g(h) \le g(h_0) + \alpha$, por tanto $|g(h) - g(h_0)| \le \alpha$.

Observación 10.5 $\exists\ K$ que depende de f tal que $g(h) \le \min\{h^2, |h|^3\}$ ·

Demostración (observación 10.5).

$$\begin{aligned}
g(h) &= |f(x+h) - f(x) - f'(x)h - \frac{1}{2}f''(x)h^2| \\
&= |f'(x_h)h - f'(x)h - \frac{1}{2}f''(x)h^2| \qquad \text{donde } x_h \in [x, x+h] \\
&= |(f'(x_h) - f'(x))h - \frac{1}{2}f''(x)h^2| \\
&= |f''(x_{\hat{h}})h^2 - \frac{1}{2}f''(x)h^2| \qquad \text{donde } x_{\hat{h}} \in [x, x_h] \\
&= |\frac{1}{2}(f''(x_{\hat{h}}))h^2| + \\
&\qquad |\frac{1}{2}(f''(x_{\hat{h}}))h^2 - \frac{1}{2}f''(x))h^2| \qquad \text{donde } x_{\hat{h}} \in [x, x_h] \\
&\le \frac{1}{2}|M_1|h^2 + \frac{1}{2}|f''(x_{\hat{h}}) - f''(x)|h^2 \\
&= \frac{1}{2}|M_1|h^2 + \frac{1}{2}|f'''(x_{\tilde{h}})|h^2|h| \qquad \text{donde } x_{\tilde{h}} \in [x, x_{\hat{x}}] \\
&\le \frac{1}{2}|M_1|h^2 + \frac{1}{2}M_2|h^3| \qquad K/2 = \max\{\frac{|M_1|}{2}, \frac{|M_2|}{2}\} \\
&\le K/2(h^2 + |h^3|) \\
&\le K/2(h^2 + |h^3|) \\
&\le K\min\{h^2, |h^3|\}
\end{aligned}$$

Finalmente para demostrar el teorema 10.3 emplearemos una última observación, la cual nos servirá de ayuda.

Observación 10.6.

$$\begin{aligned}
&|f(x+h_1) - f(x+h_2) - [f'(x)(h_1 - h_2) + \frac{1}{2}f''(x)(h_1^2 - h_2^2)]| \\
&\le |f(x+h_1) - f(x) - f'(x)h_1 - \frac{1}{2}f''(x)h_1^2 \\
&\quad - [f(x+h_2) - f(x) - f'(x)h_2 - \frac{1}{2}f''(x)h_2^2]| \\
&\le g(h_1) + g(h_2)\ .
\end{aligned}$$

Demostración (del teorema 10.3)

Consideremos la siguiente sucesión :

$$E(f(s_n^{-1}(\xi_{n_1+}\xi_{n_2+\cdots+}\xi_{n_{k(n)}})))$$
$$E(f(s_n^{-1}(\xi_{n_1+}\xi_{n_2+\cdots+}\eta_{n_{k(n)}})))$$
$$\vdots$$
$$E(f(s_n^{-1}(\eta_{n_1}+\eta_{n_2}+\cdots+\eta_{n_{k(n)}}))$$

Donde η_{n_j} son variables aleatorias con distribución normal, media cero y varianza $\sigma_{n_j}^2$ tales que :

(*) $$\xi_{n_1}, \xi_{n_2,\ldots,}\xi_{n_{k(n)}}, \eta_{n_1}, \eta_{n_2}, \ldots, \eta_{n_{k(n)}}$$

son independientes, (siempre existen).

Consideremos $\varsigma_{n_j} = \xi_{n_1} + \xi_{n_2} + \cdots + \xi_{n_{j-1}} + \eta_{n_{j+1}} + \eta_{n_{j+2}} + \cdots + \eta_{n_{k(n)}}$

Nótese que $\varsigma_{n_{k(n)}} + \xi_{n_{k(n)}} = S_n$ y que la distribución de $\varsigma_{n_1} + \xi_{n_1}$ es la misma que la de $s_n N$, pues

$$\begin{aligned} P\{\varsigma_{n_1} + \xi_{n_1} = \eta_{n_1} + \eta_{n_2} + \cdots + \eta_{n_{k(n)}} \in A\} \\ = \left(\frac{1}{\sqrt{2\pi}}\int_A e^{u^2}\, du\right)\sqrt{\sigma_{n_1}^2 + \sigma_{n_2}^2 + \cdots + \sigma_{n_{k(n)}}^2} \\ = P\{s_n N \in A\} \end{aligned}$$

Así que la idea es probar que cada elemento de la sucesión anterior está cerca del siguiente.

$$\begin{aligned} |E(f(\frac{S_n}{s_n})) - E(f(N))| &= \Big|[E(f(\frac{\xi_{n_1} + \eta_{n_2} + \cdots + \eta_{n_{k(n)}}}{s_n})) \\ &- E(f(\frac{\eta_{n_1} + \eta_{n_2} + \cdots + \eta_{n_{k(n)}}}{s_n}))] \\ &+ [E(f(\frac{\xi_{n_1} + \xi_{n_2} + \eta_{n_3} + \cdots + \eta_{n_{k(n)}}}{s_n})) \\ &- E(f(\frac{\xi_{n_1} + \eta_{n_2} + \cdots + \eta_{n_{k(n)}}}{s_n}))] + \cdots \\ &+ [E(f(\frac{\xi_{n_1} + \xi_{n_2+\cdots+}\xi_{n_{k(n)}}}{s_n})) \\ &- E(f(\frac{\xi_{n_1} + \xi_{n_2} + \cdots + \xi_{n_{k(n)-1}} + \eta_{n_{k(n)}}}{s_n}))]\Big| \\ &\leq \sum_{i=1}^{k(n)} |E[f(\frac{\varsigma_{n_i} + \xi_{n_i}}{s_n}) - f(\frac{\varsigma_{n_i} + \eta_{n_i}}{s_n})]| \end{aligned}$$

Como $\varsigma_{n_j}, \xi_{n_j}, \eta_{n_j}$ son independientes (pues ξ_{n_j}, η_{n_j} no aparecen como sumandos en ς_{n_j}) se tiene entonces que

$$\begin{aligned} E[f'(\frac{\varsigma_{n_j}}{s_n})(\xi_{n_j} - \eta_{n_j})] &= E[f'(\frac{\varsigma_{n_j}}{s_n})\xi_{n_j}] - E[f'(\frac{\varsigma_{n_j}}{s_n})\eta_{n_j}] \\ &= E(f'(\frac{\varsigma_{n_j}}{s_n}))[E(\xi_{n_j}) - E(\eta_{n_j})] \qquad \text{(independencia)} \\ &= 0 \end{aligned}$$

y

$$\begin{aligned} E[f''(\frac{\varsigma_{n_j}}{s_n})(\xi^2_{n_j} - \eta^2_{n_j})] &= E[f''(\frac{\varsigma_{n_j}}{s_n}))E(\xi^2_{n_j}) - E(f''(\frac{\varsigma_{n_j}}{s_n}))E(\eta^2_{n_j}) \\ &= \sigma^2_{n_j}[E(f''(\frac{\varsigma_{n_j}}{s_n})) - E(f''(\frac{\varsigma_{n_j}}{s_n}))] \qquad \text{(independencia)} \\ &= 0 \quad . \end{aligned}$$

Así tenemos :

$$\begin{aligned} |E(f(\frac{S_n}{s_n}) - E(f(N))| &\le \sum_{i=1}^{k(n)} |E(f(\frac{\varsigma_{n_i} + \xi_{n_i}}{s_n}) - f(\frac{\varsigma_{n_i} + \eta_{n_i}}{s_n}))| \\ &= \sum_{i=1}^{k(n)} |E\{f(\frac{\varsigma_{n_i}}{s_n} + \frac{\xi_{n_i}}{s_n}) - f(\frac{\varsigma_{n_i}}{s_n} + \frac{\eta_{n_i}}{s_n}) \\ &\quad - [f'(\frac{\varsigma_{n_i}}{s_n})(\frac{\xi_{n_i}}{s_n} - \frac{\eta_{n_i}}{s_n}) - \frac{1}{2}f''(\frac{\varsigma_{n_i}}{s_n})(\frac{\xi^2_{n_i} - \eta^2_{n_i}}{s^2_n})]\}| \\ &\le \sum_{i=1}^{k(n)} E\left(g(\frac{\xi_{n_i}}{s_n}) + g(\frac{\eta_{n_i}}{s_n})\right) . \end{aligned}$$

(la última desigualdad es consecuencia de la observación 10.6)

Por tanto basta probar que

$$\sum_{i=1}^{k(n)} E(g(\frac{\xi_{n_i}}{s_n})) \to 0$$

y que

$$\sum_{i=1}^{k(n)} E(g(\frac{\eta_{n_i}}{s_n})) \to 0.$$

Sea $\epsilon > 0$ tenemos que

$$\begin{aligned} \int_\Omega g(\frac{\xi_{n_i}}{s_n})\, dP &= \int_{\{|\xi_{n_i}| \le \epsilon s_n\}} g(\frac{\xi_{n_i}}{s_n})\, dP + \int_{\{|\xi_{n_i}| \ge \epsilon s_n\}} g(\frac{\xi_{n_i}}{s_n})\, dP \\ &\le \int_{\{|\xi_{n_i}| \le \epsilon s_n\}} K|\frac{\xi_{n_i}}{s_n}|^3\, dP + \int_{\{|\xi_{n_i}| \ge \epsilon s_n\}} K|\frac{\xi_{n_i}}{s_n}|^2\, dP \\ &\le K\epsilon \int_{\{|\frac{\xi_{n_i}}{s_n}| \le \epsilon\}} |\frac{\xi_{n_i}}{s_n}|^2\, dP + \frac{K}{s^2_n}\int_{\{|\xi_{n_i}| \ge \epsilon s_n\}} \xi_{n_i 2}\, dP \\ &\le \frac{K\epsilon}{s^2_n}\sigma^2_{n_i} + \frac{K}{s^2_n}\int_{\{|\xi_{n_i}| \ge \epsilon s_n\}} \xi_{n_i 2}\, dP . \end{aligned}$$

(en la primera desigualdad se usó la observación 10.5)

Por tanto

$$\sum_{i=1}^{k(n)} E(g(\frac{\xi_{n_i}}{s_n})) \leq \sum_{i=1}^{k(n)} \frac{K\epsilon}{s_n^2}\sigma_{n_i}^2 + K\sum_{i=1}^{k(n)} \frac{1}{s_n^2}\int_{\{|\xi_{n_i}|\geq\epsilon s_n\}} \xi_{n_i 2}\ dP$$
$$= K\epsilon + K\sum_{i=1}^{k(n)} \frac{1}{s_n^2}\int_{\{|\xi_{n_i}|\geq\epsilon s_n\}} \xi_{n_i}^2\ dP$$

por hipótesis, el segundo sumando de esta última desigualdad tiende a 0, si $n \to \infty$ (**).

Análogamente se denuestra que

$$\sum_{i=1}^{k(n)} E(g(\frac{\eta_{n_i}}{s_n})) \leq K\epsilon + K\sum_{i=1}^{k(n)} \frac{1}{s_n^2}\int_{\{|\eta_{n_i}|\geq\epsilon s_n\}} \eta_{n_i}^2\ dP$$

Para ver que el segundo sumando tiende a cero cuando $n \to \infty$, calculemos :

$$\begin{aligned}
\int_{\{|\eta_{n_i}|\geq\epsilon s_n\}} \eta_{n_i}^2\ dP &= \int_{\{|\eta_{n_i}|\geq\epsilon s_n\}} \eta_{n_i 3}\frac{1}{|\eta_{n_i}|}\ dP \\
&\leq \int_{\{\frac{1}{\epsilon s_n}\geq\frac{1}{|\eta_{n_i}|}\}} \eta_{n_i 3}\frac{1}{\epsilon s_n}\ dP \\
&= \frac{1}{\epsilon s_n}E|\eta_{n_i 3}| \\
&\leq \frac{1}{\epsilon s_n}\sigma_{n_i}^3 E(|N|^3)
\end{aligned}$$

esto último es debido a que la distribución de η_{n_i} es la misma que la de $s_n N$. Ahora

$$\begin{aligned}
\frac{\sigma_{n_i}^2}{s_n^2} &= \frac{1}{s_n^2}\int_\Omega \xi_{n_i 2}\ dP \\
&= \frac{1}{s_n^2}\int_{\{|\xi_{n_i}|\leq\epsilon s_n\}} \xi_{n_i}^2\ dP + \frac{1}{s_n^2}\int_{\{|\xi_{n_i}|\geq\epsilon s_n\}} \xi_{n_i}^2\ dP \\
&\leq \epsilon^2 + \frac{1}{s_n^2}\int_{\{|\xi_{n_i}|\geq\epsilon s_n\}} \xi_{n_i}^2\ dP
\end{aligned}$$

Aplicando nuevamente (**), el segundo término de la última desigualdad tiende a cero, por tanto $\frac{\sigma_{n_i}^2}{s_n^2} \to 0$, $\forall i \leq k(n)$, y esto implica que $\max_{i\leq k(n)} \frac{\sigma_{n_i}}{s_n} \to 0$, de lo cual tenemos que

$$\begin{aligned}\sum_{i=1}^{k(n)} \frac{\sigma_{n_i}^3}{s_n^3} &\leq \sum_{i=1}^{k(n)} \frac{\sigma_{n_i}^2}{s_n^2}\frac{\sigma_{n_i}}{s_n} \\ &\leq \sum_{i=1}^{k(n)} \frac{\sigma_{n_i}^2}{s_n^2} \max_{i\leq k(n)} \frac{\sigma_{n_i}}{s_n} \\ &= \max_{i\leq k(n)} \frac{\sigma_{n_i}}{s_n} \sum_{i=1}^{k(n)} \frac{\sigma_{n_i}^2}{s_n^2} \\ &= \max_{i\leq k(n)} \frac{\sigma_{n_i}}{s_n} \to 0 \quad ,\end{aligned}$$

y ademas

$$\sum_{i=1}^{k(n)} \frac{1}{s_n^2} \int_{\{|\eta_{n_i}|\geq \epsilon s_n\}} \eta_{n_i}^2 \, dP \leq \sum_{i=1}^{k(n)} \frac{1}{\epsilon s_n^3} \sigma_{n_i}^3 E(|N|^3) \to 0$$

lo cual completa la demostración.

Teorema 10.7

Si para algún $\delta > 0$ se tiene que:

$$\frac{1}{s_n^{2+\delta}} \sum_{i=1}^{k(n)} E\left\{|\xi_{n_i}|^{2+\delta}\right\} \to 0$$

donde ξ_{n_i}, tienen momentos de orden $2+\delta$ entonces

$$\frac{S_n}{s_n} \xrightarrow{D} N.$$

Demostración

Calculemos

$$\begin{aligned}&\frac{1}{s_n^2} \sum_{i=1}^{k(n)} \int_{\{|\xi_{n_i}|\geq \epsilon s_n\}} \xi_{n_i}^2 \, dP \\ &= \frac{1}{s_n^2} \sum_{i=1}^{k(n)} \int_{\{|\xi_{n_i}|\geq \epsilon s_n\}} |\xi_{n_i}|^{2+\delta} \frac{1}{|\xi_{n_i}|^\delta} \, dP \\ &\leq \frac{1}{s_n^2} \sum_{i=1}^{k(n)} \int_{\{|\xi_{n_i}|\geq \epsilon s_n\}} |\xi_{n_i}|^{2+\delta} \frac{1}{\epsilon s_n^\delta} \, dP \\ &= \frac{1}{\epsilon^\delta s_n^{2+\delta}} \sum_{i=1}^{k(n)} \int_{\{|\xi_{n_i}|\geq \epsilon s_n\}} |\xi_{n_i}|^{2+\delta} \, dP \\ &\leq \frac{1}{\epsilon^\delta} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{k(n)} E\left\{|\xi_{n_i}|^{2+\delta}\right\}\end{aligned}$$

Lo cual prueba (aplicando la hipótesis (**) y el teorema 10.3) el resultado.

Teorema 10.8 TEOREMA CENTRAL DE LIMITE EN $\mathbb{R}$ *Linderberg-Leví*

Si $\xi_1, \xi_2, \dots$ son variables aleatorias independientes e idénticamente distribuidas, con media cero y varianza $\sigma^2 > 0$, entonces

$$\frac{1}{\sigma\sqrt{n}}(\xi_1 + \xi_2 + \cdots + \xi_n) \xrightarrow{D} N \ .$$

Demostración

hacemos $s_n^2 = \underbrace{\sigma^2 + \sigma^2 + \cdots + \sigma^2}_{n-\text{ veces}}$, pongamos $k(n) = n$, $\xi_{n_i} = \xi_i$ y calculemos

$$\begin{aligned}\frac{1}{s_n^2}\sum_{i=1}^{k(n)}\int_{\{|\xi_{n_i}|\geq \epsilon s_n\}} \xi_{n_i}^2 \, dP &= \frac{1}{n\sigma^2}\sum_{i=1}^{k(n)}\int_{\{|\xi_{n_i}|\geq \epsilon n\sigma^2\}} \xi_{n_i}^2 \, dP \\ &= \frac{1}{n\sigma^2}\sum_{i=1}^{k(n)}\int_{\{|\xi_{n_1}|\geq \epsilon n\sigma^2\}} \xi_{n_1}^2 \, dP = \frac{1}{\sigma^2}\int_{\{|\xi_{n_1}|\geq \epsilon n\sigma^2\}} \xi_{n_1}^2 \, dP \\ &= \int_{\{|\xi_{n_1}|\geq \epsilon n\sigma^2\}} \frac{\xi_{n_1}^2}{\sigma^2} \, dP \ .\end{aligned}$$

Aplicando el teorema de convergencia monótona y aprovechando que ξ_1 es integrable esta última parte tiende a 0, lo cual prueba el teorema.

Corolario 10.9

Bajo las mismas hipótesis del teorema 10.4 se tiene que

$$\frac{\xi_1 + \xi_2 + \cdots + \xi_n}{\sqrt{n}} \xrightarrow{D} \hat{N}$$

donde $\hat{N}$ tiene distribución normal con media cero y varianza σ.

A continuación daremos un esquema general para demostrar el teorema de límite central en $\mathbb{R}^n$; Las demostraciones de los teoremas dados en seguida, se darán despues de estos.

Si P es una medida de probabilidad sobre $\mathbb{R}^n$, podemos considerar a la función $p: \mathbb{R}^n \to \mathbb{R}$ definida como

$$p(\hat{t}) = \int_{\mathbb{R}^k} \exp(i\hat{t}\cdot\hat{x}) \, dP(\hat{x}),$$

a tal función se le llama la función característica de P.

Se puede observar que p siempre está acotada por 1, y que es continua.

Sean P_n, P y Q medidas de probabilidad sobre $\mathbb{R}^n$, y sean p_n, p y q las funciones características de P_n, P y Q respectivamente.

Teorema 10.10

Supóngase que $p(\hat{t}) = q(\hat{t}) \quad \forall \hat{t}$, entonces $P \equiv Q$.

Teorema 10.11

Supóngase que $p_n(\hat{t}) \to g(\hat{t})$ para alguna función $g(\hat{t})$, supóngase además que $\{P_n\}$ es tensa, entonces $P_n \Rightarrow P$ para alguna P.

Teorema 10.12

Supóngase que $p_n(\hat{t}) \to g(\hat{t})$ para alguna g continua en cero, entonces $\{P_n\}$ es tensa.

Teorema 10.13

Supóngase que $p_n(\hat{t}) \to p(\hat{t})$, entonces $P_n \Rightarrow P$.

Sean $X_n = (X_n^1, X_n^2, \ldots, X_n^k)$, $X = (X_1, X_2, \ldots, X_k)$,

Teorema 10.14

supóngase que

$$\sum_{i=1}^{k} t_i X_n^i \xrightarrow{D} \sum_{i=1}^{k} t_i X^i \quad \forall (t_1, t_2, \ldots, t_n) \in \mathbb{R}^k$$

entonces

$$X_n \xrightarrow{D} X.$$

Teorema 10.15 (teorema de límite central en $\mathbb{R}^n$)

Sean $X_n^j = \frac{1}{\sqrt{n}}\left(\xi_1^j + \xi_2^j + \cdots + \xi_n^j\right)$ donde ξ_i^j es una variable aleatoria con media cero y varianza σ_j, consideremos también X_j =variable aleatoria con distribución normal $0, \sigma_j$, entonces

$$X_n \xrightarrow{D} X.$$

Demostración (teorema 10.15)

Es claro que por el teorema de Linderberg-Levi $X_n^j \xrightarrow{D} X_i$, por tanto

$$t_i X_n^j \xrightarrow{D} t_i X_i$$

por tanto

$$\sum_{i=1}^{k} t_i X_n^i \xrightarrow{D} \sum_{i=1}^{k} t_i X^i$$

aplcando el teorema 10.14 tenemos el resultado.

Demostración (teorema 10.14)

Sea $\hat{t} \in \mathbb{R}^n$, probaremos que $p_n(\hat{t}) \to p(\hat{t})$, donde p_n, p es la función característica de la medida de probabilidad P_{X_n}, P_X asociada a X_n, X respectivamente(distribuciones), y en virtud del teorema (10.13) se tendrá que $P_{X_n} \Rightarrow P_X$, es decir $X_n \xrightarrow{D} X$.

En virtu de que $P_{Y_n} \Rightarrow P_Y$, donde $\hat{t} = (t_1, t_2, \ldots, t_n)$, $Y_n = \sum_{i=1}^k t_i X_n^i$, y $Y = \sum_{i=1}^k t_i X^i$, entonces

$$p_{Y_n}(s) \to p_Y(s) \quad \forall s \in \mathbb{R}$$

donde $p_{Y_n}, p_Y(s)$ es la función característica de la distribución de Y_n, Y respectivamente, en particular para $s = 1$ se tiene que

$$\int_{\mathbb{R}} \exp(ix)\, dP_{Y_n}(x) \to \int_{\mathbb{R}} \exp(ix)\, dP_Y(x),$$

por el teorema de cambio de variable obtenemos que

$$\int_\Omega \exp\left(i(t_1 X_n^1 + t_2 X_n^2 + \ldots + t_k X_n^k)\right)\, dP \to \int_\Omega \exp\left(i(t_1 X_1 + t_2 X_2 + \ldots + t_k X_k)\right)\, dP$$

es decir

$$\int_\Omega \exp\left(i\hat{t} \cdot X_n\right)\, dP \to \int_\Omega \exp\left(i\hat{t} \cdot X\right)\, dP,$$

por tanto $p_n(\hat{t}) \to p(\hat{t})$. como se quería probar.

Demostración (teorema 10.13)

En virtud de que p es la función característica de P, entonces p es continua, y por el teorema (10.12) se tiene que $\{P_n\}$ es tensa, y por el teorema (10.11) se tiene que $P_n \Rightarrow P$.

Demostración (teorema 10.12)

Mostraremos que cada una de las k–distribuciones marginales es tensa, ya que esto implica que la sucesión $\{P_n\}$ es tensa. Nótese que la función característica $p_n(s,0,\ldots,0)$ correspondientes a las distribución marginal en la primera cordenada, converge puntualmente a la función $g(s,0,\ldots,0)$ continua en cero, y de igual forma para las otras, por tanto cada sucesión de distribuciones marginales $\{p_n(0,0,\ldots,s,\ldots,0)\}$ satisface la hipótesis del teorema, así que solo mostraremos el resltado en el caso $k=1$.

Dado $u \in \mathbb{R}^+$ tenemos que $(u>0)$

$$
\begin{aligned}
\frac{1}{u}\int_{-u}^{u}(1-p_n(t))\,dt &= \frac{1}{u}\int_{-u}^{u}\left(1-\left(\int_{\mathbb{R}}\exp(it\cdot x)\,dP_n(x)\right)\right)\,dt\\
&= \frac{1}{u}\int_{-u}^{u}1\,dt-\frac{1}{u}\int_{-u}^{u}\left(\int_{\mathbb{R}}\exp(it\cdot x)\,dP_n(x)\right)\,dt\\
&= \frac{1}{u}\int_{-u}^{u}1\,dt-\frac{1}{u}\int_{\mathbb{R}}\left(\int_{-u}^{u}\exp(it\cdot x)\,dt\right)\,dP_n(x) \quad \text{por Fubini}\\
&= \frac{1}{u}\int_{-u}^{u}1\,dt-\int_{\mathbb{R}}\left(\frac{1}{u}\int_{-u}^{u}\exp(it\cdot x)\,dt\right)\,dP_n(x)\\
&= \frac{1}{u}\int_{-u}^{u}1\,dt\int_{\mathbb{R}}dP_n(x)-\int_{\mathbb{R}}\left(\frac{1}{u}\int_{-u}^{u}\exp(it\cdot x)\,dt\right)\,dP_n(x)\\
&= \int_{\mathbb{R}}\left(\frac{1}{u}\int_{-u}^{u}1\,dt\right)\,dP_n(x)-\int_{\mathbb{R}}\left(\frac{1}{u}\int_{-u}^{u}\exp(it\cdot x)\,dt\right)\,dP_n(x)\\
&= \int_{\mathbb{R}}\left(\frac{1}{u}\int_{-u}^{u}(1-\exp(it\cdot x))\,dt\right)\,dP_n(x)\\
&= 2\int_{\mathbb{R}}\left(1-\frac{sin(ux)}{ux}\right)\,dP_n(x)\\
&\geq 2\int_{\{|x|\geq 2/u\}}\left(1-\frac{1}{|\,ux\,|}\right)\,dP_n(x)\\
&\geq 2\int_{\{|x|\geq 2/u\}}\frac{1}{2}\,dP_n(x)\\
&\geq P_n\{|\,x\,|\geq 2/u\}\ .
\end{aligned}
$$

Por tanto

$$
(**) \qquad \frac{1}{u}\int_{-u}^{u}(1-p_n(t))\,dt \geq P_n\{|\,x\,|\geq 2/u\}\ .
$$

Ya que g es continua en cero, podemos para un $\epsilon>0$, elegir una vecindad del cero $[-u,u]$ tal que

$$
\frac{1}{u}\int_{-u}^{u}|\,(1-g(t))\,|\ dt<\epsilon.
$$

Dado que $\mid 1 - p_n(t) \mid \to \mid 1 - g(t) \mid$ y dado que $\mid 1 - g(t) \mid$ está acotada, por un teorema de convergencia (acotada) tenemos que

$$\frac{1}{u}\int_{-u}^{u} \mid (1 - p_n(t)) \mid \ dt \to \frac{1}{u}\int_{-u}^{u} \mid (1 - g(t)) \mid \ dt$$

en particular existe un n_0 tal que $\frac{1}{u}\int_{-u}^{u} \mid (1 - p_n(t)) \mid \ dt < 2\epsilon$ si $n \geq n_0$, y por (**) se tiene que $P_n\{\mid x \mid \geq \frac{2}{u}\} < 2\epsilon$, si $n \geq n_0$ pero podemos suponer que vale para todo natural (haciendo mas grande la u). Por tanto la sucesión es tensa.

Demostración (teorema 10.11)

Si $\{P_n\}$ es tensa, entonces por el teorema de Prohorov, $\{P_n\}$ es relativamente compacto, cualquier subsucesión de esta $\{P_{n'}\}$, debe tener una subsucesión $\{P_{n''}\}$ convergente a un límite P, y por tanto, las funciones características de tal subsucesión, $p_{n''}$ deben converger a la función característica de P, por otra parte estas convergen a g, lo cual muestra (teorema (10.10)) que tal límite es único para todas las subsucesiones. Por tanto $P_n \Rightarrow P$ para alguna P.

Demostración (teorema 10.10)

Basta ver que P y Q coinciden sobre los rectángulos

$$A = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$$

para lo cual es suficiente probar que

$$\int_\Omega f \ dP = \int_\Omega f \ dQ \quad \dagger$$

con $f(\hat{x}) = f_1(x_1) \cdot f_2(x_2) \cdots f_k(x_k)$, $f_j(w) = ud(w, [a_j, b_j])$ como en el lema de Urysohn.

Sea $u \in \mathbb{N}$, dado $0 < \epsilon < 1$, elejimos r tan grande que: (1) f_j se anule fuera de $[-r, r]$ y (2) el cubo(I_r) de lado r en $\mathbb{R}^k$ sea tal que su complemento tenga probabilidad P y Q menor que ϵ; Dado que $f_j(-r) = f_j(r) = 0$ entonces podemos aplicarle el teorema de Stone-Weirstrass, para aproximar a f_j uniformemente en $[-r, r]$ por una suma finita de la forma

$$c_1 \exp(i\pi w/r) + c_2 \exp(2i\pi w/r) + \cdots + c_{n_j} \exp\left(n_j i\pi w/r\right)$$

de periodo $2r$, de tal suerte que multiplicando estas expresiones barriendo los valores de j, podemos aproximar a f por una función suma trigonométrica finita g, en donde cada sumando tiene la forma :

$$d_l \exp\left(m_1^l i\pi x_1/r + m_2^l i\pi x_2/r + \cdots + m_k^l i\pi x_k/r\right)$$

† ver lema de Urysohon

o

$$d_l \exp\left(i(x_1, x_2, \ldots, x_k) \cdot (t_1, t_2, \ldots, t_n)\right)$$

donde $m_1^l, m_2^l, \ldots, m_k^l \in \{1, \ldots, k\}$ es decir g tiene la forma

$$\sum_{l=1}^{L} d_l \exp\left(i\hat{x} \cdot \hat{t}_l\right)$$

de periodo $2r$ en cada variable, podemos elegir g de tal forma que $| f(\hat{x}) - g(\hat{x}) | < \epsilon$ dentro del cubo de lado r. Ya que f es acotada(por 1) en el cubo, g también(por $1 + \epsilon$), y por periodicidad en todo $\mathbb{R}^k$. Por tanto $| f - g |$ es acotada dentro del cubo por ϵ y fuera por $2 + \epsilon$, por tanto

$$\int_{\mathbb{R}^k} | f - g | \; dP < \epsilon + (2 + \epsilon) P(I_r^c) < 4\epsilon$$

análogamente

$$\int_{\mathbb{R}^k} | f - g | \; dP < 4\epsilon$$

Por tanto

$$\begin{aligned}
\left| \int_{\mathbb{R}^k} f \; dP - \int_{\mathbb{R}^k} f \; dQ \right| &\leq \left| \int_{\mathbb{R}^k} f \; dP - \int_{\mathbb{R}^k} g \; dP \right| + \left| \int_{\mathbb{R}^k} g \; dP - \int_{\mathbb{R}^k} f \; dQ \right| \\
&\leq 4\epsilon + \left| \int_{\mathbb{R}^k} g \; dP - \int_{\mathbb{R}^k} g \; dQ \right| + \left| \int_{\mathbb{R}^k} g \; dQ - \int_{\mathbb{R}^k} f \; dQ \right| \\
&\leq 4\epsilon + \left| \int_{\mathbb{R}^k} g \; dP - \int_{\mathbb{R}^k} g \; dQ \right| + 4\epsilon \\
&\leq 8\epsilon + \left| \int_{\mathbb{R}^k} g \; dP - \int_{\mathbb{R}^k} g \; dQ \right|
\end{aligned}$$

Pero aplicando la hipótesis, el segundo sumando de la última desigualdad, es cero, por tanto se tiene lo que bastaba probar.

11. TEOREMA DE DONSKER

Este teorema es muy importante ya que es una generalización de algunos teoremas anteriores. Aparte es muy clásico en la teoría de convergencia débil para procesos estocásticos. Lo pongo aquí porque creo que es muy conveniente tener un panorama general de los teoremas de convergencía débil, y porque es una buena aplicación del teorema 8.2. El teorema de Donsker es conocido como el teorema de límite central funcional. También incluyo su demostración.

Sea $X_t: C \to \mathbb{R}$ definida como $X_t(x) = x(t)$. Dado que para cada t la función X_t es medible, (es la proyección sobre C) la función $X: \Omega = [\,0,1\,] \to C$, $X(t) = X_t$ es un proceso estocástico.

Aceptemos que existe una medida de probabilidad W sobre $(C, \mathcal{C})$ tal que para cada $t \in (0,1]$

$$W\{x \in C | X_t(x) \leq \alpha\} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\alpha} \exp\left(-u^2/2t\right)\, du$$

si $t = 0$ interpretamos esto como $W\{x_0\} = 0$, y que además satisface que : el proceso estocástico definido arriba con

$$0 \leq t_0 \leq t_1 \leq \cdots \leq t_n \leq 1$$

cumple con que las variables aleatorias

$$X(t_1) - X(t_0), X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$$

son independientes respecto de W. A tal medida le llamamos *medida de Wiener.*

Teorema 11.1 Donsker

Sea X_n definidas como en (8.6), y supongamos además que las variables aleatorias ξ_i son independientes e identicamente distribuidas con media cero y varianza finita σ^2.

Entonces

$$X_n \xrightarrow{D} M \ .$$

Donde M es un proceso estocástico con distribución W .

Observación 11.2

Antes de probar el teorema notemos que esto es realmente una generalización del teorema de Linderberg-Leví o teorema central de limite en $\mathbb{R}$.

En efecto:

Dado que $P_n \Rightarrow W$ con P_n distribución de X_n entonces las distribuciones finito-dimensionales también convergen, en particular :

$$P_n \circ \pi_1^{-1} \Rightarrow W \circ \pi_1^{-1} .$$

Ahora si $A = (-\infty, \alpha]$ entonnces

$$\begin{aligned} P_n \circ \pi_1^{-1}(A) &= \mathrm{P}\left\{X_n \in \pi_1^{-1}(A)\right\} \\ &= \mathrm{P}\left\{\pi_1(X_n) \in A\right\} \\ &= \mathrm{P}\left\{\omega \in \Omega |\ \pi_1(X_n(\omega)) \in A\right\} \\ &= \mathrm{P}\left\{\omega \in \Omega |\ (X_n(1,\omega)) \in A\right\} \\ &= \mathrm{P}\left\{\omega \in \Omega |\ (X_n(\frac{n}{n},\omega)) \in A\right\} \\ &= \mathrm{P}\left\{\omega \in \Omega |\ \frac{1}{\sigma\sqrt{n}} S_n(\omega) \in A\right\} \\ &= \mathrm{P}\left\{\frac{S_n}{\sigma\sqrt{n}} \in A\right\} . \end{aligned}$$

Por tanto $P_n \circ \pi_1^{-1} =$ distribución de $\frac{S_n}{\sigma\sqrt{n}}$. Por otra parte

$$\begin{aligned} W \circ \pi_1^{-1}(A) &= W\{x \in C|\ x \in \pi_1^{-1}(A)\} \\ &= W\{x \in C|\ \pi_1(x) \in A\} \\ &= W\{x \in C|\ x(1) \in A\} \\ &= W\{x \in C|\ x(1) \leq \alpha\} \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-u^2/2}\, du \\ &= \mathrm{P}\{N \in A\} . \end{aligned}$$

Por tanto $W \circ \pi_1^{-1} =$ distribución de N. Es decir :

$$\frac{1}{\sigma\sqrt{n}}(\xi_1 + \xi_2 + \cdots + \xi_n) \xrightarrow{D} N.$$

Demostración (teorema de Donsker.)

Mostraremos que las distribuciones finito dimensionales de X_n convergen a alguna de M y después mostraremos que $\{X_n\}$ es tensa y finalmente una aplcación del teorema 7.1 terminará la prueba.

Sea $s \in [\,0,1\,]$. Debemos probar que

$$X_n(s) \xrightarrow{D} M(s);$$

M es una función de Ω en C, de la definición de X_n se tiene que :

$$|X_n(s,\omega) - \frac{1}{\sigma\sqrt{n}}S_{[ns]}(\omega)| \leq (ns - [ns])\frac{1}{\sigma\sqrt{n}}\xi_{[ns]+1}(\omega)$$
$$\leq \frac{1}{\sigma\sqrt{n}}\xi_{[ns]+1}(\omega) \ .$$

Usando la desigualdad de Chevychev :

$$\mathrm{P}\left\{|\frac{1}{\sigma\sqrt{n}}\xi_{[ns]+1}| \geq \epsilon\right\} \leq \frac{1}{\epsilon^2}E(\frac{1}{\sigma\sqrt{n}}\xi^2_{[ns]+1})$$
$$= \frac{1}{\epsilon^2\sigma\sqrt{n}}E(\xi^2_{[ns]+1})$$
$$= \frac{\sigma}{\epsilon^2\sqrt{n}} \rightarrow 0 \ .$$

Por tanto $\frac{1}{\sigma\sqrt{n}}\xi_{[ns]+1} \xrightarrow{\mathrm{P}} 0$, entonces

$$|X_n(s) - \frac{1}{\sigma\sqrt{n}}S_{[ns]}| \xrightarrow{\mathrm{P}} 0$$

bastaria ver en virtud del teorema 4.6 que

$$\frac{1}{\sigma\sqrt{n}}S_{[ns]} \xrightarrow{D} M_s .$$

(M_s tiene media cero y varianza s.) Nótese que

$$\frac{[ns]}{n} \rightarrow s,$$

como

$$\frac{1}{\sigma\sqrt{n}}S_{[ns]} = \sqrt{\frac{[ns]}{n}}\frac{1}{\sigma\sqrt{[ns]}}S_{[ns]}$$

entonces por el teorema de Linderberg-Leví se tiene que

$$\frac{1}{\sigma\sqrt{[ns]}}S_{[ns]} \xrightarrow{D} N$$

y por la observacion anterior tenemos :

$$\sqrt{\frac{[ns]}{n}}\frac{1}{\sigma\sqrt{[ns]}}S_{[ns]} \xrightarrow{D} \sqrt{s}N ,$$

como $\sqrt{s}N$ es una variable aleatoria normalmente distribuida con media cero y varianza s entonces

$$\frac{1}{\sigma\sqrt{n}}S_{[ns]} \xrightarrow{D} M_s$$

Sean ahora $s, t \in [\,0,1\,]$ supongamos $s < t$, tenemos que probar entonces que

$$((X_n(s), X_n(t)) \xrightarrow{D} (M_s, M_t).$$

Recordemos que $(M_s, M_t): \Omega \times \Omega \to \mathbb{R} \times \mathbb{R}$. Si consideramos $h: \mathbb{R}^2 \to \mathbb{R}^2$ definida como $h(x,y) = (x, x+y)$, entonces h es continua, por tanto si D_h es el conjunto de discontinuidades de h entonces $(M_s, M_t)^{-1}(D_h) = \emptyset$. Así que si probamos que

$$(X_n(s), X_n(t) - X_n(s)) \xrightarrow{D} (M_s, M_t - M_s)$$

Por el corolario 5.2 se tendrá que

$$h((X_n(s), X_n(t) - X_n(s)) \xrightarrow{D} h(M_s, M_t - M_s)$$

es decir

$$((X_n(s), X_n(t)) \xrightarrow{D} (M_s, M_t).$$

Nótese que las variables aleatorias $\frac{1}{\sigma\sqrt{n}} S_{[ns]}$, $\frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]})$ son independientes, además por el caso anterior tenemos que

$$\frac{1}{\sigma\sqrt{n}} S_{[ns]} \xrightarrow{D} M_s \quad \text{y} \quad \frac{1}{\sigma\sqrt{n}} S_{[nt]} \xrightarrow{D} M_t$$

por tanto

$$\frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]}) \xrightarrow{D} M_t - M_s.$$

En virtud de que $\frac{1}{\sigma\sqrt{n}} S_{[ns]}$ y $\frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]})$ son independientes, y de que $\mathbb{R}^2$ es separable, por el teorema 3.2 se tiene que :

$$\left(\frac{1}{\sigma\sqrt{n}} S_{[ns]}\,,\ \frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]})\right) \xrightarrow{D} (M_s, M_t - M_s),$$

ya que la medida $(M_s, M_t - M_s)$ es la medida producto y $M_s, M_t - M_s$ tienen incrementos independientes.

Dado que

$$\left\|((X_n(s), X_n(t) - X_n(s)) - \left(\frac{1}{\sigma\sqrt{n}} S_{[ns]}, \frac{1}{\sigma\sqrt{n}}(S_{[nt]} - S_{[ns]})\right)\right\| \xrightarrow{P} 0$$

entonces nuevamente por el teorema 4.6, tenemos el resultado.

En el caso de 3 o mas puntos la demostración de hace de manera análoga.

Ahora probaremos la tensión de la familia $\{X_n\}$ empleando un lema que nos servirá de ayuda.

Lema 11.8

Sean $\xi_1, \xi_2, \ldots$ variables aleatorias, independientes, con media cero y con varianzas finitas $\sigma_i^2 > 0$, sean $S_i = \xi_1 + \xi_2 + \cdots + \xi_i$, $s_i^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_i^2$. Entonces

$$\mathrm{P}\left\{\max_{i\leq m}|S_i| \geq \lambda s_m\right\} \leq 2\mathrm{P}\left\{|S_m| \geq (\lambda - \sqrt{2})s_m\right\}$$

En efecto : Consideremos a los conjuntos

$$E_i = \left\{\max_{j<i}|S_j| < \lambda s_m \leq |S_i|\right\}$$

Mostremos que

$$\left\{\max_{i\leq m}|S_i| \geq \lambda s_m\right\} \subseteq \left\{|S_m| \geq (\lambda - \sqrt{2})s_m\right\} \bigcup_{i=1}^{m-1} \left(E_i \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}\right)$$

Sea ω en el primer conjunto. Si $\omega \in \left\{|S_m| \geq (\lambda - \sqrt{2})s_m\right\}$ entonces ya terminamos. Si no, entonces $\omega \in \left\{|S_m| < (\lambda - \sqrt{2})s_m\right\}$. Sea j_0 el menor índice, menor o igual que m tal que

$$|S_{j_0}| \geq \lambda s_m$$

es claro que $j_0 \leq m$, y además $|S_i| < \lambda s_m$, $\forall i \leq (j_0 - 1)$, en particular

$$\max_{j\leq j_0-1}|S_j| < \lambda s_m \leq |S_{j_0}|$$

por tanto $\omega \in E_{j_0-1}$, lo cual prueba la contención.

Por tanto

$$\begin{aligned}\mathrm{P}\left\{\max_{i\leq m}|S_i| \geq \lambda s_m\right\} &\leq \mathrm{P}\left\{|S_m| \geq (\lambda - \sqrt{2})s_m\right\} \\ &\quad + \sum_{i=1}^{m-1} \mathrm{P}\left(E_i \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}\right)\end{aligned}$$

Además

$$E_i \cap \left\{|S_m| < (\lambda - \sqrt{2})s_m\right\} \subseteq E_i \cap \left\{|S_m - S_i| \geq \sqrt{2}s_m\right\}$$

Ya que si tomamos ω en el conjunto del lado izquierdo, $\omega \in E_i$,

pero si $|S_m(\omega) - S_i(\omega)| < \sqrt{2}s_m$ entonces :

$$\begin{aligned}|S_m(\omega) - S_i(\omega)| \geq \sqrt{2}s_m &\geq |S_i(\omega)| - |S_m(\omega)| \\ &\geq \lambda s_m - (\lambda - \sqrt{2})s_m = \sqrt{2}s_m\end{aligned}$$

lo cual es un absurdo.

Nótese que ambos conjuntos E_i, $\left\{|S_m - S_i| \geq \sqrt{2}s_m\right\}$, son independientes.

Se tiene pues que :

$$\sum_{i=1}^{m-1} \mathrm{P}\left(E_i \cap \{|S_m| < (\lambda - \sqrt{2})s_m\}\right) \leq \sum_{i=1}^{m-1} \mathrm{P}\left(E_i \cap \left\{|S_m - S_i| \geq \sqrt{2}s_m\right\}\right)$$

$$= \sum_{i=1}^{m-1} \mathrm{P}(E_i)\mathrm{P}\left\{|S_m - S_i| \geq \sqrt{2}s_m\right\} \qquad \text{por independencia,}$$

$$= \sum_{i=1}^{m-1} \mathrm{P}(E_i)\mathrm{P}\left\{|S_m - S_i|^2 \geq 2s_m^2\right\}$$

$$= \sum_{i=1}^{m-1} \mathrm{P}(E_i)\frac{1}{2s_m^2}E\left\{|S_m - S_i|^2 \geq 2s_m^2\right\} \qquad \text{por Chebychev,}$$

$$\leq \sum_{i=1}^{m-1} \mathrm{P}(E_i)\frac{1}{2s_m^2}\left(E(\xi_{i+1} + \xi_{i+2} + \cdots + \xi_m)^2\right)$$

$$\leq \sum_{i=1}^{m-1} \mathrm{P}(E_i)\frac{1}{2s_m^2}\left(\sum_{k=i+1}^{m} E(\xi_k^2)\right) \qquad \text{por independencia,}$$

$$\leq \sum_{i=1}^{m-1} \mathrm{P}(E_i)\frac{1}{2s_m^2}\left(\sum_{k=i+1}^{m} \sigma_k^2\right)$$

$$\leq \frac{1}{2}\sum_{i=1}^{m-1} \mathrm{P}(E_i) \leq \frac{1}{2}\mathrm{P}\left\{\max_{i\leq m}|S_i| \geq \lambda s_m\right\}$$

Por tanto

$$\mathrm{P}\left\{\max_{i\leq m}|S_i| \geq \lambda s_m\right\} \leq \mathrm{P}\left\{|S_m| \geq (\lambda - \sqrt{2})s_m\right\} + \frac{1}{2}\mathrm{P}\left\{\max_{i\leq m}|S_i| \geq \lambda s_m\right\}$$

lo cual prueba el lema.

Aplicando el lema a las variables aleatorias ξ_i y con $\lambda > 2\sqrt{2}$ se tiene que

$$\mathrm{P}\left\{\max_{i\leq n}|S_i| \geq \lambda\sigma\sqrt{n}\right\} \leq 2\mathrm{P}\left\{|S_n| \geq (\lambda - \sqrt{2})\sigma\sqrt{n}\right\}$$
$$\leq 2\mathrm{P}\left\{|S_n| \geq \frac{1}{2}\lambda\sigma\sqrt{n}\right\} \quad \text{pues } \lambda - \sqrt{2} \geq \frac{1}{2}\lambda \ .$$

Por el teorema de Linderberg-Leví se tiene que

$$\mathrm{P}\left\{|S_n| \geq \frac{1}{2}\lambda\sigma\sqrt{n}\right\} \to \mathrm{P}\left\{|N| \geq \frac{\lambda}{2}\right\} \leq \frac{8}{\lambda^3}E\{|N^3| \quad \text{si } n \to \infty \ .$$

Esta última desigualdad es debida a Chebychev. Por tanto $\exists N_0 \in \mathbb{N}$ tal que $n \geq N_0$ entonces

$$\mathrm{P}\left\{\max_{i\leq n}|S_i| \geq \lambda\sigma\sqrt{n}\right\} \leq \frac{8}{\lambda^3}E\{|N^3|\} \ .$$

Eligiendo λ tal que $8E\{|N^3|\} < \lambda\epsilon$ tenemos que : $\forall\epsilon > 0$, $\exists\ \lambda > 1$, $\exists\ N \in \mathbb{N}$ tal que

$$P\left\{\max_{i\leq n}|S_i| \geq \lambda\sigma\sqrt{n}\right\} \leq \frac{\epsilon}{\lambda^2} \qquad \forall n \geq N. \tag{11.1}$$

Probaremos que se satisface la hipótesis del teorema 8.8 con lo que se tendrá que la sucesión es tensa.

Observemos que

$|S_{k+1} - S_k| = |\xi_{k+1}|$ tiene la misma distribución que $|\xi_1|$

$|S_{k+2} - S_k| = |\xi_{k+1} + \xi_{k+2}|$ tiene la misma distribución que $|\xi_1 + \xi_2|$

$|S_{k+n} - S_k| = |\xi_{k+1} + \cdots + \xi_{k+n}|$ tiene la misma distribución que $|\xi_1 + \cdots + \xi_n|$

De lo cual se deduce que

$$P\left\{\max_{i\leq n}|S_{i+k}(\omega) - S_k(\omega)| \geq \lambda\sigma\sqrt{n}\right\} = P\left\{\max_{i\leq n}|S_i| \geq \lambda\sigma\sqrt{n}\right\}$$

Por tanto (11.1) implica que se satisface la hipótesis del teorema 8.4.

12. TEOREMA CENTRAL DE LIMITE EN C[0,1]

Dentro de este capítulo se dan condiciones bajo las cuales una sucesión de procesos satisface el *teorema central de limite.*, lo cual se define al comienzo del capítulo.

Algo que también cabe destacar es que, bajo las hipótesis del teorema central de límite en $\mathbb{R}$, el hecho de que las variables aleatorias, $\xi_1, \xi_2, \ldots, \xi_n$ satisfagan la conclusion de tal teorema, no es mas que un caso particular, de que una sucesión de procesos satisfaga el *teorema central de límite*. Lo cual es mostrado al final de este capítulo.

Sea C el espacio de funciones reales sobre $[0,1]$, métrico, completo y separable por ejemplo $C[0,1]$.

☞ *Definición 12.1*

Un proceso estocástico con valores en C es una variable aleatoria C-valuada.

☞ *Definición 12.2*

Se dice que un proceso estocástico es Gaussiano si sus distribuciones finito-dimensionales son normales.

☞ *Definición 12.3*

Sea $\{X_n\}$ una sucesión de variables aleatorias-C-valuadas definidas en el mismo espacio $(\Omega, \mathcal{B}, \mathrm{P})$ tales que :

(i) son independientes e identicamente distribuidas, con la misma distribución que X.

(ii) $E(X(t)) = 0,\ E(X^2(t)) < \infty \ \forall t \in [0,1]$.

Si $\exists Z$ proceso Gaussiano con trayectorias en C tal que

$$Z_n = \left(\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} \right) \xrightarrow{D} Z$$

decimos entonces que $\{X_n\}$ *satisface el teorema central de límite en* C. lo abreviamos diciendo que $\{X_n\}$ satisface el T.C.L.

Observación 12.4

Supóngase que $\{X_n\}$ satisface el T.C.L. Sea $\{Y_n\}$ una sucesión con las propiedades (i) (ii) de la definición 12.3. Entonces $\{Y_n\}$ satisface también el T.C.L.

Demostración

Si P_n es la distribución de $Z_n = \frac{X_1+X_2+\cdots+X_n}{\sqrt{n}}$ y $\hat{P}_n$ la de $\hat{Z}_n = \frac{Y_1+Y_2+\cdots+Y_n}{\sqrt{n}}$ bastaría ver que $\hat{P}_n = P_n$ ya que $P_n \Rightarrow P$. (donde P distribución de Z).

Dado que $Z, \hat{Z}$ tienen las mismas distribuciones finito-dimensionales (por propiedades de convolución†), y ya que estas determinan de manera única a cada medida se tiene que $P_n = \hat{P}_n$.

Definición 12.5

Decimos que X (variable aleatoria C-valuada) satisface el T.C.L. si existe una sucesión $\{X_n\}_{n\in\mathbb{N}}$ de variables aleatorias (C-valuadas) con la misma distribución que X, que satisfaga el T.C.L, es decir si $\exists Z$ proceso Gaussiano tal que

$$Z_n = \left(\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}}\right) \xrightarrow{D} Z$$

Definición 12.6

Decimos que una sucesión de procesos estocásticos X_n es uniformemente equicontinua en probabilidad si : $\forall\epsilon > 0, \forall\eta > 0,\ \exists\ \delta > 0$ tal que :

$$\mathrm{P}\{\omega \in \Omega |\ w(X_n(\omega, \delta)) \geq \epsilon\} \leq \eta \qquad \forall n.$$

Observación 12.7

Sea X un proceso continuo tal que :

(i) existe un proceso Gaussiano Z con trayectorias en C y con la misma covarianza que X.

(ii) Las sumas normalizadas $\{Z_n\}$ son uniformemente equicontinuas en probabilidad.

Entonces X satisface el T.C.L. en C.

Demostración

Como $\{Z_n\}$ son uniformemente equicontinuas en probabilidad tenemos que $\forall\epsilon > 0,\ \forall\eta > 0,\ \exists\ \delta$ con $0 < \delta < 1$ tal que:

$$\mathrm{P}\{\omega \in \Omega |\ w(Z_n(\omega), \delta) \geq \epsilon\} \leq \eta \quad \forall n.$$

---

†ver notas de M.E.Caballero y B. Fernandez

Mostraremos que $\forall\eta > 0$, $\exists\ a \in \mathbb{R}$ tal que

$$P\{|Z_n(0)| > a\} \leq \eta \quad , \quad \forall n \in \mathbb{N}$$

y en virtud del teorema 6.2 se tendrá que la sucesión $\{Z_n\}$ es tensa.

Por el T.C.L en $\mathbb{R}$, para $\sigma^2 = E(X^2(t))$ se tiene que :

$$\frac{1}{\sigma^2} Z_n(0) \xrightarrow{D} N(0,\sigma) \ .$$

Ahora dado que

$$P\left\{|\frac{1}{\sigma^2} Z_n(0)| > a\right\} = P\{Z_n(0) \in B_a\},$$

donde $B_a = (-\infty, -a\sigma^2) \cup (a\sigma^2, \infty)$. Eligiendo a tal que $P\{N \in B_a\} = 0$ obtenemos por el teorema 2.10

$$P\{Z_n(0) \in B_a\} \to P\{Z(0) \in B_a\} = P\{N \in B_a\} \quad .$$

Ahora, es claro que $\forall\eta > 0$, $\exists\ a, P\{N \in B_a\} < \frac{\eta}{2}$ y $\exists N_0$ tal que si $n \geq n_0$ entonces

$$|P\{Z_n(0) \in B_a\} - P\{N \in B_a\}| < \frac{\eta}{2}.$$

Por tanto

$$P\{Z_n(0) \in B_a\} \leq P\{N \in B_a\} + (P\{Z_n(0) \in B_a\} - P\{N \in B_a\}) < \eta.$$

Ahora solo bastaría ver que las distribuciones finito dimensionales de $\{Z_n\}$ convergen a las distribuciones finito-dimensionales de Z, donde Z es el proceso Gaussiano mencionado, para que en virtud del teorema 7.2 se tenga que $Z_n \xrightarrow{D} Z$.

Nótese que la distribución finito dimensional correspondiente a t del proceso Z (por ser Gaussiano) es normalmente distribuida con media 0 y varianza $E(X^2(t))$.

Tomemos un $t \in [0,1]$. La distribución finito dimensional correspondiente a t es la distribución de $Z_n(t)$. Nótese que las variables aleatorias $X_1(t), X_2(t), \ldots, X_n(t)$ son independientes e idénticamente distribuidas con media cero y varianza $E(X^2(t))$ la cual es finita. Supongámosla positiva. Entonces por el teorema central de límite en $\mathbb{R}$ aplicado a estas variables, se tiene que

$$\frac{1}{\sqrt{E(X^2(t))}\sqrt{n}} \sum_{i=1}^{n} X_i(t) \xrightarrow{D} N$$

Por tanto

$$\frac{X_1(t) + X_2(t) + \cdots + X_n(t)}{\sqrt{n}} = Z_n(t) \xrightarrow{D} \sqrt{E(X^2(t))}N$$

Esto de debe al corolario 5.2:

$$X_n \xrightarrow{D} X \Rightarrow h(X_n) \xrightarrow{D} h(X), \qquad P\{D_h\} = 0, \quad h(s) = \sqrt{E(X^2(t))}s.$$

Pero $\sqrt{E(X^2(t))}N$ es una variable aleatoria con distribución normal con media cero y varianza $E(X^2(t))$ ya que N es una variable aleatoria con media cero y varianza 1.

Dado que Z es un proceso Gaussiano con la misma covarianza que X (por (1)) entonces Z es una variable aleatoria normalemente distribuida con media cero y varianza $E(X^2(t))$ es decir :

$$Z_n(t) \xrightarrow{D} Z(t).$$

Para $(t_1, t_2, \cdots, t_k) \in \mathbb{R}^k$ la demostración se reduce al teorema central de límite en $\mathbb{R}^k$ (teorema 10.15).

Observación 12.8

Sean $\xi_1, \xi_2, \ldots, \xi_n$ v.a.i.i.d. con media cero y varianza σ^2, como en el teorema de Linderberg-Leví.

Demostración

Consideremos la sucesión de procesos estocásticos en C definidos de la siguiente manera, $X_n: \Omega \to C$,

$$X_n(\omega): [\,0,1\,] \to \mathbb{R}, \qquad (X_n(\omega))(t) \equiv \xi_n(\omega)$$

es decir la función constante $\xi_n(\omega)$. Nótese que X_n toma sus valores en C, y que además efectivamente es un proceso estocástico, pues es constante y por tanto medible.

Para ver que $\{X_n\}$ satisface el T.C.L, hay que checar (i),(ii) de la definición.

(i) $X_1, X_2, \ldots, X_n, \ldots,$ son independientes e idénticamente distribuidas.

Sean X_m, X_n, dos procesos antes construidos, y sean $A, B \in C$. Consideremos

$$\hat{A} = \{s \in \mathbb{R} \mid \exists\, x \in A,\ x(t) = s \quad \forall t \in [\,0,1\,]\}$$

$$\hat{B} = \{s \in \mathbb{R} \mid \exists\, x \in B,\ x(t) = s \quad \forall t \in [\,0,1\,]\}$$

y consideremos

$$U_1 = \{\omega \in \Omega, \mid X_n(\omega) \in A\}, \qquad U_2 = \{\omega \in \Omega, \mid X_m(\omega) \in B\}$$

$$W_1 = \{\omega \in \Omega, \mid \xi_n(\omega) \in \hat{A}\}, \qquad W_2 = \{\omega \in \Omega, \mid \xi_m(\omega) \in \hat{B}\}$$

Probemos que $U_1 = W_1$.

$\subseteq$) Sea $\omega \in U_1$, entonces $X_n(\omega) \in A$; ya que $(X_n(\omega))(t) \equiv \xi_n(\omega)$ se tiene que $\exists x \in A$, $x(t) = s \quad \forall t \in [\,0,1\,]$ (a saber $x \equiv X_n(\omega)$), por tanto $\xi_n(\omega) \in \hat{A}$, es decir, $\omega \in W_1$.

$\supseteq$) Dada $\omega \in W_1$, tenemos que $\xi_n(\omega) \in \hat{A}$, entonces $\exists x \in A$, $x(t) = s \quad \forall t \in [\,0,1\,]$, y por definición de $X_n(\omega)$ se tiene que $X_n(\omega) = \xi_n(\omega)$, en particular $X_n(\omega) \in A$, es decir $\omega \in U_1$.

Análogamente se tiene que $U_2 = W_2$.

Por tanto

$$U_1 \bigcap U_2 = W_1 \bigcap W_2.$$

De lo cual se deduce que

$$\begin{aligned} P\{X_n \in A, X_m \in B\} &= P(U_1 \bigcap U_2) \\ &= P(W_1 \bigcap W_2) \\ &= P(W_1)P(W_2) \quad \text{por independencia de las } \xi_k \\ &= P(U_1)P(U_2) \\ &= P\{X_n \in A\}P\{X_m \in B\} \end{aligned}$$

Por otra parte es claro que los procesos sean igualmente distribuidos, y que tengan la misma varianza.

(ii) Construyamos un proceso Z con trayectorias en C.

$$Z(\omega, t) = \sigma N,$$

donde N es la normal 0,1. Sabemos que efectivamente Z tiene trayectorias en C. y además

$$\frac{X_1 + X_2 + \cdots + X_n}{\sqrt{n}} = \frac{\xi_1 + \xi_2 + \cdots + \xi_n}{\sqrt{n}} \xrightarrow{D} \sigma N$$

Por tanto hemos checado (i),(ii) de 12.3, es decir la sucesión $X_1, X_2, \ldots, X_n, \ldots,$ satisfacel el T.C.L.

13. CONDICIONES PARA EL TEOREMA CENTRAL DE LIMITE EN C[0,1]

En este capítulo se desarrollarán algunos resultados de M. Hann, ayudándose de los capítulos anteriores.

☞ *Definición 13.1*

Decimos que X es *separable* si $\exists S \subseteq [0,1]$, S denso y numerable tal que para cualquier B intervalo cerrado, y para cualquier A intervalo abierto de $[0,1]$ se tiene que

$$\{\omega \in \Omega | X_t(\omega) \in B,\ \forall t \in A\} = \{\omega \in \Omega | X_t(\omega) \in B,\ \forall t \in A \bigcap S\} \quad .$$

☞ *Definición 13.2*

Decimos que X es continuo si $\forall \omega \in \Omega$ la función (trayectoria) :

$$X(\cdot,\omega)\colon [0,1] \to \mathbb{R}$$

es continua.

Tambien en este caso diremos que X tiene trayectorias continuas.

☞ *Definición 13.3*

Decimos que X es estocásticamente continuo en un punto $t_0 \in [0,1]$, si $\forall \epsilon > 0$

$$\lim_{t \to t_0} \mathrm{P}\{\omega \in \Omega |\ |X(t,\omega) - X(t_0,\omega)| \geq \epsilon\} = 0 \ ,$$

y diremos que X es estocásticamente continuo si es estocásticamente continuo en cualquier $t \in [0,1]$.

☞ *Definición 13.4*

Dados 2 procesos X, Y decimos que son equivalentes si para cualquier t

$$P\{\omega \in \Omega | X(t,\omega) = Y(t,\omega)\} = 1.$$

Diremos tambien que Y es una versión de X o una modificación de X.

☞ *Definición 13.5*

Sea

$$D = \{f : [\,0,1\,] \to \mathbb{R} \mid f \text{ es CADLAD }\}$$

donde CADLAD significa que f es continua por la derecha y con límites por la izquierda.

Se sabe que tal espacio con la métrica de Prohorov es completo, y separable.† (esto no se probará aquí).

☞ *Definición 13.6*

X es continuo por trayectorias si $\exists Y$ proceso continuo tal que X es equivalente a Y i,e X tiene una versión continua.

Dada X variable aleatoria D-valuada, se define $A_q^X : \Omega \to \mathbb{R}$ como

$$A_q^X(\omega) = \sup_{0 \leq s < 2^{q-1}} |X(\frac{s+1}{2^{q-1}}, \omega) - X(\frac{s}{2^{q-1}}, \omega)|. \tag{I}$$

Para un ejemplo con $q = 4$, vease figurita uno, donde $A_3^X(\omega) = \max\{a_0, a_1, \ldots, a_7\}$.

figurita uno

En el transcurso de los siguientes teoremas $\tilde{X}$ denotará una versión separable de X y $\|\cdot\|_r$ denotará la norma usual en $L^r(\Omega, P)$.

Teorema 13.7

Sea $\phi(h)$ una funcion no negativa sobre $[\,0,1\,]$ creciente para h suficientemente pequeño y tal que $\lim_{h\to 0} \phi(h) = 0$. Sea X un proceso estocásticamente continuo que

† ver T.Kurtz

satisface :

$$\sum_{q=1}^{\infty} [\phi(\frac{1}{2^{q+1}})]^{-1} ||A_q^X||_r < \infty \tag{II}$$

para algún $r \in \mathbb{N}$..

Entonces $\exists A^X \in L^r(\Omega, P)$ tal que

$$|\tilde{X}(t,\omega) - \tilde{X}(s,\omega)| \leq A^X(\omega)\phi(|t-s|)$$

para casi toda $\omega \in \Omega$ y $|t-s|$ suficientemente pqequeño. Y ademas

$$A^X(\omega) = 3\sum_{q=1}^{\infty} [\phi(\frac{1}{2^{q+1}})]^{-1} A_q^X(\omega). \tag{III}$$

La demostración de este teorema no se hará aquí, por la razón principal de que desviaría la atención del tema, además por ser demasiado técnica, ya que la parte escencial de este capítulo es el desarrollo de los teoremas del artículos de M.Hann. *

Corolario 13.8

Sea $f: \mathbb{R} \to \mathbb{R}$ tal que $\lim_{h\to\infty} f(h) = 0$. Sea X un proceso estocástico tal que

$$E|X(t) - X(s)|^r \leq f(|t-s|) \tag{IV}$$

para algún $r \geq 1$ y $\forall s, t \in [\,0,1\,]$. Supóngase además que $\exists \phi: [0,1] \to \mathbb{R}$ no negativa, creciente y tal que:

$$\sum_{q=1}^{\infty} [\phi(\frac{1}{2^{q+1}})]^{-1} 2^{\frac{q}{r}} \left(f(\frac{1}{2^{q-1}})\right)^{\frac{1}{r}} < \infty \tag{V}$$

Entonces X es continuo por trayectorias y existe $A^X \in L^r(\Omega, P)$ variable aleatoria tal que

$$|\tilde{X}(t,\omega) - \tilde{X}(s,\omega)| \leq A^X(\omega)\phi(|t-s|) \quad \text{para casi toda } \omega. \tag{VI}$$

Ademas $(||A_q^X||_r)^r$ esta acotada por una constante multiplicada por la suma dada en (V).

Demostración

Observación 13.9.- X es estocásticamente continuo pues si $t_0 \in [0,1]$ entonces

$$\begin{aligned} P\{\omega\in\Omega|\ |X(t,\omega) - X(t_0,\omega)| \geq \epsilon\} &= P\{\ |X(t) - X(t_0)|^r \geq \epsilon^r\ \} \\ &\leq \frac{1}{\epsilon^r} E|X(t) - X(s)|^r \\ &\leq \frac{f(|t-s|)}{\epsilon^r} \\ &\to 0 \end{aligned}$$

---

*para tal demostración ver Delporte.

Calculemos

$$(||A_q^X||_r)^r = \int_\Omega \left(\sup_{0\le s<2^{q-1}} |X(\frac{s+1}{2^{q-1}},\omega) - X(\frac{s}{2^{q-1}},\omega)| \right)^r dP$$

$$= \int_\Omega \left(\sup_{0\le s<2^{q-1}} |X(\frac{s+1}{2^{q-1}},\omega) - X(\frac{s}{2^{q-1}},\omega)|^r \right) dP$$

$$\le \int_\Omega \sum_{s=0}^{(2^{q-1}-1)} |X(\frac{s+1}{2^{q-1}},\omega) - X(\frac{s}{2^{q-1}},\omega)|^r \, dP$$

$$= \sum_{s=0}^{(2^{q-1}-1)} E|X(\frac{s+1}{2^{q-1}},\omega) - X(\frac{s}{2^{q-1}},\omega)|^r$$

$$\le \sum_{s=0}^{(2^{q-1}-1)} f\left(\frac{1}{2^{q-1}}\right) \qquad \text{por (IV)}$$

$$= 2^{q-1} f\left(\frac{1}{2^{q-1}}\right)$$

$$\le 2^q f\left(\frac{1}{2^{q-1}}\right).$$

Así tenemos que

$$(||A_q^X||_r)^r \le 2^q f\left(\frac{1}{2^{q-1}}\right)$$

en consecuencía :

$$\sum_{q=1}^{\infty} \left[\phi\left(\frac{1}{2^{q+1}}\right)\right]^{-1} ||A_q^X||_r \le \sum_{q=1}^{\infty} \left[\phi\left(\frac{1}{2^{q+1}}\right)\right]^{-1} 2^{\frac{q}{r}} \left(f\left(\frac{1}{2^{q-1}}\right)\right)^{\frac{1}{r}} < \infty \qquad \text{por la hip. (V)}$$

Se tiene entonces que X y ϕ satisfacen las hipótesis del teorema 13.7, por tanto

$$|\tilde{X}(t,\omega) - \tilde{X}(s,\omega)| \le A^X(\omega)\phi(|t-s|)$$

para casi toda $\omega \in \Omega$ y $|t-s|$ suficientemente pqequeño, es decir $\exists \delta_1 > 0$ tal que si $|t-s| < \delta_1$ entonces

$$B = \left\{\omega \in \Omega | \; |\tilde{X}(t,\omega) - \tilde{X}(s,\omega)| \le A^X(\omega)\phi(|t-s|)\right\} \quad \text{tiene probabilidad 1 .}$$

Enseguida, mostraremos en efecto que X es continuo por trayectorias, exibiendo una versión continua.

Sea $Y: [0,1] \times \Omega \to \mathbb{R}$ definida como :

$$Y(t,\omega) = \begin{cases} \tilde{X}(t,\omega) & \text{si } \omega \in B \cap A_t, \\ 0 & \text{si no} \end{cases}$$

Con $A_t = \{\omega \in \Omega | \; \tilde{X}(t,\omega) = X(t,\omega)\}$.

Y es una versión de X pues para $t \in [0,1]$ y $\omega \in B \bigcap A_t$

$$Y(t,\omega) = \tilde{X}(t,\omega) = X(t,\omega)$$

y

$$P(B \bigcap A_t) = P(B) + P(A_t) - P(B \bigcup A_t) = 1 \ .$$

Ademas Y es continuo, dado que para $\omega \in \Omega$
si $\omega \in (B \bigcup A_t)^c$ entonces $Y(t,\omega) = 0 \quad \forall t$, por tanto $Y(\cdot,\omega)$ es continua.
si $\omega \in (B \bigcup A_t)$ entonces $Y(\cdot,\omega) = \tilde{X}(t,\omega)$, por tanto

$$|Y(t,\omega) - Y(s,\omega)| \leq A^X(\omega)\phi(|t-s|) \to 0, \quad \text{si } t \to s \ .$$

Así que X es continuo por trayectorias. Mas aún, por la definición (III) :

$$A^X(\omega) = 3\sum_{q=1}^{\infty} \left[\phi\left(\frac{1}{2^{q+1}}\right)\right]^{-1} A_q^X(\omega)$$

entonces

$$\begin{aligned} \|A^X\|_r &\leq 3\sum_{q=1}^{\infty} \left[\phi\left(\frac{1}{2^{q+1}}\right)\right]^{-1} \|A_q^X\| \\ &\leq 3\sum_{q=1}^{\infty} \left[\phi\left(\frac{1}{2^{q+1}}\right)\right]^{-1} 2^{\frac{q}{r}} \left(f\left(\frac{1}{2^{q-1}}\right)\right)^{\frac{1}{r}} . \end{aligned}$$

Lo cual prueba la afirmación del teorema.

Teorema 13.10

Sea f una función no negativa sobre $[0,1]$ la cual es creciente cerca del cero. Sea X un proceso estocástico tal que para algún $r \geq 1$

$$E|X(t) - X(s)|^r \leq f(|t-s|).$$

Si

$$(*) \qquad \int_0 y^{\frac{-(r+1)}{r}} (f(y))^{\frac{1}{r}} \, dy < \infty, \ \dagger$$

entonces $\exists \phi: [0,1] \to \mathbb{R}$ continua y creciente con $\phi(0) = 0$ la cual solo depende de f y $\exists\ A^X \in L^r(\Omega, P)$ tal que $|\tilde{X}(t) - \tilde{X}(s)| \leq A^X \phi(|t-s|)$.

Ademas $\|A_q^X\|_r{}^r$ esta acotada por una constante, la cual depende solamente de f y ϕ.

$\dagger \int_0 g(y)dy \stackrel{\text{def}}{=} \lim_{\delta \to 0} \int_\delta^t g(y)\, dy$

Demostración

Sea $\epsilon > 0$ tal que f es decreciente en $[0,\epsilon]$ y sea $\delta > 0$, entonces

$$\int_\delta^\epsilon u^{\frac{-(r+1)}{r}} (f(u))^{\frac{1}{r}}\, du = \int_{\frac{1}{\delta}}^{\frac{1}{\epsilon}} \left(\frac{1}{y}\right)^{\frac{-(r+1)}{r}} \left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}} \left(\frac{-1}{y^{-2}}\right) dy \quad \text{haciendo } u = \frac{1}{y}$$

$$= \int_{\frac{1}{\delta}}^{\frac{1}{\epsilon}} \left(y^{-1}\right)^{\frac{-(r+1)}{r}} \left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}} \left(\frac{-1}{y^{-2}}\right) dy$$

$$= \int_{\frac{1}{\epsilon}}^{\frac{1}{\delta}} y^{1+\frac{1}{r}} y^{-2} \left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}} dy$$

$$= \int_{\frac{1}{\epsilon}}^{\frac{1}{\delta}} y^{\frac{1}{r}-1} \left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}} dy$$

como por hipótesis

$$\lim_{\delta\to 0} \int_\delta^\epsilon u^{\frac{-(r+1)}{r}} (f(u))^{\frac{1}{r}}\, du < \infty$$

tenemos

$$\lim_{\delta\to 0} \int_{\frac{1}{\epsilon}}^{\frac{1}{\delta}} y^{\frac{1}{r}-1} \left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}} dy = \int_{\frac{1}{\epsilon}}^{\infty} y^{\frac{1}{r}-1} \left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}} dy < \infty \quad .$$

Sea k_0 tal que $\frac{1}{\epsilon} < 2^{k_0-2}$; Si $g(y) = y^{\frac{1}{r}-1}\left(f\left(\frac{1}{y}\right)\right)^{\frac{1}{r}}$ entonces g es decreciente en $[2^{k-2}, 2^{k-1}]$ $\forall k \geq k_0$.

En efecto, si $k \geq k_0$ entonces $2^{k_0-2} \leq 2^{k-2}$ ahora, dados $y_1, y_2 \in [2^{k-2}, 2^{k-1}]$ se tiene que

$$\frac{1}{\epsilon} \leq y_1 \leq y_2$$

y como $r \leq 1$ entonces $1 - \frac{1}{r} \geq 0$, por tanto

$$(y_1)^{1-\frac{1}{r}} \leq (y_2)^{1-\frac{1}{r}} \quad \Rightarrow \quad (y_2)^{\frac{1}{r}-1} \leq (y_2)^{\frac{1}{r}-1},$$

como $\left(f\left(\frac{1}{y_2}\right)\right)^{\frac{1}{r}} \leq \left(f\left(\frac{1}{y_1}\right)\right)^{\frac{1}{r}}$ se tiene que $g(y_2) \leq g(y_1)$.

Por tanto

$$\int_{2^{k-2}}^{2^{k-1}} g(y)\, dy \geq \int_{2^{k-2}}^{2^{k-1}} g(2^{k-1})\, dy = 2^{k-2} g(2^{k-1}) \quad . \qquad (*)$$

Ahora

$$
\begin{aligned}
\infty &> \int_{\frac{1}{\epsilon}}^{\infty} g(y)\, dy \\
&\geq \int_{2^{k_0-2}}^{\infty} g(y)\, dy \\
&\geq \sum_{k=k_0}^{\infty} \int_{2^{k-2}}^{2^{k-1}} g(y)\, dy \\
&\geq \sum_{k=k_0}^{\infty} 2^{k-2} g(2^{k-1}) \qquad \text{usamos (*)} \\
&= \sum_{k=k_0}^{\infty} \left((2^{k-1})^{\frac{1}{r}} \left(\frac{1}{2}\right)^{k-1} \left(f\left(\frac{1}{2^{k-1}}\right)\right)^{\frac{1}{r}} \right) 2^{k-2} \\
&= \sum_{k=k_0}^{\infty} 2^{\frac{k-1}{r}} \left(\frac{1}{2}\right) \left(f\left(\frac{1}{2^{k-1}}\right)\right)^{\frac{1}{r}} \\
&= \left(\frac{1}{2}\right)\left(\frac{1}{2}^{\frac{1}{r}}\right) \sum_{k=k_0}^{\infty} 2^{\frac{k}{r}} \left(f\left(\frac{1}{2^{k-1}}\right)\right)^{\frac{1}{r}} .
\end{aligned}
$$

Por tanto

$$
\sum_{k=k_0}^{\infty} 2^{\frac{k}{r}} \left(f\left(\frac{1}{2^{k-1}}\right)\right)^{\frac{1}{r}} < \infty
$$

entonces

$$
\sum_{k=1}^{\infty} 2^{\frac{k}{r}} \left(f\left(\frac{1}{2^{k-1}}\right)\right)^{\frac{1}{r}} < \infty.
$$

Aplicando el hecho de que: $\sum a_n < \infty, \Rightarrow \;\; \exists\, \{b_n\}_{n=1}^{\infty} \uparrow \infty$ tal que $\sum a_n b_n < \infty$† se tiene que $\exists \{b_n\}_{n=1}^{\infty} \uparrow \infty$ tal que

$$
\sum_{k=1}^{\infty} b_k 2^{\frac{k}{r}} \left(f\left(\frac{1}{2^{k-1}}\right)\right)^{\frac{1}{r}} < \infty
$$

Sea $\phi\colon [0,1] \to \mathbb{R}$ definida como $\phi(0) = 0$, $\phi(\frac{1}{2}^{k+1}) = \frac{1}{b_k}$, y definida linealmente en el resto, dado que $\{b_n\}_{n=1}^{\infty}$ es creciente se tiene que ϕ satisface las hipotesis del corolario 13.8, por tanto se tiene el resultado.

Nótese que ahora X es continuo por trayectorias.

Teorema 13.11

Sea f una función no negativa sobre $[0,1]$ la cual es creciente cerca de 0. Sea X un proceso estocástico con media 0 y segundos momentos finitos y con trayectorias en D que satisface:

$$
E(X(t) - X(s))^2 \leq f(|t-s|) \quad \text{para } |t-s| \text{ pequeño} .
$$

---

† ver apéndice

y

$$\int_0 y^{\frac{3}{2}} \left(f(y)\right)^{\frac{1}{2}} \ dy < \infty$$

Entonces X es continuo por trayectorias y satisface el T.C.L.

Demostración

Por el teorema 13.10, tomando $r = 2$ obtenemos que X es continuo por trayectorias.

Sea $\{X_n\}$ una sucesión de variables aleatorias independientes e identicamente distribuidas C-valuadas con distribucion $L(X)$.

Sea $Z_n = \frac{X_1+X_2+\cdots+X_n}{\sqrt{n}}$.

Observemos que

$$\begin{aligned}
n(Z_n(t) - Z_n(s))^2 &= [(X_1(t) - X_1(s)) + (X_2(t) - X_2(s)) + \cdots + (X_n(t) - X_n(s))]^2 \\
&= (X_1(t) - X_1(s))^2 + (X_2(t) - X_2(s))^2 + \cdots + (X_n(t) - X_n(s))^2 \\
&\quad + 2 \sum_{i \neq j} ((X_i(t) - X_i(s)) \left((X_j(t) - X_j(s)\right) \\
&= \sum_{i=1}^{n} ((X_i(t) - X_i(s))^2 + 2 \sum_{i \neq j} \left((X_i(t) - X_i(s))((X_j(t) - X_j(s)\right)
\end{aligned}$$

aprovechando que $X_1, X_n, \cdots, X_n$ son igualmente distribuidas tenemos que :

$$\begin{aligned}
E(Z_n(t) - Z_n(s))^2 &= \frac{1}{n} \sum_{i=1}^{n} E((X_i(t) - X_i(s))^2 \\
&\quad + \frac{2}{n} \sum_{i \neq j} E((X_i(t) - X_i(s)) E((X_j(t) - X_j(s)) \\
&= \frac{1}{n} \sum_{i=1}^{n} E((X(t) - X(s))^2 \\
&\quad + \frac{2}{n} \sum_{i \neq j} [E(X_i(t)) - E(X_i(s))][E(X_j(t)) - E(X_j(s))] \\
&= \frac{1}{n} n E((X(t) - X(s))^2 + \frac{2}{n} \sum_{i \neq j} 0 \\
&= E((X(t) - X(s))^2
\end{aligned}$$

Ahora para el proceso Z_n se tiene entonces que $\exists f$ no negativa sobre $[\,0,1\,]$, creciente cerca de 0 tal que:

$$E(Z_n(t) - Z_n(t))^2 \leq f(|t - s|),$$

por cierto que $E(Z_n(t)) = 0 \quad \forall t$ y $E(Z_n(t))^2 < \infty \quad \forall t$. Ya que

$$\int_0 y^{\frac{3}{2}} \left(f(y)\right)^{\frac{1}{2}} \ dy < \infty$$

por el teorema 13.10 obtenemos que $\exists\phi: [0,1] \to \mathbb{R}$ creciente, tal que $\phi(h) \to 0$ si $h \to 0$, la cual solo depende de f y $\exists A^n \in L^2(\Omega, P)$ tal que

$$|\tilde{Z}_n(t) - \tilde{Z}_n(s)| \leq A^{X_n}\phi(|t-s|) \quad \forall n \quad .$$

En este caso ya que X_i tiene trayectorias continuas y por el teorema (*.3) † tenemos que $Z_n = \tilde{Z}_n$.

Sea $n \in \mathbb{N}$, como $||A^{X_n}|| \leq M$ donde M depende de f y ϕ entonces

$$||A^{X_n}|| \leq M \quad \forall n.$$

Sea $\epsilon > 0, \eta > 0$, dado que :

$$|Z_n(t,\omega) - Z_n(s,\omega)| \leq A^{X_n}(\omega)\phi(|t-s|)$$

entonces :

$$\sup_{|s-t|<\delta} |Z_n(t,\omega) - Z_n(s,\omega)| \leq A^{X_n}(\omega) \sup_{|s-t|<\delta} \phi(|t-s|)$$

dado que esto se satisface para toda ω entonces

$$\begin{aligned}\int_\Omega w^2(Z_n,\delta)\, dP &\leq \int_\Omega (A^{X_n}(\omega))^2 (\sup_{|s-t|<\delta} \phi(|t-s|))^2 \\ &= (\sup_{|s-t|<\delta} \phi(|t-s|)^2 ||A^{X_n}||_2^2 \\ &\leq M^2 (\sup_{|s-t|<\delta} \phi(|t-s|)^2 \quad .\end{aligned}$$

Por Chebychev :

$$\mathrm{P}\{w(Z_n,\delta) \geq \epsilon\} \leq \frac{1}{\epsilon^2} M^2 \left(\sup_{|s-t|<\delta} \phi(|t-s|)\right)^2 .$$

Nótese que M no depende de n. Ahora, dado que $\phi(h) \to 0$, si $h \to 0$ podemos elegir δ tal que

$$\frac{1}{\epsilon^2} M^2 (\sup_{|s-t|<\delta} \phi(|t-s|)^2 \leq \eta \quad .$$

Por consiguiente :

$$\mathrm{P}\{w(Z_n,\delta) \geq \epsilon\} \leq \eta$$

Ya que esto se satisface para toda n entonces $\{Z_n\}$ es uniformemente equicontinua en probabilidad.

En vista de la observación (12.3) basta ver que $\exists \hat{Z}$ un proceso Gaussiano con la misma covarianza que X y con trayectorias continuas. Sea Z un proceso Gaussiano con la misma covarianza que X (que siempre existe)††.

---

† ver apéndice

†† ver Ash, Teorema 1.2.3

Aplicando el teorema 13.10 a Z, dado que

$$E|Z(t)-Z(s)|^2 = E|X(t)-X(s)|^2 \leq f(|t-s|)$$

se tiene que Z es continuo por trayectorias, sea $\hat{Z}$ la version continua de Z, dado que 2 procesos equivalentes tiene la misma distribución * entonces $\hat{Z}$ es el proceso buscado.

---

* ver apéndice

BIBLIOGRAFIA

[1] . Bartle R.(1966) The elements of integration, Willey.

[2] . Billingsley P.(1968), Convergence of Probability Mesuares, Willey, New York.

[3] . Caballero M.E, Fernandez B. Notas del curso-Seminario de Prob.

[4] . Delporte J. (1964) Fonctions aleatoires presque surement continues intervalle Ferme.Ann. Inst. H. Poincaré. Sec. B, {\bf 1} 11-215.

[5] . Dieudonné J. (1969) Foundations of Modern Analysis, Academic Press, New York.

[6] . Garcia M. A.(1989) Topologia General. Porrua.

[7] . Hahn M. (1977) Conditions for sample-continuity and the central limit theorem, Ann.Prob. Vol\ {\bf 5} No 3. 351-356.

[8] . Hahn M. (1972) Anote on central limit theorem for sqare-integralbe processes, Am.Math.Soc. Vol\ {\bf 64} No 2.

[9] . Kurtz T., Markov Processes.

[10]. Yeh (1973) Stochastic Processes and the Winner Integral,Pure and M Aplied. New York.